

Nicole Auferkorte-Michaelis, Maiken Bonnes, Patrick Hintze, Julia Liebscher (Hrsg.)

Prüfungen digital gestalten

Technische und didaktische Konzepte für die Hochschullehre



Verlag Barbara Budrich



Prüfungen digital gestalten

Nicole Auferkorte-Michaelis
Maiken Bonnes
Patrick Hintze
Julia Liebscher (Hrsg.)

Prüfungen digital gestalten

Technische und didaktische Konzepte
für die Hochschullehre

Verlag Barbara Budrich
Opladen • Berlin • Toronto 2025

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Gedruckt auf FSC®-zertifiziertem Papier, CO₂-kompensierte Produktion. Mehr Informationen unter <https://budrich.de/nachhaltigkeit/>. Printed in Europe.

© 2025 Dieses Werk ist beim Verlag Barbara Budrich GmbH erschienen und steht unter der Creative Commons Lizenz Attribution 4.0 International (CC BY 4.0):

<https://creativecommons.org/licenses/by/4.0/>

Diese Lizenz erlaubt die Verbreitung, Speicherung, Vervielfältigung und Bearbeitung unter Angabe der Urheber*innen, Rechte, Änderungen und verwendeten Lizenz.

Stauffenbergstr. 7 | D-51379 Leverkusen | info@budrich.de | www.budrich.de



Die Verwendung von Materialien Dritter in diesem Buch bedeutet nicht, dass diese ebenfalls der genannten Creative-Commons-Lizenz unterliegen. Steht das verwendete Material nicht unter der genannten Creative-Commons-Lizenz und ist die betreffende Handlung gesetzlich nicht gestattet, ist die Einwilligung des jeweiligen Rechteinhabers für die Weiterverwendung einzuholen. In dem vorliegenden Werk verwendete Marken, Unternehmensnamen, allgemein beschreibende Bezeichnungen etc. dürfen nicht frei genutzt werden. Die Rechte des jeweiligen Rechteinhabers müssen beachtet werden, und die Nutzung unterliegt den Regeln des Markenrechts, auch ohne gesonderten Hinweis.

Dieses Buch steht im Open-Access-Bereich der Verlagsseite zum kostenlosen Download bereit (<https://doi.org/10.3224/84743098>).

Eine kostenpflichtige Druckversion (Print on Demand) kann über den Verlag bezogen werden.

Die Seitenzahlen in der Druck- und Onlineversion sind identisch.

ISBN 978-3-8474-3098-8 (Paperback)

eISBN 978-3-8474-3233-3 (PDF)

DOI 10.3224/84743098

Druck: Libri Plureos, Hamburg

Umschlaggestaltung: Johannes Gündel, Kleinmachnow

Titelbildnachweis: [shutterstock.com](https://www.shutterstock.com)

Satz: Linda Kutzki, Berlin – www.textsalz.de

Inhalt

Konzepte für die Lehrpraxis: Einblicke in die Projektarbeit	9
<i>Nicole Auferkorte-Michaelis, Maiken Bonnes, Patrick Hintze und Julia Liebscher</i>	

Threshold Concepts und Decoding the Disciplines

Threshold Concepts als Reflexionsanlass: Transformation und Variabilität in Lernprozessen wahrnehmen und begleiten	35
<i>Cornelia Kenneweg</i>	

Integriertes Lehr-, Lern- und Prüfkonzept für eine politikwissenschaftliche Einführungsveranstaltung	47
<i>Ulrike Berendt, Martin Krybus und Till Behmenburg</i>	

Schwellenkonzepte und Lehrinnovationen in fachlichen und fachdidaktischen Veranstaltungen im Lehramtsstudium Mathematik	57
<i>Andreas Büchter, Dana Eilers und Florian Schacht</i>	

Gestaltung lern(enden)zentrierter Veranstaltungen im bildungswissenschaftlichen Studium der Bachelor-Lehramtsstudiengänge der UDE	71
<i>Ruth M. Ingendoh und Angela Heine</i>	

Kompetenzorientiert digital prüfen in der Sport- und Bewegungswissenschaft	83
<i>Thomas Mühlbauer und Johanna Lambrich</i>	

Decoding Deutsch als Zweitsprache: Was macht das Fach DaZ schwer (zu prüfen)?	91
<i>Nguyen Minh Salzmann-Hoang und Tobias Schroedler</i>	

Digital üben, prüfen und bewerten

Das Beste aus Online und Präsenz – Open-Book-Prüfungen mit handschriftlichen Lösungen und digitalem Bewertungsworkflow	103
<i>Mathias Magdowski</i>	
Digitales Lehren und Lernen mit Moodle und JACK	115
<i>Christoph Olbricht und Melanie Schypula</i>	
Auf dem Weg zur Innovation: Digitale Präsenzklausuren am Beispiel der Englischen Sprachwissenschaft	125
<i>Frauke Milne, Birte Bös und Jana Eismann</i>	
Digitale Prüfwerkzeuge in der Baudynamik – Ein Erfahrungsbericht	137
<i>Carolin Birk und Ajay Kumar Pasupuleti</i>	
Aktivität von Studierenden einfordern – Ein Erfahrungsbericht zu den Herausforderungen interaktiver (teil)digitaler Lehre	147
<i>Jochen Gönsch</i>	
Aspekte des digitalen Lehrens und Prüfens: Überblick und neuere Entwicklungen im Fach Statistik	157
<i>Christoph Hanck und Jens Klenke</i>	
Entwicklung eines formativen Assessments zur Förderung prozeduralen und konditionalen biologiedidaktischen Wissens	169
<i>Justin Timm, Nora Lösing, Christian Johannes und Philipp Schmiemann</i>	
Hilfe, ich komme nicht weiter! Digitale Lösungs- und Visualisierungshilfen für Übungs- und Klausuraufgaben der Experimentalphysik	179
<i>Heike Theyßen, Lisa Danzig, Sebastian Haase und Hermann Nienhaus</i>	
Variation von Aufgabenstellungen zu Modellierung in der Informatik	189
<i>Marcellus Siegburg und Janis Voigtländer</i>	
Erste Schritte auf dem Weg zu einer digitalen Lern- und Prüfungsumgebung für die organische Chemie	199
<i>Katrin Schüßler, Michael Giese und Maik Walpuski</i>	
Generative KI und Prüfungen: Erste Erfahrungen und offene Fragen	209
<i>Michael Striewe</i>	

Rechtliche und studentische Perspektiven auf das (digitale) Prüfen

Die rechtliche Beratung im PITCH-Projekt – Ein Erfahrungsbericht	221
<i>Henrik Noll</i>	
Die Student Voice Group im Projekt PITCH – Ein Praxisbericht	231
<i>Maiken Bonnes und Kilian Schmitt (in Zusammenarbeit mit Colin Corbach, Ronja Hensle, Marc Ludwig, Sophie Stotz Anido und Frederike Welzel)</i>	
Portraits der Autorinnen und Autoren	243

Konzepte für die Lehrpraxis: Einblicke in die Projektarbeit

Nicole Auferkorte-Michaelis, Maiken Bonnes, Patrick Hintze und Julia Liebscher

Die digitale Transformation bietet die Chance, Prüfungen nicht nur effizienter, sondern auch gerechter und transparenter zu gestalten. Doch um dieses Potenzial ausschöpfen zu können, bedarf es eines ganzheitlichen Entwicklungsprozesses, der alle relevanten Akteursgruppen einbezieht und die strukturellen Rahmenbedingungen berücksichtigt. Die Beiträge in diesem Band beschreiben die Herausforderungen, Erfahrungen, Entwicklungen geeigneter Prüfungsformate sowie die Umsetzung digitaler Prüfungen im Lehralltag – begleitet von der Motivation, Datenschutz- und Sicherheitsaspekte zu berücksichtigen.

In diesem Band werden Ideen, Erfahrungen und Ergebnisse der (Weiter-)Entwicklung des digitalen Prüfens in der Hochschullehre aus den Perspektiven der beteiligten Akteur:innen skizziert. Der Fokus liegt auf didaktischen, technischen, organisatorischen und rechtlichen Lösungen, die ineinandergreifen müssen, um digitale Prüfungen kompetenzorientiert zu gestalten und umzusetzen. Die Universität Duisburg-Essen (UDE) hat sich mit dem Projekt „PITCH – Prüfungen innovieren, Transfer schaffen, Chancengerechtigkeit fördern“¹ auf den Weg gemacht, Strategien und Rahmenbedingungen zu schaffen, die eine nachhaltige Integration digitalen Prüfens in den Hochschulalltag ermöglichen.

In dieser Einleitung wird die Entstehung der Projektidee skizziert und darauf eingegangen, wie das Design gezielt auf die Herausforderungen digitaler Transformation abgestimmt wurde. Die Beschreibung der Umsetzung verdeutlicht, wie auf technologische Herausforderungen, organisatorische Komplikationen und äußere Einflüsse reagiert werden musste. Hierzu wurde auch die Perspektive der Hochschulleitungsmitglieder eingeholt, die die Hintergründe, strategischen Zielsetzungen und Erfahrungen mit den Herausforderungen in Form eines Interviews eingeordnet haben. Abschließend werden zentrale Erkenntnisse (Lessons Learned) herausgearbeitet, die als wertvolle Impulse für zukünftige Digitalisierungsprojekte dienen können, und die Beiträge des Sammelbands vorgestellt.

1 Das Projekt wurde bzw. wird unter der Projektnummer FBM2020-EA-1190-00081 aus Mitteln der Stiftung Innovation in der Hochschullehre gefördert. Die Laufzeit des Projekts reichte ursprünglich von August 2021 bis Juli 2024, wurde später aber bis Dezember 2025 verlängert.

1 Ein Netzwerk für das Lehren und Lernen

Die Hochschulforscherin und -managerin Ada Pellert (1999) forderte bereits um die Jahrtausendwende von der Universität als Expert:innenorganisation, dass sie sich von einer Organisation des Lernens zu einer lernenden Organisation entwickeln müsse, um den zukünftigen komplexen Anforderungen und Aufgaben gerecht werden zu können. Ein Prozess des hierfür notwendigen organisationalen Lernens, der das Erkenntnis- und Leistungsvermögen von Studium und Lehre an der Hochschule erweitert, basiert auf der Entwicklung neuer Kompetenzen und geht über die Sensibilisierung der Akteursgruppen und individuelle Erfahrungen hinaus (vgl. Senge, 1997). Dieser „tiefgreifende Lernzyklus“ (Senge/Kleiner/Smith/Roberts/Ross, 2008, S. 19) ist notwendig, damit die Wissensbasis einer Organisation sich erweitert und (Weiter-)Entwicklung angestoßen wird.

Hochschullehrer:innen, die innovative Ideen erproben, wie zum Beispiel neue digitale Ansätze in ihrer Lehr- und Prüfungspraxis, dies durch Befragungen begleiten und dabei auch Kontrollgruppen-Designs nutzen, sind gleichzeitig Forschende in ihrer eigenen Lehrpraxis, entwickeln diese evidenzbasiert weiter und tragen als Multiplikator:innen in ihrer jeweiligen Fach-Community zur Verbreitung neuer Lehr-Lernansätze bei. Die Weiterentwicklung der Lehrqualität an Hochschulen erfordert die Zusammenarbeit verschiedener Expertisen aus der Hochschuldidaktik, dem Qualitätsmanagement, der Hochschulsteuerung und nicht zuletzt den Disziplinen und ihren Akteur:innen: Ausgangsfragen werden geteilt oder gemeinsam formuliert, die Verwertung der Ergebnisse festgelegt und die Zusammenarbeit sowie der Ablauf geplant. So entstand das Projekt PITCH zur Erschließung der Potenziale digitaler Prüfungen. Im interdisziplinären Austausch einer Professional Learning Community von Lehrenden, Studierenden und zentralen Stellen werden in allen Fakultäten beispielgebende Prüfungen weiterentwickelt.

Ein einzelnes Projekt, selbst wenn es hochschulweit viele Fakultäten, zentrale Einrichtungen und Verwaltungsbereiche umfasst, kann aber nur ein Baustein im Gefüge der Entwicklung für die Qualität der Lehre und ihre kontinuierliche Erneuerung sein. Vielmehr ist es hierfür notwendig, systematisch Prozesse und Akteursgruppen zu involvieren, die Expertise aufbauen und weitergeben können, sodass die Weiterentwicklung letztlich nicht von einzelnen engagierten Personen abhängig ist, wenngleich die Einzelnen durchaus für die Entwicklung unverzichtbar sind. Alle Programme und Projekte zielen daher darauf ab, ein hochschulinternes Netzwerk für „gute Lehre“ zu etablieren und die Sichtbarkeit guter Lehrpraxis in der Hochschulöffentlichkeit zu stärken.

Die Universität Duisburg-Essen fördert systematisch die aktive Auseinandersetzung mit der eigenen Hochschullehre. Mit einem Förderprogramm werden Lehrende und ihre Teams aufgefordert, ihre Lehrpraxis weiterzuentwickeln. Die Universität verfügt mit dem Dachprogramm „Lehr-Lern-Innovationen an der UDE“ über ein etabliertes Format zur Förderung von kleineren Entwicklungs-

projekten in Studium und Lehre. Das Programm bietet einen niedrigschwelligen Einstieg in die Innovationspraxis von Studium und Lehre und unterstützt das große Engagement in der Lehrentwicklung ideell und finanziell. Die stetig steigende Anzahl der Anträge verdeutlicht, dass die Weiterentwicklung und Innovierung von Lehre von immer mehr Lehrenden als wichtiges Arbeitsfeld wahrgenommen wird. Die innerhalb von sechs Jahren gewonnenen Hochschullehrer:innen und ihre Teams mit insgesamt über 130 Mitarbeiter:innen bilden für die UDE schon jetzt ein wertvolles Netzwerk, um Inspiration für gute Lehre zu schaffen, einen Resonanzraum für die strategische Entwicklung aufzubauen und Drittmittel für größer angelegte Lehrentwicklungsprojekte einzuwerben. Letzteres betrifft wiederum auch das Projekt PITCH zur Entwicklung digitaler Prüfungen an der UDE, an dem vorwiegend Hochschullehrer:innen beteiligt sind, die zuvor schon Erfahrungen im universitätseigenen Förderprogramm „Lehr-Lern-Innovationen an der UDE“ gesammelt haben.

Die Verbindung von Hochschullehrenden, Akteur:innen in Lehr- und Prüfungsorganisation und Didaktik, Qualitätsentwicklung und der Universitätsleitung, ihren Interaktionen und Perspektiven sowie der kollegiale Austausch zur Weiterentwicklung der Lehrpraxis sind grundlegend für die hochschulinterne Netzwerkbildung. Daher finden in regelmäßigen Abständen in allen Projekten und Programmen an der Universität Netzwerkkonferenzen statt, die neben dem Austausch zu den eigenen Entwicklungsvorhaben in der Lehrpraxis auch thematische Impulse, Diskussionen und Formate zur Unterstützung sowie zur eigenen Professionalisierung bzw. Weiterbildung anbieten. Die Vernetzung der Lehrenden untereinander sowie die mit den weiteren Akteursgruppen in Studium und Lehre wird zu einer Community of Practice für die Innovation von Lehre und Studium. Die Erfahrungen aus Innovationsprojekten und der Austausch in verschiedenen Formaten haben schließlich auch eine neue Projektidee befördert: das digitale Prüfen.

2 Entstehung einer Projektidee zum digitalen Prüfen

Das Vorhaben, ein Projekt zu digitalen Prüfungen aufzulegen, zeichnete sich schon wenige Tage nach der Veröffentlichung der ersten Ausschreibung der Stiftung Innovation in der Hochschullehre im November 2020 ab. Die Idee dahinter lässt sich auf zwei maßgebliche Entwicklungsstränge zurückführen: Zum einen hatte die Universität die Digitalisierung in Studium und Lehre immer wieder in größeren Entwicklungsprojekten aufgegriffen und konnte hier auf umfangreiche Vorarbeiten, insbesondere im Bereich der E-Assessments, aufbauen. Zum anderen flossen in die Entscheidung auch die Erfahrungen mit digitalen Prüfungen unter Pandemie-Bedingungen ein, die das Potenzial von Innovationen in der Prüfungspraxis verdeutlichten. Beide Aspekte sollen nachfolgend näher ausgeführt werden.

Die Bund-Länder-Programme der 2010er Jahre haben die Weiterentwicklung der Lehre an den deutschen Hochschulen maßgeblich befördert. Mit ihnen war es möglich, neue curriculare und außercurriculare Angebote zu schaffen, Unterstützungsstrukturen aufzubauen und Innovationen zu fördern, für die es den Hochschulen im Regelbetrieb schlicht an Geld fehlen würde. Die UDE hat mit ihrem Qualitätspakt-Lehre-Projekt den Fokus auf die Studieneingangsphase gelegt und in diesem Rahmen eine Reihe von Maßnahmen mit struktureller Wirkung initiiert. Neben dem Aufbau eines fakultätsübergreifenden Mentoring-Systems, einer MINT-Studieneingangsphase und der Schreib- und Sprachförderung von Studierenden wurde auch ein Bereich für E-Assessments geschaffen, um Studierende insbesondere in den großen Veranstaltungen der ersten Fachsemester mit online absolvierbaren Übungsaufgaben zu unterstützen. Das im Rahmen einer Arbeitsgruppe der Informatik entwickelte E-Assessment-System JACK wurde zu diesem Zweck zu einer universitätsweit eingesetzten Plattform um- und ausgebaut. Wurde es zu Beginn nur im Rahmen der Programmierausbildung eingesetzt, erweiterte sich das Einsatzspektrum im Laufe der Jahre systematisch um die Übungen und Testate vieler Grundlagenveranstaltungen der Mathematik, Natur-, Ingenieur- und Wirtschaftswissenschaften. Die Vorzüge des Systems lagen insbesondere in der Generierung von Aufgaben mit randomisierten Werten, der Möglichkeit, automatisiertes Feedback zu typischen Fehlerquellen vorzusehen, und der Abbildung komplexerer, mehrstufiger Aufgaben, weshalb es unter den Lehrenden und Studierenden eine hohe Verbreitung fand. E-Assessments hatten im Rahmen des Projekts eine derart hohe Verbreitung gefunden, dass sie aus dem Übungsbetrieb nicht mehr wegzudenken waren. Die damalige Prorektorin für Studium und Lehre, Prof. Dr. Isabell van Ackeren-Mindl, erläutert die strategische Dimension des Projekts:

Wir haben festgehalten, dass es nicht unser Anliegen ist, Projekte einzuwerben, sie abzuarbeiten und dann folgt nichts daraus. Uns war wichtig, die Projekte mit strategischen Entwicklungsfragen zu verknüpfen und zu überlegen, wie wir das Erreichte später in die Verstetigung bringen können. Das betrifft auch unsere Aktivitäten im Bereich der E-Assessments, die mithilfe des JACK-Systems von immer mehr Lehrenden für veranstaltungsbegleitendes, formatives Feedback eingesetzt wurden. Gerade im Zuge dieses starken Wachstums war uns wichtig, immer wieder zu klären, welche Bedarfe die Disziplinen an Aufgabentypen und Testgestaltung haben, wie wir das Alignment zwischen Lehr-Lern-Zielen, Lehr-Lern-Phasen und Prüfungen sicherstellen und wie wir die Daten im System auch für die Qualitätsentwicklung in Studium und Lehre nutzbar machen können. Daraus haben sich schließlich auch strategische Linien der Universität entwickelt. (Interview mit Isabell van Ackeren-Mindl, S. 2)

PITCH fügte sich somit nicht nur in die Strategien der UDE ein, es setzte auch eine Tradition in der Entwicklung von E-Assessments fort. Zwar beschränkte sich diese einstweilen auf den Bereich der Übungen bzw. formativen Assessments, mit Einsetzen der Corona-Pandemie ergab sich jedoch schlagartig ein Bedarf an Lösungen für summative Assessments.

Nachdem die Hochschulen zu Beginn des Sommersemesters 2020 zunächst ihre Lehrveranstaltungen auf Online-Formate umstellen mussten, wurde wenig später bereits darüber diskutiert, wie die Prüfungen angesichts der Abstands- und Versammlungsaufgaben umgesetzt werden können bzw. welche digitalen Alternativen es zu Präsenzformaten geben könnte. Das Diskussionsforum an der UDE hierfür war die Taskforce Studium und Lehre – ein Gremium, das während der Pandemie alle wesentlichen Entscheidungen vorbereitet hat, die zur Bewältigung der vielfältigen Herausforderungen notwendig waren, und hierzu Vertreter:innen aus Fakultäten, Zentralen Einrichtungen und Hochschulleitung in wöchentlichen Telefon- und Videokonferenzen zusammenbrachte. Angesichts der häufig sehr kurzfristig kommunizierten Verordnungen der Ministerien mussten Entscheidungen zur Einführung neuer Systeme, Unterstützungsangebote und Leitfäden in hoher Geschwindigkeit abgewogen und getroffen werden. Schnell wurde festgehalten, das Lernmanagementsystem Moodle und das E-Assessment-System JACK auch für distante Online-Prüfungen einzusetzen. Während sich JACK vor allem für mathematische, naturwissenschaftliche und technische Fächer eignete und mit den Automatisierungs- und Randomisierungsfunktionen bereits ein Minimum an Täuschungsprävention mitbrachte, wurde Moodle für Openbook-Klausuren geöffnet, die an der UDE als Take-Home-Exams pilotiert wurden und insbesondere die geistes-, gesellschafts- und bildungswissenschaftlichen Fachrichtungen ansprachen. Das Konzept sah vor, dass Studierende Aufgaben online in einem längeren Zeitraum (z. B. sechs Stunden) unter Heranziehung beliebiger Hilfsmittel bearbeiten konnten. Der Zeitraum wurde so gewählt, dass auch ein zwischenzeitlicher Verbindungsabbruch kein Problem darstellen sollte, was angesichts der teilweise limitierten technischen Möglichkeiten der Studierenden zu Beginn der Pandemie nicht ausgeschlossen werden konnte (Stammen/Ebert, 2021). Das Format konnte im Sommersemester in einigen Dutzend Veranstaltungen pilotiert werden und fand sowohl bei Lehrenden als auch Studierenden hohen Zuspruch. Die Prorektorin skizziert, wie die Umsetzung der Take-Home-Exams schließlich auch inhaltliche Gestaltungsfragen auf die Agenda brachte:

Die Taskforce war ein wichtiges Gremium, um gemeinsam nach Lösungen zu suchen, Transformationen vorzubereiten und zu begleiten. Es ging darum, aus einer Notlage heraus vor die Welle zu kommen. Ausgehend von der Frage, wie wir Prüfungen mit vielen Teilnehmenden stattfinden lassen können, für die es unter Beachtung der Abstandsregeln faktisch keine Räume gibt, haben wir das Format der Take-Home-Exams entwickelt und implementiert. Das Format hatte zur Folge, dass wir nicht nur

rechtliche, organisatorische und technische Herausforderungen angehen mussten, sondern verstärkt auch auf die Qualität von Aufgaben geschaut haben. Aufgaben mussten kompetenz- und problemlösungsorientierter gestellt werden, was noch einmal zur Professionalisierung in diesem Bereich beigetragen hat. Schnell war klar, dass wir das Format fortführen, in die Breite tragen und weiterentwickeln möchten, was wir mit dem PITCH-Projekt schließlich aufgegriffen haben. (Interview mit Isabell van Ackeren-Mindl, S. 3)

Die Evaluation des neuen Klausurformats zeigte die Herausforderungen und Bedarfe auf, die es bei der Gestaltung digitaler Prüfungen anzugehen gilt. Die wenig später veröffentlichte Ausschreibung der Stiftung Innovation in der Hochschullehre fiel hier auf einen fruchtbaren Boden, stand das Thema Prüfen doch im Hinblick auf die vielen Vorarbeiten, die strategische Einbettung und den Erfahrungen aus dem ersten Online-Semester weit oben auf der Agenda. Schnell wurde das Ziel festgelegt, das Thema Prüfen als strategische Entwicklungsaufgabe zu begreifen und dabei insbesondere die Praxis zu digitalen Prüfungen an der Universität deutlich auszubauen und weiterzuentwickeln. Um auch den Bedarfen nach einer Pandemie gerecht werden zu können, sollte der Fokus nicht allein auf Online-Prüfungen liegen. Stattdessen sollten von Anfang an auch digitale Prüfungen in Präsenz einbezogen werden, bei denen die Studierenden die Prüfung vor Ort mit einem standardisierten Endgerät, z. B. in einer PC-Hall der Universität, durchlaufen.

In der Planung des Vorhabens wurden fünf Entwicklungsfelder identifiziert, die nachfolgend näher skizziert werden.

1. *Stärkung der Kompetenzorientierung*: Die Auseinandersetzung mit den Take-Home-Exams hat gezeigt, wie wichtig es ist, Prüfungen kompetenzorientiert zu gestalten. Da Lehrende immer wieder über Schwierigkeiten berichteten, Aufgaben zu konzipieren, die an höhere Taxonomiestufen anknüpfen, sollte das Projekt auch Anlässe schaffen, die Entwicklung von Lehre und Prüfungen mit didaktischen Ansätzen zu unterstützen.
2. *Implementation neuer Aufgabentypen und Prüfungsformate*: Aus den bereits durchgeführten Projekten war bekannt, dass die Akzeptanz digitaler Prüfungsformate unter Prüfer:innen in hohem Maße von den technischen Möglichkeiten abhängt, fachbezogene Aufgabentypen, Techniken und Notationen verwenden zu können (z. B. zur Darstellung geometrischer Körper oder Eingabe chemischer Strukturformeln). Der flächendeckende Einsatz der Systeme Moodle und JACK im ersten Online-Semester zeigte auf, welche Funktionen Lehrende noch vermissten bzw. welche disziplinspezifischen Aufgabentypen noch abzubilden waren.
3. *Unterstützung der Studierenden mit digitalen, auf die Prüfungen abgestimmten Übungsaufgaben*: Prüfungskonzept und Lehrkonzept sind untrennbar miteinander verbunden. Ein Projekt zur Weiterentwicklung digitaler Prüfungen

musste zwangsläufig auch die darauf vorbereitende Lehre in den Blick nehmen. Neben den allgemeinen didaktischen Herausforderungen sollten hier vor allem weitere Übungsmöglichkeiten für Studierende geschaffen werden.

4. *Erhöhung der Flexibilität und Chancengerechtigkeit*: Die UDE befragte im Sommersemester 2020 ihre Studierenden zu ihrer technischen Ausstattung und ihren Erwartungen an die Online-Lehre. Es beteiligten sich mehr als 7.000 Studierende, von denen jede:r vierte angab, nicht über alle für die Online-Lehre benötigten Geräte zu verfügen und sich eine Anschaffung nicht leisten zu können. Die Hälfte der Studierenden bekundete die Sorge, durch digitale Prüfungen Nachteile zu erfahren (Stammen/Ebert, 2021, S. 5 und S. 12). Die Daten signalisierten, wie wichtig es war, bei der Prüfungsentwicklung die Diversität und Akzeptanz der Studierenden im Blick zu behalten.
5. *Förderung einer systematischen Qualitätsentwicklung von Prüfungen*: Das Thema Prüfungen wurde, abseits der pandemischen Sondersituation, nur selten systematisch diskutiert. Mit dem Projekt sollte der Austausch von Lehrenden und Studierenden gestärkt und nach Möglichkeit mit Daten bereichert werden (Assessment Analytics).

3 Projektdesign

Für die Bearbeitung der Entwicklungsfelder griff die UDE wieder auf einen Netzwerkansatz zurück, um das Thema in der Universität zu stärken bzw. ausgehend von Good Practices in die Breite zu bringen. Eine Gruppe von 20 engagierten Lehrenden sollte mit ihren Teams als „Professional Learning Community“ bis zu 30 Prüfungen prototypisch weiterentwickeln. Dabei handelte es sich um Professor:innen aus allen Fakultäten, die aufgrund früherer interner oder externer Projektförderungen bereits in der Entwicklung von Studium und Lehre erfahren waren. Durch ihre langfristige Verbundenheit zur Organisation wurde sichergestellt, dass sie die Ergebnisse auch nach Projektende in ihrer eigenen Praxis verwerten und sie in ihren Fächern als Multiplikator:innen wirken können. Die Prüfungen wurden so ausgewählt, dass möglichst viele Formate abgedeckt werden, eine hohe Anzahl der Studierenden erreicht wird und zugleich ein hohes Transferpotenzial besteht, um die Ergebnisse des Projekts auf weitere Prüfungen innerhalb der UDE übertragen zu können.

Je nach fachlicher Ausrichtung bzw. Breite sollten die Prüfungen von Lehrenden allein, zu zweit oder auch zu dritt weiterentwickelt werden. Dabei sollten sie in allen Fällen durch ihnen zugewiesene Projektmitarbeitende unterstützt werden. Die insgesamt vierzehn² Projektteams in den Fächern wurden durch Projektstel-

2 In der Antragsphase waren es fünfzehn Teams, durch Neuordnungen ergab sich später die Konstellation von vierzehn Teams.

len bzw. -teams in verschiedenen zentralen Einrichtungen unterstützt, die die Entwicklung der Prüfungen didaktisch, technisch und rechtlich begleiten sollten:

Didaktik: Prüfungen und Lehre stehen in Wechselwirkung zueinander. Lehrende sind im Sinne des Constructive Alignments gefordert, zuerst die Prüfung zu planen und festzulegen, was gelernt und überprüft werden soll, bevor dazu passende Interaktionen in der Lehrveranstaltung entwickelt werden. Lernaktivitäten sollen dabei an das Vorwissen und die Erfahrungen der Studierenden anknüpfen, um neue fachliche Inhalte erschließen zu können. Um diese Zielvorstellung im Rahmen eines auf Prüfungen fokussierten Projekts erreichen zu können, wurden zwei heuristische Zugänge aus der Literatur herangezogen. Mithilfe der Ansätze der Threshold Concepts (Meyer/Land, 2003) bzw. Decoding the Disciplines (Middendorf/Shopkow, 2018) lässt sich analysieren, an welchen Stellen Studierende häufig Schwierigkeiten haben und inwieweit hierfür das Fehlen impliziter Expertise und impliziten Handlungswissens dafür verantwortlich gemacht werden kann. Dabei handelt es sich um Hürden bzw. Schwellen, derer sich die Expert:innen des Fachs häufig nicht mehr bewusst sind, die in der erstmaligen Auseinandersetzung aber als sehr schwierig oder gar kontraintuitiv empfunden werden können (vgl. hierzu auch den Beitrag von Cornelia Kenneweg in diesem Band).³ Im Projekt sollten die beteiligten Lehrenden Gelegenheit dazu haben, mögliche Schwellen in ihren Veranstaltungen bzw. Disziplinen aufzuspüren und bei der kompetenzorientierten Weiterentwicklung ihrer Prüfung zu berücksichtigen.

Technik: Dem Technik-Team kam die Aufgabe zu, Anforderungen der Lehrenden aufzunehmen und in die Weiterentwicklung der Prüfungssysteme Moodle und JACK einfließen zu lassen. Im Sinne des oben genannten Entwicklungsfelds sollten dabei vor allem neue Aufgabentypen und Funktionen in den Blick genommen werden, die für die jeweilige Disziplin der Lehrenden wesentlich sind und deren Berücksichtigung im Funktionsumfang eine noch höhere Akzeptanz und Verbreitung der Systeme im Fach versprechen ließen. Darüber hinaus sollte in diesem Teil des Projekts auch an strukturellen Fragen gearbeitet werden, die beispielsweise Funktionen zu Korrekturworkflows, Bewertungssysteme und Archivierung betrafen.

Recht: Da die Gestaltung von Prüfungen immer wieder rechtliche Fragen, etwa zur chancengerechten Gestaltung von Aufgaben und Prüfungsformaten, aufwirft, sollte eine Stelle zur Rechtsberatung geschaffen werden, die mit Handreichungen und Expertisen auch über das Projekt hinaus wirken sollte.

3 Beispiele hierfür sind z. B. die Quellenkritik in der Geschichtswissenschaft, der Umgang mit Kontingenz und Komplexität bei der Arbeit mit Texten und Theorien in der Linguistik oder das zielgerichtete Lesen von Journal-Artikeln in den Naturwissenschaften.

Die Struktur und der Antrag des Vorhabens wurden zu Beginn des Jahres 2021 entwickelt. Nach der Entscheidung der Auswahlgremien der Stiftung im Sommer wurde bereits im August 2021 mit der Arbeit im Projekt begonnen. Es ist festzuhalten, dass die Planung stark von den Herausforderungen der Corona-Pandemie geprägt war. Die großen Unsicherheiten in Bezug auf Online-Prüfungen, die rasch voranschreitende Digitalisierung in Studium und Lehre und nicht zuletzt die sich schnell verändernden Regeln sollten mit dem Projekt ein Stück weit aufgefangen werden. Nicht absehbar war, dass das Projekt auch einen Raum für die Bearbeitung weiterer Krisen bot, die sich in der zunächst dreijährigen Laufzeit einstellen sollten.

4 Projektpraxis zwischen Prüfungsentwicklung und Krisenresilienz

Mit dem Start des Projekts im August 2021 begann die Aufbauphase, die sich aufgrund der verspäteten Bewilligung, einer teilweise drastischen Mittelkürzung und Schwierigkeiten in der Personalgewinnung bis weit in das Jahr 2022 ziehen sollte. Die Projektkoordination stand vor der Herausforderung, dass die Lehrendenteams nicht zeitgleich starteten, sondern je nach Stellensituation versetzt mit der Prüfungsentwicklung beginnen konnten. Während im Wintersemester 2021/22 zumeist noch das bisherige Prüfungsformat Verwendung fand, jedoch im Rahmen des Projekts bereits näher ausgewertet wurde, erfolgten die ersten Erprobungen neuer Elemente vorrangig im darauffolgenden Sommersemester. Eingeleitet wurde diese Phase mit einer Kick-Off-Konferenz im Februar 2022, die aufgrund der nach wie vor geltenden Versammlungsbeschränkungen in digitaler Form stattfand.

Das Konferenzformat brachte die Lehrenden fächerübergreifend zusammen, stellte Good Practices anderer Standorte vor (vgl. den Beitrag von Mathias Magdowski in diesem Band) und gab eine Einführung in die aktuellen technischen und rechtlichen Entwicklungen sowie die Ansätze der Threshold Concepts und Decoding the Disciplines. In Zusammenarbeit mit Expert:innen anderer Hochschulen erhielten die im Projekt engagierten Lehrenden die Möglichkeit, typische Hürden und Schwellen ihres Fachs mithilfe von Decoding-Interviews aufzuspüren und die dabei gewonnenen Erkenntnisse in die Weiterentwicklung ihrer Prüfungen einfließen zu lassen.

Die Decoding-Interviews wurden in den folgenden Monaten fortgesetzt und boten den Lehrendenteams einen Anlass, um den erreichten Stand in der Lehr- und Prüfungsentwicklung mithilfe einer externen Perspektive zu reflektieren. Weitere Impulse ergaben sich zudem aus Beiträgen von Studierenden. Nicht nur wurden Follow-Up-Befragungen zu einzelnen Prüfungen durchgeführt, die studentische Perspektive sollte ferner auch im Rahmen einer Student Voice Group gestärkt werden – einer Gruppe von Studierenden, die sich fächerübergreifend mit Themen der Prüfungsentwicklung befasst und in Form von Workshops und ande-

ren Beiträgen die Diskussion im Projekt und der Hochschulöffentlichkeit bereichert. Die Formate fanden Eingang in die folgenden Projektkonferenzen im Juni und Dezember 2022, wobei letztere aufgrund des kurz zuvor erfolgten Hackerangriffs auf die Universität noch ganz andere Herausforderungen bereithielt.

Der Hackerangriff stellt eine digitale Krise sui generis dar. Nachdem am 27. November 2022 nahezu alle Systeme vom Internet getrennt werden mussten, war die Kommunikation der Universitätsmitglieder zunächst empfindlich gestört. Die Infrastruktur, die im dienstlichen Alltag als selbstverständlich vorausgesetzt wird, war offline. Studierende, Lehrende und Verwaltungsmitarbeitende fanden am 28. November eine Situation vor, in der keine Telefonie, Mailserver, Webseiten, Chat-Systeme, digitalen Workflows, ERP-Systeme oder Lern- und Prüfungsplattformen erreichbar waren. Wie auch schon andere Hochschulen zuvor, wechselten die Universitätsmitglieder schnell auf alternative Kommunikationsformen, tauschten private Handynummern aus, griffen auf externe Dienste zurück oder nutzten analoge Formen zur Bekanntmachung, wie zum Beispiel schwarze Bretter. Während die Mail- und Verwaltungssysteme noch monatelang gestört sein sollten, konnten die Lern- und Prüfungssysteme Moodle und JACK innerhalb weniger Tage bis Wochen wieder aktiviert werden, gleichwohl der Zugriff noch nicht alle Funktionen umfasste und zunächst nur aus den lokalen Netzen möglich war.

Der Prorektor für Studium, Lehre und Bildung, Prof. Dr. Stefan Rumann, der das Amt wenige Monate vor dem Hackerangriff von seiner Vorgängerin übernommen hatte, erinnert sich daran, wie trotz widriger Umstände schnell Lösungen gefunden werden konnten:

Der Hackerangriff hat gezeigt, wie abhängig wir von der Digitalisierung eigentlich waren. Es war eine Zeit, in der Handynummern auf Post-Its weitergegeben wurden, da selbst die Telefonanlage nicht mehr funktioniert hat. Lehre und Prüfungen konnten nahezu ohne größere Brüche weiterlaufen. Zu verdanken haben wir dies nicht nur analogen Alternativen, die eher zur Überbrückung eingesetzt wurden, sondern vor allem dem Innovationsschub, den wir durch die Corona-Pandemie erfahren haben. Lehrende haben digitale Kompetenzen gewonnen, die dabei halfen, Lehre auch in unvorhergesehenen Situationen mit verschiedenen Medien, Werkzeugen und Systemen zu gestalten bzw. schnell mögliche Alternativen zu entwickeln. Und auch das Thema Prüfen erwies sich als robust. Wenngleich es vorstellbar gewesen wäre, anstelle digitaler Prüfungen wieder zu papierbasierten Formaten zu greifen, so bestanden viele Lehrenden und sogar ganze Fakultäten darauf, ihre Prüfungen wie geplant digital umzusetzen, auch wenn zu diesem Zeitpunkt noch nicht alle technischen Entwicklungen in der Wiederherstellung der Systeme absehbar waren. Das wäre ohne die Entwicklungsprozesse in den Jahren zuvor undenkbar gewesen. (Interview mit Stefan Rumann, S. 4f)

Die Projektkonferenz, die gerade einmal zwei Wochen nach dem Hackerangriff stattgefunden hatte, beleuchtete somit nicht nur den Stand der Entwicklungen in den Teams, sie übernahm auch eine wichtige Funktion als Ort analoger Kommunikation und des Austauschs, an dem über das weitere Vorgehen für den anstehenden Prüfungskorridor beraten werden konnte.

Ein Gesprächsthema, das auf der Konferenz am Rande bereits diskutiert wurde, in den kommenden Monaten aber noch sehr viel mehr Raum einnehmen sollte, war die erst wenige Tage zuvor erfolgte Veröffentlichung von ChatGPT 3.5. Für ein Projekt, das auf die Entwicklung digitaler Prüfungen setzte, war der Boom von KI-Anwendungen eine ernstzunehmende Herausforderung. Während digitale Klausuren unter kontrollierten Bedingungen in Räumen der Universität kaum betroffen waren, waren die Anforderungen an distante Formate, wie Essays, Hausarbeiten oder die zu diesem Zeitpunkt nur noch selten praktizierten Take-Home-Exams durch die allgemeine Verfügbarkeit von KI-Tools nochmals gestiegen. Aufgabenstellungen mussten fortan so formuliert werden, dass die Gefahr einer Täuschung von KI nachweislich vermieden wurde, sei es durch die Erschließung höherer Taxonomiestufen, den aktiven Einbezug von KI-Kompetenzen der Studierenden oder die Verbindung mit Reflexions- und Bewertungselementen, zum Beispiel im Rahmen von Portfolio-Prüfungen (vgl. Salden/Lordick/Wiethoff, 2023, S. 15–17). Gleichzeitig eröffneten die Tools auch neue Möglichkeiten für Lehrende, etwa in der Entwicklung von Lehr-Lernmaterial, der Erzeugung von Übungs- und Prüfungsfragen oder der allgemeinen Planung von Lehrveranstaltungen.

Das KI-Thema war fortan ein wichtiger Teil der Diskussionen zur didaktischen, technischen und rechtlichen Weiterentwicklung der Prüfungspraxis im Projekt. Während die engagierten Lehrenden ohnehin bereits an kompetenzorientierten Szenarien arbeiteten, wurden über technische Möglichkeiten zum Einbezug von KI diskutiert und die rechtliche Entwicklung des Felds näher untersucht. Immer wieder standen Fragen im Raum, wie mit KI als Hilfsmittel bei Prüfungen umzugehen ist, wie Studierende belegen können, eine Arbeit selbstständig verfasst zu haben, und wie Lehrende möglichen Verstößen nachgehen und diese sanktionieren können. Der Regelungsbedarf für distante Prüfungsformate war damit unübersehbar und stand im starken Kontrast zur Situation in der Corona-Pandemie 2020/21, in der aufgrund der Notlage eher offene und weite Rechtsauslegungen die Regel waren.

Verschärft wurde die Lage 2023 schließlich noch durch landesseitige Regelungen, die die Durchführung distanter Lehr- und Prüfungsformate wieder zu einer Ausnahme erklärt haben. Wurde mehr als ein Viertel der Veranstaltungszeit in Form von E-Learning erbracht, galt diese nach der neuen Verordnung⁴ nun

4 Verordnung betreffend die digitale Lehre sowie betreffend die Durchführung online gestützter Wahlen der Hochschulen und der Studierendenschaften (Hochschul-Digitalverordnung – HDVO NRW).

als Digitallehre und musste gesondert begründet werden. Für digitale Prüfungen wurden teilweise noch höhere Anforderungen definiert, mit dem Ziel, diese möglichst rechts- bzw. klagesicher zu gestalten. Lehrende, die sich für E-Learning und digitale Prüfungen einsetzten, sahen sich durch diese neuen Anforderungen teilweise ausgebremst. Der Prorektor erläutert, wie die neuen Vorgaben von der Universität aufgenommen wurden:

Als die Pandemie ausklang, entwickelte das Wissenschaftsministerium eine Digitalisierungsverordnung, die im ersten Ansatz auch als ein Roll-back gelesen werden konnte. Die Hochschulen mussten sich Gedanken dazu machen, an welchen Stellen digitale Lehre und digitale Prüfungen in einer nicht-pandemischen Situation sinnvoll sind und ihren Einsatz genau regeln und dokumentieren. Die Motivation des Ministeriums, die unmittelbare soziale Interaktion an den Hochschulen nach der Pandemie wieder zu stärken, konnten wir als Präsenzuniversität natürlich nachvollziehen, gleichwohl gab es viele Beteiligte bei uns, die die Vorzüge dieser Formate sehr zu schätzen wussten. An der UDE haben wir die HDVO zum Anlass genommen, eine Leitlinie zu entwickeln, um die digitale Lehre zu regeln, und eine Rahmenprüfungsordnung vorzubereiten, die dies für die Durchführung digitaler Prüfungen leistet. (Interview mit Stefan Rumann, S. 6)

Der Umgang mit der Digitalisierungsverordnung stellt nach dem Hackerangriff und dem KI-Boom den dritten externen Faktor dar, der das Projekt PITCH beeinflusst hat. Die Weiterentwicklung der Prüfungspraxis bewegte sich folglich in einem hochdynamischen Feld, dessen Ausmaß zu Projektbeginn noch gar nicht absehbar war. Rechtliche Normen, technische Möglichkeiten, politische Rahmenbedingungen und nicht zuletzt auch die Akzeptanz von Studierenden und Lehrenden für (distante) digitale Formate erwiesen sich als äußerst fluide. Zwar brachten die verschiedenen Krisen viele Arbeitsplanungen im Projekt immer wieder durcheinander, sie trugen aber auch zu einem starken Interesse der Hochschulöffentlichkeit an der Gestaltung von Prüfungen bei. Die engagierten Lehrenden und Mitarbeitenden im Projekt überzeugten durch einen resilierten Umgang mit den Herausforderungen. Aufgrund ihrer Expertise und Vernetzung – untereinander sowie mit verschiedenen zentralen Stellen – gelang es ihnen, immer wieder kreative Lösungen zu finden, um die Weiterentwicklung und Erprobung digitaler Prüfungen weiter voranzutreiben. Mit ihren Ergebnissen und Erfahrungen war es ihnen möglich, erste Antworten auf die neuen Herausforderungen zu finden, mit anderen Kolleg:innen zu teilen und so der durch die Krisen erzeugten Unsicherheit ein Stück weit entgegenzuwirken.

5 Fünf Einsichten für die Gestaltung von Digitalisierungsprojekten

Aus der Genese des Projekts PITCH lassen sich fünf Einsichten ableiten, die auch für andere Digitalisierungsprojekte in Studium und Lehre gelten dürften. Diese Einsichten beziehen sich nicht auf die inhaltlichen Ergebnisse, über die der vorliegende Sammelband umfassend informiert, sondern vielmehr auf die strukturelle Gestaltung und Implementation. Es handelt sich um Punkte, die insbesondere für die Planung hochschulweiter Lehrentwicklungsprojekte und die hiermit befassten Hochschulleitungen, hochschuldidaktischen Einrichtungen und Medienzentren relevant sein dürften – auch dann, wenn der Fokus nicht auf digitalen Prüfungen, sondern auf anderen Themen und Gegenstandsbereichen der Digitalisierung liegt.

Digitalisierung verläuft dynamisch und verlangt nach agilen Lösungen.

Die vielfältigen Krisenerscheinungen und disruptiven Entwicklungen, die das PITCH-Projekt gekennzeichnet haben, machen deutlich, wie fluide die Digitalisierung an den Hochschulen (noch) ist. Viele Prozesse sind kaum planbar oder zum Zeitpunkt der Antragstellung schlicht überhaupt nicht abzusehen. Diese Dynamik steht in direktem Kontrast zu den Vorgaben und Konventionen, die bei der Einwerbung von Projekten bzw. Drittmitteln im Hochschulsystem zu beachten sind. Detaillierte Meilensteinpläne mit klar definierten Arbeitspaketen verkommen hiernach zu Momentaufnahmen, die schon zu Beginn des Projekts überholt sein können. Im Fall von PITCH war es immer wieder nötig, nachzusteuern und die Projektpläne für Entwicklungen offen zu halten. Das Projektmanagement muss agil bleiben, wie es in anderen gesellschaftlichen Sektoren schon länger üblich ist. Entscheidend hierfür ist eine Projekt- und Hochschulgovernance, die dies aktiv befördert. PITCH konnte die Krisen und Herausforderungen nur aufgrund flacher Hierarchien, flexibel vereinbarter Etappenziele und einer fortwährenden Kommunikation zwischen allen Stakeholdern erfolgreich verarbeiten.

Digitalisierung erfordert eine langfristige Ressourcenbereitstellung.

Eine unmittelbare Herausforderung für das PITCH-Projekt lag und liegt darin, dass die (weiter-)entwickelten Prüfungen nach ihrer Erprobungsphase strukturell in den Hochschulalltag implementiert werden.

Im Rahmen des PITCH-Projekts standen Ressourcen zur Verfügung, die es beispielsweise ermöglichten, unterschiedliche Softwarelösungen für digitale Prüfungen auszuprobieren. So wurden Prüfungen zunächst in LPLUS und anschließend in JACK umgesetzt. Die Projektgelder waren von entscheidender Bedeutung, da die zusätzlichen Arbeiten und der damit verbundene Mehraufwand für

die Programmierung und Einrichtung ohne diese finanzielle Unterstützung nicht hätten bewältigt werden können. Dies hat nicht nur die Qualität der Prüfungen erheblich verbessert, sondern auch die technische Expertise der Lehrendenteams gestärkt. Die Komplexität der technischen Prozesse, die für die Digitalisierung von Prüfungen notwendig sind, erfordert tiefgreifendes Verständnis und praktische Erfahrung. Dieses Wissen lässt sich nicht einfach weitergeben oder schnell aneignen.

Da die Projektstellen befristet sind, bedeutet dies, dass das entsprechend qualifizierte Personal die Universität mit dem Projektende in der Regel wieder verlässt. Diese Abgänge stellen für die Digitalisierung eine erhebliche Herausforderung dar, da das erworbene Wissen und die wertvollen Erfahrungen, die im Rahmen des Projekts gesammelt wurden, mit den Mitarbeitenden verloren gehen. Dieser Verlust von Wissen und Expertise durch den *Brain Drain* hochqualifizierter Mitarbeiter:innen kann dazu führen, dass die weitere Entwicklung und auch Implementierung der innovativen Prüfungsformate in die IT-Prüfungsstruktur ausbleibt. Um dem entgegenzuwirken, wird in der letzten Projektphase der Schwerpunkt der Projektarbeit auf die Implementierung von erfolgreich erprobten Prüfungsformaten gelegt, um möglichst viele Entwicklungen aus der Projektarbeit in den Regelbetrieb zu überführen.

Gelingt es einer Hochschule, den Übergang vom Projekt zum Regelbetrieb personell abzufedern, stellen sich darüber hinaus auch Anforderungen an die technische und räumliche Infrastruktur. Das Projekt hat die Vorzüge digitaler Präsenzprüfungen verdeutlicht, was die Nachfrage in den kommenden Jahren weiter steigen lassen wird. Bei der Durchführung großer Klausuren mit über 300 Studierenden in einer extra hierfür eingerichteten PC-Hall müssen die Studierenden gleichzeitig auf das digitale Prüfungssystem zugreifen. Die Universität sieht den Bedarf und baut weitere PC-Halls, um der Nachfrage gerecht werden zu können. Die Investitionen in Räume, Hardware und Software gehen weit über das Projekt hinaus, sind jedoch eine direkte Folge des Vorhabens, die Prüfungspraxis weiter zu digitalisieren. Auch im Projekt erprobte mobile Lösungen mit Tablets ziehen Erneuerungsinvestitionen, Verwaltungs- und Wartungsaufwand nach sich, der aufgrund der dynamischen Nachfrage- und Preisentwicklung zu Beginn eines mehrjährigen Vorhabens nur schwer abgeschätzt werden kann.

Diese beispielhaft skizzierten Herausforderungen können die nachhaltige Implementierung digitaler Lösungen potenziell verhindern oder verzögern und strapazieren nicht zuletzt die Motivation der Akteur:innen vor Ort. Sie zeigen allerdings auch, wie wichtig es ist, schon während der Projektlaufzeit über die langfristige Bereitstellung von Ressourcen zu sprechen und Vereinbarungen zu treffen.

Digitalisierung braucht wissenschaftliche Experimentierräume.

Projekte schaffen die Möglichkeit, etwas Neues auszuprobieren. Sie ermöglichen Hochschulen, neue Gebiete wie das der Digitalisierung in Studium, Lehre und Prüfungen, mit der Erprobung von Konzepten, Praktiken und Technologien abseits des Tagesgeschäfts zu erschließen. Die Bewertung des Erfolgs einer Erprobung ist ein wesentlicher Teil von Projekten und an Hochschulen eng mit wissenschaftlichen Standards und Praktiken verknüpft.

Beispielsweise bleibt die Weiterentwicklung der eigenen Lehrpraxis an der Universität nicht auf die Integration digitaler Prüfungsformen begrenzt. Digitale Prüfungen sind ein wichtiger Bestandteil, stellen aber nur einen Baustein in einem viel umfassenderen Prozess dar. Dieser Prozess beginnt mit der Reflexion über die eigene Lehre, reicht über die Entwicklung neuer Lehrstrategien und -konzepte bis hin zur Erprobung innovativer Ansätze und der kritischen Evaluation ihrer Wirksamkeit. Im Mittelpunkt dieser Bestrebungen steht das Scholarship of Teaching and Learning (SoTL) (vgl. Huber, 2018 und 2011). SoTL ermutigt Lehrende, ihre Lehrpraxis wie ein Forschungsfeld zu betrachten: Es geht darum, systematisch zu beobachten, zu hinterfragen und zu erforschen, wie das Lehren und Lernen funktioniert (Huber, 2014). Auf dieser Grundlage können dann neue Ansätze entwickelt und implementiert werden. Durch diesen Forschungsansatz werden Lehrende zu aktiven Gestalter:innen ihrer Lehre, die kontinuierlich versuchen, ihre Methoden und Inhalte auf der Basis von Evidenz und Reflexion zu verbessern.

Die Analyse von Wirkungen – auch jenen der Digitalisierung – wird so zu einer wissenschaftlichen Herausforderung in der Entwicklung von Studium und Lehre. Um die Ergebnisse richtig einordnen, Schlussfolgerungen aufstellen und Maßnahmen entwickeln zu können, bedarf es auch stets eines interdisziplinären Austauschs der Lehrenden und Forschenden, wobei der Bildungsforschung hier eine besondere Rolle zuteilwird:

Das Sammeln von Ideen sowie das ständige Vorankommen, unterstützt durch Anreize – sei es durch das Land, die Hochschule oder Stiftungen – spielen eine entscheidende Rolle. Diese Institutionen bieten Räume zum Experimentieren, Ausprobieren und zur systematischen Erfassung von Wirkungen. Wichtig ist, dass dies nicht aus dem Bauch heraus geschieht, sondern dass die vorhandenen Kompetenzen im Bereich der Bildungsforschung genutzt werden, um genauer hinzuschauen und fundierte Erkenntnisse zu gewinnen. (Interview mit Isabell van Ackeren-Mindl, S. 11)

Die Lehrentwicklung an der Universität wird insgesamt zu einem dynamischen, forschungsbasierten Prozess, in dem Reflexion und interdisziplinärer Austausch eine zentrale Rolle spielen.

Digitalisierung schafft einen Bedarf an multiprofessioneller Zusammenarbeit.

Die Digitalisierung in Studium und Lehre trägt dazu bei, dass viele Leistungen nur noch durch das Zusammenspiel unterschiedlicher Akteursgruppen umgesetzt werden können. Eine digitale Prüfung kann hiernach als ein kollaboratives Projekt begriffen werden, an dem neben den Prüfer:innen auch die Entwickler:innen und Administrator:innen der Prüfungssysteme und ggf. hochschul- und prüfungsdidaktische Expert:innen beteiligt sind. Diese Komplexität ist dabei nicht nur bei Prüfungen anzutreffen. Auch die Erstellung von E-Learning-Kursen, der Einsatz spezifischer Aktivitäten in Lernmanagementsystemen oder die Entwicklung und Nachnutzung von Lernmaterialien ist ohne die mittlerweile etablierten Infrastrukturen und Supportteams undenkbar geworden. Lehre und Prüfungen werden nicht länger nur von einzelnen Lehrenden gestaltet, vielmehr handelt es sich zunehmend um Leistungen von multiprofessionellen (ad-hoc) Arbeitsgruppen.

Die Digitalisierung erzeugt daher einen Veränderungsdruck für Hochschulen. Sie sind darin gefordert, Strukturen und Verfahren für die multiprofessionelle Zusammenarbeit zu definieren, um neuen Bedarfen bestmöglich zu begegnen und Lehrende bzw. Prüfende mit den Herausforderungen nicht allein zu lassen. Digitalisierungsprojekte, wie PITCH, können einen Beitrag zur Definition dieser Strukturen und Verfahren leisten, gleichwohl es darüber hinausgehend auch einer kulturellen Veränderung bedarf, um die Gestaltung von Studium und Lehre stärker als eine Gemeinschaftsaufgabe zu begreifen. Ein etabliertes Netzwerk bzw. eine Community of Practice, wie sie an der Universität Duisburg-Essen für die Innovation in Lehre und Studium bereits existiert, kann dies unterstützen, da es nicht nur Lehrende miteinander vernetzt, sondern diese auch stets mit weiteren in der Lehrentwicklung aktiven Akteur:innen zusammenbringt.

Digitalisierung wirft Fragen der Chancengerechtigkeit auf.

Hochschulbildung qualifiziert Menschen für verantwortungsvolle Tätigkeiten in Wirtschaft, Wissenschaft, Politik und Gesellschaft. Sie verteilt Chancen, weshalb der Zugang zu ihr transparent, fair und gerecht gestaltet sein muss. Hochschulprüfungen müssen sich an diesen Kriterien messen lassen. Werden digitale Tools bzw. E-Assessment-Systeme genutzt, müssen Studierende vor der eigentlichen Prüfung die Möglichkeit erhalten, den Umgang mit diesen zu erlernen. Sie müssen auch bei automatisch mit Parametern erzeugten Aufgaben, wie sie beispielsweise in der Mathematik Verwendung finden, gleichwertige Anforderungen und Prüfungsbedingungen vorfinden und über den gleichen Zugang zu möglichen Hilfsmitteln verfügen.

Im Projekt PITCH standen Fragen dieser Art immer wieder auf der Tagesordnung. Das Beispiel der digitalen Prüfungen zeigt, wie schnell Prozesse der Digi-

alisierung mit Fragen der Chancengerechtigkeit einhergehen. Zugleich lädt es dazu ein, die Digitalisierung in der Hochschulbildung auch in einem grundsätzlicheren Rahmen zu betrachten und über die (künftige) Rolle der Hochschulen ins Gespräch zu kommen. Bemerkenswert ist, dass der Zugang zu Wissen und wissenschaftlichen Erkenntnissen nie so einfach war wie heute. Freie Ressourcen wie Wikipedia, Open-Access-Publikationen und (OER-)Online-Kurse können zu jeder Zeit und von jedem, mit Internet versorgten Ort abgerufen werden. Allerdings ist die Nutzung dieser Wissensressourcen auch sehr voraussetzungsvoll. Menschen müssen die Quellen hinsichtlich ihrer Herkunft, Seriosität und Plausibilität richtig einschätzen können. Sie müssen in der Lage sein, die zur Verfügung stehenden Technologien – darunter auch KI-Systeme – zielgerichtet und verantwortungsbewusst einzusetzen. Sie müssen ein Verständnis für die methodische Einordnung von Informationen mitbringen und sie müssen einen technischen Zugang haben, der ihnen alle Möglichkeiten der Informationsgewinnung verschafft und nicht nur einen Ausschnitt daraus präsentiert, wie es bei kommerziellen Tools mit limitierten Freikontingenten, Paywalls oder künstlich verlangsamer Bereitstellung häufig vorkommt. Der Prorektor sieht in der technischen Entwicklung sowohl Potenziale als auch Herausforderungen für die Hochschulen:

Die entscheidende Frage besteht darin, wo wir in fünf Jahren stehen. So wie das Handy bei Schüler:innen und Studierenden heute ubiquitär verfügbar ist und ihnen ermöglicht, digital zu partizipieren, so ist diese Entwicklung auch bei neuen Technologien, wie KI-Systemen, zu erwarten. Die Ungleichheit, die wir derzeit noch bei der Verfügbarkeit von Hardware und Lizenzen beobachten können, wird mittelfristig eine geringere Rolle spielen. Die Hochschulen sind dennoch gefordert, sich für breit verfügbare Lösungen einzusetzen und die Chancengerechtigkeit im Blick zu behalten. Dazu gehört auch, das Potenzial von KI für die Bildungsgerechtigkeit zu heben. Der Zugang zur Hochschulbildung ist in Deutschland leider noch immer mit der Herkunft, der Bildungsaspiration des Elternhauses, verknüpft. KI kann hier sogar eine emanzipierende Wirkung haben, wenn sie von den Hochschulen aktiv als Gestaltungsinstrument begriffen wird. (Interview mit Stefan Rumann, S. 13)

Projekte wie PITCH unterstützen die mit der Digitalisierung einhergehende Transformation in Studium, Lehre und Prüfungen auf vielerlei Ebenen. Sie sind einerseits ein Spiegel ihrer Zeit, was sowohl die Genese der Projektidee als auch der Umgang mit den multiplen Krisen zeigen, werfen andererseits aber auch generalisierbare Herausforderungen und Fragen auf. Der vorliegende Sammelband soll den in der Lehrentwicklung engagierten Akteur:innen dabei helfen, die Herausforderungen und Fragen im Feld der digitalen Prüfungen kennenzulernen und sich für eigene Vorhaben inspirieren zu lassen.

Die digitale Transformation bietet die Chance, Prüfungen effektiver, effizienter und gerechter zu gestalten und gleichzeitig Daten für die Erforschung der eigenen Lehrpraxis zu gewinnen. Die Beiträge in diesem Band thematisieren Herausforderungen, Erfahrungen, die Entwicklung geeigneter Prüfungsformate und die Umsetzung digitaler Prüfungen im Lehralltag und darüber hinaus. Nachfolgend werden die einzelnen Beiträge kurz skizziert und vorgestellt.

6 Die Beiträge in diesem Sammelband

Der erste Teil des Buches widmet sich dem Thema „*Threshold Concepts and Decoding the Disciplines*“ und vereint eine Reihe von Beiträgen, die sich mit der Integration und Weiterentwicklung von Schwellenkonzepten in verschiedenen akademischen Disziplinen befassen. Diese Konzepte, auch als „*Threshold Concepts*“ bekannt, sind zentrale Ideen in einem Fachgebiet, die für Studierende oft schwer zu durchdringen sind, aber eine transformative Wirkung auf deren Verständnis haben können.

Den Auftakt macht *Cornelia Kenneweg* mit ihrem Beitrag „*Threshold Concepts als Reflexionsanlass: Transformation und Variabilität in Lernprozessen wahrnehmen und begleiten*“. Sie beschreibt die Bedeutung von Threshold Concepts in der Hochschuldidaktik und beleuchtet, wie diese Konzepte genutzt werden können, um Lernprozesse zu verbessern. Kenneweg, die im Rahmen des Projekts PITCH mit verschiedenen Lehrendenteams zusammengearbeitet hat, reflektiert über die didaktische Weiterentwicklung von Lehrveranstaltungen und die daraus resultierenden Veränderungen im Lehr- und Lernprozess.

Ulrike Berendt, *Martin Krybus* und *Till Behnenburg* stellen in ihrem Beitrag „*Integriertes Lehr-, Lern- und Prüfkonzept für eine politikwissenschaftliche Einführungsveranstaltung*“ ein umfassendes Konzept für die Einführungsvorlesung „Grundlagen der Politikwissenschaft“ an der Universität Duisburg-Essen vor. Sie kombinieren Threshold Concepts mit dem Decoding the Disciplines-Ansatz und dem Prinzip des Constructive Alignments, um strukturelle und inhaltliche Hürden im Lernprozess zu überwinden. Durch die Integration digitaler Prüfungen auf der Moodle-Plattform, unterstützt durch semesterbegleitende Übungen und Zusatzaufgaben, wird die Kompetenzorientierung der Studierenden gestärkt.

In dem Beitrag „*Schwellenkonzepte und Lehrinnovationen in fachlichen und fachdidaktischen Veranstaltungen im Lehramtsstudium Mathematik*“ analysieren *Andreas Büchter*, *Dana Eilers* und *Florian Schacht* am Beispiel einer fachlichen und einer fachdidaktischen Lehrveranstaltung typische Schwellenkonzepte im Lehramtsstudium Mathematik und gehen auf erprobte Lehrinnovationen ein, die die Studierenden beim Überwinden der Schwellen unterstützen können. Sie nehmen die Schwellenkonzepte (geometrisches) Beweisen sowie Umgang mit Fachtexten in den Blick und beziehen in ihren Lehrinnovationen Reflexionen

über den lernförderlichen Einsatz digitaler Medien und die Gestaltung kompetenzorientierter Prüfungen ein.

Ruth M. Ingendoh und *Angela Heine* beschäftigen sich in ihrem Beitrag „*Gestaltung lern(enden)zentrierter Veranstaltungen im bildungswissenschaftlichen Studium der Bachelor-Lehramtsstudiengänge der UDE*“ mit der Neuausrichtung von Lehrveranstaltungen im Bereich Bildungswissenschaften. Im Fokus steht die Neugestaltung des Moduls Psychologie, bei dem Schwellenkonzepte identifiziert und durch kompetenzorientierte Ansätze ersetzt wurden. Digitale Ressourcen wie Moodle und interaktive Werkzeuge wurden genutzt, um die Lehre praxisnäher und interaktiver zu gestalten.

Thomas Mühlbauer und *Johanna Lambrich* thematisieren in ihrem Beitrag „*Kompetenzorientiert digital prüfen in der Sport- und Bewegungswissenschaft*“ die Entwicklung digitaler Prüfungsformate im Bereich Sportwissenschaft. Sie erläutern, wie Schwellenkonzepte identifiziert und in digitale Lern- und Selbstüberprüfungsangebote integriert wurden, die in die Moodle-Plattform eingebunden sind.

Nguyen Minh Salzmann-Hoang und *Tobias Schroedler* untersuchen in ihrem Beitrag „*Decoding Deutsch als Zweitsprache: Was macht das Fach DaZ schwer (zu prüfen)?*“ die Herausforderungen studentischen Lernens im interdisziplinären Modul „Grundlagenwissen Deutsch als Zweitsprache“ mit Hilfe des Decoding the Disciplines-Ansatzes. Sie gehen kritisch auf die gängige Praxis von Multiple-Choice-Prüfungen ein und erläutern laufende und mögliche Maßnahmen, die Prüfungsgestaltung besser auf die Lernziele abzustimmen.

Der zweite Teil des Buches, „*Digital üben, prüfen und bewerten*“ widmet sich der technischen Umsetzung digitaler Prüfungen und den dabei gewonnenen Erkenntnissen aus acht Lehrprojekten.

Zu Beginn des zweiten Teils steht der Beitrag „*Das Beste aus Online und Präsenz – Open-Book-Prüfungen mit handschriftlichen Lösungen und digitalem Bewertungsworkflow*“ von *Mathias Magdowski*, in dem er die Entwicklung und Implementierung alternativer Prüfungsformate während und nach der Coronapandemie thematisiert. Magdowski schildert seine Erfahrungen mit einer Bandbreite an Prüfungsformaten, die von Open-Book-Prüfungen mit handschriftlichen Lösungen bis hin zu digitalen Bewertungsworkflows reichen. Er schließt seinen Artikel mit einem positiven Ausblick auf die Weiterentwicklung digitaler Prüfungsformate, betont jedoch auch die Notwendigkeit, Medienbrüche zu vermeiden und den administrativen Aufwand weiter zu optimieren.

Im weiteren Verlauf berichten die Lehrenden über ihre technischen Lösungen, insbesondere unter Einsatz des Learning Management Systems Moodle und des E-Assessment-Systems JACK sowie über die Anwendung weiterer Tools wie beispielsweise tet.folio.

Melanie Schypula und *Christoph Olbricht* stellen in ihrem Beitrag „*Digitales Lehren und Lernen mit Moodle und JACK*“ die beiden Systeme zur Durchführung elektronischer Prüfungen vor und analysieren die Unterschiede zwischen

innen. Die Autor:innen gelangen zu dem Schluss, dass eine kontinuierliche Weiterentwicklung der Systeme notwendig ist, um den steigenden Anforderungen gerecht zu werden. Besonders wichtig erscheinen dabei eine intuitive Bedienbarkeit, umfassende Dokumentation und gezielte Schulungsangebote, um Lehrende optimal zu unterstützen.

Das erste Lehrprojekt wird im Beitrag „*Auf dem Weg zur Innovation: Digitale Präsenzklausuren am Beispiel der Englischen Sprachwissenschaft*“ von *Frauke Milne, Birte Bös* und *Jana Eismann* vorgestellt. Hier beschreiben die Autorinnen die Implementierung und Herausforderungen digitaler Klausuren im Rahmen des Einführungsmoduls „Introduction to Linguistics“. Sie beleuchten die organisatorischen Rahmenbedingungen sowie die Herausforderungen, die mit den unterschiedlichen Prüfungsformaten einhergehen. Ihr Fokus liegt auf der Optimierung der Aufgabengestaltung und der Weiterentwicklung digitaler Prüfungen. Es wird deutlich, dass die Digitalisierung von Prüfungen nicht nur organisatorische und didaktische Verbesserungen erfordert, um Chancengleichheit und Qualität sicherzustellen, sondern auch nachhaltiger Strukturen und personeller, zeitlicher und räumlicher Ressourcen bedarf.

Im Beitrag „*Digitale Prüfwerkzeuge in der Baudynamik – Ein Erfahrungsbericht*“ von *Carolin Birk* und *Ajay Kumar Pasupuleti* untersuchen die Autor:innen die Erprobung des digitalen Prüfungssystems JACK im Modul „Baudynamik“ des Masterstudiengangs Bauingenieurwesen. Nach einer Analyse traditioneller Prüfungsformate leiten sie Anforderungen für digitale Prüfungen ab, bevor sie JACK in verschiedenen Szenarien erproben. Sie erkennen, dass digitale Prüfungen im Ingenieurwesen Potenzial haben, allerdings erhebliche Vorbereitungszeit sowie didaktische und technische Anpassungen erfordern.

In seinem Beitrag „*Aktivierung von Studierenden – Ein Erfahrungsbericht zu den Herausforderungen interaktiver und digitaler Lehre*“ präsentiert *Jochen Gönsch* die kontinuierliche Weiterentwicklung der Lehrveranstaltung „Software Skills“. Er setzt verschiedene digitale Unterstützungsangebote wie Bildschirmvideos mit Interaktionselementen, ein Wiki und ein Gamification-Konzept ein. Abschließend stellt er fest, dass trotz dieser digitalen und gamifizierten Elemente die Eigenverantwortung und das Engagement der Studierenden entscheidend für den Erfolg der Veranstaltung sind.

Im Beitrag „*Aspekte des digitalen Lehrens und Prüfens: Überblick und neuere Entwicklungen*“ schildern *Christoph Hanck* und *Jens Klenke* anhand ihrer Erfahrungen aus den Lehrveranstaltungen „Deskriptive und Induktive Statistik“, wie digitale Lehrangebote in JACK gestaltet werden und welche Vorteile sich daraus insbesondere für eine moderne Unterrichtsgestaltung in der Statistik ergeben. Sie betonen, dass die eingesetzten verschiedenen Aufgabentypen den Studierenden helfen, Konzepte zu verstehen, anstatt sich nur auf das Auswendiglernen von Ergebnissen zu verlassen. Die Autoren heben hervor, dass die kontinuierliche Auseinandersetzung mit dem Lernstoff durch regelmäßige Testate entscheidend für den Prüfungserfolg ist.

Justin Timm, Nora Lösing, Christian Johannes und Philipp Schmiemann berichten in ihrem Beitrag „*Entwicklung eines formativen Assessments zur Förderung prozeduralen und konditionalen biomedizinischen Wissens*“ vom Einsatz digitaler formativer Assessments, die darauf abzielen, die prozess- und handlungsorientierten Kompetenzen der Studierenden zu stärken. Ihr theoretischer Rahmen stützt sich dabei auf verschiedene Modelle und Konzepte der Lehrkräfteprofessionalisierung, insbesondere auf das fachdidaktische Wissen (PCK).

Heike Theyßen, Lisa Danzig, Sebastian Haase und Hermann Nienhaus erläutern in ihrem Beitrag „*Hilfe, ich komme nicht weiter! Digitale Lösungs- und Visualisierungshilfen für Übungs- und Klausuraufgaben der Experimentalphysik*“ die Einführung sogenannter „gestufter Lernhilfen“ für Übungs- und Klausuraufgaben. Diese sollen Studierende bei der Bearbeitung typischer Probleme in der Experimentalphysik unterstützen. Für die technische Umsetzung wurde die Plattform „tet.folio“ der FU Berlin genutzt, die es ermöglicht, diese Hilfen digital bereitzustellen und ihre Nutzung während der Klausuren nachzuverfolgen. Erste Erfahrungen zeigen, dass Studierende die Hilfen unterschiedlich und selektiv nach ihrem individuellen Bedarf nutzen.

In ihrem Beitrag „*Variation von Aufgabenstellungen zu Modellierung in der Informatik*“ beschreiben *Janis Voigtländer* und *Marcellus Siegburg* die Entwicklung und Anwendung eines Frameworks zur automatischen Generierung und Bewertung von Aufgaben in der Informatiklehre. Das Framework verwendet die E-Learning-Plattform Autotool, um verschiedene Aufgabentypen zu erstellen, die auf unterschiedlichen Kompetenzniveaus basieren, wie sie in der Taxonomie von Bloom definiert sind. Die automatisch generierten Aufgaben bieten den Lernenden die Möglichkeit, ihr Wissen und Verständnis auf verschiedenen Ebenen zu testen.

Abschließend thematisieren *Maik Walpuski, Michael Giese* und *Katrin Schüßler* in ihrem Beitrag „*Erste Schritte auf dem Weg zu einer digitalen Lern- und Prüfungsumgebung für die organische Chemie*“ die Entwicklung und Erprobung eines digitalen Tools zur Darstellung und Auswertung von Molekülen in der organischen Chemie. Sie stellen fest, dass die digitale Bearbeitung von Aufgaben zusätzliche technische Herausforderungen mit sich bringt, was zu einer geringeren Erfolgsquote im Vergleich zur analogen Bearbeitung führte. Die Autor:innen betonen die Notwendigkeit weiterer technischer Verbesserungen, um digitale Prüfungen effizienter und zuverlässiger zu gestalten.

Der zweite Teil schließt mit dem Beitrag „*Generative KI und Prüfungen: Erste Erfahrungen und offene Fragen*“ von *Michael Striewe*. Er untersucht die Auswirkungen und Herausforderungen der Nutzung generativer KI, insbesondere von Modellen wie ChatGPT, im Hochschulkontext, speziell bei Prüfungen, und erörtert unterschiedliche Szenarien, in denen generative KI in Prüfungen eingesetzt werden kann.

Im dritten Teil des Buches „*Rechtliche und studentische Perspektiven auf das (digitale) Prüfen*“ widmet sich *Henrik Noll* den rechtlichen Aspekten der Digita-

lisierung von Prüfungen in seinem Beitrag „*Die rechtliche Beratung im PITCH-Projekt – Ein Erfahrungsbericht*“. Er erläutert, wie rechtliche Unsicherheiten im Projekt PITCH geklärt wurden und betont die Notwendigkeit, zukünftige Herausforderungen wie den Einsatz Künstlicher Intelligenz, rechtlich abzusichern.

Maiken Bonnes und *Kilian Schmitt* (in Zusammenarbeit mit *Colin Corbach*, *Ronja Hensle*, *Marc Ludwig*, *Sophie Stotz Anido* und *Frederike Welzel*) stellen in ihrem Beitrag „*Die Student Voice Group im Projekt PITCH – Ein Praxisbericht*“ die Arbeit der Student Voice Group (SVG) vor. Der Beitrag beleuchtet die Zusammenarbeit mit den Studierenden im Rahmen des Projekts und liefert Einblicke in die wichtigsten Erkenntnisse und Standpunkte der Studierenden bzgl. anonymisiertem Prüfen an Hochschulen, Täuschungsversuchen und innovativen Prüfungsformaten. Abschließend werden sowohl die Vorteile als auch die Herausforderungen der gemeinsamen Projektarbeit erläutert.

Dieser Sammelband bietet eine umfassende Auseinandersetzung mit innovativen Lehr- und Prüfungskonzepten im Kontext der Hochschulbildung. Durch die interdisziplinären Perspektiven und praxisnahen Beispiele wird deutlich, wie essenziell die kontinuierliche Weiterentwicklung didaktischer Konzepte und technischer Werkzeuge für den Lernerfolg der Studierenden ist. Der Band stellt nicht nur theoretische Überlegungen an, sondern liefert auch konkrete Impulse für die praktische Anwendung und Reflexion in der Lehre. Die enge Verzahnung von Forschung und Praxis ermöglicht es, sowohl bestehende Herausforderungen zu erkennen als auch zukunftsweisende Lösungsansätze zu entwickeln. So bietet der Sammelband eine wertvolle Ressource für Lehrende, Forschende und alle, die sich mit der Weiterentwicklung der Hochschulbildung auseinandersetzen.

Wir bedanken uns bei allen Autor:innen, die die Perspektivvielfalt, die sich auch im Projekt als große Ressource herausgestellt hat, in diesen Sammelband eingebracht haben. Den Leser:innen wünschen wir interessante Erkenntnisse und eine angenehme Lektüre.

Nicole Auferkorte-Michaelis, Maiken Bonnes,
Patrick Hintze und Julia Liebscher

Literatur

- Huber, L. (2011). Forschen zum Lehren – Hochschuldidaktische Studien aus den Fächern heraus. In M. Eger/B. Gondani/R. Kröber (Hrsg.), *Verantwortungsvolle Hochschuldidaktik. Gesellschaftliche Herausforderungen, Nachhaltigkeitsanspruch und universitärer Alltag* (S. 75–92). Berlin: Lit Verlag.
- Huber, L. (2014). Scholarship of Teaching and Learning: Konzept, Geschichte, Formen, Entwicklungsaufgaben. In: L. Huber/A. Pilniok/R. Sethe/B. Szczyrba/M. Vogel (Hrsg.), *Forschendes Lehren im eigenen Fach. Scholarship of Teaching and Learning in Beispielen* (S. 19–36). Bielefeld: Bertelsmann.
- Huber, L. (2018). SoTL weiterdenken! Zur Situation und Entwicklung des Scholarship of Teaching and Learning (SoTL) an deutschen Hochschulen. In: *Das Hochschulwesen*, 66 (1/2), 33–41.
- Meyer, J. H. F./Land, R. (2003). Threshold Concepts and Troublesome Knowledge: Linkages to Ways of Thinking and Practicing within the Disciplines. In: C. Rust (Hrsg.), *ISL 10: Improving Student Learning: Theory and Practice Ten Years On* (S. 412–424). Oxford: Oxford Brooks University.
- Middendorf, J./Shopkow, L. (2018). *Overcoming Student Learning Bottlenecks. Decode the Critical Thinking of Your Discipline*. New York: Routledge.
- Pellert, A. (1999). *Die Universität als Organisation. Die Kunst, Experten zu managen*. Wien, Köln, Graz: Böhlau.
- Senge, P. (1997). *Die fünfte Disziplin. Kunst und Praxis der lernenden Organisation*. Stuttgart: Cotta.
- Salden, P./Lordick, N./Wiethoff, M. (2023). KI-basierte Schreibwerkzeuge in der Hochschullehre. In: P. Salden/J. Leschke (Hrsg.), *Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung* (S. 4–21). <https://doi.org/10.13154/294-9734>
- Senge, P./Kleiner, A./Smith, B./Roberts, C./Ross, R. (2008). *Das Fieldbook zur Fünften Disziplin* (5. Aufl.). Stuttgart: Schäffer-Poeschel.
- Stammen, K.-H./Ebert, A. (2021). *Noch online? Studierendenbefragung zur medientechnischen Ausstattung – Gesamtbericht: Ergebnisse der universitätsweiten UDE-Umfrage im Sommersemester 2020*. <https://nbn-resolving.org/urn:nbn:de:hbz:464-20210914-101738-0>

Threshold Concepts und Decoding the Disciplines

Threshold Concepts als Reflexionsanlass: Transformation und Variabilität in Lernprozessen wahrnehmen und begleiten

Cornelia Kenneweg

1 Ausgangspunkt und Einleitung

Im Laufe ihres Studiums begegnen Studierende fachlichen Inhalten und Konzepten, die ihnen neue Perspektiven eröffnen oder für sie neue Zusammenhänge herstellen und so das Potenzial haben, ihr Denken weitreichend zu verändern. Solche Inhalte und Konzepte sind einerseits entscheidend, um Fachkompetenzen überhaupt entwickeln zu können, andererseits aber häufig schwer zu lernen, unter anderem, weil dabei bisherige Überzeugungen überdacht werden müssen oder Vertrautes neu zu betrachten ist. In einführenden Modulen zur Linguistik lernen Studierende beispielsweise, Sprache zum Untersuchungsgegenstand bzw. zum Gegenstand von Analyse und Reflexion zu machen (vgl. Beiträge Milne/Bös/Eismann und Salzmann-Hoang/Schroedler in diesem Band), was für viele eine ungewohnte Perspektive auf etwas eigentlich Vertrautes, nämlich Sprache, ist. In der Biologie ist es wichtig, gut zu verstehen, was genau biologische Variabilität ist, um Phänomene in Teilgebieten wie Genetik oder Evolutionstheorie zu erklären (Batzli et al., 2016; Walck-Shannon et al., 2019). Solche Fachinhalte werden wegen ihres Schwellencharakters im Lernprozess auch als Threshold Concepts bezeichnet.

Den Begriff Threshold Concepts haben Jan H. F. Meyer und Ray Land gemeinsam mit Projektbeteiligten im Rahmen eines Projektes zur fachbezogenen Verbesserung hochschulischer Lernumgebungen (*Enhancing Teaching and Learning Environments in Undergraduate Courses* oder ETL-Projekt¹) geprägt (Meyer/Land, 2003). In diesem Projekt wurden anhand von ausgewählten Fächern (u. a. Biologie, Ökonomie und Geschichte) und im engen Austausch zwischen Hochschuldidaktiker*innen und Fachlehrenden mehrerer Universitäten disziplinäre Denk- und Arbeitsweisen, also *ways of thinking and practising*, (Hounsell/Hounsell, 2007; Hounsell/Anderson, 2009) und ihre Relevanz für die Lehre betrachtet. Kernziel war es, Lernumgebungen von hoher Qualität zu identifizieren und daraus Vorschläge für die Entwicklung guter Lehre in den Disziplinen zu erarbeiten. Die Projektbeteiligten stellten unter anderem fest, dass die didaktischen Modelle

1 Die Zwischenberichte und der Abschlussbericht sowie weitere Publikationen können nach wie vor über die Projekthomepage abgerufen werden: <http://www.etl.tla.ed.ac.uk>. (letzter Zugriff: 10.11.2023).

und Konzepte, von denen sie zunächst ausgingen, wie beispielsweise das verbreitete Constructive Alignment, allein nicht ausreichen, um alle Aspekte der Lernumgebungen hoher Qualität ausreichend zu erfassen (z. B. Hounsell/Hounsell, 2007, S. 100). Die Idee der Threshold Concepts war eines der Projektergebnisse, mit denen Lehrentwicklung fachsensibel erweitert werden sollte.

Lehrentwicklungsprojekte bringen nicht nur neue didaktische Ideen hervor, sie bieten auch einen Rahmen, um mit vorhandenen didaktischen Ansätzen zu experimentieren. So war im Arbeitsfeld „Disziplin und Didaktik“ des Innovationsprojekts PITCH von vornherein vorgesehen, mit Threshold Concepts und Decoding the Disciplines zwei im deutschsprachigen Raum erst seit einigen Jahren diskutierte und erprobte Ansätze fachsensibler Lehrentwicklung aufzugreifen. Zu diesem Zweck wurde bei verschiedenen Gelegenheiten, beispielsweise beim Kick-Off am 10./11. Februar 2022 und der Veranstaltung PITCHen im Hof am 28. Juni 2022, in diese Ansätze eingeführt. Außerdem gab es das Angebot, mit Unterstützung von mehreren Coaches, darunter die Verfasserin dieses Beitrags, in Coachinggesprächen Threshold Concepts zu identifizieren beziehungsweise Decoding-Interviews zu führen. Dieses Angebot wurde in unterschiedlichem Maße und unterschiedlicher Form von den Teams der Teilprojekte angenommen. Wie die didaktischen Impulse in einigen Teilprojekten aufgegriffen wurden, zeigen die entsprechenden Beiträge dieses Bandes.

Der vorliegende Beitrag führt kurz in das Threshold Concept Framework (nachfolgend als TCF abgekürzt) als hochschuldidaktischen Ansatz ein. Zudem diskutiert er – literatur- und erfahrungsbasiert² – das Potenzial des Ansatzes, systematische, fachbezogene didaktische Lehrreflexion anzuregen und Transformation sowie Variabilität in Lernprozessen sichtbar zu machen. Es soll verdeutlicht werden, wie die Identifikation und Reflexion von Threshold Concepts die Gestaltung kompetenzorientierter und chancengerechter Lehr- und Prüfungsformate unterstützen kann.

2 Was sind Threshold Concepts?

Als Threshold Concepts oder Schwellenkonzepte werden fachliche Konzepte bezeichnet, deren Aneignung im Lernprozess mit einer Veränderung der Sicht auf den Lerngegenstand, das Fach oder gar die Welt einhergeht. Sie werden häufig metaphorisch als Tore oder Portale beschrieben, “opening up a new and previously inaccessible way of thinking about something“ (Meyer/Land, 2003, S. 1).

Threshold Concepts sind also *transformativ*, weil sich durch die Auseinandersetzung mit ihnen im Lernprozess Perspektiven, Denk- und Handlungsmöglich-

2 Das heißt, eine empirische Begleitforschung zur Integration der hochschuldidaktischen Impulse erfolgte nicht.

keiten der Lernenden verändern. Dieser Lernprozess ist *irreversibel*, das heißt, dass diese neuen Perspektiven und Denkweisen nicht mehr ohne Weiteres wieder vergessen oder ignoriert werden können. So gilt etwa ‚legal reasoning‘, also die Fähigkeit, juristisch zu begründen und zu argumentieren, als Threshold Concept im Recht. Wie ein*e Jurist*in zu denken, ist eine Fähigkeit, die kaum wieder abzulegen ist, wenn man sie einmal erlernt hat (Åkerlind et al., 2011, S. 15). Threshold Concepts gelten als *integrativ*, in dem Sinne, dass sich durch sie neue Zusammenhänge erschließen und mehrere Basiskonzepte so verbinden, dass ein tieferes Verstehen möglich wird (Entwistle, 2008). Wer beispielsweise verinnerlicht hat, Sprache als Gegenstand von Analyse und Reflexion zu begreifen, kann etwa als Lehrkraft im Umgang mit Mehrsprachigkeit leichter linguistische, didaktische und bildungswissenschaftliche Perspektiven verschränken (vgl. Beitrag Salzmann-Hoang/Schroedler in diesem Band). Einige Autor*innen charakterisieren Threshold Concepts auch als *begrenzend* (*bounded*), weil sich durch sie auch disziplinäre Gebiete abstecken lassen, oder als *diskursiv* (*discursive*), weil mit dem Überschreiten der Schwellen auch eine Erweiterung der fachsprachlichen Ausdrucksfähigkeit einhergeht.

Kennzeichnend für Threshold Concepts ist zudem, dass es sich bei ihnen in der Regel um schwieriges Wissen (*troublesome knowledge*) handelt (Meyer/Land, 2003). Der Terminus *troublesome knowledge* geht auf David Perkins zurück (Perkins, 1999; Perkins, 2006), der sich vor dem Hintergrund konstruktivistischer Positionen zum Lehren und Lernen mit unterschiedlichen Arten von Wissen auseinandersetzt, die aus jeweils spezifischen Gründen Schwierigkeiten beim Lernen bereiten. Lerngegenstände sind beispielsweise *troublesome knowledge*, wenn sie konterintuitiv, ungewohnt oder konzeptionell komplex sind oder das Hinterfragen bisheriger Auffassungen erfordern (Perkins, 2006).

Um zu beschreiben, dass die Veränderungsprozesse beim Überschreiten von Schwellen individuell unterschiedlich und nicht linear verlaufen, wird im TCF der Begriff der Liminalität genutzt. Es werden präliminale, liminale und postliminale Phasen im Lernprozess unterschieden, die jeweils von einer hohen Variabilität gekennzeichnet sind, je nachdem, mit welchen Vorerfahrungen, Voraussetzungen, aber auch Motivationen und Intentionen Lernende an ein Threshold Concept herangehen (Calduch/Ratray, 2022; Land/Meyer, 2010).

Praktisch alle Publikationen zu Threshold Concepts verweisen auf die Kerncharakteristika, und es gibt immer wieder Bemühungen, noch besser zu klären, was unter Threshold Concepts zu verstehen ist und wie sie sich am besten identifizieren lassen. Die methodologische Offenheit des TCF bringt es aber mit sich, dass sich trotz der fast ritualisierten Verweise auf grundlegende Texte (insbesondere: Land/Meyer, 2003) verschiedene Interpretationen, Adaptionen und Verwendungsweisen des Frameworks herausgebildet haben, je nachdem mit welchem Ziel und in welchem Zusammenhang der Ansatz aufgegriffen und angewandt wird.

Klar ist: Die Identifikation von und das Arbeiten mit Threshold Concepts erfordert Aushandlung und fördert im besten Fall Dialog (Lucas/Mladenovic,

2007), und zwar Dialog von Lehrenden untereinander, von Lehrenden und Studierenden sowie von Lehrenden, Studierenden und Hochschuldidaktiker*innen. Student Voice Groups, wie es sie auch im PITCH-Projekt gibt, das Einholen von studentischem Feedback (vgl. Beitrag von Milne/Bös/Eismann in diesem Band) oder das direkte Einbeziehen von Studierenden in die Identifikation möglicher Schwellen (vgl. Beitrag Mühlbauer/Lambrich in diesem Band) sind Möglichkeiten, Studierende an diesem Dialog zu beteiligen. Ein solcher Austausch sollte eine Verständigung über Erwartungen und das Explizieren von impliziten Annahmen über Hochschullehre, über relevante disziplinäre Inhalte und über das Lernen im Fach und dabei auftretende Hürden sowie (potenzielle) Durchbrüche und Erkenntnismomente (Kenneweg/Wiemer, 2021) einschließen.

3 Threshold Concepts im hochschuldidaktischen Diskurs

Über die letzten Jahre hat sich vor allem im englischsprachigen Raum eine breite Publikationstätigkeit zum TCF entwickelt. Neben konzeptionellen Weiterentwicklungen des Ansatzes und Kritik daran, sind es vor allem Beispiele zur Nutzung von Threshold Concepts in einer Vielzahl von Fächern, die diese Publikationstätigkeit ausmachen. Im deutschsprachigen Raum finden die Schwellenkonzepte erst langsam Eingang in den Diskurs über Hochschuldidaktik und Hochschullehre – anders als das etwa zur gleichen Zeit entstandene *Decoding the Disciplines* (Pace, 2017). Beide Ansätze haben gemeinsam, dass es ihnen darum geht, fachliche Expertise zu entschlüsseln und Lernschwierigkeiten zu verstehen, um Lernprozesse von Studierenden nachhaltig zu fördern. *Decoding the Disciplines* wird gelegentlich sogar als Methode verstanden, um Threshold Concepts zu identifizieren und besser zu verstehen, worin für Studierende Herausforderungen bestehen (Shopkow, 2010).

Was Threshold Concepts für viele Lehrende attraktiv macht, ist das Erschließen von Fachwissenschaft und disziplinären Gegenständen mit der Aufmerksamkeit für individuelle Lernprozesse und die subjektive Aneignung von Wissenschaft zu verbinden (Cousin, 2010; Kenneweg/Wiemer, 2022). Gerade für die Prüfungsgestaltung ist es wichtig, Lerngegenstände und Lernprozesse so zu analysieren, dass Prüfungen und andere Formen des Feedbacks konzipiert werden, die für Lehrende wie Lernende sichtbar machen, ob das Gelernte ausreichend verstanden wurde, um später sinnvoll darauf aufbauen zu können (Land/Meyer, 2010). So wurde auch in den Gesprächen mit den Lehrendenteams von PITCH häufig die Frage diskutiert, wie statt eines oberflächlichen Lernens von deklarativem Wissen ein ‚tieferes‘ Verstehen und Anwenden fachwissenschaftlicher Inhalte gefördert und geprüft werden kann. Im PITCH-Lehrendenteam aus den Sport- und Bewegungswissenschaften ergab beispielsweise die Analyse von Klausurergebnissen zum ‚Dehnungs-Verkürzungs-Zyklus‘, dass nur wenige Studierende ein ausreichendes Verständnis dieses Konzepts erreicht hatten. Daraus

folgten didaktische Überlegungen dazu, wie Studierende mit mehr Variation in der präliminalen Phase – etwa durch Eigenerfahrungen in Bewegungsübungen – an die Schwelle herangeführt werden können (vgl. Beitrag Mühlbauer/Lambrich in diesem Band).

Eine Herausforderung ist dabei, dass fachliche Expertise sich erst im Umgang mit und im Zusammenspiel von verschiedenen Konzepten entfaltet. Über Threshold Concepts wird dabei ein Zugang zu dem eröffnet, was Perkins das „underlying game“ (Perkins, 2006) oder *episteme* eines Faches nennt, also den fachbezogenen Formen der Erkenntnisgewinnung und -kommunikation. Wie dieses Zusammenspiel funktioniert und wie sich Disziplinen dabei auch im Verständnis von Threshold Concepts unterscheiden, wird im fachübergreifenden Austausch besonders leicht deutlich (Carmichael, 2010), weshalb sich der Ansatz auch für interdisziplinäre Lehre gut eignet.

Threshold Concepts markieren häufig die Grenzen des gesicherten Wissens einer Disziplin, was für Lehrende wie Lernende herausfordernd sein kann. Denn: „The troublesomeness of a threshold concept may include its own inherent instability“ (Cousin, 2008, S. 263). Gerade im Rahmen von Curriculumentwicklung kann die Identifikation und Diskussion von Threshold Concepts auch disziplinäre Machtfragen und Mechanismen der Ein- und Ausschließung aufdecken (Ratray, 2023).

Schwierig sind außerdem oft weniger die Konzepte selbst als vielmehr die Lernerfahrung, die Studierende im Rahmen fachwissenschaftlicher Erkenntnisprozesse mit ihnen machen (Stopford, 2021), oder wie Julie Ratray es ausdrückt: „it is the experience that is threshold, not simply the knowledge“ (Ratray, 2023, o. S.).

Dass es in vielen Disziplinen keine Einigung darüber gibt, welche fachlichen Konzepte als Schwellenkonzepte gelten, hat Fragen bezüglich ihres epistemologischen Status aufgeworfen. So kritisiert Håkan Salwén (2019), dass der Begriff Threshold Concepts nicht gut genug definiert sei, um empirisch festzustellen, ob ein Konzept ein Threshold Concept ist. Dies mache das TCF als Theorie wissenschaftlich weitgehend unbrauchbar. Calduch/Ratray entgegnen auf diese Art der Kritik, dass ihr eine reduktionistische Vorstellung von Wissen zugrunde liege, die nicht ausreichend beachte, dass es in einem Fach mehrere anerkannte *ways of thinking and practicing* geben kann (Calduch/Ratray, 2022). Die Diskussion von Threshold Concepts sollte vielmehr zur Reflexion darüber führen, was als Wissen zählt und wie es jeweils bewertet wird (Ratray, 2023).

Obwohl ich mich der Entgegnung von Calduch und Ratray grundsätzlich anschließe, ist die Kritik an der begrifflichen Unschärfe mit Blick darauf ernst zu nehmen, wie Threshold Concepts für die Lehr- und Prüfungsgestaltung genutzt werden. Denn mit den Fragen nach der eindeutigen Identifizierbarkeit von Schwellenkonzepten ist noch eine weitere Kritik am TCF verbunden, die Schwierigkeiten betrifft, das Überschreiten der mit den Threshold Concepts verbundenen Schwellen zu messen und für didaktische Zwecke zu operationalisieren

(Nicola-Richmond et al., 2018). Herausforderungen bei der Operationalisierung erschweren es, Threshold Concepts unmittelbar für die Prüfungsentwicklung zu nutzen. Eine Studie zum Threshold Concept ‚biologische Variation‘ zeigte, dass sich zwar ein Überschreiten der Schwelle grundsätzlich nachweisen lässt, nicht jedoch der exakte Zeitpunkt (Walck-Shannon et al., 2019). Die Autor*innen untersuchten studentische Erklärungen von biologischer Variation zu unterschiedlichen Zeitpunkten in einem dreisemestrigen Laborcurriculum. Sie konnten unter anderem Unterschiede zwischen verschiedenen Dimensionen des Threshold Concepts herausarbeiten: Während es den meisten Studierenden zum Beispiel rasch leichtfiel, biologische Fachterminologie zu benutzen (*diskursiv*), gelang es vielen auch nach mehreren Semestern nicht gut, genetische, phänotypische Variation und Variation auf der Ebene von Populationen und Arten angemessen zu verknüpfen (*integrativ*) (Walck-Shannon et al., 2019).

Grundsätzlich sollte die Identifikation von beziehungsweise die Verständigung über Threshold Concepts im Rahmen von Curriculum- oder zumindest Modulentwicklung stattfinden und damit dem Prüfungsdesign vorausgehen. Eine solche curriculare Verankerung erleichtert die Entscheidung, ob, beziehungsweise inwieweit Threshold Concepts gezielt durch Prüfungen adressiert werden können und sollen. Dazu erscheint es sinnvoll, Threshold Concepts und ihre Aneignung nicht isoliert zu betrachten, sondern mit dem für die Prüfungsentwicklung maßgeblichen Paradigma der Kompetenzorientierung und mit hochschuldidaktisch wirkmächtigen Prinzipien wie dem Constructive Alignment gemeinsam zu diskutieren.

4 Threshold Concepts, Kompetenzorientierung und Constructive Alignment

Das Paradigma der Kompetenzorientierung hat durch die Bologna-Reform und durch bildungspolitische Rahmungen wie den Hochschulqualifikationsrahmen HQR (HRK/KMK, 2017) starke normative Kraft entfaltet und bestimmt weitreichend, wie an Hochschulen und Universitäten Studiengänge, Module und damit auch Prüfungen gestaltet werden. Daher liegt es nahe, im Zusammenhang mit einem Projekt zum akademischen Prüfen bei der Diskussion von Threshold Concepts auch an der Kompetenzorientierung anzuknüpfen. Das TCF ist einerseits gut mit Kompetenzorientierung kompatibel, lässt sich andererseits aber auch an Diskussionen anschließen, die über das Kompetenzparadigma hinausweisen oder mit dem Paradigma verknüpfte Prinzipien wie das Constructive Alignment kritisieren.

Für die Entwicklung von Hochschullehre und insbesondere von Prüfungen erhofft man sich von der Kompetenzorientierung unter anderem, dass sie mit einer genauen Beschreibung von Lernzielen beziehungsweise angestrebten Lernergebnissen (Intended Learning Outcomes) einhergeht. Um Expert*innenwis-

sen zielgruppengerecht aufzubereiten, braucht es – so die Annahme – ein gut formuliertes „Learning Outcome, das die Grundlage für alle Entscheidungen in Bezug auf die Semesterplanung, die Prüfungsaufgaben und die Bewertungskriterien darstellt“ (Wunderlich/Szczyrba, 2018, S. 4). Dieser mit dem Paradigma der Kompetenzorientierung verknüpfte Anspruch ist nicht leicht einzulösen. In der Praxis der Hochschullehre und in der hochschuldidaktischen Weiterbildung läuft das Formulieren von Lernzielen bisweilen Gefahr, zur schematisierten Pflichtübung zu geraten, die einer sinnvollen didaktischen Modellierung und Überprüfung des Kompetenzerwerbs eben gerade nicht gerecht wird. „A one-size-fits-all statement of intended learning outcomes will simply not work“ (Land/Meyer, 2009, S. 66).

Andererseits geht auch die Idee der Threshold Concepts auf einen Versuch zurück, angemessene Lernziele zu bestimmen. Die Idee wurde eingeführt, um zwischen Lernzielen zu unterscheiden, die Perspektiverweiterungen beinhalten („seeing things in a new way“) und solchen, die das nicht tun (Meyer/Land, 2003, S. 1). Wenn Threshold Concepts helfen, in diesem Sinne Lerngegenstände auszumachen, denen als „jewels of the curriculum“ (Land et al., 2005, S. 57) besondere Aufmerksamkeit zukommen sollte, dienen sie als Instrument der didaktischen Reduktion. Das Framework wurde beispielsweise in einem neuseeländischen Projekt der Lehr- und Prüfungsentwicklung in den Wirtschaftswissenschaften genutzt, um sicherzustellen, dass in grundlegenden Lehrveranstaltungen mit großen Studierendengruppen ökonomische Konzepte und Theorien fokussiert werden, deren Verständnis für die Entwicklung wirtschaftswissenschaftlicher Denkweisen aus Sicht der Beteiligten besonders wichtig ist (Hedges/Pacheco, 2015).

So verstanden lässt sich mit Threshold Concepts ein roter Faden im kompetenzorientierten Curriculum (Reis/Ruschin, 2007) herstellen und das „komplexe, realistische, an vielen Stellen mehrdimensionale Expertenwissen“ (Wunderlich/Szczyrba, 2018, S. 1) heranführen. Eine Betrachtung von Transformation und Variabilität, wie sie im TCF konzeptionalisiert werden, schärft dabei das Bewusstsein dafür, dass die Lernwege von Studierenden weniger geradlinig sind, als in Lehrplanung und Curriculumentwicklung häufig angenommen wird (Land/Meyer, 2009). Etliche Beiträge der PITCH-Lehrendenteams machen deutlich, dass es nicht einfach ist, aus diesen Erkenntnissen im Lehralltag, gerade für große und heterogene Lehrendengruppen, gute Lehr- und Prüfungsdesigns zu entwickeln, und reflektieren auch die dafür notwendigen personellen und zeitlichen Ressourcen (z. B. Beitrag von Milne/Bös/Eismann in diesem Band).

Mit der Ausrichtung von Lehrgestaltung auf kompetenzorientierte Lernziele beziehungsweise intendierte Lernergebnisse hat das Constructive Alignment (Biggs/Tang, 2011) als hochschuldidaktisches Prinzip große Verbreitung erfahren und wird in der Weiterbildung Lehrenden vielfach als grundlegendes Orientierungsmodell angeboten. Die Annahme, dass sich die Qualität von Lehre verbessern lässt, wenn Ziele, Lernaktivitäten und Prüfungsformen sinnvoll aufeinander abgestimmt werden, ist einfach zu vermitteln und erscheint plausibel,

da viele Lehrende wahrnehmen, dass Studierende ihr Lernverhalten auf Prüfungen ausrichten. Obwohl sich das Prinzip in der Praxis vielfach als nützliche Heuristik erweist, gibt es auch deutliche Kritik daran. So weist etwa Eugster (2012) auf die Ereignishaftigkeit von Bildung und dabei hilfreiche Brüche und Spannungen in Curricula hin, die durch Constructive Alignment eher verdeckt als sichtbar gemacht werden. Reinmann bescheinigt dem Constructive Alignment einen „technokratischen Zug“ (Reinmann, 2018, S. 5) und äußert die Sorge, dass Lehrende bei einem Übermaß von institutioneller Ergebnisorientierung aus ihrer didaktischen Verantwortung entlassen würden (ebd., S. 6). Sie fordert eine Rückbesinnung auf didaktische Urteilskraft und Wissenschaftsdidaktik als Grundlagen für die Gestaltung von Hochschullehre.

Mit einer ähnlichen Stoßrichtung wird im TCF der praktische Nutzen des Constructive Alignments zwar anerkannt, aber auch zu bedenken gegeben, dass es die Transformationsprozesse der Lernenden nicht angemessen erfasse (Land/Meyer, 2010, S. 73). Der Blick auf transformatives Lernen und präliminale, liminale und postliminale Variabilität (Calduch/Ratray, 2022; Meyer/Land, 2010) beim Übertritt über Schwellen sensibilisiert für den „unverfügbaren Rest“ (Wiemer/Kenneweg, 2021, S. 19) der Gestaltung von Lehr-Lern-Prozessen. Gemeint sind damit erwünschte Prozesse, die sich aber einer einfachen didaktischen Planung entziehen und verlangen, dem Eigensinn von Lehrenden und Lernenden Raum zu geben.

5 Fazit: Didaktische Lehrreflexion als Potenzial

Das didaktische Potenzial von Threshold Concepts liegt darin, Reflexion über Lehre zu fördern, und zwar sowohl mit Blick auf spezifische Inhalte sowie Denk- und Arbeitsweisen des jeweiligen Faches als auch mit Blick auf Lern- und Entwicklungsprozesse von Studierenden gerade in ihrer Variabilität. Bildungsziele, Lerngegenstände und Beobachtungen zu tatsächlichen Lernprozessen von Studierenden werden in einen Zusammenhang gebracht. Um dieses Potenzial zu nutzen, bedarf es einer systematischen und wissenschaftsbasierten Auseinandersetzung mit Lehre, die durch Hochschuldidaktiker*innen begleitet werden kann, aber nicht muss. Dabei geht es nicht nur darum, didaktische Gestaltungsideen zu entwickeln und umzusetzen, sondern auch und gerade darum, zu verstehen, was und wie Studierende lernen, welche Schwierigkeiten sie haben, welche Durchbrüche sie erleben und wie sich relevante Aspekte des Lernens sichtbar machen lassen.

Ergebnisse einer solchen Reflexion können sein:

- Zielsetzungen, curriculare Schwerpunkte und Bildungsvorstellungen werden konkretisiert, indem man danach fragt, welche Schwellen überquert werden sollten und mithilfe welcher Lernangebote sich dieses Überschreiten am ehesten sicherstellen lässt.

- Lernprozesse von Studierenden werden besser verstanden, weil sowohl Hürden und Hindernisse als auch Aha-Momente und Durchbrüche in den Blick rücken.
- Unterschiede und Ungleichzeitigkeiten in Lern- und Bildungsprozessen können über Beobachtungen zu prä- und postliminaler Variabilität erfasst und unterschiedliche Voraussetzungen, Lernergebnisse und Anwendungskontexte besser berücksichtigt werden.

Für diese didaktische Lehrreflexion sind Anlässe zu schaffen oder sich bietende Anlässe zu nutzen. Es sind unterschiedliche Formate denkbar:

Im Rahmen von Workshops zum Threshold Concept Framework in hochschuldidaktischen Weiterbildungsprogrammen können Lehrende den Ansatz kennenlernen und prüfen, inwieweit er für sie einen Zugang zur Weiterentwicklung ihrer Lehre bietet, gegebenenfalls sogar die Beforschung der eigenen Lehre in einem Vorhaben im Sinne von Scholarship of Teaching and Learning. Der Vorteil solcher Workshops ist, dass dort oft Lehrende unterschiedlicher Fachrichtungen aufeinandertreffen und einander helfen können, die Selbstverständlichkeiten ihres Faches zu hinterfragen und Lerngegenstände neu zu betrachten. Vielfach erprobt ist es zudem, die Identifikation und Diskussion von Threshold Concepts in Prozesse der Curriculum- und Studiengangsentwicklung zu integrieren (Cousin, 2008) und so Strukturierung und Ausgestaltung von Curricula auf für das Verständnis des jeweiligen Faches und die Kompetenzentwicklung bedeutsame Lerngegenstände zu fokussieren.

Auch das im PITCH-Projekt eingesetzte Gesprächsformat ist eine Möglichkeit, Threshold Concepts für die didaktische Lehrreflexion zu nutzen. Die Erfahrungen und vereinzelte Rückmeldungen aus dem Projekt zeigten, dass diese Gespräche einerseits interessante Impulse darstellten, andererseits aber nicht immer gut in den Projektablauf zu integrieren waren, weil beispielsweise viele Kapazitäten durch technische Fragen bei der Digitalisierung von Prüfungen gebunden waren. Am Beispiel der Threshold Concepts lässt sich also auch (selbst-)kritisch reflektieren, unter welchen Bedingungen es sinnvoll ist, (neue) didaktische Konzepte in Lehrentwicklungsprojekte einzubringen. Denn die Beobachtungen aus dem PITCH-Projekt führen zu der Annahme, dass gründlich zu prüfen ist, inwiefern ein didaktischer Ansatz an die konkreten Ziele und Probleme anschlussfähig ist, die im Rahmen der Lehrentwicklung bearbeitet werden, und wann der richtige Zeitpunkt dafür ist, neue didaktische Ideen einzuführen.

In der Literatur wird bisweilen darauf verwiesen, dass Threshold Concepts als Perspektive auf Lehren und Lernen selbst ein Threshold Concept sind (Entwistle, 2008; O'Brian, 2008). Nach dieser Auffassung verändert das Arbeiten an und mit Threshold Concepts die Sicht von Lehrenden auf studentisches Lernen und wirkt so transformativ für die Lehrpraxis. Integrierend sind Schwellenkonzepte, wenn sie Fachwissen, disziplinäre Denk- und Arbeitsweisen sowie die Lernprozesse und Herausforderungen von Studierenden zusammenbringen. Dies

ist allerdings oft mühevoll und beschwerlich, weil es unter Umständen gewohnte Vorgehensweisen in der Lehre in Frage stellt und ein Loslassen bisheriger Vorstellungen vom ‚richtigen‘ Lehren und Lernen erfordert. Eine Mühe, die sich durchaus lohnt.

Literatur

- Åkerlind, G./McKenzie, J./Lupton, M. (2011). A threshold concepts focus to curriculum design: Supporting student learning through application of variation theory. Sydney, Australia: Office for Learning and Teaching.
- Batzli, J. M./Knight, J.K./Hartley, L.M./Maskiewicz, A. C./Desy, E. A. (2016). Crossing the Threshold: Bringing Biological Variation to the Foreground. In: CBE-Life Science Education 15(4). <https://doi.org/10.1187/cbe.15-10-0221>
- Biggs, J. B./Tang, C. (2011). Teaching for Quality Learning at University (4. Aufl.). Buckingham, England: Open University Press.
- Calduch, I./Ratray, J. (2022). Revisiting postliminal variation in threshold concepts: issues of unexpected transformation and legitimization. In: Studies in Higher Education, 47(7), 1453–1463, <https://doi.org/10.1080/03075079.2021.1910651>
- Carmichael, P. (2010). Threshold concepts, disciplinary differences and cross-disciplinary discourse. In: Learning and Teaching in Higher Education: Gulf Perspectives, 7(2), 53–72.
- Cousin, G. (2008). Threshold Concepts: Old Wine in New Bottles or New Forms of Transactional Inquiry. In R. Land/J. Meyer/J. Smith (Hrsg.): Threshold concepts within the disciplines (S. 261–272), Rotterdam: Sense Publishers.
- Cousin, G. (2010). Neither teacher-centred nor student-centred: threshold concepts and research partnerships. In: Journal of Learning Development in Higher Education, 2, <https://doi.org/10.47408/jldhe.v0i2.64>
- Entwistle, N. (2008). Threshold Concepts and Transformative Ways of Thinking within Research into Higher Education. In R. Land/J. Meyer/J. Smith (Hrsg.): Threshold concepts within the disciplines (S. 21–35). Rotterdam: Sense Publishers.
- Eugster, B. (2012). Leistungsnachweise und ihr Ort in der Studiengangentwicklung. Überlegungen zu einer Kritik des curricularen Alignments. In T. Brinker/P. Tremp (Hrsg.): Einführung in die Studiengangentwicklung (S. 45–62). Bielefeld: Bertelsmann.
- Hedges, M. A./Pacheco, G. A. (2015). Student engagement and exam performance: It's (still) ability that matters most. In: Australasian Journal of Economics Education, 12(2), 50–74.
- Hounsell, D./Anderson, C. (2009). Ways of Thinking and Practicing in Biology and History: Disciplinary aspects of Teaching and Learning Environments. In C. Kreber (Hrsg.): The University and its Disciplines: Teaching and Learning within and beyond disciplinary boundaries (S. 71–83). New York / London: Routledge.
- Hounsell, D./Hounsell, J. (2007). Teaching-learning environments in contemporary mass higher education. In: Entwistle, N. J./Tomlinson, P. (Hrsg.): Student Learning and University Teaching (S. 91–111). Leicester: British Psychological Society.
- HRK/KMK (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

- Kenneweg, A. C./Wiemer, M. (2022). Threshold Concepts. Zugang zur Fachwissenschaft und Ansatzpunkt für die Wissenschaftsdidaktik. In: G. Reinmann/R. Rhein (Hrsg.), *Wissenschaftsdidaktik 1. Einführung* (S. 245–266). Bielefeld: transcript.
- Land, R./Cousin, G./Meyer, J. H./Davies, P. (2005). Threshold concepts and troublesome knowledge (3)*: implications for course design and evaluation. In C. Rust (Hrsg.), *Improving student learning: diversity and inclusivity* (S. 53–64). Oxford: OCSLD.
- Land, R./Meyer, J. H. F. (2010). Threshold Concepts and Troublesome Knowledge 5: Dynamics of assessment. In Meyer, J. H. F./Land, R./Baillie, C. (Hrsg.): *Threshold Concepts and Transformative Learning* (S. 61–79). Rotterdam: Sense Publishers.
- Lucas, U./Mladenovic, R. (2007). The potential of threshold concepts: An emerging framework for educational theory and practice. In: *London Review of Education*, 5(3), S. 237–248. <https://doi.org/10.1080/14748460701661294>
- Meyer, J. H. F./Land, R. (2003). Threshold Concepts and Troublesome Knowledge: Linkages to Ways of Thinking and Practising within the Disciplines. In: C. Rust (Hrsg.), *Improving Student Learning: Improving Student Learning Theory and Practice – Ten Years On* (S. 1–12). Oxford: Oxford Centre for Staff and Learning Development.
- Nicola-Richmond, K./Pépin, G./Larkin, H./Taylor, C (2018). Threshold concepts in higher education: a synthesis of the literature relating to measurement of threshold crossing. In: *Higher Education Research / Development*, 37(1), 101–114. <https://doi.org/10.1080/07294360.2017.1339181>
- O’Brian, M. (2008). Threshold Concepts for University Teaching and Learning. In: R. Land; J. Meyer/J. Smith (Hrsg.): *Threshold concepts within the disciplines* (S. 261–272). Rotterdam: Sense Publishers.
- Pace, D. (2017). *The Decoding the Disciplines Paradigm. Seven Steps to increased student learning*. Bloomington: Indiana University Press.
- Perkins, D. (1999). The many faces of constructivism. In: *Educational Leadership*. 57(3), 6–11.
- Perkins, D. (2006). Constructivism and troublesome knowledge. Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge. In J. Meyer/R. Land (Hrsg.), *Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge* (S. 33–47). New York / London: Routledge.
- Ratray, J. (2023). On the affective threshold of power and privilege. In: *Higher Education*, 87, 1829–1843. <https://doi.org/10.1007/s10734-023-01093-x>
- Reinmann, G. (2018). Shift from Teaching to Learning und Constructive Alignment: Zwei hochschuldidaktische Prinzipien auf dem Prüfstand. In: *Impact Free*, 14. Hamburg.
- Reis, O./Ruschin, S. (2007). Kompetenzorientiertes Prüfen als zentrales Element gelungener Modularisierung. In: *Journal Hochschuldidaktik*, 18(2), 6–9.
- Rowbottom, D. P. (2007). Demystifying Threshold Concepts. In: *Journal of Philosophy of Education*, 41(2), 263–270. <https://doi.org/10.1111/j.1467-9752.2007.00554.x>
- Salwén, H. (2021). Threshold concepts, obstacles or scientific dead ends? In: *Teaching in Higher Education*, 26(1), 36–49. <https://doi.org/10.1080/13562517.2019.1632828>
- Shopkow, L. (2010). What ‘Decoding the Disciplines’ has to offer ‘Threshold Concepts’. In J.H.F. Meyer/R. Land/ C. Baillie (Hrsg.), *Threshold Concepts and Transformative Learning* (S. 317–32). Rotterdam: Sense Publishers.
- Walck-Shannon, E. M./Batzli, J./Pultorak, J./Boehmer, H. (2019). Biological Variation as a Threshold Concept: Can We Measure Threshold Crossing? In: *CBE Life Sciences Education*, 18(3). <https://doi.org/10.1187/cbe.18-12-0241>

- Wiemer, M./Kenneweg, A. C. (2021). Threshold Concepts: Übergänge zu disziplinären Denkweisen und transformative Lernprozesse in der Fachlehre verstehen und begleiten. In B. Berendt/A. Fleischmann/ G. Salmhofer/N. Schaper/B. Szczyrba/M. Wiemer/J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre*, 100 (S. 43–66). Berlin: DUZ- Verlags- und Medienhaus. <https://doi.org/10.36197/DUZOPEN.029>
- Wunderlich, A./Szczyrba, B. (2018). Kompetenzorientiertes Prüfen – transparent, komplex und fair. In B. Berendt/A. Fleischmann/N. Schaper/B. Szczyrba/M. Wiemer/J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre* (S. 77–100). Berlin: DUZ Verlags- und Medienhaus

Integriertes Lehr-, Lern- und Prüfkonzept für eine politikwissenschaftliche Einführungsveranstaltung

Ulrike Berendt, Martin Krybus und Till Behmenburg¹

1 Prüfungsdigitalisierung als didaktische Herausforderung

Eine breite Zielgruppe und ein weiter Themenhorizont sind kennzeichnend für universitäre Einführungsveranstaltungen, so auch für die Vorlesung „Grundlagen der Politikwissenschaft“ an der Universität Duisburg-Essen. Über 300 Studierende aus den Bachelorstudiengängen Politikwissenschaft, Soziologie, Globale und Transnationale Soziologie sowie Moderne Ostasienstudien zählen zur Zielgruppe, die sich aus einer – etwa in puncto Zuwanderungsgeschichte oder Bildungshintergrund – zunehmend diverseren Studierendenschaft rekrutiert. Als Einführungsveranstaltung vermittelt die Vorlesung Studierenden im ersten Fachsemester grundlegende Begriffe und Konzepte der gesamten Politikwissenschaft. Der thematische Parforceritt gewährt Einblicke in politikwissenschaftliche Teilbereiche wie Politische Theorie, Analyse politischer Systeme oder Internationale Beziehungen und verbindet diese zu einem stimmigen Gesamtüberblick. Die in jedem Wintersemester abgehaltene Vorlesung gliedert sich in vierzehn Inhaltseinheiten, deren Kenntnis in einer neunzigminütigen Klausurprüfung ermittelt wird. Um Studierende bei Wiederholung und Anwendung des frontal vermittelten Lehrstoffs zu unterstützen, wird die Vorlesung von studentischen Tutorien flankiert, deren Besuch freiwillig ist.

Wie sich das bestehende Lehr- und Lern-Angebot mittels Digitalisierung der Klausurprüfung optimieren lässt, wird im Folgenden dargelegt. Ein gedanklicher Ausgangspunkt ist die Diversität der Studierendenschaft, welche Lehrende vor die besondere Herausforderung stellt, den Studienerfolg und die Chancengerechtigkeit für alle Studierenden zu gewährleisten. *Threshold Concepts* und *Decoding the Disciplines* (Middendorf/Pace, 2004) sowie das Prinzip des *Constructive Alignment* (Biggs/Tang, 2011, S. 52–55) bieten vielversprechende Ansätze, um diesen Herausforderungen zu begegnen. Der vorliegende Artikel führt diese Konzepte zusammen und erläutert anhand der Vorlesung „Grundlagen der Politikwissenschaft“, wie sich die drei Ansätze zur Prüfungsvorbereitung einsetzen lassen, um Lernerfolg und Chancengerechtigkeit zu steigern. Ebenfalls

¹ Unter Mitarbeit von Jessica Kuhlmann und Jann Tholen.

wird gezeigt, welche Synergien sich dabei mit der Digitalisierung bestehender Übungs- und Prüfungsformate erzielen lassen.

Im ersten Schritt sind die bestehenden Herausforderungen der Lehrveranstaltung zu identifizieren und zu systematisieren, um eine Problemdefinition vorzunehmen. Dabei stiften *Threshold Concepts* und der Ansatz des *Decoding the Disciplines* den theoretischen Rahmen. Mit ihrer Hilfe können einerseits klassische *Thresholds*² im Sinne inhaltlicher Hürden bestimmt werden, deren Überwindung sich für Studierende potenziell schwierig gestalten kann, jedoch für einen Fortschritt im politikwissenschaftlichen Verständnis und ebenso für das Bestehen der Klausur notwendig ist. Andererseits können *strukturelle* sowie *technische Hürden*³ identifiziert werden, die zumindest bei einem Teil der Studierenden zusätzlich den Prüfungserfolg erschweren.

In einem zweiten Schritt werden Ziele definiert, die sich einerseits aus den ermittelten inhaltlichen, strukturellen und technischen Hürden und andererseits aus den fünf Entwicklungsfeldern des PITCH-Projekts – Stärkung der Kompetenzorientierung, Implementation neuer Aufgabentypen und Prüfungsformate, Unterstützung der Klausurvorbereitung mit darauf abgestimmten Übungsaufgaben, Erhöhung der Flexibilität und Chancengerechtigkeit sowie Förderung einer systematischen Qualitätsentwicklung – ableiten lassen.

Auf Basis der definierten Ziele werden im dritten Schritt adäquate *Instrumente* zur didaktischen Weiterentwicklung der Veranstaltung zu einem integrierten Vorlesungs-, Tutorien- und Prüfungskonzepts entwickelt. Handlungsleitend ist neben den PITCH-Entwicklungsfeldern auch die Berücksichtigung eines didaktischen Handlungsfelds (Schulz, 2021, S. 21–22), da sich auf dieser Grundlage Maßnahmen für digitale Übungs- und Prüfungsformate weiter konkretisieren lassen.⁴

Die inhaltliche Konzeption der Instrumente wird nach dem Prinzip des *Constructive Alignment* vorgenommen, welches besagt, dass die Kompetenzziele die Lehrmethoden bestimmen sollen (Lambach, 2016, S. 13). Dementsprechend müssen sowohl Inhalt als auch Format von prüfungsvorbereitenden Übungen mit der Prüfung verbunden sein. Auch die Prüfung selbst muss auf das ausgerichtet sein, was von Studierenden in ihrem weiteren Studium, aber auch in

2 Vgl. hierzu Meyer/Land/Baillie (2010) und den Beitrag von Cornelia Kenneweg in diesem Band.

3 Strukturelle und technische Hürden sind im Ansatz des *Decoding the Disciplines* zwar ursprünglich nicht inkludiert, da der Ansatz explizit als „offen“ charakterisiert wurde (Middendorf/Pace, 2004, S. 4) und die identifizierten strukturellen und technischen Hürden ebenfalls konkrete Handlungsbedarfe hervorrufen, werden diese gleichrangig in die Weiterentwicklung der Veranstaltung einbezogen.

4 Schulz (2021, S. 18) unterscheidet bei der Umsetzung digitaler Prüfung die vier Handlungsfelder Recht, Technik, Didaktik und Organisation, welche bei jeder digitalen Prüfung zu beachten sind. In diesem Beitrag erfolgt jedoch eine Fokussierung auf das didaktische Handlungsfeld.

der Berufspraxis erwartet wird (Biggs/Tang, 2011, S. 52). Dies soll ein tieferes und nachhaltigeres Lernen anregen, welches dabei helfen kann, auch kognitiv anspruchsvolle Schwellenkonzepte einer Lehrveranstaltung zu durchdringen (Lambach, 2016, S. 10). Sowohl von den Entwicklungsfeldern des PITCH-Projekts als auch von den verwendeten didaktischen Ansätzen wird die Bedeutung einer Erfolgsmessung hinsichtlich angestrebter Ergebnisse (Hattie, 2009, zitiert nach Lambach, 2016, S. 13) betont, weshalb abschließend Instrumente zur Qualitätssicherung skizziert werden.

2 Identifikation von Hürden und Ableitung von Zielen

Dem Ansatz des *Decoding the Disciplines* folgend, wonach die Lernleistung von Studierenden durch gezielte Beseitigung bestehender *Hürden* gesteigert werden kann (Middendorf/Pace, 2004), werden im Folgenden die identifizierten Hürden und darauf abgestimmten Handlungsbedarfe und Instrumente aufgezeigt. Zur Identifikation dieser Hürden wurden verschiedene Pfade beschritten: Eine Klausurnachbefragung – bestehend aus einem standardisierten Fragebogen und einem von der PITCH-Projektkoordination durchgeführten Kleingruppengespräch – gab Aufschluss über bestehende Prüfungsprobleme der Studierenden. Parallel hierzu erfolgte eine statistische Analyse von Klausurergebnissen vergangener Jahre, um Prüfungsinhalte und Aufgabenstellungen mit unterdurchschnittlichen Resultaten zu identifizieren. Während diese Erhebungsinstrumente der Prüfung nachgelagert waren, fand semesterbegleitend ein kontinuierliches Monitoring des Lernfortschritts der Studierenden durch Auswertung der Übungsaufgaben statt. Ein Decoding-Interview trug ebenso zur Sensibilisierung für bestehende Hürden bei wie ein regelmäßiger Austausch im Kontext des PITCH-Projekts, etwa im Rahmen von Etappengesprächen mit der Projektkoordination oder im bilateralen Austausch mit anderen Projektteams. Insgesamt wurden verschiedene Hürden identifiziert, die sich in drei Kategorien zusammenfassen lassen: Erstens *fachspezifisch-inhaltliche Hürden*, zweitens eher *strukturell-didaktische Hürden* des Studierens, die zum Studienbeginn besonders ins Gewicht fallen, sowie drittens *technische Hürden*, die freilich nicht trennscharf voneinander separierbar sondern zum Teil eng miteinander verwoben sind.

Fachspezifisch-inhaltliche Hürden betreffen den inhaltlichen Zugang der Studierenden zum Studienfach Politikwissenschaft im Allgemeinen und zur Vorlesung „Grundlagen der Politikwissenschaft“ im Besonderen. Von grundlegender Bedeutung ist der zum Studienbeginn notwendige Spagat, einerseits an vorhandenes Politikwissen anzuknüpfen, aber andererseits das voruniversitäre politische Alltagsverständnis zu überwinden, um eine akademische Perspektive auf Politik einnehmen zu können (Mols, 2020, S. 30–31). Den Studierenden ist darum zuvorderst der Nutzen der grundlegenden politikwissenschaftlichen Definitionen und Modelle zu verdeutlichen, deren Verständnis sich zunächst schwierig gestal-

ten kann, jedoch für einen Lernfortschritt sowie das Bestehen der Klausur unverzichtbar ist. Eine solche inhaltliche Hürde stellt in der Vorlesung „Grundlagen der Politikwissenschaft“ etwa das Modell des politischen Systems dar. Die daran anknüpfenden Vorlesungsinhalte – das intermediäre System mit Parteien, Verbänden und Medien als prägenden Akteur:innen – bauen auf dem vorangehenden Verständnis des politischen Systems unmittelbar auf, sind ohne dieses also ihrerseits nicht zu erfassen.

Strukturell-didaktische Hürden betreffen jenseits des konkreten Studienfachs das Studieren allgemein. Aufgrund der Verortung einer Einführungsvorlesung am Beginn des Studiums betreffen sie jedoch insbesondere den Studieneinstieg und korrespondieren dabei mit heterogenen Ausgangsbedingungen der Studierenden, welche den Studienerfolg beeinflussen können. Über die Diversität der Studierenden geben Daten eines universitätsinternen Diversity Monitoring⁵ hinreichend Aufschluss: Für die Gesellschaftswissenschaften⁶ fällt auf, dass sich der Anteil der Studienanfänger:innen mit Zuwanderungsgeschichte in den vergangenen zehn Jahren mehr als verdoppelt hat. Bedingt durch unterschiedliche Entwicklungen ist die absolute Mehrheit der Studienanfänger:innen mittlerweile unter zwanzig Jahre. Auf ein knappes Drittel angestiegen ist der Anteil derer mit niedrigem Bildungshintergrund, während weniger als die Hälfte einen mindestens gehobenen Bildungshintergrund aufweisen.⁷ Die Note der Hochschulzugangsberechtigung hat sich im selben Zeitraum tendenziell verbessert; dennoch kann knapp die Hälfte hier nur eine befriedigende Note vorweisen. Diese Entwicklungstrends erlauben freilich keine konkreten Rückschlüsse über Fertigkeiten oder Vorkenntnisse der Studierenden, skizzieren aber eine junge sowie in puncto Zuwanderung, Bildungshintergrund und Abiturnote ausgeprägt heterogene Zielgruppe der Vorlesung „Grundlagen der Politikwissenschaft“. Angesichts der inhaltlichen und zeitlichen Verdichtung von Bachelor-Studiengängen kommt dem Gelingen des Studieneinstiegs eine erhebliche Bedeutung zu, da sich schulisches Lernen nicht bloß didaktisch fundamental von universitärem Studieren unterscheidet (Koch, 2015, S. 97–107). Die Motivation zum regelmäßigen Besuch einer Vorlesung sowie eines begleitenden Tutoriums ist im Rahmen eines

5 Das Diversity Monitoring der Universität erhebt seit dem Wintersemester 2012/13 jährlich und systemisch Strukturdaten von Studienanfänger:innen und Absolvent:innen aller Fakultäten, die unter Diversitätsaspekten relevant sind. Dazu zählen u. a. Aspekte von Diversität und Hochschulzugang (Universität Duisburg-Essen, 2023).

6 Die Studierendenschaft der Gesellschaftswissenschaften im ersten Semester ist im Wesentlichen identisch mit der Zielgruppe der Vorlesung „Grundlagen der Politikwissenschaft“, weshalb die Daten, die für die Fakultät für Gesellschaftswissenschaften erhoben wurden, ebenso für die Vorlesung gültig sind.

7 Mindestens ein Elternteil ohne beruflichen Abschluss ist dabei als niedriger Bildungshintergrund definiert, während ein gehobener Bildungsabschluss mindestens einen Elternteil mit Hochschulabschluss erfordert (Universität Duisburg-Essen, 2023).

freiwillig aufgenommenen Studiums ebenso essenziell wie die Etablierung einer funktionierenden Selbstorganisation der Studierenden (Rost, 2018, S. 109–143). Prüfungen verpflichtender Einführungsveranstaltungen sind mangels Klausurerfahrungen zudem in besonderer Weise mit Unsicherheit bezüglich Prüfungsinhalten, -struktur und -ablauf betroffen.

Technische Hürden betreffen die Umsetzung digitaler Prüfungen an Hochschulen und aufgrund verschiedener Vorkenntnisse sowohl bei Studierenden als auch Lehrenden – wenn auch in unterschiedlicher Ausprägung: Sind auf Studierendenseite die anzunehmende Heterogenität bei der Verteilung digitaler Kompetenzen sowie der Verfügbarkeit leistungsfähiger digitaler Endgeräte wesentliche Hürden, so sind die begrenzten Erfahrungen mit digitalen Lernplattformen und Prüfungen an der Fakultät ein mögliches Hemmnis bei der Realisierung digitaler Prüfungen.

Die vorangehend identifizierten Hürden korrespondieren auffallend mit den bereits benannten Entwicklungsfeldern des PITCH-Projekts, sodass sich insgesamt sieben konkrete Ziele ableiten lassen, die im Folgenden kurz vorgestellt werden: *Aktivierung, Anwendungsbezug, Augenhöhe, Angleichung der Lern- und Prüfungssituation, Arbeitsvereinfachung, Auswertung der Ergebnisse und Austausch*. Der handlungsleitenden Hauptzielsetzung verpflichtet, den Lern- und Prüfungserfolg der Studierenden nachhaltig zu verbessern, ist zur Überwindung fachspezifischer sowie allgemein didaktischer Hürden zuvorderst eine *Aktivierung der Studierenden* im Sinne eigenständigen und motivierten Lernens anzustreben. Diese Aktivierung von Lernpotenzialen soll mithilfe eines vielfältigen Übungsangebots erreicht werden. Das Ziel eines hohen *Anwendungsbezugs* leitet sich einerseits aus der identifizierten Hürde der Entwicklung einer wissenschaftlichen Perspektive auf Politik ab, andererseits aus dem ersten Entwicklungsfeld des PITCH-Projekts, welches der Steigerung der Kompetenzorientierung gilt, also der Anwendung erlernter Begriffe und Konzepte. Um den Studierenden gerade am Anfang des Studiums einen offenen Raum für Rückfragen zu eröffnen, sind Begegnungsmöglichkeiten auf *Augenhöhe* unerlässlich, wobei Peer-Mentoring eine herausgehobene Rolle spielt. Im Sinne der Chancengerechtigkeit und des *Constructive Alignment* ist die *Angleichung der Lern- und Prüfungssituation* ein weiteres Kernanliegen. Die Digitalisierung der Klausurprüfung erhöht zwar den Aufwand der Vorbereitung, sorgt aber bei der Durchführung der Veranstaltung und der Prüfung für erkennbare *Arbeitsvereinfachung*.

Ebenfalls ist eine umfangreiche *Auswertung der Ergebnisse* anzustreben – sowohl der erzielten Lernfortschritte als auch der zur Anwendung kommenden Instrumente. Einerseits wird diese Ergebnisauswertung vom fünften PITCH-Entwicklungsfeld, welches eine systematische Qualitätsentwicklung von Prüfungen anregt, dem didaktischen Handlungsfeld sowie dem *Constructive Alignment* gefordert, andererseits lässt sich der Erfolg von Maßnahmen zur Überwindung identifizierter Hürden nur durch konsequente Evaluation beurteilen. Ein *Austausch* mit Studierenden aber auch mit Lehrenden sorgt dafür, dass einerseits Stu-

dierende die Möglichkeit erhalten, ihren Lernfortschritt kontinuierlich beurteilen zu können und andererseits Lehrende bezüglich der Konzipierung und Digitalisierung ihrer Prüfungsformen neue Impulse erhalten.

3 Instrumentarium des Lehr-, Lern- und Prüfkonzepts

Nachdem die Ausgangslage beschrieben, der theoretische Rahmen geschaffen, die identifizierten Hürden vorgestellt und die daraus resultierenden Ziele formuliert wurden, gilt es nun, die Maßnahmen zur Optimierung der Veranstaltung und ihrer Abschlussklausur zu erläutern. Da die Kompetenzziele die Lehrmethoden bestimmen sollten (Lambach, 2016, S. 13), erscheint es sinnvoll, das Augenmerk zunächst auf die Klausur zu richten. Für die Weiterentwicklung der Klausur stiftet das erste PITCH-Entwicklungsfeld, die Stärkung der Kompetenzorientierung in digitalen Prüfungen, die notwendige inhaltliche Orientierung. Aus dem Postulat, dass Gegenstand der Prüfung das sein sollte, was von den Studierenden in ihrem weiteren Studium, aber auch in der Berufspraxis erwartet wird (Biggs/Tang, 2011, S. 52), lässt sich die Notwendigkeit eines hohen Anwendungsbezugs der Prüfung formulieren, welche im Sinne des zweiten PITCH-Entwicklungsfelds durch Implementation neuer Aufgabentypen und Prüfungen realisiert werden soll. Zu diesem Zweck wurde die in der Vergangenheit bereits digital durchgeführte Prüfung vollständig auf Moodle umgestellt. Dies ermöglicht die Verwendung einer breiteren Palette an komplexeren Fragetypen, die sich auf verschiedene Ebenen des Verständnisses der Studierenden konzentrieren können. Auch kann in Moodle die vom *Constructive Alignment* geforderte Angleichung der Lern- und Prüfungssituation erreicht werden, was eine Erhöhung der Chancengerechtigkeit bedeutet, da alle Studierenden der Veranstaltung mit Moodle als digitaler Lehr- und Lernumgebung bereits vertraut sind. Um auch während der Prüfung für Chancengerechtigkeit zu sorgen, wird die Moodle-Klausur in Präsenz unter kontrollierten Bedingungen in der für digitale Prüfungen ausgelegten PC-Hall der Universität geschrieben.

Aus der Perspektive der Prüfungsorganisation ergeben sich nach der erfolgreichen Implementierung ebenfalls einige klare Vorteile. So konnte der Aufwand für die Vorbereitung und Korrektur der Klausur erheblich reduziert werden. Zum einen ermöglicht Moodle das unkomplizierte Zusammenstellen von Klausuren. Gleichzeitig verringert sich der Korrekturaufwand dank der in Moodle integrierten Auswertungs- und Bewertungsmöglichkeiten.

Um den Prüfungsablauf im Vorfeld transparent darzulegen, wird den Studierenden eine Beispielprüfung zur Verfügung gestellt. Bereits in der ersten Vorlesungssitzung erfolgt die Bekanntgabe von Ort und Termin sowie von Format und Ablauf der Klausur. Am Ende der Vorlesungszeit erhalten alle Studierenden zudem die Möglichkeit, sich durch Teilnahme an einer Probeklausur ein genaueres Bild über den Aufbau der Klausur zu machen.

Das Gelingen einer digitalen Prüfung erfordert unter didaktischen Gesichtspunkten eine Verzahnung von Lehr- und Prüfungsinhalten (Schulz, 2021, S. 21–22), was gleichfalls für die angestrebte Stärkung des Anwendungsbezugs gegenüber einer reinen Wissensabfrage gilt. Orientiert an den PITCH-Entwicklungsfeldern besteht die Notwendigkeit, die Studierenden mit darauf abgestimmten Übungsaufgaben zu unterstützen, die ihrerseits einen starken Anwendungsbezug aufweisen, und das Vorlesungs-, Tutorien- und Prüfungskonzept um zusätzliche Elemente zu erweitern. Dies macht es wiederum erforderlich, die identifizierten fachspezifisch-inhaltlichen sowie strukturell-didaktischen Hürden bei der Konzeption der Übungsformate zu berücksichtigen.

Zur Überwindung dieser Hürden ist zuvorderst eine *Aktivierung der Studierenden* im Sinne eigenständigen und motivierten Lernens anzustreben. Diese Aktivierung von Lernpotenzialen soll mithilfe eines vielfältigen Lern- und Übungsangebots erreicht werden. Teil dieses Angebots sind freiwillige semesterbegleitende Lernzielkontrollen, die im Rahmen der Vorlesung durchgeführt werden. Diese Lernzielkontrollen – zum Teil konzipiert im Quizformat – werden als kontinuierliche Tests über das Semester verteilt durchgeführt. Diese sind in Moodle eingebunden, über mobile Endgeräte zu bearbeiten und fragen das Verständnis aktueller Vorlesungsinhalte sowie die Anwendung vergangener Sitzungsinhalte ab. Dem Ziel der Angleichung von Lern- und Prüfungssituation verpflichtet, werden die Lernzielkontrollen ausschließlich in Aufgabenformaten gestaltet, die später auch bei der Klausur zum Einsatz kommen. Dadurch erhalten die Studierenden die Möglichkeit, sich frühzeitig und kontinuierlich mit der Prüfungs Umgebung vertraut zu machen.

In didaktischer Hinsicht erfüllen die Lernzielkontrollen neben der Aktivierung noch zwei weitere Funktionen: Zum einen wird durch derartige Übungen die *Long Term Retention* der Vorlesungsinhalte verbessert, zum anderen erhalten sowohl Studierende als auch Lehrende Feedback zum aktuellen Lernstand (Schulz, 2021, S. 21–22). Die Studierenden können dadurch beurteilen, welche Inhalte sie bereits hinreichend erschlossen haben und an welchen Stellen sie ihre Lernaktivitäten noch intensivieren müssen. Gleichzeitig erhalten die Lehrenden durch diese Feedback-Schleife Anhaltspunkte für die (Um-)Gestaltung von Lehr- und Prüfungsformaten.

Das Lern- und Übungsangebot der Veranstaltung komplettieren studentisch geführte Tutorien, welche die Vorlesung „Grundlagen der Politikwissenschaft“ inhaltlich begleiten. Die für jeweils 20 bis 30 Teilnehmende ausgelegten und wöchentlich stattfindenden Tutorien sind durch vorbereitende Workshops für die Tutor:innen inhaltlich weitgehend standardisiert, um möglichst vergleichbare Lernumstände für Studierende bei der Erschließung der Vorlesungsinhalte zu garantieren. Auch diese Tutorien haben eine aktivierende Funktion, zeichnen sich darüber hinaus aber vor allem durch *Niedrigschwelligkeit und eine intensive Betreuung der Studierenden* aus. Um den Studierenden gerade am Anfang des Studiums einen Raum für Rückfragen zu eröffnen, der im Frontalformat

der Vorlesung notwendigerweise nur eingeschränkt vorhanden ist, begegnen die Tutor:innen den Studierenden im Sinne des *Peer-Mentoring auf Augenhöhe*. Dies trägt dazu bei, die Chancengerechtigkeit zu erhöhen, da in den Tutorien noch stärker Rücksicht auf individuelle Belange genommen werden kann, um die Studierenden bei der Überwindung inhaltlicher Hürden zu unterstützen und durch niedrigschwelliges Peer-Mentoring gegebenenfalls auch studienpraktische Fragen zu beantworten. Dies ist insbesondere vor dem Hintergrund der hohen Diversität der Studierendenschaft in puncto Alter, Zuwanderungsgeschichte, Bildungshintergrund und Abiturnote von Relevanz.

Für das Ziel, die Studierenden auf den hohen *Anwendungsbezug der Prüfung* vorzubereiten und die beschriebenen positiven Effekte der Tutorien noch zu steigern, hat sich eine Weiterentwicklung der Tutorien parallel zu Vorlesung und Klausur angeboten. Die Stärkung der Kompetenzorientierung gelingt einerseits durch Bearbeitung von Aufgaben und Materialien, welche – identisch zur Vorlesung – via Moodle bereitgestellt werden und während der Tutoriumssitzungen Anwendung finden, andererseits durch das Angebot von *freiwilligen Zusatzaufgaben*, die von Studierenden im Laufe des Semesters zusätzlich zum Besuch von Vorlesung und Tutorien bearbeitet werden. Der Anwendung von Gelerntem dienen *Fallstudien*, die als Gruppenarbeiten angeboten werden, wobei vier Themen zur Wahl stehen, für die konkrete Szenarien mit Aufgabenstellung und Arbeitsmaterial zu zentralen Inhalten der Vorlesung via Moodle bereitgestellt werden. Im Sinne des *Constructive Alignment* entsprechen die inhaltlichen Leistungsanforderungen der Fallstudien dem, was von den Studierenden in der Klausur erwartet wird, wie etwa die Anwendung des Politikzyklus auf eine vorgegebene Policy. Die Bearbeitung der Fallstudien erfolgt durch die Studierenden in Eigenverantwortung, was die Kompetenz zur Selbstorganisation stärkt. Die Einreichung der Fallstudien sowie die entsprechende Ergebnismeldung erfolgen ebenfalls über Moodle. Das Angebot der freiwilligen Zusatzaufgaben komplettiert schließlich die gleichfalls in die Tutorien integrierte Möglichkeit, als Einzelleistung einen kurzen Podcast aufzunehmen, der Vorlesungseinheiten anhand vorgegebener Leitfragen aufbereitet und exemplarisch auf das politische Tagesgeschehen eines jeweiligen Lehrthemas übertragen soll. Der dadurch trainierte politikwissenschaftliche Blick auf politische Prozesse hilft den Studierenden bei der Überwindung des bloßen Alltagsverständnisses von Politik. Ein ausführliches Handbuch bietet Orientierung für eine Podcast-Produktion, welche Kompetenzen in den Bereichen Planung, Strukturierung und Umsetzung eines eigenen Medienprojekts erfordert, wozu die Ausarbeitung eines Skripts, dessen Audioaufnahme und technische Bearbeitung zählen. Die Teilnahme an den verschiedenen Übungsaufgaben ist freiwillig und unabhängig voneinander möglich. In jedem Fall erfolgt eine zeitnahe individuelle Ergebnismeldung, die rechtzeitig vor der Klausur Aufschluss über den jeweiligen Lernfortschritt gibt.

4 Ergebnisauswertung und Erfahrungsaustausch

Wie bereits konstatiert, ist zudem eine systematische Qualitätsentwicklung von Prüfungen zu erreichen. Zwei konkrete Ziele lassen sich hiermit verbinden: Die kontinuierliche Auswertung durchgeführter Prüfungen sowie der Austausch mit anderen Lehrenden über erprobte Instrumente.

Der Einsatz von Moodle als Lern- und Prüfungsplattform ermöglicht eine kontinuierliche Auswertung der Ergebnisse von Übungsaufgaben und Prüfungen. Dies erlaubt vielfältige Ergebnisinterpretationen durch *Learning Analytics* im Sinne des Erhebens und Auswertens von Daten zum Lernfortschritt, die den Studierenden direkt und indirekt zugutekommen: Direkt durch fortlaufendes Feedback über ihren Lernfortschritt durch schnelle Ergebnisauswertung der freiwilligen Übungsaufgaben und indirekt durch eine regelmäßige Verbesserung des bestehenden Angebots aufgrund von Auswertungsergebnissen. Dadurch lassen sich Erkenntnisse über bestehende Schwierigkeiten im Umgang mit bestimmten Prüfungsthemen und Prüfungsaufgaben aufzeigen, um auf diese künftig verstärkt Augenmerk zu richten und das Lehr- und Lernangebot hinsichtlich auftretender Herausforderungen weiter zu optimieren. Auch hilft die Ergebnisauswertung von Übungen und Klausurprüfung dabei, angestrebte Positivfolgen von regelmäßigem Vorlesungsbesuch und aktiver Tutoriumsteilnahme zu messen. Zwar werden für beide Veranstaltungen keine Teilnehmendenzahlen erfasst, aber durch die Integration verschiedener Übungsaufgaben in Moodle kann zumindest ein mittelbarer Zusammenhang hergestellt werden. Eine regelmäßige und erfolgreiche Teilnahme an den Lernzielkontrollen korreliert dabei ebenso mit besseren Ergebnissen in der Klausurprüfung, wie die erfolgreiche Bearbeitung der Fallstudien und der Podcasts.

Regelmäßige Klausurnachbefragungen via Moodle und unmittelbar nach der Prüfung dienen zudem einer permanenten Qualitätskontrolle, indem etwaige technische oder inhaltliche Schwierigkeiten abgefragt werden, um eventuell gegensteuern zu können.

In kommunikativer Hinsicht ist zuletzt ein *Austausch* in mehrere Richtungen, der sowohl Studierenden bezüglich ihrer Lernfortschritte adressiert, als auch andere PITCH-Projektteams sowie Lehrende in den Gesellschaftswissenschaften, um Bewährtes zu bewerben und neue Impulse zu erhalten, ein erreichtes Ergebnis. Das Erstellen eines umfangreichen Handbuchs für Moodle-Prüfungen⁸

⁸ Das Handbuch wurde für Lehrende der Politikwissenschaft erstellt und beschreibt die typischen Einsatzmöglichkeiten von Moodle bei der Umsetzung von Prüfungen und semesterbegleitenden Übungen im politikwissenschaftlichen Kontext. Es dient dem Wissens- und Erfahrungstransfer innerhalb der Politikwissenschaft und ist anwendungsnah gestaltet, was den Lehrenden einen einfachen Einstieg in die Nutzung von Moodle als Übungs- und Prüfungsplattform ermöglicht. Gleichzeitig klärt es über die Besonderheiten auf, die es bei einer digitalen Prüfung über Moodle zu beachten gilt.

gehört ebenso dazu wie der vorliegende Beitrag und das Angebot, andere Lehrende der Politikwissenschaft bei der Prüfungsdigitalisierung via Moodle aktiv zu unterstützen.

Literatur

- Biggs, J. B./Tang, Catherine S.-K. (2011). *Teaching for Quality Learning at University*. Maidenhead: Open University Press.
- Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London: Routledge.
- Koch, L. (2015). *Lehren und Lernen: Wege zum Wissen*. Paderborn: Schöningh.
- Lambach, D. (2016). Die Lehre der Internationalen Beziehungen. In C. Masala/F. Sauer (Hrsg.), *Handbuch Internationale Beziehungen* (S. 1–22). Wiesbaden: Springer VS.
- Meyer, J. H. F./Land, R./Baillie, C. (2010). *Threshold Concepts and Transformational Learning*. Rotterdam: Sense Publishers.
- Middendorf, J./Pace, D. (2004). Decoding the disciplines: A model for helping students learn disciplinary ways of thinking. In: *New Directions for Teaching and Learning*, 98, 1–12.
- Mols, M. (2020): Politik als Wissenschaft: Zur Definition, Entwicklung und Standortbestimmung einer Disziplin. In H.-J. Lauth/Chr. Wagner (Hrsg.): *Politikwissenschaft. Eine Einführung*. 10., aktualisierte Auflage (S. 23–61). Paderborn: Ferdinand Schöningh.
- Rost, F. (2018). *Lern- und Arbeitstechniken für das Studium*. 8., überarbeitete und aktualisierte Auflage. Wiesbaden: Springer VS.
- Schulz, A. (2021): Handlungsfelder digitaler Prüfungen: Recht, Technik, Didaktik und Organisation. In *Hochschulforum Digitalisierung* (Hrsg.), *Digitale Prüfungen in der Hochschule*. Whitepaper einer Community Working Group aus Deutschland, Österreich und der Schweiz (S. 18–23). Essen: Hochschulforum Digitalisierung.
- Universität Duisburg-Essen (2023): *Diversity Monitoring*. Update 2022. Hg. v. Zentrum für Hochschulqualitätsentwicklung (ZHQE). Universität Duisburg-Essen. <https://zhqe-ude.de/public/dim/index.html> (zuletzt geprüft am 31.08.2023).

Schwellenkonzepte und Lehrinnovationen in fachlichen und fachdidaktischen Veranstaltungen im Lehramtsstudium Mathematik

Andreas Büchter, Dana Eilers und Florian Schacht

1 Einleitung

Wissen über Schwellenkonzepte kann für alle Arten von Lehr-Lern-Situationen die Grundlage für Lehrinnovationen bilden (siehe Beitrag von Cornelia Kenneweg in diesem Band). In diesem Beitrag betrachten wir Schwellenkonzepte von Studierenden in Mathematiklehramtsstudiengängen. In Deutschland besteht ein Mangel an Mathematiklehrkräften und Forschungsergebnisse deuten darauf hin, dass gerade im *Mathematiklehramtsstudium* der Grund für einen Studienabbruch oftmals eine inhaltliche Überforderung ist (vgl. Herfter, 2015, S. 70). Somit bildet die Beschäftigung mit Schwellenkonzepten in der universitären Mathematiklehramtsausbildung das Potenzial, einen positiven Unterschied zu machen. Viele Beiträge zu Schwierigkeiten von angehenden Mathematiklehrkräften im Studium beschäftigen sich ausschließlich mit fachmathematischen Herausforderungen. Wir beziehen in diesem Beitrag auch den mathematikdidaktischen Teil der universitären Lehrkräftebildung ein, der, wie in Abschnitt 3 zu lesen, andere Herausforderungen mit sich bringt.

Im Rahmen des Projektes PITCH haben wir über den Zeitraum von 2,5 Jahren zwei Veranstaltungen – „Elementare Geometrie“ und „Lern- und kognitionspsychologische Grundlagen des Mathematikunterrichts“ – vertieft betrachtet und dabei unter anderem Schwellenkonzepte identifiziert, um die es in diesem Beitrag geht. Die Veranstaltung „Elementare Geometrie“ ist eine fachmathematische Veranstaltung, die an der Universität Duisburg-Essen von Studierenden des Lehramts an Grundschulen in der Regel im dritten Bachelorsemester belegt wird. Die Veranstaltung „Lern- und kognitionspsychologische Grundlagen des Mathematikunterrichts“ ist eine mathematikdidaktische Veranstaltung, die an der Universität Duisburg-Essen von Studierenden des Lehramts an Gymnasien und Gesamtschulen bzw. Berufskollegs in der Regel im fünften Bachelorsemester belegt wird. Wir gehen im Folgenden genauer auf (geometrisches) Beweisen als Schwellenkonzept in der Fachveranstaltung und den Umgang mit pädagogisch-psychologischen und mathematikdidaktischen Texten als Schwellenkonzept in der Didaktikveranstaltung ein.

2 (Geometrisches) Beweisen als Schwellenkonzept

Bevor wir auf konkrete Schwellen beim (geometrischen) Beweisen zu sprechen kommen, thematisieren wir Geometrie als mathematische Disziplin und ihre Bedeutung in der Schule und der Lehrkräftebildung, wobei wir unsere Betrachtungen bereits auf geometrisches Beweisen fokussieren. Schließlich legen wir anhand eines konkreten Beweises potenzielle Schwellen dar und diskutieren, wie Studierende, insbesondere durch den Einsatz digitaler Medien, bei der Überwindung dieser Schwellen unterstützt werden können.

2.1 Geometrie zwischen Anschauung und abstrakter Deduktion

Die Mathematik ist eine beweisende Disziplin, in der neue Erkenntnisse immer aus bisher Bewiesenem oder aus Grundannahmen („Axiomen“) hergeleitet (also bewiesen) werden (Davis/Hersh, 1985, S. 348). So verhält es sich auch in der Geometrie. Sie gilt als ein Musterbeispiel einer deduktiven Theorie. Damit ist gemeint, dass sich die ganze Disziplin der Geometrie auf einzelne Grundbegriffe und Axiome reduzieren lässt, aus denen sich mittels logisch-geordneter Schlüsse und den damit verbundenen Sätzen und Beweisen das gesamte geometrische Wissen ableiten lässt. „Eine konsequente Realisierung dieses Aspektes erfordert (...) die Loslösung vom Anschauungsraum als einer inhaltlichen Interpretation der geometrischen Grundbegriffe und Axiome“ (Holland, 1996, S. 8f.).

Für den Mathematikunterricht in der Schule spielt eine solche abstrakte Auffassung der Geometrie nur in sehr geringem Umfang eine Rolle – etwa im Rahmen der Wissenschaftspropädeutik. Stattdessen werden eher inhaltlich-anschauliche Ansätze realisiert, gerade weil „man im Wechselspiel von Geometrie und Wirklichkeit (...) relativ schnell eine große Vielfalt beziehungsreicher (...) Themen entwickeln“ kann (Wittmann, 1987, S. VII). Auch wenn eine stark abstrakte Auffassung von Geometrie in der Schule keine bzw. nur eine geringe Rolle spielt, so sollte sie Studierenden neben inhaltlich-anschaulichen Zugängen dennoch bekannt sein.

Neben der Frage nach der Anschaulichkeit spielt – gerade mit Blick auf den Mathematikunterricht – das Spannungsfeld von statischen und dynamischen Aspekten in der Geometrie eine wichtige Rolle. Der Unterschied lässt sich an einem Beispiel verdeutlichen: So lässt sich etwa der Hypotenusensatz aus der Satzgruppe des Pythagoras entweder statisch beweisen mittels algebraischer Umformungen oder dynamisch beweisen, indem über entsprechende mathematische Abbildungen (z. B. Scherungen) das Quadrat der beiden Kathetenquadrate in das Hypotenusenquadrat überführt wird (vgl. Absatz 2.2). Gerade für die dynamische Erfahrung der geometrischen Zusammenhänge bietet sich die Arbeit mit einer Dynamischen-Geometrie-Software (im Folgenden: DGS) an (vgl. z. B. Helmerich/Lengnink, 2015). In dem Zusammenhang ist es bei der Nut-

zung von DGS im Mathematikunterricht (und auch im Lehramtsstudium Mathematik) wichtig, sich die Rolle von DGS im Zusammenhang mit dem Beweisen in der Geometrie bewusst zu machen.

Mit DGS können geometrische Zusammenhänge in der Regel nicht bewiesen werden, vielmehr kann eine DGS dazu dienen, mit geometrischen Objekten zu arbeiten, sie zu manipulieren und hinsichtlich ihrer Eigenschaften, Muster und Strukturen zu erkunden, um auf diese Weise etwa Hypothesen zu bilden, die dann im Anschluss ohne Zuhilfenahme entsprechender DGS bewiesen werden (müssen) (vgl. 2.2(4)). Die Visualisierungen können Studierende somit bei der Erschließung von geometrischen Sachverhalten unterstützen. In diesem Zusammenhang weist Wittmann (2014) auf ein Spannungsfeld hin: „Einerseits besteht die Gefahr, dass das Beweisbedürfnis weiter sinkt, weil eine dynamische Darstellung an sich schon eindrucksvoll erscheint und zu überzeugen vermag, so dass sie nicht mehr hinterfragt wird (...). Andererseits belegen empirische Studien, dass der Einsatz von DGS in Lernumgebungen zur Geometrie vielfältige Anlässe und Impulse zum Argumentieren liefern kann“ (Wittmann, 2014, S. 45). Mit Blick auf die Hochschullehre ist hierbei zu konstatieren, dass der Umgang mit digitalen Werkzeugen beim Beweisen weiter erforscht werden muss und digitale Werkzeuge grundsätzlich viel Potenzial für die Unterstützung von Studierenden beim Beweisen bieten (vgl. Stechemesser, 2023). Im Folgenden gehen wir auch exemplarisch darauf ein, wie wir dieses Potenzial im Rahmen der Veranstaltung nutzen.

2.2 Schwellen beim (geometrischen) Beweisen und Begegnung dieser Schwellen im Rahmen der Veranstaltung

Das geometrische Beweisen stellt einen zentralen Inhalt der Veranstaltung „Elementare Geometrie“ dar und ist gleichzeitig ein Schwellenkonzept für Studierende. Im Rahmen von Decoding-Interviews, Gesprächen im Veranstaltungsteam sowie der Analyse von Lernendendaten haben wir eine Vielzahl potenzieller Schwellen beim geometrischen Beweisen identifiziert. Einige dieser Schwellen diskutieren wir im Folgenden anhand eines exemplarischen Beweises. Als Beispiel dient uns der Beweis des Hypothenusensatzes aus der Satzgruppe des Pythagoras. Bevor wir auf die Schwellen zu sprechen kommen, thematisieren wir diesen Satz kurz.

Die Satzgruppe des Pythagoras umfasst drei bedeutsame Sätze: Kathetensatz (Satz von Euklid), Höhensatz und Hypotenusensatz (Satz von Pythagoras). Der Hypotenusensatz gilt als der bedeutsamste der drei Sätze, weswegen die Satzgruppe entsprechend benannt ist. „Die Satzgruppe des Pythagoras beschreibt ursprünglich Flächenbeziehungen, wird aber häufig auch zur Längenberechnung verwendet. So kann man mit Hilfe des Hypotenusensatzes stets die Länge der dritten Seite eines rechtwinkligen Dreiecks bestimmen, wenn die Längen der bei-

den anderen Seiten bekannt sind“ (Hefendehl-Hebeker, 2002, S. 84). Der Hypotenusensatz besagt:

Am rechtwinkligen Dreieck sind die Kathetenquadrate zusammen so groß wie das Hypotenusenquadrat: $a^2 + b^2 = c^2$.

Beweis (vgl. Hefendehl-Hebeker, 2002, S. 92):

Man wähle ein beliebiges Quadrat, dessen Seiten man in zyklischer Reihenfolge in Teilstrecken der Längen a und b zerlegt (vgl. Abb. 1). Auf diese Weise erhält man am Rand des Quadrates vier deckungsgleiche rechtwinklige Dreiecke, jeweils mit dem Flächeninhalt $0,5 \cdot ab$. Im Inneren des Quadrats erhält man auf diese Weise ein weiteres Quadrat mit der Seitenlänge c . Der Flächeninhalt des innenliegenden Quadrates lässt sich wie folgt in Abhängigkeit von a und b berechnen:

$$c^2 = (a + b)^2 - 4 \cdot 0,5 \cdot ab = a^2 + 2ab + b^2 - 2ab = a^2 + b^2, \\ \text{also } c^2 = a^2 + b^2$$

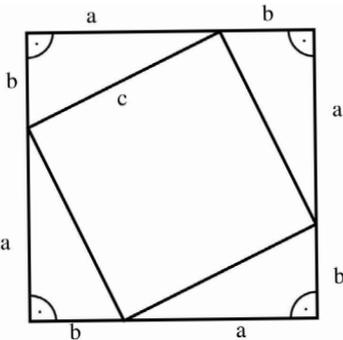


Abbildung 1: Beweisidee zum Hypotenusensatz (Quelle: Eigene Abbildung)

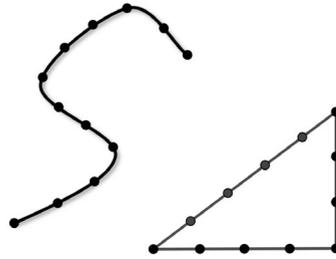


Abbildung 2: Zwölfknotenschnur, die zu einem rechtwinkligen Dreieck gespannt ist (Quelle: Eigene Abbildung)

Der Hypotenusensatz (Satz des Pythagoras) übt seit jeher eine große Faszination auf die Menschen aus – Gerwig (2021) etwa hat 365 verschiedene Beweise des Satzes in einem Buch veröffentlicht. Für eine Geometrieveranstaltung im Rahmen des Lehramtsstudiums Mathematik ist es vor diesem Hintergrund bedeutsam, etwas von der Faszination, der kulturgeschichtlichen Errungenschaft sowie von der geistigen Strenge zu vermitteln, die der Geometrie zugrunde liegt. Exemplarisch lässt sich dies sehr gut anhand der Satzgruppe des Pythagoras zeigen.

Im Folgenden werden ausgewählte Schwellen diskutiert, die im Rahmen (des Erlernens) von Beweisprozessen im Lehramtsstudium Mathematik eine

Rolle spielen – exemplarisch verdeutlicht am Beispiel des Hypotenusensatzes. Im Rahmen der Veranstaltung „Elementare Geometrie“ werden digitale Werkzeuge in Form einer DGS kontinuierlich eingesetzt. Die digitalen Werkzeuge werden grundsätzlich als heuristische Werkzeuge verwendet, gleichzeitig wurde die Veranstaltung so konzipiert, dass die DGS bei der Unterstützung des Übertretens entsprechender Schwellen unterstützen kann. Wir konkretisieren nun vier Schwellen beim (geometrischen) Beweisen.

1. Zugänge: Eine der besonderen Herausforderungen beim geometrischen Beweisen besteht in der Regel bereits darin, dass die Studierenden einen Zugang zum zugrundeliegenden Begriffsfeld (in dem Fall zur Betrachtung rechtwinkliger Dreiecke) finden müssen. Solche Begriffsfelder sind für Studierende oftmals aufgrund der abstrakten Natur mathematischer Objekte konzeptionell schwierig (vgl. „conceptually difficult knowledge“, Perkins, 2006, S. 38). Aus hochschuldidaktischer Sicht besteht daher zunächst der Anspruch, dass die Studierenden sich den mathematischen Kontext im Sinne Wittmanns (1987) inhaltlich-anschaulich erschließen. Häufig wird der Hypotenusensatz in Vorlesungen als unhinterfragter Zusammenhang postuliert, um ihn im nächsten Schritt zu beweisen – was bei Studierenden oftmals weder ein Beweisbedürfnis weckt noch inhaltlich-anschauliche Fundierungen ermöglicht. Eine sinnhafte Veranschaulichung im Sinne der bewussten Adressierung dieser Schwelle ist die Erzeugung rechter Winkel mit einer Zwölfknotenschnur (vgl. Abb. 2). Bereits die alten Ägypter haben auf diese Weise rechte Winkel beim Pyramidenbau erzeugt. Indem man die Schnur an drei der Knoten fixiert, lassen sich Dreiecke spannen. Fixiert man die Zwölfknotenschnur so, dass das Dreieck Seitenlängen mit jeweils 3, 4 und 5 Abschnitten hat, so erhält man ein Dreieck mit einer besonderen Eigenschaft: Es ist rechtwinklig (vgl. Abb. 2). Die Begründung, dass dies in dem Fall immer so ist, liefert der Hypotenusensatz, denn: Aus $3^2 + 4^2 = 5^2$ folgt mittels des Hypotenusensatzes, dass das Dreieck rechtwinklig ist. Studierende führen im Rahmen der entsprechenden Präsenzübungen solche Anwendungen mit realen Schnüren durch, sowohl um entsprechende Zugänge zu dem mathematischen Gegenstand zu bekommen als auch um die (kulturgeschichtliche) Bedeutung der mathematischen Gegenstände zu erfahren. An dieser Stelle besteht der Anspruch jedoch noch nicht darin, eine entsprechende (gar formale) Begründung für dieses Phänomen zu finden.

2. Hypothesen: Von großer Bedeutung ist in einem zweiten Schritt das Aufstellen von Hypothesen, die anschließend bewiesen werden. Eine solche Hypothesenfindung wird von Studierenden häufig als sehr herausfordernd wahrgenommen. Sie wissen oftmals nicht, wie sie beginnen können und es scheint ihnen möglicherweise, als würden Hypothesen für Mathematiker*innen vom Himmel fallen. Dies hängt damit zusammen, dass die Hypothesengenerierung oftmals stillschweigend verläuft und die mit ihr zusammenhängenden mentalen Prozesse von außen nicht sichtbar werden (vgl. „tacit knowledge“, Perkins, 2006, S. 40). In diesem Kontext bietet es sich an, eine DGS zu nutzen, die die Nutzen-

den bei der Hypothesengenerierung unterstützt. Das Ausprobieren mittels DGS kann ansonsten rein mental stattfindende Handlungen sichtbar und somit greifbarer machen. Ein Beispiel: Dargestellt sind ein rechtwinkliges Dreieck $\triangle ABC$ sowie die drei Quadrate über den Seiten a , b und c des Dreiecks. Die Studierenden können das rechtwinklige Dreieck an den Ecken manipulieren – auf diese Weise verändert es seine Form, allerdings sind die Einstellungen so vorgenommen, dass im Punkt C stets ein rechter Winkel ist, d. h. das Dreieck bleibt rechtwinklig (vgl. Abb. 3). Die Flächeninhalte der Quadrate über den Seiten a , b und c werden automatisch ausgerechnet und innerhalb der Quadrate angegeben. Die Studierenden haben nun die Aufgabe, Beobachtungen in Form entsprechender Hypothesen zu formulieren. Eine in diesem Zusammenhang sehr wichtige Beobachtung besteht darin, dass der Flächeninhalt des Quadrates über der Seite c (Hypotenuse) der Summe der Flächeninhalte der Quadrate über den Seiten a und b (Katheten) gleich – also: $c^2 = a^2 + b^2$. Dadurch, dass die Flächeninhalte numerisch in den jeweiligen Quadraten angegeben sind, wird eine entsprechende Beobachtung in der Regel von vielen Studierenden nach der Auseinandersetzung mit der DGS gemacht.

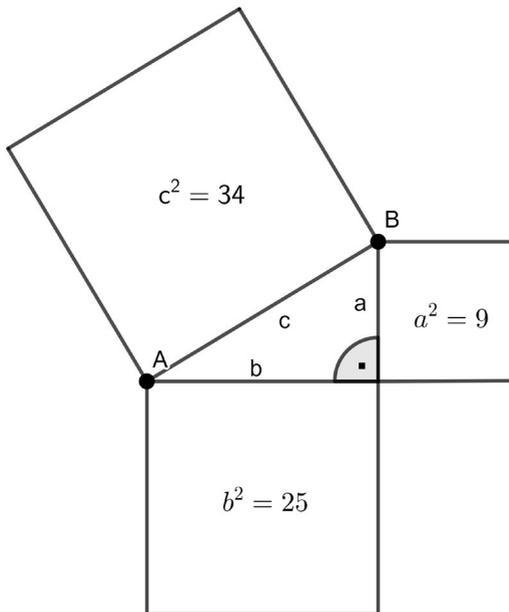


Abbildung 3: Erkundung von Zusammenhängen an rechtwinkligen Dreiecken mit einer DGS (Quelle: Eigene Abbildung)

3. *Beweisstruktur*: Wenn Studierende nun eine entsprechende Hypothese aufgestellt haben, kann die Beweisführung starten. Wichtig ist, dass man sich zunächst die Zeit nimmt, um einige Vorüberlegungen anzustellen: Welche Voraussetzungen sind gegeben? Was soll gezeigt werden? Mit welchen Argumenten lässt sich in dem Zusammenhang arbeiten? Eine zentrale Herausforderung besteht für Studierende in der Regel darin, zu verstehen, wodurch sich Voraussetzungen und die zu zeigenden Behauptungen unterscheiden. Nicht selten werden dann Zusammenhänge gezeigt, die im Rahmen des Kontextes bereits bekannt sind, weil sie eine der Voraussetzungen sind. Oder das, was gezeigt werden soll, wird während des Beweises (und somit bevor es gezeigt wurde) genutzt, als wäre es bereits gezeigt (ein sogenannter Zirkelschluss). Die Unterscheidung zwischen dem, was zu zeigen ist, und dem, was als Voraussetzung im Zuge der Argumentation genutzt werden kann, ist für Studierende zu Beginn in der Regel konzeptionell schwierig (vgl. „conceptually difficult knowledge“, Perkins, 2006, S. 38). Diese zentrale Schwelle beim mathematischen (nicht nur geometrischen) Beweisen wird im Rahmen der Veranstaltung sehr bewusst adressiert, beispielsweise indem die Studierenden ein Beweisschema nutzen können, bei dem sie die drei oben genannten Fragen Schritt für Schritt beantworten. Dies ist darüber hinaus Teil der mathematischen Enkulturation, da sich mathematische Beweise (immer) so strukturieren lassen und eine entsprechende Struktur hilft, um den Beweis dann auch konkret durchzuführen.

4. *Beweisidee*: In diesem Kontext zentral ist auch das Finden der Beweisidee. Nicht immer kann und muss die Beweisidee komplett selbstständig gefunden werden. Beweisen kann und muss erlernt werden. Das Finden der Beweisidee läuft oftmals, ähnlich wie das Aufstellen der Hypothesen, stillschweigend ab (vgl. „tacit knowledge“, Perkins, 2006, S. 40). Eine Möglichkeit, einen solchen Lernprozess (im Rahmen des Studiums, aber auch im Mathematikunterricht) zu unterstützen, besteht z. B. darin, Abbildungen (z. B. Abb. 1) zur Verfügung zu stellen und die Studierenden zu ermuntern, auf der Grundlage der Abbildung eine entsprechende Beweisidee zu finden. Wichtig ist hierbei zu betonen, dass im ersten Zugriff weder der Anspruch an eine formale Strenge besteht („Formulieren Sie die Beweisidee in eigenen Worten!“) noch der Anspruch an eine unmittelbar ausformulierte Argumentationskette. Mathematische Beweise entstehen im Rahmen eines kreativen Prozesses, in dem man Argumente prüft, wieder verwirft oder neue Zusammenhänge herstellt. Ist die Beweisidee einmal gefunden, ist es nicht mehr weit bis zum ausformulierten Beweis.

Die Diskussion der hier genannten Schwellen beim (geometrischen) Beweisen macht deutlich, inwiefern sich redliches mathematisches Arbeiten (ohne etwa auf geometrische Beweise zu verzichten) und die Berücksichtigung der Lernprozesse, die die Studierenden im Studium durchlaufen (mathematische Beweiskompetenz muss im Studium erlernt werden), nicht nur nicht ausschließen, sondern z. T. auch zur inhaltlichen Weiterentwicklung der Gestaltung der Lehrveranstaltungen führen können. Dabei können gerade im Fall der Geometrie

digitale Medien ein wichtiges Werkzeug sein, um das Schwellenkonzept (geometrisches) Beweisen im Lehramtsstudium Mathematik bewusst zu adressieren.

3 Umgang mit Fachtexten als Schwellenkonzept

Fachdidaktische Lehrveranstaltungen – wie „Lern- und kognitionspsychologische Grundlagen des Mathematikunterrichts“ – sollen, gemeinsam mit den fachlichen (und den bildungswissenschaftlichen) Lehrveranstaltungen, insgesamt zu einer professionsorientierten universitären Lehrkräftebildung beitragen. Bevor wir die spezifischen Herausforderungen der hier konkret betrachteten fachdidaktischen Lehrveranstaltung darstellen und reflektieren, ordnen wir die Zielsetzung der Lehrveranstaltung in die grundsätzliche Konzeption der universitären Lehrkräftebildung ein. Vor diesem Hintergrund wird dann deutlich, dass sich die Herausforderungen in fachlichen und fachdidaktischen Lehrveranstaltungen (in der Regel grundlegend) unterscheiden dürften.

3.1 Allgemeine Zielsetzungen fachdidaktischer Lehrveranstaltungen

Die wesentliche Orientierung für die aktuelle Konzeption der universitären Phase der Lehrkräftebildung in Nordrhein-Westfalen basiert auf den Empfehlungen einer wissenschaftlichen Kommission („Baumert-Kommission“, vgl. MIWFT, 2007). Dabei werden die Zielsetzungen der ersten und zweiten Phase der Lehrkräftebildung klar voneinander abgegrenzt:

Der Erwerb konzeptionell-analytischer Kompetenz in der universitären Ausbildung ist ebenso wenig austauschbar wie der Erwerb von reflexiv gesteuerter Handlungskompetenz in der Zweiten Phase. (ebd., S. 32)

Eine umfangreiche und auf eine potenziell mehrere Jahrzehnte dauernde Berufstätigkeit vorbereitende Wissensgrundlage bildet dabei die Basis „konzeptionell-analytischer Kompetenz“:

In Deutschland soll in der ersten wissenschaftlichen Ausbildungsphase das Berufs- und Fachwissen erworben werden, das Voraussetzung kompetenten Unterrichtens und Grundlage des Starts in eine reflexive Unterrichts- und Berufspraxis ist. Bezogen auf die professionelle Tätigkeit wird ein entwicklungsfähiges begriffliches Instrumentarium zur Verfügung gestellt, das Unterrichtsplanungen anleiten und Berufserfahrungen ordnen kann, ohne bereits die Handlungskompetenz zu vermitteln, die erst in der formativen Novizenzeit des Vorbereitungsdienstes (sog. Zweite Phase) erworben wird. (ebd., S. 14)

Damit wird einerseits betont, dass eine entsprechende Wissensgrundlage erworben, andererseits aber auch der spätere Verwendungszusammenhang während der Berufstätigkeit berücksichtigt werden soll. Dies bedeutet für die Lehre, dass die Bedeutung des Wissens für die spätere Berufstätigkeit immer wieder reflektiert und die Verwendung im Beruf angebahnt werden sollte. Dabei wird zumindest grundlegendes Fachwissen als Basis fachdidaktischer Lehrveranstaltungen vorausgesetzt, die dann eine Scharnierfunktion zwischen fachlichen Lehrveranstaltungen und der späteren Berufstätigkeit übernehmen. In ihnen soll sich insbesondere häufiger explizit auf die Berufstätigkeit bezogen werden (ebd., S. 40).

Während die Themen und Formate der fachlichen Lehrveranstaltungen in Lehramtsstudiengängen Mathematik standortübergreifend weitgehend gleich sind, unterscheiden sich Titel, Inhalte und Formate der fachdidaktischen Lehrveranstaltungen häufig deutlich. Für die Realisierung relativ abstrakter Zielsetzungen (z. B. in den „Ländergemeinsame[n] inhaltlichen Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung“ der Kultusministerkonferenz) haben sich vor Ort jeweils sehr eigene – häufig in Abhängigkeit von fachdidaktischen Forschungsschwerpunkten entwickelte – Wege herausgebildet.

3.2 Konkrete Zielsetzungen und grundsätzliche Gestaltung der betrachteten Lehrveranstaltung

Die Lehrveranstaltung „Lern- und kognitionspsychologische Grundlagen des Mathematikunterrichts“ ist im fünften Bachelorsemester des Lehramtsstudiengangs für Gymnasien und Gesamtschulen bzw. Berufskollegs verortet. Für die ersten vier Semester sind sechs fachliche und zwei fachdidaktische Lehrveranstaltungen vorgesehen. Die beiden fachdidaktischen Veranstaltungen sind „stoffdidaktisch“ ausgerichtet, beschäftigen sich also mit der didaktischen Sachanalyse bzw. fachdidaktischen Rekonstruktion von mathematischen Inhalten. Das Lehren von Mathematik erscheint hier vorrangig als auf die Sache bezogenes Handeln.

In der Lehrveranstaltung „Lern- und kognitionspsychologische Grundlagen des Mathematikunterrichts“ erscheint das Lehren hingegen vorrangig als auf das (individuelle) Lernen (der Schüler*innen) bezogenes Handeln. Damit die angehenden Lehrkräfte während ihrer späteren Berufstätigkeit die Lerntätigkeiten ihrer Schüler*innen optimal anregen können, sollen sie typische fachbezogene Denk-, Lern- und Bearbeitungsprozesse kennen und die individuellen Ausprägungen dieser Prozesse bei den einzelnen Schüler*innen berücksichtigen können. Dementsprechend werden in der Lehrveranstaltung zahlreiche hierfür nützliche pädagogisch-psychologische und mathematikdidaktische Theorien thematisiert:

- Taxonomie der Wissensformen, Prozesse des Wissenserwerbs
- Situiertheitsperspektive, Transfer und Anwendung

- Lernstrategien und selbstreguliertes Lernen
- Kognitive Stile (nach Schwank)
- Lernstilmodelle (insbesondere nach Kolb)
- Theorie der kognitiven Entwicklung (nach Piaget)
- Theorie der Repräsentationsmodi (nach Bruner)
- Prinzip der operativen Durcharbeitung
- Sprachliches und fachliches Lernen im Mathematikunterricht
- Räumliches Denken
- Algebraisches Denken
- Mathematisches Denken
- Begriffsbildung aus psychologischer und mathematikdidaktischer Sicht
- Grundvorstellungen zu mathematischen Inhalten und individuelle Vorstellungen
- Fehlerbegriff, Fehlerarten, Umgang mit Fehlern und „negatives“ Wissen
- Handlungsleitende Diagnose

Die Lehrveranstaltung besteht wöchentlich aus einer zweistündigen Vorlesung, einer schriftlichen Hausübung sowie einer einstündigen Präsenzübung. Die Aneignung der o. g. Theorien soll zunächst über die vorbereitende individuelle Lektüre einschlägiger Fachtexte jeweils vor der Vorlesung geschehen. Die Vorlesung dient dann, nach einem Inverted-Classroom-Ansatz, der Diskussion und Vertiefung der Texte. Insbesondere bei pädagogisch-psychologischen Theorien und Texten steht dabei die Konkretisierung für typische Situationen des Mathematikunterrichts im Vordergrund. Die weitergehende Anwendung und Reflexion der Theorien erfolgt dann in den schriftlichen Hausübungen und den Präsenzübungen. Typische Übungsaufgaben sind (vor dem Hintergrund einer benannten Theorie):

- Analyse von Schulbuchauszügen
- Analyse von Schüler*innenbearbeitungen
- Angabe von konkretisierenden Beispielen
- Konstruktion von Aufgaben- bzw. Lernumgebungen
- Herstellen von Bezügen zu anderen Theorien

Sowohl die Diskussion und Vertiefung der Texte in der Vorlesung als auch fast alle Übungsaufgaben (bis auf „Herstellen von Bezügen zu anderen Theorien“) sind dabei konsequent auf die spätere Berufstätigkeit orientiert. Durch die Anwendung der Theorien auf reale Schulbuchauszüge und Schüler*innenbearbeitungen sowie fiktive Unterrichtssituationen wird die o. g. Zielsetzung (optimale Anregung der Lernfähigkeit in der späteren Berufstätigkeit) durch die Entwicklung „konzeptionell-analytischer Kompetenz“ verfolgt.

3.3 Das rekonstruierte Schwellenkonzept

Die zuvor dargestellten Zielsetzungen und die grundsätzliche Gestaltung der Lehrveranstaltung stellen bereits die Ausgangssituation zu Beginn des Projektes PITCH dar. Im Rahmen des Projektes sollte einerseits die Kompetenzorientierung der Prüfung weiter gesteigert werden. Die Anforderungen sind nun fast ausschließlich durch die Anwendung oder Reflexion der Theorien geprägt. Andererseits sollte wiederkehrend auftretenden Herausforderungen und Überforderungen von Studierenden durch ein angepasstes Lehr-Lern-Arrangement begegnet werden.

Sowohl in Decoding-Interviews mit den Lehrenden als auch bei der Reflexion der Lehrveranstaltung mit Hilfskräften, die bei der lernförderlichen Rückmeldung zu den schriftlichen Hausübungen und der Korrektur der Klausur (Prüfungsleistung) mitwirken, wurde die Textarbeit als eine zentrale Herausforderung für Studierende identifiziert. Genauer geht es hier um die individuelle Aneignung der Theorien auf der Basis einschlägiger Fachtexte, die einer Anwendung der jeweiligen Theorie vorausgeht. Die Diskussionen während der Vorlesungen und der Präsenzübungen sowie die Bearbeitungen von schriftlichen Hausübungen und Klausuraufgaben deuteten darauf hin, dass zahlreiche Studierende den eigentlichen Kern der jeweils aufgerufenen Theorie höchstens oberflächlich erfasst haben. Eine weitgehend stimmige Anwendung oder Reflexion der Theorie scheiterte unseren Beobachtungen nach oftmals nicht am Anwenden oder Reflektieren selbst, sondern am vorgelagerten Theorieverständnis. Dabei ist klar, dass sich ein vertieftes Theorieverständnis auch im Wechselspiel mit (stimmiger) Anwendung und Reflexion im Diskurs (mit sich selbst, mit weiteren Texten, mit anderen Personen) entwickelt. Als Schwelle erscheint hier aber schon das erste grobe Verständnis der Theorie. Bei der Reflexion dieses Schwellenkonzepts im Lehrveranstaltungs- und Projektteam wurde deutlich, dass die Lernausgangslagen der Studierenden hinsichtlich des selbstständigen Aneignens von Inhalten aus Texten zuvor deutlich überschätzt wurden. In der gymnasialen Oberstufe werden heute überwiegend deutlich kürzere Texte selbstständig durchgearbeitet. Je nach Kombination der beiden studierten Unterrichtsfächer wird sich im Studium zuvor auch nicht unbedingt mit längeren Fachtexten auseinandergesetzt – und wenn es der Fall ist, findet die Textarbeit häufig ohne Anleitung statt (vgl. „tacit knowledge“, Perkins, 2006, S. 40).

3.4 In der Lehrveranstaltung erprobte Lösungsansätze

Im Laufe von drei Durchgängen der Lehrveranstaltung wurden mehrere Elemente entwickelt und erprobt, die einerseits die Studierenden bei der Aneignung von Theorien aus Fachtexten unterstützen und andererseits – im Sinne eines constructive alignment (Biggs/Tang, 2011) – die Schwerpunktsetzung auf

Anwendung und Reflexion von Theorien verstärken sollen. Die folgenden Elemente haben sich bewährt und werden weiterentwickelt:

- Die Fokussierung auf die wesentlichen Aspekte der Theorien wird durch Lektüreaufträge unterstützt. Diese konkretisieren das Erkenntnisinteresse, vor dessen Hintergrund der Text gelesen werden soll und geben dadurch Orientierung beim Lesen. So soll Wichtiges einfacher von Unwichtigem abgegrenzt werden können. Unsere Erfahrungen zeigen, dass dadurch Nebensächliches seltener Gegenstand von Diskussionen oder Aufgabebearbeitungen wird.
- Zusätzlich zu den im Abschnitt 3.2 beschriebenen Elementen der Lehrveranstaltung wurde ein optionales Tutorium eingeführt. In diesem können Studierende unter Moderation und Mitwirkung einer wissenschaftlichen Mitarbeiterin Verständnisfragen oder Konkretisierungs- und Anwendungsüberlegungen zum Text diskutieren. Das Tutorium bietet somit neben der Vorlesung einen weiteren Rahmen, in dem der Prozess des Aneignens von Fachtexten expliziert wird, und führt bei den (freiwillig) Teilnehmenden zu einem tieferen Verständnis.
- Das selbstständige Finden von konkretisierenden Beispielen bzw. die Konstruktion von Aufgaben- oder Lernumgebungen trägt, sofern dies tragfähig geschieht, zum Theorieverständnis bei, ist aber alles andere als leicht. Daher werden zu Beginn des Semesters Beispiele vorgegeben, die auf der Grundlage einer Theorie analysiert werden sollen. Im weiteren Verlauf des Semesters sollen dann zunehmend selbst Konkretisierungen vorgenommen werden. Dabei ist erkennbar, dass häufig auf zuvor kennengelernte Beispiele zurückgegriffen wird und diese – mal mehr, manchmal aber weniger passend – zum jeweiligen Arbeitsauftrag angepasst werden.
- Das Herstellen von Bezügen zwischen Theorien wird zunächst vorrangig in der Vorlesung in einem gemeinsamen Diskurs oder exemplarisch durch den Dozenten geleistet. Gegen Ende des Semesters sollen die Studierenden dies zunehmend selbst leisten. Die „Krönung“ stellt das offene Herstellen von Bezügen zu anderen Theorien dar (vgl. Abb. 4). Mit diesem offenen Arbeitsauftrag werden die Studierenden auch als Teilnehmende am mathematikdidaktischen Diskurs ernstgenommen.
- In der Klausur wird auf die Anwendung und Reflexion der Theorien fokussiert. Reproduktion wird höchstens implizit angesprochen. Diese Anforderungen werden in den Übungen entsprechend vorbereitet. Die Entlastung von der Reproduktion (und dem vorangehenden Auswendiglernen) von Faktenwissen erfolgt dadurch, dass die Studierenden fünf individuell (auch beidseitig) gestaltete DIN-A4-Merkblätter in die Klausur mitnehmen dürfen. Es ist beobachtbar, dass sich bei vielen Studierenden das Lernverhalten bei der Vorbereitung auf die Klausur dadurch in eine wünschenswerte Richtung verändert, indem beim Lernen mehr Tiefenstrategien und weniger Oberflächenstrategien angewendet werden.

Aufgabe 15 (8 Punkte)

In Text 10* und in Vorlesung 12 wurden zahlreiche inhaltliche Bezüge zwischen dem Grundvorstellungskonzept und anderen Theorien bzw. Konzepten, die in der Lehrveranstaltung bereits thematisiert wurden, hergestellt.

Wählen Sie zwei bereits in der Lehrveranstaltung thematisierte Theorien bzw. Konzepte, zu denen das Grundvorstellungskonzept inhaltliche Bezüge aufweist, aus und erläutern Sie diese Bezüge. Geben Sie dabei auch konkret an, auf welche Stellen in welchen Quellen Sie sich beziehen.

Abbildung 4: Schriftliche Übung zum Herstellen von Bezügen zwischen Theorien (Quelle: Eigene Abbildung)

4 Zusammenfassung

Ausgangspunkt der Betrachtungen waren zwei konkrete Veranstaltungen im Mathematiklehramtsstudium an der Universität Duisburg-Essen. Trotz der spezifischen Kontexte dieser Veranstaltungen können die identifizierten Schwellenkonzepte als typisch für die Mathematiklehrkräftebildung angesehen werden. Mathematisches Beweisen kann als ein Schwellenkonzept der fachmathematischen Lehrkräftebildung und der Umgang mit Fachtexten als ein Schwellenkonzept der mathematikdidaktischen Lehrkräftebildung angesehen werden.

Zur Unterstützung der Überwindung der Schwellenkonzepte haben wir in beiden Veranstaltungen den Studierenden die Wichtigkeit der entsprechenden Inhalte (Beweisen, Umgang mit Fachtexten) kontinuierlich verdeutlicht – nicht zuletzt dadurch, dass die Inhalte von Beginn an als Teil des summativen Assessments (in diesen Fällen: Klausur) kommuniziert wurden. Wir haben auch die formativen Assessments angepasst, zum Beispiel durch eine schrittweise Heranführung an das Anwenden von Theorien auf selbstentwickelte Praxisbeispiele in der Veranstaltung „Lern- und kognitionspsychologische Grundlagen des Mathematikunterrichts“.

Beide Schwellenkonzepte zeichnen sich durch implizites Wissen (im Sinne von Perkins „tacit knowledge“ aus. Wir haben in beiden Veranstaltungen Angebote für Studierende geschaffen, die dieses implizite Wissen sichtbar machen – in „Lern- und kognitionspsychologische Grundlagen des Mathematikunterrichts“ durch das gemeinsame Verhandeln des Textverständnisses in der Vorlesung und im Zusatztutorial und in „Elementare Geometrie“ unter anderem durch die Bereitstellung einer asynchronen digitalen Vorlesung, in der interaktive Elemente schrittweise zur eigenen Beweisführung hinleiten.

Die Überwindung dieser Schwellenkonzepte fordert von den Studierenden jeweils ganz unterschiedliche Denk- und Arbeitsweisen. Das mathematische Beweisen fußt auf abstraktem Denken und logischem Schließen und der

Umgang mit mathematikdidaktischen und pädagogisch-psychologischen Texten steht in der Tradition der Geisteswissenschaften. Somit verdeutlichen diese beiden Schwellenkonzepte die vielfältigen Anforderungen, die das Mathematiklehramtsstudium auch durch die Unterschiedlichkeit zwischen Mathematik und Mathematikdidaktik bereithält.

Literatur

- Biggs, J. B./Tang, C. S. (2011). *Teaching for Quality Learning at University: What the Student Does* (4. Aufl.). Buckingham, England: Open University Press.
- Davis, P./Hersh, R. (1985). *Erfahrung Mathematik*. Basel: Birkhäuser Verlag.
- Gerwig, M. (2021). *Der Satz des Pythagoras in 365 Beweisen. Mathematische, kulturgeschichtliche und didaktische Überlegungen zum vielleicht berühmtesten Theorem der Mathematik*. Berlin: Springer Spektrum.
- Hefendehl-Hebeker, L. (2002). *Maße und Funktionen im Geometrieunterricht der Sekundarstufe I*. Augsburg: mathematisch-naturwissenschaftliche Schriften 41. Augsburg: Wißner-Verlag.
- Helmerich, M./Lengnink, K. (2016). *Einführung Mathematik Primarstufe – Geometrie*. Berlin: Springer Spektrum.
- Herfter, C. (2015). Der Abbruch des Lehramtsstudiums – Zahlen, Gründe und Emotionserleben. In: *Zeitschrift für Evaluation*, 14(1), 57–82.
- Holland, G. (1996). *Geometrie in der Sekundarstufe. Didaktische und methodische Fragen* (2. Aufl.). Heidelberg, Berlin, Oxford: Spektrum Akademischer Verlag.
- MIWFT (Ministerium für Innovation, Wissenschaft, Forschung und Technologie) (Hrsg.) (2007). *Ausbildung von Lehrerinnen und Lehrern in Nordrhein-Westfalen. Empfehlungen der Expertenkommission zur Ersten Phase*. https://www.aqas.de/downloads/Lehrerbildung/Bericht_Baumert-Kommission.pdf
- Perkins, D. (2006). *Constructivism and troublesome knowledge. Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge*. In J. Meyer/R. Land (Hrsg.), *Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge* (S. 33–47). New York / London: Routledge.
- Stechemesser, J. M. (2023). *Geometrische Beweisprozesse von Lehramtsstudierenden bei der Arbeit mit interaktiven Büchern*. In: IDMI-Primar Goethe-Universität Frankfurt (Hrsg.), *Beiträge zum Mathematikunterricht 2022* (S. 1245–1249). <https://eldorado.tu-dortmund.de/handle/2003/41430>
- Wittmann, E. C. (1987). *Elementargeometrie und Wirklichkeit: Einführung in geometrisches Denken*. Braunschweig, Wiesbaden: Vieweg.
- Wittmann, G. (2014). *Beweisen und Argumentieren*. In H.-G. Weigand/A. F. Filler/R. Hölzl/S. Kuntze/M. Ludwig/J. Roth/B. Schmidt-Thieme/G. Wittmann (Hrsg.) *Didaktik der Geometrie für die Sekundarstufe I* (2. Auflage) (S. 35–54). Berlin: Springer Spektrum.

Gestaltung lern(enden)zentrierter Veranstaltungen im bildungswissenschaftlichen Studium der Bachelor-Lehramtsstudiengänge der UDE

Ruth M. Ingendoh und Angela Heine

Die *Universität Duisburg-Essen (UDE)* zählt – neben den Universitäten der Städte Köln und Münster – zu den größten lehrkräftebildenden Hochschulen Nordrhein-Westfalens.¹ Unabhängig von der gewählten Schulform², Schulstufe und Fächerkombination absolvieren alle Lehramtsstudierenden der UDE im Rahmen ihrer Bachelor- und Master-Studiengänge obligatorische bildungswissenschaftliche Studienanteile, die im Fall der Bachelorstudiengänge durchschnittlich ungefähr einem Fünftel³ des zu erbringenden Gesamtarbeitsaufwands eines Lehramtsstudiums an der UDE entsprechen. Daraus ergibt sich der Stellenwert von Lehrveranstaltungen wie den im Folgenden betrachteten Vorlesungen im Fach Psychologie, die von der *Fakultät für Bildungswissenschaften* der UDE in die Lehramtsstudiengänge eingebracht werden.

1 Ausgangslage

Das zu absolvierende *bildungswissenschaftliche* Studium der Bachelorstudiengänge aller Lehramtsstudierenden der UDE gliedert sich in vier Module, von denen eines – das Modul *Psychologie (MY)*⁴ – auf die Vermittlung pädagogisch-psychologischer und psychodiagnostischer Grundlagen abzielt, die für ein erfolgreiches Handeln von Lehrkräften in der schulischen Praxis unabdingbar sind (KMK, 2004). Das MY wiederum entspricht annähernd einem Fünftel der zu erwerbenden Leistungspunkte des bildungswissenschaftlichen Anteils der Bachelor-Lehramtsstudiengänge der UDE. Den Kern des MY bilden Grundgenvorlesungen⁵ in mindestens vier psychologischen Teildisziplinen, die Studie-

1 Vgl. dazu die für die UDE typischen Bachelor-Kohortengrößen in den vom *Monitor Lehrerbildung (2023)* bereitgestellten Statistiken für die Erhebung im ersten Halbjahr 2022.

2 Die UDE bildet Lehrkräfte für *Grundschulen (GS)*, *Haupt-/Real-/Sekundar-/Gesamtschulen (HRSGe)*, *Gymnasium* und *Gesamtschulen (GyGE)* sowie *Berufskollegs (BK)* aus.

3 Zwischen 13 % der Leistungspunkte für die BK- und 23 % für die HRSGe-Studiengänge.

4 Die Lehramtsoption BK sieht im Bachelorstudium ein Modul *Psychologie & Soziologie* vor.

5 Thematische Schwerpunkte der vier Kernvorlesungen sind *Einführung in die Psychologie*, *Lehr-/Lernpsychologie*, *Pädagogisch-psychologische Diagnostik* und *Inklusion*.

rende *aller* Lehramtsstudiengänge durchlaufen und mit einer schriftlichen Prüfung abschließen müssen.

Bestimmte Aspekte, die zunächst vor allem strukturelle Gegebenheiten der UDE-Lehramtsstudiengänge betreffen, aber auch Faktoren, die mit der ursprünglichen Ausgestaltung der Lehre im MY zusammenhängen, sind mit Blick beispielsweise auf die Empfehlung des Wissenschaftsrats (2008) als der Qualität von Hochschulausbildung nicht zuträglich zu werten. Als Herausforderung auf *struktureller* Ebene wäre insbesondere zu nennen:

- Mit regelmäßig zwischen 700 und 900 Teilnehmer:innen pro Semester⁶ handelt es sich bei den Kernvorlesungen im MY um Massenvorlesungen mit „ungünstigen Betreuungsrelationen“ (Wissenschaftsrat, 2022, S. 78).
- Es gibt im MY keine schulformspezifischen Studienangebote, vielmehr besuchen künftige Grundschullehrkräfte diese Vorlesungen ebenso wie Studierende, die zukünftig an Berufskollegs tätig sein werden.
- Das MY ist, in Abhängigkeit von der gewählten Schulform, curricular in unterschiedlichen Fachsemestern verankert, weshalb die Vorlesungen grundsätzlich Studierende mit unterschiedlichem studiumsbezogenen Horizont adressieren müssen – von Erstsemesterstudierenden, die gerade ein Studium aufgenommen haben, bis zu Studierenden im sechsten Semester, die bereits mit dem Schreiben ihrer Studienabschlussarbeiten befasst sind.
- Es gibt keine relevanten curricular verankerten Lehrangebote, die die Kernvorlesungen im MY flankieren, um Studierende bei der Auseinandersetzung mit den zunächst grundsätzlich eher als fachfern wahrgenommenen Inhalten der Psychologie zu unterstützen.⁷

Aber auch die Ausgestaltung der Lehre im MY vor Beginn des *PITCH*-Projekts ließ unmittelbar Ansatzpunkte für Verbesserung der Lehre erkennen:

- Abhängig von der Schulform sind im Modul vier bis sechs einstündige Vorlesungen konzipiert, die den Stoff unterschiedlicher psychologischer Teildisziplinen behandeln. Dadurch war der Lernstoff ausgesprochen umfangreich und stark auf Vermittlung von Grundlagenwissen ausgerichtet.
- Eine inhaltliche Integration des im Rahmen der Lehrveranstaltungen im MY vermittelten Lernstoffs war nicht gegeben, vielmehr standen die inhaltlichen Schwerpunkte unverbunden nebeneinander.
- Die schriftlichen Modulabschlussprüfungen setzten sich aus den Schwerpunkten generierten separaten Frageteilen zusammen.

Diese Vorlesungen werden in einzelnen Lehramtsstudiengängen ergänzt durch Vorlesungen wie *Differenzielle* (HRSGe) und *Entwicklungspsychologie* (GS) bzw. *Berufliche Sozialisation* (BK).

- 6 Im aktuell laufenden Wintersemester 2023/24 besuchen mehr als 950 Studierende die Vorlesungen im MY. In Sommersemestern liegt die Zahl typischerweise etwas niedriger.
- 7 Lediglich ein Teil der Studierenden (HRSGe) absolviert zusätzlich ein 1SWS-Seminar.

- Primär prüfungsökonomischen Erwägungen sowie der Ausrichtung der Lehre auf Vermittlung von Grundlagenwissen der psychologischen Teildisziplinen geschuldet, spiegeln die Formate der Modulabschlussprüfungen ein Prüfungskonzept wider, das v. a. die Selektionsfunktion von Leistungsmessung in den Vordergrund stellt (Reinmann, 2022).

Vor dem Hintergrund dieser Studienbedingungen ist nicht nur das chronisch unbefriedigende Leistungsniveau der Studierenden bei den Modulabschlussprüfungen nachvollziehbar – bzw. *erwartbar* –, sondern auch die Ergebnisse von Befragungen der Studierenden im MY, die erstmals im Jahr 2019 umgesetzt und in der Folge regelmäßig weitergeführt wurden.

2 Ist-Zustand bei Start des PITCH-Projekts

Im Rahmen erster Fragebogenerhebungen, die in den Vorlesungen im MY durchgeführt wurden, um die studentische Perspektive auf die Gestaltung der Lehre zu beleuchten, wurden Daten zu leistungsrelevanten Faktoren in Bezug auf einerseits den *bildungswissenschaftlichen* Teil des Lehramtsstudiums und andererseits die *fachspezifischen* Studienanteile erhoben. Die Ergebnisse dieser ersten Befragungen, die durch Erhebungen von Studierendenkohorten in den Folgesemestern bestätigt wurden, zeichnen ein eindeutiges Bild.

Einschätzungen der Studierenden, beispielsweise in Bezug auf die zu bewältigende *Lernstoffmenge* oder die *Komplexität* der Studieninhalte, fielen signifikant ungünstiger für das *bildungswissenschaftliche* Studium im Vergleich zum Studium der Fächer aus [$t(319) = -8,94, p < ,001$ bzw. $t(319) = -2,96, p = ,003$]. Diese Effekte blieben selbst dann weitestgehend stabil, wenn bestimmte Fächerkombinationen (z. B. MINT-Fächer) bei der Auswertung gesondert betrachtet wurden. Unterschiedliche motivationale Faktoren wie *akademische Kontrollüberzeugungen* (Perry et al., 2001) oder *wertbezogene Einschätzungen* der Studieninhalte (Eccles/Wigfield, 2002) fielen signifikant schwächer für die *bildungswissenschaftlichen* Studienanteile im direkten Vergleich zum Studium der Fächer aus [akademische Kontrollüberzeugungen: $t(345) = -3,40, p = ,001$; Wertkomponenten *intrinsisch*: $t(343) = -11,09, p < ,001$ bzw. *nutzenbezogen*: $t(347) = -5,29, p < ,001$]. Auch in Hinblick auf das Erleben spezifischer *akademischer Emotionen* als relevante Determinanten studiumsbezogener Leistung und Leistungsbereitschaft (Pekrun, 2006) ergaben sich auffällige Ergebnisse, die auf eine insgesamt deutlich negativ gefärbte affektive Befindlichkeit der Studierenden im Zusammenhang mit ihrem *bildungswissenschaftlichen* im Vergleich zum Fächerstudium hinwiesen. Besonders erwähnenswert sind wiederholt festgestellte, im Vergleich hohe Werte für Selbsteinschätzungen in Bezug auf *Ärger* als eine negativ-aktivierende lernbezogene Emotion [$t(346) = 6,75, p < ,001$; Pekrun (2018)], die intrinsische Lernmotivation und Leistungsbereitschaft nachweislich ungünstig beeinflusst (Pekrun et al., 2002).

Dass solche Befragungsergebnisse auf Handlungsbedarf hinweisen, steht außer Frage. Und in Anbetracht der Tatsache, dass Intervention auf *struktureller* Ebene schwer unmittelbar umsetzbar ist, lag es nahe, die konkrete Gestaltung von Lehre und Prüfungen zunächst für das MY kritisch zu hinterfragen, um auf dieser Ebene Veränderung planen und umsetzen zu können.

3 Innovation der Lehre im Rahmen des PITCH-Projekts

Das Ziel der Psychologie besteht darin, uns eine völlig neue Sichtweise auf die Dinge zu vermitteln, die wir am besten kennen (*Paul Valery, 1943*).

Die Basis für die erfolgte Neugestaltung der Lehrveranstaltungen im MY wurde im Rahmen auf die inhaltliche Ausrichtung und die jeweiligen Arbeitsphasen des Teilprojekts zugeschnittener Workshops mit Expert:innen im Bereich der Hochschullehre erarbeitet. Besonders zwei Schwerpunkte der Diskussionen mit den assoziierten externen Berater:innen des *PITCH*-Vorhabens erwiesen sich als richtungweisend für die Projektarbeit.

3.1 Konzeptuelle Vorbereitung des PITCH-Teilprojekts

Identifikation von Schwellenkonzepten. Den Ausgangspunkt aller Überlegungen, wie die Vorlesungen des MY studierendenzentriert gestaltet werden können, bildeten *Decoding-Interviews*, die auf die Sichtbarmachung von lern- und entwicklungshinderlichen Threshold Concepts abzielen (vgl. Beitrag von Cornelia Kenneweg in diesem Band). Ein als motivations- und leistungsbeschränkend identifizierter Faktor, der durch diesen spezifischen methodischen Zugang zur Analyse des Status-Quo in der Lehre herausgearbeitet werden konnte, war, dass Studierenden der – zumindest eingangs des Studiums häufig genug primär als *berufsvorbereitend* verstanden – Lehramtsstudiengänge, die Relevanz eines fundierten psychologischen Fach- und Methodenwissens für Handlungszusammenhänge in der schulischen Praxis nicht hinreichend nachvollziehbar gemacht wurde. Aber gerade im Kontext *psychologischer* Erklärungszusammenhänge, die auf Begrifflichkeiten, umschriebene inhaltliche Konzepte und methodische Zugänge verweisen, die durch ein laientheoretisches Vorverständnis besonders tief geprägt sind (Furnham, 1988), müssen explizit Gelegenheiten sowie hinreichend Freiräume für Reflexion geschaffen werden, um zunächst implizite Annahmen über behandelte Sachverhalte bewusst und damit potenziell veränderbar zu machen.

Kompetenzentwicklung statt Stoffvermittlung. Konzeptueller Dreh- und Angelpunkt der neu gestalteten Vorlesungen im MY sind nachvollziehbare übergeordnete *kompetenzorientierte Lernzieldefinitionen*, die den Studierenden eingangs der Veranstaltungsreihe vorgestellt und als für die Lehre verbindliche – einforderbare – Grundorientierung an die Hand gegeben werden. Ausgangspunkt

dieser neu formulierten Lernziele ist der Kompetenzbegriff von Weinert (2001), nach dem Handlungsdisposition durch

die bei Individuen *verfügbaren* oder durch sie erlernbaren 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 (S. 27)

determiniert ist. In diesem Sinne setzt professionelles Handeln einer Lehrkraft im komplexen Umfeld Schule neben einem Fundus an *verfügbarem* handlungsbezogenem Sach-/Fachwissen und *aktivierbaren* umsetzungsrelevanten Fertigkeiten auch eine adäquate motivationale, volitionale, affektive und auch (berufs-)ethische Grundausrichtung voraus. Das heißt, dass sich eine kompetenzorientierte Hochschullehre, die auf den Aufbau situationsangepasster professioneller Verhaltensdispositionen künftiger Lehrkräfte ausgerichtet ist, nicht allein auf die Vermittlung von *Kenntnissen* kaprizieren kann, dass vielmehr – unabhängig letztlich vom konkreten Veranstaltungsformat – die Entwicklung von prozeduralem, respektive Handlungswissen sowie wert- und haltungsbezogenen Aspekten zum Ziel universitärer Lehr-/Lernprozesse gemacht werden muss. Für die Lehre im MY, die initial primär auf die Vermittlung von Theorie- und Faktenwissen ausgerichtet war (Bachmann, 2018), und die Entwicklung über deklaratives Wissen hinausgehender domänenbezogener Kompetenzaspekte bestenfalls den Studierenden selbst überließ, implizierte das die Notwendigkeit einer konsequenten Abkehr von der Bereitstellung rein wissens- bzw. prüfungszentrierter Lernumgebungen zugunsten der Gestaltung lern(enden)zentrierter Veranstaltungen (Bransford et al., 1999).

3.2 Umgestaltung der Lehrveranstaltungen des Moduls Psychologie im bildungswissenschaftlichen Studium der Bachelor-Lehramtsstudiengänge an der UDE

Seidels (2022) Konzeption *Professioneller Unterrichtswahrnehmung* als ein relevantes kompetenzbezogenes Globalziel der Lehrkräftebildung unterscheidet den Kompetenzaspekt *Noticing* – die Kapazität einer Lehrperson „auf der Basis ihrer professionellen Wissensbestände selektiv ihre Aufmerksamkeit auf solche Ereignisse im Klassenraum zu richten, die für Lehr-Lern-Prozesse besonders relevant sind“ (S. 20) – von *Reasoning* – das Informationsverarbeitungsprozesse beschreibt, wodurch „auf der Basis der professionellen Wissensbestände Interpretationen vollzogen werden“ (S. 20). Seidels Übertragung des psychologischen Expertisebegriffs auf den Lehrerberuf wurde als Basis für die Formulierung der Lernziele auf Modulebene herangezogen:

In diesem Sinne ist ein *erstes* Lernziel der Lehre im MY, dass die zukünftigen Lehrkräfte in die Lage versetzt sind, professionelle Handlungsanforderungen –

in sowohl Routine- als auch Problemsituationen in der schulischen Praxis (Seidel, 2022) – vor dem Hintergrund psychologischer Wissensbestände wahrzunehmen (*Noticing*) und unter Rückgriff auf psychologische Modelle und Theorien zu *beschreiben*, zu *erklären* bzw. zu *prognostizieren* (*Reasoning*).

Ein *zweites* Lernziel ist – über den Kompetenzaspekt der *Professionellen Unterrichtswahrnehmung* hinausgehend – auf den vierten klassischen Aufgabenbereich der empirisch arbeitenden Psychologie ausgerichtet. Vor dem Hintergrund Professioneller Unterrichtswahrnehmung sind die künftigen Lehrkräfte in die Lage versetzt, für komplexe schulische Situationen *zielführende Handlungs- bzw. Interventionsoptionen* unter Rückgriff auf psychologische Modelle, Theorien und methodische Zugänge abzuleiten.

Aus diesen herausfordernden, kompetenzorientierten Lernzielen ergeben sich umfangreiche Anforderungen an die Neugestaltung der Lehre sowohl auf der Ebene des MY insgesamt als auch auf Ebene jeder Lehrveranstaltung.

Zentrale Aspekte der Gestaltung auf der Ebene des Moduls *Psychologie*

Während die Vorlesungen ursprünglich auf ein lineares Abarbeiten des jeweils etablierten *Wissenskanons* der einzelnen psychologischen Teilfächer – wie Lehr-/Lernpsychologie oder Psychodiagnostik – ausgerichtet waren, wird in der neu konzipierten Lehre im MY ein fallbasierter Zugang umgesetzt. Ausgehend von nachvollziehbaren Fallbeispielen aus der schulischen Praxis werden fachliche Schwerpunktsetzungen der Erschließung des jeweiligen Falls systematisch untergeordnet. Das impliziert, dass einerseits die Trennung der vier Vorlesungsteile (vgl. Fußnote 5) aufgehoben und andererseits die traditionelle Abfolge behandelter Inhalte je Teilfach aufgebrochen wurde, um den behandelten Stoff an den behandelten Fall – nicht an die jeweilige Fächerlogik – anzupassen. Mit dieser Ausrichtung der Lehrinhalte auf den sich aus den Fallbeispielen ergebenden Bedarf an theoretischen und methodischen Konzepten und Zugängen geht unabdingbar eine merkliche Reduktion der behandelten Stoffmenge einher – ein Aspekt kompetenzorientierter Gestaltung von Hochschullehre, der von Dozent:innen Umdenken in Bezug auf die eigene Aufgabe und Rolle erfordert (Viebahn, 2004).

Zentrale Aspekte der Gestaltung auf der Ebene der einzelnen Veranstaltung

Auch für die Gestaltung der einzelnen Lehreinheiten ergeben sich Veränderungen, die eine Abkehr vom klassischen Vorlesungskonzept bedeuten.

Raum für ein „Einschwingen“ ins Sitzungsthema. Ausgangspunkt jedes inhaltlichen Schwerpunkts ist die Einführung eines konkreten Fallbeispiels, das zunächst unter Bezugnahme auf Vorwissen der Studierenden analysiert wird. Diese Eingangsreflexion wird durch allgemeine Fragen gelenkt (Was passiert hier? Welche Aspekte der Situation/Ereignisse erscheinen bedeutsam? o. ä.). Die Ergebnisse der Fallreflexion werden im Plenum weitestgehend unkommen-

tiert zusammengetragen. Erst nach dieser Ausrichtung der Teilnehmer:innen auf ein thematisches Gebiet wie beispielsweise *Lernen*, *Gedächtnis* oder *Emotion*, erfolgt die explizite Klärung der Lernziele für die jeweilige Sitzung, deren Formulierung sich an der Bloomischen Lernzieltaxonomie orientieren (Krathwohl, 2002), die den Studierenden bereits vertraut gemacht wurde.

Inputphasen. Die darauffolgende eigentliche Vorlesung ist in Untereinheiten untergliedert, die in der Regel nicht mehr als 20 Minuten Lehrvortrag umfassen. Diese Inputphasen reduzieren sich auf die Vermittlung zentraler inhaltlicher und methodischer Konzepte. Der Detaillierungsgrad ist so gehalten, dass nur ausgeführt wird, was für das Verstehen der Inhalte unabdingbar ist. Der Anspruch auf Vollständigkeit muss zugunsten von Rezipierbarkeit durch Studierende aufgegeben werden. Jede Inputphase wird abgeschlossen durch Reflexionsphasen bzw. Bearbeitung von Vertiefungs- oder Transferaufgaben (Format: Flüstergruppen). Die Ergebnisse dieser regelmäßigen Aktivierungsphasen werden wiederum ins Plenum getragen und in der Gruppe ausführlich besprochen. Klärung von Fragen und Diskussionsbedarf wird so viel Raum gegeben, wie benötigt wird. Erst wenn sich eine Sättigung einstellt, folgt die nächste Inputphase.

Abschluss. Jede Sitzung schließt mit einem Rückbezug auf die Lernziele für die abgelaufene Sitzung ab. Die Studierenden werden in diesem Zusammenhang angehalten, den durch die Vorlesung veränderten Blick auf das eingangs präsentierte Fallbeispiel zu reflektieren, um Zuwachs an Kenntnissen oder Fertigkeiten explizit zu machen – in der Annahme, dass auch motivationale Kompetenzaspekte profitieren (Kusurkar et al., 2011).

Mit dieser Umstellung, die die aktive Teilnahme der Studierenden forciert, ändert sich der Charakter der Vorlesung deutlich. Damit sich auch im Kontext von Massenveranstaltungen Studierende in hinreichend großer Zahl einbringen können und *wollen*, sind entsprechende Freiräume, eine klar kommunizierte, verlässliche Haltung der Lehrenden und die Abkehr vom Konzept der Stoffvermittlung notwendig – wobei vor allem letzteres die Selbstdisziplin Lehrender herausfordert.

3.3 Nutzung digitaler Ressourcen zur Unterstützung der Lehre

Ohne die Einbeziehung digitaler Lernumgebungen und -tools wäre eine solche Umgestaltung der Lehre kaum leistbar. Aktuell sind es vier Bereiche, in denen die Vorlesungen durch digitale Ressourcen unterstützt werden:

Videomaterialien *Fallbasiertes Lehren und Lernen – FaLeLe* (Clausen et al., 2023): Im Rahmen eines Teilprojekts des *ProViel*-Vorhabens der UDE⁸ wurde ein umfangreiches in die Moodle-Umgebung der UDE eingebundenes Reposito-

8 UDE-Projekt *Professionalisierung für Vielfalt (ProViel)* im Rahmen der *Qualitätsoffensive Lehrerbildung von Bund und Ländern*; vgl. <https://www.uni-due.de/proviel/>

rium dokumentierter Unterrichtsvideos aufgebaut. Diese Ressource ist für Lehrende der UDE nutzbar und liefert einen Großteil des Videomaterials, das für die unterrichtsbezogenen Fallbeispiele, die Ausgangspunkt jeder Vorlesungseinheit sind, benötigt wird.

Polling-Tool *PINGO*⁹: Die Abstimmungssoftware PINGO ermöglicht es, auch im Kontext großer Veranstaltungen durch Live-Abfragen Information von und über Teilnehmer:innen zu erhalten. Beispielsweise wird eingangs jeder Veranstaltung im MY eine Übungsaufgabe präsentiert, die die Lernziele der jeweils letzten Sitzung aufgreift. Die Studierenden sind aufgefordert, in Zweiergruppen eine Lösung zu erarbeiten, die sie dann mit einem Lösungsvorschlag der Dozentin abgleichen können. Über PINGO kann abgefragt werden, wie angemessen die Studierenden ihre eigenen Lösungen im Vergleich zum Lösungsvorschlag der Dozentin einschätzen. Über PINGO-Abfragen, die die Studierenden über Handys bedienen können, erhält Gruppe und Dozent:in Einblicke in die leistungsbezogenen Selbsteinschätzungen der Studierenden, auf die direkt reagiert werden kann.

Moodle-basierte *H5P-Umgebungen*¹⁰: Zur Vertiefung von Vorlesungsinhalten oder zum Einüben methodischer Zugänge (z. B. Prozeduren, um psychodiagnostische Testrohre in Standardwerte zu überführen) werden den Studierenden vorlesungsbegleitend im gemeinsamen Moodle-Raum Übungsaufgaben mit Lösungsvorschlägen zur Verfügung gestellt. Dieses formative Element der Überprüfung von Leistungsentwicklung gibt Studierenden die Möglichkeit, eigene Lernprozesse zu überwachen.

Flankierende digitale *Selbstlernressourcen*: Primär, um datenbezogene Kompetenzen zu entwickeln, die zwingende Voraussetzung für die Erschließung der Inhalte der Psychodiagnostik sind (v. a. Deskriptivstatistik), wurden im Kontext des Teilprojekts der Autorinnen im UDE-Vorhaben *DataCampus*¹¹ neben einer R-Shiny-basierten Webapplikation umfangreiche animierte Lernvideos erstellt. Diese Materialien sollen Lehramtsstudierenden einen Einstieg in die Auseinandersetzung mit datenbasierten Entscheidungsprozessen ermöglichen, um die in der Vorlesung vermittelten theoretischen Konzepte, methodischen Zugänge, handlungsleitenden diagnostischen Prinzipien und Prozessmodelle zu illustrieren. Vor dem Hintergrund einer alltagsnahen Fragestellung („Lohnt es sich, selbst zu kochen?“) werden den Studierenden betont niedrigschwellig grundlegende methodische Kompetenzen vermittelt. Darüber hinaus werden komplexe inhaltliche Schwerpunkte in Form zusätzlicher erklärender Vorlesungsvideos aufbereitet und im Moodle-Kurs bereitgestellt.

9 *PINGO – Peer Instruction for Very Large Groups*; <https://imt.uni-paderborn.de/nachricht/paderborner-abstimmsoftware-pingo-ist-nun-dauerangebot-fuer-hochschulen/>

10 Vgl. <https://h5p.org/>

11 Vgl. <https://www.uni-due.de/ub/datacampus/>

4 Evaluation der Maßnahmen

Um die Wirkung der Maßnahmen bewerten zu können, wurde zu mehreren Messzeitpunkten – jeweils gegen Ende eines laufenden Semesters – schriftliches Feedback der Studierenden eingeholt. Im Sommersemester 2022 – vor Beginn der Umsetzung der Lehrinnovation – fand die erste PITCH-Erhebung statt (Baseline, T0). Im folgenden Wintersemester 2022/23 wurde die erste Evaluation nach einer partiellen Umgestaltung der Lehre durchgeführt (T1).¹² Neben dem an der UDE etablierten Instrument für die regelmäßige Lehrevaluation¹³, werden psychologische Standardskalen eingesetzt, um Einblicke in lern- und leistungsrelevante, v. a. *motivationale* Faktoren zu gewinnen.¹⁴

Beispielsweise ist das Konstrukt *Situational Interest* ausgerichtet auf situationsbezogene Faktoren, die das Interesse Studierender im akademischen Kontext beeinflussen. Skalen, die *Triggered Interest* – ein Aspekt von *Situational Interest* – operationalisieren (Linnenbrink-Garcia et al., 2010), messen, inwieweit akademische Inhalte das Potenzial haben, das Interesse Studierender zu wecken.¹⁵ Im Prä-Post-Vergleich zeigte sich bei den Lehramtsstudierenden für *Triggered Interest* ein signifikanter Unterschied durch die umgestaltete Lehre [$t(437) = -2,10, p = ,018$]. Dies ist insofern bedeutsam, als die im Abschnitt 1.2 berichteten Vorerhebungen aus dem Jahr 2019 signifikante Unterschiede in Bezug auf *Langeweile* als hemmende negative akademische Emotion im direkten Vergleich von bildungswissenschaftlichem und Fächerstudium zu Ungunsten der Bildungswissenschaften belegten [$t(342) = 1,50, p < ,001$]. Eine entsprechende Tendenz zeigt sich auch in Bezug auf die Subskala *Inhalte* aus dem Lehrevaluationsfragebogen, die mit drei Items interessens- und aufmerksamsbezogene Aspekte abfragt.¹⁶ Auch hier konnte eine signifikante Verbesserung gefunden werden [$t(248,12) = -3,79, p < ,001$].

12 Noch ausstehend ist eine Follow-Up-Evaluation im Wintersemester 2023/24, die insofern besonders relevant ist, als die wesentlichen Veränderungen auf Ebene der Lehre im *Modul* – v. a. die Zusammenlegung der Teilvorlesungen und die konsequente Integration der Inhalte der Teilgebiete – erst mit dem Wintersemester 2023/24 umgesetzt werden konnten.

13 Für die klassische Lehrevaluation wurde das vom ZHQE der UDE bereitgestellte *Feedbackinstrument zur studentischen Lehrveranstaltungsbewertung* verwendet; vgl. <https://www.uni-due.de/zhqe/lehrevaluation.php/>

14 Vergleiche der Leistung der Studierendenkohorten erfolgten bisher nicht, da psychometrische Äquivalenz der Modulabschlussprüfungen nicht gewährleistet werden kann.

15 Items „Die Lehrperson ist in der Lage, mein Interesse zu wecken“, „In der Vorlesung tut die Lehrperson Dinge, die meine Aufmerksamkeit wecken“, „Die Vorlesung ist oft unterhaltsam“ und „Die Vorlesung ist so spannend, dass es einfach ist, aufmerksam zu sein“.

16 Items „Die Inhalte werden verständlich vermittelt und aufbereitet“, „Dem Tempo bei der Vermittlung bzw. Erarbeitung kann ich gut folgen“ und „Es werden Bezüge zum Berufsfeld und Anwendungsmöglichkeiten aufgezeigt“.

Studierende ($n = 53$), die die Veranstaltungen im MY nach nicht bestandener Erstprüfung zum wiederholten Mal besuchten, wurden zusätzlich um einen Vergleich von ursprünglicher und neugestalteter Lehre gebeten.¹⁷ Neben quantitativen Abfrageformaten standen auch Freitextfelder für eine Rückmeldung von positiven und negativen Aspekten der umgestellten Lehre zur Verfügung. Die Qualität der Veranstaltungen in geänderter Form wurde insgesamt positiv eingeschätzt [Rating: 5-stufige Likert-Skalen; $M = 4,42$; $SD = 0,77$]. 86,8 Prozent dieser Studierenden gaben an, die Veranstaltungen im neuen Format als *viel besser* oder *eher besser* zu beurteilen. Auch die verbalen Rückmeldungen der Studierenden unterfüttern diese Zahlen. Rückmeldungen zu positiven Veränderungen beziehen sich primär auf die geänderte *Vermittlung der Inhalte*:

Sehr gut hat mir gefallen, dass die Inhalte so reduziert wurden, dass man gut mit dem Tempo klarkam und tief in die Materie eindringen konnte. Die Erklärungen orientierten sich stets an den Studierenden und wurden so auch anhand von Beispielen verständlich.¹⁸

Die positiven Anmerkungen bezogen sich insgesamt auf v. a. das *Tempo* der Inhaltsvermittlung (49 Nennungen), die Untermauerung mit *berufsbezogenen* Beispielen (49 Nennungen), die *Aufbereitung* der Inhalte (40 Nennungen) und den Raum für Fragen und *Interaktion* (38 Nennungen). Auch die Verbesserungswünsche der Studierenden, die die Veranstaltungen in alter und neuer Form kannten, deuten darauf hin, dass Ausrichtung und Schwerpunkte der Veränderung der Lehrveranstaltungen zielführend sind. Die Studierenden nannten am häufigsten den Wunsch nach weiteren *Fallbeispielen*, die Praxisbezug herstellen (19 Nennungen). Auch der Wunsch nach mehr *Übungsmöglichkeiten* bestand (15 Nennungen). Diese ersten Evaluationsergebnisse geben Hinweise darauf, dass sich die Umgestaltung der Lehre im MY generell positiv auf das Erleben der Studierenden auswirkt. Die Neuausrichtung scheint den Bedürfnissen der Lehramtsstudierenden entgegenzukommen.

5 Fazit

Der Rahmen des Projekts *Prüfungen innovieren, Transfer schaffen, Chancengerechtigkeit fördern – PITCH* bot die Möglichkeit, mit Schwachstellen der curricular verankerten Lehre im bildungswissenschaftlichen Studium der lehr-

¹⁷ Erhoben durch ein Item: „*Falls Sie diese Veranstaltung in einem vergangenen Semester schon einmal besucht haben: Ich empfinde die Veranstaltung dieses Semester als...*“.

¹⁸ Originalzitat einer/s Studierenden aus einer Freitextabfrage im Zusammenhang durchgeführten Evaluation (vgl. *Feedbackinstrument zur studentischen Lehrveranstaltungs-bewertung* der UDE; <https://www.uni-due.de/zhqe/lehrevaluation.php/>).

amtsbezogenen Bachelorstudiengänge der UDE systematisch im Sinne einer Entwicklung von Studienqualität umzugehen. Erste Evaluationsergebnisse belegen, dass der eingeschlagene Weg, die Lehre im Modul *Psychologie* zu innovieren, grundsätzlich in eine richtige Richtung weist. Nun gilt es einerseits das neugestaltete Lehrangebot zu konsolidieren und andererseits die sich notwendig ergebenden Schritte in *Bezug auf die Umgestaltung der Modulabschlussprüfungen und den Ausbau formativer Prüfungselemente* umzusetzen.

Literatur

- Bachmann, H. (2018). Hochschullehre neu definiert – shift from teaching to learning. In H. Bachmann (Hrsg.), *Kompetenzorientierte Hochschullehre: Die Notwendigkeit von Kohärenz zwischen Lernzielen, Prüfungsformen und Lehr-Lern-Methoden* (3. Aufl.) (S. 14–33). Bern: hep, der Bildungsverlag.
- Biggs, J. B./Tang, C. S. (2011). *Teaching for quality learning at university: What the student does* (4. Aufl.). Buckingham, England: Open University Press.
- Bransford, J. D./Brown, A. L./Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington D.C.: National Academy Press.
- Clausen, M./Jahn, S./Diebig, K. (2023). Portal Fallbasiertes Lehren und Lernen—Schulmaterial und Lernmodule. [dataset]. Universitätsinterne Veröffentlichung auf der E-Learning-Plattform Moodle der Universität Duisburg-Essen.
- Eccles, J. S./Wigfield, A. (2002). Motivational beliefs, values, and goals. In: *Annual Review of Psychology*, 53(1), 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>
- Furnham, A. (1988). *Lay Theories*. Pergamon Press. <https://doi.org/10.1016/C2009-0-14697-4>
- Gür-Seker, D./Hintze, P./Tasche, B./Wefelnberg, M./Velibeyoglu, N. (2023). Künstliche Intelligenz in Studium und Lehre – Empfehlungen zum Umgang an der UDE. Internes Positionspapier der Universität Duisburg-Essen.
- KMK (2004). Standards für die Lehrerbildung: Bildungswissenschaften (Beschluss der Kultusministerkonferenz vom 16.12.2004 i. D. F. vom 16.05.2019). http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf
- Krathwohl, D. R. (2002). A Revision of Bloom’s Taxonomy: An Overview. In: *Theory Into Practice*, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2
- Kusurkar, R. A./Ten Cate, T. J./Van Asperen, M./Croiset, G. (2011). Motivation as an independent and a dependent variable in medical education: A review of the literature. In: *Medical Teacher*, 33(5), e242–e262. <https://doi.org/10.3109/0142159X.2011.558539>
- Linnenbrink-Garcia, L./Durik, A. M./Conley, A. M./Barron, K. E./Tauer, J. M./Karabenick, S. A./Harackiewicz, J. M. (2010). Measuring Situational Interest in Academic Domains. In: *Educational and Psychological Measurement*, 70(4), 647–671. <https://doi.org/10.1177/0013164409355699>
- Monitor Lehrerbildung. (2023). Hochschulvergleich Nordrhein-Westfalen 2022. https://www.monitor-lehrerbildung.de/wp-content/uploads/2023/03/Hochschulvergleich_Nordrhein-Westfalen_2022.xlsx

- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. In: *Educational Psychology Review*, 18(4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9>
- Pekrun, R. (2018). Emotion, Lernen und Leistung. In M. Huber/S. Krause (Hrsg.), *Bildung und Emotion* (S. 215–231). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-18589-3_12
- Pekrun, R./Goetz, T./Titz, W./Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement. In: *Educational Psychologist*, 37(2), 91–105. https://doi.org/10.1207/S15326985EP3702_4
- Perry, R. P./Hladkyj, S./Pekrun, R./Pelletier, S. T. (2001). Academic control and action control in the achievement of college students: A longitudinal field study. In: *Journal of Educational Psychology*, 93(4), 776–789. <https://doi.org/10.1037/0022-0663.93.4.776>
- Reinmann, G. (2022). Prüfung oder Assessment an Hochschulen? Thesen für einen Wandel der Prüfungskultur. In J. Gerik/A. Sommer/G. Zimmermann (Hrsg.), *Kompetent Prüfungen gestalten* (S. 22–36). Münster, New York: Waxmann.
- Seidel, T. (2022). Professionelle Unterrichtswahrnehmung als Teil von Expertise im Lehrberuf. Weiterentwicklungsperspektiven für die videobasierte Lehrerforschung. In R. Junker/V. Zucker/M. Oellers/T. Rauterberg/S. Konjer/N. Meschede/M. Holodynski (Hrsg.), *Lehren und Forschen mit Videos in der Lehrkräftebildung* (S. 17–35). Münster, New York: Waxmann. <https://doi.org/10.31244/9783830995111>
- Spoden, C./Frey, A. (Hrsg.). (2021). *Psychometrisch fundierte E-Klausuren für die Hochschule*. Lengrich: Pabst Science Publishers.
- 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, Digitalisierung* (1. Aufl.). Münster, New York: Waxmann. <https://doi.org/10.31244/9783830992462>
- Viebahn, P. (2004). *Hochschullehrerpsychologie: Theorie- und empiriebasierte Praxisanregungen für die Hochschullehre*. Bielefeld: UVW, UniversitätsVerlagWebler.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – Eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–32). Weinheim Basel: Beltz.
- Wissenschaftsrat (2008). *Empfehlungen zur Qualitätsverbesserung von Lehre und Studium* (Drs. 8639-08). https://www.wissenschaftsrat.de/download/archiv/8639-08.pdf?__blob=publicationFile&v=1
- Wissenschaftsrat (2022). *Empfehlungen für eine zukunftsfähige Ausgestaltung von Studium und Lehre* (Drs. 9699-22). https://www.wissenschaftsrat.de/download/2022/9699-22.pdf?__blob=publicationFile&v=17

Kompetenzorientiert digital prüfen in der Sport- und Bewegungswissenschaft

Thomas Mühlbauer und Johanna Lambrich

Ausgehend von der herkömmlichen Prüfungsdurchführung mittels gedruckter Klausurbögen besteht für die Lehreinheit Sport- und Bewegungswissenschaft (vgl. nachstehendes Kurzprofil) die Herausforderung darin, kompetenzorientierte digitale Prüfungsformate auf Basis des „Threshold-Concepts“-Ansatzes¹ (Meyer/Land, 2003) zu entwickeln, zu erproben und auszuwerten. Im Konkreten geht es darum, sog. „Schwellen“ zu identifizieren, d. h. komplexe, schwierig zu verstehende Kernkonzepte (theoretische Konstrukte), deren Verständnis aber für den erfolgreichen Studienverlauf und die spätere Berufstätigkeit (z. B. als Sportlehrer:in) sehr bedeutsam ist (Wiemer/Kennweg, 2021). Nach der Identifikation sollen diese Kernkonzepte verstehend erarbeitet und in digitalen Prüfungsformaten unter Nutzung verschiedener Kompetenzstufen (1. Stufe: Wissen/Verstehen, 2. Stufe: Anwenden/Analysieren, 3. Stufe: Synthetisieren/Evaluieren) aufbereitet werden. Abschließend erfolgt mittels Assessment Analytics eine Auswertung der Prüfungsbearbeitung.

Kurzprofil der Lehreinheit:

- Universität Duisburg-Essen, Fakultät für Bildungswissenschaften, Institut für Sport- und Bewegungswissenschaften, Lehreinheit Sport- und Bewegungswissenschaft
- Lehrende: 4 Professuren, ca. 20 Mitarbeitende
- Bachelor- (Bachelor of Arts) und Masterstudiengänge (Master of Education) für die Lehramtsoptionen Berufskolleg, Grundschule, Gymnasium und Gesamtschule, Haupt-, Real-, Sekundar- und Gesamtschule, Sonderpädagogik
- ca. 680 Bachelor- und ca. 230 Masterstudierende

¹ Meint die anfangs mühevolle Auseinandersetzung mit einem zunächst unzugänglichen Gegenstand (z. B. Theorie, Modell, Konzept) und dem Durchschreiten „konzeptioneller Tore“ oder „Portale“, um im Ergebnis, eine neue Art des Verstehens, der Interpretation oder der Betrachtung zu erreichen.

1 Vorgehen: Thresholds identifizieren und Lehr-, Lern- und Selbstüberprüfungsangebote gestalten

Das Modul C „Grundlagen der naturwissenschaftlichen Teildisziplinen der Sportwissenschaft“ befindet sich in der Studieneingangsphase (2. Fachsemester) des Bachelorstudiums für alle Lehramtsoptionen und wird von ca. 140 bis 160 Studierenden besucht. Das Modul C besteht aus den Vorlesungen „C1: Grundlagen der Anatomie und Physiologie“ und „C2: Grundlagen der Bewegungs- und Trainingswissenschaft“ sowie dem Seminar „C3: Angewandte Bewegungs- und Trainingswissenschaft“. Als Modulprüfung ist ein summatives Assessment in Form einer Klausur vorgesehen.

In einem ersten Schritt wurde den Studierenden in der Auftaktveranstaltung der beiden Vorlesungen der „Threshold-Concepts“-Ansatz vorgestellt (Abbildung 1). Unmittelbar danach wurde ein Etherpad im Moodle-Kursraum erstellt, in dem die Studierenden individuell wahrgenommene „Schwellen“ anonym über den 15-wöchigen Veranstaltungsverlauf eintragen konnten. Zwar haben sich in Relation zur Gesamtkohorte nur ca. 5–10 Prozent der Studierenden am Etherpad beteiligt. Jedoch wurden pro Vorlesung drei bzw. vier voneinander unabhängige Schwellen genannt. Wahrgenommene Schwellen – wie zum Beispiel der Dehnungs-Verkürzungs-Zyklus² – wurden am Ende der Vorlesungsreihe in einem prüfungsvorbereitenden Repetitorium mittels Dozierenden-Studierenden-Gespräch thematisiert. Im konkreten Fall wurde mittels zusätzlichen Anschauungsmaterials und Praxisbeispielen versucht, ein besseres Verständnis zu erreichen.

Zudem erfolgte im Veranstaltungsverlauf von 15 Wochen eine Sammlung von zusätzlichen Lehrinhalten durch die Lehrenden im Modul, die aus ihrer Sicht mögliche „Schwellen“ darstellen, deren Verständnis aber für den weiteren Verlauf des Studiums und die anschließende Tätigkeit als Sportlehrer:in wichtig sind. Die identifizierten Inhalte wurden daraufhin in Modulkonferenzen besprochen sowie mit Hochschuldidaktikerinnen und -didaktikern unter Verwendung der „Decoding the discipline“-Methode³ (Middendorf/Shopkow, 2017) den Dozierenden transparent gemacht (Abbildung 2).

- 2 Der Dehnungs-Verkürzungs-Zyklus (DVZ) meint eine aktive Muskeldehnung (exzentrische Kontraktion), gefolgt von einer sofortigen Verkürzung (konzentrische Kontraktion) desselben Muskels und hat das Ziel, durch die Voraktivierung eine stärkere Muskelspannung und damit höhere Kraft zu generieren (Gollhofer/Taube/Gruber, 2009). Für die Tätigkeit als Sportlehrer:in lässt sich daraus ableiten, dass z. B. größere Sprunghöhen erreicht werden, wenn die Bewegungsausführung mit (Hockstretksprung) anstatt ohne (Stretksprung) Ausholbewegung erfolgt.
- 3 Meint einen festgelegten Prozess bestehend aus einem etabliertem mehrschrittigem Verfahren, mit dem zunächst Lernherausforderungen (sog. „bottlenecks“) identifiziert und nachfolgend didaktische Strategien entwickelt werden.

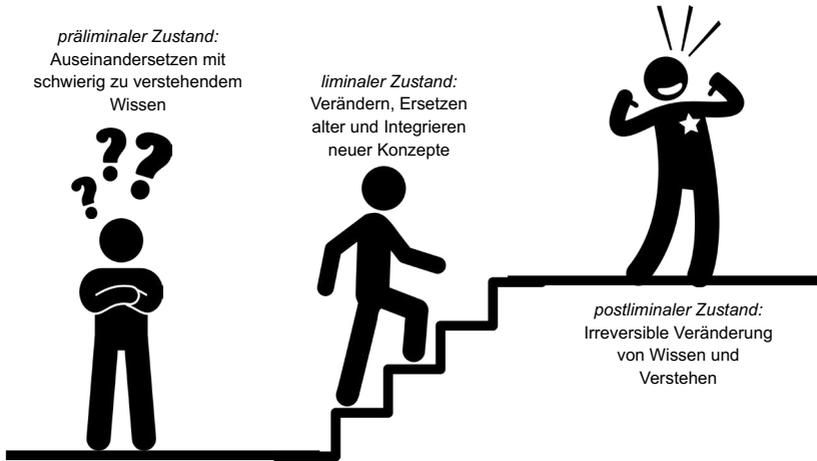


Abbildung 1: Darstellung des „Threshold-Concepts“-Ansatzes (mod. nach Meyer/Land, 2003)

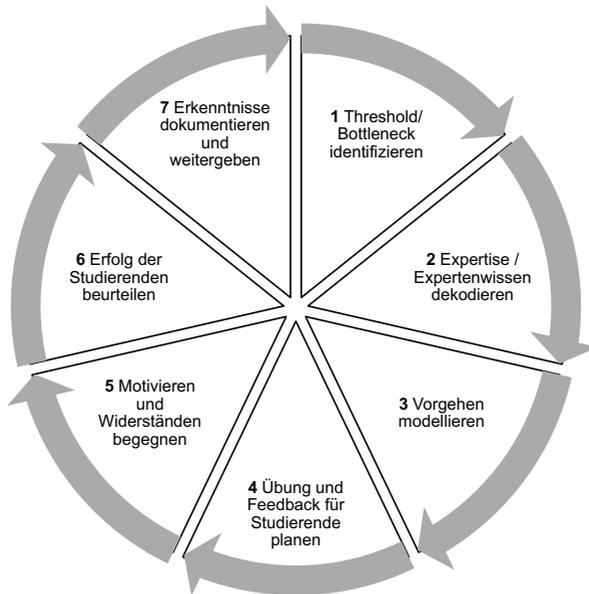
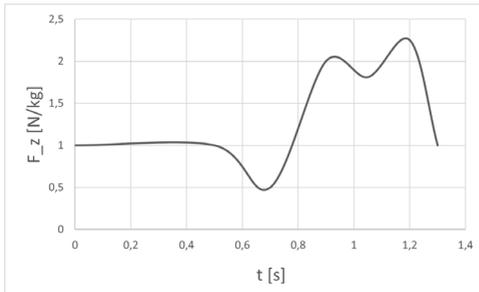


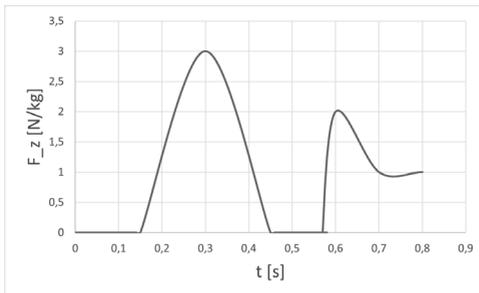
Abbildung 2: Darstellung der „Decoding the Disciplines“-Methode (mod. nach Middendorf/Shopkow, 2017)

**Die folgende Grafik zeigt die Kraft-Zeitverläufe von drei Sprungarten.
Ordnen Sie jedem Verlauf die entsprechende Bezeichnung zu! Nutzen Sie hierfür
die Drag-and-Drop-Funktion!**



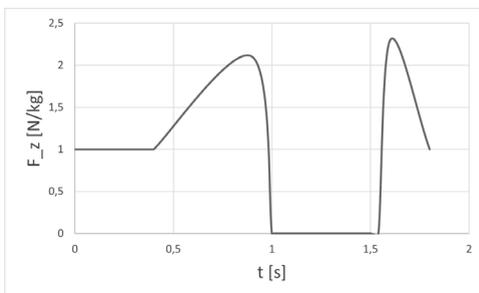
Standing Long
Jump

Squat Jump



Drop Jump

Counter-
movement Jump



Lunge Jump

Hop Jump

Abbildung 3: Darstellung einer exemplarischen Prüfungsfrage (maximal 3 Punkte) zum Dehnungs-Verkürzungs-Zyklus (Quelle: Eigene Abbildung)

Im zweiten Schritt wurden die „Schwellen“ mittels **HTML5-Package (H5P)** in Form von Lehr-, Lern- und Selbstüberprüfungsangeboten ($N=49$, Stand September 2023) didaktisch aufbereitet und in die Lehrveranstaltungen integriert. Die Verwendung von H5P-Elementen ermöglicht es, verschiedene Übungsvarianten (z. B. Lückentext, Kreuzworträtsel, Zuordnungsaufgaben) in den bestehenden Moodle-Kurs der Veranstaltung zu integrieren, ohne dass die Studierenden eine weitere Plattform nutzen müssen. Die Lehr-, Lern- und Selbstüberprüfungsangebote wurden den Studierenden zur Vorlesungsnachbereitung und Prüfungsvorbereitung zur Verfügung gestellt und von ihnen 12.098-mal aufgerufen. Darüber hinaus wurde kontinuierlich eine Moodle-basierte (<https://moodle.exam.uni-due.de/>) Sammlung von Prüfungsfragen ($N=283$, Stand Oktober 2023) aufgebaut. Hierbei wurden verschiedene Kompetenzstufen (Bloom, 1976) berücksichtigt sowie unterschiedliche Prüfungstypen (Einfach- oder Mehrfach-Entscheidung), -formate (textlich, tabellarisch, grafisch) und -varianten (z. B. Zuordnung oder Lückentext per Drag-and-Drop-Funktion, Auswahl per Drop-Down-Funktion) verwendet (Abbildung 3).

Im dritten Schritt erfolgte die Einrichtung und Erprobung (zunächst mit Masterstudierenden, um Erfahrungen zwischen digitaler versus nicht-digitaler Durchführung zu erhalten) einer Moodle-basierten Probeklausur. Anhand von Rückmeldungen (z. B. Fragen präziser zu formulieren, Antwortbereiche eindeutiger auszuweisen) wurde die Klausur überarbeitet und konnte anschließend von den aktuellen Bachelorstudierenden mit dem Ziel der Gewöhnung ausprobiert werden. Die dabei geäußerten Rückmeldungen (z. B. Präzisierung verschiedener Arten der Fragestellung, Veränderung der grafischen Darstellung) wurden für die finale Erstellung der Klausur berücksichtigt und beispielsweise mittels farblicher Hervorhebungen (z. B. von verschiedenen Antwortkategorien in Lückentextaufgaben) oder der Generierung von Antwortkästen realisiert.

2 Zwischenfazit und Ausblick

Die Auswertung der Prüfungsergebnisse ergab, dass sowohl auf der ersten Kompetenzstufe „Wissen/Verstehen“ als auch auf der zweiten Stufe „Anwenden/Analysieren“ die Mehrzahl der Studierenden mehr als die Hälfte aller Punkte erreichten (Abbildung 4). Lediglich acht Prozent der Studierenden haben weniger als die Hälfte der Punkte erreicht. In einem nächsten Schritt wurden einerseits Fragen, die von 90 bis 100 Prozent der Studierenden richtig beantwortet wurden, als „zu leicht“ eingestuft und für zukünftige Prüfungen eliminiert. Andererseits wurden Fragen, die nur von 30 Prozent der Studierenden korrekt beurteilt wurden, als „zu schwer“ klassifiziert und für kommende Prüfungen überarbeitet. Dies erfolgte durch die Erstellung vereinfachter grafischer Darstellungen sowie die Erläuterung von Fachbegriffen.

Bezogen auf die in Abbildung 3 dargestellte Frage haben drei Prozent der Studierenden drei Punkte, acht Prozent zwei Punkte, 43 Prozent einen Punkt und 46 Prozent keinen Punkt erhalten. Demzufolge wurde der von den Dozierenden als „Schwelle“ deklarierte Inhalt (Dehnungs-Verkürzungs-Zyklus) und der damit angestrebte *postliminale Zustand* (d. h. irreversible Veränderung von Wissen und Verstehen) nur durch relativ wenige Personen erfolgreich absolviert. Neben einer Überarbeitung dieser Prüfungsaufgabe resultiert als weitere Konsequenz das Bemühen um eine verstärkte verstehende Erarbeitung des dazugehörigen Wissens in den entsprechenden Veranstaltungen des Moduls C. Spezifischer ausgedrückt, muss ein stärkerer Fokus auf die Absolvierung des *präliminalen* (d. h. Auseinandersetzen mit schwierig zu verstehendem Wissen) und *liminalen* (d. h. Verändern, Ersetzen alter und Integrieren neuer Konzepte) *Zustands* gelegt werden. Dies könnte dadurch erreicht werden, dass zum Beispiel nicht nur Erkenntnisse zum Dehnungs-Verkürzungs-Zyklus erarbeitet, sondern auch praktische Eigenerfahrungen beim Anwenden von Bewegungen mit (Hockstrecksprung) und ohne (Strecksprung) Ausholbewegung sowie der Analyse resultierender Sprunghöhenunterschiede gesammelt werden.

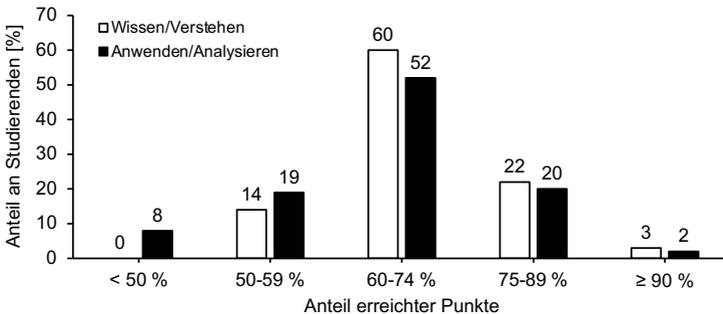


Abbildung 4: Darstellung des Anteils von Studierenden pro erreichter Punktekatgorie für die Kompetenzstufen „Wissen/Verstehen“ sowie „Anwenden/Analysieren“ (Quelle: Eigene Abbildung)

Darüber hinaus stellt die Entwicklung von Prüfungsfragen für die dritte Kompetenzstufe „Synthetisieren/Evaluieren“ eine Herausforderung für den weiteren Projektverlauf dar. Im Konkreten geht es darum, komplexe Fragestellungen mit Entscheidungsbäumen zu kreieren, die einen Gegenstand (z. B. Return to Sport) aus mehreren Perspektiven (physiologisch, psychologisch, sozial, ökonomisch etc.) betrachten und einen Ergebniskorridor mit unterschiedlichen Lösungen zulassen. Zudem wird geprüft, ob die Entwicklung, Erprobung und Auswertung kompetenzorientierter digitaler Prüfungsformate auf Basis des „Threshold-Concepts“-Ansatzes auch auf andere Module (z. B. Modul A „Einführung in die

Sportwissenschaften“ und Modul B „Grundlagen der sozial- und geisteswissenschaftlichen Teilgebiete der Sportwissenschaft“) übertragen werden kann.

Literatur

- Bloom, B. S. (1976). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim: Beltz Verlag.
- Gollhofer, A./Taube, W./Gruber, M. (2009). Biomechanik der Skelettmuskulatur. In: A. Gollhofer/E. Müller (Hrsg.), *Handbuch Sportbiomechanik* (S. 28–51). Schorndorf: Karl Hofmann Verlag.
- Meyer, J. H. F./Land, R. (2003). Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising. In C. Rust (Hrsg.), *Improving student learning: Theory and practice ten years on* (S. 412–424). Oxford: Oxford University Press.
- Middendorf, J./Shopkow, L. (2017). *Overcoming student learning bottlenecks. Decode the critical thinking of your discipline*. New York: Routledge.
- Wiemer, M./Kenneweg, A. C. (2021). Threshold Concepts: Übergänge zu disziplinären Denkweisen und transformative Lernprozesse in der Fachlehre verstehen und begleiten. In B. Berendt/A. Fleischmann/G. Salmhofer/N. Schaper/B. Szczyrba/M. Wiemer/J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre*, 100 (S. 43–66). Berlin: DUZ- Verlags- und Medienhaus. <https://doi.org/10.36197/DUZOPEN.029>

Decoding Deutsch als Zweitsprache: Was macht das Fach DaZ schwer (zu prüfen)?

Nguyen Minh Salzman-Hoang und Tobias Schroedler

Dieser Beitrag diskutiert Erkenntnisse und Erfahrungen aus dem Projekt „Prüfungen innovieren, Transfer schaffen, Chancengerechtigkeit fördern (PITCH)“, die am Institut für Deutsch als Zweit- und Fremdsprache an der Universität Duisburg-Essen (UDE) gemacht wurden. Im Zentrum steht das Ziel, die Prüfung im Modul „Grundlagenwissen Deutsch als Zweitsprache“ (im Folgenden: „DaZ-Modul“) inhaltlich und technisch weiterzuentwickeln. Dazu wird im Folgenden nach einer einleitenden Vorstellung des Moduls (1) die eigene Disziplin in den Blick genommen, um zentrale Fachspezifika und Hürden für studentische Lernprozesse herauszuarbeiten (2). Der letzte Teil (3) des Beitrags thematisiert Herausforderungen der Prüfungskonzeption und schließt mit einem Fazit der Projektarbeit mit Implikationen für die Zukunft.

1 Das DaZ-Modul an der Universität Duisburg-Essen

Deutsch als Zweitsprache (DaZ) ist ein vergleichsweise junges, interdisziplinäres Forschungsfeld, das bei sich verändernden migrationsgesellschaftlichen Rahmenbedingungen einem hohen Aktualisierungsbedarf unterliegt und sich seit seinen frühen Anfängen in den 1970er Jahren (Barkowski, 2003) im stetigen Wandel befindet. Das DaZ-Modul an der Universität Duisburg-Essen, um das es in diesem Beitrag geht, ist als Resultat eines bildungspolitischen Wandels entstanden: Mit der am 12.05.2009 beschlossenen Reform der Lehramtsausbildung wurde gesetzlich verankert, dass alle Lehramtsstudierende in Nordrhein-Westfalen Grundlagenwissen im Bereich „Deutsch für Schülerinnen und Schüler mit Zuwanderungsgeschichte“ (DSSZ) bzw. „Deutsch als Zweitsprache“ (DaZ) im Umfang von mindestens sechs Leistungspunkten erwerben (Ministerium des Innern des Landes NRW, 2009). Mit Blick auf jahrzehntelang gut dokumentierte Bildungsungleichheiten zwischen Schüler:innen mit Migrationsgeschichte und solchen ohne ist dementsprechend die Professionalisierung aller angehenden Lehrkräfte im Umgang mit Mehrsprachigkeit und DaZ im Klassenraum ein Ziel des DaZ-Moduls. Was gelernt (und daher auch geprüft) werden soll, ist somit auch mit übergeordneten gesellschaftlichen Anforderungen an Lehrkräftehandeln verbunden.

Das Modul wird an der UDE¹ für Lehramtsstudierende der Schulformen a) Grundschule, b) Haupt-, Real- und Sekundarschulen (HRSGe), c) Berufskolleg sowie d) Gymnasien und Gesamtschulen angeboten. Alle Studierende dieser Schulformen belegen es zwischen dem zweiten und vierten Bachelorsemester in Form einer Vorlesung, eines Seminars und einer Abschlussklausur (pro Semester 350–700 Studierende). Daraus ergeben sich zwei Fragenkomplexe, denen dieser Beitrag nachgeht:

- „*Decoding DaZ*“: Was sollen Studierende auf dem Weg zur angestrebten Professionalisierung vorrangig lernen? Welche disziplinspezifischen Denkweisen sind herausforderungsvoll zu erwerben?
- Was bedeutet das für die (didaktische) Konzeption einer Klausur, deren Gestaltung auch organisatorischen Sachzwängen unterworfen ist?

2 Sprache als Gegenstand von Analyse und Reflexion

Für die inhaltlich-konzeptionelle Arbeit im PITCH-Projekt sind primär zwei Ansätze aus der Hochschuldidaktik impulsgebend. Zum einen ist das der Ansatz der *Threshold Concepts* von Meyer/Land (2003), die in Anschluss an Perkins‘ (1999) konstruktivistischen Überlegungen zu *troublesome knowledge* das studentische Lernen als mühsam, integrativ und irreversibel theorisieren und im Raum von Liminalität, d. h. eines schwellenartigen Übergangszustands, verorten (Meyer/Land, 2005; Wiemer/Kenneweg, 2021). Zum anderen ist dies *Decoding the Discipline* (kurz: *Decoding*) von Middendorf und Pace (2004), welche darin sowohl eine pädagogische Theorie um sogenannte *bottlenecks* (Stellen, an denen Studierende in ihrem Lernprozess stagnieren) als auch eine methodologische Prozessbeschreibung zur Verfügung stellen (Middendorf et al., 2023, S. 181).

Ein Ausgangspunkt für das *Decoding* ist, „das nicht-explizite Denken und Handeln von Expert:innen der Disziplin sichtbar zu machen“ (Riegler, 2020, S. 357). Für den kompetenten Umgang von Lehrkräften mit Deutsch als Zweitsprache und Mehrsprachigkeit sind linguistische Inhalte zentral, doch sind diese mit didaktischen und bildungswissenschaftlichen Perspektiven verschränkt – Interdisziplinarität ist ein Kern des Fachs. Welche Fachinhalte dafür benötigt werden, wie diese zu priorisieren sind und welche Lern- und Übungsformen dafür angemessen und effektiv sind, wird fachintern und -übergreifend diskutiert und ist Gegenstand diverser hochschuldidaktischer und professionstheoretischer Beiträge (z. B. Busse, 2020; Schroedler/Grommes, 2019). Demnach lässt sich konstatieren, dass es den *einen* als etabliert angesehenen Wissenskanon innerhalb der Fachcommunity nicht gibt, und das Fach DaZ bzw. das DaZ-Modul an

1 Die inhaltliche und curriculare Ausgestaltung des Moduls obliegt den lehrkräftebildenden Universitäten und wird unterschiedlich gehandhabt (s. Putjata et al., 2016, für eine vergleichende Übersicht).

der Universität Duisburg-Essen vielmehr standortspezifisch aus einer langen Tradition erwachsen ist.² Folgende Überlegungen beziehen sich daher explizit auf Aspekte des hiesigen DaZ-Modul-Curriculums.

Zu Beginn des Projekts wurden Interviews mit erfahrenen Lehrenden durchgeführt, um die Lehre auf *threshold concepts* und *bottlenecks* hin zu untersuchen: Dabei sollten mentale Prozesse, die elementar für das DaZ-relevante professionelle Handeln von Lehrkräften sind, deren Erwerb jedoch mit Hürden verbunden ist, identifiziert werden. Im Verlauf der Gespräche rückte Sprache als Gegenstand von Analyse und Reflexion in den Fokus. Gemeint ist hier die Schwierigkeit von Studierenden, mit Sprache in all ihren Facetten analytisch und reflexiv zu operieren. Zwar würde es ihnen gelingen, unter Anwendung von sprachlicher Intuition, Normabweichungen (vermeintliche ‚Fehler‘) in Äußerungen und Schriftprodukten von Schüler:innen zu erkennen. Eine solche Defizitzentrierung und linguistisch unfundierte Haltung zu sprachlicher Variation würde im DaZ-Kontext jedoch als ‚laienhaftes‘ Handeln gelten und unterscheidet sich stark von einer systematisch sprachförderlichen Unterrichtsgestaltung, wie im Folgenden an zwei Themen in Orientierung an den Ansätzen *Decoding* und *Threshold Concepts* diskutiert wird.

Besonders deutlich wird die erwähnte Schwierigkeit im Lerngegenstand der Sprachdiagnostik, der die Anwendung und Ergebnisinterpretation von Verfahren zur Feststellung des Sprachstands ein- und mehrsprachiger Schüler:innen umfasst. Die dahinterliegenden Grundlagen über Sprache lassen sich als „mühevoll“ (*troublesome knowledge*, s. o.) charakterisieren: Zunächst ist das linguistische Wissen zur analytischen Differenzierung der Ebenen von Sprache für Studierende ohne philologisches Studienfach nicht abgesichert oder gänzlich neu. So können sich Ergebnisse von syntaktischen Analysen, wie sie für das Verfahren der Profilanalyse (Grießhaber, 2012) notwendig sind, alltagssprachlichen Verständnissen von sprachlicher Komplexität widersetzen und kontraintuitiv sein (*conceptually difficult knowledge*, Perkins, 1999), wenn z. B. aufgrund von orthografischen oder morphologischen Fehlern die Progression im Spracherwerb nicht (an-)erkannt wird. Die professionelle (unterrichts-)alltagspraktische Einbettung geht zudem über eine bloße regelmäßige Durchführung von Testterminen hinaus. Tatsächlich wird sprachförderlicher Unterricht von einer Lehrkraft auf mehreren Ebenen gleichzeitig konzipiert, wie der sogenannte Scaffolding-Ansatz (Gibbons, 2002) vorsieht: Schon die in einer geplanten Unterrichtsreihe eingesetzten Lehrmaterialien werden hinsichtlich der enthaltenen Textarten und sprachlichen Strukturen analysiert (Bedarfsanalyse), um sie mit den Ergebnissen von Sprachstandserhebungen (Lernstandserfassung) zu vergleichen. Die darauf aufbauende Planung von einzelnen Unterrichtsstunden

2 Zur bewegten Entstehungsgeschichte des DaZ-Moduls und verwandten Lehrangeboten am Institut sei hier aus Platzgründen auf die bei Baur und Scholten-Akoun (2009) und Cantone et al. (2022) versammelten Beiträge verwiesen.

verschränkt fachliche und sprachliche Aspekte und bildet den Rahmen für weitere Handlungen in der tatsächlichen Unterrichtsinteraktion (Kniffka, 2012). Deshalb sollte Sprachdiagnostik vielmehr als ein regelmäßig mitlaufender Prozess verstanden werden und diagnostische Instrumente situationsadäquat angewendet werden. Dies wiederum erfordert ein geschultes „diagnostisches Auge“, das bestimmte auftretende Muster und vorhandene Ressourcen schon im Alltag erkennt und ein breites Wissen zur Abwägung von verschiedenen Formen der Intervention voraussetzt (Schroedler et al., 2024).

Ähnlich verhält es sich mit der Auseinandersetzung mit Spracherwerbsverläufen und Sprachbiographien von Schüler:innen. Hierbei spielen psycholinguistische (Zwei-/ Mehrsprachigkeit, Erstsprache, Zweitsprache etc.) und soziolinguistische Kategorisierungen (Herkunftssprache/Familiensprache, Mehrheits-/ Minderheitensprache, aber auch Dialekt, Register, Fachsprachen etc.) sowie die damit verbundenen Merkmale eine wichtige Rolle. Der souveräne und informierte Umgang mit solchen Begrifflichkeiten, die in der Schulpraxis auch zu undifferenzierten Zuschreibungen führen können, erfordert ein hohes Maß an Aufmerksamkeit und Reflexivität in Bezug auf Sprache. Schließlich ist es im Angesicht von z. T. emotionalisierten medialen Debatten um die Wechselwirkung von Migrationsgesellschaft und Bildung(-spolitik) als Lehrkraft auch notwendig, eine eigene, fachwissenschaftlich fundierte kritische Haltung zu solchen Diskursen entwickeln zu können.

Die Vermittlung dieser disziplinspezifischen Denkmuster ist ein wichtiges Anliegen der Lehre und nimmt bereits einen hohen Stellenwert in der Gestaltung ein. Dennoch hat es sich für die Projektarbeit als äußerst produktiv erwiesen, den Blick gerade auf die (vermuteten) Hürden und Schwierigkeiten für die Studierenden ‚auf dem Weg zur Expertise‘ zu richten. Dazu gehören auch Aspekte wie Motivation und Aktivierung. So könnten Zweifel an der Relevanz von DaZ für den Lehrberuf schon die Bereitschaft zur Aneignung des Fachwissens beeinflussen. In diesem Zusammenhang ist eine stellenweise auftretende Fehlkonzeption, Sprache als reines Mittel zum Zweck (der Wissensvermittlung) zu begreifen. Hier lässt sich vermuten, dass auch ein Selektionsbias vorliegt: Wer ein Lehramtsstudium aufnimmt, hat die eigene Schullaufbahn bereits erfolgreich mit der Hochschulreife abgeschlossen – wurden dabei während der eigenen Schulzeit keine (positiven) Erfahrungen mit Mehrsprachigkeitsdidaktik oder sprachsensiblen Fachunterricht gesammelt, werden diese womöglich als entbehrlich wahrgenommen. So kann einerseits die Sensibilität für die exkludierenden Mechanismen des Schulsystems (im Sinne von „Ich habe es ja irgendwie geschafft“), andererseits auch das Vorstellungsvermögen von produktivem Umgang mit dem mehrsprachigen Klassenraum eingeschränkt sein. Studien zeigen jedoch, dass Lehramtsstudierende positive Überzeugungen zu Mehrsprachigkeit und deren Relevanz im Fachunterricht im Laufe ihres Studiums entwickeln (Schroedler/Fischer, 2020).

Doch auch wenn eine positive Haltung zum DaZ-Modul und dessen Relevanz für den eigenen Professionalisierungsprozess schon im Bachelorstudium vor-

liegt, so bleibt die Auseinandersetzung mit linguistischen Beschreibungskategorien mühsam, gerade bei Lernenden mit nichtphilologischen Studienfächern. Auf der Suche nach den richtigen Methoden und Ansätzen, um eine durchaus empfundene Wertschätzung für Mehrsprachigkeit auszudrücken und Bildungsungerechtigkeiten mit dem eigenen Handeln zu adressieren, wirken Lernziele wie das korrekte Identifizieren von Wortarten und Verbstellungsmustern zuweilen mechanisch, starr und demotivierend. In Mangel eines linguistischen Propädeutikums im Modul sind die in der Studierendenschaft unterschiedlich stark ausgeprägten Wissenslücken, teilweise aus Schulzeiten, somit in Eigenarbeit auszugleichen.

Ein weiteres Hindernis lässt sich im Prüfungsformat selbst sehen. Zum Abschluss des DaZ-Moduls wird eine Studienleistung in den Seminaren und eine Klausur zu den Vorlesungsinhalten absolviert. Die von den Lehrenden der Seminare gewählte (i. d. R. schriftliche) Studienleistung bietet grundsätzlich die Gelegenheit, sich vertieft und kritisch mit Fachliteratur auseinanderzusetzen, um zu einer bestimmten Fragestellung eine eigene professionelle Haltung zu entwickeln und das Begründen dieser zu üben. Jedoch ist diese obligatorische Leistung unbetont, während das Ergebnis in der Klausur die Modulabschlussnote bestimmt.

Während die ganzheitliche Entwicklung von Fachkompetenzen durch die Arbeit im Seminar bzw. an der dazugehörigen Studienleistung wenig incentiviert wird, ermutigt die Gewichtung der Klausur Studierende dazu, sich mit Vorlesungsfolien und Altklausuren auf die Prüfung vorzubereiten. Im folgenden Abschnitt werden die Ursachen dieses Umstands näher beleuchtet und daraus abgeleitete Implikationen vorgestellt.

3 Prüfungsmodalitäten und ihre Auswirkungen auf das studentische Lernen

Wie eingangs beschrieben, ist ein wesentliches Merkmal der Klausur im DaZ-Modul die große Anzahl der Teilnehmenden in jedem Semester (circa 350–700 Studierende). Die Klausur wurde bereits vor den Herausforderungen der COVID-19-Pandemie als PC-Klausur umgesetzt und stetig verändert. Die Umsetzung als PC-Klausur birgt zwar effizienzsteigernde Möglichkeiten hinsichtlich der Durchführung und Auswertung der Prüfung für die Modulverantwortlichen. Der Aufwand der *guten* Vorbereitung der Klausur kann jedoch vermutlich als ähnlich zu einer papierbasierten Klausur gewertet werden. Aufgrund der hohen Teilnehmendenzahl sowie aus Gründen der Gerechtigkeit bezüglich der objektiven Bewertung wird auf Freitextaufgaben verzichtet. Die Klausur besteht also ausschließlich aus geschlossenen Items (z. B. *single choice*, *multiple choice*). Konkret heißt dies, dass nach jeder Fragestellung entweder eine oder mehrere richtige Antworten (Attraktoren) sowie eine oder mehrere falsche Antworten (Distraktoren) eingeblendet werden. Durch die langjährige Arbeit kann auf eine große, thematisch gruppierte Sammlung von Klausuraufgaben zurück-

gegriffen werden, die die tatsächliche Zahl bei einer konkreten Klausur eingesetzter Items bei weitem übersteigt.

Wie weiter oben beschrieben, werden in der Vorlesung des DaZ-Moduls die Inhalte für die hier beschriebene Prüfung vermittelt. Dabei handelt es sich um fachübergreifende Grundlagenkenntnisse in unterschiedlichen Bereichen von DaZ/Mehrsprachigkeit. Ohne epistemologische Grundlagendiskurse zu eröffnen, können die vermittelten Inhalte als primär deklaratives Wissen (Konerding, 2015, S. 61) bezeichnet werden. Dies schließt den Erwerb unterschiedlicher Terminologie, Theorie, aber auch (theoretischer) Kenntnisse über Abläufe (z. B. die Durchführung der o. g. Profilanalyse innerhalb der Sprachstandsdiagnostik) mit ein. Der Erwerb prozeduralen Wissens – nach Konerding (2015) die Fähigkeit, zweckorientiert und erfolgreich zu handeln – wäre im Kontext der Lehrkräftebildung erst später im Professionalisierungsprozess angesiedelt.

Eine gute Prüfung soll, testtheoretisch betrachtet, valide, reliabel und objektiv sein (z. B. Przyborski/Wohlrab-Sahr, 2021). Hierzu gehört die Herstellung einer bestmöglichen Chancengleichheit zwischen den Klausurteilnehmenden. Diese schwierige Frage nach der Fairness impliziert neben vielen Faktoren auch das Entgegenwirken gegen studierendenseitig unlautere Möglichkeiten der Klausurvorbereitung sowie der Täuschung. Bekannt ist im Kontext der hier beschriebenen Prüfung, dass weitreichende Kenntnisse über die Inhalte früherer Prüfungen über diverse Social-Media-Kanäle kursieren. Diese Kanäle stehen jedoch nicht allen Studierenden gleichermaßen zur Verfügung, was wiederum eine schwer zu überblickende Vorwissensheterogenität und somit Chancenungerechtigkeit generiert. Da es weder sinnvoll noch umsetzbar ist, jedes Semester eine von Grund auf neue Klausur zu konzipieren, werden mittels des genannten Aufgabenpools die tatsächlich eingesetzten Aufgaben von Klausur zu Klausur variiert. Neben der Variation durch Austausch von ganzen Klausuritems wird auch die Zusammenstellung von Attraktoren und Distraktoren innerhalb von Aufgaben semesterweise verändert. Dabei handelt es sich bei der Variationsstrategie lediglich um eine Abmilderung des zugrunde liegenden Problems, die negative Nebeneffekte erzeugt. Zum einen lässt sich nur eine begrenzte Anzahl an Variationen von korrekten Antwortmöglichkeiten generieren, da es sich um Grundlagenwissen handelt. Diese können ebenso verbreitet werden, was ein an der sprachlichen Oberfläche orientiertes Auswendiglernen potenziell weiter belohnt. Die Variation der Distraktoren dahingegen ist deutlich produktiver, erzeugt aber unterschiedlich stark ‚ablenkende‘ Auswahlmöglichkeiten. Auch wenn der Effekt auf das Gesamtklausurergebnis vernachlässigbar ist, werden Aufgabenvarianten in ihrer Schwierigkeit weniger vergleichbar.

Diese lediglich ausschnitthaft dargestellten Herausforderungen hinsichtlich der Modulabschlussprüfung haben Auswirkungen auf den Professionalisierungsprozess der angehenden Lehrkräfte im Fachanteil DaZ. Während in der Lehre im Modul an stetiger Innovation hinsichtlich der Vermittlung der Inhalte gearbeitet wird, erscheint die von einem gewissen Pragmatismus geprägte Art und

Weise der studierendenseitigen Prüfungsvorbereitung (Fokus auf das Auswendiglernen) wenig förderlich für die tiefere Auseinandersetzung mit den Lerngegenständen. Das Verhalten ist gleichwohl nachvollziehbar: In grober Orientierung an den für das schulische Prüfungen gängigen drei Anforderungsbereichen von Aufgaben (I. Wissen/Kennen, II. Anwenden/Übertragen und III. Urteilen/Bewerten, z. B. KMK, 2002) dargestellt, gelingt es in nur wenigen Themenbereichen der Klausur, den zweiten oder dritten Anforderungsbereich geltend zu machen. Zum einen lassen sich in geschlossenen Items nur sehr schwer (praktische) Anwendungskompetenzen und Urteilskompetenzen modellieren. Zum anderen kann davon ausgegangen werden, dass neue Items, die einmal verwendet werden, nicht mehr gänzlich ihren eigentlichen Zweck erfüllen, wie folgendes Beispiel illustriert.

Wenn z. B. in der Lehrveranstaltung die Zusammenhänge zwischen der statistischen Kategorie ‚Migrationshintergrund‘, der linguistischen Kategorie ‚individuelle, migrationsbedingte Mehrsprachigkeit‘ und der bildungswissenschaftlichen Kategorie ‚Bildungsbenachteiligung‘ adressiert werden, müssen dahinterliegende Phänomene wie statistische Überrepräsentationen, Entstehungs- und Wirkmechanismen verstanden und richtig beurteilt werden, um Implikationen für ‚gutes Lehrkräftehandeln‘ abzuleiten. Dieser Themenbereich würde in der Klausur jedoch lediglich durch eine Abfrage im Schema ‚Welche der folgenden Aussagen ist richtig?‘ geprüft. Um diesen Wissensbestand sinnvoll zu fassen, kann lediglich eine begrenzte Anzahl von Aussagen zu den Zusammenhängen als Attraktoren formuliert werden. Sobald diese Attraktoren einmal verwendet wurden, kann davon ausgegangen werden, dass sie in die studentische Sammlung von Altklausuren übergehen. So würde für diejenigen Studierenden, die am Klausurtag erstmalig mit dem Item konfrontiert werden, eine Aufgabe entstehen, die eine Abstraktions- bzw. Transferkompetenz erfasst. Für Studierende, die eine Altklausur für ihre Vorbereitung nutzen, ginge es jedoch um die Reproduktion der auswendig gelernten Attraktoren.

Im weiter oben eingeführten Themenbereich der Sprachdiagnostik findet sich ein weiterer Klausurgegenstand, der an dieser Stelle kurz dargestellt und beurteilt werden soll. Das Erstellen einer sog. Profilanalyse der sprachlichen Fähigkeiten von Schüler:innen ist ein prominenter Gegenstand in den meisten DaZ-Modul-Lehrveranstaltungen an der UDE. Die Fähigkeit, eine solche Profilanalyse durchzuführen, ist in der Klausur über eine Sammlung geschlossener Items darstellbar, indem einzelne Textsegmente vorgegeben werden und einer Profilstufe zugeordnet werden müssen. Hierbei handelt es sich um eine Aufgabe, die das korrekte Anwenden von Wissen prüft. Textprodukte von Schüler:innen können für die Klausur mit relativ wenig Aufwand regelmäßig erneuert werden. Analog hierzu gibt es weitere Verfahren der Sprachdiagnostik, für die aufwandsarm ständig neue Beispiele mit ähnlichem Schwierigkeitsgrad erstellt werden können. So handelt es sich im Bereich der Sprachstandsdiagnostik um verhältnismäßig gut funktionierende Klausuritems, insofern als dass ‚geleakte‘ Items

den Lernprozess der Studierenden kaum beeinflussen und die Aufgaben für jede Kohorte von Prüfungsteilnehmenden ähnlich valide sind.

Veränderungen und Implikationen

Das Projekt PITCH ermöglichte den Verantwortlichen des DaZ-Moduls eine substanzielle Verbesserung der hier beschriebenen Abschlussklausur. Dazu war es zunächst wichtig, die disziplinspezifischen Denk- und Handlungsmuster im Modul zu entschlüsseln und damit verbundene bestimmte Hürden studentischen Lernens zu identifizieren. Im Sinne einer Verbesserung des *Constructive Alignments* (z. B. Biggs, 1996), also der Übereinstimmung von Lehr-Lerninhalten und Prüfungsinhalten, wurden diese zunächst einer umfassenden Überprüfung unterzogen. Daraus resultierte eine Abkehr von Prüfungsfragen, welche nur unzureichend zu den Lehr-Lerninhalten passten sowie die Erstellung neuer Items auf Basis von Lehr-Lerninhalten, die bislang nur unzureichend als Prüfungsinhalt repräsentiert waren. Für die Vorbereitung in einem der Kernthemen, der oben beschriebenen Sprachdiagnostik, wurde ein Übungskurs entwickelt und als OER veröffentlicht (Schroedler et al., 2024).

Darüber hinaus hat dieser Beitrag eine Reihe bestehender Herausforderungen adressiert. So wurde gezeigt, dass im ‚Fach‘ DaZ (als lediglich kleines, jedoch verpflichtendes Modul für alle angehenden Lehrkräfte) disziplinär heterogenes Wissen vermittelt wird. Die unterschiedlichen Themen sind mit unterschiedlichen Schwierigkeiten sowohl im studentischen Lernprozess (z. B. Unterschiede zwischen Studierenden philologischer Fächer und solchen ohne beim Erlernen linguistischer Konzepte) als auch in ihrer Abbildbarkeit in einer Abschlussklausur mit hoher Teilnehmendenzahl verbunden. In Bezug auf die Chancengerechtigkeit ist der Tatsache, dass unterschiedliche Studierende auf unterschiedlich viele Altklausur-Items Zugriff haben, nur schwierig entgegenzuwirken.

Zusammenfassend lässt sich mit Blick auf die bisherige Projektarbeit und den angestoßenen *Decoding*-Prozess infolgedessen festhalten, dass Aufgabenformate in der Prüfung und ihr Verhältnis zu unbenoteten Studienleistungen möglicherweise noch viel radikaler überdacht werden können. Wie eine bessere Kompetenzorientierung und Orientierung an den unterschiedlichen Anforderungsbereichen mit den begrenzten zur Verfügung stehenden Ressourcen hergestellt werden kann, bleibt ebenfalls offen und wird weiter bearbeitet. Ferner könnte der Übergang von deklarativem Wissen in prozedurales Wissen im Professionalisierungsprozess angehender Lehrkräfte (hier: für den Bereich DaZ) intensiver beleuchtet werden.

Literatur

- Barkowski, H. (2003). 30 Jahre Deutsch als Zweitsprache – Rückblick und Ausblick. In: Informationen Deutsch als Fremdsprache, 30(6), 521–540.
- Baur, R. S./Scholten-Akoun, D. (Hrsg.) (2009). Deutsch als Zweitsprache in der Lehrerausbildung. Bedarf – Umsetzung – Perspektiven: Dokumentation der Fachtagungen zur Situation in Deutschland und in Nordrhein-Westfalen am 10. und 11. Dezember 2009. <https://www.stiftung-mercator.de/de/publikationen/deutsch-als-zweitsprache-in-der-lehrerausbildung-bedarf-umsetzung-perspektiven/>
- Biggs, J. (1996). Enhancing teaching through constructive alignment. In: Higher Education, 32(3), 347–364.
- Busse, V. (2020). Qualifizierung von Lehramtsstudierenden zum Umgang mit Mehrsprachigkeit. In I. Gogolin/A. Hansen/S. McMonagle/D. Rauch (Hrsg.), Handbuch Mehrsprachigkeit und Bildung (S. 287–292). Wiesbaden: Springer Fachmedien.
- Cantone, K. F./Gürsoy, E./Lammers, I./Roll, H. (Hrsg.). (2022). Fachorientierte Sprachbildung und sprachliche Vielfalt in der Lehrkräftebildung: Hochschuldidaktische Formate an der Universität Duisburg-Essen. Münster, New York: Waxmann.
- Gibbons, P. (2002). Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom. Portsmouth: Heinemann.
- Griebhaber, W. (2012). Die Profilanalyse. In B. Ahrenholz (Hrsg.), Einblicke in die Zweitspracherwerbsforschung und ihre methodischen Verfahren (173–193). Berlin: DE GRUYTER.
- KMK (2002). Einheitliche Prüfungsanforderungen in der Abiturprüfung Deutsch: Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 24.05.2002.
- Kniffka, G. (2012). Scaffolding – Möglichkeiten, im Fachunterricht sprachliche Kompetenzen zu vermitteln. In M. Michalak/M. Kuchenreuther (Hrsg.), Grundlagen der Sprachdidaktik Deutsch als Zweitsprache (1. Aufl.) (S. 208–225). Baltmannsweiler: Schneider-Verlag Hohengehren.
- Konerding, K.-P. (2015). 3. Sprache Und Wissen. In E. Felder/A. Gardt (Hrsg.), Handbuch Sprache und Wissen (S. 57–80). Berlin: DE GRUYTER.
- Meyer, J. H. F./Land, R. (2003). Threshold Concepts and Troublesome Knowledge: Linkages to Ways of Thinking and Practising within the Disciplines. In: C. Rust (Hrsg.), Improving Student Learning: Improving Student Learning Theory and Practice – Ten Years On (S.1–12). Oxford: Oxford Centre for Staff and Learning Development.
- Meyer, J. H. F./Land, R. (2005). Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. In: Higher Education, 49(3), 373–388.
- Middendorf, J./Pace, D. (2004). Decoding the disciplines: A model for helping students learn disciplinary ways of thinking. In: New Directions for Teaching and Learning, 2004(98), 1–12.
- Middendorf, J./Shopkow, L./Bernstein, D. (2023). Overcoming student learning bottlenecks: Decode the critical thinking of your discipline. New York: Routledge.
- Ministerium des Innern des Landes NRW (2009). Gesetz- und Verordnungsblatt (GV. NRW.) Ausgabe 2009 Nr. 14 vom 25.5.2009 (S. 307–320).
- Perkins, D. N. (1999). The Many Faces of Constructivism. In: Educational Leadership, 57(3), 6–11.

- Przyborski, A./Wohlrab-Sahr, M. (2021). *Qualitative Sozialforschung: Ein Arbeitsbuch* (5. Aufl.). Lehr- und Handbücher der Soziologie. Berlin, Boston: De Gruyter Oldenbourg. <https://doi.org/10.1515/9783110710663>
- Putjata, G./Olfert, H./Romano, S. (2016). Mehrsprachigkeit als Kapital – Möglichkeiten und Grenzen des Moduls »Deutsch für Schülerinnen und Schüler mit Zuwanderungsgeschichte« in Nordrhein-Westfalen. In: *ÖDaF-Mitteilungen*, 32(1), 34–44.
- Riegler, P. (2020). Einflüsse von Decoding the Disciplines auf die Gestaltung von Lehr- und Lernprozessen. In: *die hochschullehre*, 6(23). <https://doi.org/10.3278/HSL2023W>
- Schroedler, T./Enzenbach, C./Kania, T./Roll, H. (2024,). Authentische Lerner*innentexte als Gegenstand fallorientierter Hochschullehre zur Vorbereitung angehender Lehrkräfte für den Umgang mit DaZ und Mehrsprachigkeit im Fachunterricht. In: *Herausforderung Lehrer*innenbildung*, 7(2), 185–201. <https://doi.org/10.11576/hlz-5358>
- Schroedler, T./Fischer, N. (2020). The Role of Beliefs in Teacher Professionalisation for Multilingual Classroom Settings. In: *European Journal of Applied Linguistics*, 8(1), 49–72.
- Schroedler, T./Grommes, P. (2019). Learning about Language: Preparing pre-service subject teachers for multilingual classroom realities. In: *Language Learning in Higher Education*, 9(1), 223–240.
- Schroedler, T./Salzmann-Hoang, N. M./Schmitz, L./Zindel, L./Gerhardt, S./Kania, T. (2024). OpenÜsK – OER zur Übung sprachdiagnostischer Kompetenzen. https://duepublico2.uni-due.de/receive/uepublico_mods_00081874
- Wiemer, M./Kenneweg, A. C. (2021). Threshold Concepts: Übergänge zu disziplinären Denkweisen und transformative Lernprozesse in der Fachlehre verstehen und begleiten. In B. Berendt/A. Fleischmann/ G. Salmhofer/N. Schaper/B. Szczyrba/M. Wiemer/J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre*, 100 (S. 43–66). Berlin: DUZ- Verlags- und Medienhaus. <https://doi.org/10.36197/DUZOPEN.029>

Digital üben,
prüfen und bewerten

Das Beste aus Online und Präsenz – Open-Book-Prüfungen mit handschriftlichen Lösungen und digitalem Bewertungsworkflow¹

Mathias Magdowski

Die Gestaltung einer adäquaten, fairen, transparenten, authentischen, effizienten, skalierbaren und (einigermaßen) rechtssicheren Prüfung ist aus meiner Sicht mindestens so komplex und aufwändig wie die entsprechende Gestaltung und Konzeption der zugehörigen Lehrveranstaltung selbst. Ausgelöst durch die Kontaktbeschränkungen der Corona-Pandemie habe ich den seit meiner Studienzeit und meinen Anfängen als Lehrender/Prüfer nie wirklich in Zweifel gezogenen Status quo von schriftlichen und beaufsichtigten Präsenzklausuren in Frage gestellt und verschiedene Varianten offenerer Prüfungsformate mit (teil-)digitalen Bewertungsprozessen entwickelt und erprobt, über die ich in diesem Beitrag berichten möchte. Dabei gehe ich zunächst auf die Ziele und Herausforderungen von Prüfungen insbesondere in den Ingenieurwissenschaften ein, beschreibe dann die während und nach der Pandemie eingesetzten Prüfungsformate und schließe mit einem Ausblick auf zukünftige Prüfungen.

1 Was sind Herausforderungen und Zielstellungen von Grundlagen-Prüfungen in den Ingenieurwissenschaften?

Eine summative Prüfung bildet häufig den Abschluss eines Moduls oder einer Lehrveranstaltung, insbesondere in den Grundlagenfächern. Bei größeren Kohorten ab ca. 20 Studierenden sind schriftliche Prüfungen aufgrund ihrer besseren Skalierbarkeit gegenüber zeitaufwändigeren mündlichen Prüfungen vorteilhaft. Inhaltlich muss die Prüfung im Sinne des Constructive Alignments natürlich zu den Lernzielen und Lernmethoden des Moduls passen und sollte den Studierenden auch ein wertvolles und konstruktives Feedback zu ihrem Lernprozess geben, selbst wenn dieses bei einer summativen Prüfung am Ende eines Moduls nur für darauf folgende Module anwendbar ist. Neben der eigentlichen Messung und Kontrolle des Lernerfolgs dient eine Prüfung aus Sicht der Studienorganisation aber vor allem der individuellen Bewertung, Bepunktung, Benotung oder dem Vergleich der Leistung der Studierenden.

¹ Der Beitrag wurde am 02.11.2023 auf dem Blog von Mathias Magdowski veröffentlicht: <https://mathiasmagdowski.wordpress.com/2023/11/02/digitale-praesenzpruefung/>

In den Ingenieurwissenschaften sind Prüfungen dabei zum überwiegenden Teil schon immer recht kompetenzorientiert gestaltet gewesen. Meist bestehen Prüfungsaufgaben nicht aus reinen Wissensabfragen, sondern sind immer anwendungs- und problemorientiert gehalten. Es kommt für einen Prüfling also nicht nur darauf an, eine Methode zu kennen, zu beschreiben oder erklären zu können. Stattdessen ist es wichtig, eine Methode für bzw. auf eine vorgegebene spezifische Problemstellung anwenden zu können. Bei einfacheren Aufgaben ist die Methode dabei fest vorgegeben. Bei komplexeren Aufgaben und Problemen geht es auch darum, aus einer Sammlung bekannter Methoden und Verfahren eine möglichst passende und effiziente Variante für die vorgegebene Problemstellung auszuwählen und dann zielgerichtet und korrekt anzuwenden. Mit solchen Prüfungsaufgaben lässt sich eine Handlungskompetenz in ergebnisoffenen Situationen relativ gut prüfen, die bei erfolgreichen Prüflingen auch ein großes fachspezifisches und gut vernetztes Domänenwissen voraussetzt.

Prüfungsaufgaben, insbesondere in der Elektrotechnik und verwandten Fächern, bestehen dabei meist aus einem kurzen Text zur Erklärung der Problemstellung, einem zugehörigen Schaltbild oder Diagramm und einer oder mehrerer konkreter Aufgaben oder Fragestellungen zu gesuchten Größen. Zur Lösung notieren Studierende einen Ansatz und Rechenweg, schreiben Formeln auf, stellen diese in Richtung der gesuchten Größen um und setzen sie ineinander ein, verrechnen Zahlen sowie Einheiten miteinander und skizzieren entsprechend vereinfachte Schaltungen oder erhaltene mathematische Funktionen in neuen Diagrammen. All das funktioniert kaum mit Single-Choice- bzw. Multiple-Choice-Fragen oder den anderen typischen Online-Aufgabenformaten, jedoch sehr einfach und intuitiv mittels Handschrift, einem herkömmlichen Stift und analogem Papier.

Aufgabe 80: Ladung und Strom (9 Punkte)

Gegeben ist der folgende Zeitverlauf der elektrischen Ladung $Q(t)$ im Bereich $0 \leq t \leq 50$ ms.

- Man berechne den zugehörigen Stromverlauf $i(t)$.
- Man stelle $i(t)$ graphisch dar.

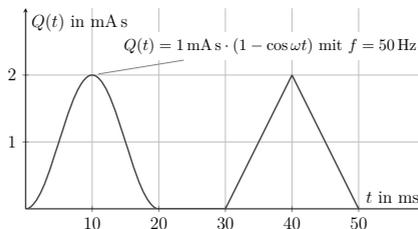
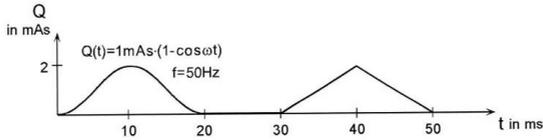


Abbildung 1: Beispielhafte Aufgabe (Quelle: Eigene Abbildung)

klade 80

Gegeben ist der folgende Zeitverlauf der elektrischen Ladung $Q(t)$ im Bereich $0 \leq t \leq 50\text{ms}$.



- a) Berechnen Sie den zugehörigen Stromverlauf $i(t)$.
- b) Stellen Sie $i(t)$ graphisch dar.

zu a) $0 \leq t < 20\text{ms}$: $Q(t) = 1\text{mAs} \cdot (1 - \cos \omega t)$
 $i(t) = \frac{dQ}{dt} = 1\text{mAs} \cdot (-\omega \cdot (-\sin \omega t)) \quad 1$
 $i(t) = \frac{1\text{mAs} \cdot 2 \cdot \pi}{T} \cdot \sin \omega t \quad 1$
 $i(t) = 0,314\text{A} \cdot \sin \omega t \quad // \quad \frac{1}{2}$

$20\text{ms} \leq t < 30\text{ms}$: $Q = \text{constant} \rightarrow i = \frac{dQ}{dt} = 0 \quad // \quad \frac{1}{2}$

$30\text{ms} \leq t < 40\text{ms}$: $Q(t) = \frac{2\text{mAs}}{10\text{ms}} \cdot t + K_1 \quad 1$
 $i(t) = \frac{dQ}{dt} = \frac{2\text{mAs}}{10\text{ms}} = 0,2\text{A} \quad // \quad 1$

$40\text{ms} \leq t < 50\text{ms}$: $Q(t) = -\frac{2\text{mAs}}{10\text{ms}} \cdot t + K_2 \quad \frac{1}{2}$
 $i(t) = \frac{dQ}{dt} = -\frac{2\text{mAs}}{10\text{ms}} = -0,2\text{A} \quad // \quad \frac{1}{2}$

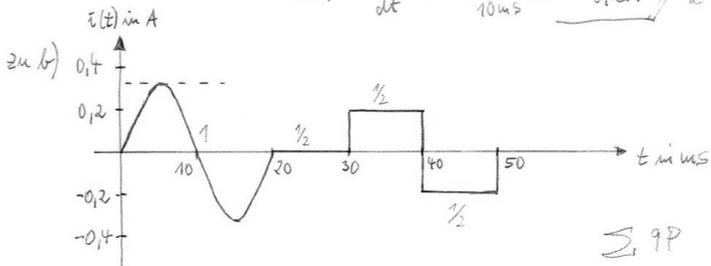


Abbildung 2: Handschriftliche Lösung zur beispielhaften Aufgabe (siehe Abbildung 1) (Quelle: Eigene Abbildung)

Natürlich kann man Formeln auch in Formeditoren oder Textsatzprogrammen wie LaTeX schreiben, Gleichungen in Computeralgebrasystemen wie Maple oder Maxima umstellen, eine Zahlenrechnung in Numerikprogrammen wie MATLAB oder GNU Octave durchführen, Diagramme per Computerprogramm wie Gnuplot plotten oder Schaltungen in Netzwerksimulatoren wie LTspice simulieren, und all das machen berufstätige Ingenieur*innen auch täglich. Diese weiterführenden Kompetenzen gehen aber etwas über das Ziel einer Grundlagenlehrveranstaltung hinaus und werden von Studierenden eher in Laborpraktika, Projektseminaren oder Abschlussarbeiten entwickelt.

Außerhalb des Zielbereichs klassischer Grundlagenprüfungen sind auch überfachliche Kompetenzen wie Kollaboration und Kooperation, effektive Gruppenarbeit oder Kommunikation, die im Rahmen des Studiums ebenso in Laborpraktika oder Projektseminaren entwickelt, gefördert und geprüft werden. Neben der summativen Prüfung gibt es im Modul „Grundlagen der Elektrotechnik“ ein formatives Assessment-Format zur Prüfungszulassung durch semesterbegleitende personalisierte Aufgaben mit anonymen Peer Review (vgl. Magdowski, 2020) sowie einer Online-Leistungskontrolle in jedem Semester (vgl. Magdowski/Siegert, 2021).

2 Wie liefen die Prüfungen unter Pandemiebedingungen?

2.1 Erstes Corona-Sommersemester 2020 – Alles wie immer

Unter Pandemiebedingungen konnten beaufsichtigte Präsenzprüfungen bzw. Klausuren aufgrund der Kontaktbeschränkungen nicht ohne weiteres stattfinden. Am Ende des ersten Corona-Sommersemesters 2020 waren Präsenzprüfungen unter Einhaltung gewisser Hygienemaßnahmen jedoch noch möglich und wurden an der Otto-von-Guericke-Universität in Magdeburg, z. B. mit sehr großen Abständen zwischen den Prüflingen in einer Messehalle, durchgeführt, ohne viel am bisherigen beaufsichtigten Prüfungsformat zu ändern. Studierende durften dabei im Modul „Grundlagen der Elektrotechnik“ wie in bisherigen Prüfungen als einzige Hilfsmittel eine eigene Formelsammlung auf einem A4-Blatt sowie einen Taschenrechner nutzen.

2.2 Erstes Corona-Wintersemester 2020/2021 – Take Home ist der Hit

Ganz anders war dann die Situation am Ende des ersten Corona-Wintersemesters 2020/2021. Die Fallzahlen waren sehr hoch und aufgrund der Kontaktbeschränkungen waren Präsenzprüfungen nahezu unmöglich. Aufgrund der Alternativlosigkeit und einer gewissen Vorhersagbarkeit dieser Situation, einer

dementsprechend langen Vorbereitungszeit für ein pandemiekonformes Online-Prüfungsformat, einer hohen Gewöhnung der Studierenden an Online-Lehrveranstaltungen und den entsprechenden Umgang mit digitalen Werkzeugen sowie dem Lernmanagementsystem der Universität, erschienen Online-Take-Home-Prüfungen als plausible Lösung.

Bisherige Herausforderungen bei Online-Prüfungen wie Authentizität und Originalität (Prüft man die richtige Person?), Datenschutz (Können dritte Personen unerlaubt an Informationen aus der Prüfung gelangen?), Transparenz und Chancengleichheit (Ist das Prüfungsverfahren für alle Beteiligten klar, insbesondere bei Problemfällen?) sowie Verfügbarkeit und Integrität (Kann die Prüfung sicher, valide, fair und reproduzierbar durchgeführt werden? Sind die Ergebnisse nachvollziehbar und sicher gespeichert? Was passiert bei Serverabstürzen und Verbindungsabbrüchen?) wurden dabei zum Teil etwas ausgeblendet oder angesichts der Krisensituation zumindest weniger kritisch gesehen.

Da wir in den Grundlagen der Elektrotechnik schon immer einen recht umfangreichen sowie kompetenzorientierten Katalog von Prüfungsaufgaben hatten und dieser in der Dokumentbeschreibungssprache LaTeX verfügbar war, konnte ich mit etwas Programmieraufwand die Erstellung von randomisierten Prüfungsbögen und deren E-Mail-Versand an die Studierenden automatisieren. Damit erübrigte sich die ethisch sowieso sehr zweifelhafte, wenig zielführende und auf verschiedenen Wegen angreifbare Online-Fernüberwachung bzw. das Proctoring der Prüflinge per Videokamera und Safe-Exam-Browser.

Stattdessen erlaubten wir den Studierenden einfach die komplette Nutzung aller möglichen analogen sowie digitalen Werkzeuge inklusive des Internets und kompensierten den sich daraus ergebenden Vorteil durch eine Erhöhung der Aufgabenanzahl von 9 auf 10 unter Beibehaltung der Bearbeitungszeit von 180 Min. in der Prüfung. Die Studierenden sollten jedoch weiterhin handschriftliche Lösungen mit Stift auf Papier schreiben, diese abfotografieren oder einscannen und im Moodle-Kurs mit ihrem persönlichen Universitätslogin als primären Identifizierungsschlüssel und Sicherheitsfaktor hochladen, wofür insgesamt weitere 30 Min. Zeit zur Verfügung standen. Die Nutzung von elektronischen Stiften auf beschreibbaren Tablet-PCs oder Notebooks haben wir aufgrund der Chancengleichheit ausgeschlossen, weil sich dort handschriftliche Inhalte viel einfacher und schneller austauschen, korrigieren, verschieben und verändern lassen. Als zweiten Identifizierungsschlüssel und Sicherheitsfaktor gegen übermäßiges Contract Cheating und die Hilfe durch dritte Personen sollte auf jedem abfotografierten oder eingescannten Blatt auch der Studierendenausweis des Prüflings sichtbar sein.

Die anschließende Korrektur, Bewertung, Einsichtnahme und Archivierung liefen dann komplett digital über das Lernmanagement Moodle ab und funktionierten relativ gut. Herausfordernd war, dass sich die Korrektor*innen bei jeder Aufgaben nicht nur in einen anderen Lösungsweg, sondern auch in eine andere Aufgabenstellung und Musterlösung hineindenken mussten, was den Aufwand

signifikant erhöhte. Interessanterweise änderte sich der Notenspiegel trotz der vielen Änderungen und Anpassungen gegenüber den bisherigen Prüfungen kaum. Zusammenfassend wurde das Open-Book- und Open-Web-Format von den Studierenden gelobt. Gleichzeitig wurde aber der Wunsch geäußert, Prüfungen in Zukunft trotzdem wieder unter Aufsicht an der Universität durchzuführen, um die Chancengleichheit zu erhöhen und insbesondere die unerlaubte Unterstützung durch Dritte zu erschweren. Außerdem stellte das technisch nicht übermäßig komplizierte aufgabenweise Abfotografieren und Hochladen der eigenen handschriftlichen Lösung einige Studierende vor ungeahnte Herausforderungen, obwohl es im Laufe des Semesters mehrfach geübt und praktiziert wurde. Weitere Informationen zu diesem Prüfungsformat sind in Magdowski (2021b) zu finden.

2.3 Zweites Corona-Sommersemester 2021 – Open-Web in Präsenz mit analoger Bewertung

Zum Ende des zweiten Corona-Sommersemesters 2021 rückten dann Präsenzprüfungen wieder in den Bereich der Möglichkeiten. Aufgrund der guten Erfahrungen mit dem Open-Book- und Open-Web-Format sowie den randomisierten Aufgabenbögen blieben wir bei diesem sehr offenen Prüfungsformat, luden die Studierenden aber wieder in eine Messehalle in der Nähe der Universität ein, wohin sie ihre eigenen digitalen Endgeräte im Sinne eines BYOD-Konzepts mitbringen konnten. Zur Verminderung des Korrekturaufwands und zur weiteren Erhöhung der Chancengleichheit wurde der Aufgabenpool zur Zusammenstellung der individuellen Prüfungsbögen jedoch stark reduziert, auch weil ein Austausch der Studierenden in diesem Setting stark eingeschränkt ist. Weitere Informationen zu dieser Open-Book-Präsenzklausur sind in Magdowski (2021a) zu finden.

Auch dieses Format hat gut und relativ problemlos funktioniert, vor allem, weil die handschriftlichen Antwortbögen der Studierenden zum Ende der Prüfung einfach eingesammelt wurden und nicht individuell abfotografiert und hochgeladen werden mussten. Negativ blieb mir in Erinnerung, wie aufwändig die anschließende Logistik für die ebenso handschriftliche Korrektur der studentischen Lösungen war. Durch die Pandemie arbeiteten viele der an der Korrektur beteiligten Kolleg*innen mobil oder aus dem Home-Office, was den Austausch der analogen Prüfungsbögen drastisch erschwerte und verzögerte. Dafür musste man sich bei der Korrektur nicht mit unscharfen Fotografien mit zu geringer Auflösung, störenden Schatten oder inkompatiblen Dateiformaten auseinandersetzen. Gleichzeitig mussten die Studierenden aber sehr lange auf ihr Feedback warten, weil ich die korrigierten Lösungen erst nach Ende der kompletten Bewertung einscannen und an die Studierenden zurückschicken konnte. Dadurch war eine Online-Prüfungseinsicht mit einer raschen Diskussion zu konkreten Aufgaben-

bewertungen auch nur sehr zeitverzögert und damit ineffizient möglich. Außerdem muss man sich als Korrektor*in den kompletten Lösungsbogen der Studierenden anschauen, auch wenn man nur eine bestimmte Aufgabe korrigiert, was bei vielen Studierenden auch eine gewisse Zeit kostet. Zusammenfassend lässt sich zu dieser Variante aber festhalten, dass die früher als ganz selbstverständlich wahrgenommenen Implikationen eines analogen Bewertungsworkflows gegenüber der vorherigen rein digitalen Variante etwas negativ in Erscheinung traten.

2.4 Wintersemester 2021/2022, Sommersemester 2022 und Wintersemester 2022/2023 – BYOD-Prüfung mit digitalem Bewertungsworkflow

Für die Prüfungen in den drei darauffolgenden Semestern, dem Wintersemester 2021/2022, dem Sommersemester 2022 und dem Wintersemester 2022/2023, die alle auch noch unter einem gewissen Pandemieeinfluss standen, setzten wir deshalb wieder auf einen rein digitalen Bewertungsworkflow in Moodle, blieben aber weiterhin bei einem Open-Book- und Open-Web-Format als Präsenzvariante. Studierende in Quarantäne oder mit gesundheitlichen Einschränkungen konnten trotzdem online an der Prüfung teilnehmen, mussten dann aber aufgrund der Gleichbehandlung videoüberwacht werden. Die Aufgaben waren auch deshalb weiterhin randomisiert. Die Studierenden mussten ihre handschriftlichen Lösungen erneut abfotografieren und in Moodle hochladen. Dort wurden sie dann aufgabenspezifisch von einzelnen Korrektor*innen bewertet. Studierende erhalten somit sehr schnell und individuell Feedback. Der Bewertungsworkflow in Moodle ist außerdem recht effizient und intuitiv nutzbar. Als störend wurden von den Korrektor*innen aber weiterhin der erhöhte Aufwand durch die unterschiedlichen Aufgabenstellungen und zugehörigen Musterlösungen sowie die teilweise schlechte Bildqualität der Einreichungen wahrgenommen.

Außerdem stellte sich heraus, dass die Studierenden ihre zur Prüfung mitgebrachten Smartphones, Tablet-PCs oder Notebooks nicht wirklich effizient und zielführend nutzten. Statt diese Geräte als „sehr mächtigen Taschenrechner“ zu nutzen, Ableitungen und Stammfunktionen analytisch mit einem Computeralgebrasystem zu bestimmen, eine Rechnung mit komplexen Zahlen bzw. das Lösen von Gleichungssystem in Numerikprogrammen durchzuführen oder Netzwerksimulatoren zur Probe einzusetzen, wurden die Computer hauptsächlich als schnelle Nachschlagewerke zweckentfremdet, zur Volltextsuche im Skript oder Buch, zur Wikipedia-Recherche, zum schnellen Scrollen durch Übungsmitschriften oder den Klausuraufgabenkatalog der Fachschaft, mit dem Ziel, dort eventuell eine vorgefertigte, kochrezeptartige Lösung zu finden, die irgendwie zur vorgegebenen Aufgabe passt und dann mit geringen Änderungen und Anpassungen repliziert werden kann.

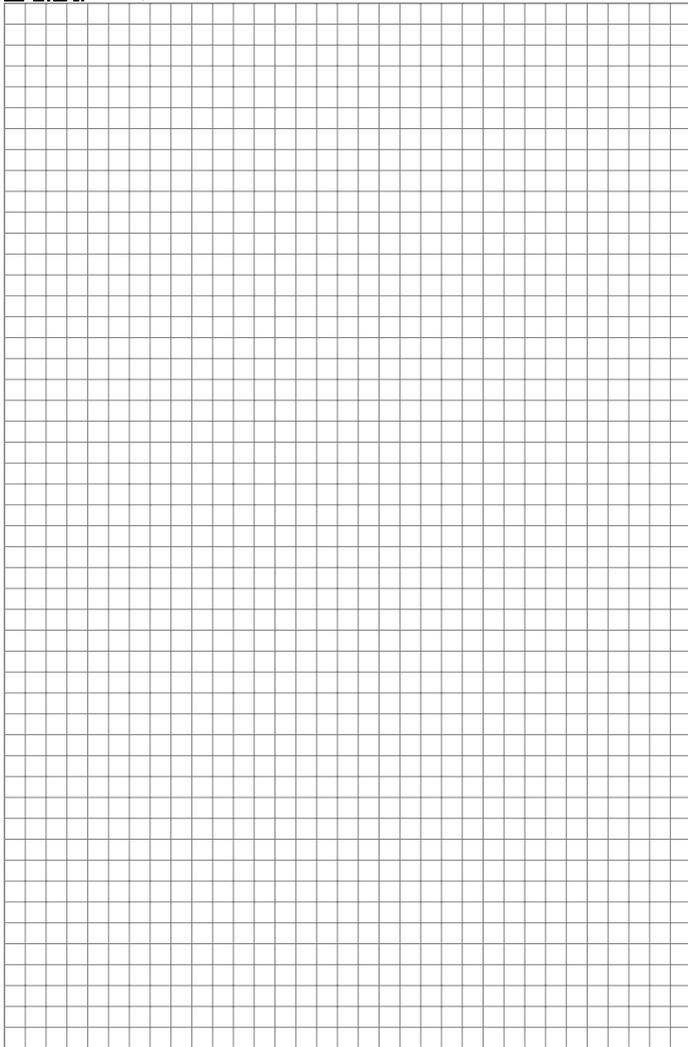
2.5 Sommersemester 2023 – Open-Book-Format, analoge Einreichung und digitale Bewertung

Aufgrund dieser Erfahrungen und durch den aufkommenden Einfluss von KI-gestützten großen Sprachmodellen und Assistenten wie ChatGPT wurde das Konzept der Prüfung im letzten Sommersemester 2023 nochmals angepasst, auch weil kaum noch Pandemieeinschränkungen vorhanden waren. Statt eines Open-Web-Formats mit freier Nutzung digitaler Endgeräte durften die Studierenden im Sinne eines Open-Book-Konzepts nur noch beliebig viele analoge Hilfsmittel wie Bücher, gedruckte Skripte, Aufgabensammlungen oder eigene Übungsmitschriften mitbringen und nutzen. Der Aufgabenumfang wurde deshalb auch wieder von 10 auf 9 Aufgaben reduziert. Da ein studentischer Austausch unter Beaufsichtigung und ohne digitale Endgeräte sehr erschwert ist, wurden die Aufgaben auch nicht mehr randomisiert, was die Vergleichbarkeit erhöht und den Korrekturaufwand deutlich vermindert.

Um weiterhin bei einem voll-digitalen Bewertungsworkflow mit schnellem Feedback zu bleiben, gleichzeitig aber eine gute Bildqualität und aufgabenspezifische Bewertung zu ermöglichen, habe ich Klausurpapier (siehe: <https://mathiasmagdowski.wordpress.com/2017/11/24/qr-code-papier/>) mit vorgedruckten QR-Codes genutzt, die ich in einer ähnlichen Variante bereits in den Jahren zuvor eingesetzt habe (siehe Magdowski, 2018). Durch die in den QR-Codes enthaltenen Informationen kann jedem Lösungsblatt der studentische Name, die zugehörige E-Mail-Adresse und die Aufgabennummer zugeordnet werden. Nach der Abgabe habe ich dann zunächst alle studentischen Einreichungen automatisiert eingescannt und über ein weiteres selbstgeschriebenes Programm mit Hilfe der QR-Codes nach Personen und Aufgabennummern sortiert. Die so strukturierten Lösungen wurden dann auf einem universitätsinternen Cloudspeicher hochgeladen und den Korrektor*innen als PDF-Datei zur Verfügung gestellt. Die Korrektur fand digital statt, wofür viele Kolleg*innen einfache Zeichentablets für handschriftliche Annotationen oder Kommentare in den PDF-Dateien nutzten. Nach der Korrektur wurden diese Dateien dann über Moodle direkt den Studierenden zugänglich gemacht, sodass zeitnahes Feedback und Rückfragemöglichkeiten vorhanden waren.

Dieses Format und der zugehörige Prozess haben sich sehr gut bewährt und bieten sich auch für zukünftige Prüfungen an. Nachteilig sind der Zeitaufwand zum Einscannen und für die manuelle Nacharbeit bei nicht oder falsch erkannten QR-Codes. Außerdem sind die Medienbrüche zwischen der Speicherung der unkorrigierten Lösungen in der Cloud, dem händischen Herunterladen, der lokalen Kommentierung und Zwischenspeicherung sowie dem anschließenden manuellen Hochladen in den Moodle-Kurs etwas störend, da pro zu korrigierender Lösung etwa 20 Klicks nötig sind, selbst wenn ein Prüfling nur ein leeres Blatt eingereicht hat. Beide Prozesse, das Einscannen sowie das Herunter- und Hochladen der Korrekturen, ließen sich sicherlich noch etwas optimieren.

 Name: Max Mustermann
Matrikelnummer: 123456
Aufgabe: 1



Klausurpapier mit QR-Code erstellt von: Mathias Magdowski (mathias.magdowski@ovgu.de)

Abbildung 3: Klausurpapier mit QR-Code (Quelle: Eigene Abbildung)

3 Was bleibt nach der Pandemie und in Zeiten von ChatGPT?

In den ersten frei verfügbaren Versionen 3 und 3.5 konnte ChatGPT nur mit Text umgehen. Die bisher ausschließlich kostenpflichtig nutzbare Version 4 kann jedoch auch Bilder einlesen und ausgeben, was die Interpretation sowie Ausgabe von Schaltbildern oder Diagrammen zur Bearbeitung typischer elektrotechnischer Aufgabenstellungen prinzipiell möglich macht. Die Entwicklung ist dabei sehr rasant und gleichzeitig schwer absehbar. Selbstverständlich müssen Studierende solche Werkzeuge auch im Studium nutzen, um einen souveränen Umgang damit zu erlernen. Gleichzeitig stellt sich bei der unreflektierten studentischen Anwendung von Diensten wie ChatGPT in Prüfungen die Frage, welcher Teil einer Lösung von ChatGPT und welcher vom Prüfling erstellt wurde. Bisher sind KI-Assistenten wie ChatGPT oder Google Bard nach eigener Erfahrung in ingenieurwissenschaftlichen Prüfungen eher weniger hilfreich und wenn ein Chatbot eine Aufgabe korrekt lösen kann, sagt das meist mehr über die Aufgabe als über den Chatbot selbst. Dieser Zustand kann sich jedoch schnell ändern.

Aus Prüfer*innen-Sicht kann man nun die Nutzung von KI-Werkzeugen wie ChatGPT in Prüfungen erlauben und dafür das Niveau der Aufgaben, deren Anzahl oder den Bewertungshorizont anpassen. Alternativ kann man versuchen, die Nutzung von ChatGPT in Prüfungen zu verbieten oder einzuschränken, was jedoch nicht mit einem BYOD-Konzept ohne Safe-Exam-Browser kompatibel ist.

Eine weitere Möglichkeit zur einschränkbaren Nutzung, z. B. nur für einen Teil der Prüfung, sind E-Prüfungszentren mit entsprechend konfigurierten PCs. Hier kann man durch entsprechende Rechteverwaltung gewisse Software freigeben (z. B. Computeralgebrasysteme, Numerikprogramme, Netzwerksimulatoren, Unterlagen im Moodle-Kurs, Wikipedia etc.), gleichzeitig aber die sonstige Nutzung des Internets einschränken (z. B. aktive Forennutzung, Messengerdienste, ChatGPT etc.) oder nur für einen Teil der Prüfung erlauben. Da in einem solchen E-Prüfungszentrum alle Studierenden ein funktionstüchtiges Endgerät mit der entsprechenden Software zur Verfügung haben, könnte man dort auch verpflichtende Aufgaben mit Softwarenutzung (z. B. von MATLAB oder LTspice) in Prüfungen vorsehen, was mit einem BYOD-Konzept nur eingeschränkt möglich ist. Werden die Rechner in einem E-Prüfungszentrum auch mit einem eigenen Scanner bzw. einem Digitizer ausgestattet, können Prüflinge ihre handschriftlichen Lösungen oder Lösungsansätze auch selbst einscannen und direkt zur Bewertung in Moodle hochladen und einreichen, was den vorher beschriebenen Scan-aufwand und die Medienbrüche minimiert bzw. verhindert.

Mehr verfügbare Technik, günstigere Hardware, gut ausgestattete E-Prüfungszentren oder zuverlässig nutzbare Zugänge über Lernmanagementsysteme wie Moodle ermöglichen viele Optionen zur flexiblen Gestaltung, Optimierung oder Anpassung von Prüfungsformaten an sich ändernde gesellschaftliche Randbedingungen. Gleichzeitig ermöglichen handschriftliche Lösungen gerade in ingenieurwissenschaftlichen oder mathematischen Grundlagenfächern eine sehr

schnelle und intuitive Verschriftlichung eines Lösungsweges mit Formeln, Symbolen, Skizzen und Diagrammen. Nicht jede E-Prüfung muss deshalb nur aus Multiple-Choice-Fragen bestehen und eine Prüfung ist nicht notwendigerweise analog, nur weil ein Prüfling einen Stift in der Hand hat.

Lehrpersonen und Prüfer*innen stehen deshalb in Zukunft vor deutlich mehr Wahl- und Gestaltungsmöglichkeiten von Prüfungsformaten als vor der Pandemie: analog vs. digital, handschriftlich vs. textbasiert, online vs. offline, als Take-Home-Variante oder in Präsenz, überwacht vs. unüberwacht, mündlich vs. schriftlich, Einzel- vs. Gruppenprüfung, automatisch vs. manuell korrigierbar etc. Eine Herausforderung liegt darin, aus der Vielfalt der Möglichkeiten das passendste Format auszuwählen und effizient umzusetzen. Dabei hilft aus eigener Erfahrung auch eine gewisse Informations- und Programmierkompetenz zur Automatisierung wiederkehrender Aufgaben. Die temporeiche Entwicklung von Werkzeugen wie ChatGPT zwingt einen dabei immer wieder kurz innezuhalten, gedanklich einen Schritt zurück zu treten und sich zu überlegen, was die Studierenden eigentlich lernen und welche Kompetenzen sie entwickeln sollen, welche Methode und Prozesse geeignet sind, um das dazu nötige Wissen zu vermitteln, wie die Studierenden gut beim eigenen Kompetenzerwerb unterstützt werden können, wie man den Lernfortschritt und die Kompetenzen gut (formativ und summativ) prüfen kann und welche Fehlanreize sich dabei ergeben.

Literatur

- Magdowski, M. (2018). QR-Code-Papier zum automatischen Korrekturversand. In M. Miglbauer/L. Kieberl/S. Schmid (Hrsg.), Hochschule digital.innovativ | #digiPH Tagungsband zur 1. Online-Tagung (S. 341–348). Norderstedt: Books on Demand GmbH, <https://www.fnma.at/content/download/1529/5759>
- Magdowski, M. (2020). Personalisierbare Aufgaben und anonymer Peer Review in den Grundlagen der Elektrotechnik. In K. Marx/J. Kiesendahl/S. Borukhovich-Weis/P. Glawe/J. Hafer/G. Lisek/M. Kuhnhehn/M. Schöner/U. Gochemann/K. Lisek (Hrsg.), Greifswalder Beiträge zur Hochschullehre – Hochschullehre im digitalen Zeitalter (11. Aufl., S. 75–85). <https://www.uni-greifswald.de/gbzh>
- Magdowski, M. (2021a). Open-Book-Präsenzklausur in den Grundlagen der Elektrotechnik. <https://mathiasmagdowski.wordpress.com/2021/06/28/open-book-prasenzklausur/>
- Magdowski, M. (2021b). Konzeption, Durchführung und Auswertung einer Online-Take-Home-Prüfung in den Grundlagen der Elektrotechnik. e-Prüfungs-Symposium zum Thema „Elektronisches Prüfen in Pandemiezeiten“. <https://e-pruefungs-symposium.de/wp-content/uploads/2021/11/Magdowski-Poster.pdf>
- Magdowski, M./Siegert, I. (2021). Experience with an Online Assessment in a Lecture about Fundamentals of Electrical Engineering. In G. Ubachs/S. Meuleman/A. Antonaci (Hrsg.), Proceedings of the Innovating Higher Education Conference. European Association of Distance Teaching Universities (EADTU). <https://www.slideshare.net/Mathias-Magdowski/experience-with-an-online-assessment-in-a-lecture-about-fundamentals-of-electrical-engineering>

Digitales Lehren und Lernen mit Moodle und JACK

Christoph Olbricht und Melanie Schypula

In der Pandemie ist distantes Lehren und Lernen in den Fokus gerückt. Prüfungen mussten plötzlich wesentlich häufiger mit E-Assessment-Systemen durchgeführt werden, da eine Prüfung vor Ort nicht möglich war. Somit wurden viele Lehrende von außen gezwungen, distanz zu prüfen. Im Zuge dessen wurde die digitale Infrastruktur während der Pandemie schnell ausgebaut, um den hohen Bedarf decken zu können. Wie sieht es aber jetzt nach der Pandemie aus? Hat sich das Verhalten der Lehrenden langfristig geändert oder ist der Großteil zur Papierklausur zurückgekehrt?

An der Universität Duisburg-Essen stehen den Lehrenden das Learning Management System (LMS) Moodle (<https://moodle.uni-due.de>) und das E-Assessment-System JACK (<https://jack-demo.s3.uni-due.de/>) für digitale Prüfungen zur Verfügung. Im Zuge des PITCH-Projekts an der Universität Duisburg-Essen werden nun Moodle und JACK als Prüfungssysteme weiter ausgebaut. Im Folgenden möchten wir den Fragen nachgehen, in welche Richtungen sich Moodle und JACK grundsätzlich entwickeln. Nähern sich beide Systeme an oder bedienen sie unterschiedliche Gruppen? Welchen Anforderungen müssen E-Prüfungssysteme heutzutage gerecht werden? Und wie ist mit der Komplexität der Systeme umzugehen? Welchen Gestaltungsprinzipien folgt die Softwareentwicklung, damit auf der einen Seite möglichst vielen Lehrenden ein einfacher Einstieg in die Lehre mit den Systemen ermöglicht wird und auf der anderen Seite die Systeme eine ausreichende Komplexität besitzen, um den Wünschen der Lehrenden nach komplexen Aufgaben gerecht zu werden?

1 Moodle

Mit dem LMS Moodle an der Universität Duisburg-Essen können auch elektronische Prüfungen durchgeführt werden. Für Prüfungen gibt es einen separaten Prüfungsserver, der über mehr Sicherheit verfügt. Unter anderem werden tägliche Datenbankbackups erstellt, um einen eventuellen Datenverlust gering zu halten, und entsprechend starke Server sorgen für Stabilität bei einer hohen Last. In Moodle werden Dinge, die Lehrenden für Interaktionen zur Verfügung stehen, Aktivität genannt. Für elektronische Prüfungen stehen die zwei Aktivitäten *Aufgabe* und *Test* zur Verfügung, die zunächst etwas genauer beschrieben werden.

1.1 Aktivität Aufgabe

Die Aktivität *Aufgabe* ist als Take-Home-Klausur oder Hausarbeit zu sehen. Die Studierenden laden die Aufgabenstellung herunter und bearbeiten diese außerhalb von Moodle. Haben sie ihre Bearbeitung der *Aufgabe* abgeschlossen, werden die Ergebnisse in Moodle hochgeladen, entweder als Datei oder in einem Textfeld direkt in Moodle. Es muss bei der Bearbeitung keine durchgehende Verbindung zu Moodle bestehen, sodass für die Bearbeitung der *Aufgabe* auch Tage oder Wochen eingeplant werden können.

Dieser offene Aufgabentyp wird nicht von Moodle direkt korrigiert, sondern muss von den Lehrenden manuell bewertet werden. Zur Unterstützung bietet Moodle ein paar Funktionen zum Korrekturworkflow an.

1.2 Aktivität Test

Die Aktivität *Test* ist ein E-Assessment, bestehend aus verschiedenen Aufgabentypen. An der Universität Duisburg-Essen stehen die folgenden Aufgabentypen in verschiedenen Varianten zur Verfügung: Multiple-Choice, Lückentext, Numerisch, Drop-Down, Drag & Drop, Wahr/Falsch und Freitext. Zusätzlich stehen die Aufgabentypen Multiple-Choice und Numerisch auch als Variable Aufgaben zur Verfügung, d. h., die Zahlen in der Aufgabe werden zufällig gezogen. Die verschiedenen Aufgabentypen können auch miteinander kombiniert werden. Bis auf Freitext sind dies alles Aufgabentypen, die direkt vom System bewertet werden können. Bei dieser Aktivität werden die Aufgaben direkt bearbeitet, d. h., die Studierenden bearbeiten die Aufgaben direkt und geben ihre Lösung in Moodle ein.

Die Aufgaben, in Moodle Fragen genannt, werden in einem Moodlekurs innerhalb einer Fragensammlung erstellt und können in diesem Kurs überall genutzt werden. Nach der Erstellung müssen sie einer Aktivität *Test* zugewiesen werden. Die Aufgaben können mit verschiedenen Kategorien geordnet und mit Schlagwörtern versehen werden. Dadurch lassen sie sich gut nach Themen sortieren. Bei der Zusammenstellung der Fragen für eine Prüfung können entweder die Fragen direkt ausgewählt werden oder Moodle zieht die Fragen aus einer bestimmten Menge. Die Reihenfolge der Aufgaben und Antwortoptionen kann wahlweise zufällig sein. Um Aufgaben in einem anderen Moodlekurs zu nutzen, müssen diese zuerst exportiert und anschließend in den anderen Moodlekurs importiert werden.

Die geschlossenen Aufgabentypen werden direkt von Moodle bewertet, sodass für diese Aufgabentypen sofort die Ergebnisse zur Verfügung stehen. Die offenen Aufgabentypen müssen auch hier manuell von den Lehrenden bewertet werden.

1.3 Prüfen mit Moodle

Um elektronische Prüfungen an der Universität Duisburg-Essen durchzuführen, sollen die Lehrenden auf dem Moodle-Prüfungsserver arbeiten. Sie können dabei entscheiden, ob sie in Moodle nur die Organisation der Prüfung machen, wie z. B. die Liste der zugelassenen Teilnehmer*innen verwalten und über einen Link die Studierenden z. B. nach JACK schicken, wo die Prüfungsaufgaben liegen, oder ob sie die Prüfung auch in Moodle mit der entsprechenden Aktivität durchführen möchten. Durch das Hochladen der Teilnehmer*innenliste der Prüfung durch die Prüfer*innen können nur die Studierende den Moodlekurs betreten, die auch für die Prüfung angemeldet und zugelassen sind.

Die Authentifizierung findet über die Anmeldung in Moodle mittels der Userkennung statt. Unterschiedliche Klausuren lassen sich über verschiedene Moodlekurse abbilden. Nach einer Prüfung können die Ergebnisse im Bereich Bewertungen eingesehen werden. Am Ende kann auch die Einsicht der Prüfung in Moodle stattfinden. Diese Funktionen stehen einem Lehrenden auch zur Verfügung, wenn Moodle nur zur Organisation genutzt und JACK als externes Tool eingebunden wird. Hier ist der einzige Unterschied, dass die Einsicht in JACK direkt stattfinden muss.

Für die Prüfung in Moodle stehen verschiedene Einstellungen zur Verfügung. Hier ist zu beachten, dass die beiden Aktivitäten *Aufgabe* und *Test* teilweise dieselben Einstellungen haben, aber auch Unterschiede aufweisen. Für jede Prüfung kann ein Zeitintervall eingestellt werden, in welchem die Aktivität für die Studierenden einsehbar und nutzbar ist. Zusätzlich können Lehrende angeben, wie lange und unter welchen Bedingungen nach Ablauf der Zeit noch Lösungen eingereicht werden können. Es gibt verschiedene Möglichkeiten, die Bewertungen festzulegen, wie die Bestehensgrenze, die Anzahl der Versuche einer Aufgabe und welcher Versuch gezählt werden soll. Auch das Layout lässt sich einstellen. So kann ein*e Lehrende*r jede Frage auf eine eigene Seite stellen und die Studierenden können selbstständig durch die Prüfung navigieren. Zusätzlich kann eingestellt werden, wann die Studierenden Feedback für ihre Lösung angezeigt bekommen.

Dadurch, dass die Aktivität *Aufgabe* ein offener Aufgabentyp ist, gibt es ein paar mehr Möglichkeiten bei dem Korrekturmanagement.

2 JACK

Das E-Assessment-System JACK wird seit 2006 durch die Arbeitsgruppe S3 an der UDE entwickelt. Es kann sowohl zu Übungszwecken als auch zur Durchführung von Klausuren verwendet werden. Die Basis stellt hierbei immer die Aufgabe dar. Mehrere Aufgaben lassen sich zu einem Kurs zusammenfassen. Dieser Kurs kann den Studierenden zu Übungszwecken zur Verfügung gestellt oder

als Klausur verwendet werden. Die unterschiedlichen Einstellungsmöglichkeiten für beide Optionen werden in Kursangeboten verwaltet. Die Stärke von JACK liegt in den Möglichkeiten zur Variabilisierung von Aufgaben und Kursen. Eine einzelne Aufgabe kann mittels Variablen eine große Menge an Varianten bereitstellen. Somit kann ein*e Studierende*r dieselbe Aufgabe mehrfach bearbeiten, um das gegebene Thema zu üben und zu festigen. Alternativ kann einer Menge an Studierenden eine jeweils individuelle Klausur vorgelegt werden. Dies kann genauso durch das Ziehen einer zufälligen Teilmenge aus einem Aufgabenpool erreicht werden. Der Begriff Variablen ist hier nicht auf den mathematischen Raum beschränkt, vielmehr geht es um variable Teile einer Aufgabenstellung.

Eine Aufgabe kann aus mehreren Teilen, in JACK Aufgabenteile genannt, bestehen. Jeder Aufgabenteil hat einen Typ, der die Art des Aufgabenteils bestimmt. Aktuell werden in JACK die allgemeinen Aufgabenteiltypen Multiple-Choice, Multiple-Choice Tabelle, Freitext, Dateupload sowie Drop-Down-Menu & Fill-In-Felder angeboten. Zudem gibt es spezifische Aufgabentypen für die Chemie und verschiedene Programmiersprachen. Bis auf Freitext und Dateupload können alle Aufgabenteiltypen mit Prüffregeln versehen werden, durch welche eine detaillierte automatische Bewertung durch JACK möglich ist. Die Aufgabenteile können nicht nur linear miteinander verknüpft und kombiniert werden. Vielmehr kann die Wahl des nächsten Aufgabenteils anhand einer Variablenbelegung oder dem Korrektheitsgrad einer Lösung entschieden werden. Somit kann der nächste Lerninhalt automatisch individuell an den Wissensstand des Studierenden angepasst werden.

Im Folgenden soll der Lebenszyklus einer Aufgabe in JACK exemplarisch durchlaufen werden, um ein Bild vom System JACK zu zeichnen.

2.1 Aufgabenentwicklung

Nach dem Login in JACK mittels Unikennung oder E-Mail soll eine neue Aufgabe im Bereich Mathematik zum Thema Kurvendiskussion erstellt werden. Statt einer spezifischen Funktion werden allerdings Platzhalter für die einzelnen Faktoren der quadratischen Gleichung definiert, sodass theoretisch jede beliebige quadratische Funktion zufällig erzeugt werden könnte. Da hierbei zu komplizierte Rechenwege entstehen könnten, wird der Zahlenraum entsprechend eingeschränkt. Die Möglichkeit, LaTeX zu verwenden, macht es möglich, auch komplexere Formeln ansprechend darzustellen.

Als erste Teilaufgabe sollen die Studierenden einige grundlegende Aussagen zur gegebenen Funktion bzw. dessen Graphen treffen. Daher wird ein Aufgabenteiltyp mit Drop-Down-Menüs, den korrekten Antwortoptionen sowie passenden Distraktoren erstellt. Hierbei werden die korrekten Antwortoptionen je nach erstellter Variante korrekt berechnet und die Distraktoren aus einer gegebenen Liste von Möglichkeiten gezogen.

Um den Studierenden einen besseren Lernerfolg zu ermöglichen, werden sogenannte Feedbackregeln definiert, mit deren Hilfe je nach Lösung der Studierenden eine Rückmeldung gegeben wird und sie ggf. in einen zusätzlichen Aufgabenteil geleitet werden, wo sie gezielte Informationen zu ihrem Fehler erhalten und sich erneut an einer ähnlichen Fragestellung, diesmal als Multiple-Choice gestaltet, versuchen können.

Ist ein Fehler bei einer der Feedbackregeln gemacht worden, so kann im JACK-Wiki (<https://wiki.uni-due.de/jack>) nachgeschlagen werden, wie die Regel korrekt zu formulieren ist. Das JACK-Wiki wurde mit einem ausführlichen Benutzerhandbuch versehen. Somit wird Lehrenden ein guter Einstiegspunkt in die Arbeit mit JACK geboten. Lehrende finden im Wiki weiterhin Informationen zu allen Features von JACK sowie eine umfassende Liste der in JACK nutzbaren Evaluatorkfunktionen zur Erstellung von detailliertem Feedback der studentischen Lösungen.

Als nächster Aufgabenteil wird ein Fill-In-Typ erstellt. Als Aufgabenstellung gilt es, die Extremstellen der Funktion zu berechnen. Die korrekten Werte für jede mögliche Variante berechnet hierbei JACK. Zudem wird weiteres Feedback für den Fall erstellt, dass ein*e Studierende*r versehentlich die Nullstellen und nicht Extremstellen berechnet hat, was einen üblichen Fehlerfall darstellt. Zusätzlich werden Feedbackregeln für Teillösungen definiert, um feingranulare Teilpunkte zu vergeben und fehlerspezifisches Feedback zu liefern, welches den Studierenden ihre gemachten Fehler erläutert.

Da eventuell noch Schwierigkeiten beim Lösungsweg bestehen könnten, kommen Hinweistexte hinzu, die auf Klick das korrekte Vorgehen beschreiben, dafür aber einen kleinen Punktabzug mit sich bringen. Wenn diese Aufgabe später in der Klausur verwendet wird, können diese Hinweise für die gesamte Klausur ausgeschaltet werden.

Im nächsten Aufgabenteil vom Typ Freitext wird jeweils eines von zehn verschiedenen Bildern eingebunden. Die Studierenden sollen in eigenen Worten den Sachverhalt des Bilds beschreiben und erläutern. Die Bildinformationen werden hierbei direkt für die Studierenden anonymisiert, sodass diese keine Rückschlüsse über den Dateinamen auf die Lösung ziehen können. Da dieser Aufgabenteiltyp nicht automatisch bewertet werden kann, wird er für den Übungsbetrieb aus der Aufgabe herausgelassen und lediglich für die bevorstehende Klausur eingebunden.

Abschließend wird ein Aufgabenteil mit dem Typ Dateiapload erstellt, um in der Klausur den von den Studierenden gezeichneten Graphen auf Korrektheit überprüfen zu können. In diesem Typ lässt sich klar definieren, welche Dateien von den Studierenden hochzuladen sind, also sowohl Menge, Typ als auch Dateinamen. Somit ist man sowohl auf das Szenario vorbereitet, dass die Studierenden eine händische Lösung zeichnen, als auch, mittels eines anderen Programms die Zeichnung digital erstellen.

Zur besseren Übersicht kann die Aufgabe mit Tags versehen werden, um sie in eine beliebige Menge von Kategorien einzusortieren. Somit können Aufgaben

gleichzeitig mit dem behandelten Thema, den Punkten, dem Schwierigkeitsgrad oder beliebigen weiteren Merkmalen markiert werden. Damit ist auch bei größeren Aufgabenmengen die schnelle Suche und Auswahl von spezifischen Aufgaben möglich.

2.2 Prüfungsvorbereitung und -durchführung

Soll nun eine Prüfung erstellt werden, bietet JACKs Zugriffsverwaltung und Rechtekonzept eine feingranulare Steuerung, wem welche Inhalte zur Verfügung stehen.

Im ersten Schritt werden alle eigenen erstellten Aufgaben sowie der zugehörige Kurs, der die Aufgaben gruppiert und definiert, wie viele Punkte welche Aufgabe insgesamt gibt, in einen Ordner abgelegt. Diesen Ordner kann ich nun für beliebige Personen freigeben.

Zunächst soll eine studentische Hilfskraft die Aufgaben testen, jedoch nicht verändern können, weshalb ihr lediglich Leserecht erteilt wird. Im Anschluss kann ein*e wissenschaftliche*r Mitarbeiter*in die Korrekturen einpflegen und eigene Änderungen vornehmen, da ihr zusätzlich das Leserecht erteilt wurde. Sollte die Person bei der Korrektur helfen, wird zudem noch das Bewertungsrecht gegeben.

Wird mit einer anderen Lehrkraft kollaboriert, erstellt man einen weiteren Ordner, auf welchem der Lehrkraft das volle Verwaltungsrecht erteilt wird, sodass die Lehrkraft den Ordner verwalten kann, als wäre es ihr eigener Ordner. So kann sie weitere Personen für ihren Klausurteil hinzufügen, falls gewünscht. Sollte die Lehrkraft bereits Aufgaben auf einem anderen JACK-Server erstellt haben, kann sie diese dort problemlos exportieren und direkt in den bereitgestellten Ordner importieren.

Nachdem alle Aufgaben in einem Kurs zusammengefügt wurden und die Punkte jeder Aufgabe angepasst sind, kann der Kurs zu einem Kursangebot hinzufügen werden.

Das Kursangebot stellt eine Instanz des Kurses, also der darin enthaltenen Aufgaben dar. Somit können die Aufgaben zunächst für eine Klausur verwendet, und im nächsten Semester den Studierenden mit anderen Einstellungen während des Semesters zur Übung zur Verfügung gestellt werden, ohne die Aufgaben noch einmal anfassen zu müssen.

Für die Klausur wird definiert, dass das Kursangebot nur von für die Klausur angemeldeten Studierenden gesehen und betreten werden darf. Während der Klausur werden den Studierenden keine Hinweise und kein Feedback gegeben. Mittels eines Zeitlimits zur Bearbeitung erhalten alle Studierenden dieselbe Zeit, auch wenn sie nicht in derselben Minute beginnen. Mittels eines Bearbeitungsstarts und -endes kann gewährleistet werden, dass selbst distante Klausuren nur im definierten Zeitraum geschrieben werden können.

2.3 Korrekturmanagement und Einsicht

Die Korrektur der meisten Aufgabenteiltypen geschieht automatisch durch die definierten Feedbackregeln von JACK. Lediglich Freitext und Dateiupload müssen von den Lehrenden manuell korrigiert werden. Dies ist bisher noch etwas umständlich, allerdings wurden im PITCH-Projekt mit Hilfe aller beteiligten Lehrstühle die Anforderungen an ein Korrekturmanagement erhoben. Perspektivisch soll es die Möglichkeit geben, direkt in JACK die Lösungen von Studierenden effizient zu sichten und manuell zu korrigieren. Dabei sollen die Lehrenden die Möglichkeit haben, die Klausurkorrektur nach ihren Wünschen und Vorstellungen organisieren zu können. Hierbei sollen Möglichkeiten geboten werden, die Menge an Klausuren nach selbst gewählten Kriterien, etwa nach Aufgaben, Varianten oder einfach einer Teilmenge der Lösungen, den jeweiligen Korrigierenden zuzuordnen. Den Korrigierenden soll es ermöglicht werden, Kommentare für die Studierenden sowie interne Kommentare für andere Korrigierende zu hinterlassen. Die Prüfungen sollen anonymisiert oder pseudonymisiert werden, damit die Korrigierenden eine Lösung nicht Studierenden zuordnen können.

Im Anschluss an den Korrekturprozess können die Studierenden ihre Klausur direkt in JACK einsehen. Jetzt wird ihnen das erstellte Feedback automatisch angezeigt, sodass sie direkt ihre Fehler nachvollziehen können.

3 Entwicklung elektronischer Prüfungen in den letzten Jahren

An der Universität Duisburg-Essen wurde Moodle zunächst als LMS genutzt: Lehrende haben Material für die Studierenden zur Verfügung gestellt, Informationen an die Studierenden weitergegeben und man konnte über Foren in den Austausch miteinander treten (Lehrende und Lernende). Als dann in der Pandemie Präsenzlehre teilweise gar nicht mehr möglich war oder nur unter sehr strengen Voraussetzungen, mussten neue Möglichkeiten geschaffen werden, die Studierenden online zu prüfen, und das Prüfen mit Moodle rückte stark in den Fokus. In kürzester Zeit wurden die Assessment-Möglichkeiten intensiv ausgebaut, sodass auch mit Moodle Prüfungen abgehalten werden konnten. Diese Möglichkeit wurde dann auch während der Pandemie stark genutzt. Nach der Pandemie ist die Nutzung wieder zurückgegangen. Das liegt zum einen daran, dass in der Pandemie spezielle Prüfungsordnungen galten und zum anderen, dass einige in ihr altes Muster zurück gegangen sind, aber eben nicht alle. Bei einigen ist die Bereitschaft gestiegen, sich nach der Pandemie intensiver mit elektronischen Prüfungen auseinanderzusetzen, wie man am PITCH-Projekt erkennen kann. Einige der Projektteilnehmer*innen haben vorher noch keine elektronischen Prüfungen abgehalten. Durch die Möglichkeiten in diesem Projekt steigt auch die Anzahl der durchgeführten E-Assessments. Es lässt sich hier also ein langfristiger positiver Effekt erkennen.

Die Universität Duisburg-Essen besitzt mit Moodle und JACK zwei E-Assessment-Systeme. Dadurch stellt sich die Frage, ob es sinnvoll ist, beide Prüfungssysteme zu betreiben, weil die Systeme unterschiedliche Nutzungsszenarien bedienen oder sie letztendlich redundant sind? Beide Systeme bieten allgemein gesprochen E-Assessment an. Allerdings richten sie sich an unterschiedliche Benutzer*innen. Mittels Moodle können einfache Aufgaben erstellt und Prüfungen durchgeführt werden. JACK hingegen bietet den Nutzer*innen wesentlich komplexere Möglichkeiten an, Aufgaben zu erstellen, diese variabel zu gestalten, sodass automatisch eine Vielzahl von Varianten einer Aufgabe erstellt werden können, sowie Teilaufgaben zu verknüpfen, um den Lernfluss der Studierenden zu steuern und auf Eingaben der Studierenden einzugehen. Die beiden Systeme nähern sich also nicht an, sondern bedienen unterschiedliche Gruppen.

4 Praxisbeispiel Vorlesung Programmierung

Ein Praxisbeispiel ist die Vorlesung *Programmierung* am Campus Essen der Universität Duisburg-Essen. Vor der Pandemie war die Prüfung am Ende des Semesters immer eine Stift-Papierklausur, die vor Ort in den Räumlichkeiten der Universität geschrieben wurde. In der Zeit, während in der Pandemie Klausuren nicht vor Ort geschrieben werden konnten, musste sich der zuständige Lehrstuhl eine Alternative überlegen, wie die Prüfung in Distanz sinnvoll durchgeführt werden kann. Für diese Veranstaltung wurde das E-Assessment-System JACK als Lösung gewählt. Die Studierenden konnten die Prüfung von zu Hause aus durchführen und ihre Lösungen in JACK einreichen und korrigieren lassen. Das hat sehr gut funktioniert, sodass auch nach der Pandemie die Studierenden ihre Prüfung digital absolvierten. Der Lehrstuhl ist lediglich dazu übergegangen, die Klausur unter Aufsicht in sogenannten PC-Halls (Räume, bei denen jeder Platz über einen eigenen PC verfügt) an der Universität zu schreiben, um möglichen Täuschungsversuchen vorzubeugen. Hier gab es durch die Pandemie ein verändertes Verhalten, welches einen Effekt über die Pandemie hinaus hatte. Es war schon lange notwendig, die Studierenden nicht mehr auf Papier programmieren zu lassen, da dies nicht widerspiegelt, welche Kompetenzen man in einer Programmierungsvorlesung lehren und abfragen möchte. An dieser Stelle sei noch angemerkt, dass die Studierenden zudem eine große Menge an Übungsaufgaben durch JACK zur Verfügung hatten, die sie während des Semesters bearbeiten konnten. Somit war dieser Teil der Lehre bereits vor der Pandemie distant und konnte ohne weiteren Aufwand auch in den Pandemie Jahren durch die Studierenden genutzt werden, um sich auf die Prüfung vorzubereiten. An diesem Beispiel lässt sich erkennen, dass die Pandemie einen positiven Effekt auf die Nutzung von E-Assessment-Systemen in der Lehre und zur Prüfung von Studierenden hatte.

5 Fazit

Das elektronische Prüfen hat sich nach der Pandemie etabliert. Einige Lehrende sind auch nach der Pandemie dageblieben, andere nehmen sich nun die Zeit, in Ruhe sinnvolle Konzepte zu erarbeiten, um elektronische Prüfungen abzuhalten. Ein immer größer werdender Kund*innenstamm bringt jedoch auch immer neue Anforderungen an die Systeme mit sich. Dadurch müssen diese stetig weiterentwickelt werden, um den Anforderungen zu genügen.

Besonderes Augenmerk ist darauf zu legen, dass ein E-Assessment-System einfach und intuitiv zu bedienen ist und dabei trotzdem viele verschiedene Funktionen besitzt. Dies gelingt beiden Systemen durch eine aufgeräumte und gut strukturierte Benutzeroberfläche. Die in JACK gegebene Einstellungsvielfalt führt allerdings bei neuen Benutzer*innen teilweise zu Überforderung.

Ein weiterer wichtiger Punkt ist eine gute Dokumentation des Systems. Somit können sich Nutzer*innen viel schneller selbst helfen und müssen nicht jedes Mal den Support fragen. Moodle bietet einige Handreichungen und JACK bietet mit dem oben genannten Wiki eine gute Dokumentation, die stetig verbessert wird. Zudem sind an vielen Stellen im UI von Moodle und JACK Infotexte eingebettet, welche bei Bedarf aufgerufen werden können.

Lehrende kennen häufig nicht alle Funktionen eines komplexen E-Assessment-Systems. Auch hier hilft eine gute Dokumentation, sie von den Möglichkeiten in Kenntnis zu setzen. Moodle bietet hier zudem an der UDE eine Sprechstunde an, während JACK den Kund*innen auf sie zugeschnittene Workshops anbietet, wo den Kund*innen die für sie nötigen komplexen Funktionen erläutert werden.

Ein weiterer Aspekt ist, Aufgaben effizient erstellen zu können. Hierbei gilt die Anforderung, schnell eine Aufgabe erstellen zu können und später, je nach Wunsch, weitere Komplexität hinzuzufügen. Hier bietet Moodle die Möglichkeit, komfortabel einfache Aufgaben und ganze Klausuren zu erstellen, gelangt jedoch schnell an die Grenze seiner Möglichkeiten. Mittels JACK lassen sich einfache Aufgaben zwar ähnlich komfortabel umsetzen, allerdings sorgen die weiteren Einstellungsmöglichkeiten und Konzepte hier schneller für Verunsicherung. Dafür können auch, wie oben beschrieben, komplexe Aufgaben mit mehreren Aufgabenteilen, Verzweigungen und Varianten erstellt werden.

Auf dem Weg zur Innovation: Digitale Präsenzklausuren am Beispiel der Englischen Sprachwissenschaft

Frauke Milne, Birte Bös und Jana Eismann

1 Einleitung

Modulprüfungen sind nicht nur ein entscheidendes Instrument der Leistungsevaluation, sondern auch Schlüsselkomponenten für die akademische und persönliche Weiterentwicklung der Studierenden. Allerdings bleibt im universitären Alltag der Prüfenden zumeist wenig Zeit für eine intensive konzeptuelle Auseinandersetzung mit Prüfungsformen und -inhalten. Das PITCH-Projekt eröffnet hier zeitliche und personelle Spielräume, um die Potenziale von digitalen Prüfungen im Spannungsfeld von technologischem Fortschritt, rechtlichen Rahmenbedingungen und einer heterogenen Studierendenschaft auszuloten.

Aus Sicht des Projektteams des Instituts für Anglophone Studien stellt sich dabei besonders die Umsetzung der Klausur zum Einführungsmodul „Introduction to Linguistics“ (Ling1) als zentrale Herausforderung dar. Das Modul macht die Studierenden mit den Grundlagen der englischen Sprachwissenschaft vertraut und bildet insofern einen wesentlichen Threshold, den es für die Studierenden zu bewältigen gilt.

Der Studienbeginn ist als Statuspassage für viele junge Erwachsene eine komplexe Bewährungsprobe (Friebertshäuser, 2004; Kossack/Lehmann, 2012). Die Linguistik ist zudem eine Wissenschaft, mit der viele Studierende bis dahin zumeist wenig Berührungspunkte hatten, welche darüber hinaus noch seltener fachterminologisch verankert wurde. Tatsächlich stellt sich in Gesprächen mit den Studierenden heraus, dass sie nicht selten als einzigen Anknüpfungspunkt den Grammatikunterricht nennen, der oft emotional negativ behaftet ist. Außerdem erschließt sich vielen Lehramtsstudierenden die Bedeutung der Sprachwissenschaften für den späteren Beruf zunächst häufig noch nicht.

In der Modul- und Klausurgestaltung geht es also vor allem um die Sicherung eines grundlegenden Verständnisses der Disziplin Linguistik als solcher mit ihren Forschungsgegenständen, Teilbereichen und Grundbegriffen. Darüber hinaus steht im Besonderen die englische Sprache im Fokus, die aufgrund ihrer ereignisreichen Geschichte durch großen sprachlichen Wandel geprägt ist. Schlussendlich geht es in der Lehre wie in der Prüfung darum, die Omnipräsenz sprachlicher Phänomene und die Bedeutung der Sprachwissenschaft in deren Erfassung herauszuarbeiten und positiv zu konnotieren.

Der folgende Beitrag beleuchtet anhand dieser Modulprüfung exemplarisch die Rahmenbedingungen, Herausforderungen und Chancen digitaler Klausuren und widmet sich dabei vor allem verschiedenen Facetten der Aufgabengestaltung.

2 Rahmenbedingungen der Klausur

2.1 Modulinhalte und -ziele

Das Modul Ling1 ist für alle BA-Studiengänge der Anglophonen Studien in den ersten zwei Fachsemestern angesiedelt und umfasst die Vorlesung *Introduction to Linguistics I: Synchronic and Descriptive Perspectives* (1 SWS) inklusive dazugehöriger obligatorischer Übung (1 SWS) sowie, im Folgesemester, die Vorlesung *Introduction to Linguistics II: Diachronic and Applied Perspectives* (2 SWS). Auch für die zweite Vorlesung wird ein Tutorium angeboten, das jedoch nicht modular verankert ist. Die Inhalte der beiden Veranstaltungen werden in einer Klausur mit zwei gleichwertigen Teilen geprüft, die gemäß Regelstudienplan am Ende des zweiten Fachsemesters absolviert wird.

Das aktuelle Modulhandbuch (Wintersemester 2018/19) für Bachelor Lehramt an Gymnasium und Gesamtschule, Studienfach Englisch (Universität Duisburg-Essen, 2024, S. 12) definiert für Ling1 folgende zu erreichende Lernergebnisse und Kompetenzen:

Die Studierenden haben einen Überblick über die Grundlagen und wichtigsten Teilgebiete der englischen Linguistik. Sie sind sowohl mit der synchron-deskriptiven Sprachbeschreibung als auch der diachron-angewandten Perspektive der Sprachwissenschaft vertraut. Sie erwerben theoretische und methodische Grundkenntnisse sowie die Fähigkeit, auf der Basis dieser Kenntnisse, sprachliche Phänomene zu beschreiben und zu analysieren.¹

Da es sich für den Großteil der Studierenden um die erste Begegnung mit der Sprachwissenschaft handelt, liegen die Kompetenzen eher in der unteren Hälfte der Lehrzieltaxonomie (Bloom, 1972). Als Basis dient das *Wissen* über linguistische Fachbegriffe und Phänomene, erweitert durch deren *Verständnis*, um diese Termini und Erscheinungen erklären zu können. Vor allem in den Übungen werden die Studierenden zur *Anwendung* des Gelernten, etwa zum Entdecken eigener Beispiele und zu metasprachlichen Beobachtungen, ermutigt. Zudem werden im gezielt definierten Rahmen auch erste *Analysen* vorgenommen.

1 Dieselbe Kompetenzdefinition findet sich ebenfalls in den Modulhandbüchern der anderen Lehramtsstudiengänge mit Bachelor-Abschluss (ab Wintersemester 2018/19), sowie der 2-Fach Bachelor Anglophone Studien (ab Wintersemester 2018/19).

2.2 Organisation und Umsetzung der Klausur

Die Ling1-Klausur findet jedes Semester statt. Der Regelstudienplan sieht die Modulprüfung am Ende des zweiten Fachsemesters vor, wobei es jedoch gelegentlich auch vorkommt, dass Studierende erst im höheren Fachsemester an der Klausur teilnehmen. Das erfolgreich bestandene Modul Ling1 ist allerdings für alle Studierenden Voraussetzung für die Teilnahme an weiteren Prüfungsleistungen im Bereich der Englischen Sprachwissenschaft. Aufgrund des regulären Beginns der BA-Studiengänge jeweils zum Wintersemester ist die durchschnittliche Teilnehmendenzahl für die Ling1-Prüfung im Sommersemester deutlich höher als im Wintersemester (vgl. Tab. 1). Die dreistelligen Teilnehmendenzahlen haben Auswirkungen auf die praktische Umsetzung der Klausur (Raum- und Personalbedarf) und deren Auswertung (Korrektur, Evaluation und Klausureinsichten), die bereits bei der Prüfungsgestaltung berücksichtigt werden müssen.

Semester	Teilnehmendenzahlen	Prüfungsform der Klausur	Besondere Bedingungen
WiSe 2019/20	90	Digitale Präsenzprüfung (L-Plus)	Keine
SoSe 2020	420	Pen and Paper	Pandemie
WiSe 2020/21	214	Digitale Distanzprüfung (L-Plus)	Pandemie
SoSe 2021	441	Digitale Distanzprüfung (L-Plus)	Pandemie
WiSe 2021/22	137	Digitale Distanzprüfung (L-Plus)	Pandemie
SoSe 2022	211	Digitale Präsenzprüfung (L-Plus)	Pandemie
WiSe 2022/23	60	Pen and Paper	Cyberangriff
SoSe 2023	201	Digitale Präsenzprüfung (Moodle)	Cyberangriff

Tabelle 1: Klausur Ling1: Teilnehmendenzahlen, Prüfungsformen, Bedingungen (Quelle: Eigene Darstellung)

Einen weiteren Faktor in der Gestaltung der Prüfung stellt die Vielzahl zum Teil wechselnder Akteure dar, da die Vorlesungen und Übungen des Moduls von verschiedenen Lehrenden des Bereichs Englische Sprachwissenschaft angeboten werden. Entsprechend sind an der Erstellung, Testung, Durchführung und Auswertung der Klausur bis zu 10 Lehrende und Mitarbeitende beteiligt, was einen erhöhten Koordinations- und Kommunikationsaufwand bedeutet.

Zusätzlich zu den regulären, durch die Prüfungsordnungen und die Modulgestaltung gesetzten Konditionen hielten die letzten Semester weitere besondere Herausforderungen bereit. Zum einen erforderten die ab dem Frühjahr 2020 einsetzenden Einschränkungen der Covid-Pandemie nicht nur ein Ausweichen der universitären Lehre auf Onlinestrukturen innerhalb kürzester Zeit, sondern auch

die Neuorganisation der Präsenzprüfungen. Zum anderen forcierte im Herbst 2022 ein massiver Cyberangriff auf die Universität Duisburg-Essen weitere Änderungen in der Umsetzung der Klausur. Wie Tab. 1 zeigt, wurden deshalb sowohl wiederholte Wechsel der Prüfungsform als auch ein Plattformwechsel notwendig.

3 Gestaltung digitaler Klausuren

3.1 Klausurformate im Vergleich

Unter Standardrahmenbedingungen handelt es sich bei der Ling1-Prüfung um eine digitale Präsenzklausur. Während die Veränderungen externer Rahmenbedingungen (Pandemie, Cyberattacke, Plattformwechsel) von Lehrenden und Studierenden gleichermaßen ein hohes Maß an Flexibilität erforderten, haben sie aber auch wichtige Erfahrungen erbracht und Denkprozesse zur Prüfungsgestaltung und -umsetzung angestoßen. Tab. 2 fasst Vorteile und Herausforderungen der erprobten Klausurformen zusammen und blickt dabei auf Aspekte wie Planbarkeit, Raum- und Technikbedarf, Personal- und Korrekturaufwand, rechtliche Rahmenbedingungen, aber auch den Punkt Chancengleichheit. Während diese angesichts der begrenzten Seitenzahl hier nicht detailliert betrachtet werden können, soll der Fokus im Folgenden vor allem auf dem Aufgabenspektrum liegen, wobei insbesondere die digitale Präsenz als Standardoption in den Blick genommen wird.

Standard-konditionen	Pandemie		Cyberangriff
<i>Format</i>	Digitale Präsenzklausur	Digitale Distanzprüfung	Präsenzklausur (Pen and Paper)
<i>Vorteile des Prüfungsformats</i>	<ul style="list-style-type: none"> – Hohe Flexibilität bei veränderter TN-Zahl – Teilautomatisierte Korrektur – Hohe Objektivität – Vertrautheit Studierender mit digitalen Formaten – Höhere Chancengleichheit (Durchführung) – Klarer rechtlicher Rahmen 	<ul style="list-style-type: none"> – Hohe Flexibilität bei veränderter TN-Zahl – Teilautomatisierte Korrektur – Hohe Objektivität – Vertrautheit Studierender mit digitalen Formaten – Kein Raumbedarf – Geringer Personalaufwand (Aufsicht) 	<ul style="list-style-type: none"> – Kein technisches Personal nötig – Höhere Chancengleichheit (Durchführung) – Klarer rechtlicher Rahmen

Standardkonditionen	Pandemie		Cyberangriff
<i>Herausforderungen des Prüfungsformats</i>	<ul style="list-style-type: none"> – Hoher Personalaufwand (Aufsicht, technischer Support) – Hoher Bedarf an Räumen mit entsprechender techn. Ausstattung 	<ul style="list-style-type: none"> – Technischer Support nötig – Geringere Chancengleichheit (techn. Ausstattung, räumliche und soziale Rahmenbedingungen) – Hohes Risiko von Täuschungsversuchen u. erschwerte Nachweisbarkeit – Unklare rechtliche Rahmenbedingungen 	<ul style="list-style-type: none"> – Hoher Personalaufwand (Aufsicht) – Hoher Raumbedarf – Eingeschränkte Flexibilität bei veränderter Teilnehmerzahl – Hoher Korrekturaufwand – Potenziell geringere Objektivität – Geringere Nachhaltigkeit (materielle Ressourcen) – Entgegen Präferenzen und Gewohnheiten von digital Natives
<i>Auswirkungen auf das Aufgabenspektrum</i>	<ul style="list-style-type: none"> – Aufgabentypen durch techn. Affordanzen (spez. Prüfungsplattformen) vorgegeben – Aufgabeninhalte: Breiteres Spektrum an Aufgaben möglich – Gruppenvariation leichter umsetzbar, aber optional 	<ul style="list-style-type: none"> – Aufgabentypen durch techn. Affordanzen (spez. Prüfungsplattformen) vorgegeben – Aufgabeninhalte: Fokus auf anwendungsorientierten Aufgaben – Gruppenvariation leichter umsetzbar und nötig 	<ul style="list-style-type: none"> – Aufgabentypen flexibler umsetzbar (z. B. Modelle und Grafiken zeichnen) – Aufgabeninhalte: Breiteres Spektrum an Aufgaben möglich – Gruppenvariation schwerer umsetzbar, je nach räumlichen Gegebenheiten nötig

Tabelle 2: Vorteile und Herausforderungen erprobter Klausurformate (Quelle: Eigene Darstellung)

3.2 Auf dem Weg zur Innovation

Der Wechsel zu digitalen Prüfungsformaten hält für Lehrende sowohl unter alltäglichen Konditionen als auch bei außergewöhnlichen Rahmenbedingungen Herausforderungen bereit. Zu Anfang war der Fokus des Projekts zunächst auf die Verbesserung der Organisationsstrukturen der Prüfung gerichtet. Der eigentliche Schwerpunkt im PITCH-Projekt liegt jedoch vor allem auf den Themen

Chancengleichheit und Innovation. Handke/Schäfer (2012, S. 157) betonen, dass die Sinnhaftigkeit und Akzeptanz von E-Assessment vor allem davon abhängen, ob nicht nur der organisatorische Aufwand reduziert, sondern tatsächlich eine Qualitätsverbesserung erzielt werden kann.

Um die Ling1-Klausur auf verschiedenen Ebenen nicht nur zu digitalisieren, sondern zu innovieren, ergaben sich im Rahmen des PITCH-Projekts folgende Ziele, die in den folgenden Abschnitten genauer betrachtet werden: (1) Kontinuierliche Erweiterung des Aufgabenpools, (2) Verbesserung der Variation der Aufgabentypen, (3) Erhöhung des Aufgabenanteils mit Anwendungs- und Transferleistungen, (4) Steigerung der Authentizität und (5) Einbezug studentischen Feedbacks.

3.2.1 Erweiterung des Aufgabenpools

Unabhängig vom Klausurformat empfiehlt sich für Klausuren ein möglichst großer Aufgabenpool. Dieser ermöglicht eine höhere Zufallsauswahl, sowohl was den Aufgabenwechsel von Semester zu Semester als auch inhaltliche Gruppenvariationen innerhalb einer Klausur betrifft. Darüber hinaus reduziert ein großer Aufgabenpool die Gefahr, dass Prüfungsfragen bekannt und unter Studierenden ausgetauscht werden (Handke/Schäfer, 2012, S. 161). Im Rahmen des Projekts wird der Aufgabenpool für beide Klausurteile kontinuierlich erweitert und ist von 60 auf nunmehr 194 Aufgaben gewachsen, Gruppenvariationen inbegriffen.

Nach der Systematisierung des Aufgabenkatalogs und der Bereinigung inhaltlicher Überschneidungen sind die jeweiligen Aufgabentypen sowohl im Umfang als auch inhaltlich variabler gestaltet worden. So kann z. B. über eine variierende Anzahl zu klassifizierender Beispiele für bestimmte Aufgaben die Klausurzusammenstellung flexibler und die Punkteverteilung auf die einzelnen Teilbereiche genauer erfolgen, was wiederum die Validität der Klausur erhöht. Der Aufgabenkatalog wird zudem laufend durch die Ergänzung weiterer Beispiele vergrößert, welche die Anzahl möglicher Gruppenvarianten erhöhen. Erleichtert wird die Variation durch die inhaltliche Eingrenzung der einzelnen Aufgaben, um sie im Sinne eines Mix-and-Match-Systems zu kombinieren, ohne dass es zu inhaltlichen Überschneidungen kommt.

Besonders für Anwendungsaufgaben, in denen mit sprachlichem Beispielmaterial (etwa Wörtern, Satzbeispielen oder Kurzdialogen) gearbeitet wird, ist die Erweiterung des Aufgabenkatalogs leicht möglich. Mittlerweile wurden aber z. B. auch verschiedene Varianten von Grafiken ergänzt, die der Exemplifizierung bestimmter Phänomene und Theorien dienen. So wird etwa die semantische Prototypentheorie in der Linguistik oft anhand der Kategorie ‚Vögel‘ illustriert, und eine entsprechende Grafik findet sich auch im Reader zur Veranstaltung. Im Rahmen des Projekts wurden Alternativversionen dieser Grafik für weitere Kategorien (‚Werkzeuge‘, ‚Musikinstrumente‘, ‚Blumen‘ usw.) erstellt, die nicht nur weitere Gruppenvariationen ermöglichen, sondern auch ein höheres Maß an Transferkompetenz bei den Studierenden voraussetzen (s. 3.2.3).

3.2.2 Verbesserung der Variation der Aufgabentypen

Ein weiterer Schwerpunkt des Projekts lag auf der Verbesserung der Variation der Aufgabentypen. Ein Prüfungsformwechsel kann nicht allein in der bloßen Übertragung in ein digitales Format liegen (Handke/Schäfer, 2012, S. 165). Neben dem Ziel der Effizienz und zumindest teilweise automatisierter Korrektur muss auch echte Innovation stattfinden, die „über ein reines ‚Ankreuzen‘ als Wissensreproduktion hinausgeh[t].“ (Erren/Keil, 2007, S. 31). In diesem Sinne wurde die Auswahl der Aufgabentypen innerhalb des Projekts erweitert.²

Im Rahmen des Projekts liegt das Augenmerk vor allem auf der Verbesserung der Variation der Aufgabentypen. Im ersten Klausurteil war bereits bei Projektbeginn ein gutes Spektrum vorhanden. Die Klausur umfasst vor allem geschlossene Aufgabenformate (vgl. Vogt/Schneider, 2009, S. 8ff.), um den hohen Teilnehmendenzahlen und dem damit verbundenen Korrekturaufwand gerecht zu werden. Des Weiteren ermöglicht eine weitgehend automatisierte Korrektur eine schnelle und einheitliche Bearbeitung der Klausur und eine besonders hohe Objektivität (Handke/Schäfer, 2012, S. 160).

Im zweiten Klausurteil war die Diversität der Aufgabentypen hingegen zunächst noch recht eingeschränkt und beruhte vor allem auf den Typen Single-Choice, Multiple-Choice, Drag & Drop und True/False. Vor allem der letztere Typ war in Teil 2 überdurchschnittlich oft zu finden, ermöglichte so eine hohe Ratewahrscheinlichkeit und nur wenig Anwendungsorientierung (s. dazu auch 3.2.3.). Nach umfassender Überarbeitung finden sich nun folgende weiter ausdifferenzierte Aufgabentypen: Single-Choice, Multiple-Choice, True/False bzw. Yes/No, Freitextfelder in Lückentexten, Dropdown Menü in Lückentexten, Grafiken und Tabellen, Drag & Drop mit Text oder Grafik sowie eine Kombination mehrerer Typen.

Das erweiterte Spektrum an Aufgabentypen ermöglicht es, den verschiedenen Inhalten und Lehrzielen gerecht zu werden. So sind z. B. reine Definitionsaufgaben nur in Präsenzklausuren sinnvoll und können ohnehin nur der reinen Wissenswiedergabe dienen. In Distanzformaten und zum Testen tieferen Verständnisses bieten sich vor allem Anwendungs- und Analyseaufgaben und die Arbeit mit Beispielmaterial an, die besonders in Teil 1 bereits erweitert wurden und nun auch zunehmend im zweiten Klausurteil ergänzt werden.

3.2.3 Erhöhung des Aufgabenanteils mit Anwendungs- und Transferleistungen

Entsprechend der zweigliedrigen Anlage des Moduls mit seinen Modulzielen sind in Teil 1 der Modulprüfung vor allem der sichere Umgang mit grundlegenden Termini sowie die korrekte Beschreibung linguistischer Phänomene nachzuweisen. Vor allem hier ist es möglich, die Klausur sehr anwendungsorientiert zu

² Für eine detaillierte Besprechung von Testformen vgl. z. B. Handke/Schäfer (2012, S. 183 ff.).

gestalten. Es finden sich leicht Beispiele, über deren korrekte Klassifikation und Analyse sich sehr gut ein entsprechendes Verständnis der Materie nachweisen lässt. Im zweiten Teil der Klausur kommen sowohl Aufgaben zur historischen Dimension der englischen Sprache als auch weitere angewandte, etwa diskursanalytische, Anteile hinzu. Auch hier steht das handlungsorientierte Prüfen mehr und mehr im Fokus. Dabei bewegt sich E-Assessment immer im Spannungsfeld von Effizienz und hoher Objektivität auf der einen und Komplexität und höheren Lernzielen auf der anderen Seite.

In der Ling1-Klausur zeigte sich vor allem in den Distanzklausuren während der Pandemie, dass einfache Definitionsaufgaben noch stärker durch anwendungs- und transferorientierte Aufgaben abgelöst werden sollten. Exemplarisch war das etwa in abgegebenen Antworten sichtbar, die wortgleich mit Passagen aus dem Reader der Veranstaltung waren, aber nicht exakt zur spezifischen Fragestellung passten. Im Laufe der Weiterentwicklung der Klausur setzten wir so vermehrt auf Aufgaben mit Anwendungs-/Transferanteil. So kann beispielsweise über einfache Zählaufgaben (etwa: ‚How many noun phrases does the following sentence contain?‘), Drag-and-Drop-Typen (z. B. mit Grafik des Vokaltrapez: ‚Please drag the sample words into the box where they were situated before the Great Vowel Shift‘) oder Drop-down-Menüs (z. B. ‚The grammaticalization of the future auxiliary *going to* consists of the mechanisms: [reanalysis, phonological reduction and layering].‘) die Anwendungskompetenz der Studierenden evaluiert werden.

3.2.4 Steigerung der Authentizität der verwendeten Beispiele

Gerade zu Beginn des Studiums stellt das Verständnis von Linguistik als Wissenschaft eine Herausforderung für viele Studierende dar. Während sich in der Literaturwissenschaft häufiger ein Anknüpfungspunkt an den Schulunterricht bietet, ist das Einnehmen einer metasprachlichen Perspektive oft weniger vertraut. Wissenschaftliches Denken im linguistischen Kontext im Allgemeinen und das Verständnis von Sprache als Untersuchungsgegenstand im Speziellen ist deshalb ein wichtiges Lernziel für Studierende in der Studieneingangsphase (Anderson, 2016, S. e274). Dabei handelt es sich genau um den Schritt im individuellen Lernfortschritt, der eine neue Art von Denken über Sprache ermöglicht, die essenziell für den weiteren Lernfortschritt im Studium ist. In diesem Sinne ist das Verstehen von Sprache als Untersuchungsgegenstand ein klassisches Schwellenkonzept (Meyer/Land, 2003, S. 1).

Ein Weg, den Studierenden das Überschreiten dieser Schwelle zu erleichtern, ist die Anbindung der Studieninhalte an authentische Lerngegenstände, um so einen Alltagsbezug herzustellen und abstrakte Konzepte leichter greifbar zu machen. In der Lehrpraxis hat sich dazu die Verwendung authentischen Materials wie Comics, Zeitungsartikel, Werbung und Memes bereits bewährt. Auch in die Klausur soll zukünftig verstärkt authentisches Sprachmaterial (z. B. Fotos

aus der Alltagswelt) eingebunden werden. Wünschenswert wäre zudem die Nutzungsmöglichkeit von Hörbeispielen oder audiovisuellem Material (etwa kurzen Videoclips), die aber eine entsprechende Erweiterung der technischen Ausstattung der Computerpools voraussetzt.

3.2.5 Einbezug studentischen Feedbacks

Bereits seit Projektbeginn besteht ein essenzieller Schritt zur Weiterentwicklung der Ling1-Klausur im Einbezug studentischen Feedbacks. Dieses wurde zunächst bei der Klausureinsicht mündlich und im weiteren Verlauf des Projekts regelmäßig und systematisch mithilfe eines Online-Fragebogens eingeholt. Handke/Schäfer (2012, S. 166f.) betonen die Mitarbeit aller involvierten Gruppen im E-Assessment. Demnach sollen auch Anregungen Studierender aufgegriffen und idealerweise umgesetzt werden.

Da eine multiprofessionelle Autorenumgebung (Handke/Schäfer, 2012, S. 161f.) zunächst nicht zur Verfügung stand, hatten Lehrende vor Projektbeginn freilich zusammengearbeitet, um die Klausurteile zu erstellen, jedoch hatte es bisher an einer Instanz gefehlt, welche die Gesamtklausur sowohl mit kritischem Abstand als auch mit linguistischem Fachwissen betrachten konnte. Das PITCH-Projekt ermöglicht hier die, wenn auch nur vorübergehende, Schaffung einer Schnitt- und Evaluationsstelle. So kann studentisches Feedback gezielt eingeholt und umgesetzt werden.

Im Fall der Ling1-Klausur bezog sich einer der am häufigsten geäußerten Feedbackpunkte auf den Zeitfaktor in der Bearbeitung der Klausur. Bei der nochmaligen Überprüfung der Klausur wurde offensichtlich, dass beide Teile zwar die gleiche Anzahl an Aufgaben hatten, die Komplexität und die zu erreichenden Punkte im zweiten Teil jedoch umfangreicher waren. Hier wurden nun gezielt längere Aufgaben gekürzt und teilweise in Alternativaufgaben gewandelt und damit auch die zu erreichende Punktzahl in beiden Klausurteilen ausbalanciert. Der Erfolg dieser Anpassung spiegelt sich in den Umfrageergebnissen. Während 83 % der Studierenden in der Feedbackumfrage im Sommersemester 2020/21 das Zeitlimit zur Bearbeitung der Klausur als nicht bis gar nicht angemessen einstufen, war im Sommersemester 2023 die gesamte Kohorte der Ansicht, das Zeitlimit sei sehr angemessen oder angemessen. Auch das Mock-Exam und die digitalen Lernumgebungen für die Veranstaltungen wurden aufgrund der studentischen Rückmeldungen angepasst.

Eine Schwierigkeit ergibt sich dennoch aus den tendenziell eher niedrigen Rücklaufquoten bei den Feedbackumfragen und der ggf. daraus resultierenden, eingeschränkten Repräsentativität der Ergebnisse. So ist erwartbar (und zeigt sich z. B. auch in den Angaben zur Eigenleistung), dass Studierende, die mit dem Klausurergebnis unzufrieden waren, eher Feedback geben als jene, die die Klausur mit aus ihrer Sicht zufriedenstellendem Ergebnis bestanden haben. Die Umfrageergebnisse werden deshalb auch innerhalb des Lehrendenteams reflek-

tiert und Anregungen auf die Notwendigkeit und Möglichkeit zu implementierender Maßnahmen überprüft.

4 Fazit

Der vorliegende Beitrag bietet neben einem Vergleich der Vor- und Nachteile verschiedener Klausurformate Einblick in die Aufgabengestaltung für eine Einführungsklausur ‚Englische Sprachwissenschaft‘, wobei insbesondere die Umsetzung der Projektziele – Erweiterung des Aufgabenpools, Verbesserung der Variation und Erhöhung des Aufgabenanteils mit Anwendungs- und Transferleistungen, Steigerung der Authentizität sowie Einbezug studentischen Feedbacks – im Mittelpunkt steht. Allerdings unterliegt die Aufgabengestaltung als eigentlicher Schwerpunkt der Klausurerstellung den praktischen Rahmenbedingungen, die Voraussetzung von Innovation sind. Unser Fazit rückt deshalb diese Voraussetzungen in den Blick.

Innovation braucht eine starke Basis in Form von zeitlichen, personellen und räumlichen Ressourcen. Zwar sind materielle Mittel wie Leitfäden und Handreichungen grundsätzlich hilfreich, jedoch nur begrenzt erfolgreich, wenn Lehrenden nicht auch ein entsprechendes Zeitkontingent eingeräumt wird. Der zeitliche Aufwand für die Klausurerstellung, Entwicklung von Strukturen und Metadokumenten, Absprachen, Evaluation und Qualitätssicherung und nicht zuletzt einschlägige Fortbildungen ist nicht zu unterschätzen.

Ebenso braucht Innovation spezifische personelle Ressourcen. Die Schnittstelle von Fachkompetenz, Lehrerfahrung, technischem Knowhow und Mitarbeitendenkommunikation hat ein Anforderungsprofil, dem studentische Hilfskräfte kaum gerecht werden können. Hinzu kommt der technische Support durch ein entsprechend aufgestelltes IT-Team. Des Weiteren müssen auch die räumlichen Gegebenheiten der Innovation dienlich sein. Abhängig vom Prüfungs- und spezifischen Aufgabentypen entstehen Bedarfe nach modernen PC-Pools mit adäquater Hard- und Software.

Weiterhin braucht Innovation Nachhaltigkeit, um dem ständigen technischen Wandel gerecht werden zu können. Lehrende müssen sich auf funktionsfähige Prüfungsplattformen verlassen können. Nur so ist es möglich, ein System effizient und langfristig so zu nutzen, dass zeitliche Ressourcen für die kontinuierliche Erweiterung und Updates der Aufgabenkataloge und eine innovative Aufgabengestaltung aufgewendet werden können.

Auch in personeller Hinsicht ist Nachhaltigkeit wichtig. Es kann Jahre dauern, die Schwerpunkte und Herausforderungen einer Prüfung kennenzulernen, Kommunikationsstrukturen mit den involvierten Lehrenden aufzubauen und die Chancen und Grenzen einzelner Prüfungsplattformen zu durchdringen. Hier gilt es, personelle Strukturen zu schaffen, die diesem diversen Anforderungsprofil langfristig gerecht werden. Ebenso wie Handke/Schäfer (2012, S. 161f.) plä-

dieren wir hier ausdrücklich für eine multiprofessionelle Autorenumgebung, die Lehrende und kompetentes Fachpersonal für die spezifische technische Umsetzung einer Klausur zusammenführt. Wenn es multiprofessionellen Teams ermöglicht wird, eng und dauerhaft zusammenzuarbeiten, können Projekte wie PITCH der Katalysator für wirkliche Innovation sein.

Nicht zuletzt braucht Innovation aber auch Agilität. Die vergangenen Semester, die nicht nur durch den digitalen Wandel, sondern zusätzlich auch durch die Pandemie und einen Cyberangriff Studierende und Lehrende vor bislang ungekannte Herausforderungen gestellt haben, machten dies mehr als deutlich. Um den immer wieder neuen Gegebenheiten gerecht zu werden, müssen Universitäten, ihre Mitarbeitenden und Studierenden dynamisch und proaktiv agieren. Dafür benötigt es eine starke gemeinsame Basis.

Literatur

- Anderson, C. (2016). Learning to think like linguists: A think-aloud study of novice phonology students. In: *Language*, 92(4), S. e274–e291.
- Bloom, B. S. (1972). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim: Beltz.
- Erren, P./Keil, R. (2007). *Medi@rena – ein Ansatz für neue Lernszenarien*. In M. Merkt/K. Mayberger/R. Schulmeister/A. Sommer/I. van den Berk (Hrsg.): *Studieren neu erfinden – Hochschule neu denken*. GMW: Band 44. (S. 21–31). Münster: Waxmann Verlag.
- Friebertshäuser, B. (2008). Statuspassage von der Schule ins Studium. In W. Helsper/J. Böhme (Hrsg.): *Handbuch der Schulforschung* (S. 611–627). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Handke, J./Schäfer, A. M. (2012). *E-Learning, E-Teaching und E-Assessment in der Hochschullehre – Eine Anleitung*. München: Oldenbourg Wissenschaftsverlag.
- Kossack, P./Lehmann/U. Ludwig, J. (2012). *Die Studieneingangsphase. Analyse, Gestaltung und Entwicklung*. Bielefeld: Universitätsverlag Webler.
- Meyer, J. H. F./Land, R. (2003). Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines. In: C. Rust (Hrsg.): *ISL10 Improving Student Learning: Theory and Practice Ten Years On* (S.412–424). Oxford: Oxford Brookes University.
- Universität Duisburg-Essen (2024). *BA Gymnasium und Gesamtschule, Studienfach Englisch, Stand Februar 2024*. https://www.uni-due.de/imperia/md/images/anglistik/mhb_la_ba_gyge_05_2024.pdf [Zugriff: 04.02.2024].
- Vogt, M./Schneider S. (2009). *E-Klausuren an Hochschulen*. Koordinationsstelle Multimedia. Gießen: JLU Gießen.

Digitale Prüfwerkzeuge in der Baudynamik – Ein Erfahrungsbericht

Carolin Birk und Ajay Kumar Pasupuleti

Während der COVID-19-Pandemie wurden Lehrende und Hochschulen plötzlich mit der Herausforderung konfrontiert, Prüfungen im Distanzformat durchführen zu müssen. Dabei standen zunächst rechtliche und technische Unsicherheiten im Vordergrund. Bald wurden jedoch auch die zusätzlichen Potenziale sichtbar, die digitale Prüfungen bieten. Vor- und Nachteile digitaler Prüfungen werden beispielsweise in Knoke (2021) zusammengefasst. Für Lehrende ist insbesondere der durch automatisierte Auswertung potenziell deutlich geringere Korrekturaufwand von Vorteil. Studierende schätzen die schnellere Verfügbarkeit der Ergebnisse. Zudem kann die automatisierte Prüfungsauswertung zur Anonymisierung und damit zur Erhöhung der Chancengleichheit beitragen. Demgegenüber stehen gegebenenfalls der höhere Aufwand zur Prüfungserstellung sowie Bedenken, inwiefern digitales kompetenzorientiertes Prüfen in allen Fachgebieten gleichermaßen möglich ist.

Am Beispiel des Moduls „Baudynamik“ des Masterstudiengangs Bauingenieurwesen setzen sich die Autor:innen dieses Beitrags damit auseinander, inwiefern kompetenzorientiertes digitales Prüfen gelingen kann. Im Folgenden wird von den Erfahrungen bei der Entwicklung und dem Einsatz digitaler Prüfwerkzeuge berichtet. Gleichzeitig werden die sich aus Sicht der Autor:innen ergebenden Grenzen und Möglichkeiten aufgezeigt. Hierzu wird zunächst die Ausgangslage erläutert und anschließend die gewählte Form der digitalen Prüfung beschrieben. Danach werden konkrete Ergebnisse und Erfahrungen zusammengefasst. Der Beitrag schließt mit einem Fazit.

1 Ausgangslage: Das Modul „Baudynamik“

„Baudynamik“ ist ein Pflichtfach in der Vertiefungsrichtung Konstruktiver Ingenieurbau des Masterstudiengangs Bauingenieurwesen. Pro Wintersemester absolvieren etwa 50–80 Studierende dieses Modul. Gleichzeitig existiert eine englischsprachige Variante des Faches, „Structural Dynamics“, welche fortlaufend für Studierende des internationalen Masterstudiengangs „Computational Mechanics“ als Wahlpflichtmodul angeboten wird. Es wird durchschnittlich von 10 bis 15 Studierenden pro Semester besucht.

Es handelt sich um ein theoretisch anspruchsvolles Fach an der Schnittstelle von Mathematik, Numerik und Tragwerksplanung. Ein Hauptanliegen der Lehrveranstaltungen ist es, dass die Studierenden ein tiefes Verständnis der Einflussfaktoren entwickeln, die das dynamische Verhalten von Tragwerken bestimmen. Bei Abschluss des Moduls sollen Studierende nicht nur typische dynamische Belastungen und grundlegende Konzepte der Schwingungsanalyse kennen, sondern die Methoden der Baudynamik auch auf realistische Probleme der Tragwerksplanung anwenden und Ergebnisse kritisch bewerten können. Um sie hierzu zu befähigen, sind Präsenzlehrveranstaltungen insbesondere darauf gerichtet, mit den Studierenden einerseits theoretisch schwierige Zusammenhänge nachhaltig zu erarbeiten und andererseits den Praxisbezug herzustellen.

Traditionell werden klassische Fächer des konstruktiven Ingenieurbaus, wie „Baudynamik“, schriftlich geprüft. Bisher werden dabei im Rahmen einer zwei-stündigen Präsenzklausur typischerweise drei Aufgaben aus den Themenkomplexen Idealisierung, freie und erzwungene Schwingung des Einmassenschwingers, menscheninduzierte Schwingungen, modale Analyse und Erdbebenanalyse bearbeitet. Die einzelnen Aufgaben zeichnen sich durch eine gewisse Komplexität aus. Sie sind derart konzipiert, dass Studierende das Gelernte auf neue Situationen übertragen und dabei für den Arbeitsprozess von Ingenieur:innen typische Kompetenzen nachweisen müssen. Hierzu gehört es, dass der Lösungsprozess nicht immer klar vorgezeichnet ist. Studierende müssen selbst in der Lage sein, entsprechend geeignete Verfahren auszuwählen, anzuwenden und die Ergebnisse zu interpretieren. Hierbei lassen manche Aufgaben einen gewissen Entscheidungsspielraum zu. Gerade aus dieser ingenieurtypischen Art der Prüfung ergeben sich jedoch spezielle Herausforderungen für die Konzipierung und Umsetzung digitaler Prüfungen.

2 Methodik: Kriterien und Anforderungen an digitale Prüfungsformate ableiten

Im Rahmen des Projekts wurde daher zunächst im Sinne einer Bestandsaufnahme die vorherrschende Situation der Präsenzprüfung gründlich analysiert, um hieraus Kriterien für und Anforderungen an mögliche digitale Prüfungsformate abzuleiten. Dabei wurden unter anderem die folgenden Merkmale traditioneller Prüfungsformen in den Fächern des konstruktiven Ingenieurbaus identifiziert:

- Typische Aufgaben setzen sich aus offenen Fragen, Anwendungsfragen und Problemlösefragen zusammen (Ulrich, 2016).
- Im Korrekturverlauf werden häufig Folgefehler berücksichtigt, da die erwarteten Lösungen komplex sind und angesichts der Prüfungssituation richtige Vorgehensweisen wichtiger sind als exakte Endergebnisse.
- Endergebnisse sind teilweise grafisch darzustellen, wie beispielsweise Eigenformen, Verschiebungs- und Schnittkraftverläufe.

Gleichzeitig wurden die an der Universität Duisburg-Essen verfügbaren digitalen Prüfwerkzeuge betrachtet und hinsichtlich ihrer Eignung für das Vorhaben bewertet. Da rein Multiple-Choice-basierte Werkzeuge keine Fehlerverfolgung zulassen, wurden diese geringer priorisiert. Hohe Bedeutung wurde jedoch der Möglichkeit der Randomisierung von Aufgaben beigemessen, um Studierende bereits im Rahmen der Lehrveranstaltungen an die entsprechenden digitalen Prüfwerkzeuge heranführen zu können. Schlussendlich schien das UDE-eigene Übungs- und Prüfungssystem JACK (vgl. Beitrag von Christoph Olbricht und Melanie Schypula in diesem Band) die spezifischen Anforderungen der Prüfungskultur in den Ingenieurwissenschaften am besten zu erfüllen. Die Nutzung von JACK im Rahmen des hier beschriebenen Teilprojekts wird im Folgenden näher beschrieben.

2.1 Einsatz von JACK im Modul „Baudynamik“

Das Übungs- und Prüfungssystem JACK wird seit 2006 an der Universität Duisburg-Essen entwickelt, wobei es ursprünglich für den Bereich der Programmierausbildung vorgesehen war. Im Zusammenhang mit dem Einsatz in der Baudynamik sind folgende Merkmale hervorzuheben: JACK erlaubt die Gestaltung mehrstufiger Aufgaben. Es sind vielfältige Fragetypen möglich, wie beispielsweise Multiple-Choice, Lückentext und Eingabefelder. Eingegeben und bewertet werden können sowohl numerische Werte als auch Formelteile. Die Aufgaben lassen sich randomisieren. Es ist möglich, das Bewertungsschema an die Anzahl der Versuche anzupassen. JACK erlaubt die gezielte und an die Ergebnisseingabe angepasste Rückgabe von Feedback und Hinweisen.

Im Rahmen des Projekts wurde JACK in zwei unterschiedlichen Szenarien eingesetzt. Um einerseits die Studierenden an die Nutzung dieses Werkzeugs heranzuführen und andererseits Machbarkeit und Grenzen genauer auszuloten, diente es im Sommersemester 2022 zunächst als reines Übungstool im Fach „Structural Dynamics“. Im Verlauf eines Semesters werden für gewöhnlich ca. 10 Präsenzübungen durchgeführt, innerhalb derer jeweils mehrere Aufgaben behandelt werden. Diese wurden vollständig in JACK-Aufgaben überführt, wobei jeweils nur dann von der Originalaufgabe abgewichen wurde, wenn der Umsetzung technische Grenzen gesetzt waren – beispielsweise bei Abfrage grafischer Ergebnisse. Die so entstandenen JACK-Aufgaben wurden jeweils direkt im Anschluss an die zugehörige Präsenzvorlesung über den entsprechenden Moodlekurs zur Verfügung gestellt. Die Studierenden wurden zur eigenständigen Bearbeitung der Aufgaben innerhalb einer Woche ermutigt, mit der Erwartung, verbleibende Fragen in der dann folgenden Präsenzübung diskutieren zu können. Didaktisch wurde so das Ziel verfolgt, das Selbstlernverhalten der Studierenden zu verbessern.

Dieses Vorgehen wurde im Wintersemester 2022/23 auf den deutschsprachigen Kurs „Baudynamik“ übertragen. Zusätzlich wurde in diesem Semester eine

erste digitale Klausur für das Modul konzipiert und in JACK umgesetzt. Hierzu diente eine typische Altklausur als Vorlage. Hierbei konnten die existierenden Prüfungsaufgaben jedoch nicht wie die oben erwähnten Übungsaufgaben lediglich originalgetreu digitalisiert werden. Wie eingangs erläutert, sind typische Prüfungsaufgaben relativ offen formuliert. Der tatsächliche Lösungsweg kann sich bei einer traditionellen Klausur gelegentlich über mehr als eine Seite hinziehen. Bei der händischen Korrektur wird dieser gesamte Lösungsweg nachvollzogen und gegebenenfalls werden Punkte auf Teilschritte und Vorgehensweisen vergeben. Eine reine „Übersetzung“ der vorliegenden Prüfungsaufgaben in JACK-Aufgaben hätte möglicherweise dazu geführt, dass lediglich Endergebnisse unmittelbar abgefragt werden. Hierbei war zu erwarten, dass eine große Anzahl der Studierenden sehr schlecht abschnitten würde, da Rechenfehler häufig auftreten und es keine Möglichkeit gäbe, den Rechenweg in Betracht zu ziehen. Eine Differenzierung zwischen Studierenden, die trotz kleiner Rechenfehler ein gutes Verständnis demonstrieren können und denen, die falsch vorgehen, wäre so nicht mehr möglich. Aus diesem Grund wurden die Prüfungsaufgaben in Orientierung an die Altklausur für JACK neu konzipiert, wobei einerseits stärker Zwischenergebnisse abgefragt, andererseits zusätzliche Verständnisfragen aufgenommen wurden. Hierbei bestand eine wesentliche Herausforderung darin, die richtige Balance zwischen Vorgeben des Lösungsweges und Prüfen der ingenieurtypischen Kompetenzen zu finden.

Die digitale Prüfung wurde in der letzten Woche des Wintersemesters 2022/23 als „Testklausur“ in einer für digitale Klausuren ausgestatteten PC-Hall der Universität Duisburg-Essen durchgeführt. Die Teilnahme war für Studierende freiwillig. Sie wurde jedoch als Ersatz für die alternativ als Zulassungsvoraussetzung für die schriftliche Klausur zu erbringende Hausübung anerkannt, um den Studierenden einen Anreiz zur Teilnahme zu bieten. Vor Beginn der offiziellen Prüfungszeit wurde den Studierenden die Nutzung des Prüfungstools an einem Beispiel demonstriert. Außerdem wurden umfangreiche Hinweise zu den Rahmenbedingungen und technischen Details, wie beispielsweise zur erforderlichen Eingabe der numerischen Werte mit Trennzeichen Punkt, gegeben. Eine der drei Prüfungsaufgaben war sehr umfangreich, sodass der mehrseitige Aufgabentext zusätzlich in Papierform zur Verfügung gestellt wurde.

Wie erwartet haben alle Studierenden die Aufgaben „klassisch“ auf Papier gelöst und anschließend die erhaltenen Ergebnisse in JACK eingegeben. Die Studierenden wurden gebeten, ihre Notizen und Zwischenrechnungen abzugeben. In der Nachbearbeitung der Klausur wurden diese Arbeitsblätter durchgesehen. Diese Vorgehensweise ermöglichte den Vergleich der automatischen Prüfungsbewertung durch JACK mit der traditionellen manuellen Korrektur. In Einzelfällen ließ sich so evaluieren, inwiefern den Studierenden durch die beschränkte Möglichkeit der Fehlernachverfolgung ggf. Nachteile entstanden. Im direkten Anschluss an die digitale Testklausur wurden die Studierenden gebeten, einen anonymen Fragebogen auszufüllen, der über Moodle bereitge-

stellt wurde. Die detaillierte Auswertung der Ergebnisse der Testklausur sowie des zugehörigen Feedbacks der Studierenden erfolgt in Abschnitt 4.

2.2 Umsetzung am Beispiel einer typischen Aufgabe

Abbildung 1 zeigt eine typische Aufgabe aus der papierbasierten Klausur „Baudynamik“. Sie umfasst drei Teilaufgaben, die jeweils voneinander abhängen. Wie erläutert, lässt sich die Aufgabe nicht identisch in JACK umsetzen, ohne dass den Studierenden Nachteile bei der Bewertung entstehen oder der Lösungsweg vorgegeben wird. In die JACK-Aufgabe wurden daher zusätzliche Teilaufgaben zum Verständnis aufgenommen. Außerdem wurden die Teilaufgaben voneinander entkoppelt, indem jeweils anzunehmende Zwischenergebnisse vorgegeben wurden.

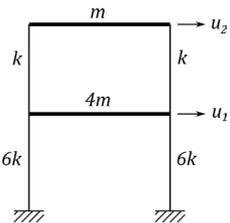


Bild 1: Vereinfachtes Modell eines zweistöckigen Rahmentragwerks

Bild 1 zeigt das vereinfachte Modell eines Rahmentragwerks. Die horizontal angeordneten Balken können als biegesteif idealisiert werden. Die Stützen seien masselos. Dämpfung kann vernachlässigt werden.

1. Geben Sie die Steifigkeitsmatrix und die Massenmatrix des Systems an.
2. Berechnen Sie die Eigenfrequenzen und Eigenformen des idealisierten Systems. Skizzieren Sie die Eigenformen.
3. Ermitteln Sie mittels modaler Analyse die Antwort des Systems auf die folgenden Anfangsbedingungen:

$$\mathbf{u}(t = 0) = \begin{Bmatrix} -2 \\ -8 \end{Bmatrix} [cm], \quad \dot{\mathbf{u}}(t = 0) = \begin{Bmatrix} -2 \\ +2 \end{Bmatrix} \left[\frac{cm}{s} \right]$$

Geben Sie die Lösung als Funktion von k und m an.

Abbildung 1: Aufgabe zum Thema „Modale Analyse“ (papierbasiert)
 (Quelle: Eigene Abbildung)

Abbildung 2 zeigt die Gliederung der digitalen Aufgabe. Die Teilaufgaben 3 und 5 prüfen das Verständnis grundlegender Konzepte zum Thema „Modale Analyse“.

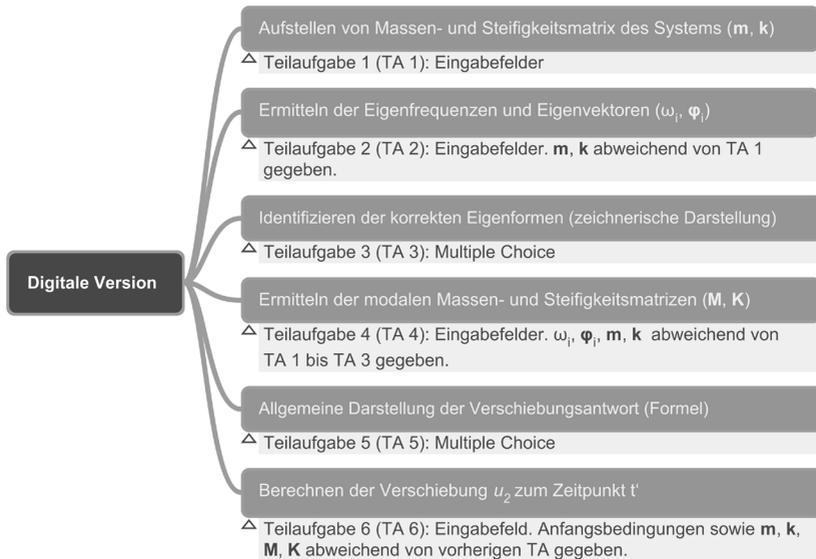


Abbildung 2: Struktur der digitalen Aufgabe zum Thema „Modale Analyse“
 (Quelle: Eigene Abbildung)

3 Ergebnisse

Die im Sommersemester 2022 und im Wintersemester 2022/23 in den Kursen „Structural Dynamics“ und „Baudynamik“ wöchentlich angebotenen Übungsaufgaben (JACK) wurden gut angenommen. In den entsprechenden Moodlekursen waren 16 beziehungsweise 74 Teilnehmer und Teilnehmerinnen eingeschrieben. Etwa 60 % dieser Personen haben die JACK-Aufgaben bearbeitet. Das entspricht auch ungefähr der Anzahl der Studierenden, die erfahrungsgemäß tatsächlich innerhalb des Semesters aktiv mitarbeiten und sich schließlich zur schriftlichen Prüfung anmelden. Beide Klausuren sind besser ausgefallen als in vorherigen Semestern. Aufgrund der erstmaligen Umsetzung des Konzepts und der geringen Teilnehmendenzahlen im Fach „Structural Dynamics“ lassen sich diese Beobachtungen jedoch noch nicht verallgemeinern. Nicht beobachten ließ sich leider die erhoffte aktivere Teilnahme der Studierenden an den Präsenzübungen in Form von Diskussionsbeiträgen.

An der digitalen Testklausur im Fach Baudynamik im Januar 2023 haben 41 Studierende teilgenommen. Die Ergebnisse der automatischen Bewertung durch JACK werden in Tabelle 1 zusammengefasst.

Ergebnistyp	Wert
Durchschnittliche Bewertung	22 %
Bestes Einzelergebnis	46 %
Schlechtestes Einzelergebnis	5 %

Tabelle 1: Ergebnisse der digitalen Testklausur „Baudynamik“ WiSe 2022/23
(Quelle: Eigene Darstellung)

Bei der Beurteilung dieser Ergebnisse ist zu beachten, dass die Testklausur innerhalb der Vorlesungszeit stattfand. Außerdem war die Bewertung nicht notenwirksam. Daher kann davon ausgegangen werden, dass die Mehrzahl der Studierenden nicht im gleichen Maße wie für eine reale Klausur vorbereitet war. Dennoch können die in Tabelle 1 erkennbaren niedrigen Ergebniswerte teilweise auf die automatische Bewertung zurückgeführt werden. Bei der manuellen Korrektur der zusätzlich eingereichten Lösungswege hat sich im Mittel ein Punktgewinn von 20 % bis 30 % ergeben.

Während der Durchführung der digitalen Testklausur sind zahlreiche Fragen aufgetreten, die von den Autor:innen dieses Beitrags individuell beantwortet wurden. Diese waren überwiegend technischer Natur und betrafen beispielsweise die korrekte Eingabe von Zahlenwerten und Formelaustrücken sowie die Wiederholbarkeit von Versuchen. Einige inhaltliche Rückfragen zu Aufgabenteilen waren ebenfalls zu verzeichnen, jedoch in vergleichbarem Maße zu einer Präsenzklausur.

Der anonyme Fragebogen wurde von 40 Studierenden ausgefüllt. Das erhobene Feedback ist äußerst umfangreich und wird im Folgenden in Auszügen wiedergegeben. Einige repräsentative Fragen und Durchschnittswerte der auf einer Skala erfassten Antworten werden in Tabelle 2 zusammengefasst.

Obwohl die auf die technische Umsetzung und didaktische Konzeption gerichteten Fragen überwiegend positiv (Durchschnittswerte zwischen 3,3 und 4,0) beantwortet wurden, favorisieren die Studierenden aktuell die klassische Prüfungsform. Betont wird, dass den Studierenden zum Zeitpunkt der Beantwortung des Fragebogens das Prüfungsergebnis noch nicht bekannt war.

Umfangreiche Hinweise wurden zu technischen Problemen gegeben. Mehrfach wurde beklagt, dass Aufgaben relativ langsam laden und sich Formularfelder erst nach erneutem Aktualisieren ausfüllen ließen. Zahlreiche Studierende hatten Schwierigkeiten bei der Verwendung des Formeleditors. Einige Aufgaben haben sich als nicht ausreichend robust gegenüber falschen Eingaben erwiesen. Beispielsweise führte die Eingabe einer einfachen Dezimalzahl an einer Stelle,

an der eine Matrix erwartet wurde, zu einer Fehlermeldung. Diese Hinweise können teilweise von den Entwickler:innen der Aufgaben adressiert werden; sie fließen aktuell in die Überarbeitung ein. Erläuternd soll an dieser Stelle gesagt werden, dass die im Übungsbetrieb verwendeten JACK-Aufgaben auf der älteren Plattform JACK2 entwickelt wurden, während die digitale Testklausur auf der neueren Plattform JACK3 umgesetzt wurde. Insbesondere den Umgang mit dem Formeleditor hatten die Studierenden also nicht vorab geübt. Aktuell werden alle Übungsaufgaben in das Format von JACK3 überführt und zusätzlich weitere Aufgaben konzipiert. Das zukünftige Üben mit JACK3 wird aus Sicht der Autor:innen entscheidend zum Erfolg und zur Akzeptanz des digitalen Prüfungsformates beitragen.

Das weitere qualitative Feedback lässt sich wie folgt zusammenfassen: Einige Studierende empfinden die Präsentation längerer Aufgaben als unübersichtlich und wünschen sich, dass alle Aufgaben zusätzlich auf Papier ausgeteilt werden. Das trage auch dazu bei, sich zu Beginn einen besseren Überblick über den gesamten Arbeitsaufwand zu verschaffen. Als negativ wurde auch empfunden, dass nicht alle Teilaufgaben unabhängig voneinander waren. Beim Zurückkehren zu einer früheren Teilaufgabe werden die Ergebnisse aller nachfolgenden (abhängigen) Teilaufgaben überschrieben. Einige Studierende gaben an, dass es schwieriger sei, sich während der digitalen Klausur zu konzentrieren. Mehrfach bemängelt wurde der geringe verfügbare Platz zur Ablage der mitgebrachten Materialien sowie zur Anfertigung handschriftlicher Zwischenrechnungen. Mehrere Studierende haben ausdrücklich die gute Vorbereitung der Testklausur sowie die logische Aufgabenführung gelobt. Auch die Anreicherung mit schnell zu beantwortenden Verständnisfragen wurde positiv hervorgehoben. Insgesamt ergab sich der Eindruck, dass viele Studierende der digitalen Klausur gegenüber positiv eingestellt und offen für Veränderung sind, sofern sie mit der technischen Umgebung und dem Ablauf vorab gut vertraut gemacht werden.

Frage	Antwort
Die technischen Anforderungen waren mir vor der Prüfung bekannt.	Ja: 88 % / Nein: 12 %
Die Vorgehensweise zur Durchführung der Prüfung war mir vor der Prüfung bekannt.	Ja: 88 % / Nein: 12 %
Die Art der Prüfungsfragen war mir vor der Prüfung bekannt (z. B. Single-Choice, Multiple-Choice, Freitext).	Ja: 90% / Nein: 10%
Mir waren Ansprechpartner:innen für technische Probleme bekannt.	Ja: 98 % / Nein: 2 %
Mir waren Ansprechpartner:innen für inhaltliche Fragen bekannt.	Ja: 98 % / Nein: 2 %
Ich hatte technische Probleme während der Prüfung.	Ja: 35 % / Nein: 65 %

Frage	Antwort
Die Prüfungsfragen/Aufgabenstellungen waren für mich verständlich formuliert. [(1) trifft überhaupt nicht zu ... (5) trifft voll und ganz zu]	Durchschnittswert: 3,7
Die Zeit für das Beantworten der Fragen/das Lösen der Aufgaben war für mich... [(1) überhaupt nicht ausreichend ... (5) angemessen]	Durchschnittswert: 3,3
Die Prüfungsfragen waren gut auf die Inhalte der Lehrveranstaltung abgestimmt. [(1) trifft überhaupt nicht zu ... (5) trifft voll und ganz zu]	Durchschnittswert: 4,0
Insgesamt war der Schwierigkeitsgrad der Prüfung für mich angemessen. [(1) trifft überhaupt nicht zu ... (5) trifft voll und ganz zu]	Durchschnittswert: 3,6
Die Prüfungsaufgaben ähnelten den Aufgaben, die wir im Rahmen der Lehrveranstaltung geübt haben. [(1) trifft überhaupt nicht zu ... (5) trifft voll und ganz zu]	Durchschnittswert: 3,8
Ich empfand die Prüfung insgesamt als fair. [(1) trifft überhaupt nicht zu ... (5) trifft voll und ganz zu]	Durchschnittswert: 3,5
Wie beurteilen Sie das Gesamtergebnis der Prüfung? [(1) sehr gut... (5) sehr schlecht]	Durchschnittswert: 2,5
War es angenehmer, mit gedruckten und digitalen Fragen zu arbeiten, als mit rein digitalen Fragen?	Ja: 75 % / Nein: 25 %
Welche Prüfungsform würden Sie zukünftig für Statik 5 (Baudynamik) bevorzugen?	Digital mit JACK: 23 % Auf Papier, wie bisher: 78 %

Tabelle 2: Fragebogen zur digitalen Testklausur – ausgewählte Ergebnisse
(Quelle: Eigene Darstellung)

4 Fazit

Im Rahmen des Projekts wurde eine digitale Testklausur für das Fach „Baudynamik“ erarbeitet, in JACK3 umgesetzt und erfolgreich durchgeführt. Die Erfahrungen im Projektverlauf, die Klausurergebnisse sowie das äußerst umfangreiche Feedback der Studierenden lassen darauf schließen, dass digitale Prüfungsformate in den Ingenieurwissenschaften großes Potenzial haben. Dennoch gibt es gewisse Grenzen, die nicht außer Acht gelassen werden können.

Der Aufwand zur Erstellung der Aufgaben ist erheblich. Entwickler:innen müssen nicht nur das zu prüfende Fach beherrschen, sondern auch Kenntnisse des Prüfungstools vorweisen und typische Fehlerquellen vorhersehen können. Insgesamt zeichnet sich jedoch ab, dass auch auf Seiten der Lehrenden ein Lern-

prozess stattfindet, der schließlich in einen effizienten Arbeitsablauf münden und zu robusteren Aufgaben führen kann. Wesentlich für die Akzeptanz digitaler Prüfungen in den Bauwissenschaften ist, wie gut die Studierenden das System und die entsprechenden Abläufe kennen. Empfehlenswert ist bei digitaler Prüfung mit JACK eine Verlängerung der Bearbeitungszeit.

Die größte Herausforderung stellt die Aufgabenkonzeption derart dar, dass nicht nur Endergebnisse geprüft werden, ohne gleichzeitig den Lösungsweg vorzugeben. Obwohl dies bei der Erstellung der Testklausur ein Anliegen war, hätten die Studierenden in der automatisch bewerteten Klausur schlechter abgeschnitten als bei manueller Korrektur und Bewertung der Rechenwege. Aus Sicht der Autor:innen ist es schwer vorstellbar, die Aufgaben kleinteiliger aufzubauen, da so ingenieurtypische Kompetenzen nicht mehr abgefragt würden.

Regelmäßig in der Lehre eingesetzte JACK-Aufgaben können einen großen Gewinn darstellen. Sie tragen dazu bei, dass Studierende das Gelernte kurzfristig üben und stärken damit allgemein das Selbstlernverhalten. Das Konzept der wochenweisen JACK-Übungsaufgaben wird daher aktuell auf Grundlagenfächer des Bachelorstudienganges Bauingenieurwesen übertragen. Alternativ zur digitalen Abschlussprüfung sehen die Autor:innen großes Potenzial in der Anwendung auf freiwillige Kurztests im Semesterverlauf.

Literatur

- Knoke, M. (2021). Fluch und Segen der digitalen Prüfung. Hochschulforum Digitalisierung. Abgerufen am 14.03.2025 von <https://hochschulforumdigitalisierung.de/fluch-und-segen-der-digitalen-pruefung/>
- Ulrich, I. (2016). Gute Lehre in der Hochschule. Praxistipps zur Planung und Gestaltung von Lehrveranstaltungen. Wiesbaden: Springer Fachmedien.

Aktivität von Studierenden einfordern – Ein Erfahrungsbericht zu den Herausforderungen interaktiver (teil)digitaler Lehre

Jochen Gönsch

Die Veranstaltung „Software Skills“ wird kontinuierlich mit interaktiven, digitalen Unterstützungsangeboten angereichert, die dazu beitragen sollen, das Lernen der Studierenden zu unterstützen und ihnen Feedback über ihre Leistungen anzubieten. Letztlich dienen die teilweise gamifizierte Elemente auch dazu, sich optimal auf die Klausurprüfung am Ende des zweiten Semesters vorzubereiten und die Bedeutung des kompetenten Umgangs mit Software für den späteren professionellen beruflichen Alltag zu begreifen. Nachfolgend werden die Entwicklung der Lehr-Lernangebote, der Prüfung sowie die Erfahrungen und Herausforderungen in der Entwicklungsarbeit dargestellt.

1 Rahmen und Inhalte der Lehrveranstaltung

In den Bachelorstudiengängen Betriebswirtschaftslehre (BWL), Wirtschaftspädagogik und Kulturwirt bildet die Veranstaltung „Software Skills“ (SWS) gemeinsam mit der Veranstaltung „Operations Research“ (OR) ein Pflichtmodul. Es findet im 2. Bachelorsemester statt und muss spätestens im dritten Versuch bestanden werden. Das gesamte Modul wird mit einer Papierklausur geprüft.

Dabei entspricht der Anteil von „Software Skills“ zwei Credits. Ziel der Veranstaltung „Software Skills“ ist es, die Studierenden in die Lage zu versetzen, grundlegende, in Operations Research erlernte, Verfahren softwaregestützt auf praxisrelevante Problemstellungen anzuwenden. Das Team aus Professor (Vorlesung), Mitarbeiterin (Übung) und Hilfskräften (Tutorien) verwendet beispielhaft die weithin verfügbaren Standardsoftwares *OpenOffice-Base* (Datenbank), *Excel* und *Excel-VBA* (Programmiersprache). Grundlegende Inhalte sind Abfragen aus einer Datenbank in der Datenbanksprache SQL, Berechnungen und Funktionen in Excel, Nutzung des Excel-Solvers sowie die Formulierung kleiner Programme in Excel-VBA. Diese Kenntnisse sind für das weitere Studium, aber auch für den späteren Berufseinstieg, außerordentlich hilfreich. So ermöglichen sie beispielsweise den Einstieg in die automatisierte Datenverarbeitung („Big Data“). Damit werden auch die Grundlagen für spätere Veranstaltungen zum immer wichtigeren Thema Künstliche Intelligenz (KI) gelegt.

Die Veranstaltung „Software Skills“ findet jedes Sommersemester für Bachelor-Studierende der Studienrichtungen BWL (Jahrgangsgröße ca. 400) sowie zusätzlich für einige Dutzend Studierende der Wirtschaftspädagogik und Kulturwirtschaft statt. Für Studierende, die die Prüfung wiederholen müssen oder im Wintersemester schreiben möchten, werden viele Bestandteile (siehe Details im folgenden Abschnitt) auch im Wintersemester angeboten. Dort liegen die Teilnehmendenzahlen bei 15–30 % des Sommersemesters, das wir im Folgenden fokussieren.

Das Modul „Operations Research und Software Skills“ ist eines der wenigen Module im Bachelorprogramm, welches aus zwei Fächern besteht.

Inhaltlich ist das Fach SWS anwendungsorientiert ausgerichtet. Die Studierenden sollen Aufgaben selbst am Computer lösen und Varianten ausprobieren. Das setzt zwingend einen Computer voraus; viele unserer Online-Angebote mindestens ein Tablet mit großem Display. Im Wintersemester 2022/23 haben wir die Teilnehmenden einer Pflichtklausur im ersten Semester („Beschaffung und Produktion“) befragt, welche Endgeräte ihnen zur Verfügung stehen (Abbildung 1).

Konkret wurden abgefragt (in Klammern Anteil der Teilnehmenden): Smartphone (97 %), Tablet (58 %), Tablet mit Tastatur (24 %), Notebook Windows (25 %), Notebook Mac (38 %), Stationärer PC Windows (29 %) und Stationärer PC Mac (3 %). An der Umfrage nahmen 217 Personen von 269 Klausurteilnehmenden teil. Wichtig war uns zum einen, dass die Umfrage gerade nicht elektronisch oder in der Veranstaltung SWS durchgeführt wurde, damit wir auch

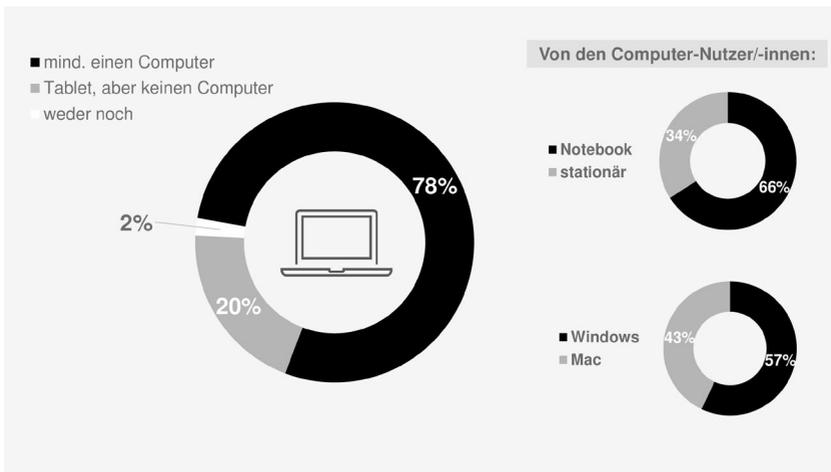


Abbildung 1: Endgeräte der Klausurteilnehmenden im Wintersemester 2022/23, Befragung der Klausurteilnehmenden der Klausur „Beschaffung und Produktion“ im Wintersemester 2022/23 (Anzahl der Befragten: 217) (Quelle: Eigene Abbildung)

Studierende erreichen, die sich aktuell vielleicht gerade wegen ihrer Hardwareausstattung gegen einen Besuch der Veranstaltung entscheiden. So zeigte sich, dass 78 % über mindestens einen Computer verfügen, 20 % über ein Tablet, aber keinen Computer, und 2 % weder noch. Bei den Computern sind zwei Drittel Notebooks und etwas mehr als die Hälfte Windows. In 2024 wiederholten wir diese Abfrage in der Präsenzvorlesung und erhielten ähnliche Ergebnisse. Das bedeutet, dass etwa 20 % der Studierenden auf die PC-Pools angewiesen sind, deren Verfügbarkeit unseres Wissens gut ist. So gibt es beispielsweise am Campus Duisburg (wo wir unterrichten) zahlreiche Arbeitsplätze mit der notwendigen Softwareausstattung, von denen auch praktisch immer einige frei sind.

2 Lehrkonzept & Gamification

Die Vorlesung wird seit dem Sommersemester 2017 gehalten und wird kontinuierlich weiterentwickelt. Dabei war von Anfang an wichtig, dass die Veranstaltung „Software Skills“ so konzipiert wird, dass die Studierenden mit Hilfe des Skripts und der Vorlesung den Inhalt verstehen und nachvollziehen können. Alle weiteren Elemente (siehe Abbildung 2), im *sich stetig erweiternden Lehrkonzept*, bieten zusätzliche Unterstützung für das Lernen der Studierenden, die sie bei Bedarf nutzen können. Unser Ziel ist, dass die Studierenden die Inhalte auch auf neue Aufgabenstellungen anwenden können. Dazu muss man nach unserer Einschätzung mit den Programmen experimentieren, Varianten der Lösungen ausprobieren. Gleichzeitig ist dies ein Lernansatz, der für unsere Studierenden eher ungewohnt ist. Wir müssen also erst einmal die notwendige Motivation erzeugen, Zeit und Kraft über das Semester hinweg in das Fach zu stecken. Und die Studierenden müssen diese aufbringen können.

Im Sommersemester 2017 stellten wir den Studierenden während der Präsenzvorlesung Kontrollfragen zum Lernfortschritt mit Hilfe des Umfragetools Mentimeter, beantworteten zunächst nur Fragen von Studierenden in einem Moodle-Forum (gefragt wurden hauptsächlich organisatorische und Verständnisfragen) und stellten Aufgabenblätter zu den einzelnen Themen zur Verfügung, deren Ergebnisse mit Hilfe von Selbsttests in Moodle überprüft werden konnten. Erweitert wurde dieses Angebot in den darauffolgenden Semestern durch 22 Bildschirmvideos mit Interaktionselementen (H5P), welche die einzelnen Schritte im jeweiligen Programm zum Lösen der im Skript enthaltenen Beispielaufgaben aufzeigen und auf einen entsprechenden Wunsch der Studierenden zurückgehen. Interaktion bedeutet hier, dass das Video an vorgegebenen Stellen stoppt und Multiple-Choice-Fragen beantwortet werden müssen, etwa dazu, welches Element als nächstes in der gezeigten Programmoberfläche angeklickt werden muss. Ein weiteres Ausbauelement ist ein Wiki, in dem wir ein Glossar von Fachbegriffen und teils auch weitergehende Hintergrundinformationen (die aber für den Lernerfolg nicht notwendig sind) bereitstellen. Dazu gibt es Tutorien in



Abbildung 2: Bestandteile der Veranstaltung „Software Skills“ (Quelle: Eigene Abbildung)

Kleingruppen (Präsenz und Online) und Sprechstunden, die jeweils im Sommer sowie Wintersemester (für Wiederholer:innen) angeboten werden.

Die Vielzahl von unterschiedlichen Elementen in Moodle wurde in ein *ganzheitliches Gamification-Konzept* eingebunden, welches für die einzelnen Elemente wie die Diskussionsforen, das Wiki, die interaktiven Bildschirmvideos und die Selbsttests einen Rahmen bildet. Hierbei werden spielerische Elemente berücksichtigt. So können die Studierenden durch Aktivitäten in den Foren, das Lösen der Selbsttests oder interaktive Elemente (HSP) in den Bildschirmvideos *Erfahrungspunkte sammeln*. Durch die Erfahrungspunkte können sie immer neue Level erreichen. Ein weiteres spielerisches Element ist die Suche nach sog. *Stickers* im Moodle-Kurs. Diese Suche soll den Studierenden den Umfang und die Möglichkeiten des Moodle-Kurses deutlich machen. Sie sollen mit Hilfe einer anonymen Rangliste aller Studierenden, der Anzeige des Entwicklungsstands im Dashboard und der erfolgreichen Absolvierung einzelner Elemente mit grünen Häkchen motiviert werden, die einzelnen Elemente zu bearbeiten. Dabei war es uns wichtig, dass die Studierenden ausschließlich eine Plattform nutzen. Daher sind die Beschreibung und alle Elemente des Gamification-Konzepts sowie alle organisatorischen Inhalte in einem gut strukturierten Moodle-Kurs zu finden.

Die Studierenden können durch die Bearbeitung der unterschiedlichen Aufgaben bis zu zwei Bonuspunkte für die abschließende Klausur erreichen (8 % der Gesamtpunkte in „Software Skills“). Im Sommersemester 2022 kam eine Übung hinzu, die den Einsatz der unterschiedlichen Softwareprogramme für *praxisnahe Datenmengen* verdeutlicht. Des Weiteren gibt es seitdem im ersten Veranstaltungstermin einen *Praxisvortrag durch eine:n Unternehmensvertreter:in*, der den

Einsatz eines in der Vorlesung genutzten Programms zeigt.

Das gesamte Lehrkonzept wird kontinuierlich getestet (freiwillige jährliche universitätsweite Lehrevaluation, Umfragen zu Veranstaltungselementen in Moodle und/oder Mentimeter, Gruppendiskussionen) und weiterentwickelt. Hierbei hat uns das Zentrum für Hochschulqualitätsentwicklung (ZHQE) der UDE in vielfältiger Weise unterstützt.

3 Teilnahmeverhalten und Prüfungserfolge der Studierenden

Studierende berichten uns regelmäßig, dass das Modul das anspruchsvollste im Bachelor BWL sei. Dazu beobachten wir eine (für eine Pflichtveranstaltung im zweiten Semester) geringe Teilnahme an der Prüfung sowie vergleichsweise niedrige Bestehensquoten. Das Modul stellt also eine erhebliche Hürde im Studienverlauf dar.

Zunächst lässt sich festhalten, dass die Teilnahme an der Vorlesung im Semesterverlauf erst ab- und dann wieder zunimmt (Abbildung 3).

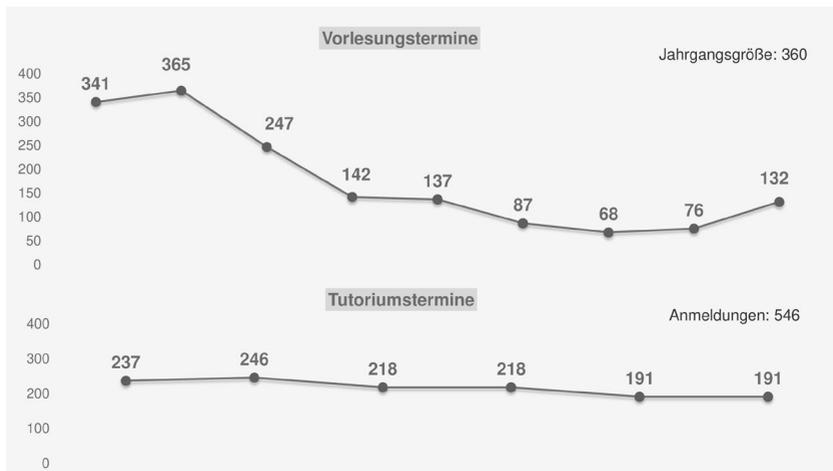


Abbildung 3: Entwicklung der Teilnehmendenzahlen im Sommersemester 2023
(Quelle: Eigene Abbildung)

Abbildung 4 stellt Klausuranmeldungen und -erfolg sowie Teilnahme am Bonuspunkteprogramm im Sommersemester 2023 dar.

Von den ursprünglich 1017 für die Klausur angemeldeten Studierenden nahmen an dem Klausurteil SWS nur knapp 20 % teil. Die Bestehensquote betrug 59 %. Von den 199 zur Klausur erschienen Studierenden hatten 167 Studierende

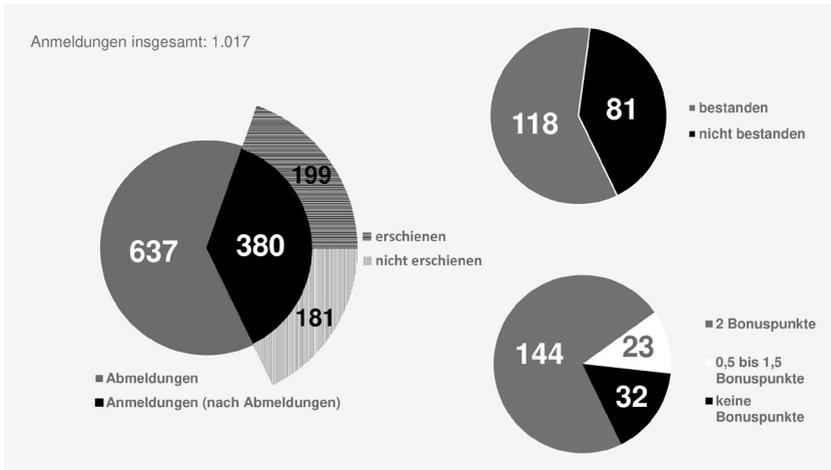


Abbildung 4: Daten zum Klausurteil „Software Skills“ (SWS) im Sommersemester 2023 (Quelle: Eigene Abbildung)

zuvor Bonuspunkte erworben. Die Korrelation zwischen Anzahl der Bonuspunkte und erreichten Punkten im SWS-Klausurteil betrug 0,3. Umgekehrt haben von den nicht erschienenen 818 Studierenden nur 54 mindestens 0,5 Bonuspunkte gesammelt.

Dass die Anmeldezahlen deutlich über den Teilnehmezahlen an Klausuren liegen, ist bei uns üblich und vermutlich darauf zurückzuführen, dass Anmeldungen nur zu Beginn des Semesters, Abmeldungen aber bis eine Woche vor der Klausur möglich sind, zurückzuführen. Die Teilnehmenden haben auch (fast) alle Bonuspunkte erreicht, ihre Zahl entspricht auch recht genau der Teilnahme an den Tutorien.

Weiter sehen wir (hier nicht dargestellt), dass Studierende häufig die Klausur im 2. Studiensemester erstmals nicht bestehen (im Sommersemester rund 41 %). Der nächste Versuch folgt jedoch oft erst im 6. oder 7. Semester, sodass diese Klausur bei vielen Studierenden als eine der letzten im Studienverlauf bestanden wird (oder, vergleichsweise selten, im 3. Versuch endgültig nicht bestanden wird).

4 Studentisches Feedback zur Veranstaltung und zur Klausur

Das Feedback der Studierenden zu der Veranstaltung zeigt, dass die Vorlesung und die unterstützenden Angebote insgesamt sehr positiv wahrgenommen werden. Bei den einzelnen Elementen sind die Nutzungsquoten unterschiedlich (Abbildung 5), jedoch wird jedes Element mindestens von der Hälfte der Befragten genutzt, teils

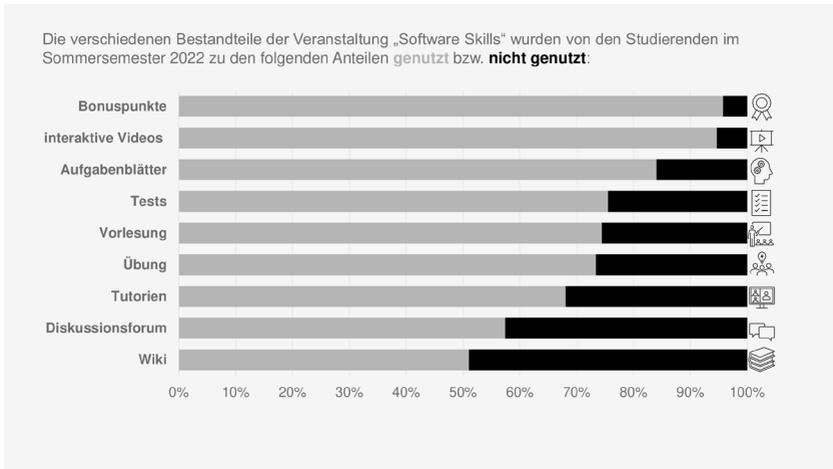


Abbildung 5: Teilnehmendenzahlen nach Elementen der Veranstaltung „Software Skills“, Studentische Lehrveranstaltungsbeurteilung (LVB) im Sommersemester 2022 (Anzahl der Antworten: 94), (Quelle: Eigene Abbildung)

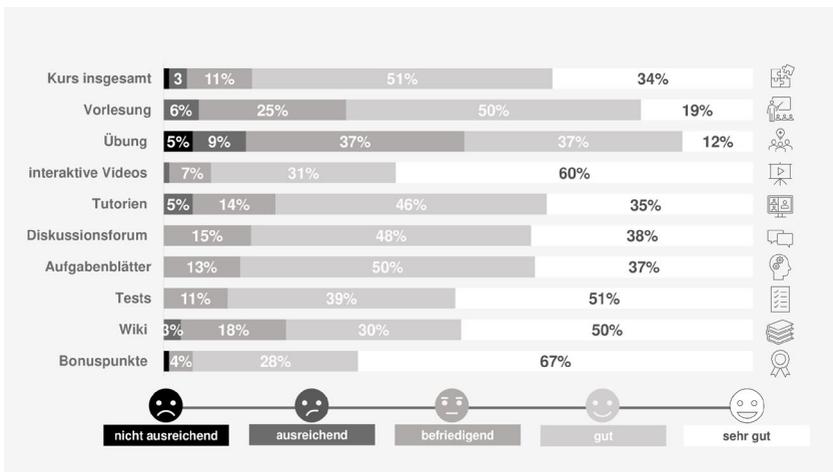


Abbildung 6: Beurteilung der verschiedenen Elemente der Veranstaltung „Software Skills“, Studentische Lehrveranstaltungsbeurteilung (LVB) im Sommersemester 2022 (Anzahl der Befragten: 94), (Quelle: Eigene Abbildung)

von fast allen. Die Zufriedenheit der Studierenden mit den Elementen der Veranstaltung ist durchweg hoch (Abbildung 6), wobei Vorlesung (nur 69 % gut oder sehr gut) und Übung (nur 49 % gut oder sehr gut) etwas abfallen.

Im Sommersemester 2022 führten wir im Rahmen von PITCH eine Prüfungsevaluation nach der Klausur durch. Hier nahmen 80 von 285 Klausurteilnehmenden teil. Der Rücklauf beträgt entsprechend 28 %. Von den antwortenden Studierenden fühlte sich nur die Hälfte vor der Klausur gut vorbereitet. Gleichzeitig gibt in Bezug auf Vorlesung, Übung und Tutorium jeweils die Hälfte an, immer oder fast immer teilgenommen zu haben. Nicht abgefragt wurde die eigene Vor- und Nachbereitung. Die Prüfungsfragen scheinen insgesamt verständlich gestellt und die Bearbeitungszeit wurde als angemessen empfunden. Die übrigen Multiple-Choice-Fragen sind für uns ohne intertemporalen Vergleich schwer zu interpretieren.

Die Freitextkommentare kritisierten etwa zur Hälfte eine Frage zu einem in den Organisationsfolien als klausurrelevant angekündigten Gastvortrag. Offensichtlich war es uns nicht gelungen, seine Relevanz zu transportieren.

5 Überlegungen zur Weiterentwicklung der Lehrveranstaltung

Uns stellt sich nun die Frage, inwiefern wir die Veranstaltung so weiterentwickeln können, dass noch mehr Studierende das Lernziel erreichen, und das vielleicht sogar in weniger Klausurversuchen. Hier erscheint uns ein zentraler Aspekt das Engagement der Studierenden.

Dabei sind die grundlegenden Argumente bekannt: Studierende sind nicht wie Lehrende. Wir müssen auf ihre sozialen, motivationalen, organisatorischen und technischen Bedürfnisse und Rahmenbedingungen eingehen.

Wird die parallele Aufbereitung von Materialien in unterschiedlichen Formaten vorgeschlagen (siehe z. B. Auferkorte-Michaelis/Haschke, 2022), so kommen Lehrenden oft schnell die eigenen begrenzten zeitlichen Ressourcen in den Sinn. Allerdings bewerten auch Studierende eine größere Materialvielfalt keineswegs durchweg positiv. Egal, wie die Aufbereitung ist, es kostet immer mentale Ressourcen, sich einen Überblick zu verschaffen – vermutlich noch verstärkt durch das Phänomen des fear-of-missing-out. Auch ist die Grenze zwischen Inklusion und kleinster gemeinsamer Nenner oft fließend.

Schließlich, und das wird seltener diskutiert, beeinflussen Aufbereitung und Vermittlung immer auch das faktische Lehrziel oder besser gesagt, das, was Studierende leisten müssen, um die Prüfung (gut) zu bestehen.

Ein Klassiker ist der Wunsch nach asynchronen, distanten Formaten. Ihre Vorteile liegen auf der Hand und werden von Studierenden gerne angeführt. Auf leistungsstarke, insbesondere engagierte Studierende trifft das sicherlich zu. Bei den anderen gibt es unserer Erfahrung nach eine große Gruppe, die eher Nachteile erfährt. Ihnen ist es nicht möglich, kontinuierlich mitzuarbeiten, sodass sie

entweder irgendwann aufgeben oder versuchen, kurz vor der Klausur zu lernen. Studierende äußern hier als Gründe meist äußere Zwänge, im vertraulicheren Gespräch auch die „normale Prokrastination“. Schon die eigene Erfahrung zeigt, dass eine Anwesenheits- oder Abgabepflicht (beispielsweise von Übungsaufgaben) – sowohl von sich selbst wie auch dem Umfeld – viel einfacher akzeptiert wird. Klar ist: Präsenzlehre ist aufwendiger, aber Lernen ist eben aufwendig. Eming und Philipowski (2022) haben das sehr schön diskutiert.

Auch die Schaffung eines Gruppengefühls ist asynchron deutlich schwieriger. Und schließlich haben sich die Studierenden – hoffentlich – bewusst für ein Vollzeitstudium an einer Präsenzuniversität entschieden. Die Nachteile anderer Formate wurden im Zusammenhang mit der Coronapandemie ausführlich diskutiert.

In SWS kommt auch nach Corona die Hardwareausstattung der Studierenden hinzu. Aktuell setzt die Veranstaltung einen Computer voraus, um selbst mitzuarbeiten. Während wir hier anfangs nur Windows supporten konnten, wird macOS inzwischen gleichwertig unterstützt. Aber 20 % der Studierenden verfügen nicht über derartige Geräte. Häufig werden in dieser Hinsicht knappe finanzielle Mittel in der Literatur diskutiert. In vertraulichen Gesprächen äußern Studierende und Hilfskräfte eher, dass ein Computer nur für SWS nötig sei. In allen anderen Fächern arbeiten sie mit dem Tablet, auch die Bachelorarbeit lasse sich mit Tablet und Tastatur schreiben.

Anerkennen müssen wir wohl, dass sich Studierende nicht wegen uns einen PC kaufen werden und gerade die ohne PC auch häufig nicht für SWS in den PC-Pool gehen werden. Alternativen, sei es webbasiert oder für Tablets, sind jedoch deutlich unkomfortabler und/oder eingeschränkter als die von uns verwendeten Standard-Office-Programme.

Von Unternehmen werden (fortgeschrittene) PC-Kenntnisse nach wie vor sehr wertgeschätzt, aber offenbar nicht mehr zwingend vorausgesetzt, zumindest bei Bachelorabsolvent:innen.

6 Fazit

Wir haben uns entschieden, bei Office-PC-Programmen zu bleiben, werden aber in Zukunft eine komfortablere Datenbanksoftware (LibreOffice) nutzen.

Natürlich müssen wir uns immer auch fragen, woran es liegt, wenn Lernangebote nicht angenommen werden. Hier ist die Bewertung der Teilnehmenden sehr positiv, aber natürlich auch nur eingeschränkt aussagekräftig. In der genannten Nach-Klausur-Befragung wurde die Abstimmung der Prüfungsfragen auf die Veranstaltungsinhalte als mittel bewertet. Das ist schwierig zu interpretieren, weil wir ja bewusst eine anspruchsvolle Klausur mit immer neuen Fragen stellen (so muss beispielsweise manchmal Code vervollständigt, korrigiert, oder selbst geschrieben werden), was auch so angekündigt wird. Vom Wissen/Fähigkeiten her wird genau das Behandelte abgefragt.

Von daher bleiben wir bei dem Leitbild, die Veranstaltung für mitarbeitende, teilnehmende Studierende zu gestalten.

Unser Lehrziel – und auch der Prüfungsgegenstand – ist das Erledigen neuer, unbekannter Aufgaben am PC mit den behandelten Mitteln. Ob und wie sie dieses Ziel zu erreichen versuchen, können die Studierenden selbstverantwortlich entscheiden. Einigen ist der gesamte Stoff schon aus der Schule bekannt. Wir können (und dürfen) hier nur Angebote machen. Und die so gut wie möglich.

Literatur

- Auferkorte-Michaelis, N./Haschke, H. (2022). „Ich gehe online, wer kommt mit?“ – Feed_in Befragungen vor Veranstaltungsbeginn. In: Diversität konkret: Handreichungen für das Lehren und Lernen an Hochschulen, 02/2022. KomDiM, Duisburg-Essen. https://www.uni-due.de/diversity/prodiversitaet_diversitaetkonkret.shtml
- Eming, J./Philipowski, K. (2022, 11. Mai). Wie Corona die akademische Lehre dauerhaft verändert. In: Forschung und Lehre. <https://www.forschung-und-lehre.de/lehre/wie-corona-die-akademische-lehre-dauerhaft-veraendert-4678>

Aspekte des digitalen Lehrens und Prüfens: Überblick und neuere Entwicklungen im Fach Statistik

Christoph Hanck und Jens Klenke

1 Einleitung

Der Lehrstuhl für Ökonometrie der wirtschaftswissenschaftlichen Fakultät der Universität Duisburg-Essen (UDE) nutzt in den Bachelorgrundlagenveranstaltungen *Deskriptive* und *Induktive Statistik* seit 2014 das E-Assessment-System JACK und übernimmt dabei einen aktiven Part bei der Weiterentwicklung der Umgebung. In beiden Modulen wird JACK sowohl formativ in Form eines Begleitkurses sowie summativ in Testaten und Klausuren verwendet. Der Vorteil von E-Assessment-Systemen in der „modernen“ Lehre von Statistik ist erheblich. Lehrende können in JACK sowohl theoretische Konzepte vermitteln als auch praxisnahe Anwendungsbeispiele bieten, welche mit einer Programmiersprache gelöst werden. Dass Studierende kein externes Programm installieren müssen, erleichtert den Einstieg.

Seit dem Wintersemester 2022/2023 nutzen wir die neueste Entwicklungsstufe JACK3. In dem Semester haben in der *Deskriptiven Statistik* insgesamt 379 Studierende JACK3 genutzt. Im darauffolgenden Sommersemester (SoSe 2023) verwendeten 392 Studierende JACK im Modul *Induktive Statistik*.

In Abschnitt 2 stellen wir die Ausgestaltung unseres digitalen Lehrangebots vor. Anschließend beschreiben wir unsere Erfahrungen beim digitalen Prüfen und skizzieren unsere Prüfungsplanung. Kapitel 4 beinhaltet eine Datenanalyse des Kurses *Induktive Statistik* SoSe 2023. Die Analyse zeigt, dass frühzeitiges und regelmäßiges Lernen entscheidend für den Klausurerfolg ist. Dadurch kann der Klausurerfolg von Studierenden relativ früh im Semester mit einer hohen Genauigkeit vorhergesagt werden. Studierende, die sich erst relativ spät mit den Inhalten befassen, können die Klausur durch einen vergleichswisen hohen Arbeitsaufwand im letzten Semesterdrittel trotzdem bestehen. Der Beitrag schließt mit einem Fazit.

2 Ausgestaltung des digitalen Lehrangebots

Das Lehrangebot in JACK3 ist in Übungsaufgaben sowie in Testaten und Klausuren gegliedert. Ab der 7. Semesterwoche werden 14-tägig Online-Tests, soge-

nannte Testate, angeboten, mit denen Studierende Bonuspunkte sammeln können. Die Bearbeitungszeit beträgt jeweils 40 Minuten. Am Ende des Semesters wird auch die Prüfungsleistung in Form einer Klausur in JACK durchgeführt. Dies erfolgt unter Aufsicht in den PC-Halls der Universität.

JACK bietet die Möglichkeit, Aufgaben in Stages (Teilaufgaben) zu unterteilen, welche unterschiedliche Einreichungsformen haben können. Im Rahmen unserer Lehrveranstaltungen werden primär folgende Typen verwendet:

- *Fill-In-Stage*: Hier erfolgt die Einreichung als Zahl oder Text in einem oder mehreren vordefinierten Input-Feldern (vgl. Abbildung 1).
- *Multiple-Choice-Stage*: Die Einreichung besteht aus der Auswahl vorgegebener Antwortmöglichkeiten.
- *R-Stage*: Die Einreichung erfolgt als R-Code in einem Textfeld, genannt Konsole, mit dem die Software gesteuert wird. Studierende können ihren Code vor der Einreichung in JACK testen.

IS10-01 Konfidenzintervalle 1

Aktuelles Gesamtergebnis: 0%

Stage 1 —

Aufgabenstellung:

Sei $X \sim \mathcal{N}(\mu, \sigma^2 = 0.91)$.

Eine Stichprobe der Größe $n = 15$ ergibt: $\bar{x} = 9.53$.

Berechnen Sie das Konfidenzintervall für μ zum Konfidenzniveau $1 - \alpha = 0.91$.

Einreichung:

[,]

(Runden Sie, falls nötig, auf die vierte Nachkommastelle und verwenden Sie einen Dezimalpunkt.)



Abbildung 1: Beispiel einer Fill-In-Stage aus dem Begleitkurs für *Induktive Statistik* (Quelle: Eigene Abbildung)

In JACK werden bei jedem Aufgabenstart neue Zahlen generiert. Studierende können so nicht einfach Ergebnisse auswendig lernen, sondern müssen die Konzepte verstehen. Eingaben von Studierenden werden anhand von Evaluierungsregeln bewertet, bei denen u. a. Rundungstoleranzen berücksichtigt werden können. Neben der Hauptevaluierungsregel können zusätzlich optionale Feedbacks erstellt werden, um typische Fehler aufzugreifen und Punkte für teilrichtige Antworten zu vergeben. Optionales Feedback kann auch genutzt werden, um Folgefehler zu berücksichtigen, wenn die fehlerhafte Einreichung einer Stage in einer späteren Stage relevant ist. Die Evaluierung von *R*-Stages unterscheidet sich gegenüber den anderen Aufgabentypen. Der eingereichte Code wird anhand von zuvor angelegten *Checkern* evaluiert. *Statische Checker* bewerten Eingaben hinsichtlich Syntax und Semantik. *Dynamische Checker* bewerten den Code anhand des numerischen Endergebnisses.

Für jede Evaluation¹ der Stages lässt sich ein individueller Feedbacktext festlegen, der Studierende bspw. auf individuelle Fehler aufmerksam macht, Hilfestellungen gibt oder eine Musterlösung, z. B. für die Einsicht einer Klausur, ausgibt. Eine weitere intensiv genutzte Funktion von JACK3 ist das Bereitstellen von Lösungshinweisen, welche von den Studierenden genutzt werden können. Für Klausuren kann das Verwenden von Hinweisen bepreist werden. Die Bepreisung erfolgt in Prozent für die jeweilige Teilaufgabe. Hierbei kann der prozentuale Abzug von den möglichen oder den tatsächlich erreichten Punkten erfolgen.

3 Planung und Ablauf beim digitalen Prüfen: Ein Erfahrungsbericht

Der Ablauf von Kursen mit digitalen Lehrangeboten und Prüfungen unterscheidet sich von herkömmlichen Veranstaltungen. Bei der Planung der digitalen Angebote bedarf es größerer Vorlaufzeiten im gesamten Prozess. Insbesondere in den Anfängen des digitalen Lehrens und Prüfens gilt es, mögliche Risiken frühzeitig zu erkennen. Entscheidend hierfür ist das ausgiebige Testen von erstellten Inhalten.

Insgesamt haben wir im Allgemeinen sehr positive Erfahrungen mit der Digitalisierung unserer Lehre und Prüfungen gemacht. Der Umstieg auf die neueste Entwicklungsstufe – JACK3 – war, wie bei der Einführung wesentlicher technischer Neuerungen, nicht unüblich, von einigen Startschwierigkeiten begleitet. Eine weitere, zeitlich überschneidende, Herausforderung waren die Auswirkungen des Hackerangriffs auf die Informationssysteme der UDE im November 2022 (Vierjahn, 2022). Unter anderem waren zum Zeitpunkt der Prüfungen etliche Server und Dienste der Universität nicht verfügbar, wodurch z. B. die Einrichtungen der Computer manuell erfolgen mussten. Weiterhin waren Verbin-

¹ Die möglichen Evaluationen in JACK sind: richtig, falsch, optionales Feedback (Teillösungen) und überspringen.

dungen zu externen Servern nicht möglich, wodurch u. a. ein digitaler Taschenrechner in JACK nicht zur Verfügung stand.

3.1 Planung der Klausur

Digitale Prüfungen bergen viele Vorteile wie Zeitersparnis und ein einheitliches, vordefiniertes Bewertungsschema. Um diese vollumfänglich nutzen zu können, bedarf es, verglichen mit dem herkömmlichen Vorgehen, einer veränderten Planung. Unsere Erfahrungen zeigen, dass der Gesamtaufwand vor allem bei Prüfungen in Großveranstaltungen wesentlich geringer ist. Der Korrekturaufwand, welcher den Großteil des Aufwands bei herkömmlichen Klausuren ausmacht, entfällt bspw. bei digitalen Prüfungen weitestgehend. Ein Teil der Zeitersparnis muss bei der Klausurerstellung in die Entwicklung und das Testen der Evaluierungsregeln investiert werden.

Die Klausurplanung und -organisation sollte möglichst frühzeitig im Semester erfolgen. So kommunizieren wir die ersten (groben) Studierendenzahlen schon im Laufe der ersten beiden Semesterwochen mit den zuständigen Stellen², damit frühzeitig Entscheidungen über technische Aspekte erfolgen können und deren Umsetzung beginnen kann.

Während des Semesters haben Studierende die Gelegenheit, an Testaten teilzunehmen. Dies bietet die Möglichkeit, die Last auf den Servern zu beobachten und wichtige Erkenntnisse für die Prüfungsvorbereitung zu sammeln. Die Analyse zeigte, dass primär der Start kritisch ist. Zum einen müssen viele Ressourcen geladen werden und zum anderen arbeiten die Studierenden noch relativ synchron, was zu vielen Anfragen zum gleichen Zeitpunkt führt.

Gegenüber herkömmlichen Prüfungen erfolgt die Erstellung der Klausur wesentlich früher im Semester. Insgesamt wird für die Erstellung und Testung, mit zeitlichem Puffer, acht bis zehn Wochen Vorbereitungsaufwand eingeplant. Als Beispiel sei hier die Zeitachse für den 1. Prüfungstermin für *Induktive Statistik* des SoSe 2023 (28.7.2023) erläutert. Die Klausuren werden Studierenden gegenwärtig zusätzlich in gedruckter Form ausgeteilt.³ Daher müssen die Aufgabenstellungen zum Drucktermin (03.07.2023) in der finalen Version stehen. In den ersten beiden Wochen dieser Periode werden die Aufgaben in JACK implementiert. Vor allem bei der Programmierung der Aufgaben kann es immer wieder dazu kommen, dass einzelne Teilaufgaben angepasst werden müssen, um technische Gegebenheiten zu berücksichtigen.

2 Für JACK besteht ein direkter Ansprechpartner beim Ruhr Institute for Software Technology (Paluno), den Entwicklern des E-Assessment-Systems JACK.

3 Dies erfolgt noch aus zwei Gründen. Studierenden soll ein besserer Überblick über die Klausur gewährt werden. Zum anderen umfasst die Klausur eine handschriftliche Aufgabe. In Zukunft wollen wir auf Papierklausuren in Gänze verzichten.

Circa zwei Wochen vor der Druckfrist testen alle zur Verfügung stehenden Mitarbeiter:innen des Lehrstuhls die Klausur. In dieser Testrunde stehen zwei Aspekte im Mittelpunkt: Erstens, dass die Aufgabenstellung in der Klausur eindeutig ist und zweitens die Aufgaben technisch einwandfrei funktionieren. Dies beinhaltet vor allem, dass alle Folgefehler bei der Programmierung beachtet wurden und die Aufgabenstellung so gestellt ist, dass jede Stage auch bearbeitet werden kann, wenn Stages zuvor nicht oder nicht richtig gelöst wurden. Zusätzlich achten alle Mitarbeiter:innen während der Testphase darauf, dass die Hinweise hilfreich und alle möglichen und sinnvollen Teillösungen implementiert sind.

Die finale Prüfungsanmeldezahl erhalten wir sechs Tage vor der Klausur. Jedoch kann schon ab der vierten Semesterwoche⁴ die jeweils aktuelle Anmeldezahl beim Dekanat erfragt werden. Wichtig ist, dass die maximale Anzahl unmittelbar mit der IT kommuniziert wird, damit frühzeitig genügend Ressourcen eingerichtet werden können.

Nachdem die Klausur auf alle Server kopiert wurde, werden im nächsten Schritt die Serververteilungseinstellungen vorgenommen, um die Last möglichst gleichmäßig zu verteilen. Die Verteilung erfolgt über individuelle Links im Moodle-System. Zusätzlich werden die Studierenden ausschließlich auf den entsprechenden Servern freigeschaltet. In den letzten Wochen vor dem Prüfungstermin sollte die Klausur außerdem direkt in den PC-Halls getestet werden.

3.2 Ablauf am Klausurtag

Am Tag der Klausur werden abermals alle möglichen Fehlerquellen überprüft. In Moodle werden dementsprechend alle Links und Zeiteinstellungen für die einzelnen Klausurgruppen geprüft. Anschließend werden noch einmal alle Zeit- und Zugangsbeschränkungen auf den verschiedenen JACK-Servern kontrolliert. Außerdem sollten alle anderen Kursangebote, wie bspw. der Begleitkurs, offline sein. Sobald die PC-Halls betreten werden können, sollten auf einigen Clients noch einmal die Einstellungen getestet werden.

4 Kursanalyse und Vorhersagemodell

4.1 Kursübersicht *Induktive Statistik SoSe 2023*

In der Veranstaltung *Induktive Statistik* im SoSe 2023 wurde ein breites digitales Lehrangebot in JACK angeboten. Im Laufe des Semesters haben insgesamt 392 Studierende Aufgaben im Begleitkurs bearbeitet. 344 Studierende haben an

⁴ Hier endet die Anmeldephase für den 1. Prüfungstermin.

mindestens einem Testat teilgenommen. Den 1. Prüfungstermin haben 174 Studierende wahrgenommen, 86 den 2. Prüfungstermin, wobei 34 Studierende an beiden teilgenommen, aber die 1. Prüfung nicht bestanden haben.

4.2 Analyse des Lernverhaltens

Abbildung 2 zeigt die tägliche Anzahl an Einreichungen im Begleitkurs für *Induktive Statistik*. Der Kurs wurde am 2. Übungstermin freigeschaltet und war bis zum Tag des zweiten Prüfungstermins online. Das Verhalten der Studierenden ist vergleichbar mit früheren Auswertungen (Massing et al., 2018). Die Einreichungen schnellen an Tagen vor Testaten und Klausuren hoch. Zwischen den Testaten flaut die allgemeine Aktivität ab, wobei der Trend bis zum ersten Prüfungstermin positiv ist. Darüber hinaus lässt sich beobachten, dass die Anzahl an aktiven Studierenden bis zum 1. Testat steigt, nach dem Testat relativ stark abnimmt und anschließend auf einem gleichbleibenden Niveau bis zur Klausur verbleibt.

Die Testate führen dazu, dass Studierende sich frühzeitig und kontinuierlich mit den Inhalten auseinandersetzen. Ein frühes und kontinuierliches Lernverhalten ist besonders in Veranstaltungen, welche thematisch aufeinander aufbauen, von enormer Bedeutung. Dies trifft sicherlich allgemein zu, aber insbesondere auf Einführungsveranstaltungen in Statistik.

Interessant ist, ob ein erhöhter Aufwand im Zusammenhang mit der Prüfungsleistung steht. Hierfür wurden die einzelnen Notenstufen in fünf verschiedene Notengruppen⁵ zusammengefasst. Studierende, die keine Prüfung absolviert haben, bilden eine eigene sechste Gruppe. Die Analyse beruht nicht auf den an das Prüfungsamt übermittelten Noten, sondern den theoretischen Noten der Studierenden, welche sie ohne Bonuspunkte und Verschiebung der Bestehensgrenze⁶ erreicht hätten. Dies erfolgt aus zwei Gründen: Erstens ermöglicht es, die Ergebnisse mit früheren Untersuchungen zu vergleichen und zweitens wäre es verzerrend, die Prüfungsleistung mit den Testatsergebnissen zu erklären, wenn diese selbst Teil der Noten sind.

Tabelle 1 zeigt, dass sich der in JACK beobachtbare Aufwand deutlich für die einzelnen Notengruppen unterscheidet und dass ein positiver Zusammenhang zwischen Aufwand und Prüfungsleistung besteht. Diese Vermutung wird durch einen Rangkorrelationskoeffizient nach Spearman von 0,9429 und einem p-Wert kleiner als 5 % gestützt. Den größten Aufwand haben die Studierenden der Notengruppe „gut“ betrieben, welcher ca. 1,5 Einheiten über dem Schnitt

5 Die Notengruppen sind wie folgt definiert: „sehr gut“ – 1,0 und 1,3; „gut“ – 1,7, 2,0 und 2,3; „befriedigend“ – 2,7, 3,0 und 3,3; „ausreichend“ – 3,7 und 4,0; „nicht ausreichend“ – 5,0.

6 Die Bestehensgrenze wurde von 25 auf 21 von 60 Punkten für beide Klausuren herabgesetzt, da während des ersten Prüfungstermins technische Probleme bei der Verbindung zu einem externen Server auftraten.

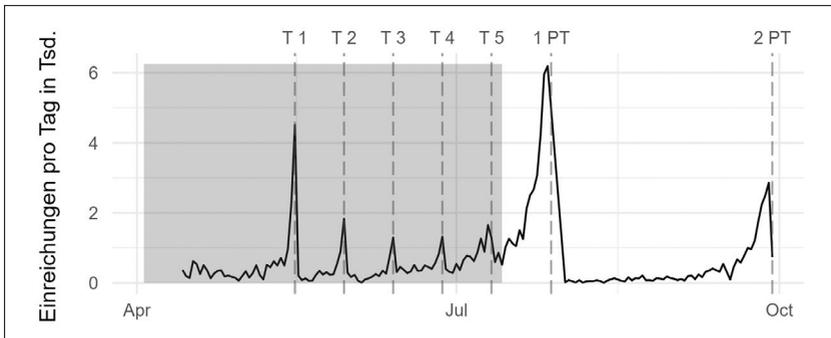


Abbildung 2: Gesamteinreichungen pro Tag im Begleitkurs *Induktive Statistik* auf JACK. Die eingefärbte Fläche spiegelt die Vorlesungszeit wider. T 1 bis T 5 indizieren die fünf Testate und 1 PT bzw. 2 PT die Prüfungstermine (Quelle: Eigene Abbildung)

der „sehr guten“ Studierenden liegt. Der Unterschied zwischen den Notengruppen „befriedigend“ und „ausreichend“ ist relativ gering. Auch der Unterschied zu den Studierenden, die die Prüfung nicht bestanden haben, beträgt nur circa 0,5 Einreichungen pro Tag. Ein Wert von 0,85 bei den Studierenden, die keine Prüfung abgelegt haben, verdeutlicht, dass sie nur sporadisch am Kurs teilgenommen haben.

Notengruppe	Anzahl Studierende	Aufwand in JACK (Einreichungen pro Tag)
Sehr gut	5	5,55
Gut	16	7,13
Befriedigend	39	4,00
Ausreichend	34	3,74
Nicht ausreichend	126	3,18
Keine Prüfung	172	0,85

Tabelle 1: Verteilung der Studierenden über den theoretischen Notengruppen und der durchschnittliche Aufwand pro Tag in JACK.

Neben dem beobachtbaren Aufwand ist die Güte der Einreichungen ein weiterer Faktor für das Bestehen der Klausur. Um dies zu überprüfen, wurde für jeden Studierenden ein wöchentlicher Score berechnet, definiert als Quotient aus der Summe der Punkte der jeweils letzten Einreichung jeder Teilaufgabe und der maximal erreichbaren Punktzahl über alle Aufgaben. Durch Normierung liegt der Score zwischen 0 und 1.

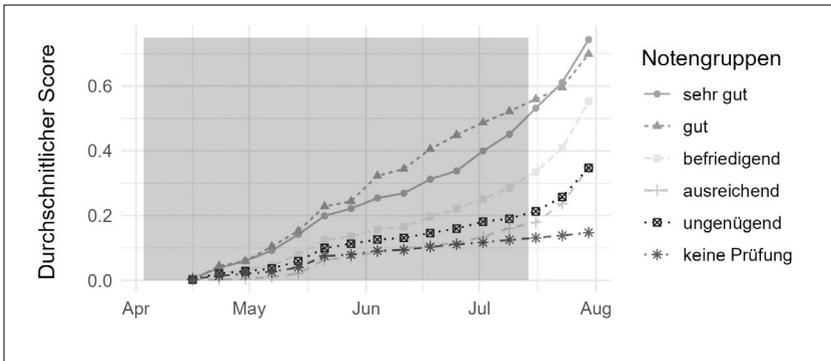


Abbildung 3: Verlauf des durchschnittlichen Scores im Semester nach Notengruppen. Die eingefärbte Fläche spiegelt die Vorlesungszeit wider (Quelle: Eigene Abbildung)

Abbildung 3 beschreibt den Verlauf des Scores über das Semester für die verschiedenen Notengruppen und die Studierenden, welche am Ende des Semesters keine Prüfung abgelegt haben. Aus der Grafik lassen sich interessante Schlüsse ziehen. Die durchschnittlichen Scores der Gruppen „sehr gut“, „gut“ und „befriedigend“ setzen sich verhältnismäßig früh im Semester von den anderen Notengruppen ab. Die Scores der Notengruppen „ausreichend“ und „nicht ausreichend“, also jene, die knapp die Klausur bestehen, und jene, die die Klausur nicht bestehen, sind vom Start des Semesters bis Mitte Juni auf gleichem Niveau. Erst in den letzten 6 Wochen setzt sich der durchschnittliche Score der Notengruppe „ausreichend“ ab. Dies bedeutet, dass die Klausur auch durch relativ spätes Lernen bestanden werden kann. Abbildung 3 suggeriert jedoch, dass eine gute Note für das Gros der Studierenden nur durch kontinuierliches Lernen möglich ist. Der vermehrte Aufwand, den die Studierenden der Gruppe „ausreichend“ in den letzten Wochen vor der Klausur erbracht haben, könnte zudem auch zulasten anderer Fächer gehen.

4.3 Vorhersage des Klausurbestehens

Ziel dieses Abschnitts ist es, das Bestehen der Klausur möglichst früh zuverlässig vorherzusagen. Dies kann genutzt werden, um Studierende, die Gefahr laufen, die Klausur nicht zu bestehen, in künftigen Kursen auf Grundlage dieser Modelle zu warnen (Klenke et al., 2023). Zur Vorhersage wird wöchentlich ein Logit-Modell geschätzt, das die Wahrscheinlichkeit des Bestehens der Klausur anhand des Aufwands, Scores und der Testatsergebnisse schätzt. Der Aufwand ist definiert als Anzahl an eingereichten Teilaufgaben in einer Woche. Die Testatsergebnisse fließen in die Modelle ein, sobald diese zur Verfügung

stehen. Für Testate, an denen nicht teilgenommen wurde, wird das Ergebnis auf 0 gesetzt.⁷

Das Modell wird in jeder Woche mit den Scores und dem Aufwand aller zurückliegenden Wochen gefüttert. Die geschätzte Wahrscheinlichkeit wird dann zur Klassifizierung genutzt. Ist sie größer als 0,5, ist die Vorhersage, dass die Klausur bestanden wird (Cameron/Trivedi, 2005, S. 469ff.). Für die Schätzung der insgesamt 17 Modelle⁸ wurden vorab die Daten der 392 Studierenden in einen Test- und Trainingsdatensatz unterteilt, wobei der Testdatensatz 20 % (79 Studierende) und der Trainingsdatensatz 80 % der Daten (313 Studierende) umfasst. Studierende, die keine Prüfung wahrgenommen haben, wurden als nicht bestanden kodiert.

Diese Klassifizierungsmodelle werden an zwei Metriken evaluiert: Accuracy und Negative Predictive Value (NPV). Die Accuracy beschreibt den Anteil an richtigen Vorhersagen an allen Vorhersagen, also dem Quotienten aus richtig vorhergesagtem „Bestehen“ und „Nicht-Bestehen“, dividiert durch die Anzahl der Studierenden. Bei einer perfekten Vorhersage für alle Studierenden läge die Accuracy bei 1. Der NPV beschreibt den Anteil an richtigen Vorhersagen „Nicht-Bestehen“ unter allen „Nicht-Bestehens“-Vorhersagen. Die Metriken werden anhand des Testdatensatzes berechnet, um Overfitting zu vermeiden.

Die Verläufe von Accuracy und NPV zeigen ein ähnliches Muster über das Semester. Accuracy und NPV starten beide bei circa 0,76 und steigen bis Semesterwoche neun recht stetig an, mit der Ausnahme von Woche vier. Dies könnte daran liegen, dass in der Zeit viele Studierende erstmalig in JACK aktiv wurden. In Semesterwoche drei und vier haben insgesamt 122 Studierende ihre erste Einreichung in JACK vorgenommen. Ein ähnliches Muster lässt sich nach der zehnten Woche beobachten. Bei einem Vergleich mit Abbildung 2 fällt auf, dass die schwächere Accuracy und NPV mit dem Zeitpunkt korreliert, an dem sich die Notengruppe „ausreichend“ von der Gruppe „nicht ausreichend“ absetzt. Dies ist gerade der Unterschied zwischen Bestehen und Nicht-Bestehen und wirkt sich dadurch negativ auf die Vorhersagekraft der Modelle aus. In der Grafik lässt sich erkennen, dass es circa 3 Wochen dauert, bis diese Dynamik in den Modellen berücksichtigt wird.

Grundsätzlich ist die Vorhersagequalität des Modells schon relativ früh im Semester gut und könnte für ein potenzielles Frühwarnsystem genutzt werden. In Zukunft soll das Modell weiter verfeinert werden. Zum einem könnte der Zeitpunkt der ersten Bearbeitung mit in das Modell aufgenommen werden. Zum anderen könnte der Score mit dem Schwierigkeitslevel der Aufgaben gewichtet werden. Dies könnte vor allem die Vorhersagekraft in den letzten Semester-

7 Andere Datenaggregationen und Modelle wurden untersucht, lieferten allerdings deutlich schlechtere Ergebnisse.

8 Ab Woche zwei wird bis zur Woche der Klausur jeweils ein eigenes Modell mit den neuesten Daten geschätzt.

wochen weiter verbessern. Besonders die ersten Aufgaben sind Einführungs-
aufgaben, deren erfolgreiche Bearbeitung nicht hinreichend zum Bestehen der
Klausur ist. Allerdings wird es naturgemäß herausfordernd bleiben, die Leistung
von Studierenden, die relativ spät im Semester anfangen zu lernen, frühzeitig
präzise vorherzusagen.

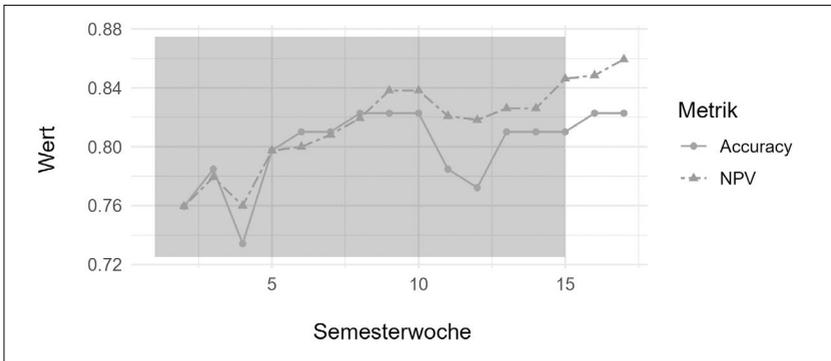


Abbildung 4: Verlauf der Modellperformance über das Semester bis zum ersten Prüfungstermin. Die eingefärbte Fläche spiegelt die Vorlesungszeit wider (Quelle: Eigene Abbildung)

5 Fazit

Grundsätzlich hat der Lehrstuhl äußerst positive Erfahrungen mit der Digitalisierung der Lehre gemacht und wird diesen Weg auch in Zukunft konsequent weitergehen. Dabei stehen neben der technischen Weiterentwicklung von JACK, um noch vielfältigere Aufgaben anbieten zu können, vor allem die Kommunikation mit den Studierenden und Open Educational Resources im Vordergrund. Unter anderem durch die Analyse von Lerndaten, wie in Abschnitt 4, sollen Studierende in Zukunft auch in Großveranstaltungen besser begleitet werden.

Literatur

- Cameron, A. C./Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo: Cambridge University Press.
- Klenke, J./Massing, T./Reckmann, N./Langerbein, J./Otto, B./Goedicke, M./Hanck, C. (2023). Effects of Early Warning Emails on Student Performance. In: *Proceedings of the 15th International Conference on Computer Supported Education – Volume 1: CSEDU*, 225–232.

- Massing, T./Schwinning, N./Striewe, M./Hanck, C./Goedicke, M. (2018). E-Assessment Using Variable-Content Exercises in Mathematical Statistics. In: *Journal of Statistics Education* 2018, 26(3), 174–189.
- Vierjahn, B. (2022). Wiederherstellung digitaler Systeme läuft auf Hochtouren. <https://www.uni-due.de/2022-12-07-wiederherstellung-digitaler-systeme-laeuft-auf-hochtouren> [Zugriff: 15.10.2023].

Entwicklung eines formativen Assessments zur Förderung prozeduralen und konditionalen biologiedidaktischen Wissens

Justin Timm, Nora Lösing, Christian Johannes und Philipp Schmiemann

Das Ziel der Lehramtsausbildung ist die Professionalisierung von Lehrer:innen vor dem Hintergrund multipler Herausforderungen. Die Erwartungen von Schüler:innen, Eltern, Politik und Gesellschaft an Lehrer:innen sind umfassend und gehen weit über die zentrale Aufgabe der Kompetenzentwicklung bei Lernenden hinaus (Bonnet/Hericks, 2014). Die Professionalisierung von Lehrer:innen ist ein kontinuierlicher Prozess, der sich über beide Phasen der Lehramtsausbildung erstreckt und mit Eintritt in den regulären Schuldienst nicht als abgeschlossen gelten kann (Terhart, 2001).

Im Rahmen des vorliegenden Beitrags wird die Entwicklung eines formativen Assessments zur Förderung prozeduraler und konditionaler fachdidaktischer Kompetenzen im Rahmen eines Biologiedidaktik-Moduls beschrieben. Ziel dieses Ansatzes ist die stärkere Ausrichtung auf die Anforderungen der Schulpraxis, insbesondere in Bezug auf Prozess- und Handlungsorientierung.

1 Theoretische Herleitung: fachliches, fachdidaktisches und curriculares Wissen

Die Struktur und Genese der Profession von Lehrkräften kann im Rahmen eines kompetenzorientierten Ansatzes durch verschiedene Modelle und Theorien beschrieben werden (Bonnet/Hericks, 2014). Die Typologie von Shulman (1986, 1987) beschreibt eine Reihe von Wissenskomponenten, die die Profession von Lehrer:innen bedingen. Die Wissenskomponenten, die in beiden Arbeiten aufgeführt werden, sind das fachliche Wissen (*content knowledge*, CK), das pädagogische Wissen (*pedagogical knowledge*, PK), das fachdidaktische Wissen (*pedagogical content knowledge*, PCK) sowie das curriculare Wissen (*curricular knowledge*). Im Modell von Blömeke/Gustafsson/Shavelson (2015) repräsentieren PCK, CK und PK zusammen mit affektiv-motivationalen Einstellungen dispositionale Merkmale, die Einfluss auf die inhalts- und situationspezifischen Fähigkeiten von Lehrkräften im Bereich der Wahrnehmung, Interpretation und Entscheidungsfindung haben, welche wiederum die Performanz in Form eines gezeigten Verhaltens beeinflussen. Die Bedeutung von fachdidaktischem Wissen und Fachwissen für eine gelingende Professionalisierung wird durch Stu-

dien untermauert, die indirekte Effekte von PCK und CK auf das Lernen von Schüler:innen nachweisen (Blömeke et al., 2022; Förtsch et al., 2016), wobei bezüglich des Zusammenhangs von fachdidaktischem Wissen und Lernerfolg insgesamt widersprüchliche Ergebnisse vorliegen (einen ersten Überblick liefern Chan/Hume, 2019). Im Rahmen dieses Beitrags orientieren wir uns dabei für fachdidaktisches Wissen an der Konzeptualisierung von Park/Oliver (2008):

PCK is teachers' understanding and enactment of how to help a group of students understand specific subject matter using multiple instructional strategies, representations, and assessments while working within the contextual, cultural, and social limitations in the learning environment (ebd.: 264, Hervorhebung im Orig.).

Für das Konstrukt PCK existieren vielfältige Konzeptualisierungen, wobei vielen gemein ist, dass das Konstrukt in Komponenten ausdifferenziert wird (Chan/Hume, 2019). Einen systematischen Überblick über mögliche Komponenten liefern Park/Oliver (2008) auf der Basis der Analyse von van Driel/Verloop/Vos (1998). In empirischen Arbeiten mit dem Fokus auf das Fach Biologie wurden bisher insbesondere, aber nicht ausschließlich, die zwei (Teil-)Komponenten Instruktionsstrategien und Schüler:innenvorstellungen untersucht (z. B. van Dijk/Kattmann, 2010; Großschedl/Welter/Harms, 2019; Jüttner et al., 2013; Schmelzing et al., 2013). Fachdidaktisches Wissen kann außerdem nach der Art des Wissens klassifiziert werden. Typisch ist dabei die Unterscheidung in deklaratives und prozedurales Wissen (Großschedl/Welter/Harms 2019). Tepner et al. (2012) grenzen darüber hinaus konditionales Wissen von den vorausgegangenen Kategorien ab (siehe auch Jüttner et al., 2013; Jüttner/Neuhaus, 2013a).

2 Kompetenzentwicklung im Modul Didaktik der Biologie II

Da deklaratives Wissen zu den zentralen Theorien der Biologiedidaktik bereits in einer vorangegangenen fachdidaktischen Veranstaltung etabliert wird, zielt das hier vorgestellte Modul Didaktik der Biologie II maßgeblich auf die Kompetenzentwicklung im Bereich des prozeduralen und konditionalen fachdidaktischen Wissens ab (Tepner et al., 2012; Jüttner et al., 2013). Das Modul umfasst Seminare zu „Methoden und Unterrichtskonzepte[n] in der Biologie“ und der „Planung und Analyse von Biologieunterricht“, die laut Studienplan im vierten bzw. fünften Semester belegt werden sollen (u. a. Universität Duisburg-Essen, 2019). In der Regel werden die Seminare semesterweise alternierend angeboten und von etwa 160 Studierenden pro Semester belegt.

Das Modul ist Bestandteil der Bachelor-Lehramtsstudiengänge aller Schulformen für das Fach Biologie. Zwar liegen bereits validierte und empirisch erprobte Instrumente zur Erfassung des biologiespezifischen fachdidaktischen Wissens

vor, diese Instrumente besitzen allerdings keine hinreichende Passung für den Einsatz als formative Assessments im Rahmen des hier betrachteten Moduls, da sie 1) als Testinstrumente entwickelt wurden, sie 2) nicht den gewünschten Umfang aufweisen und 3) nicht zu allen im vorliegenden Modul behandelten Inhalten komplexe Handlungssituationen liefern, die den Einsatz prozeduralen und konditionalen fachdidaktischen Wissens anregen.

3 Entwicklung formativer Aufgaben zur Förderung des prozeduralen und konditionalen fachdidaktischen Wissens

Im Rahmen dieses Beitrags werden formative Aufgaben zur Förderung des prozeduralen und konditionalen fachdidaktischen Wissens entwickelt (vgl. Tepner et al., 2012). Dazu wurden zunächst im Rahmen eines Abgleichs der Studienverlaufspläne mit Bildungsstandards und Kernlehrplänen (MSW NRW, 2013; KMK, 2005) Themenfelder identifiziert. Das Ziel war die Identifikation von Themenfeldern mit hoher schulischer Relevanz, für die vor dem Hintergrund vorangegangener universitärer Lehre bereits ein elaboriertes Domänenwissen existieren sollte. Ausgewählt wurden die Themenfelder Genetik, Botanik, Zoologie und Zellbiologie.

Fachdidaktische Konzepte und Theorien wurden auf der Basis der fachlichen Seminarinhalte ausgewählt, um eine hohe interne Konsistenz zu gewährleisten. Die Komponente des Wissens über fach- und inhaltspezifische Lehr- und Lernstrategien nach Park/Oliver (2008) wird beispielsweise durch Inhalte zu den Methoden der naturwissenschaftlichen Erkenntnisgewinnung (Wellnitz/Mayer, 2013), zum *Systems Thinking* im Bereich der Zellbiologie und Botanik (Mambrey et al., 2020; Verhoeff/Waarlo/Boersma, 2008) und das *Yo-Yo-Learning* in den Bereichen der Botanik und Genetik (Knippels, 2002) abgedeckt. Das Wissen über das Verständnis von Schüler:innen (Park/Oliver, 2008) wird unter anderem im Kontext von Schüler:innenvorstellungen, Kontexteffekten, Lerner:inneninteressen, mangelnder Vernetzung und Fachsprache formativ erfasst.

Die für ausgewählte Sitzungen in einem offenen Format konstruierten Aufgaben (Hammann/Jördens, 2014) werden den Studierenden online über die Lernplattform *Moodle* zur Verfügung gestellt. Die Studierenden erhalten ein elaboriertes Feedback (vgl. Narciss, 2008; Shute, 2008), welches Anlass zur Evaluation der eigenen Antwort geben kann und dafür, je nach Aufgabe, fachliche, pädagogische und/oder fachdidaktische Inhalte adressiert. Auf der Basis bisher zur Verfügung gestellter Übungsaufgaben ist davon auszugehen, dass die Studierenden die Aufgaben insbesondere zur Vorbereitung auf die Modulabschlussprüfung bearbeiten, selbst wenn sie im Rahmen des Seminars dazu angehalten werden, die Aufgaben zur Nachbereitung der jeweiligen Sitzungen zu bearbeiten.

3.1 Aufgabenbeispiel zum Einsatz biologiespezifischer Lehr-Lern-Methoden

Mit Abschluss der Sekundarstufe I sollten Lerner:innen ein grundlegendes Verständnis der Fotosynthese als Prozess der Stoff- und Energieumwandlung besitzen (KMK, 2005; MSW NRW, 2013). Die Schüler:innen sollten in der Lage sein, zur Erklärung biologischer Phänomene zwischen den Systemebenen zu wechseln (KMK, 2005). Dieser Systemebenenwechsel wird allerdings innerhalb von Lehr-Lernprozessen häufig nicht transparent gemacht, sondern implizit vollzogen. Dadurch können sie den Lernprozess behindern oder erschweren. Eine Möglichkeit, diese Systemebenenwechsel bewusst zu vollziehen, wurde von Knipfels (2002) beschrieben. Ein Aufgabenbeispiel verdeutlicht die Bedeutung einer sinnvollen Unterrichtsreihe beim Wechsel von Systemebenen zur Sicherung vertikaler Kohärenz. Darüber hinaus bietet das Aufgabenbeispiel Anlass, um die Bedeutung einer Kontextualisierung und Problemorientierung zu reflektieren.

Umplanung einer Unterrichtsreihe zur Fotosynthese

Aufgabenstellung: Erläutern Sie mögliche fachdidaktische Gründe für die Umplanung der Unterrichtsreihe.

Im letzten Schuljahr hat die Lehrkraft in der ersten Doppelstunde der Unterrichtseinheit zur Fotosynthese mithilfe eines Arbeitsblatts die Wortgleichung zur Fotosynthese aufgestellt. In der zweiten Doppelstunde haben die Schüler:innen einen Blattquerschnitt mikroskopiert. In der dritten Unterrichtsstunde wurde der Wasser- und Stofftransport innerhalb der Pflanze untersucht.

In diesem Jahr hat die Lehrkraft die Unterrichtsreihe umgestellt und geht in den ersten Stunden der Frage nach, wie die Stärke in die Kartoffel kommt. Dazu bestimmen die Schüler:innen zunächst experimentell Edukte und Produkte der Fotosynthese und identifizieren die Blätter der Kartoffel als Ort der Fotosynthese. Anschließend werden Wasser und Stofftransport untersucht, bevor der Blattquerschnitt mikroskopiert wird.

Abbildung 1: Aufgabe zur Erfassung des konditionalen Wissens zu Lehr-Lernstrategien im Themengebiet Botanik, genauer in der Pflanzenphysiologie. Struktur in Anlehnung an Budde et al. (2015). Aufgabe(nstellung) in Anlehnung an Jüttner/Neuhaus (2013b) und Großschedl/Welter/Harms (2019).

Im Aufgabenbeispiel 1 (Abb. 1) werden zwei Unterrichtseinheiten skizziert, in Verbindung mit der Aufgabe, mögliche Gründe für die Umplanung der Unterrichtsreihe zu benennen und zu erläutern (vgl. Jüttner/Neuhaus, 2013b). Durch den Vergleich der beiden Unterrichtseinheiten sollten den Lehramtsstudierenden mehrere Unterschiede auffallen, die fachdidaktisch begründbar erscheinen. Zunächst ist die Unterrichtsreihe in ihrer ursprünglichen Form nicht kontextualisiert (Gilbert, 2006; Bennett/Lubben/Hogarth, 2007). Die Kontextualisierung der

überarbeiteten Unterrichtsreihe erfüllt mehrere Kontextmerkmale nach Schmiemann et al. (2011): Sie ermöglicht ein situiertes Lernen und schafft Alltagsbezüge. Die Fragestellung kann unter Zuhilfenahme von Methoden der naturwissenschaftlichen Erkenntnisgewinnung von den Schüler:innen selbstständig beantwortet werden. Sie ist anschlussfähig an den Themenbereich gesunde Ernährung und besitzt damit eine hohe individuelle und gesellschaftliche Relevanz.

Durch die Modifikation der Reihung der Unterrichtsinhalte wird darüber hinaus vertikale Kohärenz durch den planvollen Wechsel der Systemebenen von der Phänomen- zur Organebene und von dort zur Ebene einzelner Gewebe erzeugt. Die Systemebenenwechsel erfolgen im Rahmen eines schrittweisen Hineinzoomens (Knippels, 2002). Leitend bleibt während des gesamten Prozesses die einführende Fragestellung. Der Unterricht folgt damit einer problemorientierten Struktur (vgl. Hmelo-Silver, 2004).

3.2 Aufgabenbeispiel zum Umgang mit Schüler:innenvorstellungen in der Genetik

Die Kenntnis von und ein produktiver Umgang mit Schüler:innenvorstellungen können dabei helfen, Lernhindernisse im Unterricht zu bewältigen (Hammann/Asshoff, 2023). Schüler:innenvorstellungen können allerdings auch (unbeabsichtigt) durch den Unterricht hervorgerufen werden (ebd.). Ein typischer Kontext zur Erarbeitung und Reflexion von Schüler:innenvorstellungen ist die Genetik. Das Themengebiet der Genetik wird von Lehrer:innen als abstrakt und komplex wahrgenommen (Knippels, 2002). Es liegt eine umfangreiche Beschreibung von Schüler:innenvorstellungen in diesem Bereich vor; einen Einstieg in dieses Thema, inklusive einer unterrichtspraktischen Übersicht, liefern Hammann/Asshoff (2023). Ein Aufgabenbeispiel zeigt die Bedeutung von Schüler:innenvorstellungen in der Genetik.

In der Unterrichtssituation in Aufgabenbeispiel 2 (Abb. 2) wird ein typisches Unterrichtsdilemma geschildert. Eine Schülerin soll einen humangenetischen Stammbaum analysieren. Die Schülerin identifiziert den richtigen Erbgang, basierend allerdings auf inkorrekten Annahmen in Form von Schüler:innenvorstellungen (vgl. Timm/Wools/Schmiemann, 2022). Darüber hinaus geht die Schülerin nicht hypothetisch-deduktiv vor. Diese Situation wird den Studierenden nun vorgelegt, und sie werden aufgefordert, zu beschreiben, wie sie auf die Antwort der Schülerin reagieren und den Unterricht fortsetzen würden (vgl. Jüttner/Neuhaus, 2013b). Verschiedene Ansätze erscheinen in diesem Zusammenhang zielführend. Stellvertretend seien einige Optionen geschildert, denen gemein ist, dass sie einen kognitiven Konflikt hervorrufen und dadurch Lernprozesse anstoßen sollen (Duit/Treagust, 2003; Limón, 2001): 1) Die Schüler:innen könnten aufgefordert werden, den von der Schülerin vorgestellten Lösungsweg auf andere Stammbäume anzuwenden. Dabei sollten unter diesen Stammbäumen auch solche sein, die nicht zu den Schüler:innenvorstellungen passen (vgl.

Timm/Oberste/Schmiemann, 2023). 2) Die Schüler:innen könnten im Rahmen eines Wettbewerbs herausgefordert werden, einen Stammbaum zu einem autosomal rezessiven Erbgang zu konstruieren, in dem so viele Personen wie möglich oder überwiegend Personen eines Geschlechts betroffen sind. Anhand der resultierenden Stammbäume könnte dann das Potenzial des Lösungswegs der Schüler:innen evaluiert werden. 3) Darüber hinaus könnten die Lösungswege weiterer Schüler:innen gesammelt werden, mit dem Ziel, die Lösungswege gegenüberzustellen und Potenziale und Grenzen ebendieser kritisch zu reflektieren.

Umgang mit Schüler:innenvorstellungen zur Genetik

Aufgabenstellung: Erläutern Sie, wie Sie in der beschriebenen Situation reagieren und den Unterricht fortsetzen würden.

Die Schüler:innen der zehnten Klasse der städtischen Gesamtschule haben in Einzelarbeit einen humangenetischen Stammbaum analysiert. Die Schüler:innen sollten den vorliegenden Erbgang identifizieren. Es handelt sich um einen autosomal rezessiven Erbgang. Die Lehrkraft bittet Lisa, das Ergebnis ihrer Analyse und ihren Lösungsweg vorzustellen.

Lisa: „Ich glaub das wird autosomal rezessiv vererbt.“

Lehrkraft: „Das ist richtig. Kannst du uns erklären, wie du vorgegangen bist?“

Lisa: „Ich habe mir den Stammbaum einfach angeschaut. Dabei ist mir aufgefallen, dass nur ein paar Leute betroffen sind. Muss dann ja rezessiv sein. Und weil ja Männer und Frauen betroffen sind, ist's autosomal.“

Abbildung 2: Aufgabe zur Erfassung des prozeduralen Wissens zum Umgang mit Schüler:innenvorstellungen im Themengebiet Genetik. Schülervorstellungen nach Timm/Wools/Schmiemann (2022). Aufgabenstellung in Anlehnung an Jüttner/Neuhaus (2013b)

4 Diskussion

Der vorliegende Beitrag beschreibt exemplarisch die Entwicklung formativer Aufgaben zur Erfassung und Förderung prozeduraler und konditionaler fachdidaktischer Handlungskompetenz. Die Aufgaben können damit der Förderung der Kompetenz angehender Biologielehrkräfte dienen und einen wichtigen Beitrag zur Professionalisierung von Lehramtsstudierenden leisten. Empirische Befunde belegen die Bedeutung des fachdidaktischen Wissens für die Wirksamkeit von Lehrkräften (Blömeke et al., 2022; Förtsch et al., 2016). Die Bearbeitung der entwickelten und in den Beispielen eins und zwei dargestellten Aufgaben erfordert neben fachdidaktischem Wissen auch umfangreiches fachbiologisches konzeptuelles Wissen. Dadurch verlieren die Aufgaben allerdings nicht an ökologischer

Validität, denn „Fachwissen ist die Grundlage, auf der fachdidaktische Beweglichkeit entstehen kann“ (Baumert/Kunter, 2013, S. 308, Hervorhebung des Originals aufgehoben). Die Aufgaben sollen die Studierenden zur Anwendung ihres fachdidaktischen Wissens in komplexen, unterrichtspraktischen Situationen animieren. Es ist davon auszugehen, dass die Schwierigkeit der Aufgaben maßgeblich von der Passung zu den Seminarinhalten abhängig ist. Eine hierarchische Organisation der Aufgaben in Abhängigkeit von der Wissensart – prozeduralen oder konditionalen – ist nicht zu erwarten (Jüttner et al., 2013).

Im Rahmen des formativen Einsatzes erhalten die Studierenden Rückmeldungen, die die Einschätzung und Reflexion ihrer fachdidaktischen Handlungskompetenz ermöglicht. Hierzu erhalten die Studierenden ein elaboriertes Feedback (vgl. Narciss, 2008). Dieses Feedback adressiert explizit einzelne Wissensdimensionen der Aufgabe. Je nach Aufgabe erhalten die Studierenden Hinweise zum fachlichen Inhalt und/oder es werden die relevanten fachdidaktischen Konzepte benannt, die relevant für die Beantwortung der Aufgabe sind. Soweit möglich, erhalten die Studierenden Impulsfragen, die sie zur Reflexion der in den Aufgaben und ihren Antworten demonstrierten Vorstellungen von Lernen hinsichtlich transmissiver und konstruktivistischer Überzeugungen anregen sollen. Dabei streben Modul und Feedback die Förderung einer konstruktivistischen Überzeugung von Lernen an, da eine konstruktivistische Überzeugung möglicherweise in einem positiven Zusammenhang mit Instruktionsqualität und Lernerfolg steht (Voss et al., 2013; Staub/Stern, 2002; kein sig. Zusammenhang bei Seidel et al., 2009).

Seit dem Wintersemester 2023/2024 wird an der Universität Duisburg-Essen auch das Studienfach Biologie im Bachelorstudiengang mit der Lehramtsoption sonderpädagogische Förderung angeboten. Obwohl das Seminar bereits inklusionsorientierte Inhalte enthält, erscheint es sinnvoll, das Angebot um weitere formative Aufgaben mit dezidiert inklusionsbezogenem, fachdidaktischem Fokus zu erweitern. Auch eine Erweiterung hinsichtlich des technologischen Wissens erscheint vor dem Hintergrund der Herausforderungen des 21. Jahrhunderts und der Relevanz des TPACK-Modells vielversprechend (Mishra/Koehler, 2006; Groß et al., 2022).

Eine empirische Überprüfung der Wirksamkeit der formativen Aufgaben ist bisher nicht erfolgt. Diese kann im Rahmen einer Lehrevaluation mithilfe validierter und empirisch erprobter Instrumente erfolgen (z. B. Großschedl/Welter/Harms, 2019). Dies setzt allerdings eine valide Bewertung der Antworten von Studierenden sowie eine adäquate Passung von Aufgabenschwierigkeit und fachdidaktischem Wissen der Studierenden voraus. Darüber hinaus wäre eine Analyse der Wirksamkeit einer regelmäßigen Aufgabenbearbeitung in Bezug auf die Modulabschlussnote wünschenswert.

Literatur

- Baumert, J./Kunter, M. (2013). Professionelle Kompetenz von Lehrkräften. In I. Gogolin/H. Kuper/H.-H. Krüger/J. Baumert (Hrsg.), *Stichwort: Zeitschrift für Erziehungswissenschaft* (S. 277–337). Wiesbaden: Springer Fachmedien Wiesbaden.
- Bennett, J./Lubben, F./Hogarth, S. (2007). Bringing science to life: A synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. In: *Science Education*, 91(3), 347–370.
- Blömeke, S./Gustafsson, J.-E./Shavelson, R. J. (2015). Beyond Dichotomies. In: *Zeitschrift für Psychologie*, 223(1), 3–13.
- Blömeke, S./Jentsch, A./Ross, N./Kaiser, G./König, J. (2022). Opening up the black box. Teacher competence, instructional quality, and students' learning progress. In: *Learning and Instruction*, 79, 101600.
- Bonnet, A./Hericks, U. (2014). Professionalisierung und Deprofessionalisierung im Lehrer/innenberuf. Ansätze und Befunde aktueller empirischer Forschung. In: *Zeitschrift für interpretative Schul- und Unterrichtsforschung*, 3, 3–13.
- Budde, J./Janßen, H.-J./Jeuck, J./Schaal, H./Schönke, M./Schulz, J./Schwerdtfeger, Y./Wolf, A. (2015). *Fokus Biologie 7/8. Niedersachsen G9*. Berlin: Cornelsen.
- Chan, K. K. H./Hume, A. (2019). Towards a Consensus Model: Literature Review of How Science Teachers' Pedagogical Content Knowledge Is Investigated in Empirical Studies. In A. Hume/Cooper, R./Borowski, A. (Hrsg.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (S. 3–76). Singapore: Springer.
- Duit, R./Treagust, D. F. (2003). Conceptual change. A powerful framework for improving science teaching and learning. In: *International Journal of Science Education*, 25(6), 671–688.
- Förtsch, C./Werner, S./von Kotzebue, L./Neuhaus, B. J. (2016). Effects of biology teachers' professional knowledge and cognitive activation on students' achievement. In: *International Journal of Science Education*, 38(17), 2642–2666.
- Gilbert, J. K. (2006). On the Nature of "Context" in Chemical Education. In: *International Journal of Science Education*, 28(9), 957–976.
- Groß, J./Lude, A./Nerdel, C./Paul, J./Schaal, S./Schmiemann, P./Thyssen, C. (2022). Biologische Bildung in der digitalen Welt. In V. Frederking/R. Romeike (Hrsg.), *Fachliche Bildung in der digitalen Welt. Digitalisierung, Big Data und KI im Forschungsfokus von 15 Fachdidaktiken. Fachdidaktische Forschungen, Band 14*, S. 47–81. Münster: Waxmann.
- Grötschedl, J./Welter, V./Harms, U. (2019). A new instrument for measuring pre-service biology teachers' pedagogical content knowledge: The PCK-IBI. In: *Journal of Research in Science Teaching*, 56(4), 402–439.
- Hammann, M./Asshoff, R. (2023). *Schülervorstellungen im Biologieunterricht. Ursachen für Lernschwierigkeiten* (5. Auflage). Seelze: Klett/Kallmeyer.
- Hammann, M./Jördens, J. (2014). Offene Aufgaben codieren. In D. Krüger/I. Parchmann/H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 169–178). Berlin, Heidelberg: Springer.
- Hmelo-Silver, C. E. (2004). Problem-Based Learning: What and How Do Students Learn? In: *Educational Psychology Review*, 16(3), 235–266.
- Jüttner, M./Boone, W./Park, S./Neuhaus, B. J. (2013). Development and use of a test instrument to measure biology teachers' content knowledge (CK) and pedagogical content

- knowledge (PCK). In: *Educational Assessment, Evaluation and Accountability*, 25(1), 45–67.
- Jüttner, M./Neuhaus, B. J. (2013a). Das Professionswissen von Biologielehrkräften. Ein Vergleich zwischen Biologielehrkräften, Biologen und Pädagogen. In: *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 31–49.
- Jüttner, M./Neuhaus, B. J. (2013b). Validation of a Paper-and-Pencil Test Instrument Measuring Biology Teachers' Pedagogical Content Knowledge by Using Think-Aloud Interviews. In: *Journal of Education and Training Studies*, 1, 2.
- Knippels, M.-C. P. J. (2002). Coping with the abstract and complex nature of genetics in biology education. The yo-yo learning and teaching strategy. Dissertation. CD- β Series on Research in Science Education, Band 43. Utrecht: CD- β .
- Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: a critical appraisal. In: *Learning and Instruction*, 11 (4–5), 357–380.
- Mambrey, S./Timm, J./Landskron, J. J./Schmiemann, P. (2020). The impact of system specifics on systems thinking. In: *Journal of Research in Science Teaching*, 57(10), 1632–1651.
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (MSW NRW) (Hrsg.) (2013). Kernlehrplan für die Gesamtschule – Sekundarstufe I in Nordrhein-Westfalen [Core curriculum for the schools of the type Gesamtschule – lower secondary level in North Rhine-Westphalia]. *Naturwissenschaften, Biologie, Chemie, Physik [Sciences, biology, chemistry, physics]*.
- Mishra, P./Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. In: *Teachers College Record: The Voice of Scholarship in Education*, 108(6), 1017–1054.
- Narciss, S. (2008). Feedback Strategies for Interactive Learning Tasks. In D. Jonassen/M. J. Spector/M. Driscoll/M. D. Merrill/ J. J. G. van Merriënboer/M. P. Driscoll (Hrsg.), *Handbook of Research on Educational Communications and Technology* (S. 125–144). New York, NY: Routledge.
- Park, S./Oliver, J. S. (2008). Revisiting the Conceptualisation of Pedagogical Content Knowledge (PCK). PCK as a Conceptual Tool to Understand Teachers as Professionals. In: *Research in Science Education*, 38(3), 261–284.
- Schmelzing, S./van Driel, J. H./Jüttner, M./Brandenbusch, S./Sandmann, A./Neuhaus, B. J. (2013). Development, Evaluation, and Validation of a Paper-and-Pencil Test for Measuring Two Components of Biology Teachers' Pedagogical Content Knowledge Concerning the "Cardiovascular System". In: *International Journal of Science and Mathematics Education*, 11(6), 1369–1390.
- Schmiemann, P./Linsner, M./Wenning, S./Sandmann, A. (2011). Kontextorientiertes Lernen in Biologie. Aufgaben und Arbeitsmaterialien. In P. Schmiemann/A. Sandmann (Hrsg.), *Aufgaben im Kontext. Biologie* (S. 4–12). Seelze: Friedrich.
- Seidel, T./Schwindt, K./Rimmele, R./Prenzel, M. (2009). Konstruktivistische Überzeugungen von Lehrpersonen: Was bedeuten sie für den Unterricht? In M. A. Meyer/M. Prenzel/S. Hellekamps (Hrsg.), *Perspektiven der Didaktik* (S. 259–276). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) (Hrsg.) (2005). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*.
- Shulman, L. S. (1986). Those Who Understand. Knowledge Growth in Teaching. In: *Educational Researcher*, 15(2), 4–14.

- Shulman, L. S. (1987). Knowledge and Teaching. Foundations of the New Reform. In: *Harvard Educational Review*, 57(1), 1.23.
- Shute, V. J. (2008). Focus on Formative Feedback. In *Review of Educational Research*, 78(1), 153–189.
- Staub, F. C./Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. In: *Journal of Educational Psychology*, 94(2), 344–355.
- Tepner, O./Borowski, A./Dollny, S./Fischer, H. E./Jüttner, M./Kirschner, S./Leutner, D./Neuhaus, B. J./Sandmann, A./Sumfleth, E./Thillmann, H./Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. In: *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7–28.
- Terhart, E. (2001). *Lehrerberuf und Lehrerbildung. Forschungsbefunde, Problemanalysen, Reformkonzepte*. Weinheim, Basel: Beltz.
- Timm, J./Oberste, N./Schmiemann, P. (2023). Which factors influence the success in pedigree analysis? In: *International Journal of Science Education* 45(3), 1–19.
- Timm, J./Wools, K./Schmiemann, P. (2022). Secondary Students' Reasoning on Pedigree Problems. In: *CBE – Life Sciences Education*, 21(1), ar14.
- Universität Duisburg-Essen (2019). *Fachprüfungsordnung für das Studienfach Biologie im Bachelorstudiengang mit der Lehramtsoption Gymnasien und Gesamtschulen an der Universität Duisburg-Essen*.
- van Dijk, E. M./Kattmann, U. (2010). Evolution im Unterricht. Eine Studie über fachdidaktisches Wissen von Lehrerinnen und Lehrern. In: *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 7–21.
- van Driel, J. H./Verloop, N./de Vos, W. (1998). Developing science teachers' pedagogical content knowledge. In: *Journal of Research in Science Teaching* 35(6), 673–695.
- Verhoeff, R. P./Waarlo, A. J./Boersma, K. T. (2008). Systems Modelling and the Development of Coherent Understanding of Cell Biology. In: *International Journal of Science Education*, 30(4), 543–568.
- Voss, T./Kleickmann, T./Kunter, M./Hachfeld, A. (2013). Mathematics Teachers' Beliefs. In M. Kunter/J. Baumert/W. Blum/U. Klusmann/S. Krauss/M. Neubrand (Hrsg.), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers* (S. 249–271). Boston, MA: Springer.
- Wellnitz, N./Mayer, J. (2013). Erkenntnismethoden in der Biologie. Entwicklung und Evaluation eines Kompetenzmodells. In: *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 315–345.

Hilfe, ich komme nicht weiter! Digitale Lösungs- und Visualisierungshilfen für Übungs- und Klausuraufgaben der Experimentalphysik

Heike Theyßen, Lisa Danzig, Sebastian Haase und Hermann Nienhaus

Zur individuellen Förderung der Studierenden in der Studieneingangsphase der Physik wird das Konzept der gestuften Lernhilfen aus dem Physikunterricht auf die klassischen Übungsaufgaben zu Vorlesungen übertragen. Nach Etablierung in den Übungen können diese Hilfen gegen Verrechnung mit Klausurpunkten auch in Klausuren bereitgestellt werden.

1 Übungsaufgaben im Physikstudium: Lerngelegenheiten mit Herausforderungen

Im Physikstudium sind insbesondere in der Studieneingangsphase hohe Abbruchquoten zu verzeichnen (Heublein et al., 2022). Zahlreiche Befunde deuten darauf hin, dass neben der Abiturdurchschnittsnote insbesondere das physikbezogene Vorwissen und die Rechenfähigkeit prädiktiv für den Studienerfolg sind (z. B. Sorge et al. 2016; Buschhüter et al., 2017; Woitkowski, 2019). Die Lerngruppen sind allerdings durchaus heterogen in ihren Voraussetzungen und diese Heterogenität kann durch die etablierten Lehr-Lern-Formate nicht wirksam verringert werden. So lernen zwar alle Studierenden während des ersten Semesters signifikant hinzu, aber das Drittel mit der geringsten Abiturdurchschnittsnote erreicht bei dem Konzeptverständnis und der Fähigkeit zur Wissensanwendung am Ende des ersten Semesters im Mittel bei Weitem nicht das Einstiegsniveau des Drittels mit den höchsten Abiturdurchschnittsnoten (Binder et al., 2019). Eine bestmögliche Unterstützung der Studierenden muss daher diese Heterogenität anerkennen und die Studierenden individuell fördern (vgl. Hußmann et al., 2012). Dabei erreichen optionale begleitende Unterstützungsangebote, wie Tutorien und Klausurtrainings, oft gerade nicht diejenigen Studierenden mit den größten Unterstützungsbedarfen (Haak, 2017, S. 297). Dies spricht dafür, neben Zusatzangeboten, auch Unterstützungsangebote in den regulären Lehrbetrieb zu integrieren. Dieser ist in der Studieneingangsphase neben einigen Experimentalpraktika in der Regel durch die Lehr-Lern-Formate Vorlesung und Übung charakterisiert. Begleitend zu den Vorlesungen bearbeiten die Studierenden in Einzel oder Gruppenarbeit eigenständig Übungsaufgaben, die abgegeben, korrigiert und anschließend in Übungsgruppen besprochen werden. Die Mehrzahl der Übungsaufgaben sind

fachspezifische Problemlöseaufgaben (Friege, 2001), die mit Hilfe physikalischer Konzepte und mathematischer Methoden zu lösen sind. Abbildung 1 zeigt ein Beispiel für eine solche Aufgabenstellung.

Abgelenkte Elektronen

Mit dem Aufbau in Abbildung 1 wird ein Elektron aus der Ruhe mit einer Spannung U_a auf die Geschwindigkeit v in x -Richtung beschleunigt. Mit dieser Geschwindigkeit fliegt es mittig zwischen den Platten in einen Plattenkondensator. Dieser Plattenkondensator hat eine Länge von $L = 10$ cm und einen Plattenabstand von $d = 2$ cm.

- a. Berechnen Sie den Betrag der Geschwindigkeit v , mit der das Elektron in den Plattenkondensator eintritt, wenn es mit einer Spannung von $U_a = 4000$ V beschleunigt wird. Rechnen Sie dazu nicht relativistisch.

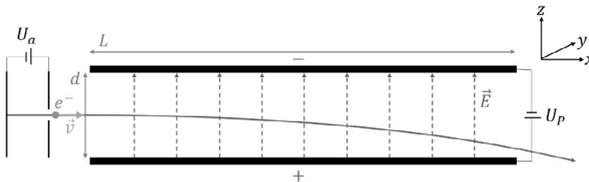


Abbildung 1 Aufbau zur Beschleunigung eines Elektrons und Ablenkung zur positiv geladenen Platte

- b. An die Platten des Plattenkondensators wird eine Spannung U_P angelegt und damit ein homogenes elektrisches Feld \vec{E} zwischen den Platten aufgebaut, wie in Abbildung 1 eingezeichnet (von der positiven zur negativen Platte). Leiten Sie die maximale Spannung U_P am Plattenkondensator her, bei der das Elektron den Plattenkondensator noch ungehindert durchfliegen kann, d.h. ohne auf eine Platte zu treffen.

Abbildung 1: Ausschnitt aus einer typischen Aufgabenstellung für Übungen in der Experimentalphysik (Screenshot aus tet.folio, Quelle: Eigene Abbildung)

Die Übungsaufgaben stellen einen wesentlichen Anteil der eigenständigen Auseinandersetzung mit den Vorlesungsinhalten dar und bereiten auf die Klausur vor, in der ähnliche Aufgaben unter Prüfungsbedingungen zu bearbeiten sind. Allerdings fällt den Studierenden die Bearbeitung dieser Übungsaufgaben oft schwer (Haak, 2017, S. 285), und die Schwierigkeiten liegen vielfach bereits darin, einen Ansatz zur Bearbeitung der Aufgabe zu finden (Pusch 2014; Woitkowski, 2021). Ohne einen geeigneten Ansatz können jedoch die weiteren Lern- und Übungsgelegenheiten, welche die Aufgabenbearbeitung bietet, nicht im intendierten Maße genutzt werden. Gleiches gilt für Vereinfachungen, z. B. durch Symmetrieüberlegungen, ohne die eine Aufgabe ggf. eine sehr viel aufwändigere Bearbeitung erfordert.

Im zweiten Fachsemester wird üblicherweise die Elektrodynamik behandelt. Hier kommt zu den oben beschriebenen themenübergreifenden Schwierigkeiten hinzu, dass zum Verständnis und zur Bearbeitung der Aufgaben oft dreidimensionale Vorstellungen von Feldern oder Ladungsverteilungen sowie deren

raum-zeitlichen Änderungen nötig sind. Diese Vorstellungen können durch statische, zweidimensionale Projektionen auf Übungsblättern nur unzureichend unterstützt werden.

Zusammenfassend stellen die Übungsaufgaben eine wichtige Lerngelegenheit innerhalb des regulären Lehrangebotes und gleichzeitig eine curricular valide Vorbereitung auf die Prüfungsanforderungen dar. Aufgrund der vielfältigen Herausforderungen können sie jedoch nicht von allen Studierenden zielführend bearbeitet werden und scheinen insbesondere nicht geeignet, die Studierenden mit schlechteren Eingangsvoraussetzungen optimal zu fördern. Auch wenn es in den Klausuraufgaben nicht um Förderung, sondern um Kompetenzmessung geht, stellt sich hier ein ähnliches Problem, wenn Studierende bei komplexeren Aufgaben den Ansatz nicht finden und ihre Fähigkeiten im weiteren Verlauf des Problemlöseprozesses nicht adäquat zeigen können. Unterstützungsmaßnahmen, die an den Übungsaufgaben und perspektivisch den Klausuraufgaben ansetzen, könnten somit nicht nur die Mehrheit der Studierenden erreichen, sondern sind auch ohne tiefgreifende Eingriffe in die etablierten Lehr-Lern-Formate umsetzbar.

2 Zielgruppe, Zielsetzung und Ansatz

Vor dem geschilderten Hintergrund ist es Ziel des Projekts PITCH in der Physik, die Bearbeitung der Übungsaufgaben individuell und differenziert zu unterstützen, damit die Studierenden trotz heterogener Lernvoraussetzungen jeweils bestmöglich von der Bearbeitung profitieren. Dies soll nicht im Rahmen optionaler Zusatzangebote geschehen, sondern innerhalb des regulären Übungsbetriebs, um möglichst viele Studierende zu erreichen.

Zielgruppe sind sowohl die „Fachstudierenden“ (B.Sc. Physik und B.Sc. Energy Science), die am Campus Duisburg die Vorlesung „*Grundlagen der Physik 2*“ besuchen, als auch die Studierenden für das Lehramt Physik an Gymnasien und Gesamtschulen sowie an Berufskollegs, die am Campus Essen eine eigens für das Lehramt konzipierte Vorlesung „*Experimentalphysik 2*“ besuchen. In beiden Gruppen sind etwa 25 bis 40 Studierende pro Jahrgang. Für die individuelle Förderung der Studierenden dieser Zielgruppen werden zwei Ansätze verfolgt.

Lösungshilfen: Den Studierenden werden zu den zentralen Lösungsschritten der Übungsaufgaben Hilfen angeboten, die sie bei Bedarf nutzen können. Das Konzept orientiert sich an den (gestuften) Lernhilfen für den Physikunterricht (Francke-Braun et al., 2008). Die Hilfen sind entlang des fachspezifischen Problemlöseprozesses (Friege, 2001) strukturiert und enthalten z. B. Hinweise zu sinnvollen Vereinfachungen und Symmetrieüberlegungen oder geben einen geeigneten Lösungsansatz an. Sie können während der Aufgabenbearbeitung bei Bedarf genutzt werden. In den Lern- und Übungsphasen während des Semesters stehen die Hilfen zur Nutzung frei zur Verfügung. In der Klausur kann die Nutzung von Hilfen mit den erreichbaren Punkten verrechnet werden. Dafür müssen

die Studierenden vor dem Aufruf einer Hilfe eine möglichst genaue Vorstellung davon haben, inwieweit die jeweilige Hilfe ihnen für den eigenen Lösungsprozess nützlich sein wird, um diesen Nutzen gegen die dafür aufgegebenen Punkte abzuwägen. Dies erfordert zum einen eine aussagekräftige Beschriftung der Hilfen hinsichtlich des zu erwartenden Inhalts sowie der dafür einzulösenden Punkte, zum anderen müssen die Studierenden zuvor mit dem Konzept der Lösungshilfen im Übungsbetrieb vertraut gemacht werden.

Visualisierungshilfen: Als eine spezifische Form der Lösungshilfen werden den Studierenden interaktive Visualisierungen der oft dynamischen, dreidimensionalen Situationen in den Aufgabenstellungen angeboten. Damit können sie die Situation aus verschiedenen, selbst gewählten Perspektiven betrachten sowie zeitliche Abläufe oder die Auswirkungen von selbst vorgenommenen Variablenänderungen beobachten. Damit die Visualisierungshilfen auch zur Prüfung der eigenen Lösung genutzt werden können, liegen ihnen quantitative Simulationen der physikalischen Zusammenhänge zugrunde.

Der Einsatz der Visualisierungshilfen erfordert ein digitales Format für die Bereitstellung. Dies gilt für den Übungsbetrieb ebenso wie für die Klausur. Auch für das Nachhalten und „Abrechnen“ der Hilfenutzung während einer Klausur bietet sich ein digitales Format an. Dabei soll die Klausur selbst nicht auf ein digitales Format umgestellt werden, sondern im klassischen, für die Physik typischen „Papier-und-Stift“-Format bleiben. Hauptgrund hierfür ist, dass die typischen Übungs- und Klausuraufgaben mit ihren vielfältigen fachspezifischen Anforderungen an mathematische Umformungen, Skizzen und Diagramme zurzeit noch nicht adäquat auf den zur Verfügung stehenden Plattformen abgebildet werden können. Sie können jedoch sehr wohl durch digitale Lösungs- und Visualisierungshilfen ergänzt werden.

Im Folgenden wird die Wahl der Plattform für die Bereitstellung der Hilfen und die technische Umsetzung erläutert (Abschnitt 1.3), bevor konkrete Einblicke in die Abfolge und Gestaltung der Hilfen für die Beispielaufgabe aus Abbildung 1 gegeben werden (Abschnitt 1.4).

3 tet.folio als Plattform für die technische Umsetzung

Für die oben skizzierten Lösungs- und Visualisierungshilfen wird eine technische Lösung benötigt, die die folgenden Anforderungen erfüllt:

- Die Plattform muss einfach und intuitiv bedienbar sein, und zwar sowohl für die Studierenden, als auch perspektivisch für die Lehrenden, um eine breite und nachhaltige Nutzung zu erreichen.
- Die Hilfen müssen abgestuft dargeboten werden, in dem Sinne, dass man zuerst eine Information erhält, was sich hinter der Hilfe verbirgt, dann ggf., wie viele Punkte in der Klausur für die Nutzung abgezogen werden, bevor mit einem weiteren Klick die eigentliche Hilfe geöffnet wird.

- Die Nutzung der Hilfen muss in Klausursituationen eindeutig personenbezogen nachverfolgbar sein.
- Die interaktiven Simulationen müssen eingebunden werden können. Hierbei ist eine Flexibilität der nutzbaren Formate wünschenswert, damit verschiedene Programme zur Erstellung der Simulationen genutzt werden können oder perspektivisch als Open Educational Resources (OER) bereits vorhandene Simulationen eingebunden werden können.

Nach Abwägung dieser Anforderungen wurde die in der Physikdidaktik bereits für Lehr-Lern-Angebote etablierte Plattform tet.folio der FU Berlin für die technische Umsetzung gewählt (Haase et al., 2016, 2021).

tet.folio ist eine digitale Lehr-Lern-Plattform mit integrierter Autorenumgebung. Sie basiert auf den offenen Internetstandards, HTML5, Javascript und CSS und lässt sich ohne Installation auf allen üblichen Web-Browsern nutzen. Ein besonderer Fokus liegt auf dem Einsatz interaktiver Inhalte. Die Plattform ist so konzipiert, dass neue Funktionalitäten ohne umfangreiche Programmierkenntnisse entwickelt werden können. Dies wurde hier genutzt, um die Hilfe-Funktion zu implementieren. Das dabei erzeugte „tet.tool“ wurde in Javascript geschrieben und ist ein parametrisiertes Plug-In, welches in Zukunft auch in anderen Lehr-Lern-Szenarien eingesetzt werden kann. Die Darstellung für die Lernenden ist in der tet.folio-Umgebung mit allgemeinen HTML5-Elementen nach didaktischen Gesichtspunkten gestaltet. Alle Elemente einer Lernumgebung können gemeinsam von angepassten Format- und Style-Vorlagen in CSS gesteuert werden, und so kann ein einheitliches Erscheinungsbild für die gesamte Lernumgebung gewährleistet werden.

Des Weiteren wurde das tet.tool so ergänzt, dass der Aufruf jeder Hilfe protokolliert werden kann. In Kombination mit dem in tet.folio bereits vorhandenen Plug-In „Action-Logger“ kann damit der gewählte Lernpfad der Lernenden dokumentiert werden. Hierbei sind alle Einzelaktivitäten, wie der Aufruf einer Seite oder einer Hilfe, mit der Uhrzeit des Aufrufs festgehalten. Die gesammelten Navigationsdaten werden auf dem tet.folio-Server gespeichert und lassen sich später im JSON-Format exportieren. Dieses ist ein universelles Format, das zur weiteren Auswertung, z. B. mit Excel oder mit R, eingelesen werden kann.

Die Zugriffsrechte für jede Aufgabe können separat festgelegt werden. So ist es insbesondere möglich, für eine Klausursituation temporär eine Gruppe von Benutzernamen zu definieren, die darauf zugreifen können. Die Benutzernamen werden mit den Navigationsdaten zusammen gespeichert.

In den Übungsphasen ist der Zugriff auf die Aufgaben ohne zeitliche Beschränkung und ohne spezielles Login für alle Studierenden möglich, die über den passenden Link verfügen. Die Navigationsdaten werden dann anonym gesammelt und es gibt bei jeder Aufgabe einen Button, mit dem die Studierenden freiwillig die gesammelten Daten als „Anonymes Interaktionsprotokoll“ zur Auswertung freigeben können. Damit dennoch mehrere Bearbeitungen einer Person zusam-

mengeführt werden können, können die Studierenden ihr Interaktionsprotokoll ebenfalls freiwillig mit einem selbst gewählten Pseudonym versehen, das sie immer wieder verwenden.

4 Einblick: Lösungs- und Visualisierungshilfen zu einer Beispielaufgabe

Die hier exemplarisch vorgestellten Lösungs- und Visualisierungshilfen beziehen sich auf die in Abbildung 1 vorgestellte Aufgabe zur Beschleunigung und Ablenkung von Elektronen in elektrischen und magnetischen Feldern.

Hier ist die erste Hilfe eine Visualisierungshilfe zur Aufgabenstellung. Die Studierenden können eine interaktive Simulation aufrufen, in der sie die in der Aufgabenstellung relevanten Parameter einstellen und variieren können (Abbildung 2). Damit können sie die Wirkung des elektrischen Feldes auf unterschiedlich schnelle Elektronen (Variation der Beschleunigungsspannung) beobachten. Auch eine Variation des magnetischen Feldes für eine weitere Teilaufgabe (in Abbildung 1 nicht dargestellt) ist möglich. Aufgrund der quantitativen Konzeption dieser Visualisierungshilfe erhalten die Studierenden, wenn sie die Beschleunigungsspannung aus Aufgabenteil a einstellen und die elektrische Feldstärke variieren, auch eine Näherungslösung für Aufgabenteil b, mit der sie die Größenordnung ihrer eigenen, rechnerischen Lösung überprüfen können.

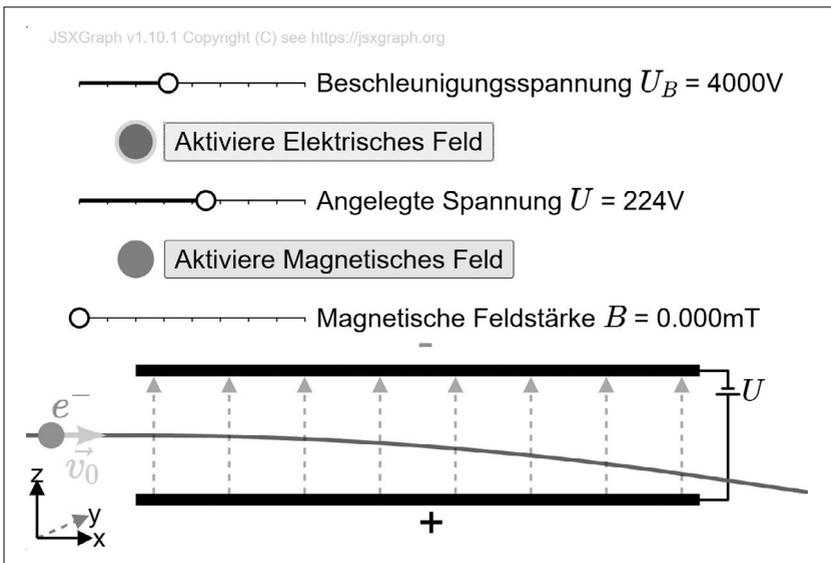


Abbildung 2: Visualisierungshilfe zur Aufgabe aus Abbildung 1 (Screenshot aus tet.folio, Quelle: Eigene Abbildung)

Die weiteren Hilfen folgen weitgehend einem einheitlichen Schema mit aufgabenspezifischen Besonderheiten. Sie bestehen jeweils aus einer Frage und einer dazu passenden Antwort:

- *Welche Werte sind gesucht und welche sind gegeben oder bekannt?* Diese Lösungshilfe enthält keine zusätzlichen Informationen gegenüber der Aufgabenstellung. Sie dient der Klärung der Aufgabenstellung und der Fokussierung auf die wesentlichen Angaben darin.
- *Welchen Ansatz kann ich nutzen?* Hier erhalten die Studierenden einen für die Aufgabe geeigneten Lösungsansatz mit einem Hinweis, wie dieser hier konkret umzusetzen ist. Für den Aufgabenteil a aus Abbildung 1 wäre das z. B. „Energieerhaltung. Die Gesamtenergie setzt sich zusammen aus der kinetischen Energie und der potenziellen elektrischen Energie des Elektrons.“
- *Welche Formeln werden benötigt?* Hier werden die für die Anwendung des Ansatzes nötigen Formeln angegeben. Dabei wird der Ansatz nicht mehr explizit genannt, sodass die Hilfen weitgehend unabhängig voneinander genutzt werden können (Abbildung 3). Dies ist insbesondere für den Klausureinsatz wichtig.

Aufgabenspezifische Lösungshilfen, die diese Grundstruktur ergänzen, bestehen z. B. in Hinweisen auf Analogien (z. B. „Mit welchem Problem aus der klassischen Mechanik ist die Flugbahn des Elektrons vergleichbar?“) oder der Angabe von Randbedingungen.

Zur Evaluation der Hilfen in den Übungsphasen wurden zusätzlich Feedback-Buttons implementiert, mit denen die Studierenden rückmelden können, ob die erhaltene Hilfe ihrer Einschätzung nach zu der Frage gepasst hat und ob sie hilfreich war (Abbildung 4). Das Feedback wird gespeichert und kann als Teil des Interaktionsprotokolls freiwillig eingereicht werden.

Formeln:	$E_{ges} = E_{kin} + E_{el}$
Welche Formeln werden benötigt?	$E_{kin} = \frac{1}{2} m_e \cdot v^2$
	$E_{el} = e \cdot U_a$
	E_{ges} : Gesamtenergie
	E_{kin} : kinetische Energie
	E_{el} : potentielle elektrische Energie
	m_e : Elektronenmasse
	v : gesuchter Betrag der Geschwindigkeit
	e : Elementarladung
	U_a : Beschleunigungsspannung

Abbildung 3: Lösungshilfe zu Aufgabenteil a aus Abbildung 1 (Screenshot aus tet.folio, Quelle: Eigene Abbildung)

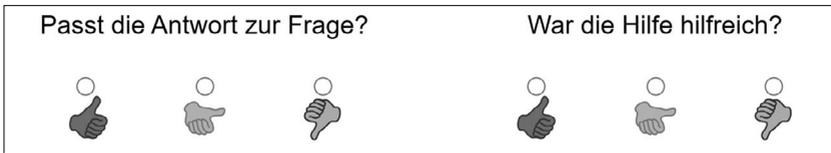


Abbildung 4: Kurzabfragen zur Evaluation der Hilfen (Screenshot aus tet.folio, Quelle: Eigene Abbildung)

Bei jeder Hilfe ist zunächst nur die Frage sichtbar, die den Studierenden einen möglichst präzisen Eindruck vermitteln soll, was für Informationen sie beim Aufruf der Hilfe erhalten. Daneben steht ein Button „Hilfe nutzen“, mit dem sie sich die Antwort anzeigen lassen können.

Für den Einsatz in einer Prüfungssituation wird ein Rückfrage-Button zwischengeschaltet, auf dem angegeben ist, wie viele Punkte für die Nutzung der jeweiligen Hilfe verrechnet werden (Abbildung 5).



Abbildung 5: In Prüfungssituationen müssen die Hilfen gegen Punkte „gekauft“ werden (Screenshot aus tet.folio, Quelle: Eigene Abbildung)

Damit soll auch verhindert werden, dass Studierende versehentlich eine Hilfe aufrufen und Punkte dafür abgeben. Die pro Hilfe verrechnete Punktzahl variiert von Hilfe zu Hilfe. Ihre Summe über eine gesamte Klausur ist dadurch limitiert, sodass man auch bei Nutzung aller Hilfen noch die Möglichkeit haben muss, die Klausur mit „ausreichend“ zu bestehen.

5 Erste Erfahrungen und Ausblick: Perspektiven für den Einsatz im Übungsbetrieb und in Prüfungen

Bisher wurden insgesamt mehr als 10 Aufgaben zu verschiedenen Themen der Elektrostatik, Elektrodynamik und Optik entwickelt, u. a. zum Aufladen eines Plattenkondensators, zur Ablenkung von Elektronen durch elektrische und magnetische Felder (Abbildung 1), zu Widerstandskombinationen und zum elektromagnetischen Schwingkreis. Diese Aufgaben wurden regelmäßig im Übungsbetrieb und zur Vorbereitung auf die Klausuren in den Vorlesungen der Zielgruppen

in Duisburg und Essen (vgl. Abschnitt 1.2) eingesetzt.

Hierbei wurde auch der Einsatz des oben beschriebenen anonymen Interaktionsprotokolls erprobt, in dem aufgezeichnet ist, welche Hilfen die Studierenden wann in Anspruch genommen haben und wie sie diese bewerten (Abbildung 4). Die bislang vorliegenden Interaktionsprotokolle lassen nur erste qualitative Schlüsse auf die Nutzung des Angebots zu. Die Studierenden nutzen die Hilfen individuell sehr unterschiedlich, nach unterschiedlichen Zeiten (bezogen auf den Start der Aufgabenbearbeitung) und in unterschiedlichen Reihenfolgen. Dies lässt vermuten, dass sie die Hilfen nicht lediglich „gesichtet“, sondern selektiv gemäß subjektiv empfundenem Bedarf genutzt haben. Es gibt jedoch auch Studierende, die nach kurzer Zeit alle Hilfen öffnen, evtl. um sich einen Überblick zu verschaffen. Dies hat in der Übungssituation keine Konsequenzen. In Prüfungssituationen ist jedoch eine selektivere Nutzung sinnvoll und sollte in weiteren Erprobungen mit den Studierenden trainiert werden. Die Bewertung der Hilfen fällt durchweg positiv aus („grüner Daumen“, Abbildung 4). Die Fragen scheinen somit zum einen ihren Zweck zu erfüllen, den Studierenden zu vermitteln, welche Informationen sie bei Aufruf der Hilfe zu erwarten haben. Zum anderen werden sie als hilfreich bewertet.

Diese Schlussfolgerungen stehen unter dem Vorbehalt, dass sich nur sehr wenige Studierende beteiligt haben und die Rückmeldungen mutmaßlich eher von denjenigen stammen, die sich ernsthaft mit der Aufgabe und den Hilfen befasst haben. Für den weiteren Einsatz muss daher die Abgabe der Interaktionsprotokolle intensiver beworben und die Evaluation intensiviert werden.

Für die Folgesemester soll das Angebot erweitert und als fester Bestandteil von Beginn an in den regulären Übungsbetrieb integriert werden. Dazu wird zum einen das Angebot auf die Vorlesung des ersten Fachsemesters (Experimentalphysik 1: Mechanik und Wärmelehre) erweitert und zum anderen auf der Plattform tet.folio eine Sammlung der Aufgaben angelegt, auf welche die regelmäßig wechselnden Dozierenden der entsprechenden Veranstaltungen zugreifen können.

In den Klausuren am Ende des Sommersemesters 2024 wurden die Lösungshilfen erstmals eingesetzt. Dies erfolgte aus technischen Gründen zunächst papierbasiert. Bei der Aufgabe war angegeben, welche Hilfen zur Verfügung stehen und wie viele Punkte dafür abgezogen werden. Die Studierenden bekamen die angeforderten Hilfen auf Papier ausgehändigt.

Literatur

Binder, T./Schmiemann, P./Theyßen, H. (2019). Knowledge Acquisition of Biology and Physics University Students – the Role of Prior Knowledge. In *Education Sciences*, 9(4), Article number: 281.

- Buschhüter, D./Spoden, C./Borowski, A. (2017). Studienerfolg im Physikstudium: Inkrementelle Validität physikalischen Fachwissens und physikalischer Kompetenz. In *Zeitschrift für Didaktik der Naturwissenschaften*, 23(1), 127–141.
- Franke-Braun, G./Schmidt-Weigand, F./Stüdel, L./Wodzinski, R. (2008). Aufgaben mit gestuften Lernhilfen- ein besonderes Aufgabenformat zur kognitiven Aktivierung der Schülerinnen und Schüler und zur Intensivierung der sachbezogenen Kommunikation. In Kasseler Forschergruppe (Hrsg.), *Kasseler Forschergruppe Empirische Bildungsforschung: Lehren – Lernen – Literacy*, Band 2 (S. 27–42). Kassel: kassel university press.
- Friege, G. (2001). Wissen und Problemlösen: eine empirische Untersuchung des wissenszentrierten Problemlösens im Gebiet der Elektrizitätslehre auf der Grundlage des Experten-Novizen-Vergleichs. Berlin: Logos Verlag.
- Haase, S./Kirstein, J./Nordmeier, V. (2016). tet. folio: Neue Ansätze zur digitalen Unterstützung individualisierten Lernens. In *PhyDid B – Didaktik Der Physik – Beiträge zur DPG-Frühjahrstagung*. <https://ojs.dpg-physik.de/index.php/phydid-b/article/view/737> [Zugriff: 31.10.2023].
- Haase, S./Sommerer, M./Kirstein, J./Nordmeier, V. (2021). tet.folio: Eine Online-Plattform für die Produktion innovativer Lehr-Lern-Angebote. In: *PhyDid B – Didaktik Der Physik – Beiträge zur DPG-Frühjahrstagung 1*. <https://ojs.dpg-physik.de/index.php/phydid-b/article/view/1117> [Zugriff: 31.10.2023].
- Haak, I. (2017). Maßnahmen zur Unterstützung kognitiver und metakognitiver Prozesse in der Studieneingangsphase. Berlin: Logos Verlag.
- Heublein, U./Hutzsch, C./Schmelzer, R. (2022). Die Entwicklung der Studienabbruchquoten an den deutschen Hochschulen. In: *DZHW Brief 2022*, 5, 1–16.
- Hußmann, S./Melle, I./Selter, C./Theyßen, H. (2012). dortMINT: Diagnose und individuelle Förderung in der Lehrerbildung. In S. Bernholt, Sascha (Hrsg.), *Konzepte fachdidaktischer Strukturierung für den Unterricht* (S. 431–433). Berlin: LIT Verlag.
- Pusch, A. (2014). Fachspezifische Instrumente zur Diagnose und individuellen Förderung von Lehramtsstudierenden der Physik. Berlin: Logos Verlag.
- Sorge, S./Petersen, S./Neumann, K. (2016). Die Bedeutung der Studierfähigkeit für den Studienerfolg im 1. Semester in Physik. In: *Zeitschrift für Didaktik der Naturwissenschaften*, 22, 165–180.
- Woitkowski, D. (2021). Fachwissen und Problemlösen im Physikstudium. In *PhyDid B – Didaktik Der Physik – Beiträge zur DPG-Frühjahrstagung*. <https://ojs.dpg-physik.de/index.php/phydid-b/article/view/1187> [Zugriff: 31.10.2023].

Variation von Aufgabenstellungen zu Modellierung in der Informatik

Marcellus Sieburg und Janis Voigtländer

Zur Unterstützung der Lehre und des Lernens setzen wir seit einigen Jahren in der Veranstaltung „Modellierung“ in der Lehrereinheit Informatik die E-Learning-Plattform Autotool mit eigenen Aufgabentypen ein. Die Aufgaben werden größtenteils automatisch generiert und bewertet (Kafa et al., 2020; Sieburg/Voigtländer, 2020; Brandt et al., 2022). Im Folgenden werden Aufgabentypen aus einer der unterstützten Domänen vorgestellt und hinsichtlich des durch sie adressierten Kompetenzniveaus beurteilt. Zur Einordnung wird die Taxonomie von Bloom (1976) verwendet. Es wird angestrebt, mit den Aufgaben ein möglichst hohes Kompetenzniveau zu üben und zu prüfen. Im Rahmen des Projekts wurde unser eigens entwickeltes Framework zur Aufgabengenerierung in vielerlei Hinsicht erweitert, um mehr Themen aus der Lehrveranstaltung abzubilden, mehr Diversität in den Aufgabenstellungen zu erreichen, spezielle Lehrwünsche genauer abzubilden, und zusätzliche Feedbackmöglichkeiten zu schaffen. Um die Stärken des Frameworks aufzuzeigen, stellt dieser Artikel einzelne Aufgabentypen vor und entwickelt schließlich einen Lernpfad über eine Abfolge dieser hinweg.

1 Die Domäne: Übungsaufgaben zu Klassendiagrammen

In der Informatik werden beim Systementwurf oft Klassendiagramme verwendet. Zum Beispiel könnte ein Ausschnitt einer Campusverwaltungssoftware wie in Abbildung 1 modelliert werden. Dabei stehen die Kästen für sogenannte Klassen und die Kanten dazwischen für bestimmte Arten von Beziehungen. Es gibt syntaktische Regeln, die bei der Verwendung dieser grafischen Modellierungssprache eingehalten werden müssen. Außerdem ist die inhaltliche Passung zu beachten, die von der Semantik abhängt. Die Semantik, also Bedeutung, eines Klassendiagramms lässt sich kurz wie folgt fassen: Ein Klassendiagramm beschreibt eine bestimmte Menge von Objektdiagrammen. Solche Objektdiagramme wiederum entsprechen konkreten Weltausschnitten. Wir werden sie erst bei der Diskussion von Feedbackmöglichkeiten an Übenende gegen Ende des Artikels wieder thematisieren. Unser Fokus liegt zunächst auf der syntaktischen Form von Klassendiagrammen, die wie in Abbildung 1 vereinfacht dargestellt

werden, indem sie bestimmte Modellierungselemente der Klassen weglassen, um den Blick gezielt auf die Beziehungen zu richten.

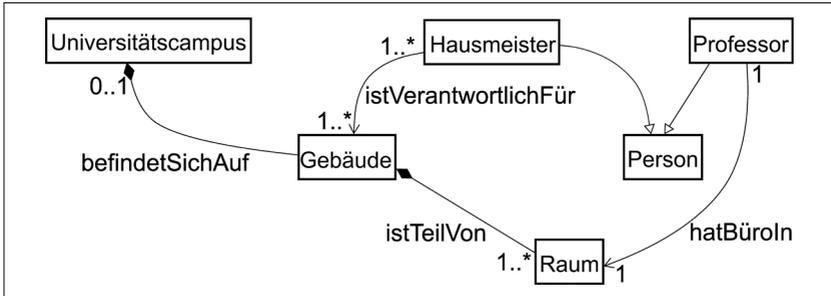


Abbildung 1: Ein Klassendiagramm (Quelle: Eigene Abbildung)

Im Folgenden werden Aufgabentypen zur (vorwiegend syntaktischen) Validität von Klassendiagrammen betrachtet. Das Lösen unserer Aufgaben dreht sich um das Erkennen von Fehlerursachen. Mögliche syntaktische Fehler können zum Beispiel unerlaubte Zyklen oder ungültige Beschriftungen sein. Um die Gültigkeit oder Nichtgültigkeit festzustellen, kann eine Strategie darin bestehen, alle für Klassendiagramme geltenden Syntaxregeln zu kennen und diese einzeln, nacheinander zu überprüfen. In diesem Sinne werden die Syntaxregeln in der Vorlesung eingeführt. Parallel dienen sie als Grundlage für unsere Aufgabengeneratoren in Gestalt von logischen Formalisierungen in Alloy, einem Modelchecker-Tool, das auch verwendet werden kann, um Instanzen zu generieren, die gezielt bestimmte Konstellationen erfüllen (Jackson, 2006). Dies setzen wir ein, indem mittels formaler Logik die Konzepte „Klassendiagramm“, „gültige und ungültige Eigenschaften eines Klassendiagramms“ sowie „Änderungen an einem Klassendiagramm“ präzise definiert werden. Eine Eigenschaft ist dabei so etwas wie „Zyklenfreiheit“ oder „hat eine Beziehung mit ungültiger Beschriftung“. Eine Änderung wäre das Hinzufügen, Löschen oder Austauschen einer Beziehungskante. Die im Folgenden beschriebenen Aufgabentypen (bis auf „Szenario“) erlauben auf dieser Basis die automatische Erzeugung von Aufgabenstellungen.

2 Diverse Aufgabentypen

Der erste Aufgabentyp „Korrekte auswählen“ realisiert die direkteste Fragestellung zur Gültigkeit von Klassendiagrammen, nämlich ob vorgegebene Kandidaten gültige Klassendiagramme sind oder nicht. Wie exemplarisch in Abbildung 2 zu sehen, werden den Lernenden vier verschiedene Klassendiagrammkandidaten angezeigt, für die sie jeweils entscheiden müssen, ob es sich um ein gülti-

ges Klassendiagramm handelt oder nicht. Dass hierbei abstrakt gehaltene Namen der Klassen und Beziehungen verwendet werden, ist Absicht und auch bei allen weiteren automatisch erzeugbaren Aufgabenstellungen der Fall. Dadurch folgt eine Bearbeitung von Fragen zur syntaktischen Korrektheit einzig den Regeln hierzu, d. h., die Lernenden sind gezwungen, sich von inhaltlichen bzw. realweltlichen Interpretationen zu lösen und unabhängig davon zu Einschätzungen der Validität zu kommen.

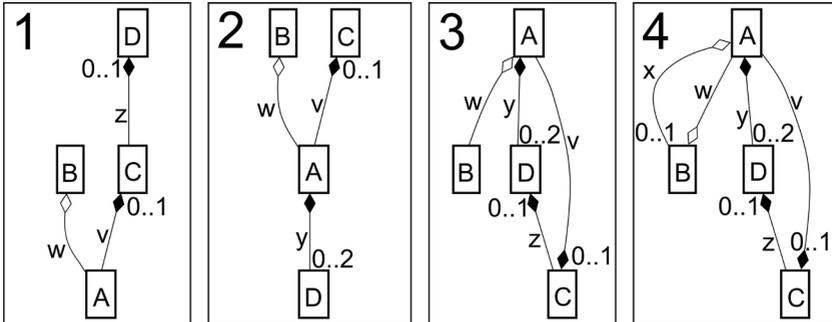


Abbildung 2: Beispielinstanz für Aufgabentyp „Korrekte auswählen“ (Quelle: Eigene Abbildung)

Der in Abbildung 3 schemenhaft dargestellte Generator nutzt die Formalisierungen zu Klassendiagrammen, ihren Eigenschaften und abstrakten Änderungen, um auf verschiedene Weisen ungültige und gültige Klassendiagramme zu erzeugen, und zwar derart, dass darunter auch Distraktoren existieren (Sieburg/Voigtländer, 2020). Dabei wird die gesamte Aufgabeninstanz samt für die Bewertung und Feedback nutzbarer Musterlösungen und Erklärungsstücke erstellt. Dem Generator werden fünf Eigenschaftsbeschreibungen übergeben. Unter strikter Einhaltung dieser erzeugt er ein Ausgangsklassendiagramm und vier Änderungsbeschreibungen, woraus zusammen er wiederum die vier Klassendiagrammkandidaten ableitet, die anzuzeigen sind.

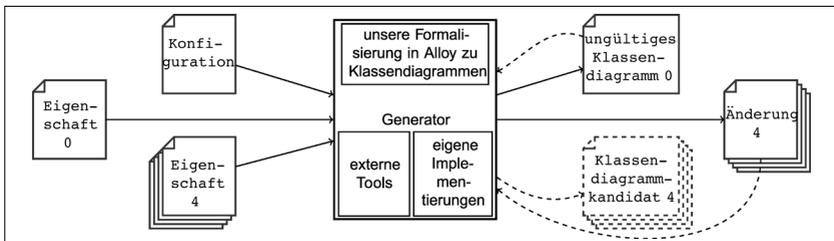


Abbildung 3: Prinzipielle Funktionsweise des Generators (Quelle: Eigene Abbildung)

Damit die Klassendiagrammkandidaten nicht völlig verschieden voneinander sind und damit Distraktoren existieren, werden sie aus demselben, ungültigen Ursprungsklassendiagramm abgeleitet. Die an den Generator kommunizierten Eigenschaften determinieren die Gültigkeit der Kandidaten. In Abbildung 2 sind der erste und zweite Kandidat gültige Klassendiagramme. Der dritte Klassendiagrammkandidat besitzt einen nicht zulässigen Zyklus bestimmter Beziehungen, hier v , y und z . Der vierte besitzt denselben Zyklus und zusätzlich die zulässige Eigenschaft, in entgegengesetzte Richtungen zeigende Beziehungen einer anderen Art zwischen zwei Klassen zu haben (hier x und w , ein erlaubter Zyklus).

Durch die einfache Fragestellung ist der Aufgabentyp „Korrekte auswählen“ eher einem niedrigen Kompetenzniveau zuzuordnen. Zwar ist es aus Lehrendensicht zunächst naheliegend, dass zur Bearbeitung der Aufgabe Verständnis erforderlich sei, wann ein Klassendiagramm ungültig wird, und daher folgerbar, dass hier das Kompetenzniveau Verstehen gefragt sei. Allerdings zeigt sich bei näherer Betrachtung, dass zum Lösen der Aufgaben solch ein Verständnis gar nicht unbedingt vorliegen muss. Denn es lässt sich durch Üben antrainieren, zu erkennen, bei welchen auftretenden optischen Mustern das Klassendiagramm ungültig ist, sodass reines Wiedererkennen ausreicht. Daher ist die Aufgabe eher im Kompetenzniveau Wissen anzusiedeln. Allerdings können wir den Aufgabentyp anpassen, um ein höheres Kompetenzniveau zu erreichen: Es kann Verständnis statt Fakten geprüft werden, indem nach dem Warum gefragt wird (Bücking, 2014, S. 14).

Die Anwendung dieses Prinzips führt zum Aufgabentyp „Fehler benennen“. Dass der Generator für jeden Klassendiagrammkandidaten die vorab angegebenen Eigenschaften kennt, wird nun ausgenutzt, um diese Information auch bei der Fragestellung zu verwenden. Die Generierung wird insofern angepasst, dass nur noch einer statt vier Kandidaten angefordert wird. Die Aufgabenstellung wird weiter verfeinert, indem außerdem nach den am Fehler beteiligten Beziehungen gefragt wird. Wie die zulässigen und unzulässigen Eigenschaften konkret mit den einzelnen Beziehungen interagieren, spielte beim vorherigen Aufgabentyp nur implizit eine Rolle, hier wird nun explizit damit gearbeitet.

In Abbildung 4 ist ein Beispiel einer Instanz für diesen Aufgabentyp dargestellt. Hier müssten die Lernenden zum korrekten Lösen feststellen, dass die Fehlerursache ein Kompositionszyklus ist und dass die Beziehungen y und z diesen verursachen. Die Fragestellung erfordert nun außer Kenntnis über die Komponenten eines Klassendiagramms auch Verständnis zu deren korrekter Verwendung, weil nach dem direkten Zusammenhang gefragt wird. Dementsprechend liegt das Kompetenzniveau hier im Bereich Verstehen und damit höher als beim Aufgabentyp „Korrekte auswählen“.

Es ist aber auch möglich, den Aufgabentyp „Korrekte auswählen“ auf eine andere Weise anzupassen, unter kaum veränderter Verwendung des Generators. Es werden wieder andere Teile zur Darstellung verwendet, dadurch die Perspektive variiert. Um die Analogie zu verdeutlichen, wird in Abbildung 5 genau die-

Betrachten Sie das folgende Klassendiagramm, welches leider ungültig ist.

```

classDiagram
    class C
    class D
    class B
    class A
    C "2..*" -- "1" C : w
    C "1" -- "1..2" D : x
    D "0..1" -- "2..*" B : y
    D "1" -- "2..*" B : z
    D "1..2" -- "1..2" A : v
    C "1" *-- "1..2" D : x
    
```

Es enthält die folgenden Beziehungen zwischen Klassen:

- 1) Assoziation w
- 2) Komposition y
- 3) Aggregation x
- 4) Komposition v
- 5) Komposition z

Wählen Sie aus, was Sie für den einen Grund halten, dass dieses Klassendiagramm ungültig ist, und nennen Sie alle Beziehungen, die definitiv zum Problem beitragen, d.h. deren Entfernung das Problem jeweils beheben würde.

Mögliche Gründe sind: Das Klassendiagramm ...

- a) enthält mindestens eine Selbstvererbung.
- b) enthält mindestens eine Selbstbeziehung, die keine Vererbung ist.
- c) enthält mindestens eine ungültige Multiplizität am Ganzen einer Komposition.
- d) enthält mindestens einen Vererbungszyklus.
- e) enthält mindestens eine Mehrfachvererbung.

... [weitere fünf Gründe]

Abbildung 4: Beispielinstanz für Aufgabentyp „Fehler benennen“ (Quelle: Eigene Abbildung)

jenige Instanz dieses neuen Aufgabentyps „Klassendiagramm reparieren“ dargestellt, die der bereits zum Aufgabentyp „Korrekte auswählen“ gezeigten Instanz entspricht. Es wird das ungültige Klassendiagramm 0 angezeigt, das Abbildung 2 nicht darstellt, das aber im Hintergrund vom Generator erzeugt worden war (siehe Abbildung 3). Außerdem werden die dort zum Erzeugen der vier anzuzeigenden Klassendiagrammkandidaten verwendeten Änderungen hier nun verbalisiert und als Reparaturangebote ausgegeben. Die Aufgabenstellung ändert sich, weil das nun allein vorgegebene Klassendiagramm definitiv ungültig ist und alle gültigen Reparaturen ausgewählt werden müssen. Auch bei dieser Veränderung des Aufgabentyps „Korrekte auswählen“ erhöht sich das Kompetenzniveau. Zur Beantwortung der Fragestellung sind die genannten Änderungen zu verstehen und auf das ungültige Klassendiagramm anzuwenden, um dann das jeweils resultierende Klassendiagramm bezüglich der Gültigkeit einzuschätzen. Es gibt einfache Fälle, in denen zur Lösung der Aufgabe statt einer Gesamtbetrachtung die reine Einschätzung der Gültigkeit der hinzugefügten bzw. gelöschten Beziehungen ausreicht, um zum korrekten Ergebnis zu gelangen. Werden solche Fälle jedoch ausgeschlossen, ist dieser Aufgabentyp mindestens dem Kompetenzniveau Verstehen

zuzuordnen. Durch die Notwendigkeit, die Änderungen am Klassendiagramm zumindest gedanklich (und eventuell sogar mit Stift und Papier) durchzuführen, lässt sich argumentieren, dass sogar das Niveau Anwenden erreicht werde.

Betrachten Sie das folgende Klassendiagramm, welches leider ungültig ist.

Welche der folgenden Änderungen würden das Klassendiagramm reparieren?

- 1) entferne Komposition y
- 2) entferne Komposition z
- 3) ersetze Aggregation w durch eine Beziehung, die A eine Aggregation aus Bs macht, wobei A 0..*-mal und B 0..1-mal beteiligt ist
- 4) ergänze eine Beziehung, die A eine Aggregation aus Bs macht, wobei A 0..*-mal und B 0..1-mal beteiligt ist

Abbildung 5: Beispielinstanz für Aufgabentyp „Klassendiagramm reparieren“
(Quelle: Eigene Abbildung)

Zum Aufgabentyp „Fehler benennen“ zurückkommend, können wir diesen als Grundlage nehmen, um durch das Einfordern einer Beurteilung einer vorgegebenen Antwort das Kompetenzniveau nochmals zu steigern (Bücking, 2014, S. 14). Dies gelingt, indem die Domäne der Entscheidung der Gültigkeit von Klassendiagrammen auf die semantische Ebene erweitert wird.

Abbildung 6 zeigt einen möglichen Vorgabetext zum neuen Aufgabentyp „Szenario“. Er wird zusammen mit dem Klassendiagramm aus Abbildung 1 präsentiert, wobei gesagt wird, dass dieses als Lösungsversuch zum Szenario-Text entstanden sei. Die nun zu bearbeitende Aufgabenstellung befindet sich in Abbildung 7. Das präsentierte Klassendiagramm ist zu analysieren und zu untersuchen, ob es zum gegebenen Szenario passt. Dabei ist auf sowohl syntaktische als auch semantische Korrektheit zu achten. Im konkreten Fall ist zum Lösen festzustellen, dass der letzte Teil des letzten Satzes des Szenario-Textes durch die *I* neben *Professor* an der Beziehung *hatBüroIn* verletzt wird. Die korrekte Antwort zu Abbildung 7 ist also „d, [2]“. Das dafür aufzubringende Verständnis für den

Gesamtzusammenhang dieser Aufgabe hat zur Folge, dass das überprüfte Kompetenzniveau im Bereich Analysieren liegt.

Eine Studentin hat folgendes Szenario erhalten:
 Ein Universitätscampus besteht aus verschiedenen Gebäuden. Jedes Gebäude wird von einem Hausmeister betreut, der für es zuständig ist. Ein Hausmeister ist eine Person. Eine andere Art von Personen sind die Professoren, die jeweils einen bestimmten Raum als eigenes Büro haben. Ein Gebäude besteht aus verschiedenen Räumen, von denen nicht jeder ein Professorenbüro ist.

Abbildung 6: Szenario-Text für eine Aufgabe vom Typ „Szenario“ (Quelle: Eigene Abbildung)

Analysieren Sie das und beziehen Sie Stellung zum entworfenen Klassendiagramm hinsichtlich der ursprünglichen Aufgabe!
 Das Klassendiagramm ...

- a) enthält einen Syntaxfehler
- b) hat gar kein passendes Objektdiagramm
- c) ist vollständig richtig modelliert
- d) hält eine Vorgabe des Szenarios nicht ein

Nennen Sie die konkreten Beziehungen, die geändert werden müssten, falls solche existieren:

- 1) Assoziation ist Verantwortlich Für
- 2) Assoziation hat Büro In
- 3) Komposition befindet Sich Auf
- 4) Komposition ist Teil Von
- 5) eine Vererbung, bei der Professor von Person erbt
- 6) eine Vererbung, bei der Hausmeister von Person erbt

Abbildung 7: Aufgabenstellung für eine Aufgabe vom Typ „Szenario“ (Quelle: Eigene Abbildung)

3 Gemeinsame Betrachtungen und Zusammenführung

Der Aufgabentyp „Szenario“ erfordert manuelles Erstellen eines Szenarios und eines Klassendiagramms. Unsere logische Formalisierung reicht nicht aus, um eine solche Aufgabe automatisch zu erstellen. Bei allen anderen Aufgabentypen hingegen kann die Formalisierung sogar genutzt werden, um erklärendes Feedback zu Antworten Übender zu geben. Dabei können entweder schon bei der Generierung der Aufgabe erzeugte oder verwendete Daten direkt genutzt oder die Formalisierung mit ihnen gespeist werden, um nützliche Informationen zu erhalten. Beispielsweise fallen beim Aufgabentyp „Klassendiagramm reparieren“ die vier resultierenden Klassendiagramme mit an. Beim Aufgabentyp „Korrekte aus-

wählen“ würden sie direkt angezeigt. Sie können aber auch bei „Klassendiagramm reparieren“ verwendet werden, um sie bei einer fehlerhaften Antwort als Rückmeldung auszugeben, mit dem Hinweis, dass die Antwort zum entsprechenden Reparaturversuch nicht korrekt ist, da das nunmehr angezeigte Klassendiagramm resultieren würde.

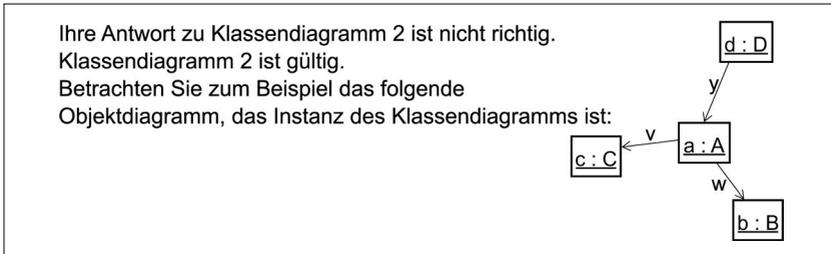


Abbildung 8: Feedback mit Objektdiagramm (Quelle: Eigene Abbildung)

Bei allen besprochenen Aufgabentypen wiederum gibt es im Fall von in Einreichungen fälschlicherweise als syntaktisch ungültig kategorisierten Klassendiagrammen die Möglichkeit, ein jeweils passendes Objektdiagramm anzuzeigen, mit der Meldung, dass das Klassendiagramm doch gültig ist und diese Objekt-Ausprägung existiert. Objektdiagramme sind eine weitere grafische Modellierungssprache. Für den Kontext hier reicht es aus, zu wissen, dass ein Objektdiagramm als eine konkrete Ausprägung eines Klassendiagramms angesehen werden kann. Passende Objektdiagramme können automatisch erzeugt werden, weil auch dafür eine Formalisierung existiert¹ (Kafa et al., 2020). Ein Beispiel für die Rückmeldung mit so einem Objektdiagramm ist in Abbildung 8 zu sehen, wobei der Fall dargestellt wird, dass in der Beispielinstant zu „Korrekte auswählen“ das zweite Klassendiagramm als ungültig kategorisiert wurde. Wenn ein Klassendiagramm hingegen fälschlicherweise als gültig kategorisiert wurde, kann ein Hinweis angezeigt werden, dass es (erst) gültig wäre, wenn eine bestimmte Beziehung nicht existieren würde. Dies weist direkt auf die problematische Stelle, wodurch eine andere Sichtweise und somit die Möglichkeit gegeben wird, fehlerhafte Antworten besser zu reflektieren. Auch hier lässt sich der Hinweis aus bereits vorhandenen Zutaten ableiten, da die laut Abbildung 3 gelieferten Änderungsbeschreibungen genau solche „problematischen Beziehungen“ berühren oder ins Spiel bringen.

Alle Aufgabentypen können gezielt über Konfigurationsparameter angesteuert werden, wodurch sich unter anderem der Schwierigkeitsgrad der Aufgaben

¹ Es gibt auch diverse Aufgabentypen zur Verbindung zwischen Klassen- und Objektdiagrammen (Siegburg/Voigtländer, 2020). Zu diesen können ähnliche Betrachtungen angestellt werden wie die hier gezeigten für den Bereich Validität von Klassendiagrammen.

anpassen lässt. So können etwa quantitative Angaben eingestellt werden, wie die Anzahl der Klassen und Beziehungen durch obere und untere Schranken, auch jeweils nach Beziehungstyp. Bestimmte Beziehungstypen ganz auszuschließen, ist eine weitere Möglichkeit, den Schwierigkeitsgrad zu regulieren oder die Aufgaben dem jeweiligen Stand in der Vorlesung anzupassen. Eine andere wäre, bestimmte gültige oder ungültige Eigenschaften eines Klassendiagramms auszuschließen. Teilweise können eigene Distraktoren definiert und beim Aufgabentyp „Szenario“ sogar der Aufgabentext komplett frei gewählt werden. Zusätzlich sind bestimmte optionale Darstellungsweisen, etwa die Ausgabe von Navigationsrichtungen (Pfeilspitzen) bei bestimmten Beziehungen, ansteuerbar. Manche zu treffenden Auswahlen, wie die Deaktivierung optionaler Namensangaben an Beziehungen, führen neben Änderung der Darstellung im Klassendiagramm automatisch auch zu geänderten Formulierungen im eigentlichen Aufgabentext (etwa bei verbalen Beschreibungen in „Klassendiagramm reparieren“), wodurch sich wieder eine leicht andere Perspektive ergibt. Es kann auch eingestellt werden, ob es in bestimmten Fällen erweitertes Feedback geben und ob nach einer Einreichung die korrekte Lösung angezeigt werden soll. Ob die Einreichung selbst korrekt ist, wird immer vermeldet, wobei sich dieses Feedback für prüfungsimmanente Aufgaben bis zur Einreichungsfrist verzögern lässt.

Die Konfigurierbarkeit der Aufgaben und die Feedback-Möglichkeiten helfen dabei, die Aufgabentypen gezielt in einem Lernpfad einzusetzen, um bei fortschreitendem Lernen im zu übenden Themenfeld zu unterstützen. Ein solcher Pfad könnte beispielsweise zunächst den Aufgabentyp „Korrekte auswählen“ einsetzen, damit sich die Übenden im Einschätzen von Korrektheit versuchen, wobei jeweils Feedback bei Fehleinschätzungen von ungültigen Klassendiagrammen angezeigt wird, das explizit auf den konkret vorliegenden Fehler hinweist. Danach wird der Aufgabentyp „Fehler benennen“ verwendet, um sie dieselben Beobachtungen, die es zuvor als Feedback gab, jetzt selbst treffen zu lassen. Bei Falschangabe des Fehlergrunds könnte das ursprüngliche Diagramm, wenn möglich, vom vermeintlichen Fehler bereinigt werden unter der Meldung, dass das Diagramm trotz Behebung des vermeintlichen Fehlers noch nicht korrekt ist. Nach dieser Aufgabe, die im Feedback gezeigt hat, wie Klassendiagramme durch gezielte Änderungen umgestaltet werden, könnte dann eine Aufgabe des Typs „Klassendiagramm reparieren“ gestellt werden, bei der es nun darum geht, die von der Aufgabe vorgeschlagenen Änderungen des Klassendiagramms eigenständig durchzuführen und dann die Gültigkeit der resultierenden Klassendiagramme zu beurteilen. Schließlich könnte der Aufgabentyp „Szenario“ verwendet werden, um Bezug zu einem realweltlichen Beispiel statt der vorher bearbeiteten, abstrakten Beispiele herzustellen und somit die zuvor gewonnenen Erfahrungen anzuwenden.

Die vorgestellten Aufgabentypen haben gezeigt, wie unser Framework verwendet werden kann, um verschiedene Kompetenzniveaus im Bereich Gültigkeit von Klassendiagrammen zu adressieren. Die Möglichkeit von erweitertem

Feedback hilft bei der Gestaltung von Lernpfaden wie im vorigen Absatz skizziert. Darüber hinaus bestehen zahlreiche Einsatzmöglichkeiten, die noch weiter exploriert werden können, zum Beispiel die Gestaltung von fairen Prüfungen (Metzger/Nüesch, 2004). Eine gute Eignung dafür suggerieren geringer Aufwand beim Erstellen und Bewerten von Aufgaben (Ökonomie) und reproduzierbare Bewertungen (Objektivität).

Literatur

- Bloom, B. S. (1976). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim: Beltz.
- Brandt, A./Siegburg, M./Voigtländer, J./Wang, K. (2022). A Report on Automatic Generation of Petri Net Exercise and Exam Task Instances. In J. Michael/J. Pfeiffer/A. Wortmann (Hrsg.), *Modellierung 2022 – Workshop Proceedings*, Hamburg, Germany, June 27–July 1, 2022 (S. 197–204). Gesellschaft für Informatik e. V.
- Bücking, J. (2014). Workshopunterlagen, Gestaltung geschlossener Fragen für Übungen und Prüfungen, 19.05.2014. http://www.eassessment.uni-bremen.de/documents/HandoutWorkshopTU-Darmstadt2014_buecking.pdf [Zugriff: 30.10.2023]
- Jackson, D. (2006). *Software Abstractions. Logic, Language, and Analysis*. Cambridge, Massachusetts: The MIT Press.
- Kafa, V./Siegburg, M./Voigtländer, J. (2020). Exercise Task Generation for UML Class/Object Diagrams, via Alloy Model Instance Finding. In B. Tait et al. (Hrsg.), *ICT Education. SACLA 2019. Communications in Computer and Information Science*, Band 1136 (S. 112–128). Cham: Springer.
- Metzger, C./Nüesch, C. (2004). Fair Prüfen. Ein Qualitätsleitfaden für Prüfende an Hochschulen. In D. Euler/C. Metzger (Hrsg.), *Hochschuldidaktische Schriften*, Band 6. St. Gallen: Universität St. Gallen, Institut für Wirtschaftspädagogik.
- Siegburg, M./Voigtländer, J. (2020). Generating Diverse Exercise Tasks on UML Class and Object Diagrams, Using Formalisations in Alloy. In J. Michael et al. (Hrsg.), *Companion Proceedings of Modellierung 2020 Short, Workshop and Tools & Demo Papers co-located with Modellierung 2020*, Vienna, Austria, February 19–21, 2020. *CEUR Workshop Proceedings*, Band 2542 (S. 89–100). CEUR-WS.org.

Erste Schritte auf dem Weg zu einer digitalen Lern- und Prüfungsumgebung für die organische Chemie

Katrin Schübler, Michael Giese und Maik Walpuski

Die Auseinandersetzung mit komplexen organischen Molekülen ist zentral für Lehre und Prüfungen in der organischen Chemie. Ziel der Einführungsveranstaltung „Organische Chemie 1“ ist es, dass Studierende den Aufbau dieser Moleküle verstehen, sie anhand struktureller Charakteristika prominenten Stoffklassen zuordnen und charakteristische Reaktionsmechanismen formulieren können. Gängige Praxis ist, dass Studierende und Lehrende Moleküle im Rahmen der Übung und Prüfung handschriftlich und analog darstellen. Existierende Programme zur digitalen Darstellung von Molekülen sind auf die Nutzung durch Expert:innen (z. B. für die Darstellung von Synthesen oder Syntheseprodukten im Rahmen von Publikationen) und nicht für die Nutzung im Rahmen von Lehre und Prüfungen ausgelegt, da (1) sie in der Regel lizenzpflichtig sind, (2) eine Integration in Lernplattformen wie Moodle nicht vorgesehen ist und (3) sie keine Möglichkeit zur automatischen Auswertung von studentischen Lösungen sowie keinen Raum für Korrekturen und Rückmeldungen bieten. Im besten Fall können mithilfe dieser Programme Bilder erzeugt werden, die von Lehrenden händisch ausgewertet werden. Ziel war daher, ein Tool zu entwickeln, (1) dass eine digitale Eingabe von typischen Aufgaben der organischen Chemie lizenzfrei ermöglicht, (2) sich in bestehende Lernplattformen wie Moodle integrieren lässt und (3) gegenüber bisherigen Lösungen den Mehrwert einer automatisierten Auswertung bietet. Im Rahmen des Beitrags werden die bisherigen Meilensteine der Entwicklung dargestellt, bestehende Schwierigkeiten beschrieben, Lösungsansätze diskutiert und ein Ausblick auf ausstehende Innovationen gegeben.

1 Ausgangslage

Übliche Praxis in der Lehre der organischen Chemie ist, dass Studierende zuhause Aufgaben bearbeiten und diese in der Präsenzübung von den Studierenden oder Lehrenden vorgestellt werden. Die Aufgaben werden dabei meist handschriftlich gelöst (Bleistift & Papier oder Kreide & Tafel). Zentrale Funktionen der Präsenzveranstaltung sind die Präsentation der Musterlösung, der Abgleich von Studierendenlösung und Musterlösung sowie das Klären von Fragen. Die Prüfung der Einführungsveranstaltung erfolgt meist in Form einer Klausur, in deren Rahmen

die Studierenden handschriftlich Aufgaben bearbeiten, die von Lehrenden kontrolliert und bewertet werden. Die Klausur zur Organischen Chemie 1 wird am Standort Essen in der Regel von 30 bis 50 Studierenden geschrieben. Prüfungen zu weiterführenden Veranstaltungen der organischen Chemie erfolgen zum Teil auch als mündliche Einzelprüfungen.

2 Entwicklung einer ersten Eingabemöglichkeit für organische Moleküle

Die Einrichtung einer digitalen Eingabemöglichkeit für organische Moleküle ist eine zentrale Voraussetzung für die Entwicklung einer digitalen Lern- und Prüfungsumgebung. Eine Anforderung hierbei ist, dass die Moleküle vom System als solche erkannt werden, sodass eine automatische Überprüfung der Eingaben der Studierenden möglich ist und nicht lediglich Bilddateien erzeugt werden, die von Lehrenden ausgewertet werden müssen.

2.1 Technische Umsetzung

Um eine den Anforderungen entsprechende Eingabemöglichkeit zu schaffen, wurde der javaskriptbasierte Moleküleditor Kekule.js (Jiang et al., 2016), der unter MIT-Lizenz verfügbar ist, in das E-Learning- und E-Assessment-Tool JACK (Striwe, 2016) integriert. Lehrende haben so die Möglichkeit, über den Aufgabentyp „Molekül“ Aufgaben anzulegen, die das Zeichnen von Molekülen erfordern (z. B. „Zeichnen Sie ein (*S*)-Pentan-2-ol-Molekül“). Für diese Aufgaben kann die gewünschte Lösung hinterlegt werden. Hierfür öffnet sich Kekule.js, so dass das gewünschte Molekül gezeichnet werden kann. Das Molekül wird in einen International Chemical Identifier (InChI) und somit in eine standardisierte Zeichenkette übersetzt, in der auch stereochemisch relevante Informationen abgebildet werden (für (*S*)-Pentan-2-ol beispielsweise die (*S*)-Konfiguration). Wenn Studierende diese Aufgabe bearbeiten, erfolgt ihre Eingabe ebenfalls über den Moleküleditor Kekule.js (siehe Abbildung 1). Auch die Studierendeneingabe wird als standardisierte Zeichenkette, in Form des InChI-Codes (Heller et al., 2015), gespeichert und dieser wird von JACK mit dem hinterlegten InChI-Code abgeglichen, sodass Studierendeneingaben automatisch als richtig oder falsch bewertet werden können. Falls Studierende beispielsweise (*R*)-Pentan-2-ol zeichnen oder in ihrer Darstellung keine Informationen zur dreidimensionalen Anordnung der Atome im Molekül (Stereoinformation) geben, unterscheiden sich der InChI-Code der Musterlösung und der Studierendeneingabe und die Eingabe wird als falsch bewertet. Ob Studierende die gleiche Schreibrichtung für ihr Molekül wählen wie in der Musterlösung, ist für die Auswertung unerheblich.

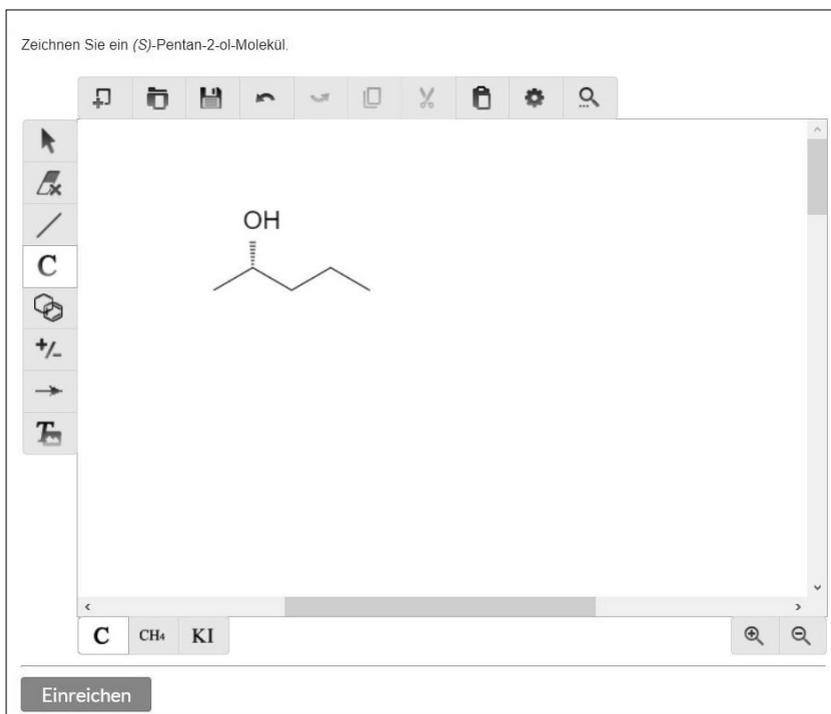


Abbildung 1: Beispiel für eine Molekül-Aufgabe in JACK (Quelle: Eigene Abbildung)

2.2 Erster Praxistest

Das entwickelte Tool wurde am Ende des Sommersemesters 2022 mit Studierenden der Studiengänge B. Sc. Chemie und Water Science ($N = 22$), die die Veranstaltung Organische Chemie 1 bereits besucht haben, getestet. Die Studierenden erhielten hierzu eine kurze Einführung in das Tool mit einer Übungsaufgabe. Im Anschluss bearbeiteten sie zunächst acht papierbasierte Molekül-Zeichen-Aufgaben (z. B. „Zeichnen Sie ein Propanol-Molekül“) und danach acht digitale Molekül-Zeichen-Aufgaben (z. B. „Zeichnen Sie ein Ethanol-Molekül“), die Grundlagen der IUPAC-Nomenklatur abprüfen.

Von den maximal möglichen 176 Antworten (8 Aufgaben x 22 Studierende) zu den analogen Aufgaben waren 104 Antworten der Studierenden richtig (59,09 %), bei den digitalen Aufgaben waren es lediglich 72 Antworten (40,91 %). Die deskriptiv höhere Lösungswahrscheinlichkeit für die analogen Aufgaben deutet darauf hin, dass die digitale Aufgabenbearbeitung neben der fachlichen Schwie-

rigkeit zusätzliche technische Herausforderungen umfasst. Rückmeldungen sowie Fragen und Anmerkungen der Studierenden während der digitalen Aufgabebearbeitung deuten ebenfalls in diese Richtung.

2.3 Erstes Zwischenfazit

Dieser erste Entwicklungsschritt zeigt, dass Kekule.js und JACK geeignet sind, um grundlegende Aufgaben der organischen Chemie umzusetzen. Die Evaluation mit Studierenden zeigt darüber hinaus, dass die digitale Eingabe für Studierende, die das papierbasierte Zeichnen von Molekülen gewohnt sind, nicht trivial ist und eine kurze Einführung in das Tool nicht ausreichend zu sein scheint. Mit Blick auf Prüfungen scheint es daher ratsam, die Studierenden bereits im Rahmen der Lehrveranstaltung an das digitale Format zu gewöhnen, damit das Prüfungsergebnis nicht durch den Formatwechsel beeinflusst wird.

3 Weiterentwicklung der automatisierten Auswertung

Basierend auf den Erkenntnissen aus dem Sommersemester wurde das Tool im Wintersemester 2022/23 systematisch in die Lehre der organischen Chemie eingebunden, um die Studierenden an das digitale Zeichnen der Moleküle zu gewöhnen. Gleichzeitig wurden die Möglichkeiten zur automatischen Auswertung der Studierendeneingaben erweitert.

3.1 Teillösungen, typische Fehler und alternative Lösungen

Um neben der erwarteten richtigen Lösung auch alternative Lösungen, Teillösungen und typische falsche Antworten hinterlegen zu können, wurde das bestehende System um Evaluator-Regeln erweitert, die diese Fälle abprüfen können.

Über die Funktion „sizeOfList“ kann beispielsweise geprüft werden, ob Studierende überhaupt ein Molekül ($\text{sizeOfList}([\text{input}=\text{inchi}])=0$), zu viele Moleküle (>1) oder zu wenige Moleküle (<1) gezeichnet haben. Beide Funktionen sind hilfreich, um grundlegende Fehler auszuschließen. Beispielsweise ziehen Anfänger:innen häufig versehentlich einen Strich, der vom System als Ethanmolekül interpretiert wird. Zusätzlich geben Anfänger:innen Reagenzien zum Teil über die Textfeld-Funktion ein, sodass diese nicht als Molekül erkannt und daher nicht ausgewertet werden können.

Mithilfe der Funktion „isElementOf“ kann geprüft werden, ob ein bestimmtes Molekül Teil der Lösung ist, sodass für dieses dann Teilpunkte vergeben werden können oder den Studierenden rückgemeldet werden kann, dass es sich hierbei um einen typischen Fehler handelt. Im Beispiel des (*S*)-Pentan-2-ols können so

(*R*)-Pentan-2-ol und Pentan-2-ol als erwartete Studierendeneingaben hinterlegt werden und den Studierenden im Rahmen der Übungsaufgaben rückgemeldet werden, dass sie das Enantiomer gezeichnet haben oder ihre Eingabe stereochemisch nicht eindeutig ist. Die Studierenden können so sofortiges fehlerspezifisches Feedback erhalten und in ihrem Lernprozess unterstützt werden (Johnson/Priest, 2014).

Über und- (&&), oder- (||) und nicht-Befehle (!) können darüber hinaus komplexere Evaluatorausdrücke erstellt werden, die eine Teilbepunktung an die Erfüllung mehrere Bedingungen knüpfen.

3.2 Integration der digitalen Aufgaben in die Übung

Um die Studierenden systematisch an digitale Aufgaben in der organischen Chemie zu gewöhnen, wurde die erste Sitzung der Übung Organische Chemie für eine ausführliche Einführung mit mehreren Übungsaufgaben und Raum für Fragen genutzt. Über das Semester erhielten die Studierenden im Rahmen der wöchentlichen Übungsaufgaben, neben den analogen Übungsaufgaben, auch immer einige digitale Übungsaufgaben. Nach drei Wochen wurden auch die Studierenden im Wintersemester ($N = 22$, Lehramt Chemie B. Sc.) gebeten, zunächst 17 analoge und anschließend 17 digitale Molekül-Zeichen-Aufgaben zu bearbeiten. Von insgesamt 374 Antworten (22 Studierende \times 17 Aufgaben) wurden analog 111 (29,68 %) und digital 108 (28,88 %) richtige Antworten gegeben. Es zeigen sich also keine deutlichen Unterschiede in der Lösungswahrscheinlichkeit zwischen digitalen und analogen Aufgaben.

3.3 Zweites Zwischenfazit

Die Evaluator-Funktionen ermöglichen eine differenzierte Auswertung komplexerer Aufgaben. Für den Übungsbetrieb ermöglichen sie darüber hinaus differenziertes fehlerspezifisches Feedback an die Studierenden. Die Erfahrungen in der Übung und das Ergebnis der Evaluation deutet darauf hin, dass die ausführlichere Einführung und die systematische Integration der digitalen Aufgaben in die Übung sinnvoll war und die Ergebnisse der Studierenden bei den digitalen Aufgaben nicht mehr durch technische Schwierigkeiten verzerrt wurden. Das insgesamt schlechtere Abschneiden der Studierenden verglichen mit dem Sommersemester ist vermutlich mit dem deutlich früheren Termin im Semester zu erklären. Eine wichtige Rückmeldung der Studierenden im Wintersemester war, dass sie den Mix aus digitalen und analogen Übungsaufgaben für ihre Selbstorganisation und Dokumentation ungünstig fanden.

4 Umsetzung einer digitalen Probeklausur

Da die Evaluator-Regeln die Umsetzung einer Vielzahl von Aufgaben erlaubt, wurden im Sommersemester die Übungsaufgaben komplett auf digitale Übungsaufgaben umgestellt und die Realisierung einer digitalen Probeklausur am Ende des Semesters angestrebt. Hierzu wurde nicht nur der Aufgabentyp „Molekül“ in JACK verwendet, sondern ergänzend auch Multiple-Choice- Aufgaben (z. B. „Geben Sie an, ob das Molekül chiral oder achiral ist“) konstruiert. Reaktionsmechanismen wurden für die Übung als Aufgaben mit mehreren Teilaufgaben angelegt. Den Studierenden wird die übergeordnete Aufgabenstellung angezeigt („Formulieren Sie den Reaktionsmechanismus der Reaktion von (*S*)-Pentan-2-ol mit Bromwasserstoff“) und zusätzlich die Teilaufgabenstellung für die aktuelle Stufe, für die eine Eingabe in Kekule.js erwartet wird (Stufe 1: „Zeichnen Sie die Edukte der Reaktion“). Nach der Eingabe öffnet sich dann die nächste Aufgabenstufe mit neuer Teilaufgabenstellung und neuem Editorfenster (Stufe 2: „Zeichnen Sie die Produkte des ersten Reaktionsschritts“). Die vorherige Teilaufgabe, ihre Eingabe und (im Rahmen der Übung) das Feedback hierzu können die Studierenden während der Bearbeitung folgender Teilaufgaben weiterhin einsehen, indem sie auf ihrem Bildschirm nach oben scrollen. Studierende können so auch Reaktionsmechanismen, die über mehrere Schritte verlaufen, bearbeiten (Stufe 3: „Zeichnen Sie die Produkte des zweiten Reaktionsschritts“, Stufe 4: „Zeichnen Sie die Produkte der Reaktion“). Zu den Übungsaufgaben wurde umfangreiches Feedback angelegt, sodass die Studierenden selbstständig lernen konnten. Aufgaben, die sich digital nicht sinnvoll umsetzen ließen (z. B. Benennen von Molekülen, Zeichnen von Reaktionsenergiediagrammen, Sesselkonformationen oder mesomeren Grenzformeln), wurden im Rahmen der Präsenzübung analog bearbeitet.

4.1 Aufgabenspezifische Anpassung des Editors und Versuchszähler

Um zu vermeiden, dass Studierende sich in der Übung „einfach durchklicken“ ohne Moleküle zu zeichnen oder Lösungen auszuprobieren, wurde die Variable „maxAttempts“ als Versuchszähler integriert (`[meta=stageCurrentAttempt]=[var=maxAttempts]`), der eine definierte Anzahl Lösungsversuche (in der Regel 3) erlaubt, bevor eine ausführliche Musterlösung (Renkl, 2021) präsentiert wird. Eine weitere zentrale technische Weiterentwicklung im Sommersemester war die Ermöglichung einer aufgabenspezifischen Konfiguration des Editors, sodass unnötige Templates (z. B. die Texteingabefunktion) und Funktionen, die die Studierenden verwirren könnten (Elektronen mit Angabe der Spinorientierung), nicht angezeigt werden. Da diese Funktion erst im laufenden Semester hinzukam, konnte sie nicht mehr für den Übungsbetrieb übernommen und getestet werden; dies ist für das Wintersemester 2023/24 geplant.

4.2 Replikation, Chiralitätsaufgaben und Erstellung der Probeklausur

Im Verlauf des Sommersemesters wurde die Evaluation der digitalen Aufgaben aus dem Wintersemester wiederholt, um die Ergebnisse zu replizieren. Hierfür haben 26 Studierende (Chemie B. Sc. und Water Science B. Sc.) erneut 14 Molekül-Zeichen-Aufgaben bearbeitet. In diesem Durchgang wurden zunächst die digitalen Aufgaben bearbeitet und anschließend die analogen Aufgaben. Zusätzlich waren die Studierenden darüber informiert, welche Kompetenzen relevant sein würden, sodass sie sich auf den Termin vorbereiten konnten. Von 364 Antworten (26 Studierende x 14 Aufgaben) waren im digitalen Format 187 (51,37 %) und im analogen Format 210 (57,69 %) Antworten richtig. Es zeigt sich also erneut eine etwas höhere Lösungswahrscheinlichkeit im analogen Format.

Zusätzlich wurden in diesem Semester einige Wochen später Aufgaben, bei denen es um Chiralität geht, zwischen den Formaten verglichen. Die Studierenden waren aufgefordert, chirale Moleküle stereochemisch eindeutig zu zeichnen („Zeichnen Sie folgendes Molekül“) und sollten im Multiple-Choice-Format entscheiden, ob oder über wie viele Stereozentren ein Molekül verfügt („Beurteilen Sie, ob die folgende Verbindung ein chirales Kohlenstoffatom aufweist“), ob Molekülpaare identische Moleküle zeigen oder Enantiomere („Beurteilen Sie, ob das folgende Molekülpaar identische Moleküle oder Enantiomere darstellt“) oder welche Konfiguration die Stereozentren eines Moleküls aufweisen („Ermitteln Sie die Konfiguration des folgenden Moleküls“). 19 Studierende bearbeiteten hierzu 19 analoge und 19 digitale Aufgaben. Von den insgesamt 361 Antworten (19 Studierende x 19 Aufgaben) waren im digitalen Format 194 (53,74 %) und im analogen Format 215 (59,56 %) Antworten richtig. Es zeigt sich also erneut eine etwas höhere Lösungswahrscheinlichkeit im analogen Format.

Die Entwicklung der Probeklausur erfolgte auf Basis einer bestehenden analogen Klausur. Im ersten Schritt wurden Aufgaben identifiziert, die problemlos in das digitale Format übertragen werden konnten. Im zweiten Schritt wurden Aufgaben betrachtet, die nicht ohne Weiteres in das digitale Format übertragen werden konnten. Dabei handelte es sich um Aufgaben mit Freitextantworten (z. B. „Benennen Sie folgendes Molekül“), eine Aufgabe, in der ein Reaktionsenergiediagramm gezeichnet werden sollte, sowie Aufgaben, die nach Reaktionsmechanismen fragen und bei deren Lösung die Elektronenpfeile oder Darstellungen in der Sesselkonformation bepunktet wurden.

Freitextaufgaben lassen sich über einen entsprechenden Aufgabentyp in JACK umsetzen, können dann aber nicht (zuverlässig/gut) automatisiert ausgewertet werden, weil bereits kleine Abweichungen (Groß-/Kleinschreibung, Zeichensetzung, Leerzeichen) dazu führen können, dass Antworten nicht erkannt werden. Für die Probeklausur wurden diese Aufgaben daher über den Aufgabentyp „Multiple-Choice“ umgesetzt. Als Distraktoren wurden überwiegend Studierendenantworten aus der ursprünglichen Klausur verwendet. Die Aufgabe, in der ein Reaktionsenergiediagramm gezeichnet werden sollte, wurde ebenfalls als

Multiple-Choice-Aufgabe mit falschen Studierendenantworten als Distraktoren umgesetzt. Da Kekule.js zwar das Zeichnen von Elektronenpfeilen ermöglicht, diese von JACK aber aktuell noch nicht aufgewertet werden können, musste die Bepunktung der Aufgaben, bei denen ursprünglich Elektronenpfeile bepunktet wurden, in der digitalen Version angepasst werden. Punkte werden hier lediglich für richtig gezeichnete Moleküle (Zwischenprodukte, Nebenprodukte und Produkte) vergeben. Da die Elektronenverschiebungen zentral für die Reaktionsmechanismen sind, stellt dies leider eine erhebliche Einschränkung im digitalen Format dar. Eine weitere Einschränkung besteht darin, dass JACK aktuell Reaktionspfeile nicht auswerten kann. Es kann bei Mechanismen also nur geprüft werden, welche Moleküle von Studierenden gezeichnet wurden, aber nicht, ob diese in der richtigen Reihenfolge gezeichnet wurden (Edukt, Zwischenprodukt, Produkt statt Edukt, Produkt, Zwischenprodukt oder ähnliches). Die Studierendenantworten der analogen Klausur deuten darauf hin, dass diese Einschränkung ein geringeres Problem darstellt als die fehlende Möglichkeit, Elektronenpfeile auszuwerten, da Reihenfolgefehler die Ausnahme darstellen und eher relevant werden, wenn ein alternativer Lösungsweg gewählt wurde.

4.3 Vorläufiges Fazit

Das Sommersemester 2023 hat gezeigt, dass der aktuelle Entwicklungsstand eine gute Basis bildet, mit der viele Aufgaben aus der Einführung in die organische Chemie, die die Darstellung von Molekülen umfassen, realisiert werden können, und dass diese Aufgaben bei einer guten Einführung in das digitale Tool und eine konsequente Nutzung im Rahmen der Übung vergleichbare Lösungswahrscheinlichkeiten haben wie entsprechende analoge Aufgaben. Der aktuelle Entwicklungsstand ist besonders im Rahmen der Übung für den Erwerb grundlegender Kompetenzen und das Einüben strukturierter (algorithmischer oder heuristischer) Lösungsschritte gut nutzbar. Die Erstellung der Probeklausur hat gezeigt, dass gerade bezüglich der digitalen Auswertung von Reaktionsmechanismen noch Entwicklungspotenzial besteht, um fachspezifische Charakteristika solcher Aufgaben wie im analogen Format bepunkten zu können. Auch die Rückmeldung der Studierenden gibt Hinweise darauf, dass eine Mehrheit der Studierenden aktuell eine analoge Klausur bevorzugt, weil das System noch nicht ausgereift genug für Prüfungssituationen ist und Studierende befürchten, in der Prüfung mit Eingabeschwierigkeiten oder technischen Hürden konfrontiert zu sein, die jenseits der zu prüfenden Inhalte Ressourcen beanspruchen.

5 Ausblick

Durch die Integration von Kekule.js in JACK wurde im Rahmen des PITCH-Projekts die Basis gelegt, um charakteristische Aufgaben der organischen Chemie digital umsetzen und automatisch auswerten zu können. Der aktuelle Entwicklungsstand ermöglicht bereits die Realisierung einfacher Molekül-Zeichen-Aufgaben sowie die strukturierter Reaktionsfolgen. Die Erstellung der Probeklausur hat gezeigt, dass das Tool noch nicht ausreichend entwickelt ist, um die fachspezifischen Anforderungen für Prüfungen der organischen Chemie vollständig erfüllen zu können. Für die Zukunft wäre es daher wichtig, das Tool diesbezüglich und mit Blick auf die Anforderungen, die weiterführende Veranstaltungen der organischen Chemie mit sich bringen, zu erfüllen.

Einen ersten Schritt in diese Richtung stellt die aktuelle Entwicklung einer Option zur Auswertung von Reaktionspfeilen durch das JACK-Team sowie die Ergänzung der Möglichkeit zur Darstellung von noch nicht vollständig gelösten, beziehungsweise ausgebildeten Bindungen, durch den Entwickler von Kekule.js dar. Darüber hinaus wäre es wünschenswert, wenn es zukünftig auch einen Aufgabentyp gäbe, bei dem Bereiche in Molekülen markiert werden könnten, um beispielsweise Stereozentren zu lokalisieren oder funktionelle Gruppen bestimmen zu lassen. Für komplexere Problemlöseaufgaben („Wie gelangen Sie ausgehend von Molekül X zu Molekül Y?“) ist es darüber hinaus notwendig, dass flexiblere Lösungswege und der Einsatz unterschiedlicher Reagenzien ermöglicht werden. Momentan ist es notwendig, dass in der Musterlösung alle Lösungswege, die bepunktet werden sollen, mit allen Molekülen vollständig identifiziert werden. Gerade im Fall von eingesetzten Reagenzien sind Teile des Moleküls aber häufig irrelevant (ihre konkrete Ausgestaltung also beliebig); so ist es für die Umsetzung von Cyclohexen zum Epoxid unerheblich, ob Perameisensäure, Persessigsäure oder *meta*-Chlorperoxybenzoesäure (*m*CPBA) verwendet wird. Für eine Reaktion irrelevante Molekülbereiche werden üblicherweise durch einen Buchstaben (z. B. R für Rest) dargestellt. Moleküle, die einen solchen Platzhalter enthalten, können gegenwärtig noch nicht in einen InChI-Code überführt und damit nicht ausgewertet werden. Wäre dies möglich, wäre es deutlich leichter, die Musterlösung für komplexere Aufgaben zu hinterlegen.

Auch wenn digitale Prüfungen in der organischen Chemie, angesichts der technischen Möglichkeiten, aktuell noch nicht zufriedenstellend umgesetzt werden können, sollten zwei Punkte nicht vergessen werden: Erstens kann auch die analoge Auswertung von Prüfungen in der organischen Chemie aufgrund der vielfältigen Möglichkeiten zur Darstellung von Molekülen Schwierigkeiten bereiten. Der digitale Vergleich zweier InChI-Codes ist weniger fehleranfällig, als es eine menschliche Korrektur sein kann. Wenn die digitale Umsetzung und automatische Auswertung von Aufgaben gelingen, ist diese also nicht nur hinsichtlich ihrer Effizienz, sondern auch mit Blick auf die Objektivität der analogen Prüfung zu überlegen. Zweitens ermöglicht das bestehende System bereits

zusätzliche Lern- und Feedbackgelegenheiten, weil es mehr individuelle und fehlerspezifische Rückmeldungen bietet, als es einzelne Studierende im Rahmen der regulären Lehrveranstaltung erhalten können. Gleichzeitig verändert ein solches Tool auch die Präsenzlehre, weil die Vorstellung der Musterlösung (gerade für simple Aufgaben) entfallen kann, sodass Raum für eine intensivere Auseinandersetzung, die über die Vorstellung der Musterlösung hinausgeht, geschaffen wird (van Alten et al., 2019); Studierende und Lehrende müssen daher lernen, die Präsenzveranstaltungen anders zu gestalten und zu nutzen.

Literatur

- Heller, S. R./McNaught, A./Pletnev, I./Stein, S./Tchekhovskoi, D. (2015). InChI, the IUPAC International Chemical Identifier. In: *Journal of cheminformatics*, 7(23). <https://doi.org/10.1186/s13321-015-0068-4>
- Jiang, C./Jin, X./Dong, Y./Chen, M. (2016). Kekule.js: An Open Source JavaScript Cheminformatics Toolkit. In: *Journal of chemical information and modeling*, 56(6), 1132–1138. <https://doi.org/10.1021/acs.jcim.6b00167>
- Johnson, C./Priest, A. H. (2014). The Feedback Principle in Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.
- Renkl, A. (2021). The Worked Example Principle in Multimedia Learning. In R. E. Mayer/L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 231–240). New York: Cambridge University Press. <https://doi.org/10.1017/9781108894333.023>
- Striewe, M. (2016). An architecture for modular grading and feedback generation for complex exercises. In: *Science of Computer Programming*, 129, 35–47. <https://doi.org/10.1016/j.scico.2016.02.009>
- van Alten, D. C./Phielix, C./Janssen, J./Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. In: *Educational Research Review*, 28, 1–18. <https://doi.org/10.1016/j.edurev.2019.05.003>

Generative KI und Prüfungen: Erste Erfahrungen und offene Fragen

Michael Striewe

Durch die Veröffentlichung von ChatGPT im November 2022 richtete sich plötzlich eine erhebliche öffentliche Aufmerksamkeit auf den Einsatz sogenannter generativer KI für die Erledigung von Alltagsaufgaben. Zwar gab es auch schon zuvor KI-basierte Werkzeuge, die z. B. Texte übersetzen oder Bilder generieren konnten, doch in der öffentlichen Wahrnehmung erhielten diese Werkzeuge ein deutlich geringeres Echo. Auch in der Hochschule spielten sie kaum eine Rolle und wurden wenig im Hinblick auf ihren Einfluss auf Prüfungen beleuchtet. Der Einfluss und das Potenzial künstlicher Intelligenz auf die Hochschulbildung wurde zwar in der Forschung regelmäßig diskutiert (Zawacki-Richter et al., 2019), der praktische Einfluss neuer Entwicklungen in diesem Bereich schien allerdings zunächst begrenzt zu sein. Die Verfügbarkeit von ChatGPT und wenig später auch einer Vielzahl weiterer, KI-basierter Systeme, insbesondere zur Textproduktion, löste jedoch eine umfangreiche Diskussion mit starken Emotionen zwischen Angst und Euphorie aus (Getto/Valdivia Rojas, 2023) und führte zu einer umfangreichen Debatte über Chancen und Herausforderungen für die Lehre (z. B. Kasneci et al., 2023; Rudolph et al., 2023).

Auch an der Universität Duisburg-Essen wurde im Projekt PITCH und darüber hinaus der Einfluss generativer KI auf die Gestaltung universitärer Prüfungen diskutiert. Die folgenden Abschnitte fassen nach einer kurzen technischen Einführung exemplarisch und ohne Anspruch auf Vollständigkeit einige Erfahrungen zusammen, die an der Universität Duisburg-Essen und an anderen Orten zum Einsatz generativer KI im Kontext von universitären Prüfungen gemacht wurden. Abschließend werden daraus Punkte für eine Diskussion und offene Fragen abgeleitet, die erst die zukünftigen Entwicklungen werden beantworten können. Für eine rechtliche Einordnung des Themas wird zudem auf den Beitrag von Henrik Noll verwiesen.

1 Technische Grundlagen

Die meisten generativen KI-Modelle basieren auf dem sogenannten maschinellen Lernen mithilfe künstlicher neuronaler Netze. Während des Lernprozesses verarbeitet das KI-Modell große Mengen an Eingabedaten (z. B. Texte im Fall

von Sprachmodellen oder Fotos im Fall von Bildmodellen) und erkennt darin Strukturen in Form von regelmäßig auftretenden Zusammenhängen. Das KI-Modell lernt dadurch, wie wahrscheinlich es ist, dass eine bestimmte Kombination von Elementen auftritt und kann dabei auch einen größeren Kontext berücksichtigen. Jedes KI-Modell verfolgt dabei in der Regel ein einziges Lernziel. Das Lernziel eines großen Sprachmodells (engl. large language model, LLM) wie dem von ChatGPT verwendeten GPT-Modell ist es, möglichst natürlich klingende Sätze und Texte produzieren zu können. Es lernt also, welche Worte typischerweise aufeinander folgen und wie Sätze miteinander verknüpft werden. Die inhaltliche Bedeutung einzelner Worte oder der logische Zusammenhang zwischen einzelnen Begriffen wird vom Modell dagegen nicht verstanden.

Beim Einsatz eines KI-Modells zur Produktion eines neuen Textes oder Bildes werden die gelernten Zusammenhänge genutzt, um ausgehend von einer initialen Eingabe (dem sogenannten Prompt) schrittweise ein wahrscheinliches nächstes Stück der Ausgabe zu prognostizieren. Als Kontext dient dabei zunächst nur die Eingabe an der Benutzerschnittstelle des verwendeten KI-Werkzeugs; im weiteren Verlauf der Antwortgenerierung dann auch alle bis dahin produzierten Ausgaben. Dadurch wirken die Ausgaben einer generativen KI auf den ersten Blick sehr stimmig und gelungen. Im Fall von ChatGPT entsteht beispielsweise der Eindruck eines echten Gesprächs, da die einzelnen Anfragen innerhalb eines Dialogs nicht isoliert stehen, sondern das Werkzeug auf zuvor geschriebenes zurückgreifen und es ergänzen oder sogar korrigieren kann. Da das KI-Modell wie oben beschrieben jedoch keine logischen Zusammenhänge gelernt hat, basieren alle Aussagen lediglich auf statistischen Wahrscheinlichkeiten (Pasquinelli, 2017; Jang/Lukasiewicz, 2023). Dies führt unter anderem dazu, dass derselbe Prompt nicht immer exakt dieselbe Antwort erhält, was im Extremfall dazu führen kann, dass eine Frage mal korrekt und mal falsch beantwortet wird. Zudem kommt der Formulierung des Prompts eine sehr zentrale Bedeutung zu, denn schon der Austausch eines einzigen Wortes kann das KI-Modell in Richtung eines anderen Kontexts und damit zu einer völlig anderen Antwort führen.

Die Nutzung eines KI-Modells erfolgt in der Praxis in der Regel nicht über einen direkten Zugriff, sondern über die Benutzerschnittstelle eines KI-basierten Werkzeugs. Auch ChatGPT ist ein solches Werkzeug und in der Lage, die Nutzereingabe vorzuverarbeiten oder die vom KI-Modell produzierte Ausgabe zu filtern, zum beispielsweise als unangemessen geltende Formulierungen zu entfernen. Im Falle der Suchmaschine Bing verarbeitet diese die Anfragen der Nutzenden zunächst als normale Internetsuche und nutzt die einschlägigsten Suchtreffer anschließend als Kontext, der bei der Erzeugung einer KI-generierten Antwort berücksichtigt wird. Auf diese Weise kann die KI stets aktuelle Informationen nutzen, auch wenn das KI-Modell zu einem Zeitpunkt trainiert wurde, zu dem diese Informationen noch gar nicht vorlagen.

2 Generative KI als Werkzeug zur Lösung von Prüfungsaufgaben

Der Einsatz generativer KI zur Erstellung von Prüfungsleistungen ist mutmaßlich das Szenario, das am häufigsten und kritischsten diskutiert wird und das in der Regel mit rechtlichen Fragestellungen nach Plagiaten oder anderen Formen von Täuschungsversuchen verknüpft wird (Cotton et al., 2023). Ein Teil dieser Diskussionen ist möglicherweise den Erfahrungen aus der Corona-Pandemie geschuldet, in der in großem Umfang unüberwachte Prüfungen von zu Hause ermöglicht wurden, deren Sinnhaftigkeit nun mit der Verfügbarkeit eines leistungsstarken und universell zur Textproduktion verwendbaren, technischen Hilfsmittels abrupt in Frage gestellt wurde. An zahlreichen Universitäten und anderen Bildungsinstitutionen wurden seit der Veröffentlichung von ChatGPT kleine und große Experimente durchgeführt um zu überprüfen, ob bestimmte Prüfungsaufgaben durch KI-gestützte Werkzeuge gelöst werden können. Das Ergebnis vieler solcher Studien lautet, dass ChatGPT in der Tat in der Lage ist, auch anspruchsvolle Prüfungen zu bestehen – zum Teil sogar mit beachtlichen Noten (z. B. Wood et al., 2023; Buchmann/Thor, 2023; Bhayana et al., 2023). Gleichzeitig kommen andere Studien zu dem Ergebnis, dass die Leistung von ChatGPT zumindest teilweise nicht zum Bestehen einer Prüfung reichen würde (z. B. Jalil et al., 2023) und dass Studierende durch die Nutzung von ChatGPT keinen positiven Effekt auf ihre Note erzielen können (Shoufan, 2023).

Eine pauschale Aussage über die Leistungsfähigkeit von ChatGPT oder anderen Werkzeugen als Hilfsmittel zur Lösung von Prüfungsaufgaben lassen die vorliegenden Studien somit nicht zu. Zudem kann jede neue Version eines LLMs und auch jede kleine, technische Änderung an einem KI-basierten Werkzeug dazu führen, dass für dieselben Aufgaben völlig andere Ergebnisse erzielt werden. Die Halbwertszeit vorliegender Erfahrungen ist dementsprechend kurz. Lehrende werden daher wohl auf absehbare Zeit stets eine sorgfältige Einzelfallprüfung vornehmen müssen, wenn sie Auskunft über die Lösbarkeit einer konkreten Klausuraufgabe für KI-basierte Werkzeuge erhalten wollen. Gleichwohl lassen sich aus den vorliegenden Studien in Verbindung mit dem technischen Aufbau generativer KI-Werkzeuge Erfahrungswerte ableiten, die auch ohne eine Einzelfallprüfung erlauben, die Lösbarkeit einer Prüfungsaufgabe durch KI-gestützte Werkzeuge abzuschätzen.

Überwiegend gut schneiden KI-gestützte Werkzeuge bei Aufgaben auf niedrigen bis mittleren Kompetenzniveaus ab, beispielsweise bei der Reproduktion bekannter Thesen und Konzepte, dem Vergleich bekannter Quellen oder der Erfindung einfacher Beispiele. Immer dann, wenn die zugrundeliegenden Fakten und Begriffe mit hoher Wahrscheinlichkeit in großer Anzahl in den Trainingsdaten eines LLMs enthalten waren und wenige Stichworte ausreichen, um den Kontext der Aufgabe und die Arbeitsanweisung treffend zu beschreiben, überwiegen die Stärken eines KI-basierten Werkzeugs bezüglich der Produktion eines inhaltlich konsistenten, gut formulierten Texts. Weniger gut schneiden KI-gestützte

Werkzeuge dagegen bei Aufgaben auf höheren Kompetenzniveaus ab, bei denen sehr spezifische Details genutzt werden müssen oder die Korrektheit einer Antwort stark vom logischen Zusammenhang einzelner Aussagen abhängt. In solchen Fällen sind möglicherweise nicht genügend Beispiele in den Trainingsdaten enthalten gewesen, um eine passende Antwort alleine auf Basis der statistischen Wahrscheinlichkeit formulieren zu können, zumal eine solche Herangehensweise logische Zusammenhänge nicht zuverlässig abbilden und somit schnell Widersprüche produzieren kann.

Bei geschlossenen Aufgabenformaten kommt eine generative KI teilweise zu schlechteren Ergebnissen als bei frei formulierten Texten (Mahat et al., 2023). Da ein Aufgabentyp wie Multiple-Choice keine frei formulierte Antwort zulässt, kann ein LLM seine Stärke hier nicht ausspielen. Stattdessen muss ein KI-basiertes Werkzeug die Antwortoptionen in Relation zur Fragestellung analysieren. Sind die falschen Antwortoptionen im gegebenen Kontext ähnlich wahrscheinlich wie die korrekte Antwort, bleibt dem KI-Modell nur der Rückgriff auf seine inhärente Fähigkeit zum Raten. Lückentextaufgaben liegen einem LLM zwar näher, da bei diesen im Wesentlichen aus dem Kontext auf ein fehlendes Wort geschlossen werden muss, aber auch hier ist keine freie Formulierung möglich, was das LLM in seiner Stärke bescheidet.

Als Gegenmaßnahme zum Einsatz generativer KI bei der Lösung von Freitextaufgaben in Prüfungen ist der Einsatz weiterer, KI-basierter Werkzeuge möglich, die KI-generierte Texte erkennen sollen. Die bisherigen Ergebnisse sind jedoch nicht überzeugend (Weber-Wulff et al., 2023) und ziehen wiederum andere Nachteile mit sich, wenn die dahinterliegenden Verfahren keine faire Einschätzung erzeugen (Liang et al., 2023).

3 Generative KI als Werkzeug zur Erstellung von Prüfungsaufgaben

Augenscheinlich deutlich weniger beachtet als der Einsatz von KI durch Studierende zur Erstellung von Prüfungsleistungen ist der Einsatz von KI durch Lehrende zur Erstellung von Prüfungsinhalten. Dies liegt möglicherweise darin begründet, dass an den Stellen, an denen die massenhafte Erstellung von Prüfungsfragen überhaupt ein relevantes Thema ist, die Nutzung automatisierter Verfahren zur Erstellung von Prüfungsfragen bereits ein etabliertes Vorgehen ist. Dies erstreckt sich von der einfachen Erzeugung von Zufallszahlen für Mathematikaufgaben bis hin zur Erzeugung fachlich anspruchsvoller Multiple-Choice-Fragen in der Medizin. Dafür liegen sowohl technische Konzepte und Werkzeuge als auch umfangreiche theoretische – vornehmlich psychometrische – Grundlagen zur Bewertung der Qualität der erzeugten Prüfungsfragen vor (Gierl/Haladyna, 2012). Dementsprechend konnten schon vor der Verfügbarkeit großer KI-Modelle qualitativ hochwertige Ergebnisse erzielt werden (Gierl/Lai, 2013). Auch der spezifische Einsatz von automatischer Textgenerierung (engl. Natural

Language Generation, NLG), für die auch die LLMs ausgelegt sind, wurde schon vor 20 Jahren beispielsweise zur Erzeugung von mathematischen Textaufgaben erprobt (Deane/Sheehan, 2003). Die Verfügbarkeit von KI-Modellen ermöglichte in diesem Bereich daher nicht erstmalig eine Automatisierung, sondern erweiterte lediglich das Portfolio an möglichen Techniken und hatte folglich auch ein viel geringeres, disruptives Potenzial. Außerdem unterscheidet sich das Einsatzszenario grundlegend: Die KI kommt bei der Erstellung von Prüfungen primär als unterstützendes Werkzeug zum Einsatz, das Lehrenden Vorschläge in Form eines Aufgabenpools vorlegt, aus denen diese dann die tatsächlichen Inhalte für eine Prüfung auswählen bzw. aus dem sie ungeeignete Vorschläge entfernen können. Die fachliche Leistung, eine potenzielle Prüfungsfrage auf ihre inhaltliche Korrektheit sowie die Angemessenheit für eine Lerngruppe zu prüfen, verbleibt damit bei den Lehrenden.

Schon vor der Veröffentlichung von ChatGPT, das zunächst auf Version 3.5 und ab März 2023 auf Version 4.0 des GPT-Modells beruhte, wurden im Hochschulkontext Experimente mit der Vorgängerversion GPT 2 durchgeführt, um beispielsweise automatisiert Verständnisfragen zu Texten zu generieren (Kasakowskij et al., 2022). Aus technischer Sicht ist der Einsatz in diesem Bereich sehr naheliegend und vielversprechend, da es zu den Stärken eines LLMs gehört, längere Texte zusammenzufassen, sowie Aussagesätze in Fragesätze zu verwandeln. Erfahrungen mit verschiedenen KI-basierten Fragengeneratoren derselben Art zeigen jedoch, dass es den Werkzeugen nicht immer gelingt, die zentralen Aussagen eines Textes tatsächlich zu identifizieren, sodass generierte Fragen möglicherweise auf Nebensächlichkeiten und Randbemerkungen abzielen. Die Kernaussage eines Textes, die sich möglicherweise erst aus der Zusammenschau mehrerer Sätze ergibt, ist für die Werkzeuge deutlich schwieriger zu identifizieren und damit auch deutlich schwieriger in eine sinnvolle Frage umzuwandeln. Auch GPT 3.5 wurde inzwischen intensiv zur automatischen Generierung von Aufgabeninhalten beispielsweise für Sprachprüfungen erprobt und hat sich dabei als leistungsfähig, aber nicht fehlerfrei erwiesen (Rüdian/Pinkwart, 2023). Auch hier ist die Leistungsfähigkeit aus technischer Sicht wenig überraschend, da es im Kontext von Aufgaben in Sprachprüfungen vor allem auf die sprachliche Korrektheit und die inhaltliche Konsistenz eines Textes ankommt, während die thematische Komplexität eines Textes gering bleiben darf und somit keine besonderen Anforderungen an einen KI-basierten Generator stellt. Gleichwohl zeigen die bisherigen Ergebnisse, dass auch vermeintlich einfache Sätze nicht immer fehlerfrei generiert werden und ein unüberwachter Einsatz daher nicht zu empfehlen ist.

Noch keine einschlägigen Erfahrungen liegen zur KI-basierten Generierung von Bildern für den Prüfungskontext vor. Während beispielsweise der Einsatz von KI zur Erzeugung von Gemälden im Stil eines bestimmten Künstlers aus kunsthistorischer Sicht kontrovers diskutiert wird (Hassine/Neeman, 2019), liegen zum didaktischen Potenzial einer solchen Erzeugung und zur Nutzbarkeit

solcher Bilder in Prüfungen, beispielsweise als Antwortoptionen in Multiple-Choice-Fragen, bisher keine Studien vor.

4 Generative KI als Werkzeug im Korrektur- und Bewertungsprozess

Auch als unterstützendes Werkzeug für Lehrende bei der Korrektur und Bewertung von Prüfungsleistungen hat die KI bisher wenig breite Resonanz erfahren. Dies liegt möglicherweise ebenfalls darin begründet, dass entsprechende Werkzeuge in der Forschung schon seit einiger Zeit erprobt wurden (z. B. Gao, 2021). Auch unabhängig von aktuellen LLMs wurde maschinelles Lernen bereits erfolgreich genutzt, um automatisch Feedback-Vorschläge aufgrund von Ähnlichkeiten zu zuvor manuell bewerteten Einreichungen zu erzeugen (Bernius et al., 2021). Aus technischer Sicht ist ein solches Vorgehen naheliegend, da ähnliche Einreichungen in der Regel auch ähnliche Bewertungen erhalten sollten. Dazu müssen insbesondere die entscheidenden, möglicherweise subtilen Unterschiede zwischen korrekten und inkorrekten Lösungen identifiziert werden können. Entscheidende Bedeutung kommt hier also weniger einem vortrainierten KI-Modell zu, als vielmehr der sinnvollen Gruppierung von Einreichungen im Lösungsraum, zu der im Kontext automatisierter Tutorensysteme für den Übungsbetrieb schon seit vielen Jahren geforscht wird (z. B. Gross et al., 2012). Maschinelle Lernverfahren sind zwar grundsätzlich geeignet, Strukturen im Lösungsraum einer Aufgabe zu erkennen und so sinnvolle Gruppierungen zu berechnen, müssen dazu jedoch auf großen Mengen an Abgaben für jede einzelne Aufgabe separat trainiert werden, was einen erheblichen Aufwand bedeutet, der sich möglicherweise nur in sehr wenigen Fällen lohnt. Allgemeine LLMs sind für diese Aufgabe dagegen kaum geeignet.

Generative KI kann auch genutzt werden, um Feedback zu einer Einreichung zu geben. Dazu werden die studentische Einreichung und eine Menge von Bewertungskriterien als Prompt in einem KI-basierten Werkzeug eingegeben. Dem Prompt kommt dabei eine erhebliche Bedeutung zu, denn es macht einen Unterschied, ob lediglich korrekte von falschen Antworten unterschieden werden sollen oder ob auch Fehler erläutert werden sollen. ChatGPT konnte in einzelnen Studien bisher zwar durchaus zutreffende Einschätzungen zur Korrektheit abgeben, allerdings teilweise nur wenig Übereinstimmung mit manuell erstelltem, detailliertem Feedback erreichen oder hat gar Fehler erläutert, die gar nicht gemacht wurden (z. B. Dai et al., 2023; Azaiz et al., 2023).

5 Diskussion und offene Fragen

Aus den bisherigen Erfahrungen lassen sich sowohl allgemeine als auch konkrete Diskussionspunkte und offene Fragen ableiten. Allgemein darf davon ausgegangen werden, dass der Einsatz generativer KI in der näheren Zukunft eher zunehmen als abnehmen wird und der Zugriff auf derartige Werkzeuge noch einfacher wird. Dies dürfte insbesondere dazu führen, dass es in Zukunft leichter wird, Texte zu produzieren, die in enger Kooperation zwischen Mensch und Maschine entstanden sind. Es ist zu diskutieren, inwiefern es möglich und didaktisch sinnvoll ist, der Nutzung derartiger Werkzeuge bei der elektronischen Durchführung von Prüfungen gänzlich zu verbieten, um sicherzustellen, dass Prüfungsleistungen auch tatsächlich von den jeweiligen Studierenden persönlich erstellt worden sind. Auch eine entsprechende Kennzeichnungspflicht wird an Hochschulen diskutiert, könnte bei der potenziellen, allgemeinen Verfügbarkeit von KI-Assistenz in jedem Textverarbeitungsprogramm aber schnell zu einem nichtssagenden Allgemeinplatz werden. Umgekehrt ist jedoch auch nicht davon auszugehen, dass bei einer kooperativen Textproduktion der Einfluss KI-basierter Werkzeuge mit automatischen Verfahren zuverlässig detektierbar ist. Benötigt wird daher eine kritische Auseinandersetzung, welcher Einsatz KI-basierter Werkzeuge bei der Lösung von Prüfungsaufgaben aus fachlicher und didaktischer Sicht unproblematisch ist, und welcher Einsatz durch geeignete Maßnahmen unterbunden werden muss. Dies schließt auch eine Diskussion darüber ein, welche Art von Prüfungsaufgaben überhaupt geeignet ist, die in der Lehre vermittelten Kompetenzen zuverlässig abzurufen. Zahlreiche Hochschulen arbeiten dazu derzeit bereits mit entsprechenden Handreichungen zur Orientierung für die Lehrenden, sodass in der näheren Zukunft weitere Erfahrungen zur Umsetzbarkeit einzelner Maßnahmen zu erwarten sind. An der Universität Duisburg-Essen wurden beispielsweise Empfehlungen zum Umgang mit künstlicher Intelligenz in Studium und Lehre an der UDE veröffentlicht.

Auf der wesentlich konkreteren Ebene stellt sich die Frage, für welche Aufgaben im Rahmen des Prüfungsprozesses genau generative KI tatsächlich hilfreich ist und wo sie schlechter abschneidet als ein Mensch. Während einem unvorbereiteten Studierenden möglicherweise bereits eine durchschnittlich leistungsfähige KI eine Hilfe bei der Lösung einer Prüfungsaufgabe ist, werden sich anspruchsvolle Lehrende kaum auf die Unterstützung durch eine mittelmäßige KI bei der Erstellung von Aufgaben oder Bewertungen verlassen wollen. Andererseits darf durchaus nicht vergessen werden, dass beispielsweise auch von Menschen durchgeführte Bewertungen auf großen Mengen von Prüfungsleistungen keineswegs immer konsistent sind. Es bleibt daher eine offene Frage für die Forschung, ob und wann KI-Werkzeuge eine vergleichbare Fehlerrate erreichen werden.

Neben den fachlichen und didaktischen Fragen bleiben auch in der technischen Entwicklung Fragen offen. Eine Frage ist, wie schnell sich generative KI-Werkzeuge weiterentwickeln und dann auch andere Ausgaben als Texte und

Bilder in angemessener Qualität produzieren können. Insbesondere zu Diagrammen und Zeichnungen aus technischen und ingenieurwissenschaftlichen Fächern liegen derzeit noch nicht genügend Erfahrungen vor. Die zurückliegenden Entwicklungen der KI lassen allerdings erwarten, dass auch in diesem Bereich eine weit fortgeschrittene Generierungsqualität erreicht werden wird. Daran schließt sich unmittelbar wieder die fachliche Frage an, welche Rolle KI-basierte Werkzeuge in der Praxis des jeweiligen Faches in Zukunft erhalten werden und ob es dann nicht sogar erforderlich ist, die Kompetenzen der Studierenden im Umgang mit generativer KI zur Lösung fachlicher Aufgabenstellungen in einer Prüfung abzu prüfen.

Ebenfalls offen ist, wie sich das Potenzial generativer KI zur Erstellung individueller, auf die Fähigkeiten, Bedürfnisse und Lernkontexte einzelner Studierende zugeschnittener Prüfungsaufgaben samt der automatischen Erzeugung passenden Feedbacks tatsächlich nutzen lässt. Für diesen Zweck wird es unvermeidbar sein, Informationen über individuelle Studierende als Teil des Prompts in ein KI-basiertes Werkzeug einzugeben. Aus datenschutzrechtlichen Gründen ist kaum vorstellbar, dass eine solche Generierung mit einem allgemein verfügbaren, auf Servern außerhalb einer Hochschule betriebenen System durchgeführt werden kann. Es wird daher notwendig sein, spezialisierte Werkzeuge beziehungsweise Funktionen innerhalb elektronischer Prüfungssysteme zu entwickeln, die von den Hochschulen selbst betrieben werden können. Da maschinelles Lernen erst bei einem großen Volumen an Lerndaten zu zuverlässigen Ergebnissen führt, stellt sich hier für die Forschung zudem die Frage, wie eine qualitativ hochwertige Unterstützung auch für kleine Zielgruppen gewährleistet werden kann. Gerade wenn es das Ziel sein soll, durch den Einsatz generativer KI individuelle Unterstützung zu leisten und die Bildungsgerechtigkeit zu erhöhen, sind solche Fragen von zentraler Bedeutung.

Literatur

- Azaiz, I./Deckarm, O./Strickroth, S. (2023). AI-Enhanced Auto-Correction of Programming Exercises: How Effective is GPT-3.5?. In: *International Journal of Engineering Pedagogy (IJEP)*, 13(8), 67–83. <https://doi.org/10.3991/ijep.v13i8.45621>
- Bernius, J. P./Krusche, S./Bruegge, B. (2021). A Machine Learning Approach for Suggesting Feedback in Textual Exercises in Large Courses. In: *Proceedings of the Eighth ACM Conference on Learning @ Scale (L@S '21)*. ACM, New York, NY, USA, 173–182. <https://doi.org/10.1145/3430895.3460135>
- Bhayana, R./Krishna, S./Bleakney, R. R. (2023). Performance of ChatGPT on a Radiology Board-style Examination: Insights into Current Strengths and Limitations. *Radiology*, 307(5), 1–8. <https://doi.org/10.1148/radiol.230582>
- Buchmann, E./Thor, A. (2023). Online Exams in the Era of ChatGPT. 21. Fachtagung Bildungstechnologien (DELFI), S. 79–84. <https://doi.org/10.18420/delfi2023-15>

- Cotton, D. R. E./Cotton, P. A./Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. In: *Innovations in Education and Teaching International*, 31(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Dai, W./Lin, J./Jin, H./Tongguang, L./Tsai, Y.-S./ Gašević, D. (2023). Can Large Language Models Provide Feedback to Students? A Case Study on ChatGPT. In: *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)*, Orem, UT, USA, 323–325. <https://doi.org/10.1109/ICALT58122.2023.00100>
- Deane, P./Sheehan, K. (2003). Automatic Item Generation via Frame Semantics: Natural Language Generation of Math Word Problems. <https://eric.ed.gov/?id=ED480135>
- Gao, J. (2021). Exploring the feedback quality of an automated writing evaluation system Pigai. In: *International Journal of Emerging Technologies in Learning (IJET)*, 16(11), 322–330. <https://doi.org/10.3991/ijet.v16i11.19657>
- Getto, B./Valdivia Rojas, S. (2023). ChatGPT und KI im Bildungswesen – Disruption, Revolution oder hatten wir alles schon? <https://blog.phzh.ch/zhe/chatgpt-und-ki-im-bildungswesen-disruption-revolution-oder-hatten-wir-alles-schon/> (zuletzt abgerufen am 20.09.2023)
- Gierl, M. J./Haladyna, T. M. (Hrsg.) (2012). *Automatic Item Generation*. New York: Routledge.
- Gierl, M. J./Lai, H. (2013). Evaluating the quality of medical multiple-choice items created with automated processes. In: *Medical Education*, 47, 726–733. <https://doi.org/10.1111/medu.12202>
- Gross, S./Mokbel, B./Hammer, B./Pinkwart, N. (2012). Feedback Provision Strategies in Intelligent Tutoring Systems Based on Clustered Solution Spaces. In Desel, J./Haake, J. M./Spannagel, C./Gesellschaft für Informatik (Hrsg.), *DeLFI 2012: Die 10. e-Learning Fachtagung Informatik* (S. 27–38) Hagen: Köllen.
- Hassine, T./Neeman, Z. (2019). The Zombification of art history: how AI resurrects dead masters, and perpetuates historical biases. In: *Journal of Science and Technology of the Arts*, 11(2), 28–35. <https://doi.org/10.7559/citarj.v11i2.663>
- Jalil, S./Rafi, S./LaToza, T./Moran, K./Lam, W. (2023). ChatGPT and Software Testing Education: Promises & Perils. In: *2023 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, 4130–4137. <https://doi.org/10.1109/ICSTW58534.2023.00078>
- Jang, M./Lukasiewicz, T. (2023). Consistency Analysis of ChatGPT. In: *The 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)*, Preprint arXiv:2303.06273. <https://doi.org/10.48550/arXiv.2303.06273>
- Kasakowskij, R./Kasakowskij, T./Seidel, N. (2022). Generation of Multiple True False Questions. In: *20. Fachtagung Bildungstechnologien (DELFI)*, 147–152. <https://doi.org/10.18420/delfi2022-026>
- Kasneci, E. /Sessler, K./Küchemann, S./Bannert, M./Dementieva, D./Fischer, F./Gasser, U./Groh, G./Günemann, S./Hüllermeier, E./Krusche, S./Kutyniok, G./Michaeli, T./Nerdel, C./Pfeffer, J./Poquet, O./Sailer, M./Schmidt, A./Seidel, T./.../Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. In: *Learning and Individual Differences*, 103. <https://doi.org/10.1016/j.lindif.2023.102274>
- Liang, W./Yuksekonul, M./Mao, Y./Wu, E./Zou, J. (2023). GPT detectors are biased against non-native English writers. In: *Patterns* 4(7). <https://doi.org/10.1016/j.patter.2023.100779>

- Mahat, R. K./Jantikar, A. M./Rathore, V./Panda, S. (2023). Assessing the performance of ChatGPT to solve biochemistry question papers of university examination. In: *Advances in Physiology Education*, 47, 528–529. <https://doi.org/10.1152/advan.00076.2023>
- Parasinelli, M. (2017). *Machines that Morph Logic: Neural Networks and the Distorted Automation of Intelligence as Statistical Inference*. <https://www.glass-bead.org/article/machines-that-morph-logic/?lang=enview> (zuletzt abgerufen am 16.08.2023)
- Rudolph, J./Tan, S./Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? In: *Journal of Applied Learning and Teaching*, 6(1), 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Rüdian, S./Pinkwart, N. (2023). Auto-generated language learning online courses using generative AI models like ChatGPT. In: 21. Fachtagung Bildungstechnologien (DELFI), 65–76. <https://doi.org/10.18420/delfi2023-14>
- Shoufan, A. (2023). Can Students without Prior Knowledge Use ChatGPT to Answer Test Questions? An Empirical Study. In: *ACM Transactions on Computing Education*, 23(4), 1–29. <https://doi.org/10.1145/3628162>
- Weber-Wulff, D./Anohina-Naumeca, A./Bjelobaba, S./Foltýnek, T./Guerrero-Dib, J./Popoola, O./Šigut, P./Waddington, L. (2023). Testing of Detection Tools for AI-Generated Text. In: *International Journal for Educational Integrity*, 19(26). <https://doi.org/10.1007/s40979-023-00146-z>
- Wood, D. A./Achhpilia, M. P./Adams, M. T./Aghazadeh, S./Akinyele, K./Akpan, M./Allee, K. D./Allen, A. M./Almer, E. D./Ames, D./Arity, V./Barr-Pulliam, D./Basoglu, K. A./Belnap, A./Bentley, J. W. Berg, T./Berglund, N. R./Berry, E./Bhandari, A./.../Zoet, E. (2023). The ChatGPT Artificial Intelligence Chatbot: How Well Does It Answer Accounting Assessment Questions? In: *Issues in Accounting Education*, 38(4), 1–28. <https://doi.org/10.2308/ISSUES-2023-013>
- Zawacki-Richter, O./Marín, V. I./Bond, M./Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? In: *International Journal of Educational Technology in Higher Education*, 16(1), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>

Rechtliche und studentische Perspektiven auf das (digitale) Prüfen

Die rechtliche Beratung im PITCH-Projekt – Ein Erfahrungsbericht

Henrik Noll

1 Einleitung

Überall dort, wo Innovationen entstehen, werden diese von rechtlichen Unsicherheiten begleitet. Solche Unsicherheiten gab es auch im PITCH-Projekt, insbesondere im prüfungsrechtlichen Kontext. So hatten viele Lehrenden Fragen bei der Digitalisierung ihrer Prüfungen, aber sie hatten auch anderweitig Ideen entwickelt, wie sie ihre Prüfungsformate verbessern oder neu gestalten können. Hierbei konnten beispielsweise Fragen zu digitalen Prüfungen erörtert werden, wobei es unter anderem um den technischen Support bei einer elektronischen Prüfung ging oder ob es zulässig ist, eine Klausur in Papierform als Reserve bereitzuhalten, falls eine elektronische Prüfung kurzfristig aus technischen Gründen nicht möglich ist. Darüber hinaus konnten auch Fragen jenseits von digitalen Prüfungen beantwortet werden. Dies betraf Fragen zu Bonuspunkten oder zur Einsicht in die Prüfungsakte. Die wichtigsten Ergebnisse der rechtlichen Beratung wurden andersorts veröffentlicht und können auf der Webseite des Projekts (<https://www.uni-due.de/pitch/>) eingesehen werden (Noll, 2024).

Darüber hinaus bestand auch im PITCH-Technikteam ein Bedarf an rechtlicher Unterstützung. Dort ging es beispielsweise um die Frage, welche Daten gespeichert werden müssen, um eine elektronische Prüfung „gerichtsfest“ zu machen.

Im Projekt stand regelmäßig die Klärung der Rechtslage im Vordergrund. Im Fokus stand demnach die Frage, ob das Vorhaben zu diesem Zeitpunkt rechtlich zulässig ist. Teilweise ging es aber nicht nur darum, die Rechtslage zu klären, sondern darüber hinaus auch die Rechtslage durch das Schaffen neuer Rechtsgrundlagen zu verändern. So wurde im Rahmen des Projekts überlegt, einen Entwurf vorzulegen, der es ermöglichen würde, Online-Prüfungen auch mit Videoaufsicht durchzuführen. Dieses Vorhaben wurde dann aber letztlich verworfen, da das Land NRW durch die Neugestaltung der Hochschul-Digitalverordnung bereits selbst entsprechende Rechtsgrundlagen geschaffen hat.

In diesem Beitrag soll jedoch weniger über die gefundenen Ergebnisse berichtet werden, sondern es soll in den Fokus gerückt werden, welche Erfahrungen auf dem Weg zu diesen Ergebnissen gesammelt wurden und welche Schwierigkeiten dabei aufgetreten sind. Hierzu wird zunächst das Ziel erklärt, welches mit

der rechtlichen Beratung verfolgt wurde (s. Kapitel 2). Darauffolgend werden die einzelnen Schritte beschrieben, die unternommen wurden, um dieses Ziel zu erreichen (s. Kapitel 3). Innerhalb dieser Schritte soll anhand von Beispielen aufgezeigt werden, welche Hürden oder Hindernisse im Rahmen des Projekts aufgetreten sind.

Ziel

Zu den Aufgaben des Justitiariats einer Hochschule gehört nicht nur die rechtliche Begleitung bereits bestehender Rechtsstreitigkeiten, sondern auch präventiv die Vermeidung solcher Rechtsstreitigkeiten. Im Rahmen des vorliegenden Projekts stand die Prävention im Vordergrund. Vor allem sollen die neuen Prüfungsformate möglichst rechtssicher gestaltet werden. Hierzu wird wie folgt vorgegangen:

- Ermittlung des Sachverhalts (s. Kapitel 3.1)
- Verifizierung der Rechtsgrundlagen (s. Kapitel 3.2)
- Recherche der Rechtsprechung und Literatur (s. Kapitel 3.3)
- Finale Abwägung aller ermittelten Argumente (s. Kapitel 3.4)

Arbeitsweise

In diesem Beitrag wird die Arbeitsweise anhand der vier bereits zuvor genannten Schritte erläutert. Die vier Schritte können zwar chronologisch abgearbeitet werden, jedoch bietet es sich in manchen Fällen an, die Schritte parallel zu bearbeiten oder die Schritte zu wiederholen, falls im späteren Verlauf neue Erkenntnisse auftauchen, die beispielsweise eine neue Recherche in einem anderen Themengebiet erforderlich machen. In diesem Zusammenhang sei noch darauf hingewiesen, dass die Bearbeitung der rechtlichen Fragen in Abstimmung mit dem Justitiariat erfolgte. Hierdurch konnten bereits durch das Justitiariat bearbeitete Themen weiter verwertet werden und es wird sichergestellt, dass die hier gefundenen Ergebnisse durch das Justitiariat auch nach Ende des Projekts weiter genutzt werden können.

2 Sachverhaltsermittlung

Als Erstes ist der vollständige Sachverhalt zu ermitteln. Denn nur aufgrund des vollständigen Sachverhalts lässt sich die genaue Rechtslage ermitteln. Dazu gehört, dass die Fragen der Lehrenden richtig verstanden werden und mögliche Missverständnisse bereits vorher ausgeräumt werden können. Hierzu waren auch technische Fachfragen mit den entsprechenden Experten des Technikteams

zu erörtern oder sich Wissen über gewisse Funktionsweisen anzueignen, da aufgrund von falschen tatsächlichen Annahmen schnell eine falsche rechtliche Beurteilung erfolgen kann. So ist es beispielsweise bei der rechtlichen Bearbeitung im Themenkomplex der KI-Generatoren im Kontext von ChatGPT elementar, sich zuerst mit dessen Funktionsweise vertraut zu machen. Hinzu besteht natürlich immer die Gefahr, dass technische Annahmen im Zuge der Digitalisierung schon innerhalb weniger Monate überholt sein können und somit die rechtlichen Einschätzungen in diesem Bereich innerhalb einer kurzen Zeitspanne schon wieder obsolet sein können. So weit ist es im Rahmen des Projekts allerdings nicht gekommen.

3 Rechtsgrundlagen

Wie in amerikanischen Filmen oft zu sehen ist, stützen sich die Entscheidungen des Gerichts dort hauptsächlich auf Präzedenzfälle. Das heißt, die Anwälte:innen und Richter:innen argumentieren dort eher anhand anderer Gerichtsentscheidungen, die den ihnen vorliegenden Fällen vergleichbar sind. Im deutschen Recht haben die Präzedenzfälle keine derart ausgeprägte Hauptrolle in der rechtlichen Argumentation. Denn hier wird sich primär an den vorhandenen Rechtsnormen orientiert. Daher wird im ersten Schritt die Gesetzeslage verifiziert. Hierzu zählen bei Prüfungen insbesondere die Prüfungsordnungen, das Hochschulgesetz NRW (HG NRW), in selteneren Fällen auch die Datenschutz-Grundverordnung (DS-GVO) und natürlich das Grundgesetz (GG). Das Grundgesetz und die darin niedergeschriebenen Grundrechte spielen im prüfungsrechtlichen Rahmen eine besondere Rolle, da jedes Prüfungsvorhaben mit den Grundrechten vereinbar sein muss, wobei der Grundsatz der Chancengleichheit, der aus Art. 3 i. V. m. Art. 12 GG abgeleitet wird (BVerwG, Beschl. v. 30.06.2015 – 6 B 11/15 = NVwZ-RR 2015, 858), bei Prüfungen nochmals besonders hervorzuheben ist. Durch den Grundsatz der Chancengleichheit sollen die gleichen Startchancen gewährleistet werden, soweit die Studierenden denselben Beruf anstreben (vgl. Fischer/Jeremias/Dieterich, Prüfungsrecht, Rn. 402). Dieser Grundsatz bildet demnach das Fundament des Prüfungsrechts, und daher muss jede Maßnahme kritisch hinterfragt werden, ob sie dem Grundsatz der Chancengleichheit noch gerecht werden kann.

Im Rahmen des Projekts taten sich drei Problemfelder auf dieser Ebene auf. Das erste Problemfeld betrifft die Prüfungsordnungen. Diese enthalten die detailliertesten Regelungen hinsichtlich des Prüfungsverfahrens und sind somit einer der wichtigsten Rechtsquellen bei der rechtlichen Beurteilung von Prüfungsvorhaben. An der Universität Duisburg-Essen gibt es zunächst eine Rahmenprüfungsordnung (RPO), die vom Senat der Universität erlassen wird. Darüber hinaus gibt es fachspezifische Prüfungsordnungen, die von den Fakultäten für den jeweiligen Studiengang erlassen werden. Zwar sind die Fakultäten dazu ange-

halten, ihre Prüfungsordnungen an der Rahmenprüfungsordnung auszurichten, dennoch gibt es aus verschiedenen Gründen oft Unterschiede zwischen der RPO und fachspezifischen Prüfungsordnungen. Unmittelbar anwendbar sind allerdings ausschließlich die fachspezifischen Prüfungsordnungen der einzelnen Studiengänge.

Ein zweites Problemfeld betrifft die Frage, ob für bestimmte Prüfungsvorhaben überhaupt rechtliche Regelungen notwendig sind, falls weder in den Prüfungsordnungen noch in anderen Gesetzen solche zu finden sind. Hierzu sind dann die Literatur und Rechtsprechung heranzuziehen. Als Beispiel für dieses Problemfeld kann die Online-Prüfung angeführt werden. Denn um überhaupt eine Online-Prüfung durchzuführen, braucht es zunächst eine entsprechende Rechtsgrundlage (vgl. BeckOK HochschulR NRW/Birnbaum HG, § 64, Rn. 51). Zwar gibt es in § 14 Abs. 6 RPO eine Rechtsgrundlage für Online-Prüfungen, diese wurde jedoch noch nicht in allen fachspezifischen Prüfungsordnungen umgesetzt. Darüber hinaus bedarf es für Klausuren nach der herrschenden Meinung einer Aufsicht (VG Frankfurt (Oder) Beschl. v. 11.5.2021 – VG 1 L 124/2; Fischer/Jeremias/Dieterich, Prüfungsrecht, Rn. 28c). Um jedoch eine Online-Aufsicht durchzuführen, bedarf es wiederum einer gesetzlichen Grundlage. An einer solchen fehlte es lange, sodass Online-Prüfungen nicht durchgeführt werden durften. Dies führte dazu, dass bis zum jetzigen Zeitpunkt die Online-Prüfungen noch wenig beleuchtet werden konnten, da es bisher an den entsprechenden Rechtsgrundlagen fehlte.

Im dritten Problemfeld geht es um den Wandel der Gesetzeslage im Laufe der Zeit. Der Gesetzgeber kann die Gesetze ändern, abschaffen oder neue Gesetze verkünden. Als Beispiel kann auch hier wieder die Online-Prüfung herangezogen werden. Durch die neue Hochschul-Digitalverordnung (HDVO) hat das Land NRW neue gesetzliche Grundlagen für die digitale Lehre, aber auch für digitale Prüfungen geschaffen. So wurde unter anderem eine rechtliche Grundlage für die Online-Aufsicht eingeführt. Hierbei ist jedoch anzumerken, dass die neuen Regelungen hinsichtlich digitaler Prüfungen erst für das Wintersemester 2024/2025 und die folgenden Lehrveranstaltungen und Prüfungen Anwendung finden.

4 Recherche

4.1 Rechtsprechung

Eine wichtige Quelle für die Frage nach der Rechtslage stellen Gerichtsentscheidungen dar. Die Digitalisierung ist für die Rechtsprechung im Prüfungsrecht allerdings doch eher noch Neuland. Das heißt, dass es schlicht noch nicht so viele gerichtliche Entscheidungen gibt, die digitale Prüfungen betreffen. Gerade im Rahmen des PITCH-Projekts, in dem es eher um neue Prüfungsformate ging, war es daher unwahrscheinlich, Gerichtsentscheidungen zu finden, die genau

diese Thematik behandelten. Dennoch konnten aus älteren Gerichtsentscheidungen Schlüsse gezogen werden. Beispielsweise wurde von den Lehrenden angefragt, ob es zulässig wäre, wenn die Studierenden sich während einer Prüfung Tipps „erkaufen“ können. Dies bedeutet, dass die Studierenden sich für bestimmte Aufgaben Tipps ansehen können, aber dafür einen Punktabzug für diese Aufgabe hinnehmen müssen. Punktabzug oder sog. Maluspunkte gibt es ebenfalls bei sog. Multiple-Choice-Klausuren. Nach der Rechtsprechung dürfen bei Multiple-Choice-Klausuren den Studierenden zunächst keine Punkte wieder abgezogen werden für die Leistungen, die bereits erbracht wurden. Bei einem Prüfungsverfahren, durch dessen Ergebnis die Freiheit der Berufswahl eingeschränkt wird, dürfen fehlerfrei erbrachte Prüfungsleistungen nicht als nicht oder schlecht erbracht gewertet werden, weil andere Prüfungsfragen nicht richtig beantwortet wurden. Die Vorgehensweise wird als nicht geeignet angesehen, um festzustellen, welche berufsbezogenen Erkenntnisse die Studierenden erlangt haben (OVG Münster, Urteil vom 16.12.2008 – 14 A 2154/08). Diese Rechtsprechung kann auch auf den vorliegenden Fall angewendet werden. Der Punktabzug für einen erkauften Tipp darf demnach keine Auswirkungen auf die anderen Aufgaben haben.

Teilweise spielten auch aktuelle Entwicklungen in der Rechtsprechung bei der Arbeit im Projekt eine Rolle. Im Zusammenhang mit der DS-GVO war eine solche Entwicklung zu beobachten, welche Auswirkungen auf die Einsicht in Prüfungsakten durch die Studierenden hat. Das VG Gelsenkirchen hat mit Urteil vom 27.10.2020 (NVwZ-RR 2020, 1070) entschieden, dass gemäß Art. 15 Abs. 1 u. 3 DS-GVO die Studierenden Anspruch auf eine kostenlose Kopie ihrer Prüfungsarbeiten samt Prüfungsgutachten haben (ausführlich Noll, 2024, S. 29f.). Vorher konnten die Prüfungsbehörden selbst festlegen, wo und wie die Einsicht in die Prüfungsakte erfolgt. Durch die neue Rechtslage müssen Teile der Prüfungsakte als Kopie kostenlos an die Studierenden zugesandt werden, wenn diese einen Anspruch aus Art. 15 Abs. 1 u. 3 DS-GVO geltend machen.

4.2 Literatur

In der Literatur hingegen gibt es bereits mehrere Werke, die sich mit digitalen Prüfungen beschäftigt haben. Insbesondere ist hier das Standardwerk zum Prüfungsrecht von Fischer, Jeremias und Dieterich zu nennen. Dieses Lehrbuch wird regelmäßig auch von den Gerichten in ihren Entscheidungen herangezogen. Daher wird dort oft die sog. herrschende Meinung im Prüfungsrecht wiederzufinden sein. So finden sich dort bereits einschlägige Definitionen zu Prüfungen in elektronischer Form sowie zu den Prüfungen in elektronischer Kommunikation, wobei mit letzteren die Online-Prüfung gemeint sind, welche dann für das hiesige Projekt als Grundlage herangezogen werden konnten. Dazu gibt es noch Gutachten, Gesetzeskommentare oder fachspezifische Aufsätze zu dieser Thema-

tik. Hierbei ist darauf zu achten, dass die Literatur veraltet sein kann. Zum Beispiel, weil sie an technische Voraussetzungen knüpfen kann, die bereits überholt sind, oder weil sich die Gesellschaft hinsichtlich der Digitalisierung weiterentwickelt hat. So wird in der älteren Literatur noch des Öfteren diskutiert, ob Prüfungen in elektronischer Form an sich noch mit dem Grundsatz der Chancengleichheit vereinbar seien, denn schließlich können einige Studierende schneller am PC schreiben als andere Studierende (vgl. Forgó et al., 2016, S. 35f.). Dies dürfte aus heutiger Sicht bei den Gerichten kaum noch diskutiert werden, da davon ausgegangen werden kann, dass jede:r Studierende Zugang zu einem Computer hat und dementsprechend das Schreiben auf einer Tastatur üben kann.

Zum Teil gab es sogar Bereiche, in denen ausschließlich auf Literatur zurückgegriffen werden konnte. Dies war beispielsweise im Bereich der individualisierten Klausur der Fall. In diesem Bereich gibt es zwar kaum Rechtsprechung, aber die Literatur hatte schon Voraussetzungen entwickelt, die erfüllt sein müssten, um den rechtlichen Anforderungen gerecht zu werden. Da diese Voraussetzungen sich schon zu einer gefestigten herrschenden Ansicht in der Literatur entwickelt haben, ist davon auszugehen, dass auch die Gerichte der Literatur folgen werden. Somit konnte ohne Risiko die Literatur hier herangezogen werden.

5 Abwägung

Zum Schluss müssen alle gesammelten Argumente gegeneinander abgewogen werden, um anschließend zu einer fundierten rechtlichen Einschätzung zu gelangen. Hierbei kommt es darauf an, wie überzeugend die Argumente sind, aber auch von welcher Seite die Argumentation kommt. So ist es deutlich wahrscheinlicher, dass sich ein Gericht der Argumentation eines Oberverwaltungsgerichts oder sogar des Bundesverwaltungsgerichts anschließt, als derjenigen, die lediglich in einem Aufsatz vertreten wird. Außerdem kann es vorkommen, dass es zu einem Thema kaum Material in der Rechtsprechung oder Literatur gibt. Hierbei wird auf die klassischen juristischen Werkzeuge zurückgegriffen, nämlich die Auslegung von Rechtsnormen. Ein solcher Fall lag bei der Frage nach der Zulässigkeit von einer „Back-Up-Klausur“ vor. Mit Back-Up-Klausur ist gemeint, dass bei einem Technikausfall am Prüfungstag kurzfristig von einer Prüfung in elektronischer Form zu einer Prüfung mit Stift und Papier umgeschwenkt werden soll. Für diese Konstellation gab es bisher keine Rechtsprechung, noch ist zu dieser Thematik etwas in der Literatur zu finden gewesen. Anknüpfungspunkt ist hier eine Regelung der RPO (§ 14 Abs. 7 S. 3), wonach die Prüfungsform grundsätzlich mit Beginn der Lehr-/Lernform nicht mehr gewechselt werden darf. Würde man sich strikt an den Wortlaut der Norm orientieren, dürfte eine Back-Up-Klausur nicht zulässig sein. Nachdem jedoch der Sinn und Zweck dieser Norm herausgearbeitet werden konnte, konnten daraus Voraussetzungen abgeleitet werden, nach denen eine Back-Up-Klausur doch noch zulässig sein könnte.

Hieraus ergab sich aber noch ein weiterer Punkt, der bei rechtlichen Stellungnahmen zu beachten ist. So ist immer auch auf die Risiken hinzuweisen. Im Fall der Back-Up-Klausur beispielsweise kann nach der hier vertretenen Ansicht zwar davon ausgegangen werden, dass ein solches Vorhaben unter bestimmten Voraussetzungen zulässig ist. Jedoch könnte das Gericht auch eine andere Ansicht vertreten. Beispielsweise könnte es im extremsten Fall ein solches Vorhaben generell für unzulässig erklären. Daher wäre die „rechtssicherste“ Variante im Falle eines kurzfristigen Technikausfalls, die Prüfung abzubrechen und die Prüfung zu einem späteren Zeitpunkt zu wiederholen.

Teilweise konnte aber auch die womöglich herrschende Meinung in der Literatur nicht vollends überzeugen. Es stellte sich im Rahmen des Projekts die Frage, wie eine Klausur von einer Hausarbeit zu unterscheiden ist. Es gibt zur Frage der Abgrenzung kaum Rechtsprechung und auch die rechtswissenschaftliche Literatur beschäftigt sich kaum mit der Thematik. Das mag zum einen daran liegen, dass die Hochschulen in ihren Prüfungsordnungen selbst die Prüfungsformen festlegen, definieren und zum Teil sogar Eigenkreationen entwickeln können, und zum anderen bisher einfach in den meisten Fällen die „klassischen“ Prüfungsformate genutzt wurden, sodass es eindeutig war, welche Prüfungsform vorliegt und es keiner genaueren Definition oder Abgrenzung bedurfte.

Dies änderte sich jedoch mit dem Beginn der Corona-Pandemie. In dieser Zeit haben viele Lehrende auf das sog. Take-Home-Examen gesetzt. Hierbei wurden die Prüfungsaufgaben online zum Download zur Verfügung gestellt und die Studierenden mussten dann bis zu einem gewissen Zeitpunkt ihre Prüfungsleistung auf einer Datei hochladen. Eine Videoaufsicht wurde dabei nicht durchgeführt. Hier kam es dann darauf an, welche Prüfungsform vorliegt. In Betracht kam zunächst die Klausur. Allerdings sagt die herrschende Meinung in der Literatur und Rechtsprechung, dass eine Klausur mit einer Aufsichtsarbeit gleichzusetzen ist und eine Klausur daher nur eine Klausur darstellt, wenn sie unter Aufsicht geschrieben wird. (VG Frankfurt (Oder) Beschl. v. 11.5.2021 – VG 1 L 124/2; Fischer/Jeremias/Dieterich, Prüfungsrecht, Rn. 28c. a.A. Birnbaum, Bildungsrecht in der Corona-Krise, 1. Auflage 2021, § 4, Rn. 62 f.). Eine Rechtsansicht will daher ohne nähere Begründung diese Take-Home-Examen als Hausarbeit ansehen (Fischer/Jeremias/Dieterich, Prüfungsrecht, Rn. 28c). Diese Ansicht versäumt es jedoch, in dieser Sache weiter zu differenzieren, und kann daher nicht überzeugen. Es mangelt an weiteren Abgrenzungsmerkmalen, da andernfalls die Hausarbeit in gewisser Weise als Auffangprüfungsformat angesehen werden könnte. In Betracht kommen eine Unterscheidung in der zeitlichen Komponente sowie eine Unterscheidung beim Bearbeitungsort. Auf der zeitlichen Ebene ist es überzeugend, dass eine Hausarbeit immer dann vorliegt, wenn der den Studierenden tatsächlich zur Verfügung stehende Zeitraum größer ist als die angedachte Bearbeitungszeit der Prüfung (Escher-Weingart, 2021, S. 18f). Die Studierenden können demnach innerhalb dieses Zeitslots ihre Bearbeitungszeit frei wählen und ihre Arbeitszeit frei aufteilen. Bei einer Klausur hingegen

ist die Bearbeitungszeit festgelegt und kann nicht selbstständig aufgeteilt werden. So zeichnet sich eine Klausur durch eine begrenzte Bearbeitungszeit aus, § 17 Abs. 1 RPO.

Ein weiterer Unterschied zwischen Klausur und Hausarbeit besteht bei der Wahl des Bearbeitungsortes. Während dieser bei der Hausarbeit durch die Studierenden frei gewählt und beliebig oft gewechselt werden darf, ist der Bearbeitungsort bei einer Klausur vorher festgelegt. Ob es sich beim Take-Home-Examen um eine Klausur oder Hausarbeit handelt, ist daher vom Einzelfall und den jeweiligen Umständen abhängig. Wenn weder eine Klausur noch eine Hausarbeit vorliegt, könnte man eine neue Prüfungsform in der Prüfungsordnung festlegen, § 13 Abs. 6 lit. g RPO.

In diesem Fall konnte daher die vermutlich herrschende Ansicht der Literatur weniger überzeugen und es musste vermehrt auf eigene Argumentation gesetzt werden, welche jedoch mit Argumenten aus einem anderen Aufsatz untermauert werden konnte.

6 Fazit und Ausblick

Mit neuen technologischen Entwicklungen ergeben sich für die Prüfenden immer mehr Möglichkeiten, ihre Prüfungen an diese anzupassen und diese zu verbessern. Im Rahmen des PITCH-Projekts konnten gerade im Bereich der E-Prüfungen einige Hürden genommen werden. Doch nicht nur im Bereich der E-Prüfungen konnten einige Fragen geklärt werden, sondern auch dort, wo sich im Zuge der voranschreitenden Digitalisierung neue Möglichkeiten bei Prüfungen ergeben haben, konnten einige Antworten gegeben werden.

Durch die neue HDVO wurde der Spielraum für die Lehrenden bei Prüfungen noch einmal vergrößert, sodass sich bald neue Fragen im Bereich der Online-Prüfung stellen werden. Eine weitere Herausforderung dürfte die Entwicklung im Bereich der Künstlichen Intelligenz darstellen. Denn durch KI-Textgeneratoren wie beispielsweise ChatGPT gibt es nicht nur für die Prüfenden neue Möglichkeiten für die Erstellung von Prüfungen, sondern auch für die Studierenden bei diesen Prüfungen zu täuschen. Es bleibt abzuwarten, wie sich sowohl die tatsächlichen als auch die rechtlichen Rahmenbedingungen in Zukunft entwickeln werden.

Durch die enge Abstimmung mit dem Justitiariat ist gewährleistet, dass die herausgearbeiteten Ergebnisse auch nach Ende des Projekts erhalten bleiben und somit die Arbeit am Projekt nachhaltig ist.

Literatur

- Birnbaum, C. (2021). *Bildungsrecht in der Corona-Krise*, 1. Auflage, München: C. H. BECK
- von Coelln, C./Schemmer, J. (2023). *BeckOK Hochschulrecht Nordrhein-Westfalen*, 29. Edition, München: C. H. BECK
- Escher-Weingart, C. (2021). *Die Prüfung – das unbekannte Wesen*, <https://doi.org/10.5445/IR/1000143052>
- Fischer, E./Jeremias, C./Dieterich, P. (2022). *Prüfungsrecht*, 8. Auflage, München: C. H. BECK
- Forgó, N./Graupe, S./Pfeiffenbring, J. (2016). *Gutachten über rechtliche Aspekte von E-Assessments an Hochschulen*, Essen, <https://doi.org/10.17185/dupublico/42871>
- Noll, H. (2024). *Prüfungen und Recht – Digitale Zeiten*. https://www.uni-due.de/imperia/md/content/zhqe/handout_henriknoll_final.pdf

Die Student Voice Group im Projekt PITCH – Ein Praxisbericht

Maiken Bonnes und Kilian Schmitt (in Zusammenarbeit mit Colin Corbach, Ronja Hensle, Marc Ludwig, Sophie Stotz Anido und Frederike Welzel)

Die aktive Einbindung der studentischen Perspektive ist ein zentraler Baustein des Projekts PITCH, der auch im Projektantrag verankert ist. Partizipativen Ansätzen wie dem Students-as-Partners-Ansatz folgend wurde mit Projektbeginn die Student Voice Group (SVG) gegründet, deren Mitglieder als studentische Hilfskräfte im Zentrum für Hochschulqualitätsentwicklung (ZHQE) der UDE beschäftigt werden. In Form einer moderierten, interdisziplinären Arbeitsgruppe bringt die SVG eine unabhängige studentische Perspektive auf die Weiterentwicklung von Prüfungen an Hochschulen ein und begleitet die Projektarbeit. Die studentischen Mitarbeiter:innen betonen Bedarfe, Erfahrungen und Erkenntnisse von Studierenden und verfassen Handreichungen mit Empfehlungen für (digitale) Prüfungen für Lehrende und Studierende. Das Ziel dieses Beitrags ist es, die Zusammenarbeit mit den Studierenden im Projekt PITCH zu erläutern, die zentralen Erkenntnisse und Positionen der Studierenden zu beschreiben sowie die Vorteile, aber auch die Herausforderungen in der Projektarbeit zu beleuchten.

1 Die Idee hinter dem Konzept der Student Voice Group

Studierende stellen die zentrale und größte Mitgliedergruppe an Hochschulen dar. Befragungen zeigen, dass sie auch grundsätzlich bereit sind, sich bei der Weiterentwicklung von Studium und Lehre sowie der Hochschulstrukturen einzubringen. Eine breite Beteiligungskultur ist derzeit jedoch weder in der studentischen Selbstverwaltung noch in Prozessen der Hochschul- bzw. Studium- und Lehrentwicklung zu beobachten. Als Gründe für das ausbleibende Engagement werden u. a. der Mangel an zeitlichen Ressourcen, an Informationen oder an dem Gefühl, tatsächlich etwas bewirken zu können, genannt (vgl. Ditzel/Bergt, 2013, S. 181ff.; Raffaele/Rediger, 2021, S. 16 ff.; Schrader, 2023, S. 13).

Jedoch nehmen Studierende in ihrem Studienalltag Probleme, Konflikte und Erschwernisse, aber auch positive (Lehr-)Praxisbeispiele wahr und können somit als „lebendige Monitoring-Akteure“ (Raffaele/Rediger, 2021, S. 5) angesehen werden. Ihr Erfahrungshorizont ist für das Projekt PITCH von besonderer Bedeutung. Deshalb wurden für die Student Voice Group eigene Projektmittel beantragt und Stellen für studentische Hilfskräfte mit einem Stundenumfang von

bis zu zehn Stunden pro Woche ausgeschrieben. Dieser materielle Anreiz sollte dem o. g. Hinderungsfaktor Zeitmangel entgegenwirken, da die Mitglieder der SVG so die studentische Erwerbstätigkeit und ihr Engagement innerhalb der UDE kombinieren können. Wie sich ihr Wirken im Rahmen des Projekts PITCH gestaltet, wird nachfolgend beschrieben.

2 Die praktische Umsetzung

2.1 Die Aufgaben und Zusammenarbeit im Team

Die SVG bestand in der ersten Förderperiode aus sechs Studierenden unterschiedlicher Bachelor- und Masterstudiengänge. So konnten ihre Erfahrungen aus den Sozial- und Gesellschaftswissenschaften, der (Wirtschafts-)Informatik, den Geisteswissenschaften und der lehramtsbildenden Studiengänge, aber auch aus ihrem Erleben des Hochschulalltags in das Projekt PITCH einfließen. Die SVG arbeitete eng angebunden an die Projektkoordination, die das Ziel verfolgte, den von Sather et al. (2024, S. 6f.) als kollaborativ und gegenseitig beschriebenen Prozess im Arbeitsalltag umzusetzen. Die SHK-Stellen waren, wie auch die Projektkoordination, organisatorisch im Zentrum für Hochschulqualitätsentwicklung (ZHQE) angesiedelt.

Da die praktische Umsetzung der studentischen Beteiligung im Projektantrag nicht konkret in Meilensteinen formuliert wurde, hatten die Mitglieder der SVG große Freiheitsgrade, ihre Partizipationsmöglichkeiten im Projekt mitzugestalten. Die Hauptaufgabe der SVG bestand darin, sich gemeinsam mit selbst ausgewählten Themen aus dem Projektkontext wissenschaftlich fundiert auseinanderzusetzen. Dabei sollten sowohl Handreichungen für Lehrende und Beiträge auf Tagungen entstehen, als auch Workshops auf den projektinternen Konferenzen durchgeführt werden. Unterstützt wurden die SVG-Mitglieder von ihren Ansprechpartner:innen in der PITCH-Projektkoordination. Um der anspruchsvollen Aufgabe des wissenschaftlichen Arbeitens gerecht werden zu können, wurden mehrere Schreibwerkstätten in Form von Tagesworkshops zu unterschiedlichen Zeitpunkten des Schreibprozesses für die SVG von einer Schreibdidaktikerin aus dem ZHQE durchgeführt. In diesem Rahmen arbeiteten die SVG-Mitglieder auch ihre Schwerpunktthemen heraus, mit denen sie sich theoretisch und empirisch vertieft beschäftigten: anonymisiertes Prüfen, Täuschungsversuche und innovative Prüfungsformate.

2.2 Inhaltliche Einblicke in die Arbeit der Student Voice Group

2.2.1 Anonymisiertes Prüfen

Im gegenwärtigen wissenschaftlichen Diskurs ist eine zunehmende Thematisierung von Diversität im Zusammenhang mit Diskriminierungskritik an Hochschulen festzustellen. Diese zeigt sich u. a. in der Auseinandersetzung mit diversitätssensibler, diskriminierungskritischer Lehre oder auch in Diversity-Politiken auf der organisationalen Ebene von Hochschulen (Breiwe/Ebert, 2023; Resch/Raschauer, 2019; Thompson/Vorbrugg, 2018). Auch an der UDE werden bspw. durch die Einrichtung eines Diversity Support Centers als zentraler Serviceeinrichtung, die Etablierung einer Antidiskriminierungsrichtlinie sowie die Benennung einer Antidiskriminierungsbeauftragten Strukturen geschaffen und Maßnahmen ergriffen, universitäre Prozesse frei von Diskriminierung zu gestalten und die Chancengerechtigkeit zu erhöhen (Hausess, 2023). Auch das Projekt PITCH macht sich dafür stark, Chancengerechtigkeit zu fördern und wirft einen Blick auf die Umsetzung von digitalen Prüfungen, die dem Anspruch einer „zeitgemäße[n] diversitätssensible[n] Hochschullehre“ (Kergel/Heidkamp, 2019, S. 3) gerecht werden. Die SVG sieht ein hohes Potenzial im anonymisierten Prüfen und beschäftigte sich daher intensiv mit den Vor- und Nachteilen aus studentischer Sicht.

Beim anonymisierten Prüfen werden die Identitäten der Prüfungsteilnehmenden während des Prüfungs- und Bewertungsprozesses bewusst verborgen und die personenbezogenen Daten erst nach Abschluss der Korrektur wieder mit der Identität der Studierenden zusammengeführt. Es ermöglicht den Studierenden, dass ihre Prüfungsleistungen unter gleichen Bedingungen bewertet werden und beugt eventueller Diskriminierung vor. Digitales Prüfen kann anonymisiertes Prüfen unterstützen, indem es technologische Mittel bereitstellt, die die Anonymität gewährleisten. Zum Beispiel können Prüfungsplattformen so konfiguriert werden, dass sie die Identität der Prüflinge während der Auswertung nicht offenlegen. Im Projekt PITCH werden solche Technologien eingesetzt.

Die SVG interviewte Lehrende der UDE, die solche Verfahren anwenden, und führte anschließend im April und Mai 2023 explorative Campusgespräche mit 54 Studierenden der UDE zu ihrer Einstellung gegenüber anonymisiertem Prüfen. Dabei wurden auch Diskriminierungserfahrungen thematisiert, da Studien zeigen, dass Diskriminierung an akademischen Bildungseinrichtungen ein reales Problem darstellt. So zeigte „Die Studierendenbefragung in Deutschland“ (Meyer et al., 2022, S. 4; Beuße et al., 2022, S. 1), verantwortet vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) und der AG Hochschulforschung an der Universität Konstanz, für das Befragungsjahr 2021, dass 26 % der befragten Studierenden selbst Diskriminierung erfahren haben, während 46 % angaben, dass ihre Kommiliton:innen diskriminiert wurden. Weitere internationale Studien untersuchten mögliche Diskriminierung bei der Notenver-

gabe und stellen dabei leistungsunabhängige Bewertungsunterschiede zwischen verschiedenen Gruppen fest. Besonders häufig wird Diskriminierung aufgrund des Geschlechts oder der Ethnie festgestellt, aber auch Faktoren wie persönliche Bekanntheit mit der Lehrkraft oder Attraktivität spielen eine Rolle bei der Bevorzugung bzw. Benachteiligung (Malouff/Thorsteinsson, 2016; Zanga/Giannakis, 2023). Diese Effekte werden nicht von allen Studien bestätigt (Bischoff et al., 2021; Sedlacek, 2021). Trotzdem stellt die Hochschule noch keinen diskriminierungsfreien Raum dar.

In den Campusgesprächen an der UDE zeigte sich ein weniger gravierendes Bild. Nur wenige Studierende berichten, selbst Diskriminierungserfahrungen an der Hochschule gemacht zu haben bzw. Bevorzugungen oder Benachteiligungen bspw. aufgrund der ethnischen Herkunft bzw. des Namens bei anderen Studierenden beobachtet zu haben. Benachteiligungen aufgrund persönlicher Aspekte berichten Studierende ebenfalls nur im Ausnahmefall, wie bspw. eine Person, die die faire Benotung ihrer Prüfungsleistung in Frage stellte, da der Dozent im Vorfeld „viel offener“ auf die Fragen ihrer Kommiliton:innen eingegangen sei, als auf ihre (Student:in #4 [Essen], 2023). Doch auch die Studierenden, die keine eigenen Erfahrungen mit Diskriminierung gemacht haben, sehen zumindest ein Risiko. Das anonymisierte Prüfen erscheint ihnen als geeignetes Mittel, die Chancengerechtigkeit zu erhöhen. Sorgen bereitet den Studierenden allerdings die Geschwindigkeit der Korrektur sowie die korrekte Zuordnung der Noten zu den Prüflingen. Beides sollte bei korrekter Anwendung der Anonymisierung jedoch kein Problem darstellen.

Die Lehrenden der UDE, die v. a. in stark nachgefragten Fächern wie den Wirtschafts- oder Bildungswissenschaften lehren und prüfen, beschreiben, dass sich aufgrund der großen Prüfungskohorten mit teilweise mehr als 500 Prüfungsteilnehmer:innen, eine Anonymisierung von sich aus ergibt, da sie nicht alle Studierenden kennen. Den Kritikpunkt, dass eine systematische Anonymisierung dazu beiträgt, dass die Universität unpersönlicher wird, teilen sie (vgl. Pitt & Winstone, 2018, S. 1191). In kleineren Fächern fällt die Anonymisierung trotz technischer Möglichkeiten dann schwerer, wenn offene Prüfungsfragen gestellt werden. Lehrende berichten, dass sie im engen Kontakt mit den Studierenden mitunter einen Eindruck gewinnen, welche für sie typischen Fehler die Studierenden machen bzw. wie sie an die Aufgabenlösung herangehen und können so teilweise Rückschlüsse auf die Person ziehen. Dies lässt sich nicht verhindern, ebenso wie bei Prüfungsformaten wie mündlichen Prüfungen oder auch Hausarbeiten, die eine Betreuung durch die Lehrperson voraussetzen und bei der eine Anonymisierung sogar hinderlich wäre. Es ist daher notwendig, selbstreflexiv ein Bewusstsein für mögliche „unconscious bias“ (Nalty, 2016, S. 45–46) zu gewinnen und so die Objektivität bei der Benotung herzustellen.

Für die SVG überwiegen die Vorteile des anonymisierten Prüfens, wenn die Prüfungsform es zulässt. Sie empfiehlt daher die Einführung von anonymisierten Prüfungen als wenig aufwändigen Beitrag dazu, Prüfungen fairer zu gestalten.

2.2.2 Täuschungsversuche

Akademisches Fehlverhalten wie Täuschungsversuche in Prüfungen sind sowohl für die Lehrenden als auch für die Studierenden ein emotionales Thema. Das International Center for Academic Integrity (ICAI) geht aktuell davon aus, dass ca. 30 % der Studierenden an US-amerikanischen Hochschulen in mindestens einer Prüfung getäuscht haben (International Center for Academic Integrity, o. J.). Vergleichbar umfassende Daten für Deutschland fehlen, aber zu einem ähnlichen Ergebnis kam auch die Untersuchung der Universität Bielefeld im Rahmen des Projekts „FAIRUSE“: 37 % der befragten Studierenden der Universität Bielefeld haben mindestens einmal in einer Klausur abgeschrieben, fast jede:r Dritte hatte Spickzettel dabei, auch wenn diese nicht immer genutzt wurden. Mehr als ein Drittel der Befragten hatte schon Arbeitsaufgaben wie Protokolle oder Programmieraufgaben abgeschrieben, fast ein Viertel hatte Daten gefälscht, und 15 % haben ein Attest vorgelegt, ohne krank zu sein, um bspw. Abgabefristen zu verlängern (Sattler/Diewald, 2013). Dabei verletzen Täuschungsversuche die Chancengleichheit, weil unehrliche Prüfungsteilnehmer:innen sich einen ungerechtfertigten Vorteil gegenüber den ehrlichen Prüfungsteilnehmer:innen verschaffen (vgl. Jantos et al., 2022, S. 5). Auch auf individueller Ebene der Studierenden zeigen sich ebenfalls negative Effekte. So werden durch Betrugsverhalten Lerngelegenheiten und der Zuwachs an Wissen und Fähigkeiten verpasst (vgl. Janke et al., 2023, S. 208).

Digitale Prüfungen haben die Möglichkeiten für betrügerisches Verhalten erweitert, bieten aber auch vielfältige Möglichkeiten, Täuschungen zu vermeiden. So kann bspw. durch die Individualisierung von Prüfungsaufgaben oder die vollautomatische Erstellung mehrerer Varianten derselben Klausur das Kopieren von Lösungen erschwert werden. Außerdem erlaubt der Einsatz datenanalytischer Verfahren das (teil-)automatisierte Aufdecken unerlaubter Zusammenarbeit (vgl. Persike, 2021, S.339f.).

In den Workshops der SVG auf der PITCH-Projektkonferenz im Dezember 2022 zu Täuschungsversuchen in digitalen Prüfungen stellte sich heraus, dass sich sowohl Studierende als auch Lehrende in einem Spannungsfeld zwischen der (ungerechtfertigten) Unterstellung akademischer Unehrllichkeit und dem berechtigten Ziel, Täuschungsversuche zu unterbinden, befinden. Teilweise äußerten Lehrende große Frustration, da sie in digitalen Klausuren vermehrte und umfangreiche Täuschungsversuche feststellten, die noch über das o. g. Ausmaß hinausgingen. Die persönliche Enttäuschung, aber auch der Mehraufwand im Korrekturprozess waren immens. Seitens der Studierenden wurde die Sorge geäußert, unter „Generalverdacht“ zu geraten.

Doch warum unternehmen Studierende Täuschungsversuche und welche Maßnahmen können Hochschulen ergreifen, um dies zu verhindern? Hierzu führte die SVG im April und Mai 2023 explorative Campusgespräche mit Studierenden der UDE durch. In diesen 17 Einzel- und Gruppeninterviews mit ins-

gesamt 23 Studierenden wurden unterschiedliche und vielfältige Perspektiven diskutiert. Zunächst zeigte sich, dass die Mehrheit der befragten Studierenden von Täuschungsversuchen wusste, die durch Studierende in ihrem Bekanntenkreis unternommen wurden. Eine Minderheit hatte bereits selbst in Prüfungssituationen getäuscht.

Besonders das Täuschen bei Online-Klausuren erscheint den Studierenden einfacher als in Präsenzklausuren. Ein:e Student:in erklärte, dass „es in den Corona-Jahren einfach einfacher war zu täuschen“ und sie daher schneller in Versuchung gekommen sei (Studierendengruppe #3 [Duisburg], 2023). Ein anderer Student berichtete, dass er von anderen Studierenden gehört habe, dass Täuschen bei Online-Klausuren sehr einfach sei (Student:in #5 [Duisburg], 2023). Mobile Endgeräte erleichtern den Zugang zu und die Replikation von Informationen besonders in E-Klausuren, die oft ohne hinreichende Beaufsichtigung durchgeführt werden (Bilen/Matros 2020). Außerdem steigt nach Aussagen der befragten Studierenden die Wahrscheinlichkeit, in Prüfungen zu täuschen, deren Sinn sich ihnen nicht erschließt und die als reine Wissensabfragen wahrgenommen werden. Auch persönliche Gründe wie Faulheit oder schlechte Vorbereitung werden als begünstigende Faktoren für Täuschungsversuche genannt. In diesen Fällen fehlt den Studierenden die intrinsische Motivation zum Lernen, das Fehlverhalten stellt hingegen einen Gewinn dar (vgl. Sattler/Diewald, 2013, S. 29). In Prüfungen zu täuschen, kann aber auch als Bewältigungsstrategie bei Leistungs- und Prüfungsangst oder Stress bspw. durch zu viele Prüfungen bzw. einem zu hohen Schwierigkeitsgrad der Prüfung angesehen werden, wie vereinzelt berichtet wird.

Die Studierenden wurden abschließend nach präventiven Maßnahmen befragt. In ca. der Hälfte der Gespräche wurde darauf hingewiesen, dass gut vorbereitete Studierende in der Regel keinen Anlass sehen, zu täuschen. Ebenso oft wurde der Einfluss der Prüfungsformate diskutiert, wobei keine Einigkeit darüber bestand, welche Formate präventiv wirken. Präsenzklausuren wurden wegen der Aufsicht als präventiv betrachtet, während auch Take-Home-, Open-Book- und Multiple-Choice-Klausuren sowie mündliche Prüfungen erwähnt wurden. Interessanterweise wurden die ersten drei dieser Formate in anderen Gesprächen als fördernd für Täuschungsversuche bewertet. Ein weiterer mehrfach angesprochener Aspekt war die Wahrscheinlichkeit der Entdeckung. Studierende äußerten die Vermutung, dass weniger getäuscht wird, wenn die Entdeckungswahrscheinlichkeit höher ist, etwa durch verstärkte Aufsicht während der Prüfungen. In dieser Hinsicht werden die viel diskutierten automatisierten Proctoring-Systeme, wie sie teilweise an anderen Universitäten, nicht jedoch an der UDE, in Onlineprüfungen eingesetzt werden, von der SVG kritisch beurteilt, obwohl sie durchaus effektiv darin sein können, Täuschungsversuche zu unterbinden (Ackermann, 2021). Darüber hinaus wurden weitere, weniger häufig genannte Faktoren thematisiert, wie bspw. die Sympathie zur Lehrperson, die Art des Lehrmaterials und die Möglichkeit einer freieren Modulwahl, die die Tendenz, in Prüfungen zu täuschen, senken.

Insgesamt deuten die Gespräche auf dem Campus darauf hin, dass die Anpassung der Lehre ein geeigneter Hebel zur Verhinderung von Täuschungsversuchen sein kann. Durch die Schaffung einer Kultur der Integrität, die Unehrlichkeit institutionell ablehnt, könnte die Einstellung der Studierenden durch die Hochschule positiv beeinflusst werden, da Studien ebenfalls zeigen, dass Studierende häufiger täuschen, wenn sie annehmen, dass auch ihre Kommiliton:innen dies tun (vgl. Kroher, 2020, S. 224f.; Stephens, 2019, S. 11f). Die Student Voice Group sieht darüber hinaus die Notwendigkeit, präventive Maßnahmen in die Lehre selbst zu integrieren. Sie plädiert für eine stärkere Einbindung der Studierenden in die Entwicklung von Prüfungsformaten und für eine transparente Kommunikation der Regeln und Erwartungen, um unabsichtlichen Verstößen vorzubeugen (Stone, 2023; Waltzer et al., 2023). Aus ihrer Sicht würde dies so nicht nur das Verständnis für akademische Integrität fördern, sondern auch die Bereitschaft der Studierenden erhöhen, diese Werte in ihrem Studium zu leben.

Mit Blick auf die Weiterentwicklung der Lehre sieht die SVG ein großes Potenzial in praxisorientierten und individuellen Lehr- und Prüfungsformaten. Das Ziel sollte eine Lernzielorientierung sein, bei der der intrinsische Wert des Lernens im Mittelpunkt steht. Maßnahmen wie die Nutzung von Hilfsmitteln in Prüfungen durch Open-Book-Klausuren oder Aufgaben auf höheren Kompetenzniveaus könnten aus ihrer Sicht maßgeblich dazu beitragen, dass übliche Täuschungsmethoden unwirksam werden (vgl. Noorbehahani et al., 2022, S. 8439). Vor allem mit innovativen (digitalen) Prüfungsformaten setze sich die SVG auch weiter auseinander (siehe das nachfolgende Kapitel 2.2.3).

2.2.3 Innovative Prüfungsformate

Bei der Auseinandersetzung mit innovativen Prüfungsformaten konnten die Mitglieder der SVG auf ihre unterschiedlichen Erfahrungen mit Prüfungen an der Universität Duisburg-Essen zurückgreifen. Auch das sog. „Bulimie-Lernen“ war ihnen bekannt und wurde als belastend und nicht nachhaltig kritisiert. Im Rahmen ihres Workshops zu Täuschungsversuchen auf der PITCH-Projektkonferenz im Dezember 2022 wurde dann die Grundsatzfrage aufgeworfen, ob die generelle Abkehr von Prüfungen an Hochschulen vor allem vor dem Hintergrund der generativen Künstlichen Intelligenz ein zukunftsfähiger Weg sei.

Doch sieht die SVG vor allem die Vorteile für den studentischen Lernfortschritt in Prüfungsformaten, die die Kompetenzorientierung in den Vordergrund stellen und gleichzeitig zum Lernen motivieren. Sie setzte sich daraufhin mit guten Beispielen formativer und summativer (digitaler) Prüfungen aus ihrem Studium auseinander und erarbeitete folgende zwei Beiträge für Kurzworkshops für Konferenzen, die sie im Jahr 2023 auch durchführten:

- Online-Kurzworkshop „Prüfungswerkstatt – Digitale Prüfungen an Hochschulen realisieren“ auf der Werkstatt-Konferenz „Digitale Transformationen in hochschulischer Bildung gestalten“ der Universität Vechta am 23.02.2023.

- Kurzworkshop „Prüfungsformate aus studentischer Sicht – Wie nachhaltig ist das erworbene Wissen aus Prüfungen für Studierende?“ im Rahmen der Pitch-ches „KI, Kollaboration & Prüfung“ im Workshop Decoding auf der TURN Conference 2023 an der Technischen Hochschule Köln am 14.09.2023.

Auf den Konferenzen diskutierten die Mitglieder der SVG ihre Erfahrungen und Vorschläge mit Hochschullehrenden, Studierenden und Akteur:innen aus dem Bereich der Hochschuldidaktik. Sie kamen zu dem Schluss, dass u. a. projektbegleitende Prüfungen, z. B. Portfolio-Prüfungen, Prüfungen mit Transferaufgaben, z. B. Take-Home-Exams oder Open-Book-Klausuren, Prüfungen, die sich an künftigen Berufen orientieren, also eine konkrete Vorbereitung auf das Berufsleben darstellen sowie Prüfungsformen, die auf Bedürfnisse, die Persönlichkeit und die Fähigkeiten des Lernenden abzielen und das gemeinsame Arbeiten fördern zum einen geeignet sind, um die unterschiedlichen Kompetenzstufen prüfen zu können und zum anderen zum Lernen motivieren.

Im Projekt PITCH werden kompetenzorientierte und alternative digitale Prüfungsformate erprobt. Nach den Prüfungen werden die Prüfungen mittels Fragebogen evaluiert. Fokusgruppen ergänzten die Ergebnisse um die qualitative Perspektive. Die SVG wurde in die Konzeption des Fragebogens und die Reflexion der Ergebnisse einbezogen, sodass ihre Erfahrungen hier direkt einfließen konnten. Insbesondere die Weiterentwicklung der Prüfungen vor dem Hintergrund der Kompetenzorientierung ist Teil der zweiten Förderperiode des Projekts PITCH und auch der thematische Fokus der SVG.

3 Resümee: Stärken und Stolpersteine

Die SVG ermöglichte allen Projektbeteiligten einen Perspektivwechsel. Als Sprachrohr der Studierenden hatten die SVG-Mitglieder die teilweise unbequeme Rolle, Lösungsansätze des Projekts in Frage zu stellen und auf Probleme hinzuweisen. Da dies in konstruktiver und lösungsorientierter Weise geschah, wurde ihre Sicht als Bereicherung seitens der Lehrendenteams und der Projektkoordination empfunden und somit eine starke Ressource des Projekts PITCH. Im Arbeitsalltag ergaben sich jedoch einige Herausforderungen.

Eine große Herausforderung in der Zusammenarbeit bestand darin, eine Besprechungskultur aufzubauen, in der sowohl ein enger Austausch zwischen den Mitgliedern der SVG untereinander und mit der PITCH-Projektkoordination möglich war, aber auch die flexiblen Arbeitszeiten der studentischen Mitarbeiter:innen eingehalten wurden, die zu ihrem jeweiligen Studienplan passen sollten. Umsetzen ließ sich dies in Form eines wöchentlichen digitalen „SVG-JourFixe“, an dem alle Mitglieder der SVG und eine Vertreterin der Projektkoordination teilnahmen. Außerdem nahmen Vertreter:innen der SVG an den wöchentlichen Meetings der Projektkoordination teil. Der gegenseitige Informa-

tionsfluss wurde so gewährleistet und die Studierenden hatten die Möglichkeit, Feedback zu aktuellen Fragestellungen, Entwicklungen und Entscheidungen im Projekt PITCH zu geben und bei der Gestaltung mitzuwirken. Der Koordinationsaufwand zwischen den SVG-Mitgliedern war jedoch hoch und erforderte deren Eigeninitiative und auch Eigenverantwortung. Ebenfalls waren die Anforderungen an die Selbstorganisationsfähigkeiten der Einzelpersonen nicht zu unterschätzen.

Inhaltlich kam es der SVG zugute, dass ihre Mitglieder aus verschiedenen Fachrichtungen stammten und sich in unterschiedlichen Phasen ihres Studiums befanden. Gleichzeitig stellte sich die Unterschiedlichkeit der Mitglieder aber auch als Herausforderung für die Zusammenarbeit dar. Es zeigte sich, dass zwischen den verschiedenen Fachrichtungen Unterschiede in Bezug auf die Arbeitspraxis bestanden und auch die Erfahrungen mit dem wissenschaftlichen Arbeiten sehr unterschiedlich ausfielen. Die Schreibwerkstätten und eine engmaschige Betreuung durch die Projektkoordination waren zwingend notwendig, um die SVG-Mitglieder zu befähigen, im Projektkontext als Partner:innen auf Augenhöhe zu handeln. Auch wurden teilweise existierende Wissenslücken bspw. über das Hochschulsystem in Deutschland sowie Strukturen der Hochschullehre innerhalb der Hochschulen aufgedeckt und mussten geschlossen werden, damit das Feedback und die Empfehlungen der SVG über das Niveau der persönlichen Erfahrungen aus ihrer bisherigen (teils erst kurzen) Studienlaufbahn und ihren Präferenzen hinausgehen konnten. Aber auch ihre persönlichen Erfahrungen waren von großem Wert. Die SVG-Mitglieder haben aufgrund der Vielfalt ihrer Studienfächer sowohl traditionelle Prüfungen wie Klausuren als auch digitale und innovative Prüfungen abgelegt. In der Reflexion und Bewertung von Prüfungsformaten auf Basis der oben genannten Prüfungsevaluationen konnten sie somit maßgeblich beitragen.

Der bereits thematisierte Perspektivwechsel war aber auch der Rolle der SVG-Mitglieder immanent: In den Campusgesprächen konnten die SVG-Mitglieder mit ihren Kommiliton:innen in einer vertrauensvollen Atmosphäre über sensible und kritische Themen wie Diskriminierungserfahrungen und Täuschungsversuche sprechen. Aufgrund der Nähe zu den Studierenden öffneten sich die Gesprächspartner:innen und gaben für das PITCH-Projekt sehr wertvolle Informationen preis. Aber auch die SVG-Mitglieder berichteten über „Aha-Momente“. So wurde ihnen im Rahmen eines Workshops zum Thema Täuschungsversuche im Gespräch mit den Lehrenden klar, dass seitens der Prüfenden großer Frust auch auf zwischenmenschlicher Ebene entsteht, wenn sie vermehrte Täuschungsaktivitäten beobachten. Dieser Perspektivwechsel auf Seiten der Studierenden führte zu einem gesteigerten Verständnis für die Lehrenden und es entstand das Gefühl, im gleichen Boot zu sitzen, statt auf der Gegenseite zu stehen.

Als „Highlight“ ihrer Tätigkeit beschrieben die SVG-Mitglieder ihre Teilnahme an der TURN Conference 2023. Begleitet durch die Projektkoordination, gingen sie alle Schritte von der Einreichung ihres Beitrags, über die Planung der

An- und Abreise sowie der Vorbereitung und Durchführung ihrer Präsentation und des anschließenden Kurzworkshops.

4 Fazit

Die Partizipation von Studierenden im Projekt PITCH hat sich als gewinnbringend und anforderungsreich herausgestellt. Um ernst genommen zu werden, mussten die SVG-Mitglieder in die Lage versetzt werden, ihre Meinung (wissenschaftlich) fundiert vorzubringen. Gerade bei Studierenden in frühen Phasen des Studiums kann nicht vorausgesetzt werden, dass das wissenschaftliche Arbeiten souverän durchgeführt wird und vertiefte Kenntnisse über das Hochschulsystem und die Lehr- und Prüfungspraxis bereits vorliegen. Eine engmaschige Betreuung und Unterstützung sind hier notwendig. Auch der Koordinations- und Kommunikationsaufwand mit einer Gruppe studentischer Mitarbeiter:innen mit flexiblen Arbeitszeiten ist nicht zu unterschätzen. Dieser Ressourcenaufwand ist schwierig zu quantifizieren, aber bei der Planung zu berücksichtigen. Auch die SVG-Mitglieder beschreiben ihr Zusammenwachsen zu einem Team als langwierig. Es gelang ihnen aber schlussendlich, eine vertrauensvolle Zusammenarbeit und Kommunikation aufzubauen, in der sie ihre Aufgaben gemeinsam angehen konnten und sich auch auf persönlicher bzw. beruflicher Ebene weiterentwickeln konnten.

Das Projekt PITCH kommt somit zu dem Schluss, dass es sich lohnt, Studierende als Partner:innen an den Entwicklungen im Projekt PITCH teilhaben zu lassen. Allen Akteur:innen wurde ein Perspektivwechsel ermöglicht und wertvolle Einsichten über die Bedürfnisse, Möglichkeiten und Restriktionen der jeweils anderen Seite gegeben. Vermeintlich Selbstverständliches wurde expliziert und der Auftrag gestärkt, die Weiterentwicklung von (digitalen) Prüfungen gemeinsam voranzubringen.

Literatur

- Ackermann, L. (2021). Prüfungen in der Pandemie: Online-Proctoring ist keine Lösung. Hochschulforum Digitalisierung. <https://hochschulforumdigitalisierung.de/blog/pruefungen-in-der-pandemie-online-proctoring-ist-keine-loesung/>
- Beuß, M./Kroher, M./Becker, K./Ehrhardt, M.-C./Isleib, S./Koopmann, J./Steinkühler, J./Völk, D./Buchholz, S./Meyer, J./Multrus, F./Hinz, T./Marczuk, A./Strauß, S. (2022). Die Studierendenbefragung in Deutschland: Eine neue, integrierte Datenbasis für Forschung. In: Bildungs- und Hochschulpolitik. (DZHW Brief 06|2022). Hannover: DZHW. https://doi.org/10.34878/2022.06.dzhw_brief
- Bilen, E./Matros, A. (2020). Online cheating amid COVID-19. In: Journal of Economic Behavior & Organization, 182, 196 – 211. <https://doi.org/10.1016/j.jebo.2020.12.004>

- Bischoff, C. S./Ejrnæs, A./Rubin, O. (2021). A Quasi-experimental Study of Ethnic and Gender Bias in University Grading. In: PLoS ONE, 16(7). <https://doi.org/10.1371/journal.pone.0254422>
- Breiwe, R./Ebert, S. C. (2023). Zeitgemäße Hochschullehre in den spätmodernen Kulturen der Diversität und Digitalität!? (Selbst-)kritische Diskussion eines Praxisbeispiels. In: Kölner Online Journal für Lehrer*innenbildung, 282–301. <https://doi.org/10.18716/ojs/kON/2023.14>
- Ditzel, B./Bergt, T. (2013). Studentische Partizipation als organisationale Herausforderung – Ergebnisse einer explorativen Studie. In S. M. Weber/M. Göhlich/A. Schröer/C. Fahrwald/H. Macha (Hrsg.), Organisation und Partizipation. Beiträge der Kommission Organisationspädagogik (S. 177–186). https://doi.org/10.1007/978-3-658-00450-7_15
- Hauses, R. (2023). Vielfalt@UDE. Hochschulrektorenkonferenz. <https://www.hrk.de/themen/hochschulsystem/diversitaet/initiative-vielfalt-an-deutschen-hochschulen/karte-der-projektstandorte/universitaet-duisburg-essen/>
- International Center for Academic Integrity (o. J.). Facts and Statistics. <https://academicintegrity.org/resources/facts-and-statistics>
- Janke, S./Fritz, T./González Cruz, H./Rudert, S./Daumiller, M. (2023). Entwicklung und Überprüfung einer Kurzsкала zur Messung akademischen Betrugsverhaltens im Selbstbericht. In: Diagnostica, 69(4), 207–217. <https://doi.org/10.1026/0012-1924/a000317>
- Jantos, A./Böhm, C./Lauber, K. (2022). Betrugsversuche in der Hochschullehre: Einordnung der aktuellen Situation an der TU Dresden und abgeleitete Handlungsmöglichkeiten. In: Lessons Learned, 2(1). <https://doi.org/10.25369/ll.v2i1.48>
- Kergel, D./Heidkamp, B. (2019). Plädoyer für eine fluide Lehr-/Lernkultur: Hochschullehre zwischen Homogenisierung und Vielfalt. In D. Kergel & B. Heidkamp (Hrsg.), Prekarisierung und soziale Entkopplung – transdisziplinäre Studien. Praxishandbuch Habituussensibilität und Diversität in der Hochschullehre (S. 3–21). Springer VS. https://doi.org/10.1007/978-3-658-22400-4_1
- Kroher, M. (2020). Akademisches Fehlverhalten: Wie ehrlich berichten Studierende über Täuschungen? In I. Krumpal/R. Berger (Hrsg.), Kriminalität und Gesellschaft. Devianz und Subkulturen: Theorien, Methoden und empirische Befunde (S. 207–240). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-27228-9_8
- Malouff, J./Thorsteinsson, E. (2016). Bias in Grading: A Meta-analysis of Experimental Research Findings. In: Australian Journal of Education, 60(3), 245–256. <https://doi.org/10.1177/00049441166664618>
- Meyer, J./Strauß, S./Hinz, T. (2022). Die Studierendenbefragung in Deutschland: Fokusanalysen zu Diskriminierungserfahrungen an Hochschulen. DHZW Brief, Artikel 8. Vorab-Onlinepublikation. https://doi.org/10.34878/2022.08.DHZW_BRIEF
- Nalty, K. (2016). Strategies for Confronting Unconscious Bias. In: The Colorado Lawyer, 45, 45–52.
- Noorbehbahani, F./Mohammadi, A./Aminazadeh, M. (2022). A Systematic Review of Research on Cheating in Online Exams from 2010 to 2021. In: Education and Information Technologies, 27(6), 8413–8460. <https://doi.org/10.1007/s10639-022-10927-7>
- Persike, M. (2021). Digitales Prüfen: Didaktik, Umsetzung und Evidenz für die neue Prüfungsnormalität an Hochschulen. In I. Neiske/J. Osthusenrich/N. Schaper/U. Trier/ N. Vöing (Hrsg.), Hochschulbildung: Lehre und Forschung: Bd. 3. Hochschule auf Abstand: Ein multiperspektivischer Zugang zur digitalen Lehre (S. 327–353). Bielefeld: transcript.

- Pitt, E./Winstone, N. (2018). The Impact of Anonymous Marking on Students' Perceptions of Fairness, Feedback and Relationships with Lecturers. In: *Assessment & Evaluation in Higher Education*, 43(7), 1183–1193. <https://doi.org/10.1080/02602938.2018.1437594>
- Raffaële, C./Rediger, P. (2021). Die Partizipation Studierender als Kriterium der Qualitätssicherung in Studium und Lehre (HoF-Arbeitsbericht Nr. 117). Institut für Hochschulforschung an der Martin-Luther-Universität Halle-Wittenberg. https://www.hof.uni-halle.de/web/dateien/pdf/ab_117.pdf
- Resch, K./Raschauer, A. (2019). Kritische Ereignisse in der Hochschullehre: Ein Beitrag zur Weiterentwicklung von Diversitätskompetenz. In: *Die Hochschullehre*, 5, 115–134.
- Sather, A./Bovill, C./Felten, P. (2014). *Engaging Students as Partners in Learning and Teaching: A Guide for Faculty*. Jossey-Bass.
- Sattler, S./Diewald, M. (2013). FAIRUSE – Fehlverhalten und Betrug bei der Erbringung von Studienleistungen: Individuelle und organisatorisch-strukturelle Bedingungen: Schlussbericht nach Muster Nr. 3.2 zum Projekt; Projektlaufzeit: 01.04.2009 bis 30.04.2012. Technische Informationsbibliothek u. Universitätsbibliothek. <https://doi.org/10.2314/GBV:773897283>
- Schrader, S. (2023). Bedeutung und Potenziale studentischer Partizipation. In: *API Magazin*, 4(2), 1–16. <https://doi.org/10.15460/apimagazin.2023.4.2.155>
- Sedlacek, Q. (2021). Experimental Evidence of Disparities in Biology Teachers' Responses to Student Writing. In: *International Journal of Science Education*, 43(14), 2351–2373. <https://doi.org/10.1080/09500693.2021.1963499>
- Stephens, J. (2019). Natural and Normal, but Unethical and Evitable: The Epidemic of Academic Dishonesty and How We End It. In: *Change: The Magazine of Higher Learning*, 51(4), 8–17. <https://doi.org/10.1080/00091383.2019.1618140>
- Stone, A. (2023). Student Perceptions of Academic Integrity: A Qualitative Study of Understanding, Consequences, and Impact. In: *Journal of Academic Ethics*, 21(3), 357–375. <https://doi.org/10.1007/s10805-022-09461-5>
- Student:in #3 (Duisburg) (4. Mai 2023). Interview durch R. Hensle, und K. Schmitt.
- Student:in #4 (Essen) (27. April 2023). Interview durch S. Stotz Anido.
- Student:in #5 (Duisburg) (20. April 2023). Interview durch C. Corbach.
- Studierendengruppe #3 (Duisburg) (20. April 2023). Interview durch C. Corbach.
- Thompson, V. E./Vorbrugg, A. (2018). Rassismuskritik an der Hochschule: Mit oder trotz Diversity-Policies? In M. Laufenberg/M. Erlemann/M. Norkus/G. Petschick (Hrsg.), *Prekäre Gleichstellung: Geschlechtergerechtigkeit, soziale Ungleichheit und unsichere Arbeitsverhältnisse in der Wissenschaft* (S. 79–99). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-11631-6_4
- Waltzer, T./Bareket-Shavit, C./Dahl, A. (2023). Teaching the What, Why, and How of Academic Integrity: Naturalistic Evidence From College Classrooms. In: *Journal of College and Character*, 24(3), 261–284. <https://doi.org/10.1080/2194587X.2023.2224575>
- Zanga, G./Gioannis, E. de (2023). Discrimination in Grading: A Scoping Review of Studies on Teachers' Discrimination in School. In: *Studies in Educational Evaluation*, 78, 1–9. <https://doi.org/10.1016/j.stueduc.2023.101284>

Portraits der Autorinnen und Autoren

Dr. **Nicole Auferkorte-Michaelis** ist Geschäftsführerin des Zentrums für Hochschulqualitätsentwicklung der Universität Duisburg-Essen. Sie ist Erziehungswissenschaftlerin und Lehrende, hochschuldidaktische und schreibdidaktische Moderatorin. Ihre Arbeits- und Forschungsschwerpunkte sind die Konzeptentwicklung für Studium und Lehre und Institutional Research. Sie ist Mitglied im Netzwerk LehreN e. V. und leitet gemeinsam mit Prof. Dr. Frank Linde von der Technischen Hochschule Köln das hochschulübergreifende Zentrum für Kompetenzentwicklung für Diversity Management für Studium und Lehre www.komdim.de. nicole.auferkorte-michaelis@uni-due.de

Till Behmenburg ist wissenschaftlicher Mitarbeiter der Professur für Praxisorientierte Sozialwissenschaften am Institut für Politikwissenschaft der Universität Duisburg-Essen. Er verfügt über langjährige Erfahrungen in der Vorbereitung, Durchführung und Korrektur digitaler Prüfungen und befasst sich ansonsten schwerpunktmäßig mit Fragen der Politischen Theorie. till.behmenburg@uni-due.de

Prof. Dr. **Ulrike Berendt** ist Professorin für Praxisorientierte Sozialwissenschaften am Institut für Politikwissenschaft der Universität Duisburg-Essen. Schwerpunkte ihrer Lehre sind Policy-Analyse, Steuerungstheorien und -instrumente sowie das politische System. Ausgehend von langjährigen Erfahrungen mit digitalen Prüfungen ist sie in verschiedenen Funktionen auf Instituts- und Fakultäts-ebene auch mit Fragen der Prüfungsdigitalisierung befasst. ulrike.berendt@uni-due.de

Prof. Dr.-Ing. **Carolin Birk** leitet das Fachgebiet Statik und Dynamik der Tragwerke in der Abteilung Bauwissenschaften, Fakultät Ingenieurwissenschaften der Universität Duisburg-Essen. In der Forschung befasst sie sich mit nicht-klassischen Diskretisierungsverfahren für Aufgabenstellungen der Festkörpermechanik sowie für gekoppelte Probleme. Sie verfügt über umfangreiche Lehrerfahrung in Grundlagen- und Vertiefungsfächern des konstruktiven Ingenieurbaus. Im Projekt PITCH widmet sie sich der Entwicklung digitaler Prüfwerkzeuge für die Fächer Baustatik und Baudynamik. carolin.birk@uni-due.de

Maiken Bonnes ist wissenschaftliche Mitarbeiterin am Zentrum für Hochschulqualitätsentwicklung (ZHQE) der Universität Duisburg-Essen. Dort ist sie im Projekt PITCH in der Projektkoordination tätig. Darüber hinaus ist sie für die qualitative Lehrveranstaltungsevaluation zuständig und unterstützt Lehrende beim Einholen von studentischem Feedback für die Qualitätsentwicklung in der Lehre. maiken.bonnes@uni-due.de

Prof. Dr. **Birte Bös** ist Professorin für Englische Sprachwissenschaft am Institut für Anglophone Studien der Universität Duisburg-Essen. Neben ihren Forschungsaktivitäten im Bereich der angewandten Linguistik engagiert sie sich, u. a. als Leiterin des Projektteams Anglophone Studien im PITCH-Projekt, für eine studierendengerechte und kompetenzorientierte Ausgestaltung von Lehr-Lern- und Prüfungsszenarien.
birte.boes@uni-due.de

Prof. Dr. **Andreas Büchter** ist Professor für Didaktik der Mathematik an der Universität Duisburg-Essen. Im Rahmen des PITCH-Projekts arbeitet er zusammen mit Dana Eilers und Prof. Dr. Florian Schacht an der Weiterentwicklung des Lehramtsstudiums im Fach Mathematik. Im Fokus steht bei ihm die mathematikdidaktische Lehrveranstaltung „Lern- und kognitionspsychologische Grundlagen des Mathematikunterrichts“ im Studiengang für Gymnasien und Gesamtschulen bzw. Berufskollegs.
andreas.buechter@uni-due.de

Lisa Danzig ist wissenschaftliche Mitarbeiterin in der Didaktik für Physik an der Universität Duisburg-Essen. Sie arbeitet mit Prof. Dr. Heike Theyßen und Prof. Dr. Hermann Nienhaus im Teilprojekt Physik des PITCH-Projekts. Dort entwickelt sie hauptsächlich digitalen Hilfen zu experimentalphysikalischen Übungsaufgaben in der Studieneingangsphase Physik.
lisa.danzig@uni-due.de

Dana Eilers ist wissenschaftliche Mitarbeiterin in der Mathematikdidaktik an der Universität Duisburg-Essen. Im Rahmen des PITCH-Projekts entwickelt sie zusammen mit Prof. Dr. Andreas Büchter und Prof. Dr. Florian Schacht eine mathematische und eine mathematikdidaktische Veranstaltung weiter. Dabei beschäftigt sie sich unter anderem mit Innovationen des formativen und summativen Assessments, dem sinnvollen Einbezug digitaler Medien in der Lehre und der konstruktiven Nutzung von Learning Analytics.
dana.eilers@uni-due.de

Jana Eismann ist studentische Hilfskraft im Department of Anglophone Studies der Universität Duisburg-Essen. Innerhalb des Projekts PITCH unterstützt sie das Team um Prof. Dr. Birte Bös bei der Weiterentwicklung digitaler Lernangebote und Klausuren in der englischen Linguistik. Dabei bereichert sie das Team vor allem als Schnittstelle zwischen Lehrenden und Studierenden.
jana.eismann@stud.uni-due.de

Prof. Dr. **Michael Giese** ist Professor für Organische Chemie an der Universität Duisburg-Essen und forscht an Supramolekularen Materialien. Im Rahmen des PITCH-Projekts arbeitet er zusammen mit Dr. Katrin Schüßler und

Prof. Dr. Maik Walpuski an Möglichkeiten, Übungsaufgaben der organischen Chemie in digitale Formate zu überführen. Hier ist er insbesondere für fachspezifische Fragen verantwortlich sowie für die Testung der digitalen Aufgaben im Rahmen der Übung zu den Grundlagen der Organischen Chemie.

michael.giese@uni-due.de

Prof. Dr. **Jochen Gönsch** ist Professor für BWL, insb. Service Operations, an der Mercator School of Management der Universität Duisburg-Essen. Dort lehrt er u. a. zu den Themen Modellierung, Optimierung und Softwareeinsatz. Im Projekt PITCH werden zwei seiner Lehrveranstaltungen um digitale Begleit- und Prüfungsangebote ergänzt.

jochen.goensch@uni-due.de

Dr. **Sebastian Haase** ist wissenschaftlicher Mitarbeiter an der Freien Universität Berlin im Fachbereich Erziehungswissenschaft und Psychologie. Er entwickelt seit 2013 die Lehr-Lernplattform *tet.folio*. Vorher studierte er Physik und hat im Bereich der Biophysik mit der Entwicklung eines hochauflösenden Fluoreszenzmikroskops promoviert.

sebastian.haase@fu-berlin.de

Prof. Dr. **Christoph Hanck** ist seit August 2012 Professor für Ökonometrie an der Universität Duisburg-Essen. Nach seiner Promotion bei Prof. Dr. Walter Krämer an der TU Dortmund/RGS im Jahr 2007 war er zunächst im DFG-Sonderforschungsbereich „Komplexitätsreduktion in multivariaten Datenstrukturen“ tätig. 2008 wechselte er als Postdoktorand an die Universität Maastricht und war von 2009 bis 2012 erst Assistenzprofessor und später assoziierter Professor für Statistik und Ökonometrie an der Rijksuniversiteit Groningen. Seit 2014 widmet er sich zudem der Digitalisierung der Lehre und ist aktiv im Bereich Learning Analytics tätig.

christoph.hanck@vwl.uni-due.de

Prof. Dr. **Angela Heine** ist seit 2017 Professorin im Institut für Psychologie der Universität Duisburg-Essen. Sie leitet dort den Arbeitsbereich Pädagogische Psychologie mit Schwerpunkt Lern- und Leistungsstörungen des Kindes- und Jugendalters. Das von ihr bearbeitete PITCH-Teilprojekt widmete sich der kompetenzorientierten Umgestaltung der Lehre sowie der Digitalisierung der Prüfung im Modul Psychologie, das Teil des bildungswissenschaftlichen Bachelorstudiums aller Lehramtsstudierenden der UDE ist.

angela.heine@uni-due.de

Dr. **Patrick Hintze** ist Geschäftsführer des Instituts für wissenschaftliche Schlüsselkompetenzen an der Universität Duisburg-Essen und war bis Juni 2023 Mitglied der Projektkoordination von PITCH. Er hat Politikmanagement studiert,

über Wissenschaftspolitik promoviert und sich auf die Qualitätsentwicklung und das Wissenschaftsmanagement an Hochschulen spezialisiert. Seine Arbeitsschwerpunkte liegen in der Entwicklung von Lehr-Lernangeboten für Schlüsselkompetenzen und der Begleitung universitätsweiter Innovationsprozesse, Projekte und Strategien.

patrick.hintze@uni-due.de

Ruth M. Ingendoh (M.Sc.) war von 2020 bis 2024 als wissenschaftliche Mitarbeiterin am Arbeitsbereich Pädagogische Psychologie mit Schwerpunkt Lern- und Leistungsstörungen des Kindes- und Jugendalters der Universität Duisburg-Essen tätig. Ab Dezember 2021 wirkte sie im PITCH-Teilprojekt des Arbeitsbereichs mit. Im Fokus ihrer Tätigkeit stand die Konzeption und Gestaltung der digitalen Prüfung im Bachelormodul Psychologie der Lehramtsstudiengänge der UDE.
ruth.ingendoh@uni-due.de

Prof. Dr. **Christian Johannes** ist apl. Professor für Genetik in der Fakultät Biologie der Universität Duisburg-Essen. Im Projekt PITCH befasst er sich mit der Weiterentwicklung biologiebezogener digitaler Lern- und Prüfungsformate. Sein genetischer Forschungsschwerpunkt liegt auf der Untersuchung von strahleninduzierten Chromosomenschäden und den Einsatzmöglichkeiten der Chromosomenanalyse in der Biologischen Dosimetrie.

christian.johannes@uni-due.de

Dr. **Cornelia Kenneweg** ist freie Hochschuldidaktikerin und Lehrentwicklerin in Leipzig. Sie berät Lehrende und hochschulische Einrichtungen und bietet als Trainerin hochschuldidaktische Workshops an, insbesondere zu fachbezogenen Ansätzen der Hochschuldidaktik und zur Verbindung von Forschung und Lehre.
kenneweg@agentur fuer lehrkultur.de

Jens Klenke ist wissenschaftlicher Mitarbeiter am Lehrstuhl für Ökonometrie der Universität Duisburg-Essen. Seinen Bachelor in Staatswissenschaften absolvierte er 2017 an der Universität Erfurt, gefolgt von einem Master in Volkswirtschaftslehre, den er 2022 an der Universität Duisburg-Essen abschloss. Im Rahmen des PITCH-Projekts ist er am Lehrstuhl unter anderem für die Digitalisierung der Lehre und die Umsetzung von Learning Analytics verantwortlich.
jens.klenke@vwl.uni-due.de

Martin Krybus ist als PITCH-Projektmitarbeiter an der Professur für Praxisorientierte Sozialwissenschaften am Institut für Politikwissenschaft der Universität Duisburg-Essen tätig. Er verfügt über langjährige Erfahrungen in der Vorbereitung, Durchführung und Korrektur digitaler Prüfungen und widmet sich ferner dem Feld der Parteienforschung.

martin.krybus@uni-due.de

Dr. **Johanna Lambrich** ist wissenschaftliche Mitarbeiterin am Institut für Sport- und Bewegungswissenschaft der Universität Duisburg-Essen und dort im Projekt PITCH tätig, in dem sie an der Entwicklung und Umsetzung kompetenzorientierter Prüfungsformate mitwirkt. Darüber hinaus beschäftigt sie sich in ihrer Forschung insbesondere mit der Analyse plantarer Druckverteilungen im Tennissport und Fragestellungen rund um das Gleichgewicht.
johanna.lambrich@uni-due.de

Dr. **Julia Liebscher** ist wissenschaftliche Mitarbeiterin am Zentrum für Hochschulqualitätsentwicklung (ZHQE) der Universität Duisburg-Essen. Dort ist sie im Projekt PITCH in der Projektkoordination tätig.
julia.liebscher@uni-due.de

Nora Lösing ist wissenschaftliche Hilfskraft am Biology Education Research and Learning Lab (BERLL) der Universität Duisburg-Essen. Sie unterstützt das BERLL bei der Vor- und Nachbereitung eines Biologiedidaktik-Moduls in den Bachelor-Lehramtsstudiengängen Biologie, unter anderem durch die (Weiter-)Entwicklung von Übungsaufgaben und Rückmeldungen für Studierende.
nora.loesing@stud.uni-due.de

Dr. **Mathias Magdowski** hat an der Otto-von-Guericke-Universität in Magdeburg Elektrotechnik studiert und promoviert. Er arbeitet dort als wissenschaftlicher Mitarbeiter am Lehrstuhl für elektromagnetische Verträglichkeit und beschäftigt sich gern mit alternativen Lehr- und Lernmethoden. Außerdem koordiniert er eine universitätsweite E-Learning-Arbeitsgruppe, engagiert sich in der MINT-Sensibilisierung von Kindern und Jugendlichen und übt sich als Wissenschaftskommunikator im Science Slam.
mathias.magdowski@ovgu.de

Frauke Milne ist wissenschaftliche Mitarbeiterin am Institut für Anglophone Studien der Universität Duisburg-Essen. Dort ist sie neben ihrer Promotion im Projekt PITCH tätig und kümmert sich unter anderem um die Planung, Erstellung und Weiterentwicklung digitaler Klausuren sowie die Betreuung von online-gestützten Prüfungsplattformen in der englischen Sprachwissenschaft.
frauke.milne@uni-due.de

Prof. Dr. **Thomas Mühlbauer** leitet die Professur für Bewegungs- und Trainingswissenschaft/Biomechanik des Sports am Institut für Sport- und Bewegungswissenschaften der Universität Duisburg-Essen. Dort ist er an der Ausbildung von Studierenden in Bachelor- und Masterstudiengängen für alle Lehramtsoptionen beteiligt und verantwortet Forschungsarbeiten zu den Themen Kontrolle und Lernen von Bewegungen sowie Diagnostik und Training von insbesondere Gleichgewicht und Kraft in der Lebensspanne.
thomas.muehlbauer@uni-due.de

apl. Prof. Dr. **Hermann Nienhaus** ist Professor in der Experimentalphysik der Universität Duisburg-Essen. Er ist seit vielen Jahren in der Lehre der Fachphysik innerhalb der Lehramtsstudiengänge tätig. Im PITCH-Projekt leitet er zusammen mit Prof. Dr. Heike Theyßen das Teilprojekt Physik.
hermann.nienhaus@uni-due.de

Henrik Noll ist Volljurist und war von Oktober 2022 bis zum Dezember 2024 im Justitiariat der Universität Duisburg-Essen beschäftigt. Dort war er für die rechtliche Betreuung des Projekts PITCH zuständig und hat sich im Rahmen dessen intensiv mit juristischen Fragestellungen zum digitalen Prüfungsrecht auseinandergesetzt.

Christoph Olbricht ist wissenschaftlicher Mitarbeiter in der Arbeitsgruppe Spezifikation von Softwaresystemen des paluno-Instituts der Universität Duisburg-Essen. Dort ist er im Projekt PITCH für die Schulung und Kundenbetreuung für JACK3 tätig. Darüber hinaus ist er für die Entwicklung von JACK3 und die Verwaltung der virtuellen Serverinfrastruktur von JACK3 zuständig.
christoph.olbricht@uni-due.de

Ajay Kumar Pasupuleti ist wissenschaftlicher Mitarbeiter am Fachgebiet Statik und Dynamik der Tragwerke in der Abteilung Bauwissenschaften, Fakultät Ingenieurwissenschaften der Universität Duisburg-Essen. Dort war er neben der Forschung im Bereich Computational Mechanics bis zum 31.07.2024 im Projekt PITCH tätig. Hierbei entwickelte er insbesondere Prüfungs- und Übungsaufgaben für das Fach Baudynamik mittels des Übungs- und Prüfungssystems JACK.
ajay.pasupuleti@uni-due.de

Nguyen Minh Salzmänn-Hoang war bis 07/2024 wissenschaftlicher Mitarbeiter am Institut für Deutsch als Zweit- und Fremdsprache der Universität Duisburg-Essen und arbeitete im Projekt PITCH an der Modulprüfung des Lehramtsstudiums (BA). Seit 10/2024 ist er wissenschaftlicher Mitarbeiter am Herder-Institut für Deutsch als Fremd- und Zweitsprache der Universität Leipzig. Seine Arbeitsschwerpunkte sind Migrationspädagogik, postkoloniale Sprachverhältnisse und Empowerment.
minh.salzmänn-hoang@uni-leipzig.de

Prof. Dr. **Florian Schacht** ist Professor für Didaktik der Mathematik an der Universität Duisburg-Essen. Im Rahmen des PITCH-Projekts arbeitet er zusammen mit Dana Eilers und Prof. Dr. Andreas Büchter an der Weiterentwicklung des Lehramtsstudiums im Fach Mathematik, insbesondere im Zusammenhang mit der Entwicklung digitaler Lehr-Lern-Materialien für die Hochschullehre. Im Rahmen von PITCH wurde das Veranstaltungskonzept für die Grundvorlesung Elementare Geometrie in den Mathematik-Lehramtsstudiengängen Grundschule, Haupt-,

Real-, Gesamt- und Sekundarschule sowie Sonderpädagogik weiterentwickelt.
florian.schacht@uni-due.de

Prof. Dr. **Philipp Schmiemann** ist Professor für Didaktik der Biologie am Biology Education Research and Learning Lab (BERLL) der Universität Duisburg-Essen. Im Projekt PITCH befasst er sich mit der Weiterentwicklung biologiebezogener digitaler Lern- und Prüfungsformate. Daneben forscht er u. a. zum Systemischen Denken und phylogenetischen Stammbäumen.
philipp.schmiemann@uni-due.de

Kilian Schmitt war im Rahmen des Projekts PITCH bis zum 30.04.2025 als studentische Hilfskraft am Zentrum für Hochschulqualitätsentwicklung (ZHQE) tätig. Als Teil der Student Voice Group (SVG) repräsentiert er innerhalb des Projekts die Interessen und Perspektiven der Studierenden in Bezug auf Prüfungen.
kilian.schmitt@stud.uni-due.de

Prof. Dr. **Tobias Schroedler** leitet die Arbeitsgruppe „Mehrsprachigkeit und gesellschaftliche Teilhabe“ am Institut für Deutsch als Zweit- und Fremdsprache in der Fakultät für Geisteswissenschaften der Universität Duisburg-Essen. Zu seinen Forschungsschwerpunkten zählen Mehrsprachigkeit im Bildungswesen und institutionelle Mehrsprachigkeit. Im Projekt PITCH verantwortet er das Teilprojekt DaZ.
tobias.schroedler@uni-due.de

Dr. **Katrin Schüßler** ist wissenschaftliche Mitarbeiterin der Chemiedidaktik an der Universität Duisburg-Essen. Im PITCH-Projekt ist sie gemeinsam mit Prof. Dr. Maik Walpuski für chemiedidaktische Aspekte bei der Digitalisierung von Übungsaufgaben in der organischen Chemie verantwortlich. Ein Schwerpunkt der Arbeit liegt in der Auseinandersetzung mit traditionellen Aufgabenformaten der organischen Chemie, gemeinsam mit Prof. Dr. Michael Giese, und der digitalen Umsetzung dieser Aufgaben unter Beibehaltung der fachspezifischen Charakteristika und Anforderungen.
katrin.schuessler@uni-due.de

Melanie Schypula ist wissenschaftliche Mitarbeiterin beim Zentrum für Informations- und Mediendienste (ZIM) an der Universität Duisburg-Essen. Dort war sie bis zum 31.12.2024 im Projekt PITCH für die Schulung und Kundenbetreuung für das LMS Moodle tätig. Darüber hinaus ist sie an der Entwicklung des e-Assessment Systems JACK3 beteiligt.
melanie.schypula@uni-due.de

Marcellus Sieburg ist wissenschaftlicher Mitarbeiter am Lehrstuhl Formale Methoden der Informatik an der Universität Duisburg-Essen. Im Rahmen des Projekts PITCH erforscht er die automatische Generierung, Bewertung und

Erzeugung von Feedback für E-Learning-Aufgaben. Dabei setzt er solche Aufgaben in der Lehrveranstaltung Modellierung ein, für die er im Rahmen seiner Lehrverpflichtung Präsenzlehre übernimmt.

marcellus.siegburg@uni-due.de

Prof. Dr. **Michael Striwe** ist Professor für Softwaretechnik und Bildungstechnologie an der Hochschule Trier. Er forscht unter anderem zu automatischen Verfahren für die Generierung von Aufgaben und kompetenzorientiertem Feedback und ist Sprecher des Arbeitskreises E-Assessment der Fachgruppe Bildungstechnologien in der Gesellschaft für Informatik e. V.. Bis September 2024 leitete er an der Universität Duisburg-Essen die Entwicklung des E-Assessment-Systems JACK.

m.striewe@inf.hochschule-trier.de

Prof. Dr. **Heike Theyßen** ist Professorin für Didaktik der Physik an der Universität Duisburg-Essen. Sie forscht u. a. zur Studieneingangsphase Physik und zu individueller Förderung beim Lernen von Physik. Im PITCH-Projekt leitet sie zusammen mit Prof. Dr. Hermann Nienhaus das Teilprojekt Physik.

heike.theysen@uni-due.de

Justin Timm ist wissenschaftlicher Mitarbeiter am Biology Education Research and Learning Lab (BERLL) der Universität Duisburg-Essen. Im Projekt PITCH arbeitete er bis zum 31.07.2024 an der Weiterentwicklung digitaler Lern- und Prüfungsformate für biologiedidaktische Themen. Darüber hinaus forscht er zum systemischen Denken und dem Vorgehen von Schüler:innen bei der Analyse humangenetischer Stammbäume.

justin.timm@uni-due.de

Prof. Dr. **Janis Voigtländer** ist Professor für Formale Methoden der Informatik an der Universität Duisburg-Essen. Im Rahmen des Projekts PITCH erprobt er den Einsatz von E-Learning-Aufgaben in zwei seiner Vorlesungs-Module, hauptsächlich der Lehrveranstaltung Modellierung, die in drei verschiedenen informatikbezogenen Bachelor-Studiengängen zum Pflichtprogramm gehört.

janis.voigtlaender@uni-due.de

Prof. Dr. **Maik Walpuski** ist Chemiedidaktiker an der Universität Duisburg-Essen. Einer seiner Forschungsschwerpunkte ist die Leistungsmessung im Fach Chemie. Im Projekt PITCH ist er gemeinsam mit Dr. Katrin Schübler für chemiedidaktische Aspekte bei der Digitalisierung von Übungsaufgaben in der organischen Chemie verantwortlich.

maik.walpuski@uni-due.de

Prüfungen digital gestalten

Hochschulen stehen vor erheblichen Herausforderungen, wenn es um digitales Prüfen geht. Die Pandemie und die Diskussion um die Nutzung künstlicher Intelligenz in Lehre und Studium haben verdeutlicht, dass es an didaktischen, technischen, organisatorischen und rechtlichen Lösungen mangelt, um Prüfungen digital und gleichzeitig kompetenzorientiert umzusetzen. Die Autor*innen geben fächerspezifische Einblicke und zeigen Lösungsansätze für diese Herausforderungen auf.

Die Herausgeber*innen:

Dr. Nicole Auferkorte-Michaelis, Geschäftsführung, Team- und Projektleitung, Zentrum für Hochschulqualitätsentwicklung, Universität Duisburg-Essen

Maiken Bonnes, Wissenschaftliche Mitarbeiterin, Zentrum für Hochschulqualitätsentwicklung, Universität Duisburg-Essen

Dr. Patrick Hintze, Geschäftsführung, Institut für wissenschaftliche Schlüsselkompetenzen, Universität Duisburg-Essen

Dr. Julia Liebscher, Wissenschaftliche Mitarbeiterin, Zentrum für Hochschulqualitätsentwicklung, Universität Duisburg-Essen

ISBN 978-3-8474-3098-8



www.budrich.de

Titelbildnachweis: Arthur Lambillotte @ unsplash