

# Beiträge zur HOCHSCHULFORSCHUNG

4 | 2018

Thema: Digitalisierung der Hochschulen

Franzen: Die digitale Transformation der Wissenschaft

Gottburgsen/Willige: Mehr Mobilitätserfahrungen durch digitale Medien?

Förster/Heiß/Klinke/Maur/Schank/Weiser:  
Implementation und Evaluation eines Flipped Classrooms

Riedel/Möbius: E-Assessment an sächsischen Hochschulen

Wolff-Grosser: Online-Self-Assessments an der Technischen Hochschule Nürnberg



# Beiträge zur HOCHSCHULFORSCHUNG

## 4 | 2018

Thema: Digitalisierung der Hochschulen

Franzen: Die digitale Transformation der Wissenschaft

Gottburgsen/Willige: Mehr Mobilitätserfahrungen durch digitale Medien?

Förster/Heiß/Klinke/Maur/Schank/Weiser:  
Implementation und Evaluation eines Flipped Classrooms

Riedel/Möbius: E-Assessment an sächsischen Hochschulen

Wolff-Grosser: Online-Self-Assessments an der Technischen Hochschule Nürnberg

# Impressum

## **Beiträge zur Hochschulforschung**

erscheinen viermal im Jahr

ISSN (Print) 0171-645X

ISSN (Online) 2567-8841

**Herausgeber:** Bayerisches Staatsinstitut für Hochschulforschung  
und Hochschulplanung, Prinzregentenstraße 24, 80538 München  
Tel.: 089/2 1234-405, Fax: 089/2 1234-450

E-Mail: [Sekretariat@ihf.bayern.de](mailto:Sekretariat@ihf.bayern.de)

Internet: <http://www.ihf.bayern.de>

### **Herausgeberbeirat:**

Prof. Dr. Katrin Auspurg (Ludwig-Maximilians-Universität München)

Dr. Dr. Lutz Bornmann (Max-Planck-Gesellschaft)

Prof. Dr. Monika Jungbauer-Gans (Universität Hannover, DZHW)

Prof. Dr. Georg Krücken (Universität Kassel, INCHER)

Prof. Dr. Isabell Welpke (Technische Universität München, IHF)

Dr. Lydia Hartwig (IHF)

**Redaktion:** Dr. Lydia Hartwig (V.i.S.d.P.)

Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung

E-Mail: [Hartwig@ihf.bayern.de](mailto:Hartwig@ihf.bayern.de)

Die abgedruckten Beiträge geben die Meinung der Verfasser wieder.

**Graphische Gestaltung:** Haak & Nakat, München

**Satz:** Dr. Ulrich Scharmer, München

**Druck:** Steinmeier, Deiningen

# Ausrichtung, Themenspektrum und Zielgruppen

Die „Beiträge zur Hochschulforschung“ sind eine der führenden wissenschaftlichen Zeitschriften im Bereich der Hochschulforschung im deutschen Sprachraum. Sie zeichnen sich durch hohe Qualitätsstandards, ein breites Themenspektrum und eine große Reichweite aus. Kennzeichnend sind zudem die Verbindung von Wissenschaftlichkeit und Relevanz für die Praxis sowie die Vielfalt der Disziplinen und Zugänge. Dabei können die „Beiträge“ auf eine lange Tradition zurückblicken. Die Zeitschrift erscheint seit ihrer Gründung 1979 viermal im Jahr und publiziert Artikel zu Veränderungen in Universitäten, Fachhochschulen und anderen Einrichtungen des tertiären Bildungsbereichs sowie Entwicklungen in Hochschul- und Wissenschaftspolitik in nationaler und internationaler Perspektive.

Wichtige Themenbereiche sind:

- Governance von Hochschulen und Forschungseinrichtungen,
- Steuerung und Optimierung von Hochschulprozessen,
- Hochschulfinanzierung,
- Qualitätssicherung und Leistungsmessung,
- Studium und Studierende, Umsetzung des Bologna-Prozesses,
- Übergänge zwischen Schule, Hochschule und Arbeitsmarkt,
- Forschung und wissenschaftlicher Nachwuchs, akademische Karrieren,
- Geschlechterungleichheiten in der Wissenschaft,
- Wissenschaft und Wirtschaft,
- International vergleichende Hochschulforschung,
- Wissenschaftsforschung.

Die Zeitschrift veröffentlicht quantitative und qualitative empirische Analysen, Vergleichsstudien, Überblicksartikel und Einblicke in die Praxis, die ein anonymes Peer Review-Verfahren (double blind) durchlaufen haben. Sie bietet die Möglichkeit zum Austausch von Forschungsergebnissen und stellt ein Forum für Hochschulforscher und Experten aus der Praxis dar. Neben Ausgaben, die das gesamte Spektrum der Hochschulforschung abbilden, erscheinen in regelmäßigen Abständen Themenhefte. Hierfür erfolgt in der Regel ein Call for Papers. Manuskripte können jederzeit in deutscher und englischer Sprache eingereicht werden.

Die „Beiträge“ richten sich an Wissenschaftler, die sich mit Fragen des Hochschulwesens und seiner Entwicklung befassen, aber auch an politische Entscheidungsträger, Hochschulleitungen, Mitarbeiter in Hochschulverwaltungen, Ministerien sowie Wissenschafts- und Hochschulorganisationen.

Alle Ausgaben der „Beiträge zur Hochschulforschung“ erscheinen in gedruckter Form und werden auf der Homepage unter [www.bzh.bayern.de](http://www.bzh.bayern.de) veröffentlicht, die einzelnen Artikel sind nach verschiedenen Kategorien recherchierbar.



# Inhalt

Editorial	4
Abstracts	6
<b>Überblick</b>	
Martina Franzen: Die digitale Transformation der Wissenschaft	8
<b>Forschungsartikel</b>	
Anja Gottburgsen, Janka Willige: Mehr Mobilitätserfahrungen durch digitale Medien? Zu den Effekten von studentischer Diversität und Lernumweltsmerkmalen auf die internationale Mobilität	30
Manuel Förster, Florian Heiß, Sigbert Klinke, Andreas Maur, Thorsten Schank, Constantin Weiser: Die Implementation und Evaluation eines Flipped Classrooms in einer Großveranstaltung der Statistik	50
Jana Riedel, Kathrin Möbius: Bestandsaufnahme, Hindernisse und Möglichkeiten des Einsatzes von E-Assessment an sächsischen Hochschulen	68
<b>Einblicke in die Praxis</b>	
Diana Wolff-Grosser: Online-Self-Assessments an der Technischen Hochschule Nürnberg – Bilanzierung, Reflexion, Zukunftsvision	88
Verleihung des Ulrich Teichler-Preises und des Nachwuchspreises der Gesellschaft für Hochschulforschung – Abstracts der preisgekrönten Arbeiten	106
Buchvorstellungen	112
Jahresindex 2018	114
Index 2018	116
Hinweise für Autorinnen und Autoren	119

# Editorial

Die Digitalisierung beeinflusst mittlerweile nahezu alle gesellschaftlichen Bereiche. Der Einsatz digitaler Technologien und damit verbundener neuer Medien erzeugt eine Dynamik, die das Gros sowohl der Kommunikations- und Interaktions- als auch der Produktionsprozesse in der Arbeits- und Lebenswelt verändert. Die auch als *digitale Revolution* charakterisierte Transformation betrifft nicht zuletzt das Hochschulsystem. Weder Forschung und Lehre noch Administration sind heute an den Hochschulen ohne digitale Infrastrukturen und Operationen denkbar. Wie aber werden die medialen und technologischen Veränderungsimpulse im Hochschul- und Wissenschaftssystem aufgegriffen und bearbeitet? Welche Auswirkungen ergeben sich für Infrastrukturbedarfe, Arbeitsprozesse, Kontrollmechanismen oder Leistungsdimensionen in Forschung, Lehre und Administration? Welche Strukturen bilden Hochschulen und Einrichtungen der außeruniversitären Forschung aus, um die Chancen der Digitalisierung zu nutzen und ihre Risiken beherrschbar zu machen?

Um Gelegenheit zur Standortbestimmung und wissenschaftlichen Diskussion zu geben, befasste sich die 12. Jahrestagung der Gesellschaft für Hochschulforschung, die vom 30. bis 31. März 2017 an der Leibniz Universität Hannover stattfand, unter dem Titel „Digitalisierung der Hochschulen: Forschung, Lehre und Administration“, mit diesem Umbruch. In diesem Heft werden ausgewählte Beiträge veröffentlicht, die alle ein externes Begutachtungsverfahren durchlaufen haben.

Mit Fragen der „Digitalen Transformation der Wissenschaft“ befasst sich der Überblicksartikel von *Martina Franzen*, der die erweiterte Fassung ihrer Keynote der Jahrestagung darstellt. Martina Franzen geht davon aus, dass die gesellschaftliche Umstellung auf digitale Verbreitungstechnologien für die Wissenschaft folgenreiche Implikationen hat, da sich nicht nur wissenschaftliche Publikationsformen pluralisieren, sondern neuartige Rezeptions-, Bewertungs- und Produktionspraktiken entstehen. Die digitalen Produktionspraktiken wiederum verändern, so die These, auch die Rolle von Wissenschaft in der Gesellschaft.

Mit der Frage, ob durch digitale Medien auch mehr Mobilitätserfahrungen gemacht werden, versuchen *Anja Gottburgsen* und *Janka Willige* die bislang unverbundenen Forschungsstränge *Diversität* und *Digitalisierung* zu verbinden. Auf Grundlage eines theoretisch durch die beiden Forschungsstränge entwickelten analytischen Modells können die Autorinnen auf der Basis von Daten des DZHW-Online-Access-Panel HISBUS zum Teil deutliche Effekte sozialer und lernumweltlicher Merkmale auf die internationale Mobilität aufzeigen.

Um ein immer noch sehr spezifisches lernumweltliches Merkmal, den Flipped Classroom, geht es im darauf folgenden Beitrag von *Manuel Förster, Florian Heiß, Sigbert Klinke, Andreas Maur, Thorsten Schank* und *Constantin Weiser*. In dem Artikel wird das theoretisch fundierte Design eines Flipped Classroom als ein Lösungsansatz zum Umgang mit Herausforderungen einer Großveranstaltung zur Statistik entwickelt. Zudem wird ein an empirischen und methodischen Forschungslücken anknüpfendes Evaluationskonzept vorgestellt, das die kognitive und motivationale Entwicklung sowie den Workload der Studierenden im Längsschnitt in der klassischen und geflippten Veranstaltung berücksichtigt.

Mit Erkenntnissen des Einsatzes von E-Assessments beschäftigen sich schließlich gleich zwei Beiträge. Im Mittelpunkt des Beitrags von *Jana Riedel* und *Kathrin Möbius* stehen Ergebnisse einer Online-Befragung sächsischer Hochschulen und von zwei Fokusgruppeninterviews zur Verbreitung und zu den Hindernissen des Einsatzes von E-Assessments sowie zur Einschätzung von Chancen, Visionen und Limitationen. Zielsetzung ist die Qualitätsverbesserung von Lehren und Prüfen.

Eine Bilanzierung und Reflexion des neunjährigen Einsatzes von Online-Self-Assessments an der Technischen Hochschule Nürnberg nimmt *Diana Wolff-Grosser* vor. Die Assessments dienen der Feststellung von Studierfähigkeit und Passung einerseits sowie der Reduzierung von Studienabbruch andererseits. Diana Wolff-Grosser hält das Potenzial der Self-Assessments hinsichtlich ihrer prognostischen Validität zur Vorhersage von Studienabbrüchen für bei weitem noch nicht ausgeschöpft. Sie hebt insbesondere das Ziel der Verbesserung der Validität von Zulassungsverfahren sowie die Chance der Entwicklung diversitätsgerechter Unterstützungsmaßnahmen hervor.

Neben den in diesem Heft veröffentlichten Aufsätzen stehen zahlreiche Präsentation der 12. Jahrestagung unter <https://www.gfhf.net/aktivitaeten/jahrestagungen/12-jahrestagung-2017/> open access zur Verfügung. Insgesamt ist zu konstatieren, dass die Forschung zur Digitalisierung von Hochschulen sehr interdisziplinär angelegt ist und noch viel Potenzial für vertiefende Analysen bietet.

Auf der Jahrestagung 2017 wurden erneut drei Nachwuchswissenschaftlerinnen und -wissenschaftler für ihre Leistungen auf dem Gebiet der Hochschulforschung mit dem Ulrich Teichler-Preis ausgezeichnet. Die „Beiträge zur Hochschulforschung“ veröffentlichen die Abstracts der Dissertationen von Heike Spangenberg zur Entwicklung der Studienentscheidungen und nachschulischen Bildungsverläufe seit 1990 in Ost- und Westdeutschland und von Christian Schmid zur managerialen Governance akademischer Lehrtätigkeit sowie der Masterarbeit von Klara Reichenbach zu Determinanten der Teilnahmemotivation an studentischer Lehrveranstaltungsevaluation.

*Walburga Freitag, Lydia Hartwig, Monika Jungbauer-Gans, Bernd Kleimann*

# Abstracts

## **Martina Franzen: The digital transformation of science**

Transition to digital distribution technology has many implications for science. With increasing digitalisation, not only scientific forms of publication are pluralising; new forms of reception, evaluation and production practices are emerging. Web-based infrastructures promote collaborative forms of knowledge production, which now also include non-scientific actors (crowd science), and big data enable novel recombination of knowledge. Looking at these development dynamics together, the question arises to what extent digitalisation does not only change the production of knowledge, but also the role of science in society.

## **Anja Gottburgsen, Janka Willige: More mobility through digitalisation? On the effects of students' diversity and the characteristics of the learning environment on study-related international mobility**

Despite high expectations for the inclusion and learning success of a diverse student body through digitisation, the two research strands „diversity“ and „digitisation“ have so far remained unconnected. In the present article, „diversity“ and „digitisation“ are merged in a theoretical model for empirical investigation. Through a secondary analysis of survey data from the German Centre for Higher Education Research and Science Studies (DZHW) online access panel „HISBUS student panel“ (n = 4,375) the effects both of the students' diversity and the characteristics of the learning environment on study-related international mobility (digital, actual) of students are analysed. Results demonstrate effects on „international mobility“ due to social attributes of the students (e.g., gender, non-academic background or parenthood) and characteristics of the learning environment (e.g., type of institute, field of study, digitalized learning environment).

## **Manuel Förster, Florian Heiß, Sigbert Klinke, Andreas Maur, Thorsten Schank, Constantin Weiser: The implementation and evaluation of a flipped classroom within a large statistics lecture**

Particularly in statistics and mathematics, students' motivational and cognitive characteristics vary significantly depending on sociodemographic characteristics, such as gender, and previous knowledge. In this regard, the limited and mostly undifferentiated possibilities of traditional teaching in large statistics lectures often lead to further negative developments in students' cognitive and motivational learning conditions. With a flipped classroom design, this paper introduces an approach to meet these challenges along with its theoretical foundations and the underlying evaluation concept. Building on existing methodological gaps in research, students' academic achievement

and its sustainability, this assessment framework considers the longitudinal development of students' academic achievement and its sustainability, motivation and self-concept in statistics and their workload in both a traditional and in a flipped classroom lecture.

### **Jana Riedel, Kathrin Möbius: Current state and possibilities to use e-assessment at Saxon universities**

The paper presents the results of a survey identifying the current state and obstacles of e-assessment scenarios at Saxon universities. It shows that e-assessments are currently not widespread at Saxon universities and that they are mainly used by lecturers for exercise purposes (formative assessment) and only rarely for examination purposes (summative assessment). In addition to discipline-specific differences in the use of the different scenarios, there are also differences in regard to the amount of participants taking part in the courses. Electronic tests as one of the four e-assessment scenarios examined are used primarily in sciences and in large lectures. The primary goal of university lecturers is to increase the quality of teaching. Potential for development is seen in the didactic integration of examination and teaching activities and in the possibilities of individualising learning processes. Many teachers see obstacles, especially for summative e-assessment, in the unclear legal situation and in the lack of infrastructure. The study results allow the conclusion that an improvement in the framework conditions can contribute to a further spread of e-assessment scenarios.

### **Diana Wolff-Grosser: Online-Self-Assessments at the Technische Hochschule Nürnberg – conclusions, reflection, perspectives**

Having implemented Online Self-Assessments for nearly nine years in an increasing amount of degree courses, the Technische Hochschule Nürnberg has considerable experience in different kinds of hosting, structures and test-taking options for psychological tests. In order to ensure quality and processes, it has turned out to be crucial to host an own portal, to consider psychological expertise and to continuously refine the test to adapt to technical and social change. Unfortunately, current legislation in Bavaria impedes exploiting the Online Self-Assessments full potential with regard to its prognostic validity concerning academic drop-out rates, which would make substantial amendment desirable.

# Die digitale Transformation der Wissenschaft<sup>1</sup>

Martina Franzen

---

Die gesellschaftsweite Umstellung auf digitale Verbreitungstechnologie hat für die Wissenschaft folgenreiche Implikationen. Mit der zunehmenden Digitalisierung pluralisieren sich nicht nur die wissenschaftlichen Publikationsformen, sondern es entstehen neuartige Rezeptions-, Bewertungs- und Produktionspraktiken in der Wissenschaft. Webbasierte Infrastrukturen befördern kollaborative Formen der Wissensproduktion, die inzwischen auch außerwissenschaftliche Akteure miteinschließen (Stichwort: Crowd Science) und im Zuge von Big Data neuartige Rekombinationen von Wissen ermöglichen. Fasst man diese Entwicklungsdynamiken zusammen, stellt sich die Frage, inwiefern sich mit dem digitalen Wandel nicht nur die Produktion von Wissen verändert, sondern auch die Rolle von Wissenschaft in der Gesellschaft.

---

## 1 Einleitung

*„Wenn sich in allen Punkten die neuen Ideale gegen die alten durchsetzen würden, dann hätten wir es tatsächlich mit einer paradigmatisch anderen Form von Wissensgenerierung und -zirkulation zu tun, bei der es sich lohnt, darüber nachzudenken, ob der Begriff Wissenschaft dann überhaupt noch tauglich und die Universität noch der richtige Ort wäre, solche Kulturtechniken zu vermitteln.“*  
(Hagner & Hirschi 2013: 8)

Mit den Schlagworten *Open Science* oder *Science 2.0* werden aktuelle Entwicklungen und Bestrebungen einer Transformation der Wissenschaft durch Digitalisierung bezeichnet (Fecher & Friesike 2014, Europäische Kommission 2014). Bereits die Existenz gleich zwei so unterschiedlicher Begriffe signalisiert, dass es sich bei der Digitalisierung um einen soziotechnischen Prozess handelt, auch wenn im öffentlichen Diskurs entweder die soziale Seite (*Open Science*) oder die technische Seite (*Science 2.0*) stärker betont wird. Während die Debatte um die Digitalisierung von Wissenschaft lange Zeit auf den Aspekt von *Open Access* verkürzt war, setzt sich langsam auch hierzulande die Einsicht durch, dass die digitale Wende weitreichendere Veränderungen als die freie Verfügbarkeit wissenschaftlicher Publikationen bedeutet (Weingart & Taubert 2016).

---

<sup>1</sup>Dieser Beitrag ist die schriftlich erweiterte Fassung der gleichnamigen Keynote auf der 12. Jahrestagung der Gesellschaft für Hochschulforschung 2017 „Digitalisierung der Hochschulen: Forschung, Lehre und Administration“ in Hannover.

Ein zentraler Treiber der digitalen Transformation der Wissenschaft ist die Europäische Kommission, die Open Science zur Governance-Leitlinie der europäischen Forschungspolitik erklärt hat (European Commission 2014). Die digitale Wende hin zu Open Science wird dabei nicht nur als tiefgreifend, sondern auch als irreversibel betrachtet:

*“These [changes] have an impact on the entire research cycle, from the inception of research to its publication, as well as on the way in which this cycle is organised. The institutions involved in science are affected (research organisations, research councils, funding bodies), as is the way in which science is disseminated and assessed e.g. the rise of new scientific disciplines, innovative pathways in publishing (among them a substantial rise of Open Access journals), new scientific reputation systems, and changes in the way the quality and impact of research are evaluated. These trends are irreversible and they have already grown well beyond individual projects.”* (European Commission 2014: 2)

Vor diesem Hintergrund stellt sich aus wissenschaftssoziologischer Perspektive die Frage, ob die digitale Wende tatsächlich auf die Strukturen durchgreift, unter denen Wissenschaft operiert, und welche Konsequenzen sich daraus ergeben. Gesellschaftstheoretisch formuliert steht somit zur Debatte, ob mit der Digitalisierung gesellschaftliche Strukturtransformationen einhergehen, die sich am Fall der Wissenschaft exemplarisch aufzeigen lassen. Wenn man der evolutionstheoretischen Annahme folgt, dass Gesellschaftsstrukturen eng an die Verfügbarkeit von Verbreitungsmedien gekoppelt sind (Luhmann 2005), ist es plausibel anzunehmen, dass mit der Genese des digitalen Verbreitungsmediums die gegenwärtige Struktur der funktional differenzierten Gesellschaft aufbricht bzw. durch eine neue Struktur überlagert wird (Baecker 2007; Dickel und Franzen 2015). Digitalisierung würde somit das Entstehen einer „nächsten Gesellschaft“ (Baecker 2007), respektive der „nächsten Wissenschaft“ markieren (vgl. Hagner & Hirschi 2013).

Ziel des Beitrags kann es nicht sein, die These eines digital basierten gesellschaftsstrukturellen Wandels empirisch zu überprüfen. Vielmehr geht es im Folgenden darum, den Rahmen zu spannen, der eine systematische Untersuchung der digitalen Transformation der Wissenschaft anleiten kann. Die zugrundeliegende Definition von Wissenschaft als Produzent neuen, gesicherten Wissens folgt der Prämisse der Luhmannschen Systemtheorie, dass es sich bei der Wissenschaft wie bei allen sozialen Systemen um einen selbstreferentiellen, operational geschlossenen Kommunikationszusammenhang handelt (Luhmann 1990).

## 2 Dimensionen des digitalen Wandels in der Wissenschaft

Die Digitalisierung, verstanden als Umstellung vom Buchdruck auf elektronische Verbreitungstechnologie und aller damit einhergehenden technischen Weiterentwicklungen, erfasst die Gesellschaft als Ganze. Mit ihr verändert sich die Art der gesellschaftsweiten Kommunikation. Für die Wissenschaft, deren basales Element die Publikation ist (Stichweh 1987), beginnt dieser Umbruch in den 1980er Jahren mit der Entwicklung „digitaler Doppelgänger“ der klassischen Printzeitschrift (Keller 2001). Sukzessive entstehen weitere wissenschaftliche Publikationsformen. Das Internet überwindet die ökonomisch bedingten Veröffentlichungslimitationen, die noch dem Buchdruck anhafteten. Als veröffentlichungswürdig werden inzwischen nicht mehr nur die textbasierten Ergebnisse der Forschung gehandelt, sondern auch die dahinterliegenden Daten. Web 2.0-Technologien schaffen wiederum die Voraussetzungen für einen interaktiveren Umgang mit Daten und Publikationen. Digitale Forschungsinfrastrukturen ermöglichen organisationsübergreifende Kollaborationen im Bereich virtueller Wissensproduktion, an der mitunter auch nicht-zertifizierte Akteure beteiligt werden (Stichwort: Crowdsourcing). Die Digitalisierung betrifft somit nicht allein die Publikationsmodalitäten, sondern greift auch auf die sozialen Strukturen der Wissenschaft durch.

Zur näheren Identifikation der Veränderungsdynamiken der Wissenschaft qua Digitalisierung werden im Folgenden die Bereiche: (1) Wissensproduktion, (2) Wissensdissemination, (3) Wissensrezeption und (4) Wissensbewertung in ihren rezenten Entwicklungen separat beleuchtet.

### 2.1 Datafizierung der Wissensproduktion

Ein entscheidender Effekt von Digitalisierung ist die Ausbreitung von Big Data. Für die Wissenschaft bedeutet dies zunächst, dass für die Forschung nicht nur mehr Daten zur Verfügung stehen, sondern auch neue Methoden der Datenauswertung Anwendung finden. Die Auswertung riesiger Datenmengen bedarf einer gesteigerten Rechenkapazität, die erst durch die Entwicklung von Supercomputern und Cloud Computing möglich geworden ist. Big Data bildet zudem den Grundstein für die Entwicklung selbstlernender Algorithmen der Datenauswertung (Machine Learning).

Eines der Paradebeispiele für Big Data Science ist die Entzifferung des menschlichen Genoms. 1990 wurde das internationale Humane Genome Project gestartet, an dem rund 1000 Wissenschaftlerinnen und Wissenschaftler aus 40 Ländern beteiligt waren. Die Aufgabe war es, die rund drei Milliarden Basenpaare der menschlichen DNA vollständig zu entziffern. Im Ergebnis zeigte sich, dass der Mensch mit seinen noch nicht einmal 30.000 Genen über vergleichsweise weniger Gene verfügt als bisher angenommen. Dies warf die wissenschaftlich interessante Frage auf, ob die bis dato

gültige These vom genetischen Determinismus und damit das Konzept des Gens nicht grundsätzlich zu überdenken sei. Damit war zugleich das Anschlussprojekt definiert: ENCODE (ENCyclopedia Of DNA Elements) widmet sich seit 2003 nun der Kartierung des Transkriptoms, um die funktionalen Abschnitte im Genom bzw. Transkriptom zu identifizieren und in einer Online-Enzyklopädie zu sammeln.

Die epistemologisch relevante Frage ist, welche Art von Wissen solche datengetriebenen Ansätze genau liefern. Mit Big Data Analysen lassen sich zwar Muster identifizieren, wie in diesem Falle Gene auf Basis der Abfolge der Basenpaare der DNA auf einzelnen Chromosomen – nicht aber Aussagen darüber treffen, welche Funktionen diese Gene ausbilden (Alberts 2012). Läutet Big Data also das „Ende der Theorie“ ein, indem Korrelationen gegen Kausalitäten eingetauscht werden (Anderson 2008)? Wenn es zutrifft, dass Big Data Science Projekte von staatlicher Hand inzwischen bevorzugt zulasten von erklärenden Ansätzen der Small Science gefördert werden (Alberts 2012; Ratti 2016), findet sich hier ein empirisch interessanter Bezugspunkt für eine noch ausstehende Untersuchung des digitalen Wandels der Wissensproduktion im Fächervergleich. Eine neuerliche Hinwendung zu Big Data Research ist der Annahme geschuldet, dass mehr Daten zu mehr Erkenntnissen führen, auch wenn diese Erwartungshaltung sicher überzogen ist (z. B. boyd und Crawford 2013).

Forschungspolitisch geht es darüber hinaus inzwischen verstärkt darum, die vorhandene Vielfalt an Daten für die Forschung zu öffnen, Open Data heißt die Devise. Mit Open Data verbindet sich einerseits die Hoffnung auf eine Validierung wissenschaftlicher Ergebnisse und andererseits eine Entlastung der staatlichen Ausgaben im Sinne einer effizienten Forschungsförderpolitik, in der einmal erhobene Daten für Zweitverwertungen<sup>2</sup> zur Verfügung stehen. So wurden in den letzten Jahren nationale Forschungsdatenzentren errichtet und Online-Datenrepositorien aufgebaut. Die Europäische Kommission hat Open Data als Teil der Open Science Policy zu einem vorrangigen Ziel ihrer Politik erklärt. Mit deutschem Vorsitz erarbeitet die sogenannte EU Open Science Policy Group derzeit konkrete Richtlinien. Auf dem Open Science Monitor, der seit März 2017 online ist<sup>3</sup>, lässt sich der Fortschritt von Open Science in den verschiedenen Bereichen von Open Access bis zu Open Data sogar mitverfolgen.

Wenn, wie allgemein behauptet, Daten die „neue Währung“ sind, werden Zugriffsmöglichkeiten auf Daten zu einem weiteren Wettbewerbselement in der Forschung. So zieht die unterschiedliche Teilungsbereitschaft von Daten eine Grenze zwischen akademischer Forschung und Industrieforschung (Haeussler 2011). Aber auch innerhalb der akademischen Wissenschaft wird die Verfügung über Daten zum Spiegel von

<sup>2</sup>Die sogenannten FAIR-Prinzipien: *F*indable, *A*ccessible, *I*nteroperable, *R*eusable (Wilkinson et al. 2016) gelten dabei als Richtschnur für die nachhaltige Veröffentlichung von Forschungsdaten.

<sup>3</sup><http://ec.europa.eu/research/openscience/index.cfm?pg=home&section=monitor>

Machtasymmetrien. Angesichts von Big Data spricht Lev Manovich von der Entstehung dreier Klassen von Menschen: Die, „*die Daten (bewusst oder unbewusst, etwa wenn sie Datenspuren hinterlassen) erzeugen; die, die über die Möglichkeiten verfügen, Daten zu sammeln; und die, die wissen wie man diese Daten auswertet*“ (Manovich 2011, zitiert nach boyd und Crawford 2013: 212).

Angesichts des schieren Volumens an digitalen Daten stellt sich neben der Frage der für die Auswertung riesiger Datensätze erforderlichen Expertise ein Ressourcenproblem für die Wissenschaft. Die Datenaufbereitung braucht in der Regel Zeit und Personal. Crowdsourcing heißt eine der Lösungen des digitalen Zeitalters, um angesichts enger Personalkapazitäten solche Aufgaben trotzdem bewältigen zu können. Allgemein beinhaltet Crowdsourcing die Delegation von definierten Aufgaben an die anonyme Masse. Im wirtschaftlichen Kontext gehört Amazon's Mechanical Turk zu den bekanntesten Crowdsourcing-Plattformen dieser Art. In der Wissenschaft findet das Prinzip des Crowdsourcing unter den Begriffen Crowd Science bzw. Citizen Science inzwischen Anwendung (Franzoni & Sauerermann 2014).

Was in der Wissenschaft lange Zeit als undenkbar galt, und zwar die Inklusion von Nicht-Wissenschaftlerinnen und Nicht-Wissenschaftlern in die akademische Wissensproduktion, wird nun über digitale Infrastrukturen ermöglicht und als „Citizen Science“<sup>4</sup> auch politisch gefördert (Dickel & Franzen 2015). Zu den internationalen Erfolgsprojekten gehört ein Beispiel aus der Astrophysik namens Galaxy Zoo. Kernelement von Galaxy Zoo ist ein Gamification-Ansatz, der eine Anreizstruktur zur unbezahlten Teilnahme von Freiwilligen schafft. Analog zur Mission von Computerspielen besteht die konkrete Forschungsaufgabe darin, neue Bilder von Galaxien nach bestimmten Merkmalen zu klassifizieren. Sowohl die Anzahl der vorgenommenen Klassifikationen wie auch die erzielten Übereinstimmungen mit anderen Usern gehen dabei auf das Spielerkonto ein. Straub (2016) hat die Art der Beteiligung der Teilnehmerschaft in den verschiedenen Etappen des Forschungsprozesses am Beispiel von Galaxy Zoo im Hinblick auf die Generierung neuen Wissens minutiös nachgezeichnet. Das Innovationspotenzial dieser Art von Kollaboration zwischen Wissenschaft und Nicht-Wissenschaft wird durch die Rekonstruktion des iterativen Prozesses verdeutlicht, der unter anderem zur Entdeckung einer neuen Galaxie führte (Green Peas).

Aber selbst weitaus komplexere Aufgaben als die Klassifikation von Daten können an die Crowd delegiert werden, wenn nur das Forschungsdesign entsprechend gestaltet ist. Im Online-Projekt Foldit kann beispielsweise an der effizienten Faltung von Proteinen gearbeitet werden, ohne über spezielles biochemisches Fachwissen verfügen zu müssen. Ähnlich wie bei Galaxy Zoo wurde hierfür eine spielerische

---

<sup>4</sup>Anschauungsbeispiele aus ganz unterschiedlichen Forschungsfeldern finden sich auf den Internetplattformen [zooniverse.org](http://zooniverse.org), [buergerschaffenwissen.de](http://buergerschaffenwissen.de) oder [zentrumfuercitizenscience.at](http://zentrumfuercitizenscience.at).

Umgebung geschaffen, bei der die Teilnehmerinnen und Teilnehmer zusätzlich durch das Wettbewerbselement öffentlich sichtbarer High Scores zur fortlaufenden Partizipation angeregt werden. Auch hier wurden über die soziale Öffnung des Forschungsprozesses neue Erkenntnisse generiert, wie z. B. die Entschlüsselung der Struktur des monomerischen Proteins M-PMV-Protease.

Kurz gefasst reicht die digitale Beteiligung von nicht-zertifizierten Akteurinnen und Akteuren an der wissenschaftlichen Wissensproduktion je nach Projekt von der Datensammlung über die Aufbereitung von Daten bis hin zur Mitwirkung an der Dissemination von Wissen. So werden Citizen Scientists mitunter sogar, zumindest im Kollektiv, auf der Liste der Autorinnen und Autoren der entsprechenden Publikationen geführt – wie z. B. bei Foldit (Eiben et al. 2012).

## 2.2 Digitale Dissemination

Die Dissemination von Wissen über Printmedien hin zur elektronischen Distribution im Internet ist eine der offenkundigsten Veränderungen der Wissenschaft durch Digitalisierung. Damit einher geht erstens eine raumzeitlich entgrenzte Verfügbarkeit an wissenschaftlichen Produkten, zweitens eine Steigerung des Informationswerts von Publikationen durch Verlinkung oder Multimedialität und drittens eine Beschleunigung des Veröffentlichungsprozesses (Franzen 2011: 82ff).

Im Zuge der Entwicklung des WWW wurden die konventionellen Printzeitschriften bereits Mitte der 1990er Jahre für tot erklärt: *„We are not dealing with journal articles any more – we are dealing with research communications among scientists. Thus, the terms ‘article’, ‘paper’, and ‘publication’ should die“* (LaPorte et al. 1995: 1389). Von einem Tod des klassischen Zeitschriftenartikels kann heute zwar keine Rede sein. Beobachtbar ist aber einerseits eine Pluralisierung wissenschaftlicher Publikationsformen und andererseits ein Trend von einem nutzer- zu einem autorfinanzierten Geschäftsmodell wissenschaftlicher Zeitschriften im Modus von Open Access.

Die Europäische Kommission hat das ehrgeizige Ziel formuliert, bis 2020 alle wissenschaftlichen Artikel frei verfügbar zu machen (Enserink 2016). Die Open Access-Bewegung hat mit den Verpflichtungserklärungen der Wissenschaftspolitik und Forschungsförderung auch hierzulande an Fahrt aufgenommen (z. B. BMBF 2016). Gemäß dem Goldenen Weg des Open Access Publizierens, d. h. einer Erstveröffentlichung wissenschaftlicher Werke im Modus von Open Access, wurden in den letzten Jahren auch in Deutschland neue Open Access-Zeitschriften gegründet (z. B. das Open Gender Journal). Die großen Wissenschaftsverlage reagieren auf die Open Access-Forderung ihrerseits mit der Schaffung von Optionen, bei Bezahlung einer Gebühr einzelne Artikel auf Open Access zu schalten (z. B. Springer Choice) oder ganze Zeit-

schriften nach dem author-pays-Modell auf Open Access umzustellen. Angesichts solcher Entwicklungen steht derzeit zur Debatte, ob die Open Access-Bewegung insofern gescheitert sei, als kommerzielle Verlage die Idee der Allmende durch die Einführung neuer Bezahlschranken auf der Autorensseite unterhöheln (Herb 2017). Als (rechtswidrige) Antwort auf die verzögerte Realisierung von Open Access lässt sich Guerilla Open Access wie die riesige Schattenbibliothek Sci-Hub lesen, die die Bezahlschranken der Verlage aushebelt (Bohannon 2016). Sci-Hub, entwickelt 2011 von der Neurowissenschaftlerin Alexandra Elbakyan, stellt über verteilte Rechner die pdfs von Artikeln selbst hinter der paywall auf Abruf zur Verfügung. Guerilla Open Access bildet somit die radikalste Form, um Beschränkungen des Zugangs zu wissenschaftlicher Literatur zu umgehen. Jüngsten Studien zufolge umfasst die Abdeckung von Sci-Hub inzwischen rund 70 Prozent aller 81,6 Millionen wissenschaftlichen Artikel, die in der Datenbank Crossref registriert sind (Himmelstein et al. 2018).

Die freie Verfügbarkeit von Wissen, in diesem Falle von publizierten Forschungsergebnissen, betrifft aber nur die eine Seite des digitalen Wandels der Wissensdissemination. Mit den gesteigerten Veröffentlichungsmöglichkeiten im Internet pluralisieren sich auch die Formen wissenschaftlicher Publikationen. Neben die Standardform des wissenschaftlichen Zeitschriftenartikels sind zahlreiche neue Veröffentlichungsmöglichkeiten getreten. Eine der alternativen Formen zum wissenschaftlichen Artikel ist die sogenannte Nanopublikation als die kleinste Einheit publizierbarer Information. Nanopublikationen „enthalten die Behauptung genau einer Tatsache“ (Heißbrüggen-Walter 2013: 149). Ihre Funktion besteht darin, wissenschaftliche Ergebnisse einerseits schneller produzieren zu können. Andererseits, und das unterscheidet sie von der „Salamipublikation“, geht es darum, sie schneller rezipieren zu können und, vor allem, maschinenlesbar zu machen.

Eine weitere wissenschaftliche Ausdrucksform stellen Weblogs dar, die vor allem in den Geistes- und Sozialwissenschaften eine weite Verbreitung findet, individuell oder auch institutionell (z. B. LSE blogs). Ihre Funktion geht über die reine Vermittlung von Ergebnissen hinaus. Vielmehr ergeben sich neue Möglichkeiten, multimediale Verknüpfungen herzustellen und den Forschungsprozess transparent zu gestalten (Bruns & Burgess 2013). Bloggen in der Wissenschaft wird so zum Ausweis von „digital scholarship“ (Hecker-Stampehl 2013). In den Geisteswissenschaften kommen im Bereich Digital Humanities *Online-Editionsplattformen* hinzu. Hier wie auch in anderen Fällen lösen soziale Netzgemeinschaften die Figur des adressierbaren Autors ab (Hagner & Hirschi 2013) – mit potenziell weitreichenden Implikationen für die Reputationsbildung in der Wissenschaft. Ob sich alternative Publikationsformen im Wissenschaftsbetrieb insgesamt oder in einzelnen Disziplinen durchsetzen, hängt nicht zuletzt davon ab, inwiefern sie zitierfähig gemacht werden, so z. B. über Digital Object Identifier (DOI).

Mit der Nutzung sozialer Medien für die Wissenschaftskommunikation geht eine Verwischung der Grenze zwischen wissenschaftsinterner und wissenschaftsexterner Kommunikation einher (Bucchi 2013). Soziale Medien verändern den Publikumsbezug der Wissenschaft von den peers hin zu einer wissenschaftlich interessierten Öffentlichkeit (Könneker & Lugger 2013). Dabei sind die Verwendungsweisen der sozialen Medien in der Wissenschaft und ihre Frequenz in den Fachkulturen und Ländern unterschiedlich ausgeprägt (König und Nentwich 2016).

Eine der Funktionen der sozialen Medien besteht darin, selektionsverstärkend zu wirken. Eugene Garfield, Informationswissenschaftler, Erfinder des Journal Impact Faktors und Gründer des Institute for Scientific Information (ISI), rechnete vor einigen Jahren vor, dass von insgesamt 38 Millionen Zeitschriftenartikeln, die zwischen 1900 und 2005 veröffentlicht wurden, nur 0,5 Prozent mehr als 200-mal zitiert worden sind, 50 Prozent hingegen kein einziges Mal (Garfield 2006). Damit ist ein Umstand angesprochen, dass ein gewichtiger Teil an Veröffentlichungen schlicht nicht wissenschaftlich wahrgenommen wird. Um die Rezeptionswahrscheinlichkeit für ihre Produkte zu erhöhen, gestaltet sich mithilfe der sozialen Medien nicht nur die traditionelle Pressearbeit der Verlage neu. Jenseits der organisationsbezogenen PR ermöglichen die sozialen Medien wie Facebook oder Twitter es jetzt jeder einzelnen Wissenschaftlerin und jedem einzelnen Wissenschaftler, Aufmerksamkeit auf neue Veröffentlichungen oder bestimmte Themen zu lenken. Der Mikroblogging-Dienst Twitter liefert ein anschauliches Beispiel dafür, wie Information und Werbung kaum mehr auseinander zu halten sind. Ein PR-Service zu wissenschaftlichen Neuerscheinungen wird hier zum Teil von social bots ausgeführt, die jedoch nicht-selektiv agieren (Haustein et al. 2016).

### 2.3 Digitale Rezeptionsweisen

Die Anzahl wissenschaftlicher Publikationen nimmt bis heute kontinuierlich zu. 2016 wurden weltweit rund 2,3 Millionen Beiträge veröffentlicht, was einer Wachstumsrate von insgesamt 3,9 Prozent zwischen 2006 und 2016 entspricht<sup>5</sup>. Auch wenn die Digitalisierung mit gesteigerten Veröffentlichungsoptionen einhergeht und damit zum überbordenden Informationswachstum beiträgt, war bereits schon lange zuvor der Punkt erreicht, dass das schiere Volumen an Publikationen die individuelle Rezeptionskapazität übersteigt. So hatte Eugene Garfield einst den Journal Impact Factor als Selektionshilfe für Wissenschaftlerinnen und Wissenschaftler entwickelt, um sich im Dickicht wissenschaftlicher Informationen besser zurechtzufinden (Garfield 1955). Im Rahmen der Einführung des New Public Managements erfuhr der Journal Impact

<sup>5</sup>Bemerkenswert an diesen neuen Zahlen der National Science Foundation ist übrigens, dass China den Spitzenreiter an Produktivität, die USA, erstmalig überholt hat und nun einen Top-Anteil von 18,6% zu 17,9% an Publikationen unterhält (Tollefson 2018).

Factor jedoch einen Bedeutungswandel als wissenschaftlicher Leistungsindikator – mit all seinen bekannten Nebenfolgen (vgl. Fleck 2013).

In der Buchdruckgesellschaft bildeten wissenschaftliche Zeitschriften eine relevante Instanz der Informationshierarchisierung. Über ein Qualitätsprüfungsverfahren, das Peer Review System, sollte gewährleistet werden, dass die publizierten Beiträge den Standards des Faches entsprechen<sup>6</sup>. Mit der Pluralisierung wissenschaftlicher Publikationsformen erweitert sich nicht nur die wissenschaftliche Informationsbasis (z. B. über die begleitende Veröffentlichung von Daten). Auch neue Rezeptionsmuster entstehen, die sich nicht mehr allein an den Auswahlentscheidungen und redaktionellen Amplifikationen der Fachzeitschriften orientieren, sondern nutzergenerierten Algorithmen folgen.

Zu den Fachzeitschriften haben sich soziale Netzwerke der Wissenschaft gesellt, die wie Academia.edu und ResearchGate autorenzentrierte Publikationsportfolios enthalten und unterschiedliche Nutzerbedürfnisse bedienen (Van Noorden 2014). Auf den Online-Plattformen finden sich nach derzeitigem Stand (Oktober 2017) 56.673.413 Wissenschaftlerinnen und Wissenschaftler (Academia.edu) und über zehn Millionen Artikel (ResearchGate).<sup>7</sup> Jede Nutzerin und jeder Nutzer wird über personalisierte Algorithmen regelmäßig über Neuerscheinungen informiert. Analog zum bekannten Empfehlungsdienst von Amazon – „Kunden, die diesen Artikel gekauft haben, kauften auch“ – operiert also auch in der Wissenschaft ein auf nutzergenerierten (Meta-)Daten und Nutzungsstatistiken aufgebautes algorithmisches System, um Personen und Inhalte zusammenzuführen (wie zuvor über Zitationsnetzwerke).

Mit der Digitalisierung verändern sich aber auch die Rezeptionsweisen selbst. Das Lesen richtet sich nicht mehr allein am Papier aus, sondern findet häufig am Computerbildschirm oder mithilfe von Mobile Devices statt. Entsprechend werden laufend neue digitale Anwendungen entwickelt, um sich im wissenschaftlichen Informationsdickicht zu orientieren. Wie Wissenschaftlerinnen und Wissenschaftler angesichts der Informationsexplosion ihr eigenes Rezeptionsverhalten mithilfe von technischen Hilfsmitteln strukturieren, hat Pain (2016) genauer erfragt. Augenscheinlich besteht vorrangig ein Zeitproblem, um mit der wachsenden Forschungsliteratur im Feld Schritt zu halten. Die Entwicklung von Automatisierungstechniken wie das Distant Reading im Bereich der Literaturwissenschaft (Morretti 2016) tragen diesem Umstand zwar

---

<sup>6</sup>Das Peer Review-Verfahren ist auch vor Fehlern nicht gefeit und unterliegt zahlreichen sozialen Verzerrungsmechanismen. Angesichts der aufscheinenden Reproduzierbarkeitsprobleme mehren sich gegenwärtig die Krisendiagnosen zum Peer Review (Franzen 2016).

<sup>7</sup>Jüngst wurde bekannt, dass die Großverlage eine Klage gegen die unrechtmäßige Zweitveröffentlichung auf diesen Plattformen eingereicht haben, um die Betreiber dazu zu zwingen, den lizenzierten Content zu löschen (Van Noorden 2017).

Rechnung, sie werden als Erkenntnisverfahren aber ebenso kontrovers diskutiert (Hagner & Hirschi 2013).

Eine speziell für die Lebenswissenschaften entwickelte App ist z.B. Papr, die dem Prinzip der Dating-App Tinder folgt. Ihre Funktion besteht darin, Preprints auf Basis von Abstracts zu sondieren (McKenzie 2017). Selbst eine solch flüchtige digitale Klassifikation von wissenschaftlichen Beiträgen gibt theoretisch gesehen Auskunft über wissenschaftliches Bewertungshandeln und kann so zu vergleichenden, externen Analysen zur Klassifikation wissenschaftlicher Veröffentlichungen herangezogen werden.

## 2.4 Digitale Bewertungspraktiken

Bewertungen sind in der Wissenschaft an der Tagungsordnung, ob es um die Relevanz von Publikationen, den gesellschaftlichen Impact der Forschung oder Berufungsentscheidungen geht. Ihren Ursprung haben wissenschaftliche Bewertungspraktiken im Peer Review der Fachzeitschriften. Das Peer Review, dessen Prototyp die Begutachtung wissenschaftlicher Beiträge vor ihrer Veröffentlichung in Fachzeitschriften ist, ist zugleich der Kern wissenschaftlicher Autonomie<sup>8</sup>. Daneben existieren externe Klassifizierungsversuche, die weniger auf explizite als auf implizite individuelle Bewertungen bauen, um daraus Wertigkeiten z.B. von Publikationen abzuleiten (Zitationsstatistiken). Mit der Digitalisierung wird dieser Trend einer datenbasierten Bewertungslogik verstärkt.

Bekanntermaßen wird jedes Nutzungsverhalten im Internet automatisch aufgezeichnet. Dies betrifft nicht nur das Kaufverhalten bei Amazon, die Informationssuche bei Google oder das Leseverhalten bei Spiegel Online, sondern schließt auch den individuellen Umgang mit wissenschaftlichen Publikationen mit ein. Auch die Wissenschaftlerin und der Wissenschaftler werden auf diese Weise getrackt und das Rezeptionsverhalten somit datafiziert:

„Heute kann das System für *jede einzelne Nutzung* einer elektronischen Ressource protokollieren, *welche* Ressource verwendet wurde, *wer* sie benutzt hat, *wo* diese Person sich befand, *wann* sie verwendet wurde, *welche Art von Anforderung* gestellt wurde, *welche Art von Aufzeichnung* es war und *von wo* sie verwendet wurde.“ (Kurtz & Bollen 2010: 4, Hervorhebung und Übersetzung MF).

<sup>8</sup>Historisch betrachtet ist das Peer Review-Verfahren, wie Biagioli (2002) im Detail herausgearbeitet hat, aus der staatlichen Zensur hervorgegangen. Die proklamierte Autonomie der Wissenschaft speist sich somit aus einer Befreiung von staatlichen Zwängen hin zur Selbstdisziplinierung im Auftrag der Wahrheitssuche.

Diese digital erfassten Nutzungsstatistiken sind es, die für die Bewertung von Publikationen nun zunehmend in klassifizierender Absicht herangezogen werden (vgl. Franzen 2015). Die Public Library of Science (PLOS) unternahm hierbei den ersten Vorstoß mit der Einführung der sogenannten Article-Level Metrics (ALMs). Das Herausgebersteam der neu gegründeten Open Access-Plattform PLOS kritisierte am bestehenden System vor allem die Abhängigkeit vom Journal Impact Faktor: Für eine faire Leistungsbewertung sei die Substanz des einzelnen Artikels höher zu gewichten als der Publikationsort. Der Hype um die wenigen Publikationsplätze in einem der Top Journale habe einen unlauteren Wettbewerb in Gang gesetzt, der für die Wissenschaft zu unliebsamen Konsequenzen führe (PLOS Medicine Editors 2006)<sup>9</sup> – deswegen brauche es eine artikelzentrierte Auswertung, so die Argumentation für die Einführung der sogenannten Article-Level Metrics (ALM).

Die Article-Level Metrics fungieren somit in zweierlei Hinsicht als Gegenmodell zum Journal Impact Faktor: Zum einen handelt es sich um eine artikelzentrierte und nicht um eine journalzentrierte Auswertung; zum anderen geht das Modell über eine eindimensionale, zitationsbasierte Bewertung hinaus. Statt nur die Zitationen (in den sogenannten Quellzeitschriften des Web of Science oder Scopus) zu zählen, werden bei den ALMs auch die Verwendungsweisen von wissenschaftlichen Artikeln unterhalb der Zitationsschwelle als Bewertungsdimensionen berücksichtigt. Hierzu gehören die Rezeptionspraktiken von Viewed, Discussed, Recommended und Cited (Fenner 2013). ALMs wurden für alle PLOS-Fachzeitschriften 2009 eingeführt.

Noch weitreichender angelegt ist das Altmetrics-Konzept, das mit dem sogenannten Altmetrics Manifesto 2010 eine hohe Aufmerksamkeit erreichte und das Ziel verfolgt, eine Änderung der wissenschaftlichen Evaluationspraxis zu bewirken (Priem et al. 2010). Seither haben sich mehrere Anbieter am Markt etabliert, die spezielle Altmetrics-Dienste für die Wissenschaft offerieren. Das bekannteste Produkt ist der sogenannte Altmetric-Donut des gleichnamigen Start-ups Altmetric.com mit Sitz in London. Analog zu den ALMs von PLOS bilden Altmetrics ein breites Resonanzspektrum pro wissenschaftlichem Artikel ab. Anders als bei den multidimensionalen ALMs werden die erfassten Nutzungsdaten zu dem sogenannten Altmetric-Score verdichtet, der auf dieser numerischen Basis Vergleiche und Rankings von Artikeln ermöglicht. So veröffentlicht Altmetric.com z. B. ein jährliches Top 100 Ranking. Die internationalen börsenorientierten Großverlage wie Wiley, Springer und Elsevier haben den Altmetric-Donut bereits in ihr Online-Portfolio integriert. Auch die High-Impact Zeitschriften Science und Nature gehören zu den early adopters. Für jeden neu erscheinenden Artikel wird auf den Websites heute der Altmetric-Score ausgewiesen.

---

<sup>9</sup>Eine ähnlich gelagerte Kritik ist Gegenstand der im Jahre 2013 veröffentlichten San Francisco Declaration (DORA), die international breite Unterstützung fand.

Der Altmetric Donut gibt mit seiner jeweiligen Farbgebung Auskunft über die Art der Resonanz, die ein Artikel<sup>10</sup> erzielt hat. Die Quellen, anhand derer die Online-Aufmerksamkeit gemessen wird, sind vielseitig und beliebig erweiterbar. Aktuell gehören unter anderem dazu: journalistische Online-Medien, soziale Medien wie Facebook und Twitter, Wikipedia, YouTube oder das Literaturverwaltungsprogramm Mendeley. Im Unterschied zu Zitationszählungen in Fachzeitschriften erscheinen die Altmetrics-Nennungen in Echtzeit und sind auf den entsprechenden Webseiten der Zeitschriften öffentlich sichtbar.

Als Marketingtool von Zeitschriften hat das Altmetrics-Konzept sicher seine Berechtigung. Wenn es jedoch darum geht, mit Altmetrics das wissenschaftliche Reputationssystem nachhaltig zu verändern, so die Forderung der Proponenten (Priem 2013), dann wird ein Problem offenkundig: Es ist bislang nicht klar, was Altmetrics genau abbilden. Erforscht werden Altmetrics bislang vor allem im Rahmen der Bibliometrie (Taylor 2013). Die bibliometrische Forschung beschränkt sich allerdings mehrheitlich darauf, Korrelationen zwischen Zitationsraten und Altmetrics zu errechnen (z. B. Costas, Zahedi & Wouters 2015). Ergebnisse dieser Art von Forschung besagen z. B., dass Mendeley-Bookmarks Zitationen am ehesten widerspiegeln (Li et al. 2011) oder die Anzahl von Downloads im Verhältnis 70:1 zu Zitationen von Artikeln steht (Lin & Fenner 2013: 23). Altmetrics entfalten jedoch eine Wirkung in der Wissenschaft, die über die Prädiktion von Zitationen weit hinausgeht. Sie fungieren insbesondere als Narzissmustechnologie (Wouters & Costas 2012). Dies wird aus der Werbebotschaft des Dienstes Altmetric.com bspw. besonders deutlich:

*“Authors love article level metrics. It doesn’t matter if it’s their first or hundredth journal article – researchers want their work to be shared, discussed, and applied. They’re curious about who is talking about their work and what is being said. Using the Altmetric data and API, publishers can deliver real value to their authors and readers through powerful article level metrics pages.”*

Die Frage ist jedoch, wie sich die Einstellung von Wissenschaftlerinnen und Wissenschaftlern ändert, wenn sich Altmetrics in eine Kontrolltechnologie verwandeln. Schon jetzt wird über die Möglichkeit einer Inkorporation von Altmetrics in die institutionalisierte Leistungsbewertung in der Wissenschaft nachgedacht. Insbesondere im Rahmen der quantitativen Bewertung des gesellschaftlichen Impacts der Forschung, als einer der neu eingeführten Bewertungsdimension für Forschungsqualität, die z. B. in das britische Research Excellence Framework integriert wurde, wird über die mögliche Verwendung von Altmetrics anstelle des Peer Reviews debattiert (Bornmann 2014). So hat die

<sup>10</sup>Das Altmetrics-Konzept ist anders als zitationsbasierte Maße nicht auf den wissenschaftlichen Artikel festgelegt, sondern kann auf beliebige wissenschaftliche Outputs (Foliensets, Datensets, Patente) angewendet werden.

zuständige Behörde, das Higher Education Funding Council for England (HEFCE), ein Gutachten zur Abschätzung der Eignung von Altmetrics für die Evaluation des gesellschaftlichen Impacts in Auftrag gegeben. Die Autoren kommen darin zu dem Schluss, dass das Instrument noch nicht ausgereift sei und noch weiterer Standardisierungen bedarf, um für die Forschungsevaluation eingesetzt zu werden. Die prinzipielle Anwendbarkeit wird zukünftig jedoch nicht ausgeschlossen (Wouters et al. 2015).<sup>11</sup>

Analog zur Auswertung der nutzergenerierten Daten durch Altmetrics wird derzeit an verschiedenen Stellen an der Entwicklung von weiteren digitalen Diensten gearbeitet, um das redaktionelle Peer Review-System mithilfe von Big Data Analytics zu unterstützen. So soll ein neues Produkt namens Bibliometric Intelligence (Meta) das redaktionelle Entscheidungsverfahren zwar (noch) nicht ersetzen, aber zumindest informieren.

### 3 Zusammenfassung der sich abzeichnenden Trends

Entlang der vier betrachteten Dimensionen lässt sich leicht erkennen, dass sich die Wissenschaft in einem digitalen Umbruch befindet, der über den im öffentlichen Diskurs dominierenden Aspekt von Open Access und die damit verbundene Rolle von Wissenschaftsverlagen innerhalb der wissenschaftlichen Kommunikation weit hinausgeht.

Vielmehr zeigen die dargestellten Entwicklungen, dass im Zuge der Digitalisierung etablierte Differenzen zwischen Wissenschaft und Nichtwissenschaft brüchig werden. Dies betrifft auf der *Ebene der Wissensproduktion* die Rolle des Wissensproduzenten (professionelle Wissenschaft vs. Citizen Science) sowie die Art des Wissens, das mithilfe von Big Data produziert wird (Information vs. Wissen). Auf der *Ebene der Wissensdissemination* lässt sich mit den gestiegenen Veröffentlichungsmöglichkeiten eine Pluralisierung der wissenschaftlichen Publikationsformen erkennen. Im Bereich Wissenschaftskommunikation 2.0 verschwimmt die etablierte Grenze zwischen wissenschaftsinterner und -externer Kommunikation, wenn soziale Medien unter anderem zur Aufmerksamkeitsgenerierung für eigene Publikationen genutzt werden oder Weblogs sich an eine breite Öffentlichkeit richten. Dieser Umstand erzeugt auf der *Ebene der Wissensrezeption* eine neue Unübersichtlichkeit. Was als wissenschaftlich relevantes Wissen gilt, ist eine Frage, die nicht mehr allein über die Qualitätssicherungsverfahren wissenschaftlicher Zeitschriften geklärt wird. Klassische Zeitschriften existieren zwar nach wie vor – auch im Internet. An ihre Seite gesellen sich jedoch zunehmend wissenschaftliche Publikationsplattformen, die anderen Veröffentlichungsregeln folgen. Dies sind zum einen sogenannte Megajournals wie PLOS ONE,

---

<sup>11</sup>Dies hätte – wie bereits aus anderen existierenden Indikatorenbasierten Verfahren bekannt – vermutlich ein reaktives Verhalten zur Folge. Medialisierungseffekte würden so noch verstärkt (Franzen 2015).

die mithilfe von Peer Review Light nach dem Prinzip „Publish-First, Filter-Later“ funktionieren (vgl. Spezi et al. 2017). Zum anderen sind Zweitverwertungsinstanzen wie die Academic Network Sites ResearchGate oder Academia.edu zu nennen, die ein Post-Publication Peer Review in ihr Portfolio integriert haben. Dabei handelt es sich um erfolgreiche Unternehmungen einer auf Big Data basierenden Plattformökonomie, die wissenschaftliche Publikationen zwar zentral online verfügbar machen, jedoch mit dem Open Access-Gedanken brechen (Hall 2015). An diesem Beispiel wird nochmals deutlich, dass das Diktum von Open Science die fortschreitende digitale Transformation der Wissenschaft nur unzureichend erfassen kann. Allein die Konnotation einer Demokratisierung von Wissenschaft erscheint irreführend.

Mit dem Digitalisierungsschub geht einher, dass das wissenschaftliche Nutzungsverhalten im Umgang mit Literatur und Daten automatisch getrackt wird und hieraus neue Geschäftsmodelle entstehen. Dieser technische Umstand bildet unter anderem den Grundstein für die Genese von Altmetrics, die derzeit als neue Form der Leistungsbewertung in Betracht gezogen werden und sich theoretisch auf alle wissenschaftlichen Produkte anwenden lassen. Altmetrics bilden die Resonanz wissenschaftlicher Beiträge im digitalen Raum ab und werden zunehmend relevant für die Bewertung von Forschung und deren Impact. Die digitale Anzeige der Echtzeitwirkung wissenschaftlicher Publikationen fordert den Wissenschaftler, seine Institution sowie deren Kommunikationsabteilungen heraus, sich entsprechend in den sozialen Medien zu positionieren, um die je eigene Resonanzquote zu erhöhen. Wenn Twitter, Facebook oder YouTube jene Kanäle sind, die von Altmetrics-Diensten ausgelesen werden, dann impliziert dies für Wissenschaftlerinnen und Wissenschaftler, sie auch entsprechend zu bedienen. Noch 2014 notierte die National Information Standards Organization, die an der Standardisierung des Altmetrics-Instruments arbeitet, dass Altmetrics 95 Prozent der Wissenschaftler noch nicht bekannt sei (NISO 2014). Dabei hat sich das Prinzip von Altmetrics auch in den wissenschaftlichen Netzwerken wie ResearchGate längst eingeschrieben.

Schon länger gilt, dass Scorings und Rankings dem Wissenschaftler direkt oder indirekt ein kontinuierliches Impression-Management abverlangen, das nun auch die Aktivität in den sozialen Medien einschließt (Franzen 2015, 2017). So lässt sich konstatieren, dass nirgendwo sonst „das metrische Wir“ (Mau 2017) so virulent ist wie in der Wissenschaft. Im Zuge der Digitalisierung werden Statuspositionen online einsehbar und wie bei Altmetric.com z. B. mithilfe von Badges ausgeflaggt. Quantitative Leistungsindikatoren verstärken bekanntermaßen die extrinsische und nicht die intrinsische Motivation. „Je mehr die Statusvisibilisierung durch quantifizierende Bewertungsformen an Fahrt gewinnt, desto mehr werden wir zu status seekers (...) in einem System der differenziellen Wertigkeit.“ (Mau 2017: 64)

Insofern steht auch zur Debatte, inwiefern die Digitalisierung auf die wissenschaftliche Reputationsordnung durchgreift. Eine der möglichen nichtintendierten Folgen für die Wissenschaft, die das Altmetrics-Konzept nahelegt, ist die Entkopplung von Erfolg (Reputationserwerb) und Leistung (wissenschaftliche Erkenntnisproduktion).

#### 4 **Ausblick**

Bislang wird den Hochschulen in Deutschland hinsichtlich der Digitalisierung eine geringe Transformationsgeschwindigkeit bescheinigt (Scheer 2015). Bezogen auf die Frage der Digitalisierung der Hochschule sind die Funktionsbereiche Forschung, Lehre und Administration einzeln, aber auch im Zusammenspiel zu betrachten. Der vorliegende Beitrag widmete sich allein der Forschung, die im Unterschied zu Hochschulen nicht an lokale Kontexte gebunden ist. Ziel war eine multidimensionale Sondierung der sich abzeichnenden digitalen Entwicklungen im Wissenschaftssystem, die jedoch aus Platzgründen exemplarisch bleiben muss. Die vorgenommene Analyse stellt vielmehr einen heuristischen Rahmen für die Beschreibung der digital induzierten Veränderungen in den Bereichen Produktion, Rezeption, Dissemination und Bewertung von Wissen zur Verfügung, um weitere empirische Forschung anzuregen. Den Ausgangspunkt der Analyse bildete der digitale Wandel des Publikationssystems.

Die Genese und Expansion des Wissenschaftssystems, wie wir es heute kennen, ist eng an die Publizität geknüpft. Ein wichtiger Impuls für die Ausdifferenzierung von Wissenschaft ging im 17. Jahrhundert von der Gründung der ersten wissenschaftlichen Zeitschriften aus, die den bis dato gängigen adressatenbeschränkten Briefverkehr ablösten. Über die Veröffentlichung wissenschaftlicher Erkenntnisse in periodisch erscheinenden Zeitschriften wurde ein breiter Wissenszugang ermöglicht und die Kumulation von Wissen durch Zertifizierungsinstanzen (Peer Review) vorangetrieben. Wenn sich aber, wie gegenwärtig, das zentrale Verbreitungsmedium gesellschaftlicher Kommunikation vom Buchdruck hin zum Internet verändert, entstehen zugleich neuartige Strukturen, die die bisherigen zwar nicht ersetzen, aber zumindest überlagern. Ein anschauliches Beispiel dafür ist die Differenz von Pre-Publication Peer Review und Post-Publication Peer Review. Aktuelle Visionen vermuten einen kompletten Funktionswandel der wissenschaftlichen Zeitschrift (Kriegeskorte 2012). Mit einer zeitlichen Verschiebung der wissenschaftlichen Qualitätskontrolle sind nicht zuletzt Rollenverschiebungen verbunden. Post-Publication Peer Review ist nicht mehr allein dem ausgewiesenen Experten vorbehalten, sondern schließt grundsätzlich jedermann mit ein (z.B. *PubPeer*). Weniger das subjektive Urteil, sondern die Aggregation vieler Einzelbewertungen soll damit für die wissenschaftliche Qualitätssicherung sorgen. Dies sind Mechanismen, die der Plattformökonomik von Amazon & Co. entstammen. Aber lassen sich solche Prinzipien umstandslos auf die Wissenschaft übertragen – und mit welchen Folgen? Welche Funktion kommt zukünftig der wissenschaftlichen Fach-

expertise zu und wie und auf welcher Basis wird epistemische Autorität zukünftig zugewiesen (Dickel und Franzen 2016)?

Wissenschaftspolitisch gewendet, scheint es unabdingbar, die Prozesse der Digitalisierung umfassend auch auf ihre weitreichenden Implikationen hin zu reflektieren, statt sie auf einzelne Aspekte wie die Entwicklung von technischen Infrastrukturen für Forschung (und Lehre) zu reduzieren (Schlagwort Science 2.0) oder die Digitalisierung aus einer normativen Position heraus vorschnell mit der Demokratisierung von Wissen gleichzusetzen (Schlagwort Open Science). Um die viel diskutierten Chancen und Risiken der digitalen Wende überhaupt abschätzen zu können, braucht es einen theoriegeleiteten empirischen Einblick in die rezenten Dynamiken der fortschreitenden Digitalisierung von Forschungspraktiken. Zu den Herausforderungen der Digitalisierung gehört demnach nicht nur die Frage, wie Wissen künftig mit welchen digitalen Tools hergestellt, verbreitet, rezipiert und bewertet wird, sondern auch, wer zukünftig das Wissen der Gesellschaft produziert – Wissenschaftlerinnen und Wissenschaftler, Crowdsources und/oder Maschinen? Klar benannt ist zumindest die wissenschaftspolitische Stoßrichtung der digitalen Transformation: Daten der Forschung maschinenlesbar zu machen, um sie mit statistischen Verfahren zu validen Informationen zu verdichten und so maschinelles Lernen zu ermöglichen (European Commission 2014; Wilkinson et al. 2016).

Die Vermutung liegt nahe, dass sich mit computergesteuerten Erkenntnis- und Bewertungsverfahren auch der Charakter des Wissens verändert (Hagner & Hirschi 2013). Versuche einer datenbasierten Automatisierung verschiedener Produktionsstufen finden aber nicht nur in der Wissenschaft, sondern in allen Gesellschaftsbereichen gleichzeitig statt und stellen somit einen zentralen Effekt der digitalen Wende dar. Ein differenzierungstheoretischer Zugang auf die Frage nach den Implikationen der Digitalisierung scheint der geeignete Ansatz, um aus einer vergleichenden Perspektive Reichweite und Tiefe des digitalen Wandels empirisch genauer bestimmen und nachhaltig gestalten zu können.

## Literatur

- Alberts, B. (2012). Editorial: The End of „Small Science“? *Science* (337), 1583
- Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired Magazine* vom 23.06.2008 Abgerufen am 27.09.2017 von <https://www.wired.com/2008/06/pb-theory/>
- Baecker, D. (2007). *Studien zur nächsten Gesellschaft*. Frankfurt am Main: Suhrkamp
- Biagioli, Mario (2002). From Book Censorship to Academic Peer Review. *Emergences: Journal for the Study of Media & Composite Cultures*, 12 (1), 11–45

Bohannon, J. (2016). Who's downloading pirated papers? Everyone. *Science*, 352 (6285), 508–512

Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8, 895–903

boyd, d.; Crawford; K. (2013). Big Data als kulturelles, technologisches und wissenschaftliches Phänomen. Sechs Provokationen. In: H. Geiselberger & Moorstedt, T. (Hrsg.), *Big Data. Das neue Versprechen der Allwissenheit (187–218)*. Berlin: Suhrkamp, edition unseld

Bruns, A. & Burgess, J. (2013). Blogforschung: Der ‚Computational Turn‘. In: P. Haber & E. Pfanzelter (Hrsg.), *historyblogosphere. Bloggen in den Geschichtswissenschaften (S. 135–148)*. München: Oldenbourg

Bucchi, M. (2013). Style in science communication. *Public Understanding of Science*, 22 (8), S. 904 –915

Bundesministerium für Bildung und Forschung (BMBF) (2016). Open Access in Deutschland. Die Strategie des Bundesministeriums für Bildung und Forschung. Abgerufen am 27.09.2017 von [https://www.bmbf.de/pub/Open\\_Access\\_in\\_Deutschland.pdf](https://www.bmbf.de/pub/Open_Access_in_Deutschland.pdf)

Costas, R.; Zahedi, Z. & Wouters, P. (2015). Do ‚altmetrics‘ correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66 (10), 2003–2019

Dickel, S. & Franzen, M. (2015). Digitale Inklusion: Zur sozialen Öffnung des Wissenschaftssystems. *Zeitschrift für Soziologie*, 44 (5), 330–347

Dickel, S. & Franzen, M. (2016). The “Problem of Extension” revisited: new modes of digital participation in science. *JCOM*, 15 (01), A06\_en

DORA (2013). American Society for Cell Biology: San Francisco declaration on research assessment. Abgerufen am 27.09.2017 von [http://www.ascb.org/files/SFDeclaration\\_FINAL.pdf](http://www.ascb.org/files/SFDeclaration_FINAL.pdf)

Eiben, C.; Siegel, J. B.; Bale, J. B.; Cooper, S.; Khatib, F.; Shen, B. W.; Foldit Players; Stoddard, B. L.; Popovic, Z. & Baker, D. (2012). Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. *Nature Biotechnology*, 30, 190–192

Enserink, M. (2016). In dramatic statement, European leaders call for ‘immediate’ open access to all scientific papers by 2020. *Science*, doi:10.1126/science.aag0577

European Commission (2014). Background document: Public Consultation ‘Science 2.0’: Science in transition. Directorates-General for Research and Innovation (RTD) and Communication Networks, Content and Technology (CONNECT) Abgerufen am 27.09.2017 von <https://ec.europa.eu/research/consultations/science-2.0/background.pdf>

Fecher, B. & Friesike, S. (2013). Open Science: One term, five schools of thought. In: S. Bartling & S. Friesike (Hrsg.), *Opening Science. The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing* (S. 7–47). Heidelberg et al.: Springer Open

Fenner, M. (2013). What Can Article-Level Metrics Do for You?. *PLoS Biol* 11(10): e1001687

Fleck, C. (2013). Der Impact Faktor-Fetischismus. *Leviathan* 41 (4), 611–646

Franzen, M. (2011). *Breaking News. Wissenschaftliche Zeitschriften im Kampf um Aufmerksamkeit*. Baden-Baden: Nomos

Franzen, M. (2015). Der Impact Faktor war gestern. Altmetrics und die Zukunft der Wissenschaft. Themenheft: Der impact des impact factors. *Soziale Welt*, 66 (2), 225–242

Franzen, M. (2016). Science between Trust and Control: Non-Reproducibility in Scholarly Publishing. In: H. Atmanspacher & S. Maasen (Hrsg.): *Reproducibility: Principles, Problems, Practices and Prospects* (S. 468–485). Hoboken, New Jersey: John Wiley & Sons, Inc

Franzen, M. (2017). Digitale Resonanz Neue Bewertungskulturen fordern die Wissenschaft heraus, *WZB Mitteilungen* 155, 30–33

Franzoni, C. & Sauermann, H. (2014). Crowd science. The organization of scientific research in open collaborative projects. *Research Policy* 43 (1), 1–20

Garfield, E. (1955). Citation Indexes for Science. A New Dimension in Documentation through Association of Ideas. *Science*, 128, 108–111

Garfield, E. (2006). The History and Meaning of the Journal Impact Factor. *JAMA* 295 (1), 90–93

Haeussler, C. (2011). Information-sharing in academia and the industry: A comparative study. *Research Policy*, 40 (1), 105–122

Hagner, M. & Hirschi, C. (2013). Editorial. In: D. Gugerli; M. Hagner; C. Hirschi; A. B. Kilcher; P. Purtschert; P. Sarasin & J. Tanner (Hrsg.), *Nach Feierabend. Digital Humanities* (S. 7–10). Zürich, Berlin: diaphanes

Hall, G. (2015). What does Academia\_edu's success mean for Open Access? The data-driven world of search engines and social networking. *LSE Impact Blog*. Abgerufen am 27.03.2018 von <http://blogs.lse.ac.uk/impactofsocialsciences/2015/10/22/does-academia-edu-mean-open-access-is-becoming-irrelevant/>

Haustein, Stefanie; Bowman, Timothy D.; Holmberg, Kim; Tsou, Andrew; Sugimoto, Cassidy R. & Vincent Larivière. (2016). Tweets as impact indicators: Examining the implications of automated “bot” accounts on twitter. *Journal of the Association for Information Science and Technology* 67(1), 232–238

Hecker-Stampehl, J. (2013). Bloggen in der Geschichtswissenschaft als Form des Wissenstransfers. In: P. Haber & E. Pfanzelter (Hrsg.), *historyblogosphere. Bloggen in den Geschichtswissenschaften* (S. 37–50). München: Oldenbourg

Herb, U. (2017). Open Access: Von Inklusion zu Exklusivität? DOI:10.5281/zenodo.1001901

Heißbrüggen-Walter, S. (2013). Tatsachen im semantischen Web: Nanopublikationen in den digitalen Geisteswissenschaften? In: P. Haber & E. Pfanzelter (Hrsg.): *historyblogosphere. Bloggen in den Geschichtswissenschaften* (S. 149–160). München: Oldenbourg

Himmelstein, D. S.; Rodriguez Romero, A.; Levernier, J. G., Munro, T. A.; McLaughlin, S.R.; Tzovaras, B. G.; Greene, C. S. (2018). Research: Sci-Hub provides access to nearly all scholarly literature. *Elife* 1, S. 1–22. doi: 10.7554/eLife.32822

Keller, A. D. (2001). *Zeitschriften in der Krise: Entwicklung und Zukunft elektronischer Zeitschriften*. Dissertation Humboldt-Universität zu Berlin, Berlin

König, R. & Nentwich, M. (2016). Soziale Medien in der Wissenschaft. In: J.-H. Schmidt; & M. Taddicken (Hrsg.), *Handbuch Soziale Medien* (S. 1–20). Wiesbaden: Springer Fachmedien

Könneker, C. & Lugger, B. (2013). Public Science 2.0 – Back to the Future. *Science*, 342 (6154), 49–50

Kriegeskorte, N. (2012). Open Evaluation: A Vision for Entirely Transparent Post-Publication Peer Review and Rating for Science. *Frontiers in Computational Neuroscience*, 6 (79). doi: 10.3389/fncom.2012.00079

Kurtz, M. J. & Bollen, J. (2010). Usage bibliometrics. *Annual Review of Information Science and Technology*, 44 (1), 1–64

LaPorte, R. E.; Marler, E.; Akazawa, S. et. al. (1995). The death of biomedical journals. *British Medical Journal*, 310 (6991), 1387–1390

Li, X.; Thelwall, M.; Giustini, D. (2012). Validating Online Reference Managers for Scholarly Impact Measurement. *Scientometrics*, 91 (2), 461–471

Lin, J. & Fenner, M. (2013). Altmetrics in Evolution: Defining and Redefining the Ontology of Article-Level Metrics. *ISQ*, 25, 20–26

Luhmann, N. (1990). *Die Wissenschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp

Luhmann, N. (2005). Veränderungen im System gesellschaftlicher Kommunikation und die Massenmedien. In: ders. (Hg.), *Soziologische Aufklärung 3* (S. 355–368). Wiesbaden: VS Verlag, 4. Auflage

Mau, S. (2017). *Das metrische Wir. Über die Quantifizierung des Sozialen*. Berlin: Suhrkamp

McKenzie, L. (2017). Swipe right for science. Papr app is 'Tinder for preprints'. *Nature*. DOI: 10.1038/nature.2017.22163

Morretti, F. (2016). *Distant Reading*. Konstanz: Konstanz University Press

National Information Standards Organization (NISO) (2014). *Alternative Metrics Initiative Phase I White Paper* vom 6. Juni 2014. Abgerufen am 27.09.2017 von [https://groups.niso.org/apps/group\\_public/download.php/13809/Altmetrics\\_project\\_phase1\\_white\\_paper.pdf](https://groups.niso.org/apps/group_public/download.php/13809/Altmetrics_project_phase1_white_paper.pdf)

Pain, E. (2016). How to keep up with the scientific literature. *Science Careers* doi:10.1126/science.caredit.a1600159

PLoS Medicine Editors (2006). Editorial: The Impact Factor Game. *PLoS Medicine*, 3, e291

Priem, J. (2013). Scholarship: Beyond the paper. *Nature*, 495, 437–440

Priem, J.; Taraborelli, D.; Groth, P. & Neylon, C. (2010). *Altmetrics: A manifesto*. Abgerufen am 27.09.2017 <http://altmetrics.org/manifesto/>

Ratti, E. (2016). The end of 'small biology'? Some thoughts about biomedicine and big science. *Big Data & Society*, 3 (2), 1–6

Scheer, A.-W. (2015). *Hochschule 4.0. Whitepaper No. 8*. Abgerufen am 27.09.2017 von <https://hochschulforumdigitalisierung.de/sites/default/files/dateien/Hochschule-4.0-Whitepaper-Professor-Scheer.pdf>

Spezi, V.; Wakeling, S.; Pinfield, S.; Creaser, C.; Fry, J. & P. Willett (2017). Open-access mega-journals: The future of scholarly communication or academic dumping ground? A review. *Journal of Documentation*, 73(2), 263–283

Stichweh, R. (1987). *Die Autopoiesis der Wissenschaft*. In: D. Baecker et al. (Hrsg.): *Theorie als Passion* (S. 447–481). Frankfurt am Main: Suhrkamp

Straub, M. C. P. (2016). Giving Citizen Scientists a Chance: A Study of Volunteer-led Scientific Discovery. *Citizen Science: Theory and Practice*, 1(1): 5, 1–10

Taylor, M. (2013). Exploring the boundaries. How altmetrics can expand our vision of scholarly communication and social impact. *Information Standards Quarterly*, 25, 27–32

Tollefson, J. (2018). China declared largest source of research articles. *Nature*, 553, 390

Van Noorden, R. (2014). Online collaboration: Scientists and the social network. *Nature*, 512 (7513), 126–129

Van Noorden, Richard (2017). Publishers threaten to remove millions of papers from ResearchGate. *Nature News* doi:10.1038/nature.2017.22793

Weingart, P. & Taubert, N. (Hrsg.) (2016). *Wissenschaftliches Publizieren: zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung*.

Berlin/Boston: De Gruyter Akademie Forschung. (Forschungsberichte/ Interdisziplinäre Arbeitsgruppen, Berlin-Brandenburgische Akademie der Wissenschaften; 38)

Wilkinson, M. D.; Dumontier, M.; Aalbersberg, J. J.; Appleton, G.; Axton, M.; Baak, A. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3, 160018 EP

Wouters, P. & Costas, R. (2012). *Users, narcissism and control – tracking the impact of scholarly publications in the 21st century*, Utrecht: SURFfoundation. Abgerufen am 27.09.2017 von <http://research-acumen.eu/wp-content/uploads/Users-narcissism-and-control.pdf>

Wouters, P. et al. (2015). *The Metric Tide: Literature Review (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management)*. HEFCE. DOI: 10.13140/RG.2.1.5066.3520

Manuskript eingereicht: 31.10.2017  
Manuskript angenommen: 22.06.2018

**Anschrift der Autorin:**

Dr. Martina Franzen  
Wissenschaftszentrum Berlin für Sozialforschung  
Forschungsgruppe Wissenschaftspolitik  
Reichpietschufer 50  
10785 Berlin  
E-Mail: [martina.franzen@wzb.eu](mailto:martina.franzen@wzb.eu)

Dr. phil. Martina Franzen ist wissenschaftliche Mitarbeiterin der Forschungsgruppe Wissenschaftspolitik am Wissenschaftszentrum Berlin für Sozialforschung (WZB). Ihr gegenwärtiger Forschungsschwerpunkt liegt auf den Implikationen einer fortschreitenden Datafizierung der Gesellschaft respektive der wissenschaftlichen Wissensproduktion.



# Mehr Mobilitätserfahrungen durch digitale Medien? Zu den Effekten von studentischer Diversität und Lernumweltsmerkmalen auf die internationale Mobilität

Anja Gottburgsen, Janka Willige

---

Trotz der hohen Erwartungen, durch Digitalisierung die Teilhabe einer diversen Studierendenschaft an hochschulischer Bildung zu erhöhen, sind die beiden Forschungsstränge „Diversität“ und „Digitalisierung“ bislang unverbunden. Im vorliegenden Beitrag werden diese auf theoretischer Ebene in einem Analysemodell zusammengeführt. Das Modell bildet die Grundlage für eine Sekundäranalyse von Befragungsdaten des DZHW-Online-Access-Panel „HISBUS“ (n = 4.375), in der untersucht wird, welche Effekte studentische Diversität einerseits und Charakteristika der Lernumwelt andererseits auf die studienbezogene internationale Mobilität (digitale und tatsächliche) von Studierenden haben. Die Ergebnisse belegen zum Teil deutliche Effekte auf die „internationale Mobilität“ entlang sozialer Merkmale der Studierenden (z. B. Gender, nicht-akademisches Elternhaus, Elternschaft) sowie lernumweltlicher Charakteristika (z. B. Hochschultyp, Studienfach, digitalisiertes Lernumfeld).

---

## 1 Einführung

Flexibilisierung, Mobilität und Individualisierung sind die zentralen Schlüsselemente der Europäischen Studienreform 2.0 (KMK/HRK, 2015-16). Eine umfassende Digitalisierung der studentischen Lernumwelten verspricht, ein wichtiger Baustein bei der Erreichung dieser Ziele zu sein. Durch den Einsatz digitaler Lern-/Lehrformate sollen nicht nur Chancen auf ein flexibles, zeit- und ortsunabhängiges Lernen der Einzelnen eröffnet, sondern es soll gleichzeitig auch die Internationalisierung der Hochschulen befördert werden (Wissenschaftsrat, 2018). Angesichts expandierender Studierendenzahlen (mit rd. 2,8 Millionen eingeschriebenen Studierenden im Wintersemester 2017/2018; vgl. Statistisches Bundesamt, 2018) verwundert es nicht, dass auch Hochschulleitungen und -verwaltungen hohe Erwartungen formulieren und günstige Effekte der Digitalisierung für die Inklusion vielfältiger Studierendengruppen, das individualisierte Lernen, den Lernerfolg oder auch die (internationale) Mobilität von Studierenden vermuten (Schmid et al., 2017, S. 24). Digitalisierung als Chance zur Verbesserung der Teilhabechancen diverser Studierendengruppen – seit 2009 unter dem Stichwort „soziale Dimension des Studiums“ auf der europäischen Agenda (Europäische Hochschulministerinnen/-minister, 2009; HRK, 2013, S. 21f) – wird heute

international und national propagiert (Dahlstrom & Bichsel, 2014; Henderson et al., 2015; Hochschulforum Digitalisierung, 2016; Mayrberger, 2016; Schmid et al., 2017). Trotz hoher Erwartungen und vermuteter Chancen durch Digitalisierung für die erfolgreiche Teilhabe vielfältiger Studierendengruppen an hochschulischer Bildung werden bislang jedoch die empirischen Forschungen zu Diversität und Digitalisierung nicht systematisch miteinander verknüpft.

Im vorliegenden Beitrag werden die jeweiligen Perspektiven beider Forschungsstränge im Rahmen einer Sekundäranalyse zusammengeführt. Es wird untersucht, ob die vielfältigen Merkmale von Studierenden auf individueller Ebene (Geschlecht, Migrationsstatus, Bildungsherkunft, Alter, Elternschaft u. a.) sowie die Charakteristika der Lernumwelt (Hochschultyp, Fächergruppen, digitalisiertes Lernumfeld) ihre digitale und auch tatsächliche studienbezogene internationale Mobilität beeinflussen.

Der Beitrag gliedert sich wie folgt: Auf Grundlage des Forschungsstandes der bisher unverbundenen Forschungsstränge zu Digitalisierung und Diversität (2) wird das empirische Untersuchung zugrundeliegende Analysemodell zum Einfluss von studentischer Diversität und digitalisierter Lernumwelt auf die internationale Mobilität (digital und tatsächlich) vorgestellt (3). Der Präsentation von Datengrundlage und Operationalisierungen (4) folgt die ausführliche Darstellung der Ergebnisse (5). In der abschließenden Zusammenfassung mit Ausblick werden entlang der Limitationen der vorliegenden Befunde weitere Forschungsbedarfe identifiziert (6).

## 2 Forschungsstand zu Digitalisierung und Diversität

International und national weit verbreitet sind Erhebungen zum Mediennutzungsverhalten von Studierenden (Dahlstrom & Bichsel, 2014; Brooks, 2016; Persike & Friedrich, 2016; Schmid et al., 2017; Willige, 2016; Zawacki-Richter et al., 2016; für ein Review siehe Steffens et al., 2017). Beispielsweise zeigt Willige (2016, S. 27f.) in ihrer Befragung von 4.375 Studierenden, dass diese an ihrer eigenen deutschen Hochschule sehr häufig einzelne digitale Lern- und Lehrelemente nutzen, wie z. B. online zur Verfügung gestelltes Lehrveranstaltungsbegleitendes Material (96 Prozent der Befragten) oder das Online-Portal ihrer Hochschule zur Studienorganisation (86 %) oder ihre Lektüre in digitaler Form lesen (78 Prozent). Rund ein Drittel der Befragten nutzen Online-Vorlesungen (30 Prozent), E-Lectures (28 Prozent) oder Online-Prüfungen und Online-Tests (26 Prozent). Deutlich weniger genutzt werden Online-Seminare im Rahmen eines regulären Studiengangs (14 Prozent), Online-Betreuung (12 Prozent), Mobiles Lernen (12 Prozent), Inverted Teaching/Flipped Classroom (12 Prozent) oder Online-Kurse außerhalb des Studiengangs (10 Prozent). Nur außerordentlich selten werden von den befragten Studierenden digitale Formate wie Game-Based-Learning und E-Portfolio

(jeweils 5 Prozent), Online-Praktikum/-Exkursion oder ein Online-Studiengang (jeweils 2 Prozent) angegeben.

In den Studien zum studentischen Mediennutzungsverhalten sind nur wenige Aussagen zum Einfluss individueller Merkmale von Studierenden auf ihr Mediennutzungsverhalten enthalten (Gender: Grosch &, Gideon 2011; Karapanos & Fendler, 2015; Alter: Lai & Hong, 2015). Die wachsende Zahl der Studierenden und ihre Diversität sind jedoch Faktoren, die bei der Gestaltung von Lehr-/Lernprozessen einbezogen werden müssen. Die Ergebnisse der 21. Sozialerhebung (Middendorff et al., 2017) belegen eine erhebliche Vielfalt der Studierenden. So sind aktuell 68 Prozent der Studierenden erwerbstätig, 48 Prozent der Studierenden verfügen über eine nicht-akademische Bildungsherkunft, 33 Prozent sind im Erststudium älter als 25 Jahre, 28 Prozent verfügen über weniger als den BAföG-Höchstsatz von 735 Euro/Monat, 20 Prozent haben selbst oder über ihre (Groß-)Eltern Migrationserfahrungen (und ihre Hochschulzugangsberechtigung in Deutschland erworben), elf Prozent der Studierenden geben eine Beeinträchtigung an, neun Prozent sind internationale Studierende mit einer nicht in Deutschland erworbenen Hochschulzugangsberechtigung (DAAD & DZHW, 2017) und sechs Prozent betreuen Kinder im eigenen Haushalt.

Zugangs- und Verbleibschancen strukturieren sich nicht nur entlang eines einzelnen sozialen Merkmals, sondern soziale Merkmale wirken sich in intersektionaler Verschränkung auf diese Chancen aus (Gross, Gottburgsen, & Phoenix, 2016; für den schulischen Kompetenzerwerb Gottburgsen & Gross, 2012; Gross & Gottburgsen, 2013; für die Teilhabe von wissenschaftlichen Beschäftigten an der Organisation Hochschule Buche & Gottburgsen, 2012). International (z. B. Alba & Waters, 2011; Heath & Brinbaum, 2016) wie national zeigen sich beispielsweise systematische Zusammenhänge zwischen Migrationserfahrung in Verbindung mit einer nicht-akademischen Bildungsherkunft und der Studiendauer (Autorengruppe Bildungsberichterstattung, 2016; Kristen, 2014) bzw. dem Studienabbruch (Ebert & Heublein, 2017). Aktuelle Daten zu internationalen Studierenden belegen im Bachelorstudium eine höhere Abbruchquote von 41 Prozent gegenüber 29 Prozent bei Studierenden mit deutscher Staatsangehörigkeit. Im Masterstudium liegt der Studienabbruch ebenfalls über dem Niveau der inländischen Studierenden (Heublein et al., 2017, S. 261ff).

Dass digitalisierte Lernumwelten der diversen Zusammensetzung der Studierendenschaft sowie ihren daraus resultierenden sehr unterschiedlichen Lebenslagen, ihren differierenden Lernvoraussetzungen und -erfahrungen besser gerecht werden und somit Lernerfolge unterstützen können, ist eine international und national häufig propagierte Erwartung (Dahlstrom & Bichsel, 2014; Henderson, Selwyn, & Aston, 2015; Hochschulforum Digitalisierung, 2016; Schmid et al., 2017). Belegt ist, dass durch die mit dem Einsatz digitaler Lernformate einhergehende Flexibilisierung von

Zeit und Lernort die Vereinbarkeit des Studiums mit Erwerbstätigkeit und Care-Aufgaben erhöht (Zawacki-Richter, 2015), ein selbstbestimmtes Lerntempo gemäß individuellen Lernvoraussetzungen und Bedürfnissen fördert (Mayrberger, 2016; mit Bezug auf Beeinträchtigung siehe Seale, Georgeson, Mamas, & Swain, 2015; van Rooij & Zirkle, 2016) und die Qualität und Häufigkeit der Interaktion von Lehrenden und Lernenden sowie der Studierenden untereinander steigert (Willige, 2016; Henderson et al., 2015).

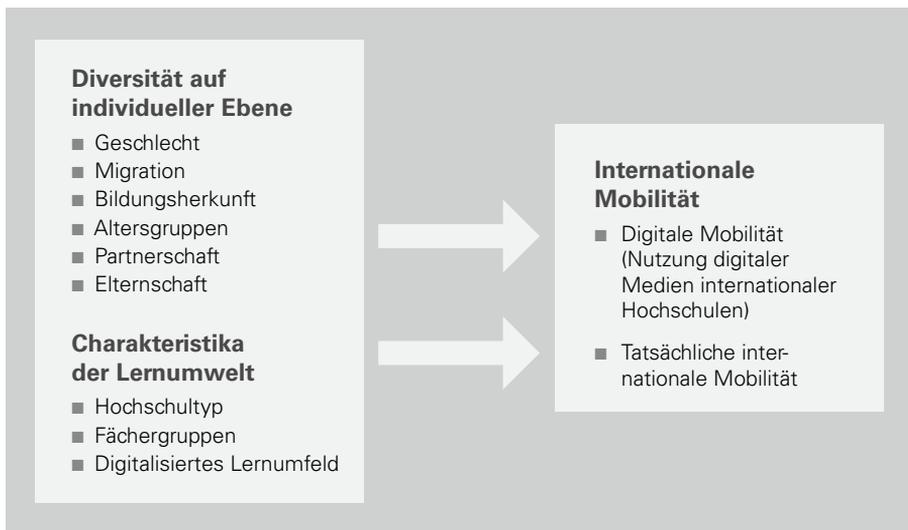
Bezüglich der Lernerfolge in digitalen Lernformaten fallen die Befunde uneinheitlich aus: Während Means et al. (2013) in ihrer Meta-Analyse von 22 Studien feststellen, dass Studierende in Blended-Learning-Formaten deutlich besser abschneiden als in Face-to-Face-Lernformaten (nicht jedoch in reinen Online-Formaten im Vergleich zu Face-to-Face-Lernformaten), finden Bowen et al. (2014) keine Unterschiede in den Lernerfolgen. Jedoch benötigten die Lernenden in Blended-Lernformaten zum Erreichen ihrer Ergebnisse 25 Prozent weniger Zeit im Vergleich zu Face-to-Face-Lernformaten. Zudem ließen sich Dropout-Raten durch Blended Learning im Vergleich zu reinen Online-Lernformaten reduzieren (López-Pérez et al., 2011; Porter et al. 2014). Auffällig ist auch hier der „blinde Fleck“ bezüglich der vielfältigen sozialen Merkmale der Studierenden. Ob und inwieweit die hohen Dropout-Raten in reinen Online-Formaten wie beispielsweise in MOOCs (Eriksson et al., 2017) oder in Studienangeboten von Distance-Learning-Universitäten (wie z. B. British Open University, Spanish National Distance Teaching University, Athabasca University in Canada; vgl. Simpson, 2013; Gregori et al., 2018) in Zusammenhang mit den sozialen Merkmalen der Studierenden stehen, untersuchen unseres Wissens erstmals Stoessel et al. (2015). Sie fanden in ihrer Analyse von Studierendendaten der Fernuniversität Hagen (n= 4.599), dass das Risiko eines Studienabbruchs für Frauen, für Studierende mit Migrationserfahrungen und für vollzeiterwerbstätige Studierende höher ist (im Vergleich zur jeweiligen Referenzgruppe), dass aber ältere Studierende und Studierende mit Kindern ein geringeres Risiko aufweisen, ihr Studium abzubrechen.

Zusammenfassend lässt sich festhalten: Nicht nur fehlen strukturierte Vergleiche der Mediennutzungsmuster verschiedener studentischer Zielgruppen. Auch die systematische Überprüfung, wie die vielfältigen sozialen Merkmale der Studierenden zusammen mit einer digitalisierten Lernumwelt ihre Lernerfolge und weitere Learning Outcomes (wie z. B. ihre internationale Mobilität) beeinflussen, steht noch aus (Lack, 2015; Sclater et al. 2016; Stoessel et al., 2015).

### 3 **Analysemodell zu den Effekten studentischer Diversität und den Charakteristika der Lernumwelt auf eine studienbezogene internationale Mobilität**

Das der sekundäranalytischen Auswertung zugrundeliegende Analysemodell (Abbildung 1) postuliert, dass die sozialen Merkmale der Studierenden auf individueller Ebene gemeinsam mit den Charakteristika der Lernumwelt die Learning Outcomes (Kompetenzen, Lernerfolge, Zertifikate etc.) beeinflussen. Internationale Mobilität wird im Modell als ein Learning Outcome verstanden, also als das mögliche und gewünschte Ergebnis eines Lernprozesses, zu dem die Ausgestaltung der Lernumwelt ganz maßgeblich beiträgt. Ausgegangen wird davon, dass die Diversitätsmerkmale in der Studierendenschaft, wie sie heute im Kontext Hochschule als zentral gelten (Museus & Griffin, 2011; Gottburgsen & Arbeitskreis, 2015; Middendorff et al., 2017; Leicht-Scholten, 2012) und die in der bisherigen Diversitätsforschung als relevant identifiziert wurden (vgl. Gross, Gottburgsen, & Phoenix, 2016), im Kontext unterschiedlicher Lernumwelten differierende Effekte auf die internationale Mobilität von Studierenden haben (zum Effekt der Bildungsherkunft Lörz et al., 2016; Netz, 2015). Die Diversität *auf individueller Ebene* wird erfasst über die sozialen Merkmale Geschlecht, Migration, Bildungsherkunft, Zugehörigkeit zu verschiedenen Altersgruppen, Partnerschaft und Elternschaft.

**Abbildung 1:** Analysemodell „Effekte von studentischer Diversität und von Charakteristika der Lernumwelt auf studienbezogene internationale Mobilität“



Quelle: eigene Darstellung

Dass die Charakteristika der *Lernumwelten* die Learning Outcomes von Studierenden beeinflussen können, ist gut dokumentiert (Braun et al. 2014; Schneider & Preckel, 2017; Schaeper & Weiß, 2016). Wie in Abbildung 1 veranschaulicht, wird die Teilhabe

an internationalen Mobilitätserfahrungen der Studierenden als potenzielles Learning Outcome ebenfalls durch die Merkmale der jeweiligen Lernumwelt mit beeinflusst. Verschiedene Lernumwelten sind internationaler Mobilität jedoch in unterschiedlichem Maße förderlich. Wie Woisch & Willige (2015, S. 8f) in ihrer Untersuchung der Mobilitätsquoten von Studierenden zeigen, variieren diese hinsichtlich sowohl des besuchten Hochschultyps (Universitäten bzw. Fachhochschulen/Hochschulen für Angewandte Wissenschaften) als auch der Fachausrichtung des Studiengangs. Ein weiteres Charakteristikum der Lernumwelt stellt das digitale Lern- und Lehrumfeld des Studiengangs dar. Digitale Elemente, spezifische Formate wie auch Online-Studiengänge können u. a. die digitale „Vernetzung“ mit Studierendengruppen im In- wie im Ausland befördern, sie können internationale Erfahrungen ermöglichen (auch als Ersatz für tatsächliche physische Mobilität), und sie können der Vor- und Nachbereitung von studienbezogenen Auslandsaufenthalten dienen (vgl. Wissenschaftsrat, 2018, S. 97). Das digitale Lern- und Lehrumfeld im Studiengang wird für die folgende Analyse wie folgt bestimmt (ausführlicher dazu in Abschnitt 4; siehe auch Wannemacher, 2016): Ein sogenanntes „grundständiges“ digitales Lern- und Lehrumfeld liegt vor, wenn nur einzelne digitale Lehr- und Lernelemente zum Einsatz kommen (z. B. Lehrveranstaltungsbegleitendes Material online, wie beispielsweise Skripts, Protokolle). Darüber hinaus können spezifische digitale Lern- und Lehrformate, wie beispielsweise Inverted Teaching/Flipped Classroom (Stoffaneignung im Online-Selbststudium und anschließende Vertiefung im Präsenzstudium), hinzukommen. Ergänzend finden sich in geringerem Umfang eigenständige digitale Lehrveranstaltungen und Studiengänge. Es ist davon auszugehen, dass ein digitales Lern- und Lehrumfeld der Hochschule bzw. des Studiengangs, das neben der Basisausstattung mit digitalen Elementen spezifischere digitale Formate und sogar eigenständige digitale Studiengänge bereitstellt, günstige Bedingungen für digitale wie auch faktisch-räumliche internationale Mobilität schafft. Dem Modell liegt damit ein weiter Begriff von „Mobilität“ zugrunde, demzufolge sowohl die digitale Nutzung von Angeboten internationaler Hochschulen als auch tatsächliche studienbezogene Auslandsaufenthalte als internationale Mobilität aufgefasst werden.

#### **4 Daten und Operationalisierungen**

Zur Überprüfung der Frage, ob und inwieweit soziale Merkmale auf individueller Ebene (Geschlecht, Migration, Bildungsherkunft, Alter, Partner- und Elternschaft) sowie Charakteristika der Lernumwelt (Hochschultyp, Fachausrichtung, digitalisiertes Lernumfeld) die digitale und tatsächliche internationale Mobilität der befragten Studierenden beeinflussen, werden die im Rahmen des Projekts „Auslandsmobilität und digitale Medien“ (Willige, 2016) mit dem DZHW-Online-Access-Panel „HISBUS-Studierendenpanel“ im Frühjahr 2016 generierten Daten sekundäranalytisch ausgewertet.

## 4.1 Beschreibung der Stichprobe

Der zugrundeliegende Datensatz enthält 4.375 Fälle, von denen der überwiegende Anteil das Merkmal „weiblich“ aufweist (61 Prozent weiblich, 39 Prozent männlich). Nur ein Prozent der Fälle hat eine nicht-deutsche Staatsangehörigkeit und die Hochschulzugangsberechtigung (HZB) in Deutschland erworben (sogenannte Bildungsinländerinnen bzw. -inländer). Die Variable „Bildungsinland“ wird trotz der geringen Fallzahl in die Analyse aufgenommen, da es sich um das einzige verfügbare Merkmal im Datensatz zur Abbildung des Migrationshintergrunds handelt. Weitere differenziertere Aussagen zu familiären Migrationserfahrungen beispielsweise über die (Groß-) Eltern erlaubt der Datensatz nicht. Aufgrund der außerordentlich geringen Fallzahl sind die Ergebnisse zum Merkmal „Migrationshintergrund“ entsprechend nur eingeschränkt interpretierbar und ihre Aussagekraft nur von begrenzter Gültigkeit. Die Mehrheit der Fälle in der Stichprobe hat ein akademisches Elternhaus (57 Prozent vs. 43 Prozent ohne akademische Bildungsherkunft). Der größte Teil der Angaben stammt von Studierenden, die bis zu 30 Jahre alt sind (38 Prozent sind bis 25 Jahre alt, 42 Prozent sind 26 bis 30 Jahre, 20 Prozent älter als 31 Jahre). 93 Prozent der Fälle in der Stichprobe haben keine Kinder, 60 Prozent leben in einer Partnerschaft. 77 Prozent der befragten Studierenden sind an einer Universität eingeschrieben. Die Verteilung auf die Fächergruppen stellt sich wie folgt dar: 30 Prozent studieren einen Studiengang innerhalb der Fächergruppe Sprach- und Kulturwissenschaften, 23 Prozent in der Fächergruppe Mathematik und Naturwissenschaften, 13 Prozent in den Sozialwissenschaften und Sozialwesen, 14 Prozent in den Ingenieurwissenschaften, neun Prozent in den Wirtschaftswissenschaften, sieben Prozent in der Medizin und vier Prozent in den Rechtswissenschaften.

## 4.2 Methode und Operationalisierungen

Wie im Analysemodell (Abbildung 1) dargestellt, werden für die abhängigen Variablen digitale und tatsächliche internationale Mobilität voneinander getrennte Regressionsmodelle berechnet. Dem Analysemodell entsprechend werden in beiden Modellen eine Reihe von *unabhängigen Variablen* zur Diversität auf individueller Ebene sowie zur Lernumwelt berücksichtigt.

Für die Beschreibung der *Diversität auf individueller Ebene* werden in die Analysen die folgenden Variablen einbezogen: neben dem *Geschlecht* (dichotom, Referenz weiblich) das Merkmal *Bildungsinland* (dichotom, Referenz Studierende mit deutscher Staatsangehörigkeit) sowie die *Bildungsherkunft* der Studierenden (dichotom, Referenz Studierende aus akademischem Elternhaus, d. h. mindestens ein Elternteil verfügt über einen akademischen Abschluss). Die kategoriale Variable *Altersgruppe* weist drei Gruppen aus, zunächst die Studierendengruppe bis 25 Jahre (Referenz), dann eine

mittlere Gruppe von 26 bis 30 Jahren sowie die Gruppe der Personen, die 31 Jahre und älter sind. Des Weiteren geht die (vorhandene bzw. nicht vorhandene) *Partnerschaft* der Studierenden ein (dichotome Variable, Referenzkategorie Studierende mit Partnerschaft). Für die Bildung der Variable *Elternschaft* wurde aus den Angaben zu Kindern eine dichotome Variable gebildet; die Referenzgruppe bilden hier Studierende, die keine Kinder haben. Als *Faktoren der Lernumwelt* gehen die *Hochschulart* (dichotom, Referenz Universitäten) sowie eine Fächergruppendifferenzierung orientiert an der Kategorisierung des Statistischen Bundesamts in die Analyse ein; als Referenzkategorie werden die Wirtschaftswissenschaften gesetzt, die anderen Fächergruppen sind die Sprach-, Kulturwissenschaften, die Rechtswissenschaften, die Sozialwissenschaften, Sozialwesen, Psychologie, Pädagogik, die Mathematik, Naturwissenschaften, die Medizin und Gesundheitswissenschaften sowie die Ingenieurwissenschaften. Das *digitale Lern- und Lehrumfeld im Studiengang* geht mit einer Variable mit drei Ausprägungen in die Analyse ein. Dafür werden die Angaben der Studierenden zum Vorhanden- bzw. Nicht-Vorhandensein digitaler Medien im Lern- und Lehrumfeld des Studiengangs wie folgt kategorisiert: Der Referenzwert steht für ein grundständiges digitalisiertes Lern- und Lehrumfeld im Studiengang. Zu dieser Basisausstattung im Lern- und Lehrumfeld zählen digitale *Lern- und Lehrelemente*, wie z. B. online gestellte Lehrveranstaltungs begleitende Materialien, Online-Vorlesungen, soziale Medien sowie Blogs Lehrender. Eine weitere Kategorie umfasst die Angaben zu spezifischen digitalen *Lehr- und Lernformaten*, wie z. B. Inverted Teaching, Flipped Classroom und Game-Based Learning. Angaben zu eigenständigen digitalen *Lehrveranstaltungen und Studiengängen* werden als weitere Kategorie aufgefasst, zu nennen sind z. B. Open Course bzw. MOOC, Online-Praktikum, Online-Exkursion, Online-Studiengang (vgl. zum Erhebungsinstrument Willige, 2016; zur Differenzierung Handke, 2015).

Die zwei eingesetzten *abhängigen Variablen digital international mobil* und *tatsächliche studienbezogene Auslandsmobilität* wurden wie folgt operationalisiert: Auf Basis der Mehrfachangaben zur Frage, welche digitalen Medien und Anwendungen aus dem Angebot ausländischer Hochschulen genutzt wurden, wurde erfasst, ob Studierende *digital international mobil* sind. In der Regressionsanalyse steht der Wert 1 für die Nutzung digitaler Medien im Angebot ausländischer Hochschulen, die 0 für die Nicht-Nutzung. Die *tatsächliche studienbezogene Auslandsmobilität* wird über eine dichotome Variable erhoben; der Wert 1 steht für studienbezogene Auslandsmobilität im bisherigen Studienverlauf (z. B. ein durchgeführtes Auslandssemester, eine Studienreise, Praktikum, Sprachkurs, Summer School), die Nicht-Mobilität ist mit dem Wert 0 codiert.

## 5 Ergebnisse

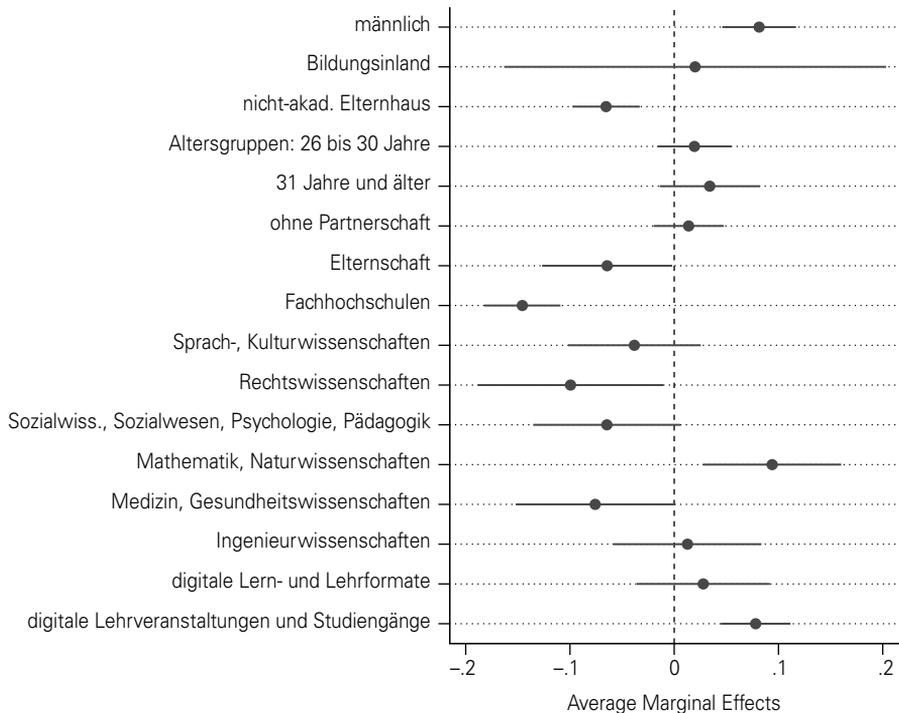
Im Folgenden werden die Ergebnisse der Berechnungen von zwei logistischen Regressionsmodellen vorgestellt, um aufzuzeigen, welche Diversitätsmerkmale und welche Merkmale der Lernumwelt einen Einfluss auf digitale internationale Mobilität und die tatsächliche Mobilität haben. Mittels der Modelle wird jeweils untersucht, welchen Einfluss die einzelnen unabhängigen Variablen haben, wenn parallel dazu alle anderen Faktoren kontrolliert werden. Im folgenden Ergebnisteil werden die Modelle grafisch dargestellt, interpretiert werden die Average Marginal Effects (AME). Die Nennung aller Befunde (Odds-Ratios/OR, Average Marginal Effects/AME und Predictive Margins/PM) erfolgt in der detaillierten Darstellung im Anhang. Die AME lassen sich anschaulich als *Prozentpunktdifferenz* zwischen der Wahrscheinlichkeit des Eintritts eines Ereignisses im Vergleich zur jeweils gesetzten Referenzkategorie beschreiben.

### 5.1 Digitale internationale Mobilität

Digitale internationale Mobilität erfasst die Angaben Studierender zur Nutzung digitaler Angebote von Hochschulen im Ausland. Digital mobil können Studierende von der heimischen Hochschule aus sein – ohne tatsächlich studienbezogen ins Ausland zu reisen. Für die individuelle Ebene (Abbildung 2) erweist sich von den untersuchten Diversitätsmerkmalen im Modell das Geschlecht als bedeutsam: Die befragten Studenten geben signifikant häufiger als die Studentinnen an, digital international mobil zu sein (AME: 0,08). Ein nicht-akademisches Elternhaus hat einen signifikant negativen Effekt auf die digitale internationale Mobilität (AME:-0,07). In Bezug auf die Altersgruppen der älteren Studierenden finden sich keine signifikanten Unterschiede. Ebenso finden sich für den Status „Partnerschaft“ keine Effekte auf die digitale internationale Mobilität, wohingegen die Elternschaft die Wahrscheinlichkeit für digitale internationale Mobilität deutlich verringert (AME: -0,06). Im Bereich der Lernumwelt erweist sich der Hochschultyp als einflussstark. Es zeigt sich, dass Studierende an Fachhochschulen signifikant seltener digital international mobil sind als Studierende an Universitäten (die Differenz beträgt 14 Prozentpunkte). Die Zugehörigkeit zu unterschiedlichen Fächergruppen scheint ebenfalls eine Rolle im Zusammenhang mit digitaler internationaler Mobilität zu spielen: Negative Effekte im Vergleich zur Referenz der Wirtschaftswissenschaften finden sich für die Fächergruppe Rechtswissenschaften (AME: -0,10), positive Effekte finden sich für die Fächergruppe Mathematik, Naturwissenschaften (AME: 0,09). Die Effekte eines digitalisierten Lern- und Lernumfeldes auf die digitale internationale Mobilität sind gering. Es zeigt sich jedoch, dass Studierende, die eigenständige digitale Lehrveranstaltungen und Studiengänge vorfinden, eine signifikant höhere

Wahrscheinlichkeit aufweisen (AME:0,08), digitale Angebote internationaler Hochschulen zu nutzen, also digital international mobil zu werden, als Studierende, die lediglich in einem grundständig digitalen Lern- und Lehrumfeld studieren.

**Abbildung 2:** Digitale internationale Mobilität



Anmerkungen:

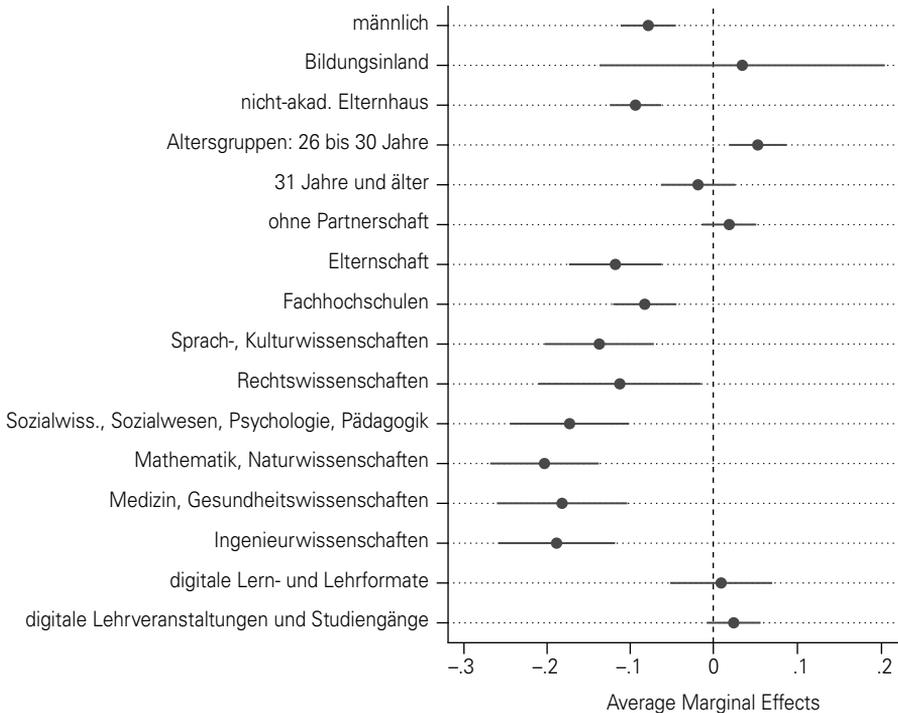
Average Marginal Effects einer logistischen Regression, 95%-Konfidenzintervall, d.h. die Irrtumswahrscheinlichkeit für ein signifikantes Ergebnis ist kleiner als 5%. Die horizontalen Linien geben das Konfidenzintervall an, in dem sich der Schätzer für die einzelnen Variablen befindet. Rechts der vertikalen 0-Achse liegende Intervallschätzer zeigen signifikant positive, links befindliche Schätzer signifikant negative Effekte an. Schneidet der Schätzer die Null-Linie liegen keine signifikanten Unterschiede vor. Referenzen (in Klammern) sind jeweils: männlich (*weiblich*); Bildungsinland (*deutsche Studierende*); nicht-akad. Elternhaus (*akad. Elternhaus*); Altersgruppen (*bis 25 Jahre*); ohne Partnerschaft (*mit Partnerschaft*); Elternschaft (*ohne Kinder*); Fachhochschulen (*Universitäten*); Sprach- und Kulturwissenschaften bis Ingenieurwissenschaften (*Wirtschaftswissenschaften*); digitale Lern- und Lehrformate und digitale Lehrveranstaltungen und Studiengänge (*grundständiges digitales Lernumfeld im Studiengang*)

## 5.2 Tatsächliche internationale Mobilität

In einem weiteren Modell wurde untersucht, inwieweit die verschiedenen Merkmale der Diversität auf individueller Ebene sowie Merkmale der Lernumwelt Effekte auf die Wahrscheinlichkeit haben, tatsächlich einen studienbezogenen Auslandsaufenthalt

durchzuführen (Abbildung 3). Die in das Modell aufgenommenen Merkmale stellen in Bezug auf die generell bedingenden Faktoren für studienbezogene Auslandsmobilität eine Auswahl dar (siehe ausführlicher zu den verschiedenen Faktoren Willige, Sudheimer, & Grützmaker, 2017; Woisch & Willige, 2015). Im untersuchten Modell erweist sich auf individueller Ebene das Geschlecht als bedeutsam für die Durchführung studienbezogener Auslandsmobilität: Studenten weisen im Vergleich zu Studentinnen eine signifikant geringere Wahrscheinlichkeit auf, auslandsmobil zu werden (AME:  $-0,08$ ). Bildungsinländerinnen und Bildungsinländer haben eine im Vergleich zu Studierenden mit deutscher Staatsangehörigkeit um drei Prozentpunkte höhere Wahrscheinlichkeit, tatsächlich im Verlaufe ihres Studiums auslandsmobil zu werden (AME:  $0,03$ ). Allerdings ist dieser Befund nicht signifikant und aufgrund der geringen Fallzahl lediglich ein erster Hinweis, der durch weitergehende Studien, die zudem den Migrationshintergrund differenzierter erheben müssten, zu prüfen. Von den weiteren Merkmalen auf individueller Ebene erweisen sich ein nicht-akademisches Elternhaus (AME:  $-0,09$  sowie die Elternschaft (AME:  $-0,12$ ) als hemmende Faktoren für tatsächliche Auslandsmobilität. Der Effekt der Bildungsherkunft ist signifikant, ebenso der der Elternschaft. Eine im Vergleich zu Studierenden der Altersgruppe bis 25 Jahre um fünf Prozentpunkte höhere Wahrscheinlichkeit, im Studium auslandsmobil zu werden, findet sich für die „mittlere“ Altersgruppe der 26 bis 30 Jahre alten Studierenden, kein signifikanter Unterschied zwischen den 31 Jahre und älteren Studierenden (AME:  $-0,02$ ) und der jüngsten Gruppe.

Von den Faktoren der Lernumwelt hat die besuchte Hochschulart einen recht deutlichen Effekt auf das tatsächliche Mobilitätsverhalten; Studierende an Fachhochschulen haben eine signifikant geringere Wahrscheinlichkeit, auslandsmobil zu werden, als Studierende an Universitäten (AME:  $-0,08$ ). Studierende aller untersuchten Fächergruppen weisen im Vergleich zur Referenzkategorie der Wirtschaftswissenschaften eine signifikant geringere Wahrscheinlichkeit auf, auslandsmobil zu werden (die Prozentpunktedifferenzen betragen bis zu 20 Prozentpunkte, etwa für die Studierenden in der Fächergruppe Mathematik und Naturwissenschaften). Ein digitales Lern- und Lehrumfeld im eigenen Studiengang mit spezifischen digitalen Lern- und Lehrformaten und auch ein Angebot an eigenständigen digitalen Lehrveranstaltungen und Studiengängen hat dagegen im Vergleich zur Referenzgruppe, die über ein grundständiges digitales Lern- und Lehrumfeld verfügt, keinen signifikanten Effekt auf die Durchführung studienbezogener Auslandsaufenthalte.

**Abbildung 3:** Tatsächliche internationale Mobilität

Anmerkungen:

Average Marginal Effects einer logistischen Regression, 95%-Konfidenzintervall, d.h. die Irrtumswahrscheinlichkeit für ein signifikantes Ergebnis ist kleiner als 5%. Die horizontalen Linien geben das Konfidenzintervall an, in dem sich der Schätzer für die einzelnen Variablen befindet. Rechts der vertikalen 0-Achse liegende Intervallschätzer zeigen signifikant positive, links befindliche Schätzer signifikant negative Effekte an. Schneidet der Schätzer die Null-Linie liegen keine signifikanten Unterschiede vor. Referenzen (in Klammern) sind jeweils: männlich (*weiblich*); Bildungsinland (*deutsche Studierende*); nicht-akad. Elternhaus (*akad. Elternhaus*); Altersgruppen (*bis 25 Jahre*); ohne Partnerschaft (*mit Partnerschaft*); Elternschaft (*ohne Kinder*); Fachhochschulen (*Universitäten*); Sprach- und Kulturwissenschaften bis Ingenieurwissenschaften (*Wirtschaftswissenschaften*); digitale Lern- und Lehrformate und digitale Lehrveranstaltungen und Studiengänge (*grundständiges digitales Lernumfeld im Studiengang*)

## 6 Zusammenfassung und Ausblick

Ergeben sich durch digitale Medien an der Hochschule neue Perspektiven für mehr (digitale und tatsächliche) internationale Mobilität für Studierende? Diese Frage wurde in diesem Beitrag im Hinblick auf die Effekte ausgewählter sozialer Merkmale sowie verschiedener Merkmale der Lernumwelt der Hochschule auf die digitale internationale Mobilität und auf die tatsächliche Auslandsmobilität von Studierenden untersucht.

Für die Diversitätsmerkmale Geschlecht, Bildungsinland, Bildungsherkunft, Altersgruppenzugehörigkeit, Partnerschaft und Elternschaft zeigen sich folgende Effekte: *Männliche Studierende* haben eine höhere Wahrscheinlichkeit, Angebote digitaler internationaler Mobilität zu nutzen. Dagegen zeigt sich für sie ein negativer Effekt bei

der Wahrscheinlichkeit zur Durchführung tatsächlicher studienbezogener Auslandsaufenthalte. Die Effekte des Geschlechts sind statistisch signifikant. Für das Merkmal *Bildungsinland* als das im Datensatz einzige zur Verfügung stehende Merkmal im Zusammenhang „Migrationshintergrund“ zeigt sich kein Effekt für Bildungsinländerinnen und Bildungsinländer für die Teilhabe an digitaler und tatsächlicher Mobilität. Die Befunde sind nicht signifikant und zudem aufgrund der geringen Fallzahl und deutlichen Streuung der Werte nur eingeschränkt zu interpretieren, sie sind allenfalls ein erster Hinweis auf die Notwendigkeit vertiefter, weiterer Analysen mit anderen Datensätzen – zumal diese Gruppe im Kontext der Diskussion über Diversität, Aspekte der Digitalisierung und Teilhabe an Auslandsmobilität von zentraler Bedeutung ist. Für die *Bildungsherkunft* finden sich in Bezug auf digitale internationale Mobilität sowie die Durchführung tatsächlicher studienbezogener Auslandsaufenthalte statistisch signifikante negative Effekte für Studierende aus einem nicht-akademischen Elternhaus. Positiv signifikante Effekte des *Alters* finden sich lediglich im Zusammenhang mit der Durchführung studienbezogener Auslandsaufenthalte für die Studierenden der mittleren im Vergleich zur jüngeren Altersgruppe. Dagegen zeigen sich für das Merkmal *Partnerschaft* tendenziell positive (jedoch nicht signifikante) Effekte für ungebundene Studierende – ohne Partnerschaft – im Hinblick auf digitale und tatsächliche Mobilität. Wie zu erwarten, erweist sich das Merkmal *Elternschaft* als hemmender Faktor für digitale und tatsächliche Mobilität, Studierende mit zu betreuenden Kindern im eigenen Haushalt haben eine geringere Wahrscheinlichkeit, digital wie tatsächlich Auslands-mobilitätserfahrungen zu sammeln als Studierende ohne Kinder.

Die Ergebnisse zur Diversität auf individueller Ebene weisen auf die Notwendigkeit hin, in einer von den Autorinnen geplanten Analyse von Interaktionseffekten zu prüfen, ob und inwieweit die verschiedenen sozialen Merkmale in intersektionaler Verschränkung (Gross, Gottburgsen, & Phoenix, 2016) die digitale internationale Mobilität sowie die tatsächliche Auslandsmobilität von Studierenden beeinflussen.

Für alle Merkmale der Lernumwelt – die Hochschulart, die Fächergruppen wie auch das digitale Lern- und Lehrumfeld – finden sich im Vergleich zur jeweiligen Referenzgruppe vorwiegend negative Effekte sowohl auf die digitale internationale Mobilität als auch die tatsächliche Auslandsmobilität. Bei Teilhabe an digitaler internationaler Mobilität lassen sich lediglich für die Fächergruppen Rechtswissenschaften signifikante negative sowie für Mathematik und Naturwissenschaften signifikante positive Effekte im Vergleich zu den Wirtschaftswissenschaften feststellen. Bei der konkreten Durchführung studienbezogener Auslandsaufenthalte finden sich für alle Fächergruppen signifikant negative Effekte im Vergleich zur Referenz der Wirtschaftswissenschaften. Das digitale Lern- und Lehrumfeld geht über verschiedene Abstufungen in die Analyse ein: Es zeigt sich, dass spezifische digitale Lern- und Lehrformate wie auch eigenständige digitale Lehrveranstaltungen im Vergleich zu einer „grundständigen“ Basis-

ausstattung keine Effekte auf digitale und tatsächliche internationale Mobilität haben. Eine Ausnahme stellen eigenständige digitale Lehrveranstaltungen und Studiengänge dar, sie zeigen einen signifikant positiven Effekt auf digitale internationale Mobilität.

Dass digitale Angebote im Lern- und Lehrumfeld vor allem auch im administrativen Bereich die Bereitschaft zu studienbezogenen Auslandsaufenthalten erhöhen könnten, belegen deskriptive Befunde (Willige, 2016). Danach geben 68 Prozent der Studierenden an, dass ihre Mobilitätsbereitschaft in hohem Maße durch eine Online-Datenbank mit hochschulgenauen Informationen über die Anerkennungspraxis der jeweiligen Studienaufenthalte an der eigenen Hochschule gesteigert würde. Für 45 Prozent würde die Option einer Online-Prüfungsteilnahme an der Heimathochschule während eines Aufenthalts ihre Mobilitätsbereitschaft erhöhen. Immerhin noch 37 Prozent sprechen Online-Formaten zur Veranstaltungsteilnahme an der Heimathochschule während des Auslandsaufenthalts eine solche Wirkung zu.

Die Ergebnisse unterstreichen die Fruchtbarkeit des entwickelten Analysemodells, das die Diversität der Studierenden auf individueller Ebene gemeinsam mit den Charakteristika ihrer Lernumwelten analysiert. Erstmalig werden zum Teil sehr deutliche Effekte auf das Learning Outcome *internationale Mobilität* entlang sozialer sowie lernumweltlicher Merkmale belegt. Die sich auf Basis des Modells ergebenden systematischen Vergleiche zeichnen ein bereits recht detailliertes Bild im Vergleich zur bisherigen Forschung im Feld „Digitalisierung“. Ergänzt werden sollte dieses jedoch durch eine differenzierte Prüfung potenzieller Interaktionseffekte, wie sie eine intersektionale Perspektive nahelegt. Künftige Forschung zu den Chancen der Digitalisierung sollte insbesondere darauf ausgerichtet werden, die Learning Outcomes (Kompetenzen, Lernerfolge, Zertifikate etc.) verschiedener studentischer Zielgruppen in verschiedenen digitalen Lernformaten zu untersuchen.

## Literatur

Alba, R. D., & Waters, M. C. (2011). *The next generation: Immigrant youth in a comparative perspective*. New York: New York University

Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bildung in Deutschland: Vol. 2016. Bielefeld: W. Bertelsmann Verlag

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2014). Interactive Learning Online at Public Universities: Evidence from a Six-Campus Randomized Trial. *Journal of Policy Analysis and Management*, 33(1), 94–111

Braun, E., Weiß, T., & Seidel, T. (2014). Lernumwelten in der Hochschule. In A. Krapp & T. Seidel (2014) (Hrsg.), *Pädagogische Psychologie*. 6., vollst. überarb. Aufl. (S. 433–454). Weinheim u.a.: Beltz

Brooks, D.C. (2016). *ECAR Study of Undergraduate Students and Information Technology*. Research report. Louisville, CO: Abgerufen am 28.08.2018 von [https://ccit.clemson.edu/wp-content/uploads/2016/11/ERS1605\\_ECAR\\_STUDENT\\_SUMMARY\\_2016.pdf](https://ccit.clemson.edu/wp-content/uploads/2016/11/ERS1605_ECAR_STUDENT_SUMMARY_2016.pdf)

Buche, A., & Gottburgsen, A. (2012). Migration, soziale Herkunft und Gender: „Intersektionalität“ in der Hochschule. In Pielage, P., Pries, L., & Schultze, G. (Hrsg.), *Soziale Ungleichheit in der Einwanderungsgesellschaft. Kategorien, Konzepte, Einflussfaktoren*. Tagungsdokumentation im Auftrag der Abteilung Wirtschafts- und Sozialpolitik der Friedrich-Ebert-Stiftung (WISO Diskurs) (S. 113–126). Bonn: Friedrich-Ebert-Stiftung

DAAD, & DZHW (2017). *Wissenschaft Weltoffen 2017: Daten und Fakten zur Internationalität von Studium und Forschung in Deutschland*. Fokus: Akademische Mobilität und Kooperation im Ostseeraum. Bielefeld: Bertelsmann Verlag

Dahlstrom, E., & Bichsel, J. (2014). *ECAR Study of Undergraduate Students and Information Technology, 2014*. Research report. Louisville, CO: ECAR. Abgerufen am 31.08.2018 von <https://library.educause.edu/~media/files/library/2014/10/ers1406-pdf.pdf?la=en>

Ebert, J., & Heublein, U. (2017). *Ursachen des Studienabbruchs bei Studierenden mit Migrationshintergrund*. Hannover

Eriksson, T., Adawi, T., & Stöhr, C. (2017). „Time is the bottleneck“. A qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education*, 29(1), 133–146

Europäische Hochschulministerinnen/-minister (2009). *The Bologna Process 2020 -The European Higher Education Area in the new decade*. Abgerufen am 27.08.2018 von [http://www.ehea.info/media.ehea.info/file/2009\\_Leuven\\_Louvain-la-Neuve/06/1/Leuven\\_Louvain-la-Neuve\\_Communique\\_April\\_2009\\_595061.pdf](http://www.ehea.info/media.ehea.info/file/2009_Leuven_Louvain-la-Neuve/06/1/Leuven_Louvain-la-Neuve_Communique_April_2009_595061.pdf)

Gottburgsen, A., & Arbeitskreis (2015). *Empfehlungen zur Förderung einer gender- und diversitätssensiblen Lehr- und Lernkultur*. Working Paper. Erlangen-Nürnberg: Büro für Gender und Diversity. Abgerufen am 31.08.2018 von [https://www.gender-und-diversity.fau.de/files/2017/11/2015\\_0405\\_empfehlungen\\_gender\\_diversittsensibile\\_lehrlernkultur.pdf](https://www.gender-und-diversity.fau.de/files/2017/11/2015_0405_empfehlungen_gender_diversittsensibile_lehrlernkultur.pdf)

Gottburgsen, A., & Gross, C. (2012). Welchen Beitrag leistet „Intersektionalität“ zur Klärung von Kompetenzunterschieden bei Jugendlichen? *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (Sonderband „Soziologische Bildungsforschung“, hrsg. von H. Solga & R. Becker), 86–109

Gross, C., & Gottburgsen, A. (2013). Gender, soziale Herkunft und Migration: „Intersektionalität“ im Kompetenzerwerb. In Hadjar, A., & Hupka-Brunner, S. (Hrsg.).

Geschlecht, Migration und Bildungserfolg (S. 188–212). Weinheim/München: Juventa Verlag

Gregori, P., Martínez, V., & Moyano-Fernández, J.J. (2018). Basic actions to reduce dropout rates in distance learning. *Evaluation and program planning*, 66, 48–52

Grosch, M., & Gidion, G. (2011). Mediennutzungsgewohnheiten im Wandel. Ergebnisse einer Befragung zur studiumsbezogenen Mediennutzung. Karlsruhe: KIT Scientific Publishing. Abgerufen am 30.08.2018 von <http://www.ksp.kit.edu/9783866446656>

Gross, C., Gottburgsen, A., & Phoenix, A. (2016). Education systems and intersectionality. In Hadjar, A., & Gross, C. (Hrsg.). *Educational systems and inequalities. International Comparisons* (S. 51–72). Bristol, UK: Policy Press

Handke, J. (2015). *Handbuch Hochschullehre Digital. Leitfaden für eine moderne und mediengerechte Lehre*. Marburg: Tectum Verlag

Heath, A., & Brinbaum, Y. (2016). Explaining ethnic inequalities in educational attainment. *Ethnicities*, 7(3), 291–304. <https://doi.org/10.1177/1468796807080230>

Henderson, M., Selwyn, N., & Aston, Rachel (2015). What works and why? Student perceptions of 'useful' digital technology in university teaching and learning. *Studies in Higher Education*, 42(8), 1567–1579

Heublein, U., Ebert, J., Hutsch, C., Isleib, S., König, R., Richter, J., & Woisch, A. (2017). Zwischen Studierenerwartungen und Studienwirklichkeit. Ursachen des Studienabbruchs, beruflicher Verbleib der Studienabbrecherinnen und Studienabbrecher und Entwicklung der Studienabbruchquote an deutschen Hochschulen (Forum Hochschule No. 1/2017). Hannover

Hochschulforum Digitalisierung (2016). *The Digital Turn – Hochschulbildung im digitalen Zeitalter*. Arbeitspapier Nr. 27. Berlin: Hochschulforum Digitalisierung. Abgerufen am 30.08.2018 von <https://hochschulforumdigitalisierung.de/sites/default/files/dateien/Abschlussbericht.pdf>

HRK (2013). *Europäische Studienreform. Empfehlung der 15. Mitgliederversammlung der Hochschulrektorenkonferenz* (Karlsruhe, 19.11.2013). Abgerufen am 27.08.2018 von [https://www.hrk.de/fileadmin/\\_migrated/content\\_uploads/Empfehlung\\_Europaeische\\_Studienreform\\_19112013.pdf](https://www.hrk.de/fileadmin/_migrated/content_uploads/Empfehlung_Europaeische_Studienreform_19112013.pdf)

Karapanos, M., & Fendler, J. (2015). Lernbezogenes Mediennutzungsverhalten von Studierenden der Ingenieurwissenschaften. Eine geschlechterkomparative Studie. *Journal of Technical Education*, 3(1), 39–55

KMK/HRK (2015-16). *Europäische Studienreform. Gemeinsame Erklärung von Kultusministerkonferenz und Hochschulrektorenkonferenz*. Beschluss der Hochschulrektorenkonferenz vom 10.11.2015 sowie der Kultusministerkonferenz vom 08.07.2016. Abgerufen am 27.08.2018 von [https://www.hrk.de/uploads/tx\\_szconvention/EUStudienreform\\_GemErklaerung\\_KMK\\_HRK\\_2015\\_2016.pdf](https://www.hrk.de/uploads/tx_szconvention/EUStudienreform_GemErklaerung_KMK_HRK_2015_2016.pdf)

Kristen, C. (2014). Migrationspezifische Ungleichheiten im deutschen Hochschulsystem. *Journal für Bildungsforschung*, 6(2), 113–134

Lack, K. A. (2015). Current Status of Research on Online Learning in Postsecondary Education. New York: Ithaka S+R. Retrieved from <http://www.sr.ithaka.org/wp-content/uploads/2015/08/ithaka-sr-online-learning-postsecondary-education-may2012.pdf>

Lai, K.-W., & Hong, K.-S. (2015). Technology use and learning characteristics of students in higher education. Do generational differences exist? *British Journal of Educational Technology* 46(4), 725–738

Leicht-Scholten, C. (2012). Diversity Management an deutschen Hochschulen – eine Annäherung. In Hochschulrektorenkonferenz-Nexus (Hrsg.), *Chancen erkennen – Vielfalt gestalten. Konzepte und gute Praxis für Studium und Lehre* (S. 8–12). Bonn: HRK

López-Pérez, M. V., Pérez-López, M. C., & Rodríguez-Ariza, L. (2011). Blended learning in higher education: Students' perceptions and their relation to outcomes. *Computers & Education*, 56(3), 818–826

Lörz, M., Netz, N., & Quast, H. (2016). Why do students from underprivileged families less often intend to study abroad? *Higher Education*, 72(2), 153–174

Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The Effectiveness of Online and Blended Learning: A Meta-Analysis of the Empirical Literature. *Teachers College Record*, 115(3), 1–47

Mayrberger, K. (2016). Lehren mit digitalen Medien – divers und lernendenorientiert. *synergie*, 01, 11–17

Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S., & Poskowsky, J. (2017). Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2016. 21. Sozialerhebung des Deutschen Studentenwerks – durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung. Berlin: Bundesministerium für Bildung und Forschung (BMBF)

Museus, S.D., & Griffin, K.A. (2011). Mapping the margins in higher education. On the promise of intersectionality frameworks in research and discourse. In K.A. Griffin & S.D. Museus (Hrsg.), *Using mixed-methods approaches to study intersectionality in higher education* (S. 5–13). San Francisco: Jossey Bass

Netz, N. (2015). What deters students from studying abroad? Evidence from four European countries and its implications for higher education policy. *Higher Education Policy*, 28(2), 151–174

Persike, M., & Friedrich, J.-D. (2016). Lernen mit digitalen Medien aus Studierendenperspektive. Arbeitspapier. Nr. 17. Berlin: Hochschulforum Digitalisierung. Abgerufen am 31.08.2018 von [https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD\\_AP\\_Nr\\_17\\_Lernen\\_mit\\_digitalen\\_Medien\\_aus\\_Studierendenperspektive.pdf](https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_Nr_17_Lernen_mit_digitalen_Medien_aus_Studierendenperspektive.pdf)

Porter, W. W., Graham, C. R., Spring, K. A., & Welch, K. R. (2014). Blended learning in higher education: Institutional adoption and implementation. *Computers & Education*, 75, 185–195

Slater, N., Peasgood, A., & Mullen, J. (2016). *Learning Analytics in Higher Education: A review of UK and international*. Bristol: jisc

Seale, J., Georgeson, J., Mamas, C., & Swain, J. (2015). Not the right kind of 'digital capital'? An examination of the complex relationship between disabled students, their technologies and higher education institutions. *Computers & Education*, 82, 118–128

Simpson, O. (2013). Student retention in distance education. Are we failing our students? *Open Learning: The Journal of Open, Distance and e-Learning*, 28(2), 105–119

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, 143(6), 565–600

Schaeper, H., & Weiß, T. (2016). The conceptualization, development, and validation of an instrument for measuring the formal learning environment in higher education. In H.-P. Blossfeld, J. von Maurice, M. Bayer, & J. Skopek (Hrsg.), *Methodological issues of longitudinal surveys: The example of the National Educational Panel Study* (S. 269–292). Wiesbaden: Springer VS

Schmid, U., Goertz, L., Radomski, S., Thom, S., & Behrens, J. (2017). Monitor Digitale Bildung. Die Hochschulen im digitalen Zeitalter. CHE – Centrum für Hochschulentwicklung; Bertelsmann Stiftung. Gütersloh

Stoessel, K., Ihme, T. A., Barbarino, M.-L., Fisseler, B., & Stürmer, S. (2015). Socio-demographic Diversity and Distance Education: Who Drops Out from Academic Programs and Why? *Research in Higher Education*, 56(3), 228–246

Statistisches Bundesamt (2018). Studierende an Hochschulen. Fachserie 11, Reihe 4.1, WS 2017/2018, Vorbericht

Steffens, Y., Schmitt, I. L., & Aßmann, S. (2017). Mediennutzung Studierender: über den Umgang mit Medien in hochschulischen Kontexten – Systematisches Review nationaler und internationaler Studien zur Mediennutzung Studierender. Bochum. Abgerufen am 28.08.2018 von <https://doi.org/10.13154/rub.106.95>

van Rooij, S. W., & Zirkle, K. (2016). Balancing pedagogy, student readiness and accessibility: A case study in collaborative online course development. *The Internet and Higher Education*, 28, 1–7

Wannemacher, K. (2016). Digitale Lernszenarien im Hochschulbereich. Arbeitspapier Nr. 15. Berlin: Hochschulforum Digitalisierung. Abgerufen am 31.08.2018 von [https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD%20AP%20Nr%2015\\_Digitale%20Lernszenarien.pdf](https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD%20AP%20Nr%2015_Digitale%20Lernszenarien.pdf)

Willige, J. (2016). Auslandsmobilität und digitale Medien. Arbeitspapier. Nr. 23. Berlin: Hochschulforum Digitalisierung. Abgerufen am 31.08.2018 von [https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD\\_AP\\_Nr23\\_Digitale\\_Medien\\_und\\_Mobilitaet.pdf](https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_Nr23_Digitale_Medien_und_Mobilitaet.pdf)

Willige, Sudheimer, Grützmacher (2017). Internationale Mobilität im Studium 2017. Ergebnisse der sechsten Befragung deutscher Studierender zur studienbezogenen Auslandsmobilität. Hannover: DAAD/DZHW (unveröffentlicht)

Wissenschaftsrat (2018). Empfehlungen zur Internationalisierung von Hochschulen. Abgerufen am 27.08.2018 von <https://www.wissenschaftsrat.de/download/archiv/7118-18.pdf>

Woisch, A., & Willige, J. (2015). Internationale Mobilität im Studium 2015. Ergebnisse der fünften Befragung deutscher Studierender zur studienbezogenen Auslandsmobilität. Hannover: DAAD/DZHW

Zawacki-Richter, O. (2015). Zur Mediennutzung im Studium – unter besonderer Berücksichtigung heterogener Studierender. Zeitschrift für Erziehungswissenschaft, 18(3), 527–549

Zawacki-Richter, O., Kramer, C., & Müskens, W. (2016). Studiumbezogene Mediennutzung im Wandel. Querschnittdaten 2012 und 2015 im Vergleich. Schriftenreihe zum Bildungs- und Wissensmanagement, 1,1–42

## Anhang

Modelle	Digitale internationale Mobilität			Tatsächliche internationale Mobilität		
	OR	Average Marginal Effects (AME)	Predictive Margins (PM)	OR	Average Marginal Effects (AME)	Predictive Margins (PM)
<b>Merkmale auf individueller Ebene und der Lernumwelt</b>						
<b>Diversität auf individueller Ebene</b>						
Geschlecht (Referenz: weiblich)	1,45*	0,08	0,36	0,69*	-0,08	0,27
Bildungsland (Referenz: deutsche Studierende)	1,10	0,02	0,33	1,17	0,03	0,35
Bildungsherkunft (Referenz: akad. Elternhaus)	0,74*	-0,07	0,28	0,64*	-0,09	0,27
Altersgruppen (Referenz: bis 25 Jahre)						
– 26 bis 30 Jahre	1,10	0,02	0,32	1,27*	0,05	0,35
– 31 Jahre und älter	1,17	0,03	0,33	0,91	-0,02	0,28
ohne Partnerschaft (Referenz: mit Partnerschaft)	1,07	0,01	0,32	1,09	0,02	0,33
Elternschaft (Referenz: ohne Kinder)	0,73*	-0,06	0,25	0,54*	-0,12	0,21
<b>Lernumwelt</b>						
Hochschulart Fachhochschulen (Referenz: Universitäten)	0,48*	-0,15	0,20	0,67*	-0,08	0,26
Fächergruppen (Referenz: Wirtschaftswissenschaften)						
– Sprach-, Kulturwiss.	0,83	-0,04	0,28	0,56*	-0,14	0,34
– Rechtswiss.	0,60*	-0,10	0,22	0,63*	-0,11	0,36
– Sozialwiss., Sozialwesen, Psychologie, Pädagogik	0,73	-0,06	0,26	0,48*	-0,17	0,30
– Mathematik, Naturwiss.	1,50*	0,09	0,42	0,41*	-0,20	0,27
– Medizin, Gesundheitswiss.	0,69	-0,08	0,25	0,46*	-0,18	0,29
– Ingenieurwiss.	1,06	0,01	0,33	0,44*	-0,19	0,29

**Anhang, Fortsetzung**

Modelle	Digitale internationale Mobilität			Tatsächliche internationale Mobilität		
	OR	Average Marginal Effects (AME)	Predictive Margins (PM)	OR	Average Marginal Effects (AME)	Predictive Margins (PM)
<b>Merkmale auf individueller Ebene und der Lernumwelt</b>						
Digitales Lern- und Lehrumfeld im Studiengang (Referenz: Grundständiges digitales Lern- und Lehrumfeld im Studiengang)						
– Spezifische digitale Lern- und Lehrformate im Studiengang	1,15	0,03	0,30	1,04	0,01	0,32
– Eigenständige digitale Lehrveranstaltungen und Studiengänge	1,44*	0,08	0,35	1,12	0,02	0,33

\* 95%-Konfidenzintervall, d.h. die Irrtumswahrscheinlichkeit für ein signifikantes Ergebnis ist kleiner als 5%

Manuskript eingereicht: 07.09. 2018  
Manuskript angenommen: 15.10. 2018

**Anschrift der Autorinnen:**

Dr. Anja Gottburgsen  
Janka Willige  
Deutsches Zentrum für Hochschul- und  
Wissenschaftsforschung GmbH (DZHW)  
Lange Laube 12  
30159 Hannover  
E-Mail: gottburgsen@dzhw.eu  
E-Mail: willige@dzhw.eu

Beide Autorinnen sind am Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) in Hannover tätig.

Dr. Anja Gottburgsen ist Referentin für Forschung und Change Management. Zu ihren Forschungsschwerpunkten gehören Soziale Ungleichheit, intersektionale Diversität und deren Effekte auf Teilhabe in Schule und Hochschule.

Janka Willige ist Projektleiterin des HISBUS-Studierendenpanels (Online-Access-Panel) und des Studienqualitätsmonitors (SQM). Ihre Forschungsschwerpunkte liegen in den Bereichen Digitalisierung, Auslandsmobilität und Studienqualität.

# Die Implementation und Evaluation eines Flipped Classrooms in einer Großveranstaltung der Statistik

Manuel Förster, Florian Heiß, Sigbert Klinke, Andreas Maur, Thorsten Schank, Constantin Weiser

---

Unter den üblichen Arrangements statistischer Großveranstaltungen verschärfen sich aufgrund eingeschränkter Maßnahmen der Lernprozessindividualisierung nicht selten die heterogen ausgeprägten kognitiven und motivationalen Eingangskonstellationen, die insbesondere im Fach Statistik unter anderem durch das Geschlecht, das Vorwissen und andere soziodemographische Faktoren bedingt sind. In dem Artikel wird das theoretisch fundierte Design eines Flipped Classrooms als ein Lösungsansatz zum Umgang mit dieser Herausforderung in einer statistischen Großveranstaltung entwickelt. Zudem wird ein an empirischen und methodischen Forschungslücken anknüpfendes Evaluierungskonzept vorgestellt, welches die kognitive und motivationale Entwicklung sowie den Workload der Studierenden im Längsschnitt in der klassischen und geflippten Veranstaltung berücksichtigt.

---

## 1 **Notwendigkeit und Potenziale innovativer Lehr- und Lernkonzepte für die universitäre Statistikausbildung**

Die zunehmende Wichtigkeit des versierten Umgangs mit stetig wachsenden quantitativen Datenmengen zeigt sich sowohl in alltäglichen als auch insbesondere in beruflichen Lebenssituationen. Dementsprechend werden in einem Großteil der Stellenausschreibungen aus wirtschafts- und sozialwissenschaftlichen Anforderungsbereichen explizit statistische Kompetenzen vorausgesetzt (Kauermann, 2015; Lippe & Kladroba 2008), da diese eine Grundvoraussetzung zur Überwachung der für den Organisationserfolg relevanten Qualitätsmanagementprozesse darstellen. Die Relevanz des empirisch-quantitativen Forschungsparadigmas schlägt sich auch in den Curricula der wirtschaftswissenschaftlichen Bachelorstudiengänge nieder, in denen die statistische und mathematische Methodenausbildung deutschlandweit mit bis zu 20 Prozent der Kreditpunkte einen wesentlichen Bestandteil bildet und an nahezu allen Universitäten und Fachhochschulen gelehrt wird (Zlatkin-Troitschanskaia, Förster, Brückner, Hansen & Happ, 2013). Insbesondere in dem hier betrachteten Kontext der Wirtschaftswissenschaften, in welchem Statistik als Großveranstaltung mit mehreren Hundert Studierenden in den Vorlesungen gelehrt wird, gerät die Förderung eines ganzheitlichen Kompetenzerwerbs mittels individualisierter Lehr-Lernangebote und Gruppenarbeiten bei den in aller Regel dozentenzentriert gehaltenen Vorlesungen

schnell an ihre Grenzen. Solche individualisierten Lehr- und Lernarrangements sind allerdings notwendig, um der steigenden Heterogenität der Studierenden (Happ & Zlatkin-Troitschanskaia, 2015) adäquat begegnen zu können.

Neben der bereits mehrfach nachgewiesenen Bedeutung des Geschlechts bei statistischen Lernprozessen (Förster & Maur, 2015; Tempelaar & Schim van der Loeff, 2011) hat die gewachsene Beteiligung an der tertiären Bildung zu einer breiten Streuung der für den statistischen Wissenserwerb hoch bedeutsamen mathematisch-statistischen Vorkenntnisse, aber auch der soziodemographischen Rahmenbedingungen (v. a. familiärer Hintergrund, Bildungsaspiration, Subsistenz) geführt (Ramirez, Schau & Emmioglou 2012). Dazu hat unter anderem der steigende Anteil an Studierenden mit Migrationshintergrund beigetragen. Unter den üblichen Arrangements statistischer Großveranstaltungen mit eingeschränkten Möglichkeiten zur Berücksichtigung dieser individuellen Anforderungen verschärfen sich nicht selten im Laufe einer klassischen Methodenveranstaltung die Differenzen in den kognitiven und motivationalen Eingangskonstellationen (Förster & Maur, 2016). Dazu existiert bei den Studierenden der Wirtschafts- und Sozialwissenschaften oftmals die trügerische Hoffnung, mathematische und statistische Inhalte mit ihrer Studienwahl umgehen zu können, sodass die Statistikveranstaltungen letztlich zumeist nur mit Abneigung und geringer Motivation absolviert werden (Macher, Paechter, Papousek & Ruggeri, 2012). Außerdem weisen Studierende im Fach Statistik zahlreiche Fehlkonzeptionen auf, welche verhindern, dass statistische Konzepte adäquat kognitiv verknüpft oder auf neue Kontexte angewendet werden können (Garfield & Ben-Zvi 2007). Durch diese passive Form der Wissensvermittlung und aus der geringen Motivation heraus werden die Studierenden dazu verleitet, erst kurz vor den Klausuren fragmentarisches Faktenwissen ohne Kontext- und Anwendungsbezug aufzunehmen und kurze Zeit später wieder zu vergessen (Römmer-Nossek, Peschl & Zimmermann, 2013).

Ein Lösungsansatz zum Umgang mit dieser Heterogenität in einer statistischen Großveranstaltung kann in der Implementation eines Flipped Classroom-Designs<sup>1</sup> gesehen werden, mit welchem den Studierenden durch wiederverwendbare Lernvideos zunächst ein individualisiertes Lernangebot ermöglicht werden kann. Das Hauptmerkmal des Flipped Classroom besteht darin, dass Wissensvermittlung und Anwendung bzw. Übung insofern vertauscht werden, dass sich die Studierenden das neue Wissen selbstständig mit Hilfe von bereitgestellten Lernmaterialien, wie z. B. Lernvideos, interaktiven Visualisierungen, ausgewählten Lehrtexten und Aufgaben außerhalb der Lehrveranstaltung aneignen. In der Präsenzzeit liegt der Fokus auf der aktiven Bearbeitung komplexer Aufgabenstellungen und deren Diskussion mit

<sup>1</sup>Anstelle von „Flipped Classroom“ wird auch oftmals der Begriff des „Inverted Classroom“ verwendet. In der internationalen Literatur werden diese beiden Ansätze in der Regel synonym verwendet, so dass auch in diesem Artikel keine Differenzierung stattfindet.

individueller Betreuung der Studierenden (Fischer & Spannagel, 2012). Das Potential des Flipped Classroom liegt darin, dass verschiedenartige Lernmaterialien in der Selbstlernzeit entsprechend des individuellen Lerntyps ausgewählt, beliebig oft wiederholt und die Lernzeit und -geschwindigkeit an die eigenen Lernbedürfnisse angepasst werden können. Dadurch wird erwartet, dass sich die Studierenden früher aktiv mit den Inhalten auseinandersetzen, eigene Fehlkonzeptionen schneller erkennen und diese durch Hilfestellungen von Kommilitonen oder den Lernhelfern direkt in der Veranstaltung aufgelöst werden können.

Bislang gibt es kaum Forschungsbefunde zu den Wirkungsweisen eines Flipped Classroom im Vergleich zu klassischen Großveranstaltungen, wobei empirisch abgesicherte Erkenntnisse über hochschuldidaktische Zugewinne mittels eines Flipped Classroom in Anbetracht des aus der Implementation resultierenden Ressourcenaufwands zur weiteren Verbreitung dieses Ansatzes beitragen können. In diesem skizzierten Umfeld wird in dem seit Beginn des Jahres 2017 laufenden Projekt „Förderung statistischer Lehr- und Lernprozesse mittels eines Flipped Classroom-Designs“ [FLIPPS] untersucht, ob und inwieweit das Arrangement eines Flipped Classroom im Vergleich zu traditionellen Veranstaltungsformen in der Lage ist, die Lernprozesse der heterogenen Studierendenschaft so zu fördern, dass in der Statistik insgesamt positivere und stabilere Wissensstrukturen erworben werden. Weiterhin wird untersucht, ob durch das Flipped Classroom-Design auch das Selbstkonzept und die Motivation im Fach Statistik gefördert werden können.

In diesem Artikel wird auf Grundlage dieser Vorüberlegungen ein Assessment Framework entwickelt, welches zum Ziel hat, den potentiellen didaktischen Mehrwert eines Flipped Classroom empirisch zu prüfen. Dazu werden aus bestehenden Forschungsdesiderata sowohl theoretische und methodische Analyseansätze als auch forschungsleitende Fragestellungen abgeleitet. Nachfolgend werden das zur Analyse der Forschungsfragen verwendete Studiendesign, die Messinstrumente, Erhebungsmethoden und in Kürze die geplanten empirischen Analyseverfahren vorgestellt. Schließlich werden mit dem Assessment-Framework verbundene messmethodische Herausforderungen und die Generalisierbarkeit der Studienergebnisse kritisch diskutiert.

## **2 Inhaltlicher und messmethodischer Forschungsstand zu Lernprozessen in Flipped Classroom-Kontexten**

In der Forschung zum Erwerb statistischer Kenntnisse kristallisieren sich neben den grundlegenden kognitiven Fähigkeiten und der Wissensanwendung auch motivationale Größen und das statistikbezogene Selbstkonzept als zentrale Einflussgrößen auf den Kompetenzerwerb heraus (Ramirez, Schau & Emmioglu, 2012), welche auch in dieser

Studie genauer untersucht werden. Zunächst liefern Forschungsergebnisse im tertiären Bildungssektor und hier insbesondere im Bereich Statistik bisher nur ein widersprüchliches Bild darüber, wie das Flipped Classroom-Konzept auf das *Wissen* bzw. den *Lernerfolg* wirkt (Bishop & Verleger, 2013; Abeysekera & Dawson, 2015). Zudem wurde der kognitive Lernerfolg und dessen Nachhaltigkeit oftmals nur mit Selbstauskünften und nicht mittels objektiv gewonnener Leistungsdaten erfasst (Bishop & Verleger, 2013) sowie der tatsächlich erbrachte Kompetenzerwerb eher undifferenziert und ohne lernpsychologische Grundlage (z. B. Blooms Taxonomie; O’Flaherty & Phillips, 2015) erhoben. Mit Blick auf die *Motivation und das Selbstkonzept* im Vergleich zu klassischen Designs verweisen einige Studien auf Zugewinne (Moore & Chung, 2015; Fischer & Spannagel, 2012), während andere Studien Motivationsverluste berichten, die durch den höheren Verpflichtungscharakter und die offeneren Strukturen des Flipped Classroom ausgelöst werden (Abeysekera & Dawson, 2015). Ein zusätzlicher Mangel liegt darin, dass nur wenige Längsschnittdaten über die Veränderung von in ein Gesamtmodell überführten kognitiven und affektiven Merkmale vorliegen (O’Flaherty & Phillips, 2015). Dies schränkt die Möglichkeit ein, empirisch belastbare Aussagen über die Kausalität von Veränderungen in *Wissen*, *Motivation* und *Selbstkonzept* der Studierenden und somit zur Wirkungsweise eines solchen Lehrkonzepts zu treffen. In den bisherigen Designs bleiben auch die wöchentliche Lernaktivität bzw. der *Workload* hinsichtlich angebotener Lerngelegenheiten weitgehend unberücksichtigt (Bishop & Verleger, 2013). So wird bislang z. B. kaum untersucht, ob das Flipped Classroom-Design zu einem höheren *Workload* im Vergleich zu einem anderen Lehr- und Lerndesign führt, was gegebenenfalls auch einen möglichen höheren Kompetenzgewinn erklären könnte. Auch Studien aus anderen Hochschulkontexten beschäftigen sich kaum mit dem tatsächlichen *Workload* über ein gesamtes Semester, sondern eher mit Fragen nach dessen Operationalisierbarkeit (Rožman, Lešer, Širca, Dermol & Skrbinjek, 2014). Bislang existieren zudem nur wenige und uneinheitliche Befunde dazu, inwieweit das Flipped Classroom-Konzept eine förderliche kognitive und affektive Differenzierung des Lernens bei den oben angesprochenen heterogenen Gruppen (bzgl. Vorwissen, Migrationshintergrund, Geschlecht, etc.) ermöglicht, so dass sich diese Gruppen im Hinblick auf *Lernerfolg*, *Motivation* und *Selbstkonzept* bestmöglich entwickeln (Bishop & Verleger, 2013).

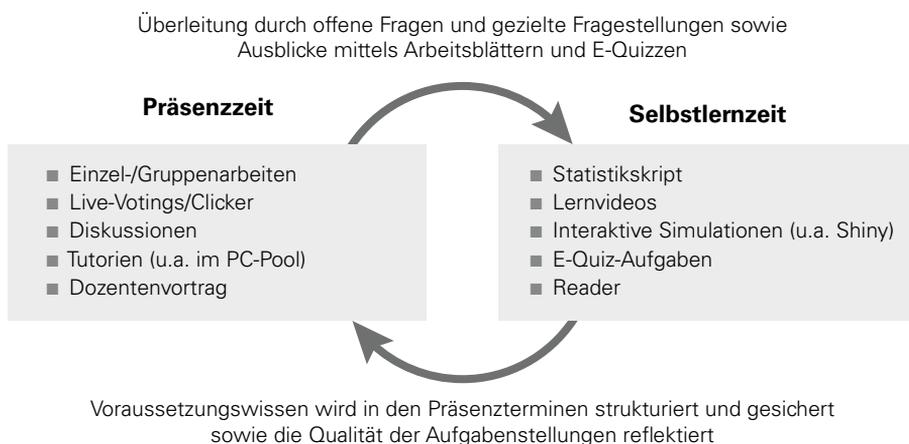
Anknüpfend an diese Forschungslücken wird mit dem Assessment-Design das Ziel angestrebt, die Effekte eines Flipped Classroom sowohl auf den kognitiven, objektiv erfassten Lernerfolg als auch auf kognitiv-affektive Lernprozessvariablen wie Motivation, Selbstkonzept und Wahrnehmungen des Lernprozesses in einem lerntheoretisch validierten Gesamtmodell im Längsschnitt zu erfassen und ihren Zusammenhang sowie die Wirksamkeit hinsichtlich der Förderung heterogener Studierendengruppen im Vergleich zu traditionellen Vorlesungen zu untersuchen. Die konkrete Ausgestaltung eines Assessment-Designs, welches die genannten Desiderate aufgreift, wird in den folgenden Kapiteln erörtert.

### 3 Ausgestaltung des Flipped Classroom-Konzepts

Die neu konzipierte Flipped Classroom-Veranstaltung wird an zwei Universitätsstandorten (Düsseldorf und Mainz) mit jeweils ca. 600 Studierenden angeboten. An beiden Standorten wurde ein Jahr vor der Einführung des Flipped Classroom eine konventionelle, dozentenorientierte Vorlesung mit begleitenden Übungen angeboten. Die traditionelle Veranstaltung setzt sich aus wöchentlichen Vorlesungs- und Übungsterminen (insgesamt ca. 600 Teilnehmer) zusammen. In den Vorlesungen orientiert sich der Lehrende an dem Vorlesungsskript, während in den Tutorien Aufgabenblätter bearbeitet werden. Das Lernmaterial der traditionellen Veranstaltung besteht aus dem Vorlesungsskript und den Übungsaufgaben. Die Art der Interaktion beschränkt sich jedoch meistens darauf, dass die Studierenden vom Lehrpersonal dazu angeregt werden, Rückfragen zu stellen. Dabei konnte festgestellt werden, dass ein Großteil der Studierenden die Aufgabenblätter im Vorfeld nicht bearbeitet und in den Tutorien nur die „Musterlösung“ notiert. Die klassische Veranstaltung umfasst 5 Semesterwochenstunden, während die Flipped Classroom-Veranstaltung aufgrund der Auslagerung der Wissensvermittlung auf die Lernvideos auf 3,5 Semesterwochenstunden reduziert wird.

Abbildung 1 zeigt nunmehr die im Flipped Classroom-Design implementierten, operativen Elemente, auf Grund derer eine Variation in den Lerneffekten im Vergleich zur traditionellen Veranstaltung erwartet werden kann:

**Abbildung 1:** Verflechtung der Veranstaltungsbestandteile



Zunächst sollen sich die Studierenden mittels konzeptioneller Lernvideos die statistischen Inhalte selbst erschließen, indem diese zunächst eine lebensweltnahe Ausgangsproblematik schaffen, das dazu notwendige statistische Verfahren dynamisch

erläutern und eine Verbindung zur Präsenzveranstaltung herstellen. In dieser werden die Themen vertiefend mit einem formelleren Zugang in Kleingruppen angewendet. Von den Lernvideos wird direkt auf entsprechende interaktive Simulationstools zur Auseinandersetzung mit den statistischen Inhalten verlinkt (MM\*Stat und GrASP), mit denen sich die Studierenden durch Experimentieren und eigene Manipulationen die Funktionsweise statistischer Konzepte erschließen können. Angelehnt an das Inverted Classroom Mastery Model (Handke, 2015) werden regelmäßig verpflichtende E-Quiz-Aufgaben bereitgestellt, die den Studierenden eine Einschätzung ermöglichen, ob sie die jeweiligen Inhalte verstanden haben und ob sie für die Präsenzteilnahme bereit sind. Gewährleistet wird dies dadurch, dass die Studierenden unmittelbar nach dem Ausfüllen des E-Quizzes eine automatisierte Rückmeldung zu ihren Aufgabenlösungen erhalten. Die aus wöchentlichen Groß- und Kleingruppenübungen inklusive praktischen Softwaretutorials bestehenden Präsenztermine haben eine moderierend-strukturierende Funktion und werden durch Lernhelfer bzw. Tutoren begleitet, sodass in diesen die Inhalte aus Selbstlernzeit als Voraussetzungswissen kontrolliert und gesichert werden und auf nachfolgende Inhalte referenziert wird. Zudem werden Audio-Response-Systeme und Smartphones in die Veranstaltungen integriert, mit denen die Studierenden Wissensfragen live beantworten können, so dass auch der Dozent eine Rückmeldung bekommt, in welchem Ausmaß die Aufgaben richtig beantwortet werden konnten.

#### **4 Forschungsleitende Annahmen**

Ausgehend von dem Forschungsziel, den aus den Forschungsdesiderata abgeleiteten Zielkonstrukten und den konkreten operativen Designelementen fokussiert sich die Studie auf die folgenden vier Arbeitshypothesen:

1. Die Studierenden weisen im Flipped Classroom eine bessere und nachhaltigere kognitive Zielerreichung auf als Studierende im klassischen Design.

Der Selbstregulationstheorie von Bandura (2015) folgend, gehen die systematisierte Vorbereitung für die Präsenzveranstaltungen sowie die formativen Rückmeldungen zu den individuellen Lernleistungen durch die E-Quizze und die intensivere Interaktion mit Lehrenden und Peers mit einer realistischeren Selbsteinschätzung einher. Solches Feedback wirkt sich gemäß der Metaanalyse von Hattie und Timperley (2007) positiv auf den kognitiven Wissenserwerb aus, da Lernende so ihr Lernverhalten in Hinblick auf die erwarteten Lernziele regulieren können. Zudem werden durch die (erwartete) frühere und intensivere Auseinandersetzung mit dem Lernstoff dauerhaftere Wissensstrukturen gebildet.

2. Studierende im Flipped Classroom weisen einen über das Semester hinweg höheren, aber gleichmäßiger über das Semester verteilten Workload auf, da sich mehr Studierende bereits während der frühen Phasen der Veranstaltungen aktiv mit den Fachinhalten auseinandersetzen.

Ziel im Rahmen dieser Hypothese ist es, das sogenannte „Bulimielernen“ zu verringern und Studierende dazu zu bewegen, sich frühzeitig mit den statistischen Inhalten und Aufgaben auseinanderzusetzen. Vereinzelt Befunde deuten darauf hin, dass Studierende den Workload in Flipped Classroom-Arrangements als zu hoch erachten, wofür jedoch meist keine längsschnittlichen Stundenangaben, sondern Ratings verfasst wurden (z. B. Lo, Hew & Cheng, 2017). Dagegen sollte die Zeit, in der sich die Studierenden aktiv in den Präsenzveranstaltungen mit den Inhalten auseinandersetzen („time on task“) steigen und sich gleichmäßiger über das Semester verteilen. Dadurch, dass in den Präsenzveranstaltungen auf Lernaktivitäten in der Selbstlernzeit (u. a. Rezeption der Lernvideos, Bearbeitung von Aufgabenblättern und Quizen) aufgebaut wird und diese somit vorausgesetzt werden, ist anzunehmen, dass sich Studierende im Flipped Classroom-Design frühzeitiger und in der Präsenzveranstaltung intensiver mit den Lerninhalten auseinandersetzen.

3. Studierende im Flipped Classroom weisen eine positivere Motivations- und Selbstkonzeptentwicklung auf als Studierende im klassischen Design.

Diese Hypothese fokussiert die nicht-kognitiven Lernziele und setzt an der Annahme an, dass Studierende früher Feedback durch die Kommilitonen, die Tutoren und die E-Quizze erhalten, was zu einer Korrektur des oftmals negativen Selbstkonzepts führen sollte. Der Selbstbestimmungstheorie von Ryan und Deci (2016) folgend verzeichnen Lerner insbesondere dann ein hohes Interesse, Engagement und eine hohe intrinsische Motivation, wenn die Lehrveranstaltung das Autonomie- und Kompetenzbedürfnis sowie die soziale Eingebundenheit der Studierenden befriedigt. Dies ergibt sich im Flipped Classroom daraus, dass Studierende die Lerninhalte örtlich und zeitlich flexibler wahrnehmen können und die Lernmaterialien hinsichtlich ihres Vorwissens und der individuellen Verarbeitungsgeschwindigkeit angepasst rezipieren können.

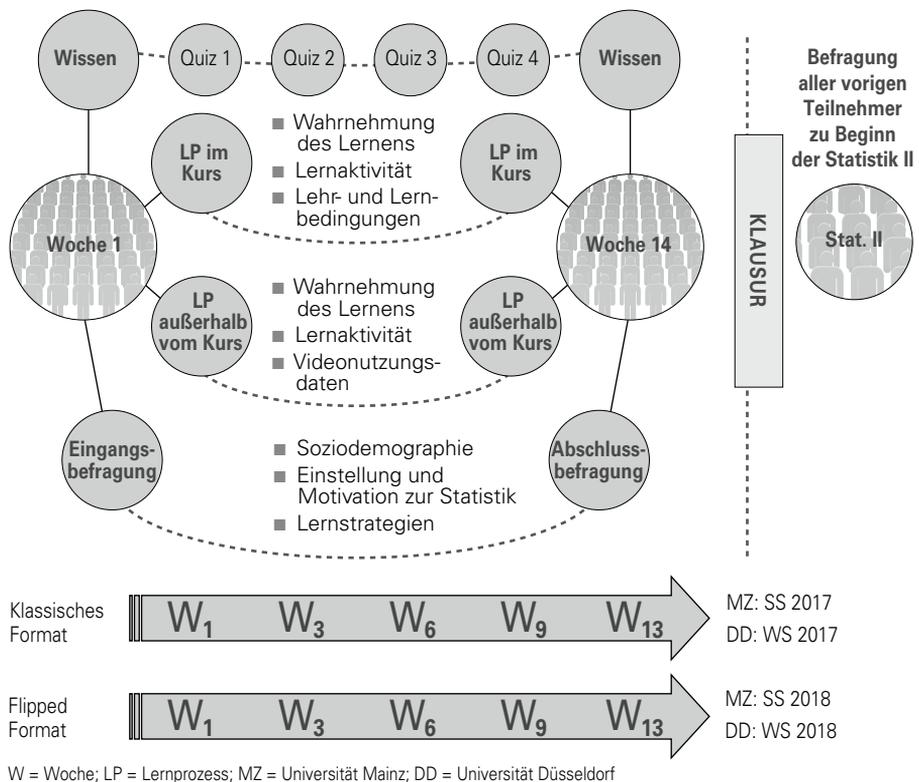
4. Es wird erwartet, dass das hohe Flexibilitätspotential des Flipped Classroom höhere Anpassungsleistungen an Adressaten mit unterschiedlichen Merkmalen (Geschlecht, Vorwissen, Migrationshintergrund) als das klassische Veranstaltungskonzept zu erbringen vermag, so dass alle diese Teilgruppen kognitiv und affektiv vom Flipped Classroom-Konzept profitieren.

Es wird angenommen, dass das Flipped Classroom-Konzept geeignet ist, alle Studierendengruppen in Hinblick auf die untersuchten kognitiven und affektiven Einflussgrößen besser zu fördern als die klassische Veranstaltung. Aufgrund der Methodenvielfalt und Multimedialität des Flipped Classroom bietet dieser heterogenen Studierendengruppen flexible Zugänge zu den Inhalten. Insbesondere in statistischen Veranstaltungen zeigt sich, dass weibliche Studierende und Studierende mit geringerem Vorwissen eine geringere Motivation in Statistik aufweisen und dadurch gegebenenfalls schlechter darin abschneiden (Ramirez et al., 2012).

## 5 Assessment-Framework

Um diese Hypothesen prüfen zu können, wird ein Vergleichsdesign an zwei Hochschulstandorten realisiert (siehe Abbildung 2):

**Abbildung 2:** Assessment Framework im Überblick



Zunächst wurde im Sommersemester 2017 an der Johannes Gutenberg-Universität in Mainz und im Wintersemester 2017/2018 an der Heinrich-Heine-Universität in Düsseldorf das konventionelle Lehr- und Lernkonzept evaluiert, bevor zum Sommersemester 2018 bzw. Wintersemester 2018/2019 die Veranstaltungen an beiden Standorten auf das Flipped- Classroom-Design umgestellt und evaluiert werden. Da in Überblicksstudien (vgl. Bishop & Verleger, 2013) der Mangel an qualitativ verschiedenen Datenquellen und deren Triangulation als Forschungsdesiderat benannt wurde (siehe Kapitel 2), werden im Projekt Ratingdaten zum Lernverhalten im Längsschnitt sowie objektive Nutzungs- und Performancedaten der E-Materialien, der E-Quiz und Klausuren im Quer- und Längsschnitt erfasst und verknüpft. Tabelle 1 gibt einen Überblick über die erfassten Zielkonstrukte und Kontrollvariablen.

**Tabelle 1:** Übersicht der erfassten Ziel- und Kontrollvariablen im Längsschnitt mit Beispielitems

Zielvariablen											
Vorlesungs- woche	1	3	6	7	8	9	10	12	13	KL	NB
Befragungs- form	PP	PP	OB/ Q1	PP	OB/ Q2	PP	OB/ Q3	Q4	PP	KL	PP
H1: Kognitive Lernziel- erreicherung			X		X		X	X		X	X
	Beispielitem aus einem E-Quiz: s. Abbildung 3										
		X		X		X			X	X	
H2: Workload	Angabe des Workload in Stunden u. a. für folgende Lerngelegenheiten: Vor- und Nachbereitung des Skriptes, Bearbeitung von Übungsaufgaben, Rezeption von Lernvideos, Nutzung von Literatur, Nutzung von Lehrbüchern										
	Im Flipped Classroom werden zudem Videonutzungsdaten (probandenbezogene Anzahl, Dauer und Datum der Aufrufe pro Lernvideo) erhoben.										
H3: Motiva- tion & Selbst- konzept (SATS)	X				X				X		X
	Konstrukte mit Beispielitems; Ratingskala von „trifft nicht zu“ – „trifft voll zu“: – <b>Affect:</b> Ich mag Statistik. ( $\alpha = 0,822-0,757$ ) – <b>Cognitive Competence:</b> Ich kann Statistik erlernen. ( $\alpha = 0,831-0,761$ ) – <b>Value:</b> Statistik ist für mein Studium wichtig. ( $\alpha = 0,828-0,813$ ) – <b>Interest:</b> Ich habe Interesse daran, Statistik anzuwenden. ( $\alpha = 0,853-0,825$ ) – <b>Difficulty:</b> Statistik ist ein kompliziertes Fach. ( $\alpha = 0,722-0,672$ ) – <b>Effort:</b> Ich investiere viel Arbeit in die Statistikveranstaltung. ( $\alpha = 0,833-0,712$ )										
H4: Hetero- genitäts- kriterien	X										
	U. a. Geschlecht, Note in vorausgegangenen Mathematikveranstaltungen, Abiturnote, Abiturnote in Mathematik, Muttersprache, Herkunft der Eltern										

Fortsetzung Tabelle 1 nächste Seite

**Tabelle 1, Fortsetzung**

Kontrollvariablen											
Vorlesungs- woche	1	3	6	7	8	9	10	12	13	KL	NB
Befragungs- form	PP	PP	OB/ Q1	PP	OB/ Q2	PP	OB/ Q3	Q4	PP	KL	PP
Lernen und Emotionen innerhalb und außerhalb der Veranstaltung (AEQ)		X	X	X		X	X		X		
Lernstrategien (MSLQ)	X								X		
Lernbedingun- gen (SEEQ)		X		X		X					

PP = Paper-Pencil; OB = Onlinebefragung; Q= E-Quiz; KL = Klausur; NB = Nachbefragung;  
 SATS = Survey of Attitudes Towards Statistics; AEQ = Achievement Emotion Questionnaire;  
 MSLQ = Motivated Strategies for Learning Questionnaire; SEEQ = Students' Evaluations of Educational Quality;  
 $\alpha$  = Cronbach's Alpha über die verschiedenen Befragungszeitpunkte

Die Erfassung der *kognitiven Lehrziele* (H1) wird mit Hilfe von vier Online-Befragungen im Rahmen der in der Veranstaltung angebotenen E-Quiz und damit im selben Modus wie die semesterabschließende E-Klausur (KL) vollzogen.<sup>2</sup> Die Quizfragen wurden bereits von sachverständigen Experten mittels eines Assessment-Frameworks unter integrativer Zugrundelegung sowohl der aktualisierten Bloom-Taxonomie als auch des SOLO-Modells unter Berücksichtigung einer repräsentativen inhaltlichen und kognitiven Abdeckung aller relevanten statistischen Inhaltbereiche konzipiert (Beispielaufgabe s. Abbildung 3). Während die Bloom-Stufen einen Indikator für die analytische Durchdringung und Verarbeitungstiefe der statistischen Inhalte darstellen, spiegeln die SOLO-Stufen die Verknüpfung verschiedener statistischer Konzepte wider (Garfield & Ben-Zvi, 2007).

<sup>2</sup>Lediglich die mit der Nachbefragung im nachfolgenden Semester gemessene Nachhaltigkeit des Wissenserwerbs wird aufgrund der erwarteten höheren Teilnehmersausbeute im Rahmen der Lehrveranstaltung als Paper-Pencil-Befragung durchgeführt.

**Abbildung 3:** Beispielaufgabe aus einem E-Quiz

**Regression Gebrauchtwagen (6 Punkte)**

Andreas interessiert sich für den Kauf eines Gebrauchtwagens vom Typ „Smart“. In der Samstagsausgabe der Regionalzeitung findet er acht passende Angebote:

<b>Tachostand</b> [1000 km]	14	9	30	37	17	6	20	20
<b>Preis</b> [1000 €]	11.7	12	9.6	9.3	10.8	12.3	10.2	10.5

- a) Andreas nimmt an, dass der Preis (Y) linear vom Tachostand (X) abhängt. Berechnen Sie die Regressionsgerade nach der Methode der kleinsten Quadrate.  
 (Benutzen Sie hierfür  $\bar{y} = 10.8$ ;  $\bar{x} = 20$ ;  $S_y^2 = 1.08$ ;  $S_x^2 = 100$ ;  $c_{xy} = -9.9$ )

$\hat{Y}_i =$  \_\_\_\_\_

- b) Berechnen Sie das Bestimmtheitsmaß:

$R^2 =$  \_\_\_\_\_

- c) Andreas erfährt von seinem Freund Simon, dass dieser ihm seinen gebrauchten „Smart“ mit dem Tachostand 40.000 km zum Freundschaftspreis von 9.000 € verkaufen würde. Berechnen Sie den prognostizierten Wert für einen Gebrauchtwagen mit 40.000 km.

$\hat{Y}_{(40)} =$  \_\_\_\_\_ [in 1000 €]

Zur Erfassung des Lernverhaltens werden Paper-Pencil-Befragungen während den Vorlesungen sowie örtlich flexibel bearbeitbare Onlinebefragungen eingesetzt (siehe Tabelle 1). Für Hypothese H2 soll neben dem Workload (Angabe retrospektiv für die letzten sieben Tage vor der jeweiligen Befragung) im Flipped Classroom-Design zudem die Nutzung der Medien pseudoanonymisiert erfasst werden, indem die Daten zwar Nutzern zugeordnet werden, diese aber direkt eine neue anonymisierte ID erhalten. Hierbei wird sowohl die Nutzung der Lernvideos und Zusatzmaterialien anhand der Abrufdaten, -dauer und -häufigkeiten, als auch der Erfolg bei der Bearbeitung der E-Quizze erfasst. Dies ermöglicht vor allem innerhalb des Flipped Classroom eine Analyse des mediengestützten Lernverhaltens. *Motivation und Selbstkonzept* (H3) werden mittels des „Survey of Attitudes Towards Statistics-36“ (SATS-36; Schau, 2003) erhoben, welches auf der Grundlage des Erwartungs-Mal-Wert-Motivationsmodells von Eccles und Wigfield (2001) konzipiert wurde. Da es sich um die erste Statistikveranstaltung der Studierenden handelt, wird im Rahmen der Heterogenitätskriterien (H4) das Vorwissen über die Note in der vorausgehenden Mathematik-Vorlesung, die Abiturnote in Mathematik und den Besuch eines Mathematikleistungs- oder Grundkurses erfasst. Zur Verknüpfung aller Individualdaten der verschiedenen Befragungen geben die Studierenden in allen Befragungen ihre Nutzerkennung für die Online-Plattform an. Schließlich werden zur Erfassung eines ganzheitlichen Lernprozesses verschiedene Kontrollvariablen, wie unter anderem die Lernstrategien (Memotechniken, Selbstregulation, Elaborationsstrategien, kritisches Denken,

Prokrastination, Peer Learning, Ressourcenmanagement) als auch die Angst vor Klausuren mittels des Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia & McKeachie, 1991) erhoben. Außerdem wurden die Studierenden im Laufe des Semesters zu ihrer inner- und außerinstitutionellen Wahrnehmung des Lernens befragt. Fokussiert wurden hierbei im Rahmen des Achievement Emotion Questionnaire (AEQ; Pekrun, Goetz, Titz & Perry, 2005) die wahrgenommene Langeweile, Frustration, Motivation als auch das Selbstkonzept.<sup>3</sup> Zusätzlich wurden zur Kontrolle veranstaltungsmethodischer Einflüsse Konstrukte zur wahrgenommenen Instruktionklarheit, Interaktion und Organisation der Veranstaltung mittels des Students' Evaluations of Educational Quality (SEEQ; Marsh, 2001) erfasst.<sup>4</sup> Die Intervalle für die Erfassung der verschiedenen Konstrukte wurden möglichst äquidistant gesetzt, während bei Konstrukten, die per Annahme im Zeitverlauf persistenter im Vergleich zu situationsabhängigen Emotionen sind, wie beispielsweise den Lernstrategien (MSLQ) oder den globaleren Einstellungen zur Statistik, breitere Intervalle gewählt wurden.

Um die Veränderungen in Motivation, Selbstkonzept und Workload zu modellieren und Aussagen zu veranstaltungsbezogenen und methodischen Ursachen dieser Veränderungen treffen zu können (siehe Kapitel 2), werden bei den kommenden Analysen Latent-Change-Modelle und Wachstumskurvenmodelle eingesetzt. Diese ermöglichen in einem ersten Schritt, die Datenverläufe des traditionellen und Flipped Classroom im Längsschnitt einzeln zu untersuchen (Within-Design-Effekte) und dann miteinander zu vergleichen (H1-H3). Weiterhin erfolgt mittels autoregressiver Strukturgleichungsmodelle eine Schätzung, in welchem Maße sich die verschiedenen latenten Variablen der Motivation, Selbstkonzept und kognitiven Lernzielerreichung im Laufe des Semesters gegenseitig bedingen und inwieweit die unterschiedlichen Kursdesigns dazu beigetragen haben. Diese Modellierungen erlauben auch gleichzeitig, die Entwicklungen verschiedener Subgruppen (Heterogenitätsgruppen, H4) zu vergleichen.

## 6 Diskussion des Projekts

Bei den untersuchten Treatments handelt es sich um Großveranstaltungen mit 600 bis 700 Studierenden, so dass das Projekt erstmalig belastbare und vorsichtig generalisierbare Befunde über die aus der Digitalisierung der Lehre resultierenden Änderungen des Lernverhaltens in Flipped Classroom-Großveranstaltungen im Ver-

<sup>3</sup>Die Wahrnehmung der außerinstitutionellen Lernprozesse wird dabei kontextadäquat mittels Onlinebefragungen außerhalb der Veranstaltung erfasst, während die innerinstitutionellen Lernprozesse während der Veranstaltung mittels Paper-Pencil erfasst werden.

<sup>4</sup>Diese wurden ebenfalls aus Gründen der Kontextunmittelbarkeit im Rahmen der Veranstaltung im Paper-Pencil-Format erfasst.

gleich zu traditionellen Veranstaltungen liefert – auch wenn Replikationsstudien in weiteren Fächern oder an weiteren Standorten erforderlich bleiben. Ein weiterer Diskussionspunkt liegt im quasi-experimentellen Design der Studie, welches aufgrund der fehlenden Randomisierung keine eindeutig kausale Zuordnung von beobachteten Effekten zu den zahlreichen variierenden Elementen in dem Flipped Classroom-Design zulässt (Means, Bakia, & Murphy, 2014). Ein Vorteil des quasi-experimentellen Designs liegt demgegenüber darin, dass dieses durch die Berücksichtigung der natürlichen Umgebung der Studierenden ein gutes Abbild der tatsächlichen Lernprozesse liefert. Eine Randomisierung von verschiedenen Studierendengruppen mit unterschiedlichen Treatments innerhalb eines Semesters hätte zu problematischen Spill-Over Effekten führen können, da Studierende sich ungleich behandelt fühlen könnten. Eine mögliche Abhilfe in dem gegebenen Kontext würde gegebenenfalls darin bestehen, lediglich die Art der Wissensvermittlung (Lernvideos an Stelle von frontalem Dozentenvortrag) zu variieren und die anderen didaktischen Elemente konstant zu halten. Jedoch würde dieses Vorgehen an der Praxis vorbeiziele, da gerade durch den Einsatz von Lernvideos die Chance gegeben wird, die Präsenzlehre didaktisch zu verändern. Zudem sollten die Wahl desselben Dozenten, der weitgehend gleichen Tutoren und die Zugrundelegung desselben Curriculums für beide Settings diesbezüglich keine weitere Variation in den Effekten erwarten lassen.

Durch das Längsschnittdesign und die daraus resultierende hohe Anzahl an Befragungen während des Semesters ist eine Panelsterblichkeit durch Abwesenheit oder Teilnahmeverweigerung vonseiten der Studierenden nicht auszuschließen. Des Weiteren können Paneleffekte dadurch auftreten, dass die Bewusstseinsprozesse der Studierenden durch die Durchführung der Befragung und die sich oft wiederholenden Fragen explizit auf die Optimierung des Lehrdesigns der Veranstaltung konditioniert werden und sich normative Antwortmuster ergeben. Im Rahmen der Befragung im Sommersemester 2017 an der Johannes Gutenberg-Universität Mainz wurden die Befragungen mit der Verlosung von Incentives (iPads)<sup>5</sup> fest in die Veranstaltung eingebunden, was dazu führte, dass kaum befragungsbedingter Drop-Out vor Ort festgestellt werden konnte.

Dadurch, dass die Studierenden aus ihrem Studium eher den frontalen Vorlesungsstil und die Ballung des Workloads gegen Semesterende gewöhnt sind, ist es möglich, dass sie im Flipped Classroom-Design durch die verlagerte Lernaktivität und die erhöhten Freiheitsgrade während des Semesters eine erhöhte kognitive Belastung und Selbstregulationsverantwortung empfinden (Horz, 2015), welche das Antwortverhalten und die Teilnahmemotivation beeinträchtigen könnte. Daher wurde mit der Statistik I-Veranstaltung eine Veranstaltungsform gewählt, die sehr früh im Bachelorstudiengang an beiden Standorten besucht wird (im ersten oder zweiten Fach-

---

<sup>5</sup>Um die Vergleichbarkeit beider Messzeitpunkte zu gewährleisten, wird im Sommersemester 2018 das Konzept der Incentivierung analog und in gleicher Form genutzt.

semester), wodurch die Umgewöhnung etwas leichter fallen sollte. Mögliche Methodeneffekte (Leitgöb, 2017) durch die unterschiedlichen Befragungsformen (Paper-Pencil und Onlinebefragung) können dahingehend abgeschwächt werden, dass der Grad der Anonymisierung bei der Befragung in einer Massenveranstaltung von Studierenden womöglich im Vergleich zu einer sozial gänzlich entkontextualisierten Onlinebefragung ähnlich wahrgenommen wird, sodass Effekte aufgrund sozialer Erwünschtheit eher ausgeschlossen werden können. Zudem war auch in den Paper-Pencil-Befragungen das Ausmaß der Einbindung des Testleiters eingeschränkt und nur auf eine kurze Vorstellung des Projekts und eine kurze, anfängliche Instruktion begrenzt.

Bei der Untersuchung der Between-Design-Effekte ist es aufgrund der unterschiedlichen Jahrgänge denkbar, dass es kleine Unterschiede in den Eingangsvoraussetzungen der Studierenden zwischen der herkömmlichen Vorlesung und dem Flipped Classroom gibt. Zunächst gibt es keine Annahme, warum sich die Studierendenschaft in den Studiengängen der beiden Universitäten innerhalb eines Jahres signifikant verändern sollte. Dennoch werden bei der Untersuchung zum Vergleich beider Lehrdesigns Analysen mit Propensity Score Matching durchgeführt, um neben der Gruppenzugehörigkeit auch andere individuelle Eingangsvoraussetzungen der Probanden sowie institutionell-strukturelle und zeitliche Einflüsse möglichst gut kontrollieren zu können. Durch die inhaltlichen Überschneidungen zwischen den Methodenlehren verschiedener anderer Studiengänge, wie Politikwissenschaften, Soziologie und Psychologie, ist es möglich, Folgestudien in weiteren Studiengängen im gesamten deutschsprachigen Raum durchzuführen und damit die Befunde auf ihre Robustheit hin zu erproben oder sie unter Beleuchtung weiterer Perspektiven zu modifizieren. Das Flipped Classroom-Design sollte zum Teil auch auf die mathematische Propädeutik in den Wirtschaftswissenschaften (und ebenfalls in anderen Disziplinen) übertragbar sein, da deren Struktur (Aufbau, Übungskonzepte, numerischer Bezug) derjenigen der statistischen Kurse stark ähnelt. Insoweit können die zu generierenden Befunde nicht nur die bislang stark vernachlässigte hochschuldidaktische Lehr- und Lernforschung beleben, sondern bei Bewährung des Flipped Classroom-Treatments, sei es im Ganzen oder in Teilen, auch in einigen Bereichen einen Paradigmenwechsel in den Lehrmethoden anstoßen, wie er im Zuge der Bologna-Reform angestrebt worden ist.

## Literatur

Abeysekera, L. & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: definition, rationale and call for research. *Higher Education Research & Development*, 34(1), 1–14

Bandura, A. (2015). Self-regulation of motivation and action through internal standards and goal systems. In L. A. Pervin (Hrsg.), *Goal Concepts in Personality and Social Psychology* (S. 19–86). Hove u. New York: Psychology Press

Bishop, J. L. & Verleger, M. (2013, June). The Flipped Classroom: A Survey of the Research. Paper presented at the 120th ASEE Annual Conference & Exposition, Atlanta, Georgia

Eccles, J. & Wigfield, A. (2002). Motivational Beliefs, Values, and Goals. *Annual Review of Psychology*, 53, 109–132

Fischer, M. & Spannagel, C. (2012). Lernen mit Vorlesungsvideos in der umgedrehten Mathematikvorlesung. In J. Desel, J. M. Haake & C. Spannagel (Hrsg.), *DeLFI 2012 – Die 10. e-Learning Fachtagung Informatik* (S. 225–236). Bonn: Gesellschaft für Informatik

Förster, M. & Maur, A. (2015). Statistics Anxiety and Self-Concept of Beginning Students in the Social Sciences. *Zeitschrift für Hochschulentwicklung (ZfHE)*, 10(4), 67–90

Förster, M. & Maur, A. (2016, April): Analyzing Change in Students' Statistics Self-Concept and Anxiety. Discussion Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC

Garfield, J. & Ben-Zvi, D. (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review*, 75(3), 372–396

Handke, J. (2015). Shift Learning Activities – vom Inverted Classroom Mastery Model zum xMOOC. In N. Nistor & S. Schirlitz (Hrsg.), *Digitale Medien und Interdisziplinarität* (S. 113–123). Münster u. New York: Waxmann

Happ, R. & Zlatkin-Troitschanskaia, O. (2015). Vergleichende Analysen zur Heterogenität der Studierenden in wirtschaftswissenschaftlichen Studiengängen. In S. Harris-Hümmert, L. Mitterauer & P. Pohlenz (Hrsg.), *Heterogenität der Studierenden: Herausforderung für die Qualitätsentwicklung in Studium und Lehre, neuer Fokus für die Evaluation?* (S. 149–165). Bielefeld: UVW Universitätsverlag Webler

Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112

Horz, H. (2015). Medien. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (2. Aufl., S. 121–147). Berlin u. Heidelberg: Springer

Kauermann, G. (2015). Anwendungsorientierte Studiengänge und Fächerkombinationen – Statistik. In Springer Spektrum (Hrsg.), *Studien- und Berufsplaner Mathematik. Schlüsselqualifikationen für Technik, Wirtschaft und IT* (5. Aufl., S. 120–123). Wiesbaden: Springer Spektrum

Leitgöb, H. (2017). Ein Verfahren zur Dekomposition von Mode-Effekten in eine mess- und eine repräsentationsbezogene Komponente. In S. Eifler & F. Faulbaum (Hrsg.),

Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung (S. 51–98). Wiesbaden: Springer VS

Lippe, P. von der & Kladroba, A. (2008). Der unaufhaltsame Niedergang der Fächer Statistik und Ökonometrie in den Wirtschaftswissenschaften. *AStA Wirtschafts- und Sozialstatistisches Archiv*, 2(1), 21–40

Lo, C. K., Hew, K. F. & Chen, G. (2017). Toward a set of design principles for mathematics flipped classrooms. *Educational Research Review*, 22, 50–73

Macher, D., Paechter, M., Papousek, I. & Ruggeri, K. (2012). Statistics Anxiety, Trait Anxiety, Learning Behavior, and Academic Performance. *European Journal of Psychology of Education*, 27(4), 483–498

Marsh, H. (2007). Students' Evaluations of University Teaching: Dimensionality, Reliability, Validity, Potential Biases and Usefulness. In R. P. Perry & J. C. Smart (Hrsg.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective* (S. 319–383). Dordrecht, Netherlands: Springer

Means, B., Bakia, M. & Murphy, R. (2014). *Learning Online: What Research Tells Us About Whether, When, and How*. New York: Routledge

Moore, C. & Chung, C.-J. (2015). Students' Attitudes, Perceptions, and Engagement within a Flipped Classroom Model as Related to Learning Mathematics. *Journal of Studies in Education*, 5(3), 286–208

O'Flaherty, J. & Phillips, C. (2015). The Use of Flipped Classrooms in Higher Education: A Scoping Review. *Internet and Higher Education*, 25, 85–95

Pekrun, R., Goetz, T., Titz, W. & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement. *Educational Psychologist*, 37(2), 91–106

Pintrich, P. R., Smith, D. A. F., Garcia, T. & McKeachie, W. J. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor, Michigan: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning

Ramirez, C., Schau, C. & Emmioglu, E. (2012). The Importance of Attitudes in Statistics Education. *Statistics Education Research Journal*, 11(2), 57–71

Römmer-Nossek, B., Peschl, M. F. & Zimmermann, E. (2013). Kognitionswissenschaft. Ihre Perspektive auf Lernen und Lehren mit Technologien. In M. Ebner & S. Schön (Hrsg.), *Lehrbuch für Lernen und Lehren mit Technologien* (2. Aufl., S. 374–386). Berlin: epubli

Rožman, L., Lešer, V. J., Širca, N. T., Dermol, V. & Skrbinjek, V. (2014, June). Assessing student workload. Discussion Paper presented at the Management, Knowledge and Learning International Conference, Portorož, Slovenia

Ryan, R. M. & Deci, E. L. (2016): Facilitating and Hindering Motivation, Learning, and Well-Being in Schools. In K. R. Wentzel & D. B. Miele (Hrsg.), *Handbook of Motivation at School* (2. Aufl., S. 96–119). New York: Routledge

Schau, C. (2003). Survey of Attitudes Toward Statistics (SATS-36). Abgerufen am 15.10.2017 von <http://www.evaluationandstatistics.com/bizwaterSATS36monkey.pdf>

Tempelaar, D. T., Schim van der Loeff, S. (2011, August). The Development of Students' Subject Attitudes when Taking a Statistics Course. Paper presented at 58th World Statistics Congress of the International Statistical Institute, Dublin, Ireland

Zlatkin-Troitschanskaia, O., Förster, M., Brückner, S., Hansen, M. & Happ, R. (2013). Modellierung und Erfassung der wirtschaftswissenschaftlichen Fachkompetenz bei Studierenden im deutschen Hochschulbereich. *Lehrerbildung auf dem Prüfstand (Sonderheft)*, 6(1), 108–133

Manuskript eingereicht: 31.10.2017  
Manuskript angenommen: 04.07.2018

### **Anschrift der Autoren:**

Jun.-Prof. Dr. Manuel Förster  
Juniorprofessur für Wirtschaftspädagogik  
Johannes Gutenberg-Universität Mainz  
Jakob-Welder-Weg 9  
55099 Mainz  
E-Mail: [manuel.foerster@uni-mainz.de](mailto:manuel.foerster@uni-mainz.de)

Prof. Dr. Florian Heiß  
Lehrstuhl für Statistics und Econometrics  
Heinrich-Heine-Universität Düsseldorf  
Universitätsstr. 1  
40225 Düsseldorf  
E-Mail: [florian.heiss@hhu.de](mailto:florian.heiss@hhu.de)

Dr. Sigbert Klinke  
Lehrstuhl für Statistik  
Humboldt Universität zu Berlin  
Unter den Linden 6  
10099 Berlin  
E-Mail: [sigbert@wiwi.hu-berlin.de](mailto:sigbert@wiwi.hu-berlin.de)

Andreas Maur, M.Sc.  
Johannes-Gutenberg-Universität Mainz  
Lehrstuhl für Wirtschaftspädagogik  
Jakob-Welder-Weg 9  
55099 Mainz  
E-Mail: [anmaur@uni-mainz.de](mailto:anmaur@uni-mainz.de)

Prof. Dr. Thorsten Schank  
Johannes Gutenberg-Universität Mainz  
Lehrstuhl für Angewandte Statistik und Ökonometrie  
Jakob-Welder-Weg 4  
55099 Mainz  
E-Mail: schank@uni-mainz.de

Dr. Constantin Weiser  
Fachbereich Rechts- und Wirtschaftswissenschaften  
Jakob-Welder-Weg 4  
55099 Mainz  
E-Mail: constantin.weiser@uni-mainz.de

Jun.-Prof. Dr. Manuel Förster beschäftigt sich im Rahmen des Projekts mit individuellen und strukturellen Einflussfaktoren auf den Kompetenzerwerb im Fach Statistik unter Berücksichtigung der Chancen und Herausforderungen der digitalisierten Lehre in der akademischen Ausbildung.

Prof. Dr. Florian Heiß implementiert und beforscht semesterbegleitende, Online-Materialien (Lernvideos, Online-Tools, E-Quizze etc.), die es den Studierenden ermöglichen, regelmäßig ein Feedback zu deren fachstatistischen Lernstand zu erhalten.

Dr. Sigbert Klinke widmet sich der Forschung zur Unterstützung der Lehre durch multimedial-interaktive Lehrinhalte, beispielsweise mittels Wikis (MM\*Stat), elektronischen Aufgabensammlungen und Webcasts.

Andreas Maur befasst sich mit instruktionellen Bedingungsfaktoren motivationaler, volitionaler und emotionaler Facetten des Wissenserwerbs im Laufe von traditionellen und auf Blended-Learning basierenden Statistikveranstaltungen.

Prof. Dr. Thorsten Schank beforscht im Gebiet der Methodenlehre die Implementierung von E-Quizen, interaktiven Tutorien sowie fachstatistischen Aufgabensammlungen in wirtschaftswissenschaftlichen Großveranstaltungen.

Dr. Constantin Weiser erforscht im Rahmen der Didaktik der Statistik den Einsatz und die Integration handlungsorientierter Methoden (u. a. Lernen mit Videos, forschendes Lehren) sowie verschiedene Mediennutzungstypologien.

# Bestandsaufnahme, Hindernisse und Möglichkeiten des Einsatzes von E-Assessment an sächsischen Hochschulen

Jana Riedel, Kathrin Möbius

---

Der Beitrag stellt die Ergebnisse einer Studie zum E-Assessment an sächsischen Hochschulen vor. E-Assessment-Szenarien sind an sächsischen Hochschulen aktuell noch wenig verbreitet und werden von den Lehrenden überwiegend zu Übungszwecken (formatives Assessment) und nur selten zu Prüfungszwecken (summatives Assessment) eingesetzt. Dabei existieren neben fachspezifischen Unterschieden im Einsatz der unterschiedlichen Szenarien auch Unterschiede in Bezug auf die Veranstaltunggröße. Elektronische Tests als eines von vier untersuchten E-Assessment-Szenarien werden vor allem in der Mathematik, in den Naturwissenschaften und in Massenveranstaltungen eingesetzt. Vorrangiges Ziel der Lehrenden ist die Steigerung der Lehrqualität. Potenziale werden in der didaktischen Integration von Prüfungs- und Lehrtätigkeit und den Möglichkeiten der Individualisierung von Lernprozessen gesehen. Die größten Hindernisse vor allem für das summative E-Assessment sehen viele Lehrende in der unklaren Rechtslage und der mangelnden Infrastruktur. Die Studienergebnisse lassen den Schluss zu, dass eine Verbesserung der Rahmenbedingungen zu einer weiteren Verbreitung von E-Assessment-Szenarien beitragen kann.

---

## 1 E-Assessment in der Hochschullehre

Digitale Medien können Lehrprozesse in verschiedenen Bereichen unterstützen. Neben der Digitalisierung von Wissensbeständen und der medial vermittelten Darbietung von Inhalten stehen auch Werkzeuge zur Kommunikation und Zusammenarbeit sowie zur Umsetzung (teil-)digitalisierter Prüfungsformen zur Verfügung. Letztere gewinnen in der Hochschullehre fortlaufend an Bedeutung. Seit den 1990er Jahren werden eigenständige E-Assessment-Formate erprobt und zunehmend eingeführt (vgl. Michel & Goertz 2015, S. 12).

Unter E-Assessment-Szenarien werden hier auf elektronischen Informations- und Kommunikationstechnologien basierende Verfahren verstanden, die der lehrzielbezogenen Bestimmung, Beurteilung, Bewertung, Dokumentation und Rückmeldung von Lernvoraussetzungen, Lernstand und Lernergebnissen dienen (vgl. Bloh 2006, zit. in Seufert & Brahm 2007) und damit das Lehren und Lernen durch wichtige Feedback-Prozesse bereichern (vgl. u. a. Hattie 2013). E-Assessments begleiten nach dieser Auffassung daher den gesamten Lernprozess und umfassen nicht nur die Prüfung am Ende eines Semesters.

Im Sinne des *diagnostischen* Assessments werden die Voraussetzungen der Lernenden *vor Beginn* eines Lernprozesses bspw. in Zulassungstests sowie Eignungs- und Einstufungsverfahren erfasst. An deutschen Hochschulen steigt in diesem Bereich auch die Verbreitung von fachspezifischen Online-Self-Assessments zur Studienorientierung. Die Begleitung und Förderung von Lernprozessen durch kontinuierliche Rückmeldungen zum Lernfortschritt wird durch *formative* Assessment-Szenarien unterstützt. In diesem Bereich erhält in der Hochschullehre unter anderem der Einsatz von Live-Abstimmungen mit Audience-Response-Systemen (ARS) und Clickern zunehmend Aufmerksamkeit. Die Prüfung von Lernleistungen *nach Abschluss* eines Lernprozesses wird als *summatives* Assessment bezeichnet. Die wachsende Zahl an Assessment-Centern unterstützt diese Form des Assessments unter den Stichworten E-Klausuren, E-Prüfungen und E-Examinations.

Zur Gestaltung von E-Assessment-Szenarien können dabei unterschiedliche Werkzeuge, wie elektronische Tests, Audience-Response-Systeme (ARS) zur Realisierung von Live-Abstimmungen oder E-Portfolio-Werkzeuge genutzt werden (vgl. z. B. Michel & Goertz 2015). Auch die Nutzung digitaler Werkzeuge (bspw. des Web 2.0) durch Studierende zur Erstellung eigener Medienprodukte (wie bspw. eigene Webseiten/Blogs, Videos, Podcasts, Lernmodule) wird hier mit Reinmann (2007) als eine Möglichkeit des „*Erfassens, Sammelns, Einschätzens und Interpretierens von Leistungen oder Wissensartefakten*“ verstanden (Reinmann 2007, S. 18).

Die Umsetzungsvarianten des E-Assessments sind dementsprechend vielfältig und unterschiedlich stark erforscht. Es ist unstrittig, dass Assessment-Angebote (ob elektronisch oder nicht) vielfältige Funktionen wie bspw. Aufmerksamkeitssteuerung, Feedback, Motivation, Orientierung im Lernprozess übernehmen können (vgl. z. B. Boud 1995, Reeves 2006, Roediger & Karpicke 2006, Cilliers, Schuwirth, Adendorff, Herman & van der Vleuten 2010, Gruttmann 2010). Die Konzeption von Assessment-Szenarien sollte daher als fester Bestandteil der Lehrveranstaltungsplanung erachtet und Prüfungsformen und Lehrmethoden aufeinander abgestimmt werden (vgl. z. B. Biggs 2003, Wass, van der Vleuten, Shatzer, Jones 2001, Reinmann 2007, Biggs & Tang 2011). Hierfür verwendet und verbreitet die Hochschuldidaktik zunehmend das Konzept des *Constructive Alignment* (vgl. Biggs & Tang 2011, Baumert & May 2013).

Die elektronisch gestützte Umsetzung der Assessment-Angebote bietet Vorteile in der Vielfalt und Realitätsnähe der Aufgabenformate sowie der Gestaltung kompetenzorientierter alternativer Prüfungsformen (wie bspw. E-Portfolios oder das Assessment studentischer Medienprodukte) (vgl. z. B. Reinmann 2007). Diese sind jedoch erst in Ansätzen erforscht, wobei ein Schwerpunkt auf der Bewertung von Szenarien mit E-Portfolios liegt (vgl. z. B. Baumgartner 2012, Bisovsky & Schaffert 2009).

Vielfältigere Forschungsergebnisse existieren hingegen zum Assessment mit elektronischen Tests, die überwiegend auf Aufgaben im Antwort-Wahl-Verfahren basieren und automatisiert ausgewertet werden können. Hierfür werden verschiedene Vor- und Nachteile zusammengetragen (zusammengefasst z. B. bei Jurecka & Hartig 2007) und die konkrete Gestaltung von Multiple-Choice-Aufgaben erforscht (vgl. z. B. Haladyna 2004, Lindner, Strobel & Köller 2015). Aus ökonomischer Perspektive wird positiv beurteilt, dass durch automatisierte Korrekturen vor allem im diagnostischen und formativen Bereich in der Hochschullehre häufiger bzw. überhaupt erst Assessment-Angebote geschaffen werden können (zusammengefasst bei Gruttmann 2010, S. 2). Bedenken betreffen vor allem die rechtliche Beurteilung und Zulässigkeit elektronischer Aufgaben in summativen Prüfungsszenarien.

Die aktuelle Studie „Digitales Prüfen und Bewerten im Hochschulbereich“ trägt für die verschiedenen Assessment-Formen Beispiele an deutschen und internationalen Hochschulen zusammen und beschreibt damit den Möglichkeitsraum digitaler Assessments im Hochschulbereich. Die Aussagen zur Verbreitung einzelner Szenarien werden in dieser Studie nicht quantifiziert, dennoch verweisen sie darauf, dass die Verbreitung sowohl von Hochschule zu Hochschule und als auch innerhalb der Hochschulen sehr stark differiert und dass der Einsatz häufig vom Engagement einzelner Lehrender oder projektbasierter Initiativen abhängig ist (vgl. Michel & Goertz 2015, S. 12).

Ziel des Projekts „E-Assessment in Sachsen. Ist-Stand und Bedarf“ (Laufzeit: 01.09.2015–31.12.2016, Förderung: Sächsisches Staatsministerium für Wissenschaft und Kunst) war es daher, die Verbreitung des E-Assessments an sächsischen Hochschulen zu erheben und vor allem den Unterstützungsbedarf der Lehrenden zu identifizieren, um daraus Handlungsempfehlungen für eine Weiterentwicklung und flächendeckende Verbreitung von E-Assessment-Angeboten mit speziellem Fokus auf E-Prüfungen abzuleiten.

Die nachfolgend beschriebenen Untersuchungen umfassen daher sowohl diagnostische, formative als auch summative Assessment-Verfahren, die mit Hilfe von unterschiedlichen digitalen Werkzeugen wie elektronischen Tests, Audience-Response-Systemen (ARS), E-Portfolio-Systemen oder Werkzeugen zur Erstellung eigener Medienprodukte durch die Studierenden realisiert werden können.

## **2 Stichprobe und methodisches Vorgehen**

Die Erfassung des Ist-Stands und des Unterstützungsbedarfs zum Einsatz von E-Assessment-Szenarien in der Hochschullehre im Freistaat Sachsen erfolgte zweistufig. Zunächst wurde eine Online-Befragung Hochschullehrender als Vollerhebung an allen sächsischen Hochschulen durchgeführt. Anschließend erfolgte eine qualitative

Vertiefung der quantitativen Daten mittels zwei Fokusgruppeninterviews ( $N = 7$ ). Die Gestaltung von Lehr- und Prüfungsszenarien ist zentrale Aufgabe der Hochschullehrenden, die Gestaltung digital gestützter Szenarien gründet sich aktuell zu einem wesentlichen Teil auf Eigeninitiative der Hochschullehrenden (vgl. Schmid, Goertz, Radomski, Thom & Behrens 2017, S. 6). Die Erhebungen stützen sich daher auf die Aussagen dieser Untersuchungsgruppe.

Um möglichst viele Hochschullehrende für die Befragung zu gewinnen, wurde der Feldzugang über die E-Mailverteiler für die Mitarbeitenden an den jeweiligen Hochschulstandorten gewählt. Die Mitarbeitenden wurden folglich über eine Einladung zur Teilnahme an der Umfrage aufgerufen und um ihre Unterstützung gebeten, sofern sie in der Lehre aktiv sind bzw. es waren. Die Online-Befragung erfolgte im Zeitraum von Februar bis April 2016 im Freistaat Sachsen. Insgesamt haben 545 Lehrende den Fragebogen beendet.<sup>1</sup> Die Stichprobe setzt sich aus 38 Prozent weiblichen und 55 Prozent männlichen Studienteilnehmenden zusammen. Damit entspricht die Geschlechterverteilung der Befragung ungefähr der Verteilung der Geschlechter der Mitarbeitenden im Freistaat Sachsen (60 Prozent männlich, 40 Prozent weiblich). Es haben Lehrende aller Fachbereiche an der Befragung teilgenommen. Dennoch ist zu betonen, dass der Bereich der Humanmedizin/Gesundheitswissenschaften im Vergleich zur Fachbereichsverteilung an den Hochschulen in Sachsen unterrepräsentiert und der Bereich der Geistes-, Sprach-, Literatur- und Kulturwissenschaften etwas überrepräsentiert ist.

Die Teilnehmenden der Fokusgruppen wurden im Rahmen der Online-Befragung identifiziert. Sie haben zum Abschluss der Online-Befragung angegeben, sich zum Thema E-Assessment weiter vernetzen zu wollen und können bereits auf eigene Erfahrungen in diesem Bereich zurückgreifen. Daraufhin wurden diese Experten zu zwei Fokusgruppenterminen im Sommer 2016 eingeladen, wovon ein Termin mit vier Teilnehmenden an einem Hochschulstandort und ein zweiter Termin mit drei Teilnehmenden unterschiedlicher Hochschulen online in einem virtuellen Klassenzimmer von Adobe Connect stattfand.

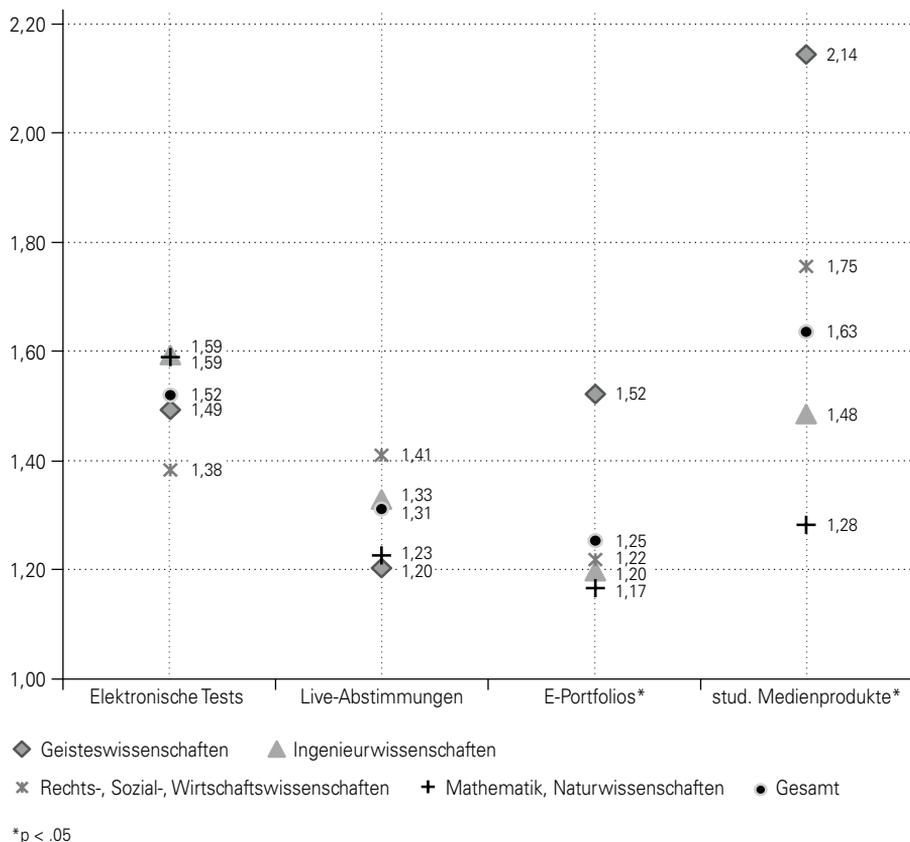
### **3 Verbreitung von E-Assessment-Szenarien an sächsischen Hochschulen**

Die quantitative Untersuchung hat gezeigt, dass die Verbreitung von E-Assessment-Formaten in der Hochschullehre derzeit noch gering ist. So vergeben nur knapp 11 Prozent der Teilnehmenden der Online-Befragung ( $N = 545$ ) unter sächsischen Hochschullehrenden das Schlagwort E-Assessment zur Charakterisierung ihrer Lehrveranstaltung. Die Schlagworte Peer-Assessment (8,2 Prozent) und Live-Feedback (5,2 Prozent) werden noch seltener vergeben.

<sup>1</sup>Die Nettobeteiligung lag insgesamt bei 1171 Personen. Die Rücklaufquote kann aufgrund der fehlenden Aussagen über die Anzahl der Personen in den Mitarbeiterverteilern und der Anzahl der in der Lehre aktiven Mitarbeitenden nicht bestimmt werden.

Darüber hinaus wurden die Lehrenden befragt, welche Formate und Werkzeuge sie für die Umsetzung von E-Assessment-Szenarien nutzen (siehe Kap. 1). Knapp ein Viertel der Befragten setzt bereits elektronische Tests ein, 17 Prozent nutzen Live-Abstimmungen und 14 Prozent verwenden E-Portfolios. Das häufigste Szenario (35 Prozent) ist die Erstellung eigener Medienprodukte durch Studierende. Dieses Ergebnis überrascht in Bezug auf die Tatsache, dass dieses Szenario in der Forschung bisher nur wenig Aufmerksamkeit erhalten hat (s. Kap. 1).

**Abbildung 1:** Nutzung von E-Assessment-Einsatzmöglichkeiten nach Fachbereichen (N = 408 – 412); Mittelwerte einer Skala von 1 (=nie) bis 5 (=immer)

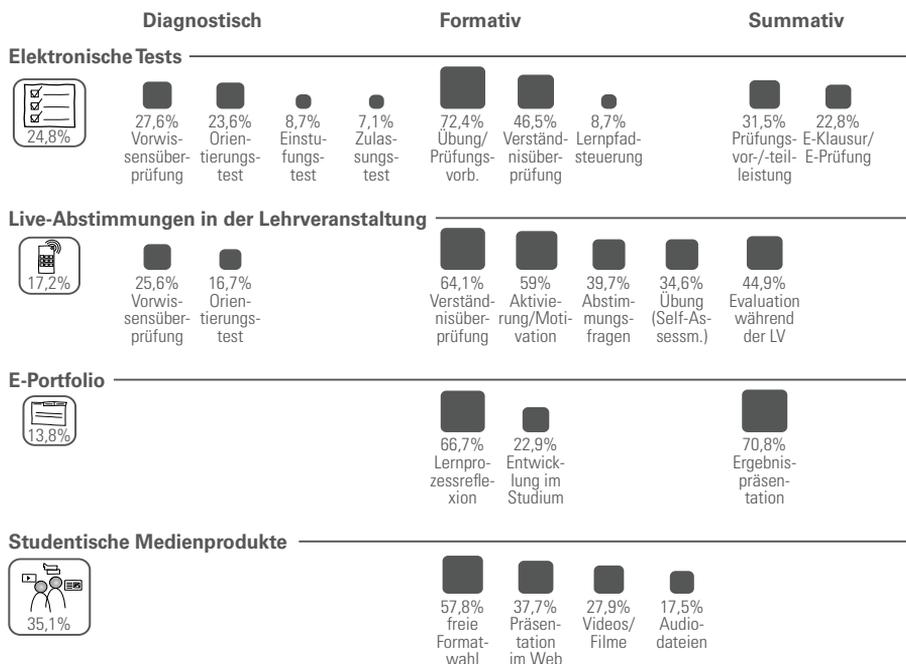


Es zeigten sich hoch signifikante Unterschiede in der Nutzungshäufigkeit dieser E-Assessment-Szenarien in Bezug auf den Fachbereich und die Veranstaltungsgröße. So sollen die Studierenden vor allem in den Geistes- sowie Rechts-/Sozial-/Wirtschaftswissenschaften häufiger E-Portfolios ( $\chi^2 = 13.51$ ,  $df = 3$ ,  $N = 408$ ,  $p = .004$ ) führen oder eigene Medienprodukte erstellen ( $\chi^2 = 31.02$ ,  $df = 3$ ,  $N = 408$ ,  $p = .000$ ). Elektronische Tests sind das einzige digital gestützte Szenario, das von den befragten

Lehrenden aus dem Bereich Mathematik/Naturwissenschaften am häufigsten genutzt wird (siehe Abb. 1). In Bezug auf die Veranstaltungsgröße konnten hoch signifikante Unterschiede bei der Nutzung von elektronischen Tests und Live-Abstimmungen (mit ARS) festgestellt werden. Demnach nimmt die Wahrscheinlichkeit für eine häufigere Nutzung mit steigender Veranstaltungsgröße zu ( $\tau_b = .156$ ,  $p = .000$ ,  $N = 446$  für elektronische Tests bzw.  $\tau_b = .141$ ,  $p = .001$ ,  $N = 442$  für Live-Abstimmungen). Vor allem die elektronischen Tests werden überwiegend (von 66,7 Prozent der Nutzenden dieses Formates) in Massenveranstaltungen mit über 300 Studierenden genutzt.

Die Nutzung von E-Assessment-Formaten erwies sich hingegen als unabhängig von Alter, Status und Lehrerfahrung der befragten Lehrenden. Auch die Anzahl der zu betreuenden Lehrveranstaltungen und der Hochschultyp zeigten keinen signifikanten Einfluss auf den Einsatz von E-Assessment-Szenarien.

**Abbildung 2:** E-Assessment-Szenarien im Detail (Angaben in Prozent, Mehrfachantworten möglich,  $N = 48\text{--}154$ )



E-Assessment findet bei den befragten Lehrenden vor allem im formativen Bereich statt (siehe Abb. 2). Bei den elektronischen Tests überwiegt der Einsatz zur Übung/Prüfungsvorbereitung (72,4 Prozent) und Verständnisüberprüfung (46,5 Prozent). Der Einsatz für das summative Prüfen in E-Klausuren liegt weit dahinter zurück (22,8 Prozent). Als didaktische Funktion des Einsatzes elektronischer Tests benennen die

befragten Lehrenden daher an erster Position (65,3 Prozent) die Festigung des Wissens und erst danach (56,8 Prozent) die Kontrolle und Bewertung, gefolgt von der Entwicklung von Fähigkeiten und Fertigkeiten (Übung) (41,1 Prozent). Dabei haben Lehrende aus den Bereichen Geisteswissenschaften sowie Mathematik/Naturwissenschaften diese Funktion signifikant häufiger ausgewählt als Lehrende der Fachbereiche Rechts-/Sozial-/Wirtschaftswissenschaften sowie Ingenieurwissenschaften ( $\chi^2 = 12.46$ ,  $df = 1$ ,  $N = 95$ ,  $p = .006$ ).

Für Live-Abstimmungen und die Erstellung studentischer Medienprodukte wählen die befragten Lehrenden ebenfalls formative Einsatzziele, in denen die Kontrolle und Bewertung nicht vordergründig ist. Nur für die E-Portfolios wählen sie die didaktische Funktion „Kontrolle und Bewertung“ an erster Stelle aus (41,5 Prozent), gefolgt von Systematisierung des Wissens und Könnens (39,6 Prozent) und Festigung des Wissens (34 Prozent).

## **4 Potenziale des E-Assessment-Einsatzes**

### **4.1 Motivationsgründe für den Einsatz von elektronischen Tests**

Befragte, die bereits elektronische Tests nutzen ( $N = 197$ ), wurden zusätzlich zu den Gründen für die Nutzung dieses Formats befragt. Hierbei wurde die Zustimmung zu in der Literatur benannten Vorteilen des E-Assessments in dichotomer Form und mit Möglichkeit zur Mehrfach-Nennung abgefragt (zusammengefasst bspw. bei Jurecka & Hartig 2007). Dabei kommen sowohl didaktische als auch ökonomische Aspekte zum Tragen. Anhand einer Faktorenanalyse wurde ermittelt, inwiefern sich die Gründe zur Nutzung elektronischer Tests in übergeordneten Nutzungsmotiven zusammenführen lassen. Es konnten insgesamt elf Items anhand einer Hauptkomponentenanalyse mit orthogonaler Varimax-Rotation und Kaiser-Normalisierung in drei Faktoren zusammengefasst werden.<sup>2</sup> Die Tabelle zeigt die rotierten Faktorladungen, die aufgrund der Inhalte der Items wie folgt interpretiert werden:

Faktor 1 umfasst Items, die sich auf eine Qualitätsverbesserung des Angebots für Studierende beziehen und das Lehrveranstaltungsangebot als Ganzes betrachten, er wird daher mit „Lehrqualität steigern“ bezeichnet. Die Items dieses Faktors werden von den Befragten am häufigsten als Gründe für den Einsatz elektronischer Tests angegeben (Lehre verbessern: 53,7 Prozent, Feedback: 62,6 Prozent, Übungsmöglichkeiten: 63,4 Prozent). Die Items, die in Faktor 2 gebündelt werden, beziehen sich auf Qualitätsmerkmale eines konkreten Testangebots und werden unter dem Titel „Testqualität steigern“ zusammengefasst. In Faktor 3 werden Items gruppiert, die die Durchführungseffizienz von Testszenarien betreffen. Er wird demnach mit „Effizienz

<sup>2</sup>Die Analyse zur Eigenwertbestimmung der Items ergab, dass drei Komponenten Eigenwerte von  $> 1$  (Kaiser-Kriterium) hatten und zusammen 55,18 Prozent der Varianz aufklärten.

steigern“ betitelt. Aus diesem Cluster sticht das Item „Korrekturzeit minimieren“ in Bezug auf die Häufigkeit der Nennungen heraus. Es wird mit 52 Prozent an vierter Stelle der Nutzungsgründe genannt.

**Tabelle:** Rotierte Komponentenmatrix

	1	2	3
Lehre verbessern	<b>,697</b>	,215	,101
Studierenden Feedback geben	<b>,744</b>	,161	–,114
(mehr) Übungsmöglichkeiten anbieten	<b>,756</b>	–,113	,221
Qualität der Tests und Prüfungen verbessern	,442	<b>,563</b>	,184
Lesbarkeit von Antworten verbessern	–,094	<b>,668</b>	,186
Auswertungsobjektivität verbessern	,025	<b>,566</b>	,380
Statistische Item-Analyse erleichtern	,186	<b>,764</b>	–,046
Fragen besser verwalten (Fragenpools anlegen)	,121	<b>,627</b>	,175
Vielfalt der Fragetypen erhöhen	–,026	,426	<b>,677</b>
Korrekturzeit minimieren	,320	–,251	<b>,640</b>
langfristig Kosten sparen	,259	,163	<b>,658</b>

Extraktionsmethode: Hauptkomponentenanalyse.

Rotationsmethode: Varimax mit Kaiser-Normalisierung. ( $N = 136$ ,  $s^2 = 55,18\%$ )

In der Einschätzung dieser Potenziale lassen sich ebenfalls fachspezifische Tendenzen feststellen. Vor allem Lehrende aus den Geisteswissenschaften geben häufiger an, mehr Übungsmöglichkeiten anbieten zu wollen als diejenigen der anderen Bereiche (87,5 Prozent der Lehrenden im Bereich Geisteswissenschaften gegenüber 68,2 Prozent im Bereich Rechts-, Sozial- und Wirtschaftswissenschaften, 61,1 Prozent im Bereich Mathematik/Naturwissenschaften und 48,7 Prozent im Bereich Ingenieurwissenschaften). Lehrende aus dem Bereich Ingenieurwissenschaften nennen am häufigsten die Einsparung von Korrekturzeiten als Motivationsgrund für den Einsatz von elektronischen Tests (56,4 Prozent Ingenieurwissenschaften gegenüber 44,4 Prozent Mathematik/Naturwissenschaften, 40,9 Prozent Rechts-/Sozial- und Wirtschaftswissenschaften und 31,3 Prozent Geisteswissenschaften).<sup>3</sup>

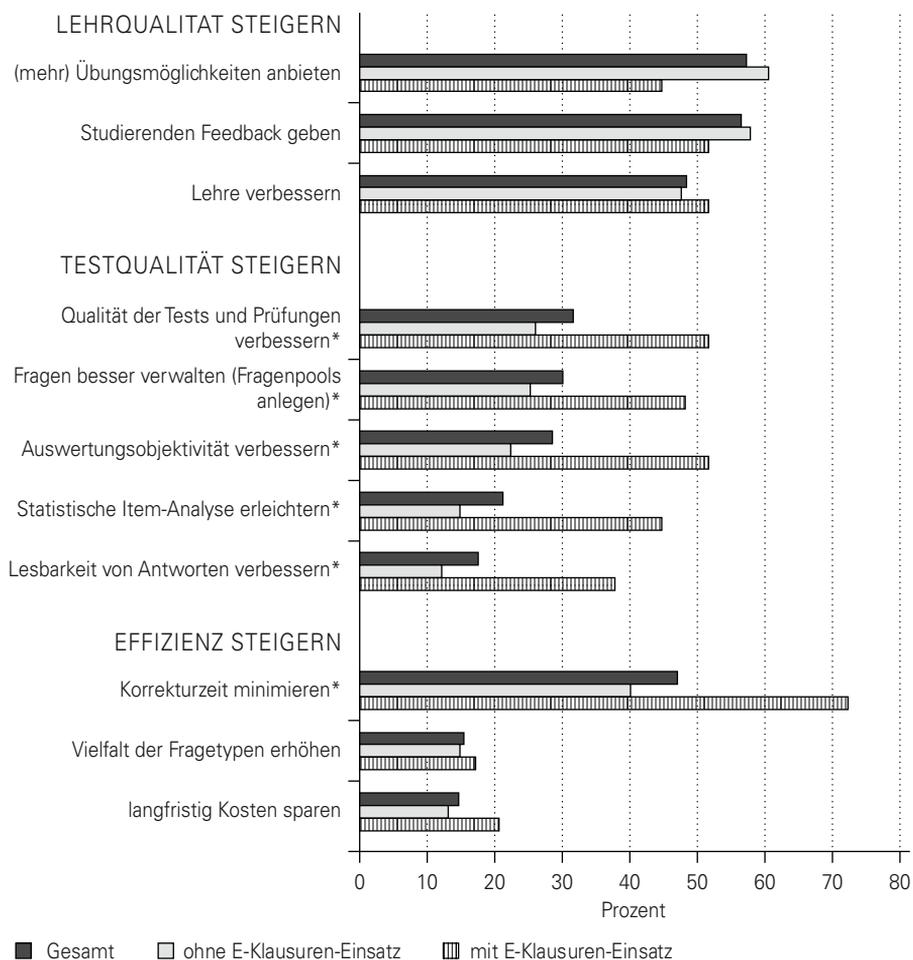
Deutliche Unterschiede zeigen sich bei der Bewertung zwischen Lehrenden, die die elektronischen Tests bereits in E-Klausuren<sup>4</sup> einsetzen, und denjenigen Befragten, die zwar elektronische Tests einsetzen, jedoch noch nicht elektronisch prüfen. Erstere zielen deutlich häufiger auf die Verbesserung der Testqualität. Für diesen Faktor zeigen alle Items signifikante Unterschiede in der Bewertung der Nutzungsmotivation

<sup>3</sup>Aufgrund zu geringer Fallzahlen konnten hier keine Signifikanzen berechnet werden.

<sup>4</sup>Die Lehrenden nutzen für diese E-Klausuren sowohl elektronisch auswertbare Aufgaben (93,1 Prozent) als auch Freitext-Aufgaben (72,4 Prozent) (Mehrfachnennungen möglich).

( $\chi^2 = [5.75, 12.14]$ ,  $df = 1$ ,  $N = 136$ ,  $p \leq .016$ ). Der Hauptanreiz für das Durchführen elektronischer Prüfungen ist für die Nutzenden dieses Formats jedoch, die Korrekturzeit zu minimieren (72,4 Prozent) (siehe Abb. 3). Bezüglich der Korrekturzeiten sei an dieser Stelle darauf hingewiesen, dass vor allem die bessere Lesbarkeit von Freitext-Aufgaben zu einer höheren Zeitersparnis führt als die ohnehin schon effiziente Form des Prüfens mit Multiple-Choice-Fragen (vgl. Schulz & Apostolopoulos 2011). In ihrer Studie berechneten die Autoren eine Zeitersparnis von 33 Prozent für die Korrektur von Freitext-Aufgaben (vgl. ebd., S. 38).

**Abbildung 3:** Bewertung von Potenzialen elektronischer Tests nach Befragten mit und ohne Einsatz von E-Klausuren (Angaben in Prozent, Mehrfachantworten möglich,  $N = 136$ )



\* $p < .05$

Durch die qualitative Erhebung mit Erfahrungsträgern aus dem Bereich des E-Assessments können die in der Online-Befragung erhobenen Chancen bestätigt und konkretisiert werden. Die von den Experten genannten Potenziale lassen sich denselben drei übergeordneten Kategorien (Lehrqualität steigern, Testqualität steigern, Effizienz steigern) zuordnen. Den Faktor „Lehrqualität steigern“ ergänzen die Interviewten um Aspekte, die vor allem auch hochschuldidaktische Forderungen adressieren. Gerade die Chancen zur Realisierung des Konzepts des „Constructive Alignment“ werden mehrfach betont. Das Konzept fordert eine Kohärenz von Lernprozessen und Assessments, wobei letztere durch zusätzliche elektronisch gestützte Übungen ermöglichen, *„(...) dass es mehr auch um die Verknüpfung von Lernen und Prüfen geht, dass Prüfen auch vielmehr ein lernförderliches Instrument wird“* (Fokusgruppe1, Z. 75ff.).

Auch die Forderung nach Individualisierung von Lerngelegenheiten für eine immer heterogener werdende Studierendenschaft wird nach Ansicht der Experten durch E-Assessment-Angebote unterstützt, indem diese eine Orientierungs- und Leitfunktion übernehmen und den Lehrenden die Möglichkeit zur Einschätzung des Leistungsstands ihrer Studierenden geben, auf die sie spezifisch reagieren könnten. Außerdem soll durch E-Assessment-Angebote die Eigeninitiative und -aktivität sowie das kontinuierliche Lernen angeregt werden, indem *„(...) die Studenten aktiviert werden schon im Rahmen des Semesters etwas zu machen und dann ganz anders in die Seminare oder in die Vorlesungen reingehen“* (Fokusgruppe 2, Z. 368f.).

Weiterhin verweisen die Experten auf eine Erhöhung von Transparenz, Fairness und Chancengleichheit. Dieser Aspekt wird gleichzeitig auch als Vision für zukünftige Einsatzmöglichkeiten elektronisch gestützter Assessments gesehen: *„Dass man also sagen kann, wenn man in Sachsen studiert, dann hat man Chancengleichheit. Vielleicht würde man sogar mit Hochschulen in anderen Bundesländern oder mit Partnerhochschulen im Ausland dahin kommen, dass man sowas für einzelne Lehrveranstaltungen definiert. Das wäre dann auch ein weiterer großer Gewinn für die Studierenden.“* (Fokusgruppe 1, Z. 166ff.).

## 4.2 Visionen zum Einsatz von E-Assessment

Hochschulübergreifende, gemeinsame Aufgabenpools werden hier noch als Vision behandelt, obwohl sich gerade in Sachsen bereits ein hochschulübergreifender Aufgabenpool für Mathematik-Aufgaben entwickelt, der von Lehrenden unterschiedlicher Hochschulen erweitert und in formativen Settings genutzt wird. Die beteiligten Lehrenden selbst nehmen die Zusammenarbeit dabei als Erleichterung war. Die hochschulübergreifende Entwicklung von Aufgabenpools birgt ebenfalls ein Potenzial für

die Qualitätssicherung von Tests. Dennoch erwachsen aus der Praxis geteilter Aufgabenpools gerade im Hinblick auf den Einsatz in summativen Prüfungsszenarien rechtliche Fragestellungen, die sowohl urheberrechtliche als auch prüfungsrechtliche Belange betreffen. Die Nutzung geteilter Aufgabenpools in Prüfungsszenarien bedarf daher einer sorgfältigen Prüfung, bevor diese an den Hochschulen praktisch umgesetzt werden kann. Hinzu kommt die Bereitschaft der Lehrenden, Inhalte zu teilen und weiter zu verwenden, die aktuell noch bezweifelt wird: *„das wird aber auch von den meisten Dozenten einfach kritisch gesehen, weil Austausch nicht nur Nehmen, sondern auch Geben heißt, und letztendlich der Austausch von Übungsaufgaben letztendlich einen Einblick in die Lehre und damit auch Lehrqualität der einzelnen Dozenten, ja, ermöglicht“* (Fokusgruppe 2, Z. 509 ff.).

Als weitere Vision wird der Einsatz von E-Examinations-on-Demand genannt, die in örtlichen E-Assessment-Centern realisiert werden können: *„Und meine Vorstellung wäre jetzt, wenn man ein E-Assessment-Center hätte und ausreichend Personal und diese ganzen E-Assessments laufen, dann könnte ich mir vorstellen, wäre es organisatorisch vorstellbar und machbar, dass die Studierenden sagen: ‚Jetzt fühle ich mich fit. Alle Voraussetzungen habe ich erbracht, jetzt suche ich mir einen geeigneten Prüfungstermin für die Prüfung‘“* (Fokusgruppe 1, Z. 128ff.). Der visionäre Charakter dieser Idee wird nicht nur in Bezug auf die rechtlichen Rahmenbedingungen deutlich, in Sachsen fehlt hierzu auch die notwendige Infrastruktur (siehe Kapitel 5.1).

Für eine weitere Vision müssen nicht nur rechtliche und organisatorische Hürden überwunden, sondern auch theoretische Grundlagen geschaffen werden. *„Was mir natürlich dann gleich noch einfällt (...) ist natürlich die Anknüpfung an die probabilistische Testtheorie „adaptives Prüfen“. (...) Also wie kann es mir gelingen, den Studierenden mit möglichst wenig Fragen möglichst gut zu diagnostizieren – nein das ist zu viel gesagt, aber zu prüfen. Und das, denke ich, ist noch eine große Herausforderung“* (Fokusgruppe 1, Z. 203ff.).

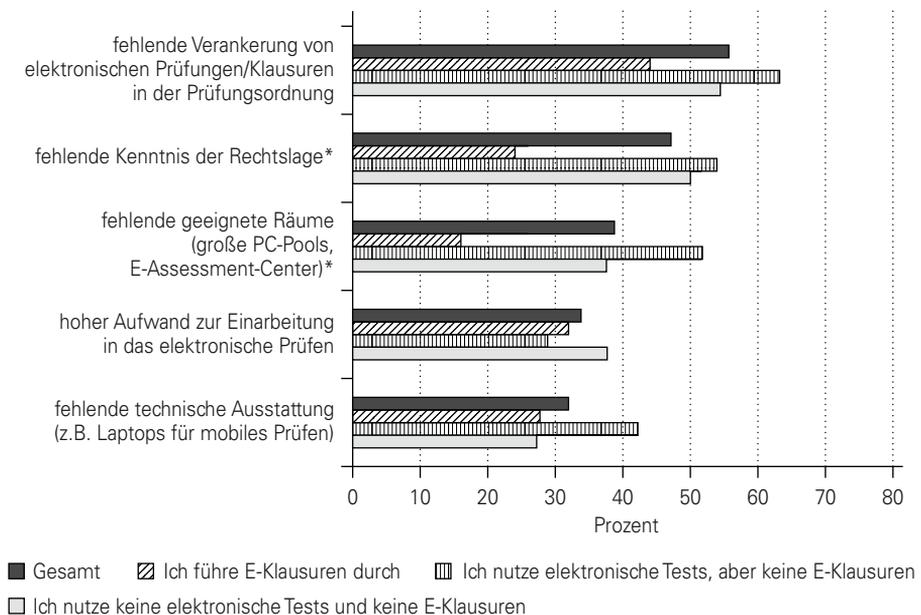
## **5 Herausforderungen für den Einsatz von E-Prüfungen**

### **5.1 Aktuelle Rahmenbedingungen als Hindernis**

Da gerade für das summative Prüfen die Potenziale hinsichtlich einer Erhöhung der Kosteneffizienz und Steigerung der Prüfungsqualität besonders hoch eingeschätzt werden (s. Kapitel 4.1), wird im Folgenden diskutiert, warum dieses Szenario bisher eher selten eingesetzt wird. Die Hindernisse für den Einsatz elektronischer Prüfungen

lassen sich mit drei Schlagworten zusammenfassen: Recht, Infrastruktur und Zeit. Die fünf am häufigsten genannten Herausforderungen, weswegen die befragten Lehrenden noch keine E-Prüfungen durchführen, sind demnach die fehlende Verankerung elektronischer Prüfungen in den Prüfungsordnungen (55,8 Prozent), die fehlende Kenntnis der Rechtslage (47,3 Prozent), fehlende geeignete Räume (38,8 Prozent), der hohe Aufwand zur Einarbeitung in das elektronische Prüfen (33,9 Prozent) und die fehlende technische Ausstattung (z. B. mit Laptops) (32,1 Prozent). Dabei zeigen sich hinsichtlich der Bewertung dieser Aspekte deutliche Unterschiede zwischen den Befragten, die bereits E-Klausuren einsetzen und denjenigen, die diese noch nicht nutzen (siehe Abb. 4).

**Abbildung 4:** Herausforderungen beim Einsatz elektronischer Prüfungen nach Befragten mit und ohne Einsatz elektronischer Klausuren (Angaben in Prozent, Mehrfachantworten möglich, N = 160)



\*  $p < .04$

Diese Befunde lassen die Vermutung zu, dass eine Beseitigung dieser Probleme eine flächendeckende Verbreitung des Einsatzes elektronischer Prüfungen ermöglichen wird. Diese Vermutung bestätigten alle Experten in den Fokusgruppen uneingeschränkt: „Und ich denke, sobald die Infrastrukturen gegeben sind, kriegt man es breitenwirksam hin und ich denke, das wäre der erste Schritt, Lehrenden und Studierenden die Angst

zu nehmen.“ (Fokusgruppe 1, Z. 238ff.), „*Ich würde einfach mal beginnen und sagen: dann werden die gemacht.*“ (Fokusgruppe 1, Z. 46), „*Ja, wenn diese Bedingungen gegeben wären, dann könnte man mit Sicherheit solche Prüfungen ins Auge fassen, insbesondere dort, wo man Studiengänge hat mit hohen Zahlen von Teilnehmern*“ (Fokusgruppe 2, Z. 58ff.).

Die Rahmenbedingungen sehen in Sachsen jedoch anders aus. Bisher ist das elektronische Prüfen nur an einzelnen Fakultäten einer sächsischen Hochschule in den Prüfungsordnungen verankert. Eine Überarbeitung der Prüfungsordnungen ist an einigen Hochschulen zukünftig vorgesehen, wobei noch unsicher ist, wann veränderte Regelungen in Kraft treten werden.

Das einzige E-Assessment-Center in Sachsen mit 75 Prüfungsplätzen wird an der Hochschule betrieben, die ihre Prüfungsordnungen bereits angepasst hat. Dieses wird unter anderem von den Experten der Fokusgruppen als zu klein beschrieben. Die Diskussionen mit den Akteuren zeigen, dass vor allem die Hoffnung besteht, in hinreichend großen Centern mit 500 und mehr Plätzen zu prüfen. In der Praxis sind hier jedoch deutliche Grenzen gesetzt, da durch rechtliche Vorgaben in den Versammlungsstättenverordnungen der Länder E-Assessment-Center mit mehr als 200 Plätzen als unrentabel erachtet werden. Erste Anläufe zur Erprobung eines mobilen E-Assessment-Centers an der genannten Hochschule konnten zum aktuellen Zeitpunkt noch nicht in eine praktische Nutzung überführt werden. Eine im hier vorgestellten Projekt erstellte Expertise zum Vergleich unterschiedlicher E-Assessment-Center-Typen rät derzeit aufgrund verschiedener Probleme (Rüstzeiten, Stabilität und Sicherheit des WLANs, ...) vom Betrieb eines mobilen Centers ab (Schulz 2017, S. 25).

Interessanterweise konnten in der Erhebung der Nutzungshäufigkeiten elektronischer Prüfungen keine Unterschiede in der Verbreitung von E-Klausuren an der Hochschule mit E-Assessment-Center im Vergleich zu den anderen Hochschulen des Bundeslands festgestellt werden. Alle Lehrenden, die elektronisch prüfen ( $N = 22$ ), nutzen hierfür PC-Pools. Einige nutzen zusätzlich auch das E-Assessment-Center ( $N = 4$ ) oder einen Hörsaal/Seminarraum mit Laptops ( $N = 1$ ). In den meisten Fällen reicht hierfür ein Raum aus, um alle Studierenden gleichzeitig zu prüfen ( $N = 13$ ). Das Prüfen findet aber auch zeitgleich in mehreren Räumen ( $N = 8$ ), zeitversetzt im gleichen Raum ( $N = 8$ ) und zeitversetzt in unterschiedlichen Räumen ( $N = 9$ ) statt. Hierfür werden die Fragen für die Klausuren randomisiert, also die Reihenfolge und Auswahl der Klausurfragen zufällig ausgewählt ( $N = 17$ ), woraus sowohl äquivalente ( $N = 14$ ) als auch identische ( $N = 8$ ) Klausurversionen für jeden Durchgang erstellt werden (Mehrfachantworten möglich).

## 5.2 Limitationen des E-Assessment-Einsatzes

Lehrende, die noch keine elektronischen Prüfungen einsetzen, äußern allgemeine Vorbehalte, die sie vom Einsatz von E-Klausuren abhalten. Am häufigsten zweifeln sie an der Übertragbarkeit ihrer Aufgaben in automatisiert auswertbare Aufgabenformate (51,4 Prozent) und an den Vorteilen des Einsatzes elektronischer Prüfungen (48,1 Prozent).<sup>5</sup> Dabei zeigen sich in der Online-Befragung auch fachspezifische Unterschiede. Vor allem die befragten Lehrenden aus dem Bereich Mathematik/Naturwissenschaften sehen nur begrenzte Möglichkeiten zur Übertragung ihrer Aufgabenstellungen in ein automatisiert auswertbares Verfahren (76 Prozent). In den anderen Fachbereichen stimmten dieser Aussage knapp die Hälfte der Befragten zu.<sup>6</sup> Gerade die Lehrenden der Mathematik/Naturwissenschaften sowie Ingenieurwissenschaften sehen signifikant häufiger keine Vorteile im Einsatz von elektronischen Prüfungen ( $\chi^2 = 9.77$ ,  $df = 1$ ,  $N = 151$ ,  $p = .021$ ). Interessant ist dieser Befund vor allem, weil er in einem scheinbaren Widerspruch zur stärkeren Verbreitung elektronischer Tests in diesen Wissenschaftsdisziplinen steht (siehe Kapitel 3). Auch von den sieben Experten der Fokusgruppen kamen fünf aus diesen Wissenschaftsdisziplinen. Die diesbezüglich aufgestellte Vermutung, dass die Bewertung der Vorteile bzw. der Übertragbarkeit von Aufgaben abhängig von bereits besuchten Weiterbildungen ist, konnte anhand des Datenmaterials nicht bestätigt werden. Auffällig ist diesbezüglich, dass diejenigen, die keine elektronischen Tests (für formative oder diagnostische Assessments) einsetzen, auch häufiger keine Vorteile in elektronischen Prüfungen sehen ( $\chi^2 = 12.34$ ,  $df = 1$ ,  $N = 183$ ,  $p = .000$ ).

Teilweise geben die Experten in den Fokusgruppeninterviews diesen Vorbehalten Recht. So verweisen auch sie darauf, „(...) dass Sie nicht alle Kompetenzen, nicht alle Fragen die Sie prüfen wollen, wirklich automatisch an der Stelle bewerten können.“ (Fokusgruppe 2, Z. 221f.) und „diese programmierte Logik der Bewertung, weil die eben sehr trivial ist, die reicht eben nicht aus, um hohe Kompetenzstufen-Fragen der Aufgaben fair zu bewerten (...).“ (Fokusgruppe 2, Z. 154f.). Die Experten ergänzen jedoch, dass daher eine Kombination von automatisch auswertbaren Aufgaben mit Freitext-Aufgaben angeraten sei („und es ist auch nicht nötig weil ich am Ende, wenn ich 200 Studenten habe, und ich habe 80 Prozent der Aufgaben die automatisch ausgewertet werden, das ist 80 Prozent der Zeitersparnis bei meiner Auswertung (...) [und] schon allein die Tatsache, dass das Ganze mir digital in Textform vorliegt, macht

<sup>5</sup>Dabei kann vermutet werden, dass E-Klausuren von den Lehrenden häufig fälschlicherweise mit MC-Klausuren gleichgesetzt werden. Die Erhebung zeigt jedoch, dass sowohl Freitext- als auch automatisiert auswertbare Aufgaben eingesetzt werden.

<sup>6</sup>Geisteswissenschaften 45,5 Prozent, Rechts-/Sozial-/Wirtschaftswissenschaften 50 Prozent, Ingenieurwissenschaften 50,8 Prozent.

*das Ganze wesentlich einfacher, erstmal zu lesen und (...) weil ich natürlich nicht mehr diesen Papierstapel mit hin nehmen kann, ich kann das auch abends um zehn im Wohnzimmer noch machen“ (Fokusgruppe 2, Z. 251ff.).*

Außerdem verweisen die Experten darauf, dass sich die erhoffte Zeitersparnis aufgrund des hohen Initialaufwandes bei der Erstellung von elektronischen Tests relativiere. Daher dauere es eine Zeit, bis sich der Zeitaufwand amortisiere. Darüber hinaus bestehen Zweifel, ob es beim elektronischen im Vergleich zum papierbasierten Prüfen zu modalitätsspezifischen Verzerrungen kommen kann (Test-mode-effect).

## **6 Handlungsempfehlungen zur Stärkung des E-Assessments in Sachsen**

Ein wichtiger Entwicklungsschritt wird von den befragten Lehrenden sowohl in der Online-Befragung als auch in den Fokusgruppen zunächst darin gesehen, die Rahmenbedingungen für den E-Assessment-Einsatz (siehe Kapitel 5.1) zu verbessern.

Dies betrifft die nötige Verankerung des elektronischen Prüfens in den Prüfungsordnungen. Die Lösung wird dabei nicht wie in anderen Bundesländern im Erlassen einer Rahmenprüfungsordnung für die gesamte Hochschule gesehen, da das sächsische Hochschulfreiheitsgesetz eine solche Regelung nicht vorsieht. Vielmehr werden dort Prüfungsangelegenheiten als Aufgabe der Fakultäten angesehen (siehe §13 Abs. 4 SächsHSFG). Die Verankerung muss demnach an allen Fakultäten individuell erfolgen. Eine entsprechende Musterprüfungsordnung an den Hochschulen kann hier Unterstützung bieten.

Auch die Einrichtung von Prüfungsräumen wird als notwendig erachtet. Hierbei sind verschiedene Umsetzungsvarianten denkbar (vgl. Schulz 2017). Deutschlandweit entstehen immer mehr generische E-Assessment-Center, die nur für das Prüfen genutzt werden. Darüber hinaus existieren Lösungen zur Nutzung von PC-Pools für das Prüfen. Weitere Lösungen wie die Einrichtung eines temporären (saisonalen) Centers sowie das mobile Prüfen mit Laptops befinden sich noch im Erprobungsstatus. Die Existenz eines E-Assessment-Centers mit entsprechenden personellen Ressourcen zur Betreuung und Unterstützung der Prüfungsorganisation schafft ein Sicherheitsgefühl für die Lehrenden (vgl. Schulze-Achatz & Riedel 2016).

Darüber hinaus können zur Reduktion des initialen Aufwands Unterstützungsangebote bereitgestellt werden. Diese können Beratungs- und Weiterbildungsangebote ebenso umfassen wie spezifische Dienstleistungen. Darunter fällt beispielsweise die Migration

und Erstellung von Prüfungsaufgaben in der Prüfungssoftware. In vielen Fällen entspricht diese aufgrund höherer technischer und funktionaler Anforderungen nicht den für das formative Assessment verwendeten Testwerkzeugen. Schnittstellen hierfür wären wünschenswert.

Auch das in Sachsen an den meisten Hochschulen verbreitete Testwerkzeug ONYX entspricht noch nicht den Anforderungen unter anderem zur Bewertung von automatisiert auswertbaren Aufgaben, wie sie z. B. in Multiple-Choice-Ordnungen vorgegeben werden. Gerade die Akteure, die bereits große Aufgabenpools mit diesem System erstellt haben, fordern die Möglichkeit zur Weiternutzung ihrer Aufgaben. Weitere Entwicklungen der Software sind daher erforderlich und werden regelmäßig durch die genannten Akteure beauftragt.

Ähnliche Empfehlungen formuliert auch die Themengruppe „Innovationen in Lern- und Prüfungsszenarien“ des Hochschulforums Digitalisierung (HFD). Zur Entwicklung des E-Assessments an deutschen Hochschulen wird auf die strategische Einbettung des E-Assessments, die Schaffung von geeigneten Infrastrukturen und die Bereitstellung von Weiterbildungsangeboten verwiesen (vgl. Geschäftsstelle Hochschulforum Digitalisierung 2015, S. 8ff.). Besonders hervorzuheben ist hierbei auch die Empfehlung zur Vernetzung der Akteure, angefangen beim regelmäßigen Erfahrungsaustausch und der Sichtbarkeit von Best Practices bis hin zur Erstellung gemeinsamer Aufgabenpools (vgl. ebd., S. 9). Vielfältige Projekte in Sachsen (Facharbeitskreise, hochschulübergreifende Verbundprojekte, Informationskampagnen, u. a.) unterstützen diese Anliegen temporär, konnten jedoch noch nicht verstetigt werden. Auch wenn in Sachsen bereits Bemühungen um eine Weiterentwicklung des E-Assessments und vor allem der dafür notwendigen Rahmenbedingungen sichtbar werden, ist der Weg für die flächendeckende Verankerung des E-Assessments in der Hochschullehre noch weit.

Trotz des Fokus auf die Rahmenbedingungen bleibt zu bedenken, dass allein durch die Bereitstellung von Infrastrukturen und rechtlicher Voraussetzungen die digitalisierte Lehre noch nicht zur Selbstverständlichkeit wird (vgl. Riedel; Dubrau; Köhler; Halgasch; Meinhold; Hamann; Heise; Schneider; Sieler; Kawalek; Pengel; Wollersheim; Tittmann & Schumann 2016). Hierfür ist darüber hinaus auch ein Wandel der Lehr-, Lern- sowie Prüfungskultur notwendig.

## Literatur

Baumert, Britta & May Dominik (2013): Constructive Alignment als didaktisches Konzept. In: *journal hochschuldidaktik* 1-2/2013. Abgerufen am 12.04.2018 von [http://www.zhb.tu-dortmund.de/hd/journal-hd/2013\\_1-2/journal\\_hd\\_2013\\_artikel\\_baumert\\_may.pdf](http://www.zhb.tu-dortmund.de/hd/journal-hd/2013_1-2/journal_hd_2013_artikel_baumert_may.pdf)

Baumgartner, Peter (2012): Eine Taxonomie für E-Portfolios. Teil II des BMWF Abschlussberichts E-Portfolio an Hochschulen: GZ 51.700/0064-VII/10/2006. Forschungsbericht. Unter Mitarbeit von K. Himpsl und S. Kleindienst. Department für Interaktive Medien und Bildungstechnologien, Donau Universität Krems

Bloh, E. (2006): Methodische Formen des E-/Online-Assessments. Unveröffentlichtes Manuskript, Kaiserslautern

Biggs, John (2003): Aligning teaching for constructing learning. Abgerufen am 11.04.2018 von: [https://www.heacademy.ac.uk/sites/default/files/resources/id477\\_aligning\\_teaching\\_for\\_constructing\\_learning.pdf](https://www.heacademy.ac.uk/sites/default/files/resources/id477_aligning_teaching_for_constructing_learning.pdf)

Biggs, John & Tang, Catherine (2011): Teaching for quality learning at university. What the student does. Maidenhead: McGraw-Hill and Open University Press

Bisovsky, Gerhard & Schaffert, Sandra (2009): Learning and Teaching With E-Portfolios: Experiences in and Challenges for Adult Education. In: *International Journal of Emerging Technologies in Learning (iJET)*, 4 (1), 13–15. Abgerufen am 11.04.2018 von <http://online-journals.org/i-jet/article/view/822>

Boud, David (1995): Enhancing Learning Through Self Assessment, Kogan Page, London

Cilliers, Francois. J., Schuwirth, Lambert W., Adendorff, Hanelie J., Herman, Nicoline & van der Vleuten, Cees P. (2010): The mechanism of impact of summative assessment on medical students' learning. *Advances in health sciences education*, 15(5), 695–715. Abgerufen am 29.03.2018 von [http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2995206/pdf/10459\\_2010\\_Article\\_9232.pdf](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2995206/pdf/10459_2010_Article_9232.pdf)

Geschäftsstelle Hochschulforum Digitalisierung (2015): E-Assessment als Herausforderung für Hochschulen. Handlungsempfehlungen. Abgerufen am 30.10.2017 von [http://www.che.de/downloads/HFD\\_E\\_Assessment\\_als\\_Herausforderung\\_Handlungsempfehlungen\\_fuer\\_Hochschulen.pdf](http://www.che.de/downloads/HFD_E_Assessment_als_Herausforderung_Handlungsempfehlungen_fuer_Hochschulen.pdf)

Gruttmann, Susanne (2010): Formatives E-Assessment in der Hochschullehre – Computerunterstützte Lernfortschrittskontrollen im Informatikstudium. Münster: MV-Verlag

Haladyna, Thomas M. (2004): Developing and Validating Multiple-Choice Test Items. 3. Auflage, Mahwah, N.J.: Lawrence Erlbaum Associates

Hattie, John (2013): Lernen sichtbar machen. Baltmannsweiler: Schneider

Jurecka, Astrid & Hartig, Johannes (2007): Computer- und netzwerkasiertes Assessment. In: Hartig, Johannes & Klieme, Eckhard (Hrsg.): Möglichkeiten und Voraus-

setzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung. Bonn, Berlin: BMBF, S. 37–48. Abgerufen am 11.04.2018 von [https://www.bmbf.de/pub/Bildungsforschung\\_Band\\_20.pdf](https://www.bmbf.de/pub/Bildungsforschung_Band_20.pdf)

Lindner, Marlit A., Strobel, Benjamin & Köller, Olaf (2015): Multiple-Choice-Prüfungen an Hochschulen? Ein Literaturüberblick und Plädoyer für mehr praxisorientierte Forschung. In: Zeitschrift für Pädagogische Psychologie 29 (3–4), S. 133–149. Abgerufen am 11.04.2018 unter <https://econtent.hogrefe.com/doi/pdf/10.1024/1010-0652/a000156>

Michel, Lutz P. & Goetz, Lutz (2015): Digitales Prüfen und Bewerten im Hochschulbereich. Gütersloh: Geschäftsstelle Hochschulforum Digitalisierung. Abgerufen am 30.10.2017 von [http://www.che.de/downloads/HFD\\_Studie\\_DigitalesPruefen.pdf](http://www.che.de/downloads/HFD_Studie_DigitalesPruefen.pdf)

Reeves, Thomas C. (2006): How do we know they are learning? The importance of alignment in higher education. *International Journal of Learning Technology*, 2 (4), 294–309. Verfügbar unter: Abgerufen am 29.03.2018 von <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.880.4817&rep=rep1&type=pdf>

Reinmann, Gabi (2007): Bologna in Zeiten des Web 2.0. Assessment als Gestaltungsfaktor (Arbeitsbericht Nr. 16). Augsburg: Universität Augsburg, Medienpädagogik. Abgerufen am 11.04.2018 von <https://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/deliver/index/docId/643/file/Arbeitsbericht16.pdf>

Riedel, Jana; Dubrau, Marlen.; Köhler, Thomas; Halgasch, Jana; Meinhold, Michael; Hamann, Marco; Heise, Linda; Schneider, André; Sieler, Oliver; Kawalek, Jürgen; Pengel, Norbert; Wollersheim, Heinz-Werner; Tittmann, Claudia & Schumann, Christian-Andreas (2016): Auf dem Weg zu einer neuen Lernkultur? Von Standorten und Stolpersteinen. In: Kawalek, Jürgen; Hering, Klaus, Schuster, Enrico: Tagungsband 14. Workshop on e-Learning (WeL ,16), Görlitz: Hochschule Zittau/Görlitz, S. 89–104

Roediger, Henry L. & Karpicke, Jeffrey D. (2006): Test-Enhanced Learning. Taking Memory Tests Improves Long-Term Retention. In: *Psychological Science*, Vol. 17, Nr. 3, S. 249–255

Schulz, Alexander (2017): E-Assessment-Center im Vergleich. Voraussetzungen und Kosten für die Einrichtung verschiedener E-Assessment-Center im Vergleich. Verfügbar unter: <http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-224532> (30.10.2017)

Schulz, Alexander & Apostolopoulos, Nicolas (2011): eExaminations Put To Test. Potenziale computergestützter Prüfungen. *Hamburger eLMagazin*, 7, S. 38–40. Verfügbar unter: <http://www.uni-hamburg.de/elearning/hamburger-elearning-magazin-07.pdf> (30.10.2017)

Schulze-Achatz, Sylvia & Riedel, Jana (2016): E-Assessment an Hochschulen gemeinsam stärken: Erfahrungen und Erfordernisse. In: *HDS Journal* 1/2016, S. 15–18. Verfügbar unter: [http://ul.qucosa.de/fileadmin/data/qucosa/documents/20260/hds\\_journal\\_1\\_2016-2.pdf](http://ul.qucosa.de/fileadmin/data/qucosa/documents/20260/hds_journal_1_2016-2.pdf) (30.10.2017)

Schmid, Ulrich, Goertz, Lutz, Radomski, Sabine, Thom, Sabrina, Behrens, Julia & Bertelsmann Stiftung (2017): Monitor Digitale Bildung. Bertelsmann Stiftung. Verfügbar unter: <https://doi.org/10.11586/2017014> (30.10.2017)

Seufert, Sabine & Brahm, Taiga (2007): „Ne(x)t Generation Learning“: E-Assessment und E-Portfolio: halten sie, was sie versprechen? Themenreihe II zur Workshop-Serie SCIL-Arbeitsbericht 13. Verfügbar unter: <https://www.alexandria.unisg.ch/45249/1/2007-03-brahm-seufert-next-generation-learning.pdf> (30.05.2017)

Wass, Val, van der Vleuten, Cees, Shatzer, John & Jones, Roger (2001): Assessment of clinical competence. In: The Lancet. Vol 357, S. 945–949. Abgerufen am 11.04.2018 von <http://acmd615.pbworks.com/f/Wass.pdf>

Manuskript eingegangen: 31.10.2017  
Manuskript angenommen: 15.06.2018

### **Anschrift der Autorinnen:**

Jana Riedel, M.A.  
Technische Universität Dresden  
Medienzentrum (MZ)  
Strehleener Straße 22/24  
01069 Dresden  
E-Mail: [jana.riedel@tu-dresden.de](mailto:jana.riedel@tu-dresden.de)

Kathrin Möbius, M.A.  
Technische Universität Dresden  
Medienzentrum (MZ)  
Strehleener Straße 22/24  
01069 Dresden  
E-Mail: [Kathrin.Moebius@tu-dresden.de](mailto:Kathrin.Moebius@tu-dresden.de)

Jana Riedel, M.A. ist stellvertretende Leiterin der Abteilung Digitales Lehren und Lernen am Medienzentrum der Technischen Universität Dresden und war Koordinatorin des Projekts „E-Assessment in Sachsen“. Ihre Forschungsschwerpunkte liegen im Bereich des digital gestützten Lehrens und Lernens in der Hochschullehre und dem selbstgesteuerten Lernen.

Kathrin Möbius, M.A. ist wissenschaftliche Mitarbeiterin der Abteilung Digitales Lehren und Lernen am Medienzentrum der Technischen Universität Dresden und im Projekt „E-Assessment in Sachsen“. Ihre Schwerpunkte liegen in der Entwicklung von Untersuchungs- und Testverfahren sowie in der Unterstützung selbstgesteuerter Lernprozesse.



# Online-Self-Assessments an der Technischen Hochschule Nürnberg – Bilanzierung, Reflexion, Zukunftsvision

Diana Wolff-Grosser

---

Da Online Self-Assessments an der Technischen Hochschule Nürnberg seit neun Jahren in einer zunehmenden Anzahl von Studiengängen eingesetzt werden, verfügt diese über einschlägige Erfahrungen mit unterschiedlichen Arten der Testbetreuung, der Teststruktur und Bearbeitungsoptionen. Zur Sicherstellung der Abläufe und Qualität bewähren sich in der Praxis die Bereitstellung eines eigenen Portals zum Hosten der Online Self-Assessments, das Einbeziehen testpsychologischer Sachkenntnis sowie die stetige Überprüfung und Weiterentwicklung der Tests, um sich gesellschaftlichen und technischen Änderungen anzupassen. Aufgrund der aktuellen Gesetzeslage in Bayern ist jedoch das Potenzial der Online Self-Assessments hinsichtlich ihrer prognostischen Validität in Bezug auf die Vorhersage von Studienabbrüchen bei weitem noch nicht ausgeschöpft, weshalb einige Änderungen wünschenswert wären.

---

## 1 Ausgangslage

Das Thema Studienabbruch ist bereits seit einigen Jahren in den Fokus der Politik geraten und beschäftigt Hochschulen nach wie vor. Gerade in MINT-Fächern finden sich je nach Studiengang hohe Abbruchquoten zwischen 22 und 51 Prozent (Marsch, 2017; Heublein, Ebert, Hutzsch, Isleib, König, Richter, & Woisch, 2017). Studienabbrüche sind mit hohen Kosten für die Gesellschaft, aber auch das betreffende Individuum verbunden: Eine Studie des Stifterverbands für die Deutsche Wissenschaft (Studienabbruch: Staat setzt jährlich 2,2 Mrd. Euro in den Sand, 2007) errechnete 2007 2,2 Milliarden Euro als direkte und sogar 7,6 Milliarden indirekte Kosten. Die häufigsten Ursachen für Studienabbruch sind gemäß der aktuellen Studie des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW) (Heublein et al., 2017) unbewältigte Leistungsanforderungen, mangelnde Studienmotivation sowie der Wunsch nach mehr Praxisnähe. Aufgrund der hohen Kosten dieses Phänomens gibt es auf Seiten der Hochschulen zahlreiche Bemühungen, dem Abbruch entgegenzuwirken; eine der initiierten Maßnahmen stellen Online-Self-Assessments bzw. Studierfähigkeitstests (beide Begriffe werden hier synonym verwendet) dar, die sich an der Technischen Hochschule Nürnberg seit inzwischen neun Jahren gut etabliert haben. In dieser Zeit erfolgten einige Änderungen und Entwicklungen, die beispielsweise Anzahl, Teststruktur und technischen Hintergrund des Instruments betreffen. Somit sind auf Seiten der Institution einige Erfahrungen bezüglich Einführung, Hindernissen und dem Umgang damit sowie technische und

testpsychologische Instandhaltung vorhanden, die in diesem Artikel durch praktische Einblicke dargestellt werden sollen, um anderen Institutionen, die selbst die Einführung eines Self-Assessments erwägen, eine Entscheidungshilfe an die Hand zu geben. Darüber hinaus sollen die bisher festgestellten Effekte bilanziert und reflektiert sowie wünschenswerte Entwicklungen aufgezeigt werden.

## **2 Entwicklungsgeschichte der Online Self-Assessment-Einführung**

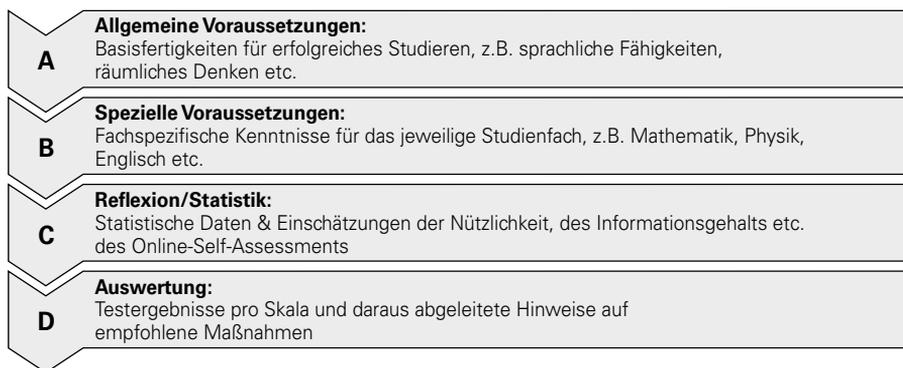
### **2.1 Technische und teststrukturelle Entwicklung**

Zum besseren Verständnis wird im Folgenden kurz die Entstehungsgeschichte der Online-Self-Assessments an der Technischen Hochschule Nürnberg umrissen. Im Rahmen des vom Bayerischen Staatsministerium für Wissenschaft und Kunst und von der Vereinigung der Bayerischen Wirtschaft e. V. geförderten Projekts „MINT – Wege zu mehr MINT-Absolventen“ wurden 2009 an der Technischen Hochschule Nürnberg sukzessive Online-Self-Assessments für sieben (inzwischen 14) Studiengänge eingeführt, basierend auf der Annahme, dass häufige Gründe für einen Studienabbruch wie mangelnde Motivation, größerer Wunsch nach Praxis, aber auch fehlende fachliche Voraussetzungen letztlich auf einer mangelhaften Auseinandersetzung mit dem Studienfach und daraus resultierender Uninformiertheit beruhen. Daraus leiteten sich folgende Ziele ab, die mittels der Einführung von Self-Assessments erreicht werden sollten: Die Hochschule suchte die Informiertheit Studieninteressierter über ein bestimmtes Studienfach sowie die Induktion einer sinnvollen Selbstselektion der Studieninteressierten durch Konfrontation mit fachlichen Stärken und Schwächen zu verbessern, so dass diese durch Aufzeigen des individuellen Fähigkeitsniveaus und Anstoßen von Selbstreflexion fundierte Entscheidungen hinsichtlich ihres zukünftigen Studiengangs treffen. Letztlich sollten die Online Self-Assessments als ein Mittel von mehreren zur Reduktion vorzeitiger Studienabbrüche beitragen.

Die Online Self-Assessments wurden gemäß der klassischen Testtheorie unter der fachlichen Leitung eines Testpsychologen konstruiert. Zur damaligen Zeit wiesen sie eine vierteilte Teststruktur auf: Teil A maß die allgemeine Befähigung zu einem technischen Studiengang. Neben Aufgaben zu Interessenschwerpunkten enthielt dieser Teil auch Fragen zur Studienmotivation und kognitive (Leistungs-)Aufgaben zu rechnerischem, räumlichem oder sprachlichem Denken. Teil B enthielt je nach Studiengang unterschiedlich viele Skalen mit fachspezifischen Aufgaben (überwiegend im Multiple-Choice-Format), die die Fähigkeiten in für den Studiengang relevanten Bereichen wie z. B. Mathematik, Physik, Chemie etc. aufzeigten. Zur Auflockerung enthielten die Tests studiengangspezifische Informationsseiten mit Hinweisen zum Studiengang und Erklärungen der Praxisrelevanz der fachlichen Aufgaben. Teil C schließlich umfasste Fragen nach persönlichen und demographischen Angaben, nach der Akzeptanz des Tests sowie nach Anmerkungen und Kommentaren. In Teil D

erhielten die Teilnehmer eine individuelle Rückmeldung ihrer Ergebnisse, eine Interpretationshilfe sowie Verlinkungen zu möglichen Unterstützungsmaßnahmen wie Brückenkursen etc. (siehe Abb. 1). Eine konkrete Empfehlung, den gewünschten Studiengang zu studieren oder nicht, erfolgte bewusst nicht und wird (aufgrund nicht abgeschlossener und theoretisch problematischer Validitätsnachweise, siehe Kubinger, 2015) auch heute nicht ausgesprochen.

**Abbildung 1:** Teststruktur der „alten“ Online-Self-Assessments



Die Testaufgaben wurden von Professorinnen und Professoren sowie Dozierenden der jeweiligen Fachgebiete zur Verfügung gestellt und entsprachen vom Schwierigkeitsniveau her den Anforderungen, die an die jeweiligen Studienanfängerinnen und -anfänger gestellt werden. Alle Testaufgaben wurden im Vorfeld in Pretests an Stichproben der späteren Zielgruppe geprüft, um die wichtigen Testgütekriterien (Item-Schwierigkeit, Trennschärfe, Varianz; auf Skalenebene Reliabilität) zu ermitteln und somit die Qualität der Testaufgaben sicherzustellen. Tabelle 1 zeigt beispielhaft die Kennwerte einiger erster Skalen vor und nach der (ersten) Aufgabenrevision.

**Tabelle 1:** Kennwerte verschiedener Skalen vor und nach der ersten Aufgabenrevision

Module	Statistische Kennwerte der Aufgabenanalyse		
	Schwierigkeitsindex von bis	Trennschärfeindex von bis	Reliabilität Cronbach's Alpha
Mathematik	0.09–0.98	–0.16–0.46	0.53
Neu	0.13–0.93	0.14–0.50	0.67
Physik	0.04–0.77	–0.08–0.43	0.66
Neu	0.38–0.77	0.16–0.43	0.67
Chemie	0.34–1.00	0.16–0.45	0.71
Neu	0.34–0.88	0.19–0.44	0.72
Englisch	0.57–0.96	0.07–0.49	0.51
Neu	0.69–0.84	0.22–0.35	0.46
Empfohlener Bereich	0.05–0.95	0.20–0.80	0.60–1.00

Um das Angebot möglichst niedrigschwellig zu halten und somit die Teilnahme zu erleichtern, mussten sich zur damaligen Zeit die Teilnehmerinnen und Teilnehmer nicht registrieren; lediglich bei den Studiengängen, die die Bearbeitung (und nur diese!) des Tests als Voraussetzung für die Immatrikulation verlangten, musste nach der Bearbeitung eine E-Mail-Adresse angegeben werden. Zwar war dadurch die Anonymität der Teilnehmenden gewährleistet, allerdings zog die damalige Teststruktur einige Nachteile mit sich, wie etwa eine Bearbeitungsdauer von 90 Minuten bei fehlender Möglichkeit der Pausierung des Tests, was von vielen Teilnehmern als zu lang empfunden wurde.

Im Jahr 2012 wurde die technische Umsetzung der Online Self-Assessments von der Fakultät Informatik übernommen und auf einen eigenen Server sowie ein anderes Onlineportal übertragen, wodurch die technische Zuverlässigkeit erheblich gesteigert wurde; weiterhin gab es fortan feste Ansprechpartner. 2015 erfolgte eine grundlegende Änderung der Teststruktur zur Korrektur der auffälligsten Schwachstellen: mit der Schaffung eines eigenen Portals, in dem sich Teilnehmenden vor der Bearbeitung registrieren müssen, konnte die mehrfache Bearbeitung von einzelnen Testteilen vermieden werden, was vorher aufgrund des bei allen Tests identischen Teils A ein Problem darstellte und zu Ermüdungserscheinungen und Motivationsdefiziten führte. Dieser Umstand spiegelte sich in den hohen Abbruchquoten der Tests wider, die zuvor dreimal so hoch waren wie die Quoten abgeschlossener Tests, was sich nach der Umstellung quasi umdrehte (siehe Tabelle 2).

**Tabelle 2:** Abgebrochene und abgeschlossene Tests vor und nach Portalumstellung

Nutzungsjahr	2013/2014/2015*	2015**/2016/2017
Abgebrochene Tests	24402	3747
Abgeschlossene Tests (12/14 Studiengänge)	8472	9548

\* bis Juni 2015; \*\*ab Juli 2015

Weiterhin ermöglicht die Schaffung des Portals durch die Wiedererkennung der teilnehmenden Person eine Testbearbeitung zu verschiedenen Zeitpunkten. Teil A der „alten“ Test-Struktur – allgemeine Studierfähigkeiten – wurde von den fachspezifischen Skalen getrennt und um Skalen zu lernorganisatorischen Fähigkeiten, die selbst (ebenefalls gemäß klassischer Testtheorie) konstruiert und in einem Pretest mit verschiedenen Schulen überprüft wurden, ergänzt (Details siehe Tab. 3), und stellt nun einen eigenständigen allgemeinen Studierfähigkeitstest dar. Aktuell hat die Technische Hochschule Nürnberg für 14 Studiengänge fachspezifische Studierfähigkeitstests, deren Aufbau abgesehen von der Abspaltung von Teil A nach wie vor dem in Abbildung 1 entspricht, sowie den allgemeinen Studierfähigkeitstest. Alle Tests sind sowohl auf klassischen Computern wie auch mobilen Endgeräten abruf- und bearbeitbar.

**Tabelle 3:** Reliabilität der vier neu entwickelten Skalen des Allgemeinen Studierfähigkeitstests

Skala	N	Reliabilität $r_{tt}$ (Cronbach's Alpha)
Selbstreguliertes Lernen	3664	.797
Selbstwirksamkeit	3669	.836
Volition	3616	.813
Erwartungen an ein Studium	3585	.744

## 2.2 Rahmenbedingungen der Testbearbeitung

Da es sich auch bei den verpflichtenden Tests nicht um verbindliche Eignungstests, sondern um Studierfähigkeitstests handelt, die primär der eigenen Information dienen, ist die Testdurchführung nicht auf einen bestimmten Zeitraum und Ort beschränkt (wie etwa bei einer Prüfung), sondern kann prinzipiell wann immer privat erfolgen, inzwischen auch von mobilen Endgeräten aus. Lediglich der (in etwa zweimonatige) Anmeldezeitraum stellt eine Art limitierenden Faktor für die Bearbeitung eines verpflichtenden Tests dar, da bei Studiengängen mit verpflichtenden Online Self-Assessments die Testteilnahme bei der Immatrikulation vom System geprüft wird. Ansonsten ist die Bearbeitung der Tests jedoch völlig frei, wenn auch eine ungestörte Bearbeitung am Computer zu Hause mit einem Blatt Papier als Hilfsmittel aufgrund mancher Aufgabenarten sinnvoll ist. Die Vorteile einer völlig freien Bearbeitungsmöglichkeit liegen in der zeitlichen und örtlichen Flexibilität, die aufgrund der Freiheit die Motivation erhöhen und mögliche Reaktanzeffekte reduzieren. Nachteile bestehen jedoch darin, dass die Teilnahme prinzipiell unkontrolliert erfolgt: Weder kann von Seiten der Hochschule gesichert werden, dass der Test von der sich einschreibenden Person durchlaufen wird, noch dass der Test ohne unerlaubte Hilfsmittel wie Internet etc. durchgeführt wird, was die prognostische Validität reduzieren kann. Da jedoch als primäres Ziel eine bessere Informiertheit der Teilnehmenden durch Konfrontation mit im Studiengang gängigen Aufgabenstellungen postuliert wird, wird diese Unsicherheit toleriert. Der einzig (sinnvolle) Weg, die Ungewissheit der persönlichen Teilnahme auszuräumen, bestünde in einer verbindlichen Teilnahme an den Tests an einem Ort an der Hochschule, wofür jedoch aktuell nicht die nötigen Rahmenbedingungen und Gesetzesgrundlagen (s. u.) gegeben sind.

## 2.3 Welche testpsychologischen Aspekte spielen eine Rolle?

Die Online Self-Assessments wurden von Beginn an durch Testpsychologen geprüft und begleitet, was vom Projektträger zweckmäßigerweise als obligate Voraussetzung für die Implementierung von Tests angesehen wurde. Die testpsychologische Begleitung umfasst Beratung bei der Konzeption und Formulierung von Testaufgaben nach test-

psychologischen Maßstäben, die Berechnung der in der psychologischen Diagnostik gängigen Testgütekriterien (Trennschärfe, Schwierigkeit, Reliabilität), die Entscheidung für eine bestimmte Art der Ergebnisrückmeldung inklusive Normierung sowie die Evaluation und Maßnahmengenerierung anhand der erhaltenen Testergebnisse.

## **2.4 Aktuelle Entwicklungen**

Aktuell erfolgt die Förderung der mit Studierfähigkeitstests zusammenhängenden Projekte erneut durch das Bayerische Staatsministerium für Wissenschaft und Kunst im Rahmen des Projekts „**MINTer Aktiv**“. Das Online Self-Assessment-Projektteam der Technischen Hochschule Nürnberg ist gegenwärtig bestrebt, den Fokus von der Vorstudiumsphase auf die Studieneingangsphase auszuweiten und neben der Diagnostik, die durch die Tests recht zuverlässig stattfindet, auch das Thema Intervention stärker zu berücksichtigen, wozu einzelne Skalen des allgemeinen Studierfähigkeitstests herangezogen werden. Auch sollen künftig die Studierfähigkeitstests nicht nur defizit-, sondern auch ressourcenorientiert eingesetzt werden, z. B. als Mittel zur Identifizierung sehr begabter Teilnehmer, was eine Erweiterung der bisherigen Teststruktur erfordert.

## **3 Bilanzierung, Reflexion, Zukunftsvision**

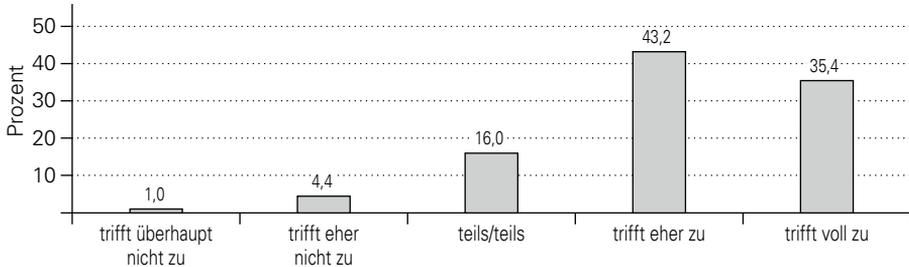
Welche Bilanz zu Studierfähigkeitstests kann die THN nun nach guten neun Jahren ziehen? Auf diese Frage wird zunächst anhand der im Erstprojekt anvisierten Ziele eingegangen, anschließend werden weitere wichtige Erkenntnisse aus der Test-Nutzung dargestellt.

### **3.1 Informiertheit**

Das primäre Ziel bei der Einführung der Tests war, sowohl durch Präsentation der im Studiengang relevanten fachlichen Aufgaben als auch durch im Online Self-Assessment enthaltene Informationsseiten zum betreffenden Studium eine bessere Informiertheit der Studieninteressierten zu erreichen, da Studienabbrüche in den ersten Semestern auf eine Diskrepanz zwischen persönlichen Erwartungen und den vorgefundenen Begebenheiten hinweisen (Fellenberg & Hannover, 2006). Zur Erfassung dieses Ziels werden die Teilnehmenden eines Online Self-Assessments direkt im Anschluss an die fachlichen Teile zur Einschätzung des Grads ihrer Informiertheit über das Studium befragt, was folgende Ergebnisse erbrachte:

**Abbildung 2:** Prozentuale Antworten aller Teilnehmer auf Reflexionsfrage Nr. 6  
(N = 7058)

Ich bin jetzt über das Studium gut informiert und weiß, was in den ersten Semestern auf mich zukommt.

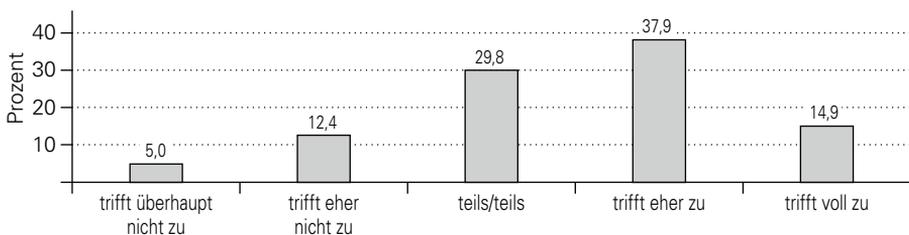


Wie anhand Abbildung 2 ersichtlich, fühlen sich über drei Viertel der Teilnehmenden nach dem Test gut über den Studiengang informiert. Um zu überprüfen, ob sich durch die Testbearbeitung eine *Verbesserung* der Informiertheit ergeben hat, wäre zukünftig die Erfassung des Vorher-Zustands dieses Konstrukts in Form einer Selbsteinschätzung sinnvoll oder ggf. eine anderweitige Überprüfung des neugewonnen Wissens über den Studiengang.

In Abbildung 3 zeigt sich weiterhin, dass die Teilnahme an einem Online Self-Assessment auch Effekte auf das konkrete Verhalten (zumindest im Selbstbericht) ausüben kann:

**Abbildung 3:** Prozentuale Antworten aller Teilnehmer auf Reflexionsfrage Nr. 5  
(N = 7062)

Ich werde das Testergebnis bei meiner Studien- und Berufswahl sicher berücksichtigen.



Interessant ist auch folgendes Phänomen: Sowohl vor als auch nach der Bearbeitung der fachspezifischen Fragen werden die Teilnehmenden gebeten, ihre Passung zum gewählten Studiengang auf einer Rating-Skala einzuschätzen. Um festzustellen, ob sich an der individuellen Einschätzung der Teilnehmer bezüglich ihrer Passung etwas nach Durchlaufen des Tests verändert hat, wurde ein Vorher-Nachher-Vergleich durchgeführt.

**Tabelle 4:** Vorher-Nachher-Vergleich der selbsteingeschätzten Passung zum Studiengang (N = 8472)

	Passung höher (%)	Passung niedriger (%)	Passung gleich (%)
MW (9 Studiengänge)	9.66	23.56	66.78
SD (9 Studiengänge)	4.27	6.08	6.51

MW = Mittelwert; SD = Standardabweichung

Aus Tabelle 4 sind folgende Tendenzen ersichtlich: Der Großteil der Teilnehmenden gibt vor und nach der Testbearbeitung die gleiche (subjektive) Passung zum Studiengang an. Ein kleinerer Teil empfindet nach der Testbearbeitung die individuelle Passung zum Studiengang besser. Im mittleren Bereich befinden sich Teilnehmerinnen und Teilnehmer, die ihre persönliche Passung zum Studiengang nach Durchlaufen des Tests als niedriger empfinden. Das lässt erkennen, dass die Bearbeitung des Tests durchaus Effekte auf die Teilnehmenden hat. Aus Sicht des Projektziels „Reduzierung des Studienabbruchs“ interessieren vor allem die Teilnehmenden, die ihre Passung niedriger beurteilen als vor der Bearbeitung. Bislang konnten sowohl aus datenschutzrechtlichen als auch technischen Gründen keine Aussagen darüber getroffen werden, ob aus der wahrgenommenen niedrigeren Passung auch auf der Verhaltensebene Konsequenzen gezogen werden – z.B. mit der Wahl eines anderen Studiengangs, dem Beseitigen von Wissenslücken oder der Beibehaltung des Studiengangs bei Wahl einer anderen Hochschule, was jedoch zukünftig größtenteils geschehen soll.

### 3.2 Senkung der Abbruchquoten

Können Studierfähigkeitstests zur Reduzierung von Studienabbrüchen beitragen? Diese Frage wird sowohl von Test-Kritikern als auch Befürwortern immer wieder gestellt, lässt sich aber leider nicht einfach beantworten. Die Bereitstellung und Qualitätssicherung von Tests ist durchaus mit einigen Personalkosten verbunden, so dass die Antwort auf diese Frage wichtig erscheint. Offenbar herrscht an einigen Hochschulen (Kubinger, 2015), teilweise auch an der eigenen Institution, die Vorstellung vor, dass die von Studierfähigkeitstests intendierten Selbstselektionsprozesse einerseits vollständig gelingen (sprich: „ungeeignete“ Bewerberinnen und Bewerber lassen sich perfekt identifizieren, handeln streng rational und sehen von ihrer Studienwahl ab) und andererseits auch, dass dies für eine Reduzierung von Studienabbrüchen ausreicht, nach dem Motto: „Wenn nur die richtigen Personen immatrikuliert sind, läuft der Rest von allein“. In der Praxis finden sich jedoch in den meisten Studiengängen einige „ungeeignete“, also wenig erfolgreich Studierende trotz Vorhandenseins von Online Self-Assessments, und es finden nach wie vor Studienabbrüche statt. In ähnlicher Weise lässt sich dieses Phänomen auch bei Bewerbungsprozessen beobachten, wo trotz Intention der Stellenanzeigen, durch entsprechende Formulierungen nur geeignete Bewerberinnen und Bewerber anzusprechen, sich eine bestimmte Anzahl ungeeigneter

bewirbt, aus den unterschiedlichsten Gründen. Dieser Effekt im Bereich Studium kann zahlreiche Gründe haben: So könnte es bedeuten, dass Studierfähigkeitstests die Funktion der Selbstselektion nicht oder nur gering erfüllen, z. B., weil fachlich (zunächst?) ungeeignete Teilnehmende es wider besseren Wissens „trotzdem probieren“ möchten (was auch legitim wäre, wenn sie entsprechende Maßnahmen zur Reduzierung von Defiziten ergreifen würden), oder das Ergebnis aufgrund nicht ernsthafter Bearbeitung nicht handlungsleitend ist. Es könnte aber auch sein, dass es zwar zu einer Selbstselektion kommt, den geeigneten Teilnehmenden jedoch zwischenzeitlich etwas geschieht, was sie trotz prinzipieller Eignung zum Abbruch veranlasst (z. B. familiäre, finanzielle Probleme, Demoralisierung während des Studiums etc.). Denkbar ist ferner eine Vielzahl weiterer Gründe, die sich aus ökonomischen Gründen schwer systematisch erforschen lassen. So lässt sich kaum feststellen, ob tatsächlich so genannte „ungeeignete“ Teilnehmende durch den Test vom Studium Abstand genommen haben, da diese ja im Studiengang nicht auftauchen und daher auch keine Rückschlüsse über deren Eignung bzw. Erfolg im Studium getroffen werden können. Generell muss man konstatieren, dass die Vorstellung, ein komplexes und multikausal bedingtes Problem wie Studienabbruch durch eine monokausale (?) Einmal-Intervention wie ein Online Self-Assessment nahezu vollständig zu beheben, unrealistisch anmutet, und von vornherein als Illusion klargestellt werden sollte.

Empirisch jedoch gut abgesichert ist die Prognosekraft der Ergebnisse von Studierfähigkeitstests kombiniert mit Schulabschlussnoten (Troost, 2003). Eine Kombination des Testergebnisses mit anderen gängigen Methoden der Zulassung, wie beispielsweise das Heranziehen von Abschlussnoten oder Noten einzelner Fächer, kann im Sinne der inkrementellen Validität die Prognosekraft der Schulabschlussnote erhöhen, wie an Auswahlverfahren in medizinischen Studiengängen wiederholt gezeigt wird (Gold & Souvignier, 2005). Eine Metaanalyse von Hell, Trapmann, & Schuler (2007) ergab, dass (fachspezifische) Studierfähigkeitstests „zu den validesten Einzelprädiktoren zur Vorhersage des Studienerfolgs“ (Hell et al., 2007, S. 263) gehören, so dass eine stärkere Etablierung dieses Instruments sehr sinnvoll erscheint. Allerdings muss die Durchführung eines solchen Tests dann zwingend unter kontrollierten Bedingungen stattfinden, um die persönliche Durchführung der betreffenden Person ohne unerlaubte Hilfsmittel zu gewährleisten. Auch an der Technischen Hochschule Nürnberg sind nach langjähriger Abklärung der Datenschutz-relevanten Angelegenheiten sowie entsprechender Programmierung zukünftig Korrelationsrechnungen zwischen Test-Ergebnissen und Studienerfolg-relevanten Kriterien (z. B. erzielte ECTS-Punkte, angetretene Prüfungen, Studierdauer etc.) vorgesehen, um Aussagen zur prognostischen Validität des Instruments treffen zu können. Sollte sich hierbei herausstellen, dass die eigenen Tests über eine ebenso gute prognostische Validität wie ähnliche Instrumente verfügen, kann das als solide Basis für weitergehende Maßnahmen dienen. Denkbar wäre dann in etwa eine bessere Nutzung der Testergebnisse bei der Konzeption von Hilfs-

angeboten oder auch bei der Studierendenauswahl. Dass eine reflektierte Bewerberauswahl durch die Hochschulen tatsächlich mit niedrigeren Studienabbruchquoten einhergeht, zeigt das Beispiel der Technischen Universität München in zahlreichen Studiengängen (Marsch, 2017). Sie hat hierzu in zahlreichen Studiengängen ein aufwändiges Verfahren etabliert, wonach Studierende, die aufgrund ihrer Abiturnote weder eindeutig zugelassen noch abgelehnt werden, in persönlichen Auswahlgesprächen auf ihre Eignung geprüft werden, was hinsichtlich der Drop-Out-Raten beträchtliche Erfolge zeigt. Die Methode hat den großen Vorteil, dass auf diese Art Bewerberinnen und Bewerber, die für das betreffende Studienfach eine hohe Motivation oder großes Interesse aufweisen, was sich aber möglicherweise in der Abiturnote nicht widerspiegelt, dennoch einen Studienplatz erhalten. Der Nachteil des Verfahrens liegt jedoch im hohen zeitlichen und personellen (und somit finanziellen) Aufwand: Nach Marsch (2017) führt die Technische Universität München allein zu den Wintersemestern 5000 bis 6000 Auswahlgespräche. Trotz dieses Aufwands und der bürokratischen Hürden, die für ein solches Verfahren überwunden werden müssen, hält die Universität aufgrund ihrer sehr positiven Erfahrungen daran fest. Somit wäre es durchaus auch für andere Hochschulen bedenkenswert, sich dieses Zulassungsverfahren bzw. das zugrundeliegende Prinzip einer zusätzlichen Passungsüberprüfung über die Hochschulzugangsberechtigungsnote hinaus als Vorbild zu nehmen und ggf. dabei empirisch zu erforschen, ob sich hinsichtlich Abbruchquoten auch mit weniger aufwändigen Methoden wie etwa Studierfähigkeitstests ähnliche Ergebnisse erzielen lassen. Hell et al. (2007) weisen darauf hin, dass fachspezifische Studierfähigkeitstests die Validität von Auswahlgesprächen übertreffen; somit könnten Studierfähigkeitstests durchaus eine gute Alternative zu Auswahlgesprächen im Sinne eines gerechteren und dennoch ökonomisch vertretbaren Verfahrens der Studienplatzvergabe darstellen.

Nach diesen Ausführungen bleibt die Frage bestehen: Können Online-Self-Assessments Studienabbruch reduzieren? Die Antwort aus Sicht des Projektteams dazu lautet: Ja, können sie, wenn die Testergebnisse sinnvoll genutzt werden, z.B. für frühzeitige Interventionen bei gefährdeten Studierenden, oder als Zusatzkriterium bei der Bewerberwahl (unter anderen rechtlichen Rahmenbedingungen als aktuell vorhanden). Die Etablierung von Tests im Sinne eines einmaligen „Hindernisses“ zur Abschreckung ungeeigneter Bewerberinnen und Bewerber, verbunden mit der Annahme, das Studium sei für die Verbleibenden ein Selbstläufer, muss jedoch als unrealistisch klar benannt werden – so können Studierfähigkeitstests nicht (nennenswert) zum Studienabbruch beitragen.

## 4 Lessons learned auf Seiten des Projektteams und daraus abgeleitete Empfehlungen

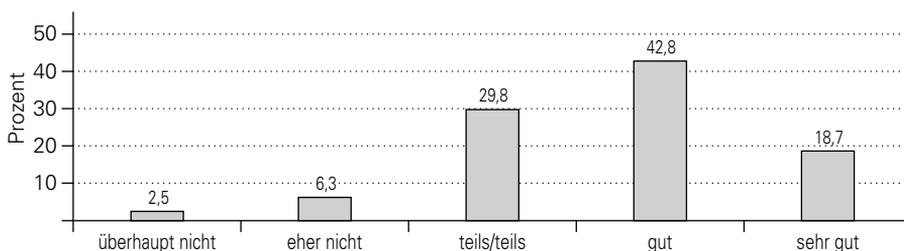
Im Folgenden soll auf Erfahrungen eingegangen werden, die das Projektteam jenseits der angestrebten Ziele gewinnen konnte, und entsprechende Empfehlungen für die Praxis daraus abgeleitet werden.

### 4.1 Akzeptanz der Online-Self-Assessments durch die Teilnehmer

Eine Befürchtung, die häufig im Hinblick auf Tests angeführt wird, ist, dass diese potenziell attraktive und geeignete Studieninteressierte abschrecken, insbesondere wenn andere Hochschulen keine Teilnahme an einem Test bei der Einschreibung verlangen. Bei Betrachtung der Akzeptanz der Tests fällt jedoch auf, dass nur sehr wenige Teilnehmende das Instrument an sich völlig ablehnen:

**Abbildung 4:** Prozentuale Antworten aller Teilnehmer auf Reflexionsfrage Nr. 1 (N = 7088)

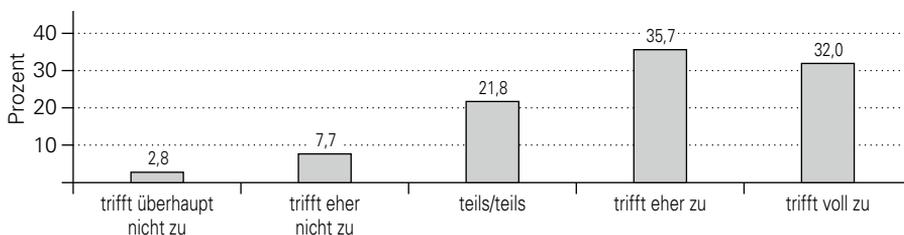
Wie hat Ihnen das Selbsteinstufungsverfahren insgesamt gefallen?



Lediglich neun Prozent der Teilnehmenden geben an, dass ihnen der Studierfähigkeitstest eher nicht oder überhaupt nicht gefallen hat. Ähnlich verhält es sich bei der Frage nach einer prinzipiellen Empfehlung der Teilnahme an einem Test, wie in Abbildung 5 ersichtlich:

**Abbildung 5:** Prozentuale Antworten aller Teilnehmer auf Reflexionsfrage Nr. 8 (N = 7058)

Insgesamt betrachtet würde ich jedem Interessenten an einem technischen Studium das Online-Self-Assessment zur Vorbereitung auf das Studium empfehlen.



## 4.2 Freiwilligkeit versus Verpflichtung der Teilnahme

Eine weitere häufige Kontroverse im Zusammenhang mit den Tests an der eigenen Institution ist das Thema der Freiwilligkeit der Teilnahme, die eng mit der Akzeptanz der Tests verknüpft ist. Bereits seit Einführung der Tests im Jahr 2009 legen manche Studiengänge die Teilnahme am fachspezifischen Online Self-Assessment als verpflichtend für die Immatrikulation fest. Da es sich dabei ausdrücklich nicht um Eignungs- und/oder Zulassungstests handelt, spielt das jeweils erzielte Ergebnis für die Zulassung zum Studium keine Rolle (theoretisch reicht es, wenn sich Studieninteressierte durch die Aufgaben „durchklicken“ und damit am Ende keine Punkte erhalten). Diejenigen Bewerberinnen und Bewerber, die den Test absolvieren, sollen über die Teilnahme nicht zuletzt im Sinne einer „realistic job preview“ einen realistischen Eindruck davon bekommen, welche (Leistungs-) Anforderungen auf sie zukommen werden.

Während von Verpflichtungs-Befürwortern der Aspekt der Chance auf eine bessere (wenngleich erzwungene) Informiertheit betont wird, verweisen Freiwilligkeitsverfechter auf mögliche Reaktanz- und Abschreckungseffekte durch eine verpflichtende Teilnahme. Aus psychologischer Sicht sind Reaktanzeffekte bei erzwungenen Teilnahmen durchaus plausibel, da Teilnehmende sich hierdurch in ihrer persönlichen Freiheit zur Teilnahme eingeschränkt fühlen können, weshalb dieser Aspekt nicht ohne weiteres ignoriert werden sollte. Über die Frage „Wie hat Ihnen das Selbsteinstufungsverfahren insgesamt gefallen“ (mit 1 = überhaupt nicht und 5 = sehr gut), wird versucht, einen Hinweis auf mögliche Reaktanzphänome im Zusammenhang mit freiwilliger vs. verpflichtender Teilnahme zu bekommen.

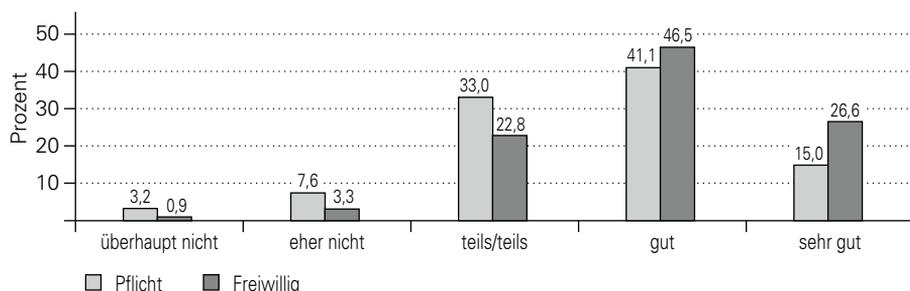
In diesem Zusammenhang wurde weiterhin analysiert, ob es einen signifikanten Mittelwertsunterschied in der Akzeptanz der Tests zwischen verpflichteten und freiwilligen Teilnehmenden gibt. In der Tat ist der Mittelwert bei dieser Frage (und somit der Akzeptanz-Wert) bei Teilnehmenden, die den Test freiwillig absolvieren, größer ( $M = 3,94$ ,  $SD = .840$ ) als bei verpflichteten Teilnehmenden ( $M = 3,57$ ,  $SD = .944$ ) und wird mit  $t(4854.95) = -16.74$ ,  $p = .001$  statistisch signifikant. Die Effektstärke weist mit  $d_{Cohen} = .41$  auf einen mittleren Effekt hin.

Abbildung 6 zeigt einen Vergleich der Einzelantworten zwischen freiwilligen und verpflichteten Teilnehmenden. Antworten der Kategorien „überhaupt bzw. eher nicht“ sind bei verpflichteten Teilnehmenden mit knapp elf Prozent häufiger vorhanden als bei freiwilligen Teilnehmenden (gut vier Prozent), dennoch machen diese beiden Kategorien bei beiden Gruppen einen insgesamt sehr niedrigen Anteil aus. Weiterhin erscheint das Ergebnis angesichts der Annahme, dass ein Test auf freiwilliger Basis ohnehin eher von Personen mit höherer Akzeptanz gegenüber psychologischen Tests absolviert wird, sehr plausibel, so dass bei der Interpretation des Ergebnisses als

Reaktanzphänomen zumindest Vorsicht, wenn nicht gar Zweifel angebracht sind, und es somit erst recht nicht sinnvoll erscheint, anhand der Ergebnisse auf Abschreckungseffekte zu schließen.

**Abbildung 6:** Akzeptanz der Tests nach Verpflichtung der Teilnahme aufgeschlüsselt  
( $N_{\text{Pflicht}} = 4847$ ;  $N_{\text{Freiwillig}} = 2241$ )

Wie hat Ihnen das Selbsteinstufungsverfahren insgesamt gefallen?



Um dennoch mögliche Reaktanzphänome zu minimieren, die im Zusammenhang mit dem verpflichtenden Absolvieren des „allgemeinen Studierfähigkeitstest“ von Fakultätsseite befürchtet wurden, entschied man sich dafür, den allgemeinen Studierfähigkeitstest von den fachlichen Tests zu trennen. Bemerkenswerter Weise fällt hier auf, dass der allgemeine Studierfähigkeitstest – dessen Bearbeitung nun komplett auf freiwilliger Basis erfolgt – sehr gute Teilnahmezahlen aufzuweisen hat, nämlich etwa ein Drittel der Summe der Teilnehmerzahlen aller anderen 14 fachspezifischen Tests (siehe Tabelle 3). Von daher erscheinen Einwände gegen den allgemeinen Studierfähigkeitstest offensichtlich unbegründet. Da mit dem „allgemeinen Studierfähigkeitstest“ für den Studienerfolg wichtige Kompetenzen und Einstellungen wie selbst-reguliertes Lernen, Selbstwirksamkeit, Volition und Erwartungen an das Studium erfasst werden, wird die relativ hohe Quote freiwilliger Teilnahmen positiv bewertet im Sinne der Zielsetzung einer besseren Informiertheit bezüglich der mit der Aufnahme eines Studiums verbundenen Anforderungen.

### 4.3 Notwendige Personalressourcen

Sinnvollerweise sollten mindestens zwei Mitarbeitende exklusiv oder zumindest hauptberuflich für die Tests zuständig sein, und zwar einerseits für technische Angelegenheiten, andererseits für testpsychologische Belange. Die Inbetriebnahme erfordert – möglicherweise auch in Abhängigkeit vom verwendeten Portal – eine ständige technische Überwachung und Erneuerung, um einerseits die ständige Verfügbarkeit, andererseits die Aktualität zu gewährleisten. Ein Modell, wonach ein technischer Mitarbeitender sich eher nebenberuflich um das Hosting kümmert oder die Tests „neben-

bei mitlaufen“ lässt, hat sich an der eigenen Institution nicht bewährt, weshalb letztlich der „Umzug“ der Studierfähigkeitstest zur Fakultät Informatik erfolgte, wo diese Bedingungen gegeben sind.

Aus fachlicher Sicht sollten Online-Self-Assessments nicht ohne testpsychologische Kenntnisse eingeführt oder ausgewertet werden. Der Einführung eines (testpsychologisch und messtechnisch) qualitativ hochwertigen Tests gehen nahezu immer eine längere Pretest-Phase und mehrere Iterationen voraus, bis die eingesetzten Aufgaben den Qualitätsmaßstäben – d. h., das Erreichen akzeptabler messtechnischer Kennwerte – genügen. Das Fehlen eines solchen vorher stattfindenden Prozesses ist nicht nur ein theoretisches Desiderat, sondern führt tatsächlich auch in der Praxis zu unbefriedigenden Ergebnissen (z. B. aufgrund mangelhafter Reliabilitäten). Aufgrund des Veraltens oder Bekanntwerdens von Testfragen sind regelmäßige Kontrollen der gängigen messtechnischen Parameter und daraus folgende Eliminierungen bzw. Testrevisionen obligat. (Test-) Psychologische Kenntnisse und Aspekte sollten ebenfalls bei der Frage nach einer – in Abhängigkeit von der Zielsetzung der Testung – sinnvollen Ergebnisrückmeldung berücksichtigt werden, um Fehlerwartungen und Missverständnissen bei den Teilnehmenden vorzubeugen.

#### **4.4 Herausforderungen im Alltagsbetrieb – Kooperation mit Fakultäten und unterschiedliche Interessenlagen je nach Studiengang**

Als eine nicht unerhebliche Schwierigkeit bei der (Weiter-)Entwicklung der Tests an der Technischen Hochschule Nürnberg hat sich die Zulieferung von fachspezifischen Aufgaben durch die Fakultäten heraus gestellt. Generell bedarf es bei der Konzeption von fachspezifischen Aufgaben, die dann in Testitems „übersetzt“ werden, der Zuarbeit von Fachexpertinnen und -experten. Der Aufwand für Dozierende sowie Professorinnen und Professoren ist dabei durchaus nicht trivial, und ein unmittelbarer Nutzen für diesen Mehraufwand allenfalls mittelfristig erkennbar. Insofern empfiehlt es sich, entweder den Mehrwert der Tests für die Fakultäten deutlicher zu machen, oder Anreizsysteme zu installieren.

Wie sich über die Jahre gezeigt hat, ist die Motivation der Studiengänge, einen Studierfähigkeitstest einzuführen, von sehr unterschiedlichen Interessenlagen geprägt. Sehr große Studiengänge mit mehreren hundert Studienanfängern stehen dem Instrument offener gegenüber als kleinere Studiengänge, die eventuell zusätzlich mit anderen Studiengängen der gleichen oder anderer Hochschulen um Bewerber konkurrieren und deswegen sehr sensibel hinsichtlich potenzieller Abschreckungseffekte sind. Generell sollten solche Vorbehalte, auch wenn sie sich in der Realität nicht als zutreffend erweisen, von Seiten der Test-Verantwortlichen ernst genommen und nicht bagatellisiert werden, da die Akzeptanz des Tests von Seiten der Fakultät langfristig

sehr wichtig ist. Sollten sich Vorbehalte und Befürchtungen trotz empirischer Gegenbeweise nicht ausräumen lassen, bietet sich eine Kompromissfindung dergestalt an, dass die Test-Bearbeitung auf freiwilliger Basis erfolgt, Anreize in Form von vorab erzielbaren ECTS-Punkten bietet oder sogar eine andere Art von Test realisiert wird, um beispielsweise den Informationscharakter (wie etwa beim Online Self-Assessment Informatik an der Technischen Hochschule Nürnberg) stärker zu betonen.

Anhand dieser Ausführungen lässt sich erahnen, dass von Seiten des Online Self-Assessment-Teams gegenüber den Fakultäten, aber durchaus auch intern eine Haltung nötig ist, die versucht, theoretische Aspekte der „reinen Lehre“ mit Erfordernissen der Praxis zu vereinen. Dabei ist wichtig, zwar einerseits auf die unumgänglichen diagnostisch-wissenschaftlichen Mindeststandards zur Qualitätssicherung zu bestehen, dafür jedoch bei wünschenswerten, aber nicht obligaten Aspekten Entgegenkommen zu zeigen und eine praktikable Lösung zu akzeptieren, um die Widerstände möglichst gering zu halten oder auch ökonomisch vertretbar zu arbeiten. Die Erfahrung zeigt, dass trotz Aufklärungsarbeit wenig Wissen und Verständnis für die Relevanz psychologischer Diagnostik vorhanden ist, so dass mit rigider Orthodoxie in der Praxis wenig erreichbar ist, insbesondere, wenn die erste Begeisterung des neuen Projekts „Studierfähigkeitstests“ vergangen ist und der Test zum Alltagsgeschäft gehört. Dazu zählt beispielsweise auch die Akzeptanz der (Praxis-)Relevanz nicht-fachlicher, aber durchaus positiver Nebeneffekte der Test-Einführung, die etwa der Studiengang Betriebswirtschaft bemerken konnte: Seit der verpflichtenden Testteilnahme in diesem Studiengang sanken die Bewerbungen etwas, dafür stiegen die Annahmquoten. Das heißt, dass Teilnehmende, die sich seit der Test-Einführung an der Technischen Hochschule Nürnberg um einen Studienplatz in Betriebswirtschaft bewerben, diesen (möglicherweise aus Gründen der kognitiven Dissonanz-Reduktion?) zuverlässiger annehmen als vor der Einführung, was für die Fakultät hinsichtlich Planungssicherheit bezüglich der benötigten Ressourcen einen großen Vorteil darstellt und die Online Self-Assessments hier entsprechend beliebt sind, völlig unabhängig von den diagnostischen Qualitäten des Tests. Man kann in der Praxis also sehen, dass Online Self-Assessments nicht abschreckend wirken müssen, sondern das Engagement auf Seiten der Teilnehmenden als „Nebenwirkung“ erhöhen können; ein Effekt, den sich bekanntermaßen viele Vereinigungen zunutze machen (Aronson, Wilkert & Akert, 2006).

## **5 Zusammenfassung und Zukunftsaussichten**

Studierfähigkeitstests haben für die Prognose von Studienerfolg und somit für die Reduktion von Studienabbrüchen durchaus großes Potenzial, wenn die Rahmenbedingungen geschaffen werden, dieses vollumfänglich ausnutzen zu können. Neben der prognostischen Validität der Tests sind durchaus auch deren positiven „Nebeneffekte“ zu berücksichtigen und zu beachten: So ist die Informiertheit der Studien-

interessierten über die sie im Studium erwartenden Aufgabenstellungen ein nicht zu vernachlässigender Faktor, den informelle Schnuppertage oder Schulbesuche in diesem Detaillierungsgrad kaum leisten können. Auch eine (bei ernsthafter Bearbeitung) individuelle Rückmeldung über den für das Studienfach relevanten Fähigkeitsstand ist eine Information, über die viele Studieninteressierte ohne Studierfähigkeitstest lediglich implizit, oft nur in Form einer aggregierten Schulnote, verfügen. Und schließlich vermitteln Online Self-Assessments nicht nur die für ein Studium benötigten Fähigkeiten, sie können auch – z.B. bei Wiederholung des Tests nach einer festgelegten Zeitspanne – verdeutlichen, dass solche Fähigkeiten nicht notwendigerweise stabil, sondern veränderbar sind, sie können also neben der Status- auch der Prozessdiagnostik dienen; ein Ansatz, den die Technische Hochschule Nürnberg aktuell näher anhand unterschiedlicher Szenarien verfolgt. Zugleich bedeuten Tests aber auch wie erwähnt einen hohen Kosten- und Personalaufwand. Die Einführung von Studierfähigkeitstests lässt sich am ehesten über eine eindeutige bildungspolitische Zielsetzung begründen.

Perspektivisch sind nun in Bezug auf die Online Self-Assessments vor allem zwei Dinge von politischer Seite wünschenswert: Erstens wäre es wichtig, die rechtlichen Rahmenbedingungen zu schaffen, so dass Zulassungsverfahren, wie sie an der Technischen Universität München (s.o.) stattfinden, ohne größeren bürokratischen und zeitlichen Vorlauf stattfinden können, und zwar für alle Studiengänge, die das Verfahren einführen möchten (und nicht nur solche mit bestimmten Charakteristika). Der zweite Punkt zielt auf eine bildungspolitische Grundsatzfrage. So begrüßenswert prinzipiell eine höhere Durchlässigkeit ins Studium ist, so wichtig ist es, die daraus resultierenden Konsequenzen zu bedenken. Die Zunahme der Abiturientenquote ist nicht gleichbedeutend mit der Anzahl studierfähiger junger Menschen. Wenn Hochschulen hier nicht Maßnahmen zum Umgang mit der inzwischen sehr heterogenen Gruppe von Studienanfängern entwickeln, sind hohe Studienabbrecherquoten auch eine mögliche Folge. Hochschulen können – und wollen – sich aber nicht lediglich auf diejenigen Studieninteressierten fokussieren, die nach traditionellen Kriterien die höchste Erfolgswahrscheinlichkeit aufweisen. Das wäre weder sozialpolitisch gerecht noch ökonomisch zu rechtfertigen. Sinnvoller erscheint vielmehr eine dialektische Synthese von „höherer Durchlässigkeit“ und „gute Passung zwischen Studiengang und Studierendem“ im Sinne eines Ergänzungsverhältnisses (sensu Wertequadrat von Schulz von Thun, 1989). In diesem Zusammenhang spielen Eignungsfeststellungsverfahren oder Studierfähigkeitstests eine wichtige Rolle. Denn sie stellen sicher, dass das Ausmaß der Passung zwischen den Anforderungen des Studiengangs sowie den fachlichen und motivationalen Eigenschaften des jeweiligen Studienanfängers entsprechend frühzeitig diagnostiziert werden kann. Damit können in der Folge diversitätsgerechte Unterstützungsmaßnahmen konzipiert werden, die sich auf valide diagnostischen Daten stützen, statt auf plausible Überlegungen, was Studienanfänger brauchen könnten. In diesem Sinne leisten Tests

einen nicht zu unterschätzenden Beitrag, wenn es um die Entscheidung zur Mittelverteilung für entsprechende Maßnahmen geht. Online Self-Assessments liefern datengestützte Erkenntnisse über eine zunehmend heterogene Gruppe. Aus diesen Daten können Hochschulen also kontinuierlich lernen sowie Supportstrukturen und -maßnahmen entsprechend anpassen. Damit lohnt sich die Investition in Tests – plausible, aber letztlich wirkungslose Maßnahmen aufgrund nicht überprüfter Vorannahmen ließen sich dadurch möglicherweise minimieren.

## Literatur

Aronson, E., Wilson, T., & Akert, R. (2004). Sozialpsychologie. München: Pearson Verlag

Fellenberg, F., & Hannover, B. (2006). Kaum begonnen, schon zerronnen? Psychologische Ursachenfaktoren für die Neigung von Studienanfängern, das Studium abzubrechen oder das Fach zu wechseln. *Empirische Pädagogik*, 20, 381–399

Gold, A., & Souvignier, E. (2005). Prognose der Studierfähigkeit. Ergebnisse aus Längsschnittanalysen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37, 214–222

Hell, B., Trapmann, S., & Schuler, H. (2007). Eine Metaanalyse der Validität von fachspezifischen Studierfähigkeitstest im deutschsprachigen Raum. *Empirische Pädagogik*, 21, 251 – 270

Heublein, U., Ebert, J., Hutzsch, C., Isleib, S., König, R., Richter, J., & Woisch, A. (2017). Zwischen Studiererwartungen und Studienwirklichkeit. *Forum Hochschule*, abgerufen von [http://www.dzhw.eu/pdf/pub\\_fh/fh-201701.pdf](http://www.dzhw.eu/pdf/pub_fh/fh-201701.pdf)

Kubinger, K.D. (2015). Kritische Reflexion zu Self-Assessments im Rahmen der Studienberatung. *Das Hochschulwesen*, 63, 76–80

Marsch, U. (2017, 29. November). TU München steigert Studiererfolg mit Eignungsprüfungen. Abgerufen von <https://idw-online.de/de/news685544/>

Osel, J. (2016, 11. Oktober). Experten plädieren für verbindliche Uni-Eignungstests. Abgerufen von <http://www.sueddeutsche.de/bayern/hochschulen-in-bayern-experten-plaedieren-fuer-verbindliche-eignungstests-1.3196885/>

Schulz von Thun, F. (1989). *Miteinander reden: Band 2. Stile, Werte und Persönlichkeitsentwicklung: Differentielle Psychologie der Kommunikation*. Hamburg: Rowohlt Verlag

Studienabbruch: Staat setzt jährlich 2,2 Mrd. Euro in den Sand. (2007, 02. Oktober). Abgerufen von <https://www.presseportal.de/pm/18931/1057835/>

Trost, G. (2003). *Deutsche und internationale Studierfähigkeitstests. Arten, Brauchbarkeit, Handhabung. Dokumentationen & Materialien, Band 51*. Bonn: Deutscher Akademischer Austauschdienst

Manuskript eingereicht: 03.07.2018  
Manuskript angenommen: 08.10.2018

**Anschrift der Autorin:**

Diana Wolff-Grosser  
Diplom-Psychologin  
Mitarbeiterin Testentwicklung Online Self-Assessments  
Technische Hochschule Nürnberg Georg Simon Ohm  
Hohfederstraße 40  
90489 Nürnberg  
E-Mail: [diana-wolff-grosser@th-nuernberg.de](mailto:diana-wolff-grosser@th-nuernberg.de)

## Verleihung des Ulrich Teichler-Preises und des Nachwuchspreises der Gesellschaft für Hochschulforschung

Am 30. März 2017 wurden zwei Nachwuchswissenschaftlerinnen und ein Nachwuchswissenschaftler für ihre Leistungen auf dem Gebiet der Hochschulforschung ausgezeichnet. Die Würdigung und feierliche Preisverleihung erfolgte anlässlich der zwölften Jahrestagung der Gesellschaft für Hochschulforschung, die das Deutsche Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) vom 30. bis 31. März 2017 in Hannover organisierte.

Der Ulrich-Teichler-Preis wurde 2008 zum ersten Mal vergeben. Der Kasseler Hochschulforscher Professor Dr. Ulrich Teichler stiftete den Preis, um junge Wissenschaftlerinnen und Wissenschaftler anzuregen, sich mit Fragen der Hochschulforschung wissenschaftlich auseinanderzusetzen. Zugleich sollten herausragende Nachwuchsarbeiten zu dieser Thematik stärker in der Öffentlichkeit bekannt gemacht werden. Die Initiative wurde bis 2013 vom Bundesministerium für Bildung und Wissenschaft unterstützt.

Aus den 19 eingereichten Arbeiten wählte eine sechsköpfige Jury die folgenden Preisträgerinnen aus:

- Heike Spangenberg:  
Dissertation „Konvergenzen oder dauerhafte Unterschiede? Zur Entwicklung der Studienentscheidungen und nachschulischen Bildungsverläufe seit 1990 in Ost- und Westdeutschland “
- Christian Schmid  
Dissertation „Die soziale Organisiertheit und Organisierbarkeit von Interessen(freiheit) – Der Fall der managerialen Governance akademischer Lehrtätigkeit“
- Klara Reichenbach  
Masterarbeit „Besprechung der Evaluationsergebnisse und Evaluationszeitpunkt als Determinanten der Teilnahmemotivation an studentischer Lehrveranstaltungsevaluation“

Die Jury war der Ansicht, dass alle drei Arbeiten einen fundierten Beitrag zur fachlichen und professionellen Weiterentwicklung der Hochschulforschung leisten. Die „Beiträge zur Hochschulforschung“ veröffentlichen nachfolgend die Abstracts der drei prämierten Nachwuchsarbeiten.

# Konvergenzen oder dauerhafte Unterschiede? Zur Entwicklung der Studienentscheidungen und nachschulischen Bildungsverläufe seit 1990 in Ost- und Westdeutschland

Heike Spangenberg (Dissertation, Humboldt-Universität zu Berlin)

Seit der Wiedervereinigung Deutschlands im Jahr 1990 ist mehr als ein Vierteljahrhundert vergangen. Zu Beginn des damit für die fünf „neuen Länder“ verbundenen Transformationsprozesses wurde angenommen, dass nach einer schwierigen Phase des Umbruchs in Ostdeutschland nicht nur eine schnelle Angleichung auf der makro- und mesogesellschaftlichen Ebene an westdeutsche Verhältnisse erfolgen würde, sondern auch auf der Mikroebene der Denk- und Verhaltensmuster der Menschen und damit unter anderem auch bei anstehenden Bildungsentscheidungen.

Vor dem Hintergrund der zwischen Ost- und Westdeutschland nach wie vor und zum Teil erheblich differierenden Anteile junger Männer und Frauen, die eine schulische Hochschulzugangsberechtigung erlangen (Studienberechtigtenquote) und diese anschließend durch den Übergang an eine Hochschule einlösen (Studierquote), werden mittels eines selbst entwickelten Modells, das sich an die soziologische Wert-Erwartungs-Theorie von Erikson & Jonsson (1996) sowie den lebensverlaufstheoretischen Ansatz von Mayer (1990) anlehnt, verschiedene individuelle und kontextuelle Einflussfaktoren der Studienentscheidung erstmals in einem Kohortenvergleich seit 1990 regressions- und mehrebenenanalytisch betrachtet. Formale Bildungszertifikate sind in modernen Gesellschaften von zentraler Bedeutung, unter anderem für berufliche Karrieren und die soziokulturelle Teilhabe. Das individuelle Bildungsverhalten der Studienberechtigten und seine Konsequenzen für spezifische monetäre und nicht-monetäre Bildungserträge sowie für die erworbenen Kompetenzen sind wiederum von erheblicher Bedeutung für die Wirtschaft und die Gesellschaft insgesamt.

Der Übergang in eine Hochschule stellt nur *eine* Schwelle im gesamten nachschulischen Bildungsverlauf von Studienberechtigten dar. Daher werden erstmals die Bildungsverläufe der Studienberechtigtenkohorten 1990 und 1999 auch in ihrer Gesamtheit, also unter Berücksichtigung von Mehrfachqualifizierung, Fortbildung und Promotion, sowie die Erwerbs- und Familienverläufe über einen Zeitraum von zehneinhalb Jahren nach Erwerb der Hochschulzugangsberechtigung vergleichend in den Blick genommen und mittels Sequenzmusteranalysen jeweils typische Verlaufsmuster für ost- und westdeutsche Studienberechtigte ermittelt.

Zur Untersuchung der zentralen Forschungsfrage nach Konvergenzen, Divergenzen und dauerhaften Unterschieden in den individuellen und kontextuellen Einflussfaktoren der Studienentscheidung sowie den nachschulischen Bildungsverläufen seit 1990 in Ost- und Westdeutschland werden Daten der Studienberechtigtenpanel des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW) 1990, 1994, 1999, 2002 und 2006 verwendet.

Die Arbeit identifiziert zusammenfassend zahlreiche Konvergenzen und Gemeinsamkeiten, insbesondere bei den individuellen Einflussfaktoren für eine Studienentscheidung. Bereits bei der Kohorte 1994 hatten sich bezüglich des Geschlechtereinflusses die Verhältnisse in Ostdeutschland an diejenigen in Westdeutschland angeglichen und Frauen wiesen eine signifikant geringere Studierwahrscheinlichkeit auf als Männer. Insgesamt lässt sich hinsichtlich der Einflussfaktoren der Studienentscheidung aber häufig keine gerichtete Entwicklung, kein klarer Trend über alle Kohorten hinweg beobachten. Es gab zum Teil Unterbrechungen von zuvor konvergenten oder auch divergenten Entwicklungen, zunächst vorhandene Gemeinsamkeiten mündeten in kurzzeitige oder auch längere Ost-West-Unterschiede, um sich anschließend wieder einzustellen. Es lässt sich in der Konsequenz kein einheitliches Gesamtbild zeichnen. Bei aller Varianz der bedeutsamen Einflussfaktoren zwischen den Kohorten lassen sich aber Effekte individueller Faktoren sowie Kontext- und Kompositionseffekte erkennen, die weitgehend stabil Ost-West-Unterschiede erklärten. Dies gilt beispielsweise für die individuellen Faktoren der antizipierten Opportunitätskosten und bisherigen Bildungsbiographie, den Kontextfaktor der Hochschulentfernungen sowie die Kompositionseffekte überdurchschnittlicher Anteile von ostdeutschen Studienberechtigten, die hohe Opportunitätskosten für ein Studium antizipierten, und die aus Arbeiterhaushalten kamen.

Die nachschulischen Bildungs- und Lebensverläufe weisen bereits bei der Kohorte 1990 erhebliche Gemeinsamkeiten zwischen ost- und westdeutschen Studienberechtigten auf. Darüber hinaus bestehen besondere Spezifika. Sowohl in Ost- als auch in Westdeutschland war anhand der identifizierten Verlaufstypen zwischen den Kohorten 1990 und 1999 eine Bedeutungsverschiebung von einer beruflichen Fortbildung hin zu einer Promotion und einem berufsbegleitendem Studium festzustellen. Für die ostdeutschen Studienberechtigten zeigt sich zudem ein erheblicher Bedeutungszuwachs der Doppelqualifizierung. Des Weiteren waren die ostdeutschen Studienberechtigten der Kohorte 1999 zu vergleichsweise hohen Anteilen und bei vier der insgesamt sieben dort identifizierten Verlaufstypen überdurchschnittlich lang von Arbeitslosigkeit betroffen. Auch der Anteil von familienbezogenen Tätigkeiten war in der Mehrzahl der bei ihnen vorkommenden Verlaufstypen relativ hoch. Bezüglich dieser zwei Aspekte haben sich bei der Kohorte 1999 neue Unterschiede zwischen ost- und westdeutschen Studienberechtigten herausgebildet.

## Literatur

Erikson, R. & Jonsson, J. O. (Hrsg.) (1996): Can education be equalized? The Swedish case in comparative perspective. Boulder, Colo: Westview Press

Mayer, K. U. (Hrsg.) (1990): *Lebensverläufe und sozialer Wandel*. Opladen: Westdt. Verlag

Spangenberg, Heike (2017): Konvergenzen oder dauerhafte Unterschiede? Zur Entwicklung der Studienentscheidungen und nachschulischen Bildungsverläufe seit 1990 in Ost- und Westdeutschland. <https://edoc.hu-berlin.de/handle/18452/18336> (Zugriff: 23.10.2018)

### **Anschrift der Autorin:**

Dr. Heike Spangenberg

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW)

Abteilung Bildungsverläufe und Beschäftigung

Lange Laube 12

30159 Hannover

E-Mail: [spangenberg@dzhw.eu](mailto:spangenberg@dzhw.eu)

## Die soziale Organisiertheit und Organisierbarkeit von Interessen(freiheit) – Der Fall der managerialen Governance akademischer Lehrtätigkeit

Christian Johann Schmid (Dissertation, Technische Universität Dortmund)

Welche intendierten und nicht-intendierten Einflusspotenziale haben Organisationen auf das Interesse ihrer Mitglieder, Dinge erstens überhaupt (Handlungs-Motivation) und zweitens in bestimmten Erledigungs-Modi (Handlungs-Weisen) zu tun? Diese organisationssoziologische Fragestellung wurde in der Form einer theoretischen Reinterpretation bzw. ex post-Reflektion mehrjähriger Drittmittel-Forschung zur managerialen Governance akademischer Lehrtätigkeit an deutschen Hochschulen beantwortet. Dazu hat der Autor umfangreiches empirisches Datenmaterial (Statistiken, Online-Befragungen, Interviews) zusammengestellt und mit verschiedenen Methoden der quantitativen und qualitativen Datenanalyse ausgewertet.

Im primären Rekurs auf die bourdieusche Sozial- bzw. Praxis-Theorie im Allgemeinen und dessen Interesse-Begriff im Speziellen wurde diese Empirie dann dahingehend

gedeutet, wie die akademische Lehr-Praxis schon immer sozial organisiert ist; das heißt ohne, jenseits oder trotz gezielter Interventionsmaßnahmen durch das Hochschul-Management. Darauf aufbauend wurden dann ganz konkrete, anwendungsbezogene Management-Vorschläge für Hochschulleitungsorgane gefolgert, um das Lehr-Engagement sowie die Lehr-Ansätze von Professorinnen und Professoren möglichst handlungswirksam zu reorganisieren. Ganz im Sinne Bourdieus gilt es dabei, die axiomatischen Reduktionen und Auslassungen jener ökonomischen Theorien zu identifizieren, zu kompensieren und zu vermeiden, welche ideologisch dafür instrumentalisiert wurden, die bisherige Reform der Hochschulbinnenorganisation nach primärer Maßgabe des New Public Managements (NPM) zu gestalten.

Mit der vorliegenden Abhandlung zielt der Autor letztlich auf organisationssoziologische Verallgemeinerungsprofite durch Denkfiguren ab, welche zwar am spezifischen Fall der Hochschulorganisation entwickelt bzw. auf diesen angewendet wurden, aber dennoch darüber hinwegverweisen sollen: eine Soziologie der Organisation von Interessenfreiheit.

## **Literatur**

Schmid, Christian Johann (2016): Die soziale Organisiertheit und Organisierbarkeit von Interessen(freiheit) – Der Fall der managerialen Governance akademischer Lehrtätigkeit. <https://eldorado.tu-dortmund.de/handle/2003/34952> (Zugriff: 16.10.2018)

## **Anschrift des Autors:**

Dr. Christian Johann Schmid  
Technische Universität Dortmund  
Center for Higher Education  
Hohe Straße 141  
44139 Dortmund  
E-Mail: christian.schmid@tu-dortmund.de

# Besprechung der Evaluationsergebnisse und Evaluationszeitpunkt als Determinanten der Teilnahmemotivation an studentischer Lehrveranstaltungsevaluation

Klara Reichenbach (Masterarbeit, Humboldt-Universität zu Berlin)

Ziel der Masterarbeit war es, relevante situationsbezogene Einflussfaktoren auf die Teilnahmemotivation an studentischer Lehrveranstaltungsevaluation zu identifizieren sowie diese detailliert zu beschreiben und systematisch zu untersuchen, um daraus konkrete Handlungsempfehlungen für die Hochschulpraxis abzuleiten. Als relevante Determinanten wurden die Besprechung der Evaluationsergebnisse mit den Studierenden (Ergebnisbesprechung vs. keine Ergebnisbesprechung) sowie der Evaluationszeitpunkt (Semestermitte vs. -ende) betrachtet. Es wurde ein positiver Effekt der Ergebnisbesprechung und der Befragung zur Semestermitte auf die lehrevaluationsbezogene, internale Kontrollüberzeugung – definiert als das Gefühl der Studierenden, durch ihre Teilnahme etwas bewirken zu können – sowie auf die Teilnahmemotivation angenommen. Die Prädiktoren wurden mittels vier verschiedener, randomisiert präsentierter Vignetten im Rahmen eines Online-Experiments (N = 239 Studierende) variiert. Bei der Berechnung zweier multipler Regressionsanalysen für die beiden abhängigen Variablen ergab sich die Ergebnisbesprechung mit einer kleinen Effektstärke *sensu* Cohen (1988) als signifikanter Prädiktor. Darüber hinaus konnte eine Mediationsanalyse zeigen, dass der Zusammenhang zwischen der Ergebnisbesprechung und der Teilnahmemotivation durch die lehrevaluationsbezogene, internale Kontrollüberzeugung vermittelt wird. Die Ergebnisse betonen die Relevanz der Berücksichtigung situationsbezogener Einflussfaktoren für die Evaluationspraxis an Hochschulen. Abschließend werden konkrete praktische Implikationen im Umgang mit Studierenden und Dozierenden diskutiert.

## Literatur

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates

## Kontakt:

Klara Reichenbach

E-Mail: [klara.reichenbach@web.de](mailto:klara.reichenbach@web.de)

## Buchvorstellungen

---

Dittler, Ullrich; Kreidl, Christian (Hrsg.) (2018): Hochschule der Zukunft. Beiträge zur zukunftsorientierten Gestaltung von Hochschulen. Wiesbaden: Springer, ISBN 978-3-658-20402-0, 307 Seiten

---

Vor dem Hintergrund aktueller Entwicklungen wie Digitalisierung, Arbeitswelt 4.0 und Erhöhung der Studierendenquote widmet sich dieser Sammelband der Frage, wie Hochschulen zukünftig gestaltet werden können, um ihren vielfältigen Aufgaben in Lehre, Forschung, Third Mission und den Erwartungen verschiedenster Akteure gerecht zu werden. Nach einem einleitenden Kapitel der Herausgeber zur Entwicklung des Hochschulwesens und einer Bestandsaufnahme zur hochschulischen Lehre finden sich insgesamt 14 Beiträge von deutschen und österreichischen Autoren und Autorinnen aus vier unterschiedlichen Perspektiven. Zunächst werden aus der Sicht von Hochschulleitungen die Rolle von Innovation und Digitalisierung sowie das Konzept des lebenslangen Lernens und die vielfältigen Stellschrauben zur Ausbildung eines Zukunftskonzepts für Hochschulen dargestellt. Aus bildungsdidaktischer Perspektive werden sodann das Verständnis akademischer Bildung in Abgrenzung zur Berufsbildung, das Spannungsfeld Forschung und Lehre, die Diversität von Lehr-/Lernveranstaltungen sowie die Digitalisierung im Bereich der Lehre diskutiert. Der abschließende Beitrag formuliert Wünsche und Anforderungen aus Sicht von Unternehmen in Verbindung mit Vorschlägen, welche Lehrinhalte daraus hervorgehen sollten und benennt sieben Aspekte, mit denen die Hochschulausbildung den Anforderungen des Arbeitsmarkts entgegen kommen kann.

---

Schulz, Jens (2017): Hochschulentwicklung und E-Learning. Digitalisierung als organisationale Herausforderung. Glückstadt: vvh-Verlag, ISBN 978-3-86488-124-4, 243 Seiten

---

Das Buch analysiert die Entwicklungspotenziale der deutschen Hochschulen auf dem Gebiet der Organisation des E-Learnings. Auf der Basis einer Dokumentenanalyse und einer schriftlichen Befragung von Expertinnen und Experten aus Hochschulpolitik und Hochschulen, insbesondere aus den Bereichen Hochschulentwicklungsplanung, E-Learning-Unterstützungsstrukturen und -initiativen, werden wichtige Aspekte der Organisation und Weiterentwicklung des E-Learning im Zusammenhang aktueller Hochschuldebatten identifiziert und diskutiert. Ein wesentlicher Aspekt der Dissertation liegt auf den strukturellen Anforderungen des E-Learnings an die Arbeitsebene und auf Handlungsempfehlungen für Dienstleistungen und Support, Steuerung und Kooperationen.

---

Welpel, Isabell M.; Brosi, Prisca; Schwarzmüller, Tanja (2018): Digital Work Design. Die Big Five für Arbeit, Führung und Organisation im digitalen Zeitalter. Frankfurt/New York: Campus Verlag, ISBN 978-3-593-50851-1, 248 Seiten

---

Ausgehend von den Ergebnissen des Projekts „Digital Work Design – Turning Risks into Chances“ veranschaulicht das Buch, wie Digitalisierung die tägliche Arbeit sowie die spezifischen Aufgabengebiete von Mitarbeiterinnen und Mitarbeitern mit und ohne Leitungsfunktion verändert sowie welche Chancen und Risiken damit verbunden sind. Auf der Basis von Interviews mit Digitalisierungsexperten aus Wirtschaft, Wissenschaft und Politik, Interviews mit Beschäftigten und Tagebuchstudien werden fünf Erfolgsfaktoren für den digitalen Wandel von Arbeitsorganisationen herausgearbeitet sowie Handlungsempfehlungen für Organisationen, Führungskräfte und Mitarbeiter formuliert. Die fünf Erfolgsfaktoren („Big Five“) sind: (1) Kernkompetenzen im Umgang mit Unbeständigkeit, Unsicherheit, Komplexität und Mehrdeutigkeit, (2) neue Arten von Teamarbeit, (3) Demokratisierung von Organisationen, (4) Bedeutung von Beziehungen, (5) Bedeutung von Gesundheit. Ein 10-Punkte-Plan, ein Selbst-Diagnose-Check und Handlungsempfehlungen runden dieses informative und praxisnahe Handbuch ab.

---

Falk, Susanne; Reimer, Maike; Schmidt, Uwe (Hrsg.) (2018): Absolventenstudien und Qualitätsmanagement. Best Practices an deutschen und österreichischen Hochschulen. Münster: Waxmann, ISBN: 978-3-8309-3885-9, 188 Seiten

---

Qualitätssicherungsverfahren nutzen zunehmend Ergebnisse aus Absolventenbefragungen, um die Perspektive von Absolventinnen und Absolventen für die Bewertung von Qualitätszielen im Bereich Studium und Lehre nutzen zu können. Wie aber werden diese Daten erhoben, wie werden sie in die Qualitätsprozesse an Hochschulen eingebunden, und welche Formen der Rückmeldung erweisen sich als hilfreich? Die elf Beiträge dieses Bands zeigen, wie Universitäten und Hochschulen für angewandte Wissenschaften in Deutschland und Österreich tragfähige Wege entwickelt haben, um produktiv und erfolgreich mit diesen Rückmeldungen zu arbeiten. Sie thematisieren auch die Leitbilder der Qualitätsmanagementsysteme, die Reichweite und Grenzen der Aussagekraft von Absolventenstudien sowie die persönlichen, organisationalen und wissenschaftlichen Voraussetzungen, die gegeben sein müssen, damit die Daten tatsächlich in konkrete Verbesserungen münden können. Als zentral erweist sich ein intensiver Kommunikationsprozess, bei dem die Verantwortlichen für die Organisation und Durchführung von Lehre eingebunden werden. Die Akzeptanz der Rückmeldungen bei den Lehrenden und Verantwortlichen für die Studiengänge ist Voraussetzung dafür, dass ihre Expertise und Erfahrung genutzt werden können, um die notwendigen Maßnahmen zur Verbesserung zu planen.

## Jahresindex 2018

	Heft	Seite
Behle, Heike; Maher, Sunil: Measuring "Teaching Excellence" and "Learning Gain" in the United Kingdom	2	48
Bittmann, Felix: Über den Nutzen von Pflichtpraktika. Eine Replikation der Studie von Klein & Weiss (2011) mit Daten des Bayerischen Absolventenpanels	3	78
Braun, Edith; Athanassiou, Georgios; Pollerhof, Kathleen; Schwabe, Ulrike: Wie lassen sich kommunikative Kompetenzen messen? – Konzeption einer kompetenzorientierten Prüfung kommunikativer Fähigkeiten von Studierenden	3	34
Buß, Imke: Erfolgreich studieren mit Beeinträchtigung durch Interaktionen im Studium	3	56
Falk, Susanne; Reimer, Maike: Die „Bayerischen Absolventenstudien“ (BAS): Ein Instrument zur Unterstützung der Qualitätssicherung an Hochschulen	3	96
Förster, Manuel; Heiß, Florian; Klink, Sigbert; Maur, Andreas; Schank, Thorsten; Weiser, Constantin: Die Implementation und Evaluation eines Flipped Classrooms in einer Großveranstaltung der Statistik	4	50
Franke, Ray: Drinking, Smoking, Partying – And Still Graduate On Time? Eine Mehrebenenanalyse zum Einfluss von adversen Studierverhalten auf den Bachelorabschluss in den USA	1	34
Franzen, Martina: Die digitale Transformation der Wissenschaft	4	8
Gottburgsen, Anja; Willige, Janka: Mehr Mobilitätserfahrungen durch digitale Medien? Zu den Effekten von studentischer Diversität und Lernumweltsmerkmalen auf die internationale Mobilität	4	30
Kleimann, Bernd; Hückstädt, Malte: Auswahlkriterien in Berufungsverfahren: Universitäten und Fachhochschulen im Vergleich	2	20
Mallwitz, Michelle; Adler, Philipp; Benzler, Guido: Der Einfluss von Evaluierungen auf Forschungseinrichtungen am Beispiel der Leibniz-Gemeinschaft	2	8
Möhringer, Jutta; Baumgartner, Raphaela: Begabtenförderprogramme für Studierende – Eine Bestandsaufnahme an deutschen Universitäten	1	8

Overberg, Jasmin: Beratung statt Kontrolle – Externe Qualitätssicherung an finnischen Hochschulen und ihre Wahrnehmung aus der Perspektive des Hochschulpersonals	1	80
Riedel, Jana; Möbius, Kathrin: Bestandsaufnahme, Hindernisse und Möglichkeiten des Einsatzes von E-Assessment an sächsischen Hochschulen	4	68
Stöver, Britta; Sibbertsen, Philipp: Die räumliche Flexibilität von Studierenden – Gründe für das Wanderungsverhalten von Studienanfängerinnen und Studienanfängern zwischen den Bundesländern	3	8
Tafertshofer, Lorenz; Werner, Evamaria; Schmidt-Hertha, Bernhard: Grundlagen der Reputation von Studienstandorten: Bewertungsmaßstäbe für die Qualität von Hochschulstandorten und Studiengängen aus der Sicht von sozialwissenschaftlichen Professorinnen und Professoren	2	68
Wolff-Grosser, Diana: Online-Self-Assessments an der Technischen Hochschule Nürnberg – Bilanzierung, Reflexion, Zukunftsvision	4	88
Zawacki-Richter, Olaf; Müskens, Wolfgang; Garz, Detlef; Gierke, Willi B.: Profile von Studienbewerberinnen und -bewerbern ohne Abitur – Ergebnisse einer Latenten Klassenanalyse auf der Basis von Prüfungsakten von 1971 bis 2010	1	58

## Index 2018

English abstracts of all articles can be found on the first pages of the respective issue. The page numbers below refer to the full articles.

	issue	page
Behle, Heike; Maher, Sunil: Measuring "Teaching Excellence" and "Learning Gain" in the United Kingdom	2	48
Bittmann, Felix: About the effect of mandatory internships. A replication of the study of Klein and Weiss (2011) using data from the Bavarian Graduate Panel	3	78
Braun, Edith; Athanassiou, Georgios; Pollerhof, Kathleen; Schwabe, Ulrike: How can we scale communication skills? – Conception of a competence-based testing of communication skills of students	3	34
Buß, Imke: Study success of disabled students – integration as an important factor in higher education institutions	3	56
Falk, Susanne; Reimer, Maïke: The Bavarian Graduate Studies (BAS): An instrument to support quality assurance in Higher Education	3	96
Förster, Manuel; Heiß, Florian; Klinke, Sigbert; Maur, Andreas; Schank, Thorsten; Weiser, Constantin: The implementation and evaluation of a flipped classroom within a large statistics lecture	4	50
Franke, Ray: Drinking, smoking, partying – And still graduate on time? Using a multilevel analysis to examine the effect of adverse study behavior on 4-year degree completion	1	34
Franzen, Martina: The digital transformation of science	4	8
Gottburgsen, Anja; Willige, Janka: More mobility through digitalisation? On the effects of students' diversity and the characteristics of the learning environment on study-related international mobility	4	30
Kleimann, Bernd; Hückstädt, Malte: Selection criteria in appointment procedures for professorships: a comparison of universities and universities of applied sciences	2	20
Mallwitz, Michelle; Adler, Philipp; Benzler, Guido: How institutional evaluations influence research institutions using the example of the Leibniz Association	2	8
Möhringer, Jutta; Baumgartner, Raphaela: Fostering the talented – Assistance programs to particularly gifted students in German universities	1	8

Overberg, Jasmin: Enhancement instead of control – External quality assurance in Finnish higher education institutions and its perception by university staff	1	80
Riedel, Jana; Möbius, Kathrin: Current state and possibilities to use e-assessment at Saxon universities	4	68
Stöver, Britta; Sibbertsen, Philipp: Drivers for the spatial flexibility and the migration of first-year students	3	8
Tafertshofer, Lorenz; Werner, Evamaria; Schmidt-Hertha, Bernhard: Basics of a study of the reputation of university locations: Evaluation criteria for the quality of university locations and study programs from social science professors' perspective	2	68
Wolff-Grosser, Diana: Online-Self-Assessments at the Technische Hochschule Nürnberg – conclusions, reflection, perspectives	4	88
Zawacki-Richter, Olaf; Müskens, Wolfgang; Garz, Detlef; Gierke, Willi B.: Profiles of non-traditional students with vocational qualifications. Results from a latent class analysis based on examination files from 1971 to 2010	1	58

Wir danken dem Herausgeberbeirat  
sowie allen anonymen Gutachterinnen und Gutachtern  
für ihre wertvolle Arbeit.

## Hinweise für Autorinnen und Autoren

### **Konzept:**

Die Zeitschrift „Beiträge zur Hochschulforschung“ bietet Hochschulforschenden und Akteuren im Hochschulbereich die Möglichkeit zur Erstveröffentlichung von Artikeln, die wichtige Entwicklungen im Hochschulbereich aus unterschiedlichen methodischen und disziplinären Perspektiven behandeln. Dabei wird ein Gleichgewicht zwischen quantitativen und qualitativen empirischen Analysen, Vergleichsstudien, Überblicksartikeln und Einblicken in die Praxis angestrebt.

Eingereichte Artikel sollten klar und verständlich formuliert, übersichtlich gegliedert sowie an ein Lesepublikum aus unterschiedlichen Disziplinen mit wissenschaftlichem und praxisbezogenem Erwartungshorizont gerichtet sein.

### **Review-Verfahren:**

Wie für eine wissenschaftliche Zeitschrift üblich, durchlaufen alle eingereichten Manuskripte eine externe Begutachtung durch anonyme Sachverständige (double blind). Dabei kommen je nach Ausrichtung des Artikels folgende Kriterien zum Tragen: Relevanz des Themas, Berücksichtigung des hochschulpolitischen Kontexts, Praxisbezug, theoretische und methodische Fundierung, Qualität der Daten und empirischen Analysen, Berücksichtigung der relevanten Literatur, klare Argumentation und Verständlichkeit für ein interdisziplinäres Publikum. Die Autorinnen und Autoren werden über das Ergebnis schriftlich informiert und erhalten gegebenenfalls Hinweise zur Überarbeitung.

### **Umfang und Form der eingereichten Manuskripte:**

Manuskripte sollten bevorzugt per E-Mail eingereicht werden und einen Umfang von 20 Seiten/50 000 Zeichen mit Leerzeichen nicht überschreiten (Zeilenabstand 1,5, Arial 11). Ergänzend sollten je ein Abstract (maximal 1000 Zeichen mit Leerzeichen) in deutscher und in englischer Sprache sowie Anschrift und Angaben zur beruflichen Funktion des Autors beigefügt sein. Die Druckfassung wird extern von einem Graphiker erstellt.

Bitte beachten Sie in jedem Fall bei Einreichung eines Manuskripts die ausführlichen verbindlichen Hinweise für Autoren unter <http://www.bzh.bayern.de>.

### **Kontakt:**

Dr. Lydia Hartwig

E-Mail: [Beitraege@ihf.bayern.de](mailto:Beitraege@ihf.bayern.de)





## Aus dem Inhalt

Martina Franzen: Die digitale Transformation der Wissenschaft

Anja Gottburgsen, Janka Willige: Mehr Mobilitätserfahrungen durch digitale Medien? Zu den Effekten von studentischer Diversität und Lernumweltsmerkmalen auf die internationale Mobilität

Manuel Förster, Florian Heiß, Sigbert Klinke, Andreas Maur, Thorsten Schank, Constantin Weiser: Die Implementation und Evaluation eines Flipped Classrooms in einer Großveranstaltung der Statistik

Jana Riedel, Kathrin Möbius: Bestandsaufnahme, Hindernisse und Möglichkeiten des Einsatzes von E-Assessment an sächsischen Hochschulen

Diana Wolff-Grosser: Online-Self-Assessments an der Technischen Hochschule Nürnberg – Bilanzierung, Reflexion, Zukunftsvision