

SoTL PROJEKTBERICHTE



EFFEKTE STRUKTURELLER GAMIFIZIERUNG EINES
INFORMATIKMODULS AUF KOMPETENZERWERB
UND BESTEHENSQUOTE

TIM PIDUN

Hochschule Ansbach
Fakultät Technik, Studiengangleiter Angewandte Ingenieurwissenschaften

tim.pidun@hs-ansbach.de

ABSTRACT

In den Grundlagenfächern der Ingenieurwissenschaften sind immer noch recht hohe Abbruchquoten zu beobachten, die oft mit Überforderung und mangelnder Motivation der Studierenden im Übergang zwischen schulischem und hochschulischem Lernen zu erklären sind. Um den Lernenden einen eher eigenverantwortlichen und damit intrinsisch motivierenderen Ansatz zu bieten, werden in der Hochschuldidaktik aktivierende Methoden verwendet. Im Falle großer Gruppen bieten sich hier insbesondere gamifizierende Elemente an, die ohne größere Umbrüche in der Konzeption von Veranstaltungen eingesetzt werden können und spielerische Herausforderungen in den Lernansatz einführen. In der vorliegenden Untersuchung wurden strukturell gamifizierende Interventionen in ein Informatikmodul eingeführt und nach Durchführung der Veranstaltungen die Bestehensquoten und studentischen Evaluationen hinsichtlich des Kompetenzerwerbs der Studierenden miteinander verglichen. Erste Ergebnisse der noch laufenden Untersuchung weisen darauf hin, dass die eingesetzten Elemente die Quoten und Evaluationen positiv beeinflussen. Lassen sich die Ergebnisse auch über weitere Vergleichsgruppen bestätigen, sollten die Treiber dieser Veränderungen genau untersucht werden,

Schlagwörter: Gamification, Grundlagenfächer, Ingenieurwissenschaften, Bestehensquote, Kompetenzerwerb

damit die Anwendung dieser Methoden auch für den Einsatz in anderen Grundlagenfächern oder Hochschulengeneralisiert empfohlen werden kann – um damit letztlich zu verminderten Abbruchquoten in den Ingenieurwissenschaften beizutragen.

1. HINTERGRUND

Die Erfahrungen der Hochschulen der vergangenen Jahre zeigen, dass trotz aller Bemühungen die Abbruchquote von ingenieurtechnischen Studiengängen in Deutschland noch immer um die 30% liegt (Heublein & Schmelzer, 2018, S. 10). Bedingende Faktoren sind hierbei unter anderem, dass Eigenheiten der Lehre an der Hochschule auf das aus der Schule gewohnte Lernen sowie neue und schwierige Inhalte auf sehr heterogenes und mitunter zeitlich entferntes Vorwissen der neuen Studierenden prallen (Heublein, Hutzsch, Schreiber, Sommer & Besuch, 2010, 53ff). Daraus entstehende Leistungsprobleme durch Überforderung (25% nach Heublein et al., 2010, S. 168) führen zu niedrigen Bestehensquoten vor allem von Grundlagenveranstaltungen in den ersten Semestern. Dennoch ist das Bestehen von Grundlagenfächern und die Vermittlung eines einheitlichen darin transportierten niveauvollen Wissens essenziell, um überhaupt weitere Inhalte eines Studiums erfolgreich darauf aufbauen zu können. Daher sollten die Inhalte von Modulen und Studiengängen nicht kurzfristig, quasi beliebig inhaltlich adaptiert werden, um damit bestmöglich die individuellen Anforderungen der Lernenden zu befriedigen, sondern vielmehr das

vorhandene Lehr-Lernkonzept so gestaltet werden, dass den Studierenden ein individuelles Lernen im Sinne des „shift from teaching to learning“ ermöglicht wird. Darunter versteht man die Änderung des rein stoffvermittelnden Lernens hin zu kompetenzorientiertem Lernen (Berendt, 1999). Dies wird oft durch den Einsatz aktivierender Lernmethoden versucht, die den persönlichen Interaktionsgrad der Lernenden mit dem Lehrstoff erhöhen und damit rein rezeptives Lernen um individuelles aktives Lernen erweitern. Aktives Lernen dient letztlich dem Aufbau von Handlungs- und metakognitiven Kompetenzen, indem es Ansätze heteronomen und autonomen Lernens nutzt (Bretschneider & Wildt, 2007, 51f). Damit wird im Ergebnis ein lernförderliches Lehren angeboten.

2. BETRACHTETES LEHR-LERNKONZEPT

Im vorliegenden Falle des Moduls „Informatik“ im Grundlagenblock der Ingenieurausbildung wird zwar ein aktivierendes Lehr-Lernkonzept angewendet, dennoch ist die Effektivität des Moduls gemessen an Bestehensquoten und Evaluationsergebnissen der Studierenden noch vergleichsweise gering. Im Modul wird in erster Linie elementares Wissen über Programmierlogik in Form einer Vorlesung im Umfang von 2 SWS vermittelt, in der grundlegende Elemente wie Variablen oder Berechnungen, Algorithmen wie Schleifen oder Bedingungen sowie Konzepte wie Klassen und Objekte vorgestellt werden. In der darauf abgestimmten Übung werden ebenfalls in 2 SWS die vermittelten Lehrinhalte

te durch die Studierenden an einfachen Beispielen aktiv angewandt, um das Verständnis für den Aufbau von Algorithmen und die Nutzung der Elemente und Konzepte zu fördern sowie deren sinnhafte Anwendung aufzuzeigen. Das Modul ist auf einen Teilnehmendenkreis von 100-150 Personen ausgelegt und wird im ersten Fachsemester für Studiengang A und im zweiten Fachsemester für Studiengang B gelehrt. Die korrekte Anwendung der vermittelten Inhalte wird in Form einer Programmierübung in einem Zwischen-Testat abgefragt; das bestandene Testat ist im Studiengang A fakultativ, im Studiengang B Voraussetzung für die Klausurteilnahme. Die Kenntnis und das Verständnis aller vermittelten Elemente und Konzepte wird in der das Modul abschließenden schriftlichen Klausur abgeprüft. Studierende, die die Klausur nicht bestehen, können die Prüfung also im folgenden Semester direkt wiederholen.

Das Modul „Informatik“ wurde über viele Jahre fest in den Curricula der beiden Studiengänge verankert und wird über alle daran Beteiligten effizient und reproduzierbar umgesetzt. Vorlesung und Übung sind nahe an einem etablierten Lehrbuch ausgerichtet, das die Studierenden auch zum Selbststudium und zur Vertiefung nutzen können. Zudem sind die bestehenden Quellen, Unterlagen, Lehrformate und Prüfungsdesigns aufeinander abgestimmt, denn die Studierenden (insbesondere die Wiederholer) verlassen sich auf die Erfahrungen der Vorsemester und Ergebnisse der Probeklausuren. Vor allem aber bleiben damit die Effektivitätsmaße des Moduls mit den in der Vergangenheit erreichten vergleichbar.

Der fachliche Inhalt des Moduls soll – wie in der Einleitung dargestellt – nicht reduziert oder verändert werden, um den Studierenden vermeintlich inhaltlich entgegenzukommen und damit direkt die Bestehensquote zu steuern; das Niveau des Moduls soll weiterhin den erhöhten Ansprüchen einer Hochschulausbildung im Vergleich zur Oberstufe genügen. Im Ergebnis wird also das gesamte Modul auf der Makro-Ebene der Veranstaltungsarten und der Übermittlung des Inhalts genau wie bisher fortgeführt. Im Vergleich zur eher rezeptiven Vorlesung ist die Übung dabei bereits ein eher aktivierender und interaktiver Teil des gesamten Moduls und wird damit weiterhin bewusst als Lehrformat eingesetzt, um das individuelle Lernen zu befördern. Daher wird im Rahmen dieser Forschungsarbeit auch eher der *verstärkte* Einsatz aktivierender Methoden gegenüber der alternativen Einführung komplett *neuer* Methoden angestrebt.

Um den Anteil aktivierender Elemente im vorliegenden Modul also weiter zu steigern, bieten sich vor allem in Grundlagenfächern Konzepte der Gamifizierung an. Gamifizierung bedeutet in diesem Zusammenhang die Anwendung von Elementen des Spieledesigns in einem Nicht-Spiel-Kontext (Deterding, Dixon, Khaled & Nacke, 2011), strukturelle Gamifizierung den nicht auf die Inhalte, sondern nur auf die Form der Darbietung beschränkten Einsatz der Elemente (Kapp, 2013). Die Gamifizierung kann über das Vehikel gesteigerten Engagements, Vergnügens und damit der Motivation (Hamari, Koivisto & Sarsa, 2014, S. 4; Sailer, 2016) den Übergang von schu-

lischem zu akademischem Lernen erleichtern und damit einer o.a. drohenden Überforderung durch den Aufbau einer spielerischen *Herausforderung* entgegenwirken. Dieser Versuch der Beeinflussung der Motivation der Studierenden resultiert aus der Erkenntnis, dass motivationelle Faktoren die größte Gruppe der Treiber für erfolgreiches studentisches Lernen sind (Richardson, Abraham & Bond, 2012, S. 355) und außerdem die direkte, spielerische Wechselwirkung zwischen Stoff und dem Studierenden den Teilnehmenden direktes Feedback gibt, womit ihnen die Wirksamkeit ihrer Interaktionen unmittelbar aufgezeigt wird. Diese „Selbstwirksamkeitsüberzeugung“ wiederum ist tatsächlich die stärkste den Lernerfolg der Studierenden beeinflussende Maßnahme (Schneider & Mustafić, 2015, S. 186).

3. FORSCHUNGSVORHABEN

Ziel dieser Untersuchung ist es, die Auswirkungen strukturell gamifizierender Elemente auf die Effektivität der Veranstaltung insgesamt zu untersuchen, wobei diese anhand der Veränderung der Bestehensquoten und Evaluationsergebnisse im Hinblick auf den Kompetenzerwerb gemessen werden soll.

Das Modul wurde im WS 2018/2019 von mir unverändert vom vorherigen Modulverantwortlichen übernommen und dann erstmals im SS 2019 anhand der Richtlinien von Škuta & Kostolányová, (2018, S. 373) durch die Einführung von bestimmten Interventionen modifiziert, um die in den Rei-

hen der Studierenden vermuteten Spielertypen „Explorer“ (Teilnehmende, die nach Informationen und Hintergründen forschen) und „Achiever“ (Teilnehmende, die sich über Erfolg und Leistung anhand einer externen Herausforderung definieren) anzusprechen (Typologie nach Bartle, 2004). Die verwendeten Interventionen und die damit verfolgten Ziele sind in Tabelle 1 zusammengefasst:

Intervention	Ausprägung	Ziel	verwandte Interventionen
Visible Status	Fortschrittsbalken in Skripten	Erkennen des Fortschritts der Mission	Storyline: Der Weg des Helden durch die Mission
Goals and Challenges	Interaktiver, anonymer Wettbewerb: Verständnisfrage während der Vorlesung	Herausforderung, den Stoff im Laufe der Vorlesung richtig aufzufassen	Time Restriction: Aufbau von Druck als Spielregel zur Erzeugung einer Testsituation
	Interaktive, anonyme Möglichkeit des Feedbacks von noch nicht verstandenen Themen und Behandlung in der Übung	Offene Benennung der Probleme unter dem Schutz der Anonymität und Transport in die Übung zur Vertiefung	Storyline: Die Schlachtbesprechung und das Sammeln von Erfahrung
Freedom to fail	Wechsel der Übungsgruppe	Wechsel zu der Gruppe, bei der das Lernen am leichtesten fällt oder zur Wiederholung	Storyline: Der Held darf auch scheitern und es erneut versuchen

Tabelle 1: Verwendete strukturell gamifizierende Interventionen

Nach der Durchführung der jeweiligen Veranstaltungen in den Semestern wurden die Ergebnisse der studentischen Evaluation und die Anzahl der bestandenen Testate und Klausuren miteinander verglichen. Dies geschieht in zwei Untersuchungsreihen (vgl. Tabelle 2):

1. direkter Vergleich der beiden Kohorten für die zwei Studiengänge A und B jeweils mit und ohne Gamifizierung (Vergleich Gruppe 1 mit Gruppe 3 sowie Gruppe 2 mit Gruppe 4)
2. Vergleich der Wiederholer (Vergleich von Gruppe 2 mit Gruppe 3 und 4), ob eine signifikante Ände-

rung im Lernen mit gamifiziertem LLkonzept oder ohne es erfahren wurde

Im Rahmen dieser Untersuchung liegt also ein quasiexperimentelles Untersuchungsdesign mit Kontrollgruppen im Ex-Post-Facto-Design vor, da die Ursachensuche nach der Durchführung des Experiments durchgeführt wird. Nach Sarris,

(1992, S. 148) sind bei Quasiexperimenten auch insbesondere Sekundärfehlerquellen methodisch nicht befriedigend kontrolliert bzw. kontrollierbar, da der Untersuchungsgegenstand nicht unter reproduzierbaren Laborbedingungen betrachtet wird, sondern im tatsächlich auftretenden Feld, welches je nach Kontrollgruppe

mehr oder weniger in seiner Zusammensetzung variieren kann und damit Verzerrungen in der Auswahl der betrachteten Objekte beinhaltet. So ist auch in diesem Falle der Untersuchungsgegenstand (die einzelnen Kohorten) inhärent heterogen. Diese prinzipielle Heterogenität hinsichtlich

Gruppe	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4
Semester	SS 2018	WS 2018-2019	SS 2019	WS 2019-2020
Studiengang	Studiengang B	Studiengang A	Studiengang B	Studiengang A
Veränderung	ohne Gamifizierung	ohne Gamifizierung	mit Gamifizierung	mit Gamifizierung
Untersuchungsreihe				
Untersuchungsreihe 1			1: Vergleich mit Gruppe 1	1: Vergleich mit Gruppe 2
Untersuchungsreihe 2			2: Vergleich mit Gruppe 2	2: Vergleich mit Gruppe 2

Tabelle 2: Untersuchungsgruppen und -reihen

Gruppengröße, -zusammensetzung und individuellem Wissensstand der Gruppenmitglieder ist zwar nicht antizipierbar oder zu vereinheitlichen, über alle betrachteten Kohorten allerdings immer gleich. So bestehen die Gruppen, die die Veranstaltungen besuchen, regelmäßig aus folgenden Teilnehmendenarten:

1. Reguläre Teilnehmende der jeweiligen Studiengänge A oder B, die das Modul genau in dem im Curriculum vorgesehenen Semester hören und abschließen werden – was der Mehrzahl der Teilnehmenden entspricht
2. Wiederholende der beiden Studiengänge A und B, die die Klausur wegen Nichtbestehens später in ihrem Studium erneut absolvieren müssen und den Stoff daher wiederholen
3. Wiederholende des Studienganges B, die in den Vorsemestern das Testat, nicht aber die Klausur bestanden hatten, die das Modul also ebenfalls nicht in dem im Curriculum vorgesehenen Semester abschließen
4. Andere Teilnehmende, die aus anderen Studiengängen gewechselt sind, bisher also ein komplett anderes Curriculum absolviert haben, und die das Modul mit der Klausur abschließen möchten.

Diese Heterogenitätsaspekte werden im weiteren Verlauf der Untersuchung auch – wo notwendig – für die Kommentierung der Resultate eingesetzt.

4. UNTERSUCHUNGSERGEBNISSE

Im Folgenden werden die Ergebnisse der Auswertung der Klausuren und der studentischen Evaluationen bis zur Gruppe 3 dargestellt, die dieses Semester erstmals die Klausur nach einer gamifizierten Veranstaltung absolviert hat (siehe Tabelle 3 auf der Folgeseite). Gruppe 4 wird die Veranstaltung ab dem kommenden Wintersemester absolvieren; danach wird die Auswertung aktualisiert und ggf. erweitert. Die Daten für die Bestehensquoten wurden für die

Tabelle 3: Untersuchungsergebnisse

Gruppe	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4
Semester	SS 2018	WS 2018-2019	SS 2019	WS 2019-2020
Studiengang	Studiengang B	Studiengang A	Studiengang B	Studiengang A
Veränderung	ohne Gamifizierung	ohne Gamifizierung	mit Gamifizierung	mit Gamifizierung
Bestehensquote				
Teilnehmende Modul	195	94	112	n.v.
Testatanmeldung	99	75	103	n.v.
Testat teilgenommen	99	75	103	n.v.
Testat bestanden ¹	46	28	67	n.v.
Testat nicht bestanden	53	47	36	n.v.
Klausuranmeldung	106	94	83	n.v.
Klausur teilgenommen	87	76	71	n.v.
Klausur bestanden ²	63	60	64	n.v.
Klausur nicht bestanden	24	16	7	n.v.
Testat-Bestehensquote ³	46,46%	37,33%	65,05%	n.v.
Klausur-Bestehensquote ⁴	72,41%	78,95%	90,14%	n.v.
Kombinierte Bestehensquote aus Testat und Klausur ⁵	32,31%	63,83%	57,14%	n.v.

Gruppe	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4
Evaluationen				
Rückmeldungen	51	58	24	n.v.
Rücklaufquote	63%	92%	69%	n.v.
Kompetenzerwerb durch Lernen ⁶	2,8	2,5	2,27	n.v.
Standardabweichung	1,1	0,84	0,83	n.v.
Kompetenzerwerb durch Anwenden ⁷	2,29	2,00	1,65	n.v.
Standardabweichung	0,76	0,62	0,57	n.v.
Wiedererkennung Gamifizierungselemente ⁸	n.v.	n.v.	2	n.v.
Kompetenzerwerb durch Gamifizierung ⁹	n.v.	n.v.	1,25	n.v.
Standardabweichung	n.v.	n.v.	0,5	n.v.

- Das Testat ist bestanden, wenn das erstellte Programm höchstens mit einem unkritischen Fehler ausführbar ist.
- Die Klausur ist bestanden, wenn die Korrektur der Prüfung keine Note „5,0“ ergab.
- Ermittelt aus der Menge der bestandenen Testate in Bezug auf die Menge der Testateilnehmenden.
- Ermittelt aus der Menge der bestandenen Klausuren in Bezug auf die Menge der Klausurteilnehmenden.
- Ermittelt aus der Menge aller bestandenen Klausuren in Bezug auf die Menge der Teilnehmenden der Gruppe.
- Aussage im Fragebogen: „Ich habe in dieser Lehrveranstaltung sehr viel gelernt“ auf einer Skala von 1 – „sehr viel“ bis 5 – „zu wenig“.
- Aussage im Fragebogen: „Aufgrund der Lehrveranstaltung kann ich wichtige Begriffe und Sachverhalte zu diesem Thema wiedergeben bzw. einsetzen“ auf einer Skala von 1 – „stimme voll und ganz zu“ bis 4 – „stimme gar nicht zu“.
- Zeigt an, wie viele Wiederholende die Gamifizierungselemente bewusst wahrgenommen haben.
- Aussage im Fragebogen: „Die Veränderungen halfen mir beim Lernen und Verstehen“ auf einer Skala von 1 – „stimme voll und ganz zu“ bis 4 – „stimme gar nicht zu“.

Gruppe 1 aus historischen anonymisierten Notenlisten des Studierendenservice und Aufzeichnungen der Übungsleiter erhoben, für die beiden anderen betrachteten Gruppen direkt aus den Anmeldezahlen zu den Veranstaltungen in der hochschulweiten Kollaborationsplattform sowie den konkreten Korrekturergebnissen. Nicht vorhandene Daten sind mit „nicht vorhanden“ (n.v.) gekennzeichnet.

5. AUSWERTUNG

Die Darstellung der Daten erfolgt in rein deskriptiver Statistik mit dem Ziel, Trends und erste Erklärungsansätze zu finden, die in weiteren Untersuchungen in der nächsten Kontrollgruppe überprüft und ggf. einer Kausalitätsanalyse zugeführt werden sollen. Der folgende Abschnitt führt zuerst generell die speziellen Umstände ein, die zum Bestehen des Testates und der Klausur in diesem Lernumfeld aus verschiedenen Studiengängen und –semestern führen, und zeigt dann die entsprechenden Erkenntnisse aus den Daten auf. Danach werden je nach Untersuchungsreihe spezifische signifikante Ergebnisse dargestellt.

6. GENERELLE ERKENNTNISSE

Die Gruppen unterscheiden sich zwar erwartungsgemäß nach der Menge der Teilnehmenden, im Verlauf der Veranstaltungen zeigt sich jedoch in jeder Gruppe ein ähnlicher, deutlicher Teilnehmendenschwund; nicht alle Teilnehmenden absolvieren auch die zum Abschluss notwendigen Prüfungen. Die Zahl derer, die sich im Verlauf des Semesters zum Testat anmelden, ist kleiner als die, die sich generell zu Beginn des Semesters für das Modul angemeldet haben, und die Zahl der Klausuranmeldungen übertrifft die Zahl der tatsächlichen Klausurteilnehmerinnen und -teilnehmer. Die Tatsache, dass die Menge der Testatanmeldungen identisch mit der Menge der Testateilnehmenden ist, kann dadurch erklärt werden, dass das Testat als erster, mehr oder weniger folgenloser Versuch

angenommen wird, um die eigene Prüfungsreife festzustellen. Dazu kommt, dass im Studiengang A das Testat fakultativ ist, also auch ein Nichtbestehen kein Hindernis zur Klausurzulassung ist. Der auf den ersten Blick unvermutete Zuwachs zwischen den Testateilnehmenden und den Klausuranmeldungen lässt sich allerdings deutlich mit den Heterogenitätsaspekten 2 bis 4 erklären: Es melden sich hier wahrscheinlich zusätzlich zu den regulären Teilnehmenden Wiederholende aus Vorsemestern an.

Treten Teilnehmende allerdings dann nicht zur der Klausur an, für die sie sich schon angemeldet haben, wird dieser Versuch mit „nicht teilgenommen“ und damit als misslungener Versuch bewertet, tatsächlich absolvierte ungenügende Prüfungen hingegen mit „5.0“. Relevant für die Testat- und Klausurbestehensquoten sind mithin nur die tatsächlich Teilnehmenden, nicht die Anmeldungen, um hier die Gründe auszuschließen, die aus anderen als ungenügenden Prüfungsleistungen resultieren – von plötzlicher Krankheit bis zum Ziehen der sprichwörtlichen „Notbremse“. Um zusätzlich zu analysieren, wie viele Studierende diese Kohorte über das Semester nicht direkt nachvollziehbar verliert, wird eine kombinierte Bestehensquote aus Testat und Klausur angegeben, die die erfolgreiche Klausurteilnahme der ursprünglich angemeldeten Teilnehmendenzahl gegenüberstellt (Abbildung 1). Die Gründe für diesen Verlust über alle Schritte der Veranstaltung sind vielfältig, aber leider nur mutmaßlich: vom Verschieben des Moduls im eigenen Studienplan über die o.a. Überforderung

und daraus folgende Resignation bis ggf. Umorientierung. Hier können vermutlich nur gezielte Evaluationen Erklärungsansätze bieten. Dennoch kann bei allen Gruppen eine wahrnehmbare Tendenz abgelesen werden, die in folgenden Untersuchungen noch geklärt werden muss.

Die Bestehensquoten für das Testat schwanken sehr stark von 46,46% über 37,33% bis 65,05% (Abbildung 1 - Schwankungsbreite rund 27%). Insofern ist zu vermuten, dass hier die Zusammensetzung der Gruppen über die Zeit unterschiedlich ist – insbesondere, dass sich in Gruppe 2 viele reguläre Teilnehmende (Heterogenitätsaspekt 1) befanden, die das Testat zwar nicht erfolgreich, aber folgenlos absolviert haben, und die somit ohne einen überprüfenden Filter oder Erkenntnisse über ihre Prüfungsreife an der Klausur teilgenommen haben. Insofern sind die Testatquoten vermutlich nur zwischen den einzelnen Studiengängen wirklich vergleichbar, da dort zu diesem Zeitpunkt der Veranstaltung ähnliche Gruppenzusammensetzungen vorherrschen.

Die Bestehensquoten für die Klausur sind ähnlich stark schwankend, von 72,41% über 78,95% bis zu 90,14% (Abbildung 1 - Schwankungsbreite rund 17%). Die durchweg zahlenmäßig höheren und weniger schwankenden Werte könnten dadurch erklärt werden, dass zu diesem Zeitpunkt in der Veranstaltung in der Regel nur diejenigen Teilnehmenden auch an der Klausur teilnehmen, die z.B. durch eine (erfolgreiche) Teilnahme am Testat oder ausreichende Mitarbeit und Vorbereitung auch prüfungssicher sind. Dennoch unterschei-

den sich die jeweiligen Differenzen zur Testatquote auch über die Studiengänge. In Studiengang B schließen rund 25% mehr Teilnehmerinnen und Teilnehmer die Klausur ab als Teilnehmerinnen und Teilnehmer das Testat erfolgreich abschließen, in Studiengang A beträgt dieser Überhang Klausur- zu Testatbestehensquote rund 40%. Dies mag damit zu erklären sein, dass nicht alle Teilnehmenden des Studienganges A das Testat überhaupt als Prüfungshürde wahrnehmen und daher die Testatquoten sehr gering ausfallen, s.o.

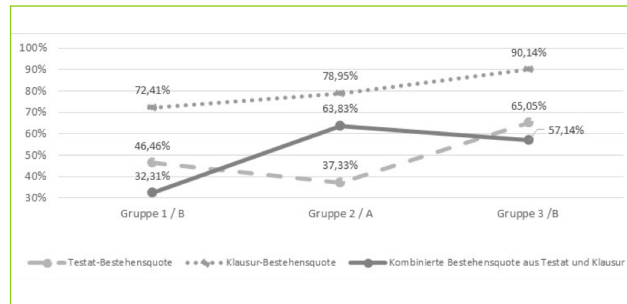


Abbildung 1: Bestehensquoten im Vergleich zu den Gruppen

Im Ergebnis kann also aus dieser Datenlage kein direkter und durchgängig erklärbarer Trend in der *Bestehensquote* über alle Gruppen im Verlauf der Forschung abgeleitet werden.

Der zweite Teil der Betrachtungen der Effektivität der eingeführten Interventionen betrifft die Veränderung des Kompetenzerwerbs. Die Einschätzung der durch das Modul gewonnenen Kompetenzen wird in den studentischen Evaluationen am Ende des Moduls in zwei Standardfragen für alle Teilnehmerinnen und Teilnehmer und zum Zweck der Auswertung der Erfahrungs-

gen und Meinung der Wiederholer und Wiederholerinnen in einer zusätzlichen spezifischen Frage erhoben. Die Rücklaufquoten der Evaluationsergebnisse zeigen, dass wie allgemein üblich nicht alle Teilnehmenden an den Prüfungen auch eine Evaluation abgegeben haben. Unter den 24 Rückmeldungen für Gruppe 3 haben sich jedoch elf Wiederholende gemeldet, was aus den Heterogenitätsaspekten 2 bis 3 zu vermuten war. Drei Teilnehmende stammten aus Studiengang A, 21 aus Studiengang B, was der Annahme aus Heterogenitätsaspekt 1 entspricht. Die Rücklaufquoten betragen außerdem zwischen 63% und 92%, was ausreicht, um den Daten eine gewisse Validität zuzubilligen. Generell sind in allen Datenreihen an den Werten und der Standardabweichung Tendenzen in Richtung besserer und präziserer Bewertungen des empfundenen Kompetenzerwerbs zu erkennen.

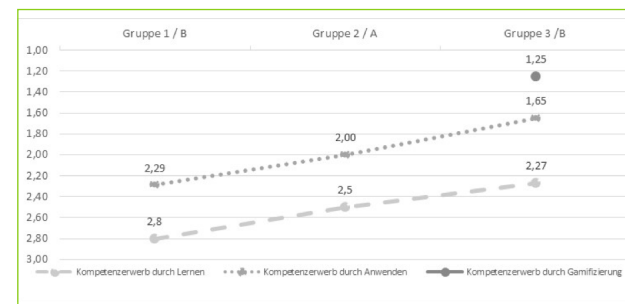


Abbildung 2: Kompetenzerwerb im Vergleich zwischen den Gruppen (Werte von 1-5, siehe Fußnoten 4-9)

Im Ergebnis kann aus dieser Datenlage eine Art Trend im *Kompetenzerwerb* bei allen Gruppen im Verlauf der Forschung abgeleitet werden. Daraus lässt sich aber derzeit keine Aussage herlei-

ten, dass sich für beide Studiengänge gleichermaßen eine Änderung ergeben hätte, die direkt als Ergebnis der Interventionen erkennbar wäre, denn derzeit können noch nicht je zwei Vergleichsgruppen der beiden Studiengänge untersucht werden. Ebenso wären Vermutungen über weitere Kausalitäten verfrüht. Hier bleiben die Ergebnisse der nächsten Kohorte des Studienganges A abzuwarten.

7. UNTERSUCHUNGSREIHE 1: VERGLEICH DER KOHORTEN

Der in dieser Untersuchungsreihe 1 vorgesehene Vergleich von Gruppe 2 mit Gruppe 4 des Studienganges A kann mangels Erhebung noch nicht dargestellt werden. Der Vergleich von Gruppe 1 mit Gruppe 3 dient dazu, festzustellen, ob mit der Einführung der Gamifizierung signifikante Unterschiede in der Bestehensquote und im Kompetenzerwerb der verschiedenen Kohorten des Studienganges B einhergehen. Dies ist zu *bejahen*, wenngleich damit sicherlich noch keine kausale Beziehung abgeleitet werden kann (Tabelle 4). Als Hauptindikator ist hierbei die Klausur-Bestehensquote zu bevorzugen, da sie den ultimativen Abschluss des Moduls aller Teilnehmenden, die die notwendigen Prüfungs Voraussetzungen erreicht haben, kennzeichnet. Da die Einführung der Interventionen für diese Gruppe auch gleichzeitig mit der Änderung des Dozenten für die Vorlesung einherging, ist hier auch ein gewisser Effekt auf die Studentinnen und Studenten durch die Änderung in Art

und Stil des Vortragenden zu vermuten. Die Freitextangaben in der studentischen Evaluation belegen dies tatsächlich – aber insbesondere bei den Fragen, die die Veranstaltung bezüglich der Didaktik und Kommunikation bewerten. Ein direkter Bezug eines Dozentenwechsels zu den Bestehensquoten ist allerdings quasi ausgeschlossen, da die Vorlesung und der komplette Testataufgabenpool sowie die Klausurversionen in ihrem Aufbau, Schwierigkeitsgrad und Bewertungsmaßstäben sehr ähnlich gehalten wurden. In beiden Fällen wurde außerdem eine Durchschnittsnote von 2,7 erzielt, so dass dies ebenso die Vergleichbarkeit der gestellten Anforderungen belegt. Die Änderungen in der selbst wahrgenommenen Bewertung des Kompetenzerwerbs sind außerdem eher individuell-studierendenbezogen zu sehen und damit relativ unabhängig von den Bewertungen, die sich auf die Person des Vortragenden beziehen. Da die Gruppe 3 die erste Gruppe ist, die an einer Veranstaltung mit Gamifizierung teilgenommen hat, sind hier noch keine direkten Vergleichsdaten für einen dementsprechenden Kompetenzerwerb verfügbar. Die zwei Nennungen der Gamifizierungselemente beziehen sich des Weiteren auf die meist sichtbare Intervention, den Fortschrittsbalken.

Damit bleibt abzuwarten, ob sich diese Tendenz bei den folgenden Kohorten verstetigt, was auf die Wirksamkeit der Interventionen, unabhängig von den so sanft wie möglich gestalteten sonstigen Änderungen innerhalb der jeweiligen Studiengänge, hinweisen würde. Dennoch besteht noch keine Klarheit über eine direkte Kausalität.

Gruppe	Gruppe 1	Gruppe 3	Änderung prozentual
Semester	SS 2018	SS 2019	Bei Kompetenzerwerb: Differenz im Wert bezogen auf Skalenendwert
Studiengang	Studiengang B	Studiengang B	
Veränderung	ohne Gamifizierung	mit Gamifizierung	
Testat-Bestehensquote	46,46%	65,05%	+18,59%
Klausur-Bestehensquote	72,41%	90,14%	+17,73%
Kombinierte Bestehensquote aus Testat und Klausur	32,31%	57,14%	+24,83%
Kompetenzerwerb durch Lernen	2,80	2,27	+10,60%
Kompetenzerwerb durch Anwenden	2,29	1,65	+16,00%

Tabelle 4: Veränderung der Bestehensquote und des Kompetenzerwerbs in Studiengang B nach Gamifizierung

UNTERSUCHUNGSREIHE 2: VERGLEICH DER WIEDERHOLERINNEN UND WIEDERHOLER

Wie in der Untersuchungsreihe 1 kann derzeit noch kein Vergleich zu Gruppe 4 hergestellt werden, so dass sich der Vergleich auf Gruppe 2 zu 3 beschränkt. Ziel ist hier festzustellen, ob eine signifikante Änderung im Lernen mit gamifiziertem Lehr-Lernkonzept erfahren wurde oder nicht – vorausgesetzt es sind in der Gruppe genügend Wiederholende, die das beurteilen können. Ausgehend von einem Verhältnis von 11 Wiederholenden zu 24 Evaluierten wird eine Wiederholendenquote von knapp 50% unterstellt, was ausreichend scheint, um prinzipiell valide Ergebnisse insbesondere in den Evaluationsergebnissen darzustellen.

len (Tabelle 5). Dennoch muss erneut die unterschiedliche Heterogenität der Gruppen auch als wahrscheinliche Erklärung für insbesondere die hohen und auch negativen Differenzen in den Quoten erwähnt werden. Die Klausur-Bestehensquote hat sich allerdings prinzipiell ebenso *positiv* geändert. Ebenso ist hier wieder zu bemerken, dass noch nicht abschließend geklärt ist, ob die

Gruppe	Gruppe 2	Gruppe 3	Änderung prozentual
Semester	WS 2018-2019	SS 2019	Bei Kompetenzerwerb: Differenz im Wert bezogen auf Skalenendwert
Studiengang	Studiengang A	Studiengang B	
Veränderung	ohne Gamifizierung	mit Gamifizierung	
Testat-Bestehensquote	37,33%	65,05%	+27,67%
Klausur-Bestehensquote	78,95%	90,14%	+11,19%
Kombinierte Bestehensquote aus Testat und Klausur	63,83%	57,14%	-6,69%
Kompetenzerwerb durch Lernen	2,50	2,27	+4,60%
Kompetenzerwerb durch Anwenden	2,00	1,65	+8,75%

Interventionen die tatsächlichen Treiber der Veränderung sind.

Tabelle 5: Veränderung der Bestehensquote und des Kompetenzerwerbs von Wiederholenden in Gruppe 3 nach Gamifizierung

8. ERGEBNIS

Die Klausurbestehensquote im Vergleich Prä-Post-Gamifizierung ist in den Untersuchungsreihen 1 und 2 gestiegen, und zwar um knapp 10% beziehungsweise 5%. Dieser Unterschied ist auch im Kompetenzerwerb sichtbar, hier beträgt die Verbesserung durchschnittlich knapp 12% beziehungsweise 6%. Der Effekt fällt also im Vergleich zwischen den verschiedenen Kohorten des gleichen Studienganges (Untersuchungsreihe 1) ungefähr doppelt so hoch aus als im Vergleich verschiedener Studiengruppen mit dem einzigen verbindenden Element der Wiederholenden. Im Falle der Untersuchungsreihe 1 ist mithin noch zu prüfen, ob sich diese Resultate auch im Vergleich zwischen den Gruppen 2 und 4 nachweisen las-

sen, was die strukturelle Wirksamkeit der Interventionen auch über die Vergleichsgruppen der unterschiedlichen Studiengänge belegen würde. Im Falle der Untersuchungsreihe 2 werden insbesondere die Urteile der Wiederholer und Wiederholerinnen in Gruppe 4 von Interesse dafür sein, ob sie die Interventionen als hilfreich erkennen und damit der

Kompetenzerwerb steigt. Des Weiteren ist von Interesse, ob sich dann zwischen den Gruppen 2 und 4 ein ähnlicher Trend in den Bestehensquoten wie zwischen den Gruppen 1 und 3 abzeichnet, was wiederum für eine gleichförmige Wirksamkeit in den verschiedenen Studiengängen sprechen würde.

9. KRITISCHE BETRACHTUNG UND AUSBLICK

Obwohl die Untersuchung auf eine längerfristige Überprüfung der Wirksamkeit der eingesetzten Elemente ausgerichtet ist und derzeit noch keine abschließenden Ergebnisse vorliegen, sollten möglicherweise beschränkende Aspekte bereits zu diesem Zeitpunkt diskutiert werden. Je nach Rücklaufquote der studentischen Evaluation und der Qualität der Eingaben kann erstens die Aussagekraft dieser Daten gegebenenfalls beschränkt

sein, zweitens bedingt der Vergleich zweier Gruppen in Untersuchungsreihe 1 immer als grundlegende Annahme, dass der Wissensstand und die Lernwege der Studierenden in unterschiedlichen Kohorten trotzdem annähernd gleich sind, und damit Sekundärfehler auf die Bestehensquoten und den Kompetenzerwerb aus dieser Richtung geringer ausgeprägt sind. Drittens zielt der Vergleich in Untersuchungsreihe 2 zwar auf die gleichen Personen – die Wiederholenden ab, die aber tatsächlich leicht unterschiedliche Lehr-Lernangebote angeboten bekamen, und nach der Änderung einer Veranstaltung im laufenden Lehrbetrieb (und dazu zählt auch die strukturelle Änderung des Lernangebotes) wird oft insgesamt eine geringe Verschlechterung der Evaluationsergebnisse beobachtet, so dass hier mittelfristig auch eine entsprechende Reflektion in den Ergebnissen erwartet wird. Außerdem ist eine Kausalanalyse unbedingt notwendig, um die Treiber der Veränderung in den unzweifelhaft beeinflussten Bestehensquoten und Evaluationen zum Kompetenzerwerb deutlicher zu identifizieren. Des Weiteren könnte auch eine Fortführung der Erhebung der Bestehensquoten und des Kompetenzerwerbs in Erwägung gezogen werden, um die Qualität des angebotenen Moduls zu erheben, zu sichern und damit langfristig die Zufriedenheit der Studierenden mit dem Studienangebot zu erhöhen. Dazu müssten allerdings die verschiedenen Heterogenitätsaspekte reduziert werden, z.B. dadurch, dass auch für den Studiengang A das Testat verpflichtend wird. Damit lassen sich die Differenzen zwischen den einzelnen Bestehensquoten für die verschiedenen Studiengänge nur noch auf den Wissensstand der

Teilnehmenden reduzieren; sie wären dann weniger von taktischen Überlegungen zur Prüfungsvorbereitung einzelner Studiengänge beeinflusst. Ein weiterer Ansatz für künftige sich dem Problem der mangelhaften Bestehensquoten nähernden Forschung wären beispielsweise direkte Evaluationen, die die Gründe des Verlusts von Studierenden über das Semester untersuchen, was sich in der kombinierten Testat-Klausur-Bestehensquote widerspiegelt. Dies würde aber sicherlich einen von dieser Problemstellung getrennten Beitrag erfordern.

LITERATUR

- Bartle, R. A. (2004):** Designing virtual worlds. Indianapolis, Ind: New Riders Pub. Retrieved from <http://proquest.tech.safaribooksonline.de/0131018167>
- Berendt, B. (1999):** Academic Staff Development in Europe. Relevance, Types of Programmes and Suggestions for Discussion. In: Unesco (Hrsg.), World Conference on Higher Education in theXX-Ist Century (IV, S. 30–40). Paris.
- Bretschneider, F. & Wildt, J. (2007):** Handbuch Akkreditierung von Studiengängen. Eine Einführung für Hochschule, Politik und Berufspraxis (GEW-Materialien aus Hochschule und Forschung - Band 110, 1. Aufl.). s.l.: Bertelsmann W. Verlag. Verfügbar unter <http://gbv.eblib.com/patron/Full-Record.aspx?p=1322619>
- Deterding, S.; Dixon, D.; Khaled, R. & Nacke, L. (2011):** From game design elements to gamefulness. In: Lugmayr, A.; Franssila, H.; Safran, C. & Hammouda, I. (eds.), Proceedings of the 15th International Academic MindTrek Conference Envisioning Future Media Environments (S. 9). New York, NY: ACM.
- Hamari, J.; Koivisto, J. & Sarsa, H. (2014):** Does Gamification Work?-A Literature Review of Empirical Studies on Gamification. In: Proceedings of the 47th Hawaii International Conference on System Sciences .
- Heublein, U.; Hutzsch, C.; Schreiber, J.; Sommer, D. & Besuch, G. (2010):** Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen. HIS: Forum Hochschule, (2). Zugriff am 03.07.2019. Verfügbar unter https://www.dzhw.eu/publikationen/index_html
- Heublein, U. & Schmelzer, R. (2018):** Die Entwicklung der Studienabbruchquoten an den deutschen Hochschulen. DZHW-Projektbericht, (Oktober). Zugriff am 03.07.2019. Verfügbar unter https://www.dzhw.eu/publikationen/index_html
- Kapp, K. M. (2013):** Two Types of Gamification. Zugriff am 04.09.2018. Verfügbar unter <http://karlkapp.com/two-types-of-gamification/>
- Richardson, M.; Abraham, C. & Bond, R. (2012):** Psychological correlates of university students' academic performance: a systematic review and meta-analysis. Psychological Bulletin, 138(2), 353–387. <https://doi.org/10.1037/a0026838>.
- Sailer, M. (2016):** Die Wirkung von Gamification auf Motivation und Leistung. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-14309-1>
- Sarris, V. (1992):** Versuchsplanung und Stadien des psychologischen Experiments (UTB Große Reihe, Lehrbuch in 2 Bd. für Studenten der Psychologie, Medizin und Pädagogik mit praktischen Demonstrationsbeispielen /Viktor Sarris ; 2). München: UTB /BRO.
- Schneider, M. & Mustafić, M. (2015):** Gute Hochschullehre: Eine evidenzbasierte Orientierungshilfe. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-45062-8>
- Škuta, P. & Kostolányová, K. (2018):** Adaptive Approach to the Gamification in Education. DIVAI 2018, 367–377.