

DANIEL KAIMANN

Benutzerfreundlichkeit von Lernüberprüfungen - Ein Vergleich von papierbasierten und elektronischen Prüfungen

Zusammenfassung

Lernprüfungen spielen im Lern-/Lehrprozess des Constructive Alignments eine entscheidende Rolle. Dabei besteht die Aufgabe Prüfungen kompetenzorientiert und integrativ in den Lern-/Lehrprozess einzubinden. Die vorliegende Studie stellt einen Vergleich der Benutzerfreundlichkeit von etablierten papierbasierten und aufkommenden elektronischen Prüfungen an. Mit Hilfe einer Befragung von 142 Studierenden und der Berechnung eines Benutzerfreundlichkeitswertes anhand des System Usability Scores (SUS), wird dem elektronischen gegenüber dem papierbasierten Prüfungssystem eine um 8% höher empfundene Benutzerfreundlichkeit nachgewiesen. Die Benutzerfreundlichkeit wird vor allem durch geschlechtsspezifische Unterschiede und Erfahrungen mit vorherigen Prüfungsformen getrieben. Insbesondere Männer zeigen eine höhere Akzeptanz von elektronischen Prüfungen. Außerdem spielen Technikaffinität und Anzahl der Fachsemester eine Rolle bei der subjektiven Bewertung der Benutzerfreundlichkeit von Lernüberprüfungssystemen.

1 Einleitung

Der Lern-/Lehrprozess schließt mit der Überprüfung der erreichten Lernziele ab. Im Zuge des Constructive Alignment sind auch Prüfungen kompetenzorientiert auszurichten und zu stellen (siehe Abbildung 1). Insbesondere Studierende verlagern ihren Lehrschwerpunkt auf die Vorbereitung der Abschlussprüfung. Daher sollten Lernüberprüfungen bereits im Anfangsstadium der Lern-/Lehrprozesse so konzipiert werden, dass die erwünschten Lernergebnisse von den Studierenden erreicht und sinnvoll festgestellt werden können (Schermutzki 2008). Die Lernüberprüfungen sind durch einen überwiegend summativen und obligatorischen Charakter geprägt, die meist zur Erlangung eines Qualifikationsnachweises vorgeschrieben sind. Dennoch sollten sie, neben der Gewährleistung eines formal erforderlichen Leistungsnachweises, Lernenden und Lehrenden die Möglichkeit zum Feedback geben. Durch das gewonnene Feedback können Lernende und Lehrende ihre eigenen Lern-/Lehrprozesse weiter optimieren. Die Abfrage und das Erreichen einer oder mehrerer Taxonomiestufen von Lernzielen nach Bloom (1976) ist unabhängig von der eingesetzten Prüfungsform und wird durch die kompetenzorientierte Gestaltung von Lernüberprüfungsaufgaben bestimmt.

Gerade in Massenveranstaltungen ist es nicht nur schwierig, eine kompetenzorientierte Lehr-/Lernbeziehung aufzubauen, sondern die gelernten Fähigkeiten auch kompetenzorientiert zu prüfen. Aufgrund der Anzahl der beteiligten Studierenden an diesen Veranstaltungen kommen meist Multiple-Choice-Prüfungen zum Einsatz. Diese Art der Prüfung stellt an die Lehrenden und Dozenten nicht nur ein hohes Maß an konzeptioneller Kompetenz, sondern auch hohe Anforderungen und Schwierigkeiten an die Durchführung und Auswertung der Lernüberprüfungen. Außerdem besteht bei Multiple-Choice-Prüfungen die Schwierigkeit des kompetenzorientierten Prüfens, da Multiple-Choice-Klausuren insbesondere die WissensEbene bedienen.

Damit Multiple-Choice-Prüfungen Bestandteil des Constructive Alignment sein können und die Abfrage von Fachkompetenzen ermöglichen, dürfen sie keine Methodenkompetenzanforderungen an Studierende stellen, die in vorherigen Lehrveranstaltungen weder gelehrt, noch geübt oder aufgezeigt worden sind. Bisherig eingesetzte papierbasierte Prüfungstechnologien, wie zum Beispiel das „Scantron-Verfahren“, weisen hier einige Schwierigkeiten und Schwächen auf. Damit stellt alleine der Einsatz der Abfragetechnologie ein erhöhtes Maß an Methodenkompetenzfähigkeiten, also den Fähigkeiten, die zur Anwendung bestimmter Arbeitsmethoden von Nöten sind, an die Studierenden, da es ein „früher gelerntes Lösungsmodell“ nicht gegeben hat.

Eine Abhilfe könnten elektronische Prüfungen geben. Zu elektronischen Prüfungen, oder kurz ePrüfungen, zählen Assessments, Testate oder Klausuren, die nicht papierbasiert, sondern am Bildschirm bearbeitet und mit Hilfe von Rechnern ausgewertet werden. Kern der ePrüfung bildet ein elektronisches Prüfungssystem. Das Prüfungssystem muss dabei den gesamten Prüfungsprozess, von der Erstellung von Prüfungsfragen bis hin zur automatisierten Auswertung und Benotung der Prüfung, abdecken. Bunderson, Inouye und Olsen kündigten bereits 1989 eine Revolution der Leistungsbewertung durch den Einsatz von computergestützten Tests im Bildungskontext an. Diese Revolution wird in vier verschiedenen Generationen computergestützter Messinstrumente und -prozeduren beschrieben, wobei jede eine Erhöhung an Raffinesse und Leistung repräsentiert. In der ersten Generation werden konventionelle Prüfungen am Computer gestellt. Bei der zweiten Generation handelt es sich um computerbasierte anpassungsfähige Prüfungen, d.h. die Schwierigkeit oder der Inhalt der nächsten Aufgaben oder die Zeitsteuerung können auf die vorherige Antwort des Prüflings zugeschnitten werden. Die dritte Generation weist fortlaufende Messungen auf. Es werden kalibrierte Maßzahlen in den Lehrplan eingebettet, um kontinuierlich und unauffällig dynamische Veränderungen in der Leistungskurve und dem Lernprofil des Prüflings zu messen. Damit stellt die dritte Generation die Integration von elektronischen Lernüberprüfungen in den integrativen Lernprozess des Constructive Alignment dar. Die vierte Generation wird als intelligente Messung definiert, d.h. das mit Hilfe von deduktiven Prozeduren und Informationsdatenbanken Erkenntnisse über die individuellen Lernprofile der Prüflinge gewonnen werden können. Gerade im Hinblick auf das angestrebte kompetenzorientierte

Prüfen bieten computerbasierte Prüfungen Möglichkeiten, die über die der papierbasierten Prüfungsformen hinausgehen. Zum einen kann die Validität der elektronischen Prüfungen höher sein als die der Papierprüfungen. Zum anderen erlauben interaktive Elemente und das Einfügen von Rich-Media-Dateien (z.B. Videos oder Animationen) weiterführende Kompetenzabfragen. Wenn der Sprung von den heutigen konventionellen Tests am Computer in die nächsten Generationen gelingt, entstehen Chancen, mit denen Prüflinge effektiver und aufwandärmer kompetenzorientiert geprüft werden können.

Ehlers et al. (2009) untersuchten in ihrer Studie, durchgeführt von 2005 bis 2008 an der Tierärztlichen Hochschule Hannover, die Akzeptanz und Qualität der durchgeführten elektronischen Prüfungen mit Hilfe einer Umfrage von 658 Teilnehmer(innen). Die Befragten zeigten mit 86,7% eine hohe Akzeptanz gegenüber den elektronischen Prüfungen. Außerdem wurde vor allem die Benutzerfreundlichkeit¹ als hoch bewertet. Ein Zustand, der insbesondere auf die ausgeprägte private Nutzung von elektronischen Geräten zurückzuführen ist. Glowalla et al. (2009) untersuchten in drei Kursevaluationen, durchgeführt von WS 2003/2004 bis WS 2004/2005 im Radiologiekurs der Universität Marburg, unter anderem die Akzeptanz der zuvor durchgeführten elektronischen Prüfungen mit Hilfe einer Umfrage von insgesamt 297 Teilnehmer(innen). Der Großteil der Studierenden hatte kaum Schwierigkeiten sich auf die elektronische Klausur umzustellen und bereits 96% gaben an, sich an die neue Prüfungsform gewöhnt zu haben. Dieses Ergebnis ist auf die Vertrautheit zurückzuführen, die sie bereits durch die intensive Nutzung der universitären Online-Lernplattformen erlangt haben. Krieschke (2005) führte im Rahmen des Pilotprojektes „eExamination“ bereits 1300 ePrüfungen an der Universität Bremen durch. Insbesondere in den Bereichen Gleichberechtigung, Zeit, Aufwand und Niveau der Prüfungsaufgaben konnte man bei elektronischen Prüfungen eine deutliche Effizienzsteigerung erkennen.

Die vorliegende Studie stellt im Gegensatz zu vorangehenden Studien eine Vergleichsstudie dar und untersucht die Benutzerfreundlichkeit von papierbasierten und elektronischen Prüfungen mit Hilfe einer Studierendenbefragung. Die Befragung wurde im Sommersemester 2013 an der Universität Paderborn an der Fakultät Wirtschaftswissenschaften in der Veranstaltung „Grundzüge der VWL“ durchgeführt. Ziel der Befragung ist die Beantwortung folgender Hypothesen:

H₁: Ein elektronisches Prüfungssystem ist benutzerfreundlicher als ein papierbasiertes Lernüberprüfungssystem.

H₂: Ein elektronisches Prüfungssystem unterstützt kompetenzorientiertes Prüfen besser als ein papierbasiertes Lernüberprüfungssystem.

¹ Die DIN-Norm EN ISO 9241-11 definiert die Benutzerfreundlichkeit über die drei Dimensionen Effektivität, Effizienz und Zufriedenheit. Dabei ist vor allem das subjektive Kriterium der Zufriedenheit als wesentliche Komponente der Benutzerfreundlichkeit anzusehen. Zufriedenheit ist das subjektive Empfinden eines Anwenders bei der Benutzung einer Technologie. Dabei sind insbesondere die Dimensionen "Schwierige Bedienbarkeit" und "Einfache Bedienbarkeit" zu unterscheiden.

Hypothese H_1 untersucht neben dem Vergleich der Benutzerfreundlichkeit beider Technologien auch die nachhaltige Verbesserung der Bedienbarkeit von Prüfungen. Hypothese H_2 geht der Frage nach, ob elektronische Prüfungen besser als papierbasierte Prüfungen geeignet sind, die Abfrage von Fach- und Methodenkompetenzen zu gewährleisten. Außerdem ist der Frage nachzugehen, ob die Validität der elektronischen Multiple-Choice-Prüfungen höher ist als die der papierbasierten Multiple-Choice-Klausuren, da das Ausfüllen der Papierprüfungen zusätzliche Kompetenzen an die Studierende stellt. Damit könnten elektronische Prüfungen integrativer Bestandteil des Constructive Alignment werden.

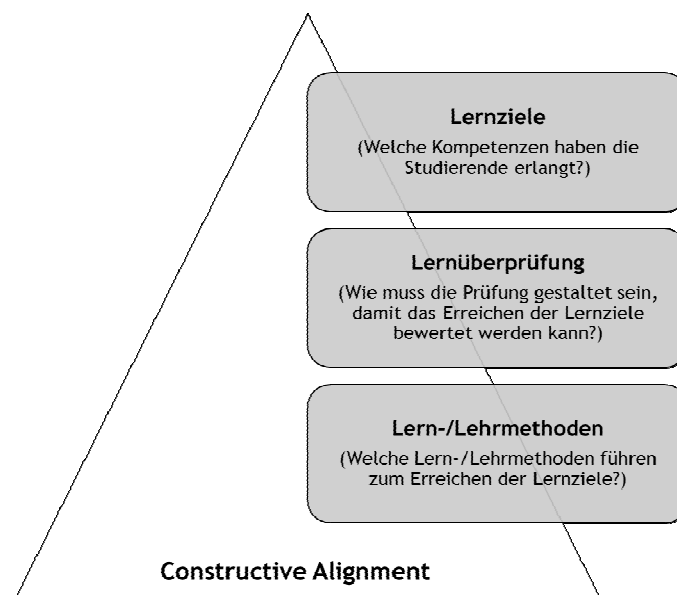


Abb. 1: Integrativer Prozess des Constructive Alignments

2 Methodik

Die Evaluierung der beiden Lernüberprüfungssysteme erfolgte mithilfe der System Usability Scale (SUS) während des Sommersemesters 2013 (April 2013 bis September 2013). Diese Evaluierungsmethode wurde bereits von Studien, die die Benutzerfreundlichkeit von eLearning Systemen untersucht haben, eingesetzt (Kundisch et al. 2012). Die System Usability Scale ist eine sogenannte Quick-and-Dirty-Methode zur Messung der Benutzerfreundlichkeit einer Anwendung oder Technologie, die 1986 von Brooke entwickelt wurde. Sie gehört zu den etablierten und anerkannten Erhebungsverfahren in der quantitativen Nutzerforschung. Die System Usability Scale befragt Nutzer zur der Benutzerfreundlichkeit einer Anwendung und stellt das Ergebnis als prozentualen Benutzerfreundlichkeitswert (oder Usability Score) der Applikation dar.

Der Fragebogen besteht aus insgesamt 10 Fragen, die mit Hilfe einer Likert-Skala mit fünf Optionswerten bewertet werden können. Der System Usability Scale Fragebogen besteht dabei aus fünf negativ und fünf positiv formulierten Aussagen, die

die Benutzerfreundlichkeit der zu bewertenden Applikation oder Technologie erfassen. Brooke gibt folgende zehn Fragetypen vor, die in Ihrer Ausgestaltung für die Befragung und eingesetzte Prüfungsform angepasst und durch offene Fragen ergänzt wurden:

Fragen zur Benutzerfreundlichkeit von Lernüberprüfungssystemen

- Ich denke, ich würde das Prüfungssystem gern häufig benutzen.
- Ich fand das Prüfungssystem unnötig komplex.
- Ich fand, das Prüfungssystem war einfach zu benutzen.
- Ich denke, ich bräuchte eine Einweisung, um das Prüfungssystem zu benutzen.
- Ich fand die verschiedenen Funktionen des Prüfungssystems gut integriert.
- Ich fand das Prüfungssystem zu inkonsistent.
- Ich denke, die meisten Menschen würden den Umgang mit dem Prüfungssystem schnell erlernen.
- Ich fand das Prüfungssystem sehr umständlich zu benutzen.
- Ich fühlte mich sehr sicher im Umgang mit dem Prüfungssystem.
- Ich musste viele Dinge lernen, bevor ich mit dem Prüfungssystem umgehen konnte.

Offene Fragen

- Bitte beschreiben Sie kurz mit Ihren eigenen Worten, was Ihnen bei den Lernüberprüfungen gut gefallen hat.
- Was könnte man Ihrer Meinung nach verbessern?

Aus allen Antworten und Ergebnissen des Fragebogens errechnet sich ein sogenannter System Usability Score, der eine Ausprägung zwischen 0 (schlechteste Benutzerfreundlichkeit des Prüfungssystems) und 100 (beste Benutzerfreundlichkeit des Prüfungssystems) annehmen kann. Der System Usability Score ist ein Durchschnittswert, der als Prozentwert interpretiert werden kann. Demnach entsprechen 100% einem perfekten System ohne Benutzerfreundlichkeitsprobleme, ein Wert über 80% deutet auf ein System mit guter bis exzellenter Benutzerfreundlichkeit hin, Werte zwischen 60 und 80% sind als grenzwertig bis gut zu interpretieren und ein Benutzerfreundlichkeitswert unter 60% weist auf erhebliche Probleme in der Bedienung des Prüfungssystems hin.

Der System Usability Score ist als Tendenz zu verstehen. Falls eine Benutzerfreundlichkeitsumfrage einen System Usability Score von 82% ergeben hat und eine andere Umfrage zu einem ähnlichen System einen Wert von 85% als Ergebnis aufweist, dann sind diese beiden Werte nicht ganz trennscharf voneinander zu interpretieren. Ein Score von 82% im Vergleich zu einem Usability Score von 54% zeigt aber dennoch eine deutliche Varianz in der Nutzbarkeit der vergleichenden Systeme.

Die Umfrage wurde an der Universität Paderborn an der Fakultät für Wirtschaftswissenschaften während des Sommersemesters 2013 durchgeführt. Dafür eignete sich insbesondere die Veranstaltung „Grundzüge der Volkswirtschaftslehre“. Bei

dem Modul „Grundzüge der Volkswirtschaftslehre“ handelt es sich um eine Pflichtveranstaltung in der Einführungsphase der Studiengänge Wirtschaftswissenschaften, International Business Studies sowie der Wirtschaftsinformatik als auch der Wirtschaftsingenieurwissenschaften. Die Teilnehmerzahlen der Veranstaltung liegen in der Regel zwischen 1000 und 1200 Studierenden. Alle Teilnehmer schließen das Modul mit einer Multiple-Choice-Klausur ab.

Ich erwarte für die Untersuchung valide Ergebnisse, da durch die betrachtete Stichprobe eine ausreichend große und heterogene Gruppe abgedeckt wird. Außerdem erwarte ich einen höheren System Usability Score bei den elektronischen Prüfungen, da diese Benutzertechnologie bei den Studierenden vertrauter ist und einfacher erlernbar.

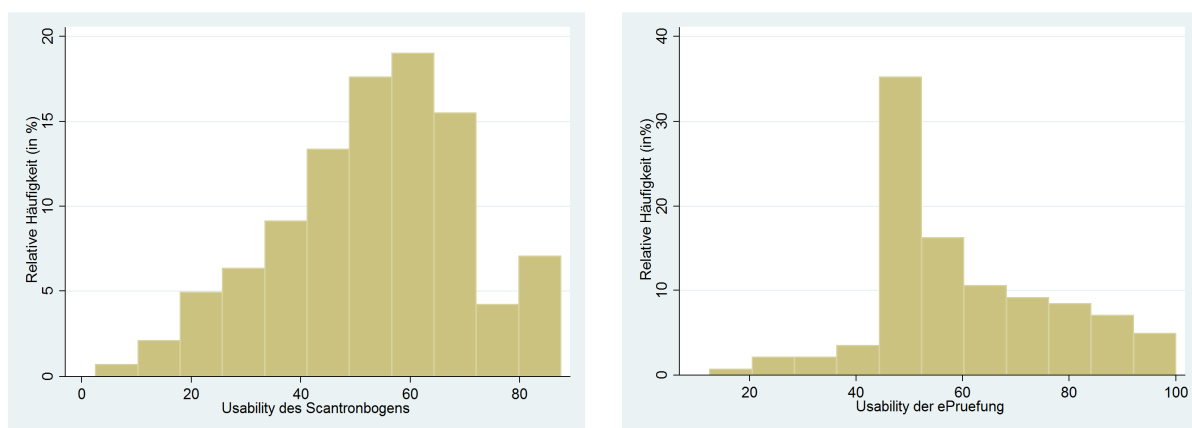
3 Empirische Ergebnisse

Im Rahmen der Befragung zur Benutzerfreundlichkeit von Lernüberprüfungen wurden im Mai 2013 in der Veranstaltung „Grundzüge der VWL“ eine Zufallsstichprobe von 142 Studierenden befragt. Um den Effekt der Prüfungstechnologie zwischen papierbasierten und elektronischen Klausuren zu selektieren, wurden den Studierenden dieselben Prüfungsfragen gestellt. Zur Auswahl standen dabei zwei Prüfungsverfahren. Die erste Methode wird durch das sogenannte „Scantron-Verfahren“ repräsentiert. Dabei müssen die Studierenden die richtigen Antworten auf einem Lösungsbogen markieren (siehe Anhang A1: Muster eines Scantron-Bogens). Dieses Verfahren weist allerdings einige Schwierigkeiten und Schwächen auf: Das Ausfüllen der Scantron-Bögen darf lediglich mit Bleistift und ohne jegliche Radierung erfolgen. Damit stellt das Scantron-Verfahren an die Studierenden Methodenkompetenzanforderungen, die in vorherigen Lehrveranstaltungen weder gelehrt, noch geübt oder aufgezeigt wurden. Nach dem Lernzielstufenmodell nach Bloom (1976) befinden wir uns damit bereits auf der dritten, der Anwendungsstufe. In der Anwendungsstufe wendet ein Studierender das bereits Gelernte in einer neuen Situation an. Das Problem im vorliegenden Prüfungssystem liegt allerdings darin, dass es das „früher gelernte Lösungsmodell“ nicht gegeben hat. Damit stellt alleine der Einsatz der Abfragetechnologie ein erhöhtes Maß an Kompetenzfähigkeiten an die Studierenden dar. Die zweite Prüfungsmethode repräsentiert die elektronische Prüfung. Dabei müssen alle Aufgaben am Bildschirm bearbeitet werden. Die Studierenden können dafür ihre eigenen mobilen Endgeräte, wie z.B. Notebooks, Tablets oder auch Smartphones benutzen.

36% der Studierenden gaben an ein Notebook zu besitzen, 14% der Befragten besitzen ein Tablet und 44% der Studierenden haben ein Smartphone. Außerdem sind 63% der Befragten weiblich und 37% der beteiligten Studierenden männlich. Bei den Studiengängen gibt es einen Schwerpunkt auf den Bachelorstudiengang Wirtschaftswissenschaften (56%), gefolgt von B.Sc. International Business Studies (38%) und B.Sc. Wirtschaftsingenieurwesen (5%). 59% der Studierenden gaben an, bereits

Erfahrung mit dem Scantron-Verfahren zu haben. Elektronische Prüfungen wurden bisher von 1% der Befragten genutzt.

Aus allen Antworten der Studierenden errechnet sich ein System Usability Score von 52,62% für papierbasierte und ein System Usability Score von 60,35% für elektronische Prüfungen. Damit wird diese elektronische Prüfung tendenziell von den Studierenden benutzerfreundlicher als papierbasierte Scantron-Prüfung eingeschätzt. Elektronische Prüfungen können mit einem Benutzerfreundlichkeitswert von über 60% noch als gut interpretiert, wohingegen die Benutzerfreundlichkeit von papierbasierten Prüfungen insgesamt als grenzwertig bewertet werden muss. Abbildung 2 illustriert die SUS-Werte der beiden Lernüberprüfungsmethoden und stellt diese vergleichend dar.



SUS des Scantron-Verfahrens: 52,62%
(Standardabweichung: 16,78)

SUS der ePrüfung: 60,35%
(Standardabweichung: 16,93)

Abb. 2: System Usability Score von papierbasierten und elektronischen Prüfungen

Die Korrelationskoeffizienten geben genauere Hinweise auf den Zusammenhang von Einflussfaktoren und Benutzerfreundlichkeitswerten. Insbesondere ist der Einfluss des Alters, des Geschlechtes, die Anzahl der Fachsemester sowie die Erfahrung mit mobilen Endgeräten und verschiedenen Prüfungsformen von Interesse. Tabelle 1 zeigt die Korrelation dieser Variablen mit den SUS von Scantron und elektronischen Prüfungen. Dabei weisen vor allem die Variablen „Männlich“ und „Prüfungsform“ eine hohe Korrelation mit der Benutzerfreundlichkeit von elektronischen Prüfungen auf. Die anderen weisen weder hohe noch signifikante Korrelationswerte mit den System Usability Scales des Scantron-Verfahrens oder der elektronischen Prüfungen auf.

	SUS Scantron-Verfahren	SUS ePrüfung
SUS Scantron-Verfahren	1	0.06 (0.51)
SUS ePrüfung	0.06 (0.51)	1
Alter	0.03 (0.72)	0.06 (0.49)
Männlich	0.02 (0.85)	0.14 (0.10)
Fachsemester	-0.04 (0.62)	0.04 (0.60)
Mobiles Endgerät	0.01 (0.92)	0.07 (0.43)
Prüfungsform	-0.01 (0.94)	0.11 (0.18)

Tabelle 1: Korrelationen der Einflussvariablen mit dem SUS des Scantron-Verfahrens und der ePrüfung (in den Klammern befindet sich das Signifikanzniveau einer jeden Korrelation)

Um ein detaillierteres Bild über den Zusammenhang des Einflusses der beschriebenen Variablen auf die Benutzerfreundlichkeit der papierbasierten und elektronischen Prüfung zu erhalten, führe ich in einem weiteren Schritt die linearen Regressionen für diese Variablen durch. Die Usability Scores der papierbasierten und elektronischen Prüfungen bilden dabei die abhängigen Variablen und die Variablen Alter, Männlich, Fachsemester, Mobiles Endgerät und Prüfungsform die unabhängigen Variablen. Die Ergebnisse der beiden linearen Regressionen sind in Tabelle 2 illustriert.

Die Regressionen bestätigen den ersten Eindruck der Korrelationskoeffizienten und zeigen signifikante Zusammenhänge zwischen den Faktoren Männlich und Prüfungsform und dem System Usability Score einer elektronischen Prüfung. Demnach steigt der abgefragte Benutzerfreundlichkeitswert einer elektronischen Prüfung um 5%, wenn die Erfahrung mit Prüfungsformen im Allgemeinen oder die Anzahl der männlichen Teilnehmer um 1% steigt. Diese positiven Zusammenhänge könnten insbesondere auf zwei Einflüssen beruhen. Anekdotische Evidenz bezeichnet Männer gemeinhin technikaffiner als Frauen. Durch ihre Erfahrungen mit technischen Systemen und ihre Affinität für technische Systeme bewerten sie elektronische Prüfungssysteme im Durchschnitt besser als papierbasierte Klausuren. Zum anderen steigt die empfundene Akzeptanz von elektronischen Prüfungen, wenn Studierende bereits Erfahrungen mit anderen Prüfungsformen haben. Außerdem zeigen die Faktoren Alter und Mobiles Endgerät einen positiven Zusammenhang mit dem Usability Score von elektronischen Prüfungen. Ein weiterer Hinweis, dass sich ein höherer Erfahrungswert und Lernkurve von Studierenden positiv auf die Bewertung der Benutzerfreundlichkeit von elektronischen Prüfungen auswirkt. Dass Erfahrung auch einen negativen Zusammenhang mit dem SUS von elektronischen Prüfungen haben kann, zeigt der Faktor Fachsemester. Hier zeigt sich, dass die Erfahrungen aus vergangenen Studiensemestern und die Probleme bei vorherigen Prüfungen zu einem Frustrationslevel aufgebaut werden können, die sich mit jedem weiteren Semester negativ auf Bewertung von ePrüfungen auswirken. Allerdings ist darauf hinzuweisen, dass die Faktoren Alter, Fachsemester und Mobiles Endgerät keine signifikanten Werte zeigen.

Die Regression des System Usability Scores des Scantron-Verfahrens weist ähnliche Zusammenhänge der unabhängigen Variablen mit Benutzerfreundlichkeitwert der papierbasierten Prüfung auf, wie die Zusammenhänge der Faktoren mit dem SUS der elektronischen Prüfung. Allerdings besteht ein negativer Zusammenhang zwischen der Erfahrung mit vorherigen Prüfungsformen und der empfundenen Benutzerfreundlichkeit von papierbasierten Prüfungen. Dies ist, wie zuvor bereits beschrieben, insbesondere mit einem höheren Frustrationslevel der Studierenden mit zuvor absolvierten Prüfungen zu erklären. Allerdings sind jegliche unabhängigen Variablen der SUS Scantron-Verfahren Regression nicht signifikant. Weitere Untersuchungen und Befragungen, die zu einer Erhöhung der untersuchten Stichprobe führen, sollten die Validität der Ergebnisse in zukünftigen Studien verbessern.

Unabhängige Variablen	SUS Scantron-Verfahren	SUS ePrüfung
Alter	.3204 (1.051)	.5580 (1.021)
Männlich	2.128 (3.066)	5.962** (2.978)
Fachsemester	-.4351 (1.989)	-1.347 (1.933)
Mobiles Endgerät	2.663 (3.979)	2.158 (3.865)
Prüfungsform	-.0503 (3.006)	5.179* (2.920)
N	128	128
R²	0.01	0.06

*Tabelle 2: Regression der Benutzerfreundlichkeit einer papierbasierten und elektronischen Prüfung (***, ** und * stehen für ein Signifikanzniveau von 1%, 5% und 10%)*

4 Zusammenfassung und Diskussion

Der Vergleich der Benutzerfreundlichkeit von papierbasierten und elektronischen Prüfungen hat durch eine Befragung eine als höher empfundene Benutzerfreundlichkeit von elektronischen gegenüber papierbasierten Prüfungen auf Seiten der Studierenden gezeigt. Dabei wurde mit Hilfe des System Usability Scores, einem etablierten und anerkannten Erhebungsverfahren in der quantitativen Nutzerforschung, die Benutzerfreundlichkeit von beiden Prüfungssystemen gemessen. Die Befragung von 142 Studierenden ergab eine höhere Benutzerfreundlichkeit des elektronischen gegenüber dem papierbasierten Prüfungssystem um ca. 8%.

Diese Zufriedenheit seitens der Studierenden gegenüber elektronischen Prüfungen ist auf mehrere Gesichtspunkte zurückzuführen. Die Vertrautheit mit elektronischen Endgeräten und die Anlehnung des „Looks and Feels“ der ePrüfung an bereits

eingesetzte Systeme und Applikationen trägt unter anderem zu einem hohen Benutzerfreundlichkeitsempfinden bei. Dieser Zusammenhang ist vor allem bei den männlichen Testpersonen zu beobachten. Außerdem könnte der Neuheitswert der elektronischen Prüfungen für ein besseres Image gesorgt haben. Die Erfahrungen mit anderen Prüfungsformen dürfte ein weiterer Gesichtspunkt zum guten Abschneiden der elektronischen gegenüber papierbasierten Prüfungen in den Bewertungen der Studierenden sein. Sieht man die Benutzerfreundlichkeit des Scantron-Verfahrens als Nullgrenze an, dann können Verbesserungen in diesem Verfahren nur zu höheren Benutzerfreundlichkeitsempfinden führen.

Die vorliegenden Ergebnisse liefern erste Anhaltspunkte, dass elektronische Prüfungen kompetenzorientiertes Prüfen besser als andere Prüfungsformen abbilden können. Papierbasierte Prüfungen, wie zum Beispiel Scantronbögen, erfordern oft eine zusätzliche Kompetenz von den Studierenden und können nicht durch Rich-Media-Formate erweitert werden. Dieser empirische Hauptbefund sollte in zukünftigen Untersuchungen, bspw. durch T-Tests für unabhängige Stichproben, ergänzt werden, um zu schauen, inwieweit sich die befragten Gruppen voneinander unterscheiden. Außerdem könnten 2-Faktorielle ANOVA Analysen dazu beitragen, den Befund auch für Subgruppen, wie zum Beispiel technikaffine Studierende oder verschiedene Altersgruppen, zu belegen.

Nun stellt sich unweigerlich die Frage, mit welchem Technologieeinsatz Multiple-Choice-Klausuren in Zukunft gestellt werden sollten. Elektronische Prüfungen stellen aufgrund mehrerer Faktoren eine nennenswerte Alternative dar. Die vorliegende Studie belegt eine gute Benutzerfreundlichkeit seitens der Studierende. Mit Hilfe von Rich-Media-Formaten lassen sich auch interaktive Frageelemente integrieren und sogar Freitextaufgaben stellen. Damit können Fachkompetenzen außerhalb der Wissensebene abgefragt werden. Außerdem sinkt durch den Einsatz von ePrüfungen die Korrekturzeit, was wiederum zu einer Verbesserung der Studiensituation für Studierende führt. Daher sollte der stärkere Einsatz von elektronischen Prüfungen, insbesondere in Massenveranstaltungen und Multiple-Choice-Klausuren, diskutiert und angestoßen werden.

Danksagung

Ich danke insbesondere Nadja Maraun für Ihre Unterstützung bei der Datenerhebung.

Literatur

- Bloom, B.S. (1976): Taxonomie von Lernzielen im kognitiven Bereich. Weinheim, Basel: Beltz.
- Brooke, J. (1996): SUS - A quick and dirty usability scale. In Jordan, Thomas, Weerdmeester and McClelland. Usability Evaluation in Industry. London: Taylor and Francis.
- Bunderson, V. C., Inouye, D. K., & Olsen, J. B. (1989). The four generations of computerized educational measurement. In R. L. Linn (Ed.), Educational measurement: Third edition (pp. 367- 407). New York: Macmillan.
- Ehlers, J. P., Carl, T., Windt, K-H., Möbs, D., Rehage, J., Tipold, A. (2009): Blended Assessment: Mündliche und elektronische Prüfungen im klinischen Kontext. Zeitschrift für Hochschulentwicklung, 4 (3).
- Glowalla, U., Schneider, S., Siegert, M., Gotthardt, M., Koolman, J. (2005): Einsatz wissensdiagnostischer Module in elektronischen Prüfungen. In Proceeding of DeLFI (2005): 3. Deutsche e-Learning Fachtagung Informatik, der Gesellschaft für Informatik e.V. (GI) 13.-16. September 2005, Rostock.
- Kundisch, D., Herrmann, P., Sievers, M., Zoyke, A., Whittaker, M., Beutner, M., Fels, G., Reinhardt, W., Magenheimer, J. (2012): Designing a Web-Based Application to Support Peer Instruction for Very Large Groups. In Proceedings of the International Conference of Information Systems, Orlando, USA.
- Krieschke, W. (2005): Der elektronische Prüfer. Die Zeit, 2005/32.
- Schermutzki, M. (2008): Learning-Outcomes - Lernergebnisse: Begriffe, Zusammenhänge, Umsetzung und Erfolgsvermittlung, Lernergebnisse und Kompetenzvermittlung als elementare Orientierungen des Bologna-Prozesses. Online: http://opus.bibliothek.fhaachen.de/opus/volltexte/2008/248/pdf/schermutzki_learning_outcomes.pdf.

Anhang

A1: Muster eines Scantron-Bogens



Universität Paderborn

Name:

Vorname:

Wichtig! Die folgenden Angaben unbedingt eintragen ...

Matrikel-Nr.

0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0

Code-Nr.

0	0	0	0	0	0
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0
7	0	0	0	0	0
8	0	0	0	0	0
9	0	0	0	0	0

**Klausur-
alternative**

A

B

**Bogen nicht knicken, falten oder verschmutzen.
Nur mit Bleistift markieren.**

**Die Antworten
so markieren –**



Gewertet werden nur die Antworten, die auf diesem Lösungszettel an der vorgegebenen Stelle und in der vorgegebenen Weise mit Bleistift markiert worden sind.

1 = A B C D E

2 = A B C D E

3 = A B C D E

4 = A B C D E

5 = A B C D E

6 = A B C D E

7 = A B C D E

8 = A B C D E

9 = A B C D E

10 = A B C D E

11 = A B C D E

12 = A B C D E

13 = A B C D E

14 = A B C D E

15 = A B C D E

16 = A B C D E

17 = A B C D E

18 = A B C D E

19 = A B C D E

20 = A B C D E

21 = A B C D E

22 = A B C D E

23 = A B C D E

24 = A B C D E

25 = A B C D E

26 = A B C D E

27 = A B C D E

28 = A B C D E

29 = A B C D E

30 = A B C D E

31 = A B C D E

32 = A B C D E

33 = A B C D E

34 = A B C D E

35 = A B C D E

36 = A B C D E

37 = A B C D E

38 = A B C D E

39 = A B C D E

40 = A B C D E

41 = A B C D E

42 = A B C D E

43 = A B C D E

44 = A B C D E

45 = A B C D E

46 = A B C D E

47 = A B C D E

48 = A B C D E

49 = A B C D E

50 = A B C D E

51 = A B C D E

52 = A B C D E

53 = A B C D E

54 = A B C D E

55 = A B C D E

56 = A B C D E

57 = A B C D E

58 = A B C D E

59 = A B C D E

60 = A B C D E

61 = A B C D E

62 = A B C D E

63 = A B C D E

64 = A B C D E

65 = A B C D E

66 = A B C D E

67 = A B C D E

68 = A B C D E

69 = A B C D E

70 = A B C D E

71 = A B C D E

72 = A B C D E

73 = A B C D E

74 = A B C D E

75 = A B C D E

76 = A B C D E

77 = A B C D E

78 = A B C D E

79 = A B C D E

80 = A B C D E

81 = A B C D E

82 = A B C D E

83 = A B C D E

84 = A B C D E

85 = A B C D E

86 = A B C D E

87 = A B C D E

88 = A B C D E

89 = A B C D E

90 = A B C D E

91 = A B C D E

92 = A B C D E

93 = A B C D E

94 = A B C D E

95 = A B C D E

96 = A B C D E

97 = A B C D E

98 = A B C D E

99 = A B C D E

100 = A B C D E

101 = A B C D E

102 = A B C D E

103 = A B C D E

104 = A B C D E

105 = A B C D E

106 = A B C D E

107 = A B C D E

108 = A B C D E

109 = A B C D E

110 = A B C D E

111 = A B C D E

112 = A B C D E

113 = A B C D E

114 = A B C D E

115 = A B C D E

116 = A B C D E

117 = A B C D E

118 = A B C D E

119 = A B C D E

120 = A B C D E

Kaimann, D. (2014). Benutzerfreundlichkeit von Lernüberprüfungen - Ein Vergleich von papierbasierten und elektronischen Prüfungen. In R. Kordts-Freudinger, D. Urban & N. Schaper (Hrsg.), Lehr- und Lernpraxis im Fokus - Forschungs- und Reflexionsbeiträge aus der Universität Paderborn.

(abrufbar unter www.zhw.uni-hamburg/almanach)

ISSN: 2192-1466

