

Digitale Studien- und Prüfungsleistungen¹

Masha Mae Arakeljan & Kai-Uwe Schnapp, Universität Hamburg

(Version 2.0, 11.06.2021)

1 Einleitung

Hybride und Onlinelehrformen sind Teil des akademischen Alltags. Dafür standen Begriffe wie Blended Learning, Flipped Classroom und Online Assessment schon lange, bevor SARS CoV-2 mit der ganzen auch die akademische Welt in einen anderen Operationsmodus zwang. Am Fachbereich Sozialwissenschaften in Hamburg sammeln wir seit sechs Jahren in zwei großen einführenden Veranstaltungen Erfahrungen mit Onlinetests. Diese Erfahrungen werden wir hier vorstellen und reflektieren, um sie für aktuelle und künftige Lehrsituationen nutzbar zu machen.

Für die Formung der eigenen Einstellung zu Onlinetests ist es u.E. relevant, zu wissen, warum wir uns bereits vor Längerem für diese Form der Leistungsüberprüfung entschieden haben. Wir wollten weg von Klausuren, die am Ende einer Vorlesungszeit vor allem Faktenwissen abfragen. Stattdessen wollten wir eine Form der Leistungsprüfung etablieren, die eine lernbegleitende und lernunterstützende Selbstkontrolle im laufenden Semester ermöglicht und von uns aktiv in den Lernprozess eingebunden werden kann. Das Ganze hat positive Nebeneffekte: Das fortlaufende Prüfen erleichtert es den Lehrenden, ihre Lehrinhalte im Sinne des *Constructive Alignment* vom Ende her zu entwickeln.² Außerdem sinkt die geballte Prüfungslast am Vorlesungsende bei den Studierenden.

Im Folgenden werden wir zunächst Informationen zum Lehrkontext geben, in den die Onlineprüfung eingebettet ist und auf die technischen sowie rechtlichen Rahmenbedingungen eingehen. Danach steht das Prüfungsformat selbst im Vordergrund, wobei wir explizit auf die unterschiedlichen Frageformate, die Bewertung sowie die didaktische Qualitätssicherung Bezug nehmen.

Abschließend nennen wir aus unserer Sicht benötigte Ressourcen und empfehlen einen organisatorischen Ablauf für die Vorbereitung von Onlinetests. Zur Unterstützung für Lehrende, die in das Online-Prüfen einsteigen wollen, haben wir Checklisten für die Praxis verfasst.

¹ Wir danken Brian Dietrich sowie Olga Herzog für Ergänzungen und Anregungen zu diesem Beitrag und Saskia Mestern für das wie immer sehr umsichtige Lektorat.

Eine kurze Fassung dieses Textes, in der die Liste von Frageformaten sowie die Beispielsammlung im Anhang fehlen, ist erschienen als Schnapp (2020).

² Das Prinzip des „constructive alignment“ besagt, die Lehre bei der Planung von den gewünschten Lernergebnissen her zu denken: Was sollen die Studierenden nach einer Lehrepisode, also z.B. einer Vorlesung, wissen und können? Wie muss ich prüfen, damit ich feststellen kann, ob dieses Wissen und Können vorliegt? Und wie muss ich lehren, damit die Studierende dieses Wissen und Können erwerben? Siehe <https://www.lehren.tum.de/themen/lehre-gestalten-didaktik/erfolgskriterien-guter-lehre/constructive-alignment/> (14.07.2020).

2 Anwendungskontext der Onlineprüfungen

2.1 Lehrkontext

Onlinetests finden am Fachbereich Sozialwissenschaften der Universität Hamburg in zwei Erstsemesterveranstaltungen statt: Der Ringvorlesung „Einführung in die Politikwissenschaft“, an der alle Professor:innen der Politikwissenschaft mit mindestens einem Termin beteiligt sind, und der Vorlesung „Einführung in die empirischen Methoden der Sozialforschung“. Der Test wird erstens genutzt, um in den Vorlesungen vermitteltes Wissen abzufragen. Wir nutzen ihn zweitens auch, um Studierende in die ergänzende Literatur zu „drängen“ oder ihnen Webseiten nahezubringen, die für das Fach und das Studium relevant sind. So gibt es in der Methodenvorlesung Testfragen, die ohne Nutzung des Messinstrumentenrepositoriums der GESIS³ nicht beantwortet werden können. Das Ziel dieser Art von Fragen ist nicht die Wissensabfrage, sondern das Kennenlernen von und das Anregen zur Auseinandersetzung mit diesen für das Studium wichtigen Quellen. Die Prüfungssituation wird also explizit zur Fortsetzung des Lern- und Auseinandersetzungsprozesses genutzt. Die Methodenvorlesung arbeitet während der Präsenzsitzungen außerdem regelmäßig mit einem Classroom Response System (bzw. einem Onlinequiztool für die Onlinelehre). Die mit diesen Systemen in den Lehrveranstaltungen bearbeiteten Fragen fungieren teilweise als Vorbereitung für die Online-Tests. So wird nebenher ein Anreiz zur aktiven Beteiligung an den Vorlesungen geschaffen.

Die Onlinetests sind in erster Linie obligatorische Studienleistungen. Die erreichte Gesamtpunktzahl über alle Tests ist entscheidend für das Bestehen der Studienleistung. Da aber sowohl an der Ringvorlesung „Einführung in die Politikwissenschaft“ als auch an der Methodenvorlesung Studierende mit unterschiedlichen Prüfungsordnungen teilnehmen (Hauptfach- und Nebenfach, Lehramtsstudierende), müssen einige Studierende nicht nur eine Studien-, sondern eine Prüfungsleistung erbringen. Die Onlinetests werden dann als benotete Prüfungsleistung gewertet. Die Note wird am Ende der Veranstaltung auf der Basis der über alle Tests erreichten Gesamtpunktzahl ermittelt.

2.2 Technische Umgebung

Die Universität Hamburg nutzt das E-Learning-System (ELS) OpenOLAT.⁴ In OpenOLAT werden für die oben genannten Veranstaltungen Kurse erstellt. Diese enthalten Lehrmaterialien (Folien und ergänzende Literatur, Zugang zu Vorlesungsvideos) und dienen der Kommunikation innerhalb des Kurses und mit den Teilnehmenden. Gleichzeitig werden die Onlinetests innerhalb dieser Kurse bereitgestellt und verwaltet.

OpenOLAT ist mit einem Testeditor ausgestattet, der Onlinetests im QTI 2.1-Format erzeugt. Im Testmodul können Tests erstellt werden, es ist aber auch möglich, bereits bestehende Fragebibliotheken einzulesen, wenn diese im IMS-QTI-Format gespeichert wurden.⁵

Das Testmodul wertet die Tests in den meisten Fällen automatisch aus (mehr dazu in Abschnitt 3.1), informiert die Studierenden über die pro Test erreichte Punktzahl und erlaubt es, den Studierenden inhaltliche Rückmeldungen auf jeden Test zu geben. Bei Bedarf können auch Musterlösungen zum nachträglichen Einsehen eingestellt werden. Zusätzlich bietet das

³ <https://zis.gesis.org/> (16.07.2020).

⁴ <https://www.openolat.com/open-source-lms/> (14.07.2020).

⁵ Erläuterungen zum QTI-Standard finden sich u.a. hier: https://www.e-teaching.org/materialien/glossar/ims_qti (14.07.2020).

System die Möglichkeit, die erstellten Testfragen in einem Fragenpool zu hinterlegen. So können Fragebibliotheken für die Erstellung zukünftiger Tests genutzt und innerhalb des Systems auch mit anderen Lehrenden ausgetauscht werden.

2.3 Rechtliche Rahmenbedingungen

Der rechtliche Rahmen von Onlineprüfungen wird neben den Prüfungsordnungen vor allem durch die europäische Datenschutzgrundverordnung (DSGVO) bestimmt, da es bei jeder Form der digitalen Prüfung immer auch zu einer Verarbeitung von Daten kommt (Schwartzmann 2020). Zu regeln ist bei dem hier vorgestellten Typ der Online-Prüfung vor allem der Zugriff auf die Leistungen der Studierenden. Außerdem sind konkrete Regeln für das Speichern und Löschen der Daten festzulegen, die sich allerdings auch am Prüfungsrecht orientieren müssen (ebd.: 503). Die Hochschulen sind verpflichtet, die Einhaltung von Datenschutzrechten bei der Software-Nutzung zu gewährleisten und schließen hierfür Vereinbarungen mit den anbietenden Firmen ab. Damit Lehrpersonen sich bei datenschutzrechtlichen Fragen nicht in die alleinige Verantwortung stellen (ebd. 502), sollten sie sich unbedingt an den Software-Vorgaben der eigenen Hochschule orientieren.

Prüfungsrechtlich sind Onlinetests nach wie vor ein unsicheres Terrain. Bei ihrer Nutzung sollte daher in jedem Falle die Rechtsinterpretation der eigenen Hochschule beachtet werden. Eine unhintergehbare Notwendigkeit ist die Einhaltung des auch sonst bei der Bewertung von Prüfungen geforderten Vieraugenprinzips. Dies wird bei Onlinetests dadurch sichergestellt, dass Fragen, Antwortoptionen und richtige Antworten sowie die Punktvergaben von einer Lehrperson entworfen werden, und dann von einer zweiten Person geprüft werden müssen, ehe die Inhalte im System programmiert werden. Zentrales Prüfkriterium für die Rechtssicherheit der Fragen ist, dass diese möglichst eindeutig beantwortbar sein müssen.

Je nachdem, ob Onlinetests benotete Prüfungsleistungen sind oder als Studienleistungen genutzt werden, sind strengere oder weniger strenge rechtliche Voraussetzungen zu erfüllen. Immer muss aber sichergestellt sein, dass die geltende Prüfungsordnung (PO) und das Landeshochschulgesetz Onlinetests zulassen, die von den Studierenden in Eigenregie und ohne Aufsicht von zu Hause durchgeführt werden. Im Falle von Studienleistungen lässt die PO am Fachbereich Sozialwissenschaften der Universität Hamburg hier freie Hand. Haben die Onlinetests Prüfungscharakter, so werden sie als Take-Home-Prüfungen verstanden. Diese sind in der PO ausdrücklich geregelt. So legt diese zeitliche Rahmendaten fest (etwa die Zeit zwischen frühestem Start und spätestem Ende eines Tests oder die Zeit, die die einzelne Studierende mit dem Test maximal zubringen darf).⁶ Solche Regeln muss das genutzte Prüfungssystem natürlich auch implementieren können, wenn sie etwas nützen sollen.

3 Die Onlineprüfung selbst⁷

3.1 Fragen und Frageformate

Eine gute Onlineprüfung braucht Fragen, die 1) die Erreichung der gesetzten Lernziele prüfen, 2) unterschiedliche Denkstufen abdecken, 3) die gewünschten Arbeitsschritte, etwa den Besuch einer Webseite oder das Nachlesen in einem Text, notwendig machen, 4) die vorhandene

⁶ Detailliertere rechtliche Hinweise gibt e-teaching.org: <https://www.e-teaching.org/lehrszenarien/pruefung/pruefungsform/e-pruefung> (14.07.2020).

⁷ Ausführliche Hinweise zur Entwicklung und Durchführung von Onlinetests finden sich hier https://www.lehren.tum.de/fileadmin/w00bmo/www/Downloads/Themen/Studiengaenge_gestalten/Dokumente/MCEmpfehlungen_Stand_Oktober_2012_final.pdf (11.06.2021) und hier http://www.bris.ac.uk/esu/media/e-learning/tutorials/writing_e-assessments/index.htm (11.06.2021).

Palette an Frageformaten ausnutzen, und 5) die Ressourcen zur Überprüfung der Antworten nicht überfordern – also möglichst auf automatische Auswertbarkeit setzen (siehe unten die Anmerkungen bei den Fragetypen Freitext und Lückentext). Im Folgenden werden einige Frageformate vorgestellt:⁸⁹

- Single-Choice-Fragen bestehen aus einer Frage und mindestens zwei Antwortoptionen, von denen nur eine richtig ist.
- Multiple-Choice Fragen bestehen aus einer Frage und mindestens zwei Antwortoptionen, von denen mehrere richtig sein können.¹⁰
- Kprim-Fragen bestehen immer aus einer Frage und genau vier Antwortoptionen. Für jede dieser vier Optionen muss die Testperson entscheiden, ob sie zutrifft (richtig ist) oder nicht. Es können 0-4 Antworten richtig sein. Kprim-Fragen sind testtheoretisch den SC- und MC-Fragen deutlich überlegen und führen zu einer gerechteren Punktevergabe.¹¹
- In Matrixfragen sind die Antwortoptionen in einer Tabelle angeordnet. Die Aufgabe besteht dann darin, die Kategorien in den Zeilen richtig den Kategorien in den Spalten zuzuordnen. Dabei sind innerhalb von Zeilen und Spalten Einfach- und Mehrfachzuordnungen denkbar.
- Bei Zuordnungsaufgaben (Drag & Drop) müssen Begriffe, Aussagen oder auch Darstellungen anderen Begriffen, Aussagen oder Darstellungen durch verschieben mit der Maus zugeordnet werden. Zuordnungsaufgaben eignen sich gut für die Prüfung höherer Denkstufen.
- Wahr/Falsch-Aufgaben bestehen aus einem sinnvollen Text mit aufeinander aufbauenden Aussagen. In diesen Text wird an mehreren Stellen die Möglichkeit eingebaut, Aussagen des Textes als wahr oder falsch zu kennzeichnen. Auch dieser Aufgabentyp eignet sich gut für die Prüfung höherer Denkstufen, weil Zusammenhänge darstellbar sind und das Zusammenhangsverständnis geprüft werden kann.
- Eine Lückentextfrage besteht aus einem Text in den Lücken integriert sind. Diese Lücken sind durch die Studierenden mit eigenen Formulierungen oder auch Zahlen zu füllen. Lückentextfragen eignen sich sehr gut für die Prüfung höherer Denkstufen. Gleichwohl sind sie in größeren Lerngruppen nur zu empfehlen, wenn das Prüfungssystem einen intelligenten Auswertungsalgorithmus beinhaltet, der etwa Tippfehler und Wortvarianten erkennt. Steht ein solches System nicht zur Verfügung, sollte auf Fragen mit offener Texteingabe unbedingt verzichtet werden, da die Auswertung von Lückentextfragen dann sehr (personal)aufwändig ist.

Eine Ausnahme von dieser Regel bilden Lückentextfragen, die numerische Antworten erfordern, bei denen die Antwort also ausschließlich aus Zahlen besteht. Wenn das System in der Lage ist, alphanumerische Eingaben (Buchstaben) zu erkennen und zurückzuweisen, und wenn es erlaubt, Wertebereiche anstelle fixer Werte vorzugeben, dann kann die Auswertung solcher Lückentextfragen automatisiert erfolgen. Ihre Anwendung ist dann auch für große Lerngruppen weitgehend ressourcenneutral möglich.

⁸ Die hier vorgestellte Liste ist der Webseite des in Hamburg genutzten ELS OpenOLAT entnommen: <https://confluence.openolat.org/display/OO130DE/Testeditor+QTI+2.1+im+Detail> (15.07.2020).

⁹ Konkrete Fallbeispiele zu den genannten Fragetypen finden Sie im Anhang. Sie stammen aus vergangenen Onlineprüfungen der oben genannten Ringvorlesung für Erstsemester der Politikwissenschaft.

¹⁰ Hinweise für die Formulierung guter SC- und MC-Fragen finden sich hier: http://theelearning-coach.com/elearning_design/rules-for-multiple-choice-questions/ (14.07.2020).

¹¹ Siehe auch <https://www.databay.de/unternehmen/newsarchiv/multiple-choice-fragentyp-kprim-in-iliast-integriert/> (14.07.2020).

- Eine Alternative zum Lückentext ist der sogenannte Hottext. Hier werden nicht Textlücken gelassen, sondern im Text alternative Formulierungen vorgeschlagen, von denen die jeweils richtige auszuwählen ist. Diese Testvariante eignet sich ebenfalls sehr gut für die Prüfung des Denkens höherer Ordnung und ist in jedem Falle automatisch auswertbar.
- Eine Möglichkeit Abbildungen zur Wissensprüfung zu nutzen, ist die Hotspot-Frage. Hier müssen die Studierenden in einer Abbildung (Karte, Ablaufschema eines Prozesses, räumliches Politikmodell etc.) die als Antwort auf die Frage passende richtige Stelle mit der Maus anklicken. Dieser Fragetyp eignet sich gut für die Prüfung des Denkens höherer Ordnung, wie etwa Anwendungsfragen.
- Der Vollständigkeit halber seien Freitextantworten erwähnt, die aber für große Lerngruppen noch weniger zu empfehlen sind als Lückentextfragen, weil hier der Auswertungsaufwand nicht nur groß ist, sondern auch noch sehr schwer zu kalkulieren.

3.2 Die Bewertung der Tests und der einzelnen Fragen

Damit die Studierenden sich über die einzelnen Testrunden hinweg gut orientieren können, soll die Punktzahl für jeden Einzeltest bzw. jedes Thema (annähernd) gleich sein. Die Punktzahl, die für eine Frage vergeben wird, soll den Schwierigkeitsgrad der Frage deutlich machen und soll daher auch für die Studierenden sichtbar sein. Eine formale Rückmeldung über das Testergebnis sollte erfolgen, sobald die einzelne Studierende den Test abgeschlossen hat. Dies ist allerdings nur bei einer automatischen Auswertung der Fragen durch das System möglich. Die Rückmeldung kann dann entweder konkret über die erreichte Punktzahl oder über die Erreichung einer „Bestehensgrenze“ erfolgen. Wenn der Onlinetest nicht nur als Leistungsüberprüfung dienen soll, sondern als Teil des Lernprozesses verstanden wird, dann muss nach Abschluss einer Testrunde zusätzlich eine inhaltliche Rückmeldung durch Vorlage der korrekten Antworten erfolgen. Durch das inhaltsbezogene Feedback über die Punktzahl hinaus, erhalten Studierende so die Möglichkeit, sich mit falschen Antworten auseinanderzusetzen und Inhalte nachzuarbeiten.

Eine wichtige rechtliche Vorgabe ist, dass keine Übertragung negativer Ergebnisse von einer Aufgabe auf eine andere erfolgen darf. Anders ausgedrückt muss die Mindestpunktzahl, die bei einer Aufgabe erreicht werden kann, immer 0 Punkte betragen. Bei Multiple-Choice-Fragen dürfen *innerhalb* einer Frage falsche von richtigen Antworten abgezogen werden. Das Gleiche gilt für Wahr/Falsch-Fragen. Minuspunkte, etwa für mehr falsche als richtige Antworten, dürfen dabei allerdings nicht herauskommen, denn sie verletzen das Verbot der Übertragung negativer Ergebnisse.

Es gibt für die meisten Fragetypen unterschiedliche Wege, eine Punktzahl zu ermitteln. Daher ist es wichtig, vor der Testerstellung festzulegen, nach welchem System gearbeitet werden soll. Dieses muss dann konsistent durchgehalten werden, damit die Studierenden sich darauf einstellen können.¹²

3.3 Qualitätssicherung

Eine gesicherte didaktische Qualität von Onlinetests setzt voraus, dass die Lehrenden sich genau überlegen, was sie prüfen und daher fragen wollen. Sie müssen sich also über den abzu-

¹² Hinweise auf Punktsysteme finden sich etwa auf dieser Wikipediaseite: https://de.wikipedia.org/wiki/Multiple_Choice#Bewertung_von_Testleistungen (16.07.2020). Hier wird unter anderem auch erläutert, wie dem Erraten von Antworten durch das Punktesystem begegnet werden kann.

fragenden Inhalt, den jeweils geeigneten Fragetyp und die zu erreichende Punktezahl Gedanken machen. Die einzelnen Fragen sollen möglichst unterschiedliche Schwierigkeitsgrade aufweisen, die Fragetypen sollen variieren (Stykwow 2018).

Die Qualität von Tests kann erhöht werden, wenn es ein Team von Kolleg:innen gibt, das gemeinsam Fragen entwickelt und prüft. Unabhängig davon muss, wie bereits erwähnt, das Vier-Augen-Prinzip gesichert sein.¹³ Insbesondere bei vollständiger Neuentwicklung sollte für jeden einzelnen Test ein technischer Probelauf durchgeführt werden, bei dem studentische Hilfskräfte oder Tutor:innen den Test bearbeiten.

Sollen Fragen in einem weiteren Semester erneut genutzt werden, ist zu empfehlen, die Fragenqualität nach der Testdurchführung statistisch zu prüfen. Fragen mit sehr hoher Erfolgsquote sind wahrscheinlich zu leicht, Fragen mit sehr niedriger Erfolgsquote zu schwer. Bei Fragen mit annähernder Gleichverteilung der Antworten sind die Antwortoptionen möglicherweise nicht ausreichend eindeutig, um ein sicheres Erkennen der richtigen Antwort(en) zu ermöglichen.

3.4 Sicherung der Autorenschaft der Antworten

Führen Studierende einer großen Studiengruppe, also etwa einer Vorlesung, Onlinetests ohne Aufsicht am heimischen PC/Laptop durch, stellt sich die Frage, wie die Autor:innenschaft der Antworten gesichert werden kann. In einem Fach wie der Politikwissenschaft, ist die Hauptprüfungsform die Hausarbeit. Diese wird grundsätzlich allein und ohne Aufsicht erstellt. Vor diesem Hintergrund kann sich zunächst also auch überlegt werden, für wie schwerwiegend das Problem überhaupt gehalten wird. Ein Schritt zur rechtlichen Sicherung kann es dann sein, wie unter Hausarbeiten eine „Ehrenerklärung“ zu verlangen. Es gibt aber auch technische und inhaltliche Möglichkeiten, das „Abschreiben“ von Lösungen zu behindern.

Der erste Schritt ist das randomisierte Vorlegen der Antwortoptionen einer Frage, die von den meisten Testsystemen unterstützt wird. Die einfache Weitergabe der Nummer der richtigen Antwortoption wird so unterbunden. Steht ein großer Fragenpool zur Verfügung, dann kann dieser genutzt werden, um vom Testsystem individuelle Tests zusammenstellen zu lassen, die für alle Studierenden anders aussehen. Viele Systeme erlauben Frageblöcke. Diese bestehen aus mehreren inhaltsähnlichen punktgleichen Fragen. Allen Testteilnehmenden wird eine zufällige Auswahl der Fragen aus einem Block vorgelegt, sodass über mehrere Blöcke hinweg individuelle Tests entstehen. Das einfache Abschreiben von „Musterlösungen“ wird so effektiv unterbunden. Die Entstehung eines für die Nutzung dieser Option hinreichend großen Fragenpools ist oft vor allem eine Frage der Zeit. Wird bei jedem neuen Durchlauf des Tests die Menge der vorhandenen Fragen erweitert, entsteht ein Schritt für Schritt wachsender Fragenpool.

Gehen die Ideen für neue Fragen aus, ist es eine Möglichkeit, Studierende an der Entwicklung neuer Fragen zu beteiligen. Ein Anreizsystem können Zusatzpunkte für das Testpunktekonto sein. Für so ein Vorgehen müssen transparente Anerkennungsregeln vorliegen, die etwa klären, wie mit ähnlichen oder gar gleichen eingereichten Fragen umgegangen wird oder inwiefern die unmittelbare Testtauglichkeit einer Frage Voraussetzung für eine Punktevergabe ist.

¹³ Diese Webseite <https://confluence.openolat.org/display/OO130DE/Fragenpool+Beurteilungsprozess> (11.06.2021) stellt ein sehr detailliertes Prozessmodell für die Erstellung, Prüfung und Implementierung von Onlinetests zur Verfügung.

Was kann getan werden, wenn Studierende sich diesen Vorkehrungen zum Trotz bei der Bearbeitung der Tests beraten? Aus unserer Sicht stellt das vor allem dann kein Problem dar, wenn der Test als Studienleistung vorgenommen wird. Wir sehen das sogar als produktiv an, weil es eine Form der kollektiven Auseinandersetzung mit dem Thema ist. Wird Prüfen als Teil des Lernprozesses verstanden und nicht ausschließlich als „Assessment“, so ist das eher ein wünschenswerter denn ein zu unterbindender Effekt.

Checkliste 1: Der einzelne Test

- Sind ausreichend viele und unterschiedliche Fragen für den Test vorhanden?
- Sind die vorhandenen Fragen technisch umsetzbar?
- Sind die Fragen inhaltlich und formal vielfältig?
- Entspricht der Test den Vorgaben für Gesamtpunktzahl und Schwierigkeitsgrad?
- Ist die Prüfung getestet worden?
- Sind Start- und Endtermin sowie die Bearbeitungszeit festgelegt, programmiert und kommuniziert?
- Ist eine verständliche und korrekte Rückmeldung über die erreichte Punktzahl vorbereitet?
- Sind Musterantworten vorbereitet, an denen die Studierenden ihr Antwortverhalten überprüfen können?

4 Benötigte Ressourcen

Welche Ressourcen werden für die erfolgreiche Implementation von Onlinetests benötigt? Natürlich bedarf es eines E-Learning-Systems (ELS), das Onlinetests zur Verfügung stellt und das möglichst im regulären Besitz der eigenen Hochschule ist. Tests sollten in dem ELS leicht umsetzbar sein und über geeignete Ex- und Importmöglichkeiten verfügen, damit vor allem am Anfang auf Fragenpools von Kolleginnen zurückgegriffen werden kann. Idealerweise wird das ELS an der Hochschule bereits breit genutzt. Für Lehrende senken sich die Hürden, wenn es erfahrene Kolleg:innen oder eine gute Unterstützung durch E-Learning-Büros oder vergleichbare Einrichtungen gibt. Letztere sind zusammen mit dem Hochschulrechenzentrum wichtige Akteur:innen für den (technischen) Support.

In jedem Falle werden für die Umsetzung von Onlinetests ausreichend Hilfskraftstunden benötigt. Das gilt unabhängig davon, ob ein Test zum ersten oder zum wiederholten Male durchgeführt wird. Die Zahl der benötigten Stunden sinkt bei der wiederholten Durchführung, aber jeder Testdurchlauf braucht Personalressourcen. Dieser Bedarf steigt, wenn mehrere Personen in einer Lehrveranstaltung koordiniert werden müssen, wie es beispielsweise bei dem Format der Ringvorlesung der Fall ist. Die Umsetzung der Tests durch eine studentische Hilfskraft (SHK) beinhaltet mindestens folgende Aufgaben:

- 1) Sammlung der Fragen für die einzelnen Testrunden
- 2) Prüfung der Fragen und gegebenenfalls Rücksprache und Überarbeitung
- 3) Programmierung und Testung der Prüfung
- 4) Organisation und Verwaltung der Prüfung
- 5) Technische Unterstützung für die Studierenden während des laufenden Tests. Hier ist insbesondere wichtig, dass die Studierenden die SHK als Ansprechperson für Technisches und Organisatorisches anerkennen, damit diese der Lehrperson den entsprechenden Kommunikationsaufwand abnehmen kann.

- 6) Aufbereitung und Aggregation der Testergebnisse
- 7) Idealerweise Analyse der Testergebnisse als Unterstützung für die Weiterentwicklung des Fragenpools

5 Organisation und zeitlicher Ablauf

Für Vorbereitung und Durchführung eines Tests sollte großzügig Zeit eingeplant werden. Für den ersten Versuch empfehlen wir folgenden Ablauf, später kann gegebenenfalls knapper kalkuliert werden:

- 3 Wochen vor dem Testtermin: Lehrperson gibt Fragen bei der SHK ab
- Sofort im Anschluss:
 - SHK prüft Vollständigkeit und Umsetzbarkeit
 - gegebenenfalls Rücksprache von SHK mit Lehrperson
- 2 Wochen vor dem Einsatz: Test wurde von SHK programmiert
- Direkt im Anschluss: Test
 - bei Erstanwendung: mit 3-5 Studierenden, etwa Tutor:innen; bei wiederholter Anwendung sind Tests durch die SHK gegebenenfalls ausreichend
- 2 Tage vor dem Einsatz: Korrekturen am Test wurden vorgenommen, Test ist startklar

Checkliste 2: Einführung und Umsetzung von Onlinetests

- Welche didaktischen Ziele möchte ich mit dem Test erreichen?
 - Sind auf dem zu prüfenden Studien- und Wissensniveau standardisierte Tests als Prüfungsinstrument geeignet?
 - Sind die zu prüfenden Themen für standardisierte Tests geeignet?
- Lohnt sich der Aufwand, d.h. wird der zu entwickelnde Test von einer großen Zahl von Studierenden und in mehreren Runden genutzt werden?
- Ist an meiner Hochschule ein geeignetes ELS verfügbar?
 - Ist das ELS geeignet, meine didaktischen Ziele umzusetzen?
 - Ist das ELS bei Studierenden bekannt und leicht bedienbar?
 - Gibt es Ansprechpartner:innen für didaktische und technische Unterstützung der Lehrenden?
 - Gibt es Kurse, die Lehrende in das ELS und das Onlineprüfungsmodul einführen?
- Verfüge ich über die notwendigen Personalressourcen (SHK)?

Für den Test selbst geben wir den Studierenden eine Woche Zeit. Innerhalb dieser Laufzeit haben sie beliebig viel Zeit für die Bearbeitung des Tests. Die Bearbeitung kann jederzeit und mehrfach unterbrochen werden, die Ergebnisse müssen dafür zwischengespeichert werden. Aus technischen Gründen geben wir den Studierenden grundsätzlich zwei Testversuche. Ein Testversuch ist erfolgreich, wenn über 50% der Punkte erreicht worden sind. Geht beim ersten Versuch etwas schief, das den Studierenden nicht anzulasten ist, muss so nicht sofort durch die Testadministration eingegriffen und ein Ersatzversuch freigeschaltet werden. Zusätzlich werden Studierende, die im ersten Versuch unter 50% der Punkte erhalten haben, dazu ange-regt, sich in einem zweiten Versuch erneut mit dem abgefragten Thema zu beschäftigen. Hierzu erhalten die Studierenden direkt nach Ihrem ersten Testversuch neben der formalen

Rückmeldung über die Punktzahl den Hinweis, dass Ihnen noch ein zweiter Versuch zur Verfügung steht.

6 Fazit

Fällt der Begriff Onlineprüfung, dann assoziieren sowohl Studierende als auch Lehrende zunächst vermutlich stupide Single- oder Multiple-Choice-Fragen, mit denen allenfalls ein wenig Faktenwissen geprüft werden kann. Klug eingesetzt können solche Tests allerdings sehr viel mehr sein und Lehr-Lern-Prozesse sinnvoll ergänzen oder gar fördern. Das gilt vor allem dann, wenn auch die Prüfung selbst durch geeignete Frageinhalte und -formate die Studierenden zur Wissenserweiterung anregt. Gelingt dies, sind Onlinetests ein weiteres wertvolles Instrument im didaktischen Werkzeugkoffer. Eine rechtliche Absicherung kann über die Vorgaben der Hochschulen erfolgen und auch die zu überwindenden technischen Hürden sind angesichts weit entwickelter E-Learning-Systeme heute nicht mehr groß. Das gilt umso mehr, wenn an der eigenen Hochschule ein Mindestmaß an technischer Unterstützung gegeben ist.

Checkliste 3: Studierende

- Gibt es eine Einführung in die Nutzung des ELS und des Onlineprüfungsmoduls für die Studierenden?
- Wird die technische Bedienung der Tests im Detail erklärt und kennen Studierende die technischen, formalen und inhaltlichen Spielregeln?
- Haben die Studierenden einen guten zeitlichen Überblick über alle Tests?
- Haben die Studierenden eine Ansprechperson bei technischen Schwierigkeiten und für inhaltliche Fragen?
- Kennen die Studierenden die Vorgehensweise bei der Bewertung der einzelnen Tests und bei der Zusammenfassung der Testergebnisse?
- Wissen die Studierenden, wann und wie die Ergebnisse der einzelnen Tests und das Gesamtergebnis kommuniziert werden?
- Gibt es klare Regelungen für eine Wiederholung bei Nichtbestehen?
- Sind die Antworten auf all diese Fragen an geeigneter Stelle (ELS) für die Studierenden nachlesbar?

7 Literaturverzeichnis

Schnapp, Kai-Uwe. 2020. Digitale Studien- und Prüfungsleistungen. In *Bausteine digitaler Hochschullehre in der Politikwissenschaft*, Hrsg. Daniel Lambach, 43–52. Frankfurt a. M.: Wochenschau Verlag.

Schwartmann, Rolf. 2020. Prüfungsrechtliches Harakiri. Was beim digitalen Prüfen rechtlich zu beachten ist. *Forschung & Lehre* 27: 502–503.

Stykov, Petra. 2018. *Prüfen in politikwissenschaftlichen Studiengängen*. Frankfurt a. M.: Wochenschau Verlag.

Anhang: Beispiele für Frageformate

Frage 1: Single Choice

Klassische Politikbereiche - Single Choice	2 Punkte	Nicht beantwortet
Welcher der folgend genannten Politikbereiche gehört nicht zu den „klassischen“ Politikbereichen?		
Markieren Sie die entsprechende Antwort.		
<input type="radio"/> Innenpolitik		
<input type="radio"/> Finanzpolitik		
<input type="radio"/> Justiz		
<input type="radio"/> Wirtschaftspolitik		
<input type="radio"/> Verteidigungspolitik		
<input type="radio"/> Außenpolitik		
<input checked="" type="checkbox"/> Antwort speichern		Nächste Frage >

Frage 2: Multiple Choice

Anmerkung:

Hier muss auf eine Vergabe von Minuspunkten für falsche Antworten geachtet werden, da die Studierenden sonst alle Antworten auswählen könnten. Dennoch darf die Punktzahl am Ende der Frage null nicht unterschreiten.

Policy-Cycle nach Blum-Schubert - Multiple Choice	3 Punkte	Nicht beantwortet
Welche Kritiken am Modell des Policy-Cycle wird in Kapitel 5 von Blum/Schubert formuliert?		
Bitte markieren Sie die entsprechende(n) Aussage(n).		
<input type="checkbox"/> Der Policyprozess wird durch das Modell übermäßig komplex dargestellt.		
<input type="checkbox"/> Der Policyprozess wird durch das Modell zu stark vereinfacht dargestellt.		
<input type="checkbox"/> Das Modell ist zu stark bottom-up orientiert.		
<input type="checkbox"/> Das Modell kann keine Kausalitäten aufzeigen.		
<input checked="" type="checkbox"/> Antwort speichern		Nächste Frage >

Frage 3: Kprim

Begriffe des Konstruktivismus - Kprim		4 Punkte	Nicht beantwortet
Sind die untenstehenden Begriffe für den Konstruktivismus entscheidend?			
Markieren Sie richtig/falsch zu der jeweiligen Aussage.			
Richtig	Falsch		
<input type="checkbox"/>	<input type="checkbox"/>	Der Begriff der "Ideen" ist für den Konstruktivismus entscheidend.	
<input type="checkbox"/>	<input type="checkbox"/>	Der Begriff der "Identität" ist für den Konstruktivismus entscheidend.	
<input type="checkbox"/>	<input type="checkbox"/>	Der Begriff der "Demokratie" ist für den Konstruktivismus entscheidend.	
<input type="checkbox"/>	<input type="checkbox"/>	Der Begriff der "Legitimität" ist für den Konstruktivismus entscheidend.	
		<input checked="" type="checkbox"/> Antwort speichern	Nächste Frage >

Frage 4: Matrix

Gütekriterien - Matrix		3 Punkte	Nicht beantwortet
Die folgende Tabelle listet Gütekriterien für sozialwissenschaftliche Messungen und Eigenschaften dieser Gütekriterien auf.			
Markieren Sie bitte durch Häkchen, welche Eigenschaften zu welchen Kriterien gehören.			
	Objektivität	Validität	Reliabilität
Eine Messung misst das Konstrukt, das auch tatsächlich gemessen werden soll.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eine Prüfung des Kriteriums kann erfolgen, indem mehrere Personen den gleichen Zustand messen und die Ergebnisse verglichen werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Wenn der gleiche Zustand erneut vermessen wird, ist das Messergebnis gleich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Messung ist von der durchführenden Person unabhängig.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eine Prüfung des Kriteriums erfolgt durch wiederholte Messung des gleichen Zustandes und Vergleich der Messergebnisse.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eine Prüfung des Kriteriums erfolgt durch Vergleich des Messergebnisses eines Messverfahrens mit den Ergebnissen anderer Verfahren zur Messung des gleichen Sachverhalts.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		<input checked="" type="checkbox"/> Antwort speichern	Nächste Frage >

Frage 5: Drag and Drop

Anmerkung:

Bei dieser Zuordnung werden Studierende angehalten, eine Rangfolge zu bilden. Die Zuordnung kann aber beispielsweise auch auf Basis inhaltlicher Aussagen erfolgen.

Fiktive Ministerien 1 - Drag and Drop 2 Punkte Nicht beantwortet

Ordnen Sie die folgend genannten (fiktiven) Ministerien nach der zeitlichen Reihenfolge, in der sie wahrscheinlich entstanden wären.

Ziehen Sie dazu die gelben Kästchen in die blauen Felder links.

Auf die 1 ziehen Sie das Ministerium, dass zuerst entstanden wäre, auf die 3 das Ministerium, welches zuletzt entstanden wäre.

Ministerium für Bahn und Post

Sozialministerium

Ministerium für Inneres

1

2

3

✓ Antwort speichern Nächste Frage >

Frage 6: Reihenfolge

Anmerkung:

Bei diesem Fragetyp muss eine Reihenfolge erstellt werden. Da bei diesem Frageformat die Antworten allerdings beliebig in das blaue Feld gezogen werden können, ist eine genaue Beschreibung über die Zuordnung in der Aufgabe obligatorisch (siehe unten).

Fiktive Ministerien 1 - Reihenfolge 2 Punkte Nicht beantwortet

Ordnen Sie die folgend genannten (fiktiven) Ministerien nach der zeitlichen Reihenfolge, in der sie wahrscheinlich entstanden wären.

Ziehen Sie dazu die gelben Kästchen in die blauen Felder links.

Ganz oben sollte das Ministerium stehen, dass Ihrer Meinung nach zuerst entstanden wäre, auf die letzte Position das Ministerium, welches zuletzt entstanden wäre.

Ziehen Sie die nicht verwendete Elemente von hier...

Sozialministerium

Ministerium für Inneres

Ministerium für Bahn und Post

Sortieren Sie die ausgewählten Elemente hier...

✓ Antwort speichern Nächste Frage >

Frage 7: True/False bzw. Wahr/Falsch

Normenforschung in der IB Theorie - True/False			1 Punkt	Nicht beantwortet
Bitte markieren Sie die richtige und die falsche Aussage zu Normenforschung in der IB Theorie:				
Unbeantwortet	Richtig	Falsch		
<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Die Legitimation von Normen in der internationalen Politik ist abhängig von sozialer Praxis.	
<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Relevante Normen in der internationalen Politik werden allein durch Rechtsetzung im Völkerrecht bestimmt.	
			<input type="button" value="Antwort speichern"/>	<input type="button" value="Nächste Frage >"/>

Frage 8: Lückentext

Anmerkung:

Lückentext-Fragen finden bei uns aufgrund der aufwändigen Programmierung eine seltenere Verwendung als andere Frageformate und werden nur für Ein-Wort-Antworten genutzt. Um eine automatische Auswertung durch das System zu gewährleisten, werden unterschiedliche Schreibweisen als „gültige Antworten“ hinterlegt. Beispielsweise hier: „Föderalismus“ und „Foederalismus“.

Function Follows Form - Lückentext		2 Punkte	Nicht beantwortet
Tragen Sie die richtige Antwort (nur ein Wort!) in das freie Feld ein.			
Das Motto „function follows form“ steht für welchen Ansatz der politikwissenschaftlichen Analyse der EU?			
Den	<input type="text" value="bitte ausfüllen"/>	.	
		<input type="button" value="Antwort speichern"/>	<input type="button" value="Nächste Frage >"/>

Frage 9: Numerische Frage

Anmerkung:

In diesem Beispiel sehen Sie außerdem die Rückmeldung über die ungültige Eingabe (rot), die den Studierenden zurückgemeldet wird und somit verhindert, dass die Punkte durch ungültige Eingaben verloren gehen.

Artikel im GG - Numerische Frage		2 Punkte	Achtung
In welchem Artikel des Grundgesetzes ist geregelt, wie der Bundesrat an der Beschlussfassung des Bundestages zu beteiligen ist?			
Schreiben sie bitte lediglich die Nummer des Artikels auf, also etwa „1“, wenn Artikel 1 die richtige Stelle wäre.			
Richtige Antwort:			
<input type="text"/>	Ihre Eingabe muss einen Zahlenwert haben, Dezimalzahlen mit Komma oder Punkt.		
		<input type="button" value="Antwort speichern"/>	<input type="button" value="Nächste Frage >"/>

Frage 10: Hottext

Anmerkung:

In diesem Beispiel wurden immer nur „richtig“ und „falsch“ als mögliche Antworten angeboten. Es besteht darüber hinaus die Möglichkeit, auch andere Begriffe als mögliche Antworten zu definieren.

Regieren in Deutschland - Hottext	2.5 Punkte	Nicht beantwortet
Der folgende Textausschnitt enthält Aussagen zum Thema "Verfassung der Bundesrepublik Deutschland" sowie zum Thema "Kommunen". Klicken Sie bitte an, ob der genannte Inhalt "richtig" oder "falsch" ist.		
Die Bundesrepublik Deutschland ist laut Verfassung ein demokratischer <input type="checkbox"/> richtig <input type="checkbox"/> falsch, sozialer <input type="checkbox"/> richtig <input type="checkbox"/> falsch und partizipativer <input type="checkbox"/> richtig <input type="checkbox"/> falsch Staat. Die Kommunen als Teil des politischen Mehrebenensystems der Bundesrepublik Deutschland haben eine eigenständige in der Verfassung abgesicherte Rolle <input type="checkbox"/> richtig <input type="checkbox"/> falsch und verfügen über eine eigene autonome Staatlichkeit <input type="checkbox"/> richtig <input type="checkbox"/> falsch.		
<input checked="" type="checkbox"/> Antwort speichern		Nächste Frage >

Frage 11: Hotspot

Anmerkung:

Hier markieren die unterschiedlichen Kreise die Antwortmöglichkeiten. Bei diesem Fragetyp kann aber je nach Grafik auch eine andere Form gewählt werden, um so ein genaues Anwählen der Antworten durch die Studierenden zu garantieren.

Policyprozess und Steet-Level-Bureaucracy - Hotspot	3 Punkte	Nicht beantwortet
Für die Beschreibung welcher Phase des Policyprozesses spielt der Begriff "street level bureaucracy" eine wichtige Rolle? Bitte markieren Sie die entsprechende Phase in der Abbildung. (Abbildung aus Blum/Schubert 2009: Politikfeldanalyse, S. 102)		
Schaubild 5-1: Der Policy-Cycle		
<pre>graph TD; A((Politik-Terminierung)) --> B((Problem (Re)Definition)); B --> C((Agenda Setting)); C --> D((Politik-Formulierung)); D --> E((Implementierung)); E --> F((Politik-Evaluierung)); F --> A; F --> B; F --> C; F --> D; F --> E;</pre>		
Quelle: Jann/Wegrich 2009		
<input checked="" type="checkbox"/> Antwort speichern		