

Erfahrungen mit Online-Prüfungen und zukünftige Entwicklungen: technische Umsetzungen und emotionale Aspekte

Ruedi Seiler, Dirk Langemann

integral learning, Technische Universität Braunschweig

Stud.IP-Tagung, Göttingen, 13./14. September 2023

Überblick

1. Organisatorische Informationen
2. Erfahrungen der Teilnehmer:innen und der Lehrenden
3. Spezifische Aspekte von online-Prüfungen
 - Aufgabenstellungen und Kompetenzen, Vergleichbarkeit
 - Technische Ausrüstung und technische Schwierigkeiten
 - Betrug, Aufsicht, rechtliche Fragen
4. Emotionale Aspekte
 - Individualität, Distanz und Information
 - Finalisierung von Entscheidungen

Igm, R. Seiler, Online-Prüfungen im Fach Math. mit integral-learning an der TU BS: Erfahrungen und Entwicklungen, *Mathematikinformation* **78**, p. 33-53, 2023

Igm, R. Seiler, Online-Prüfungen, *Forschung & Lehre* **6**, p. 420-421, 2023

Rückblick

- Distanz-Regeln während der Corona-Epidemie \rightsquigarrow digitale Lösungen
- online-Prüfungen an der TU Braunschweig u. a. mit integral-learning
- integral-learning GmbH: vor 10 Jahren Spin-off der TU Berlin, Lernplattform MUMIE, Einbindung in Moodle, Ilias, Stud.IP, ...
- langjährige Erfahrung mit online-Kursen, oft Zusatzmaterial
- on-Campus-Prüfungen an der TU Delft, \approx 1000 Teilnehmer:innen
- **Literatur:** technische Machbarkeit, Best-Practice-Beispiele, allgemeine Digitalisierung, Randomisierung von Aufgaben, online-Kurse, Recht: Chancengleichheit vs. Privatsphäre
- widersprüchliche Aussagen tzum Lernerfolg
- widersprüchliche Aussagen zur Vergleichbarkeit
- Nachteile dominieren öffentlichen Diskurs

Online-Prüfungen an der TU Braunschweig

- vorher OER-Aufgaben auf MUMIE als Zusatzangebot
- ab WiSe 2020/21 ca. 130 online-Prüfungen mit ≥ 8000 Studierenden
- vorrangig in Ingenieur- und Naturwissenschaften
- Begleitung durch IL: Konzeption und Erstellung der Aufgaben, Durchführung der Prüfung, Behandlung von technischen Problemen, (automatische) Korrektur, Einsicht, Archivierung
- Feb./ März 2021: 41 Prüfungen, 3031 Studierenden (4000 Anmeldungen bis Vortag), **sehr viele** Abmeldungen, s. Freichuss- und Verbesserungsregelungen
- größte Prüfung Ingenieurmathematik mit 736 Teilnehmer:innen

Erfahrungen der Prüfungsteilnehmer:innen

- eher positiv mit Einschränkungen
- durch direkte Rückmeldung in z. B. Sprechstunden, Chat, durch schriftliche Evaluation, durch Fragen im Umfeld der Prüfung, über Dritte (Prüfungsämter, Studiengangskoordinator:innen, Anwalt)
- Unsicherheiten zum technischen und organisatorischen Ablauf
- Nachfragen zu denkbaren technischen Unregelmäßigkeiten
- Sorgen wegen denkbaren Internet-Ausfällen
- Schwierigkeiten mit Web-Formularen, z. B. zur Anmeldung
- Zeitknappheit

Trotz retrospektiver positiver Bewertungen wurde die **Sorge** um die technischen und organisatorischen Abläufe, z. B. Räume, immer wieder geäußert, vgl. Distanzierung und Entfremdung.

Erfahrungen der Lehrenden

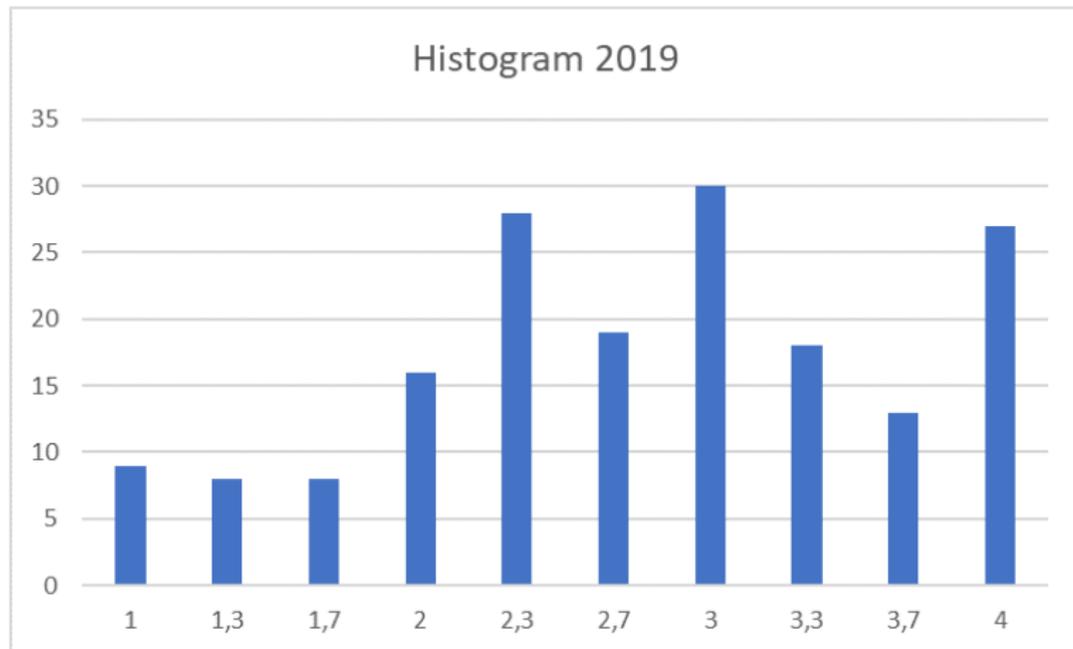
- gesteigener Kommunikationsbedarf
- zeitliche Verlagerung des Arbeitsaufwands
- neue Anforderungen an Aufgabenstellungen
- scheinbare Objektivität knapper Entscheidungen
- Schnelligkeit und objektive Leistungsbeurteilung positiv

- fehlende Identitätsfeststellung
- Angst vor Internetrecherche und vor allem **Betrug**,
s. Buchungsversuche in Hausaufgaben-Hilfe-Portal,
s. Zusammenarbeiten in What's-App-Gruppe

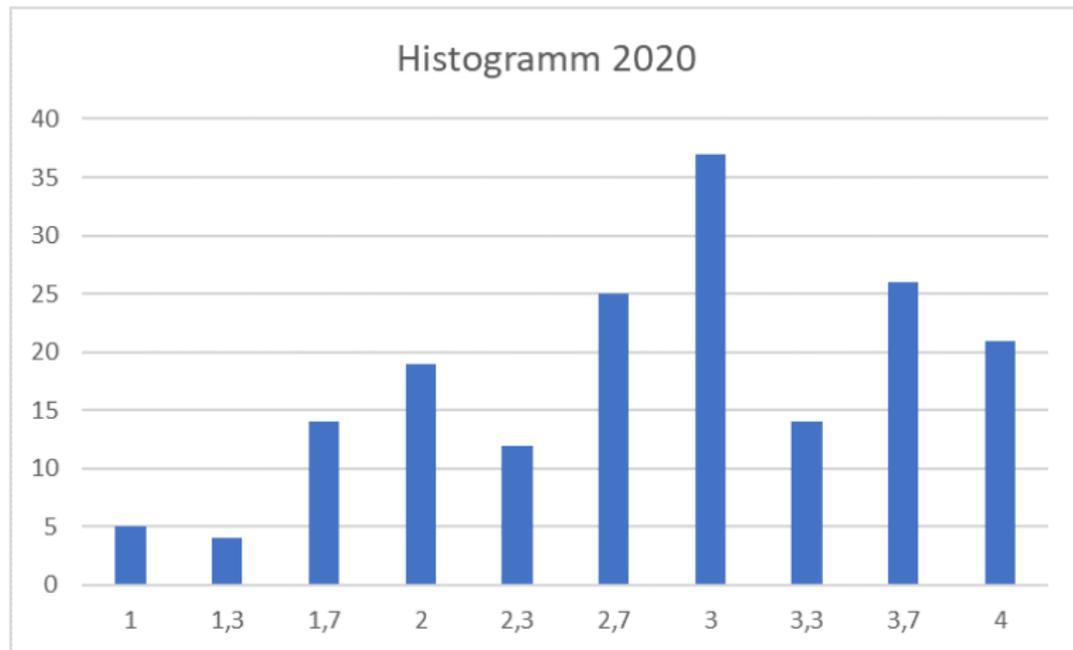
Aufgabenstellungen

- offline: als freie Entscheidung gemäß Vermittlung empfunden
- online (Epidemie): als aufgezwungene Entscheidung empfunden
- online-Prüfungen eher eine Klasse von Prüfungsformen
- teilweise intensiver Diskurs zu abprüfbaren Kompetenzen
- sehr variabel
 - Multiple-Choice-Aufgaben, variable Korrekturmodi
 - Rechenaufgaben mit Zahlen oder algebraischen Ausdrücken, Zwischenergebnisse, Entscheidung über Struktur, Verzweigung
 - Lückentexte mit gegebenen Möglichkeiten oder Freitextfeldern
 - interaktive Graphiken, z. B. Ablesen von Eigenschaften
 - freie handschriftliche Aufgaben ohne automatische Korrektur
- umfassend parametrisierbar und randomisierbar
- zahlreiche kleine Fragen zum Nachweis von Verständnis
- realistische Aufgaben, nahe am Arbeitsalltag

Vergleichbarkeit der Ergebnisse



Vergleichbarkeit der Ergebnisse



Technische Voraussetzungen

- **leicht erfüllbar** für Teilnehmer:innen
- heimischer Laptop oder PC
- besser nicht mit dem Mobiltelefon oder Tablet
- Scan-App für auf Papier bearbeitete Aufgaben

- leicht erfüllbar für die Lehrenden
- Web-Zugang, Aufgaben in WebMiau, vgl. LaTeX

- anspruchsvoll bzgl. der Server
- gleichzeitiger Zugriff am Beginn und Ende
- große Datenmengen

Technische Durchführung

- Vorbereitung: Planung der Zuständigkeiten, Informationstexte
- Aufgabenerstellung, Planung des gesamten Ablaufs
- Aktivierungslink und Test: **erste große Hürde**

- Registrierung und Testseite zu Elementen: **Erinnerungsmails**

- hohe Server-Belastung zur Einwahl
- inhaltlicher und technischer Support während der Prüfung

- automatische Korrektur, Veröffentlichung, Einsicht
- Archivierung der Daten, s. Serverkapazität, Cloud und Testlauf
- Anonymisierung und Löschung

Automatische Korrigierbarkeit

- nur algorithmische Entscheidbarkeit notwendig
- implementierbar im WebMiau
- gegeben, wenn Musterlösung möglich

Technische Ausfälle

- erhebliche Sorge
- schnellstmögliche Meldung erforderlich, aber vgl. Voice-over-IP
- leicht vortäuschbar, z. B. Totalausfall
- unbestimmter Begriff, s. heimisches W-LAN

- Prüfungsordnung: Prüfung gilt als nicht angetreten
- trotzdem erhebliches Konfliktpotential

- **Realität:** (SoSe 2021, WiSe 2021/22)
insgesamt 600 Support-Anfragen an IL von 8000 Studierenden
- davon 50 Anfragen zu technischen Dingen, davon 1/4 Probleme
- 6% technische Anfragen, $\leq 2\%$ technische Probleme, vgl. offline

Technische Ausfälle II

- **Beispiel:** WiSe 2020/21: ca. 300 Anfragen
 - 110 im weitesten Sinne technisch
 - 87 mit Bedienfehlern: 9 falsche Prüfung, 17 falsche Mail, 36 Link im Spam, 5 Passwort vergessen, 5 Bildformate
 - 8 Korrektur-Probleme, kurzfristig behebbar
 - 6 Hochladen von Bildern, nicht reproduzierbar
 - 2 Aufgabe lädt nicht, keine Rückmeldung seitens der Studierenden meist in ≤ 5 min bearbeitet
 - **1 ernstes:** 3 min vor Ende 10 % ohne Zugriff \rightsquigarrow mündliche Prüfung
- technische Probleme wesentlich seltener als angenommen
- **aber** Teilnehmer:innen überschätzen heimisches W-LAN
- W-LAN-Upload langsam, mehrfache Versuche
- Upload unverkleinerter Bilder, 20 MB pro Bild \rightsquigarrow bis zu 150 GB

Betrugsversuche

- Proctoring stößt an praktische und sittliche Grenzen
- auch in Präsenz Ausschluss von Betrug schwierig
- Verhinderung von Absprachen Gebot der Fairness, z. B. durch randomisierte und parametrisierte Aufgaben, aber als unterschiedlich empfundene Schwierigkeit begrenzt Variation
- Upload handgeschriebener Notizen, telefonische Nachfragen, Vergleich mit Vorleistungen, Ehrenerklärung, . . .

Nachforschungen im Internet

- on-Campus verhinderbar
- off-Campus kaum ausschließbar, s. zweites Endgerät
- wähle Aufgabenstellungen, die auch mit Internet herausfordernd sind
- Zusammenhänge ebenfalls nachschlagbar, aber bei Unwissen weniger schnell umsetzbar, vgl. literarische Texte
- nur wenig Faktenwissen in modernen Prüfungen
- unabhängiger Test der Schwierigkeit bei zugänglichem Internet
- **aber** www.ableitungsrechner.de u. ä.
- reine Übernahmen nicht bepunktet, ggf. als Notlösung aufgetreten
- relativ knappe Prüfungsdauer, vgl. Erfahrungen

Rechtliche Aspekte

- Persönlichkeitsrechte vs. Gleichheit, valide Prüfung
- Zulässigkeit von KI-Systemen fragwürdig
- uneinheitliche Rechtsprechung, s. föderale Struktur
- häusliche Umgebung zur Prüfung vorbereitbar \rightsquigarrow andere Balance
- Identität der Teilnehmer:innen ??
- technische Ausfälle unabhängig vom Verschulden, kaum Missbrauch
- Regelung zu MC-Aufgaben, Adaption an Durchschnitt:
bestanden bei 22 % unter Durchschnitt der MC-Punkte,
bei online-Prüfungen häufige Absenkung

Raumproblem: offline und online

- Finden des Raums teilweise vergleichbar mit offline-Raumsuche
- online-Verfehlen facettenreicher und weniger kontrollierbar
- Informationen zur online-Prüfung über einen ähnlichen Kanal
- Zurückgeworfenheit auf sich selbst
- umfangreicher und wiederholter Support nötig
- **Illustration**
 - Bushaltestelle
 - Haltestelle am Stadtrand von Neapel
 - abendliche Haltestelle am Stadtrand von Neapel
 - abendliche Haltestelle am Stadtrand von Neapel, kein Bus
 - ich bin allein
 - gegenüber sitzt jemand in einem unbeleuchteten Fahrzeug
- Bin ich richtig?

Internet-Formulare

- Institution A erstellt \neq Personen B_1, B_2 bearbeiten (verantwortlich)
- Abtrennung der Bearbeitenden B_i vom Expertenwissens bei A
- empfundenes und tatsächliches Fehlen des Feedbacks
- unveränderlicher Rahmen
- Kontrollverlust
- akzeptiert: Bahn, Shopping, Alltags-Banking, Ferienwohnungen, ...
- angstbesetzt: (Grund)-Steuer, Versicherungen, digitale Maut, digitale Forschungsanzeige, ...

Finalisierung von Entscheidungen

- finale Bestätigung mit jeder Eingabe erforderlich
- zusätzliche Entscheidung: Entschlussfreude und Selbstbewertung
- bei schriftlichen und mündlichen Prüfungen weniger empfunden: toleranter gegen kleine Fehler
- Eingabemasken noch nicht tolerant gegen kleine allgemeine Ungenauigkeiten
- Finalisierung wirkt wie kritische Nachfrage oder öffentliches Eintreten, \neq sehr wohlwollendes Feedback in der Schule
- Unwohlsein mit Finalisierung unabhängig von der objektiven Alternativlosigkeit des ermittelten Ergebnisses
- Verantwortlichkeit auf die Seite der Studierenden verlagert
- ähnlich bei alltäglichen online-Formularen

Kommunikation

- kaum neue, rein virtuelle Kommunikationsformen
- Rückgriff auf etablierte Techniken: Fortführung von Kontakten, eher praktische Diskussionen, bestehende Beziehungen
kaum Kontakthanbahnung, kaum unverbindliche Nebengespräche
- niederschwelliges Nachfragen in der Gruppe erschwert
- Lehrperson kommuniziert mit vielen einzelnen Studierenden, weniger mit der Gruppe \neq Präsenz
- wesentliche Elemente des Campus-Lebens nicht abgebildet

Online-Kommunikation ist möglich, und es gibt viele gelungene Beispiele. Aber während der Corona-Epidemie haben viele das Fehlen des weniger benennbaren kommunikativen Miteinanders beklagt.

Vor- und Nachteile von online-Prüfungen

- Nachhaltigkeit, spätere Verwendung, gleiches Niveau
- Objektivität und Schnelligkeit der automatische Korrektur
- neue Aufgabenformate möglich, z. B. interaktive Graphiken
- berufsbegleitende und nebenberufliche Studiengänge

- höherer Aufwand bei der Konzeption und Erstellung
- komplexer Ablauf
- Qualitätskontrolle, Betrugsversuche
- Unsicherheitsempfinden
- Technik induziert Stress

Der Mensch ist gegenüber der Technik auf sich allein gestellt. Die Illusion einer gemeinschaftlichen Lösung praktischer Probleme ist aufgelöst.

Zukünftige Entwicklungen

- KI-Systeme zur Überwachung und zur automatischen Korrektur
- z. B. GradeScope zur Handschrifterkennung bei freien Aufgaben
- online-Prüfungen immer näher an Präsenzprüfungen
- wachsende Akzeptanz

Zusammenfassung

- online-Prüfung: Klasse von Prüfungsformen
- technisch vieles handhabbar
- öffentlicher Diskurs: Betrug und Proctoring
- tatsächliches Feedback: Technik-induzierter Stress
- besonders für Lehrveranstaltungen mit online-Angeboten, s. heterogene Studierende, berufsbegleitende Angebote, ...
- Hoffnung auf zukünftig freie Entscheidung der Lehrenden

Danke

Cordula Reisch

Janina Dierkes

Marko Stautz

Antje Klopp

Katja Tüting

Frank Konieczny

Ruedi Seiler

Michael Heimann

Petrus Tan



© L. Prötzel

Technisch ist vieles möglich, aber wir sollten die emotionalen Herausforderungen von online-Prüfungen nicht zu gering schätzen.

- OECD (2020), "Remote online exams in higher education during the COVID-19 crisis", OECD Education Policy Perspectives **6**, OECD Publishing, Paris
- M. Bandtel, M. Baume, E. Brinkmann et al. (2021) Digitale Prüfungen in der Hochschule. Whitepaper einer Community Working Group aus D, A und CH
- B. Chen, M. West, C. Zilles (2018) How much randomization is needed to deter collaborative cheating on asynchronous exams? Proc. 5th ACM Conf. on Learning at Scale **62**: 1-10
- M.R. Santos, V. Richman, J. Jiang (2019) Online teaching: a study for the effectiveness of randomized exams. J. Instructional Pedagogies **22**: 1-9
- C. Modiz, A. Edthofer, A. Körner (2021) Heuristische Untersuchung von Online-Prüfungen an der TU Wien. CEA-Wismar, Asim Tagung
- S. Büchele, M. Liebendörfer, E. Lankeit (2021) Increasing the effect of a remedial mathematics course by switching to an online format during the COVID-19 crisis: evidence from a German university. Teaching Math. and its Appl. **40**: 478-496
- T. Hoeren, Universität Münster (2020): Gutachten zur datenschutzrechtlichen Zulässigkeit von Überwachungsfunktionen bei Online-Klausuren
- Fernprüfungen mit Protorio, TU München (2022)
www.prolehre.tum.de/prolehre/angebote/infrastruktur/e-pruefungen/fernpruefungen/

Literatur II

- Igm, V. Sommer, Große Grundlagenvorlesungen und die Initiierung von inneren Fragen, Proc. Hanse-Kolloquium zur Hochschuldidaktik 2015, WTM Münster, 163-176, 2016
- A. Singh, S. Karayev, K. Gutowski, P. Abbeel (2017). Gradescope: A Fast, Flexible, and Fair System for Scalable Assessment of Handwritten Work. In Proc. 4th (2017) ACM Conference on Learning, Association for Computing Machinery, New York, 81–88
- Handreichung zu Multiple-Choice-Prüfungen <https://www.rwth-aachen.de/global/>
- E. Stern (2012) Jenseits von Lerntypen: Unterschiede als Herausforderung - wie kann Differenzierung aussehen, in Hellbrügge, Schneeweiß (Hrsg.) Kinder im Schulalter
- <https://www.mumie.net> und <https://www.mumie.net/publications>