



UNIVERSITÄT  
LEIPZIG

MMKH / ELAN Netzwerk Landesinitiativen

# Online-Prüfungen erfolgreich gestalten

Prof. Dr. Heinz-Werner Wollersheim



1. Prüfen in der „digitalen Präsenz“
2. Kompetenz-orientiert prüfen: Was allen Prüfungen gemeinsam ist
3. Open-book Prüfungen
4. Digital – Mündliche Prüfung
5. Digital – schriftliche Prüfungen



Bild: Pexels CC0

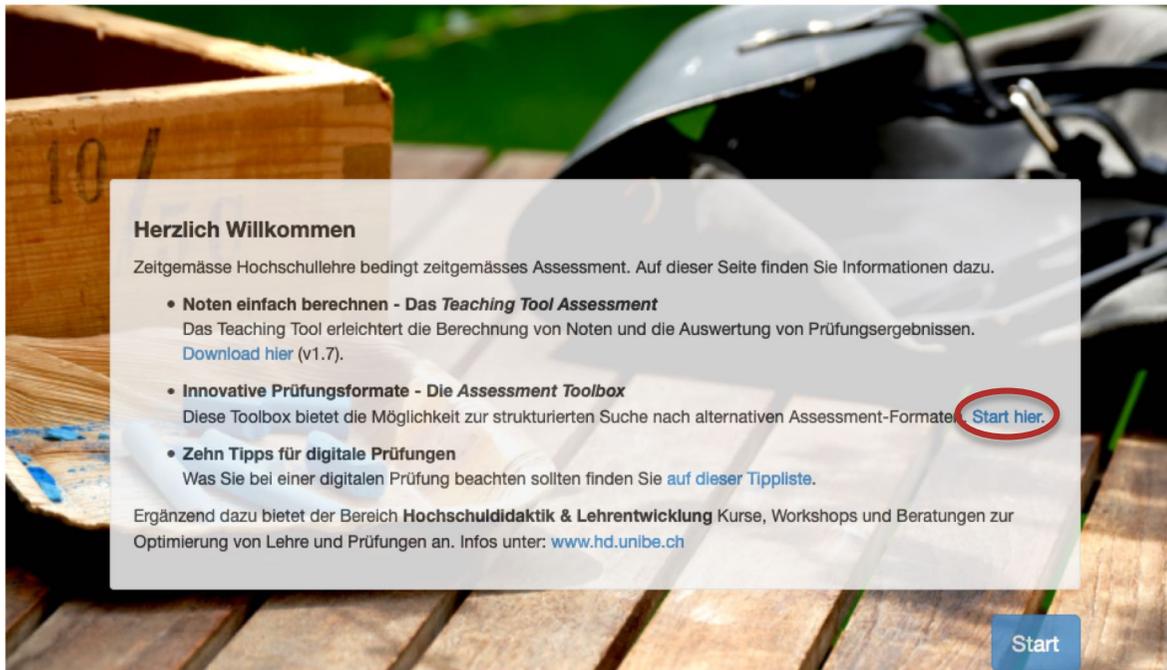
1. Die Notwendigkeit:  
"die Präsenz"

## ONLINE-PRÜFUNGEN ERFOLGREICH GESTALTEN

- Viele tolle Ideen zum digitalen Prüfen.
- sehr hilfreich:

AssessmentToolbox

Hilfe Über uns und die Toolbox



- AssessmentToolbox der Uni Bern:

<https://www.assessment.unibe.ch/>

## ONLINE-PRÜFUNGEN ERFOLGREICH GESTALTEN

Assessment-Toolbox Uni Bern, wählbar:

- Gruppengröße
- Format
- Feedbackquelle
- Taxonomiestufe

liefert Vorschläge für geeignete Prüfungsformate

und Hinweise zur Durchführung

The screenshot displays the 'Auswahl' (Selection) configuration page of the Assessment-Toolbox. It is divided into four main sections, each with a blue header and a list of items with toggle switches:

- Anzahl Prüflinge** (Number of examinees):
  - 1:
  - 2 - 5:
  - 6 - 15:
  - 16 - 30:
  - > 30:
- Ergebnisse** (Results):
  - Schriftliche Darlegung:
  - Mündliche Erörterung:
  - Praktische Handlung:
  - Fachspezifisches Produkt:
- Feedbackquellen** (Feedback sources):
  - Selbstbeurteilung:
  - Lehrperson / Dozierende:
  - Coach / Mentor / Praktikumsleitung:
  - Tutorin / Teaching Assistant:
  - Experte / Assessor:
  - Mitstudierende / Gruppenmitglieder:
  - Patienten / Kundinnen / Auftraggebende:
  - Computer / Software:
- Aufgabentypen** (Task types):
  - Reproduktion (Fakten erinnern):
  - Interpretation (Konzepte/Modelle beschreiben):
  - Anwendung (Verfahren ausführen / Erkenntnisse übertragen):
  - Analyse (Zusammenhänge ermitteln):
  - Evaluation (Kriterienorientiert beurteilen):
  - Entwicklung (Neues entwerfen):
  - Realisierung (Praktische Umsetzung):

Below the configuration options is the 'Assessments' section. It states: 'Leider sind keine Assessments erfasst, welche genau mit Ihren Anforderungen übereinstimmen.' (Unfortunately, no assessments are recorded that exactly match your requirements.)

It then lists 'Folgende Assessments stimmen teilweise mit Ihren Anforderungen überein:' (The following assessments partially match your requirements):

Assessment	Match Percentage
Planspiel	82.5%
Take-Home-Klausuren	77.5%
Projektarbeit mit Dokumentation	75%
Computerized Dynamic Assessment	71.25%
Merkmals- oder Funktionsmatrix	70%
Multiple Choice Prüfung	67.5%



Bild: Pexels CC0

## 2. Kompetenz-orientiert Prüfen:

Was allen Prüfungen  
gemeinsam ist

## WAS BEDEUTET „PRÜFEN“?

Feststellen, inwieweit ein Prüfobjekt eine Forderung erfüllt (DIN 1319).

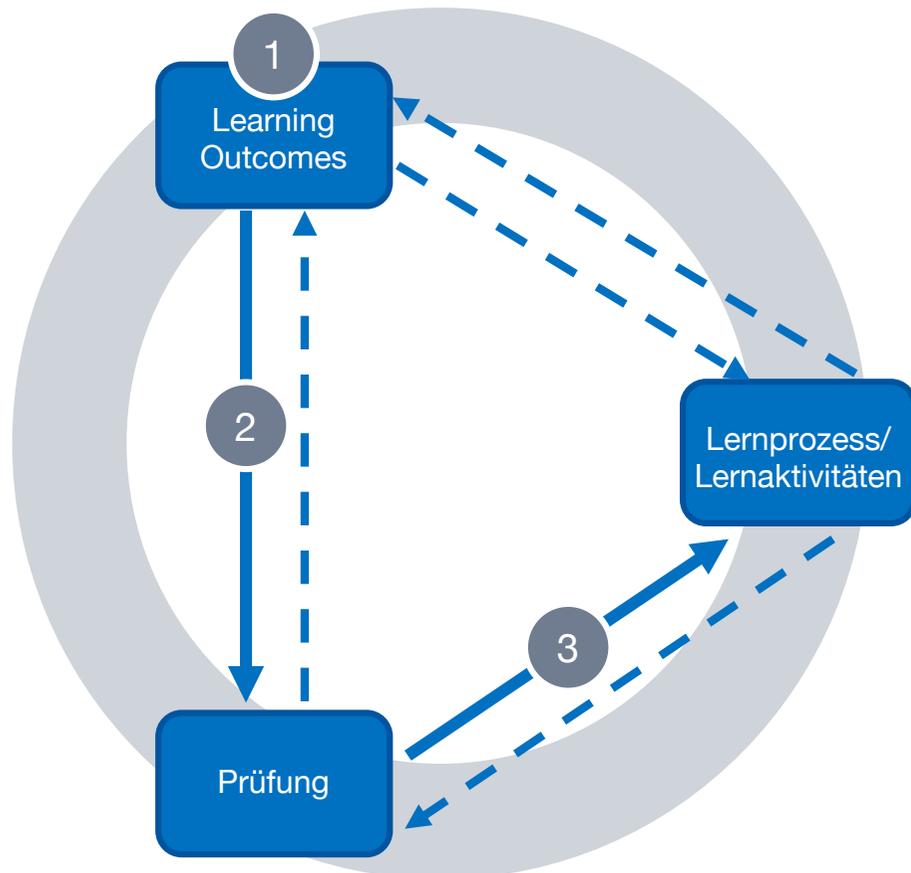
- Prüf - „objekte“: Personen
- „Forderung“: Kompetenz
- Kompetenz(en) operationalisiert durch Learning Outcomes
- „inwieweit“: gradueller Abstufung von Kompetenz modellieren



## „PRÜFEN“ IM AKADEMISCHEN SINN:

- methodisch kontrollierte **Produktion von Beobachtungsdaten**, aus denen man
- methodisch kontrolliert auf die **Erreichung der Learning Outcomes** schließen kann, durch die eine
- **Kompetenz** in unterschiedlichen Stufen operationalisiert wird.

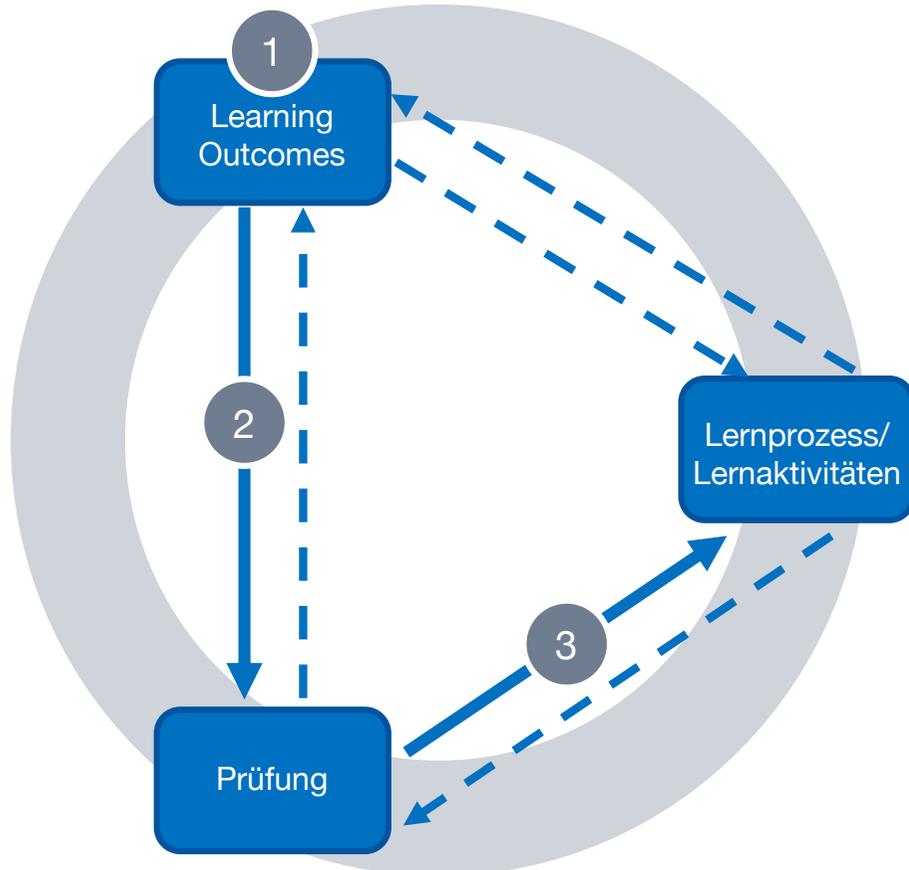
## WAS WOLLEN, WAS SOLLTEN WIR EIGENTLICH PRÜFEN?



### Constructive Alignment: Von den Learning Outcomes zur Prüfung

- Was sollen Studierende können?
- Was will / kann ich prüfen?
- Welche Lernprozesse sind erforderlich?
- Wie kann ich Lernprozesse anregen und begleiten?

## WARUM LEARNING OUTCOMES? – „DIE SOLLEN DOCH EINFACH NUR RECHNEN KÖNNEN“



### 1 Learning Outcomes

- Learning Outcomes modellieren die Zielkompetenz(en)
- Learning Outcomes vor der LV festlegen
- Learning Outcomes beeinflussen Prüfung und Lernprozess

### 2 Prüfungsentwicklung

- Entwicklung der Prüfung ausgehend von Learning Outcomes
- Anpassung von Learning Outcomes möglich

### 3 Lernprozessplanung

- Planung des Lernprozesses in Abstimmung auf LO's und Prüfung

## WARUM DAS WICHTIG IST

„*Learning Outcomes* sind zentrale Elemente, um nicht zu sagen der Dreh- und Angelpunkt einer kompetenzorientierten Studiengangsgestaltung“  
(Schaper 2012, 46)

Die Reflexion der Learning Outcomes kann helfen, Prüfungen zu retten und in anderen als den gewohnten Formaten möglich zu machen.

## WAS HEISST ALSO PRÜFEN:

Worum geht es?

- Ziel muss sein, die **Lösungsqualität** zu beurteilen
- Lösungsqualität ist im Normalfall ein **mehrdimensionales** Konstrukt.

# Lösungsqualität

Definiert durch Prüfer  
nach Kriterien der Scientific Community und  
Anforderungen des Berufsfeldes

## Inhaltliche Kriterien:

- Richtigkeit der Darstellung (Begriffe, Konzepte, Prozesse)
- Logik der Argumentation,
- Rechenweg
- Aussprache
- Hörverstehen

## Kommunikative Kriterien:

- Notation
- Fachsprachlichkeit
- Zitation
- Adressatenorientierung
- Rhetorik
- Medieneinsatz

## Kreativität der Lösung:

- Innovation
- Originalität
- ...

## Weitere Kriterien:

- ...
- ...
- ...
- ...

## WAS HEISST ALSO PRÜFEN:

Worum geht es?

- Ziel muss sein, die **Lösungsqualität** zu beurteilen
- Lösungsqualität ist im Normalfall ein **mehrdimensionales** Konstrukt.

Wie lässt sich das erreichen:

- methodische Gewinnung von vergleichbaren Beobachtungsdaten
- Identische Beobachtungskriterien für jeden Prüfling,  
gleichförmige Auswertung  
gleicher Bewertungsschlüssel

## LÖSUNGSQUALITÄT BEURTEILEN

### Vorgehen im praktischen Fall:

1. Festlegung von Beobachtungskriterien
2. Definition einer Rating- (Beurteilungs-)Skala  
(Kriterium: gute Handhabbarkeit, 0 – 4 oder 0 - 5)
3. Kombination der Skalen zu einem Beobachtungsinstrument
4. Festlegung der Gewichtungsfaktoren in einzelnen Kriteriumsdimensionen
5. Verrechnungsmodell: gewichtetes Mittel
6. Interpretation im Bewertungsmodell

## BEOBACHTUNGSKRITERIEN FESTLEGEN, SKALEN BAUEN

Kriterium	%	0	1	2	3	4
Inhalt, Richtigkeit	30				x	
Inhalt, Struktur	20					
Vollständigkeit	-			x		
Klarheit	-				x	
...	-					
Argumentation, Logik	15				x	
Fachsprachlichkeit	15					x
weitere Kriterien ...	?					

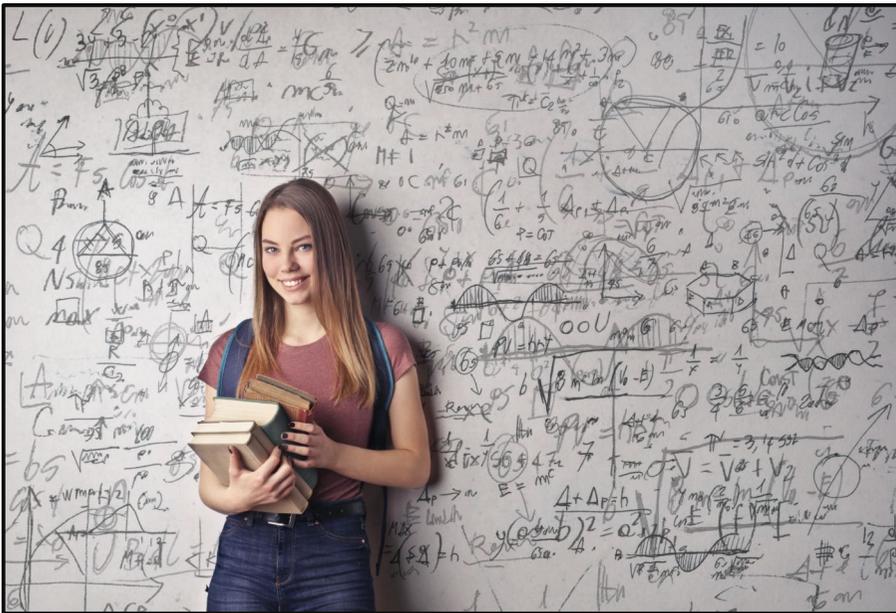


Bild: Pexels CC0

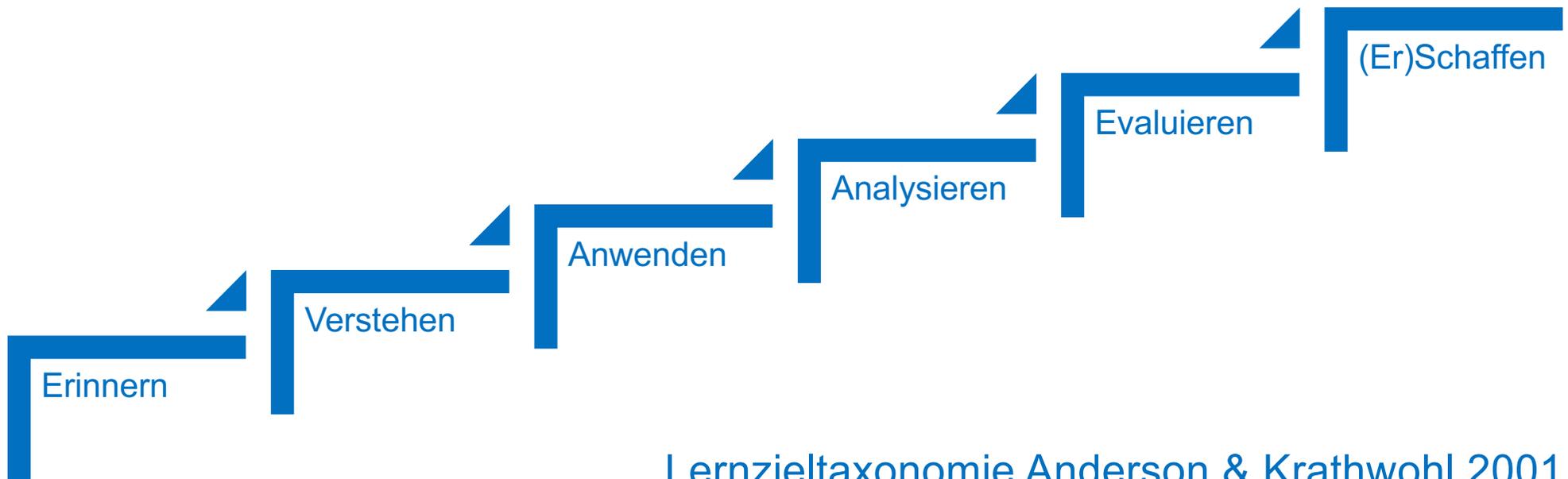
# 3. Open-book Prüfungen

## — ONLINE-PRÜFUNGEN ERFOLGREICH GESTALTEN

- (Manche) Prüfer:innen sorgen sich um Eigenständigkeit der Prüfungsleistung, wenn Prüfungen nicht unter Aufsicht im Prüfungsraum abgelegt werden.
- Kooperation als sozialer Störfaktor?
- Für Prüfungen:
  - „Kosten“ der Kooperation: Zeit
  - Eigenständigkeit in klassischen Prüfungsformaten per Erklärung
- Hoffnung: komplexe, hochwertige Open-book-Prüfungen

Aber:

Jede Prüfung, die außerhalb offizieller Prüfungsräume abgelegt wird, ist eine Open-book-Prüfung!



Lernzieltaxonomie Anderson & Krathwohl 2001  
basierend auf Bloom 1956

## Beschreibung der Niveaustufen durch Superverben

	<b>Stufe</b>	<b>Superverben</b>
1	Erinnern	<i>nennen, aufzählen, beschriften, auswählen, identifizieren,</i>
2	Verstehen	<i>beschreiben (in eigenen Worten), erläutern, erklären, begründen, verallgemeinern, (eigene) Beispiele geben, übersetzen, erörtern, interpretieren</i>
3	Anwenden	<i>durchführen, anwenden, berechnen, lösen, bestimmen, zeigen, übertragen, verwenden, testen</i>
4	Analysieren	<i>analysieren, kategorisieren, unterteilen, vergleichen, gegenüberstellen, unterscheiden, zuordnen</i>
5	Evaluieren	<i>kritisieren, diskutieren, interpretieren, verteidigen, bewerten, beurteilen, entscheiden</i>
6	(Er-)Schaffen	<i>erstellen, modellieren, produzieren, aufbauen, konzipieren, kombinieren, gestalten, entwerfen, konstruieren</i>

## Kooperations-“sensitive“ Niveaustufen bei Distanzprüfung

	<b>Stufe</b>	<b>Superverben</b>
1	Erinnern	<i>nennen, aufzählen, beschriften, auswählen, identifizieren,</i>
2	Verstehen	<i>beschreiben (in eigenen Worten), erläutern, erklären, begründen, verallgemeinern, (eigene) Beispiele geben, übersetzen, erörtern, interpretieren</i>
3	Anwenden	<i>durchführen, anwenden, berechnen, lösen, bestimmen, zeigen, übertragen, verwenden, testen</i>
4	Analysieren	<i>analysieren, kategorisieren, unterteilen, vergleichen, gegenüberstellen, unterscheiden, zuordnen</i>
5	Evaluieren	<i>kritisieren, diskutieren, interpretieren, verteidigen, bewerten, beurteilen, entscheiden</i>
6	(Er-)Schaffen	<i>erstellen, modellieren, produzieren, aufbauen, konzipieren, kombinieren, gestalten, entwerfen, konstruieren</i>

## ONLINE-PRÜFUNGEN ERFOLGREICH GESTALTEN

- Learning Outcomes auf Stufe 1 und 2 sind besonders „kooperations-sensitiv“
- Es gibt Lösungsvorschläge
  - kürzere Bearbeitungszeit
  - Aufgaben erstellen lassen statt lösen lassen
- Deutlich: Problem der Skalierung
  - Kleine Gruppen: Viele Möglichkeiten
  - $N > 30$ : Skalierung problematisch – andere Taxonomiestufen adressieren!



Bild: Pexels CC0

4.

Mündliche  
Prüfungen:

valide(r) und digital  
unterstützt

## MEHR VALIDITÄT

- Strukturierte mündliche Prüfung
- Anspruch: vergleichbare Güte wie schriftlich
- Benötigt: Blueprint in Orientierung an den Learning Outcomes der Lehrveranstaltung (gleiche Chancen und Schwierigkeit für alle TN)
- Aufgabensammlungen (Fallbücher), keine Lehrbuchfragen, keine Beispiele aus der Veranstaltung

## AUCH IM MÜNDLICHEN BEREICH: PRÄSENTATIONEN MIT PEER-ASSESSMENTS



Bild: Pexels CC0

- Einbezug der Studierenden (Ziel: Wissenschaftskommunikation)
- Peer-Review
- Gesamtnote entsteht aus zwei Komponenten:
  - Beurteilung der Präsentationsleistung
  - Beurteilung der Qualität der studentischen Reviews

**Umsetzung? Papier?**

# AUS DER EIGENEN ARBEIT

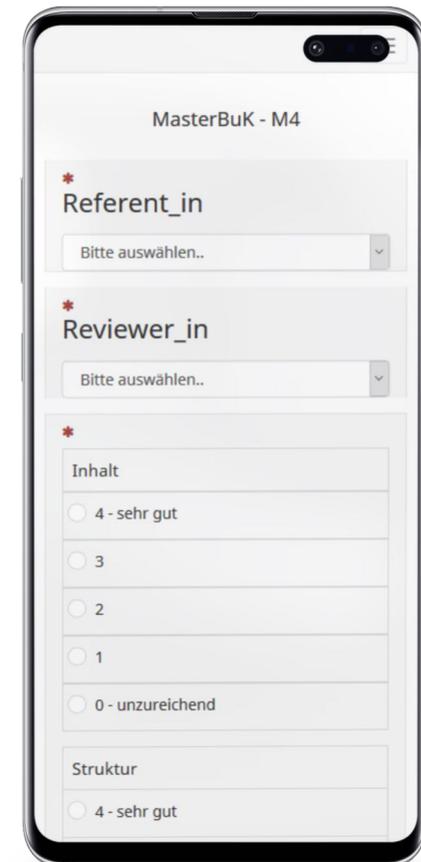
Prüfungspräsentation Master BuK Modul 4									
Name:									
Datum:									
Uhrzeit:									
	Gewichtung	sehr gut	4	3	2	1	0	zu verbessern	
Inhalt	60 %	- sachlich richtig, - Richtigkeit und Vollständigkeit von PSP, PAP und Risikomanagement, - eigener Ansatz erkennbar, - angemessene Gewichtung von Haupt- und Nebenspunkten							- sachliche Fehler, - Lücken in PSP und PAP, - fehlendes Risikomanagement, - wichtige Punkte zu kurz, nebensächliche Punkte zu ausführlich
Struktur		- Gliederung klar erkennbar, - zielorientiert, - Einhaltung des Zeitlimits							- Gliederung intransparent, - nicht nachvollziehbar, - Überschreitung des Zeitlimits
Rhetorik	15 %	Sprache	- Wissenschaftssprache, - verständlich in Satzbau und Wortwahl, - sicher im Ausdruck						- unverständlich, - unständig, - unsicher, - unangemessen
		Sprechweise, Stimme	- deutlich, - angemessen in Lautstärke und Betonung, - variiert						- undeutlich, - zu leise oder zu laut, monoton
		Sprechtempo	- ausgeglichen, - dynamisch, - gute Pausentechnik						- zu schnell, - keine Pausen, - stockend, Blackouts
		Stilmittel	- effektiv, - spannend, - interessant						- eintönig, - ohne Akzente
Körpersprache	15 %	Blickkontakt	- jeder fühlt sich angesprochen, - Vortrag möglichst frei						- fehlt, - unsicher, - von der Vorlage abgelesen
		Gestik/Haltung	- unterstreicht die Aussage offen und freundlich, - wendet sich an das Publikum						- blockiert, - abgewandt, - steif, - übertrieben
		Mimik	- freundlich, - entspannt						- verkrampt
Visualisierung	15 %	- aussagekräftige und übersichtliche Darstellung des PSP und PAP, - klare Bezeichnungen						- keine oder überladene Darstellung des PSP und PAP, - Darstellungen ohne Aussagewert	
Medieneinsatz		- routinierter Umgang mit Technik, - vortragsunterstützend						- nicht leserlich, - hemmt flüssigen Ablauf, - nicht vortragsunterstützend	
Kreativität	10 %	- Übertragung des Inhalts in eine geschickte Form, - Pointiertheit des Kerns							- phantasielos, - löst wenig Zuhörerinteresse aus
Wirkung		- Zuhörer_innen werden erreicht, - Interesse wird geweckt, - auf Fragen wird eingegangen						- ohne Bezug zu den Zuhörer_innen, - Fragen werden unzureichend beantwortet	

Eigene Weiterentwicklung nach: ANAK, GÄTTIG, NORDI, MLENEK, Akademie für Lehrerfortbildung Esslingen

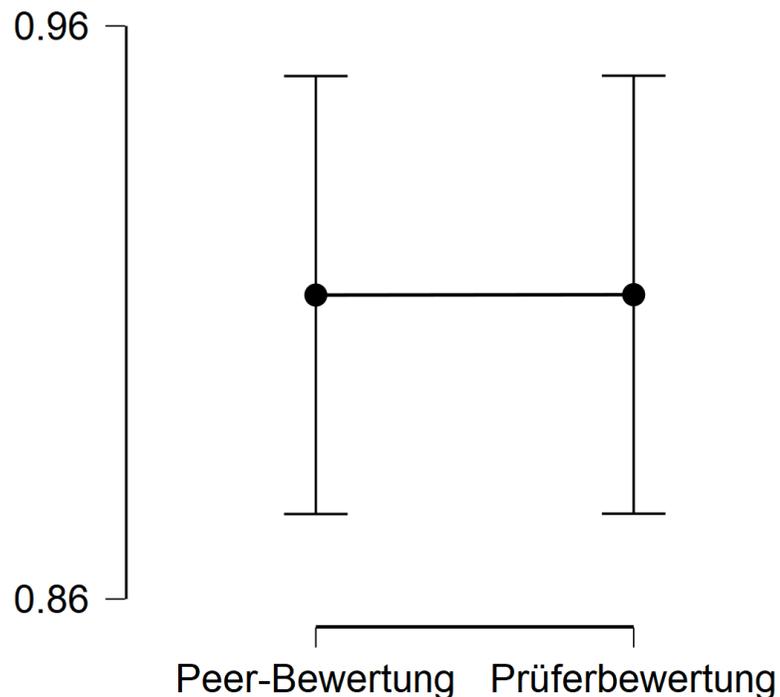
- Der eingesetzte Beobachtungsbogen
- 13 Beobachtungskriterien in vier Gruppen
- einfache Rating-Skala 0 bis 4
- Rohpunkte werden verrechnet zu einem gewichteten Mittelwert
- Kriterien transparent, Bekanntgabe bei Semesterbeginn

## WORKFLOW MIT LINESURVEY & EXCEL

- Studierende bewerten ihre Peers mittels **Onlinefragebogen** auf dem Smartphone
- nach Abschluss der Vorträge werden Daten in ein präpariertes **Excel-Dokument** importiert
- in **weniger als 10 Minuten** ist gesamte Datenauswertung abgeschlossen
- Ergebnisse können noch in der Prüfungsveranstaltung mit den Studierenden diskutiert werden



## UNTERSCHIEDE ZWISCHEN PEER- UND PRÜFERBEWERTUNG



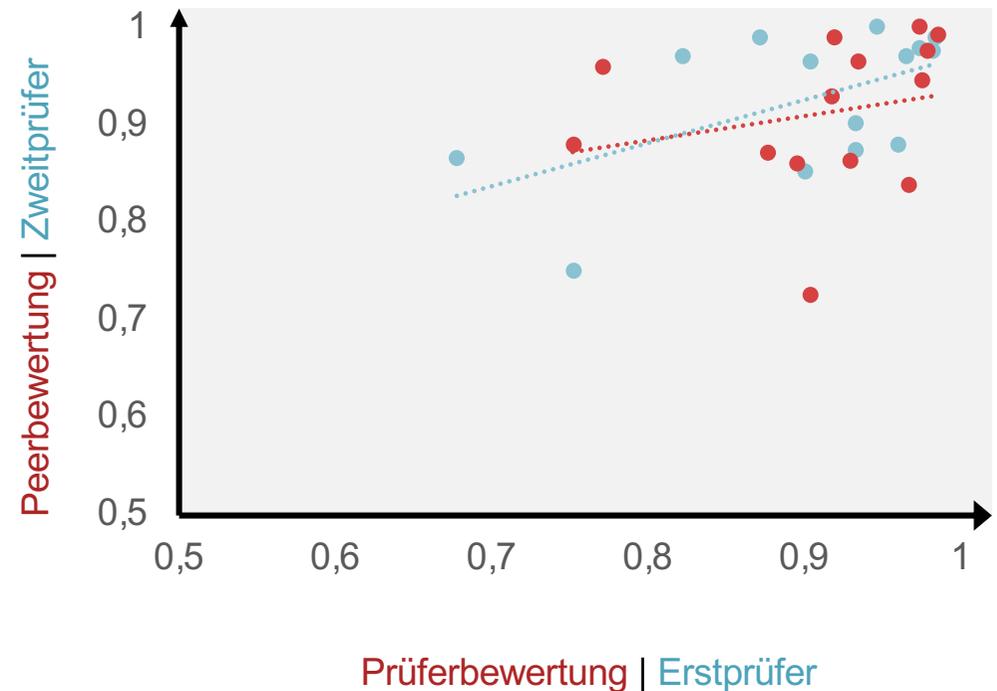
Studierende bewerten einander fair, sofern die Qualität ihrer Reviews in die Endnote eingeht.

Ein relevanter Milde-Effekt tritt nicht auf ( $t(13) = -0.993$ ,  $p = .998$ ,  $d = 0.007$ , 95% CI = [-0.524, 0.523]).

Peer (M = 0.91, SD = 0.021) und Prüferbewertungen (M = 0.91, SD = 0.019) sind praktisch identisch.

## KORRELATION ZWISCHEN PEER- UND PRÜFERBEWERTUNG

- Peerbewertungen und Prüferbewertung korrelieren moderat ( $\rho = .49$ ,  $p = .078$ , 95% CI [-.06, .81]).
- Prüferbewertungen untereinander korrelieren vergleichbar ( $\rho = .54$ ,  $p = .046$ , 95% CI [.13, .83])



## Aufwandvergleich:

- Kein Mehraufwand gegenüber Präsenzprüfung
- 30 Prüflinge erfordern 10 Zeitstunden (3 x 15 Minuten, Pause)



# 5. Digital: schriftliche Prüfungen effektiv unterstützen

## BEOBACHTUNGSKRITERIEN FESTLEGEN, SKALEN BAUEN

Kriterium	0	1	2	3	4
Inhalt, Richtigkeit					
Inhalt, Struktur					
Argumentation, Logik					
Fachsprachlichkeit					
ggf. Originalität & Ökonomie					
...					
...					

## ODER BLUEPRINT FESTLEGEN IM E-ASSESSMENT (E-KLAUSUR)

Aufgabenart	Anzahl	P/Aufg	max erreichb.
Single Choice	50	3	150
Multiple Choice	15	5	75
Zuordnung	5	5	25
Reihenfolge	10	5	50
Lückentext	10	5	50
<b>Punktsumme</b>			<b>300</b>
<b>Bestehensgrenze</b>			<b>150</b>

## DIGITALE UNTERSTÜTZUNG

1. Aufgabensammlungen anlegen, annotieren und verwalten
2. Durchführung der Prüfung auf Distanz (Lernplattform, Einreichung)
3. E-Klausur (nicht in Distanz)
4. Open-book-Prüfung (als Distanzprüfung möglich, Auswertung aufwendig)

- Auch maschinell korrigierbare Aufgaben können höhere Taxonomiestufen adressieren
- nachfolgend ein Beispiel mit Master-Verb „analysieren“, also Level 4
- (Parametrisierten) Text analysieren, Fehler finden
- Systematische Variation der Parameter ergibt nicht 1, sondern  $2^5 = 32$  Aufgaben, die gleichen Schwierigkeitsgrad haben

In einer Zeitungsnotiz über die KMK findet sich folgender Text:

„Die KMK wurde 1952 gegründet. Sie hatte im Nachkriegsdeutschland die wesentliche Aufgabe, das erreichbare Höchstmaß an Mobilität zu sichern. Wegen der Wichtigkeit dieser Aufgabe kann die KMK Beschlüsse nur mit Zweidrittel-Mehrheit. Seit 1990 sind auch die ostdeutschen Bundesländer in der KMK vertreten. Die KMK kann nur Empfehlungen geben, die jeweils nach Landesrecht umgesetzt werden müssen. Dennoch ist die KMK das bei weitem wichtigste Kooperationsgremium der Länder im Kulturföderalismus.“

Dieser Text enthält einen oder mehrere Fehler.

Analysieren Sie den Text. Welche der im Text enthaltene(n) Aussage(n) oder Angabe(n) ist/sind falsch?

1. Gründungsjahr
2. Aufgabenbeschreibung
3. Modus der Beschlussfassung
4. Qualität der KMK-Beschlüsse
5. Aufnahmejahr der O-Länder



Bild: Pexels CC0

## 6. Kurzes Fazit: Online-Prüfungen erfolgreich gestalten

## DIGITAL UNTERSTÜTZTE PRÜFUNGEN

- sind heute bereits vielfältig möglich
- ausgehen von Learning Outcomes: Was sollen Studieren können und wie kann ich mich davon überzeugen, dass und in welcher Qualität sie das können
- Das A & O: methodisch kontrollierte Generierung von Beobachtungsdaten und deren Verarbeitung zu einer Bewertung

## EINSTIEG IN DIGITAL UNTERSTÜTZTE PRÜFUNGEN

- vertraute Arbeitsumgebungen kreativ nutzen
- Aufgabenstellungen anpassen, z.B. nach Fehlern in Berechnungen / Argumentationen suchen lassen, parametrisierte Aufgabenstellungen nutzen
- Kumulativ semesterbegleitend prüfen (z.B. für Fächer, die bislang wöchtl. Hausaufgaben plus Klausur verlangen)

## STUDIERENDE EINBEZIEHEN

- auch Korrekturleistungen bei Peers können zeigen, ob die Ziele erreicht werden
- bestehendes Know-How nutzen: Peer-Assessment-Tool steht zur Verfügung für beliebige Artefakte
- funktioniert nur, wenn die Qualität des studentischen Reviews in die Endnote eingeht