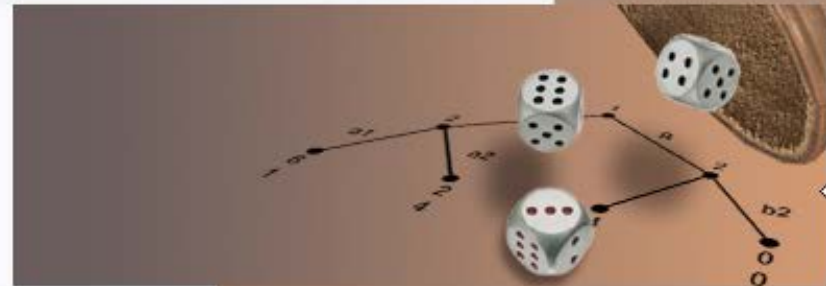


Innovative Prüfungsformate: Peer-Review-Verfahren formativ und summativ

Evelyn Korn
Philipps-Universität Marburg



Verantwortung für den Lernprozess

- Hessisches Hochschulgesetz (§13) sieht Befähigung der Studierenden zu wissenschaftlichem Diskurs, beruflicher Tätigkeit, Übernahme gesellschaftlicher Verantwortung vor.
- Übertragung auf Hörsaal?
- Verantwortung bei Studierenden für (wissenschaftlich fundierten) Schreib- und Leseprozess.

Grundidee des Formats

- Studierende schreiben Kurzaufsätze (bis 600 Worte)
- Sie geben Bewertungen für verschiedene Aufsätze (andere Studierende, den eigenen, ggf. Aufsätze der Lehrenden) ab
- Beide Leistungen gehen in die Reflexion oder abschließende Bewertung ein.
- Je nach – formativem oder summativen – Ziel sind verschiedene Variationen möglich und denkbar.

Reflexion der eigenen Erwartungen

- Was macht einen überzeugenden wissenschaftlichen Aufsatz aus?
- 5 Minuten Austausch mit Nachbar*in

Einbeziehung der Studierenden

- Kennen Ihre Studierenden Ihre Kriterien?
- Entwurf von Schreib- und Leserastern:

<input type="checkbox"/>	Needs Continued Development	Good	Exemplary	Comments
Introductory sentence	No sentence. <input type="checkbox"/>	Introductory sentence. <input type="checkbox"/>	Concise sentence that introduces topic. <input type="checkbox"/>	
Guiding questions/ structure	The planned thread of the essay needs to be more visible for the reader. <input type="checkbox"/>	The planned structure of the essay is visible to the reader; a main research question has been stated. <input type="checkbox"/>	Topic has been broken down to a thread of research questions; these are clearly articulated, aligned to the methods used and well structured. The reader is guided through the essay. <input type="checkbox"/>	
••••	••••	••••	•••••	

Implementierung der Verantwortungsteilung

- „Fachlich fundiert Schreiben“, „Kritisch lesen“, „Kritik fachlich begründen“ als Ziel
- Daher Schreiben und gegenseitige Bewertung als Teil des Semesterablaufs
- Damit auch Teil der Prüfung

Checks vor dem Einsatz des Verfahrens

- Wie viel kann und will ich in die Hand der Studierenden geben?
Was erlaubt die PO?
- Wie flexibel bin ich in der Auswahl des Inhalts der Veranstaltung?
- Welche Kriterien sind „gesetzt“, wo wäre Spielraum, um mit den Studierenden über neue Kriterien nachzudenken?
- Passt die Verantwortungsteilung in meine Sicht auf die Aufgaben von Lehrenden?

Einführung des Verfahrens

- Studierende schreiben einen ersten Aufsatz (20 min. im Hörsaal) mithilfe des Schreibrasters und bewerten ihn mithilfe des Leserasters.
- Wiederholung, Bewertung nun in Partnerarbeit.

Reaktionen der Studierenden

- Erkenntnis über Komplexität von Bewertung
- Sorge über fehlende Kompetenz
- Zögern vor Machtausübung / Empfinden fehlender Legitimation
- Misstrauen in Bewertungsleistung der anderen

Benotung durch Peers (1)

1. Schreibphase: Studierende planen ihren Aufsatz mit Hilfe des Schreibrasters.
Dafür bis zu 15 Punkte von Dozentin (Wunsch der Studierenden)
2. Schreibphase: Studierende schreiben einen Essay mit bis zu 600 Worten.
3. Zuordnung von Peer-Reviewer*innen: Studierenden erhalten jeweils drei Essays von Kommiliton*innen, die Kalibrierungssessays und den eigenen.

Benotung durch Peers (2)

Bis zu 30 Punkte für „Güte“ der Bewertung

4. Bewertungsphase – Kalibrierung: Studierende bewerten drei von der Dozentin verfasste Musteressays auf einer Notenskala von 1-15.
 - Anhand dieser Bewertung wird eine Kalibrierung vorgenommen, mit der die Zuverlässigkeit der jeweiligen studentischen Bewertungen durch einen Qualitätsindex ermittelt wird.
 - Dieser Index errechnet sich aus dem Kehrwert der Summe der Abweichungen der Notenpunkte von der Bewertung der Dozentin für die Kalibrierungssessays.
Studierende erhalten damit einen hohen Index, wenn sie die Kalibrierungssessays ähnlich wie die Dozentin bewerten.

Benotung durch Peers (3)

5. Bewertungsphase – Peer Review:

- Studierende bewerten die Essays der Kommiliton*innen und den eigenen Essay.
- Berechnung der Noten für jeden Essay.
Für jede abgegebene Bewertung wird ein proportionales Gewicht ermittelt, das vom eigenen Qualitätsindex und denjenigen der beiden anderen Bewertenden abhängt.
- Die Notenpunkte für jeden Essay werden aus den proportional gewichteten Noten der Bewertenden ermittelt.

Diese Note als Teil der Bewertung, d.h. bis zu 15 Punkte

Benotung durch Peers (4)

Bis zu 30 Punkte für Nähe am Durchschnitt

6. Bewertung der Peer-Review-Leistung:
Für jeden Essay wird ermittelt, wie nahe die Bewertung eines Reviewers am ermittelten Durchschnitt ist.
7. Selbsteinschätzung: Bis zu 10 Punkte für Nähe am Durchschnitt
Studierende benoten ihren eigenen Aufsatz.
Auch hier wird ermittelt, wie nahe die Bewertung am gewichteten Durchschnitt der von den Reviewer*innen vergebenen Noten liegt.

Für die Bewertung der Review-Leistung wird folgendes Schema genutzt:
10 Punkte für 0-1 Notenpunkt Abweichung von der Durchschnittsbewertung,
5 für 1-3 Notenpunkte Abweichung und
0 Punkte für eine Abweichung von drei oder mehr Notenpunkten.

Beispielrechnung Noten

	Abgegebene Bewertungen für Kalibrierungsaufsätze C ₁ -C ₃ (Noten der Dozentin in Klammern) Aufsatz von Student*in D				Summe der Abweichungen	Qualitätsindex	Proportionale Gewichtung der Qualitätsindizes
	C ₁ (4)	C ₂ (8)	C ₃ (12)				
A	3	6	13	8	1+2+1=4	1/4	$\frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{3} + \frac{1}{7}} = 0,344$
B	5	9	13	11	1+1+1=3	1/3	$\frac{\frac{1}{3}}{\frac{1}{4} + \frac{1}{3} + \frac{1}{7}} = 0,459$
C	1	10	14	14	3+2+2=7	1/7	$\frac{\frac{1}{7}}{\frac{1}{4} + \frac{1}{3} + \frac{1}{7}} = 0,197$
Daraus ermittelte Note für D: $0,344 \cdot 8 + 0,459 \cdot 11 + 0,197 \cdot 14 = 10,56$							
Abweichungen für Reviewer*innen vom Mittel:					A: $10,56-8=2,56$	B: $11-10,56=0,44$	C: $14-10,56=3,44$

Noten im summativen Verfahren

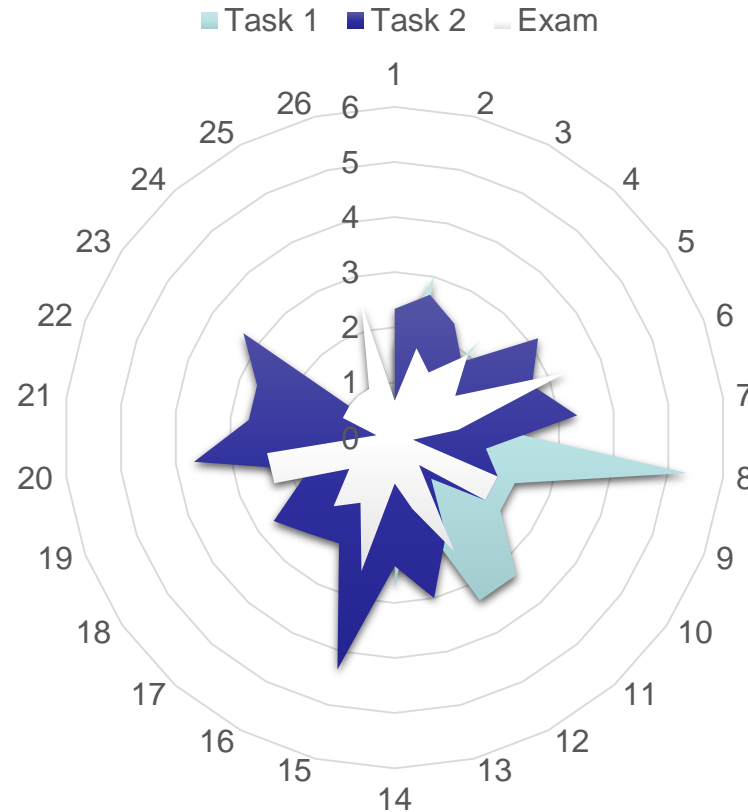
- Studierende konnten damit bis zu 100 Punkte erhalten.
- Abschließend Eingreifen der Dozentin möglich (angekündigt, wenn Qualität aller drei Bewertenden aus Kalibrierung zu schlecht).
- Noten für den Kurs dann auf Verteilung über 100 Punkte – damit auch Korrektur von „zu guten“ Bewertungen möglich.

Qualitätssicherung: Einpassung in Prüfungslogik

Wie soll die Qualitätssicherung erfolgen?

- Wie sieht der Beitrag der Lehrperson aus?
- Ist er für alle sichtbar?
- Für die Sicherung der Noten sind zwei Wege denkbar: Ein Eingreifen der Lehrperson (durch Notenkorrektur, Beratung, Wiederholung der Aufgabe, etc.), wenn keiner der bewertenden Peers eine verlässliche Bewertung der Kalibrierungssessays vorgenommen hat; oder eine Beteiligung der Lehrperson am Bewertungsprozess.

Qualitätssicherung: Erkennbarer Lernfortschritt



Notenabweichungen für Kalibrierungsesays

Einsatz im formativen Verfahren

- Anonymität zentraler Bestandteil des summativen Einsatzes
- In kleineren Gruppen Arbeit mit Rastern immer noch sinnvoll für Reflexion – gegenseitige Bewertung dann möglich, wenn Umgang unter den Studierenden respektvoll.
Das braucht in der Einführung Zeit und gute Begleitung
- Mit Ilias-Tool regelmäßiger formativer Einsatz möglich

Design Ilias

9.1 Erläuterung zu 9.

Falls Sie in 9. "ja" gewählt haben, erklären Sie dies hier bitte kurz

10. Wird emotional aufgeladene Sprache verwendet? Bitte führen Sie ggf. Belege an. *

11. Ist der Essay klar und verständlich geschrieben? *

-- Bitte auswählen --

-- Bitte auswählen --

Ja

Nein

Bitte führen Sie hier an, woran Sie feststellen, dass der Essay klar und verständlich oder unklar und unverständlich geschrieben ist.

12. Wie ist der Essay sprachlich zu beurteilen (Grammatik, Ausdruck, Rechtschreibung) *



Mehr Sterne stehen für gute Sprache

Ausschnitt aus Kriterienkatalog

Chancen und Risiken

- Formativer Einsatz nimmt Druck der Benotung und Misstrauen gegenüber der Leistung anderer Studierender
- Gleichzeitig fehlt ein Anreiz zur ernsthaften Befassung.
- Hier kann Qualitätssicherung auf verschiedenen Wegen hergestellt werden:
 - Teilnahme an Verfahren nur nach erfolgreichem Wissenstest
 - Abgabe von Bewertungen für andere Studierende nur, wenn Kalibrierung hinreichend erfolgreich
 - Stichprobenkontrolle durch Dozent/in

Wo passt welche Variante?

Wie erfahren sind die Studierenden in Bezug auf Schreiben, Lesen, Benoten?

- Für Anfänger*innen:
Welche Form von (kurzen) Schreibübungen sind sinnvoll?
Welche Regeln können klar kommuniziert werden?
Wie lässt sich die Anwendung von Bewertungskriterien einüben?
- Für erfahrene Studierende:
Wie können die Studierenden sinnvoll und glaubwürdig an der Verfahrensentwicklung beteiligt werden?
- Für alle:
Über wieviel Erfahrung mit dem Notensystem verfügen Studierende?
Wie kann Anonymität der Bewertung glaubhaft zugesagt werden?

Rahmenbedingungen

- Einsatz in Masterkurs
 - Studierende mit unterschiedlichen (wissenschafts-)kulturellen Hintergründen — brauchte Zeit für Gewöhnung an Noten
- Anpassungen für weniger erfahrene Studierende nötig
 - Mehr Begleitung durch Dozent*in
 - Arbeit näher am Text
 - Benotung vermutlich schwierig

Checks vor dem Einsatz des Verfahrens

- Wie viel kann und will ich in die Hand der Studierenden geben?
Passt die Verantwortungsteilung in meine Sicht auf die Aufgaben von Lehrenden?
- Welche Teile des bisher vermittelten „Inhalts“ bin ich bereit zu ersetzen?
Das Verfahren braucht Übung im Semester: 40 Minuten für 500-600-Worte-Text, 20 Minuten für jedes Feedback; d.h. pro Durchlauf mit Kalibrierung 180 Minuten für die Studierenden (dieser Teil kann Arbeitszeit zuhause sein).
Es braucht Zeit für Reflexion und offene Fragen.
- Wie wird mit den Studierenden thematisiert, wo die Lehrperson eingreift? Wie wird Vertrauen in den Prozess hergestellt?

Vielen Dank für Ihre Aufmerksamkeit!

Ich freue mich auf die Diskussion.