

ePS2017 - Knowledge Café 1-1

Ergebnisprotokoll der Diskussion

Moderation: Martina Salm

Qualitätssicherung von Testfragen – Teststatistiken auswerten, Fragen und Aufgaben verbessern

Silke Kirberg, Yvonne Fischer

1. Welche Informationen sind für Lehrende interessant?
2. Aufbau der Interpretationshilfen?
3. „Liste“ für Rahmenbedingungen bei der Interpretation
4. z.B. Anzahl der Versuche (summativ vs. formativ), Unterschiede in der Lehrsituation / in der Gestaltung des Selbststudiums
5. Wie kann zu einer Reflexion im Nachgang einer Prüfung und zu einer „qualitätssichernden Überarbeitung“ angeregt werden?
6. Welche Erfahrungen gibt es im Plenum mit der Dokumentation von Prüfungsergebnissen über mehrere Kohorten hinweg?

www.eassessmentnrw.de

1. Durchgang

- ILIAS farbliche Markierungen, Sichtbarkeit für Lehrende
- Idee: Ampelsystem, Balkendiagramme (z.B. Iqul)
- graphische Darstellung , z.B.
- "Wertung" vs. ist die Frage gut /schlecht?
- offen: Auswertung je Test, nicht über mehrere moodle-Kurse hinweg
- in der Handreichung geht es um geschlossene Aufgaben

2. Durchgang

- Was ist der Grund für das schlechte Ergebnis einer Frage? War meine Lehre anders als sonst? Waren die Studierenden besser oder schlechter als sonst?
- Statistiken für eine kritische Selbstreflexion nutzen
- Peer-Review, Kollegen einbeziehen?
- Sollte die Frage "schwierig" sein?
- Trennschärfe - was bedeutet das für diese Frage / welche Intention hatte ich?
- Z.B. alle "Guten" haben eine Aufgabe falsch beantwortet - was ist da passiert?
- Graphische Darstellungen anstelle "nur Werte"
- Prozente wie z.B. Durchfallquoten, wie viele von N haben die Frage nicht beantwortet / falsch beantwortet
- Bedeutung des Kennwertes klar machen, auch ohne Wert verständlich machen
- Ab wann wird eine Statistik aussagekräftig? "Kundige Interpretation"
- Interpretation hängt vom Konzept der Prüfung ab
- Fragengesteuerte Erläuterung der Kennwerte; welche von den umfangreichen Werten sind überhaupt interessant für eine Interpretation?
- z.B. Verteilung

- Gibt es Kennwerte die einen Mindeststandard haben sollen?
- Wie wird das Interesse von Lehrenden geweckt?
- Kohortenvergleich: z.B. wegen Änderung bei der Gestaltung der Lehre - gibt es einen Einfluss?
- Bei Fragenpools wird der Kohortenvergleich leichter
- von Werten zur Reflexion über Inhalte kommen
- Sind die Anzahl der Teilnehmenden groß genug für eine objektive Bewertungen?
- <https://github.com/ilifau/ExtendedTestStatistics> Info zu ILIAS

3. Durchgang

- Werte können Anlass für Reflexion sein, aber wie Lehrende darauf ansprechen?
- Multiplikatoren / Beratungs-Einrichtungen bieten Analyse / Reflexion nach der Prüfung an, bleibt oft nicht die Zeit
- Hilfe wäre: Bedeutung der Werte z.B. Range x bis y = was könnte es bedeuten?
- Zur Interpretation ist notwendig zu wissen, welche Intentionen dahinter steckten. Z.B. ob die Frage leicht sein sollten, ob die Distraktoren zu einfach waren, versteckte Hinweise darin, Hinweise in der Vorlesung,
- "Was könnte sich hinter dem Wert verbergen?"
- interessant: Hinweis geben und in die Aufgabe einstreuen, dann analysieren
- z.B. eine Antwortmöglichkeit in einer MC-Aufgabe wurde niemals ausgewählt - das hinterfragen.
- Haben die "Guten" es erreicht, die insgesamt gut abgeschnitten haben, haben die Studierende , die insgesamt schlecht abgeschnitten haben in einer Frage Gut abgeschnitten? Was könnte der Grund gewesen sein?
- keine Auswertung für offene Fragen, die manuell von Lehrenden bewertet werden
- Erfahrung aus offenen Fragen: Korrektoren "lernen" während der manuellen Beurteilung / Verständnis wächst bei Wiederholungen
- TN-Anzahl einbeziehen