

Revision von eKlausuren auf Ebene der Fragetypen

Rheinische Friedrich-Wilhelms-Universität Bonn
Landwirtschaftliche Fakultät
Institut für Tierwissenschaften
Dr. Stephanie Hiss-Pesch

Bei welcher Herausforderung bzw. welchem Handlungsbedarf setzt Ihr Projekt an?

In Bonn ist es seit einigen Jahren möglich, elektronische Klausuren (eKlausuren) zu schreiben. Das umstellen von Papier-basierten Klausuren auf eKlausuren bietet den Lehrenden bei großen Kohorten viele Vorteile. Neben organisatorischen Vorteilen wie der elektronischen Archivierung von Klausuren, eröffnen sich bei eKlausuren eine große Auswahl an Fragetypen. Darüber hinaus sind auch Bild-Audio- oder Videodateien einsetzbar. Neben den Vorteilen für die Gestaltung der Fragen, ergeben sich auch Möglichkeiten, die Fragen zu evaluieren und zu revidieren. Das war auch bei Papier-basierten Klausuren möglich, setzte aber einen sehr großen Arbeitsaufwand voraus, um Qualitätsindikatoren für die Klausuren und Fragen zu berechnen. In der täglichen Lehrpraxis wurde das meiner Erfahrung nach kaum durchgeführt.

Mit den eKlausuren bekommt der Lehrende neben den Ergebnissen der Prüflinge auch eine detaillierte Teststatistik mit Qualitätsindikatoren für die Klausuren und die Einzelfragen. Diese wertvollen Informationen gilt es zu nutzen bei der Evaluierung und Revision von Klausurfragen. Die Fülle der Informationen ist jedoch eher abschreckend für die Lehrenden und vermutlich die Antwort darauf, weshalb meine Kollegen und ich diese Qualitätsindikatoren nur spärlich genutzt haben.

Wie lautet Ihre Antwort / Ihre Idee darauf?

Die Revision der Klausurfragen auf Ebene der Einzelfragen verhindert das Erkennen von systematischen Schwächen bestimmter Fragetypen oder Fehler bei der Konstruktion bestimmter Fragetypen von einzelnen Lehrenden. Die Revision der Klausuren auf verschiedenen Ebenen kann helfen, diese Schwächen zu erkennen in dem zunächst die Klausuren auf der Ebene der Fragetypen analysiert werden und erst danach die Einzelfragen evaluiert und revidiert werden.

Mit welcher Zielgruppe und welchen Rahmenbedingungen haben Sie es zu tun?

Die Zielgruppe ist in diesem Fall nicht die Gruppe der Studierenden, sondern die Gruppe der Lehrenden, die eKlausuren einsetzen. Es geht im Besonderen um eKlausuren, in denen geschlossene Frageformate eingesetzt werden. Bei geschlossenen Frageformaten handelt es sich um Fragen, bei denen Antwortmöglichkeiten vorgegeben werden. Diese Art von Frageformaten eignen sich für große Kohorten, da sie voll automatisiert korrigiert werden und keine manuelle Nachkorrektur notwendig ist.

In diesem Fall handelt es sich um eKlausuren des ersten Semesters Agrarwissenschaften mit ca. 400 Studierenden.

Welche Lernziele verfolgt Ihr Projekt?

Die Lehrenden nutzen die zur Verfügung stehenden Qualitätsindikatoren für Klausuren zur Revision von Fragen um die Klausuren insgesamt zu verbessern. Die Analyse der Fragen auf unterschiedlichen Ebenen strukturiert die Revision und erlaubt das Erkennen von systematischen Schwächen von bestimmten Fragetypen. Die Fülle der Information wird zielgerichtet genutzt.

Mit welchen Methoden und welcher Vorgehensweise versuchen Sie, diese Ziele zu erreichen?

Für die systematische Revision von eKlausurfragen müssen ausreichend Fragen zur Verfügung stehen. Im vorliegenden Fall stehen Qualitätsindikatoren für fünf Klausuren (WS 17/18 und 18/19) zur Verfügung. Zunächst wird der Qualitätsindikator für die gesamte Klausur überprüft. Das Bonner eKlausursystem (ILIAS) gibt hierfür die Reliabilität als Cronbach's alpha an. Ein Wert von $> 0,8$ sollte erreicht werden (vgl. Möltner et al. 2006). Die Klausuren weisen Werte zwischen 0,7 und 0,8 auf, was bedeutet, dass die Klausuren revisionswürdig sind.

Insgesamt umfassen die fünf Klausuren 115 verschiedene Fragen, die revidiert werden sollen. Diese Fragen lassen sich sechs geschlossenen Fragetypen zuordnen (Abb. 1)

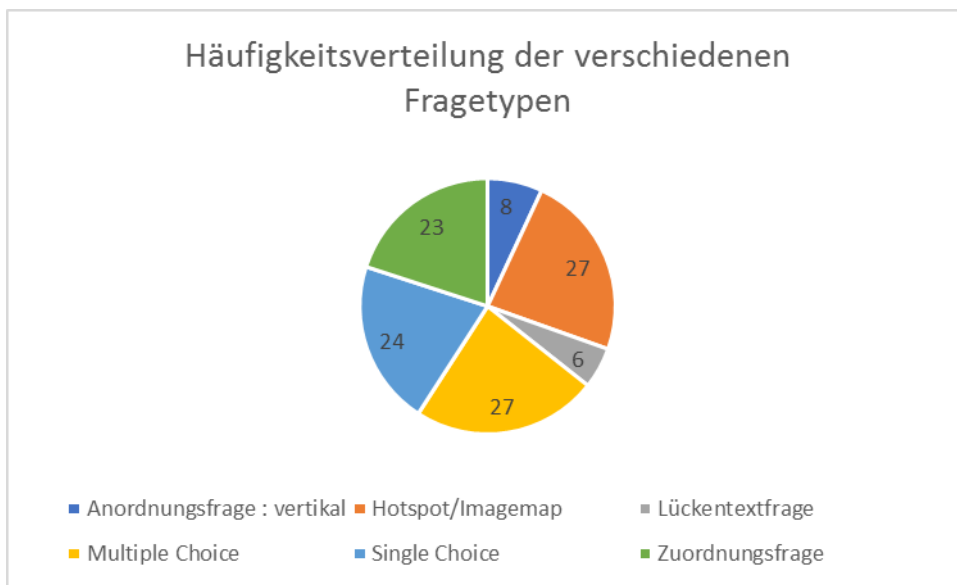


Abb. 1: Verteilung der 115 Fragen auf die verschiedenen Fragetypen

Bei allen Fragen werden zwei Qualitätsindikatoren untersucht, die das eKlausursystem automatisiert berechnet. Der Schwierigkeitsindex gibt die Aufgabenschwierigkeit an und sollte zwischen 40 % und 80 % liegen (vgl. Möltner et al. 2006); d.h. 40 bis 80 % der Prüflinge konnten die Aufgabe richtig lösen. Der zweite Qualitätsindikator, der analysiert wurde, ist die Trennschärfe. Hierbei geht es darum, zwischen guten und schlechten Prüflingen zu unterscheiden. Trennschärfen von $> 0,2$ gelten als akzeptabel (vgl. Möltner et al. 2006). Bei Werten von $< 0,2$ können schlechte Prüflinge die Frage lösen und gute Prüflinge nicht.

Die Fragen wurden zunächst auf Ebene der Fragetypen analysiert. Sowohl beim Schwierigkeitsindex als auch bei der Trennschärfe besteht ein signifikanter Zusammenhang ($p < 0,05$) zwischen leichten Fragen bzw. schlechter Trennschärfe und dem Fragentyp.

Insgesamt sind 38,3 % der Fragen zu leicht (Schwierigkeitsindex > 80 %). Bei den Anordnungs- und Zuordnungsfragen befinden sich mehr als 60 % der Fragen in der Kategorie „zu leicht“ (Abb. 2) Eine schlechte Trennschärfe ($r < 0,2$; Möltner et al. 2006) weisen 24,4 % der Fragen auf. Bei den Anordnungsfragen sind mehr als 60 % der Fragen in der Kategorie „schlechte Trennschärfe“ (Abb. 2)

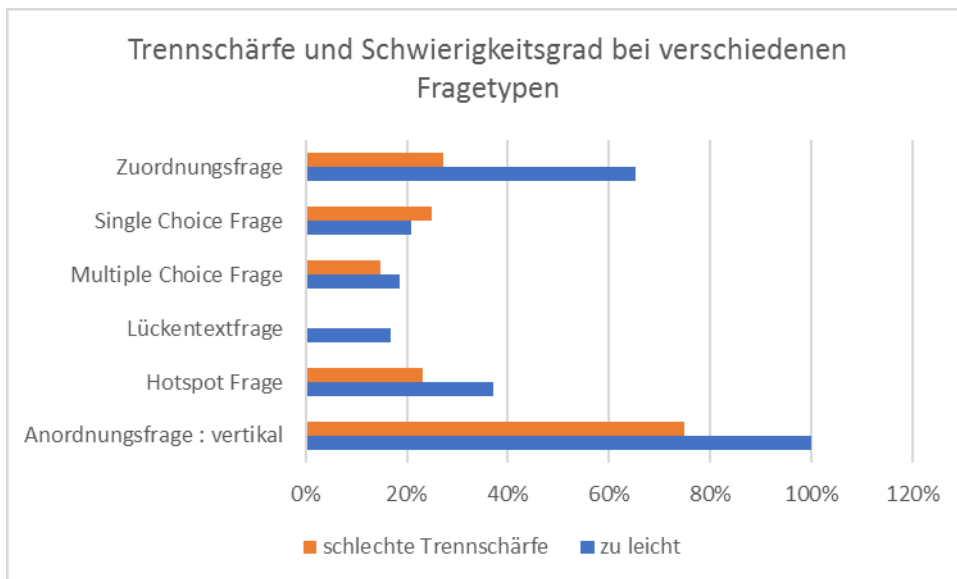


Abb. 2: Trennschärfe und Schwierigkeitsgrad bei den verschiedenen Fragetypen

Durch die Analyse der Fragen auf Ebene der Fragetypen, lässt sich die Flut der zu revidierenden Fragen eindämmen. Systematische Überlegungen bezüglich der einzelnen Fragetypen können mit der Literatur verglichen werden, um Erklärungen und Lösungsansätze zu finden.

Wie lassen sich diese Fragen verbessern, um einen adäquaten Schwierigkeitsindex und eine akzeptable Trennschärfe zu erreichen?

Sind Verbesserungen innerhalb der Fragetypen möglich oder müssen die Fragen in andere Fragetypen überführt werden?

Lösungsansätze:

Anordnungsfrage: Bei den Anordnungsfragen handelt es sich um „Alles oder nichts-Fragen“, d.h. das eKlausursystem lässt keine Teilpunkte zu. Eine Erhöhung des Schwierigkeitsgrades innerhalb dieses Fragetyps führt dazu, dass viele Punkte verloren werden können. Aus diesem Grund werden die Anordnungsfragen in Long-Menu-Fragen umgewandelt. Bei der Anordnungsfrage müssen vorgegebene Begriffe in die korrekte Reihenfolge gebracht werden. Das Umwandeln in eine Long-Menu-Frage bietet die Möglichkeit, die Begriffe nicht vorgeben zu müssen, sondern bei Eingabe der ersten korrekten Buchstaben Begriffe in einem Drop-down-Menu anzubieten. Damit wird der Schwierigkeitsgrad erhöht (Niveaustufe 1 Erkennen/Erinnern \Rightarrow Niveaustufe 2 Erklären). Außerdem können Teilpunkte vergeben werden, was das Erhöhen der Schwierigkeit insgesamt erleichtert und die Trennschärfe verbessert.

Abbildung 3 zeigt beispielhaft eine Anordnungsfrage wie sie im eKlausursystem ILIAS zur Verfügung steht. In Abb. 4 ist die gleiche Frage als Long-Menu-Frage dargestellt, bei der Teilpunkte vergeben werden können und die Begriffe selbst benannt werden müssen.

In der Literatur (einschlägige Handbücher zum Thema Erstellung von eKlausur-Fragen) wurden speziell zum Thema Anordnungsfragen keine hilfreichen Lösungsvorschläge gefunden.

Zuordnung Magen Darm 2

Der Magen-Darm-Trakt des Pferdes: Ordnen Sie die unten stehenden Begriffe entsprechend der Nummerierung in der Abbildung. Beginnen Sie mit Nummer 1 von oben nach unten.



Caecum (Blinddarm)

Colon ascendens (aufsteigender Grimmdarm)

Colon descendens (absteigender Grimmdarm)

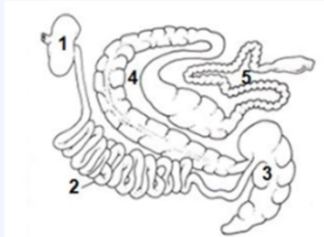
Magen

Dünndarm

Abb. 3: Anordnungsfrage horizontal in ILIAS

Magen-Darm Pferd neu

In der Abbildung ist der Verdauungsapparat des Pferdes dargestellt. Wobei handelt es sich bei den Organen 1-5.



Die bestmögliche Lösung lautet:

Organ 1 Magen oder Gaster oder einhöhliger Magen

Organ 2 Dünndarm oder Jejunum

Organ 3 Blinddarm oder Zäkum oder Cäkum

Organ 4 Colon ascendens oder aufsteigender Dickdarm oder aufsteigender Grimmdarm


Organ 5 absteigender Dickdarm oder absteigender Grimmdarm oder Colon descendens

Abb. 4: Anordnungsfrage aus Abb. 3 als Long Menu Frage in ILIAS umgewandelt

Zuordnungsfrage: Das Anheben des Schwierigkeitsgrades und der Trennschärfe bei den Zuordnungsfragen erscheint gut möglich durch das Erhöhen der Antwortwahlmöglichkeiten. Die einschlägigen Handbücher geben zu diesem Thema keine eindeutige Antwort. Lediglich Krebs (2004) äußert sich hierzu eindeutig und gibt an, dass bis zu 26 Antwortwahlmöglichkeiten eine Annäherung an freies Antworten ermöglichen.

In Abb. 5 ist eine Zuordnungsfrage mit zwei überzähligen Antwortmöglichkeiten dargestellt, wie sie in Handbüchern oft dargestellt wird. Abbildung 6 ist die geänderte Frage mit n überzähligen Antwortmöglichkeiten.

Welche histologische Abbildung passt zu welchem Begriff?



Lamellenknochen

hyaliner Knorpel

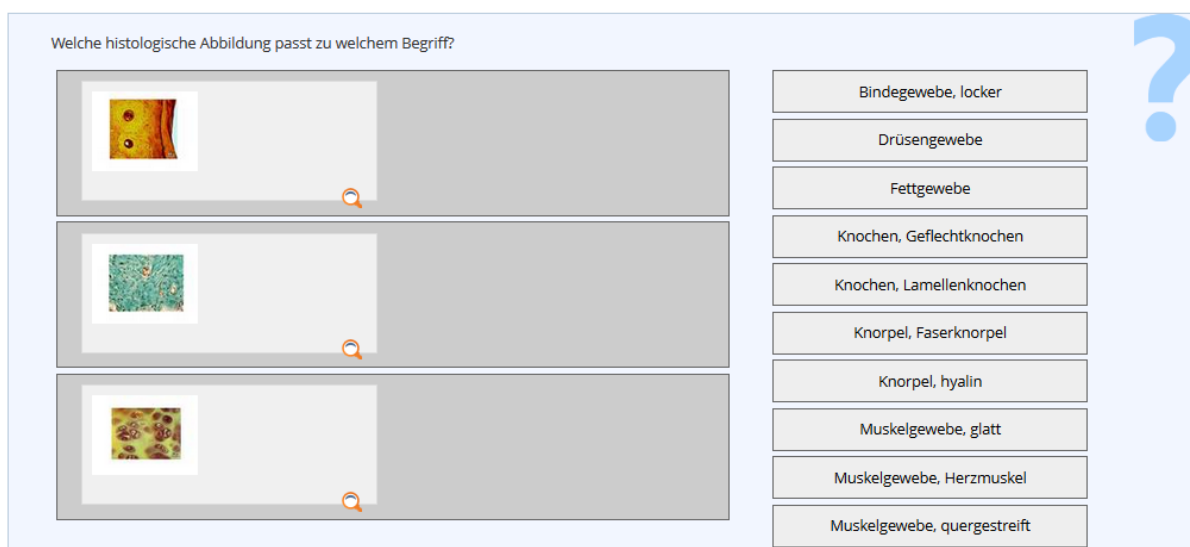
lockeres Bindegewebe

Geflechtknochen

Faserknorpel

Abb. 5: Zuordnungsfrage mit zwei Distraktoren

Welche histologische Abbildung passt zu welchem Begriff?



Bindegewebe, locker

Drüsengewebe

Fettgewebe

Knochen, Geflechtknochen

Knochen, Lamellenknochen

Knorpel, Faserknorpel

Knorpel, hyalin

Muskelgewebe, glatt

Muskelgewebe, Herzmuskel

Muskelgewebe, quergestreift

Abb. 6: Zuordnungsfrage mit 10 Antwortmöglichkeiten, die alphabetisch geordnet sind

Wie ist das Feedback seitens der Studierenden bzw. wurden die Klausuren und Fragen durch die Maßnahmen verbessert?

Das Feedback der Studierenden wird in der Form eingeschätzt, dass die revidierten Fragetypen in zukünftigen Klausuren eingesetzt werden und die Qualitätsindikatoren der Klausuren und Fragen auf Verbesserung überprüft werden. Die Ergebnisse der neuen Klausuren liegen noch nicht vor. Ein direktes Feedback der Studierenden wurde im Rahmen einer Übungsklausur eingeholt. Dabei waren die Studierenden mit dem neuen Fragetyp Long Menu Frage eher unzufrieden. Das eigenständige Eingeben von Begriffen fiel ihnen schwerer als die Auswahl von vorgegebenen Begriffen. Die Prüfungsstatistik der Übungsklausur ergab jedoch verbesserte Qualitätsindikatoren.

Was halten Sie als Fazit fest? Welche Lernerfahrungen haben Sie gemacht und welche Schlussfolgerungen ziehen Sie, z.B. in Hinsicht auf die Weiterentwicklung der Lehre in Ihrem Bereich?

Das systematische Analysieren der Fragen auf Ebene des Fragetyps hilft dabei, die Hemmschwelle der Fragenrevision zu überwinden. Der „Berg“ der zu revidierenden Fragen wird deutlich kleiner und

systematische Schwächen von bestimmten Fragetypen können erkannt werden. Es ist möglich, dass die Schwächen bestimmter Fragetypen nicht auf alle Fachdisziplinen übertragbar sind, sondern jeder Lehrende für sich selbst entscheiden muss, welche Fragetypen sinnvoll umgesetzt werden können. Die zur Verfügung stehenden Handbücher sind dabei nur partiell hilfreich. Das Nutzen der Qualitätsindikatoren, die vom eKlausursystem automatisiert zur Verfügung gestellt werden, ist in jedem Fall hilfreich. Ich werde die Informationen, die uns die Prüfungsstatistik zur Verfügung stellt, in Zukunft nutzen und meine Kollegen im Fachbereich motivieren, Verbesserungen der Lehre auf die Verbesserung der Klausuren auszuweiten und die Möglichkeiten, die sich mit den eKlausuren ergeben, auszuschöpfen.

Literatur:

- Krebs, R. (2004): Anleitung zur Herstellung von MC-Fragen und MC-Prüfungen für die ärztliche Ausbildung. Universität Bern. Bern
- Lukas J., Melzer, A. & Much, S., unter Mitarbeit von S. Eisentraut (2017). Auswertung von Klausuren im Antwort-Wahl-Format. Zugriff unter <http://nbn-resolving.de/urn:nbn:de:gbv:3:2-66099>
- Michel, L. (2015): Digitales prüfen und Bewerten im Hochschulbereich. Arbeitspapier Nr. 1. Geschäftsstelle Hochschulforum Digitalisierung.
- Möltner, A. , Schellberg, D. & Jünger, J. (2006): Grundlegende quantitative Analysen medizinischer Prüfungen. GMS Zeitschrift für Medizinische Ausbildung 23 (3): Doc53
- Vogt, M. & Schneider, S. (2009): E-Klausuren an Hochschulen. Koordinationsstelle Multimedia, JLU Gießen. 1. Auflage
- Waldschmidt-Dietz, F. (2015): E-Prüfungen aus der Sicht Lehrender. Auswertung einer Umfrage im April 2015 zu E-Klausuren und Scanner-Klausuren an der JLU Gießen