

Klausurerfolg trotz Unwissenheit

Akademische Leistungsbewertung mit dem Antwort-Wahl-Verfahren

| MARKUS KLEIN | Prüfungsordnungen halten manches Mal Überraschungen für den Prüfer parat. Da geht es um absolute und relative Bestehensgrenzen und um vorgegebene Benotungen, die dem Autor dieses Beitrags wenig einleuchten. Das Verhältnis von „Schweregrad der Prüfung“ und „Leistungsvermögen der Prüflinge“ scheint aus der Balance.

Wer kennt das nicht? Auf dem Schreibtisch stapeln sich Berge von Klausuren, deren Lektüre nur in den wenigsten Fällen Vergnügen bereitet. Die Handschrift der Studierenden ist oftmals kaum zu entziffern, und verständlich formulierte, grammatikalisch korrekte Sätze sind eher selten. Das alles gipfelt in durch Spiegelstriche gegliederte Stichwort-sammlungen, die sich um eine zusammenhängende Argumentation gar nicht erst bemühen. Befindet man sich in dieser Situation, dann erscheinen Klausuren nach dem sogenannten Antwort-Wahl-Verfahren als Verheißung. Die Korrektur beschränkt sich dann nämlich auf das Anlegen einer Lösungsschablone und kann durch das Scannen der Klausurbögen oder aber die Durchführung elektronischer Prüfungen sogar vollständig automatisiert werden. Auch der Autor dieses Beitrags war vor einiger Zeit geneigt, dieser Versuchung zu erliegen.

Dies änderte sich bei der Lektüre

der Prüfungsordnung des von mir be-
dienten Studiengangs sehr schnell. Dort
ist in § 19 (3) Folgendes festgehalten:
„Eine nach dem Antwort-Wahl-Verfahren
(z.B. Single-Choice oder Multiple-
Choice) durchgeführte Prüfung ist be-
standen, wenn der Prüfling mindestens
51 von Hundert der Fragen zutreffend
beantwortet hat (absolute Bestehens-
grenze). Abweichend davon wird in den
Fällen, in denen der Mittelwert aller

»Die Prüfungsordnung regelt nicht nur detailliert die Bestehensgrenze, sondern schreibt auch die Benotung für Prüfungen im Antwort-Wahl-Verfahren vor.«

Prüfungen abzüglich 18 von Hundert
schlechter ist als die absolute Bestehens-
grenze der so ermittelte Wert als relative
Bestehensgrenze festgelegt“.

Da diese Bestimmungen zunächst
eher technisch verklausuliert daher
kommen, versuchte ich mir die Konse-
quenzen für die Bewertung meiner
Klausuren auszumalen. Die relative Be-
stehensgrenze bedeutet zunächst, dass
Situationen eintreten können, in denen
Studierende die Klausur auch dann be-
stehen, wenn sie weniger als 51 Prozent
der Fragen richtig beantworten. Dies ist
immer dann gegeben, wenn die Prüflin-
ge im Durchschnitt weniger als 69 Pro-
zent zutreffende Antworten aufweisen.
Werden beispielsweise im Durchschnitt
aller Klausuren nur 30 Prozent der Fra-
gen richtig beantwortet, dann verschiebt
sich die relative Bestehensgrenze auf 12

Prozent. Sinkt also die durchschnittliche
Qualität der Klausurbearbeitung, dann
sinkt auch die relative Bestehensgrenze.
Nach unten sind dabei in der Prüfungs-
ordnung keine Grenzen gesetzt. Konse-
quent zu Ende gedacht gilt folglich:
Werden durchschnittlich nur 18 Prozent
aller Fragen richtig beantwortet, dann
verschiebt sich die relative Bestehens-
grenze auf 0 Prozent. In diesem Fall
würden alle Prüflinge die Klausur beste-
hen und zwar selbst dann, wenn sie alle
Fragen entweder gar nicht oder aber
falsch beantwortet haben.

Die Prüfungsordnung regelt aber
nicht nur detailliert die Bestehensgren-
ze. Auch die Benotung ist für Prüfungen
im Antwort-Wahl-Verfahren
haarklein vorgeschrieben. In
§ 19 (4) ist ein detailliertes
Bewertungsschema festge-
halten. So erhält ein Prüfling
beispielsweise die Note
3,0, wenn er mindestens 66,
aber weniger als 71 Prozent
der Fragen richtig beantwortet. Ergänzt
wird dieses Schema durch folgende Be-
stimmung aus § 19 (3): „Zur Ermittlung
der einzelnen Prüfungsergebnisse wird
die Differenz zwischen der relativen und
absoluten Bestehensgrenze bei jedem
Prüfling addiert“. Die Notenskala wird
folglich auf Räder gestellt und zwar der-
gestalt, dass eine Klausur, deren Pro-
zentanteil richtiger Antworten dem
Durchschnitt aller Klausurteilnehmer
entspricht, immer mit der Note 3,0 be-
wertet wird.

All dies fand ich erstaunlich. Ich ver-
suchte mir auszumalen, wie eine im Ant-
wort-Wahl-Verfahren gestellte Klausur
ausfallen würde, die von Studierenden
im Zustand vollkommener Unwissenheit
bearbeitet wird. Hierfür stellte ich mir
der Einfachheit halber eine Single-Choic-
e-Klausur mit 100 Fragen mit jeweils

AUTOR



Markus Klein ist
Professor für Politi-
sche Soziologie an
der Universität
Hannover. Seine
Forschungsschwer-
punkte liegen im
Bereich des Wähler-

verhaltens, des Wertewandels und der
Methoden der empirischen Sozial-
forschung.



Foto: picture-alliance

vier Antwortmöglichkeiten vor. Bei jeder Frage besteht damit eine Wahrscheinlichkeit von 25 Prozent, die richtige Antwort zu erraten. Der Erwartungswert für den Prozentanteil richtiger Antworten liegt damit sowohl für den einzelnen Studierenden als auch für die Studierendengruppe insgesamt bei 25 Prozent. Die absolute Bestehensgrenze wäre folglich deutlich verfehlt. Allerdings greift nun die relative Bestehensgrenze und die Klausur wäre mit sieben Prozent richtiger Antworten bestanden. Mittels des statistischen Modells der Binomialverteilung lässt sich die dabei zu erwartende Durchfallquote bestimmen. Diese ist erst auf der fünften Nachkommastelle von Null verschieden, d.h. alle Klausurteilnehmer bestehen die Klausur. Eine Klausur mit einem dem Erwartungswert von 25 Prozent entsprechenden Anteil gültiger Antworten wird mit der Note 3,0 bewertet.

Wie kommen diese in ihren Konsequenzen teilweise absurden Bestimmungen der Prüfungsordnung zustande? Zunächst sind diese Regelungen keine Besonderheit meiner Universität. Sie fin-

den sich in mehr oder weniger ähnlicher Form in allen Prüfungsordnungen, in denen das Antwort-Wahl-Verfahren vorgesehen ist. Hintergrund ist die geltende Rechtsprechung, nicht zuletzt ein Urteil

»Relative Bestehensgrenzen machen es letztlich unmöglich, einen substanzial definierten Mindestanspruch festzulegen.«

des Bundesverfassungsgerichts aus dem Jahr 1989 (BVerfGE 80, 1). Argumentiert wird dabei im Wesentlichen, dass der Schwierigkeitsgrad von Prüfungen im Antwort-Wahl-Verfahren im Vorhinein nicht zuverlässig bestimmt werden könne. Ihr Schwierigkeitsgrad schwanke daher von Prüfung zu Prüfung. Um die Gleichbehandlung der Prüflinge verschiedener Prüfungstermine zu gewährleisten, dürften sich die Bestehensgrenze und die Notenskala daher nicht ausschließlich an absoluten Kriterien festmachen, sondern müssten sich

auch am durchschnittlichen Klausurergebnis aller Prüflinge eines Termins orientieren. Diese Argumentation übersieht allerdings, dass das Prüfungsergebnis nicht nur vom Schweregrad der Prüfung abhängt, sondern natürlich auch vom Leistungsvermögen der Prüflinge. Relative Bewertungsmaßstäbe korrigieren daher für Unterschiede im Schwierigkeitsgrad der Prüfung – aber eben auch für Unterschiede in der akademischen Leistungsfähigkeit der Prüfungskohorten.

Relative Bestehensgrenzen machen es außerdem letztlich unmöglich, einen substanzial definierten Mindestanspruch festzulegen, dem Prüflinge unter allen Umständen genügen müssen. Im Extremfall ermöglichen sie Klausurerfolg trotz Unwissenheit. Will man akademische Prüfungen aber nicht ad absurdum führen, dann muss man Prüfern das Recht zubilligen, inhaltliche Mindestanforderungen für das Bestehen zu formulieren. Und man muss ihnen auch zutrauen, deren Erfüllung in angemessener Weise zu kontrollieren.