Wissenschaftliche Beiträge

Multiple Choice als Prüfungsform bei rechtswissenschaftlichen Fachprüfungen

Stefan Griller/Monika Stöggl[*]

Zusammenfassung: Unter allen Prüfungsformen, die im rechtswissenschaftlichen Studium im deutschsprachigen Raum angeboten werden, haben Multiple Choice (MC) Prüfungen mit Abstand den schlechtesten Ruf. Dennoch ist das Prüfen in dieser Form – nicht zuletzt wegen der wachsenden Studierendenzahlen – ein zunehmender Trend. Die wohl am kritischsten hinterfragte Prüfungsmethode steht in diesem Beitrag auf dem Prüfstand, nota bene auch für Fachprüfungen, also Abschlussprüfungen in einzelnen Fächern.

Der Beitrag widmet sich der Frage, welche Lernziele durch MC abgefragt werden können, ob und unter welchen Voraussetzungen MC als Prüfungsform im rechtswissenschaftlichen Studium geeignet ist, wie hochwertige MC-Fragen erstellt werden, sowie den Vor- und Nachteilen dieser Prüfungsform. Abschließend wird anhand von Testgütekriterien ein Vergleich mit klassischen schriftlichen Prüfungen durchgeführt.

A. Multiple Choice ist nicht gleich Multiple Choice!

Unter dem Begriff „Multiple-Choice“ werden unterschiedliche Fragetypen sogenannter „geschlossener Fragen“ zusammengefasst. Geschlossene Fragen sind generell solche, die das Ankreuzen von bereits vorgegebenen Antworten einfordern.[1] Dabei wird unterschieden zwischen „Single-Choice“ und „Multiple-Choice“.[2] Während bei „Single-Choice-Fragen“ grundsätzlich nur eine von mehreren Antworten richtig ist, können bei MC-Fragen mehrere Antworten richtig sein.[3] Die Art der gewählten Methode bestimmt nicht nur den Schwierigkeitsgrad für die Studierenden, sondern hat auch maßgeblichen Einfluss auf die Ratewahrscheinlichkeit. Neben der Art der gewählten Methode wirkt sich die Anzahl der Fragen insgesamt, die Gestaltung der Testkonstruktion, die Art und Weise, wie die Antwortalternativen (Distraktoren) vorgegeben werden, die Anzahl der vorgegebenen Teilantworten („1 aus 4“- „1 aus 6“) und die Anzahl der richtigen (eine oder mehrere) Teilantworten (“1 aus 5“ oder „x aus 5“) sowie die Auswertungsmethode (Minus-

* *Stefan Griller* ist Universitätsprofessor und *Monika Stöggl* Universitätsassistentin am Fachbereichsteil Europarecht der Universität Salzburg

1 Vgl. *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 92 (92).

2 Hier gibt es begriffliche Unterschiede im deutsch- und englischsprachigen Raum: Während im deutschsprachigen Raum zwischen Single Choice und Multiple Choice unterschieden wird, wird im englischsprachigen Raum oftmals zwischen Multiple Choice (eine richtige Antwortmöglichkeit) und Multiple Response (mehrere Antwortmöglichkeiten) differenziert.

3 Vgl. *Kubinger*, in: Weber/Rammsayer (Hrsg.), S. 158 (158 f.).

ZDRW 2/2018, DOI: 10.5771/2196-7261-2018-2-148

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

punkte innerhalb eines Fragenkomplexes)[4] maßgeblich auf die Qualität, den Schwierigkeitsgrad und die Ratewahrscheinlichkeit aus.[5] Dabei gilt grundsätzlich: Der Schwierigkeitsgrad einer MC-Prüfung erhöht und die Ratewahrscheinlichkeit verringert sich mit zunehmender Zahl an Teilantworten (zumindest 4 Teilantworten)[6], der Anzahl der Aufgaben insgesamt sowie einer unbestimmt richtigen Anzahl von Teilantworten („x aus 5"). Die Ratewahrscheinlichkeit bei MC-Aufgaben der Art „x aus 5", bei der jede der fünf Antwortmöglichkeiten unabhängig von den übrigen als richtig oder falsch erkannt werden muss, liegt bereits nach 3 Aufgaben bei 0,003 Prozent[7] und ist damit vernachlässigbar.

Der Generalverdacht eines wenig anspruchsvollen, vor allem Zufallsergebnisse begünstigenden „Kreuzerltests" kann für die neu und hochwertig entwickelten MC-Prüfungen keinesfalls bestätigt werden.

B. Multiple Choice und lernzielorientiertes Prüfen

Ob und gegebenenfalls in welchen juristischen Fächern MC-Prüfungen eingesetzt werden sollen setzt die Beantwortung der Frage voraus, welche Lernziele in den jeweiligen Fächern verfolgt werden.

Denn grundsätzlich gilt, dass das Prüfungsformat (bspw. mündliche, schriftliche, MC-Prüfungen) anhand der angezielten Lernergebnisse gewählt werden sollte und nicht aufgrund anderer Parameter, wie bspw. der Verfügbarkeit und Kosteneffizienz des jeweiligen Prüfungsinstruments oder der Studierendenzahlen.[8]

Eine umfassende Auseinandersetzung mit dem erziehungswissenschaftlichen Ansatz der Kompetenzorientierung hat innerhalb des rechtswissenschaftlichen Studiums in Österreich bisher nicht stattgefunden. Es gibt kein rechtswissenschaftliches Kompetenzmodell, welches als gesicherte Basis für die Curricula dienen könnte.[9] Eher im Gegenteil: Die österreichischen Universitäten erachten sich bei der Definition der Prüfungsanforderungen weitgehend als frei.[10] Das hat zur Folge, dass in den rechtswissenschaftlichen Studienplänen – wenn überhaupt – die angestrebten

4 Vgl. dazu: Beschluss des Oberverwaltungsgerichtes NRW vom 16. Dezember 2008, 14 A 2154/08, http://www.justiz.nrw.de/nrwe/ovgs/ovg_nrw/j2008/14_A_2154_08urteil20081216.html (1.5.2017).

5 Vgl. *Jacobs*, Richtlinien zur Erstellung von einfachen Multiple-Choice-Aufgaben nach Grönlund, S. 9.

6 Vgl. *Rütter*, Formen der Testaufgabe, S. 149 ff.

7 Die Berechnungsformel lautet: (1/2)5 (je Teilantwort richtig/falsch – insgesamt 5 Teilantworten) = 1/32. Um die dritte Potenz (3 Aufgaben) von 1/32 zu berechnen werden Zähler und Nenner jeweils zur dritten Potenz genommen, das Ergebnis ist dann 1/32768. Als Prozentwert dargestellt sind das 0,0030517578125%.

8 Vgl. *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 91 (95); *Gruber/Avallone*, Empfehlungen zum Einsatz von Multiple-Choice-Prüfungen, S. 8.

9 Vgl. *Warto*, in: Astleitner/Lagodny et al. (Hrsg.), S. 35 (41).

10 Vgl. *Lachmayer*, in: Zfhr 2012, S. 139 (141 f).

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

Lernergebnisse nur in rudimentärer Form aufgelistet sind.[11] Doch was sind Lernergebnisse?

Unter Lernergebnissen (auch Learning-Outcomes)[12] wird eine Aussage darüber verstanden, was Studierende nach dem Abschluss eines bestimmten Lernprozesses wissen, verstehen und tun können.[13] Lernergebnisse können als Wissen (die Studierende/der Studierende kennt…), Fertigkeiten (die Studierende/der Studierende beherrscht die Methode x) oder Qualifikationen (die Studierende/der Studierende ist befähigt, eine bestimmte Position einzunehmen oder Tätigkeit auszuüben) beschrieben werden.[14] Sie werden als Kompetenzen unter Einschluss von Kenntnissen und Fertigkeiten verstanden.[15] „Kompetenzen" werden dabei wie folgt definiert „…,die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können"[16].

Um die Einsatzmöglichkeiten von MC-Prüfungen beurteilen zu können, ist es notwendig, die angestrebten Lernergebnisse einer Lehrveranstaltung nach Niveaugraden zu unterscheiden. Ein sinnvolles Hilfsmittel dazu ist die Taxonomie-Tabelle der modifizierten *Bloom´schen* Lernergebnis-Taxonomie. Die Tabelle differenziert die kognitiven Lernergebnisse nach sechs Erkenntnisdimensionen: Erinnern, Verstehen, Anwenden, Analysieren, Bewerten und Entwickeln.[17]

Die mit MC-Tests überprüfbaren Lernergebnisse reichen prinzipiell von der Reproduktion bzw. dem Erinnern von Faktenwissen bis zur Entwicklung neuer Problemlösungen unter Anwendung von Problemlösungsstrategien.[18]

11 Vgl. *Warto*, in: Astleitner/Lagodny et al. (Hrsg.), S. 35 (49); Im Gegensatz dazu sind die Inhalte der rechtswissenschaftlichen Lehre in Deutschland aus den Lernzeilen der zu absolvierenden Staatsexamen abzuleiten.

12 Vgl. *Wildt/Wildt*, in: Berendt/Voss et al. (Hrsg.), S. 4.

13 Vgl. *Reis*, in: ZDRW (2013), S. 21 (27); *Kennedy/Hyland/Ryan*, in: Froment/Kohler et al. (Hrsg.), S. 1 (5).

14 Vgl. *Schermutzki*, Lernergebnisse – Begriffe, S. 5; *Gruber/Avallone*, Empfehlungen zum Einsatz von Multiple-Choice-Prüfungen, S. 8.

15 Ebd.

16 Zu den Begriffen „Lernergebnis" und „Kompetenz" vgl. ECTS-Leitfaden, S. 13 ff.; *Weinert*, Leistungsmessung in Schulen, S. 27 ff.

17 Neben den kognitiven Lernzielen werden auch affektive und psychomotorische Lernziele unterschieden Letztere sind für unsere Fragestellung unanwendbar. Affektive Lernziele werden stärker für die Geisteswissenschaften postuliert, zum Teil aber auch für die Rechtswissenschaften. Sie gliedern sich in Aufnehmen, Reagieren, Werten, Wertordnung und Bestimmtsein durch Werte. Es besteht eine starke Überlappung zur Begriffsbildung oben im Text. Von den Studierenden wird gefordert, sich eine Meinung zu einem Sachverhalt zu bilden. Dazu gehört, dass Standpunkte und Kriterien rein kognitiv aufgenommen werden, diese werden dann aber in weiteren Schritten einander zugeordnet, und die eigene Sichtweise wird mit der erlernten abgeglichen. Bsp: Sachverhaltsbeschreibung mit Befragung der Studierenden nach ihrer Einschätzung. Siehe zu dieser Begrifflichkeit *Vogt/Schneider*, E-Klausuren an Hochschulen, S. 7.

18 Vgl. *Gruber/Avallone*, Empfehlungen zum Einsatz von Multiple-Choice-Prüfungen, S. 3 ff.

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

Die Reproduktion, d.h. die Wiedergabe des erlernten Wissens, ist problemlos mit MC-Fragen zu überprüfen. Studierende müssen bei MC-Fragen in der Lage sein, das erworbene Wissen auf einfache oder komplexe Teilantworten zu übertragen. Verständnis und Anwendung des Gelernten stellt schon eine größere Herausforderung an die Aufgabenstellung. Das Anwenden und Verstehen kann im Bereich der Rechtswissenschaften dadurch abgefragt werden, dass im Fragenstamm ein Sachverhalt angegeben wird, auf den sich die Teilantworten beziehen. Ist der Sachverhalt fiktiver Natur oder wurde er gegenüber jenem etwa in einer Lernunterlage geringfügig geändert, wird ersichtlich, ob Studierende ihr erlerntes Wissen auf neue Sachverhalte bzw. in einem neuen Kontext anwenden können. Die Überprüfung von Kompetenzen im Bereich rechtlicher Beurteilungen sowie der konkreten Rechtsanwendung können auf diese Weise problemlos überprüft werden. Die Kompetenz „Rechtsanwendung" kann auch dadurch abgefragt werden, dass Urteilsauszüge oder Textpassagen im Stamm vorgegeben werden und die daraus zu ziehenden rechtlichen Schlüsse in den Teilantworten präsentiert werden.

Durch eine MC-Prüfung müssen Studierende ihre Argumente nicht schriftlich oder mündlich ausführen, jedoch ist die Denkleistung Voraussetzung für das richtige Beantworten der Frage, und manifestiert sich in einer von den Prüfenden vorgegebenen Antwort. Die kognitiven Lernprozesse des ‚Erinnerns, Verstehens, Anwendens und Analysierens', aber auch des ‚Bewertens' finden hier statt, bspw. aufgrund vernetzter oder auf einander aufbauender Fragestellungen. Darüber hinaus können die notwendigen Argumentationsschritte (z.B. zur Falllösung) in Form von MC explizit abgefragt werden.

Zu beachten ist allerdings, dass die Entwicklung der Prüfungsfragen mit zunehmenden Komplexitätsgrad des Wissens und der kognitiven Prozesse schwieriger wird.[19] Der Hauptanwendungsbereich von MC-Prüfungen sollte möglichst auf die kognitiven Prozesse Erinnern, Verstehen, Anwenden und Analysieren beschränkt werden.[20] Diese Einschränkung gilt im Übrigen nicht nur für MC, sondern auch für Fragen mit offenen mündlichen oder schriftlichen Antwortmöglichkeiten.[21] Für die Überprüfung von Lernzielen auf den Ebenen (vertiefendes) Entwickeln oder Bewerten eignen sich andere Prüfungsformen wie Forschungsprojekte, Berichte oder wissenschaftliche Abhandlungen wesentlich besser.[22] Freilich ist – nicht nur, aber auch aus rechtswissenschaftlicher Sicht – anzumerken, dass die Grenzlinie zwischen Analysieren und Bewerten fließend ist. Das Lösen von Fällen kann als Beurteilung („Bewertung") eines Sachverhalts anhand von Rechtsnormen dargestellt werden, die ihrerseits bestimmte Kriterien und Standards enthalten. Diese Beurteilung erfordert eine Analyse sowohl der anzuwendenden Normen als auch des zu beurteilenden Sachverhalts, die juristische Subsumption besteht hauptsächlich aus

19 Vgl. *Asmuth*, Prüfen mit der Multiple-Choice-Methode, S. 1.

20 Vgl. *Walzik*, Kompetenzorientiert prüfen, S. 41 f.; *Gruber/Avallone*, Empfehlungen zum Einsatz von Multiple-Choice-Prüfungen, S. 9 ff.

21 Vgl. *Astleitner/Wageneder et al.*, 12 Tipps für eine kompetenzorientierte Lehre, S. 7.

22 Ebd, S. 7 ff.

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

dem In-Beziehung-Setzen der beiden zueinander – das berühmte „Hin- und Herwandern des Blicks".[23] Der Vorgang lässt sich somit auch als Analyse im Sinne der obigen Abgrenzung darstellen. Die Fähigkeit, Fälle zu lösen, ist daher durch MC-Tests durchaus überprüfbar.

Nach testtheoretischen Erkenntnissen sind im kognitiven Bereich fast alle Taxonomiestufen durch MC-Prüfungen prüfbar.[24] Wenngleich ein Betrachter aus der Außenperspektive lediglich das Setzen von Kreuzen beobachten kann, finden, bei entsprechender Aufgabenstellung, komplexe kognitive Prozesse bei den Studierenden statt.[25] Lediglich ein „Erzeugen durch Ankreuzen" ist bei geschlossenen Antwortoptionen nicht möglich.[26] Im Ergebnis führt dies dazu, dass MC-Prüfungen grundsätzlich geeignet sind, auch im rechtswissenschaftlichen Studium als Prüfungsform eingesetzt zu werden. Ob eine MC-Prüfung in jedem Rechtsgebiet Anwendung finden kann, ist nicht pauschal zu beantworten. Vielmehr muss zunächst jedes Rechtsgebiet die konkreten Lernziele definieren. Darüber hinaus ist ein ausgewogener Einsatz aller Prüfungsformen im rechtswissenschaftlichen Studium empfehlenswert. Ein KO-Argument gegen MC-Prüfungen lässt sich aus den lerntheoretischen Grundlagen im ersten Zugang nicht ableiten.[27]

C. Das Erstellen qualitativ hochwertiger Fragen

Das Erstellen von MC-Fragen ist nicht einfach. Gute Fragen, die das bloße Erraten der richtigen Antwort unmöglich machen und so zu Denkleistungen zwingen, sind nicht in ein paar Minuten erstellt. Die adäquate Konstruktion von MC-Fragen erfordert folglich viel mehr Zeit als jene von offenen Fragen. Sie setzt zudem besondere Erfahrungen im betreffenden Sachgebiet, der angemessenen Formulierung von Fragen sowie umfassende Kenntnisse im Bereich der Testkonstruktion und -auswertung voraus. [28]

MC-Tests stellen nur dann eine zuverlässige, objektive und gegenüber den Studierenden auch faire Prüfungsformen dar, wenn sie „*richtig*" gemacht werden.[29]

Dafür sollten MC-Aufgaben einer vorgegebenen Struktur folgen. Sie bestehen aus einem Stamm, der die Problembeschreibung darstellt, einer konkreten und eindeutigen Frage/einem Arbeitsauftrag, sowie mehreren Antwortoptionen.[30] Für die Schwierigkeit einer Aufgabe sind vor allem die falschen Antwortmöglichkeiten

23 Vgl. *Engisch*, Logische Studien zur Gesetzesanwendung, S. 15: „Hin- und Herwandern des Blickes zwischen Obersatz und Lebenssachverhalt". Typisch auch – bei allen Unterschieden zu *Engisch* – etwa *Bydlinski*, Juristische Methodenlehre und Rechtsbegriff, S. 395 ff.

24 Vgl. *Krüger*, in: Brockmann/Pilniok (Hrsg.), S. 109 (113).

25 Ebd S. 109 (113).

26 Ebd S. 109 (114).

27 Zur Frage der möglichen Einsatzgebiete von MC im juristischen Studium siehe auch *Krüger*, in: Brockmann/Pilniok (Hrsg.), S. 109 (116).

28 Vgl. *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 92 (94).

29 Vgl. *Gruber/Avallone*, Empfehlungen zum Einsatz von Multiple-Choice-Prüfungen, S. 17 ff.

30 Vgl. *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 92 (94).

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

(Ablenker) entscheidend. Sie sollen den unwissenden Studierenden plausibel erscheinen, damit sie von der richtigen Antwort abgelenkt werden. Hingegen dürfen sie die Studierenden nicht verwirren.[31] Hinsichtlich der Distraktoren bietet es sich in den Rechtswissenschaften an, häufige Fehlmeinungen oder falsche Konzepte zu verwenden.[32] Es ist auch darauf zu achten, dass Distraktoren keinesfalls unsinnig, unplausibel oder trivial sein dürfen, damit sie von unwissenden Studierenden nicht als „falsch" erkannt werden können.[33]

Damit das Finden der richtigen Antwort ausschließlich vom Wissen der Studierenden abhängt und die Ratewahrscheinlichkeit minimiert wird, ist die Einhaltung einer Reihe von inhaltlichen sowie formalen und sprachlichen Regeln empfehlenswert.[34] Teilantworten sollten inhaltlich homogen, kurz und plausibel sein und sich auf die Fragestellung beziehen.[35] Die Antwortalternativen sollten klar unterscheidbar sein und sich nicht überschneiden.[36]

Von der Verwendung der Optionen „alle genannten Alternativen" bzw „keine der genannten Alternativen" und der Art „sowohl a als auch c" „nur a, sofern c"; „a oder b, wenn nicht c, doppelte Verneinungen" ist abzuraten, zumal sich dadurch Distraktoren leichter identifizieren lassen und gleichzeitig die Validität der Lernergebniserfassung reduziert wird.[37]

Um unbeabsichtigte Lösungshinweise zu vermeiden, sollten auch formale und sprachliche Kriterien beachtet werden. Nach *Krebs*[38] müssen alle Antwortoptionen grammatikalisch zum Aufgabenstamm passen, die falschen Antwortalternativen sollten in ihrer Länge und in ihrem Differenzierungsgrad den richtigen entsprechen, Ähnlichkeiten von Wörtern sollten ebenso vermieden werden wie „Lehrbuchformulierungen" oder „absolute Begriffe" (wie bspw. „niemals", „immer", „alle", „kein", „nur") und moderate Begriffe („manchmal", „möglicherweise", „gewöhnlich"). Während „absolute Begriffe" auf eine falsche Lösung hindeuten, lassen moderate Begriffe auf die richtige Lösung schließen. Lehrbuchformulierungen kommen den Studierenden bekannt vor und signalisieren Ihnen so, dass es sich um die richtige Lösung handelt.[39] Auch die Anordnung der Antwortoptionen sollte

31 Vgl. *Jacobs*, Richtlinien zur Erstellung von einfachen Multiple-Choice-Aufgaben nach Gronlund, S. 9.

32 Vgl. *Krebs*, Anleitung zur Herstellung von MC-Fragen und MC-Prüfungen, S. 16; *Jacobs*, Richtlinien zur Erstellung von einfachen Multiple-Choice-Aufgaben nach Gronlund, S. 9.

33 Vgl. *Krebs*, Anleitung zur Herstellung von MC-Fragen und MC-Prüfungen, S. 16.

34 Vgl. *Gruber/Avallone*, Empfehlungen zum Einsatz von Multiple-Choice-Prüfungen, S. 17 ff.

35 Vgl. *Krebs*, Anleitung zur Herstellung von MC-Fragen und MC-Prüfungen, S. 16; *Gruber/Avallone*, Empfehlungen zum Einsatz von Multiple-Choice-Prüfungen, S. 17 ff.

36 Vgl. *Krebs*, Anleitung zur Herstellung von MC-Fragen und MC-Prüfungen, S. 17.

37 Vgl. *Jacobs*, Richtlinien zur Erstellung von einfachen Multiple-Choice-Aufgaben nach Gronlund, S. 10.

38 Vgl. *Krebs*, Anleitung zur Herstellung von MC-Fragen und MC-Prüfungen, S. 17 ff.

39 Vgl. *Jacobs*, Richtlinien zur Erstellung von einfachen Multiple-Choice-Aufgaben nach Gronlund, S. 8.

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

nach dem Zufallsprinzip erfolgen. Richtige Antworten werden überdurchschnittlich häufig von den Prüfern in der Mitte der Antwortalternativen platziert.

Im Folgenden ein Beispiel für eine MC-Prüfungsfrage aus Europarecht an der Universität Salzburg,[40] wo diese Prüfungstechnik im Wintersemester 2016/17 eingeführt wurde. Die Frage soll die Kompetenzen Erinnern, Verstehen, Anwenden, Analysieren, Bewerten abtesten, ausgehend davon, dass der Fall nicht genauso in den Lehrveranstaltungen besprochen wurde, wohl aber ähnliche Fälle:

Art. 8 der Verordnung (EU) Nr. 1173/2011 über die wirksame Durchsetzung der haushaltspolitischen Überwachung im Euro-Währungsgebiet, ABl. Nr. 2011/L 306/1 lautet auszugweise:

> *„Artikel 8*
>
> *Sanktionen bei Manipulation von Statistiken*
>
> *(1) Der Rat, der auf Empfehlung der Kommission tätig wird, kann beschließen, gegen einen Mitgliedstaat, der Daten über Defizite und Schulden, die für die Anwendung der Artikel 121 oder 126 AEUV oder für die Anwendung des dem EUV und dem AEUV beigefügten Protokolls über das Verfahren bei einem übermäßigen Defizit von Bedeutung sind, absichtlich oder aufgrund schwerwiegender Nachlässigkeit falsch darstellt, eine Geldbuße zu verhängen.*
>
> *(2) Die Geldbußen nach Absatz 1 müssen wirksam, abschreckend und — in Bezug auf Art, Schwere und Dauer der Verfälschung der Darstellung — verhältnismäßig sein. Der Betrag der Geldbuße darf die Höhe von 0,2 % des BIP des betreffenden Mitgliedstaats nicht überschreiten.*
>
> *…*
>
> *(5) Der Gerichtshof der Europäischen Union hat die unbeschränkte Befugnis zur Überprüfung der Beschlüsse des Rates, mit denen Geldbußen gemäß Absatz 1 verhängt werden. Er kann die verhängte Geldbuße aufheben, herabsetzen oder erhöhen.“*
>
> *Markieren Sie die zutreffende(n) Aussage(n) – und bedenken Sie dabei die Art. 258-260 sowie Art. 293 AEUV!*
>
> *a) Diese Bestimmung ist Teil des „sixpack“, durch den der Stabilitäts- und Wachstumspakt geändert wurde, mit dem Zweck, die Wirtschafts- und Währungsunion krisenfester zu machen.*

40 Das Europarecht ist in den Studienplänen sowohl des rechtswissenschaftlichen Diplomstudiums als auch des Bachelorstudiums Recht und Wirtschaft mit jeweils 4 Wochenstunden (insgesamt mit 6 ECTS) vertreten; das sind Vorlesungen mit Übungsanteil; Seminare gibt es nicht; die Prüfung ist schriftlich. Nach Auffassung der Verfasser handelt es sich unter diesen Bedingungen um ein Fach, das im Kontext der Studienpläne sinnvoll mittels MC abgeprüft werden kann.
Der geneigte Leser ist eingeladen, es selbst zu probieren! Zur Auflösung des Falles siehe unten Fn 42.

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

b) Der Rat darf von der Empfehlung der Kommission nur einstimmig abweichen.
c) Die Verordnung und damit auch ihr Art. 8 gilt für alle EU-Mitgliedstaaten.
d) Die Bestimmung ist inhaltsgleich mit jenen über das Vertragsverletzungsverfahren.
e) Eine Abweichung vom Vertragsverletzungsverfahren liegt darin, dass dort die subjektive Tatseite (Vorsatz, Fahrlässigkeit) keine Rolle spielt.

D. Die automatisierte Prüfungsauswertung

MC-Prüfungen können automatisiert ausgewertet werden. Dies ist nicht nur zeit- und ressourcenschonend, sondern trägt dazu bei, dass Prüfungen schneller und prüferunabhängig (damit für Studierende fairer) benotet werden. Darüber hinaus werden durch die automatisierte Auswertung die Leistungen der Studierenden statistisch besser erfasst, was eine verlässliche Datenbasis für die Prüfungsevaluierung (Stärken-/Schwächen-Analyse) erzeugt. Bei der automatisierten Prüfungsauswertung gibt es unterschiedliche Modelle, um die jeweiligen Markierungen der Studierenden zu werten und daraus eine Note abzuleiten. Gängig ist, für jede richtige Markierung einen Punkt zu geben und/oder für jede falsche Markierung ein bis ein halber Punkt abzuziehen.

Ein Nachteil dabei ist die jedenfalls in Österreich bisher ungeklärte Rechtmäßigkeit von Minuspunkten, wenn sie fragenübergreifend vergeben werden. Das Oberverwaltungsgericht *Münster* urteilte, dass ein Bewertungsverfahren rechtsfehlerhaft sei, wenn für eine falsche Antwort „Punkte abgezogen werden, die durch eine richtige Antwort erreicht worden sind". Die Aufgabe einer Prüfung ist es, „Aussagen darüber zu gewinnen, welche berufsbezogenen Kenntnisse der Prüfling hat. Einem Bewertungsverfahren, bei dem fehlerfrei erbrachte Prüfungsleistungen als nicht oder schlecht erbracht gewertet werden, weil andere Prüfungsfragen nicht richtig beantwortet worden sind, fehlt diese Eignung."[41] Problematisch sind damit jedenfalls Minuspunkte, die eine positive Punktezahl reduzieren, die bei einem anderen Beispiel erreicht wurden. Nicht betroffen von dieser Kritik sind Punkteabzüge innerhalb eines Beispiels, durch welche die Leistung differenziert wird.

Eine andere häufig gewählte Korrekturform ist die „Ratekorrektur nach *Gronlund*"[42]. Der Korrekturschlüssel lautet dabei: (richtig gelöste Auswahlantworten–falsche gelöste Auswahlantworten)/(Anzahl der Antworten –1). Hier gilt das Nicht-Ankreuzen einer falschen Antwort als richtige Lösung.[43]

41 Vgl. dazu: Beschluss des Oberverwaltungsgerichtes NRW vom 16. Dezember 2008, 14 A 2154/08, http://www.justiz.nrw.de/nrwe/ovgs/ovg_nrw/j2008/14_A_2154_08urteil20081216.html (15.12.2107).

42 Vgl. *Gronlund*, Assessment of student achievement, S. 53.

43 Ebd, S 53.

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

Das Werten jeder Teilantwort mit einem Punkt eröffnet jedoch erfolgreiche Rateoptionen. Im schlimmsten Fall können Studierende 4 von 5 Punkten erhalten (wenn sie alle 5 Teilantworten markieren und 4 richtig sind oder umgekehrt, wenn eine Teilantwort richtig wäre und die Studierenden gar keine Lösung markieren).

Die am Fachbereichsteil Europarecht der Universität Salzburg verwendete Auswertungsmethode durchkreuzt jeden Versuch, durch Raten zum Erfolg zu kommen, indem das vollständigen Ankreuzen oder Nichtankreuzen aller Antwortoptionen automatisch zu 0 Punkten führt – ausgenommen die sehr sparsam eingesetzte Variante, dass alle Antwortoptionen richtig sein können. Diese auch an der WU Wien[44] praktizierte Methode verwendet ein Pluspunkte- und Abzugsschema innerhalb einer Aufgabe. Die Antworten werden gewichtet. Alle richtigen Antworten ergeben in Summe +100 %, alle Distraktoren geben in Summe –100 %. Für das Nicht-Ankreuzen einer falschen Antwort erhält man keine Punkte. Die Punkte für die Aufgabe ergeben sich durch den Abzug der gewichteten, falsch markierten Antworten von den gewichteten, richtig markierten Antworten. Im schlechtesten Fall wird eine Aufgabe mit 0 Punkten bewertet. Beispiel: 3 richtige und 2 falsche Antwortmöglichkeiten bedeutet: jede richtige Antwort zählt 1,667 Punkte, jede falsche -2,5. Ankreuzen und Nichtankreuzen aller Antworten ergibt jeweils 0; zwei richtige und eine falsche ergibt 2 x (1,667) - 2,5 = 0,834; drei richtige und eine falsche ergibt 5 - 2,5 = 2,5 Punkte usw.[45]

E. Die Vorteile und Nachteile von MC-Prüfungen

Grundlage einer Entscheidung für oder gegen MC als Prüfungsform sollte eine umfassende Auseinandersetzung mit den konkreten Lernzielen im jeweiligen Rechtsgebiet sein. Zudem bedarf es von den Prüfenden besonderer und umfassender Kenntnisse und Erfahrungen bei der Fragenformulierung sowie im Bereich der Testkonstruktion und -auswertung. [46] Arbeits-, Ressourcen- und Zeitaufwand für die Konzeption von qualitativ hochwertigen MC-Fragen sind definitiv erheblich größer als bei offenen Fragen. MC-Prüfungen verlagern den Arbeitsaufwand in die Erstellungsphase, während bei der Korrektur der Prüfung Zeit und Ressourcen gewonnen werden. Auch die Qualitätssicherung, statistische Auswertung und allenfalls Nachbearbeitung der Fragen nehmen viel Zeit und Ressourcen in Anspruch. Das bedeutet ferner, dass MC-Prüfungen nur dann als ökonomisch[47] eingestuft werden können, wenn sie regelmäßig und dauerhaft zur Anwendung kommen, sodass sich die Erstellung eines Fragenpools lohnt.

44 An der WU Wien ist diese Methode – nach einem früheren Rektor – als „Hansen Methode" bekannt.

45 Beurteilen Sie sich selbst: im Beispiel oben nach FN 37 sind die Antwortoptionen b, c, d, falsch, die Antwortoptionen a, e richtig.

46 Vgl. *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 91 (94).

47 Im Sinne von ressourcenschonend, nämlich zeitsparend. Andererseits braucht es leistungsfähige Scanner, die aktuell mehr als € 10.000 kosten, was bei entsprechender Auslastung aber weit unter den eingesparten Personalkosten bleibt.

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

Im Rahmen einer MC-Prüfung kann aufgrund der höheren Fragenanzahl gegenüber anderen Prüfungsformen ein breites Stoffgebiet flächendeckend abgefragt werden, sodass „Glück“ durch Lücken in der Fragenzusammenstellung weitgehend ausscheidet.

Für die Studierenden signalisieren MC-Prüfungen ein hohes Maß an Transparenz, Fairness und Gleichheit der Beurteilung. Ferner erhalten sie ihre Prüfungsergebnisse in der Regel wesentlich schneller als bei klassischen schriftlichen Prüfungen.

Die automatisierte Auswertung erzeugt weiteren Nutzen: dadurch können umfassende Prüfungsanalysen sowie statistische Auswertungen erfolgen. Die Itemanalyse oder auch ein Schwierigkeitsindex misst die relative Lösungshäufigkeit bzw. den Schwierigkeitsgrad einer Aufgabe.[48] Die Prüfungsfragen können dadurch als leichte, mittelschwere und schwere Aufgaben kategorisiert in den Fragenpool aufgenommen werden. Dieses Vorgehen ermöglicht, dass künftige Prüfungen ausgewogener erstellt werden. Ein weiterer Vorteil ist die statistische Antwortenanalyse, wodurch schlecht formulierte Fragen identifiziert werden können. MC-Tests sind damit ein sehr gutes Messinstrument von Lernleistungen.

Vorteile liegen auch in der Kosteneffizienz bei hohen Studierendenzahlen. Darüber hinaus verhindern MC-Prüfungen subjektive Beurteilungsfehler, wie sie bei herkömmlichen mündlichen oder schriftlichen Prüfungen häufig, meist unbewusst unterlaufen. Solche Fehler werden auf verschiedene Faktoren zurückgeführt, wie insbesondere: der Persönlichkeitstyp der Prüfenden (Der/Die Vorsichtige; Strenge; Großzügige; Pedantische; Objektive; Subjektive; Konstruktive) die Beziehung zwischen Prüfenden und Studierenden (z.B. der Hierarchieeffekt, wonach gute Leistungen bei „besseren“ Studierenden und Fehler bei „schlechteren“ Studierenden beobachtet werden) oder momentan wirksame situative Faktoren (wie der erste Eindruck, Sympathie/Antipathie, oder die selektive Wahrnehmung des Prüfenden).[49] Beispiele für Beobachtungs- und Beurteilungsfehler sind insbesondere Erwartungseffekte, Projektionsfehler und Halo-Effekt, Reihenfolgeeffekt, der persönliche Maßstab der Prüfenden, Tendenz zu Extremurteile sowie Tendenz zur Mitte.[50]

Nach *Gübeli*[51] können sich die positiven oder negativen Erwartungen von Lehrenden/Prüfenden an die Studierenden wie eine „selbst erfüllende Prophezeiung“ auf die Beurteilung der Leistungen auswirken. Der Erwartungseffekt kann sich aber auch in anderer Weise auf Noten auswirken, etwa wenn es Lehrenden (oft unbewusst) widerstrebt, mehrfach hintereinander ein „Sehr gut“ zu vergeben, in der impliziten Annahme, dass derartig viele „Sehr gut“ nicht hintereinander auftreten können. Beim „Halo-Effekt“ wird von einem Merkmal (z.B. Aussehen der Studie-

48 Vgl. *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 92 (96); *Lienert*, Testaufbau und Testanalyse, S. 95.

49 Siehe dazu *Gübeli*, Beobachtungs- und Beurteilungsfehler, S. 3.

50 Siehe dazu *Gübeli*, Beobachtungs- und Beurteilungsfehler, S. 3 ff.

51 Vgl. *Gübeli*, Beobachtungs- und Beurteilungsfehler, S. 2 ff.

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

renden) auf ein anderes, völlig unabhängiges Merkmal „geschlossen". Insofern kann die Wahrnehmung der Studierenden deren Leistung „überstrahlen". Auch die Reihenfolge kann Einfluss auf die Beurteilung haben. So besteht oft die Tendenz, die ersten Arbeiten strenger zu beurteilen als die letzten. Auch gibt es die Neigung, eine mäßige Arbeit besser zu beurteilen, wenn vorher einige sehr schlechte Arbeiten gelesen wurden bzw. umgekehrt. Auch der persönliche Maßstab der Prüfenden hat entscheidenden Einfluss auf die Notenvergabe. Aufgrund der unterschiedlichen Persönlichkeitstypen von Prüfenden kann es Tendenzen zu Extremurteilen oder zur Mitte geben. Im ersten Fall wird weiter zwischen „guten" und „schlechten" Leistungen unterschieden: Ist die Schwelle „gut" erreicht wird ungerechtfertigt die beste Beurteilung vergeben und umgekehrt. Insbesondere unsichere Prüfende entwickeln bisweilen eine Tendenz zur Mitte.

Darüber hinaus fördert MC – im Vergleich zu den klassischen mündlichen und schriftlichen Prüfungen – zusätzliche andere Kompetenzen bei den Studierenden, wie hohe Konzentrationsfähigkeit oder genaues und aufmerksames Lesen über einen längeren Zeitraum. Gleichermaßen wird das selbstständige Erkennen von Fehlern und damit die Kritikfähigkeit von Studierenden gefördert. Bei klassischen schriftlichen oder mündlichen Prüfungen hat das Erkennen und Benennen der „richtigen" Lösung oberste Priorität. Im Gegensatz dazu hat die Kritik an einer vertretenen Position einen vergleichsweise geringeren Stellenwert. In juristischen Berufen ist diese Kompetenz aber zweifellos wichtig. Zu denken ist an das Gegenargumentieren von Schriftsätzen, sei es bei der Klagebeantwortung oder Formulierung von Rechtsmitteln. Alle in den Schriftsätzen enthaltenen Argumente der gegnerischen Partei müssen kritisch hinterfragt und auf ihre inhaltliche Richtigkeit überprüft werden. Diese Kompetenzen werden bei MC besonders gefördert.

F. Multiple Choice und das Ende der (schriftlichen) Ausdrucksfähigkeit im rechtswissenschaftlichen Studium?

Eines der Hauptargumente für die ablehnende Grundhaltung gegenüber MC-Prüfungen ist, dass diese ungeeignet wären, die (schriftliche) Ausdrucksfähigkeit zu fördern.[52] Demnach sei diese Prüfungsform für ein rechtswissenschaftliches Studium ungeeignet, lautet der pauschale Vorwurf.

Er impliziert zunächst, dass die Vermittlung der „Beherrschung und Anwendung der deutschen Sprache" eine Aufgabe des rechtswissenschaftlichen Studiums sei. Demgegenüber wird allerdings auch vertreten, diese Sprachkompetenzen seien eine „*conditio sine qua non*" für die Zulassung zum Studium.[53] Ferner hängt der diesbezügliche „Kompetenzbedarf" zweifellos maßgeblich vom zukünftigen Einsatzgebiet der AbsolventInnen ab. Es macht einen erheblichen Unterschied aus, ob das etwa die Justiz oder die Rechtsabteilung eines international agierenden Unterneh-

52 Ähnlich wie im Folgenden kann die Argumentation zum – tendenziell wohl zutreffenden – Einwand geführt werden, kreatives Denken werde durch MC-Prüfungen nicht gefördert.

53 Vgl. *Warto*, in: Astleitner/Lagodny et al. (Hrsg.), S. 35 (44).

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

mens ist.[54] *Spezifische Ausdrucksfähigkeit* ist insofern, so wird gesagt, eine Zusatzkompetenz, die in der nachfolgenden juristischen Ausbildungsstufe (Unternehmen, Gerichts- und Anwaltspraxis) erworben werden müsse.[55]

Nun soll hier aber nicht bezweifelt werden, dass die sprachliche Ausdrucksfähigkeit ein wesentliches Rüstzeug ausgebildeter Juristen ist. Diese Kompetenz muss daher im Studium überprüft werden, gleichgültig ob sie schon zum Zulassungszeitpunkt bestehen sollte. Überdies geht es auch um die Ausdrucksfähigkeit in der juristischen Fachsprache, die erst im Laufe des Studiums erworben wird.

Auf dem Boden dieses Standpunkts fragt sich sogleich, ob diese Kompetenz in ausreichendem Maße durch herkömmliche schriftliche und mündliche Prüfungen – also ohne MC – überprüft wird. Das steht durchaus nicht fest.[56] Darüber hinaus wäre zu überlegen, welche Lehrveranstaltungen für diese spezifische Kompetenzförderung in Frage kommen. Am ehesten wohl Seminare und Übungen. Ob solche in einem Fach angeboten werden können entscheidet sich maßgeblich nach den Curricula, nicht zuletzt nach dem zur Verfügung stehenden Stundenkontingent.

Die teilweise erheblichen Unterschiede zwingen gleichzeitig zur Arbeitsteilung: ein „kleines" Fach kann nicht im selben Maße sprachliche Ausdrucksfähigkeit vermitteln wie „große" Fächer.[57] Es muss darauf vertrauen, dass diese Fähigkeit in anderen Fächern erworben wird und dann in allen Fächern angewendet werden kann. Insofern ist das Argument „Würden alle MC-Prüfungen durchführen, würde ein wichtiges Ausbildungselement verloren gehen, nämlich die Schulung im ausformulierten juristischen Argumentieren." zwar zutreffend, aber zu vereinfachend. Die Kritik, dass MC-Prüfungen einen ungünstigen Einfluss auf das Lernverhalten der Studierenden hätten, wäre bei ausschließlichem MC-Einsatz berechtigt. Ein solches System wäre daher abzulehnen. Das ist jedoch ein Problem des Prüfungsdesigns in den Curricula insgesamt, nicht ein isoliertes Problem einer MC-Prüfung.[58]

Stärker als bisher wäre, ganz unabhängig von MC-Prüfungen, zu diskutieren, inwieweit besondere Lehrveranstaltungen zur Förderung von wichtigen Kompetenzen wie sprachliche Ausdrucksfähigkeit, aber auch Kompromissfähigkeit, Verhandlungsstrategie, Überzeugungskraft, Körpersprache und Teamfähigkeit, geschaffen werden sollten.[59] Jedenfalls erscheint die Vorstellung gänzlich realitäts-

54 Vgl. zur Kompetenz Beherrschung und Anwendung der englischen Sprache: Ebd., S. 35 (44).

55 Ebd., S. 44.

56 Jeder Prüfer kennt das Problem: Darf man eine mit Rechtschreib- und Ausdrucksfehlern gespickte Arbeit negativ beurteilen, wenn die sachliche Substanz – die man sich vielleicht teilweise aus Kürzeln „zusammenreimen" muss – „an der Grenze" ist? Jeder Prüfer kennt die – teilweise wohl zutreffende – Rechtfertigung: „Das habe ich ohnedies gemeint". Überdies ist das ein Beispiel für enorme Unterschiede zwischen verschiedenen Prüfern (dazu auch oben bei Fn 45): Manche legen besonderen Wert auf den sprachlichen Ausdruck, andere versuchen davon weitgehend zugunsten der sachlichen Substanz der Antwort zu abstrahieren.

57 Bei insgesamt 4 Wochenstunden in einem Fach etwa ist es so gut wie unvertretbar, ein zweistündiges Seminar durchzuführen.

58 Vgl. *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 92 (99).

59 Vgl. *Warto*, in: Astleitner/Lagodny et al. (Hrsg.), S. 35 (44).

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

fern, dass alle rechtswissenschaftlichen Teilfächer – auch jene mit geringem Stundenkontingent – im gleichen Umfang geeignet wären, diese Zusatzkompetenzen neben den fachlichen Kernkompetenzen zu vermitteln. Nach einer intensiven Auseinandersetzung mit dieser konkreten Fragestellung wurde für die Fachprüfung Europarecht der Universität Salzburg, welche ein Stundenkontingent von 6 ECTS aufweist, die Möglichkeit verneint, diese Kompetenzen zu unterrichten; demgemäß erscheint eine MC-Prüfung für dieses Fach vertretbar und adäquat.

Im Gesamtkontext eines rechtswissenschaftlichen Studiums mit ausgewogenem Einsatz der Prüfungsformen führen MC-Prüfungen dementsprechend nicht zum Ende der schriftlichen Ausdrucksfähigkeit.

G. Multiple Choice im Vergleich mit klassischen schriftlichen Prüfungen anhand von Testgütekriterien

Ein Vergleich zwischen MC- und herkömmlichen schriftlichen Prüfungen mit offener Antwortmöglichkeit ist anhand von Testgütekriterien möglich. Ein „guter" Test besteht laut *Lienert*[60] aus drei Haupt- und vier Nebengütekriterien. Zu den drei Hauptgütekriterien zählen die Objektivität, die Validität (Zuverlässigkeit) sowie die Reliabilität (Gültigkeit). Als bedingte Forderung an einen Test nennt er noch, dass ein Test normiert, vergleichbar, ökonomisch und nützlich sein soll. *Metzger* und *Nüesch* führen für Prüfungen im Hochschulwesen vier Anforderungskriterien an: sie sollen demnach gültig, zuverlässig, ökonomisch und fair (Chancengleichheit für alle Studierenden) sein.[61]

Bei der Objektivität wird zwischen der Durchführungs-, Auswertungs- und Interpretationsobjektivität unterschieden. Das bedeutet, dass jeder Studierende unter den gleichen Testbedingungen geprüft wird, die Auswertung bei allen Studierenden standardisiert erfolgt sowie die Prüfung anhand einheitlicher Kriterien bewertet wird.[62] Diesen Anforderungen entsprechen MC-Prüfungen in der Regel, bei einer Mehrzahl von Prüfern sogar besser als bei offenen Fragen, gleich ob mündlich oder schriftlich. MC-Prüfungen stellen eine prüferunabhängige Benotung sicher. Beurteilungsfehler, wie sie – oft unbewusst – bei mündlichen oder schriftlichen Prüfungen mit offenen Fragen auftreten, sind bei MC-Prüfungen durch die standardisierte Auswertung ausgeschlossen.

Die Validität oder Gültigkeit einer Prüfung gibt den Grad der Genauigkeit an, mit dem sie dasjenige Merkmal, das sie messen soll, tatsächlich misst.[63] Man unterscheidet hierbei inhalts-, konstrukt- und kriterienbezogene Validität.[64] Inhaltsvalide Prüfungen repräsentieren den Inhalt der Lehre, während die Konstruktvalidität

60 Vgl. *Lienert*, Testaufbau und Testanalyse, S. 7.

61 Vgl. *Metzger/Nüesch*, S. 6.

62 Vgl. *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 92 (96); *Lienert*, Testaufbau und Testanalyse, S. 7.

63 Vgl. *Lienert*, Testaufbau und Testanalyse, S. 10 ff.

64 Vgl. *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 92 (97).

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

untersucht, ob die Prüfung auch alle Aspekte der Lehre beinhaltet. Bei der Kriteriumsvalidität wird zwischen Übereinstimmungs- und Vorhersagevalidität unterschieden.[65] Die Validität ist bei MC-Prüfungen nicht von vornherein gegeben, sondern bei der Erstellung der Prüfung zu beachten.[66] Die Herausforderung besteht insbesondere darin, Teilantworten so zu formulieren, dass die gewünschten Lernergebnisse tatsächlich überprüft werden.

Reliabilität bezeichnet die Beständigkeit und Zuverlässigkeit von Beurteilungsergebnissen, d.h. die Genauigkeit von Messungen, Messinstrumenten und Ergebnissen.[67] Sie gibt Auskunft darüber, wie genau etwas gemessen wird, sagt aber nichts darüber aus, was tatsächlich gemessen wird, und bezeichnet die Stetigkeit von Testergebnissen von einer Messung zur nächsten.[68]

Es gibt drei Methoden zur Bestimmung der Reliabilität: die Retest-Reliabilität, die Paralleltest-Reliabilität und die Verfahren zur Ermittlung der inneren Konsistenz.[69]

Die Reliabilität von MC-Prüfungen ist dann sehr hoch, wenn das Beurteilungsverfahren vorgegeben und genau strukturiert ist.[70]

Je konkreter die Lernziele definiert werden, desto leichter fällt es, passende Fragen zu einem Lernziel zu formulieren. Eine Aussage wie „Offene Fragen sind immer besser, weil schwieriger zu lösen" ist nicht richtig. Zum einen mag das Lernziel gar nicht verlangen, dass der Prüfling einen Begriff frei reproduzieren kann. Eventuell genügt es auch in der Praxis, einen Begriff oder Sachverhalt als falsch oder richtig zu erkennen.[71] Zum anderen kann durch geschickt formulierte Antwortalternativen getestet werden, ob Studierende häufigen Fehlmeinungen unterliegen.[72]

Die Vergleichbarkeit von offenen, schriftlichen Prüfungsfragen und MC-Aufgaben wurde durch empirische Studien – von *Kubinger, Holocher-Ertl* und *Frebort* (2006)[73] sowie von *Kubinger* und *Gottschall* (2007)[74] – geprüft. Sie kamen einerseits zum Ergebnis, dass MC-Aufgaben im Antwortformat „1 aus 6" signifikant leichter sind als MC-Aufgaben im Antwortformat „x aus 5" und Aufgaben im freien Antwortformat. Jedoch wurde andererseits auch festgestellt, dass die Aufgaben-

65 Ebd., S. 97.

66 Vgl. *Lienert*, Testaufbau und Testanalyse, S. 10 ff.

67 Ebd., S. 10.

68 Ebd., S. 10 ff.

69 Ebd. S. 10 ff., vgl. *Gronlund*, Assessment of student achievment, S. 212; *Brauns/Schubert*, in: Dany/Szczyrba et al. (Hrsg.), S. 92 (96 f.); *Lienert*, Testaufbau und Testanalyse, S. 10 f.

70 Ebd., S. 10 ff.

71 Vgl. *Vogt/Schneider*, E-Klausuren an Hochschulen, S. 11.

72 Ebd., S. 11.

73 Vgl. *Kubinger/Holocher-Ertl/Frebort*, in: Gula/Alexandrowicz et al. (Hrsg), S. 459 ff.; *Vogt/Schneider*, E-Klausuren an Hochschulen, 2009; *Gottschall/Kubinger*, Item difficulty of multiple choice tests dependant on different item response formats, S. 361 ff.

74 Vgl. *Gottschall/Kubinger*, Item difficulty of multiple choice tests dependant on different item response formats, S. 361 ff.

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

schwierigkeit im Antwortformat MC „2 aus 5“ bzw „x aus 5“ jener des freien Antwortformates entspricht.

Ein direkter Vergleich zwischen der klassischen, schriftlichen Prüfung mit offenen Antwortmöglichkeiten und MC-Prüfungen mit dem Format „x aus 5“ führt zu folgendem Ergebnis: Die Objektivität von MC-Prüfungen ist deutlich höher als bei offenen, schriftlichen oder mündlichen Prüfungen (insbesondere bei der standardisierten Auswertung). Dauerhaft und mit Fragenpool eingesetzt sind MC-Prüfungen ökonomischer als schriftliche Prüfungen. MC-Prüfungen sind prüferunabhängig, lassen keinen Spielraum für die klassischen Beurteilungsfehler und sind damit für Studierende fairer. Auch die Zuverlässigkeit von Beurteilungsergebnissen ist höher als bei schriftlichen Prüfungen einzuschätzen. Im Hinblick auf die Validität ist auf die Herausforderung bei der Konstruktion von MC-Prüfungen zu verweisen.

H. Ausblick: Multiple Choice als Chance?

Wie alle Formen von Leistungsnachweisen haben auch MC-Prüfungen ein grundlegendes Ziel: Prüfende möchten, dass ihre Studierenden gewisse Lernziele erreichen. Die Realität wird diesen Wunschvorstellungen nicht immer gerecht. Lernziele, Lernaktivitäten und Leistungskontrolle fallen oft stärker auseinander, als dies den Lehrenden und Prüfenden bewusst ist.[75] Während sich Lehrende oft an den Lehrinhalten orientieren, gehen Studierende bei der Planung ihres Lernprozesses gegenteilig häufig von den Prüfungen aus.[76] Werden Prüfungen nicht sorgfältig konzipiert, lernen Studierende meist anders und anderes als von den Lehrenden angestrebt. Im Bereich der Rechtswissenschaften sind Studierende – im Gegensatz zu der Erwartungshaltung mancher Lehrenden – oft mit dem Auswendiglernen von Faktenwissen vollständig ausgelastet.[77] Das Fehlen von grundlegenden fachdidaktischen Vorarbeiten[78] führt dazu, dass Studierende "overdetailed but underinformed"[79] sind.

Spätestens die Bologna-Reform hat, wo sie durchgeführt wurde, einen Wandel der Hochschullehre bewirkt: Nicht mehr der Aufwand der Lehrenden (Semesterwochenstunden) steht im Zentrum, „sondern der Lernende mit seinem Arbeitsaufwand und den Kompetenzen (learning outcomes), die er im Laufe des Studiums erlangt“[80].

75 Siehe zum constructive alignment: *Anderson/Krathwohl et al.*, A taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, S. 10 ff; *Biggs*, Enhancing teaching through constructive alignment, in: Higher education 1996, S. 347 ff.

76 Vgl. *Baumert/May*, Constructive Alignment als didaktisches Konzept. S. 23 (27).

77 Vgl. *Warto*, in: Astleitner/Lagodny et al. (Hrsg.), S. 35 (36).

78 Vgl auch, *Brockmann/Pilniok*, in: Bork (Hrsg.), S. 71 ff.

79 Vgl. *Warto*, in: Astleitner/Lagodny et al. (Hrsg.), S. 35 (36); zum „Overnewsed but underinformed“-Effekt *Grob/Bieletzke*, Aufbruch in die Informationsgesellschaft, 1998, S. 31.

80 Homepage der Hochschulrektorenkonferenz zum Thema Bologna, http://www.hrk-bologna.de/bologna/de/home/1923.php (17.12.2107).

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

Schlagworte wie „Rechtsdidaktik“, „kompetenzorientiertes Lernen und Lehren“ sowie „Multiple Choice-Prüfungen“ stellen auch das traditionelle, rechtswissenschaftliche Studium vor neue Herausforderungen.

Die Frage nach möglichen Einsatzgebieten von MC-Prüfungen bringt den Wandel besonders deutlich zum Vorschein: Dabei geht es um viel mehr als eine „bloße Prüfungsumstellung“, es geht um den dahinterliegenden, grundlegenden neuen Diskurs hin zu einem kompetenzorientierten Lernen und Lehren sowie einer einheitlichen und prüferunabhängigen Leistungsbeurteilung. Dies setzt jedoch eine Auseinandersetzung mit lernzielorientiertem Lernen und Lehren voraus, die in der geforderten Intensität in Österreich bisher nicht stattgefunden hat.

Vorurteile und das Festhalten an den bekannten Strukturen verhindern, den Wandel als Chance zur Qualitätsverbesserung der Lehre und Prüfungen zu begreifen.

I. Fazit

MC-Prüfungen haben – nicht nur aufgrund steigender Studierendenzahlen – auch in die Rechtswissenschaften Eingang gefunden. Für die Wahl der Prüfungsform gilt grundsätzlich, dass die zu messenden Lernergebnisse die Prüfungsform bestimmen sollten, d.h. vorab sollte klar sein, was geprüft werden soll, um zu entscheiden, welche Frageform sinnvoll ist. MC ist als Prüfinstrument grundsätzlich geeignet, wenn die Lernziele schwerpunktmäßig auf Ebene des Erinnerns, Verstehens, Anwendens, und Analysierens liegen. Aber auch die Lernergebnisse des Bewertens sind in rechtswissenschaftlichen Fächern überprüfbar. Zu beachten ist allerdings, dass die Entwicklung der Prüfungsfragen mit zunehmenden Komplexitätsgrad des Wissens und der kognitiven Prozesse schwieriger wird.

Die Fragenformulierung sowie Testkonstruktion stellt eine besondere Herausforderung für die Prüfer dar. Es ergibt sich ein sehr hoher Arbeits-, Ressourcen und Zeitaufwand. MC-Prüfungen verlagern den Arbeitsaufwand in die Erstellungsphase, während bei der Korrektur der Prüfung viel an Zeit und Ressourcen (insbesondere bei der automatisieren Ausarbeitung) gewonnen wird. Damit das Beantworten von MC-Aufgaben ausschließlich vom Wissen der Studierenden abhängt und die Ratewahrscheinlichkeit möglichst minimiert wird, müssen zudem einer Reihe von inhaltlichen sowie formalen und sprachlichen Regeln eingehalten werden.

MC-Prüfungen sind wegen der vorgegebenen Antwortalternativen nicht geeignet, die schriftliche Ausdrucksfähigkeit der Studierenden zu überprüfen. Aufgrund eines fehlenden Kompetenzmodells für das rechtswissenschaftliche Studium ist jedoch bis dato nicht hinreichend geklärt, ob und in welchem Umfang die Kompetenz „Beherrschung und Anwendung der deutschen Sprache“ im rechtswissenschaftliche Studium vermittelt und abgeprüft werden sollte.

Im Vergleich – anhand von Testgütekriterien – mit der klassischen, schriftlichen Prüfung mit offenen Antwortmöglichkeiten wird ersichtlich, dass die Zuverlässigkeit von Beurteilungsergebnissen bei MC-Prüfungen höher ist. MC bringt als Prü-

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

fungsform ein deutliches höheres und damit neues Maß an Objektivität (insbesondere bei der standardisierten Auswertung) in Form einer prüferunabhängigen Leistungsbeurteilung. MC-Prüfungen lassen keinen Spielraum für die klassischen Beurteilungsfehler und sind damit für Studierende fairer. Wenn dauerhaft eingesetzt, sind MC-Prüfungen ökonomischer als schriftliche Prüfungen. Im Hinblick auf die Validität ist auf die Herausforderung bei der Konstruktion von MC-Prüfungen zu verweisen.

Bei gezieltem Einsatz können MC-Prüfungen als neue Prüfungsform ein Gewinn sowohl für Studierende als auch Lehrende sein. Im Gesamtkontext eines Studiums mit ausgewogenem Einsatz der Prüfungsformen ist eine MC-Prüfung, die den Studierenden sehr hohe Konzentration, genaues und aufmerksames Lesen und aufgrund der hohen Fragenzahl eine umfassende Kenntnis des Stoffes abverlangt, als sinnvolle Ergänzung und damit als Gewinn einzustufen.

Literaturverzeichnis

Anderson, Lorin/Krathwohl, David, A taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, New York: Allyn & Bacon 2001.

Astleitner, Hermann/Wageneder, Günter et al., 12 Tipps für eine kompetenzorientierte Lehre, Salzburg 2015. http://www.uni-salzburg.at/fileadmin/multimedia/Qualitaetsmanagement/documents/Handbuecher/12_Tipps_f%C3%BCr_eine_kompetenzorientierte_Lehre.pdf (13.11.2017).

Asmuth, Markus, Prüfen mit der Multiple-Choice-Methode, Unterricht mit neuen Medien, Schulen ans Netz e.V., 2003. https://www.lehrer-online.de/artikel/fa/pruefen-mit-der-multiple-choice-methode/ (01.3.2017).

Baumert, Britta/May, Dominik, Constructive Alignment als didaktisches Konzept. Journal Hochschulddidaktik 1-2, 2013, S. 23–27.

Biggs, John, Enhancing teaching through constructive alignment, in: Higher education, 1996, pp. 347-364.

Brockmann, Judith/Pilniok, Arne, Prüfen in der Rechtswissenschaft: Stand und Perspektiven der disziplinären Forschung, in: Bork (Hrsg.), Prüfungsforschung, Hamburg 2014, S. 71 – 93.

Bydlinski, Franz, Juristische Methodenlehre und Rechtsbegriff, Wien 1982.

Brauns, Katrin/Schubert, Sebastian, Qualitätssicherung von Multiple-Choice-Prüfungen, in: Dany/Szczyrba/Wildt (Hrsg.), Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen Band 118, Bielefeld 2008, S. 91–102.

Engisch, *Karl,* Logische Studien zur Gesetzesanwendung, 3. Auflage, Heidelberg 1963.

Grob/Bieletzke, Aufbruch in die Informationsgesellschaft, Münster (u.a) 1998.

Gronlund, Norman, Assessment of student achievement, 7th edition, Boston: Ally & Bacon 2003.

Gottschall, Christian/Kubinger, Klaus, Item difficulty of multiple choice tests dependant on different item response formats – An experiment in fundamental research on psychological assessment, in: Psychology Science. Volume 49, 2007, S. 361–374.

Gruber, Simone/Avallone, Manuela, Empfehlungen zum Einsatz von Multiple-Choice-Prüfungen, Leipzig 2012.

https://www.lehren.tum.de/fileadmin/w00bmo/www/Downloads/Themen/Studiengaenge_gestalten/Dokumente/MCEmpfehlungen_Stand_Oktober_2012_final.pdf (13.11.2017).

Homepage der Hochschulrektorenkonferenz zum Thema Bologna, http://www.hrk-bologna.de/bologna/de/home/1923.php (27.10.2017).

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.

Gübeli, Rita, Beobachtungs- und Beurteilungsfehler, Organisation der Arbeitswelt Gesundheit, Zürich 2012, S. 1–5. http://www.z-in-a.ch/files/media/files/a105d723f6a84f5b2902969992c44c32/Beobachtung%20und%20Beurteilung.pdf (1.8.2017)

Jacobs, Bernhard, Richtlinien zur Erstellung von einfachen Multiple-Choice-Aufgaben nach Gronlund, Saarbrücken 2008.

Kennedy, Declan/Hyland, Àine/Ryan, Norma, Writing and Using Learning Outcomes, A Practical Guide, in: Froment/Kohler/Purser/Wilson (Hrsg.), EUA Bologna Handbook (BH), Berlin/Stuttgart 2006, Griffmarke C 3.4–1.

Krebs, René, Multiple Choice Fragen? Ja, aber richtig. Medizinische Fakultät; Institut für Medizinische Lehre IML, Abteilung für Assessment- und Evaluation AAE, Bern 2008. http://home.uni-leipzig.de/~allpaed/wiki/lib/exe/fetch.php?media=qualitaet:krebs-mc_aber_richtig-2008.pdf (13.10.2017)

Krüger, Heike, Ankreuzen kann ja jeder, oder?, in: Brockmann/Pilniok (Hrsg.), Prüfen in der Rechtswissenschaft, Hamburg 2013, S. 109–119.

Kubinger, Klaus, Objektive psychologisch-diagnostische Verfahren, in: Weber/Rammsayer (Hrsg.), Handbuch der Persönlichkeitspsychologie und Differentiellen Psychologie aus Handbuch der Psychologie, Göttingen: Hogrefe 2005, S. 158–165.

Kubinger, Klaus/Holocher-Ertl, Stefana/Frebort, Martina, Zur testtheoretischen Qualität von Multiple Choice-Items: 2 richtige aus 5 vs. 1 richtige aus 6 Antwortmöglichkeiten, in Gula/Alexandrowicz/Strauß/Brunner/Jenull-Schiefer/Vitouch (Hrsg.), Perspektiven psychologischer Forschung in Österreich. Proceedings zur 7. Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie, Lengerich 2006, S. 459 – 464:

Lachmayer, Konrad, Ziele des rechtswissenschaftlichen Studiums und Rechtsdidaktik, in: Zfhr 2012, S. 139–145.

Lienert, Gustav, Testaufbau und Testanalyse, 6. Auflage, Weinheim: Psychologie Verlag Union 1998.

Metzger, Christoph/Nüesch, Charlotte, Fair prüfen – Ein Qualitätsleitfaden für Prüfende an Hochschulen, in: Euler/Metzger (Hrsg.), IWP Institut für Wirtschaftspädagogik (= Hochschuldidaktische Schriften. Band 6). St. Gallen 2004.

Nickel, Sigrun, Der Bologna-Prozess aus Sicht der Hochschulforschung – Analysen und Impulse für die Praxis, Gütersloh 2011.

Reis, Oliver, Hochschuldidaktische Herausforderungen an die Rechtswissenschaft, in: ZDRW 2013, S. 21–44.

Rütter, Theodor, Formen der Testaufgabe, München 1973.

Schermutzki, Margret, Lernergebnisse – Begriffe, Zusammenhänge, Umsetzung und Erfolgsermittlung: Lernergebnisse und Kompetenzvermittlung als elementare Orientierungen des Bologna-Prozesses, Aachen: FH Aachen 2007.

Vogt, Michael/Schneider, Stefan, E-Klausuren an Hochschulen: Didaktik – Technik – Systeme – Recht – Praxis, Gießen 2009. https://elearning.hs-flensburg.de/elearningblog/wp-content/uploads/2015/09/E-klausuren-an-Hochschulen.pdf (10.10.2017).

Warto, Patrick, Kompetenzorientierung in der rechtswissenschaftlichen Ausbildung, in: Astleitner/Lagodny/Warto/Zumbach (Hrsg.), Tagungsband zur 1. Fachtagung Rechtsdidaktik in Österreich, Salzburg 2015, S. 35–50.

Wildt, Johannes/Wildt, Beatrix, Lernprozessorientiertes Prüfen im „Constructive Alignment“, in: Berendt/Voss/Wildt (Hrsg.), Neues Handbuch Hochschullehre: Lehren und Lernen effizient gestalten. (Lieferung 50, Ziffer H 6.1). Berlin: Raabe Fachverlag für Wissenschaftsinformation 2011.

https://doi.org/10.5771/2196-7261-2018-2-148
Generiert durch IP '18.118.20.118', am 12.10.2024, 11:39:59.
Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.