Qualitätsverbesserung von MC Fragen

Ein exemplarischer Weg für eine medizinische Fakultät

Quality assurance of Multiple Choice Questions

An exemplary way for a medical faculty

• Thomas Rotthoff[1] • Sibylle Soboll[2]

Zusammenfassung:
Wegen fehlender notenrelevanter Prüfungen an den Medizinischen Fakultäten in Deutschland bestand bisher keine Notwendigkeit, die Fragenqualität in schriftlichen Prüfungen konsequent zu reflektieren. Erst durch die neue Approbationsordnung sind zeugnisrelevante, fakultätsinterne Prüfungen vorgeschrieben. Es fehlen somit oftmals Strukturen und Prozesse, die zu einer Verbesserung der Fragenqualität führen. Damit die Klausuren das unterschiedliche Leistungsspektrum der Studierenden widerspiegeln, sind u.a. unterschiedliches Schwierigkeitsniveau und eine gute Trennschärfe der Prüfungsfragen notwendig.
Für eine fachübergreifende Klausur des vierten Studienjahres an der Universität Düsseldorf sollten neue MC -Fragen entwickelt werden, die anwendungsorientiert prüfen, eindeutig formuliert und weitgehend frei von formalen Fehlern sind.
Die Umsetzung erfolgte in der Konzeption und Durchführung von Workshops zur Erstellung von MC-Fragen für Fragenautoren und die Einrichtung eines interdisziplinären Reviewkomitees.
Es konnte gezeigt werden, dass eine Autorenschulung den Reviewprozess für das Reviewkomitee erleichtert und beschleunigt. Der Nutzen des Reviewprozesses spiegelt sich in einer hohen Anwendungsorientierung der einzelnen Items wider. Hochwertige, in einem Reviewverfahren erstellte und messtechnisch analysierte Prüfungsfragen könnten künftig in interuniversitäre Datenbanken eingegeben werden und damit den anfänglich erforderlichen Zeitaufwand relativieren. Durch die interdisziplinäre Zusammensetzung des Reviewkomitees bietet sich außerdem die Möglichkeit einer verstärkten Diskussion über die Inhalte der Prüfungsfragen.
Schlüsselwörter:Prüfung, MC-Fragen, Multiple-Choice-Frage, Review, Medizinische Ausbildung

Abstract:
Because of the missing relevance of graded examinations at the German medical faculties, there was no need to reflect the question quality in written examinations consistently. Through the new national legislation for medical education certification-relevant, faculty-internal examinations are prescribed. Until now, there is a lack of structures and processes which could lead to an improvement of the question quality. To reflect the different performance of the students, a different severity and a good selectivity of the test questions are necessary.
For a interdisciplinary examination for fourth year undergraduate students at the University Hospital Duesseldorf, new Multiple choice (MC)- questions which are application-orientated, clearly formulated and to a large extent free from formal errors should be developed.
The implementation took place in the conception and performance of Workshops for the construction of MC-questions and the appointment of an interdisciplinary review-committee.
It could be shown that an author training facilitates and accelerates the review-process for the committee and that a review process reflects itself in a high practise-orientation of the items.
Prospectively, high-quality questions which are created in a review-process and metrological analysed could be read into inter-university databases. Therewith the initial expenditure of time could be reduced. The interdisciplinary constitution of the review-committee offers the possibility of an intensified discussion over content and relevance of the questions.
Keywords:assessment, MC-Questions, Multiple-Choice-Question, Review, Medical Education

Einleitung

Durch die Novellierung der ärztlichen Approbationsordnung (ÄAppO) werden Änderungen in der Ausbildung und bei den Prüfungen an den medizinischen Fakultäten in Deutschland notwendig. Bisher dienten die fakultätsinternen Klausuren im klinischen Abschnitt ausschließlich dem Scheinerwerb, ohne dass eine Differenzierung der Prüfungsergebnisse erfolgen musste. Die an der Heinrich-Heine-Universität Düsseldorf (HHU) im klinischen Abschnitt eingesetzten MC-Fragen prüften überwiegend descriptives, weniger anwendungsorientiertes Wissen und in der Regel waren die Autoren in der Erstellung von MC-Fragen nicht geschult. Für die Klausuren wurde auch auf einen den Studierenden bekannten Altfragenpool zurückgegriffen, woraus Bestehensquoten von teilweise über 95% resultierten.

In der neuen ÄAppO sind nun erstmalig benotete Leistungsnachweise in allen klinischen Fächern, einschließlich der Querschnittbereiche vorgeschrieben [1]. Diese Noten werden auf dem Abschlusszeugnis gesondert ausgewiesen, womit die Prüfungen sowohl für die Studierenden als auch für die Hochschulen eine neue Qualität bekommen haben. Die Bedeutung der Noten ist bislang insbesondere im Hinblick auf die Vergleichbarkeit zwischen den

[1] *Universitätsklinikum Düsseldorf, Klinik für Endokrinologie, Diabetologie und Rheumatologie, Düsseldorf, Deutschland*

[2] *Heinrich-Heine-Universität Düsseldorf, Medizinische Fakultät, Dekanat, Düsseldorf, Deutschland*

Artikel online frei zugänglich unter http://www.egms.de/en/journals/zma/2006-23/zma000264.shtml
Bitte zitieren als: Rotthoff T, Soboll S. Qualitätsverbesserung von MC Fragen: Ein exemplarischer Weg für eine medizinische Fakultät. GMS Z Med Ausbild. 2006;23(3):Doc45.
gQp020060785 31

Hochschulen unklar. Ein strukturierter Qualitätssicherungsprozess ist aber vor diesem Hintergrund unabdingbar. Zusätzlich soll die Anfechtbarkeit der Prüfung durch eine eindeutige Formulierung der Prüfungsfragen und damit eine juristische Auseinandersetzung mit den Studierenden vermieden werden.

Projektbeschreibung

Ziel des Projektes war die Entwicklung qualitativ hochwertiger MC-Fragen, die mehr prüfen sollten als faktisches Wissen und weitgehend frei sind von formalen Fehlern wie z.B. Cues und schlechten Distraktoren [2]. Es sollte den Fragen nachgegangen werden, ob durch eine Dozentenschulung und Implementierung eines Reviewprozesses die Qualität von MC-Fragen verbessert werden kann und ob Fragen von geschulten Autoren den Reviewprozess schneller durchlaufen, als die Fragen von nicht geschulten Autoren.

Methode

Eine der ersten Prüfungen mit benoteten Leistungsnachweisen wurde im klinischen Abschnitt an der HHU am Ende des WiSe 2004/2005 als Abschluss des 2. klinischen Studienjahres durchgeführt. Für dieses Studienjahr wurde mit Beginn des SoSe 04 ein fachübergreifender Unterricht eingeführt, der sich in 24 symptombezogene Wochenmodule aufteilt. Jede Modulwoche wird von einem modulverantwortlichen Oberarzt bzw. Oberärztin betreut. An den Modulen sind die Fächer Innere Medizin, Chirurgie, klinische Chemie, Gynäkologie, Urologie und Orthopädie beteiligt. Die Abschlussprüfung setzt sich aus Fragen zu den Modulen sowie den begleitenden Vorlesungen zusammen. Dabei bilden jeweils drei Fächer- Innere Medizin, Gynäkologie und Klinische Chemie bzw. Chirurgie, Orthopädie und Urologie einen gemeinsam benoteten fachübergreifenden Leistungsnachweis. Die drei Fächer eines jeden fachübergreifenden Leistungsnachweises werden zusammen in einer Klausur geprüft. Zusätzlich wird auch das Ergebnis für jedes Einzelfach benotet und auf dem späteren universitären Abschlusszeugnis ausgewiesen. Für die Fächer Innere Medizin und Chirurgie wurden je 40 Fragen und für die übrigen Fächer je 20 Fragen gestellt. Auf Beschluss der zuständigen Unterrichtskommission sollte die Entwicklung neuer qualitätsgeprüfter Fragen zunächst als Pilotphase auf die Lerninhalte der symptombezogenen Module beschränkt werden. Die Fragen sollten dann, nach erfolgtem Review, gemäß ihres Inhaltes einzelnen Fächern für die spätere Bewertung zugeordnet werden. Die restlichen Prüfungsfragen wurden einem Altfragenpool entnommen. Allen 24 Modulleitern wurde angeboten, an einer Schulung zur Erstellung von MC-Fragen teilzunehmen (siehe Abbildung 1).

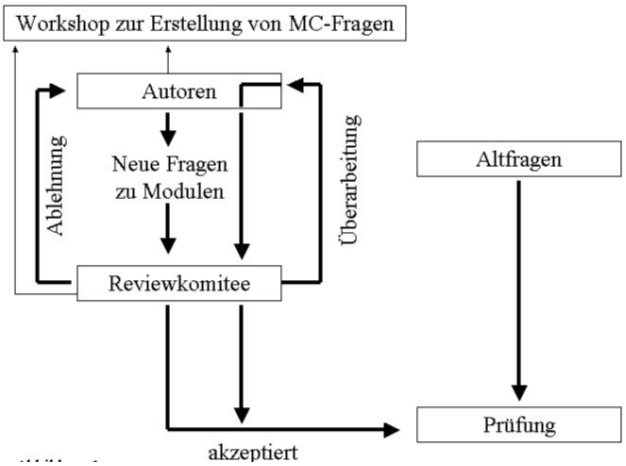

Abbildung 1: Organisationsstrukturen und Arbeitsabläufe

Zwei Dozenten der HHU entwickelten ein Workshopkonzept zur Erstellung von MC-Fragen anhand publizierter Kriterien und Standards [3], [4]. In dem Düsseldorfer Workshop wurden im Oktober 2004 unterschiedliche MC Fragen-Typen und die Prinzipien zur Fall-, Fragen- und Antwortkonstruktion vorgestellt. Die Teilnehmer (Modulleiter) sollten eigenständig Düsseldorfer Klausurfragen vom SoSe 2004 in Kleingruppen in Bezug auf Qualitätsmerkmale analysieren und die Ergebnisanalyse anschließend präsentieren. Die Teilnehmer erhielten Informationen über die Erstellung einer Fallvignette, die Konstruktion von Antworten bzw. Distraktoren und sie sollten eigene Fragen entsprechend modifizieren.

Außerdem wurde erstmalig für die Begutachtung der Fragen ein neues interdisziplinäres Review-Komitee eingerichtet. Bei der Bildung wurde auf eine interdisziplinäre Zusammensetzung unter Einbindung möglichst vieler klinischen Fächer geachtet. (Innere Medizin, Chirurgie, Gynäkologie, Urologie, Allgemeinmedizin und klinische Chemie). Insgesamt bestand das Komitee aus sieben freiwilligen Mitgliedern einschließlich des Projektleiters. Es waren sowohl Oberärzte als auch AssistenzärztInnen vertreten. Die Mitglieder des Reviewkomitees wurden ebenfalls mit dem o.g. Workshopformat geschult, erhielten jedoch darüber hinaus noch Hintergrundinformationen zu prüfungstechnischen Begriffen.

Insgesamt wurden 24 Modulleiter aufgefordert, eine Fallvignette mit zwei anwendungsorientierten Fragen zu erstellen. Allen Fragenautoren wurde ein publiziertes, auf die Düsseldorfer Situation angepasstes Online-Formular zur Fragenerstellung zur Verfügung gestellt [3]. Mit diesem Formular wurden die Fragen standardisiert per e-Mail beim Reviewkomitee eingereicht. Das Formular enthielt Angaben über Autor, Prüfziel und die Quellenangabe der richtigen Lösung sowie ein Feld für die Stellungnahme des Reviewkomitees. Die jeweiligen Fristen für die Eingabe der Fragen durch die Autoren wurden vom Reviewkomitee festgelegt. Nach Abschluss des Reviewprozesses lagen 37 Fragen vor, wovon 27 Fragen dem Fach Innere Medizin, 7 Fragen der Chirurgie und 3 Fragen der Orthopädie zugewiesen wurden. Die Fragen waren inhaltlich allerdings nicht nach einem Blueprint gewichtet . Die Umsetzung des Projektes wurde durch qualitätssichernde Maßnahmen begleitet. Jede neu entwickelte Frage wurde auf ihrem Weg durch den Reviewprozess verfolgt. Innerhalb des Reviewprozesses konnte eine Frage akzeptiert, zur Überarbeitung an den Autor zurückgegeben oder

Artikel online frei zugänglich unter http://www.egms.de/en/journals/zma/2006-23/zma000264.shtml
Bitte zitieren als: Rotthoff T, Soboll S. Qualitätsverbesserung von MC Fragen: Ein exemplarischer Weg für eine medizinische Fakultät. GMS Z Med Ausbild. 2006;23(3):Doc45.

auch abgelehnt werden (siehe Abbildung 1). Dabei waren nur dem Projektleiter die Namen aller geschulten Autoren bekannt. Nach stattgehabter Klausur sollten alle neu erstellten und in die Klausur aufgenommenen Fragen auf ihre Itemschwierigkeit und Trennschärfe analysiert werden.

Die Berechnung der korrigierten Trennschärfen und Reliabilität (Cronbach´s alpha) erfolgte mit der SAS Software [5].

Ergebnisse

Insgesamt nahmen 19 Personen, davon 13 Modulleiter und sechs Mitglieder des Reviewkomitees, an einem halbtägigen Workshop zur Erstellung von MC-Fragen teil. 52 Fragen wurden beim Review-Komitee zur Begutachtung eingereicht. Innerhalb des ersten Reviewprozesses (zwei Sitzungen) wurden 14 Fragen direkt akzeptiert, 30 an die Autoren zurückgegeben und 8 Fragen abgelehnt (siehe Abbildung 2). Es zeigte sich, dass die Fragen von geschulten Autoren schneller akzeptiert wurden als die Fragen von ungeschulten Autoren. Geschulte Autoren nutzten überwiegend Fragen mit Einfachauswahl aus fünf Antworten mit einer guten Fallvignette und weniger Cues. Ungeschulte Autoren benutzten häufig Typ K Fragen (Antwortkombinationen), die wegen ihrer messtechnischen Unschärfe international nicht mehr empfohlen werden. Alle Autoren erhielten das kommentierte Online-Formular zurück und wurden so über das Ergebnis des Reviewprozesses informiert. Gleichzeitig wurden sie aufgefordert die Fragen zu überarbeiten oder bei Ablehnung, neue Fragen zu erstellen. Bis zur Abgabefrist erhielt das Reviewkomitee 31 überarbeitete bzw. neu erstellte Fragen zurück. Nicht alle Fragenautoren kamen der Aufforderung nach, die Prüfungsfragen entsprechend der Kommentare des Reviewkomitees zu bearbeiten. Von den geschulten Autoren wurde fristgerecht nur die Hälfte der zu überarbeitenden oder neu zu formulierenden Fragen für einen zweiten Reviewprozess eingereicht. Ein nicht geschulter Autor sandte noch eine zusätzliche Frage für den zweiten Reviewprozess ein. Insgesamt konnten acht Fragen auch im zweiten Reviewverfahren nicht akzeptiert werden. Die Gründe lagen in der unklaren Gesamtkonzeption der Fragen, dem nicht erkennbaren Prüfziel, inadäquaten Distraktoren, Cues und einer unbefriedigenden Fallvignette. Insgesamt wurden 37 der eingereichten Fragen für die Klausur akzeptiert.

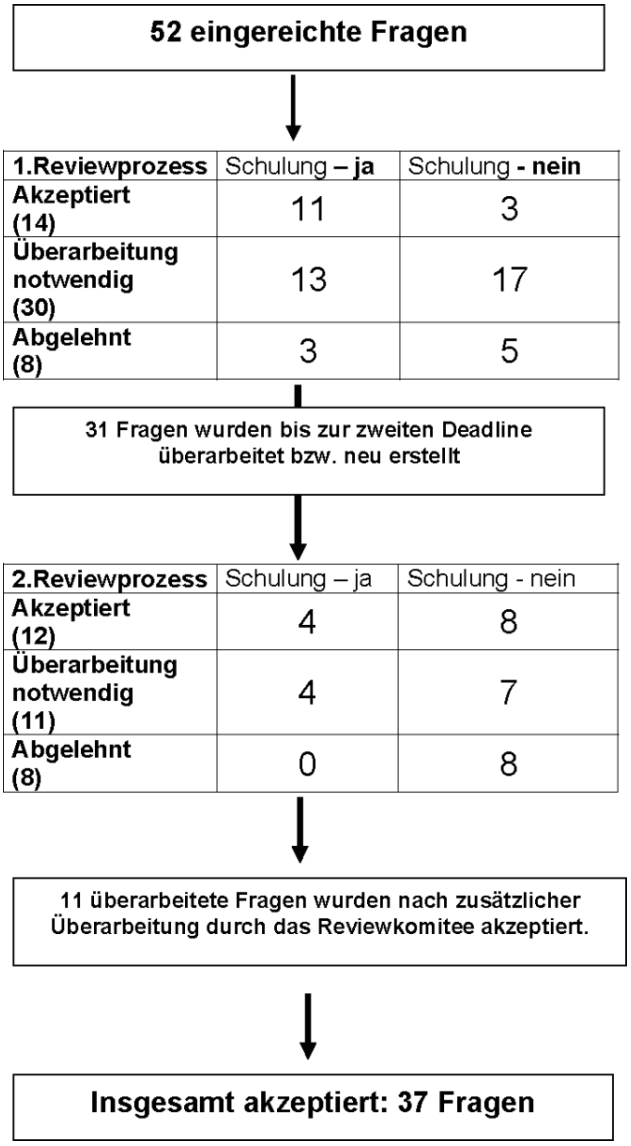

Abbildung 2: Fragen im Reviewprozess

• Trennschärfe

An der Prüfung nahmen 113 Studierende teil. Die Antwortbögen wurden elektronisch ausgewertet und für jedes Item die korrigierte Trennschärfe berechnet. Die Trennschärfe gibt die Fähigkeit eines Items an, Kandidaten mit guter und schlechter Leistung in der Gesamtprüfung zu trennen [6]. Um eine valide Aussage über die Trennschärfe der neuen Fragen zu erhalten, wurden diese zusätzlich aus der Gesamtmenge der 100 Fragen aus den Fächern Innere Medizin, Chirurgie und Orthopädie extrahiert und einer separaten Trennschärfenanalyse zugeführt. Diesen drei Fächern waren die 37 Fragen aus dem Review zugeordnet worden. Die mittlere korrigierte Trennschärfe aller 37 neuen Items lag separat gemessen bei 0,16 gegenüber 0,11 bei den Altfragen der o.g. Fächer (siehe Tabelle 1). Insgesamt hatten von den 37 neuen Items 15 Items (40,5%) eine Trennschärfe = 0,2 und je 11 Items (29,7%) eine Trennschärfe zwischen 0,1 und 0,2. bzw. < 0,1, wovon fünf Items eine Trennschärfe kleiner Null aufwiesen (siehe Abbildung 3). Bei diesen fünf Items handelte es ausschließlich um chirurgische Items. Die Reliabilität (Crohnbach´s alpha) ist mit einem Wert

Artikel online frei zugänglich unter http://www.egms.de/en/journals/zma/2006-23/zma000264.shtml
Bitte zitieren als: Rotthoff T, Soboll S. Qualitätsverbesserung von MC Fragen: Ein exemplarischer Weg für eine medizinische Fakultät. GMS Z Med Ausbild. 2006;23(3):Doc45.

von 0,71 für die Gesamtzahl der Fragen noch unbefriedigend. Eine hypothetische Testverlängerung mit den neuen Fragen könnte zu einem reliableren Prüfungsergebnis (0,83) führen (siehe Tabelle 1).

Tabelle 1: Basiskennwerte und Reliabilitätskoeffizienten im Vergleich

Innere Medizin Chirurgie Orthopädie	Mittlere (korr.) Trennschärfe				Mittlere Itemschwierigkeit				Reliabilitäts-koeffizient (Cronbach´s Alpha)	Reliabilität bei Testverlängerung auf 100 Fragen (nach Spearman-Brown)
	Mittel	Std.-abw.	Min	Max	Mittel	Std.-abw.	Min	Max		
Altfragen (n=63)	0,11	0,17	-0,11	0,36	0,69	0,23	0,13	1,0	0,61	0,71
Neue Fragen (n=37)	0,16	0,15	-0,12	0,36	0,57	0,22	0,15	0,96	0,64	0,83
Alle Fragen (n=100)	0,12	0,17	-0,13	0,39	0,65	0,23	0,15	1,0	0,71	

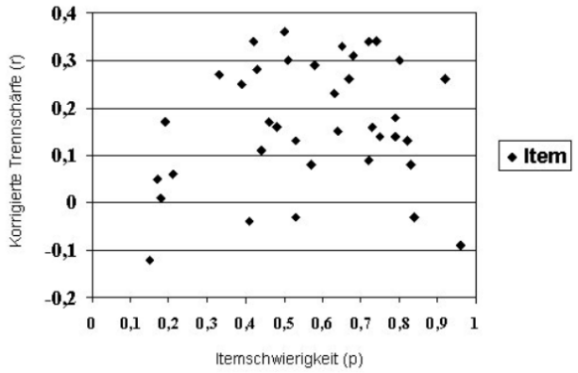

Abbildung 3: Trennschärfen und Schwierigkeiten der neuen Items

• Itemschwierigkeit

Jedes Item wurde auf seine Schwierigkeit (mittlere erreichte Punktzahl je Item durch die mögliche Gesamtpunktzahl) hin überprüft. Die mittlere Schwierigkeit der neuen Fragen lag bei 0,57 bzw. 0,69 bei den alten Fragen (Tabelle 1). Bei einer anzustrebenden Itemschwierigkeit zwischen 0,4 und 0,9 mussten 7 Items (18,9%) mit einem Wert < 0,4 als schwer und nur ein Item mit einem Wert >0,9 (2,7%) als sehr leicht gewertet werden. 29 Items (78,4%) hatten eine adäquate Schwierigkeit zwischen 0,4 und 0,9.

Diskussion

• Schulung und Reviewverfahren

Bei summativen Prüfungen wie z.B. Staatsexamina, ist ein Review der Prüfungsfragen internationaler Standard. Ein interdisziplinärer Reviewprozess war im klinischen Abschnitt der HHU bisher noch nicht implementiert. Wegen fehlender notenrelevanter Prüfungen existierten bisher wenig Anreize die Fragenqualität konsequent zu reflektierten. Es fehlten somit Strukturen und Prozesse, die zu einer Verbesserung der Fragenqualität führen könnten. Der dargestellte Prozess bestätigt, dass eine Autorenschulung zur Erstellung von MC-Fragen und die Einrichtung eines Reviewkomitees die Fragenqualität verbessert und die Einflüsse sachfremder Faktoren und Zufälligkeiten auf das Prüfungsergebnis reduziert. Einschränkend muss jedoch darauf hingewiesen werden, dass alle neu erstellten Fragen in diesem Projekt nur mit Altfragen verglichen wurden. Eine Aussage über die Qualität neuer Fragen die von ungeschulten Autoren erstellt und keinen Reviewprozess durchlaufen haben, ist nicht möglich. Es konnte aber gezeigt werden, dass die Schulung der Fragenautoren den Reviewprozess durch eine größere Akzeptanz primär eingereichter Fragen und einen geringeren Korrekturaufwand erleichtert und beschleunigt, obwohl auch bei den Autoren eine unterschiedliche Talentierung bei der Fragenerstellung erkennbar ist. Trotz einer schulungsbedingten Beschleunigung des Reviewprozesses benötigte der formale Review aller Fragen einen Zeitaufwand von ca. 6 Stunden in gemeinschaftlicher Sitzung. Der Zeitaufwand ist mit dem anderer Fakultäten vergleichbar. Der Nutzen dieses aufwendigen Reviewprozesses spiegelt sich insbesondere in einer besseren Anwendungsorientierung der einzelnen Items wider, wodurch eine höhere Taxonomie kognitiver Lernziele erreicht wird [7]. Die neuen Items prüften klinische Entscheidungsfindungen und nicht ausschließlich Faktenwissen.

• Trennschärfe

Die Angaben bezüglich einer optimalen Trennschärfe sind in der Literatur uneinheitlich. Für fakultätsinterne Prüfungen mit häufig neu generierten Fragen wird bereits eine zuverlässige Leistungsdifferenzierung ab einer Trennschärfe $\geq 0,1$ möglichst jedoch $\geq 0,2$ genannt [8]. In der psychometrischen Literatur werden dagegen Trennschärfen von 0,3 - 0,5 als mittel und erst solche > 0,5 als hoch bezeichnet [8]. Diese beziehen sich jedoch in erster Linie auf käufliche, wiederholt eingesetzte und in ihrer Fragenzusammensetzung konstante psychometrische Tests. Wegen des großen zeitlichen und personellen Aufwandes für eine Fakultät kontinuierlich neue Fragen zu generieren, ist eine Verwerfung von Fragen mit einer Trennschärfe < 0,3 kaum praktikabel. Auf den ersten Blick erscheint die Trennschärfenanalyse für einige neue Items enttäuschend niedrig. Die Trennschärfenanalyse ist aber nicht nur abhängig von der Kandidatengruppe sondern auch von der Fragenzusammensetzung einer Prüfung. Bevor gut formulierte und inhaltlich relevante Fragen eliminiert werden, erscheint es sinnvoll diese bei anderen Kandidaten und in Kombination mit anderen Items einzusetzen und die Messergebnisse zu vergleichen. Psychometrische Messergebnisse sind zwar wichtig für die Beurteilung eines Items, sollten aber auch nicht überbewertet werden, wenn das Item durch eine gute Anwendungsorientierung und klinischen Bezug zur Validität der Prüfung beiträgt [9], [10]. Im Gegensatz zu standardisierten psychologischen Tests, in denen in der Regel ein Score aus mehreren Items eine Aussage beispielsweise zu einem Symptomkomplex ermöglicht, steht in einer medizinischen MC-Prüfung jedes einzelne Item für sich. Deshalb sollte sowohl eine Überprüfung der formalen Konzeption, die Berechnung von Trennschärfe und Schwierigkeit aber auch eine inhaltliche Analyse jedes einzelnen Items erfolgen bevor über seine weitere Verwendung entschieden wird [9].

Ausblick

In einem Reviewverfahren erstellte und messtechnisch analysierte Prüfungsfragen können künftig in Prüfungsdatenbanken eingegeben werden, wie sie teilweise schon eingesetzt werden (Progress Test Berlin) oder sich momentan im Aufbau befinden (z.B. an der Universität Heidelberg oder für die medizinischen Fakultäten in NRW an der Universität Münster). Im Rahmen eines Prüfungsverbundes mehrerer Fakultäten kann so eine kritische Masse an Prüfungsfragen erstellt und genutzt werden und damit die einzelnen Fakultäten in der Erstellung von Prüfungsfragen entlasten. Neben

Artikel online frei zugänglich unter http://www.egms.de/en/journals/zma/2006-23/zma000264.shtml
Bitte zitieren als: Rotthoff T, Soboll S. Qualitätsverbesserung von MC Fragen: Ein exemplarischer Weg für eine medizinische Fakultät. GMS Z Med Ausbild. 2006;23(3):Doc45.

der qualitativen Fragenverbesserung ist zu fordern, dass der interdisziplinäre Reviewprozess auch zu einer inhaltlichen Einflussnahme auf das Curriculum führt. Mittlerweile werden in der Fortsetzung des Projektes die Prüfungsfragen vom Reviewkomitee auch inhaltlich kommentiert und nur dann für die Prüfung akzeptiert, wenn das Prüfziel ersichtlich ist und der Frageninhalt sich mit den definierten Lernzielen deckt. Die Erstellung eines Blueprints für die Prüfungsinhalte fand bisher bei den Dozenten noch keine Akzeptanz. Seit dem SoSe 2005 werden die Prüfungen zusätzlich von einem Statistiker ausgewertet und die Ergebnisse den Fachbereichen zur Verfügung gestellt. Geplant ist die Erstellung eines Prüfberichtes für jede Prüfung zur Qualitätskontrolle und Verbesserung.

Schlussfolgerung

Die Arbeit zeigt exemplarisch für eine medizinische Fakultät, dass Schulungen von Fragenautoren sowie die Einrichtung eines Reviewkomitees die Qualität von MC-Fragen formal und inhaltlich verbessert. Die Autorenschulung beschleunigt den Reviewprozess durch die Berücksichtigung formaler Vorgaben bei der Fragenerstellung. Dabei kann sich die Investition in den zeitaufwendigen Reviewprozess durch einen interuniversitären Austausch qualitätsgeprüfter Fragen künftig relativieren. Die Fortsetzung des beschriebenen Projektes wurde durch die Unterrichtskommission beschlossen und personell unterstützt. Zwischenzeitlich wurde das Projekt in anderen Unterrichtskommissionen vorgestellt und die Gründung von zwei weiteren Reviewkomitees nach gleichem Muster beschlossen. Eines ist bereits seit dem SoSe 2005 für das 3. klinische Studienjahr tätig.

Unter Sicherheitsaspekten ist das aktuelle Verfahren, die Fragen per e-Mail an die Autoren zu verschicken, problematisch. Hier ist künftig der Einsatz eines gesicherten Bereiches und eine verschlüsselte Technologie erforderlich.

Danksagung

Mein Dank gilt Prof. Werner A. Scherbaum für die Unterstützung des Projektes, Dr. Reinhart Willers und Dipl.-Ing. Gerd Kopczynski für die Unterstützung bei der Fragenanalyse, Dr. Matthias Hofer für die Unterstützung bei der Entwicklung und Durchführung der Workshops sowie Dr. Martin Fischer für die Beratung und kritische Durchsicht des Manuskriptes.

Korrespondenzadresse:

• Dr. med. Thomas Rotthoff, Universitätsklinikum Düsseldorf, Klinik für Endokrinologie, Diabetologie und Rheumatologie, Moorenstraße 5, 40225 Düsseldorf, Deutschland, Tel.: 0211/8118-713, Fax.: 0221/8118-772
rotthoff@med.uni-duesseldorf.de

Literatur:

[1] Bundesrat. Approbationsordnung für Ärzte 2002. Bonn: Bundesanzeiger Verlagsgesellschaft; 2002. Zugänglich unter: http://www.bmgs.bund.de/download/gesetze/gesundheitsberufe/AeAppO.pdf.

[2] MCCourbrie P. Improving the fairness of multiple-choice questions: a literature review. Med Teach. 2004;26(8):709-712.

[3] Bloch R, Hofer D, Krebs R, Schlöppi R, Weiss S, Westkämper R. Kompetent prüfen, Bern/Wien: Institut für Aus-, Weiter- und Fortbildung Universität Bern; 1999.

[4] Case SM, Swanson DB. Constructing written Test Questions for the Basic and Clinical Sciences. Philadelphia: National Board of Medical Examiners; 2002.

[5] SAS Institute Inc. SAS/STAT User's Guide, Version 9.1. Cary, NC: SAS Institute Inc.; 2004.

[6] Lienert G, Raatz U. Testaufbau und Testanalyse 6. Auflage. Weinheim: Verlag Beltz, Psychologie Verlags Union; 1998.

[7] Bloom BS, Engelhart MD. Taxonomie von Lernzielen im kognitiven Bereich. 4.Aufl. Weinheim u. Basel: Beltz; 1974.

[8] Krebs R. Anleitung zur Herstellung von MC-Fragen und MC-Prüfungen für die ärztliche Ausbildung. Bern: Institut für Medizinische Lehre IML, Abteilung für Ausbildungs- und Examensforschung AAE; 2004.

[9] Schuwirth L, van der Vleuten C. Merging views on assessment. Med Educ. 2004;38(12):1208-1211.

[10] Downing SM. Threats to the Validity of Locally Developed Multiple-Choice Tests in Medical Education: Construct-Irrelevant Variance and Construct Underrepresentation. Adv Health Sci Educ Theory Pract. 2002;7(3):235-241.

Artikel online frei zugänglich unter http://www.egms.de/en/journals/zma/2006-23/zma000264.shtml
Bitte zitieren als: Rotthoff T, Soboll S. Qualitätsverbesserung von MC Fragen: Ein exemplarischer Weg für eine medizinische Fakultät. GMS Z Med Ausbild. 2006;23(3):Doc45.
gQp020060705 31