

EVALUATING THE USE OF MCQ AS AN ASSESSMENT METHOD IN A MEDICAL SCHOOL FOR ASSESSING MEDICAL STUDENTS IN THE COMPETENCE-BASED CURRICULUM

Mitayani Purwoko, Trisnawati Mundijo

Fakultas Kedokteran Universitas Muhammadiyah Palembang, Palembang - INDONESIA

ABSTRACT

Background: Student's cognitive ability could be assessed using MCQ. The aim of this study was to evaluate the quality of MCQ as an assessment method in Medical Faculty of Muhammadiyah University Palembang.

Method: This study was designed as a cross sectional descriptive observational study. Sample was MCQ assessment in Genetics and Molecular Biology Module for academic year 2013/2014 until 2015/2016 for total 299 questions. Item analysis was done manually.

Results: The item analysis showed that 61.2% questions were recall-type question. This situation showed that the construction of the question was not good and only testing the lower cognitive area. There was 45.2% ideal question with 30-70% difficulty index and 23.1% questions whom distractor efficiency was 100%. Half of the questions (56.2%) should be revised. This revision-needed questions distributed equally into easy, ideal, and hard level of difficulty. Revision-needed questions had lower distractor efficiency mean compared to good questions.

Conclusion: MCQ as an assessment method do not reach maximum target yet because there were many questions that should be revised. Faculty should enhance the development of the lecturer in writing the good MCQs.

Keywords: MCQ, cognitive assessment, item analysis

ABSTRAK

Latar belakang: Ranah pembelajaran kognitif mahasiswa dinilai melalui ujian MCQ. Tujuan MCQ sebagai metode penilaian perlu ditunjang dengan soal yang ideal. Tujuan penelitian ini adalah untuk mengevaluasi MCQ sebagai metode penilaian di Fakultas Kedokteran Universitas Muhammadiyah Palembang.

Metode: Penelitian ini merupakan penelitian observasional deskriptif dengan pendekatan *cross sectional*. Sampel ujian MCQ yang diambil adalah soal MCQ Blok VI yang telah diujikan pada mahasiswa pada tahun akademik 2013/2014 hingga 2015/2016 dengan total 299 soal. *Item analysis* dilakukan secara manual.

Hasil: Hasil *item analysis* menunjukkan 61,2% dari soal MCQ adalah tipe *recall*. Hal ini menunjukkan bahwa konstruksi soal kurang baik dan hanya menguji tingkat kognitif rendah. Sebanyak 45,2% soal memiliki tingkat kesulitan ideal 30-70% dan 23,1% soal memiliki tingkat efisiensi distraktor 100%. Sebagian besar soal perlu direvisi (56,2%) dan soal yang perlu direvisi ini tersebar secara merata pada soal dengan tingkat kesulitan mudah, ideal, maupun sulit. Soal yang perlu direvisi memiliki rata-rata tingkat efisiensi distraktor yang lebih rendah dibandingkan soal kualitas baik.

Kesimpulan: MCQ sebagai salah satu metode penilaian belum mencapai tujuannya secara maksimal karena sebagian besar soal masih perlu direvisi. Oleh karena itu, fakultas perlu meningkatkan pengembangan kemampuan dosen dalam membuat soal MCQ yang berkualitas.

Kata kunci: MCQ, penilaian kognitif, *item analysis*

contact: mitayani.dr@gmail.com

PENDAHULUAN

Kurikulum pendidikan dokter di Fakultas Kedokteran Universitas Muhammadiyah Palembang (FK UMP) menggunakan model Kurikulum Berbasis Kompetensi (KBK) dengan merujuk kepada standar nasional yang ditetapkan oleh Konsil Kedokteran Indonesia (KKI). Ranah pembelajaran kognitif dapat dinilai dengan penilaian tertulis melalui ujian *multiple choice question* (MCQ). MCQ merupakan salah satu instrumen yang digunakan untuk menilai hasil pembelajaran mahasiswa. Susunan soal ujian MCQ dapat menjadi penilaian kognitif mahasiswa, dikarenakan susunan soal MCQ yang baik mengharuskan mahasiswa untuk melakukan analisis dan pemecahan masalah.¹

Tujuan penelitian ini untuk mengevaluasi kelayakan MCQ sebagai metode penilaian di Fakultas Kedokteran Universitas Muhammadiyah Palembang. Hasil penelitian ini diharapkan dapat menjadi masukan bagi Fakultas Kedokteran dalam mengevaluasi capaian ujian MCQ di era kurikulum berbasis kompetensi.

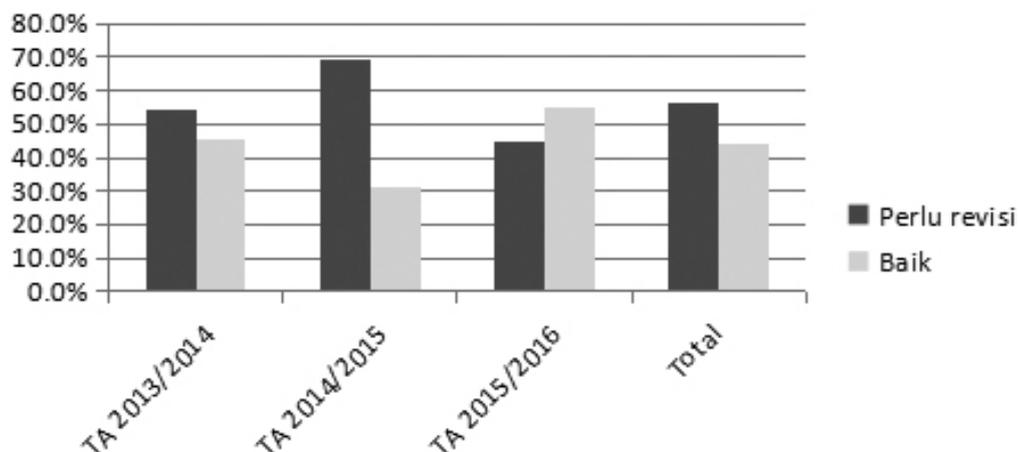
METODE

Penelitian ini merupakan penelitian deskriptif dengan desain potong lintang. Pengambilan data dilakukan di FK Universitas Muhammadiyah Palembang. Sampel yang diambil dalam penelitian ini adalah MCQ Blok VI yang telah diujikan pada mahasiswa pada tahun akademik 2013/2014 hingga

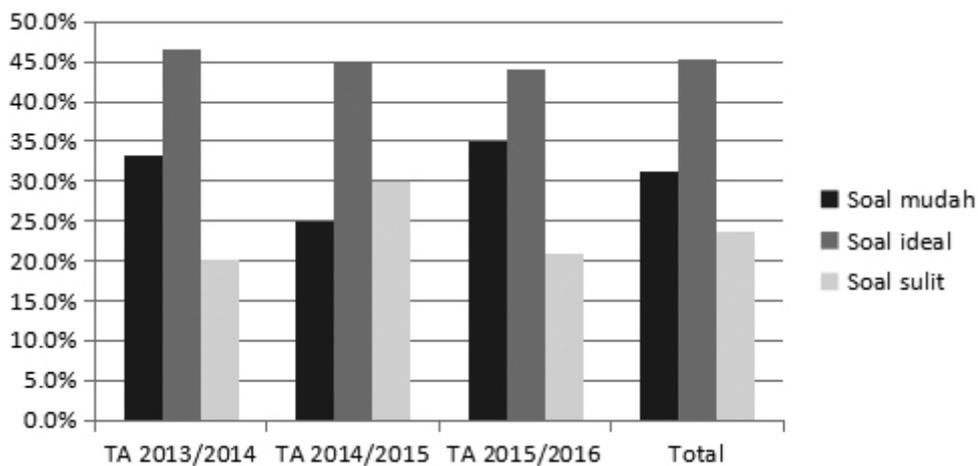
2015/2016 dengan total 299 soal. Variabel yang diteliti adalah tipe soal menurut Bloom, *difficulty index*, *discrimination index*, dan *distractor efficiency*. *Difficulty index* adalah nilai yang didapat dengan menggunakan rumus Jumlah peserta yang menjawab benar: jumlah total peserta x 100%. Soal MCQ dinyatakan sulit bila nilai <30%, ideal bila nilai 30-70%, dan sulit bila nilai >70%. *Discrimination index* adalah nilai yang didapat dengan menggunakan rumus: $2x$ (jumlah kelompok atas yang menjawab benar-jumlah kelompok bawah yang menjawab benar)/jumlah anggota kedua kelompok. Soal MCQ perlu direvisi bila nilai <0,25 dan dinyatakan mampu membedakan antara mahasiswa yang mampu dengan yang tidak mampu bila nilai $\geq 0,25$. *Distractor efficiency* adalah nilai yang diperoleh dengan menghitung persentase masing-masing distraktor dalam satu soal MCQ yang dipilih oleh peserta saat ujian. Distraktor dinyatakan berfungsi bila dipilih oleh $\geq 5\%$ peserta.

HASIL DAN PEMBAHASAN

Setiap soal MCQ dikategorikan menjadi tipe soal menurut Bloom. Pada penelitian ini didapatkan 61,2% soal tipe mengingat, 24,1% soal tipe memahami, dan 14,7% soal tipe analisis, tidak didapatkan soal tipe aplikasi, sintesis, dan evaluasi. Secara keseluruhan, soal MCQ yang digunakan selama tahun akademik 2013/2014 hingga 2015/2016 yang perlu direvisi sebanyak 56,2% (Gambar 1) dengan proporsi tingkat kesulitan yang ideal mendominasi sebesar 45,2% (Gambar 2).



Gambar 1. *Discrimination index* soal MCQ



Gambar 2. Difficulty index soal MCQ

Tingkat efisiensi distraktor soal MCQ Blok VI didominasi oleh tingkat efisiensi distraktor 75%, yang berarti ada 3 dari 4 distraktor yang berfungsi

pada satu soal, sebesar 27,4%. Namun ada 9,4% soal yang semua distraktornya tidak dipilih oleh peserta (DE 0%) (Gambar 3).



Gambar 3. Tingkat efisiensi distraktor soal MCQ (Keterangan DE = *distractor efficiency*)

Analisis kelayakan soal MCQ Blok VI FK UM Palembang dapat dinyatakan bahwa sebagian besar soal perlu direvisi (56,2%) dan soal yang perlu direvisi ini tersebar secara merata pada soal dengan

tingkat kesulitan mudah, ideal, maupun sulit. Soal yang perlu direvisi memiliki rata-rata tingkat efisiensi distraktor yang lebih rendah dibandingkan soal baik (Tabel 1).

Tabel 1. Matriks analisis kelayakan soal MCQ Blok Genetika dan Biologi Molekuler

DF	Soal perlu direvisi	Rata-rata DE	Soal baik	Rata-rata DE	Total soal
Mudah	55 (59,1%)	22%	38 (40,9%)	41%	93
Ideal	55 (40,7%)	70%	80 (59,3%)	73%	135
Sulit	58 (81,7%)	74%	13 (18,3%)	85%	71
Total soal	168 (56,2%)	55,3%	131 (43,8%)	65%	299

Keterangan: DE = *distractor efficiency*, DF = *difficulty index*

Soal MCQ blok Genetika dan Biologi Molekuler sebagian besar termasuk tipe soal *recall* (mengingat). Amin dan Eng¹ menyatakan bahwa soal MCQ yang konstruksinya tidak baik cenderung menguji *recall* pengetahuan. Ada kecenderungan bahwa dosen mengkonstruksi butir soal dengan hanya menguji *recall* pengetahuan, yang hanya menilai tingkat kognitif rendah. Hal ini terjadi karena soal MCQ yang menguji tingkat kognitif yang lebih tinggi lebih sulit untuk dibuat. Dosen cenderung membuat soal MCQ yang mudah dikembangkan tapi gagal menilai tingkat kognitif yang lebih tinggi. Hasil penelitian ini menunjukkan kesamaan dengan hasil penelitian Kim et al² yang melakukan penelitian mengenai klasifikasi Bloom dalam soal MCQ di *University of Pacific, California*. Mereka menemukan 283 dari 638 (44,4%) soal MCQ memiliki tipe *recall*. Kim et al² menyimpulkan bahwa soal MCQ dengan tipe *application*, *synthesis*, dan *evaluate* memiliki *discrimination index* secara signifikan lebih tinggi daripada soal tipe *recall* dan *understanding*. Hasil penelitian Palmer dan Devitt³ dan Baig et al⁴, juga menemukan lebih dari 50% soal MCQ yang diteliti ternyata hanya fokus pada soal tipe *recall*. Zafar-Khan dan Aljarallah⁵ menemukan hal yang berbeda yaitu 60% dari soal MCQ merupakan soal tipe analisis (level 3 dari klasifikasi Bloom) sementara soal tipe mengingat (level 1) hanya ditemukan sebesar 28%.

Dalam soal MCQ blok Genetika dan Biologi Molekuler ditemukan lebih dari separuh soal (56,2%) perlu direvisi berdasarkan *discrimination index*. Hasil penelitian ini sesuai dengan hasil penelitian Gajjar et al⁶ di fakultas kedokteran di Ahmedabad, Gujarat, India. Gajjar et al⁶ menemukan bahwa 52% soal memiliki nilai indeks diskriminasi yang tidak baik. Hal yang menyebabkan banyaknya soal yang tidak bisa membedakan antara mahasiswa yang benar-benar paham dengan mahasiswa

yang tidak paham kemungkinan disebabkan oleh kurangnya kemampuan dosen dalam menyusun soal MCQ karena tipe soal sebagian besar hanya bersifat mengingat (*recall*). Banyaknya soal MCQ Blok VI di FK UMP dengan nilai *discrimination index* yang rendah kemungkinan disebabkan bentuk soal yang lebih banyak berupa soal *recall*, karena menurut hasil penelitian Kim et al² tipe soal yang mampu membedakan antara peserta yang benar-benar mampu dengan mahasiswa yang tidak mampu adalah tipe soal *application*, *synthesis*, dan *evaluate*.

Sebaran soal MCQ untuk semua tahun akademik didominasi oleh soal dengan tingkat kesulitan ideal sebesar 45,2%. Hasil penelitian ini sama dengan hasil penelitian Sim dan Rasiah⁷ di Fakultas Kedokteran Universitas Malaya, Malaysia, dimana mereka menemukan 53% soal dari 50 pertanyaan MCQ dengan 5 pilihan jawaban adalah soal ideal berdasarkan *difficulty index*. Abdulghani et al⁸ menyatakan bahwa soal MCQ di King Saud University, Riyadh, lebih banyak pada soal dengan tingkat kesulitan ideal baik pada *pre-test* (62,5%) dan *post-test* (76,3%).

Dari 299 soal hanya ditemukan 28 soal (9,4%) yang memiliki tingkat efisiensi *distractor* 0%. Hasil penelitian ini lebih rendah dari hasil penelitian Gajjar et al⁶ dan Kilgour dan Tayyaba⁹ yang menemukan 30% dan 14,2% soal MCQ yang diteliti tidak memiliki satupun *distractor* yang berfungsi. Hal ini menunjukkan bahwa *distractor* yang dibuat oleh dosen untuk soal MCQ blok Genetika dan Biologi Molekuler sudah cukup baik karena hanya sedikit yang *distractor*-nya sama sekali tidak dipilih oleh peserta ujian. Gajjar et al⁶ menyimpulkan bahwa tingkat efisiensi *distractor* lebih tinggi pada soal dengan tingkat kesulitan tinggi daripada tingkat kesulitan rendah. Hal ini juga sesuai dengan

penelitian ini dimana soal sulit memiliki rata-rata tingkat efisiensi *distractor* 74-92%. Angka ini jauh lebih tinggi daripada rata-rata tingkat efisiensi *distractor* pada soal mudah. Hal ini terjadi karena mahasiswa sulit menentukan jawaban yang benar pada soal MCQ dengan *difficulty index* rendah sehingga hampir semua *distractor* dipilih oleh lebih dari 5% peserta ujian.

Sim dan Rasiah⁷ menyimpulkan bahwa soal MCQ dengan nilai diskriminasi indeks yang baik cenderung menjadi soal dengan tingkat kesulitan yang cukup tinggi. Hal ini tidak sesuai dengan hasil penelitian ini dimana soal dengan nilai diskriminasi indeks yang baik lebih banyak menjadi soal dengan tingkat kesulitan ideal. Situasi ini sudah sangat baik dimana soal tidak terlalu mudah ataupun terlalu sulit namun dapat membedakan antara mahasiswa yang benar-benar mampu dengan mahasiswa yang tidak mampu. Sim dan Rasiah⁷ juga menyimpulkan bahwa soal yang sangat sulit cenderung memiliki nilai diskriminasi indeks yang negatif, yang berarti bahwa soal sulit cenderung perlu direvisi. Hal ini tidak sesuai dengan hasil penelitian ini dimana dari 168 soal MCQ yang perlu direvisi, sebarannya merata pada soal tipe mudah, ideal, dan sedang.

Untuk meningkatkan kualitas soal MCQ, dosen sebaiknya menilai sendiri soal MCQ buatan mereka. Namun, penilaian internal terkadang menunjukkan bahwa semua soal sudah baik dan layak diujikan. Oleh karena itu perlu adanya penilaian silang oleh dosen lain atau Unit Pendidikan Kedokteran.¹⁰

KESIMPULAN

MCQ sebagai salah satu metode penilaian belum mencapai tujuannya secara maksimal karena sebagian besar soal masih perlu direvisi. Oleh karena itu, fakultas perlu meningkatkan pengembangan kemampuan dosen dalam membuat soal MCQ yang berkualitas.

UCAPAN TERIMA KASIH

Ucapan terima kasih kami ucapkan kepada Dekan dan Ketua Program Studi Pendidikan Kedokteran Fakultas Kedokteran Universitas Muhammadiyah Palembang atas pemberian izin pengambilan data, staf akademik atas bantuan dalam pengambilan

data, serta Lembaga Penelitian dan Pengabdian Masyarakat Universitas Muhammadiyah Palembang atas pemberian hibah internal penelitian.

DAFTAR PUSTAKA

1. Amin Z, Eng KH. Basics in medical education. Singapore: World Scientific Publishing Co. Pte. Ltd. 2003.
2. Kim MK, Patel RA, Uchizono JA, Beck L. Incorporation of Bloom's taxonomy into multiple-choice examination question for a pharmacotherapeutics course. *American Journal of Pharmaceutical Education*. 2012; 76(6): 1-8.
3. Palmer EJ dan Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. *BMC Medical Education*. 2007;7(49):1-7.
4. Baig M, Ali SK, Ali S, Huda N. Evaluation of multiple choice and short essay question items in basic medical sciences. *Pak J Med Sci*. 2014;30(1):3-6.
5. Zafar-Khan M, Aljarallah BM. Evaluation of Modified Essay Question (MEQ) and Multiple Choice Question (MCQ) as a tool for assessing the cognitive skills of undergraduate medical students. *IJHS*. 2011;5(1):39-43.
6. Gajjar S, Sharma R, Kumar P, dan Rana M. Item and test analysis to identify quality Multiple Choice Questions (MCQs) from an assessment of Medical students of ahmedabad, Gujarat. *Indian J Community Med*. 2014 Jan-Mar; 39(1):17-20. doi: 10.4103/0970-0218.126347
7. Sim SM dan Rasiah RI. Relationship Between Item Difficulty and Discrimination Indices in True/False-Type Multiple Choice Questions of a Para-clinical Multidisciplinary Paper. *Ann Acad Med Singapore*. 2006;35:67-71.
8. Abdulghani HM, Ahmad F, Irshad M, Khalil MS, Al-Shaikh GK, Syed A, et al. Faculty development programs improve the quality of Multiple Choice Questions items' writing. *Scientific Reports*. 2015;5:9556. doi:10.1038/srep09556
9. Kilgour JM, Tayyaba S. An investigation into the optimal number of distractors in single-best answer exams. *Adv Health Sci Educ Theory Pract*. 2016;21;571-85.
10. Collins J. Writing multiple-choice questions for continuing medical education activities and self-assessment modules. *RadioGraphics*. 2006;26(2):543-51. doi: http://dx.doi.org/10.1148/rg.262055145