

PHÂN TÍCH CÂU HỎI THI TRẮC NGHIỆM KHÁCH QUAN CỦA MỘT SỐ MÔN THI LÝ THUYẾT DÀNH CHO SINH VIÊN TẠI TRƯỜNG ĐẠI HỌC Y HÀ NỘI

NGUYỄN THỊ NGA, NGUYỄN HỮU TÚ,
KIM BẢO GIANG, TRẦN THỊ THÙY DUNG, TẠ THỊ MIẾN
Trường Đại học Y Hà Nội

TÓM TẮT

Mục tiêu: Mô tả kết quả và phân tích đề thi trắc nghiệm lý thuyết một số môn thi tại trường ĐHYHN năm học 2010-2011. Đối tượng và phương pháp nghiên cứu: Tất cả các câu hỏi thi của 9 môn thi trong 3822 lượt bài thi được chấm, tổng hợp và phân tích theo quy trình được trung tâm Khảo thí và Đảm bảo chất lượng áp dụng. Kết quả: Tỷ lệ bài thi đạt điểm khá giỏi từ 7 trở lên cao (67,4%), chỉ có 6,6% bài thi có điểm dưới 5. Tỷ lệ câu hỏi quá dễ chiếm 58,0%. Tỷ lệ câu hỏi chấp nhận được chiếm 38,7%. Tỷ lệ câu hỏi khó chiếm 3,3%. Tỷ lệ câu hỏi thi có độ phân biệt chưa tốt chiếm 49,3%. Tỷ lệ câu hỏi thi có phương án nhiễu < 5% chiếm 34,2%. Kết luận: Qua kết quả này, cho thấy đề thi trắc nghiệm lý thuyết còn quá dễ, các Bộ môn nên có kế hoạch rà soát và chỉnh sửa để nâng hàng câu hỏi ngày càng hoàn thiện. Phân tích đề thi cần tiếp tục được thực hiện nhằm cung cấp cho việc theo dõi và cải thiện chất lượng đào tạo của nhà trường.

Từ khóa: Trắc nghiệm khách quan (TNKQ), độ khó, độ phân biệt, chất lượng câu hỏi thi.

SUMMARY

Objectives: Describe the results and perform item analysis of written tests used for summative assessment of some training modules at the Hanoi Medical University in 2010-2011. Study subjects and methods: All items used in 9 summative assessments for 3822 students were examined, synthesised and analysed using technical procedures by the Center for Educational Testing and Quality Assurance. Results: Proportion of written tests had good and excellent marks (from 7 and above) was high (67.4%), only 6.6% of the written tests had under average marks. The proportion of too easy items was 58.0%; acceptable items accounted for 38.7% and difficult items accounted for 3.3%. Proportion of items that had low discriminant levels accounted for 49.3% and proportion of items that had a distractor selected by less than 5% of students was 34.2%. Conclusion: The items used in written tests were still too easy, academic departments should have plan to review and revise these items to

improve quality of the item bank over time. Item analysis should continue to be done to provide information for monitoring and improving quality of training at the university.

Keywords: Objective structured tests, difficulty, discriminant, quality of written tests.

ĐẶT VẤN ĐỀ

Lượng giá sinh viên là một hoạt động quan trọng trong một chương trình giáo dục. Yêu cầu đối với các đề thi là phải đảm bảo tính giá trị, sự bao phủ các nội dung học tập, đo lường mức độ hoàn thành mục tiêu học tập. Đề thi cũng phải đảm bảo tính tin cậy, tính chính xác và công bằng trong đánh giá sinh viên. Đề thi có chất lượng là đề thi không quá khó, không quá dễ và đảm bảo phân biệt được sinh viên thuộc các nhóm học lực khác nhau [1]. Sử dụng những câu hỏi thi có chất lượng không chỉ giúp cơ sở đào tạo đánh giá đúng mức độ hoàn thành mục tiêu và yêu cầu học tập của sinh viên, đo lường chất lượng giáo dục của nhà trường mà còn có vai trò thúc đẩy, khuyến khích sinh viên tích cực học tập không ngừng. Tuy nhiên, nhiều khi đánh giá độ khó dễ cũng như khả năng phân loại thí sinh trước đây chủ yếu là do chủ quan của người ra đề. Trên thực tế, nhiều khi một câu hỏi người ra đề cho là dễ nhưng lại ít thí sinh làm được và ngược lại, có câu thi người ra đề đánh giá là khó thì đa số thí sinh lại làm được. Từ khi trường Đại học Y Hà Nội thành lập trung tâm Khảo thí và Đảm bảo chất lượng Giáo dục, nhiều môn học đã được tổ chức thi theo hình thức trắc nghiệm khách quan, quét bài thi và chấm thi trên máy [2]. Việc tổ chức thi và chấm thi này giúp cho việc phân tích câu hỏi thi được thực hiện dễ dàng hơn. Nhận thức rõ được tầm quan trọng của việc xem xét chất lượng câu hỏi thi nhằm không ngừng nâng cao chất lượng đào tạo của nhà trường, Trung tâm Khảo thí và Đảm bảo chất lượng Giáo dục đã tiến hành phân tích câu hỏi thi cho những môn thi trắc nghiệm được tổ chức tại

Trung tâm. Nghiên cứu nhằm *Mục tiêu*: Mô tả kết quả thi trắc nghiệm lý thuyết một số môn tại Trung tâm Khảo thí & ĐBCLGD - trường ĐHYHN năm học 2010-2011; Phân tích chất lượng đề thi trắc nghiệm của các Bộ môn đã tổ chức thi tại Trung tâm Khảo thí & ĐBCLGD năm học 2010-2011.

ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Đối tượng nghiên cứu:

Câu hỏi và kết quả thi trắc nghiệm của 9 môn thi có số lượng 200 sinh viên trên mỗi đợt thi được phân tích.

Thiết kế nghiên cứu:

Mô tả cắt ngang

Cỡ mẫu:

Tất cả 3822 lượt bài thi và điểm kết quả bài thi của 9 môn thi được đưa vào phân tích.

Quá trình làm đề thi và phân tích kết quả thi:

Đối với mỗi môn thi, từ một đề chuẩn, phần mềm xây dựng đề thi sẽ tự hoán vị thứ tự câu hỏi và thứ tự phương án trả lời trong câu hỏi để xây dựng thành các đề hoán vị ứng với các mã đề khác nhau. Sau mỗi môn thi, Trung tâm Khảo thí & Đảm bảo chất lượng giáo dục sẽ tổ chức chấm và phân tích đề thi dựa trên phần mềm chuyên dụng theo các bước: (1) Quét, chuyển đổi dữ liệu từ dạng ảnh sang dạng text và xử lý bài thi để nhận dạng các câu trả lời của sinh viên, kiểm tra lỗi dữ liệu của các lô bài thi. Đối với các bài thi nhận dạng bị lỗi sẽ được kiểm dò lại một cách chính xác; (2) Tổ hợp các mã đề hoán vị thành một đề chuẩn, mọi thông tin phân tích sẽ dựa trên đề chuẩn; (3) Chấm thi và phân tích: sau khi có điểm thi sẽ tính toán tỷ lệ bài thi đạt các mức điểm khác nhau, mô tả đề thi gồm tổng số câu hỏi thi, số câu dạng Đúng/Sai, số câu dạng nhiều lựa chọn (MCQ); phân tích độ khó, độ phân biệt của câu hỏi thi, chất lượng của các phương án nhiễu. Nghiên cứu này cũng sử dụng mô hình RASCH để phân tích sự phù hợp của các câu hỏi thi với đo lường năng lực của sinh viên.

Một số khái niệm:

- Độ khó của câu hỏi thi (k) là tỷ lệ % sinh viên làm đúng câu hỏi trên tổng số sinh viên tham gia trả lời câu hỏi. Độ khó chỉ ra mức độ khó của câu hỏi, k được đo từ 0.00 (khó nhất) đến 1.00 (dễ nhất). Thông thường câu hỏi được cho là dễ khi có $k > 0.7$; chấp nhận được khi k dao động trong khoảng 0.3-0.7; và khó khi $k < 0.3$ [1].

- Độ phân biệt của câu hỏi thi (p) chỉ ra mức độ phân biệt được thí sinh giỏi và thí sinh kém, p được đo từ -1.00 đến 1.00 (-1.00 là độ phân biệt kém nhất, 1.00 là độ phân biệt tốt nhất). Độ phân biệt $p > 0.20$ là chấp nhận được, $p = 1$ là phân biệt hoàn hảo, $p < 0.2$ là phân biệt kém [1].

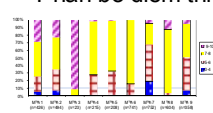
- Chất lượng của các phương án nhiễu: phương án nhiễu là phương án sai trong câu hỏi nhiều lựa chọn. Phương án nhiễu được coi là quá dễ khi chỉ có <5% sinh viên chọn; phương án nhiễu có >50% sinh viên chọn được xem là không phù hợp. Ngoài ra còn xem xét tỷ lệ % số câu có đáp án đúng nhưng ít sinh viên chọn [1].

- Phân tích RASCH: sử dụng mô hình RASCH trong phân tích câu hỏi thi giúp xác định được mức độ phù hợp của từng câu hỏi thi trong đo lường năng lực của

sinh viên, chỉ rõ câu hỏi nào phù hợp để đo lường năng lực của sinh viên và những câu hỏi không thực sự đo lường năng lực quan tâm của sinh viên. Những câu hỏi đó có thể nằm trên hoặc dưới ngưỡng năng lực của sinh viên cần phải xem lại, có thể câu hỏi đó quá khó đối với năng lực của sinh viên hoặc câu hỏi đó làm sinh viên hiểu nhầm. Một đề thi có quá nhiều câu hỏi nằm ngoài năng lực của sinh viên thì đề thi đó cần xem lại. Ưu điểm của mô hình này là kết quả phân tích độc lập với đặc điểm của nhóm sinh viên thi nên kết quả đánh giá câu hỏi thi không thay đổi nếu áp dụng trên những nhóm sinh viên khác nhau [3,4].

KẾT QUẢ

Phân bố điểm thi của các môn thi



Hình 1: Biểu đồ phân bố điểm của từng môn thi

Trong tổng số 3822 lượt sinh viên tham gia 9 môn thi tỷ lệ sinh viên đạt điểm khá giỏi chiếm tỷ lệ cao (67.4%), chỉ có 26.0% đạt điểm trung bình (5-6 điểm) và 6.6% sinh viên đạt điểm dưới 5. Đặc biệt có môn thi số 3 có tới 91.3% sinh viên đạt điểm 9-10 và không có sinh viên nào đạt điểm dưới 7.

Chất lượng câu hỏi thi

Bảng 1. Phân loại các câu hỏi thi của từng môn thi theo mức độ khó

Môn	Tổng số câu	Độ khó		
		Khó n (%)	Chấp nhận được n(%)	Rất dễ n(%)
Môn 1	50	1(2,0)	17(34,0)	32(64,0)
Môn 2	50	1(2,0)	20(40,0)	29(58,0)
Môn 3	50	2(4,0)	1(2,0)	47(94,0)
Môn 4	100	1(1,0)	47(47,0)	52(52,0)
Môn 5	50	4(8,0)	18(36,0)	28(56,0)
Môn 6	100	0(0,0)	42(42,0)	58(58,0)
Môn 7	65	6(9,2)	41(63,1)	18(27,7)
Môn 8	100	2(2,0)	25(25,0)	73(73,0)
Môn 9	50	3(6,0)	27(54,0)	20(40,0)
Tổng	615	20(3,3)	238(38,7)	357(58,0)

Kết quả cho thấy trong tổng số 615 câu hỏi thi của 9 môn thi, tỷ lệ câu hỏi quá dễ chiếm 58.0%; tỷ lệ câu hỏi chấp nhận được chiếm 38,7% và câu hỏi khó chỉ chiếm 3,3%. Đặc biệt môn thi số 3, tỷ lệ câu dễ chiếm tới 94%, tương ứng với tỷ lệ sinh viên đạt điểm giỏi ở biểu đồ 1. Môn có số câu hỏi khó với tỷ lệ cao nhất cũng chỉ chiếm 9,2%. Đặc biệt môn 6 không có câu hỏi nào khó.

Độ phân biệt

Bảng 2. Phân loại câu hỏi thi của từng môn thi theo mức độ phân biệt

Môn	Tổng số câu	Chấp nhận được n (%)	Chưa tốt n (%)
Môn 1	50	44 (88.0)	6 (12.0)
Môn 2	50	45 (90.0)	5 (10.0)
Môn 3	50	7 (14.0)	43 (86.0)
Môn 4	100	49 (49.0)	51 (51.0)
Môn 5	50	24 (48.0)	26 (52.0)
Môn 6	100	19 (19.0)	81 (81.0)
Môn 7	65	55 (84.6)	10 (15.4)
Môn 8	100	29 (29.0)	71 (71.0)
Môn 9	50	40 (80.0)	10 (20.0)
Tổng	615	312 (50.7)	303 (49.3)

Nhìn chung gần một nửa số câu hỏi thi (49.3%) có độ phân biệt chưa tốt. Đặc biệt, môn thi số 3, tỷ lệ câu hỏi có độ phân biệt chưa tốt chiếm 86%. Môn thi số 1 là môn có độ phân biệt khá tốt vì số câu hỏi có độ phân biệt chưa tốt chỉ chiếm 12%.

Chất lượng của các phương án nhiễu

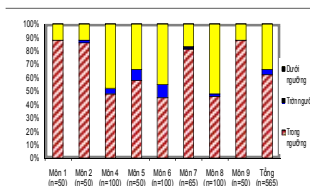
Bảng 3. Chất lượng các phương án nhiễu của câu hỏi thi

Môn	Tổng số câu	Số p/án nhiễu được chọn < 5% (n, %)	Số p/án nhiễu được chọn ≥ 50% (n, %)	Số câu Đ/A đúng ít được chọn (n, %)
Môn 1	50	72 (48,0)	0 (0,0)	0 (0,0)
Môn 2	50	59 (39,3)	0 (0,0)	0 (0,0)
Môn 3	50	140 (93,3)	0 (0,0)	1 (0,7)
Môn 4	100	32 (10,7)	18 (6,0)	30 (10,0)
Môn 5	50	81 (54,0)	5 (3,3)	2 (1,3)
Môn 6	100	31 (10,3)	21 (7,0)	39 (13,0)
Môn 7	65	29 (14,9)	3 (1,5)	3 (1,5)
Môn 8	100	117 (53,2)	2 (0,9)	1 (0,5)
Môn 9	50	43 (28,7)	2 (1,3)	2 (1,3)
Tổng	615	604 (34,2)	51 (2,9)	78 (12,7)

Kết quả cho thấy, mặc dù số phương án nhiễu có trên 50% sinh viên chọn chỉ chiếm 2,9%, có tới 1/3 số phương án nhiễu (34,2%) quá dễ và cần phải điều chỉnh để nâng cao chất lượng đề thi. Đặc biệt môn thi số 3 có tới 93,3% số phương án nhiễu quá dễ.

Sự phân phối của các câu hỏi thi qua phân tích Rasch

Do môn số 3 có nhiều câu hỏi quá dễ nên không được đưa vào phân tích.



Hình 2. Biểu đồ phân bố câu hỏi theo năng lực của sinh viên

Kết quả phân tích 8 môn còn lại cho thấy, trong tổng số 565 câu hỏi thi, tỷ lệ câu hỏi nằm trong ngưỡng và phản ánh được năng lực của sinh viên là 61,9%. Tỷ lệ câu hỏi nằm ngoài ngưỡng năng lực của sinh viên là 38,1%, trong đó có 34,2% nằm dưới ngưỡng và 3,9% số câu hỏi nằm trên ngưỡng năng lực của sinh viên. Môn thi số 1, 12, 7, 9 có chất lượng tốt hơn với trên 80% câu hỏi nằm trong ngưỡng năng lực của sinh viên.

BÀN LUẬN

Nghiên cứu của chúng tôi cho thấy phân bố điểm thi là phân bố lệch về phía điểm khá giỏi, đồng thời số câu hỏi rất dễ (58%) và độ phân biệt chưa tốt (49,3%) chiếm tỷ lệ khá cao. Đặc biệt có một số đề thi còn có số câu hỏi quá dễ chiếm trên 90%, phương án nhiễu có tỷ lệ chọn <5% chiếm trên 50%. Mặc dù cấu trúc các câu hỏi khó dễ của một đề thi cần sắp xếp tùy thuộc vào mức độ kiến thức, nhận thức và thực hành theo yêu cầu của môn học, theo khuyến cáo của các chuyên gia lượng giá, một đề thi hợp lý thường nên có 33% câu dễ, 33% câu khó và 34% câu ở mức trung bình. Ngoài ra, các câu hỏi nhiễu lựa chọn cũng không nên có phương án nhiễu chỉ có <5% hoặc có >50% sinh viên chọn [1]. Như vậy, theo khuyến cáo này, nhận định chung là các đề thi tại trường còn dễ và cần phải được chỉnh sửa để nâng cao chất lượng của đề thi. Từ học kỳ II năm học 2010 - 2011, đối với mỗi môn thi trắc nghiệm được tổ chức tại Trung tâm Khảo thí & ĐBCLGD, cùng với kết quả thi, Trung tâm đã gửi kèm kết quả phân tích từng câu hỏi thi cho Bộ môn để các Bộ môn tham khảo và có kế hoạch sửa chữa, bổ sung hoặc loại bỏ nếu không còn phù hợp. Nếu những nỗ lực này tiếp tục được thực hiện và được các bộ môn quan tâm và kiên quyết thực hiện, chất lượng đề thi sẽ từng bước được cải thiện.

KẾT LUẬN

Kết quả phân tích đề thi của một số môn thi trắc nghiệm lý thuyết của một số Bộ môn cho thấy hơn một nửa số câu hỏi thi còn đang thuộc mức quá dễ, tỷ lệ phương án nhiễu chưa phù hợp còn cao. Vì vậy các bộ môn cần tiếp tục sửa chữa, bổ sung để nâng cao chất lượng đề thi. Hoạt động phân tích đề thi cần tiếp tục được thực hiện để cung cấp thông tin cho các bộ môn và nhà trường trong theo dõi và đảm bảo chất lượng đào tạo.

TÀI LIỆU THAM KHẢO

1. Đơn vị Đào tạo và tư vấn Giáo dục Y học, trường Đại học Y Hà Nội (2010). *Tài liệu tập huấn Lượng giá sinh viên*. Hà Nội, 2010.
2. Trung tâm Khảo thí và Đảm bảo chất lượng Giáo dục, trường Đại học Y Hà Nội. *Báo cáo hoạt động Trung tâm năm học 2010-2011*. Hà Nội 2011.
3. Chong Ho Yu. *A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling*. December 2010: <http://www.creative-wisdom.com>
4. K.-Y. Chang, M.-Y. Tsou, K.-H. Chan, S.-H. Chang, J. J. Tai and H.-H. Chen (2010). *Item analysis for the written test of Taiwanese board certification examination in anaesthesiology using the Rasch model*. *British Journal of Anaesthesia* 104 (6): 717—22.