

Summary

STUDY ON CHARACTERISTICS OF MULTIPLE CHOICE QUESTIONS BY RESPONSE QUESTION THEORY

Modern test theory is now widely applied to assess the correlation between the ability of candidates and the difficulty of each question. **Objectives:** (1) Defining the form and marks of the test. (2) Analyzing the correlation between the questions and the candidates based on response question theory model. 3. Analyzing the harmony between the candidates and the model. **Method:** 402 tests of Hanoi medical 2th year students. **Result:** Most MCQs fit Rasch model. Most questions have high distinction level. Most abilities of candidates are suitable for question difficulty. Some candidates have unusual features. Measurement capabilities of candidates with difficulty level, distinction level of each question will help to review, edit, standardize question banks better and better. **Conclusion:** Rasch model definitely fits to analyze the distribution of question difficulty, evaluating the specific parameters of questions, topic test. Model also detects the unusual response in correlation of candidate ability and question difficulty.

Keywords: response question theory, standardize question banks

ỨNG DỤNG MÔ HÌNH HAI THAM SỐ TRONG NGHIÊN CỨU MỘT SỐ ĐẶC TÍNH TRẮC NGHIỆM TÂM LÝ ĐÁNH GIÁ TRẠNG THÁI TÂM TRÍ TỐI THIỂU (MMSE) Ở NGƯỜI BÌNH THƯỜNG TUỔI 50 - 59

Trần Thúy Liễu¹, Bùi Mỹ Hạnh²

¹Trường Cao đẳng Y tế Hà Nội; ²Trường Đại học Y Hà Nội

Trắc nghiệm đánh giá trạng thái tâm trí tối thiểu (Mini Mental State Examination - MMSE) được chuyển thể sang nhiều phiên bản với những ngôn ngữ khác nhau trong đó có Việt Nam và được sử dụng như một công cụ sàng lọc chức năng nhận thức. **Mục tiêu:** ứng dụng mô hình hai tham số để tìm ra một số đặc tính các câu hỏi và toàn bộ bộ trắc nghiệm MMSE trên người bình thường tuổi 50 - 59. **Đối tượng và phương pháp nghiên cứu:** 120 người tuổi từ 50 - 59, gồm 60 nam và 60 nữ được thực hiện trắc nghiệm MMSE. Các điểm số thu được được đưa vào chương trình BILOG để xử lý và phân tích. **Kết quả:** sau phase 1: 9/30 câu bị loại bỏ khỏi chương trình xử lý do không phù hợp (hai câu có độ phân biệt < 0, bảy câu có độ phân biệt bằng 0). Sau phase 2: trong hai mươi một câu được phân tích mười câu có độ khó không phù hợp, mười một câu có độ khó phù hợp. Điểm thô tối thiểu của 21 câu được xác định từ hàm thông tin câu hỏi là 18. **Bàn luận:** trắc nghiệm MMSE có nhiều câu tỏ ra quá dễ so với năng lực của đối tượng, chỉ có một số ít câu cần thiết thể hiện năng lực nhận thức của đối tượng nghiên cứu. **Kết luận** sử dụng mô hình hai tham số cho thấy có thể đơn giản hóa bộ trắc nghiệm MMSE, việc loại bỏ một số câu không phù hợp không những làm công cụ này ngắn lại, dễ tiến hành, mà còn duy trì được tính giá trị và độ tin cậy của trắc nghiệm để có thể sử dụng như một test sàng lọc ở cộng đồng.

Từ khóa: trắc nghiệm tâm lý, trạng thái tâm trí tối thiểu

I. ĐẶT VẤN ĐỀ

Trắc nghiệm MMSE từ lâu được sử dụng như một công cụ để chẩn đoán, sàng lọc đối với các thể bệnh như sa sút nhận thức nhẹ, Alzheimer [3]. Quy trình ứng dụng MMSE được tiến hành theo một nguyên tắc chung trong đó các nhà nghiên cứu, lâm sàng thường đặc biệt lưu ý đến sự phù hợp hay không phù hợp của các câu hỏi trắc nghiệm với từng nhóm đối tượng bệnh nhân và người bình thường. Bên cạnh những phương pháp phân tích thống kê thông thường, nhiều nhóm nghiên cứu đã tích hợp phương pháp đánh giá các câu hỏi bằng các chương trình máy tính dựa trên lý thuyết đáp ứng câu hỏi của nhà toán học Rasch. Kết quả thu được sẽ cung cấp những thông tin có tính giá trị cho việc xác định độ khó, độ phân biệt, tính phù hợp của từng câu hỏi cũng như của cả bộ câu hỏi trắc nghiệm. Thực tế hiện nay có nhiều trắc nghiệm thần kinh tâm lý trong đó có bộ trắc nghiệm MMSE đang được sử dụng tại Việt Nam mà đa số là sử dụng các câu trắc nghiệm nguyên bản của nước ngoài dịch sang tiếng Việt. Rất ít công trình nghiên cứu về tính thích ứng của các câu hỏi này theo hướng công bố tính giá trị, độ tin cậy mà kết quả mới chỉ đưa ra các thông số đánh giá mang tính thống kê về điểm số trung bình, độ lệch, tìm hệ số tương quan tuyến tính hay không với các yếu tố như tuổi, giới, trình độ học vấn... Chưa thấy có những phân tích về mối tương quan giữa năng lực của từng cá thể, nhóm cá thể với tính chất của từng câu hỏi trắc nghiệm như độ khó, độ phân biệt,...

Xuất phát từ những lý do trên **mục tiêu** của nghiên cứu này là:

Ứng dụng mô hình hai tham số nghiên cứu một số đặc tính tâm lý của trắc nghiệm đánh giá trạng thái tâm trí tối thiểu trên người bình thường tuổi 50 - 59.

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

1. Đối tượng nghiên cứu

Phân tích số liệu thứ cấp lấy từ kết quả thực hiện trắc nghiệm MMSE cho 120 người tuổi từ 50 - 59 gồm 60 nam, 60 nữ chia làm hai nhóm có trình độ học vấn đại học và dưới đại học[1].

2. Phương pháp nghiên cứu

Thiết kế nghiên cứu: mô tả cắt ngang.

Cách tiến hành: sử dụng các bộ câu hỏi MMSE đã được dịch sang tiếng Việt và đang sử dụng tại Viện Lão khoa Quốc gia. Trắc nghiệm được tiến hành riêng với từng đối tượng trong điều kiện phòng riêng yên tĩnh, ánh sáng rõ ràng. Để tạo tâm lý hợp tác của đối tượng, chúng tôi giải thích rõ mục đích đánh giá và sơ bộ cách tiến hành trắc nghiệm. Các đối tượng sẽ được thực hiện trắc nghiệm sau khi đã kí vào bản thỏa thuận đồng ý tham gia nghiên cứu. Đánh giá chức năng nhận thức tổng quát bằng trắc nghiệm đánh giá trạng thái tâm trí tối thiểu (MMSE) của Folstein. Trắc nghiệm này gồm 30 câu hỏi nhị phân bao gồm: 5 câu định hướng về thời gian; 5 câu định hướng về không gian; 3 câu Ghi nhớ tức thì 3 từ; 5 câu cho khả năng chú ý và tính toán. 3 câu nhớ lại, 8 câu về ngôn ngữ (gọi tên đồ vật, nhắc lại câu, làm theo mệnh lệnh viết, làm theo mệnh lệnh 3 giai đoạn, viết câu) và 1 câu vẽ hình theo mẫu. Tổng điểm tối đa của trắc nghiệm MMSE là 30.

Xử lý số liệu: số liệu được làm sạch lỗi và mã hoá nhập vào máy tính. Sử dụng chương trình máy tính BILOG theo mô hình hai tham số để đánh giá các câu hỏi và cả bộ trắc nghiệm.

Bước 1: chuẩn hoá các câu hỏi

- Câu hỏi phù hợp là câu có độ phân biệt dương (> 0): gồm hai loại:

Câu hỏi điều kiện: là những câu hỏi bắt buộc

đối tượng phải trả lời được, đây là những câu có độ khó thấp.

Câu hỏi phân loại: là những câu hỏi có độ khó nằm trong phạm vi phân bố chuẩn (từ - 3 đến +3), thể hiện qua đường cong đặc trưng của từng câu (hàm đặc trưng câu hỏi). Những câu hỏi này dùng để phân biệt năng lực của từng đối tượng (thể hiện qua mức tương quan năng lực - độ khó, độ phân biệt).

Câu hỏi không phù hợp: gồm những câu có hệ số tương quan dưới mức 0, bằng 0, những câu

100% đối tượng đều đáp ứng như nhau hoặc những câu không đối tượng nào trả lời được. Những câu hỏi này sẽ bị phần mềm loại nên không có hàm thông tin.

Bước 2: chuẩn hoá trắc nghiệm: Một trắc nghiệm tốt thể hiện qua độ tin cậy trắc nghiệm và đường cong hàm thông tin trắc nghiệm; điểm cực đại của hàm thông tin trắc nghiệm (đây là điểm phân ngưỡng của trắc nghiệm và từ điểm này có thể tính ra điểm thô tối thiểu là tổng điểm của các câu hỏi có độ khó nằm phía dưới giá trị điểm cực đại).

III. KẾT QUẢ

1. Một số đặc điểm của đối tượng nghiên cứu

Bảng 1. Một số đặc điểm của đối tượng nghiên cứu

		$\bar{X} \pm SD$	95% độ tin cậy	p
Giới	Nam	27,37 ± 1,74	26,92 ÷ 27,82	0,273
	Nữ	27,68 ± 1,40	27,32 ÷ 28,04	
Trình độ	Đại học	28 ± 1,28	27,67 ÷ 28,33	0,0008
	Dưới đại học	27,05 ± 1,71	26,61 ÷ 27,49	

Kết quả ở bảng 1 cho thấy điểm trung bình của nhóm đối tượng nghiên cứu dao động từ 26,92 - 28,04 và không có sự khác biệt giữa nam và nữ nhưng có sự khác biệt có ý nghĩa giữa nhóm có trình độ đại học và dưới đại học với $p < 0,001$.

2. Các câu hỏi thành phần của bộ câu hỏi MMSE

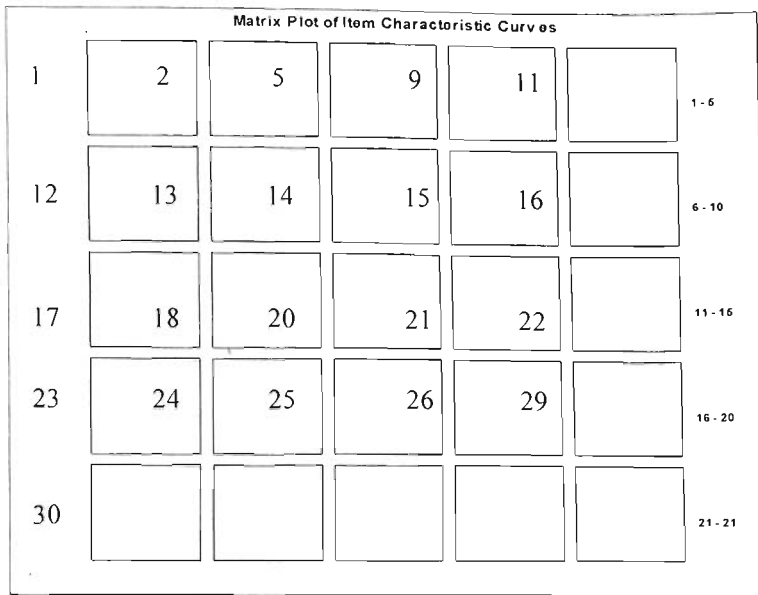
Độ khó, độ phân biệt, hệ số tương quan của các câu hỏi được trình bày trong bảng 2.

Bảng 2 cho thấy độ khó các CH dao động từ - 4,436 đến 0,037, chín câu không xuất hiện trong bảng độ khó. Độ phân biệt dao động từ 0,417 đến 1,554. Hệ số tương quan của các câu hỏi dao động từ - 0,303 đến 0,496. Hai câu có hệ số tương quan âm là câu số 4, 19 và bảy câu có hệ số tương quan bằng 0 là câu 3, 6, 7, 8, 10, 27, 28.

Đường cong đặc trưng của từng câu hỏi trong biểu đồ 1 được sắp xếp theo thứ tự câu hỏi trong bảng độ khó (bảng 2). Chín câu không có đường cong đặc trưng là câu 3, 4, 6, 7, 8, 10, 19, 27, 28.

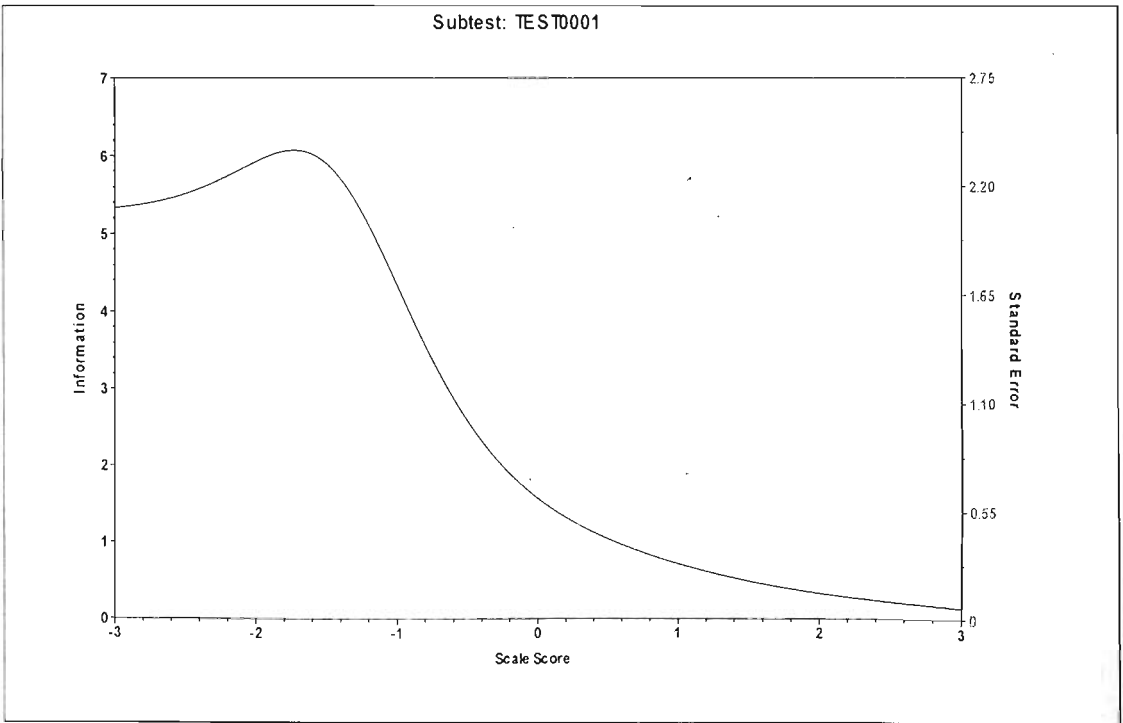
Bảng 2. Độ khó, độ phân biệt của các câu hỏi trong MMSE

Câu hỏi	Độ phân biệt	Độ khó	Hệ số tương quan
ITEM0001	0,914	- 3,747	0,239
ITEM0002	0,762	- 3,379	0,058
ITEM0003			0,000
ITEM0004			- 0,199
ITEM0005	0,903	- 3,017	0,532
ITEM0006			0,000
ITEM0007			0,000
ITEM0008			0,000
ITEM0009	0,777	- 4,203	0,045
ITEM0010			0,000
ITEM0011	1,219	- 2,253	0,438
ITEM0012	1,554	- 1,542	0,496
ITEM0013	0,603	- 2,618	0,185
ITEM0014	0,555	- 3,515	0,164
ITEM0015	0,686	- 0,090	0,364
ITEM0016	0,484	- 1,783	0,172
ITEM0017	0,478	- 3,416	0,155
ITEM0018	0,686	- 2,057	0,433
ITEM0019			- 0,303
ITEM0020	1,319	- 1,335	0,462
ITEM0021	0,479	0,037	0,222
ITEM0022	0,417	- 2,588	0,044
ITEM0023	0,562	- 2,049	0,245
ITEM0024	0,724	- 4,436	0,045
ITEM0025	0,914	- 3,747	0,239
ITEM0026	0,843	- 3,961	0,239
ITEM0027			0,000
ITEM0028			0,000
ITEM0029	0,699	- 3,596	0,293
ITEM0030	0,417	- 1,722	0,203



Biểu đồ 1. Hàm đặc trưng câu hỏi trong MMSE

2. Bộ trắc nghiệm đánh giá trạng thái tâm trí tối thiểu MMSE



Biểu đồ 2. Hàm thông tin của bộ trắc nghiệm MMSE

Biểu đồ 2 cho thấy hàm thông tin có điểm "cực đại" ứng với giá trị - 1,65; điểm số kỳ vọng của nhóm đối tượng có mức năng lực là - 1,65 là 17 trên 21 câu hỏi.

IV. BÀN LUẬN

1. Một số đặc tính chung

Kết quả ở bảng 1. chỉ ra không có sự khác biệt về điểm thô trung bình của trắc nghiệm MMSE theo giới nhưng lại có sự khác biệt theo trình độ học vấn. Kết quả này cũng phù hợp với nhiều nghiên cứu trong và ngoài nước khi nhận xét rằng trình độ học vấn luôn ảnh hưởng đến chức năng nhận thức ở mọi lớp tuổi, trên cả hai giới [5, 7].

2. Phân tích các câu hỏi thành phần của bộ câu hỏi MMSE

Kết quả phân tích các câu hỏi của trắc nghiệm MMSE cho thấy mô hình phân tích đã loại ra chín câu hỏi không phù hợp với nhóm đối tượng nghiên cứu. Hai câu hệ số tương quan dưới mức 0 là câu số 4, 19. Câu số 4, 19 là câu hỏi *hôm nay là ngày bao nhiêu và nhắc lại tên ba từ* là hai câu được đánh giá là khó với hầu hết các độ tuổi, trên cả người bình thường và bệnh nhân. Như vậy nếu dùng câu hỏi này sẽ không phân biệt được người bình thường và bệnh lý. Bảy câu có hệ số tương quan bằng 0 là câu 3, 6, 7, 8, 10, 27, 28. Đây cũng là những câu có những vấn đề không phù hợp. thường là những câu quá dễ so với năng lực của nhóm đối tượng này với tỷ lệ trả lời đúng là 99 - 100%. Trên biểu đồ hàm đặc trưng câu hỏi, chín câu này không có đường cong đặc trưng. Kết quả này tương tự như một số nghiên cứu khác, các tác giả cho rằng đây là những câu quá dễ, không đòi hỏi một mức năng lực nhận thức cao và sẽ tạo ra nhiều sai số. Các đối tượng thường làm không đúng những câu hỏi yêu cầu cao hơn như câu nhắc lại 3 từ, cấu trúc thị giác. Như vậy, rất cần có những nghiên cứu khác theo mô hình này để có thể đi đến thống nhất nên sửa đổi hay loại bỏ những câu trên như thế nào cho phù hợp với người Việt Nam. Một nghiên cứu lượng tự nhưng tiến hành trên nhóm đối tượng

bình thường có độ tuổi cao hơn cũng cho thấy nếu phân tích theo mô hình hai tham số thì có 14 câu hỏi bị loại khỏi mô hình phân tích, để lại các câu mà đáp ứng đúng dao động từ 66 - 97%, hệ số tương quan nhị phân câu hỏi - tổng điểm dao động từ 0,22 - 0,58 [4]. Các câu còn lại có tương quan trên mức 0 ở biểu đồ 3.1 cho thấy có mười câu hỏi đường cong đặc trưng có độ dốc thấp là câu: 1, 2, 5, 9, 14, 17, 24, 25, 26, 29. Tương ứng trên bảng độ khó các câu này đều có độ khó thấp hơn so với mức năng lực của nhóm đối tượng (độ khó nằm ngoài phạm vi từ - 3 đến + 3). Một số câu có độ dốc cao hơn gồm câu 11, 12, 13, 15, 16, 18, 20, 21, 22, 23, 30, là những câu có độ khó nằm trong phạm vi phân bố chuẩn, những câu này tạo nên điểm cực đại trên hàm thông tin của trắc nghiệm và là những "câu hỏi phân loại" của trắc nghiệm.

Như vậy qua phân tích từng câu hỏi chúng tôi nhận thấy rằng các câu hỏi trong bộ trắc nghiệm MMSE có các độ khó khác nhau mặc dù cách cho điểm lại giống nhau là đúng được 1 sai được 0 điểm. Điều này đồng nghĩa với việc nếu tiếp tục sử dụng cách cho điểm như vậy sẽ khó xác định được người có rối loạn hay không rối loạn nhận thức. Các nghiên cứu tiếp theo nên tìm hiểu yếu tố nào ảnh hưởng đến sự trả lời được hay không được của các nhóm đối tượng. Kết quả phân tích 21 câu hỏi còn lại bao gồm những câu đo lường khả năng định hướng, tính toán, ngôn ngữ, cấu trúc thị giác. Nhiều nghiên cứu trên thế giới cũng cho kết quả và những khuyến nghị tương tự. Bộ câu hỏi 30 câu MMSE bao trùm 6 lĩnh vực của nhận thức nhưng chỉ có rất ít câu cần thiết để cung cấp thông tin về khả năng nhận thức thực tế của đối tượng. Việc loại bỏ những câu ít thông tin sẽ làm công cụ này ngắn hơn mà vẫn đảm bảo độ tin cậy trong đánh giá chức năng nhận thức.

2. Phân tích bộ trắc nghiệm đánh giá trạng thái tâm trí tối thiểu MMSE

Độ tin cậy của bộ trắc nghiệm MMSE là 0,60. Kết quả này thấp hơn so với nhóm tuổi > 60 của một số tác giả trong và ngoài nước [2], [8]. Phải chăng độ tin cậy và độ giá trị của trắc nghiệm MMSE giảm ở nhóm tuổi trẻ hơn. Điểm cực đại trên hàm thông tin tương ứng với giá trị - 1,65, qua điểm này chúng tôi lấy điểm thô tối thiểu của bộ trắc cho nhóm đối tượng trong độ tuổi này là 17 - 18 điểm trên 21 câu hỏi được phân tích. Nếu so với mốc 24 trên 30 câu hỏi được coi là mốc bắt đầu có rối loạn nhận thức ở những đối tượng > 65 tuổi như một số nơi đang áp dụng thì kết quả tính toán từ mô hình Rasch cũng tương tự ở nhóm tuổi 50 - 59 nhưng với số lượng câu hỏi ít hơn. Theo kết quả phân tích này, sẽ có một tỷ lệ đối tượng dưới điểm thô tối thiểu khi làm bộ trắc nghiệm MMSE 30 câu nhưng lại có thể có điểm số > 17 câu khi làm trắc nghiệm 21 câu. Những đối tượng này sẽ là những trường hợp "âm tính giả" [5]. Nếu tăng điểm thô tối thiểu lên 18 thì số trường hợp âm tính giả sẽ tăng lên. Do mục đích của trắc nghiệm là cố gắng phát hiện sớm những trường hợp bắt đầu có biểu hiện rối loạn nhận thức nên theo nhiều đề nghị từ những nghiên cứu nước ngoài nên chọn điểm cutoff là 18 nhằm lưu ý được nhiều hơn số người trong diện cần theo dõi vì âm tính giả.

V. KẾT LUẬN

Trong tổng số 30 câu trắc nghiệm MMSE có nhiều mức độ khó khác nhau bao trùm nhiều lĩnh vực của chức năng nhận thức. 2 câu có tương quan dưới 0, 7 câu tương quan bằng 0 bị loại khỏi mô hình phân tích. Trong 21/30 câu hỏi thành phần của MMSE được phân tích, 10 câu có độ khó không phù hợp, 11 câu có độ khó nằm trong phạm vi phân bố chuẩn. Nhiều câu hỏi tỏ ra quá dễ so với năng

lực của nhóm đối tượng, cần xem xét để thay thế hoặc chỉnh sửa. Số lượng câu hỏi có tính phân loại ít (11/30 câu). Điểm thô tối thiểu trắc nghiệm MMSE là 18/21. Tuy nhiên chỉ có một số ít câu thể hiện cung cấp nhiều thông tin cho đánh giá năng lực của nhóm đối tượng trong nghiên cứu này. Với mô hình hai tham số có thể cho thấy có thể đề xuất bộ MMSE với những câu hỏi đơn giản, số lượng ít hơn cho những nhóm đối tượng có độ tuổi khác nhau mà vẫn đảm bảo được tính giá trị cũng như độ tin cậy của cả bộ câu hỏi trắc nghiệm.

TÀI LIỆU THAM KHẢO

1. **Trần Thúy Liễu (2008)**, "Góp phần chuẩn hóa một số trắc nghiệm tâm lý thần kinh trong đánh giá chức năng nhận thức ở người Việt Nam bình thường tuổi 50 - 59," Luận văn Thạc sỹ Y học.
2. **Lương Linh Ly, Bùi Mỹ Hạnh (2009)**, "Điểm số trắc nghiệm trí nhớ ở người cao tuổi Việt Nam", Nghiên cứu Y học, 61 (2): tr. 73 - 83
3. **K. Schultz - Larsen, S. Kreiner and R.K. (2006)**. Lomholt, Mini - Mental Status Examination: Mixed Rasch model item analysis derived two different cognitive dimensions of the MMSE, J Clin Epidemiol 60, pp. 256 - 259.
4. **K. Schultz - Larsen, R.K. Lomholt and S. Kreiner (2006)**, Mini - Mental Status Examination: a short form of MMSE was as accurate as the original MMSE in predicting dementia, J Clin Epidemiol 60, pp. 260 - 267.
5. **M.F. Folstein, S.E. Folstein and P.R. McHugh (1975)**, Mini - mental state. A practical method for grading the cognitive state of patients for the clinician, J Psychiatr Res 12, pp. 189 - 198.
6. **P.K. Crane, L.E. Gibbons, L. Jolley, G. van Belle, R. Salleri and E. Dalmonte et al., (2006)** Differential item functioning related to education

and age in the Italian version of the Mini - mental State Examination, *Int Psychogeriatr*, pp. 18 [Epub ahead of print].

7. **P.R. De Brito - Marques and J.E. (2004)** Cabral - Filho, The role of education in mini - mental state examination: a study in northeast

Brazil, *Arq Neuropsiquiatr* 62, pp. 206 – 211.

8. **A.A. Kucukdeveci, S. Kutlay, A. (2005).** Elhan and A. Tennant, Preliminary study to evaluate the validity of the mini - mental state examination in a normal population in Turkey, *Int J Rehabil Res* 28, pp. 77 – 79.

Summary

PSYCHOMETRIC PROPERTIES OF MINI - MENTAL STATE EXAMINATION (MMSE) OF HEALTHY PEOPLE AGED 50 - 59: USING THE TWO - PARAMETER RESPONSE MODEL

Mini - Mental State Examination (MMSE) is arguably the most widely used screening instrument for detection of an individual's cognitive impairment. Validation studies have very important role for developing Vietnamese translation. **Objectives:** This study aimed to assess the psychometric properties of the Vietnamese translation of MMSE. **Methods:** The original US English scales were translated into Vietnamese (orientation to time, place, registration, attention/calculation, recall and language) were administered for 120 subjects to the Vietnamese version, aged 50 - 59. Item response two - parameter polytomous model was used to analyse the data. **Results:** The analysis using the Rasch model showed that 21/30 of items fit the model's expectation. These items included mainly the measures of orientation, recall, attention/calculation, visual copy. The internal consistency of the 21 items was 0,60. The proposed new cut - off the 21 item MMSE was 18. The new cut - off point was determined for the purpose of over - identifying patients at risk so as ensure detection of prevention from onset of cognitive disturbance. **Conclusion:** The results demonstrate the feasibility, validity, item reliability, item fit, person fit of the Vietnamese version of MMSE Scales for population health research in Vietnam. However, only few items are need to dicribe the subject's cognitive status. Using two parameter response theory analysis found that the MMSE could be simplified. Deleting some items makes MMSE not only shorter, easier to administer but also enables one to maintain validity as a cognitive function test for clinical setting.

Keywords: the Mini Mental Stage Examination, item response theory