

CHUYỂN ĐỔI CÔNG CỤ ĐO LƯỜNG SỰ TỰ TIN VÀO NĂNG LỰC HỢP TÁC CHUYÊN MÔN LIÊN NGÀNH THÀNH PHIÊN BẢN TIẾNG VIỆT: DỊCH NGHĨA VÀ TÍNH GIÁ TRỊ CỦA IPECC-SET 27 VÀ IPECC-SET 9

Nguyễn Thị Thùy Linh¹, Trần Thụy Khánh Linh², Nguyễn Anh Vũ³, Phạm Dương Uyển Bình⁴, Trần Diệp Tuấn⁵, Phạm Lê An⁶

TÓM TẮT:

Giới thiệu: Hiện nay, nhu cầu phát triển một công cụ đo lường hiệu quả cho chương trình giáo dục liên ngành ngày càng tăng. Cả hai bộ công cụ đo lường IPECC-SET 9 gồm 9 mục và IPECC-SET 27 gồm 27 mục được nghiên cứu và ứng dụng rộng rãi để đánh giá khả năng tự tin của sinh viên ngành y khoa trong việc hợp tác chuyên môn liên ngành với độ tin cậy cao. Bộ công cụ đánh giá IPECC-SET 27 đã được chuyển ngữ sang tiếng Việt, được hội đồng chuyên gia công nhận và phân tích đánh giá tính giá trị với mẫu cỡ nhỏ. Tuy nhiên, tính giá trị chưa được đánh giá dựa trên Bộ khung tiêu chuẩn trắc nghiệm giáo dục học và tâm lý học của AERA, APA & NCME. **Mục đích:** Nghiên cứu này nhằm đánh giá tính giá trị của IPECC-SET 27 và IPECC-SET 9 để có thể sử dụng được trong các chương trình giáo dục liên ngành tại Việt Nam. **Phương pháp:** Số liệu được phân tích lại trên mẫu 156 sinh viên ngành Y, Dược, Điều dưỡng và Vật lý trị liệu trước khi học module giáo dục liên ngành. Phân tích thống kê được thực hiện dựa trên một khung bằng chứng giá trị của Messick và sử dụng mô hình đo lường Rasch. Các thành phần giá trị của khung được tương ứng với các phân tích Rasch khác nhau. Dữ liệu được quản lý bằng Excel 2013, phân tích bằng R 4.3.1 và Jamovi 2.3.28. **Kết quả:** Về tổng thể các công cụ có tính giá trị chấp nhận được. Mô hình Rating Scale RSM được sử dụng. Đối với bộ IPECC 27 mục, thu được Cronbach's $\alpha = 0,984$, overall KMO = 0,961, explained variance = 71,5%, person reliability = 0,982. Trong SET 27 có ba mục TT3, RR4 và RR2 cần được xem xét và loại bỏ. Sau khi loại bỏ, infit biến thiên từ 0,568 – 1,341 và outfit dao động trong khoảng 0,624 – 1,363. Đối với SET 9, Cronbach's $\alpha = 0,947$, overall KMO = 0,909, explained variance = 70,4%, person reliability = 0,938, infit dao động trong khoảng 0,759 – 1,443 và outfit dao động 0,727 – 1,441. Các kết quả phân tích phiên bản tiếng Việt tương đồng khá rõ với kết quả phân tích trên nguyên bản tiếng Anh. Phân tích DIF không phát hiện sự khác biệt nào đủ rõ có liên quan đến giới tính và kinh nghiệm làm việc trước đó. **Kết luận:** Trên tổng thể, phiên bản tiếng Việt của IPECC-SET 9 và IPECC-SET 27 có đủ tính giá trị và có thể được sử dụng để đánh giá sự tự tin về năng lực bản thân trong hợp tác chuyên môn liên ngành giữa các sinh viên khối ngành sức khỏe. Tuy nhiên, cần thử nghiệm các công cụ này trên mẫu lớn hơn bao gồm nhiều thành viên nhiều chuyên ngành hơn đủ đại diện cho cộng đồng chuyên môn chăm sóc sức khỏe.

Từ khóa: giáo dục liên ngành, thực hành cộng tác liên ngành, tự tin vào năng lực bản thân, mô hình đo lường Rasch, IPECC-SET 9, IPECC-SET 27.

1. Ngành Điều dưỡng, Khoa Y, Đại học Quốc gia TP Hồ Chí Minh

2. Khoa Khoa học Cơ bản, Đại học Y Dược TP Hồ Chí Minh

3. Khoa Điều dưỡng – Kỹ thuật Y học, Đại học Y Dược TP Hồ Chí Minh

4. Phòng Đảm bảo Chất lượng Giáo dục và Khảo thí, Đại học Y Dược TP Hồ Chí Minh

5. Hội đồng trường, Đại học Y Dược TP Hồ Chí Minh

6. Khoa Y, Đại học Y Dược TP Hồ Chí Minh

Tác giả liên hệ: Nguyễn Anh Vũ - 0909 090 838 . Email: nguyenganhvu@ump.edu.vn.

ADAPTING MEASUREMENT OF SELF-EFFICACY FOR COMPETENCE IN INTERPROFESSIONAL COLLABORATION INTO VIETNAMESE VERSION: TRANSLATION AND VALIDATION OF IPECC-SET 27 AND IPECC-SET 9

SUMMARY

Introduction: There is currently an increasing need to develop an effective measurement tool for the Interprofessional Education (IPE) programs. The both 9-item and 27-item Interprofessional Education Collaborative Competency Self Efficacy Tools (IPECC-SET) were well-known measurement instruments to evaluate health professions students' self-efficacy in interprofessional collaborative competency with robust validity. The assessment tool IPECC-SET 27 has been translated into Vietnamese versions, accredited and pilot analyzed with a small sample. However, their validation is still not fullfilled based on *Standards for Educational and Psychological Testing* of AERA-APA-NCME.

Purpose: This study is to validate IPECC-SET 27 and IPECC-SET 9 assessment toolkits for using them in Vietnamese IPE programs.

Methods: The data was re-analyzed on a sample of 152 students from the fields of medicine, pharmacy, nursing, and physical therapy before taking the interdisciplinary education module. The statistical analysis was conducted based on Messick's evidence framework and used the Rasch measurement model. The framework's value components were mapped to different Rasch analyses. The data was managed using Excel 2013, analyzed using R 4.3.1 and Jamovi 2.3.28.

Results: The overall results suggested that the instruments displayed acceptable validity. Rating Scale (RSM) model was used for analysis across the subsequent steps. For SET 27 we get Cronbach's α 0.984, overall KMO 0.961, explained variance 71.5%, person reliability 0.982. In the same item set there were three items TT3, RR4 and RR2 that needed to consider. After removing these, infit ranges 0.568 – 1.341 and outfit ranges 0.624 – 1.363. For SET 9, we get Cronbach's α 0.947, overall KMO 0.909, and explained variance 70.4%, person reliability 0.938, infit ranges 0.759 – 1.443 and outfit ranges 0.727 – 1.441. These results on Vietnamese version are quite similar to the analysis of the original English measurement tools. The DIF analysis did not reveal any systematic item differences in relation to gender and earlier work experience.

Conclusion: According to the findings, the Vietnamese IPECC-SET 9 and IPECC-SET 27 demonstrate valid psychometric properties and can potentially be used to assess self-efficacy for competence in interprofessional collaborative practice among health profession students. However, it is recommended to conduct further testing of these instruments in larger samples that include a diverse range of healthcare team members to represent the healthcare community.

Keywords: Interprofessional education, interprofessional collaborative practice, self-efficacy, Rasch measurement model.

ĐẶT VẤN ĐỀ

Hiện nay, nhu cầu phát triển một công cụ đo lường hiệu quả cho chương trình giáo dục liên ngành ngày càng tăng. Mức độ tự tin vào năng lực bản thân là một thước đo quan trọng xác định cách cảm nhận, suy nghĩ, có động lực và hành vi như thế nào. Cấu trúc này cũng ảnh hưởng đến mức độ nỗ lực của người học cũng như sự kiên trì kháng bại trước những chướng ngại và trải nghiệm tiêu cực (Kathryn và Miele, 2016). Tất cả những yếu tố này đều rất quan trọng đối với sự hợp tác liên ngành giữa các chuyên gia chăm sóc sức khỏe. Nhiều bộ công cụ đánh giá cấu trúc này đã được xây dựng, để khám phá mức độ tự tin về năng lực bản thân của sinh viên ngành y tế trong thực hành cộng tác chuyên môn liên ngành (Kottorp và cộng sự, 2018). Hasnain và các cộng sự (2017) đã thiết kế và đánh giá bộ công cụ IPECC-SET 38 gồm 38 mục, dựa trên bộ tiêu chuẩn các năng lực cốt lõi trong hợp tác chuyên môn liên ngành của IPEC năm 2011. Phiên bản sửa đổi năm 2016 của IPEC cho thấy sự hợp tác liên ngành là một cấu trúc đơn chiều. Điều này dẫn đến việc xây dựng trên nền tảng IPECC-SET 38 một thang đo đơn chiều IPECC-SET 27 gồm 27 mục, và một phiên bản IPECC-SET 9 ngắn hơn chỉ gồm 9 mục mà vẫn đáp ứng các tiêu chí về tính đơn chiều và năng lực đánh giá (Kottorp và cộng sự, 2018; Axelsson và cộng sự, 2022). Nguyễn Thị Thùy Linh và nhóm nghiên cứu (2022) đã thực hiện chuyển ngữ IPECC-SET 27. Phiên bản tiếng Anh được dịch sang tiếng Việt, sau đó được dịch ngược trở lại tiếng Anh bởi các phiên dịch viên độc lập. Sau đó bản dịch được xem xét và đánh giá bởi một nhóm chuyên gia Việt Nam và một nhóm chuyên gia tiếng Anh. Mỗi mục và toàn bộ công cụ đều được đánh giá về tính thích hợp và rõ ràng, sử dụng chỉ số giá trị nội dung CVI và thống kê Kappa. Kết quả công nhận IPECC-SET 27 dẫn đến IPECC-SET 9 cũng đã được chuyển ngữ và công nhận. IPECC-SET 27 đã được đánh giá tính giá trị giá qua phân tích thăm dò với mẫu cỡ nhỏ gồm 52 sinh viên. Tuy nhiên các nguồn bằng chứng về tính giá trị chưa được kiểm chứng theo các khung giá trị hiện đại. Quan điểm hiện nay coi tính giá trị là một giả thuyết, do đó không thể được chứng minh nhưng bằng chứng có thể hỗ trợ hoặc bác bỏ luận điểm tính giá trị. Khung đánh giá tính giá trị được đề xuất năm 1989 (Messick, 1989, 1994, 1995) và được chấp nhận làm tiêu chuẩn cho lĩnh vực này vào năm 1999 và một lần nữa vào năm 2014 (Cook, Beckman, 2006). Theo Messick (1995), mọi giá trị đều được thống nhất như những khía cạnh của giá trị cấu trúc, nói lên mức độ phù hợp với khuôn khổ của cấu trúc lý thuyết được nghiên cứu của các kết quả đo lường trên cấu trúc lý thuyết đó. Mặc dù giá trị là một khái niệm thống nhất được biểu hiện trong giá trị cấu trúc, trong đó xác định sáu khía cạnh riêng biệt của giá trị cấu trúc: nội dung, thực chất, kết cấu, tổng quát hóa, bên ngoài và hệ quả. Nghiên cứu này có mục đích đánh giá lại tính giá trị của IPECC-SET 27 và IPECC-SET 9 theo khung tiêu chuẩn giá trị của Messick để có thể sử dụng được hai công cụ này trong chương trình giáo dục liên ngành tại Việt Nam.

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP

Số liệu thứ cấp được sử dụng và phân tích lại trên 156 sinh viên Đại học Y Dược Thành phố Hồ Chí Minh, chuyên ngành Y đa khoa năm 4, Dược năm 4, Điều dưỡng năm 3, Vật lý trị liệu năm 3 chuẩn bị học phần Giáo dục liên ngành 2021-2022. Số liệu gốc được Nguyễn Thị Thùy Linh thu thập (Nguyễn Thị Thùy Linh và cộng sự, 2022). Bộ công cụ IPECC-SET 27 và SET 9 được tách riêng phần bản dịch, sau đó được phân tích riêng. Phân tích thống kê sử dụng các mô hình Partial Credit Model (PCM) và Rating Scale Model (RSM). Kết quả phân tích được xác định

tiêu chuẩn bằng chứng theo khung giá trị Messick và được sắp xếp vào khung tiêu chuẩn. Số liệu được quản lý bằng Excel 2013, phân tích bằng R 4.3.1 và Jamovi 2.3.28.

Bảng 1. Các chiều giá trị theo Messick và các phân tích Rasch tương ứng.

Chiều	Ý nghĩa liên quan	Phân tích Rasch
Nội dung	Mức độ phản ánh tốt cấu trúc được đo của điểm số đo lường.	Point-measure correlations, Item fit, Wright map, Item separation index
Thực chất	Mức độ phản ánh sự có ý nghĩa về mặt lý thuyết của điểm số đo lường.	Person fit, Rating scale structure analysis
Kết cấu	Mức độ phù hợp giữa cấu trúc trắc nghiệm và mô hình điểm số.	Dimensionality analysis
Tổng quát hóa	Mức độ tổng quát hóa ý nghĩa và diễn giải điểm số đo lường trên các nhóm dân số, bối cảnh, nhiệm vụ.	Differential item functioning (DIF), Person reliability index
Bên ngoài	Mức độ tương quan giữa số đo với các số đo bên ngoài khác của cùng một cấu trúc, hoặc có liên quan	Wright map, Person separation index
Hệ quả	Mức độ quan trọng của việc giải thích điểm số, kết quả và quyết định dựa trên các phép đo và quyết định được đưa ra dựa trên các phép đo.	Items fit, Differential item functioning, Wright map

KẾT QUẢ NGHIÊN CỨU

Bảng 2. Đặc điểm của sinh viên tham gia nghiên cứu.

Biến số	Chung (n = 156)	Y (n = 50)	Dược (n = 64)	Điều dưỡng (n = 28)	Vật lý trị liệu (n = 14)
Giới tính					
Nam, n(%)	61 (39,1)	33 (54,1)	23 (37,7)	2 (3,3)	3 (4,9)
Nữ, n(%)	95 (60,9)	17 (17,9)	41 (43,2)	26 (27,4)	11 (11,5)
Tuổi (năm), Q50 (Q25-Q75)	22 (21-22)	22 (21-22)	22 (22-22)	21(20,7-21)	20,5(20-21)
Thực hành lâm sàng					
Có, n(%)	63 (40,4)	35 (55,6)	0 (0)	28 (44,4)	0 (0)
Không, (n%)	93 (59,6)	15 (16,1)	64 (68,8)	0 (0)	14 (15,1)
Thời gian thực hành lâm sàng,					
Trung vị (Q25 – Q75)	12 (6 - 32,5)	32 (2-42,5)	0	9 (6-12)	0

Nhận xét: Sinh viên các ngành có độ tuổi gần như đồng nhất. Nói chung tỉ lệ giới tính chênh lệch rõ rệt ($p < 0,0001$; KTC95% = 15,6% – 29,1%) do sinh viên nữ chiếm đa số trong các ngành Dược, Vật lý trị liệu và Điều dưỡng. Trong khi sinh viên ngành Y có tỉ lệ sinh viên nam cao hơn nữ ($p = 0,023$, KTC95% = 24,8%-56,7%). Thời gian thực hành lâm sàng của nhóm sinh viên Y phân tán khá mạnh so với Điều dưỡng, cũng có trung vị cao hơn rõ rệt.

Đối với bộ IPECC-SET27, kết quả thu được Cronbach's $\alpha = 0,984$, overall KMO = 0,961, explained variance = 71,5%, person reliability = 0,982. Vì không có sự khác biệt đáng kể giữa các mô hình Partial Credit (PCM) và Rating Scale (RSM) ($p = 0,950$), đồng thời RSM có tiêu chí thông tin (AIC = 12.669, BIC = 12.779) tương đối tốt hơn so với PCM (AIC = 12.910, BIC = 13.654), nên RSM được chọn để áp dụng trong các bước tiếp theo.

Bảng 3. Độ khó câu hỏi và độ phù hợp mô hình của bộ công cụ IPECC-SET27

Thống kê	Điểm thô	Số câu hỏi	Độ khó logit	SE độ khó logit	Infit MnSq	Outfit MnSq
Trung bình	158	27	-1,073	0,076	0,973	0,984
Sai số chuẩn	3,47	27	0,053	0,0003	0,061	0,062
Trung vị	165	27	-1,069	0,076	0,948	0,962
Lớn nhất	233	27	-0,518	0,081	1,862	1,906
Nhỏ nhất	36	27	-1,739	0,073	0,586	0,6

Nhận xét: Có dấu hiệu không phù hợp mô hình, cần rà soát bộ câu hỏi. Trong SET27 có ba mục câu hỏi TT3, RR4 và RR2 cần được xem xét và loại bỏ. Sau khi loại bỏ, infit biến thiên từ 0,568 – 1,341 và outfit dao động trong khoảng 0,624 – 1,363.

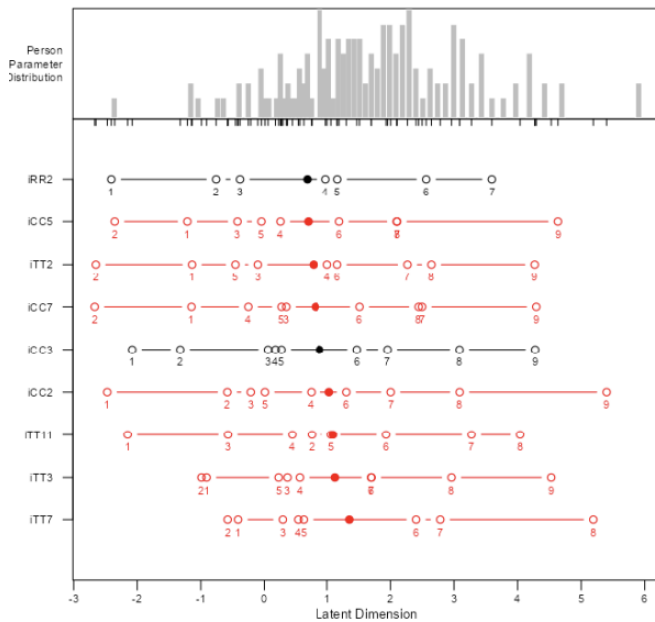
Đối với SET 9, không có sự khác biệt đáng kể giữa các mô hình Partial Credit (PCM) và Rating Scale (RSM) không đáng kể ($p = 0,379$). Tiêu chí thông tin của mô hình RSM (AIC = 4.594, BIC = 4.649) có tốt hơn PCM (AIC = 4.656, BIC = 4.906), vì vậy mô hình RSM được sử dụng.

Bảng 4. Độ khó câu hỏi và độ phù hợp mô hình của bộ công cụ IPECC-SET9

Thống kê	Điểm thô	Số câu hỏi	Độ khó logit	SE độ khó logit	Infit MnSq	Outfit MnSq
Trung bình	53,20	9	-0,928	0,073	1,002	1,003
Sai số chuẩn	1,15	9	0,080	0,0007	0,084	0,087
Trung vị	55,00	9	-0,867	0,072	0,897	0,905
Lớn nhất	81,00	9	-0,707	0,078	1,443	1,441
Nhỏ nhất	9,00	9	-1,509	0,071	0,759	0,727

Nhận xét: Có dấu hiệu phù hợp với mô hình.

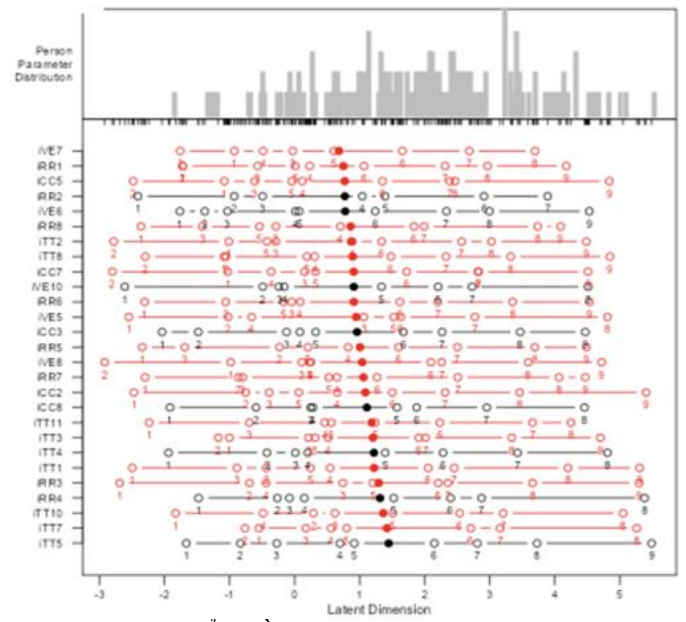
SET 9 có Cronbach's $\alpha = 0,947$, overall KMO = 0,909, and explained variance = 70,4%, person reliability = 0,938, infit dao động trong khoảng 0,759 – 1,443 và outfit dao động 0,727 – 1,441.



Hình 1. Biểu đồ Wright của IPECC-SET9

Nhận xét: Biểu đồ Wright cho thấy phân bố trên thang logit của cá nhân và câu hỏi khá đều, đủ hiển thị sự phân biệt giữa những mức tự tin cao so với những mức thấp hơn. Mức độ xáo trộn ngưỡng mạnh hơn ở SET27 so với SET9. Không phát hiện có dấu hiệu tách nhóm đủ rõ.

Phân tích DIF cũng không phát hiện sự khác biệt nào đủ rõ có liên quan đến giới tính và kinh nghiệm làm việc trước đó.



Hình 2. Biểu đồ Wright của IPECC-SET27

IV. BÀN LUẬN: Kết quả phân tích IPECC-SET9 và SET29 cho thấy đặc tính của phiên bản tiếng việt tương đồng khá rõ với các đặc tính của nguyên bản tiếng Anh (Hasnain, Gruss et al, 2017; Kottorp, Keehn et al, 2018; Axelsson, Kottorp et al, 2022). Kết quả cũng cho thấy mô hình Rasch RSM có thể cung cấp bằng chứng thực nghiệm để hỗ trợ sáu khía cạnh của khung giá trị Messick cho công cụ đo lường IPECC-SET. Rasch RSM không sử dụng điểm thô như mô hình trắc nghiệm cổ điển CTT, các mục được phân loại dựa trên mức độ khó để cung cấp điểm số logit cho phân tích thang đo tuyến tính. Nghiên cứu hiện tại sử dụng Rasch RSM như một mô hình phù hợp hơn CTT trong đánh giá tính giá trị đo lường của công cụ tâm lý giáo dục học IPECC-SET 9 và 27. Sáu khía cạnh của khung giá trị Messick đã được so khớp tương ứng với các phân tích khác nhau trong Rasch RSM để tổ chức hệ thống nguồn bằng chứng thực nghiệm phục vụ đánh giá tính giá trị đo lường. Dựa trên hệ thống bằng chứng này, kết quả tổng thể cho thấy IPECC-SET 7 và 27 thể hiện tính giá trị ở mức độ tương đối chấp nhận được, khá tương đồng với Kottorp et al. (2022) và Axelsson et al. (2018). Do đó, sử dụng Rasch RSM là đủ để cung cấp bằng chứng hỗ trợ các lập luận về giá trị dựa trên khung giá trị Messick. Quyết định sử dụng một công cụ đánh giá dựa trên việc xem xét kỹ lưỡng cơ sở lập luận logic và bằng chứng thực nghiệm để xác định xem chúng có hỗ trợ các suy luận, diễn giải và kết luận dựa trên kết quả đánh giá hay không. Các khung lý thuyết giá trị hiện đại coi luận điểm giá trị là một giả thuyết và cần có bằng chứng thực nghiệm xác thực để hỗ trợ hoặc bác bỏ. Theo Downing (2003), kết luận về tính giá trị đòi hỏi phải có chuỗi bằng chứng liên kết rõ ràng các diễn giải về kết quả trắc nghiệm với mạng lưới các lý thuyết, giả thuyết và logic được trình bày để hỗ trợ hoặc bác bỏ tính hợp lý của các diễn giải được xem xét. Tính giá trị là đặc tính của điểm số, diễn giải kết quả và quyết định, không phải đặc tính của công

cụ. Tính giá trị cũng có mức độ, không phải là quyết định có hoặc không. Hơn nữa, xác nhận về tính giá trị là một quá trình, không phải là điểm kết thúc (Cook, Hatala, 2016). Tính giá trị có thể được đánh giá nhưng không bao giờ có thể được chứng minh. Điều đó dẫn đến một chuỗi chu kỳ kiểm định dựa trên các nguồn bằng chứng theo Messick cho đến khi tất cả các luận điểm về tính giá trị được chấp nhận là hợp lý hoặc bị bác bỏ.

V.KẾT LUẬN: Trên tổng thể, phiên bản tiếng Việt của IPECC-SET 9 và IPECC-SET 27 có đủ tính giá trị và có thể được sử dụng để đánh giá sự tự tin về năng lực bản thân trong hợp tác chuyên môn liên ngành giữa các sinh viên khối ngành sức khỏe. Tuy nhiên, vẫn cần thử nghiệm các công cụ này trên mẫu lớn hơn bao gồm nhiều thành viên nhiều chuyên ngành hơn đủ để đại diện cho cộng đồng chuyên môn chăm sóc sức khỏe.

TÀI LIỆU THAM KHẢO

- 1.American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- 2.Baghaei, P., & Amrahi, N. (2011). Validation of a multiple choice English vocabulary test with the Rasch model. *Journal of Language Teaching and Research*, 2, 1052–1060.
- 3.Cook, A., Beckman, T.J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application David. *The American Journal of Medicine*, 119, 166.e7-16.
- 4.Cook, D.A., Hatala, R. (2016). Validation of educational assessments: a primer for simulation and beyond. *Advances in Simulation*, 1 - 31.
- 5.Downing SM (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education*, 37, 830-837.
- 6.Hasnain, M., Gruss, V., Keehn, M., Peterson, E., Valenta, A. L., & Kottorp, A. (2017). Development and validation of a tool to assess self-efficacy for competence in interprofessional collaborative practice. *Journal of Interprofessional Care*, 31, 255–262.
- 7.Kottorp, A., Keehn, M., Hasnain, M., Gruss, V., Peterson, E. (2018) Instrument Refinement for Measuring Self-Efficacy for Competence in Interprofessional Collaborative Practice: Development and Psychometric Analysis of IPECC-SET 27 and IPECC-SET. *Journal of Interprofessional Care*, 1-10.
8. Axelsson, M., Kottorp, A., Carlson, E., Gudmundsson, P., Kumlien, C. & Jakobsson, J. (2022): Translation and validation of the Swedish version of the IPECC-SET 9 item version, *Journal of Interprofessional Care*, 900-907.
- 9.Messick, S. (1994).Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning. Research Report RR-94-45. *Educational Testing Service*, Princeton.
- 10.Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.
- 11.Murphy, K.R. (2009). Validity, Validation and Values. *The Academy of Management Annals*, 3(1), 421-461.

12. Nguyễn Thị Thùy Linh , Trần Thụy Khánh Linh, Nguyễn Anh Vũ, Hà Thị Như Xuân, Dương Duy Khoa, Katrina Einhellig, Trần Diệp Tuấn. (2022). Chuyển ngữ và tính giá trị của công cụ tự đánh giá về năng lực hợp tác trong giáo dục liên ngành 27 mục (IPECC-SET 27) phiên bản tiếng Việt. *Tạp chí Y học Việt Nam*, Tập 519 – Tháng 10 – Số chuyên đề, 91-102.

13. Kathryn R. Wentzel, K.R., Miele D.B. (2016). *Handbook of Motivation at School*, 2nd Ed. Routledge, Educational Psychology Handbook Series.

PHẦN MỀM ĐƯỢC SỬ DỤNG

1. R Core Team (2021). R: A Language and environment for statistical computing. (Version 4.1) Retrieved from <https://cran.r-project.org>.

2. Revelle, W. (2019). psych: Procedures for Psychological, Psychometric, and Personality Research. [R package]. Retrieved from <https://cran.r-project.org/package=psych>

3. The jamovi project (2022). jamovi. (Version 2.3) Retrieved from <https://www.jamovi.org>.

4. Robitzsch, A., Kiefer, T., & Wu, M. (2020). TAM: Test Analysis Modules. (Version 4.1.4). Retrieved from <https://CRAN.R-project.org/package=TAM>.

5. Robitzsch, A., Kiefer, T., George, A., & Uenlue, A. (2022). CDM: Cognitive Diagnosis Modeling. (Version 8.2.6). Retrieved from <https://CRAN.R-project.org/package=CDM>.

6. Martinkova, P., & Drabinova, A. (2018). ShinyItemAnalysis: for teaching psychometrics and to enforce routine analysis of educational tests. (Version 1.4.2. Retrieved from <https://CRAN.R-project.org/package=ShinyItemAnalysis>.

7. Mair, P., Hatzinger, R., Maier, M., Rusch, T., & Debelak R. (2021). eRm: Extended Rasch Modeling. (Version 1.0.2)[R package]. Retrieved from <https://CRAN.R-project.org/package=eRm>.

8. Seol, H. (2022). snowIRT: Item Response Theory for jamovi. (Version 4.8.1)[jamovi module]. URL <https://github.com/hyunsooseol/snowIRT>