

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC NGOẠI NGỮ
KHOA SAU ĐẠI HỌC**

NGUYỄN THỊ MAI HỮU

**AN INVESTIGATION INTO THE COGNITIVE VALIDITY OF THE
SPEAKING SECTION OF THE VIETNAMESE STANDARDIZED TEST
OF ENGLISH PROFICIENCY (VSTEP.3-5)**

*Nghiên cứu giá trị xác thực đối với quá trình tư duy của thí sinh khi thi phần thi nói
bài thi chuẩn hóa đánh giá năng lực tiếng Anh từ bậc 3 đến bậc 5 theo Khung năng
lực ngoại ngữ 6 bậc dùng cho Việt Nam (VSTEP.3-5)*

Chuyên ngành: Lý luận và phương pháp dạy học bộ
môn tiếng Anh

Mã số: 9140231.01

Cán bộ hướng dẫn: GS. Hoa Nguyen

GS. Fred Davidson

TÓM TẮT LUẬN ÁN TIẾN SĨ

HÀ NỘI, 2021

CHƯƠNG 1: DẪN NHẬP

1.1. Mục đích nghiên cứu

Nghiên cứu này được thực hiện nhằm mục đích:

- Mô tả quá trình tư duy của thí sinh thi kỹ năng nói bài thi đánh giá năng lực tiếng Anh bậc 3 đến bậc 5 theo Khung năng lực ngoại ngữ 6 bậc dùng cho Việt Nam (VSTEP.3-5) để xác định giá trị xác thực của quá trình này của bài thi VSTEP.3-5, và
- Xác định giá trị xác thực của quá trình tư duy của thí sinh thi bài thi nói VSTEP.3-5.

Với mục đích như vậy, các câu hỏi nghiên cứu sau đã được xây dựng:

- **CH1:** Bài thi VSTEP.3-5 có phản ánh quá trình tư duy của thí sinh khi nói không?
- **CH2:** Các yêu cầu về tư duy được sắp xếp như thế nào trong phần thi nói bài thi VSTEP.3-5 (theo từng bậc năng lực tiếng Anh của thí sinh)?
- **CH3:** Quá trình tư duy của thí sinh khi thi bài thi nói VSTEP.3-5 có giống với quá trình này khi thí sinh nói tiếng Anh trong bối cảnh không phải bài thi không?

1.2. Cấu trúc của nghiên cứu

Nghiên cứu được thực hiện theo 8 chương như sau:

Chương 1: Dẫn nhập

Chương 2: Cơ sở lý luận

Chương 3: Bài thi chuẩn hóa đánh giá năng lực tiếng Anh bậc 3 đến bậc 5 theo Khung năng lực ngoại ngữ 6 bậc dùng cho Việt Nam

Chương 4: Phương pháp nghiên cứu

Chương 5: Quá trình tư duy của thí sinh phản ánh trong phần thi nói bài thi VSTEP.3-5

Chương 6: Sự sắp xếp các yêu cầu về quá trình tư duy của thí sinh khi thi nói bài thi VSTEP.3-5

Chương 7: Quá trình tư duy của thí sinh khi thi bài thi nói VSTEP.3-5 và khi nói tiếng Anh không trong bối cảnh làm bài thi

Chương 8: Kết luận

CHƯƠNG 2: CƠ SỞ LÝ LUẬN

2.1. Khái niệm giá trị xác thực

Thuật ngữ *giá trị xác thực* thu hút sự quan tâm của không chỉ người sử dụng kết quả thi mà cả các nhà nghiên cứu vì thuật ngữ này liên quan nhiều tới chất lượng một bài thi.

Theo Messick (1989), *giá trị xác thực* là “đánh giá tổng thể về mức độ đầy đủ và phù hợp của việc hiểu và sử dụng kết quả thi hoặc kết quả đánh giá khác thể hiện trên một hệ thống những minh chứng và căn cứ khoa học” (S Messick, 1989, p. 13). Khái niệm cốt lõi là “ý nghĩa của kết quả”, được định nghĩa như là “cơ cấu giúp hiểu năng lực của thí sinh thể hiện trên điểm thi và mối quan hệ giữa điểm thi và các biến số khác, cơ cấu này được hiểu như giá trị xác thực của nội dung kiểm tra” (Samuel Messick, 1996, p. 245). Quan điểm này của Messick rất quan trọng bởi đây là một cách hiểu khác về giá trị xác thực so sánh với khái niệm truyền thống của thuật ngữ này.

Quan điểm của Messick đã được Bachman (1990) phát triển. Bachman cho rằng giá trị xác thực của việc sử dụng kết quả kiểm tra là kết quả của một quá trình phức tạp bao gồm “việc phân tích các minh chứng giúp hiểu các kết quả thi, giá trị đạo đức của không chỉ việc sử dụng kết quả thi như thế nào mà còn là năng lực của thí sinh” (Bachman, 1990, p. 237). Tiếp nối thuật ngữ “ma trận lũy tiến” của Messick, Bachman tập trung vào giá trị xác thực của nội dung kiểm tra và “giá trị sử dụng của việc hiểu kết quả thi theo một hướng cụ thể” bằng cách nghiên cứu cơ sở lý thuyết của ngôn ngữ và các giá trị giáo dục và xã hội tương ứng của việc hiểu kết quả thi (Bachman 1990, p. 243). Bachman dựa trên lý thuyết của Messick và kết quả phân tích về “cơ sở của giá trị xác thực” có minh chứng để đưa ra việc thu thập các loại minh chứng cho quá trình xác định giá trị xác thực để giải thích mối liên hệ giữa điểm thi và việc hiểu và sử dụng kết quả này. Ngay khi đưa ra quan điểm về giá trị xác thực được xây dựng trên cơ sở hậu quả của điểm thi, Bachman đã cho rằng bài thi không được thiết kế và sử dụng tách rời khỏi “giá trị lượng hóa của kết quả thi” và cho rằng bài thi phải đáp ứng các nhu cầu của hệ thống giáo dục hoặc của

toàn xã hội nơi mặc định là nơi kết quả thi được sử dụng.

2.2. Quá trình xác định giá trị xác thực của bài thi ngôn ngữ

Trong cuốn *Đo lường trong giáo dục*, Cronbach (1971) định nghĩa “quá trình xác định giá trị xác thực là quá trình nghiên cứu tính chính xác của một dự đoán khoa học hoặc một suy luận từ điểm thi... Theo nghĩa rộng hơn, quá trình xác định giá trị xác thực nghiên cứu độ tin cậy của việc hiểu điểm thi bằng cách mô tả hay giải thích cũng như nghiên cứu sự suy đoán có liên quan đến bối cảnh (Cronbach, 1971, p. 443).

Vào năm 1996, Messick định nghĩa rằng “xác định giá trị xác thực là quá trình đánh giá bằng trải nghiệm ý nghĩa và kết quả của hoạt động đo lường, có tính đến các nhân tố khác trong bối cảnh sử dụng có thể ảnh hưởng tới giá trị xác thực của việc hiểu và sử dụng điểm thi” (Samuel Messick, 1996, p. 245).

Quá trình xác định giá trị xác thực được xây dựng và giới thiệu bởi nhiều học giả trong đó có Cronbach and Meehl (1955), Cronbach (1971), Messick (1989) trong các ấn phẩm được trích dẫn bởi nhiều nghiên cứu về xác định giá trị xác (Cronbach, 1971) (Cronbach & Meehl, 1955) (Loevinger, 1957) (S Messick, 1989).

Trong những năm 2000, mô hình tư duy - xã hội được Weir (2005) giới thiệu cho xây dựng và xác định giá trị xác thực của bài thi ngôn ngữ và đạt được sự chú ý của nhiều học giả. Mô hình xã hội – tư duy giúp các nhà nghiên cứu xây dựng lập luận về giá trị xác thực một cách nhất quán và rõ ràng với hệ thống minh chứng, trong khi vẫn đảm bảo sự tương tác giữa các loại minh chứng. Mô hình này bao gồm các cấu phần khác nhau mà nhà nghiên cứu cần quan tâm tới trong quá trình xây dựng đề thi trong nhiều hơn một khâu của các quá trình xây dựng, thực hiện và xây dựng giá trị xác thực. Các cấu phần này gồm đặc điểm của thí sinh dự thi, giá trị xác thực về quá trình tư duy, bối cảnh, nội dung, so sánh, kết quả.

Trong các mô hình xây dựng giá trị xác thực khác nhau, mô hình so sánh thường được áp dụng với các bài thi khi việc so sánh là khả thi, mô hình nội dung thường được sử dụng cho các bài kiểm tra cuối khóa học, các mô hình giá trị xác thực về nội dung kiểm tra, mô hình tổng hợp, mô hình dựa trên lập luận rất phổ biến và được sử dụng bởi nhiều học giả, mô hình xã hội – tư duy cung cấp cơ sở lý luận vững chắc để xây dựng lập luận về giá trị xác thực một cách tổng thể bao gồm các khía cạnh khác nhau của giá trị xác thực của bài

thi. Mô hình xã hội – tư duy đưa ra mô tả chi tiết về quá trình tư duy mà thí sinh có thể trải qua trong khi thực hiện các bài thi gồm cả 4 kỹ năng trong đó một khía cạnh quan trọng là đặc điểm của thí sinh dự thi. Với những đặc điểm như vậy, mô hình xã hội – tư duy được đánh giá là nỗ lực đầu tiên xây dựng một đường hướng nhất quán, hệ thống để xây dựng đề thi và xác định giá trị xác thực của đề thi, kết hợp các khía cạnh xã hội, tư duy, và đánh giá khi giải quyết vấn đề về việc sử dụng ngôn ngữ và kết nối các khía cạnh đó và các khía cạnh về bối cảnh và kết quả sử dụng đề thi, đặc biệt khi liên quan đến khía cạnh tư duy của một bài thi ngôn ngữ. Cũng với những đặc điểm này, mô hình tư duy – xã hội của Weir đã được lựa chọn sử dụng như khung lý thuyết của nghiên cứu này.

2.3. Khái niệm giá trị xác thực về quá trình tư duy

Trong kiểm tra và đánh giá, khái niệm giá trị xác thực về quá trình tư duy được Weir phát triển trong đường hướng xã hội – tư duy trong xác định giá trị xác thực của một đề thi. Giá trị xác thực của quá trình tư duy theo Weir mô tả tương tự như giá trị truyền thụ trong nghiên cứu của Bachman và Palmer (1996). Thuật ngữ này mô tả mức độ và hình thức tham gia của năng lực ngôn ngữ của thí sinh, kiến thức về chủ đề, và những tác động lên việc làm bài thi ngôn ngữ của thí sinh (Bachman & Palmer, 1996). Một vấn đề quan trọng là quá trình tư duy trong thực tế của thí sinh cần được phản ánh càng nhiều càng tốt trong tình huống thi nếu bài thi này được công bố là có giá trị xác thực về quá trình tư duy. Giá trị xác thực của quá trình tư duy, theo Field (Field, 2013), thu hút được sự quan tâm của các nhà nghiên cứu đặc biệt trong các bài thi có kết quả được sử dụng để dự đoán năng lực của thí sinh phù hợp với một bối cảnh cụ thể như là việc học ở bậc đại học hay khi làm việc cụ thể nào đó.

Những minh chứng về giá trị xác thực của quá trình tư duy thường được tập hợp qua việc nghiên cứu hành vi của thí sinh bằng các hình thức phản ánh lời nói khác nhau (ví dụ như bằng nói ngay khi nghĩ, hồi tưởng ngay sau khi nghĩ, hay hồi tưởng sau một thời gian nhất định) để đưa ra nhận xét thí sinh thực tế đã làm gì trong khi làm bài thi nói (Field, 2011), nghe (Field, 2013), đọc (Khalifa & Weir, 2009), hoặc viết (Shaw, Shaw, & Weir, 2007).

Nhìn chung, theo Field (2013), giá trị xác thực của quá trình tư duy cũng nghiên cứu “các quá trình liên quan được sắp xếp theo các bậc năng lực khác nhau theo mức độ yêu cầu về tư duy mà các bài thi áp lên thí sinh.” Thuật ngữ “yêu cầu về tư duy” sử dụng trong

nghiên cứu này là một thuật ngữ sử dụng trong khoa học tư duy, chỉ yêu cầu đối với khả năng tư duy qua các khía cạnh khác nhau như sự phức tạp, sự rõ ràng, sự hàm ẩn, mức độ trừu tượng (Edwards & Dall'Alba, 1981) hoặc được định nghĩa theo Stein (2009) như “loại và mức độ suy nghĩ cần có để người sinh viên có thể thành công trong việc giải quyết một vấn đề” (Stein, Smith, Henningsen, & Silver, 2009). Cũng được nêu ra bởi Field (2013), giá trị xác thực của quá trình tư duy nghiên cứu mức độ tư duy mà thí sinh có thể trải qua khi làm bài thi. Trong bối cảnh này, thuật ngữ, mức độ tư duy được dùng để chỉ mức độ yêu cầu về tư duy của quá trình xử lý thông tin.

2.4. Giá trị xác thực của quá trình tư duy và việc đánh giá kỹ năng nói

2.4.1. Đánh giá kỹ năng nói

Theo Louma, khi đánh giá kỹ năng nói, các thành tố của năng lực nói cần được nghiên cứu chi tiết. Các nhân tố này là âm thanh, ngữ pháp và khẩu ngữ, từ vựng và ngôn ngữ nói, đặc điểm của sản sinh ngôn ngữ và chứng năng của lời (Luoma, 2004). Việc kiểm tra phát âm (hàm ngôn và ngoại ngôn), khẩu ngữ, và thậm chí là ngôn ngữ xã hội của lời nói là những nội dung cần kiểm tra khi đánh giá kỹ năng nói. Những cấu thành này là nền tảng để thiết kế và phát triển bài kiểm tra kỹ năng nói, hình thành lên nội dung cần kiểm tra.

Lado viết “năng lực nói một ngoại ngữ không nghi ngờ gì là kỹ năng thú vị nhất.. tuy nhiên khảo thí việc nói một ngoại ngữ không nghi ngờ gì được phát triển ít nhất và ít được thực hành trong lĩnh vực khảo thí ngôn ngữ.” (Lado, 1961, p. 239)

Nói là việc sử dụng ngôn ngữ lời nói để giao tiếp với người khác. Theo Fulcher, nội dung kiểm tra trong khảo thí kỹ năng nói gồm đặc điểm âm thanh (phát âm và ngữ điệu, độ chính xác và độ trôi chảy), khía cạnh tâm lý, chiến lược nói (tạo từ và sửa lỗi), khía cạnh cấu trúc lời nói gồm mở và kết và lượt lời, khía cạnh ngữ dụng, từ vựng, và sự tạo lời (Fulcher, 2003).

Field (2004) xây dựng mô hình sản sinh lời nói theo mô hình được Levelt (1989, 1999) (Levelt, 1989; W. Levelt, 1999; W. J. Levelt, 1999), mô hình này làm sáng tỏ rằng dù bất cứ mô hình sản sinh lời nói nào (như ngôn ngữ thứ nhất hay ngôn ngữ thứ 2) đều cần có các giai đoạn sau:

- Giai đoạn khái niệm: đây là giai đoạn các ý niệm lời nói bắt đầu hình thành

- trong đầu người nói
- Giai đoạn hệ thống: là giai đoạn khi người nói chọn một khung phù hợp để tạo lời, trong đó nhấn mạnh sự hòa hợp giữa động từ, v.v.
 - Giai đoạn từ vựng: là khi người nói tìm kiếm từ hoặc từ vựng trong kho từ vựng của mình, quá trình này được hỗ trợ bởi các dấu hiệu trong hình thức như từ hay âm tiết.
 - Giai đoạn tạo âm: đây là giai đoạn thông tin được tập hợp để có thể được chuyển thể thành lời nói
 - Giai đoạn phát âm: đây là giai đoạn khi các đặc điểm của âm thanh được hình thành
 - Giai đoạn tạo lời: ở giai đoạn này, lời nói được hình thành

2.4.2. Giá trị xác thực của quá trình tư duy trong đánh giá kỹ năng nói

Một khía cạnh trong số những khía cạnh khác nhau được mô tả trong mô hình xã hội – tư duy của Weir là giá trị xác thực của quá trình tư duy, theo Weir (2005), giá trị xác thực của quá trình tư duy được hình thành bởi minh chứng giai đoạn trước khi thi thể hiện bởi việc thể hiện quá trình tư duy trong đề thi và minh chứng sau khi thi được thể hiện qua việc phân tích nội dung kiểm tra thể hiện trên điểm thi của thí sinh. “Những người thiết kế đề thi ngôn ngữ cần nắm được lý luận về quá trình tư duy của thí sinh trong đó khẳng định quá trình này cần phản ánh quá trình tư duy trong sử dụng ngôn ngữ trong đời thật” (Taylor, 2011). Dựa trên nghiên cứu của Weir, Field (2011) áp dụng mô hình của Levelt về sản sinh lời nói (1999) để xây dựng khung giá trị xác thực của quá trình tư duy cho kỹ năng nói.

Các quá trình tư duy được nghiên cứu trong mô hình này gồm:

- Hình thành khái niệm: là giai đoạn tạo ý tưởng lời nói
- Mã hóa ngữ pháp: là giai đoạn hình thành khung cú pháp và sắp xếp các âm từ sẽ được sử dụng
- Mã hóa âm thanh: là giai đoạn chuyển những khái niệm trừu tượng của giai đoạn trước thành chuỗi từ và âm thanh
- Mã hóa phát âm: là giai đoạn trật tự âm thanh để việc phát âm được thực hiện dễ dàng; kết nối các âm thành một chuỗi các chỉ dẫn và phát âm; lưu trữ các chỉ dẫn vào khu vực chờ trong khi mệnh đề chuẩn bị được phát âm

- Phát âm: là giai đoạn sản sinh lời nói
- Tự kiểm soát: là giai đoạn người nói tập trung vào thông điệp nói ra ngay trước khi nói và ngay sau khi nói để kiểm tra sự chính xác, rõ ràng và phù hợp của thông điệp được nói ra

CHƯƠNG 3: ĐỀ THI ĐÁNH GIÁ NĂNG LỰC TIẾNG ANH BẬC 3 ĐẾN BẬC 5 THEO KHUNG NĂNG LỰC NGOẠI NGỮ 6 BẬC DÙNG CHO VIỆT NAM VSTEP.3-5

Đề thi VSTEP.3-5 được xây dựng tuân thủ chặt chẽ theo 4 bước phát triển đề thi đưa ra trong Manual for language test development and examining introduced by the Europe Council in 2011 (de l'Europe, 2011), gồm lập kế hoạch, thiết kế, thử nghiệm, và thông báo các bên liên quan. Sản phẩm của các bước thực hiện này là định dạng đề thi VSTEP.3-5, đặc tả kỹ thuật của đề thi VSTEP.3-5, tài liệu hướng dẫn viết tiêu mục đề thi của đề thi, và tài liệu hướng dẫn bồi dưỡng cán bộ chấm thi nói của đề thi VSTEP.3-5.

Thông tin chung của đề thi VSTEP.3-5 được trình bày như sau:

Bài thi	Thời gian	Số câu hỏi/nhiệm vụ bài thi	Dạng câu hỏi/nhiệm vụ bài thi	Mục đích
Nghe hiểu	Khoảng 40 phút, bao gồm thời gian chuyển câu trả lời sang phiếu trả lời.	3 phần, 35 câu hỏi đa lựa chọn (MCQ)	Thí sinh nghe các đoạn trao đổi ngắn, hướng dẫn, thông báo, các đoạn hội thoại và các bài nói chuyện, bài giảng, sau đó trả lời câu hỏi đa lựa chọn (MCQ) đã in sẵn trong đề thi.	Kiểm tra các tiêu kỹ năng Nghe khác nhau, có độ khó từ bậc 3 đến bậc 5: nghe thông tin chi tiết, nghe hiểu thông tin chính, nghe hiểu ý kiến, mục đích của người nói và suy ra từ thông tin trong bài.
Đọc hiểu	60 phút, bao gồm thời gian chuyển câu trả lời sang phiếu trả lời.	4 bài đọc, 40 câu hỏi đa lựa chọn	Thí sinh đọc 4 văn bản về các vấn đề khác nhau, độ khó của văn bản tương đương bậc 3-5 với tổng số từ dao động từ 1900-2050 từ. Thí sinh trả lời các câu hỏi đa lựa chọn sau mỗi bài đọc.	Kiểm tra các tiêu kỹ năng Đọc khác nhau, có độ khó từ bậc 3 đến bậc 5: đọc hiểu thông tin chi tiết, đọc hiểu ý chính, đọc hiểu ý kiến, thái độ của tác giả, suy ra từ thông tin trong bài và đoán nghĩa của từ trong văn cảnh.
Viết	60 phút	2 bài viết	Bài 1: Viết một bức thư/ thư	Kiểm tra kỹ năng Viết

			điện tử có độ dài khoảng 120 từ Bài 1 chiếm 1/3 tổng số điểm của bài thi Viết.	trương tác và Viết sản sinh.
			Bài 2: Thí sinh viết một bài luận khoảng 250 từ về một chủ đề cho sẵn, sử dụng lý do và ví dụ cụ thể để minh họa cho các lập luận. Bài 2 chiếm 2/3 tổng số điểm của bài thi Viết.	

Nói	12 phút	3 phần	Phần 1: <i>Tương tác xã hội</i>	
			Thí sinh trả lời 3-6 câu hỏi về 2 chủ đề khác nhau.	
			Phần 2: <i>Thảo luận giải pháp</i>	
			Thí sinh được cung cấp một tình huống và 3 giải pháp đề xuất. Thí sinh phải đưa ra ý kiến về giải pháp tốt nhất trong 3 giải pháp được đưa ra và phân biện các giải pháp còn lại.	Kiểm tra các kỹ năng Nói khác nhau: tương tác, thảo luận và trình bày một vấn đề.
			Phần 3: <i>Phát triển chủ đề</i>	
			Thí sinh nói về một chủ đề cho sẵn, có thể sử dụng các ý được cung cấp sẵn hoặc tự phát triển ý của riêng mình. Phần 3 kết thúc với một số câu hỏi thảo luận về chủ đề trên.	

Đối với kỹ năng nói, đề thi VSTEP.3-5 phần thi nói được mô tả như sau:

Thời gian: 12 phút (bao gồm 2 phút chuẩn bị, 1 phút cho phần 2 và 1 phút cho phần 3)

Miêu tả chung các phần: Đề thi năng lực tiếng Anh – kỹ năng nói gồm ba phần: (1) Giao tiếp xã hội, (2) Thảo luận giải pháp, (3) Phát triển chủ đề.

Miêu tả chung về nội dung/ngôn ngữ yêu cầu đối với ngôn ngữ sản sinh: diễn ngôn hội thoại (giao tiếp xã hội, thảo luận) và bài trình bày.

Miêu tả mỗi phần:

- Phần 1: Giao tiếp xã hội Trong phần này, giám khảo hỏi thí sinh 3-6 câu hỏi thuộc hai chủ đề khác nhau. Thí sinh trả lời câu hỏi. Các câu hỏi được thiết kế dưới dạng câu hỏi có từ để hỏi. Khi sử dụng câu hỏi nghi vấn, một câu hỏi có từ để hỏi cần được bổ sung để gợi mở cho thí sinh trả lời câu hỏi.

- Phần 2: Thảo luận giải pháp Trong phần này, thí sinh được cung cấp một tình huống với 3 lựa chọn để giải quyết vấn đề nêu ra trong tình huống. Thí sinh đưa ra lập luận về giải pháp thí sinh cho là tối ưu cho vấn đề nêu ra trong tình huống và lập luận tại sao không chọn các giải pháp còn lại. Phần này được thiết kế dưới hình thức một đoạn văn ngắn gồm hai tới ba câu trong đó có 1-2 câu mô tả tình huống và 1 câu đưa ra 3 giải pháp để giải quyết vấn đề nêu ra trong tình huống.

- Phần 3: Phát triển chủ đề Trong phần này, thí sinh trình bày phát triển ý về một chủ đề cho trước. Thí sinh sử dụng dàn ý cho trước dưới dạng sơ đồ tư duy hoặc/và có thể sử dụng dàn ý riêng của mình để phát triển chủ đề. Sau khi thí sinh trình bày xong bài nói, giám khảo hỏi 1-3 câu hỏi liên quan tới nội dung chủ đề. Chủ đề để phát triển ý được trình bày dưới dạng một câu hoàn chỉnh. Dàn ý gợi ý cho thí sinh sử dụng để phát triển ý được trình bày dưới dạng một sơ đồ tư duy. Mỗi ý của dàn ý và chủ đề trong sơ đồ tư duy được để dưới dạng danh từ hoặc cụm danh từ. Mỗi ý của dàn ý cũng có thể được trình bày dưới dạng động từ hoặc tính từ. Ba câu hỏi được ưu tiên thiết kế dưới dạng câu hỏi dùng từ để hỏi. Câu hỏi một được thiết kế cho thí sinh có thể đạt bậc 3 của Khung năng lực ngoại ngữ Việt nam. Câu hỏi hai được thiết kế cho thí sinh có thể đạt bậc 4 của Khung năng lực ngoại ngữ Việt nam. Câu hỏi ba được thiết kế cho thí sinh có thể đạt bậc 5 và bậc cao hơn của Khung năng lực ngoại ngữ Việt nam.

Tổng số tác vụ thi: 3

Tổng điểm: Điểm thô trên tổng điểm 10 được chuyển sang mức điểm tương ứng của thang điểm công bố, trong đó:

Dưới 4 điểm: Dưới bậc B1

4-6 điểm: bậc năng lực B1

6-8 điểm: bậc năng lực B2

8.5-10: bậc năng lực C1

Ví dụ:

Part 1: Social Interaction (3')

Talk about the climate in your area.

- *What is the weather like in your area at this time of the year?*
- *Which season do you like the best? Why?*
- *Do you prefer to live in a cold region or hot region? Why?*

Talk about how your travelling experience.

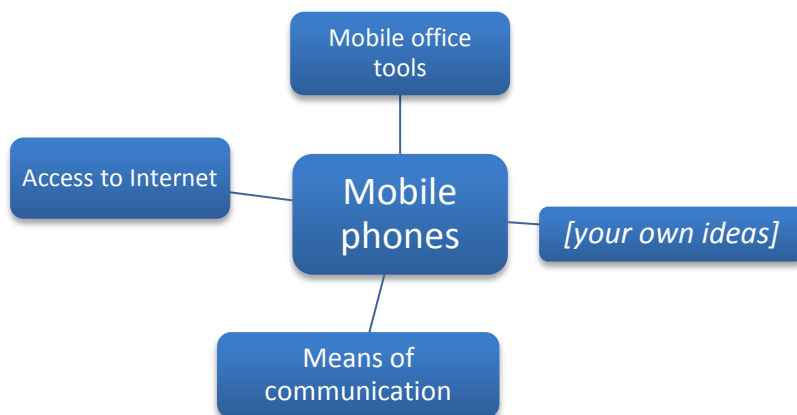
- *What was the last place you traveled to?*
- *Have you ever travelled alone?*
- *Which city in Vietnam do you like the most?*

Part 2: Solution Discussion (4')

Situation: You are having a birthday party and many of your friends are invited. Three locations are suggested: at home, in a restaurant, and in a karaoke bar. Which do you think is the best place for the party?

Part 3: Topic Development (5')

Topic: Mobile phones are useful tools at schools.



- *Do you think people will continue using mobile phones in the future?*
- *What negative effects do you think mobile phones have on young children?*
- *Do young people use mobile phones differently from old people in your country? How?*

CHƯƠNG 4: PHƯƠNG PHÁP NGHIÊN CỨU

4.1. Câu hỏi nghiên cứu

- **CH1:** Bài thi VSTEP.3-5 có phản ánh quá trình tư duy của thí sinh khi thí sinh thi phần thi nói không?

- **CH2:** Các yêu cầu về tư duy được sắp xếp như thế nào trong phần thi nói bài thi VSTEP.3-5 theo từng bậc năng lực tiếng Anh của thí sinh?
- **CH3:** Quá trình tư duy của thí sinh khi thi bài thi nói VSTEP.3-5 có giống với quá trình này khi thí sinh nói tiếng Anh trong bối cảnh không phải bài thi không?

4.2. Phương pháp nghiên cứu

4.2.1. Phương pháp nghiên cứu

Phương pháp nghiên cứu hỗn hợp là phương pháp nghiên cứu được sử dụng phổ biến gần đây trong nghiên cứu ngôn ngữ học. Nghiên cứu này sử dụng cả các phương pháp nghiên cứu định lượng và các phương pháp nghiên cứu định tính trong một nghiên cứu hoặc một loạt các nghiên cứu. Theo Creswell & Clark (2017), “phương pháp nghiên cứu hỗn hợp là nghiên cứu có giả thiết khoa học và nhiều cách thu thập thông tin. Như một phương pháp nghiên cứu, phương pháp này bao gồm các giả thiết khoa học định hướng cho thu thập và phân tích dữ liệu và việc sử dụng hỗn hợp phương pháp nghiên cứu định lượng và định tính trong các khâu của quá trình nghiên cứu. Như một kỹ thuật, phương pháp này tập trung vào thu thập dữ liệu, phân tích và hòa trộn dữ liệu định lượng và định tính trong một nghiên cứu hoặc một loạt các nghiên cứu. Việc kết hợp phương pháp định lượng và định tính cho phép hiểu rõ hơn vấn đề nghiên cứu so với chỉ áp dụng một trong hai phương pháp” (Creswell & Clark, 2017).

Phương pháp nghiên cứu đối chiếu 3 phương pháp và thiết kế lồng ghép được áp dụng cho nghiên cứu này để thu thập những minh chứng về giá trị xác thực của quá trình tư duy theo các bước khác nhau của một chu trình thi.

Trước tiên, bảng phỏng vấn nhóm với 4 chuyên gia xây dựng đề thi, 5 cán bộ viết tác vụ đề thi, 5 cán bộ chấm thi nói, yêu cầu về tư duy của đề thi nói VSTEP.3-5 theo các bước tiền thi của bài thi này. Sau đó, điểm thi và các bậc năng lực của 288 thí sinh thi bài thi VSTEP.3-5 được phân tích để nghiên cứu mức độ yêu cầu về quá trình tư duy của thí sinh được phân bố như thế nào theo các bậc năng lực khác nhau của thang chấp điểm nói VSTEP.3-5. Khảo sát 288 thí sinh này và phỏng vấn sử dụng khuyến khích hồi tưởng 30 thí sinh trong số 288 thí sinh này để nghiên cứu quá trình tư duy của các thí sinh này trong bối cảnh bài thi và bối cảnh ngoài bài thi.

4.3. Công cụ thu thập dữ liệu

Để thu thập được dữ liệu phù hợp với nghiên cứu, các công cụ sử dụng để thu thập dữ liệu gồm định dạng bài thi VSTEP.3-5 và các tài liệu liên quan, câu hỏi phỏng vấn nhóm, câu hỏi khảo sát, câu hỏi phỏng vấn khuyến khích hồi tưởng như được trình bày trong bảng dưới đây:

Câu hỏi nghiên cứu	Công cụ thu thập dữ liệu
CH1: Bài thi VSTEP.3-5 có phản ánh quá trình tư duy của thí sinh khi thí sinh thi phần thi nói không?	Báo cáo xây dựng VSTEP.3-5, định dạng bài thi VSTEP.3-5, đặc tả kỹ thuật đề thi VSTEP.3-5 và các tài liệu liên quan khác Thảo luận nhóm
CH2: Các yêu cầu về tư duy được sắp xếp như thế nào trong phần thi nói bài thi VSTEP.3-5 theo từng bậc năng lực tiếng Anh của thí sinh?	Điểm thi VSTEP.3-5
CH3: Quá trình tư duy của thí sinh khi thi bài thi nói VSTEP.3-5 có giống với quá trình này khi thí sinh nói tiếng Anh trong bối cảnh không phải bài thi không?	Câu hỏi khảo sát Phỏng vấn sử dụng khuyến khích hồi tưởng

4.4. Phân tích dữ liệu

Khía cạnh quan trọng nhất và cũng là khía cạnh khó nhất của nghiên cứu hỗn hợp là tích hợp dữ liệu định lượng và dữ liệu định tính. Theo Creswell & Clark (2017), một đường hướng là phân tích hai nhóm dữ liệu này riêng lẻ và sau đó kết quả nghiên cứu từ hai nhóm dữ liệu này được đối chiếu và phân tích. Dữ liệu cho hai nhóm phương pháp nghiên cứu này được lưu trữ riêng và được phân tích sử dụng kỹ thuật phân tích phù hợp với từng loại; ví dụ, các kỹ thuật về thống kê được sử dụng để phân tích dữ liệu khảo sát trong khi phân tích theo chủ điểm để phân tích dữ liệu phỏng vấn. Theo đường hướng này, việc tích hợp dữ liệu được thực hiện sau khi đạt kết quả nghiên cứu với từng công cụ thu thập dữ liệu. Một đường hướng phân tích dữ liệu hỗn hợp khác là kỹ thuật phân tích

tích hợp. Với phương pháp sau, dữ liệu định lượng được chuyển thành dữ liệu định tính và ngược lại.

Kỹ thuật phân tích tích hợp được sử dụng để phân tích dữ liệu của nghiên cứu này. Trước tiên, dữ liệu định tính từ thảo luận nhóm được sử dụng để phân tích tìm minh chứng cho giá trị xác thực của quá trình tư duy từ các bước xây dựng đề thi trước khi thi. Kết quả này sau đó được đối chiếu với minh chứng thực nghiệm theo kết quả thi của 288 thí sinh. Kết quả khảo sát và kết quả phỏng vấn khuyến khích hồi tưởng được phân tích tích hợp với nhau để xác định quá trình tư duy của thí sinh khi họ thi bài thi nói VSTEP.3-5 và khi họ nói trong bối cảnh không phải thi. Kết quả của các kỹ thuật thu thập dữ liệu khác nhau sau đó được tích hợp để đưa ra kết quả nghiên cứu chung.

Khi phân tích dữ liệu định lượng, điểm thi của 288 thí sinh và số liệu khảo sát được phân tích sử dụng các công cụ gồm EXCEL, SPSS, và FACETS với việc sử dụng các công thức tính gồm thống kê, tương quan, t-test, ANOVA, Cronbach's alpha, phân tích nhân tố và phân tích mô hình Rasch và Rasch nhiều nhân tố. Với dữ liệu định lượng, phân tích hội thoại được áp dụng với việc ghi chép lại toàn bộ dữ liệu thảo luận nhóm và phỏng vấn khuyến khích hồi tưởng. Bằng cách phân tích theo phương pháp này, người nghiên cứu sẽ tránh được dữ liệu thứ cấp, phân tích ví dụ mẫu và kết quả phân tích có thể sử dụng để đưa ra những minh chứng phản biện hỗ trợ cho giả thuyết đã được đưa ra và khuyến khích người nghiên cứu khác sử dụng kết quả nghiên cứu ban đầu này với những nguồn dữ liệu khác (Markee, 2000).

CHƯƠNG 8: KẾT LUẬN

Các chương 5, 6, 7 trình bày kết quả thu thập dữ liệu, phân tích dữ liệu và kết quả phân tích dữ liệu. Trong phạm vi tài liệu tóm tắt này, nội dung ba chương này không được đưa vào vì các nội dung tóm tắt đã được đưa ra trong phần kết luận này.

8.1. Tóm tắt kết quả nghiên cứu

Bằng việc quá dụng khung xã hội – tư duy của Weir dùng cho xây dựng và xác định giá trị xác thực của đề thi ngôn ngữ, quá trình tư duy của thí sinh thi bài thi VSTEP.3-5 được nghiên cứu và đưa ra minh chứng cho phần thi nói. Minh chứng về giá trị xác thực được chia thành ba nhóm gồm quá trình tư duy được lồng ghép vào đề thi, việc sắp xếp mức độ

yêu cầu về tư duy theo từng bậc năng lực khác nhau của đề thi, và sự giống và khác nhau giữa quá trình tư duy của thí sinh khi thi bài thi nói VSTEP.3-5 và khi nói tiếng Anh trong bối cảnh ngoài bài thi.

8.1.1. Quá trình tư duy thí sinh trải qua khi thi bài thi nói VSTEP.3-5

Quá trình tư duy thí sinh có thể trải nghiệm khi thi bài thi nói VSTEP.3-5 được nghiên cứu theo cả hai bước là phát triển đề thi và tổ chức thi. Minh chứng về giá trị xác thực của quá trình tư duy của thí sinh khi thi bài thi nói VSTEP.3-5 được xây dựng dựa trên mô hình sản sinh lời nói của Levelt (1989, 1999) và sau đó được áp dụng bởi Weir (2005) trong mô hình xã hội – tư duy cho phát triển và xây dựng giá trị xác thực của đề thi ngôn ngữ. Mô hình này được chia thành các bước sau: hình thành khái niệm, mã hóa ngữ pháp, mã hóa âm thanh, mã hóa phát âm, phát âm và sửa lỗi.

8.1.1.1. Hình thành khái niệm

Bước hình thành khái niệm của quá trình tư duy trong sản sinh lời nói của thí sinh thi bài thi VSTEP.3-5 được hình thành theo hai nhóm chính gồm cung cấp ý tưởng và tích hợp thông tin thành khung diễn ngôn.

Cung cấp ý tưởng

Thông tin được đưa ra khác nhau từ bậc 3 đến bậc 5 của bài thi nói VSTEP.3-5 với các mức độ thân thuộc giảm dần từ bậc năng lực cao đến bậc năng lực thấp đối với từng thí sinh, điều này đảm bảo mức độ tư duy thí sinh có trình độ năng lực tiếng Anh khác nhau trải nghiệm không làm ảnh hưởng tới việc sản sinh lời nói của thí sinh.

Nhân tố thứ hai có thể ảnh hưởng tới việc cung cấp ý tưởng cho thí sinh thi bài thi nói VSTEP.3-5 là việc tăng mức độ khó của đề thi. Có thể dễ dàng thấy rằng một cách tăng độ khó của bài thi là giảm mức độ hỗ trợ thí sinh khi làm bài thi khi độ khó của bài thi tăng lên. Tuy nhiên, đây không phải là chiến lược được sử dụng với đề thi nói VSTEP.3-5. Mức độ hỗ trợ cho thí sinh thi bài thi nói VSTEP.3-5 được cung cấp để đảm bảo độ tương thích giữa bài làm của thí sinh ở từng trình độ khác nhau vì khái niệm và lĩnh vực từ vựng sử dụng tương đồng nhau. Bên cạnh đó, mức độ hỗ trợ cho thí sinh giúp yêu cầu về tư duy thí sinh trải qua trong bài thi liên quan đến việc hình thành khái niệm không ảnh hưởng đến kết quả bài làm của thí sinh.

Hỗ trợ thí sinh thi nói bài thi VSTEP.3-5 được thể hiện theo nhiều cách. Trước hết là câu hỏi thi, câu hỏi thi nói VSTEP.3-5 bắt đầu từ các câu hỏi có chủ đề quen thuộc đến của Phần 1 đến tình huống của Phần 2 và chủ đề và thông tin hỗ trợ thí sinh trả lời của Phần 3. Nhân tố thứ 2 hỗ trợ thí sinh trong thi phần nói bài thi VSTEP.3-5 là thí sinh có thời gian để chuẩn bị câu trả lời hay không có. Với bài thi nói VSTEP.3-5, thí sinh có 3 phút để chuẩn bị Phần 2 và Phần 3, trong đó 1 phút cho Phần 2 và 2 phút cho Phần 3. Một nhân tố khác ảnh hưởng đến độ khó của ngôn ngữ sử dụng trong câu hỏi thi. Nếu như đa số các câu hỏi thi của bài thi nói VSTEP.3-5 được biên soạn với độ khó của từ vựng và ngữ pháp ở mức độ bậc 3 để thí sinh có năng lực bậc 3, 4, hay 5 đều có thể hiểu nội dung được hỏi, từ đó việc hình thành khái niệm để trả lời câu hỏi của thí sinh không bị ảnh hưởng, hay đây chính là hỗ trợ tới thí sinh để đảm bảo thí sinh có thể làm bài thi. Trong số các câu hỏi thi, chỉ có câu hỏi số 2 và 3 của Phần 3 có độ khó đạt bậc 4 và bậc 5, tuy nhiên, hai câu hỏi này được dành riêng cho các thí sinh có thể đạt được đến năng lực bậc 4 và bậc 5, không kỳ vọng tất cả các thí sinh có thể trả lời được. Các câu hỏi này giúp giám khảo có thể tự tin trong việc xếp bậc năng lực của thí sinh là bậc 4 hay bậc 5 về năng lực ngoại ngữ.

Tích hợp lời nói thành khung diễn ngôn

Trong số các nhân tố Levelt xác định có ảnh hưởng tới quá trình hình thành khái niệm là nhận thức về chủ đề, hình thành chủ đề của thông tin, và nhận thức thông chung và khác nhau với người nghe, hiểu ý và từ dùng của giám khảo hỏi thi và các nguyên tắc cơ bản quyết định thông tin sẽ được sắp xếp như thế nào. Phần thi nói bài thi VSTEP.3-5 nghiên cứu tất cả các yếu tố trên liên quan tới khung thoại của giám khảo, thang đánh giá đối với tiêu chí “discourse management”.

Bài thi nói VSTEP.3-5 sử dụng mô hình một hỏi một. Mặc dù được mô tả trong nội dung kiểm tra, tương tác chỉ được hình thành giữa giám khảo và thí sinh, điều này phần nào làm ảnh hưởng tới tính tương tác tự nhiên của người nói. Một thuận lợi của mô hình một hỏi một là thể hiện của thí sinh trong bài thi không bị ảnh hưởng bởi bài làm của thí sinh khác. Sự thể hiện của thí sinh thay vào đó sẽ bị ảnh hưởng bởi sự hợp tác của giám khảo. Sự hợp tác của giám khảo là cách làm việc cảm thông của giám khảo giúp thí sinh tham gia tương tác tốt hơn với giám khảo. Đặc điểm giọng nói của giám khảo cũng có thể ảnh hưởng tới thí sinh như giọng nói, sự rõ ràng, liền mạch của giám khảo... Bên cạnh đó các

yếu tố ngoại ngôn ngữ trong tương tác với thí sinh cũng có thể giúp trao đổi thông tin thuận lợi hơn. Năng lực chung và năng lực giao tiếp của giám khảo thể hiện ở khả năng giám khảo hiểu các vấn đề đang được quan tâm chung trong xã hội, kiến thức của giám khảo về vấn đề nêu trong tác vụ thi... Các yếu tố trên đều có thể tác động tới mức độ tư duy của thí sinh khi làm bài thi. Vì vậy, phần thi nói của bài thi VSTEP.3-5 được thiết kế sao cho giám khảo sử dụng cùng kịch bản chuẩn bị trước đối với thí sinh thi cùng bộ đề thi. Cũng được mô tả trong đặc tả kỹ thuật bài thi, giám khảo không được nói các nội dung khác ngoài kịch bản đã chuẩn bị sẵn cho họ. Những giám khảo không đảm bảo thực hiện điều này có thể dẫn tới tình huống tạo ra các mức độ tư duy khác nhau cho các thí sinh khác nhau.

8.1.1.2. Mã hóa ngữ pháp

Nội dung ngôn ngữ của các bài kiểm tra nói thường không được quy định về cấu trúc ngữ pháp mà là về các chức năng ngôn ngữ mà người dự thi bắt buộc phải thực hiện. Nó có thể được coi như một câu hỏi ánh xạ từ chức năng mà người dự thi muốn thực hiện đến cách thức thể hiện tốt nhất chức năng đó. Vấn đề này cần được thảo luận khi xem xét giá trị xác thực của quá trình tư duy không phải là độ phức tạp của ngôn ngữ được hình thành mà là mức độ dễ dàng khi thí sinh làm bài thi. Có hai cách để giảm tải yêu cầu về tư duy cho thí sinh ở bậc năng lực thấp hơn với nguồn ngôn ngữ hạn chế. Một nằm ở việc hạn chế số lượng chức năng mà người dự thi phải thực hiện (đặc biệt liên quan đến một nhiệm vụ duy nhất). Phương thức kia tính đến khả năng có thể truy cập của một chức năng nhất định. Các chức năng quen thuộc, thường xuyên và cụ thể rõ ràng sẽ được lập bản đồ nhanh chóng và đáng tin cậy hơn những chức năng không có.

Với bài thi VSTEP.3-5, mức độ yêu cầu về tư duy khác nhau giữa ba tác vụ. Sự đa dạng thể hiện ở số lượng chức năng cần thiết và khả năng tiếp cận của các chức năng được đề cập trong bản đặc tả kỹ thuật đề thi. Số lượng các chức năng khác nhau giữa ba tác vụ. Trong đó, tác vụ 3 chiếm số lượng chức năng nhiều nhất. Có thể thấy rõ sự tiến triển của nhu cầu nhận thức trong ba tác vụ của bài thi nói. Có thể thấy rằng mức độ yêu cầu về tư duy tăng từ tác vụ 1 đến tác vụ 3 của bài thi nói VSTEP.3-5. Ngoài ra còn có sự giảm dần theo chức năng từ nhiệm vụ 1 đến nhiệm vụ 3 theo yêu cầu đặt ra đối với người dự thi: chuyển từ mô tả về các tương tác xã hội sang trình bày một chủ đề phức tạp và trừu tượng ở mức độ trừu tượng cao nhất tóm tắt của câu hỏi số 3 của tác vụ số 3.

8.1.1.3. Mã hóa âm thanh

Một phần của quá trình tư duy khi nói là mã hóa âm thanh, được xác định là bao gồm các đặc điểm sau đây góp phần làm cho việc nói trôi chảy ngôn ngữ thứ hai: sử dụng các cụm từ ghi nhớ sẵn, độ dài của bài nói không bị ngắt quãng, thời gian tạm ngừng ở ranh giới cú pháp, tần suất do dự tạm ngừng, và sự dễ dàng truy xuất và lắp ráp các từ.

Liên quan đến các cụm từ ghi nhớ sẵn, các cụm từ được đề cập trong VSTEP.3-5 dưới dạng các thành ngữ và các từ ngữ thông tục; và được đề cập trong các đặc tả bậc 5 của thang đánh giá (tiêu chí Vocabulary). Người dự thi VSTEP.3-5 đã đạt đến trình độ thông thạo ngôn ngữ nhất định, nơi họ đã thiết lập được kho các từ có thể được lấy ra với nỗ lực tối thiểu. Tuy nhiên, các thành ngữ và từ ngữ thông tục được đề cập trong thang đánh giá khả năng nói của VSTEP như một phần của tiêu chí Vocabulary, không phải là một phần của tiêu chí Fluency. Theo Field (2011), cụm từ không chỉ bao gồm các biểu thức thành ngữ và từ thông tục, mà còn bao gồm các loại ngôn ngữ công thức khác.

Liên quan đến nội dung kiểm tra, có thể thấy rằng các từ ngữ thông tục không được đề cập trong mô tả của các mức độ thông thạo khác nhau. Việc sử dụng các từ ngữ thông tục chỉ được đề cập ở mức độ thông thạo C2 của CEFR. Liên quan đến các cấp độ thông thạo khác được đề cập trong cấu trúc của bài nói VSTEP.3-5, các biểu thức công thức được đề cập từ cấp độ 1 đến cấp độ 5. Với bậc 1, nó là “very short, isolated, mainly pre-packaged utterances”; với bậc 2, đó là “basic sentence patterns with memorized phrases, groups of few words and formulae”; với bậc 3, nó là “a repertoire of frequently used routines”. Mặc dù được đề cập trong nội dung kiểm tra, các đặc tả trên không phải là một phần của thang chấm điểm kỹ năng nói VSTEP.3-5. Vấn đề có lẽ một phần là ngôn ngữ ghi nhớ sẵn có thể vừa là một chỉ báo tiêu cực vừa là một chỉ báo tích cực cho đến nay. Một mặt, nó có thể chỉ ra sự phụ thuộc vào một số công thức hạn chế được quy ước hóa cao, một số công thức học thuộc lòng; mặt khác (như đã thảo luận trước đây), nó có thể cho thấy rằng người dự thi đã tiến bộ như một diễn giả đến mức họ đã thiết lập được một kho các chuỗi từ có thể được tạo ra với nỗ lực tối thiểu.

Một yếu tố khác có thể ảnh hưởng đến sự trôi chảy là độ dài của bài nói không bị ngắt quãng. Mức độ yêu cầu về tư duy bị ảnh hưởng rất nhiều từ độ dài yêu cầu của bài nói. Các đặc tả kỹ thuật của VSTEP.3-5 cho thấy rằng độ dài yêu cầu của bài nói khác nhau giữa ba phần của bài nói. Ngoài ra, các đặc tả của thang đánh giá VSTEP.3-5 cho thấy

cấu trúc của VSTEP.3-5 liên quan đến độ dài của bài nói không bị gián đoạn. Có thể thấy rõ trong thang đánh giá, các đặc tả từ mức điểm 2 đến mức điểm 8 đề cập đến các phản hồi mở rộng và các khoảng ngôn ngữ mở rộng; tuy nhiên, với mức điểm 9 và mức điểm 10, vấn đề độ dài của bài nói được đề cập dưới dạng tạm ngừng để cho thấy rằng người dự thi ở mức độ thông thạo này nói chung nói được đoạn dài.

Khoảng thời gian lập kế hoạch tạm ngừng ở ranh giới cú pháp và tần suất tạm ngừng, do dự cũng được mô tả như một phần của mã hóa âm thanh. Trong thang đánh giá nói VSTEP.3-5, các đặc điểm như vậy cũng được mô tả. Như có thể quan sát thấy trong thang VSTEP.3-5, sự tạm ngừng và ngập ngừng được đề cập từ mức độ thông thạo thấp hơn đến mức cao hơn của thang điểm. Giám khảo cần xác định sự khác biệt giữa tạm ngừng và ngập ngừng vì vốn ngôn ngữ hạn chế do thí sinh tìm kiếm các từ và cụm từ thích hợp và chính xác để sử dụng và ngập ngừng để sửa lỗi hay tìm từ ở mức độ khó hơn phù hợp với bối cảnh.

8.1.1.4. Mã hóa phát âm, phát âm

Levelt (1989) đã chỉ ra rằng một người nói một ngôn ngữ không phải là tiếng mẹ đẻ của mình có thể bị ảnh hưởng bởi ngôn ngữ đầu tiên xét theo một tập hợp các biểu diễn âm vị học trong tâm trí vốn dùng để xác định các giá trị âm vị của ngôn ngữ đầu tiên và một tập hợp các quy trình tự động cao, phù hợp với cài đặt khớp của ngôn ngữ đầu tiên và các chuyển động liên kết giữa chúng với nhau. Sự phù hợp của hai yếu tố này với các giá trị không quen thuộc của ngôn ngữ đích tạo thành cơ sở cho bất kỳ tài mức độ tư duy nào về cách thụ đắc ngôn ngữ thứ 2. Người ta không nên quên thực tế rằng việc phát âm L2 kém không chỉ do không có khả năng hình thành các âm đích mà còn do sự thể hiện không đầy đủ các âm thanh đó trong tâm trí và / hoặc không có khả năng truyền đạt các tín hiệu thích hợp đến việc phát âm. Điều này giải thích rõ việc trình bày yêu cầu đối với mức độ dễ hiểu của cách phát âm của người dự thi VSTEP.3-5 từ không rõ đến dễ hiểu với mức độ dễ hiểu đối với người bản ngữ làm tăng các đặc tả về mức độ thông thạo thấp hơn đến mức độ thông thạo cao hơn. Các mô tả của VSTEP.3-5 cho thấy một số lượng giảm dần từ cấp độ thành thạo cao đến cấp độ thành thạo thấp của ngôn ngữ của các nhân tố phản ánh tác động của các giá trị âm vị được lưu trữ của ngôn ngữ thứ nhất của người dự thi và các quá trình phát âm tự động hóa liên quan đến chúng.

8.1.1.5. Tự kiểm soát

Tự kiểm soát và tự sửa lỗi là các khía cạnh về hiệu suất của người nói mà khó nắm bắt được trong các thông số kỹ thuật của bài kiểm tra biểu mẫu. Như đã đề cập bởi Field (2011), bản chất của việc giám sát và sửa chữa được giới hạn ở mức độ thông thạo thấp hơn của người nói ngôn ngữ thứ 2 và mức năng lực cao hơn có thể được đặc trưng bởi bằng chứng giám sát về hiệu quả thực dụng cũng như độ chính xác về ngôn ngữ. Liên quan đến việc tự kiểm soát và sửa lỗi của thí sinh, các đặc tả nhất định được tìm thấy trong đặc tả kỹ thuật VSTEP.3-5 trong thang đánh giá. Có thể thấy từ các đặc tả của thang đánh giá VSTEP.3-5 rằng mức độ sửa lỗi giảm dần từ mức độ thông thạo thấp đến mức độ thông thạo cao hơn. Bằng chứng về việc giám sát hiệu quả thực dụng đối với độ chính xác ngôn ngữ cũng rõ ràng trong thang đánh giá VSTEP.3-5 trong các đặc tả Vocabulary. Tóm lại, mặc dù mô hình ntur duy - xã hội của Weir không phải là khung lý thuyết được áp dụng khi phần nói VSTEP.3-5 được phát triển, các quy trình tư duy của mô hình cung cấp một khung tốt để nghiên cứu giá trị xác thực của bài thi. Tất cả các mẫu được mô tả trong mô hình có thể được tìm thấy trong phần nói của VSTEP.3-5, cho thấy nỗ lực của nhóm phát triển VSTEP.3-5 trong việc giải quyết các nhu cầu tư duy mà người dự thi có thể gặp phải khi làm bài thi. Bên cạnh đó, khi áp dụng mô hình để thiết lập minh chứng về giá trị xác thực của quá trình tư duy cho phần nói, một số vấn đề đã được xác định bao gồm:

- (1) Việc sử dụng các từ thông tục chỉ được đề cập ở mức độ thông thạo C2 của khung CEFR và CEFR-VN; tuy nhiên, chúng được tìm thấy trong các đặc tả của mức độ thành thạo 9 và 10 của thang đánh giá nói VSTEP.3-5.
- (2) Các điểm tạm ngừng và do dự được tìm thấy trong các đặc tả của hầu hết các nhóm của thang đánh giá VSTEP.3-5; tuy nhiên, không có sự giải thích về sự khác biệt của những khoảng ngừng và ngập ngừng như vậy giữa các mức điểm.
- (3) Tải trọng tư duy của người đối thoại và người đánh giá khi phỏng vấn và đánh giá người dự thi đã được mô tả trong CEFR và CEFR-VN, khung lý thuyết được áp dụng khi thiết kế bài thi nói VSTEP.3-5; tuy nhiên, tải trọng đó không được đề cập trong đặc tả kỹ thuật hoặc sổ tay đào tạo cán bộ chấm thi nói.

8.1.2. Việc sắp xếp các mức độ yêu cầu về quá trình tư duy theo các bậc năng lực khác nhau của bài thi nói VSTEP.3-5

Nhìn chung, các đặc tả của thang đánh giá nói VSTEP.3-5 được sắp xếp hợp lý với mức độ khó tăng dần từ mức điểm 1 đến mức điểm 10, ngoại trừ một số đặc tả như sau:

(1) Khi chạy Mô hình thang đánh giá của phân tích Rasch, các đặc tả theo các mức điểm khác nhau của thang đánh giá nên được quan tâm nhiều hơn bao gồm các mức Vocabulary 4,5, 9,10 và Fluency 6 và 7.

(2) Khi chạy t-test về sự khác biệt giữa điểm nói và điểm tổng trong các nhóm thí sinh thuộc các mức độ thông thạo khác nhau, sự khác biệt đáng kể được tìm thấy giữa điểm nói và điểm tổng của học sinh ở trình độ C1. Kết quả phân tích cho thấy cần nghiên cứu các dải điểm 9 và 10 của tất cả các tiêu chí nói và người chấm thi vẫn đáp phải được thông báo về mẫu điểm này của thang chấm điểm.

(3) Sắp xếp độ khó của các đặc tả trên cùng một biểu đồ, các các đặc tả nhất định nổi bật, trong số đó là các mức điểm Grammar 3, 4, 6, 7; Vocabulary 4, 5, 9, 10; Discourse Management 3, 9, 10; Pronunciation 7, 8; Fluency 6, 7. Đây là các đặc tả có độ khó được sắp xếp vào các vị trí đặc biệt so với độ khó của các đặc tả khác của cùng một mức điểm.

Nhìn chung, các đặc tả liên kế nhau và tương ứng với cùng mức độ thông thạo của người dự thi thì có mức độ khó tương đối như nhau. Bên cạnh đó, khi nghiên cứu kỹ các đặc tả, dường như khó nhận thấy sự khác biệt giữa kết quả hoạt động của thí sinh tương ứng với các điểm lân cận của cùng một tiêu chí của thang đo. Một vấn đề khác được xác định là các đặc tả của các mức điểm Vocabulary 9 và 10 bao gồm yêu cầu về khả năng sử dụng các cách diễn đạt thành ngữ và các từ thông tục, các đặc tả này không được đề cập trong CEFR hoặc CEFR-VN của bậc trình độ năng lực tương ứng. Sau đó, có thể nhận thấy rằng mức điểm 9 và 10 của tất cả các tiêu chí dường như được đặt ở mức yêu cầu tư duy cao hơn so với tất cả các mức điểm khác. Một số mức điểm như điểm từ vựng mô tả năng lực không được đề cập trong nội dung kiểm tra; tuy nhiên, các đặc tả này lại được mô tả trong thanh chấm điểm.

8.1.3. Quá trình tư duy của thí sinh khi làm bài thi nói VSTEP.3-5 và khi nói tiếng Anh trong bối cảnh khác ngoài bài thi

Bảng câu hỏi khảo sát và kết quả phỏng vấn khuyến khích hồi tưởng cho thấy rằng:

(1) Người dự thi đã trải qua tất cả các giai đoạn của quá trình tư duy khi thi bài thi nói VSTEP.3-5. Tất cả năm giai đoạn của quá trình tư duy bao gồm hình thành khái niệm, mã hóa âm thanh, mã hóa ngữ pháp, mã hóa ngữ âm, phát âm và tự kiểm soát.

(2) Hơn một nửa số người dự thi VSTEP.3-5 đã trải qua quá trình tư duy khi nói tương tự giữa điều kiện làm bài thi và điều kiện không phải bài thi. Tất cả năm giai đoạn của quá trình được các thí sinh chia sẻ đã trải qua khi nói tiếng Anh trong bối cảnh ngoài bài thi gồm hình thành khái niệm, mã hóa âm thanh, mã hóa ngữ pháp, mã hóa ngữ âm, phát âm và tự kiểm soát.

(3) Đối với những người nói rằng quy trình tư duy không giống nhau cho rằng chúng có thể nói tốt hơn những gì họ đã thể hiện trong bài thi. Họ tuyên bố rằng thông tin về phần nói VSEP.3-5 bao gồm thang điểm đánh giá phần nói và các bài thi mẫu chỉ được tiếp cận với họ một cách hạn chế. Những thí sinh chuẩn bị tốt hơn cho bài kiểm tra có xu hướng đạt điểm cao hơn so với những người không được chuẩn bị.

(4) Một tình huống của Phần 2 của bài thi khó đối với một nhóm thí sinh so với các nhóm khác. Điều này cho thấy nên thử nghiệm các đề thi một cách kỹ lưỡng hơn.

8.2. Đề xuất đối với bài thi VSTEP/3-5, phát triển và xác định giá trị xác thực đề thi ngôn ngữ, và mô hình xã hội – tư duy của Weir

8.2.1. Phát triển và tổ chức thi bài thi VSTEP.3-5

Mặc dù các bước của quy trình tư duy mà người dự thi có thể trải qua khi làm bài thi VSTEP.3-5 đã được nghiên cứu ở cả giai đoạn phát triển và tổ chức thi của bài thi, một số vấn đề nhất định đã được xác định bao gồm:

(1) Việc sử dụng các từ thông tục chỉ được đề cập ở mức độ thông thạo C2 của khung CEFR, tương đương bậc 6 của khung CEFR-VN; tuy nhiên, những đặc tả này được tìm thấy trong các mức điểm 9 và 10 của thang đánh giá nói VSTEP.3-5. Để người kiểm tra nhận thức rõ về thực tế này, thông tin về mô tả của bậc C2 được sử dụng trong thang chấm điểm nói VSTEP.3-5 đối với bậc 4 nên được bổ sung vào tài liệu bồi dưỡng giám khảo chấm thi. Một cách khác để giải quyết các đặc tả này là đưa các đặc tả này ra khỏi thang đánh giá. Để làm như vậy, các đặc tả kỹ thuật của bài thi VSTEP.3-5 cần được sửa đổi, dẫn đến việc sửa đổi Quyết định số 729 / QĐ-BGDĐT ngày 11 tháng 3 năm 2015 của Bộ Giáo dục và Đào tạo Việt Nam.

(2) Sự ngừng lời và ngập ngừng khi nói được tìm thấy trong các đặc tả của hầu hết các mức điểm của thang đánh giá VSTEP.3-5; tuy nhiên, không có giải thích nào khác được đưa ra về sự khác biệt của những khoảng dừng và ngập ngừng như vậy giữa các mức điểm. Do vậy, cần có giải thích thêm trong tài liệu bồi dưỡng giám khảo chấm thi nói để giám khảo không bị nhầm lẫn về các dấu hiệu ngừng lời hoặc ngập ngừng khi chấm điểm kết quả thi của thí sinh.

(3) Yêu cầu về tư duy của cán bộ chấm thi nói khi hỏi thi và chấm thi đã được mô tả trong khung CEFR và CEFR-VN, khung lý thuyết được áp dụng khi thiết kế bài thi nói VSTEP.3-5; tuy nhiên, tải trọng đó không được đề cập trong đặc tả kỹ thuật hoặc tài liệu đào tạo cán bộ chấm thi. Mức độ tư duy mà cán bộ chấm thi nói trải qua khi phỏng vấn và đánh giá người dự thi có thể tác động đến tính công bằng trong việc phỏng vấn và chấm điểm của họ, và do đó ảnh hưởng đến kết quả thi của người dự thi và điểm số của họ. Chúng tôi đặc biệt khuyến nghị rằng các yếu tố ảnh hưởng đến mức độ tư duy của cán bộ chấm thi vào khi phỏng vấn và đánh giá người dự thi nói VSTEP.3-5 nên được đưa vào tài liệu bồi dưỡng cán bộ chấm thi nói để người chấm thi nhận thức được những vấn đề đó khi phỏng vấn và đánh giá. Một gợi ý cụ thể là cán bộ chấm thi nói hiện tại không được cung cấp thời gian để làm công việc chấm điểm giữa những người dự thi gần nhau. Khi được bổ sung thêm khoảng một phút để chấm điểm giữa hai người dự thi liên kế, các cán bộ chấm thi có thể sẽ chấm thi hiệu quả hơn.

(4) Các đặc tả gần nhau tương ứng với cùng mức độ thông thạo của người dự thi có độ khó tương đối bằng nhau. Bên cạnh đó, khi nghiên cứu kỹ các đặc tả và độ khó của các đặc tả, khó nhận ra sự khác biệt giữa năng lực của thí sinh tương ứng với các mức điểm lân cận của cùng một tiêu chí của thang chấm điểm VSTEP.3-5. Sự khó phân biệt này có thể gây khó khăn cho cán bộ chấm thi khi xếp thí sinh vào đúng các mức điểm tương ứng của thang đánh giá. Để cán bộ chấm thi có thể xếp chính xác kết quả bài nói của thí sinh vào các mức điểm của thang đánh giá, các đặc điểm định lượng của các mức điểm cần được đưa vào tài liệu bồi dưỡng cán bộ chấm thi nói hoặc cần có bài làm mẫu của thí sinh dưới dạng âm thanh và/hoặc video cho tất cả các mức điểm khác nhau đối với từng tiêu chí đánh giá của thang đánh giá. Một cách khác để giải quyết tình huống này có thể là hợp nhất các đặc tả của các dải tương ứng với cùng một bậc trình độ tiếng Anh, làm như vậy, số lượng các mức điểm của thang đánh giá VSTEP.3-5 sẽ giảm xuống

còn 5 mức điểm. Điều này sẽ đơn giản hóa công việc chấm điểm của các giám khảo và có thể sẽ dẫn đến độ tin cậy của kết quả đánh giá tốt hơn.

(5) Tất cả các đặc tả của mức điểm 9 và 10 có độ khó cao hơn đáng kể so với tất cả các mức điểm 8 của thang đánh giá VSTEP.3-5, điều này có thể dẫn đến tỷ lệ thí sinh được xếp loại C1 thấp so với các mức độ thông thạo được đánh giá bằng bài kiểm tra. Để điểm nói đạt mức độ tin cậy hơn đối với những thí sinh đó, các giám khảo phải được cung cấp những dữ liệu này khi họ được tập huấn. Một lý do có thể giải thích cho độ khó cao đáng kể của các đặc tả tương ứng với các mức điểm đó là do các giám khảo có xu hướng ít cho điểm cao nhất cho người dự thi. Để giải quyết vấn đề này, chúng tôi khuyến nghị rằng mức điểm nhất của thang đánh giá phải phản ánh kết quả hoạt động của mức độ thông thạo CEFR C2, bằng cách làm như vậy, giám khảo sẽ thấy rõ kỳ vọng về mức độ thông thạo CEFR C2 và xếp thí sinh vào mức điểm CEFR C1 với độ tin cậy cao hơn.

(6) Các thí sinh cho rằng thông tin về phần thi nói VSEP.3-5 bao gồm thang điểm đánh giá kỹ năng nói và các bài kiểm tra mẫu chỉ được tiếp cận với họ một cách hạn chế. Và những thí sinh chuẩn bị tốt hơn cho bài kiểm tra có xu hướng đạt điểm cao hơn những người không chuẩn bị. Các đơn vị tổ chức thi VSTEP.3-5 nên cung cấp cho người dự thi quyền truy cập không chỉ vào định dạng đề thi, mà còn cả thang đánh giá và các bài làm mẫu, đề thi mẫu để người dự thi hiểu rõ những gì họ mong đợi sẽ thực hiện ở các mức độ thông thạo khác nhau, và vì vậy họ có thể chuẩn bị tốt hơn cho bài thi.

(7) Một tình huống của Phần 2 của bài thi dường như khó đối với một nhóm thí sinh so với các nhóm khác. Điều này cho thấy nên thử nghiệm các đề thi một cách kỹ lưỡng hơn. Chúng tôi đặc biệt khuyến nghị rằng tất cả các đề thi nên được xây dựng theo đúng 12 giai đoạn phát triển đề thi do Bộ Giáo dục và Đào tạo Việt Nam ban hành kèm theo Thông tư 23/2017/TT-BGDĐT, trong đó mỗi bài kiểm tra nên được thi thử trước hai lần. Bên cạnh đó, để điểm số được phân tích đúng với độ tin cậy cao, đặc biệt cho phần thi nói và viết, các ứng dụng như FACETS và R, trên đó có thể chạy Mô hình thang điểm đánh giá và / hoặc Mô hình phân tích nhân tố, nên được sử dụng để phân tích điểm thi.

8.2.2. Công tác phát triển và xây dựng giá trị xác thực đề thi ngôn ngữ

Nghiên cứu xác định giá trị xác thực của quá trình tư duy khi thí sinh thi bài thi nói VSTEP.3-5 lần nữa đã minh chứng rằng mô hình tư duy xã hội của Weir cung cấp một khung lý thuyết toàn diện để phát triển và xây dựng giá trị xác thực cho bài thi ngôn ngữ.

Đặc biệt đối với giá trị xác thực của một khía cạnh nhất định của bài kiểm tra ngôn ngữ như giá trị xác thực về quá trình tư duy, mô hình tư duy xã hội của Weir không chỉ giúp phát triển lập luận xác thực có hệ thống cho một bài kiểm tra ngôn ngữ cụ thể mà còn giúp xác định vấn đề mà bài kiểm tra ngôn ngữ có thể gặp phải ở cả hai các giai đoạn phát triển và xây dựng giá trị xác thực. Do đó, chúng tôi đặc biệt khuyến nghị rằng mô hình tư duy xã hội của Weir nên được áp dụng khi một bài kiểm tra ngôn ngữ mới được phát triển và xác thực.

8.2.3. Weir's socio cognitive framework

Khung nhận thức xã hội của Weir được chứng minh là có tính ứng dụng cao trong việc phát triển và xác thực bài kiểm tra ngôn ngữ, tuy nhiên khi áp dụng khung này trong việc xác thực bài kiểm tra ngoại ngữ như VSTEP.3-5, các đặc điểm cụ thể của việc học tiếng Anh như một ngoại ngữ cần được nghiên cứu kỹ lưỡng để có thể được áp dụng tốt hơn. Trong trường hợp về giá trị xác thực, các quá trình tư duy mà người dự thi có thể trải qua khi làm bài kiểm tra tuân theo các quy trình tư duy của Levelt (1988, 1989) về sản sinh lời nói cho người nói ngôn ngữ thứ nhất và quy trình tư duy của Weir (2005) về sản sinh lời nói cho người nói ngôn ngữ thứ hai nên hai mô hình này gần như giống nhau. Tuy nhiên, khi nghiên cứu tính xác thực về quá trình tư duy của bài thi nói VSTEP.3-5, cần thiết phải phát triển một mô hình tư duy cập nhật cho quá trình sản sinh lời nói của những người nói một ngôn ngữ như ngoại ngữ, vấn đề này được nảy phát sinh sau thực tế là một số người dự thi bài thi nói VSTEP.3-5 đã suy nghĩ bằng tiếng Việt, ngôn ngữ thứ nhất của họ. Việc suy nghĩ bằng ngôn ngữ thứ nhất khi nói một ngoại ngữ ảnh hưởng nhất định đến khả năng nói của họ khi làm bài thi và có thể cả trong điều kiện không làm bài thi, đặc biệt trong trường hợp của người dự thi có trình độ CEFR B1 trở xuống.

8.3. Limitations and further studies

Đầu tiên, nhà nghiên cứu khuyến nghị rằng Quyết định số 729/QĐ-BGDĐT ngày 11 tháng 3 năm 2015 do Bộ Giáo dục và Đào tạo Việt Nam ban hành có thể cần được sửa đổi liên quan đến các đặc tả kỹ thuật của bài thi VSTEP.3-5. Để thuyết phục các nhà nghiên cứu và các nhà hoạch định chính sách thay đổi Quyết định như vậy, cần thiết phải tiến hành các nghiên cứu sâu hơn để đưa ra minh chứng rõ ràng hơn về sự cần thiết phải sửa đổi bản đặc tả kỹ thuật của bài kiểm tra này và phần nào của đặc tả kỹ thuật này cần được sửa đổi.

Thứ hai, nghiên cứu ở đây được thực hiện trên quy mô thu thập dữ liệu nhỏ (288 người dự thi). Để dữ liệu thu thập được có giá trị và mức độ tin cậy cao hơn, các nghiên cứu tương tự với quy mô thu thập dữ liệu lớn hơn được khuyến khích tiến hành.

Thứ ba, hiện tại, khung nhận thức xã hội của Weir được chứng minh là có tính ứng dụng cao để thực hiện các nghiên cứu về giá trị xác thực; tuy nhiên, mô hình xử lý nhận thức trong sản sinh lời nói hiện tại chỉ được phát triển cho người nói ngôn ngữ thứ nhất và ngôn ngữ thứ 2. Một mô hình có xem xét đến các đặc điểm của việc thụ đắc ngoại ngữ nên được phát triển để cung cấp mô hình hoàn hảo hơn để thiết lập giá trị xác thực của một bài kiểm tra ngoại ngữ, cụ thể là khi bài kiểm tra được thiết kế để đánh giá người dự thi có năng lực CEFR B1 và các cấp độ thông thạo thấp hơn. Do đó, các nghiên cứu tiếp theo có thể được tiến hành để xây dựng mô hình tư duy cho người nói ngoại ngữ, điều này có thể khả thi bằng cách áp dụng mô hình của Levelt cho người nói L1 hoặc một mô hình như vậy có thể được phát triển mới vì cho đến nay mô hình nói L1 của Levelt là mô hình duy nhất có thể áp dụng trong việc phát triển và xây dựng giá trị xác thực cho đề kiểm tra ngôn ngữ.

Cuối cùng, các ý kiến phê bình và đóng góp cho nghiên cứu này được đánh giá cao.

THESIS RELATED PUBLICATIONS

1. Nguyen Thi Mai Huu (2019). *Stimulated recall – A practical data collection technique for cognitive studies in language testing*. Journal of Foreign Languages Studies, 58/2019, 38-50.
2. Nguyen Thi Mai Huu (2020). *Application of the socio-cognitive framework in language test development and validation*. 2020 International graduate research symposium ND 10TH East Asia Chinese teaching forum, Volume 1, 727-735.
3. Phuong Tran, Hoa Nguyen, Trang Dang, Minh Nguyen, Lan Nguyen, Tuan Huynh, Ha Do, Huu Nguyen, Fred Davidson (2015). *A validation study on the newly-developed Vietnam standardized English proficiency test*. 37th Language Testing Research Colloquium, 142.
4. Victoria Clark, Nguyen Thi Mai Huu, Jessica Wu, Jamie Dunlea, Richard West (2016). *Test centers and standardized Testing: Challenges, issues and benefits*. 4th British Council New Directions in English language assessment conference, 19-20.
5. Nathan Carr, Quynh Nguyen, Huu Nguyen, Yen Nguyen, Thao Nguyen (2016). *Systematic support for a communicative standardized proficiency test in Vietnam*. 4th British Council New Directions in English language assessment conference, 26-27.
6. Huu Nguyen (2018). *An Investigation into the Cognitive Validity of the Speaking Section of the Vietnam's Standardized Test of English Proficiency*. 40th Language Testing Research Colloquium, 114.
7. Huu Nguyen (2019). *An investigation into the cognitive processes reflected in the VSTEP speaking scoring rubrics*. 7th British Council New Directions in English language assessment conference, 22.
8. Barry O'Sullivan, Josep Lo Bianco, Huu Nguyen, Mitsuharu Ota, Yoshinori Watanabe (2019). *Language Assesment Policy*. 7th British Council New Directions in English language assessment conference, 18.