



การพัฒนาแบบทดสอบ MMPI-2-RF ฉบับภาษาไทย: การแปลและการทดสอบแบบวัดซ้ำกับผู้ที่ชำนาญสองภาษา MMPI-2-RF Thai Translation Development: A Translation and Bilingual Test-Retest Study

วรวัชร ด้งจิตจรเจริญ*, บุรชัย อัศวทวิบุญ*, ชลาลัย แต่ศลปะสาดิต*, ดำรงค์ อดุลยฤทธิกุล**,
พนิดา ชิมารณ**, สวตริ แฮมมอนด**, ยพิน โภคสิทธิตยกุล**, สุชีรา ภัทรายุตวรอน***
Worawach Tungjitcharoen*, Burachai Asawathaweeboon*,
Chalalai Taesilapasatit*, Dumrong Adunyarittigun**, Phanida Simaporn**,
Sawitri Hammond**, Yupin Pokhithiyuk**, Sucheera Phattharayuttawat***

* สาขาจิตวิทยา คณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์

** สาขาวิชาภาษาอังกฤษ คณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์

*** ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล

* Department of Psychology, Faculty of Liberal Arts, Thammasat University

** Department of English, Faculty of Liberal Arts, Thammasat University

*** Department of Psychiatry, Faculty of Medicine, Siriraj Hospital, Mahidol University

บทคัดย่อ

วัตถุประสงค์ เพื่อแปล และตรวจสอบคุณภาพของแบบทดสอบด้วยการทดสอบแบบวัดซ้ำในผู้ที่
ชำนาญสองภาษา

วิธีการศึกษา การแปลแบบทดสอบใช้การกระบวนการแปลย้อนกลับเพื่อตรวจสอบคุณภาพเบื้องต้น
หลังจากนั้นทำการหาค่าสหสัมพันธ์แบบวัดซ้ำกับผู้ชำนาญสองภาษาเพื่อเปรียบเทียบค่าความเชื่อมั่น
ที่ได้จากการทดสอบข้ามภาษากับค่าความเชื่อมั่นจากการวัดซ้ำในภาษาอังกฤษจากต้นฉบับ โดยมีผู้ตอบ
แบบทดสอบครบทั้งสองครั้งทั้งสิ้นจำนวน 40 ราย ผู้ตอบแบบทดสอบทั้งหมดเป็นผู้ที่มีความชำนาญ
ทั้งภาษาไทยและภาษาอังกฤษ

ผลการศึกษา ผลที่ได้จากการแปลย้อนกลับได้ความใกล้เคียงกับต้นฉบับในระดับน่าพอใจ ผลการ
วิเคราะห์ข้อมูลจากการศึกษาแบบวัดซ้ำพบว่า ในสเกลทั้งหมด 51 สเกล มีสเกลที่ค่าสหสัมพันธ์ในกรณี
วัดซ้ำข้ามภาษาและวัดซ้ำในภาษาอังกฤษในคู่มือไม่มีความแตกต่างกันอย่างมีนัยสำคัญจำนวนทั้งสิ้น
36 สเกล และมีสเกลที่ค่าสหสัมพันธ์ในสองกรณีนี้มีความแตกต่างกันจำนวน 15 สเกล โดยสเกลที่ค่า
สหสัมพันธ์แบบวัดซ้ำกรณีทดสอบไทย-อังกฤษสูงกว่าค่าสหสัมพันธ์แบบวัดซ้ำในภาษาอังกฤษนั้นมีอยู่
3 สเกล ส่วนสเกลที่ค่าสหสัมพันธ์แบบวัดซ้ำกรณีทดสอบไทย-อังกฤษต่ำกว่าค่าสหสัมพันธ์แบบวัดซ้ำ
ฉบับในภาษาอังกฤษนั้นมีอยู่ 12 สเกล เมื่อทำการเปรียบเทียบรายข้อพบข้อคำถามที่มีความแตกต่าง
รายข้อในระดับที่ต้องนำมาพิจารณาจำนวน 41 ข้อ ผู้วิจัยได้ทำการปรับปรุงข้อคำถามจำนวนทั้งสิ้น
18 ข้อ ซึ่งเป็นข้อคำถามที่มีความสัมพันธ์กับสเกลที่มีค่าสหสัมพันธ์ต่ำ

สรุป แบบทดสอบ MMPI-2-RF ฉบับแปลมีคุณภาพเหมาะสมและได้รับการอนุมัติไปใช้เก็บค่า
บรรทัดฐานจากประชากรไทยได้

คำสำคัญ MMPI-2-RF การแปลแบบทดสอบทางจิตวิทยา การศึกษาแบบวัดซ้ำ

Corresponding author: วรวัชร ด้งจิตจรเจริญ

วารสารสมาคมจิตแพทย์แห่งประเทศไทย 2560; 62(4): 349-358

ABSTRACT

Objectives : To translate MMPI-2-RF into Thai, and to assess its quality using a test-retest study.

Materials and Methods : Back-translation technique is used for the preliminary adequacy assessment of the translation. We conducted a test-retest study to compare the correlations of each scale from the test and the test-retest reliability coefficients provided in the MMPI-2-RF Technical Manual. The test was counterbalanced. 40 bilinguals completed the test.

Results : The translation adequacy is acceptable. The analyses found no significant difference in 36 scales from a total of 51 scales in MMPI-2-RF. In 15 scales that showed significant differences, the cross-language Thai-English correlations were significantly higher than the correlations in the Technical Manual in 3 scales, whereas the cross-language Thai-English correlations were significantly lower than the correlations in the Technical Manual in 12 scales. When comparing item-level discrepancies, the analyses found 41 items with high discrepancies. We examined these items and made minor adjustments to 18 items that related to scales with low correlations. The reviewer at the University of Minnesota approved the analyses and the translation for further uses in the next phase.

Conclusion : This Thai version of MMPI-2-RF is adequacy in terms of its psychometric properties and is approved for the collection of normative data in the Thai population.

Keywords : MMPI-2-RF, psychological assessment translation, test-retest study

Corresponding author: Worawach Tungjitcharoen

J Psychiatr Assoc Thailand 2017; 62(4): 349-358

บทนำ

นับตั้งแต่ปี ค.ศ. 1940 เป็นต้นมา แบบทดสอบ MMPI (The Minnesota Multiphasic Personality Inventory) ได้ถูกนำมาใช้ในฐานะของแบบทดสอบบุคลิกภาพในทางคลินิกและจากสำรวจการใช้แบบทดสอบทางจิตวิทยา แบบทดสอบ MMPI ถือว่าเป็นแบบทดสอบทางจิตวิทยาที่มีการใช้มากที่สุดในประเทศสหรัฐอเมริกา รองจากแบบทดสอบวัดเชาวน์ปัญญาในผู้ใหญ่ของเวคสเลอร์ (Wechsler Adult Intelligence Scale)^{1,4,10}

จนถึงปัจจุบันแบบทดสอบ MMPI ได้ถูกพัฒนาขึ้นหลายครั้งและปรับปรุงข้อคำถามรวมถึงวิธีการคิดคะแนน แบบทดสอบ MMPI ฉบับดั้งเดิมนั้นใช้วิธีการเชิงประจักษ์ในการสร้างแบบทดสอบ โดยการทำการสร้างข้อคำถามจำนวนมากและนำข้อคำถามเหล่านี้ไปวัดกับกลุ่มตัวอย่างจากผู้ป่วยที่ถูกวินิจฉัยว่ามีอาการผิดปกติทางจิตในรูปแบบต่างๆ และบุคคลธรรมดาเพื่อเปรียบเทียบความแตกต่างบุคคลสองกลุ่มนี้ ได้ข้อคำถามทั้งสิ้นจำนวน 383 ข้อ โดยเริ่มแรก สเกลในการวัดประกอบไปด้วย clinical scales จำนวน 8 สเกล อย่างไรก็ตามการใช้ clinical scales ในการวินิจฉัยและเป็นตัวทำนายผล (predictor) ในช่วงแรกนั้นกลับไม่มีความเที่ยงตรงเท่าใดนัก⁹แม้กระทั่งมีการพบว่าเพื่อพิจารณา clinical scales หลายๆ ตัวประกอบกันมีความสัมพันธ์กับคุณลักษณะทางบุคลิกภาพบางประการ ผู้ใช้และผู้วิจัยจึงหันมาสนใจลักษณะของคะแนนในสเกลต่างๆ ประกอบกัน เป็นตัวชี้วัดความผิดปกติทางจิต การใช้ลักษณะของคะแนนในสเกลต่างๆ นี้ถูกเรียกว่า profile type หรือ code type โดยการแปลผลส่วนใหญ่เกิดจากข้อมูลงานวิจัยเชิงประจักษ์เกี่ยวกับลักษณะการกระจายของค่าในสเกลหลายตัวประกอบกันและอาการทางจิตแบบต่างๆ^{6,11,12}

ต่อมาได้มีการพัฒนามาเป็นแบบทดสอบ MMPI-2 และ MMPI-A (ใช้ในวัยรุ่นอายุระหว่าง 14-18 ปี) โดยมีเป้าหมายสำคัญเพื่อการเก็บข้อมูลจากกลุ่มตัวอย่างบรรทัดฐาน (normative sample) ใหม่ และเพื่อแก้ไขและปรับปรุงข้อคำถาม

ต่อมามีการพัฒนา RC Scales (restructured clinical scales) ซึ่งเป็นสเกลหลักตัวใหม่ เนื่องจากสเกลต่างๆ ใน clinical scales ของ MMPI-2 นั้นมีความเที่ยงตรงเชิงบรรจบ (convergent validity) มาก ข้อคำถามแต่ละข้อวัดและสะท้อนค่าในสเกลต่างๆ จำนวนหลายสเกล² ดังนั้น RC scales จึงถูกพัฒนาขึ้นเพื่อลดปัญหาความหลากหลายในเชิงโครงสร้าง (structural heterogeneity) ของข้อคำถาม โดย RC scales ที่ผ่านกระบวนการพัฒนาจำนวน 9 สเกลนั้นได้ถูกนำมาใช้ควบคู่กับ MMPI-2 ในระยะแรก ต่อมาในปี ค.ศ. 2008 แบบทดสอบ MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2-Restructured Form) ได้รับการตีพิมพ์เพื่อนำมาใช้แทนที่แบบทดสอบ MMPI-2 ต่อไป โดยการพัฒนา MMPI-2-RF นี้ เป็นการนำข้อคำถามจาก MMPI-2 มาปรับปรุงแก้ไขและตัดข้อคำถามบางข้อที่ไม่สอดคล้องกับการวินิจฉัยในปัจจุบัน พร้อมทั้งการปรับโครงสร้างของสเกลต่างๆ ยกเลิกการใช้ clinical scales เดิม และใช้ RC scales แทน ซึ่งข้อดีของแบบทดสอบฉบับใหม่นี้คือ ส่วนของการแปลผลที่ง่ายและแม่นยำมากขึ้น⁷

ในขณะที่การศึกษาที่เกี่ยวข้องกับแบบทดสอบ MMPI ได้ถูกตีพิมพ์ทั่วโลกมากกว่า 10,000 บทความ⁸ แต่เมื่อสำรวจการศึกษาในประเทศไทยด้วยคำค้นหาคำว่า “MMPI” “Minnesota Multiphasic Personality Inventory” และ “มินเนโซต้า” พบข้อมูลที่เกี่ยวข้องจากฐานข้อมูลงานวิจัย (โครงการเครือข่ายห้องสมุดในประเทศไทย) ฐานข้อมูลวิทยานิพนธ์อิเล็กทรอนิกส์ ฐานข้อมูลวิจัยทางสุขภาพจิตและจิตเวชของกรมสุขภาพจิต

ข้อมูลจากห้องสมุดกลางโรงพยาบาลตำรวจ และในวารสารสมาคมจิตแพทย์แห่งประเทศไทย พบการศึกษาที่เกี่ยวข้องทั้งหมดเพียง 11 งานวิจัยเท่านั้น โดยการศึกษาส่วนใหญ่ยังคงใช้แบบทดสอบ MMPI ฉบับแรก และในการแปลผลยังคงอ้างอิงกับค่าคะแนนมาตรฐานจากประเทศสหรัฐอเมริกา

นอกจากนี้ นับตั้งแต่ประเทศไทยใช้แบบทดสอบ MMPI มาจนถึงปัจจุบัน ยังไม่ปรากฏว่ามีการศึกษาเพื่อเก็บข้อมูลคะแนนมาตรฐานของประชากรไทย มีเพียงการสำรวจในกลุ่มประชากรบางกลุ่มเท่านั้น ส่งผลให้เกิดปัญหาด้านความน่าเชื่อถือในการนำแบบทดสอบ MMPI มาใช้ในการศึกษาหรือในทางคลินิก

ในการศึกษาคั้งนี้จึงมีวัตถุประสงค์เพื่อพัฒนาแบบทดสอบ MMPI-2-RF ฉบับภาษาไทยในขั้นตอนการแปลภาษาอังกฤษต้นฉบับมาเป็นภาษาไทยและการหาค่าคุณภาพของแบบทดสอบ ซึ่งจะทำได้แบบทดสอบ MMPI-2-RF ฉบับภาษาไทยที่พร้อมจะนำไปใช้ในการเก็บตัวอย่างเพื่อสร้างข้อมูลบรรทัดฐานต่อไป

วิธีการศึกษา

แบบทดสอบ MMPI-2-RF เป็นแบบทดสอบในตระกูลแบบทดสอบ MMPI ฉบับล่าสุด มีวัตถุประสงค์หลักเพื่อใช้ในการตรวจวัดความผิดปกติทางจิตในหลายๆ กลุ่มอาการ ใช้กับบุคคลที่มีอายุตั้งแต่ 18 ปีขึ้นไป มีความสามารถทางการอ่านเขียนในระดับ grade 6 ขึ้นไป มีสเกลรวมทั้งสิ้นถึง 51 สเกล สามารถแบ่งเป็นกลุ่มได้ 6 กลุ่ม ดังนี้ กลุ่มสเกลที่ใช้ในการตรวจสอบความเที่ยงตรงของการวัด (validity scales) จำนวน 9 สเกล กลุ่มสเกลที่ใช้วัดความผิดปกติกลุ่มใหญ่ (higher-order scales) จำนวน 3 สเกล กลุ่มสเกลหลักทางคลินิก (restructured clinical scales: RC scales) จำนวน 9 สเกล กลุ่มสเกลวัดปัญหาเฉพาะด้าน 23 สเกล กลุ่มสเกลวัดความสนใจ (interest scales) 2 สเกล และ

กลุ่มสเกลอ้างอิงตามแนวคิดพยาธิสภาพทางจิตหลัก 5 ด้าน อีกจำนวน 5 สเกล (personality psychopathology five (PSY-5) scales) แบบทดสอบเป็นแบบเลือกตอบจริง-เท็จ มีจำนวนข้อคำถามทั้งสิ้น 338 ข้อ คะแนนจากข้อคำถามแต่ละข้ออาจมีความสัมพันธ์กับคะแนนในสเกลมากกว่า 1 สเกล²

แบบทดสอบ MMPI-2-RF นี้มีเจ้าของลิขสิทธิ์คือ มหาวิทยาลัยแห่งมินนิโซตา (University of Minnesota) ประเทศสหรัฐอเมริกา การพัฒนาแบบทดสอบในทุกขั้นตอนจะต้องได้รับการอนุมัติโดยเจ้าของลิขสิทธิ์ และมีการตรวจสอบและรายงานผลในทุกๆ ขั้นตอนการวิจัย ซึ่งทางคณะผู้วิจัยได้ทำบันทึกข้อตกลงกับทางเจ้าของลิขสิทธิ์เป็นที่เรียบร้อยแล้ว

ขั้นตอนหลักของการวิจัยเริ่มจากการแปลแบบทดสอบจากฉบับมาตรฐานภาษาอังกฤษให้เป็นภาษาไทย จากนั้นจึงทำการแปลย้อนกลับ (back translation) แล้วนำผลมาเปรียบเทียบเพื่อหาความแตกต่างและปรับปรุงแก้ไขข้อคำถาม จากนั้นผู้วิจัยจะต้องรายงานผลการแปลและส่งให้หน่วยงานบริการด้านภาษาของเจ้าของลิขสิทธิ์ตรวจสอบความถูกต้องอีกครั้งหนึ่ง

ภายหลังการอนุมัติการแปลแล้วจึงทำการทดสอบแบบวัดซ้ำข้ามภาษากับผู้ที่มีความชำนาญสองภาษา (bilingual test-retest study) เพื่อตรวจสอบคุณภาพของแบบทดสอบฉบับแปล โดยกลุ่มตัวอย่างในการทดสอบแบบวัดซ้ำนี้เป็นผู้ที่มีความชำนาญทั้งในภาษาไทยและภาษาอังกฤษ โดยอาศัยเกณฑ์ประกอบด้วย คะแนนจากการทดสอบภาษาอังกฤษ โดยมีคะแนนการสอบภาษาอังกฤษมาตรฐาน TOEFL 100 ขึ้นไป IELTS 7.0 ขึ้นไป หรือเทียบเคียง มีพื้นฐานทางการศึกษาและการดำรงชีวิตที่มีการใช้ชีวิตในสภาพแวดล้อมที่มีการใช้ภาษาอังกฤษตั้งแต่เด็ก และการสัมภาษณ์ ซึ่งอาสาสมัครจะต้องสนทนาได้อย่างเป็นธรรมชาติ โดยกำหนดกลุ่มตัวอย่างมีขนาด 50 ราย โดย

ผู้รับการทดสอบจะต้องรับการทดสอบทั้งสองภาษาในช่วงที่แยกจากกันเป็นระยะเวลา 2 สัปดาห์ และทำการถ่วงสมดุล (counterbalancing) ดังนั้นกลุ่มตัวอย่างครึ่งหนึ่งจะได้ทำฉบับภาษาอังกฤษก่อน และกลุ่มตัวอย่างอีกกลุ่มหนึ่งจะได้ทำฉบับแปลก่อน

ค่าที่ได้จากการทดสอบแบบวัดซ้ำข้ามภาษานี้ จะถูกนำไปประเมินความแตกต่างของคะแนนระดับสเกล และความแตกต่างในระดับข้อคำถาม โดยฐานในการประเมินความแตกต่างระดับสเกลเป็นค่าสหสัมพันธ์ในการทดสอบวัดซ้ำ (test-retest correlation) ด้วย Pearson's correlation coefficient ที่ทำในกลุ่มบรรทัดฐาน ซึ่งแสดงในคู่มือฉบับของสหรัฐอเมริกาและเอกสารสนับสนุนอื่นๆ ส่วนฐานในการประเมินความแตกต่างระดับข้อคำถามให้ประเมินตามความเหมาะสม ไม่มีเกณฑ์ประเมินชัดเจนที่กำหนดมาโดยให้ทางคณะผู้วิจัยวิเคราะห์และกำหนดเกณฑ์เอง

การศึกษานี้ผ่านการอนุมัติด้านจริยธรรมการวิจัยในคน จากคณะอนุกรรมการจริยธรรมการวิจัยในคน มหาวิทยาลัยธรรมศาสตร์ ชุดที่ 2

ผลการศึกษา

คุณภาพการแปล

การวิจัยในส่วนของ การแปลในเบื้องต้นซึ่งมีการแปลเป็นภาษาไทย และทำการแปลย้อนกลับนั้น ได้รับการตรวจสอบคุณภาพโดยหน่วยงานบริการด้านภาษาของมหาวิทยาลัยมหินนิโชตา ได้ผลอยู่ในระดับ ยอมรับได้ (acceptable) ซึ่งเป็นระดับสูงสุดลำดับที่สองจากระดับคุณภาพ 5 ลำดับ อย่างไรก็ตามหน่วยงานบริการด้านภาษาได้นำเสนอข้อปรับปรุงสำหรับข้อคำถามในบางข้อ ซึ่งผู้วิจัยได้ปรับปรุงแก้ไขตามข้อเสนอที่ให้มา

ผลการวิเคราะห์ระดับสเกล

ในส่วนของการทดสอบแบบวัดซ้ำ จากผู้เข้าร่วมงานวิจัยทั้งหมด 50 ราย ปรากฏว่า ผู้วิจัยจำเป็นจะต้องลดขนาดกลุ่มตัวอย่างเหลือเพียง 40 ราย เนื่องจากมี

ผู้เข้าร่วมการวิจัยจำนวน 10 ราย ไม่ได้มาทำแบบทดสอบตามกำหนดจำนวน 2 ครั้ง

ผลการทดสอบค่าสหสัมพันธ์แบบวัดซ้ำด้วย Pearson's correlation coefficient ในระดับสเกลที่ได้จากการวิจัยซึ่งเป็นการวัดซ้ำแบบข้ามภาษาไทย-อังกฤษ ($r_{\text{Thai-English}}$) เมื่อนำมาเปรียบเทียบกับข้อมูลค่าความเชื่อมั่นแบบวัดซ้ำในคู่มือทางเทคนิค³ ($r_{\text{US sample}}$) ซึ่งมีกลุ่มตัวอย่างที่ใช้ในการวัดซ้ำจำนวน 193 ราย ด้วยการแปลงคะแนนค่า r ให้เป็นค่า z ด้วย Fisher's z transformation จากนั้นนำมาทดสอบความแตกต่างด้วยการทดสอบค่า z เป็นตามตารางที่ 1

จากตารางที่ 1 จะพบว่าค่าความแตกต่างระหว่างค่าสหสัมพันธ์แบบวัดซ้ำกรณีทดสอบไทย-อังกฤษกับค่าความเชื่อมั่นแบบวัดซ้ำในภาษาอังกฤษฉบับมาตรฐานที่ปรากฏในคู่มือทางเทคนิค³ ซึ่งเป็นค่าสหสัมพันธ์จากการวัดซ้ำอังกฤษ-อังกฤษมีความแตกต่างกันในหลายสเกล ซึ่งในบางสเกลค่าสหสัมพันธ์แบบวัดซ้ำกรณีทดสอบไทย-อังกฤษจะมีค่าสูงกว่า แต่ในบางสเกล ค่าสหสัมพันธ์แบบวัดซ้ำฉบับมาตรฐานอังกฤษ-อังกฤษจะมีค่าสูงกว่า เมื่อทำการเปรียบเทียบความแตกต่างของค่าสหสัมพันธ์แบบ Pearson พบว่า มีสเกลจำนวน 15 สเกลที่มีความแตกต่างของค่าความเชื่อมั่นอย่างมีนัยสำคัญทางสถิติ อันได้แก่ L-r K-r THD RCd RC4 RC6 GIC SUI NFC STW MSF SUB AGG DISC-r และ NEGE-r

โดยเมื่อพิจารณาจากค่าความแตกต่างระหว่างค่าสหสัมพันธ์แล้วพบว่า สเกลที่ค่าสหสัมพันธ์แบบวัดซ้ำกรณีทดสอบไทย-อังกฤษสูงกว่าค่าความสหสัมพันธ์แบบวัดซ้ำฉบับมาตรฐานอังกฤษ-อังกฤษนั้นมีอยู่ 3 สเกล คือ THD RC6 และ SUI ส่วนสเกลที่ค่าสหสัมพันธ์แบบวัดซ้ำกรณีทดสอบไทย-อังกฤษต่ำกว่าค่าสหสัมพันธ์แบบวัดซ้ำฉบับมาตรฐานอังกฤษ-อังกฤษนั้นมีอยู่ 12 สเกล คือ L-r K-r RCd RC4 GIC NFC STW MSF SUB AGG DISC-r และ NEGE-r

ตารางที่ 1 ตารางแสดงผลค่าสหสัมพันธ์แบบวัดซ้ำเปรียบเทียบการวัดซ้ำข้ามภาษาไทย-อังกฤษ และค่าความเชื่อมั่นแบบวัดซ้ำจากคู่มือทางเทคนิค

ชื่อสเกล	r _{Thai-English}	r _{US sample}	ความแตกต่าง	p value (Two-tailed)
Validity Scales				
VRIN-r	0.25	0.52	-0.27	0.0735
TRIN-r	0.64	0.40	0.24	0.0628
F-r	0.83	0.82	0.01	0.8650
Fp-r	0.83	0.71	0.12	0.0949
Fs	0.72	0.51	0.21	0.0549
FBS-r	0.76	0.72	0.04	0.6241
RBS	0.74	0.71	0.03	0.7263
L-r	0.51	0.79	-0.28	0.0047**
K-r	0.63	0.84	-0.21	0.0076**
Higher-Order Scales				
EID	0.81	0.90	-0.09	0.0549
THD	0.86	0.71	0.15	0.0238*
BXD	0.83	0.91	-0.08	0.0588
Restructured Clinical Scales				
RCd	0.77	0.88	-0.11	0.0477*
RC1	0.71	0.79	-0.08	0.3030
RC2	0.65	0.76	-0.11	0.2187
RC3	0.79	0.82	-0.03	0.6312
RC4	0.77	0.89	-0.12	0.0257*
RC6	0.82	0.64	0.18	0.0264*
RC7	0.83	0.88	-0.05	0.2983
RC8	0.82	0.74	0.08	0.2501
RC9	0.83	0.86	-0.03	0.5552
Specific Problems and Interest Scales				
MLS	0.69	0.82	-0.13	0.0854
GIC	0.54	0.75	-0.21	0.0404*
HPC	0.69	0.78	-0.09	0.2713
NUC	0.41	0.54	-0.13	0.3472
COG	0.81	0.74	0.07	0.3271
SUI	0.83	0.68	0.15	0.0455*
HLP	0.61	0.65	-0.04	0.7114
SFD	0.70	0.81	-0.11	0.1471
NFC	0.69	0.84	-0.15	0.0375*
STW	0.39	0.77	-0.38	0.0007**
AXY	0.69	0.71	-0.02	0.8259
ANP	0.72	0.81	-0.09	0.2225
BRF	0.71	0.67	0.04	0.6672
MSF	0.66	0.85	-0.19	0.0099**
JCP	0.83	0.85	-0.02	0.7039
SUB	0.75	0.87	-0.12	0.0455*
AGG	0.55	0.78	-0.23	0.0173*
ACT	0.76	0.77	-0.01	0.8966
FML	0.78	0.78	0.00	1.0000
IPP	0.69	0.78	-0.09	0.2713
SAV	0.84	0.84	0.00	1.0000
SHY	0.82	0.88	-0.06	0.2225
DSF	0.46	0.60	-0.14	0.2757
AES	0.78	0.86	-0.08	0.1676
MEC	0.88	0.92	-0.04	0.2340
PSY-5 Scales				
AGGR-r	0.78	0.84	-0.06	0.3271
PSYC-r	0.86	0.76	0.10	0.0989
DISC-r	0.84	0.93	-0.09	0.0151*
NEGE-r	0.62	0.85	-0.23	0.0031**
INTR-r	0.83	0.84	-0.01	0.8572

ผลการวิเคราะห์รายข้อ

สำหรับความแตกต่างในรายข้อนั้น ทางผู้วิจัยได้คำนวณออกมาเป็นเปอร์เซ็นต์การตอบซ้ำในการตอบคำถามทั้งสองครั้งในแต่ละข้อและจากการวิเคราะห์ข้อมูลคุณภาพแบบทดสอบในระดับข้อคำถามเบื้องต้นพบว่า โดยทั่วไปความสอดคล้องของการตอบคำถามในรายข้อโดยรวมมีค่าเฉลี่ยร้อยละ 77.21 ซึ่งถือว่ามากกว่าการตอบอย่างสุ่ม (ซึ่งความสอดคล้องรายข้อในการตอบอย่างสุ่มจะอยู่ที่ร้อยละ 50 เนื่องจากคำตอบเป็นแบบ 2 ตัวเลือก) อย่างมีนัยสำคัญ ($z = 3.422, p < .01$)

การวิเคราะห์ข้อคำถามรายข้อเพื่อปรับปรุงข้อคำถามที่มีปัญหานั้น ทาง University of Minnesota นั้นได้แนะนำเบื้องต้นว่า ไม่มีเกณฑ์พิจารณาความแตกต่างในระดับข้อคำถาม (item discrepancy) ที่เป็นทางการสำหรับการตรวจสอบความบกพร่องของข้อคำถามและให้ทางผู้วิจัยวิเคราะห์ได้ตามเกณฑ์ที่กำหนดเอง (arbitrary standard)

จากการปรึกษานักสถิติและผู้เชี่ยวชาญเพิ่มเติมผู้วิจัยจะทำการเลือกข้อคำถามที่เป็นปัญหา จากการวิเคราะห์เปรียบเทียบว่า ข้อคำถามแต่ละข้อจะมีโอกาสตอบสอดคล้องมากกว่าโอกาสที่จะตอบแบบเดาสุ่มอย่างมีนัยสำคัญหรือไม่ กล่าวอีกนัยหนึ่งคือ พิจารณาว่าความสอดคล้องรายข้อมากกว่าร้อยละ 50 อย่างมีนัยสำคัญหรือไม่ ซึ่งสามารถคำนวณจุดตัดสำหรับการวิเคราะห์ได้จากการทดสอบความแตกต่างของสัดส่วนแบบหางเดียว ซึ่งจุดตัดที่ได้คือร้อยละ 63 ($z = 1.644, p = .05$) โดยมีระดับนัยสำคัญที่ .05 โดยเมื่อนำจุดตัดดังกล่าวไปเปรียบเทียบกับข้อคำถามทุกข้อแล้ว พบว่ามีข้อคำถามที่ผู้ตอบไม่ได้ตอบสอดคล้องกันมากกว่าการสุ่มตอบอย่างมีนัยสำคัญจำนวน 41 ข้อ ซึ่งผู้วิจัยได้ทำการวิเคราะห์ข้อคำถามดังกล่าวและได้ทำการปรับแก้ข้อคำถามโดยพิจารณาจากภาษาที่ใช้ในข้อคำถามและ

คะแนนในสเกลที่ข้อคำถามส่งผลถึงเป็นสำคัญ ในส่วนของการปรับแก้ ผู้วิจัยได้ทำการปรับแก้ทั้งสิ้นจำนวน 18 ข้อ โดยปรับเปลี่ยนคำศัพท์และขยายความประโยคให้ง่ายต่อการเข้าใจมากขึ้น

ภายหลังการปรับแก้ผู้วิจัยได้ส่งบทวิเคราะห์และการปรับแก้แก่ University of Minnesota เพื่อปรึกษากับผู้พัฒนาแบบทดสอบ และได้รับการอนุมัติจากทางผู้พัฒนาแบบวัด พร้อมคำแนะนำให้ดูเพิ่มเติมในข้อคำถามที่ใช้วัด STW ทางผู้วิจัยจึงทำการพิจารณาข้อคำถามเหล่านั้นและพิจารณาปรับแก้เพิ่มเติม

วิจารณ์

เมื่อทำการเปรียบเทียบผลการวิจัยชิ้นนี้กับงานวิจัยอื่นที่ทำการตรวจสอบความเที่ยงตรงของการแปลแบบทดสอบ MMPI-2 และ MMPI-2-RF นั้นพบว่า Chung และคณะ⁵ ซึ่งทำการทดสอบกับฉบับที่แปลเป็นภาษาไทยเกาหลี ได้ผลค่าสหสัมพันธ์ในการทดสอบแบบวัดซ้ำแบบข้ามภาษาเกาหลี-อังกฤษอยู่ในระดับปานกลาง เมื่อเปรียบเทียบค่าสหสัมพันธ์ในการทดสอบแบบวัดซ้ำแบบข้ามภาษาเกาหลี-อังกฤษกับค่าสหสัมพันธ์การทดสอบแบบวัดซ้ำในการทำแบบทดสอบในฉบับภาษาอังกฤษมาตรฐานพบว่าค่าสหสัมพันธ์ในการทดสอบแบบวัดซ้ำแบบข้ามภาษาดำกว่าค่าสหสัมพันธ์การทดสอบแบบวัดซ้ำในการทำแบบทดสอบในฉบับภาษาอังกฤษมาตรฐาน ในการเปรียบเทียบรายข้อพบว่าข้อคำถามที่เป็นประโยคความเดียว กระชับ จะมีคะแนนความสอดคล้องรายข้อ (item agreement) ที่สูง ในทางตรงกันข้าม ข้อคำถามที่มีความซับซ้อนทางภาษารวมไปถึงมีการใช้ปฏิเสธซ้อนและการใช้สำนวนอเมริกัน จะมีคะแนนความสอดคล้องรายข้อต่ำ ในขณะที่ในฉบับภาษาลัตเวีย¹³ เมื่อทำการวัดซ้ำแบบข้ามภาษาลัตเวีย-อังกฤษ มีค่าสหสัมพันธ์ที่ไม่น้อยกว่าค่าสหสัมพันธ์การทดสอบแบบวัดซ้ำในการทำแบบทดสอบในฉบับ

ภาษาอังกฤษมาตรฐาน และสำหรับในภาษาสเปน (ลาตินอเมริกา)¹⁴ ก็พบว่าโดยรวม ค่าสหสัมพันธ์มีความสอดคล้องในกันในระดับที่ยอมรับได้ เมื่อเปรียบเทียบการวัดซ้ำแบบข้ามภาษากับการวัดซ้ำในฉบับมาตรฐานภาษาอังกฤษ

ในการศึกษาครั้งนี้ยังพบว่า ผลการวิเคราะห์รายข้อในเชิงคุณภาพ มีความสอดคล้องกับการวิเคราะห์ผลของ Chung และคณะ⁵ นั่นคือ ข้อคำถามที่มีการใช้ภาษาที่สั้นและง่าย เป็นประโยคความเดียว ส่วนใหญ่จะมีคะแนนระดับความสอดคล้องรายข้อที่สูงกว่าข้อคำถามที่เป็นประโยคซับซ้อน และมีการใช้สำนวนภาษาอังกฤษแบบอเมริกันในต้นฉบับ ทั้งนี้อาจเกิดจากปัจจัยสำคัญสองประการ คือ ความชำนาญทางภาษาของผู้เข้าร่วมการการศึกษา และ ความแตกต่างทางภาษาของภาษาไทยและภาษาอังกฤษ

ในประการแรก ถึงแม้ว่าผู้เข้าร่วมการการศึกษาในการทดสอบครั้งนี้จะเป็นผู้ที่มีความชำนาญทั้งสองภาษา และผ่านเกณฑ์พื้นฐานที่กำหนดข้างต้นในการคัดเลือกอาสาสมัครแล้วนั้น อย่างไรก็ตามผู้เข้าร่วมการศึกษาก็มีภาษาที่ถนัดมากกว่า ซึ่งในกลุ่มตัวอย่างมีทั้งผู้ที่ยังรายงานว่าถนัดภาษาไทยมากกว่าและผู้ที่ยังรายงานว่าถนัดภาษาอังกฤษมากกว่าตามพื้นหลังของแต่ละบุคคล ซึ่งทำให้ผู้เข้าร่วมวิจัยอาจจะไม่เข้าใจหรือเกิดการตีความหมายคำถามที่ไม่ถูกต้องในบางข้อ โดยเฉพาะข้อที่ใช้ศัพท์ยากหรือศัพท์ที่เป็นทางการ รวมไปถึงข้อที่มีการใช้สำนวนในภาษานั้นๆ นอกจากนี้กลุ่มตัวอย่างยังรายงานว่าศัพท์บางคำเป็นศัพท์ที่เขาไม่ค่อยพบเจอในชีวิตประจำวันและไม่ทราบความหมายของศัพท์เหล่านั้น เช่น คำว่า “สัปดน” “สบล” เป็นต้น

สาเหตุของความแตกต่างอีกประการหนึ่งคือ ความแตกต่างทางภาษา เนื่องจากภาษาไทยและภาษาอังกฤษมีความแตกต่างกันมากทั้งในด้านหลักทางไวยากรณ์ คำศัพท์ และลักษณะทางวัฒนธรรมที่เป็นตัว

กำหนดการใช้ภาษา ทำให้การแปลข้อคำถามจากภาษาอังกฤษเป็นภาษาไทยนั้นทำได้ยาก ตัวอย่างเช่น คำว่า “nervous” ภาษาอังกฤษ เมื่อทำการแปลเป็นภาษาไทยจะมีความหมายที่หลากหลาย อาทิ หงุดหงิด กังวล ประสาท เป็นต้น ทำให้ผู้แปลจะต้องเลือกคำที่สอดคล้องกับความหมายในเชิงบริบทของข้อคำถามด้วย ซึ่งหากข้อคำถามวัดปัจจัยที่เกี่ยวข้องกับความวิตกกังวลก็ควรเลือกแปลว่า กังวลใจ การแปลในลักษณะนี้ทำให้ผู้ตอบแบบทดสอบอาจมีความคิดและการตีความที่ต่างกันในแต่ละภาษา เช่น เมื่ออ่านฉบับภาษาอังกฤษ ผู้อ่านอาจจะตีความไปในเชิงหงุดหงิด แต่เมื่ออ่านฉบับภาษาไทยก็ตีความไปในลักษณะของความวิตกกังวล เป็นต้น

อย่างไรก็ตาม การแปลคำศัพท์บางคำที่มีคำแปลตายตัวก็อาจสร้างความสับสนแก่ผู้อ่านได้ โดยคำแปลภาษาไทยในข้อคำถามบางข้อ (โดยเฉพาะในสเกล RC1) ถูกแปลในเชิงปฏิเสธค่อนข้างมาก เช่น “never” ซึ่งแปลว่า “ไม่เคย” หรือ “seldom” ซึ่งแปลว่า “แทบจะไม่” เมื่อผู้อ่านต้องการตอบ โดยเฉพาะในการตอบเท็จซึ่งผู้อ่านมักจะตีความคำว่า “เท็จ” ว่า “ไม่จริง” เนื่องจากคำหลังเป็นภาษาพูดและเข้าใจได้ง่ายกว่า การตีความดังกล่าวมีผลทำให้เกิดการใช้ปฏิเสธซ้อนได้

นอกจากนี้ การแปลประโยคบางประโยคยังไม่สามารถทำได้อย่างสมบูรณ์เนื่องจากมีการใช้สำนวนภาษา โดยการใช้สำนวนภาษาอังกฤษแบบอเมริกันในข้อคำถามบางข้อเป็นอีกอุปสรรคหนึ่งที่สำคัญต่อการแปล ตัวอย่างเช่น “... on the top of the world.” “... my head ... feels tender.” “My hands ... become clumsy ...” เป็นต้น ซึ่งวลีและประโยคเหล่านี้ไม่สามารถแปลอย่างตรงไปตรงมาได้ ต้องเทียบเคียงกับความหมายที่ใกล้เคียงในภาษาไทย

ดังนั้นจึงกล่าวได้ว่า ในการแปลแบบทดสอบทางจิตวิทยา MMPI-2-RF นั้น ไม่สามารถถ่ายทอดนัยยะทางภาษาได้อย่างสมบูรณ์ ซึ่งข้อจำกัดในด้าน

ต่างๆ ที่ได้กล่าวไปในข้างต้นเป็นปัจจัยที่มีความสำคัญต่อผลการวิจัยอย่างหลีกเลี่ยงไม่ได้ อย่างไรก็ตามภายหลังจากการแปลและตรวจสอบคุณภาพของการแปลแบบทดสอบแล้ว การเก็บค่าบรรทัดฐานจากกลุ่มตัวอย่างในประชากรเป้าหมายและกลุ่มตัวอย่างเฉพาะจะเป็นอีกกลไกหนึ่งที่จะแสดงให้เห็นถึงความแตกต่างทางภาษาและวัฒนธรรมที่ส่งผลต่อค่าคะแนนมาตรฐานที่ (T-Score) ในแต่ละสเกลและโครงสร้างของแบบทดสอบในอีกระดับหนึ่ง

ข้อเสนอแนะ

ในการศึกษาครั้งนี้ ผู้วิจัยพบว่า การคัดเลือกกลุ่มตัวอย่างซึ่งมีคุณสมบัติตามเกณฑ์ที่กำหนดไว้ ทำได้ยากและต้องใช้เวลา รวมไปถึงต้องการมีตรวจสอบคุณสมบัติของอาสาสมัครอย่างรัดกุม อาจจะมีการวัดอาสาสมัครด้วยแบบทดสอบทางภาษาอย่างเข้มข้นก่อนเข้าการวิจัยอีกครั้งหนึ่งเพื่อเป็นการยืนยันถึงลักษณะความเป็นผู้ชำนาญทั้งสองภาษาของอาสาสมัครอีกครั้งหนึ่ง นอกจากนี้เนื่องจากการวิจัยมีการวัดซ้ำจึงทำให้อาสาสมัครที่ผ่านเกณฑ์การคัดเลือกเป็นผู้ร่วมการวิจัยจำเป็นต้องทำการทดสอบสองครั้งในเวลาห่างกันถึง 2 สัปดาห์ การศึกษาเช่นนี้มีโอกาสของการยกเลิกการเข้าร่วมการวิจัยสูง ซึ่งในงานวิจัยนี้มีผู้ไม่มาทำการทดสอบในครั้งที่สองมากถึง 10 ราย ดังนั้นในงานวิจัยลักษณะนี้ ผู้วิจัยควรเผื่อจำนวนของผู้เข้าร่วมการวิจัยไว้มากพอประมาณ

ผู้วิจัยยังเห็นว่า ในอนาคตควรมีการศึกษาแบบวัดซ้ำในฉบับภาษาไทยภายหลังจากที่ได้ค่าบรรทัดฐานแล้ว ซึ่งจะทำให้ได้ผลมาเปรียบเทียบกับค่าความเชื่อมั่นที่ได้จากการวัดซ้ำข้ามภาษาที่ได้จากวิจัยนี้ และการวัดซ้ำในต้นฉบับภาษาอังกฤษ ผลที่ได้ก็น่าจะมาวิเคราะห์และปรับปรุงข้อคำถามในอีกขั้นตอนหนึ่งได้

สรุป

การวิจัยเพื่อแปลแบบทดสอบ MMPI-2-RF ส่วนนี้ เป็นส่วนของการแปลภาษาของแบบทดสอบจากภาษาอังกฤษเป็นภาษาไทย การตรวจสอบคุณภาพและความเที่ยงตรงของการแปลด้วยกระบวนการแปลย้อนกลับ และการหาค่าความเชื่อมั่นของแบบทดสอบฉบับแปลโดยการวัดซ้ำไทย-อังกฤษในผู้ที่ชำนาญสองภาษา โดยเริ่มจากคณะผู้วิจัยทำการแปลแบบวัดพร้อมทำการแปลย้อนกลับ (back-translation) เพื่อตรวจสอบความเที่ยงตรงของการแปล หลังจากนั้น จึงทำการหาค่าความเชื่อมั่นแบบวัดซ้ำกับผู้ชำนาญสองภาษาเพื่อเปรียบเทียบค่าความสหพันธ์ที่ได้จากการทดสอบนี้กับค่าความเชื่อมั่นจากการวัดซ้ำในภาษาอังกฤษฉบับมาตรฐานที่ปรากฏในคู่มือทางเทคนิคโดยผลการวิเคราะห์และแก้ไขข้อคำถามได้รับการอนุมัติจาก University of Minnesota เป็นที่เรียบร้อยแล้ว และแบบทดสอบได้รับการอนุมัติเพื่อใช้เก็บค่าบรรทัดฐานในประเทศไทยได้ในระยะต่อไป

กิตติกรรมประกาศ

การศึกษานี้ได้รับทุนวิจัยจากคณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์สัญญาที่ 5/2559 ทั้งนี้ผู้วิจัยขอขอบคุณคณาจารย์สาขาจิตวิทยา มหาวิทยาลัยธรรมศาสตร์ ที่คอยให้ความสนับสนุนและคำปรึกษาในการทำวิจัย ผู้ช่วยวิจัยทุกท่านในการเก็บรวบรวมข้อมูล และผู้เข้าร่วมการวิจัยทุกท่านที่สละเวลาเพื่อเข้าร่วมการวิจัย

เอกสารอ้างอิง

1. Archer RP, Buffington-Vollum JK, Stredny RV, Handel RV. A survey of psychological test use patterns among forensic psychologists. *J Pers Assess* 2006; 87: 84-94.

2. Ben-Porath YS, Tellegen A. MMPI-2-RF manual for administration, scoring and interpretation. Minneapolis: University of Minnesota Press; 2008.
3. Ben-Porath YS, Tellegen A. MMPI-2-RF technical manual. Minneapolis: University of Minnesota Press; 2008.
4. Camara WJ, Nathan JS, Puente AE. Psychological test usage: Implications in professional psychology. *Prof Psychol Res Pr* 2000; 31: 141-54. doi:10.1037/0735-7028.31.2.141.
5. Chung JJ, Weed NC, Han K. Evaluating cross-cultural equivalence of the Korean MMPI-2 via bilingual test-retest. *Int J Intercult Relat* 2006; 30: 531-43.
6. Gilberstadt H, Duker J. A handbook for clinical and actuarial MMPI interpretation. Oxford, England: Saunders; 1965.
7. Greene RL. The MMPI-2/MMPI-2-RF: An interpretive manual. 3rd ed. Boston: Allyn & Bacon; 2011.
8. Groth-Marnat G. Handbook of psychological assessment. 4th ed. New York: John Wiley & Sons; 2003.
9. Hathaway SR. Foreword. In: Dahlstorm WG, Welsh GS, editors. An MMPI handbook: A guide to use in clinical practice and research. Minneapolis: University of Minnesota Press; 1960.
10. Lubin B, Lasren RM, Matarazzo JD. Pattern of psychological test usage in United States: 1935-1982. *Am Psychol* 1984; 39: 451-54.
11. Marks PA, Seeman W. The actuarial description of abnormal personality: An atlas for use with the MMPI-2. Baltimore: Williams & Wilkins; 1963.
12. Marks PA, Seeman W. The actuarial use of the MMPI with adolescents and adults. Baltimore: Williams & Wilkins; 1974.
13. Sarma ZM. Adaptation of the Minnesota Multiphasic Personality Inventory-2 to Latvia. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 2006; 66(10-B): 5694.
14. Velasquez RJ, Chavira DA, Karle HR, Callahan WJ, Garcia JA, Castellanos J. Assessing bilingual and monolingual Latino students with translations of the MMPI-2: Initial data. *Cultur Divers Ethnic Minor Psychol* 2000; 6: 65-72.