

การสร้างและทดสอบแบบประเมินการทำงานของมือฉบับภาษาไทยด้วย Rasch Model

สุภารัตน์ สุขโท^{1,2}, แสง วัชรธนกิจ³

¹นักศึกษาปริญญาเอก คณะเภสัชศาสตร์ มหาวิทยาลัยอุบลราชธานี

²อาจารย์ประจำคณะแพทย์แผนไทยและแพทย์ทางเลือก มหาวิทยาลัยราชภัฏอุบลราชธานี

³กลุ่มวิชาเภสัชกรรมปฏิบัติ คณะเภสัชศาสตร์ มหาวิทยาลัยอุบลราชธานี

บทคัดย่อ

วัตถุประสงค์: เพื่อสร้างและทดสอบแบบประเมินการทำงานของมือฉบับภาษาไทย **วิธีการ:** แบบประเมินฉบับร่างสร้างขึ้นจากแบบสอบถามบริเวณมือฉบับภาษาไทยและต่างประเทศ ตลอดจนการสัมภาษณ์นักกายภาพบำบัดและนักกิจกรรมบำบัด จากนั้นตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ 3 ท่าน พบดัชนีความสอดคล้องของข้อคำถามกับวัตถุประสงค์มีค่าอยู่ระหว่าง 0.67-1.00 แบบประเมินฉบับสุดท้ายมีคำถาม 33 ข้อ บนมาตราวัด 5 ระดับ จากนั้นเก็บข้อมูลด้วยแบบประเมินในผู้ป่วยจำนวน 180 ราย จากแผนกกายภาพบำบัดของ 10 โรงพยาบาล วิเคราะห์ข้อมูลโดยใช้ Rasch Model **ผลการวิจัย:** แบบประเมินที่สร้างมีความเป็นเอกมิติจากค่าความแปรปรวนที่อธิบายได้จากโมเดลการวัด (มีค่าร้อยละ 54.96) และความแปรปรวนขององค์ประกอบแรกของส่วนเหลือ (มีค่าร้อยละ 3.61) มีค่ามากกว่าร้อยละ 40 และน้อยกว่าร้อยละ 15 ตามลำดับ ค่าถามทุกข้อมีความเป็นอิสระของข้อคำถาม โดย person reliability และ person separation มีค่า 0.93 และ 3.59 ตามลำดับ ส่วนค่า item reliability และ item separation มีค่า 0.95 และ 4.29 ซึ่งทุกค่าอยู่ในเกณฑ์เป็นที่ยอมรับ การทดสอบความเหมาะสมรายข้อพบว่า ค่าถามทั้ง 33 ข้อ มีค่า infit mean-squares และค่า outfit mean-squares อยู่ในเกณฑ์เหมาะสมซึ่งอยู่ระหว่าง 0.50-1.50 ยกเว้นข้อคำถามที่ 18 มีค่า outfit mean-squares เท่ากับ 1.53 หลังจากปรับปรุงตัวเลือกเหลือ 3 รายการคำตอบและตัดกลุ่มตัวอย่างที่ไม่เหมาะสมกับ Rasch Model จำนวน 29 ราย แล้วทดสอบความเหมาะสมรายข้อ พบว่าค่าถามทั้ง 33 ข้อ มีค่า infit mean-squares และค่า outfit mean-squares อยู่ในเกณฑ์เหมาะสม ข้อคำถามมีระดับความยากไม่กระจายตัวเมื่อเทียบกับการกระจายตัวของระดับความสามารถในการทำงานของมือของกลุ่มตัวอย่าง โดยพบว่าตัวอย่างส่วนหนึ่งมีความสามารถในการทำงานของมือน้อยกว่าข้อคำถามที่ง่ายที่สุด และตัวอย่างอีกส่วนหนึ่งมีระดับความสามารถในการทำงานของมือน้อยกว่าข้อคำถามที่ยากที่สุด โดยข้อคำถามมีระดับความยากระหว่าง -1.19 ถึง 1.55 logits ในขณะที่คะแนนระดับความสามารถในการทำงานของมือระหว่าง -6.62 ถึง 5.12 logits **สรุป:** แบบประเมินการทำงานของมือฉบับภาษาไทยมีความเป็นเอกมิติและความเป็นอิสระของข้อคำถาม มีความตรงและความเที่ยงที่ดี แต่ในการศึกษาครั้งต่อไปควรเพิ่มข้อคำถามที่วัดระดับการทำงานของมือทั้งระดับยากและง่ายขึ้นกว่าเดิมเพื่อให้ครอบคลุมระดับความสามารถในการทำงานของมือของกลุ่มตัวอย่าง

คำสำคัญ: แบบประเมินการทำงานของมือ ภาษาไทย Rasch Model

รับต้นฉบับ: 24 ก.ค. 2562, ได้รับบทความฉบับปรับปรุง: 2 ม.ค. 2563, รับลงตีพิมพ์: 5 ม.ค. 2563

ผู้ประสานงานบทความ: แสง วัชรธนกิจ คณะเภสัชศาสตร์ มหาวิทยาลัยอุบลราชธานี อำเภอวารินชำราบ จังหวัดอุบลราชธานี 34190 **E-mail:** sawaeng.w@ubu.ac.th

Development and Validation of Thai Hand Function Assessment Questionnaire Using the Rasch Model

Suparat Sooktho^{1,2}, Sawaeng Watcharathanakij³

¹Graduate student in Doctor of Philosophy Program in Pharmaceutical Sciences, Ubon Ratchathani University

²Lecturer, Faculty of Thai Traditional and Alternative Medicine, Ubon Ratchathani Rajabhat University

³Department of Pharmacy Practice, Faculty of Pharmaceutical Sciences, Ubon Ratchathani University

Abstract

Objective: To develop and validate the Thai Hand Function Assessment Questionnaire. **Methods:** The first draft of the Thai Hand Function Assessment Questionnaire was developed based upon Thai and foreign questionnaires and from an interview with physical therapists and occupational therapists. Its face validity was assessed by 3 experts' opinion. Index of item-objective congruence ranged 0.67-1.00. The final version of the Thai Hand Function Assessment Questionnaire comprised 33 items with a 5-point scale. It was administered in to 180 patients in the physical therapy departments of 10 hospitals. Data were analyzed by the Rasch model. **Results:** The questionnaire showed unidimensionality from the variance explained by measures (54.96%) and the variance of residuals in the first contrast (3.61%) was greater than 40% and less than 15% respectively. All items displayed local independence. Person reliability and person separation index were 0.93 and 3.59 whereas item reliability and item separation were 0.95 and 4.29 respectively, which were regarded as acceptable. Test of item fit statistics revealed that 33 items had optimal infit mean-squares and outfit mean-squares ranging between 0.50-1.50 except for item 18 that outfit mean-squares was 1.53. After combining 5 response categories to 3 categories and excluding 29 mis-fitted persons from the Rasch model, the test of item fit statistics found that all 33 items had acceptable infit mean-squares and outfit mean-squares. Item difficulty was not well distributed when compared with the distribution of ability of subjects' hand function. Some groups of subjects had their hand function ability lower than the easiest item while others had their hand function ability higher than the most difficult item. Item difficulty ranged between -1.19 to 1.55 logits while ability of hand function ranged between -6.62 to 5.12 logits. **Conclusion:** The Thai Hand Function Assessment Questionnaire has unidimensionality and local independence with high validity and reliability. However, further studies should include additional items with more difficult and easier items to capture all hand function levels of the subjects.

Keywords: Thai Hand Function Assessment Questionnaire, Thai version, Rasch model

บทนำ

มือเป็นอวัยวะที่สำคัญของร่างกายและมีความสำคัญกับคุณภาพชีวิตของทุก ๆ คน เนื่องจากทุกคนใช้มือในการทำกิจกรรมประจำวันทั้งระดับง่ายและยาก ได้แก่ การหยิบจับสิ่งของขนาดเล็ก การจับแก้วน้ำเพื่อดื่ม การอาบน้ำทำความสะอาดร่างกาย การขยับข้อนิ้วหรือข้อนิ้วข้อมือเพื่อเดินทางไปทำงาน หรือการหิ้วของยกของที่มีน้ำหนักน้อยจนถึงน้ำหนักมาก หากมีความผิดปกติของมือ ย่อมส่งผลกระทบต่อการทำงานในชีวิตประจำวัน ดังนั้นแบบประเมินการทำงานของมือฉบับภาษาไทยจึงมีความสำคัญ เพราะผู้ประเมินสามารถใช้แบบประเมินบอกระดับความสามารถในการทำงานของมือได้

นักกายภาพบำบัดเป็นวิชาชีพที่มีความสำคัญและมีบทบาทหน้าที่ในการประเมินและให้การรักษาผู้ป่วยที่มีความผิดปกติบริเวณมือ อีกทั้งในปัจจุบันยังไม่มีข้อกำหนดแบบประเมินการทำงานของมือที่เป็นมาตรฐาน ทำให้นักกายภาพบำบัดใช้เครื่องมือที่หลากหลายประเมินการทำงานของมือ ดังนั้นแบบประเมินการทำงานของมือฉบับภาษาไทย ที่ใช้ประเมินกิจกรรมประจำวันหรือความสามารถในการใช้งานมือของผู้ป่วยที่เป็นมาตรฐาน ซึ่งหากมีการใช้ร่วมกันในวงกว้าง อาจทำให้คุณภาพการรักษานักกายภาพบำบัดดีขึ้น โดยเฉพาะในกรณีที่มีการส่งตัวผู้ป่วยไปรักษาที่สถานบริการอื่น ๆ ทำให้มีการใช้ข้อมูลร่วมกันระหว่างสถานบริการในการดูแลผู้ป่วย

จากการทบทวนวรรณกรรมพบว่า มีแบบประเมินการทำงานของร่างกาย Upper Limb Functional Index (ULFI) (1) แบบประเมินการทำงานของข้อศอก Patient Rated Elbow Evaluation (PREE) (2) และแบบประเมินการทำงานของข้อไหล่ Shoulder Pain and Disability Index (SPADI) (3) ซึ่งเป็นแบบประเมินฉบับต่างประเทศที่มีข้อคำถามที่เกี่ยวข้องกับหลายข้อต่อ ส่วนใหญ่ไม่ได้แยกประเมินการทำงานเฉพาะส่วนมือ สำหรับ British Measure of Activity Performance in the Hand (MAP-Hand) (4), 6 item CTS symptom scale (5), Patient-Rated Wrist Evaluation (PRWE) (6) และ Munich Wrist Questionnaire (MWQ) (7) เป็นแบบประเมินส่วนมือ แต่ข้อคำถามที่ประเมินกิจกรรมประจำวันส่วนใหญ่ไม่เข้ากับบริบทของคนไทย และยังไม่พบแบบประเมินการทำงานเฉพาะส่วนมือฉบับภาษาไทยที่ครอบคลุมกิจกรรมประจำวันของผู้ป่วย (8-9)

สำหรับการประเมินการทำงานของมือในปัจจุบันของนักกายภาพบำบัด ยังไม่มีประกาศข้อกำหนดจากสภากายภาพบำบัดให้แต่ละแผนกในโรงพยาบาลใช้แบบประเมินเป็นรูปแบบเดียวกัน จึงมีการใช้แบบประเมินที่หลากหลาย และแบบประเมินมือส่วนใหญ่มีการวิเคราะห์โดยอิงตามทฤษฎีดั้งเดิม (6-7) ซึ่งยังมีข้อจำกัด ดังนั้นการศึกษานี้มีวัตถุประสงค์เพื่อสร้างแบบประเมินการทำงานของมือฉบับภาษาไทยสำหรับผู้ป่วยที่มีพยาธิสภาพบริเวณมือ พร้อมตรวจสอบคุณสมบัติความตรงและความเที่ยงโดยใช้วิธี Rasch Model เพื่อทดสอบคุณสมบัติของแบบประเมินที่สร้างขึ้น

วิธีการวิจัย

การศึกษานี้ได้ผ่านการพิจารณาจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยอุบลราชธานี เลขที่ UBU-REC-20/2561

ตัวอย่าง

ตัวอย่าง คือ ผู้ป่วยระบบกระดูกและกล้ามเนื้อหรือระบบประสาทที่มีพยาธิสภาพที่มือจำนวน 180 ราย ซึ่งมารับบริการที่โรงพยาบาล 10 แห่ง ได้แก่ โรงพยาบาลศรีเทพ โรงพยาบาลวิเชียรบุรี โรงพยาบาลหล่มสัก โรงพยาบาลท่าปลา โรงพยาบาลแม่สอด โรงพยาบาลหนองไผ่ โรงพยาบาลสตึก โรงพยาบาลลำทับ โรงพยาบาลวารินชำราบ และโรงพยาบาลสรรพสิทธิประสงค์ ทั้งนี้ต้องเป็นผู้ที่สื่อสารได้ เข้าใจ และสามารถทำตามคำแนะนำได้

เครื่องมือ

ผู้วิจัยสร้างแบบประเมินการทำงานของมือฉบับภาษาไทยจากการทบทวนแบบประเมินที่เกี่ยวข้องกับการทำกิจกรรมประจำวันบริเวณมือหรือมีการใช้มือร่วมด้วย โดยเลือกกิจกรรมประจำวันที่เกี่ยวข้องกับวิถีชีวิตคนไทย และปรับให้เข้ากับบริบทของคนไทยรวมทั้งรวม 9 แบบประเมิน ได้แก่ 1) Upper Limb Functional Index (1) 2) Patient Rated Elbow Evaluation (2) 3) Shoulder Pain and Disability Index (3) 4) British Measure of Activity Performance in the Hand (4) 5) 6-item CTS Symptom Scale (5) 6) Patient-Rated Wrist Evaluation (6) 7) Munich Wrist Questionnaire (7) 8) Disabilities of the Arm, Shoulder and Hand (DASH) ฉบับภาษาไทย (8) และ 9) Boston Questionnaire ฉบับภาษาไทย (9) จากนั้นตรวจสอบความตรงเชิงเนื้อหาของแบบประเมินการทำงาน

ของมือฉบับภาษาไทยฉบับร่างโดยใช้ผู้เชี่ยวชาญ 3 ท่าน ได้แก่ นักกายภาพบำบัด 2 ท่านและนักกิจกรรมบำบัด 1 ท่านที่มีประสบการณ์ในการทำงาน 5 ปีขึ้นไป พิจารณาความสอดคล้องของคำถามที่ใช้ประเมินการทำงานของมือของผู้ป่วยว่าตรงตามคุณลักษณะหรือองค์ประกอบที่กำหนดหรือไม่

จากคำถามทั้งหมด 52 ข้อ ผู้เชี่ยวชาญได้เสนอแนะให้ปรับคำถามที่มีความซ้ำซ้อนที่อาจทำให้ผู้ตอบสับสน และตัดข้อคำถามที่ไม่เหมาะสมกับบริบทของคนไทย จำนวน 19 ข้อออก คงเหลือข้อคำถาม 33 ข้อ การศึกษานี้ประเมินความตรงเชิงเนื้อหาจากค่าดัชนีความสอดคล้อง (Item-Objective Congruence Index or IOC) คำถามมีค่า IOC อยู่ระหว่าง 0.67-1.00 ซึ่งอยู่ในเกณฑ์ที่เหมาะสม ($IOC \geq 0.5$) (10) จึงถือว่าคำถามสอดคล้องกับคุณลักษณะหรือองค์ประกอบที่ต้องการวัด

แบบประเมินการทำงานของมือฉบับภาษาไทยแบ่งเป็น 2 ส่วน ได้แก่ ส่วนที่ 1 เป็นคำถาม 5 ข้อที่วัดข้อมูลส่วนบุคคล และส่วนที่ 2 เป็นคำถาม 33 ข้อที่วัดความสามารถในการทำงานของมือของผู้ป่วย ดังแสดงในตารางที่ 3 แบบประเมินเป็นมาตราส่วนประมาณค่า 5 ตัวเลือกจากทำไม่ยาก (1) ไปยังทำไม่ได้เลย (5) การคำนวณคะแนนต้องกลับหัวคะแนนก่อนนำคะแนนแต่ละข้อมารวมกัน เช่น จาก 5 คะแนนเป็น 1 คะแนน หรือ 2 คะแนน เป็น 4 คะแนน เพื่อให้ค่าคะแนนสูงหมายถึงความสามารถของมือในการทำกิจกรรมประจำวันมากกว่าค่าคะแนนต่ำ

การเก็บรวบรวมข้อมูล

การศึกษานี้เก็บข้อมูลโดยนักกายภาพบำบัดในทั้ง 10 โรงพยาบาลที่มีประสบการณ์การทำงานอย่างน้อย 1 ปีและเป็นผู้ที่ดูแลรักษาผู้ป่วย ผู้เก็บข้อมูลได้รับการชี้แจงจากคู่มือการใช้แบบประเมิน นักกายภาพบำบัดประเมินความสามารถในการทำกิจกรรมในชีวิตประจำวันของผู้ป่วยเมื่อเก็บข้อมูลเสร็จสิ้น จะส่งแบบประเมินกลับมายังผู้วิจัยทางไปรษณีย์โดยวิธีลงทะเบียนหรือตัวนพิเศษ

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลส่วนบุคคลของกลุ่มตัวอย่างใช้สถิติเชิงพรรณนา และใช้โปรแกรม WINSTEPS version 3.9.1.0 วิเคราะห์ความตรงของแบบประเมินด้วยวิธี Rasch Model โดยมีรายละเอียดดังนี้

เส้นโค้งการเลือกการตอบ (category response curves or CRC) เป็นแผนภาพแสดงความน่า

จะเป็นในการเลือกการตอบของตัวอย่างที่มีระดับความสามารถในการทำงานของมือที่แตกต่างกัน ดังนั้นผู้ป่วยที่มีระดับความสามารถในการทำงานของมือมาก ย่อมมีโอกาสในการเลือกตัวเลือกการตอบ “ไม่ยาก” มากกว่าผู้ป่วยที่มีความสามารถน้อยกว่า ในขณะที่ผู้ป่วยที่มีระดับความสามารถในการทำงานของมือน้อย ย่อมมีโอกาสเลือกการตอบ “ทำไม่ได้เลย” มากกว่าผู้ป่วยที่มีความสามารถมากกว่า

Fit statistics เป็นค่าที่แสดงถึงความสอดคล้องภายในแต่ละข้อคำถามว่าสอดคล้องกับราสช์โมเดลที่กำหนดไว้หรือไม่ แบ่งเป็น person fit เป็นการตรวจสอบรูปแบบของคะแนน (item-score) ที่เป็นไปตามโมเดลของกลุ่มตัวอย่าง (11) และ item fit เป็นการตรวจสอบความเหมาะสมของคำถามรายข้อ ซึ่งมีค่าสถิติที่ใช้คือ OUTFIT MNSQ (outfit mean-square) และ INFIT MNSQ (infit mean-square) ทั้งสองค่าขึ้นอยู่กับแนวคิดของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนมาตรฐานระหว่างผลการตอบข้อคำถามและโมเดลการวัดนั้น (12) การศึกษานี้กำหนด ค่า OUTFIT MNSQ ของ Person fit ที่ > 2.00 เพื่อตัดกลุ่มตัวอย่างที่มีรูปแบบไม่ตรงตามโมเดลออก สำหรับค่า item fit ใช้ค่า OUTFIT MNSQ และ INFIT MNSQ ระหว่าง 0.50-1.50 (13) เป็นเกณฑ์ในการตัดสินคุณภาพของข้อคำถาม

การทดสอบความเที่ยงเป็นการทดสอบความคงเส้นคงวาของการวัด และตรวจสอบความสอดคล้องภายในของแบบสอบถาม ซึ่งเกณฑ์ที่ยอมรับคือ ค่า person reliability > 0.80 (เทียบเท่ากับสัมประสิทธิ์แอลฟา) item reliability > 0.90 , person separation > 2 และ item separation > 3 (14)

การทดสอบความเป็นเอกมิติ (unidimensionality) ทำโดยการทดสอบ principal components analysis (PCA) ซึ่งเป็นการวิเคราะห์เศษเหลือ (residuals) หลังจากประเมินมิติหลัก (15) โดยเกณฑ์ในการพิจารณาลักษณะความเป็นเอกมิติมี 3 ข้อ ได้แก่ 1) ค่าความแปรปรวนที่อธิบายโดยมิติของการวัด (variance explained by the measurement dimension) ต้องไม่น้อยกว่าร้อยละ 40 (16) 2) ค่าความแปรปรวนที่อธิบายได้โดยองค์ประกอบแรกของเศษเหลือ (variance explained by first principal component of the residuals) ต้องไม่เกินร้อยละ 15 และ 3) อัตราส่วนของความแปรปรวนในข้อ 1 และข้อ 2 ต้อง $\geq 3:1$ (17)

ความเป็นอิสระของข้อคำถาม (local independence) หมายความว่า การเลือกตอบข้อคำถามใด ๆ จะไม่มีผลต่อการเลือกตอบข้อคำถามอื่นๆ ซึ่งสามารถพิจารณาได้จากค่า intra-item correlations ของ standardized residuals ของแบบสอบถาม ซึ่งควรมีค่า < 0.70 (18)

แผนภาพบุคคลและข้อคำถาม (item-person map) แสดงความสัมพันธ์ของระดับความสามารถในการทำงานของมือของผู้ป่วย (person ability) และความยาก-ง่ายของข้อคำถาม (item hierarchies) แผนภาพนี้จะนำคุณสมบัติทั้งสองมาวางบนแกนที่มีมาตรวัดเดียวกันโดยมีหน่วยการวัดเป็น logits ทำให้สามารถเปรียบเทียบได้ว่า ข้อคำถามนั้นยากหรือง่ายหรือพอดีกับความสามารถในการทำงานของมือของกลุ่มตัวอย่าง

การทำหน้าที่ต่างกันของข้อคำถาม (differential item functioning or DIF) หมายถึง ข้อคำถามทำให้ผลการตอบคำถามระหว่างกลุ่มย่อยของตัวอย่างที่มีความสามารถในการทำงานของมือระดับเดียวกัน มีโอกาสในการตอบข้อคำถามได้แตกต่างกัน เช่น กลุ่มตัวอย่างเพศชายและหญิงที่มีความสามารถในการทำงานของมือของผู้ป่วยระดับเดียวกัน มีความน่าจะเป็นในการเลือกการตอบแตกต่างกัน ถือว่าข้อคำถามนั้นมีการทำหน้าที่ต่างกัน ดังนั้นการวิเคราะห์ DIF ช่วยในการตรวจสอบว่า ข้อคำถามมีความแตกต่างกันระหว่างกลุ่มย่อยของกลุ่มตัวอย่างหรือไม่ เกณฑ์ที่ใช้ในการตัดสิน โดยทั่วไปมี 2 ข้อ ได้แก่ 1)

มีความแตกต่างของผลตอบสนองต่อข้อคำถามระหว่างกลุ่มของกลุ่มตัวอย่าง ≥ 0.50 logits และ 2) เมื่อทดสอบ chi-squared ค่า $P < 0.05$ ถือว่าข้อคำถามนั้นมีการทำหน้าที่ต่างกัน (19)

ผลการวิจัย

ลักษณะของตัวอย่าง

จากตัวอย่าง 180 ราย ผู้วิจัยตัดตัวอย่าง 14 ราย ออกจากการวิเคราะห์เนื่องจากรูปแบบของคะแนนไม่เหมาะสม (person misfit) กับราสซ์โมเดล หลังจากวิเคราะห์ความเหมาะสมของการวัดระดับความสามารถในการใช้มือทำกิจวัตรประจำวันรายบุคคล (person measure) พบความไม่เหมาะสมของคะแนน จึงตัดตัวอย่างออกอีก 15 ราย ตารางที่ 1 แสดงคุณลักษณะของตัวอย่าง 151 รายที่คงเหลือ ตัวอย่างส่วนใหญ่เป็นเพศหญิง (ร้อยละ 60.90) มีอายุระหว่าง 60-90 ปี (ร้อยละ 64.20) ประกอบอาชีพเกษตรกร (ร้อยละ 35.80) ระยะเวลาในการประกอบอาชีพไม่เกิน 10 ปี (ร้อยละ 45.70) และโรคที่ได้รับการวินิจฉัย คือ โรคอัมพฤกษ์/อัมพาตที่มีสาเหตุอื่นนอกจากโรคหลอดเลือดสมองขาดเลือด (hemiplegia) (ร้อยละ 33.10)

โครงสร้างของการเลือกรายการคำตอบ

ตารางที่ 2 แสดงผลการวิเคราะห์โครงสร้างของการเลือกรายการคำตอบ (rating scale response structures)

ตารางที่ 1. ลักษณะของกลุ่มตัวอย่าง (n=151 ราย)

	ลักษณะทั่วไป	จำนวน (ร้อยละ)
เพศ		
หญิง		92 (60.90)
ชาย		59 (39.10)
อายุ (ปี)		
10-59		54 (35.80)
60-90		97 (64.20)
อาชีพ		
เกษตรกร		54 (35.80)
รับราชการ/รัฐวิสาหกิจ		23 (15.20)
รับจ้าง		19 (12.60)
ค้าขาย/ธุรกิจ		16 (10.60)
นักศึกษา		9 (6.00)
อื่น ๆ ได้แก่ ข้าราชการบำนาญ ผู้ป่วยไม่ได้ประกอบอาชีพ		20 (22.20)

ตารางที่ 1. ลักษณะของกลุ่มตัวอย่าง (n=151 ราย) (ต่อ)

ลักษณะทั่วไป	จำนวน (ร้อยละ)
ระยะเวลาในการประกอบอาชีพ (ปี)	
≤ 10	69 (45.70)
11-20	33 (21.90)
21-30	16 (10.60)
31-40	23 (15.20)
41-50	9 (6.00)
> 51	1 (0.70)
โรคที่ได้รับการวินิจฉัย	
โรคอัมพฤกษ์/อัมพาตที่มีสาเหตุอื่นนอกจากโรคหลอดเลือดสมองขาดเลือด (hemiplegia)	50 (33.10)
โรคหลอดเลือดสมองขาดเลือด	18 (11.90)
โรคนิ้วล็อก (trigger finger)	15 (9.90)
ภาวะกระดูกหัก (fracture)	14 (9.30)
โรคการกดทับเส้นประสาทบริเวณข้อมือ (carpal tunnel syndrome)	11 (7.30)
โรคปลอกหุ้มเอ็นนิ้วหัวแม่มืออักเสบ (de quervain)	2 (1.30)
ข้อติด (joint stiffness)	1 (0.70)
อื่น ๆ ได้แก่ intracerebral hematoma, cerebral infection, head injury, brain tumor, spinal cord injury, tennis elbow syndrome radial nerve palsy, muscle strain	40 (26.50)

ผลการวิเคราะห์แสดงความถี่ของการเลือกตัวเลือกรายการคำตอบที่ไม่เท่าเทียม ซึ่งสะท้อนให้เห็นถึงความไม่สม่ำเสมอของการกระจายคำตอบ แต่ไม่พบตัวเลือกใดที่มีความถี่เป็นศูนย์ ตัวเลือกที่มีการตอบมากที่สุดคือ “ยากเล็กน้อย” (ร้อยละ 24.73) และรายการคำตอบ “ทำไม่ได้เลย”, “ยากปานกลาง”, “ไม่ยาก”, “ยากมาก” มีผู้ตอบร้อยละ 23.10, 21.90, 18.13 และ 12.14 ตามลำดับ

ความเหมาะสมของข้อคำถาม

จากตารางที่ 3 การทดสอบความเหมาะสมรายข้อ (item fit) พบว่า ก่อนการตัดตัวอย่างที่ไม่เหมาะสมกับราสซ์โมเดล คำถามทั้ง 33 ข้อ มีค่า infit mean-squares และค่า outfit mean-squares อยู่ในเกณฑ์เหมาะสม (0.50-1.50)

(13) ยกเว้นข้อคำถามที่ 18 “การล้างรถและเช็ดทำความสะอาด” มีค่า outfit mean-squares มากกว่าเกณฑ์ที่เหมาะสมเพียงเล็กน้อย เท่ากับ 1.53 อย่างไรก็ตาม หลังจากตัดตัวอย่างที่ไม่เหมาะสมกับราสซ์โมเดลจำนวน 29 รายการออก พบว่า คำถามมีค่า infit mean-squares และค่า outfit mean-squares อยู่ในเกณฑ์เหมาะสมทุกข้อ

ความเป็นเอกมิติ

การทดสอบความเป็นเอกมิติในการศึกษาครั้งนี้ หมายถึง แบบวัดความสามารถในการทำงานของมือในการทำกิจวัตรประจำวันว่ามีเพียงมิติเดียวหรือไม่ โดยทดสอบด้วย PCA ตารางที่ 4 แสดงผลการทดสอบความเป็นเอกมิติเนื่องจากความแปรปรวนที่อธิบายได้โดยมิติของการวัดคือ

ตารางที่ 2. การกระจายของการเลือกรายการคำตอบ

	เมื่อตัวเลือกเป็น 5 รายการคำตอบ					เมื่อปรับตัวเลือกเป็น 3 รายการคำตอบ		
	ไม่ยาก	ยากเล็กน้อย	ยากปานกลาง	ยากมาก	ทำไม่ได้เลย	ไม่ยาก	ยากเล็กน้อย-ยากมาก	ทำไม่ได้เลย
ความถี่	1077	1469	1301	721	1372	1077	3491	1372
ร้อยละ	18.13	24.73	21.90	12.14	23.10	18.13	58.77	23.10

ตารางที่ 3. ความเหมาะสมรายข้อ item fit (infit mean-squares และค่า outfit mean-squares)

คำถาม	MNSQ ¹ (n=180)		MNSQ ¹ (n=151)	
	5 รายการคำตอบ		3 รายการคำตอบ	
	infit	outfit	infit	outfit
1. การเปิดฝากระป๋องแบบใช้นิ้วกด	0.94	0.90	0.87	0.80
2. การเปิดฝาขวดที่ต้องใช้มือหมุน	0.94	0.92	0.87	0.82
3. การเขียนหนังสือ	1.34	1.36	1.29	1.32
4. การทำอาหาร				
4.1 การใช้ตะหลิว	0.70	0.69	0.82	0.69
4.2 การใช้มีด	0.84	0.80	0.96	0.90
4.3 การจับหม้อมีด้าม	0.99	0.96	0.83	0.69
4.4 การจับหุ้หม้อไม่มีด้าม	0.96	1.29	0.91	1.05
5. การถือจานข้าว	1.02	1.02	0.98	1.00
6. การหยิบแก้ว				
6.1 แก้วน้ำปกติ (ไม่มีหู) เช่น แก้วน้ำดื่มทั่วไป	0.90	0.90	0.88	0.89
6.2 แก้วน้ำแบบมีหู เช่น แก้วเบียร์	0.78	0.87	0.85	0.83
7. การใช้ช้อน – ส้อม	0.85	0.89	0.87	0.81
8. การล้างจาน	0.66	0.67	0.74	0.64
9. การกวาดบ้าน/ถูบ้าน	0.89	0.87	0.90	0.84
10. การซักผ้าด้วยมือ	0.99	0.94	0.91	0.83
11. การบิดยาสีฟัน	1.05	1.01	1.02	1.02
12. การแปรงฟัน	0.92	0.94	0.88	0.83
13. การทำความสะอาดร่างกาย (อาบน้ำ)	0.90	0.96	0.94	0.92
14. การสระผม	0.90	0.96	0.98	1.07
15. การสวมเสื้อผ้า				
15.1 สวมเสื้อยืดคอกลม	1.36	1.28	1.27	1.21
15.2 สวมเสื้อเชิ้ต กางเกงหรือกระโปรงที่มีกระดุม	1.13	1.19	0.97	0.91
15.3 กางเกงหรือกระโปรงที่มีซิป	1.26	1.36	1.11	0.97
16. การใส่รองเท้า	1.10	1.01	1.14	1.18
17. การขับรถ				
17.1 การจับพวงมาลัยรถยนต์/การจับพวงมาลัยรถมอเตอร์ไซด์	1.15	1.14	1.13	1.30
17.2 การหมุนพวงมาลัยรถยนต์	1.26	1.13	1.23	1.22
17.3 การบิดพวงมาลัยรถมอเตอร์ไซด์	1.04	0.97	0.98	0.93
18. การล้างรถ และเช็ดทำความสะอาดรถ	0.95	1.53*	0.99	1.25
19. การจับและไขลูกบิดประตู	0.76	0.77	0.83	0.67
20. การหยิบเม็ดยา	0.88	0.95	0.92	0.94
21. การหิ้วของ	1.14	1.32	1.18	1.14
22. การกดเครื่องคิดเลข	0.96	0.94	0.98	0.94

ตารางที่ 3. ความเหมาะสมรายข้อ item fit (infit mean-squares และค่า outfit mean-squares) (ต่อ)

คำถาม	MNSQ ¹ (n=180)		MNSQ ¹ (n=151)	
	5 รายการคำตอบ		3 รายการคำตอบ	
	infit	outfit	infit	outfit
23. การใช้เมาส์	1.29	1.34	1.14	1.02
24. การพิมพ์คอมพิวเตอร์	1.28	1.35	1.06	0.92
25. การกรตริโมท	1.01	0.93	0.81	0.76

1: MNSQ = mean square *: คำถามที่มีค่า outfit mean-square ไม่อยู่ในเกณฑ์ที่เหมาะสม

ร้อยละ 54.96 ค่าความแปรปรวนของเศษเหลือที่อธิบายได้ โดยองค์ประกอบแรก มีค่าเท่ากับร้อยละ 3.61 และ อัตราส่วนของความแปรปรวนทั้งสองเท่ากับ 15.22 : 1

ผลการทดสอบหลังจากตัดตัวอย่างที่ไม่เหมาะสม กับรหัสโมเดล 29 รายการออก มิติของการวัดสามารถอธิบาย ค่าความแปรปรวนได้ร้อยละ 41.73 ค่าความแปรปรวนของ เศษเหลือที่สามารถอธิบายได้โดยองค์ประกอบแรก เท่ากับ ร้อยละ 3.51 และอัตราส่วนของความแปรปรวนทั้งสอง คือ 11.89 : 1 ซึ่งให้เห็นว่า แบบประเมินการทำงานของมือฉบับ ภาษาไทยมีความเป็นเอกมิติ

ความเที่ยงและอำนาจการจำแนก

ความเที่ยงและอำนาจการจำแนกของแบบประเมิน สำหรับกลุ่มตัวอย่างจำนวน 151 ราย พบว่า person

reliability มีค่า 0.95 และมี person separation มีค่า 4.46 ส่วนค่า item reliability มีค่า 0.94 และมีค่า item separation มีค่า 3.86 ซึ่งอยู่ในเกณฑ์ที่เหมาะสม (14) (ตารางที่ 5)

การเรียงลำดับความยากง่าย

การทดสอบการเรียงลำดับความยากง่ายของ ตัวเลือก 3 รายการคำตอบในแต่ละข้อคำถาม พบว่า ทุกข้อ มีการเรียงลำดับความสามารถในการทำงานของมือจาก รายการคำตอบ “ทำไม่ได้เลย” (1), “ยากเล็กน้อย-ยากมาก” (2) ไปหา “ไม่ยาก” (3) ตามความน่าจะเป็นของการ เรียงลำดับรายการคำตอบ เช่น ข้อคำถามที่ 3 มีค่าความ ยากง่ายของรายการคำตอบเท่ากับ -3.12, 0.33 และ 2.90 ตามลำดับ และข้อคำถามที่ 17.1 มีค่าความยากง่ายของ รายการคำตอบ -2.53, 0.90 และ 3.28 ตามลำดับ (รูปที่ 1)

ตารางที่ 4. การทดสอบความเป็นเอกมิติ

ค่าสถิติ	เกณฑ์	MNSQ ¹ (180 ราย)	MNSQ ¹ (151 ราย)
	(ร้อยละ)	5 รายการคำตอบ	3 รายการคำตอบ
ค่าความแปรปรวนที่อธิบายโดยมิติของการวัด	≥ 40	54.96	41.73
ค่าความแปรปรวนของเศษเหลือที่อธิบายได้โดยองค์ประกอบแรก	< 15	3.61	3.51
อัตราส่วนของความแปรปรวนในข้อ 1 และ 2	3 : 1	15.22 : 1	11.89 : 1

¹MNSQ = mean square

ตารางที่ 5. person reliability, item reliability และ separation statistics

การทดสอบความเที่ยง	เกณฑ์	180 ราย, 33 ข้อคำถาม	151 ราย, 33 ข้อคำถาม
		5 รายการคำตอบ	3 รายการคำตอบ
person reliability	≥ 0.80	0.93	0.95
item reliability	≥ 0.90	0.95	0.94
person separation	≥ 2.00	3.59	4.46
item separation	≥ 3.00	4.29	3.86

ITEM CATEGORY/OPTION/DISTRACTOR FREQUENCIES: MISFIT ORDER

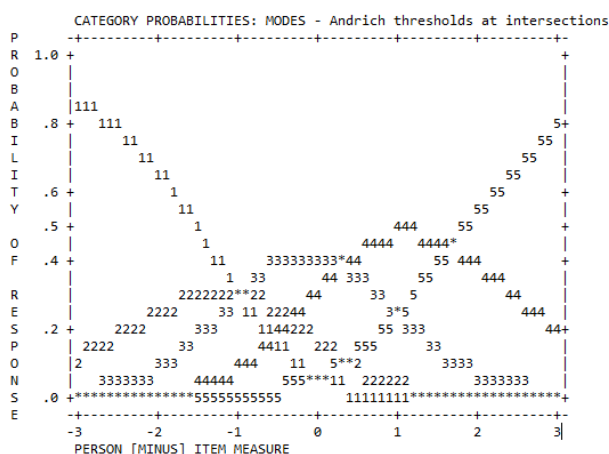
ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	ABILITY MEAN	S.E. P.SD	S.E. MEAN	INFT MNSQ	OUTF MNSQ	PTMA CORR.	ITEM
3	A 1	1	24	16	-3.12	1.75	.37	1.0	.8	-.57	item 3
	2	2	103	68	.33	1.91	.19	1.2	1.2	.08	
	3	3	24	16	2.90	1.59	.33	1.2	1.2	.47	
23	B 1	1	43	28	-2.53	2.05	.32	1.0	1.0	-.69	item 17.1
	2	2	91	60	.90	1.56	.16	.7	.5	.35	
	3	3	17	11	3.28	1.11	.28	1.2	1.2	.44	

รูปที่ 1. ตัวอย่างการเรียงลำดับความยากง่ายของตัวเลือก 3 รายการคำตอบในแต่ละข้อคำถาม

เส้นโค้งการเลือกรายการคำตอบ

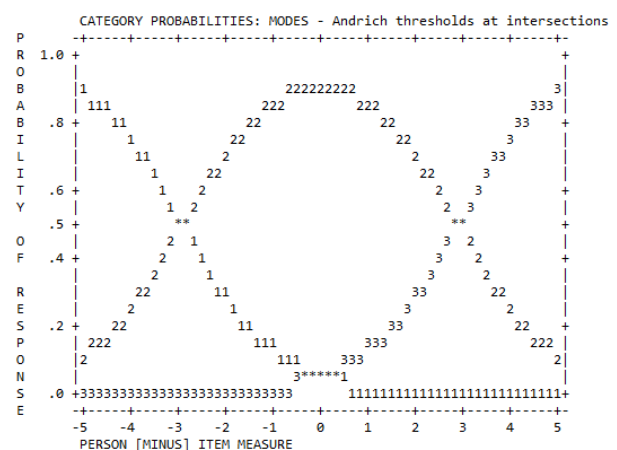
จากเส้นโค้งการเลือกรายการคำตอบ (CRC) จำนวน 5 รายการคำตอบของตัวอย่าง พบว่า ความน่าจะเป็นในการเลือกรายการคำตอบลำดับที่ 2 (ยากมาก) มีค่าต่ำกว่ารายการคำตอบอื่น ๆ โดยโอกาสสูงสุดเพียงร้อยละ 30 และมีความน่าจะเป็นในการเลือกรายการคำตอบลำดับที่ 2 ซ้อนทับกับตัวเลือกที่ 4 (ยากน้อย) และตัวเลือกที่ 3 (ยากปานกลาง) ซ้อนทับตัวเลือกที่ 4 ดังแสดงในรูปที่ 2 ดังนั้นจึงปรับปรุงรายการคำตอบ โดยรวมรายการคำตอบลำดับที่ 2, 3 และ 4 เข้าด้วยกัน และปรับเกณฑ์ให้คะแนน ดังนี้ โดย “ทำไม่ได้เลย” เท่ากับ 1 คะแนน “ยากเล็กน้อย, ยากปานกลาง, ยากมาก” เท่ากับ 2 คะแนน และ “ไม่ยาก” เท่ากับ 3 คะแนน

หลังจากนั้นทำการวิเคราะห์อีกครั้งพบว่า CRC มีความน่าจะเป็นในการเลือกรายการคำตอบแยกออกจากกัน อย่างชัดเจน กลุ่มตัวอย่างที่มีระดับความสามารถในการใช้



รูปที่ 2. เส้นโค้งการเลือกรายการคำตอบของคำถามที่มีตัวเลือก 5 รายการคำตอบ

มือทำกิจกรรมประจำวันระดับต่ำที่ -5.00 logits มีโอกาสตอบว่า “ทำไม่ได้เลย” มากที่สุด ประมาณร้อยละ 90 ทั้งนี้พบโอกาสที่จะเลือกตอบ “ยากเล็กน้อย, ยากปานกลาง, ยากมาก” ประมาณร้อยละ 10 และไม่พบโอกาสที่จะเลือกตอบว่า “ไม่ยาก” สำหรับกลุ่มตัวอย่างที่มีระดับความสามารถสูงที่ระดับคะแนน 5.00 logits พบโอกาสที่จะเลือกตอบว่า “ไม่ยาก” มากที่สุด ประมาณร้อยละ 90 พบโอกาสที่จะเลือกตอบ “ยากเล็กน้อย, ยากปานกลาง, ยากมาก” ประมาณร้อยละ 10 และไม่พบโอกาสที่จะเลือกตอบ “ทำไม่ได้เลย” ดังแสดงในรูปที่ 3



รูปที่ 3. เส้นโค้งการเลือกรายการคำตอบของคำถามที่มีตัวเลือก 3 รายการคำตอบ

แผนภาพบุคคลและคำถาม

แผนภาพบุคคลและคำถาม (item-person map) ของแบบประเมินการทำงานของมือฉบับภาษาไทย แสดงในรูปที่ 4 แผนภาพแนวดิ่งนี้มีระดับมาตรฐานวัดตั้งแต่ -7 จนถึง +5 โดยด้านขวาแสดงระดับความยาก-ง่ายของข้อคำถาม

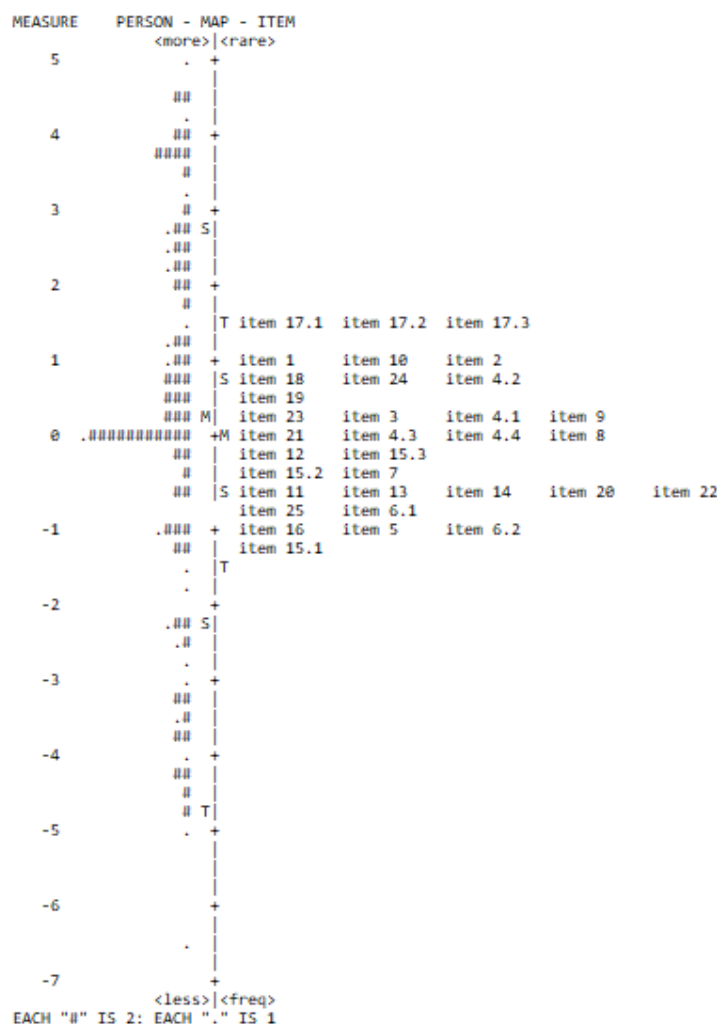
(item hierarchies) โดยคำว่า “rare” คือ ข้อคำถามที่ยาก และคำว่า “freq” คือ ข้อคำถามที่ง่าย ส่วนด้านซ้ายแสดงระดับความสามารถของกลุ่มตัวอย่าง (person ability) โดยคำว่า “more” คือ ความสามารถสูง คำว่า “less” คือ ความสามารถต่ำ คำ “M” คือค่าเฉลี่ยของกลุ่มตัวอย่างและข้อคำถาม “S” คือหนึ่งเท่าของส่วนเบี่ยงเบนมาตรฐานจากค่าเฉลี่ยของกลุ่มตัวอย่างและข้อคำถาม และคำ “T” คือสองเท่าของส่วนเบี่ยงเบนมาตรฐานจากค่าเฉลี่ยของตัวอย่างและข้อคำถาม เครื่องหมาย “#” คือจำนวนตัวอย่าง จำนวน 2 ราย และเครื่องหมาย “.” คือ จำนวนตัวอย่าง จำนวน 1 ราย

จากรูปที่ 4 กลุ่มตัวอย่างมีค่าเฉลี่ยความสามารถของมือในการทำกิจวัตรประจำวัน (person mean) สูงกว่าค่าเฉลี่ยระดับความยากของข้อคำถาม (item mean) เท่ากับ 0.18 logit (แสดงด้วยตำแหน่ง “M” บนสเกลด้านขวาและซ้ายตามลำดับ) ข้อคำถามมีระดับความยาก-ง่ายกระจุกตัว

เมื่อเทียบกับความสามารถในการทำงานของมือของกลุ่มตัวอย่างที่มีการกระจายตัวกว้างกว่า ตัวอย่างเช่น ข้อคำถามที่ 17.2 “การหมุนพวงมาลัยรถยนต์” และ 17.3 “การบิดปลดคันเร่งมอเตอร์ไซด์” เป็นข้อที่กลุ่มตัวอย่างเลือกรายการคำตอบ “ทำไม่ได้เลย” มากที่สุด ส่วนข้อคำถามที่ 15.1 “สวมเสื้อยืดคอกลม” เป็นข้อที่กลุ่มตัวอย่างเลือกรายการคำตอบ “ไม่ยาก” มากกว่าข้อคำถามอื่นๆ

ความเหมาะสมของบุคคล

ความเหมาะสมของบุคคล (person fit) เป็นการตรวจจึรูปแบบของคะแนน (item-score) ที่ไม่เป็นไปตามราสซ์โมเดลของกลุ่มตัวอย่าง โดยพบค่า infit mean-squares และ outfit mean-squares ที่ไม่เหมาะสมกับราสซ์โมเดลจำนวน 14 ราย (ร้อยละ 7.78) จากกลุ่มตัวอย่างจำนวน 180 ราย และถูกตัดออกจากการวิจัยครั้งนี้



รูปที่ 4. แผนภาพบุคคลและข้อคำถามที่มีตัวเลือก 3 รายการคำตอบ

ความเหมาะสมของการวัดระดับความสามารถในการใช้มือทำกิจวัตรประจำวันรายบุคคล

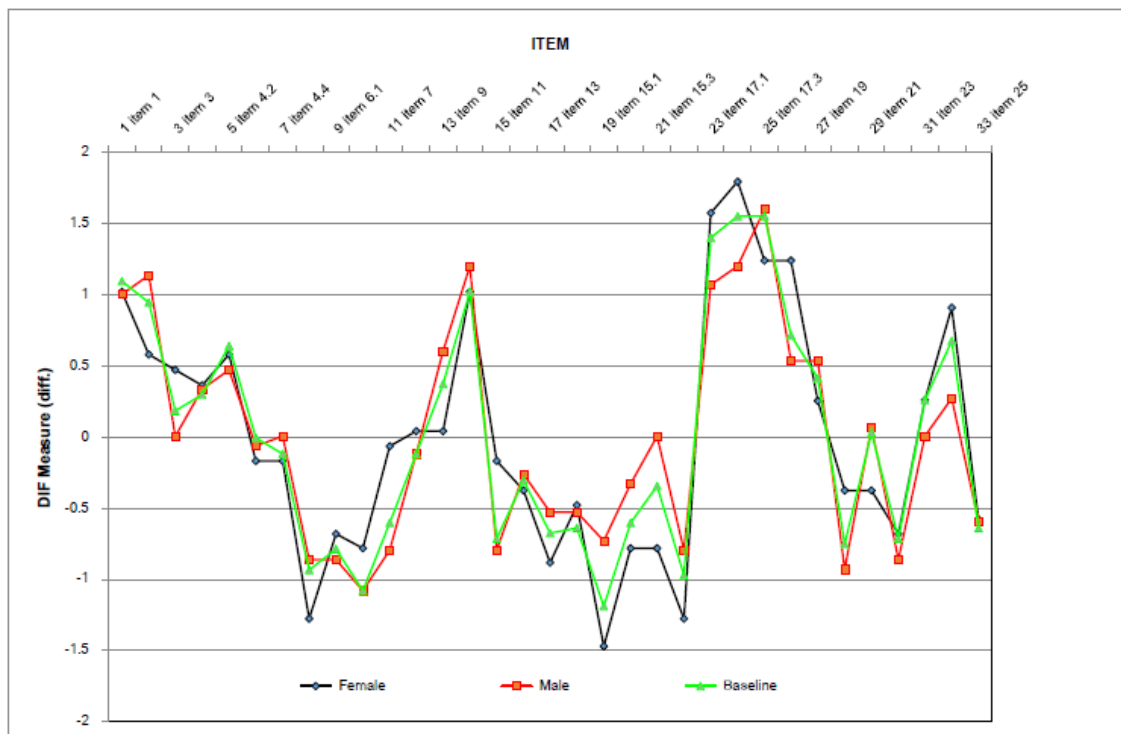
จากการวิเคราะห์ความสามารถในการใช้มือทำกิจวัตรประจำวันรายบุคคล (person measure) โดยค่า infit mean-squares และค่า outfit mean-squares พบว่ามีกลุ่มตัวอย่างที่มีความสามารถในการใช้มือทำกิจวัตรประจำวันในระดับคะแนนต่ำสุดของแบบประเมิน (minimum measure) จำนวน 14 ราย และความสามารถในการใช้มือทำกิจวัตรประจำวันในระดับคะแนนสูงสุดของแบบประเมิน (maximum measure) จำนวน 1 ราย ดังนั้น มีตัวอย่างที่มีความสามารถในการใช้มือทำกิจวัตรประจำวันรายบุคคลที่ไม่เหมาะสมกับรหัสโมเดลจำนวน 15 ราย (ร้อยละ 8.33) จากกลุ่มตัวอย่างจำนวน 180 ราย และถูกตัดออกจากการวิจัยครั้งนี้

ความเป็นอิสระของข้อคำถาม

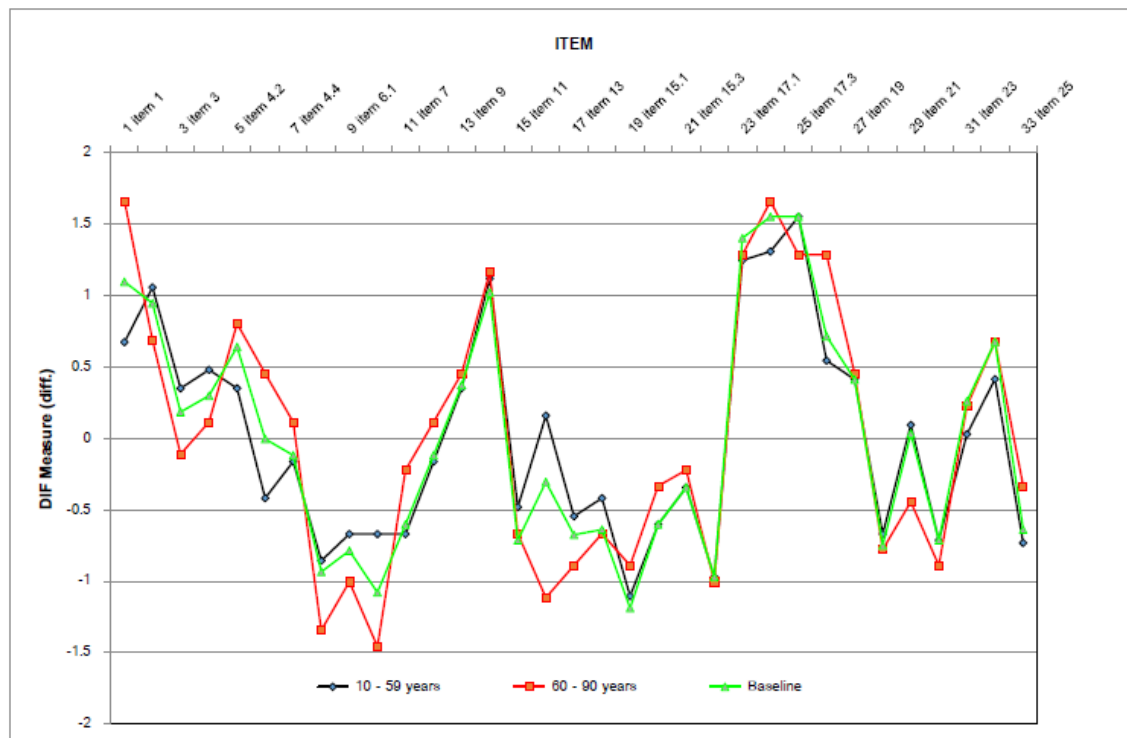
ความเป็นอิสระของข้อคำถามพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ของส่วนที่เหลือที่ได้รับการปรับค่าให้เป็นมาตรฐาน (intra-item correlations of standardized residuals) ของข้อคำถามหลังจากตัดตัวอย่างที่ไม่เหมาะสมกับรหัสโมเดล โดยค่าสหสัมพันธ์ดังกล่าวมีค่าระหว่าง -0.33 ถึง 0.71 ซึ่งแสดงให้เห็นว่า ข้อคำถามมีอิสระต่อกัน ถึงแม้บางข้อจะมีค่ามากกว่าเกณฑ์ที่เหมาะสมเพียงเล็กน้อย (เกณฑ์ค่าสหสัมพันธ์ < 0.70)

การทำหน้าที่ต่างกันของข้อคำถาม

การทดสอบการทำหน้าที่ต่างกันของข้อคำถามระหว่างกลุ่มตัวอย่าง (DIF contrast) จำแนกตามเพศ ส่วนใหญ่พบความแตกต่างของผลตอบสนองต่อข้อคำถามระหว่างเพศ < 0.50 logits ยกเว้นข้อคำถามที่ 2, 7, 9, 11, 15.1, 15.3, 17.1, 17.2, 18, 20, 24 ที่มีค่า ≥ 0.50 logits และเมื่อทดสอบด้วยสถิติ chi-squared พบว่ามีเพียงข้อคำถามที่ 15.1 “สวมเสื้อยืดคอกลม” มีการทำหน้าที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($P < 0.05$) ดังแสดงในรูปที่ 5 สำหรับการจำแนกตามกลุ่มอายุ ส่วนใหญ่พบความแตกต่างของผลตอบสนองต่อข้อคำถามระหว่างกลุ่มของตัวอย่าง < 0.50 logits ยกเว้นข้อคำถามที่ 1, 4.3, 6.2, 12, 18, 21 มีค่า ≥ 0.50 logits และเมื่อทดสอบด้วยสถิติ chi-squared พบว่ามีเพียงข้อคำถามที่ 16 “การใส่รองเท้า” และข้อคำถามที่ 19 “การจับและไขลูกบิดประตู” มีการทำหน้าที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($P < 0.05$) ดังแสดงในรูปที่ 6 ดังนั้นหากทดสอบการทำหน้าที่ต่างกันของข้อคำถามระหว่างกลุ่มตัวอย่างโดยตัวแปรตัวอื่น ๆ เช่น ศาสนา รายได้ ก็อาจพบข้อคำถามบางข้อที่แตกต่างจากข้างต้นแสดงค่านัยสำคัญทางสถิติ ซึ่งถือว่ามีจำนวนน้อยเมื่อเทียบกับข้อคำถามทั้งหมดและเทียบกับจำนวนตัวแปรต้นที่ใช้เปรียบเทียบ โดยรวมจึงถือว่า ข้อคำถามไม่แสดงการทำหน้าที่ที่แตกต่างกันของข้อคำถามระหว่างกลุ่ม



รูปที่ 5. การทำหน้าที่ต่างกันของข้อคำถามที่มีตัวเลือก 3 รายการคำตอบจำแนกตามเพศ



รูปที่ 6. การทำหน้าที่ต่างกันของข้อคำถามที่มีตัวเลือก 3 รายการคำตอบจำแนกตามอายุ

การอภิปรายผล

แบบประเมินการทำงานของมือฉบับภาษาไทยที่สร้างขึ้น มีการปรับข้อคำถามให้เข้ากับกิจวัตรประจำวันของคนไทย เพื่อให้สอดคล้องกับความเข้าใจของผู้ป่วยในการตอบแบบประเมินที่สอบถามโดยนักกายภาพบำบัด และมีความเที่ยงที่ดีโดยมีค่า person reliability เท่ากับ 0.93 เมื่อเทียบกับเกณฑ์ที่ระบุว่าไม่ควรน้อยกว่า 0.80 (14) นอกจากนี้แบบประเมินแสดงความเป็นเอกมิติและมีความเป็นอิสระของข้อคำถาม สำหรับค่า item fit statistics พบว่ามีข้อคำถามที่มีค่า infit mean-squares และ outfit mean-squares อยู่ในเกณฑ์เหมาะสม (0.50-1.50) จำนวน 32 ข้อ (ร้อยละ 96.97) อย่างไรก็ตามหลังจากตัดตัวอย่างที่ไม่เหมาะสมกับราชสีโมเดลออก และปรับตัวเลือกเป็น 3 รายการคำตอบ พบว่าข้อคำถามที่มีค่า infit mean-squares และค่า outfit mean-squares จัดอยู่ในเกณฑ์ที่เหมาะสมทุกข้อ และการเรียงลำดับความยากง่ายของตัวเลือก 3 รายการคำตอบในแต่ละข้อคำถามมีการเรียงลำดับตามที่คาดหวังทุกข้อ

การศึกษามีการปรับปรุงรายการคำตอบโดยรวมตัวเลือก “ยากเล็กน้อย, ยากปานกลาง และยากมาก” เข้าด้วยกัน เนื่องจากการวิเคราะห์การเลือกคำตอบแสดง

ความถี่การเลือกรายการคำตอบที่ไม่เท่าเทียม พบความถี่การตอบ “ยากมาก” ต่ำที่สุด และจาก CRC พบความน่าจะเป็นในการเลือกคำตอบ “ยากมาก” มีค่าต่ำกว่าตัวเลือกอื่นๆ และมีความน่าจะเป็นในการเลือกคำตอบซ้อนทับกับ “ยากเล็กน้อย” “ยากปานกลาง” และ “ยากมาก” เป็นส่วนใหญ่นั่นในการศึกษาครั้งต่อไปอาจพิจารณาลดจำนวนตัวเลือกรายการคำตอบของข้อคำถามลง ซึ่งจะทำให้สอดคล้องกับความเป็นจริงที่ว่า ผู้ป่วยมีความสามารถในการทำงานของมือในกิจวัตรประจำวันต่าง ๆ ระดับใด

ข้อคำถามทั้ง 33 ข้อถูกคงไว้ในแบบประเมินแม้ว่ามีคำถามที่ไม่เหมาะสมกับราชสีโมเดลจำนวน 1 ข้อ ได้แก่ ข้อคำถามที่ 18 “การล้างรถ และเช็ดทำความสะอาดรถ” ซึ่งมีค่า outfit mean-squares = 1.53 แต่ยังคงเก็บคำถามนี้ไว้ในแบบประเมิน เนื่องจากเมื่อตัดข้อคำถามทิ้งไปจะมีผลต่อความเที่ยงของแบบประเมิน และค่า mean-square ของข้อคำถามนี้ยังไม่มากกว่า 2.0 ดังนั้นการคงข้อคำถามนี้ไว้ก็ไม่ทำให้เกิดผลเสียหลายต่อผลลัพธ์ที่ประเมินได้จากข้อคำถามอื่น ๆ อย่างไรก็ตาม หลังจากตัดตัวอย่างที่ไม่เหมาะสมกับราชสีโมเดล 29 ราย พบว่า ข้อคำถามทุกข้ออยู่ในเกณฑ์ที่เหมาะสม

จากแผนภาพบุคคลและคำถาม ข้อคำถามมีระดับความยากระหว่าง -1.19 ถึง 1.55 logits และระดับ

ความสามารถในการทำงานของมือระหว่าง -6.62 ถึง 5.12 logits ดังนั้นความสามารถในการทำงานของมือของผู้ป่วยมีระดับสูงกว่าระดับความยากของข้อคำถามข้อที่ยากที่สุด และต่ำกว่าระดับความง่ายของข้อคำถามข้อที่ง่ายที่สุด สำหรับข้อคำถามที่มีระดับความยาก 2 ข้อ ได้แก่ ข้อคำถามที่ 17.2 “การหมุนพวงมาลัยรถยนต์” และข้อคำถามที่ 17.3 “การบิดปลดคันเร่งมอเตอร์ไซค์” ซึ่งผู้ป่วยที่มีความสามารถในการใช้มือสูงกว่าระดับความยากของข้อคำถามถูกจำกัดการตอบคำถามโดยเลือกรายการคำตอบ “ไม่ยาก” ในข้อคำถามนั้น ๆ โดยแท้ที่จริงมีความสามารถสูงกว่ากิจวัตรประจำวันที่ระบุ สำหรับข้อคำถามที่มีระดับความง่าย 1 ข้อ ได้แก่ ข้อคำถามที่ 15.1 “สวมเสื้อยืดคอกลม” ผู้ป่วยที่มีความสามารถในการใช้มือต่ำกว่าระดับความยากของข้อคำถามถูกจำกัดการตอบคำถามโดยเลือกรายการคำตอบ “ทำไม่ได้เลย” ในข้อคำถามนั้น ๆ โดยแท้ที่จริงมีความสามารถต่ำกว่ากิจวัตรประจำวันที่ระบุ อย่างไรก็ตามอาจมีผลจากข้อคำถามที่ยังไม่ครอบคลุมกับทุกอาชีพ เช่น ข้อคำถามที่ 22 “การกดเครื่องคิดเลข” ข้อคำถามที่ 23 “การใช้เมาส์” และข้อคำถามที่ 24 “การพิมพ์คอมพิวเตอร์” ซึ่งผู้ป่วยส่วนใหญ่ประกอบอาชีพเกษตรกรรมและไม่ได้ทำกิจวัตรประจำวันเหล่านี้จะคาดเดาการเลือกรายการคำตอบ หรือหากผู้ป่วยที่ไม่สามารถขยับได้จะคาดเดาการเลือกรายการคำตอบสำหรับข้อคำถามที่ 17.2 “การหมุนพวงมาลัยรถยนต์” และข้อคำถามที่ 17.3 “การบิดปลดคันเร่งมอเตอร์ไซค์” ดังนั้นการศึกษาครั้งต่อไปควรเพิ่มข้อคำถามที่ระบุกิจวัตรประจำวันที่มีความยาก-ง่ายเพิ่มขึ้นเพื่อให้มีความสัมพันธ์กับความสามารถของผู้ป่วย โดยคำนึงถึงกิจวัตรประจำวันที่ครอบคลุมทุกกลุ่มอาชีพ ทุกเพศ และทุกวัย

การทดสอบการทำหน้าที่ต่างกันของข้อคำถามระหว่างกลุ่มตัวอย่าง (DIF contrast) จำแนกตามเพศ ส่วนใหญ่พบความแตกต่างของผลตอบสนองต่อข้อคำถามระหว่างกลุ่มของตัวอย่าง < 0.50 logits ยกเว้นข้อคำถามที่ 15.1 “สวมเสื้อยืดคอกลม” พบความแตกต่างอย่างมีนัยสำคัญทางสถิติ อาจเนื่องมาจากผู้ป่วยไม่ได้ใช้มือทำกิจวัตรชีวิตประจำวันตามข้อคำถาม จึงอาจคาดเดาคำตอบส่งผลให้มีความแตกต่างกันระหว่างเพศชายและหญิง ดังนั้นการศึกษาครั้งต่อไปอาจต้องปรับหรือตัดข้อคำถามนี้ทิ้งไปหรือยกตัวอย่างในรายละเอียดของข้อคำถามเพื่อให้แบบประเมินสามารถใช้สำหรับผู้ป่วยทุกเพศ สำหรับการจำแนกตามกลุ่มอายุ ส่วนใหญ่พบความแตกต่างของผลตอบสนอง

ต่อข้อคำถามระหว่างกลุ่มของตัวอย่าง < 0.50 logits ยกเว้นข้อคำถามที่ 16 “การใส่รองเท้า” และข้อคำถามที่ 19 “การจับและไขลูกบิดประตู” พบความแตกต่างอย่างมีนัยสำคัญทางสถิติ ทั้งนี้อาจมีสาเหตุจากการแบ่งกลุ่มช่วงอายุเป็น 2 ช่วง ได้แก่ 10-59 ปี และ 60-90 ปี ซึ่งกลุ่มผู้สูงอายุ 60-90 ปี มีการทำกิจวัตรประจำวันที่แตกต่างกันกับกลุ่มช่วงอายุ 10-59 ปี สำหรับการใส่รองเท้าอาจมีลักษณะรองเท้าที่แตกต่างกัน หรือลักษณะของลูกบิดประตูที่ไม่เหมือนกันทำให้การจับลูกบิดแตกต่างกัน ส่งผลทำให้พบความแตกต่างระหว่างกลุ่มอายุ ดังนั้นการศึกษาครั้งต่อไปอาจต้องปรับข้อคำถาม เพื่อให้แบบประเมินสามารถใช้ประเมินสำหรับผู้ป่วยทุกช่วงวัย นอกจากนี้ควรวัดความสามารถในการทำงานของมือของผู้ป่วยซ้ำอีกครั้งเพื่อศึกษาการตอบสนองต่อการเปลี่ยนแปลง (responsiveness) ของแบบประเมินซึ่งเป็นอีกคุณสมบัติหนึ่งของแบบประเมิน นอกจากนี้ ในการใช้แบบวัด ควรให้ผู้ป่วยปฏิบัติตามกิจวัตรประจำวันต่าง ๆ ดังกล่าว พร้อมให้นักกายภาพบำบัดประเมินแทนการสอบถามจากผู้ป่วยรายงาน

สรุปผล

แบบประเมินการทำงานของมือฉบับภาษาไทย มีคำถามจำนวน 33 ข้อ การตรวจสอบคุณสมบัติการวัดด้วยการวิเคราะห์รหัสโมเดล พบว่า แบบประเมินการทำงานของมือฉบับภาษาไทยมีความเป็นเอกมิติ และความเป็นอิสระของข้อคำถาม มีความตรงและความเที่ยงที่ดี การศึกษาครั้งต่อไปควรเพิ่มข้อคำถามทั้งระดับยากขึ้นและง่ายขึ้นเพื่อให้ครอบคลุมระดับความสามารถในการทำงานของมือของกลุ่มตัวอย่าง

กิตติกรรมประกาศ

ขอขอบพระคุณนักกายภาพบำบัดและอาสาสมัครทุกคนที่ให้ความร่วมมืออย่างดีในการศึกษานี้ และขอบคุณทุนสนับสนุนเพื่อการศึกษาจากกองทุนพัฒนาบุคลากรของมหาวิทยาลัยราชภัฏอุบลราชธานี ที่สนับสนุนการวิจัยในการศึกษานี้ด้วย

เอกสารอ้างอิง

1. Gabel CP, Michener LA, Burkett B, Neller, A. The Upper Limb Functional Index: development and

- determination of reliability, validity, and responsiveness. *J Hand Ther* 2006; 19: 328-48.
2. Vincent JI, MacDermid JC, King GJ, Grewal R. Rasch analysis of the Patient Rated Elbow Evaluation questionnaire. *Health Qual Life Outcomes* 2015; 13 : 2-13.
3. Phongamwong C, Choosakde A. Reliability and validity of the Thai version of the Shoulder Pain and Disability Index (Thai SPADI). *Health Qual Life Outcomes* 2015; 13: 2-4.
4. Prior Y, Tennant A, Tyson S, Kjekken L, Hammond A. Measure of activity performance of the hand (MAP-Hand) questionnaire: linguistic validation, cultural adaptation and psychometric testing in people with rheumatoid arthritis in the UK. *BMC Musculoskeletal Disorders* 2018; 19: 1-12.
5. Rosales RS, Hidalgo YM, Morales LR, Atroschi I. Reliability and construct validity of the Spanish version of the 6-item CTS symptoms scale for outcomes assessment in carpal tunnel syndrome. *BMC Musculoskeletal Disorder* 2016; 17: 2-4.
6. John M, Angst F, Awiszus F, Pap G, MacDermid JC, Simmen BR. The Patient-Rated Wrist Evaluation (PRWE): cross-cultural adaptation into German and evaluation of its psychometric properties. *Clin Exp Rheumatol* 2008; 26: 1047-58.
7. Beirer M, Serly J, Vester H, Pförringer D, Crönlein M, Deiler S, et al. The Munich Wrist Questionnaire (MWQ)—development and validation of a new patient-reported outcome measurement tool for wrist disorders. *BMC Musculoskeletal Disord* 2016; 17: 1-8.
8. Rapipong J. Disabilities of the Arm, Shoulder and Hand (DASH) Thai version [online]. 2006 [cited May 13, 2019]. Available from: www.dash.iwh.on.ca/sites/dash/public/translations/DASH_Thai.pdf.
9. Upatham S, Kumnerddee W. Reliability of Thai Version Boston Questionnaire. *J Med Assoc Thai* 2008 ; 8: 1250-6.
10. Guntayoung C, Chinchai S. The content validity and test-retest reliability of the developmental visual perception test (DTVP-2) in Thai children. *Int J Med Pharm Sci* 2013; 3: 1-6.
11. Meijer RR, Sijtsma K. Methodology review: Evaluating person fit. *Appl Psychol Meas* 2001; 25: 107-35.
12. DeMars C. Measuring higher education outcomes with a multidimensional Rasch model. *J Appl Meas* 2004; 5: 350-61.
13. Linacre JM. Misfit diagnosis: Infit outfit mean-square standardized [online]. 2012 (cited May 7, 2019). Available from: www.winsteps.com/winman/diagnosingmisfit.htm.
14. Linacre JM. Reliability and separation of measures [online]. 1999 [cited May 7, 2019]. Available from: www.winsteps.com/winman/reliability.htm.
15. Smith EV. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 2002; 3: 205-31.
16. Linacre JM. Data variance explained by Rasch measures. *Rasch Measurement Transactions* 2006 ; 20: 1045.
17. Embetson SE, Reise SP. Item response theory for psychologists. Mahwah: New Jersey, L. Erlbaum Associates, 2000.
18. Wang Z, Zhou J, Luo X, Xu Y, She X, Chen L, et al. Rasch analysis of the Adult Strabismus Quality of Life Questionnaire (AS-20) among Chinese adult patients with strabismus. *PLOS ONE* 2015; 10: e0142188. doi:10.1371/journal.pone.0142188.
19. Bond T, Fox C. Applying the Rasch model: Fundamental measurement in the human sciences. 2nd ed. Mahwah: Lawrence Erlbaum Associates; 2007