

การประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ 4 พารามิเตอร์ในการพัฒนาคลังข้อสอบดนตรีไทย APPLYING 4-PARAMETER ITEM RESPONSE THEORY FOR DEVELOPING A TEST ITEM BANK OF THAI MUSIC

ราชนันท์ เจริญแก่นทราย* สมประสงค์ เสนารัตน์ และ เบญจมาภรณ์ เสนารัตน์

Rachun Charoenkaensai*, Somprasong Senarat and Benjamaporn Senarat

สาขาวิชาวิจัยและประเมินผลการศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏร้อยเอ็ด จังหวัดร้อยเอ็ด 45120

Program in Educational Research and Evaluation, Faculty of Education, Roi Et Rajabhat University,
Roi Et 45120, Thailand

*Corresponding author: E-mail: nametaey160@gmail.com

รับบทความ 15 กรกฎาคม 2563 แก้ไขบทความ 23 พฤศจิกายน 2563 ตอบรับบทความ 23 พฤศจิกายน 2563 เผยแพร่บทความ กรกฎาคม 2564

บทคัดย่อ

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อพัฒนาคลังข้อสอบดนตรีไทย ตัวอย่าง คือ นักศึกษาดนตรีศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏร้อยเอ็ด ปีการศึกษา 2562 จำนวน 200 คน ได้มาจากการสุ่มอย่างง่าย เครื่องมือที่ใช้ในการวิจัย คือ แบบทดสอบแบบเลือกตอบชนิด 4 ตัวเลือก วิเคราะห์ข้อมูลหาค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบด้วยโปรแกรม R แพคเกจ mirt, Nortest, Lavaan, Psych และ GPArotation ก่อนที่จะคัดเลือกข้อสอบเข้าคลัง ผลจากการวิจัย พบว่า คลังข้อสอบสร้างขึ้นโดยใช้โปรแกรมการจัดการฐานข้อมูล ด้วยโปรแกรม MySQL และ PhpMyAdmin สามารถเก็บข้อมูลได้ทั้งรูปภาพและตัวอักษร มีข้อสอบที่ผ่านเกณฑ์การคัดเลือกเข้าคลังข้อสอบ จำนวน 19 ข้อ ประกอบด้วย ข้อสอบที่มีความยากเหมาะสมตามระดับของทักษะที่ต้องการวัดจากง่ายไปหายาก คลังข้อสอบมีค่าพารามิเตอร์ของข้อสอบอยู่ในเกณฑ์ดี มีอำนาจจำแนกค่อนข้างสูง ข้อสอบมีระดับความยากอยู่ในเกณฑ์ปานกลางถึงง่าย มีค่าโอกาสการเดาต่ำ และค่าความสะเพร่าต่ำ ข้อสอบมีความตรงตามนิยามเนื้อหา และความตรงเชิงโครงสร้าง

คำสำคัญ: ดนตรีไทย, คลังข้อสอบ, ทฤษฎีการตอบสนองข้อสอบ 4 พารามิเตอร์

ABSTRACT

The purpose of this research was to develop a test item bank of Thai music. The samples, obtained through simple random sampling, consisted of 200 music education students in the Faculty of Education, Roi Et Rajabhat University in the academic year 2019. The instrument was a multiple-choice type of test with four options. Data were analyzed for test parameters according to item response theory, responding to the tests with the R package mirt, Nortest, Lavaan, Psych, and GPArotation package before selecting the items for the test item bank. Results showed that the test item bank was built using MySQL and PhpMyAdmin database management program which included images and text data. The 19 items were selected within acceptable ranges to their degrees of difficulty level that were able to measure the skills needed from easy to difficult levels. The test item bank had a good level of parameter which contained high discrimination. Item difficulties were between moderate to easy levels with low guessing possibility and low mistake possibility. The test also had content validity and construct validity.

Keywords: Thai Music, Test Item Bank, 4-Parameter Item Response Theory

บทนำ

การพัฒนาการเรียนรู้ของผู้เรียน นับว่าเป็นกระบวนการที่สำคัญยิ่งในการจัดการเรียนสอน การที่จะพัฒนาผู้เรียนได้นั้นจะต้องทราบถึงความรู้ความสามารถของผู้เรียนว่าบรรลุตามจุดมุ่งหมายการเรียนรู้มากน้อยเพียงใด วิธีการวัดและประเมินผลการเรียนรู้จึงเป็นการตรวจสอบการเรียนรู้ของผู้เรียน เป็นเครื่องมือชนิดหนึ่งที่น่านำมาใช้ในการวัดและประเมินผลการเรียนรู้ คือ แบบทดสอบ และในการวิเคราะห์คุณภาพตามทฤษฎีการตอบสนองข้อสอบ ด้วยโมเดลการตอบสนองข้อสอบ ซึ่งเป็นโมเดลที่วัดความเป็นระบบความสัมพันธ์ระหว่างตัวแปรอิสระที่ร่วมกันสำหรับทำนายตัวแปรตาม (ศิริชัย กาญจนวาสิ, 2555, หน้า 52)

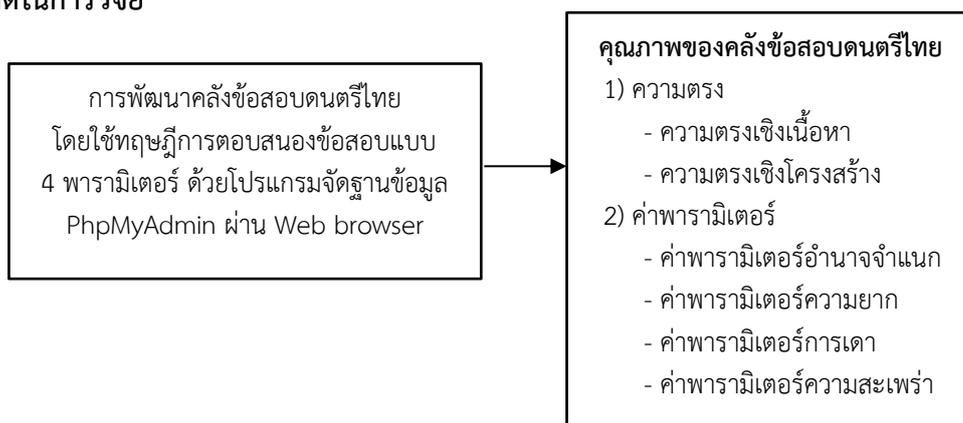
ทฤษฎีการตอบสนองข้อสอบอาจจะประกอบด้วยโมเดลการตอบสนองข้อสอบแบบหนึ่งพารามิเตอร์ แบบสองพารามิเตอร์ และแบบสามพารามิเตอร์ ซึ่งโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ เป็นโมเดลที่อธิบายความสัมพันธ์ระหว่างความสามารถที่มีอยู่ภายในบุคคล (Latent trait or Ability) กับผลการตอบสนองข้อสอบ หรือข้อคำถามโดยใช้โค้งลักษณะข้อสอบ (Item Characteristic Curve; ICC) ซึ่งมีการกำหนดลักษณะของข้อสอบด้วยพารามิเตอร์ความยาก (b) อำนาจจำแนก (a) โอกาสการเดาข้อสอบ (c) (ศิริชัย กาญจนวาสี, 2555, หน้า 52-53) ในปี 1981 Barton & Lord, (1981, p. 2) ได้เสนอแนวคิดเกี่ยวกับพารามิเตอร์ที่ 4 คือ ความสะเพร่า แต่ยังไม่ได้รับความสนใจ เนื่องจากยังไม่มีเครื่องมือที่ช่วยในการประมาณค่าความสามารถ ต่อมาแนวคิดของทฤษฎีการตอบสนองข้อสอบแบบพารามิเตอร์ที่ 4 ได้มีการนำมาพิจารณาอีกครั้งเนื่องจากวิทยาการทางคอมพิวเตอร์ที่ทันสมัยขึ้น ทำให้สามารถพัฒนาโปรแกรมทางสถิติที่มีความแม่นยำในการคำนวณค่าความสามารถในเรื่องของการคำนวณความสะเพร่าได้ (Magis, 2013, p. 304) ซึ่งการประมาณค่าความสามารถของผู้สอบด้วยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive testing: CAT) ในโมเดล 4 พารามิเตอร์ เหนือกว่าโมเดล 3 พารามิเตอร์ภายใต้จุดเริ่มต้นที่เหมือนกัน (Liao, Ho & Yen, 2012, p. 1686) อีกทั้งในยุคปัจจุบันเป็นยุคของการนำเอาคอมพิวเตอร์เข้ามาช่วยในการทำงานเก็บรวบรวมข้อมูล การค้นหาข้อมูล และการสร้างคลังข้อสอบ (คมสัน เอี่ยมจรัส, 2547, หน้า 28) ซึ่งคลังข้อสอบเป็นที่รวมของข้อสอบที่มีคุณภาพจำนวนมากสำหรับใช้วัดความรู้ความสามารถ และทักษะอันเป็นผลการเรียนรู้ของผู้เรียนตามจุดมุ่งหมายของวิชาและหลักสูตร (ศิริชัย กาญจนวาสี, 2552, หน้า 246) นอกจากนี้คลังข้อสอบยังเป็นเครื่องมือที่จำเป็นสำหรับการบริหารการสอบแบบเลือกตอบที่ถูกต้องและเชื่อถือได้ สามารถเก็บรายการข้อสอบไว้ และสามารถนำกลับมาใช้ได้ตามต้องการ (McDonald, 2002, p. 200 อ้างอิงจาก Rudner : 1998) ทำให้ไม่จำเป็นต้องเขียนข้อสอบใหม่ทั้งหมดสำหรับการสอบในทุก ๆ ครั้ง (McDonald, 2002, p. 200) ดังนั้น การประมาณค่าความสามารถของผู้สอบด้วยโมเดลการตอบสนองข้อสอบแบบ 4 พารามิเตอร์ จึงดีกว่าโมเดลแบบเดิม โดยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สามารถลดจำนวนข้อสอบ ให้เวลาน้อยลง การรายงานผลการทดสอบได้ทันที แต่การทดสอบจำเป็นที่จะต้องมีการสร้างคลังข้อสอบเพื่อรองรับการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และยังมีความสะดวกในการนำกลับมาใช้งานง่ายต่อการสืบค้น และเพิ่มเติมข้อสอบภายในหลัง

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องข้างต้น จึงเป็นเหตุผลให้ผู้วิจัยสนใจที่จะศึกษาการพัฒนาคลังข้อสอบดนตรีไทยสำหรับนักศึกษาสาขาวิชาดนตรีศึกษา มหาวิทยาลัยราชภัฏร้อยเอ็ด ซึ่งผลการวิจัยจะทำให้ได้ข้อสอบที่มีคุณภาพ จัดเก็บไว้ในคลังข้อสอบด้วยคอมพิวเตอร์แบบออนไลน์ เพื่อเตรียมรองรับการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ต่อไป

ความมุ่งหมายของการวิจัย

เพื่อพัฒนาคลังข้อสอบดนตรีไทย สำหรับนักศึกษาสาขาวิชาดนตรีศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏร้อยเอ็ด โดยใช้ทฤษฎีการตอบสนองข้อสอบแบบ 4 พารามิเตอร์

กรอบแนวคิดในการวิจัย



ภาพประกอบ 1 กรอบแนวคิดในการวิจัย การพัฒนาคลังข้อสอบดนตรีไทย

วิธีดำเนินการวิจัย

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าในครั้งนี้ เป็นนักศึกษาสาขาวิชาดนตรีศึกษาชั้นปีที่ 1-5 คณะครุศาสตร์ มหาวิทยาลัยราชภัฏร้อยเอ็ด ในปีการศึกษา 2562 จำนวน 218 คน

ตัวอย่าง

ตัวอย่างในการศึกษาค้นคว้าในครั้งนี้ เป็นนักศึกษาสาขาวิชาดนตรีศึกษา ชั้นปีที่ 1-5 คณะครุศาสตร์ มหาวิทยาลัยราชภัฏร้อยเอ็ด ในปีการศึกษา 2562 จำนวน 200 คน กำหนดขนาดตัวอย่างโดยใช้แนวคิดการกำหนดตัวอย่างของ Comrey & Lee (1992, p. 217) ตัวอย่าง จำนวน 200 คน เป็นจำนวนตัวอย่างที่อยู่ในเกณฑ์ที่รับได้ (Comrey & Lee, 1992, p. 217; Dennis, 2006, p. 50) ได้มาจากการสุ่มอย่างง่าย โดยการจับฉลาก

ขั้นตอนการพัฒนาคลังข้อสอบ

การพัฒนาคลังข้อสอบมีวิธีดำเนินการ ดังภาพประกอบ 2



ภาพประกอบ 2 ลำดับขั้นตอนการพัฒนาคลังข้อสอบ

จากภาพประกอบ 2 การพัฒนาคลังข้อสอบ แบบเลือกตอบ 4 ตัวเลือก มีวิธีสร้างเครื่องมือ ดังนี้

1. ศึกษาเอกสาร ทฤษฎี และงานวิจัยที่เกี่ยวข้องกับการสร้างแบบทดสอบ วิเคราะห์ สังเคราะห์ แนวคิด ทฤษฎี เอกสาร ตำรา บทความ งานวิจัยทั้งในประเทศและต่างประเทศ รวมทั้งฐานข้อมูลต่าง ๆ ในอินเทอร์เน็ตที่เกี่ยวข้องกับเนื้อหาวิชาดนตรีไทย ขั้นพื้นฐาน และได้ศึกษาการหาคุณภาพของข้อสอบ

2. สร้างแบบทดสอบ โดยการวิเคราะห์เนื้อหา เพื่อกำหนดนิยามเนื้อหาที่ต้องการวัดซึ่งผู้วิจัยได้กำหนดเนื้อหาตามความสำคัญของเนื้อหาดนตรีไทยซึ่งได้วิเคราะห์หลักสูตรดนตรีศึกษา มหาวิทยาลัยราชภัฏร้อยเอ็ด และเขียนข้อสอบตามนิยามเนื้อหาเป็นแบบเลือกตอบ 4 ตัวเลือก จำนวน 80 ข้อ

3. การหาคุณภาพของข้อสอบ การตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ จำนวน 5 ท่าน เป็นผู้เชี่ยวชาญด้านดนตรีศึกษา 3 ท่าน และด้านวัดและประเมินผล 2 ท่าน เพื่อพิจารณาความสอดคล้องและความเหมาะสมระหว่างข้อสอบและนิยามเนื้อหาวิชาดนตรีไทย คัดเลือกข้อสอบที่มีค่า IOC ตั้งแต่ 0.50 ขึ้นไป (เบญจมาภรณ์ เสนารัตน์ และสมประสงค์ เสนารัตน์, 2561, หน้า 110-111) นำไปพิมพ์เป็นแบบทดสอบฉบับสมบูรณ์

4. นำแบบทดสอบไปทดสอบกับตัวอย่าง จำนวน 200 คน ในวันที่ 22 เมษายน พ.ศ. 2563 โดยดำเนินการจัดการทดสอบออนไลน์โดยใช้ Google form และกำหนดเวลาในการทำข้อสอบจำนวน 1 ชั่วโมง เมื่อสิ้นสุดการทดสอบหรือครบกำหนดเวลา 1 ชั่วโมงแล้วระบบจะปิดอัตโนมัติ จากนั้นนำผลการทดสอบของนักศึกษา มาตรวจสอบความสมบูรณ์ของข้อมูล เพื่อนำไปใช้ในการวิเคราะห์ข้อมูลต่อไป

5. การวิเคราะห์ข้อมูล

5.1 ตรวจสอบการแจกแจงเป็นปกติของตัวอย่าง ด้วยสถิติ Lilliefors (Kolmogorov-Smirnov) normality test ด้วยโปรแกรม R (R Core Team, 2018) แพ็กเกจ Nortest (Gross & Ligges, 2015, pp. 4-5)

5.2 วิเคราะห์ค่าพารามิเตอร์ของข้อสอบ ด้วยทฤษฎีการตอบสนองข้อสอบแบบ 4 พารามิเตอร์ (IRT 4 PL) ด้วยโปรแกรม R แพ็กเกจ MIRT (Chalmers, 2020, pp. 95-117) วิเคราะห์ค่าพารามิเตอร์ของข้อสอบ โดยพิจารณาค่าอำนาจจำแนก (a) ค่าความยาก (d) ค่าโอกาสการเดา (g) และ ค่าความสับสน (u)

5.3 นำข้อสอบที่ผ่านเกณฑ์มาตรวจสอบหาความเป็นเอกมิติ โดยการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) โปรแกรม R แพ็กเกจ Psych (Revelle, 2020, pp. 111-122) และ GPArotation (Bernaards & Jennrich, 2015, p. 2) ด้วยวิธีสกัดองค์ประกอบขั้นต้น Minimum Likelihood หมุนแกนองค์ประกอบด้วยวิธี Quartimax

5.4 จากการวิเคราะห์องค์ประกอบเชิงสำรวจ นำมาตรวจสอบเป็นเอกมิติอีกครั้ง โดยวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) โดยใช้โปรแกรม R แพ็กเกจ Lavaan (Rossee, 2020, pp. 24-25)

6. การจัดทำคลังข้อสอบออนไลน์ โดยการออกแบบฐานข้อมูลสำหรับจัดเก็บข้อสอบและพารามิเตอร์ข้อสอบไว้ในคลังข้อสอบในระบบออนไลน์ ด้วยโปรแกรม MySQL และ PhpMyAdmin

7. คัดเลือกข้อสอบเข้าคลังโดยพิจารณาจากค่าพารามิเตอร์ของข้อสอบตามเกณฑ์ที่กำหนด โดยพิจารณาจากค่าพารามิเตอร์ของข้อสอบกับเกณฑ์ที่กำหนดไว้ คือ ค่าอำนาจจำแนก ต้องไม่ติดลบ มีค่าตั้งแต่ 0.800 ถึง 5.000 ค่าความยาก มีค่าตั้งแต่ -3.000 ถึง 3.000 ค่าโอกาสการเดา มีค่าไม่เกิน 0.300 และ ความสับสน มีค่าไม่ต่ำกว่า 0.800 สอดคล้องงานวิจัยของ Loken and Rulison (2010, p. 522)

ผลการวิเคราะห์ข้อมูล

การนำเสนอผลการวิจัยจะนำเสนอตามลำดับดังต่อไปนี้

1. ผลการวิเคราะห์ความสอดคล้องระหว่างข้อสอบกับนิยามเนื้อหาคนตรีไทย ผลการพิจารณา พบว่า ค่าความสอดคล้องระหว่างข้อสอบกับนิยามเนื้อหา ผ่านเกณฑ์ จำนวน 50 ข้อ มีค่า IOC อยู่ระหว่าง 0.60-1.00 มีค่า IOC เท่ากับ 1.00 จำนวน 39 ข้อ มีค่า IOC เท่ากับ 0.80 จำนวน 4 ข้อ และมีค่า IOC เท่ากับ 0.60 จำนวน 7 ข้อ ไม่ผ่านเกณฑ์ จำนวน 30 ข้อ

2. ผลตรวจสอบการวิเคราะห์ค่าสถิติพื้นฐาน การกระจายตัวของข้อมูล จากตัวอย่าง 200 คน ด้วยสถิติ (Kolmogorov-Smirnov) Normality Test วิเคราะห์การกระจายข้อมูลมีค่าพารามิเตอร์ดังนี้ ค่า D เท่ากับ 0.062 และ p-value เท่ากับ 0.061 และค่าสถิติเบื้องต้น พบว่า มีค่าเฉลี่ยเท่ากับ 28.77 ค่ามัธยฐานเท่ากับ 29.00 และฐานนิยมเท่ากับ 29.00 คะแนนตามลำดับ ค่าความเบ้เท่ากับ -0.003 ค่าความโด่ง เท่ากับ -0.420 เมื่อพิจารณาค่าสถิติเบื้องต้นมีค่าเกือบเท่ากัน ลักษณะของกราฟค่อนข้างสมมาตร แสดงว่าข้อมูลที่ได้จากอย่างมีการแจกแจงอย่างเป็นปกติ (เบญจมาภรณ์ เสนารัตน์ และสมประสงค์ เสนารัตน์, 2552, หน้า 33) ดังตาราง 1

ตาราง 1 ค่าสถิติเบื้องต้นของการกระจายของข้อมูล

n	Mean	Median	Mode	Min	Max	Skewness	Kurtosis
200	28.77	29.00	29.00	11	47	- 0.003	- 0.420

3. ผลการประมาณค่าพารามิเตอร์ของข้อสอบ พบว่า ข้อสอบ จำนวน 50 ข้อ มีค่าอำนาจจำแนกอยู่ระหว่าง -29.94 ถึง 48.31 ค่าเฉลี่ย เท่ากับ 8.36 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.05 ค่าความยาก อยู่ระหว่าง -33.28 ถึง 31.28 ค่าเฉลี่ย เท่ากับ -3.03 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 1.75 ค่าโอกาสการเดา อยู่ระหว่าง 0.00 ถึง 0.88 ค่าเฉลี่ย เท่ากับ 0.29 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.38 ค่าความสับสนอยู่ระหว่าง 0.21 ถึง 1.00 ค่าเฉลี่ย เท่ากับ 0.92 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.07 ดังตาราง 2

ตาราง 2 ตารางการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบแบบ 4 พารามิเตอร์

Item	ค่าพารามิเตอร์				item	ค่าพารามิเตอร์			
	a	d	g	u		a	d	g	u
1*	0.00	$-\infty$	1.00	0.00	26*	26.94	-11.32	0.40	1.00
2*	2.22	-2.52	0.78	1.00	27	2.90	2.78	0.00	1.00
3*	-21.03	31.28	0.88	1.00	28	1.30	-0.20	0.21	1.00
4*	33.10	-10.86	0.77	1.00	29*	3.25	-0.80	0.63	1.00
5	1.61	-1.00	0.00	1.00	30	1.16	-0.40	0.00	1.00
6	1.83	0.02	0.00	1.00	31	1.75	1.84	0.00	1.00
7*	-29.94	-26.55	0.35	0.56	32	2.48	0.30	0.17	0.91
8*	30.57	-3.21	0.86	1.00	33*	5.02	-1.95	0.41	0.78

ตาราง 2 (ต่อ)

Item	ค่าพารามิเตอร์				item	ค่าพารามิเตอร์			
	a	d	g	u		a	d	g	u
9*	2.05	-0.30	0.86	1.00	34	1.94	-0.12	0.00	0.91
10*	13.32	27.25	0.00	0.96	35*	4.50	-0.59	0.38	0.94
11*	3.50	3.47	0.57	1.00	36	4.92	-1.52	0.22	0.88
12*	18.71	19.10	0.34	0.76	37	1.61	1.66	0.00	0.91
13*	6.72	3.85	0.00	0.90	38	1.10	1.04	0.00	0.96
14*	23.66	-33.28	0.31	1.00	38*	45.60	-32.04	0.38	0.86
15	1.74	0.51	0.29	1.00	40*	3.32	2.17	0.55	0.96
16	2.33	-0.45	0.11	1.00	41*	30.67	-25.03	0.24	0.86
17*	43.47	-29.59	0.21	0.89	42*	48.31	-23.67	0.23	0.75
18*	14.31	22.18	0.29	0.57	43*	2.39	-1.93	0.35	1.00
19*	2.08	-0.14	0.47	1.00	44*	1.39	-1.11	0.63	1.00
20*	14.39	-29.93	0.09	1.00	45*	1.65	-0.31	0.37	1.00
21	1.75	1.03	0.00	1.00	46*	5.24	-3.62	0.54	1.00
22*	2.50	0.49	0.32	1.00	47	3.99	0.03	0.26	0.90
23	4.68	-0.68	0.20	0.80	48	1.43	0.17	0.00	1.00
24	1.07	-0.34	0.00	1.00	49*	4.17	6.17	0.00	0.21
25*	25.52	-30.70	0.27	0.90	50	2.30	-0.04	0.25	0.89
Minimum						-29.94	-33.28	0.00	0.21
Maximum						48.31	31.28	0.88	1.00
Mean						8.36	-3.03	0.29	0.92
Standard Deviation						0.05	1.75	0.38	0.07

a แทนค่าอำนาจจำแนก, d แทนค่าความยาก, g แทนค่าโอกาสการเดา, u แทนค่าความสะเพร่า

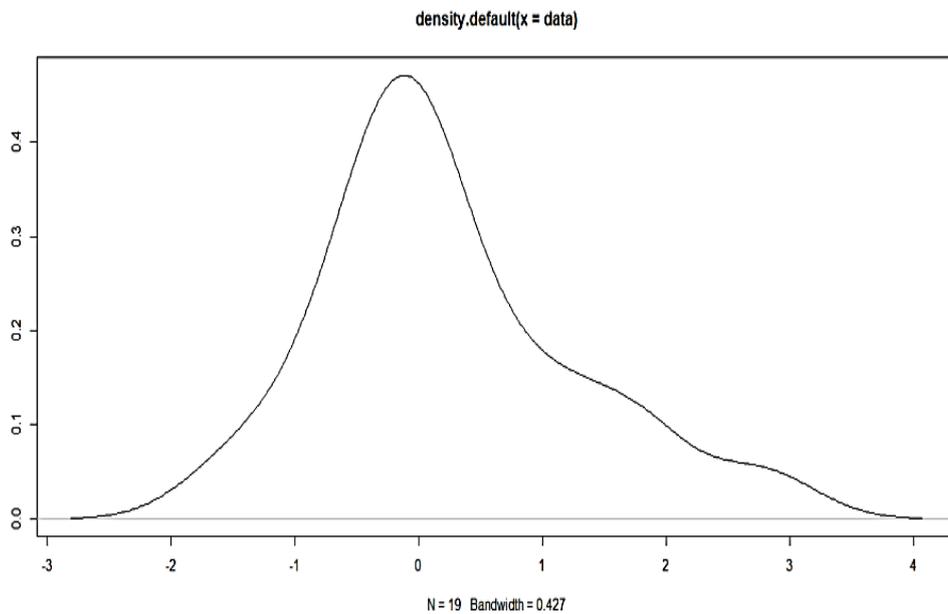
* ข้อสอบที่ไม่เป็นไปตามเกณฑ์ และถูกคัดออก

4. ผลการคัดเลือกข้อสอบ พบว่า มีข้อสอบที่เป็นไปตามเกณฑ์ที่ตั้งไว้ จำนวน 19 ข้อ โดยมีค่าอำนาจจำแนก อยู่ระหว่าง 1.07 ถึง 4.93 ค่าเฉลี่ย เท่ากับ 2.21 แสดงว่า ข้อสอบมีอำนาจจำแนกค่อนข้างสูง ค่าจุดตัดความยาก อยู่ระหว่าง -1.53 ถึง 2.79 ค่าเฉลี่ย เท่ากับ 0.24 แสดงว่า ข้อสอบมีระดับความยากอยู่ในเกณฑ์ปานกลางค่อนข้างง่าย และค่าโอกาสการเดา อยู่ระหว่าง 0 ถึง 0.29 ค่าเฉลี่ยเท่ากับ 0.09 และค่าความสะเพร่า อยู่ระหว่าง 0.81 ถึง 1 ค่าเฉลี่ยเท่ากับ 0.96 ดังตาราง 3

ตาราง 3 ตารางพิจารณาคุณภาพของข้อสอบดนตรีไทย

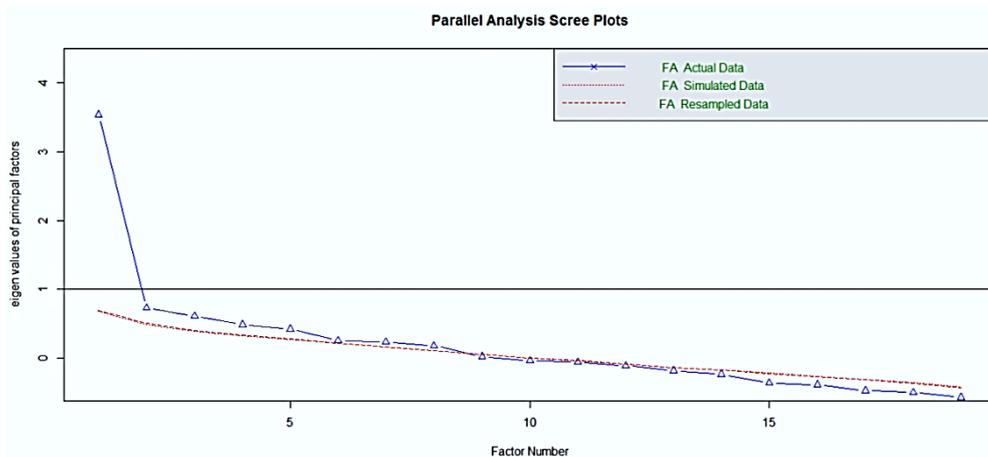
item	a	d	g	u	item	a	d	g	u
5	1.61	-1.00	0.00	1.00	31	1.75	1.84	0.00	1.00
6	1.83	0.02	0.00	1.00	32	2.48	0.30	0.17	0.91
15	1.74	0.51	0.29	1.00	34	1.94	-0.12	0.00	0.91
16	2.33	-0.45	0.11	1.00	36	4.92	-1.52	0.22	0.88
21	1.75	1.03	0.00	1.00	37	1.61	1.66	0.00	0.91
23	4.68	-0.68	0.20	0.80	38	1.10	1.04	0.00	0.96
24	1.07	-0.34	0.00	1.00	47	3.99	0.03	0.26	0.90
27	2.90	2.78	0.00	1.00	48	1.43	0.17	0.00	1.00
28	1.30	-0.20	0.21	1.00	50	2.30	-0.04	0.25	0.89
30	1.16	-0.40	0.00	1.00					
Minimum		1.07		-1.53		0.00		0.81	
Maximum		4.93		2.79		0.29		1.00	
Mean		2.21		0.24		0.09		0.96	
Standard Deviation		1.15		1.05		0.11		0.05	

เมื่อนำค่าพารามิเตอร์ความยากของข้อสอบที่เป็นไปตามเกณฑ์ มีค่าอยู่ระหว่าง -1.53 ถึง 2.79 ค่าความยากเฉลี่ย เท่ากับ 0.24 และมี ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 1.05 มาเขียนกราฟเพื่อแสดงถึงลักษณะการกระจายของค่าความยากของข้อสอบ พบว่า โค้งการแจกแจงของกราฟมีลักษณะค่อนข้างสมมาตร แสดงว่าข้อสอบในคลังข้อสอบมีความง่าย สามารถนำไปวัดทักษะพื้นฐานการเรียนรู้ดนตรีไทยได้ ดังภาพประกอบ 3



ภาพประกอบ 3 ค่าความยากพารามิเตอร์ของข้อสอบที่เป็นไปตามเกณฑ์

5. ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) ด้วยโปรแกรม R แพ็คเกจ Psych และ GPArotation โดยการนำข้อสอบที่ผ่านเกณฑ์มาตรวจสอบเชิงโครงสร้าง พบว่า ได้ผลการตรวจสอบความตรงเชิงโครงสร้างของข้อสอบด้วยการหาความเป็นเอกมิติด้วยวิธี Minimum Likelihood การหมุนแกนแบบ Quartimax ผลการตรวจสอบความตรงเชิงโครงสร้าง โดยพิจารณาค่า ERR ที่คำนวณจากอัตราส่วนของค่าไอเกน (Ratio of eigen ratio) โดยใช้ค่าตัวเลขที่ได้จากอัตราส่วนของอัตราส่วนค่าไอเกนขององค์ประกอบที่ 1 และ 2 กับค่าไอเกนขององค์ประกอบที่ 2 และ 3 โดยมีเกณฑ์การประเมินความเป็นเอกมิติจากค่า ERR ที่มีค่าตั้งแต่ 4 ขึ้นไป แบบสอปนั้น จึงจะมีความเป็นเอกมิติ (สุวิมล ตีรภานันท์, 2548, หน้า 11) จากการคำนวณค่า ERR จากค่า $E_1 = 3.54$, $E_2 = 0.73$ และ $E_3 = 0.61$ ได้ค่า $ERR = 4.03$ แสดงว่าแบบทดสอบมีความเป็นเอกมิติ และเป็นหลักฐานสนับสนุนเพื่อบ่งชี้ว่ามีความตรงเชิงโครงสร้าง (ศิริชัย กาญจนวาสี, 2552, หน้า 136) ดังภาพประกอบ 4



ภาพประกอบ 4 ภาพกราฟ Scree Plot

คลังข้อสอบดนตรีไทยโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบ 4 พารามิเตอร์
Item bank of Thai Music Examination using 4 parameter response theory
สาขาวิชาดนตรีศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏสุรินทร์

ออกแบบโดย: นายราชนันท์ ตรีคุณเกษียร นักศึกษาระดับปริญญาเอก สาขาวิชาวิจัยและประเมินผลการศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏสุรินทร์

ติดต่อสอบถาม: นายสำเริง คำมิ่งขันธ์ เจ้าหน้าที่ระบบ

หน้าหลัก | ข้อมูล | แสดงข้อสอบ | ผู้ใช้งานระบบ | ออกจากระบบ

แสดงข้อสอบ

id1=1 id2=0 กล้องเสียงของเครื่องดนตรีไทยที่ทำเป็นกระพุ้งจะพบในเครื่องดนตรีประเภทใดเป็นหลัก
ก. เครื่องพิณ
ข. เครื่องเสี่
ค. เครื่องพี่
ง. เครื่องเป่า

id1=2 id2=0 ถ้าได้รับเลือกให้เล่นดนตรีไทยแล้วเล่นเป็นทำนองหลักของวง เครื่องดนตรีที่จะได้บรรเลงคืออะไร
ก. อ้อวงใหญ่
ข. อ้อวงเล็ก
ค. ระนาดเอก
ง. ระนาดทุ้ม

ภาพประกอบ 6 ตัวอย่างหน้าจอคลังข้อสอบดนตรีไทย

สรุปผลการวิจัย

การพัฒนาคลังข้อสอบทักษะดนตรีไทย ด้วยทฤษฎีการตอบสนองข้อสอบแบบ 4 พารามิเตอร์ พบว่า คลังข้อสอบที่สร้างขึ้นมีคุณภาพ ในด้านความตรงเชิงเนื้อหา ความตรงเชิงโครงสร้าง และมีค่าพารามิเตอร์เป็นไปตามเกณฑ์ที่กำหนดไว้ จำนวน 19 ข้อ สามารถนำมาใช้เป็นฐานข้อมูลในการพัฒนาโปรแกรมทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้ คลังข้อสอบที่สร้างขึ้นใช้โปรแกรมการจัดการฐานข้อมูล MySQL และ PhpMyAdmin ผ่าน Web browser สามารถจัดเก็บข้อมูลได้ทั้งรูปภาพและตัวอักษร ประกอบด้วยข้อสอบที่มีความเหมาะสมตามระดับของทักษะที่ต้องการวัดจากง่ายไปหายาก และอัตราส่วนของข้อสอบ คลังข้อสอบมีค่าพารามิเตอร์ของข้อสอบอยู่ในเกณฑ์ดี มีอำนาจจำแนกค่อนข้างสูง ข้อสอบมีระดับความยากอยู่ในเกณฑ์ปานกลางถึงง่าย มีค่าโอกาสการเดา และค่าโอกาสสะเพร่าต่ำ

อภิปรายผลการวิจัย

1. ผลการวิเคราะห์ความสอดคล้องระหว่างข้อสอบกับนิยามเนื้อหาดนตรีไทย ผลการพิจารณา พบว่า ค่าความสอดคล้องระหว่างข้อสอบกับนิยามเนื้อหา ผ่านเกณฑ์ จำนวน 50 ข้อ ไม่ผ่านเกณฑ์ จำนวน 30 ข้อ ข้อสอบที่ผ่านเกณฑ์มีค่า IOC อยู่ระหว่าง (0.60-1.00) เนื่องจากกระบวนการสร้างข้อสอบของผู้วิจัยเอง อาจจะไม่ละเอียดเพียงพอ เพราะผู้วิจัยเริ่มฝึกในการเขียนข้อสอบ ทำให้ขาดความชำนาญในการเขียน และความรอบคอบในการตรวจสอบ ไม่สามารถเขียนข้อสอบให้สอดคล้องกับนิยามเนื้อหา ถึงทำให้ข้อสอบไม่ผ่านตามเกณฑ์จำนวนหลายข้อ แต่อย่างไรก็ตามผู้วิจัยได้เขียนสอบไว้มากกว่าที่ต้องการร้อยละ 60 ซึ่งสอดคล้องกับแนวคิดของ เบญจมาภรณ์ เสนารัตน์ และสมประสงค์ เสนารัตน์ (2561, หน้า 66) ที่กล่าวไว้ว่าจำนวนข้อสอบที่ต้องการควรรอกข้อสอบเมื่อไว้อย่างน้อยร้อยละ 25

2. ผลการวิเคราะห์คุณภาพของข้อสอบ จำนวน 50 ข้อ พบว่า มีอำนาจจำแนก อยู่ระหว่าง -29.94 ถึง 48.31 ค่าอำนาจจำแนกเฉลี่ย เท่ากับ 8.36 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.05 แสดงว่า ข้อสอบมีอำนาจจำแนกค่อนข้างต่ำ ส่วนค่าความยาก มีค่าอยู่ระหว่าง -33.28 ถึง 31.28 ค่าความยากเฉลี่ย -3.03 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 1.75 แสดงว่าข้อสอบค่อนข้างง่ายถึงปานกลาง ค่าโอกาสการเดา มีค่าอยู่ระหว่าง 0.00 ถึง 0.88 ค่าโอกาสการเดาเฉลี่ย 0.29 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.38 แสดงว่าข้อสอบมีค่าการเดาต่ำ สำหรับค่าความสะเพร่า มีค่าอยู่ระหว่าง 0.21 ถึง 1.00 ค่าสะเพร่าเฉลี่ย 0.92 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.07 แสดงว่าข้อสอบมีค่าความสะเพร่าต่ำ ทั้งนี้เนื่องจากโมเดลการตอบสนองข้อสอบแบบ 4 พารามิเตอร์ ให้สารสนเทศมากกว่าโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ซึ่งสอดคล้องกับผลการวิจัยของ Liao, Ho & Yen (2012,

p. 1686) ที่พบว่า การประมาณค่าความสามารถของผู้สอบ ที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้คลังข้อสอบสร้างขึ้นด้วยโมเดลตอบสนองข้อสอบแบบ 4 พารามิเตอร์ จะประมาณค่าความสามารถได้ถูกต้องมากกว่าการทดสอบแบบปรับเหมาะที่ใช้คลังข้อสอบที่สร้างขึ้นด้วยโมเดลตอบสนองข้อสอบแบบ 3 พารามิเตอร์

3. ผลการคัดเลือกข้อสอบเข้าคลังข้อสอบ พบว่า ข้อสอบที่ผ่านการวิเคราะห์ค่าพารามิเตอร์ จำนวน 50 ข้อ ถูกคัดเลือกเข้าคลัง จำนวน 19 ข้อ ค่าพารามิเตอร์ของข้อสอบ มีค่าอำนาจจำแนก อยู่ระหว่าง 1.07 ถึง 4.93 ค่าเฉลี่ย เท่ากับ 0.209 แสดงว่า ข้อสอบมีอำนาจจำแนกค่อนข้างสูง ค่าจุดตัดความยาก อยู่ระหว่าง -1.53 ถึง 2.79 ค่าเฉลี่ย เท่ากับ 0.243 แสดงว่า ข้อสอบมีระดับความยากอยู่ในเกณฑ์ปานกลางค่อนข้างง่าย ทั้งนี้เนื่องจาก เนื้อหาดนตรีไทยเนื้อหาส่วนมากเป็นเนื้อหาที่ไม่ยากหรือง่ายเกินไป ผู้สอบทั่วไปสามารถทำข้อสอบได้ ซึ่งสอดคล้องกับงานวิจัยของ ประยุทธ์ ไทยธานี (2546, หน้า 282) ที่ได้ทำวิจัยเรื่อง การสร้างและพัฒนาแบบสอบความถนัดทางดนตรีไทย พบว่า แบบทดสอบมีค่าความยากอยู่ระหว่าง 0.21-0.78 มีค่าความยากเฉลี่ยเท่ากับ 0.61 ซึ่งถือว่าเป็นข้อสอบที่มีความยากปานกลางถึงค่อนข้างง่าย

4. ผลการจัดทำคลังข้อสอบ พบว่า คลังข้อสอบประกอบด้วยรายการข้อคำถาม ตัวเลือก คำตอบ ค่าพารามิเตอร์ ได้แก่ ค่าอำนาจจำแนก ค่าความยาก ค่าการโอภาสการเดา และค่าความสะเพร่า การบริหารคลังข้อสอบผ่าน Web Browser ได้ด้วย Google Chrome ที่พัฒนาด้วยโปรแกรมการจัดการฐานข้อมูล PhpMyAdmin ใช้ภาษา PHP เพื่อใช้ในการบริหารจัดการฐานข้อมูล MySQL สามารถบันทึก แก้ไข เพิ่ม แสดงฐานข้อมูล และจำนวนข้อสอบได้ สะดวกและเหมาะสมต่อการใช้งาน เป็นโปรแกรมที่จัดการฐานข้อมูลที่ได้รับความนิยมอย่างแพร่หลายเนื่องจากเป็นโปรแกรมฟรี และใช้งานได้สะดวก ซึ่งสอดคล้องกับงานวิจัยของ เบญจมาภรณ์ เสนารัตน์ (2558, หน้า 126-134) ที่ได้ศึกษาวิจัย เรื่อง การพัฒนาเครื่องมือประเมินความสามารถทางการวิจัยการศึกษาของนักศึกษาครู มหาวิทยาลัยราชภัฏในภาคตะวันออกเฉียงเหนือ พบว่า เครื่องมือที่พัฒนาด้วยโปรแกรมการจัดการฐานข้อมูล PhpMyAdmin ใช้ภาษา PHP เพื่อใช้ในการบริหารจัดการฐานข้อมูล MySQL สามารถดูข้อสอบผ่าน Web Browser ได้ด้วย Google Chrome สามารถบันทึก แก้ไข เพิ่ม แสดงฐานข้อมูล และจำนวนข้อสอบได้

ข้อเสนอแนะ

1. ข้อเสนอแนะในการวิจัยครั้งต่อไป

1.1 ในการวิจัยครั้งต่อไปควรขยายขอบเขตของประชากรเพิ่มเติมในกลุ่มภาคตะวันออกเฉียงเหนือ หรือในระดับประเทศ และพัฒนาการใช้เกณฑ์พิจารณาจากคุณภาพของทฤษฎีตอบสนองข้อสอบ 4 พารามิเตอร์ เพิ่มเป็น 5 พารามิเตอร์

1.2 ควรมีการทำวิจัยเพื่อสร้างคลังข้อสอบเรื่องการกำหนดเกณฑ์การประเมินทักษะปฏิบัติทางด้านดนตรีไทย หรือในเนื้อหาอื่น ๆ เกี่ยวกับดนตรีไทย

2. ข้อเสนอแนะสำหรับการนำไปใช้ประโยชน์

ครูผู้สอนสามารถนำเอาข้อสอบที่ผู้วิจัยพัฒนาขึ้นไปประยุกต์ใช้ในการทดสอบทั้งในรูปแบบใช้กระดาษเขียนตอบ (Paper Pencil Test) ได้ หรือจะใช้การทดสอบบนคอมพิวเตอร์ได้ ทั้งนี้สามารถดูค่าความยากของข้อสอบประกอบการตัดสินใจนำไปใช้งานให้เหมาะสมกับความสามารถของนักศึกษา

เอกสารอ้างอิง

- คมสัน เอี่ยมจรัส. (2547). *การสร้างและพัฒนาโปรแกรมระบบการทดสอบ*. วิทยานิพนธ์ วท.ม. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.
เบญจมาภรณ์ เสนารัตน์. (2558). *การพัฒนาเครื่องมือประเมินความสามารถทางการวิจัยการศึกษาของนักศึกษาครูมหาวิทยาลัยราชภัฏในภาคตะวันออกเฉียงเหนือ*. วิทยานิพนธ์. ป.ด. มหาสารคาม: มหาวิทยาลัยมหาสารคาม.
เบญจมาภรณ์ เสนารัตน์ และ สมประสงค์ เสนารัตน์. (2552). *แผนแบบการทดลองแบบสุ่มสมบูรณ์*. *วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม*, 15, 25-36.
_____. (2561). *หลักการวัดและประเมินผลทางการศึกษา*. (พิมพ์ครั้งที่ 4). มหาสารคาม: อภิชาติการพิมพ์.
ประยุทธ์ ไทยธานี. (2546). *การสร้างและพัฒนาแบบสอบความถนัดทางดนตรีไทย*. วิทยานิพนธ์. ป.ด. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.
สุวิมล ติरणานนท์. (2548). *การตรวจสอบความเป็นเอกมิตติ*. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
ศิริชัย กาญจนวาสี. (2552). *ทฤษฎีการทดสอบแบบดั้งเดิม*. (พิมพ์ครั้งที่ 6). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
_____. (2555). *ทฤษฎีทดสอบแนวใหม่*. (พิมพ์ครั้งที่ 4). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

- Barton, M. A., & Lord, F. M. (1981). *An upper asymptote for the three-parameter logistic item-response model*. Princeton, NJ: Educational Testing Service.
- Bernaards, C. & Jennrich, R. (2015). *Package 'GPARotation'*. Retrieved June, 3, 2020, from <https://cran.r-project.org/web/packages/GPARotation/GPARotation.pdf>
- Chalmer, R. P. (2020). *Package 'mirt'*. from <http://cran.rproject.org/web/package/mirt.pdf>. June 3rd, 2020.
- Comrey, A. L. and Lee, H. B., (1992), *A first course in factor analysis*, Hillsdale, New Jersey: Erlbaum.
- Dennis, C. (2006). *The Essentials of Analysis. Third edition published by Continuum*, https://books.google.co.th/books?id=rQ2vdJgohH0C&pg=PA50&lpg=PA50&dq=comrey+and+lee+1992+sample+size&source=bl&ots=mX9lHskT3Q&sig=ACfU3U0Y2JTPlBZOWY2O_nNEBE6oeJnkhQ&hl=th&sa=X&ved=2ahUKEwj64e6F_uXpAhXK7XMBHQDsCKAQ6AEwEXoECAoQAAQ#v=onepage&q=comrey%20and%20lee%201992%20sample%20size&f=false. June 3rd, 2020.
- Gross, J. & Ligges, U. (2015). *Package 'nortest'*. from <http://cran.rproject.org/web/packages/nortest/nortest.pdf>. June 3rd, 2020.
- Liao, W.-W., Ho, R.-G., Yen, Y.- C., & Cheng, H.-C. (2012). The four-parameter logistic item response theory model as a robust method of estimating ability despite aberrant responses. *Social Behavior and Personality*, 40(10), 1679-1694.
- Loken, E. & Rulison, K. L. (2010). Estimation of a 4-parameter Item Response Theory model. *The British Journal of Mathematical and Statistical Psychology*, 63(3), 509-525.
- Magis, D. (2013). *A Note on the Item Information Function of the Four-Parameter Logistic Model*. Department of Education (B32), University of Lie`ge, Boulevard du Rectorat 5, B-4000 Lie`ge, Belgium. *Applied Psychological Measurement*, 37(4) 304–315.
- McDonald, M. E. (2002). *Systematic Assessment of Learning Outcomes: Developing Multiple-Choice Exams*. Published by Jones and Bartlett Publishers.
- R Core Team. (2018). *R: A Language and environment for statistical computing. [Computer software]*. from <https://cran.r-project.org/>. June 12th, 2020.
- Revelle, W. (2020). *Package 'psych'*. from <http://personality-project.org/r/psych-manual.pdf>. June 3rd, 2020.
- Rossee, Y. (2020). *Package 'lavaan'*. from <http://cran.r-project.org/web/packages/lavaan/lavaan.pdf>. June 3rd, 2020.