การจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยากของข้อสอบ : โดยใช้ทฤษฎีการตอบสนองข้อสอบ

Item Bank Classified by Content and Difficulty Levels: Using Item Response Theory

สุรชัย รักสมบัติ[1*] ปิยะทิพย์ ประดุจพรม[2] และ กนก พานทอง[2]
Surachai Raksombat[1*] Piyathip Pradujphom[2] and Kanok Panthong[2]
(Received: January 22, 2020 ; Revised: March 5, 2020 ; Accepted: March 16, 2020)

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์คุณภาพข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ และจัดคลังข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 จำแนกตามเนื้อหาและระดับความยาก วิธีการดำเนินการวิจัยแบ่งเป็น 4 ขั้นตอน คือ 1) วิเคราะห์คุณภาพข้อสอบตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ 2) จำแนกระดับความยากของข้อสอบตามค่าพารามิเตอร์ความยากของข้อสอบ 3) ประเมินความตรงเชิงโครงสร้างและพฤติกรรมที่วัดของข้อสอบ และ 4) จัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยากของข้อสอบ

ผลการวิจัยพบว่า

1. ข้อสอบที่ผ่านเกณฑ์การวิเคราะห์คุณภาพข้อสอบตามหลักการทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ มีค่าพารามิเตอร์ความยาก ตั้งแต่ -1.19 ถึง 2.49 (Mean = 1.20, SD = 0.81) ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ ตั้งแต่ 0.51 ถึง 2.42 (Mean = 1.15, SD = 0.40) ค่าพารามิเตอร์การเดาของข้อสอบ ตั้งแต่ 0.08 ถึง 0.30 (Mean = 0.23, SD = 0.03)

2. ข้อสอบที่ผ่านเกณฑ์การวิเคราะห์คุณภาพจำแนกเป็น 3 ระดับ ดังนี้ ง่าย ปานกลาง และยาก ระดับแรก ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ -1.19 ถึง 0.50 เป็นข้อสอบที่ง่าย ระดับที่สอง ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ 0.51 ถึง 1.50 เป็นข้อสอบปานกลาง และระดับสุดท้าย ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ 1.51 ถึง 2.49 เป็นข้อสอบที่ยาก

[1] นิสิตปริญญาเอก สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

[2] อาจารย์ วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

[1] Doctoral degree Student, Research and Statistics in Cognitive Science, College of Research Methodology and Cognitive Science, Burapha University

[2] Lecturer, College of Research Methodology and Cognitive Science, Burapha University

* Corresponding Author E-mail: 59810027@go.buu.ac.th, su_ra_ch@yahoo.co.th

3. ข้อสอบผ่านเกณฑ์การประเมินความตรงเชิงโครงสร้างและด้านพฤติกรรมที่วัด มีความตรงเชิงโครงสร้าง (Mean = 4.69, SD = 0.12) และพฤติกรรมที่วัด (Mean = 4.76, SD = 0.10) โดยรวมมีความสอดคล้องอยู่ในระดับมากที่สุด

4. คลังข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 จำแนกตามเนื้อหาและระดับความยากของข้อสอบ มีจำนวนข้อสอบทั้งหมด 345 ข้อ (ข้อสอบที่นำมาใช้ในการวิจัยทั้งสิ้น 480 ข้อ) จำแนกได้ 3 สาระ คือ สาระที่ 1 จำนวนและพีชคณิต มีจำนวน 238 ข้อ (ข้อสอบที่ง่าย 32 ข้อ ข้อสอบปานกลาง 105 ข้อ และข้อสอบที่ยาก 101 ข้อ) สาระที่ 2 การวัดและเรขาคณิต มีจำนวน 80 ข้อ (ข้อสอบที่ง่าย 20 ข้อ ข้อสอบปานกลาง 32 ข้อ และข้อสอบที่ยาก 28 ข้อ) และสาระที่ 3 สถิติและความน่าจะเป็น มีจำนวน 27 ข้อ (ข้อสอบที่ง่าย 6 ข้อ ข้อสอบปานกลาง 10 ข้อ และข้อสอบที่ยาก 11 ข้อ)

คำสำคัญ คลังข้อสอบ ระดับความยาก ทฤษฎีการตอบสนองข้อสอบ

Abstract

This study aimed to analyze the quality of test items in mathematics using the item response theory with the three-parameter model in mathematics of grade 7, and to organize the item bank for grade 7, classified by content and difficulty levels. The research methodology comprised 4 steps: 1) analysis of the quality of test items using the item response theory with the three-parameter model; 2) classification of difficulty levels according to the item difficulty parameter; 3) assessment of the construct validity and behavioral measures; and 4) organizing the item bank, classified by content and difficulty levels.

The results showed that:

1. The successful questions analyzed with the item response theory with three-parameter model had the item difficulty parameters ranging from -1.19 to 2.49 (Mean = 1.20, SD = 0.81), the item discrimination parameters ranging from 0.51 to 2.42 (Mean = 1.15, SD = 0.40), and the item guessing parameters ranging from 0.08 to 0.30 (Mean = 0.23, SD = 0.03).

2. The quality questions were divided into three levels: easy, moderate and difficult. In the first level, the easy level, the item difficulty parameters ranged from -1.19 to 0.50. In the second level, the moderate level, the item difficulty parameters ranged from 0.51 to 1.50. And finally, in the third level, the difficult level, the item difficulty parameters ranged from 1.51 to 2.49.

3. The successful questions that had been assessed on the construct validity and behavioral measures yielded the following data: the construct validity (Mean = 4.69, SD = 0.12) and behavioral measures (Mean = 4.76, SD = 0.10); on the whole, the consistency was in the highest level.

4. The item bank of grade 7 mathematics, classified by content and difficulty levels, had a total of 345 successful items (The total test items used in the research was 480). They could be divided into three learning substances: the first learning substance was number and algebra which had 238 items (32 easy, 105 moderate and 101 difficult); the second learning substance was measurement and geometry which had 80 items (20 easy, 32 moderate and 28 difficult); and the third learning substance was statistics and probability which had 27 items (6 easy, 10 moderate and 11 difficult).

Keywords: item bank, difficulty level, item response theory

บทนำ

หลักสูตรการศึกษาทั้งในประเทศ และต่างประเทศให้ความสำคัญกับวิชาคณิตศาสตร์เป็นอย่างมาก เด็กทุกคนเริ่มเรียนคณิตศาสตร์ตั้งแต่เข้าศึกษาในระบบ และเรียนอย่างต่อเนื่องยาวนานจนเข้ามหาวิทยาลัย สิ่งสำคัญที่เกิดควบคู่กับการเรียนการสอนคณิตศาสตร์ คือ การวัดและประเมินผลการเรียนรู้ (ศศิธร จันทรมหา, 2560) จากผลการศึกษาคะแนนผลการสอบ O-NET ในวิชาคณิตศาสตร์ ที่ทดสอบความรู้ทางการศึกษาระดับชาติขั้นพื้นฐาน ปีการศึกษา 2561 ทั่วประเทศ ชั้นประถมศึกษาปีที่ 6 มีคะแนนเฉลี่ยร้อยละ 37.50 ชั้นมัธยมศึกษาปีที่ 3 มีคะแนนเฉลี่ยร้อยละ 30.04 และชั้นมัธยมศึกษาปีที่ 6 มีคะแนนเฉลี่ยร้อยละ 30.72 (สถาบันทดสอบการศึกษาแห่งชาติ (องค์การมหาชน), 2562) ปัจจัยเชิงสาเหตุที่ทำให้คะแนน O-NET ของนักเรียนชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 6 ต่ำกว่าที่ควรจะเป็น มีสาเหตุมาจากปัจจัยด้านต่าง ๆ หลายด้าน เช่น ตัวผู้เรียน ครูผู้สอน ผู้บริหาร และตัวข้อสอบ ในส่วนของปัจจัยด้านครูผู้สอนถือว่ามีส่วนสำคัญอย่างยิ่ง เช่น ครูมีภาระงานมาก คุณวุฒิของครูไม่ตรงตามวิชาที่สอน การเข้าถึงข้อมูลของครูมีน้อย ซึ่งเห็นได้ว่าผลสัมฤทธิ์ทางการเรียนของนักเรียนไม่ได้เกี่ยวข้องกับนักเรียนฝ่ายเดียว แต่ครูถือได้ว่ามีบทบาทสำคัญในการช่วยเหลือเด็ก หน้าที่สำคัญประการหนึ่งของครูคณิตศาสตร์ คือ การสร้างเครื่องมือสำหรับวัดและประเมินผลตามตัวชี้วัด และมาตรฐานการเรียนรู้ เครื่องมือที่นิยมใช้กันมาก คือ แบบทดสอบประเภทเลือกตอบชนิดหลายตัวเลือก ซึ่งมักเป็นปัญหาสำหรับครูที่มีภาระงานมาก เพราะต้องใช้เวลาในการสร้างข้อสอบให้มีคุณภาพ และคณิตศาสตร์เป็นวิชาที่ต้องอาศัยการฝึกฝนบ่อย ๆ จากทฤษฎีฝึกสมอง (mental discipline) กล่าวว่า จิต สมอง หรือสติปัญญา สามารถพัฒนาให้ปราดเปรื่องได้โดยการฝึกสมองให้เป็นระเบียบ ฝึกโดยการคิด เรียนรู้ในสิ่งยาก ๆ ซ้ำ ๆ ซึ่งยิ่งยากมากเท่าไร สมองก็จะแข็งแกร่งมากขึ้นเท่านั้น (เอื้อมพร หลินเจริญ และคณะ, 2552)

แบบทดสอบที่มีคุณภาพจะต้องสร้างจากข้อสอบได้คุณภาพ การจัดทำคลังข้อสอบเป็นขั้นตอนหนึ่งที่มีความสำคัญ คลังข้อสอบควรบรรจุด้วยข้อสอบที่มีคุณภาพ (นุภาพรรณ ปลื้มใจ และคณะ, 2558) ข้อสอบที่นำมาบรรจุในคลังข้อสอบต้องผ่านการวิเคราะห์คุณภาพของข้อสอบตามหลักเกณฑ์ที่กำหนด สำหรับการวิเคราะห์คุณภาพของข้อสอบที่นำมาใช้กับคลังข้อสอบควรมีความละเอียดมากที่สุด ข้อสอบปรนัยแบบเลือกคำตอบ (Multiple choice) ควรใช้วิธีการวิเคราะห์คุณภาพที่ถูกต้อง โดยใช้โมเดลของทฤษฎีการตอบสนอง

ข้อสอบแบบ 3 พารามิเตอร์ เพราะสามารถวิเคราะห์คุณภาพของข้อสอบได้อย่างละเอียด ได้แก่ ค่าพารามิเตอร์ความยากค่าพารามิเตอร์อำนาจจำแนก และค่าพารามิเตอร์การเดา (ญานิศรา มุนินทร์สาคร, 2558)

การสร้างแบบทดสอบในการวัดและประเมินผลเป็นขั้นตอนที่มีความสำคัญมากต้องสร้างแบบทดสอบที่มีเนื้อหาสาระ ผลการเรียนรู้ที่คาดหวัง เหมาะสม และอยู่ในกรอบของมาตรฐานการเรียนรู้ของหลักสูตร ซึ่งจะส่งผลให้แบบทดสอบที่สร้างขึ้นมีคุณภาพ ข้อสอบแบบเลือกตอบเป็นเครื่องมือวัดผลซึ่งเป็นที่รู้จักและใช้กันอย่างแพร่หลายเหมาะสำหรับการประเมินด้านพุทธิพิสัยหรือความรู้ (cognitive domain) ของผู้เรียน มีความเป็นปรนัยสูง กล่าวคือ สามารถให้คะแนนได้ใกล้เคียงหรือเหมือนกัน แม้ว่าผู้ประเมินจะต่างคนกัน นอกจากนั้นยังสามารถบริหารจัดการสอบได้ง่าย เหมาะกับผู้เรียนจำนวนมาก ในการวัดผลด้านความรู้ Anderson et al. (2001) ได้ปรับการจำแนกจุดประสงค์ทางการศึกษาด้านพุทธิปัญญา (cognitive domain) ตามแนวคิดของ Bloom et al. (1956) ออกเป็นลำดับการเรียนรู้ที่เกิดขึ้นในสมองสามารถวัดได้หลายระดับตามการจำแนกระดับความรู้ โดยแบ่งออกเป็น 6 ขั้น ได้แก่ จำ (remember) เข้าใจ (understand) ประยุกต์ (apply) วิเคราะห์ (analyze) ประเมินค่า (evaluate) และสร้างสรรค์ (create) ซึ่งจำแนกตามกระบวนการพุทธิปัญญา (the cognitive process dimension) เป็น 3 ระดับ คือ ระดับที่ 1 (Level 1) ได้แก่ จำ (remember) ระดับที่ 2 (level 2) ได้แก่ เข้าใจ (understand) ระดับที่ 3 (level 3) ได้แก่ ประยุกต์ (apply) วิเคราะห์ (analyze) ประเมินค่า (evaluate) และสร้างสรรค์ (create)

จากที่กล่าวมาข้างต้น ผู้วิจัยจึงสนใจที่จะศึกษาการจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยากของข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบ จำแนกเนื้อหาตามโครงสร้างหลักสูตรวิชาคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ประกอบด้วย 3 สาระ คือ สาระที่ 1 จำนวนและพีชคณิต สาระที่ 2 การวัดและเรขาคณิต และสาระที่ 3 สถิติและความน่าจะเป็น (กระทรวงศึกษาธิการ, 2560) และจำแนกระดับความยากของข้อสอบตามหลักการทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ แบ่งเป็น 3 ระดับ คือ ง่าย ปานกลาง และยาก งานวิจัยนี้จะทำให้ได้ข้อสอบจำแนกตามเนื้อหาและระดับความยากเพื่อนำไปใช้ในการพัฒนาระบบสร้างข้อสอบอัตโนมัติ ซึ่งระบบดังกล่าวสามารถที่จะสร้างข้อสอบได้จำนวนมากในเวลาที่รวดเร็ว เป็นการช่วยลดภาระงานของครูในการสร้างข้อสอบวิชาคณิตศาสตร์

ความมุ่งหมายของการวิจัย

1. เพื่อวิเคราะห์คุณภาพข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์
2. เพื่อจัดคลังข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 จำแนกตามเนื้อหาและระดับความยาก

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ผู้วิจัยรวบรวมข้อสอบและผลการตอบข้อสอบวิชาคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ที่เข้าสอบแข่งขันวัดความสามารถทางคณิตศาสตร์ (top mathematics test) เป็นข้อมูลแบบทุติยภูมิ

(secondary data) ระหว่างปี 2557 – 2561 ที่ได้มาจาก บริษัท ท็อป เทสท์ เซ็นเตอร์ จำกัด ซึ่งมีข้อสอบทั้งหมด 480 ข้อ จากจำนวนนักเรียนที่เข้าสอบ 27,886 คน

2. เครื่องมือที่ใช้ในการวิจัย

แบบประเมินความตรงเชิงโครงสร้าง และพฤติกรรมที่วัดของข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

3. การเก็บรวบรวมข้อมูล

3.1 รวบรวมข้อสอบวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 ระหว่างปี 2557 - 2561 จำนวน 480 ข้อ เพื่อมาทำการวิเคราะห์คุณภาพตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ ด้วยโปรแกรม Xcalibre™ 4.2.2

3.2 วิเคราะห์คุณภาพข้อสอบตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ มีขั้นตอน ดังนี้

3.2.1 จัดเตรียมข้อมูลผลการตอบข้อสอบ พร้อมกำหนดเงื่อนไขของการวิเคราะห์คุณภาพของข้อสอบตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์

3.2.2 วิเคราะห์คุณภาพของข้อสอบตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ โดยใช้โปรแกรมสำเร็จรูป Xcalibre™ 4.2.2

3.2.3 คัดเลือกข้อสอบจากผลการวิเคราะห์คุณภาพของข้อสอบ โดยคัดเลือกข้อสอบที่ตรงตามเงื่อนไขที่กำหนดไว้ ดังนี้

เกณฑ์การคัดเลือกข้อสอบ (Urry, 1977 อ้างถึงใน จารุจิตร สิทธิปรุ และคณะ, 2559) ค่าพารามิเตอร์ความยาก (b-Parameter) มีค่าตั้งแต่ -2.50 ถึง 2.50 ค่าพารามิเตอร์อำนาจจำแนก (a-Parameter) มีค่าตั้งแต่ 0.50 ถึง2.50 และค่าพารามิเตอร์การเดา (c-Parameter) มีค่าไม่เกิน 0.30

3.3 จำแนกระดับความยากของข้อสอบตามพารามิเตอร์ค่าความยากของข้อสอบ

นำข้อคำถาม ตัวเลือก เฉลย และค่าพารามิเตอร์ของข้อสอบ บันทึกลงในฐานข้อสอบที่พัฒนาขึ้นโดยใช้ฐานข้อมูล MySQL โดยการจำแนกค่าความยาก (b-Stratified Model) สำหรับข้อสอบที่คัดเลือกไว้จะมีค่าความยากของข้อสอบตั้งแต่ -2.50 ถึง +2.50

การจำแนกข้อมูลค่าพารามิเตอร์ความยากของข้อสอบ จะใช้ค่าพารามิเตอร์ความยากของข้อสอบ จากผลการวิเคราะห์ข้อสอบทั้งหมดผ่านเกณฑ์การวิเคราะห์มาทำการหาค่าเฉลี่ย และจำแนกออกเป็น 3 ช่วง ตามระดับค่าความยากของข้อสอบ ดังนี้ ช่วงแรก บรรจุข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ -2.50 ถึง 0.50 เป็นข้อสอบที่ง่าย ช่วงที่สอง บรรจุข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ 0.51 ถึง 1.50 เป็นข้อสอบปานกลาง ช่วงสุดท้าย บรรจุข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ 1.51 ถึง 2.50 เป็นข้อสอบที่ยาก

3.4 ประเมินความตรงเชิงโครงสร้างและพฤติกรรมที่วัดของข้อสอบ

นำข้อสอบในแต่ละข้อที่จำแนกระดับความยากออกเป็น 3 ระดับ คือ ง่าย ปานกลาง และยาก มาประเมินว่าข้อสอบแต่ละระดับความยากประกอบด้วยองค์ประกอบ 3 สาระการเรียนรู้ คือ สาระที่ 1 จำนวน

และพีชคณิต สาระที่ 2 การวัดและเรขาคณิต และสาระที่ 3 สถิติและความน่าจะเป็น สำหรับการตรวจสอบความตรงเชิงโครงสร้าง จิตติรัตน์ แสงเลิศอุทัย (2560) ได้แนะนำไว้ว่า วิธีที่จะใช้ตรวจสอบความตรงเชิง โครงสร้างได้วิธีหนึ่ง คือ การประเมินความสอดคล้องของข้อสอบกับองค์ประกอบของทฤษฎีที่ต้องการวัดโดยพิจารณาข้อสอบแต่ละข้อว่ามีความสอดคล้องกับทฤษฎีหรือพิจารณาว่าข้อสอบหรือประเด็นคำถามมีความเหมาะสมที่จะเป็นตัวแทนลักษณะพฤติกรรมตามโครงสร้างทฤษฎีนั้นหรือไม่ เกณฑ์การพิจารณาให้ใช้เกณฑ์การประเมินมาตรประมาณค่า 5 ระดับ โดยข้อความมีความสอดคล้องมากที่สุดมีค่าคะแนนเท่ากับ 5 และ ข้อความมีความสอดคล้องน้อยที่สุดมีค่าคะแนนเท่ากับ 1 จากนั้นให้นำคะแนนที่ผู้เชี่ยวชาญประเมินมาคำนวณหาความสอดคล้อง และพิจารณาคะแนนค่าความสอดคล้องเฉลี่ยมากกว่า 3.50 ขึ้นไป และมีค่าส่วนเบี่ยงเบนมาตรฐานน้อยกว่า 1.00 แสดงว่าข้อสอบหรือประเด็นคำถามนั้นใช้ได้ แต่หากไม่เป็นไปตามประเด็นดังกล่าว ให้ตัดทิ้งไปหรือปรับปรุงแก้ไขก่อนนำไปใช้ ส่วนด้านพฤติกรรมที่วัดจะให้ผู้เชี่ยวชาญประเมินระดับความรู้ตามทฤษฎีการเรียนรู้ของบลูม ฉบับปรับปรุง (A Revision of Bloom’s Taxonomy) (Anderson et al., 2001) ที่แบ่งเป็น 6 ขั้น ได้แก่ 1) จำ (remember) 2) เข้าใจ (understand) 3) ประยุกต์ (apply) 4) วิเคราะห์ (analyze) 5) ประเมินค่า (evaluate) และ 6) สร้างสรรค์ (create) เกณฑ์การพิจารณาให้ใช้เกณฑ์การประเมินมาตรประมาณค่า 5 ระดับ และพิจารณา คะแนนค่าความสอดคล้องเฉลี่ยมากกว่า 3.50 ขึ้นไป

การประเมินความตรงเชิงโครงสร้างของข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ได้ดังนี้

3.4.1 นำข้อสอบตรวจสอบความครอบคลุมของโครงสร้าง และพฤติกรรมที่วัด

3.4.2 นำข้อสอบให้ผู้เชี่ยวชาญ จำนวน 5 ท่าน ประกอบด้วยผู้ทรงคุณวุฒิที่มีประสบการณ์การสอนตั้งแต่ 10 ปี ขึ้นไป เพื่อประเมินความตรงเชิงโครงสร้าง ว่ากรอบแนวคิดหรือทฤษฎีที่นำไปใช้กำหนดเป็นโครงสร้างในการวัดมีความสอดคล้องหรือไม่ กำหนดระดับการแสดงความคิดเห็นในลักษณะมาตรประมาณค่า 5 ระดับ (Wiersma and Jurs, 1990) ซึ่งแต่ละระดับมีความหมาย ดังนี้

5 หมายถึง ข้อสอบมีความสอดคล้องมากที่สุด

4 หมายถึง ข้อสอบมีความสอดคล้องมาก

3 หมายถึง ข้อสอบมีความสอดคล้องปานกลาง

2 หมายถึง ข้อสอบมีความสอดคล้องน้อย

1 หมายถึง ข้อสอบมีความสอดคล้องน้อยที่สุด

เกณฑ์แปลความหมาย ดังนี้

คะแนนเฉลี่ย 4.51 – 5.00 แปลความว่า มีความสอดคล้องในระดับมากที่สุด

คะแนนเฉลี่ย 3.51 – 4.50 แปลความว่า มีความสอดคล้องในระดับมาก

คะแนนเฉลี่ย 2.51 – 3.50 แปลความว่า มีความสอดคล้องในระดับปานกลาง

คะแนนเฉลี่ย 1.51 – 2.50 แปลความว่า มีความสอดคล้องในระดับน้อย

คะแนนเฉลี่ย 1.00 – 1.50 แปลความว่า มีความสอดคล้องในระดับน้อยที่สุด

3.5 จัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยากของข้อสอบ

4. การวิเคราะห์ข้อมูล

4.1 วิเคราะห์คุณภาพของข้อสอบตามการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ คือ ค่าพารามิเตอร์ความยาก ค่าพารามิเตอร์อำนาจจำแนก และค่าพารามิเตอร์การเดา ด้วยโปรแกรม Xcalibre™ 4.2.2 และวิเคราะห์สถิติพื้นฐาน คือ ค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (SD) ด้วยโปรแกรมสำเร็จรูป

4.2 วิเคราะห์ค่าความสอดคล้องของความตรงเชิงโครงสร้างและพฤติกรรมที่วัดของข้อสอบจากคะแนนความคิดเห็นของผู้เชี่ยวชาญ โดยใช้สถิติพื้นฐานในการวิเคราะห์ข้อมูล คือ ค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (SD) ด้วยโปรแกรมสำเร็จรูป

ผลการวิจัย

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อเพื่อวิเคราะห์คุณภาพข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ แสดงผลดังตาราง 1 และจัดคลังข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 จำแนกตามเนื้อหาและระดับความยาก แสดงผลดังตาราง 3, 4, 5 และ 6

ตาราง 1 ผลการวิเคราะห์คุณภาพตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ ข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ตั้งแต่ ระหว่างปี 2557-2561

พารามิเตอร์	ช่วงค่าที่ผ่านเกณฑ์การคัดเลือก	Mean	SD
ค่าความยาก	-1.19 ถึง 2.49	1.20	0.81
ค่าอำนาจจำแนก	0.51 ถึง 2.42	1.15	0.40
ค่าการเดา	0.08 ถึง 0.30	0.23	0.03

จากตาราง 1 ข้อสอบที่ผ่านเกณฑ์การคัดเลือกมีค่าพารามิเตอร์ความยาก ตั้งแต่ -1.19 ถึง 2.49 ค่าเฉลี่ยพารามิเตอร์ความยากของข้อสอบ เท่ากับ 1.20 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.81 มีค่าพารามิเตอร์อำนาจจำแนกของข้อสอบตั้งแต่ 0.51 ถึง 2.42 ค่าเฉลี่ยพารามิเตอร์อำนาจจำแนกของข้อสอบเท่ากับ 1.15 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.40 มีค่าพารามิเตอร์การเดาของข้อสอบตั้งแต่ 0.08 ถึง 0.30 ค่าเฉลี่ยพารามิเตอร์การเดาของข้อสอบ เท่ากับ 0.23 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.03

ตาราง 2 การจำแนกระดับความยากข้อสอบตามพารามิเตอร์ค่าความยากของข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

พารามิเตอร์ค่าความยาก	ระดับความยากของข้อสอบ	จำนวน (ข้อ)
-1.19 ถึง 0.50	ง่าย	65
0.51 ถึง 1.50	ปานกลาง	155
1.51 ถึง 2.49	ยาก	148
	รวม	**368**

จากตาราง 2 ข้อสอบจำแนกออกเป็น 3 ระดับ คือ ง่าย ปานกลาง และยาก คือ ระดับแรก ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ -1.19 ถึง 0.50 เป็นข้อสอบที่ง่าย จำนวน 65 ข้อ ระดับที่สอง ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ 0.51 ถึง 1.50 เป็นข้อสอบปานกลาง จำนวน 155 ข้อ และระดับสุดท้าย ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ 1.51 ถึง 2.49 เป็นข้อสอบที่ยาก จำนวน 148 ข้อ รวมข้อสอบทั้ง 3 ระดับ มีจำนวนทั้งสิ้น 368 ข้อ

ตาราง 3 ผลการประเมินความตรงเชิงโครงสร้างและพฤติกรรมที่วัดของข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

ระดับความยากของข้อสอบ	จำนวน (ข้อ)	ความตรงเชิงโครงสร้าง		พฤติกรรมที่วัด		ระดับความสอดคล้อง
		Mean	SD	Mean	SD	มากที่สุด
ง่าย	58	4.71	0.13	4.73	0.12	มากที่สุด
ปานกลาง	147	4.69	0.12	4.76	0.09	มากที่สุด
ยาก	140	4.68	0.11	4.77	0.09	มากที่สุด
รวม	**345**	**4.69**	**0.12**	**4.76**	**0.10**	**มากที่สุด**

จากตาราง 3 ข้อสอบที่ผ่านเกณฑ์การประเมินความตรงเชิงโครงสร้างและด้านพฤติกรรมที่วัด จำนวน 345 ข้อ โดยความตรงเชิงโครงสร้างมีความสอดคล้องเฉลี่ย เท่ากับ 4.69 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.12 และพฤติกรรมที่วัดมีความสอดคล้องเฉลี่ย เท่ากับ 4.76 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.10 โดยรวมมีความสอดคล้องอยู่ในระดับมากที่สุด

ตาราง 4 เปรียบเทียบโครงสร้างสาระการเรียนรู้ในแต่ละระดับความยากของข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

โครงสร้างสาระการเรียนรู้คณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551	ระดับความยากของข้อสอบ		
	ง่าย (ข้อ)	ปานกลาง (ข้อ)	ยาก (ข้อ)
สาระที่ 1 จำนวนและพีชคณิต	32	105	101
สาระที่ 2 การวัดและเรขาคณิต	20	32	28
สาระที่ 3 สถิติและความน่าจะเป็น	6	10	11
รวม	**58**	**147**	**140**

จากตาราง 4 ข้อสอบระดับง่าย มีจำนวนทั้งหมด 58 ข้อ ประกอบด้วย สาระที่ 1 จำนวนและพีชคณิต จำนวน 32 ข้อ สาระที่ 2 การวัดและเรขาคณิต จำนวน 20 ข้อ และสาระที่ 3 สถิติและความน่าจะเป็น จำนวน 6 ข้อ ข้อสอบระดับปานกลาง มีจำนวนทั้งหมด 147 ข้อ ประกอบด้วย สาระที่ 1 จำนวนและพีชคณิต จำนวน 105 ข้อ สาระที่ 2 การวัดและเรขาคณิต จำนวน 32 ข้อ และสาระที่ 3 สถิติและความน่าจะเป็น จำนวน 10 ข้อ และข้อสอบระดับยาก มีจำนวนทั้งหมด 140 ข้อ ประกอบด้วย สาระที่ 1 จำนวนและพีชคณิต จำนวน 101 ข้อ สาระที่ 2 การวัดและเรขาคณิต จำนวน 28 ข้อ และสาระที่ 3 สถิติและความน่าจะเป็น จำนวน 11 ข้อ

ตาราง 5 เปรียบเทียบพฤติกรรมที่วัดในแต่ละระดับความยากของข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

พฤติกรรมที่วัดโดยจำแนกลำดับการเรียนรู้ด้านพุทธิปัญญา (cognitive domain) ตามแนวคิดของบลูม และคณะ ฉบับปรับปรุง (Anderson et al., 2001)	**ระดับความยากของข้อสอบ**		
	ง่าย (ข้อ)	ปานกลาง (ข้อ)	ยาก (ข้อ)
ขั้นที่ 1 จำ (remember)	0	0	0
ขั้นที่ 2 เข้าใจ (understand)	40	76	41
ขั้นที่ 3 ประยุกต์ (apply)	3	45	78
ขั้นที่ 4 วิเคราะห์ (analyze)	15	20	18
ขั้นที่ 5 ประเมินค่า (evaluate)	0	5	1
ขั้นที่ 6 สร้างสรรค์ (create)	0	1	2
รวม	**58**	**147**	**140**

จากตาราง 5 ข้อสอบระดับง่าย มีจำนวนทั้งหมด 58 ข้อ ประกอบด้วยการเรียนรู้ด้านพุทธิปัญญา ขั้นที่ 2 เข้าใจ (understand) จำนวน 40 ข้อ ขั้นที่ 3 ประยุกต์ (apply) จำนวน 3 ข้อ และขั้นที่ 4 วิเคราะห์ (analyze) จำนวน 15 ข้อ ข้อสอบระดับปานกลาง มีจำนวนทั้งหมด 147 ข้อ ประกอบด้วยการเรียนรู้ด้าน พุทธิปัญญา ขั้นที่ 2 เข้าใจ (Understand) จำนวน 76 ข้อ ขั้นที่ 3 ประยุกต์ (apply) จำนวน 45 ข้อ ขั้นที่ 4 วิเคราะห์ (analyze) จำนวน 20 ข้อ ขั้นที่ 5 ประเมินค่า (evaluate) จำนวน 5 ข้อ และขั้นที่ 6 สร้างสรรค์ (create) จำนวน 1 ข้อ และข้อสอบระดับยาก มีจำนวนทั้งหมด 140 ข้อ ประกอบด้วยการเรียนรู้ด้านพุทธิปัญญา ขั้นที่ 2 เข้าใจ (understand) จำนวน 41 ข้อ ขั้นที่ 3 ประยุกต์ (apply) จำนวน 78 ข้อ ขั้นที่ 4 วิเคราะห์ (analyze) จำนวน 18 ข้อ ขั้นที่ 5 ประเมินค่า (evaluate) จำนวน 1 ข้อ และขั้นที่ 6 สร้างสรรค์ (create) จำนวน 2 ข้อ

อภิปรายผล

1. การวิเคราะห์คุณภาพข้อสอบวิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 1

การวิเคราะห์คุณภาพข้อสอบวิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 1 ตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ ด้วยโปรแกรม Xcalibre Version 4.2.2 สำหรับประมาณค่าพารามิเตอร์ของข้อสอบ ซึ่งมีเกณฑ์การคัดเลือกข้อสอบ คือ ค่าความยากของข้อสอบมีค่าตั้งแต่ - 2.50 ถึง 2.50 ค่าอำนาจจำแนกของข้อสอบ มีค่าตั้งแต่ 0.50 ถึง 2.50 และค่าการเดาของข้อสอบมีค่าไม่เกิน 0.30 (จารุจิตร สิทธิปรุ และคณะ, 2559 อ้างถึงใน Urry, 1977) มีข้อสอบที่ผ่านเกณฑ์การคัดเลือก จำนวน 368 ข้อ จากข้อสอบทั้งหมดที่นำมาวิเคราะห์ 480 ข้อ คิดเป็นร้อยละ 76.67 โดยข้อสอบมีค่าพารามิเตอร์ความยาก ตั้งแต่ -1.19 ถึง 2.49 ค่าเฉลี่ยพารามิเตอร์ความยากของข้อสอบ เท่ากับ 1.20 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.81 มีค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ ตั้งแต่ 0.51 ถึง 2.42 ค่าเฉลี่ยพารามิเตอร์อำนาจจำแนกของข้อสอบเท่ากับ 1.15 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.40 มีค่า พารามิเตอร์การเดาของข้อสอบ ตั้งแต่ 0.08 ถึง 0.30 ค่าเฉลี่ยพารามิเตอร์การเดาของข้อสอบ เท่ากับ 0.23 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.03 สอดคล้องกับงาน วิจัยของ นุภาพรรณ ปลื้มใจ และคณะ (2558) กล่าวว่า การจัดทำคลังข้อสอบเป็นขั้นตอนหนึ่งที่มีความสำคัญ เนื่องจากคลังข้อสอบจะส่งผลต่อการจัดการแบบทดสอบควรจัดทำคลังข้อสอบที่มีประสิทธิภาพ ซึ่งข้อสอบที่นำมาบรรจุในคลังข้อสอบต้องผ่านการวิเคราะห์คุณภาพของข้อสอบตามหลักเกณฑ์และเงื่อนไขของทฤษฎีการตอบสนองข้อสอบ การวิเคราะห์คุณภาพของข้อสอบหลายวิธี เช่น การวิเคราะห์คุณภาพข้อสอบด้วยโมเดลการตอบสนองข้อสอบ แบบ 1 พารามิเตอร์ แบบ 2 พารามิเตอร์ และแบบ 3 พารามิเตอร์ สำหรับการวิเคราะห์คุณภาพของข้อสอบที่นำมาใช้กับคลังข้อสอบในการทดสอบ ควรมีความละเอียดมากที่สุด ดังนั้นการวิเคราะห์คุณภาพของข้อสอบจึงใช้โมเดลของทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ เพราะสามารถวิเคราะห์คุณภาพของข้อสอบได้ทั้งค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ สอดคล้องกับงานวิจัยของ ญานิศรา มุนินทร์สาคร (2558); จารุจิตร สิทธิปรุ และคณะ (2559); สุชาดา กรเพชร ปาณี และคณะ (2559); ปิยะทิพย์ ประดุจพรม และกนก พานทอง (2562) พบว่าใช้การวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ประกอบด้วย ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ สำหรับประมาณค่าพารามิเตอร์ของข้อสอบซึ่งมีเกณฑ์การคัดเลือกข้อสอบ คือ ค่าอำนาจจำแนกของข้อสอบ ตั้งแต่ 0.50 ถึง 2.50 ค่าความยากของข้อสอบ ตั้งแต่ -2.50 ถึง 2.50 และค่าการเดาของข้อสอบ ไม่เกิน 0.30 (Urry, 1977)

2. การจำแนกระดับความยากข้อสอบตามพารามิเตอร์ค่าความยากของข้อสอบวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1

การจำแนกระดับความยากข้อสอบตามพารามิเตอร์ค่าความยากของข้อสอบวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 โดยการนำข้อสอบจำนวน 368 ข้อ ที่ผ่านเกณฑ์การวิเคราะห์ข้อสอบมาทำการแจกแจงแบบปกติ และจำแนกข้อมูลค่าพารามิเตอร์ความยาก คือ ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบที่อยู่ใกล้ -2.50 เป็นข้อสอบที่ง่าย ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบที่อยู่ใกล้ค่าเฉลี่ยของข้อสอบทั้งหมดที่ผ่านเกณฑ์การวิเคราะห์ เป็นข้อสอบปานกลาง ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบที่อยู่

ใกล้ +2.50 เป็นข้อสอบที่ยาก ได้ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ -1.19 ถึง 0.50 เป็นข้อสอบที่ง่าย ข้อสอบที่มีค่าพารามิเตอร์ความยากของข้อสอบตั้งแต่ 0.51 ถึง 1.50 เป็นข้อสอบปานกลาง ข้อสอบที่มีค่า พารามิเตอร์ความยากของข้อสอบตั้งแต่ 1.51 ถึง 2.49 เป็นข้อสอบที่ยาก สอดคล้องกับงานวิจัยของ จารุจิตร สิทธิปรุ และคณะ (2559) พบว่าการวิเคราะห์คุณภาพข้อสอบ O-NET ชั้นมัธยมศึกษาปีที่ 3 ใช้ค่าเฉลี่ยความยากข้อสอบ จากผลการวิเคราะห์ข้อสอบทั้งหมด แบ่งออกเป็น 7 ช่วง คือ ช่วงที่ 1) ค่าความยากเฉลี่ยมากกว่า 2.0000 หมายถึง ข้อสอบที่ยากมาก ช่วงที่ 2) ค่าความยากเฉลี่ยตั้งแต่ 1.0001 ถึง 2.0000 หมายถึง ข้อสอบที่ยาก ช่วงที่ 3) ค่าความยากเฉลี่ยตั้งแต่ 0.5001 ถึง 1.0000 หมายถึง ข้อสอบค่อนข้างยาก ช่วงที่ 4) ค่าความยากเฉลี่ยตั้งแต่ -0.4999 ถึง 0.5000 หมายถึง ข้อสอบปานกลาง ช่วงที่ 5) ค่าความยากเฉลี่ยตั้งแต่ -0.9999 ถึง -0.5000 หมายถึง ข้อสอบค่อนข้างง่าย ช่วงที่ 6) ค่าความยากเฉลี่ยตั้งแต่ -2.0000 ถึง -1.0000 หมายถึง ข้อสอบค่อนข้างง่าย และช่วงที่ 7) ค่าความยากเฉลี่ยน้อยกว่า -2.0000 หมายถึง ข้อสอบที่ง่ายมาก สอดคล้องกับงานวิจัยของ สุนันทา ศิริเบญจา และคณะ (2556) พบว่าทำการคัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์เข้าสู่ระบบคลังข้อสอบออนไลน์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ตามหลักการทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โดยใช้วิธีหาค่าเฉลี่ยความยากจากผลการวิเคราะห์ข้อสอบทั้งหมดที่นำมาวิเคราะห์ สอดคล้องกับงานวิจัยของ Thomas (2016) พบว่าได้นำค่าพารามิเตอร์ความยากของข้อสอบ (b-Parameter) ในอดีตมาทำการศึกษาย้อนรอยเพื่อเปรียบเทียบความสัมพันธ์กับค่าพารามิเตอร์ของข้อสอบตามหลักทฤษฎีการตอบ สนองข้อสอบ ผลการวิจัยแสดงให้เห็นถึงความสัมพันธ์เชิงบวกอย่างมีนัยสำคัญทางสถิติ ($p = .001$) ระหว่างค่า พารามิเตอร์ความยากของข้อสอบ (b-Parameter) เชิงประจักษ์กับแบบจำลองความซับซ้อนมากกว่า .800

3. การประเมินความตรงเชิงโครงสร้างและพฤติกรรมที่วัดของข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

การประเมินความตรงเชิงโครงสร้างและพฤติกรรมที่วัดของข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 โดยผู้เชี่ยวชาญจำนวน 5 ท่าน ข้อสอบที่ผ่านเกณฑ์การประเมินความตรงเชิงโครงสร้าง และด้านพฤติกรรมที่วัด มีจำนวน 345 ข้อ จากข้อสอบที่ผ่านคุณภาพการวิเคราะห์ตามหลักการทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ จำนวน 368 ข้อ คิดเป็นร้อยละ 93.75 มีความตรงเชิงโครงสร้างมีความสอดคล้องเฉลี่ยเท่ากับ 4.69 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.12 และพฤติกรรมที่วัดมีความสอดคล้องเฉลี่ยเท่ากับ 4.76 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.10 สรุปผลโดยรวมมีความสอดคล้องอยู่ในระดับมากที่สุด สอดคล้องกับงานวิจัยของ จิตติรัตน์ แสงเลิศอุทัย (2560) กล่าวว่า ความตรงเชิงโครงสร้างเป็นคุณสมบัติของข้อสอบที่วัดโครงสร้างหรือคุณลักษณะตามโครงสร้างหรือทฤษฎีของสิ่งนั้นหรือตรงตามลักษณะที่จะวัดได้ครบถ้วน การตรวจสอบความสอดคล้องของข้อสอบกับองค์ประกอบของทฤษฎีที่ต้องการวัด โดยพิจารณาข้อสอบแต่ละข้อว่ามีความสอดคล้องกับทฤษฎีหรือพิจารณาว่าข้อสอบหรือประเด็นคำถามมีความเหมาะสมที่จะเป็นตัวแทนลักษณะพฤติกรรมตามโครงสร้างทฤษฎีนั้นหรือไม่ เกณฑ์การพิจารณาให้ใช้เกณฑ์การประเมินมาตรประมาณ 5 ระดับ โดยพิจารณาคะแนนค่าความสอดคล้องเฉลี่ยมากกว่า 3.50 ขึ้นไป และมีค่าส่วนเบี่ยงเบนมาตรฐานน้อยกว่า 1.00 แสดงว่าข้อสอบหรือประเด็นคำถามนั้นใช้ได้ สอดคล้องกับงานวิจัยของ ขนิษฐา ราศรี และ

ประภาพร ศรีตระกูล (2553) พบว่าการสร้างและหาคุณภาพของแบบทดสอบโดยการทำการตรวจสอบคุณภาพทั้งฉบับด้านความตรงเชิงโครงสร้างของแบบทดสอบวัดการคิดวิเคราะห์ทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 โดยเสนอต่อผู้เชี่ยวชาญเพื่อพิจารณาความเหมาะสมของการคัดเลือกข้อสอบ การที่ข้อสอบมีคุณภาพผ่านเกณฑ์เนื่องจากผู้วิจัยได้สร้างแบบทดสอบแบบอิงมาตรฐานการเรียนรู้กลุ่มสาระการเรียนรู้คณิตศาสตร์ โดยการวิเคราะห์พฤติกรรมจากมาตรฐานการเรียนรู้ และนำผลการวิเคราะห์มาตรฐานการเรียนรู้ดังกล่าวเสนอต่อผู้เชี่ยวชาญเพื่อพิจารณาปรับปรุงคุณภาพทำให้ได้แบบทดสอบที่สามารถวัดได้ตรงกับระดับการคิดวิเคราะห์ทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มุ่งวัด

4. การจัดทำคลังข้อสอบจำแนกตามเนื้อหาและระดับความยากของข้อสอบวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1

ข้อสอบที่ผ่านเกณฑ์การคัดเลือกเข้าสู่คลังข้อสอบจำแนกตามเนื้อหาและระดับความยากของข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 มีจำนวนข้อสอบ 345 ข้อ จากข้อสอบจำนวนทั้งหมด 480 ข้อ ที่นำมาใช้ในการวิจัยในครั้งนี้ คิดเป็นร้อยละ 71.88 การจำแนกเนื้อหาตามโครงสร้างหลักสูตรวิชาคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 และจำแนกระดับความยากของข้อสอบตามหลักการทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ผลการวิจัยพบว่า สาระที่ 1 จำนวนและพีชคณิต มีจำนวนข้อสอบทั้งหมด 238 ข้อ เรียงลำดับจากมากไปน้อย คือ ระดับปานกลาง จำนวน 105 ข้อ คิดเป็นร้อยละ 44.12 ระดับยากจำนวน 101 ข้อ คิดเป็นร้อยละ 42.44 และระดับง่าย จำนวน 32 ข้อ คิดเป็นร้อยละ 13.44 ตามลำดับ สาระที่ 2 การวัดและเรขาคณิต มีจำนวนข้อสอบทั้งหมด 80 ข้อ เรียงลำดับจากมากไปน้อย คือ ระดับปานกลาง จำนวน 32 ข้อ คิดเป็นร้อยละ 40.00 ระดับยาก จำนวน 28 ข้อ คิดเป็นร้อยละ 35.00 และระดับง่าย จำนวน 20 ข้อ คิดเป็นร้อยละ 25.00 ตามลำดับ และสาระที่ 3 สถิติและความน่าจะเป็น มีจำนวนข้อสอบทั้งหมด 27 ข้อ เรียงลำดับจากมากไปน้อย คือ ระดับยาก จำนวน 11 ข้อ คิดเป็นร้อยละ 40.74 ระดับปานกลาง จำนวน 10 ข้อ คิดเป็นร้อยละ 37.04 และระดับง่าย จำนวน 6 ข้อ คิดเป็นร้อยละ 22.22 ตามลำดับ

ผลการวิจัยแสดงให้เห็นถึงร้อยละของข้อสอบความยากระดับง่าย ปานกลาง และยาก ที่จำแนกเนื้อหาตามโครงสร้างหลักสูตรวิชาคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 คือ ข้อสอบที่มีค่าความยากระดับปานกลาง (พารามิเตอร์ค่าความยากตั้งแต่ 0.51 ถึง 1.50) และระดับยาก (พารามิเตอร์ค่าความยากตั้งแต่ 1.51 ถึง 2.49) มีร้อยละของข้อสอบในสัดส่วนที่มากใกล้เคียงกัน คือ สาระที่ 1 (ระดับปานกลาง ร้อยละ 44.12 และระดับยาก ร้อยละ 42.44) สาระที่ 2 (ระดับปานกลาง ร้อยละ 40.00 และระดับยาก ร้อยละ 35.00) และสาระที่ 3 (ระดับปานกลาง ร้อยละ 37.04 และระดับยาก ร้อยละ 40.74) ข้อสอบที่มีค่าความยากระดับง่าย (พารามิเตอร์ค่าความยากตั้งแต่ -1.19 ถึง 0.50) มีร้อยละของข้อสอบในสัดส่วนที่น้อย คือ สาระที่ 1 (ระดับง่าย ร้อยละ 13.44) สาระที่ 2 (ระดับง่าย ร้อยละ 25) และสาระที่ 3 (ระดับง่าย ร้อยละ 22.22) สาเหตุเพราะข้อสอบที่นำมาใช้ในการวิจัยในครั้งนี้ เป็นข้อสอบที่ใช้ในการสอบแข่งขันวัดความสามารถทางคณิตศาสตร์ ดังนั้นข้อสอบจึงประกอบด้วยข้อสอบที่มีค่าความยากตั้งแต่ระดับปานกลางถึงระดับยาก (พารามิเตอร์ค่าความยากตั้งแต่ 0.51 ถึง 2.49) สอดคล้องกับงานวิจัยของ

สุชาดา กรเพชรปาณี และคณะ (2559) พบว่าจัดทำคลังข้อสอบสำหรับพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และชั้นมัธยมศึกษาปีที่ 6 จำนวน 8 กลุ่มสาระการเรียนรู้ โดยการนำข้อสอบ O–NET จากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ซึ่งเป็นข้อมูลแบบทุติยภูมิ ตั้งแต่ ปี 2551-2553 มาทำการวิเคราะห์คุณภาพตามทฤษฎีการตอบสนองข้อสอบ (IRT) แบบ 3 พารามิเตอร์ ผลการวิจัยพบว่า ข้อสอบวิชาคณิตศาสตร์ที่อยู่ในคลังข้อสอบที่ผ่านเกณฑ์การคัดเลือกของชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และชั้นมัธยมศึกษาปีที่ 6 มีค่าความยากของข้อสอบเฉลี่ยเท่ากับ 0.5932, 1.3047, 1.0478 ตามลำดับ โดยใช้ค่าเฉลี่ยความยากของข้อสอบจากผลการวิเคราะห์ข้อสอบทั้งหมด แบ่งออกเป็น 7 ช่วง คือ ช่วงที่ 1) ค่าความยากเฉลี่ยมากกว่า 2.0000 หมายถึง ข้อสอบที่ยากมาก ช่วงที่ 2) ค่าความยากเฉลี่ยตั้งแต่ 1.0001 ถึง 2.0000 หมายถึง ข้อสอบที่ยาก ช่วงที่ 3) ค่าความยากเฉลี่ยตั้งแต่ 0.5001 ถึง 1.0000 หมายถึง ข้อสอบค่อนข้างยาก ช่วงที่ 4) ค่าความยากเฉลี่ยตั้งแต่ -0.4999 ถึง 0.5000 หมายถึง ข้อสอบปานกลาง ช่วงที่ 5) ค่าความยากเฉลี่ยตั้งแต่ -0.9999 ถึง -0.5000 หมายถึง ข้อสอบค่อนข้างง่าย ช่วงที่ 6) ค่าความยากเฉลี่ยตั้งแต่ -2.0000 ถึง -1.0000 หมายถึง ข้อสอบค่อนข้างง่าย และช่วงที่ 7) ค่าความยากเฉลี่ยน้อยกว่า -2.0000 หมายถึง ข้อสอบที่ง่ายมาก แสดงให้เห็นว่าข้อสอบ O-NET วิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 มีค่าความยากของข้อสอบเฉลี่ยค่อนข้างยาก ชั้นมัธยมศึกษาปีที่ 3 มีค่าความยากของข้อสอบเฉลี่ยยาก และชั้นมัธยมศึกษาปีที่ 6 มีค่าความยากของข้อสอบเฉลี่ยยาก ตามลำดับ

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 ใช้เป็นข้อมูลพื้นฐานในการวิเคราะห์เพื่อจำแนกเนื้อหาและระดับความยากของข้อสอบ ที่แบ่งออกเป็น 3 ระดับ คือ ยาก ปานกลาง และง่าย ตามสาระการเรียนรู้ในรายวิชาต่าง ๆ ที่สนใจ

1.2 ใช้ในการพัฒนาระบบสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก ที่แบ่งออกเป็น 3 ระดับ คือ ยาก ปานกลาง และง่าย

1.3 ครูผู้สอนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 สามารถนำข้อสอบในคลังข้อสอบที่จำแนกตามเนื้อหาและระดับความยาก ให้นักเรียนฝึกฝนเพื่อพัฒนาทักษะทางด้านคณิตศาสตร์ และจากคลังข้อสอบที่ผู้วิจัยจัดขึ้นมีการจำแนกระดับความยาก คือ ง่าย ปานกลาง และยาก ดังนั้นครูต้องจำแนกทักษะทางด้านคณิตศาสตร์ของนักเรียนออกมาก่อน เพื่อการนำข้อสอบไปใช้ให้เหมาะสมกับความสามารถของนักเรียนเพื่อให้เกิดผลสัมฤทธิ์สูงสุดในการนำไปใช้การวัดและประเมินผล

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรมีการศึกษาวิจัยในการสร้างคลังข้อสอบจำแนกตามเนื้อหาและระดับความยากของวิชาคณิตศาสตร์ในระดับชั้นต่าง ๆ และในวิชาอื่น ๆ เช่น ภาษาอังกฤษ วิทยาศาสตร์ เป็นต้น

2.2 ข้อสอบที่นำมาใช้ในการวิจัยครั้งนี้เป็นแบบเลือกตอบ ชนิดไม่เกิน 4 ตัวเลือก มีลักษณะการตรวจให้คะแนนแบบ 2 ค่า (ตอบถูกได้ 1 คะแนน และตอบผิดได้ 0 คะแนน) เท่านั้น ควรศึกษาวิจัยกับข้อสอบลักษณะอื่น ๆ ด้วย เช่น ข้อสอบอัตนัย (แบบเขียนตอบ) เป็นต้น

เอกสารอ้างอิง

กระทรวงศึกษาธิการ. (2560). *ตัวชี้วัดและสาระการเรียนรู้แกนกลาง กลุ่มสาระการเรียนรู้คณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551*. กรุงเทพฯ: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย.

ขนิษฐา ราศรี และประภาพร ศรีตระกูล. (2553). การสร้างแบบทดสอบวัดการคิดวิเคราะห์ทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 สังกัดสำนักงานเขตพื้นที่การศึกษาขอนแก่นเขต 4. *วารสารศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น, 33*(4), 60-70.

จารุจิตร สิทธิปรุ, ปิยะทิพย์ ตินวร และโสพส สุขานนท์สวัสดิ์. (2559). การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการจัดสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 3. *วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม, 22*(1), 47-62.

จิตติรัตน์ แสงเลิศอุทัย. (2560). คุณภาพของเครื่องมือที่ใช้ในการวิจัย. *วารสารวิจัยและพัฒนาหลักสูตร, 7*(1), 1-15.

ญานิศรา มุนินทร์สาคร. (2558). *การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6.* วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา.

นุภาพรรณ ปลื้มใจ, ปิยะทิพย์ ตินวร และโสพส สุขานนท์สวัสดิ์. (2558). การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการจัดสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 6. *วิทยาการวิจัยและวิทยาการปัญญา, 13*(2), 109-125.

ศศิธร จันทรมหา. (2560). *การสร้างข้อสอบอัตโนมัติวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 โดยใช้โปรแกรมคอมพิวเตอร์.* วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา.

สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2562). NIETS News. *วารสารจดหมายข่าวสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)*, (75), 1-8.

สุชาดา กรเพชรปาณี, ปิยะทิพย์ ตินวร และโสพส สุขานนท์สวัสดิ์. (2559). การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการจัดสอบ O-NET. *วิทยาการวิจัยและวิทยาการปัญญา, 14*(1), 14-31

สุนันทา ศิริเบญจา, ไชยรัตน์ ปราณี และดวงใจ สีเขียว. (2556). การพัฒนาแบบทดสอบปรับเหมาะโดยใช้คอมพิวเตอร์ สาระเทคโนโลยีสารสนเทศและการสื่อสาร โดยประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3. *วารสารวิชาการและวิจัยสังคมศาสตร์, 8*(22), 87-102.

เอื้อมพร หลินเจริญ, สิริศักดิ์ อาจวิชัย และภีรภา จันทร์อินทร์. (2552). *ปัจจัยเชิงสาเหตุที่ทำให้คะแนนการทดสอบ O-NET ของนักเรียนชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 6 ต่ำ.* กรุงเทพฯ: สถาบันทดสอบการศึกษาแห่งชาติ (องค์การมหาชน).

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., ... & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, abridged edition*. White Plains, NY: Longman.

Thomas, C. A. (2016). *A Comparison of Traditional Test Blueprinting to Assessment Engineering in a Large Scale Assessment Context.* Ph.D. Dissertation, The Graduate School, The University of North Carolina at Greensboro, U.S.A.

Urry, V. W. (1977). Tailored Testing: A Successful Application of Latent Trait Theory. *Journal of Education Measurement, 14*(2)*,* 181-196.

Wiersma, W., & Jurs, S. (1990). *Educational measurement and testing*. Needham Heights, MA: Allyn and Bacon.