

การพัฒนาคลังข้อสอบกลุ่มสาระการเรียนรู้คณิตศาสตร์ สำหรับระดับ มัธยมศึกษาปีที่ 4

THE DEVELOPMENT OF ITEM BANK FOR MATHEMATICS - GRADE 10

สุวิมล กฤษศยาสน์^{1*} สุวิมล ตีรกาณันท์²
Suwimon Kritkharuehart^{1*}, Suwimon Tirakanant²

¹ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

¹Department of Educational Measurement and Research, Faculty of Education, Srinakharinwirot University.

²ภาควิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

²Department of Evaluation and Research, Faculty of Education, Ramkhumhaeng University.

*Corresponding author, E-mail: monmeasure@gmail.com

บทคัดย่อ

การศึกษานี้มีวัตถุประสงค์หลักเพื่อพัฒนาคลังข้อสอบกลุ่มสาระการเรียนรู้คณิตศาสตร์ตามทฤษฎีการตอบสนองข้อสอบ ดำเนินการเก็บรวบรวมข้อมูลจากนักเรียนชั้นมัธยมศึกษาปีที่ 4 จำนวน 3,466 ด้วยข้อสอบปรนัย จำนวน 240 ข้อ มีดัชนีความสอดคล้องระหว่างข้อสอบกับวัตถุประสงค์ 0.67 ถึง 1.00 สถิติที่ใช้คือ การวิเคราะห์ความเป็นเอกมิติ และโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ ผลการศึกษาพบว่าข้อสอบทั้งหมดวัดคุณลักษณะเด่นเดียว จำนวนข้อสอบที่ถูกคัดเลือกเข้าคลังข้อสอบตามเกณฑ์ของทฤษฎีการตอบสนองข้อสอบ คือ 81 ข้อ โดยมีค่าพารามิเตอร์ความยากระหว่าง (-3.00) ถึง 3.00 และค่าพารามิเตอร์อำนาจจำแนกระหว่าง 0.05 – 2.50 และข้อสอบในคลังจำแนกตามเนื้อหาย่อย คือเรื่องความรู้เบื้องต้นเกี่ยวกับเซต จำนวน 30 ข้อ คิดเป็นร้อยละ 37.04 การดำเนินการของเซต จำนวน 42 ข้อคิดเป็นร้อยละ 51.85 และการแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์-ออยเลอร์ จำนวน 9 ข้อ คิดเป็นร้อยละ 11.11 นอกจากนี้คณะผู้วิจัยได้เสนอการใช้ประโยชน์จากคลังข้อสอบในเบื้องต้น คือการจัดชุดแบบทดสอบจากคลังข้อสอบแบบที่เป็นระบบ พบว่าแบบทดสอบตัวอย่างจำนวน 10 ฉบับ ที่ได้จากระบบดังกล่าวมีฟังก์ชันสารสนเทศของแบบทดสอบ 1.190 ถึง 2.379 ความคลาดเคลื่อนมาตรฐานของการประมาณค่า 0.195 ถึง 0.483 และแบบทดสอบทั้ง 10 ฉบับ มีความเชื่อมั่นของการประมาณค่าความสามารถของผู้สอบ 0.721 ถึง 0.815

คำสำคัญ: คลังข้อสอบ คณิตศาสตร์ ทฤษฎีการตอบสนองข้อสอบ

Abstract

The main purpose of this research was to develop an item bank in mathematics with item response theory. Data collection was performed with 3,466 grade 10 students by 240 multiple choice test items with IOC of 0.67 - 1.00. The statistics for data analyzing included dimensionality and 2-parameters logistic item response model. The results showed that all test items were unidimensional. An item bank included 81 test items were qualified with IRT. The difficulty parameter value and discrimination power vary from

(-3.000) - 3.000 and 0.50 - 2.50 respectively. The 30 items (37.04%) were based on basic knowledge in Set content, the 42 items (51.85%) were based on procedural in Set content, and the 9 items (11.11%) were based on problem solving with Venn-Euler diagrams content. Furthermore, the researchers have suggested that the primary utilization of an item bank should be a systematic test assembly. In this research, there were 10 sample tests including test information function of 1.190 - 2.379, and the standard errors of estimation were 0.195 - 0.483. Finally, the empirical reliability of 10 sample tests was 0.721- 0.815.

Keywords: Item Bank, Mathematics, Item Response Theory

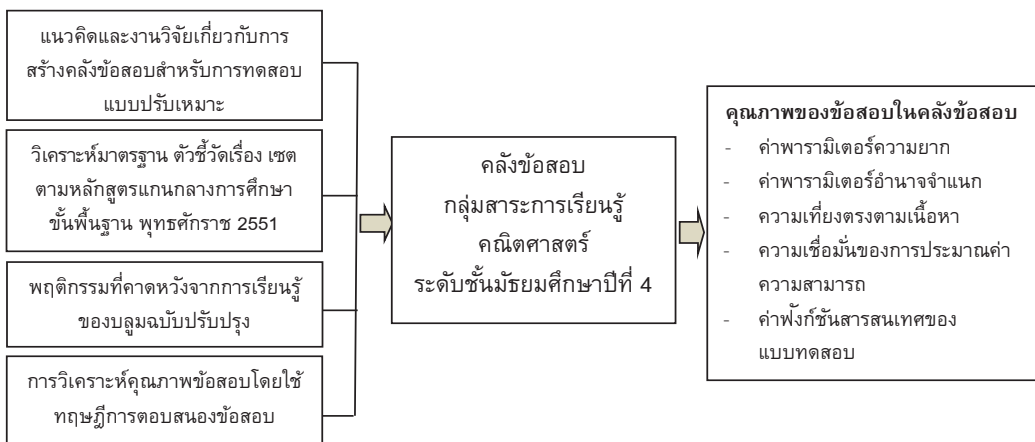
บทนำ

การวัดและประเมินผลการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 มุ่งเน้นกระบวนการเก็บรวบรวม ตรวจสอบ ตีความผลการเรียนรู้ และพัฒนาการด้านต่าง ๆ ของผู้เรียนตามมาตรฐานการเรียนรู้หรือตัวชี้วัดของหลักสูตรนำผลไปปรับปรุงพัฒนาการจัดการเรียนรู้และใช้เป็นข้อมูลสำหรับการตัดสินใจผลการเรียน สถานศึกษาต้องมีกระบวนการจัดการที่เป็นระบบ เพื่อให้การดำเนินการวัดและประเมินผลการเรียนรู้เป็นไปอย่างมีคุณภาพและประสิทธิภาพ และให้ผลการประเมินที่ตรงตามความรู้ความสามารถที่แท้จริงของผู้เรียน ถูกต้องตามหลักการวัดและประเมินผลการเรียนรู้ รวมทั้งสามารถรองรับการประเมินภายในและการประเมินภายนอกตามระบบประกันคุณภาพการศึกษาได้ [1] ดังนั้นผู้บริหารสถานศึกษาต้องถือเอางานวิชาการมาเป็นอันดับแรกเพราะหน้าที่ของสถานศึกษา คือการจัดการกระบวนการเรียนรู้ให้ผู้เรียนอย่างมีคุณภาพและประสิทธิภาพสูงสุด การเป็นผู้นำทางด้านวิชาการโดยการทำงานร่วมกับครู กระตุ้นเตือนครู และประสานงานให้ครูทุกคนทำงานร่วมกันเพื่อมุ่งจุดหมายอันเดียวกัน คือผู้เรียน คุณภาพงานวิชาการจึงเป็นงานที่พิสูจน์ความสำเร็จของสถานศึกษาหรืออีกนัยหนึ่งความสำเร็จของสถานศึกษาคือ เด็กที่เป็นผลผลิตของโรงเรียน

มีคุณภาพ และมีคุณสมบัติตามที่หลักสูตรกำหนด คือมีผลสำเร็จทางการเรียนสูง หรือเป็นไปตามมาตรฐานการเรียนรู้ที่กำหนดไว้ [2]

จากการศึกษาของสมวงษ์ แปลงประสพ โชค สมเดช บุญประจักษ์ และจรรยา ภูอุตม [3] ได้ศึกษาสำรวจความเห็นเกี่ยวกับสาเหตุที่เด็กไทยอ่อนคณิตศาสตร์ โดยเก็บข้อมูลจากการจัดสัมมนาครู หัวหน้าหมวดคณิตศาสตร์ นักการศึกษาคณิตศาสตร์ จำนวน 118 คน สัมภาษณ์ อาจารย์ นักการศึกษาและผู้เชี่ยวชาญจาก 6 มหาวิทยาลัย จำนวน 12 คน พบว่าปัญหาด้านวิธีการวัดผลและประเมินผลของครู คือครูใช้วิธีการวัดผลไม่ตรงกับเนื้อหาหรือกิจกรรมที่สอน ครูออกข้อสอบในสิ่งที่ไม่เกี่ยวข้องกับเรื่องที่สอน และข้อสอบยาก ครูขาดการนิเทศระหว่างกัน และกัน จากปัญหาดังกล่าวการสร้างคลังข้อสอบในโรงเรียนจะทำให้สถานศึกษาสามารถพัฒนาคุณภาพด้านการวัดผลและประเมินผล รวมทั้งงานด้านการเรียนการสอนได้ Gronlund [4] กล่าวว่า ประโยชน์ของคลังข้อสอบ คือช่วยให้สถานศึกษาจัดเก็บข้อสอบที่มีคุณภาพไว้ในคลังข้อสอบโดยจัดเป็นหมวดหมู่ตามเนื้อหาสาระ ระดับชั้น จุดประสงค์การเรียนรู้ ระดับความยากง่าย และระดับอำนาจจำแนกของข้อสอบ การพัฒนาคลังข้อสอบที่ดีนอกจากจะแสดงถึงความก้าวหน้าทางการบริหารวิชาการของโรงเรียนแล้วยังแสดงให้เห็นได้

อีกด้วยว่าโรงเรียนมีระบบบริหารงานวิชาการที่ดี และมีการดำเนินการเรียนการสอนที่ยึดหลักเกณฑ์ หรือหลักการในศาสตร์ทางการศึกษาโดยตรงมาก น้อยเพียงใด [5] จากความสำคัญและประโยชน์ ของคลังข้อสอบคณะผู้วิจัยจึงพัฒนาคลังข้อสอบ กลุ่มสาระการเรียนรู้คณิตศาสตร์เรื่องเซตเนื่องจาก ความรู้เรื่องเซตที่สอนในระดับชั้นมัธยมศึกษาปีที่ 4 เป็นความรู้พื้นฐานที่สามารถนำไปประยุกต์ใช้แก้ โจทย์ปัญหาเกี่ยวกับจำนวนจริง ความสัมพันธ์และ ฟังก์ชัน และการให้เหตุผล จากการประมวลเอกสาร ข้างต้น คณะผู้วิจัยได้กำหนดเป็นกรอบแนวคิดการ วิจัยดังภาพที่ 1



ภาพที่ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ดำเนินการเก็บข้อมูลจากนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 จำนวน 3,466 คน ในเขตกรุงเทพมหานคร จำนวน 10 โรงเรียน ซึ่งเป็นโรงเรียนที่ยินดีเข้าร่วมการทดลองใช้ข้อสอบ และเป็นโรงเรียนขนาดกลาง และขนาดใหญ่ คือ 1) โรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร (ฝ่ายมัธยม) 2) โรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ ปทุมวัน 3) โรงเรียนเทพศิลา 4) โรงเรียนบางกะปิ 5) โรงเรียนศรีอยุธยา ในพระอุปถัมภ์ สมเด็จพระเจ้าภคินีเธอ เจ้าฟ้าเพชรรัตนราชสุดา สิริโสภาพัณณวดี 6) โรงเรียนนนทรีวิทยา 7) โรงเรียนวัดประตู่ในทรงธรรม 8)

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์หลักเพื่อพัฒนาคลังข้อสอบกลุ่มสาระการเรียนรู้คณิตศาสตร์เนื้อหาเรื่องเซตสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4 ตามทฤษฎีการตอบสนองข้อสอบ โดยมีวัตถุประสงค์ย่อย คือ

1. เพื่อสร้างข้อสอบและจัดกลุ่มข้อสอบตามเนื้อหาการวัด
2. เพื่อตรวจสอบคุณภาพของข้อสอบในคลังข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

โรงเรียนวัดราชบพิศ 9) โรงเรียนเทพศิรินทร์ร่วมเกล้า และ 10) โรงเรียนเตรียมอุดมศึกษาสุวินทวงศ์ การใช้จำนวนตัวอย่างเป็นไปตามข้อเสนอแนะของ Baker [6] ที่กล่าวว่าควรใช้ขนาดตัวอย่างมากกว่าหรือเท่ากับ 1,000 คน สำหรับการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการตอบสนองข้อสอบ

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการสร้างคลังข้อสอบ คือ ข้อสอบปรนัยเรื่องเซต จำนวน 240 ข้อ จัดพิมพ์เป็นแบบทดสอบ 12 ฉบับ ฉบับละ 20 ข้อ โดยโครงสร้างของแบบทดสอบทั้ง 12 ฉบับมีความเป็นคู่ขนานของแบบทดสอบดังตารางที่ 1

ตารางที่ 1 โครงสร้างของแบบทดสอบเรื่องเซต

ฉบับ	เนื้อหาย่อย	ข้อที่
1	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 5
	การดำเนินการของเซต	4, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 19
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	13, 14, 20
2	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 4
	การดำเนินการของเซต	5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	11, 19, 20
3	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 3, 4, 5
	การดำเนินการของเซต	2, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	18, 19, 20
4	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 12
	การดำเนินการของเซต	4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	18, 19, 20
5	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 4
	การดำเนินการของเซต	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	18, 19, 20
6	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 4
	การดำเนินการของเซต	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	18, 19, 20
7	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 5
	การดำเนินการของเซต	4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	18, 19, 20
8	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 4
	การดำเนินการของเซต	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	18, 19, 20
9	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 4
	การดำเนินการของเซต	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	18, 19, 20

ฉบับ	เนื้อหาย่อย	ข้อที่
10	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 4
	การดำเนินการของเซต	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	17, 19, 20
11	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 4
	การดำเนินการของเซต	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	18, 19, 20
12	ความรู้เบื้องต้นเกี่ยวกับเซต	1, 2, 3, 4
	การดำเนินการของเซต	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	การแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์	18, 19, 20

ผลการตรวจสอบความเที่ยงตรงตามเนื้อหา (Content Validity) จากผู้เชี่ยวชาญจำนวน 3 ท่าน ซึ่งดำเนินการพิจารณาถึงความสอดคล้องของข้อสอบกับวัตถุประสงค์ในการวัดตามระดับความรู้และกระบวนการทางพุทธิปัญญาตามแนวคิดของบลูมฉบับปรับปรุง พบว่าคลังข้อสอบเรื่องเซต ซึ่งประกอบด้วยข้อสอบปรนัยจำนวน 240 ข้อ มีดัชนีความสอดคล้องระหว่างข้อสอบกับวัตถุประสงค์ในการวัด (IOC) ตั้งแต่ 0.67 – 1.00

ผลการวิจัย

ตอนที่ 1 ผลการตรวจสอบความสอดคล้องระหว่างข้อสอบกับการระบุความรู้และกระบวนการทางพุทธิปัญญาตามแนวคิดของบลูมฉบับปรับปรุง

ผลการตรวจสอบจากผู้เชี่ยวชาญเกี่ยวกับการระบุระดับความรู้และกระบวนการทางพุทธิปัญญาตามแนวคิดของบลูมฉบับปรับปรุง พบว่าส่วนใหญ่เป็นข้อสอบของเนื้อหาย่อยที่ 2 การดำเนินการของเซต จำนวน 156 ข้อ คิดเป็นร้อยละ 65 รองลงมาคือเนื้อหาย่อยที่ 1 ความรู้เบื้องต้น

ต้นเกี่ยวกับเซต จำนวน 48 ข้อ คิดเป็นร้อยละ 20 เมื่อพิจารณาจำนวนข้อสอบตามความรู้และกระบวนการทางพุทธิปัญญา พบว่าเนื้อหาย่อยที่ 2 การดำเนินการของเซต มีข้อสอบส่วนใหญ่วัดความรู้เชิงข้อเท็จจริงและกระบวนการทางพุทธิปัญญาด้านความเข้าใจ จำนวน 36 ข้อ คิดเป็นร้อยละ 15 รองลงมาคือวัดความรู้เชิงวิธีดำเนินการและกระบวนการทางพุทธิปัญญาด้านความเข้าใจ จำนวน 35 ข้อ คิดเป็นร้อยละ 14.6 ส่วนเนื้อหาย่อยที่ 1 ความรู้เบื้องต้นเกี่ยวกับเซต มีข้อสอบส่วนใหญ่วัดความรู้เชิงข้อเท็จจริงและกระบวนการทางพุทธิปัญญาด้านความจำ จำนวน 31 ข้อ คิดเป็นร้อยละ 12.9 รองลงมาคือวัดความรู้เชิงข้อเท็จจริงและกระบวนการทางพุทธิปัญญาด้านความเข้าใจ จำนวน 10 ข้อ คิดเป็นร้อยละ 4.2

ตอนที่ 2 ผลการตรวจสอบความเป็นคู่ขนานของแบบทดสอบ และการตรวจสอบข้อตกลงเบื้องต้นตามทฤษฎีการตอบสนองข้อสอบ

ความเป็นคู่ขนานของแบบทดสอบในการวิจัยครั้งนี้ตรวจสอบด้วยการทดสอบความแตกต่างของค่าพารามิเตอร์ความยากเฉลี่ยของแบบทดสอบ (\bar{b}) พบว่าแบบทดสอบทั้ง 12 ฉบับ

มีค่าพารามิเตอร์ความยากเฉลี่ยไม่แตกต่างกันนั้น คือแบบทดสอบทั้ง 12 ฉบับมีความเป็นคู่ขนานกัน การตรวจสอบข้อตกลงเบื้องต้นตามทฤษฎีการตอบสนองข้อสอบดำเนินการดังนี้ 1) ความเป็นเอกมิติของข้อสอบที่บรรจุในคลังข้อสอบตรวจสอบด้วยวิธี eigen value test พบว่าแบบทดสอบเรื่องเซต จำนวน 240 ข้อวัดคุณลักษณะเด่นเดียว (Single Dominant Factor) ซึ่งพิจารณาจากค่า eigen ที่ 1 ต่างจากที่เหลือมากๆ และค่า eigen ที่เหลือมีลักษณะเหมือนกัน รวมทั้งอัตราส่วนระหว่างค่า eigen ที่ 1 กับค่า eigen ที่ 2 เท่ากับ 3.61 ซึ่งอัตราส่วนดังกล่าวมีค่ามากกว่า 3 เป็นตามเกณฑ์ของ Gorsuch อ้างอิงใน Slocum [7] 2) ความเป็นอิสระในการตอบข้อสอบ เมื่อแบบทดสอบมีความเป็นเอกมิติแล้วความเป็นอิสระในการตอบข้อสอบก็จะมีเช่นกัน Hambleton and Swaminathan (อ้างอิงใน เกียรติศักดิ์ ส่องแสง) [8] 3) การตรวจสอบความสอดคล้องของข้อมูลกับโมเดล (Model Data Fit Assessment) พิจารณาทั้งความไม่แปรเปลี่ยนของค่าประมาณความสามารถ (Invariance of Ability Parameter estimate) และความไม่แปรเปลี่ยนของค่าประมาณค่าพารามิเตอร์ของข้อสอบ (Invariance of Item Parameter Estimates) [9] 4) ดำเนินการด้วยการสุ่มผู้สอบเป็น 2 กลุ่มย่อยและเปรียบเทียบโค้งของฟังก์ชันการตอบสนองข้อสอบ (Item Response Function: IRF) ของกลุ่มย่อย 2 กลุ่ม และพิจารณาค่า RMSD (Root Mean Squared Difference) ของค่าพารามิเตอร์อำนาจจำแนก (a) กับค่าพารามิเตอร์ความยาก (b) ของข้อสอบเป็นรายข้อ พบว่าข้อสอบจำนวน 240 ข้อ มีค่า RMSD อยู่ระหว่าง .001 ถึง .048 ซึ่งมีค่าน้อยกว่า .050 นั่นคือข้อมูลมีความเหมาะสมกับโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์

ตอนที่ 3 ผลการวิเคราะห์ค่าพารามิเตอร์

ของข้อสอบ และการคัดเลือกข้อสอบเข้าคลังข้อสอบ

ข้อสอบส่วนใหญ่มีค่าพารามิเตอร์ความยากเหมาะสม $(-3.000) \leq b \leq 3.000$ จำนวน 184 ข้อ คิดเป็นร้อยละ 76.70 รองลงมาคือข้อสอบยากมาก $(b \geq 3.001)$ จำนวน 55 ข้อ คิดเป็นร้อยละ 22.90 ข้อสอบส่วนใหญ่วัดเนื้อหาเกี่ยวกับการดำเนินการของเซต จำนวน 157 ข้อ คิดเป็นร้อยละ 65.40 รองลงมาวัดเนื้อหาเกี่ยวกับความรู้เบื้องต้นเกี่ยวกับเซต จำนวน 45 ข้อ คิดเป็นร้อยละ 18.80 สำหรับค่าพารามิเตอร์อำนาจจำแนกส่วนใหญ่ข้อสอบมีอำนาจจำแนกได้ต่ำ $(a < 0.50)$ จำนวน 159 ข้อ คิดเป็นร้อยละ 66.30 รองลงมาข้อสอบมีอำนาจจำแนกได้ดี $(0.50 \leq a \leq 2.50)$ จำนวน 81 ข้อ คิดเป็นร้อยละ 33.80 การคัดเลือกข้อสอบเข้าคลังข้อสอบเป็นไปตามเกณฑ์ของ ศิริชัย กาญจนวาสี [10] และ De Ayala [9] พิจารณาจากค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง 0.50 ถึง 2.50 ค่าพารามิเตอร์ความยาก (b) อยู่ระหว่าง (-3.00) ถึง 3.00 พบว่าข้อสอบที่สร้างขึ้นทั้งหมด จำนวน 240 ข้อ สามารถนำข้อสอบไปบรรจุในคลังได้ทั้งหมด 81 ข้อ จำแนกเป็นเนื้อหาย่อยเรื่องความรู้เบื้องต้นเกี่ยวกับเซต จำนวน 30 ข้อ คิดเป็นร้อยละ 37.04 การดำเนินการของเซต จำนวน 42 ข้อ คิดเป็นร้อยละ 51.85 และการแก้โจทย์ปัญหาโดยใช้แผนภาพเวนนี – ออยเลอร์ จำนวน 9 ข้อ คิดเป็นร้อยละ 11.11

ตอนที่ 4 การจัดชุดแบบทดสอบ

การใช้ประโยชน์จากคลังข้อสอบในเบื้องต้น คือ การจัดชุดแบบทดสอบจากคลังข้อสอบแบบที่เป็นระบบตามที่คณะผู้วิจัยได้จัดทำขึ้น ซึ่งมีขั้นตอนการจัดชุดแบบทดสอบดังนี้

- 1) การแจกแจงจำนวนข้อสอบตามตารางกำหนดคุณลักษณะเฉพาะและข้อสอบที่นำมาบรรจุในคลังข้อสอบ ดังตารางที่ 2

ตารางที่ 2 จำนวนข้อสอบตามตารางกำหนดคุณลักษณะเฉพาะและจำนวนข้อสอบที่นำมาบรรจุในคลังข้อสอบ

เนื้อหาย่อย	จำนวนข้อสอบต่อฉบับตามตารางกำหนดคุณลักษณะเฉพาะ	จำนวนข้อสอบที่ผ่านเกณฑ์และนำมาบรรจุในคลังข้อสอบ
ความรู้เบื้องต้นเกี่ยวกับเซต	4	30
การดำเนินการของเซต	13	42
การแก้โจทย์ปัญหาโดยใช้แผนภาพเวเนน-ออยเลอร์	3	9
รวม	20	81

2) นำเสนอจำนวนรูปแบบจัดหมู่ข้อสอบจำแนกตามเนื้อหาย่อยของเซต

คณะผู้วิจัยได้ประยุกต์ใช้หลักการของการจัดหมู่ (Combination) คือ nCr มาใช้เป็นแนวทางในการคำนวณจำนวนรูปแบบของการจัดหมู่ข้อสอบจำแนกตามเนื้อหาย่อยของเซตที่เป็นไปได้ทั้งหมดจากคลังข้อสอบโดยที่

n คือ จำนวนข้อสอบของเนื้อหาย่อยใดๆ ที่ผ่านเกณฑ์และนำมาบรรจุในคลังข้อสอบ

r คือ จำนวนข้อสอบของเนื้อหาย่อยใดๆ ตามตารางกำหนดคุณลักษณะเฉพาะ

จากหลักการข้างต้นสามารถนำมากำหนดเป็นจำนวนรูปแบบของการจัดหมู่ข้อสอบจำแนกตามเนื้อหาย่อยของเซต ดังนี้

2.1) เนื้อหาย่อยที่ 1 ความรู้เบื้องต้นเกี่ยวกับเซต จากตารางที่ 2 พบว่าเนื้อหานี้มีจำนวนข้อสอบที่ผ่านเกณฑ์ (n) คือ 30 ข้อ ส่วนจำนวนข้อสอบตามตารางกำหนดคุณลักษณะเฉพาะ (r) คือ 4 ข้อ หากผู้สอนต้องการสุ่มข้อสอบตามข้อกำหนดดังกล่าว สามารถดำเนินการจัดชุดข้อสอบ ดังนี้

ตัวอย่างที่ 1 กรณีผู้สอนต้องการข้อสอบตามตารางกำหนดคุณลักษณะเฉพาะจำนวน 4 ข้อ นั้นมีค่าพารามิเตอร์ความยากระดับ (-3.000) ถึง (-1.001) จำนวน 1 ข้อ ค่าพารามิเตอร์ความยาก

ระดับ (-1.000) ถึง (1.000) จำนวน 2 ข้อ และค่าพารามิเตอร์ความยากระดับ (1.001) ถึง (3.000) จำนวน 1 ข้อ รูปแบบในการจัดหมู่ของเนื้อหาย่อยที่ 1 ความรู้เบื้องต้นเกี่ยวกับเซต คือ ${}^30C_1 \cdot {}^{18}C_2 \cdot {}^2C_1$ ดังนั้นจำนวนรูปแบบในการจัดหมู่ของเนื้อหาย่อยที่ 1 ความรู้เบื้องต้นเกี่ยวกับเซต คือ 3,060 รูปแบบ

2.2) เนื้อหาย่อยที่ 2 ความรู้เบื้องต้นเกี่ยวกับเซต จากตารางที่ 2 พบว่าเนื้อหาย่อยที่ 2 การดำเนินการของเซต มีจำนวนข้อสอบที่ผ่านเกณฑ์ (n) คือ 42 ข้อ ส่วนจำนวนข้อสอบตามตารางกำหนดคุณลักษณะเฉพาะ (r) คือ 13 ข้อ หากผู้สอนต้องการสุ่มข้อสอบตามข้อกำหนดดังกล่าว สามารถดำเนินการจัดชุดข้อสอบ ดังนี้

ตัวอย่างที่ 2 กรณีผู้สอนต้องการข้อสอบตามตารางกำหนดคุณลักษณะเฉพาะจำนวน 13 ข้อ นั้นมีค่าพารามิเตอร์ความยากระดับ (-3.000) ถึง (-1.001) จำนวน 3 ข้อ ค่าพารามิเตอร์ความยากระดับ (-1.000) ถึง (1.000) จำนวน 9 ข้อ และค่าพารามิเตอร์ความยากระดับ (1.001) ถึง (3.000) จำนวน 1 ข้อ รูปแบบในการจัดหมู่ของเนื้อหาย่อยที่ 2 การดำเนินการของเซต คือ ${}^{42}C_3 \cdot {}^{38}C_9 \cdot {}^1C_1$ ดังนั้นจำนวนรูปแบบในการจัดหมู่ของเนื้อหาย่อยที่ 2 การดำเนินการของเซต คือ 49,252,180 รูปแบบ

2.3) เนื้อหาย่อยที่ 3 การแก้โจทย์ปัญหาโดยใช้แผนภาพเวเนน-ออยเลอร์

จากตารางที่ 2 พบว่าเนื้อหาย่อยที่ 3 การแก้ไขข้อผิดพลาดโดยใช้แผนภาพเวเน-ออยเลอร์มีจำนวนข้อสอบที่ผ่านเกณฑ์ (n) คือ 9 ข้อ ส่วนจำนวนข้อสอบตามตารางกำหนดคุณลักษณะเฉพาะ (r) คือ 3 ข้อ หากผู้สอนต้องการสุ่มข้อสอบตามข้อกำหนดดังกล่าวสามารถดำเนินการจัดชุดข้อสอบ ดังนี้

ตัวอย่างที่ 3 กรณีผู้สอนต้องการข้อสอบตามตารางกำหนดคุณลักษณะเฉพาะจำนวน 3 ข้อ นั้น มีค่าพารามิเตอร์ความยากระดับ (-1.000) ถึง (1.000) จำนวน 2 ข้อ และค่าพารามิเตอร์ความยากระดับ (1.001) ถึง (3.000) จำนวน 1 ข้อ รูปแบบในการจัดหมู่ของเนื้อหาย่อยที่ 3 การแก้ไขข้อผิดพลาดโดยใช้แผนภาพเวเน-ออยเลอร์ คือ ${}^7C_2 \cdot {}^2C_1$ ดังนั้นจำนวนรูปแบบในการจัดหมู่ของเนื้อหาย่อยที่ 3 การแก้ไขข้อผิดพลาดโดยใช้แผนภาพเวเน-ออยเลอร์ คือ 42 รูปแบบ

สรุปจากตัวอย่างที่ 1 – 3 แสดงให้เห็นว่าจำนวนรูปแบบในการจัดหมู่ข้อสอบ ทำให้ครูผู้สอนสุ่มข้อสอบจากคลังเพื่อจัดหมู่ข้อสอบตามเนื้อหาย่อยที่ 1-3 คือ เนื้อหาย่อยที่ 1 ความรู้เบื้องต้น

เกี่ยวกับเซต สามารถทำการสุ่มและจัดหมู่ข้อสอบทั้งหมด จำนวน 3,060 รูปแบบ เนื้อหาย่อยที่ 2 การดำเนินการของเซต สามารถทำการสุ่มและจัดหมู่ข้อสอบทั้งหมด จำนวน 49,252,180 รูปแบบ และเนื้อหาย่อยที่ 3 การแก้ไขข้อผิดพลาดโดยใช้แผนภาพเวเน-ออยเลอร์ สามารถทำการสุ่มและจัดหมู่ข้อสอบทั้งหมด จำนวน 42 รูปแบบ การใช้คลังข้อสอบชุดนี้จึงเป็นประโยชน์ต่อการทำแบบทดสอบทำให้ครูผู้สอนเกิดความสะดวกและประหยัดเวลาในการปฏิบัติงานในช่วงเวลาที่มีภาระงานมาก

ตอนที่ 5 ฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information) ของตัวอย่างแบบทดสอบจากการสุ่มข้อสอบ จำนวน 10 ฉบับ

จากแนวทางของการใช้ประโยชน์จากคลังข้อสอบด้วยการจัดชุดแบบทดสอบจากคลังข้อสอบแบบที่เป็นระบบตามที่คณะผู้วิจัยได้จัดทำขึ้นนั้น คณะผู้วิจัยจึงดำเนินการทดลองสุ่มข้อสอบตามตารางการกำหนดคุณลักษณะของแบบทดสอบแล้วนำจัดเป็นชุดแบบทดสอบ โดยการกำหนดจำนวนข้อสอบตามเนื้อหาย่อยแสดงดังตารางที่ 3

ตารางที่ 3 จำนวนข้อสอบตามเนื้อหาย่อยจากตารางการกำหนดคุณลักษณะของแบบทดสอบเรื่อง เซต

เนื้อหาย่อย	จำนวนข้อสอบต่อฉบับตามตารางกำหนดคุณลักษณะเฉพาะ
ความรู้เบื้องต้นเกี่ยวกับเซต	4
การดำเนินการของเซต	13
การแก้ไขข้อผิดพลาดโดยใช้แผนภาพเวเน-ออยเลอร์	3
รวม	20

การทดลองสุ่มข้อสอบตามตารางการกำหนดคุณลักษณะของแบบทดสอบครั้งนี้ดำเนินการจัดเป็นชุดแบบทดสอบ จำนวน 10 ฉบับและนำมาวิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information) ความคลาดเคลื่อนมาตรฐานของการประมาณค่า

(Standard Error of Estimation) และความเชื่อมั่นของการประมาณค่าความสามารถของผู้สอบ (Empirical Reliability) การวิเคราะห์ตามทฤษฎี IRT จะใช้แบบแผนการตอบสนองแบบทดสอบเป็นรายชื่อในการประมาณค่าความสามารถของผู้สอบ ดังนั้นการประเมินคุณภาพของแบบทดสอบ

จึงสามารถพิจารณาจากความถูกต้องแม่นยำในการประมาณความสามารถของผู้ตอบ โดยใช้ดัชนีตัวหนึ่งเรียกว่า ค่าฟังก์ชันสารสนเทศของแบบทดสอบ [Test Information: $I(\Theta)$] สำหรับความคลาดเคลื่อนมาตรฐานของการประมาณค่า [SE(Θ)] เป็นค่าเบี่ยงเบนมาตรฐานของการแจกแจงความน่าจะเป็นของค่าประมาณความสามารถที่แท้จริง (Θ) สำหรับค่าที่เหมาะสมสำหรับการคัดเลือกข้อสอบคือควรมีค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าไม่เกิน 0.50 [11] ค่าความเชื่อมั่นของการประมาณค่าความสามารถ

(Empirical Reliability) เป็นค่าที่บ่งบอกถึงความน่าเชื่อถือของการประมาณค่าความสามารถของผู้สอบ (Θ) โดยคำนวณได้จากอัตราส่วนของความแปรปรวนของคะแนนความสามารถที่ประมาณค่าได้กับผลรวมของความแปรปรวนของคะแนนความสามารถที่ประมาณค่าได้กับความแปรปรวนของความคลาดเคลื่อน และหากนำค่าความเชื่อมั่นของการประมาณค่ามายกกำลังสองและคูณด้วยร้อยแล้ว เป็นค่าที่บ่งบอกถึงประสิทธิภาพของการประมาณค่าความสามารถของผู้สอบ[9, 12] ดังตารางที่ 4

ตารางที่ 4 ฟังก์ชันสารสนเทศของแบบทดสอบ ความคลาดเคลื่อนมาตรฐานของการประมาณค่าและค่าความเชื่อมั่นของการประมาณค่าความสามารถ

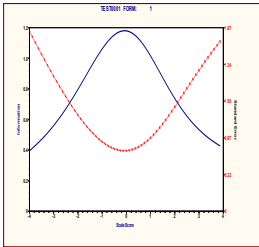
ฉบับ	ฟังก์ชันสารสนเทศของแบบทดสอบ	ความคลาดเคลื่อนมาตรฐานของการประมาณค่า	ค่าความเชื่อมั่นของการประมาณค่าความสามารถ
1	1.190	0.425	0.752
2	1.706	0.390	0.772
3	2.379	0.195	0.815
4	1.492	0.483	0.721
5	1.336	0.476	0.740
6	1.602	0.360	0.786
7	1.937	0.350	0.795
8	2.349	0.328	0.798
9	1.789	0.380	0.777
10	1.377	0.437	0.748

จากตารางที่ 4 พบว่าตัวอย่างของแบบทดสอบ จำนวน 10 ฉบับ มีฟังก์ชันสารสนเทศของแบบทดสอบอยู่ระหว่าง 1.190 ถึง 2.379 ความคลาดเคลื่อนมาตรฐานของการประมาณค่าอยู่ระหว่าง 0.195 ถึง 0.483 และแบบทดสอบทั้ง 10 ฉบับมีความเชื่อมั่นของการประมาณค่า

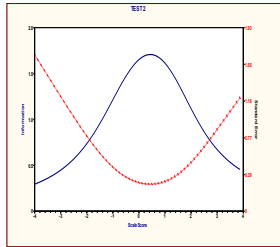
ความสามารถของผู้สอบ (Empirical Reliability) อยู่ระหว่าง 0.721 ถึง 0.815 นอกจากนี้แบบทดสอบฉบับที่ 1, 4, 7, 8 และ 10 เหมาะสำหรับการทดสอบกับผู้สอบที่มีความสามารถ (Θ) ระหว่าง (-1.00) ถึง 1.00 แบบทดสอบฉบับที่ 3 เหมาะสำหรับการทดสอบกับผู้สอบที่มีความสามารถ (Θ)

ระหว่าง (-2.50) ถึง (-0.50) แบบทดสอบฉบับที่ 5 เหมาะสำหรับการทดสอบกับผู้สอบที่มีความสามารถ (Θ) ระหว่าง (-1.50) ถึง 0.50 แบบทดสอบฉบับที่ 2 เหมาะสำหรับการทดสอบกับผู้สอบที่มีความสามารถ (Θ) ระหว่าง (-0.50) ถึง 1.50

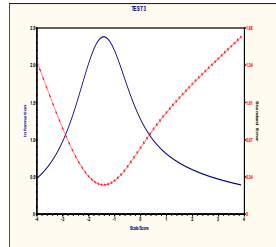
แบบทดสอบฉบับที่ 6 เหมาะสำหรับการทดสอบกับผู้สอบที่มีความสามารถ (Θ) ระหว่าง 0.00 ถึง 1.50 และแบบทดสอบฉบับที่ 9 เหมาะสำหรับการทดสอบกับผู้สอบที่มีความสามารถ (Θ) ระหว่าง 0.50 ถึง 1.50 รายละเอียดดังภาพที่ 2 ถึง 11



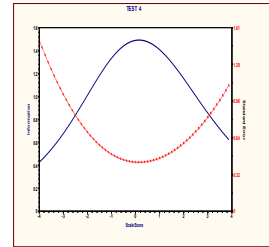
ภาพที่ 2 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 1



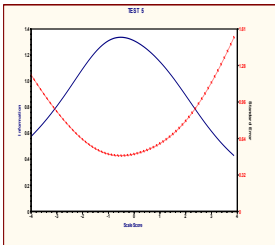
ภาพที่ 3 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 2



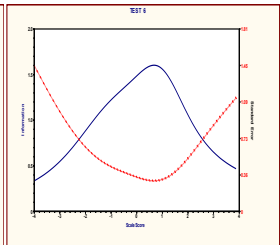
ภาพที่ 4 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 3



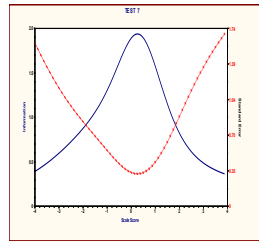
ภาพที่ 5 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 4



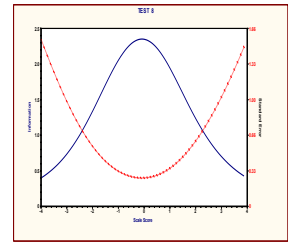
ภาพที่ 6 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 5



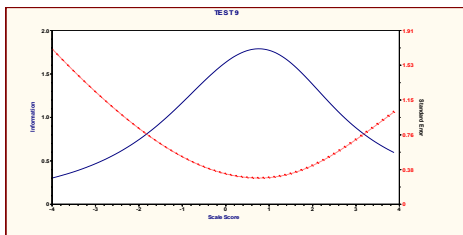
ภาพที่ 7 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 6



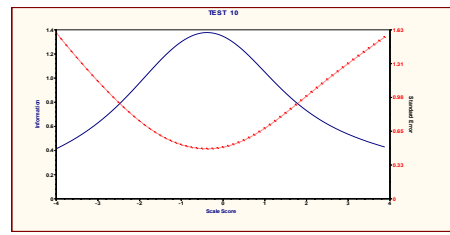
ภาพที่ 8 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 7



ภาพที่ 9 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 8



ภาพที่ 10 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 9



ภาพที่ 11 ฟังก์ชันสารสนเทศของแบบทดสอบฉบับที่ 10

ตอนที่ 6 แนวทางการนำข้อสอบที่ยากมากไปใช้ในทางปฏิบัติ

ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบที่สามารถบรรจุไว้ในคลังข้อสอบ จำนวน 81 ข้อ พบว่ามีข้อสอบ จำนวน 55 ข้อที่มีค่าพารามิเตอร์ความยาก (b) ตั้งแต่ 3.001 ถึง 4.530 ซึ่งในทางปฏิบัติถือว่าเป็นข้อสอบที่ยากมาก โดยส่วนใหญ่เป็นข้อสอบวัดที่วัดความรู้วิธีดำเนินการและกระบวนการทาง

พุทธิปัญญาด้านการประยุกต์ใช้และด้านการวิเคราะห์ ดังนั้นคณะผู้วิจัยจึงทำการประชุมกลุ่มย่อยร่วมกับครูผู้สอนคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 4 ที่มีประสบการณ์สอนมากกว่า 2 ปี จำนวน 3 ท่านเพื่อระดมความคิดเห็นเกี่ยวกับแนวทางการจัดการเรียนรู้ให้สอดคล้องกับกระบวนการทางพุทธิปัญญาที่ผู้เรียนใช้ในการแก้โจทย์ปัญหาเพื่อให้ผู้เรียนสามารถถูกวัดและประเมินด้วยข้อสอบที่ค่อนข้างยาก

ผลการประชุมกลุ่มย่อยพบว่า โดยส่วนใหญ่กระบวนการทางพุทธิปัญญาที่ผู้เรียนต้องใช้ในการตอบข้อสอบยาก คือผู้เรียนต้องสามารถจดจำและมีความเข้าใจในเรื่องราวที่เป็นสมาชิกของเซต จำนวนสมาชิก สับเซต เพาเวอร์เซต และการดำเนินการของเซต ในกรณีที่ผู้เรียนตอบผิดมักเกิดจากผู้เรียนคิดว่าเรื่องความเป็นสมาชิกของเซตและการเขียนจำนวนสมาชิกของเซตเป็นสิ่งที่ไม่ค่อยสำคัญ แต่ที่จริงแล้วสิ่งเหล่านี้เป็นพื้นฐานที่จะช่วยให้เข้าใจในเรื่องเซตมากขึ้น และผู้เรียนขาดความสามารถในการประยุกต์ใช้ความรู้ อีกทั้งไม่สามารถค้นหาหรือสรุปความสัมพันธ์ของเงื่อนไขที่โจทย์ปัญหากำหนดให้ ซึ่งหากครูผู้สอนฝึกให้ผู้เรียนเชื่อมโยงความรู้กับการประยุกต์ใช้ความรู้กับโจทย์ปัญหาด้วยการใช้แผนภาพเวนน์ – ออยเลอร์ จะทำให้ผู้เรียนสามารถนำความรู้ทั้งหมดมาใช้ร่วมกันได้ รวมทั้งหากฝึกให้ผู้เรียนค้นหา หรือสรุปความสัมพันธ์ของเงื่อนไขที่หลากหลายก็จะสามารถนำแก้โจทย์คำตอบได้ง่ายขึ้น

สรุปและอภิปรายผล

การศึกษาในครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาคลังข้อสอบกลุ่มสาระการเรียนรู้คณิตศาสตร์ สำหรับระดับชั้นมัธยมศึกษาปีที่ 4 และตรวจสอบคุณภาพของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) จากผลการศึกษาค้นคว้าคณะผู้วิจัยสรุปผลและอภิปรายผลตามลำดับดังนี้

1. จากการจัดทำตารางการกำหนดคุณลักษณะเฉพาะของแบบทดสอบ (Test Specification) ตามแนวคิดของบลูมฉบับปรับปรุง (Revision of Bloom's Taxonomy) ร่วมกับผู้เชี่ยวชาญ และคณาจารย์กลุ่มสาระการเรียนรู้คณิตศาสตร์โรงเรียนสาริตถมหาวิทยาลัยศรีนครินทร์วิโรฒ ฝ่ายมัธยม เพื่อใช้ในการกำหนดสัดส่วนของจำนวนข้อสอบในแต่ละเนื้อหาย่อย ผลจากการประชุมกลุ่มย่อยให้ข้อสรุปว่าพฤติกรรมการเรียนรู้ที่ต้องการให้เกิดขึ้นกับผู้เรียนของเนื้อหา

เรื่องเซต คือ มิตินี้ที่ 1 กระบวนการทางพุทธิปัญญาสามารถเกิดขึ้นตั้งแต่ระดับความจำถึงการวิเคราะห์เท่านั้น ส่วนมิตินี้ที่ 2 ความรู้สามารถเกิดขึ้นตั้งแต่ระดับความรู้เชิงข้อเท็จจริงถึงความรู้เชิงวิพากษ์วิจารณ์ จากผลดังกล่าวจะเห็นได้ว่าการใช้แนวคิดของบลูมฉบับปรับปรุง ทำให้เกิดข้อจำกัดเกี่ยวกับจำนวนข้อสอบที่วัดความรู้เชิงอภิमानและกระบวนการทางพุทธิปัญญาด้านการประเมินค่าและการสร้างสรรค์ ทั้งนี้อาจเนื่องมาจากธรรมชาติของวิชาคณิตศาสตร์เน้นกระบวนการประยุกต์ใช้ในการแก้ปัญหา และจากผลการวิจัยเรื่องการจัดทำยุทธศาสตร์การปฏิรูปการศึกษาขั้นพื้นฐานให้เกิดความรับผิดชอบของสมเกียรติ ตั้งกิจวานิชย์ และคณะ [13] กล่าวว่ามาตรฐานของวิชาคณิตศาสตร์ถือว่าสอดคล้องกับแนวคิดทักษะแห่งศตวรรษที่ 21 เพราะให้ความสำคัญกับการแก้ปัญหาและการประยุกต์ใช้ความรู้ในสถานการณ์จริง ฉะนั้นหากใช้แนวคิดการวัดความสามารถทางสติปัญญาในการเรียนรู้คณิตศาสตร์ตามแนวคิดของ Wilson [14] จะเป็นแนวคิดที่เหมาะสมกับธรรมชาติของวิชาคณิตศาสตร์มากกว่า โดยความสามารถทางสติปัญญาในการเรียนรู้คณิตศาสตร์ แบ่งเป็น 4 ระดับ คือ 1) ความรู้ความจำด้านการคิดคำนวณ เป็นความรู้ความจำเกี่ยวกับข้อเท็จจริง ศัพท์ และความสามารถในการใช้กระบวนการคิดคำนวณตามลำดับขั้นที่ได้เรียนรู้ 2) ความเข้าใจ เป็นความเข้าใจในโมโนทัศน์ หลักการ กฎทางคณิตศาสตร์ การสรุปเป็นกรณีทั่วไป มีความเข้าใจในโครงสร้างทางคณิตศาสตร์พิจารณาจากมีความสามารถในการแปลงรูปโครงสร้างและสามารถอ่านและตีความโจทย์ปัญหาได้ 3) การนำไปใช้ เป็นความสามารถในการแก้โจทย์ปัญหาต่างๆ ที่คล้ายกับสิ่งที่ได้เรียนรู่มาก่อน สามารถเปรียบเทียบข้อมูลและค้นพบโครงสร้างที่คล้ายกันได้ และ 4) การวิเคราะห์ เป็นความสามารถในการแก้โจทย์ปัญหาใหม่ๆ ซึ่งเป็นโจทย์ที่ซับซ้อนกว่าที่เคยเรียนแต่ยังอยู่ในขอบเขตของเนื้อหาสาระที่เรียน สามารถวิเคราะห์ความ

สัมพันธ์จากสิ่งที่โจทย์กำหนดให้เพื่อนำมาใช้ในการแก้โจทย์ปัญหา

2. ค่าพารามิเตอร์ความยากของข้อสอบพบว่าข้อสอบส่วนใหญ่มีค่าพารามิเตอร์ความยากระดับ (-3.000) ถึง 3.000 รองลงมา คือ 3.001 ถึง 4.530 ส่วนอำนาจจำแนกของข้อสอบส่วนใหญ่อยู่ในระดับต่ำกว่า 0.50 รองลงมา คือ 0.50 ถึง 2.50 สำหรับฟังก์ชันสารสนเทศของแบบทดสอบที่ได้จากการทดลองสุ่มข้อสอบตามตารางการกำหนดคุณลักษณะของแบบทดสอบ จำนวน 10 ฉบับ พบว่ามีฟังก์ชันสารสนเทศของแบบทดสอบอยู่ระหว่าง 1.190 ถึง 2.379 ความคลาดเคลื่อนมาตรฐานของการประมาณค่าอยู่ระหว่าง 0.195 ถึง 0.483 และแบบทดสอบทั้ง 10 ฉบับมีความน่าเชื่อถือของการประมาณค่าความสามารถของผู้สอบ (Empirical Reliability) อยู่ระหว่าง 0.721 ถึง 0.815 จากผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบและแบบทดสอบแสดงให้เห็นว่าเป็นค่าพารามิเตอร์ของข้อสอบที่สามารถนำไปใช้ได้ในทางปฏิบัติ และสอดคล้องกับทฤษฎีการตอบสนองข้อสอบ

3. ในการวิจัยครั้งนี้พบว่าข้อสอบจำนวน 55 ข้อที่มีค่าพารามิเตอร์ความยาก (b) ตั้งแต่ 3.001 ถึง 4.530 ซึ่งในทางปฏิบัติถือว่าเป็นข้อสอบที่ยากมาก [9-10] โดยมักเป็นข้อสอบที่วัดความรู้วิธีดำเนินการและกระบวนการทางพุทธิปัญญาด้านความเข้าใจ และการประยุกต์ใช้จากการพิจารณาถึงลักษณะโจทย์ดังกล่าวที่จัดว่าเป็นข้อสอบที่ค่อนข้างยากนั้นเพราะกระบวนการทางพุทธิปัญญาที่ผู้เรียนต้องใช้ในการแก้โจทย์ปัญหาคือ ผู้เรียนต้องเกิดเข้าใจเกี่ยวกับเซต การดำเนินการของเซต และการแก้โจทย์ปัญหาโดยใช้แผนภาพเวนน์ – ออยเลอร์ และสามารถนำความรู้ ความเข้าใจที่ได้เรียนรู้มาประยุกต์ใช้ในการแก้ปัญหาที่ซับซ้อน ดังนั้นการที่ผู้เรียนจะแก้โจทย์ปัญหาต้องมีกระบวนการคิดเป็นลำดับขั้น คือ 1) การอ่านเพื่อทำความเข้าใจกับสิ่งที่โจทย์ถาม และสิ่งที่โจทย์กำหนดให้

2) การจำแนกปัญหาเป็นข้อย่อย เพื่อค้นหาและใช้วิธีการเพื่อแก้โจทย์ ทั้งนี้จากผลการวิจัยครั้งนี้ระบุว่า มีข้อสอบจำนวนหนึ่งเป็นข้อสอบที่ค่อนข้างยากนั้น อาจเนื่องมาจากผู้เรียนมีปัญหาเกี่ยวกับความเข้าใจในโจทย์ปัญหาดังปรากฏตามรายงานผลการประเมินโครงการประเมินผลนักเรียนร่วมกับนานาชาติ (Programme for International Student Assessment: PISA) PISA 2012 คณิตศาสตร์ การอ่าน และวิทยาศาสตร์ ของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี [15] พบว่าแนวโน้มการรู้เรื่องคณิตศาสตร์ของนักเรียนไทยระหว่าง PISA 2009 - 2012 ปรากฏว่ากลุ่มที่มีคะแนนลดลงเป็นนักเรียนกลุ่มที่เคยมีคะแนนสูง นั่นคือ นักเรียนจากเขตกรุงเทพฯและปริมณฑลกับนักเรียนจากภาคเหนือตอนบน สิ่งที่ระบบการศึกษาต้องกังวล คือ นักเรียนกลุ่มสูงกลับมีคะแนนลดลง โดยนักเรียนจากเขตกรุงเทพฯและปริมณฑลมีคะแนนลดลงมาก กลุ่มที่มีคะแนนลดลงมากที่สุดเมื่อเปรียบเทียบกับ PISA 2003 เป็นนักเรียนจากภาคกลางซึ่งลดลงมาต่ำกว่านักเรียนจากภาคอีสานตอนล่างเล็กน้อย ส่วนการรู้เรื่องคณิตศาสตร์ของนักเรียนไทยหมวดกระบวนการทางคณิตศาสตร์ ค่าเฉลี่ยนักเรียนไทยมีจุดอ่อนที่สุดในกระบวนการการคิดวิธีการ หรือการคิดให้เป็นคณิตศาสตร์ นั่นคือ การคิดถึงปัญหาตามสถานการณ์ในบริบทให้เป็นวิธีการทางคณิตศาสตร์ แต่นักเรียนจากกลุ่มโรงเรียนจุฬาราชวิทยาลัย และโรงเรียนสาธิต มีจุดอ่อนด้านการตีความ ซึ่งส่วนมากเป็นการตีความ แปลความผลลัพธ์ทางคณิตศาสตร์ มาเป็นผลการแก้ปัญหาของโลกในชีวิตประจำวัน ซึ่งโรงเรียนทั้งสองกลุ่มมีจุดอ่อนเหมือนกัน

ข้อเสนอแนะ

1. ข้อเสนอแนะที่ได้จากการวิจัย

1.1 จากการทดลองสุ่มข้อสอบเพื่อจัดชุดเป็นแบบทดสอบของแบบทดสอบฉบับที่ 3 และ 5 เป็นแบบทดสอบที่เหมาะสมกับผู้สอบที่มี

ความสามารถ (Θ) ระหว่าง (-2.50) ถึง (-0.50) และ (-1.50) ถึง 0.50 ตามลำดับ กล่าวคือ เป็นแบบทดสอบที่เหมาะสมสำหรับการทดสอบเพื่อพัฒนาการเรียนรู้ (Formative Test) ซึ่งเป็นการวัดและประเมินความรู้และทักษะของนักเรียน เพื่อทราบถึงพัฒนาการของผู้เรียนและวินิจฉัยจุดบกพร่องของผู้เรียน โดยสามารถนำผลการทดสอบมาออกแบบการจัดการเรียนรู้ให้สอดคล้องกับระดับความสามารถของผู้เรียน เช่น การทดสอบท้ายบทเรียน หรือการทดสอบกลางภาค

1.2 จากการทดลองสุ่มข้อสอบเพื่อจัดชุดเป็นแบบทดสอบของแบบทดสอบฉบับที่ 2, 6 และ 9 เป็นแบบทดสอบที่เหมาะสมกับผู้สอบที่มีความสามารถ (Θ) ระหว่าง (-0.50) ถึง 1.50, 0.00 ถึง 1.50 และ 0.50 ถึง 1.50 ตามลำดับ กล่าวคือเป็นแบบทดสอบที่เหมาะสมสำหรับการทดสอบเพื่อตัดสิน (Summative Testing) ซึ่งเป็นการทดสอบสำหรับวัดความรู้หรือทักษะของผู้เรียนจากการเรียนรู้ในช่วงที่ผ่านมา ณ ช่วงเวลาหนึ่ง และนำผลการทดสอบมาตัดสิน เช่น การทดสอบปลายภาค เป็นต้น

1.3 สำหรับข้อสอบที่ค่อนข้างยาก คือมีค่าพารามิเตอร์ความยาก (b) ตั้งแต่ 3.001-4.530 จำนวน 55 ข้อ นั้น ครูผู้สอนสามารถนำไปเพิ่มเติมลงในแบบทดสอบที่มีคุณลักษณะ เช่นแบบทดสอบฉบับที่ 6 และ 9 ซึ่งเหมาะสำหรับการทดสอบกับผู้สอบที่มีความสามารถ (Θ) ระหว่าง 0.00 ถึง 1.50 เพื่อใช้เป็นส่วนหนึ่งในการนำผลการสอบเพื่อตัดสินในการทดสอบปลายภาค

1.4 จากผลการทดสอบชี้ให้เห็นว่าผู้เรียนยังขาดทักษะการอ่านเพื่อให้เกิดความเข้าใจในโจทย์ปัญหา เพื่อระบุว่าสิ่งที่โจทย์กำหนดให้ และสิ่งที่โจทย์ต้องการให้หาคำตอบ ดังนั้นครูผู้สอนควรให้ความสำคัญกับการฝึกผู้เรียนให้คุ้นเคยกับปัญหาคณิตศาสตร์ที่มีลักษณะย้อน และมีความรวบรัดในการอ่านเพื่อให้ผู้เรียนเกิดความเข้าใจผู้เรียนจำเป็นต้องใช้สมาธิ ใช้ความพยายามในการเก็บรายละเอียด ทักษะการอ่านสามารถกระทำได้

ในช่วงโมเมนต์คณิตศาสตร์ โดยเฉพาะเมื่อถึงตัวอย่างหรือแบบฝึกหัดเกี่ยวกับการแก้โจทย์ปัญหา โดยอาจฝึกเป็นรายบุคคล หรือฝึกเป็นกลุ่มด้วยการอภิปรายร่วมกันถึงข้อความ หรือสถานการณ์ที่กำหนดให้

1.5 จากผลการทดสอบเนื้อหาย่อย เรื่องการดำเนินการของเซต และการแก้โจทย์ปัญหาโดยใช้แผนภาพเวเนน-ออยเลอร์ ชี้ให้เห็นว่าผู้เรียนยังขาดการประยุกต์ใช้ความรู้ ความเข้าใจจากการเรียนที่ผ่านมา ดังนั้นครูผู้สอนควรฝึกผู้เรียนให้คุ้นเคยกับเทคนิคที่สามารถใช้วิธีช่วยเพิ่มพูนความเข้าใจ เช่น การเขียนภาพหรือสร้างแบบจำลอง เพื่อแสดงความสัมพันธ์ของข้อมูลต่างๆ ของปัญหาจะทำให้ปัญหามีความเป็นรูปธรรมขึ้น ทำความเข้าใจได้ง่ายขึ้น การปรับขนาดของปริมาณต่างๆ ที่กำหนดในตัวปัญหา เช่น ลดปริมาณลง เมื่อมีปริมาณน้อยๆ จะช่วยให้โครงสร้างของปัญหามีความชัดเจนขึ้น การลดขนาดของปริมาณนี้ ต้องกระทำในแนวทางที่ถูกต้องมีความเป็นไปได้และสมเหตุสมผล เพราะมีฉะนั้นแล้วแทนที่จะช่วยให้เข้าใจปัญหาอาจทำให้ปัญหามีความยุ่งยากเพิ่มขึ้นก็ได้

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 การจัดทำตารางการกำหนดคุณลักษณะเฉพาะของแบบทดสอบ (Test Specification) ควรใช้การวัดความสามารถทางสติปัญญาในการเรียนรู้คณิตศาสตร์ตามแนวคิดของ Wilson เพื่อให้สอดคล้องกับธรรมชาติวิชาคณิตศาสตร์ และระดับการเกิดความรู้และกระบวนการทางพุทธิปัญญา

2.2 ควรนำคลังข้อสอบไปพัฒนาเป็นการทดสอบแบบปรับเหมาะ (Computerized Adaptive Testing) ซึ่งเป็นการทดสอบที่ผู้เรียนถูกทดสอบด้วยข้อสอบที่มีค่าความยากง่ายเหมาะสมกับความสามารถของผู้สอบ และสามารถป้องกันการทุจริตในการสอบเป็นอย่างดี

2.3 การสร้างคลังข้อสอบควรเป็นข้อสอบที่สอดคล้องกับทักษะแห่งศตวรรษที่ 21 คือ

ข้อสอบควรมีลักษณะที่มุ่งเน้น 1) การวัดความเข้าใจ เพื่อให้ผู้เรียนสามารถเชื่อมโยงระหว่างแนวคิดกับปัญหาในเชิงรูปธรรม 2) การวัดความสามารถประยุกต์ใช้แนวคิดและความรู้ในสถานการณ์จริงที่มีความซับซ้อน เช่น การใช้จ่ายการเดินทาง การทำอาหาร การจัดการการเงินของตน การประเมินสถานการณ์ การตัดสินใจประเด็นปัญหา

ทางสังคมการเมือง เป็นต้น ซึ่งความรู้คณิตศาสตร์สามารถเข้ามาช่วยทำให้การมองประเด็นหรือสถานการณ์ชัดเจนขึ้น

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนจากคณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

เอกสารอ้างอิง

- [1] กระทรวงศึกษาธิการ. (2554). *แนวปฏิบัติการวัดและประเมินผลการเรียนรู้ ตามหลักสูตรแกนกลาง การศึกษาขั้นพื้นฐาน พุทธศักราช 2551*. กรุงเทพฯ: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทยจำกัด.
- [2] ดำริ บุญชู. (2546, มกราคม-มีนาคม). การใช้ประโยชน์จากแหล่งเรียนรู้ในสถานศึกษา. *วารสารวิชาการ*. 8(1): 1.
- [3] สมวงษ์ แปลงประสพโชค; สมเดช บุญประจักษ์; และ จรรยา ภูอุดม. (2549). *นวัตกรรมเพื่อพัฒนาคุณภาพการจัดการศึกษาคณิตศาสตร์ขั้นพื้นฐานของเด็กไทย : การศึกษาสาเหตุเด็กไทยอ่อนคณิตศาสตร์และแนวทางแก้ไข*. มหาวิทยาลัยราชภัฏพระนคร. สืบค้นเมื่อมิถุนายน 2556, จาก http://www.ripon-math.com/doc/25510502/child_low_math.doc
- [4] Grolund, N.E. (1998). *Assessment of Student Achievement. Sixth Edition*. Needham Heights, MA: Allyn and Bacon.
- [5] อุทัย บุญประเสริฐ. (2540). *หลักและแนวทางในการสร้างคลังข้อทดสอบสำหรับโรงเรียน*. กรุงเทพฯ: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- [6] Baker, F. B. (1992). *Item Response Theory*. New York, NY: Marcel Dekker.
- [7] Slocum, S.L.. (2005). *Assessing unidimensionality of psychological scales using individual and integrative criteria from factor analysis*. Dissertation, Ph.D. (Measurement, Evaluation and Research Methodology). Graduate school The University of British Columbia.
- [8] เกียรติศักดิ์ ส่องแสง. (2547). *ฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์. ปรินซิพการศึกษาดุษฎีบัณฑิต (การทดสอบและวัดผลการศึกษา)* กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
- [9] De Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*. New York: The Guilford Press.
- [10] ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่*. พิมพ์ครั้งที่ 4. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- [11] Hambleton, Ronald K. and Swaminathan, H.. (1991). *Fundamentals of item response theory*. California: Sage Publications Inc.
- [12] Toilt, Mathilda, D. (2003). *IRT from SSI: BILOG-MG, MULTILOG, PARSCALE, TESTFACT*. Illinois: Scientific software International.

- [13] สมเกียรติ ตั้งกิจวานิชย์; และคณะ. (2556). รายงานฉบับสมบูรณ์การจัดทำยุทธศาสตร์การปฏิรูปการศึกษาขั้นพื้นฐานให้เกิดความรับผิดชอบ. ม.ป.พ.
- [14] Wilson, J.W. (1971). *Evaluation of Learning in Secondary School Mathematics in Handbook on Formative and Summative Evaluation of Student Learning*. Edited by Benjamin S. Bloom. New York: McGraw-Hill
- [15] สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยีกระทรวงศึกษาธิการ. (2556). *ผลการประเมิน PISA 2012 คณิตศาสตร์ การอ่าน และวิทยาศาสตร์*. กรุงเทพฯ: แอดวานซ์ ฟรินติ้งเซอร์วิส.