

การเปรียบเทียบความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์วิชา  
คณิตศาสตร์รูปแบบผสมที่มีสัดส่วนของการตรวจให้คะแนนในแบบสอบแตกต่างกัน:  
การประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนความรู้บางส่วน  
และ โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนความรู้บางส่วนแบบทั่วไป  
Comparison of standard error of mixed-format mathematics  
achievement test among different proportion of mixed-format scoring:  
An application of partial credit model and generalized partial credit  
model

นรินทร์ บุญธรรมพาณิชย์<sup>1</sup>, โชติกา ภาษีผล<sup>2</sup>

Narintorn Bunthumpanich<sup>1</sup> Shotiga Pasiphol<sup>2</sup>

<sup>1</sup>นิสิตระดับปริญญาโท สาขาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

Faculty of Education, Chulalongkorn University

Corresponding author: E-mail: spoipoy@hotmail.com

<sup>2</sup>รองศาสตราจารย์ ดร. คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

Lecturer, Ph.D. Associate Professor Faculty of Education, Chulalongkorn University

E-mail: aimorn.j@chula.ac.th

## บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์รูปแบบผสมที่มีสัดส่วนของการตรวจให้คะแนนในแบบสอบแตกต่างกัน 2) เปรียบเทียบความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์รูปแบบผสมที่มีสัดส่วนของการตรวจให้คะแนนในแบบสอบแตกต่างกัน โดยวิเคราะห์ปฏิสัมพันธ์เมื่อพิจารณาในภาพรวมของแบบสอบทั้ง 15 ฉบับ และเมื่อจำแนกตามสาระการเรียนรู้ทั้ง 5 สาระ โดยแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้นแต่ละสาระการเรียนรู้มีสัดส่วนของการตรวจให้คะแนนที่แตกต่างกัน 3 สัดส่วน คือ 60:40, 70:30 และ 80:20 ใช้โมเดลการวิเคราะห์แบบ 1pl และ 2pl ผลการวิจัยสรุปได้ดังนี้

1) แบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์รูปแบบผสมที่ผู้วิจัยพัฒนาขึ้น พบว่า ด้านความเที่ยงของแบบสอบส่วนใหญ่มีค่าตั้งแต่ 0.50 ขึ้นไป และทุกฉบับเหมาะกับผู้สอบหรือนักเรียนที่มีความสามารถปานกลาง

2) ในภาพรวมของแบบสอบทั้ง 15 ฉบับ พบว่าไม่มีปฏิสัมพันธ์กันระหว่างสัดส่วนการตรวจให้คะแนนและโมเดลการวิเคราะห์ ที่มีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐาน (SE) อย่างมีนัยสำคัญ

ทางสถิติที่ระดับ .05 เมื่อศึกษาอิทธิพลหลักพบว่าสัดส่วนของการตรวจให้คะแนน มีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐาน (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยผลการเปรียบเทียบรายคู่พบว่าค่าเฉลี่ย SE ระหว่างสัดส่วน 60:40 ต่ำกว่า 80:20 และ 70:30 ต่ำกว่า 80:20 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในส่วนโมเดลการวิเคราะห์ มีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐาน (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยพบว่าค่าเฉลี่ยของโมเดล 2pl ต่ำกว่าของโมเดล 1pl อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3) เมื่อจำแนกตามสาระการเรียนรู้ พบว่ามีเพียงสาระการวัด เท่านั้นที่มีปฏิสัมพันธ์กันระหว่างสัดส่วนการตรวจให้คะแนนและโมเดลการวิเคราะห์ ที่มีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐาน (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยพบว่าค่าเฉลี่ย SE ของ 60:40(1pl) ต่ำกว่า 80:20(1pl), 60:40(2pl) ต่ำกว่า 80:20(1pl), 70:30(1pl) ต่ำกว่า 80:20(1pl), 70:30(2pl) ต่ำกว่า 80:20(1pl) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เมื่อศึกษาอิทธิพลหลักพบว่ามีเพียงสาระจำนวนและการดำเนินการที่สัดส่วนของการตรวจให้คะแนน มีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐาน (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยผลการเปรียบเทียบรายคู่พบว่าค่าเฉลี่ย ระหว่างสัดส่วน 60:40 ต่ำกว่า 70:30, 60:40 ต่ำกว่า 80:20 และ 70:30 ต่ำกว่า 80:20 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 นอกจากนี้พบว่าสาระการเรียนรู้เรขาคณิตและสาระพีชคณิตมีค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐาน (SE) ที่ใช้โมเดล 2pl ต่ำกว่า ที่ใช้โมเดล 1pl อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

**คำสำคัญ :** แบบสอบรูปแบบผสม โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนความรู้บางส่วน โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนความรู้บางส่วนแบบทั่วไป

## Abstract

This research was purposed to ( 1) develop mixed-format mathematics achievement test with different proportion of scoring format, and (2) compare a standard error of mixed-format mathematics achievement test among different proportion of mixed-format scoring by analyzed a whole 15 sets of tests and each content independently. Five contents comprised of (1) number and operation, (2) measurement, (3) geometry, (4) algebra, and (5) data analysis and probability. Researcher developed 15 sets of mixed-format mathematics achievement test with 3 different proportion of scoring format (60:40, 70:30, and 80:20) applying 1pl and 2pl model. The finding reveled that;

1) Most of mixed-format mathematics achievement test were acceptable reliability (greater than 0.5). All 15 sets of tests were fitted for moderate level students or learners.

2) In perspective of 15 sets of tests, there was no interaction between proportion of scoring format and model influencing standard error at statistical significance level of .05. According to main effect study, proportion of scoring format affected mean of

standard error (SE) at statistical significance level of .05. Researcher also made paired comparison of means of SE for each proportion of scoring format as (1) 60:40 was lower than 80:20, (2) 70:30 was lower than 80:20 at statistical significance level of .05. Model also affected SE at statistical significance of .05 as mean of SE using 2pl was lower than using 1pl at statistical significance level of .05.

3) Content of measurement was the sole content that exhibited the interaction between proportion of scoring format and model influencing SE at statistical significance level of .05. Researcher found that (1) mean of SE of 60:40(1pl) was lower than 80:20(1pl), (2) mean of SE of 60:40(2pl) was lower than 80:20(1pl), (3) mean of SE of 70:30(1pl) was lower than 80:20(1pl), and (4) mean of SE of 70:30(2pl) was lower than 80:20(1pl) at statistical significance level of .05. Main effect study revealed that content of number and operation was the only group that proportion of scoring format affected mean of SE at statistical significance level of .05. Paired comparison revealed that (1) mean of SE of 60:40 was lower than 70:30, (2) mean of SE of 60:40 was lower than 80:20 and (3) mean of SE of 70:30 was lower than 80:20 at statistical significance level of .05. Additionally, content of geometry and content of algebra exhibited mean of SE using 2pl was lower than using 1pl at statistical significance level of .05.

**Keywords :** Mixed-format test, Partial credit model, Generalized partial credit model

## บทนำ

ในปัจจุบันการประเมินผู้เรียนมีความพยายามที่จะวัดทักษะความสามารถในการวิเคราะห์ การสังเคราะห์ ความคิดสร้างสรรค์ ของผู้เรียนมากกว่าการวัดแค่เพียงความรู้ ความจำ และความเข้าใจในเนื้อหาเท่านั้น แต่การเลือกใช้ข้อสอบชนิดใดชนิดหนึ่งจะมีข้อจำกัดในการวัดเพราะทั้งข้อสอบแบบหลายตัวเลือก (multiple choice: MC) และข้อสอบแบบเติมคำตอบ(constructed-response: CR) มีทั้งจุดแข็งและจุดอ่อน การใช้ข้อสอบแบบMC ก็ยังเป็นที่ยอมรับใช้ แต่ถูกวิพากษ์วิจารณ์ว่ามักจะมุ่งเน้นไปที่ความรู้ความจำ มากเกินไป Shepard (Shepard, 2002 อ้างถึงใน Liu, Lee, & Linn, 2001) ได้เสนอแนะว่าข้อสอบแบบCR มีความสามารถในการตรวจสอบความรู้ความเข้าใจได้จากการแสดงการให้เหตุผลของในการตอบคำถามของนักเรียน จึงถือว่าวัดได้ตรงกับสภาพจริงมากยิ่งขึ้น นอกจากนี้ Thorndike & Hagen, 1977; Kubiszyn & Borich, 2003) (อ้างถึงใน อัญชลี ศรีภักขานนท์, 2552) กล่าวว่า ข้อสอบแบบ CR สามารถวัดความสามารถของผู้เรียนในการจัดระบบ การบูรณาการ และการสังเคราะห์ความรู้ ซึ่งสอดคล้องกับ นิคม บุญหลาย (2562) กล่าวว่าแบบทดสอบที่มีข้อสอบอัตนัยหรือข้อสอบแบบCR เป็นเครื่องมือวัดผลอีกชนิดหนึ่งที่นิยมนำมาวัดและประเมินผลทั้งในอดีตและปัจจุบัน นอกจากจะเหมาะสมในการวัดพฤติกรรมการเรียนรู้ในระดับพื้นฐานแล้ว ยังสามารถประเมินทักษะในการเขียน การแก้ปัญหา ตลอดจนการคิดในระดับสูงอีกด้วย แต่ข้อสอบประเภทนี้ก็มีจุดอ่อนคือใช้เวลาและงบประมาณเป็นจำนวนมาก

มากในการจัดการทดสอบแต่ละครั้ง อีกทั้งยังนับว่าเป็นการประเมินศักยภาพของผู้เรียนโดยการลงมือปฏิบัติ ไม่มีเฉลยหรือคำตอบถูกที่แน่ชัดลงไปแบบข้อสอบแบบเลือกตอบ การตรวจหรือประเมินผลงานมีความจำเป็นต้องประเมินอย่างเป็นปรนัย ซึ่งนับว่ายากที่จะทำได้ (อุไร จักร์ตรีมงคล , 2557) จึงไม่อาจสร้างข้อสอบแบบ CR ให้มีจำนวนข้อมากๆ ได้ ปัจจุบันทดสอบจึงมีการประยุกต์ใช้แบบสอบทั้ง 2 รูปแบบควบคู่กัน เรียกว่า แบบสอบรูปแบบผสม (Mixed-format Tests) ซึ่งประกอบด้วยข้อสอบแบบ MC และข้อสอบแบบ CR ซึ่งน่าจะช่วยให้วัดความรู้ความสามารถของนักเรียนได้อย่างครอบคลุมมากยิ่งขึ้น (Kinsey, 2003; Walker, & McHale, 2007; Reshetar, & Melican, 2010 อ้างถึงใน รุ่งนภา แสนอานวยผล, 2555) ซึ่งเป็นการนำจุดเด่นของข้อสอบทั้งสองชนิดมาประยุกต์ใช้ ปัจจุบันมีการนำแบบทดสอบรูปแบบผสมมาใช้ในสถานการณ์การทดสอบทั้งในระดับประเทศ และระดับนานาชาติ เช่น การทดสอบภาษาอังกฤษ TOEFL การทดสอบ GMAT การทดสอบ GRE ฯลฯ (อัญชลี ศรีกลชาญ, 2552)

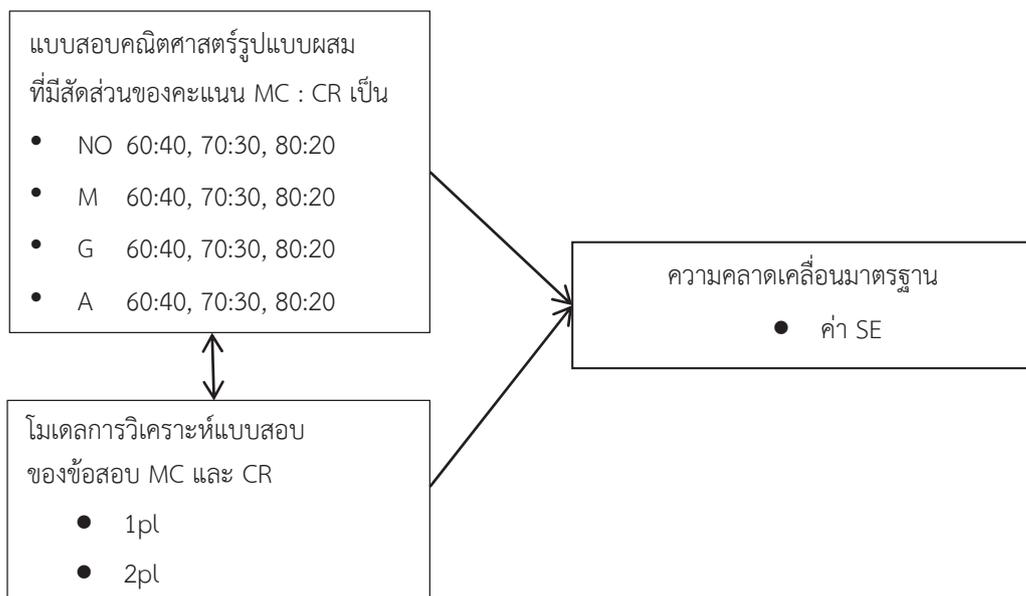
ในด้านของสัดส่วนของข้อสอบแบบ MC และ CR ยังไม่สามารถที่จะระบุได้ชัดเจนว่าสัดส่วนของข้อสอบในแบบสอบรูปแบบผสมที่ดีควรมีค่าสัดส่วนเป็นเท่าใด (Saen-amnuaiophon, 2012) Kolen and Lee (2011) ได้ทำการวิจัยเกี่ยวกับสัดส่วนของข้อสอบแบบ MC และ CR ซึ่งมีสัดส่วนของข้อสอบที่แตกต่างกันถึง 13 รูปแบบ โดยจำลองข้อมูลขึ้นพบว่า สัดส่วนของข้อสอบที่มีค่าความเที่ยงสูงที่สุดคือ 70:30 (0.915) รองลงมาคือ 60:40 (0.904) นอกจากนี้แล้ว Saen-amnuaiophon (2012) ทำการจำลองข้อมูลจำนวน 1,000 ข้อมูลเพื่อศึกษาสัดส่วน ที่แตกต่างกันทั้งสิ้น 3 รูปแบบคือ 20:80 , 50:50 และ 80:20 พบว่าสัดส่วน 20:80 ให้ค่า SE ต่ำว่าสัดส่วนอื่นๆ ถือเป็นสัดส่วนที่มีประสิทธิภาพดีกว่าสัดส่วนอื่นๆ เพราะข้อสอบแบบ MC ยังมีโอกาสในการเดาอยู่ เมื่อข้อสอบแบบ MC ลดลง โอกาสในการเดาก็ลดลง ค่า SE ก็ลดลงตามไปด้วย แต่ตามทัศนะของผู้วิจัยเองสถานการณ์ในการวัดจริงไม่สามารถที่จะสร้างข้อสอบแบบ CR ให้มีสัดส่วนสูงๆ ได้ เพราะด้วยข้อจำกัดเรื่องงบประมาณและเวลาในการทำข้อสอบแต่ละชุด ยังมีจำนวนข้อสอบแบบ CR ก็ยังใช้ระยะเวลาในการจัดสอบมากขึ้นตามไปด้วย จากการศึกษางานวิจัยรุ่งนภา แสนอานวยผล (2555) ได้ศึกษาประสิทธิภาพของแบบทดสอบรูปแบบผสม และศึกษาปฏิสัมพันธ์ระหว่างโมเดลการตรวจให้คะแนนแบบตรวจให้คะแนนแบบ 2 ค่าและมากกว่า 2 ค่า โดยแบ่งเป็น 1) โมเดลโลจิส 1 พารามิเตอร์ร่วมกับ PCM และ 2) โมเดลโลจิส 3 พารามิเตอร์ร่วมกับ GPCM กับสัดส่วนของข้อสอบพบว่าโมเดลการตรวจให้คะแนนและสัดส่วนคะแนนของแบบสอบมีปฏิสัมพันธ์กันส่งผลต่อค่า SE อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 งานวิจัยข้างต้นเป็นการจำลองข้อมูลขึ้นด้วยโปรแกรมคอมพิวเตอร์ แต่ตามสถานการณ์การวัดจากผู้เรียนจริงเรามีโอกาสคาดเดาคำตอบของคุณภาพของแบบสอบได้เลยว่าแบบสอบรูปแบบผสมที่มีสัดส่วนแตกต่างกันจะมีคุณภาพเป็นเช่นใด และจากการสังเคราะห์ข้อมูลผลการวิจัยของ ญัฐไฉไล พริ้งมาตี (2544), เวชฤทธิ์ อังกะภักทขจร (2546), ไช้มุก เลืองสุนทร (2552), สุวิมล เสวกสุริยวงศ์ (2553) และ เกวลิน ชัยณรงค์ (2554) ทำให้สรุปได้ว่า ช่วงชั้นที่ 3 คือระดับชั้นมัธยมที่ 1-3 เป็นช่วงชั้นที่นักเรียนค่อนข้างจะมีจุดบกพร่องและมีโมโนทัศน์ที่คลาดเคลื่อนมากที่สุด ผู้วิจัยจึงสนใจมิติของสัดส่วนของข้อสอบแบบ MC และ CR มาพัฒนาเป็นแบบสอบผลสัมฤทธิ์คณิตศาสตร์รูปแบบผสมระดับชั้นมัธยมศึกษาตอนต้นตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ.2551 เพื่อให้ครอบคลุมทุกเนื้อหาผู้วิจัยจึงแบ่งแบบสอบ

ออกเป็น 5 กลุ่มเนื้อหาสาระคือ 1) จำนวนและการดำเนินการ(NO) 2) การวัด(M) 3) เรขาคณิต(G) 4) พีชคณิต(A) และ 5) การวิเคราะห์ข้อมูลและความน่าจะเป็น (DP) ผู้วิจัยสร้างแบบสอบโดยแต่ละสาระการเรียนรู้มีแบบสอบที่มีสัดส่วนคะแนนของข้อสอบแบบ MC และ CR แตกต่างกัน 3 สัดส่วน คือ สัดส่วนคะแนน MC: CR 60:40 ,70:30 และ 80:20 นอกจากนี้ผู้วิจัยสนใจนำโมเดลในการตรวจให้คะแนนข้อสอบแบบ MC และ ข้อสอบแบบ CR มาเป็นตัวแปรอิสระตัวแปรหนึ่งที่ใช้ในการวิจัยโดยแบ่งออกเป็นทั้งหมด 2 โมเดล คือ (1) โมเดลโลจิส 1 พารามิเตอร์ร่วมกับโมเดล PCM และ (2) โมเดลโลจิส 2 พารามิเตอร์ร่วมกับโมเดล GPCM เพื่อให้การนำเสนอกระชับขึ้น ผู้วิจัยขอใช้ตัวย่อเป็น โมเดลการวิเคราะห์ 1pl และ 2pl

### วัตถุประสงค์ของการวิจัย

- 1) เพื่อพัฒนาแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์รูปแบบผสม ที่มีสัดส่วนของการตรวจให้คะแนนในแบบสอบแตกต่างกัน
- 2) เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์รูปแบบผสมที่มีสัดส่วนของการตรวจให้คะแนนในแบบสอบแตกต่างกัน เมื่อพิจารณาในภาพรวมของแบบสอบทั้ง 15 ฉบับ และ เมื่อจำแนกตามสาระการเรียนรู้ทั้ง 5 สาระ

### กรอบแนวคิดในการวิจัย



ภาพที่ 1 กรอบแนวคิดในการวิจัย

### วิธีดำเนินการวิจัย

ประชากรที่ใช้ในการวิจัยครั้งนี้เป็น นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 ในสำนักงานศึกษาธิการจังหวัดกรุงเทพมหานคร สังกัดสำนักงานการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ

ซึ่งตัวอย่างที่ใช้ในการวิจัยครั้งนี้ได้มาจากการสุ่มแบบหลายขั้นตอน (multistage random) นักเรียนตัวอย่างจำนวนทั้งสิ้น 3,157 คน โดยแบบสอบแต่ละฉบับ นักเรียนมีคะแนนเฉลี่ยอยู่ระหว่าง 16.40 ถึง 18.47 แสดงให้เห็นว่านักเรียนที่เป็นตัวอย่างมีความสามารถที่ใกล้เคียงกัน (รายละเอียดดังตารางที่ 1) เครื่องมือที่ใช้ในการครั้งนี้ผู้วิจัยสร้างขึ้นที่จำแนกสาระการเรียนรู้ออกเป็น 5 สาระ แต่ละสาระมีแบบสอบสาระละ 3 ฉบับ ตามสัดส่วนของคะแนน MC:CR รวมจำนวนแบบสอบทั้งสิ้น 15 ฉบับ ซึ่งภายหลังจากการดำเนินการวิเคราะห์เนื้อหา สร้างผังข้อสอบ เขียนผังข้อสอบ และทำการทดลองเครื่องมือเรียบร้อยแล้ว จากนั้นนำไปดำเนินการทดสอบกับนักเรียน จากนั้นดำเนินการตรวจให้คะแนนแบบสอบทั้งส่วนที่เป็นข้อสอบแบบ MC และข้อสอบแบบ CR

**ตารางที่ 1** นักเรียนตัวอย่างจำแนกตามสัดส่วนการตรวจให้คะแนน

สาระการเรียนรู้	สัดส่วนการตรวจให้คะแนน						รวม (คน)
	60:40	Mean	70:30	Mean	80:20	Mean	
1. สาระจำนวนและการดำเนินการ	213 (6.75%)	17.10	219 (6.94%)	17.83	217 (6.87%)	18.10	649 (20.56%)
2. สาระการวัด	220 (6.97%)	18.04	211 (6.68%)	18.47	208 (6.59%)	17.41	639 (20.24%)
3. สาระเรขาคณิต	201 (6.37%)	16.77	209 (6.62%)	16.40	211 (6.68%)	16.72	621 (19.67%)
4. สาระพีชคณิต	212 (6.72%)	17.14	216 (6.84%)	17.15	217 (6.87%)	16.98	645 (20.43%)
5. สาระการวิเคราะห์ข้อมูลและความน่าจะเป็น	200 (6.34%)	17.23	202 (6.40%)	16.79	201 (6.37%)	16.89	603 (19.10%)
<b>รวม</b>	<b>1,046 (33.13%)</b>	<b>17.26</b>	<b>1,057 (33.48%)</b>	<b>17.33</b>	<b>1,054 (33.39%)</b>	<b>17.22</b>	<b>3,157 (100%)</b>

### การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลแบ่งออกเป็น 2 ตอน ดังนี้

1. วิเคราะห์ข้อสอบในแบบสอบทั้ง 15 ฉบับ โดยรายงานคุณภาพของแบบสอบจำแนกตามแบบสอบทั้ง 15 ฉบับ โดยรายงาน ค่าความเที่ยง ค่าความยาก ค่าอำนาจจำแนก ค่าสารสนเทศของแบบสอบ
2. วิเคราะห์ปฏิสัมพันธ์ (Interaction Effect) ระหว่างตัวแปรสัดส่วนการให้คะแนนในแบบสอบและตัวแปรโมเดลการวิเคราะห์แบบสอบโดยใช้การวิเคราะห์ความแปรปรวนแบบ 2 ทาง (Two – way ANOVA) ด้วยโปรแกรมสำเร็จรูป SPSS ซึ่งทำการวิเคราะห์เมื่อพิจารณาในภาพรวมของแบบสอบทั้ง 15 ฉบับ และ เมื่อจำแนกตามสาระการเรียนรู้ทั้ง 5 สาระ

## ผลการวิจัย

1) แบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์รูปแบบผสมที่ผู้วิจัยพัฒนาขึ้น พบว่า ด้านความเที่ยงของแบบสอบส่วนใหญ่มีค่าตั้งแต่ 0.50 ขึ้นไป โดยแบบสอบที่พบว่ามีค่าความเที่ยงสูงที่สุดคือแบบสอบสาระจำนวนและการดำเนินการ ที่ใช้สัดส่วน 60:40 ร่วมกับโมเดลการวิเคราะห์แบบ 2pl (0.79) และทุกฉบับเหมาะกับผู้สอบหรือนักเรียนที่มีความสามารถปานกลาง (รายละเอียดดังตารางที่ 2 ซึ่งรายงานค่าคุณภาพของแบบสอบสาระการเรียนรู้จำนวนและการดำเนินการ เป็นสาระหนึ่งจากทั้งหมด 5 สาระ)

2) ในภาพรวมของแบบสอบทั้ง 15 ฉบับไม่พบปฏิสัมพันธ์ระหว่างสัดส่วนการตรวจให้คะแนนในแบบสอบและโมเดลการวิเคราะห์แบบสอบที่ส่งผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (ตารางที่ 3) แต่เมื่อทำการศึกษา อิทธิพลหลัก (main effect) พบว่า สัดส่วนของการตรวจให้คะแนนในแบบสอบมีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยผลการเปรียบเทียบรายคู่พบว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) ระหว่างสัดส่วน 60:40 ต่ำกว่า 80:20 และ 70:30 ต่ำกว่า 80:20 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (ตารางที่ 4) ส่วนตัวแปรโมเดลการวิเคราะห์แบบสอบมีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยพบว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) ที่ใช้โมเดล 2pl ต่ำกว่า ที่ใช้โมเดล 1pl อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (ตารางที่ 5)

**ตารางที่ 2** ผลการวิเคราะห์คุณภาพของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์สาระจำนวนและการดำเนินการจำแนกตามสัดส่วนการตรวจให้คะแนนและโมเดลการวิเคราะห์

proportion	model	reliability	Test Information ( $\theta_{max}, SE(\theta)$ )	MC		CR	
				$\bar{b}$	$\bar{a}$	$\bar{\delta}$	$\bar{\alpha}$
60:40	1pl	0.78	4.849 (-0.8, 0.45)	-0.85	-	-0.18	-
	2pl	0.79	5.261 (-0.6, 0.44)	-0.78	0.55	-0.15	1.59
70:30	1pl	0.75	4.413 (-1.0, 0.48)	-1.36	-	-0.03	-
	2pl	0.76	4.532 (-0.8, 0.47)	-1.34	0.45	-0.04	1.69
80:20	1pl	0.70	3.766 (-1.0, 0.52)	-1.46	-	-0.14	-
	2pl	0.70	3.752 (-1.0, 0.52)	-1.47	0.42	-0.13	1.63

\* หมายเหตุ  $\theta_{max}$  คือ ค่าความสามารถ ณ ตำแหน่ง Test Information สูงสุด

**ตารางที่ 3** ผลการวิเคราะห์ปฏิสัมพันธ์ของสัดส่วนของการตรวจให้คะแนนในแบบสอบและโมเดลการวิเคราะห์แบบสอบที่มีผลต่อความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ฯ (SE) ในภาพรวมของแบบสอบทั้งสิ้น 15 ฉบับ

แหล่งความแปรปรวน	Type III SS	Df	MS	F	Sig.
PROPORTION	0.259	2	0.129	7.533	0.001*
MODEL	0.263	1	0.263	15.297	0.000*
PROPORTION*MODEL	0.002	2	0.001	0.048	0.953
Error	15.867	924	0.017		
<b>รวม</b>	<b>442.744</b>	<b>930</b>			

**ตารางที่ 4** ผลการเปรียบเทียบความแตกต่างรายคู่ของค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ฯ (SE) เมื่อสัดส่วนของการตรวจให้คะแนนในแบบสอบแตกต่างกัน

สัดส่วนการตรวจให้คะแนน	Mean Difference	Std. Error	Sig.
60:40*80:20	-0.036	0.011	0.004*
70:30*80:20	-0.035	0.011	0.004*

3) เมื่อจำแนกตามสาระการเรียนรู้ทั้ง 5 สาระพบว่า สาระการเรียนรู้การวัด เพียงสาระเดียวที่มีปฏิสัมพันธ์กันระหว่างสัดส่วนการตรวจให้คะแนนในแบบสอบและโมเดลการวิเคราะห์แบบสอบที่ส่งผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (ตารางที่ 6) โดยผลการเปรียบเทียบค่าเฉลี่ยรายคู่พบว่าสัดส่วน 60:40\*1pl ต่ำกว่า 80:20\*1pl , 60:40\*2pl ต่ำกว่า 80:20\*1pl, 70:30\*1pl ต่ำกว่า 80:20\*1pl และ 70:30\*2pl ต่ำกว่า 80:20\*1pl อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (ตารางที่ 7) ขณะที่แบบสอบสาระจำนวนและการดำเนินการเรขาคณิต พีชคณิต การวิเคราะห์ข้อมูลและความน่าจะเป็น ไม่มีปฏิสัมพันธ์ระหว่างสัดส่วนการตรวจให้คะแนนในแบบสอบและโมเดลการวิเคราะห์แบบสอบต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

จากการศึกษาอิทธิพลหลัก (main effect) พบว่ามีเพียงสาระการเรียนรู้จำนวนและการดำเนินการ ที่สัดส่วนของการตรวจให้คะแนนในแบบสอบ มีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ฯ (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยพบว่าค่าเฉลี่ย ระหว่างสัดส่วน 60:40 ต่ำกว่า 70:30 สัดส่วน 60:40 ต่ำกว่า สัดส่วน 80:20 และสัดส่วน 70:30 ต่ำกว่า 80:20 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (ตารางที่ 8) ส่วนแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์สาระการเรียนรู้เรขาคณิต พีชคณิต การวิเคราะห์ข้อมูลและความน่าจะเป็น สัดส่วนของการตรวจให้คะแนนในแบบสอบ ไม่มีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) อย่างมีนัยสำคัญทางสถิติ

ที่ระดับ .05 ส่วนโมเดลการวิเคราะห์แบบสอป พบว่าสาระการเรียนรู้เรขาคณิตและสาระพีชคณิตมีค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอปผลสัมฤทธิ์ (SE) ที่ใช้โมเดล 2pl ต่ำกว่า ใช้โมเดล 1pl อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (ตารางที่ 9) ส่วนสาระการเรียนรู้ จำนวนและการดำเนินการ และการวิเคราะห์ข้อมูลและความน่าจะเป็น โมเดลการวิเคราะห์แบบสอป ไม่มีผลต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอปผลสัมฤทธิ์ (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

**ตารางที่ 5** ผลการเปรียบเทียบความแตกต่างของค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอปผลสัมฤทธิ์ (SE) เมื่อโมเดลการวิเคราะห์แบบสอปแตกต่างกัน

Model	Mean Difference	Std. Error	Levene's Test		t-test	
			F	Sig.	t	Sig.
1pl*2pl	0.037	0.009	53.166	0.000*	3.888	0.000*

**ตารางที่ 6** ผลการวิเคราะห์ปฏิสัมพันธ์ของสัดส่วนของการตรวจให้คะแนนในแบบสอปและโมเดลการวิเคราะห์แบบสอปที่มีผลต่อความคลาดเคลื่อนมาตรฐานของแบบสอปผลสัมฤทธิ์ (SE) สาระการวัด

สาระการเรียนรู้	แหล่งความแปรปรวน	Type III SS	Df	MS	F	Sig.
การวัด	PROPORTION	0.694	2	0.347	27.899	0.000*
	MODEL	0.062	1	0.062	4.999	0.027*
	PROPORTION*MODEL	0.105	2	0.052	4.216	0.016*
	Error	2.239	180	0.012		
	รวม		76.157	186		

**ตารางที่ 7** ผลการวิเคราะห์ค่าสถิติพื้นฐานของค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอปผลสัมฤทธิ์ (SE) สาระการเรียนรู้การวัดที่มีสัดส่วนของการตรวจให้คะแนนในแบบสอปและโมเดลการวิเคราะห์แตกต่างกัน

Proportion*Model	Mean Difference	Std. Error	Sig.
60:40(1pl)* 80:20(1pl)	-0.181	0.014	0.000*
60:40(2pl)* 80:20(1pl)	-0.185	0.010	0.000*
70:30(1pl)* 80:20(1pl)	-0.179	0.013	0.000*
70:30(2pl)* 80:20(1pl)	-0.180	0.011	0.000*

**ตารางที่ 8** ผลการเปรียบเทียบความแตกต่างรายคู่ของค่าเฉลี่ยของคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) เมื่อสัดส่วนของการตรวจให้คะแนนในแบบสอบแตกต่างกันสาระจำนวนและการดำเนินการ

สาระการเรียนรู้	สัดส่วน	Mean Difference	Std. Error	Sig.
จำนวนและการดำเนินการ	60:40*70:30	-0.032	0.009	0.003*
	60:40*80:20	-0.083	0.009	0.000*
	70:30*80:20	-0.051	0.010	0.000*

**ตารางที่ 9** ผลการเปรียบเทียบความแตกต่างของค่าเฉลี่ยของคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) เมื่อโมเดลการวิเคราะห์แบบสอบแตกต่างกัน จำแนกตามสาระการเรียนรู้

สาระการเรียนรู้	Mean Difference	Std. Error	Levene's Test		t-test	
			F	Sig.	t	Sig.
เรขาคณิต	0.047	0.012	75.782	0.000*	3.844	0.000*
พีชคณิต	0.052	0.015	41.239	0.000*	3.583	0.001*

## อภิปรายผล

1. จากผลการวิจัยพบว่าในสาระจำนวนและการดำเนินการแบบสอบที่มีสัดส่วน 60:40 ที่ใช้โมเดล 2pl มีแนวโน้มที่จะให้ค่าความเที่ยงสูงแบบสอบอื่นๆ ทั้งนี้เนื่องจากมีสาระสำคัญคือ สมบัติของจำนวนนับ ระบบจำนวนเต็ม เลขยกกำลัง เศษส่วน ทศนิยม การประมาณค่า อัตราส่วน สัดส่วน และร้อยละ ความรู้เบื้องต้นเกี่ยวกับจำนวนจริง ซึ่งจะเห็นได้ว่าสาระสำคัญเกี่ยวข้องกับเรื่องจำนวน อนุมานได้ว่าปัจจัยหนึ่งที่ส่งผลต่อความเที่ยงคือเนื้อหาที่มีความเป็นเอกพันธ์ ทำให้มีค่าสัมประสิทธิ์ความเที่ยงสูง (ศิริชัย กาญจนวาสิ, 2552) แต่นอกเหนือไปจากนี้แล้วยังมีปัจจัยในด้านอื่นๆอีกที่อาจส่งผลต่อความเที่ยง อาทิ เช่น ความเป็นปรนัยของข้อคำถาม ลักษณะและขนาดของตัวอย่าง จำนวนข้อคำถามและความยากง่ายของข้อคำถาม นอกจากนี้แล้วจะเห็นได้ว่าสัดส่วนการตรวจให้คะแนนในแบบสอบก็เป็นปัจจัยหนึ่งที่ส่งผลต่อคุณภาพของแบบสอบด้านความเที่ยงทั้งนี้ ในผลการวิจัยที่มีสัดส่วนของข้อสอบแบบ CR เพิ่มขึ้น ค่าความเที่ยงก็มีแนวโน้มจะเพิ่มขึ้นด้วย สอดคล้องกับงานวิจัยของ Michael และคณะ (2011)

2. จากผลการวิจัย ในภาพรวมนั้นไม่พบปฏิสัมพันธ์ระหว่างสัดส่วนการตรวจให้คะแนนในแบบสอบและโมเดลการวิเคราะห์แบบสอบ แต่เมื่อทำการศึกษาอิทธิพลหลัก พบว่าตัวแปรสัดส่วนของการตรวจให้คะแนนในแบบสอบ 60:40 และ 70:30 มีแนวโน้มที่จะให้ค่าความคลาดเคลื่อนมาตรฐานของแบบสอบ (SE) ต่ำกว่าสัดส่วน 80:20 สะท้อนให้เห็นว่าสัดส่วนของการตรวจให้คะแนนในแบบสอบ มีความสัมพันธ์กับค่าความคลาดเคลื่อนมาตรฐานของแบบสอบ (SE) โดยแบบสอบฉบับที่มีสัดส่วนของข้อสอบแบบเติมคำตอบ (CR) เพิ่มขึ้น จะมีค่าความคลาดเคลื่อนมาตรฐานของแบบสอบ (SE) ต่ำลง สอดคล้องกับ Saen-amnuaiophon (2012) ที่พบว่าสัดส่วนคะแนนของข้อสอบแบบ MC และ CR ที่

แตกต่างกันทั้งสิ้น 3 สัดส่วน คือ 20:80 , 50:50 และ 80:20 พบว่าสัดส่วน 20:80 ให้ค่า SE ต่ำว่าสัดส่วนอื่นๆ ถือเป็นสัดส่วนที่มีประสิทธิภาพดีกว่าสัดส่วนอื่นๆ นั่นก็เพราะข้อสอบแบบ MC ยังมีโอกาสในการเดาอยู่เมื่อข้อสอบแบบ MC ลดลง โอกาสในการเดาก็ลดลง ค่า SE ก็ต่ำลงตามไปด้วย ในด้านตัวแปรโมเดลการวิเคราะห์แบบสอบ 2pl มีแนวโน้มที่จะให้ค่าความคลาดเคลื่อนมาตรฐานของแบบสอบ (SE) ต่ำกว่าโมเดลการวิเคราะห์แบบสอบ 1pl ขณะที่งานวิจัยที่ผ่านมาส่วนใหญ่ไม่ได้มีการศึกษาอิทธิพลหลักของโมเดลการวิเคราะห์ต่อค่าความคลาดเคลื่อนมาตรฐานของแบบสอบ (SE) เนื่องจากพบปฏิสัมพันธ์ระหว่างโมเดลการวิเคราะห์ และสัดส่วนของการตรวจให้คะแนนสองค่าและความยาวของแบบทดสอบ ที่ส่งผลต่อค่า SE อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 นอกจากนี้แล้วยังพบปฏิสัมพันธ์รายคู่ พบว่า มีปฏิสัมพันธ์ระหว่างโมเดลการวิเคราะห์กับสัดส่วนของการตรวจให้คะแนน ระหว่างโมเดลการวิเคราะห์กับความยาวของแบบทดสอบ และระหว่างสัดส่วนของการตรวจให้คะแนนกับความยาวของแบบทดสอบ ดังเช่นงานวิจัยของรุ่งนภา แสนอำนวยผล (2555) โดยผลการวิจัยครั้งนี้ได้ข้อมูลเชิงประจักษ์ว่าโมเดลการวิเคราะห์แบบสอบ 2pl มีความคลาดเคลื่อนมาตรฐานต่ำกว่า โมเดลการวิเคราะห์แบบสอบ 1pl

3. เมื่อจำแนกตามสาระการเรียนรู้ พบว่ามีเพียงสาระการเรียนรู้การวัด ที่มีปฏิสัมพันธ์ระหว่างสัดส่วนการตรวจให้คะแนนในแบบสอบและโมเดลการวิเคราะห์แบบสอบต่อค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของแบบสอบผลสัมฤทธิ์ (SE) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยผลการเปรียบเทียบค่าเฉลี่ยรายคู่พบว่าสัดส่วน 60:40\*1pl ต่ำกว่า 80:20\* 1pl สัดส่วน 60:40\*2pl ต่ำกว่า 80:20\*1pl สัดส่วน 70:30\*1pl ต่ำกว่า 80:20\*1pl และสัดส่วน 70:30\*2pl ต่ำกว่า 80:20\*1pl อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับ รุ่งนภา แสนอำนวยผล (2555) ที่พบว่า มีปฏิสัมพันธ์ระหว่างโมเดลการตรวจให้คะแนน สัดส่วนของข้อสอบที่ตรวจให้คะแนนสองค่าและมากกว่าสองค่า ที่ส่งผลต่อค่า SE อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในขณะที่สาระการเรียนรู้อื่นกลับไม่พบปฏิสัมพันธ์ระหว่างสัดส่วนของการตรวจให้คะแนนในแบบสอบ และโมเดลการวิเคราะห์แบบสอบ สะท้อนให้เห็นว่าธรรมชาติของเนื้อหาจะเป็นปัจจัยหนึ่งที่สัมพันธ์กับความคลาดเคลื่อนมาตรฐานของแบบสอบ (SE) ไม่ใช่เพียงแต่สัดส่วนของการตรวจให้คะแนนในแบบสอบ และโมเดลการวิเคราะห์แบบสอบเท่านั้น ฉะนั้นคุณครูที่ต้องการสร้างแบบสอบในสาระการเรียนรู้การวัดควรใช้สัดส่วนของการตรวจให้คะแนนในแบบสอบเป็น 60:40 ร่วมกับโมเดลการวิเคราะห์แบบสอบแบบ 2pl แต่สาระการเรียนรู้อื่นๆที่นอกเหนือจากสาระการวัด หากจะพิจารณาถึงปฏิสัมพันธ์ระหว่างสัดส่วนของการตรวจให้คะแนนในแบบสอบและโมเดลการวิเคราะห์แบบสอบหรือปัจจัยในด้านอื่นๆ ก็ควรมีการศึกษาเพิ่มเติมต่อไป

### ข้อเสนอแนะ

1. ด้านโมเดลการวิเคราะห์แบบสอบ นักวิจัยหรือนักวัดผลการศึกษาควรให้ความสำคัญกับการเลือกใช้โมเดลในการวิเคราะห์แบบสอบ เนื่องจากข้อค้นพบในงานวิจัยครั้งนี้พบว่าโมเดลการวิเคราะห์แบบสอบ แบบ 2pl มีผลต่อคุณภาพของข้อสอบด้านความเที่ยงและความคลาดเคลื่อนมาตรฐานของแบบสอบ (SE) ดีกว่าโมเดลการวิเคราะห์แบบสอบ แบบ 1pl แต่เมื่อหากต้องการเปรียบเทียบเฉพาะโมเดลการ

วิเคราะห์อย่างเดี่ยวโดยที่ไม่ได้พิจารณาความคลาดเคลื่อนมาตรฐานของแบบสอบ (SE) อาจเลือกใช้โมเดลการวิเคราะห์โดยตรวจสอบความเหมาะสมระหว่างข้อมูลกับโมเดลการวิเคราะห์ว่ามีความสอดคล้องกันหรือไม่ ไม่ว่าจะเป็นการวิเคราะห์แบบ 1 พารามิเตอร์ 2 พารามิเตอร์ หรือ 3 พารามิเตอร์ โดยพิจารณาจากค่า fit index ประเภทต่างๆ เช่น  $-2\log\text{-likelihood}$  ค่า AIC ค่า BIC เป็นต้น จะเป็นประโยชน์ในการตัดสินใจได้ง่ายขึ้นว่าควรใช้โมเดลการวิเคราะห์แบบใด

2. ด้านเนื้อหาหรือสาระการเรียนรู้ ก็เป็นอีกสิ่งที่ต้องพิจารณาในการสร้างและพัฒนาข้อสอบ เนื่องจากข้อค้นพบในงานวิจัยพบว่า เนื้อหาหรือสาระการเรียนรู้ สัดส่วนของการให้คะแนนในแบบสอบและโมเดลการวิเคราะห์แบบสอบที่แตกต่างกัน ส่งผลต่อคุณภาพของข้อสอบด้านความเที่ยงและความคลาดเคลื่อนมาตรฐานของแบบสอบ (SE) จากผลการวิจัยพบว่าสาระการเรียนรู้การวัดพบปฏิสัมพันธ์ระหว่างสัดส่วนการตรวจให้คะแนนและโมเดลการวิเคราะห์แบบสอบ หากคุณครูในกลุ่มสาระคณิตศาสตร์ ถ้าต้องการพัฒนาแบบสอบสาระการเรียนรู้การวัดเพื่อให้มีความคลาดเคลื่อนมาตรฐานของการประมาณค่าต่างๆควรมีการพิจารณาทั้งสัดส่วนของการตรวจให้คะแนนในแบบสอบและโมเดลการวิเคราะห์ โดยรูปแบบที่ผู้วิจัยแนะนำคือ สัดส่วนของการตรวจให้คะแนนในแบบสอบเป็น 60:40 ร่วมกับโมเดลการวิเคราะห์แบบสอบแบบ 2pl ส่วนสาระการเรียนรู้อื่นสามารถพิจารณาเฉพาะสัดส่วนของการตรวจให้คะแนนหรือโมเดลการวิเคราะห์เพียงอย่างเดียวอย่างหนึ่งก็ได้ แต่หากจะพิจารณาถึงปฏิสัมพันธ์ระหว่างสัดส่วนของการตรวจให้คะแนนในแบบสอบและโมเดลการวิเคราะห์แบบสอบหรือปัจจัยในด้านอื่นๆก็ควรมีการศึกษาเพิ่มเติมต่อไป

## เอกสารอ้างอิง

- เกวลิน ชัยนรงค์. (2554). *การศึกษาสภาพและปัญหาการจัดการเรียนการสอนคณิตศาสตร์ในโรงเรียนมัธยมศึกษา สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานกรุงเทพมหานคร*. (วิทยานิพนธ์ปริญญาโทบริหารบัณฑิต) กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย คณะครุศาสตร์ สาขาการศึกษาคณิตศาสตร์.
- ไข่มุก เลื่องสุนทร. (2552). *การศึกษามโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับจำนวนของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในโรงเรียนสังกัดเขตพื้นที่การศึกษาราชบุรี เขต 1*. (วิทยานิพนธ์ปริญญาโทบริหารบัณฑิต) กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย คณะครุศาสตร์ สาขาการศึกษาคณิตศาสตร์.
- ณัฐไฉไล พริ้งมาตี. (2544). *การศึกษามโนทัศน์ทางคณิตศาสตร์เรื่องเส้นขนานของนักเรียนชั้นมัธยมศึกษาตอนต้น*. (วิทยานิพนธ์ปริญญาโทบริหารบัณฑิต) กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย คณะครุศาสตร์ สาขาการศึกษาคณิตศาสตร์.
- เวชฤทธิ์ อังกะภักทขจร. (2546). *การสังเคราะห์งานวิจัยที่เกี่ยวข้องกับมโนทัศน์ที่คลาดเคลื่อนในวิชาคณิตศาสตร์*. (วิทยานิพนธ์ปริญญาโทบริหารบัณฑิต) กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย คณะครุศาสตร์ สาขาการศึกษาคณิตศาสตร์.

- สุวิมล เสวกสุริยวงศ์. (2553). การสังเคราะห์งานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต) กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย คณะครุศาสตร์ สาขาการวัดและประเมินผลการศึกษา.
- รุ่งนภา แสนอำนวยการ. (2555). ประสิทธิภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนรูปแบบผสม : การประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนความรู้บางส่วน และทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนความรู้บางส่วนแบบทั่วไป. *Journal of Education KhonKaen University*, 35(1) : 58-66
- ศิริชัย กาญจนวาสี. (2552). *ทฤษฎีการทดสอบแบบดั้งเดิม*. พิมพ์ครั้งที่ 6. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- อัญชลี ศรีกลชาญ. (2552). *คุณภาพของการปรับเทียบคะแนนสำหรับแบบสอบรูปแบบผสม: การประยุกต์ใช้การปรับเทียบตามทฤษฎีการตอบสนองข้อสอบด้วยวิธีโค้งคุณลักษณะและการปรับค่าพารามิเตอร์พร้อมกัน*. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต) กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย คณะครุศาสตร์ สาขาการวัดและประเมินผลการศึกษา.
- อุไร จักษ์ตรีมงคล. (2557). การกำหนดค่าให้คะแนน . *วารสารการวัดผลการศึกษา สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ*, 31(89) : 17-26
- นิคม บุญหลาย. (2562). การตรวจให้คะแนนแบบทดสอบอัตนัยให้มีความตรงและความเที่ยง: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดและโมเดลหลายองค์ประกอบของราส์ซ. *วารสารการวัดผลการศึกษา สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ*, 36(99) : 154-169.
- Ou Lydia Liu , Hee-Sun Lee & Marcia C. Linn.(2011). An Investigation of Explanation Multiple-Choice Items in Science Assessment. *Journal of educational assessment*, 16:164–184.
- Kinsey, T.L. (2003). *Comparison of IRT and Rasch Procedures in A Mixed- Item Format Test*. Dissertation Prepared for the Degree of Doctor of Philosophy University of North Texas.
- Michael J. Kolen , Won-Chan Lee. (2011). Psychometric Properties of Raw and Scale Scores on Mixed-Format Tests. *Educational Measurement*, 30(2):15-24
- R. Saen-amnuaiophon. (2012). The Effect of Proportion of Mixed-Format Scoring: Mixed-Format Achievement Tests. *Procedia - Social and Behavioral Sciences*, 69:1522-1528