

การเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์ รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

A Comparison of the Item Information Function of the Mixed-Format Test for Mathematical Literacy with Different Test Lengths and Proportions of Items

นุชรากร นันตะสุยะ^{1*} จุฑาภรณ์ มาสันเทียะ² และ พนิดา พานิชวัฒนะ³

Nucharakorn Nantasuya^{1*}, Chuthaphon Masantiah² and Panida Panidvadana³

(Received: March 19, 2022; Revised: April 5, 2022; Accepted: April 11, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ 1) สร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน และ 2) เปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน กลุ่มตัวอย่างที่ใช้ในการวิจัย ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 900 คน เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน ได้แก่ (1) โมเดล 1 มีความยาว 30 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% (2) โมเดล 2 มีความยาว 30 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% (3) โมเดล 3 มีความยาว 50 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% และ (4) โมเดล 4 มีความยาว 50 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน การวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA) ผลการวิจัยพบว่า

1. ในการสร้างแบบทดสอบรูปแบบผสมได้แบบทดสอบที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน 4 โมเดล โดยโมเดล 1 และ โมเดล 2 ซึ่งมีความยาว 30 ข้อ มีความเที่ยง 0.842 และ 0.803 ตามลำดับ ส่วนโมเดล 3 และ โมเดล 4 ซึ่งมีความยาว 50 ข้อ มีความเที่ยง 0.876 และ 0.872 ตามลำดับ โดยทั้ง 4 โมเดลที่สร้างขึ้นมีข้อสอบรวมจำนวน 11 ข้อ

¹ นักศึกษาปริญญาโท สาขาวิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

² ผู้ช่วยศาสตราจารย์ สาขาวิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

³ อาจารย์ สาขาวิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

¹ Master's degree student, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

² Assistant Professor, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

³ Lecturer, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

* Corresponding Author E-mail: nucharakorn14@gmail.com

2. ผลการเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดลที่แตกต่างกัน พบว่า มีค่าเฉลี่ยสารสนเทศของข้อสอบแตกต่างกันอย่างน้อย 1 คู่ เมื่อเปรียบเทียบค่าเฉลี่ยรายคู่ พบว่า มีค่าเฉลี่ยสารสนเทศของข้อสอบแตกต่างกัน 5 คู่ โดยโมเดล 1 มีค่าสารสนเทศมากกว่าโมเดล 2, 3 และ 4 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และโมเดล 3 มีค่าเฉลี่ยสารสนเทศของข้อสอบสูงกว่าโมเดล 2 และ 4 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

คำสำคัญ: ค่าสารสนเทศของข้อสอบ ทฤษฎีการตอบสนองข้อสอบ แบบทดสอบรูปแบบผสม

Abstract

The objectives of this research were to 1) construct mixed format of mathematical literacy tests with different test lengths and proportions of items, and 2) compare the item information function of the mixed-format tests for mathematical literacy with different test lengths and proportions of items. The sample consisted of 900 grade 9 students. The research instruments were mixed-format tests for mathematical literacy with different test lengths and proportions of items. These tests consisted of the following models: (1) Model 1, with the length of thirty items and the proportion of multiple-choice items and subjective items of 30:70 percent; (2) Model 2, with the length of thirty items and the proportion of multiple-choice items and subjective items of 50:50 percent; (3) Model 3, with the length of fifty items and the proportion of multiple-choice items and subjective items of 30:70 percent; and (4) Model 4, with the length of fifty items and the proportion of multiple-choice items and subjective items of 50:50 percent. The statistics used in the data analysis were frequency, percentage, mean, standard deviation and one-way ANOVA. The findings were as follows:

1. In the construction of the mixed-format tests, 4 different test lengths and proportions of items were obtained. Model 1 and Model 2, each with the length of 30 items, had the reliability of 0.842 and 0.803 respectively. Model 3 and Model 4, each with the length of 50 items, had the reliability of 0.876 and 0.872 respectively. All 4 models of the tests that were created consisted of a combined exam of 11 items.

2. The result of the comparison of the item information function between different models showed that there was at least one pair of different means of item information function. When each individual mean pair was considered, it was found that there were five pairs of different means of item information function. Model 1 had the item information function higher than model 2, 3, and 4, with statistical significance at the .05 level. And Model

3 had the mean of item information function higher than models 2 and 4, with statistical significance at the .05 level.

Keywords: item information function, Item Response Theory, mixed-format test

บทนำ

คณิตศาสตร์มีบทบาทสำคัญต่อความสำเร็จในการเรียนรู้ในศตวรรษที่ 21 เนื่องจากคณิตศาสตร์เป็นเครื่องมือในการศึกษาด้านวิทยาศาสตร์ เทคโนโลยี และศาสตร์อื่น ๆ อันเป็นรากฐานในการพัฒนาบุคลากรของชาติให้มีคุณภาพ ซึ่งถือว่าเป็นสิ่งจำเป็นสำหรับการพัฒนาและการแข่งขันทางเศรษฐกิจของประเทศ เห็นได้จากการประเมินการรู้เรื่องคณิตศาสตร์ (Mathematical Literacy) ซึ่งเป็นจุดเน้นของโครงการประเมินผลนักเรียนร่วมกับนานาชาติ (Programme for International Student Assessment: PISA) ดำเนินการโดยองค์การเพื่อความร่วมมือทางเศรษฐกิจและการพัฒนา (Organisation for Economic Co-operation and Development: OECD) เพื่อสะท้อนว่าระบบการศึกษาของประเทศได้เตรียมเยาวชนของชาติให้มีความพร้อมสำหรับการใช้ชีวิตและการทำงานในอนาคต จากรายงานผลการประเมินโครงการ PISA ในประเทศไทย ตั้งแต่ PISA 2000 ถึงปัจจุบัน พบว่านักเรียนที่จบการศึกษาภาคบังคับ (นักเรียนอายุ 15 ปี) มีคะแนนเฉลี่ยการรู้เรื่องคณิตศาสตร์ต่ำกว่าคะแนนเฉลี่ยของทุกประเทศรวมกันและต่ำกว่าระดับพื้นฐาน (Institute for the Promotion of Teaching Science and Technology, 2018) ผู้วิจัยจึงสนใจศึกษา รูปแบบข้อสอบของ PISA ซึ่งประกอบด้วยข้อสอบแบบปรนัยและอัตนัยรวมอยู่ในฉบับเดียวกัน

ปัจจุบันการประเมินผู้เรียนมีความพยายามในการวัดทักษะความสามารถการคิดวิเคราะห์ คิดสังเคราะห์ คิดสร้างสรรค์ ของผู้เรียนมากกว่าการวัดความรู้ ความจำ และความเข้าใจในเนื้อหาเท่านั้น แต่การเลือกใช้ข้อสอบรูปแบบเดียวมีข้อจำกัดในการวัด มีจุดแข็งและจุดอ่อน (Kinsey, 2003; Kim et al., 2007; Reshetar & Melican, 2010) การทดสอบจำเป็นต้องใช้แบบทดสอบทั้ง 2 รูปแบบควบคู่กันไป เรียกว่า “แบบทดสอบรูปแบบผสม” (Mixed-format Tests) ประกอบด้วยข้อสอบแบบปรนัยและข้อสอบอัตนัย ซึ่งแบบทดสอบรูปแบบผสมสามารถลดข้อจำกัดของข้อสอบปรนัยและข้อสอบอัตนัย โดยสามารถวัดได้ครอบคลุมเนื้อหาที่ต้องการวัดมากขึ้นและวัดระดับความสามารถของผู้สอบในขั้นสูงได้ นอกจากนี้ยังเปิดโอกาสให้ผู้สอบแสดงความรู้ความสามารถของตนเองและช่วยลดโอกาสในการเดาข้อสอบปรนัยได้ (Kinsey, 2003; Reshetar & Melican, 2010; Srikolchan, 2009)

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบทดสอบรูปแบบผสมในประเด็นเกี่ยวกับความยาว พบว่า มีการสร้างแบบทดสอบที่มีความยาวแตกต่างกันตั้งแต่ 10 ข้อ ถึง 80 ข้อ (Nansombat, 1999; Kinsey, 2003; Shin, 2007; Jurich & Goodman, 2009; Saen-amnuaiphon, 2012) โดยผลการวิจัย พบว่า แบบทดสอบรูปแบบผสมไม่เหมาะสมกับแบบทดสอบที่มีความยาวน้อยกว่า 30 ข้อ เนื่องจากจะทำให้ความเที่ยงและความตรงของแบบทดสอบมีค่าต่ำลง ความยาวที่เหมาะสมจะอยู่ระหว่าง 30-50 ข้อ ซึ่งความเที่ยงและความตรงของแบบทดสอบจะมีค่าใกล้เคียงกัน

เมื่อศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบทดสอบรูปแบบผสมในประเด็นสัดส่วนของข้อสอบพบว่า มีการสร้างแบบทดสอบรูปแบบผสมที่มีสัดส่วนระหว่างข้อสอบปรนัยและอัตนัยที่แตกต่างกัน เช่น (1) 90:10 (2) 80:20 (3) 50:50 (4) 30:70 และ (5) 20:80 (Bastari, 2000; Shin, 2007; Saen-amnuaiphon, 2012) โดยผลการวิเคราะห์ พบว่า แบบทดสอบที่มีสัดส่วนของข้อสอบอัตนัยมากกว่าจะให้ค่าสารสนเทศของข้อสอบสูงกว่าและแบบทดสอบที่มีสัดส่วนข้อสอบปรนัยน้อยกว่าอัตนัย จะให้ประสิทธิภาพของแบบทดสอบคือ ความเที่ยง และความตรงสูงสุด

จากการศึกษาเอกสารและงานวิจัยในอดีตที่มีการตรวจสอบคุณภาพของแบบทดสอบรูปแบบผสม มีการศึกษาคุณสมบัติทางจิตมิติโดยใช้ทฤษฎีการตอบสนองข้อสอบ เช่น ค่าอำนาจจำแนก (a) ค่าความยาก (b) ค่าโอกาสในการเดา (c) ค่าความเที่ยง (reliability) ความตรง (validity) ค่าสารสนเทศของข้อสอบ (Item Information Function: IIF) ค่าสารสนเทศของแบบสอบ (Test Information Function: TIF) ความไม่ลำเอียงของข้อคำถาม และอิทธิพลของวิธีวัด (Chaisuwan, 2012; Sawangsri, 2015) การหาคุณภาพเครื่องมือต้องมีการตรวจสอบความเที่ยงเสมอ ซึ่งความเที่ยงเป็นคุณสมบัติที่มีความสำคัญ ถ้าแบบวัดได้มีค่าความเที่ยงสูง แสดงว่าแบบวัดนั้นมีคุณภาพสูงด้วย งานวิจัยนี้จึงมุ่งสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาว 30 ข้อ และ 50 ข้อ รวมถึงมีสัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 2 แบบ คือ 30% : 70% และ 50% : 50% โดยพิจารณาค่าความเที่ยงรวมถึงค่าสารสนเทศของข้อสอบโดยเปรียบเทียบว่าความยาวและสัดส่วนข้อสอบแบบใดจะให้ค่าสารสนเทศของข้อสอบสูงสุด โดยนำองค์ประกอบของการประเมินผลนักเรียนระดับนานาชาติ PISA มาเป็นองค์ประกอบหนึ่งในการสร้างแบบทดสอบในครั้งนี้ด้วย ผลที่ได้จากการวิจัยจะเป็นสารสนเทศสำคัญในการเลือกใช้แบบทดสอบรูปแบบผสมโดยพิจารณาความยาวของแบบทดสอบและสัดส่วนของข้อสอบปรนัยและอัตนัย เพื่อให้ได้แบบทดสอบรูปแบบผสมที่มีความถูกต้องแม่นยำในการประมาณค่า และสามารถประมาณค่าได้ใกล้เคียงกับระดับความสามารถที่แท้จริงของผู้สอบให้มากที่สุด อีกทั้งเป็นแนวทางแก่งานที่เกี่ยวข้องใช้สร้างแบบวัดผลสัมฤทธิ์ทางการเรียน แบบทดสอบการรู้เรื่องคณิตศาสตร์ของผู้เรียนที่มีรูปแบบผสมที่มีคุณสมบัติทางจิตมิติที่เหมาะสมต่อไป

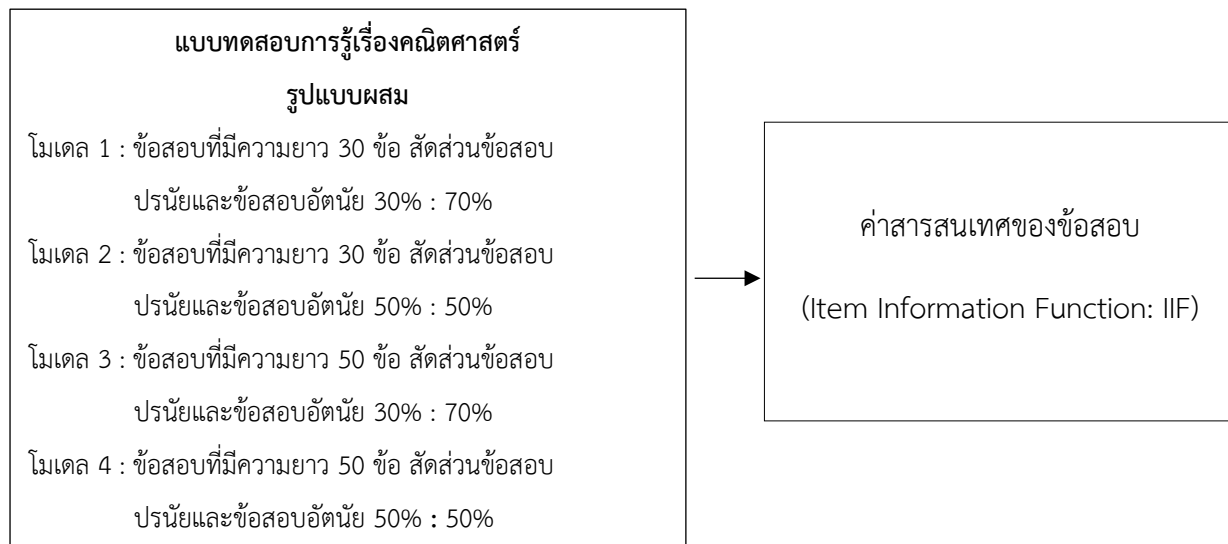
ความมุ่งหมายของการวิจัย

1. เพื่อสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน
2. เพื่อเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มี

ความยาวและสัดส่วนแตกต่างกัน

กรอบแนวคิดในการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องดังกล่าว สามารถกำหนดเป็นกรอบแนวคิดการวิจัย การเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน ได้ดังภาพ 1



ภาพ 1 กรอบแนวคิดในการวิจัย

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2563 ที่กำลังศึกษาอยู่ในจังหวัดนครนายก สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 7 ในจังหวัดนครนายก จำนวน 11 โรงเรียน จำนวน 1,849 คน

กลุ่มตัวอย่างในงานวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 7 ในจังหวัดนครนายก ปีการศึกษา 2563 ได้กลุ่มตัวอย่างจำนวน 1,020 คน เป็นกลุ่มทดลองใช้ จำนวน 120 คน และกลุ่มใช้จริง จำนวน 900 คน แบ่งเป็น 4 กลุ่มย่อย ๆ ละ 225 คน จำแนกตามขนาดโรงเรียน พบว่า เป็นโรงเรียนขนาดใหญ่พิเศษคิดเป็นร้อยละ 62.22 ขนาดใหญ่ 17.78 ขนาดกลาง 17.78 ขนาดเล็ก 2.22 และเมื่อจำแนกตามเพศพบว่าส่วนใหญ่เป็นเพศหญิงคิดเป็นร้อยละ 64.78 ซึ่งจำนวนกลุ่มตัวอย่างที่เหมาะสมสำหรับการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบและแบบทดสอบแบบ 1 พารามิเตอร์ควรมีจำนวน 200 คน (Kinsey, 2003; Srikolchan, 2009; Saen-amnuaiphon, 2012) เนื่องจากงานวิจัยนี้มีค่าพารามิเตอร์ของข้อสอบและแบบสอบแบบ 1 พารามิเตอร์ จำนวน 4 โมเดล จึงได้กลุ่มตัวอย่าง 800 คน และเพื่อป้องกันการสูญหายจึงปรับเพิ่มกลุ่มตัวอย่างเป็นโมเดลละ 225 คน ได้กลุ่มตัวอย่างทั้งหมด 900 คน ตัวอย่างที่ได้มาจากการสุ่มแบบหลายขั้นตอน (Multi-stage random sampling)

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน ซึ่งประกอบด้วยข้อสอบ 2 รูปแบบ คือ ข้อสอบปรนัย เป็นข้อสอบแบบเลือกตอบ 4 ตัวเลือก และข้อสอบอัตนัย เป็นข้อสอบแบบเขียนตอบ แบบตอบสั้น จำนวน 80 ข้อ คัดเลือกเหลือ 67 ข้อ ซึ่งข้อสอบปรนัยและอัตนัยจะเป็นแบบทดสอบที่ตรวจให้คะแนนสองค่า (ตอบถูกต้อง 1 คะแนน และตอบผิดหรือไม่ตอบได้ 0 คะแนน) รวมอยู่ในฉบับเดียวกัน จำนวน 4 โมเดล ได้แก่ (1) โมเดล 1 ข้อสอบที่มีความยาว

30 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% (2) โมเดล 2 ข้อสอบที่มีความยาว 30 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% (3) โมเดล 3 ข้อสอบที่มีความยาว 50 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% และ (4) โมเดล 4 ข้อสอบที่มีความยาว 50 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% โดยแบบทดสอบนี้สร้างขึ้นจากองค์ประกอบของการรู้เรื่องคณิตศาสตร์ ประกอบด้วย (1) สถานการณ์หรือบริบทของปัญหา ได้แก่ บริบทส่วนตัว บริบทสังคม บริบทการงานอาชีพ และบริบทวิทยาศาสตร์ (2) เนื้อหาคณิตศาสตร์ ได้แก่ เนื้อหาปริภูมิและรูปทรงสามมิติ เนื้อหาการเปลี่ยนแปลงและความสัมพันธ์ เนื้อหาปริมาณ และ เนื้อหาความไม่แน่นอน และ (3) กลุ่มสมรรถนะทางคณิตศาสตร์ ได้แก่ สมรรถนะการทำให้ใหม่ สมรรถนะการต่อเชื่อม และ สมรรถนะการสะท้อนและการสื่อสาร ในที่นี้ผู้วิจัยได้นำเสนอข้อสอบรวมที่มีค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) ที่เหมาะสม รายละเอียดตัวอย่างข้อสอบดังตาราง 1

ตาราง 1 ตัวอย่างข้อสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

ข้อสอบปรนัย (ข้อสอบรวม)	ข้อสอบอัตนัย (ข้อสอบรวม)
บริบท : สังคม	บริบท : การงานอาชีพ
เนื้อหา : การเปลี่ยนแปลงและความสัมพันธ์	เนื้อหา : ปริมาณ
สมรรถนะ : การสะท้อนและการสื่อสาร	สมรรถนะ : การทำให้ใหม่
$p = .65$	$p = .53$
$r = .30$	$r = .72$
การเดินทาง การเดินทางจากนครนายก ไปกรุงเทพฯ สามารถเลือกเดินทางได้ 2 วิธี ดังนี้ <u>วิธีที่ 1</u> เดินทางโดยรถประจำทาง รถมามาทุกๆ 10 นาที และใช้เวลาในการเดินทางไปถึงกรุงเทพฯ ๖ เป็นเวลา 80 นาที <u>วิธีที่ 2</u> เดินทางโดยรถตู้ รถมามาทุกๆ 15 นาที และใช้เวลาในการเดินทางไปถึงกรุงเทพฯ ๖ เป็นเวลา 75 นาที ถ้าต้องการเดินทางจากนครนายก ไปยังกรุงเทพฯ ต้องเลือกโดยสารรถแบบเดียวเท่านั้น และรถจะออกเที่ยวแรกพร้อมกันในเวลา 08.00 น 1. พิมพ์ต้องเดินทางจากนครนายกเพื่อไปเรียนในกรุงเทพฯ ๖ เธอมาถึงป้ายรถประจำทาง เวลา 09.10 น. เธอจะเลือกขึ้นรถประจำทาง หรือ รถตู้ จึงจะใช้เวลาในการเดินทางน้อยที่สุดและเมื่อรวมระยะเวลาในการรอรถคิดเป็นระยะเวลาเท่าใด ก. รถประจำทาง ใช้เวลา 85 นาที ข. รถประจำทาง ใช้เวลา 80 นาที ค. รถตู้ ใช้เวลา 75 นาที ง. รถประจำทาง หรือ รถตู้ ก็ได้ ใช้เวลา 80 นาที เท่ากัน	ร้านเบเกอรี่ ทนายต้องการจะเปิดร้านเบเกอรี่ โดยภายในร้านมีผลิตภัณฑ์ต่าง ๆ เช่น ขนมเค้ก ขนมปัง คุกกี้และโดนัต ทนายต้องตัดสินใจเข้าสถานที่เพื่อเปิดร้านเบเกอรี่ โดยมีข้อมูลสถานที่ทั้ง 3 แห่ง ดังนี้ ● ร้าน ก คิดค่าเช่าเดือนละ 5,000 บาท จ่ายค่าเช่าล่วงหน้า 3 เดือน ค่าประกันร้านค้า 1,500 บาท ● ร้าน ข คิดค่าเช่าปีละ 50,000 บาท จ่ายค่าเช่าก่อนครบปีหนึ่งเดือน ค่าประกันร้านค้า 10,000 บาท ● ร้าน ค คิดค่าเช่าเดือนละ 8,000 บาท จ่ายค่าเช่าทุกสิ้นเดือน ไม่คิดค่าประกันร้านค้า 2. หากคำนึงถึงค่าใช้จ่ายในการเข้าสถานที่เท่านั้น ทนายควรจะทำสัญญาเช่ากับร้านใดจึงจะประหยัดค่าใช้จ่ายได้มากที่สุด และจ่ายไปปีละเท่าใด ตอบ.....

3. การเก็บรวบรวมข้อมูล

3.1 นำแบบทดสอบที่ผ่านการคัดเลือกข้อคำถามที่มีค่า IOC ตั้งแต่ .50 ขึ้นไป จำนวน 80 ข้อ ไปทดสอบกับกลุ่มทดลองใช้ (try out) จำนวน 120 คน เพื่อหาคุณภาพของแบบทดสอบ ซึ่งมีข้อสอบที่ผ่านเกณฑ์ 67 ข้อ

3.2 นำแบบทดสอบที่ผ่านคุณภาพมาจัดชุดข้อสอบจำนวน 4 ฉบับ ไปทดสอบกับกลุ่มตัวอย่าง 900 คน ซึ่งในการทดสอบใช้กลุ่มตัวอย่างโมเดลละ 225 คน โดยกำหนดให้ใช้ระยะเวลาในการสอบ 60 นาที

3.3 นำแบบทดสอบมาตรวจให้คะแนน โดยมีเกณฑ์สำหรับการให้คะแนน คือ ตอบถูกให้ 1 คะแนน และตอบผิดหรือไม่ตอบให้ 0 คะแนน

3.4 นำผลที่ได้จากการทดสอบมาวิเคราะห์หาคุณภาพของข้อสอบและค่าสารสนเทศของข้อสอบ

4. การวิเคราะห์ข้อมูล

4.1 การวิเคราะห์คุณภาพของแบบทดสอบ

คุณภาพของแบบทดสอบ พิจารณาจาก 1) ค่าดัชนีความสอดคล้อง (IOC) ด้านความเที่ยงตรงเชิงเนื้อหา (content validity) จากผู้เชี่ยวชาญ 3 ท่าน มีค่า IOC ตั้งแต่ .67 ขึ้นไป ถือว่าผ่านเกณฑ์และได้คัดเลือกไว้ 80 ข้อ คิดเป็นร้อยละ 100.00 2) ค่าความยาก (p) และค่าอำนาจจำแนก (r) เมื่อนำไปทดสอบกับกลุ่มทดลองใช้ (try out) โดยพิจารณาคัดเลือกข้อสอบที่มีค่าความยาก (p) อยู่ระหว่าง .20 ถึง .80 ค่าอำนาจจำแนก (r) ตั้งแต่ .20 ขึ้นไป ซึ่งมีข้อสอบที่ผ่านเกณฑ์ทั้งหมด 67 ข้อ คิดเป็นร้อยละ 83.75 และ 3) ค่าความเที่ยง (reliability) ของแบบทดสอบ ใช้วิธีของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson method) KR_{20} วิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูป

4.2 การวิเคราะห์คุณภาพของข้อสอบ

คุณภาพของข้อสอบ พิจารณาจาก ค่าสารสนเทศของข้อสอบ (item information function) โดยใช้โปรแกรม MULTILOG 7.03 การวิเคราะห์ค่าสารสนเทศของข้อสอบ (Karnjanawasee, 2007) ดังสูตร

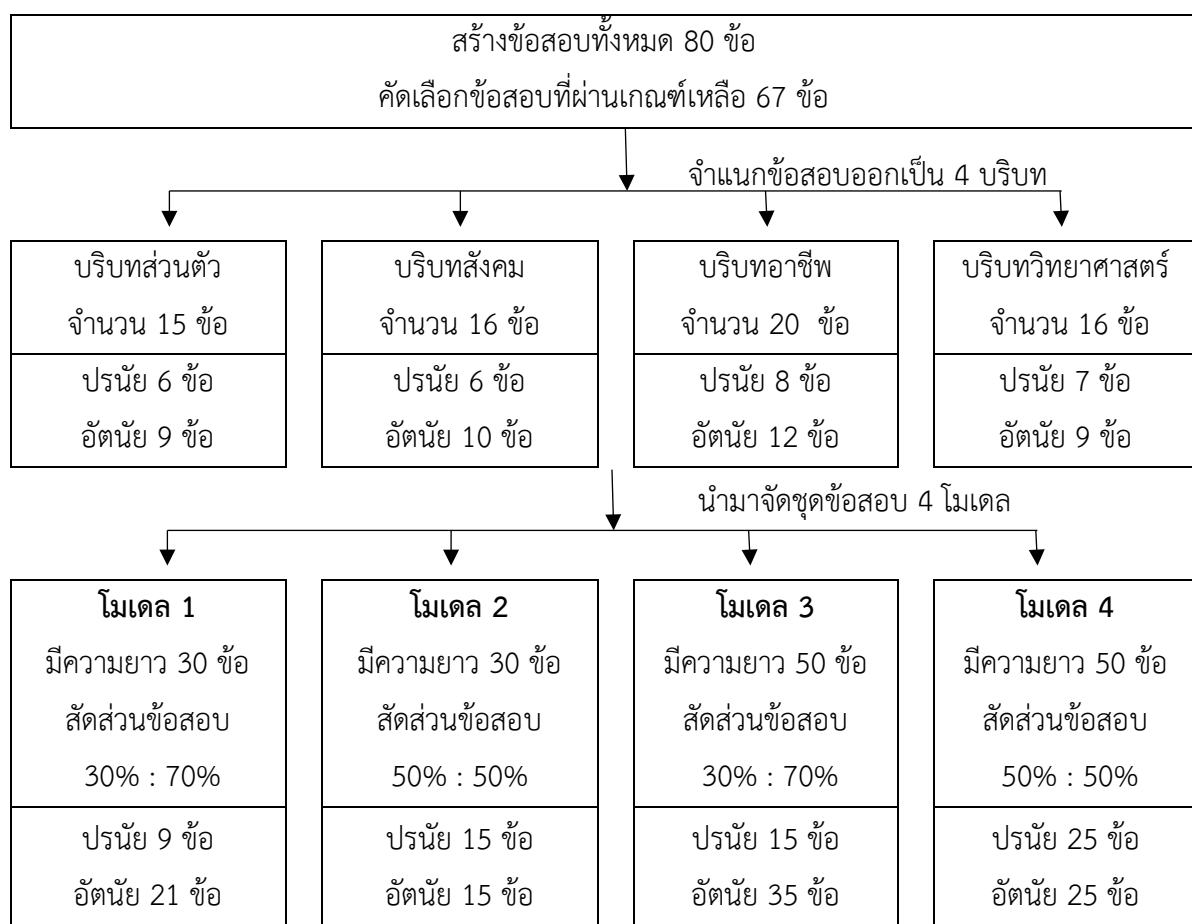
$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)} \quad , i = 1, 2, 3, \dots, k$$

เมื่อ	$I_i(\theta)$	แทน	ค่าสารสนเทศ หรือค่าสารสนเทศที่ได้รับจากข้อสอบข้อที่ i สำหรับผู้ตอบที่มีความสามารถ θ
	$P'_i(\theta) = P'_i$	แทน	ความชันของฟังก์ชันการตอบข้อสอบข้อที่ i ณ ตำแหน่งความสามารถ θ
	$P_i(\theta) = P_i$	แทน	ความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง
	$Q_i(\theta) = Q_i = 1 - P_i(\theta)$		

ผลการวิจัย

1. ผลการสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

1.1 ผลการจัดชุดข้อสอบแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมผู้วิจัยได้คัดเลือกข้อสอบที่ผ่านเกณฑ์การหาคุณภาพข้อสอบ จำนวน 67 ข้อ จากทั้งหมด 80 ข้อ นำมาจัดชุดข้อสอบจำแนกตามสถานการณ์และบริบทที่มีความยาวและสัดส่วนข้อสอบแตกต่างกันจำนวน 4 โมเดล ซึ่งแต่ละโมเดลได้กำหนดให้มีข้อสอบรวมจำนวน 11 ข้อ คิดเป็นร้อยละ 36.67 และร้อยละ 22.00 ตามที่ Srikolchan (2009) ระบุว่า แบบทดสอบรูปแบบผสมควรมีข้อสอบรวมอย่างน้อยร้อยละ 20 ของข้อสอบทั้งหมด ข้อสอบรวมดังกล่าวมีค่าความยาก (p) อยู่ระหว่าง .40 ถึง .69 และค่าอำนาจจำแนก (r) อยู่ระหว่าง .30 ถึง .72 ซึ่งถือว่าอยู่ในเกณฑ์คุณภาพที่เหมาะสม โดยผลการสร้างแบบทดสอบสรุปได้ดังภาพ 2

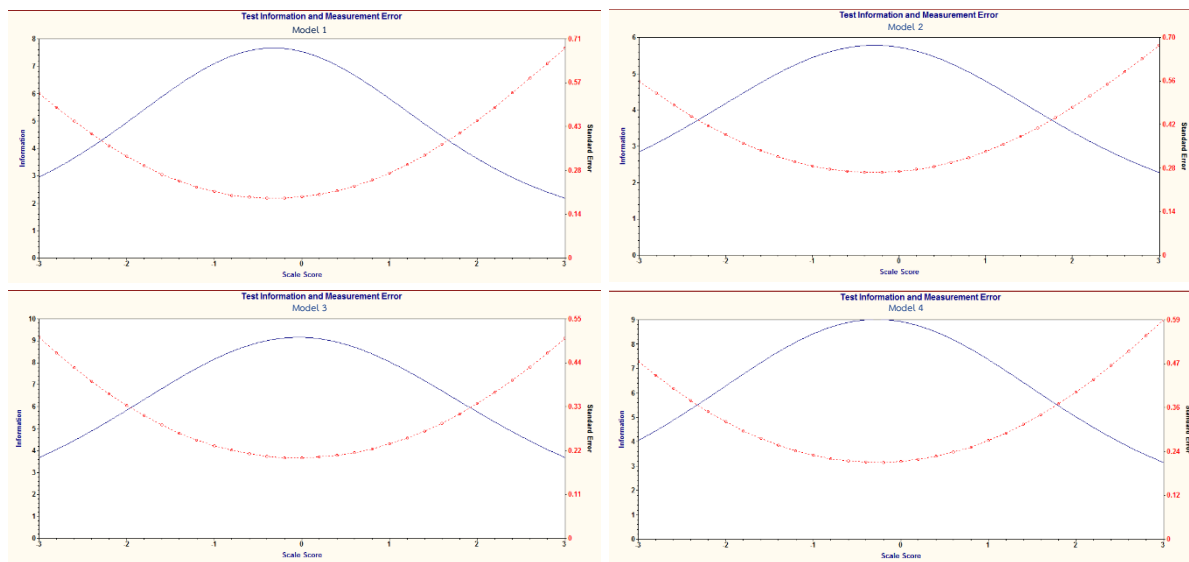


ภาพ 2 การจัดชุดข้อสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

การจัดชุดข้อสอบทั้ง 4 โมเดล มีข้อสอบรวม จำนวน 11 ข้อ เป็นข้อสอบรวมในบริบทส่วนตัว 4 ข้อ บริบทสังคม 4 ข้อ บริบทอาชีพ 1 ข้อ และบริบทวิทยาศาสตร์ 2 ข้อ

2. ผลการเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

2.1 ค่าสารสนเทศของแบบทดสอบ (test information function) การรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน จำนวน 4 โมเดล ณ ช่วงความสามารถอยู่ระหว่าง -3 ถึง 3 พบว่า (1) โมเดล 1 มีค่าสูงสุด เท่ากับ 7.652 ณ ระดับความสามารถเท่ากับ -.40 (2) โมเดล 2 มีค่าสูงสุด เท่ากับ 5.774 ณ ระดับความสามารถ เท่ากับ -.20 (3) โมเดล 3 มีค่าสูงสุด เท่ากับ 9.150 ณ ระดับความสามารถ เท่ากับ .00 และ (4) โมเดล 4 มีค่าสารสนเทศของแบบทดสอบสูงสุด เท่ากับ 8.999 ณ ระดับความสามารถ เท่ากับ -.20 โดยมีผลการวิเคราะห์ค่าสารสนเทศของแบบทดสอบ ดังภาพ 3



ภาพ 3 ค่าสารสนเทศของแบบทดสอบทั้ง 4 โมเดล

2.2 การวิเคราะห์และเปรียบเทียบค่าสารสนเทศของข้อสอบ (item information function) ระหว่างโมเดลที่แตกต่างกัน ซึ่งนำความยาวและสัดส่วนข้อสอบมาพิจารณาพร้อมกันได้ 4 โมเดล โดยการวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA) ผลการวิเคราะห์ ดังตาราง 2

ตาราง 2 ผลการวิเคราะห์และเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดลที่แตกต่างกัน

แหล่งความแปรปรวน	SS	df	MS	F	p
ระหว่างกลุ่ม	.210	3	.070	19.457	.000*
ภายในกลุ่ม	.562	156	.004		
รวม	.772	159			

*p < .05

จากตาราง 2 ผลการวิเคราะห์และเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดลที่แตกต่างกัน พบว่า มีค่าเฉลี่ยสารสนเทศของข้อสอบของโมเดลแตกต่างกันอย่างน้อย 1 คู่ อย่างมีนัยสำคัญทางสถิติ .05 ($F = 19.457, p = .000$) ดังนั้นต้องทำการวิเคราะห์ Post-hoc เพื่อเปรียบเทียบค่าเฉลี่ยรายคู่โดยใช้วิธีของ Tamhane เพราะผลการตรวจสอบของ Levene's test พบว่า มีความแปรปรวนแตกต่างกันอย่างมีนัยสำคัญทางสถิติ .05 (Levene = 6.251, $p = .000$) ดังตาราง 3

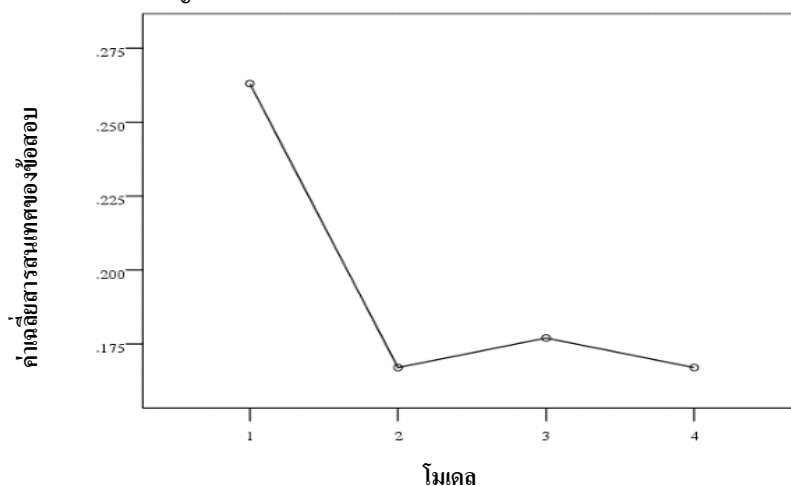
ตาราง 3 ผลการเปรียบเทียบรายคู่ (Post-hoc test) โดยวิธีของ Tamhane ในการเปรียบเทียบ

ค่าสารสนเทศของข้อสอบ (item information function) ระหว่างแต่ละโมเดลที่แตกต่างกัน

โมเดล	$\bar{x}$	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
		.2630	.1670	.1770	.1670
โมเดล 1	.2630	-	.0960*	.0860*	.0960*
			(.004)	(.012)	(.004)
โมเดล 2	.1670		-	-.0100*	.0000
				(.000)	(1.000)
โมเดล 3	.1770			-	.0100*
					(.000)
โมเดล 4	.1670				-

ตัวเลขในวงเล็บ คือ ค่า p-value * มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 3 ผลการเปรียบเทียบรายคู่ (Post-hoc test) โดยวิธีของ Tamhane ในการเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดล 4 แบบ พบว่า มีค่าเฉลี่ยของสารสนเทศของข้อสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 5 คู่ ซึ่งสรุปได้ว่า โมเดล 1 มีค่าเฉลี่ยสารสนเทศของข้อสอบสูงกว่าโมเดล 2, 3 และ 4 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และโมเดล 3 มีค่าเฉลี่ยสารสนเทศของข้อสอบสูงกว่าโมเดล 2 และ 4 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับผลในภาพ 4



ภาพ 4 ผลการเปรียบเทียบค่าเฉลี่ยสารสนเทศของข้อสอบระหว่างโมเดลที่แตกต่างกัน

อภิปรายผล

1. อภิปรายผลการสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน

จากผลการวิเคราะห์ความตรงเชิงเนื้อหาของแบบสอบรูปผสมวัดการรู้เรื่องคณิตศาสตร์ จำนวน 80 ข้อ ผ่านเกณฑ์การพิจารณาความตรงเชิงเนื้อหาจากผู้เชี่ยวชาญ โดยมีค่าดัชนีความสอดคล้องอยู่ระหว่าง 0.67 ถึง 1.00 คัดเลือกไว้จำนวน 80 ข้อ ซึ่งแสดงให้เห็นว่า ข้อสอบที่สร้างขึ้นมีความตรงเชิงเนื้อหา สอดคล้องกับ Saiyod & Saiyod (2000); Pasiphol (2016) ที่ได้กล่าวไว้ว่า ดัชนีความสอดคล้องที่คำนวณได้ต้องมีค่ามากกว่า 0.50 จึงจะมีความตรงเชิงเนื้อหา ที่เป็นเช่นนี้เพราะว่าผู้วิจัยได้ศึกษากรอบแนวคิดของ PISA เพื่อกำหนดกรอบโครงสร้าง (Test Blueprint) ที่ชัดเจน ศึกษางานวิจัยที่เกี่ยวข้องกับการสร้างแบบทดสอบรูปผสมของข้อสอบปรนัยและข้อสอบอัตนัย และตรวจสอบความตรงเชิงเนื้อหา โดยผู้เชี่ยวชาญพิจารณาความสอดคล้องระหว่างข้อคำถามกับเนื้อหาและจุดประสงค์เชิงพฤติกรรม 2 ครั้ง แล้วปรับปรุงแก้ไขตามที่ผู้เชี่ยวชาญเสนอแนะ ส่งผลให้แบบทดสอบที่ผู้วิจัยสร้างขึ้นเป็นแบบทดสอบที่สามารถนำไปใช้ได้

จากผลการวิเคราะห์ความเที่ยงของแบบทดสอบที่แตกต่างกันใน 4 โมเดล พบว่ามีค่าความเที่ยงใกล้เคียงกัน ดังนี้ โมเดล 1 และโมเดล 2 ซึ่งมีความยาว 30 ข้อ มีค่าความเที่ยง 0.842 และ 0.803 ตามลำดับ ขณะที่โมเดล 3 และโมเดล 4 ซึ่งมีความยาว 50 ข้อ มีค่าความเที่ยง 0.876 และ 0.872 ตามลำดับ จะเห็นว่าจำนวนข้อสอบที่มากขึ้นส่งผลให้แบบทดสอบมีค่าความเที่ยงสูงขึ้น สอดคล้องกับ Karnjanawasee (2007) ที่กล่าวว่าความยาวของแบบทดสอบที่ส่งผลต่อความเที่ยงของแบบทดสอบ แบบทดสอบที่มีความยาวมากจะมีผลให้ค่าความเที่ยงของแบบทดสอบสูงกว่าแบบทดสอบที่มีความยาวน้อยกว่า นอกจากนี้ยังสอดคล้องกับ Bastari (2000); Saen-amnuaiophon (2012) ที่กล่าวว่า แบบทดสอบควรมีความยาวอยู่ระหว่าง 30-50 ข้อ เพราะจะมีค่าความเที่ยงและค่าความตรงของแบบทดสอบใกล้เคียงกัน แต่แบบทดสอบที่มีความยาว 50 ข้อ มีแนวโน้มว่ามีประสิทธิภาพสูงกว่าความยาว 30 ข้อและ 10 ข้อ

2. อภิปรายผลการเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน

จากผลการวิเคราะห์ค่าสารสนเทศของแบบทดสอบ (test information function) พบว่าแบบทดสอบทั้ง 4 โมเดลมีค่าสารสนเทศของแบบทดสอบระหว่าง 5.774 ถึง 9.150 ณ ความสามารถช่วง -.40 ถึง .00 สะท้อนให้เห็นว่าแบบทดสอบที่สร้างขึ้นเหมาะสมสำหรับนำไปใช้กับผู้สอบที่มีความสามารถปานกลางไปจนถึงความสามารถต่ำ จะให้ค่าสารสนเทศของข้อสอบที่เพียงพอ สอดคล้องกับงานวิจัยของ Kawmanee (2002) ที่เปรียบเทียบค่าสารสนเทศของแบบทดสอบเลือกตอบวิชาคณิตศาสตร์ ซึ่งผลการวิจัย พบว่า แบบทดสอบที่สร้างขึ้นเหมาะสมสำหรับผู้สอบที่มีความสามารถปานกลางไปจนถึงต่ำ ที่ความสามารถระหว่าง -.40 ถึง .30 แต่อย่างไรก็ดีเมื่อพิจารณาค่าสารสนเทศของแบบทดสอบ พบว่ามีค่าต่ำกว่า 10 ซึ่ง Karnjanawasee (2007) ได้กล่าวว่า ค่าสารสนเทศของแบบทดสอบควรมีค่ามากกว่า 10 ดังนั้นอาจเกิดจากลักษณะของรูปแบบข้อสอบรวมถึงลักษณะของตัวเลือกข้อสอบ ซึ่งสอดคล้องกับ Kawmanee (2002) ที่พบว่า ลักษณะของตัวเลือก

ในข้อสอบปรนัยมีผลต่อค่าสารสนเทศของแบบทดสอบโดยเฉพาะตัวเลือกแบบปลายเปิดส่งผลให้ค่าสารสนเทศของแบบทดสอบสูงกว่าตัวเลือกแบบปกติ

ผลการวิเคราะห์และเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดลที่ต่างกันพบว่า มีค่าเฉลี่ยสารสนเทศของข้อสอบแตกต่างกันอย่างน้อย 1 คู่ เมื่อเปรียบเทียบค่าเฉลี่ยรายคู่ของโมเดลทั้ง 4 พบว่าโมเดล 1 คือความยาว 30 ข้อ สัดส่วนระหว่างข้อสอบปรนัยและอัตนัย 30% : 70% มีค่าสารสนเทศสูงกว่าโมเดล 2, 3 และ 4 และโมเดล 3 คือ ความยาว 50 ข้อ สัดส่วนระหว่างข้อสอบปรนัยและอัตนัย 30% : 70% มีค่าสารสนเทศสูงกว่าโมเดล 2 และ 4 สะท้อนให้เห็นว่า ในกรณีแบบทดสอบที่มีความยาว 30 ข้อ และมีสัดส่วนข้อสอบอัตนัยมากกว่าข้อสอบปรนัย มีค่าสารสนเทศของข้อสอบสูงกว่า อาจเป็นเพราะข้อสอบอัตนัยในการวิจัยครั้งนี้แม้จะเป็นแบบเติมคำตอบสั้น แต่ก็เดาได้ยากกว่าข้อสอบปรนัย จึงสะท้อนความสามารถของผู้สอบได้ดีกว่าปรนัย ซึ่งมีโอกาสในการเดา จึงส่งผลให้ค่าสารสนเทศเมื่อมีข้อสอบอัตนัยสูงกว่าปรนัยด้วย สอดคล้องกับ Nidehah (2000); Kim et al. (2007) ที่กล่าวว่า ถ้ามีการเดาลดลงจะทำให้ค่าสารสนเทศของข้อสอบเพิ่มขึ้น

เมื่อพิจารณาความยาวของข้อสอบ จะเห็นว่า ในงานวิจัยนี้แบบทดสอบที่มีความยาว 30 ข้อ มีค่าสารสนเทศของข้อสอบสูงกว่าแบบทดสอบที่มีความยาว 50 ข้อ ซึ่งให้ผลแตกต่างจากการศึกษาของ Kinsey (2003); Saen-amnuaiophon (2012) ที่พบว่า แบบทดสอบที่มีความยาว 50 ข้อ มีค่าสารสนเทศสูงกว่าแบบทดสอบที่มีความยาว 30 ข้อ ทั้งนี้อาจเนื่องมาจาก ลักษณะของรูปแบบข้อสอบ และอีกประเด็นหนึ่งตามทฤษฎีการตอบสนองข้อสอบ ต้องเป็นการสอบที่ไม่แข่งขันด้านเวลา ความเร็วในการตอบจะต้องไม่มีอิทธิพลต่อผลการตอบ การจัดสอบจึงต้องไม่อยู่ในสถานการณ์ที่สอบแข่งขันกันด้วยเวลา แต่ในสภาพการทดสอบจริงไม่สามารถจัดการสอบที่ไม่แข่งขันด้านเวลาได้ในทุกครั้ง ซึ่งอาจมีผลต่อค่าพารามิเตอร์ของข้อสอบ สอดคล้องกับ Karnjanawasee (2007) ที่กล่าวว่า ค่าสารสนเทศของข้อสอบมีความสอดคล้องกับค่าความยากง่าย กล่าวคือ ค่าสารสนเทศของข้อสอบจะสูงขึ้น เมื่อค่าความสามารถสูงสุดใกล้เคียงกับค่าความยากง่าย

เมื่อพิจารณาสัดส่วนของข้อสอบ พบว่า แบบทดสอบที่มีสัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% มีค่าสารสนเทศของข้อสอบสูงกว่าแบบทดสอบที่มีสัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50%:50% สอดคล้องกับ Kim et al. (2007); Reshetar & Melican (2010) ที่กล่าวถึงจุดเด่นของข้อสอบอัตนัยที่ช่วยลดโอกาสในการเดาคำตอบถูกของผู้สอบ หรือเมื่อไม่มีโอกาสในการเดา ค่าสารสนเทศของข้อสอบจะสูงขึ้น

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 จากผลการวิจัยที่พบว่า แบบทดสอบที่มีความยาว 50 ข้อ มีค่าความเที่ยงสูงกว่าความยาว 30 ข้อ ดังนั้นในบริบทของการนำแบบทดสอบไปใช้ ครูหรือผู้ที่มีส่วนเกี่ยวข้องหากนำไปใช้ในการวิจัยหรือการวัดผลมีระยะเวลาในการทดสอบที่เพียงพอ ควรเลือกใช้แบบทดสอบที่มีความยาว 50 ข้อ เนื่องจากให้ค่าความเที่ยงที่สูงกว่า 30 ข้อ

1.2 จากผลการวิจัยจำแนกตามโมเดลของแบบทดสอบ พบว่า โมเดล 1 ที่มีความยาว 30 ข้อ มีสัดส่วนระหว่างปรนัยและอัตนัย 30% : 70% มีค่าสารสนเทศของข้อสอบสูงกว่าสัดส่วนระหว่างข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% ดังนั้นในสถานการณ์ที่มีระยะเวลาจำกัด ครูอาจประหยัดเวลาโดยสร้างข้อสอบเพียง 30 ข้อได้ แต่ควรมีสัดส่วนข้อสอบอัตนัยสูงกว่าข้อสอบปรนัยจึงจะให้ค่าสารสนเทศของข้อสอบสูง

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรมีการศึกษาคุณสมบัติทางจิตมิติของแบบทดสอบรูปแบบผสมที่มีความยาวและสัดส่วนของข้อสอบที่แตกต่างกัน โดยการประยุกต์ใช้โมเดลตอบสนองข้อสอบ 2 พารามิเตอร์และ 3 พารามิเตอร์ เนื่องจากให้สารสนเทศที่เป็นประโยชน์ในการสร้างคลังข้อสอบของแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสม ซึ่งประกอบด้วยค่าอำนาจจำแนก (a) และค่าโอกาสในการเดา (c) ต้องทำการศึกษาเกี่ยวกับกลุ่มตัวอย่างที่มีขนาดใหญ่ขึ้น เพื่อลดความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ข้อสอบและความสามารถของผู้สอบ

2.2 ควรมีการศึกษาการสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมโดยประยุกต์ใช้ทฤษฎีตอบสนองข้อสอบ ที่มีรูปแบบการตรวจให้คะแนนมากกว่าสองค่า หรือการตรวจให้คะแนนบางส่วน เนื่องจากเนื้อหาวิชาคณิตศาสตร์บางสาระจำเป็นที่จะต้องให้ผู้เรียนเขียนแสดงวิธีคิดซึ่งสามารถตรวจให้คะแนนได้หลายค่า

กิตติกรรมประกาศ

บทความวิจัยเรื่อง การเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน เป็นส่วนหนึ่งของวิทยานิพนธ์ระดับปริญญาโทเรื่อง การเปรียบเทียบคุณสมบัติทางจิตมิติของแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน ซึ่งเป็นส่วนหนึ่งของการสำเร็จการศึกษาในหลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาวิชาการประเมินและการวิจัยการศึกษา มหาวิทยาลัยรามคำแหง โดยมี ผู้ช่วยศาสตราจารย์ ดร.จุฑาภรณ์ มาสันเทียะ เป็นที่ปรึกษาวิทยานิพนธ์หลัก และอาจารย์ ดร.พนิดา พานิชวัฒน์ เป็นที่ปรึกษาวิทยานิพนธ์ร่วม

References

- Bastari, B. (2000). *Linking Multiple – Choice and Constructed – Response Items to a common Proficiency Scale* [Doctoral dissertation]. Massachusetts Amherst University.
- Jurich, D., & Goodman, J. (2009). *A Comparison of IRT Parameter Recovery in Mixed Format Examinations Using PARSCALE and ICL* [Poster presented]. The Annual meeting of Northeastern Educational Research Association.

- Kim, S., Walker, M. E., & McHale, F. (2007). *Equating of Mixed-Format Tests in Large Scale Assessments* [Paper Presented]. The Annual Meeting of the American Educational Research Association (AERA) and The National Council on Measurement in Education (NCME).
- Kinsey, T.L. (2003). *Comparison of IRT and Rasch Procedures in A Mixed- Item Format Test* [Doctoral dissertation]. University of North Texas.
- Reshetar, R., & Melican, G. J. (2010). *Design and Evaluation of Mixed-Format Large Scale Assessments for the Advanced Placement Program (AP)* [Paper presented]. The Annual Meeting of the American Education Research Association in Denver.
- Shin, D. (2007). *A Comparison of Methods of Estimating Subscale Scores for Mixed Format Tests*. Pearson Educational Measurement.

Translate Thai References

- Chaisuwan, K. (2012). *A Comparison of Psychometric Properties among the Short Forms Developed Using Different Methods: An Adversity Quotient Test Study* [Master's thesis]. Chulalongkorn University. (in Thai)
- Institute for the Promotion of Teaching Science and Technology. (2018). *PISA 2015 Assessment Results Science Reading and Math Educational Excellence and Equality*. Success Public Company Limited. (in Thai)
- Karnjanawasee, S. (2007) *New Exam Theory*. (3th Ed.). Chulalongkorn University Press. (in Thai)
- Kawmanee, K. (2002) *A Comparison of Test Information Function of Mathematics Multiple Choice Test with Different Model Options* [Master's thesis]. Thaksin University. (in Thai)
- Nansombat, U. (1999). *A Study of Item Analysis According to Item Response Theory Calculated from Various Sample Sizes and Test Lengths* [Master's thesis]. Mahasarakham University. (in Thai)
- Nidehah, N. (2000). *A Comparison of Test Information Function of Mathematics Multiple Choice Test with Different Type of Choices* [Master's thesis]. Songkla University. (in Thai)
- Pasiphol, S. (2016). *Learning Measurement and Evaluation*. Chulalongkorn University Press. (in Thai)

- Saen-amnuaiphon, R. (2012). *The Efficiency of Mixed-Format Achievement Test: An Application of Partial Credit Model and Generalized Partial Credit Model* [Doctoral dissertation]. Khon Kaen University. (in Thai)
- Saiyod, L., & Saiyod, U. (2000). Techniques for measuring learning outcomes. Suweeriyasan. (in Thai)
- Sawangsi, P. (2015). *Comparison of Psychometric Properties Integrating Indicators of Reading Analytical Thinking and Writing Abilities and Indicators of Content Science of Nine Grade using The Different Item Review Methods* [Master's thesis]. Chulalongkorn University. (in Thai)
- Srikolchan, A. (2009). *The Quality of Equating for Mixed-Format Test: An Application of IRT Equating with Characteristic Curve and Concurrent Calibration Methods* [Doctoral dissertation]. Chulalongkorn University. (in Thai)