



การพัฒนาแบบสอบคู่ขนานเรื่องสมการเชิงเส้นตัวแปรเดียว

Development of Parallelism Test on Linear Equations with One Variable

ธนากร เทพนรงค์¹ และ กิตติทัศน์ หวานฉ่ำ^{2*}

Thanakorn Thepnarong¹ and Kittitas Wancham^{2*}

บทคัดย่อ

แบบสอบคู่ขนานถูกสร้างขึ้นเพื่อลดปัญหาการถูกเปิดเผยของข้อสอบ สามารถใช้ประเมินความรู้ของผู้เรียนได้ ทำให้การสอบเป็นมาตรฐานเดียวกัน และยังสามารถป้องกันข้อสอบรั่วจากการถูกเปิดเผย ด้วยประโยชน์ของแบบสอบคู่ขนานจึงถูกพัฒนาเพื่อนำมาใช้วัดความรู้ทางด้านคณิตศาสตร์เพื่อพัฒนาความรู้ ความสามารถของนักเรียนได้อย่างแท้จริง งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาและตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือกเรื่องสมการเชิงเส้นตัวแปรเดียว และ 2) พัฒนาและตรวจสอบคุณภาพของแบบสอบคู่ขนานและข้อสอบคู่ขนานเรื่องสมการเชิงเส้นตัวแปรเดียว ผู้ให้ข้อมูล คือ ผู้เชี่ยวชาญด้านการสอนคณิตศาสตร์ จำนวน 4 คน และผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา จำนวน 1 คน ตัวอย่างวิจัย คือ ผู้เรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 521 คน ที่ได้จากการสุ่มแบบหลายขั้นตอน ผู้วิจัยตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบ โดยใช้สัดส่วนความสอดคล้อง และตรวจสอบคุณภาพของแบบสอบและข้อสอบคู่ขนาน โดยพิจารณาจากพารามิเตอร์ข้อสอบความเที่ยงแบบสอดคล้องภายใน ความตรงเชิงเนื้อหา ความตรงเชิงโครงสร้าง และดัชนี RMSD ผลการวิจัย พบว่า 1) คุณลักษณะเฉพาะของข้อสอบมีสัดส่วนความสอดคล้องเท่ากับ 1.00 ทุกองค์ประกอบ โดยลักษณะเฉพาะที่พัฒนาขึ้นมีทั้งหมด 5 องค์ประกอบ ได้แก่ (1) วัตถุประสงค์ที่มุ่งวัด (2) ลักษณะคำถาม (3) ขอบข่ายคำตอบ (4) ตัวอย่างข้อสอบ และ (5) ข้อมูลเพิ่มเติม และ 2) แบบสอบเรื่องสมการเชิงเส้นตัวแปรเดียวทั้ง 2 ฉบับ สอดคล้องกับโมเดลแบบ 2 พารามิเตอร์มีความเที่ยงแบบความสอดคล้องภายในวิธีสัมประสิทธิ์แอลฟาของครอนบาคอยู่ในระดับสูง มีความตรงเชิงเนื้อหา และความตรงเชิงโครงสร้าง อีกทั้งมีความเป็นคู่ขนานแบบทั้งโครงสร้างและฟังก์ชันสารสนเทศของแบบสอบ ส่วนข้อสอบมีความเป็นคู่ขนานแบบทั้งโครงสร้างและฟังก์ชันสารสนเทศของข้อสอบทุกคู่

คำสำคัญ : ลักษณะเฉพาะของข้อสอบ, ความเป็นคู่ขนานของข้อสอบ, ความเป็นคู่ขนานของแบบสอบ, สมการเชิงเส้นตัวแปรเดียว

¹ นิสิตมหาบัณฑิต สาขาวิชาวิธีวิทยาการพัฒนานวัตกรรมทางการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อีเมล : thanakorn.34928@gmail.com
Master Student in Division of Methodology for Innovation Development in Education, Department of Educational Research and Psychology,
Faculty of Education, Chulalongkorn University. Email: thanakorn.34928@gmail.com

² ผู้ช่วยศาสตราจารย์ ดร. ประจำสาขาวิชาการวัดและการประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อีเมล : kittitas.w@chula.ac.th
Assistant Professor (Ph.D.) in Division of Educational Measurement and Evaluation, Department of Educational Research and Psychology,
Faculty of Education, Chulalongkorn University. Email: kittitas.w@chula.ac.th

Abstract

Parallel tests are constructed to mitigate the issue of test item exposure. They can be utilized to assess learners' knowledge while ensuring standard equivalence across examinations. Furthermore, they prevent test leaks resulting from exposure. Given these benefits, parallel tests have been developed to measure mathematics proficiency, aiming to truly enhance students' knowledge and abilities. This research aimed to: 1) Develop and validate the quality of item specifications for multiple-choice items on linear equations with one variable. 2) Develop and validate the quality of parallel tests and parallel test items on linear equations with one variable. The informants consisted of 4 mathematics teaching experts and 1 educational measurement and evaluation expert. The research sample comprised 521 Grade 7 students obtained through multi-stage random sampling. The researcher validated the quality of the item specifications using the proportion of agreement. The quality of the tests and parallel test items was evaluated based on item parameters, internal consistency reliability, content validity, construct validity, and the RMSD index. The research findings revealed that: 1) The item specifications achieved a proportion of agreement of 1.00 across all components. The developed specifications comprised 5 components: (1) measured objectives, (2) stimulus attributes, (3) response attributes, (4) sample items, and (5) specification supplement. 2) Both versions of the tests on linear equations with one variable fit the 2-parameter logistic model (2PLM). They exhibited high internal consistency reliability via Cronbach's alpha coefficient, as well as content validity and construct validity. In addition, parallelism was demonstrated across both the test characteristic curves (TCC) and test information functions (TIF). As for the test items, parallelism was also established across both the item characteristic curves (ICC) and item information functions (IIF) for every matched pair.

Keywords: item specifications, item parallelism, test parallelism, linear equations one variable

บทนำ

วิชาคณิตศาสตร์ช่วยพัฒนาความคิดของคน ทำให้มีความคิดสร้างสรรค์ คิดอย่างมีเหตุผล คิดอย่างเป็นระบบ มีแบบแผน สามารถวิเคราะห์ปัญหาและแก้ไขปัญหาได้อย่างมีเหตุผล (กระทรวงศึกษาธิการ, 2551) จึงจำเป็นอย่างยิ่งที่จะต้องวางรากฐานวิชาคณิตศาสตร์ตั้งแต่ระดับชั้นอนุบาล ประถมศึกษา และระดับมัธยมศึกษาให้มีความแข็งแกร่ง สามารถนำไปต่อยอดในการใช้ชีวิตประจำวันของตนเอง (อุกฤษฏ์ ทองอยู่, 2562) ผู้เรียนส่วนมากมีปัญหาเกี่ยวกับวิชาคณิตศาสตร์โดยมีความรู้ในรายวิชาคณิตศาสตร์ไม่เหมาะสมกับระดับชั้นที่เรียน ผู้เรียนมีทัศนคติที่ไม่ดีกับวิชาคณิตศาสตร์ ทำให้เกิดปัญหาในการเรียนรู้ในระดับชั้นที่สูงขึ้น (ประพันธ์ศิริ สุเสารัจ, 2551) เมื่อพิจารณาการทดสอบระดับชาติ (O-net) วิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 ตั้งแต่ปี พ.ศ. 2564 ถึง พ.ศ. 2568 ผู้เรียนชั้นประถมศึกษาปีที่ 6 ทั้งประเทศมีคะแนนเฉลี่ยตั้งแต่ร้อยละ 24.90 ถึง 36.83 (สถาบันทดสอบการศึกษาแห่งชาติ, 2569) ซึ่งมีคะแนนต่ำกว่าเกณฑ์ร้อยละ 50 เป็นสิ่งที่สามารถบ่งบอกได้ว่าผู้เรียนระดับชั้นประถมศึกษาปีที่ 6 ขาดทักษะในการคิด และการ

คำนวณ ซึ่งเมื่อจบการศึกษาในระดับชั้นประถมศึกษาจำเป็นที่จะต้องศึกษาต่อในระดับชั้นมัธยมศึกษาปีที่ 1 ซึ่งเนื้อหาในรายวิชาคณิตศาสตร์ที่เป็นปัญหาสำหรับผู้เรียนส่วนใหญ่คือ สมการเชิงเส้นตัวแปรเดียว (จุฑามณี อินทร์อุทิศ, 2564)

สมการเชิงเส้นตัวแปรเดียวนับว่ามีความสำคัญเป็นอย่างมากเพราะเมื่อผู้เรียนเลื่อนชั้นไปเรียนในระดับที่สูงขึ้น ผู้เรียนจำเป็นที่จะต้องเรียนระบบสมการเชิงเส้นสองตัวแปร อีกทั้งยังต้องประยุกต์กับเนื้อหาส่วนอื่น ๆ เช่น อัตราส่วน สัดส่วน และร้อยละ รวมไปถึงการใช้แก้ปัญหาวิชาอื่น ๆ เช่น ฟิสิกส์ หรือ เคมี (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2563) โดย สุทิน กองเงิน (2547) และ วนิดา นามโคตร (2559) พบว่าความรู้พื้นฐานเดิมเป็นตัวแปรที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ดังนั้นจึงมีความจำเป็นอย่างยิ่งที่จะต้องพัฒนาผู้เรียนให้มีความเข้าใจเรื่องสมการเชิงเส้นตัวแปรเดียว โดยสิ่งที่จะนำมาใช้ในการวัดความเข้าใจของผู้เรียนได้อย่างถูกต้องแม่นยำ คือ แบบสอบ เนื่องจาก แบบสอบเป็นเครื่องมือที่ใช้ทดสอบความรู้และความเข้าใจ ซึ่งจะประกอบไปด้วยชุดข้อสอบเพื่อให้ผู้สอบแสดงความสามารถผ่านการเลือกคำตอบหรือเขียนตอบ

การใช้แบบสอบจำนวน 1 ฉบับในการวัดการเรียนรู้ของผู้เรียนอาจเกิดปัญหา คือ แบบสอบได้รับการเปิดเผย โดยถ้าแบบสอบได้รับการเปิดเผยแล้วย่อมไม่สามารถวัดการเรียนรู้ที่แท้จริงของผู้เรียนได้ ทำให้ขาดความน่าเชื่อถือในการประเมินผล สอดคล้องกับงานของ อนุพงษ์ กันธิวงศ์ (2560) ที่มีการระบุว่า การวัดและประเมินผลในแต่ละครั้งจะต้องคิดถึงคุณภาพและความคงเส้นคงวาเพื่อนำผลสอบไปใช้ได้อย่างถูกต้อง และมีความเชื่อมั่นในข้อมูลเชิงประจักษ์ว่าผลสอบผ่านการประเมินผลอย่างเท่าเทียมกันจึงนิยมใช้แบบสอบคู่ขนานในการประเมินอย่างแพร่หลายในปัจจุบัน กล่าวคือการใช้แบบสอบแบบเดิมซ้ำ ๆ ในการศึกษาไทยทำให้ขาดความเชื่อมั่น และเกิดความไม่ยุติธรรมได้ในการประเมินผล ซึ่ง โชติกา ภาชีผล (2563) เสนอการใช้แบบสอบคู่ขนานหลายฉบับเพื่อใช้ประเมินการเรียนรู้ของผู้เรียน นั่นคือ แบบสอบคู่ขนานจะช่วยให้การดำเนินการสอบมีความเป็นมาตรฐานเดียวกัน โดยแบบสอบคู่ขนาน หมายถึง แบบสอบสองฉบับที่มีลักษณะเหมือนกัน 3 ประการ คือ พฤติกรรมการเรียนรู้ที่มุ่งจะวัด เนื้อหาที่มุ่งวัด และวัตถุประสงค์การเรียนรู้ ซึ่งพิจารณาภาพรวมของข้อคำถามทั้งฉบับ นอกจากนี้เมื่อพิจารณาตามทฤษฎีการตอบสนองข้อสอบ (item response theory) แบบสอบที่มีความเป็นคู่ขนานควรมีโค้งลักษณะแบบสอบ (test characteristic curve) และค่าสารสนเทศของแบบสอบ (test information function) ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ (โชติกา ภาชีผล, 2563; Ali & van Rijn, 2016; Huijing et al., 2022)

การตรวจสอบความเป็นคู่ขนานของแบบสอบ (parallel test) สามารถดำเนินการโดยใช้ทฤษฎีการตอบสนองข้อสอบซึ่งมีการประยุกต์ใช้ได้สอดคล้องกับสภาพจริงมากกว่าทฤษฎีการทดสอบแบบดั้งเดิม โดยทฤษฎีการทดสอบแบบดั้งเดิมมีข้อจำกัด คือ ขาดสารสนเทศในการพิจารณาความเป็นคู่ขนานของข้อสอบรายข้อ หากใช้แบบสอบเดิมกับประชากรใหม่อาจทำให้ค่าพารามิเตอร์ของข้อสอบและความเที่ยงของแบบสอบเปลี่ยนไปตามลักษณะของประชากร ซึ่งทฤษฎีการตอบสนองข้อสอบสามารถตรวจสอบความเป็นคู่ขนานของข้อสอบในระดับรายข้อและในระดับแบบสอบทั้งฉบับ โดยมีความไม่แปรเปลี่ยนพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ (Ali & Rijn, 2016; Boughton & Gierl, 2000; Sanders & Verschoor, 1998;

Thompson, 2009) ซึ่งแบบสอบจะเป็นการพิจารณาในรูปแบบของทั้งฉบับ โดยรูปแบบความเป็นคู่ขนานของแบบสอบแบ่งเป็น 3 รูปแบบ ได้แก่ 1) ความเป็นคู่ขนานแบบทั้งโค้งลักษณะแบบสอบและฟังก์ชันสารสนเทศของแบบสอบ (TCC & TIF – parallel) เป็นแบบสอบคู่ขนานที่มีทั้งโค้งลักษณะของแบบสอบ (Test Characteristic Curve : TCC) และฟังก์ชันสารสนเทศของแบบสอบ (Test Information Function : TIF) ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ 2) ความเป็นคู่ขนานแบบโค้งลักษณะแบบสอบ (TCC – parallel) เป็นแบบสอบคู่ขนานที่มีโค้งลักษณะของแบบสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยมีโค้งสารสนเทศของแบบสอบแตกต่างกัน และ 3) ความเป็นคู่ขนานแบบฟังก์ชันสารสนเทศของแบบสอบ (TIF – parallel) เป็นแบบสอบคู่ขนานที่มีฟังก์ชันสารสนเทศของแบบสอบ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยมีโค้งลักษณะแบบสอบแตกต่างกัน (Ali & van Rijn, 2016; Debeer et al., 2017; McDonald, 2011; Samejima, 1977)

ขณะที่ข้อสอบคู่ขนาน (parallel item) พิจารณาในลักษณะของข้อสอบเป็นรายข้อ เป็นข้อสอบที่มีลักษณะเหมือนกันในแง่ของจุดประสงค์การเรียนรู้ และระดับพฤติกรรมการเรียนรู้ ซึ่งจะมีความแตกต่างจากแบบสอบคู่ขนานในแง่ของการพิจารณาข้อคำถามเป็นรายข้อ นอกจากนี้ข้อสอบคู่ขนานกันจะต้องมีโค้งลักษณะข้อสอบและค่าสารสนเทศของข้อสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ (โชติกา ภาชีผล, 2563; ยุพิน วิชัยพรหม, 2541; สมพิศ กาญจนาเพ็ญกุล, 2556; Ali & van Rijn, 2016; Huijing et al., 2022) โดยสามารถแบ่งรูปแบบความเป็นคู่ขนานของข้อสอบได้เป็น 3 รูปแบบ ได้แก่ 1) ความเป็นคู่ขนานแบบทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ (ICC & IIF – parallel) เป็นข้อสอบคู่ขนานที่มีทั้งโค้งลักษณะของข้อสอบ (Item Characteristic Curve : ICC) และฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function : IIF) ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ 2) ความเป็นคู่ขนานแบบโค้งลักษณะข้อสอบ (ICC – parallel) เป็นข้อสอบคู่ขนานที่มีโค้งลักษณะของข้อสอบ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยมีฟังก์ชันสารสนเทศของข้อสอบแตกต่างกัน และ 3) ความเป็นคู่ขนานแบบฟังก์ชันสารสนเทศของข้อสอบ (IIF – parallel) เป็นข้อสอบคู่ขนานที่มีฟังก์ชันสารสนเทศของข้อสอบ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยมีโค้งลักษณะข้อสอบแตกต่างกัน (Ali & van Rijn, 2016; Debeer et al., 2017; McDonald, 2011; Samejima, 1977)

การตรวจสอบความเป็นคู่ขนานของแบบสอบและข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบสามารถพิจารณาได้จากดัชนีต่าง ๆ ได้แก่ Root mean square deviation (RMSD), Mean square deviation (MSD), Maximum score gap (MSG), Maximum relative characteristic curve gap (MRCG), Mean relative deviation (MRD) และ Maximum relative information gap (MRIG) ซึ่งดัชนีเหล่านี้พิจารณาความแตกต่างของโค้งลักษณะแบบสอบหรือข้อสอบ และจากฟังก์ชันสารสนเทศของแบบสอบหรือข้อสอบ (โชติกา ภาชีผล, 2563; Ali & van Rijn, 2016; Huijing et al., 2022) ซึ่งงานวิจัยส่วนใหญ่ตรวจสอบความเป็นคู่ขนานของแบบสอบและข้อสอบโดยใช้ดัชนี RMSD เนื่องจากมีความเข้มงวดในการตรวจสอบความเป็นคู่ขนานอย่างเหมาะสม พบความผิดพลาดได้น้อยกว่าดัชนีอื่น (โชติกา ภาชีผล, 2563; อนุพงษ์ กันธิวงศ์, 2560; Ali & van Rijn, 2016; Boughton & Gierl, 2000; Sanders & Verschoor, 1998) นอกจากนี้งาน

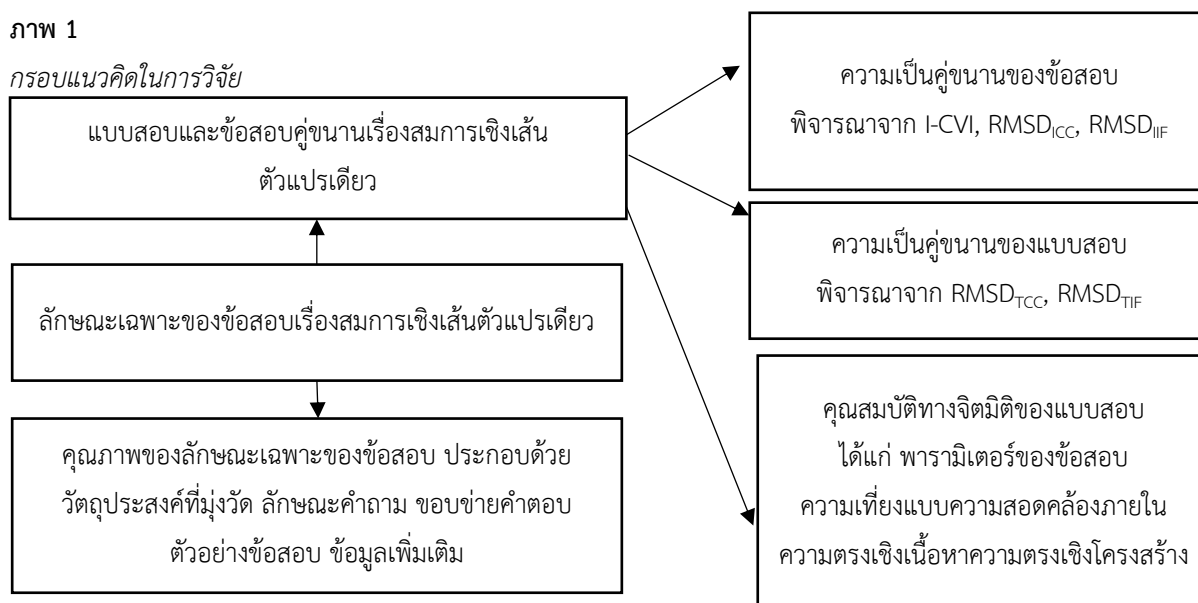
ของอนุพงษ์ กันธิวงศ์ (2560) และสาวินา บุญเสน และคณะ (2566) อธิบายว่า โค้งลักษณะข้อสอบ มีข้อจำกัดคือ อาจมีความคลาดเคลื่อนในการเปรียบเทียบความแตกต่างได้ จึงนิยมใช้สารสนเทศของข้อสอบ แต่โค้งลักษณะข้อสอบมีข้อดีคือทำให้ทราบระดับความยากของข้อสอบ ทำนายผลได้แม่นยำ อีกทั้งสามารถออกข้อสอบได้อย่างยืดหยุ่นสามารถปรับลดหรือเพิ่มข้อสอบเพื่อให้ครอบคลุมความสามารถของผู้สอบได้ ผู้วิจัยจึงเลือกที่จะศึกษาทั้งโค้งลักษณะข้อสอบและโค้งสารสนเทศของข้อสอบ เพื่อให้ตรวจสอบความเป็นคู่ขนานได้ครบถ้วนสมบูรณ์มากยิ่งขึ้น

การสร้างแบบสอบและข้อสอบคู่ขนานผู้วิจัยเลือกสร้างโดยใช้วิธีการทางภาษา และลักษณะเฉพาะของข้อสอบ เนื่องจากการใช้วิธีการทางภาษาเป็นเทคนิคที่ช่วยให้สร้างข้อสอบคู่ขนานที่มีลักษณะแตกต่างกันหลากหลายรูปแบบ ขณะที่การใช้ลักษณะเฉพาะของข้อสอบช่วยให้สามารถกำหนดการสร้างข้อสอบอย่างเป็นระบบทำให้ได้ข้อสอบคู่ขนานที่มีความหลากหลายหลาย (โชติกา ภาชีผล, 2563)

ผู้วิจัยจึงพัฒนาและตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือกรื่องสมการเชิงเส้นตัวแปรเดียวที่พัฒนาขึ้นมาได้ 5 องค์ประกอบ ได้แก่ (1) วัตถุประสงค์ที่มุ่งวัด (2) ลักษณะคำถาม (3) ขอบข่ายคำตอบ (4) ตัวอย่างข้อสอบ และ (5) ข้อมูลเพิ่มเติม ซึ่งองค์ประกอบทั้ง 5 องค์ประกอบสามารถนำไปใช้ในการสร้างลักษณะเฉพาะของบทเรียนอื่น ๆ เพื่อใช้เป็นแนวทางในการสร้างแบบสอบให้มีความเป็นคู่ขนานได้ในทุกเรื่องที่มีความต้องการในการวัดและประเมินผล พร้อมทั้งพัฒนาและตรวจสอบคุณภาพของแบบสอบและข้อสอบคู่ขนานเรื่องสมการเชิงเส้นตัวแปรเดียวที่ครอบคลุมมากกว่างานวิจัยในอดีต เพื่อให้คุณภาพของแบบสอบคู่ขนานดียิ่งขึ้น การวิจัยนี้ช่วยให้ครูและนักวิชาการสามารถนำลักษณะเฉพาะของข้อสอบที่มีคุณภาพไปใช้สร้างข้อสอบคู่ขนานเรื่องสมการเชิงเส้นตัวแปรเดียวสำหรับใช้ตามวัตถุประสงค์ที่ต้องการ รวมถึงนำแบบสอบคู่ขนานไปใช้ในการวัดความรู้ของผู้เรียนได้อย่างถูกต้องแม่นยำ โดยมีกรอบแนวคิดวิจัยดังภาพ 1 ซึ่งสามารถอธิบายได้ว่า แบบสอบและข้อสอบคู่ขนานจะสร้างมาจากลักษณะเฉพาะของข้อสอบที่มีคุณภาพโดยมีทั้ง 5 องค์ประกอบ ได้แก่ วัตถุประสงค์ที่มุ่งวัด ลักษณะคำถาม ขอบข่ายคำตอบ ตัวอย่างข้อสอบ ข้อมูลเพิ่มเติม เมื่อได้ลักษณะเฉพาะของข้อสอบแล้ว ตรวจสอบความเป็นคู่ขนานโดยมีดัชนี I-CVI, RMSD_{ICC}, RMSD_{IIF}, RMSD_{TCC}, RMSD_{TIF} รวมถึงคุณสมบัติทางจิตมิติ ได้แก่ พารามิเตอร์ของข้อสอบ ความเที่ยงแบบสอดคล้องภายใน ความตรงเชิงเนื้อหา และความตรงเชิงโครงสร้าง

ภาพ 1

กรอบแนวคิดในการวิจัย



วัตถุประสงค์

1. เพื่อพัฒนาและตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือกเรื่องสมการเชิงเส้นตัวแปรเดียว
2. เพื่อพัฒนาและตรวจสอบคุณภาพของแบบสอบถามคู่ขนานและข้อสอบคู่ขนานเรื่องสมการเชิงเส้นตัวแปรเดียว ได้แก่ พารามิเตอร์ของข้อสอบ ความเที่ยงแบบสอดคล้องภายใน ความตรงเชิงเนื้อหา ความตรงเชิงโครงสร้าง ความเป็นคู่ขนานของข้อสอบ และความเป็นคู่ขนานของแบบสอบถาม

วิธีการวิจัย

ผู้วิจัยดำเนินการตามระเบียบวิธีการวิจัย ซึ่งประกอบไปด้วยการดำเนินการวิจัย 2 ระยะ ระยะที่ 1 คือ การพัฒนาและตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือก เรื่อง สมการเชิงเส้นตัวแปรเดียว และระยะที่ 2 การพัฒนาและตรวจสอบคุณภาพของแบบสอบถามและข้อสอบคู่ขนาน เรื่อง การแก้สมการเชิงเส้นตัวแปรเดียว ดังนี้

ระยะที่ 1 การพัฒนาและตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือก เรื่อง สมการเชิงเส้นตัวแปรเดียว

การวิจัยในระยะที่ 1 มีผู้ให้ข้อมูล เครื่องมือวิจัย ขั้นตอนการสร้างลักษณะของข้อสอบ และการวิเคราะห์ข้อมูล ดังรายละเอียดต่อไปนี้

1. ผู้ให้ข้อมูล

ผู้ให้ข้อมูล ได้แก่ ผู้เชี่ยวชาญด้านการสอนคณิตศาสตร์ จำนวน 4 คน เป็นผู้สำเร็จการศึกษาตั้งแต่ระดับปริญญาตรีขึ้นไป ในสาขาวิชาที่เกี่ยวข้องกับการสอนคณิตศาสตร์ และมีประสบการณ์การสอนตั้งแต่ 5 ปีขึ้นไป และผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา จำนวน 1 คน เป็นผู้ที่สำเร็จการศึกษา

ตั้งแต่ระดับปริญญาโทขึ้นไป ในสาขาวิชาที่เกี่ยวข้องกับการวัดและการประเมินผลการศึกษา และมีประสบการณ์สอนตั้งแต่ 5 ปีขึ้นไป

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือวิจัย คือ แบบตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือก เรื่องสมการเชิงเส้นตัวแปรเดียว เป็นแบบตรวจสอบรายการ (checklist) ซึ่งมี 2 ตัวเลือก ได้แก่ เหมาะสม และไม่เหมาะสม โดยแบ่งออกเป็น 5 ประเด็น ได้แก่ (1) วัตถุประสงค์ที่มุ่งวัด (2) ลักษณะคำถาม (3) ขอบข่ายคำตอบ (4) ตัวอย่างข้อสอบ และ (5) ข้อมูลเพิ่มเติม

3. ขั้นตอนการสร้างลักษณะของข้อสอบ

3.1 กำหนดมาตรฐานและตัวชี้วัดจากเรื่องสมการเชิงเส้นตัวแปรเดียวเพื่อนำมาสร้างข้อสอบคู่ขนาน แล้ววิเคราะห์เนื้อหาที่มุ่งวัดออกมาเป็นประเด็นหลักสำหรับสร้างตารางวิเคราะห์ข้อสอบให้ตรงกับตัวชี้วัดในเรื่องสมการเชิงเส้นตัวแปรเดียว โดยมีมาตรฐาน จำนวน 1 มาตรฐาน และ 1 ตัวชี้วัด ดังนี้ มาตรฐาน ค 1.3 ใช้นิพจน์ สมการ และอสมการ อธิบายความสัมพันธ์หรือช่วยแก้ปัญหาที่กำหนดให้ และตัวชี้วัด คือ เข้าใจและใช้สมบัติของการเท่ากันและอสมการเท่ากัน เพื่อวิเคราะห์และแก้ปัญหาโดยใช้สมการเชิงเส้นตัวแปรเดียว

3.2 สร้างตารางวิเคราะห์ข้อสอบเพื่อกำหนดระดับพฤติกรรมที่มุ่งวัดและเนื้อหาสำหรับการสร้างลักษณะเฉพาะของข้อสอบ โดยมีตารางวิเคราะห์ข้อสอบดังตาราง 1

3.3 สร้างลักษณะเฉพาะของข้อสอบหลายตัวเลือกเรื่องสมการเชิงเส้นตัวแปรเดียว โดยมี 5 องค์ประกอบ ได้แก่ (1) วัตถุประสงค์ที่มุ่งวัด (2) ลักษณะคำถาม (3) ขอบข่ายคำตอบ (4) ตัวอย่างข้อสอบ และ (5) ข้อมูลเพิ่มเติม

3.4 ให้ผู้เชี่ยวชาญด้านการสอนคณิตศาสตร์ และผู้เชี่ยวชาญด้านการวัดและการประเมินผล การศึกษาตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบ โดยมีองค์ประกอบที่ใช้ในการตรวจสอบ ได้แก่ (1) วัตถุประสงค์ที่มุ่งวัด (2) ลักษณะคำถาม (3) ขอบข่ายคำตอบ (4) ตัวอย่างข้อสอบ และ (5) ข้อมูลเพิ่มเติม

ตาราง 1

ตารางวิเคราะห์ข้อสอบ เรื่อง สมการเชิงเส้นตัวแปรเดียว

เนื้อหา	ระดับพฤติกรรมพุทธิพิสัย						รวม
	ความ จำ	ความ เข้าใจ	การนำไปใช้	การ วิเคราะห์	การประ เมิน ค่า	การสร้าง สรรค์	
การหาค่านิพจน์ พีชคณิต	0	2	0	0	0	0	2
การเขียน พีชคณิต	0	2	0	0	0	0	2

ตาราง 1 (ต่อ)

ตารางวิเคราะห์ข้อสอบ เรื่อง สมการเชิงเส้นตัวแปรเดียว

เนื้อหา	ระดับพฤติกรรมพุทธิพิสัย						รวม
	ความ จำ	ความ เข้าใจ	การนำไปใช้	การ วิเคราะห์	การประ เมิน ค่า	การสร้าง สรรค์	
สมการและ คำตอบของ สมการ	1	4	1	0	0	0	6
การแก้สมการ เชิงเส้นตัวแปร เดียว	1	3	0	0	0	0	4
โจทย์ปัญหา เรื่องสมการเชิง เส้นตัวแปรเดียว	0	0	6	0	0	0	6
รวม	2	11	7	0	0	0	20

4. การวิเคราะห์ข้อมูล

ตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือกเรื่อง สมการเชิงเส้นตัวแปรเดียว โดยใช้สัดส่วนความสอดคล้อง (proportion of agreement : Pi) โดยเป็นค่าที่ใช้วัดความเห็นที่ตรงกันของผู้ประเมินตั้งแต่ 2 คนขึ้นไป ซึ่งควรมีค่ามากกว่า 0.61 (Landis & Koch, 1977) คำนวณโดยนำจำนวนครั้งที่ผู้ประเมินให้ผลเหมือนกันหารด้วยจำนวนครั้งที่ทั้งหมดที่ประเมิน

ระยะที่ 2 การพัฒนาและตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบหลายตัวเลือกเรื่องสมการเชิงเส้นตัวแปรเดียว

การวิจัยในระยะที่ 2 มีผู้ให้ข้อมูล ประชากรและตัวอย่าง เครื่องมือวิจัย การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล ดังรายละเอียดต่อไปนี้

1. ผู้ให้ข้อมูล

ผู้ให้ข้อมูล ได้แก่ ผู้เชี่ยวชาญด้านการสอนคณิตศาสตร์ จำนวน 4 คน และผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา จำนวน 1 คน ซึ่งมีคุณสมบัติเหมือนผู้ให้ข้อมูลในระยะที่ 1

2. ประชากรและตัวอย่าง

2.1 ประชากร คือ ผู้เรียนชั้นมัธยมศึกษาปีที่ 1 โรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษา มัธยมศึกษากรุงเทพมหานคร เขต 1 และ เขต 2 จำนวน 38,481 คน

2.2 ตัวอย่าง คือ ผู้เรียนชั้นมัธยมศึกษาปีที่ 1 โรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษา มัธยมศึกษากรุงเทพมหานคร เขต 1 และ เขต 2 จำนวน 600 คน เนื่องจากกำหนดขนาดตัวอย่างวิจัยตามการวิเคราะห์ข้อมูลโดยใช้ทฤษฎีการตอบสนองข้อสอบ พบว่าต้องใช้ตัวอย่างขั้นต่ำ 500 คน (DeMars, 2010)

แต่เนื่องจากป้องกันการขาดหายของตัวอย่างวิจัย จึงกำหนดเป็น 600 คน โดยได้มาจากการสุ่มแบบหลายขั้นตอน โดยขั้นตอนที่ 1 ดำเนินการสุ่มเขตโดยใช้การสุ่มอย่างง่าย ได้เขตตอนเมืองกับเขตพญาไท ขั้นตอนที่ 2 สุ่มโรงเรียนโดยใช้การสุ่มอย่างง่ายได้โรงเรียนสีกัน (วัฒนานคร์อุปถัมภ์) และโรงเรียนสามเสนวิทยาลัย และในขั้นตอนที่ 3 สุ่มนักเรียนในระดับชั้นมัธยมศึกษาปีที่ 1 โรงเรียนละ 300 คน

3. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือวิจัย คือ (1) แบบตรวจสอบความตรงเชิงเนื้อหา โดยเป็นมาตรฐานค่า 4 ระดับ ได้แก่ 1 หมายถึง ข้อคำถามไม่สอดคล้องกับวัตถุประสงค์การเรียนรู้ 2 หมายถึง ข้อคำถามค่อนข้างสอดคล้องกับวัตถุประสงค์การเรียนรู้ 3 หมายถึง ข้อคำถามสอดคล้องกับวัตถุประสงค์การเรียนรู้ 4 หมายถึง ข้อคำถามสอดคล้องอย่างยิ่งกับวัตถุประสงค์การเรียนรู้ และ (2) แบบตรวจสอบความเป็นคู่ขนานของแบบสอบและข้อสอบคู่ขนานเรื่อง สมการเชิงเส้นตัวแปรเดียว เป็นมาตรฐานค่า 5 ระดับ ได้แก่ 1 หมายถึง เห็นด้วยอย่างยิ่งว่าข้อสอบคู่ขนานกัน 2 หมายถึง เห็นด้วยว่าข้อสอบคู่ขนานกัน 3 หมายถึง ไม่แน่ใจว่าข้อสอบคู่ขนานกัน 4 หมายถึง ไม่เห็นด้วยว่าข้อสอบคู่ขนานกัน 5 หมายถึง ไม่เห็นด้วยอย่างยิ่งว่าข้อสอบคู่ขนานกัน

4. การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูล ประกอบด้วย แบบสอบหลายตัวเลือก 2 ฉบับที่มีความเป็นคู่ขนานกัน (ฉบับ A และ ฉบับ B) ซึ่งสร้างมาจากลักษณะเฉพาะของข้อสอบที่พัฒนาขึ้นในระยะที่ 1 เวลาในการทำข้อสอบ 50 นาที โดยตัวอย่างวิจัย 1 คน จะทำแบบสอบทั้งหมด 2 ฉบับ ซึ่งแบ่งตัวอย่างวิจัยเป็น 2 กลุ่ม ดังนี้

(1) ตัวอย่างวิจัยกลุ่มที่ 1 จำนวน 300 คน ทำแบบสอบฉบับ A ก่อน หลังจากนั้นทำแบบสอบฉบับ B

(2) ตัวอย่างวิจัยกลุ่มที่ 2 จำนวน 300 คน ทำแบบสอบหลายตัวเลือกฉบับ B ก่อน หลังจากนั้นทำแบบสอบฉบับ A

การสลับลำดับในการตอบคำถามมีวัตถุประสงค์เพื่อลดความคลาดเคลื่อนจากลำดับ และความเหนื่อยล้าในการทำข้อสอบ ทั้งนี้ ตัวอย่างวิจัยจำนวน 521 คน มีผลการตอบข้อสอบครบถ้วนทั้ง 2 ฉบับ สำหรับนำไปวิเคราะห์ข้อมูล ส่วนตัวอย่างวิจัยอีก 79 คน คัดออกจากการวิจัย เนื่องจากมีผลการตอบไม่ครบถ้วน ซึ่งข้อมูลสูญหาย (missing data) ไม่เกิดผลกระทบต่อความแม่นยำในการวิเคราะห์เนื่องจากตัวอย่างวิจัยมีขนาดมากกว่าตัวอย่างขั้นต่ำ

ตาราง 2

ตารางแสดงการแบ่งตามระดับความสามารถของผู้เรียน

โรงเรียน	ระดับ	จำนวนเพศชาย	คิดเป็นร้อยละ	จำนวนเพศหญิง	คิดเป็นร้อยละ
ความสามารถ					
โรงเรียนที่ 1 (261 คน)	กลุ่มเก่ง	44 คน	8.45	43 คน	8.25
	กลุ่มกลาง	43 คน	8.25	44 คน	8.45
	กลุ่มอ่อน	44 คน	8.45	43 คน	8.25

ตาราง 2 (ต่อ)

ตารางแสดงการแบ่งตามระดับความสามารถของผู้เรียน

โรงเรียน	ระดับ	จำนวนเพศชาย	คิดเป็นร้อยละ	จำนวนเพศหญิง	คิดเป็นร้อยละ
ความสามารถ					
โรงเรียนที่ 2 (260 คน)	กลุ่มเก่ง	43 คน	8.25	44 คน	8.45
	กลุ่มกลาง	44 คน	8.45	43 คน	8.25
	กลุ่มอ่อน	43 คน	8.25	43 คน	8.25
รวม (521 คน)	กลุ่มเก่ง	87 คน	16.70	87 คน	16.70
	กลุ่มกลาง	87 คน	16.70	87 คน	16.70
	กลุ่มอ่อน	87 คน	16.70	86 คน	16.50

5. การวิเคราะห์ข้อมูล

5.1 ตรวจสอบความตรงเชิงเนื้อหาของแบบสอบหลายตัวเลือกเรื่องสมการเชิงเส้นตัวแปรเดียว ฉบับ A และ B โดยใช้ดัชนี I-CVI และ S-CVI ซึ่งควรมีค่ามากกว่า 0.80 (Polit & Beck, 2006)

5.2 ตรวจสอบความเป็นคู่ขนานของข้อสอบหลายตัวเลือก เรื่อง สมการเชิงเส้นตัวแปรเดียว ก่อนการทดลองใช้ โดยใช้ดัชนี I-CVI ซึ่งควรมีค่ามากกว่า 0.80 (Polit & Beck, 2006)

5.3 ตรวจสอบคุณสมบัติทางจิตมิติของแบบสอบแต่ละฉบับดังนี้

1) วิเคราะห์ค่าพารามิเตอร์ข้อสอบซึ่งประมาณค่าโดยใช้ทฤษฎีการตอบสนองข้อสอบ วิเคราะห์โดยใช้โปรแกรม R โดยเลือกโมเดลการตอบสนองข้อสอบที่สอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ซึ่งมีโมเดลที่พิจารณา 3 โมเดล ได้โมเดล 1PL, 2PL และ 3PL ผู้วิจัยเปรียบเทียบความสอดคล้องของโมเดล โดยพิจารณาจาก AIC, BIC, $-2 \log \text{likelihood}$, likelihood ratio test และพิจารณาความสอดคล้องเชิงสมบูรณ์ SRMR ซึ่งจะมีค่าน้อยกว่า 0.08 (Hu & Bentler, 1999) โดยโมเดลแบบ 2 พารามิเตอร์สอดคล้องกับผลการตอบ จึงประมาณค่าพารามิเตอร์ข้อสอบ 2 ค่า ได้แก่ ความยาก และอำนาจจำแนก

2) วิเคราะห์ความเที่ยงแบบความสอดคล้องภายในวิธีสัมประสิทธิ์แอลฟาครอนบาค โดยใช้โปรแกรม R

3) วิเคราะห์ความตรงเชิงโครงสร้างโดยใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน โดยใช้โปรแกรม LISREL

5.4 ตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ดัชนี RMSD_{TCC} และ RMSD_{TIF} สามารถแบ่งความเป็นคู่ขนานของแบบสอบได้เป็น 4 กรณี คือ 1) ความเป็นคู่ขนานแบบทั้งโด่งลักษณะแบบสอบและฟังก์ชันสารสนเทศของแบบสอบ 2) ความเป็นคู่ขนานแบบโด่งลักษณะแบบสอบ 3) ความเป็นคู่ขนานแบบฟังก์ชันสารสนเทศของแบบสอบ และ 4) ไม่มีความเป็นคู่ขนานทั้งแบบโด่งลักษณะแบบสอบและฟังก์ชันสารสนเทศของแบบสอบ โดยดัชนีมีค่า 0-1 ซึ่งอยู่ในรูปของคะแนนดิบ

5.5 ตรวจสอบความเป็นคู่ขนานของข้อสอบโดยใช้ดัชนี RMSD_{ICC} และ RMSD_{IIF} สามารถแบ่งความเป็นคู่ขนานของข้อสอบได้เป็น 4 กรณี คือ 1) ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชัน

สารสนเทศของข้อสอบ 2) ความเป็นคู่ขนานแบบโค้งลักษณะข้อสอบ 3) ความเป็นคู่ขนานแบบฟังก์ชัน
สารสนเทศของข้อสอบ และ 4) ไม่มีความเป็นคู่ขนานทั้งแบบโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของ
ข้อสอบ โดยดัชนีมีค่า 0-1 ซึ่งอยู่ในรูปของคะแนนดิบ

ผลการวิจัย

ผลการวิจัยแบ่งเป็น 2 ส่วน ดังรายละเอียดต่อไปนี้

1. ผลการพัฒนาและตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือกเรื่องสมการ เชิงเส้นตัวแปรเดียว

จากการตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือกเรื่อง สมการเชิงเส้นตัว
แปรเดียว โดยใช้สัดส่วนความสอดคล้อง พบว่า ลักษณะเฉพาะของข้อสอบทั้ง 5 องค์ประกอบ ได้แก่
1) วัตถุประสงค์ที่มุ่งวัด 2) ลักษณะคำถาม 3) ลักษณะคำตอบ 4) ตัวอย่างข้อสอบ และ 5) ข้อมูลเพิ่มเติม
มีสัดส่วนความสอดคล้อง เท่ากับ 1.00 ดังตาราง 3

ตาราง 3

ผลการตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบโดยใช้สัดส่วนสอดคล้อง

ข้อที่	วัตถุประสงค์ที่มุ่งวัด	ลักษณะคำถาม	ลักษณะคำตอบ	ตัวอย่างข้อสอบ	ข้อมูลเพิ่มเติม
1	1.00	1.00	1.00	1.00	1.00
2	1.00	1.00	1.00	1.00	1.00
3	1.00	1.00	1.00	1.00	1.00
4	1.00	1.00	1.00	1.00	1.00
5	1.00	1.00	1.00	1.00	1.00
6	1.00	1.00	1.00	1.00	1.00
7	1.00	1.00	1.00	1.00	1.00
8	1.00	1.00	1.00	1.00	1.00
9	1.00	1.00	1.00	1.00	1.00
10	1.00	1.00	1.00	1.00	1.00
11	1.00	1.00	1.00	1.00	1.00
12	1.00	1.00	1.00	1.00	1.00
13	1.00	1.00	1.00	1.00	1.00
14	1.00	1.00	1.00	1.00	1.00
15	1.00	1.00	1.00	1.00	1.00
16	1.00	1.00	1.00	1.00	1.00
17	1.00	1.00	1.00	1.00	1.00
18	1.00	1.00	1.00	1.00	1.00
19	1.00	1.00	1.00	1.00	1.00
20	1.00	1.00	1.00	1.00	1.00

พารามิเตอร์อย่างมีนัยสำคัญทางสถิติที่ระดับ.05 [$\Delta\chi^2(19,n=521)= 41.79, p=.002$] ทั้งนี้โมเดล 2 พารามิเตอร์มีความสอดคล้องกับข้อมูลเชิงประจักษ์ไม่แตกต่างจาก 3 พารามิเตอร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 [$\Delta\chi^2(20,n=521)= 28.75, p=.09$] โดยมีดัชนีความสอดคล้องของโมเดลดังตาราง 5

จากผลการประมาณค่าพารามิเตอร์ข้อสอบโดยใช้โมเดล 2 พารามิเตอร์ พบว่า ข้อสอบมีพารามิเตอร์ความยาก (b) ตั้งแต่ -2.14 ถึง -1.39 ซึ่งถือว่าเป็นเกณฑ์ที่ยอมรับได้ตั้งแต่ -2.50 ถึง 2.50 โดยค่าพารามิเตอร์ความยากตั้งแต่ -2.5 ถึง 1.0 เป็นข้อสอบระดับค่อนข้างง่าย หมายถึง ผู้สอบส่วนใหญ่ทำข้อสอบได้ถูกต้อง และมีพารามิเตอร์อำนาจจำแนก (a) ตั้งแต่ 0.42 ถึง 1.52

แบบสอบฉบับ B มีความสอดคล้องกับโมเดล 2 พารามิเตอร์มากที่สุด ซึ่งพิจารณาจากผลวิเคราะห์ likelihood ratio test พบว่าโมเดล 2 พารามิเตอร์สอดคล้องกับข้อมูลเชิงประจักษ์มากกว่าโมเดล 1 พารามิเตอร์อย่างมีนัยสำคัญทางสถิติที่ระดับ.05 [$\Delta\chi^2(19,n=521)= 44.83, p < .001$] ทั้งนี้โมเดล 2 พารามิเตอร์มีความสอดคล้องกับข้อมูลเชิงประจักษ์ไม่แตกต่างจาก 3 พารามิเตอร์ อย่างมีนัยสำคัญทางสถิติที่ระดับ.05 [$\Delta\chi^2(20,n=521)= 23.88, p = .25$]

จากผลการประมาณค่าพารามิเตอร์ข้อสอบโดยใช้โมเดล 2 พารามิเตอร์ พบว่า ข้อสอบมีพารามิเตอร์ความยาก (b) ตั้งแต่ -2.22 ถึง -1.39 และมีพารามิเตอร์อำนาจจำแนก (a) ตั้งแต่ 0.42 ถึง 1.75

ตาราง 5

ดัชนีความสอดคล้องของโมเดลสำหรับแบบสอบฉบับ A

ดัชนีความสอดคล้อง	1PL	2PL	3PL
-2 log likelihood	9255.332	9213.520	9211.700
AIC	9297.332	9293.520	9331.700
BIC	9386.703	9463.750	9587.050
SRMSR	0.07	0.06	0.06

ดัชนีความสอดคล้องของโมเดลสำหรับแบบสอบเรื่องสมการเชิงเส้นตัวแปรเดียว แสดงดังตาราง 6

ตาราง 6

ดัชนีความสอดคล้องของโมเดลสำหรับแบบสอบฉบับ B

ดัชนีความสอดคล้อง	1PL	2PL	3PL
-2 log likelihood	9212.446	9167.608	9144.626
AIC	9254.446	9247.608	9264.627
BIC	9343.817	9417.838	9519.972
SRMSR	0.07	0.06	0.06

2.2 ผลการวิเคราะห์ความเที่ยงแบบความสอดคล้องภายในวิธีสัมประสิทธิ์แอลฟาของครอนบาค และความตรงเชิงเนื้อหา

แบบสอบเรื่องสมการเชิงเส้นตัวแปรเดียวฉบับ A และ B มีความเที่ยงแบบความสอดคล้องภายในวิธีสัมประสิทธิ์แอลฟาของครอนบาค เท่ากับ 0.70 และ 0.71 ตามลำดับ ซึ่งมีความเที่ยงอยู่ในระดับสูง (Taber, 2018) สำหรับความตรงเชิงเนื้อหา ข้อสอบมีค่า I-CVI เท่ากับ 1.00 ทุกข้อ ทั้งฉบับ A และ B และมีค่า S-CVI เท่ากับ 1.00 ทั้งฉบับ A และ B

2.3 ผลการวิเคราะห์ความตรงเชิงโครงสร้าง

แบบสอบเรื่องสมการเชิงเส้นตัวแปรเดียวมีความตรงเชิงโครงสร้างทั้ง 2 ฉบับ เนื่องจากผลการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่าสำหรับแบบสอบฉบับ A โมเดลการวัดมีความสอดคล้องกับข้อมูลเชิงประจักษ์ [$\chi^2(77, n=521) = 94.28, p=.09$] โดยน้ำหนักองค์ประกอบของข้อสอบทุกข้อมีนัยสำคัญทางสถิติที่ระดับ .05 ทุกข้อ มีค่าตั้งแต่ 0.16 ถึง 0.70 ขณะที่แบบสอบฉบับ B โมเดลการวัดมีความสอดคล้องกับข้อมูลเชิงประจักษ์ [$\chi^2(80, n=521) = 88.30, p=.25$] โดยน้ำหนักองค์ประกอบของข้อสอบทุกข้อมีนัยสำคัญทางสถิติที่ระดับ .05 ทุกข้อ มีค่าตั้งแต่ 0.30 ถึง 0.65

2.4 ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบหลายตัวเลือกเรื่องสมการเชิงเส้นตัวแปรเดียว

ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบหลายตัวเลือกเรื่องสมการเชิงเส้นตัวแปรเดียว ก่อนการทดลองใช้ พบว่าข้อสอบมีค่าดัชนี I-CVI = 1.00 ทุกคู่ ส่วนผลการตรวจสอบความเป็นคู่ขนานหลังการทดลองใช้ โดยพิจารณา $RMSD_{ICC}$ และ $RMSD_{IIF}$ พบว่ามีค่าไม่เกิน 0.50 ทุกข้อ (Ali & van Rijn, 2016) แสดงว่าข้อสอบทั้ง 20 คู่ มีความเป็นคู่ขนานกัน ดังตาราง 7

ตาราง 7

ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบ

ข้อ	$RMSD_{IIF}$	$RMSD_{ICC}$	การแปลผล
1	0.0005	0.1416	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
2	0.0379	0.0473	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
3	0.0543	0.0215	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
4	0.0918	0.0705	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
5	0.0262	0.0214	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
6	0.0365	0.0643	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
7	0.0321	0.0447	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
8	0.0259	0.0140	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
9	0.0109	0.0084	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
10	0.0181	0.0374	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
11	0.0185	0.0109	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
12	0.0262	0.0216	ความเป็นคู่ขนานแบบทั้งโด่งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ

ตาราง 7 (ต่อ)

ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบ

ข้อ	RMSD _{III}	RMSD _{ICC}	การแปลผล
13	0.0364	0.0365	ความเป็นคู่ขนานแบบทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
14	0.0074	0.0051	ความเป็นคู่ขนานแบบทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
15	0.0022	0.0056	ความเป็นคู่ขนานแบบทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
16	0.1273	0.0463	ความเป็นคู่ขนานแบบทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
17	0.0176	0.0189	ความเป็นคู่ขนานแบบทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
18	0.0042	0.0099	ความเป็นคู่ขนานแบบทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
19	0.0121	0.0182	ความเป็นคู่ขนานแบบทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ
20	0.0314	0.0250	ความเป็นคู่ขนานแบบทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ

2.5 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบหลายตัวเลือก เรื่องสมการเชิงเส้นตัวแปรเดียว

ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบหลายตัวเลือกเรื่องสมการเชิงเส้นตัวแปรเดียว พบว่า แบบสอบมีความเป็นคู่ขนานแบบทั้งโค้งลักษณะแบบสอบและฟังก์ชันสารสนเทศของแบบสอบ เนื่องจาก $RMSD_{TCC}$ มีค่าเท่ากับ 0.0705 ซึ่งมีค่าไม่เกิน 0.50 และ $RMSD_{TIF}$ มีค่าเท่ากับ 0.1266 ซึ่งมีค่าไม่เกิน 0.50

อภิปรายผล

ผู้วิจัยแบ่งการอภิปรายผลการวิจัยเป็น 2 ประเด็น ดังนี้

1. การอภิปรายการพัฒนาและตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบหลายตัวเลือก เรื่องสมการเชิงเส้นตัวแปรเดียว

จากการตรวจสอบคุณภาพของลักษณะเฉพาะของข้อสอบเรื่องสมการเชิงเส้นตัวแปรเดียว พบว่า ทุกองค์ประกอบมีสัดส่วนความสอดคล้องเท่ากับ 1.00 ซึ่งหมายความว่า ลักษณะเฉพาะของข้อสอบควรมีองค์ประกอบที่สำคัญ 5 องค์ประกอบ ได้แก่ 1) วัตถุประสงค์ที่มุ่งวัด 2) ลักษณะคำถาม 3) ลักษณะคำตอบ 4) ตัวอย่างข้อสอบ และ 5) ข้อมูลเพิ่มเติม กล่าวคือ ลักษณะเฉพาะของข้อสอบเรื่องสมการเชิงเส้นตัวแปรเดียวที่พัฒนาขึ้นแสดงให้เห็นถึงขอบเขตเนื้อหาของการประเมิน รูปแบบการประเมิน จำนวนคำถาม และประเภทของคำถามที่จะใช้ในการประเมิน ทำให้สามารถใช้เขียนข้อคำถาม ให้มีความรอบคอบ รัดกุม เกิดความชัดเจน มีความเป็นปรนัย อีกทั้งเป็นแนวทางสำหรับผู้สร้างคำถาม โดยข้อกำหนดเฉพาะของเนื้อหาการประเมินจะแสดงให้เห็นถึงความมั่นใจและความสอดคล้องกันของคะแนนที่คาดหวังกับคะแนนที่เกิดขึ้นจริง ซึ่งความสอดคล้องดังกล่าวจะเกี่ยวข้องกับความตรงของแบบสอบ (โชติกา ภาชีผล, 2563; Johnson & Johnson, 2009) ดังนั้นการที่มีความสอดคล้องเท่ากับ 1.00 แสดงถึงความสำคัญกับทุกองค์ประกอบที่ทำให้ได้ข้อคำถามเกี่ยวกับเรื่องสมการเชิงเส้นตัวแปรเดียวที่ดี ซึ่งจะเป็นแนวทางในการนำไปสร้างข้อสอบคู่ขนาน ซึ่งแต่ละองค์ประกอบควรมีลักษณะดังต่อไปนี้ 1) วัตถุประสงค์ที่วัด จะต้องมีการอธิบายถึงรายละเอียดของ

ข้อสอบเพื่อใช้ในการชี้แจงให้เข้าใจจุดมุ่งหมายของข้อสอบที่ต้องการให้ผู้สอบบรรลุ 2) ลักษณะคำถาม จะต้องมีความชี้แจงจากรายละเอียดของข้อสอบที่ชัดเจนเพื่อให้สอดคล้องกับวัตถุประสงค์ที่มุ่งวัด 3) ลักษณะคำตอบ จะต้องมีการแสดงขอบข่ายของคำตอบที่ถูกต้องซึ่งอาจเป็นตัวเลือกในข้อสอบหรือคำตอบที่ผู้สอบเขียนขึ้น 4) ตัวอย่างข้อสอบ เป็นแนวทางสำหรับผู้ที่ต้องการจะนำข้อสอบไปใช้และสำหรับผู้ที่ต้องการแนวทางในการออกข้อสอบ และ 5) ข้อมูลเพิ่มเติม จะต้องมีการแสดงรายละเอียดอื่น ๆ ที่เป็นประโยชน์ต่อผู้สอบที่จำเป็นต้องรู้ ดังนั้นเมื่อองค์ประกอบเป็นไปตามที่บรรยายลักษณะเหล่านี้ ทำให้การสร้างข้อสอบคู่ขนานเรื่องสมการเชิงเส้นตัวแปรเดียวสามารถทำได้อย่างมีคุณภาพ (Davidson & Lynch, 2002; Foghahae, 2019; Spaan, 2006; Owen, 2018)

2. การอภิปรายการพัฒนาและตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบหลายตัวเลือก เรื่องสมการเชิงเส้นตัวแปรเดียว

แบบสอบหลายตัวเลือกเรื่องสมการเชิงเส้นตัวแปรเดียวทั้ง 2 ฉบับ มีคุณภาพอยู่ในระดับที่ยอมรับได้ เมื่อพิจารณาจากพารามิเตอร์ของข้อสอบ ความเที่ยงแบบสอดคล้องภายใน ความตรงเชิงเนื้อหา และความตรงเชิงโครงสร้าง แสดงว่าแบบสอบประกอบด้วยข้อสอบที่มีความยากเหมาะสม สามารถจำแนกความแตกต่างระหว่างผู้สอบได้อย่างดี นอกจากนี้แบบสอบยังสามารถใช้ในการวัดการเรียนรู้ของผู้เรียนได้อย่างถูกต้องแม่นยำและมีความคลาดเคลื่อนต่ำ (โชติกา ภาชีผล, 2563 ; Boughton & Gierl, 2000; Gronlund & Waugh, 2009) ขณะที่ประเด็นค่าของผลน้ำหนักร้อยประกอบของข้อสอบจากการวิเคราะห์องค์ประกอบเชิงยืนยันมีค่าน้อยในบางข้อ พบว่าอาจเกิดจากหลายสาเหตุ เช่น มีการใช้ภาษาที่ตีความได้หลากหลายรูปแบบ จึงเลือกตอบแบบสุ่ม ทำให้ข้อมูลไม่มีน้ำหนักมากพอ และมีความคลาดเคลื่อนที่สูง (ไชยันต์ สกฤตศรีประเสริฐ, 2556)

จากผลการตรวจสอบความเป็นคู่ขนานพบว่า แบบสอบทั้ง 2 ฉบับ มีความเป็นคู่ขนานทั้งโค้งลักษณะแบบสอบและฟังก์ชันสารสนเทศของแบบสอบ โดยความเป็นคู่ขนานแบบทั้งโค้งลักษณะแบบสอบและฟังก์ชันสารสนเทศของแบบสอบ จะเป็นแบบสอบที่ตรวจสอบแล้วพบว่า $RMSD_{TCC}$ และ $RMSD_{TIF}$ มีค่าไม่เกิน 0.50 (Ali & van Rijn, 2016) อีกทั้งข้อสอบทุกคู่มีความเป็นคู่ขนานทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ ซึ่งการที่ข้อสอบจะมีความเป็นคู่ขนานทั้งโค้งลักษณะข้อสอบและฟังก์ชันสารสนเทศของข้อสอบ จะต้องมีความค่าดัชนี $RMSD_{ICC}$ และ $RMSD_{IIF}$ ไม่เกิน 0.50 (Ali & van Rijn, 2016) ความเป็นคู่ขนานรูปแบบนี้จะทำให้การเปรียบเทียบคะแนนระหว่างแบบสอบมีความคลาดเคลื่อนต่ำทำให้สามารถนำผลการเปรียบเทียบคะแนนระหว่างแบบสอบไปใช้ได้อย่างถูกต้องแม่นยำ (Ali & van Rijn, 2016; McDonald, 2011; Samejima, 1977)

ดังนั้นการที่จะสร้างแบบสอบให้ดีและมีคุณภาพอีกทั้งมีความเป็นคู่ขนานกันจำเป็นที่จะต้องปฏิบัติดังนี้ ต้องสร้างและพัฒนาลักษณะเฉพาะของข้อสอบขึ้นมาให้มีองค์ประกอบทั้ง 5 องค์ประกอบ โดย 1) วัตถุประสงค์ที่มุ่งวัด 2) ลักษณะคำถาม 3) ลักษณะคำตอบ 4) ตัวอย่างข้อสอบ และ 5) ข้อมูลเพิ่มเติม (Davidson & Lynch, 2002; Foghahae, 2019; Spaan, 2006; Owen, 2018) ต่อมาเป็นวิธีการทางภาษาที่จะช่วยสร้างข้อคำถามให้มีความเป็นคู่ขนานมากยิ่งขึ้น โดยมี 5 วิธี ได้แก่ การสลับคำตอบมาเป็นคำถาม

การเปลี่ยนคำถามตรงกันข้ามหรือเป็นการปฏิเสธ การใช้คำถามเหมือนกันแต่ตัวเลือกมีความแตกต่างกัน การใช้คำถามต่างกันแต่ตัวเลือกเหมือนกัน และการเปลี่ยนถ้อยคำในคำถามแต่ใจความของคำถามยังเหมือนเดิม (โชติกา ภาชีผล, 2563) ซึ่งเมื่อสร้างเสร็จจะให้ผู้เชี่ยวชาญทางด้านการวัดและประเมินผลกับผู้เชี่ยวชาญในวิชาดังกล่าวเป็นผู้ทรงคุณวุฒิในการวัดว่าข้อสอบมีความเป็นคู่ขนานกันหรือไม่ เมื่อทำตามขั้นตอนดังกล่าว จะทำให้ข้อคำถามแต่ละข้อสามารถวัดในประเด็นเนื้อหาเดียวกัน วัดคุณสมบัติและพฤติกรรมที่วัดเดียวกัน ซึ่งเป็นไปตามแนวคิดของแบบสอบคู่ขนาน นอกจากนี้หากนำไปใช้กับผู้เรียนในกลุ่มอื่นที่มีความสามารถต่างกัน แต่เมื่อวัดในประเด็นเนื้อหาเดียวกัน วัดคุณสมบัติและพฤติกรรมการเรียนรู้เดียวกันย่อมทำให้สามารถเปรียบเทียบคะแนนระหว่างกันได้ ช่วยลดปัญหาการใช้ข้อสอบซ้ำและการเปิดเผยข้อสอบ (โชติกา ภาชีผล, 2563) ซึ่งสามารถนำไปใช้ได้ทั้งระดับโรงเรียนและเขตพื้นที่การศึกษา

ข้อเสนอแนะ

ข้อเสนอแนะสำหรับการนำผลวิจัยไปใช้

1.1 การวิจัยนี้ได้พัฒนาลักษณะเฉพาะของข้อสอบเรื่องสมการเชิงเส้นตัวแปรเดียวที่มีคุณภาพสามารถนำไปใช้เป็นต้นแบบในการสร้างข้อสอบคู่ขนานเพื่อใช้ในการทดสอบที่มีความเป็นคู่ขนานและลดปัญหาการเปิดเผยข้อสอบ

1.2 การวิจัยนี้ได้พัฒนาแบบสอบคู่ขนานเรื่องสมการเชิงเส้นตัวแปรเดียว สามารถนำไปใช้ในการทดสอบผู้เรียนในกรณีที่มีการสอบซ้ำ และสามารถเปรียบเทียบคะแนนระหว่างแบบสอบโดยมีความคลาดเคลื่อนต่ำ

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

การวิจัยครั้งต่อไปควรพัฒนาลักษณะเฉพาะของข้อสอบเป็นโมเดลข้อสอบ เพื่อประยุกต์ใช้เทคโนโลยีในการสร้างข้อสอบคู่ขนานอัตโนมัติ ทำให้ได้ข้อสอบที่มีคุณภาพในเวลาอันรวดเร็ว

รายการอ้างอิง

ภาษาไทย

กระทรวงศึกษาธิการ. (2551). *หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2551*.

จุฑามณี อินทร์อุทิศ. (2564). *การจัดการเรียนรู้โดยใช้กิจกรรมเป็นฐานร่วมกับการเรียนรู้แบบร่วมมือที่มีต่อความสามารถในการแก้ปัญหาทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียน เรื่อง สมการเชิงเส้นตัวแปรเดียวของนักเรียนชั้นมัธยมศึกษาปีที่ 1* [วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ, มหาวิทยาลัยศรีนครินทรวิโรฒ]. DSpace at Srinakharinwirot University. <https://ir-thesis.swu.ac.th/items/fbac9720-2c0c-4cef-a492-d8f56e74a2ee>

โชติกา ภาชีผล. (2563). *การพัฒนาแบบสอบคู่ขนาน: การสร้างและวิเคราะห์ความเป็นคู่ขนาน*. สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

- ไชยันต์ สกกุลศรีประเสริฐ. (2556). การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory factor analysis). *วารสารจิตวิทยาคลินิก*, 44(1), 1-16. https://www.researchgate.net/publication/308692590_karwikheraahxngkhpraxbcheingyunyan
- ประพันธ์ศิริ สุเสารัจ. (2551). *การพัฒนาการคิด* (พิมพ์ครั้งที่ 2). 9119 เทคนิคพรินต์ติ้ง.
- ยุพิน วิชัยพรหม. (2541). *การศึกษาความเป็นคู่ขนานของแบบทดสอบ ที่สร้างจากลักษณะเฉพาะของข้อสอบ วิชาวิทยาศาสตร์ ว 101* [วิทยานิพนธ์ปริญญาโทบริหารบัณฑิต, มหาวิทยาลัยเชียงใหม่]. สำนักงานการวิจัยแห่งชาติ. https://doi.nrct.go.th/ListDoi/listDetail?Resolve_DOI=10.14457/CMU.the.1998.53
- วนิดา นามโคตร. (2559). การศึกษาปัจจัยที่ส่งผลต่อปัญหาการเรียนวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษา ปีที่ 3 โรงเรียนราชสีมาวิทยาลัย. *วารสารออนไลน์บัณฑิตศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง*, 3(1), 1-17. http://www.edu-journal.ru.ac.th/AbstractPdf/2559-3-1_1534429567_5814620022.pdf
- สถาบันทดสอบการศึกษาแห่งชาติ. (2569, 27 มีนาคม). รายงานสรุปผลการทดสอบ O-NET ป.6 ปีการศึกษา 2568. NIETS. <https://www.niets.or.th/th/content/view/27073>
- สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2563). *หนังสือรายวิชาพื้นฐาน คณิตศาสตร์ เล่ม 2 ชั้นมัธยมศึกษาปีที่ 1 กลุ่มสาระการเรียนรู้คณิตศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560)* (พิมพ์ครั้งที่ 3).
- สาวินา บุญเสน, จุฑาภรณ์ มาสันเทียะ, และ พนิดา พานิชวัฒนะ. (2566). การพัฒนาแบบสอบคู่ขนานวิชาคณิตศาสตร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนสังกัดกรุงเทพมหานคร โดยการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์. *วารสารการวัดผลการศึกษามหาวิทยาลัยมหาสารคาม*, 29(1), 245-258. <https://so02.tci-thaijo.org/index.php/jemmsu/article/view/253938>
- สุทิน กองเงิน. (2547). *ปัจจัยที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 6 จังหวัดมหาสารคาม* [วิทยานิพนธ์ปริญญาโทบริหารบัณฑิต, มหาวิทยาลัยมหาสารคาม]. Mahasarakham University Library. <https://opac.msu.ac.th/bibitem?bibid=b00007904>
- สมพิศ กาญจนาเพ็ญกุล. (2556). การสร้างข้อสอบคู่ขนานอิงโดเมนโดยใช้รูปแบบฟาเซท วิชาคณิตศาสตร์ เรื่องวงกลม ชั้นมัธยมศึกษาปีที่ 3. *วารสารวิชาการศรีปทุม ชลบุรี*, 9(4), 56-60. https://www.chonburi.spu.ac.th/journal/booksearch/upload/1255-56-60_construction.pdf

อนุพงษ์ กันธิวงศ์. (2560). การเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน: การประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ [วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ, จุฬาลงกรณ์มหาวิทยาลัย]. Chula Digital Collections. <https://digital.car.chula.ac.th/cgi/viewcontent.cgi?article=2227&Context=chulaetd>

อุกฤษฏ์ ทองอยู่. (2562). การพัฒนาความสามารถการแก้โจทย์ปัญหาทางคณิตศาสตร์ โดยใช้การเรียนรู้แบบร่วมมือเทคนิค TGT ชั้นประถมศึกษาปีที่ 2 [วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ, มหาวิทยาลัยธุรกิจบัณฑิต]. DPU Library. <https://libdoc.dpu.ac.th/thesis/Ukrit.Tho.pdf>

ภาษาอังกฤษ

Ali, U. S., & van Rijn, P.W. (2016). An evaluation of different statistical targets for assembling parallel forms in item response theory. *Applied Psychological Measurement*, 40(3), 163–179. <https://doi.org/10.1177/0146621615613308>

Boughton, K. A., & Gierl, M. J. (2000, April 24). *Automated test assembly procedures for criterion-referenced testing using optimization heuristics* [Paper presentation]. Annual Meeting of the American Educational Research Association, New Orleans, Louisiana.

Davidson, F., & Lynch, B. K. (2002). *Test craft: A teacher's guide to writing and using language test specifications*. Yale University Press.

Debeer, D., Ali, U. S., & van Rijn, P. W. (2017). Evaluating statistical targets for assembling parallel mixed-format test forms. *Journal of Educational Measurement*, 54(2), 218-242. <https://doi.org/10.1111/jedm.12142>

DeMars, C. (2010). *Item response theory: Understanding statistics measurement*. Oxford University Press.

Foghahae, Z. (2019). A review of standardizing an English second language test-item through reverse engineering. *Language Teaching Research Quarterly*, 11, 31-42. <https://api.eurokd.com/Uploads/Article/213/ltrq.2019.11.04.pdf>

Gronlund, N. E., & Waugh C. K. (2009). *Assessment of student achievement* (9th ed.). Pearson.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>

- Huijing, Z., Junjie, L., Pingfei, Z., & Chunhua, K. (2022). Two test assembly methods with two statistical targets. *Frontiers in Psychology*, 13, Article 786772. <https://doi.org/10.3389/fpsyg.2022.786772>
- Johnson, S., & Johnson, R. (2009). *Conceptualising and interpreting reliability*. Ofqual. <https://assets.publishing.service.gov.uk/media/5a814e92ed915d74e33fd7c7/2010-02-05-conceptualising-and-interpreting-reliability.pdf>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. <https://doi.org/10.2307/2529310>
- McDonald, R. P. (2011). *Test theory: A unified treatment*. Lawrence Erlbaum Associates. https://api.pageplace.de/preview/DT0400.9781135675318_A25036655/preview-9781135675318_A25036655.pdf
- Owen, N. (2018). Test specifications. In J. I. Lontas (Ed.), *The TESOL encyclopedia of english language teaching* (pp.1-6). John Wiley & Sons.
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489-497. <https://doi.org/10.1002/nur.20147>
- Sanders, P. F., & Verschoor, A. J. (1998). Parallel test construction using classical item parameters. *Applied Psychological Measurement*, 22(3), 212-223. <https://doi.org/10.1177/01466216980223002>
- Samejima, F. (1977). A method of estimating item characteristic functions using the maximum likelihood estimate of ability. *Psychometrika*, 42(2), 163-191. <https://doi.org/10.1007/BF02294047>
- Spaan, M. (2006). Test and item specifications development. *Language Assessment Quarterly*, 3(1), 71-79. https://doi.org/10.1207/s15434311laq0301_5
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in science education*, 48, 1273-1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Thompson, N. A. (2009). *Ability estimation with item response theory*. Assessment Systems Corporation. [https://assess.com/docs/Thompson_\(2009\)_-_Ability_estimation_with_IRT.pdf](https://assess.com/docs/Thompson_(2009)_-_Ability_estimation_with_IRT.pdf)