

การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบ อัตโนมัติสำหรับการศึกษาทางไกล

The Development of Automatic Parallel Multiple-Choice Question Generation System Using Automatic Item Generation for Distance Education

รัชกฤษ ธนพัฒน์ดล^{1*}Ratchakrit Tanapattanadol^{1*}

(Received: December 2, 2024; Revised: January 8, 2025; Accepted: January 14, 2025)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมายเพื่อ 1) สร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี 2) พัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ และ 3) ตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยแบ่งวิธีดำเนินการวิจัยออกเป็น 3 ระยะ ได้แก่ ระยะที่ 1 การสร้างและประเมินคุณภาพโมเดลข้อสอบ จำนวนทั้งสิ้น 42 โมเดล พร้อมทั้งออกแบบคลังข้อมูลสำหรับจัดเก็บค่าตัวแปรของโมเดลข้อสอบที่สร้างขึ้น เพื่อให้ทุกโมเดลข้อสอบสามารถใช้ข้อมูลที่จัดเก็บร่วมกันได้ ระยะที่ 2 การพัฒนาระบบในรูปแบบเว็บแอปพลิเคชัน ให้สามารถสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติได้ พร้อมทั้งประเมินคุณภาพของระบบที่พัฒนาขึ้น ระยะที่ 3 การตรวจสอบความเป็นคู่ขนานของแบบทดสอบ โดยนำแบบทดสอบต้นแบบและแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น ส่งให้ผู้เชี่ยวชาญประเมินความเป็นคู่ขนานของข้อสอบรายข้อ และนำไปทดลองสอบกับนักศึกษาสาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช ที่ลงทะเบียนชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 30 คน

ผลการวิจัยพบว่า 1) โมเดลข้อสอบชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยีที่สร้างขึ้นมีคุณภาพผ่านเกณฑ์คุณภาพที่กำหนดทุกโมเดล 2) ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติที่พัฒนาขึ้น สามารถสร้างข้อสอบคู่ขนานอัตโนมัติ ได้สอดคล้องกับวัตถุประสงค์ของการพัฒนาระบบ โดยมีคุณภาพของระบบในภาพรวมอยู่ใน ระดับมาก 3) ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบรายข้อโดยผู้เชี่ยวชาญ พบว่าข้อสอบทุกข้อของแบบทดสอบทั้ง 2 ฉบับ สอดคล้องกับเกณฑ์ความเป็นคู่ขนานที่กำหนด และผลการวิเคราะห์ข้อมูลจากการทดลองสอบ พบว่า แบบทดสอบต้นแบบ มีค่าความยากเฉลี่ยทั้งฉบับเท่ากับ 0.52 ค่าอำนาจจำแนกเฉลี่ยทั้งฉบับเท่ากับ 0.41 ค่าความแปรปรวน

¹ อาจารย์ สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช

¹ Lecturer, Office of Registration Records and Evaluation, Sukhothai Thammathirat Open University
ทุนวิจัยการศึกษาทางไกล มหาวิทยาลัยสุโขทัยธรรมาธิราช

Research grants for distance education, Sukhothai Thammathirat Open University

* Corresponding Author E-mail: ratchakrit.tan@stou.ac.th

ทั้งฉบับเท่ากับ 0.24 ค่าเฉลี่ยของคะแนนแบบทดสอบเท่ากับ 21.90 ค่าความเชื่อมั่นเท่ากับ 0.89 และค่าความคลาดเคลื่อนมาตรฐานของการวัดเท่ากับ 2.98 แบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์มีค่าความยากเฉลี่ยทั้งฉบับเท่ากับ 0.51 ค่าอำนาจจำแนกเฉลี่ยทั้งฉบับเท่ากับ 0.40 ค่าความแปรปรวนทั้งฉบับเท่ากับ 0.25 ค่าเฉลี่ยของคะแนนแบบทดสอบเท่ากับ 21.00 ค่าความเชื่อมั่นเท่ากับ 0.90 และค่าความคลาดเคลื่อนมาตรฐานของการวัดเท่ากับ 2.84 เมื่อทำการทดสอบความแตกต่างข้อมูลดังกล่าว ของแบบทดสอบทั้ง 2 ฉบับปรากฏว่าไม่แตกต่างกัน จากผลการตรวจสอบความเป็นคู่ขนานโดยผู้เชี่ยวชาญและผลจากการทดลองสอบจึงสรุปได้ว่า แบบทดสอบต้นแบบและแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์มีความเป็นคู่ขนานกัน

คำสำคัญ โมเดลข้อสอบ การสร้างข้อสอบอัตโนมัติ ข้อสอบคู่ขนาน การศึกษาทางไกล

Abstract

The aims of this research were to: 1) create item models and design a data warehouse for mathematics and statistics for science and technology; 2) develop an automatic parallel multiple-choice question generation system using automatic item generation; and 3) investigate parallelism between the prototype test and computer-generated test. The research methodology was divided into three phases: Phase 1 involved item model creation and evaluation to create 42 models, alongside the design of a data warehouse to store parameter values of the created item models for all item models to share stored data. Phase 2 focused on system development in web application form to create automatic parallel multiple-choice questions, together with quality evaluation of the developed system. Phase 3 consisted of parallelism investigation by having experts evaluate the parallelism of each item between the prototype test and computer-generated test, followed by experimental testing with 30 students enrolled in Mathematics and Statistics for Science and Technology at the School of Science and Technology, Sukhothai Thammathirat Open University.

The results revealed that: 1) All created item models of mathematics and statistics for science and technology passed the required criteria. 2) The automatic parallel multiple-choice question generation system successfully created parallel questions that aligned with the system's design objectives, with the overall system quality at a high level. 3) Expert evaluation confirmed that all items in both tests (the prototype and computer-generated tests) met the parallelism criteria. The experimental results showed that: (1) The prototype test had an average difficulty of 0.52, average discrimination of 0.41, variance of 0.24, mean score of 21.90, reliability of 0.89, and standard error of measurement of 2.98. (2) The computer-generated test had an average difficulty of 0.51, average discrimination of 0.40, variance of 0.25, mean

score of 21.00, reliability of 0.90, and standard error of measurement of 2.84. Statistical analysis indicated no significant differences between the two test versions. Based on both expert evaluation and experimental results, it was concluded that the prototype test and computer-generated test were parallel.

Keywords: item model, automatic item generation, parallel test, distance education

บทนำ

การสอบเป็นวิธีการวัดและประเมินผลการศึกษาที่สามารถใช้ได้กับผู้เรียนในทุกระดับ และเป็นวิธีการวัดและประเมินผลที่ใช้กันอย่างแพร่หลายในสถาบันการศึกษา โดยเฉพาะในระบบการศึกษาทางไกล อย่างมหาวิทยาลัยสุโขทัยธรรมมาธิราช ซึ่งมีรูปแบบการเรียนการสอนผ่านทางไกลเป็นรูปแบบหลัก มีการเผชิญหน้าระหว่างผู้เรียนกับผู้สอนเป็นส่วนน้อย วิธีการวัดและประเมินผลเพื่อตัดสินผลการเรียนรู้ที่สอดคล้องกับรูปแบบการเรียนการสอนดังกล่าว จึงเป็นการสอบโดยการจัดการแบบรวมศูนย์เพื่อให้เป็นมาตรฐานเดียวกันทั้งมหาวิทยาลัย ซึ่งเครื่องมือหลักที่ใช้ในการวัดและประเมินผลคือ แบบทดสอบ ประกอบด้วยข้อสอบแบบเลือกตอบ (5 ตัวเลือก) หรือข้อสอบแบบเลือกตอบผสมแบบเขียนตอบ หรือเป็นข้อสอบแบบเขียนตอบล้วน ทั้งนี้ขึ้นอยู่กับข้อกำหนดของแต่ละสาขา และระดับการศึกษาของผู้สอบ

ระบบคลังข้อสอบมาตรฐานที่ใช้สำหรับการวัดและประเมินผลการเรียนรู้ในระบบการศึกษาทางไกลของมหาวิทยาลัยสุโขทัยธรรมมาธิราช หรือเรียกสั้น ๆ ว่าระบบคลังข้อสอบ นับเป็นกลไกสำคัญในการทำหน้าที่เก็บรวบรวมข้อสอบสำหรับสร้างเป็นแบบทดสอบดังกล่าว เพื่อใช้ในการสอบในรูปแบบต่างๆ ของมหาวิทยาลัย นอกจากนี้ระบบคลังข้อสอบยังมีวัตถุประสงค์เพื่อพัฒนาคุณภาพแบบทดสอบ ลดภาระในการออกข้อสอบใหม่ รวมถึงการลดค่าใช้จ่ายในการสร้างข้อสอบใหม่อีกด้วย นอกจากบทบาทสำคัญในการสร้าง และพัฒนาคุณภาพแบบทดสอบดังกล่าวแล้ว ระบบคลังข้อสอบยังมีข้อควรระวังที่สำคัญประการหนึ่ง ซึ่งอาจกระทบต่อประสิทธิภาพของการวัดประเมินผลการศึกษาได้ คือถ้าหากระบบคลังข้อสอบมีจำนวนข้อสอบไม่มากพอ หรือมีจำนวนไม่สอดคล้องกับความต้องการใช้งานในปัจจุบัน หรืออนาคต อาจส่งผลให้ข้อสอบในคลังข้อสอบมีโอกาสถูกนำมาใช้งานซ้ำบ่อยครั้งขึ้น และการที่ข้อสอบถูกนำไปใช้ในการทดสอบบ่อยครั้งเกินไป ย่อมทำให้เกิดผลกระทบต่อความปลอดภัยของการทดสอบ กล่าวคือเมื่อมีข้อสอบที่ถูกนำไปใช้ในการทดสอบบ่อยครั้งเกินไป จนรู้กันในกลุ่มผู้สอบ เป็นผลทำให้ผู้สอบในรุ่นถัดมาสามารถคาดการณ์ล่วงหน้าได้ว่าตนเองจะได้ข้อสอบข้อใด ทำให้สามารถตอบข้อสอบข้อดังกล่าวได้โดยไม่ต้องใช้ความสามารถของตนเอง ส่งผลให้คะแนนสอบมีความคลาดเคลื่อน ผู้สอบสามารถทำคะแนนได้มากเกินไปความสามารถของตนเอง กระทั่งคุณภาพของการวัดและประเมินผลการศึกษา (Premthongsuk et al., 2017) ดังนั้นการมีข้อสอบที่มีคุณภาพจำนวนมากในระบบคลังข้อสอบ หรือการสร้างข้อสอบใหม่ที่มีคุณภาพเทียบเท่ากับข้อสอบที่อยู่ในระบบคลังข้อสอบได้จำนวนมาก ในระยะเวลารวดเร็ว และทันเวลาต่อความต้องการใช้งาน จึงเป็นสิ่งสำคัญและจำเป็น ในขณะที่การสร้างข้อสอบเพื่อเพิ่มจำนวนข้อสอบลงในระบบคลังข้อสอบ โดยผ่านขั้นตอนการจัดทำคลังข้อสอบที่ใช้อยู่ในปัจจุบันของ

มหาวิทยาลัยนั้น ไม่เอื้อให้สามารถสร้างข้อสอบจำนวนมาก ๆ ในระยะเวลาที่รวดเร็วได้ เนื่องจากต้องใช้ผู้เชี่ยวชาญด้านเนื้อหาในการออกข้อสอบ รวมถึงการดำเนินการในแต่ละขั้นตอนมีรายละเอียดที่ต้องใช้ระยะเวลาค่อนข้างนานในการดำเนินการ รวมทั้งมีผู้เกี่ยวข้องหลายฝ่าย จากข้อจำกัดดังกล่าวย่อมเป็นอุปสรรคต่อการสร้างข้อสอบจำนวนมากที่พร้อมใช้งานได้ในระยะเวลาอันสั้น จากปัญหาข้างต้นหากมีวิธีที่สามารถสร้างข้อสอบแบบอัตโนมัติเพื่อบรรจุไว้ในคลังข้อสอบได้จำนวนมาก รวดเร็ว และมีความเป็นคู่ขนานกับข้อสอบที่อยู่ในระบบคลังข้อสอบ หรือข้อสอบต้นฉบับ ย่อมทำให้ระบบคลังข้อสอบสามารถสร้างแบบทดสอบที่วัดผลการเรียนรู้เดียวกันได้หลากหลาย ส่งผลให้การวัดและประเมินผลการศึกษาของมหาวิทยาลัยมีประสิทธิภาพมากยิ่งขึ้น นอกจากนี้ยังช่วยลดภาระหรือเวลาในการออกข้อสอบใหม่ของผู้เชี่ยวชาญด้านเนื้อหา และลดค่าใช้จ่ายในการสร้างข้อสอบใหม่อีกด้วย

ปัจจุบันมีการพัฒนาเทคโนโลยีคอมพิวเตอร์มาช่วยในการสร้างข้อสอบที่เรียกว่าการสร้างข้อสอบอัตโนมัติ (Automatic Item Generation : AIG) ซึ่งการสร้างข้อสอบอัตโนมัติดังกล่าว เป็นกระบวนการสร้างข้อสอบภายใต้โครงสร้างประกอบของโมเดลข้อสอบที่สร้างขึ้นตามหลักการของ Gierl et al. (2008) ที่ช่วยให้สามารถสร้างข้อสอบได้อย่างรวดเร็ว เมื่อเปรียบเทียบกับวิธีการอื่น (Gierl & Haladyna, 2013) รวมทั้งมีงานวิจัยอีกหลายเรื่องที่เป็นต้นแบบ ในการนำหลักการสร้างข้อสอบอัตโนมัติดังกล่าว ไปใช้ดำเนินการวิจัย เช่น งานวิจัยของ Chantharamaha & Chadcham (2018); Gierl et al. (2012); Gierl & Lai (2012; 2013); Pradujprom & Pantong (2019); Raksombat et al. (2022) เป็นต้น ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะพัฒนาระบบสร้างข้อสอบคู่ขนานอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ ภายใต้หลักการการสร้างโมเดลข้อสอบของ Gierl et al. (2008) เพื่อเป็นอีกหนึ่งแนวทางที่สามารถสร้างข้อสอบคู่ขนานได้จำนวนมาก รวดเร็ว และข้อสอบมีคุณภาพเท่าเทียมกับข้อสอบต้นแบบ (Chantharamaha & Chadcham, 2018) โดยความเท่าเทียมดังกล่าวจะอยู่ในบริบทของความเป็นคู่ขนาน คือข้อสอบมีเนื้อหาเดียวกัน หรือวัดวัดคุณสมบัติและระดับพฤติกรรมเดียวกัน มีความยากง่ายใกล้เคียงกัน เพื่อให้สามารถนำข้อสอบของแต่ละวัดคุณสมบัติการเรียนรู้ของชุดวิชาที่สร้างจากระบบที่พัฒนาขึ้นไปใช้ในการทดสอบรูปแบบต่าง ๆ หรือจัดเก็บเพิ่มเติมไว้ในระบบคลังข้อสอบได้ทันที โดยการพัฒนาระบบสร้างข้อสอบคู่ขนานอัตโนมัติดังกล่าว ผู้วิจัยได้มุ่งเน้นพัฒนาระบบคอมพิวเตอร์ให้มีความสามารถในการสร้างข้อสอบคู่ขนานแบบเลือกตอบโดยเฉพาะ เนื่องจากข้อสอบแบบเลือกตอบดังกล่าว ได้ถูกใช้เป็นเครื่องมือหลักสำหรับวัดและประเมินผลการศึกษาในระบบการศึกษาทางไกลของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ในระดับปริญญาตรีซึ่งเป็นนักศึกษาส่วนใหญ่ของมหาวิทยาลัย

สำหรับการวิจัยเพื่อพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติที่นำเสนอ นั้น จะดำเนินการภายใต้ชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี ซึ่งเป็นชุดวิชาที่ใช้ในการเรียนการสอนในระบบการศึกษาทางไกลของมหาวิทยาลัยสุโขทัยธรรมมาธิราช เนื่องจากเป็นชุดวิชาที่มีเนื้อหาเอื้อต่อการนำข้อสอบมาสร้างโมเดลข้อสอบแล้วคาดว่ามีแนวโน้มที่จะได้โมเดลข้อสอบหลากหลายรูปแบบ มากกว่าการดำเนินการภายใต้ชุดวิชาอื่น ๆ ที่อาจมีเนื้อหาของข้อสอบประกอบด้วยส่วนที่แปรเปลี่ยนได้น้อยกว่า ส่งผลให้ได้โมเดลข้อสอบรูปแบบซ้ำ ๆ ไม่หลากหลาย ดังนั้นการมีรูปแบบของโมเดลข้อสอบที่หลากหลายดังกล่าว ย่อมสามารถใช้เป็นแนวทางในการสร้างโมเดลข้อสอบของชุด

วิชาอื่น ๆ ที่เนื้อหาชุดวิชามีบริบทใกล้เคียงกับกับชุดวิชาคณิตศาสตร์ และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยีดังกล่าวต่อไปในอนาคตได้ นอกจากนี้ข้อสอบที่นำมาใช้เป็นข้อสอบต้นแบบ เพื่อนำไปใช้ในกระบวนการสร้างข้อสอบอัตโนมัติตามวิธีที่นำเสนอดังกล่าว ผู้วิจัยจะดำเนินการสร้างข้อสอบต้นแบบตามหลักการสร้างและตรวจสอบคุณภาพข้อสอบ โดยเนื้อหาข้อสอบต้นแบบที่สร้างเป็นของชุดวิชาดังข้างต้น แทนการใช้ข้อสอบจริงที่อยู่ในคลังข้อสอบ ด้วยเหตุผลเรื่องความปลอดภัยของข้อสอบในระหว่างการทำวิจัย และการเผยแพร่เนื้อหาข้อสอบในรายงานการวิจัย เมื่อดำเนินการวิจัยเสร็จสิ้น ระบบสร้างข้อสอบที่พัฒนาขึ้นจากงานวิจัยนี้ ยังสามารถขยายผลให้กับข้อสอบจริงที่จัดเก็บอยู่ในระบบคลังข้อสอบมาตรฐานของมหาวิทยาลัย เพื่อใช้สำหรับวัดและประเมินผลการศึกษาในระบบการศึกษาทางไกลได้ต่อไป นอกจากนี้สถาบันการศึกษาอื่น ๆ ยังสามารถนำวิธีดำเนินการวิจัย ไปประยุกต์ใช้เป็นแนวทางในการพัฒนาระบบสร้างข้อสอบคู่ขนานอัตโนมัติสำหรับใช้วัดและประเมินผลการศึกษาในสถาบันของตนเองได้เช่นเดียวกัน

ความมุ่งหมายของการวิจัย

1. เพื่อสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี
2. เพื่อพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ
3. เพื่อตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์

วิธีดำเนินการวิจัย

การวิจัยนี้แบ่งวิธีการดำเนินการวิจัยออกเป็น 3 ระยะ ดังนี้

ระยะที่ 1 การสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี

การวิจัยในระยะนี้มีวัตถุประสงค์เพื่อสร้างโมเดลข้อสอบ และออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี มีวิธีดำเนินการวิจัยที่เกี่ยวข้องดังนี้

1. การสร้างข้อสอบต้นแบบเพื่อนำไปสร้างโมเดลข้อสอบมีขั้นตอนดังนี้

1) กำหนดวัตถุประสงค์การเรียนรู้ ระดับพฤติกรรม และเนื้อหา สำหรับใช้สร้างข้อสอบแบบเลือกตอบ 5 ตัวเลือก ซึ่งเนื้อหาดังกล่าวเป็นส่วนหนึ่งของเนื้อหาชุดวิชา 96102 คณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 6 เรื่องหลัก ได้แก่ (1) อนุพันธ์ (2) เซต (3) สมการ (4) อสมการ (5) เรขาคณิตวิเคราะห์ และ (6) เมทริกซ์และดีเทอร์มิแนนต์

2) เขียนลักษณะเฉพาะของข้อสอบ (Item Specification) ของเนื้อหาข้อสอบทุกเรื่องที่ถูกกำหนดไว้สำหรับนำมาใช้สร้างข้อสอบ เรื่องละ 7 ข้อ รวมทั้งสิ้น 42 ข้อ ซึ่งรายละเอียดของลักษณะเฉพาะของข้อสอบประกอบด้วย ชื่อเนื้อหา วัตถุประสงค์ ระดับพฤติกรรม ลักษณะคำถาม ลักษณะตัวถูก และตัวอย่างข้อสอบ โดยไม่เขียนลักษณะตัวถูก

3) สร้างข้อสอบเฉพาะคำถาม และตัวถูก ตามลักษณะเฉพาะของข้อสอบที่กำหนด พร้อมทั้งนำข้อสอบเฉพาะคำถาม และคำตอบ ที่สร้างขึ้นไปสัมภาษณ์ผู้สอนวิชาคณิตศาสตร์ จำนวน 10 คน จากหลากหลายสถาบัน โดยสัมภาษณ์ในประเด็นเกี่ยวกับความเข้าใจคลาดเคลื่อน หรือความผิดพลาดของผู้เรียนในการทำข้อสอบจากประสบการณ์ที่ผ่านมาของผู้สอน เพื่อนำไปใช้เป็นข้อมูลส่วนหนึ่งในการสร้างตัวถูก

4) เขียนลักษณะเฉพาะของข้อสอบ (item specification) ของเนื้อหาข้อสอบแต่ละเรื่องที่ถูกกำหนดไว้สำหรับนำมาใช้สร้างข้อสอบให้สมบูรณ์ ซึ่งประกอบด้วย ชื่อเนื้อหา วัตถุประสงค์ ระดับพฤติกรรม ลักษณะคำถาม ลักษณะตัวถูก ลักษณะตัวถูก และตัวอย่างข้อสอบ พร้อมทั้งสร้างข้อสอบตามลักษณะเฉพาะของข้อสอบ (item specification) จำนวน 60 ข้อ และนำข้อสอบที่สร้างขึ้นไปสัมภาษณ์ผู้สอนวิชาคณิตศาสตร์ จำนวน 10 คน อีกครั้ง ในประเด็นความเหมาะสมของ คำถาม ตัวถูก และตัวถูกที่สร้างขึ้น พร้อมทั้งปรับข้อสอบตามข้อเสนอแนะจากการสัมภาษณ์

5) ตรวจสอบคุณภาพด้านความตรงเชิงเนื้อหา (content validity) ของข้อสอบที่สร้างขึ้น โดยผู้เชี่ยวชาญทางด้านคณิตศาสตร์ จำนวน 3 คน เพื่อพิจารณาความสอดคล้องของข้อสอบกับวัตถุประสงค์การเรียนรู้ และความสอดคล้องของข้อสอบกับลักษณะเฉพาะของข้อสอบ จากนั้นนำไปทดลองใช้กับกลุ่มตัวอย่าง ได้แก่ นักศึกษามหาวิทยาลัยสุโขทัยธรรมาธิราช ที่ลงทะเบียนชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 30 คน และทำการวิเคราะห์คุณภาพข้อสอบรายข้อ และรายฉบับ พร้อมทั้งคัดเลือกข้อสอบที่มีคุณภาพจำนวน 42 ข้อ เพื่อนำไปใช้เป็นข้อสอบต้นแบบสำหรับสร้างโมเดลข้อสอบต่อไป

2. การสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยีมีขั้นตอนการดำเนินการดังนี้

1) นำข้อสอบต้นแบบและลักษณะเฉพาะของข้อสอบที่สร้างขึ้น จำนวน 42 ข้อ ไปใช้เป็นข้อมูลสำหรับสร้างโมเดลข้อสอบ โดยเลือกรูปแบบโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) ที่ต้องการสร้าง โดยวิเคราะห์เนื้อหาของข้อสอบต้นแบบ และลักษณะเฉพาะของข้อสอบแต่ละข้อว่า มีความสอดคล้องกับประเภทของโมเดลข้อสอบโมเดลใดใน 10 โมเดล (โมเดลที่ 1 – โมเดลที่ 10) ดังตาราง 1 จากนั้นจึงดำเนินการสร้างโมเดลข้อสอบตามโครงสร้างของโมเดลนั้น ๆ

ตาราง 1 การจัดหมวดหมู่โมเดลข้อสอบ (item model taxonomy)

ตัวเลือก (Options)	โจทย์ (Stem)			
	อิสระ (independent)	ไม่อิสระ (dependent)	ผสม (mixed)	คงที่ (fixed)
รูปแบบการสุ่ม (randomly selected)	โมเดลที่ 1	โมเดลที่ 4	โมเดลที่ 7	โมเดลที่ 10
รูปแบบจำกัด (constrained)	โมเดลที่ 2	โมเดลที่ 5	โมเดลที่ 8	N/A
รูปแบบคงที่ (fixed)	โมเดลที่ 3	โมเดลที่ 6	โมเดลที่ 9	N/A

จากตาราง 1 โจทย์ของโมเดลข้อสอบ แบ่งออกเป็น 4 รูปแบบ ได้แก่ (1) องค์ประกอบที่แปรเปลี่ยนได้อย่างอิสระ (Independent Element) เมื่อองค์ประกอบใดในโจทย์มีการเปลี่ยนแปลงจะไม่ส่งผลกระทบต่อองค์ประกอบอื่น ๆ (2) องค์ประกอบที่แปรเปลี่ยนได้อย่างไม่อิสระ (dependent element) เมื่อองค์ประกอบใด

องค์ประกอบหนึ่งในโจทย์มีการเปลี่ยนแปลงจะส่งผลให้องค์ประกอบอื่นๆ ที่สัมพันธ์กันเปลี่ยนแปลงตามไปด้วย (3) การผสมกันระหว่างองค์ประกอบที่แปรเปลี่ยนได้อย่างอิสระ และแปรเปลี่ยนได้อย่างไม่อิสระ (mixed independent/dependent element) โจทย์ประกอบไปด้วยองค์ประกอบที่แปรเปลี่ยนได้อย่างอิสระ ซึ่งเมื่อองค์ประกอบใดในโจทย์มีการเปลี่ยนแปลงจะไม่ส่งผลกระทบต่อองค์ประกอบอื่น ๆ แต่ในขณะเดียวกัน โจทย์ดังกล่าวก็ยังมีองค์ประกอบที่แปรเปลี่ยนได้อย่างไม่อิสระ เมื่อองค์ประกอบใดขององค์ประกอบหนึ่งในโจทย์ มีการเปลี่ยนแปลงจะส่งผลให้องค์ประกอบอื่นๆ ที่สัมพันธ์กันเปลี่ยนแปลงตามไปด้วย และ (4) องค์ประกอบคงที่ (fixed element) ลักษณะของโจทย์คงที่ไม่มีการเปลี่ยนแปลง ในขณะที่ตัวเลือกของโมเดลข้อสอบ แบ่งออกเป็น 3 รูปแบบ ได้แก่ (1) ตัวเลือกที่ได้จากการสุ่ม (randomly selected options) คือรูปแบบที่หาคำตอบที่ถูกต้องและตัวลวงถูกเลือกแบบสุ่มจากขอบเขตข้อมูลที่มีเนื้อหาสอดคล้องกัน (2) ตัวเลือกแบบจำกัด (constrained options) คือรูปแบบที่ทั้งคำตอบที่ถูกต้องและตัวลวงถูกสร้างตามข้อกำหนด เช่น สูตร หรือรูปแบบการคำนวณ เป็นต้น และ (3) ตัวเลือกแบบคงที่ (fixed options) คือรูปแบบที่ทั้งคำตอบที่ถูกต้องและตัวลวงไม่มีการเปลี่ยนแปลง

2) สร้างโมเดลข้อสอบโดยการเชื่อมโยงเนื้อหาข้อสอบต้นแบบเข้ากับองค์ประกอบทั้ง 5 ส่วนของโมเดลข้อสอบ แต่ละประเภทที่เลือกไว้ พร้อมทั้งกำหนดรายละเอียดเนื้อหาของโมเดลข้อสอบทั้ง 5 ส่วน ดังนี้ (1) โจทย์ (stem): นำเนื้อหาโจทย์ของข้อสอบต้นแบบมาวิเคราะห์เพื่อกำหนดส่วนคงที่ และส่วนที่สามารถเปลี่ยนค่าได้ (2) ส่วนประกอบ (elements): กำหนดรายละเอียดเนื้อหาโจทย์ของโมเดลข้อสอบ ในส่วนที่สามารถเปลี่ยนแปลงค่าได้ 3 ส่วน ได้แก่ 1. ชื่อตัวแปร และชนิดของตัวแปร 2. ค่าตัวแปร ขอบเขตของค่าตัวแปร และค่ายกเว้นตัวแปร 3. ค่าเงื่อนไขระหว่างตัวแปร (3) ตัวเลือก (options): กำหนดรายละเอียดเนื้อหาตัวเลือกทั้ง 5 ตัวเลือก ของแต่ละโมเดล ซึ่งตัวเลือกดังกล่าวจะมีเนื้อหาในรูปแบบใดขึ้นอยู่กับเนื้อหาของข้อสอบต้นแบบ และประเภทของโมเดลข้อสอบที่สอดคล้องกับข้อสอบต้นแบบที่เลือกไว้ (4) ข้อมูลเสริม (auxiliary information): ข้อมูลเสริมสามารถกำหนดได้ที่ตำแหน่งโจทย์และตัวเลือกของโมเดลข้อสอบ ซึ่งแต่ละโมเดลข้อสอบจะมีข้อมูลเสริมหรือไม่ หรือมีในรูปแบบใดขึ้นอยู่กับเนื้อหาของข้อสอบต้นแบบ โดยข้อมูลเสริมดังกล่าว ได้แก่ ภาพ และตาราง และ (5) เฉลย (key): การกำหนดคำตอบที่ถูกต้องของข้อสอบข้อนั้น ๆ ตามที่โจทย์กำหนด

3) ตรวจสอบคุณภาพโมเดลข้อสอบที่สร้างขึ้น โดยดำเนินการดังนี้ (1) สร้างแบบประเมินคุณภาพโมเดลข้อสอบ โดยประยุกต์ใช้ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ (Index of Item Objective Congruence : IOC) มาเป็นเกณฑ์การประเมินคุณภาพโมเดลข้อสอบของแต่ละรายการประเมิน โดยคะแนน -1 หมายถึง ไม่สอดคล้อง คะแนน 0 หมายถึง ไม่แน่ใจ และคะแนน 1 หมายถึง สอดคล้อง พร้อมทั้งนำแบบประเมินคุณภาพโมเดลข้อสอบ ส่งให้ผู้เชี่ยวชาญทางด้านการวัดและประเมินผลจำนวน 3 คน ตรวจสอบคุณภาพของแบบประเมินก่อนนำไปใช้ประเมินจริง (2) ประเมินคุณภาพของโมเดลข้อสอบ โดยให้ผู้เชี่ยวชาญซึ่งประกอบด้วย 1) ประธานชุดวิชาหรืออาจารย์ผู้สอนที่เป็นผู้แทนคณะกรรมการบริหารชุดวิชา คณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 1 คน 2) ผู้เชี่ยวชาญทางด้านคณิตศาสตร์ จำนวน 1 คน และ 3) ผู้เชี่ยวชาญทางด้านการวัดและประเมินผล ซึ่งนอกจากจะมีคุณวุฒิทางด้านการวัดและประเมินผลแล้ว ยังต้องมีคุณวุฒิด้านคณิตศาสตร์ด้วย จำนวน 1 คน รวมทั้งหมด 3 คน ซึ่งเป็นจำนวน

ผู้เชี่ยวชาญที่เหมาะสมในการประเมินความสอดคล้องตามคำแนะนำของ Srisa-ard (1995) ทำการตรวจสอบคุณภาพของโมเดลข้อสอบพร้อมทั้งทำแบบประเมินคุณภาพโมเดลข้อสอบ ซึ่งรายการประเมินคุณภาพโมเดลข้อสอบ มีประเด็นพิจารณาทั้งสิ้น 5 ประเด็น ได้แก่ ประเด็นที่ 1 ความสอดคล้องระหว่างประเภทของโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) ที่กำหนดกับโมเดลข้อสอบที่สร้างขึ้น ประเด็นที่ 2 ความสอดคล้องระหว่างเนื้อหาข้อสอบต้นแบบกับเนื้อหาของโมเดลข้อสอบที่สร้างขึ้น ประเด็นที่ 3 ความสอดคล้องระหว่างวัตถุประสงค์การเรียนรู้ที่นำมาสร้างข้อสอบต้นแบบกับโมเดลข้อสอบที่สร้างขึ้น ประเด็นที่ 4 ความสอดคล้องระหว่างระดับพฤติกรรมการเรียนรู้ที่นำมาสร้างข้อสอบต้นแบบกับโมเดลข้อสอบที่สร้างขึ้น และประเด็นที่ 5 ความสอดคล้องระหว่างความยากของข้อสอบต้นแบบกับความยากของโมเดลข้อสอบที่สร้างขึ้น และ (3) วิเคราะห์ผลการประเมินคุณภาพโมเดลข้อสอบจากผู้เชี่ยวชาญ โดยการคำนวณค่าดัชนีความสอดคล้อง ซึ่งโมเดลข้อสอบทุกข้อหรือทุกโมเดล ต้องมีค่าดัชนีความสอดคล้องของทุกรายการประเมินเป็นไปตามเกณฑ์ที่กำหนดคือ ตั้งแต่ 0.5 ขึ้นไป พร้อมทั้งสรุปผลการประเมินคุณภาพโมเดลข้อสอบ

3. การออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี

เป็นการออกแบบพื้นที่จัดเก็บข้อมูลกลาง สำหรับนำไปสร้างข้อสอบคู่ขนานของชุดวิชาดังกล่าวผ่านโครงสร้างของโมเดลข้อสอบ ซึ่งคลังข้อมูลจะถูกออกแบบให้เป็นองค์ประกอบหนึ่งของการสร้างโมเดลข้อสอบ ผ่านระบบคอมพิวเตอร์ที่พัฒนาขึ้น โดยคลังข้อมูลจะเก็บข้อมูลในรูปแบบของค่าตัวแปรชนิดตัวอักษร (character) หรือตัวแปรชนิดข้อความ (string) ที่นำไปใช้สร้างโมเดลข้อสอบ ไว้ร่วมกันที่ส่วนกลาง เพื่อบริการให้ทุกโมเดลข้อสอบสามารถใช้ข้อมูลที่ถูกจัดเก็บไว้ในคลังข้อมูลดังกล่าวร่วมกันได้ โดยที่ผู้สร้างโมเดลข้อสอบไม่ต้องเสียเวลาในการสร้างค่าตัวแปรของข้อมูลใหม่ทุกครั้งให้กับโมเดลข้อสอบที่กำลังสร้างขึ้น หากตัวแปรใดถูกสร้างไว้ในคลังข้อมูลแล้ว ผู้ใช้งานก็สามารถนำไปใช้เป็นองค์ประกอบของโมเดลข้อสอบนั้น ๆ ได้ทันที เพื่อให้คลังข้อมูลมีคุณสมบัติดังกล่าวข้างต้น การออกแบบคลังข้อมูลสามารถดำเนินการได้ดังนี้ 1) ศึกษาองค์ประกอบข้อมูลและรูปแบบการจัดเก็บข้อมูลในคลังข้อมูลข้อสอบ 2) ศึกษาวิธีการเชื่อมโยงระหว่างข้อมูลในคลังข้อมูลกับข้อมูลตัวแปรสำหรับใช้สร้างโมเดลข้อสอบ 3) ออกแบบองค์ประกอบของคลังข้อมูล 4) ออกแบบรูปแบบการจัดเก็บ การลบ การแก้ไข ข้อมูลในคลังข้อมูล 5) ออกแบบขั้นตอนการสร้างคลังข้อมูลผ่านระบบสร้างข้อสอบคู่ขนานที่พัฒนาขึ้น และ 6) ออกแบบขั้นตอนการนำข้อมูลจากคลังข้อมูลไปใช้งาน อย่างไรก็ตาม การออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยีดังกล่าวเป็นเพียงการออกแบบเพื่อบริการการสร้างโมเดลข้อสอบในระบบสร้างข้อสอบคู่ขนานอัตโนมัติให้มีความสะดวกในการสร้างโมเดลข้อสอบมากยิ่งขึ้นเท่านั้น ซึ่งผู้ใช้งานอาจจะใช้งานข้อมูลที่มีในคลังข้อมูลหรือไม่ก็ได้ ทั้งนี้ขึ้นอยู่กับความประสงค์ และบริบทของเนื้อหาข้อสอบที่นำมาสร้างโมเดลข้อสอบนั้น ๆ ด้วย

ระยะที่ 2 การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ

การวิจัยในระยะนี้มีวัตถุประสงค์เพื่อพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ ซึ่งระบบดังกล่าวจะถูกพัฒนาขึ้นสำหรับใช้สร้างข้อสอบคู่ขนานแบบเลือกตอบ

อัตโนมัติ ภายใต้หลักการสร้างข้อสอบอัตโนมัติและการสร้างโมเดลข้อสอบของ Gierl et al. (2008) ทั้ง 10 ประเภท และนำโมเดลข้อสอบที่สร้างในการวิจัยระยะที่ 1 จำนวน 42 ข้อ รวมทั้งคลังข้อมูลข้อสอบที่ได้ ออกแบบไว้ มาเป็นข้อมูลในการพัฒนาและทดสอบระบบไปด้วย ซึ่งมีวิธีดำเนินการวิจัยดังนี้

1. วิเคราะห์และออกแบบระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ

การวิเคราะห์และออกแบบระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ มีขั้นตอนการดำเนินการดังนี้ 1) วิเคราะห์และออกแบบระบบ ด้วยการศึกษางานวิจัย และเทคโนโลยีที่เกี่ยวข้อง เพื่อให้ระบบที่พัฒนาขึ้นสามารถรองรับการทำงานหลักได้แก่ (1) สามารถสร้างโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) ได้ครบทั้ง 10 ประเภท (2) สามารถสร้างคลังข้อมูลข้อสอบเพื่อใช้จัดเก็บข้อมูลค่าตัวแปรต่าง ๆ ของโมเดลข้อสอบได้ และ (3) สามารถสร้างข้อสอบคู่ขนานอัตโนมัติจากโมเดลข้อสอบและเงื่อนไขที่กำหนดได้ 2) ออกแบบระบบฐานข้อมูล เพื่อรองรับหรือจัดเก็บข้อมูลที่เกี่ยวข้องกับระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ และ 3) ออกแบบส่วนติดต่อกับผู้ใช้งาน เพื่อให้ผู้ใช้งานสามารถโต้ตอบหรือสื่อสารกับระบบที่พัฒนาขึ้นได้อย่างสะดวก

2. พัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ

การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ มีขั้นตอนการดำเนินการที่เกี่ยวข้องดังนี้ 1) พัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติในรูปแบบเว็บแอปพลิเคชัน (web application) ให้สอดคล้องกับผลการวิเคราะห์และออกแบบระบบ 2) ทดสอบและแก้ไขข้อผิดพลาดทุกฟังก์ชันการทำงานของระบบที่พัฒนาขึ้นโดยใช้เทคนิคการทดสอบแบบกล่องดำ (black box testing) และ 3) จัดทำคู่มือการใช้งาน

3. ตรวจสอบคุณภาพของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติที่พัฒนาขึ้น

การตรวจสอบคุณภาพของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติที่พัฒนาขึ้น มีขั้นตอนการดำเนินการดังนี้ 1) สร้างแบบประเมินคุณภาพระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ ที่มีรูปแบบเป็นมาตรประเมินค่า (rating scale) 5 ระดับ 2) นำแบบประเมินคุณภาพระบบที่สร้างขึ้น ส่งให้ผู้เชี่ยวชาญทางด้านการวัดและประเมินผลจำนวน 3 คน ตรวจสอบคุณภาพของแบบประเมินก่อนนำไปใช้ประเมินจริง 3) ประเมินคุณภาพระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น โดยให้ผู้เชี่ยวชาญจำนวน 5 คน ประกอบด้วย (1) ผู้เชี่ยวชาญด้านคอมพิวเตอร์ จำนวน 3 คน (2) ผู้เชี่ยวชาญด้านการวัดและประเมินผล จำนวน 1 คน และ (3) ประธานชุดวิชาหรืออาจารย์ผู้สอนที่เป็นผู้แทนคณะกรรมการบริหารชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 1 คน ทำการทดสอบการใช้งานระบบที่พัฒนาขึ้น พร้อมทั้งทำแบบประเมินคุณภาพระบบ ซึ่งมีประเด็นพิจารณาทั้งสิ้น 4 ด้าน ได้แก่ (1) ด้านการออกแบบระบบ (2) ด้านการทำงานของระบบ (3) ด้านการใช้งานระบบ และ (4) ด้านคู่มือการใช้งานระบบ 4) วิเคราะห์ผลการประเมินคุณภาพระบบ จากผู้เชี่ยวชาญโดยการคำนวณค่าเฉลี่ย (mean) และความเบี่ยงเบนมาตรฐาน (standard deviation) จากนั้นนำค่าเฉลี่ยดังกล่าว ไปเทียบ

กับเกณฑ์ระดับคุณภาพว่ารายการประเมินนั้นมีคุณภาพอยู่ที่ระดับใด ซึ่งระบบที่พัฒนาขึ้นต้องมีผลการประเมินคุณภาพ ตั้งแต่ ระดับมากขึ้นไป ทุกรายการประเมินจึงจะถือว่าผ่านเกณฑ์การประเมิน 5) สรุปผลการประเมินคุณภาพระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้นพร้อมทั้งแก้ไขตามข้อเสนอแนะของผู้เชี่ยวชาญ และ 6) นำระบบไปให้ผู้ใช้งานที่เกี่ยวข้องทดลองใช้งาน

ระยะที่ 3 การตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์

การวิจัยในระยะนี้มีวัตถุประสงค์เพื่อตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบจำนวน 2 ฉบับ ได้แก่ (1) แบบทดสอบต้นแบบ และ (2) แบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ (ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่ผู้วิจัยพัฒนาขึ้น) ซึ่งมีวิธีดำเนินการวิจัยดังนี้

1. สร้างแบบทดสอบด้วยระบบคอมพิวเตอร์ที่พัฒนาขึ้น

การสร้างแบบทดสอบด้วยระบบคอมพิวเตอร์ที่พัฒนาขึ้น ผู้วิจัยดำเนินการโดยกำหนดให้ระบบคอมพิวเตอร์สร้างข้อสอบจำนวน 42 ข้อ ตามรูปแบบของโมเดลข้อสอบที่ได้สร้างไว้แล้วจากข้อสอบต้นแบบทุกข้อในการวิจัยระยะที่ 1 โดยกำหนดเงื่อนไขให้โมเดลข้อสอบแต่ละข้อ สามารถนำไปสร้างข้อสอบได้จำนวน 1 ข้อต่อหนึ่งโมเดลข้อสอบ เพื่อใช้เป็นตัวแทนของข้อสอบทั้งหมดที่โมเดลข้อสอบนั้น ๆ สามารถสร้างได้ ซึ่งตัวแทนข้อสอบของแต่ละโมเดลข้อสอบดังกล่าว ระบบจะสร้างด้วยวิธีการสุ่มค่าตัวแปรต่าง ๆ ภายใต้ขอบเขตตัวแปรที่กำหนดไว้ในแต่ละโมเดล จากนั้นนำข้อสอบทั้ง 42 ข้อ มาจัดฉบับเป็นแบบทดสอบเพื่อนำไปตรวจสอบความเป็นคู่ขนานเทียบกับข้อสอบต้นแบบต่อไป

2. ตรวจสอบความเป็นคู่ขนานของข้อสอบรายข้อระหว่างข้อสอบจากแบบทดสอบต้นแบบกับข้อสอบจากแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยผู้เชี่ยวชาญ

การตรวจสอบความเป็นคู่ขนานระหว่างข้อสอบจากแบบทดสอบต้นแบบ (ฉบับ A) กับข้อสอบจากแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น (ฉบับ B) ผู้วิจัยดำเนินการในรูปแบบให้ผู้เชี่ยวชาญประเมินความเป็นคู่ขนานของข้อสอบเป็นรายข้อ จากแบบทดสอบทั้ง 2 ฉบับดังกล่าว โดยมีขั้นตอนการดำเนินการดังนี้ 1) สร้างแบบประเมินความเป็นคู่ขนานของข้อสอบ โดยประยุกต์ใช้ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ มาเป็นเกณฑ์การประเมินคุณภาพโมเดลข้อสอบของแต่ละรายการประเมิน พร้อมทั้งนำแบบประเมินความเป็นคู่ขนานของข้อสอบ ส่งให้ผู้เชี่ยวชาญทางด้านการวัดและประเมินผลจำนวน 3 คน ตรวจสอบคุณภาพของแบบประเมินก่อนนำไปใช้ประเมินจริง 2) ประเมินความเป็นคู่ขนานของข้อสอบเป็นรายข้อ โดยให้ผู้เชี่ยวชาญกลุ่มเดียวกันกับที่ประเมินคุณภาพโมเดลข้อสอบ ทำการตรวจสอบความเป็นคู่ขนานของแบบทดสอบทั้ง 2 ฉบับ พร้อมทั้งทำแบบประเมินความเป็นคู่ขนานของข้อสอบ ซึ่งมีประเด็นพิจารณาทั้งสิ้น 4 ประเด็น ได้แก่ (1) ความสอดคล้องของเนื้อหาข้อสอบ (2) ความสอดคล้องของวัตถุประสงค์การเรียนรู้ (3) ความสอดคล้องของระดับพฤติกรรมการเรียนรู้ และ (4) ความสอดคล้องของความยาก และ 3) วิเคราะห์ผลการประเมินความเป็นคู่ขนานของข้อสอบจากผู้เชี่ยวชาญ โดยการคำนวณค่าดัชนีความสอดคล้อง ซึ่งข้อสอบที่มีความเป็นคู่ขนานกัน ต้องมีค่าดัชนีความสอดคล้องของทุกรายการประเมิน เป็นไปตามเกณฑ์ที่กำหนดคือ ตั้งแต่ 0.5 ขึ้นไป และสรุปผลการประเมินความเป็นคู่ขนานของข้อสอบ

3. ตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น โดยการทดลองสอบ

การตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น โดยการทดลองสอบ มีขั้นตอนการดำเนินการดังนี้ 1) นำแบบทดสอบต้นแบบ จำนวน 1 ฉบับ และแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น จำนวน 1 ฉบับ ให้นักศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช ที่ลงทะเบียนชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 30 คน ทำการทดสอบ โดยใช้เวลาทดสอบฉบับละ 2 ชั่วโมง 2) ตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น โดยใช้ผลจากการทำแบบทดสอบทั้ง 2 ฉบับ ของนักศึกษา มาดำเนินการตรวจสอบความเป็นคู่ขนานของแบบทดสอบทั้ง 2 ฉบับ ดังกล่าว ในด้านต่าง ๆ ดังนี้ (1) ความแตกต่างของค่าความยากรายข้อและค่าความยากเฉลี่ยรายฉบับ ทดสอบด้วยสถิติ Dependent t-test (2) ความแตกต่างของค่าอำนาจจำแนกเฉลี่ยรายฉบับ ทดสอบด้วยสถิติ Dependent t-test (3) ความแตกต่างของค่าความแปรปรวนรายข้อและรายฉบับ ทดสอบด้วยสถิติ F-test (4) ความแตกต่างของค่าเฉลี่ยของคะแนนแบบทดสอบ ทดสอบด้วยสถิติ dependent t-test (5) ค่าความเชื่อมั่นแบบสมมูลของแบบทดสอบ โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนแบบทดสอบ (6) ค่าความเชื่อมั่นของแบบทดสอบ โดยใช้สูตรของ Kuder-Richardson 20 (KR_{20}) และ (7) ค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบ และ 3) สรุปผลการตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น

ผลการวิจัย

1. ผลการสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์ฯ

ผู้วิจัยแบ่งผลการวิจัยออกเป็น 3 ส่วนได้แก่ (1) ผลการสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ จำนวน 42 ข้อ (2) ผลการประเมินคุณภาพโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ และ (3) ผลการออกแบบคลังข้อมูลข้อสอบชุดวิชาคณิตศาสตร์ฯ โดยแต่ละส่วนมีรายละเอียดดังนี้

1.1 ผลการสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ แสดงตัวอย่างดังตาราง 2

ตาราง 2 ตัวอย่างผลการสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ เรื่อง อนุพันธ์

ข้อสอบต้นแบบ: ข้อที่ 1	
อนุพันธ์ของฟังก์ชัน $f(x) = 2x^5 + 5x^3 - 8$ มีค่าเท่ากับข้อใด	
1. $10x^4 + 15x^2$	4. $10x^4 + 15x^2 - 8$
2. $10x^5 + 15x^3$	5. $10x^6 + 15x^4 - 8$
3. $10x^6 + 15x^4$	
เฉลย : 1	
โมเดลข้อสอบ : ข้อที่ 1	

ตาราง 2 (ต่อ)

ประเภทโมเดลข้อสอบ : โมเดลที่ 5

โจทย์ : อนุพันธ์ของฟังก์ชัน $f(x) = I_1x^{I_4} + I_2x^{I_5} + I_3$ มีค่าเท่ากับข้อใด

ส่วนประกอบ :

I_1, I_2, I_3 : มีขอบเขตข้อมูลเป็นจำนวนเต็มในช่วง -9 ถึง 9 และ $I_1, I_2 \neq 0, 1, -1$ และ $I_1 \neq I_2 \neq I_3$

I_4, I_5 : มีขอบเขตข้อมูลเป็นจำนวนเต็มบวกในช่วง 2 ถึง 9 และ $I_4 > I_5$

ตัวเลือก :

- | | |
|--|--|
| 1. $(I_1I_4)x^{(I_4-1)} + (I_2I_5)x^{(I_5-1)}$ | 4. $(I_1I_4)x^{(I_4-1)} + (I_2I_5)x^{(I_5-1)} + I_3$ |
| 2. $(I_1I_4)x^{I_4} + (I_2I_5)x^{I_5}$ | 5. $(I_1I_4)x^{(I_4+1)} + (I_2I_5)x^{(I_5+1)} + I_3$ |
| 3. $(I_1I_4)x^{(I_4+1)} + (I_2I_5)x^{(I_5+1)}$ | |

ข้อมูลเสริม : ไม่มี

เฉลย : 1

1.2 ผลการตรวจสอบคุณภาพโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ โดยผู้เชี่ยวชาญดังกล่าว พบว่า โมเดลข้อสอบที่สร้างขึ้นจำนวน 42 ข้อ ทุกข้อผ่านเกณฑ์คุณภาพที่กำหนดไว้ คือโมเดลข้อสอบทุกข้อหรือทุกโมเดล ต้องมีค่าดัชนีความสอดคล้องของทุกรายการประเมิน เป็นไปตามเกณฑ์ที่กำหนดคือ ตั้งแต่ 0.5 ขึ้นไป

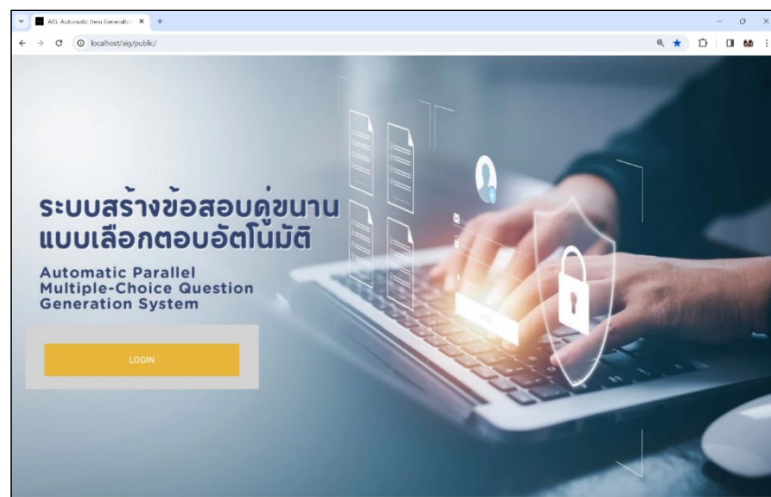
1.3 ผลการออกแบบคลังข้อสอบชุดวิชาคณิตศาสตร์ฯ มีดังนี้ 1) องค์ประกอบของคลังข้อสอบที่ออกแบบขึ้น ประกอบด้วยองค์ประกอบของข้อมูล แบ่งเป็น 2 กลุ่ม ได้แก่ หมวดข้อมูล และรายการข้อมูลที่อยู่ภายใต้หมวดข้อมูลที่สร้างขึ้น โดยข้อมูลที่จัดเก็บไว้ในคลังข้อสอบต้องเป็นข้อมูลชนิดตัวอักษร หรือข้อมูลชนิดข้อความเท่านั้น 2) ขั้นตอนการสร้างคลังข้อสอบผ่านระบบสร้างข้อสอบคู่ขนานที่พัฒนาขึ้น ดำเนินการได้ดังนี้ (1) กำหนดข้อมูลที่ต้องการจัดเก็บในคลังข้อสอบ (2) เลือกเมนูหรือฟังก์ชันจัดการคลังข้อสอบ (3) ระบบแสดงหน้าจอให้กรอกข้อมูล ชื่อหมวดข้อมูล และรายการข้อมูลภายใต้หมวดข้อมูลแต่ละหมวด (4) กรอกชื่อหมวดข้อมูล พร้อมทั้งกดปุ่มเพิ่มหมวดข้อมูล จากนั้นให้กรอกรายการข้อมูลภายใต้หมวดข้อมูลที่สร้างขึ้นแต่ละหมวด พร้อมทั้งกดปุ่มเพิ่มรายการข้อมูล เพื่อบันทึกข้อมูลคลังข้อสอบ (5) ระบบแสดงรายการข้อมูลที่ถูกจัดเก็บไว้ในคลังข้อสอบ (6) กรณีที่ต้องการ ลบ/แก้ไขข้อมูลคลังข้อสอบ สามารถกดปุ่ม ลบ/แก้ไขได้ที่ส่วนแสดงข้อมูลรายการข้อมูลในคลังข้อสอบที่ได้บันทึกไว้ในระบบ และ 3) ผลการออกแบบขั้นตอนการนำข้อมูลจากคลังข้อสอบไปใช้งาน เป็นดังนี้คือ เมื่อผู้ใช้งานมีการกำหนดค่าตัวแปรเป็นชนิดตัวอักษร หรือกำหนดค่าตัวแปรเป็นชนิดข้อความ ให้กับตัวแปรใดตัวแปรหนึ่งของโมเดลข้อสอบ ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น จะต้องแสดงข้อมูลในคลังข้อสอบให้ผู้ใช้งานพิจารณาเลือกนำไปใช้ กำหนดค่าตัวแปรให้กับโมเดลข้อสอบทุกครั้ง จากผลการออกแบบคลังข้อสอบชุดวิชาคณิตศาสตร์ฯ ดังกล่าวข้างต้น ได้ถูกนำไปใช้เป็นข้อมูลสำหรับพัฒนาค้นคว้าคลังข้อสอบขึ้นใช้งานจริง เพื่อเป็นฟังก์ชันการทำงาน ฟังก์ชันหนึ่งของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น

2. ผลการพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ

ผู้วิจัยแบ่งการนำเสนอผลการวิจัยออกเป็น 3 ส่วน ได้แก่ (1) ผลการออกแบบระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ (2) ผลการพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ และ (3) ผลการตรวจสอบคุณภาพของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ มีรายละเอียดดังนี้

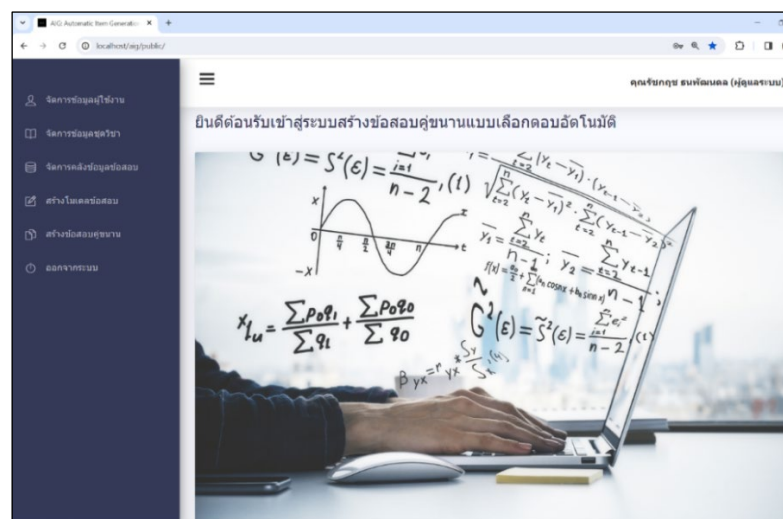
2.1 ผลการออกแบบระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ มีรายละเอียด ประกอบด้วย ฟังก์ชันการทำงานทั้งหมดจำนวน 6 ฟังก์ชัน ได้แก่ (1) ฟังก์ชันการเข้าสู่ระบบ (2) ฟังก์ชันจัดการข้อมูลผู้ใช้งาน (3) ฟังก์ชันจัดการข้อมูลชุดวิชา (4) ฟังก์ชันจัดการคลังข้อมูลข้อสอบ (5) ฟังก์ชันสร้างโมเดลข้อสอบ และ (6) ฟังก์ชันสร้างข้อสอบคู่ขนาน

2.2 ผลการพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ แสดงตัวอย่างหน้าจอของบางฟังก์ชันการทำงาน ได้ดังภาพ 1-4



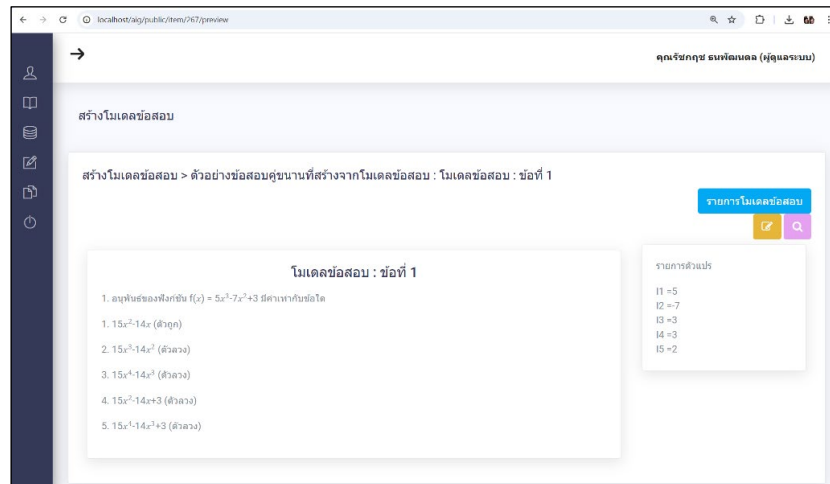
ภาพ 1 หน้าจอแรกเมื่อผู้ใช้งานเรียกใช้งานระบบ

จากภาพ 1 หน้าจอแรกเมื่อผู้ใช้งานเรียกใช้งานระบบ ประกอบด้วย (1) ชื่อระบบ และ (2) ปุ่ม Login เพื่อเชื่อมโยงไปยังหน้าจอเข้าสู่ระบบ สำหรับให้ผู้ใช้งานกรอกรหัสผู้ใช้และรหัสผ่านเพื่อเข้าสู่ระบบ



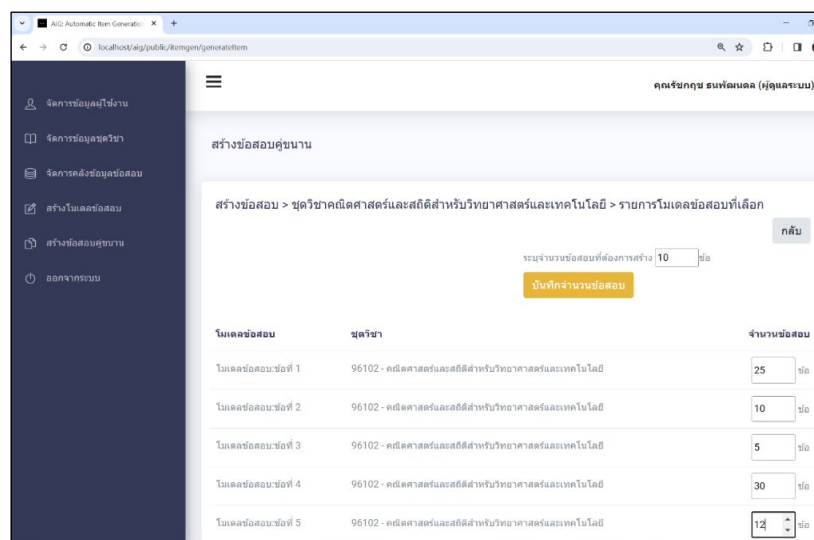
ภาพ 2 หน้าจอหลักของระบบ

จากภาพ 2 หน้าจอหลักของระบบ ประกอบด้วย (1) แถบเมนูให้คลิกเลือกฟังก์ชันการทำงานของระบบ ได้แก่ ฟังก์ชันจัดการข้อมูลผู้ใช้งาน ฟังก์ชันจัดการข้อมูลชุดวิชา ฟังก์ชันจัดการคลังข้อมูลข้อสอบ ฟังก์ชันสร้างโมเดลข้อสอบ ฟังก์ชันสร้างข้อสอบคู่ขนาน และออกจากระบบ (2) ข้อความแสดง ชื่อ – นามสกุล ผู้ใช้งานระบบ และ (3) ส่วนแสดงผลข้อมูลเมื่อผู้ใช้งานคลิกเลือกฟังก์ชันการทำงานของระบบฟังก์ชันใดฟังก์ชันหนึ่งในแถบเมนู ซึ่งหน้าจอดังกล่าวพัฒนาขึ้นสำหรับให้ผู้ใช้เลือกเมนูที่ต้องการใช้งานระบบ



ภาพ 3 หน้าจอแสดงตัวอย่างข้อสอบคู่ขนานที่สร้างจากโมเดลข้อสอบ

จากภาพ 3 หน้าจอแสดงตัวอย่างข้อสอบคู่ขนานที่สร้างจากโมเดลข้อสอบข้อที่ 1 ในตาราง 2 ซึ่งหน้าจอประกอบด้วยส่วนสำคัญหลัก 2 ส่วน ได้แก่ (1) ส่วนแสดงเนื้อหาข้อสอบที่สร้างจากโมเดลข้อสอบ (2) ส่วนแสดงรายการตัวแปรและค่าตัวแปรของโมเดลข้อสอบที่ระบบสุ่มค่าได้ ซึ่งหน้าจอดังกล่าวพัฒนาขึ้นสำหรับ ใช้ตรวจสอบความถูกต้องโครงสร้าง และเนื้อหาของข้อสอบคู่ขนานที่สร้างจากโมเดลข้อสอบก่อนนำไปสร้างข้อสอบคู่ขนาน



ภาพ 4 หน้าจอระบุจำนวนข้อสอบคู่ขนานที่ต้องการสร้างจากโมเดลข้อสอบ

จากภาพ 4 หน้าจอสำหรับให้ผู้ใช้งานระบุจำนวนข้อสอบคู่ขนานที่ต้องการสร้างจากโมเดลข้อสอบที่สร้างไว้ สามารถกำหนดให้ทุกโมเดลข้อสอบสร้างข้อสอบจำนวนข้อเท่ากัน หรือต่างกันก็ได้ เมื่อระบบสร้างข้อสอบคู่ขนานครบตามจำนวนที่ผู้ใช้งานกำหนดเสร็จสิ้นแล้ว จะทำการบันทึกไฟล์ข้อสอบคู่ขนานที่ถูกสร้างลงในเครื่องคอมพิวเตอร์ที่ใช้งานระบบในรูปแบบของไฟล์ word (.docx)

2.3 ผลการตรวจสอบคุณภาพในภาพรวม ของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ โดยผู้เชี่ยวชาญ แสดงรายละเอียดดังตาราง 3

ตาราง 3 ผลการตรวจสอบคุณภาพในภาพรวมของระบบโดยผู้เชี่ยวชาญ

รายการ	<i>M</i>	<i>SD</i>	ระดับคุณภาพ
ภาพรวมของระบบ			
1. การออกแบบระบบ	4.37	0.55	มาก
2. การทำงานของระบบ	4.56	0.54	มากที่สุด
3. การใช้งานระบบ	4.53	0.51	มากที่สุด
4. คู่มือการใช้งานระบบ	4.10	0.55	มาก
เฉลี่ยรวม	4.39	0.56	มาก

จากตาราง 3 ผลการวิเคราะห์ข้อมูลการตรวจสอบคุณภาพในภาพรวม ของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติฯ โดยผู้เชี่ยวชาญ พบว่าระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ มีคุณภาพเฉลี่ยในภาพรวมอยู่ในระดับมาก และนอกจากนี้ระบบที่พัฒนาขึ้นยังมีผลประเมินการใช้งานระบบจากผู้ใช้งานเฉลี่ยในภาพรวมอยู่ในระดับมากอีกด้วย

3. ผลการตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์

3.1 ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบรายข้อระหว่างข้อสอบจากแบบทดสอบต้นแบบกับข้อสอบจากแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยผู้เชี่ยวชาญ จำนวน 42 ข้อ พบว่า ข้อสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้นมีความเป็นคู่ขนานกับข้อสอบต้นแบบทุกข้อ เนื่องจากผ่านเกณฑ์ความเป็นคู่ขนานของข้อสอบที่กำหนดไว้ คือข้อสอบที่มีความเป็นคู่ขนานกัน ต้องมีค่าดัชนีความสอดคล้องของทุกรายการประเมิน เป็นไปตามเกณฑ์ที่กำหนดคือ ตั้งแต่ 0.5 ขึ้นไป

3.2 ผลการตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยการทดลองสอบ มีผลการวิเคราะห์ข้อมูลที่เกี่ยวข้องกับการตรวจสอบความเป็นคู่ขนานของแบบทดสอบ ดังนี้

1) ค่าความยากและผลการทดสอบความแตกต่างของค่าความยาก พบว่าค่าความยากรายข้อของแบบทดสอบต้นแบบ (p_A) อยู่ระหว่าง 0.33-0.77 ค่าความยากเฉลี่ยทั้งฉบับเท่ากับ 0.52 ค่าความยากรายข้อของแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น (p_B) อยู่ระหว่าง 0.27-0.77 ค่าความยากเฉลี่ยทั้งฉบับเท่ากับ 0.51 และผลการทดสอบความแตกต่างของค่าความยากเฉลี่ยทั้งฉบับ แสดงดังตาราง 4

ตาราง 4 ผลการทดสอบความแตกต่างของค่าความยากเฉลี่ยทั้งฉบับของแบบทดสอบ 2 ฉบับ

ค่าความยาก	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i> – value
p_A	42	0.52	0.12	1.25	41	0.24
p_B	42	0.51	0.13			

จากตาราง 4 เมื่อพิจารณาจากค่าสถิติที่ใช้ทดสอบความแตกต่างของข้อมูลแล้ว สามารถสรุปได้ว่าค่าความยากเฉลี่ยทั้งฉบับของแบบทดสอบทั้ง 2 ฉบับ ไม่แตกต่างกัน

2) ค่าอำนาจจำแนกและผลการทดสอบความแตกต่างของค่าอำนาจจำแนก พบว่าค่าอำนาจจำแนกรายข้อของแบบทดสอบต้นแบบ (r_A) อยู่ระหว่าง 0.20-0.80 และค่าอำนาจจำแนกรายข้อของแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น (r_B) อยู่ระหว่าง 0.20-0.73 ผลการทดสอบความแตกต่างของค่าอำนาจจำแนกเฉลี่ยทั้งฉบับแสดงดังตาราง 5

ตาราง 5 ผลการทดสอบความแตกต่างของค่าอำนาจจำแนกเฉลี่ยทั้งฉบับของแบบทดสอบ 2 ฉบับ

ค่าอำนาจจำแนก	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i> – value
r_A	42	0.41	0.19	0.10	41	0.91
r_B	42	0.40	0.15			

จากตาราง 5 เมื่อพิจารณาจากค่าสถิติที่ใช้ทดสอบความแตกต่างของข้อมูลแล้ว สามารถสรุปได้ว่าค่าอำนาจจำแนกเฉลี่ยทั้งฉบับของแบบทดสอบทั้ง 2 ฉบับ ไม่แตกต่างกัน

3) ค่าความแปรปรวนและผลการทดสอบความแตกต่างของค่าความแปรปรวน พบว่าค่าความแปรปรวนรายข้อของแบบทดสอบต้นแบบ อยู่ระหว่าง 0.19-0.26 ค่าความแปรปรวนทั้งฉบับเท่ากับ 0.24 ค่าความแปรปรวนรายข้อของแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น อยู่ระหว่าง 0.19-0.26 ค่าความแปรปรวนทั้งฉบับเท่ากับ 0.25 และเมื่อพิจารณาจากค่าสถิติที่ใช้ทดสอบความแตกต่างของข้อมูลที่มีค่า $F = 1.04$ ($\alpha = 0.05$, $df_1 = 29$, $df_2 = 29$, $F = 1.85$) สามารถสรุปได้ว่าค่าความแปรปรวนทั้งฉบับของแบบทดสอบทั้ง 2 ฉบับไม่แตกต่างกัน

4) ค่าเฉลี่ยและผลการทดสอบความแตกต่างของค่าเฉลี่ยของคะแนนแบบทดสอบระหว่างแบบทดสอบต้นแบบ ($\bar{T}_A$) และแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น ($\bar{T}_B$) แสดงดังตาราง 6

ตาราง 6 ค่าเฉลี่ยและผลการทดสอบความแตกต่างของค่าเฉลี่ยของคะแนนแบบทดสอบ 2 ฉบับ

คะแนนแบบทดสอบ	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i> – value
$\bar{T}_A$	30	21.9	9.26	0.52	29	0.60
$\bar{T}_B$	30	21.0	9.28			

จากตาราง 6 เมื่อนำแบบทดสอบไปใช้ทดสอบกับนักศึกษาจำนวน 30 คน พบว่าค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบต้นแบบและแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้นเท่ากับ $M=21.9$, $SD=9.26$ และ $M=21.0$, $SD=9.28$ ตามลำดับ เมื่อพิจารณาจากค่าสถิติที่ใช้ทดสอบความแตกต่างของข้อมูลแล้ว สามารถสรุปได้ว่าค่าเฉลี่ยของคะแนนแบบทดสอบทั้ง 2 ฉบับ ไม่แตกต่างกัน

5) ค่าความเชื่อมั่นแบบสมมูลของแบบทดสอบ แสดงดังตาราง 7

ตาราง 7 ผลการหาค่าความเชื่อมั่นแบบสมมูลของแบบทดสอบ 2 ฉบับ

คะแนนแบบทดสอบ	<i>n</i>	Correlation	<i>p</i> – value
แบบทดสอบต้นแบบ	30	0.98	0.00
แบบทดสอบสร้างจากระบบคอมพิวเตอร์			

จากตาราง 7 พบว่าค่าความเชื่อมั่นแบบสมมูลของแบบทดสอบโดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนแบบทดสอบต้นแบบและแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้นเท่ากับ 0.98 ดังนั้นคะแนนแบบทดสอบต้นแบบและแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้นจึงมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

6) ค่าความเชื่อมั่นของแบบทดสอบ (r_{tt}) โดยใช้สูตรของ Kuder-Richardson 20 (KR_{20}) และค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบ (SE) แสดงได้ดังตาราง 8

ตาราง 8 ผลการหาค่าความเชื่อมั่นของแบบทดสอบ 2 ฉบับ

แบบทดสอบ	k (จำนวนข้อสอบ)	n	r_{tt}	SE
แบบทดสอบต้นแบบ	42	30	0.89	2.98
แบบทดสอบสร้างจากระบบคอมพิวเตอร์	42	30	0.90	2.84

จากตาราง 8 พบว่าแบบทดสอบทั้ง 2 ฉบับมีค่าความเชื่อมั่นและความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบใกล้เคียงกัน ซึ่งเป็นสิ่งหนึ่งที่บ่งบอกได้ว่าแบบทดสอบทั้ง 2 ฉบับมีความเป็นคู่ขนานกัน

อภิปรายผล

ผลการสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์ฯ การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติฯ และการตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ มีประเด็นอภิปรายดังนี้

1. การสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์ฯ อภิปรายผลการวิจัยได้ดังนี้
1) ประเภทของโมเดลข้อสอบที่สามารถสร้างขึ้นในงานวิจัยนี้มีจำนวนทั้งสิ้น 2 ประเภท ได้แก่ โมเดลที่ 5 และโมเดลที่ 8 โดยไม่สามารถสร้างโมเดลข้อสอบภายใต้หลักการการสร้างโมเดลข้อสอบของ Gierl et al. (2008) ได้ครบทั้ง 10 ประเภท เนื่องจากข้อจำกัดทางด้านเนื้อหาของข้อสอบต้นแบบที่ไม่สามารถจัดเข้ากับโครงสร้างของโมเดลข้อสอบอีก 8 ประเภท ที่เหลือได้ซึ่งสอดคล้องกับงานวิจัยของ Chantharamaha & Chadcham (2018); Jantana & Pradujprom (2020); Pradujprom & Pantong (2019) ที่ไม่สามารถสร้างโมเดลข้อสอบวิชาคณิตศาสตร์ได้ครบทั้ง 10 โมเดล เนื่องจากข้อจำกัดทางด้านเนื้อหาของข้อสอบต้นแบบดังกล่าว อย่างไรก็ตามการสร้างโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) ก็ได้มีข้อกำหนดไว้ว่าต้องสร้างโมเดลข้อสอบให้ครบทั้ง 10 ประเภทเช่นกัน และ 2) คลังข้อมูล หรือคลังข้อมูลข้อสอบ เป็นคลังข้อมูลที่ออกแบบขึ้น

เพื่อพัฒนาเป็นฟังก์ชันการทำงานฟังก์ชันหนึ่งของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ ที่สามารถรองรับการสร้างโมเดลข้อสอบในระบบสร้างข้อสอบคู่ขนานอัตโนมัติ ให้มีความสะดวกในการสร้างโมเดลข้อสอบมากยิ่งขึ้นเท่านั้น ซึ่งผู้ใช้งานอาจจะใช้งานข้อมูลที่มีในคลังข้อมูลหรือไม่ก็ได้ ทั้งนี้ขึ้นอยู่กับความประสงค์ และบริบทของเนื้อหาข้อสอบที่นำมาสร้างโมเดลข้อสอบนั้น ๆ ด้วยว่าถูกกำหนดให้มีตัวแปรชนิดตัวอักษร หรือตัวแปรชนิดข้อความหรือไม่ และตัวแปรที่ต้องการใช้งานถูกสร้างไว้ในคลังข้อมูลแล้วหรือไม่

2. การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติฯ อภิปรายผลการวิจัยได้ดังนี้

1) ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น สามารถสร้างข้อสอบคู่ขนานอัตโนมัติได้เป็นผลจากการบูรณาการปัจจัยดังต่อไปนี้ (1) หลักการสร้างข้อสอบอัตโนมัติผ่านโมเดลข้อสอบของ Gierl et al. (2008) (2) การสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ ที่คำนึงถึงหลักการความเป็นคู่ขนานของข้อสอบ และ (3) หลักการพัฒนาระบบตามวงจรพัฒนาระบบ SDLC ตามเอกสารของ Preechapanich (2014). และหลักการพัฒนาโปรแกรม นอกจากนี้ระบบสร้างข้อสอบคู่ขนานที่ดังกล่าว ยังถูกพัฒนาให้สามารถรองรับการสร้างข้อสอบคู่ขนาน ของข้อสอบวิชาคณิตศาสตร์ เรื่องอื่น ๆ ที่นอกเหนือจากที่แสดงในงานวิจัยนี้ได้ รวมทั้งสามารถรองรับเนื้อหาข้อสอบวิชาอื่น ๆ ที่บริบทเนื้อหาวิชาดังกล่าวสามารถประยุกต์เข้ากับหลักการสร้างโมเดลข้อสอบของ Gierl et al. (2008) ได้ 2) ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น ผู้วิจัยสามารถออกแบบและพัฒนาระบบ ให้ผู้ใช้งานสร้างโมเดลข้อสอบได้ภายใต้ฟังก์ชันสร้างโมเดลข้อสอบที่พัฒนาขึ้น ตามหลักการของ Gierl et al. (2008) ได้ครบทั้ง 10 โมเดล ซึ่งแตกต่างจากระบบสร้างข้อสอบอัตโนมัติในงานวิจัยอื่น เช่นงานวิจัยของ Chantharamaha & Chadcham (2018); Jantana & Pradujprom (2020); Pradujprom & Pantong (2019); Raksombat et al. (2022) ที่มีข้อจำกัดที่ผู้ใช้งานไม่สามารถสร้างโมเดลข้อสอบเพิ่มเติมหรือแก้ไขโมเดลข้อสอบได้ แต่อย่างไรก็ตามฟังก์ชันสร้างโมเดลข้อสอบที่พัฒนาขึ้นดังกล่าว ยังมีข้อจำกัดที่ผู้ใช้งานต้องเป็นผู้สร้างโมเดลข้อสอบในระบบเอง ยังไม่สามารถพัฒนาให้ระบบสามารถสร้างโมเดลข้อสอบได้อย่างอัตโนมัติ

3. การตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ อภิปรายผลการวิจัยได้ว่า ข้อสอบจากแบบทดสอบต้นแบบ และข้อสอบจากแบบทดสอบสร้างขึ้นจากระบบคอมพิวเตอร์ทั้ง 42 ข้อ มีความเป็นคู่ขนานกัน ทั้งเชิงเนื้อหาและเชิงสถิติ อาจเป็นผลจากในกระบวนการสร้างโมเดลข้อสอบ ได้นำข้อมูลของลักษณะเฉพาะของข้อสอบ และข้อสอบต้นแบบ มาเป็นขอบเขตข้อมูลในการสร้างโมเดลข้อสอบ จากนั้นจึงนำโมเดลข้อสอบที่สร้างขึ้นไปใช้สร้างข้อสอบคู่ขนานผ่านระบบที่พัฒนาขึ้นต่อไป จึงส่งผลให้ข้อสอบคู่ขนานที่สร้างขึ้น มีรูปแบบเนื้อหา และคุณภาพของข้อสอบใกล้เคียงกับข้อสอบต้นแบบหรือข้อสอบต้นฉบับ ทำให้การวัดและประเมินผลการเรียนรู้ด้วยแบบทดสอบคู่ขนานดังกล่าวมีความเท่าเทียมกัน หรือมีมาตรฐานเดียวกัน และมีความน่าเชื่อถือ นอกจากนี้การตรวจสอบตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยอาศัยดุลยพินิจของผู้เชี่ยวชาญจำนวน 3 คน นั้นมีความเพียงพอที่ทำให้ผลการทดสอบแสดงให้เห็นว่าแบบทดสอบทั้ง 2 ฉบับ มีความเป็นคู่ขนานกัน เนื่องจากแบบทดสอบที่นำมาให้ผู้เชี่ยวชาญตรวจสอบความเป็นคู่ขนานนั้น ได้สร้างจากโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) จึงส่งผลให้ข้อสอบที่ได้มีแนวโน้มของความเป็นคู่ขนานกันอยู่ก่อนแล้ว ซึ่งสอดคล้องกับ Chantharamaha & Chadcham (2018)

และด้วยเหตุผลดังกล่าวยังส่งผลให้การทดลองสอบมีค่าสถิติที่บ่งบอกถึงความเป็นคู่ขนานของแบบทดสอบ โดยมีความสอดคล้องกับงานวิจัยของ Charuthat (2003); Kanchanapenkul (2011); Sawangboon et al. (2012) อีกด้วย

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 อาจารย์ผู้สอนสามารถนำระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น ไปปรับใช้สร้างข้อสอบคู่ขนานอัตโนมัติของวิชาอื่น ๆ ที่มีบริบทเนื้อหาใกล้เคียง หรือที่มีบริบทเนื้อหาที่สามารถประยุกต์เข้ากับหลักการสร้างโมเดลข้อสอบที่ใช้ในงานวิจัยนี้ได้ เช่น วิชาคณิตศาสตร์ระดับชั้นต่าง ๆ วิชาฟิสิกส์ วิชาเคมี เป็นต้น

1.2 อาจารย์ผู้สอนสามารถนำระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น ไปปรับใช้สร้างแบบฝึกหัด หรือข้อคำถาม เพื่อนำไปใช้วัดและประเมินผลการเรียนรู้ของผู้เรียนในรูปแบบต่าง ๆ ได้

1.3 ผู้ดูแลระบบคลังข้อสอบมาตรฐานสามารถนำข้อสอบคู่ขนานที่สร้างจากระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น ไปจัดเก็บไว้ในระบบคลังข้อสอบมาตรฐานเพื่อใช้วัดและประเมินผลการเรียนรู้ของผู้เรียน ในภาคการศึกษาต่าง ๆ ได้ต่อไป

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 จากวิธีการวิเคราะห์หรือตรวจสอบความเป็นคู่ขนานของแบบทดสอบในครั้งนี้อาศัยดุลยพินิจของผู้เชี่ยวชาญและการตรวจสอบความเป็นคู่ขนานโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม ดังนั้นในการวิจัยครั้งต่อไปอาจประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบมาตรวจสอบความเป็นคู่ขนานของข้อสอบ

2.2 จากชนิดของข้อสอบที่สร้างขึ้นในครั้งนี้เป็นแบบเลือกตอบ 100% ดังนั้นในการวิจัยครั้งต่อไปควรมีการศึกษาเพื่อพัฒนาระบบสร้างข้อสอบคู่ขนานรูปผสม หรือแบบอัตนัยอัตโนมัติ เพื่อให้สามารถใช้ในเนื้อหาวิชาอื่นได้ต่อไป

References

- Gierl, M. J., & Haladyna, T. (Eds.) (2013). *Automatic Item Generation: Theory and Practice*. Routledge.
- Gierl, M. J., & Lai, H. (2012). The Role of Item Models in Automatic Item Generation. *International Journal of Testing*, 12(3), 273-298.
- Gierl, M. J., & Lai, H. (2013). Instructional Topics in Educational Measurement (ITEMS) Module: Using Automated Processes to Generate Test Items. *Educational Measurement: Issues and Practice*, 32(3), 36-50.
- Gierl, M. J., Lai, H., & Turner, S. (2012). Using Automatic Item Generation to Create Multiple-Choice Items for Assessments in Medical Education. *Medical Education*, 46(8), 757-765.

Gierl, M.J., Zhou, J., & Alves, C. (2008). Developing a Taxonomy of Item Model Types to Promote Assessment Engineering. *Journal of Technology, Learning, and Assessment*, 7(2), 1-51.

Translate Thai Reference

Chantharamaha, S., & Chadcham, S. (2018). Automatic Item Generation in Mathematics for Grade Six Students Using Computer Software. *Research Methodology & Cognitive Science*, 16(1), 138-149. (in Thai)

Charuthat, S. (2003). Study of Parallel of Prathomsuksa 6 Mathematics Achievement Test Constructed with Facet Design [Master's thesis]. Chiang Mai University. (in Thai)

Jantana, S., & Pradujprom, P. (2020). Development of Automatic Item Generation Program in Mathematics for Grade Nine Level. *Journal of Educational Measurement*, 37(102), 162-177. (in Thai)

Kanchanapenkul, S. (2011) A Construction of Parallel Domain-Referenced Test by Using Facet Design on "Circles" for Matthayomsuksa 3. *Sripatum Chonburi Academic Journal*, 9(4), 56-60. (in Thai)

Pradujprom, P., & Panthong, K. (2019). A Computer Program for Developing Multiple Choice Test Items Using Automatic Item Generation Method. *Journal of Southern Technology*, 12(2), 74-87. (in Thai)

Preechapanich, O. (2014). *Guide book: System Analysis and Design, Complete version*. IDC Premier. (in Thai)

Premthongsuk, P., Chadcham, S., & Pradujprom, P. (2017). Development of the Next Item Selection Procedure Using Hurwicz Criterion and Item Exposure Control for Computerized Adaptive Testing. *Veridian E-Journal, Science and Technology Silpakorn University*, 4(6), 32-50. (in Thai)

Raksombat, S., Pradujphom, P., & Panthong, K. (2022). Development of an Automatic Item Generation System for Item Bank Classified by Content and Difficulty Levels: Application of Assessment Engineering Concepts. *Journal of Educational Measurement, Mahasarakham University*, 28(1), 311-330. (in Thai)

Sawangboon, N., Phattiyathani, S., & Wongrattana, C. (2012). A Construction of Mathematics Parallel Tests Using Facet Design on "Factoring of Degree-2 Polynomials" for Student in Matthayomsuksa 2. *Journal of Educational Measurement, Mahasarakham University*, 17(1), 369-378. (in Thai)

Srisa-ard, O. (1995). Verification of the validity of measurement tools by experts. *Journal of Educational Measurement, Mahasarakham University*, 1(1), 45-49. (in Thai)