

การพัฒนากระบวนสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหา
และระดับความยาก: การประยุกต์แนวคิดวิศวกรรมการประเมิน
Development of an Automatic Item Generation System for Item Bank
Classified by Content and Difficulty Levels:
Application of Assessment Engineering Concepts

สุรัชชัย รักสมบัติ^{1*} ปิยะทิพย์ ประดุงพรม² และ กนก พานทอง³

Surachai Raksombat^{1*} Piyathip Pradujphom² and Kanok Panthong³

(Received: March 13, 2021 ; Revised: April 19, 2021 ; Accepted: April 26, 2021)

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) สร้างโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 และ 2) พัฒนาระบบสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 วิธีการดำเนินการวิจัยแบ่งเป็น 2 ระยะ คือ 1) การสร้างโมเดลข้อสอบ วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 และ 2) การพัฒนาระบบสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบ จำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

ผลการวิจัยพบว่า

1. การสร้างโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ประกอบด้วยส่วนสำคัญ 5 ส่วน ได้แก่ 1) โจทย์ 2) ส่วนประกอบ 3) ตัวเลือก 4) ข้อมูลเสริม และ 5) เฉลย ข้อสอบที่เหมาะสมในการสร้างเป็น โมเดลข้อสอบจำนวนทั้งหมด 203 ข้อ และสร้างเป็นโมเดลข้อสอบจำนวน 10 รูปแบบ โดยรวมโมเดลข้อสอบ มีความถูกต้องผ่านเกณฑ์

¹ นิสิตปริญญาเอก สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

² ผู้ช่วยศาสตราจารย์ หน่วยวิจัยวิทยาการปัญญาและนวัตกรรม วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

³ อาจารย์ หน่วยวิจัยวิทยาการปัญญาและนวัตกรรม วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

¹ Doctoral degree student, Research and Statistics in Cognitive Science, College of Research Methodology and Cognitive Science, Burapha University

² Assistant Professor, Cognitive Science and Innovation Research Unit: CSIRU, College of Research Methodology and Cognitive Science, Burapha University

³ Lecturer, Cognitive Science and Innovation Research Unit: CSIRU, College of Research Methodology and Cognitive Science, Burapha University

*Corresponding Author E-mail: piyatip10@hotmail.com

2. ระบบสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยากวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ที่พัฒนาขึ้นในรูปแบบของเว็บแอปพลิเคชัน สามารถใช้งานผ่านระบบอินเทอร์เน็ตได้ที่ <https://www.aig-system.com> โปรแกรมในภาพรวมมีความเหมาะสมอยู่ในระดับมากเป็นที่ยอมรับของผู้เชี่ยวชาญ และการประเมินประสิทธิภาพในการใช้งานของโปรแกรมอยู่ในระดับดีมากเป็นที่พึงพอใจของผู้ทดลองใช้โปรแกรม

คำสำคัญ: การพัฒนาระบบ คลังข้อสอบ การสร้างข้อสอบอัตโนมัติ โมเดลข้อสอบ

Abstract

The objectives of this research were 1) to create a mathematics item model in mathematics of grade 7, and 2) to develop an automatic item generation system for item bank classified by content and difficulty levels in mathematics of grade 7. The research was conducted in 2 phases. Phase 1 dealt with creating a mathematics item model in mathematics of grade 7. Phase 2 dealt with development of an automatic item generation system and to organize the item bank for grade 7, classified by content and difficulty levels in mathematics of grade 7.

The results showed that:

1. The mathematics item model in mathematics of grade 7 that was created consisted of 5 main parts: 1) stems, 2) elements, 3) options, 4) auxiliary Information and 5) keys. The items appropriate for creating an item model were 203 questions and were used to create 10 item model formats. On the whole, the item model was valid and passed the criteria.

2. The development of an automatic item generation system for item bank classified by content and difficulty levels in mathematics for grade 7 was carried out in the form of a web application which can be used via the Internet at <https://www.aig-system.com>. The overall program was suitable at a high level and accepted by the experts. The evaluation of the efficiency of the program was at a very good level, satisfying the program's users.

Keywords: system development, item bank, automatic item generation, item model

บทนำ

ปัจจุบันความก้าวหน้าอย่างรวดเร็วทางด้านเทคโนโลยีคอมพิวเตอร์และอินเทอร์เน็ตได้เข้ามา มีบทบาทอย่างมากในการวัดผลทางการศึกษา ทั้งนี้เนื่องมาจากการจัดการการทดสอบที่มีประสิทธิภาพและการได้รับความนิยมน้อย่างกว้างขวาง เพราะเทคโนโลยีคอมพิวเตอร์และอินเทอร์เน็ตได้ทำให้เกิดการรูปแบบการทดสอบที่หลากหลาย Gierl and Lai (2012) กล่าวว่า การนำเทคโนโลยีทางคอมพิวเตอร์มาช่วยในการทดสอบ จะทำให้เกิดรูปแบบการทดสอบใหม่ ๆ และส่งผลถึงการพัฒนาด้านข้อสอบโดยตรง ซึ่งจะต้องมีการสร้างข้อสอบ

ที่มีคุณภาพขึ้นเป็นจำนวนมากเพื่อนำข้อสอบเข้าสู่คลังข้อสอบทดแทนข้อสอบที่ถูกใช้ไป เป็นการป้องกันปัญหาการใช้ข้อสอบซ้ำกัน และเพื่อรักษาความเชื่อมั่นในการทดสอบ (Gierl & Lai, 2013) การสร้างข้อสอบโดยวิธีการแบบดั้งเดิม คือ ให้ผู้เชี่ยวชาญเป็นผู้ดำเนินการสร้างข้อสอบ มักจะพบปัญหาหลายอย่างตามมา ได้แก่ ค่าใช้จ่ายที่สูงในการดำเนินงาน ระยะเวลาการสร้างข้อสอบ การไม่ปฏิบัติตามหลักเกณฑ์ในการสร้างข้อสอบ รวมถึงการที่ไม่สามารถสร้างข้อสอบในระดับที่ซับซ้อนมาก ๆ เพราะจำกัดด้วยความสามารถของผู้สร้างข้อสอบ (Arendasy & Sommer, 2012)

Gierl and Lai (2012) กล่าวว่า การสร้างข้อสอบอัตโนมัติ (Automatic Item Generation: AIG) เป็นแนวทางการสร้างข้อสอบโดยใช้เทคโนโลยีคอมพิวเตอร์ที่กำลังได้รับความนิยมอย่างมากในปัจจุบัน เพราะสามารถสร้างข้อสอบเป็นจำนวนมากในระยะเวลาสั้น ซึ่งเป็นการแก้ปัญหาการมีข้อสอบไม่เพียงพอ การสร้างข้อสอบอัตโนมัติเป็นงานวิจัยค่อนข้างใหม่แต่มีการพัฒนาอย่างรวดเร็ว โดยเฉพาะด้านที่มีการใช้ทฤษฎีทางการปัญญา (cognitive) และจิตวิทยา (psychometric) หลักการสร้างข้อสอบอัตโนมัติใช้สองขั้นตอนหลัก คือ ขั้นตอนหนึ่ง ผู้เชี่ยวชาญด้านการพัฒนาข้อสอบจะสร้างแบบจำลองข้อสอบเป็นข้อสอบต้นแบบ โดยเน้นคุณลักษณะที่ต้องการประเมิน และขั้นที่สอง ทำการจัดการเพื่อสร้างข้อสอบใหม่โดยใช้อัลกอริทึมทางคอมพิวเตอร์ การสร้างข้อสอบอัตโนมัติสามารถสร้างข้อสอบขึ้นใหม่เป็นจำนวนมากจากข้อสอบต้นแบบเพียงรายการเดียว จากงานวิจัยของ Gierl et al. (2012) กล่าวว่า การสร้างข้อสอบอัตโนมัตียังมีจุดอ่อนสองประการสำคัญ คือ ประการแรก ข้อสอบจำนวนมากที่ถูกสร้างขึ้นยังไม่สามารถนำเข้าสู่คลังข้อสอบเพื่อเตรียมนำไปใช้ ในการทดสอบได้ทันทีจะต้องถูกนำไปวิเคราะห์คุณภาพข้อสอบว่าผ่านเกณฑ์ตามที่กำหนดไว้หรือไม่ โดยใช้กลุ่มตัวอย่างที่มีความเหมาะสมกับข้อสอบ จากนั้นจะนำผลการทดสอบไปวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบด้วยโปรแกรมวิเคราะห์คุณภาพ ข้อสอบ โดยคุณภาพของข้อสอบจะแยกเป็นสองส่วน คือ ส่วนแรกเป็นแบบรายข้อ ได้แก่ ค่าพารามิเตอร์ความยาก (b-parameter) ค่าพารามิเตอร์อำนาจจำแนก (a-Parameter) และค่าพารามิเตอร์การเดา (c-parameter) ส่วนที่สอง เป็นแบบทั้งฉบับ ได้แก่ ค่าความตรง (validity) และค่าความเที่ยง (reliability) เพื่อให้แน่ใจว่าข้อสอบมีคุณภาพเพียงพอในการทำให้ผลการทดสอบมีมาตรฐานและน่าเชื่อถือ นอกจากนี้ผลของงานวิจัยยังพบว่า ข้อสอบที่ถูกสร้างแบบอัตโนมัติถูกคัดออกจำนวนมากเนื่องจากไม่ผ่านการวิเคราะห์คุณภาพ และประการที่สอง ข้อสอบจำนวนมากที่ถูกสร้างขึ้นไม่ได้จำแนกระดับความยากของข้อสอบ ทำให้ข้อสอบที่ถูกสร้างขึ้นไม่เหมาะสมกับ ผู้เข้าทดสอบในด้านความสามารถ จากงานวิจัยของ Luecht (2013) ได้กล่าวถึง การสร้างข้อสอบอัตโนมัติ (AIG) ภายใต้แนวคิดวิศวกรรมการประเมิน (AE) เป็นหลักการที่ทำให้การวัดผลการทดสอบหรือการให้คะแนนเป็นไปอย่างถูกต้อง และสามารถประหยัดค่าใช้จ่ายสำหรับการออกแบบและพัฒนาระบบทดสอบ วิศวกรรมการประเมินเป็นการผสมผสานกันระหว่างการสร้างแบบจำลองความรู้ความเข้าใจ และหลักการออกแบบทางวิศวกรรม คือ การใช้แนวทางใหม่เกี่ยวกับลักษณะของการให้น้ำหนักการให้คะแนนในด้านความยากของข้อสอบและด้านเนื้อหาของข้อสอบ

จากเหตุผลที่กล่าวมาข้างต้น ผู้วิจัยจึงสนใจที่จะพัฒนาระบบสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก โดยการประยุกต์แนวคิดวิศวกรรมการประเมิน (AE) ของ Luecht (2013) ด้วยการสร้างโมเดลข้อสอบจำแนกตามเนื้อหาและระดับความยากของข้อสอบ โดยการจำแนก

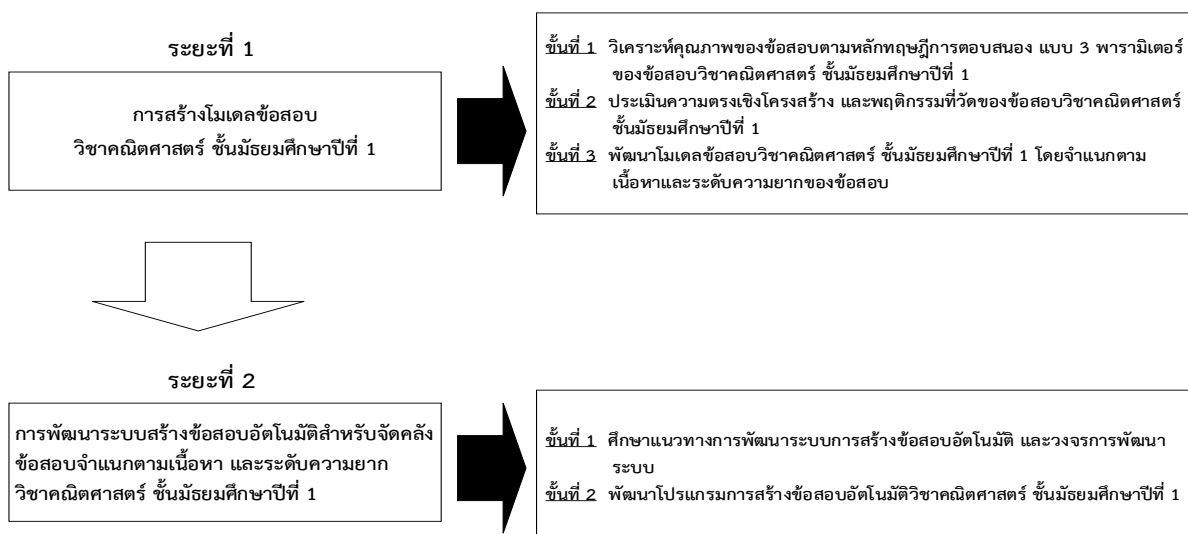
เนื้อหาตามโครงสร้างหลักสูตรวิชาคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ประกอบด้วย 3 สาระ คือ สาระที่ 1 จำนวนและพีชคณิต สาระที่ 2 การวัดและเรขาคณิต และสาระที่ 3 สถิติและความน่าจะเป็น และจำแนกระดับความยากของข้อสอบตามหลักการทฤษฎีการตอบ สมองข้อสอบ (Item Response Theory: IRT) แบบ 3 พารามิเตอร์ แบ่งเป็น 3 ระดับ คือ ง่าย ปานกลาง และยาก ซึ่งการสร้างข้อสอบอัตโนมัติ (AIG) ของ Gierl et al. (2008) จะเป็นการสร้างระบบข้อสอบอัตโนมัติ วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ที่จำแนกตามเนื้อหาและระดับความยาก ซึ่งจะสามารถสร้างข้อสอบได้จำนวนมากในเวลาทีรวดเร็ว เป็นการช่วยลดภาระงานของครูในการสร้างข้อสอบวิชาคณิตศาสตร์ ทำให้การวัดผลการเรียนรู้มีคุณภาพ รวมทั้งได้ข้อสอบที่มีคุณภาพที่ได้แบ่งระดับความยากง่ายของข้อสอบเก็บไว้ในคลังข้อสอบจำนวนมากไว้ฝึกฝน จะช่วยเพิ่มทักษะทางคณิตศาสตร์ให้กับนักเรียน

ความมุ่งหมายของการวิจัย

1. เพื่อสร้างโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1
2. เพื่อพัฒนาระบบสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยากวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

วิธีดำเนินการวิจัย

การวิจัยนี้แบ่งวิธีดำเนินการวิจัยเป็น 2 ระยะ ได้แก่ ระยะที่ 1 การสร้างโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 และระยะที่ 2 การพัฒนาระบบสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 แสดงดังภาพ 1



ภาพ 1 วิธีดำเนินการวิจัย

ระยะที่ 1 การสร้างโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

1. กลุ่มตัวอย่าง

ผู้วิจัยรวบรวมข้อสอบและผลการตอบข้อสอบวิชาคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ที่เข้าสอบแข่งขันวัดความสามารถทางคณิตศาสตร์ (Top Mathematics Test) เป็นข้อมูลแบบทุติยภูมิ (secondary data) ตั้งแต่ปี พ.ศ. 2557 – 2561 ที่ได้มาจาก บริษัท ท็อป เทสต์ เซ็นเตอร์ จำกัด ซึ่งมีข้อสอบทั้งหมด 480 ข้อ จากจำนวนนักเรียนที่เข้าสอบทั้งหมด 27,886 คน

2. เครื่องมือที่ใช้ในการวิจัย

แบบประเมินความถูกต้องของโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 โมเดลข้อสอบจะแบ่งรายละเอียดเป็น 5 ส่วน คือ 1) โจทย์ (stem) 2) ส่วนประกอบ (element) 3) ตัวเลือก (options) 4) ข้อมูลเสริม (auxiliary Information) และ 5) เฉลย (key) แสดงดังภาพ 2

สาระที่ 1 จำนวนและพีชคณิต ข้อที่ 31				
ประเมินความถูกต้องของโมเดลข้อสอบ				
ความคิดเห็น ของผู้วิจัย	โมเดลข้อสอบ			โมเดลที่ 7
	รูปแบบของโจทย์และตัวเลือก ในโมเดลข้อสอบ			โจทย์มีลักษณะเป็นแบบผสม (Mixed) ตัวเลือกแบบเลือกจากการสุ่ม (Randomly Selected)
ผู้เชี่ยวชาญ พิจารณา	ระดับความถูกต้อง			
	1	2	3	4
ข้อเสนอแนะ				
Stem				
<p>ให้ S1 เป็นคำตอบของสมการ $\frac{3A+1}{2} = \frac{7A-2}{5}$ และ S2 เป็นคำตอบของสมการ $7(B+3)-5 = 4B+8$ แล้วค่าของ S1 - I1S2 ตรงกับข้อใด</p>				
Elements				
<p>S1 Range: "A", "X", "M" S2 Range: "B", "Y", "N" As_S1 = "A" then S2 = "B" As_S1 = "X" then S2 = "Y" As_S1 = "M" then S2 = "N" I1 Value Range: [-30] - [30] by 3 I2 = $-9 - \left[(11) * \left(-\frac{8}{3} \right) \right]$</p>				
Options				
<p>Key: I2 Distractors: Value range: [-100] - [100]</p>				
Auxiliary Information				
ไม่มี				
Key				
3				

ภาพ 2 ตัวอย่างแบบประเมินความถูกต้องของโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

จากภาพ 2 แบบประเมินความถูกต้องของโมเดลข้อสอบผู้วิจัยได้แสดงความคิดเห็นไว้ในแต่ละข้อของโมเดลข้อสอบหลังจากได้ทำการวิเคราะห์ตามรูปแบบของโจทย์ และตัวเลือกลักษณะต่าง ๆ โดยใช้รูปแบบของ Gierl et al. (2008) จำนวน 10 รูปแบบ เป็นโมเดลเปรียบเทียบ และผู้เชี่ยวชาญจะทำการพิจารณา มาตรฐานประเมินความถูกต้อง 4 ระดับ ในแต่ละข้อของโมเดลข้อสอบ โดยใช้รูปแบบของ Gierl et al. (2008) จำนวน 10 รูปแบบ เป็นโมเดลเปรียบเทียบ และมีข้อเสนอแนะเพิ่มเติมเพื่อการพัฒนาโมเดลข้อสอบให้ดียิ่งขึ้น

3. การเก็บรวบรวมข้อมูล

3.1 รวบรวมผลการตอบข้อสอบมาวิเคราะห์คุณภาพข้อสอบตามหลักการทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ ด้วยโปรแกรม Xcalibre™ 4.2.2 (Guyer & Thompson, 2014) และทำการคัดเลือกข้อสอบตามเกณฑ์การคัดเลือกข้อสอบ (Urry, 1977) ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a) มีค่าตั้งแต่ 0.50 ถึง 2.50 ค่าความยากของข้อสอบ (b) มีค่าตั้งแต่ -2.50 ถึง 2.50 และค่าพารามิเตอร์การเดาของข้อสอบ (c) มีค่าไม่เกิน 0.30

3.2 จำแนกระดับความยากของข้อสอบที่ผ่านเกณฑ์การคัดเลือกเป็น 3 ระดับ คือ ง่าย ปานกลาง และยาก โดยให้ผู้เชี่ยวชาญทำการประเมินความตรงเชิงโครงสร้างตามหลักสูตรวิชาคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ประกอบด้วย 3 สาระ คือ สาระที่ 1 จำนวนและพีชคณิต สาระที่ 2 การวัดและเรขาคณิต และสาระที่ 3 สถิติและความน่าจะเป็น

3.3 วิเคราะห์ข้อสอบแต่ละข้อเกี่ยวกับความเหมาะสมในการพัฒนาเป็นโมเดลข้อสอบโดยพิจารณา เนื้อความของข้อสอบแต่ละข้อว่ามีส่วนที่สามารถกำหนดเป็นส่วนคงที่ และส่วนที่แปรเปลี่ยนได้หรือไม่ ถ้าข้อสอบข้อใด ไม่เหมาะสมจะคัดข้อสอบข้อนั้นออกไป โดยใช้หลักการของ Gierl et al. (2008) ที่ได้แบ่งประเภทโมเดลของข้อสอบเป็น 2 ประเภท ดังนี้

ประเภทโจทย์ หรือข้อคำถาม ในส่วนนี้สามารถแบ่งออกเป็น 4 รูปแบบ ดังนี้

1) องค์ประกอบที่แปรเปลี่ยนได้อย่างอิสระ (independent element) มีจำนวน 1 องค์ประกอบ เมื่อองค์ประกอบใดในโจทย์มีการเปลี่ยนแปลง จะไม่ส่งผลกระทบต่อองค์ประกอบอื่น ๆ

2) องค์ประกอบที่แปรเปลี่ยนได้อย่างไม่อิสระ (dependent element) องค์ประกอบมีความแตกต่างขึ้นอยู่กับสถานะองค์ประกอบอื่น ๆ มีจำนวน 2 องค์ประกอบแต่ละองค์ประกอบมีการเชื่อมโยงกับ องค์ประกอบอื่น

3) การผสมกันระหว่างองค์ประกอบที่แปรเปลี่ยนได้อย่างอิสระและแปรเปลี่ยนได้อย่างไม่อิสระ (mixed independent/ dependent element (s) โดยมีจำนวนองค์ประกอบอิสระ 1 องค์ประกอบ และ จำนวนองค์ประกอบไม่อิสระ 1 องค์ประกอบ

4) องค์ประกอบค่าคงที่ (fixed element) มีลักษณะคงที่ จะไม่มีการเปลี่ยนแปลง

ประเภทที่ 2 ตัวเลือกของโมเดลข้อสอบ แบ่งออกเป็น 3 รูปแบบ ดังนี้

1) ตัวเลือกที่ได้จากการสุ่ม (randomly selected options) คือ รูปแบบที่ทั้งคำตอบที่ถูกต้อง และตัวลวงถูกเลือกแบบสุ่มจากคลังที่มีเนื้อหาตรงกัน

2) ตัวเลือกแบบจำกัด (constrained options) คือ รูปแบบที่ทั้งคำตอบที่ถูกต้อง และตัวลวงถูกสร้างตามข้อกำหนด เช่น สูตร และการคำนวณ เป็นต้น

3) ตัวเลือกแบบคงที่ (fixed options) คือ รูปแบบที่ทั้งคำตอบที่ถูกต้อง และตัวลวงไม่มีการเปลี่ยนแปลง

การสลับกันของประเภทของโจทย์ และตัวเลือก ทำให้เกิดโมเดลข้อสอบรูปแบบต่าง ๆ ดังตาราง 1

ตาราง 1 รูปแบบของโจทย์ และตัวเลือกลักษณะต่าง ๆ ในโมเดลข้อสอบ

ตัวเลือก (options)	โจทย์ (stem)			แบบคงที่ (fixed)
	แบบอิสระ (independent)	แบบไม่อิสระ (dependent)	แบบผสม (mixed)	
แบบเลือกจากการสุ่ม (randomly selected)	✓	✓	✓	✓
แบบจำกัด (constrained)	✓	✓	✓	รูปแบบที่ไม่สามารถสร้างได้ (N/A)
แบบคงที่ (fixed)	✓	✓	✓	รูปแบบที่ไม่สามารถสร้างได้ (N/A)

จากตาราง 1 จะเห็นว่ามีการสร้างโมเดลข้อสอบได้จำนวน 10 รูปแบบ จากการจับคู่ระหว่างโจทย์กับตัวเลือกประเภทต่าง ๆ และมี 2 รูปแบบที่ไม่สามารถสร้างได้ (N/A)

3.4 ตรวจสอบความถูกต้องของโมเดลข้อสอบที่สร้างขึ้นโดยผู้วิจัยทำหนังสือขอความอนุเคราะห์ถึงผู้เชี่ยวชาญ 3 ท่าน เป็นผู้ทรงคุณวุฒิ มีวุฒิการศึกษาไม่ต่ำกว่าปริญญาเอก หรือมีประสบการณ์เกี่ยวกับการสอนคณิตศาสตร์ ตั้งแต่ 30 ปี ขึ้นไป การประเมินความถูกต้องของโมเดลข้อสอบจะพิจารณาข้อสอบที่ละข้อว่ามีความถูกต้อง เน้นที่ระดับความเห็นด้วยของผู้เชี่ยวชาญต่อข้อสอบข้อนั้น ๆ

4. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลเป็นการนำผลของการประเมินความถูกต้องของโมเดลข้อสอบที่ผู้เชี่ยวชาญพิจารณาข้อสอบที่ละข้อว่ามีความถูกต้อง โดยเน้นที่ระดับความเห็นด้วยของผู้เชี่ยวชาญต่อข้อสอบข้อนั้น ๆ โดยเปรียบเทียบกับรูปแบบของโจทย์ และตัวเลือกลักษณะต่าง ๆ ในโมเดลข้อสอบตามรูปแบบของ Gierl et al. (2008) จำนวน 10 รูปแบบ เป็นโมเดลเปรียบเทียบ แล้วนำมาคำนวณค่าดัชนีความตรงเชิงเนื้อหาของข้อสอบ (content validity index) ค่าดัชนีความตรงเชิงเนื้อหา มี 2 ประเภท คือ ค่าดัชนีความตรงเชิงเนื้อหาของข้อคำถามรายข้อ (item-level CVI) และค่าดัชนีความตรงเชิงเนื้อหาของเครื่องมือทั้งชุด (scale-level CVI) ผู้เชี่ยวชาญจะพิจารณามาตรฐานประเมินความถูกต้อง 4 ระดับ แต่ละระดับมีความหมาย ดังนี้ 4 หมายถึง โมเดลข้อสอบมีความถูกต้อง 3 หมายถึง โมเดลข้อสอบต้องปรับปรุงเล็กน้อยจึงถูกต้อง 2 หมายถึง โมเดลข้อสอบต้องปรับปรุงมากจึงถูกต้อง และ 1 หมายถึง โมเดลข้อสอบไม่ถูกต้อง

เกณฑ์การพิจารณาค่าดัชนีความตรงเชิงเนื้อหาของข้อคำถามรายข้อ (item-level CVI) และค่าดัชนีความตรงเชิงเนื้อหาของเครื่องมือทั้งชุด (scale-level CVI) มีดังนี้

Polit and Beck (2006) ใช้คำย่อแทนค่าทั้งสองว่า I-CVI และ S-CVI โดยค่า I-CVI ที่ดีควรมีค่ามากกว่า 0.80 และข้อสอบที่มีความสมบูรณ์ ค่า I-CVI จะเท่ากับ 1 (Polit & Beck, 2006) เมื่อผู้เชี่ยวชาญประเมินเรียบร้อยแล้ว ผู้วิจัยจะทำการนับข้อสอบที่ผู้เชี่ยวชาญแต่ละคนให้คะแนน โดยนับเฉพาะข้อที่ได้ 3 และ 4 คะแนน เท่านั้น มาคำนวณหาค่า I-CVI ดังนี้

$$I - CVI = \frac{N_c}{N}$$

โดย N_c หมายถึง จำนวนผู้เชี่ยวชาญที่ประเมินข้อคำถามในระดับความถูกต้องระดับ 3 และ 4

N หมายถึง จำนวนผู้เชี่ยวชาญทั้งหมด

ข้อสอบที่ได้คะแนน 1 และ 2 คะแนน ผู้วิจัยจะทำการแก้ไขปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญหรือตัดออกไป หลังจากการพิจารณาของผู้ทรงคุณวุฒิจนได้ค่า I-CVI ตามเกณฑ์ทุกข้อแล้ว จึงคำนวณค่า S-CVI โดยหาค่าเฉลี่ยจากผลรวมของค่า I-CVI หารด้วยจำนวนข้อคำถาม ซึ่งค่า S-CVI ควรได้อย่างน้อย 0.90 (Polit & Beck, 2006)

ระยะที่ 2 การพัฒนาระบบสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

1. กลุ่มตัวอย่าง

ประชากร คือ ครูกลุ่มสาระการเรียนรู้คณิตศาสตร์ ชั้นมัธยมศึกษา ในโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน

กลุ่มตัวอย่าง ครูกลุ่มสาระการเรียนรู้คณิตศาสตร์ ชั้นมัธยมศึกษา ในโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน จำนวน 40 คน ได้มาโดยการเลือกแบบเจาะจง (purposive sampling) เป็นการเลือกกลุ่มตัวอย่างโดยพิจารณาจากการตัดสินใจของผู้วิจัย โดยลักษณะของกลุ่มตัวอย่างที่เลือกเป็นไปตามวัตถุประสงค์ของการวิจัย คือ เป็นครูกลุ่มสาระการเรียนรู้คณิตศาสตร์ที่สอนในชั้นมัธยมศึกษา ได้แก่ โรงเรียนขนาดเล็ก 1 แห่ง ได้แก่ โรงเรียนห้วยผึ้งพิทยา จังหวัดกาฬสินธุ์ ครูกลุ่มตัวอย่าง 3 คน โรงเรียนขนาดปานกลาง 1 แห่ง ได้แก่ โรงเรียนวิทยาศาสตร์จุฬารามรราชวิทยาลัย เพชรบุรี จังหวัดเพชรบุรี ครูกลุ่มตัวอย่าง 8 คน และโรงเรียนขนาดใหญ่ 2 แห่ง ได้แก่ โรงเรียนสุรวิทยาคาร จังหวัดสุรินทร์ ครูกลุ่มตัวอย่าง 11 คน และโรงเรียนสรรพวิทยาคม จังหวัดตาก ครูกลุ่มตัวอย่าง 18 คน รวมครูกลุ่มตัวอย่างจำนวน 40 คน

2. เครื่องมือที่ใช้ในการวิจัย

2.1 แบบประเมินความเหมาะสมของโปรแกรมโดยผู้เชี่ยวชาญ

แบบประเมินความเหมาะสมของโปรแกรมโดยผู้เชี่ยวชาญ ประกอบด้วย 5 ด้าน ได้แก่ 1) ความตรงตามความต้องการของโปรแกรมฯ 2) การทำงานของโปรแกรมฯ 3) การใช้งาน 4) การรักษาความปลอดภัยและการตรวจสอบความถูกต้องในการเข้าใช้โปรแกรมฯ และ 5) ความชัดเจนของคู่มือการใช้โปรแกรมฯ

แบบประเมินที่พัฒนาขึ้นมีลักษณะเป็นมาตราประมาณค่า 5 ระดับ โดยมีเกณฑ์การให้คะแนน ดังนี้ 5 หมายถึง โปรแกรมมีความเหมาะสมมากที่สุด 4 หมายถึง โปรแกรมมีความเหมาะสมมาก 3 หมายถึง โปรแกรมมีความเหมาะสมปานกลาง 2 หมายถึง โปรแกรมมีความเหมาะสมน้อย และ 1 หมายถึง โปรแกรมมีความ

เหมาะสมน้อยที่สุด จากนั้นนำข้อมูลวิเคราะห์โดยใช้โปรแกรมสำเร็จรูป SPSS เพื่อหาค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD) และรายงานผลการประเมิน

แปลความหมาย ดังนี้ คะแนนเฉลี่ย 4.5–5.00 แปลความว่า มีความเหมาะสมในระดับมากที่สุด คะแนนเฉลี่ย 3.51–4.50 แปลความว่า มีความเหมาะสมในระดับมาก คะแนนเฉลี่ย 2.51–3.50 แปลความว่า มีความเหมาะสมในระดับปานกลาง คะแนนเฉลี่ย 1.51–2.50 แปลความว่า มีความเหมาะสมในระดับน้อย และคะแนนเฉลี่ย 1.00–1.50 แปลความว่า มีความเหมาะสมในระดับน้อยที่สุด

2.2 แบบประเมินประสิทธิภาพของโปรแกรมโดยผู้ใช้งาน

แบบประเมินประสิทธิภาพของโปรแกรมโดยผู้ใช้งาน ประกอบด้วย 2 ด้าน ได้แก่ 1) ลักษณะทั่วไปของโปรแกรมฯ และ 2) ความสะดวกในการใช้โปรแกรมฯ โดยแบบประเมินที่พัฒนาขึ้นมีลักษณะเป็นมาตราประมาณค่า 5 ระดับ โดยมีเกณฑ์การให้คะแนน ดังนี้ 5 หมายถึง โปรแกรมมีประสิทธิภาพมากที่สุด 4 หมายถึง โปรแกรมมีประสิทธิภาพมาก 3 หมายถึง โปรแกรมมีประสิทธิภาพปานกลาง 2 หมายถึง โปรแกรมมีประสิทธิภาพน้อย และ 1 หมายถึง โปรแกรมมีประสิทธิภาพน้อยที่สุด จากนั้นนำข้อมูลวิเคราะห์โดยใช้โปรแกรมสำเร็จรูป SPSS เพื่อหาค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (standard deviation: SD) และรายงานผลการประเมิน

แปลความหมาย ดังนี้ คะแนนเฉลี่ย 4.51–5.00 แปลความว่า โปรแกรมดีมาก คะแนนเฉลี่ย 3.51–4.50 แปลความว่า โปรแกรมดี คะแนนเฉลี่ย 2.51–3.50 แปลความว่า โปรแกรมพอใช้ คะแนนเฉลี่ย 1.51–2.50 แปลความว่า โปรแกรมควรปรับปรุง และคะแนนเฉลี่ย 1.00–1.50 แปลความว่า โปรแกรมควรปรับปรุงอย่างยิ่ง

3. การเก็บรวบรวมข้อมูล

3.1 วิเคราะห์และออกแบบระบบ โดยจัดทำแผนภาพบริบท (context diagram) เพื่อวิเคราะห์ภาพรวมของระบบทั้งหมด และจัดทำผังการไหลของข้อมูล (Data Flow Diagram: DFD) เพื่อแสดงรายละเอียดของกระบวนการทำงานหลักของระบบ

3.2 ออกแบบฐานข้อมูล ในการจัดเก็บข้อมูลที่มีความสัมพันธ์กันอย่างเป็นระบบ เพื่อให้ผู้ใช้ระบบสามารถเรียกใช้งานข้อมูลเหล่านั้นร่วมกันได้

3.3 พัฒนาโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก ในรูปแบบของเว็บแอปพลิเคชัน เป็นขั้นตอนที่นำรูปแบบโครงสร้างของหน้าจอต่าง ๆ ที่ออกแบบไว้ และข้อสอบที่สร้างเป็นโมเดลข้อสอบไว้แล้วมาพัฒนาโปรแกรมตามวัตถุประสงค์ ผู้วิจัยได้พัฒนาโปรแกรมคอมพิวเตอร์ ตามหลักการของวงจรการพัฒนาาระบบ (System Development Life Cycle: SDLC) ของ Preechapanich (2014) คือ 1) การกำหนด และวิเคราะห์ปัญหา และ 2) การเขียนผังงาน และชุดโค้ด

3.4 ทดสอบและแก้ไขข้อผิดพลาดของโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก ผู้วิจัยทดสอบโปรแกรมโดยใช้เทคนิคการทดสอบแบบกล่องดำ (black box testing) ซึ่งเป็นการทดสอบที่เน้นผลลัพธ์ที่เกิดขึ้น (output) จากการประมวลผลโปรแกรม (process)

โดยไม่เน้นรูปแบบการเขียนโปรแกรม เพื่อหาจุดบกพร่องเบื้องต้นก่อน หลังจากนั้นจึงนำโปรแกรมดังกล่าวไปทำการประเมินความเหมาะสมของโปรแกรมโดยผู้เชี่ยวชาญ เพื่อตรวจสอบความถูกต้อง และค้นหาจุดบกพร่องที่อาจเกิดขึ้น จากนั้นจึงปรับปรุงแก้ไขโปรแกรมจนมีความสมบูรณ์มากขึ้น

3.5 จัดทำคู่มือการใช้โปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก เป็นขั้นตอนการจัดทำคู่มือการใช้โปรแกรม เพื่ออธิบายวิธีการใช้โปรแกรมและเป็นแนวทางการนำไปใช้จริง

3.6 การประเมินความเหมาะสมของโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก เป็นขั้นตอนที่นำโปรแกรมและคู่มือการใช้งานไปให้ผู้เชี่ยวชาญประเมินความเหมาะสมของโปรแกรม ผู้วิจัยทำหนังสือขอความอนุเคราะห์ถึงผู้เชี่ยวชาญ 4 ท่าน โดยผู้เชี่ยวชาญมีวุฒิการศึกษาไม่ต่ำกว่าปริญญาเอก มีประสบการณ์ไม่น้อยกว่า 10 ปี

3.7 การประเมินประสิทธิภาพในการใช้งานของโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก โดยการประเมินความคิดเห็นของผู้ทดลองใช้โปรแกรมกลุ่มตัวอย่างเป็นครูกลุ่มสาระการเรียนรู้คณิตศาสตร์ ชั้นมัธยมศึกษา จำนวน 40 คน

4. การวิเคราะห์ข้อมูล

วิเคราะห์ข้อมูลการประเมินความเหมาะสมของโปรแกรมโดยผู้เชี่ยวชาญ และวิเคราะห์ข้อมูลการประเมินประสิทธิภาพของโปรแกรมโดยผู้ใช้งาน โดยการนำคะแนนที่ได้จากการประเมินโปรแกรมมาทำการวิเคราะห์ข้อมูลโดยใช้สถิติพื้นฐาน คือ ค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (SD) ด้วยโปรแกรมสำเร็จรูป และแปลความหมาย

ผลการวิจัย

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ 1) การสร้างโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 และ 2) พัฒนาระบบสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

1. ผลการสร้างโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

โมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ที่สร้างขึ้นประกอบด้วยรายละเอียด 5 ส่วน ได้แก่ 1) โจทย์ (stem) 2) ส่วนประกอบ (element) 3) ตัวเลือก (options) 4) ข้อมูลเสริม (auxiliary information) และ 5) เฉลย (key) รายละเอียด ดังนี้

1) โจทย์ (stem) เป็นส่วนสำคัญที่ประกอบด้วยข้อมูล เนื้อหา ข้อคำถามที่จำเป็นกับผู้ทดสอบในการตอบแบบทดสอบ โดยแบ่งข้อมูลเป็น 2 ส่วน คือ ส่วนคงที่ และส่วนที่แปรเปลี่ยนได้ โดยวิเคราะห์หาข้อมูลใดที่จะกำหนดให้เป็นส่วนที่แปรเปลี่ยนได้ ให้กำหนดเป็นตัวแปร ดังนี้

1.1) ข้อความ กำหนดเป็น ตัวแปร S และถ้าโจทย์มีข้อความหลายส่วนให้กำกับด้วยตัวเลข เช่น S1, S2 หรือ S3 เป็นต้น อาจมีบางกรณีที่ต้องมีเงื่อนไขเกี่ยวกับข้อความ ซึ่งจะตั้งเป็นตัวแปร As_S และกำกับด้วยตัวเลข เช่น As_S1, As_S2 หรือ As_S3 เป็นต้น

1.2) จำนวน กำหนดเป็น ตัวแปร I และถ้าโจทย์มีจำนวนหลายส่วน ให้กำกับด้วยตัวเลข เช่น I1, I2 หรือ I3 เป็นต้น อาจมีบางกรณีที่ต้องมีเงื่อนไขหรือช่วงของจำนวน ซึ่งจะตั้งเป็นตัวแปร As_I และกำกับด้วยตัวเลข เช่น As_I1, As_I2 หรือ As_I3 เป็นต้น

1.3) รูปภาพประกอบ ให้กำหนดเป็นตัวแปร P และถ้าในโจทย์มีรูปภาพหลายส่วน ให้กำกับด้วยตัวเลข เช่น P1, P2 หรือ P2 เป็นต้น อาจมีบางกรณีที่ต้องมีเงื่อนไขเกี่ยวกับรูปภาพ ซึ่งจะตั้งเป็นตัวแปร As_P และกำกับด้วยตัวเลข เช่น As_P1, As_P2 หรือ As_P3 เป็นต้น

รายละเอียดของการกำหนดเกี่ยวกับค่าทั้งหมดที่สามารถแปรเปลี่ยนได้ ได้แก่ ค่าของตัวแปร S ตัวแปร I ตัวแปร P ตัวแปร As_S ตัวแปร As_I และตัวแปร As_P นั้น จะระบุไว้ในช่วงของ Elements

2) ส่วนประกอบ (element) จะมีรายละเอียดของการกำหนดเกี่ยวกับค่าทั้งหมดที่สามารถแปรเปลี่ยนได้ ได้แก่ ค่าของตัวแปร S ตัวแปร I ตัวแปร P ตัวแปร As_S ตัวแปร As_I และตัวแปร As_P โดยกำหนดเป็นช่วงข้อมูลที่ผู้วิจัยเห็นว่าเหมาะสมสำหรับโจทย์แต่ละข้อ เช่น S1 เป็นตัวแปรแทนชื่อคน อาจกำหนดให้ S1 สุ่มเลือกจาก ‘เปา’ ‘ไก่อ’ ‘กิก’ และ ‘วิน’ เป็นต้น หรือ I1 เป็นตัวแปรแทนค่าเงิน ในโจทย์อาจกำหนดให้ I1 สุ่มเลือกค่าได้ตั้งแต่ 100 ถึง 10,000 เป็นต้น หรือ P1 เป็นตัวแปรแทนรูปภาพ อาจกำหนดให้ P1 สุ่มเลือกจาก ‘Px’ ‘Py’ ‘Pz’ และ ‘Pw’ เป็นต้น

3) ตัวเลือก (Options) ประกอบด้วยคำตอบจำนวน 4 ตัวเลือก โดยมีเพียงหนึ่งข้อที่ถูกต้องที่สุด และที่เหลือเป็นตัวลวง ถ้าตัวเลือกนั้นเป็นจำนวนจะกำหนดเป็นตัวแปร I ถ้ามีหลายส่วนจะมีตัวเลขกำกับ เช่น I1, I2, I3 หรือเป็นตัวแปร As_I เมื่อมีเงื่อนไขที่ซับซ้อน เป็นต้น และในทำนองเดียวกัน ถ้าตัวเลือกนั้นเป็นข้อความ หรือรูปภาพ จะกำหนดเป็นตัวแปร S หรือตัวแปร P ตามลำดับ

4) ข้อมูลเสริม (auxiliary Information) เป็นข้อมูลส่วนย่อยที่อาจใส่ไว้ในส่วนของโจทย์ (item) หรือส่วนของตัวเลือก (options) ซึ่งเป็นได้ทั้ง ข้อความ รูปภาพ ตาราง แผนภาพ เสียง หรือ วิดีโอ สำหรับงานวิจัยนี้ข้อสอบในการนำมาสร้างเป็นโมเดลข้อสอบนั้น ส่วนใหญ่จะเป็นรูปภาพจึงกำหนดเป็นตัวแปร P และถ้ามีรูปร่างหลายส่วน ให้กำกับด้วยตัวเลข เช่น P1, P2, P3

5) เฉลย (key) เป็นตัวเลือกที่ถูกต้องที่สุด โดยระบุเพียงตัวเลือกเดียวในข้อคำถามแต่ละข้อ

ผลการสร้างโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ผู้วิจัยนำเสนอผลการวิจัยเป็น 2 ประเด็น ดังนี้ 1) การวิเคราะห์ข้อสอบแต่ละข้อเพื่อความเหมาะสมในการสร้างเป็นโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 และ 2) การตรวจสอบความถูกต้องของโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 มีรายละเอียด ดังนี้

1.1 การวิเคราะห์ข้อสอบแต่ละข้อเพื่อความเหมาะสมในการสร้างเป็นโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 แสดงผลดังตาราง 2

ตาราง 2 ผลการวิเคราะห์ความเหมาะสมในการสร้างเป็นโมเดลข้อสอบ วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

โมเดล	รูปแบบของโจทย์ และตัวเลือก ในโมเดลข้อสอบ ตามหลักการ ของ Gierl et al. (2008)	ระดับความยากของข้อสอบ									รวม (ข้อ)
		ระดับง่าย 54 ข้อ			ระดับปานกลาง 71 ข้อ			ระดับยาก 78 ข้อ			
		สาระการเรียนรู้			สาระการเรียนรู้			สาระการเรียนรู้			
		สาระที่ 1 จำนวน(ข้อ)	สาระที่ 2 จำนวน(ข้อ)	สาระที่ 3 จำนวน(ข้อ)	สาระที่ 1 จำนวน(ข้อ)	สาระที่ 2 จำนวน(ข้อ)	สาระที่ 3 จำนวน(ข้อ)	สาระที่ 1 จำนวน(ข้อ)	สาระที่ 2 จำนวน(ข้อ)	สาระที่ 3 จำนวน(ข้อ)	
1	โจทย์มีลักษณะเป็นแบบอิสระ ตัวเลือกเป็นแบบเลือกจากการสุ่ม	7	14	3	5	14	5	4	17	4	73
2	โจทย์มีลักษณะเป็นแบบอิสระ ตัวเลือกเป็นแบบจำกัด	12	-	2	23	5	3	16	6	2	69
3	โจทย์มีลักษณะเป็นแบบอิสระ ตัวเลือกเป็นแบบคงที่	-	-	-	-	1	-	-	-	-	1
4	โจทย์มีลักษณะเป็นแบบไม่อิสระ ตัวเลือกเป็นแบบเลือกจากการสุ่ม	1	1	-	-	-	-	1	1	1	5
5	โจทย์มีลักษณะเป็นแบบไม่อิสระ ตัวเลือกเป็นแบบจำกัด	4	-	-	1	-	-	1	1	-	7
6	โจทย์มีลักษณะเป็นแบบไม่อิสระ ตัวเลือกเป็นแบบคงที่	-	1	-	-	-	-	-	-	-	1
7	โจทย์มีลักษณะเป็นแบบผสม ตัวเลือกแบบเลือกจากการสุ่ม	-	-	-	2	1	-	3	2	3	11
8	โจทย์มีลักษณะเป็นแบบผสม ตัวเลือกเป็นแบบจำกัด	5	-	-	8	-	1	11	-	-	25
9	โจทย์มีลักษณะเป็นแบบผสม ตัวเลือกเป็นแบบคงที่	-	-	-	-	-	-	1	-	-	1
10	โจทย์มีลักษณะเป็นแบบคงที่ ตัวเลือกเป็นแบบเลือกจากการสุ่ม	2	2	-	1	1	-	3	1	-	10
รวม		31	18	5	40	22	9	40	28	10	203

จากตาราง 2 ข้อสอบระดับง่าย จำนวน 54 ข้อ มีความเหมาะสมในการสร้างเป็นโมเดลข้อสอบ ประกอบด้วย สารที่ 1 จำนวนและพีชคณิต จำนวน 31 ข้อ สารที่ 2 การวัดและเรขาคณิต จำนวน 18 ข้อ และสารที่ 3 สถิติและความน่าจะเป็น จำนวน 5 ข้อ และสามารถสร้างโมเดลข้อสอบได้ 7 รูปแบบ ได้แก่ โมเดลที่ 1, 2, 4, 5, 6, 8 และ 10 ข้อสอบระดับปานกลาง จำนวน 71 ข้อ มีความเหมาะสมในการสร้างเป็นโมเดลข้อสอบ ประกอบด้วย สารที่ 1 จำนวนและพีชคณิต จำนวน 40 ข้อ สารที่ 2 การวัดและเรขาคณิต จำนวน 22 ข้อ และสารที่ 3 สถิติและความน่าจะเป็น จำนวน 9 ข้อ และสามารถสร้างโมเดลข้อสอบได้ 7 รูปแบบ ได้แก่ โมเดลที่ 1, 2, 3, 5, 7, 8, และ 10 และข้อสอบระดับยาก จำนวน 78 ข้อ มีความเหมาะสม ในการสร้างเป็นโมเดลข้อสอบประกอบด้วย สารที่ 1 จำนวนและพีชคณิต จำนวน 40 ข้อ สารที่ 2 การวัดและเรขาคณิต จำนวน 28 ข้อ และสารที่ 3 สถิติและความน่าจะเป็น จำนวน 10 ข้อ และสร้างโมเดลข้อสอบได้ 8 รูปแบบ ได้แก่ โมเดลที่ 1, 2, 4, 5, 7, 8, 9 และ 10 รวมทั้ง 3 ระดับ มีจำนวนโมเดลข้อสอบทั้งสิ้น 203 ข้อ และมีจำนวนรูปแบบของโมเดลข้อสอบทั้งสิ้น 10 รูปแบบ

1.2 การตรวจสอบความถูกต้องของโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 แสดงผลดังตาราง 3

ตาราง 3 ผลการตรวจสอบความถูกต้องของโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

ดัชนีตรวจสอบ ความถูกต้อง ของโมเดลข้อสอบ	ค่าดัชนีตรวจสอบความถูกต้องของโมเดลข้อสอบ			เกณฑ์การ พิจารณา	ผลการตรวจสอบ ความถูกต้อง
	ข้อสอบ ระดับง่าย (54 ข้อ)	ข้อสอบ ระดับปานกลาง (71 ข้อ)	ข้อสอบ ระดับยาก (78 ข้อ)		
I-CVI	1.00	1.00	1.00	มากกว่า 0.80 (> 0.80)	ผ่านเกณฑ์
S-CVI	1.00	1.00	1.00	ไม่น้อยกว่า 0.90 (≥ 0.90)	ผ่านเกณฑ์

จากตาราง 3 ข้อสอบระดับง่าย มีจำนวนทั้งหมด 54 ข้อ โดยมีค่าดัชนีความตรงเชิงเนื้อหาของข้อคำถามรายข้อ (I-CVI) เท่ากับ 1.00 และค่าดัชนีความตรงเชิงเนื้อหาของเครื่องมือทั้งหมด (S-CVI) เท่ากับ 1.00 ผลการตรวจสอบโมเดลข้อสอบมีความถูกต้องผ่านเกณฑ์ ข้อสอบระดับปานกลาง มีจำนวนทั้งหมด 71 ข้อ โดยมีค่าดัชนีความตรงเชิงเนื้อหาของข้อคำถามรายข้อ (I-CVI) เท่ากับ 1.00 และค่าดัชนีความตรงเชิงเนื้อหาของเครื่องมือทั้งหมด (S-CVI) เท่ากับ 1.00 ผลการตรวจสอบโมเดลข้อสอบมีความถูกต้องผ่านเกณฑ์ และข้อสอบระดับยาก มีจำนวนทั้งหมด 78 ข้อ โดยมีค่าดัชนีความตรงเชิงเนื้อหาของข้อคำถามรายข้อ (I-CVI) เท่ากับ 1.00 และค่าดัชนีความตรงเชิงเนื้อหาของเครื่องมือทั้งหมด (S-CVI) เท่ากับ 1.00 ผลการตรวจสอบโมเดลข้อสอบมีความถูกต้องผ่านเกณฑ์ โดยรวมโมเดลข้อสอบมีความถูกต้องผ่านเกณฑ์

2. ผลการพัฒนากระบวนการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

การพัฒนากระบวนการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหา และระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ผู้วิจัยนำเสนอผลการวิจัยเป็น 2 ประเด็น ดังนี้ 1) ผลการพัฒนากระบวนการในรูปแบบเว็บแอปพลิเคชัน โดยใช้งานผ่านระบบอินเทอร์เน็ต และ 2) ผลการประเมินโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 โดยผู้ใช้งาน

2.1 ผลการพัฒนากระบวนการในรูปแบบเว็บแอปพลิเคชัน โดยใช้งานผ่านระบบอินเทอร์เน็ต

การใช้งานสามารถใช้งานผ่านระบบอินเทอร์เน็ตได้ที่ <https://www.aig-system.com> มีรายละเอียด ดังนี้

1) รูปแบบของเว็บแอปพลิเคชัน ใช้ภาษา HTML (Hypertext Markup Language) เป็นภาษาหลักที่ใช้ในการสร้างโฮมเพจร่วมกับภาษา CSS (Cascading Style Sheets) ที่เป็นภาษาที่มีรูปแบบการเขียน syntax ที่มีความเฉพาะในการกำหนดรูปแบบเอกสารเว็บหรือตกแต่งเอกสาร HTML ให้มีหน้าตา สี สัน ตัวอักษร ตำแหน่ง เส้นขอบ พื้นหลัง ระยะห่าง โดยการกำหนดคุณสมบัติให้กับ element ต่าง ๆ ของ HTML และใช้ JavaScript สำหรับงานการคำนวณ การแสดงผลการรับ-ส่งข้อมูล และสามารถโต้ตอบกับผู้ใช้ได้อย่างทันที

2) โปรแกรมจะทำงานบนคลาวด์คอมพิวติ้ง (cloud computing) ในรูปแบบ Platform as a Service (PaaS) ทั้งในส่วนของเว็บแอปพลิเคชัน และส่วนของการจัดการดูแลข้อมูลต่าง ๆ นอกจากนี้ โปรแกรมได้ถูกออกแบบมาให้เหมาะสำหรับการใช้งานกับระบบปฏิบัติการทางคอมพิวเตอร์ เช่น คอมพิวเตอร์ตั้งโต๊ะ (desktop computer) คอมพิวเตอร์แบบพกพา (notebook) และแท็บเล็ตคอมพิวเตอร์ (tablet computer)

3) การใช้งานโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก ได้จัดทำคู่มือการใช้งานของโปรแกรมฯ โดยแบ่งออกเป็น 7 ส่วน ได้แก่ 1) การลงทะเบียน 2) การสร้างข้อสอบ 3) การรายงานผลการสร้างข้อสอบ 4) คู่มือการใช้งานโปรแกรม 5) ข้อมูลอ้างอิง 6) ติดต่อสอบถาม และ 7) ข้อมูลผู้ใช้ แสดงดังภาพ 3



ภาพ 3 หน้าจอโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก

4) การประเมินความเหมาะสมของโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบ
จำแนกตามเนื้อหาและระดับความยากโดยผู้เชี่ยวชาญ แสดงผลดังตาราง 4

ตาราง 4 ผลการประเมินความเหมาะสมของโปรแกรมโดยผู้เชี่ยวชาญ

รายการประเมิน	Mean	SD	ระดับความเหมาะสม
1. ด้านความตรงตามความต้องการของโปรแกรม	4.50	0.52	มาก
2. ด้านการทำงานของโปรแกรม	4.25	0.57	มาก
3. ด้านการใช้งาน	4.18	0.48	มาก
4. ด้านการรักษาความปลอดภัยและการตรวจสอบ ความถูกต้องในการเข้าใช้โปรแกรม	4.29	0.97	มาก
5. ด้านความชัดเจนของคู่มือการใช้โปรแกรม	4.45	0.56	มาก
สรุป	4.33	0.54	มาก

จากตาราง 4 แสดงผลการประเมินโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบ
จำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 โดยผู้เชี่ยวชาญ แสดงให้เห็นว่า
โปรแกรมในภาพรวมมีความเหมาะสมอยู่ในระดับมาก (Mean = 4.33, SD = 0.54) เมื่อพิจารณาเป็นรายด้าน
พบว่า ทั้ง 5 ด้าน มีความเหมาะสมอยู่ในระดับมาก ได้แก่ ด้านความตรงตามความต้องการของโปรแกรม (Mean =
4.50, SD = 0.52) ด้านความชัดเจนของคู่มือการใช้โปรแกรม (Mean = 4.45, SD = 0.57) การรักษาความปลอดภัย
และการตรวจสอบความถูกต้องในการเข้าใช้โปรแกรม (Mean = 4.29, SD = 0.48) ด้านการทำงานของโปรแกรม
(Mean = 4.25, SD = 0.97) และด้านการใช้งาน (Mean = 4.18, SD = 0.56)

2.2 การประเมินประสิทธิภาพของโปรแกรมโดยผู้ใช้งาน

การประเมินประสิทธิภาพของโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบ
จำแนกตามเนื้อหาและระดับความยากใช้โดยผู้ใช้งาน แสดงผลดังตาราง 5

ตาราง 5 ผลการประเมินประสิทธิภาพของโปรแกรมโดยผู้ใช้งาน

รายการประเมิน	Mean	SD	ระดับความคิดเห็น
1. ด้านลักษณะทั่วไปของโปรแกรม	4.54	0.55	ดีมาก
2. ด้านความสะดวกในการใช้โปรแกรม	4.53	0.56	ดีมาก
สรุป	4.54	0.56	ดีมาก

จากตาราง 5 แสดงผลการประเมินประสิทธิภาพโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลัง
ข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 โดยผู้ใช้งาน แสดงให้เห็นว่า
โปรแกรมในภาพรวมมีประสิทธิภาพอยู่ในระดับดีมาก (Mean = 4.54, 0.56) เมื่อพิจารณาเป็นรายด้าน พบว่า
ทั้ง 2 ด้าน มีประสิทธิภาพอยู่ในระดับดีมาก ได้แก่ 1) ด้านลักษณะทั่วไปของโปรแกรม (Mean = 4.54, SD =
0.55) และ 2) ด้านความสะดวกในการใช้โปรแกรม (Mean = 4.53, SD = 0.56)

อภิปรายผล

1. การวิเคราะห์ข้อสอบแต่ละข้อเพื่อความเหมาะสมในการสร้างเป็นโมเดลข้อสอบวิชาคณิตศาสตร์
ชั้นมัธยมศึกษาปีที่ 1

ข้อสอบที่เหมาะสมในการสร้างเป็นโมเดลข้อสอบ มีจำนวนทั้งสิ้น 203 ข้อ และสามารถสร้างได้ครบ 10 รูปแบบ ตามหลักการของ Gierl et al. (2008) พบว่า รูปแบบโมเดลที่ 3 โมเดลที่ 6 และโมเดลที่ 9 มีลักษณะของตัวเลือกที่เหมือนกัน คือ ตัวเลือกเป็นแบบคงที่ ซึ่งเป็นรูปแบบโมเดลที่สร้างยาก เนื่องจากโมเดลจะทำการแปรเปลี่ยนองค์ประกอบของโจทย์เท่านั้น โดยตัวเลือกยังคงที่ไม่มีการเปลี่ยนแปลง สอดคล้องกับงานวิจัยของ Chantharamaha and Chadcham (2018) ได้ศึกษาการสร้างข้อสอบอัตโนมัติวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 โดยใช้โปรแกรมคอมพิวเตอร์ โดยข้อสอบที่นำมาสร้างเป็นโมเดลข้อสอบ มีจำนวน 44 ข้อ และสร้างโมเดล ข้อสอบตามหลักการของ Gierl et al. (2008) ได้ 5 รูปแบบ จาก 10 รูปแบบ พบว่า ไม่สามารถสร้างโมเดลที่ 6 และโมเดลที่ 9 สอดคล้องกับงานวิจัยของ Jantana (2019) ได้ศึกษาการพัฒนาโปรแกรมการสร้างข้อสอบอัตโนมัติวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 โดยข้อสอบที่นำมาสร้างเป็นโมเดลข้อสอบ มีจำนวน 50 ข้อ และสร้างโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) ได้ 8 รูปแบบ จาก 10 รูปแบบ โดยพบว่า ไม่สามารถสร้างโมเดลที่ 9 และสอดคล้องกับงานวิจัยของ Pradujprom and Panthong (2019) ได้ศึกษาการพัฒนาโปรแกรมการสร้างข้อสอบแบบหลายตัวเลือก โดยใช้ข้อสอบอัตโนมัติวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 6 โดยข้อสอบที่นำมาสร้างเป็นโมเดลข้อสอบ มีจำนวน 22 ข้อ และใช้วิธีการสร้างข้อสอบอัตโนมัติตามหลักการกำหนดรูปแบบของโจทย์และตัวเลือกตามหลักการของ Gierl et al. (2008) พบว่า ไม่สามารถสร้างโมเดลได้ครบ 10 รูปแบบ จากงานวิจัยที่กล่าวมาข้างต้น จะเห็นได้ว่าจำนวนข้อสอบที่นำมาสร้างเป็นโมเดลข้อสอบมีจำนวนข้อสอบไม่มากซึ่งเป็นสาเหตุหนึ่งที่ทำให้ไม่สามารถสร้างโมเดลได้ครบ 10 รูปแบบ และจากผลการวิจัยในครั้งนี้สามารถสร้างโมเดลได้ครบ 10 รูปแบบ มีสาเหตุมาจากปัจจัยหลายประการ ได้แก่ 1) การมีข้อสอบจำนวนมากที่ผ่านการวิเคราะห์เพื่อนำมาสร้างเป็นโมเดลข้อสอบ 2) ข้อสอบที่นำมาใช้ในการวิจัยในครั้งนี้เป็นข้อสอบที่ใช้ในการสอบแข่งขันวัดความสามารถทางคณิตศาสตร์ ทำให้ข้อสอบมีรูปแบบของโจทย์และตัวเลือกที่หลากหลาย 3) ผู้วิจัยได้ศึกษาค้นคว้าและพบว่า รูปแบบของโจทย์และตัวเลือกลักษณะต่าง ๆ ในโมเดลข้อสอบเกิดจากการสลับกันของประเภทของโจทย์และตัวเลือกเพื่อทำให้เกิดโมเดลข้อสอบรูปแบบต่าง ๆ ต้องอาศัยข้อมูล 2 ส่วน คือ ส่วนคงที่ และส่วนที่แปรเปลี่ยนได้ของโจทย์และตัวเลือกในการสร้างโมเดลข้อสอบ

2. การตรวจสอบความถูกต้องของโมเดลข้อสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

การตรวจสอบความถูกต้องของโมเดลข้อสอบที่พัฒนาขึ้นของข้อสอบวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 ให้ผู้เชี่ยวชาญประเมินความถูกต้องของโมเดลข้อสอบโดยการพิจารณาข้อสอบทีละข้อว่ามีความถูกต้องเน้นที่ระดับความเห็นด้วยของผู้เชี่ยวชาญต่อข้อสอบข้อนั้น ๆ แล้วนำมาคำนวณค่าดัชนีความตรงเชิงเนื้อหาของข้อสอบ (Content Validity Index) จากการประเมินความถูกต้องของโมเดลข้อสอบ ปรากฏว่า ข้อสอบทั้ง 3 ระดับ คือ ง่าย ปานกลาง และยาก จำนวนทั้งสิ้น 203 ข้อ มีค่าดัชนีความตรงเชิงเนื้อหาของข้อคำถามรายข้อ (I-CVI) เท่ากับ 1.00 และค่าดัชนีความตรงเชิงเนื้อหาของเครื่องมือทั้งชุด (S-CVI) เท่ากับ 1.00 โดยรวมโมเดลข้อสอบมีความถูกต้องผ่านเกณฑ์ สอดคล้องกับงานวิจัยของ Tiansawad (2007) กล่าวว่า การคำนวณค่าดัชนี

ความตรงเชิงเนื้อหาควรกระทำสองระดับ คือ หาความตรงของข้อความรายข้อ (I-CVI) และของเครื่องมือทั้งฉบับ (S-CVI) ซึ่งหากข้อความได้ถูกปรับปรุงจนกระทั่งมีค่าดัชนีความตรงเชิงเนื้อหาตามเกณฑ์ขั้นต่ำแล้ว เครื่องมือทั้งฉบับก็จะมีค่าดัชนีความตรงเชิงเนื้อหาตามเกณฑ์ได้เอง สำหรับการนำเสนอค่าดัชนี ในรายงานการวิจัยนั้น ผู้พัฒนาเครื่องมือควรรายงานทั้งดัชนีความตรงเชิงเนื้อหาของรายข้อ และภาพรวมของเครื่องมือทั้งฉบับ เพื่อให้เป็นหลักฐานแสดงถึงคุณภาพของเครื่องมือที่น่าเชื่อถือ สอดคล้องกับงานวิจัยของ Polit and Beck (2006) กล่าวว่า ค่าดัชนีความตรงเชิงเนื้อหา ของข้อความรายข้อ (I-CVI) ที่ดีควรมีค่ามากกว่า 0.80 และค่าดัชนีความตรงเชิงเนื้อหาของเครื่องมือทั้งชุด (S-CVI) ควรได้อย่างน้อย 0.90

3. การพัฒนาระบบการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

ผลการประเมินความเหมาะสมของโปรแกรมโดยผู้เชี่ยวชาญ พบว่า ทั้ง 5 ด้าน มีความเหมาะสมอยู่ในระดับมาก โดยด้านที่มีความเหมาะสมมากเป็นอันดับที่ 1 คือ ด้านความตรงตามความต้องการของโปรแกรมฯ ซึ่งเป็นส่วนที่เกี่ยวข้องกับความสามารถของระบบตรงตามความต้องการของผู้ใช้ แสดงให้เห็นว่าโปรแกรมฯ ได้ถูก ออกแบบผลลัพธ์ของระบบได้สอดคล้องกับความต้องการของผู้ใช้ระบบ และด้านที่มีความเหมาะสมเป็นอันดับที่ 5 คือ ด้านการใช้งาน ซึ่งเป็นส่วนที่เกี่ยวข้องกับความสามารถของระบบในการติดต่อระหว่างระบบและผู้ใช้ แสดงให้เห็นว่าโปรแกรมฯ ยังมีข้อจุดอ่อนในด้านการออกแบบส่วนต่อประสานกับผู้ใช้ (user interface design) ที่จะต้องมีการพัฒนาให้ดียิ่งขึ้น

4. การประเมินประสิทธิภาพในการใช้งานของโปรแกรมการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1

การประเมินความคิดเห็นโดยผู้ทดลองใช้โปรแกรม ผู้วิจัยเลือกวิธีการตรวจสอบความคิดเห็นของครูที่ทดลองใช้โปรแกรมซึ่งเป็นครูกลุ่มสาระการเรียนรู้คณิตศาสตร์ ชั้นมัธยมศึกษา จำนวน 40 คน เนื่องจากการสำรวจความคิดเห็นเป็นวิธีตรวจสอบข้อบกพร่องของโปรแกรมที่นิยมใช้ในปัจจุบัน ผลการประเมินความคิดเห็น พบว่า ทั้ง 2 ด้าน มีประสิทธิภาพของโปรแกรมอยู่ในระดับดีมาก ได้แก่ 1) ด้านลักษณะทั่วไปของโปรแกรม คือ โปรแกรมมีความสะดวกเมื่อเรียกใช้งาน การออกแบบโปรแกรมมีความน่าสนใจ รูปแบบของโปรแกรมเข้าใจง่าย โปรแกรมง่ายต่อการใช้งาน การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว และโปรแกรมมีระบบป้องกันการทำงานผิดพลาดของผู้ใช้งาน และ 2) ด้านความสะดวกในการใช้งานโปรแกรม คือ โปรแกรมมีเมนูที่เข้าถึงส่วนต่าง ๆ ของโปรแกรมได้ง่ายและสะดวก ผู้ใช้งานสามารถเลือกสร้างข้อสอบด้วยวิธีการสุ่มข้อสอบและสร้างข้อสอบด้วยวิธีการกำหนดข้อสอบได้ตามความต้องการ ผู้ใช้งานสามารถเลือกสร้างข้อสอบตามเนื้อหาและระดับความยากได้ตามความต้องการ โปรแกรมสามารถสร้างข้อสอบได้เร็วกว่าการสร้างข้อสอบด้วยมือ โปรแกรมสามารถแสดงผลการสร้างข้อสอบบนหน้าจอภาพได้ครบถ้วน ผู้ใช้งานสามารถบันทึกรายงานผลการสร้างข้อสอบและส่งพิมพ์แบบทดสอบพร้อมเฉลยคำตอบได้อย่างสะดวก และเมื่อมีข้อสงสัยในการใช้โปรแกรม ผู้ใช้สามารถดูคำชี้แจงจากคู่มือการใช้โปรแกรมและสามารถปฏิบัติตามได้ จากผลการประเมินแสดงให้เห็นว่าโปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นช่วยลดภาระงานการสร้างข้อสอบของครูได้มาก เนื่องจากใช้เวลาประมาณ 1 นาที ก็สามารถสร้างข้อสอบที่มีคุณภาพ และจำนวนข้อที่มาก โดยที่ครูสามารถเลือกสร้างข้อสอบได้

ตามหลักสูตรวิชาคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ประกอบด้วย 3 สาระ คือ สาระที่ 1 จำนวนและพีชคณิต สาระที่ 2 การวัดและเรขาคณิต และสาระที่ 3 สถิติและความน่าจะเป็น และในแต่ละสาระได้จัดแยกข้อสอบออกเป็น 3 ระดับ ได้แก่ ข้อสอบระดับง่าย ข้อสอบระดับปานกลาง และข้อสอบระดับยาก ซึ่งในคลังข้อสอบมีโมเดลข้อสอบต้นแบบทั้งหมด 203 ข้อ นอกจากนี้โปรแกรมยังออกแบบให้ครูสามารถเลือกสร้างข้อสอบได้ 2 วิธี คือ 1) สร้างจากการสุ่มข้อสอบ และ 2) สร้างจากการกำหนดข้อสอบ

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 ระบบการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ที่พัฒนาขึ้นนี้ เป็นโปรแกรมในรูปแบบเว็บแอปพลิเคชัน การใช้งานสามารถใช้ผ่านระบบอินเทอร์เน็ตได้ที่ <https://www.aig-system.com> โดยโปรแกรมได้ถูกออกแบบมาให้เหมาะสำหรับการใช้งานกับระบบปฏิบัติการทางคอมพิวเตอร์ เช่น คอมพิวเตอร์ตั้งโต๊ะ (desktop computer) คอมพิวเตอร์แบบพกพา (notebook) และแท็บเล็ตคอมพิวเตอร์ (tablet computer) ซึ่งครูผู้สอน และผู้สนใจทั่วไปสามารถเข้ามาใช้โปรแกรมนี้ได้สะดวกทุกที่และทุกเวลาที่สามารถเชื่อมต่ออินเทอร์เน็ตได้

1.2 การนำระบบการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ไปใช้ ควรนำไปทดลองใช้ในบริบทเล็ก ๆ ก่อน เพื่อศึกษาปัญหาและความเหมาะสมของการนำไปใช้ในบริบทที่แตกต่างกัน ซึ่งอาจมีปัญหาดังกล่าวต่างกัน ก่อนที่จะนำไปขยายผลต่อไป เช่น ประสิทธิภาพของการเข้าถึงอินเทอร์เน็ตที่แตกต่างกันในแต่ละพื้นที่ เป็นต้น

1.3 ครูที่สอนกลุ่มสาระการเรียนรู้คณิตศาสตร์มัธยมศึกษาปีที่ 1 สามารถใช้ระบบการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 สร้างข้อสอบเพื่อเป็นแบบฝึกหัดให้นักเรียนได้ทบทวนเนื้อหาเกี่ยวกับวิชาคณิตศาสตร์ โดยผู้ใช้งานสามารถดาวน์โหลดคู่มือการใช้งานได้จากเว็บไซต์ <https://www.aig-system.com> และปฏิบัติตามคำชี้แจงในคู่มือการใช้งานโปรแกรม และหากมีข้อสงสัยสามารถติดต่อผู้วิจัยผ่านทางเครือข่ายสังคมออนไลน์ที่แจ้งไว้บนเว็บไซต์ข้างต้น

1.4 การพัฒนาระบบการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 เป็นการพัฒนาโปรแกรมฯ ตามหลักการวงจรการพัฒนาแบบที่มีความเหมาะสมของโปรแกรมฯ ในระดับมาก และมีการออกแบบผลลัพธ์ของระบบสอดคล้องกับความต้องการของผู้ใช้ในเกณฑ์ดี ผู้ที่สนใจในการพัฒนาโปรแกรมฯ อื่น ๆ สามารถนำหลักการวงจรการพัฒนาแบบไปประยุกต์ใช้ในงานวิจัยได้

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรมีการพัฒนาระบบการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 เพื่อให้ครูสามารถพิมพ์ข้อสอบใส่เข้าไปได้เลย เพื่อให้ นักเรียนเข้าไปทดสอบออนไลน์และสามารถประมวลผลได้ทันที

2.2 ควรพัฒนาระบบการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ให้โปรแกรมสามารถใช้งานได้สะดวกทั้งในระบบปฏิบัติการทางคอมพิวเตอร์ และระบบปฏิบัติการบนสมาร์ตโฟน

2.3 ควรมีการพัฒนากระบวนการสร้างข้อสอบอัตโนมัติสำหรับจัดคลังข้อสอบจำแนกตามเนื้อหาและระดับความยาก วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ในที่ส่วนเกี่ยวข้องกับความสามารถของระบบด้านการติดต่อระหว่างระบบและผู้ใช้ เช่น ด้านการออกแบบส่วนต่อประสานกับผู้ใช้ (user interface design) ให้โปรแกรมมีความน่าสนใจมากยิ่งขึ้น

Reference

- Arendasy, M. E., & Sommer, M. (2012). Using automatic item generation to meet the increasing item demands of high-stakes educational and occupational assessment. *Learning and Individual Differences, 22*(1), 112-117. doi:10.1016/j.lindif.2011.11.005
- Gierl, M. J., & Lai, H. (2012). The Role of Item Models in Automatic Item Generation. *International Journal of Testing, 12*(3), 273-298. doi:10.1080/15305058.2011.635830
- 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. R. (2012). Using automatic item generation to create multiple-choice test items. *Med Educ, 46*(8), 757-765. doi:10.1111/j.1365-2923.2012.04289.x
- Gierl, M. J., Zhou, J., & Alves, C. (2008). Developing a Taxonomy of Item Model Types to Promote Assessment Engineering. *The Journal of Technology, Learning, and Assessment, 7*(2), 1-51.
- Guyer, R., & Thompson, N.A. (2014). *User's Manual for Xcalibre item response theory calibration software, version 4.2.2 and later*. Assessment Systems Corporation.
- Luecht, R. M. (2013). Assessment Engineering Task Model Maps, Task Models and Templates as a New Way to Develop and Implement Test Specifications. *Journal of Applied Testing Technology, 14*, 1-38.
- Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health, 29*(5), 489-497. doi:10.1002/nur.20147
- Urry, V. W. (1977). Tailored Testing: A Successful Application of Latent Trait Theory. *Journal of Education Measurement, 14*(2), 181-196.

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)
- Jantana, S. (2019). *Development of Automatic Item Generation in Mathematics Assessment for Grade Nine Level*. [Master's thesis]. Burapha University. (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)
- Tiansawad, S (2007). Content validity index: Critique and recommendation for computation. *Nursing Journal, 34*(4), 1-9. (in Thai)