



การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์



วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้
ปีการศึกษา 2569
ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์



วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้
ปีการศึกษา 2569
ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

วิทยานิพนธ์ เรื่อง “การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์

แบบออนไลน์”

ของ รัฐภาพรรณ บุญมี

ได้รับการพิจารณาให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการสอบวิทยานิพนธ์

(รองศาสตราจารย์ ดร.พัชรี จันทรเพ็ญ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(รองศาสตราจารย์ ดร.กฤษฎากาญจน์ ไตพิทักษ์)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

(ผู้ช่วยศาสตราจารย์ ดร.สายฝน วิบูลรังสรรค์)

.....กรรมการผู้ทรงคุณวุฒิภายใน

(ผู้ช่วยศาสตราจารย์ ดร.เอี่ยมพร หลินเจริญ)

.....กรรมการผู้ทรงคุณวุฒิภายใน

(ผู้ช่วยศาสตราจารย์ ดร.ชำนาญ ปาณาวงษ์)

อนุมัติ

.....

(รองศาสตราจารย์ ดร.เกสัชกรประยุทธ์ ภูวรรตนาวิธ)

คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง	การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์
ผู้วิจัย	ฐิภาพรรณ บุญมี
อาจารย์ที่ปรึกษาหลัก	รองศาสตราจารย์ ดร.กฤษยากาญจน์ โตพิทักษ์
อาจารย์ที่ปรึกษาร่วม	ผู้ช่วยศาสตราจารย์ ดร.สายฝน วิบูลรังสรรค์
ประเภทสารนิพนธ์	วิทยานิพนธ์ ปริญญาโท สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้, มหาวิทยาลัยนเรศวร, 2569
คำสำคัญ	การคิดเชิงคำนวณ, แบบวัดเชิงสถานการณ์, คุณสมบัติทางจิตมิติ

บทคัดย่อ

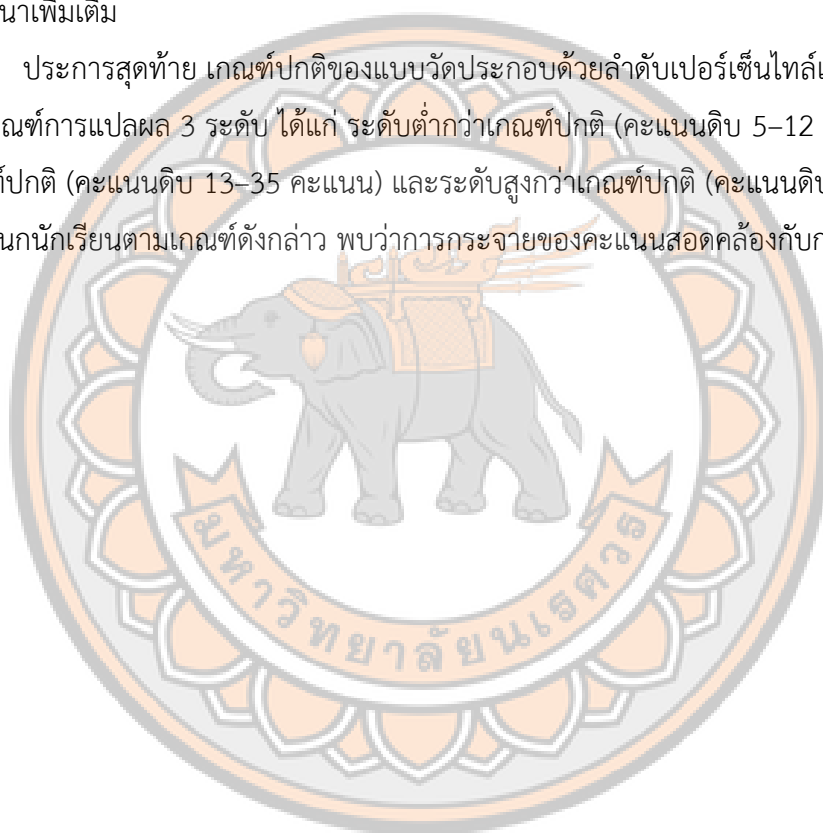
การวิจัยครั้งนี้มีวัตถุประสงค์ 3 ประการ ได้แก่ (1) เพื่อพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น (2) เพื่อตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด ภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ และ (3) เพื่อสร้างเกณฑ์ปกติของแบบวัดสำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร การวิจัยใช้ระเบียบวิธีวิจัยและพัฒนา ดำเนินการ 3 ขั้นตอน ได้แก่ การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ การตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด และการสร้างเกณฑ์ปกติของแบบวัด กลุ่มตัวอย่างคัดเลือกแบบเจาะจง แบ่งเป็น 2 กลุ่ม คือ กลุ่มทดลองใช้ จำนวน 150 คน และกลุ่มเก็บข้อมูลหลัก จำนวน 1,265 คน ซึ่งเป็นนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร เครื่องมือที่ใช้ในการวิจัย คือ แบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ การวิเคราะห์ข้อมูลประกอบด้วยสถิติบรรยาย การวิเคราะห์คุณภาพรายข้อตามทฤษฎีการทดสอบแบบดั้งเดิม การวิเคราะห์องค์ประกอบเชิงยืนยัน การวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบภายใต้โมเดล 2 พารามิเตอร์ และการสร้างเกณฑ์ปกติ

ผลการวิจัยพบว่า แบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ฉบับสมบูรณ์จำนวน 40 ข้อ จาก 10 สถานการณ์ ครอบคลุมองค์ประกอบของทักษะการคิดเชิงคำนวณ 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม ด้านละ 10 ข้อ โดยให้การให้คะแนนแบบทวิภาค แบบวัดพัฒนาจากฉบับร่าง 80 ข้อ ซึ่งผ่านการตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ 7 คน มีค่าดัชนีความสอดคล้องอยู่ในช่วง 0.57 ถึง 1.00 ผ่านเกณฑ์การยอมรับทุกข้อ

ประการที่สอง แบบวัดมีคุณสมบัติทางจิตมิติเป็นไปตามเกณฑ์ที่กำหนดทุกด้าน ภายใต้ทฤษฎีการทดสอบแบบดั้งเดิม ค่าความยากของข้อคำถามอยู่ในช่วง 0.37 ถึง 0.81 ค่าอำนาจจำแนก

อยู่ในช่วง 0.26 ถึง 0.65 และค่าความเชื่อมั่น KR-20 เท่ากับ 0.84 จากกลุ่มทดลองใช้ และเท่ากับ 0.93 จากกลุ่มเก็บข้อมูลหลัก ภายใต้ทฤษฎีการตอบสนองข้อสอบ ข้อมูลสอดคล้องกับข้อตกลงเบื้องต้นด้านความเป็นเอกมิติและความเป็นอิสระเฉพาะที่ โดยโมเดล 2 พารามิเตอร์มีความเหมาะสมกับข้อมูลมากกว่าโมเดล 1 พารามิเตอร์ ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบอยู่ระหว่าง 0.629 ถึง 3.367 ค่าพารามิเตอร์ความยากอยู่ระหว่าง -1.161 ถึง 0.716 และฟังก์ชันสารสนเทศของแบบวัดมีค่าสูงสุดที่ระดับความสามารถ $\theta \approx -1.00$ แสดงให้เห็นว่าแบบวัดมีความแม่นยำสูงในการประเมินผู้เรียนที่มีระดับความสามารถค่อนข้างต่ำ จึงเหมาะสำหรับใช้คัดกรองนักเรียนที่ต้องการการส่งเสริมและพัฒนาเพิ่มเติม

ประการสุดท้าย เกณฑ์ปกติของแบบวัดประกอบด้วยลำดับเปอร์เซ็นต์และคะแนนที่ปกติ พร้อมเกณฑ์การแปลผล 3 ระดับ ได้แก่ ระดับต่ำกว่าเกณฑ์ปกติ (คะแนนดิบ 5–12 คะแนน) ระดับอยู่ในเกณฑ์ปกติ (คะแนนดิบ 13–35 คะแนน) และระดับสูงกว่าเกณฑ์ปกติ (คะแนนดิบ 36–40 คะแนน) เมื่อจำแนกนักเรียนตามเกณฑ์ดังกล่าว พบว่าการกระจายของคะแนนสอดคล้องกับการแจกแจงปกติ



Title THE DEVELOPMENT OF SITUATIONAL ONLINE COMPUTATIONAL THINKING SKILL TEST

Author Thipaphan Boonmee

Advisor Associate Professor Dr. Krityakarn Topitak, Ph.D.

Co-Advisor Assistant Professor Dr. Saifon Vibulrangson, Ph.D.

Academic Paper Ph.D. Thesis in Innovation of Learning Measurement, Naresuan University, 2026

Keywords computational thinking, situational assessment, psychometric properties

ABSTRACT

This study had three objectives: (1) to develop a situational online computational thinking assessment for lower-secondary school students; (2) to examine the psychometric properties of the assessment under the frameworks of Classical Test Theory (CTT) and Item Response Theory (IRT); and (3) to construct test norms for lower-secondary school students in Bangkok. The study employed a research and development methodology comprising three phases: development of the situational online computational thinking assessment, examination of the psychometric properties of the assessment, and construction of test norms. Purposive sampling was used to recruit two groups of participants: a try-out group of 150 students and a main data collection group of 1,265 students, all of whom were lower-secondary school students in Bangkok. The research instrument was the situational online computational thinking assessment. Data were analyzed using descriptive statistics, CTT-based item analysis, confirmatory factor analysis, IRT analysis under the two-parameter logistic model, and norm construction.

The findings revealed three major results. First, a situational online computational thinking assessment consisting of 40 dichotomously scored items across 10 situations was successfully developed, covering four computational thinking components: decomposition, pattern recognition, abstraction, and algorithm design, with 10 items per component. The assessment was developed from an 80-item draft form, in which

content validity evaluated by seven experts yielded Index of Item–Objective Congruence (IOC) values ranging from 0.57 to 1.00, with all items meeting the acceptance criterion.

Second, the assessment demonstrated satisfactory psychometric properties across all evaluated dimensions. Under the Classical Test Theory framework, item difficulty indices ranged from 0.37 to 0.81, item discrimination indices ranged from 0.26 to 0.65, and the KR-20 reliability coefficients were 0.84 for the try-out group and 0.93 for the main data collection group. Under the Item Response Theory framework, the data satisfied the assumptions of unidimensionality and local independence, and the two-parameter logistic model provided a better fit to the data than the one-parameter model. Item discrimination parameters ranged from 0.629 to 3.367, whereas item difficulty parameters ranged from -1.161 to 0.716 . The test information function reached its maximum at approximately $\theta = -1.00$, indicating that the assessment provides the greatest measurement precision for students with relatively low ability levels and is therefore particularly suitable for identifying students who require additional instructional support.

Finally, the normative scores consisted of percentile ranks and normalized T-scores, accompanied by three interpretive categories: below normal (raw scores 5–12), within the normal range (raw scores 13–35), and above normal (raw scores 36–40). Classification based on these criteria demonstrated a score distribution consistent with the normal distribution.

ประกาศคุณูปการ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความรู้และการสนับสนุนอย่างดียิ่งจากบุคคลหลายท่าน ผู้วิจัยขอกราบขอบพระคุณทุกท่านด้วยความเคารพและซาบซึ้งใจอย่างสูง ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. กฤษยาภาณุจันทร์ โดพัทธ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ที่ได้กรุณาให้คำแนะนำอย่างใกล้ชิด เอาใจใส่ และอุทิศเวลาอันมีค่าในการตรวจสอบและให้ข้อเสนอแนะที่เป็นประโยชน์ต่องานวิจัยตลอดระยะเวลาการศึกษา รวมทั้งให้กำลังใจและสนับสนุนผู้วิจัยในทุกขั้นตอนจนงานวิจัยสำเร็จลุล่วงได้ด้วยดี

ขอกราบขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ทุกท่าน ที่ได้กรุณาสละเวลาอ่านวิทยานิพนธ์ ให้ข้อคิดเห็น ข้อเสนอแนะ และข้อสังเกตที่มีคุณค่า ซึ่งช่วยให้ผู้วิจัยสามารถปรับปรุงวิทยานิพนธ์ฉบับนี้ให้มีความสมบูรณ์ยิ่งขึ้น

ขอกราบขอบพระคุณผู้เชี่ยวชาญทั้ง 7 ท่าน ที่ได้กรุณาสละเวลาอันมีค่าในการตรวจสอบความตรงเชิงเนื้อหา (Content Validity) ของแบบวัด และให้ข้อเสนอแนะเชิงคุณภาพที่เป็นประโยชน์อย่างยิ่งในการปรับปรุงข้อคำถามและสถานการณ์ในแบบวัด

ขอขอบคุณผู้บริหารโรงเรียน คณาจารย์ผู้สอนวิชาวิทยาการคำนวณ และนักเรียนระดับมัธยมศึกษาตอนต้นจากโรงเรียนสาธิตสังกัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม (อว.) โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษากรุงเทพมหานคร (สพม. กทม.) และโรงเรียนในสังกัดสำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.) ในกรุงเทพมหานคร รวมทั้งสิ้น 1,415 คน ที่ให้ความร่วมมือในการเก็บข้อมูลเป็นอย่างดี ทำให้การวิจัยในครั้งนี้ประสบความสำเร็จ

ขอขอบพระคุณโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ ศูนย์วิจัยและพัฒนาการศึกษา ที่ให้โอกาสผู้วิจัยได้ศึกษาต่อในระดับที่สูงขึ้น และสนับสนุนทุนการศึกษาในครั้งนี้

ขอกราบขอบพระคุณบิดา มารดา และสมาชิกทุกคนในครอบครัว ที่คอยเป็นกำลังใจ ให้การสนับสนุน ตลอดระยะเวลาการศึกษาในระดับปริญญาเอก

คุณประโยชน์ทั้งปวงที่เกิดจากคุณิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชาพระคุณของบิดา มารดา ครูบาอาจารย์ทุกท่านที่ให้ความรู้และอบรมสั่งสอนผู้วิจัยมาตลอดชีวิตการศึกษา

วิทยานิพนธ์ฉบับนี้ได้รับทุนอุดหนุนการวิจัยในการทำวิทยานิพนธ์ของนิสิตบัณฑิตศึกษาที่มีศักยภาพสูง ประจำปีงบประมาณ 2568 จากบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร

ฐิภาพรรณ บุญมี

สารบัญ

หน้า

บทคัดย่อภาษาไทย	ค
บทคัดย่อภาษาอังกฤษ.....	จ
ประกาศคุุณูปการ	ช
บทที่ 1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
วัตถุประสงค์การวิจัย.....	7
ความสำคัญของการวิจัย.....	7
ขอบเขตการวิจัย.....	8
นิยามศัพท์เฉพาะ	9
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	12
1. แนวคิดเชิงคำนวณ.....	12
2. แบบวัดเชิงสถานการณ์	21
3. การวัดและประเมินผลในรูปแบบออนไลน์.....	30
4. กรอบแนวคิดทางจิตมิติสำหรับการพัฒนาแบบวัด.....	36
5. งานวิจัยที่เกี่ยวข้อง	50
6. กรอบแนวคิดในการวิจัย	56
บทที่ 3 วิธีดำเนินงานวิจัย.....	59
ตอนที่ 1 การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์... 60	
ตอนที่ 2 การตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด.....	69
ตอนที่ 3 การสร้างเกณฑ์ปกติของแบบวัด.....	75
บทที่ 4 ผลการวิจัย.....	79
ตอนที่ 1 ผลการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์.....	79
ตอนที่ 2 ผลการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด.....	83
ตอนที่ 3 ผลการสร้างเกณฑ์ปกติของแบบวัด	96

บทที่ 5 บทสรุป.....	101
สรุปผลการวิจัย	101
อภิปรายผลการวิจัย	103
ข้อเสนอแนะ.....	106
บรรณานุกรม	109
ภาคผนวก	118
ประวัติผู้วิจัย	235



สารบัญตาราง

	หน้า
ตาราง 1 การสังเคราะห์องค์ประกอบของแนวคิดเชิงคำนวณจากกรอบแนวคิดสำคัญ	15
ตาราง 2 เกณฑ์การแปลผลค่าดัชนีความสอดคล้อง (IOC).....	37
ตาราง 3 เกณฑ์การแปลผลค่าอำนาจจำแนก (r).....	39
ตาราง 4 เกณฑ์การแปลผลค่าความเชื่อมั่น KR-20.....	40
ตาราง 5 ดัชนีความสอดคล้องของโมเดลและเกณฑ์การยอมรับ	42
ตาราง 6 เกณฑ์การแปลผลค่า Q3 สำหรับการตรวจสอบความเป็นอิสระเฉพาะที่	43
ตาราง 7 เกณฑ์การแปลผลค่าพารามิเตอร์อำนาจจำแนก (a).....	44
ตาราง 8 เกณฑ์การแปลผลคะแนนที่เป็น 3 ระดับ	49
ตาราง 9 ภาพรวมวิธีดำเนินการวิจัยจำแนกตามวัตถุประสงค์.....	59
ตาราง 10 โครงสร้างแบบวัดฉบับร่าง จำนวน 80 ข้อ.....	62
ตาราง 11 สรุปกระบวนการพัฒนาแบบวัดและจำนวนข้อคำถามในแต่ละขั้น	68
ตาราง 12 เกณฑ์การแปลผลดัชนีคุณภาพข้อสอบและความเชื่อมั่นตามทฤษฎีการทดสอบ แบบดั้งเดิม.....	70
ตาราง 13 เกณฑ์การแปลผลค่า Q3 สำหรับการตรวจสอบความเป็นอิสระเฉพาะที่	71
ตาราง 14 ดัชนีความสอดคล้องของโมเดลในการวิเคราะห์องค์ประกอบเชิงยืนยันและเกณฑ์ การยอมรับ.....	72
ตาราง 15 ดัชนีการเปรียบเทียบโมเดลทฤษฎีการตอบสนองข้อสอบ	73
ตาราง 16 เกณฑ์การแปลผลค่าพารามิเตอร์อำนาจจำแนก (a)	74
ตาราง 17 เกณฑ์การแปลผลระดับทักษะการคิดเชิงคำนวณ	76
ตาราง 18 ผลการประเมินความตรงเชิงเนื้อหาของแบบวัดฉบับร่าง 80 ข้อ โดยผู้เชี่ยวชาญ 7 คน ...	80
ตาราง 19 ผลการคัดเลือกข้อคำถามเข้าสู่แบบวัดฉบับสมบูรณ์	82
ตาราง 20 ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของข้อคำถามรายข้อ จำนวน 40 ข้อ จากกลุ่มเก็บข้อมูลหลัก 1,265 คน	83
ตาราง 21 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันภายใต้โมเดลองค์ประกอบเดียว	88
ตาราง 22 ผลการเปรียบเทียบโมเดลทฤษฎีการตอบสนองข้อสอบระหว่างโมเดล 1 พารามิเตอร์ และโมเดล 2 พารามิเตอร์.....	89

ตาราง 23 ค่าพารามิเตอร์อำนาจจำแนก (a) และค่าความยาก (b) ของข้อคำถามรายข้อ จำนวน 40 ข้อ ภายใต้โมเดล 2 พารามิเตอร์	90
ตาราง 24 สรุปผลการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดฉบับสมบูรณ์ 40 ข้อ	95
ตาราง 25 เกณฑ์ปกติของคะแนนแบบวัดทักษะการคิดเชิงคำนวณสำหรับนักเรียนระดับมัธยมศึกษา ตอนต้นในกรุงเทพมหานคร	97
ตาราง 26 จำนวนและร้อยละของนักเรียนจำแนกตามระดับทักษะการคิดเชิงคำนวณ	99



สารบัญภาพ

หน้า

ภาพ 1 กรอบแนวคิดในการวิจัย	57
ภาพ 2 ฟังก์ชันสารสนเทศของแบบวัดทักษะการคิดเชิงคำนวณ.....	93



บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ในศตวรรษที่ 21 ความก้าวหน้าของเทคโนโลยีดิจิทัลและการขยายตัวของเศรษฐกิจฐานความรู้ (knowledge-based economy) ได้ส่งผลให้ระบบการศึกษาทั่วโลกต้องปรับเปลี่ยนแนวทางการจัดการศึกษาอย่างต่อเนื่อง เพื่อให้สอดคล้องกับบริบทของสังคมและตลาดแรงงานที่เปลี่ยนแปลงอย่างรวดเร็ว องค์การเพื่อความร่วมมือและการพัฒนาทางเศรษฐกิจ (Organisation for Economic Co-operation and Development [OECD], 2018) ระบุว่า การศึกษาในยุคปัจจุบันควรมุ่งพัฒนาทักษะการคิดขั้นสูง (higher-order thinking skills) ได้แก่ การคิดวิเคราะห์ การแก้ปัญหาที่ซับซ้อน การคิดอย่างมีวิจารณญาณ และการคิดสร้างสรรค์ เพื่อเตรียมผู้เรียนให้สามารถปรับตัวและดำรงชีวิตได้อย่างมีประสิทธิภาพในโลกที่มีความผันผวน แนวโน้มดังกล่าวสะท้อนว่า การศึกษาในปัจจุบันมิได้มุ่งเน้นเพียงการถ่ายทอดองค์ความรู้ แต่ให้ความสำคัญกับการพัฒนาศักยภาพของผู้เรียนในการประยุกต์ใช้ความรู้เพื่อแก้ปัญหาในสถานการณ์จริง

ในบรรดาทักษะการคิดขั้นสูงดังกล่าว แนวคิดการคิดเชิงคำนวณ (Computational Thinking: CT) เป็นทักษะที่ได้รับการยอมรับอย่างกว้างขวางว่าจำเป็นสำหรับผู้เรียนในศตวรรษที่ 21 Wing (2006) ผู้เสนอแนวคิดนี้ ได้อธิบายว่าการคิดเชิงคำนวณเป็นกระบวนการคิดที่ใช้ในการกำหนดปัญหา วิเคราะห์ปัญหา และออกแบบแนวทางแก้ไขอย่างเป็นระบบและเป็นขั้นตอน แนวคิดนี้ไม่ได้จำกัดอยู่เพียงการเขียนโปรแกรมคอมพิวเตอร์ แต่เป็นทักษะการคิดที่ประยุกต์ใช้ได้ในทุกสาขาวิชา และในสถานการณ์ชีวิตจริง (Shute, Sun, & Asbell-Clarke, 2017) ต่อมา Wing (2008) ได้เสนอเพิ่มเติมว่าการคิดเชิงคำนวณควรได้รับการพัฒนาให้แก่ผู้เรียนทุกคน ไม่จำกัดเฉพาะผู้ที่ประกอบอาชีพด้านวิทยาการคอมพิวเตอร์ ซึ่งสอดคล้องกับ Barr และ Stephenson (2011) ที่เห็นว่าทักษะนี้มีความสำคัญเทียบเท่าการอ่าน การเขียน และคณิตศาสตร์ ในฐานะพื้นฐานของการศึกษา ในด้านกรอบแนวคิด Brennan และ Resnick (2012) ได้เสนอกรอบการคิดเชิงคำนวณที่ครอบคลุม 3 มิติ ได้แก่ แนวคิดเชิงคำนวณ (computational concepts) กระบวนการเชิงคำนวณ (computational practices) และมุมมองเชิงคำนวณ (computational perspectives) ซึ่งถูกนำไปใช้ในการพัฒนาหลักสูตรและการประเมินผลการเรียนรู้ในหลายประเทศ

ความสำคัญของการคิดเชิงคำนวณทำให้หลายประเทศบรรจุแนวคิดนี้ไว้ในหลักสูตรการศึกษาขั้นพื้นฐานอย่างเป็นรูปธรรม สหราชอาณาจักรกำหนดให้วิชาคอมพิวเตอร์เป็นวิชาบังคับใน

หลักสูตรแห่งชาติตั้งแต่ปี ค.ศ. 2014 (Department for Education, 2013) สหรัฐอเมริกาดำเนินโครงการ Computer Science for All พร้อมพัฒนามาตรฐานการเรียนรู้วิทยาการคอมพิวเตอร์สำหรับผู้เรียนทุกระดับ สิงคโปร์บูรณาการการคิดเชิงคำนวณเข้ากับวิชาคณิตศาสตร์และวิทยาศาสตร์ (Ministry of Education Singapore, 2018) และเกาหลีได้กำหนดให้การศึกษาด้านซอฟต์แวร์เป็นวิชาบังคับในระบบการศึกษาขั้นพื้นฐาน (Kim & Lee, 2020) นอกจากการรับรองในระดับนโยบายแล้ว ความสำคัญของทักษะนี้ยังมีหลักฐานเชิงประจักษ์สนับสนุนอย่างต่อเนื่อง Grover และ Pea (2013) ทบทวนวรรณกรรมด้านการคิดเชิงคำนวณในระดับการศึกษาขั้นพื้นฐานอย่างเป็นระบบ และพบว่าทักษะการคิดเชิงคำนวณมีความสัมพันธ์ใกล้ชิดกับความสามารถในการแก้ปัญหา เนื่องจากทั้งสองกระบวนการต่างอาศัยการแบ่งย่อยปัญหา การจดจำรูปแบบ และการออกแบบขั้นตอนการแก้ปัญหาเช่นเดียวกัน ในทำนองเดียวกัน Tang และคณะ (2020) ทบทวนงานวิจัยด้านการวัดและประเมินทักษะการคิดเชิงคำนวณจำนวน 96 เรื่อง และพบว่าการพัฒนาทักษะนี้มีความสัมพันธ์เชิงบวกกับผลสัมฤทธิ์ทางการเรียนในกลุ่มสาระวิทยาศาสตร์ เทคโนโลยี วิศวกรรม และคณิตศาสตร์ ส่วน Bocconi และคณะ (2016) รายงานว่าประเทศที่บูรณาการการคิดเชิงคำนวณเข้าในหลักสูตรอย่างชัดเจน มีแนวโน้มพัฒนาทักษะการแก้ปัญหาของผู้เรียนได้ดีกว่าประเทศที่ยังไม่มีนโยบายดังกล่าว หลักฐานเชิงนโยบายและผลการวิจัยข้างต้นสะท้อนว่า การคิดเชิงคำนวณเป็นทักษะพื้นฐานที่มีความสำคัญต่อการพัฒนาความสามารถในการคิดวิเคราะห์และแก้ปัญหาของผู้เรียนในทุกระดับการศึกษา

สำหรับประเทศไทย กระทรวงศึกษาธิการได้ให้ความสำคัญกับการพัฒนาผู้เรียนให้มีทักษะที่จำเป็นสำหรับการดำรงชีวิตในศตวรรษที่ 21 โดยกำหนดนโยบายส่งเสริมความสามารถด้านการคิดวิเคราะห์ การคิดสร้างสรรค์ การใช้เทคโนโลยี และการอยู่ร่วมกับผู้อื่นในสังคม (กระทรวงศึกษาธิการ, 2560) เพื่อรองรับนโยบายดังกล่าว กระทรวงศึกษาธิการได้ประกาศใช้หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560) โดยบรรจุสาระเทคโนโลยี (วิทยาการคำนวณ) เข้าเป็นส่วนหนึ่งของกลุ่มสาระการเรียนรู้วิทยาศาสตร์ มีเป้าหมายเพื่อพัฒนาผู้เรียนให้คิดวิเคราะห์และแก้ปัญหาย่อยอย่างเป็นขั้นตอนและเป็นระบบ รวมถึงประยุกต์ใช้เทคโนโลยีในการแก้ปัญหาชีวิตจริง (สำนักงานเลขาธิการสภาการศึกษา, 2564) ทั้งนี้ สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท., 2560) ซึ่งเป็นหน่วยงานหลักในการพัฒนาหลักสูตรวิทยาการคำนวณ ได้กำหนดองค์ประกอบของทักษะการคิดเชิงคำนวณไว้ 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา (decomposition) การพิจารณารูปแบบ (pattern recognition) การคิดเชิงนามธรรม (abstraction) และการออกแบบอัลกอริทึม (algorithm design) องค์ประกอบทั้ง 4 ด้านนี้สอดคล้องกับกรอบการสังเคราะห์งานวิจัยระดับนานาชาติของ Grover และ Pea (2013) ซึ่งแสดงว่าหลักสูตรวิทยาการคำนวณของไทยมีรากฐานทางทฤษฎีที่สอดคล้องกับมาตรฐานสากล หลักสูตรฉบับปรับปรุงนี้เริ่มใช้ในโรงเรียนตั้งแต่

ปีการศึกษา 2561 และขยายครบทุกระดับชั้นในปีการศึกษา 2563 ส่งผลให้ปัจจุบันนักเรียนระดับมัธยมศึกษาตอนต้นได้รับการพัฒนาทักษะการคิดเชิงคำนวณผ่านการเรียนรู้ตามหลักสูตรอย่างต่อเนื่อง ภายใต้บริบทดังกล่าว โรงเรียนและครูผู้สอนจึงมีบทบาทสำคัญในการติดตามและประเมินพัฒนาการด้านการคิดเชิงคำนวณของผู้เรียน เพื่อให้ทราบว่าผู้เรียนสามารถพัฒนาทักษะได้ตามเป้าหมายของหลักสูตรมากน้อยเพียงใด ข้อมูลจากการประเมินมีความสำคัญต่อการวางแผนจัดการเรียนรู้ การปรับปรุงกิจกรรมการเรียนการสอน การส่งเสริมผู้เรียนที่ยังมีจุดอ่อน ตลอดจนการตัดสินใจทางการศึกษาในระดับต่าง ๆ ดังนั้น การมีเครื่องมือวัดทักษะการคิดเชิงคำนวณที่สามารถประเมินระดับความสามารถของผู้เรียนได้อย่างถูกต้องและน่าเชื่อถือ จึงเป็นประเด็นที่มีความสำคัญทั้งในเชิงวิชาการและเชิงปฏิบัติ

อย่างไรก็ตาม การวัดและประเมินทักษะการคิดเชิงคำนวณยังคงเป็นประเด็นที่ท้าทาย เนื่องจากการคิดเชิงคำนวณมิได้เป็นเพียงความรู้เชิงเนื้อหาที่สามารถวัดได้ด้วยแบบทดสอบข้อเขียนแบบดั้งเดิม แต่เป็นทักษะการคิดและการปฏิบัติที่ปรากฏผ่านกระบวนการวิเคราะห์และแก้ปัญหาในสถานการณ์ต่าง ๆ Pellegrino และ Hilton (2012) ชี้ว่าการประเมินทักษะแห่งศตวรรษที่ 21 ควรมุ่งวัดความสามารถในการประยุกต์ใช้ความรู้และทักษะเพื่อจัดการกับสถานการณ์ใหม่ที่มีความซับซ้อนมากกว่าการวัดเพียงความจำหรือความเข้าใจเชิงเนื้อหา แบบทดสอบที่เน้นการวัดความรู้เชิงเนื้อหา (knowledge-based tests) แม้จะสะท้อนระดับความรู้พื้นฐานของผู้เรียนได้ แต่ยังไม่เพียงพอที่จะยืนยันว่าผู้เรียนสามารถนำความรู้ไปประยุกต์ใช้ในการแก้ปัญหาจริงได้หรือไม่ นอกจากนี้ การประเมินที่แยกวัดองค์ประกอบของทักษะออกจากกัน (fragmented assessment) ยังไม่สอดคล้องกับลักษณะของการแก้ปัญหาในชีวิตจริง ซึ่งผู้เรียนต้องใช้การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึมร่วมกันในเวลาเดียวกัน การวัดแบบแยกส่วนจึงมีข้อจำกัดในการสะท้อนสมรรถนะการคิดเชิงคำนวณของผู้เรียนอย่างครบถ้วน

จากข้อจำกัดของการวัดที่เน้นความรู้เชิงเนื้อหาและการประเมินแบบแยกองค์ประกอบดังกล่าว แนวคิดการประเมินเชิงสถานการณ์ (situational judgment tests: SJTs) จึงได้รับความสนใจเพิ่มขึ้นในฐานะเครื่องมือที่มีศักยภาพในการประเมินทักษะการคิดขั้นสูงและสมรรถนะเชิงปฏิบัติ เนื่องจากการนำเสนอสถานการณ์จำลองที่ใกล้เคียงกับบริบทจริง และเปิดโอกาสให้ผู้ตอบแสดงการตัดสินใจหรือเลือกแนวทางแก้ปัญหาภายใต้เงื่อนไขของสถานการณ์ที่กำหนด (พิชิต ฤทธิ์จรูญ, 2556) งานวิจัยจำนวนมากสนับสนุนประสิทธิภาพของการประเมินลักษณะดังกล่าว โดย Lievens, Peeters และ Schollaert (2008) พบว่าแบบวัดเชิงสถานการณ์มีความตรงเชิงพยากรณ์ในระดับที่ดี และสามารถทำนายพฤติกรรมหรือผลการปฏิบัติงานจริงได้ในหลากหลายบริบท ขณะที่ Whetzel และ McDaniel (2009) รายงานว่าแบบวัดประเภทนี้มีประสิทธิภาพในการประเมินความสามารถด้านการตัดสินใจและการแก้ปัญหาในบริบทเฉพาะ นอกจากนี้ Wiggins (1990) ยังเสนอแนวคิดการ

ประเมินตามสภาพจริง (authentic assessment) ที่มุ่งประเมินความสามารถของผู้เรียนผ่านสถานการณ์ที่มีความหมายและใกล้เคียงกับชีวิตจริง เพื่อให้ผู้เรียนได้แสดงการประยุกต์ใช้ความรู้และทักษะในการจัดการกับปัญหาที่พบในชีวิตประจำวัน

แนวคิดดังกล่าวสอดคล้องกับธรรมชาติของทักษะการคิดเชิงคำนวณ ซึ่งมีได้เกิดจากการใช้ทักษะย่อยเพียงด้านใดด้านหนึ่ง แต่เป็นกระบวนการบูรณาการการแบ่งย่อยปัญหา การพิจารณา รูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึมเข้าด้วยกันในการแก้ปัญหา ดังนั้น การใช้แบบวัดเชิงสถานการณ์จึงน่าจะเป็นแนวทางที่เหมาะสมในการประเมินทักษะการคิดเชิงคำนวณให้สะท้อนสมรรถนะของผู้เรียนได้ใกล้เคียงกับการนำไปใช้ในสถานการณ์จริงมากกว่าการวัดแบบแยกส่วน

แม้แนวคิดการประเมินเชิงสถานการณ์จะได้รับการยอมรับว่าเป็นแนวทางที่เหมาะสมสำหรับการประเมินทักษะการคิดขั้นสูงและการแก้ปัญหาในบริบทจริง แต่เมื่อพิจารณาเครื่องมือวัดทักษะการคิดเชิงคำนวณที่มีการพัฒนาขึ้นในปัจจุบัน พบว่ายังคงมีข้อจำกัดบางประการที่ควรได้รับการพัฒนาเพิ่มเติม โดยเฉพาะในด้านการสะท้อนการใช้ทักษะการคิดเชิงคำนวณอย่างบูรณาการภายใต้สถานการณ์ที่ใกล้เคียงกับชีวิตจริง

ตัวอย่างเช่น Computational Thinking Test (CTt) ที่พัฒนาโดย Román-González, Pérez-González และ Jiménez-Fernández (2017) ซึ่งใช้ปริศนาเชิงภาพในการวัดทักษะการคิดเชิงคำนวณ แม้จะผ่านการตรวจสอบคุณภาพจนได้รับการยอมรับอย่างกว้างขวางในวงวิชาการ แต่ลักษณะข้อสอบยังเป็นการวัดแบบแยกข้อ (discrete items) ที่มุ่งประเมินทักษะแต่ละด้านเป็นรายข้อ จึงอาจยังไม่สะท้อนกระบวนการแก้ปัญหาที่เกิดขึ้นอย่างต่อเนื่องในสถานการณ์จริงได้อย่างครบถ้วน ส่วน Bebras Challenge แม้จะมีข้อสอบที่หลากหลายและท้าทายความคิดของผู้เรียน แต่เน้นการแก้ปัญหาในลักษณะปริศนา (puzzle-like problems) มากกว่าการประยุกต์ใช้ทักษะในบริบทชีวิตประจำวัน ขณะที่ Dr. Scratch ซึ่งพัฒนาโดย Moreno-León, Robles และ Román-González (2015) มุ่งวิเคราะห์ผลงานโปรแกรมที่ผู้เรียนสร้างขึ้นแล้ว จึงสะท้อนผลลัพธ์ของการเขียนโปรแกรมมากกว่ากระบวนการคิดและการตัดสินใจในสถานการณ์ที่หลากหลาย

สำหรับบริบทประเทศไทย งานวิจัยด้านการพัฒนาเครื่องมือวัดทักษะการคิดเชิงคำนวณยังมีจำนวนไม่มากนัก เครื่องมือส่วนใหญ่อยู่ในรูปแบบแบบทดสอบวัดความรู้ หรือเป็นการประยุกต์ใช้เครื่องมือที่พัฒนาจากต่างประเทศ ซึ่งอาจยังไม่สอดคล้องกับบริบททางการศึกษา ภาษา และประสบการณ์การเรียนรู้ของผู้เรียนไทย นอกจากนี้ ยังพบงานวิจัยจำนวนน้อยที่พัฒนาแบบวัดในรูปแบบเชิงสถานการณ์ซึ่งสามารถวัดองค์ประกอบของทักษะการคิดเชิงคำนวณทั้ง 4 ด้านร่วมกันภายใต้สถานการณ์เดียว และผ่านการตรวจสอบคุณภาพด้วยกรอบทางจิตมิติอย่างครอบคลุมนอกจาก รูปแบบของเครื่องมือแล้ว การตรวจสอบคุณภาพของเครื่องมือวัดก็เป็นประเด็นสำคัญ งานวิจัยด้านการพัฒนาแบบทดสอบส่วนใหญ่ยังอาศัยกรอบทฤษฎีการทดสอบแบบดั้งเดิม (classical test

theory: CTT) เพียงกรอบเดียว ซึ่งมีข้อจำกัดสำคัญคือ คุณสมบัติของข้อสอบขึ้นอยู่กับกลุ่มตัวอย่างที่ใช้ทดสอบ (sample-dependent) และความสามารถของผู้สอบขึ้นอยู่กับความยากของข้อสอบ (test-dependent) อีกทั้งไม่สามารถให้ข้อมูลความแม่นยำของการวัดในแต่ละระดับความสามารถได้ (Embretson & Reise, 2000; Hambleton, Swaminathan, & Rogers, 1991) ทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) ได้รับการพัฒนาขึ้นเพื่อแก้ข้อจำกัดดังกล่าว โดยมีคุณสมบัติเด่นด้านความคงที่ของพารามิเตอร์ (parameter invariance) กล่าวคือ คุณสมบัติของข้อสอบไม่เปลี่ยนแปลงตามกลุ่มผู้สอบ และความสามารถของผู้สอบไม่เปลี่ยนแปลงตามชุดข้อสอบที่ใช้ (Lord, 1980) การใช้ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบร่วมกันช่วยให้ตรวจสอบคุณภาพของแบบวัดได้ครอบคลุมยิ่งขึ้น เนื่องจากทฤษฎีทั้งสองให้ข้อมูลที่เสริมกัน ทฤษฎีการทดสอบแบบดั้งเดิมให้ภาพรวมคุณภาพของแบบวัดทั้งฉบับ ส่วนทฤษฎีการตอบสนองข้อสอบให้ข้อมูลคุณภาพรายข้อที่เป็นอิสระจากกลุ่มตัวอย่าง และระบุช่วงระดับความสามารถที่แบบวัดวัดได้แม่นยำที่สุด การใช้ทั้งสองทฤษฎีจึงให้หลักฐานคุณภาพของแบบวัดที่หนักแน่นกว่าการใช้ทฤษฎีใดทฤษฎีหนึ่งเพียงอย่างเดียว

เมื่อได้เครื่องมือที่มีคุณภาพแล้ว การตีความคะแนนที่ได้จากการประเมินก็มีความสำคัญไม่น้อยไปกว่าการพัฒนาเครื่องมือเอง การตีความแบบอ้างอิงกลุ่ม (norm-referenced interpretation) ซึ่งใช้เกณฑ์ปกติ (norms) เป็นเกณฑ์เปรียบเทียบ ช่วยให้ทราบว่าผู้สอบมีความสามารถอยู่ในระดับใดเมื่อเทียบกับกลุ่มอ้างอิงที่มีลักษณะใกล้เคียงกัน (Crocker & Algina, 2008) โดยเฉพาะเกณฑ์ปกติเฉพาะพื้นที่ (local norms) มีความสำคัญเป็นพิเศษ เนื่องจากแต่ละพื้นที่มีบริบททางการศึกษา ทรัพยากร และโอกาสทางการเรียนรู้ที่แตกต่างกัน การเปรียบเทียบกับกลุ่มที่มาจากบริบทเดียวกันจึงให้การตีความที่มีความหมายและเป็นธรรมกว่าการเปรียบเทียบกับกลุ่มที่มีเงื่อนไขแตกต่างกัน (Linn & Gronlund, 2000; Reynolds, Livingston, & Willson, 2009)

นอกจากการพัฒนาเครื่องมือที่มีคุณภาพและการจัดทำเกณฑ์ปกติสำหรับการแปลผลแล้ว การนำผลการประเมินไปใช้ประโยชน์อย่างรวดเร็วและมีประสิทธิภาพก็เป็นอีกประเด็นสำคัญของการวัดและประเมินผลทางการศึกษาในปัจจุบัน โดยเฉพาะในยุคดิจิทัลที่การประเมินผู้เรียนจำนวนมากจำเป็นต้องอาศัยระบบที่สามารถจัดการข้อมูลได้อย่างรวดเร็วและแม่นยำ ด้วยเหตุนี้ การพัฒนาแบบวัดในรูปแบบออนไลน์ (online assessment) จึงได้รับความสนใจเพิ่มขึ้น เนื่องจากช่วยเพิ่มประสิทธิภาพของการวัดและประเมินผลในหลายด้าน ทั้งการลดเวลาและค่าใช้จ่ายในการดำเนินงาน การตรวจให้คะแนนอัตโนมัติที่ลดความคลาดเคลื่อนจากการตรวจด้วยมือ การรายงานผลที่รวดเร็วซึ่งเอื้อต่อการนำผลไปใช้พัฒนาผู้เรียน และการเพิ่มโอกาสการเข้าถึงของผู้เรียน (Alruwais, Wills, & Wald, 2018; อนุวัติ คุณแก้ว, 2566; Bunderson, Inouye, & Olsen, 1989; Whitelock & Brasher, 2006)

จากการทบทวนวรรณกรรมและสภาพปัญหาข้างต้น ผู้วิจัยสรุปช่องว่างการวิจัยได้ 5 ประการ ประการแรก ยังขาดเครื่องมือวัดที่ประเมินทักษะทั้ง 4 องค์ประกอบอย่างบูรณาการในสถานการณ์เดียว แทนการวัดแยกเป็นองค์ประกอบย่อย ประการที่สอง เครื่องมือส่วนใหญ่วัดความรู้เชิงเนื้อหา มากกว่าความสามารถในการประยุกต์ใช้ทักษะเพื่อแก้ปัญหาในสถานการณ์ที่ใกล้เคียงกับสถานการณ์จริง ประการที่สาม ยังมีข้อจำกัดด้านเครื่องมือวัดที่พัฒนาขึ้นภายใต้บริบททางการศึกษาของประเทศไทย โดยเฉพาะ ประการที่สี่ การตรวจสอบคุณภาพเครื่องมือส่วนใหญ่อาศัยเพียงกรอบทฤษฎีการทดสอบ แบบดั้งเดิม โดยยังมีการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบร่วมกันในระดับจำกัด และประการที่ห้า ยังขาดเกณฑ์ปกติที่ช่วยสนับสนุนการแปลผลคะแนนให้มีความชัดเจนและสอดคล้องกับบริบทของผู้เรียน ช่องว่างการวิจัยดังกล่าวสะท้อนความจำเป็นในการพัฒนาเครื่องมือวัดที่สามารถตอบสนองประเด็นเหล่านี้ได้อย่างเป็นระบบ

การวิจัยครั้งนี้เลือกกรุงเทพมหานครเป็นพื้นที่ศึกษา เนื่องจากกรุงเทพมหานครมีโรงเรียนครบทั้ง 3 สังกัดหลักที่พบในระดับประเทศ ได้แก่ สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา (สพม.) สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.) และกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม (อว.) ซึ่งสะท้อนความหลากหลายของบริบททางการศึกษาในเชิงโครงสร้าง นอกจากนี้ กรุงเทพมหานครยังมีความพร้อมด้านโครงสร้างพื้นฐานทางเทคโนโลยีที่เอื้อต่อการดำเนินการวิจัยและการใช้แบบวัดในรูปแบบออนไลน์

ส่วนการเลือกนักเรียนระดับมัธยมศึกษาตอนต้นทั้ง 3 ระดับชั้นเป็นกลุ่มเป้าหมาย มีเหตุผลสนับสนุน 4 ประการ ประการแรก นักเรียนกลุ่มนี้ได้เรียนวิชาวิทยาการคำนวณมาอย่างต่อเนื่องตั้งแต่ระดับประถมศึกษา จึงมีพื้นฐานแนวคิดเชิงคำนวณติดตัวมา ประการที่สอง ตามทฤษฎีพัฒนาการทางสติปัญญาของ Piaget (1972) นักเรียนวัย 13-15 ปีอยู่ในขั้นการคิดเชิงนามธรรม (formal operational stage) ที่สามารถคิดเชิงตรรกะ ตั้งสมมติฐาน และแก้ปัญหาที่ซับซ้อนได้ ซึ่งสอดคล้องกับลักษณะของแบบวัดเชิงสถานการณ์ ประการที่สาม การครอบคลุมนักเรียนทั้ง 3 ระดับชั้นช่วยให้สามารถศึกษาการใช้แบบวัดกับผู้เรียนที่มีประสบการณ์การเรียนรู้แตกต่างกัน และสนับสนุนการสร้างเกณฑ์ปกติที่ครอบคลุมกลุ่มเป้าหมาย และประการที่สี่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 เป็นผู้เรียนในช่วงปลายของระดับมัธยมศึกษาตอนต้น จึงเหมาะสมต่อการประเมินผลลัพธ์การเรียนรู้ที่เกิดขึ้นภายใต้หลักสูตรในช่วงชั้นดังกล่าว

ด้วยเหตุผลข้างต้น ผู้วิจัยจึงพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร เพื่อให้ได้เครื่องมือที่สามารถสะท้อนความสามารถในการประยุกต์ใช้ทักษะการคิดเชิงคำนวณในการแก้ปัญหาที่ใกล้เคียงกับสถานการณ์จริงได้อย่างเหมาะสม

การวิจัยครั้งนี้มีวัตถุประสงค์ที่เชื่อมโยงกัน 3 ประการ ได้แก่ (1) พัฒนาแบบวัดทักษะการคิดเชิงคำนวณที่ประเมินองค์ประกอบทั้ง 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม อย่างบูรณาการผ่านสถานการณ์ที่ใกล้เคียงกับชีวิตจริงในรูปแบบออนไลน์ที่สามารถรายงานผลกลับแก่ผู้เรียนได้ (2) ตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดภายใต้ทั้งทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ เพื่อให้ได้หลักฐานสนับสนุนคุณภาพของเครื่องมืออย่างรอบด้าน และ (3) สร้างเกณฑ์ปกติสำหรับใช้ประกอบการแปลผลคะแนนของนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร

ผลที่ได้จากการวิจัยครั้งนี้จะช่วยเพิ่มองค์ความรู้ด้านการพัฒนาเครื่องมือวัดทักษะการคิดเชิงคำนวณ ตลอดจนเป็นแนวทางในการพัฒนาและตรวจสอบคุณภาพเครื่องมือวัดทักษะการคิดขั้นสูงในบริบทการศึกษาไทย อีกทั้งยังช่วยสนับสนุนการประเมินผู้เรียนให้สอดคล้องกับธรรมชาติของทักษะการคิดเชิงคำนวณ ซึ่งมุ่งเน้นการประยุกต์ใช้ความรู้และกระบวนการคิดในการแก้ปัญหามากกว่าการวัดความรู้เชิงเนื้อหาเพียงอย่างเดียว

วัตถุประสงค์การวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ 3 ประการ ดังต่อไปนี้

1. เพื่อพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น
2. เพื่อตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดทักษะการคิดเชิงคำนวณ ภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)
3. เพื่อสร้างเกณฑ์ปกติของแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร

ความสำคัญของการวิจัย

ผลของการวิจัยครั้งนี้มีความสำคัญทั้งในด้านวิชาการและด้านการนำไปใช้ ดังนี้

1. ความสำคัญด้านวิชาการ
 - 1.1 ได้แบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น ที่ครอบคลุมองค์ประกอบของทักษะการคิดเชิงคำนวณทั้ง 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม

ภายใต้สถานการณ์เดียวกัน ซึ่งเป็นการนำเสนอแนวทางการวัดที่มุ่งประเมินการใช้ทักษะแบบบูรณาการมากกว่าการวัดแยกเป็นองค์ประกอบย่อย

1.2 ได้หลักฐานเชิงประจักษ์เกี่ยวกับคุณภาพของแบบวัดจากการตรวจสอบภายใต้ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ ซึ่งช่วยสนับสนุนความน่าเชื่อถือของผลการวัดและเพิ่มองค์ความรู้ด้านการพัฒนาเครื่องมือวัดทางการศึกษาในบริบทของทักษะการคิดเชิงคำนวณ

1.3 ได้แนวทางการพัฒนาและตรวจสอบคุณภาพแบบวัดเชิงสถานการณ์แบบออนไลน์ที่สามารถใช้เป็นข้อมูลอ้างอิงสำหรับการพัฒนาเครื่องมือวัดทักษะหรือสมรรถนะด้านอื่นในบริบททางการศึกษา

2. ความสำคัญด้านการนำไปใช้

2.1 ได้เครื่องมือวัดที่ผ่านการตรวจสอบคุณภาพทางจิตมิติ ซึ่งสามารถนำไปใช้ประเมินทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้นได้ตามวัตถุประสงค์ที่กำหนดไว้

2.2 ผลการวัดที่ได้จากแบบวัดสามารถใช้เป็นข้อมูลประกอบการพิจารณาระดับทักษะการคิดเชิงคำนวณของผู้เรียน ทั้งในภาพรวมและรายองค์ประกอบ เพื่อสนับสนุนการติดตามและประเมินผลการเรียนรู้

2.3 ได้เกณฑ์ปกติสำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร ซึ่งช่วยสนับสนุนการแปลผลคะแนนให้มีความชัดเจนและสอดคล้องกับบริบทของกลุ่มเป้าหมายที่ศึกษา

2.4 กระบวนการพัฒนาแบบวัด การตรวจสอบคุณภาพ และการสร้างเกณฑ์ปกติที่ได้จากการวิจัยครั้งนี้ สามารถใช้เป็นกรอบอ้างอิงสำหรับการดำเนินการวิจัยหรือพัฒนาเครื่องมือวัดในบริบทที่มีลักษณะใกล้เคียงกันต่อไป

ขอบเขตการวิจัย

การวิจัยครั้งนี้กำหนดขอบเขตไว้ 4 ด้าน ดังนี้

1. ขอบเขตด้านเนื้อหา

การวิจัยครั้งนี้มุ่งพัฒนาแบบวัดทักษะการคิดเชิงคำนวณ ตามองค์ประกอบที่สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท., 2560) กำหนด จำนวน 4 องค์ประกอบ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม โดยใช้สถานการณ์ที่จำลองจากบริบทชีวิตประจำวันของนักเรียนระดับมัธยมศึกษาตอนต้นเป็นฐานของข้อคำถาม

2. ขอบเขตด้านประชากรและกลุ่มตัวอย่าง

ประชากรในการวิจัย คือ นักเรียนระดับมัธยมศึกษาตอนต้น ชั้นมัธยมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 3 ในกรุงเทพมหานคร กลุ่มตัวอย่างแบ่งเป็น 2 กลุ่มตามระยะของการวิจัย ได้แก่ กลุ่มตัวอย่างในการทดลองใช้แบบวัด เพื่อวิเคราะห์คุณภาพรายข้อตามทฤษฎีการทดสอบแบบดั้งเดิม และกลุ่มตัวอย่างในการเก็บข้อมูลหลัก เพื่อตรวจสอบคุณสมบัติทางจิตมิติและสร้างเกณฑ์ปกติ ทั้งสองกลุ่มเป็นนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร

3. ขอบเขตด้านตัวแปร

ตัวแปรที่ศึกษาในการวิจัยครั้งนี้ คือ ทักษะการคิดเชิงคำนวณ ซึ่งผู้วิจัยศึกษาตามกรอบองค์ประกอบ 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม

4. ขอบเขตด้านการตรวจสอบคุณภาพเครื่องมือ

การตรวจสอบคุณภาพของแบบวัด ครอบคลุมการตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ การวิเคราะห์คุณภาพรายข้อและความเชื่อมั่นตามทฤษฎีการทดสอบแบบดั้งเดิม การวิเคราะห์ความตรงเชิงโครงสร้างและพารามิเตอร์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ และการสร้างเกณฑ์ปกติในรูปคะแนนที่ปกติ

นิยามศัพท์เฉพาะ

เพื่อให้เกิดความเข้าใจตรงกันในการวิจัยครั้งนี้ ผู้วิจัยกำหนดนิยามศัพท์เฉพาะที่ใช้ในการวิจัย ดังนี้

1. ทักษะการคิดเชิงคำนวณ

ทักษะการคิดเชิงคำนวณ หมายถึง ความสามารถของนักเรียนในการวิเคราะห์และแก้ปัญหาอย่างเป็นระบบ โดยใช้กระบวนการคิดในการแยกปัญหา ค้นหาแบบ สร้างข้อสรุปเชิงนามธรรม และกำหนดลำดับขั้นตอนการแก้ปัญหา เพื่อให้ได้แนวทางหรือคำตอบที่เหมาะสมต่อสถานการณ์ที่กำหนด ในการวิจัยครั้งนี้ ทักษะการคิดเชิงคำนวณประกอบด้วย 4 องค์ประกอบ ดังนี้

1.1 การแบ่งย่อยปัญหา

การแบ่งย่อยปัญหา หมายถึง ความสามารถในการวิเคราะห์ปัญหาที่มีความซับซ้อน โดยแยกออกเป็นส่วนย่อยที่สามารถพิจารณาและจัดการได้ง่ายขึ้น ครอบคลุมความสามารถในการ (1) ระบุปัญหาหลักที่ต้องแก้ไข (2) แยกปัญหาออกเป็นส่วนย่อยที่มีความชัดเจน (3) พิจารณาความสัมพันธ์ระหว่างส่วนย่อยของปัญหา (4) กำหนดแนวทางจัดการปัญหาในแต่ละส่วน และ (5) เชื่อมโยงผลการแก้ปัญหาแต่ละส่วนเข้าด้วยกันเพื่อแก้ปัญหาโดยรวม

1.2 การพิจารณารูปแบบ

การพิจารณารูปแบบ หมายถึง ความสามารถในการวิเคราะห์ข้อมูล เหตุการณ์ หรือสถานการณ์ เพื่อค้นหาความเหมือน ความแตกต่าง ความสัมพันธ์ หรือรูปแบบที่เกิดขึ้นซ้ำ ซึ่งสามารถนำไปใช้ในการอธิบาย คาดการณ์ หรือแก้ปัญหาได้ ครอบคลุมความสามารถในการ (1) สังเกตและรวบรวมข้อมูลที่เกี่ยวข้อง (2) เปรียบเทียบข้อมูลเพื่อค้นหาความเหมือนและความแตกต่าง (3) จำแนกข้อมูลตามลักษณะร่วมที่สำคัญ (4) ระบุรูปแบบหรือแนวโน้มที่เกิดขึ้นซ้ำ และ (5) ประยุกต์ใช้รูปแบบที่ค้นพบกับสถานการณ์ใหม่

1.3 การคิดเชิงนามธรรม

การคิดเชิงนามธรรม หมายถึง ความสามารถในการวิเคราะห์สถานการณ์หรือปัญหาโดยมุ่งเน้นสาระสำคัญ ตัดทอนรายละเอียดที่ไม่จำเป็น และสร้างข้อสรุปหรือหลักการทั่วไปที่สามารถนำไปใช้ในบริบทอื่นได้ ครอบคลุมความสามารถในการ (1) ระบุสาระสำคัญของปัญหา (2) คัดกรองข้อมูลหรือรายละเอียดที่ไม่เกี่ยวข้อง (3) สรุปเป็นกฎเกณฑ์หรือหลักการทั่วไป (4) สร้างแบบจำลองหรือแนวคิดที่ช่วยอธิบายปัญหาได้ง่ายขึ้น และ (5) ประยุกต์ใช้หลักการที่ได้กับสถานการณ์อื่นที่มีลักษณะคล้ายคลึงกัน

1.4 การออกแบบอัลกอริทึม

การออกแบบอัลกอริทึม หมายถึง ความสามารถในการกำหนดลำดับขั้นตอนการแก้ปัญหาอย่างเป็นระบบ ชัดเจน และสามารถนำไปปฏิบัติได้ ครอบคลุมความสามารถในการ (1) กำหนดจุดเริ่มต้นและผลลัพธ์ที่ต้องการ (2) จัดลำดับขั้นตอนการดำเนินงานอย่างเป็นเหตุเป็นผล (3) กำหนดเงื่อนไขหรือทางเลือกที่อาจเกิดขึ้นในแต่ละขั้นตอน (4) ตรวจสอบและปรับปรุงขั้นตอนการดำเนินงานให้มีประสิทธิภาพยิ่งขึ้น และ (5) พัฒนาแนวทางการแก้ปัญหาที่สามารถนำไปประยุกต์ใช้กับสถานการณ์ลักษณะเดียวกันได้

ในการวิจัยครั้งนี้ ทักษะการคิดเชิงคำนวณ หมายถึง คะแนนที่นักเรียนได้รับจากการทำแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ที่ผู้วิจัยพัฒนาขึ้น ซึ่งครอบคลุมองค์ประกอบ 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม

2. แบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์

แบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ หมายถึง เครื่องมือวัดที่ผู้วิจัยพัฒนาขึ้นเพื่อประเมินทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้น โดยประกอบด้วยสถานการณ์ที่จำลองจากบริบทชีวิตประจำวันของนักเรียน จำนวน 10 สถานการณ์ แต่ละสถานการณ์มีข้อคำถามแบบปรนัยชนิด 4 ตัวเลือก จำนวน 4 ข้อ รวมทั้งสิ้น 40 ข้อ

ครอบคลุมองค์ประกอบของทักษะการคิดเชิงคำนวณ 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณา รูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม ด้านละ 10 ข้อ แบบวัดดำเนินการผ่านระบบออนไลน์และสามารถรายงานผลคะแนนทั้งในภาพรวมและรายองค์ประกอบของผู้สอบได้

3. การให้คะแนนแบบทวิภาค

การให้คะแนนแบบทวิภาค หมายถึง วิธีการให้คะแนนที่กำหนดให้ผู้ตอบได้รับ 1 คะแนน เมื่อเลือกคำตอบถูกต้อง และได้รับ 0 คะแนน เมื่อเลือกคำตอบไม่ถูกต้อง เหมาะสำหรับข้อสอบปรนัย ที่มีคำตอบถูกต้องเพียงคำตอบเดียว และสอดคล้องกับการวิเคราะห์ข้อมูลภายใต้ทฤษฎีการตอบสนองข้อสอบสำหรับข้อมูลแบบสองค่า

4. คุณสมบัติทางจิตมิติ

คุณสมบัติทางจิตมิติ หมายถึง คุณลักษณะที่ใช้เป็นหลักฐานสนับสนุนคุณภาพของแบบวัดทั้งในด้านความตรง ความเชื่อมั่น และความเหมาะสมของข้อสอบ โดยในการวิจัยครั้งนี้ตรวจสอบภายใต้ 2 กรอบทฤษฎี ดังนี้

4.1 ทฤษฎีการทดสอบแบบดั้งเดิม

การตรวจสอบภายใต้ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) ครอบคลุมการตรวจสอบ (1) ความตรงเชิงเนื้อหาโดยใช้ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ (IOC) จากผู้เชี่ยวชาญ (2) ค่าความยากของข้อสอบ (3) ค่าอำนาจจำแนกของข้อสอบ และ (4) ค่าความเชื่อมั่นของแบบวัดทั้งฉบับด้วยสูตร KR-20

4.2 ทฤษฎีการตอบสนองข้อสอบ

การตรวจสอบภายใต้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ครอบคลุมการตรวจสอบ (1) ความเป็นเอกมิติ (Unidimensionality) ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน (2) ความเป็นอิสระเฉพาะที่ (Local Independence) ด้วยสถิติ Q3 (3) การเปรียบเทียบและคัดเลือกโมเดลที่เหมาะสม (4) การประมาณค่าพารามิเตอร์ข้อสอบ ได้แก่ ค่าอำนาจจำแนก (a) และค่าความยาก (b) (5) ความสอดคล้องของโมเดลทั้งในระดับโดยรวมและระดับรายข้อ และ (6) ฟังก์ชันสารสนเทศของแบบวัด (Test Information Function: TIF)

5. เกณฑ์ปกติ

เกณฑ์ปกติ หมายถึง เกณฑ์อ้างอิงสำหรับใช้แปลผลคะแนนของแบบวัด ซึ่งสร้างจากคะแนนของนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร ในการวิจัยครั้งนี้นำเสนอในรูปแบบลำดับเปอร์เซ็นต์และคะแนนที่ปกติ พร้อมกำหนดระดับทักษะเป็น 3 ระดับ ได้แก่ สูงกว่าเกณฑ์ปกติ อยู่ในเกณฑ์ปกติ และต่ำกว่าเกณฑ์ปกติ

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยและพัฒนาเครื่องมือวัดครั้งนี้มุ่งพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น ผู้วิจัยได้ทบทวนและสังเคราะห์เอกสารและงานวิจัยที่เกี่ยวข้อง ครอบคลุมกรอบทฤษฎี งานวิจัยเชิงประจักษ์ และมาตรฐานทางวิชาการทั้งในประเทศและต่างประเทศ เพื่อเป็นฐานความรู้สำหรับการออกแบบเครื่องมือวัดที่มีคุณภาพและนำไปใช้ได้จริงในบริบทการศึกษาไทย โดยนำเสนอตามลำดับหัวข้อ ดังนี้

1. แนวคิดเชิงคำนวณ
2. แบบวัดเชิงสถานการณ์
3. การวัดและประเมินผลในรูปแบบออนไลน์
4. กรอบแนวคิดทางจิตมิติสำหรับการพัฒนาแบบวัด
5. งานวิจัยที่เกี่ยวข้อง
6. กรอบแนวคิดในการวิจัย

1. แนวคิดเชิงคำนวณ

แนวคิดเชิงคำนวณ (computational thinking: CT) เป็นกรอบการคิดที่ได้รับความสนใจอย่างกว้างขวางในวงการการศึกษาทั่วโลกในช่วงสองทศวรรษที่ผ่านมา นักวิชาการและหน่วยงานทางการศึกษาในหลายประเทศต่างยอมรับว่า ทักษะการคิดเชิงคำนวณเป็นหนึ่งในทักษะพื้นฐานที่จำเป็นต่อการดำรงชีวิตและการทำงานในศตวรรษที่ 21 ทัดเทียมกับการอ่าน การเขียน และคณิตศาสตร์ ทักษะนี้ไม่ได้จำกัดเฉพาะผู้ที่ศึกษาในสาขาวิทยาการคอมพิวเตอร์ หากแต่เป็นวิธีคิดที่ผู้เรียนทุกคนควรได้รับการพัฒนาเพื่อใช้ในการแก้ปัญหาอย่างเป็นระบบ การศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องในหัวข้อนี้ครอบคลุมประเด็นสำคัญ ได้แก่ ความเป็นมาและพัฒนาการของแนวคิดเชิงคำนวณ ความหมายของแนวคิดเชิงคำนวณ องค์ประกอบของทักษะการคิดเชิงคำนวณที่ใช้ในการวิจัย ความสัมพันธ์ระหว่างแนวคิดเชิงคำนวณกับการคิดแก้ปัญหาและทักษะการคิดระดับสูง ความสำคัญของการพัฒนาทักษะการคิดเชิงคำนวณในนักเรียนระดับมัธยมศึกษาตอนต้น และนิยามทักษะการคิดเชิงคำนวณที่ใช้ในการวิจัยครั้งนี้

1.1 ความเป็นมาและพัฒนาการของแนวคิดเชิงคำนวณ

แนวคิดเชิงคำนวณมีจุดกำเนิดในวงการวิทยาการคอมพิวเตอร์ และต่อมาได้ขยายขอบเขตเข้าสู่การศึกษาทั่วไป จุดเริ่มต้นที่สำคัญปรากฏในงานของ Papert (1980) แห่ง Massachusetts Institute of Technology ซึ่งในหนังสือ *Mindstorms: Children, Computers, and Powerful Ideas* ได้เสนอว่ากระบวนการคิดแบบคอมพิวเตอร์สามารถเป็นเครื่องมือในการเรียนรู้และการแก้ปัญหาสำหรับเด็กและเยาวชน แนวคิดของ Papert นับเป็นรากฐานทางความคิด ที่นำไปสู่การพัฒนาแนวคิดเชิงคำนวณในเวลาต่อมา

ต่อมา Wing (2006) ได้เผยแพร่บทความเรื่อง “Computational Thinking” ในวารสาร *Communications of the ACM* ซึ่งมีบทบาทสำคัญต่อการขยายขอบเขตของแนวคิดเชิงคำนวณสู่แวดวงการศึกษา โดยเสนอว่าแนวคิดเชิงคำนวณเป็นกระบวนการคิดที่ควรได้รับการพัฒนาในผู้เรียนทุกคน มิใช่เฉพาะผู้ที่ศึกษาในสาขาวิทยาการคอมพิวเตอร์ ทั้งนี้ Wing อธิบายว่าแนวคิดเชิงคำนวณเกี่ยวข้องกับการกำหนดปัญหาและการนำเสนอวิธีแก้ปัญหาในรูปแบบที่สามารถดำเนินการได้อย่างมีประสิทธิภาพ ไม่ว่าจะโดยมนุษย์หรือเครื่องจักร แนวคิดดังกล่าวได้รับการอ้างอิงและนำไปประยุกต์ใช้ในการพัฒนาหลักสูตร การจัดการเรียนรู้ และการประเมินผลทางการศึกษาอย่างกว้างขวางในเวลาต่อมา

นอกจากนี้ Shute, Sun และ Asbell-Clarke (2017) ได้สังเคราะห์องค์ความรู้เกี่ยวกับการวัดและประเมินทักษะการคิดเชิงคำนวณในบทความเรื่อง “Demystifying Computational Thinking” ซึ่งตีพิมพ์ในวารสาร *Educational Research Review* โดยสรุปว่าการประเมินทักษะการคิดเชิงคำนวณผ่านสถานการณ์ที่มีความเชื่อมโยงกับบริบทในชีวิตจริง มีความเหมาะสมต่อการสะท้อนกระบวนการคิดของผู้เรียน พร้อมทั้งช่วยสร้างความสมดุลระหว่างความตรงเชิงนิเวศวิทยา (ecological validity) และความเป็นระบบในการให้คะแนน นอกจากนี้ รูปแบบการประเมินดังกล่าวยังเหมาะสำหรับผู้เรียนที่ไม่มีพื้นฐานด้านการเขียนโปรแกรมหรือวิทยาการคอมพิวเตอร์โดยตรง

สำหรับประเทศไทย กระทรวงศึกษาธิการได้บูรณาการแนวคิดเชิงคำนวณเข้าสู่หลักสูตรแกนกลางการศึกษาขั้นพื้นฐานผ่านรายวิชาวิทยาการคำนวณ ตามหลักสูตรฉบับปรับปรุง พ.ศ. 2560 โดยมีเป้าหมายให้สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) เป็นหน่วยงานหลักในการพัฒนาหลักสูตร สื่อการเรียนรู้ และแนวทางการวัดและประเมินผล (สสวท., 2560) การบรรจุแนวคิดเชิงคำนวณเข้าในหลักสูตรแกนกลางสะท้อนว่าประเทศไทยให้ความสำคัญกับการพัฒนาทักษะนี้ในผู้เรียนทุกระดับชั้น

1.2 ความหมายของแนวคิดเชิงคำนวณ

นักวิชาการหลายท่านได้ให้นิยามแนวคิดเชิงคำนวณไว้แตกต่างกันตามกรอบแนวคิดและวัตถุประสงค์ในการศึกษา ส่งผลให้แนวคิดเชิงคำนวณได้รับการอธิบายในมิติที่หลากหลาย ทั้งในฐานะกระบวนการแก้ปัญหา วิธีคิดเชิงระบบ และทักษะพื้นฐานสำหรับการเรียนรู้ในศตวรรษที่ 21

Wing (2006) นิยามแนวคิดเชิงคำนวณว่าเป็นกระบวนการคิดที่เกี่ยวข้องกับการแก้ปัญหา การออกแบบระบบ และการทำความเข้าใจพฤติกรรมของมนุษย์ โดยอาศัยแนวคิดพื้นฐานของวิทยาการคอมพิวเตอร์ ต่อมา Wing (2008) ได้ขยายความหมายเพิ่มเติม โดยเสนอว่าแนวคิดเชิงคำนวณคือกระบวนการคิดที่นำหลักการพื้นฐานของวิทยาการคอมพิวเตอร์มาประยุกต์ใช้ในการแก้ปัญหาในชีวิตประจำวันและในศาสตร์อื่น ตั้งแต่วิทยาศาสตร์ คณิตศาสตร์ มนุษยศาสตร์ ไปจนถึงศิลปะ Wing เน้นว่าแนวคิดเชิงคำนวณไม่ใช่การฝึกให้ผู้เรียนคิดเหมือนคอมพิวเตอร์ หากแต่เป็นการฝึกให้ผู้เรียนคิดอย่างเป็นระบบเพื่อแก้ปัญหาที่ซับซ้อน โดยใช้กระบวนการที่ชัดเจน สามารถสื่อสารและตรวจสอบได้

Grover และ Pea (2013) ได้ทบทวนวรรณกรรมอย่างเป็นระบบ และสรุปว่าแนวคิดเชิงคำนวณประกอบด้วยองค์ประกอบสำคัญ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม กรอบองค์ประกอบของ Grover และ Pea ได้รับการอ้างอิงอย่างกว้างขวางในวงการวิจัยด้านการคิดเชิงคำนวณ ส่วน Selby และ Woollard (2013) แห่ง University of Southampton ได้เสนอกรอบแนวคิดเชิงคำนวณที่ประกอบด้วย 5 มิติ ได้แก่ การคิดเชิงนามธรรม การแบ่งย่อยปัญหา การคิดเชิงอัลกอริทึม การประเมิน และการนำไปประยุกต์ใช้

สำหรับบริบทไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท., 2560) นิยามแนวคิดเชิงคำนวณว่าเป็นกระบวนการแก้ปัญหาและการคิดวิเคราะห์อย่างมีเหตุผลเป็นขั้นตอน เพื่อหาวิธีแก้ปัญหาในรูปแบบที่สามารถนำไปประมวผลได้อย่างมีประสิทธิภาพ โดยมุ่งเน้นการประยุกต์ใช้ในชีวิตประจำวันและในรายวิชาต่าง ไม่จำกัดเฉพาะวิชาวิทยาการคอมพิวเตอร์

จากนิยามข้างต้นจะเห็นได้ว่า แม้นักวิชาการแต่ละท่านจะใช้ถ้อยคำต่างกัน แต่มีจุดร่วมที่สำคัญคือ แนวคิดเชิงคำนวณเป็นกระบวนการคิดอย่างเป็นระบบเพื่อแก้ปัญหา ที่ประยุกต์ใช้ได้ในทุกสาขาวิชา ไม่จำกัดเฉพาะการเขียนโปรแกรมคอมพิวเตอร์

1.3 องค์ประกอบของทักษะการคิดเชิงคำนวณที่ใช้ในการวิจัย

จากการทบทวนวรรณกรรมพบว่า นักวิชาการและหน่วยงานทางการศึกษาหลายแห่งได้นำเสนอองค์ประกอบของทักษะการคิดเชิงคำนวณไว้แตกต่างกัน โดยมีจำนวนองค์ประกอบตั้งแต่ 4 ถึง 6 องค์ประกอบ ผู้วิจัยจึงดำเนินการสังเคราะห์องค์ประกอบจากกรอบแนวคิดสำคัญที่ได้รับการ

อ้างอิงอย่างแพร่หลาย เพื่อกำหนดกรอบแนวคิดที่เหมาะสมสำหรับใช้เป็นพื้นฐานในการพัฒนาแบบวัดในการวิจัยครั้งนี้ ดังแสดงในตาราง 1

ตาราง 1 การสังเคราะห์องค์ประกอบของแนวคิดเชิงคำนวณจากกรอบแนวคิดสำคัญ

องค์ประกอบ	Wing (2006)	Grover และ Pea (2013)	Selby และ Woollard (2013)	สสวท. (2560)	องค์ประกอบที่ ใช้ในการวิจัย
การแบ่งย่อย ปัญหา	✓	✓	✓	✓	✓
การพิจารณา รูปแบบ	✓	✓	-	-	✓
การคิดเชิง นามธรรม	✓	✓	✓	✓	✓
การออกแบบ อัลกอริทึม	✓	✓	✓	✓	✓
การประเมิน การนำไป ประยุกต์ใช้	-	✓	✓	-	-

หมายเหตุ: สังเคราะห์โดยผู้วิจัยจากกรอบแนวคิดของ Wing (2006), Grover และ Pea (2013), Selby และ Woollard (2013) และ สสวท. (2560)

จากการสังเคราะห์ข้อมูลในตาราง 1 ผู้วิจัยเลือกใช้องค์ประกอบของทักษะการคิดเชิงคำนวณตามกรอบของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท., 2560) ซึ่งประกอบด้วย 4 องค์ประกอบ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม เนื่องจากเป็นกรอบแนวคิดที่ใช้เป็นพื้นฐานในการจัดการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานของประเทศไทย และมีความสอดคล้องกับแนวคิดของ Grover และ Pea (2013) ซึ่งได้รับการยอมรับในวงวิชาการระดับสากล องค์ประกอบทั้ง 4 ด้านดังกล่าวเป็นองค์ประกอบที่ได้รับการกล่าวถึงอย่างต่อเนื่องในกรอบแนวคิดส่วนใหญ่ที่ผู้วิจัยศึกษา จึงมีความเหมาะสมสำหรับใช้เป็นกรอบในการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณสำหรับนักเรียน

ระดับมัธยมศึกษาตอนต้น สำหรับองค์ประกอบด้านการประเมิน (Evaluation) และการนำไปประยุกต์ใช้ (Generalization หรือ Transfer) ซึ่งปรากฏในกรอบแนวคิดของนักวิชาการบางราย ผู้วิจัยไม่นำมาใช้เป็นองค์ประกอบหลักในการวิจัยครั้งนี้ เนื่องจากองค์ประกอบดังกล่าวไม่ได้ถูกกำหนดไว้ในกรอบแนวคิดของ สสวท. (2560) ซึ่งเป็นกรอบอ้างอิงหลักของการวิจัย องค์ประกอบของทักษะการคิดเชิงคำนวณทั้ง 4 ด้านที่ใช้ในการวิจัยมีรายละเอียดดังต่อไปนี้

1.3.1 การแบ่งย่อยปัญหา

การแบ่งย่อยปัญหา (Decomposition) หมายถึง ความสามารถในการวิเคราะห์ปัญหาที่มีความซับซ้อนโดยแยกออกเป็นส่วนย่อยที่สามารถทำความเข้าใจและจัดการได้ง่ายขึ้น (Grover & Pea, 2013; สสวท., 2560) โดย Wing (2006) อธิบายว่า การแบ่งย่อยปัญหาเป็นกระบวนการสำคัญที่ช่วยลดความซับซ้อนของปัญหา ทำให้ผู้แก้ปัญหาสามารถพิจารณาองค์ประกอบต่าง ๆ ของปัญหาได้อย่างเป็นระบบ แทนการจัดการกับปัญหาทั้งหมดในคราวเดียว ก่อนนำผลจากแต่ละส่วนมาบูรณาการเพื่อสร้างแนวทางแก้ปัญหาโดยรวม สอดคล้องกับ Lee และคณะ (2011) ที่ระบุว่า การแบ่งย่อยปัญหาช่วยให้ผู้เรียนมองเห็นโครงสร้างและความสัมพันธ์ระหว่างองค์ประกอบของปัญหา อันเป็นพื้นฐานสำคัญของการวางแผนและการแก้ปัญหาอย่างมีประสิทธิภาพ

ในการวิจัยครั้งนี้ ผู้วิจัยกำหนดพฤติกรรมบ่งชี้ของการแบ่งย่อยปัญหา ประกอบด้วย (1) การระบุปัญหาหลักที่ต้องแก้ไข (2) การแยกปัญหาออกเป็นส่วนย่อยที่เป็นอิสระต่อกัน (3) การจัดลำดับความสำคัญของแต่ละส่วน (4) การกำหนดแนวทางจัดการสำหรับปัญหาแต่ละส่วน และ (5) การเชื่อมโยงผลจากการจัดการปัญหาแต่ละส่วนเพื่อทำความเข้าใจปัญหาโดยรวม พฤติกรรมดังกล่าวสะท้อนความสามารถในการวิเคราะห์โครงสร้างของปัญหา แยกองค์ประกอบของปัญหา ออกเป็นส่วนย่อยที่สามารถจัดการได้ และพิจารณาความสัมพันธ์ระหว่างส่วนต่าง ๆ ของปัญหาอย่างเป็นระบบ

1.3.2 การพิจารณารูปแบบ

การพิจารณารูปแบบ (Pattern Recognition) หมายถึง ความสามารถในการค้นหาความเหมือน ความแตกต่าง ความสัมพันธ์ หรือรูปแบบที่เกิดขึ้นซ้ำในข้อมูล สถานการณ์ หรือปัญหา (Grover & Pea, 2013; สสวท., 2560) ทักษะนี้ช่วยให้ผู้เรียนสามารถสรุปกฎเกณฑ์หรือแนวโน้มจากข้อมูลที่ปรากฏ และนำความรู้ที่ได้ไปใช้ในการคาดการณ์ อธิบาย หรือแก้ปัญหาในสถานการณ์ใหม่ Brennan และ Resnick (2012) อธิบายว่า การพิจารณารูปแบบเป็นกลไกสำคัญที่ช่วยเชื่อมโยงประสบการณ์เดิมกับสถานการณ์ใหม่ ทำให้ผู้เรียนสามารถประยุกต์แนวทางการแก้ปัญหาที่เคยประสบความสำเร็จมาใช้กับปัญหาที่มีลักษณะคล้ายคลึงกันได้อย่างมีประสิทธิภาพ

ในการวิจัยครั้งนี้ ผู้วิจัยกำหนดพฤติกรรมบ่งชี้ของการพิจารณารูปแบบ ประกอบด้วย (1) การสังเกตและรวบรวมข้อมูลที่เกี่ยวข้อง (2) การเปรียบเทียบข้อมูลเพื่อค้นหาความเหมือน

และความแตกต่าง (3) การจำแนกข้อมูลตามลักษณะร่วมที่สำคัญ (4) การระบุรูปแบบหรือแนวโน้มที่เกิดขึ้นซ้ำ และ (5) การประยุกต์ใช้รูปแบบที่ค้นพบกับสถานการณ์ใหม่ พฤติกรรมดังกล่าวสะท้อนความสามารถในการวิเคราะห์ข้อมูลและสังเคราะห์ความสัมพันธ์ของข้อมูลเพื่อนำไปสู่การสร้างข้อสรุปและการแก้ปัญหา

1.3.3 การคิดเชิงนามธรรม

การคิดเชิงนามธรรม (Abstraction) หมายถึง ความสามารถในการมุ่งเน้นสาระสำคัญของปัญหา โดยคัดกรองรายละเอียดที่ไม่จำเป็นออก เพื่อสร้างแบบจำลอง แนวคิด หรือหลักการทั่วไปที่ช่วยให้เข้าใจและแก้ปัญหาได้อย่างมีประสิทธิภาพ (Grover & Pea, 2013; สสวท., 2560) โดย Wing (2008) ระบุว่า การคิดเชิงนามธรรมเป็นหัวใจสำคัญของแนวคิดเชิงคำนวณ เนื่องจากช่วยให้ผู้แก้ปัญหาสามารถจัดการกับความซับซ้อนของสถานการณ์จริงผ่านการเลือกเฉพาะข้อมูลที่จำเป็นต่อการตัดสินใจและการแก้ปัญหา ต่อมา Wing (2010) อธิบายเพิ่มเติมว่า การคิดเชิงนามธรรมเป็นกระบวนการสร้างแบบจำลองทางความคิดที่สะท้อนสาระสำคัญของปัญหา และเป็นพื้นฐานของการพัฒนาทฤษฎีหรือหลักการทั่วไปที่สามารถนำไปประยุกต์ใช้ในสถานการณ์ที่หลากหลายได้

ในการวิจัยครั้งนี้ ผู้วิจัยกำหนดพฤติกรรมบ่งชี้ของการคิดเชิงนามธรรม ประกอบด้วย (1) การระบุสาระสำคัญของปัญหา (2) การคัดกรองรายละเอียดที่ไม่จำเป็น (3) การสร้างทฤษฎีหรือหลักการทั่วไป (4) การสร้างแบบจำลองหรือแนวคิดที่ช่วยอธิบายปัญหา และ (5) การประยุกต์ใช้หลักการที่ได้กับสถานการณ์อื่น พฤติกรรมดังกล่าวสะท้อนความสามารถในการสกัดสาระสำคัญจากข้อมูลหรือสถานการณ์ที่ซับซ้อน และแปลงเป็นหลักการที่สามารถนำไปใช้ได้อย่างกว้างขวาง

1.3.4 การออกแบบอัลกอริทึม

การออกแบบอัลกอริทึม (Algorithm Design) หมายถึง ความสามารถในการกำหนดลำดับขั้นตอนการแก้ปัญหาอย่างเป็นระบบ ชัดเจน และสามารถนำไปปฏิบัติได้ (Grover & Pea, 2013; สสวท., 2560) Brennan และ Resnick (2012) อธิบายว่า อัลกอริทึมเป็นลำดับขั้นตอนการดำเนินงานที่มีความชัดเจน ไม่กำกวม และมุ่งไปสู่ผลลัพธ์ที่กำหนดไว้ ทักษะการออกแบบอัลกอริทึมจึงเป็นการนำผลจากการแบ่งย่อยปัญหา การพิจารณารูปแบบ และการคิดเชิงนามธรรมมาประมวลเป็นแนวทางการแก้ปัญหาที่สามารถนำไปปฏิบัติได้จริง นอกจากนี้ Knuth (1997) ยังระบุว่า อัลกอริทึมที่ดีควรมีคุณลักษณะสำคัญ ได้แก่ ความมีที่สิ้นสุด ความชัดเจนของขั้นตอน การมีข้อมูลนำเข้าและผลลัพธ์ที่กำหนดไว้อย่างชัดเจน และความสามารถในการนำไปปฏิบัติได้จริง

ในการวิจัยครั้งนี้ ผู้วิจัยกำหนดพฤติกรรมบ่งชี้ของการออกแบบอัลกอริทึม ประกอบด้วย (1) การกำหนดจุดเริ่มต้นและผลลัพธ์ที่ต้องการ (2) การจัดลำดับขั้นตอนการดำเนินงาน

อย่างเป็นเหตุเป็นผล (3) การกำหนดเงื่อนไขหรือทางเลือกที่อาจเกิดขึ้นในแต่ละขั้นตอน (4) การตรวจสอบและปรับปรุงขั้นตอนการดำเนินงาน และ (5) การพัฒนาแนวทางการแก้ปัญหาที่สามารถนำไปประยุกต์ใช้กับสถานการณ์ลักษณะเดียวกันได้ พฤติกรรมดังกล่าวสะท้อนความสามารถในการออกแบบกระบวนการแก้ปัญหาอย่างเป็นระบบและสามารถนำไปใช้ได้จริง

1.4 ความสัมพันธ์ระหว่างแนวคิดเชิงคำนวณกับการคิดแก้ปัญหาและทักษะการคิดระดับสูง

แนวคิดเชิงคำนวณมีความสัมพันธ์ใกล้ชิดกับการคิดแก้ปัญหา เนื่องจากโดยนิยามแล้วแนวคิดเชิงคำนวณคือกระบวนการแก้ปัญหาอย่างเป็นระบบ การทำความเข้าใจรากฐานทางทฤษฎีของการแก้ปัญหา จึงช่วยให้เห็นว่าองค์ประกอบทั้ง 4 ของแนวคิดเชิงคำนวณเชื่อมโยงกับกระบวนการแก้ปัญหาอย่างไร

ในทางจิตวิทยาการรู้คิด ปัญหา (problem) หมายถึง สถานการณ์ที่บุคคลต้องการบรรลุเป้าหมายบางประการ แต่ยังไม่ทราบวิธีการหรือแนวทางที่ชัดเจนในการไปสู่เป้าหมายนั้น (Mayer, 1992) โดย Newell และ Simon (1972) อธิบายว่า ปัญหาประกอบด้วยองค์ประกอบสำคัญ 3 ส่วน ได้แก่ สถานะเริ่มต้น (initial state) สถานะเป้าหมาย (goal state) และตัวดำเนินการ (operators) ที่ใช้เปลี่ยนผ่านจากสถานะเริ่มต้นไปสู่สถานะเป้าหมาย การแก้ปัญหาจึงเป็นกระบวนการค้นหาและเลือกชุดของการกระทำที่เหมาะสมเพื่อนำไปสู่ผลลัพธ์ที่ต้องการ

แนวคิดดังกล่าวมีความสอดคล้องกับแนวคิดเชิงคำนวณ ซึ่งมุ่งเน้นการวิเคราะห์ปัญหา การจัดการความซับซ้อน และการออกแบบขั้นตอนการแก้ปัญหาอย่างเป็นระบบ ผู้เรียนจำเป็นต้องทำความเข้าใจโครงสร้างของปัญหา ค้นหารูปแบบที่เกี่ยวข้อง สกัดสาระสำคัญ และพัฒนาแนวทางการแก้ปัญหาที่เป็นลำดับขั้นตอน จึงอาจกล่าวได้ว่า การแก้ปัญหาเป็นบริบทสำคัญที่สะท้อนการประยุกต์ใช้ทักษะการคิดเชิงคำนวณในสถานการณ์ต่าง ๆ

โมเดลการแก้ปัญหาที่มีอิทธิพลและเชื่อมโยงกับแนวคิดเชิงคำนวณโดยตรง ได้แก่ โมเดลของ Polya (1957) ในหนังสือ How to Solve It ซึ่งเสนอกระบวนการแก้ปัญหา 4 ขั้นตอน คือ การทำความเข้าใจปัญหา การวางแผน การดำเนินการตามแผน และการตรวจสอบย้อนกลับ ขั้นตอนเหล่านี้สอดคล้องกับองค์ประกอบของแนวคิดเชิงคำนวณ โดยเฉพาะการแบ่งย่อยปัญหาที่ปรากฏในขั้นทำความเข้าใจปัญหาและขั้นวางแผน และการออกแบบอัลกอริทึมที่ปรากฏในขั้นดำเนินการตามแผน ในทำนองเดียวกัน Mayer (1992) เสนอโมเดลที่อธิบายกระบวนการแก้ปัญหา 4 ระยะ ได้แก่ การแปลความหมายปัญหา การบูรณาการข้อมูล การวางแผนและตรวจสอบ และการดำเนินการแก้ปัญหา ซึ่งสอดคล้องกับกระบวนการใช้ทักษะการคิดเชิงคำนวณในการแก้ปัญหาตามสถานการณ์ ความสอดคล้องระหว่างโมเดลการแก้ปัญหาที่เป็นที่ยอมรับกับองค์ประกอบของ

แนวคิดเชิงคำนวณ ยืนยันว่าแนวคิดเชิงคำนวณมีรากฐานทางทฤษฎีที่มั่นคง และการวัดทักษะการคิดเชิงคำนวณผ่านสถานการณ์ปัญหาจึงเป็นแนวทางที่สอดคล้องกับธรรมชาติของทักษะนี้

นอกจากความเชื่อมโยงกับการคิดแก้ปัญหาแล้ว แนวคิดเชิงคำนวณยังสัมพันธ์กับทักษะการคิดระดับสูง Tang และคณะ (2020) ได้ทบทวนงานวิจัยด้านการวัดและประเมินทักษะการคิดเชิงคำนวณอย่างเป็นระบบ จำนวน 96 บทความ และพบว่าการพัฒนาทักษะการคิดเชิงคำนวณมีความสัมพันธ์กับการคิดวิเคราะห์ การแก้ปัญหา และผลสัมฤทธิ์ทางการเรียน เมื่อพิจารณาความสัมพันธ์กับการจำแนกระดับพุทธิพิสัยตาม Anderson และ Krathwohl (2001) พบว่าองค์ประกอบของแนวคิดเชิงคำนวณสอดคล้องกับพุทธิพิสัยระดับสูง กล่าวคือ การแบ่งย่อยปัญหาและการพิจารณารูปแบบสอดคล้องกับระดับการวิเคราะห์ การคิดเชิงนามธรรมสอดคล้องกับระดับการประเมิน และการออกแบบอัลกอริทึมสอดคล้องกับระดับการสร้างสรรค์ ความสอดคล้องนี้แสดงว่าพัฒนาทักษะการคิดเชิงคำนวณเป็นการพัฒนาทักษะการคิดระดับสูงของผู้เรียนไปพร้อมกัน

ความสัมพันธ์ระหว่างแนวคิดเชิงคำนวณกับทักษะการคิดระดับสูงนี้มีความสำคัญต่อกลุ่มเป้าหมายของการวิจัย เนื่องจากตามทฤษฎีพัฒนาการทางสติปัญญาของ Piaget (1972) นักเรียนระดับมัธยมศึกษาตอนต้นซึ่งมีอายุ 13 ถึง 15 ปี อยู่ในขั้นการคิดเชิงนามธรรม (formal operational stage) ที่สามารถคิดเชิงตรรกะ ตั้งสมมติฐาน และแก้ปัญหาที่ซับซ้อนได้ นักเรียนกลุ่มนี้จึงมีความพร้อมทางพัฒนาการในการใช้ทักษะการคิดเชิงคำนวณ และเหมาะสมต่อการได้รับการประเมินด้วยแบบวัดที่อาศัยสถานการณ์ปัญหา

1.5 กรอบการประเมินแนวคิดเชิงคำนวณในระดับสากลและระดับชาติ

การประเมินแนวคิดเชิงคำนวณได้รับการพัฒนาทั้งในระดับสากลและระดับชาติ ซึ่งสะท้อนแนวทางการวัดที่หลากหลาย ในระดับสากล Bebras International Challenge เป็นการแข่งขันที่ประเมินแนวคิดเชิงคำนวณของผู้เรียนผ่านการแก้ปัญหาในสถานการณ์จำลอง โดยมีผู้เข้าร่วมจากหลายประเทศทั่วโลก (Dagienė & Futschek, 2008) กรอบการประเมินของ Bebras ให้ความสำคัญกับการวัดแนวคิดเชิงคำนวณผ่านบริบทที่สมจริง โดยไม่ต้องอาศัยความรู้ด้านการเขียนโปรแกรม ลักษณะการประเมินเช่นนี้แสดงให้เห็นว่าการวัดแนวคิดเชิงคำนวณผ่านสถานการณ์ปัญหาที่ใกล้เคียงกับชีวิตจริงเป็นแนวทางที่ได้รับการยอมรับในระดับสากล

ในระดับชาติ สสวท. (2560) ได้กำหนดผลลัพธ์การเรียนรู้ด้านแนวคิดเชิงคำนวณในรายวิชาการคำนวณครอบคลุมทุกระดับชั้น ตั้งแต่ชั้นประถมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 6 โดยเพิ่มความซับซ้อนของสถานการณ์ปัญหาและระดับทักษะที่คาดหวังตามช่วงวัย สำหรับระดับมัธยมศึกษาตอนต้น ผู้เรียนต้องสามารถแบ่งย่อยปัญหา พิจารณารูปแบบ คิดเชิงนามธรรม และออกแบบอัลกอริทึมในสถานการณ์ที่ซับซ้อนขึ้นได้ การกำหนดผลลัพธ์การเรียนรู้ในลักษณะนี้แสดงว่าหลักสูตรของไทยมุ่งพัฒนาทักษะการคิดเชิงคำนวณอย่างเป็นลำดับขั้นตามพัฒนาการของผู้เรียน

1.6 ความสำคัญของการพัฒนาทักษะการคิดเชิงคำนวณในนักเรียนระดับมัธยมศึกษาตอนต้น

การพัฒนาทักษะการคิดเชิงคำนวณในผู้เรียนระดับการศึกษาขั้นพื้นฐาน โดยเฉพาะระดับมัธยมศึกษาตอนต้น มีความสำคัญด้วยเหตุผลหลายประการ Wing (2008) เสนอว่าทักษะการคิดเชิงคำนวณควรเป็นทักษะพื้นฐานที่ผู้เรียนได้รับการพัฒนาตั้งแต่ระดับการศึกษาขั้นพื้นฐาน เช่นเดียวกับการอ่าน การเขียน และคณิตศาสตร์ เพื่อเตรียมความพร้อมในการแก้ปัญหาในยุคที่เทคโนโลยีดิจิทัลเข้ามามีบทบาทสำคัญในทุกด้านของชีวิต

ในด้านพัฒนาการ Bocconi และคณะ (2016) ในรายงานการศึกษาสำหรับ คณะกรรมาธิการยุโรป ระบุว่า ระดับมัธยมศึกษาตอนต้นเป็นช่วงวัยที่เหมาะสมสำหรับการพัฒนาทักษะการคิดเชิงคำนวณ เนื่องจากผู้เรียนในช่วงอายุนี้อาจเริ่มมีพัฒนาการทางการคิดเชิงนามธรรมที่ดีขึ้น สามารถเข้าใจแนวคิดและหลักการที่ซับซ้อนได้ ซึ่งสอดคล้องกับทฤษฎีพัฒนาการทางสติปัญญาของ Piaget (1972) ที่ระบุว่าผู้เรียนในช่วงวัยนี้กำลังก้าวเข้าสู่ขั้นการคิดเชิงนามธรรม อันเป็นช่วงที่สามารถคิดเชิงตรรกะ ตั้งสมมติฐาน และใช้เหตุผลกับปัญหาที่ซับซ้อนได้ การจัดประสบการณ์การเรียนรู้ที่ส่งเสริมทักษะการคิดเชิงคำนวณในช่วงวัยนี้จึงเป็นโอกาสที่เหมาะสมในการต่อยอดศักยภาพของผู้เรียน

ในบริบทของประเทศไทย หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560) ได้กำหนดให้สาระเทคโนโลยี (วิทยาการคำนวณ) เป็นสาระบังคับในระดับมัธยมศึกษาตอนต้น โดยมุ่งเน้นการพัฒนาทักษะการคิดเชิงคำนวณ ทักษะการคิดอย่างมีวิจารณญาณ และทักษะการแก้ปัญหา (สสวท., 2560) การจัดทำหลักสูตรในลักษณะนี้สะท้อนให้เห็นนโยบายของกระทรวงศึกษาธิการที่ให้ความสำคัญกับการเตรียมผู้เรียนสำหรับสังคมและเศรษฐกิจในยุคดิจิทัล

อย่างไรก็ตาม การพัฒนาผู้เรียนให้มีทักษะการคิดเชิงคำนวณตามเป้าหมายของหลักสูตร จำเป็นต้องมีเครื่องมือวัดและประเมินผลที่สามารถให้ข้อมูลเกี่ยวกับระดับทักษะของผู้เรียนได้อย่างเหมาะสม การประเมินทักษะการคิดเชิงคำนวณยังเป็นความท้าทาย เนื่องจากทักษะดังกล่าวเป็นกระบวนการคิดที่เกี่ยวข้องกับการวิเคราะห์ปัญหา การค้นหารูปแบบ การคิดเชิงนามธรรม และการออกแบบขั้นตอนการแก้ปัญหา ซึ่งไม่สามารถสะท้อนได้อย่างเพียงพอจากการวัดความรู้เชิงเนื้อหาเพียงอย่างเดียว การประเมินจึงควรเปิดโอกาสให้ผู้เรียนได้ประยุกต์ใช้ทักษะดังกล่าวผ่านสถานการณ์ที่มีลักษณะใกล้เคียงกับการแก้ปัญหาในชีวิตจริง

ด้วยเหตุนี้ การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์ที่ผ่านการตรวจสอบคุณภาพทางจิตมิติและมีเกณฑ์สำหรับการแปลผลคะแนน จึงเป็นแนวทางหนึ่งที่จะช่วยให้การประเมินทักษะการคิดเชิงคำนวณมีความชัดเจนและสามารถนำผลการประเมินไปใช้ประกอบการพิจารณาระดับทักษะของผู้เรียนได้ ซึ่งเป็นที่มาของการพัฒนาแบบวัดในการวิจัยครั้งนี้

1.7 นิยามทักษะการคิดเชิงคำนวณตามที่ผู้วิจัยกำหนด

จากการทบทวนเอกสาร แนวคิด และงานวิจัยที่เกี่ยวข้องในหัวข้อ 1.1 ถึง 1.6 ผู้วิจัยได้สังเคราะห์นิยามและกรอบองค์ประกอบของทักษะการคิดเชิงคำนวณที่ใช้เป็นกรอบแนวคิดในการวิจัยครั้งนี้

ทักษะการคิดเชิงคำนวณ ในการวิจัยครั้งนี้ หมายถึง ความสามารถของนักเรียนระดับมัธยมศึกษาตอนต้นในการแก้ปัญหาอย่างเป็นระบบภายใต้สถานการณ์ที่กำหนด โดยประยุกต์ใช้องค์ประกอบของการคิดเชิงคำนวณร่วมกัน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม เพื่อวิเคราะห์ปัญหา สังเคราะห์สารสนเทศ และกำหนดแนวทางการแก้ปัญหาอย่างมีเหตุผล

นิยามดังกล่าวสังเคราะห์จากแนวคิดของ Wing (2006, 2008, 2010) ซึ่งอธิบายการคิดเชิงคำนวณในฐานะกระบวนการแก้ปัญหาอย่างเป็นระบบ ร่วมกับกรอบแนวคิดของ Grover และ Pea (2013) และกรอบของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท., 2560) ซึ่งกำหนดองค์ประกอบสำคัญของการคิดเชิงคำนวณไว้ 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม

ในการวิจัยครั้งนี้ ผู้วิจัยกำหนดให้องค์ประกอบทั้ง 4 ด้านเป็นองค์ประกอบร่วมของทักษะการคิดเชิงคำนวณ และใช้เป็นกรอบในการพัฒนาแบบวัดเชิงสถานการณ์แบบออนไลน์ โดยมีรายละเอียดของแต่ละองค์ประกอบและพฤติกรรมบ่งชี้ตามที่นำเสนอไว้ในหัวข้อ 1.3

นิยามและกรอบองค์ประกอบทั้ง 4 ด้านนี้เป็นพื้นฐานในการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ โดยผู้วิจัยออกแบบให้แต่ละสถานการณ์ประกอบด้วยข้อคำถามที่ครอบคลุมองค์ประกอบทั้ง 4 ด้านภายใต้บริบทเดียวกัน ทั้งนี้เพื่อให้การประเมินสอดคล้องกับลักษณะของการแก้ปัญหาในสถานการณ์ที่ผู้เรียนต้องใช้ในการแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึมร่วมกัน

2. แบบวัดเชิงสถานการณ์

แบบวัดเชิงสถานการณ์ (situational assessment) เป็นแนวทางการวัดที่ได้รับความสนใจอย่างกว้างขวางในวงการการศึกษาและการประเมิน โดย Situational Judgment Test (SJT) ถือเป็นรูปแบบหนึ่งของการประเมินเชิงสถานการณ์ที่ได้รับการศึกษาและพัฒนาอย่างต่อเนื่อง แนวคิดร่วมของการประเมินลักษณะนี้คือการใช้สถานการณ์เป็นบริบทในการวัดความสามารถของผู้เรียนหรือผู้ตอบแบบวัด ภายใต้เงื่อนไขที่ใกล้เคียงกับการปฏิบัติจริง ในหัวข้อนี้ ผู้วิจัยนำเสนอเนื้อหาที่เกี่ยวข้องกับแบบวัดเชิงสถานการณ์ ประกอบด้วย ความหมายและพัฒนาการของแบบวัดเชิงสถานการณ์ ลักษณะและองค์ประกอบของแบบวัดเชิงสถานการณ์ หลักการสร้างแบบวัดเชิงสถานการณ์ที่มี

คุณภาพ ประสิทธิภาพของแบบวัดเชิงสถานการณ์ ข้อดีและข้อจำกัดของแบบวัดเชิงสถานการณ์ และการประยุกต์ใช้แนวคิดดังกล่าวในการวิจัยครั้งนี้

2.1 ความหมายและการพัฒนาการของแบบวัดเชิงสถานการณ์

แบบวัดเชิงสถานการณ์เป็นแนวทางการประเมินที่มีรากฐานทางทฤษฎีจากหลายสาขา ทั้งจิตวิทยาการรู้คิด จิตวิทยาสังคม และทฤษฎีการเรียนรู้เชิงสถานการณ์ โดยมีแนวคิดร่วมกันว่า พฤติกรรมและความสามารถของบุคคลเกิดจากปฏิสัมพันธ์ระหว่างลักษณะของบุคคลกับบริบท แวดล้อมที่บุคคลเผชิญ (Bandura, 1977) ดังนั้น การประเมินความสามารถจึงไม่ควรพิจารณาเฉพาะ ความรู้หรือศักยภาพของบุคคลเพียงลำพัง แต่ควรคำนึงถึงบริบทที่บุคคลต้องนำความรู้นั้นไปใช้ในการ ตัดสินใจหรือแก้ปัญหาด้วย การใช้สถานการณ์ที่มีลักษณะใกล้เคียงกับชีวิตจริงในการประเมินจึงช่วยให้การวัดมีความสอดคล้องกับบริบทการแสดงความสามารถของผู้เรียนมากยิ่งขึ้น

แบบวัดเชิงสถานการณ์มีพัฒนาการมาอย่างต่อเนื่อง โดยเริ่มจากการประยุกต์ใช้ใน สาขาจิตวิทยาอุตสาหกรรมและองค์การเพื่อการคัดเลือกและประเมินบุคลากร ก่อนขยายสู่บริบท ทางการศึกษา Motowidlo, Dunnette และ Carter (1990) ได้นำเสนอแนวคิดแบบวัดเชิงสถานการณ์ ในรูปแบบที่เป็นระบบ โดยอธิบายว่าเป็นเครื่องมือที่นำเสนอสถานการณ์จำลองให้ผู้สอบพิจารณา และให้เลือกหรือประเมินทางเลือกในการตอบสนองที่เหมาะสมที่สุด การประเมินลักษณะนี้ช่วยให้ ผู้ประเมินสามารถศึกษากระบวนการตัดสินใจของผู้สอบภายใต้บริบทที่กำหนดได้ ต่อมาแนวคิด ดังกล่าวได้รับการนำมาประยุกต์ใช้ในการวัดความสามารถและทักษะที่เกี่ยวข้องกับการคิด การ ตัดสินใจ และการแก้ปัญหาในบริบททางการศึกษาอย่างกว้างขวาง ซึ่งสอดคล้องกับลักษณะของ ทักษะการคิดเชิงคำนวณที่มุ่งเน้นการประยุกต์ใช้กระบวนการคิดเพื่อแก้ปัญหาในสถานการณ์ต่าง ๆ

McDaniel และคณะ (2001) ได้ทบทวนงานวิจัยเกี่ยวกับแบบวัดเชิงสถานการณ์อย่างเป็นระบบ และให้นิยามว่าเป็นวิธีการประเมินที่ใช้สถานการณ์ที่เกี่ยวข้องกับงานหรือกิจกรรมจริง เป็นสิ่งกระตุ้น และให้ผู้สอบพิจารณาหรือเลือกแนวทางการตอบสนองที่เหมาะสมต่อสถานการณ์นั้น ผลการทบทวนของผู้วิจัยดังกล่าวพบว่าแบบวัดเชิงสถานการณ์มีหลักฐานสนับสนุนด้านความตรงในการประเมินพฤติกรรมและการตัดสินใจภายใต้บริบทที่กำหนด ต่อมา Weekley และ Ployhart (2006) ได้นิยามแบบวัดเชิงสถานการณ์ว่าเป็นวิธีการประเมินที่นำเสนอสถานการณ์ให้ผู้สอบอาจเผชิญ ในบริบทจริง และให้ผู้สอบเลือกการตอบสนองที่เหมาะสมที่สุด หรือประเมินความเหมาะสมของ ทางเลือกที่กำหนดไว้ นิยามดังกล่าวสะท้อนลักษณะสำคัญของแบบวัดเชิงสถานการณ์ 2 ประการ ได้แก่ (1) การใช้สถานการณ์เป็นบริบทในการประเมิน และ (2) การให้ผู้สอบพิจารณาและตัดสินใจ เลือกแนวทางการตอบสนองต่อสถานการณ์นั้น

การประยุกต์ใช้แบบวัดเชิงสถานการณ์ได้ขยายจากบริบทการคัดเลือกบุคลากรสู่บริบท ทางการศึกษาอย่างต่อเนื่อง โดย Lievens, Peeters และ Schollaert (2008) พบว่าแบบวัดเชิงสถานการณ์

สามารถนำมาใช้ประเมินความสามารถที่เกี่ยวข้องกับการคิด การตัดสินใจ และการแก้ปัญหาได้ เนื่องจากเปิดโอกาสให้ผู้เรียนแสดงการประยุกต์ใช้ความรู้และกระบวนการคิดภายใต้บริบทที่กำหนด คุณลักษณะดังกล่าวจึงสอดคล้องกับการประเมินทักษะการคิดเชิงคำนวณ ซึ่งมุ่งวัดความสามารถในการใช้กระบวนการคิดเพื่อวิเคราะห์และแก้ปัญหาในสถานการณ์ต่าง ๆ มากกว่าการวัดความรู้เชิงเนื้อหาเพียงอย่างเดียว

สำหรับบริบทการศึกษาไทย นักวิชาการด้านการวัดและประเมินผลได้ให้นิยามแบบวัดเชิงสถานการณ์ไว้หลากหลาย พิชิต ฤทธิ์จรูญ (2556) อธิบายว่าการทดสอบแบบสถานการณ์คือการทดสอบที่กำหนดสถานการณ์ให้ผู้รับการทดสอบแสดงพฤติกรรมตามที่กำหนด ซึ่งสะท้อนความสามารถในการประยุกต์ใช้ความรู้และทักษะได้ดีกว่าการทดสอบแบบดั้งเดิม วันเพ็ญ จันทสุวรรณ (2550) นิยามว่าแบบวัดเชิงสถานการณ์เป็นแบบทดสอบที่กำหนดสถานการณ์จำลอง ข้อความ หรือภาพ ที่คล้ายกับเหตุการณ์ในชีวิตจริง แล้วให้ผู้ตอบแสดงพฤติกรรมตอบสนองที่สังเกตหรือวัดได้ สมนึก ภัททิยธนี (2562) ระบุว่าแบบวัดเชิงสถานการณ์เป็นแบบสอบที่กำหนดข้อความ ภาพ หรือตาราง ให้นักเรียนพิจารณาก่อน แล้วตั้งคำถามให้ผู้ตอบวิเคราะห์และตีความ ส่วน อนุวัติ คุณแก้ว (2566) อธิบายว่าแบบวัดเชิงสถานการณ์เป็นแบบสอบที่กำหนดเรื่องราวหรือสถานการณ์สมมติ ให้นักเรียนอ่านแล้วแสดงความคิดและเหตุผล และเพิ่มเติมว่าแบบวัดเชิงสถานการณ์ที่มีคุณภาพต้องคำนึงถึงความเป็นมิตรกับผู้ใช้ควบคู่กับคุณสมบัติทางจิตมิติ

จากการศึกษาเอกสารและงานวิจัยข้างต้น ผู้วิจัยสรุปได้ว่า แบบวัดเชิงสถานการณ์ คือ เครื่องมือวัดที่นำเสนอสถานการณ์จำลองหรือบริบทที่ใกล้เคียงกับชีวิตจริง เพื่อให้ผู้สอบวิเคราะห์ข้อมูล พิจารณาทางเลือก หรือตอบคำถามภายใต้สถานการณ์ที่กำหนด โดยมุ่งประเมินความสามารถในการประยุกต์ใช้ความรู้ ทักษะ และกระบวนการคิดในการแก้ปัญหา การวัดลักษณะนี้แตกต่างจากการวัดความรู้เชิงเนื้อหาเพียงอย่างเดียว เนื่องจากให้ความสำคัญกับการนำความรู้ไปใช้ภายใต้บริบทที่มีความหมายต่อผู้สอบ จึงเหมาะสำหรับการประเมินทักษะที่เกี่ยวข้องกับการคิดและการแก้ปัญหา เช่น ทักษะการแก้ปัญหา ทักษะการคิดอย่างมีวิจารณญาณ และทักษะการคิดเชิงคำนวณ

2.2 ลักษณะและองค์ประกอบของแบบวัดเชิงสถานการณ์

แบบวัดเชิงสถานการณ์มีลักษณะเฉพาะที่แตกต่างจากแบบทดสอบทั่วไป เนื่องจากใช้สถานการณ์ที่ใกล้เคียงกับบริบทจริงเป็นฐานในการประเมินความสามารถของผู้เรียน โดยองค์ประกอบสำคัญของแบบวัดประกอบด้วย 3 ส่วน ได้แก่ สถานการณ์กระตุ้น ข้อคำถามและทางเลือกคำตอบ และระบบการให้คะแนน ซึ่งองค์ประกอบทั้งสามส่วนต้องได้รับการออกแบบให้สอดคล้องกับวัตถุประสงค์ของการวัดและลักษณะของทักษะที่ต้องการประเมิน เพื่อให้ผลการวัดสะท้อนความสามารถของผู้สอบได้อย่างเหมาะสม (McDaniel et al., 2001; Weekley & Ployhart, 2006; วันเพ็ญ จันทสุวรรณ, 2550; อนุวัติ คุณแก้ว, 2566)

2.2.1 สถานการณ์กระตุ้น

สถานการณ์กระตุ้น (stem scenario) เป็นองค์ประกอบสำคัญของแบบวัดเชิงสถานการณ์ เนื่องจากเป็นบริบทที่กำหนดเงื่อนไขให้ผู้สอบใช้ความรู้ ทักษะ หรือกระบวนการคิด ในการตอบคำถาม Weekley และ Ployhart (2006) อธิบายว่า สถานการณ์กระตุ้นคือเหตุการณ์หรือบริบทจำลองที่นำเสนอแก่ผู้สอบก่อนการตอบคำถาม โดยสถานการณ์ที่มีคุณภาพควรมีลักษณะสำคัญหลายประการ ได้แก่ (1) ความสมจริง เพื่อให้ผู้สอบสามารถเชื่อมโยงกับประสบการณ์หรือบริบทที่คุ้นเคย (2) ความเป็นกลาง โดยหลีกเลี่ยงอคติที่อาจส่งผลต่อผลการวัด (3) ระดับความซับซ้อนที่เหมาะสม ซึ่งสามารถกระตุ้นการคิดของผู้สอบได้โดยไม่ก่อให้เกิดความสับสนเกินความจำเป็น และ (4) ความเพียงพอของข้อมูล เพื่อให้ผู้สอบสามารถใช้ข้อมูลที่ได้รับการประกอบการวิเคราะห์และตอบคำถามได้อย่างเหมาะสม ทั้งนี้ McDaniel และคณะ (2001) ระบุว่า สถานการณ์สามารถนำเสนอได้หลายรูปแบบ เช่น ข้อความ ภาพ วิดีโอ หรือสื่อมัลติมีเดีย โดยการเลือกรูปแบบขึ้นอยู่กับลักษณะของทักษะที่ต้องการวัดและบริบทของการนำไปใช้

ในการวิจัยครั้งนี้ ผู้วิจัยพัฒนาสถานการณ์จากบริบทชีวิตประจำวันของนักเรียนระดับมัธยมศึกษาตอนต้น โดยครอบคลุม 3 ด้าน ได้แก่ ชีวิตส่วนบุคคล การจัดกิจกรรมในโรงเรียน และชุมชน สังคม และสิ่งแวดล้อม ทั้งนี้ การเลือกใช้บริบทที่ใกล้เคียงกับประสบการณ์ของผู้เรียนมีจุดมุ่งหมายเพื่อเปิดโอกาสให้นักเรียนสามารถแสดงการประยุกต์ใช้ทักษะการคิดเชิงคำนวณในการวิเคราะห์และแก้ปัญหาภายใต้สถานการณ์ที่มีความหมายและสอดคล้องกับชีวิตประจำวันของตน

2.2.2 ข้อคำถามและทางเลือกคำตอบ

ข้อคำถามและทางเลือกคำตอบเป็นองค์ประกอบสำคัญของแบบวัดเชิงสถานการณ์ เนื่องจากเป็นส่วนที่ใช้ประเมินความสามารถของผู้สอบหลังจากการพิจารณาสถานการณ์ที่กำหนด ข้อคำถามควรสอดคล้องกับสถานการณ์และสะท้อนพฤติกรรมหรือกระบวนการคิดที่ต้องการวัดอย่างชัดเจน ขณะที่ทางเลือกคำตอบควรได้รับการออกแบบให้มีความสมเหตุสมผลและสอดคล้องกับบริบทของสถานการณ์ เพื่อให้สามารถจำแนกความแตกต่างของระดับความสามารถระหว่างผู้สอบได้อย่างมีประสิทธิภาพ

Haladyna และ Rodriguez (2013) อธิบายว่า ทางเลือกคำตอบที่มีคุณภาพ ควรประกอบด้วยคำตอบที่ถูกต้องเพียงคำตอบเดียว และตัวลวงที่มีความน่าเชื่อถือ สามารถดึงดูดผู้สอบที่ยังมีความเข้าใจไม่สมบูรณ์ได้ โดยหลีกเลี่ยงทางเลือกที่ผิดอย่างชัดเจนหรือมีลักษณะคลุมเครือ นอกจากนี้ ทางเลือกคำตอบควรสะท้อนระดับการคิดที่แตกต่างกัน เพื่อให้ข้อคำถามสามารถจำแนกผู้สอบที่มีความสามารถต่างระดับได้อย่างเหมาะสม

สำหรับจำนวนทางเลือก งานวิจัยของ Haladyna และ Rodriguez (2013) สนับสนุนว่าข้อสอบปรนัยแบบ 4 ตัวเลือกมีความเหมาะสมทั้งในด้านคุณภาพการวัดและประสิทธิภาพในการสร้างข้อสอบ เนื่องจากสามารถรักษาคุณภาพของตัวลงได้ดี ขณะเดียวกันยังลดความยุ่งยากในการพัฒนาทางเลือกที่มีความน่าเชื่อถือ เมื่อเปรียบเทียบกับข้อสอบที่มีจำนวนทางเลือกมากกว่า

ในการวิจัยครั้งนี้ ผู้วิจัยออกแบบข้อคำถามแบบปรนัยชนิด 4 ตัวเลือก โดยแต่ละข้อมีคำตอบที่ถูกต้องเพียงคำตอบเดียว และสร้างตัวลงจากข้อผิดพลาดหรือแนวคิดที่อาจเกิดขึ้นได้จริงในการวิเคราะห์และแก้ปัญหา เพื่อให้ข้อคำถามสามารถวัดทักษะการคิดเชิงคำนวณและจำแนกระดับความสามารถของผู้เรียนได้อย่างเหมาะสม

2.2.3 ระบบการให้คะแนน

ระบบการให้คะแนนเป็นองค์ประกอบสำคัญของแบบวัด เนื่องจากมีผลต่อความถูกต้องของการวัด การแปลผลคะแนน และการวิเคราะห์คุณภาพของเครื่องมือ โดยรูปแบบการให้คะแนนควรสอดคล้องกับลักษณะของข้อคำถามและวัตถุประสงค์ของการวัด สำหรับแบบวัดเชิงสถานการณ์ รูปแบบการให้คะแนนอาจแตกต่างกันไปตามลักษณะของคำตอบที่ต้องการวัด เช่น การให้คะแนนแบบทวิภาคสำหรับข้อคำถามที่มีคำตอบถูกต้องชัดเจน หรือการให้คะแนนแบบหลายระดับสำหรับข้อคำถามที่เปิดโอกาสให้ผู้ตอบแสดงระดับความเหมาะสมของคำตอบที่แตกต่างกัน (Weekley & Ployhart, 2006)

ในการวิจัยครั้งนี้ ผู้วิจัยใช้ระบบการให้คะแนนแบบทวิภาค (dichotomous scoring) โดยกำหนดให้คำตอบที่ถูกต้องได้ 1 คะแนน และคำตอบที่ไม่ถูกต้องได้ 0 คะแนน เนื่องจากข้อคำถามแต่ละข้อได้รับการพัฒนาให้มีคำตอบที่ถูกต้องเพียงคำตอบเดียวตามหลักการของทักษะการคิดเชิงคำนวณ การให้คะแนนในลักษณะดังกล่าวช่วยให้การตรวจให้คะแนนเป็นไปอย่างมีมาตรฐาน ลดความคลาดเคลื่อนจากผู้ตรวจ และรองรับการประมวลผลผ่านระบบออนไลน์ได้อย่างมีประสิทธิภาพ

นอกจากนี้ การให้คะแนนแบบทวิภาคยังสอดคล้องกับการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิม และเหมาะสมกับการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบสำหรับข้อมูลแบบสองค่า ซึ่งเป็นกรอบการวิเคราะห์ที่ใช้ในการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดในการวิจัยครั้งนี้

2.3 หลักการสร้างแบบวัดเชิงสถานการณ์ที่มีคุณภาพ

การพัฒนาแบบวัดเชิงสถานการณ์ให้มีคุณภาพจำเป็นต้องดำเนินการอย่างเป็นระบบ ตั้งแต่การกำหนดกรอบแนวคิดของสิ่งที่ต้องการวัด การออกแบบสถานการณ์และข้อคำถาม ไปจนถึงการตรวจสอบคุณภาพของเครื่องมือ เพื่อให้แบบวัดสามารถสะท้อนความสามารถของผู้เรียนได้อย่างตรงตามวัตถุประสงค์และให้ผลการวัดที่น่าเชื่อถือ นักวิชาการหลายท่านได้เสนอหลักการพัฒนาแบบวัด

เชิงสถานการณ์ไว้ในแนวทางที่สอดคล้องกัน (McDaniel และคณะ, 2001; Weekley & Ployhart, 2006; Lievens และคณะ, 2008) ซึ่งสามารถสรุปเป็นขั้นตอนสำคัญได้ 5 ขั้นตอน ดังนี้

ขั้นที่ 1 การกำหนดกรอบแนวคิดและองค์ประกอบของสิ่งที่ต้องการวัด

ผู้พัฒนาต้องศึกษาทฤษฎี เอกสาร และงานวิจัยที่เกี่ยวข้อง เพื่อกำหนดองค์ประกอบพฤติกรรมบ่งชี้ และขอบเขตของคุณลักษณะที่ต้องการวัดให้ชัดเจน เนื่องจากกรอบแนวคิดดังกล่าวเป็นพื้นฐานสำคัญในการออกแบบสถานการณ์และข้อคำถามให้สอดคล้องกับสิ่งที่ต้องการวัด

ขั้นที่ 2 การพัฒนาสถานการณ์

สถานการณ์ควรสะท้อนบริบทที่ใกล้เคียงกับชีวิตจริงของกลุ่มเป้าหมาย และเปิดโอกาสให้ผู้ตอบสนองกระบวนการคิดหรือความสามารถที่ต้องการวัดได้อย่างเหมาะสม McDaniel และคณะ (2001) เสนอว่าสถานการณ์ควรพัฒนาจากบริบทที่เกิดขึ้นจริงหรือจากข้อมูลที่ได้รับจากผู้เชี่ยวชาญ เพื่อเพิ่มความสมจริงและความสอดคล้องกับสภาพการณ์ที่ผู้เรียนอาจเผชิญ

ขั้นที่ 3 การพัฒนาข้อคำถามและทางเลือกคำตอบ

ข้อคำถามควรเชื่อมโยงกับสถานการณ์ที่กำหนด และมุ่งวัดพฤติกรรมหรือองค์ประกอบของคุณลักษณะที่ต้องการวัดอย่างชัดเจน ส่วนทางเลือกคำตอบควรมีความสมเหตุสมผลสอดคล้องกับบริบทของสถานการณ์ และสามารถจำแนกความแตกต่างของระดับความสามารถระหว่างผู้ตอบได้

ขั้นที่ 4 การกำหนดคำตอบและระบบการให้คะแนน

ผู้พัฒนาต้องกำหนดคำตอบที่ถูกต้องและหลักเกณฑ์การให้คะแนนให้สอดคล้องกับกรอบแนวคิดที่ใช้ในการพัฒนาเครื่องมือ พร้อมทั้งจัดทำเหตุผลรองรับคำตอบแต่ละข้ออย่างชัดเจน เพื่อใช้ในการตรวจสอบคุณภาพของข้อคำถามและเป็นหลักฐานสนับสนุนความตรงของเครื่องมือ

ขั้นที่ 5 การตรวจสอบคุณภาพของแบบวัด

แบบวัดที่พัฒนาขึ้นควรได้รับการตรวจสอบคุณภาพทั้งด้านความตรงและความเชื่อมั่น โดยเริ่มจากการตรวจสอบความตรงเชิงเนื้อหาจากผู้เชี่ยวชาญ และทดลองใช้กับกลุ่มตัวอย่างเพื่อตรวจสอบคุณภาพรายข้อและคุณสมบัติทางจิตมิติของเครื่องมือ ก่อนนำไปใช้จริง (Lievens และคณะ, 2008)

จากขั้นตอนข้างต้น ผู้วิจัยสังเคราะห์หลักการสำคัญในการสร้างแบบวัดเชิงสถานการณ์ที่มีคุณภาพได้ 3 ด้าน ดังนี้

2.3.1 หลักการด้านการกำหนดสถานการณ์

สถานการณ์ที่ใช้ในแบบวัดควรเป็นสถานการณ์ที่ใกล้เคียงกับประสบการณ์จริงของกลุ่มเป้าหมาย มีความชัดเจน เข้าใจง่าย และให้ข้อมูลเพียงพอต่อการวิเคราะห์และตัดสินใจ โดยหลีกเลี่ยงอคติทางวัฒนธรรม สังคม หรือประสบการณ์เฉพาะกลุ่มที่อาจส่งผลต่อผลการวัดโดย

ไม่เกี่ยวข้องกับทักษะที่ต้องการประเมิน (พิชิต ฤทธิ์จรูญ, 2556) สำหรับการวัดทักษะการคิดเชิงคำนวณ Shute, Sun และ Asbell-Clarke (2017) เสนอว่าการใช้สถานการณ์จากชีวิตประจำวันช่วยให้ผู้เรียนสามารถแสดงกระบวนการคิดในการแก้ปัญหาได้โดยไม่ถูกจำกัดด้วยความรู้ด้านการเขียนโปรแกรมหรือความรู้เฉพาะทางอื่น ส่งผลให้ผลการวัดสะท้อนทักษะที่ต้องการประเมินได้ชัดเจนยิ่งขึ้น

ในการวิจัยครั้งนี้ ผู้วิจัยออกแบบสถานการณ์จากบริบทชีวิตประจำวันของนักเรียนระดับมัธยมศึกษาตอนต้น ครอบคลุมด้านชีวิตส่วนบุคคล การจัดกิจกรรมในโรงเรียน และชุมชน สังคม และสิ่งแวดล้อม เพื่อให้สถานการณ์มีความสอดคล้องกับประสบการณ์ของผู้เรียนและเอื้อต่อการแสดงทักษะการคิดเชิงคำนวณในบริบทที่มีความหมายต่อผู้ตอบ

2.3.2 หลักการด้านการเขียนข้อคำถาม

ข้อคำถามควรเชื่อมโยงกับสถานการณ์ที่กำหนดอย่างชัดเจน และมุ่งวัดพฤติกรรมหรือกระบวนการคิดตามองค์ประกอบที่กำหนดไว้ในกรอบแนวคิดของการวิจัย โดยหลีกเลี่ยงการใช้ถ้อยคำที่กำกวม ชับซ้อนเกินความจำเป็น หรือเปิดโอกาสให้ตีความได้หลายความหมาย (อนุวัติ คุณแก้ว, 2566) ทั้งนี้ ข้อคำถามควรออกแบบให้ผู้ตอบต้องอาศัยข้อมูลจากสถานการณ์ในการวิเคราะห์และเลือกคำตอบ ไม่ใช่สามารถตอบได้จากความรู้ทั่วไปเพียงอย่างเดียว

สำหรับการวิจัยครั้งนี้ ผู้วิจัยกำหนดให้แต่ละสถานการณ์ประกอบด้วยข้อคำถาม 4 ข้อ โดยวัดทักษะการคิดเชิงคำนวณทั้ง 4 องค์ประกอบ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม ตามกรอบของ สสวท. (2560) เพื่อให้การวัดสะท้อนการประยุกต์ใช้ทักษะแต่ละด้านภายใต้บริบทเดียวกัน

2.4 ประเภทของแบบวัดเชิงสถานการณ์

แบบวัดเชิงสถานการณ์สามารถจำแนกได้หลายลักษณะตามวัตถุประสงค์และรูปแบบการใช้งาน ทั้งนี้ การจำแนกประเภทช่วยให้ผู้พัฒนาเครื่องมือสามารถเลือกแนวทางที่เหมาะสมกับลักษณะของคุณลักษณะที่ต้องการวัดและบริบทของกลุ่มเป้าหมาย ในการวิจัยครั้งนี้ ผู้วิจัยนำเสนอการจำแนกที่เกี่ยวข้องกับการพัฒนาเครื่องมือวัดทักษะการคิดเชิงคำนวณ 3 มิติ ได้แก่ การจำแนกตามรูปแบบการตอบ การจำแนกตามลักษณะการวัด และการจำแนกตามลักษณะการนำเสนอ

2.4.1 การจำแนกตามรูปแบบการตอบ

Whetzel และ McDaniel (2009) จำแนกแบบวัดเชิงสถานการณ์ตามรูปแบบการตอบออกเป็น 2 ประเภท ได้แก่ แบบให้เลือกตอบ (selected-response format) และแบบให้สร้างคำตอบ (constructed-response format)

แบบให้เลือกตอบเป็นรูปแบบที่ผู้ตอบเลือกคำตอบจากทางเลือกที่กำหนดไว้ล่วงหน้า มีข้อดีคือสามารถตรวจให้คะแนนได้อย่างรวดเร็ว ความเป็นปรนัย และเหมาะสำหรับการเก็บข้อมูลกับกลุ่มตัวอย่างจำนวนมาก ส่วนแบบให้สร้างคำตอบเปิดโอกาสให้ผู้ตอบอธิบายแนวคิดหรือ

เหตุผลของตนเองได้อย่างละเอียด แต่ต้องใช้ผู้ตรวจให้คะแนนและมีความซับซ้อนในการวิเคราะห์ข้อมูลมากกว่า

ในการวิจัยครั้งนี้ ผู้วิจัยเลือกใช้แบบให้เลือกตอบชนิด 4 ตัวเลือก เนื่องจากสอดคล้องกับการให้คะแนนแบบทวิภาค การประมวลผลผ่านระบบออนไลน์ และการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ

2.4.2 การจำแนกตามวัตถุประสงค์ของการวัด

Weekley และ Ployhart (2006) อธิบายว่าแบบวัดเชิงสถานการณ์อาจมีวัตถุประสงค์ในการวัดแตกต่างกัน บางเครื่องมือมุ่งวัดการตัดสินใจหรือการเลือกพฤติกรรมที่เหมาะสม ขณะที่บางเครื่องมือมุ่งวัดความสามารถในการวิเคราะห์และแก้ปัญหาภายใต้สถานการณ์ที่กำหนด

ในการวิจัยครั้งนี้ ผู้วิจัยใช้สถานการณ์เป็นบริบทในการวัดความสามารถของผู้เรียนในการประยุกต์ใช้ทักษะการคิดเชิงคำนวณเพื่อวิเคราะห์และแก้ปัญหา ดังนั้น แบบวัดที่พัฒนาขึ้นจึงอยู่ในลักษณะของการวัดความสามารถ (ability-based assessment) มากกว่าการวัดแนวโน้มพฤติกรรมหรือการตัดสินใจเชิงบุคลิกภาพ โดยผู้ตอบต้องใช้กระบวนการคิดตามองค์ประกอบของทักษะการคิดเชิงคำนวณในการเลือกคำตอบที่ถูกต้องภายใต้สถานการณ์ที่กำหนด

2.4.3 การจำแนกตามลักษณะการนำเสนอ

แบบวัดเชิงสถานการณ์สามารถจำแนกตามลักษณะการนำเสนอได้หลายรูปแบบ โดยอาจนำเสนอสถานการณ์ผ่านข้อความ ภาพ วิดีโอ หรือสื่อผสม ทั้งนี้การเลือกใช้รูปแบบใดขึ้นอยู่กับลักษณะของสิ่งที่ต้องการวัด กลุ่มเป้าหมาย และข้อจำกัดด้านการดำเนินการ (Weekley & Ployhart, 2006)

สำหรับการวิจัยครั้งนี้ ผู้วิจัยพัฒนาแบบวัดในรูปแบบออนไลน์ โดยนำเสนอสถานการณ์ผ่านข้อความบรรยายเป็นหลัก และให้ผู้สอบตอบข้อคำถามผ่านระบบออนไลน์ เหตุผลสำคัญในการเลือกใช้รูปแบบดังกล่าวมี 3 ประการ ได้แก่

ประการแรก สถานการณ์ในรูปข้อความสามารถถ่ายทอดข้อมูลที่จำเป็นต่อการวิเคราะห์และแก้ปัญหาได้อย่างครบถ้วน และสอดคล้องกับระดับพัฒนาการของนักเรียนระดับมัธยมศึกษาตอนต้น ซึ่งสามารถทำความเข้าใจสถานการณ์จากการอ่านได้อย่างเหมาะสม

ประการที่สอง การใช้ข้อความช่วยลดอิทธิพลของปัจจัยแทรกซ้อนที่อาจเกิดจากความแตกต่างด้านอุปกรณ์ ความเร็วอินเทอร์เน็ต หรือการรับรู้สื่อมัลติมีเดียของผู้สอบ ทำให้ผลการวัดสะท้อนทักษะการคิดเชิงคำนวณได้ชัดเจนยิ่งขึ้น

ประการที่สาม รูปแบบออนไลน์เอื้อต่อการบริหารจัดการการทดสอบ การรวบรวมข้อมูล และการประมวลผลคะแนน โดยผู้สอบสามารถเข้าถึงแบบวัดผ่านระบบออนไลน์ และได้รับรายงานผลการประเมินทั้งคะแนนรวมและคะแนนรายองค์ประกอบภายหลังการทดสอบเสร็จสิ้น ด้วยเหตุนี้ ผู้วิจัยจึงเลือกพัฒนาแบบวัดในรูปแบบออนไลน์ที่ใช้สถานการณ์แบบข้อความเป็นฐาน เพื่อให้สอดคล้องกับลักษณะของทักษะที่ต้องการวัดและบริบทการใช้งานของกลุ่มเป้าหมายในการวิจัยครั้งนี้

2.5 ข้อดีและข้อจำกัดของแบบวัดเชิงสถานการณ์

แบบวัดเชิงสถานการณ์มีทั้งข้อดีและข้อจำกัดที่ผู้พัฒนาและผู้ใช้ควรพิจารณา การเข้าใจทั้งสองด้านช่วยให้ประโยชน์จากแบบวัดได้อย่างเหมาะสมและตระหนักถึงข้อจำกัดในการตีความผล

2.5.1 ข้อดีของแบบวัดเชิงสถานการณ์

McDaniel และคณะ (2001) ระบุข้อดีสำคัญของแบบวัดเชิงสถานการณ์หลายประการ ประการที่หนึ่ง ความตรงเชิงเนื้อหาสูง เนื่องจากสถานการณ์ในแบบวัดสะท้อนบริบทจริง ทำให้สิ่งที่วัดมีความใกล้เคียงกับการแสดงทักษะในชีวิตจริง ประการที่สอง มีความตรงเชิงพยากรณ์ที่ยอมรับได้ในการทำนายการปฏิบัติงานหรือพฤติกรรม ประการที่สาม ลดอคติทางวัฒนธรรมเมื่อเทียบกับแบบทดสอบความรู้ทั่วไป โดย Lievens และคณะ (2008) พบว่าแบบวัดเชิงสถานการณ์มีความแตกต่างของคะแนนระหว่างกลุ่มน้อยกว่าแบบทดสอบความสามารถทั่วไป และประการที่สี่ เป็นที่ยอมรับของผู้สอบ เนื่องจากผู้สอบเห็นความเกี่ยวข้องระหว่างเครื่องมือกับสิ่งที่ต้องการประเมิน ทำให้ผู้สอบยอมรับและตั้งใจในการตอบมากขึ้น (Whetzel & McDaniel, 2009) นอกจากนี้ Christian, Edwards และ Bradley (2010) ได้สังเคราะห์งานวิจัยเชิงอภิมานเกี่ยวกับความตรงเชิงเกณฑ์ของแบบวัดเชิงสถานการณ์ และยืนยันว่าแบบวัดเชิงสถานการณ์มีความตรงเชิงเกณฑ์ในการทำนายผลการปฏิบัติงานในระดับที่น่าพอใจ

2.5.2 ข้อจำกัดของแบบวัดเชิงสถานการณ์

นอกเหนือจากข้อดีแล้ว แบบวัดเชิงสถานการณ์ยังมีข้อจำกัดที่ต้องคำนึงถึง ประการที่หนึ่ง การพัฒนาเครื่องมือใช้เวลาและทรัพยากรมาก โดยเฉพาะการพัฒนาสถานการณ์ที่มีคุณภาพและการกำหนดคำตอบที่ถูกต้องผ่านผู้เชี่ยวชาญ (Weekley & Ployhart, 2006) ประการที่สอง การวัดอาจไม่ครอบคลุมทักษะทั้งหมด เนื่องจากแต่ละสถานการณ์ครอบคลุมเพียงบริบทเฉพาะ การวัดทักษะให้ครอบคลุมจึงต้องอาศัยจำนวนสถานการณ์ที่เพียงพอ และประการที่สาม ความตรงเชิงโครงสร้างอาจถูกทำลาย เนื่องจากการตอบในสถานการณ์อาจสะท้อนทั้งทักษะที่ต้องการวัดและปัจจัยอื่น เช่น ความรู้ทั่วไป ประสบการณ์ส่วนตัว หรือทักษะการอ่าน (Whetzel & McDaniel, 2009) สำหรับการวิจัยครั้งนี้ ผู้วิจัยลดผลกระทบของข้อจำกัดดังกล่าว ด้วยการพัฒนาสถานการณ์

จำนวน 10 สถานการณ์เพื่อให้ครอบคลุมบริบทที่หลากหลาย และควบคุมความยาวของสถานการณ์ให้เหมาะสมกับระดับชั้นของกลุ่มเป้าหมาย เพื่อลดผลของทักษะการอ่านที่มีต่อผลการวัด

2.6 การประยุกต์ใช้แบบวัดเชิงสถานการณ์ในการวิจัยครั้งนี้

จากการทบทวนเอกสารและงานวิจัยเกี่ยวกับแบบวัดเชิงสถานการณ์ในหัวข้อ 2.1 ถึง 2.5 ผู้วิจัยเลือกใช้แนวคิดแบบวัดเชิงสถานการณ์เป็นกรอบในการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณ เนื่องจากแนวคิดดังกล่าวมุ่งประเมินการตัดสินใจและการแก้ปัญหาภายใต้บริบทของสถานการณ์ที่กำหนด ซึ่งสอดคล้องกับลักษณะของทักษะการคิดเชิงคำนวณที่เน้นการประยุกต์ใช้กระบวนการคิดในการแก้ปัญหาอย่างเป็นระบบ นอกจากนี้ แนวคิดแบบวัดเชิงสถานการณ์ยังเอื้อต่อการออกแบบข้อคำถามที่เชื่อมโยงกับบริบทที่ผู้เรียนคุ้นเคย และสามารถใช้เป็นฐานในการวัดองค์ประกอบต่าง ๆ ของทักษะการคิดเชิงคำนวณภายในสถานการณ์เดียวกันได้

ในการวิจัยครั้งนี้ ผู้วิจัยพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ โดยใช้ข้อความบรรยายสถานการณ์ที่อ้างอิงบริบทชีวิตประจำวันของนักเรียนระดับมัธยมศึกษาตอนต้น จำนวน 10 สถานการณ์ แต่ละสถานการณ์ประกอบด้วยข้อคำถามแบบปรนัย ชนิด 4 ตัวเลือก จำนวน 4 ข้อ รวมทั้งสิ้น 40 ข้อ เพื่อวัดองค์ประกอบของทักษะการคิดเชิงคำนวณ 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม โดยกำหนดให้แต่ละข้อคำถามวัดองค์ประกอบหนึ่งด้านภายใต้บริบทสถานการณ์เดียวกัน

แบบวัดใช้การให้คะแนนแบบทวิภาค โดยตอบถูกให้ 1 คะแนน และตอบผิดให้ 0 คะแนน ซึ่งสอดคล้องกับลักษณะของข้อคำถามที่มีคำตอบถูกต้องเพียงคำตอบเดียว และเอื้อต่อการวิเคราะห์ข้อมูลตามทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบที่ใช้ในการวิจัยครั้งนี้ ทั้งนี้ การออกแบบแบบวัดดังกล่าวมีพื้นฐานจากแนวคิดของแบบวัดเชิงสถานการณ์ที่มุ่งประเมินการใช้กระบวนการคิดในการแก้ปัญหาภายใต้บริบทที่กำหนด มากกว่าการวัดความรู้เชิงเนื้อหาเพียงอย่างเดียว

3. การวัดและประเมินผลในรูปแบบออนไลน์

การวัดและประเมินผลในรูปแบบออนไลน์ (online assessment) ได้รับการพัฒนาและขยายตัวอย่างต่อเนื่องในรอบสองทศวรรษที่ผ่านมา ตามความก้าวหน้าของเทคโนโลยีอินเทอร์เน็ต และการแพร่หลายของอุปกรณ์ดิจิทัล โดยเฉพาะหลังการแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 ที่ทำให้สถาบันการศึกษาทั่วโลกต้องปรับตัวมาใช้ในการประเมินออนไลน์อย่างแพร่หลาย ในหัวข้อนี้ ผู้วิจัยนำเสนอเนื้อหาที่เกี่ยวข้องกับการวัดและประเมินผลในรูปแบบออนไลน์ ประกอบด้วย ความหมายและประเภทของการวัดและประเมินผลในรูปแบบออนไลน์ ข้อได้เปรียบของการวัดและประเมินผลในรูปแบบออนไลน์ เครื่องมือและแพลตฟอร์มสำหรับการวัดออนไลน์ ระบบการรายงานผลอัตโนมัติ และการประยุกต์ใช้การวัดออนไลน์ในการวิจัยครั้งนี้

3.1 ความหมายและประเภทของการวัดและประเมินผลในรูปแบบออนไลน์

การวัดและประเมินผลในรูปแบบออนไลน์ (online assessment) หมายถึง กระบวนการวัดและประเมินผลที่ดำเนินการผ่านระบบดิจิทัลและเครือข่ายอินเทอร์เน็ต ตั้งแต่การนำเสนอข้อสอบ การบันทึกคำตอบ การประมวลผลคะแนน ไปจนถึงการรายงานผลการประเมินแก่ผู้สอบ (Alruwais, Wills, & Wald, 2018) การประเมินในรูปแบบดังกล่าวแตกต่างจากการประเมินแบบกระดาษและดินสอในด้านช่องทางการดำเนินการ การจัดเก็บข้อมูล และการประมวลผล ผลการประเมินโดยเทคโนโลยีดิจิทัลช่วยให้การจัดการข้อมูลเป็นไปอย่างสะดวกและเป็นระบบมากขึ้น

การวัดและประเมินผลในรูปแบบออนไลน์สามารถดำเนินการได้กับข้อสอบหลายลักษณะ ทั้งข้อสอบปรนัย ข้อสอบอัตนัย ข้อสอบเชิงสถานการณ์ และข้อสอบที่ใช้สื่อมัลติมีเดีย ทั้งนี้ รูปแบบของข้อสอบควรสอดคล้องกับลักษณะของสิ่งที่ต้องการวัดและวัดดูประสงค์ของการประเมิน

ในการวิจัยครั้งนี้ ผู้วิจัยพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบออนไลน์ โดยใช้ข้อสอบปรนัยเชิงสถานการณ์เป็นเครื่องมือในการเก็บรวบรวมข้อมูล ผู้สอบทำแบบวัดผ่านระบบออนไลน์ที่พัฒนาขึ้น และระบบจะบันทึกคำตอบ ประมวลผลคะแนน และรายงานผลการประเมินโดยอัตโนมัติ การออกแบบดังกล่าวสอดคล้องกับลักษณะของแบบวัดที่มุ่งประเมินการประยุกต์ใช้ทักษะการคิดเชิงคำนวณภายใต้สถานการณ์ที่กำหนด และรองรับการดำเนินการวัดกับกลุ่มตัวอย่างจำนวนมากได้อย่างเป็นระบบ

3.2 ข้อได้เปรียบของการวัดและประเมินผลในรูปแบบออนไลน์

การวัดและประเมินผลในรูปแบบออนไลน์มีลักษณะการดำเนินงานที่แตกต่างจากการประเมินแบบกระดาษและดินสอ โดยเทคโนโลยีดิจิทัลช่วยสนับสนุนการจัดการข้อมูล การประมวลผลคะแนน และการสื่อสารผลการประเมินได้อย่างเป็นระบบ (Alruwais, Wills, & Wald, 2018) ข้อได้เปรียบที่สำคัญประการหนึ่งของการประเมินในรูปแบบออนไลน์ คือ ความสามารถในการประมวลผลและรายงานผลการประเมินได้อย่างรวดเร็ว

3.2.1 การรายงานผลย้อนกลับทันที

ระบบการประเมินออนไลน์สามารถตรวจให้คะแนนและประมวลผลข้อมูลได้โดยอัตโนมัติ ทำให้ผู้สอบได้รับผลการประเมินภายหลังการทดสอบเสร็จสิ้นในระยะเวลาอันสั้น ซึ่งแตกต่างจากการประเมินแบบกระดาษและดินสอที่ต้องใช้เวลาในการตรวจและรวบรวมผลคะแนนก่อนรายงานผลแก่ผู้สอบ (Alruwais, Wills, & Wald, 2018)

ในการวิจัยครั้งนี้ แบบวัดทักษะการคิดเชิงคำนวณถูกพัฒนาบนระบบออนไลน์ที่สามารถตรวจให้คะแนนโดยอัตโนมัติ เนื่องจากข้อคำถามเป็นแบบปรนัยที่มีคำตอบถูกต้องเพียงคำตอบเดียว ภายหลังการทำแบบวัดเสร็จสิ้น ระบบจะประมวลผลคะแนนรวมและคะแนนรายองค์ประกอบของทักษะการคิดเชิงคำนวณ พร้อมจัดส่งผลการประเมินให้ผู้สอบผ่านระบบที่พัฒนาขึ้น

การดำเนินการดังกล่าวช่วยลดขั้นตอนการตรวจให้คะแนนด้วยมือและทำให้การจัดการข้อมูลการประเมินเป็นไปอย่างมีประสิทธิภาพมากขึ้น

3.2.2 การลดความคลาดเคลื่อนในการตรวจให้คะแนน

การวัดและประเมินผลในรูปแบบออนไลน์ช่วยลดความคลาดเคลื่อนที่อาจเกิดขึ้นในกระบวนการตรวจให้คะแนน เนื่องจากระบบสามารถประมวลผลคะแนนตามเกณฑ์ที่กำหนดไว้ล่วงหน้าได้โดยอัตโนมัติ (Alruwais, Wills, & Wald, 2018) แตกต่างจากการตรวจให้คะแนนโดยผู้ตรวจ ซึ่งอาจเกิดความแตกต่างในการพิจารณาคำตอบหรือการใช้เกณฑ์การให้คะแนน โดยเฉพาะเมื่อมีผู้เข้าสอบจำนวนมาก

ในการวิจัยครั้งนี้ ข้อคำถามเป็นแบบปรนัยที่มีคำตอบถูกต้องเพียงคำตอบเดียว และใช้การให้คะแนนแบบทวิภาค โดยตอบถูกให้ 1 คะแนน และตอบผิดให้ 0 คะแนน ระบบออนไลน์จึงสามารถตรวจให้คะแนนและประมวลผลข้อมูลได้โดยอัตโนมัติตามเกณฑ์เดียวกันสำหรับผู้สอบทุกคน ช่วยให้การจัดการข้อมูลการวัดเป็นไปอย่างเป็นระบบและลดขั้นตอนการตรวจให้คะแนนด้วยมือ

นอกจากนี้ ลักษณะข้อมูลแบบทวิภาคยังสอดคล้องกับการวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบสำหรับข้อมูลแบบสองค่า ซึ่งเป็นกรอบการวิเคราะห์ที่ใช้ในการตรวจสอบคุณภาพของแบบวัดในการวิจัยครั้งนี้

3.2.3 การเข้าถึงที่กว้างขวาง

การวัดและประเมินผลในรูปแบบออนไลน์เอื้อให้สามารถเก็บรวบรวมข้อมูลจากผู้ตอบที่อยู่ต่างสถานที่ได้ภายในช่วงเวลาเดียวกัน ช่วยลดข้อจำกัดด้านสถานที่และการเดินทาง ผู้ตอบสามารถเข้าร่วมการทำแบบวัดผ่านอุปกรณ์ที่เชื่อมต่ออินเทอร์เน็ตได้โดยตรง ทำให้การดำเนินการเก็บข้อมูลเป็นไปได้อย่างสะดวกและต่อเนื่อง

ข้อได้เปรียบดังกล่าวมีความเหมาะสมต่อการพัฒนาแบบวัดทางการศึกษา โดยเฉพาะการตรวจสอบคุณภาพเครื่องมือภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ ตลอดจนการสร้างเกณฑ์ปกติ ซึ่งต้องอาศัยข้อมูลจากกลุ่มตัวอย่างจำนวนมาก การใช้ระบบออนไลน์จึงเป็นแนวทางที่ช่วยสนับสนุนการรวบรวมข้อมูลในวงกว้าง และเอื้อต่อการดำเนินการวิจัยด้านการวัดและประเมินผลทางการศึกษา

3.2.4 การจัดเก็บและวิเคราะห์ข้อมูล

ระบบออนไลน์สามารถจัดเก็บข้อมูลคำตอบในรูปแบบดิจิทัลได้โดยอัตโนมัติ ทำให้ลดขั้นตอนการบันทึกข้อมูลด้วยมือ ซึ่งอาจก่อให้เกิดความคลาดเคลื่อนในการจัดเตรียมข้อมูลสำหรับการวิเคราะห์ (Alruwais และคณะ, 2018) ข้อมูลที่จัดเก็บในระบบยังสามารถส่งออกในรูปแบบไฟล์ที่รองรับการนำเข้าสู่โปรแกรมวิเคราะห์ทางสถิติได้โดยตรง ช่วยอำนวยความสะดวกต่อการจัดการข้อมูลและการวิเคราะห์ผลการประเมิน

3.2.5 การประหยัดทรัพยากร

การวัดและประเมินผลในรูปแบบออนไลน์ช่วยลดความจำเป็นในการจัดพิมพ์เอกสาร การจัดส่งแบบทดสอบ และการจัดเก็บเอกสารในรูปแบบกระดาษ ส่งผลให้ลดการใช้ทรัพยากรที่เกี่ยวข้องกับการดำเนินการประเมิน โดยเฉพาะในกรณีที่มีผู้ตอบจำนวนมากหรือมีการเก็บข้อมูลจากหลายสถานศึกษา

ข้อได้เปรียบของการวัดและประเมินผลในรูปแบบออนไลน์ทั้ง 5 ประการ ได้แก่ การรายงานผลที่รวดเร็ว การลดขั้นตอนการให้คะแนน การเข้าถึงผู้ตอบในวงกว้าง การจัดเก็บข้อมูลในรูปแบบดิจิทัล และการลดการใช้ทรัพยากร ล้วนเป็นคุณลักษณะที่สนับสนุนการพัฒนาเครื่องมือวัดทางการศึกษาในบริบทปัจจุบัน โดยเฉพาะการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ ซึ่งต้องอาศัยการเก็บรวบรวมข้อมูล การจัดการข้อมูล และการวิเคราะห์ผลจากกลุ่มตัวอย่างจำนวนมาก

3.3 เครื่องมือและแพลตฟอร์มสำหรับการวัดออนไลน์

ปัจจุบันมีเครื่องมือและแพลตฟอร์มสำหรับการวัดและประเมินผลในรูปแบบออนไลน์ที่หลากหลาย โดยแต่ละประเภทได้รับการออกแบบให้สอดคล้องกับวัตถุประสงค์การใช้งานที่แตกต่างกัน สามารถจำแนกได้เป็น 4 กลุ่มหลัก ได้แก่ ระบบจัดการการเรียนรู้ (Learning Management System: LMS) แพลตฟอร์มการสอบออนไลน์เฉพาะทาง เครื่องมือสร้างแบบสอบถามและแบบทดสอบทั่วไป และเครื่องมือการประเมินเชิงเกม

กลุ่มแรก คือ ระบบจัดการการเรียนรู้ เช่น Moodle และ Google Classroom ซึ่งเป็นแพลตฟอร์มที่สนับสนุนการบริหารจัดการการเรียนการสอนอย่างครบวงจร รวมถึงการสร้างแบบทดสอบ การจัดเก็บข้อมูลผู้เรียน และการติดตามผลการเรียนรู้ เหมาะสำหรับสถานศึกษาที่ต้องการระบบกลางในการจัดการเรียนรู้และการประเมินผล อย่างไรก็ตาม การติดตั้งและดูแลระบบบางประเภทอาจต้องอาศัยทรัพยากรด้านเทคนิคเพิ่มเติม

กลุ่มที่สอง คือ แพลตฟอร์มการสอบออนไลน์เฉพาะทาง ซึ่งได้รับการออกแบบเพื่อรองรับการจัดสอบในวงกว้างและการควบคุมสภาพแวดล้อมการสอบ มีการนำระบบรักษาความปลอดภัยและการกำกับดูแลการสอบมาใช้ในระดับที่แตกต่างกันตามวัตถุประสงค์ของการประเมิน จึงมักถูกนำไปใช้ในการสอบที่มีความสำคัญต่อการตัดสินผลหรือการคัดเลือก

กลุ่มที่สาม คือ เครื่องมือสร้างแบบสอบถามและแบบทดสอบทั่วไป เช่น Google Forms และ Microsoft Forms ซึ่งมีจุดเด่นด้านความสะดวกในการใช้งาน การเข้าถึงได้ง่าย และการจัดเก็บข้อมูลในรูปแบบดิจิทัลโดยอัตโนมัติ จึงได้รับความนิยมในการเก็บรวบรวมข้อมูลสำหรับการเรียนการสอนและการวิจัยทางการศึกษา

กลุ่มสุดท้าย คือ เครื่องมือการประเมินเชิงเกม เช่น Kahoot และ Quizizz ซึ่งมุ่งเน้นการสร้างแรงจูงใจและการมีส่วนร่วมของผู้เรียนผ่านกลไกของเกม เครื่องมือประเภทนี้เหมาะสำหรับการประเมินเพื่อส่งเสริมการเรียนรู้ระหว่างเรียนและการให้ข้อมูลย้อนกลับอย่างรวดเร็ว โดยมีวัตถุประสงค์การใช้งานแตกต่างจากการพัฒนาเครื่องมือวัดที่ต้องมีการตรวจสอบคุณภาพทางจิตมิติอย่างเป็นระบบ

จากเครื่องมือและแพลตฟอร์มสำหรับการวัดออนไลน์ที่มีอยู่หลากหลายประเภท ผู้วิจัยเลือกใช้ Google Forms ร่วมกับ Google Sites ในการพัฒนาแบบวัดครั้งนี้ เนื่องจากมีคุณลักษณะที่เอื้อต่อการสร้างข้อคำถาม การจัดเก็บข้อมูล การบริหารจัดการแบบวัด และการรายงานผลผ่านระบบออนไลน์

ประการแรก Google Forms เป็นเครื่องมือที่สามารถใช้งานได้โดยไม่มีค่าใช้จ่ายด้านลิขสิทธิ์ และรองรับการเผยแพร่แบบวัดผ่านระบบอินเทอร์เน็ต ทำให้สะดวกต่อการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างจำนวนมาก

ประการที่สอง Google Forms มีรูปแบบการใช้งานที่ไม่ซับซ้อน สามารถเข้าถึงได้ผ่านอุปกรณ์หลากหลายประเภท ทั้งคอมพิวเตอร์ แท็บเล็ต และสมาร์ทโฟน จึงเอื้อต่อการเข้าร่วมตอบแบบวัดของนักเรียนในบริบทการศึกษาขั้นพื้นฐาน

ประการที่สาม ระบบรองรับการพัฒนาเพิ่มเติมผ่าน Google Apps Script ซึ่งช่วยให้สามารถออกแบบกระบวนการจัดการข้อมูลและการรายงานผลให้สอดคล้องกับวัตถุประสงค์ของการวิจัย โดยเฉพาะการจัดส่งผลการประเมินในรูปแบบอัตโนมัติภายหลังการตอบแบบวัด

ประการสุดท้าย Google Sites เป็นเครื่องมือสำหรับพัฒนาเว็บไซต์ที่สามารถเชื่อมโยงกับบริการอื่นภายในระบบของกูเกิลได้อย่างสะดวก ผู้วิจัยจึงใช้เป็นพื้นที่สำหรับนำเสนอข้อมูลเกี่ยวกับโครงการวิจัย คำชี้แจงสำหรับผู้ตอบแบบวัด วิธีการเข้าถึงแบบวัด และช่องทางการติดต่อ ซึ่งช่วยสนับสนุนการบริหารจัดการแบบวัดและการสื่อสารข้อมูลกับผู้เกี่ยวข้องได้อย่างเป็นระบบ

ด้วยคุณลักษณะดังกล่าว Google Forms และ Google Sites จึงเป็นเครื่องมือที่เหมาะสมสำหรับการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ ซึ่งต้องรองรับการเก็บรวบรวมข้อมูล การจัดการข้อมูล และการรายงานผลผ่านระบบดิจิทัลตลอดกระบวนการวิจัย

3.4 ระบบการรายงานผลอัตโนมัติในการวัดออนไลน์

ระบบการรายงานผลอัตโนมัติเป็นองค์ประกอบสำคัญของการวัดและประเมินผลในรูปแบบออนไลน์ เนื่องจากช่วยให้ผู้ตอบได้รับผลการประเมินภายหลังการทำแบบวัดโดยไม่ต้องผ่านกระบวนการตรวจและรายงานผลแบบใช้เวลา ระบบดังกล่าวจึงเป็นกลไกสำคัญที่สนับสนุนการให้ข้อมูลย้อนกลับแก่ผู้ตอบอย่างรวดเร็ว

Hattie และ Timperley (2007) เสนอว่าการให้ข้อมูลย้อนกลับที่มีประสิทธิภาพควรช่วยตอบคำถามสำคัญ 3 ประการ ได้แก่ ผู้เรียนกำลังมุ่งไปสู่เป้าหมายใด ผู้เรียนอยู่ในระดับใดเมื่อเทียบกับเป้าหมายนั้น และผู้เรียนควรดำเนินการอย่างไรต่อไป ข้อมูลย้อนกลับที่ครอบคลุมทั้งสามประเด็นจะช่วยให้ผู้เรียนเข้าใจผลการประเมินของตนเองได้ชัดเจนยิ่งขึ้น ขณะที่ Black และ Wiliam (2009) เสนอว่ารายงานผลการประเมินควรนำเสนอทั้งผลการประเมินโดยรวมและผลการประเมินรายองค์ประกอบ เพื่อให้ผู้เรียนสามารถพิจารณาผลการประเมินได้อย่างละเอียดมากขึ้น

ในการวิจัยครั้งนี้ ผู้วิจัยได้พัฒนาระบบรายงานผลอัตโนมัติที่จัดส่งผลการประเมินผ่านทางอีเมลภายหลังผู้ตอบทำแบบวัดเสร็จสิ้น โดยออกแบบให้สอดคล้องกับแนวคิดการให้ข้อมูลย้อนกลับดังกล่าว รายงานผลประกอบด้วยคะแนนรวมของแบบวัด คะแนนรายองค์ประกอบทั้ง 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม ตลอดจนการแปลผลระดับทักษะพร้อมคำอธิบายที่เหมาะสมกับนักเรียนระดับมัธยมศึกษาตอนต้น ข้อมูลเหล่านี้ช่วยให้ผู้ตอบสามารถพิจารณาผลการประเมินของตนเองทั้งในภาพรวมและรายองค์ประกอบได้อย่างสะดวก

นอกจากนี้ รายงานผลที่แสดงคะแนนรายองค์ประกอบยังเป็นข้อมูลสารสนเทศที่สามารถนำไปใช้ประกอบการพิจารณาผลการประเมินของผู้เรียนในแต่ละด้านได้ ทั้งในระดับผู้เรียนและผู้สอน อันเป็นประโยชน์ต่อการทำความเข้าใจลักษณะผลการประเมินที่ปรากฏจากแบบวัด

3.5 การประยุกต์ใช้การวัดออนไลน์ในการวิจัยครั้งนี้

จากการทบทวนเอกสารและงานวิจัยเกี่ยวกับการวัดและประเมินผลในรูปแบบออนไลน์ในหัวข้อ 3.1 ถึง 3.4 ผู้วิจัยสรุปได้ว่าการวัดออนไลน์มีคุณลักษณะที่สอดคล้องกับวัตถุประสงค์และกระบวนการดำเนินการวิจัยครั้งนี้หลายประการ

ประการแรก การวัดออนไลน์เอื้อต่อการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างจำนวนมาก และสามารถบริหารจัดการข้อมูลได้อย่างเป็นระบบ ซึ่งมีความสำคัญต่อการตรวจสอบคุณภาพของแบบวัดและการสร้างเกณฑ์ปกติ ประการที่สอง ระบบตรวจให้คะแนนอัตโนมัติมีความสอดคล้องกับลักษณะการให้คะแนนแบบทวิภาคที่ใช้ในการวิจัยครั้งนี้ และช่วยลดขั้นตอนการประมวลผลคะแนนภายหลังการเก็บข้อมูล ประการที่สาม ระบบรายงานผลอัตโนมัติสามารถส่งผลการประเมินกลับไปยังผู้ตอบได้ภายหลังการทำแบบวัดเสร็จสิ้น ซึ่งสอดคล้องกับแนวคิดการให้ข้อมูลย้อนกลับที่กล่าวถึงในวรรณกรรมด้านการประเมินผล และประการสุดท้าย การใช้เครื่องมือที่สามารถเข้าถึงได้ผ่านระบบอินเทอร์เน็ตและไม่ต้องติดตั้งโครงสร้างพื้นฐานที่ซับซ้อน ช่วยสนับสนุนการพัฒนาและการบริหารจัดการแบบวัดในรูปแบบออนไลน์ได้อย่างสะดวก

จากเหตุผลดังกล่าว ผู้วิจัยจึงพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ในการวิจัยครั้งนี้ โดยใช้ Google Forms เป็นแพลตฟอร์มหลักสำหรับ

การนำเสนอข้อคำถามและเก็บรวบรวมข้อมูล ใช้ Google Sites เป็นช่องทางในการเผยแพร่ข้อมูล และคำชี้แจงเกี่ยวกับแบบวัด และใช้ Google Apps Script ในการพัฒนาระบบรายงานผลอัตโนมัติทางอีเมล ซึ่งรายงานคะแนนรวม คะแนนรายองค์ประกอบทั้ง 4 ด้าน และผลการแปลระดับทักษะ ภายหลังจากทำแบบวัดเสร็จสิ้น

4. กรอบแนวคิดทางจิตมิติสำหรับการพัฒนาแบบวัด

การพัฒนาแบบวัดที่มีคุณภาพต้องอาศัยกรอบแนวคิดทางจิตมิติ เพื่อตรวจสอบคุณสมบัติของเครื่องมือว่าสามารถวัดสิ่งที่ต้องการวัดได้อย่างถูกต้อง แม่นยำ และเชื่อถือได้ การใช้กรอบทฤษฎีทางจิตมิติเป็นสิ่งจำเป็นในการพัฒนาเครื่องมือวัดทางการศึกษา เพราะช่วยให้ผู้วิจัยมีหลักฐานเชิงประจักษ์ในการยืนยันคุณภาพของแบบวัด แทนการตัดสินจากความรู้สึกหรือประสบการณ์ส่วนตัว ในวงการวัดและประเมินผลทางการศึกษา มีกรอบทฤษฎีหลัก 2 กรอบที่ใช้ในการตรวจสอบคุณภาพแบบวัด ได้แก่ ทฤษฎีการทดสอบแบบดั้งเดิม (classical test theory: CTT) ที่เป็นกรอบพื้นฐานในการวิเคราะห์ข้อสอบ และทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) ที่เป็นกรอบทฤษฎีสมัยใหม่ที่ให้ข้อมูลคุณภาพข้อสอบที่ละเอียดและไม่ขึ้นกับกลุ่มตัวอย่าง นอกจากนี้ การพัฒนาแบบวัดที่สมบูรณ์ยังจำเป็นต้องสร้างเกณฑ์ปกติ เพื่อให้ผู้ใช้สามารถตีความคะแนนของผู้สอบในบริบทเปรียบเทียบกับกลุ่มอ้างอิงได้อย่างมีความหมาย ในหัวข้อนี้ ผู้วิจัยนำเสนอเนื้อหาเกี่ยวกับกรอบแนวคิดทางจิตมิติทั้งสามด้าน ได้แก่ ทฤษฎีการทดสอบแบบดั้งเดิม ทฤษฎีการตอบสนองข้อสอบ การสร้างเกณฑ์ปกติ และการประยุกต์ใช้กรอบแนวคิดทางจิตมิติในการวิจัยครั้งนี้

4.1 ทฤษฎีการทดสอบแบบดั้งเดิม

4.1.1 ความหมายและพื้นฐานทางทฤษฎี

ทฤษฎีการทดสอบแบบดั้งเดิมเป็นกรอบทฤษฎีพื้นฐานในการวิเคราะห์คุณภาพของแบบทดสอบและข้อสอบ ที่พัฒนามาตั้งแต่ต้นศตวรรษที่ 20 มีรากฐานจากแนวคิดเรื่องความคลาดเคลื่อนในการวัดที่ Spearman (1904) เสนอเป็นครั้งแรก ทฤษฎีการทดสอบแบบดั้งเดิมตั้งอยู่บนสมมติฐานหลักที่เรียกว่าโมเดลคะแนนจริง ซึ่งกำหนดว่าคะแนนที่ผู้สอบได้รับจากการทดสอบประกอบด้วย 2 ส่วน ตามสมการ $X = T + E$ เมื่อ X คือคะแนนที่สังเกตได้ T คือคะแนนจริงที่สะท้อนความสามารถที่แท้จริงของผู้สอบ และ E คือคะแนนความคลาดเคลื่อนที่เกิดจากปัจจัยสุ่มที่ไม่เกี่ยวข้องกับความสามารถ เช่น สภาพร่างกายของผู้สอบในวันสอบ สภาพแวดล้อมในการสอบ หรือการเดาคำตอบ

จากสมการพื้นฐานนี้ ทฤษฎีการทดสอบแบบดั้งเดิมได้พัฒนาดัชนีและวิธีการต่าง ๆ ในการประเมินคุณภาพของแบบทดสอบและข้อสอบรายข้อ ทั้งในด้านความตรงและความเชื่อมั่น ทฤษฎีการทดสอบแบบดั้งเดิมเป็นกรอบที่นิยมใช้ในการตรวจสอบคุณภาพแบบวัดเบื้องต้น เนื่องจาก

เข้าใจง่าย คำนวณด้วยสถิติพื้นฐานได้ ต้องการกลุ่มตัวอย่างขนาดใหญ่มาก และมีโปรแกรมสถิติรองรับอย่างแพร่หลาย อย่างไรก็ตาม ทฤษฎีการทดสอบแบบดั้งเดิมมีข้อจำกัดสำคัญ Hambleton, Swaminathan และ Rogers (1991) ระบุข้อจำกัด 2 ประการ ประการแรก ค่าดัชนีคุณภาพของข้อสอบที่ได้จากทฤษฎีนี้ขึ้นอยู่กับกลุ่มตัวอย่างที่ใช้ในการวิเคราะห์ ทำให้ค่าที่ได้จากกลุ่มหนึ่งอาจไม่สามารถนำไปใช้กับอีกกลุ่มหนึ่งได้โดยตรง และประการที่สอง คะแนนของผู้สอบขึ้นอยู่กับชุดข้อสอบที่ใช้ ทำให้ไม่สามารถเปรียบเทียบคะแนนของผู้สอบที่ทำชุดข้อสอบต่างกันได้โดยตรง ข้อจำกัดเหล่านี้นำไปสู่การพัฒนาทฤษฎีการตอบสนองข้อสอบ ในการวิจัยครั้งนี้ ผู้วิจัยใช้กรอบทฤษฎีการทดสอบแบบดั้งเดิมเป็นกรอบเริ่มต้นในการตรวจสอบคุณภาพแบบวัด เพื่อให้ได้ข้อมูลพื้นฐานเกี่ยวกับคุณภาพของข้อสอบรายข้อและของแบบวัดในภาพรวม ก่อนนำไปวิเคราะห์ต่อในกรอบทฤษฎีการตอบสนองข้อสอบ

4.1.2 ความตรงเชิงเนื้อหาและดัชนีความสอดคล้อง

ความตรงเชิงเนื้อหา หมายถึง ระดับที่เนื้อหาของแบบทดสอบครอบคลุมและเป็นตัวแทนของขอบเขตเนื้อหาที่ต้องการวัด การตรวจสอบความตรงเชิงเนื้อหาเป็นขั้นตอนสำคัญในการพัฒนาแบบวัด เนื่องจากเครื่องมือที่ไม่ครอบคลุมเนื้อหาที่ต้องการวัด แม้จะมีคุณสมบัติทางสถิติที่ดีก็ไม่สามารถวัดสิ่งที่ต้องการได้อย่างแท้จริง วิธีการตรวจสอบความตรงเชิงเนื้อหาที่นิยมใช้ในงานวิจัยทางการศึกษา คือการให้ผู้เชี่ยวชาญในสาขาประเมินความสอดคล้องระหว่างข้อสอบกับวัตถุประสงค์การวัด แล้วคำนวณค่าดัชนีความสอดคล้อง (index of item-objective congruence: IOC) ที่เสนอโดย Rovinelli และ Hambleton (1977)

ในการคำนวณค่าดัชนีความสอดคล้อง ผู้เชี่ยวชาญแต่ละคนจะให้คะแนนข้อสอบในระดับ +1 เมื่อข้อสอบสอดคล้องกับวัตถุประสงค์การวัด 0 เมื่อไม่แน่ใจ และ -1 เมื่อไม่สอดคล้อง แล้วคำนวณค่าดัชนีจากผลรวมของคะแนนหารด้วยจำนวนผู้เชี่ยวชาญ ค่าดัชนีความสอดคล้องมีช่วงตั้งแต่ -1.00 ถึง +1.00 โดยค่าที่เข้าใกล้ +1.00 บ่งบอกว่าผู้เชี่ยวชาญส่วนใหญ่เห็นว่าข้อสอบสอดคล้องกับวัตถุประสงค์การวัด เกณฑ์การยอมรับค่าดัชนีความสอดคล้องในงานวิจัยทางการศึกษา คือ ตั้งแต่ 0.50 ขึ้นไป ข้อสอบที่มีค่าต่ำกว่า 0.50 ควรได้รับการปรับปรุงหรือคัดออก เกณฑ์การแปลผลค่าดัชนีความสอดคล้องแสดงในตาราง 2

ตาราง 2 เกณฑ์การแปลผลค่าดัชนีความสอดคล้อง (IOC)

ช่วงค่าดัชนีความสอดคล้อง	ความหมาย	ข้อปฏิบัติ
0.67–1.00	สอดคล้องดี	ยอมรับ นำไปใช้ได้
0.50–0.66	สอดคล้องพอใช้	ยอมรับ แต่ควรปรับปรุงตามข้อเสนอแนะ

ตาราง 2 (ต่อ)

ช่วงค่าดัชนีความสอดคล้อง	ความหมาย	ข้อปฏิบัติ
น้อยกว่า 0.50	ไม่สอดคล้อง	ปรับปรุงหรือตัดออก
น้อยกว่า 0.00 (ติดลบ)	ขัดแย้งกับวัตถุประสงค์	ตัดออก

หมายเหตุ: ดัดแปลงจาก Rovinelli และ Hambleton (1977)

จำนวนผู้เชี่ยวชาญที่เหมาะสมในการตรวจสอบความตรงเชิงเนื้อหาด้วยดัชนีความสอดคล้อง (Index of Item-Objective Congruence: IOC) โดยทั่วไปควรมีตั้งแต่ 3 คนขึ้นไป เพื่อให้การพิจารณามีความรอบคอบและช่วยลดอคติจากความคิดเห็นส่วนบุคคล (Rovinelli & Hambleton, 1977) ผู้เชี่ยวชาญจะทำหน้าที่ประเมินความสอดคล้องระหว่างข้อคำถามกับนิยามหรือวัตถุประสงค์ของสิ่งที่ต้องการวัด จากนั้นนำผลการประเมินมาคำนวณเป็นค่าดัชนีความสอดคล้องโดยข้อคำถามที่มีค่า IOC ตั้งแต่ 0.50 ขึ้นไปถือว่ามีความสอดคล้องกับสิ่งที่ต้องการวัดในระดับที่ยอมรับได้

4.1.3 ค่าความยากของข้อสอบ

ค่าความยากของข้อสอบ สัญลักษณ์ p เป็นค่าที่บ่งบอกความยากของข้อสอบแต่ละข้อ สำหรับข้อสอบที่ให้คะแนนแบบทวิภาค ค่า p คือสัดส่วนของผู้สอบที่ตอบข้อสอบนั้นถูก ค่า p มีค่าระหว่าง 0.00 ถึง 1.00 โดยค่าที่เข้าใกล้ 1.00 หมายความว่าข้อสอบนั้นง่าย เพราะผู้สอบส่วนใหญ่ตอบถูก และค่าที่เข้าใกล้ 0.00 หมายความว่าข้อสอบนั้นยาก เพราะผู้สอบส่วนใหญ่ตอบผิด ในการวิจัยครั้งนี้ ผู้วิจัยคำนวณค่าความยากด้วยวิธีกลุ่มสูงและกลุ่มต่ำ ร้อยละ 27 ซึ่ง Kelley (1939) เสนอว่าให้ประสิทธิภาพสูงสุดในการแยกแยะผู้สอบ

เกณฑ์การแปลผลค่าความยากของข้อสอบที่นิยมใช้ในงานวิจัยทางการศึกษา คือ ข้อสอบที่ดีควรมีค่า p อยู่ในช่วง 0.20 ถึง 0.80 โดยข้อสอบที่มีค่า p น้อยกว่า 0.20 ถือว่ายากเกินไป เนื่องจากมีผู้ตอบถูกจำนวนน้อย ขณะที่ข้อสอบที่มีค่า p มากกว่า 0.80 ถือว่าง่ายเกินไป เพราะมีผู้ตอบถูกเป็นจำนวนมาก ส่วนข้อสอบที่มีค่า p อยู่ระหว่าง 0.20 ถึง 0.80 ถือว่ามีระดับความยากที่เหมาะสมและสามารถนำไปใช้ในการวัดได้อย่างมีประสิทธิภาพ

4.1.4 ค่าอำนาจจำแนกของข้อสอบ

ค่าอำนาจจำแนกของข้อสอบ สัญลักษณ์ r เป็นค่าที่บ่งบอกความสามารถของข้อสอบในการจำแนกผู้สอบที่มีความสามารถสูงออกจากผู้สอบที่มีความสามารถต่ำ ข้อสอบที่ดีต้องสามารถจำแนกผู้สอบได้ กล่าวคือ ผู้สอบที่มีความสามารถสูงควรตอบข้อสอบถูกในสัดส่วนที่สูงกว่า

ผู้สอบที่มีความสามารถต่ำ ค่าอำนาจจำแนกที่นิยมใช้คือสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของข้อสอบกับคะแนนรวมของแบบทดสอบ ค่า r มีค่าระหว่าง -1.00 ถึง +1.00 โดยค่าที่เข้าใกล้ +1.00 บ่งบอกว่าข้อสอบมีอำนาจจำแนกดี ค่าที่เข้าใกล้ 0 บ่งบอกว่าข้อสอบไม่มีอำนาจจำแนก และค่าติดลบ บ่งชี้ว่าข้อสอบมีปัญหา เนื่องจากผู้สอบที่มีความสามารถต่ำตอบถูกมากกว่าผู้สอบที่มีความสามารถสูง ซึ่งอาจเกิดจากการที่ข้อสอบมีกฎเกณฑ์คำตอบผิด หรือมีคำถามที่กำกวม เกณฑ์การแปลผลค่าอำนาจจำแนกตาม Ebel และ Frisbie (1991) แสดงในตาราง 3

ตาราง 3 เกณฑ์การแปลผลค่าอำนาจจำแนก (r)

ช่วงค่า r	ระดับคุณภาพ	ข้อปฏิบัติ
0.40 ขึ้นไป	ดีมาก	ยอมรับและใช้ได้ทันที
0.30–0.39	ดี	ยอมรับและอาจปรับปรุงเพิ่มเติม
0.20–0.29	พอใช้ได้	ควรปรับปรุงก่อนนำไปใช้
น้อยกว่า 0.20	ไม่ดี	ตัดออกหรือปรับปรุงอย่างมาก
น้อยกว่า 0 (ติดลบ)	ใช้ไม่ได้	ตัดออกทันที

หมายเหตุ: ดัดแปลงจาก Ebel และ Frisbie (1991)

4.1.5 ค่าความเชื่อมั่นของแบบวัด

ความเชื่อมั่น หมายถึง ระดับความคงที่หรือความสม่ำเสมอของผลการวัด เครื่องมือที่มีความเชื่อมั่นสูงจะให้ผลการวัดที่ใกล้เคียงกันเมื่อนำไปใช้ซ้ำกับผู้สอบกลุ่มเดียวกัน หรือเมื่อใช้แบบทดสอบที่เทียบเคียงกัน ความเชื่อมั่นจึงเป็นคุณสมบัติสำคัญที่บ่งบอกความน่าเชื่อถือของผลการวัด ค่าความเชื่อมั่นมีค่าระหว่าง 0.00 ถึง 1.00 โดยค่าที่เข้าใกล้ 1.00 หมายถึงแบบทดสอบมีความเชื่อมั่นสูง

วิธีการคำนวณค่าความเชื่อมั่นมีหลายวิธี ขึ้นอยู่กับลักษณะของข้อมูลและรูปแบบของแบบทดสอบ สำหรับข้อสอบที่ให้คะแนนแบบทวิภาค วิธีที่นิยมใช้คือสูตร KR-20 ที่พัฒนาโดย Kuder และ Richardson (1937) ซึ่งเหมาะสำหรับการประมาณค่าความสอดคล้องภายในของแบบทดสอบที่ให้คะแนนถูก-ผิด เกณฑ์การแปลผลค่า KR-20 แสดงในตาราง 4

ตาราง 4 เกณฑ์การแปลผลค่าความเชื่อมั่น KR-20

ช่วงค่า KR-20	ระดับคุณภาพ	ความเหมาะสมในการใช้งาน
0.90 ขึ้นไป	สูงมาก	เหมาะสำหรับการตัดสินใจเกี่ยวกับบุคคลรายคน
0.80–0.89	สูง	เหมาะสำหรับแบบวัดวิจัยและการตัดสินใจระดับปานกลาง
0.70–0.79	ยอมรับได้	เหมาะสำหรับการวัดกลุ่ม ไม่ควรใช้ตัดสินใจรายบุคคล
0.60–0.69	น่าสงสัย	ควรปรับปรุงแบบวัด
น้อยกว่า 0.60	ไม่ยอมรับ	ต้องปรับปรุงแบบวัดอย่างจริงจัง

หมายเหตุ: ดัดแปลงจาก Nunnally และ Bernstein (1994)

ปัจจัยที่ส่งผลต่อค่าความเชื่อมั่น ได้แก่ จำนวนข้อสอบ ความสอดคล้องของเนื้อหาที่วัด และความหลากหลายของกลุ่มผู้สอบ โดยทั่วไป แบบทดสอบที่มีจำนวนข้อเพียงพอ วัดคุณลักษณะเดียวกันอย่างสม่ำเสมอ และใช้กับกลุ่มผู้สอบที่มีความแตกต่างของความสามารถ จะมีแนวโน้มให้ค่าความเชื่อมั่นสูงกว่า ทั้งนี้ Nunnally และ Bernstein (1994) เสนอว่าค่าความเชื่อมั่นตั้งแต่ 0.70 ขึ้นไปถือว่าอยู่ในระดับที่ยอมรับได้สำหรับการวิจัยทางการศึกษา

4.2 ทฤษฎีการตอบสนองข้อสอบ

4.2.1 ความหมายและพื้นฐานทางทฤษฎี

ทฤษฎีการตอบสนองข้อสอบเป็นกรอบทฤษฎีสมัยใหม่ในวงการจิตมิติที่พัฒนาขึ้นเพื่อแก้ข้อจำกัดของทฤษฎีการทดสอบแบบดั้งเดิม โดยเฉพาะข้อจำกัดที่ดัชนีคุณภาพข้อสอบและคะแนนของผู้สอบขึ้นอยู่กับกลุ่มตัวอย่างและชุดข้อสอบ ทฤษฎีนี้ได้รับการพัฒนาตั้งแต่ทศวรรษ 1960 โดยมี Lord (1980) และ Rasch (1960) เป็นนักวิจัยที่มีบทบาทสำคัญในการวางรากฐาน Embretson และ Reise (2000) อธิบายว่าทฤษฎีการตอบสนองข้อสอบเป็นกรอบทฤษฎีที่อธิบายความสัมพันธ์ระหว่างความสามารถของผู้สอบ ซึ่งใช้สัญลักษณ์ θ (อ่านว่า ทีตา) กับความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบถูก ผ่านฟังก์ชันทางคณิตศาสตร์ที่เรียกว่าฟังก์ชันการตอบสนองข้อสอบ หรือเส้นโค้งลักษณะของข้อสอบ ซึ่งโดยทั่วไปมีลักษณะเป็นเส้นโค้งรูปตัว S

ค่าความสามารถของผู้สอบ θ ในทฤษฎีการตอบสนองข้อสอบมีค่าอยู่บนสเกลที่ต่อเนื่อง โดยทั่วไปอยู่ในช่วง -3 ถึง +3 ค่าที่เข้าใกล้ +3 หมายความว่าผู้สอบมีความสามารถสูง ค่าที่

เข้าใกล้ -3 หมายความว่าผู้สอบมีความสามารถต่ำ และค่า θ ที่เท่ากับ 0 แทนความสามารถระดับปานกลางในกลุ่ม จุดเด่นสำคัญของทฤษฎีการตอบสนองข้อสอบที่แตกต่างจากทฤษฎีการทดสอบแบบดั้งเดิม คือ คุณสมบัติความคงที่ของพารามิเตอร์ (parameter invariance) Hambleton, Swaminathan และ Rogers (1991) อธิบายว่า ภายใต้เงื่อนไขที่โมเดลมีความเหมาะสม ค่าพารามิเตอร์ของข้อสอบจะมีความคงที่เมื่อประมาณค่าจากกลุ่มผู้สอบที่แตกต่างกัน และความสามารถของผู้สอบจะมีความคงที่เมื่อประมาณค่าจากชุดข้อสอบที่วัดคุณลักษณะเดียวกัน คุณสมบัติดังกล่าวเป็นข้อได้เปรียบสำคัญที่ทำให้ทฤษฎีการตอบสนองข้อสอบได้รับความนิยมในการพัฒนาค้างข้อสอบและการสร้างแบบวัดมาตรฐาน

นอกจากนี้ ทฤษฎีการตอบสนองข้อสอบยังสามารถให้ข้อมูลคุณภาพของข้อสอบในระดับรายข้อผ่านค่าพารามิเตอร์ต่าง ๆ เช่น ค่าอำนาจจำแนกและค่าความยาก รวมทั้งให้ข้อมูลเกี่ยวกับความแม่นยำของการประมาณค่าความสามารถของผู้สอบในแต่ละระดับผ่านฟังก์ชันสารสนเทศของข้อสอบและฟังก์ชันสารสนเทศของแบบวัด ซึ่งเป็นข้อมูลที่ทฤษฎีการทดสอบแบบดั้งเดิมไม่สามารถให้ได้โดยตรง อย่างไรก็ตาม การวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบจะให้ผลที่มีความน่าเชื่อถือเมื่อข้อมูลเป็นไปตามข้อตกลงเบื้องต้นที่สำคัญ ได้แก่ ความเป็นเอกมิติ (unidimensionality) และความเป็นอิสระเฉพาะที่ (local independence) ดังนั้น ก่อนการประมาณค่าพารามิเตอร์ข้อสอบและค่าความสามารถของผู้สอบ จึงจำเป็นต้องตรวจสอบความสอดคล้องของข้อมูลกับข้อตกลงเบื้องต้นดังกล่าวเสียก่อน

4.2.2 ข้อตกลงเบื้องต้นด้านความเป็นเอกมิติ

ความเป็นเอกมิติ หมายถึง สถานการณ์ที่ข้อสอบทุกข้อในแบบวัดวัดความสามารถแฝงเพียงมิติเดียวเป็นหลัก (Hambleton และคณะ, 1991) การละเมิดข้อตกลงนี้จะทำให้ผลการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบไม่น่าเชื่อถือ เนื่องจากค่าพารามิเตอร์ที่ประมาณได้อาจสะท้อนหลายมิติของความสามารถผสมรวมกัน ทำให้ตีความได้ยาก อย่างไรก็ตาม Reckase (1979) ระบุว่าในทางปฏิบัติแทบไม่มีแบบวัดใดที่มีความเป็นเอกมิติสมบูรณ์ สิ่งที่สำคัญกว่าคือการที่มีปัจจัยหลักเพียงปัจจัยเดียวที่โดดเด่น

การตรวจสอบความเป็นเอกมิติสามารถดำเนินการได้หลายวิธี วิธีหนึ่งที่ได้รับ การอ้างอิงอย่างแพร่หลายคือการพิจารณาอัตราส่วนค่าไอเกนระหว่างองค์ประกอบแรกกับองค์ประกอบที่สอง (λ_1/λ_2) โดย Reckase (1979) เสนอว่าอัตราส่วนที่มีค่าประมาณ 4 ต่อ 1 หรือสูงกว่า แสดงให้เห็นว่าองค์ประกอบแรกมีอิทธิพลเด่นเหนือองค์ประกอบอื่น และสนับสนุนสมมติฐานความเป็นเอกมิติ นอกจากนี้ นักวิชาการบางส่วนเสนอว่าอัตราส่วนตั้งแต่ 3 ต่อ 1 ขึ้นไปสามารถใช้เป็นหลักฐานเบื้องต้นของความเป็นเอกมิติได้ โดยควรพิจารณาร่วมกับหลักฐานอื่นประกอบ

ในปัจจุบัน วิธีที่ได้รับความนิยมมากขึ้นคือการใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) ด้วยโมเดลองค์ประกอบเดียว ซึ่งสามารถทดสอบสมมติฐานเกี่ยวกับโครงสร้างของแบบวัดได้โดยตรง และให้ดัชนีความสอดคล้องของโมเดลหลายประเภท ดัชนีหลักที่ใช้ในการตรวจสอบความสอดคล้องของโมเดลแสดงในตาราง 5

ตาราง 5 ดัชนีความสอดคล้องของโมเดลและเกณฑ์การยอมรับ

ดัชนี	เกณฑ์การยอมรับ	ความหมาย
χ^2/df	น้อยกว่า 3.00	อัตราส่วนไคสแควร์ต่อองศาอิสระ
RMSEA	น้อยกว่า 0.06 (ดี), ไม่เกิน 0.08 (ยอมรับได้)	ความคลาดเคลื่อนของโมเดลต่อองศาอิสระ
CFI	มากกว่า 0.95 (ดี), มากกว่า 0.90 (ยอมรับได้)	เปรียบเทียบโมเดลกับโมเดลฐาน
TLI	มากกว่า 0.95 (ดี), มากกว่า 0.90 (ยอมรับได้)	คล้าย CFI แต่ปรับตามจำนวนพารามิเตอร์
SRMR	น้อยกว่า 0.08	ค่าเฉลี่ยของส่วนเหลือมาตรฐาน

หมายเหตุ: เกณฑ์ตาม Hu และ Bentler (1999)

4.2.3 ข้อตกลงเบื้องต้นด้านความเป็นอิสระเฉพาะที่

ความเป็นอิสระเฉพาะที่ (local independence) หมายถึง สถานการณ์ที่เมื่อควบคุมความสามารถของผู้สอบแล้ว การตอบข้อสอบแต่ละข้อจะต้องไม่ขึ้นอยู่กับคำตอบข้ออื่น (Hambleton และคณะ, 1991) การละเมิดข้อตกลงนี้เรียกว่า การพึ่งพิงเฉพาะที่ (local dependence) ซึ่งเกิดขึ้นเมื่อข้อสอบบางคู่มีความสัมพันธ์กันเกินกว่าที่ความสามารถแฝงจะอธิบายได้ สาเหตุหนึ่งที่ทำให้พบได้บ่อยคือการที่ข้อสอบหลายข้อใช้สถานการณ์หรือสิ่งเร้าร่วมกัน การมีการพึ่งพิงเฉพาะที่อาจส่งผลให้การประมาณค่าพารามิเตอร์ข้อสอบและค่าความสามารถของผู้สอบเกิดความคลาดเคลื่อน รวมทั้งทำให้ค่าความเชื่อมั่นของแบบวัดสูงกว่าความเป็นจริง

วิธีการตรวจสอบความเป็นอิสระเฉพาะที่ที่นิยมใช้ คือ สถิติ Q3 ที่เสนอโดย Yen (1984) ซึ่งคำนวณจากสหสัมพันธ์ของค่าคงเหลือระหว่างคู่ข้อสอบภายหลังการปรับโมเดลการ

ตอบสนองข้อสอบ เกณฑ์ที่นิยมใช้ในการแปลผลคือ ค่าสัมบูรณ์ของ Q3 ที่น้อยกว่า 0.20 แสดงว่า ข้อสอบมีความเป็นอิสระเฉพาะที่อยู่ในระดับที่ยอมรับได้ (Christensen, Makransky, & Horton, 2017) เกณฑ์การแปลผลค่า Q3 แสดงในตาราง 6

ตาราง 6 เกณฑ์การแปลผลค่า Q3 สำหรับการตรวจสอบความเป็นอิสระเฉพาะที่

ช่วงค่า Q3	การแปลผล	ข้อปฏิบัติ
น้อยกว่า 0.00	ไม่มีการพึ่งพิงเฉพาะที่	วิเคราะห์ต่อได้
0.00 ถึง 0.20	พึ่งพิงเฉพาะที่ระดับต่ำ ยอมรับได้	วิเคราะห์ต่อได้
มากกว่า 0.20 ถึง 0.30	พึ่งพิงเฉพาะที่ระดับปานกลาง	รายงานและพิจารณาผลกระทบ
มากกว่า 0.30	พึ่งพิงเฉพาะที่ระดับสูง	รวมหรือตัดข้อสอบที่มีปัญหาออก

หมายเหตุ: ดัดแปลงจาก Christensen, Makransky และ Horton (2017) และ Yen (1984)

4.2.4 โมเดลทฤษฎีการตอบสนองข้อสอบสำหรับข้อสอบแบบให้คะแนนสองค่า

สำหรับข้อสอบที่ให้คะแนนแบบทวิภาค ที่ตอบถูกได้ 1 คะแนน และตอบผิดได้ 0 คะแนน ซึ่งเป็นรูปแบบที่ใช้ในการวิจัยครั้งนี้ มีโมเดลทฤษฎีการตอบสนองข้อสอบที่นิยมใช้ 3 โมเดล ดังนี้

โมเดลแรก คือ โมเดล 1 พารามิเตอร์ พัฒนาโดย Rasch (1960) เป็นโมเดลที่ง่ายที่สุด มีพารามิเตอร์ข้อสอบเพียงตัวเดียวคือค่าความยาก โดยตั้งข้อสมมติว่าข้อสอบทุกข้อในแบบทดสอบมีค่าอำนาจจำแนกเท่ากัน ข้อสมมตินี้ทำให้โมเดลมีคุณสมบัติทางคณิตศาสตร์ที่เรียบง่ายและตีความได้ง่าย แต่เป็นข้อจำกัดเมื่อข้อสอบในแบบทดสอบมีอำนาจจำแนกที่แตกต่างกัน โมเดลที่สอง คือ โมเดล 2 พารามิเตอร์ พัฒนาโดย Birnbaum (1968) เพิ่มพารามิเตอร์ค่าอำนาจจำแนกเข้าไปในโมเดล ทำให้สะท้อนความแตกต่างของอำนาจจำแนกระหว่างข้อสอบได้ ค่าอำนาจจำแนกบ่งบอกความชันของฟังก์ชันการตอบสนองข้อสอบ ค่าที่สูงทำให้เส้นโค้งมีความชันมาก หมายความว่าข้อสอบจำแนกผู้สอบที่มีความสามารถใกล้เคียงกันได้ดี และโมเดลที่สาม คือ โมเดล 3 พารามิเตอร์ เพิ่มพารามิเตอร์ค่าการเดาเข้ามา เหมาะกับข้อสอบแบบเลือกตอบที่ผู้สอบมีโอกาสเดาถูก แต่ต้องการกลุ่มตัวอย่างขนาดใหญ่มาก และประมาณค่าไม่เสถียรหากข้อสอบง่ายเกินไป

การเลือกโมเดลที่เหมาะสมในการวิเคราะห์ข้อมูลสามารถพิจารณาได้จากระดับความสอดคล้องระหว่างโมเดลกับข้อมูลเชิงประจักษ์ โดยดัชนีที่นิยมใช้ในการเปรียบเทียบโมเดล ได้แก่ ค่า AIC (Akaike Information Criterion) และค่า BIC (Bayesian Information Criterion) ซึ่ง

เป็นดัชนีที่คำนึงถึงทั้งความสามารถในการอธิบายข้อมูลและความซับซ้อนของโมเดล โดยโมเดลที่มีค่าดัชนีต่ำกว่าจะถือว่ามีความเหมาะสมมากกว่า นอกจากนี้ ยังสามารถใช้การทดสอบอัตราส่วนภาชนะน่าจะเป็น (Likelihood Ratio Test: LRT) เพื่อเปรียบเทียบโมเดลที่มีความสัมพันธ์แบบซ้อนกัน (nested models) ว่าโมเดลที่มีความซับซ้อนมากกว่าสามารถอธิบายข้อมูลได้ดีกว่าอย่างมีนัยสำคัญทางสถิติหรือไม่ (de Ayala, 2009; Hambleton et al., 1991)

โดยทั่วไป การเลือกโมเดลควรพิจารณาร่วมกันทั้งหลักฐานเชิงสถิติและความเหมาะสมทางทฤษฎี เพื่อให้ได้โมเดลที่สามารถอธิบายข้อมูลได้อย่างมีประสิทธิภาพและมีความประหยัดในการใช้พารามิเตอร์ (parsimony)

4.2.5 ค่าพารามิเตอร์ข้อสอบและการตรวจสอบความสอดคล้องของโมเดล

ภายหลังการเลือกโมเดลทฤษฎีการตอบสนองข้อสอบที่เหมาะสม ขั้นตอนต่อไปคือการประมาณค่าพารามิเตอร์ข้อสอบและตรวจสอบความสอดคล้องของโมเดลกับข้อมูล ค่าพารามิเตอร์อำนาจจำแนก สัญลักษณ์ a บ่งบอกความสามารถของข้อสอบในการจำแนกผู้สอบที่มีความสามารถใกล้เคียงกับค่าความยากของข้อสอบ ค่า a ที่สูงหมายความว่าข้อสอบมีอำนาจจำแนกดี เกณฑ์การแปลผลค่าพารามิเตอร์อำนาจจำแนกตาม Baker (2001) แสดงในตาราง 7

ตาราง 7 เกณฑ์การแปลผลค่าพารามิเตอร์อำนาจจำแนก (a)

ช่วงค่า a	ระดับการจำแนก
น้อยกว่า 0.35	จำแนกได้น้อยมาก
0.35–0.64	จำแนกได้น้อย
0.65–1.34	จำแนกได้ปานกลาง
1.35–1.69	จำแนกได้สูง
1.70 ขึ้นไป	จำแนกได้สูงมาก

หมายเหตุ: ดัดแปลงจาก Baker (2001)

ส่วนค่าพารามิเตอร์ความยาก สัญลักษณ์ b บ่งบอกตำแหน่งของข้อสอบบนแกนความสามารถ θ โดยค่า b คือระดับความสามารถที่ผู้สอบมีโอกาสตอบข้อสอบถูกร้อยละ 50 ค่า b ตีลบหมายความว่าข้อสอบง่าย เพราะผู้สอบที่มีความสามารถต่ำก็ยังมีโอกาสตอบถูกร้อยละ 50 ค่า b เป็นบวกหมายความว่าข้อสอบยาก เพราะต้องใช้ผู้สอบที่มีความสามารถสูงจึงจะมีโอกาสตอบถูกร้อยละ 50

โดยทั่วไปค่า b ในแบบทดสอบจะอยู่ในช่วง -3 ถึง $+3$ ซึ่งสอดคล้องกับช่วงของค่าความสามารถ θ ของผู้สอบ

การตรวจสอบความสอดคล้องของโมเดลกับข้อมูล (Model Fit) เป็นขั้นตอนสำคัญในการวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบ เพื่อพิจารณาว่าโมเดลที่เลือกสามารถอธิบายรูปแบบการตอบของผู้สอบได้สอดคล้องกับข้อมูลเชิงประจักษ์หรือไม่ การตรวจสอบความสอดคล้องสามารถดำเนินการได้ทั้งในระดับโมเดลโดยรวมและระดับรายข้อ

ในระดับโมเดลโดยรวม นิยมใช้สถิติ M^2 ที่พัฒนาโดย Maydeu-Olivares และ Joe (2005) ซึ่งได้รับการพัฒนาขึ้นเพื่อแก้ไขข้อจำกัดของสถิติไคสแควร์แบบดั้งเดิมในการประเมินความสอดคล้องของโมเดลทฤษฎีการตอบสนองข้อสอบ โดยมักพิจารณาพร้อมกับดัชนี RMSEA (Root Mean Square Error of Approximation), CFI (Comparative Fit Index), TLI (Tucker-Lewis Index) และ SRMSR (Standardized Root Mean Square Residual) ซึ่งเป็นดัชนีที่ใช้กันอย่างแพร่หลายในการประเมินความสอดคล้องของโมเดล เกณฑ์ที่นิยมใช้ ได้แก่ ค่า RMSEA ไม่เกิน 0.08 ค่า CFI และ TLI ตั้งแต่ 0.95 ขึ้นไป และค่า SRMSR ไม่เกิน 0.08 ซึ่งบ่งชี้ว่าโมเดลมีความสอดคล้องกับข้อมูลในระดับที่ยอมรับได้

ในระดับรายข้อ นิยมใช้สถิติ $S-X^2$ ที่พัฒนาโดย Orlando และ Thissen (2000) เพื่อตรวจสอบว่าข้อสอบแต่ละข้อมีความสอดคล้องกับโมเดลหรือไม่ โดยพิจารณาจากระดับนัยสำคัญทางสถิติ หากข้อสอบมีค่านัยสำคัญน้อยกว่า .05 อาจบ่งชี้ว่าข้อสอบข้อนั้นไม่สอดคล้องกับโมเดลและควรได้รับการพิจารณาเพิ่มเติม

นอกจากการตรวจสอบความสอดคล้องของโมเดลแล้ว การวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบยังให้ข้อมูลเกี่ยวกับพารามิเตอร์ของข้อสอบ เช่น ค่าอำนาจจำแนก (a) และค่าความยาก (b) ซึ่งช่วยอธิบายคุณภาพของข้อสอบในแต่ละข้อได้อย่างละเอียด โดยเกณฑ์การแปลผลค่าอำนาจจำแนกที่ได้รับการอ้างอิงอย่างแพร่หลาย คือ เกณฑ์ของ Baker (2001)

4.2.6 ฟังก์ชันสารสนเทศของข้อสอบและของแบบวัด

จุดเด่นสำคัญประการหนึ่งของทฤษฎีการตอบสนองข้อสอบที่แตกต่างจากทฤษฎีการทดสอบแบบดั้งเดิม คือ ความสามารถในการแสดงความแม่นยำของการวัดในแต่ละระดับความสามารถของผู้สอบ ผ่านแนวคิดฟังก์ชันสารสนเทศ (Information Function) ซึ่งช่วยให้ทราบว่าแบบวัดหรือข้อสอบแต่ละข้อให้ข้อมูลได้ดีเพียงใดในแต่ละระดับความสามารถ

ฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function: IIF) แสดงปริมาณสารสนเทศที่ข้อสอบแต่ละข้อให้ในแต่ละระดับความสามารถ โดยข้อสอบที่มีค่าอำนาจจำแนกสูงจะให้สารสนเทศมากกว่าข้อสอบที่มีค่าอำนาจจำแนกต่ำ ขณะที่ฟังก์ชันสารสนเทศของแบบวัด (Test

Information Function: TIF) เป็นผลรวมของสารสนเทศจากข้อสอบทุกข้อในแบบวัด จึงสะท้อนความแม่นยำของแบบวัดทั้งฉบับในแต่ละระดับความสามารถ

Embretson และ Reise (2000) อธิบายว่า ฟังก์ชันสารสนเทศมีความสัมพันธ์ผกผันกับความคลาดเคลื่อนมาตรฐานของการวัด (Standard Error of Measurement: SEM) กล่าวคือ ในช่วงระดับความสามารถที่แบบวัดให้สารสนเทศสูง ความคลาดเคลื่อนมาตรฐานของการวัดจะต่ำ และการประมาณค่าความสามารถของผู้สอบจะมีความแม่นยำสูง ในทางตรงกันข้าม หากสารสนเทศมีค่าน้อย ความคลาดเคลื่อนของการวัดจะสูงขึ้น

โดยทั่วไป กราฟฟังก์ชันสารสนเทศของแบบวัดมักมีลักษณะเป็นโค้งที่มีจุดสูงสุด ณ ระดับความสามารถช่วงใดช่วงหนึ่ง ซึ่งสะท้อนว่าแบบวัดมีความเหมาะสมในการวัดผู้สอบที่มีระดับความสามารถในช่วงดังกล่าวมากที่สุด ข้อมูลจากฟังก์ชันสารสนเทศจึงมีประโยชน์ต่อการประเมินคุณภาพของแบบวัด การเลือกใช้แบบวัดให้เหมาะสมกับกลุ่มเป้าหมาย และการพัฒนาคลังข้อสอบในอนาคต

4.2.7 การประมาณค่าความสามารถของผู้สอบรายบุคคล

ภายหลังการประมาณค่าพารามิเตอร์ของข้อสอบแล้ว ทฤษฎีการตอบสนองข้อสอบสามารถนำมาใช้ประมาณค่าความสามารถแฝงของผู้สอบรายบุคคลได้ โดยค่าความสามารถดังกล่าวมักแสดงในรูปของค่า θ (theta) ซึ่งสะท้อนระดับความสามารถของผู้สอบบนมาตรวัดเดียวกับพารามิเตอร์ของข้อสอบ

วิธีการประมาณค่าความสามารถของผู้สอบมีหลายวิธี วิธีที่นิยมใช้ ได้แก่ วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation: MLE) และวิธีคาดหวังภายหลัง (Expected a Posteriori: EAP) วิธีภาวะน่าจะเป็นสูงสุดเป็นการประมาณค่าความสามารถที่ทำให้ความน่าจะเป็นของรูปแบบการตอบที่เกิดขึ้นมีค่าสูงที่สุด มีข้อดีคือไม่ต้องอาศัยการกำหนดการแจกแจงล่วงหน้า แต่มีข้อจำกัดคือไม่สามารถประมาณค่าได้เมื่อผู้สอบตอบถูกทุกข้อหรือผิดทุกข้อ

ส่วนวิธีคาดหวังภายหลังเป็นการประมาณค่าความสามารถโดยใช้ค่าเฉลี่ยของการแจกแจงภายหลัง (posterior distribution) ซึ่งพิจารณาทั้งข้อมูลการตอบของผู้สอบและการแจกแจงล่วงหน้าของความสามารถ วิธีนี้สามารถประมาณค่าได้ในทุกกรณี รวมถึงกรณีที่ผู้สอบตอบถูกทุกข้อหรือผิดทุกข้อ และมีความเสถียรในการประมาณค่ามากกว่าวิธีภาวะน่าจะเป็นสูงสุดในกลุ่มตัวอย่างบางลักษณะ

Bock และ Mislevy (1982) รายงานว่าวิธีคาดหวังภายหลังให้ค่าประมาณความสามารถที่มีความเสถียรและมีความคลาดเคลื่อนต่ำ จึงได้รับความนิยมในการวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบ โดยเฉพาะในการวิจัยทางการศึกษาและจิตวิทยาที่มุ่งประมาณค่าความสามารถของผู้สอบรายบุคคล

4.3 การสร้างเกณฑ์ปกติ

การสร้างเกณฑ์ปกติเป็นขั้นตอนสำคัญในการพัฒนาเครื่องมือวัดที่สมบูรณ์ เพื่อให้ผู้ใช้สามารถตีความคะแนนของผู้สอบในบริบทเปรียบเทียบกับกลุ่มอ้างอิงได้อย่างมีความหมาย แบบทดสอบที่ผ่านการตรวจสอบคุณภาพทางจิตมิติแล้วแม้จะมีคุณภาพดี แต่หากไม่มีเกณฑ์ปกติ ผู้ใช้ก็ไม่สามารถนำคะแนนของผู้สอบมาเปรียบเทียบกับกลุ่มอ้างอิงได้ ทำให้คะแนนที่ได้มีความหมายเพียงในเชิงคะแนนดิบ ในหัวข้อนี้ ผู้วิจัยนำเสนอเนื้อหาเกี่ยวกับการสร้างเกณฑ์ปกติ ประกอบด้วย ความหมายและประเภทของเกณฑ์ปกติ ประเภทของคะแนนมาตรฐาน ขั้นตอนการสร้างเกณฑ์ปกติ และเกณฑ์การแปลผลระดับทักษะ

4.3.1 ความหมายและประเภทของเกณฑ์ปกติ

เกณฑ์ปกติ (Norms) หมายถึง ค่าทางสถิติที่ได้จากการวัดหรือการทดสอบกับกลุ่มอ้างอิงที่เป็นตัวแทนของประชากรเป้าหมาย ซึ่งใช้เป็นกรอบในการแปลความหมายคะแนนของผู้สอบ โดยช่วยให้สามารถเปรียบเทียบผลการประเมินของบุคคลกับผลการประเมินของกลุ่มอ้างอิงได้ เกณฑ์ปกติจึงเป็นองค์ประกอบสำคัญที่ช่วยให้คะแนนดิบมีความหมายมากยิ่งขึ้นในการนำผลการประเมินไปใช้

Anastasi และ Urbina (1997) อธิบายว่า เกณฑ์ปกติสามารถจำแนกได้หลายประเภทตามลักษณะของกลุ่มอ้างอิง เช่น เกณฑ์ปกติระดับชาติ (National Norms) ซึ่งพัฒนาจากกลุ่มตัวอย่างที่เป็นตัวแทนของประชากรทั้งประเทศ และเกณฑ์ปกติเฉพาะพื้นที่ (Local Norms) ซึ่งพัฒนาจากกลุ่มตัวอย่างในพื้นที่หรือกลุ่มประชากรเฉพาะ เช่น ระดับจังหวัด เขตพื้นที่การศึกษา หรือสถานศึกษา การเลือกใช้เกณฑ์ปกติแต่ละประเภทขึ้นอยู่กับวัตถุประสงค์ของการประเมิน ลักษณะของกลุ่มเป้าหมาย และขอบเขตการนำผลการประเมินไปใช้

โดยทั่วไป เกณฑ์ปกติเฉพาะพื้นที่มีข้อได้เปรียบในการสะท้อนลักษณะของกลุ่มอ้างอิงที่ใกล้เคียงกับผู้ได้รับการประเมินมากกว่า จึงช่วยให้การแปลผลคะแนนมีความเหมาะสมกับบริบทของกลุ่มเป้าหมาย ขณะที่เกณฑ์ปกติระดับชาติมีข้อได้เปรียบในด้านการเปรียบเทียบผลการประเมินกับประชากรในวงกว้าง

4.3.2 ประเภทของคะแนนมาตรฐานและการเลือกใช้ในการวิจัย

คะแนนดิบไม่สามารถสะท้อนตำแหน่งสัมพัทธ์ของผู้สอบเมื่อเปรียบเทียบกับกลุ่มอ้างอิงได้โดยตรง จึงจำเป็นต้องแปลงคะแนนให้อยู่ในรูปคะแนนแสดงตำแหน่งหรือคะแนนมาตรฐานเพื่อใช้ในการแปลผล เกณฑ์ปกติสามารถจัดทำได้จากคะแนนหลายรูปแบบ โดยคะแนนที่นิยมใช้ ได้แก่ ลำดับเปอร์เซ็นต์ไทล์และคะแนนทีปกติ

ลำดับเปอร์เซ็นต์ไทล์ (Percentile Rank) เป็นการแสดงตำแหน่งของผู้สอบเมื่อเปรียบเทียบกับกลุ่มอ้างอิง โดยบ่งบอกร้อยละของผู้สอบที่มีคะแนนต่ำกว่าหรือเท่ากับคะแนน

ที่พิจารณา ข้อดีของลำดับเปอร์เซ็นต์คือสามารถทำความเข้าใจได้ง่ายและเหมาะสำหรับการสื่อสารผลการประเมินแก่ผู้ใช้ทั่วไป อย่างไรก็ตาม ช่วงคะแนนเปอร์เซ็นต์ไม่ได้มีระยะห่างที่เท่ากันตลอดช่วงการแจกแจง จึงมีข้อจำกัดในการนำไปใช้เปรียบเทียบความแตกต่างระหว่างคะแนน

ชวาล แพรัตกุล (2552) อธิบายว่า การสร้างเกณฑ์ปกติควรใช้คะแนนที่ปกติ (Normalized T-score) ซึ่งได้จากการแปลงลำดับเปอร์เซ็นต์ให้เป็นคะแนนซีภายใต้การแจกแจงปกติมาตรฐาน แล้วจึงแปลงเป็นคะแนนที่มีค่าเฉลี่ยเท่ากับ 50 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 วิธีการดังกล่าวช่วยให้คะแนนมีการแจกแจงสอดคล้องกับโค้งปกติ ลดผลกระทบจากการแจกแจงที่ไม่เป็นปกติของคะแนนดิบ และเอื้อต่อการกำหนดเกณฑ์การแปลผลในลักษณะของระดับความสามารถ

จากข้อดีของคะแนนทั้งสองรูปแบบ นักวิชาการด้านการวัดและประเมินผลจึงนิยมใช้ลำดับเปอร์เซ็นต์ร่วมกับคะแนนที่ปกติในการสร้างเกณฑ์ปกติ โดยลำดับเปอร์เซ็นต์ใช้แสดงตำแหน่งสัมพัทธ์ของผู้สอบในกลุ่มอ้างอิง ขณะที่คะแนนที่ปกติใช้เป็นพื้นฐานในการกำหนดระดับผลการประเมินและการเปรียบเทียบผลการวัดระหว่างบุคคล (ชวาล แพรัตกุล, 2552; สมนึก ภัททิยธนี, 2562)

4.3.3 ขั้นตอนการสร้างเกณฑ์ปกติ

สมนึก ภัททิยธนี (2562) เสนอแนวทางการสร้างเกณฑ์ปกติในรูปคะแนนที่ปกติ (T-score) โดยเริ่มจากการจัดเรียงคะแนนดิบและคำนวณความถี่สะสม จากนั้นแปลงคะแนนดิบเป็นลำดับเปอร์เซ็นต์ และแปลงลำดับเปอร์เซ็นต์เป็นคะแนนซีภายใต้การแจกแจงปกติมาตรฐาน ก่อนแปลงเป็นคะแนนที่มีค่าเฉลี่ยเท่ากับ 50 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 ขั้นตอนสุดท้ายคือการจัดทำตารางเกณฑ์ปกติที่แสดงความสัมพันธ์ระหว่างคะแนนดิบ ลำดับเปอร์เซ็นต์ และคะแนนที่ปกติ เพื่อใช้เป็นเกณฑ์ในการแปลผลคะแนน

แนวทางดังกล่าวเป็นวิธีที่ได้รับการยอมรับในการพัฒนาเกณฑ์ปกติสำหรับการประเมินผลทางการศึกษา เนื่องจากช่วยให้สามารถเปรียบเทียบผลการวัดของผู้เรียนกับกลุ่มอ้างอิงได้อย่างเป็นระบบ และเอื้อต่อการตีความความหมายของคะแนนที่ได้รับ

4.3.4 เกณฑ์การแปลผลระดับทักษะ

นอกเหนือจากการจัดทำเกณฑ์ปกติในรูปของลำดับเปอร์เซ็นต์และคะแนนที่ปกติแล้ว การแปลผลคะแนนควรกำหนดระดับผลการประเมินเพื่อให้การตีความคะแนนมีความชัดเจนและเป็นระบบ สมนึก ภัททิยธนี (2562) เสนอการแปลผลคะแนนที่ปกติเป็น 3 ระดับ โดยอาศัยหลักการแจกแจงปกติและกำหนดค่า ± 1 ส่วนเบี่ยงเบนมาตรฐานจากค่าเฉลี่ยเป็นจุดแบ่งระดับ ดังแสดงในตาราง 8

ตาราง 8 เกณฑ์การแปลผลคะแนนที่เป็น 3 ระดับ

ระดับคะแนน	ระดับการแปลผล	ความหมาย
$T \geq 60$ สูงกว่าเกณฑ์ปกติ	สูงกว่าเกณฑ์ปกติ	สูงกว่าค่าเฉลี่ย 1 ส่วนเบี่ยงเบนมาตรฐานขึ้นไป ครอบคลุมประมาณร้อยละ 16 บนของกลุ่ม
$T 40-59$	อยู่ในเกณฑ์ปกติ	อยู่ในช่วง ± 1 ส่วนเบี่ยงเบนมาตรฐานจากค่าเฉลี่ย ครอบคลุมประมาณร้อยละ 68 กลางของกลุ่ม
$T < 40$	ต่ำกว่าเกณฑ์ปกติ	ต่ำกว่าค่าเฉลี่ย 1 ส่วนเบี่ยงเบนมาตรฐานลงไป ครอบคลุมประมาณร้อยละ 16 ล่างของกลุ่ม

หมายเหตุ: ดัดแปลงจาก สมนึก ภัททิยธนี (2562)

แนวทางดังกล่าวสอดคล้องกับคุณลักษณะของการแจ่มแจ้งปกติ ซึ่งคะแนนส่วนใหญ่มักจะกระจุกตัวอยู่บริเวณค่าเฉลี่ย ขณะที่คะแนนที่สูงหรือต่ำกว่าค่าเฉลี่ยอย่างชัดเจนจะอยู่ในสัดส่วนที่น้อยกว่า การกำหนดระดับผลการประเมินตามหลักการนี้จึงช่วยให้การจำแนกผลการวัดเป็นไปอย่างมีหลักเกณฑ์และสอดคล้องกับลักษณะการกระจายของข้อมูล

แนวทางการแปลผลคะแนนดังกล่าวสามารถนำมาใช้เป็นกรอบในการกำหนดระดับผลการประเมินของแบบวัดทักษะการคิดเชิงคำนวณ โดยจำแนกผลการประเมินออกเป็น 3 ระดับ ได้แก่ ระดับต่ำกว่าเกณฑ์ปกติ ระดับอยู่ในเกณฑ์ปกติ และระดับสูงกว่าเกณฑ์ปกติ

4.4 การประยุกต์ใช้กรอบแนวคิดทางจิตมิติในการวิจัยครั้งนี้

จากการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องในหัวข้อ 4.1 ถึง 4.3 พบว่า การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณจำเป็นต้องอาศัยหลักฐานเชิงประจักษ์จากการตรวจสอบคุณภาพเครื่องมือในหลายมิติ เพื่อให้สามารถสนับสนุนความน่าเชื่อถือของผลการวัดได้อย่างรอบด้าน ด้วยเหตุนี้ การวิจัยครั้งนี้จึงบูรณาการแนวคิดจากทฤษฎีการทดสอบแบบดั้งเดิม ทฤษฎีการตอบสนองข้อสอบ และแนวทางการสร้างเกณฑ์ปกติ มาใช้เป็นกรอบในการพัฒนาและตรวจสอบคุณภาพของแบบวัด

ภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิม มีการตรวจสอบความตรงเชิงเนื้อหา คุณภาพรายข้อ และความเชื่อมั่นของแบบวัด เพื่อประเมินคุณภาพพื้นฐานของข้อคำถามและแบบวัด โดยรวม ส่วนกรอบทฤษฎีการตอบสนองข้อสอบใช้ในการตรวจสอบข้อตกลงเบื้องต้นของโมเดล

การประมาณค่าพารามิเตอร์ข้อสอบ และการประเมินความเหมาะสมของโมเดล เพื่อศึกษาคุณภาพของข้อสอบในระดับรายข้อและความสัมพันธ์ระหว่างข้อสอบกับความสามารถของผู้ตอบ สำหรับการสร้างเกณฑ์ปกติ ใช้ลำดับเปอร์เซ็นต์ไทล์และคะแนนที่ปกติในการจัดทำเกณฑ์อ้างอิงสำหรับการแปลผลคะแนน

การบูรณาการกรอบแนวคิดทั้งสามส่วนช่วยให้การตรวจสอบคุณภาพของแบบวัดครอบคลุมทั้งด้านความตรง ความเชื่อมั่น คุณภาพของข้อสอบ และการแปลผลคะแนน ซึ่งเป็นองค์ประกอบสำคัญของการพัฒนาเครื่องมือวัดทางการศึกษาตามหลักจิตมิติ

5. งานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้เกี่ยวข้องกับองค์ความรู้หลายด้าน ได้แก่ การวัดทักษะการคิดเชิงคำนวณ การพัฒนาแบบวัดเชิงสถานการณ์ การวัดและประเมินผลในรูปแบบออนไลน์ และการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบในการพัฒนาเครื่องมือวัดทางการศึกษา ผู้วิจัยจึงทบทวนงานวิจัยที่เกี่ยวข้องตามประเด็นสำคัญของการวิจัย เพื่อศึกษาพัฒนาการขององค์ความรู้ แนวทางการพัฒนาเครื่องมือวัดและประเด็นที่ยังต้องการการศึกษาเพิ่มเติม อันเป็นพื้นฐานในการกำหนดแนวทางการวิจัยครั้งนี้

5.1 งานวิจัยที่เกี่ยวข้องกับการวัดทักษะการคิดเชิงคำนวณ

ในช่วงทศวรรษที่ผ่านมา การพัฒนาเครื่องมือวัดทักษะการคิดเชิงคำนวณได้รับความสนใจอย่างต่อเนื่องในหลายประเทศ โดยมีการพัฒนาเครื่องมือสำหรับผู้เรียนในระดับการศึกษาที่แตกต่างกัน และใช้แนวทางการประเมินที่หลากหลาย ทั้งการทดสอบแบบปรนัย การประเมินจากผลงาน การสังเกตพฤติกรรม และการสัมภาษณ์ สะท้อนให้เห็นว่าทักษะการคิดเชิงคำนวณเป็นโครงสร้างเชิงพหุมิติที่ต้องอาศัยวิธีการวัดที่เหมาะสมและสอดคล้องกับลักษณะของทักษะที่ต้องการประเมิน

Román-González, Pérez-González และ Jiménez-Fernández (2017) ได้พัฒนาแบบทดสอบการคิดเชิงคำนวณสำหรับนักเรียนระดับประถมศึกษาและมัธยมศึกษาในประเทศสเปน โดยใช้ข้อสอบปรนัยที่นำเสนอปัญหาผ่านภาพและสัญลักษณ์ ผลการวิจัยพบว่าแบบทดสอบมีความสัมพันธ์กับความสามารถทางการคิดด้านอื่น เช่น การคิดเชิงตรรกะและการคิดเชิงพื้นที่ ซึ่งสนับสนุนความตรงเชิงโครงสร้างของแบบวัดและแสดงให้เห็นว่าทักษะการคิดเชิงคำนวณสามารถวัดได้ด้วยเครื่องมือที่มีการพัฒนาอย่างเป็นระบบ

ในอีกแนวทางหนึ่ง Brennan และ Resnick (2012) เสนอกรอบการประเมินทักษะการคิดเชิงคำนวณผ่านการใช้โปรแกรม Scratch โดยใช้ข้อมูลจากผลงานของผู้เรียน การสังเกต และการสัมภาษณ์ ผลการวิจัยนำไปสู่การเสนอองค์ประกอบของทักษะการคิดเชิงคำนวณใน 3 มิติ ได้แก่ แนวคิด

เชิงคำนวณ แนวปฏิบัติเชิงคำนวณ และมุมมองเชิงคำนวณ ซึ่งสะท้อนว่าการประเมินทักษะการคิดเชิงคำนวณสามารถดำเนินการได้หลากหลายรูปแบบ ขึ้นอยู่กับกรอบแนวคิดและวัตถุประสงค์ของการวัด

นอกจากนี้ Korkmaz, Çakir และ Özden (2017) ได้พัฒนามาตรวัดการคิดเชิงคำนวณสำหรับนักศึกษาระดับอุดมศึกษา และตรวจสอบคุณภาพของเครื่องมือด้วยการวิเคราะห์องค์ประกอบ ซึ่งแสดงให้เห็นแนวทางการพัฒนาเครื่องมือวัดที่อาศัยหลักการทางจิตมิติในการยืนยันคุณภาพของแบบวัด ขณะที่ Tang และคณะ (2020) ได้ทบทวนงานวิจัยด้านการวัดและประเมินทักษะการคิดเชิงคำนวณจำนวน 96 เรื่อง และพบว่าเครื่องมือส่วนใหญ่มุ่งวัดความสามารถด้านการเขียนโปรแกรมหรือกิจกรรมที่เกี่ยวข้องกับการเขียนโปรแกรมเป็นหลัก ขณะที่เครื่องมือที่ใช้สถานการณ์ในชีวิตประจำวันและไม่อาศัยทักษะการเขียนโปรแกรมยังมีจำนวนจำกัด

จากการทบทวนงานวิจัยข้างต้น พบว่าแม้จะมีการพัฒนาเครื่องมือวัดทักษะการคิดเชิงคำนวณอย่างต่อเนื่อง แต่เครื่องมือส่วนใหญ่ยังเน้นการวัดในบริบทของการเขียนโปรแกรม หรือใช้ข้อคำถามที่แยกวัดแต่ละสถานการณ์เป็นรายข้อ ขณะที่การพัฒนาแบบวัดเชิงสถานการณ์ที่สามารถประเมินองค์ประกอบของทักษะการคิดเชิงคำนวณได้อย่างครอบคลุมโดยไม่ต้องอาศัยความรู้ด้านการเขียนโปรแกรมยังมีไม่มากนัก ประเด็นดังกล่าวจึงเป็นพื้นฐานสำคัญในการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในการวิจัยครั้งนี้

ในบริบทของประเทศไทย มีงานวิจัยหลายเรื่องที่ศึกษาการส่งเสริมและพัฒนาทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้น โดยใช้สถานการณ์ปัญหาและกิจกรรมการเรียนรู้ที่เชื่อมโยงกับบริบทในชีวิตประจำวัน พิเชฐ ศรีสังข์งาม และชัยศ เดชสุระ (2564) ได้พัฒนากิจกรรมการเรียนรู้ตามแนวทางสะเต็มศึกษาเพื่อส่งเสริมทักษะการคิดเชิงคำนวณ โดยออกแบบกิจกรรมที่เปิดโอกาสให้ผู้เรียนใช้การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึมในการแก้ปัญหา ผลการวิจัยพบว่าผู้เรียนมีทักษะการคิดเชิงคำนวณสูงขึ้นหลังการจัดกิจกรรม สะท้อนให้เห็นว่าองค์ประกอบของการคิดเชิงคำนวณทั้ง 4 ด้านสามารถประยุกต์ใช้กับสถานการณ์ปัญหาในชีวิตจริงของผู้เรียนระดับมัธยมศึกษาตอนต้นได้

ในทำนองเดียวกัน ยุภารัตน์ พิเชษฐ์ (2564) ศึกษาการพัฒนาทักษะการคิดเชิงคำนวณโดยใช้การจัดการเรียนรู้แบบใช้ปัญหาเป็นฐาน และพบว่าวิธีการดังกล่าวช่วยส่งเสริมทักษะการคิดเชิงคำนวณของผู้เรียนได้อย่างมีประสิทธิภาพ ข้อค้นพบนี้สะท้อนถึงความเหมาะสมของการใช้สถานการณ์ปัญหาเป็นบริบทในการส่งเสริมและประเมินทักษะการคิดเชิงคำนวณ

นอกจากนี้ วัชรภรณ์ พิงโพธิ์ (2567) ได้ศึกษาแนวทางการจัดการเรียนการสอนเพื่อส่งเสริมการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้น ซึ่งให้ข้อมูลเกี่ยวกับลักษณะและระดับทักษะการคิดเชิงคำนวณของผู้เรียนในบริบทไทย อันเป็นข้อมูลพื้นฐานที่ช่วยให้เห็นลักษณะของกลุ่มเป้าหมายที่ใช้ในการวิจัยครั้งนี้

เมื่อพิจารณาร่วมกัน จะเห็นได้ว่างานวิจัยในบริบทไทยส่วนใหญ่มุ่งศึกษาการพัฒนาทักษะการคิดเชิงคำนวณผ่านกิจกรรมการเรียนรู้และสถานการณ์ปัญหาในชีวิตประจำวัน ขณะที่งานวิจัยที่มุ่งพัฒนาเครื่องมือวัดทักษะการคิดเชิงคำนวณโดยตรงยังมีจำนวนไม่มากนัก โดยเฉพาะเครื่องมือที่นำเสนอปัญหาในรูปแบบเชิงสถานการณ์และพัฒนาเป็นระบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น ประเด็นดังกล่าวจึงเป็นพื้นฐานสำคัญในการกำหนดแนวทางการวิจัยครั้งนี้

5.2 งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบวัดเชิงสถานการณ์

สำหรับงานวิจัยต่างประเทศ การพัฒนาแบบวัดเชิงสถานการณ์มีพื้นฐานมาจากแนวคิดที่มุ่งประเมินพฤติกรรมหรือกระบวนการคิดของผู้ตอบภายใต้บริบทที่ใกล้เคียงกับสถานการณ์จริง โดย Motowidlo, Dunnette และ Carter (1990) เป็นกลุ่มนักวิจัยสำคัญที่เสนอแนวคิดของแบบวัดเชิงสถานการณ์อย่างเป็นระบบ และชี้ให้เห็นว่าการนำเสนอปัญหาผ่านสถานการณ์จำลองสามารถสะท้อนพฤติกรรมและการตัดสินใจของผู้ตอบได้ใกล้เคียงกับบริบทจริงมากกว่าการวัดแบบดั้งเดิม

ต่อมา McDaniel และคณะ (2001) ได้สังเคราะห์งานวิจัยเกี่ยวกับแบบวัดเชิงสถานการณ์และสรุปองค์ประกอบสำคัญของเครื่องมือประเภทนี้ไว้ 3 ส่วน ได้แก่ สถานการณ์ (Situation) ทางเลือกคำตอบ (Response Options) และระบบการให้คะแนน (Scoring System) ซึ่งเป็นองค์ประกอบพื้นฐานที่ได้รับการนำไปใช้ในการออกแบบแบบวัดเชิงสถานการณ์ในงานวิจัยจำนวนมาก

นอกจากนี้ Christian, Edwards และ Bradley (2010) ได้สังเคราะห์งานวิจัยเชิงอภิมานเกี่ยวกับความตรงเชิงเกณฑ์ของแบบวัดเชิงสถานการณ์ และพบว่าเครื่องมือประเภทนี้มีความตรงเชิงเกณฑ์อยู่ในระดับที่น่าพอใจ สะท้อนให้เห็นว่าแบบวัดเชิงสถานการณ์สามารถใช้ประเมินทักษะและคุณลักษณะที่มีความซับซ้อนได้อย่างมีประสิทธิภาพ

จากการทบทวนงานวิจัยต่างประเทศ พบว่ามีการพัฒนาองค์ความรู้เกี่ยวกับแบบวัดเชิงสถานการณ์อย่างต่อเนื่อง ทั้งในด้านการออกแบบแนวคิด องค์ประกอบของเครื่องมือ และหลักฐานเชิงประจักษ์ด้านคุณภาพการวัด ซึ่งเป็นพื้นฐานสำคัญสำหรับการออกแบบและพัฒนาแบบวัดเชิงสถานการณ์ในการวิจัยทางการศึกษา

ในบริบทการศึกษาไทย การพัฒนาแบบวัดเชิงสถานการณ์ได้รับความสนใจอย่างต่อเนื่อง เนื่องจากเป็นรูปแบบการวัดที่นำเสนอปัญหาหรือเหตุการณ์ที่ใกล้เคียงกับสถานการณ์จริง ช่วยให้ผู้ตอบแสดงกระบวนการคิดและการตัดสินใจภายใต้บริบทที่มีความหมาย ส่งผลให้การวัดสามารถสะท้อนทักษะหรือพฤติกรรมที่ต้องการประเมินได้ใกล้เคียงกับการนำไปใช้จริงมากกว่าการวัดด้วยข้อคำถามที่ปราศจากบริบท

สุพัชญา เจริรัตน์ (2560) ได้พัฒนาแบบวัดทักษะการแก้ปัญหาเชิงสถานการณ์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น และเปรียบเทียบคุณสมบัติทางจิตมิติระหว่างรูปแบบเขียนตอบกับ

รูปแบบออนไลน์ ผลการวิจัยพบว่าแบบวัดทั้งสองรูปแบบมีคุณภาพใกล้เคียงกัน ขณะที่รูปแบบออนไลน์มีความสะดวกในการบริหารการทดสอบและการรายงานผลมากกว่า

ไพลิน จันทิษฐ (2563) ได้พัฒนาแบบวัดทักษะการคิดแก้ปัญหาในรูปแบบเชิงสถานการณ์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น โดยใช้สถานการณ์เป็นฐานในการสร้างข้อคำถามและตรวจสอบคุณภาพของแบบวัด ผลการวิจัยพบว่าแบบวัดมีความตรงเชิงเนื้อหาและคุณภาพรายข้ออยู่ในเกณฑ์ที่ยอมรับได้

นอกจากนี้ วรรณกร พรประเสริฐ (2564) ได้พัฒนาแบบวัดเชิงสถานการณ์การรู้เท่าทันดิจิทัลของนิสิตนักศึกษาครู พร้อมทั้งตรวจสอบคุณภาพของเครื่องมือและสร้างเกณฑ์ปกติสำหรับการแปลผลคะแนน ขณะที่ แพรวนภา แก้วกุก พนิดา ศกุนตนาค และอรอุมา เจริญสุข (2566) ได้พัฒนาแบบวัดเชิงสถานการณ์เพื่อวัดความเป็นพลเมืองดิจิทัลของนักเรียนระดับมัธยมศึกษาตอนต้น ซึ่งผลการวิจัยพบว่าแบบวัดมีความตรงและความเชื่อมั่นอยู่ในระดับที่ยอมรับได้

เมื่อพิจารณางานวิจัยในประเทศไทยร่วมกัน พบว่างานส่วนใหญ่มุ่งพัฒนาแบบวัดเชิงสถานการณ์เพื่อประเมินทักษะการคิดหรือคุณลักษณะทางการศึกษาเฉพาะด้าน และผลการวิจัยโดยทั่วไปสนับสนุนว่าแบบวัดเชิงสถานการณ์สามารถพัฒนาให้มีคุณภาพทางจิตมิติอยู่ในเกณฑ์ที่ยอมรับได้ ทั้งในรูปแบบกระดาษและรูปแบบออนไลน์ อย่างไรก็ตาม งานวิจัยที่มุ่งพัฒนาแบบวัดเชิงสถานการณ์สำหรับการประเมินทักษะการคิดเชิงคำนวณโดยตรงยังมีจำนวนจำกัด โดยเฉพาะในกลุ่มนักเรียนระดับมัธยมศึกษาตอนต้นและในรูปแบบออนไลน์ที่มีระบบรายงานผลอัตโนมัติ จึงยังปรากฏช่องว่างการวิจัยที่ควรได้รับการพัฒนาเพิ่มเติม

5.3 งานวิจัยที่เกี่ยวข้องกับการวัดและประเมินผลในรูปแบบออนไลน์

การวัดและประเมินผลในรูปแบบออนไลน์ได้รับความสนใจเพิ่มขึ้นอย่างต่อเนื่องจากความก้าวหน้าของเทคโนโลยีดิจิทัล โดยงานวิจัยในสาขานี้มุ่งศึกษาทั้งประสิทธิภาพของระบบการประเมินออนไลน์ การจัดการข้อมูล และการรายงานผลการประเมิน

Black และ Wiliam (2009) ได้ศึกษาแนวทางการประเมินเพื่อการเรียนรู้ และชี้ให้เห็นว่าการรายงานผลการประเมินควรนำเสนอข้อมูลที่มีความหมายต่อผู้เรียนและสามารถนำไปใช้ปรับปรุงการเรียนรู้ได้จริง แนวคิดดังกล่าวสะท้อนความสำคัญของการรายงานผลที่ไม่จำกัดเฉพาะคะแนนรวม แต่ควรแสดงข้อมูลที่ช่วยให้ผู้เรียนและครูเข้าใจจุดเด่นและจุดที่ควรพัฒนาในแต่ละด้านของการเรียนรู้

นอกจากนี้ Alruwais, Wills และ Wald (2018) ได้ทบทวนวรรณกรรมเกี่ยวกับการประเมินออนไลน์และสรุปว่าระบบการประเมินออนไลน์มีข้อได้เปรียบหลายประการ ได้แก่ การตรวจให้คะแนนอัตโนมัติ การรายงานผลได้อย่างรวดเร็ว การจัดเก็บข้อมูลในรูปแบบดิจิทัล และความ

สะดวกในการนำข้อมูลไปวิเคราะห์ต่อ ซึ่งช่วยเพิ่มประสิทธิภาพของกระบวนการวัดและประเมินผล เมื่อเปรียบเทียบกับ การดำเนินการในรูปแบบดั้งเดิม

เมื่อพิจารณาร่วมกัน จะเห็นได้ว่างานวิจัยที่เกี่ยวข้องกับการประเมินออนไลน์ ให้ความสำคัญกับการจัดการข้อมูล การรายงานผล และการนำผลการประเมินไปใช้ประโยชน์ในการ พัฒนาการเรียนรู้ ขณะเดียวกัน เทคโนโลยีออนไลน์ยังช่วยเพิ่มประสิทธิภาพในการให้คะแนน การจัดเก็บ ข้อมูล และการสื่อสารผลการประเมินแก่ผู้สอบ ประเด็นดังกล่าวจึงเป็นพื้นฐานสำคัญในการพัฒนาแบบวัด ทักษะการคิดเชิงคำนวณในรูปแบบออนไลน์และระบบรายงานผลอัตโนมัติในการวิจัยครั้งนี้

5.4 งานวิจัยที่เกี่ยวข้องกับการพัฒนาเครื่องมือวัดด้วยทฤษฎีการตอบสนองข้อสอบและ การสร้างเกณฑ์ปกติ

การพัฒนาเครื่องมือวัดทางการศึกษาในปัจจุบันมีการนำทฤษฎีการตอบสนองข้อสอบ มาใช้เพิ่มขึ้นอย่างต่อเนื่อง เนื่องจากทฤษฎีดังกล่าวสามารถให้สารสนเทศเกี่ยวกับคุณภาพของข้อสอบ และความสามารถของผู้สอบได้อย่างละเอียดมากกว่าทฤษฎีการทดสอบแบบดั้งเดิม อีกทั้งยัง สนับสนุนการพัฒนาเกณฑ์อ้างอิงสำหรับการแปลผลคะแนนได้อย่างมีประสิทธิภาพ

ในด้านกรอบแนวคิดทางทฤษฎี Lord (1980) ได้นำเสนอแนวคิดฟังก์ชันสารสนเทศ ของข้อสอบและแบบวัด ซึ่งเป็นแนวคิดสำคัญที่แสดงให้เห็นระดับความแม่นยำของการวัดในแต่ละ ช่วงความสามารถของผู้สอบ ต่อมา Hambleton, Swaminathan และ Rogers (1991) ได้อธิบาย ข้อตกลงเบื้องต้นที่สำคัญของทฤษฎีการตอบสนองข้อสอบ ได้แก่ ความเป็นเอกมิติและความเป็นอิสระ เฉพาะที่ ซึ่งถือเป็นเงื่อนไขพื้นฐานก่อนการประมาณค่าพารามิเตอร์ของโมเดล

นอกจากนี้ Embretson และ Reise (2000) ได้นำเสนอแนวทางการประยุกต์ใช้ทฤษฎี การตอบสนองข้อสอบในการพัฒนาเครื่องมือวัดทางการศึกษาและจิตวิทยา ขณะที่ Orlando และ Thissen (2000) พัฒนาสถิติ $S-X^2$ สำหรับการตรวจสอบความสอดคล้องของข้อสอบรายข้อ และ Maydeu-Olivares และ Joe (2005) พัฒนาสถิติ M2 สำหรับการประเมินความสอดคล้องของโมเดล โดยรวม ส่วน Baker (2001) ได้เสนอแนวทางการแปลความหมายค่าพารามิเตอร์ข้อสอบที่ได้รับการ นำมาใช้อย่างแพร่หลายในงานวิจัยด้านการวัดผล

ในด้านเครื่องมือวิเคราะห์ข้อมูล Chalmers (2012) ได้พัฒนาแพ็คเกจ mirt สำหรับการ วิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบ และ Rosseel (2012) ได้พัฒนาแพ็คเกจ lavaan สำหรับการวิเคราะห์องค์ประกอบเชิงยืนยัน ซึ่งเป็นเครื่องมือที่ได้รับความนิยมในการ ตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดในงานวิจัยทางการศึกษา

องค์ความรู้ดังกล่าวสะท้อนให้เห็นพัฒนาการของแนวคิด วิธีการ และเครื่องมือที่ใช้ใน การวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบ ซึ่งเป็นพื้นฐานสำคัญในการตรวจสอบคุณสมบัติ ทางจิตมิติของแบบวัดและการสร้างเกณฑ์ปกติในการวิจัยครั้งนี้

ในบริบทการวัดผลทางการศึกษาไทย มีเอกสารวิชาการหลายฉบับที่อธิบายแนวทางการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบในการพัฒนาเครื่องมือวัด ศิริชัย กาญจนวาสี (2556) อธิบายความสัมพันธ์ระหว่างทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ พร้อมเสนอแนวทางการเลือกและรายงานผลการวิเคราะห์ให้สอดคล้องกับลักษณะของข้อมูลและวัตถุประสงค์ของการวัด ขณะที่ สมนึก ภัททิยธนี (2562) และ อนุวัติ คุณแก้ว (2566) ได้เสนอแนวทางการตรวจสอบคุณภาพเครื่องมือวัดที่ผสมผสานการใช้ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบร่วมกัน เพื่อให้ได้หลักฐานคุณภาพของแบบวัดที่ครอบคลุมทั้งในระดับข้อสอบและระดับแบบวัด

สำหรับงานวิจัยเชิงประยุกต์ พิสิฐพงศ์ เจริมเมฆ, เทตศักดิ์ สุพันธุ์ และ สมประสงค์ เสนารัตน์ (2567) ได้พัฒนาคำสั่งข้อสอบวิชาวิทยาศาสตร์และเทคโนโลยีสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้ทฤษฎีการตอบสนองข้อสอบในการวิเคราะห์คุณภาพข้อสอบและประมาณค่าพารามิเตอร์ ผลการวิจัยแสดงให้เห็นว่าทฤษฎีการตอบสนองข้อสอบสามารถนำมาใช้พัฒนาข้อสอบและจัดทำคำสั่งข้อสอบที่มีคุณภาพได้อย่างมีประสิทธิภาพในบริบทการศึกษาไทย งานวิจัยดังกล่าวมีความเกี่ยวข้องกับการวิจัยครั้งนี้ เนื่องจากใช้กลุ่มผู้เรียนระดับมัธยมศึกษาตอนต้นเช่นเดียวกัน และสะท้อนให้เห็นความเหมาะสมของการใช้ทฤษฎีการตอบสนองข้อสอบในการตรวจสอบคุณภาพเครื่องมือวัดทางการศึกษา

เมื่อพิจารณาร่วมกับงานวิจัยและเอกสารวิชาการทั้งในและต่างประเทศ จะเห็นได้ว่าทฤษฎีการตอบสนองข้อสอบได้รับการยอมรับอย่างกว้างขวางในการพัฒนาและตรวจสอบคุณภาพเครื่องมือวัดทางการศึกษา ทั้งในด้านการวิเคราะห์คุณภาพข้อสอบ การประเมินความเหมาะสมของโมเดล และการศึกษาความแม่นยำของการวัด ซึ่งเป็นแนวทางสำคัญที่ผู้วิจัยนำมาประยุกต์ใช้ในการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดทักษะการคิดเชิงคำนวณในการวิจัยครั้งนี้

จากการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง สามารถสังเคราะห์ประเด็นสำคัญและช่องว่างของการวิจัยได้ 4 ประการ ดังนี้

ประการแรก งานวิจัยด้านการวัดทักษะการคิดเชิงคำนวณส่วนใหญ่มุ่งวัดองค์ประกอบของทักษะแยกจากกัน หรือเน้นการประเมินความสามารถด้านการเขียนโปรแกรมเป็นหลัก ขณะที่สถานการณ์การแก้ปัญหาในชีวิตจริงมักต้องอาศัยการใช้ทักษะการคิดเชิงคำนวณหลายองค์ประกอบร่วมกัน จึงยังมีความจำเป็นในการพัฒนาเครื่องมือวัดที่สามารถประเมินการใช้ทักษะทั้ง 4 องค์ประกอบภายใต้บริบทสถานการณ์เดียวกัน

ประการที่สอง งานวิจัยเกี่ยวกับการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในบริบทการศึกษาไทยที่ใช้รูปแบบเชิงสถานการณ์แบบออนไลน์ และตรวจสอบคุณภาพเครื่องมือภายใต้ทั้ง

กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบยังมีจำนวนจำกัด ทำให้หลักฐานเชิงประจักษ์เกี่ยวกับคุณภาพของเครื่องมือวัดในลักษณะดังกล่าวยังมีอยู่น้อย

ประการที่สาม แม้ว่าการวัดและประเมินผลในรูปแบบออนไลน์จะได้รับการพัฒนาอย่างต่อเนื่อง แต่การรายงานผลการประเมินส่วนใหญ่มีก่นำเสนอคะแนนรวมเป็นหลัก ขณะที่ข้อมูลผลการประเมินรายองค์ประกอบของทักษะสามารถช่วยให้ผู้เรียนและครูเข้าใจจุดเด่นและประเด็นที่ควรพัฒนาได้ชัดเจนยิ่งขึ้น จึงยังมีช่องว่างในการพัฒนารูปแบบการรายงานผลที่สะท้อนข้อมูลการประเมินในระดับรายองค์ประกอบ

ประการที่สี่ จากการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง พบว่างานวิจัยที่รายงานเกณฑ์ปกติสำหรับการแปลผลคะแนนทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้นยังมีจำนวนจำกัด จึงยังมีช่องว่างในการพัฒนาเกณฑ์อ้างอิงสำหรับการตีความผลการประเมิน

จากช่องว่างการวิจัยดังกล่าว การวิจัยครั้งนี้จึงมุ่งพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น พร้อมตรวจสอบคุณภาพของเครื่องมือภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ รวมทั้งสร้างเกณฑ์ปกติสำหรับการแปลผลคะแนน เพื่อเพิ่มหลักฐานเชิงประจักษ์ด้านการวัดและประเมินทักษะการคิดเชิงคำนวณในบริบทการศึกษาไทย

6. กรอบแนวคิดในการวิจัย

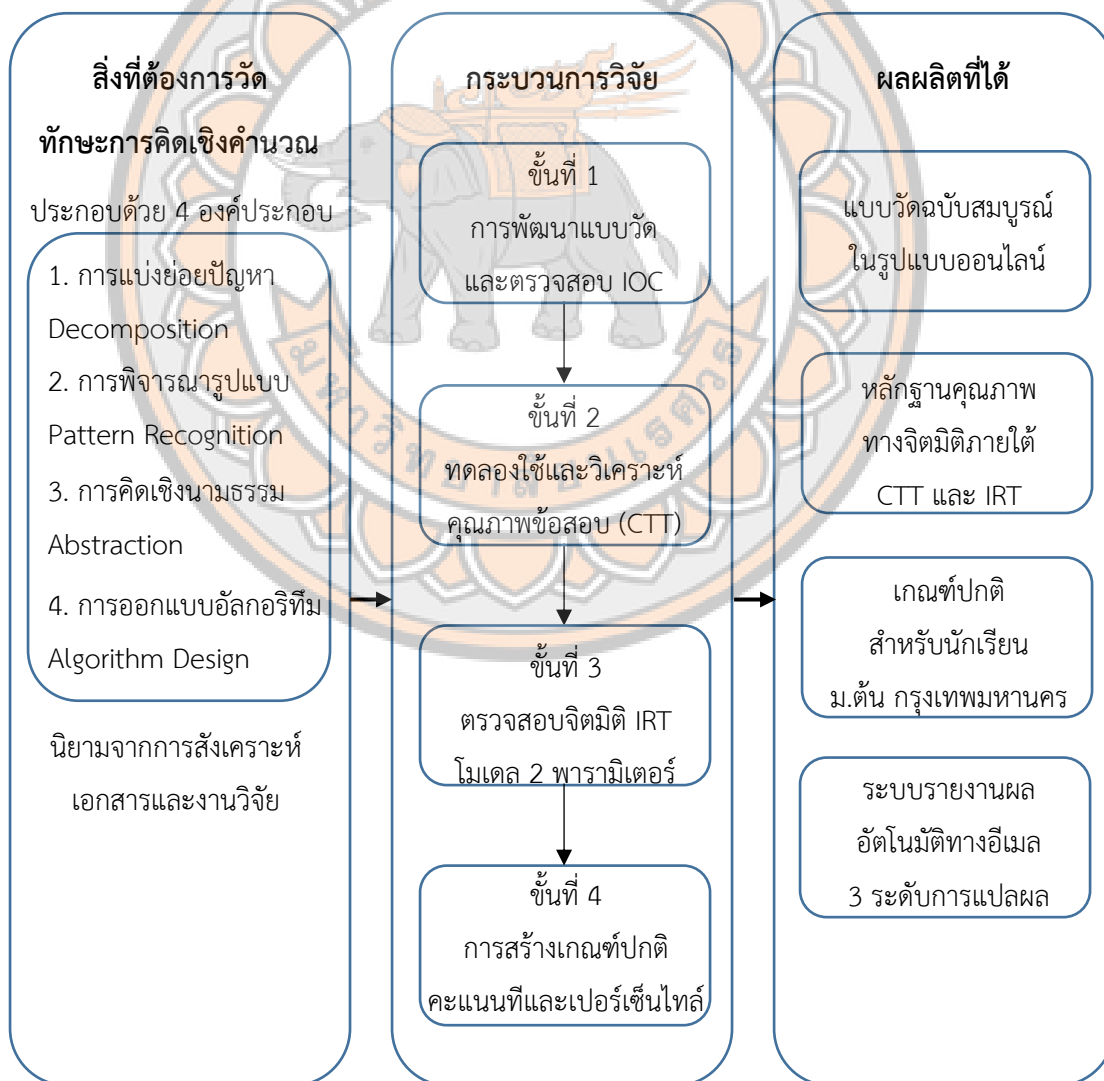
จากการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องในบทที่ 2 ซึ่งครอบคลุมแนวคิดเกี่ยวกับทักษะการคิดเชิงคำนวณ แบบวัดเชิงสถานการณ์ การวัดและประเมินผลในรูปแบบออนไลน์ กรอบแนวคิดทางจิตมิติ และงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้บูรณาการแนวคิดดังกล่าวเพื่อพัฒนาเป็นกรอบแนวคิดในการวิจัย โดยแสดงความเชื่อมโยงระหว่างสิ่งที่ต้องการวัด กระบวนการพัฒนาและตรวจสอบคุณภาพแบบวัด และผลผลิตที่ได้จากการวิจัย

ส่วนที่ 1 สิ่งที่ต้องการวัด ผู้วิจัยกำหนดตัวแปรที่ศึกษา คือ ทักษะการคิดเชิงคำนวณ ซึ่งได้จากการสังเคราะห์แนวคิดและเอกสารที่เกี่ยวข้อง ประกอบด้วย 4 องค์ประกอบ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม องค์ประกอบทั้ง 4 ด้านใช้เป็นกรอบในการกำหนดเนื้อหาและพัฒนาข้อคำถามของแบบวัดให้ครอบคลุมพฤติกรรมที่ต้องการวัด

ส่วนที่ 2 กระบวนการวิจัย เป็นการพัฒนาและตรวจสอบคุณภาพแบบวัดตามลำดับขั้นตอน ประกอบด้วย (1) การพัฒนาแบบวัดฉบับร่างและตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ (2) การทดลองใช้และวิเคราะห์คุณภาพรายข้อตามทฤษฎีการทดสอบแบบดั้งเดิม (3) การตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดตามทฤษฎีการตอบสนองข้อสอบภายใต้โมเดล 2 พารามิเตอร์ และ (4) การสร้างเกณฑ์ปกติในรูปลำดับเปอร์เซ็นต์ไทล์และคะแนนที่ปกติ กระบวนการทั้งหมดดำเนินการ

ภายใต้แนวคิดของแบบวัดเชิงสถานการณ์และกรอบแนวคิดทางจิตมิติที่ใช้เป็นพื้นฐานในการพัฒนาตรวจสอบคุณภาพ และแปลผลคะแนน

ส่วนที่ 3 ผลผลิตที่ได้จากการวิจัย ประกอบด้วย แบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ฉบับสมบูรณ์ หลักฐานคุณภาพทางจิตมิติภายใต้ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ เกณฑ์ปกติสำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร และระบบรายงานผลอัตโนมัติทางอีเมลที่แสดงคะแนนและผลการแปลผลระดับทักษะ ซึ่งใช้เป็นข้อมูลอ้างอิงในการประเมินและตีความผลการวัดทักษะการคิดเชิงคำนวณของนักเรียน ผลผลิตดังกล่าวประกอบด้วยแบบวัดฉบับสมบูรณ์ หลักฐานคุณภาพทางจิตมิติ เกณฑ์ปกติ และระบบรายงานผลอัตโนมัติ ซึ่งใช้เป็นข้อมูลอ้างอิงในการประเมินและแปลผลทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้น ความสัมพันธ์ของส่วนทั้งสามแสดงในภาพ 1



ภาพ 1 กรอบแนวคิดในการวิจัย

จากภาพ 1 กรอบแนวคิดในการวิจัยแสดงลำดับความสัมพันธ์ขององค์ประกอบสำคัญในการดำเนินการวิจัย ตั้งแต่การกำหนดสิ่งที่ต้องการวัด ได้แก่ ทักษะการคิดเชิงคำนวณ 4 องค์ประกอบ การพัฒนาและตรวจสอบคุณภาพแบบวัดตามกรอบแนวคิดทางจิตมิติ ตลอดจนการสร้างเกณฑ์ปกติสำหรับการแปลผลคะแนน ซึ่งกระบวนการดังกล่าวนำไปสู่ผลผลิตที่ได้จากการวิจัย ได้แก่ แบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ที่มีหลักฐานคุณภาพทางจิตมิติ เกณฑ์ปกติสำหรับการแปลผล และระบบรายงานผลอัตโนมัติ เพื่อใช้เป็นเครื่องมือในการประเมินทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้นได้อย่างมีประสิทธิภาพ



บทที่ 3

วิธีดำเนินงานวิจัย

การวิจัยครั้งนี้เป็นการวิจัยและพัฒนาเครื่องมือวัด มีวัตถุประสงค์ 3 ประการ คือ (1) เพื่อพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น (2) เพื่อตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด ภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ และ (3) เพื่อสร้างเกณฑ์ปกติของแบบวัดสำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร ผู้วิจัยนำเสนอวิธีการดำเนินการวิจัยตามลำดับวัตถุประสงค์ทั้ง 3 ข้อ โดยภาพรวมของวิธีดำเนินการแสดงในตาราง 9

ตาราง 9 ภาพรวมวิธีดำเนินการวิจัยจำแนกตามวัตถุประสงค์

วัตถุประสงค์	แหล่งข้อมูล/ กลุ่มตัวอย่าง	เครื่องมือ	วิธีเก็บข้อมูล	การวิเคราะห์
ข้อ 1 พัฒนา แบบวัด	ผู้เชี่ยวชาญ 7 คน และนักเรียน ม.ต้น 150 คน (ทดลองใช้)	แบบวัดฉบับ ร่าง 80 ข้อ และแบบ ประเมิน ความตรง เชิงเนื้อหา	ส่งแบบประเมิน ให้ผู้เชี่ยวชาญ และทดลองใช้ แบบวัดออนไลน์	ค่าดัชนีความสอดคล้อง (IOC) และการวิเคราะห์คุณภาพราย ข้อตามทฤษฎีการทดสอบแบบ ดั้งเดิม
ข้อ 2 ตรวจสอบ คุณสมบัติ ทางจิตมิติ	นักเรียน ม.ต้น ในกรุงเทพ มหานคร 1,265 คน	แบบวัดฉบับ สมบูรณ์ 40 ข้อ	เก็บข้อมูลด้วย แบบวัดออนไลน์	การวิเคราะห์ความตรง ความ เชื่อมั่น และพารามิเตอร์ ข้อสอบตามทฤษฎีการทดสอบ แบบดั้งเดิมและทฤษฎีการ ตอบสนองข้อสอบ
ข้อ 3 สร้างเกณฑ์ ปกติ	นักเรียนกลุ่ม เดียวกับข้อ 2	แบบวัดฉบับ สมบูรณ์ 40 ข้อ	ใช้ข้อมูลคะแนน จากการเก็บ ข้อมูลในข้อ 2	การสร้างเกณฑ์ปกติในรูปแบบ คะแนนที่ปกติ

จากตาราง 9 แสดงให้เห็นว่า วิธีดำเนินการวิจัยแบ่งออกเป็น 3 ส่วนตามวัตถุประสงค์การวิจัยที่มีความเชื่อมโยงต่อเนื่องกัน โดยวัตถุประสงค์ข้อที่ 1 เป็นการพัฒนาแบบวัดและตรวจสอบคุณภาพเบื้องต้น เพื่อให้ได้แบบวัดทักษะการคิดเชิงคำนวณฉบับสมบูรณ์ที่มีความเหมาะสมสำหรับนำไปใช้เก็บข้อมูล จากนั้นจึงนำแบบวัดดังกล่าวไปใช้ดำเนินการตามวัตถุประสงค์ข้อที่ 2 เพื่อศึกษาคุณสมบัติทางจิตมิติของแบบวัดภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ และวัตถุประสงค์ข้อที่ 3 เพื่อสร้างเกณฑ์ปกติสำหรับใช้ในการแปลผลคะแนนของผู้สอบ โดยใช้ข้อมูลจากกลุ่มตัวอย่างเดียวกัน รายละเอียดวิธีดำเนินการวิจัยในแต่ละวัตถุประสงค์มีดังต่อไปนี้

ตอนที่ 1 การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์

1.1 การกำหนดกรอบเนื้อหาและโครงสร้างของแบบวัด

ผู้วิจัยศึกษาและสังเคราะห์แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องกับทักษะการคิดเชิงคำนวณ แบบวัดเชิงสถานการณ์ การวัดและประเมินผลในรูปแบบออนไลน์ และการวัดทางจิตมิติจากเอกสารและงานวิจัยทั้งในและต่างประเทศที่เกี่ยวข้อง ซึ่งได้นำเสนอรายละเอียดไว้ในบทที่ 2 เพื่อใช้เป็นพื้นฐานในการกำหนดกรอบเนื้อหาและโครงสร้างของแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์

จากการสังเคราะห์เอกสารและงานวิจัย ผู้วิจัยกำหนดกรอบเนื้อหาของแบบวัดโดยใช้แนวคิดทักษะการคิดเชิงคำนวณตามกรอบของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท., 2560) ซึ่งสอดคล้องกับแนวคิดของ Grover และ Pea (2013) ประกอบด้วย 4 องค์ประกอบ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม โดยใช้องค์ประกอบทั้ง 4 ด้านเป็นกรอบในการกำหนดพฤติกรรมที่ต้องการวัดและสร้างข้อคำถามของแบบวัด

สำหรับรูปแบบของเครื่องมือ ผู้วิจัยพัฒนาแบบวัดในลักษณะเชิงสถานการณ์ตามแนวคิดของ Weekley และ Ployhart (2006) และ Wiggins (1990) โดยนำเสนอสถานการณ์ที่ใกล้เคียงกับบริบทในชีวิตประจำวันของนักเรียนระดับมัธยมศึกษาตอนต้น เพื่อเปิดโอกาสให้ผู้สอบประยุกต์ใช้ทักษะการคิดเชิงคำนวณในการวิเคราะห์และตัดสินใจภายใต้สถานการณ์ที่กำหนด นอกจากนี้ แบบวัดถูกพัฒนาในรูปแบบออนไลน์ตามแนวคิดของ Alruwais, Wills และ Wald (2018) ซึ่งเอื้อต่อการเก็บรวบรวมข้อมูล การตรวจให้คะแนนอัตโนมัติ และการรายงานผลการประเมินอย่างมีประสิทธิภาพ

ในด้านการวัดและวิเคราะห์ข้อมูล ผู้วิจัยกำหนดให้ข้อคำถามเป็นแบบเลือกตอบที่ให้คะแนนแบบทวิภาค โดยใช้กรอบการวิเคราะห์คุณภาพเครื่องมือทั้งภายใต้ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ ตามแนวทางของ Crocker และ Algina (2008) และ

Embretson และ Reise (2000) เพื่อให้ได้หลักฐานคุณภาพของแบบวัดที่ครอบคลุมทั้งระดับข้อสอบ และระดับแบบวัด

จากกรอบแนวคิดดังกล่าว ผู้วิจัยกำหนดโครงสร้างของแบบวัดให้ 1 สถานการณ์ ประกอบด้วยข้อคำถามจำนวน 4 ข้อ โดยแต่ละข้อวัดองค์ประกอบของทักษะการคิดเชิงคำนวณ จำนวน 1 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม ตามลำดับ ส่งผลให้แต่ละสถานการณ์สามารถวัดทักษะการคิดเชิงคำนวณได้ครบ ทั้ง 4 องค์ประกอบภายใต้บริบทเดียวกัน โครงสร้างดังกล่าวถูกใช้เป็นกรอบในการพัฒนาแบบวัด ฉบับร่างในขั้นตอนต่อไป

ผู้วิจัยสร้างแบบวัดทักษะการคิดเชิงคำนวณฉบับร่าง โดยกำหนดจำนวนข้อคำถามเป็น 2 เท่าของจำนวนข้อที่ต้องการในแบบวัดฉบับสมบูรณ์ เพื่อให้มีข้อคำถามสำรองเพียงพอสำหรับการคัดเลือกข้อที่มีคุณภาพในขั้นตอนการวิเคราะห์ข้อสอบ กล่าวคือ กำหนดเป้าหมายแบบวัดฉบับสมบูรณ์จำนวน 40 ข้อ จึงสร้างข้อคำถามฉบับร่างจำนวน 80 ข้อ จัดเป็น 20 สถานการณ์ โดยแต่ละสถานการณ์ประกอบด้วยข้อคำถามจำนวน 4 ข้อ ซึ่งวัดองค์ประกอบของทักษะการคิดเชิงคำนวณทั้ง 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม

ผู้วิจัยกำหนดสถานการณ์จากบริบทที่นักเรียนระดับมัธยมศึกษาตอนต้นคุ้นเคย และสามารถพบได้ในชีวิตประจำวัน โดยแบ่งบริบทของสถานการณ์ออกเป็น 3 ด้าน ได้แก่ (1) ด้านชีวิตส่วนบุคคล (2) ด้านการจัดกิจกรรมในโรงเรียน และ (3) ด้านชุมชน สังคม และสิ่งแวดล้อม เพื่อให้สถานการณ์มีความหลากหลายและครอบคลุมประสบการณ์ของผู้เรียนในบริบทต่าง ๆ สถานการณ์ทั้งหมดนำเสนอในรูปแบบข้อความบรรยายที่เหมาะสมกับระดับการอ่านและประสบการณ์ของนักเรียนระดับมัธยมศึกษาตอนต้น

ข้อคำถามทุกข้อเป็นแบบปรนัยชนิดเลือกตอบ 4 ตัวเลือก โดยมีคำตอบที่ถูกต้องเพียง 1 ตัวเลือก และสร้างตัวลวงให้มีความสมเหตุสมผล สอดคล้องกับความคลาดเคลื่อนหรือความเข้าใจที่อาจเกิดขึ้นในการแก้ปัญหา ทั้งนี้ ผู้วิจัยดำเนินการสร้างข้อคำถามและตัวเลือกตามหลักการสร้างข้อสอบปรนัยของ Haladyna และ Rodriguez (2013) เพื่อให้ข้อคำถามมีคุณภาพและเหมาะสมสำหรับการนำไปตรวจสอบคุณภาพในขั้นตอนต่อไป

ภายหลังการสร้างข้อคำถาม ผู้วิจัยจัดทำตารางวิเคราะห์ข้อสอบ (Table of Specifications) เพื่อแสดงความสอดคล้องระหว่างองค์ประกอบของทักษะการคิดเชิงคำนวณ สถานการณ์ที่ใช้วัด และจำนวนข้อคำถามในแต่ละองค์ประกอบ โดยใช้เป็นเครื่องมือควบคุมความครอบคลุมของเนื้อหาและตรวจสอบความสมดุลของแบบวัดฉบับร่าง ก่อนนำไปตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ ดังแสดงในตาราง 10

ตาราง 10 โครงสร้างแบบวัดฉบับร่าง จำนวน 80 ข้อ

สถานการณ์ที่	การแบ่งย่อย ปัญหา	การพิจารณา รูปแบบ	การคิดเชิง นามธรรม	การออกแบบ อัลกอริทึม	รวม
1	1	1	1	1	4
2	1	1	1	1	4
3	1	1	1	1	4
4	1	1	1	1	4
5	1	1	1	1	4
6	1	1	1	1	4
7	1	1	1	1	4
8	1	1	1	1	4
9	1	1	1	1	4
10	1	1	1	1	4
11	1	1	1	1	4
12	1	1	1	1	4
13	1	1	1	1	4
14	1	1	1	1	4
15	1	1	1	1	4
16	1	1	1	1	4
17	1	1	1	1	4
18	1	1	1	1	4
19	1	1	1	1	4
20	1	1	1	1	4
รวม	20	20	20	20	80

หมายเหตุ: ข้อคำถามแต่ละสถานการณ์วัดองค์ประกอบของทักษะการคิดเชิงคำนวณครบทั้ง 4 ด้าน
ด้านละ 1 ข้อ

1.3 การตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ

ผู้วิจัยตรวจสอบความตรงเชิงเนื้อหาของแบบวัดฉบับร่าง โดยการส่งแบบวัดให้ผู้เชี่ยวชาญพิจารณาความสอดคล้องระหว่างข้อสอบกับองค์ประกอบของทักษะการคิดเชิงคำนวณที่ต้องการวัด มีรายละเอียดดังนี้

1.3.1 การคัดเลือกผู้เชี่ยวชาญ

ผู้วิจัยคัดเลือกผู้เชี่ยวชาญจำนวน 7 คน ด้วยวิธีการเลือกแบบเจาะจง (Purposive Sampling) ตามคุณสมบัติและความเชี่ยวชาญที่เกี่ยวข้องกับการพัฒนาแบบวัด ประกอบด้วย ผู้เชี่ยวชาญด้านการคิดเชิงคำนวณ จำนวน 3 คน ซึ่งมีประสบการณ์ด้านการสอน การวิจัย หรือการพัฒนาหลักสูตรที่เกี่ยวข้องกับการคิดเชิงคำนวณ และผู้เชี่ยวชาญด้านการวัดและประเมินผลทางการศึกษา จำนวน 4 คน ซึ่งมีประสบการณ์ด้านการพัฒนาเครื่องมือวัด การวิเคราะห์คุณภาพข้อสอบ หรือจิตมิติทางการศึกษา

การกำหนดจำนวนผู้เชี่ยวชาญ 7 คน สอดคล้องกับข้อเสนอของ Lynn (1986) ที่ระบุว่า การตรวจสอบความตรงเชิงเนื้อหาควรใช้ผู้เชี่ยวชาญอย่างน้อย 5 คนขึ้นไป เพื่อให้การตัดสินใจมีความน่าเชื่อถือและลดความคลาดเคลื่อนจากความคิดเห็นรายบุคคล นอกจากนี้ การคัดเลือกผู้เชี่ยวชาญที่มีความเชี่ยวชาญทั้งด้านเนื้อหาที่ต้องการวัดและด้านการวัดและประเมินผล ช่วยให้การพิจารณาความสอดคล้องของข้อคำถามครอบคลุมทั้งความถูกต้องของเนื้อหา ความเหมาะสมของพฤติกรรมที่ต้องการวัด และความเหมาะสมของรูปแบบข้อคำถาม อันส่งผลให้การตรวจสอบความตรงเชิงเนื้อหา มีความรอบด้านและน่าเชื่อถือยิ่งขึ้น

1.3.2 เครื่องมือและวิธีการเก็บข้อมูลจากผู้เชี่ยวชาญ

ผู้วิจัยจัดทำแบบประเมินความตรงเชิงเนื้อหา (Content Validity Assessment Form) สำหรับใช้ในการพิจารณาความสอดคล้องระหว่างข้อคำถามกับองค์ประกอบของทักษะการคิดเชิงคำนวณที่ต้องการวัด โดยแบบประเมินประกอบด้วย (1) คำชี้แจงและวัตถุประสงค์ของการประเมิน (2) นิยามเชิงปฏิบัติการของทักษะการคิดเชิงคำนวณทั้ง 4 องค์ประกอบ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม (3) แบบวัดฉบับร่างจำนวน 80 ข้อ พร้อมระบุองค์ประกอบที่มุ่งวัดในแต่ละข้อ (4) ตารางสำหรับประเมินความสอดคล้องของข้อคำถาม และ (5) ช่องสำหรับบันทึกข้อเสนอแนะเพิ่มเติมของผู้เชี่ยวชาญ

การประเมินความสอดคล้องของข้อคำถามใช้เกณฑ์การให้คะแนนตามแนวทางของ Rovinelli และ Hambleton (1977) ดังนี้ +1 หมายถึง ข้อคำถามสอดคล้องกับองค์ประกอบที่ต้องการวัด 0 หมายถึง ไม่แน่ใจว่าข้อคำถามสอดคล้องกับองค์ประกอบที่ต้องการวัด และ -1 หมายถึง ข้อคำถามไม่สอดคล้องกับองค์ประกอบที่ต้องการวัด

1.3.3 การคำนวณค่าดัชนีความสอดคล้องและเกณฑ์การยอมรับ

ภายหลังได้รับผลการประเมินจากผู้เชี่ยวชาญครบทั้ง 7 คน ผู้วิจัยคำนวณค่าดัชนีความสอดคล้อง (Index of Item-Objective Congruence: IOC) ของข้อสอบแต่ละข้อ ตามสูตรของ Rovinelli และ Hambleton (1977) ดังนี้

$$IOC = \Sigma R / N$$

เมื่อ ΣR คือ ผลรวมของคะแนนความสอดคล้องที่ผู้เชี่ยวชาญทุกคนให้ และ N คือ จำนวนผู้เชี่ยวชาญ ในการวิจัยครั้งนี้ N เท่ากับ 7 ผู้วิจัยกำหนดเกณฑ์การยอมรับไว้ที่ค่าดัชนีความสอดคล้องตั้งแต่ 0.50 ขึ้นไป ตามที่ Rovinelli และ Hambleton (1977) เสนอ ซึ่งเป็นเกณฑ์ที่นิยมใช้ในงานวิจัยทางการศึกษา ข้อคำถามที่มีค่าดัชนีความสอดคล้องต่ำกว่า 0.50 แสดงว่าผู้เชี่ยวชาญส่วนใหญ่เห็นว่าข้อคำถามไม่สอดคล้องกับองค์ประกอบที่ต้องการวัด และต้องได้รับการปรับปรุงตามข้อเสนอแนะของผู้เชี่ยวชาญหรือคัดออกจากแบบวัด

นอกจากการคำนวณค่าดัชนีความสอดคล้องแล้ว ผู้วิจัยยังรวบรวมข้อเสนอแนะเชิงคุณภาพจากผู้เชี่ยวชาญ เพื่อนำมาปรับปรุงข้อคำถามใน 2 ลักษณะ คือ การปรับปรุงเชิงภาษาและความชัดเจน เช่น การปรับถ้อยคำให้เข้าใจง่ายขึ้น และการปรับปรุงเชิงเทคนิคของข้อคำถาม เช่น การปรับตัวลงให้สมเหตุสมผลมากขึ้นและการตรวจสอบความถูกต้องของคำตอบ การพิจารณาทั้งข้อมูลเชิงปริมาณและเชิงคุณภาพประกอบกันช่วยให้ได้แบบวัดที่มีคุณภาพสมบูรณ์ทั้งในเชิงโครงสร้างและเชิงเนื้อหา

1.4 การพัฒนาแบบวัดในรูปแบบออนไลน์

ภายหลังปรับปรุงแบบวัดฉบับร่างตามข้อเสนอแนะของผู้เชี่ยวชาญแล้ว ผู้วิจัยพัฒนาแบบวัดให้อยู่ในรูปแบบออนไลน์ โดยใช้แบบฟอร์มออนไลน์ของกูเกิล (Google Forms) เป็นเครื่องมือหลักในการนำเสนอสถานการณ์และเก็บคำตอบ ร่วมกับเว็บไซต์ที่สร้างด้วยกูเกิลไซต์ (Google Sites) เป็นศูนย์กลางข้อมูลสำหรับผู้สอบ และระบบรายงานผลอัตโนมัติที่พัฒนาด้วยกูเกิลแอปสคริปต์ (Google Apps Script) มีรายละเอียดดังนี้

1.4.1 การพัฒนาแบบวัดออนไลน์ด้วยแบบฟอร์มออนไลน์

ผู้วิจัยพัฒนาแบบวัดออนไลน์โดยออกแบบให้มียอดประกอบสำคัญ ดังนี้

ด้านระบบการยืนยันตัวตน ผู้สอบต้องลงชื่อเข้าใช้ด้วยบัญชีของกูเกิล และกรอกข้อมูลที่จำเป็นก่อนทำแบบวัด ได้แก่ ชื่อ-นามสกุล ระดับชั้น โรงเรียน สังกัด และที่อยู่อีเมล โดยระบบกำหนดให้ทำแบบวัดได้เพียงครั้งเดียวต่อหนึ่งบัญชี เพื่อป้องกันการทำซ้ำและรักษาความเที่ยงตรงของข้อมูล

ด้านการนำเสนอสถานการณ์ แบบวัดออนไลน์แสดงสถานการณ์ในรูปข้อความ พร้อมข้อความย่อย 4 ข้อต่อหนึ่งสถานการณ์ โดยนำเสนอสถานการณ์ไว้ที่ส่วนบนของหน้า และให้ผู้สอบเลือกคำตอบของข้อความทั้ง 4 ข้อในหน้าเดียวกัน เพื่อให้อ่านสถานการณ์ประกอบการตอบได้อย่างต่อเนื่อง

ด้านการกำหนดเวลา ผู้วิจัยกำหนดเวลาในการทำแบบวัดประมาณ 50 นาที ซึ่งเฉลี่ยประมาณ 1 นาทีต่อข้อสำหรับแบบวัดฉบับสมบูรณ์ 40 ข้อ ระยะเวลาที่สอดคล้องกับการเป็นแบบทดสอบที่ใช้ในชั้นเรียน

ด้านระบบรายงานผลอัตโนมัติ ภายหลังการทำแบบวัดเสร็จสิ้น ระบบจะประมวลผลและส่งผลคะแนนไปยังอีเมลของผู้สอบโดยอัตโนมัติ ทั้งคะแนนรวมและคะแนนรายองค์ประกอบทั้ง 4 ด้าน

ด้านการเก็บรวบรวมข้อมูล ระบบบันทึกคำตอบของผู้สอบทุกคนลงในตารางคำนวณออนไลน์ (Google Sheets) โดยอัตโนมัติ พร้อมบันทึกข้อมูลกำกับ ได้แก่ วันและเวลาที่ทำแบบวัด และที่อยู่อีเมลของผู้สอบ

1.4.2 การพัฒนาเว็บไซต์ศูนย์รวมข้อมูล

ผู้วิจัยพัฒนาเว็บไซต์ด้วยยูเกิลไซต์ เพื่อทำหน้าที่เป็นศูนย์กลางข้อมูลสำหรับผู้สอบและผู้เกี่ยวข้อง ใช้เป็นจุดเริ่มต้นในการเข้าถึงแบบวัด เว็บไซต์ประกอบด้วยส่วนสำคัญ ดังนี้ ส่วนที่หนึ่ง หน้าแนะนำการวิจัยและวัตถุประสงค์ ซึ่งนำเสนอข้อมูลพื้นฐานเกี่ยวกับการวิจัย ที่มาของแบบวัด และประโยชน์ที่ผู้เข้าร่วมจะได้รับ ส่วนที่สอง คำชี้แจงและคู่มือการทำแบบวัด ซึ่งอธิบายรูปแบบของแบบวัด จำนวนข้อ วิธีการตอบ เวลาที่ใช้ และระบบการรายงานผล รวมทั้งคำชี้แจงสำหรับครูและผู้ปกครองเกี่ยวกับการนำผลการประเมินไปใช้และการคุ้มครองข้อมูลส่วนบุคคล ส่วนที่สาม ลิงก์เข้าสู่แบบวัดออนไลน์ และส่วนที่สี่ ข้อมูลติดต่อของผู้วิจัย เพื่อให้ผู้สอบหรือครูสามารถสอบถามข้อมูลเพิ่มเติมได้

การมีเว็บไซต์ศูนย์รวมข้อมูลก่อนเข้าทำแบบวัด มีประโยชน์สำคัญในการเตรียมความพร้อมของผู้สอบและสร้างความเข้าใจร่วมกันระหว่างผู้สอบ ครู และผู้วิจัย ทำให้ผู้สอบสามารถเข้าทำแบบวัดได้อย่างมั่นใจและให้ข้อมูลที่สะท้อนความสามารถของตนเองอย่างแท้จริง การพัฒนาเว็บไซต์ร่วมกับแบบฟอร์มออนไลน์ ทำให้ระบบการทดสอบมีลักษณะเป็นระบบบูรณาการที่ครอบคลุมทั้งการให้ข้อมูล การรับคำตอบ และการรายงานผลในแพลตฟอร์มเดียวกัน

1.4.3 การพัฒนาระบบรายงานผลอัตโนมัติ

ผู้วิจัยพัฒนาระบบรายงานผลอัตโนมัติด้วยยูเกิลแอปสคริปต์ ซึ่งเป็นภาษาเขียนโปรแกรมที่อิงกับภาษาจาวาสคริปต์และทำงานบนแพลตฟอร์มของยูเกิล ระบบนี้ทำงานเชื่อมโยงกับแบบฟอร์มออนไลน์ เมื่อผู้สอบส่งคำตอบเสร็จสิ้น ระบบจะดำเนินการอัตโนมัติตามขั้นตอน ดังนี้ ขั้นที่หนึ่ง

ตรวจให้คะแนนคำตอบแต่ละข้อ โดยเปรียบเทียบกับคำตอบที่ถูกต้องที่กำหนดไว้ ให้คะแนน 1 เมื่อตอบถูก และให้คะแนน 0 เมื่อตอบผิด ขั้นที่สอง คำนวณคะแนนรวมของผู้สอบจากผลรวมของคะแนนทุกข้อ และคำนวณคะแนนรายองค์ประกอบจากผลรวมของคะแนนข้อสอบที่วัดแต่ละองค์ประกอบ ขั้นที่สาม แปลผลระดับทักษะของผู้สอบ และจัดทำข้อความรายงานผลที่ครอบคลุมคะแนนรวม คะแนนรายองค์ประกอบ และระดับทักษะ พร้อมคำอธิบายในภาษาที่เหมาะสมกับนักเรียนระดับมัธยมศึกษาตอนต้น และขั้นที่สี่ ส่งรายงานผลทางอีเมลให้ผู้สอบที่อีเมลที่กรอกไว้ในแบบวัด

1.4.4 การทดสอบระบบ

ผู้วิจัยทดสอบการทำงานของระบบทุกส่วนก่อนนำไปใช้กับกลุ่มทดลองใช้เครื่องมือ โดยทดสอบการเข้าสู่เว็บไซต์ การกรอกแบบวัด การส่งคำตอบ การตรวจให้คะแนน การคำนวณคะแนน และการส่งรายงานผลทางอีเมล เพื่อให้มั่นใจว่าระบบทำงานได้อย่างถูกต้องและสมบูรณ์ก่อนการเก็บข้อมูลจริง

1.5 การทดลองใช้แบบวัดและการวิเคราะห์คุณภาพรายข้อ

ผู้วิจัยนำแบบวัดฉบับร่าง 80 ข้อ ในรูปแบบออนไลน์ ไปทดลองใช้กับกลุ่มทดลองใช้เครื่องมือ เพื่อวิเคราะห์คุณภาพรายข้อตามทฤษฎีการทดสอบแบบดั้งเดิม และคัดเลือกข้อคำถามที่มีคุณภาพ มีรายละเอียดดังนี้

1.5.1 กลุ่มตัวอย่างในการทดลองใช้แบบวัด

กลุ่มตัวอย่างในการทดลองใช้แบบวัด คือ นักเรียนระดับมัธยมศึกษาตอนต้น จำนวน 150 คน จากโรงเรียนสาธิตในสังกัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม ในกรุงเทพมหานคร ผู้วิจัยเลือกกลุ่มตัวอย่างด้วยวิธีการเลือกแบบเจาะจง จำนวน 150 คน เป็นจำนวนที่เพียงพอต่อการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิม ซึ่ง ศิริชัย กาญจนวาสี (2556) ระบุว่า การทดลองใช้เครื่องมือควรมีก่อนมีกลุ่มตัวอย่างอย่างน้อย 100 ถึง 200 คน เนื่องจากการทดลองใช้ในขั้นตอนนี้มีวัตถุประสงค์เพื่อตรวจสอบคุณภาพของเครื่องมือเป็นหลัก มิใช่การสรุปผลเพื่ออ้างอิงถึงประชากร

กลุ่มตัวอย่างในขั้นตอนนี้ถูกแยกออกจากกลุ่มเก็บข้อมูลหลักในตอนที่ 2 อย่างสิ้นเชิง แม้จะมาจากโรงเรียนในสังกัดเดียวกัน เพื่อป้องกันการปนเปื้อนของข้อมูลที่อาจเกิดขึ้นหากผู้สอบเคยทำแบบวัดมาก่อน

1.5.2 การเก็บข้อมูลและการคัดกรองข้อมูล

ผู้วิจัยจัดทำหนังสือขอความอนุเคราะห์เก็บข้อมูลจากสถานศึกษา พร้อมแนบรายละเอียดของแบบวัด เมื่อสถานศึกษาให้ความยินยอม ผู้วิจัยเปิดระบบแบบวัดออนไลน์ให้นักเรียนเข้าทำแบบวัด ภายใต้การกำกับดูแลของครูผู้สอน โดยข้อมูลการตอบถูกบันทึกอัตโนมัติในระบบ จากนั้นผู้วิจัยคัดกรองข้อมูลเบื้องต้น โดยตรวจสอบความครบถ้วนของการตอบ และตัดข้อมูลของผู้ที่ตอบ

ไม่ครบทั้ง 80 ข้อออกจากการวิเคราะห์ ข้อมูลที่ผ่านการคัดกรองถูกนำเข้าสู่โปรแกรมภาษา R เวอร์ชัน

4.5.3 ผ่านเว็บแอปพลิเคชัน Posit Cloud เพื่อวิเคราะห์คุณภาพรายข้อ

1.5.3 การวิเคราะห์คุณภาพรายข้อตามทฤษฎีการทดสอบแบบดั้งเดิม

ผู้วิจัยวิเคราะห์คุณภาพรายข้อของข้อสอบทั้ง 80 ข้อ ตามทฤษฎีการทดสอบแบบดั้งเดิม โดยพิจารณาดัชนี 3 ประการ ดังนี้

ดัชนีที่ 1 ค่าความยากของข้อสอบ สัญลักษณ์ p ผู้วิจัยคำนวณด้วยวิธีกลุ่มสูง-กลุ่มต่ำ ร้อยละ 27 ตามแนวคิดของ Kelley (1939) โดยแบ่งกลุ่มตัวอย่างเป็นกลุ่มสูงและกลุ่มต่ำ กลุ่มละ ร้อยละ 27 ของจำนวนทั้งหมด ตามสูตร

$$p = (RU + RL) / 2N$$

เมื่อ RU คือ จำนวนผู้ตอบถูกในกลุ่มสูง RL คือ จำนวนผู้ตอบถูกในกลุ่มต่ำ และ N คือ จำนวนผู้สอบในแต่ละกลุ่ม ผู้วิจัยกำหนดเกณฑ์การยอมรับไว้ที่ค่าความยากระหว่าง 0.20 ถึง 0.80 ตามเกณฑ์ของ Crocker และ Algina (2008) เนื่องจากข้อสอบที่ง่ายเกินไปหรือยากเกินไปไม่สามารถจำแนกความสามารถของผู้สอบได้

ดัชนีที่ 2 ค่าอำนาจจำแนกของข้อสอบ สัญลักษณ์ r ผู้วิจัยคำนวณตามสูตร

$$r = (RU - RL) / N$$

เมื่อ RU คือ จำนวนผู้ตอบถูกในกลุ่มสูง RL คือ จำนวนผู้ตอบถูกในกลุ่มต่ำ และ N คือ จำนวนผู้สอบในแต่ละกลุ่ม ผู้วิจัยกำหนดเกณฑ์การยอมรับไว้ที่ค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป ตามเกณฑ์ของ Ebel และ Frisbie (1991) ซึ่งระบุว่าข้อสอบที่มีค่าอำนาจจำแนกต่ำกว่า 0.20 ไม่สามารถจำแนกผู้สอบที่มีความสามารถต่างกันได้

ดัชนีที่ 3 ค่าความเชื่อมั่นของแบบวัดทั้งฉบับ ผู้วิจัยคำนวณด้วยสูตร Kuder-Richardson Formula 20 (KR-20) ของ Kuder และ Richardson (1937) ซึ่งเหมาะสมกับข้อสอบที่ให้คะแนนแบบทวิภาค ตามสูตร

$$KR-20 = [k / (k-1)] \times [1 - (\sum p_i q_i / \sigma_x^2)]$$

เมื่อ k คือ จำนวนข้อสอบ p_i คือ สัดส่วนของผู้ตอบถูกในข้อที่ i q_i คือ สัดส่วนของผู้ตอบผิดในข้อที่ i (q_i เท่ากับ $1 - p_i$) $\sum p_i q_i$ คือ ผลรวมของผลคูณ p_i และ q_i ของทุกข้อ และ σ_x^2 คือ ความแปรปรวนของคะแนนรวม

1.6 การคัดเลือกข้อคำถามเข้าสู่แบบวัดฉบับสมบูรณ์

ภายหลังการวิเคราะห์คุณภาพรายข้อ ผู้วิจัยคัดเลือกข้อคำถามที่มีคุณภาพจากแบบวัดฉบับร่าง 80 ข้อ เข้าสู่แบบวัดฉบับสมบูรณ์ 40 ข้อ โดยดำเนินการเป็น 2 ขั้นตอน ดังนี้

ขั้นที่ 1 การคัดเลือกตามเกณฑ์คุณภาพเชิงสถิติ ผู้วิจัยคัดเลือกข้อคำถามที่ผ่านเกณฑ์คุณภาพ 3 ประการพร้อมกัน คือ ค่าดัชนีความสอดคล้องตั้งแต่ 0.50 ขึ้นไป ค่าความยากระหว่าง 0.20 ถึง 0.80 และค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป จากแบบวัดฉบับร่าง 80 ข้อ มีข้อคำถามที่ผ่านเกณฑ์คุณภาพเชิงสถิติ จำนวน 50 ข้อ

ขั้นที่ 2 การคัดเลือกตามเกณฑ์ความสมดุลของโครงสร้างแบบวัด เนื่องจากแบบวัดฉบับสมบูรณ์ต้องคงโครงสร้างหนึ่งสถานการณ์ประกอบด้วยข้อคำถามครบทั้ง 4 องค์ประกอบ ผู้วิจัยจึงคัดเลือกจากข้อคำถาม 50 ข้อที่ผ่านเกณฑ์ในขั้นที่ 1 โดยเลือกเฉพาะสถานการณ์ที่มีข้อคำถามผ่านเกณฑ์ครบทั้ง 4 องค์ประกอบ ผลการคัดเลือกได้สถานการณ์ที่มีข้อคำถามครบทั้ง 4 องค์ประกอบ จำนวน 10 สถานการณ์ รวมเป็นแบบวัดฉบับสมบูรณ์ จำนวน 40 ข้อ ที่วัดองค์ประกอบของทักษะการคิดเชิงคำนวณครบทั้ง 4 ด้าน ด้านละ 10 ข้อ

ผู้วิจัยสรุปกระบวนการพัฒนาแบบวัดในตอนต้นที่ 1 ดังแสดงในตาราง 11

ตาราง 11 สรุปกระบวนการพัฒนาแบบวัดและจำนวนข้อคำถามในแต่ละขั้น

ขั้นตอน	การดำเนินการ	ผลที่ได้
1	สร้างแบบวัดฉบับร่าง (2 เท่าของเป้าหมาย)	80 ข้อ ใน 20 สถานการณ์
2	ตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ 7 คน	ข้อคำถามผ่านเกณฑ์ IOC ทุกข้อ
3	ทดลองใช้และวิเคราะห์คุณภาพรายข้อตามทฤษฎีการทดสอบแบบดั้งเดิม	ผ่านเกณฑ์เชิงสถิติ 50 ข้อ
4	คัดเลือกให้ครบ 4 องค์ประกอบในแต่ละสถานการณ์	แบบวัดฉบับสมบูรณ์ 40 ข้อ ใน 10 สถานการณ์

หมายเหตุ: แบบวัดฉบับสมบูรณ์ 40 ข้อ ใช้เป็นเครื่องมือในการดำเนินการตามวัตถุประสงค์ข้อ 2 และข้อ 3

ตอนที่ 2 การตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด

การดำเนินการตามวัตถุประสงค์ข้อ 2 มุ่งตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดฉบับสมบูรณ์ 40 ข้อ ภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ มีรายละเอียดดังนี้

2.1 กลุ่มตัวอย่างในการเก็บข้อมูล

กลุ่มตัวอย่างในการเก็บข้อมูลเพื่อตรวจสอบคุณสมบัติทางจิตมิติ คือ นักเรียนระดับมัธยมศึกษาตอนต้น ชั้นมัธยมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 3 ในกรุงเทพมหานคร จำนวน 1,265 คน จากโรงเรียน 3 สังกัด ได้แก่ สังกัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม จำนวน 975 คน สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษากรุงเทพมหานคร จำนวน 270 คน และสังกัดสำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน จำนวน 20 คน

ผู้วิจัยเลือกกลุ่มตัวอย่างด้วยวิธีการเลือกแบบเจาะจง โดยพิจารณาความพร้อมและความร่วมมือของโรงเรียนในการให้นักเรียนเข้าร่วมการวิจัย ร่วมกับการพิจารณาให้ครอบคลุมโรงเรียนทั้ง 3 สังกัด เพื่อให้กลุ่มตัวอย่างสะท้อนความหลากหลายของบริบทการศึกษาในเขตกรุงเทพมหานคร จำนวนกลุ่มตัวอย่าง 1,265 คน เป็นจำนวนที่เพียงพอต่อการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ ซึ่ง Embretson และ Reise (2000) ระบุว่าค่าพารามิเตอร์ภายใต้โมเดล 2 พารามิเตอร์ต้องการกลุ่มตัวอย่างอย่างน้อย 500 คน เพื่อให้ค่าพารามิเตอร์มีความเสถียร และเพียงพอต่อการวิเคราะห์องค์ประกอบเชิงยืนยัน ซึ่ง Kline (2016) แนะนำให้มีสัดส่วนกลุ่มตัวอย่างต่อพารามิเตอร์ในโมเดลอย่างน้อย 10 ต่อ 1 รวมทั้งเพียงพอต่อการสร้างเกณฑ์ปกติ ซึ่ง Crocker และ Algina (2008) แนะนำให้ใช้กลุ่มตัวอย่างขนาดใหญ่เพียงพอที่จะสะท้อนความหลากหลายของประชากร กลุ่มตัวอย่างในตอนนี้แยกออกจากกลุ่มทดลองใช้แบบวัดในตอนต้นที่ 1 อย่างสิ้นเชิง เพื่อป้องกันการปนเปื้อนของข้อมูลจากการที่ผู้สอบเคยทำแบบวัดมาก่อน

2.2 การตรวจสอบคุณสมบัติทางจิตมิติภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิม

ผู้วิจัยตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดฉบับสมบูรณ์ 40 ข้อ จากข้อมูลของกลุ่มเก็บข้อมูลหลัก ภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิม ครอบคลุม 3 ด้าน ได้แก่ ค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบ และค่าความเชื่อมั่นของแบบวัด

ค่าความยากของข้อสอบ คือ สัดส่วนของผู้สอบที่ตอบข้อสอบนั้นถูก ผู้วิจัยใช้เกณฑ์การแปลผลตามแนวทางของ Crocker และ Algina (2008) โดยกำหนดเกณฑ์การยอมรับว่าข้อสอบมีคุณภาพดีเมื่อมีค่าความยากอยู่ในช่วง 0.20 ถึง 0.80 ค่าอำนาจจำแนกของข้อสอบ คือ ค่าที่บ่งบอกความสามารถของข้อสอบในการจำแนกผู้สอบที่มีความสามารถสูงออกจากผู้สอบที่มีความสามารถต่ำ ผู้วิจัยคำนวณในรูปของสหสัมพันธ์พอยต์ไบซีเรียลระหว่างคะแนนของข้อสอบกับคะแนนรวมของแบบวัด และใช้เกณฑ์การแปลผลตามแนวทางของ Ebel และ Frisbie (1991) โดยกำหนดเกณฑ์การยอมรับว่า

ข้อสอบมีอำนาจจำแนกที่ดีพอเมื่อมีค่าตั้งแต่ 0.20 ขึ้นไป ส่วนค่าความเชื่อมั่นของแบบวัดทั้งฉบับ ผู้วิจัยคำนวณด้วยสูตร KR-20 และใช้เกณฑ์การแปลผลตามแนวทางของ Nunnally และ Bernstein (1994) เกณฑ์การแปลผลของดัชนีทั้ง 3 ด้านแสดงในตาราง 12

ตาราง 12 เกณฑ์การแปลผลดัชนีคุณภาพข้อสอบและความเชื่อมั่นตามทฤษฎีการทดสอบแบบดั้งเดิม

ดัชนี	ช่วงค่า	การแปลผล
ค่าความยาก (p)	0.00–0.19	ยากเกินไป ควรปรับปรุงหรือคัดออก
	0.20–0.39	ค่อนข้างยาก ยอมรับได้
	0.40–0.60	เหมาะสม (ดีที่สุด)
	0.61–0.80	ค่อนข้างง่าย ยอมรับได้
	0.81–1.00	ง่ายเกินไป ควรปรับปรุงหรือคัดออก
ค่าอำนาจจำแนก (r)	≥ 0.40	ดีมาก (ดีที่สุด)
	0.30–0.39	ดี
	0.20–0.29	พอใช้
	< 0.20	ต่ำ ควรปรับปรุง
	ติดลบ	มีปัญหา ต้องคัดออก
ค่าความเชื่อมั่น (KR-20)	≥ 0.90	สูงมาก
	0.80–0.89	สูง
	0.70–0.79	ยอมรับได้
	< 0.70	ต่ำ ควรปรับปรุงแบบวัด

หมายเหตุ: เกณฑ์ค่าความยากดัดแปลงจาก Crocker และ Algina (2008) เกณฑ์ค่าอำนาจจำแนกจาก Ebel และ Frisbie (1991) และเกณฑ์ค่าความเชื่อมั่นจาก Nunnally และ Bernstein (1994)

2.3 การเก็บรวบรวมและการคัดกรองข้อมูล

ผู้วิจัยจัดทำหนังสือขอความอนุเคราะห์เก็บข้อมูลจากสถานศึกษาทั้ง 3 สังกัด เมื่อสถานศึกษาให้ความยินยอม ผู้วิจัยเปิดระบบแบบวัดออนไลน์ให้นักเรียนเข้าทำแบบวัด ภายใต้การกำกับดูแลของครูผู้สอน โดยข้อมูลการตอบถูกบันทึกอัตโนมัติในตารางคำนวณออนไลน์ จากนั้นผู้วิจัยคัดกรองข้อมูลเบื้องต้น โดยตรวจสอบความครบถ้วนของการตอบ คัดข้อมูลที่ไมครบถ้วนออก และ

ตรวจสอบรูปแบบการตอบที่ผิดปกติ เช่น การตอบรูปแบบเดียวกันตลอดทั้งฉบับ ข้อมูลที่ผ่านการคัดกรองถูกนำเข้าสู่อุปกรณ์ภาษา R เวอร์ชัน 4.5.3 ผ่านเว็บแอปพลิเคชัน Posit Cloud เพื่อทำการวิเคราะห์

2.4 การตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ก่อนการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ ผู้วิจัยตรวจสอบข้อตกลงเบื้องต้น 2 ประการ ได้แก่ ความเป็นเอกมิติ และความเป็นอิสระเฉพาะที่

2.4.1 การตรวจสอบความเป็นเอกมิติ

ความเป็นเอกมิติ หมายถึง ข้อตกลงเบื้องต้นที่ระบุว่าแบบวัดต้องวัดความสามารถแฝงเพียงมิติเดียวเป็นหลัก ผู้วิจัยตรวจสอบความเป็นเอกมิติด้วย 2 วิธีประกอบกัน วิธีที่หนึ่งคือ การวิเคราะห์อัตราส่วนค่าไอเกนระหว่างองค์ประกอบแรกกับองค์ประกอบที่สอง ตามแนวคิดของ Reckase (1979) ที่ระบุว่าหากอัตราส่วนมีค่ามากกว่า 3 ต่อ 1 ถือว่าข้อมูลมีความเป็นเอกมิติในระดับที่ยอมรับได้ และมีปัจจัยหลักปัจจัยเดียวที่โดดเด่น วิธีที่สอง คือ การวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโมเดลองค์ประกอบเดียว ที่กำหนดให้ข้อสอบทั้ง 40 ข้อบรรจุอยู่ในปัจจัยแฝงเดียวกัน แล้วพิจารณาดัชนีความสอดคล้องของโมเดล

2.4.2 การตรวจสอบความเป็นอิสระเฉพาะที่

ความเป็นอิสระเฉพาะที่ หมายถึง ข้อตกลงเบื้องต้นที่ระบุว่า เมื่อควบคุมความสามารถของผู้สอบแล้ว การตอบข้อสอบแต่ละข้อต้องไม่ขึ้นอยู่กับคำตอบข้อสอบอื่น ผู้วิจัยตรวจสอบความเป็นอิสระเฉพาะที่ด้วยสถิติ Q3 ที่เสนอโดย Yen (1984) ซึ่งคำนวณจากสหสัมพันธ์ของค่าคงเหลือระหว่างคู่ข้อสอบ หลังจากปรับโมเดลการตอบสนองข้อสอบแล้ว สำหรับแบบวัดเชิงสถานการณ์ที่ข้อสอบ 4 ข้อในแต่ละสถานการณ์อาศัยบริบทเดียวกัน การตรวจสอบความเป็นอิสระเฉพาะที่มีความสำคัญเป็นพิเศษ เนื่องจากข้อสอบในสถานการณ์เดียวกันมีความเสี่ยงที่จะสัมพันธ์กันเกินกว่าที่ความสามารถแฝงจะอธิบายได้ ผู้วิจัยใช้เกณฑ์การแปลผลตามแนวทางของ Christensen, Makransky และ Horton (2017) ดังแสดงในตาราง 13

ตาราง 13 เกณฑ์การแปลผลค่า Q3 สำหรับการตรวจสอบความเป็นอิสระเฉพาะที่

ช่วงค่าสัมบูรณ์ของ Q3	การแปลผล	ข้อปฏิบัติ
$Q3 \leq 0.20$	ความเป็นอิสระเฉพาะที่อยู่ในระดับที่ยอมรับได้	ดำเนินการวิเคราะห์ต่อได้
$0.20 < Q3 \leq 0.30$	มีการพึ่งพิงเฉพาะที่ระดับปานกลาง	รายงานและพิจารณาผลกระทบ

ตาราง 13 (ต่อ)

ช่วงค่าสัมบูรณ์ของ Q3	การแปลผล	ข้อปฏิบัติ
$Q3 > 0.30$	มีการพึ่งพิงเฉพาะที่ระดับสูง	พิจารณารวมหรือตัดข้อสอบที่มีปัญหา

หมายเหตุ: ดัดแปลงจาก Christensen, Makransky และ Horton (2017) และ Yen (1984)

ในการวิจัยครั้งนี้ ผู้วิจัยกำหนดเกณฑ์การยอมรับที่ค่าสัมบูรณ์ของ Q3 น้อยกว่า 0.20 และพิจารณาค่าเฉลี่ยของ Q3 ในแบบวัดทั้งฉบับ ที่ควรใกล้ 0 เพื่อแสดงว่าโดยรวมแบบวัดมีความเป็นอิสระเฉพาะที่

2.5 การวิเคราะห์ความตรงเชิงโครงสร้าง

ผู้วิจัยตรวจสอบความตรงเชิงโครงสร้างของแบบวัดด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน เนื่องจากข้อมูลเป็นตัวแทนแบบทวิภาค ผู้วิจัยประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดถ่วงน้ำหนักที่ปรับค่าเฉลี่ยและความแปรปรวน (Weighted Least Squares Mean and Variance Adjusted: WLSMV) ซึ่งเหมาะสมกับข้อมูลแบบจัดประเภท และพิจารณาความสอดคล้องของโมเดลด้วยดัชนีหลายตัวร่วมกัน ตามเกณฑ์ของ Hu และ Bentler (1999) และ Kline (2016) ดังแสดงในตาราง 14

ตาราง 14 ดัชนีความสอดคล้องของโมเดลในการวิเคราะห์องค์ประกอบเชิงยืนยันและเกณฑ์การยอมรับ

ดัชนี	ชื่อเต็ม	เกณฑ์การยอมรับ
χ^2/df	อัตราส่วนไคสแควร์ต่อองศาอิสระ	น้อยกว่า 3.00
RMSEA	ความคลาดเคลื่อนของการประมาณค่า	ไม่เกิน 0.06 (ดี); ไม่เกิน 0.08 (ยอมรับได้)
CFI	ดัชนีความสอดคล้องเชิงเปรียบเทียบ	มากกว่า 0.95
TLI	ดัชนีทักเกอร์-ลูอิส	มากกว่า 0.95
SRMR	ค่าเฉลี่ยของส่วนเหลือมาตรฐาน	ไม่เกิน 0.08

หมายเหตุ: เกณฑ์ตาม Hu และ Bentler (1999) และ Kline (2016)

2.6 การคัดเลือกโมเดลทฤษฎีการตอบสนองข้อสอบ

ผู้วิจัยเปรียบเทียบโมเดลทฤษฎีการตอบสนองข้อสอบสำหรับข้อสอบแบบให้คะแนนทวิภาค จำนวน 3 โมเดล ได้แก่ โมเดล 1 พารามิเตอร์ ซึ่งประมาณเฉพาะค่าความยาก โมเดล 2 พารามิเตอร์ ซึ่งประมาณค่าความยากและค่าอำนาจจำแนก และโมเดล 3 พารามิเตอร์ ซึ่งเพิ่มค่าการเดาสติติ ได้แก่ ค่า AIC และค่า BIC โดยโมเดลที่มีค่าต่ำกว่าถือว่าเหมาะสมกว่า และการทดสอบอัตราส่วน

ภาวะน่าจะเป็น สำหรับเปรียบเทียบโมเดลที่ซ้อนกัน ด้านที่สอง เกณฑ์เชิงระเบียบวิธี ได้แก่ ความสามารถในการประมาณค่าพารามิเตอร์ของโมเดล และความสอดคล้องเชิงทฤษฎีของโมเดลกับลักษณะของแบบวัด ดัชนีการเปรียบเทียบโมเดลแสดงในตาราง 15

ตาราง 15 ดัชนีการเปรียบเทียบโมเดลทฤษฎีการตอบสนองข้อสอบ

ดัชนี	เกณฑ์การเลือกโมเดล
AIC	โมเดลที่มีค่า AIC ต่ำกว่า เหมาะสมกว่า
BIC	โมเดลที่มีค่า BIC ต่ำกว่า เหมาะสมกว่า
การทดสอบอัตราส่วนภาวะน่าจะเป็น	ทดสอบว่าโมเดลที่ซับซ้อนกว่ามีความสอดคล้องกับข้อมูลดีกว่าโมเดลที่เรียบง่ายขึ้นอย่างมีนัยสำคัญทางสถิติหรือไม่

หมายเหตุ: ดัดแปลงจาก De Ayala (2009)

2.7 การประมาณค่าพารามิเตอร์และการตรวจสอบความสอดคล้องของโมเดล

ภายหลังคัดเลือกโมเดลที่เหมาะสมแล้ว ผู้วิจัยประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีภาวะน่าจะเป็นสูงสุดส่วนเพิ่ม ผ่านอัลกอริทึมการคาดหวังและการหาค่าสูงสุด โดยพารามิเตอร์ที่ประมาณ ได้แก่ ค่าอำนาจจำแนก สัญลักษณ์ a ซึ่งแสดงความสามารถของข้อสอบในการจำแนกผู้สอบที่มีความสามารถต่างกัน และค่าความยาก สัญลักษณ์ b ซึ่งแสดงระดับความยากของข้อสอบบนสเกลความสามารถแฝง ผู้วิจัยใช้เกณฑ์การแปลผลค่าพารามิเตอร์อำนาจจำแนกตามแนวทางของ Baker (2001) ดังแสดงในตาราง 16

ตาราง 16 เกณฑ์การแปลผลค่าพารามิเตอร์อำนาจจำแนก (a)

ช่วงค่า a	ระดับการจำแนก
น้อยกว่า 0.35	จำแนกได้น้อยมาก
0.35–0.64	จำแนกได้น้อย
0.65–1.34	จำแนกได้ปานกลาง
1.35–1.69	จำแนกได้สูง
1.70	ขึ้นไปจำแนกได้สูงมาก

หมายเหตุ: ดัดแปลงจาก Baker (2001)

ส่วนค่าพารามิเตอร์ความยาก สัญลักษณ์ b บ่งบอกตำแหน่งของข้อสอบบนแกนความสามารถ โดยทั่วไปค่า b อยู่ในช่วง -3 ถึง $+3$ ค่า b ติดลบหมายความว่าข้อสอบง่าย และค่า b เป็นบวกหมายความว่าข้อสอบยาก

ผู้วิจัยตรวจสอบความสอดคล้องของโมเดลกับข้อมูลใน 2 ระดับ ระดับที่หนึ่ง คือ ระดับโมเดลโดยรวม ผู้วิจัยใช้สถิติ M^2 ที่พัฒนาโดย Maydeu-Olivares และ Joe (2005) ร่วมกับดัชนี RMSEA, CFI, TLI และ SRMSR โดยใช้เกณฑ์การยอมรับที่ RMSEA ไม่เกิน 0.08 ค่า CFI และ TLI มากกว่า 0.95 และ SRMSR ไม่เกิน 0.08 ระดับที่สอง คือ ระดับรายข้อ ผู้วิจัยใช้สถิติ $S-X^2$ ที่พัฒนาโดย Orlando และ Thissen (2000) ซึ่งตรวจสอบว่าข้อสอบแต่ละข้อมีความสอดคล้องกับโมเดลที่เลือกหรือไม่ โดยข้อสอบที่มีค่านัยสำคัญตั้งแต่ 0.05 ขึ้นไปถือว่าสอดคล้องกับโมเดล ทั้งนี้ เมื่อกลุ่มเก็บข้อมูลหลักมีขนาดใหญ่ สถิติ $S-X^2$ อาจมีนัยสำคัญทางสถิติแม้ความไม่สอดคล้องจะมีน้อย ดังนั้น ผู้วิจัยพิจารณาทั้งค่านัยสำคัญและขนาดของความไม่สอดคล้องร่วมกัน

นอกจากนี้ ผู้วิจัยวิเคราะห์ฟังก์ชันสารสนเทศของแบบวัด เพื่อตรวจสอบความแม่นยำของการประมาณความสามารถผู้สอบในแต่ละระดับความสามารถ ค่าสารสนเทศที่สูงในช่วงความสามารถใดบ่งบอกว่าแบบวัดสามารถประมาณความสามารถของผู้สอบในช่วงนั้นได้แม่นยำสูง พร้อมคำนวณค่าความเชื่อมั่นเชิงประจักษ์ในกรอบทฤษฎีการตอบสนองข้อสอบ ซึ่งมีค่าระหว่าง 0 ถึง 1 และตีความเช่นเดียวกับค่าความเชื่อมั่นในทฤษฎีการทดสอบแบบดั้งเดิม และประมาณค่าความสามารถของผู้สอบรายบุคคลด้วยวิธีคาดหวังภายหลัง ตามคำแนะนำของ Bock และ Mislevy (1982) เพื่อนำค่าความสามารถไปใช้ในการสร้างเกณฑ์ปกติต่อไป

ตอนที่ 3 การสร้างเกณฑ์ปกติของแบบวัด

การดำเนินการตามวัตถุประสงค์ข้อ 3 มุ่งสร้างเกณฑ์ปกติของแบบวัดทักษะการคิดเชิงคำนวณ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร เพื่อใช้เป็นเกณฑ์อ้างอิงในการตีความคะแนนของผู้สอบ ที่ครูและผู้ที่เกี่ยวข้องสามารถนำไปใช้ในการประเมินทักษะการคิดเชิงคำนวณของผู้เรียนได้อย่างเป็นรูปธรรม ผู้วิจัยใช้ข้อมูลคะแนนจากกลุ่มตัวอย่าง 1,265 คน ที่เก็บในตอนที่ 2 และสร้างเกณฑ์ปกติตามแนวทางของ ซวาล แพร์ตกุล (2552) และ สมนึก ภัททิยธนี (2562)

3.1 ขั้นตอนการสร้างเกณฑ์ปกติ

ผู้วิจัยสร้างเกณฑ์ปกติในรูปคะแนนที่ปกติ เป็นเกณฑ์ภาพรวมชุดเดียว สำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร โดยดำเนินการ 4 ขั้นตอน ดังนี้

ขั้นที่ 1 การวิเคราะห์การแจกแจงของคะแนนดิบ ผู้วิจัยวิเคราะห์การแจกแจงของคะแนนดิบ โดยพิจารณาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด ค่าสูงสุด ค่าความเบ้ และค่าความโด่ง เพื่อให้เข้าใจลักษณะการแจกแจงของคะแนนก่อนการสร้างเกณฑ์ปกติ ซึ่งเป็นเงื่อนไขสำคัญในการสร้างคะแนนที่ปกติ

ขั้นที่ 2 การคำนวณคะแนนเปอร์เซ็นต์ไทล์ ผู้วิจัยแปลงคะแนนดิบเป็นลำดับเปอร์เซ็นต์ไทล์ ซึ่งสะท้อนตำแหน่งสัมพัทธ์ของผู้สอบในกลุ่มอ้างอิง ตามสูตร

$$PR = [(cf_below + 0.5f) / N] \times 100$$

เมื่อ cf_below คือ ความถี่สะสมของคะแนนที่ต่ำกว่าคะแนนที่พิจารณา f คือ ความถี่ของคะแนนที่พิจารณา และ N คือ จำนวนผู้สอบทั้งหมดในกลุ่มอ้างอิง

ขั้นที่ 3 การแปลงคะแนนเปอร์เซ็นต์ไทล์เป็นคะแนนที่ปกติ ผู้วิจัยแปลงลำดับเปอร์เซ็นต์ไทล์เป็นคะแนนซีโดยใช้ฟังก์ชันการแจกแจงปกติมาตรฐานย้อนกลับในโปรแกรม R แล้วแปลงคะแนนซีเป็นคะแนนทีปกติด้วยสูตร $T = 50 + 10Z$ การใช้คะแนนที่ปกติทำให้คะแนนมีค่าเฉลี่ย 50 และส่วนเบี่ยงเบนมาตรฐาน 10 ซึ่งเข้าใจง่ายและไม่มีค่าติดลบ

ขั้นที่ 4 การจัดทำตารางเกณฑ์ปกติ ผู้วิจัยจัดทำตารางเกณฑ์ปกติที่แสดงความสัมพันธ์ระหว่างคะแนนดิบ ลำดับเปอร์เซ็นต์ไทล์ และคะแนนที่ปกติ เพื่อให้ครู ผู้บริหาร และผู้เกี่ยวข้องนำไปใช้อ้างอิงในการตีความคะแนนของนักเรียน

3.2 การกำหนดเกณฑ์การแปลผลระดับทักษะ

ผู้วิจัยกำหนดเกณฑ์การแปลผลระดับทักษะการคิดเชิงคำนวณเป็น 3 ระดับ ตามแนวทางของ สมนึก ภัททิยธนี (2562) ซึ่งอิงหลักการทางสถิติของการแจกแจงปกติ โดยใช้ค่า ± 1 ส่วนเบี่ยงเบนมาตรฐานเป็นจุดแบ่งระดับ ดังแสดงในตาราง 17

ตาราง 17 เกณฑ์การแปลผลระดับทักษะการคิดเชิงคำนวณ

ระดับ	ช่วงคะแนนที่	ลำดับเปอร์เซ็นต์ไทล์	ความหมาย
สูงกว่าเกณฑ์ปกติ	$T \geq 60$	$PR \geq 84$	มีทักษะสูงกว่าค่าเฉลี่ยของกลุ่มอ้างอิง 1 ส่วนเบี่ยงเบนมาตรฐานขึ้นไป
อยู่ในเกณฑ์ปกติ	$T 40-59$	$PR 16-83$	มีทักษะอยู่ในระดับปกติ ครอบคลุมประมาณร้อยละ 68 กลางของกลุ่ม
ต่ำกว่าเกณฑ์ปกติ	$T < 40$	$PR < 16$	มีทักษะต่ำกว่าค่าเฉลี่ยของกลุ่มอ้างอิง 1 ส่วนเบี่ยงเบนมาตรฐานลงไป

หมายเหตุ: เกณฑ์การแบ่ง ± 1 ส่วนเบี่ยงเบนมาตรฐาน อิงหลักการทางสถิติของการแจกแจงปกติที่ครอบคลุมประชากรประมาณร้อยละ 68 ในช่วงปกติ (สมนึก ภัททิยธนี, 2551)

ผู้วิจัยเลือกใช้เกณฑ์การแปลผล 3 ระดับ เนื่องจากเป็นเกณฑ์ที่เข้าใจง่ายและใช้ได้จริง ในบริบทการเรียนการสอน สอดคล้องกับหลักการทางสถิติของการแจกแจงปกติ และเพียงพอต่อการจำแนกนักเรียนเพื่อวางแผนการพัฒนา โดยไม่ซับซ้อนเกินไปจนทำให้การตีความยากขึ้น

3.3 การสร้างเกณฑ์ปกติของคะแนนรวมและคะแนนรายองค์ประกอบ

ผู้วิจัยสร้างเกณฑ์ปกติ 2 ชุด ตามที่ใช้ในระบบรายงานผล ดังนี้ ชุดที่หนึ่ง เกณฑ์ปกติของคะแนนรวม สร้างจากคะแนนรวมของผู้สอบทั้ง 40 ข้อ จากกลุ่มเก็บข้อมูลหลัก 1,265 คน ครอบคลุมตารางลำดับเปอร์เซ็นต์ไทล์ ตารางคะแนนที่ปกติ และเกณฑ์การแปลผล 3 ระดับ และชุดที่สอง เกณฑ์ปกติของคะแนนรายองค์ประกอบ สร้างจากคะแนนของผู้สอบในแต่ละองค์ประกอบ ซึ่งมี 10 ข้อต่อองค์ประกอบ จากกลุ่มเดียวกัน ครอบคลุมตารางลำดับเปอร์เซ็นต์ไทล์และคะแนนที่ปกติ สำหรับการตีความทักษะของผู้สอบในแต่ละด้าน

เนื่องจากจำนวนข้อสอบรายองค์ประกอบ ซึ่งมี 10 ข้อ น้อยกว่าจำนวนข้อสอบทั้งฉบับ ซึ่งมี 40 ข้อ ความเชื่อมั่นในการประมาณความสามารถผู้สอบในรายองค์ประกอบจึงต่ำกว่าในระดับคะแนนรวม ผู้วิจัยจึงระบุในระบบรายงานผลและในเล่มวิทยานิพนธ์ว่า เกณฑ์ปกติรายองค์ประกอบใช้สำหรับการแปลผลในเชิงสังเกตเพื่อระบุจุดเด่นและจุดที่ควรพัฒนา ส่วนการตัดสินใจหลักเกี่ยวกับระดับทักษะให้ใช้เกณฑ์ปกติของคะแนนรวมเป็นหลัก

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

ผู้วิจัยใช้สถิติในการวิเคราะห์ข้อมูล จำแนกตามวัตถุประสงค์การวิจัย ดังนี้

ประการที่หนึ่ง สถิติพื้นฐาน ใช้ในการอธิบายลักษณะของกลุ่มตัวอย่างและลักษณะของคะแนนของผู้สอบ ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้ และค่าความโด่ง

ประการที่สอง สถิติสำหรับการตรวจสอบคุณภาพแบบวัดภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิม ได้แก่ ค่าดัชนีความสอดคล้องสำหรับการตรวจสอบความตรงเชิงเนื้อหา ค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบ และค่าความเชื่อมั่นของแบบวัดด้วยสูตร KR-20

ประการที่สาม สถิติสำหรับการตรวจสอบคุณภาพแบบวัดภายใต้กรอบทฤษฎีการตอบสนองข้อสอบ ได้แก่ การวิเคราะห์องค์ประกอบเชิงยืนยันพร้อมดัชนีความสอดคล้องของโมเดล ได้แก่ χ^2/df , RMSEA, CFI, TLI และ SRMR สถิติ Q3 สำหรับการตรวจสอบความเป็นอิสระเฉพาะที่ ค่า AIC และ BIC สำหรับการเปรียบเทียบโมเดล ค่าพารามิเตอร์ข้อสอบ ได้แก่ ค่าอำนาจจำแนกและค่าความยาก สถิติ M2 และดัชนีความสอดคล้องของโมเดลสำหรับความสอดคล้องในระดับโมเดลโดยรวม สถิติ $S-\chi^2$ สำหรับความสอดคล้องในระดับรายข้อ และฟังก์ชันสารสนเทศของแบบวัดพร้อมค่าความเชื่อมั่นเชิงประจักษ์

ประการที่สี่ สถิติสำหรับการสร้างเกณฑ์ปกติ ได้แก่ ลำดับเปอร์เซ็นต์ไทล์ และคะแนนที่ปกติ เครื่องมือทางสถิติและซอฟต์แวร์ที่ใช้

ผู้วิจัยวิเคราะห์ข้อมูลด้วยโปรแกรมภาษา R เวอร์ชัน 4.5.3 ผ่านเว็บแอปพลิเคชัน Posit Cloud ซึ่งเป็นโปรแกรมที่ใช้งานได้โดยไม่มีค่าใช้จ่ายและเป็นที่ยอมรับในวงการวิจัยทางสถิติ โดยใช้แพ็คเกจหลัก ดังนี้ แพ็คเกจ psych สำหรับการคำนวณค่าสถิติพื้นฐาน ค่าความเชื่อมั่น KR-20 และการวิเคราะห์ปัจจัย แพ็คเกจ lavaan สำหรับการวิเคราะห์องค์ประกอบเชิงยืนยัน แพ็คเกจ mirt สำหรับการวิเคราะห์ภายใต้กรอบทฤษฎีการตอบสนองข้อสอบ ครอบคลุมการประมาณค่าพารามิเตอร์ การเปรียบเทียบโมเดล การตรวจสอบความสอดคล้อง และการคำนวณฟังก์ชันสารสนเทศ และแพ็คเกจ ggplot2 สำหรับการสร้างกราฟและภาพประกอบ

นอกจากโปรแกรม R แล้ว ผู้วิจัยยังใช้กูเกิลเวิร์กสเปซ (Google Workspace) ในการพัฒนาระบบนำเสนอแบบวัดและระบบรายงานผลอัตโนมัติ ครอบคลุมแบบฟอร์มออนไลน์ เว็บไซต์ ตารางคำนวณออนไลน์ และกูเกิลแอปสคริปต์ ดังที่อธิบายในตอนต้นที่ 1 ทั้งนี้ รหัสคำสั่งภาษา R ที่ใช้ในการวิเคราะห์ข้อมูล นำเสนอในภาคผนวก เพื่อความโปร่งใสในการวิจัยและเพื่อให้สามารถทำซ้ำผลการวิเคราะห์ได้

จริยธรรมการวิจัย

ผู้วิจัยดำเนินการวิจัยภายใต้หลักจริยธรรมการวิจัยในมนุษย์ โดยคำนึงถึงสิทธิและความปลอดภัยของผู้เข้าร่วมการวิจัย ครอบคลุมประเด็นสำคัญ 4 ประการ ดังนี้

ประการที่หนึ่ง การขอความยินยอม ผู้วิจัยขอความยินยอมจากโรงเรียนและครูผู้สอนก่อนการเก็บข้อมูล โดยอธิบายวัตถุประสงค์ของการวิจัย กระบวนการเก็บข้อมูล ระยะเวลาที่ใช้ในการทำแบบวัด และการนำผลการวิจัยไปใช้ ในขณะเดียวกัน ผู้สอบทุกคนได้รับการแจ้งให้ทราบว่าการเข้าร่วมการวิจัยเป็นไปโดยสมัครใจ และสามารถถอนตัวได้ทุกขณะโดยไม่มีผลกระทบใด ๆ

ประการที่สอง การคุ้มครองข้อมูลส่วนบุคคล ข้อมูลส่วนบุคคลของผู้สอบ ได้แก่ ชื่อ-นามสกุล อีเมล โรงเรียน และระดับชั้น ใช้เพื่อการส่งรายงานผลให้ผู้สอบเท่านั้น และเก็บรักษาไว้อย่างเป็นความลับ ผู้วิจัยไม่เปิดเผยข้อมูลส่วนบุคคลแก่บุคคลภายนอก และในการนำเสนอผลการวิจัย ผู้วิจัยใช้ข้อมูลในรูปของกลุ่ม ไม่ใช่รายบุคคล

ประการที่สาม การคืนประโยชน์ให้ผู้เข้าร่วมการวิจัย ผู้สอบทุกคนได้รับรายงานผลทักษะการคิดเชิงคำนวณของตนเองทางอีเมลโดยอัตโนมัติภายหลังการทำแบบวัดเสร็จสิ้น ครอบคลุมคะแนนรวม คะแนนรายองค์ประกอบ และระดับทักษะ พร้อมคำอธิบายที่เข้าใจง่าย เพื่อให้ผู้สอบและครูสามารถนำผลไปใช้ในการพัฒนาการเรียนรู้ต่อไป

ประการที่สี่ การจัดเก็บข้อมูล ผู้วิจัยจัดเก็บข้อมูลการวิจัยในรูปแบบดิจิทัล โดยมีการสำรองข้อมูลและการป้องกันการเข้าถึงจากบุคคลที่ไม่เกี่ยวข้อง ข้อมูลจะถูกเก็บไว้ตามระยะเวลาที่กำหนดในหลักจริยธรรมการวิจัย และถูกทำลายอย่างปลอดภัยเมื่อสิ้นสุดการใช้งาน



บทที่ 4

ผลการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ 3 ประการ ได้แก่ (1) เพื่อพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น (2) เพื่อตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด ภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ และ (3) เพื่อสร้างเกณฑ์ปกติของแบบวัด ผู้วิจัยนำเสนอผลการวิจัยตามลำดับวัตถุประสงค์ทั้ง 3 ข้อ เป็น 3 ตอน ดังนี้

ตอนที่ 1 ผลการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์

ตอนที่ 2 ผลการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด

ตอนที่ 3 ผลการสร้างเกณฑ์ปกติของแบบวัด

เพื่อความสะดวกในการอ่าน ผู้วิจัยกำหนดสัญลักษณ์ที่ใช้ในบทนี้ ดังนี้

N แทน จำนวนกลุ่มตัวอย่าง

M แทน ค่าเฉลี่ย

SD แทน ส่วนเบี่ยงเบนมาตรฐาน

p แทน ค่าความยากของข้อสอบ

r แทน ค่าอำนาจจำแนกของข้อสอบ

a แทน ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบในกรอบทฤษฎีการตอบสนองข้อสอบ

b แทน ค่าพารามิเตอร์ความยากของข้อสอบในกรอบทฤษฎีการตอบสนองข้อสอบ

θ แทน ค่าความสามารถของผู้สอบในกรอบทฤษฎีการตอบสนองข้อสอบ

IOC แทน ค่าดัชนีความสอดคล้องของข้อสอบกับวัตถุประสงค์การวัด

PR แทน ลำดับเปอร์เซ็นต์ไทล์

และ T แทน คะแนนที่ปกติ

ตอนที่ 1 ผลการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์

ผลการดำเนินการตามวัตถุประสงค์ข้อ 1 ผู้วิจัยพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำเร็จ มีรายละเอียดผลการพัฒนาในแต่ละขั้น ดังนี้

1.1 ผลการสร้างแบบวัดฉบับร่าง

ผู้วิจัยพัฒนาแบบวัดฉบับร่างได้จำนวน 80 ข้อ ใน 20 สถานการณ์ โดยแต่ละสถานการณ์ประกอบด้วยข้อคำถามจำนวน 4 ข้อ ซึ่งครอบคลุมองค์ประกอบของทักษะการคิดเชิงคำนวณทั้ง 4 ด้าน

ได้แก่ การแบ่งย่อยปัญหา (Decomposition) การพิจารณารูปแบบ (Pattern Recognition) การคิดเชิงนามธรรม (Abstraction) และการออกแบบอัลกอริทึม (Algorithm Design) ภายใต้บริบทของสถานการณ์เดียวกัน

สถานการณ์ทั้ง 20 สถานการณ์ได้รับการพัฒนาจากบริบทในชีวิตประจำวันของนักเรียนระดับมัธยมศึกษาตอนต้น โดยครอบคลุม 3 ด้าน ได้แก่ ด้านชีวิตส่วนบุคคล ด้านการจัดกิจกรรมในโรงเรียน และด้านชุมชน สังคม และสิ่งแวดล้อม เพื่อให้การวัดสะท้อนความสามารถในการประยุกต์ใช้ทักษะการคิดเชิงคำนวณในการแก้ปัญหาในบริบทที่หลากหลายและใกล้เคียงกับสถานการณ์จริงที่ผู้เรียนพบในชีวิตประจำวัน

การสร้างข้อคำถามจำนวน 80 ข้อ ซึ่งเป็น 2 เท่าของจำนวนข้อคำถามในแบบวัดฉบับสมบูรณ์ที่กำหนดไว้ มีวัตถุประสงค์เพื่อให้มีข้อคำถามสำรองเพียงพอสำหรับการคัดเลือกข้อคำถามที่มีคุณภาพในขั้นตอนการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิม ส่งผลให้แบบวัดฉบับสมบูรณ์ประกอบด้วยข้อคำถามที่ผ่านเกณฑ์คุณภาพและสามารถสะท้อนองค์ประกอบของทักษะการคิดเชิงคำนวณได้อย่างครบถ้วน

1.2 ผลการตรวจสอบความตรงเชิงเนื้อหา

ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบวัดฉบับร่าง 80 ข้อ โดยผู้เชี่ยวชาญ 7 คน พบว่าข้อคำถามทุกข้อมีค่าดัชนีความสอดคล้องผ่านเกณฑ์การยอมรับที่ 0.50 ดังแสดงในตาราง 18

ตาราง 18 ผลการประเมินความตรงเชิงเนื้อหาของแบบวัดฉบับร่าง 80 ข้อ โดยผู้เชี่ยวชาญ 7 คน

ช่วงค่าดัชนีความ สอดคล้อง	จำนวนข้อ	ร้อยละ	การแปลผล
1.00	75	93.75	ผู้เชี่ยวชาญเห็นพ้องกันทุกคน
0.86	4	5.00	ผู้เชี่ยวชาญ 6 จาก 7 คนเห็นพ้อง
0.57	1	1.25	ผ่านเกณฑ์ แต่มีผู้เชี่ยวชาญบางคนไม่แน่ใจ
รวม	80	100.00	ผ่านเกณฑ์ทุกข้อ

หมายเหตุ: เกณฑ์การยอมรับที่ค่าดัชนีความสอดคล้องตั้งแต่ 0.50 ขึ้นไป

จากตาราง 18 ข้อคำถามทั้ง 80 ข้อมีค่าดัชนีความสอดคล้องอยู่ในช่วง 0.57 ถึง 1.00 ผ่านเกณฑ์ทุกข้อ ผลนี้มีความหมายว่าข้อคำถามที่ผู้วิจัยสร้างขึ้นสอดคล้องกับองค์ประกอบของทักษะการคิดเชิงคำนวณที่ต้องการวัด โดยข้อคำถามจำนวน 75 ข้อ หรือร้อยละ 93.75 ได้รับการยืนยันจาก

ผู้เชี่ยวชาญทุกคนอย่างเป็นเอกฉันท์ ความเห็นพ้องในระดับสูงเช่นนี้สะท้อนว่านิยามเชิงปฏิบัติการของทักษะการคิดเชิงคำนวณทั้ง 4 องค์ประกอบที่ผู้วิจัยสังเคราะห์ขึ้นมีความชัดเจนเพียงพอ และผู้วิจัยสามารถแปลงนิยามดังกล่าวเป็นข้อคำถามที่วัดองค์ประกอบเป้าหมายได้ตรงตามเจตนา ผลการตรวจสอบความตรงเชิงเนื้อหาเป็นหลักฐานชิ้นแรกที่สนับสนุนคุณภาพของแบบวัด ซึ่งเป็นพื้นฐานสำคัญก่อนการตรวจสอบคุณสมบัติทางจิตมิติด้านอื่นในลำดับต่อไป

นอกจากการประเมินเชิงปริมาณแล้ว ผู้วิจัยรวบรวมข้อเสนอแนะเชิงคุณภาพจากผู้เชี่ยวชาญ และนำมาปรับปรุงข้อคำถามใน 2 ลักษณะ คือ การปรับถ้อยคำของสถานการณ์และข้อคำถามให้ชัดเจนและเหมาะสมกับระดับชั้นของนักเรียน และการปรับตัวเลือกให้มีความสมเหตุสมผล และสะท้อนระดับการคิดที่แตกต่างกัน

1.3 ผลการพัฒนาแบบวัดในรูปแบบออนไลน์

ภายหลังปรับปรุงข้อคำถามตามข้อเสนอแนะของผู้เชี่ยวชาญแล้ว ผู้วิจัยพัฒนาแบบวัดให้อยู่ในรูปแบบออนไลน์ โดยใช้แบบฟอร์มออนไลน์ของกูเกิลเป็นเครื่องมือนำเสนอสถานการณ์และเก็บคำตอบ ร่วมกับเว็บไซต์ที่สร้างด้วยกูเกิลไซต์เป็นศูนย์กลางข้อมูล ผลการพัฒนาได้ระบบแบบวัดออนไลน์ที่ประกอบด้วยส่วนการทำงานสำคัญ ดังนี้

ส่วนที่ 1 เว็บไซต์ศูนย์รวมข้อมูลการวิจัย ผู้วิจัยพัฒนาเว็บไซต์ที่ทำหน้าที่ชี้แจงวัตถุประสงค์ของการวิจัย คำชี้แจงการทำแบบวัด คู่มือการใช้งาน และช่องทางติดต่อผู้วิจัย เว็บไซต์นี้เผยแพร่ที่ <https://sites.google.com/view/comthinkprob/homepage>

ส่วนที่ 2 ระบบยืนยันตัวตนและการเก็บข้อมูลพื้นฐาน ผู้สอบลงชื่อเข้าใช้ด้วยบัญชีของกูเกิล และกรอกข้อมูลพื้นฐาน ได้แก่ ชื่อ-นามสกุล ระดับชั้น โรงเรียน สังกัด และที่อยู่อีเมล โดยระบบกำหนดให้ทำแบบวัดได้ครั้งเดียวต่อหนึ่งบัญชี

ส่วนที่ 3 ระบบนำเสนอสถานการณ์และข้อคำถาม แบบวัดฉบับสมบูรณ์ 40 ข้อ นำเสนอในรูปแบบข้อความ โดยแสดงสถานการณ์พร้อมข้อคำถามย่อย 4 ข้อในหน้าเดียวกัน

ส่วนที่ 4 ระบบรายงานผลอัตโนมัติ ภายหลังการทำแบบวัดเสร็จสิ้น ระบบประมวลผลและส่งผลคะแนนไปยังอีเมลของผู้สอบโดยอัตโนมัติ ทั้งคะแนนรวมและคะแนนรายองค์ประกอบทั้ง 4 ด้าน

ผลการพัฒนาในขั้นนี้แสดงให้เห็นว่าแบบวัดที่พัฒนาขึ้นเป็นระบบที่นำไปใช้ได้จริง โดยมีคุณลักษณะที่เอื้อต่อการนำไปใช้ในวงกว้าง ได้แก่ การเข้าถึงได้จากทุกสถานที่ที่มีอินเทอร์เน็ต การตรวจให้คะแนนอัตโนมัติที่ลดความคลาดเคลื่อนจากการตรวจด้วยมือ และการรายงานผลที่รวดเร็วซึ่งช่วยให้นักเรียนและครูได้รับข้อมูลย้อนกลับทันที

1.4 ผลการคัดเลือกข้อคำถามเข้าสู่แบบวัดฉบับสมบูรณ์

ผู้วิจัยนำแบบวัดฉบับร่าง 80 ข้อ ไปทดลองใช้กับนักเรียนระดับมัธยมศึกษาตอนต้นจำนวน 150 คน แล้ววิเคราะห์คุณภาพรายข้อตามทฤษฎีการทดสอบแบบดั้งเดิม ผลการวิเคราะห์ใน

ขั้นตอนการใช้ครอบคลุมค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบ และค่าความเชื่อมั่นของแบบวัดทั้งฉบับ ดังนี้

ค่าความยากของข้อสอบทั้ง 80 ข้อ จากการคำนวณด้วยวิธีกลุ่มสูง-กลุ่มต่ำ ร้อยละ 27 พบว่ามีข้อคำถามจำนวน 50 ข้อ ที่มีค่าความยากอยู่ในช่วง 0.20 ถึง 0.80 ตามเกณฑ์ของ Crocker และ Algina (2008) ส่วนข้อคำถามอีก 30 ข้อ มีค่าความยากต่ำกว่า 0.20 หรือสูงกว่า 0.80 ซึ่งไม่ผ่านเกณฑ์เนื่องจากง่ายเกินไปหรือยากเกินไป

ค่าอำนาจจำแนกของข้อสอบทั้ง 80 ข้อ พบว่ามีข้อคำถามที่มีค่าตั้งแต่ 0.20 ขึ้นไป ตามเกณฑ์ของ Ebel และ Frisbie (1991) สอดคล้องกับข้อคำถามที่ผ่านเกณฑ์ค่าความยาก

ค่าความเชื่อมั่นของแบบวัดฉบับร่าง 80 ข้อ จากกลุ่มทดลองใช้ 150 คน ด้วยสูตร KR-20 มีค่าเท่ากับ 0.84 ซึ่งอยู่ในระดับสูง ตามเกณฑ์ของ Nunnally และ Bernstein (1994) ค่าความเชื่อมั่นในระดับนี้สะท้อนว่าแบบวัดฉบับร่างที่พัฒนาขึ้นมีความสม่ำเสมอภายในที่ดี เป็นพื้นฐานที่มั่นคงสำหรับการคัดเลือกข้อคำถามที่มีคุณภาพเข้าสู่แบบวัดฉบับสมบูรณ์

จากการวิเคราะห์คุณภาพรายข้อ ผู้วิจัยคัดเลือกข้อคำถามที่ผ่านเกณฑ์คุณภาพทั้งด้านค่าความยากและค่าอำนาจจำแนก จำนวน 50 ข้อ จากนั้นพิจารณาความครบถ้วนขององค์ประกอบในแต่ละสถานการณ์ และได้สถานการณ์ที่มีข้อคำถามครบทั้ง 4 องค์ประกอบ จำนวน 10 สถานการณ์ รวมเป็นแบบวัดฉบับสมบูรณ์ 40 ข้อ ที่วัดองค์ประกอบของทักษะการคิดเชิงคำนวณครบทั้ง 4 ด้าน ด้านละ 10 ข้อ ดังแสดงในตาราง 19

ตาราง 19 ผลการคัดเลือกข้อคำถามเข้าสู่แบบวัดฉบับสมบูรณ์

ขั้นตอน	จำนวนข้อคำถาม	จำนวนสถานการณ์
แบบวัดฉบับร่าง	80	20
ผ่านเกณฑ์คุณภาพเชิงสถิติ (ค่าความยากและค่าอำนาจจำแนก)	50	–
คัดเลือกให้ครบ 4 องค์ประกอบในแต่ละสถานการณ์ (แบบวัดฉบับสมบูรณ์)	40	10

หมายเหตุ: แบบวัดฉบับสมบูรณ์ 40 ข้อ เผยแพร่ในรูปแบบออนไลน์ และใช้เป็นเครื่องมือในการดำเนินการตามวัตถุประสงค์ข้อ 2 และข้อ 3

ผลการคัดเลือกในตอนที 1 สรุปได้ว่า ผู้วิจัยพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้นได้สำเร็จ เป็นแบบวัดฉบับสมบูรณ์ 40 ข้อ ใน 10 สถานการณ์ ที่ผ่านการตรวจสอบความตรงเชิงเนื้อหาและการวิเคราะห์คุณภาพรายข้อ ซึ่งตอบวัตถุประสงค์ข้อ 1 ของการวิจัย

ตอนที่ 2 ผลการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด

ผู้วิจัยตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดฉบับสมบูรณ์ 40 ข้อ กับกลุ่มเก็บข้อมูลหลัก ซึ่งเป็นนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร จำนวน 1,265 คน ภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ ผลการตรวจสอบมีรายละเอียดดังนี้

2.1 ผลการวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิม

ผู้วิจัยวิเคราะห์คุณภาพของแบบวัดฉบับสมบูรณ์ 40 ข้อ ตามทฤษฎีการทดสอบแบบดั้งเดิมกับกลุ่มเก็บข้อมูลหลัก 1,265 คน ใน 3 ด้าน ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าความเชื่อมั่น

2.1.1 ผลการวิเคราะห์ค่าความยากและค่าอำนาจจำแนกรายข้อ

ผลการวิเคราะห์ค่าความยากของข้อคำถามทั้ง 40 ข้อ จากกลุ่มเก็บข้อมูลหลัก 1,265 คน พบว่ามีค่าอยู่ในช่วง 0.37 ถึง 0.81 ผ่านเกณฑ์การยอมรับที่กำหนด คือ 0.20 ถึง 0.80 ทุกข้อ และค่าอำนาจจำแนกอยู่ในช่วง 0.26 ถึง 0.65 ผ่านเกณฑ์การยอมรับ คือ ตั้งแต่ 0.20 ขึ้นไป ทุกข้อ ผลการวิเคราะห์รายข้อแสดงในตาราง 20

ตาราง 20 ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของข้อคำถามรายข้อ จำนวน 40 ข้อ จากกลุ่มเก็บข้อมูลหลัก 1,265 คน

ข้อที่	องค์ประกอบที่วัด	ค่าความยาก	ค่าอำนาจจำแนก	การแปลผล
1	การแบ่งย่อยปัญหา	0.734	0.405	ค่อนข้างง่าย จำแนกได้ดีมาก
2	การพิจารณา รูปแบบ	0.695	0.417	ค่อนข้างง่าย จำแนกได้ดีมาก
3	การคิดเชิง นามธรรม	0.736	0.509	ค่อนข้างง่าย จำแนกได้ดีมาก
4	การออกแบบ อัลกอริทึม	0.703	0.504	ค่อนข้างง่าย จำแนกได้ดีมาก

ตาราง 20 (ต่อ)

ข้อที่	องค์ประกอบที่วัด	ค่าความยาก	ค่าอำนาจจำแนก	การแปลผล
5	การแบ่งย่อยปัญหา	0.668	0.475	ค่อนข้างง่าย จำแนกได้ดีมาก
6	การพิจารณา รูปแบบ	0.369	0.376	ค่อนข้างยาก จำแนกได้ดี
7	การคิดเชิง นามธรรม	0.550	0.442	ยากง่ายปานกลาง จำแนกได้ดีมาก
8	การออกแบบ อัลกอริทึม	0.751	0.433	ค่อนข้างง่าย จำแนกได้ดีมาก
9	การแบ่งย่อยปัญหา	0.732	0.483	ค่อนข้างง่าย จำแนกได้ดีมาก
10	การพิจารณา รูปแบบ	0.807	0.504	ง่ายมาก จำแนกได้ดีมาก
11	การคิดเชิง นามธรรม	0.768	0.610	ค่อนข้างง่าย จำแนกได้ดีมาก
12	การออกแบบ อัลกอริทึม	0.661	0.547	ค่อนข้างง่าย จำแนกได้ดีมาก
13	การแบ่งย่อยปัญหา	0.538	0.544	ยากง่ายปานกลาง จำแนกได้ดีมาก
14	การพิจารณา รูปแบบ	0.748	0.626	ค่อนข้างง่าย จำแนกได้ดีมาก
15	การคิดเชิง นามธรรม	0.749	0.611	ค่อนข้างง่าย จำแนกได้ดีมาก
16	การออกแบบ อัลกอริทึม	0.742	0.517	ค่อนข้างง่าย จำแนกได้ดีมาก
17	การแบ่งย่อยปัญหา	0.542	0.422	ยากง่ายปานกลาง จำแนกได้ดีมาก
18	การพิจารณา รูปแบบ	0.692	0.452	ค่อนข้างง่าย จำแนกได้ดีมาก

ตาราง 20 (ต่อ)

ข้อที่	องค์ประกอบที่วัด	ค่าความยาก	ค่าอำนาจจำแนก	การแปลผล
19	การคิดเชิงนามธรรม	0.708	0.576	ค่อนข้างง่าย จำแนกได้ดีมาก
20	การออกแบบอัลกอริทึม	0.450	0.393	ยากง่ายปานกลาง จำแนกได้ดี
21	การแบ่งย่อยปัญหา	0.615	0.544	ค่อนข้างง่าย จำแนกได้ดีมาก
22	การพิจารณารูปแบบ	0.664	0.525	ค่อนข้างง่าย จำแนกได้ดีมาก
23	การคิดเชิงนามธรรม	0.755	0.649	ค่อนข้างง่าย จำแนกได้ดีมาก
24	การออกแบบอัลกอริทึม	0.397	0.258	ยากง่ายปานกลาง จำแนกได้พอใช้
25	การแบ่งย่อยปัญหา	0.726	0.638	ค่อนข้างง่าย จำแนกได้ดีมาก
26	การพิจารณารูปแบบ	0.621	0.479	ค่อนข้างง่าย จำแนกได้ดีมาก
27	การคิดเชิงนามธรรม	0.517	0.544	ยากง่ายปานกลาง จำแนกได้ดีมาก
28	การออกแบบอัลกอริทึม	0.522	0.337	ยากง่ายปานกลาง จำแนกได้ดี
29	การแบ่งย่อยปัญหา	0.474	0.290	ยากง่ายปานกลาง จำแนกได้พอใช้
30	การพิจารณารูปแบบ	0.589	0.600	ยากง่ายปานกลาง จำแนกได้ดีมาก
31	การคิดเชิงนามธรรม	0.611	0.411	ค่อนข้างง่าย จำแนกได้ดีมาก
32	การออกแบบอัลกอริทึม	0.595	0.455	ค่อนข้างง่าย จำแนกได้ดีมาก

ตาราง 20 (ต่อ)

ข้อที่	องค์ประกอบที่วัด	ค่าความยาก	ค่าอำนาจจำแนก	การแปลผล
33	การแบ่งย่อยปัญหา	0.651	0.506	ค่อนข้างง่าย จำแนกได้ดีมาก
34	การพิจารณา รูปแบบ	0.419	0.372	ยากง่ายปานกลาง จำแนกได้ดี
35	การคิดเชิง นามธรรม	0.685	0.466	ค่อนข้างง่าย จำแนกได้ดีมาก
36	การออกแบบ อัลกอริทึม	0.591	0.444	ยากง่ายปานกลาง จำแนกได้ดีมาก
37	การแบ่งย่อยปัญหา	0.674	0.559	ค่อนข้างง่าย จำแนกได้ดีมาก
38	การพิจารณา รูปแบบ	0.693	0.492	ค่อนข้างง่าย จำแนกได้ดีมาก
39	การคิดเชิง นามธรรม	0.427	0.444	ยากง่ายปานกลาง จำแนกได้ดีมาก
40	การออกแบบ อัลกอริทึม	0.455	0.274	ยากง่ายปานกลาง จำแนกได้ดี

หมายเหตุ: N = 1,265 คน เกณฑ์ค่าความยาก 0.20–0.80 และค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป การแปลผลค่าอำนาจจำแนกตาม Ebel และ Frisbie (1991)

จากตาราง 20 ผลการวิเคราะห์รายข้อสะท้อนให้เห็นว่าแบบวัดมีคุณภาพอยู่ในเกณฑ์ที่ยอมรับได้ทั้งด้านค่าความยากและค่าอำนาจจำแนก

ในด้านค่าความยาก ข้อคำถามทั้ง 40 ข้อมีค่าความยากอยู่ในช่วง 0.369 ถึง 0.807 ซึ่งอยู่ในช่วงที่เหมาะสมตามเกณฑ์ที่กำหนด แสดงว่าข้อคำถามไม่มีข้อใดที่ง่ายหรือยากจนเกินไป เมื่อพิจารณาการกระจายตัวของค่าความยาก พบว่ามีทั้งข้อคำถามที่ค่อนข้างง่ายและข้อคำถามที่มีความยากระดับปานกลาง โดยข้อ 10 มีค่าความยากสูงสุดเท่ากับ 0.807 ขณะที่ข้อ 6 มีค่าความยากต่ำสุดเท่ากับ 0.369 การกระจายของระดับความยากดังกล่าวช่วยให้แบบวัดสามารถวัดผู้เรียนที่มีระดับความสามารถแตกต่างกันได้ในระดับหนึ่ง

ในด้านค่าอำนาจจำแนก ข้อคำถามมีค่าอำนาจจำแนกอยู่ในช่วง 0.258 ถึง 0.649 โดยส่วนใหญ่มีค่าอำนาจจำแนกอยู่ในระดับดีถึงดีมาก ข้อคำถามที่มีค่าอำนาจจำแนกสูงสุดคือ ข้อ 23 ซึ่งวัดองค์ประกอบการคิดเชิงนามธรรม มีค่าอำนาจจำแนกเท่ากับ 0.649 สะท้อนความสามารถในการแยกความแตกต่างระหว่างผู้เรียนที่มีทักษะการคิดเชิงคำนวณในระดับสูงและระดับต่ำได้อย่างชัดเจน ขณะที่ข้อ 24 มีค่าอำนาจจำแนกต่ำสุดเท่ากับ 0.258 แต่ยังอยู่ในระดับที่ยอมรับได้ตามเกณฑ์ที่ใช้ในการวิจัย

เมื่อพิจารณาในภาพรวม ข้อคำถามส่วนใหญ่มีระดับความยากค่อนข้างง่ายถึงปานกลาง และมีค่าอำนาจจำแนกอยู่ในระดับดี สะท้อนว่าแบบวัดมีความเหมาะสมสำหรับใช้ประเมินและคัดกรองระดับทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้น โดยสามารถให้ข้อมูลเกี่ยวกับความแตกต่างของผู้เรียนได้อย่างมีประสิทธิภาพ และสนับสนุนการนำผลการประเมินไปใช้ในการวางแผนพัฒนาการเรียนรู้ของผู้เรียนต่อไป

2.1.2 ผลการวิเคราะห์ค่าความเชื่อมั่น

ผลการวิเคราะห์ค่าความเชื่อมั่นของแบบวัดฉบับสมบูรณ์ 40 ข้อ จากกลุ่มเก็บข้อมูลหลัก 1,265 คน ด้วยสูตร KR-20 พบว่ามีค่าเท่ากับ 0.93 ซึ่งสูงกว่าเกณฑ์ที่กำหนดสำหรับแบบวัดที่ใช้ตัดสินใจเกี่ยวกับบุคคลรายคน คือ ตั้งแต่ 0.80 ขึ้นไป แสดงให้เห็นว่าแบบวัดมีความเชื่อมั่นอยู่ในระดับสูง

ค่าความเชื่อมั่น 0.93 สะท้อนว่าข้อคำถามทั้ง 40 ข้อมีความสอดคล้องกันในการวัดทักษะการคิดเชิงคำนวณ และให้ผลการวัดที่มีความคลาดเคลื่อนจากการวัดในระดับต่ำ ค่าความเชื่อมั่นในระดับสูงนี้สนับสนุนว่าแบบวัดสามารถนำไปใช้ประเมินผู้เรียนรายบุคคลได้อย่างเหมาะสม เนื่องจากคะแนนที่ได้มีความน่าเชื่อถือเพียงพอสำหรับการแปลผลและการรายงานผลการประเมิน ทั้งในภาพรวมและเพื่อใช้เป็นข้อมูลประกอบการพัฒนาการเรียนรู้ของผู้เรียน

2.2 ผลการตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ก่อนการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ ผู้วิจัยตรวจสอบข้อตกลงเบื้องต้น 2 ประการ ได้แก่ ความเป็นเอกมิติและความเป็นอิสระเฉพาะที่

2.2.1 ผลการตรวจสอบความเป็นเอกมิติ

ผู้วิจัยตรวจสอบความเป็นเอกมิติด้วย 2 วิธีประกอบกัน

วิธีที่ 1 การวิเคราะห์องค์ประกอบเชิงยืนยันภายในได้โมเดลองค์ประกอบเดียว

ผลการวิเคราะห์พบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ โดยมีค่าดัชนีตามตาราง 21

ตาราง 21 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันภายใต้โมเดลองค์ประกอบเดียว

ดัชนีความสอดคล้อง	ค่าที่พบ	เกณฑ์การยอมรับ	ผลการประเมิน
χ^2/df	3.88	< 5.00	ผ่านเกณฑ์
RMSEA	0.048	< 0.06	ผ่านเกณฑ์
CFI	0.977	> 0.95	ผ่านเกณฑ์
TLI	0.976	> 0.95	ผ่านเกณฑ์
SRMR	0.071	< 0.08	ผ่านเกณฑ์

หมายเหตุ: N = 1,265 คน ประเมินค่าพารามิเตอร์ด้วยวิธี WLSMV

วิธีที่ 2 การพิจารณาอัตราส่วนค่าไอเกน ผลการวิเคราะห์พบว่าค่าไอเกนของปัจจัยแรกต่อปัจจัยที่สองคิดเป็นอัตราส่วน 7.67 ต่อ 1 ซึ่งสูงกว่าเกณฑ์ 3 ต่อ 1 ที่ Reckase (1979) เสนอไว้ ผลดังกล่าวแสดงให้เห็นว่าปัจจัยแรกสามารถอธิบายความแปรปรวนของข้อมูลได้เด่นชัดกว่าปัจจัยอื่น และสนับสนุนสมมติฐานความเป็นเอกมิติของแบบวัด

ผลการตรวจสอบความเป็นเอกมิติด้วยทั้ง 2 วิธีให้ผลสอดคล้องกัน และสนับสนุนว่าแบบวัดมีความเป็นเอกมิติในระดับที่ยอมรับได้ ผลดังกล่าวสะท้อนว่าแม้แบบวัดจะพัฒนาขึ้นจากกรอบแนวคิดที่ประกอบด้วยองค์ประกอบย่อย 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม แต่ข้อคำถามทั้งหมดสามารถสะท้อนความสามารถแฝงร่วมกันได้ในระดับที่เหมาะสม สอดคล้องกับแนวคิดที่มองว่าทักษะการคิดเชิงคำนวณเป็นโครงสร้างที่ประกอบด้วยองค์ประกอบย่อยซึ่งทำงานร่วมกันในการแก้ปัญหา นอกจากนี้ ผลการตรวจสอบดังกล่าวยังสนับสนุนความเหมาะสมของการนำทฤษฎีการตอบสนองข้อสอบแบบมิตินี้มาใช้ในการประมาณค่าพารามิเตอร์ข้อสอบและการวิเคราะห์ในขั้นตอนต่อไป

2.2.2 ผลการตรวจสอบความเป็นอิสระเฉพาะที่

ผู้วิจัยตรวจสอบความเป็นอิสระเฉพาะที่ด้วยสถิติ Q3 ของ Yen (1984) ผลการวิเคราะห์จากคู่ข้อสอบทั้งหมด 780 คู่ พบว่าค่าเฉลี่ยของ Q3 เท่ากับ -0.023 ซึ่งมีค่าใกล้ศูนย์ และเมื่อพิจารณาค่า Q3 ที่ปรับฐานแล้วตามแนวทางของ Christensen, Makransky และ Horton (2017) พบว่ามีคู่ข้อสอบ 4 คู่ จากทั้งหมด 780 คู่ ที่มีค่าเกินเกณฑ์ $|Q3| > 0.20$ คิดเป็นร้อยละ 0.51 ของจำนวนคู่ข้อสอบทั้งหมด

ประเด็นที่มีความสำคัญสำหรับแบบวัดเชิงสถานการณ์ คือ ความเป็นอิสระระหว่างข้อคำถามที่อยู่ภายในสถานการณ์เดียวกัน เนื่องจากข้อคำถาม 4 ข้อในแต่ละสถานการณ์ ใช้

บริบทร่วมกัน จึงอาจเกิดความสัมพันธ์ระหว่างข้อคำถามที่ไม่ได้เกิดจากความสามารถแฝงที่ต้องการวัด ผลการวิเคราะห์พบว่าค่า Q3 ของคู่ข้อคำถามภายในสถานการณ์เดียวกันมีค่าสูงสุดเพียง 0.153 ซึ่งต่ำกว่าเกณฑ์ 0.20 แสดงให้เห็นว่าการใช้สถานการณ์ร่วมกันไม่ได้ก่อให้เกิดความสัมพันธ์ส่วนเกินระหว่างข้อคำถามในระดับที่น่ากังวล

นอกจากนี้ คู่ข้อคำถามที่มีค่า Q3 เกินเกณฑ์ทั้ง 4 คู่ ล้วนเป็นข้อคำถามที่วัดองค์ประกอบการแบ่งย่อยปัญหา (Decomposition) แต่ปรากฏอยู่คนละสถานการณ์ จึงไม่ได้เกิดจากการใช้สถานการณ์ร่วมกันโดยตรง ผลการวิเคราะห์ดังกล่าวสะท้อนว่าข้อคำถามส่วนใหญ่มีความเป็นอิสระต่อกันในระดับที่ยอมรับได้ และสอดคล้องกับสมมติฐานความเป็นอิสระเฉพาะที่ของทฤษฎี การตอบสนองข้อสอบ

ผลการตรวจสอบนี้สนับสนุนว่าข้อมูลมีความเหมาะสมสำหรับการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ และแสดงให้เห็นว่าโครงสร้างแบบวัดที่กำหนดให้แต่ละสถานการณ์ประกอบด้วยข้อคำถาม 4 ข้อซึ่งวัดองค์ประกอบย่อยที่แตกต่างกัน เป็นแนวทางการออกแบบที่เหมาะสมในเชิงจิตมิติ

2.3 ผลการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ

2.3.1 ผลการคัดเลือกโมเดล

ผู้วิจัยเปรียบเทียบโมเดลทฤษฎีการตอบสนองข้อสอบสำหรับข้อมูลแบบทวิภาค เพื่อเลือกโมเดลที่เหมาะสมที่สุด ผลการวิเคราะห์พบว่าโมเดล 3 พารามิเตอร์ไม่สามารถลู่อเข้าได้ แม้กำหนดจำนวนรอบการประมวลผลสูงถึง 2,000 รอบ ซึ่งสอดคล้องกับข้อสังเกตของ Baker (2001) ที่ระบุว่าโมเดล 3 พารามิเตอร์ต้องการข้อสอบที่มีระดับความยากสูงและมีผู้สอบความสามารถต่ำในสัดส่วนที่มากพอ จึงจะประมาณค่าการเดาได้ ผู้วิจัยจึงเปรียบเทียบเฉพาะโมเดล 1 พารามิเตอร์และโมเดล 2 พารามิเตอร์ ผลการเปรียบเทียบแสดงในตาราง 22

ตาราง 22 ผลการเปรียบเทียบโมเดลทฤษฎีการตอบสนองข้อสอบระหว่างโมเดล 1 พารามิเตอร์ และโมเดล 2 พารามิเตอร์

โมเดล	ค่า AIC	ค่า BIC	ผลการเปรียบเทียบ
โมเดล 1 พารามิเตอร์	53,775.30	53,986.20	ค่า AIC และ BIC สูงกว่า
โมเดล 2 พารามิเตอร์	52,832.50	53,243.90	ค่า AIC และ BIC ต่ำกว่า
ผลต่าง	942.80	742.30	โมเดล 2 พารามิเตอร์เหมาะสมกว่า

หมายเหตุ: N = 1,265 คน ผลการทดสอบอัตราส่วนภาวะน่าจะเป็น $\chi^2 = 1,020.82$, df = 39, $p < .001$ โมเดล 3 พารามิเตอร์ไม่สามารถลู่อเข้าได้

จากตาราง 22 โมเดล 2 พารามิเตอร์มีค่า AIC และ BIC ต่ำกว่าโมเดล 1 พารามิเตอร์อย่างชัดเจน และผลการทดสอบอัตราส่วนภาวะน่าจะเป็น (Likelihood Ratio Test) มีนัยสำคัญทางสถิติที่ระดับ .001 ผู้วิจัยจึงเลือกใช้โมเดล 2 พารามิเตอร์ในการวิเคราะห์ข้อมูล

การที่โมเดล 2 พารามิเตอร์มีความเหมาะสมมากกว่าโมเดล 1 พารามิเตอร์ สะท้อนว่าข้อคำถามในแบบวัดมีความแตกต่างกันในด้านอำนาจจำแนก กล่าวคือ โมเดล 1 พารามิเตอร์ ตั้งอยู่บนสมมติฐานว่าข้อคำถามทุกข้อมีอำนาจจำแนกเท่ากัน ขณะที่โมเดล 2 พารามิเตอร์เปิดโอกาส ให้ข้อคำถามแต่ละข้อมีค่าอำนาจจำแนกแตกต่างกันตามข้อมูลเชิงประจักษ์ การเลือกใช้โมเดล 2 พารามิเตอร์จึงสามารถอธิบายลักษณะของข้อคำถามได้สอดคล้องกับข้อมูลมากกว่า และเอื้อให้การ ประเมินค่าความสามารถของผู้สอบมีความเหมาะสมยิ่งขึ้น

ผลการวิเคราะห์ดังกล่าวยังสอดคล้องกับผลการวิเคราะห์ตามทฤษฎีการทดสอบ แบบดั้งเดิมที่พบว่าข้อคำถามมีค่าอำนาจจำแนกแตกต่างกันในช่วง 0.258 ถึง 0.649 แสดงให้เห็นว่า ข้อคำถามแต่ละข้อมีความสามารถในการจำแนกผู้สอบไม่เท่ากัน จึงสนับสนุนความเหมาะสมของการ เลือกใช้โมเดล 2 พารามิเตอร์ในการวิเคราะห์ครั้งนี้

2.3.2 ผลการประมาณค่าพารามิเตอร์ข้อสอบรายข้อ

ผลการประมาณค่าพารามิเตอร์ข้อสอบทั้ง 40 ข้อ ภายใต้โมเดล 2 พารามิเตอร์ พบว่าค่าอำนาจจำแนก (a) อยู่ในช่วง 0.629 ถึง 3.367 และค่าความยาก (b) อยู่ในช่วง -1.161 ถึง 0.716 ค่าพารามิเตอร์รายข้อแสดงในตาราง 23

ตาราง 23 ค่าพารามิเตอร์อำนาจจำแนก (a) และค่าความยาก (b) ของข้อคำถามรายข้อ จำนวน 40 ข้อ ภายใต้โมเดล 2 พารามิเตอร์

ข้อที่	องค์ประกอบที่วัด	ค่าอำนาจจำแนก (a)	ค่าความยาก (b)
1	การแบ่งย่อยปัญหา	1.194	-1.096
2	การพิจารณารูปแบบ	1.202	-0.898
3	การคิดเชิงนามธรรม	1.686	-0.929
4	การออกแบบอัลกอริทึม	1.580	-0.814
5	การแบ่งย่อยปัญหา	1.402	-0.710
6	การพิจารณารูปแบบ	0.964	0.647
7	การคิดเชิงนามธรรม	1.196	-0.245
8	การออกแบบอัลกอริทึม	1.335	-1.114
9	การแบ่งย่อยปัญหา	1.508	-0.960

ตาราง 23 (ต่อ)

ข้อที่	องค์ประกอบที่วัด	ค่าอำนาจจำแนก (a)	ค่าความยาก (b)
10	การพิจารณารูปแบบ	2.038	-1.161
11	การคิดเชิงนามธรรม	2.783	-0.911
12	การออกแบบอัลกอริทึม	1.724	-0.623
13	การแบ่งย่อยปัญหา	1.592	-0.179
14	การพิจารณารูปแบบ	2.930	-0.832
15	การคิดเชิงนามธรรม	2.619	-0.854
16	การออกแบบอัลกอริทึม	1.826	-0.923
17	การแบ่งย่อยปัญหา	1.112	-0.213
18	การพิจารณารูปแบบ	1.291	-0.852
19	การคิดเชิงนามธรรม	2.159	-0.747
20	การออกแบบอัลกอริทึม	1.053	0.213
21	การแบ่งย่อยปัญหา	1.647	-0.459
22	การพิจารณารูปแบบ	1.673	-0.642
23	การคิดเชิงนามธรรม	3.367	-0.838
24	การออกแบบอัลกอริทึม	0.629	0.716
25	การแบ่งย่อยปัญหา	2.738	-0.765
26	การพิจารณารูปแบบ	1.334	-0.525
27	การคิดเชิงนามธรรม	1.671	-0.100
28	การออกแบบอัลกอริทึม	0.756	-0.145
29	การแบ่งย่อยปัญหา	0.762	0.142
30	การพิจารณารูปแบบ	1.987	-0.343
31	การคิดเชิงนามธรรม	1.060	-0.548
32	การออกแบบอัลกอริทึม	1.275	-0.428
33	การแบ่งย่อยปัญหา	1.510	-0.619
34	การพิจารณารูปแบบ	1.037	0.365
35	การคิดเชิงนามธรรม	1.362	-0.793
36	การออกแบบอัลกอริทึม	1.243	-0.413
37	การแบ่งย่อยปัญหา	1.887	-0.653

ตาราง 23 (ต่อ)

ข้อที่	องค์ประกอบที่วัด	ค่าอำนาจจำแนก (a)	ค่าความยาก (b)
38	การพิจารณารูปแบบ	1.486	-0.797
39	การคิดเชิงนามธรรม	1.239	0.284
40	การออกแบบอัลกอริทึม	0.650	0.296

หมายเหตุ: N = 1,265 คน ประเมินค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดส่วนเพิ่ม

จากตาราง 23 พบว่าข้อคำถามทั้ง 40 ข้อมีค่าอำนาจจำแนก (a) เป็นบวกทุกข้อ โดยมีค่าอยู่ในช่วง 0.629 ถึง 3.367 เมื่อพิจารณาตามเกณฑ์ของ Baker (2001) พบว่าข้อคำถามส่วนใหญ่อยู่ในระดับจำแนกได้ปานกลางถึงสูงมาก และมีข้อคำถามจำนวน 5 ข้อที่มีค่าอำนาจจำแนกมากกว่า 2.50 ได้แก่ ข้อ 11 ข้อ 14 ข้อ 15 ข้อ 23 และข้อ 25 แสดงให้เห็นว่าข้อคำถามดังกล่าวสามารถจำแนกความแตกต่างระหว่างผู้สอบที่มีระดับความสามารถแตกต่างกันได้เป็นอย่างดี โดยข้อ 23 ซึ่งวัดองค์ประกอบของการคิดเชิงนามธรรม มีค่าอำนาจจำแนกสูงที่สุด ($a = 3.367$)

สำหรับค่าความยาก (b) พบว่าข้อคำถามทั้ง 40 ข้อมีค่าอยู่ในช่วง -1.161 ถึง 0.716 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ตามเกณฑ์การพิจารณาค่าพารามิเตอร์ความยากของข้อสอบ โดยค่าความยากเฉลี่ยของข้อคำถามอยู่ที่ประมาณ -0.45 และข้อคำถามส่วนใหญ่มีค่าความยากเป็นลบ สะท้อนว่าข้อคำถามโดยรวมมีระดับความยากต่ำกว่าระดับความสามารถเฉลี่ยของกลุ่มผู้สอบเล็กน้อย

นอกจากนี้ แนวโน้มของค่าความยากภายใต้ทฤษฎีการตอบสนองข้อสอบ ยังสอดคล้องกับผลการวิเคราะห์ค่าความยากตามทฤษฎีการทดสอบแบบดั้งเดิม ซึ่งพบว่าข้อคำถามส่วนใหญ่อยู่ในระดับค่อนข้างง่าย แสดงให้เห็นว่าผลการวิเคราะห์จากทั้งสองกรอบแนวคิดให้ข้อค้นพบไปในทิศทางเดียวกัน

ผลการวิเคราะห์พารามิเตอร์ข้อสอบดังกล่าวสะท้อนว่าแบบวัดฉบับนี้ประกอบด้วยข้อคำถามที่มีคุณภาพ สามารถจำแนกผู้เรียนที่มีระดับทักษะการคิดเชิงคำนวณแตกต่างกันได้อย่างมีประสิทธิภาพ และมีระดับความยากที่เหมาะสมกับกลุ่มนักเรียนระดับมัธยมศึกษาตอนต้นที่เป็นกลุ่มเป้าหมายของการวิจัย จึงสนับสนุนความเหมาะสมของแบบวัดสำหรับการนำไปใช้ประเมินและคัดกรองระดับทักษะการคิดเชิงคำนวณของผู้เรียนในบริบทการศึกษาได้อย่างน่าเชื่อถือ

2.3.3 ผลการตรวจสอบความสอดคล้องของโมเดล

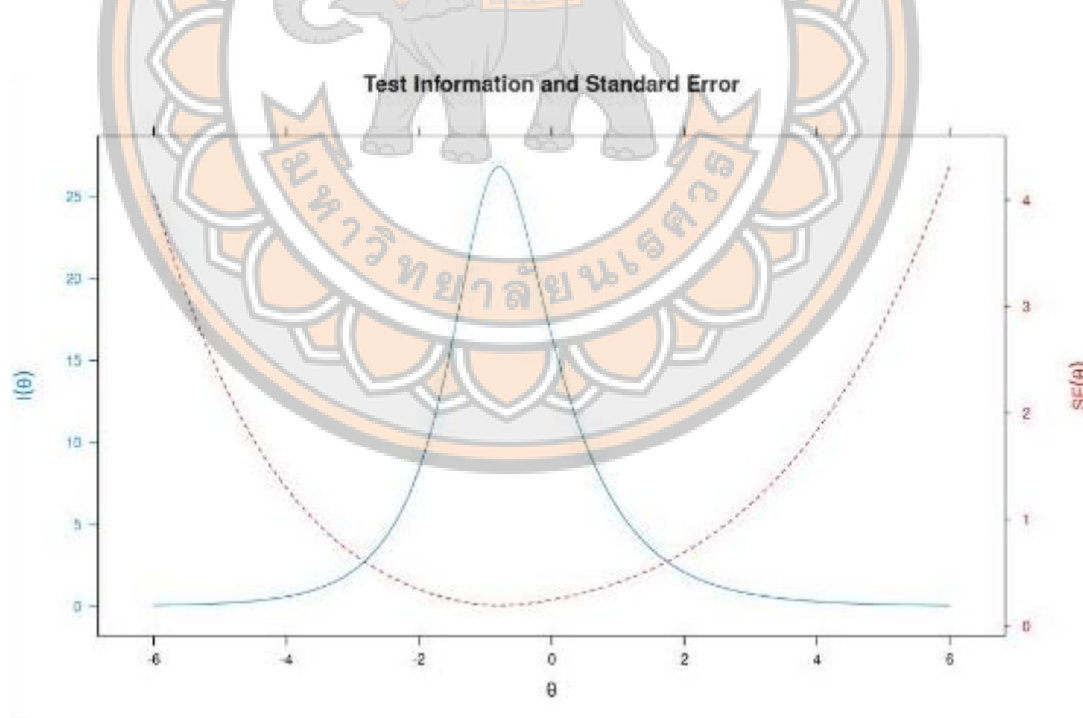
ผู้วิจัยตรวจสอบความสอดคล้องของโมเดล 2 พารามิเตอร์กับข้อมูลทั้งในระดับโมเดลโดยรวมและระดับรายข้อ ผลการตรวจสอบในระดับโมเดลโดยรวมด้วยสถิติ M^2 พบว่าโมเดล

มีความสอดคล้องกับข้อมูลเชิงประจักษ์ โดยมีค่า RMSEA เท่ากับ 0.055 ซึ่งต่ำกว่าเกณฑ์ 0.06 ค่า CFI เท่ากับ 0.958 และค่า TLI เท่ากับ 0.956 ซึ่งสูงกว่าเกณฑ์ 0.95 และค่า SRMSR เท่ากับ 0.049 ซึ่งต่ำกว่าเกณฑ์ 0.08 แสดงให้เห็นว่าโมเดลสามารถอธิบายข้อมูลได้ในระดับที่ดี

ผลการตรวจสอบดังกล่าวสนับสนุนว่าโมเดล 2 พารามิเตอร์มีความเหมาะสมสำหรับการวิเคราะห์ข้อมูลชุดนี้ และสามารถอธิบายความสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับรูปแบบการตอบข้อคำถามได้อย่างสอดคล้องกับข้อมูลเชิงประจักษ์ จึงมีความเหมาะสมสำหรับการประมาณค่าพารามิเตอร์ข้อสอบและการวิเคราะห์ตามกรอบทฤษฎีการตอบสนองข้อสอบในขั้นต่อไป

2.3.4 ผลการวิเคราะห์ฟังก์ชันสารสนเทศของแบบวัด

ผู้วิจัยวิเคราะห์ฟังก์ชันสารสนเทศของแบบวัดเพื่อศึกษาความแม่นยำของการวัดในแต่ละระดับความสามารถ ผลการวิเคราะห์พบว่าแบบวัดให้ค่าสารสนเทศสูงสุดเท่ากับ 25.49 ที่ระดับความสามารถประมาณ -1.00 บนสเกลความสามารถแฝง และมีค่าความคลาดเคลื่อนมาตรฐานของการวัดต่ำที่สุดที่ระดับความสามารถดังกล่าว แสดงให้เห็นว่าแบบวัดมีความแม่นยำในการวัดสูงสุดในช่วงนี้ ผลการวิเคราะห์ฟังก์ชันสารสนเทศของแบบวัดแสดงในภาพ 2



ภาพ 2 ฟังก์ชันสารสนเทศของแบบวัดทักษะการคิดเชิงคำนวณ

จากภาพ 2 ผลการวิเคราะห์ฟังก์ชันสารสนเทศของแบบวัดสะท้อนระดับความแม่นยำของการวัดในแต่ละช่วงความสามารถของผู้สอบ โดยสามารถสรุปผลการวิเคราะห์ได้ดังนี้

ประการที่ 1 แบบวัดให้ความแม่นยำสูงสุดในกลุ่มผู้สอบที่มีความสามารถต่ำกว่าค่าเฉลี่ย ค่าสารสนเทศสูงสุดปรากฏที่ระดับความสามารถประมาณ -1.0 ซึ่งบนสเกลความสามารถมาตรฐานที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 หมายถึงกลุ่มผู้สอบที่มีความสามารถต่ำกว่าค่าเฉลี่ยประมาณ 1 ส่วนเบี่ยงเบนมาตรฐาน เมื่อพิจารณารูปทรงของฟังก์ชันสารสนเทศพบว่า แบบวัดให้สารสนเทศในระดับสูงบริเวณช่วงความสามารถต่ำกว่าค่าเฉลี่ย โดยเฉพาะในช่วงประมาณ -1.5 ถึง 0 แสดงว่าแบบวัดมีความแม่นยำในการวัดนักเรียนที่มีทักษะการคิดเชิงคำนวณอยู่ในระดับค่อนข้างอ่อนถึงต่ำกว่าค่าเฉลี่ย

ประการที่ 2 ช่วงที่แบบวัดให้ความแม่นยำสูงสุดสอดคล้องกับวัตถุประสงค์การนำไปใช้ แบบวัดมีความสามารถในการจำแนกความแตกต่างของนักเรียนที่มีระดับทักษะการคิดเชิงคำนวณต่ำกว่าค่าเฉลี่ยได้อย่างแม่นยำ จึงเหมาะสำหรับการใช้คัดกรองหรือระบุผู้เรียนที่อาจต้องการการส่งเสริมหรือพัฒนาเพิ่มเติม โดยผลการวัดที่ได้สามารถใช้เป็นข้อมูลประกอบการวางแผนพัฒนาการเรียนรู้ได้อย่างเหมาะสม

ประการที่ 3 ค่าความคลาดเคลื่อนมาตรฐานของการวัดต่ำในช่วงที่แบบวัดให้สารสนเทศสูง ฟังก์ชันสารสนเทศของแบบวัดมีความสัมพันธ์ผกผันกับค่าความคลาดเคลื่อนมาตรฐานของการวัด กล่าวคือ บริเวณที่แบบวัดให้ค่าสารสนเทศสูง โดยเฉพาะที่ระดับความสามารถประมาณ -1.0 ซึ่งเป็นจุดที่สารสนเทศมีค่าสูงสุด จะมีค่าความคลาดเคลื่อนมาตรฐานของการวัดต่ำที่สุด แสดงว่าแบบวัดสามารถประมาณความสามารถของนักเรียนในช่วงดังกล่าวได้อย่างแม่นยำ และผลการวัดมีความน่าเชื่อถือในระดับสูง

อย่างไรก็ตาม แบบวัดให้ค่าสารสนเทศลดลงอย่างชัดเจนในกลุ่มผู้สอบที่มีความสามารถสูง โดยเฉพาะตั้งแต่ระดับความสามารถประมาณ 1.5 ขึ้นไป สะท้อนว่าแบบวัดมีความแม่นยำในการจำแนกความแตกต่างของนักเรียนที่มีทักษะการคิดเชิงคำนวณสูงค่อนข้างจำกัด ลักษณะดังกล่าวอาจสัมพันธ์กับการกระจายตัวของค่าความยากข้อสอบ ซึ่งส่วนใหญ่อยู่ในระดับต่ำถึงปานกลาง ดังนั้น หากมีวัตถุประสงค์เพื่อวัดหรือจำแนกผู้เรียนที่มีความสามารถสูง ควรพัฒนาและเพิ่มข้อคำถามที่มีระดับความยากสูงขึ้น เพื่อขยายช่วงการวัดให้ครอบคลุมระดับความสามารถที่กว้างขึ้น

2.4 สรุปผลการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด

ผลการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดฉบับสมบูรณ์ 40 ข้อ จากกลุ่มเก็บข้อมูลหลัก 1,265 คน ทั้งภายใต้ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ สรุปได้ดังตาราง 24

ตาราง 24 สรุปผลการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดฉบับสมบูรณ์ 40 ข้อ

ประเด็นการตรวจสอบ	ผลการวิเคราะห์	ผลการประเมิน
ค่าความยาก (CTT)	0.369–0.807	ผ่านเกณฑ์ทุกข้อ
ค่าอำนาจจำแนก (CTT)	0.398–0.798	ผ่านเกณฑ์ทุกข้อ
ค่าความเชื่อมั่น KR-20	0.93	ระดับสูง
ความเป็นเอกมิติ	RMSEA = 0.048, CFI = 0.977, TLI = 0.976; อัตราส่วนค่าไคเอน 7.67 ต่อ 1	ผ่านเกณฑ์
ความเป็นอิสระเฉพาะที่	ค่า Q3 เฉลี่ย -0.023; คูในสถานการณ์เดียวกันสูงสุด 0.153	ผ่านเกณฑ์
โมเดล IRT ที่เหมาะสม	โมเดล 2 พารามิเตอร์	เลือกใช้ 2 พารามิเตอร์
ค่าอำนาจจำแนก (IRT)	0.629–3.367	ผ่านเกณฑ์
ค่าความยาก (IRT)	-1.161 ถึง 0.716	ผ่านเกณฑ์ทุกข้อ
ฟังก์ชันสารสนเทศสูงสุด	25.49 ที่ระดับความสามารถ -1.00	เหมาะสำหรับการประเมินผู้เรียนที่มีความสามารถต่ำกว่าค่าเฉลี่ย

หมายเหตุ: N = 1,265 คน ทุกค่าในตารางนี้คำนวณจากแบบวัดฉบับสมบูรณ์ 40 ข้อ กับกลุ่มเก็บข้อมูลหลัก 1,265 คน

จากตาราง 24 สรุปได้ว่าแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ที่พัฒนาขึ้นมีคุณสมบัติทางจิตมิติเป็นไปตามเกณฑ์ที่กำหนดทั้งภายใต้ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ โดยข้อคำถามทั้ง 40 ข้อมีค่าความยากและค่าอำนาจจำแนกอยู่ในเกณฑ์ที่ยอมรับได้ และแบบวัดมีความเชื่อมั่นอยู่ในระดับสูง (KR-20 = 0.93) สะท้อนถึงความสม่ำเสมอของการวัดและความสามารถในการจำแนกความแตกต่างระหว่างผู้เรียนได้อย่างเหมาะสม

ผลการวิเคราะห์ภายใต้ทฤษฎีการตอบสนองข้อสอบพบว่า ข้อมูลเป็นไปตามข้อตกลงเบื้องต้นของโมเดลทั้งด้านความเป็นเอกมิติและความเป็นอิสระเฉพาะที่ อีกทั้งโมเดล 2 พารามิเตอร์มีความเหมาะสมกับข้อมูลมากกว่าโมเดล 1 พารามิเตอร์ จึงถูกเลือกใช้ในการประมาณ

ค่าพารามิเตอร์ของข้อสอบ โดยข้อคำถามทุกข้อมีค่าพารามิเตอร์อำนาจจำแนกและค่าพารามิเตอร์ความยากอยู่ในช่วงที่เหมาะสม และผ่านเกณฑ์ความสอดคล้องกับโมเดลที่กำหนด

นอกจากนี้ ฟังก์ชันสารสนเทศของแบบวัดให้ค่าสูงสุดเท่ากับ 25.49 ที่ระดับความสามารถประมาณ -1.00 ซึ่งหมายถึงกลุ่มผู้เรียนที่มีความสามารถต่ำกว่าค่าเฉลี่ย ส่งผลให้แบบวัดมีความแม่นยำในการวัดสูงเป็นพิเศษในช่วงระดับความสามารถดังกล่าว และมีความเหมาะสมสำหรับการใช้คัดกรองหรือระบุผู้เรียนที่อาจต้องการการพัฒนาเพิ่มเติมด้านทักษะการคิดเชิงคำนวณ

โดยสรุป ผลการตรวจสอบคุณสมบัติทางจิตมิติจากทั้งสองกรอบแนวคิดให้หลักฐานสนับสนุนว่าแบบวัดที่พัฒนาขึ้นมีคุณภาพทั้งด้านความตรง ความเชื่อมั่น คุณภาพของข้อคำถาม และความแม่นยำของการวัด ซึ่งสอดคล้องกับวัตถุประสงค์ข้อที่ 2 ของการวิจัย

ตอนที่ 3 ผลการสร้างเกณฑ์ปกติของแบบวัด

ผู้วิจัยสร้างเกณฑ์ปกติของแบบวัดทักษะการคิดเชิงคำนวณ เป็นเกณฑ์ภาพรวมชุดเดียวสำหรับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร โดยใช้ข้อมูลคะแนนจากกลุ่มเก็บข้อมูลหลัก 1,265 คน

3.1 ผลการวิเคราะห์การแจกแจงของคะแนน

ผลการวิเคราะห์สถิติพื้นฐานของคะแนนรวมจากแบบวัด 40 ข้อ ของกลุ่มตัวอย่าง 1,265 คน พบว่ามีค่าเฉลี่ยเท่ากับ 25.02 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 9.62 คะแนน คะแนนต่ำสุดเท่ากับ 5 คะแนน และคะแนนสูงสุดเท่ากับ 40 คะแนน จากคะแนนเต็ม 40 คะแนน

ค่าเฉลี่ย 25.02 คะแนน คิดเป็นประมาณร้อยละ 62.5 ของคะแนนเต็ม สะท้อนว่าคะแนนของกลุ่มตัวอย่างโดยรวมอยู่ในระดับปานกลางเมื่อเทียบกับคะแนนเต็มของแบบวัด นอกจากนี้ ช่วงคะแนนที่พบตั้งแต่ 5 ถึง 40 คะแนน แสดงให้เห็นว่ากลุ่มตัวอย่างมีระดับคะแนนที่หลากหลายและครอบคลุมช่วงคะแนนของแบบวัดค่อนข้างกว้าง

การที่คะแนนมีการกระจายตัวในช่วงกว้างดังกล่าวเป็นลักษณะที่เหมาะสมสำหรับการสร้างเกณฑ์ปกติ เนื่องจากช่วยให้สามารถจัดลำดับและเปรียบเทียบผลการวัดระหว่างผู้เรียนที่มีระดับคะแนนแตกต่างกันได้อย่างชัดเจนมากขึ้น

3.2 ผลการสร้างเกณฑ์ปกติ

ผู้วิจัยแปลงคะแนนดิบเป็นลำดับเปอร์เซ็นต์ไทล์ (Percentile Rank) และคะแนนที (T-score) เพื่อจัดทำเกณฑ์ปกติสำหรับการแปลผลคะแนนของผู้เรียน ดังแสดงในตาราง 25

ตาราง 25 เกณฑ์ปกติของคะแนนแบบวัดทักษะการคิดเชิงคำนวณสำหรับนักเรียนระดับ
มัธยมศึกษาตอนต้นในกรุงเทพมหานคร

คะแนน	ความถี่	ความถี่	Percentile	T-score	ระดับทักษะ
5	6	6	0.24	21.80	ต่ำกว่าเกณฑ์ปกติ
6	9	15	0.83	26.05	ต่ำกว่าเกณฑ์ปกติ
7	17	32	1.86	29.16	ต่ำกว่าเกณฑ์ปกติ
8	24	56	3.48	31.86	ต่ำกว่าเกณฑ์ปกติ
9	33	89	5.73	34.22	ต่ำกว่าเกณฑ์ปกติ
10	34	123	8.38	36.20	ต่ำกว่าเกณฑ์ปกติ
11	53	176	11.82	38.16	ต่ำกว่าเกณฑ์ปกติ
12	27	203	14.98	39.63	ต่ำกว่าเกณฑ์ปกติ
13	39	242	17.59	40.69	อยู่ในเกณฑ์ปกติ
14	17	259	19.80	41.51	อยู่ในเกณฑ์ปกติ
15	18	277	21.19	42.00	อยู่ในเกณฑ์ปกติ
16	25	302	22.89	42.58	อยู่ในเกณฑ์ปกติ
17	31	333	25.10	43.29	อยู่ในเกณฑ์ปกติ
18	18	351	27.04	43.88	อยู่ในเกณฑ์ปกติ
19	22	373	28.62	44.36	อยู่ในเกณฑ์ปกติ
20	34	407	30.83	44.99	อยู่ในเกณฑ์ปกติ
21	33	440	33.48	45.73	อยู่ในเกณฑ์ปกติ
22	24	464	35.73	46.34	อยู่ในเกณฑ์ปกติ
23	40	504	38.26	47.01	อยู่ในเกณฑ์ปกติ
24	29	533	40.99	47.72	อยู่ในเกณฑ์ปกติ
25	40	573	43.72	48.42	อยู่ในเกณฑ์ปกติ
26	40	613	46.88	49.22	อยู่ในเกณฑ์ปกติ

ตาราง 25 (ต่อ)

คะแนน	ความถี่	ความถี่	Percentile	T-score	ระดับทักษะ
26	40	613	46.88	49.22	อยู่ในเกณฑ์ปกติ
27	51	664	50.47	50.12	อยู่ในเกณฑ์ปกติ
28	61	725	54.90	51.23	อยู่ในเกณฑ์ปกติ
29	50	775	59.29	52.35	อยู่ในเกณฑ์ปกติ
30	48	823	63.16	53.36	อยู่ในเกณฑ์ปกติ
31	56	879	67.27	54.47	อยู่ในเกณฑ์ปกติ
32	47	926	71.34	55.63	อยู่ในเกณฑ์ปกติ
33	53	979	75.30	56.84	อยู่ในเกณฑ์ปกติ
34	48	1027	79.29	58.16	อยู่ในเกณฑ์ปกติ
35	46	1073	83.00	59.54	อยู่ในเกณฑ์ปกติ
36	30	1103	86.01	60.81	สูงกว่าเกณฑ์ปกติ
37	55	1158	89.37	62.46	สูงกว่าเกณฑ์ปกติ
38	51	1209	93.56	65.19	สูงกว่าเกณฑ์ปกติ
39	28	1237	96.68	68.36	สูงกว่าเกณฑ์ปกติ
40	28	1265	98.89	72.87	สูงกว่าเกณฑ์ปกติ

หมายเหตุ: N = 1,265 คน; เกณฑ์ปกติสร้างจากกลุ่มตัวอย่างนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร โดยนำเสนอในรูปคะแนนดิบ ลำดับเปอร์เซ็นต์ไทล์ (Percentile Rank) และคะแนนที่ (T-score) เพื่อใช้ในการแปลผลคะแนนรายบุคคล; การแปลผลระดับทักษะอ้างอิงจากคะแนนที่ โดยกำหนดให้ $T < 40$ = ต่ำกว่าเกณฑ์ปกติ, $T = 40-59$ = อยู่ในเกณฑ์ปกติ และ $T \geq 60$ = สูงกว่าเกณฑ์ปกติ

จากตาราง 25 เกณฑ์ปกติที่สร้างขึ้นสามารถใช้แปลผลคะแนนทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้นได้เป็น 3 ระดับ ได้แก่ ระดับต่ำกว่าเกณฑ์ปกติ ระดับอยู่ในเกณฑ์ปกติ และระดับสูงกว่าเกณฑ์ปกติ โดยกำหนดจากคะแนนที่ (T-score) ซึ่งเป็นคะแนนมาตรฐานที่ช่วยให้สามารถเปรียบเทียบผลการประเมินของผู้เรียนกับกลุ่มอ้างอิงได้

ระดับต่ำกว่าเกณฑ์ปกติ คือ ผู้ที่มีคะแนนที่ต่ำกว่า 40 ตรงกับคะแนนดิบตั้งแต่ 5-12 คะแนน ระดับอยู่ในเกณฑ์ปกติ คือ ผู้ที่มีคะแนนที่ตั้งแต่ 40-59 ตรงกับคะแนนดิบ 13-35 คะแนน และระดับสูงกว่าเกณฑ์ปกติ คือ ผู้ที่มีคะแนนที่ตั้งแต่ 60 ขึ้นไป ตรงกับคะแนนดิบ 36-40 คะแนน

3.3 ผลการจำแนกนักเรียนตามระดับทักษะ

เมื่อใช้เกณฑ์ปกติจำแนกนักเรียนกลุ่มตัวอย่าง 1,265 คน ออกเป็น 3 ระดับ พบผลดังตาราง 26

ตาราง 26 จำนวนและร้อยละของนักเรียนจำแนกตามระดับทักษะการคิดเชิงคำนวณ

ระดับทักษะ	ช่วงคะแนนดิบ	จำนวน (คน)	ร้อยละ
สูงกว่าเกณฑ์ปกติ	36-40	192	15.2
อยู่ในเกณฑ์ปกติ	13-35	870	68.8
ต่ำกว่าเกณฑ์ปกติ	5-12	203	16.0
รวม		1,265	100.0

หมายเหตุ: เกณฑ์การแปลผลอ้างอิงจากคะแนนที่ (T-score) ตามแนวทางการสร้างเกณฑ์ปกติของ สมนึก ภัททิยธนี (2562)

จากตาราง 26 พบว่านักเรียนส่วนใหญ่อยู่ในระดับอยู่ในเกณฑ์ปกติ จำนวน 870 คน คิดเป็นร้อยละ 68.8 รองลงมาคือระดับต่ำกว่าเกณฑ์ปกติ จำนวน 203 คน คิดเป็นร้อยละ 16.0 และระดับสูงกว่าเกณฑ์ปกติ จำนวน 192 คน คิดเป็นร้อยละ 15.2 การกระจายของผู้เรียนในทั้ง 3 ระดับ มีสัดส่วนใกล้เคียงกับแนวคิดการแบ่งระดับตามคะแนนมาตรฐานที่ ซึ่งสะท้อนว่าเกณฑ์ปกติที่สร้างขึ้นสามารถใช้จำแนกผู้เรียนออกเป็นกลุ่มที่มีระดับทักษะแตกต่างกันได้อย่างเหมาะสม

ผลการวิเคราะห์ในส่วนนี้ต่อบัณฑิตยุคประสงค์ข้อ 3 ของการวิจัย โดยได้เกณฑ์ปกติในรูปแบบลำดับเปอร์เซ็นต์ไทล์และคะแนนที่ พร้อมเกณฑ์การแปลผลสำหรับใช้จำแนกระดับทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้น

สรุปผลการวิจัย

ผลการวิจัยตามวัตถุประสงค์ทั้ง 3 ประการ สรุปได้ดังนี้

ประเด็นที่ 1 ผลการพัฒนาแบบวัด

ผู้วิจัยพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น โดยเริ่มจากการสร้างแบบวัดฉบับร่างจำนวน 80 ข้อ ใน 20 สถานการณ์ และตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญจำนวน 7 คน ซึ่งข้อคำถามทุกข้อมีค่าดัชนีความสอดคล้องอยู่ในเกณฑ์ที่ยอมรับได้ จากนั้นนำไปทดลองใช้กับนักเรียนจำนวน 150 คน เพื่อวิเคราะห์คุณภาพรายข้อ ผลการวิเคราะห์พบว่ามีค่าความเชื่อมั่น KR-20 เท่ากับ 0.84 และมีข้อคำถามที่ผ่านเกณฑ์จำนวน 50 ข้อ ผู้วิจัยจึงคัดเลือกข้อคำถามให้ครอบคลุมองค์ประกอบของทักษะการคิดเชิงคำนวณทุกด้านในแต่ละสถานการณ์ จนได้แบบวัดฉบับสมบูรณ์จำนวน 40 ข้อ ใน 10 สถานการณ์ และพัฒนาเป็นแบบวัดออนไลน์ที่มีระบบรายงานผลอัตโนมัติ

ประเด็นที่ 2 ผลการตรวจสอบคุณสมบัติทางจิตมิติ

ผลการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดฉบับสมบูรณ์จำนวน 40 ข้อ จากกลุ่มตัวอย่างจำนวน 1,265 คน พบว่าแบบวัดมีคุณภาพผ่านเกณฑ์ที่กำหนดทั้งตามกรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ ภายใต้ทฤษฎีการทดสอบแบบดั้งเดิม พบว่าค่าความยากอยู่ในช่วง 0.369–0.807 ค่าอำนาจจำแนกอยู่ในช่วง 0.398–0.798 และค่าความเชื่อมั่น KR-20 เท่ากับ 0.93 ซึ่งอยู่ในระดับสูง

ภายใต้ทฤษฎีการตอบสนองข้อสอบ พบว่าข้อมูลเป็นไปตามข้อตกลงเบื้องต้นด้านความเป็นเอกมิติและความเป็นอิสระเฉพาะที่ โมเดล 2 พารามิเตอร์เป็นโมเดลที่มีความเหมาะสมกับข้อมูลมากที่สุด ค่าอำนาจจำแนกของข้อคำถามอยู่ในช่วง 0.629–3.367 และค่าความยากอยู่ในช่วง -1.161 ถึง 0.716 นอกจากนี้ แบบวัดให้ค่าสารสนเทศสูงสุดที่ระดับความสามารถประมาณ -1.00 แสดงว่าแบบวัดมีความแม่นยำในการวัดค่อนข้างสูงสำหรับผู้เรียนที่มีระดับความสามารถต่ำกว่าค่าเฉลี่ยเล็กน้อย

ประเด็นที่ 3 ผลการสร้างเกณฑ์ปกติ

ผู้วิจัยสร้างเกณฑ์ปกติของแบบวัดในรูปลำดับเปอร์เซ็นต์ไทล์และคะแนนที่ปกติ เพื่อใช้ในการแปลผลคะแนนของนักเรียนระดับมัธยมศึกษาตอนต้น โดยกำหนดเกณฑ์การแปลผลเป็น 3 ระดับ ได้แก่ ระดับต่ำกว่าเกณฑ์ปกติ (คะแนนดิบ 5–12 คะแนน) ระดับอยู่ในเกณฑ์ปกติ (คะแนนดิบ 13–35 คะแนน) และระดับสูงกว่าเกณฑ์ปกติ (คะแนนดิบ 36–40 คะแนน) เมื่อจำแนกนักเรียนตามเกณฑ์ดังกล่าว พบว่าการกระจายของผู้เรียนในแต่ละระดับมีลักษณะสอดคล้องกับการแจกแจงปกติ และสามารถนำไปใช้เป็นข้อมูลอ้างอิงในการแปลผลคะแนนทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้นได้

บทที่ 5

บทสรุป

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น 2) ตรวจสอบคุณภาพของแบบวัดตามกรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ และ 3) สร้างเกณฑ์ปกติสำหรับใช้ในการแปลผลคะแนนของผู้เรียน กลุ่มตัวอย่างในการเก็บข้อมูลหลักเป็นนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร จำนวน 1,265 คน เครื่องมือที่ใช้ในการวิจัยคือแบบวัดทักษะการคิดเชิงคำนวณเชิงสถานการณ์แบบออนไลน์ จำนวน 40 ข้อ ซึ่งพัฒนาขึ้นตามกรอบแนวคิดทักษะการคิดเชิงคำนวณ 4 องค์ประกอบ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม

ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลโดยใช้ทั้งกรอบทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) รวมทั้งสร้างเกณฑ์ปกติในรูปแบบเปอร์เซ็นต์ไทล์และคะแนนที่ปกติ ผลการวิจัยสรุป อภิปราย และนำเสนอข้อเสนอแนะ ดังต่อไปนี้

สรุปผลการวิจัย

1. ผลการพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น

ผู้วิจัยพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์โดยยึดกรอบแนวคิดทักษะการคิดเชิงคำนวณ 4 องค์ประกอบ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม แบบวัดฉบับร่างประกอบด้วยข้อคำถามจำนวน 80 ข้อ จาก 20 สถานการณ์ และผ่านการตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญจำนวน 7 คน พบว่าข้อคำถามทุกข้อมีค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์อยู่ในเกณฑ์ที่ยอมรับได้

ภายหลังการทดลองใช้กับนักเรียนจำนวน 150 คน ผู้วิจัยวิเคราะห์คุณภาพรายข้อและคัดเลือกข้อคำถามที่มีคุณสมบัติเหมาะสม โดยคำนึงถึงความครอบคลุมขององค์ประกอบที่ต้องการวัดในแต่ละสถานการณ์ร่วมด้วย จนได้แบบวัดฉบับสมบูรณ์จำนวน 40 ข้อ จาก 10 สถานการณ์

ซึ่งครอบคลุมองค์ประกอบของทักษะการคิดเชิงคำนวณทั้ง 4 ด้านอย่างสมดุล และพัฒนาให้อยู่ในรูปแบบออนไลน์ที่สามารถตรวจให้คะแนน ประมวลผล และรายงานผลการประเมินได้โดยอัตโนมัติ

2. ผลการตรวจสอบคุณภาพของแบบวัด

ผลการตรวจสอบคุณภาพของแบบวัดตามกรอบทฤษฎีการทดสอบแบบดั้งเดิม พบว่าข้อคำถามทั้ง 40 ข้อมีความยากอยู่ระหว่าง 0.37 ถึง 0.81 และมีค่าอำนาจจำแนกอยู่ระหว่าง 0.26 ถึง 0.65 โดยผ่านเกณฑ์การพิจารณาทุกข้อ นอกจากนี้ แบบวัดมีค่าความเชื่อมั่น KR-20 เท่ากับ 0.93 แสดงถึงความสม่ำเสมอภายในของแบบวัดในระดับสูง

ผลการตรวจสอบตามกรอบทฤษฎีการตอบสนองข้อสอบ พบว่าข้อมูลเป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ ได้แก่ ความเป็นเอกมิติและความเป็นอิสระเฉพาะที่ โดยผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลปัจจัยเดียวให้ค่าดัชนีความสอดคล้องอยู่ในเกณฑ์ที่ดี และอัตราส่วนค่าไอเกนของปัจจัยที่หนึ่งต่อปัจจัยที่สองเท่ากับ 7.67 ต่อ 1 ซึ่งสนับสนุนสมมติฐานความเป็นเอกมิติของแบบวัด ขณะเดียวกัน ผลการตรวจสอบความเป็นอิสระเฉพาะที่ด้วยสถิติ Q3 พบว่าคู่ข้อสอบส่วนใหญ่มีความสัมพันธ์ของเศษเหลืออยู่ในระดับต่ำ และคู่ข้อสอบที่อยู่ในสถานการณ์เดียวกันไม่มีคู่ใดเกินเกณฑ์ที่กำหนด

เมื่อเปรียบเทียบความเหมาะสมของโมเดล พบว่าโมเดล 2 พารามิเตอร์มีความเหมาะสมกับข้อมูลมากกว่าโมเดล 1 พารามิเตอร์ โดยมีค่า AIC และ BIC ต่ำกว่า และผลการทดสอบอัตราส่วนภาวะน่าจะเป็นมีนัยสำคัญทางสถิติ จึงเลือกใช้โมเดล 2 พารามิเตอร์ในการวิเคราะห์ข้อมูล

ผลการประมาณค่าพารามิเตอร์ข้อสอบภายใต้โมเดล 2 พารามิเตอร์ พบว่าค่าอำนาจจำแนกของข้อสอบอยู่ระหว่าง 0.629 ถึง 3.367 และค่าความยากอยู่ระหว่าง -1.161 ถึง 0.716 โดยข้อสอบทุกข้อมีค่าอำนาจจำแนกเป็นบวกและอยู่ในเกณฑ์ที่ยอมรับได้ นอกจากนี้ แบบวัดมีค่าสารสนเทศสูงสุดเท่ากับ 25.49 ที่ระดับความสามารถประมาณ -1.00 แสดงว่าแบบวัดมีความแม่นยำในการวัดสูง โดยเฉพาะสำหรับผู้เรียนที่มีระดับความสามารถต่ำกว่าค่าเฉลี่ยของกลุ่มอ้างอิง ซึ่งเป็นกลุ่มที่ควรได้รับการส่งเสริมและพัฒนาทักษะการคิดเชิงคำนวณเพิ่มเติม

โดยสรุป แบบวัดฉบับสมบูรณ์มีคุณภาพทางจิตมิติอยู่ในระดับที่น่าพอใจทั้งตามกรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ โดยข้อคำถามมีคุณภาพรายข้ออยู่ในเกณฑ์ที่ยอมรับได้ มีความเชื่อมั่นในระดับสูง และสอดคล้องกับข้อตกลงเบื้องต้นของการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ จึงมีความเหมาะสมสำหรับการนำไปใช้ประเมินทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้น รวมทั้งใช้เป็นข้อมูลประกอบการวางแผนส่งเสริมและพัฒนาผู้เรียนได้อย่างเหมาะสม

3. ผลการสร้างเกณฑ์ปกติสำหรับการแปลผลคะแนน

ผู้วิจัยสร้างเกณฑ์ปกติจากข้อมูลของนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร จำนวน 1,265 คน โดยแปลงคะแนนดิบเป็นลำดับเปอร์เซ็นต์ไทล์และคะแนนที่ปกติ ผลการวิเคราะห์พบว่าคะแนนรวมของนักเรียนมีค่าเฉลี่ยเท่ากับ 25.02 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 9.62 คะแนน มีคะแนนต่ำสุด 5 คะแนน และคะแนนสูงสุด 40 คะแนน

จากการสร้างเกณฑ์ปกติ สามารถกำหนดเกณฑ์การแปลผลออกเป็น 3 ระดับ ได้แก่ ระดับต่ำกว่าเกณฑ์ปกติ ระดับอยู่ในเกณฑ์ปกติ และระดับสูงกว่าเกณฑ์ปกติ โดยระดับต่ำกว่าเกณฑ์ปกติสอดคล้องกับคะแนนดิบ 5-12 คะแนน ระดับอยู่ในเกณฑ์ปกติสอดคล้องกับคะแนนดิบ 13-35 คะแนน และระดับสูงกว่าเกณฑ์ปกติสอดคล้องกับคะแนนดิบ 36-40 คะแนน

เมื่อจำแนกนักเรียนตามเกณฑ์ดังกล่าว พบว่านักเรียนร้อยละ 68.8 อยู่ในระดับปกติ ร้อยละ 16.0 อยู่ในระดับต่ำกว่าเกณฑ์ปกติ และร้อยละ 15.2 อยู่ในระดับสูงกว่าเกณฑ์ปกติ ซึ่งมีลักษณะการกระจายใกล้เคียงกับการแจกแจงปกติ และสะท้อนว่าเกณฑ์ปกติที่พัฒนาขึ้นสามารถใช้เป็นกรอบอ้างอิงสำหรับการแปลผลคะแนนทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้นได้

อภิปรายผลการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น ตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด และสร้างเกณฑ์ปกติสำหรับการแปลผลคะแนน ผลการวิจัยนำไปสู่ประเด็นสำคัญที่ควรอภิปรายทั้งในเชิงทฤษฎี เชิงวิธีวิทยา และเชิงการประยุกต์ใช้ ดังนี้

1. การพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์

ผลการวิจัยพบว่าสามารถพัฒนาแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้นได้สำเร็จ โดยแบบวัดฉบับสมบูรณ์ประกอบด้วยข้อคำถาม 40 ข้อ จาก 10 สถานการณ์ ครอบคลุมองค์ประกอบของทักษะการคิดเชิงคำนวณ 4 ด้าน ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม

ผลดังกล่าวสนับสนุนแนวคิดที่ว่าทักษะการคิดเชิงคำนวณเป็นกระบวนการคิดที่ควรได้รับการประเมินผ่านบริบทของการแก้ปัญหา มากกว่าการวัดความรู้หรือทักษะย่อยแบบแยกส่วน ทั้งนี้ เนื่องจากการนำเสนอข้อคำถามในรูปแบบสถานการณ์ทำให้ผู้เรียนต้องใช้กระบวนการคิดหลายด้าน ร่วมกันในการวิเคราะห์ข้อมูล ตัดสินใจ และเลือกแนวทางแก้ปัญหา ซึ่งสอดคล้องกับแนวคิดของ Wing (2006) ที่อธิบายว่าการคิดเชิงคำนวณเป็นกระบวนการคิดเพื่อแก้ปัญหาอย่างเป็นระบบ และสอดคล้อง

กับกรอบแนวคิดของ Brennan และ Resnick (2012) ที่มองว่าการคิดเชิงคำนวณเป็นความสามารถที่เกิดจากการบูรณาการทักษะย่อยหลายด้านเข้าด้วยกัน

นอกจากนี้ ผลการวิจัยยังสนับสนุนแนวคิดของ Motowidlo, Dunnette และ Carter (1990) ที่เสนอว่าแบบวัดเชิงสถานการณ์สามารถสะท้อนพฤติกรรมหรือกระบวนการคิดของผู้ตอบได้ใกล้เคียงกับบริบทจริงมากกว่าการวัดแบบดั้งเดิม เนื่องจากผู้ตอบต้องพิจารณาสถานการณ์ที่มีความหมายก่อนตัดสินใจเลือกคำตอบ ส่งผลให้การวัดมีความสอดคล้องกับลักษณะการใช้ทักษะในชีวิตจริงมากขึ้น

เมื่อเปรียบเทียบกับงานวิจัยที่ผ่านมา พบว่างานวิจัยในประเทศไทยจำนวนมากมุ่งพัฒนาแบบวัดเชิงสถานการณ์สำหรับทักษะหรือคุณลักษณะเฉพาะด้าน เช่น ทักษะการแก้ปัญหา การรู้เท่าทันดิจิทัล และความเป็นพลเมืองดิจิทัล ขณะที่งานวิจัยที่พัฒนาแบบวัดเชิงสถานการณ์สำหรับการคิดเชิงคำนวณโดยตรงยังมีจำนวนจำกัด โดยเฉพาะในระดับมัธยมศึกษาตอนต้น ดังนั้น งานวิจัยครั้งนี้จึงช่วยขยายองค์ความรู้ด้านการวัดและประเมินผลการคิดเชิงคำนวณในบริบทการศึกษาไทย และแสดงให้เห็นความเป็นไปได้ในการพัฒนาเครื่องมือที่เชื่อมโยงแนวคิดการคิดเชิงคำนวณกับการวัดเชิงสถานการณ์อย่างเป็นระบบ

ประเด็นสำคัญอีกประการหนึ่งคือการพัฒนาแบบวัดในรูปแบบออนไลน์พร้อมระบบรายงานผลอัตโนมัติ ซึ่งสะท้อนแนวโน้มการประเมินผลทางการศึกษาในยุคดิจิทัลที่มุ่งลดข้อจำกัดด้านเวลาและการจัดการข้อมูล ผลการวิจัยจึงไม่ได้ให้เพียงเครื่องมือวัดเท่านั้น แต่ยังนำเสนอแนวทางการประเมินที่สามารถนำไปประยุกต์ใช้ในสภาพจริงของสถานศึกษาได้อย่างมีประสิทธิภาพ

2. คุณสมบัติทางจิตมิติของแบบวัด

ผลการตรวจสอบคุณสมบัติทางจิตมิติพบว่าแบบวัดมีคุณภาพผ่านเกณฑ์ทั้งภายใต้ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ ซึ่งเป็นหลักฐานสำคัญที่สนับสนุนความน่าเชื่อถือของการวัด

ภายใต้ทฤษฎีการทดสอบแบบดั้งเดิม ข้อคำถามทุกข้อมีค่าความยากและค่าอำนาจจำแนกอยู่ในเกณฑ์ที่เหมาะสม อีกทั้งมีค่าความเชื่อมั่น KR-20 เท่ากับ 0.93 ซึ่งอยู่ในระดับสูง สะท้อนว่าข้อคำถามมีความสอดคล้องกันในการวัดคุณลักษณะเดียวกัน ผลดังกล่าวแสดงให้เห็นว่ากระบวนการสร้างข้อคำถามที่อิงกรอบแนวคิดเชิงทฤษฎีอย่างชัดเจนช่วยให้ได้เครื่องมือที่มีคุณภาพและมีความคงเส้นคงวาในการวัด

เมื่อพิจารณาภายใต้ทฤษฎีการตอบสนองข้อสอบ ผลการวิเคราะห์พบว่าข้อมูลเป็นไปตามข้อตกลงเบื้องต้นทั้งด้านความเป็นเอกมิติและความเป็นอิสระเฉพาะที่ แม้แบบวัดจะออกแบบให้ครอบคลุมองค์ประกอบย่อย 4 ด้าน แต่ผลการวิเคราะห์ยืนยันว่าข้อคำถามทั้งหมดสะท้อนความสามารถแฝงร่วมเดียวกัน ผลดังกล่าวมีนัยสำคัญในเชิงทฤษฎี เนื่องจากสนับสนุนมุมมองที่ว่าความคิด

เชิงคำนวณเป็นโครงสร้างเชิงบูรณาการ (integrated construct) มากกว่าการเป็นเพียงการรวมกันของทักษะย่อยที่แยกขาดจากกัน

ประเด็นที่น่าสนใจอีกประการหนึ่งคือผลการตรวจสอบความเป็นอิสระเฉพาะที่ ซึ่งพบว่าคู่ข้อความภายในสถานการณ์เดียวกันไม่มีความสัมพันธ์ส่วนเกินเกินกว่าเกณฑ์ที่ยอมรับได้ ทั้งที่ข้อความใช้บริบทสถานการณ์ร่วมกัน ผลดังกล่าวแสดงให้เห็นว่าการออกแบบให้ข้อความแต่ละข้อวัดองค์ประกอบที่ต่างกันสามารถลดผลกระทบจากการใช้สถานการณ์ร่วมกันได้อย่างมีประสิทธิภาพ และเป็นหลักฐานเชิงประจักษ์ที่สนับสนุนการออกแบบแบบวัดเชิงสถานการณ์ภายใต้กรอบทฤษฎีการตอบสนองข้อสอบ

ผลการเปรียบเทียบโมเดลยังพบว่าโมเดล 2 พารามิเตอร์มีความเหมาะสมกับข้อมูลมากกว่าโมเดล 1 พารามิเตอร์อย่างมีนัยสำคัญทางสถิติ สะท้อนว่าข้อความแต่ละข้อมีอำนาจจำแนกแตกต่างกัน การอนุญาตให้ข้อความมีค่าอำนาจจำแนกเฉพาะของตนเองจึงสามารถอธิบายข้อมูลได้ดีกว่า ผลดังกล่าวสอดคล้องกับธรรมชาติของข้อความเชิงสถานการณ์ซึ่งมีความซับซ้อนและกระบวนการคิดที่แตกต่างกันในแต่ละสถานการณ์

นอกจากนี้ แบบวัดให้สารสนเทศสูงสุดที่ระดับความสามารถประมาณ -1.00 ซึ่งสะท้อนว่าแบบวัดมีความแม่นยำในการประมาณค่าความสามารถของผู้เรียนสูงที่สุดในช่วงที่ระดับทักษะการคิดเชิงคำนวณต่ำกว่าค่าเฉลี่ยของกลุ่มอ้างอิง ผลดังกล่าวชี้ให้เห็นถึงศักยภาพของแบบวัดในการใช้เพื่อการคัดกรองและระบุผู้เรียนที่อาจต้องได้รับการส่งเสริมหรือพัฒนาทักษะเพิ่มเติมได้อย่างมีประสิทธิภาพ

โดยสรุป แบบวัดฉบับสมบูรณ์มีคุณภาพทางจิตมิติที่สอดคล้องกับเกณฑ์ที่ยอมรับได้ทั้งภายใต้กรอบทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ และมีความเหมาะสมสำหรับการนำไปใช้ประเมินทักษะการคิดเชิงคำนวณของนักเรียนระดับมัธยมศึกษาตอนต้นในบริบทการศึกษาไทย

3. การสร้างเกณฑ์ปกติสำหรับการแปลผลคะแนน

ผลการวิจัยสามารถสร้างเกณฑ์ปกติของแบบวัดในรูปลำดับเปอร์เซ็นต์ไทล์และคะแนนที่ปกติสำหรับนักเรียนระดับมัธยมศึกษาตอนต้นได้สำเร็จ และสามารถจำแนกระดับทักษะการคิดเชิงคำนวณออกเป็น 3 ระดับ ได้แก่ ระดับต่ำกว่าเกณฑ์ปกติ ระดับอยู่ในเกณฑ์ปกติ และระดับสูงกว่าเกณฑ์ปกติ

ผลดังกล่าวมีความสำคัญทั้งในเชิงการวัดและเชิงการนำไปใช้ เนื่องจากการประเมินทางการศึกษาจะมีความหมายมากขึ้นเมื่อผู้ใช้สามารถตีความคะแนนของผู้เรียนโดยเปรียบเทียบกับกลุ่มอ้างอิงที่เหมาะสม การสร้างเกณฑ์ปกติจึงช่วยเปลี่ยนคะแนนดิบที่มีความหมายจำกัดให้เป็นสารสนเทศที่สามารถใช้ประกอบการตัดสินใจทางการศึกษาได้

ในเชิงวิชาการ ผลการวิจัยครั้งนี้มีส่วนช่วยขยายองค์ความรู้ด้านการประเมินทักษะการคิดเชิงคำนวณ เนื่องจากงานวิจัยที่ผ่านมาโดยส่วนใหญ่มักมุ่งเน้นการพัฒนาเครื่องมือวัดและการตรวจสอบคุณภาพทางจิตมิติของเครื่องมือ ขณะที่การพัฒนาเกณฑ์ปกติสำหรับการใช้ในการแปลผลคะแนนยังได้รับความสนใจค่อนข้างน้อย การวิจัยครั้งนี้จึงนำเสนอแนวทางการประเมินที่ครอบคลุมตั้งแต่การพัฒนาเครื่องมือ การตรวจสอบคุณภาพทางจิตมิติ ไปจนถึงการสร้างเกณฑ์อ้างอิงสำหรับการตีความผลการประเมิน อันเป็นการเชื่อมโยงกระบวนการวัดผลกับการนำผลไปใช้ทางการศึกษาได้อย่างเป็นระบบมากยิ่งขึ้น

ในเชิงปฏิบัติ เกณฑ์ปกติที่พัฒนาขึ้นช่วยให้ครูสามารถตีความผลการประเมินได้อย่างเป็นระบบ สามารถระบุผู้เรียนที่ควรได้รับการส่งเสริมเพิ่มเติมหรือผู้เรียนที่มีศักยภาพสูงได้อย่างชัดเจน ขณะเดียวกันผู้เรียนสามารถใช้ผลการประเมินเป็นข้อมูลสะท้อนกลับเพื่อวางแผนพัฒนาตนเองได้อย่างเหมาะสม

โดยสรุป ผลการวิจัยครั้งนี้สนับสนุนแนวคิดที่ว่าทักษะการคิดเชิงคำนวณสามารถพิจารณาในลักษณะของความสามารถเชิงบูรณาการที่สะท้อนผ่านการแก้ปัญหาในสถานการณ์ที่มีบริบทใกล้เคียงกับการใช้งานจริง พร้อมทั้งแสดงให้เห็นถึงความเป็นไปได้ในการพัฒนาแบบวัดเชิงสถานการณ์แบบออนไลน์ที่มีคุณภาพทางจิตมิติอยู่ในเกณฑ์ที่ยอมรับได้ นอกจากนี้ การสร้างเกณฑ์ปกติสำหรับการแปลผลคะแนนยังช่วยขยายขอบเขตการใช้ประโยชน์ของผลการประเมินจากรายงานคะแนนไปสู่การตีความผลการวัดและการนำข้อมูลไปใช้ประกอบการพัฒนาผู้เรียนได้อย่างเป็นระบบมากยิ่งขึ้น

ข้อเสนอแนะ

จากผลการวิจัย ผู้วิจัยมีข้อเสนอแนะ 2 ส่วน คือ ข้อเสนอแนะในการนำผลการวิจัยไปใช้ และข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 การนำแบบวัดไปใช้เพื่อการคัดกรองและติดตามผู้เรียน

ผลการวิเคราะห์ฟังก์ชันสารสนเทศของแบบวัดพบว่า แบบวัดให้สารสนเทศสูงสุดที่ระดับความสามารถต่ำกว่าค่าเฉลี่ยของกลุ่มอ้างอิง จึงมีความแม่นยำในการวัดเป็นพิเศษสำหรับผู้เรียนในช่วงความสามารถดังกล่าว ผลการวิจัยนี้สะท้อนว่าแบบวัดมีความเหมาะสมสำหรับการใช้คัดกรองผู้เรียนที่อาจต้องการการส่งเสริมหรือพัฒนาทักษะการคิดเชิงคำนวณเพิ่มเติม โดยสถานศึกษาสามารถนำแบบวัดไปใช้ประเมินระดับทักษะของผู้เรียนในช่วงต้นปีการศึกษา เพื่อประกอบการวางแผนจัดกิจกรรมหรือแนวทางพัฒนาที่เหมาะสมกับผู้เรียนแต่ละกลุ่ม นอกจากนี้ ผลการประเมินยังสามารถใช้เป็นข้อมูลอ้างอิงในการติดตามพัฒนาการของผู้เรียนเมื่อมีการประเมินซ้ำในช่วงเวลาต่อมาได้

1.2 การนำผลการประเมินไปใช้เพื่อวางแผนพัฒนาผู้เรียนรายบุคคล

แบบวัดที่พัฒนาขึ้นสามารถรายงานผลได้ทั้งคะแนนรวมและคะแนนรายองค์ประกอบ ได้แก่ การแบ่งย่อยปัญหา การพิจารณารูปแบบ การคิดเชิงนามธรรม และการออกแบบอัลกอริทึม ดังนั้น ผู้สอนสามารถใช้ผลการประเมินเพื่อวิเคราะห์จุดแข็งและจุดที่ควรพัฒนาของผู้เรียนแต่ละคนได้อย่างเฉพาะเจาะจง และนำข้อมูลดังกล่าวไปใช้ในการออกแบบกิจกรรมการเรียนรู้หรือการให้ความช่วยเหลือที่สอดคล้องกับความต้องการของผู้เรียนแต่ละราย

1.3 การนำผลการประเมินไปใช้ประกอบการพัฒนาการจัดการเรียนการสอน

นอกจากการใช้ผลการประเมินในระดับรายบุคคลแล้ว ผลการประเมินในภาพรวมยังสามารถใช้เป็นข้อมูลสะท้อนคุณภาพการจัดการเรียนการสอนด้านการคิดเชิงคำนวณได้ ผู้สอนและสถานศึกษาสามารถวิเคราะห์รูปแบบคะแนนของผู้เรียนในแต่ละองค์ประกอบ เพื่อระบุประเด็นที่ควรได้รับการส่งเสริมเพิ่มเติม และใช้เป็นข้อมูลประกอบการวางแผนพัฒนาหลักสูตร กิจกรรมการเรียนรู้ หรือการพัฒนาศักยภาพครูผู้สอนในประเด็นที่เกี่ยวข้อง

1.4 การใช้เกณฑ์ปกติเพื่อสนับสนุนการแปลผลคะแนน

เกณฑ์ปกติที่พัฒนาขึ้นจากข้อมูลของนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร สามารถใช้เป็นข้อมูลอ้างอิงสำหรับการแปลผลคะแนนของผู้เรียน โดยช่วยให้การตีความคะแนนมีความหมายมากกว่าการพิจารณาคะแนนดิบเพียงอย่างเดียว ผู้ใช้แบบวัดสามารถพิจารณาระดับความสามารถของผู้เรียนเมื่อเปรียบเทียบกับกลุ่มอ้างอิง และใช้ข้อมูลดังกล่าวประกอบการให้คำแนะนำ การวางแผนพัฒนา หรือการติดตามผลการเรียนรู้ได้อย่างเหมาะสม

1.5 การใช้แบบวัดในระบบการประเมินออนไลน์

แบบวัดที่พัฒนาขึ้นได้รับการออกแบบให้อยู่ในรูปแบบออนไลน์และมีระบบรายงานผลอัตโนมัติ ซึ่งช่วยลดภาระในการตรวจให้คะแนนและการแปลผลของผู้ใช้แบบวัด ทั้งยังเอื้อต่อการนำไปใช้กับผู้เรียนจำนวนมากในระยะเวลาอันสั้น ดังนั้น สถานศึกษาหรือหน่วยงานทางการศึกษาสามารถประยุกต์ใช้แบบวัดเป็นส่วนหนึ่งของระบบการประเมินทักษะการคิดเชิงคำนวณ เพื่อสนับสนุนการติดตามและพัฒนาผู้เรียนได้อย่างมีประสิทธิภาพมากขึ้น

2. ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

2.1 การพัฒนาแบบวัดให้ครอบคลุมระดับความสามารถที่หลากหลายยิ่งขึ้น

ผลการวิเคราะห์ฟังก์ชันสารสนเทศของแบบวัดพบว่า แบบวัดให้ความแม่นยำสูงเป็นพิเศษในช่วงระดับความสามารถต่ำกว่าค่าเฉลี่ยของกลุ่มอ้างอิง ดังนั้น การวิจัยครั้งต่อไปอาจพัฒนาข้อคำถามเพิ่มเติมที่มีระดับความยากสูงขึ้น เพื่อเพิ่มความแม่นยำในการประเมินผู้เรียนที่มีระดับทักษะการคิดเชิงคำนวณสูง และทำให้แบบวัดสามารถครอบคลุมช่วงความสามารถของผู้เรียนได้กว้างยิ่งขึ้น

2.2 การตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดในกลุ่มผู้เรียนที่หลากหลาย

การวิจัยครั้งนี้ดำเนินการกับนักเรียนระดับมัธยมศึกษาตอนต้นในกรุงเทพมหานคร ดังนั้น การวิจัยครั้งต่อไปควรตรวจสอบความคงที่ของโครงสร้างการวัดและคุณสมบัติทางจิตมิติของแบบวัดในกลุ่มผู้เรียนที่มีบริบทแตกต่างกัน เช่น ภูมิภาคอื่นของประเทศ โรงเรียนในสังกัดที่ต่างกัน หรือกลุ่มผู้เรียนที่มีลักษณะทางสังคมและเศรษฐกิจแตกต่างกัน เพื่อเพิ่มหลักฐานสนับสนุนความสามารถในการนำแบบวัดไปใช้กับประชากรที่หลากหลายมากขึ้น

2.3 การศึกษาพัฒนาการของทักษะการคิดเชิงคำนวณจากการจัดการเรียนรู้

แม้ว่าการวิจัยครั้งนี้จะมุ่งพัฒนาและตรวจสอบคุณภาพของแบบวัดเป็นสำคัญ แต่แบบวัดที่พัฒนาขึ้นสามารถนำไปใช้เป็นเครื่องมือในการศึกษาพัฒนาการของทักษะการคิดเชิงคำนวณของผู้เรียนได้ในอนาคต ดังนั้น การวิจัยครั้งต่อไปอาจนำแบบวัดไปใช้ติดตามการเปลี่ยนแปลงของทักษะการคิดเชิงคำนวณก่อนและหลังการจัดการเรียนรู้ หรือในช่วงเวลาต่าง ๆ ของการศึกษา เพื่อศึกษาลักษณะการพัฒนาของผู้เรียนในแต่ละองค์ประกอบของทักษะการคิดเชิงคำนวณ อันจะช่วยให้เกิดความเข้าใจเชิงลึกเกี่ยวกับกระบวนการพัฒนาทักษะดังกล่าว และเป็นประโยชน์ต่อการออกแบบกิจกรรมการเรียนรู้ที่สอดคล้องกับพัฒนาการของผู้เรียนมากยิ่งขึ้น

2.4 การศึกษาความตรงเชิงเกณฑ์และความสัมพันธ์กับตัวแปรทางการศึกษาอื่น

การวิจัยครั้งนี้มุ่งเน้นการพัฒนาแบบวัดและตรวจสอบคุณภาพของเครื่องมือเป็นสำคัญ ดังนั้น การวิจัยครั้งต่อไปควรศึกษาความสัมพันธ์ระหว่างคะแนนจากแบบวัดทักษะการคิดเชิงคำนวณกับตัวแปรอื่นที่เกี่ยวข้อง เช่น ผลสัมฤทธิ์ทางการเรียนด้านวิทยาการคำนวณ ความสามารถในการแก้ปัญหา การคิดเชิงวิพากษ์ หรือสมรรถนะด้านดิจิทัล เพื่อสร้างหลักฐานเพิ่มเติมเกี่ยวกับความตรงเชิงเกณฑ์ของแบบวัด และขยายองค์ความรู้เกี่ยวกับบทบาทของทักษะการคิดเชิงคำนวณในบริบททางการศึกษา

2.5 การต่อยอดสู่การพัฒนากระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

เนื่องจากการวิจัยครั้งนี้ได้ประมาณค่าพารามิเตอร์ข้อสอบภายใต้กรอบทฤษฎีการตอบสนองข้อสอบ และแสดงให้เห็นว่าข้อคำถามมีคุณภาพทางจิตมิติอยู่ในเกณฑ์ที่ยอมรับได้ การวิจัยครั้งต่อไปอาจขยายจำนวนข้อคำถามและพัฒนาคลังข้อสอบที่ครอบคลุมระดับความสามารถที่หลากหลายมากขึ้น เพื่อเป็นพื้นฐานสำหรับการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) ซึ่งสามารถคัดเลือกข้อคำถามให้สอดคล้องกับระดับความสามารถของผู้เรียนแต่ละคนได้อย่างเหมาะสม การพัฒนาในแนวทางดังกล่าวจะช่วยเพิ่มประสิทธิภาพของการวัด ลดภาระในการตอบข้อสอบ และสนับสนุนการประเมินทักษะการคิดเชิงคำนวณรายบุคคลในสภาพแวดล้อมดิจิทัลได้อย่างมีประสิทธิภาพมากยิ่งขึ้น



บรรณานุกรม

บรรณานุกรม

- กระทรวงศึกษาธิการ. (2560). *หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560)*. กรุงเทพฯ: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย.
- ชวาล แพรัตกุล. (2552). *เทคนิคการวัดผล* (พิมพ์ครั้งที่ 7). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- พิชิต ฤทธิ์จรรยา. (2556). *หลักการวัดและประเมินผลการศึกษา* (พิมพ์ครั้งที่ 8). กรุงเทพฯ: แฮสส์ ออฟ เคอร์มิสท์.
- พิสิฐพงศ์ เหมเมฆ, เทิดศักดิ์ สุพันธุ์, และ สมประสงค์ เสนารัตน์. (2567). การพัฒนาล้างข้อสอบวิชาวิทยาศาสตร์และเทคโนโลยีสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้ทฤษฎีการตอบสนองข้อสอบ. *วารสารสถาบันวิจัยและพัฒนา มหาวิทยาลัยราชภัฏชัยภูมิ*, 6(1), 1–15.
- พิเชฐ ศรีสังข์งาม, และ ชัยยศ เดชสุระ. (2564). การพัฒนากิจกรรมการเรียนรู้ตามแนวทางสะเต็มศึกษาเพื่อส่งเสริมทักษะการคิดเชิงคำนวณของผู้เรียนระดับชั้นมัธยมศึกษาตอนต้น. *วารสารวิจัยราชภัฏกรุงเก่า สาขามนุษยศาสตร์และสังคมศาสตร์*, 8(2), 1–15.
- แพรวนภา แก้วกุก, พนิดา ศกุนตนาค, และ อรุมา เจริญสุข. (2566). การพัฒนาแบบวัดเชิงสถานการณ์เพื่อวัดความเป็นพลเมืองดิจิทัลของนักเรียนชั้นมัธยมศึกษาตอนต้น. *Interdisciplinary Academic and Research Journal*, 3(4), 1–20.
- ไพลิน จันทิษฐ์. (2563). การพัฒนาแบบวัดทักษะการคิดแก้ปัญหาในรูปแบบเชิงสถานการณ์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น (วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต). มหาวิทยาลัยศรีนครินทรวิโรฒ, กรุงเทพฯ.
- ยุภารัตน์ พิชสิงห์. (2564). การพัฒนาทักษะการคิดเชิงคำนวณโดยใช้การจัดกิจกรรมการเรียนรู้แบบใช้ปัญหาเป็นฐาน (วิทยานิพนธ์ปริญญาโท). มหาวิทยาลัยมหาสารคาม, มหาสารคาม.
- วรรณกร พรประเสริฐ. (2564). การวิจัยและพัฒนาแบบวัดเชิงสถานการณ์การรู้เท่าทันดิจิทัลของนิสิตนักศึกษาครูในสถาบันอุดมศึกษา. *Journal of Education and Innovation*, 23(2), 1–15.
- วัชรภรณ์ พิงโพธิ์. (2567). *แนวทางการจัดการเรียนการสอนเพื่อส่งเสริมการคิดเชิงคำนวณของนักเรียนมัธยมศึกษาตอนต้น* (ดุษฎีนิพนธ์ปริญญาดุษฎีบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.
- วันเพ็ญ จันทสุวรรณ. (2550). *การวัดและประเมินผลทางการศึกษา*. กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยรามคำแหง.

- ศิริชัย กาญจนวาสี. (2556). *ทฤษฎีการทดสอบแบบดั้งเดิม* (พิมพ์ครั้งที่ 7). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2560). *คู่มือการใช้หลักสูตรรายวิชาพื้นฐานวิทยาศาสตร์ กลุ่มสาระการเรียนรู้วิทยาศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 สาระเทคโนโลยี (วิทยาการคำนวณ) ระดับมัธยมศึกษาตอนต้น*. กรุงเทพฯ: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.
- สมนึก ภัททิยธนี. (2562). *การวัดผลการศึกษา* (พิมพ์ครั้งที่ 9). กทม.: ประสานการพิมพ์.
- สำนักงานเลขาธิการสภาการศึกษา. (2564). *แผนการศึกษาแห่งชาติ พ.ศ. 2560–2579*. กรุงเทพฯ: พรักหวานกราฟฟิค.
- สุพัญญา เจริรัตน์. (2560). *การพัฒนาแบบวัดทักษะการแก้ปัญหาเชิงสถานการณ์สำหรับนักเรียนมัธยมศึกษาตอนต้น (คู่มือฉบับปฐมฤกษ์)*. จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.
- อนุวัติ คุณแก้ว. (2566). *การวัดและประเมินผลการศึกษา* (พิมพ์ครั้งที่ 2). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- Alruwais, N., Wills, G., & Wald, M. (2018). Advantages and challenges of using e-assessment. *International Journal of Information and Education Technology*, 8(1), 34–37. <https://doi.org/10.18178/ijiet.2018.8.1.1008>
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.
- Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). College Park, MD: ERIC Clearinghouse on Assessment and Evaluation.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12: What is involved and what is the role of the computer science education community? *ACM Inroads*, 2(1), 48–54. <https://doi.org/10.1145/1929887.1929905>
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397–479). Reading, MA: Addison-Wesley.

- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. doi:10.1007/s11092-008-9068-5
- Bock, R. D., & Mislevy, R. J. (1982). Adaptive EAP estimation of ability in a microcomputer environment. *Applied Psychological Measurement*, 6(4), 431–444. <https://doi.org/10.1177/014662168200600405>
- Bocconi, S., Chiocciariello, A., Dettori, G., Ferrari, A., & Engelhardt, K. (2016). *Developing computational thinking in compulsory education: Implications for policy and practice*. Luxembourg: Publications Office of the European Union. <https://doi.org/10.2791/792158>
- Brennan, K., & Resnick, M. (2012). *New frameworks for studying and assessing the development of computational thinking*. Paper presented at the Annual Meeting of the American Educational Research Association, Vancouver, Canada.
- Bunderson, C. V., Inouye, D. K., & Olsen, J. B. (1989). The four generations of computerized educational measurement. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 367–407). New York, NY: American Council on Education & Macmillan.
- Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6), 1–29. <https://doi.org/10.18637/jss.v048.i06>
- Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology*, 63(1), 83–117. <https://doi.org/10.1111/j.1744-6570.2009.01163.x>
- Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement*, 41(3), 178–194. <https://doi.org/10.1177/0146621616677520>
- Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory*. Mason, OH: Cengage Learning.

- Dagienė, V., & Futschek, G. (2008). Bebras international contest on informatics and computer literacy: Criteria for good tasks. In R. T. Mittermeir & M. M. Sysło (Eds.), *Informatics education: Supporting computational thinking* (Lecture Notes in Computer Science, Vol. 5090, pp. 19–30). Berlin, Germany: Springer.
https://doi.org/10.1007/978-3-540-69924-8_2
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: Guilford Press.
- Department for Education. (2013). *Computing programmes of study: Key stages 3 and 4. National curriculum in England*. London, UK: Department for Education.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. *Educational Researcher*, 42(1), 38–43.
<https://doi.org/10.3102/0013189X12463051>
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. New York, NY: Routledge.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
<https://doi.org/10.1080/10705519909540118>
- Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology*, 30(1), 17–24.
<https://doi.org/10.1037/h0057123>

- Kim, S., & Lee, Y. (2020). Computational thinking education in South Korea: Policy and practice in K-12 software education. *Informatics in Education*, 19(3), 1–18.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford Press.
- Knuth, D. E. (1997). *The art of computer programming, Volume 1: Fundamental algorithms* (3rd ed.). Reading, MA: Addison-Wesley.
- Korkmaz, Ö., Çakir, R., & Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). *Computers in Human Behavior*, 72, 558–569. <https://doi.org/10.1016/j.chb.2017.01.005>
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151–160. <https://doi.org/10.1007/BF02288391>
- Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., ... Werner, L. (2011). Computational thinking for youth in practice. *ACM Inroads*, 2(1), 32–37. <https://doi.org/10.1145/1929887.1929902>
- Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review*, 37(4), 426–441. <https://doi.org/10.1108/00483480810877598>
- Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Prentice Hall.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385.
- Maydeu-Olivares, A., & Joe, H. (2005). Limited- and full-information estimation and goodness-of-fit testing in 2^n contingency tables: A unified framework. *Journal of the American Statistical Association*, 100(471), 1009–1020. <https://doi.org/10.1198/016214504000002069>
- Maydeu-Olivares, A. (2013). Goodness-of-fit assessment of item response theory models. *Measurement: Interdisciplinary Research and Perspectives*, 11(3), 71–101. <https://doi.org/10.1080/15366367.2013.831680>

- Maydeu-Olivares, A., & Joe, H. (2014). Assessing approximate fit in categorical data analysis. *Multivariate Behavioral Research*, 49(4), 305–328.
<https://doi.org/10.1080/00273171.2014.911075>
- Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York, NY: W. H. Freeman.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, 86(4), 730–740.
<https://doi.org/10.1037/0021-9010.86.4.730>
- Ministry of Education Singapore. (2018). *Computational thinking and coding initiatives in Singapore schools*. Singapore: Ministry of Education.
- Moreno-León, J., Robles, G., & Román-González, M. (2015). Dr. Scratch: Automatic analysis of Scratch projects to assess and foster computational thinking. *RED: Revista de Educación a Distancia*, 46, 1–23. <https://doi.org/10.6018/red/46/10>
- Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology*, 75(6), 640–647. <https://doi.org/10.1037/0021-9010.75.6.640>
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Organisation for Economic Co-operation and Development. (2018). *The future of education and skills: Education 2030*. Paris, France: OECD Publishing.
- Orlando, M., & Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous item response theory models. *Applied Psychological Measurement*, 24(1), 50–64. <https://doi.org/10.1177/01466216000241003>
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York, NY: Basic Books.
- Pellegrino, J. W., & Hilton, M. L. (Eds.). (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: National Academies Press.

- Piaget, J. (1972). *The psychology of the child*. New York, NY: Basic Books.
- Polya, G. (1957). *How to solve it: A new aspect of mathematical method* (2nd ed.). Princeton, NJ: Princeton University Press.
- Posit Team. (2026). *Posit Cloud: Integrated development environment for R* [Web-based software]. Boston, MA: Posit Software, PBC. Retrieved from posit.cloud
- R Core Team. (2026). *R: A language and environment for statistical computing* (Version 4.5.3) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from r-project.org
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danish Institute for Educational Research.
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4(3), 207–230. <https://doi.org/10.3102/10769986004003207>
- Reynolds, C. R., Livingston, R. B., & Willson, V. (2009). *Measurement and assessment in education* (2nd ed.). Upper Saddle River, NJ: Pearson.
- Román-González, M., Pérez-González, J.-C., & Jiménez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the Computational Thinking Test. *Computers in Human Behavior*, 72, 678–691. <https://doi.org/10.1016/j.chb.2016.08.047>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Rovinelli, R. J., & Hambleton, R. K. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, 2, 49–60.
- Selby, C., & Woollard, J. (2013). *Computational thinking: The developing definition*. Southampton, UK: University of Southampton.
- Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, 22, 142–158. <https://doi.org/10.1016/j.edurev.2017.09.003>

- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101.
<https://doi.org/10.2307/1412159>
- Tang, X., Yin, Y., Lin, Q., Hadad, R., & Zhai, X. (2020). Assessing computational thinking: A systematic review of empirical studies. *Computers & Education*, 148, 103798.
<https://doi.org/10.1016/j.compedu.2019.103798>
- Weekley, J. A., & Ployhart, R. E. (Eds.). (2006). *Situational judgment tests: Theory, measurement, and application*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*, 19(3), 188–202.
<https://doi.org/10.1016/j.hrmr.2009.03.007>
- Whitelock, D., & Brasher, A. (2006). Roadmap for e-assessment: Which way now? In M. Danson (Ed.), *Proceedings of the 10th International Computer Assisted Assessment Conference* (pp. 487–501). Loughborough, UK: Loughborough University.
- Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation*, 2(2), 1–3. <https://doi.org/10.7275/ffb1-mm19>
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35.
<https://doi.org/10.1145/1118178.1118215>
- Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1881), 3717–3725.
<https://doi.org/10.1098/rsta.2008.0118>
- Wing, J. M. (2010). *Computational thinking: What and why?* Pittsburgh, PA: Carnegie Mellon University.
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125–145. <https://doi.org/10.1177/014662168400800201>



ภาคผนวก

มหาวิทยาลัยรัตนโกสินทร์

ภาคผนวก ก รายนามผู้เชี่ยวชาญตรวจสอบคุณภาพเครื่องมือวิจัย

1. ผู้ช่วยศาสตราจารย์ ดร.ณัฐวิจิตร เลิศพงษ์จักร อาจารย์ประจำสาขาวิชาเทคโนโลยีดิจิทัล
เพื่อการศึกษา
คณะศึกษาศาสตร์
มหาวิทยาลัยเกษตรศาสตร์
2. อาจารย์ ดร.คมกริช แม่นยำ อาจารย์กลุ่มสาระการเรียนรู้วิทยาศาสตร์และเทคโนโลยี (เทคโนโลยี)
โรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์
ศูนย์วิจัยและพัฒนาการศึกษา
3. อาจารย์ ดร.มุกดา หวังคิด อาจารย์กลุ่มสาระการเรียนรู้วิทยาศาสตร์และเทคโนโลยี (เทคโนโลยี)
โรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์
ศูนย์วิจัยและพัฒนาการศึกษา
4. ดร.รัชฎา เทียมธรรม ผู้อำนวยการ
โรงเรียนประถมสาธิต
มหาวิทยาลัยราชภัฏพระนคร
5. ผู้ช่วยศาสตราจารย์ ดร.สิริกร โตสติ อาจารย์ประจำมหาวิทยาลัย
คณะครุศาสตร์
มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา
6. ผู้ช่วยศาสตราจารย์ ดร.นรินทร์ ตาก้อนทอง รองผู้อำนวยการฝ่ายวิชาการ
สำนักทดสอบทางการศึกษาและจิตวิทยา
มหาวิทยาลัยศรีนครินทรวิโรฒ
7. ดร.ปิยะพงศ์ คล้ายคลึง รองผู้อำนวยการฝ่ายบริหารและแผนงาน
สำนักทดสอบทางการศึกษาและจิตวิทยา
มหาวิทยาลัยศรีนครินทรวิโรฒ

หมายเหตุ: ลำดับที่ 1-3 ผู้เชี่ยวชาญด้านแนวคิดเชิงคำนวณจำนวน 3 ท่าน

ลำดับที่ 4-7 ผู้เชี่ยวชาญด้านการวัดและประเมินผลทางการศึกษาจำนวน 4 ท่าน

ภาคผนวก ข ผลการตรวจสอบความตรงเชิงเนื้อหาของการพัฒนาแบบวัดทักษะการ
คิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์

ตาราง ข.1 ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ (IOC) ของข้อคำถามจำนวน
80 ข้อ

ข้อ ที่	ตัวชี้วัด	ผู้เชี่ยวชาญ							ผลรวม	IOC	ผล การ ประเมิน
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	คนที่ 6	คนที่ 7			
		1	2	3	4	5	6	7			
1	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
2	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
3	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
4	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
5	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
6	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
7	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
8	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
9	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
10	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้

ตาราง ข.1 (ต่อ)

ข้อ ที่	ตัวชี้วัด	ผู้เชี่ยวชาญ							ผลรวม	IOC	ผล การ ประเมิน
		คนที่	คนที่	คนที่	คนที่	คนที่	คนที่	คนที่			
		1	2	3	4	5	6	7			
11	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
12	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
13	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
14	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
15	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
16	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
17	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
18	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
19	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
20	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
21	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
22	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้

ตาราง ข.1 (ต่อ)

ข้อ ที่	ตัวชี้วัด	ผู้เชี่ยวชาญ							ผลรวม	IOC	ผล การ ประเมิน
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	คนที่ 6	คนที่ 7			
23	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
24	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
25	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
26	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
27	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
28	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
29	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
30	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
31	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
32	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
33	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
34	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้

ตาราง ข.1 (ต่อ)

ข้อ ที่	ตัวชี้วัด	ผู้เชี่ยวชาญ							ผลรวม	IOC	ผล การ ประเมิน
		คนที่	คนที่	คนที่	คนที่	คนที่	คนที่	คนที่			
		1	2	3	4	5	6	7			
35	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
36	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
37	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
38	การพิจารณา รูปแบบ	0	1	1	1	1	1	1	6	0.86	ใช้ได้
39	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
40	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
41	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
42	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
43	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
44	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
45	การแบ่งย่อย ปัญหา	1	0	1	1	1	1	1	6	0.86	ใช้ได้
46	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้

ตาราง ข.1 (ต่อ)

ข้อ ที่	ตัวชี้วัด	ผู้เชี่ยวชาญ							ผลรวม	IOC	ผล การ ประเมิน
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	คนที่ 6	คนที่ 7			
47	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
48	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
49	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
50	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
51	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
52	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
53	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
54	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
55	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
56	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
57	การแบ่งย่อย ปัญหา	1	0	1	1	1	1	1	6	0.86	ใช้ได้
58	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้

ตาราง ข.1 (ต่อ)

ข้อ ที่	ตัวชี้วัด	ผู้เชี่ยวชาญ							ผลรวม	IOC	ผล การ ประเมิน
		คนที่	คนที่	คนที่	คนที่	คนที่	คนที่	คนที่			
		1	2	3	4	5	6	7			
59	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
60	การออกแบบ อัลกอริทึม	1	0	1	1	1	1	1	6	0.86	ใช้ได้
61	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
62	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
63	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
64	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
65	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
66	การพิจารณา รูปแบบ	0	0	1	0	1	1	1	4	0.57	ใช้ได้
67	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
68	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
69	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
70	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้

ตาราง ข.1 (ต่อ)

ข้อ ที่	ตัวชี้วัด	ผู้เชี่ยวชาญ							ผลรวม	IOC	ผล การ ประเมิน
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	คนที่ 6	คนที่ 7			
71	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
72	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
73	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
74	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
75	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
76	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
77	การแบ่งย่อย ปัญหา	1	1	1	1	1	1	1	7	1.00	ใช้ได้
78	การพิจารณา รูปแบบ	1	1	1	1	1	1	1	7	1.00	ใช้ได้
79	การคิดเชิง นามธรรม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
80	การออกแบบ อัลกอริทึม	1	1	1	1	1	1	1	7	1.00	ใช้ได้
รวม	80 ข้อ	-	-	-	-	-	-	-		0.57– 1.00	ผ่าน ทั้งหมด

ภาคผนวก ค ตัวอย่างสถานการณ์และข้อคำถามของแบบวัดทักษะการคิดเชิงคำนวณ ในรูปแบบเชิงสถานการณ์แบบออนไลน์

สถานการณ์ที่ 1: การเตรียมตัวสอบปลายภาค

ปริศนาต้องสอบปลายภาคจำนวน 8 วิชา ใช้ระยะเวลา 2 สัปดาห์ และแต่ละวิชามีเนื้อหา
จำนวน 5-10 บท

ข้อ	แนวคิดที่วัด	แบบวัดทักษะ
1	Decomposition	ปริศนาต้องเตรียมสอบปลายภาค 8 วิชา ภายในเวลา 2 สัปดาห์ และแต่ละวิชามีหลายบทวิธีใดเหมาะสมที่สุดในการเริ่มต้นเตรียมสอบ? ก. เลือกอ่านเฉพาะวิชาที่ยากก่อน แล้วจึงอ่านวิชาอื่นภายหลัง ข. อ่านหนังสือทุกวิชาสลับกันไป โดยไม่แบ่งหัวข้อหรือวางแผนเวลา ค. แยกเนื้อหาแต่ละวิชาเป็นหัวข้อย่อย และวางแผนเวลาอ่านแต่ละส่วน ง. อ่านหนังสือตามลำดับวิชาในตารางเรียน โดยไม่แยกเนื้อหาเป็นส่วนย่อย <u>เฉลย: ค</u>
2	Pattern Recognition	จากการทำข้อสอบเก่า ปริศนาพบว่าข้อสอบวิชาวิทยาศาสตร์มักออกเกี่ยวกับการทดลองในบท 3, 5, 7 และ 9 ซึ่งเป็นบทที่มีการทดลอง ปริศนาควรวางแผนทบทวนอย่างไร? ก. อ่านทุกบทด้วยเวลาที่ใกล้เคียงกัน โดยไม่พิจารณาข้อมูลจากข้อสอบเดิม ข. เน้นทบทวนบทที่มีการทดลอง และฝึกทำโจทย์ลักษณะเดียวกับข้อสอบ ค. ท่องจำชื่อการทดลองทั้งหมด โดยไม่เน้นการฝึกทำโจทย์เพิ่มเติม ง. เลือกอ่านเฉพาะบทที่เข้าใจง่าย เพื่อให้ทบทวนได้ครบทุกบท <u>เฉลย: ข</u>
3	Abstraction	ในการจัดสรุปเนื้อหาเพื่อทบทวน ปริศนาควรเลือกบันทึกข้อมูลแบบใด? ก. จัดเนื้อหาทั้งหมดตามลำดับที่เรียน โดยไม่ตัดรายละเอียดใดออก ข. จัดแนวคิดหลัก สูตรสำคัญ และตัวอย่างที่จำเป็นต่อการทำความเข้าใจ ค. จัดชื่อบทและเลขหน้า เพื่อใช้เปิดอ่านจากหนังสือเรียนภายหลัง ง. จัดทุกสิ่งที่ครูอธิบายในคาบเรียน โดยไม่คัดเลือกข้อมูลสำคัญ <u>เฉลย: ข</u>

ข้อ	แนวคิดที่วัด	แบบวัดทักษะ
4	Algorithm	ขั้นตอนที่มีประสิทธิภาพในการท่องจำสูตรคณิตศาสตร์คืออะไร? ก. ทำโจทย์ทันที → ดูเฉลย → จำคำตอบจากตัวอย่าง ข. อ่านสูตรหลายครั้ง → ท่องจำสูตร → ทำแบบฝึกหัดตามตัวอย่าง ค. ท่องสูตรทุกวัน → ทำข้อสอบเก่า → หยุดอ่านเมื่อจำได้ ง. ทำความเข้าใจสูตร → เชื่อมโยงกับตัวอย่าง → ฝึกใช้ → ทบทวน เฉลย: ง

สถานการณ์ที่ 8: การจัดการปัญหาการบูลลี่ในโรงเรียน โรงเรียนแห่งหนึ่งพบว่ามึนักเรียนบางกลุ่มถูกล้อเลียน กลั่นแกล้ง และคุกคามทั้งทางวาจาและทางออนไลน์ ทำให้ผู้ถูกรังแกไม่อยากมาโรงเรียน ครูแนะแนวได้แต่งตั้ง “ทีมเพื่อนช่วยเพื่อน” ให้มีหน้าที่ ร่วมกันวิเคราะห์ปัญหา วางแผนป้องกัน และเสนอแนวทางแก้ไขปัญหาการบูลลี่ในโรงเรียนอย่างเป็นระบบ นักเรียนในทีมต้องใช้ข้อมูลจริงจากโรงเรียน เช่น รายงานเหตุการณ์ ความเห็นของเพื่อน และการสังเกตพฤติกรรม เพื่อออกแบบแนวทางจัดการปัญหานี้

ข้อ	แนวคิดที่วัด	แบบวัดทักษะ
29	Decomposition	เมื่อทีมเพื่อนช่วยเพื่อนเริ่มดำเนินงานเพื่อจัดการปัญหาการบูลลี่ในโรงเรียน แนวทางใดต่อไปนี้จะช่วยการ แยกปัญหาออกเป็นส่วนย่อยตามแนวคิดเชิงคำนวณได้เหมาะสมที่สุด? ก. กำหนดมาตรการลงโทษผู้กระทำ กำหนดกฎระเบียบของโรงเรียน และประกาศนโยบายห้ามบูลลี่ให้ทุกคนรับทราบพร้อมกัน ข. แยกพิจารณาประเภทของการบูลลี่ ผู้ที่เกี่ยวข้อง สถานที่และช่วงเวลาที่เกิดเหตุ รวมถึงผลกระทบที่เกิดกับผู้ถูกรังแก ค. จัดกิจกรรมรณรงค์ต่อต้านการบูลลี่ จัดอบรมคุณธรรมจริยธรรม และเชิญผู้ปกครองเข้าร่วมรับฟังแนวทางของโรงเรียน ง. รับฟังความคิดเห็นของนักเรียนทุกระดับชั้น สอบถามความรู้สึกของผู้เกี่ยวข้อง และรายงานปัญหาให้ฝ่ายบริหารโรงเรียนทราบ เฉลย: ข

ข้อ	แนวคิดที่วัด	แบบวัดทักษะ
30	Pattern Recognition	จากการรวบรวมข้อมูล ทีมเพื่อนช่วยเพื่อนพบว่า เหตุการณ์บูลลี่มักเกิดซ้ำในช่วงพักกลางวัน บริเวณหลังอาคารเรียน และมีนักเรียนกลุ่มเดิมตกเป็นผู้ถูกระทำบ่อยครั้ง ข้อใดเป็นการสรุปข้อมูลตามแนวคิดการพิจารณารูปแบบได้เหมาะสมที่สุด ? ก. ปัญหาการบูลลี่เกิดขึ้นจากพฤติกรรมเฉพาะบุคคลและไม่สามารถคาดการณ์รูปแบบการเกิดเหตุได้ ข. ปัญหาการบูลลี่เป็นเหตุการณ์ทั่วไปในโรงเรียนและไม่จำเป็นต้องวิเคราะห์เชิงข้อมูลเพิ่มเติม ค. การบูลลี่เกิดจากความขัดแย้งส่วนตัวของนักเรียนและควรแก้ไขเป็นรายกรณีเท่านั้น ง. การบูลลี่มีรูปแบบการเกิดซ้ำที่สัมพันธ์กับช่วงเวลา สถานที่ และกลุ่มนักเรียนบางกลุ่มอย่างชัดเจน เฉลย: ง
31	Abstraction	หากทีมเพื่อนช่วยเพื่อนต้องออกแบบแนวทางป้องกันการบูลลี่ที่สามารถใช้ได้กับทั้งโรงเรียน ข้อใดควรถูกเลือกเป็นสาระสำคัญเพื่อใช้เป็นหลักในการดำเนินงาน? ก. รายชื่อผู้ถูกรู้ในแต่ละเหตุการณ์ พร้อมรายละเอียดพฤติกรรมของผู้กระทำในแต่ละครั้ง ข. หลักการสำคัญในการป้องกัน เช่น การสร้างพื้นที่ปลอดภัย ช่องทางการแจ้งเหตุ และระบบช่วยเหลือเพื่อนอย่างต่อเนื่อง ค. ประวัติส่วนตัว ภูมิหลังครอบครัว และผลการเรียนของนักเรียนที่เกี่ยวข้องกับเหตุการณ์ ง. คำพูดหรือข้อความที่ใช้ในการบูลลี่ พร้อมวันเวลาและสถานที่ที่เกิดเหตุอย่างละเอียด เฉลย: ข
32	Algorithm	ทีมเพื่อนช่วยเพื่อนต้องเสนอขั้นตอนการจัดการเมื่อพบเหตุบูลลี่ เพื่อให้สามารถนำไปใช้ได้อย่างเป็นระบบและต่อเนื่อง ขั้นตอนใดต่อไปนี้เรียงลำดับได้เหมาะสมที่สุด? ก. ให้การช่วยเหลือผู้ถูกรู้ทันที รับแจ้งเหตุจากนักเรียน ตรวจสอบข้อมูล และติดตามผลในภายหลัง ข. รับแจ้งเหตุจากนักเรียน ตรวจสอบและประเมินข้อมูล ให้การช่วยเหลือผู้ถูกรู้ และติดตามผลเพื่อป้องกันการเกิดซ้ำ ค. ตรวจสอบข้อมูลจากครูและเพื่อน แจ้งฝ่ายบริหาร ให้การช่วยเหลือและรายงานผลเมื่อสิ้นภาคเรียน

ข้อ	แนวคิดที่วัด	แบบวัดทักษะ
		ง. รับฟังความคิดเห็นของนักเรียนทั้งหมด ให้คำแนะนำทั่วไป จัดกิจกรรมรณรงค์ และติดตามผลเป็นครั้งคราว เฉลย: ข

สถานการณ์ที่ 9 : การจัดการน้ำท่วมในชุมชนริมแม่น้ำ ชุมชนนิเวศน์ตั้งอยู่ใกล้แม่น้ำและประสบปัญหาน้ำท่วมเป็นประจำในช่วงฤดูฝน คณะกรรมการชุมชนร่วมกับนักเรียนในพื้นที่ที่ต้องการวางแผนรับมือกับสถานการณ์น้ำท่วม โดยใช้ข้อมูลระดับน้ำ ปริมาณฝน และประสบการณ์ที่ผ่านมา เพื่อเตรียมการแจ้งเตือนและลดความเสียหายต่อบ้านเรือนและทรัพย์สินของประชาชน

ข้อ	แนวคิดที่วัด	แบบวัดทักษะ
33	Decomposition	หากต้องการวางแผนรับมือปัญหาน้ำท่วมในชุมชนอย่างเป็นระบบ นักเรียนควรแบ่งงานหลักของการจัดการน้ำท่วมออกเป็นส่วนย่อยใดจึงจะเหมาะสมที่สุด? ก. เผื่อระวังระดับน้ำ แจ้งเตือนประชาชน เตรียมอพยพ และประเมินผลหลังน้ำลด ข. ซ่อมแซมถนน ปลุกต้นไม้ เก็บขยะ และจัดกิจกรรมรณรงค์สิ่งแวดล้อม ค. แจกถุงยังชีพทุกครัวเรือน ตั้งโรงครัวเคลื่อนที่ และจัดงานช่วยเหลือผู้ประสบภัย ง. สร้างเขื่อนขนาดใหญ่ และย้ายชุมชนไปยังพื้นที่ใหม่ทั้งหมด เฉลย: ก
34	Pattern Recognition	จากข้อมูลย้อนหลังพบว่าทุกครั้งที่ปริมาณฝนสะสมเกินเกณฑ์หนึ่งระดับน้ำในแม่น้ำจะเพิ่มสูงและเกิดน้ำท่วมในชุมชน นักเรียนควรใช้ข้อมูลนี้เพื่อวางแผนรับมืออย่างไร? ก. รอให้เกิดน้ำท่วมก่อนแล้วจึงเริ่มแจ้งเตือนประชาชน โดยผ่านระบบ SMS เตือนภัย ข. กำหนดเกณฑ์ปริมาณฝนเพื่อใช้เป็นสัญญาณเตือนล่วงหน้าแก่ชุมชน ค. ลดการเก็บข้อมูลในปีถัดไปเพราะรูปแบบเกิดซ้ำทุกปี ง. แจ้งเตือนเฉพาะครัวเรือนที่อยู่ใกล้แม่น้ำมากที่สุดเท่านั้น เฉลย: ข

ข้อ	แนวคิดที่วัด	แบบวัดทักษะ
35	Abstraction	หากต้องออกแบบระบบแจ้งเตือนน้ำท่วมสำหรับชุมชน ข้อมูลใดควรถูกเลือกเป็นข้อมูลหลักของระบบเพื่อให้การแจ้งเตือนมีประสิทธิภาพ? ก. สีของแม่น้ำ กลิ่นน้ำ และเสียงการไหลของน้ำ ข. ระดับน้ำ ปริมาณฝน ตำแหน่งพื้นที่เสี่ยง และเวลา ค. ประวัติการทอ่งเทียวมแม่น้ำของชุมชน ง. ชื่อผู้อยู่อาศัย จำนวนสมาชิกในครอบครัว และอาชีพ เฉลย: ข
36	Algorithm	36. เพื่อให้การแจ้งเตือนน้ำท่วมในชุมชนเป็นไปอย่างเป็นระบบ ควรจัดลำดับขั้นตอนการดำเนินงานอย่างไรให้เหมาะสมที่สุด? ขั้นตอนที่กำหนด 1. ตรวจสอบข้อมูลปริมาณฝนและระดับน้ำ 2. เปรียบเทียบข้อมูลกับเกณฑ์ที่กำหนด 3. แจ้งเตือนประชาชนในพื้นที่เสี่ยง 4. ติดตามสถานการณ์และปรับการแจ้งเตือน ก. 1 → 2 → 3 → 4 ข. 2 → 1 → 3 → 4 ค. 3 → 1 → 2 → 4 ง. 4 → 3 → 2 → 1 เฉลย: ก

ภาคผนวก ง ผลการคัดเลือกข้อคำถามตามทฤษฎีการทดสอบแบบดั้งเดิม

ตาราง ง.1 ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) ของข้อคำถามจำนวน 80 ข้อ จากกลุ่มทดลองใช้เครื่องมือ จำนวน 150 คน

ข้อ ที่	สถานการณ์ ที่	องค์ประกอบ	p (ความยาก)	r (อำนาจจำแนก)	ผลการ ตัดสินใจ
1	1	การแบ่งย่อยปัญหา	0.70	0.50	คัดเก็บ
2	1	การพิจารณารูปแบบ	0.70	0.35	คัดเก็บ
3	1	การคิดเชิงนามธรรม	0.68	0.35	คัดเก็บ
4	1	การออกแบบอัลกอริทึม	0.80	0.35	คัดเก็บ
5	2	การแบ่งย่อยปัญหา	0.63	0.60	คัดเก็บ
6	2	การพิจารณารูปแบบ	0.98	0.05	ตัดทิ้ง
7	2	การคิดเชิงนามธรรม	0.78	0.15	ตัดทิ้ง
8	2	การออกแบบอัลกอริทึม	0.54	0.33	คัดเก็บ
9	3	การแบ่งย่อยปัญหา	0.69	0.63	คัดเก็บ
10	3	การพิจารณารูปแบบ	0.66	0.53	คัดเก็บ
11	3	การคิดเชิงนามธรรม	0.63	0.60	คัดเก็บ
12	3	การออกแบบอัลกอริทึม	0.64	0.38	คัดเก็บ
13	4	การแบ่งย่อยปัญหา	0.76	0.28	คัดเก็บ
14	4	การพิจารณารูปแบบ	0.64	0.68	คัดเก็บ
15	4	การคิดเชิงนามธรรม	0.68	0.60	คัดเก็บ
16	4	การออกแบบอัลกอริทึม	0.73	0.25	คัดเก็บ
17	5	การแบ่งย่อยปัญหา	0.60	0.55	คัดเก็บ
18	5	การพิจารณารูปแบบ	0.76	0.38	คัดเก็บ
19	5	การคิดเชิงนามธรรม	0.79	0.23	คัดเก็บ
20	5	การออกแบบอัลกอริทึม	0.49	0.28	คัดเก็บ
21	6	การแบ่งย่อยปัญหา	0.98	0.00	ตัดทิ้ง
22	6	การพิจารณารูปแบบ	0.86	0.28	ตัดทิ้ง
23	6	การคิดเชิงนามธรรม	0.45	0.10	ตัดทิ้ง

ตาราง ง.1 (ต่อ)

ข้อ ที่	สถานการณ์ ที่	องค์ประกอบ	p (ความยาก)	r (อำนาจจำแนก)	ผลการ ตัดสินใจ
24	6	การออกแบบอัลกอริทึม	0.78	0.20	คัดเก็บ
25	7	การแบ่งย่อยปัญหา	0.76	0.38	คัดเก็บ
26	7	การพิจารณารูปแบบ	0.66	0.53	คัดเก็บ
27	7	การคิดเชิงนามธรรม	0.73	0.40	คัดเก็บ
28	7	การออกแบบอัลกอริทึม	0.56	0.38	คัดเก็บ
29	8	การแบ่งย่อยปัญหา	0.60	0.45	คัดเก็บ
30	8	การพิจารณารูปแบบ	0.70	0.50	คัดเก็บ
31	8	การคิดเชิงนามธรรม	0.66	0.43	คัดเก็บ
32	8	การออกแบบอัลกอริทึม	0.70	0.40	คัดเก็บ
33	9	การแบ่งย่อยปัญหา	0.71	0.53	คัดเก็บ
34	9	การพิจารณารูปแบบ	0.71	0.28	คัดเก็บ
35	9	การคิดเชิงนามธรรม	0.73	0.25	คัดเก็บ
36	9	การออกแบบอัลกอริทึม	0.98	0.05	ตัดทิ้ง
37	10	การแบ่งย่อยปัญหา	0.64	0.38	คัดเก็บ
38	10	การพิจารณารูปแบบ	0.70	0.45	คัดเก็บ
39	10	การคิดเชิงนามธรรม	0.54	0.33	คัดเก็บ
40	10	การออกแบบอัลกอริทึม	0.80	0.20	คัดเก็บ
41	11	การแบ่งย่อยปัญหา	0.70	0.35	คัดเก็บ
42	11	การพิจารณารูปแบบ	0.71	0.38	คัดเก็บ
43	11	การคิดเชิงนามธรรม	0.74	0.33	คัดเก็บ
44	11	การออกแบบอัลกอริทึม	0.64	0.38	คัดเก็บ
45	12	การแบ่งย่อยปัญหา	0.96	0.08	ตัดทิ้ง
46	12	การพิจารณารูปแบบ	0.93	0.15	ตัดทิ้ง
47	12	การคิดเชิงนามธรรม	0.93	0.15	ตัดทิ้ง
48	12	การออกแบบอัลกอริทึม	0.96	0.08	ตัดทิ้ง
49	13	การแบ่งย่อยปัญหา	0.99	0.03	ตัดทิ้ง
50	13	การพิจารณารูปแบบ	0.95	0.05	ตัดทิ้ง

ตาราง ง.1 (ต่อ)

ข้อ ที่	สถานการณ์ ที่	องค์ประกอบ	p (ความยาก)	r (อำนาจจำแนก)	ผลการ ตัดสินใจ
51	13	การคิดเชิงนามธรรม	0.99	0.03	ตัดทิ้ง
52	13	การออกแบบอัลกอริทึม	0.95	0.10	ตัดทิ้ง
53	14	การแบ่งย่อยปัญหา	0.98	0.05	ตัดทิ้ง
54	14	การพิจารณารูปแบบ	0.99	0.03	ตัดทิ้ง
55	14	การคิดเชิงนามธรรม	0.88	0.05	ตัดทิ้ง
56	14	การออกแบบอัลกอริทึม	0.98	0.05	ตัดทิ้ง
57	15	การแบ่งย่อยปัญหา	0.98	0.05	ตัดทิ้ง
58	15	การพิจารณารูปแบบ	0.91	0.13	ตัดทิ้ง
59	15	การคิดเชิงนามธรรม	0.96	0.03	ตัดทิ้ง
60	15	การออกแบบอัลกอริทึม	0.93	0.10	ตัดทิ้ง
61	16	การแบ่งย่อยปัญหา	0.50	0.20	คัดเก็บ
62	16	การพิจารณารูปแบบ	0.75	0.40	คัดเก็บ
63	16	การคิดเชิงนามธรรม	0.73	0.45	คัดเก็บ
64	16	การออกแบบอัลกอริทึม	0.70	0.35	คัดเก็บ
65	17	การแบ่งย่อยปัญหา	0.73	0.35	คัดเก็บ
66	17	การพิจารณารูปแบบ	0.75	0.30	คัดเก็บ
67	17	การคิดเชิงนามธรรม	0.70	0.45	คัดเก็บ
68	17	การออกแบบอัลกอริทึม	0.64	0.38	คัดเก็บ
69	18	การแบ่งย่อยปัญหา	0.98	0.00	ตัดทิ้ง
70	18	การพิจารณารูปแบบ	0.75	0.15	ตัดทิ้ง
71	18	การคิดเชิงนามธรรม	0.98	0.00	ตัดทิ้ง
72	18	การออกแบบอัลกอริทึม	0.86	0.28	ตัดทิ้ง
73	19	การแบ่งย่อยปัญหา	0.96	0.03	ตัดทิ้ง
74	19	การพิจารณารูปแบบ	0.96	0.08	ตัดทิ้ง
75	19	การคิดเชิงนามธรรม	0.96	-0.03	ตัดทิ้ง
76	19	การออกแบบอัลกอริทึม	0.96	-0.03	ตัดทิ้ง
77	20	การแบ่งย่อยปัญหา	0.76	0.33	คัดเก็บ

ตาราง ง.1 (ต่อ)

ข้อ ที่	สถานการณ์ ที่	องค์ประกอบ	p (ความยาก)	r (อำนาจจำแนก)	ผลการ ตัดสินใจ
78	20	การพิจารณารูปแบบ	0.73	0.40	คัดเก็บ
79	20	การคิดเชิงนามธรรม	0.71	0.43	คัดเก็บ
80	20	การออกแบบอัลกอริทึม	0.54	0.33	คัดเก็บ

หมายเหตุ: 1. เกณฑ์การคัดเลือกข้อคำถามรายข้อ ได้แก่ ค่าความยาก (p) อยู่ระหว่าง 0.20–0.80 และค่าอำนาจจำแนก (r) ตั้งแต่ 0.20 ขึ้นไป (Ebel & Frisbie, 1991)

2. ผลการวิเคราะห์คุณภาพรายข้อพบว่า มีข้อคำถามที่ผ่านเกณฑ์จำนวน 50 ข้อ จากทั้งหมด 80 ข้อ

ตาราง ง.2 สรุปผลการคัดเลือกข้อคำถามของแบบวัดทักษะการคิดแก้ปัญหาตามแนวคิด
เชิงคำนวณจำแนกตามองค์ประกอบ

องค์ประกอบ	จำนวนข้อก่อน คัดเลือก	จำนวนข้อหลัง คัดเลือก	หมายเลขข้อที่ได้รับการคัดเลือก
การย่อปัญหา	20	10	1,9,13,17,25,37,41,61,65,77
การพิจารณารูปแบบ	20	10	2,10,14,18,26,38,42,62,66,78
การคิดเชิงนามธรรม	20	10	3,11,15,19,27,39,43,63,67,79
การออกแบบอัลกอริทึม	20	10	4,12,16,20,28,40,44,64,68,80
รวม	80	40	

หมายเหตุ: การคัดเลือกข้อคำถามจากแบบวัดฉบับร่างจำนวน 80 ข้อ เพื่อพัฒนาเป็นแบบวัดฉบับสมบูรณ์จำนวน 40 ข้อ ดำเนินการ 2 ขั้นตอน ได้แก่

(1) การคัดเลือกตามเกณฑ์คุณภาพรายข้อ ประกอบด้วย ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ (IOC) ค่าความยาก (p) และค่าอำนาจจำแนก (r) ซึ่งส่งผลให้คงเหลือข้อคำถามที่ผ่านเกณฑ์จำนวน 50 ข้อ

(2) การคัดเลือกตามโครงสร้างของแบบวัด โดยพิจารณาให้แต่ละสถานการณ์ประกอบด้วย ข้อคำถามที่ครอบคลุมทักษะการคิดเชิงคำนวณครบทั้ง 4 องค์ประกอบ ได้แก่ การแยกส่วนปัญหา การหารูปแบบ การคิดเชิงนามธรรม และการออกแบบขั้นตอนวิธี

ผลการคัดเลือกในขั้นตอนสุดท้าย ทำให้ได้แบบวัดฉบับสมบูรณ์จำนวน 40 ข้อ ครอบคลุม 10 สถานการณ์ โดยแต่ละสถานการณ์ประกอบด้วยข้อคำถามจำนวน 4 ข้อ ซึ่งวัดทักษะการคิดเชิงคำนวณครบทั้ง 4 องค์ประกอบภายใต้บริบทของสถานการณ์เดียวกัน

ตาราง ง.3 ผลการวิเคราะห์ค่าความเชื่อมั่นของแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ทั้งฉบับ

เครื่องมือวิจัย	จำนวน ข้อคำถาม	จำนวนกลุ่ม ทดลอง	ค่าความเชื่อมั่น (KR-20)
แบบวัดทักษะการคิด แก้ปัญหาตามแนวคิด เชิงคำนวณ	40	150	0.84

จากตาราง ง.3 พบว่า ผลการวิเคราะห์ค่าความเชื่อมั่นของแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์จำนวน 40 ข้อ กับกลุ่มทดลองใช้เครื่องมือจำนวน 150 คน ด้วยสูตรคูเดอร์-ริชาร์ดสัน 20 (KR-20) มีค่าความเชื่อมั่นเท่ากับ 0.844 แสดงว่าแบบวัดฉบับนี้มีความเชื่อมั่นอยู่ในระดับสูง มีคุณภาพและเสถียรภาพเพียงพอที่จะนำไปใช้ในการเก็บรวบรวมข้อมูลเพื่อการวิจัยต่อไปได้

ภาคผนวก จ ผลการตรวจสอบคุณภาพของแบบวัดฉบับสมบูรณ์ตามทฤษฎีการทดสอบแบบดั้งเดิม

ตารางที่ จ.1 ผลการวิเคราะห์ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) รายข้อตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ของแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์แบบออนไลน์ จำนวน 40 ข้อ

ข้อที่	ค่าความยากง่าย (p)	ค่าอำนาจจำแนก (r)	การแปล ความหมาย	คุณภาพ
1	0.7335968	0.4045625	ค่อนข้างง่าย	จำแนกได้ดีมาก
2	0.6948617	0.4172383	ค่อนข้างง่าย	จำแนกได้ดีมาก
3	0.7359684	0.5093574	ค่อนข้างง่าย	จำแนกได้ดีมาก
4	0.7027668	0.5042387	ค่อนข้างง่าย	จำแนกได้ดีมาก
5	0.6679842	0.4750278	ค่อนข้างง่าย	จำแนกได้ดีมาก
6	0.3691700	0.3760158	ค่อนข้างยาก	จำแนกได้ดี
7	0.5501976	0.4417047	ยากง่ายปานกลาง	จำแนกได้ดีมาก
8	0.7509881	0.4329715	ค่อนข้างง่าย	จำแนกได้ดีมาก
9	0.7320158	0.4825995	ค่อนข้างง่าย	จำแนกได้ดีมาก
10	0.8071146	0.5037912	ง่ายมาก	จำแนกได้ดีมาก
11	0.7675889	0.6102881	ค่อนข้างง่าย	จำแนกได้ดีมาก
12	0.6608696	0.5470010	ค่อนข้างง่าย	จำแนกได้ดีมาก
13	0.5383399	0.5443535	ยากง่ายปานกลาง	จำแนกได้ดีมาก
14	0.7478261	0.6257794	ค่อนข้างง่าย	จำแนกได้ดีมาก
15	0.7486166	0.6114786	ค่อนข้างง่าย	จำแนกได้ดีมาก
16	0.7415020	0.5168274	ค่อนข้างง่าย	จำแนกได้ดีมาก
17	0.5415020	0.4222021	ยากง่ายปานกลาง	จำแนกได้ดีมาก
18	0.6924901	0.4522081	ค่อนข้างง่าย	จำแนกได้ดีมาก
19	0.7083004	0.5767090	ค่อนข้างง่าย	จำแนกได้ดีมาก
20	0.4498024	0.3934307	ยากง่ายปานกลาง	จำแนกได้ดี
21	0.6150198	0.5435166	ค่อนข้างง่าย	จำแนกได้ดีมาก

ตารางที่ จ.1 (ต่อ)

ข้อที่	ค่าความยากง่าย (p)	ค่าอำนาจจำแนก (r)	การแปล ความหมาย	คุณภาพ
22	0.6640316	0.5253305	ค่อนข้างง่าย	จำแนกได้ดีมาก
23	0.7549407	0.6494860	ค่อนข้างง่าย	จำแนกได้ดีมาก
24	0.3968379	0.2583627	ยากง่ายปานกลาง	จำแนกได้พอใช้
25	0.7256917	0.6381127	ค่อนข้างง่าย	จำแนกได้ดีมาก
26	0.6205534	0.4786487	ค่อนข้างง่าย	จำแนกได้ดีมาก
27	0.5169960	0.5439506	ยากง่ายปานกลาง	จำแนกได้ดีมาก
28	0.5217391	0.3371427	ยากง่ายปานกลาง	จำแนกได้ดี
29	0.4735178	0.2903817	ยากง่ายปานกลาง	จำแนกได้พอใช้
30	0.5889328	0.5997730	ยากง่ายปานกลาง	จำแนกได้ดีมาก
31	0.6110672	0.4113438	ค่อนข้างง่าย	จำแนกได้ดีมาก
32	0.5952569	0.4552285	ค่อนข้างง่าย	จำแนกได้ดีมาก
33	0.6513834	0.5056726	ค่อนข้างง่าย	จำแนกได้ดีมาก
34	0.4189723	0.3723663	ยากง่ายปานกลาง	จำแนกได้ดี
35	0.6845850	0.4660112	ค่อนข้างง่าย	จำแนกได้ดีมาก
36	0.5905138	0.4437103	ยากง่ายปานกลาง	จำแนกได้ดีมาก
37	0.6743083	0.5587518	ค่อนข้างง่าย	จำแนกได้ดีมาก
38	0.6932806	0.4916315	ค่อนข้างง่าย	จำแนกได้ดีมาก
39	0.4268775	0.4435076	ยากง่ายปานกลาง	จำแนกได้ดีมาก
40	0.4545455	0.2738534	ยากง่ายปานกลาง	จำแนกได้ดี

จากตารางที่ จ.1 พบว่า ผลการวิเคราะห์สถิติรายข้อตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ของแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์ จำนวน 40 ข้อ มีค่าความยากง่าย (p) รายข้ออยู่ระหว่าง 0.37 ถึง 0.81 แสดงว่าข้อสอบมีความยากอยู่ในระดับยากปานกลางถึงค่อนข้างง่าย และมีค่าอำนาจจำแนก (r) รายข้ออยู่ระหว่าง 0.26 ถึง 0.65 แสดงว่าข้อสอบทุกข้อมีประสิทธิภาพในการจำแนกผู้สอบได้ดีถึงดีมาก ซึ่งอยู่ในเกณฑ์ที่เหมาะสมและมีคุณภาพตามมาตรฐานทุกข้อคำถาม

ตาราง จ.2 ผลการวิเคราะห์ค่าความเชื่อมั่นของแบบวัดทักษะการคิดแก้ปัญหาตามแนวคิด
เชิงคำนวณในรูปแบบเชิงสถานการณ์ ทั้งฉบับ

เครื่องมือวิจัย	จำนวน ข้อคำถาม	จำนวนกลุ่ม ทดลอง	ค่าความเชื่อมั่น (KR-20)
แบบวัดทักษะการคิด แก้ปัญหาตามแนวคิด เชิงคำนวณ	40	1,265	0.93

จากตาราง จ.2 พบว่า ผลการวิเคราะห์ค่าความเชื่อมั่นของแบบวัดทักษะการคิดเชิงคำนวณ
ในรูปแบบเชิงสถานการณ์ ทั้งฉบับ จำนวน 40 ข้อ กับกลุ่มทดลองหลักจำนวน 1,265 คน ด้วยสูตรคู
เตอร์-ริชาร์ดสัน 20 (KR-20) มีค่าความเชื่อมั่นเท่ากับ 0.93 แสดงว่าแบบวัดฉบับนี้มีความเชื่อมั่นอยู่
ในระดับสูงมาก มีความสอดคล้องภายในและความคงเส้นคงวาทางสถิติในระดับดีเยี่ยม พร้อมสำหรับ
นำผลคะแนนที่ได้ไปดำเนินการวิเคราะห์ขั้นสูงต่อไป

ภาคผนวก ฉ ผลการวิเคราะห์คุณภาพแบบวัดตามทฤษฎีการตอบสนองข้อสอบ (IRT)

ตาราง ฉ.1 ผลการเปรียบเทียบโมเดลการวัดตามทฤษฎีการตอบสนองข้อสอบ IRT (1PL vs 2PL)

โมเดล	-2 Log Likelihood	AIC	BIC	จำนวนพารามิเตอร์
1PL		53,774.55	53,985.41	41
2PL		52,829.64	53,241.07	80
3PL	ไม่ลู่อเข้า	ไม่ลู่อเข้า	ไม่ลู่อเข้า	120

หมายเหตุ: Likelihood Ratio Test (LRT) ระหว่าง 1PL และ 2PL: $\chi^2 = 1,022.912$, $df = 39$, $p < .001$ — โมเดล 2PL ดีกว่าอย่างมีนัยสำคัญ

ตาราง ฉ.2 ค่าพารามิเตอร์ (a, b) และผลการตรวจสอบ Item Fit ($S-\chi^2$, RMSEA, p-Holm) ของข้อคำถามทั้ง 40 ข้อ ภายใต้โมเดล IRT 2PL

ข้อ	องค์ประกอบ	a	b	$S-\chi^2$	df	RMSEA	p-Holm	Item Fit
1	Decomp.	1.194	-1.096	48.601	32	0.0203	0.393	ผ่าน
2	Pattern Rec.	1.202	-0.898	79.014	32	0.0341	< .001	ผ่าน
3	Abstraction	1.686	-0.929	36.263	29	0.0141	0.729	ผ่าน
4	Algorithm	1.580	-0.814	46.489	30	0.0209	0.391	ผ่าน
5	Decomp.	1.402	-0.710	53.363	31	0.0239	0.158	ผ่าน
6	Pattern Rec.	0.964	0.647	46.237	32	0.0188	0.545	ผ่าน
7	Abstraction	1.196	-0.245	31.959	31	0.0049	0.729	ผ่าน
8	Algorithm	1.335	-1.114	77.160	31	0.0343	< .001	ผ่าน
9	Decomp.	1.508	-0.960	59.063	30	0.0277	0.030	ผ่าน
10	Pattern Rec.	2.038	-1.161	38.288	28	0.0170	0.729	ผ่าน

ตาราง จ.2 (ต่อ)

ข้อ	องค์ประกอบ	a	b	S- χ^2	df	RMSEA	p-Holm	Item Fit
11	Abstraction	2.783	-0.911	41.524	25	0.0229	0.303	ผ่าน
12	Algorithm	1.724	-0.623	66.549	30	0.0310	0.004	ผ่าน
13	Decomp.	1.592	-0.179	48.512	30	0.0221	0.282	ผ่าน
14	Pattern Rec.	2.930	-0.832	34.647	24	0.0187	0.664	ผ่าน
15	Abstraction	2.619	-0.854	34.241	27	0.0146	0.729	ผ่าน
16	Algorithm	1.826	-0.923	70.606	28	0.0347	< .001	ผ่าน
17	Decomp.	1.112	-0.213	73.167	31	0.0328	0.001	ผ่าน
18	Pattern Rec.	1.291	-0.852	70.613	32	0.0309	0.003	ผ่าน
19	Abstraction	2.159	-0.747	38.399	28	0.0171	0.729	ผ่าน
20	Algorithm	1.053	0.213	61.488	32	0.0270	0.032	ผ่าน
21	Decomp.	1.647	-0.459	51.392	31	0.0228	0.218	ผ่าน
22	Pattern Rec.	1.673	-0.642	69.781	30	0.0324	0.002	ผ่าน
23	Abstraction	3.367	-0.838	26.373	24	0.0088	0.729	ผ่าน
24	Algorithm	0.629	0.716	51.983	32	0.0222	0.242	ผ่าน
25	Decomp.	2.738	-0.765	34.455	26	0.0160	0.729	ผ่าน
26	Pattern Rec.	1.334	-0.525	104.665	31	0.0434	< .001	ผ่าน
27	Abstraction	1.671	-0.100	51.242	30	0.0237	0.184	ผ่าน
28	Algorithm	0.756	-0.145	98.961	33	0.0398	< .001	ผ่าน
29	Decomp.	0.762	0.142	74.056	32	0.0322	0.001	ผ่าน
30	Pattern Rec.	1.987	-0.343	63.624	29	0.0307	0.006	ผ่าน
31	Abstraction	1.060	-0.548	60.521	32	0.0266	0.039	ผ่าน

ตาราง จ.2 (ต่อ)

ข้อ	องค์ประกอบ	a	b	S- χ^2	df	RMSEA	p-Holm	Item Fit
32	Algorithm	1.275	-0.428	45.619	31	0.0193	0.526	ผ่าน
33	Decomp.	1.510	-0.619	72.242	31	0.0324	0.001	ผ่าน
34	Pattern Rec.	1.037	0.365	124.136	32	0.0477	< .001	ผ่าน
35	Abstraction	1.362	-0.793	45.671	32	0.0184	0.555	ผ่าน
36	Algorithm	1.243	-0.413	54.898	31	0.0247	0.113	ผ่าน
37	Decomp.	1.887	-0.653	90.208	29	0.0409	< .001	ผ่าน
38	Pattern Rec.	1.486	-0.797	51.126	30	0.0236	0.184	ผ่าน
39	Abstraction	1.239	0.284	63.928	31	0.0290	0.012	ผ่าน
40	Algorithm	0.650	0.296	42.980	33	0.0155	0.729	ผ่าน

หมายเหตุ: 1. $n = 1,265$ คน วิเคราะห์ด้วยแพ็คเกจ mirt (Chalmers, 2012) ในโปรแกรมภาษา R เวอร์ชัน 4.5.3 (R Core Team, 2026) ผ่านระบบปฏิบัติการคลาวด์ Posit Cloud (Posit Team, 2026) ค่า a = ค่าอำนาจจำแนก (Discrimination Parameter), b = ค่าความยาก (Difficulty Parameter), S- χ^2 = สถิติความสอดคล้องระดับรายข้อของ Orlando และ Thissen (2000), RMSEA = Root Mean Square Error of Approximation ของ S- χ^2 , p-Holm = p-value หลังการปรับค่าด้วยวิธี Holm-Bonferroni (Holm, 1979)

2. เกณฑ์การประเมิน Item Fit ตามคำแนะนำของ Maydeu-Olivares (2013) และ Maydeu-Olivares และ Joe (2014) ใช้ดัชนี $RMSEA \leq 0.05$ เป็นเกณฑ์หลัก เนื่องจากดัชนีนี้ไม่ไวต่อขนาดกลุ่มตัวอย่างเช่นเดียวกับ p-value ซึ่งเหมาะสมสำหรับการวิเคราะห์ในกลุ่มตัวอย่างขนาดใหญ่ ($n = 1,265$) ที่สถิติ S- χ^2 มีอำนาจการทดสอบสูงเกินไป

3. ผลการประเมินข้อคำถามทั้ง 40 ข้อมีค่า RMSEA อยู่ในช่วง 0.005–0.048 (ค่าเฉลี่ย 0.025) ต่ำกว่าเกณฑ์ 0.05 ทุกข้อ จึงผ่านเกณฑ์ Item Fit

ตาราง น.3 ผลการตรวจสอบความสอดคล้องของโมเดลในระดับโมเดลโดยรวม (Global Fit)

ดัชนีความสอดคล้อง	ค่าที่ได้	เกณฑ์
M^2	3,561.32	—
df	740	—
RMSEA	0.055	< 0.06
CFI	0.958	> 0.95
TLI	0.956	> 0.95
SRMSR	0.049	< 0.08

ตาราง น.4 ผลการตรวจสอบความเป็นเอกมิติ (CFA Single-Factor Model)

ดัชนีความสอดคล้อง	ค่าที่ได้	เกณฑ์
χ^2	2,869.738	—
df	740	—
χ^2/df	3.88	< 5.00
RMSEA	0.048	< 0.06
CFI	0.977	> 0.95
TLI	0.976	> 0.95
SRMR	0.071	< 0.08
Eigenvalue Ratio	7.67 : 1	> 3 : 1

หมายเหตุ: วิเคราะห์ด้วยเมทริกซ์ Tetrachoric Correlation สำหรับข้อมูลทวิภาค (0–1) พบค่า Eigenvalue ลำดับที่ 1 และ 2 เท่ากับ 16.944 และ 2.209 ตามลำดับ คิดเป็นอัตราส่วน (Ratio) เท่ากับ 7.67 : 1 ซึ่งสูงกว่าเกณฑ์ 3 : 1 จึงสนับสนุนสมมติฐานความเป็นเอกมิติของแบบวัด

ตาราง น.5 ผลการตรวจสอบความเป็นอิสระเฉพาะที่ (Q3 Statistics)

สถิติพรรณนาของค่า Q3	ค่าที่ได้
จำนวนคู่ข้อสอบทั้งหมด	780 คู่
ค่าเฉลี่ย (Mean)	-0.023
ค่ามัธยฐาน (Median)	-0.025
ค่าควอไทล์ที่ 3 (Q3 Quartile)	≤ 0.014
ค่าสูงสุด (Maximum)	0.475
จำนวนคู่ที่เกินเกณฑ์ ($ Q3 > 0.20$)	4 คู่ (0.51%)

ตาราง น.6 รายละเอียดคู่ข้อสอบที่มีค่า Q3 เกินเกณฑ์ 0.20

คู่ข้อสอบ	ค่า Q3	สถานการณ์ของ ข้อแรก	สถานการณ์ของ ข้อที่สอง	องค์ประกอบที่วัด
Q21 $\times$ Q13	+0.475	สถานการณ์ที่ 6	สถานการณ์ที่ 4	Decomposition
Q21 $\times$ Q5	+0.454	สถานการณ์ที่ 6	สถานการณ์ที่ 2	Decomposition
Q25 $\times$ Q13	-0.219	สถานการณ์ที่ 7	สถานการณ์ที่ 4	Decomposition
Q37 $\times$ Q5	-0.205	สถานการณ์ที่ 10	สถานการณ์ที่ 2	Decomposition

หมายเหตุ: ค่า Q3 สูงสุดจากการวิเคราะห์จริงเท่ากับ 0.4746625 ซึ่งเมื่อปัดเป็นทศนิยม 3 ตำแหน่ง
จะได้ 0.475 จึงสอดคล้องกับค่าที่รายงานในตาราง

ตาราง น.7 สรุปค่าพารามิเตอร์ IRT 2PL จำแนกตามองค์ประกอบของแบบวัด

องค์ประกอบ	จำนวน ข้อ	ช่วงค่า a	ช่วงค่า b	RMSEA สูงสุด	Item Fit
Decomposition	10	0.762 ถึง 2.738	-1.096 ถึง 0.142	0.0409	ผ่านทุกข้อ
Pattern Recognition	10	0.964 ถึง 2.930	-1.161 ถึง 0.647	0.0477	ผ่านทุกข้อ
Abstraction	10	1.060 ถึง 3.367	-0.929 ถึง 0.284	0.0290	ผ่านทุกข้อ
Algorithm Design	10	0.629 ถึง 1.826	-1.114 ถึง 0.716	0.0398	ผ่านทุกข้อ
ภาพรวมทั้งฉบับ	40	0.629 ถึง 3.367	-1.161 ถึง 0.716	0.0477	40/40 ผ่าน

หมายเหตุ: ค่า a อยู่ในช่วง 0.629–3.367 (Baker, 2001 กำหนดเกณฑ์ที่ยอมรับได้ส่วนใหญ่ 0.50–2.50 และ $a > 2.00$ ถือว่าอยู่ในระดับดีมาก) ค่า b อยู่ในช่วง -1.161 ถึง 0.716 ผ่านเกณฑ์ที่ยอมรับได้ ($-3.00 \leq b \leq 3.00$) ทุกข้อ ค่า b ที่เป็นลบส่วนใหญ่สะท้อนว่าข้อสอบมีระดับความยากต่ำกว่าค่าเฉลี่ยของกลุ่มเล็กน้อย เหมาะสมกับกลุ่มเป้าหมายที่เป็นนักเรียนมัธยมศึกษาตอนต้น

ภาคผนวก ข เกณฑ์ปกติและแนวทางการแปลผลคะแนนของแบบวัด

ตาราง ข.1 เกณฑ์ปกติ (Norms) และระดับทักษะของแบบวัดฉบับสมบูรณ์

คะแนน	ความถี่	ความถี่	Percentile	T-score	ระดับทักษะ
0	-	-	-	-	
1	-	-	-	-	
2	-	-	-	-	
3	-	-	-	-	
4	-	-	-	-	
5	6	6	0.24	21.80	ต่ำกว่าเกณฑ์ปกติ
6	9	15	0.83	26.05	ต่ำกว่าเกณฑ์ปกติ
7	17	32	1.86	29.16	ต่ำกว่าเกณฑ์ปกติ
8	24	56	3.48	31.86	ต่ำกว่าเกณฑ์ปกติ
9	33	89	5.73	34.22	ต่ำกว่าเกณฑ์ปกติ
10	34	123	8.38	36.20	ต่ำกว่าเกณฑ์ปกติ
11	53	176	11.82	38.16	ต่ำกว่าเกณฑ์ปกติ
12	27	203	14.98	39.63	ต่ำกว่าเกณฑ์ปกติ
13	39	242	17.59	40.69	อยู่ในเกณฑ์ปกติ
14	17	259	19.80	41.51	อยู่ในเกณฑ์ปกติ
15	18	277	21.19	42.00	อยู่ในเกณฑ์ปกติ
16	25	302	22.89	42.58	อยู่ในเกณฑ์ปกติ
17	31	333	25.10	43.29	อยู่ในเกณฑ์ปกติ
18	18	351	27.04	43.88	อยู่ในเกณฑ์ปกติ
19	22	373	28.62	44.36	อยู่ในเกณฑ์ปกติ
20	34	407	30.83	44.99	อยู่ในเกณฑ์ปกติ

ตาราง ข.1 (ต่อ)

คะแนน	ความถี่	ความถี่	Percentile	T-score	
21	33	440	33.48	45.73	อยู่ในเกณฑ์ปกติ
22	24	464	35.73	46.34	อยู่ในเกณฑ์ปกติ
23	40	504	38.26	47.01	อยู่ในเกณฑ์ปกติ
24	29	533	40.99	47.72	อยู่ในเกณฑ์ปกติ
25	40	573	43.72	48.42	อยู่ในเกณฑ์ปกติ
26	40	613	46.88	49.22	อยู่ในเกณฑ์ปกติ
27	51	664	50.47	50.12	อยู่ในเกณฑ์ปกติ
28	61	725	54.90	51.23	อยู่ในเกณฑ์ปกติ
29	50	775	59.29	52.35	อยู่ในเกณฑ์ปกติ
30	48	823	63.16	53.36	อยู่ในเกณฑ์ปกติ
31	56	879	67.27	54.47	อยู่ในเกณฑ์ปกติ
32	47	926	71.34	55.63	อยู่ในเกณฑ์ปกติ
33	53	979	75.30	56.84	อยู่ในเกณฑ์ปกติ
34	48	1027	79.29	58.16	อยู่ในเกณฑ์ปกติ
35	46	1073	83.00	59.54	อยู่ในเกณฑ์ปกติ
36	30	1103	86.01	60.81	สูงกว่าเกณฑ์ปกติ
37	55	1158	89.37	62.46	สูงกว่าเกณฑ์ปกติ
38	51	1209	93.56	65.19	สูงกว่าเกณฑ์ปกติ
39	28	1237	96.68	68.36	สูงกว่าเกณฑ์ปกติ
40	28	1265	98.89	72.87	สูงกว่าเกณฑ์ปกติ

หมายเหตุ: สถิติพรรณนา: Mean = 25.02, SD = 9.62, Min = 5, Max = 40

ตาราง ข.2 เกณฑ์การจำแนกระดับทักษะด้วยคะแนนที่ปกติ (Normalized T-score)

ระดับ	ช่วง T-score	Percentile Rank	ความหมาย
สูงกว่าเกณฑ์ปกติ	$T \geq 60$	$PR \geq 84$	สูงกว่าค่าเฉลี่ยตั้งแต่ 1 SD ขึ้นไป
อยู่ในเกณฑ์ปกติ	$T 40-59$	$PR 16-83$	ภายใน ± 1 SD จากค่าเฉลี่ย
ต่ำกว่าเกณฑ์ปกติ	$T < 40$	$PR < 16$	ต่ำกว่าค่าเฉลี่ยมากกว่า 1 SD

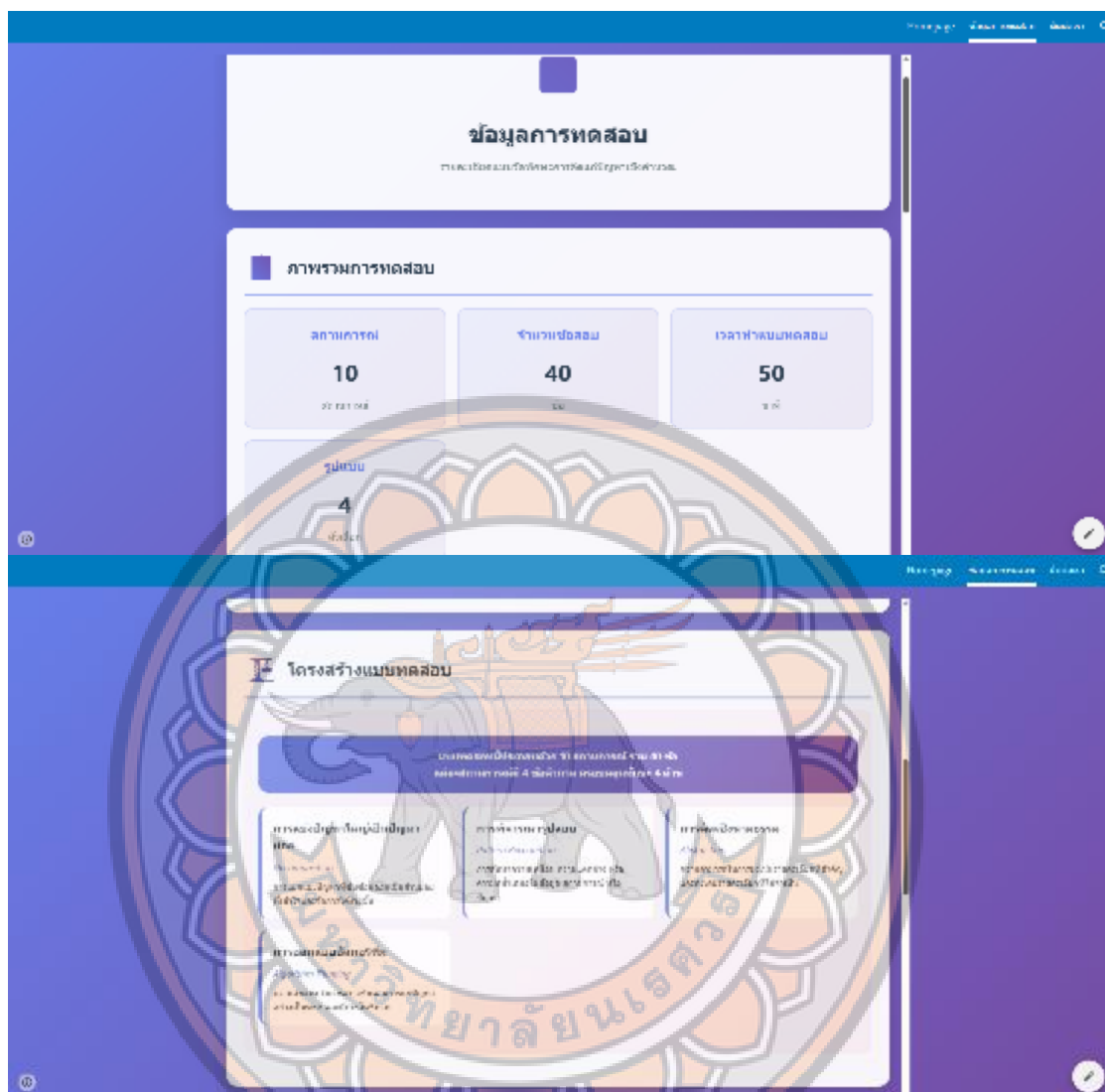
ตาราง ข.3 เกณฑ์การประเมินระดับทักษะจำแนกตามช่วงคะแนนดิบของแบบวัดฉบับสมบูรณ์

ช่วงคะแนนดิบ	ระดับทักษะ	ความหมาย
36 – 40	สูงกว่าเกณฑ์ปกติ	สูงกว่าค่าเฉลี่ยตั้งแต่ 1 SD ขึ้นไป
13 – 35	อยู่ในเกณฑ์ปกติ	ภายใน ± 1 SD จากค่าเฉลี่ย
5 – 12	ต่ำกว่าเกณฑ์ปกติ	ต่ำกว่าค่าเฉลี่ยมากกว่า 1 SD

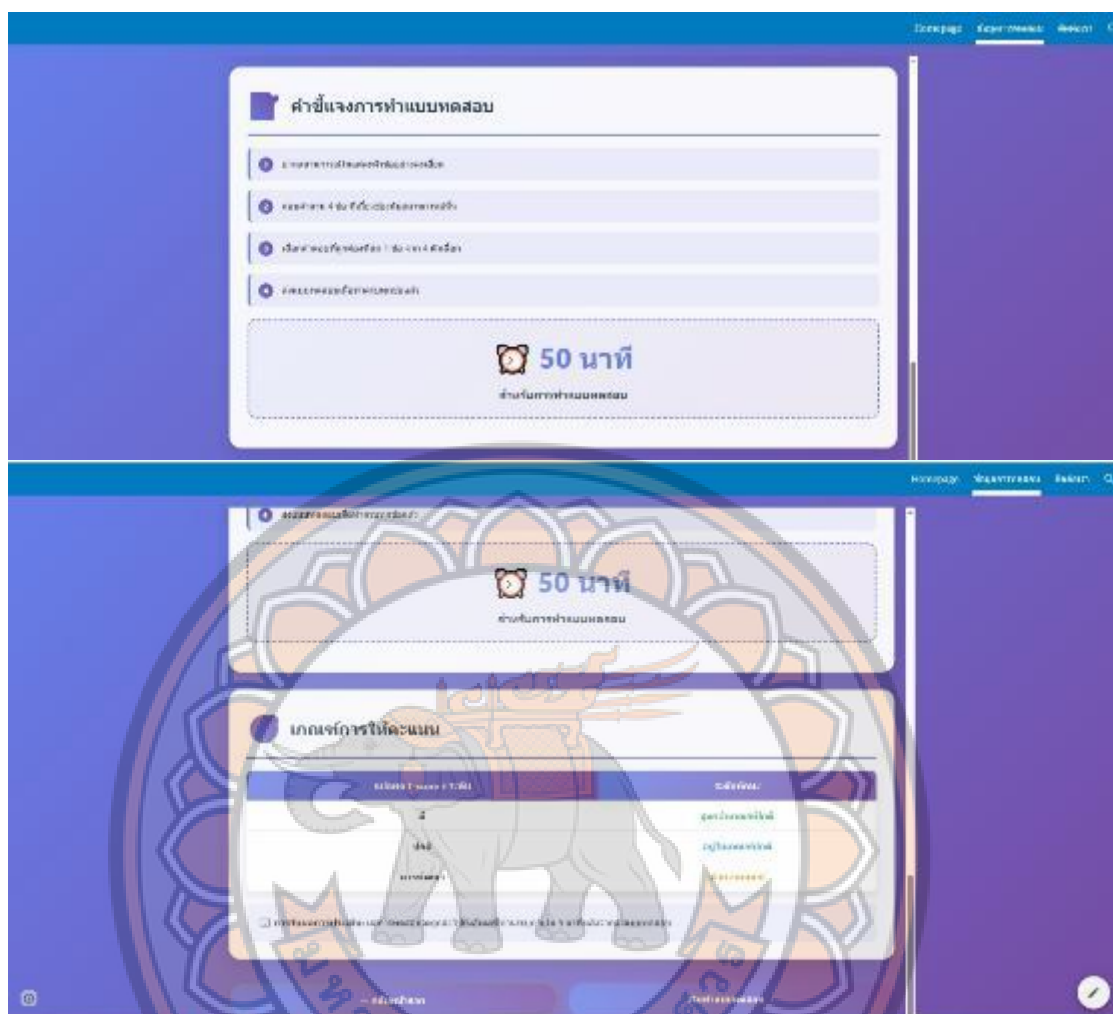
ภาคผนวก ช ระบบการประเมินแบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิง
สถานการณ์แบบออนไลน์



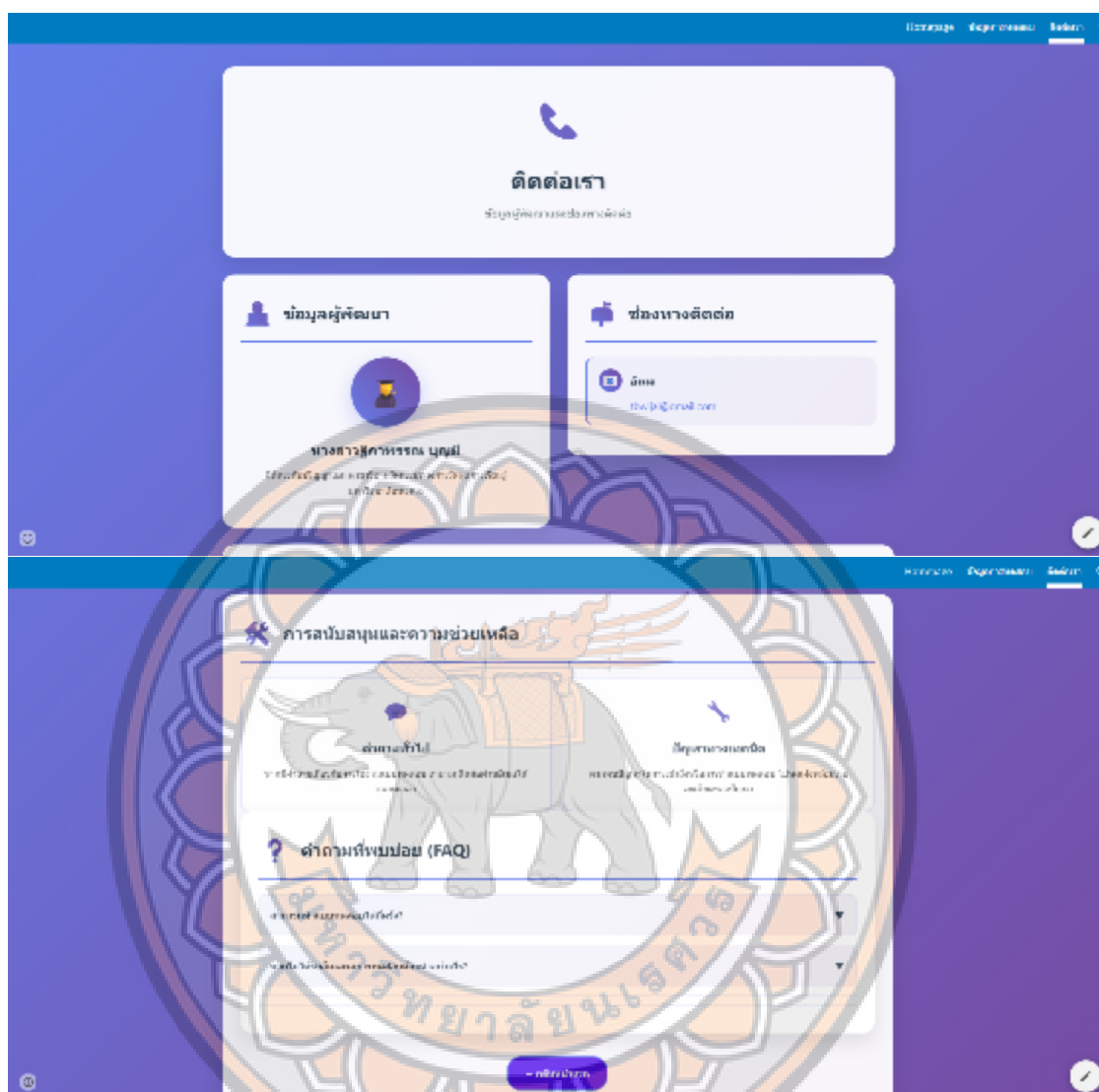
ภาพ ช.1 หน้าเว็บไซต์ Homepage



ภาพ ซ.2 หน้าเว็บไซต์ข้อมูลการทดสอบ



ภาพ ซ.2 (ต่อ)



ภาพ ซ.3 หน้าเว็บไซต์ติดต่อเรา

แบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์

tbwijai@gmail.com [Switch account](#) 

* Indicates required question

Email *

Your email

ชั้น ม.1/ม.2/ม.3 และห้อง (ตัวอย่าง ม.2/3) *

Your answer

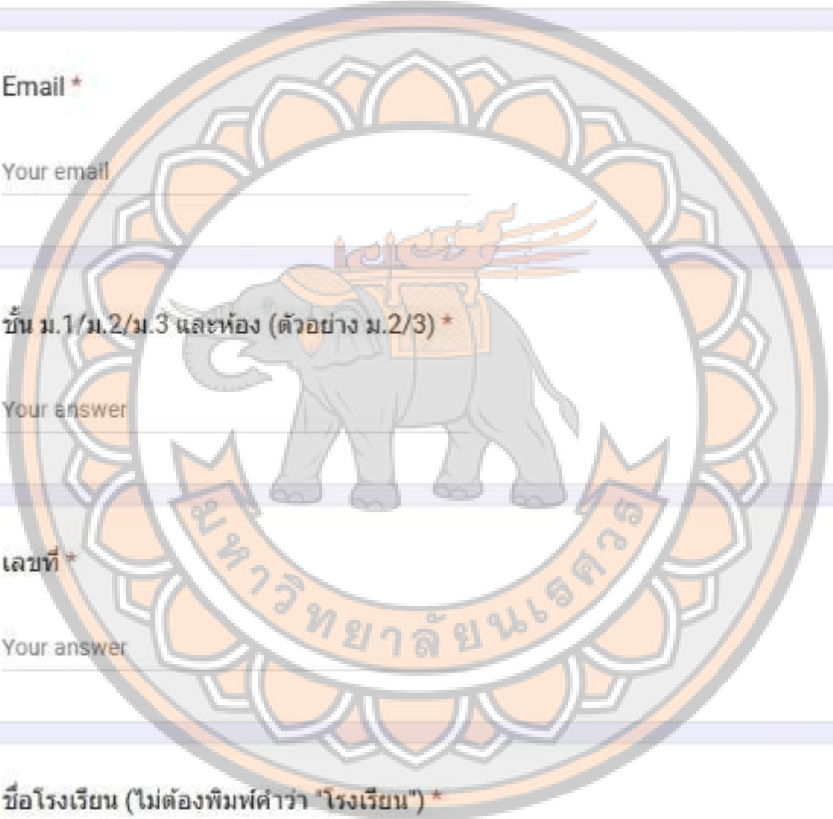
เลขที่ *

Your answer

ชื่อโรงเรียน (ไม่ต้องพิมพ์คำว่า "โรงเรียน") *

Your answer

[Next](#) [Clear form](#)



ภาพ ข.4 หน้ากรอกข้อมูลเพื่อเข้าสู่แบบวัดออนไลน์

2. จากการทำข้อสอบเก่า ประเมินพบว่าข้อสอบวิชาวิทยาศาสตร์มักออกเกี่ยวกับการทดลอง *
ในบท 3, 5, 7 และ 9 ซึ่งเป็นบทที่มีการทดลอง แนวทางใด **เหมาะสมที่สุด** หากประเมินมี
เวลาจำกัดและต้องการเพิ่มโอกาสทำคะแนนในวิชาวิทยาศาสตร์ให้สูงที่สุด ?

- ☐ ก. ทบทวนทุกบทโดยให้เวลาบทที่มีการทดลองมากขึ้นเล็กน้อย โดยไม่เน้นการวิเคราะห์รูปแบบข้อสอบ
- ☐ ข. ทบทวนบทที่มีการทดลองเป็นหลัก โดยพิจารณารูปแบบคำถามจากข้อสอบเก่า
- ☐ ค. ทบทวนบทที่มีการทดลองทั้งหมดก่อน แล้วอ่านบทอื่นแบบย่อโดยไม่เชื่อมโยงรูปแบบคำถาม
- ☐ ง. เลือกทบทวนเฉพาะบทที่ออกสอบบ่อยจากประสบการณ์ โดยไม่วิเคราะห์รูปแบบข้อสอบอย่างเป็นระบบ

3. ในการจัดสรุปเนื้อหาเพื่อทบทวน **ปรึศนาควรเลือกบันทึกข้อมูลแบบใด ? ***

- ☐ ก. จัดเนื้อหาตามลำดับบทเรียน โดยคงรายละเอียดส่วนใหญ่ไว้เพื่อความครบถ้วน
- ☐ ข. จัดแนวคิดหลัก สูตรสำคัญ และข้อมูลที่จำเป็นต่อการทำความเข้าใจเนื้อหา
- ☐ ค. จัดประเด็นสำคัญของแต่ละบท พร้อมตัวอย่างบางส่วนโดยไม่สรุปเป็นภาพรวม
- ☐ ง. จัดข้อมูลจากคำอธิบายของครูเป็นหลัก โดยไม่พิจารณาว่าข้อมูลใดจำเป็น

4. ขั้นตอนที่มีประสิทธิภาพในการท่องจำสูตรคณิตศาสตร์คืออะไร? *

- ☐ ก. อ่านสูตรหลายครั้ง → ทำตัวอย่างทุกวัน → ตู้อะไร → จัดจำขั้นตอนการคำนวณที่ถูกต้อง
- ☐ ข. ท่องจำสูตรทุกวัน → ทำแบบฝึกหัด → ตรวจคำตอบ → ทบทวนเฉพาะข้อที่ผิด
- ☐ ค. ทำความเข้าใจสูตร → ฝึกใช้กับตัวอย่าง → ทำแบบฝึกหัด → หยุดเมื่อทำได้
- ☐ ง. ทำความเข้าใจสูตร → เชื่อมโยงกับตัวอย่าง → ฝึกใช้หลายรูปแบบ → ทบทวนเป็นระยะ

Back

Next

Clear form

40. เพื่อให้การบันทึกข้อมูลขยะรีไซเคิลของคุณและกรรมการธนาคารขยะเป็นไปอย่างถูกต้อง และตรวจสอบได้ ควรจัดลำดับขั้นตอนการทำงานอย่างไรให้มีความเป็นระบบ *

ขั้นตอนที่กำหนด

1. ชั่งน้ำหนักขยะรีไซเคิล
2. บันทึกข้อมูลลงในสมุดหรือโปรแกรม
3. แยกขยะตามประเภท
4. รวมยอดน้ำหนักและคำนวณเป็นเงิน

- ☐ ก. 3 → 1 → 2 → 4
- ☒ ข. 3 → 1 → 4 → 2
- ☐ ค. 1 → 3 → 2 → 4
- ☐ ง. 3 → 2 → 1 → 4



Send me a copy of my responses.

Back

Submit

Clear form

แบบวัดทักษะการคิดเชิงคำนวณในรูปแบบเชิงสถานการณ์

Your response has been recorded.

[View accuracy](#)

[Submit another response](#)

This content is neither created nor endorsed by Google. - [Contact form owner](#) - [Terms of Service](#) - [Privacy Policy](#)

Does this form look suspicious? [Report](#)

Google Forms

ภาคผนวก ณ รายงานผลการประเมินและการให้ข้อมูลป้อนกลับอัตโนมัติ



ภาพ ณ.2 ตัวอย่างการแปลผลคะแนน


Abstraction 5/10

การคิดเชิงนามธรรม

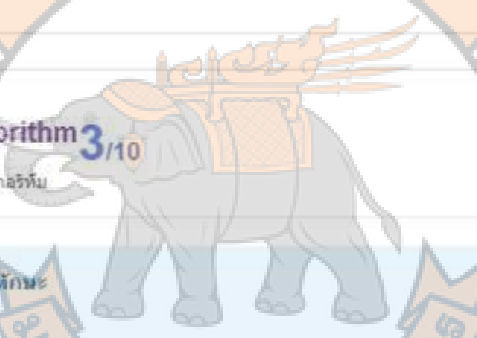
 ระดับทักษะ

ปกติ

คุณสามารถมองภาพใหญ่และเข้าใจหลักการได้พอสมควร และสามารถใช้ในการอธิบายได้

 แนวทางการพัฒนา

- ฝึกการสร้างแผนผังความคิด (Mind Map)
- ลองสรุปบทเรียนเป็นแนวคิดสั้น ๆ ของตัวเอง
- ฝึกแยกแยะข้อมูลสำคัญและไม่สำคัญ




Algorithm 3/10

การออกแบบอัลกอริทึม

 ระดับทักษะ

ปกติ

คุณสามารถสร้างขั้นตอนการทำงานได้ด้วยตนเอง และเข้าใจการใช้เงื่อนไขในการตัดสินใจ

 แนวทางการพัฒนา

- ฝึกเขียนผังงาน (Flowchart) สำหรับกิจกรรมต่างๆ
- ลองเขียนโปรแกรมง่ายๆ ด้วย Scratch หรือ Code.org
- ฝึกออกแบบเกมหรือกิจกรรมที่มีขั้นตอนชัดเจน


ขอบคุณที่ทำตามทดสอบ

รายงานนี้เป็นข้อมูลส่วนบุคคลของคุณ 1

หากมีข้อสงสัยเพิ่มเติม กรุณาติดต่อผู้ดูแลระบบ

 สงวนลิขสิทธิ์โดยมีลิขสิทธิ์จากระบบประเมินทักษะการคิดเชิงคำนวณ

ภาพ ณ.2 (ต่อ)

ภาคผนวก ญ การพัฒนาระบบและการให้ข้อมูลป้อนกลับอัตโนมัติการประเมินออนไลน์ด้วย Google Apps Script

สิ่งที่ได้นี้ทำ

1. เมื่อนักเรียนกดส่ง Google Forms → ระบบทำงานอัตโนมัติ
2. ตรวจให้คะแนนข้อสอบ ทั้ง 40 ข้อ (ตอบถูก = 1, ตอบผิด = 0)
3. คำนวณคะแนนรวม + คะแนนรายองค์ประกอบ 4 ด้าน
4. คำนวณ T-score จากเกณฑ์ปกติ ($M = 25.02$, $SD = 9.62$)
5. แปลผลตามเกณฑ์ 3 ระดับ
6. ส่งอีเมลด้วยภาษาเข้าใจง่าย

ขั้นตอนที่ 1: เข้า Apps Script

1. เปิด Google Forms ของแบบวัด
2. คลิกที่เมนู Extensions
3. คลิกเลือก App Script
4. ระบบจะเปิด Apps Script Editor



ภาพ ญ.1 การเข้า Apps Script

ขั้นตอนที่ 2: paste โค้ด

1. ลบโค้ดเดิมทั้งหมด ในไฟล์ Code.gs
2. paste โค้ดด้านล่างนี้แทน
3. กด Save (Ctrl + S หรือ Cmd + S)

```

1 // =====
2 // ระบบรายงานผลแบบทดสอบทักษะการคิดเชิงคำนวณ
3 // คำนวณ T-score + แปลผล 3 ระดับ
4 //
5 // =====
6 // ค่าคงที่อ้างอิงทางทฤษฎี (ตามแนวทาง สมนึก ภัททิยธนี, 2551)
7 // =====
8
9 // เกณฑ์อ้างอิงทฤษฎีของคะแนนรวม
10 // ใช้ค่ากึ่งกลางของสเกล (คะแนนเต็ม 40)
11 const NORM_TOTAL_MEAN = 20;
12 const NORM_TOTAL_SD = 8;
13
14 // เกณฑ์อ้างอิงทฤษฎีรายองค์ประกอบ
15 // ใช้ค่ากึ่งกลางของสเกล (คะแนนเต็ม 10 ต่อด้าน)
16 const NORM_COMPONENT = {
17   decomposition: { mean: 5, sd: 2 },
18   patternRecognition: { mean: 5, sd: 2 },
19   abstraction: { mean: 5, sd: 2 },
20   algorithm: { mean: 5, sd: 2 }
21 };
22
23 // ค่าคงที่สำหรับการคำนวณ T-score
24 const T_MEAN = 50;
25 const T_SD = 10;
26
27 // เกณฑ์การแปลผล 3 ระดับ (ตามแนวทาง สมนึก ภัททิยธนี, 2551)
28 const T_HIGH = 60; // T ≥ 60 → สูงกว่าเกณฑ์ปกติ (ดี)
29 const T_LOW = 40; // T < 40 → ต่ำกว่าเกณฑ์ปกติ (ควรพัฒนา)
30

```

```

1 // =====
2 // ระบบรายงานผลแบบทดสอบทักษะการคิดเชิงคำนวณ
3 // คำนวณ T-score + แปลผล 3 ระดับ
4 //
5 // =====
6 // ค่าคงที่อ้างอิงทางทฤษฎี (ตามแนวทาง สมนึก ภัททิยธนี, 2551)
7 // =====
8
9 // เกณฑ์อ้างอิงทฤษฎีของคะแนนรวม
10 // ใช้ค่ากึ่งกลางของสเกล (คะแนนเต็ม 40)
11 const NORM_TOTAL_MEAN = 20;
12 const NORM_TOTAL_SD = 8;
13
14 // เกณฑ์อ้างอิงทฤษฎีรายองค์ประกอบ
15 // ใช้ค่ากึ่งกลางของสเกล (คะแนนเต็ม 10 ต่อด้าน)
16 const NORM_COMPONENT = {
17   decomposition: { mean: 5, sd: 2 },
18   patternRecognition: { mean: 5, sd: 2 },
19   abstraction: { mean: 5, sd: 2 },
20   algorithm: { mean: 5, sd: 2 }
21 };
22
23 // ค่าคงที่สำหรับการคำนวณ T-score
24 const T_MEAN = 50;
25 const T_SD = 10;
26
27 // เกณฑ์การแปลผล 3 ระดับ (ตามแนวทาง สมนึก ภัททิยธนี, 2551)
28 const T_HIGH = 60; // T ≥ 60 → สูงกว่าเกณฑ์ปกติ (ดี)
29 const T_LOW = 40; // T < 40 → ต่ำกว่าเกณฑ์ปกติ (ควรพัฒนา)
30

```

ภาพ ญ.2 การใส่โค้ดในไฟล์ Code.gs ของ Apps Script

```

31 // =====
32 // คำตอบที่ถูกต้องของแบบทดสอบ 40 ข้อ
33 // =====
34
35 const CORRECT_ANSWERS = {
36   // สถานการณ์ที่ 1
37   9: "ค. แยกเนื้อหาแต่ละวิชาเป็นหัวข้อย่อย และวางแผนเวลาอ่านแต่ละส่วน",
38   10: "ข. ทบทวนบทที่มีการทดลองเป็นหลัก โดยพิจารณารูปแบบคำถามจากข้อสอบเก่า",
39   11: "ข. จัดแนวคิดหลัก สูตรสำคัญ และข้อมูลที่จำเป็นต่อการทำความเข้าใจเนื้อหา",
40   12: "ง. ทำความเข้าใจสูตร → เชื่อมโยงกับตัวอย่าง → ฝึกใช้หลายรูปแบบ → ทบทวนเป็นระยะ",
41
42   // สถานการณ์ที่ 2
43   13: "ค. แบ่งเวลาออกเป็นช่วงสำหรับการบ้าน การเล่นเกม และการอ่านหนังสือ และปฏิบัติตามแผนที่กำหนด",
44   14: "ง. เล่นกีตาร์ก่อนแล้วค่อยทำการบ้าน เพราะช่วยให้ใช้เวลาว่างน้อยลง",
45   15: "ก. ระบุกิจกรรม เวลาเริ่ม-จบ และระยะเวลาของแต่ละกิจกรรม เพื่อให้เห็นโครงสร้างการใช้เวลา",
46   16: "ข. อ่านโจทย์ทั้งหมด → วางแผนการทำ → ลงมือทำตามลำดับที่กำหนด",
47
48   // สถานการณ์ที่ 3
49   17: "ก. แยกเสื้อผ้าออกเป็นกลุ่มตามประเภทการใช้งาน ก่อนจัดเก็บลงตู้แต่ละช่อง",
50   18: "ข. จัดเสื้อผ้าที่ใช้น้อยไว้ตำแหน่งที่หยิบง่าย และแยกจากเสื้อผ้าที่ใช้น้อย",
51   19: "ง. ประเภทเสื้อผ้า สี ตำแหน่งที่เก็บ และความถี่ในการใช้งาน",
52   20: "ง. แยกเสื้อผ้าตามประเภท → คัดเสื้อผ้าที่ไม่ใช่ → พับ → จัดเก็บตามหมวดและความถี่การใช้",
53
54   // สถานการณ์ที่ 4
55   21: "ง. แบ่งงานบ้านออกเป็นงานย่อย และมอบหมายให้สมาชิกแต่ละคนตามความเหมาะสม",
56   22: "ก. เพิ่มผู้รับผิดชอบล้างจานหลังมื้อเย็น และเพิ่มรอบการซักผ้าในวันหยุด",
57   23: "ข. รายการงานบ้าน ผู้รับผิดชอบ วันเวลาที่ทำ และสถานะความสำเร็จของงาน",
58   24: "ก. เก็บของให้เป็นระเบียบ → บัดฝุ่นจากที่สูงลงต่ำ → ดูดฝุ่นหรือกวาดพื้น → ถูพื้น",
59
60   // สถานการณ์ที่ 5
61   25: "ข. แยกอาหารตามประเภททั้ง 3 ประเภท ก่อนนำไปจัดเก็บและใช้งาน",
62   26: "ค. เตรียมพื้นที่จัดเก็บและวางแผนซื้ออาหารเพิ่มในวันหยุดหรือวันที่มีการทำอาหารมากกว่าปกติ",
63   27: "ก. ประเภทอาหาร ปริมาณ วันที่จัดเก็บ วันหมดอายุ และตำแหน่งที่เก็บอาหารแต่ละประเภท",
64   28: "ค. แยกประเภทอาหารสด → ทำความสะอาด → จัดเก็บให้เหมาะสม → ปั่นปิ้งอาหาร",
65
66   // สถานการณ์ที่ 6
67   29: "ค. แยกหนังสือออกเป็นหมวดหมู่ตามประเภท เพื่อจัดการหนังสือแต่ละกลุ่ม",
68   30: "ค. หนังสือที่อ่านง่ายและให้ความบันเทิงถูกยืมมากกว่าหนังสือที่มีเนื้อหาเชิงวิชาการ",
69   31: "ก. ชื่อหนังสือ ชื่อผู้ยืม วันที่ยืม และวันที่คืน",
70   32: "ค. ตรวจสอบสถานะ → ลงชื่อยืม → บันทึกข้อมูล → ส่งมอบหนังสือ",
71
72   // สถานการณ์ที่ 7
73   33: "ข. แบ่งงานออกเป็นงานการแข่งขัน และวางแผนฝึกซ้อมให้เพียงพอ",
74   34: "ค. เตรียมน้ำดื่ม และเจ้าหน้าที่เพิ่มเป็นพิเศษในวันแรกของการแข่งขัน",
75   35: "ง. เวลาแข่งขัน สนามแข่งขัน ผู้เข้าแข่งขัน และลำดับการแข่งขัน",
76   36: "ข. จัดการแข่งขันเป็นรอบคัดเลือก รอบรองชนะเลิศ และรอบชนะเลิศตามลำดับ",
77
78   // สถานการณ์ที่ 8
79   37: "ก. แยกพิจารณาประเภทของการบุลลี ผู้ที่เกี่ยวข้อง สถานที่และช่วงเวลาที่เกิดเหตุ รวมถึงผลกระทบที่เกิดขึ้นกับผู้ถูกรุณภัย",
80   38: "ง. การบูลลี่มีรูปแบบการเกิดขึ้นที่สัมพันธ์กับช่วงเวลา สถานที่ และกลุ่มนักเรียนบางกลุ่มอย่างชัดเจน",
81   39: "ข. หลักการสำคัญในการป้องกัน เช่น การสร้างพื้นที่ปลอดภัย ขอบทางการแจ้งเหตุ และระบบช่วยเหลือเพื่อนอย่างคอบบ์",
82   40: "ค. รับแจ้งเหตุจากนักเรียน ตรวจสอบและประเมินข้อมูล ให้การช่วยเหลือผู้ถูกรุณภัย และติดตามผลเพื่อป้องกันการเกิดซ้ำ",
83
84   // สถานการณ์ที่ 9
85   41: "ก. เลื่อนระดับน้ำ แรงดันประสาชน เครื่องมือแพทย์ และประเมินผลหลังใช้",
86   42: "ข. กำหนดเกณฑ์ปริมาณแผ่นสไลด์เพื่อให้เป็นสัญญาณเตือนล่วงหน้าแก่แพทย์",
87   43: "ค. ระดับน้ำ ปริมาณแผ่น ตำแหน่งพื้นที่เสี่ยง และเวลา",
88   44: "ก. 1 → 2 → 3 → 4",
89

```

```

90 // สถานการณ์ที่ 18
91 45: "ก. รวบรวมขยะ คัดแยก จำหน่าย บันทึกข้อมูล และประเมินผลการดำเนินงานของโครงการ",
92 46: "ข. เพิ่มจุดรับขยะ เคียวพื้นที่จัดเก็บให้เพียงพอ ประสานผู้รับซื้อให้สอดคล้องกับปริมาณที่เพิ่มขึ้น",
93 47: "ง. รับสมัครเรียน ประเภทขยะ น้ำหนัก วันที่นำมา และคะแนนสะสม",
94 48: "ก.  $3 + 1 + 2 + 4$ ",
95 };
96
97 // =====
98 // ฟังก์ชันหลัก: ทำงานเมื่อ submit form
99 // =====
100
101 function onFormSubmit(e) {
102   const sheet = SpreadsheetApp.getActiveSpreadsheet().getActiveSheet();
103   const lastRow = sheet.getLastRow();
104   const lastCol = sheet.getLastColumn();
105
106   try {
107     const row = sheet.getRange(lastRow, 1, 1, lastCol).getValues()[0];
108
109     const timestamp = row[0];
110     const studentEmail = row[1];
111     const formScore = row[2];
112     const studentPre = row[3];
113     const studentName = row[4];
114     const studentM3 = row[5];
115     const studentthong = row[6];
116
117     const scores = calculateScores(row);
118     sendResultEmail(studentthong, studentEmail, scores);
119     sheet.getRange(lastRow, lastCol).setValue("SENT");
120     Logger.log("✅ ส่งอีเมลสำเร็จ: " + studentEmail);
121   } catch (error) {
122     sheet.getRange(lastRow, lastCol).setValue("ERROR: " + error.message);
123     Logger.log("❌ ส่งอีเมลไม่สำเร็จ: " + error.toString());
124   }
125 }
126
127 // =====
128 // คำนวณคะแนน
129 // =====
130
131
132 function calculateScores(row) {
133   let decomposition = 0;
134   let patternRecognition = 0;
135   let abstraction = 0;
136   let algorithm = 0;
137
138   for (let col = 9; col <= 48; col++) {
139     const userAnswer = row[col - 1];
140     const correctAnswer = CORRECT_ANSWERS[col];
141
142     if (userAnswer === correctAnswer) {
143       const questionIndex = (col - 9) % 4;
144

```

```

145     switch(questionIndex) {
146         case 0: decomposition++; break;
147         case 1: patternRecognition++; break;
148         case 2: abstraction++; break;
149         case 3: algorithm++; break;
150     }
151 }
152 }
153
154 const totalScore = decomposition + patternRecognition + abstraction + algorithm;
155
156 return {
157     decomposition: decomposition,
158     patternRecognition: patternRecognition,
159     abstraction: abstraction,
160     algorithm: algorithm,
161     total: totalScore
162 };
163 }
164
165 // =====
166 // ฟังก์ชันคำนวณ T-score
167 // =====
168
169 function calculateTScore(rawScore, mean, sd) {
170     if (sd === 0) return T_MEAN;
171     const zScore = (rawScore - mean) / sd;
172     const tScore = T_MEAN + (T_SD * zScore);
173     return Math.round(tScore * 100) / 100;
174 }
175
176 // =====
177 // แปลผล T-score เป็น 3 ระดับ
178 // =====
179
180 function interpretLevelByTScore(tScore) {
181     if (tScore >= T_HIGH) {
182         return {
183             friendly: "ดี",
184             formal: "สูงกว่าเกณฑ์ปกติ",
185             color: "#4caf50",
186             bgColor: "#e8f5e9",
187             borderColor: "#4caf50",
188             titleColor: "#2e7d32",
189             textColor: "#1b5e20"
190         };
191     } else if (tScore >= T_LOW) {
192         return {
193             friendly: "ปกติ",
194             formal: "อยู่ในเกณฑ์ปกติ",
195             color: "#2196f3",
196             bgColor: "#e3f2fd",
197             borderColor: "#2196f3",
198             titleColor: "#1565c0",
199             textColor: "#0d47a1"
200         };
201     };

```

ภาพ ญ.2 (ต่อ)

```

201 } else {
202     return {
203         friendly: "ควรพัฒนา",
204         formal: "ต่ำกว่าเกณฑ์ปกติ",
205         color: "#ff9800",
206         bgColor: "#fff3cd",
207         borderColor: "#ff9800",
208         titleColor: "#c55108",
209         textColor: "#bf366c"
210     };
211 }
212 }
213
214 // -----
215 // ส่วนรวมความสามารถ + ข้อเสนอแนะ
216 // -----
217
218 function getComponentDescription(component, level) {
219     const descriptions = {
220         decomposition: {
221             "ดี": {
222                 ability: "คุณเก่งในการแบ่งปัญหาใหญ่ให้เป็นปัญหย่อย ๆ และใช้ความสามารถเชื่อมโยงระหว่างส่วนต่าง ๆ ได้ดี",
223                 suggestion: "• ลองนำงานชิ้นเดิมมาศึกษาให้ละเอียดขึ้น เช่น ไปเจอจุดที่ตัวเองยังสงสัยแล้วหาคำตอบ<br>• คิดเป็นผู้นำในการวางแผนงานกลุ่ม<br>• ลองนำความสามารถไปใช้กับงานอื่น ๆ เช่น ช่วยทำงานของเพื่อนคนอื่น"
224             },
225             "ปกติ": {
226                 ability: "คุณสามารถแบ่งปัญหาย่อยเป็นส่วนย่อยได้พอสมควร และเริ่มเข้าใจความเชื่อมโยงระหว่างส่วนต่าง ๆ",
227                 suggestion: "• คิดเป็นผู้นำได้ 5-8 ชิ้นต่อเป็นประจำวัน<br>• ลองนำความสามารถไปใช้ในการเรียนหรือทำงาน<br>• ลองนำความสามารถไปใช้กับงานอื่น ๆ เช่น ช่วยงานของเพื่อนคนอื่น"
228             },
229             "ควรพัฒนา": {
230                 ability: "คุณกำลังเริ่มต้นในการเรียนรู้การแบ่งปัญหา ลองฝึกเขียนโครงร่างเพื่อพัฒนาทักษะ",
231                 suggestion: "• เริ่มจากกิจกรรมง่าย ๆ ในชีวิตประจำวัน เช่น การจัดการเป็นกิจวัตร<br>• ลองถามตัวเองว่า 'ทำไมเราถึงไม่คิด'<br>• ลองวาดแผนที่ความคิด (Mind Map) เพื่อช่วยการคิด"
232             }
233         },
234         patternRecognition: {
235             "ดี": {
236                 ability: "คุณสามารถเชื่อมโยงองค์ความรู้ที่เรียนมาในวิชาต่าง ๆ ได้ดี และใช้ความรู้ที่เรียนมาเพื่อแก้ปัญหาได้",
237                 suggestion: "• ศึกษาปัญหาที่เจอแล้วลองวิเคราะห์ แล้วลองใช้ความรู้ที่เรียนมาเพื่อแก้ปัญหา<br>• ลองนำความรู้ไปใช้กับงานอื่น ๆ ในการวิจัยหรือเขียนบทความ<br>• ลองนำความรู้ไปใช้กับงานอื่น ๆ เช่น ช่วยงานของเพื่อนคนอื่น"
238             },
239             "ปกติ": {
240                 ability: "คุณสามารถใช้ความรู้ที่เรียนมาเพื่อแก้ปัญหาได้บ้าง แต่ยังขาดความเชื่อมโยงระหว่างวิชา",
241                 suggestion: "• ลองสังเกตปัญหาในชีวิตประจำวัน เช่น เวลาซื้อข้าว เวลาตัก<br>• ลองนำความรู้ไปใช้กับงานอื่น ๆ เช่น ช่วยงานของเพื่อนคนอื่น"
242             },
243             "ควรพัฒนา": {
244                 ability: "คุณยังขาดการเชื่อมโยงความรู้ที่เรียนมา ลองเริ่มจากสิ่งง่าย ๆ รอบตัวเพื่อพัฒนาทักษะ",
245                 suggestion: "• ลองสังเกตปัญหาในชีวิตประจำวัน ๆ ในชีวิตประจำวัน<br>• ลองนำความรู้ไปใช้กับงานอื่น ๆ เช่น ช่วยงานของเพื่อนคนอื่น"
246             }
247         },
248         abstraction: {
249             "ดี": {
250                 ability: "คุณสามารถมองภาพรวมและเข้าใจความสำคัญได้ดี รวมทั้งสามารถนำความรู้ไปใช้กับงานอื่นได้",
251                 suggestion: "• ศึกษาจากปัญหาที่เจอแล้วลองวิเคราะห์ แล้วลองใช้ความรู้ที่เรียนมาเพื่อแก้ปัญหา<br>• ลองนำความรู้ไปใช้กับงานอื่น ๆ เช่น ช่วยงานของเพื่อนคนอื่น"
252             },
253             "ปกติ": {
254                 ability: "คุณสามารถมองภาพรวมและเข้าใจความสำคัญได้บ้าง แต่ยังขาดความเชื่อมโยงระหว่างวิชา",
255                 suggestion: "• ศึกษาจากปัญหาที่เจอแล้วลองวิเคราะห์ แล้วลองใช้ความรู้ที่เรียนมาเพื่อแก้ปัญหา<br>• ลองนำความรู้ไปใช้กับงานอื่น ๆ เช่น ช่วยงานของเพื่อนคนอื่น"
256             },
257             "ควรพัฒนา": {
258                 ability: "คุณยังขาดการเชื่อมโยงความรู้ที่เรียนมา ลองเริ่มจากสิ่งง่าย ๆ รอบตัวเพื่อพัฒนาทักษะ",
259                 suggestion: "• ลองสังเกตปัญหาในชีวิตประจำวัน ๆ ในชีวิตประจำวัน<br>• ลองนำความรู้ไปใช้กับงานอื่น ๆ เช่น ช่วยงานของเพื่อนคนอื่น"
260             }
261         }
262     };

```

ภาพ ญ.2 (ต่อ)

```

261   },
262   algorithm: {
263     "ดี": {
264       ability: "คุณสมารถออกแบบขั้นตอนการทำงานที่ชัดเจนและมีการสังเกตภาพได้ดี",
265       suggestion: "• ลองศึกษาอัลกอริทึมพื้นฐานในวิชาวิทยาการคำนวณ<br>• ศึกษาเว็บไซต์ของเมืองด้วย Scratch หรือ Python<br>• ลองเขียนหาพื้นที่ในข้อที่ 4 กันเถอะ ท้าทายใจหน่อย"
266     },
267     "ปกติ": {
268       ability: "คุณสามารถอธิบายขั้นตอนการทำงานได้ด้วยตนเอง และเข้าใจภาพในงานที่คิดค้นได้",
269       suggestion: "• ศึกษาขั้นตอนการ (Flowchart) สำหรับกิจกรรมต่าง ๆ <br>• ลองเขียนโปรแกรมด้วย ๆ ด้วย Scratch หรือ Code.org<br>• ศึกษาออกแบบแผนหรือกิจกรรมที่ขึ้นตอนชัดเจน"
270     },
271     "ควรพัฒนา": {
272       ability: "คุณยังต้องฝึกการคิดเป็นขั้นตอน ลองเริ่มจากกิจกรรมที่มีขั้นตอนชัดเจนเพื่อพัฒนาทักษะ",
273       suggestion: "• ศึกษากิจกรรมที่มีขั้นตอนชัดเจน เช่น ทำอาหารตามสูตร<br>• ศึกษาการปฏิบัติคำสั่งและลำดับที่ระบุ<br>• ศึกษาเรื่องลำดับเหตุการณ์หรือกิจกรรม"
274     }
275   }
276 }
277
278 return descriptions[component][level];
279 }
280
281 function getOverallDescription(level) {
282   const overall = {
283     "ดี": "คุณมีทักษะการคิดเชิงคำนวณที่ต่ำกว่าส่วนอื่น ๆ ของคุณ คุณสมารถวางโครงข่ายไปใช้แก้ปัญหาในโปรแกรมง่ายและการทำงานได้อย่างมีประสิทธิภาพ",
284     "ปกติ": "คุณมีทักษะการคิดเชิงคำนวณในระดับปกติ โดยได้ฝึกฝนเพิ่มเติมเกี่ยวกับ ทักษะการคิดเป็นขั้นตอน และทักษะการสังเกตภาพ",
285     "ควรพัฒนา": "ทักษะการคิดเชิงคำนวณของคุณยังอยู่ในระดับเริ่มต้น ในเบื้องต้นได้ ศึกษาทำความเข้าใจเกี่ยวกับ ทักษะการสังเกตภาพและลำดับที่ระบุ และทำอาหารหรือเล่นเกมตามขั้นตอนที่กำหนด"
286   };
287   return overall[level];
288 }
289
290 // =====
291 // ส่วนอื่น
292 // =====
293
294 function sendResultEmail(name, email, scores) {
295   const subject = "ผลการประเมินผลสัมฤทธิ์ทางการเรียน";
296
297   const tTotal = calculateTScore(scores.total, NORM_TOTAL_MEAN, NORM_TOTAL_STD);
298   const tDecomp = calculateTScore(scores.decomposition, NORM_COMPONENT.decomposition.mean, NORM_COMPONENT.decomposition.sd);
299   const tPattern = calculateTScore(scores.patternRecognition, NORM_COMPONENT.patternRecognition.mean, NORM_COMPONENT.patternRecognition.sd);
300   const tAbstract = calculateTScore(scores.abstraction, NORM_COMPONENT.abstraction.mean, NORM_COMPONENT.abstraction.sd);
301   const tAlgo = calculateTScore(scores.algorithm, NORM_COMPONENT.algorithm.mean, NORM_COMPONENT.algorithm.sd);
302
303   const overallLevel = interpretLevelByScore(tTotal);
304   const decompLevel = interpretLevelByScore(tDecomp);
305   const patternLevel = interpretLevelByScore(tPattern);
306   const abstractLevel = interpretLevelByScore(tAbstract);
307   const algoLevel = interpretLevelByScore(tAlgo);
308
309   const decompDesc = getComponentDescription("decomposition", decompLevel, friendly);
310   const patternDesc = getComponentDescription("patternRecognition", patternLevel, friendly);
311   const abstractDesc = getComponentDescription("abstraction", abstractLevel, friendly);
312   const algoDesc = getComponentDescription("algorithm", algoLevel, friendly);
313   const overallDesc = getOverallDescription(overallLevel, friendly);
314
315   const htmlBody = `
316   <html>
317   <head>
318     <style>
319       body { font-family: 'Sarabun', 'Tahoma', Arial, sans-serif; line-height: 1.8; color: #333; background-color: #f5f5f5; margin: 0; padding: 0; }
320       .container { max-width: 700px; margin: 20px auto; background: white; border-radius: 10px; box-shadow: 0 4px 8px rgba(0,0,0,0.1); overflow: hidden; }
321       .header { background: linear-gradient(135deg, #6677aa 0%, #778899 100%); color: white; padding: 40px 30px; text-align: center; }
322       .header h1 { margin: 0 0 10px 0; font-size: 28px; font-weight: bold; }
323       .header h2 { margin: 0 0 10px 0; font-size: 18px; font-weight: bold; }
324       .content { padding: 20px; }
325
326       .overall-box { background: linear-gradient(135deg, #f5f5f5 0%, #f5f5f5 100%); color: white; padding: 30px; text-align: center; border-radius: 10px; margin-bottom: 10px; box-shadow: 0 4px 8px rgba(0,0,0,0.1); }
327       .overall-box .score { font-size: 24px; font-weight: bold; margin-bottom: 10px; }
328       .overall-box .percentage { font-size: 16px; font-weight: bold; margin-bottom: 10px; }
329       .overall-box .level-name { font-size: 18px; font-weight: bold; margin-bottom: 10px; }
330       .overall-box .level-desc { font-size: 14px; line-height: 1.7; margin-top: 10px; opacity: 0.9; padding: 0 10px; }
331
332       .section-title { font-size: 24px; font-weight: bold; color: #6677aa; border-bottom: 3px solid #6677aa; padding-bottom: 10px; margin: 20px 0 0 0; }
333
334       .score-card { background: #ffffff; border: 2px solid #ccc; border-radius: 10px; padding: 20px; margin-bottom: 20px; box-shadow: 0 4px 8px rgba(0,0,0,0.1); }
335

```

ภาพ ญ.2 (ต่อ)

```

386 .card-header { display: flex; justify-content: space-between; align-items: center; margin-bottom: 20px; padding-bottom: 15px;
387 border-bottom: 2px solid #f0f0f0; }
388 .card-title { font-size: 28px; font-weight: bold; color: #764ba2; }
389 .card-title .icon { font-size: 28px; margin-right: 10px; vertical-align: middle; }
390 .card-title .small { font-size: 18px; font-weight: normal; color: #666; display: block; margin-top: 8px; }
391 .card-score { font-size: 36px; font-weight: bold; color: #6677aa; }
392 .card-score .small { font-size: 18px; }
393
394 .level-badge { padding: 15px 20px; border-radius: 8px; margin: 15px 0; border-left: 4px solid; }
395 .level-badge .badge-title { font-size: 16px; font-weight: bold; margin-bottom: 8px; }
396 .level-badge .badge-level { font-size: 22px; font-weight: bold; margin-bottom: 8px; }
397 .level-badge .badge-desc { font-size: 14px; line-height: 1.7; }
398
399 .suggestion-box { background: linear-gradient(135deg, #fff8e1 0%, #ffe0b2 100%); border-left: 4px solid #ffa726; padding: 15px 20px;
400 border-radius: 8px; margin: 15px 0; }
401 .suggestion-box .sug-title { font-size: 16px; font-weight: bold; color: #e57373; margin-bottom: 10px; }
402 .suggestion-box .sug-content { font-size: 14px; color: #424242; line-height: 1.8; }
403
404 .legend { background: #f5f5f5; padding: 20px; border-radius: 18px; margin: 20px 0; font-size: 14px; }
405 .legend h3 { margin: 0 0 12px 0; color: #555; font-size: 16px; }
406 .legend-item { padding: 5px 0; }
407 .legend-tag { display: inline-block; padding: 3px 12px; border-radius: 12px; color: white; font-weight: bold; margin-right: 8px;
408 font-size: 12px; }
409
410 .footer { text-align: center; padding: 30px; background: #f5f5f5; border-top: 1px solid #e0e0e0; }
411 .footer p { margin: 8px 0; color: #555; font-size: 13px; }
412 .footer .highlight { color: #6677aa; font-weight: bold; }
413
414 @media only screen and (max-width: 600px) {
415   .container { margin: 10px; border-radius: 10px; }
416   .header { padding: 30px 20px; }
417   .header h1 { font-size: 22px; }
418   .content { padding: 20px; }
419   .overall-box .score { font-size: 44px; }
420   .overall-box .level-name { font-size: 24px; padding: 8px 20px; }
421   .card-header { flex-direction: column; align-items: flex-start; }
422   .card-score { margin-top: 10px; }
423 }
424
425 </style>
426 </head>
427 <body>
428   <div class="container">
429     <div class="header">
430       <h1>🐘 ผลการทดสอบทักษะการคิดเชิงคำนวณ</h1>
431       <p>สวัสดีคุณ <strong>${name}</strong></p>
432       <p style="font-size: 14px; opacity: 0.9; margin-top: 10px;">รายงานผลคะแนนรายบุคคล</p>
433     </div>
434     <div class="content">
435       <div class="overall-box">
436         <div style="font-size: 17px; opacity: 0.9;">คะแนนรวมทั้งหมด</div>
437         <div class="score">${scores.total}<span style="font-size: 32px;">/40</span></div>
438
439         <div class="level-name">${overallLevel.friendly}</div>
440         <div class="level-desc">${overallDesc}</div>
441       </div>
442
443       <div class="legend">
444         <h3>🧠 ความหมายของระดับทักษะ</h3>
445         <div class="legend-item">
446           <span class="legend-tag" style="background: #4caf50;">ดี</span>
447           ทักษะของคุณดีกว่าเพื่อนวัยเดียวกันส่วนใหญ่
448         </div>
449         <div class="legend-item">
450           <span class="legend-tag" style="background: #2196f3;">ปกติ</span>
451           ทักษะของคุณอยู่ในระดับปกติ คล้ายเพื่อนวัยเดียวกัน
452         </div>
453         <div class="legend-item">
454           <span class="legend-tag" style="background: #ff9800;">ควรพัฒนา</span>
455           ทักษะของคุณยังพัฒนาเพิ่มเติมได้ ลองทำตามข้อเสนอนะ
456         </div>
457       </div>
458     </div>
459   </body>
460

```

ภาพ ญ.2 (ต่อ)


```

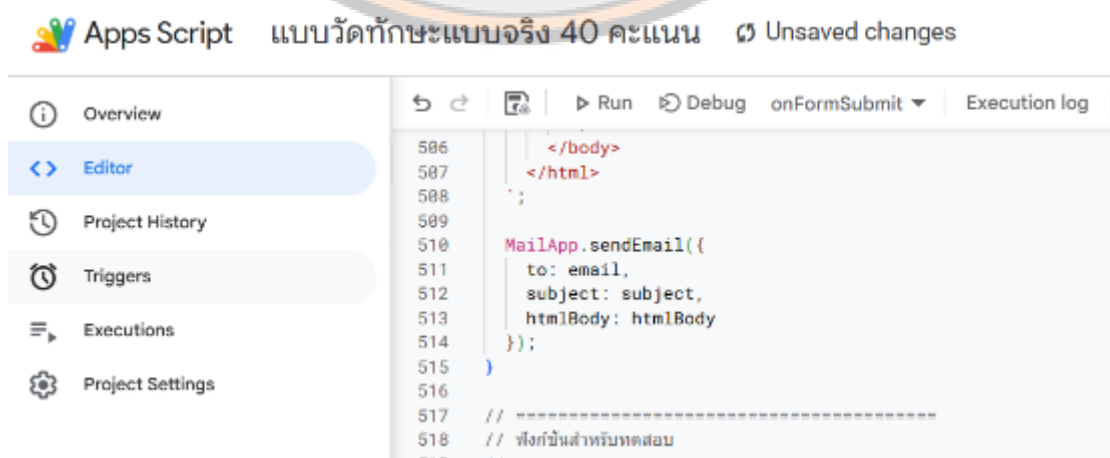
499     <div class="footer">
500       <p> <span class="highlight">ผลงานที่นำมาทดสอบ</span></p>
501       <p>รายงานนี้เป็นโดยส่วนกลางของครู <{name}</p>
502       <p>หากมีข้อสงสัยเพิ่มเติม กรุณาติดต่อผู้ดูแลระบบ</p>
503       <p style="margin-top: 15px; font-size: 12px; color: #999;"> <img alt="Logo of the Ministry of Education, Culture and Sport of Thailand" data-bbox="280 160 300 180"/> เว็บไซต์นี้จัดทำขึ้นโดยระบบประมวลผลข้อมูลการคัดเลือก</p>
504     </div>
505   </div>
506 </body>
507 </html>
508
509
510 MailApp.sendEmail({
511   to: email,
512   subject: subject,
513   htmlBody: htmlBody
514 });
515
516
517 // =====
518 // ฟังก์ชันสำหรับทดสอบ
519 // =====
520
521 function testEmail() {
522   const testScores = {
523     decomposition: 8,
524     patternRecognition: 6,
525     abstraction: 7,
526     algorithms: 7,
527     total: 28
528   };
529
530   sendResultEmail("ทดสอบ ระบบ", "tbwijai@gmail.com", testScores);
531
532   const tTotal = calculateTScore(testScores.total, NORM_TOTAL_MEAN, NORM_TOTAL_SD);
533   const overallLevel = interpretLevelByTScore(tTotal);
534
535   Logger.log("==== ผลการทดสอบ =====");
536   Logger.log("คะแนนรวม: " + testScores.total + "/40");
537   Logger.log("T-score รวม: " + tTotal);
538 }

```

ภาพ ญ.2 (ต่อ)

ขั้นตอนที่ 3: ตั้ง Trigger

1. ในเมนูซ้าย คลิก "Triggers" (รูปนาฬิกา)



ภาพ ญ.3 การเลือกคำสั่ง Triggers

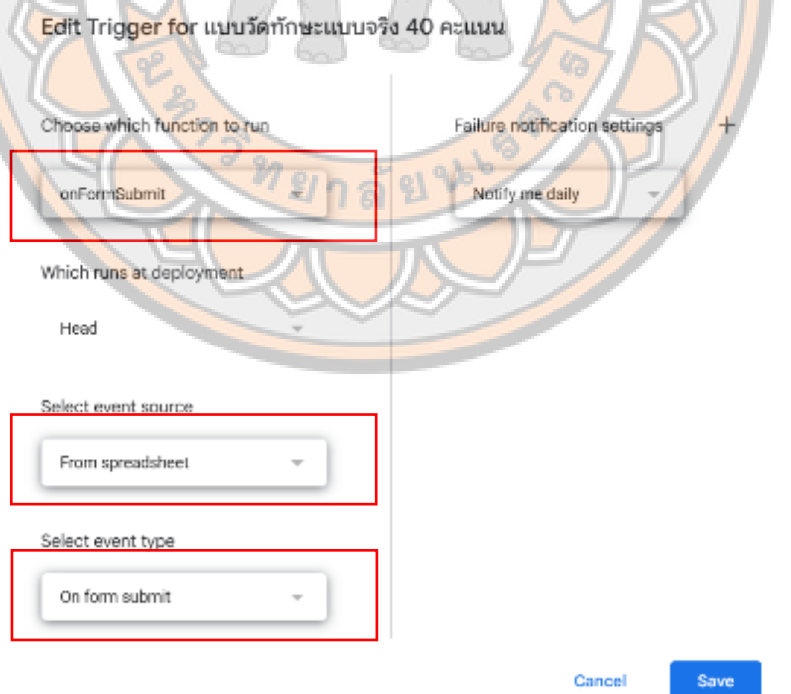
2. คลิก "+ Add Trigger"



ภาพ ญ.4 การคลิกเลือก "+ Add Trigger"

3. ตั้งค่าดังนี้:

- Choose which function to run: **onFormSubmit**
- Select event source: **Form spreadsheet**
- Select event type: **On form submit**



ภาพ ญ.5 การตั้งค่า Triggers

4. กด Save → อนุญาตสิทธิ์เมื่อ Google ถาม

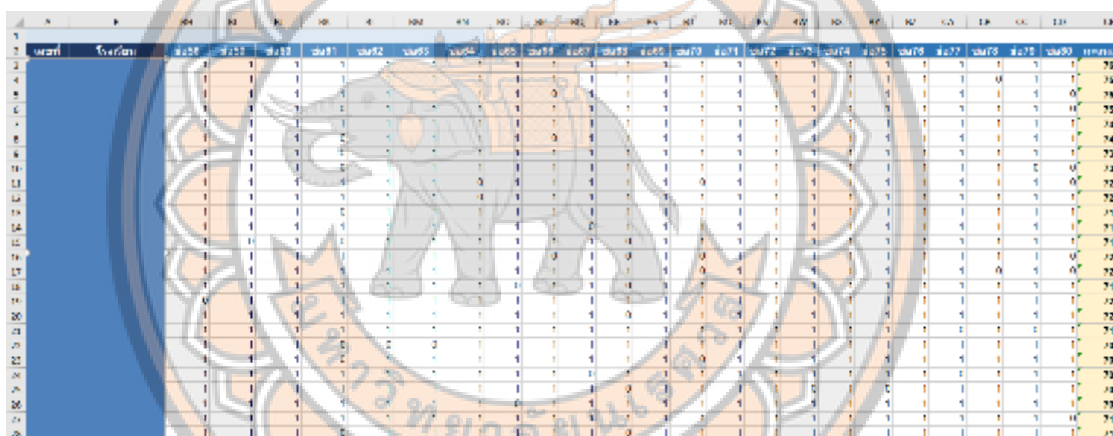
ภาคผนวก ก ตัวอย่างคำสั่งและผลลัพธ์การวิเคราะห์ข้อมูลค่าความเชื่อมั่น KR-20 ของกลุ่มทดลองใช้เครื่องมือ 150 คนด้วยโปรแกรมภาษา R เวอร์ชัน 4.5.3 ผ่านเว็บแอปพลิเคชัน Posit Cloud

1. การเตรียมข้อมูล

ผู้วิจัยจัดเก็บข้อมูลคำตอบของผู้ตอบแบบวัดในรูปแบบไฟล์ Microsoft Excel (.xlsx) โดยกำหนดให้

- ตอบถูก = 1 คะแนน
- ตอบผิด = 0 คะแนน

และบันทึกข้อมูลไว้ในไฟล์ CTT_Real.xlsx



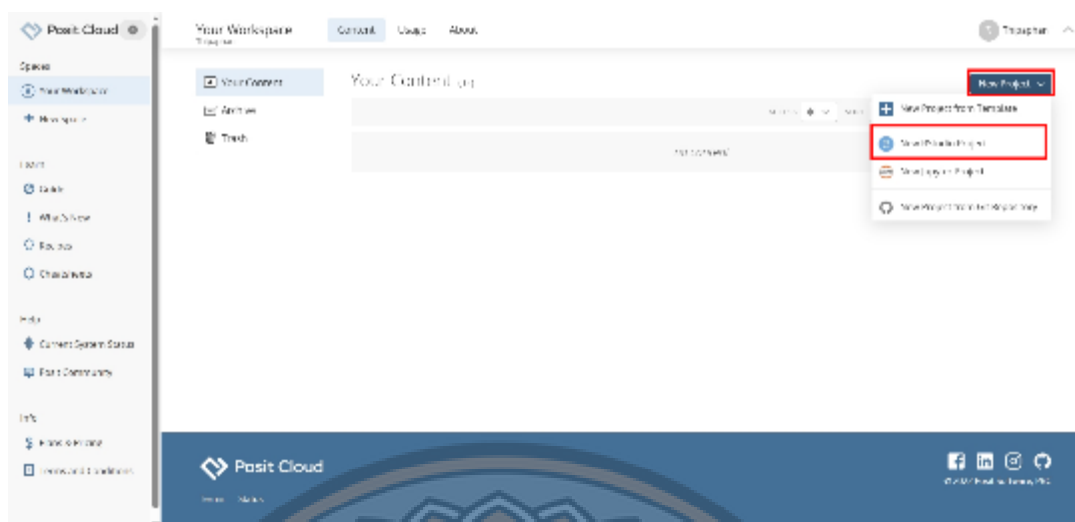
ภาพ ก.1 แสดงข้อมูลคำตอบของผู้ตอบแบบวัดจำนวน 150 คน โดยข้อสอบแต่ละข้อถูกบันทึกในรูปแบบคะแนน 0 และ 1 เพื่อใช้ในการวิเคราะห์ค่าความเชื่อมั่นด้วยวิธี KR-20

2. การเข้าสู่ระบบ Posit Cloud

- 2.1 เปิดเว็บเบราว์เซอร์เข้าเว็บไซต์ <https://posit.cloud>
- 2.2 เข้าสู่ระบบ > เลือก Sign In > กรอกอีเมลและรหัสผ่าน

3. การสร้าง Project ใหม่

- 3.1 เลือก New Project
- 3.2 เลือก New RStudio Project



ภาพ ฎ.2 การสร้าง Project ใหม่

4. การนำเข้าข้อมูล

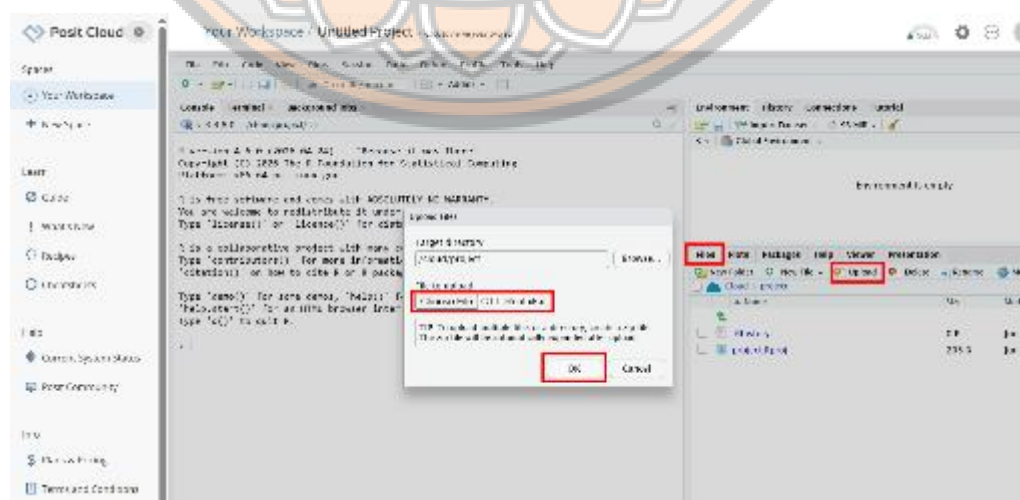
4.1 เลือกแท็บ Files

4.2 กด Upload

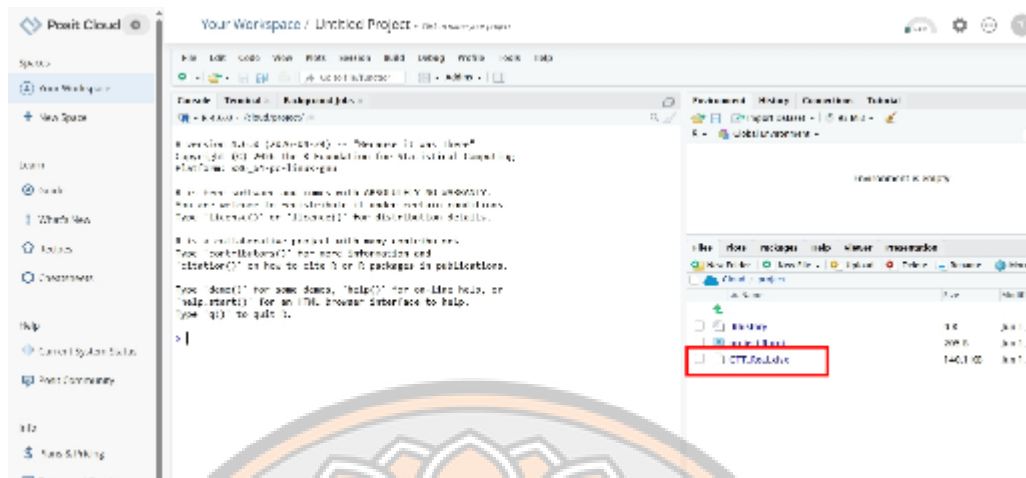
4.3 จะปรากฏหน้าต่าง Target directory ให้เลือก File to upload และคลิก Choose File เลือกไฟล์ที่จัดเตรียมไว้ (ไฟล์ที่ใช้ในตัวอย่างคือ CTT_Real.xlsx)

4.4 จะปรากฏชื่อไฟล์ที่เลือก

4.5 กด OK เพื่ออัปโหลดไฟล์ข้อมูล



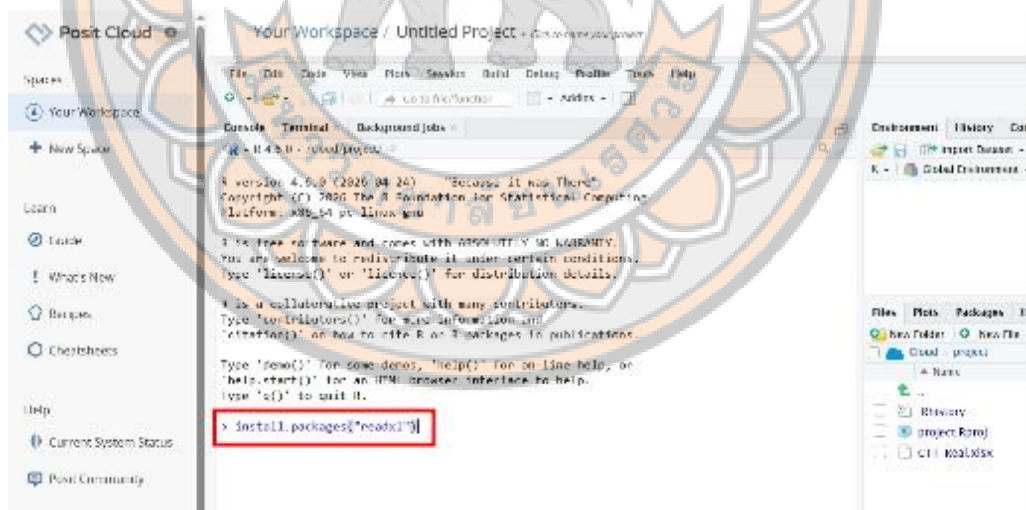
ภาพ ฎ.3 การอัปโหลดไฟล์



ภาพ ฎ.4 เมื่ออัปโหลดสำเร็จ ไฟล์ CTT_Real.xlsx จะปรากฏในหน้าต่าง Files

5. การติดตั้ง Package

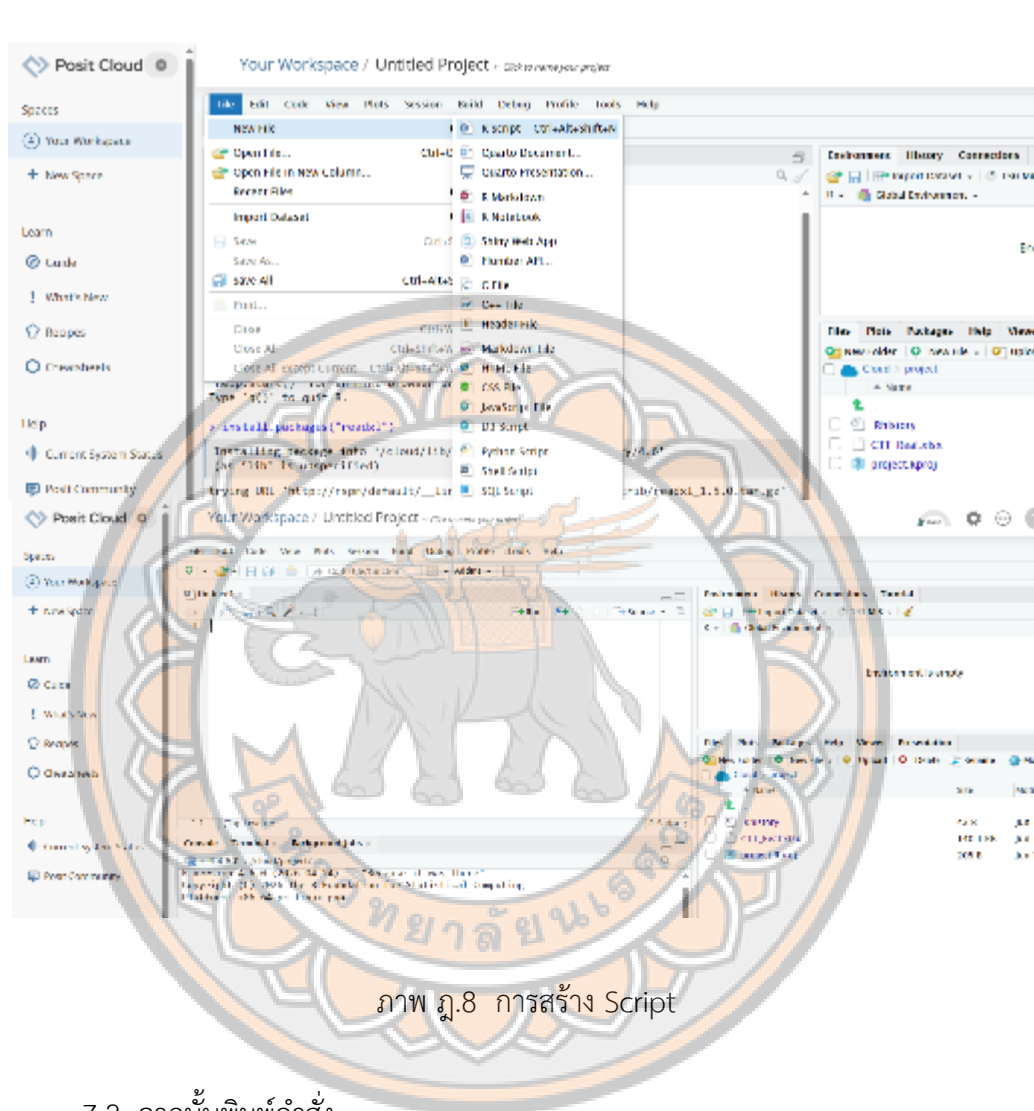
ที่ Console พิมพ์คำสั่ง `install.packages("readxl")` จากนั้นกด Enter



ภาพ ฎ.5 การติดตั้ง Package readxl

7. การอ่านข้อมูลจาก Excel

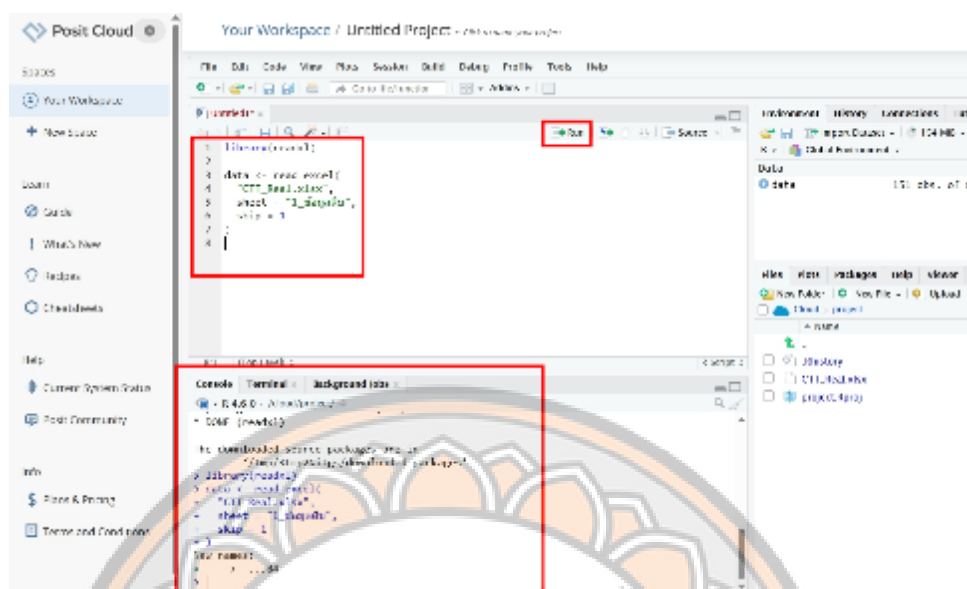
7.1 สร้าง Script ใหม่ > เลือก File → New File → R Script



ภาพ ๗.8 การสร้าง Script

7.2 จากนั้นพิมพ์คำสั่ง

```
library(readxl)
data <- read_excel(
  "CTT_Real.xlsx",
  sheet = "1_ข้อมูลดิบ",
  skip = 1
)
```



ภาพ ฏ.9 การอ่านข้อมูลจาก Excel

8. การตัดแถวสรุปผลรวมออกจากข้อมูล

เนื่องจากแถวสุดท้ายเป็นข้อมูลสรุปจำนวนผู้ตอบ จึงต้องตัดออกก่อนวิเคราะห์พิมพ์คำสั่ง

```
data <- data[-nrow(data), ]
```



ภาพ ฏ.10 การเตรียมข้อมูลก่อนวิเคราะห์

9. การเลือกเฉพาะข้อสอบ 40 ข้อ

พิมพ์คำสั่ง

```
items <- data[, paste0("ข้อ", 1:40)]
```



ภาพ ฏ.11 การเลือกตัวแปรข้อสอบ

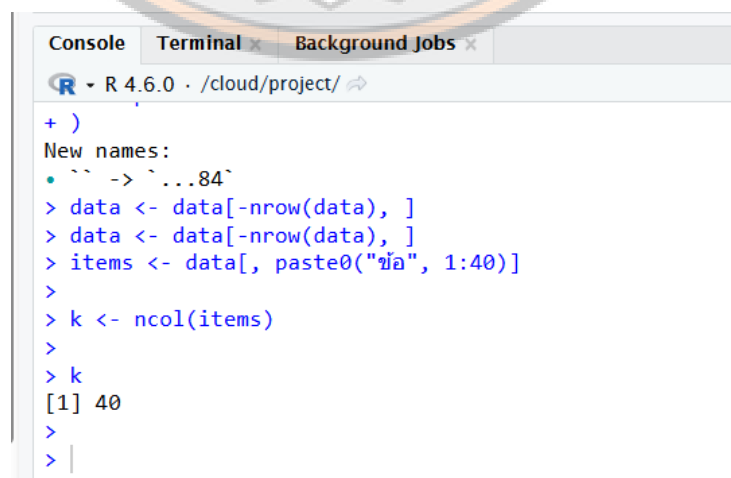
10. การตรวจสอบจำนวนข้อสอบ

ที่ Console พิมพ์คำสั่ง

```
k <- ncol(items)
```

```
k
```

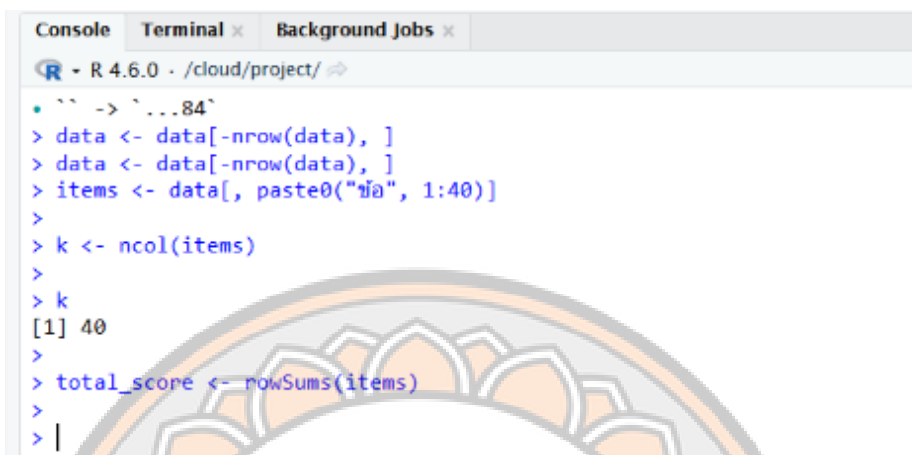
ผลลัพธ์ที่ได้ คือ 40



ภาพ ฏ.12 การตรวจสอบจำนวนข้อสอบ

11. การคำนวณคะแนนรวม

ที่ Console พิมพ์คำสั่ง `total_score <- rowSums(items)`



```

R - R 4.6.0 - /cloud/project/
> `` -> `...84`
> data <- data[-nrow(data), ]
> data <- data[-nrow(data), ]
> items <- data[, paste0("ข้อ", 1:40)]
>
> k <- ncol(items)
>
> k
[1] 40
>
> total_score <- rowSums(items)
>
> |

```

ภาพ ฎ.13 การคำนวณคะแนนรวม

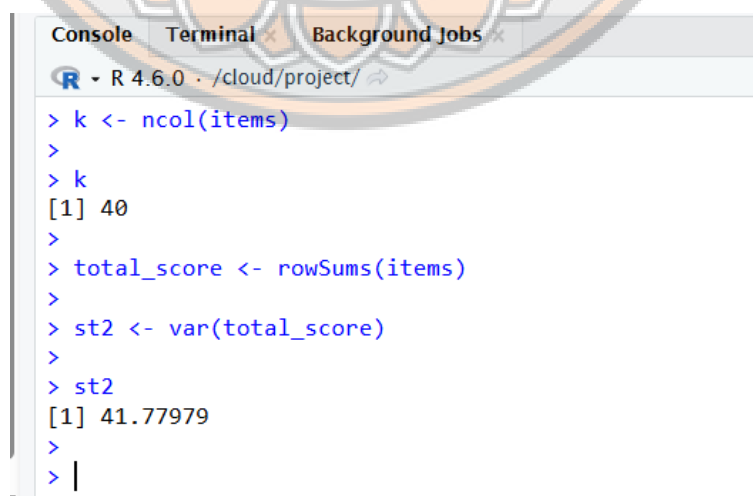
12. การคำนวณความแปรปรวนของคะแนนรวม

ที่ Console พิมพ์คำสั่ง

`st2 <- var(total_score)`

`st2`

ผลลัพธ์ที่ได้ 41.77979



```

R - R 4.6.0 - /cloud/project/
> k <- ncol(items)
>
> k
[1] 40
>
> total_score <- rowSums(items)
>
> st2 <- var(total_score)
>
> st2
[1] 41.77979
>
> |

```

ภาพ ฎ.14 การคำนวณความแปรปรวนคะแนนรวม

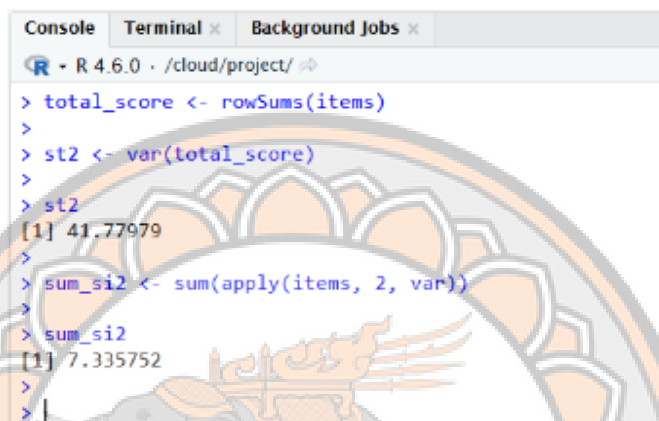
13. การคำนวณผลรวมความแปรปรวนรายข้อ

ที่ Console พิมพ์คำสั่ง

```
sum_si2 <- sum(apply(items, 2, var))
```

```
sum_si2
```

ผลลัพธ์ที่ได้ 7.335752



```

R - R 4.6.0 - /cloud/project/
> total_score <- rowSums(items)
>
> st2 <- var(total_score)
>
> st2
[1] 41.77979
>
> sum_si2 <- sum(apply(items, 2, var))
>
> sum_si2
[1] 7.335752
>
> |

```

ภาพ ฏ.15 การคำนวณผลรวมความแปรปรวนรายข้อ

14. การคำนวณค่าความเชื่อมั่น KR-20

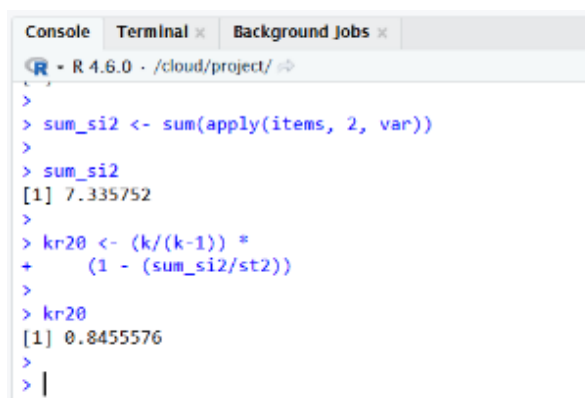
ที่ Console พิมพ์คำสั่ง

```
kr20 <- (k/(k-1)) *
```

```
(1 - (sum_si2/st2))
```

```
kr20
```

ผลลัพธ์ที่ได้ 0.8455576



```

R - R 4.6.0 - /cloud/project/
>
> sum_si2 <- sum(apply(items, 2, var))
>
> sum_si2
[1] 7.335752
>
> kr20 <- (k/(k-1)) *
+   (1 - (sum_si2/st2))
>
> kr20
[1] 0.8455576
>
> |

```

ภาพ ฏ.16 การคำนวณค่า KR-20

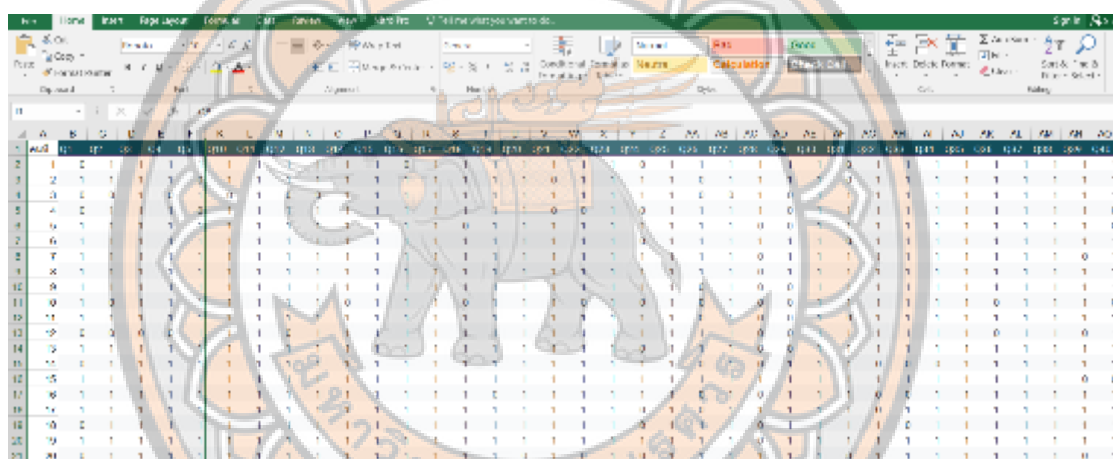
ภาคผนวก ก ตัวอย่างคำสั่งและผลลัพธ์การวิเคราะห์ข้อมูลคุณภาพแบบวัดตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ด้วยโปรแกรมภาษา R เวอร์ชัน 4.5.3 ผ่านเว็บแอปพลิเคชัน Posit Cloud

1. เตรียมไฟล์ข้อมูล

จัดเตรียมข้อมูลในรูปแบบ Excel (.xlsx)

กำหนดให้

- ตอบถูก = 1
- ตอบผิด = 0



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

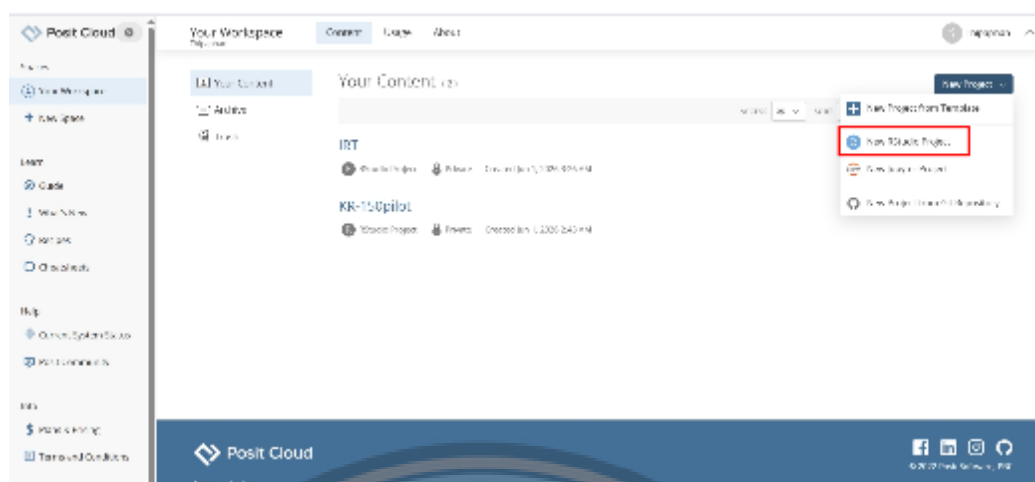
ภาพ ก.1 ไฟล์ข้อมูลดิบใน Excel

2. การเข้าสู่ระบบ Posit Cloud

- 2.1 เปิดเว็บเบราว์เซอร์เข้าเว็บไซต์ <https://posit.cloud>
- 2.2 เข้าสู่ระบบ > เลือก Sign In > กรอกอีเมลและรหัสผ่าน

3. การสร้าง Project ใหม่

- 3.1 เลือก New Project
- 3.2 เลือก New RStudio Project



ภาพ ฎ.2 การสร้าง Project ใหม่

4. การนำเข้าข้อมูล

4.1 เลือกเก็บ Files

4.2 กด Upload

4.3 จะปรากฏหน้าต่าง Target directory ให้เลือก File to upload และคลิก Choose File เลือกไฟล์ที่จัดเตรียมไว้ (ไฟล์ที่ใช้ในตัวอย่างคือ analyz_IRT.xlsx)

4.4 จะปรากฏชื่อไฟล์ที่เลือก

4.5 กด OK เพื่ออัปโหลดไฟล์ข้อมูล



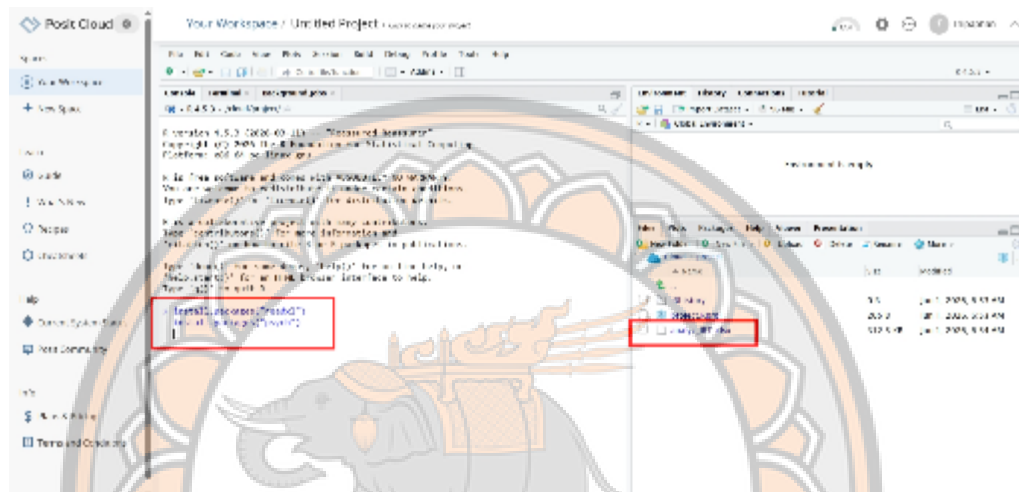
ภาพ ฎ.3 การอัปโหลดไฟล์

4. การติดตั้ง Package ที่จำเป็น ที่ Console พิมพ์คำสั่ง

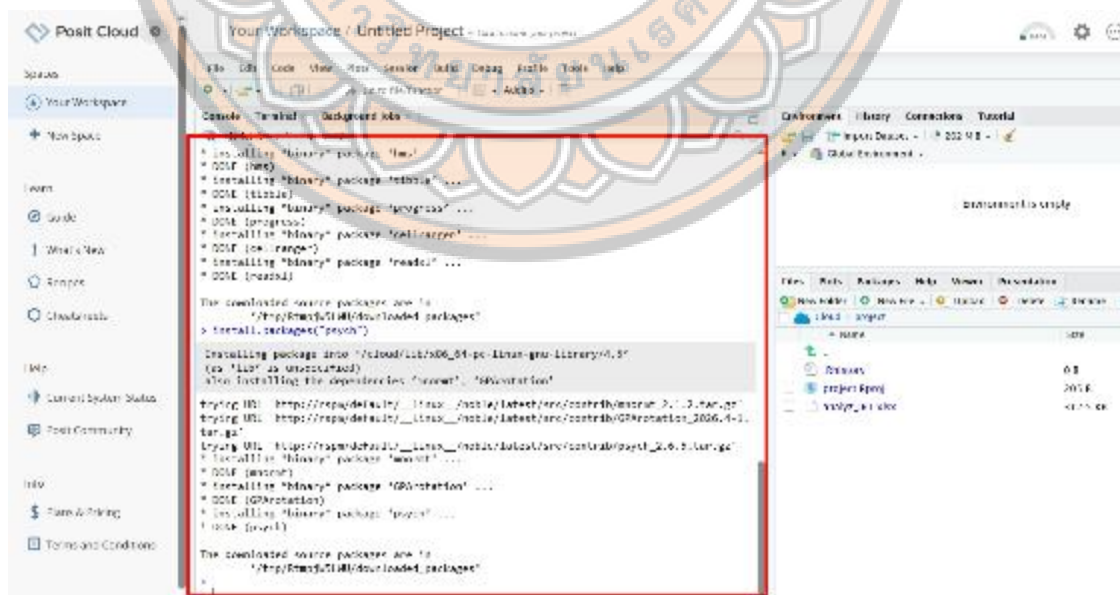
`install.packages("readxl")`

`install.packages("psych")`

จากนั้นกด Enter



ภาพ ฏ.4 การติดตั้ง Package



ภาพ ฏ.5 Package ที่จำเป็นเตรียมพร้อมสำหรับใช้งาน

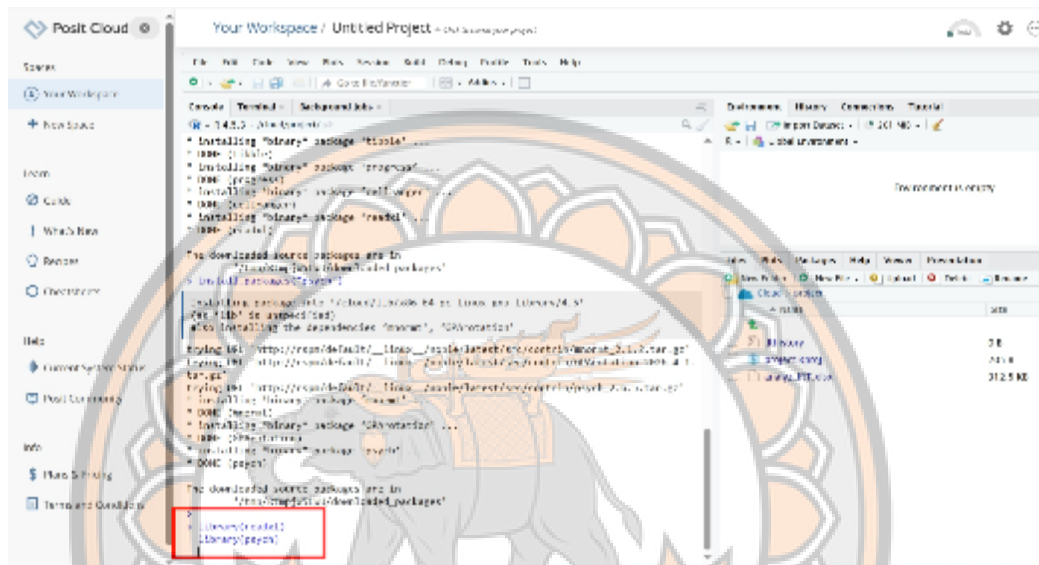
5. การเรียกใช้ Package

ที่ Console พิมพ์คำสั่ง

```
library(readxl)
```

```
library(psych)
```

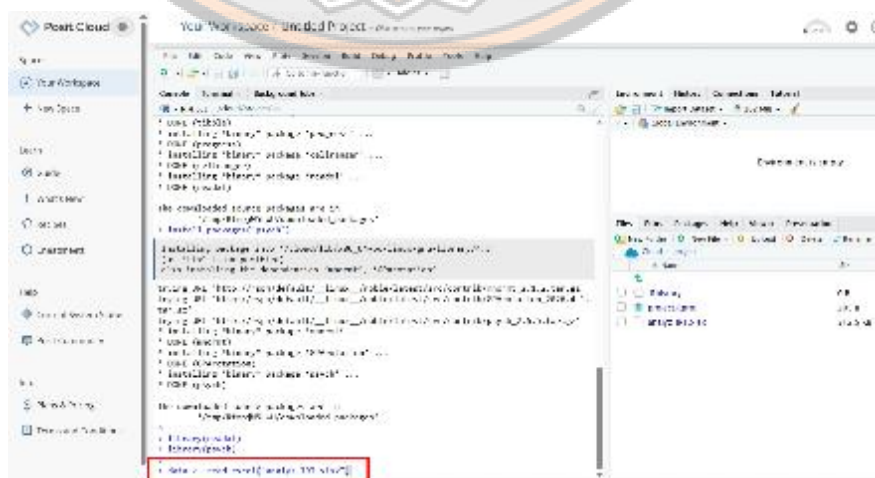
จากนั้นกด Enter



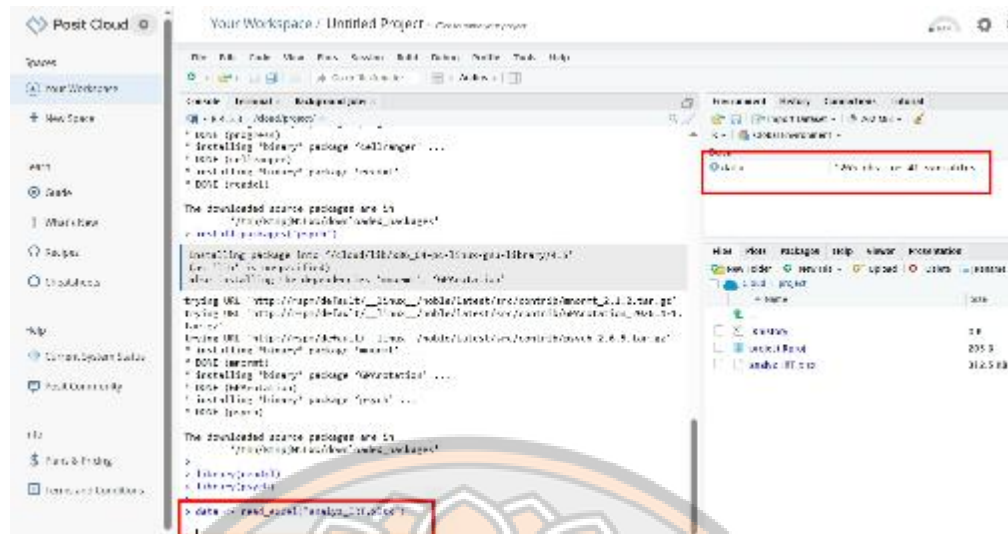
ภาพ ฎ.6 การเรียกใช้ Package

6. ให้โปรแกรมอ่านไฟล์ Excel

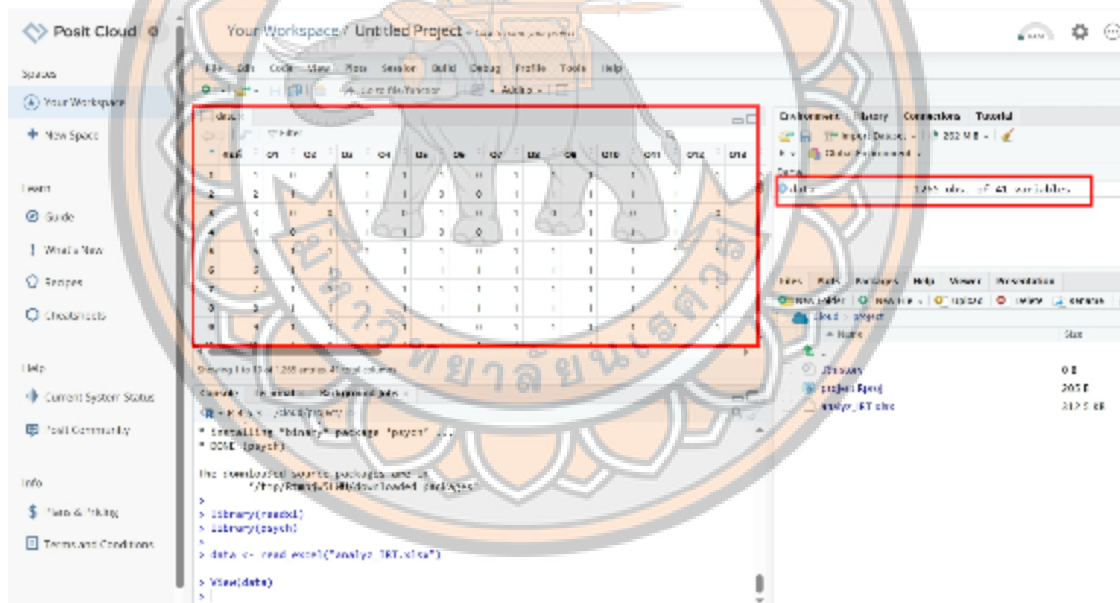
ที่ Console พิมพ์คำสั่ง `data <- read_excel("analyz_IRT.xlsx")` จากนั้นกด Enter



ภาพ ฎ.8 การอ่านข้อมูลจาก Excel



ภาพ ฏ.9 โปรแกรมอ่านข้อมูลจาก Excel และแสดงใน data

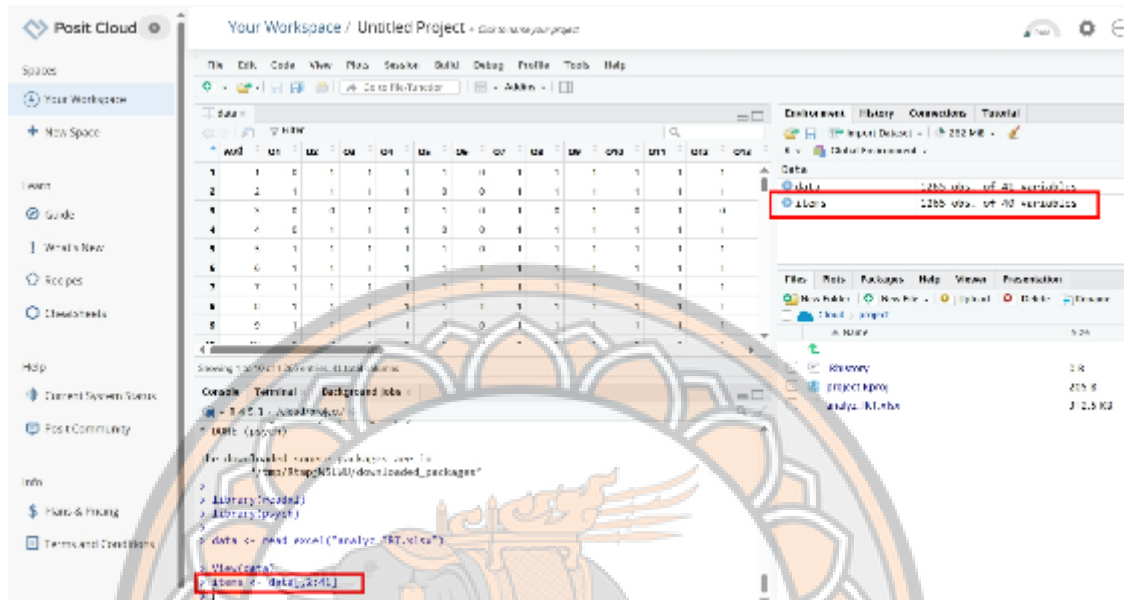


ภาพ ฏ.10 เมื่อคลิกที่ data จะปรากฏข้อมูลใน data เพื่อแสดงข้อมูลค่าใน Excel

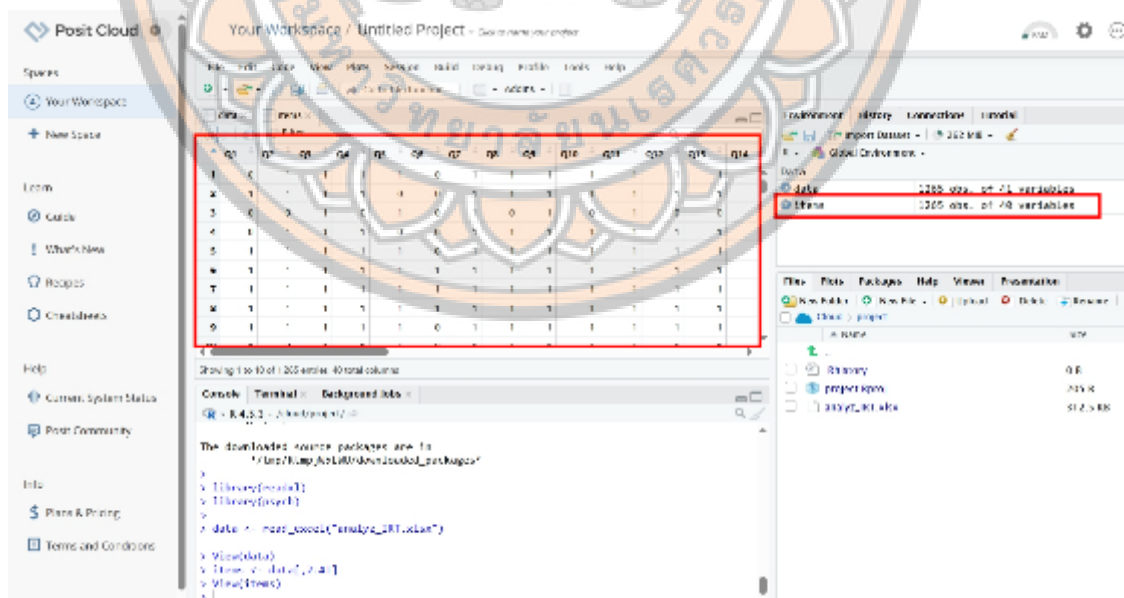
7. คำสั่งให้โปรแกรมเลือกเฉพาะข้อสอบ

หากคอลัมน์แรกเป็นรหัสผู้ตอบ ที่ Console พิมพ์คำสั่ง

```
items <- data[,2:41]
```



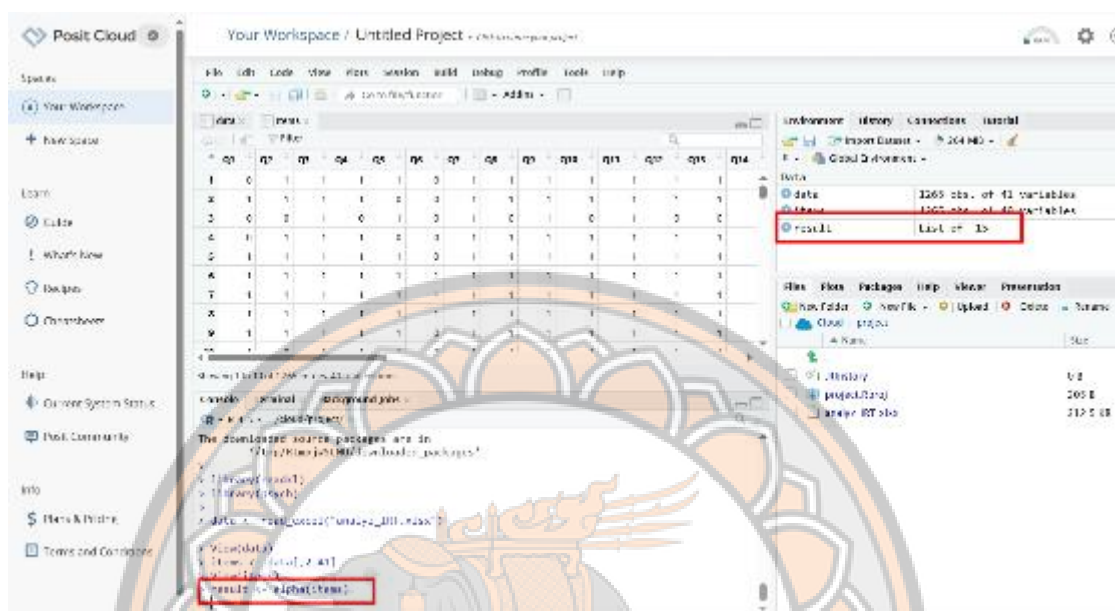
ภาพ ฏ.11 โปรแกรมจะสร้างข้อมูล Item และแสดงใน data



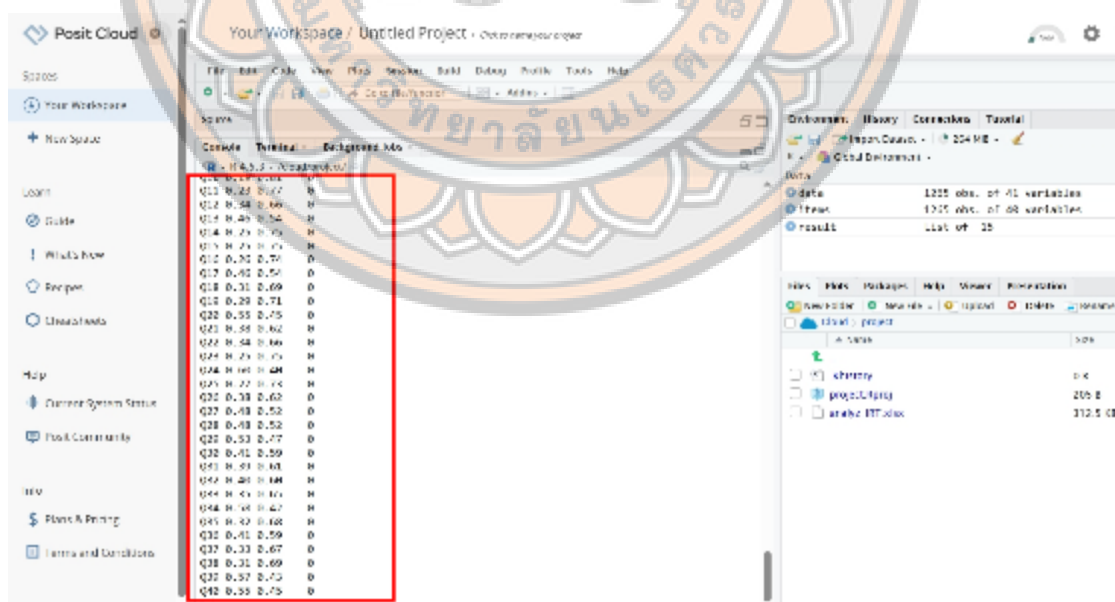
ภาพ ฏ.12 เมื่อคลิกที่ Item จะปรากฏข้อมูลใน data เพื่อแสดงข้อมูลค่าใน Excel

8. วิเคราะห์ค่า CTT

ที่ Console พิมพ์คำสั่ง `result <- alpha(items)` แล้วกด Enter



ภาพ ฎ.13 โปรแกรมจะสร้างข้อมูล result และแสดงใน data



ภาพ ฎ.14 Console ที่แสดงข้อมูล result

ผลลัพธ์ที่ได้จากการใช้คำสั่ง result\$total

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean
0.9277752	0.9290912	0.9381627	0.2467416	13.10263	0.002891842	0.6255138
sd						
0.2405548						
median_r						
0.2432609						

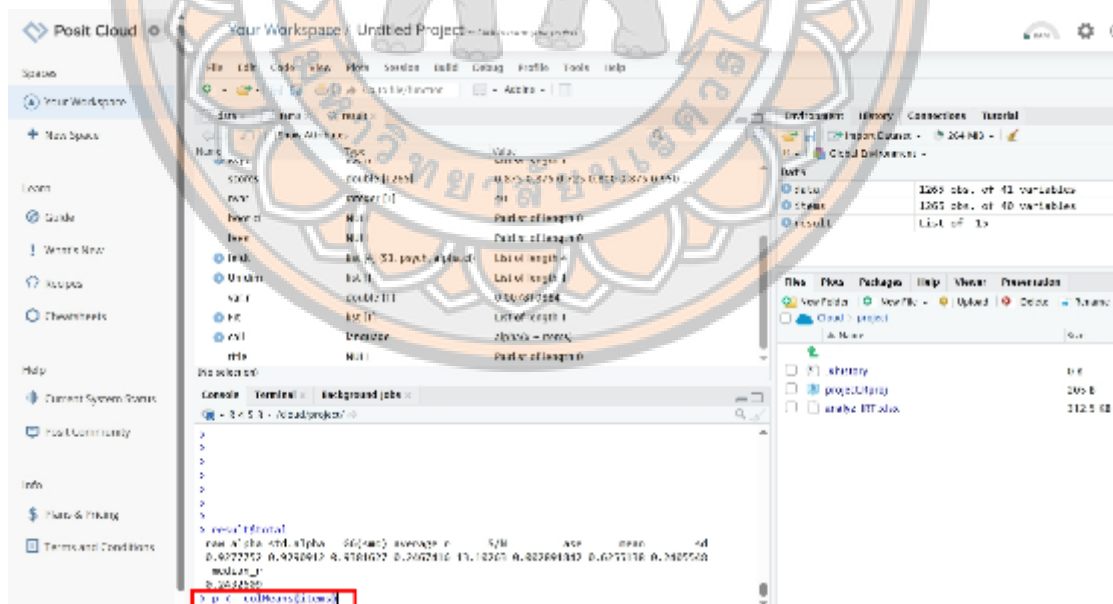
10. วิเคราะห์ค่าความยาก (p)

ที่ Console พิมพ์คำสั่ง

```
p <- colMeans(items)
```

p

จากนั้นกด Enter



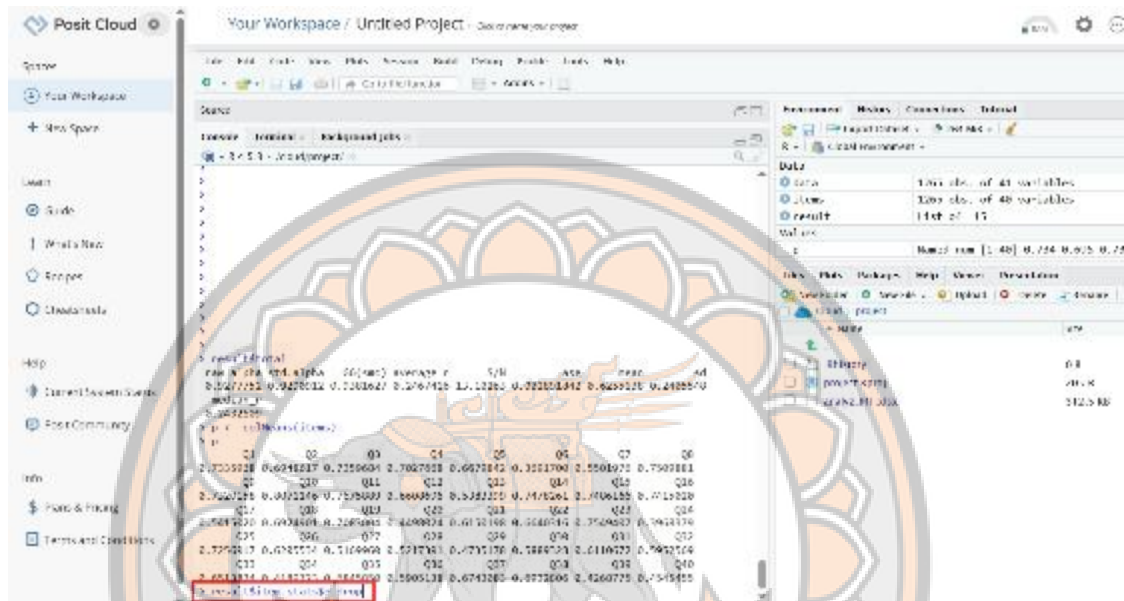
ภาพ ฏ.17 พิมพ์คำสั่งเพื่อแสดงผลค่าความยาก (p)

11. วิเคราะห์ค่าอำนาจจำแนก (r)

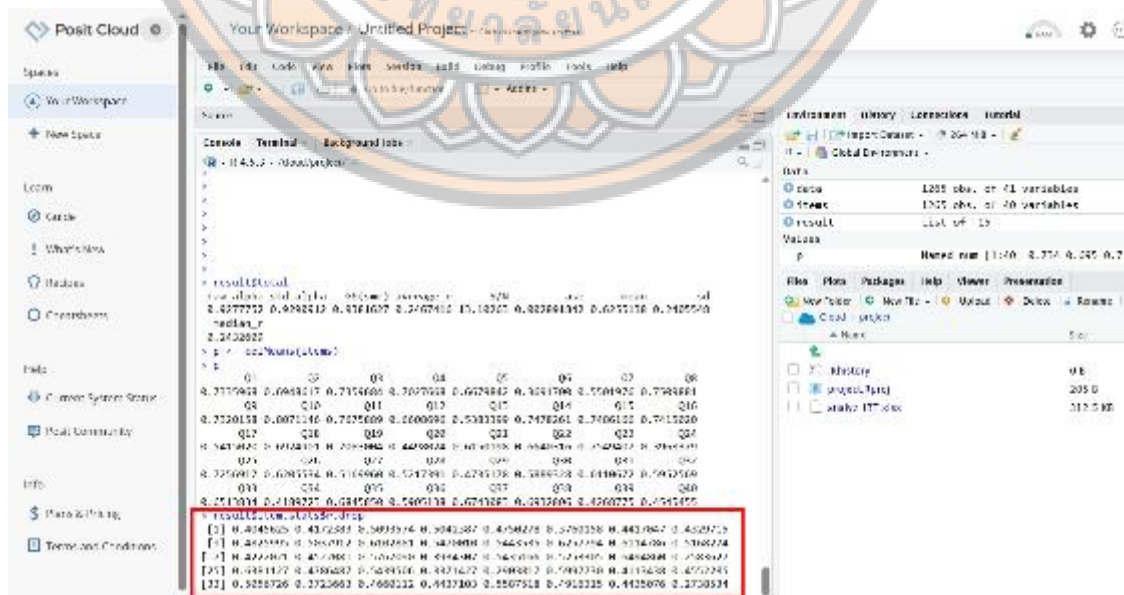
ที่ Console พิมพ์คำสั่ง

result\$item.stats หรือ result\$item.stats\$r.drop

จากนั้นกด Enter



ภาพ ฏ.19 พิมพ์คำสั่งเพื่อแสดงผลค่าอำนาจจำแนก (r)

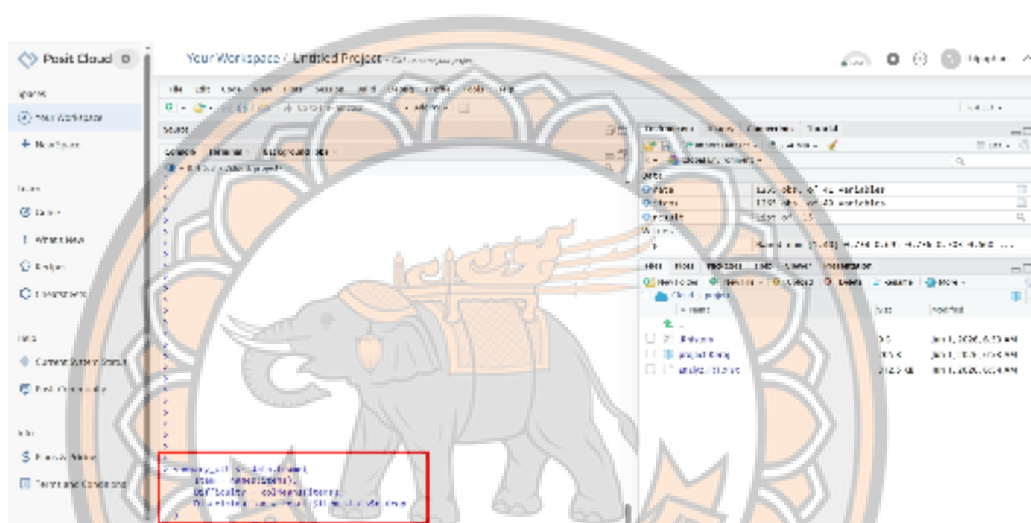


ภาพ ฏ.20 โปรแกรมแสดงผลค่าอำนาจจำแนก (r)

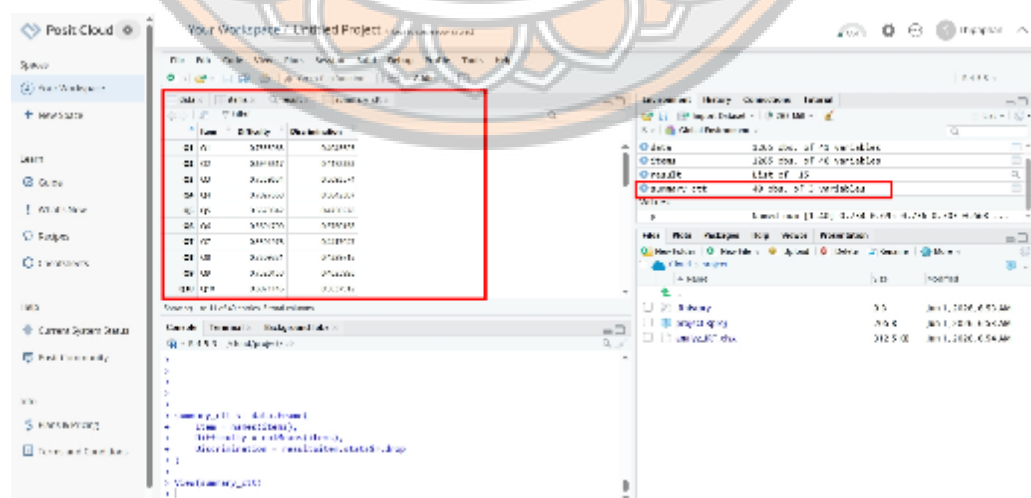
12. สร้างตารางสรุปค่าความยากและค่าอำนาจจำแนก

ที่ Console พิมพ์คำสั่ง

```
summary_ckt <- data.frame(
  Item = names(items),
  Difficulty = colMeans(items),
  Discrimination = result$item.stats$r.drop
)
```



ภาพ ฎ.21 พิมพ์คำสั่งเพื่อสร้างตารางสรุป



ภาพ ฎ.22 โปรแกรมสร้างตารางสรุปค่าความยากและค่าอำนาจจำแนก

ภาคผนวก ร ตัวอย่างคำสั่งและผลลัพธ์การวิเคราะห์ข้อมูลคุณภาพแบบวัดตามทฤษฎีการตอบสนองข้อสอบ (IRT) ด้วยโปรแกรมภาษา R เวอร์ชัน 4.5.3 ผ่านเว็บแอปพลิเคชัน Posit Cloud

1. การเข้าสู่ระบบ Posit Cloud

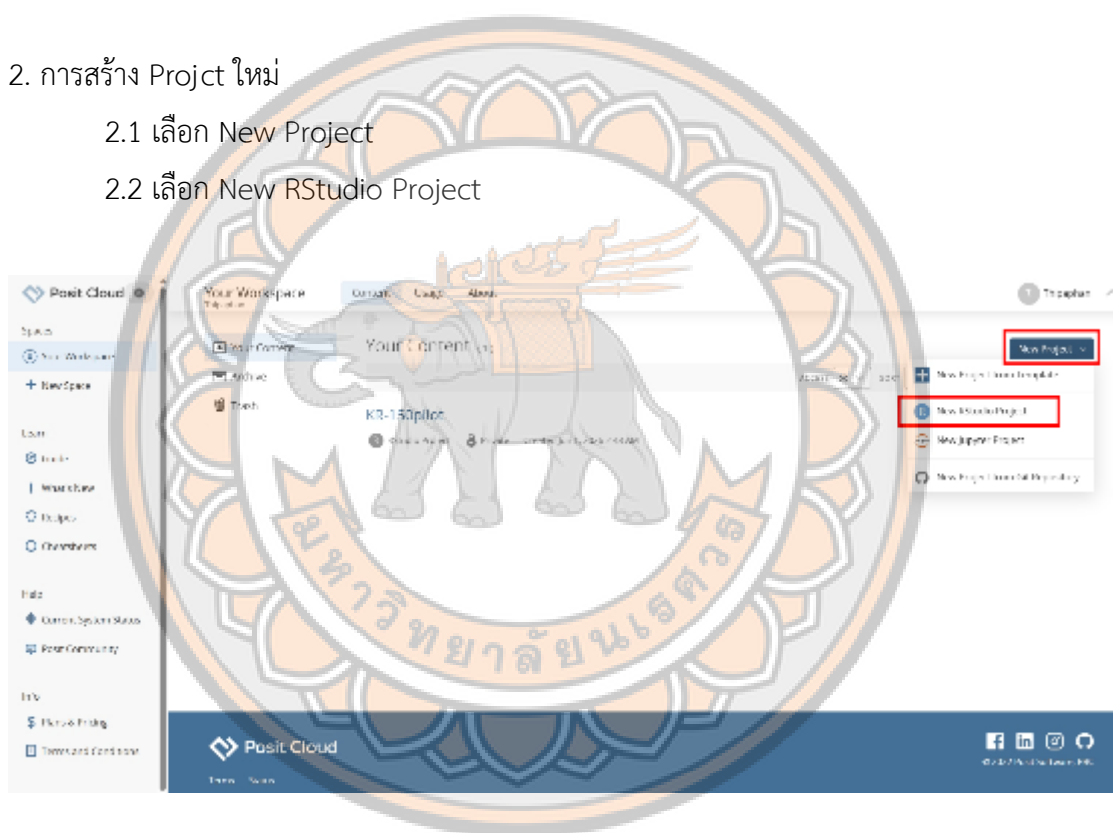
1.1 เปิดเว็บเบราว์เซอร์เข้าเว็บไซต์ <https://posit.cloud>

1.2 เข้าสู่ระบบ > เลือก Sign In > กรอกอีเมลและรหัสผ่าน

2. การสร้าง Project ใหม่

2.1 เลือก New Project

2.2 เลือก New RStudio Project



ภาพ ร.1 การสร้าง Project ใหม่

3. การนำเข้าข้อมูล

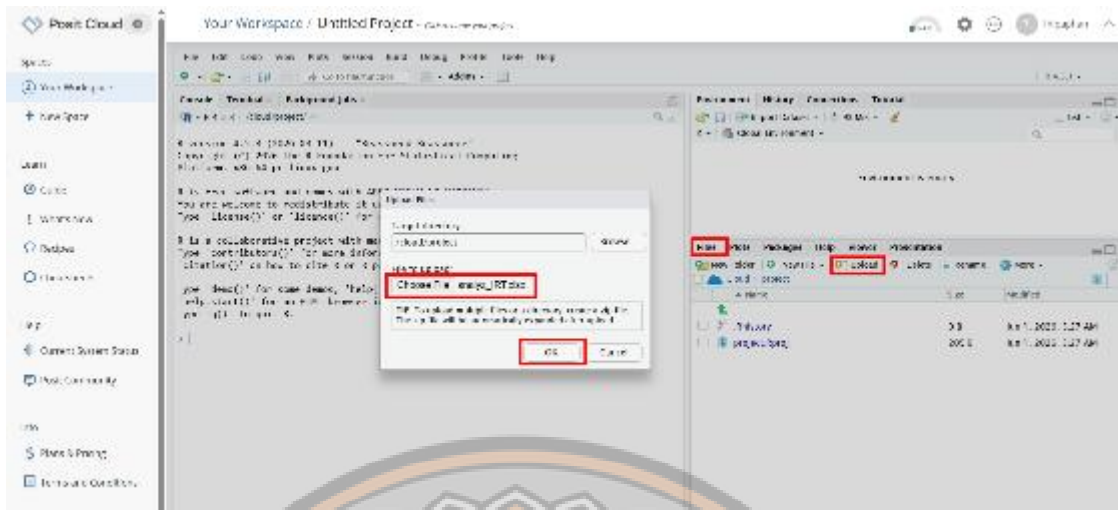
3.1 เลือกแท็บ Files

3.2 กด Upload

3.3 จะปรากฏหน้าต่าง Target directory ให้เลือก File to upload และคลิก Choose File เลือกไฟล์ที่จัดเตรียมไว้ (ไฟล์ที่ใช้ในตัวอย่างคือ analyz_IRT.xlsx)

3.4 จะปรากฏชื่อไฟล์ที่เลือก

3.5 กด OK เพื่ออัปโหลดไฟล์ข้อมูล



ภาพ ร.2 การอัปโหลดไฟล์

4. การติดตั้ง Package

ที่ Console พิมพ์คำสั่ง

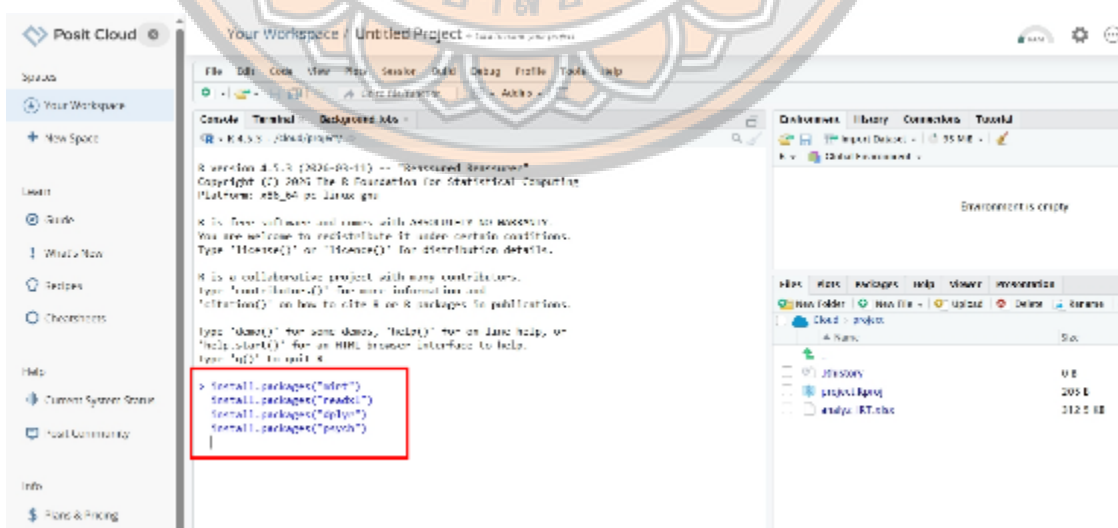
```
install.packages("mirt")
```

```
install.packages("readxl")
```

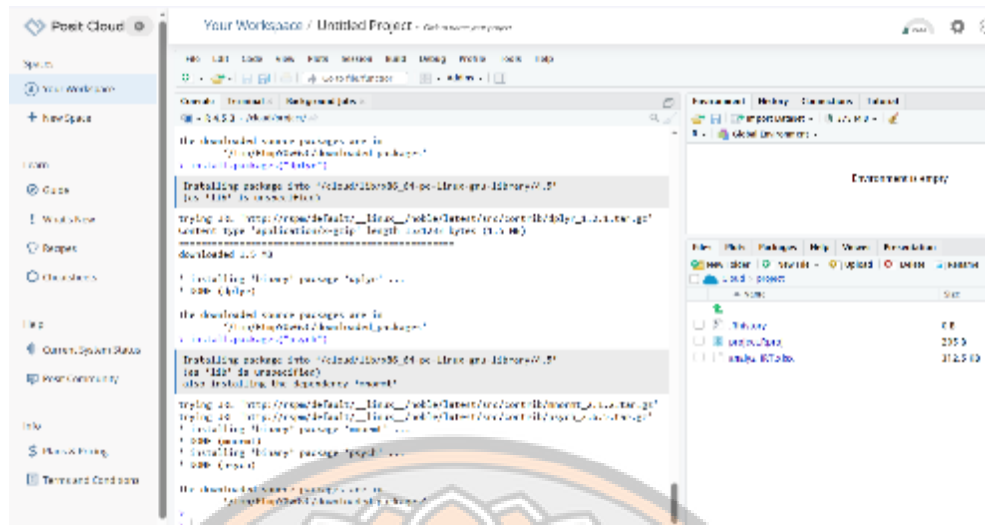
```
install.packages("dplyr")
```

```
install.packages("psych")
```

จากนั้นกด Enter



ภาพ ร.3 การติดตั้ง Package



ภาพ รูป.4 package mirt สำหรับการวิเคราะห์ IRT
และ package readxl สำหรับนำเข้าข้อมูลจาก Excel

5. การเรียกใช้ Package

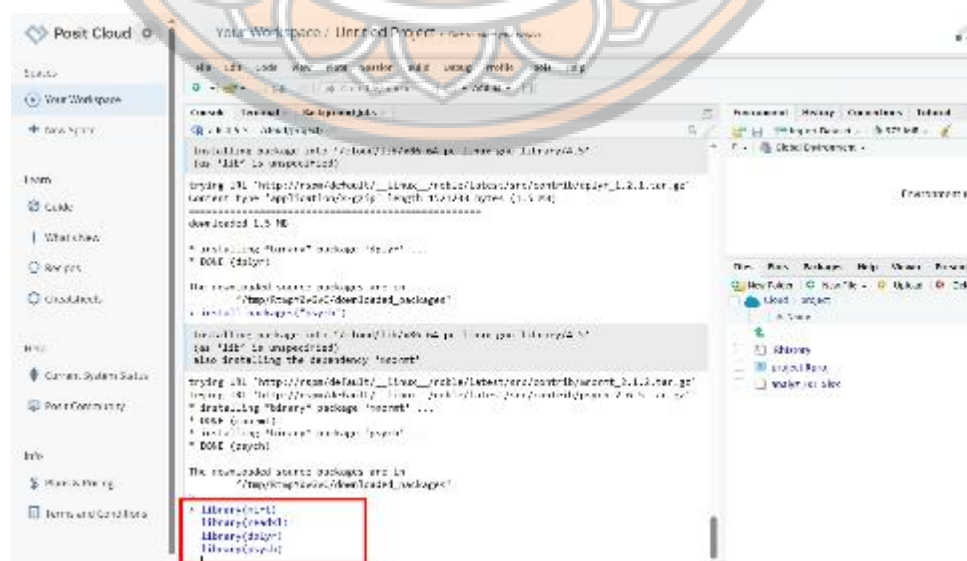
ที่ Console พิมพ์คำสั่ง

library(mirt)

library(readxl)

library(dplyr)

library(psych)



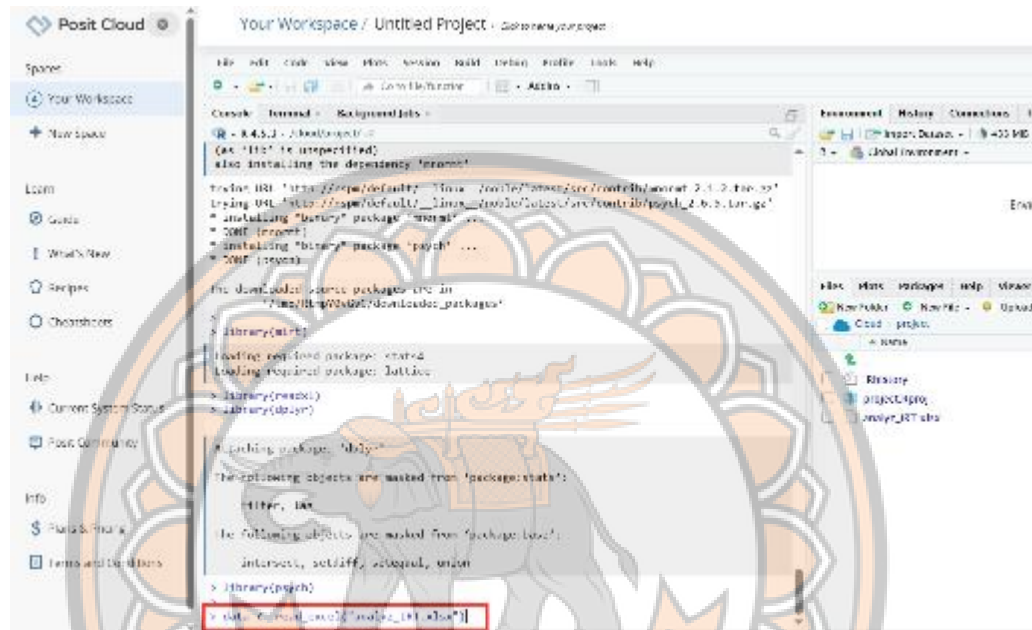
ภาพ รูป.5 การเรียกใช้ Package

6. การอ่านข้อมูล

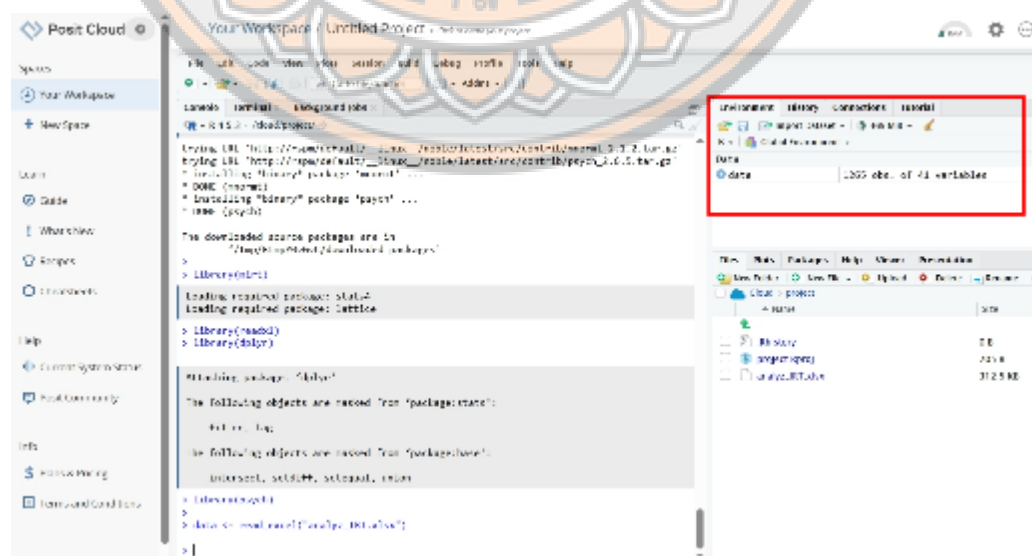
6.1 ที่ Console พิมพ์คำสั่ง

```
data <- read_excel("analyz_IRT.xlsx")
```

จากนั้นกด Enter



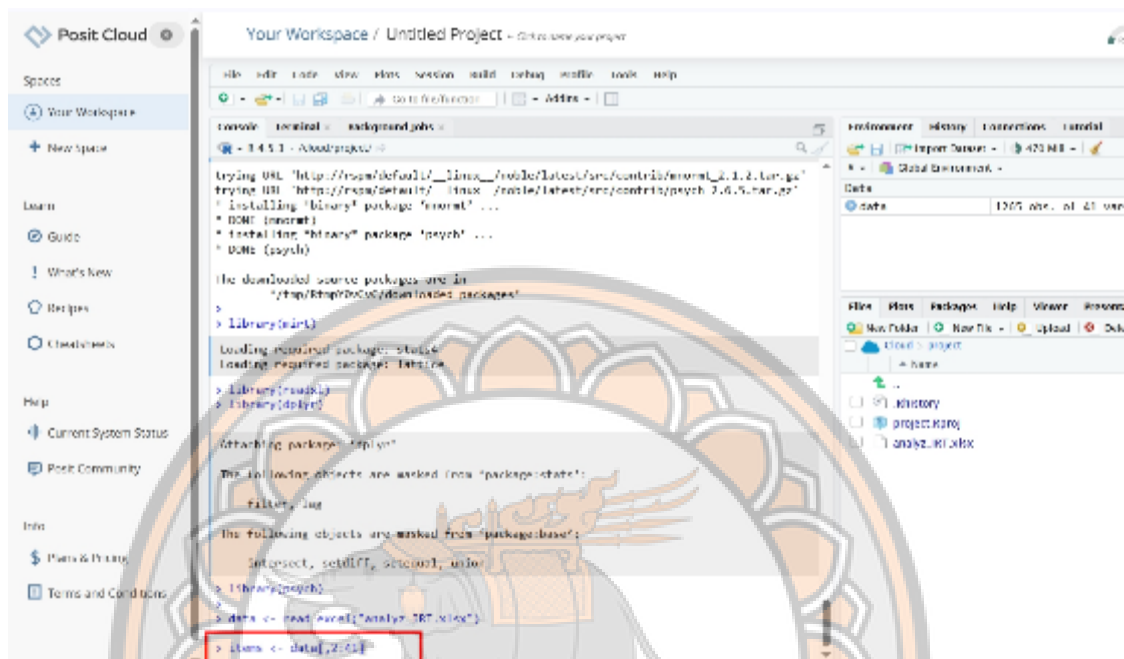
ภาพ ฐ.6 การอ่านข้อมูล



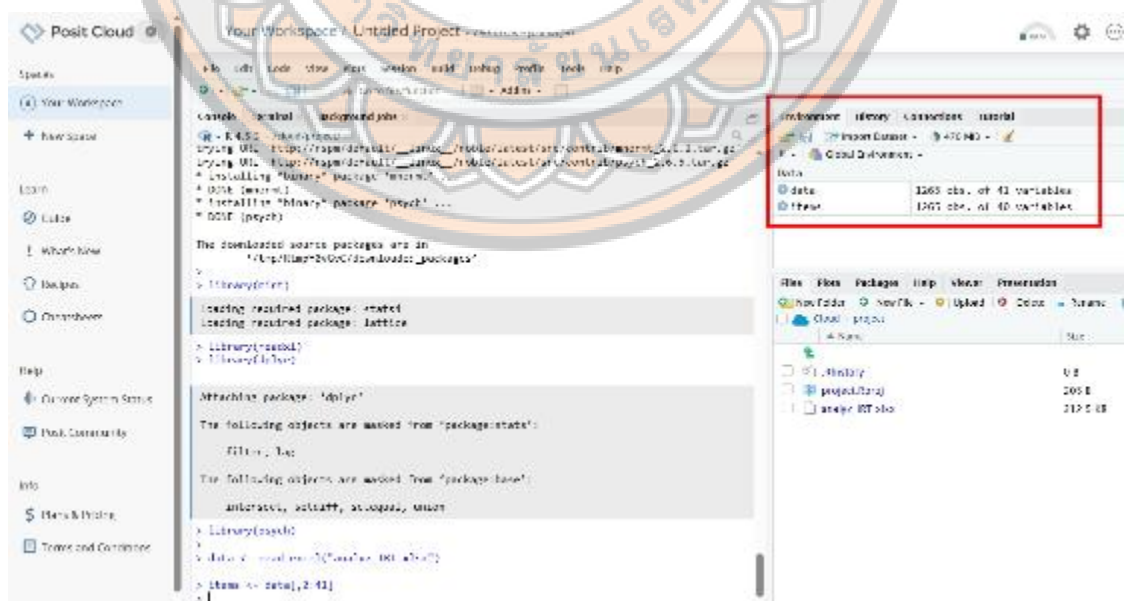
ภาพ ฐ.7 ตรวจสอบข้อมูล

6.2 เขียนคำสั่งเลือกเฉพาะข้อสอบ

ที่ Console พิมพ์คำสั่ง `items <- data[,2:41]` จากนั้นกด Enter



ภาพ ร.8 คำสั่งเลือกเฉพาะข้อสอบ



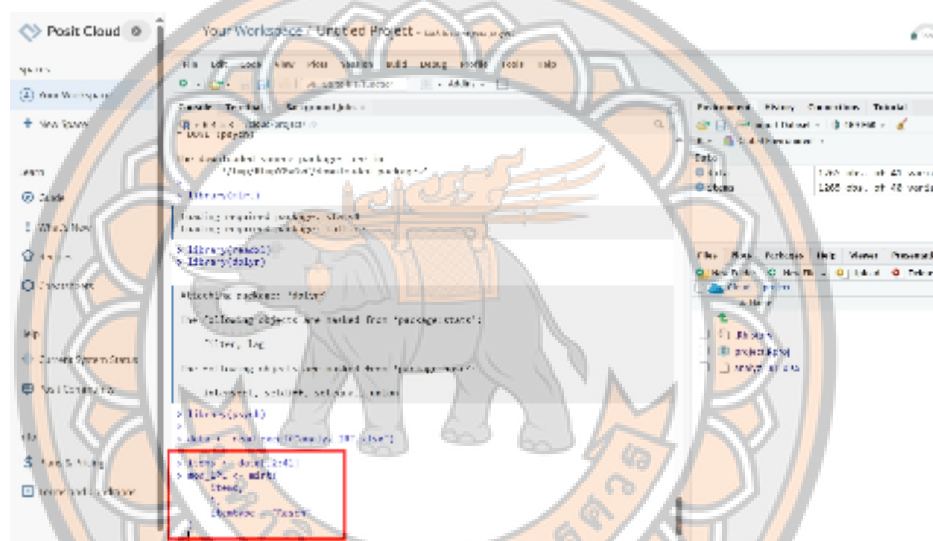
ภาพ ร.9 ผลจะปรากฏในช่อง Data

7. เปรียบเทียบโมเดล 1PL 2PL 3PL

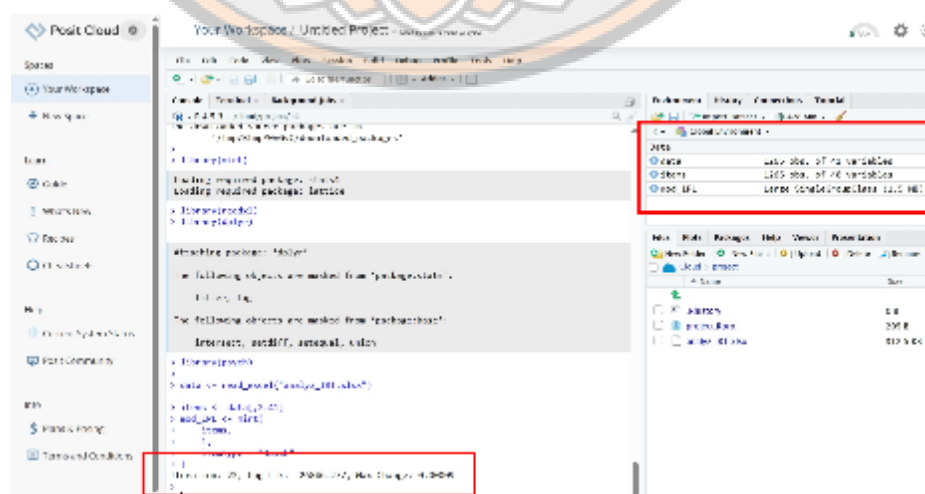
7.1 โมเดล 1 PL

ที่ Console พิมพ์คำสั่ง

```
mod_1PL <- mirt(
  items,
  1,
  itemtype = "Rasch"
)
```



ภาพ ร.10 คำสั่งโมเดล 1 PL

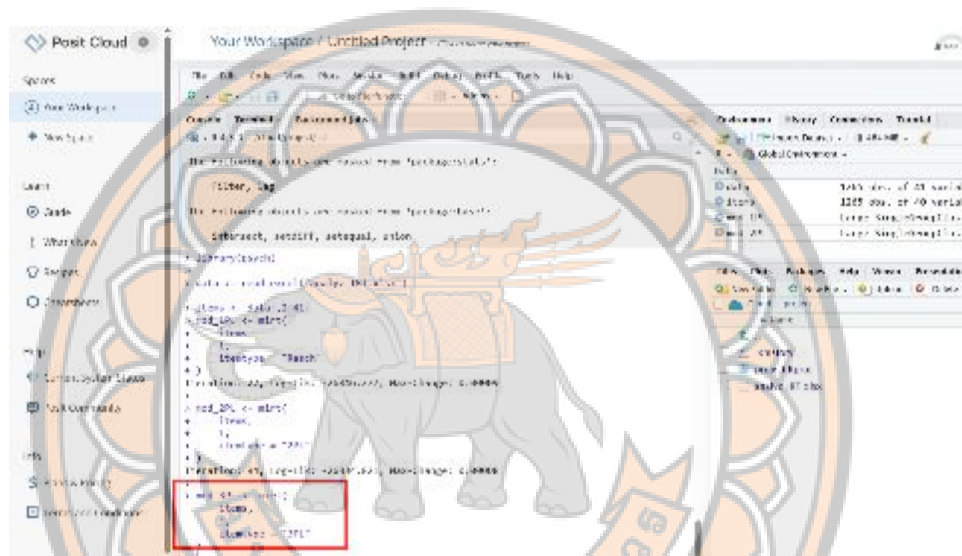


ภาพ ร.11 ผลที่ได้จากคำสั่ง โมเดล 1 PL

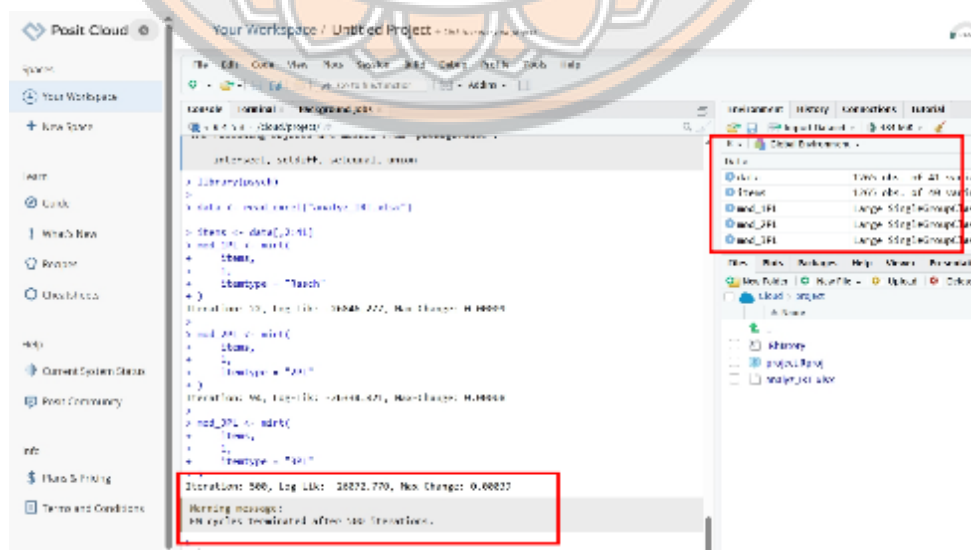
7.3 โมเดล 3PL

ที่ Console พิมพ์คำสั่ง

```
mod_3PL <- mirt(
  items,
  1,
  itemtype = "3PL"
)
```



ภาพ ฐ.14 คำสั่งโมเดล 3 PL



ภาพ ฐ.15 ผลที่ได้จากคำสั่งโมเดล 3 PL

7.3 เปรียบเทียบโมเดล

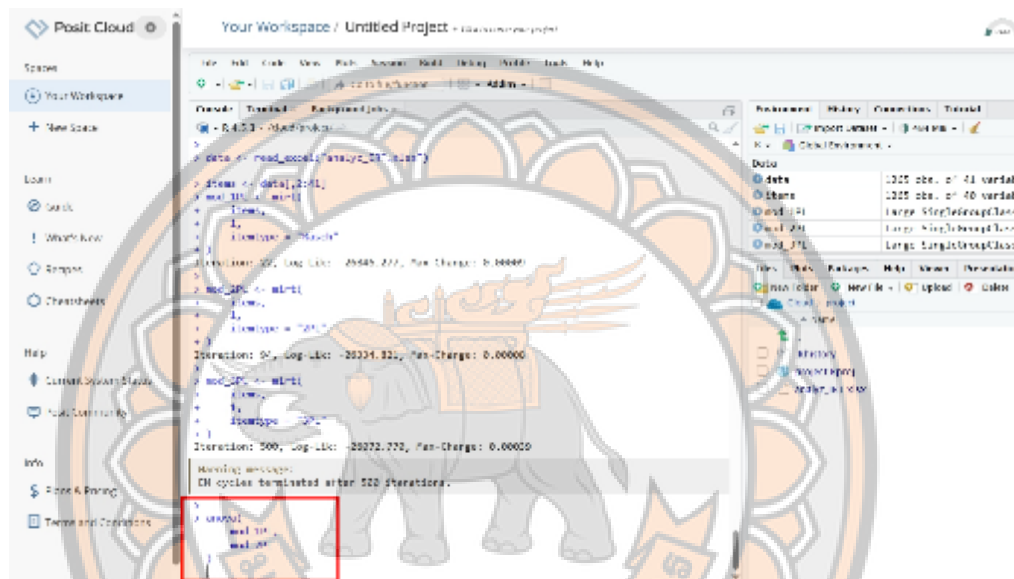
ที่ Console พิมพ์คำสั่ง

anova(

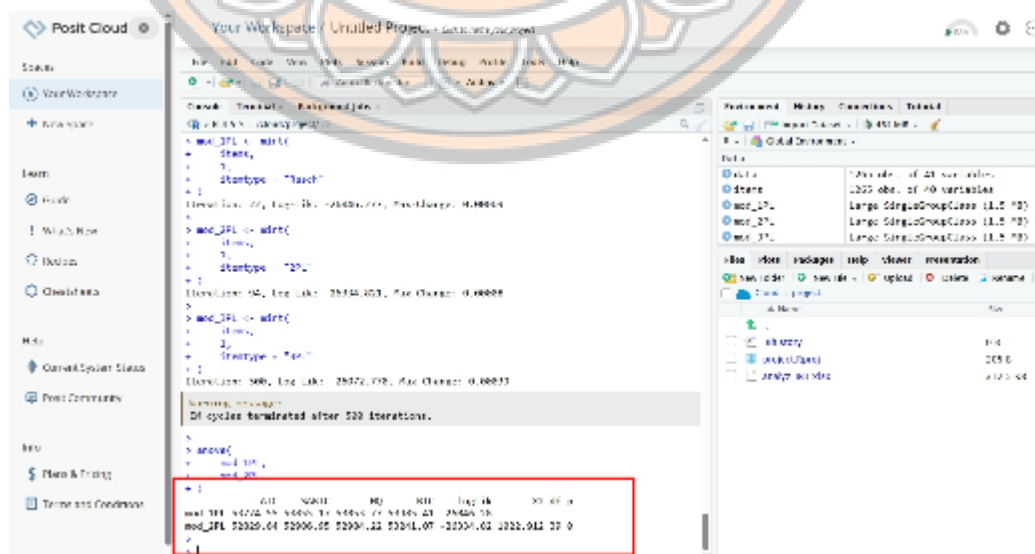
mod_1PL,

mod_2PL

)



ภาพ ฐ.16 คำสั่งเปรียบเทียบโมเดล



ภาพ ฐ.17 ผลที่ได้จากคำสั่งเปรียบเทียบโมเดล

8. การเขียนคำสั่ง ดึงค่าพารามิเตอร์ a (อำนาจจำแนก) และ ค่าพารามิเตอร์ b (ความยาก)

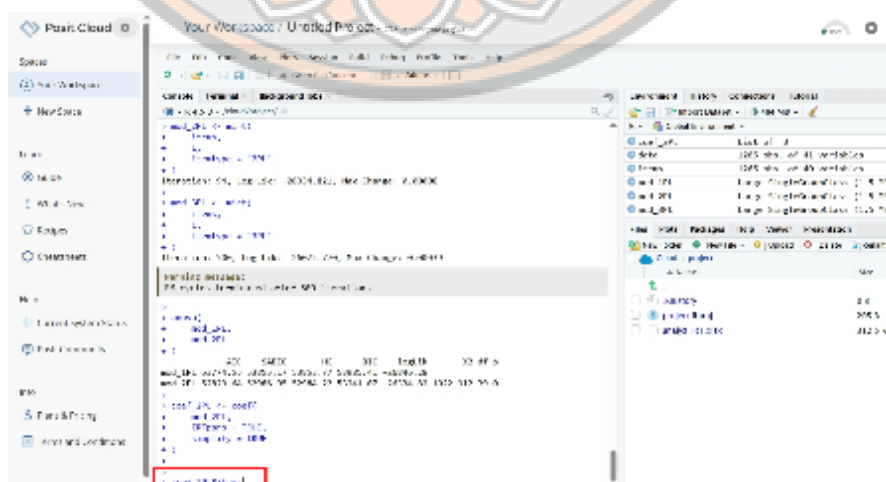
8.1 ที่ Console พิมพ์คำสั่งเพื่อเตรียมดึงค่าพารามิเตอร์ a และ b

```
coef_2PL <- coef(
mod_2PL,
IRTpars = TRUE,
simplify = TRUE
)
```

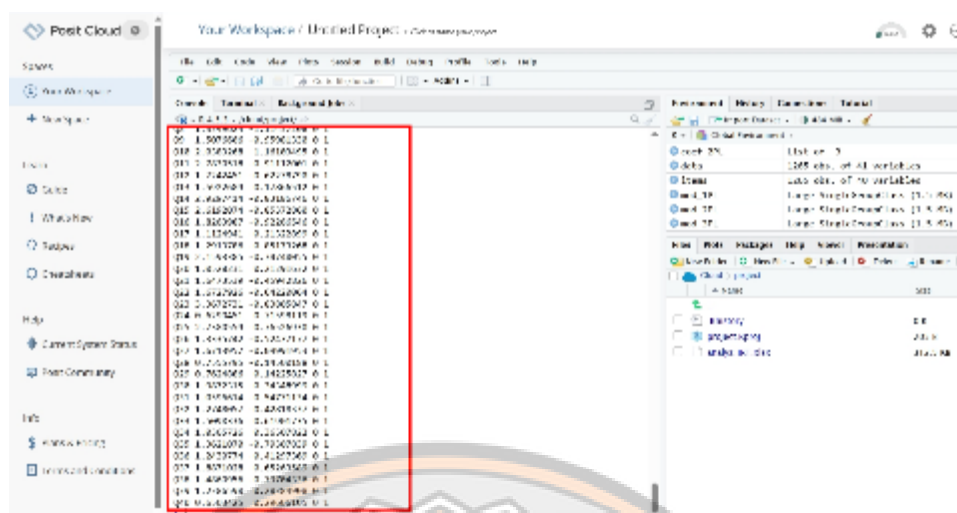


ภาพ ฐ.18 คำสั่งเพื่อเตรียมดึงค่าพารามิเตอร์ a (อำนาจจำแนก) และ ค่าพารามิเตอร์ b (ความยาก)

8.2 จากนั้นที่ Console พิมพ์คำสั่ง coef_2PL\$items เพื่อดึงค่าพารามิเตอร์ a และ b



ภาพ ฐ.19 คำสั่งดึงค่าพารามิเตอร์ a (อำนาจจำแนก) และ ค่าพารามิเตอร์ b (ความยาก)

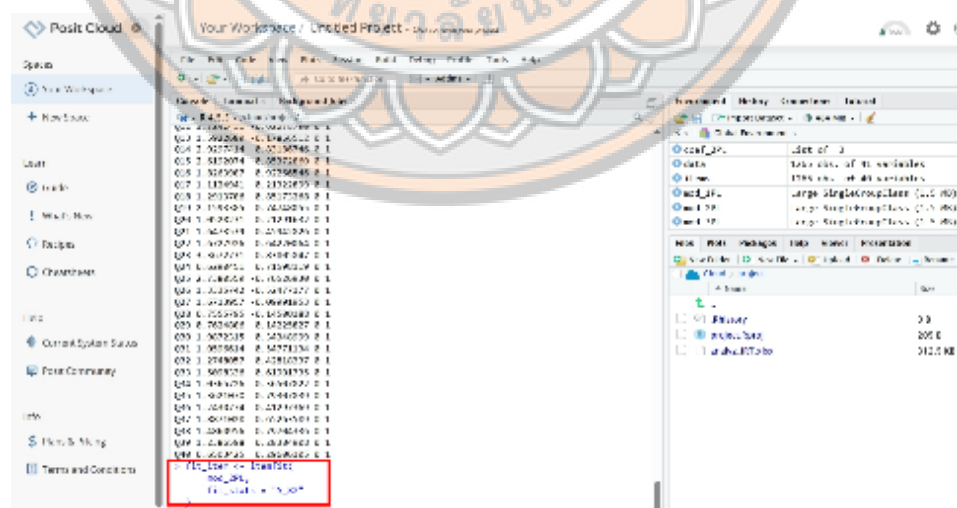


ภาพ รูป.20 ผลลัพธ์ที่ได้ค่าพารามิเตอร์ a (อำนาจจำแนก) และ ค่าพารามิเตอร์ b (ความยาก)

9. การเขียนคำสั่งเพื่อให้แสดงค่า Item Fit

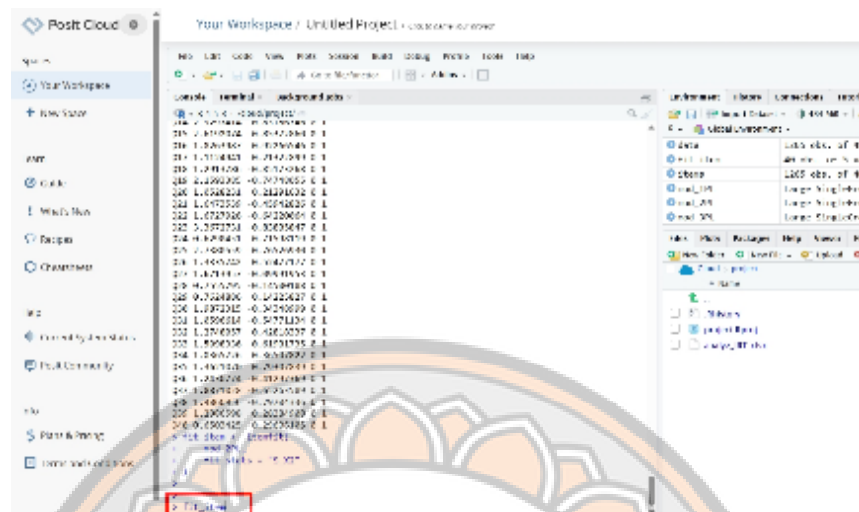
9.1 ที่ Console พิมพ์คำสั่ง

```
fit_item <- itemfit(
  mod_2PL,
  fit_stats = "S_X2"
)
```

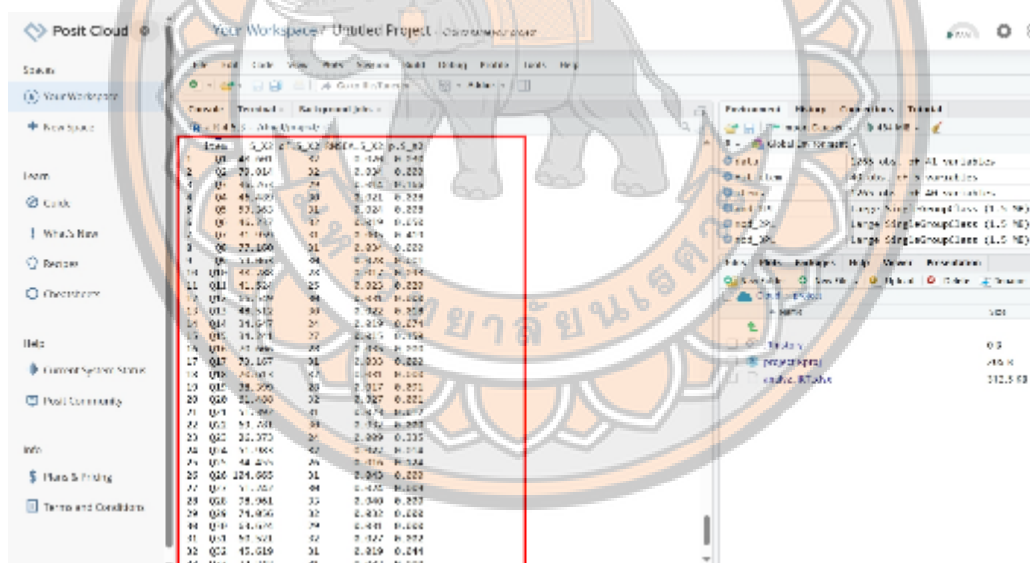


ภาพ รูป.21 คำสั่งให้แสดงค่า Item Fit

9.2 ที่ Console พิมพ์คำสั่ง fit_item เพื่อแสดงผล



ภาพ รูป.22 คำสั่ง fit_item เพื่อแสดงผล



ภาพ รูป.23 ผลลัพธ์ที่ได้จากคำสั่ง fit_item

ค่าที่ได้ประกอบด้วย

$S-X^2$

df

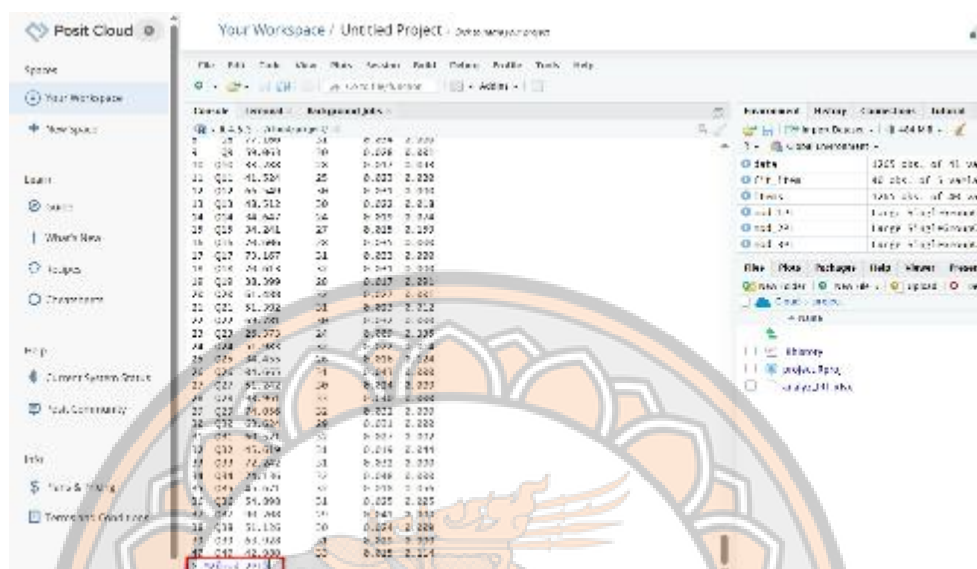
RMSEA

p-value

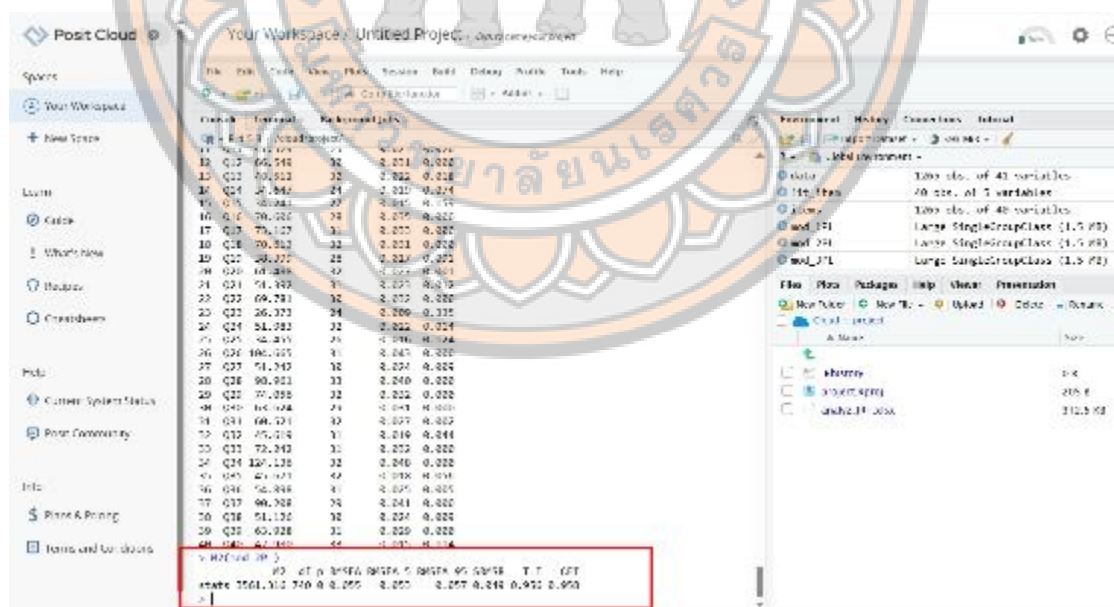
ของทุกข้อ

10. การเขียนคำสั่ง Global Fit

ที่ Console พิมพ์คำสั่ง M2(mod_2PL) จากนั้นกด Enter



ภาพ ร.24 คำสั่ง Global Fit เพื่อแสดงผล



ภาพ ร.25 ผลลัพธ์ที่ได้จากคำสั่ง Global Fit

ผลลัพธ์ที่ได้คือค่าของ M^2 df RMSEA SRMSR TLI CFI

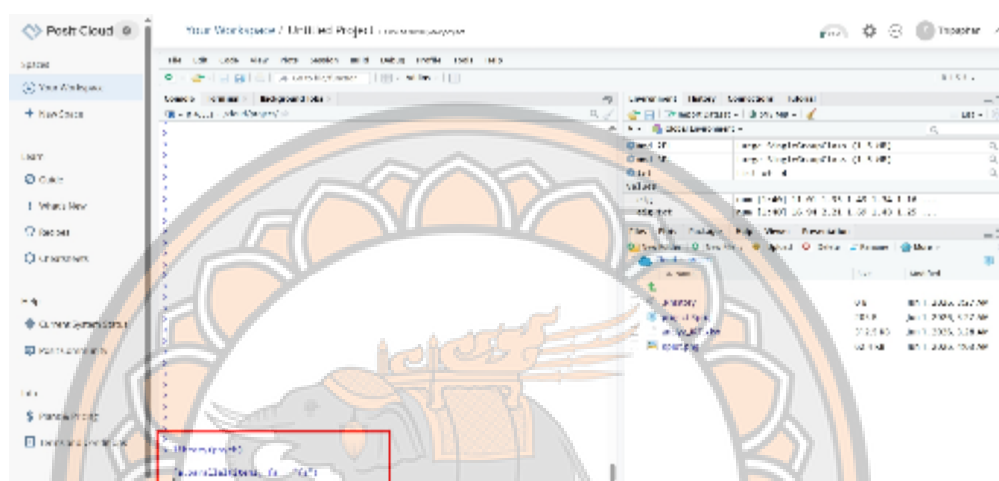
11. การเขียนคำสั่งตรวจสอบความเป็นเอกมิติ (Unidimensionality)

11.1 ตรวจสอบจำนวนปัจจัยเบื้องต้น

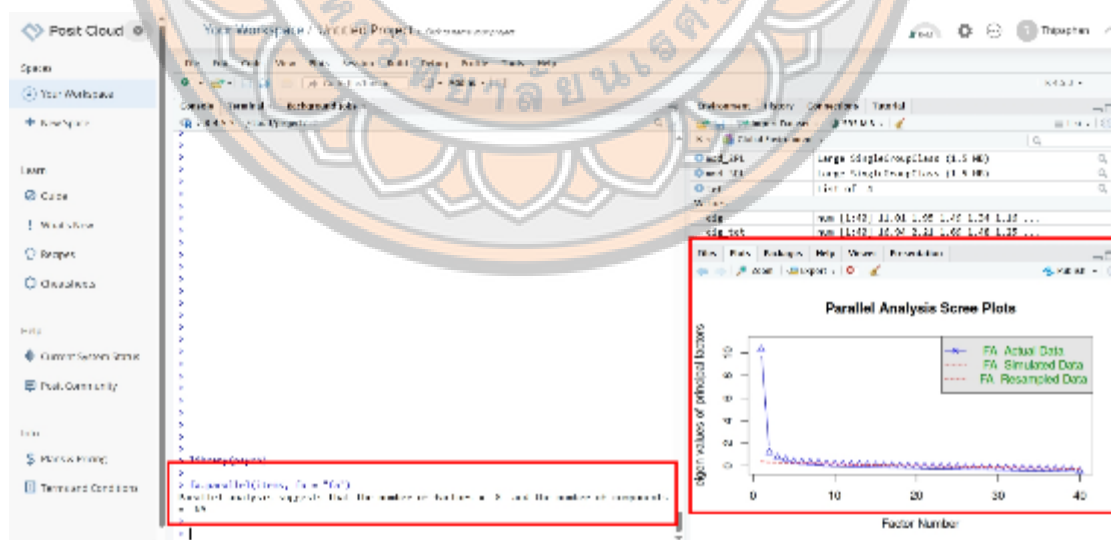
ที่ Console พิมพ์คำสั่ง

```
library(psych)
```

```
fa.parallel(items, fa = "fa")
```



ภาพ ฐ.26 คำสั่งตรวจสอบจำนวนปัจจัยเบื้องต้น



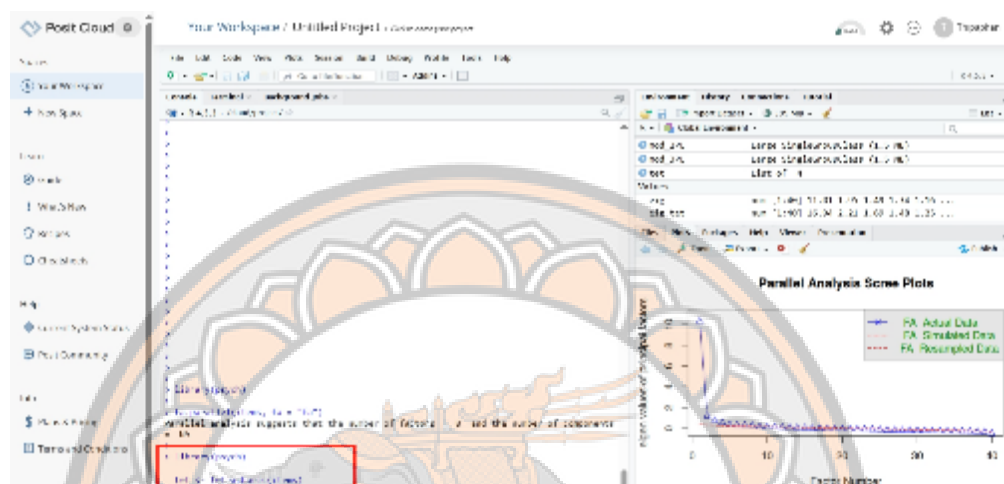
ภาพ ฐ.27 ผลลัพธ์ที่ได้จากการคำสั่งตรวจสอบจำนวนปัจจัยเบื้องต้นด้วยวิธี Parallel Analysis โดยเปรียบเทียบค่า Eigenvalue ที่ได้จากข้อมูลจริงกับค่า Eigenvalue จากข้อมูลสุ่ม

11.2 คำนวณ Tetrachoric Correlation Matrix

ที่ Console พิมพ์คำสั่ง

```
library(psych)
```

```
tet <- tetrachoric(items)
```



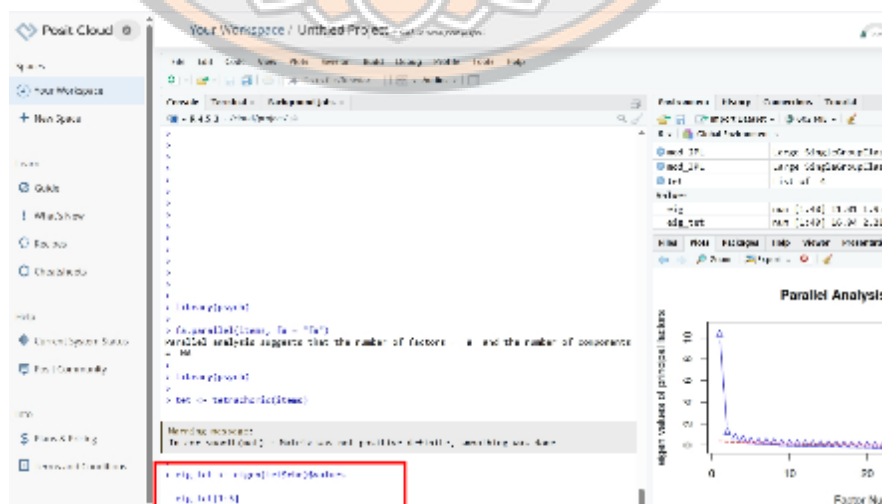
ภาพ รูป.28 คำสั่งคำนวณ Tetrachoric Correlation Matrix

11.3 ตรวจสอบค่า Eigenvalue

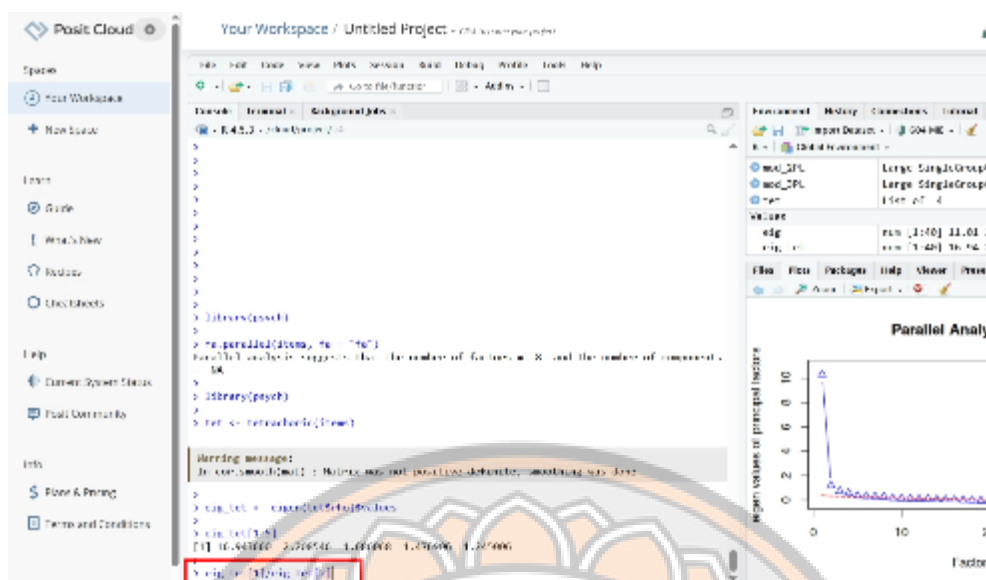
ที่ Console พิมพ์คำสั่ง

```
eig_tet <- eigen(tet$rho)$values
```

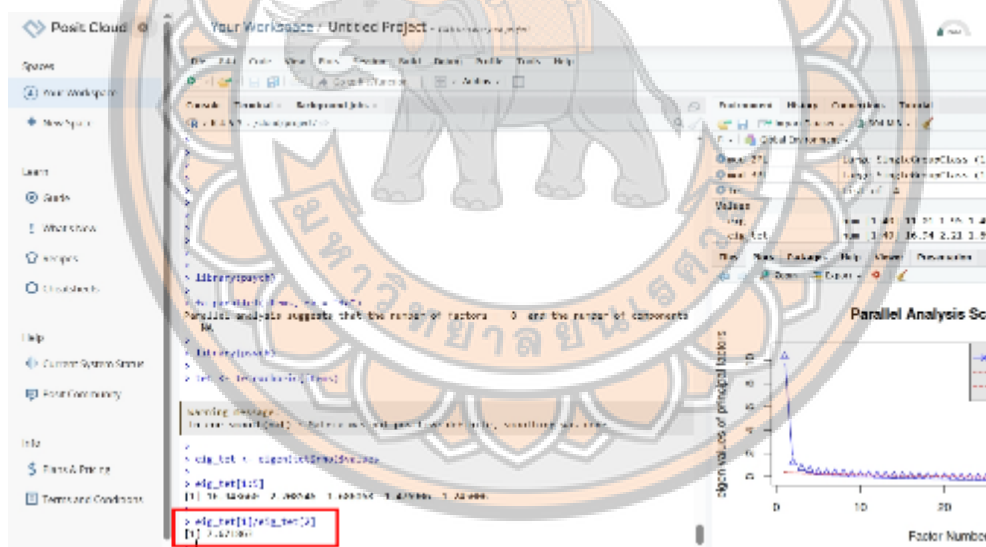
```
eig_tet[1:5]
```



ภาพ รูป.29 คำสั่งตรวจสอบค่า Eigenvalue



ภาพ ฐ.31 คำสั่งคำนวณ Eigenvalue Ratio



ภาพ ฐ.32 ผลลัพธ์ที่ได้จากคำสั่งคำนวณ Eigenvalue Ratio

ผลลัพธ์ที่ได้ เท่ากับ 7.671863

การตรวจสอบความเป็นเอกมิติของแบบวัดเพิ่มเติมโดยวิเคราะห์ค่า Eigenvalue จากเมทริกซ์สหสัมพันธ์แบบ Tetrachoric Correlation ซึ่งเหมาะสมสำหรับข้อมูลข้อสอบแบบทวิภาค (0–1) ใช้คำสั่งดังนี้

```
library(psych)
tet <- tetrachoric(items)
eig_tet <- eigen(tet$rho)$values
eig_tet[1]/eig_tet[2]
```

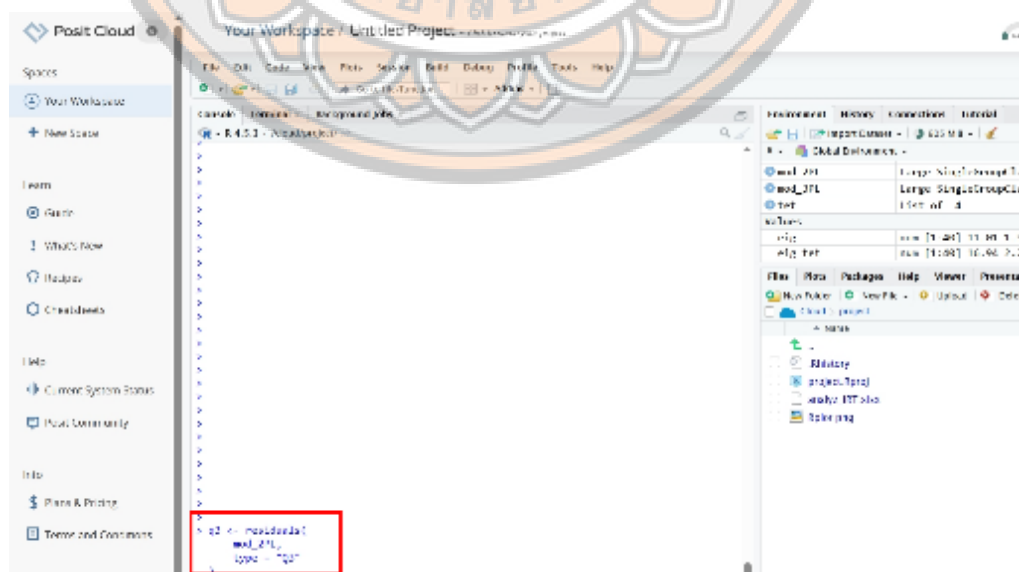
ผลการวิเคราะห์พบว่า ค่า Eigenvalue ลำดับที่ 1 เท่ากับ 16.944 และค่า Eigenvalue ลำดับที่ 2 เท่ากับ 2.209 ส่งผลให้อัตราส่วน Eigenvalue Ratio เท่ากับ 7.67 : 1 ซึ่งสูงกว่าเกณฑ์ 3 : 1 จึงสนับสนุนสมมติฐานความเป็นเอกมิติของแบบวัด

12. การตรวจสอบ Local Independence

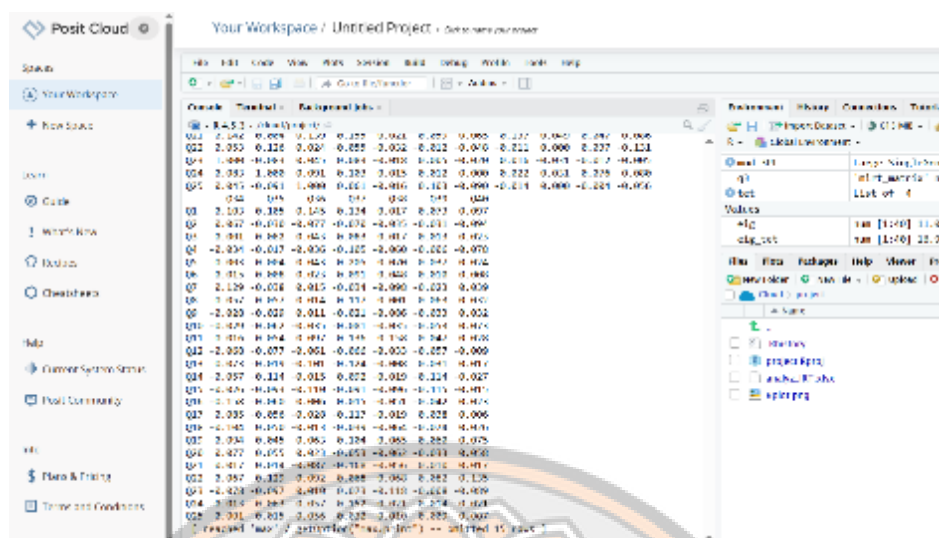
12.1 คำสั่งสร้างเมทริกซ์ Q3

ที่ Console พิมพ์คำสั่ง

```
q3 <- residuals(
  mod_2PL,
  type = "Q3"
)
```



ภาพ ฐ.33 คำสั่งสร้างเมทริกซ์ Q3



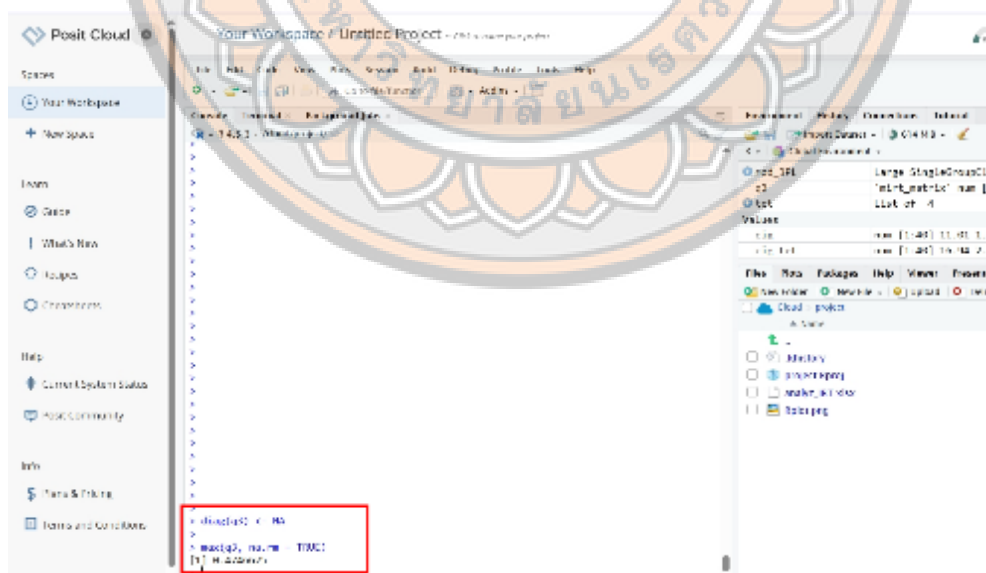
ภาพ ฐ.34 ผลลัพธ์ที่ได้จากคำสั่งสร้างเมทริกซ์ Q3

12.2 คำสั่งตรวจสอบค่า Q3 สูงสุด

ที่ Console พิมพ์คำสั่ง

```
diag(q3) <- NA
```

```
max(q3, na.rm = TRUE)
```



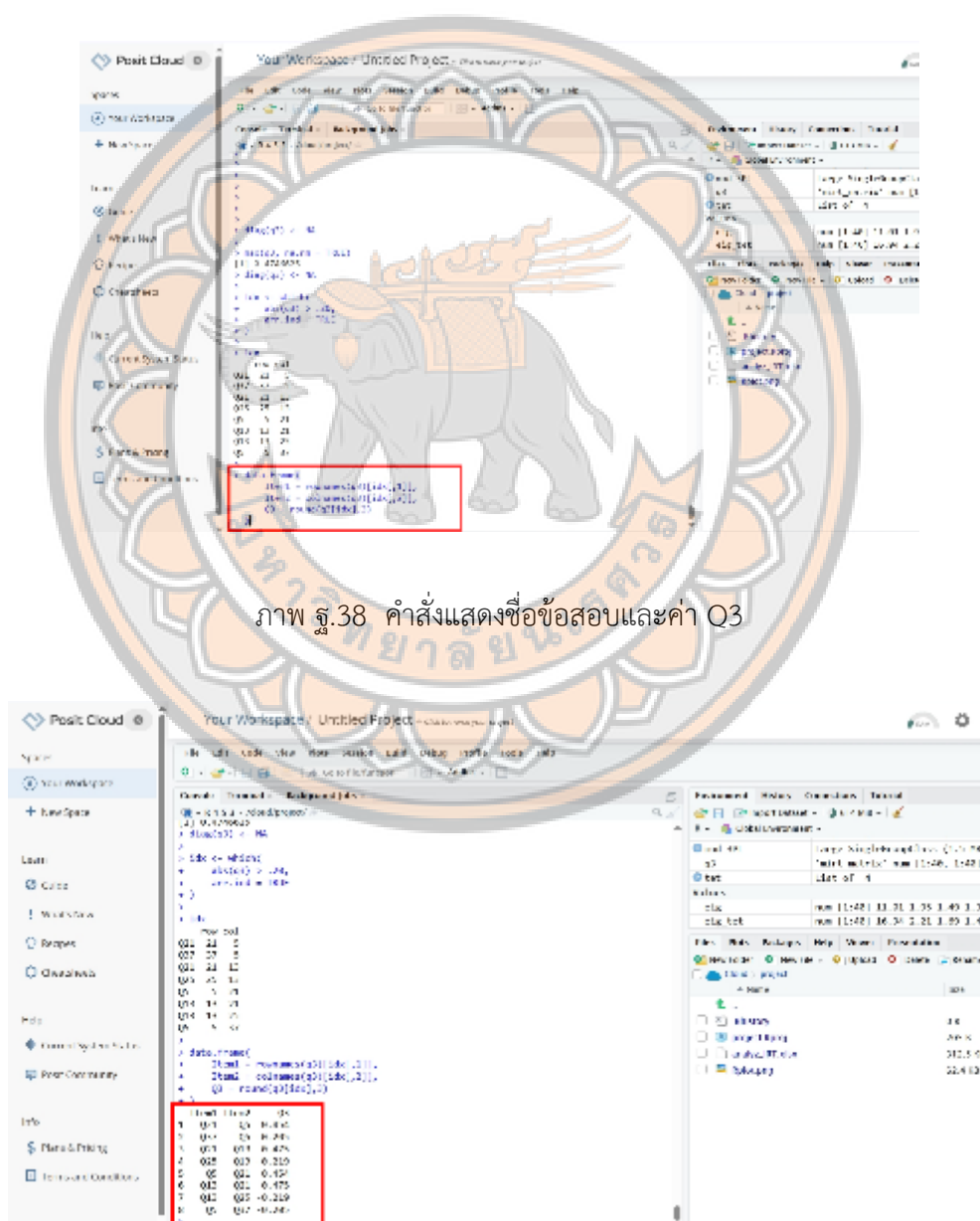
ภาพ ฐ.35 คำสั่งตรวจสอบค่า Q3 สูงสุดและผลลัพธ์ที่ได้

ผลลัพธ์ที่ได้ 0.475

12.4 คำสั่งแสดงชื่อข้อสอบและค่า Q3

ที่ Console พิมพ์คำสั่ง

```
data.frame(
  Item1 = rownames(q3)[idx[,1]],
  Item2 = colnames(q3)[idx[,2]],
  Q3 = round(q3[idx],3)
)
```

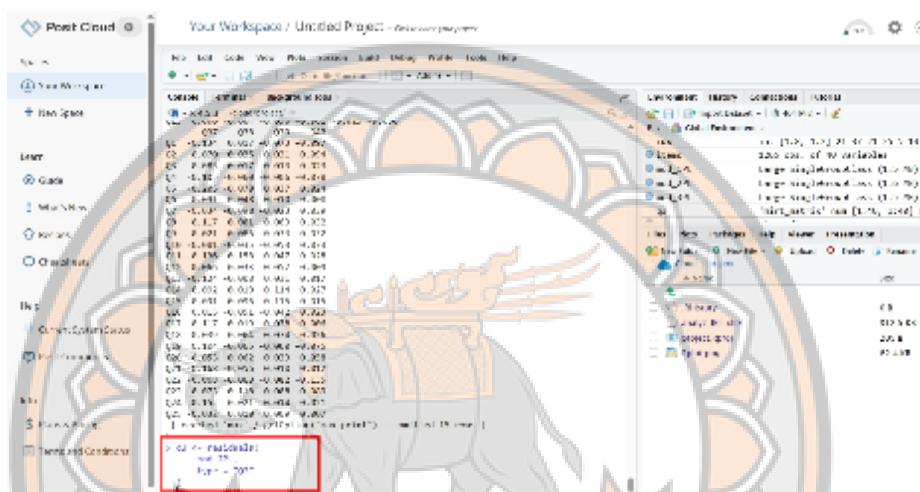


ภาพ ร.39 ผลลัพธ์ที่ได้จากคำสั่งค้นหาข้อสอบที่เกินเกณฑ์

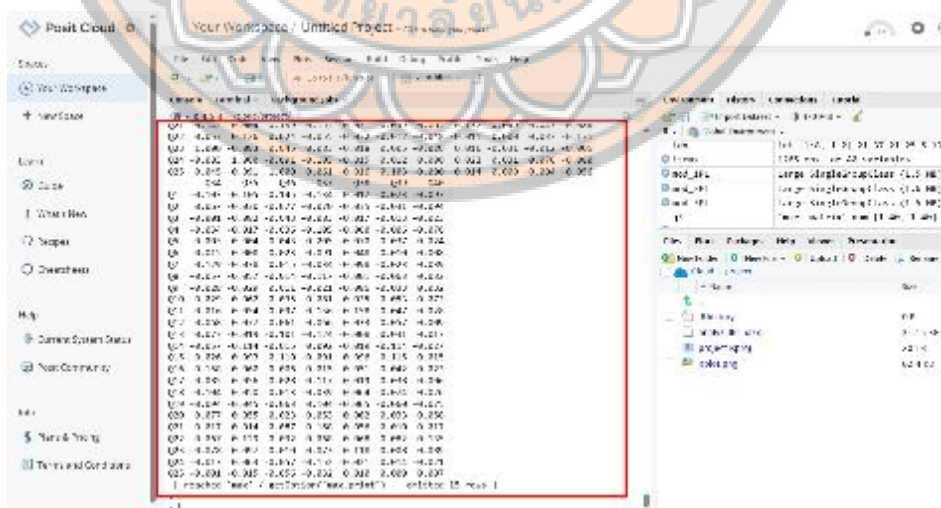
12.5 คำสั่งค่า Mean / Median / Quartile ของ Q3

ที่ Console พิมพ์คำสั่งเตรียมไฟล์ข้อมูล

```
q3 <- residuals(
  mod_2PL,
  type = "Q3"
)
```



ภาพ รุ.40 คำสั่งคำสั่งเตรียมไฟล์ข้อมูล

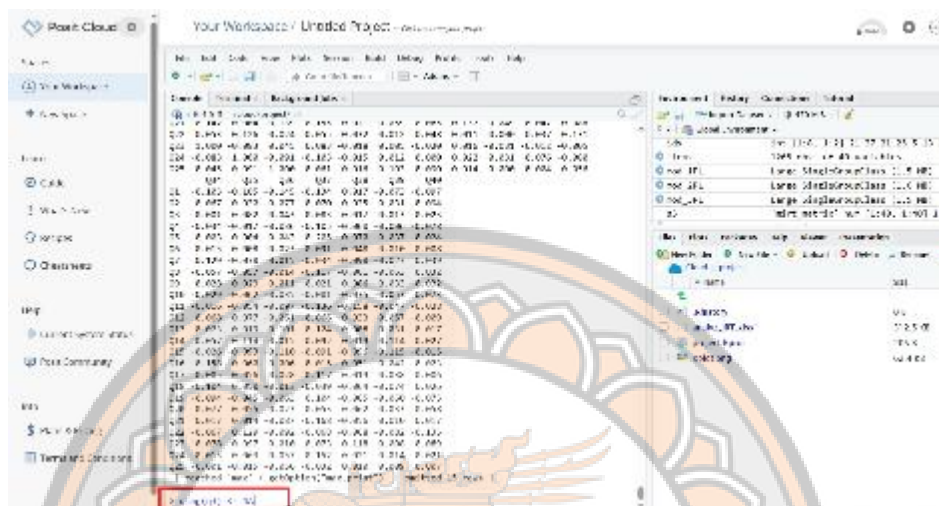


ภาพ รุ.41 ผลลัพธ์ที่ได้จากคำสั่งเตรียมไฟล์ข้อมูล

12.6 คำสั่งให้ตัด diagonal ออก

ที่ Console พิมพ์คำสั่ง

```
diag(q3) <- NA
```

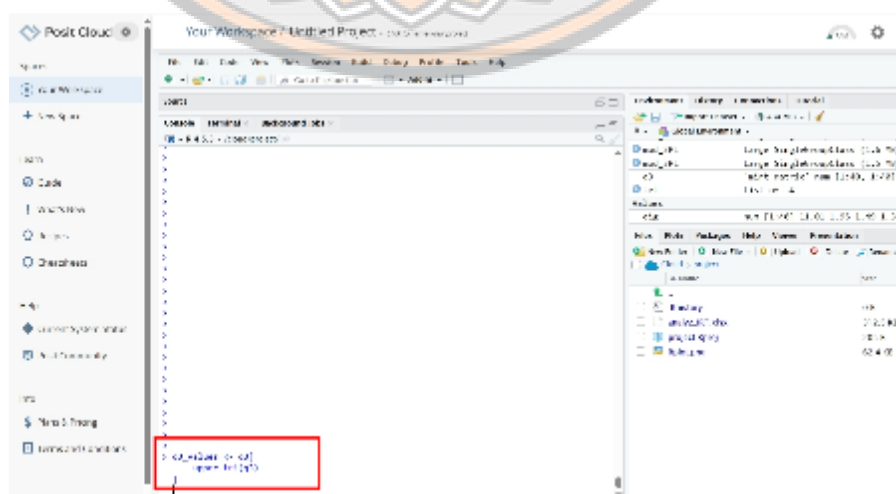


ภาพ ร.42 คำสั่งให้ตัด diagonal ออก

12.7 คำสั่งดึงเฉพาะค่าที่ไม่ใช่ NA

ที่ Console พิมพ์คำสั่ง

```
q3_values <- q3[
  upper.tri(q3)
]
```



ภาพ ร.43 คำสั่งดึงเฉพาะค่าที่ไม่ใช่ NA

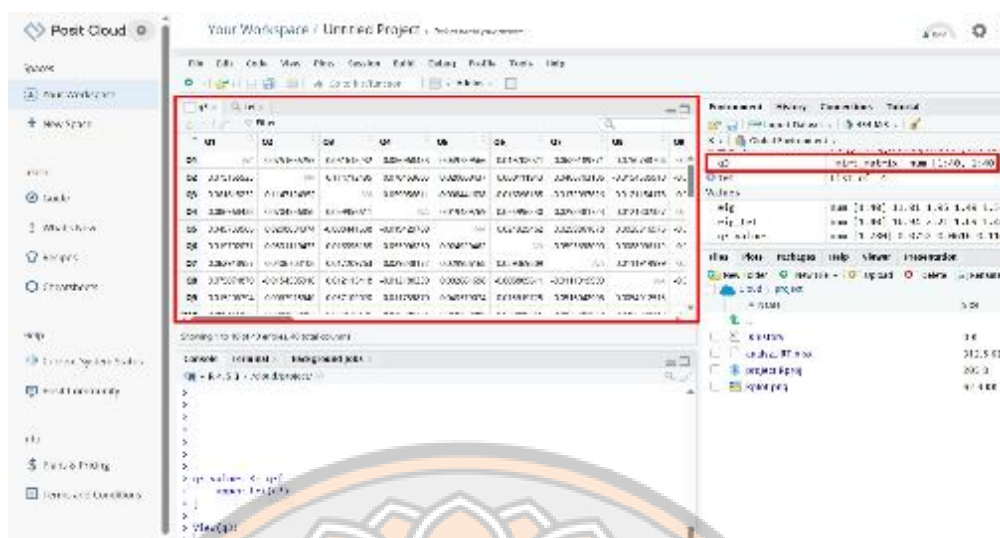


Figure 44 shows the RStudio interface. The main window displays a data frame with columns Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9, and Q10. The Q3 column is highlighted with a red box. The console shows the output of the mean function, with the value 11.57143 highlighted.

ภาพ ร.44 ผลลัพธ์ที่ได้จากคำสั่งดึงเฉพาะค่าที่ไม่ใช่ NA

12.8 คำสั่งหาค่า Mean

ที่ Console พิมพ์คำสั่ง

```
mean(
  q3_values,
  na.rm = TRUE
)
```

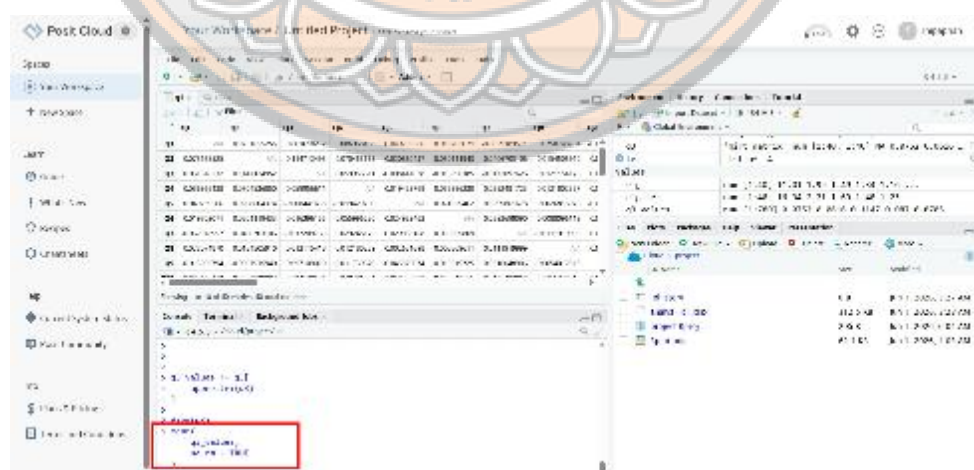
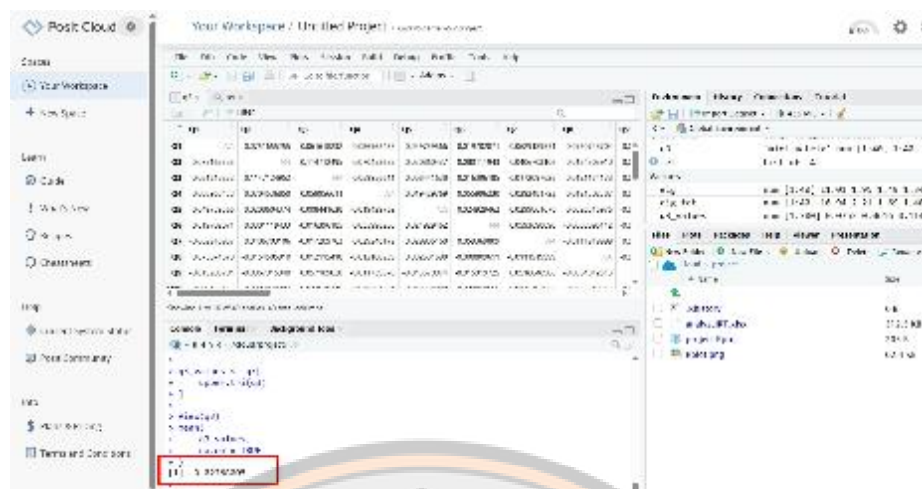


Figure 45 shows the RStudio interface. The console displays the command `mean(q3_values, na.rm = TRUE)` and its output `11.57143`. The output is highlighted with a red box.

ภาพ ร.45 คำสั่งหาค่า Mean



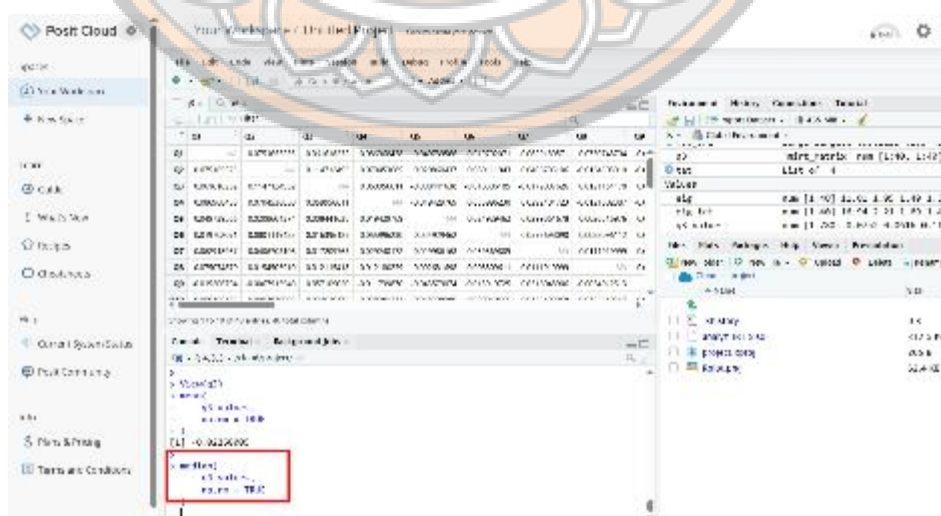
ภาพ ร.46 ผลลัพธ์ที่ได้จากคำสั่งหาค่า Mean

ผลลัพธ์ที่ได้ เท่ากับ 0.2268005

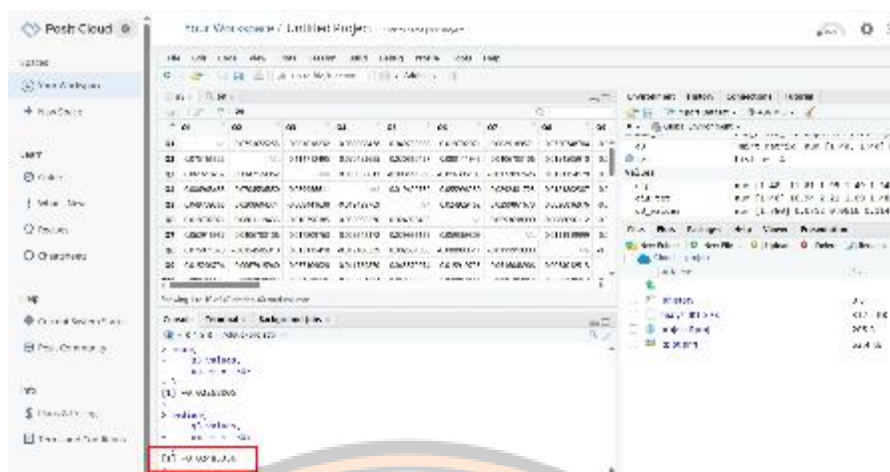
12.9 คำสั่งหาค่า Median

ที่ Console พิมพ์คำสั่ง

```
median(
  q3_values,
  na.rm = TRUE
)
```



ภาพ ร.47 คำสั่งหาค่า Median



ภาพ รูป.48 ผลลัพธ์ที่ได้จากคำสั่งหาค่า Median

ผลลัพธ์ที่ได้ เท่ากับ 0.2483356

12.10 คำสั่งหาค่า Quartile ของ Q3

ที่ Console พิมพ์คำสั่ง

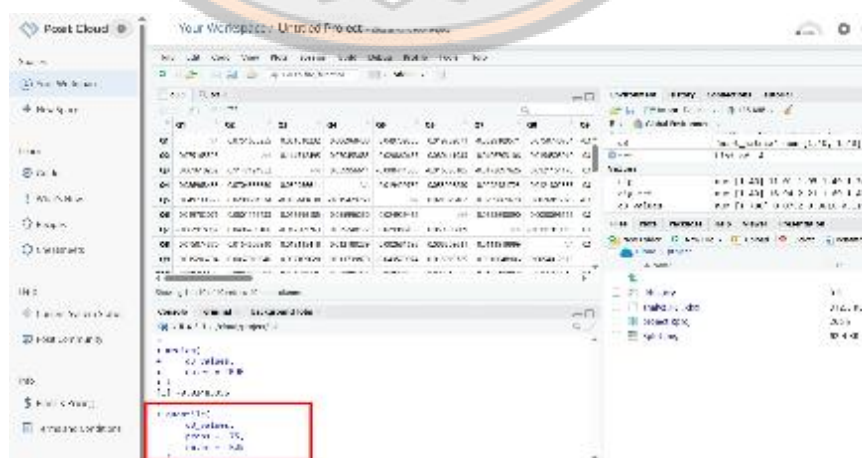
quantile(

q3_values,

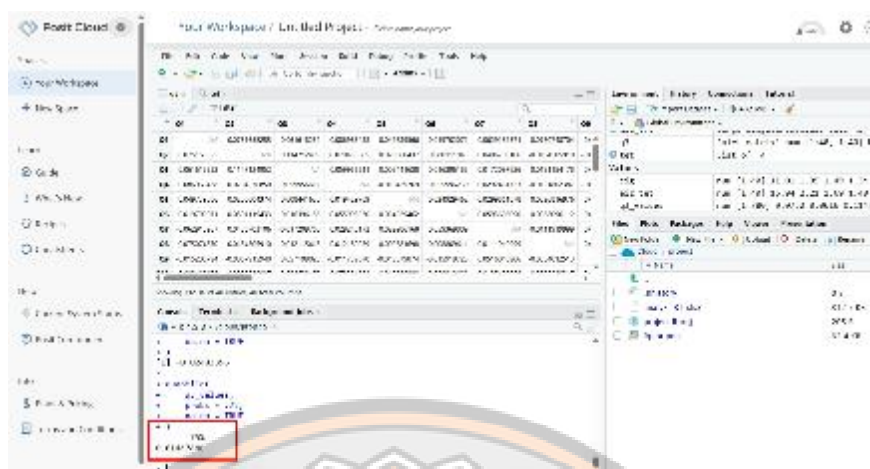
probs = .75,

na.rm = TRUE

)



ภาพ รูป.49 คำสั่งหาค่า Quartile ของ Q3



ภาพ รูป.50 ผลลัพธ์ที่ได้จากคำสั่งหาค่า Quartile ของ Q3

ผลลัพธ์ที่ได้ เท่ากับ 75% 0.014

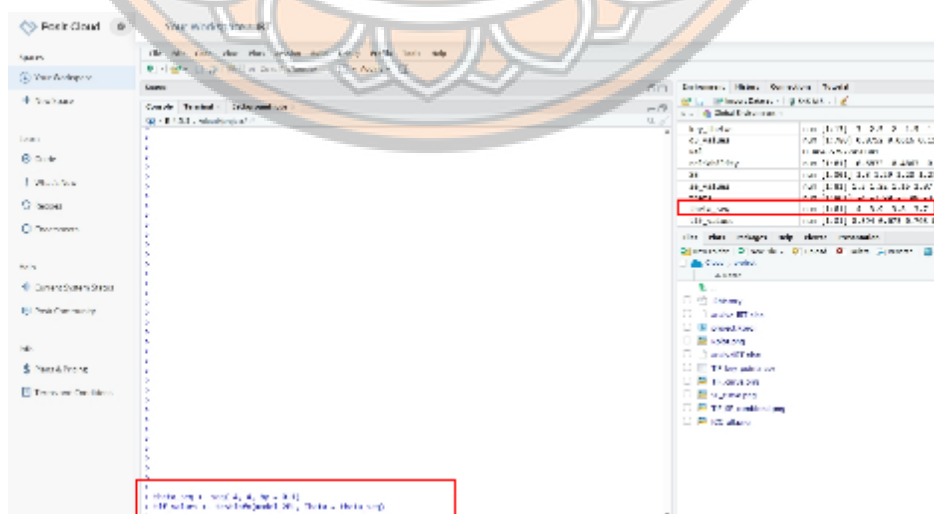
13. คำสั่งหา TIF (Test Information Function)

13.1 คำนวณค่า TIF ที่ระดับความสามารถ (theta) ตั้งแต่ -4 ถึง +4:

ที่ Console พิมพ์คำสั่ง

```
theta_seq <- seq(-4, 4, by = 0.1)
```

```
tif_values <- testinfo(model_2PL, Theta = theta_seq)
```



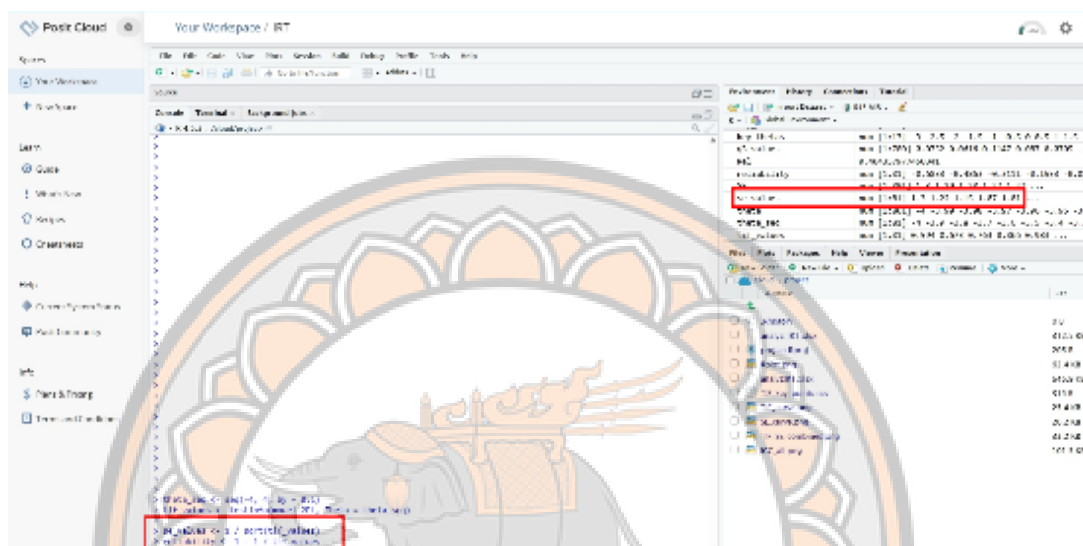
ภาพ รูป.51 คำสั่งคำนวณค่า TIF ที่ระดับความสามารถ (theta) ตั้งแต่ -4 ถึง +4 และผลลัพธ์ที่ได้

13.2 คำนวณค่า SE และ Reliability

ที่ Console พิมพ์คำสั่ง

```
se_values <- 1 / sqrt(tif_values)
```

```
reliability <- 1 - 1 / tif_values
```



ภาพ ฐ.52 คำนวณค่า SE และ Reliability และผลลัพธ์ที่ได้

13.3 แสดงค่า TIF ที่ theta = -1.0 (จุดที่แบบวัดมีความแม่นยำสูง)

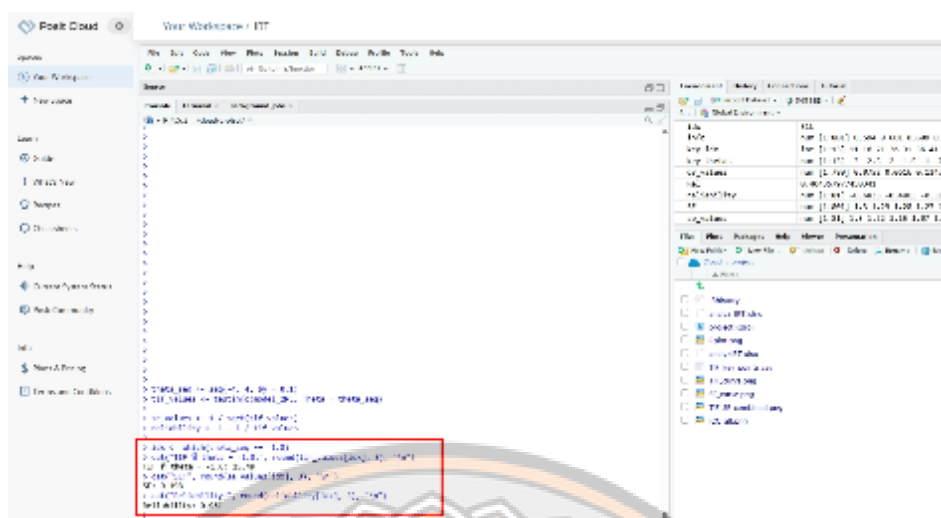
ที่ Console พิมพ์คำสั่ง

```
idx <- which(theta_seq == -1.0)
```

```
cat("TIF ที่ theta = -1.0:", round(tif_values[idx], 3), "\n")
```

```
cat("SE:", round(se_values[idx], 3), "\n")
```

```
cat("Reliability:", round(reliability[idx], 3), "\n")
```

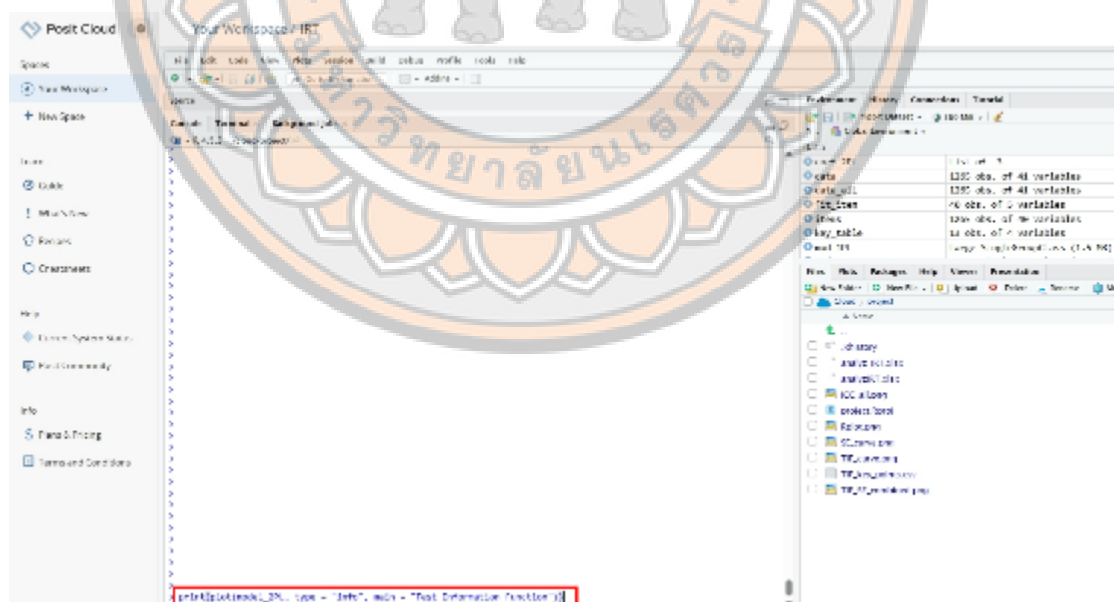


ภาพ รูป.53 คำนวณค่า SE และ Reliability และผลลัพธ์ที่ได้

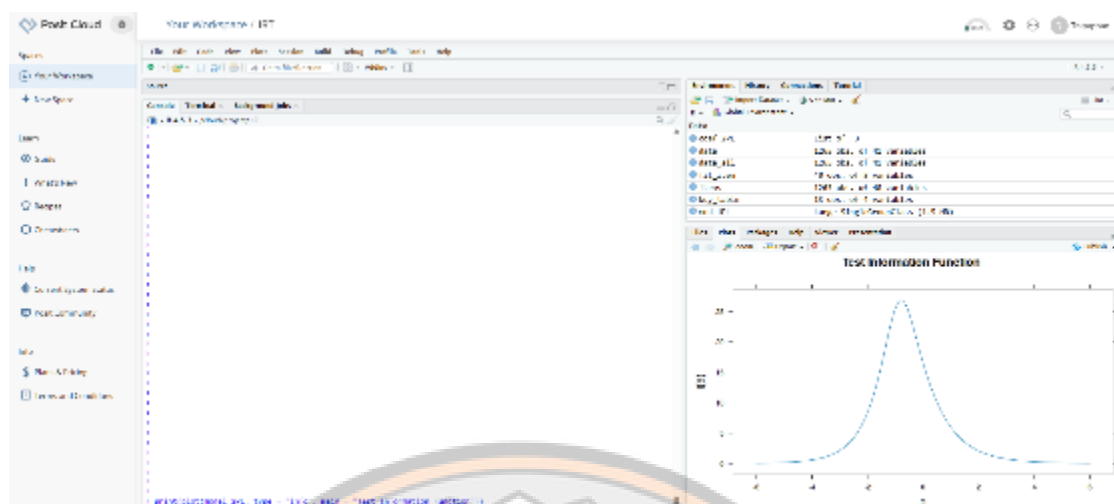
14. สร้างกราฟ Test Information Function

ที่ Console พิมพ์คำสั่ง

```
print(plot(model_2PL, type = "info", main = "Test Information Function"))
```



ภาพ รูป.54 สร้างกราฟ Test Information Function

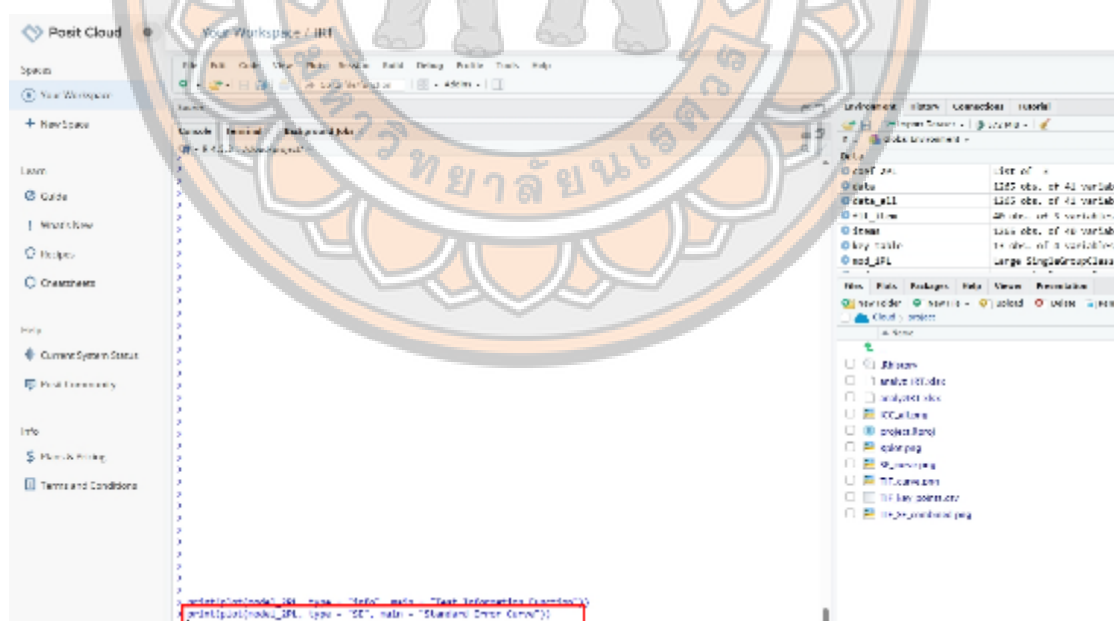


ภาพ ร.55 กราฟ Test Information Function

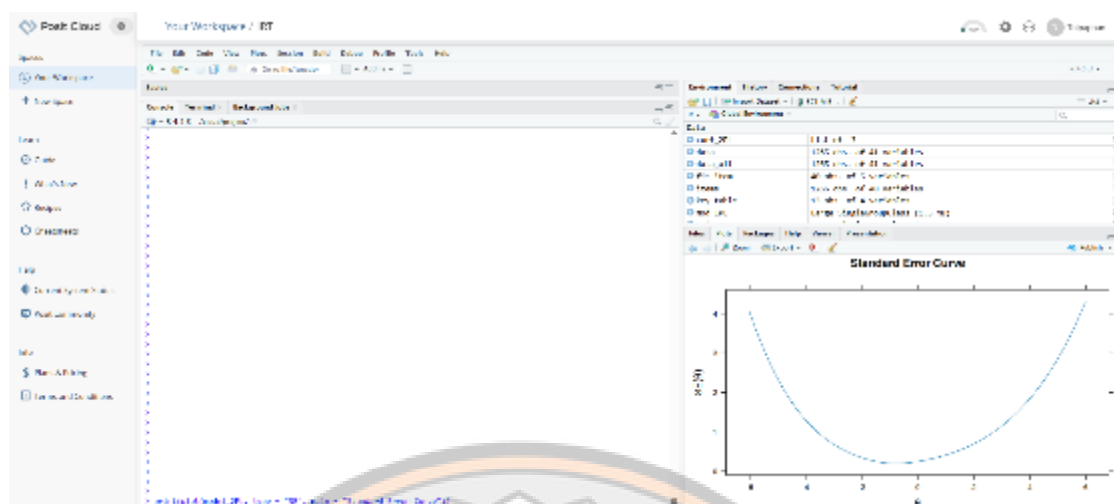
15. สร้างกราฟ Standard Error

ที่ Console พิมพ์คำสั่ง

```
print(plot(model_2PL, type = "SE", main = "Standard Error Curve"))
```



ภาพ ร.56 สร้างกราฟ Standard Error

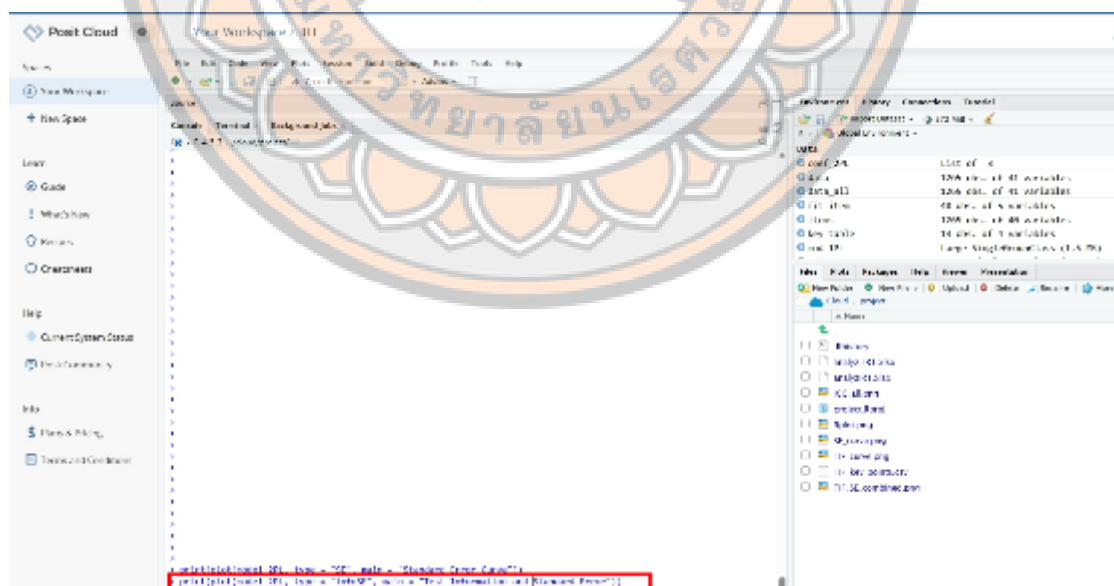


ภาพ ฐ.57 กราฟ Standard Error

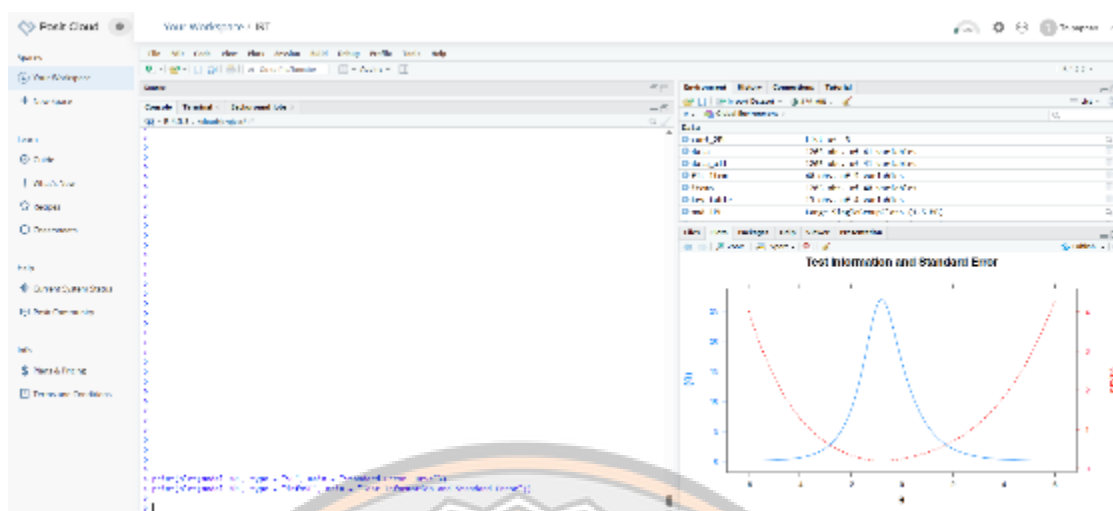
16. สร้างกราฟ TIF และ SE รวมในภาพเดียว

ที่ Console พิมพ์คำสั่ง

```
print(plot(model_2PL, type = "infoSE", main = "Test Information and  
Standard Error"))
```



ภาพ ฐ.58 สร้างกราฟ TIF และ SE รวมในภาพเดียว

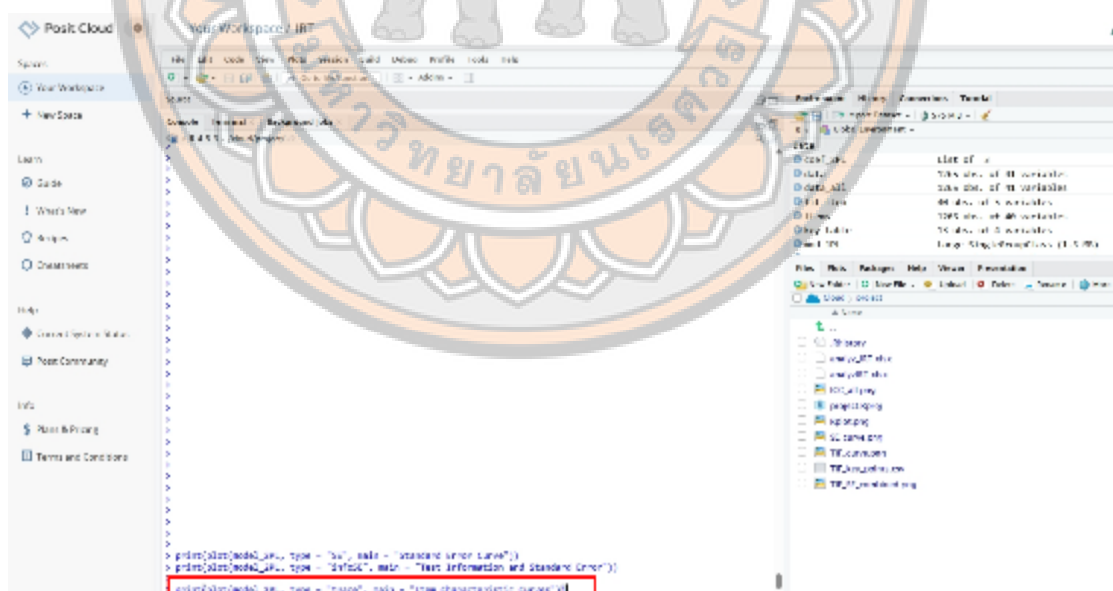


ภาพ รูป.59 กราฟ TIF และ SE รวมในภาพเดียว

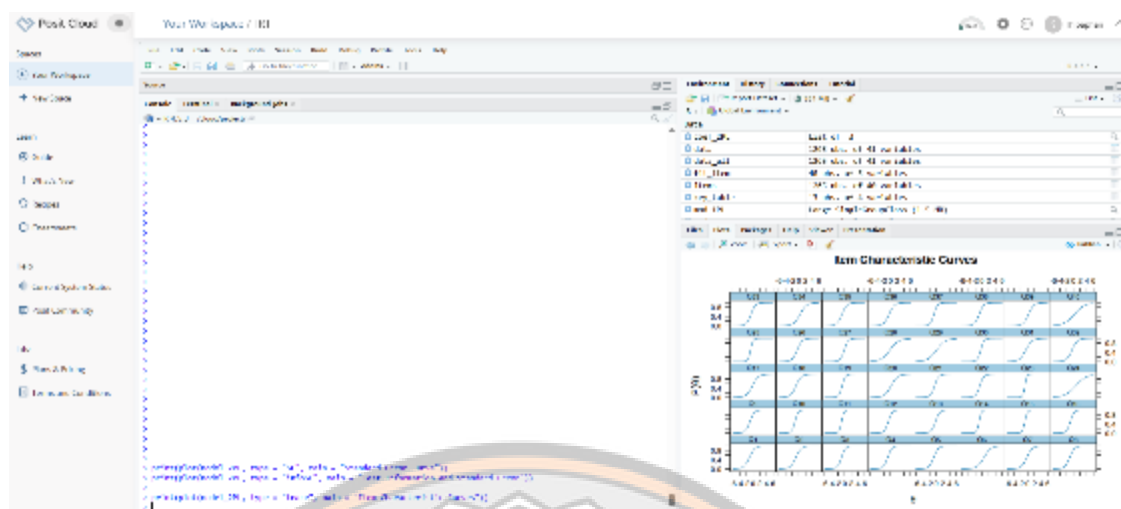
17. สร้างกราฟ Item Characteristic Curves (ICC)

ที่ Console พิมพ์คำสั่ง

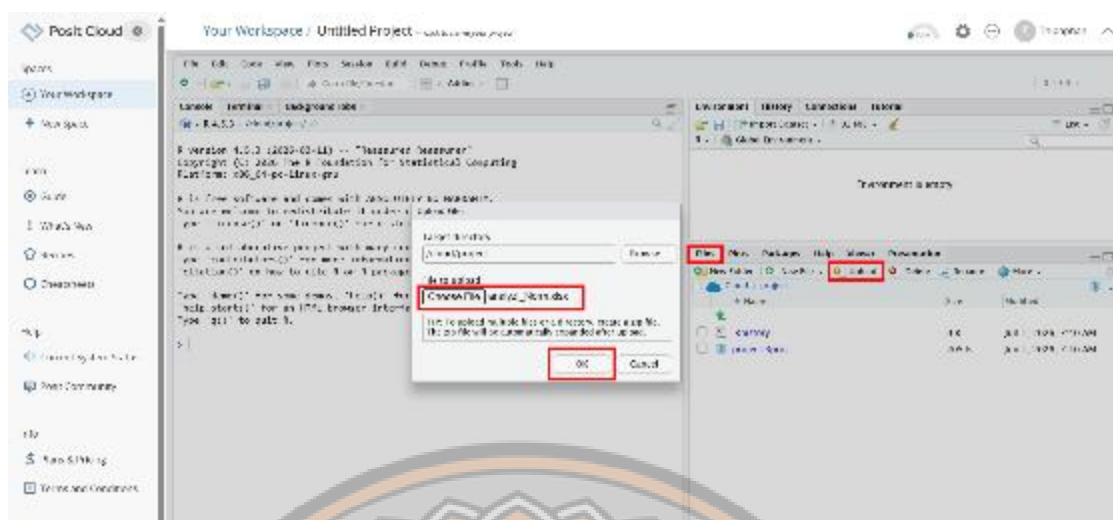
```
print(plot(model_2PL, type = "trace", main = "Item Characteristic Curves"))
```



ภาพ รูป.60 Item Characteristic Curves (ICC)



ภาพ 6.61 Item Characteristic Curves (ICC)



ภาพ จ.2 การอัปโหลดไฟล์

4. ติดตั้ง Package

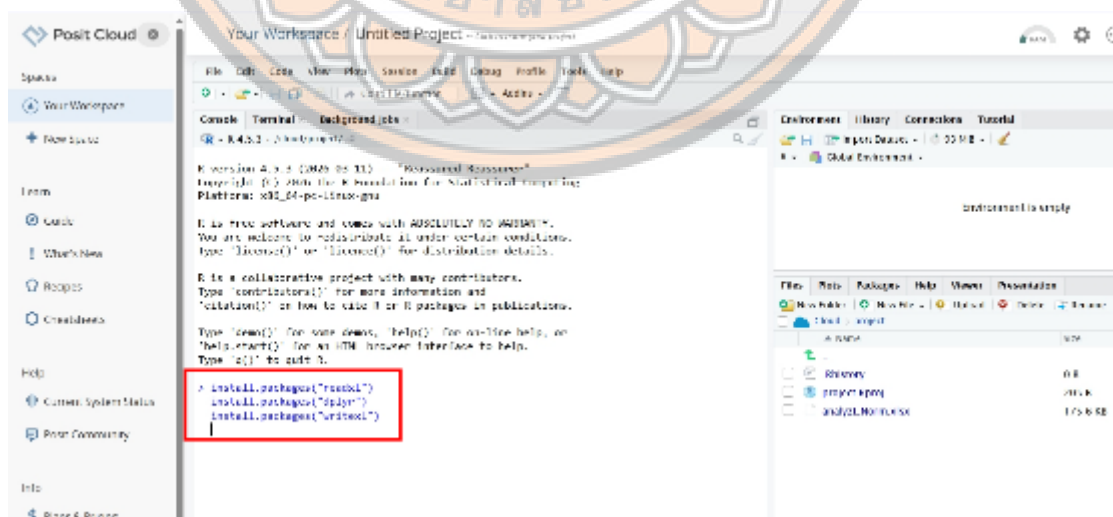
ที่ Console พิมพ์คำสั่ง

```
install.packages("readxl")
```

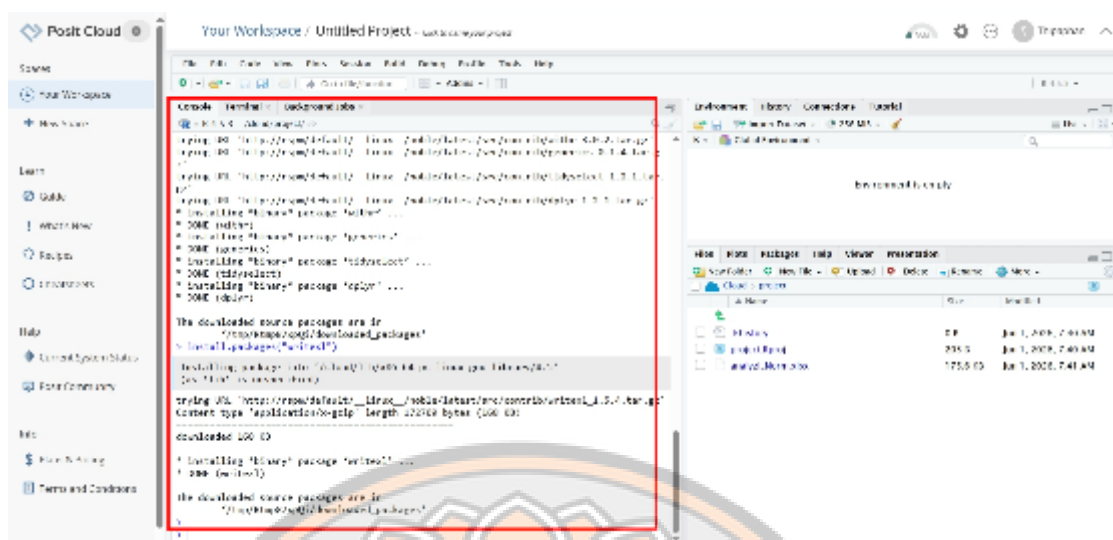
```
install.packages("dplyr")
```

```
install.packages("writexl")
```

จากนั้นกด Enter



ภาพ จ.3 การติดตั้ง Package



ภาพ ท.4 Package ที่จำเป็นเตรียมพร้อมสำหรับใช้งาน

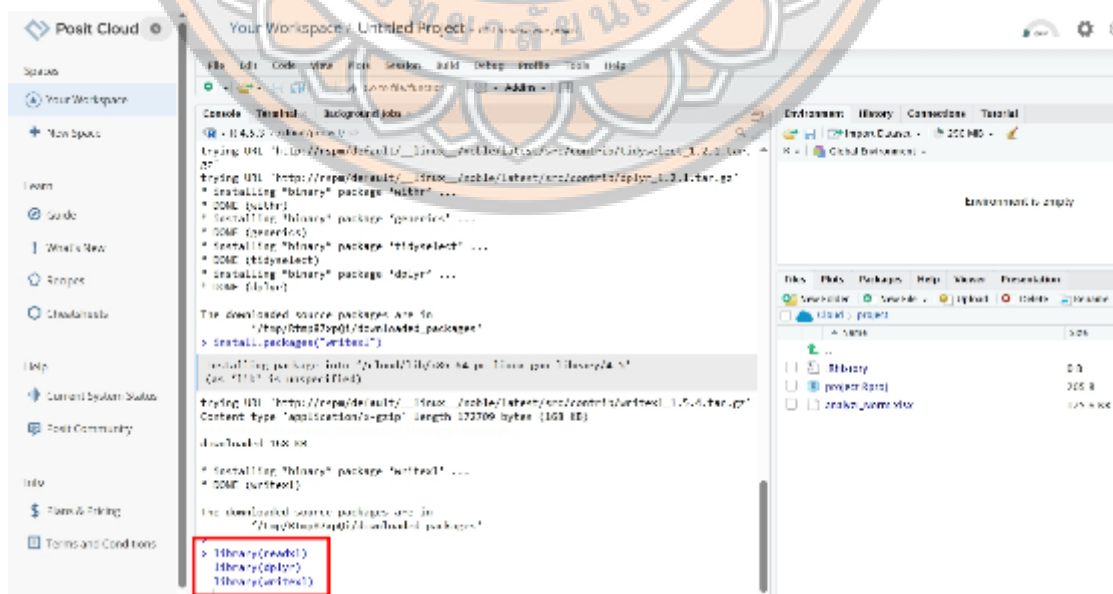
5. เรียกใช้ Package

ที่ Console พิมพ์คำสั่ง

`library(readxl)`

`library(dplyr)`

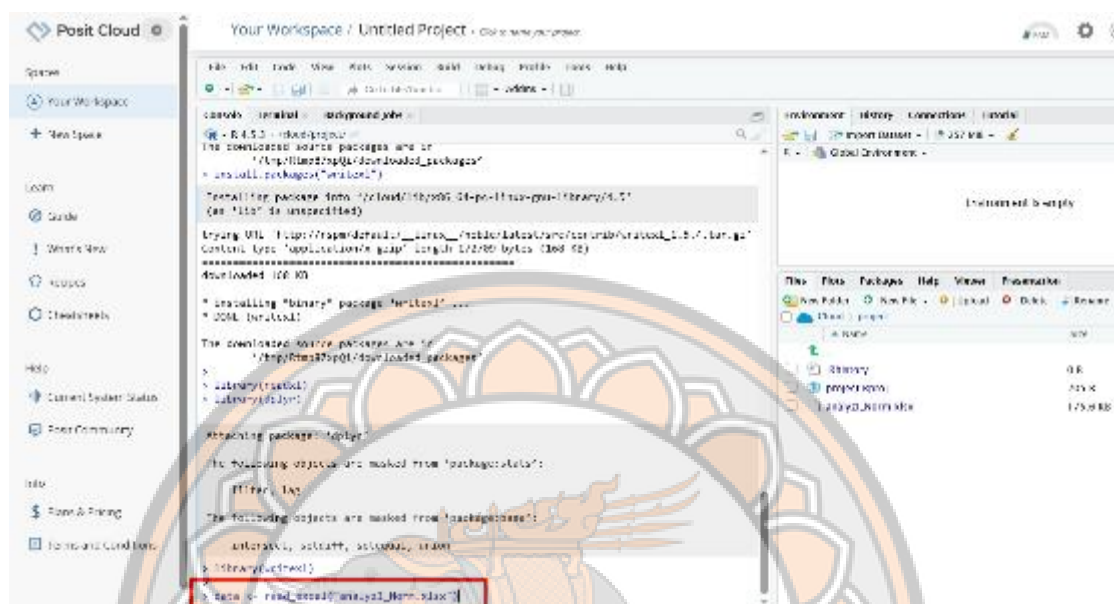
`library(writextl)`



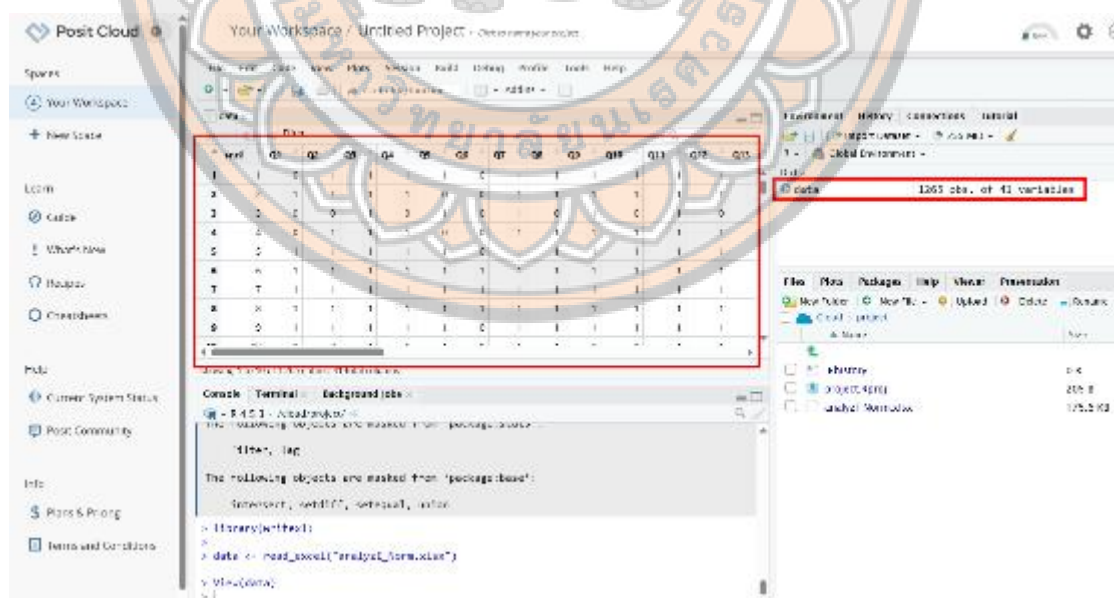
ภาพ ท.5 การเรียกใช้ Package

6. อ่านไฟล์ข้อมูล

ที่ Console พิมพ์คำสั่ง `data <- read_excel("analyzl_Norm.xlsx")` จากนั้นกด Enter



ภาพ ท.6 การอ่านข้อมูลจาก Excel



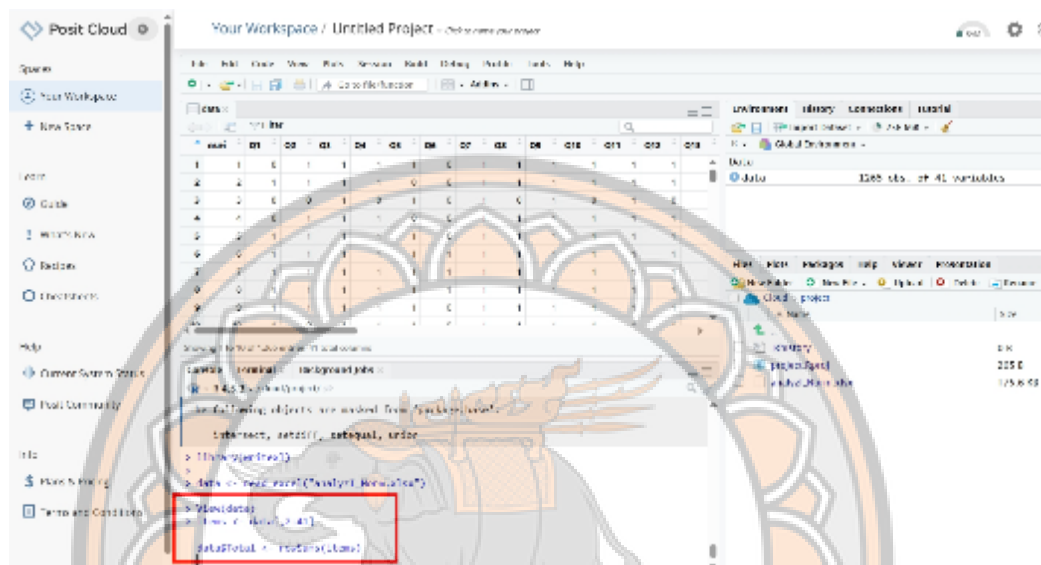
ภาพ ท.7 โปรแกรมจะสร้าง data เก็บไว้ เมื่อคลิกที่ data จะแสดงข้อมูลใน Excel

7. คำนวณคะแนนรวม

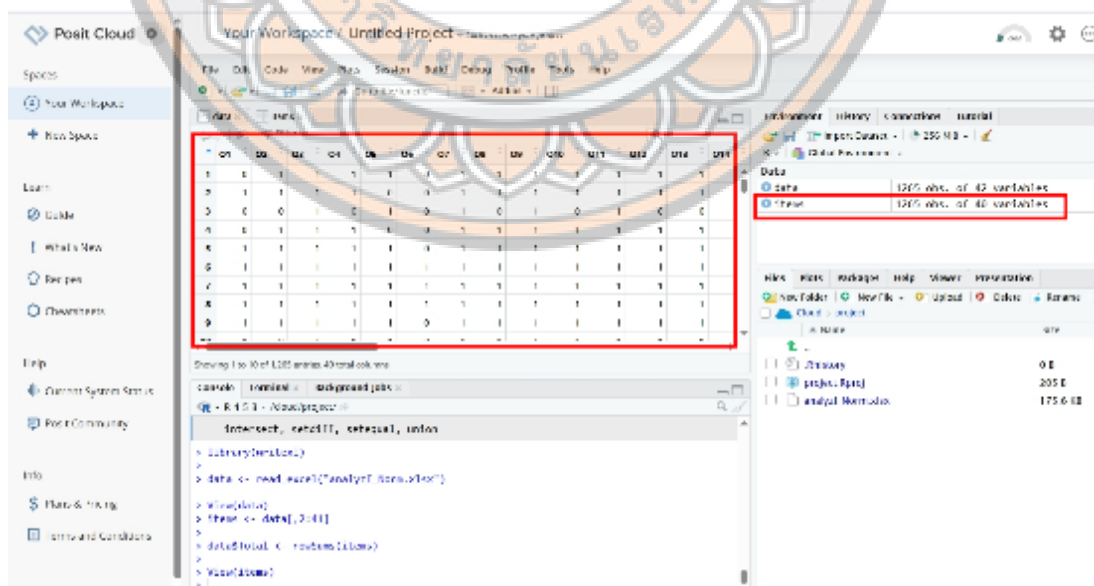
ที่ Console พิมพ์คำสั่ง

```
items <- data[,2:41]
```

```
data$Total <- rowSums(items)
```



ภาพ ๗.8 คำสั่งคำนวณคะแนนรวม



ภาพ ๗.9 โปรแกรมจะสร้าง Item เก็บไว้ที่ Data เมื่อคลิกที่ Item จะแสดงข้อมูลใน Excel

8. การหาค่าเฉลี่ยและ SD

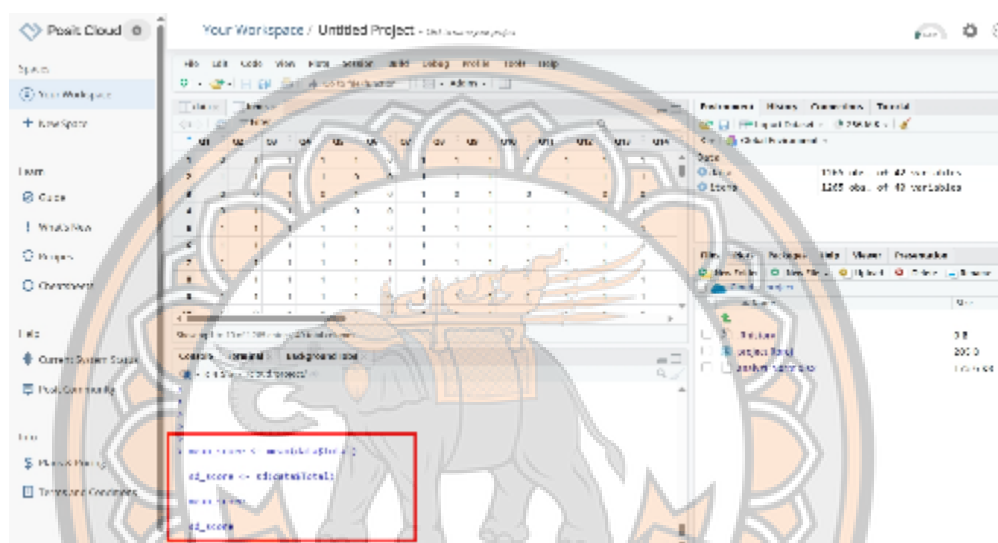
ที่ Console พิมพ์คำสั่ง

```
mean_score <- mean(data$Total)
```

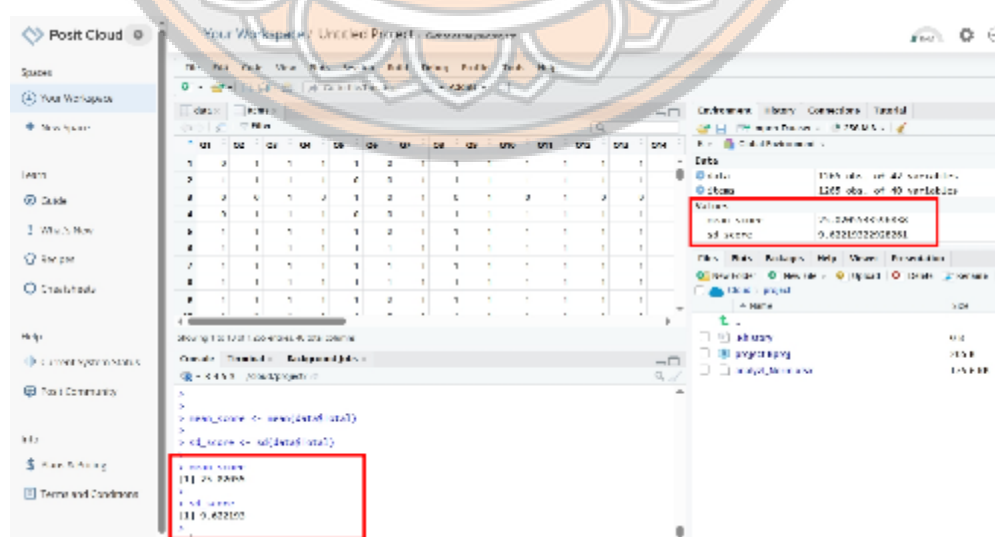
```
sd_score <- sd(data$Total)
```

```
mean_score
```

```
sd_score
```



ภาพ ท.10 คำสั่งหาค่าเฉลี่ยและ SD

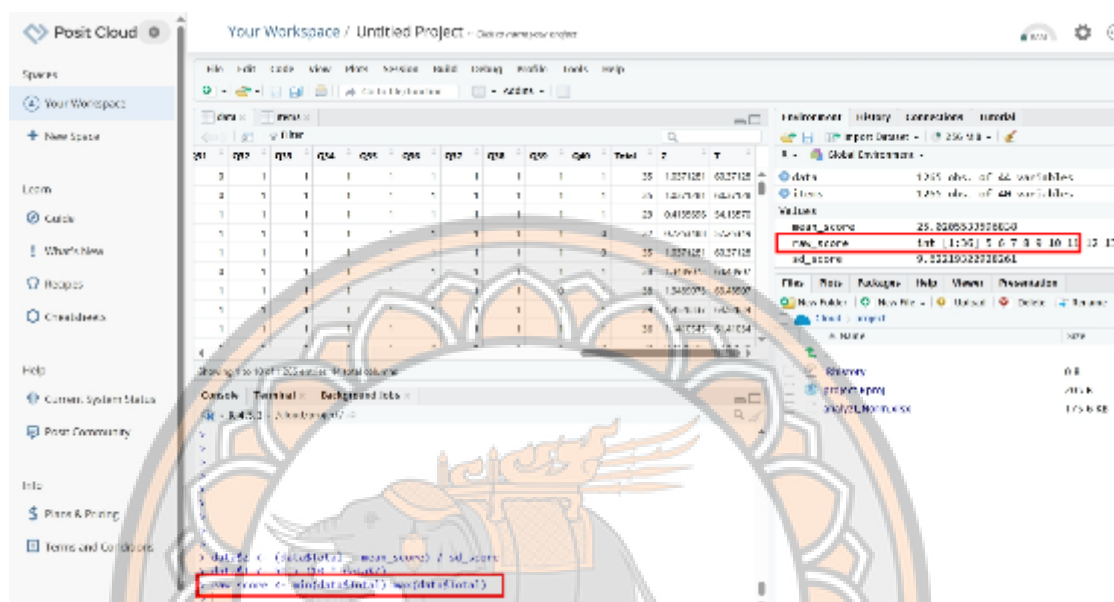


ภาพ ท.11 แสดงผลค่าเฉลี่ยและ SD

11. สร้างตาราง Norm

11.1 สร้างคะแนนดิบทุกคะแนน

ที่ Console พิมพ์คำสั่ง `raw_score <- min(data$Total):max(data$Total)`



ภาพ ข.14 สร้างคะแนนดิบทุกคะแนน

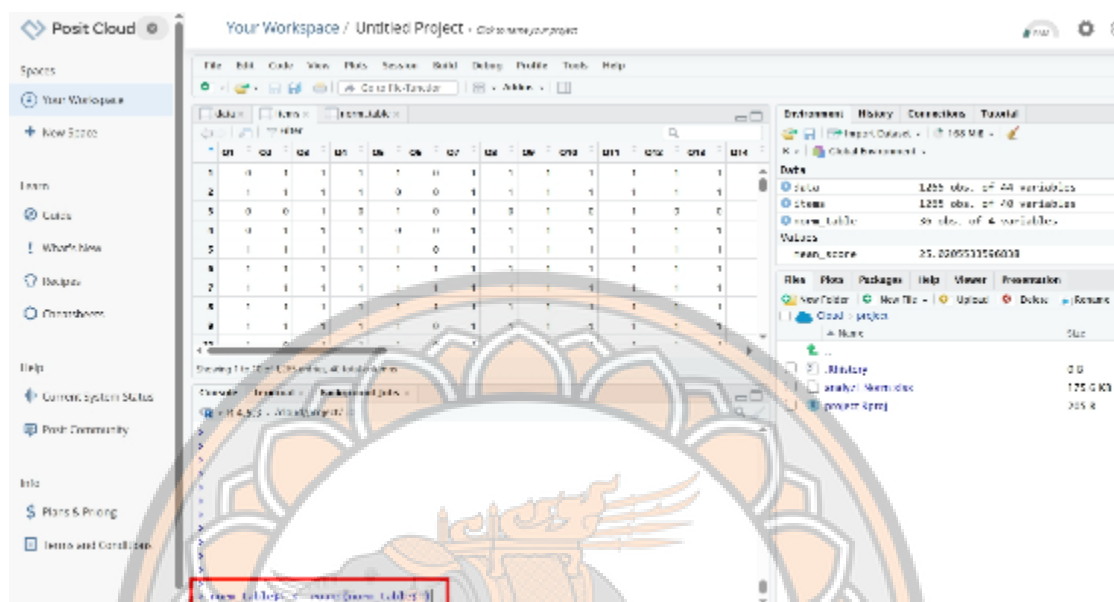
11.2 สร้างตาราง Norm

ที่ Console พิมพ์คำสั่ง

```
norm_table <- data.frame(
  Raw = raw_score
)
norm_table$Z <- (norm_table$Raw - mean_score)/sd_score
norm_table$T <- 50 + 10*norm_table$Z
```


12. ปัดเศษ T-score

ที่ Console พิมพ์คำสั่ง `norm_table$T <- round(norm_table$T)`

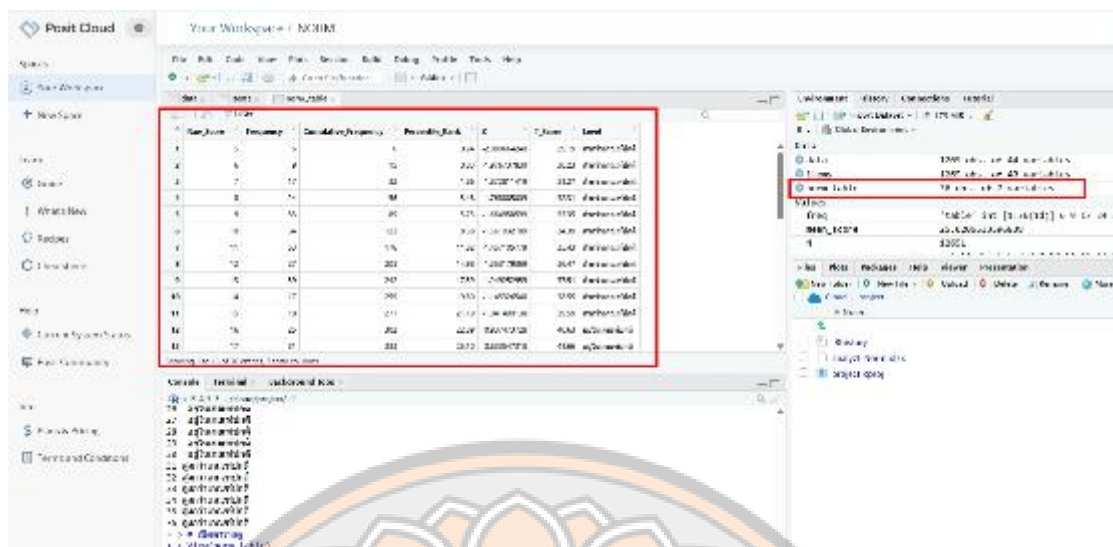


ภาพ ท.17 ปัดเศษ T-score

13. แพลตฟอร์มทักษะ (3 ระดับ)

ที่ Console พิมพ์คำสั่ง

```
norm_table$Level <- cut(
  norm_table$T,
  breaks = c(-Inf, 39, 59, Inf),
  labels = c(
    "ต่ำกว่าเกณฑ์ปกติ",
    "อยู่ในเกณฑ์ปกติ",
    "สูงกว่าเกณฑ์ปกติ"
  )
)
```



Rank	Frequency	Cumulative Frequency	Percentile Rank	Z	T-Score	Level
1	5	5	5	0.00000000	0.00	อันดับที่ 1
2	6	11	11	0.00000000	0.00	อันดับที่ 2
3	7	18	18	0.00000000	0.00	อันดับที่ 3
4	8	26	26	0.00000000	0.00	อันดับที่ 4
5	9	35	35	0.00000000	0.00	อันดับที่ 5
6	10	45	45	0.00000000	0.00	อันดับที่ 6
7	11	56	56	0.00000000	0.00	อันดับที่ 7
8	12	68	68	0.00000000	0.00	อันดับที่ 8
9	13	81	81	0.00000000	0.00	อันดับที่ 9
10	14	95	95	0.00000000	0.00	อันดับที่ 10
11	15	110	110	0.00000000	0.00	อันดับที่ 11
12	16	126	126	0.00000000	0.00	อันดับที่ 12
13	17	143	143	0.00000000	0.00	อันดับที่ 13
14	18	161	161	0.00000000	0.00	อันดับที่ 14
15	19	180	180	0.00000000	0.00	อันดับที่ 15
16	20	200	200	0.00000000	0.00	อันดับที่ 16
17	21	221	221	0.00000000	0.00	อันดับที่ 17
18	22	243	243	0.00000000	0.00	อันดับที่ 18
19	23	266	266	0.00000000	0.00	อันดับที่ 19
20	24	290	290	0.00000000	0.00	อันดับที่ 20
21	25	315	315	0.00000000	0.00	อันดับที่ 21
22	26	341	341	0.00000000	0.00	อันดับที่ 22
23	27	368	368	0.00000000	0.00	อันดับที่ 23
24	28	396	396	0.00000000	0.00	อันดับที่ 24
25	29	425	425	0.00000000	0.00	อันดับที่ 25
26	30	455	455	0.00000000	0.00	อันดับที่ 26
27	31	486	486	0.00000000	0.00	อันดับที่ 27
28	32	518	518	0.00000000	0.00	อันดับที่ 28
29	33	551	551	0.00000000	0.00	อันดับที่ 29
30	34	585	585	0.00000000	0.00	อันดับที่ 30
31	35	620	620	0.00000000	0.00	อันดับที่ 31
32	36	656	656	0.00000000	0.00	อันดับที่ 32
33	37	693	693	0.00000000	0.00	อันดับที่ 33
34	38	731	731	0.00000000	0.00	อันดับที่ 34
35	39	770	770	0.00000000	0.00	อันดับที่ 35
36	40	810	810	0.00000000	0.00	อันดับที่ 36
37	41	851	851	0.00000000	0.00	อันดับที่ 37
38	42	893	893	0.00000000	0.00	อันดับที่ 38
39	43	936	936	0.00000000	0.00	อันดับที่ 39
40	44	980	980	0.00000000	0.00	อันดับที่ 40
41	45	1025	1025	0.00000000	0.00	อันดับที่ 41
42	46	1071	1071	0.00000000	0.00	อันดับที่ 42
43	47	1118	1118	0.00000000	0.00	อันดับที่ 43
44	48	1166	1166	0.00000000	0.00	อันดับที่ 44
45	49	1215	1215	0.00000000	0.00	อันดับที่ 45
46	50	1265	1265	0.00000000	0.00	อันดับที่ 46
47	51	1316	1316	0.00000000	0.00	อันดับที่ 47
48	52	1368	1368	0.00000000	0.00	อันดับที่ 48
49	53	1421	1421	0.00000000	0.00	อันดับที่ 49
50	54	1475	1475	0.00000000	0.00	อันดับที่ 50
51	55	1530	1530	0.00000000	0.00	อันดับที่ 51
52	56	1586	1586	0.00000000	0.00	อันดับที่ 52
53	57	1643	1643	0.00000000	0.00	อันดับที่ 53
54	58	1701	1701	0.00000000	0.00	อันดับที่ 54
55	59	1760	1760	0.00000000	0.00	อันดับที่ 55
56	60	1820	1820	0.00000000	0.00	อันดับที่ 56
57	61	1881	1881	0.00000000	0.00	อันดับที่ 57
58	62	1943	1943	0.00000000	0.00	อันดับที่ 58
59	63	2006	2006	0.00000000	0.00	อันดับที่ 59
60	64	2070	2070	0.00000000	0.00	อันดับที่ 60
61	65	2135	2135	0.00000000	0.00	อันดับที่ 61
62	66	2201	2201	0.00000000	0.00	อันดับที่ 62
63	67	2268	2268	0.00000000	0.00	อันดับที่ 63
64	68	2336	2336	0.00000000	0.00	อันดับที่ 64
65	69	2405	2405	0.00000000	0.00	อันดับที่ 65
66	70	2475	2475	0.00000000	0.00	อันดับที่ 66
67	71	2546	2546	0.00000000	0.00	อันดับที่ 67
68	72	2618	2618	0.00000000	0.00	อันดับที่ 68
69	73	2691	2691	0.00000000	0.00	อันดับที่ 69
70	74	2765	2765	0.00000000	0.00	อันดับที่ 70
71	75	2840	2840	0.00000000	0.00	อันดับที่ 71
72	76	2916	2916	0.00000000	0.00	อันดับที่ 72
73	77	2993	2993	0.00000000	0.00	อันดับที่ 73
74	78	3071	3071	0.00000000	0.00	อันดับที่ 74
75	79	3150	3150	0.00000000	0.00	อันดับที่ 75
76	80	3230	3230	0.00000000	0.00	อันดับที่ 76
77	81	3311	3311	0.00000000	0.00	อันดับที่ 77
78	82	3393	3393	0.00000000	0.00	อันดับที่ 78
79	83	3476	3476	0.00000000	0.00	อันดับที่ 79
80	84	3560	3560	0.00000000	0.00	อันดับที่ 80
81	85	3645	3645	0.00000000	0.00	อันดับที่ 81
82	86	3731	3731	0.00000000	0.00	อันดับที่ 82
83	87	3818	3818	0.00000000	0.00	อันดับที่ 83
84	88	3906	3906	0.00000000	0.00	อันดับที่ 84
85	89	3995	3995	0.00000000	0.00	อันดับที่ 85
86	90	4085	4085	0.00000000	0.00	อันดับที่ 86
87	91	4176	4176	0.00000000	0.00	อันดับที่ 87
88	92	4268	4268	0.00000000	0.00	อันดับที่ 88
89	93	4361	4361	0.00000000	0.00	อันดับที่ 89
90	94	4455	4455	0.00000000	0.00	อันดับที่ 90
91	95	4550	4550	0.00000000	0.00	อันดับที่ 91
92	96	4646	4646	0.00000000	0.00	อันดับที่ 92
93	97	4743	4743	0.00000000	0.00	อันดับที่ 93
94	98	4841	4841	0.00000000	0.00	อันดับที่ 94
95	99	4940	4940	0.00000000	0.00	อันดับที่ 95
96	100	5040	5040	0.00000000	0.00	อันดับที่ 96
97	101	5141	5141	0.00000000	0.00	อันดับที่ 97
98	102	5243	5243	0.00000000	0.00	อันดับที่ 98
99	103	5346	5346	0.00000000	0.00	อันดับที่ 99
100	104	5450	5450	0.00000000	0.00	อันดับที่ 100

ภาพ ท.18 โปรแกรมสร้าง norm_table ไว้ที่ Data เมื่อคลิกจะแสดงข้อมูลทั้งหมด

