

ความเที่ยงภายใต้เงื่อนไขการทดสอบที่ต่างกันของแบบสอบอัตโนมัติวัดทักษะการคิด
วิจารณ์ญาณทางคณิตศาสตร์: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต
สาขาวิชาวิธีวิทยาการพัฒนานวัตกรรมการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2564

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Reliability of Modified Essay Question Test for Measuring Critical Thinking Skills in
Mathematics under Different Testing Conditions: An Application of Generalizability
Theory



Miss Pakjira Jongsooksai

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Education in Methodology for Innovation Development in
Education

Department of Educational Research and Psychology

FACULTY OF EDUCATION

Chulalongkorn University

Academic Year 2021

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	ความเที่ยงภายใต้เงื่อนไขการทดสอบที่ต่างกันของแบบ
	สอบอัตนัยประยุกต์วัดทักษะการคิด วิจัยรณญาณทาง
	คณิตศาสตร์:การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง
โดย	น.ส.ภัคจิรา จงสุกใส
สาขาวิชา	วิธีวิทยาการพัฒนานวัตกรรมการศึกษา
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.กมลวรรณ ตั้งธนกานนท์

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต

----- คณบดีคณะครุศาสตร์
(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

คณะกรรมการสอบวิทยานิพนธ์

----- ประธานกรรมการ
(รองศาสตราจารย์ ดร.สุวิมล กฤษศยาสา)

----- อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.กมลวรรณ ตั้งธนกานนท์)

----- กรรมการ
(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

ภคจิรา จงสุกใส : ความเที่ยงภายใต้เงื่อนไขการทดสอบที่ต่างกันของแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิด วิจัยรณญาณทางคณิตศาสตร์:การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง. (Reliability of Modified Essay Question Test for Measuring Critical Thinking Skills in Mathematics under Different Testing Conditions: An Application of Generalizability Theory) อ.ที่ปรึกษาหลัก : รศ. ดร.กมลวรรณ ตั้งธนนานนท์

การวิจัยนี้มีวัตถุประสงค์เพื่อ (1) พัฒนาแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจัยรณญาณทางคณิตศาสตร์และวิธีตรวจให้คะแนนรูปริกสองชั้นสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 (2) ตรวจสอบคุณภาพวิธีตรวจให้คะแนนรูปริกสองชั้นในแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจัยรณญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ (3) เปรียบเทียบความเที่ยงของแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจัยรณญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนที่แตกต่างกันโดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง ตัวอย่างวิจัย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 90 คน เครื่องมือวิจัย คือ แบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจัยรณญาณทางคณิตศาสตร์ และเกณฑ์การตรวจให้คะแนน

ผลการวิจัย พบว่า (1) แบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจัยรณญาณทางคณิตศาสตร์แต่ละข้อคำถามสอดคล้องกับกระบวนการคิดวิจัยรณญาณทางคณิตศาสตร์ และวิธีตรวจให้คะแนนรูปริกสองชั้น ชั้นที่ 1 ประเมินตามข้อรายการย่อยให้คะแนนแบบมาตราประมาณค่า 3 ระดับ ชั้นที่ 2 แปลงคะแนนชั้นที่ 1 (2) วิธีตรวจ ให้คะแนนรูปริกสองชั้นมีความตรงเชิงเนื้อหาสอดคล้องกับแนวคำตอบ ความเที่ยงในการตรวจให้คะแนนภายในผู้ประเมินและความเที่ยงในการตรวจให้คะแนนระหว่างผู้ประเมินมีค่าสูง และ (3) ผลการเปรียบเทียบความเที่ยงของแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจัยรณญาณทางคณิตศาสตร์ด้วยค่าสัมประสิทธิ์การสรุปอ้างอิง สำหรับทุกวิธีตรวจให้คะแนนแบบสอบที่มีความเที่ยงตั้งแต่ 0.7 ขึ้นไป เมื่อพิจารณาภายใต้ผู้ประเมินจำนวน 1 คน จะใช้แบบสอบ 6 เหตุการณ์ ภายใต้ผู้ประเมินจำนวน 2 คน จะใช้แบบสอบ 5 เหตุการณ์ พบว่า วิธีตรวจให้คะแนนรูปริกสองชั้นให้ค่าความเที่ยงสูงสุด รองลงมา คือ วิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ตามลำดับ และผลการเปรียบเทียบความแตกต่างความเที่ยงของแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจัยรณญาณทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนนรูปริกสองชั้น โดยเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox พบว่า วิธีตรวจให้คะแนนรูปริกสองชั้นมีความเที่ยงสูงกว่าวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox อย่างมีนัยสำคัญที่ระดับ 0.1 ตามลำดับ

สาขาวิชา	วิธีวิทยาการพัฒนานวัตกรรมทางการศึกษา	ลายมือชื่อ นิสิต
ปีการศึกษา	2564	ลายมือชื่อ อ.ที่ปรึกษาหลัก

6380123827 : MAJOR METHODOLOGY FOR INNOVATION DEVELOPMENT IN EDUCATION

KEYWORD:

Pakjira Jongsooksai : Reliability of Modified Essay Question Test for Measuring Critical Thinking Skills in Mathematics under Different Testing Conditions: An Application of Generalizability Theory. Advisor: Assoc. Prof. Dr. KAMONWAN TANGDHANAKANOND

This thesis was aimed at (1) developing a modified essay question test for measuring mathematical critical thinking skills and double layer scoring rubric for ninth grade students, (2) examining the quality of the double layer scoring rubric for ninth grade students, and (3) comparing the reliability of the modified essay question for measuring mathematical critical thinking skills with different scoring methods by an application of generalizability theory. Samples comprised of 90 ninth grade students. Research instruments included the essay question test for mathematical critical thinking skills and three types of scoring rubrics.

Result were as follows: (1) Each question of the modified essay questions test was in accordance with the process of mathematical critical thinking. The first step of the double -layer scoring rubric concentrated on a detailed assessment of validity, an understanding of what was being measured, and communication in mathematics. Scores obtained from the first step were converted into those of the second step according to the assigned criteria. (2) The double layer scoring rubric yielded content validity, which conformed to the answering guideline and reliability of scoring of an individual assessor as well as between assessors was at a high level. (3) Comparison of reliability of the modified essay question test for measuring mathematical critical thinking skills conducted under the G-coefficient for test questions with validity of 0.7 or above, the comparison, done by an assessor using 6-situation testing and 2 assessors using 5-situation testing, showed that the double layer scoring rubric was of the highest precision, followed by the analytic method and Knox scoring method, respectively. The difference in the precision of the modified essay question test for measuring mathematical critical thinking skills, in which the double layer scoring rubric was significantly higher precision than the analytic method and Knox scoring method with value of 0.1, respectively.

Field of Study:	Methodology for Innovation Development in Education	Student's Signature
Academic Year:	2021	Advisor's Signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความกรุณาและเมตตาของรองศาสตราจารย์ ดร.กมลวรรณ ตังธนากานนท์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่กรุณาให้เวลาและคำปรึกษาที่เป็นประโยชน์อย่างยิ่งแก่ผู้วิจัย ช่วยประสานเหตุการณ์ดำเนินงาน ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ ทั้งยังให้กำลังใจแก่ผู้วิจัยเป็นอย่างดีมาโดยตลอด ผู้วิจัยรู้สึกซาบซึ้งและขอกราบขอบพระคุณท่านอาจารย์เป็นอย่างสูงมา ณ โอกาสนี้

ขอกราบขอบพระคุณรองศาสตราจารย์ ดร. สุวิมล กฤษศยาสา ประธานคณะกรรมการสอบวิทยานิพนธ์ และรองศาสตราจารย์ ดร. ศิริเดช สุชีวะ คณะกรรมการสอบวิทยานิพนธ์ที่กรุณาให้คำชี้แนะแนวทางในการแก้ไขข้อบกพร่องต่าง ๆ ของงานวิจัย ทำให้งานวิจัยสมบูรณ์มากยิ่งขึ้น

ขอกราบขอบพระคุณรองศาสตราจารย์ ดร.ณัฐภรณ์ หลาวทอง และอาจารย์ ดร.ณภัทร ชัยมงคล คณาจารย์ภาควิชาและจิตวิทยาการศึกษาที่ประสิทธิ์ประสาทวิชาความรู้อันมีคุณค่าแก่ผู้วิจัย และสละเวลาให้คำปรึกษาในการวิจัยครั้งนี้

ขอกราบขอบพระคุณผู้ช่วยศาสตราจารย์ ดร.จنگล ทำสวน นายมานัส ทิพย์สัมฤทธิ์กุล และนางสาวชนิสรา สงวนไ้ว คณะผู้ทรงคุณวุฒิที่เสียสละเวลาเพื่อตรวจสอบคุณภาพเครื่องมือและให้ข้อเสนอแนะที่เป็นประโยชน์แก่ผู้วิจัยในครั้งนี้

ขอกราบขอบพระคุณนายอารี แก้วปาน ที่สละเวลาให้คำปรึกษา เรียบเรียงและตรวจสอบงานวิจัยภาคภาษาอังกฤษและข้อเสนอแนะที่เป็นประโยชน์แก่ผู้วิจัยในครั้งนี้

ขอกราบขอบพระคุณผู้อำนวยการ คณาจารย์โรงเรียนนรรณสูตศึกษาลัย โรงเรียนสงวนหญิงโรงเรียนกาญจนาภิเษกวิทยาลัยสุพรรณบุรี โรงเรียนบางลี่วิทยาและโรงเรียนสระยายโสมวิทยาทุกท่าน และขอขอบคุณนักเรียนทุกคนที่เป็นกลุ่มตัวอย่างซึ่งให้ความร่วมมือกับผู้วิจัยในการเก็บรวบรวมข้อมูลในครั้งนี้เป็นอย่างดี

ขอขอบพระคุณทุกท่านสำหรับความรัก ความห่วงใย และกำลังใจที่มีให้แก่ผู้วิจัย และมีส่วนช่วยให้วิทยานิพนธ์เล่มนี้สำเร็จลุล่วงไปได้ด้วยดี โดยเฉพาะครอบครัวจงสุกใส และครอบครัวสุคติ ที่ส่งเสริมสนับสนุนและคอยให้กำลังใจแก่ผู้วิจัยตลอดมา

ภัคจิรา จงสุกใส

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ค
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ	ฉ
สารบัญตาราง.....	ฎ
สารบัญภาพ.....	ฏ
บทที่ 1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
คำถามวิจัย.....	5
วัตถุประสงค์ของการวิจัย	5
สมมุติฐานการวิจัย	6
ขอบเขตของการวิจัย.....	7
คำจำกัดความที่ใช้ในการวิจัย	7
ประโยชน์ที่ได้รับจากการวิจัย.....	11
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	12
ตอนที่ 1 ทักษะการคิดวิจารณ์.....	12
1.1 ทักษะการคิดวิจารณ์ตามหลักสูตรฐานสมรรถนะ.....	12
1.2 ขั้นตอนการคิดวิจารณ์.....	14
1.3 งานวิจัยที่เกี่ยวข้องกับทักษะการคิดวิจารณ์.....	19
ตอนที่ 2 แบบสอบอัตนัยประยุกต์.....	21
2.1 ความหมายของแบบสอบอัตนัยประยุกต์.....	21

2.2 ลักษณะของแบบสอบอัตนัยประยุกต์.....	22
2.3 ขั้นตอนการสร้างแบบสอบอัตนัยประยุกต์.....	25
2.4 ขั้นตอนการดำเนินการจัดสอบแบบสอบอัตนัยประยุกต์.....	26
ตอนที่ 3 คุณภาพของการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์.....	27
3.1 คุณภาพของแบบสอบอัตนัยประยุกต์.....	27
3.2 วิธีการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์.....	30
3.3 จำนวนเหตุการณ์ของแบบสอบอัตนัยประยุกต์.....	35
3.4 งานวิจัยที่เกี่ยวข้องกับแบบสอบอัตนัยประยุกต์.....	37
ตอนที่ 4 ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หรือ G – Theory.....	40
4.1 ความเป็นมาของ G – Theory	40
4.2 แนวคิดพื้นฐานของ G – Theory.....	41
4.3 การออกแบบฟาเซท	46
4.4 งานวิจัยที่เกี่ยวข้องกับการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด.....	54
ตอนที่ 5 กรอบแนวคิดในการวิจัย.....	56
บทที่ 3 วิธีดำเนินการวิจัย	57
ขั้นตอนที่ 1 การสร้างและตรวจสอบคุณภาพของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์.....	58
1.1 กำหนดขอบเขตเนื้อหาวิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น.....	58
1.2 การเลือกเหตุการณ์ที่จะนำมาสร้างแบบสอบและข้อคำถาม	59
1.3 การกำหนดรูปแบบของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทาง คณิตศาสตร์.....	61

1.4 การสร้างแนวคำตอบ เกณฑ์การให้คะแนน การตรวจให้คะแนนและการบันทึกคะแนน	64
1.5 การหาความตรงเชิงเนื้อหา	66
1.6 การคัดเลือกและฝึกอบรมผู้ประเมิน	80
1.7 การทดลองใช้แบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์	81
ขั้นตอนที่ 2 การเก็บรวบรวมข้อมูล	86
2.1 การกำหนดประชากรและตัวอย่างวิจัย	86
2.2 การสุ่มตัวอย่างวิจัย	87
2.3 เครื่องมือที่ใช้ในงานวิจัย	88
2.4 การดำเนินการเก็บข้อมูลจากการออกแบบการวัดโดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง	89
ขั้นตอนที่ 3 การวิเคราะห์ข้อมูล	91
3.1 การวิเคราะห์ค่าสถิติพื้นฐานของผลการตรวจให้คะแนนแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์	91
3.2 การวิเคราะห์คุณภาพของการตรวจให้คะแนนด้วยวิธีตรวจให้คะแนนรูปริกสองชั้นในแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์	91
3.3 การวิเคราะห์ความเที่ยงของแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกัน ภายใต้จำนวนเหตุการณ์จำนวน ข้อคำถาม และจำนวนผู้ประเมินที่ต่างกัน โดยใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (G-Theory)	92
บทที่ 4 ผลการวิเคราะห์ข้อมูล	94
ตอนที่ 1 ผลการพัฒนาแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ และวิธีตรวจให้คะแนนแบบรูปริกสองชั้น สำหรับนักเรียนชั้นมัธยมศึกษา ปีที่ 3	94

ตอนที่ 2 ผลการตรวจสอบคุณภาพของวิธีการตรวจให้คะแนนแบบรูบริกสองชั้นในแบบสอบ อัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้น มัธยมศึกษาปีที่ 3.....	106
ตอนที่ 3 ผลการเปรียบเทียบความเที่ยงของแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิด วิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีการตรวจให้ คะแนนที่แตกต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง	116
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ.....	129
สรุปผลการวิจัย	131
อภิปรายผลการวิจัย.....	136
ข้อเสนอแนะ	141
บรรณานุกรม	146
ภาคผนวก.....	155
ภาคผนวก ก รายชื่อผู้ทรงคุณวุฒิ.....	156
ภาคผนวก ข คู่มือการใช้แบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ทาง คณิตศาสตร์.....	158
ภาคผนวก ค แบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับ นักเรียนชั้นมัธยมศึกษาปีที่ 3	186
ภาคผนวก ง แนวคำตอบ วิธีตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric) .	206
ภาคผนวก จ แบบบันทึกคะแนน แบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ทาง คณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric) ...	244
ภาคผนวก ฉ วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) วิธีตรวจให้คะแนน Knox	246
ภาคผนวก ช แบบบันทึกคะแนนแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ทาง คณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) แบบบันทึก คะแนนแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ โดยใช้วิธี ตรวจให้คะแนน Knox	279

ประวัติผู้เขียน..... 282



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญตาราง

	หน้า
ตาราง 1 การสังเคราะห์กระบวนการคิดวิจารณ์ญาณ.....	18
ตาราง 2 การสังเคราะห์จำนวนเหตุการณ์ของแบบสอบอัตนัยประยุกต์.....	35
ตาราง 3 ผลการเรียนรู้และเนื้อหาวิชาคณิตศาสตร์กลุ่มสาระการเรียนรู้คณิตศาสตร์ หลักสูตร แกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 (ฉบับปรับปรุง พ.ศ. 2561).....	58
ตาราง 4 โครงสร้างของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ 61	
ตาราง 5 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์ โดยวิเคราะห์ความสอดคล้องของข้อคำถามกับขั้นของกระบวนการ คิดวิจารณ์ญาณ.....	68
ตาราง 6 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์โดยวิเคราะห์ความสอดคล้องของแนวคำตอบกับขั้นของ กระบวนการคิดวิจารณ์ญาณ.....	69
ตาราง 7 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์ โดยวิธีตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric)	71
ตาราง 8 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์ โดยวิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method).....	75
ตาราง 9 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์ โดยวิธีการตรวจให้คะแนน Knox	79
ตาราง 10 เกณฑ์การแปลความหมายของค่าความยาก (p)	82
ตาราง 11 ค่าความยาก ค่าอำนาจจำแนก และค่าความเที่ยงของแบบสอบอัตนัยประยุกต์วัด ทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์โดยใช้วิธีการตรวจให้คะแนนที่ต่างกัน จากกลุ่ม ทดลองใช้ (try out)	83

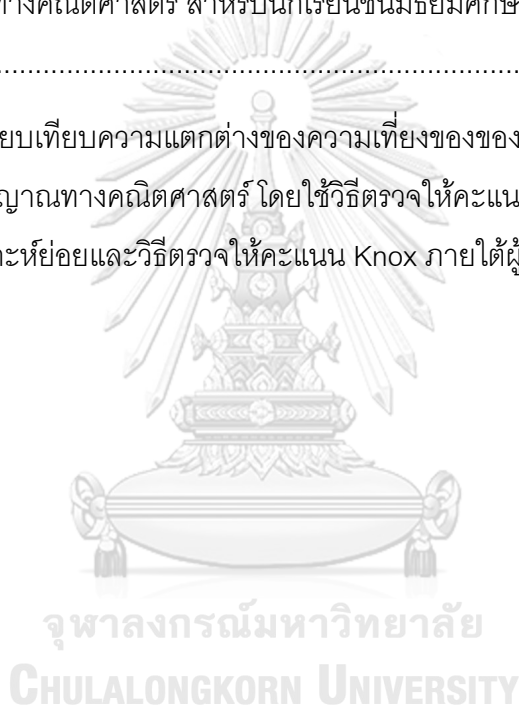
ตาราง 12 ค่าสถิติพื้นฐานของคะแนนรายชื่อของแบบสอบอัตโนมัติวัดทักษะการคิด วิจารณ์ทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนที่ต่างกัน จากผลการทดลองใช้แบบสอบ	85
ตาราง 13 จำนวนนักเรียนที่ใช้เป็นตัวอย่างวิจัยในการเก็บข้อมูล	88
ตาราง 14 เกณฑ์การแปลความหมายของค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (ICC)	92
ตาราง 15 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบอัตโนมัติวัดทักษะการคิด วิจารณ์ทางคณิตศาสตร์โดยวิธีตรวจให้คะแนนรูปrikสองชั้น (double layer scoring rubric)	108
ตาราง 16 ค่าสถิติพื้นฐานของคะแนนรายชื่อของแบบสอบอัตโนมัติวัดทักษะการคิด วิจารณ์ทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนนรูปrikสองชั้น (double layer scoring rubric) จากผู้ประเมินคนที่ 1 และผู้ประเมินคนที่ 2 ที่ตรวจให้คะแนน 2 ครั้ง	112
ตาราง 17 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ประเมินคนที่ 1 ที่ตรวจ ให้คะแนน 2 ครั้ง	114
ตาราง 18 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ประเมินคนที่ 2 ที่ตรวจ ให้คะแนน 2 ครั้ง	115
ตาราง 19 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ประเมิน 2 คน	116
ตาราง 20 ผลการวิเคราะห์ความแปรปรวน G-Study for p x r x o ของแบบสอบอัตโนมัติวัด ทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจ ให้คะแนนรูปrikสองชั้น.....	118
ตาราง 21 ผลการวิเคราะห์ความแปรปรวน G-Study for p x r x o ของแบบสอบอัตโนมัติวัด ทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจ ให้คะแนนวิเคราะห์ย่อย	119
ตาราง 22 ผลการวิเคราะห์ความแปรปรวน G-Study for p x r x o ของแบบสอบอัตโนมัติวัด ทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจ ให้คะแนน Knox.....	120

ตาราง 23 ผลการศึกษา D (D-Study of p x r x o design) ของแบบสอบถามอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนรูปกรงสองชั้น 121

ตาราง 24 ผลการศึกษา D (D-Study of p x r x o design) ของแบบสอบถามอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนวิเคราะห์ย่อย 122

ตาราง 25 ผลการศึกษา D (D-Study of p x r x o design) ของแบบสอบถามอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนน Knox..... 123

ตาราง 26 ผลการเปรียบเทียบความแตกต่างของความเที่ยงของแบบสอบถามอัตโนมัติวัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนนรูปกรงสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้ผู้ประเมิน 1 และ 2 คน..... 128



สารบัญภาพ

หน้า

ภาพที่ 1 ตัวอย่างลักษณะสถานการณ์และข้อความของแบบสอบอัตนัยประยุกต์ (มาฆพันธ์ อ๋านาคิล และกมลวรรณ ตั้งธนานนท์, 2557)	23
ภาพที่ 2 ตัวอย่างวิธีตรวจให้คะแนน Knox และวิธีตรวจให้คะแนน3 Feletti (มาฆพันธ์ อ๋านาคิล และกมลวรรณ ตั้งธนานนท์, 2557)	31
ภาพที่ 3 ตัวอย่างวิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) (ณัฐณี ศิริโชติ และกมลวรรณ ตั้งธนานนท์, 2559)	33
ภาพที่ 4 การแยกส่วนความคลาดเคลื่อนใน G-Theory	41
ภาพที่ 5 แหล่งความคลาดเคลื่อนและองค์ประกอบของความแปรปรวนสำหรับหนึ่งองค์ประกอบแบบไขว้ $p \times i$ (one-facet crossed design, $p \times i$ design)	48
ภาพที่ 6 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับสองฟาเซทแบบไขว้ $p \times r \times o$ (two-facet crossed design, $p \times r \times o$ design).....	50
ภาพที่ 7 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับสองฟาเซทแบบแฝง $p \times (i:o)$ (two-facet nested design, $p \times (i:o)$ design)	51
ภาพที่ 8 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับสามฟาเซทแบบแฝง $p \times o \times (i:o)$ (three-facet, $p \times r \times (i:o)$ design)	53
ภาพที่ 9 กรอบแนวคิดในการวิจัย	56
ภาพที่ 10 ตัวอย่างแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์.....	95
ภาพที่ 11 ตัวอย่างวิธีให้คะแนนรูปริกสองชั้น (Double layer scoring rubric) ในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์	101
ภาพที่ 12 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การรูปร่างของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมพัทธ์ (relative coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปริกสองชั้น เมื่อเทียบ	

กับวิธีตรวจให้คะแนน วิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 1 คน และจำนวนเหตุการณ์ 2 – 6 เหตุการณ์..... 124

ภาพที่ 13 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมพัทธ์ (relative Coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปrikสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนน วิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 2 คน และจำนวนเหตุการณ์ 2 – 6 เหตุการณ์..... 124

ภาพที่ 14 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปrikสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนน วิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 1 คน และจำนวนเหตุการณ์ 2 – 6 เหตุการณ์..... 125

ภาพที่ 15 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์ (absolute Coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปrikสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 2 คน และจำนวนเหตุการณ์ 2 – 6 เหตุการณ์..... 126

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

หลักสูตรฐานสมรรถนะ (competency – based curriculum) หมายถึง หลักสูตรที่ยึดทักษะการเรียนรู้ของผู้เรียนเป็นองค์ประกอบหลัก ส่วนความรู้ เจตคติและคุณลักษณะต่าง ๆ เป็นองค์ประกอบรองที่ผลักดันให้ผู้เรียนสามารถปฏิบัติงาน ๆ หนึ่งได้ประสบผลสำเร็จและเป็นประโยชน์ต่อชีวิตประจำวัน ดังนั้นการศึกษาจำเป็นต้องมีเครื่องมือที่ดีในการวัดทักษะต่าง ๆ ของผู้เรียนให้เป็นที่ไปตามเป้าหมายที่ตั้งไว้ หลักสูตรฐานสมรรถนะจึงมีความสำคัญและจำเป็นเร่งด่วนที่ต้องเข้ามามีบทบาทแทนหลักสูตรอิงมาตรฐานที่เน้นเนื้อหาจำนวนมหาศาลแต่ขาดการส่งเสริมให้ผู้เรียนเกิดทักษะการเรียนรู้ ทางสำนักงานเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ(2563) ได้วางแนวทางการจัดการเรียนรู้หลักสูตรฐานสมรรถนะไว้ 5 สมรรถนะ ประกอบด้วย สมรรถนะที่ 1 การจัดการตนเอง สมรรถนะที่ 2 การสื่อสาร สมรรถนะที่ 3 การรวมพลังทำงานเป็นทีม สมรรถนะที่ 4 การคิดขั้นสูง และสมรรถนะที่ 5 การเป็นพลเมืองที่เข้มแข็ง (สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2563)

เมื่อพิจารณาสมรรถนะการคิดขั้นสูง (Higher Order Thinking: HOT) เป็นการคิด ไตร่ตรอง คิดวิพากษ์ คิดอย่างเป็นระบบ สามารถใช้ข้อมูลในการตัดสินใจแก้ปัญหา มีความยืดหยุ่นทางความคิด ริเริ่มความคิดใหม่ ๆ โดยคำนึงถึงผลกระทบรอบด้าน ประกอบด้วย 4 ด้านย่อย ได้แก่ การคิดอย่างมีวิจารณญาณ การคิดเป็นระบบการคิดแก้ปัญหา และการคิดสร้างสรรค์ ซึ่งจะเห็นได้ว่า ในทักษะด้านการคิดอย่างมีวิจารณญาณ และทักษะในการแก้ปัญหา (critical thinking and problem solving) เป็นทักษะการเรียนรู้ที่จำเป็นของผู้เรียนในศตวรรษที่ 21 (3Rs8Cs) ตามแผนการศึกษาแห่งชาติ พ.ศ.2560-2579 เนื่องจากเป็นกระบวนการคิดทางปัญญาที่ผ่านการพิจารณา ไตร่ตรองปัญหา โดยใช้ข้อมูลหรือสมมติฐานที่ตั้งอยู่บนตรรกะอย่างมีเหตุและผล ผ่านการประเมินและตรวจสอบความถูกต้องน่าเชื่อถือของข้อมูล ก่อนที่จะลงข้อสรุปหรือข้อสมมติฐานที่ตั้งไว้ (อัมพร ม้าคะนอง, 2554; สิริภักตร์ ศิริโท, 2558)

การคิดวิจาร์ณญาณ (critical thinking) เป็นการคิดที่ต้องใช้เหตุผลในการพิจารณาอย่างรอบคอบผ่านการศึกษาค้นคว้าข้อมูล หลักฐาน จนสามารถแยกแยะข้อเท็จจริงหรือข้อคิดเห็นได้ ตลอดจนพิจารณาความน่าเชื่อถือของข้อมูล ตัดสินใจอย่างรอบคอบรอบด้าน เพื่อให้เห็นว่าสิ่งใด

ควรเชื่อหรือควรทำ และเพราะเหตุใด จนนำไปสู่แนวทางในการแก้ไขปัญหา สำหรับการคิด
 วิจารณ์ญาณในการเรียนการสอนคณิตศาสตร์มุ่งเน้นให้ผู้เรียนสามารถวิเคราะห์วิพากษ์และ
 ประเมินข้อมูลทางคณิตศาสตร์โดยเฉพาะข้อมูลที่เป็นการให้เหตุผล เพื่อตัดสินใจหาข้อสรุปทาง
 คณิตศาสตร์ซึ่งต้องตั้งอยู่บนพื้นฐานของการเชื่อมโยงความรู้ต่าง ๆ ทางคณิตศาสตร์อย่างรอบ
 ด้าน และจากการสังเคราะห์แนวคิดกระบวนการในการคิดวิจารณ์ญาณ (Woolfolk, 1995;
 Gagne, 1985; Dressel & Mayhew, 1957; Watson & Glaser, 2002) พบว่ามีนักการศึกษา
 เสนอกระบวนการคิดวิจารณ์ญาณที่เหมาะสมสำหรับนำไปใช้ในทางคณิตศาสตร์มี 5 ขั้นตอน
 ประกอบด้วย ขั้นตอนปัญหา ขั้นตอนสมมติฐาน ขั้นรวบรวมข้อมูล ขั้นวิเคราะห์ข้อมูล และขั้นสร้าง
 ข้อสรุป

โดยส่วนใหญ่เครื่องมือในการประเมินการคิดขั้นสูงนิยมใช้แบบสอบถามเรียงหรือแบบ
 สอบอัตนัย (essay question) ที่มักจะมีการให้บทความยาวมาในครั้งเดียว ไม่มีตัวเลือก ผู้สร้าง
 แบบสอบจะเขียนข้อคำถามแล้วให้ผู้สอบเขียนตอบด้วยตัวเอง แม้ว่าจะข้อสอบอัตนัยจะวัด
 ความคิดขั้นสูงแต่ก็ยังไม่บรรลุวัตถุประสงค์เนื่องจากผู้สร้างข้อสอบตั้งเป้าหมายในการตรวจให้
 คะแนนได้ง่ายเป็นสำคัญ ทำให้แบบสอบอัตนัยนี้มักวัดความรู้ที่อยู่ในระดับท่องจำหรือเข้าใจพื้นฐาน
 เท่านั้น และมีความเป็นปรนัยในการตรวจให้คะแนนค่อนข้างต่ำ จากข้อจำกัดของแบบสอบอัตนัย
 แบบดั้งเดิมจึงได้รับการพัฒนาเป็นแบบสอบอัตนัยประยุกต์ (Modified Essay Question test:
 MEQ)

แบบสอบอัตนัยประยุกต์ (MEQ) เป็นแบบสอบถามเรียงที่บรรยายเหตุการณ์ที่เกิดขึ้น
 จริงตามลำดับของเหตุการณ์นั้น ๆ แต่จะไม่นำเสนอเหตุการณ์ทั้งหมดเพียงครั้งเดียวเพื่อเป็นการ
 เพิ่มข้อมูลที่อาจจะเกี่ยวข้องหรือไม่เกี่ยวข้องกับคำถาม ผู้สอบจะต้องนำข้อมูลมาพิจารณาเอง ใน
 แต่ละข้อคำถามจะเป็นอิสระจากกันและผู้สอบไม่มีโอกาสกลับไปแก้ไขคำตอบของข้อที่ผ่านมาได้
 โดยผู้สอบจะต้องใช้ข้อมูลที่กำหนดมาให้เท่านั้นตอบคำถาม จึงทำให้แบบสอบอัตนัยประยุกต์
 แตกต่างกับแบบสอบอัตนัยแบบดั้งเดิม จึงจัดได้ว่าแบบสอบอัตนัยประยุกต์เป็นแบบสอบในการวัด
 การคิดวิจารณ์ญาณในการแต่ละข้อมูลที่ได้รับจากเหตุการณ์ต่าง ๆ ก่อนจะตั้งสมมติฐานหรือลง
 ข้อสรุปได้ จากงานวิจัยที่ผ่านมา มีเพียงแบบสอบประเภทเลือกตอบหลายตัวเลือกที่ใช้กับการวัด
 การคิดวิจารณ์ญาณ ทางคณิตศาสตร์

แม้ว่าแบบสอบอัตนัยประยุกต์จะเป็นแบบสอบที่มีคุณภาพ ใช้เวลาในการสร้างแบบสอบ
 น้อย ประหยัดค่าใช้จ่าย แต่จากการศึกษางานวิจัยที่เกี่ยวข้องกับแบบสอบอัตนัยประยุกต์ พบว่า
 แบบสอบอัตนัยประยุกต์ มีปัญหาในเรื่องจำนวนเหตุการณ์ที่ใช้ ถ้าแบบสอบมีความยาวมาก

เกินไป จะช่วยเพิ่มความเที่ยงของแบบสอบแต่ก็จะทำให้ผู้สอบเกิดความเมื่อยล้าในการตอบ ในทางกลับกันถ้าจำนวนเหตุการณ์มีน้อยเกินไป ทำให้ไม่สามารถครอบคลุมองค์ความรู้ที่ต้องการ ประเมินได้ครบและความเที่ยงของแบบสอบก็ลดลงเช่นกัน (ศิริชัย กาญจนวาสี, 2563) อย่างไรก็ตาม ผลการประเมินจะเป็นอย่างไรนั้นก็ขึ้นอยู่กับผู้ประเมิน จากงานวิจัย พบว่า เมื่อเพิ่มจำนวนผู้ประเมินจะเพิ่มความเที่ยงสูงขึ้น (Brennan et al., 1995; Sudweeks, Reeve, & Bradshaw, 2004; Carl, W. S. et al., 1999) ดังนั้นหากมีผู้ประเมินมากกว่าหนึ่งคน ก็ต้องมีการตรวจสอบ ความสอดคล้องระหว่างผลการประเมินด้วยเครื่องมือการประเมินเดียวกัน (กมลวรรณ ตังธนกานนท์, 2563) แต่ทั้งนี้ ทั้งจำนวนผู้ประเมินที่เหมาะสมและจำนวนเหตุการณ์ที่ใช้อย่างไม่สามารถสรุป ได้แน่ชัด

ประการสุดท้ายที่ต้องพิจารณาเนื่องจากแบบสอบอัตนัยประยุกต์เป็นแบบสอบประเภท ความเรียงรูปแบบหนึ่ง ดังนั้นจึงต้องพิจารณาเรื่องความเป็นปรนัยในการตรวจให้คะแนนที่เกิดจากการเลือกใช้วิธีการตรวจให้คะแนนที่เหมาะสม สำหรับวิธีการตรวจให้คะแนนแบบสอบอัตนัย ประยุกต์มี 2 วิธี คือ วิธีการตรวจให้คะแนนตามแบบ Knox มีรูปแบบการให้คะแนนเป็น 0, 1, 2, 3, ... และวิธีการตรวจให้คะแนนตามแบบ Feletti มีรูปแบบในการให้คะแนนเป็น 0, 1 นอกจากนี้ยังมี วิธีการตรวจให้คะแนนแบบสอบประเภทความเรียงที่นิยมอีก 2 วิธี คือ วิธีการให้คะแนนแบบองค์รวม (holistic method) และวิธีการให้คะแนนแบบวิเคราะห์ย่อย (analytic method) (Charles & Richard, 1990; Ebe, R. L., & Frisbie, D. A., 1972; Albert, O., 2003; Abatihun, A. S., 2013)

จากการศึกษาของงานวิจัยของณัฐธณี ศิริโชติ และกมลวรรณ ตังธนกานนท์ (2559) พบว่า วิธีการให้คะแนนแบบรูปริกสองชั้น (double layer scoring rubric) จัดเป็นเกณฑ์การให้คะแนน รูปริกแบบแยกองค์ประกอบ (analytic scoring rubric) อีกประเภทหนึ่งที่ทำให้ผลการประเมิน ละเอียดยิ่งขึ้นและง่ายต่อการประเมิน เนื่องจากอธิบายลักษณะของงานหรือการปฏิบัติงานใน แต่ละความสามารถ 2 ระดับชั้น ระดับชั้นที่ 1 เป็นเกณฑ์การให้คะแนนที่อธิบายลักษณะงานหรือ การปฏิบัติงานในข้อรายการย่อย ๆ ลดหลั่นตามระดับความสามารถโดยผสมผสานระหว่างวิธีการ ให้คะแนนรูปริกแบบแยกองค์ประกอบ (analytic scoring rubric) กับมาตราประมาณค่า (Rating scale) ส่วนระดับชั้นที่ 2 เป็นการแปลงระดับคะแนน (rescale) ที่ได้ในระดับชั้นที่ 1 ออกมาให้ เป็นไปตามช่วงคะแนนที่กำหนด (Hamzah et al., 2014) สำหรับจุดเด่นของวิธีการให้คะแนนแบบ รูปริกสองชั้น คือ เกณฑ์การให้คะแนนมีการแบ่งมิติหรือข้อรายการย่อยมากขึ้นเพิ่มความปรนัย ในการตรวจให้คะแนน และ ผลคะแนนในข้อรายการย่อยที่ได้ในแต่ละด้าน ยังทำให้ทราบถึง จุดแข็งจุดอ่อนในองค์ประกอบด้านนั้น ๆ (กมลวรรณ ตังธนกานนท์, 2563)

นอกจากวิธีตรวจให้คะแนนแบบสอบอัตนัยข้างต้น ยังพบวิธีการให้คะแนนรูบริกสองชั้น (double layer scoring rubric) ซึ่งเกณฑ์การให้คะแนนรูบริกแบบแยกองค์ประกอบอีกประเภทหนึ่งที่ทำให้ผลการประเมินละเอียดมากยิ่งขึ้นและง่ายต่อการประเมินจากการทำงานของนักวิจัยของ ณัฐฐณี ศิริโชติ และ กมลวรรณ ตังธนกานนท์ (2559) โดยในชั้นที่ 1 ประเมินตามข้อรายการย่อย ๆ โดยให้คะแนนแบบมาตราประมาณค่า (rating scale) ชั้นที่ 2 แปลงระดับคะแนน (rescale) จากชั้นที่ 1 ให้เป็นไปตามช่วงคะแนนที่กำหนด (Hamzah et al., 2014) และจากงานวิจัยของมาฆพันธ์ อ่อนาคิล และ กมลวรรณ ตังธนกานนท์ (2557) พบว่า วิธีการตรวจให้คะแนนวิเคราะห์ย่อยมีความเที่ยงมากกว่าวิธีการตรวจให้คะแนน Knox และวิธีการตรวจให้คะแนน Feletti เนื่องจากวิธีวิเคราะห์ย่อยมีความละเอียดและให้ข้อมูลย้อนกลับได้ ในขณะที่วิธีของ Knox และวิธีของ Feletti ยังไม่สามารถให้ข้อมูลย้อนกลับได้ แต่ยังไม่พบวิธีตรวจให้คะแนนแบบรูบริกสองชั้นกับแบบสอบอัตนัยประยุกต์ ทั้งที่มีขั้นตอนที่ละเอียดเหมาะแก่การนำไปใช้และสามารถใช้เป็นข้อมูลย้อนกลับให้แก่ผู้สอบได้ด้วยเช่นกัน

ในขณะที่งานวิจัยส่วนหนึ่งได้ประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory: G-Theory) ศึกษาเชิงจำแนกแหล่งความคลาดเคลื่อนของการวัดในสถานการณ์ต่าง ๆ ซึ่งขยายแนวคิดจากทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) ที่เชื่อว่าความคลาดเคลื่อนรวมมีเพียงแหล่งเดียว มิได้ให้ความสนใจต่อสถานการณ์หรือเงื่อนไขการวัด และปฏิสัมพันธ์ระหว่างผู้สอบกับเงื่อนไขของการวัดที่สามารถส่งผลต่อความน่าเชื่อถือของผลการวัด ในขณะที่ G-Theory ได้ให้แนวคิดแยกส่วน ความคลาดเคลื่อน (error) จากหลายแหล่ง (multiple error sources) ประกอบด้วย แหล่งความคลาดเคลื่อนอย่างเป็นระบบ (systematic source) อาทิเช่น ความยาวของแบบสอบ จำนวนครั้งในการทดสอบ จำนวนผู้ประเมิน เป็นองค์ประกอบของความคลาดเคลื่อนหรือฟาเซท (facet) และความคลาดเคลื่อนสุ่ม (random source) G-Theory จึงเป็นทั้งวิธีประเมินความเที่ยงของเครื่องมือวัดผล และออกแบบการวัดให้มีความเที่ยงสูงถึงระดับที่ต้องการ เพื่อนำผลไปใช้ในเป็นสารสนเทศสำหรับการตัดสินใจได้อย่างมีประสิทธิภาพ (ศิริชัย กาญจนวาสี, 2563) ดังนั้นจะเห็นได้ว่าทั้งจำนวนเหตุการณ์ และจำนวนผู้ประเมิน ล้วนเป็นฟาเซทที่ต้องใช้ G-Theory ในการวิเคราะห์แหล่งความแปรปรวนที่ส่งผลต่อความเที่ยงของแบบสอบอัตนัยประยุกต์

จากความสำคัญของการวัดทักษะการคิดขั้นสูงตามหลักสูตรฐานสมรรถนะ และงานวิจัยที่เกี่ยวข้องกับแบบสอบอัตนัยประยุกต์ ผู้วิจัยจึงพัฒนาแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ตามเกณฑ์ตามหลักสูตรฐานสมรรถนะที่ใช้วิธีตรวจให้คะแนนรูบริกสองชั้น โดย

เปรียบเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้จำนวนผู้ประเมินและจำนวนเหตุการณ์ที่ต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดในการวิเคราะห์ความน่าเชื่อถือของผลการวัดในสถานการณ์ต่าง ๆ และควบคุมแหล่งความคลาดเคลื่อนได้ตรงประเด็น เพื่อนำผลที่ได้ไปใช้ในการตัดสินใจการสรุปอ้างอิงแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ภายใต้จำนวนผู้ประเมินและจำนวนเหตุการณ์ที่ต่างกัน ว่าวิธีตรวจให้คะแนนใดจะออกแบบการวัดในการนำไปใช้จริงได้เหมาะสมจนได้แบบสอบที่มีความเที่ยงในระดับที่ต้องการ เพื่อให้ได้แบบสอบที่มีคุณภาพต่อการนำไปใช้และเป็นประโยชน์ต่อผู้เกี่ยวข้องกับการจัดการเรียนการสอนและระบบการศึกษาไทยต่อไป

คำถามวิจัย

1. แบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์และวิธีการตรวจให้คะแนนรูปกรสองชั้นสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ควรมีลักษณะอย่างไร
2. คุณภาพของวิธีตรวจให้คะแนนรูปกรสองชั้นสำหรับแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 เป็นอย่างไร
3. แบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกันสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จะมีความเที่ยงของแบบสอบต่างกันหรือไม่ อย่างไร

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์และวิธีตรวจให้คะแนนรูปกรสองชั้นสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3
2. เพื่อตรวจสอบคุณภาพวิธีตรวจให้คะแนนรูปกรสองชั้นสำหรับแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3
3. เพื่อเปรียบเทียบความเที่ยงของแบบสอบอัตรณ์ประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนที่แตกต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง

สมมุติฐานการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบสอบอัตนัยประยุกต์เพื่อศึกษาวิธีการตรวจให้คะแนน พบว่า วิธีการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์มี 2 วิธี คือ วิธีการตรวจให้คะแนนแบบ Knox ที่มีรูปแบบการให้คะแนนเป็น 0,1,2,3... คะแนน และวิธีการตรวจให้คะแนนแบบ Feletti ที่มีรูปแบบการให้คะแนนเป็น 0,1 คะแนน (สุไรยา หมัดหมั่น, 2549) นอกจากนี้ยังมีวิธีการตรวจให้คะแนนแบบสอบประเภทความเรียง 2 วิธี คือ วิธีการประเมินรวมและวิธีวิเคราะห์ย่อย จากงานวิจัยของมาฆพันธ์ อ๋านาคิล และกมลวรรณ ตังธนกานนท์ (2557) พบว่า วิธีตรวจให้คะแนนวิเคราะห์ย่อยมีความเที่ยงมากกว่าวิธีตรวจให้คะแนน Knox และวิธีตรวจให้คะแนน Feletti ตามลำดับ เนื่องจากวิธีตรวจให้คะแนนวิเคราะห์ย่อยมีความละเอียดเหมาะแก่การนำไปใช้และให้ข้อมูลย้อนกลับในการพัฒนาผู้สอบได้ แต่วิธีตรวจให้คะแนน Knox และวิธีตรวจให้คะแนน Feletti ยังไม่สามารถใช้เป็นข้อมูลย้อนกลับได้ถึงแม้ว่าวิธีตรวจให้คะแนน Knox จะมีความละเอียดมากกว่าวิธีตรวจให้คะแนน Feletti ก็ตาม และจากศึกษาวิธีการตรวจให้คะแนนเพิ่มเติมจากงานวิจัยของณัฐฐณี ศิริโชติ และกมลวรรณ ตังธนกานนท์ (2559) พบว่า วิธีการให้คะแนนแบบรูปกรีกสองชั้นจัดเป็นเกณฑ์การให้คะแนนรูปกรีกที่พัฒนามาจากวิธีตรวจให้คะแนนรูปกรีกแยกองค์ประกอบอีกประเภทหนึ่งที่ทำให้ผลการประเมินละเอียดมากยิ่งขึ้นและง่ายต่อการประเมิน มีจุดเด่นในเกณฑ์การให้คะแนนที่แบ่งมิติหรือข้อรายการย่อยมากขึ้นเพิ่มความเป็นปรนัยในการตรวจให้คะแนน

จากเกณฑ์ที่ Nunnally (1978) เสนอไว้ว่า แบบสอบที่ใช้ในการศึกษาควรมีความเที่ยงอย่างน้อย 0.7 และจากการศึกษางานวิจัยของสุภชิต ผดุงผล และกมลวรรณ ตังธนกานนท์ (2559) ที่ศึกษาในเรื่องจำนวนผู้ประเมินที่ใช้สำหรับแบบสอบอัตนัยประยุกต์ พบว่า ผลการประเมินจะเป็นอย่างไรนั้นขึ้นอยู่กับผู้ตรวจให้คะแนน เมื่อเพิ่มจำนวนผู้ตรวจให้คะแนนจะทำให้ความเที่ยงมีค่าสูงขึ้น (Brennan et al., 1995; Sudweeks et al., 2004; Carl, W. S. et al., 1999) และจากการศึกษางานวิจัยของมาฆพันธ์ อ๋านาคิล และกมลวรรณ ตังธนกานนท์ (2557) ที่ศึกษาในเรื่องจำนวนเหตุการณ์ที่ใช้สำหรับแบบสอบอัตนัยประยุกต์ที่มีความเที่ยงตั้งแต่ 0.7 ขึ้นไป พบว่า แบบสอบส่วนใหญ่ใช้จำนวนเหตุการณ์มากเกินไป ส่งผลให้แบบสอบมีจำนวนข้อย่อยมากตามไปด้วย ตัวอย่างเช่นแบบสอบอัตนัยประยุกต์ที่ Feletti (1980) พัฒนาขึ้นมีจำนวน 6 สถานการณ์ แบบสอบอัตนัยประยุกต์ที่สุพัฒตา ภูสอดสี (2553) พัฒนาขึ้นเพื่อวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ มีจำนวน 8 เหตุการณ์ แบบสอบอัตนัยประยุกต์ที่พรพิมล ปิลันธนาภรณ์ และชิตชนก เชิงเยาว์ (2562) พัฒนาขึ้นเพื่อวัดความสามารถในการตัดสินใจทางคลินิกของนักศึกษา

แพทย์มีจำนวน 5 เหตุการณ์ ซึ่งเป็นสาเหตุหนึ่งส่งผลต่อประสิทธิภาพในการวัด คือ ถ้าแบบสอบมีความยาวมากเกินไปจะทำให้ผู้สอบเกิดความเมื่อยล้าในการตอบ ดังนั้นผู้วิจัยต้องการหาข้อสรุปของจำนวนเหตุการณ์ภายใต้จำนวนผู้ประเมินที่ต่างกันในแต่ละวิธีตรวจให้คะแนนสำหรับแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ให้แบบสอบมีความเที่ยงในระดับที่ต้องการ จึงกำหนดสมมุติฐานดังนี้

แบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนนที่ต่างกัน ในระดับความเที่ยงที่ยอมรับ 0.7 ขึ้นไป วิธีตรวจให้คะแนนรูปกรอกสองชั้นจะมีความเที่ยงสูงสุดเมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้จำนวนผู้ประเมินและจำนวนเหตุการณ์ที่ต่างกัน

ขอบเขตของการวิจัย

1. ประชากรที่ใช้ในการวิจัยครั้งนี้ ได้แก่ นักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปีที่ 3 โรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา สุพรรณบุรี
2. แบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีแหล่งความคลาดเคลื่อนที่สำคัญ 2 แหล่งภายใต้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ดังนี้
 - 1) จำนวนผู้ประเมิน (r) ได้แก่ ผู้ประเมินที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 1 และ 2 คน
 - 2) จำนวนเหตุการณ์ (o) ได้แก่ 2, 3, 4, 5 และ 6 เหตุการณ์ โดยสุ่มข้อสอบตามสาระทางคณิตศาสตร์ จำนวน 2 สาระ ได้แก่ จำนวนและพีชคณิต การวัดและเรขาคณิต
3. เนื้อหาวิชาที่ใช้ ได้แก่ วิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น ตามสาระทางคณิตศาสตร์ จำนวน 2 สาระ ได้แก่ จำนวนและพีชคณิต (เรื่อง อัตราส่วน) และการวัดและเรขาคณิต (เรื่อง ทฤษฎีบทพีทาโกรัส และพื้นที่ผิวและปริมาตร)

คำจำกัดความที่ใช้ในการวิจัย

แบบสอบอัตโนมัติประยุกต์ หมายถึง แบบสอบประเภทความเรียงที่ใช้เหตุการณ์จำลองเป็นกรณีศึกษาที่มีเนื้อหาตามที่ต้องการออกข้อสอบให้ครอบคลุมตามวัตถุประสงค์ที่ตั้งไว้และมีการเสนอเหตุการณ์จำลองตามลำดับเหตุการณ์ และเสนอแยกออกเป็นข้อ ๆ ที่เป็นอิสระจากกัน แทรกคำถามตามวัตถุประสงค์ของการวัดเป็นระยะ ๆ ผู้สอบต้องนำข้อมูลจากเหตุการณ์ที่เสนอมา

ใช้ตัดสินใจเพื่อตอบคำถามปลายเปิดเอง และต้องทำข้อสอบให้เสร็จที่ละหน้าเนื่องจากผู้สอบไม่สามารถย้อนกลับไปแก้ไขคำตอบในข้อที่ทำไปแล้วได้

แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ตามหลักสูตรฐานสมรรถนะ หมายถึง แบบสอบประเภทความเรียงที่เป็นการวัดผลแบบอิงกลุ่มที่ผู้วิจัยสร้างขึ้นเป็นเหตุการณ์จำลองที่ใกล้เคียงกับชีวิตจริงของผู้สอบ เพื่อลดความได้เปรียบเสียเปรียบของรายละเอียดข้อมูลในแต่ละเหตุการณ์ โดยมีเนื้อหาตามรายวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 ซึ่งจะเสนอเหตุการณ์จำลองตามลำดับเหตุการณ์ และเสนอแยกออกเป็นข้อ ๆ ที่เป็นอิสระจากกัน แทรกคำถามตามวัตถุประสงค์ของการวัดเป็นระยะ ๆ แต่ละเหตุการณ์ประกอบด้วยข้อคำถามจำนวน 5 ข้อ ตามขั้นตอนการคิดวิจารณ์ญาณเพื่อฝึกการจำแนกข้อเท็จจริงจากแหล่งข้อมูลที่กำหนดให้จนสามารถนำมาวิเคราะห์ และสร้างข้อสรุปของปัญหาในเหตุการณ์นั้นได้ สำหรับการทำให้แบบทดสอบผู้สอบต้องนำข้อมูลจากเหตุการณ์ที่เสนอมาใช้ตัดสินใจเพื่อตอบคำถามปลายเปิดเอง และต้องทำข้อสอบให้เสร็จที่ละหน้า เนื่องจากผู้สอบไม่สามารถย้อนกลับไปแก้ไขคำตอบในข้อที่ทำไปแล้วได้

การคิดวิจารณ์ญาณทางคณิตศาสตร์ หมายถึง กระบวนการคิดทางคณิตศาสตร์ของนักเรียนในระดับชั้นมัธยมศึกษาตอนต้นผ่านแบบสอบอัตนัยประยุกต์อาศัยเหตุผลในการพิจารณาข้อมูลและหลักฐานที่มีอยู่อย่างรอบคอบ ตรวจสอบความน่าเชื่อถือของข้อมูล ความสัมพันธ์ของข้อมูลที่เกี่ยวข้องกับปัญหาหรือเหตุการณ์ เพื่อสร้างกระบวนการจัดระบบข้อมูลในการแก้ปัญหาหรือตั้งสมมติฐาน นำมาสู่การสร้างข้อสรุปที่มีหลักฐานมาสนับสนุนความคิด ตั้งอยู่บนตรรกะอย่างมีเหตุและผล จนได้คำตอบที่เหมาะสม ซึ่งกระบวนการคิดวิจารณ์ญาณ มี 5 ขั้นตอน ได้แก่ 1) ขั้นนิยามปัญหา 2) ขั้นตั้งสมมติฐาน 3) ขั้นรวบรวมข้อมูล 4) ขั้นวิเคราะห์ข้อมูล และ 5) ขั้นสร้างข้อสรุป ซึ่งแต่ละขั้นตอน มีความหมายดังนี้

1) ขั้นนิยามปัญหา หมายถึง การกำหนดประเด็นปัญหาจากการพิจารณาข้อมูลหรือข้อโต้แย้ง ภายใต้ข้อตกลงเบื้องต้นในเหตุการณ์ที่กำหนดให้ โดยปราศจากการให้ข้อมูลอื่นที่นอกเหนือจากบริบทของสถานการณ์หรือความคิดเห็นส่วนตัวเพิ่มเข้าไป

2) ขั้นตั้งสมมติฐาน หมายถึง การสร้างแนวทางคำตอบที่เป็นไปได้มากที่สุดหรือคาดคะเนแนวทาง ในการหาคำตอบ โดยอาศัยข้อมูลในสถานการณ์ที่มีอยู่เดิมประกอบการตัดสินใจอย่างสมเหตุสมผล

3) **ชั้นรวบรวมข้อมูล** หมายถึง การคัดเลือกข้อมูลที่มีความสัมพันธ์กับประเด็นปัญหา ข้อโต้แย้ง หรือข้อมูลที่คลุมเครือ โดยดึงข้อมูลหรือความรู้จากประสบการณ์เดิมมาใช้ เพื่อแสวงหาข้อมูลที่ถูกต้องและชัดเจนมากยิ่งขึ้น

4) **ชั้นวิเคราะห์ข้อมูล** หมายถึง การนำข้อมูลที่ผ่านการรวบรวมจากการพิจารณาความน่าเชื่อถือและ ความเพียงพอของข้อมูล มาเชื่อมโยงหาความสัมพันธ์โดยใช้ความรู้ทางคณิตศาสตร์ เพื่อนำไปสู่การสร้างข้อสรุป

5) **ชั้นสร้างข้อสรุป** หมายถึง การนำผลการวิเคราะห์ข้อมูล และหลักฐานที่มีความหนักแน่นน่าเชื่อถือมาใช้ในการตัดสินใจสรุปผลการทดสอบสมมติฐานอย่างเหมาะสมกับสถานการณ์ โดยตั้งอยู่บนตรรกะเหตุและ ผลทางคณิตศาสตร์

เงื่อนไขการทดสอบ หมายถึง เงื่อนไขการวัดที่มีผลต่อความคลาดเคลื่อนของการสรุปอ้างอิงคุณภาพของแบบสอบ ประกอบด้วย จำนวนผู้ประเมิน และจำนวนเหตุการณ์ ภายใต้วิธีการตรวจให้คะแนนที่ต่างกัน

วิธีการตรวจให้คะแนน หมายถึง วิธีการตรวจคำตอบแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ตามเกณฑ์ที่เฉลยคำตอบไว้ว่าตอบอย่างไร และได้กี่คะแนน มีทั้งหมด 3 วิธี คือ

วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) หมายถึง วิธีการตรวจให้คะแนนแบบสอบอัตโนมัติประยุกต์กับเกณฑ์การให้คะแนนแบบรูปริก 2 ชั้น โดยแบ่งออกเป็น 2 ระดับชั้น ชั้นที่ 1 อธิบายตามองค์ประกอบเช่นเดียวกับวิธีตรวจให้คะแนนวิเคราะห์ย่อย และในแต่ละองค์ประกอบจะประเมินตามข้อรายการย่อย โดยใช้มาตราประมาณค่า 3 ระดับ ได้แก่ 2 คะแนน คือ ถูกต้องทั้งหมด, 1 คะแนน คือ ถูกต้องบางส่วน และ 0 คะแนน คือ ไม่ถูกต้องทั้งหมด หรือไม่มีการตอบคำถาม จากนั้นนำผลรวมของคะแนนในระดับชั้นที่ 1 มาแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) หมายถึง วิธีการตรวจให้คะแนนแบบสอบอัตโนมัติประยุกต์ โดยมีองค์ประกอบของเกณฑ์การให้คะแนนข้อคำถามที่ 1 นิยามปัญหา ข้อคำถามที่ 2 ตั้งสมมติฐาน และข้อคำถามที่ 3 รวบรวมข้อมูล ของแต่ละเหตุการณ์ มี 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) และในข้อคำถามที่ 4 วิเคราะห์ข้อมูล และข้อคำถามที่ 5 สร้างข้อสรุป เพิ่มอีกหนึ่งองค์ประกอบ คือ การสื่อความหมายทางคณิตศาสตร์ (2 คะแนน) ซึ่งในแต่ละองค์ประกอบมีเกณฑ์การให้คะแนนเป็น 3 ระดับเช่นเดียวกันกับวิธีการตรวจให้คะแนนตามแบบ Knox

โดยความถูกต้องของเนื้อหา หมายถึง การระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์, การเข้าใจในสิ่งที่ต้องการวัด หมายถึง การกำหนดชั้นของการวัดภายใต้บริบทของสถานการณ์ และการสื่อความหมายทางคณิตศาสตร์ หมายถึง การเปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ และคำนวณได้อย่างถูกต้อง

วิธีตรวจให้คะแนน Knox หมายถึง วิธีการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ โดยมีองค์ประกอบของเกณฑ์การให้คะแนนแต่ละข้อคำถามของแต่ละเหตุการณ์เป็น 3 ระดับ ได้แก่ 2 คะแนน คือ ผู้ตอบตอบคำถามได้ถูกต้องและครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ 1 คะแนน คือ ผู้ตอบตอบคำถามผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ และ 0 คะแนน คือ ผู้ตอบตอบคำถามไม่ถูกต้องตามแนวคำตอบ หรือไม่มีการตอบคำถาม

จำนวนเหตุการณ์ หมายถึง ชุดของปัญหาของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

ผู้ประเมิน หมายถึง จำนวนผู้ตรวจให้คะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยครูผู้ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไปที่ได้รับการอบรม ซึ่งแจ้งวิธีการตรวจและกฎเกณฑ์การให้คะแนนจนมีความเข้าใจตรงกัน

ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory: G-Theory) หมายถึง ทฤษฎีสำหรับวิเคราะห์ความน่าเชื่อถือของผลการวัดที่แสดงประสิทธิภาพการสรุปอ้างอิงผลของการวัดที่ได้จากแบบสอบไปยังคะแนนเฉลี่ยที่ผู้สอบแต่ละคนควรได้รับภายใต้สถานการณ์การสอบหรือเงื่อนไขต่าง ๆ ของการวัด และสามารถจำแนกแหล่งความคลาดเคลื่อนได้หลายแหล่ง

ความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 หมายถึง ความเที่ยงของคะแนนที่ได้จากแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกันภายใต้จำนวนจำนวนผู้ประเมินและจำนวนเหตุการณ์ที่ต่างกัน โดยแสดงถึงระดับความแม่นยำและความน่าเชื่อถือของคะแนนที่ได้จากการวัด (level of dependability) เรียกว่า สัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Coefficient: G-Coefficient) ซึ่งคล้ายกับสัมประสิทธิ์ความเที่ยง (reliability coefficient) ในทฤษฎีการทดสอบแบบดั้งเดิม โดยพิจารณาค่าสัมประสิทธิ์การสรุป

อ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (relative coefficient) ที่บ่งบอกความเที่ยงของแบบสอบ
อัตนัยประยุกต์ที่เป็นการวัดผลแบบอิงกลุ่ม

ประโยชน์ที่ได้รับจากการวิจัย

1. ได้แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์และวิธี
ตรวจให้คะแนนแบบรูบริกสองชั้นสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีคุณภาพ
2. ได้วิธีตรวจให้คะแนนรูบริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิด
วิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3
3. ได้จำนวนเหตุการณ์ภายใต้จำนวนผู้ประเมินที่เหมาะสมกับความสามารถในการคิด
อย่างมีวิจารณ์ญาณทางคณิตศาสตร์ในแบบสอบอัตนัยประยุกต์ที่มีความเที่ยงในระดับที่ยอมรับ
ได้ และเป็นประโยชน์ในการนำเครื่องมือไปใช้ในวัดการคิดขั้นสูงประกอบกับการจัดการเรียนการ
สอน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับความเที่ยงภายใต้เงื่อนไขการทดสอบที่ต่างกันของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง โดยผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง โดยแบ่งการศึกษาออกเป็น 5 ตอน ดังนี้

ตอนที่ 1 ทักษะการคิดวิจารณ์ญาณ

ตอนที่ 2 แบบสอบอัตนัยประยุกต์

ตอนที่ 3 คุณภาพของการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์

ตอนที่ 4 ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory)

หรือ G – Theory

ตอนที่ 5 กรอบแนวคิดในการวิจัย

ตอนที่ 1 ทักษะการคิดวิจารณ์ญาณ

1.1 ทักษะการคิดวิจารณ์ญาณตามหลักสูตรฐานสมรรถนะ

หลักสูตรฐานสมรรถนะ (competency – based curriculum) หมายถึง หลักสูตรที่ยึดความสามารถของผู้เรียน มิได้มุ่งเน้นการจดจำความรู้หรือเนื้อหาที่อาจเปลี่ยนแปลงได้ตามกาลเวลาซึ่งต่างจากหลักสูตรปัจจุบันที่ให้ความรู้เป็นองค์ประกอบหลัก แต่หลักสูตรฐานสมรรถนะนี้จะให้ทักษะความสามารถเป็นองค์ประกอบหลัก ส่วนความรู้ เจตคติและคุณลักษณะต่าง ๆ เป็นองค์ประกอบรองที่ผลักดันให้ผู้เรียนสามารถปฏิบัติงาน ๆ หนึ่งได้ประสบผลสำเร็จและเป็นประโยชน์ต่อชีวิตประจำวัน รวมไปถึงในอนาคตของผู้เรียน โดยสามารถประยุกต์ใช้ความรู้ได้จริงในสถานการณ์ที่หลากหลาย หลักสูตรนี้จะมีเกณฑ์ทักษะความสามารถในด้านต่าง ๆ ที่ผู้เรียนต้องปฏิบัติได้ในแต่ละระดับการศึกษา เพื่อเป็นมาตรฐานรองรับความสามารถของผู้เรียนเมื่อจบการศึกษาว่าจะมีทักษะความสามารถในด้านต่าง ๆ ตามที่ต้องการ ซึ่งหลักสูตรฐานสมรรถนะนี้จะกำหนดการพัฒนาทักษะความสามารถในแต่ละระดับให้มีความต่อเนื่อง เพื่อเพิ่มพูนทักษะและความสามารถในการพัฒนาต่อยอดในอนาคตต่อไป (สุจิตรา ปทุมลังการ์, 2552; สำนักงาน

เลขาธิการสภาการศึกษา, 2563; อนุสรณ์ หงษ์ขุนทด, 2563, กลุ่มนิเทศ ติดตาม และประเมินผล การจัดการศึกษา สพ.น.นครปฐม เขต 2, 2564)

สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้น พื้นฐาน (2563) ได้จัดร่างระดับสมรรถนะทั้งห้า (scales of five competencies) เพื่อรองรับ หลักสูตรฐานสมรรถนะ ประกอบด้วย สมรรถนะที่ 1 การจัดการตนเอง (Self Management: SM) สมรรถนะที่ 2 การสื่อสาร (Communication: CM) สมรรถนะที่ 3 การรวมพลังทำงานเป็นทีม (Teamwork and Collaboration: TC) สมรรถนะที่ 4 การคิดขั้นสูง (Higher Order Thinking: HOT) และสมรรถนะที่ 5 การเป็นพลเมืองที่เข้มแข็ง (Active Citizen: AC)

เมื่อพิจารณาสมรรถนะที่เกี่ยวข้องกับกระบวนการทางปัญญา และสามารถเชื่อมโยง สมรรถนะอื่น ๆ ให้พัฒนาต่อยอดไปได้ นั่นก็คือ สมรรถนะที่ 4 การคิดขั้นสูง (Higher Order Thinking: HOT) เป็นการคิดวิพากษ์ คิดไตร่ตรองอย่างมีระบบ สามารถใช้ข้อมูลในการตัดสินใจ แก้ปัญหาโดยคำนึงถึงผลกระทบรอบด้าน พบว่า มีลักษณะที่สำคัญ 4 ด้าน คือ 1) การคิดอย่างมี วิจารณ์ญาณ (Critical Thinking: HOT-CTC) 2) การคิดเป็นระบบ (Systematic Thinking: HOT-STM) 3) การคิดแก้ปัญหา (Problem Solving Thinking: HOT-PRB) และ 4) การคิดสร้างสรรค์ (Creative Thinking: HOT-CRT) โดยทักษะการคิดวิจารณ์ญาณ (Critical Thinking Skills) เป็น กระบวนการคิดที่ผสมผสานทักษะการคิดแก้ปัญหาอย่างเป็นระบบโดยใช้เหตุผลวิเคราะห์ ไตร่ตรองข้อมูล เพื่อนำมาสู่การสร้างข้อสรุปอย่างสร้างสรรค์

โปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล (Programme for International Student Assessment หรือ PISA) เน้นการประเมินสมรรถนะของนักเรียนในการใช้ ความรู้และทักษะในชีวิตจริง ซึ่งเป็น การประเมินความฉลาดรู้ทางการอ่าน คณิตศาสตร์ และวิทยาศาสตร์ ผ่านลักษณะข้อสอบที่หลากหลายสถานการณ์ในชีวิตจริง (สถาบันส่งเสริมการ สอนวิทยาศาสตร์และเทคโนโลยี, 2564) จะเห็นได้ว่าผู้เรียนจำเป็นต้องมีพื้นฐานทักษะการอ่าน และคณิตศาสตร์ในการวิเคราะห์ข้อมูล จนสามารถแก้ไข และสร้างข้อสรุปเพื่อตอบปัญหา ซึ่งต้อง อาศัยทักษะการคิดขั้นสูงและหนึ่งในทักษะการคิดขั้นสูงที่สำคัญตามหลักสูตรฐานสมรรถนะ คือ ทักษะการคิดวิจารณ์ญาณ (critical thinking skills) เนื่องจากเป็นกระบวนการคิดพิจารณา ไตร่ตรองข้อมูล และหลักฐานที่มีอยู่อย่างรอบคอบ ตรวจสอบความน่าเชื่อถือของข้อมูลหา ความสัมพันธ์ และจัดระบบข้อมูลเพื่อนำมาสู่การสร้างข้อสรุปที่มีหลักฐานมาสนับสนุนความคิด โดยตั้งอยู่บนตรรกะอย่างมีเหตุและผล จนได้แนวทางข้อสรุปที่เหมาะสมกับสถานการณ์ต่าง ๆ

(Glaser, 1941; Dumke, 1987; Ennis, 1985; Warnick & Inch, 1994; Halpern, 1999; อัมพร ม้าคะนอง, 2554; สิริภักตร์ ศิริโท, 2558; สำนักงานราชบัณฑิตยสภา, 2558)

ในทักษะการคิดขั้นสูงส่วนใหญ่จะนำทักษะการแก้ปัญหามาใช้ เนื่องจากทักษะนี้สามารถประยุกต์ใช้ได้หลากหลายวิชา ไม่ว่าจะเป็นวิทยาศาสตร์ หรือคณิตศาสตร์ แต่กลับพบว่า ยังไม่เคยมีการนำทักษะการคิดวิจารณ์มาพัฒนากระบวนการคิดของผู้เรียนในรายวิชา คณิตศาสตร์ในประเทศทั้งที่ทักษะนี้ประยุกต์ใช้ทักษะการคิดขั้นสูงในหลากหลายด้าน และเป็นทักษะพื้นฐานที่สำคัญในการนำไปต่อยอดสู่สมรรถนะอื่น ๆ ได้ ผู้วิจัยสนใจพัฒนาทักษะการคิดวิจารณ์ทางคณิตศาสตร์ให้เกิดขึ้นกับผู้เรียนผ่านการสร้างเครื่องมือที่มีประสิทธิภาพต่อไป

1.2 ขั้นตอนการคิดวิจารณ์

ขั้นตอนการคิดวิจารณ์ มีนักวิชาการหลายท่านได้กล่าวถึงไว้มากมาย บางขั้นตอนเหมือนกันและต่างกัน ดังนี้

Dressel and Mayhew (1957) ได้เสนอขั้นตอนการคิดวิจารณ์ไว้ ดังนี้

1. การนิยามปัญหา เป็นความสามารถในการกำหนดปัญหา ข้อโต้แย้ง วิเคราะห์ ข้อความหรือข้อมูลที่คลุมเครือให้ชัดเจน ภายใต้ขอบเขตข้อเท็จจริงที่กำหนดให้
2. การรวบรวมข้อมูลสำหรับการแก้ปัญหา เลือกข้อมูลที่เกี่ยวข้องกับปัญหา ข้อโต้แย้ง/ข้อโต้แย้ง/ข้อมูลที่คลุมเครือ เพื่อแสวงหาข้อมูลที่ถูกต้องและชัดเจนมากยิ่งขึ้น
3. การจัดระบบข้อมูล โดยการจำแนกความแตกต่างระหว่างข้อมูลที่ชัดเจนกับข้อมูลที่คลุมเครือ ข้อเท็จจริงกับข้อคิดเห็น
4. การเลือกสมมติฐาน เลือกสมมติฐานที่เป็นไปได้ทั้งหมดจากความสัมพันธ์เชิงเหตุผล ตรวจสอบ ความสอดคล้องระหว่างสมมติฐานกับข้อมูล พิจารณาทางเลือกหลาย ๆ ทางในการแก้ปัญหา
5. การลงข้อสรุป เป็นความสามารถในการคิดพิจารณาข้อความคลุมเครือของข้อมูล โดยจำแนกข้อมูลที่เหตุผลหนักแน่นและน่าเชื่อถือ เพื่อนำไปสู่การตัดสินใจสรุป ถ้าการสรุปไม่มีเหตุผลเพียงพอต้องมีการหาเหตุผลเพิ่มเติมมาพิจารณาการสรุป

Decaroli (1973) ได้เสนอขั้นตอนการคิดวิจารณ์ไว้ ดังนี้

1. การนิยามปัญหา เป็นการกำหนดปัญหา ทำความตกลงเกี่ยวกับความหมายของคำ และข้อความ รวมไปถึงการกำหนดเกณฑ์
2. การกำหนดสมมติฐาน หาความสัมพันธ์เชิงเหตุผล หาทางเลือก และพยากรณ์

3. การประมวลข้อมูล เป็นการระบุข้อมูลที่จำเป็น รวบรวมข้อมูลที่เกี่ยวข้อง การหาหลักฐาน และจัดระบบข้อมูล

4. การตีความข้อเท็จจริง และการสรุปอ้างอิงจากหลักฐาน
5. การใช้เหตุผล โดยระบุเหตุผล และผลลัพธ์เชิงตรรกศาสตร์
6. การประมวลผล โดยอาศัยเกณฑ์ความสมเหตุสมผล
7. การประยุกต์ใช้ หรือนำไปปฏิบัติ

Gagne (1985) ได้เสนอขั้นตอนการคิดวิจารณ์ญาณไว้ ดังนี้

1. สังเกตข้อความหรือสถานการณ์ที่ได้รับ เพื่อรวบรวมข้อมูลให้เป็นความคิดรวบยอดที่เชื่อมโยงความสัมพันธ์ต่าง ๆ

2. อธิบาย/ตอบคำถาม แสดงความคิดเห็นว่าเห็นด้วย/ไม่เห็นด้วยกับข้อมูลที่กำหนดให้โดยการอ้างอิงกฎเกณฑ์หรือหลักฐานประกอบที่น่าเชื่อถือ

3. รับฟังความคิดเห็นที่แตกต่าง เน้นการปรับเปลี่ยนความคิดเห็นอย่างมีเหตุผล

4. เชื่อมโยงความสัมพันธ์ เปรียบเทียบความแตกต่างและความคล้ายคลึงของสิ่งต่าง ๆ

5. วิเคราะห์เหตุการณ์ แนวคิด/การกระทำ เพื่อหาจุดเด่น/จุดด้อย ประกอบการยกเหตุผลและหลักฐานอ้างอิง

6. สรุปผลอย่างตรงไปตรงมา โดยใช้เหตุผลประกอบว่าข้อความที่กล่าวมานั้นเชื่อถือได้หรือไม่

Ennis (1987) ได้สร้างกรอบแนวคิดการวัดการคิดวิจารณ์ญาณออกเป็น 8 ทักษะ ดังนี้

1. การทำความเข้าใจปัญหา โดยสามารถบอกรายละเอียดพื้นฐานของปัญหาที่พบ และ จำรายละเอียดของปัญหาที่พบได้

2. การสังเกตความน่าเชื่อถือของข้อมูล โดยพิจารณาความน่าเชื่อถือของแหล่งข้อมูล สังเกตรายละเอียด พร้อมทั้งแปลความ และตัดสินผลของข้อมูลโดยใช้ประสาททั้งห้า

3. การใช้เหตุผลเชิงนิรนัย โดยการนำหลักการใหญ่ไปแตกเป็นหลักการย่อย ๆ ตั้งอยู่บนเหตุผลทางตรรกศาสตร์ จนสามารถสรุปผลที่จะตามมาของข้อสรุปนั้นได้

4. การใช้เหตุผลเชิงอุปนัย โดยสามารถหาเหตุผลเพื่อหาข้อสรุปจากหลักการย่อย ๆ ไปสู่การสรุปเป็นเหตุการณ์ทั่วไปได้ อาจมีการยกตัวอย่าง หรือบอกรายละเอียดย่อย ๆ ของเนื้อหาให้ครอบคลุม

5. การตัดสินใจคุณค่า โดยพิจารณาทางเลือกที่มีข้อมูลเพียงพอ พิจารณาว่าอะไรจะเป็นผลที่เกิดขึ้น และชั่งน้ำหนักผลดีและผลเสียก่อนตัดสินใจ

6. การแปลความหมาย โดยสามารถให้คำนิยามเชิงปฏิบัติการได้ ยกตัวอย่างที่ใช่ และไม่ใช่ได้ หรือสามารถจำแนกและจัดกลุ่มได้

7. การกำหนดข้อสมมติฐาน โดยสามารถใช้เหตุผลในอธิบายสาเหตุของการกำหนดข้อสันนิษฐานได้จากข้อความที่จัดแสดง

8. การแก้ปัญหา โดยกำหนดแนวทางในการแก้ปัญหาที่ดีที่สุดเพียงวิธีเดียวได้ มีการเลือกเกณฑ์ตัดสินผลที่เป็นไปได้ และทบทวนทางเลือกอย่างสมเหตุสมผล

Woolfolk (1995) ได้เสนอขั้นตอนการคิดวิจารณ์ญาณไว้ ดังนี้

1. นิยามและทำความเข้าใจกับปัญหา ประกอบด้วย 1) การระบุประเด็นที่สำคัญ/ระบุปัญหา 2) การเปรียบเทียบความคล้ายคลึง/ความแตกต่าง 3) การตัดสินใจว่าข้อมูลใดชัดเจน/ข้อมูลใดคลุมเครือ/ข้อมูลใดจำเป็น/ข้อมูลใดไม่จำเป็น 4) การตั้งคำถามที่จะนำไปสู่ความเข้าใจที่ชัดเจนลึกซึ้งซึ่งเกี่ยวกับเรื่องราวหรือสถานการณ์นั้น ๆ

2. รวบรวมข้อมูลที่มีความสัมพันธ์กับปัญหา ประกอบด้วย 1) จำแนกความแตกต่างระหว่างข้อเท็จจริงกับข้อคิดเห็น 2) ตัดสินว่าข้อความนั้นมีความสอดคล้องสัมพันธ์กัน/สอดคล้องกันทั้งหมดหรือไม่ 3) คาดเดาหรือระบุสมมติฐานที่ไม่ได้กล่าวไว้ในข้ออ้างเหตุผล 4) ระบุความคิดที่มีอคติหรือปัจจัยด้านอารมณ์

3. แก้ปัญหาและลงข้อสรุป ประกอบด้วย 1) ระบุความเพียงพอของข้อมูลที่สามารถตัดสินใจได้ว่าข้อมูลที่มีอยู่เพียงพอหรือไม่ 2) พยากรณ์/ทำนายผลที่อาจเป็นไปได้ทั้งหมด

Facione, P. A. (1992) ได้สร้างแบบวัดความสามารถในการคิดวิจารณ์ญาณ The California Critical Thinking Skill Test (CCTST) มีองค์ประกอบ 5 ประการ ดังนี้

1. การตีความ (interpretation) เพื่อสร้างความเข้าใจ และนำไปสู่การกระทำที่หลากหลาย สถานการณ์ ข้อมูล เหตุการณ์ การตัดสินใจ การประชุม ความเชื่อ กฎ กระบวนการหรือเกณฑ์มาใช้ การตีความประกอบด้วย การถอดความ และการให้ความหมายที่ชัดเจน

2. การวิเคราะห์ (analysis) เพื่อระบุองค์ประกอบของสถานการณ์ และเห็นความสัมพันธ์ขององค์ประกอบเหล่านั้น

3. การประเมินผล (evaluation) เพื่อประเมินความน่าเชื่อถือของแหล่งข้อมูลว่ามีความน่าเชื่อถือมากหรือน้อยเพียงใด และยังใช้ตัดสินคุณภาพของการวิเคราะห์ การตีความ การอธิบาย และการอนุมาน

4. การอนุมานหรือการลงความเห็นเพื่อสร้างข้อสรุป (inference) เพื่อสร้างข้อสรุปจากเหตุผลและหลักฐานที่มีอยู่

5. การอธิบาย (explanation) เพื่อแสดงเหตุผล โดยรวบรวมหลักฐานและกระบวนการพิสูจน์มาสนับสนุน

Watson, G. (1980) ได้เสนอขั้นตอนการคิดวิเคราะห์ญาณไว้ ดังนี้

1. การระบุข้อตกลงเบื้องต้น (recognition of assumption) หมายถึง การตรวจสอบความสมเหตุสมผลของข้อตกลงเบื้องต้นจากข้อสรุปที่กำหนดให้

2. การอนุมาน (inference) หมายถึง การตรวจสอบความเป็นไปได้ของข้อสรุปจากเหตุการณ์ ที่กำหนดให้

3. การนิรนัย (deduction) หมายถึง การตรวจสอบความสมเหตุสมผลของข้อสรุปจากข้อมูลหรือเหตุ (premises) ที่กำหนดให้

4. การประเมินข้อสรุป (evaluation of argument) หมายถึง การตรวจสอบความน่าเชื่อถือและเหตุผลของการตัดสินใจจากคำถามที่กำหนดให้

5. การตีความ (interpretation) หมายถึง การตรวจสอบความสมเหตุสมผลของข้อความจากขอบเขต และเงื่อนไขต่าง ๆ ที่กำหนดให้

วรนิพิฏ์ พันธุ์หนองหว่า (2559) ได้สรุปความสามารถในการคิดวิเคราะห์ญาณของผู้เรียนร่วมกับกระบวนการแก้ปัญหาที่มีโครงสร้างไม่สมบูรณ์ทางคณิตศาสตร์ ตามแนวคิดของ Facione, P. A. (1990) และ Watson and Glaser (2002) ไว้ 5 ประการ ดังนี้

1. ความสามารถในการให้เหตุผลแบบนิรนัย เป็นความสามารถในการจำแนกว่าข้อสรุปใดสมเหตุสมผล โดยอิงความรู้ หลักการ และทฤษฎี

2. ความสามารถในการกำหนดข้อตกลงเบื้องต้น เป็นความสามารถในการจำแนกว่าข้อตกลงใดเป็นข้อตกลงหรือเงื่อนไขที่ได้จากสถานการณ์ที่กำหนด

3. ความสามารถในการวิเคราะห์ เป็นความสามารถในการระบุองค์ประกอบของสถานการณ์ และเชื่อมโยงความสัมพันธ์ขององค์ประกอบเหล่านั้น

4. ความสามารถในการอนุมานหรือการลงความเห็นเพื่อสร้างข้อสรุป เป็นความสามารถในการจำแนกระดับความน่าจะเป็นของข้อสรุปที่ได้จากการสังเกตหรือคาดคะเนความจริงจากสถานการณ์ที่กำหนดให้

5. ความสามารถในการประเมินข้อสรุป เป็นความสามารถในการจำแนกว่าข้อโต้แย้งใด มีความหนักแน่นน่าเชื่อถือ โดยพิจารณาจากเหตุผลประกอบข้อโต้แย้งนั้น

จากการศึกษาเอกสาร และงานวิจัยที่เกี่ยวข้องกับกระบวนการคิดวิเคราะห์ของนักศึกษาดังที่กล่าวไว้ข้างต้น ผู้วิจัยสามารถสังเคราะห์กระบวนการคิดวิเคราะห์ของนักศึกษาดังกล่าวออกมาได้ ดังตาราง 1

ตาราง 1 การสังเคราะห์กระบวนการคิดวิเคราะห์ของนักศึกษาดังกล่าว

ขั้นตอน	นักการศึกษา									รวม
	Dressel and Mayhew (1957)	Decaroli (1973)	Gagne (1985)	Ennis (1987)	Woolfolk (1995)	Facione (1992)	Watson (1980)	วรมณีพิฎ พันธ์ทองแห้ว (2559)		
การนิยามปัญหา	P	P		P	P					4
การระบุข้อตกลงเบื้องต้น				P	P		P	P		4
การให้เหตุผลเชิงนิรนัย				P			P	P		3
การตั้งสมมติฐาน	P	P		P	P					4
การหาความสัมพันธ์ข้อมูล		P	P	P	P	P	P	P		7
การรวบรวมข้อมูล	P		P		P	P				4
การแสดงความคิดเห็น			P	P	P					3
การจัดระบบข้อมูล	P	P								2
การวิเคราะห์ข้อมูล			P		P	P		P		4
การเชื่อมโยงความรู้			P					P		3
การแก้ปัญหา				P	P					2
การลงข้อสรุป/การให้เหตุผลเชิงอุปนัย	P	P	P	P	P	P		P		7
การประเมินข้อสรุป							P	P		2

จากการสังเคราะห์ขั้นตอนการคิดวิเคราะห์ของนักศึกษาดังกล่าวข้างต้น ผู้วิจัยได้รวบรวมขั้นตอนการคิดวิเคราะห์ของนักศึกษาดังกล่าวที่ตรงกัน 4 คนขึ้นไปจากทั้งหมด 8 คน ออกมาได้ 5 ขั้นตอน ดังนี้

1. ขั้นนิยามปัญหา หมายถึง การกำหนดประเด็นปัญหาจากการพิจารณาข้อมูลหรือข้อโต้แย้ง ภายใต้ข้อตกลงเบื้องต้นในสถานการณ์ที่กำหนดให้ โดยปราศจากการให้ข้อมูลอื่นที่นอกเหนือจากบริบทของสถานการณ์หรือความคิดเห็นส่วนตัวเพิ่มเข้าไป

2. **ขั้นตั้งสมมติฐาน** หมายถึง การสร้างแนวทางคำตอบที่เป็นไปได้มากที่สุดหรือคาดคะเนแนวทางในการหาคำตอบ โดยอาศัยข้อมูลในสถานการณ์ที่มีอยู่เดิมประกอบการตัดสินใจอย่างสมเหตุสมผล

3. **ขั้นรวบรวมข้อมูล** หมายถึง การคัดเลือกข้อมูลที่มีความสัมพันธ์กับประเด็นปัญหาข้อโต้แย้ง หรือข้อมูลที่คลุมเครือ โดยดึงข้อมูลหรือความรู้จากประสบการณ์เดิมมาใช้ เพื่อแสวงหาข้อมูลที่ถูกต้องและชัดเจนมากยิ่งขึ้น

4. **ขั้นวิเคราะห์ข้อมูล** หมายถึง การนำข้อมูลที่ผ่านการรวบรวมจากการพิจารณาความน่าเชื่อถือและความเพียงพอของข้อมูล มาเชื่อมโยงหาความสัมพันธ์โดยใช้ความรู้ทางคณิตศาสตร์ เพื่อนำไปสู่การสร้างข้อสรุป

5. **ขั้นสร้างข้อสรุป** หมายถึง การนำผลการวิเคราะห์ข้อมูล และหลักฐานที่มีความหนักแน่นน่าเชื่อถือ มาใช้ในการตัดสินใจสรุปผลการทดสอบสมมติฐานอย่างเหมาะสมกับสถานการณ์ โดยตั้งอยู่บนตรรกะเหตุและผลทางคณิตศาสตร์

1.3 งานวิจัยที่เกี่ยวข้องกับทักษะการคิดวิจารณ์ญาณ

จากการศึกษางานวิจัยในประเทศและงานวิจัยต่างประเทศที่เกี่ยวข้องกับทักษะการคิดวิจารณ์ญาณ มีดังนี้

ทวี สระน้ำคำ (2551) ศึกษาผลของวิธีสอนแบบนิรนัยและวิธีสอนแบบอุปนัยที่มีแบบฝึกหัดหลังเรียนต่างกัน โดยใช้บทเรียนบนเว็บในวิชาฟิสิกส์ที่มีต่อการคิดวิจารณ์ญาณของนักเรียนมัธยมศึกษาปีที่ 4 จากตัวอย่างวิจัย จำนวน 160 คน ที่ผ่านการใช้เครื่องมือแบบวัดการคิดวิจารณ์ญาณเนื้อหาทางการพยาบาลที่สร้างขึ้นจากกรอบแนวคิดของเอนนิสที่แบ่งการคิดวิจารณ์ญาณออกเป็น 8 ทักษะ ได้แก่ 1) การทำความเข้าใจปัญหา 2) การสังเกตความน่าเชื่อถือของข้อมูล 3) การใช้เหตุผลเชิงนิรนัย 4) การใช้เหตุผลเชิงอุปนัย 5) การตัดสินใจคุณค่า 6) การแปลความหมาย 7) การกำหนดข้อสมมติฐาน และ 8) การแก้ปัญหา ซึ่งลักษณะของแบบวัดเป็นปรนัย 4 ตัวเลือก จำนวนทั้งหมด 40 ข้อ โดยข้อคำถามหนึ่งข้ออาจใช้ความสามารถในการคิดวิจารณ์ญาณมากกว่าหนึ่งทักษะมาช่วยหาคำตอบ และจากผลการวิจัย พบว่า แบบฝึกหัดหลังเรียนในบทเรียนบนเว็บทั้ง 2 แบบ ได้แก่ แบบฝึกหัดหลังเรียนโดยวิธีสร้างโจทย์ และโดยวิธีแก้โจทย์ ทั้งสองแบบส่งผลต่อการคิดวิจารณ์ญาณของนักเรียนแตกต่างกันอย่างมีนัยสำคัญ ผู้วิจัยจึงเสนอแนะการวิจัยครั้งต่อไปให้ใช้แบบฝึกหัดหลังเรียนแบบผสม คือ ใช้ทั้งวิธีสร้างโจทย์ และแก้โจทย์มาเป็นตัวแปรที่สามเพื่อเปรียบเทียบให้เห็นว่าผลต่อการคิดวิจารณ์ญาณของนักเรียน

แตกต่างกันหรือไม่ สำหรับการวิจัยนี้ทั้งแบบฝึกหัดและแบบวัดการคิดวิจารณ์ของนักเรียนใช้รูปแบบปรนัย 4 ตัวเลือกทั้งหมด

Richard et al. (2019) ศึกษาการประเมินการคิดอย่างมีวิจารณญาณของนักศึกษาระดับมหาวิทยาลัย โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด จากกลุ่มตัวอย่างที่เป็นนักศึกษามาจากมหาวิทยาลัย German จำนวน 30 คน ใช้การประเมินงานเขียน 23 ข้อ ซึ่งแต่ละข้อต้องเขียนสรุปใจความสำคัญ แสดงเหตุผลในการให้ข้อเสนอแนะข้อโต้แย้งจากสถานการณ์ที่กำหนดให้ ภายในเวลา 1 ชั่วโมง โดยใช้เกณฑ์การให้คะแนนรูปแบบแยกองค์ประกอบ (analytic scoring rubric) จำนวน 4 ฉบับ เป็นเครื่องมือที่ใช้ในการประเมินงานเขียน ประกอบด้วย rubric ฉบับที่ 1 (ประกอบด้วย 4 องค์ประกอบ) ประเมินการความเข้าใจใจที่ rubric ฉบับที่ 2 (ประกอบด้วย 9 องค์ประกอบ) ประเมินการแยกข้อเท็จจริงและข้อคิดเห็นของใจที่ rubric ฉบับที่ 3 (ประกอบด้วย 4 องค์ประกอบ) ประเมินการตัดสินใจ และ rubric ฉบับที่ 4 (ประกอบด้วย 6 องค์ประกอบ) ประเมินการเขียน โดยมีผู้ตรวจ 4 คนที่ผ่านการฝึกการตรวจมาอย่างดี ในการวิเคราะห์และตรวจสอบคุณภาพของแบบทดสอบ ประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดที่มีสององค์ประกอบคือ จำนวนข้อ และจำนวนผู้ตรวจ พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมพันธ์และ ค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์ไม่ต่ำกว่า 0.70 กรณีที่จำนวนผู้ตรวจ 2 คน ภายใต้จำนวนข้อคำถาม 23 ข้อ และจากการศึกษาเพื่อการตัดสินใจสรุปอ้างอิง พบว่าค่าความเที่ยงจะสูงขึ้น เมื่อจำนวนข้อคำถามเพิ่มขึ้นภายใต้จำนวนผู้ตรวจ 2 คนขึ้นไป จึงจะมีค่าความเที่ยงไม่ต่ำกว่า 0.6

Mehtap and Erdinc (2020) ศึกษาความสามารถในกระบวนการคิดอย่างมีวิจารณญาณเกี่ยวกับการวิจัยทางวิทยาศาสตร์ของครูในโรงเรียนมัธยมศึกษาตอนต้น ผ่านการอ่านตัวอย่างบทความในหนังสือพิมพ์จากกลุ่มตัวอย่างเป็นครูคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 4 จำนวน 38 คน ผ่านการเขียนสะท้อนกระบวนการคิดอย่างมีวิจารณญาณตามแนวคิดของ Facione, P. A. (1990) โดยงานวิจัยสังเคราะห์กระบวนการคิดอย่างมีวิจารณญาณพร้อมทั้งทักษะย่อยของการคิดอย่างมีวิจารณญาณแต่ละขั้นตอนไว้ 5 ขั้นตอน พร้อมตัวอย่างงานในการประเมิน ได้แก่ 1) การทำความเข้าใจ (comprehending) ประกอบด้วย การระบุใจความสำคัญของเรื่อง การจัดระเบียบข้อมูล และการชี้แจง/ระบุข้อมูล 2) การเรียบเรียงข้อความ (making connections) ประกอบด้วย การตรวจสอบความเชื่อมโยงระหว่างความคิด 3) การอนุมาน (inferring) ประกอบด้วย การตรวจสอบหลักฐาน การเสนอทางเลือก และการสร้างข้อสรุป 4) การ

วิจารณ์ (critiquing) ประกอบด้วย การตรวจสอบข้อมูลที่บิดเบือน และความน่าเชื่อถือของข้อมูล และ 5) การสะท้อนตนเอง (self-reflecting) ประกอบด้วย การเขียนแสดงถึงจุดแข็งและจุดอ่อนของตัวเอง และการปรับปรุงแก้ไข ผลการวิจัย พบว่า ครูส่วนใหญ่มีปัญหาในการแยกแยะข้อเท็จจริงและข้อคิดเห็น รวมไปถึงการเขียนหาเหตุผลมาสนับสนุนข้อโต้แย้งของตนเองที่ไม่ได้อยู่ในบทความ เนื่องจากครูมีความรู้ที่ไม่เพียงพอ ดังนั้นครูต้องศึกษาหาความรู้เพิ่มเติมเพื่อเป็นข้อมูลที่ช่วยในการคิดอย่างมีวิจารณญาณในบริบทของการศึกษาต่อไป

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องข้างต้น พบว่า โดยส่วนใหญ่ถ้าเป็นแบบสอบข้อเขียน เพื่อวัดการคิดวิจาร์ณญาณก็จะเป็นเพียงการให้สถานการณ์มา แต่ไม่ได้จำแนกขั้นตอนการคิดวิจาร์ณญาณเพื่อวัดความคิดในแต่ละขั้น เป็นเพียงแค่การวัดการคิดวิจาร์ณญาณโดยภาพรวม และอีกหลายงานวิจัยที่วัดการคิดวิจาร์ณญาณในรูปแบบของแบบสอบปรนัย ซึ่งอาจไม่ได้วัดการคิดวิจาร์ณญาณได้อย่างแท้จริงเนื่องจากมีโอกาสในการเดาสอง ทั้งนี้ยังไม่พบงานวิจัยที่สังเคราะห์ขั้นตอนการคิดวิจาร์ณญาณมาใช้สร้างเครื่องมือในรูปแบบของ แบบสอบข้อเขียนหรือแบบสอบอัตนัยซึ่งเป็นแบบสอบที่ส่งเสริมการคิดขั้นสูงของผู้เรียนได้ ผู้วิจัยจึงได้สนใจที่จะพัฒนาเครื่องมือในการวัดการคิดวิจาร์ณญาณโดยใช้แบบสอบอัตนัย

ตอนที่ 2 แบบสอบอัตนัยประยุกต์

2.1 ความหมายของแบบสอบอัตนัยประยุกต์

แบบสอบอัตนัยประยุกต์ หรือแบบสอบเอ็มอีคิว (Modified Essay Question test: MEQ) หมายถึง แบบสอบความเรียงหรือแบบสอบอัตนัย (essay question) ที่มีลักษณะคำถามเป็นปลายเปิด (open-ended question) โดยบรรยายเหตุการณ์ที่เกิดขึ้นจริงตามลำดับเหตุการณ์ที่ละขั้นตอน จะไม่นำเสนอเหตุการณ์ทั้งหมดเพียงครั้งเดียวต่อเนื่องเหมือนแบบสอบความเรียง ผู้ตอบต้องอาศัยข้อมูลจากเหตุการณ์นั้นมาหาคำตอบ (นภา หลิมรัตน์, 2551)

แบบสอบอัตนัยประยุกต์ถูกสร้างขึ้นเพื่อมาใช้แทนแบบสอบอัตนัยแบบดั้งเดิมที่มีความเป็นปรนัยในการตรวจให้คะแนนค่อนข้างต่ำ และวัดความรู้ที่อยู่ในระดับท่องจำหรือเข้าใจพื้นฐานเท่านั้น แบบสอบอัตนัยประยุกต์พัฒนาขึ้นโดย Hodgkin และ Knox คณะกรรมการตรวจสอบแห่งราชวิทยาลัยเวชปฏิบัติทั่วไป สหราชอาณาจักร ในปี ค.ศ. 1965 เพื่อเป็นเครื่องมือวัดทักษะการแก้ปัญหา (problem solving skill) ซึ่งจัดเป็นทักษะการคิดขั้นสูง (higher order cognitive thinking) ของแพทย์ที่จบการศึกษาในระดับปริญญาตรีที่ต้องการสอบเข้าเป็นสมาชิกของสมาคม ต่อมา Feletti (1980) แห่งมหาวิทยาลัยนิวคาสเซิล ได้ปรับปรุงแบบสอบอัตนัย

ประยุกต์โดยเพิ่มการกำหนดเวลาในการทำข้อสอบแต่ละข้อ และผู้สอบต้องทำข้อสอบทุกข้อให้เสร็จภายในเวลาที่กำหนด นอกจากนี้ Feletti ยังสร้างคำตอบเพื่อใช้ในการให้คะแนนและกำหนดคะแนนผ่านหรือระดับความสามารถขั้นต่ำที่ผ่านด้วย (Feletti, 1980; กนกวรรณ ศรีรักษา, 2554; สุโรยา หมั่นหมัด, 2549 อ้างถึงใน มาฆพันธ์ อ๋านาคิล และกมลวรรณ ตังธนทานนท์; 2557 และ สุภชิต ผดุงผล และกมลวรรณ ตังธนทานนท์, 2559; เชิดศักดิ์ ไอรณรัตน์, 2558)

2.2 ลักษณะของแบบสอบอัตนัยประยุกต์

แบบสอบอัตนัยประยุกต์ ได้มีนักวิชาการหลายท่านได้กล่าวถึงการพัฒนาลักษณะของรูปแบบแบบสอบอัตนัยประยุกต์ ดังนี้

Hodgkin and Knox (1980) (ฐิติมา ฐิติรุ่งเรือง, 2544 อ้างถึงใน 2549) ได้กำหนดรูปแบบของแบบสอบอัตนัยประยุกต์ ดังนี้

1. แบบสอบที่มีลักษณะเป็นการศึกษาเฉพาะกรณี (Case study) เพื่อผู้ออกข้อสอบเลือกมาให้เหมาะสมกับเนื้อหาที่ต้องการออกข้อสอบ และครอบคลุมวัตถุประสงค์ที่ตั้งไว้
2. บรรยายเหตุการณ์ที่ปรากฏขึ้นจริงตามลำดับ
3. สถานการณ์ที่น่าเสนอไม่เสนอทั้งหมดในครั้งเดียว แต่ให้ข้อมูลเป็นตอน ๆ
4. ลักษณะคำถามเป็นคำถามปลายเปิด ที่ผู้สอบต้องหาคำตอบโดยอาศัยข้อมูลในสถานการณ์ที่กำหนดให้มาตอบคำถามข้อนั้น
5. ข้อมูลที่ให้เพิ่มเติม อาจเกี่ยวข้องกับคำถามมากหรือน้อย หรือบางส่วนอาจไม่เกี่ยวข้อง ผู้สอบต้องตัดสินใจพิจารณาเลือกข้อมูลมาวิเคราะห์
6. วิธีการทำแบบสอบให้ผู้สอบใช้ข้อมูลในหน้านั้น ๆ มิให้ย้อนกลับไปแก้ไขข้อที่ทำไปแล้วหรือเปิดไปดูข้อมูลข้างหน้าที่ให้เพิ่มเติม
7. การตรวจให้ค่าคะแนนเป็น 0, 1, 2, 3, ... ในแต่ละข้อ หรืออาจให้คะแนนติดลบเมื่อพบว่าคำตอบมีความเสียหายแก้ไขซ้ำ
8. การกำหนดเกณฑ์การให้คะแนนในแต่ละข้อ โดยให้ผู้เชี่ยวชาญ 10 – 12 คน ทำการตอบข้อสอบเพื่อรวบรวมคำตอบที่เป็นไปได้ การให้ผู้เชี่ยวชาญที่มีประสบการณ์สูง ทำให้แบบสอบอัตนัยประยุกต์ ฉบับนี้มีค่าความเที่ยงตรงสูง

Feletti (1980) (ฐิติมา ฐิติรุ่งเรือง, 2544 อ้างถึงใน สุโรยา หมั่นหมัด, 2549) ได้พัฒนาแบบสอบอัตนัยประยุกต์จาก Hodgkin และ Knox ดังนี้

1. การกำหนดเวลาในการทำข้อสอบแต่ละข้อ ซึ่งผู้สอบต้องปรับเวลาในการทำข้อสอบแต่ละข้อให้เหมาะสม และต้องทำข้อสอบให้เสร็จทุกข้อในเวลารวมที่กำหนดให้

2. การสร้างคำตอบเพื่อใช้ในการตรวจให้คะแนน โดยรวบรวมคำตอบทั้งหมดจากผู้เชี่ยวชาญ มีการกำหนดคะแนนให้ผ่าน หรือระดับความสามารถขั้นต่ำที่ให้ผ่านไปได้ (Mandatory Level of Competence: MLC) ของแต่ละข้อคำถามไว้ เช่น คำตอบของข้อคำถามที่ทำในข้อนั้นมีจำนวน 16 ข้อ ถ้าผู้ตอบตอบได้ 6 ข้อ ข้อใดก็ได้ในจำนวนคำตอบทั้งหมดจึงผ่านเกณฑ์ให้ 1 คะแนน ถ้าตอบได้ไม่ถึง 6 ข้อ ให้ 0 คะแนน

3. จำนวนข้อของแบบสอบอัตนัยประยุกต์ ควรมีตั้งแต่ 5 – 35 ข้อ

แนวคิดของนักวิชาการข้างต้นได้สอดคล้องกับลักษณะของแบบสอบอัตนัยประยุกต์ จากงานวิจัยของ สุโรยา หมั่นหมัด (2549) มาฆพันธ์ อ๋านาคิล และกมลวรรณ ตังธกานนท์ (2557) สุภชิต ผดุงผล และกมลวรรณ ตังธกานนท์ (2559) ที่อธิบายแบบสอบอัตนัยประยุกต์ (MEQ) ว่าเป็นแบบสอบปลายเปิดที่ใช้ศึกษาเฉพาะกรณี (case study) มีลักษณะเป็นชุดคำถาม โดยแยกชุดคำถามเป็นข้อคำถามละ 1 หน้า เพื่อนำเสนอเหตุการณ์ตามลำดับเหตุการณ์ที่เกิดขึ้นตามสภาพจริง และมีการแทรกคำถามเป็นระยะ ๆ ในแต่ละข้อคำถามจะแบ่งออกเป็น 3 ส่วน คือ ส่วนที่แสดงเวลาที่ใช้ ข้อความหรือสถานการณ์ ข้อคำถาม และส่วนที่ว่างให้เขียนแสดงคำตอบไม่มีการอนุญาตให้เปิดย้อนกลับไปแก้ไขข้อสอบข้อที่ทำไปแล้วหรือกลับไปดูข้อมูลที่ผ่านมาได้ สำหรับเกณฑ์การให้คะแนนจะถูกกำหนดโดยผู้เชี่ยวชาญ ดังภาพที่ 1

ภาพที่ 1 ตัวอย่างลักษณะสถานการณ์และข้อคำถามของแบบสอบอัตนัยประยุกต์ (มาฆพันธ์ อ๋านาคิล และกมลวรรณ ตังธกานนท์, 2557)

เวลาสำหรับทำข้อ นี้ 5 นาที	เหตุการณ์ที่ 1 ข้อคำถามที่ 1	ชื่อ - สกุล โรงเรียน
<p>สมเกียรติและสมศรีมีลูกชายด้วยกัน 1 คน ชื่อแบงค์ อายุ 16 ปี เรียนอยู่ชั้นมัธยมศึกษาปีที่ 4 ทั้งสองคนประกอบอาชีพค้าขาย โดยปกติแบงค์เป็นเด็กน่ารัก ตั้งใจเรียน และทุกวันหลังเลิกเรียนแบงค์จะกลับมาช่วยพ่อแม่ขายของอยู่เสมอ แต่ตั้งแต่แบงค์เข้าเรียนต่อในระดับชั้นมัธยมศึกษาปีที่ 4 ที่โรงเรียนแห่งใหม่ สมศรีสังเกตเห็นว่าแบงค์มีร่างกายซูบผอม ผิวหนังและริมฝีปากแห้ง และเปลี่ยนเป็นคนเจ้าอารมณ์ โมโหง่าย สมศรีจึงเล่าเรื่องทั้งหมดให้สมเกียรติฟัง</p>		

ข้อคำถามที่ 1 จากสถานการณ์ที่กำหนดให้ นักเรียนคิดว่าปัญหาของแบงค์คืออะไร

คำตอบ

.....

.....

.....

.....

.....

.....

.....

คำเตือน ก่อนทำข้อคำถามข้อต่อไป นักเรียนต้องตรวจสอบให้แน่ใจว่าทำข้อคำถามนี้เรียบร้อยแล้ว เพราะจะไม่สามารถกลับมาแก้ไขได้อีก

จากแนวคิดของนักวิชาการที่กล่าวไว้ข้างต้น ผู้วิจัยได้พัฒนาแบบสอบอัตนัยประยุกต์ วัดทักษะการคิดมีวิจารณญาณทางคณิตศาสตร์ตามหลักสูตรฐานสมรรถนะ โดยใช้แนวคิดของ Hodgkin Knox และ Feletti มีลักษณะแบบสอบอัตนัยประยุกต์ ดังนี้

1. แบบสอบอัตนัยประยุกต์จะนำเสนอเป็นเหตุการณ์ ในแต่ละเหตุการณ์จะมีข้อมูลให้เพิ่มเติมตามลำดับ และมีข้อคำถามจำนวน 5 ข้อคำถาม โดยจัดวางไว้หน้าละ 1 ข้อคำถาม
2. ข้อมูลที่ให้เพิ่มเติมอาจเกี่ยวข้องกับคำถามมากหรือน้อย ผู้ตอบต้องตัดสินใจพิจารณาข้อมูลเพื่อนำมาวิเคราะห์และเขียนแสดงคำตอบ
3. ลักษณะคำถามเป็นคำถามปลายเปิด และมีการกำหนดเกณฑ์การให้คะแนนในแต่ละข้อคำถามโดยให้ผู้เชี่ยวชาญเป็นผู้ตรวจสอบ
4. วิธีการทำแบบสอบอัตนัยประยุกต์ ให้ผู้สอบใช้ข้อมูลหน้านั้น ๆ มิให้ย้อนกลับไปแก้ไขข้อที่ทำไปแล้ว หรือเปิดไปดูข้อมูลข้างหน้าที่ให้เพิ่มเติม โดยเริ่มทำข้อสอบตั้งแต่เหตุการณ์ที่ 1 ข้อคำถามที่ 1 เมื่อทำเสร็จให้ส่งข้อสอบทั้งที่
5. กำหนดเวลาในการทำข้อสอบแต่ละเหตุการณ์อยู่ด้านบนของแบบสอบ เหตุการณ์ละ 20 นาที โดยผู้สอบต้องจัดสรรเวลาในการทำแบบสอบเอง และทำข้อสอบให้ครบทุกข้อภายในเวลาที่กำหนด

2.3 ขั้นตอนการสร้างแบบสอบอัตโนมัติประยุกต์

สำหรับขั้นตอนการสร้างแบบสอบอัตโนมัติประยุกต์โดยภาพรวม มีขั้นตอนในการสร้างและพัฒนาเหมือนกับแบบสอบโดยทั่วไป แตกต่างกันที่รายละเอียดบางขั้นตอน จากการศึกษา งานวิจัย สุไรยา หมั่นหมัด (2549) มีขั้นตอนในการสร้างแบบสอบอัตโนมัติประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ดังนี้

1. กำหนดจุดมุ่งหมายในการพัฒนาแบบสอบ
2. ศึกษาทฤษฎี หลักสูตร และเอกสารงานวิจัยที่เกี่ยวข้อง เพื่อวิเคราะห์ขอบเขตของเนื้อหาในการพัฒนา โดยศึกษาตามหลักสูตรการศึกษาขั้นพื้นฐาน มาตรฐานการเรียนรู้และตัวชี้วัดในกลุ่มสาระวิทยาศาสตร์ พร้อมทั้งศึกษากระบวนการแก้ปัญหาทางวิทยาศาสตร์
3. กำหนดกรอบปัญหา ตามขอบเขตของเนื้อหา และวิธีการแก้ปัญหาทางวิทยาศาสตร์
4. สร้างเครื่องมือตามกรอบปัญหา โดยใช้แบบสอบอัตโนมัติประยุกต์เป็นเครื่องมือในการวัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์
5. ตรวจสอบคุณภาพขั้นต้นของแบบสอบอัตโนมัติประยุกต์และเกณฑ์การตรวจให้คะแนน โดยให้ ผู้เชี่ยวชาญที่มีประสบการณ์ทางการสอนวิทยาศาสตร์อย่างน้อย 5 ปี และผู้เชี่ยวชาญที่มีประสบการณ์ทางการวัดผลทางการศึกษา เป็นผู้ตรวจสอบความตรงของเหตุการณ์และข้อคำถาม
6. ทดลองใช้แบบทดสอบ (Try out) และตรวจสอบคุณภาพหลังการทดลอง โดยหาคุณภาพในด้านความยากง่าย ค่าอำนาจจำแนก ความตรง และความเที่ยงของแบบทดสอบ
7. จัดทำคู่มือการใช้แบบทดสอบ และจัดพิมพ์แบบทดสอบเป็นรูปเล่ม

นอกจากนี้ พวงแก้ว ปุณยนก (2531) อ้างถึงใน สุภชิต ผดุงผล และกมลวรรณ ตั้งธนานนท์ (2559) เสนอขั้นตอนการสร้างโมเดลคำตอบ โดยนำแบบสอบที่สร้างขึ้นไปให้ผู้ทรงคุณวุฒิ หรือผู้เชี่ยวชาญตรวจสอบความถูกต้องและทดลองทำข้อสอบ เพื่อนำคำตอบมาสร้างเป็นโมเดลคำตอบทั้งหมด โดยจะเลือกคำตอบที่ผู้ทรงคุณวุฒิเห็นตรงกันเกินครึ่งหนึ่งมาเป็นคำตอบในการให้คะแนน ส่วนคำตอบที่ผู้เชี่ยวชาญมีความเห็นไม่ตรงกันจะนำมาอภิปรายเพื่อแก้ไขปรับปรุงหรือตัดทิ้งไป และหาเวลาที่เหมาะสมในการทำข้อสอบแต่ละข้อ

จากแนวคิดของนักวิชาการที่กล่าวไว้ข้างต้น การสร้างแบบสอบอัตโนมัติจึงเป็นขั้นตอนที่ส่งผลต่อความน่าเชื่อถือของแบบสอบหากมีขั้นตอนในการสร้างที่ถูกต้อง ผู้วิจัยจึงได้กำหนดขั้นตอนการสร้างแบบสอบอัตโนมัติประยุกต์ ดังนี้ 1) กำหนดจุดมุ่งหมาย 2) ศึกษาเอกสารที่

เกี่ยวข้อง 3) กำหนดขอบเขตของแบบสอบ 4) เลือกเหตุการณ์ที่จะนำมาสร้างสถานการณ์และข้อคำถาม 5) กำหนดรูปแบบของแบบสอบอัตโนมัติประยุกต์ ในการวัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ 6) สร้างโมเดลคำตอบ เกณฑ์การให้คะแนน การตรวจ ให้คะแนน และการบันทึกคะแนน 7) ตรวจสอบความตรง 8) ทดลองใช้ (try out) 9) คัดเลือกข้อสอบที่อยู่ในเกณฑ์ และปรับปรุงแก้ไข 10) นำแบบสอบอัตโนมัติประยุกต์ที่พัฒนาแล้วไปใช้จริง 11) วิเคราะห์ข้อสอบ และหาคุณภาพ 12) สร้างเกณฑ์ และ 13) สร้างคู่มือการใช้ และจัดทำรูปเล่ม

2.4 ขั้นตอนการดำเนินการจัดสอบแบบสอบอัตโนมัติประยุกต์

สำหรับการดำเนินการจัดสอบ เนื่องด้วยแบบสอบอัตโนมัติประยุกต์มีลักษณะเฉพาะในการนำเสนอเหตุการณ์ที่ละเหตุการณ์อย่างต่อเนื่อง โดยแยกข้อสอบออกเป็นเหตุการณ์ละ 1 แผ่น และเมื่อผู้สอบทำข้อสอบเสร็จก็จะต้องคว่ำแบบสอบลง หรือวางไว้ได้ของใส่แบบสอบ แล้วจึงสามารถหยิบข้อคำถามถัดไปขึ้นมาทำได้จากการศึกษางานวิจัยที่เกี่ยวข้องกับแบบสอบอัตโนมัติประยุกต์ มีผู้วิจัยได้พัฒนาวิธีการจัดสอบด้วยวิธีอื่น ๆ ที่น่าสนใจ ดังนี้

การจัดสอบด้วยการใช้คอมพิวเตอร์ โดยข้อคำถามจะเสนอบนจอตามเวลาที่กำหนด ข้อดีคือ ป้องกันผู้สอบกลับไปแก้ไขคำตอบในข้อก่อนหน้า ช่วยประหยัดเวลา และงบประมาณ สามารถควบคุมเวลาที่ใช้ในการสอบได้ดี แต่ข้อจำกัดคือ ถ้าผู้สอบไม่เคยทำแบบสอบผ่านคอมพิวเตอร์ จะทำให้เกิดความประหม่า และส่งผลต่อความคลาดเคลื่อนของคะแนน และในบางครั้งมักจะประสบปัญหาเรื่องความไม่เพียงพอของจำนวนเครื่องคอมพิวเตอร์ ทำให้ไม่สามารถจัดสอบพร้อมกันในครั้งเดียวได้

การนำเครื่องฉายภาพ (visualizer) และเครื่องฉายภาพข้ามศีรษะ (overhead projector) มาประยุกต์ใช้ในการจัดสอบ โดยผู้คุมสอบจะเขียนคำถามหรือแสดงข้อคำถามลงบนกระดาน หรือแผ่นใส จากนั้นผู้สอบจะตอบคำถามลงในกระดานคำตอบ ข้อดีคือ ใช้เวลาในการจัดเตรียมแบบสอบน้อย ประหยัดเวลาและงบประมาณ สามารถควบคุมเวลาที่ใช้ในการสอบได้ดี เช่นเดียวกับการจัดสอบด้วยการใช้คอมพิวเตอร์ แต่ข้อจำกัดคือ ถ้ามีผู้สอบจำนวนมาก จะมีข้อจำกัดเรื่องจำนวนอุปกรณ์

ดังนั้นการจัดสอบแบบสอบอัตโนมัติประยุกต์รูปแบบใดนั้น ต้องคำนึงถึงวัตถุประสงค์ของการสอบ จำนวนผู้สอบ ความพร้อมของผู้จัดสอบ รวมถึงวัสดุอุปกรณ์ต่าง ๆ ที่จะนำมาใช้ในการจัดสอบ ผู้วิจัยจึงจะดำเนินการจัดสอบด้วยการใช้แบบสอบอัตโนมัติประยุกต์ในลักษณะของกระดาน เนื่องจากจำนวนผู้สอบ มีปริมาณมาก ทำให้วัสดุอุปกรณ์ต่าง ๆ ที่จะนำมาใช้ในการจัดสอบอาจมี

จำนวนไม่เพียงพอต่อจำนวนผู้สอบ (ดวงมณี หลาคำแดง, 2544; นงรักษ์ ไทธานี, 2542; สุพัฒตา ภูสอดสี, 2553 อ้างถึงใน มาฆพันธ์ อานาคิล และกมลวรรณ ตั้งธนานนท์, 2557)

ตอนที่ 3 คุณภาพของการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์

3.1 คุณภาพของแบบสอบอัตนัยประยุกต์

การตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ ที่เป็นลักษณะแบบสอบปลายเปิดเป็นเครื่องมือที่เหมาะสมกับการวัดทักษะการคิดขั้นสูงของผู้เรียนได้ดี ดังนั้นจึงมีขั้นตอนในการสร้างแบบสอบที่เป็นระบบและต้องอาศัยการตรวจให้คะแนนอย่างมีคุณภาพมากที่สุดก่อนจะนำแบบสอบไปใช้ จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการหาคุณภาพของแบบสอบอัตนัยประยุกต์มีนักวิชาการหลายท่านได้ให้รายละเอียดไว้ ดังนี้

1. การตรวจสอบคุณภาพรายข้อ

แบบสอบอัตนัยประยุกต์จะมีลักษณะเฉพาะมากกว่าการสร้างแบบสอบความเรียงทั่วไป ดังเช่น Knox (1980) อ้างถึงใน สุโรยา หมั่นหมัด (2549) ที่ได้กำหนดรูปแบบของแบบสอบอัตนัยประยุกต์ไว้ว่าเป็นแบบสอบปลายเปิดที่ใช้การศึกษาเฉพาะกรณี มีบรรยายเหตุการณ์ที่เกิดขึ้นจริงตามลำดับของเหตุการณ์นั้น ๆ โดยเหตุการณ์ที่น่าเสนอจะไม่เสนอทั้งหมดเพียงครั้งเดียว แต่จะเสนอทีละตอนตามลำดับเหตุการณ์ ดังนั้นการตรวจสอบคุณภาพรายข้อจึงมีความจำเป็นอย่างมาก เพราะจะสามารถวัดได้ว่าข้อสอบแต่ละข้อมีความเหมาะสมต่อความรู้ความสามารถของผู้เรียนหรือไม่ มีความสามารถในการจำแนกผู้สอบและวัดคุณภาพของผู้สอบได้ว่าอยู่ในระดับใด โดยใช้สูตรการวิเคราะห์ข้อสอบของแบบสอบความเรียงแบบอิงกลุ่ม ทั้งค่าความยาก (p) และค่าอำนาจจำแนก (r) ดังที่ไซติกา ภาษีผล (2559) กล่าวไว้ ดังนี้

$$p = \frac{\sum X_H + \sum X_L}{I(N_H + N_L)}$$

$$r = \frac{\sum X_H - \sum X_L}{I(N_H \text{ or } N_L)}$$

เมื่อ $\sum X_H$ คือ ผลรวมคะแนนรายข้อแต่ละคนในกลุ่มสูง

$\sum X_L$ คือ ผลรวมคะแนนรายข้อแต่ละคนในกลุ่มต่ำ

N_H คือ จำนวนคนในกลุ่มสูง

N_L คือ จำนวนคนในกลุ่มต่ำ

I คือ คะแนนเต็มของข้อนั้น ๆ

2. การตรวจสอบความตรงของแบบสอบอัตนัยประยุกต์

การสร้างแบบสอบอัตนัยประยุกต์ให้มีคุณภาพควรมีการวางแผนให้ครอบคลุมพฤติกรรมการเรียนรู้ และ มีการกำหนดเกณฑ์การให้คะแนนที่ชัดเจน (พิชิต ฤทธิจรูญ, 2552) ดังนั้นผู้สร้างแบบสอบต้องมีการตรวจสอบความตรง (validity) ของแบบสอบเพื่อสร้างความน่าเชื่อถือของแบบสอบให้สามารถนำไปใช้ได้อย่างมีประสิทธิภาพ ซึ่งแบบสอบที่สร้างขึ้นควรวัดเนื้อหาได้ครอบคลุม และเป็นตัวแทนของเนื้อเรื่องหรือประสบการณ์ที่มุ่งวัด (ศิริชัย กาญจนวาสี, 2563) โดยมาฆพันธ์ อ๋านาคิล และกมลวรรณ ตังธนกานนท์ (2557) ลักษณะมี จันทราช (2544) สมคิด เทียรพิสุทธิ (2550) สายหมอก ขุนศักดิ์ (2543) อามาลา สารชาติ (2548) อ้างถึงใน สุภชิต ผดุงผล (2559) ได้ใช้การตรวจสอบความตรงเชิงเนื้อหา IOC (Index of item Objective Congruence) โดยให้ผู้เชี่ยวชาญตรวจสอบความเหมาะสมและลงความเห็นว่าคุณสมบัติของแต่ละข้อมุ่งวัดตามเนื้อหาและจุดประสงค์หรือไม่ และอยู่ในระดับใด

3. การตรวจสอบความเที่ยงของแบบสอบอัตนัยประยุกต์

สำหรับการออกแบบเงื่อนไขการทดสอบที่ต่างกัน ประกอบด้วย จำนวนเหตุการณ์ จำนวนข้อคำถาม และจำนวนผู้ประเมิน ส่งผลต่อค่าความเที่ยงของแบบสอบอัตนัยประยุกต์ ดังนั้นการวิเคราะห์ความน่าเชื่อถือของผลการวัด หรือความเที่ยงในสถานการณ์หรือเงื่อนไขต่าง ๆ ของการวัดจึงใช้แนวคิดพื้นฐานของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด จากการศึกษาเอกสารและงานวิจัย พบว่า มาฆพันธ์ อ๋านาคิล และกมลวรรณ ตังธนกานนท์ (2557) ออกแบบการวัดแบบ one-facet design โดยมีรูปแบบการวัดเป็น $p \times 1$ ชนิสรา สงวนไว้ และศิริชัย กาญจนวาสี (2558) ออกแบบการวัดแบบ two-facet design โดยมีรูปแบบการวัดเป็น $p \times i \times r$ และ $p \times (i:r)$ สุภชิต ผดุงผล และกมลวรรณ ตังธนกานนท์ (2559) ออกแบบการวัดแบบ three-facet design โดยมีรูปแบบการวัดเป็น $p \times r \times (i:o)$ ซึ่งงานวิจัยที่กล่าวมาข้างต้นใช้การวิเคราะห์ความเที่ยงของแบบสอบอัตนัยประยุกต์ โดยโปรแกรมสำเร็จรูป EduG และเปรียบเทียบค่าความเที่ยงค่าสัมประสิทธิ์การสรุปอ้างอิง G-Coefficient

4. การตรวจสอบความเที่ยงของผู้ประเมิน

แบบสอบอัตนัยประยุกต์สำหรับการวัดทักษะการคิดวิจารณ์ของนักเรียนซึ่งเป็นทักษะการคิดขั้นสูงควรจะเป็นแบบสอบที่คุณสมบัติความคงเส้นคงวาคะแนนที่ได้จากแบบสอบ ที่ไม่ว่าจะเป็นการทดสอบในครั้งใดเวลาใด ผลการทดสอบของผู้สอบคนเดิมที่ทำการทดสอบ โดยใช้แบบสอบอัตนัยประยุกต์ฉบับนี้ก็ควรได้คะแนนที่ใกล้เคียงของเดิม (ศิริชัย กาญจนวาสี, 2563) จะเห็นได้ว่าการตรวจให้คะแนนย่อมส่งผลต่อความเชื่อมั่นของการให้คะแนนเป็นอย่างมาก

แม้จะมีเกณฑ์การให้คะแนนที่แน่นอนและชัดเจน แต่ถ้าใช้ผู้ตรวจคนเดียวตรวจให้คะแนน อาจเกิดความลำเอียงในการตรวจให้คะแนนได้ (จิรายุ เกาวโท และคณะ, 2559) ส่วนใหญ่แบบสอบอัตนัย ประยุกต์จะใช้ผู้ตรวจตั้งแต่ 2 คนขึ้นไป ตรวจสอบความเที่ยงระหว่างผู้ประเมิน (inter rater reliability) โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation) ที่เป็นการตรวจสอบความสอดคล้องระหว่างผลการประเมินด้วยเครื่องมือการประเมินเดียวกันจากผู้ประเมินที่มากกว่า 1 คน (กมลวรรณ ตังธนากานนท์, 2563) ควบคู่กับการหาค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) ของ McGraw and Wong (1996) โดยเลือกใช้โมเดลแบบ two-way mixed-effects model หรือโมเดล ICC(3,2) เนื่องจากเป็นการประเมินความน่าเชื่อถือระหว่างผู้ประเมินที่มาจากกลุ่มผู้ประเมินเดียวกันโดยเลือกผู้ประเมินแบบเจาะจง โดยใช้ผู้ประเมิน 2 คน นอกจากนี้ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันยังสามารถตรวจสอบความเที่ยงในการตรวจให้คะแนนของผู้ตรวจคนเดียว (intra rater reliability) ได้อีกด้วย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการหาคุณภาพของแบบสอบอัตนัย ประยุกต์ ผู้วิจัยจะทำการตรวจสอบคุณภาพของแบบสอบอัตนัยประยุกต์ ดังนี้ 1) การตรวจสอบคุณภาพรายข้อ โดยใช้การหาค่าความยากและอำนาจจำแนก ที่ใช้สูตรการวิเคราะห์ข้อสอบของแบบสอบความเรียงแบบอิงกลุ่ม และวิเคราะห์ความเที่ยง (alpha coefficient) ของแบบสอบแต่ละฉบับ 2) การตรวจสอบความตรงเชิงเนื้อหา IOC โดยให้ผู้เชี่ยวชาญตรวจสอบความเหมาะสมของข้อสอบแต่ละข้อกับจุดประสงค์และเนื้อหาที่ตั้งไว้ และ 3) การตรวจสอบความเที่ยงในการตรวจให้คะแนนของผู้ตรวจ โดยตรวจสอบความเที่ยงภายในผู้ตรวจคนเดียว (intra rater reliability) ด้วยสถิติค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation) และความเที่ยงระหว่างผู้ประเมิน (inter rater reliability) ด้วยสถิติค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation) ควบคู่กับค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) ที่วิเคราะห์จากโมเดลอิทธิพลผสมแบบสองทาง (two-way mixed-effects model) ที่ระดับความเชื่อมั่น 95% สำหรับการวิเคราะห์ความเที่ยงของแบบสอบอัตนัยประยุกต์ภายใต้เหตุการณ์ที่ต่างกัน ประยุกต์กับทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (G-Theory) โดยใช้ค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient)

3.2 วิธีการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์

สุไรยา หมั่นหมัด (2544) ได้กล่าวถึงวิธีการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ไว้ 2 วิธี ดังนี้

1. วิธีการตรวจให้คะแนนตามแบบ Knox

สำหรับแบบทดสอบที่เป็นกรณีศึกษาเฉพาะ เพื่อผู้ออกข้อสอบสามารถเลือกมาให้เหมาะสมกับเนื้อหาที่ต้องการออกข้อสอบและครอบคลุมวัตถุประสงค์ที่ตั้งไว้ ต้องมีการบรรยายเหตุการณ์ที่ปรากฏขึ้นจริงตามลำดับ โดยที่สถานการณ์ที่น่าเสนอจะไม่แสดงรายละเอียดทั้งหมดในครั้งเดียว แต่จะเป็นการให้ข้อมูลเป็นตอน ๆ ลักษณะคำถามเป็นปลายเปิด ที่ผู้ตอบต้องหาคำตอบโดยอาศัยข้อมูลในสถานการณ์ที่กำหนดให้มาตอบคำถามข้อนั้น ข้อมูลที่ให้เพิ่มเติม อาจเกี่ยวข้องกับคำถามมากหรือน้อย หรือบางส่วนอาจไม่เกี่ยวข้อง ผู้ตอบต้องตัดสินใจพิจารณาเลือกข้อมูลมาวิเคราะห์ด้วยตนเอง ในขณะที่ทำแบบทดสอบกำหนดให้ผู้ตอบใช้ข้อมูลหน้านั้น ๆ มิให้ย้อนกลับไปแก้ไขข้อที่ทำไปแล้วเปิดไปดูข้อมูลข้างหน้าที่ให้เพิ่มเติม โดยการตรวจให้คะแนนจะเป็นระดับ 0, 1, 2, 3, ... ในแต่ละข้อ หรืออาจให้คะแนนติดลบเมื่อพบว่าคำตอบมีความเสียหายแก่วิชาชีพ โดยกำหนดเกณฑ์การให้คะแนนในแต่ละข้อจากผู้เชี่ยวชาญ 10 -12 คน ทำการตอบข้อสอบเพื่อรวบรวมคำตอบทั้งหมดที่เป็นไปได้ ซึ่งการใช้ผู้เชี่ยวชาญที่มีประสบการณ์สูง ทำให้แบบทดสอบอัตนัยประยุกต์มีความเที่ยงสูง ดังภาพที่ 1

2. วิธีการตรวจให้คะแนนตามแบบ Feletti

พัฒนาขึ้นจากวิธีการตรวจให้คะแนนตามแบบ Knox โดยมีการกำหนดเวลาในการทำข้อสอบแต่ละข้อ ซึ่งผู้ตอบต้องปรับเวลาในการทำข้อสอบแต่ละข้อให้เหมาะสม ต้องทำข้อสอบให้เสร็จทุกข้อในเวลารวมที่กำหนดให้ ซึ่งการสร้างคำตอบเพื่อใช้ในการตรวจให้คะแนนจะมาจากการรวบรวมคำตอบทั้งหมดจากผู้เชี่ยวชาญ แต่มีการกำหนดคะแนนให้ผ่านหรือระดับความสามารถขั้นต่ำที่ให้ผ่านได้ (Mandatory level of competence : MLC) ของแต่ละข้อคำถามไว้ เช่น คำตอบข้อคำถามที่ทำในข้อนั้นมีจำนวน 16 ข้อ ถ้าผู้ตอบสอบได้ 6 ข้อ ข้อใดก็ได้ในจำนวนคำตอบทั้งหมดจึงผ่านเกณฑ์ให้ 1 คะแนน ถ้าตอบได้ไม่ถึง 6 ข้อ ให้ 0 คะแนน และจำนวนข้อของแบบสอบอัตนัยประยุกต์ ควรมีตั้งแต่ 5 – 35 ข้อ ดังภาพที่ 2

ภาพที่ 2 ตัวอย่างวิธีตรวจให้คะแนน Knox และวิธีตรวจให้คะแนน3 Feletti (มาฆพันธ์ อำนาคิล และกมลวรรณ ตังชนกานนท์, 2557)

เหตุการณ์ที่ 1 แนวคำตอบข้อคำถามที่ 1 แบนด์มีร่างกายซูบผอม ผิวหนังและริมฝีปากแห้ง และเปลี่ยนเป็นคนเจ้าอารมณ์ โมโหง่าย		
เกณฑ์การตรวจให้คะแนนตามแบบของ Knox		
2 คะแนน	1 คะแนน	0 คะแนน
แบนด์มีร่างกายซูบผอม ผิวหนังและริมฝีปากแห้งและเปลี่ยนเป็นคนเจ้าอารมณ์ โมโหง่าย	แบนด์มีรูปร่างซูบผอม ผิวหนังและริมฝีปากแห้ง <u>หรือ</u> แบนด์เปลี่ยนเป็นคนเจ้าอารมณ์โมโหง่าย	ตอบคำถามไม่ถูกต้องตามแนวคำตอบ <u>หรือ</u> ไม่มีการตอบคำถาม
เกณฑ์การตรวจให้คะแนนตามแบบของ Feletti		
1 คะแนน	0 คะแนน	
แบนด์มีร่างกายซูบผอม ผิวหนังและริมฝีปากแห้งและเปลี่ยนเป็นคนเจ้าอารมณ์ โมโหง่าย	ตอบคำถามไม่ถูกต้องตามแนวคำตอบ <u>หรือ</u> ไม่มีการตอบคำถาม	

นักวิชาการหลายคนได้เสนอวิธีการตรวจให้คะแนนแบบสอบประเภทความเรียงที่นิยมมากที่สุด 2 วิธี คือ (Charles & Richard, 1990; Ebel, R. L., & Frisbie, D. A. 1972; Albert, O., 2003; Abatihun, A. S., 2013)

1. วิธีการให้คะแนนแบบองค์รวม (holistic method)

วิธีนี้คำตอบจะไม่ถูกแบ่งเป็นส่วน ๆ ต้องอาศัยความประทับใจโดยภาพรวมในการตรวจให้คะแนน ผู้ตรวจจะใช้วิธีการอ่านคำตอบอย่างรวดเร็วและใช้มาตรฐานบางอย่างในการกำหนดระดับของคำตอบ และเพื่อควบคุมความสม่ำเสมอหรือเพื่อให้การตรวจมีมาตรฐาน ก็จะมีการแบ่งกลุ่มคำตอบเป็นหลาย ๆ กอง ตามระดับคุณภาพที่ต่างกัน แล้วให้คะแนนตามลำดับคุณภาพอีกทีหนึ่ง ซึ่งการตรวจวิธีนี้จะมีประสิทธิภาพสำหรับข้อสอบที่มีจำนวนมาก เนื่องจากการตรวจจะทำได้อย่างรวดเร็ว ใช้เวลาน้อย ง่ายต่อการนำไปใช้งาน อย่างไรก็ตามแม้ว่าวิธีการตรวจ

ประเภทนี้จะทำให้การให้คะแนนมีความเที่ยงสูง แต่ยากต่อการให้ข้อมูลย้อนกลับแก่ผู้สอบเกี่ยวกับเหตุผลของการได้คะแนนในแต่ละข้อ

2. วิธีการให้คะแนนแบบวิเคราะห์ย่อย (analytic method)

วิธีนี้มีรูปแบบคำตอบที่ประกอบด้วยองค์ประกอบเฉพาะที่กำหนดไว้ล่วงหน้า โดยจะแยกออกเป็นตอนย่อย ๆ ตามองค์ประกอบที่สำคัญและมีการกำหนดการให้คะแนนในแต่ละตอน จากนั้นจะนำคะแนนของแต่ละตอนมารวมเป็นคะแนนที่ได้รับทั้งข้อ ซึ่งการกำหนดองค์ประกอบที่คาดหวังจะช่วยให้ผู้ตรวจให้คะแนนสามารถให้คะแนนได้อย่างคงที่

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องเกี่ยวกับวิธีการตรวจให้คะแนน พบว่า มีอีกหนึ่งวิธีการตรวจให้คะแนนที่พัฒนามาจากวิธีวิเคราะห์ย่อย นั่นก็คือ วิธีตรวจให้คะแนนแบบสองชั้น (double layer scoring rubric) ซึ่งเป็นเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) อีกประเภทหนึ่งที่ใช้ในการประเมินแต่ละด้าน มีการอธิบายลักษณะของงานในแต่ละระดับความสามารถ ประกอบด้วยรายการย่อย ๆ 2 ชั้น โดยชั้นที่ 1 เป็นเกณฑ์การให้คะแนนแยกองค์ประกอบที่อธิบายลักษณะของงานในข้อรายการย่อย ๆ ลดหลั่นตามระดับความสามารถ ส่วนชั้นที่ 2 เป็นการแปลงคะแนนผลรวมของคะแนนที่ได้ในระดับที่ 1 ออกมาให้เป็นไปตามช่วงคะแนนที่กำหนด จากณัฐณี ศิริโชติ และกมลวรรณ ตังธนกานนท์ (2559) ได้ออกแบบเครื่องมือในการประเมินตนเองและความสามารถทางวิทยาศาสตร์ที่มีต่อพัฒนาการปฏิบัติการทดลองของนักเรียนด้วยวิธีการให้คะแนนแบบรูปrikสองชั้น ซึ่งข้อดีของวิธีตรวจให้คะแนนรูปrikสองชั้น คือ ผู้สอนทราบถึงข้อบกพร่องของนักเรียนในทักษะย่อย ๆ ขององค์ประกอบที่ใช้ในการประเมินด้านนั้น ๆ มากกว่าวิธีการประเมินการให้คะแนนแบบรูปrikปกติ จึงกล่าวได้ว่าวิธีการให้คะแนนแบบรูปrikสองชั้น จะช่วยทำให้คะแนนมีความเป็นปรนัยขึ้นมากกว่าเกณฑ์การให้คะแนนแบบรูปrikอื่น ๆ ดังภาพที่ 3

ภาพที่ 3 ตัวอย่างวิธีตรวจให้คะแนนรูปrikสองชั้น (double layer scoring rubric) (ณัฐณี ศิริโชติ และกมลวรรณ ตั้งชนกานนท์, 2559)

ระดับชั้นที่ 1

ด้าน/ข้อรายการย่อย	ข้อรายการย่อย	ผลการประเมิน		
		2	1	0
1. เทคนิคการทดลอง 1.1 เลือกใช้อุปกรณ์และสารเคมีได้ถูกต้อง	1.1.1 ระบุชนิดของสารเคมีที่กำหนดให้ได้ถูกต้อง	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด
	1.1.2 เลือกประเภทของอุปกรณ์ได้เหมาะสมกับการใช้งาน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด
	1.1.3 เลือกขนาดภาชนะรองรับสารได้อย่างเหมาะสม	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด
	1.1.4 ติดตั้งอุปกรณ์ในการทดลองได้ถูกต้อง	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด
	1.1.5 ความพอดีของปริมาณสารที่ตวง/ชั่งไปใช้	<input type="checkbox"/> ตวง/ชั่งสารไปใช้ในปริมาณที่พอดี	<input type="checkbox"/> ตวง/ชั่งสารไปเกินปริมาณที่ใช้น้อยกว่า 1 ml	<input type="checkbox"/> ตวง/ชั่งสารไปเกินปริมาณที่ใช้ตั้งแต่ 1 ml ขึ้นไป
	คะแนนรวม = คะแนน คิดเป็น = คะแนน			
1.2 มีความปลอดภัยในการทดลอง	1.2.1 สวมแว่นตาขณะปฏิบัติการทดลอง	<input type="checkbox"/> สวมใส่ตลอดเวลา	<input type="checkbox"/> สวมใส่บางเวลา	<input type="checkbox"/> ไม่สวมใส่ตลอดเวลา
	1.2.2 ใส่เสื้อกาวน์ขณะปฏิบัติการทดลอง	<input type="checkbox"/> สวมใส่ตลอดเวลา	<input type="checkbox"/> สวมใส่บางเวลา	<input type="checkbox"/> ไม่สวมใส่ตลอดเวลา
	1.2.3 มีลำดับขั้นตอนที่ถูกต้องในการผสมสารเคมี	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด
	1.2.4 หันปากภาชนะที่มีสารเคมีไปทางผู้อื่น	<input type="checkbox"/> ไม่เกิน 1 ครั้ง	<input type="checkbox"/> มากกว่า 1 ครั้งแต่ไม่เกิน 3 ครั้ง	<input type="checkbox"/> มากกว่า 3 ครั้ง
	1.2.5 ใช้มือสัมผัส	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ด้าน/ข้อรายการย่อย	ข้อรายการย่อย	ผลการประเมิน		
		2	1	0
	สารเคมีโดยตรง	ไม่เกิน 1 ครั้ง	มากกว่า 1 ครั้งแต่ไม่เกิน 3 ครั้ง	มากกว่า 3 ครั้ง
คะแนนรวม = คะแนน คิดเป็น = คะแนน				

ภาพที่ 3 (ต่อ)

ระดับชั้นที่ 2 (คะแนนเต็ม 5 คะแนน)

ข้อที่	คะแนนในชั้นที่ 1 (ก่อนแปลงระดับคะแนน)	คะแนนในชั้นที่ 2 (หลังแปลงระดับคะแนน)
1.1 และ 1.2	9-10	5
	7-8	4
	5-6	3
	3-4	2
	0-2	1

จากงานวิจัยของมาฆพันธ์ อำนาคิล และกมลวรรณ ตังธนกานนท์ (2557) พบว่า วิธีการตรวจให้คะแนนวิเคราะห์ย่อยมีความเที่ยงมากกว่าวิธีตรวจให้คะแนน Knox และวิธีตรวจให้คะแนน Feletti ตามลำดับ เนื่องจากวิธีวิเคราะห์ย่อยมีความละเอียดเหมาะแก่การนำไปใช้และให้ข้อมูลย้อนกลับในการพัฒนาผู้สอบได้ แต่วิธีตรวจให้คะแนน Knox และวิธีตรวจให้คะแนน Feletti ยังไม่สามารถใช้เป็นข้อมูลย้อนกลับได้ถึงแม้ว่าวิธีตรวจให้คะแนน Knox จะมีความละเอียดมากกว่าวิธีตรวจให้คะแนน Feletti ก็ตาม ทั้งนี้สำหรับวิธีตรวจให้คะแนนรูปริกสองชั้นยังไม่พบการนำวิธีการตรวจนี้มาพัฒนาเพื่อใช้กับแบบสอบอัตนัยประยุกต์ ทั้งที่มีขั้นตอนที่ละเอียดเหมาะแก่การนำไปใช้และสามารถใช้เป็นข้อมูลย้อนกลับให้แก่ผู้สอบได้ด้วยเช่นกัน ผู้วิจัยจึงสนใจพัฒนาวิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) เพื่อเปรียบเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) และวิธีตรวจให้คะแนน Knox ซึ่งเป็นวิธีตรวจให้คะแนนสำหรับแบบสอบอัตนัยประยุกต์ในหลายงานวิจัยที่ผ่านมา (มาฆพันธ์ อำนาคิล และกมลวรรณ ตังธนกานนท์, 2557; ชนิศรา สงวนไว และ ศิริชัย กาญจนวาสี, 2558; สุภชิต ผดุงผล และกมลวรรณ ตังธนกานนท์, 2559) อีกทั้งวิธีตรวจให้คะแนนรูปริกสองชั้นนี้ยังไม่พบการนำวิธีตรวจให้คะแนนนี้มาเพื่อใช้กับแบบสอบอัตนัยประยุกต์

3.3 จำนวนเหตุการณ์ของแบบสอบอัตนัยประยุกต์

แบบสอบอัตนัยประยุกต์ (Modified Essay Question test: MEQ) พัฒนามาจากข้อสอบทางการแพทย์ที่เริ่มจากการให้เหตุการณ์ของผู้ป่วย แล้วให้ผู้สอบตอบคำถามเกี่ยวกับการแก้ปัญหาผู้ป่วยในสถานการณ์นั้น ๆ โดยไม่มีตัวเลือกให้ แล้วในเหตุการณ์จะมีคำถามเพิ่มเติมเพื่อเปิดเผยข้อมูลเกี่ยวกับผู้ป่วยมากขึ้นทีละน้อย ดังนั้นจะเห็นได้ว่าจำนวนเหตุการณ์ของแบบสอบอัตนัยประยุกต์มีผลต่อความเที่ยงของแบบสอบ คือ ถ้าแบบสอบมีความยาวมากเกินไป จะช่วยเพิ่มความเที่ยงของแบบสอบ แต่ก็จะทำให้ผู้สอบเกิดความเมื่อยล้าในการตอบ ในทางกลับกัน ถ้าจำนวนเหตุการณ์มีน้อยเกินไป ทำให้ไม่สามารถครอบคลุมองค์ความรู้ที่ต้องการประเมินได้ครบ และความเที่ยงของแบบสอบก็ลดลงเช่นกัน (เชิดศักดิ์ ไชรมณีรัตน์, 2558; ศิริชัย กาญจนวาสี, 2563)

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการใช้แบบสอบอัตนัยประยุกต์ของนักวิชาการหลายท่านได้ออกแบบจำนวนเหตุการณ์ในแบบสอบอัตนัยประยุกต์ไว้ ดังตาราง 2

ตาราง 2 การสังเคราะห์จำนวนเหตุการณ์ของแบบสอบอัตนัยประยุกต์

ผู้วิจัย	จุดประสงค์	จำนวน เหตุการณ์	จำนวนข้อ
Feletti (1980)	เพื่อทดสอบกระบวนการแก้ปัญหาทางการแพทย์และการจัดการผู้ป่วย	6	60 (เหตุการณ์ละ 10 ข้อ)
สุพัฒน์ ภูสอ ด สี (2553)	เพื่อวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์	10	50 (เหตุการณ์ละ 5 ข้อ)
มาฆพันธ์ อานาคิล และกมลวรรณ ดั่งธน กานนท์ (2557)	เพื่อเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์โดยใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด	5	20 (เหตุการณ์ละ 4 ข้อ)

ตารางที่ 2 (ต่อ)

ผู้วิจัย	จุดประสงค์	จำนวน เหตุการณ์	จำนวนข้อ
ชนิสรา สงวนไว้ และ ศิริชัย กาญจนวาสี (2558)	เพื่อเปรียบเทียบความเที่ยง ของแบบสอบถามวัด ความสามารถในการ แก้ปัญหาอย่างสร้างสรรค์ทาง คณิตศาสตร์ โดยใช้ทฤษฎีการ สรุปอ้างอิงความน่าเชื่อถือ ของผลการวัด	3	15 (เหตุการณ์ละ 5 ข้อ)
สุภชิต ผดุงผล และ กมลวรรณ ตั้งธนา กานนท์ (2559)	เพื่อเปรียบเทียบความเที่ยง ของแบบสอบถามวัด วัดความสามารถในการใช้ วิธีการทางวิทยาศาสตร์ในวิชา ฟิสิกส์ โดยใช้ทฤษฎีการสรุป อ้างอิงความน่าเชื่อถือของผล การวัด	4	20 (เหตุการณ์ละ 5 ข้อ)
กติกกร กมลรัตนสมบัติ และกมลวรรณ ตั้งธนา กานนท์ (2559)	เพื่อเปรียบเทียบพัฒนาการ ความสามารถในการ แก้ปัญหาทางคณิตศาสตร์ที่ ได้รับการให้ข้อมูลย้อนกลับ จากแบบสอบถามวัด ประเด็น	3	12 (เหตุการณ์ละ 4 ข้อ)
พรพิมล บิลันธนาภรณ์ และชิตชนก เริงเซาว์ (2561)	เพื่อพัฒนาแบบสอบถามวัด ประเด็นวัดความสามารถใน การตัดสินใจทางคลินิกของ นักศึกษาแพทย์ โดยใช้ทฤษฎี การสรุปอ้างอิงความน่าเชื่อถือ ของผลการวัด	5	41 (เหตุการณ์ละ 8-9 ข้อ)

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง พบว่าจำนวนเหตุการณ์ที่นิยมใช้ในการพัฒนาแบบสอบถามวัดประเด็น คือ 3-5 เหตุการณ์ เหตุการณ์ละ 4-5 ข้อคำถาม ผู้วิจัยจึง

สนใจสร้างแบบสอบอัตนัยประยุกต์ที่มีจำนวน 3 เหตุการณ์ โดยแต่ละเหตุการณ์มี 5 ข้อคำถามตามขั้นของการคิดวิจารณ์ญาณ

3.4 งานวิจัยที่เกี่ยวข้องกับแบบสอบอัตนัยประยุกต์

จากการศึกษางานวิจัยในประเทศและงานวิจัยต่างประเทศที่เกี่ยวข้องกับแบบสอบอัตนัยประยุกต์ มีดังนี้

มาฆพันธ์ อานาคิล และกมลวรรณ ตังธนากานนท์ (2557) ศึกษาการเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกันภายใต้จำนวนเหตุการณ์ โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด เครื่องมือที่ใช้ในงานวิจัยในครั้งนี้เป็นแบบสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ จำนวน 7, 9, 11, 13 และ 15 เหตุการณ์ แต่ละเหตุการณ์มีข้อคำถามจำนวน 4 ข้อ จากกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 91 คน ภายใต้วิธีการตรวจให้คะแนน 3 วิธี คือ วิธีตรวจให้คะแนน Knox วิธีตรวจให้คะแนน Feletti และวิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) วิเคราะห์และตรวจสอบคุณภาพของแบบทดสอบ โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดที่มีหนึ่งองค์ประกอบคือ จำนวนผู้สอบและจำนวนเหตุการณ์ ภายใต้วิธีการวัดที่แตกต่างกัน พบว่า แบบสอบที่ใช้วิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อย มีค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมพัทธ์ (relative coefficient) และค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) สูงที่สุด รองลงมาคือ แบบสอบที่ใช้วิธีตรวจให้คะแนน Knox และสุดท้ายคือ แบบสอบที่ใช้วิธีตรวจให้คะแนน Feletti และการศึกษาเพื่อการตัดสินใจการสรุปอ้างอิงภายใต้เงื่อนไขจำนวนเหตุการณ์ พบว่า ความเที่ยงของวิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อย มีค่าความเที่ยงสูงสุดในทุกเงื่อนไขจำนวนเหตุการณ์ และมีค่าความเที่ยงสูงเมื่อจำนวนเหตุการณ์เพิ่มขึ้น

ชนิสรา สงวนไว้ และศิริชัย กาญจนวาสิ (2558) ศึกษาการเปรียบเทียบความเที่ยงของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด เครื่องมือที่ใช้ในงานวิจัยในครั้งนี้เป็นแบบสอบอัตนัยประยุกต์ จำนวนข้อสอบ 3 ข้อ จากกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 120 คน วิเคราะห์และตรวจสอบคุณภาพของแบบทดสอบโดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดที่มีสององค์ประกอบ คือ จำนวนข้อสอบ และจำนวนผู้ตรวจ โดยแบ่งออกเป็น 2 ลักษณะ คือ ผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน (two-facet

crossed design) และผู้ตรวจตรวจข้อสอบเฉพาะข้อของผู้สอบทุกคน (two-facet nested design) พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (relative coefficient) และค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) สูงขึ้นเมื่อจำนวนข้อสอบและจำนวนผู้ตรวจเพิ่มขึ้น และจากการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง พบว่า การออกแบบโดยผู้ตรวจตรวจเฉพาะข้อของผู้สอบทุกคนมีค่าความเที่ยงสูงกว่าผู้ตรวจตรวจทุกข้อของผู้สอบทุกคน และควรเลือกใช้สถานการณ์ $p \times i \times r$ ที่มีผู้ตรวจ 1 คน ตรวจให้คะแนน 7 ข้อ หรือ $p \times (i:r)$ ที่มีผู้ตรวจ 2 คน ตรวจให้คะแนนคนละ 5 ข้อ เพื่อให้มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์มากกว่า 0.80

กติกกร กมลรัตนะสมบัติ และกมลวรรณ ดังธนกานนท์ (2559) ศึกษาผลของการให้ข้อมูลย้อนกลับจากแบบสอบอัตนัยประยุกต์ที่มีต่อพัฒนาการความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนระดับมัธยมศึกษา เครื่องมือที่ใช้ในงานวิจัยในครั้งนี้เป็นแบบสอบอัตนัยประยุกต์ จำนวน 3 เหตุการณ์ แต่ละเหตุการณ์ประกอบด้วย 4 ข้อคำถาม จำนวนรวม 12 ข้อ เกณฑ์การตรวจให้คะแนนแบบวิเคราะห์ย่อย และข้อมูลย้อนกลับแบบชี้แนะเพื่อการปรับปรุง จากกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 60 คน แบ่งเป็นกลุ่มทดลอง 30 คน และกลุ่มควบคุม 30 คน ซึ่งแต่ละกลุ่มจะแบ่งออกเป็นนักเรียนตามระดับความสามารถในการแก้ปัญหาทางคณิตศาสตร์สูง ปานกลาง และต่ำ ระดับละ 10 คน โดยใช้การวิเคราะห์ความแปรปรวนแบบวัดซ้ำ พบว่า นักเรียนมีพัฒนาการความสามารถในการแก้ปัญหาทางคณิตศาสตร์สูงที่สุดในระยะที่ 3 ทั้งกลุ่มได้รับข้อมูลย้อนกลับแบบชี้แนะเพื่อการปรับปรุงและกลุ่มที่ไม่ได้รับข้อมูลย้อนกลับ โดยนักเรียนที่ได้รับข้อมูลย้อนกลับแบบชี้แนะเพื่อการปรับปรุงมีพัฒนาการความสามารถในการแก้ปัญหาทางคณิตศาสตร์สูงกว่ากลุ่มที่ไม่ได้รับข้อมูลย้อนกลับ และนักเรียนที่มีความสามารถทางคณิตศาสตร์แตกต่างกันที่ได้รับข้อมูลย้อนกลับแบบชี้แนะเพื่อการปรับปรุงมีพัฒนาการความสามารถในการแก้ปัญหาทางคณิตศาสตร์ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สุภชิต ผดุงผล และกมลวรรณ ดังธนกานนท์ (2559) ศึกษาการเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์สำหรับวัดความสามารถในการใช้วิธีการทางวิทยาศาสตร์ในวิชาฟิสิกส์ ภายใต้จำนวนเหตุการณ์และจำนวนผู้ตรวจที่ต่างกัน โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด เครื่องมือที่ใช้ในงานวิจัยในครั้งนี้เป็นแบบสอบอัตนัยประยุกต์ จำนวน 4 เหตุการณ์ แต่ละเหตุการณ์ประกอบด้วย 5 ข้อคำถาม จำนวนรวม 20 ข้อ จากกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 จำนวน 94 คน จากโรงเรียนในสังกัด เขตพื้นที่

การศึกษาเขต 1 กรุงเทพมหานคร วิเคราะห์และตรวจสอบคุณภาพของแบบทดสอบโดย ประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดที่มีสามองค์ประกอบคือ จำนวน เหตุการณ์ จำนวนข้อคำถาม และจำนวนผู้ตรวจ โดยใช้ค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนน สำหรับไปใช้ตัดสินใจเชิงสัมพัทธ์ของแบบสอบอัตโนมัติประยุกต์วัดความสามารถในการใช้วิธีการ ทางวิทยาศาสตร์ในวิชาฟิสิกส์ พบว่า แบบสอบที่มีจำนวนเหตุการณ์ตั้งแต่ 4 เหตุการณ์ขึ้นไป ภายใต้ผู้ตรวจจำนวน 1 คน จึงจะมีค่าความเที่ยงเกิน 0.8 และค่าสัมประสิทธิ์ความน่าเชื่อถือของ คะแนนสำหรับไปใช้ตัดสินใจเชิงสัมบูรณ์ของแบบสอบอัตโนมัติประยุกต์วัดความสามารถในการใช้ วิธีการทางวิทยาศาสตร์ในวิชาฟิสิกส์ พบว่า แบบสอบที่มีจำนวนเหตุการณ์ตั้งแต่ 7 เหตุการณ์ขึ้นไป ภายใต้ผู้ตรวจจำนวน 2 คน จึงจะมีค่าความเที่ยงไม่ต่ำกว่า 0.8

Nagandla et al. (2018) ศึกษาผลการประเมินระหว่างเรียนของนักศึกษาแพทย์ โดยใช้แบบสอบอัตโนมัติประยุกต์ (MEQs) แบบสอบทักษะวิชาชีพ (OSPE) และแบบประเมินทาง คลินิก (OSCE) ว่าส่งผลต่อการทำนายการสอบใบประกอบจบวิชาชีพเวชกรรมหรือไม่ จาก นักศึกษาแพทย์ 3 กลุ่ม จำนวน 245 คน ซึ่งจะดำเนินการสอบทุก ๆ 6 เดือน ในภาคเรียนที่ 2 ของ ชั้นปีที่ 3 และภาคเรียนที่ 1 และ 2 ของชั้นปีที่ 4 ซึ่งแบบสอบอัตโนมัติประยุกต์ (MEQs) มีลักษณะ เป็นสถานการณ์ทางคลินิกสั้น ๆ ที่มีรูปแบบการตรวจให้คะแนนรูปrikที่แน่นอน สำหรับแบบสอบ ทักษะวิชาชีพ (OSPE) มีลักษณะเป็นสถานการณ์ตัวอย่าง 18 สถานการณ์เพื่อวัดทักษะการตีความ การแก้ปัญหาและการสื่อสาร รวมถึงทัศนคติของนักศึกษาแพทย์ ส่วนแบบประเมินทางคลินิก (OSCE) มีลักษณะเป็นสถานการณ์ทางคลินิกแบบสั้น 5 นาทีและแบบยาว 15 นาที จำนวน 16 สถานการณ์ พบว่า แบบสอบอัตโนมัติประยุกต์ (MEQs) แบบสอบทักษะวิชาชีพ (OSPE) และแบบ ประเมินทางคลินิก (OSCE) ไม่สามารถทำนายความสำเร็จในการสอบใบประกอบจบวิชาชีพเวช กรรมได้ที่ระดับความเชื่อมั่น 95% แต่ค่าสหสัมพันธ์ของแบบสอบอัตโนมัติประยุกต์ (MEQs) และ แบบสอบทักษะวิชาชีพ (OSPE) สูงกว่าแบบประเมินทางคลินิก (OSCE)

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องข้างต้น พบว่า มีหลากหลาย งานวิจัยที่นำแบบสอบอัตโนมัติประยุกต์ไปใช้วัดทักษะการแก้ปัญหาทางคณิตศาสตร์ และวัดทักษะ การแก้ปัญหาทางวิทยาศาสตร์ แต่ทั้งนี้ก็ยังไม่มีพบทั้งงานวิจัยในประเทศและงานวิจัยต่างประเทศที่ นำแบบสอบอัตโนมัติประยุกต์นี้ไปวัดการคิดวิจารณ์ญาณซึ่งครอบคลุมทักษะการคิดแก้ปัญหาใน วิชาคณิตศาสตร์ ผู้วิจัยจึงสนใจที่จะพัฒนาเครื่องมือในการวัดการคิดวิจารณ์ญาณทาง คณิตศาสตร์โดยใช้แบบสอบอัตโนมัติประยุกต์

ตอนที่ 4 ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หรือ G – Theory

ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory: G-Theory) ถือว่าคะแนนความคลาดเคลื่อนของการวัดสามารถเกิดได้จากหลายแหล่ง (multiple error sources) ที่สามารถประมาณค่าแยกกันภายใต้การวิเคราะห์ครั้งเดียวได้ ทฤษฎีนี้ต้องการสรุปอ้างอิงความน่าเชื่อถือของคะแนนที่สังเกตได้ จากแบบทดสอบที่ทำการวัดในเงื่อนไขหรือสถานการณ์ต่าง ๆ ของการทดสอบไปยังคะแนนจริง เงื่อนไขของสถานการณ์หรือการทดสอบที่จะใช้ไม่ได้หลากหลายขึ้นอยู่กับกรอบการออกแบบ (design) G-Theory จึงได้เสนอโมเดลของการศึกษาวิเคราะห์ความน่าเชื่อถือของผลการวัด หรือความเที่ยงในสถานการณ์หรือเงื่อนไขต่าง ๆ ของการวัด ทำให้ทราบและสามารถควบคุมแหล่งความคลาดเคลื่อนได้ตรงประเด็น เพื่อผลการวัด มีความน่าเชื่อถือ หรือมีความเที่ยงสูงถึงระดับที่ต้องการ ผู้วิจัยขอแนะนำเสนอประเด็นตามลำดับต่อไปนี้ คือ ความเป็นมาของ G-Theory แนวคิดพื้นฐานของ G-Theory แนวคิดสำคัญและข้อตกลงเบื้องต้น คำศัพท์ที่เกี่ยวข้องกับ G-Theory และการออกแบบพาทเซท (ศิริชัย กาญจนวาสี, 2563)

4.1 ความเป็นมาของ G – Theory

นักทฤษฎีการทดสอบหลายท่านพยายามศึกษาเชิงจำแนกแหล่งความคลาดเคลื่อนของการวัดโดยใช้เทคนิควิเคราะห์ความแปรปรวน (ANOVA) การนำเทคนิคการคิดวิเคราะห์ความแปรปรวนซึ่งเป็นวิธีการ ทางสถิติมาใช้ประมาณค่าความเที่ยงของแบบสอบและความคลาดเคลื่อนจากการวัด มีมานานแล้วในแวดวงวิชาการการวัดผล วิธีที่เป็นที่รู้จักกันดี ได้แก่ วิธีการวิเคราะห์ของฮอยท์ (Hoyt, 1941) นอกจากนี้ยังมีท่านอื่นที่เสนอวิธีการในการทำงานคล้าย ๆ กัน เช่น Linn (1953), Medley and Mitzel (1963) เป็นต้น ในเวลาต่อมา Cronbach et al. (1963) ได้เสนอแนวคิดทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดขึ้นเป็นครั้งแรก เพื่อประมาณค่าความเที่ยงทั่วไปของแบบสอบในสถานการณ์ของการทดสอบหรือเงื่อนไขการวัดแบบต่าง ๆ นอกจากนี้ Rajaratnam et al. (1965) Gleser et al. (1965) และ Cronbach et al. (1972) ได้พัฒนาทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดให้มีรูปแบบของโมเดล และการประยุกต์ใช้ในสถานการณ์ทดสอบที่ชัดเจนขึ้น จึงทำให้แนวคิดและการประมาณค่าความเที่ยงตามทฤษฎีการทดสอบแบบดั้งเดิม ได้ขยายกว้างไกลออกไปและมีความครอบคลุมเงื่อนไขการวัดหรือแหล่งความคลาดเคลื่อนต่าง ๆ ได้ด้วย ซึ่งจะเป็นประโยชน์ต่อการพัฒนาแบบสอบหรือแบบวัดให้มีระดับความเที่ยงที่ต้องการ ภายใต้สถานการณ์ทดสอบที่ต้องการนำแบบสอบหรือแบบวัดไปใช้ พร้อมทั้งได้มีการประยุกต์ใช้ในการวิเคราะห์ความตรงทั่วไปได้อีกด้วย

4.2 แนวคิดพื้นฐานของ G – Theory

Cronbach et al. (1972) ได้เสนอทฤษฎีสำหรับการวิเคราะห์ความน่าเชื่อถือของผลการวัดพฤติกรรม (the dependability of behavioral measurements) สำหรับสถานการณ์ของการวัดผลลักษณะต่าง ๆ ซึ่งต่อมาเป็นที่รู้จักกันในชื่อของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หรือ G-Theory ทฤษฎีนี้ได้ขยายแนวคิดของความเที่ยง (reliability) ตามทฤษฎีการทดสอบดั้งเดิม (CTT) ให้สามารถสรุปผลความเที่ยงในสถานการณ์หรือเงื่อนไขการทดสอบลักษณะต่าง ๆ ได้

ทฤษฎีการทดสอบแบบดั้งเดิมมีพื้นฐานความเชื่อว่า ความผันแปรของคะแนนที่สังเกตได้ (σ_X^2) ประกอบด้วย 2 ส่วน คือ ความผันแปรของคะแนนจริง (σ_T^2) ซึ่งเป็นความแตกต่างที่แท้จริงระหว่างบุคคล และความผันแปรของคะแนนความคลาดเคลื่อน (σ_E^2) ซึ่งเป็นความคลาดเคลื่อนรวมทุกแหล่ง ที่ไม่สามารถระบุหรือแบ่งแยกได้ (single error source) นอกจากนี้ทฤษฎีการทดสอบแบบดั้งเดิม มิได้ให้ความสนใจต่อสถานการณ์หรือเงื่อนไขการวัด และปฏิสัมพันธ์ระหว่างผู้สอบกับเงื่อนไขของการวัดที่สามารถส่งผลกระทบต่อความน่าเชื่อถือของผลการวัด ในขณะที่ G-Theory ได้ให้แนวคิดแยกส่วนความคลาดเคลื่อน (error) จากหลายแหล่ง (multiple error sources) ประกอบด้วย แหล่งความคลาดเคลื่อนอย่างเป็นระบบ (systematic source) และความคลาดเคลื่อนสุ่ม (random source) มิใช่เป็นความคลาดเคลื่อนรวมเพียงแหล่งเดียวเหมือนอย่างทฤษฎีการทดสอบแบบดั้งเดิม ดังภาพที่ 4

ภาพที่ 4 การแยกส่วนความคลาดเคลื่อนใน G-Theory

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2$$

```

graph TD
    A["σX2 = σT2 + σE2"] --- B["σT2"]
    A --- C["σE2"]
    B --- D["Systematic error variance"]
    C --- E["Random error variance"]
  
```

G-Theory ได้เสนอวิธีวิเคราะห์ค่าสัมพัทธ์ของความแปรปรวนของคะแนนความคลาดเคลื่อนอย่างเป็นระบบจากแหล่งต่าง ๆ อันเป็นสถานการณ์หรือเงื่อนไขของการวัด เช่น

ความยาวของแบบสอบ ชุดของแบบสอบ จำนวนครั้งของการทดสอบ จำนวนผู้ตรวจ เป็นต้น รวมทั้งปฏิสัมพันธ์ระหว่างผู้สอบกับเงื่อนไขของการวัด เช่น ปฏิสัมพันธ์ระหว่างผู้สอบกับชุดของแบบสอบ ผู้สอบกับครั้งของการทดสอบ เป็นต้น จึงทำให้ทราบและสามารถควบคุมแหล่งความคลาดเคลื่อนได้ตรงประเด็น ยังผลให้สัมประสิทธิ์ความน่าเชื่อถือหรือคล้ายกับสัมประสิทธิ์ความเที่ยง (จากทฤษฎีการทดสอบแบบดั้งเดิม) ของการวัดสูงขึ้น G-Theory จึงเป็นทั้งวิธีประเมินความน่าเชื่อถือหรือความเที่ยงของเครื่องมือวัดผล และกลยุทธ์ของการออกแบบการวัดให้ได้ผลของการวัดที่มีความน่าเชื่อถือ หรือความเที่ยงสูงถึงระดับที่ต้องการ เพื่อนำผลไปใช้ในเป็นสารสนเทศสำหรับการตัดสินใจได้อย่างมีประสิทธิภาพ

4.2.1 แนวคิดสำคัญและข้อตกลงเบื้องต้น

G-Theory เป็นทฤษฎีทางสถิติของการวิเคราะห์ความน่าเชื่อถือของผลการวัดในสถานการณ์ของการวัดผลลักษณะต่าง ๆ ที่เป็นเป้าหมายของการนำเครื่องมือไปใช้ ความน่าเชื่อถือของผลการวัด หมายถึง ความถูกต้องของการสรุปอ้างอิง (generalization) จากคะแนนที่สังเกตได้ไปยังคะแนนจริงของบุคคล โดยคะแนนจริงเป็นคะแนนเฉลี่ยที่พึงได้ของผู้สอบแต่ละคน จากการทดสอบภายใต้สถานการณ์หรือเงื่อนไขของการวัดที่ยอมรับได้ทั้งหมด

การศึกษาความน่าเชื่อถือของผลการวัดตามแนว G-Theory อยู่บนพื้นฐานของข้อตกลงเบื้องต้น ดังนี้

1. คุณลักษณะที่มุ่งวัดของบุคคล ไม่ว่าจะเป็นความรู้ ทักษะ หรือคุณลักษณะอื่น ๆ ซึ่งเป็นเป้าหมายของการวัด เป็นค่าที่อยู่ในสภาวะคงที่ (steady state)
2. ผู้สอบคนเดียวกันได้คะแนนแตกต่างกันจากการวัดในแต่ละสถานการณ์ หรือเงื่อนไขของการวัด เนื่องมาจากความคลาดเคลื่อนที่เป็นระบบอย่างน้อย 1 แหล่ง โดยองค์ประกอบด้านวุฒิภาวะ และการเรียนรู้ระหว่างการวัด ไม่เป็นแหล่งความคลาดเคลื่อนที่ได้จากการวัด
3. เมื่อพิจารณาผู้สอบทั้งกลุ่ม ความแปรปรวนของคะแนนที่สังเกตได้ประกอบด้วย ความแปรปรวนของคะแนนจริง ซึ่งเป็นความแตกต่างที่แท้จริงระหว่างบุคคล ประกอบด้วย ความแปรปรวนของคะแนนความคลาดเคลื่อนที่เป็นระบบอย่างน้อย 1 แหล่ง และประกอบด้วย ความแปรปรวนของคะแนนความคลาดสุ่ม

ตามแนวคิดของ G-Theory คะแนนที่ได้จากการวัดภายใต้สถานการณ์หรือเงื่อนไขเดียวด้วยแบบสอบชุดเดียว และทำการสอบครั้งเดียว คะแนนที่ได้จะไม่สามารถเชื่อถือได้

อย่างเต็มที่ เพราะคะแนนที่ได้จะไม่เป็นตัวแทนที่ดีของคะแนนจริงซึ่งเป็นคะแนนเฉลี่ยที่ได้จากการทดสอบในหลาย ๆ ครั้ง เหตุการณ์ด้วยแบบสอบหลาย ๆ ชุด และทำการสอบหลาย ๆ ครั้ง

ผู้สอบคนเดียวเมื่อทำการทดสอบหลายครั้งในหลายเหตุการณ์ มักจะได้คะแนนในแต่ละครั้งแตกต่างกัน ความไม่คงเส้นคงวาของคะแนนเกิดจากความคลาดเคลื่อนของการวัด ซึ่งทฤษฎีการทดสอบแบบดั้งเดิมให้คำตอบด้วยการแยกวิเคราะห์สำหรับการทดสอบแต่ละครั้ง โดยพิจารณาว่าความคลาดเคลื่อนของการวัดมาจากแหล่งเดียว ในขณะที่ G-Theory ได้พิจารณาถึงแหล่งความคลาดเคลื่อนพหุสามารถวิเคราะห์ความคลาดเคลื่อนหลายแหล่งพร้อมกันไป ผู้บริหารการทดสอบจึงสามารถตัดสินใจได้ว่าควรใช้เหตุการณ์หรือเงื่อนไขการวัดแบบใด จึงจะได้คะแนนที่เชื่อถือถึงระดับที่ต้องการผลการวิเคราะห์ด้วย G-Theory จะให้ค่าสัมประสิทธิ์เชิงสรูปที่แสดงถึงระดับความน่าเชื่อถือที่ได้จากการวัด (level of dependability) เรียกว่า สัมประสิทธิ์การสรูปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Coefficient: G-Coefficient) ซึ่งคล้ายกับสัมประสิทธิ์ความเที่ยง (reliability coefficient) ในทฤษฎีการทดสอบแบบดั้งเดิม

การตรวจสอบความน่าเชื่อถือของผลการวัดโดยใช้ G-Theory ทำให้ผู้บริหารการทดสอบทำการตัดสินใจความน่าเชื่อถือของผลการวัดได้ 2 ลักษณะ ได้แก่

1. ความน่าเชื่อถือของคะแนนในการนำไปใช้เชิงสัมพัทธ์ (relative) เพื่อแปลผลเปรียบเทียบระหว่างบุคคล เช่น นาย ก. มีความสามารถสูงเพียงใดเมื่อเทียบกับกลุ่มผู้สอบด้วยกัน หรือ ก. มีความสามารถสูงหรือต่ำกว่าบุคคลในกลุ่มเพียงใด หรือ ก. ได้คะแนนคิดเป็นคะแนนมาตรฐานหรือเปอร์เซ็นต์ไทล์เท่าไร เป็นต้น

2. ความน่าเชื่อถือของคะแนนในการนำไปใช้เชิงสัมบูรณ์ (absolute) เพื่อแปลผลเปรียบเทียบกับเกณฑ์มาตรฐาน เช่น นาย ก. มีความสามารถระดับใด สูงหรือต่ำเพียงใดตามเกณฑ์มาตรฐาน หรือ ก. มีความสามารถผ่านระดับมาตรฐานหรือไม่ โดยไม่ได้สนใจเปรียบเทียบกับความสามารถของเพื่อน ๆ ในกลุ่ม เป็นต้น

4.2.2 คำศัพท์เกี่ยวข้องกับ G-Theory

ในการทำความเข้าใจเกี่ยวกับทฤษฎีการสรูปอ้างอิงความน่าเชื่อถือของผลการวัด มีความจำเป็นต้องทำความเข้าใจคำศัพท์ที่เกี่ยวข้อง คำศัพท์สามัญ ดังนี้ (ศิริชัย กาญจนวาสี, 2563)

1. ประชากร (population) และเอกภพ (universe)

G-Theory ได้จำแนกคำศัพท์ “ประชากร” และ “เอกภพ” ให้มีความหมายแตกต่างกัน ดังนี้

1.1 ประชากร (population) หมายถึง สิ่งที่มีมุ่งหวังทั้งหมด ในสถานการณ์ของการทดสอบทั่วไป สิ่งที่มีมุ่งหวังมักได้แก่ บุคคลหรือผู้ทำการสอบ

1.2 เอกภพ (universe) หมายถึง เงื่อนไขของการวัดที่สนใจทั้งหมด กลุ่มเงื่อนไขของการวัด เรียกว่า ฟาเซท (facet) ซึ่งเป็นองค์ประกอบที่คาดว่าจะมีผลต่อความคลาดเคลื่อนของการวัด เช่น ความยาวของแบบสอบ รูปแบบของข้อสอบ จำนวนครั้งของการทดสอบ จำนวนผู้ตรวจให้คะแนน เป็นต้น

1.3 ฟาเซท (facet) ที่ต้องการศึกษาอาจเป็นองค์ประกอบสุ่ม (random) หรือองค์ประกอบเจาะจง (fix) ถ้าเงื่อนไขของการวัดถูกเลือกมาอย่างเจาะจงจากองค์ประกอบที่ศึกษา แสดงว่า ผู้ศึกษาสามารถทำการสรุปอ้างอิงความเที่ยงของแบบสอบ ไปยังองค์ประกอบเฉพาะระดับของเงื่อนไขที่เลือกมาศึกษาเท่านั้น แต่ถ้าเงื่อนไขของการวัดได้รับการสุ่มเพื่อเป็นตัวแทนขององค์ประกอบที่ศึกษา แสดงว่า ผู้ศึกษาสามารถทำการสรุปอ้างอิงความเที่ยงของแบบสอบไปยังระดับต่าง ๆ ขององค์ประกอบที่ศึกษาได้

2. การศึกษา G (G-Study) และการศึกษา D (D-Study)

ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ประกอบด้วยขั้นตอนที่สำคัญ 2 ขั้นตอน ได้แก่ การสรุปเชิงอ้างอิง หรือการศึกษา G (Generalizability Study: G-Study) กับการศึกษาเชิงตัดสินใจ หรือการศึกษา D (Decision Stud: D-Study)

2.1 การศึกษา G (G-Study) เป็นการสรุปอ้างอิงผลจากการศึกษาตัวอย่างการวัดตามเงื่อนไขที่สนใจ บรรยายความแปรปรวนของความคลาดเคลื่อนจากแหล่งความคลาดเคลื่อนต่าง ๆ ของการวัด

2.2 การศึกษา D (D-Study) เป็นการใช้อ้างอิงข้อมูลจากการศึกษา G ที่สอดคล้องกับจุดประสงค์เฉพาะของการตัดสินใจเลือกใช้แบบสอบในสถานการณ์ต่าง ๆ ของการวัด

3. ความแปรปรวนของคะแนนความคลาดเคลื่อนสัมบูรณ์และสัมพัทธ์ (absolute and relative error variance)

ตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) คะแนนจริง (true score : T_p) ของผู้สอบ คือ ค่าเฉลี่ยของคะแนนจากการสอบซ้ำ ๆ ด้วยแบบสอบคู่ขนาน ความแปรปรวนของคะแนนจริงจึงเป็นความแปรปรวนของค่าเฉลี่ยของการสอบซ้ำนั้น และความแปรปรวนของคะแนนที่สังเกตได้ เป็นผลรวมของความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนความคลาดเคลื่อน ดังนี้

$$X_{pi} = T_{pi} + E_{pi}$$

$$\sigma_{xp}^2 = \sigma_{Tp}^2 + \sigma_{Ep}^2$$

สำหรับ G-Theory คะแนนเอกภพ (universe score : M_p) คือ ค่าเฉลี่ยของคะแนนการวัดซ้ำ ๆ ครั้ง ตามเงื่อนไขการวัดในเอกภพของการสุ่มอย่างอิสระ สำหรับความคลาดเคลื่อนของการวัด (E_{pi}) ถูกจำแนกออกเป็นความคลาดเคลื่อนจากฟาสเซท หรือกลุ่มเงื่อนไขของการวัด (E_i) และความคลาดเคลื่อนจากแหล่งที่เหลืออื่น ๆ (e_{pi}) ซึ่งการวัดแต่ละครั้งไม่จำเป็นต้องใช้แบบสอบคู่ขนานเหมือนทฤษฎีการทดสอบแบบดั้งเดิม ส่วนความแปรปรวนของค่าคาดหวังของคะแนนที่สังเกตได้เป็นผลรวมของความแปรปรวนของคะแนนเอกภพ (σ_p^2 หรือเขียนย่อ ๆ ว่า σ_p^2) กับความแปรปรวนของคะแนนความคลาดเคลื่อนจากฟาสเซท หรือองค์ประกอบ (i) ต่าง ๆ ของการวัด ($\sigma_{E_i}^2$) และความแปรปรวนของคะแนนความคลาดเคลื่อนจากแหล่งอื่น ๆ ($\sigma_{e_p}^2$) ดังนี้

$$X_{pi} = T_{pi} + E_{pi}$$

$$\sigma_{xp}^2 = \sigma_{\mu_p}^2 + \sigma_{E_i}^2 + \sigma_{e_p}^2$$

สำหรับความแปรปรวนของคะแนนความคลาดเคลื่อนจากองค์ประกอบต่าง ๆ ของการวัด สามารถแบ่งออกเป็น 2 ประเภท ได้แก่

3.1 ความแปรปรวนของคะแนนความคลาดเคลื่อนสัมบูรณ์ (absolute error variance: σ_{ABS}^2 หรือ σ_{Δ}^2) คือ ความแปรปรวนของ $\mu_p - X_p$ ซึ่งคำนวณได้จากผลรวมของความแปรปรวนของคะแนนจากแหล่งต่าง ๆ ยกเว้น σ_p^2 หรือ $\sigma_{\mu_p}^2$

3.2 ความแปรปรวนของคะแนนความคลาดเคลื่อนสัมพัทธ์ (relative error variance: σ_{REL}^2 หรือ σ_{δ}^2) คือ ความแปรปรวนของ $\mu_p - X_p$ ซึ่งคำนวณได้จากผลรวมของความแปรปรวนของคะแนน จากแหล่งต่าง ๆ ที่มีปฏิสัมพันธ์ร่วมกับผู้สอบ (p)

4. สัมประสิทธิ์การสุ่มอย่างอิสระ (G-Coefficient)

ตามทฤษฎีการทดสอบแบบดั้งเดิม ความเที่ยง (reliability) ของแบบสอบเป็นสัดส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่สังเกตได้ ดังนี้

$$\begin{aligned} \text{Reliability} &= \frac{\sigma_T^2}{\sigma_X^2} \\ &= 1 - \frac{\sigma_E^2}{\sigma_X^2} \\ &= \rho_{XT} \\ &= \rho_{XX} \end{aligned}$$

สำหรับ G-Theory สัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient: ρ^2) เป็นสัดส่วนระหว่าง ความแปรปรวนของคะแนนเอกภพ กับความแปรปรวนของค่าคาดหวังของคะแนนที่สังเกตได้ ดังนี้

$$\text{G-Coefficient} = \frac{\sigma_p^2}{\sigma_p^2 + \text{Error Variance}}$$

เนื่องจากความแปรปรวนของคะแนนความคลาดเคลื่อนจากองค์ประกอบต่าง ๆ ของการวัดมี 2 ประเภท จึงทำให้การสรุปอ้างอิงมี 2 ประเภท ได้แก่

4.1 สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (ρ_{Abs}^2) เมื่อคะแนนความคลาดเคลื่อนสัมบูรณ์ สัมประสิทธิ์นี้บ่งบอกความเที่ยงของแบบสอบ ในสถานการณ์ของการตัดสินใจที่ขึ้นกับคะแนนของผู้สอบตามลำพัง ไม่มีการเปรียบเทียบภายในกลุ่มหรือระหว่างกลุ่ม เช่น ความเที่ยงของแบบสอบอิงเกณฑ์ ซึ่งใช้แสดงความน่าเชื่อถือสำหรับตรวจสอบความสามารถของผู้สอบเทียบกับเกณฑ์ที่กำหนดไว้ล่วงหน้า เป็นต้น

4.2 สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (ρ_{Rel}^2) เมื่อคะแนนความคลาดเคลื่อนสัมพัทธ์ สัมประสิทธิ์นี้บ่งบอกความเที่ยงของแบบสอบ ในสถานการณ์ของการตัดสินใจที่มีการเปรียบเทียบคะแนนระหว่างผู้สอบ เช่น ความเที่ยงของแบบสอบอิงกลุ่ม ซึ่งใช้แสดงความน่าเชื่อถือสำหรับเปรียบเทียบคะแนนระหว่างผู้สอบด้วยการเข้าโค้งการแจกแจงคะแนน และตัดเกรด หรือ การสอบแข่งขันหาผู้สอบที่ได้คะแนนสูงเมื่อเทียบกับกลุ่ม เพื่อเข้าศึกษาต่อในสถานการณ์ต่าง ๆ เป็นต้น

4.3 การออกแบบฟาเซท

การออกแบบฟาเซท หมายถึง การออกแบบเงื่อนไขในสิ่งที่เราต้องการศึกษาซึ่งผู้วิจัยได้ศึกษารายละเอียดการออกแบบฟาเซทไว้ดังนี้ (ศิริชัย กาญจนวาสี, 2555 อ้างถึงใน สุขจิต ผดุงผล และกมลวรรณ ตังธนากานนท์, 2559)

4.3.1 การออกแบบการวัดที่มีหนึ่งฟาเซท (one-facet design)

การออกแบบการวัดหนึ่งฟาเซท (one-facet, $p \times i$ design) คือ คะแนนที่สังเกตได้ของบุคคลในข้อสอบข้อนั้นที่ได้รับอิทธิพลมาจากส่วนต่าง ๆ ดังนี้ (Shavelson & Webb, 1991)

$$\begin{aligned}
 X_{pi} &= \mu && \text{(ค่าคงที่หรือค่าเฉลี่ยของทั้งหมด)} \\
 &+ \mu_p - \mu && \text{(อิทธิพลของบุคคล)} \\
 &+ \mu_i - \mu && \text{(อิทธิพลของข้อสอบ)} \\
 &+ X_{pi} - \mu_p - \mu_i + \mu && \text{(ส่วนที่เหลือ หรือ } e_{pi}\text{)}
 \end{aligned}$$

เมื่อ μ คือ ค่าเฉลี่ยของคะแนนทั้งหมดจากประชากรในเอกภพ (Grand mean)

μ_p คือ คะแนนเอกภพของแต่ละคน

μ_i คือ คะแนนเฉลี่ยรายข้อ

พบว่าอิทธิพลอื่นที่ต่างไปจากค่าเฉลี่ยทั้งหมด (grand mean) มีการกระจายเกิดขึ้นแต่ค่าเฉลี่ยทั้งหมดมีค่าคงที่ จึงมีความแปรปรวนเป็นศูนย์

การแจกแจงที่มีค่าเฉลี่ยเป็นศูนย์ และมีความแปรปรวน (σ^2) เรียกว่า องค์ประกอบความแปรปรวน (variance component) ประกอบด้วยส่วนแรกเป็นอิทธิพลของบุคคล ค่าเฉลี่ยของความแตกต่างของคะแนนระหว่างบุคคลเป็นศูนย์ คือ

$$E_p(\mu_p - \mu) = E_p(\mu_p) - E_p(\mu) - E_p(\mu) = \mu - \mu = 0$$

ความแปรปรวนของอิทธิพลของบุคคล ใช้สัญลักษณ์เป็น σ_p^2 เรียกว่า องค์ประกอบความแปรปรวนของบุคคล หรือความแปรปรวนของคะแนนเอกภพ เขียนสมการได้เป็น

$$\sigma_p^2 = E_p(\mu_p - \mu)^2$$

ค่าเฉลี่ยและองค์ประกอบความแปรปรวนของแบบสอบ จะมีค่าเฉลี่ยรายข้อเป็นศูนย์และองค์ประกอบความแปรปรวนของแบบสอบ ใช้สัญลักษณ์ σ_i^2 อิทธิพลส่วนสุดท้ายคือ ส่วนที่เหลือ มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนเป็น $\sigma_{pi,e}^2$

จะได้ว่าความแปรปรวนของคะแนนสังเกตได้ เป็นผลรวมของ 3 องค์ประกอบความแปรปรวนข้างต้น ดังนี้

$$\sigma_{X_{pi}}^2 = \sigma_p^2 + \sigma_i^2 + \sigma_{pi,e}^2$$

เมื่อ $\sigma_{X_{pi}}^2$ คือ ความแปรปรวนของคะแนนสังเกตได้ (X_{pi})

σ_p^2 คือ ความแปรปรวนของคะแนนเอกภพ (μ_p)

σ_i^2 คือ ความแปรปรวนของคะแนนเฉลี่ยรายข้อ (μ_i)

$\sigma_{pi,e}^2$ คือ ความแปรปรวนของคะแนนความคลาดเคลื่อน หรือส่วนที่เหลือ (e_{pi})

ความแปรปรวนของคะแนนที่สังเกตได้ในการวัดที่มีหนึ่งองค์ประกอบแบบไขว้ (one-facet crosses design) จึงขึ้นอยู่กับ 4 แหล่ง คือ 1) ความแตกต่างระหว่างบุคคลหรือสิ่งที่วัด (person effect) 2) ความแตกต่างระหว่างความยากของข้อสอบ (item effect) 3) ปฏิสัมพันธ์ระหว่างบุคคลกับข้อสอบ ($p \times i$ interaction) และ 4) ความคลาดเคลื่อนที่เกิดขึ้นอย่างสุ่ม (unsystematic of random error) แต่ความแปรปรวนจากแหล่งที่ 3 และ 4 ไม่สามารถแยกออกจากกันได้ เรียกว่า เป็นส่วนที่เหลือ (residual) จึงเป็นแหล่งความแปรปรวนร่วมกัน ดังภาพที่ 5

ภาพที่ 5 แหล่งความคลาดเคลื่อนและองค์ประกอบของความแปรปรวนสำหรับหนึ่งองค์ประกอบแบบไขว้ $p \times i$ (one-facet crossed design, $p \times i$ design)



จุฬาลงกรณ์มหาวิทยาลัย

4.3.2 การออกแบบการวัดที่มีสองฟาเซต (Two-Facet Design)

4.3.2.1 การออกแบบการวัดที่มีสองฟาเซตแบบไขว้ Two-Facet Crossed Design ($p \times r \times o$ design) คือ คะแนนที่สังเกตได้ของบุคคลในข้อสอบข้อหนึ่ง ได้รับอิทธิพลมาจากส่วนต่าง ๆ ดังนี้

$X_{pi} = \mu$	(ค่าคงที่หรือค่าเฉลี่ยของทั้งหมด)
$+ \mu_p - \mu$	(อิทธิพลของบุคคล)
$+ \mu_r - \mu$	(อิทธิพลของผู้ตรวจ)
$+ \mu_o - \mu$	(อิทธิพลของเหตุการณ์)
$+ \mu_{pr} - \mu_p - \mu_r + \mu$	(ปฏิสัมพันธ์ระหว่างบุคคลกับผู้ตรวจ)

$$\begin{aligned}
& + \mu_{po} - \mu_p - \mu_o + \mu && \text{(ปฏิสัมพันธ์ระหว่างบุคคลกับ} \\
& && \text{เหตุการณ์)} \\
& + \mu_{ro} - \mu_r - \mu_o + \mu && \text{(ปฏิสัมพันธ์ระหว่างผู้ตรวจกับ} \\
& && \text{เหตุการณ์)} \\
& + X_{pro} - \mu_{po} - \mu_{pr} - \mu_{ro} + \mu_p + \mu_r + \mu_o - \mu && \text{(อิทธิพลของ} \\
& && \text{ส่วนที่เหลือ)}
\end{aligned}$$

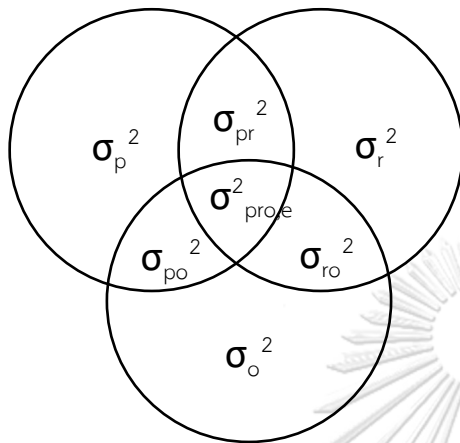
แต่ละองค์ประกอบยกเว้นค่าเฉลี่ยทั้งหมด (grand mean) นั้นมีการกระจาย นั่นคือ พฤติกรรมของผู้สอบเปลี่ยนไป ผู้ตรวจให้คะแนนแตกต่างกัน ซึ่งแต่ละการกระจายมีค่าเฉลี่ยเป็นศูนย์และมีความแปรปรวน ความแปรปรวนทั้งหมดของเอกภพและประชากร จึงเป็นผลรวมขององค์ประกอบความแปรปรวน 7 ส่วน ดังนี้

$$\sigma_{X_{pro}}^2 = \sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_{pro,e}^2$$

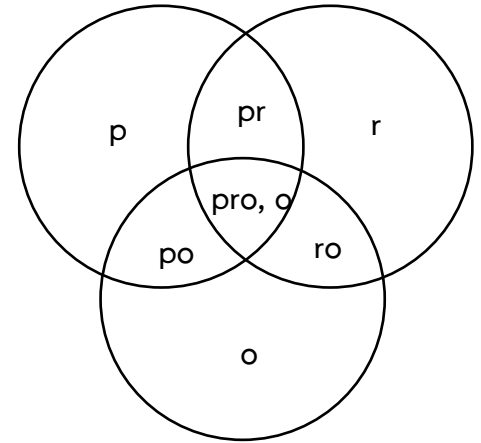
- เมื่อ
- $\sigma_{X_{pro}}^2$ คือ ความแปรปรวนของคะแนนสังเกตได้
 - σ_p^2 คือ ความแปรปรวนของคะแนนเอกภพของผู้สอบ (μ_p)
 - σ_r^2 คือ ความแปรปรวนของคะแนนเฉลี่ยของผู้ตรวจ (μ_r)
 - σ_o^2 คือ ความแปรปรวนของคะแนนเฉลี่ยของเหตุการณ์การสอบแต่ละครั้ง (μ_o)
 - σ_{pr}^2 คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับผู้ตรวจ
 - σ_{po}^2 คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับเหตุการณ์ของการสอบแต่ละครั้ง
 - σ_{ro}^2 คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้ตรวจกับเหตุการณ์ของการสอบแต่ละครั้ง
 - $\sigma_{pro,e}^2$ คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ pro กับความคลาดเคลื่อนส่วนที่เหลือ (e_{pro})

เขียนแผนภาพเวนนีได้ดังภาพที่ 6

ภาพที่ 6 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับสองฟาเซทแบบไขว้
 $p \times r \times o$ (two-facet crossed design, $p \times r \times o$ design)



องค์ประกอบความแปรปรวน
 (variance component)



แหล่งความแปรปรวน
 (source of variation)

4.3.2.2 การออกแบบการวัดที่มีสองฟาเซทแบบแฝง (two-facet nested design, $p \times (i:o)$ design)

สมมติต้องการออกแบบและวิเคราะห์ความแปรปรวนของคะแนนจริงและความแปรปรวนของคะแนนความคลาดเคลื่อนจากแหล่งต่าง ๆ ที่สนใจ โดยนักวัดผลทำการสุ่มตัวอย่างนักเรียน p คน เพื่อให้ทำข้อสอบ i ข้อ และทำการทดสอบ o ครั้ง ซึ่งแต่ละครั้งของการทดสอบใช้ข้อสอบต่างกัน ดังนี้

$$\sigma^2(X_{px(i:o)}) = \sigma_p^2 + \sigma_o^2 + \sigma_{i:o}^2 + \sigma_{po}^2 + \sigma_{pi:o,e}^2$$

เมื่อ $\sigma^2(X_{px(i:o)})$ คือ ความแปรปรวนของคะแนนสังเกตได้

σ_p^2 คือ ความแปรปรวนของคะแนนเอกภพของนักเรียน

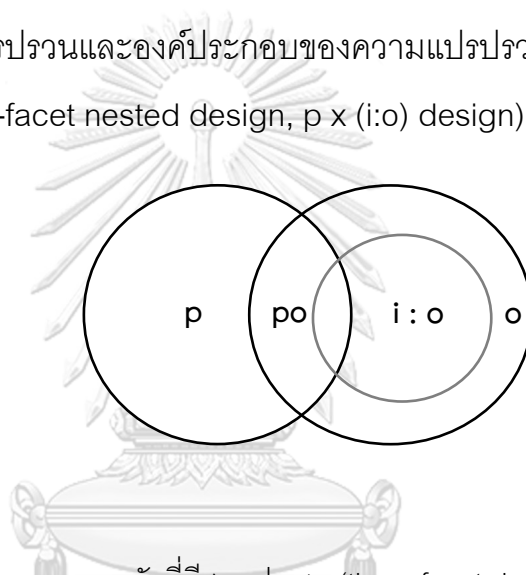
σ_o^2 คือ ความแปรปรวนของคะแนนเฉลี่ยจากจำนวนครั้งของการประเมิน

$\sigma_{i:o}^2$ คือ ความแปรปรวนของคะแนนจากจำนวนข้อสอบที่มาจาก การประเมินแต่ละครั้ง

σ_{po}^2	คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างนักเรียนกับจำนวนครั้งในการประเมิน
$\sigma_{pi:o,e}^2$	คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างนักเรียน, จำนวนข้อสอบ, ครั้งของการประเมิน และความคลาดเคลื่อนส่วนที่เหลือกับครั้งของการประเมิน

เขียนแผนภาพเวนนีได้ดังภาพที่ 7

ภาพที่ 7 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับสองฟาเซทแบบแฝง $p \times (i:o)$ (two-facet nested design, $p \times (i:o)$ design)



4.3.3 การออกแบบการวัดที่มีสามฟาเซท (three-facet design)

การออกแบบการวัดสามฟาเซทแบบแฝง (three-facet nested design, $p \times o \times (i:o)$ design)

สมมติต้องการพัฒนาคุณภาพของการประเมินนิสิตฝึกสอนจึงกำหนดสถานการณ์ดังนี้ ผู้สอบ จำนวน p คน ทำแบบสอบอัตนัยประยุกต์ o เหตุการณ์ แต่ละเหตุการณ์ ใช้ชุดข้อคำถามที่แตกต่างกัน i ข้อ โดยผู้ประเมิน r คน ตรวจทุกข้อคำถาม ดังนั้น ผลรวมขององค์ประกอบความแปรปรวนมีทั้งหมด 11 ส่วน ดังนี้

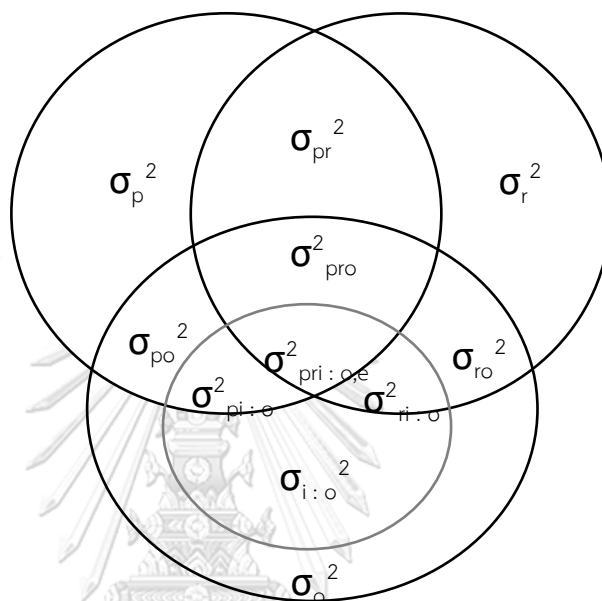
$$\sigma^2(X_{p \times r \times (i:o)}) = \sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{i:o}^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{pi:o}^2 + \sigma_{ro}^2 + \sigma_{ri:o}^2 + \sigma_{pro}^2 + \sigma_{pri:o,e}^2$$

เมื่อ

$\sigma^2(X_{p \times r \times (i:o)})$	คือ ความแปรปรวนของคะแนนสังเกตได้
σ_p^2	คือ ความแปรปรวนของคะแนนเอกภพของผู้สอบ
σ_r^2	คือ ความแปรปรวนของคะแนนเฉลี่ยจากผู้ประเมิน
σ_o^2	คือ ความแปรปรวนของคะแนนเฉลี่ยจากเหตุการณ์
$\sigma_{i:o}^2$	คือ ความแปรปรวนของคะแนนจากจำนวนข้อคำถาม ที่มาจาก เหตุการณ์เดียวกัน
σ_{pr}^2	คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่าง ผู้สอบกับผู้ประเมิน
σ_{po}^2	คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่าง ผู้สอบกับเหตุการณ์
$\sigma_{pi:o}^2$	คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่าง ผู้สอบกับข้อคำถามจากแต่ละเหตุการณ์
σ_{ro}^2	คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่าง ผู้ประเมินกับเหตุการณ์
$\sigma_{ri:o}^2$	คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่าง ผู้ประเมินกับข้อคำถามจากแต่ละเหตุการณ์
σ_{pro}^2	คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่าง ผู้สอบผู้ประเมิน และเหตุการณ์
$\sigma_{pri:o,e}^2$	คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่าง ผู้สอบผู้ประเมิน ข้อคำถามจากแต่ละเหตุการณ์ และความ คลาดเคลื่อนส่วนที่เหลือกับเหตุการณ์

เขียนแผนภาพเวนนีได้ดังภาพที่ 8

ภาพที่ 8 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวนสำหรับสามฟาเซทแบบแฝง
 $p \times o \times (i:o)$ (three-facet, $p \times r \times (i:o)$ design)



กล่าวโดยสรุปได้ว่าทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory: G-Theory เป็นทฤษฎีที่ขยายแนวคิดมาจากทฤษฎีการทดสอบแบบดั้งเดิม ดังนี้ 1) สามารถประมาณค่า ความแปรปรวนจากแหล่งต่าง ๆ ได้จากการวัดบริบทเดียว และจากการวิเคราะห์เพียงครั้งเดียว 2) สามารถประมาณค่าสัมประสิทธิ์ความเที่ยงโดยเน้นไปที่องค์ประกอบความแปรปรวนของแต่ละแหล่งความคลาดเคลื่อนของการวัด 3) สามารถแบ่งตัดสินใจออกเป็น 2 ลักษณะ คือ การตัดสินใจเชิงสัมพัทธ์ (relative decision) และการตัดสินใจเชิงสัมบูรณ์ (absolute decision) 4) สามารถประมาณน้ำหนักของแหล่งความคลาดเคลื่อนของการวัดที่สามารถเป็นไปได้ทั้งหมด (G-Study) และสามารถนำข้อมูลที่ได้มาใช้ในการออกแบบการวัดครั้งต่อไป เพื่อเป็นการลดความคลาดเคลื่อนของการวัดให้น้อยลง (D-Study) ซึ่งเป็นการออกแบบการวัดให้มีความคุ้มค่าทั้งค่าใช้จ่ายและเวลา และยังสามารถออกแบบการวัดได้หลายรูปแบบทั้งการออกแบบการวัดแบบไขว้ (crossed design) การออกแบบการวัดแบบแฝง (nested design) และกรณีการออกแบบการวัดแบบผสม (confounded design)

4.4 งานวิจัยที่เกี่ยวข้องกับการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

จากการศึกษางานวิจัยในประเทศและงานวิจัยต่างประเทศที่เกี่ยวข้องกับการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด มีดังนี้

พรพิมล ปิลันธนาภรณ์ และชิตชนก เชิงเชาว์ (2563) ศึกษาการพัฒนาแบบทดสอบอัตนัยประยุกต์ วัดความสามารถในการตัดสินใจทางคลินิกของนักศึกษาแพทยมหาวิทยาลัยสงขลานครินทร์ เครื่องมือที่ใช้ในงานวิจัยในครั้งนี้เป็นแบบทดสอบอัตนัยประยุกต์วัดความสามารถในการตัดสินใจทางคลินิกของนักศึกษาแพทย์ 5 เหตุการณ์ จำนวนรวม 35 ข้อ จากกลุ่มตัวอย่างเป็นนักศึกษาแพทย์ภาคเวชปฏิบัติ มหาวิทยาลัยสงขลานครินทร์ ปีการศึกษา 2559 จำนวน 127 คน วิเคราะห์และตรวจสอบคุณภาพของแบบทดสอบโดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดที่มีหนึ่งองค์ประกอบคือ จำนวนนักศึกษาแพทย์และจำนวนข้อคำถามที่ได้จาก D-Study พบว่า ข้อสอบมีค่าความยากง่ายตั้งแต่ 0.31 ถึง 0.79 ค่าอำนาจจำแนกมีค่าตั้งแต่ 0.21 ถึง 0.45 ค่าความเชื่อมั่นแบบสอดคล้องภายใน สัมประสิทธิ์แอลฟาเท่ากับ 0.701 ค่าความเที่ยงตรงเชิงเนื้อหา มีค่าดัชนีความสอดคล้อง (IOC) ตั้งแต่ 0.67 ถึง 1.00 ค่าความเที่ยงตรงเชิงสภาพ โดยพิจารณาคะแนนจากแบบทดสอบอัตนัยประยุกต์วัดความสามารถในการตัดสินใจทางคลินิกที่มีความสัมพันธ์กับคะแนนการฝึกภาคปฏิบัติอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 และผลประมาณค่าสัมประสิทธิ์เชิงสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมพันธ์ เท่ากับ 0.921 และค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์ เท่ากับ 0.896

ดรุณี อภัยกาวิ, ประกิตติยา ทักษิโณ และกมลวรรณ ตังธนานนท์ (2562) ศึกษาผลการตรวจให้คะแนนแบบสอบอัตนัย ประยุกต์ใช้โมเดลหลายองค์ประกอบของราล์ช และทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด เครื่องมือที่ใช้ในการวิจัยในครั้งนี้เป็นแบบสอบอัตนัยการคิดทางคณิตศาสตร์ จำนวน 3 ข้อ พบว่า ค่าความตรงตามสภาพของผลการตรวจให้คะแนนแบบสอบอัตนัย 3 ข้อ ของผู้ตรวจให้คะแนนที่ต่างกัน เทียบกับคะแนนเกณฑ์กลางซึ่งมาจากคะแนนอันดับการตรวจของผู้เชี่ยวชาญ พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ตรวจให้คะแนนเป็นกลางและเข้มงวดสูงกว่าผู้ตรวจให้คะแนนใจดี อีกทั้งค่าสัมประสิทธิ์การสรุปอ้างอิงในทุกคุณลักษณะของผู้ตรวจให้คะแนนที่ต่างกัน ภายใต้เงื่อนไขของรูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน $p \times (i:r)$ สูงกว่ารูปแบบการตรวจข้อสอบทุกข้อของผู้สอบทุกคน $p \times i \times r$

Lisa et al. (2020) ศึกษาความเที่ยงของเครื่องมือในการทำนายโอกาสในระยะเริ่มต้นของเด็กที่มีอาการทางโรคประสาทและกล้ามเนื้อในแต่ละกลุ่มอาการ โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด จากกลุ่มตัวอย่างที่เป็นผู้ป่วยเด็ก จำนวน 172 คน แบ่งออกเป็น 4 กลุ่มอาการ ได้แก่ กลุ่มอาการ Angelman syndrome, กลุ่มอาการ Prader-Willi syndrome, กลุ่มอาการ Williams syndrome และ กลุ่ม Row-risk control โดยใช้เครื่องมือแบบตรวจสอบรายการ CSBS-ITC จำนวน 24 ข้อวิเคราะห์และตรวจสอบคุณภาพของแบบทดสอบโดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดที่สององค์ประกอบ ได้แก่ จำนวนผู้ป่วย, อายุ และจำนวนข้อคำถาม ภายใต้แต่ละกลุ่มอาการ จาก D-study พบว่าเด็กที่อยู่ในกลุ่มอาการ Prader-Willi syndrome มีค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์สูงสุด ($\sigma^2_{Ab} = 0.685$) รองลงมาคือ ผู้ป่วยเด็กในกลุ่มอาการ Williams syndrome, กลุ่มอาการ Angelman syndrome และกลุ่ม Row-risk control ตามลำดับ ($\sigma^2_{Ab} = 0.672, 0.612, 0.556$) และการศึกษาเพื่อการตัดสินใจสรุปอ้างอิง โดยมีเงื่อนไขในการวัด คือ อายุ (3 เดือน, 6 เดือน, 9 เดือน และ 12 เดือน) และจำนวนข้อคำถาม พบว่าค่าความเที่ยงของเครื่องมือ CSBS-ITC สูงขึ้นเมื่อจำนวนข้อคำถามเพิ่มขึ้น

Quoc et al. (2020) ศึกษาการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดกับแบบวัด The Five Facets Mindfulness Questionnaire (FFMQ) ของผู้เข้าฝึกอบรมพัฒนาจิตจำนวน 83 คน เครื่องมือที่ใช้ในงานวิจัยเป็นแบบวัด FFMQ ซึ่งแต่ละฉบับจะ 5 องค์ประกอบ โดยวัดทั้งหมด 3 เหตุการณ์ วิเคราะห์และตรวจสอบคุณภาพของแบบทดสอบโดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดสององค์ประกอบ ได้แก่ จำนวนผู้เข้าฝึกอบรม, จำนวนครั้งในการวัดและจำนวนข้อในแต่ละองค์ประกอบ ที่ได้จาก D-Study พบว่าแบบวัด FFMQ มีค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมพัทธ์และค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์ไม่ต่ำกว่า 0.50 กรณีที่จำนวนข้อในแต่ละองค์ประกอบต้องไม่เกิน 5 ข้อ และการศึกษาเพื่อการตัดสินใจสรุปอ้างอิง โดยมีเงื่อนไขในการวัด คือ จำนวนครั้งในการวัด ได้แก่ 1, 2, 3 และจำนวนข้อในแต่ละองค์ประกอบ ได้แก่ 5, 6, 7, 8 พบว่า ค่าความเที่ยงของแบบวัด FFMQ สูงขึ้นเมื่อจำนวนครั้งในการวัดเพิ่มขึ้น แต่จำนวนข้อคำถามในแต่ละองค์ประกอบไม่ควรเกิน 5 ข้อ

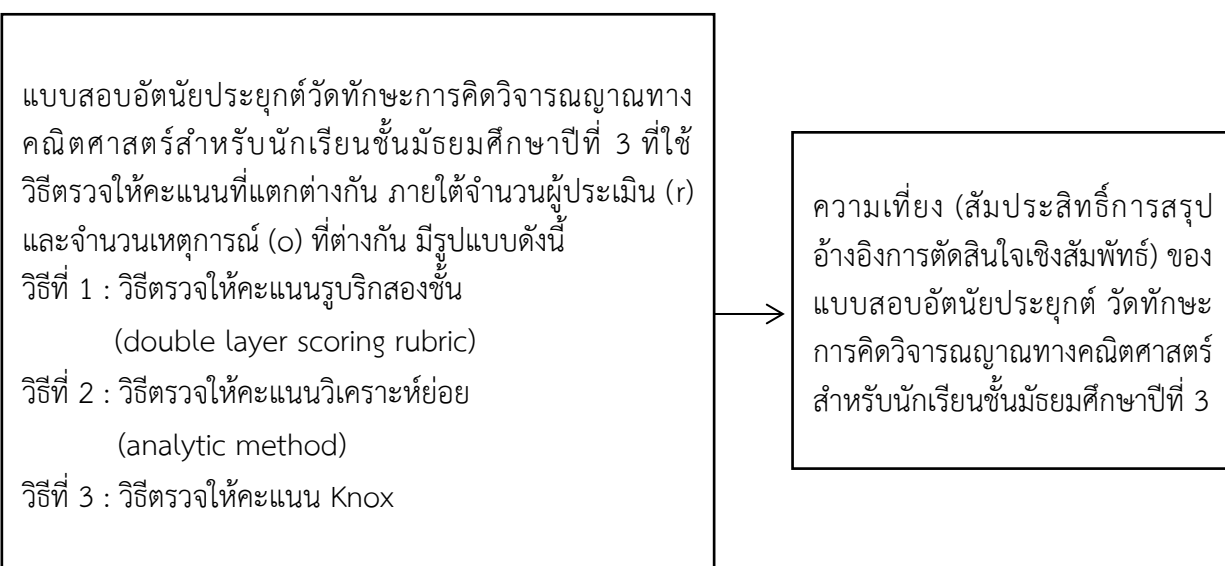
จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง พบว่า มีหลายงานวิจัยที่พัฒนาแบบสอบอัตโนมัติประยุกต์ โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดมาใช้ในการวิเคราะห์ และควบคุมแหล่งความคลาดเคลื่อน ส่วนใหญ่มีการออกแบบการวัด 1-2

องค์ประกอบ ได้แก่ จำนวนผู้ประเมิน หรือความยาวของข้อสอบ แต่ยังไม่พบงานวิจัยที่มีการออกแบบการวัด 2 องค์ประกอบ สำหรับการสร้างแบบสอบอัตนัยประยุกต์ในวิชาคณิตศาสตร์ที่ต้องพิจารณาทั้งจำนวนผู้ประเมิน และจำนวนข้อคำถามในแต่ละเหตุการณ์ที่วัดตามขั้นของการคิด วิจัยทางคณิตศาสตร์ผ่านการสังเคราะห์จากงานวิจัยที่ศึกษามาข้างต้น ผู้วิจัยจึงสนใจที่จะพัฒนาแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจจัยทางคณิตศาสตร์ โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง สำหรับการออกแบบ 2 องค์ประกอบ ได้แก่ จำนวนผู้ประเมิน (r) และจำนวนเหตุการณ์ที่ใช้ (o) ภายใต้การออกแบบการวัด $p \times r \times o$

ตอนที่ 5 กรอบแนวคิดในการวิจัย

จากการสังเคราะห์องค์ความรู้จากเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบสอบอัตนัยประยุกต์ ผู้วิจัยมุ่งเน้นการเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจจัยทางคณิตศาสตร์ ซึ่งถือว่าเป็นหนึ่งในทักษะการคิดขั้นสูงตามหลักสูตรฐานสมรรถนะที่ใช้วิธีตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric) โดยเปรียบเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) และวิธีตรวจให้คะแนน Knox ในแต่ละวิธีตรวจให้คะแนนนั้นจะศึกษาภายใต้จำนวนผู้ประเมินและจำนวนเหตุการณ์ที่แตกต่างกัน โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง เพื่อนำผลที่ได้ไปตัดสินใจการสรุปอ้างอิงแบบสอบอัตนัยประยุกต์ว่า วิธีตรวจให้คะแนนแต่ละวิธีนั้น ควรใช้จำนวนผู้ประเมินและจำนวนเหตุการณ์เท่าใด จึงจะได้แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจจัยทางคณิตศาสตร์ที่มีความเที่ยงในระดับที่ต้องการในการวัดผลแบบอิงกลุ่ม สามารถเขียนเป็นกรอบแนวคิดในการวิจัยได้ดังภาพที่ 9

ภาพที่ 9 กรอบแนวคิดในการวิจัย



บทที่ 3 วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาแบบสอบถามอัตโนมัติประยุกต์วัดทักษะการคิด
วิจารณ์ทางคณิตศาสตร์และตรวจสอบคุณภาพวิธีการตรวจให้คะแนนแบบรูบริกสองชั้นใน
แบบสอบถามอัตโนมัติ โดยเปรียบเทียบความเที่ยงของแบบสอบถามอัตโนมัติวัดทักษะการคิด
วิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูบริก
สองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนน วิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox โดย
ประยุกต์ใช้ทฤษฎีการสุ่มอ้างอิง ประกอบด้วย 2 ฟาเซท คือ จำนวนผู้ประเมิน และจำนวน
เหตุการณ์ ซึ่งมีรายละเอียดในการดำเนินการวิจัย ดังนี้

ขั้นตอนที่ 1 การสร้างและตรวจสอบคุณภาพของแบบสอบถามอัตโนมัติวัดทักษะการคิด
วิจารณ์ทางคณิตศาสตร์

1.1 กำหนดขอบเขตเนื้อหาวิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น

1.2 การเลือกเหตุการณ์ที่จะนำมาสร้างแบบสอบถามและข้อคำถาม

1.3 การกำหนดรูปแบบของแบบสอบถามอัตโนมัติวัดทักษะการคิดวิจารณ์ทาง
คณิตศาสตร์

1.4 การสร้างแนวคำตอบ เกณฑ์การให้คะแนน การตรวจให้คะแนนและการบันทึก
คะแนน

1.5 การหาความตรงเชิงเนื้อหา

1.6 การทดลองใช้แบบสอบถามอัตโนมัติวัดทักษะการคิดวิจารณ์ทาง
คณิตศาสตร์

ขั้นตอนที่ 2 การเก็บรวบรวมข้อมูล

2.1 การกำหนดประชากรและตัวอย่างวิจัย

2.2 การสุ่มตัวอย่างวิจัย

2.3 เครื่องมือที่ใช้ในงานวิจัย

2.4 การดำเนินการสอบ

ขั้นตอนที่ 3 การวิเคราะห์ข้อมูล

1.1 การวิเคราะห์ค่าสถิติพื้นฐานของผลการตรวจให้คะแนนแบบสอบถามอัตโนมัติ
วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

- 1.2 การวิเคราะห์คุณภาพของการตรวจให้คะแนนด้วยวิธีการให้คะแนนแบบรูปrik สองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทาง คณิตศาสตร์
- 1.3 การวิเคราะห์ความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกัน ภายใต้ จำนวนเหตุการณ์ และจำนวน ผู้ประเมินที่ต่างกัน โดยใช้ทฤษฎีการสรุปอ้างอิง (G-Theory)

ขั้นตอนที่ 1 การสร้างและตรวจสอบคุณภาพของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ทางคณิตศาสตร์

1.1 กำหนดขอบเขตเนื้อหาวิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษา

ตอนต้น

เนื้อหาวิชาคณิตศาสตร์ที่ใช้ในการศึกษาการวัดทักษะการคิดวิจารณ์ทาง คณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น ตามสาระทางคณิตศาสตร์ จำนวน 2 สาระ ได้แก่ จำนวนและพีชคณิต 1 เรื่อง คือ อัตราส่วน การวัดและเรขาคณิต 2 เรื่อง คือ ทฤษฎีบทพีทาโกรัส และพื้นที่ผิวและปริมาตร เนื่องจากสาระการวัดและเรขาคณิตอาศัยความรู้พื้นฐานจากสาระ จำนวนและพีชคณิต และประยุกต์เข้ากับสถานการณ์ในชีวิตจริงเพื่อส่งเสริมการคิดวิจารณ์ ซึ่งสามารถวิเคราะห์มาตรฐานการเรียนรู้ ตัวชี้วัด และสาระการเรียนรู้ทางคณิตศาสตร์ ดังตาราง 3

ตาราง 3 ผลการเรียนรู้และเนื้อหาวิชาคณิตศาสตร์กลุ่มสาระการเรียนรู้คณิตศาสตร์ หลักสูตร แกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 (ฉบับปรับปรุง พ.ศ. 2561)

เรื่อง	มาตรฐาน	ตัวชี้วัด	สาระการเรียนรู้
สาระจำนวนและพีชคณิต			
อัตราส่วน	ค 1.1 เข้าใจความ	ม.1/3 เข้าใจและประยุกต์ใช้	1. อัตราส่วนของจำนวน
หลากหลายของการแสดง		อัตราส่วน สัดส่วน และร้อยละ	หลาย ๆ จำนวน
จำนวน ระบบจำนวน การ		ละ ในการแก้ปัญหา	2. สัดส่วน
ดำเนินการของจำนวน ผล		คณิตศาสตร์ และปัญหาใน	3. การนำความรู้เกี่ยวกับ
ที่เกิดขึ้นจากการ		ชีวิตจริง	อัตราส่วน สัดส่วน และร้อยละ
ดำเนินการ สมบัติของการ			ละไปใช้ในการแก้ปัญหา
ดำเนินการ และนำไปใช้			

ตาราง 3 (ต่อ)

เรื่อง	มาตรฐาน	ตัวชี้วัด	สาระการเรียนรู้
สาระการวัดและเรขาคณิต			
พื้นที่ผิว และ ปริมาตร	ค 2.1 เข้าใจพื้นฐาน เกี่ยวกับการวัด วัดและ คาคคเนขนาดสิ่งที ต้องการวัดและนำไปใช้	ม.2/1 ประยุกต์ใช้ความรู้ เรื่องพื้นที่ผิวของปริซึม และ ทรงกระบอกในการ แก้ปัญหาคณิตศาสตร์และ ปัญหา ในชีวิตจริง	1. การหาพื้นที่ผิวของปริซึม และทรงกระบอก 2. การนำความรู้เกี่ยวกับ พื้นที่ผิวของปริซึม และ ทรงกระบอกใช้ในการ แก้ปัญหา
พื้นที่ผิว และ ปริมาตร	ค 2.1 เข้าใจพื้นฐาน เกี่ยวกับการวัด วัดและ คาคคเนขนาดสิ่งที ต้องการวัดและนำไปใช้	ม.2/2 ประยุกต์ใช้ความรู้ เรื่องปริมาตรของปริซึม และ ทรงกระบอกในการ แก้ปัญหาคณิตศาสตร์และ ปัญหา ในชีวิตจริง	1. การหาปริมาตรของปริซึม และทรงกระบอก 2. การนำความรู้เกี่ยวกับ ปริมาตรของปริซึม และ ทรงกระบอกไปใช้ ในการแก้ปัญหา
สาระการวัดและเรขาคณิต			
ทฤษฎีบท พิทาโกรัส และ บท พิทาโกรัส	ค 2.2 เข้าใจและวิเคราะห์ รูปเรขาคณิต สมบัติของรูป เรขาคณิต ความสัมพันธ์ ระหว่างรูปเรขาคณิต และ ทฤษฎีบททางเรขาคณิต และนำไปใช้	ม.2/5 เข้าใจและใช้ทฤษฎีบท พิทาโกรัส และ บทกลับในการแก้ปัญหา คณิตศาสตร์และปัญหาใน ชีวิตจริง	1. ทฤษฎีบทพีทาโกรัส และ บทกลับ 2. การนำความรู้เกี่ยวกับ ทฤษฎีบทพีทาโกรัส และบท กลับไปใช้ในชีวิตจริง

1.2 การเลือกเหตุการณ์ที่จะนำมาสร้างแบบสอบและข้อคำถาม

เนื้อหาวิชาคณิตศาสตร์ที่ผู้วิจัยได้กำหนดขอบเขตให้ครอบคลุมมาตรฐานการเรียนรู้ ตัวชี้วัด และสาระการเรียนรู้ทางคณิตศาสตร์ มีเนื้อหารายละเอียดที่แตกต่างกัน ซึ่งการสร้างแบบสอบอัตรณ์ประยุกต์ วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ เหตุการณ์ที่นำมาสร้างข้อคำถามควรเป็นเหตุการณ์ที่บูรณาการความรู้ทางคณิตศาสตร์และเชื่อมโยงกับบริบทที่ใกล้ชิดกับ ผู้เรียน สามารถพบเห็นได้ในชีวิตประจำวัน หรือสิ่งที่อยู่ในความสนใจของผู้เรียน เพื่อลดความ ได้เปรียบเสียเปรียบของผู้สอบในรายละเอียดของข้อมูลในแต่ละเหตุการณ์ที่นำมาสร้างแบบสอบ

(วีรวุฒิ ไทยขำ และ ทรงชัย อักษรคิด, 2564) ผู้วิจัยจึงได้กำหนดเหตุการณ์ให้ครอบคลุมเนื้อหาวิชาคณิตศาสตร์ โดยคัดเลือกเนื้อหาที่สนใจ และมีเนื้อหาเหมาะสมกับระดับชั้นของนักเรียนจากแหล่งข้อมูลต่าง ๆ เช่น หนังสือพิมพ์ ข่าว อินเทอร์เน็ต เป็นต้น โดยมีการปรับเนื้อหาให้เหมาะสมผ่านการตัดทอนหรือปรับข้อความ แต่ยังคงความหมายใจความเดิม

ผู้วิจัยนำเหตุการณ์มาสร้างข้อคำถามโดยออกแบบว่า เมื่อนักเรียนอ่านเหตุการณ์และข้อคำถามแล้ว นักเรียนสามารถตอบคำถามโดยใช้ความรู้และทักษะทางคณิตศาสตร์ รวมถึงการเลือกใช้ข้อมูลจากแต่ละเหตุการณ์ได้ ผู้วิจัยสร้างข้อคำถามของเหตุการณ์ตามลำดับสถานการณ์อย่างต่อเนื่อง โดยแทรกข้อคำถามที่ใช้วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ โดยผู้วิจัยสร้างแบบสอบอัตนัยประยุกต์ จำนวน 3 เหตุการณ์ และแต่ละเหตุการณ์มีข้อคำถาม จำนวน 5 ข้อ รวมทั้งแบบสอบมีข้อคำถาม จำนวน 15 ข้อ ซึ่งเป็นไปตามที่ Feletti (1980) เสนอไว้ว่า แบบสอบอัตนัยประยุกต์ควรมีข้อคำถาม จำนวน 5-35 ข้อ จึงจะสามารถวัดทักษะการคิดขั้นสูงได้

ข้อคำถามในแต่ละเหตุการณ์จะวัดขั้นตอนการคิดวิจารณ์ทางคณิตศาสตร์ทั้ง 5 ขั้นตอน คือ 1) ชั้นนิยามปัญหา 2) ชั้นตั้งสมมติฐาน 3) ชั้นรวบรวมข้อมูล 4) ชั้นวิเคราะห์ข้อมูล และ 5) ชั้นสร้างข้อสรุป โดยในข้อคำถามที่ 1 ของทุกเหตุการณ์จะวัดความสามารถในการนิยามปัญหา ข้อคำถามที่ 2 ของทุกเหตุการณ์จะวัดความสามารถในการตั้งสมมติฐาน ข้อคำถามที่ 3 ของทุกเหตุการณ์จะวัดความสามารถ ในการรวบรวมข้อมูล ข้อคำถามที่ 4 ของทุกเหตุการณ์จะวัดความสามารถในการวิเคราะห์ข้อมูล และข้อคำถามที่ 5 ของทุกเหตุการณ์ จะวัดความสามารถในการสร้างข้อสรุป ดังตาราง 4

ตาราง 4 โครงสร้างของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

ความสามารถ ที่จะวัด เหตุการณ์	นิยาม ปัญหา		ตั้งสมมติฐาน	รวบรวม ข้อมูล	วิเคราะห์ ข้อมูล	สร้าง ข้อสรุป
เหตุการณ์ที่ 1	ข้อคำถาม ที่ 1		ข้อคำถามที่ 2	ข้อคำถาม ที่ 3	ข้อคำถามที่ 4	ข้อคำถาม ที่ 5
เหตุการณ์ที่ 2	ข้อคำถาม ที่ 1		ข้อคำถามที่ 2	ข้อคำถาม ที่ 3	ข้อคำถามที่ 4	ข้อคำถาม ที่ 5
เหตุการณ์ที่ 3	ข้อคำถาม ที่ 1		ข้อคำถามที่ 2	ข้อคำถาม ที่ 3	ข้อคำถามที่ 4	ข้อคำถาม ที่ 5

1.3 การกำหนดรูปแบบของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

ผู้วิจัยออกแบบรูปแบบของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ตามรูปแบบของมาชนพันธ์ อ๋านาคิล และกมลวรรณ ตั้งธนกานนท์ (2557) และสุกชิต ผดุงผล และกมลวรรณ ตั้งธนกานนท์ (2559) แล้วนำแนวคิดเหล่านั้นมาสร้างแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยมีรายละเอียดดังต่อไปนี้

1.3.1 การกำหนดรูปแบบของแบบสอบอัตนัยประยุกต์

ผู้วิจัยออกแบบรูปแบบของแบบสอบอัตนัยประยุกต์ตามรูปแบบของสุกชิต ผดุงผล และกมลวรรณ ตั้งธนกานนท์ (2559) โดยแบ่งแบบสอบออกเป็น 3 ส่วน ดังนี้ ส่วนที่ 1 คือ ส่วนด้านบนของกระดาษ ซึ่งแบ่งออกเป็น 3 ช่องย่อย ช่องซ้าย คือ เวลาที่ใช้ในการทำเหตุการณ์นี้ ช่องกลาง คือ เลขที่ของเหตุการณ์ในแบบสอบและข้อคำถามในเหตุการณ์นั้น ช่องขวา คือ ส่วนที่นักเรียนกรอกชื่อ นามสกุล ชั้น และเลขที่ ส่วนที่ 2 คือ ส่วนที่อยู่ถัดลงมาจากส่วนที่ 1 เป็นส่วนที่ใช้แสดงสถานการณ์ ข้อคำถาม และวันที่ไว้สำหรับให้ตอบคำถาม และส่วนที่ 3 คือ ส่วนที่แสดงคำเตือนในการทำแบบสอบอัตนัยประยุกต์ ดังตัวอย่างต่อไปนี้

ตัวอย่างแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

เวลาสำหรับทำเหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ ... ข้อคำถามที่ ...	ชื่อ-สกุล ชั้น เลขที่
สถานการณ์		
ข้อคำถามข้อที่ ... คำถาม : คำตอบ		
<p><u>ข้อควรระวัง</u> ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>		

1.3.2 การจัดพิมพ์แบบสอบอัตโนมัติประยุกต์

ผู้วิจัยดำเนินการพิมพ์เหตุการณ์ลงในกระดาษขนาด A4 โดยในแต่ละเหตุการณ์ จะพิมพ์ข้อความละ 1 แผ่น และเพื่อไม่ให้ผู้สอบเกิดความสับสนในแต่ละเหตุการณ์ ผู้วิจัยได้พิมพ์ชื่อเหตุการณ์คั่นระหว่างเหตุการณ์หนึ่ง ๆ

1.3.3 การพิมพ์คำแนะนำในการทำแบบสอบอัตโนมัติประยุกต์ โดยมีรายละเอียดดังนี้

คำแนะนำในการทำแบบสอบอัตโนมัติประยุกต์
วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

1. ข้อสอบนี้มีทั้งหมด 3 เหตุการณ์ โดยแต่ละเหตุการณ์จะมีข้อความ จำนวน 5 ข้อ คำถาม รวมข้อความทั้งสิ้น จำนวน 15 ข้อ รวม 15 หน้า
2. มีเวลาในการทำแบบสอบทั้งหมด 1 ชั่วโมง โดยแบ่งเวลาในการทำแต่ละเหตุการณ์ 20 นาที ซึ่งอยู่รวมบนซ้ายของแบบสอบอัตโนมัติประยุกต์ ผู้สอบต้องจัดสรรเวลาในการทำแบบสอบอัตโนมัติประยุกต์เอง และทำข้อสอบทุกข้อให้เสร็จทันในเวลาที่กำหนด
3. นักเรียนต้องกรอกชื่อ นามสกุล ชั้น และเลขที่ ที่มุมบนขวาของกระดาษคำถามทุกแผ่น
4. เริ่มทำข้อสอบตั้งแต่เหตุการณ์ที่ 1 ข้อคำถามที่ 1 เมื่อทำเสร็จให้นักเรียนตรวจสอบให้เรียบร้อย แล้วส่งข้อสอบโดยคว่ำกระดาษลง วางไว้ใต้ช่องใส่แบบสอบอัตโนมัติประยุกต์ และไปลงเวลาให้กลับแบบที่ 1๑หรือ 1๑๑ข้อคำถามถัดไป

ผู้วิจัยเรียงหน้าของแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยเริ่มต้นที่คำแนะนำในการทำแบบสอบอัตโนมัติประยุกต์ จากนั้นจึงเรียงเหตุการณ์ของแบบสอบจากเหตุการณ์ที่ 1 จนถึงเหตุการณ์ที่ 3

1.3.5 การบรรจุแบบสอบอัตโนมัติประยุกต์

ผู้วิจัยบรรจุแบบสอบอัตโนมัติประยุกต์ลงในซองสีน้ำตาลที่มีขนาดของซองเท่ากับกระดาษ A5 เนื่องจากต้องการอำนวยความสะดวกให้กับนักเรียนเวลาดึงแบบสอบออกจากซองข้อสอบ และพิมพ์ชื่อแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ พร้อมทั้งเลขที่ตามลำดับที่หน้าซองของแบบสอบทุกซอง

1.4 การสร้างแนวคำตอบ เกณฑ์การให้คะแนน การตรวจให้คะแนนและการบันทึก

คะแนน

1.4.1 การสร้างแนวคำตอบ

ผู้วิจัยสร้างแนวคำตอบทันทีหลังจากที่สร้างข้อคำถามเสร็จ โดยคำตอบในแต่ละข้อคำถามต้องสามารถแสดงให้เห็นถึงขั้นตอนในการแก้ปัญหาที่ต้องการวัดในข้อนั้น ๆ และเป็นคำตอบที่แสดงให้เห็นถึงความรู้และความสามารถในทักษะการคิดวิเคราะห์ทางคณิตศาสตร์ด้วย

1.4.2 เกณฑ์การให้คะแนน

งานวิจัยในครั้งนี้ใช้วิธีตรวจให้คะแนน 3 วิธี คือ วิธีที่ 1 วิธีตรวจให้คะแนนแบบรูปกรีกสองชั้น (double layer scoring rubric) วิธีที่ 2 วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) และวิธีที่ 3 วิธีตรวจให้คะแนน Knox ซึ่งเกณฑ์การให้คะแนนทั้ง 3 วิธี ของแต่ละข้อคำถามในทุกเหตุการณ์ที่วัดขั้นตอน การคิดวิเคราะห์ทางคณิตศาสตร์ 5 ขั้นตอน จะมีประเด็นในการให้คะแนนที่เหมือนกัน แต่มีลักษณะของรายละเอียดคำอธิบายการให้คะแนนที่ต่างกัน ดังนี้

วิธีที่ 1 วิธีตรวจให้คะแนนแบบรูปกรีกสองชั้น หมายถึง วิธีการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์กับเกณฑ์การให้คะแนนแบบรูปกรีก 2 ชั้น โดยแบ่งออกเป็น 2 ระดับชั้น ชั้นที่ 1 อธิบายตามองค์ประกอบ และในแต่ละองค์ประกอบจะประเมินตามข้อรายการย่อย ดังนี้ 1) ความถูกต้องของเนื้อหา (4 คะแนน) แบ่งออกเป็น 1.1) ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง (2 คะแนน) 1.2) อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง (2 คะแนน) 2) การเข้าใจสิ่งที่ต้องการวัด (4 คะแนน) แบ่งออกเป็น 2.1) กำหนดขั้นของการวัดได้ถูกต้อง (2 คะแนน) 2.2) บริบทขั้นของการวัดได้อยู่ภายใต้ข้อมูลในสถานการณ์ (2 คะแนน) และ 3) การสื่อความหมายทางคณิตศาสตร์ (4 คะแนน) แบ่งออกเป็น 3.1) เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ได้ถูกต้อง (2 คะแนน) 3.2) คำนวณได้ถูกต้อง (2 คะแนน) โดยในแต่ละข้อรายการย่อยใช้มาตราประมาณค่า (rating scale) 3 ระดับ ได้แก่ 2 คะแนน คือ ถูกต้องทั้งหมด, 1 คะแนน คือ ถูกต้องบางส่วน และ 0 คะแนน คือ ไม่ถูกต้องทั้งหมด หรือไม่มีการตอบคำถาม จากนั้นนำผลรวมของคะแนนในระดับชั้นที่ 1 มาแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้

วิธีที่ 2 วิธีตรวจให้คะแนนวิเคราะห์ย่อย หมายถึง วิธีการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ โดยมีองค์ประกอบของเกณฑ์การให้คะแนนข้อคำถามที่ 1 นิยามปัญหา ข้อคำถามที่ 2 ตั้งสมมติฐาน และข้อคำถามที่ 3 รวบรวมข้อมูล ของแต่ละ

เหตุการณ์มี 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) และในข้อคำถามที่ 4 วิเคราะห์ข้อมูล และข้อคำถามที่ 5 สร้างข้อสรุปเพิ่มอีกหนึ่งองค์ประกอบ คือ การสื่อความหมายทางคณิตศาสตร์ (2 คะแนน) ซึ่งในแต่ละองค์ประกอบมีเกณฑ์การให้คะแนนเป็น 3 ระดับเช่นเดียวกันกับวิธีตรวจให้คะแนน Knox

โดยความถูกต้องของเนื้อหา หมายถึง การระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์, การเข้าใจในสิ่งที่ต้องการวัด หมายถึง การกำหนดขั้นของการวัดภายใต้บริบทของสถานการณ์ และการสื่อความหมายทางคณิตศาสตร์ หมายถึง การเปลี่ยนประโยคภาษาเป็นประโยค ทางคณิตศาสตร์ และคำนวณได้อย่างถูกต้อง

วิธีที่ 3 วิธีตรวจให้คะแนน Knox หมายถึง วิธีการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ โดยมีองค์ประกอบของเกณฑ์การให้คะแนนแต่ละข้อคำถามของแต่ละเหตุการณ์เป็น 3 ระดับ ได้แก่ 2 คะแนน คือ ผู้ตอบตอบคำถามได้ถูกต้องและครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ 1 คะแนน คือ ผู้ตอบตอบคำถามผิดบางส่วนหรือไม่ครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ และ 0 คะแนน คือ ผู้ตอบตอบคำถามไม่ถูกต้องตามแนวคำตอบ หรือไม่มีการตอบคำถาม

1.4.3 การตรวจให้คะแนน

งานวิจัยในครั้งนี้มีผู้ประเมินให้คะแนน คือ ผู้ประเมินที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน ซึ่งก่อนที่ผู้ประเมินทั้งสองจะทำการตรวจแบบสอบอัตนัยประยุกต์ ผู้วิจัยได้อธิบายวิธีการตรวจ เกณฑ์การให้คะแนนในแต่ละวิธี รวมถึงตัวอย่างการตรวจให้คะแนนแก่ผู้ประเมินทั้งสอง เพื่อลดความคลาดเคลื่อนอันเนื่องมาจากการให้คะแนนของผู้ประเมิน พร้อมทั้งให้คำแนะนำในการตรวจให้คะแนน ดังนี้

1) ตรวจคำตอบทีละข้อของนักเรียนทุกคน นั่นคือ ถ้าจะตรวจคำตอบข้อคำถามที่ 1 ของเหตุการณ์ที่ 1 ก็ให้ตรวจข้อคำถามที่ 1 ของเหตุการณ์ที่ 1 ของนักเรียนทุกคนจนจบ แล้วจึงเริ่มตรวจข้อคำถามที่ 2 ของเหตุการณ์ที่ 1 โดยผู้ประเมินห้ามตรวจข้อสอบให้เสร็จเป็นคน ๆ

2) ให้ความคงเส้นคงวาในการตรวจ เกณฑ์การให้คะแนนจะเหมือนกัน สำหรับทุก ๆ คำตอบ ครูบางคนอาจจะให้คะแนนค่อนข้างน้อยสำหรับคำตอบแผ่นแรก ๆ แต่จะให้คะแนนมากขึ้นสำหรับคำตอบแผ่นหลัง ๆ ทั้ง ๆ ที่คำตอบเหล่านั้นเป็นแนวเดียวกัน ดังนั้น ผู้ประเมินต้องอ่านคำตอบและประเมินค่าคุณภาพคำตอบไว้ล่วงหน้าก่อนตรวจจริง

3) เวลาตรวจห้ามดูชื่อนักเรียน เพื่อป้องกันไม่ให้เกิดอคติต่อการให้คะแนน เวลาตรวจ ผู้ประเมินห้ามนำลายมือของนักเรียนมาสัมผัสต่อการตรวจให้คะแนน พยายามตรวจข้อคำถามใด ข้อคำถามหนึ่งของนักเรียนทุกคนให้เสร็จ ไม่ควรหยุดพักจนกว่าจะตรวจข้อคำถามนั้นเสร็จแล้ว

1.4.4 การบันทึกคะแนน

เมื่อผู้ประเมินตรวจให้คะแนนจะต้องบันทึกคะแนนลงในแบบบันทึกคะแนน โดยถ้าตรวจให้คะแนนด้วยวิธีตรวจให้คะแนนรูปริกสองชั้น โดยก่อนการแปลงระดับคะแนนในชั้นที่ 1 ของข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 จะมีคะแนนเต็ม ข้อละ 8 คะแนน หลังแปลงคะแนนในชั้นที่ 2 แล้วแต่ละข้อคำถามจะมีคะแนนเต็ม 2 คะแนน ส่วนข้อคำถามที่ 4 และข้อคำถามที่ 5 จะมีคะแนนเต็ม ข้อละ 12 คะแนน หลังแปลงคะแนนในชั้นที่ 2 แล้วแต่ละข้อคำถามจะมีคะแนนเต็ม 3 คะแนน รวมคะแนนเต็มแต่ละเหตุการณ์ 12 คะแนน และถ้าพิจารณาคะแนนรวมทั้งฉบับ คือ 3 เหตุการณ์ แบบสอบอัตนัยประยุกต์นี้จะมีคะแนนเต็ม 36 คะแนน

ถ้าตรวจให้คะแนนด้วยวิธีตรวจให้คะแนนวิเคราะห์ย่อย ข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 จะมีคะแนนเต็ม ข้อละ 4 คะแนน ส่วนข้อคำถามที่ 4 และข้อคำถามที่ 5 จะมีคะแนนเต็ม ข้อละ 6 คะแนน รวมคะแนนเต็มแต่ละเหตุการณ์ 24 คะแนน และถ้าพิจารณาคะแนนรวมทั้งฉบับ คือ 3 เหตุการณ์ แบบสอบอัตนัยประยุกต์นี้จะมีคะแนนเต็ม 72 คะแนน

ถ้าตรวจให้คะแนนด้วยวิธีตรวจให้คะแนนตามแบบ Knox แต่ละข้อคำถามจะมีคะแนนเต็ม ข้อละ 2 คะแนน แต่ละเหตุการณ์จะมีคะแนนเต็ม 10 คะแนน และถ้าพิจารณาคะแนนรวมทั้งฉบับ คือ 3 เหตุการณ์ แบบสอบอัตนัยประยุกต์นี้จะมีคะแนนเต็ม 30 คะแนน

สำหรับแบบบันทึกคะแนนของวิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) และวิธีตรวจให้คะแนน Knox จะปรากฏอยู่ในภาคผนวก จ และภาคผนวก ข

1.5 การหาความตรงเชิงเนื้อหา

เมื่อผู้วิจัยสร้างแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ พร้อมเกณฑ์การให้คะแนนแล้ว ผู้วิจัยนำไปให้อาจารย์ที่ปรึกษาวิทยานิพนธ์เพื่อปรับปรุงแก้ไข จากนั้นนำแบบสอบอัตนัยประยุกต์ฉบับนี้ไปให้ผู้ทรงคุณวุฒิ จำนวน 3 คน เป็นผู้

ตรวจสอบความตรงเชิงเนื้อหาของเหตุการณ์ และเกณฑ์การให้คะแนน พร้อมทั้งให้ผู้เชี่ยวชาญเป็นผู้กำหนดแนวคำตอบ โดยกำหนดคุณสมบัติผู้ทรงคุณวุฒิไว้ ดังนี้

1.5.1 ผู้ทรงคุณวุฒิทางหลักสูตรและการสอนคณิตศาสตร์ คือ ผู้ที่มีคุณวุฒิทางการศึกษาระดับปริญญาเอกด้านการศึกษาคณิตศาสตร์ และมีความรู้และประสบการณ์ทางการสอนคณิตศาสตร์ ในระดับอุดมศึกษาไม่น้อยกว่า 5 ปี จำนวน 1 คน

1.5.2 ผู้ทรงคุณวุฒิทางการสอนคณิตศาสตร์ คือ ผู้ที่มีประสบการณ์ด้านการสอนคณิตศาสตร์ในระดับมัธยมศึกษาไม่น้อยกว่า 5 ปี จำนวน 2 คน

การหาความตรงเชิงเนื้อหา ผู้วิจัยสร้างแบบฟอร์มการประเมินความตรงเชิงเนื้อหา และ ให้ผู้ทรงคุณวุฒิ จำนวน 3 คน ประเมินว่าข้อคำถามของเหตุการณ์ในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิเคราะห์ทางคณิตศาสตร์สอดคล้องกับมาตรฐานการเรียนรู้ตัวชี้วัด และสาระการเรียนรู้ทางคณิตศาสตร์หรือไม่ แนวคำตอบมีความสอดคล้องกับข้อคำถามและถูกต้องหรือไม่ และเกณฑ์การให้คะแนนมีความสอดคล้องกับแนวคำตอบหรือไม่ โดยกำหนดระดับความสอดคล้องของข้อคำถามกับกระบวนการคิดวิเคราะห์ทางคณิตศาสตร์ ดังนี้ 1 = สอดคล้อง 0 = ไม่แน่ใจว่าสอดคล้อง และ -1 = ไม่สอดคล้อง จากนั้นนำมาหาค่าดัชนีความสอดคล้อง (Item-Objective Congruence: IOC) โดยกำหนดเกณฑ์ความสอดคล้องของผู้ทรงคุณวุฒิ ≥ 0.5 สำหรับข้อคำถามใดที่มีค่าระดับความสอดคล้องของผู้ทรงคุณวุฒิ และสำหรับข้อคำถามใดที่มีค่าดัชนีความสอดคล้อง < 0.5 ผู้วิจัยจะนำมาปรับปรุงตามข้อเสนอแนะของผู้ทรงคุณวุฒิต่อไป

ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิเคราะห์ทางคณิตศาสตร์ โดยวิเคราะห์ความสอดคล้องของข้อคำถาม ที่ผู้วิจัยพัฒนาขึ้น ทั้ง 3 เหตุการณ์ ซึ่งแต่ละเหตุการณ์ประกอบด้วยข้อคำถามย่อย 5 ข้อคำถาม รวมทั้งสิ้น 15 ข้อคำถาม กับขั้นของกระบวนการคิดวิเคราะห์ทางคณิตศาสตร์ พบว่า เมื่อเทียบกับเกณฑ์ ≥ 0.5 ทุกข้อคำถาม แสดงว่า ข้อคำถามมีความสอดคล้องกับเนื้อหา และขั้นของกระบวนการคิดวิเคราะห์ทางคณิตศาสตร์ และเมื่อพิจารณาคำแนะนำที่ได้จากผู้เชี่ยวชาญได้ข้อมูล โดยสรุปว่า ข้อสอบ ที่นำมาสร้างข้อคำถามควรมีการปรับภาษาให้มีความชัดเจนทุกข้อ และสอดคล้องกับความเป็นจริงมากขึ้น โดยเหตุการณ์ที่ 1 ขั้น 3) ขั้นรวบรวมข้อมูล สามารถตัดรายละเอียดข้อมูลที่ไม่จำเป็นออกได้ เพื่อลดความสับสนให้กับผู้สอบ ดังตาราง 5

ตาราง 5 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถามระดับประยุกต์วัดทักษะการคิด
 วิจัยญาณทางคณิตศาสตร์ โดยวิเคราะห์ความสอดคล้องของข้อคำถามกับขั้นของกระบวนการ
 คิดวิจัยญาณ

เหตุการณ์ (O)	ข้อคำถาม (I)	IOC	ข้อเสนอแนะ
1	1.1 (ขั้นนิยามปัญหา)	0.84	ปรับภาษาที่ใช้ให้มีความชัดเจนมากขึ้น
	1.2 (ขั้นตั้งสมมติฐาน)	1.00	-
	1.3 (ขั้นรวบรวมข้อมูล)	0.84	ควรตัดรายละเอียดข้อมูลที่ไม่จำเป็นออก
	1.4 (ขั้นวิเคราะห์ข้อมูล)	1.00	-
	1.5 (ขั้นสร้างข้อสรุป)	1.00	-
2	2.1 (ขั้นนิยามปัญหา)	0.84	ปรับตัวละครในสถานการณ์ให้มีความ สอดคล้องกับความจริงมากขึ้น
	2.2 (ขั้นตั้งสมมติฐาน)	1.00	-
	2.3 (ขั้นรวบรวมข้อมูล)	1.00	-
	2.4 (ขั้นวิเคราะห์ข้อมูล)	1.00	ปรับนัยสำคัญของตัวเลขให้สอดคล้องกับ ข้อมูลในสถานการณ์ทั้งเหตุการณ์
	2.5 (ขั้นสร้างข้อสรุป)	1.00	-
3	3.1 (ขั้นนิยามปัญหา)	1.00	ปรับภาษาให้ทางการมากขึ้น
	3.2 (ขั้นตั้งสมมติฐาน)	1.00	-
	3.3 (ขั้นรวบรวมข้อมูล)	1.00	-
	3.4 (ขั้นวิเคราะห์ข้อมูล)	1.00	ปรับภาษาให้ทางการมากขึ้น
	3.5 (ขั้นสร้างข้อสรุป)	1.00	-

ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถามอัตโนมัติประยุกต์วัดทักษะการคิด
 วิจัยทางคณิตศาสตร์ โดยวิเคราะห์ความสอดคล้องของแนวคำตอบ ที่ผู้วิจัยพัฒนาขึ้นทั้ง
 3 เหตุการณ์ ซึ่งแต่ละเหตุการณ์ประกอบด้วยข้อคำถามย่อย 5 ข้อคำถาม รวมทั้งสิ้น 15 ข้อคำถาม
 กับชั้นของกระบวนการคิดวิจัยทางคณิตศาสตร์ พบว่า เมื่อเทียบกับเกณฑ์ ≥ 0.5 ทุกข้อคำถาม แสดงว่า
 ข้อคำถามมีความสอดคล้องกับเนื้อหา และชั้นของกระบวนการคิดวิจัยทางคณิตศาสตร์ และเมื่อพิจารณา
 คำแนะนำที่ได้จากผู้เชี่ยวชาญได้ข้อมูล โดยสรุปว่า ข้อสอบที่นำมาสร้างข้อคำถามควรมีการปรับ
 ปรุงภาษาให้มีความชัดเจน โดยเหตุการณ์ที่ 2 ชั้น 2) ชั้นตั้งสมมติฐาน ควรปรับเนื้อหาในแนวคำตอบ
 ให้ใช้ความรู้ในระดับชั้นมัธยมศึกษาตอนต้น เพื่อให้สอดคล้องกับระดับความสามารถของผู้สอบ
 ดังตาราง 6

ตาราง 6 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถามอัตโนมัติประยุกต์วัดทักษะการคิด
 วิจัยทางคณิตศาสตร์โดยวิเคราะห์ความสอดคล้องของแนวคำตอบกับชั้นของ
 กระบวนการคิดวิจัยทางคณิตศาสตร์

เหตุการณ์ (O)	ข้อคำถาม (I)	IOC	ข้อเสนอแนะ
1	1.1 การนิยามปัญหา	0.67	ปรับภาษาที่ใช้ให้มีความชัดเจนมากขึ้น
	1.2 การตั้งสมมติฐาน	1.00	-
	1.3 การรวบรวมข้อมูล	1.00	-
	1.4 การวิเคราะห์ข้อมูล	1.00	-
	1.5 การสร้างข้อสรุป	1.00	-
2	2.1 การนิยามปัญหา	1.00	-
	2.2 การตั้งสมมติฐาน	1.00	ปรับภาษาที่ใช้ให้มีความชัดเจนมากขึ้น และปรับเนื้อหาในแนวคำตอบให้ใช้ ความรู้ในระดับชั้นมัธยมศึกษาตอนต้น
	2.3 การรวบรวมข้อมูล	1.00	-
	2.4 การวิเคราะห์ข้อมูล	1.00	-
	2.5 การสร้างข้อสรุป	1.00	-

ตารางที่ 6 (ต่อ)

เหตุการณ์ (O)	ข้อคำถาม (I)	IOC	ข้อเสนอแนะ
	3.1 การนิยามปัญหา	1.00	-
	3.2 การตั้งสมมติฐาน	1.00	-
3	3.3 การรวบรวมข้อมูล	1.00	-
	3.4 การวิเคราะห์ข้อมูล	1.00	-
	3.5 การสร้างข้อสรุป	1.00	-

ผลการตรวจสอบความตรงเชิงเนื้อหาของแนวคำตอบกับวิธีตรวจให้คะแนนรูปกรีกสองชั้น (double layer scoring rubric) พบว่า วิธีตรวจให้คะแนนรูปกรีกสองที่ผู้วิจัยพัฒนาขึ้นเพื่อใช้ตรวจให้คะแนนข้อคำถามของเหตุการณ์ในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ ทุกข้อมีระดับความสอดคล้องกับข้อคำถามและแนวคำตอบเมื่อเทียบกับเกณฑ์ ≥ 0.5 แสดงว่า ข้อคำถามมีความสอดคล้องกับเนื้อหา และขั้นของกระบวนการคิดวิจารณ์ญาณ และเมื่อพิจารณาคำแนะนำที่ได้จากผู้เชี่ยวชาญได้ข้อมูลโดยสรุปว่า ทุกเหตุการณ์ในความหมายของเกณฑ์ให้คะแนน 0 คะแนน ควรปรับภาษา จาก 'ไม่ถูกต้องทั้งหมด' เป็น "ไม่ถูกต้อง" โดยเหตุการณ์ที่ 1 ชั้น 4) การวิเคราะห์ข้อมูล ในด้านองค์ประกอบของความถูกต้องของเนื้อหา ควรมีการปรับคำอธิบายของเกณฑ์ให้กระชับและครอบคลุมมากขึ้น เหตุการณ์ที่ 2 ชั้น 2) การตั้งสมมติฐาน ในด้านองค์ประกอบของการเข้าใจสิ่งที่ต้องการวัด ควรปรับเนื้อหาในแนวคำตอบให้ใช้ความรู้ในระดับชั้นมัธยมศึกษาตอนต้น เพื่อให้สอดคล้องกับระดับความสามารถของผู้สอบ และชั้น 4) การวิเคราะห์ข้อมูล ในด้านองค์ประกอบของการสื่อความหมายทางคณิตศาสตร์ ควรเพิ่มกรณีแนวคำตอบทุกกรณีที่เป็นไปได้ และเหตุการณ์ที่ 3 ชั้น 4) การวิเคราะห์ข้อมูล ในด้านองค์ประกอบของการสื่อความหมายทางคณิตศาสตร์ ควรปรับคำอธิบายของ ข้อรายการย่อย ให้สอดคล้องกับการสื่อความหมายทางคณิตศาสตร์ ดังตาราง 7

ตาราง 7 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถามฉบับต้นนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ โดยวิธีตรวจให้คะแนนรูปบิกสองชั้น (double layer scoring rubric)

เหตุการณ์ (O)	ข้อคำถาม (I)	ความตรงเชิงเนื้อหาของวิธีตรวจให้คะแนนรูปบิกสองชั้น					
		ความถูกต้องของเนื้อหา	การเข้าใจในสิ่งที่ต้องการวัด	การสื่อความหมายทางคณิตศาสตร์	IOC	IOC	IOC
1	1.1 การนิยามปัญหา	0.8	0.8	-	ไม่ถูกต้องประเภทอื่น	-	-
	1.2 การตั้งสมมติฐาน	0.8	0.8	-	ไม่ถูกต้องประเภทอื่น	-	-
	1.3 การรวบรวมข้อมูล	0.8	0.8	-	ไม่ถูกต้องประเภทอื่น	-	-
	1.4 การวิเคราะห์ข้อมูล	0.8	0.8	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง' ทั้งหมด เป็น 'ไม่ถูกต้อง'	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้ครอบคลุมทุกประเด็น
	1.5 การสร้างข้อสรุป	0.8	0.8	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง' ทั้งหมด เป็น 'ไม่ถูกต้อง'	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง' ทั้งหมด เป็น 'ไม่ถูกต้อง'

ตาราง 7 (ต่อ)

เหตุการณ์ (O)	ข้อคำถาม (I)	ความตรงเชิงเนื้อหาของวิธีตรวจให้คะแนนรูปrikสองชั้น			
		ความถูกต้องของเนื้อหา	การเข้าใจในสิ่งที่ต้องการวัด	การสื่อความหมายทาง คณิตศาสตร์	ข้อเสนอแนะ
		IOC	IOC	IOC	IOC
2.1	การนิยาม ปัญหา	0.8	0.8	-	ไม่ถูกต้องครอบคลุมนี้
	การ ตั้งสมมติฐาน	0.8	0.8	-	ไม่ถูกต้องครอบคลุมนี้
2.2	การรวบรวม ข้อมูล	0.8	0.8	-	ไม่ถูกต้องครอบคลุมนี้
	การวิเคราะห์	0.8	0.8	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้ ครอบคลุมทุกประเด็น
2.5	การสร้าง ข้อสรุป	0.8	0.8	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง ทั้งหมด' เป็น 'ไม่ถูกต้อง'

ตาราง 7 (ต่อ)

เหตุการณ์ (O)	ความตรงเชิงเนื้อหาของวิธีตรวจให้คะแนนรัฐบริการของชั้น					
	ข้อคำถาม (I)	ความถูกต้องของเนื้อหา	การเข้าใจในสิ่งที่ต้องการวัด	การสื่อความหมายทาง คณิตศาสตร์	IOC	ข้อเสนอแนะ
	IOC	ข้อเสนอแนะ	IOC	ข้อเสนอแนะ	IOC	ข้อเสนอแนะ
3	3.1 การนิยาม ปัญหา	0.8	0.8	-	-	ไม่ถูกต้องประกอบนี้
	3.2 การ ตั้งสมมติฐาน	0.8	0.8	-	-	ไม่ถูกต้องประกอบนี้
	3.3 การรวบรวม ข้อมูล	0.8	0.8	0.8	0.8	ไม่ถูกต้องประกอบนี้
	3.4 การวิเคราะห์ ข้อมูล	0.8	0.8	0.8	0.8	ไม่ถูกต้องประกอบนี้
	3.5 การสร้าง ข้อสรุป	0.8	0.8	0.8	0.8	ไม่ถูกต้องประกอบนี้

ผลการตรวจสอบความตรงเชิงเนื้อหาของแนวคำตอบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) พบว่า วิธีตรวจให้คะแนนวิเคราะห์ย่อยที่ผู้วิจัยพัฒนาขึ้นเพื่อใช้ตรวจให้คะแนนข้อคำถามของเหตุการณ์ในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ ทุกข้อมีระดับความสอดคล้องกับข้อคำถามและแนวคำตอบเมื่อเทียบกับเกณฑ์ ≥ 0.5 แสดงว่า ข้อคำถามมีความสอดคล้องกับเนื้อหา และชั้นของกระบวนการคิดวิจารณ์ญาณ และเมื่อพิจารณาคำแนะนำที่ได้จากผู้เชี่ยวชาญได้ข้อมูลโดยสรุปว่า เหตุการณ์ที่ 1 ชั้น 1) การนิยามปัญหา และชั้น 4) การวิเคราะห์ข้อมูล ในด้านองค์ประกอบของความถูกต้องของเนื้อหา ควรมีการปรับคำอธิบายของเกณฑ์ให้กระชับและครอบคลุมมากขึ้น เหตุการณ์ที่ 2 ชั้น 2) การตั้งสมมติฐาน ในด้านองค์ประกอบของการเข้าใจสิ่งที่ต้องการวัด ควรปรับเนื้อหาในแนวคำตอบให้ใช้ความรู้ในระดับชั้นมัธยมศึกษาตอนต้น เพื่อให้สอดคล้องกับระดับความสามารถของผู้สอบ และชั้น 4) การวิเคราะห์ข้อมูล ในด้านองค์ประกอบของการสื่อความหมายทางคณิตศาสตร์ ควรเพิ่มกรณีแนวคำตอบทุกกรณีที่เป็นไปได้ และเหตุการณ์ที่ 3 ชั้น 4) การวิเคราะห์ข้อมูล ในด้านองค์ประกอบของการสื่อความหมายทางคณิตศาสตร์ ควรปรับคำอธิบายของข้อรายการย่อยให้สอดคล้องกับการสื่อความหมายทางคณิตศาสตร์ ดังตาราง 8

ตาราง 8 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถามย้อนนัยประยุกต์วัดทักษะการคิดวิเคราะห์ทางคณิตศาสตร์ โดยวิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method)

เหตุการณ์ (O)	ความตรงเชิงเนื้อหาของวิธีตรวจให้คะแนนวิเคราะห์ย่อย					
	IOC	ข้อเสนอนแนะ	IOC	ข้อเสนอนแนะ	IOC	ข้อเสนอนแนะ
1.1 การนิยามปัญหา	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้กระชับ เนื่องจากใช้คำซ้ำซ้อน	0.8	-	-	ไม่วัดองค์ประกอบนี้
1.2 การตั้งสมมติฐาน	0.8	-	0.8	-	-	ไม่วัดองค์ประกอบนี้
1.3 การรวบรวมข้อมูล	0.8	-	0.8	-	-	ไม่วัดองค์ประกอบนี้
1.4 การวิเคราะห์ข้อมูล	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้กระชับ ควรระบุเกณฑ์เปรียบเทียบอัตราส่วนได้	0.8	-	0.8	-
1.5 การสร้างข้อสรุป	0.8	“ดูบางส่วน ให้ชัดเจนกว่า” หรือ 3 “เปรียบเทียบอัตราส่วนได้ถูกต้อง 2” (จาก 4 สกิลเงิน)”	0.8	-	0.8	-

ตารางที่ 8 (ต่อ)

เหตุการณ์ (O)	ความตรงเชิงเนื้อหาของวิธีตรวจให้คะแนนวิเคราะห์ย่อย					
	ข้อความถูกต้องของเนื้อหา		การเข้าใจในสิ่งที่ต้องการวัด		การสื่อความหมายทางคณิตศาสตร์	
	IOC	ข้อเสนอแนะ	IOC	ข้อเสนอแนะ	IOC	ข้อเสนอแนะ
2.1 การนิยาม ปัญหา	0.8	-	0.8	-	-	ไม่วัดองค์ประกอบนี้
2.2 การ ตั้งสมมติฐาน	0.8	-	0.8	ปรับเนื้อหาในแนวคำตอบให้ใช้ ความรู้ในระดับชั้นมัธยมศึกษา ตอนต้น	-	ไม่วัดองค์ประกอบนี้
2.3 การรวบรวม ข้อมูล	0.8	-	0.8	-	-	ไม่วัดองค์ประกอบนี้
2.4 การวิเคราะห์ ข้อมูล	0.8	-	0.8	-	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้ กระชับ และปรับแยกคุณ คำตอบ เพื่อเพิ่มจุดเชื่อมต่อของ แนวคำตอบทุกกรณีที่เป็นไปได้
2.5 การสร้างข้อสรุป	0.8	-	0.8	-	0.8	-

ตารางที่ 8 (ต่อ)

เหตุการณ์ (O)	ข้อคำถาม (I)	ความตรงเชิงเนื้อหาของวิธีตรวจให้คะแนนในวิเคราะห์ย่อย					
		ความถูกต้องของเนื้อหา		การเข้าใจในสิ่งที่ต้องการวัด		การสื่อความหมายทางคณิตศาสตร์	
		IOC	ข้อเสนอแนะ	IOC	ข้อเสนอแนะ	IOC	ข้อเสนอแนะ
3	3.1 การนิยาม ปัญหา	0.8	-	0.8	-	-	ไม่วัดองค์ประกอบนี้
	3.2 การ ตั้งสมมติฐาน	0.8	-	0.8	-	-	ไม่วัดองค์ประกอบนี้
	3.3 การรวบรวม ข้อมูล	0.8	-	0.8	-	-	ไม่วัดองค์ประกอบนี้
	3.4 การวิเคราะห์ ข้อมูล	0.8	-	0.8	-	0.8	ปรับคำอธิบายของข้อรายการย่อย ให้สอดคล้องกับการสื่อความหมาย คณิตศาสตร์
	3.5 การสร้างข้อสรุป	0.8	-	0.8	-	0.8	-

ผลการตรวจสอบความตรงเชิงเนื้อหาของแนวคำตอบกับวิธีตรวจให้คะแนน Knox พบว่าวิธีตรวจให้คะแนน Knox ที่ผู้วิจัยพัฒนาขึ้นเพื่อใช้ตรวจให้คะแนนข้อคำถามของเหตุการณ์ในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ ทุกข้อมีระดับความสอดคล้องกับข้อคำถามและแนวคำตอบเมื่อเทียบกับเกณฑ์ ≥ 0.5 แสดงว่า ข้อคำถามมีความสอดคล้องกับเนื้อหา และชั้นของกระบวนการคิดวิจารณ์ และเมื่อพิจารณาคำแนะนำที่ได้จากผู้เชี่ยวชาญได้ข้อมูลโดยสรุปว่า ข้อสอบที่ควรมีการปรับคำอธิบายของเกณฑ์ให้กระชับและครอบคลุมมากขึ้น คือ เหตุการณ์ที่ 1 ชั้น 1) การนิยามปัญหา ชั้น 4) การวิเคราะห์ข้อมูล เหตุการณ์ที่ 2 ชั้น 2) การตั้งสมมติฐาน ชั้น 4) การวิเคราะห์ข้อมูล และเหตุการณ์ที่ 3 ชั้น 2) การตั้งสมมติฐาน โดยเหตุการณ์ที่ 2 ชั้น 2) การตั้งสมมติฐาน ควรปรับเนื้อหาในแนวคำตอบให้ใช้ความรู้ในระดับชั้นมัธยมศึกษาตอนต้น เพื่อให้สอดคล้องกับระดับความสามารถของผู้สอบ ดังตาราง 9



ตาราง 9 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด
 วิจัยรณญาณทางคณิตศาสตร์ โดยวิธีการตรวจให้คะแนน Knox

เหตุการณ์ (O)	ข้อคำถาม (I)	ความตรงเชิงเนื้อหาของวิธีตรวจให้คะแนน Knox	
		IOC	ข้อเสนอแนะ
1	1.1 การนิยามปัญหา	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้กระชับ เนื่องจากใช้คำซ้ำซ้อน
	1.2 การตั้งสมมติฐาน	0.8	-
	1.3 การรวบรวมข้อมูล	0.8	-
	1.4 การวิเคราะห์ข้อมูล	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้กระชับ เนื่องจากใช้คำซ้ำซ้อน และควรระบุเกณฑ์เปรียบเทียบอัตราส่วนได้ 'ถูกบางส่วน' ให้ชัดเจนว่า "เปรียบเทียบอัตราส่วนได้ถูกต้อง 2 หรือ 3 (จาก 4 สกูลเงิน)"
	1.5 การสร้างข้อสรุป	0.8	-
2	2.1 การนิยามปัญหา	0.8	-
	2.2 การตั้งสมมติฐาน	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้ครอบคลุมมากขึ้น และปรับเนื้อหาในแนวคำตอบให้ใช้ความรู้ในระดับชั้นมัธยมศึกษาตอนต้น
	2.3 การรวบรวมข้อมูล	0.8	-
	2.4 การวิเคราะห์ข้อมูล	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้กระชับ เนื่องจากใช้คำซ้ำซ้อน และปรับแยกคู่กรณีคำตอบ เพื่อเพิ่มจุดเชื่อมต่อของแนวคำตอบทุกกรณีที่เป็นไปได้
	2.5 การสร้างข้อสรุป	0.8	-
3	3.1 การนิยามปัญหา	0.8	-
	3.2 การตั้งสมมติฐาน	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้กระชับ เนื่องจากใช้คำซ้ำซ้อน และเพิ่มกรณีของแนวคำตอบให้ครอบคลุม
	3.3 การรวบรวมข้อมูล	0.8	-
	3.4 การวิเคราะห์ข้อมูล	0.8	-
	3.5 การสร้างข้อสรุป	0.8	-

สำหรับผลการตรวจสอบความตรงเชิงเนื้อหาของแนวคำตอบกับวิธีการตรวจให้คะแนน ทั้ง 3 วิธี คือ วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) และวิธีตรวจให้คะแนน Knox พบว่า ส่วนใหญ่ผู้เชี่ยวชาญได้ให้ข้อเสนอแนะสอดคล้องกัน โดยเฉพาะวิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อยและวิธีการตรวจให้คะแนนแบบรูปริกสองชั้น สามารถบ่งชี้ได้ถึงองค์ประกอบย่อยของแต่ละชั้นการวัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ ซึ่งวิธีตรวจให้คะแนนรูปริกสองชั้นจะมีแนวการเขียนคำอธิบายของข้อรายการย่อยง่ายกว่าการเขียนคำอธิบายของแต่ละองค์ประกอบในวิธีตรวจให้คะแนนวิเคราะห์ย่อย ส่วนวิธีตรวจให้คะแนน Knox จะไม่สามารถบ่งชี้ไปถึงองค์ประกอบย่อยของแต่ละชั้นการวัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

1.6 การคัดเลือกและฝึกอบรมผู้ประเมิน

เมื่อผู้วิจัยสร้างแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์พร้อมทั้งแนวคำตอบ และเกณฑ์การให้คะแนนแล้ว ผู้วิจัยคัดเลือกผู้ประเมินจากครูที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน โดยขั้นตอนการฝึกอบรมผู้ประเมิน มีรายละเอียดดังนี้

1.6.1 อธิบายหรือสาธิตการประเมินแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ เพื่อให้ผู้ประเมินเข้าใจหลักและวิธีการประเมิน ตลอดจนการใช้เกณฑ์การให้คะแนนตามวิธีการตรวจให้คะแนน 3 วิธี คือ วิธีตรวจให้คะแนนรูปริกสองชั้น วิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox

1.6.2 ให้ผู้ประเมินฝึกใช้เกณฑ์การตรวจให้คะแนนโดยผู้ประเมินตรวจให้คะแนนตัวอย่าง ผลการตอบแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่จัดไว้สำหรับฝึกการประเมิน ซึ่งตัวอย่างดังกล่าวได้ให้คะแนนไว้ก่อนแล้วโดยผู้เชี่ยวชาญ (validation committee) เพื่อตรวจสอบทักษะการประเมินของผู้ประเมินว่าได้มาตรฐานหรือไม่

1.6.3 ผู้ประเมินฝึกตรวจให้คะแนนผลงานชุดใหม่ที่ไม่ใช่ผลงานเดิมที่ฝึกตรวจให้คะแนนไปแล้ว โดยผู้วิจัยคอยกำกับติดตาม และตรวจสอบความถูกต้องแม่นยำของผลการประเมิน หากผลการประเมินแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์โดยผู้ประเมินที่เข้าอบรมสอดคล้องกับผลการประเมินโดยผู้เชี่ยวชาญที่ให้คะแนนตัวอย่างแบบสอบเดียวกันไว้ล่วงหน้าแล้วไม่ต่ำกว่าร้อยละ 70 ถือว่าผู้ประเมินนั้นมีทักษะการประเมินได้มาตรฐาน (Penny et al., 2009 อ้างถึงในกมลวรรณ ตังธนากานนท์, 2563)

1.7 การทดลองใช้แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทาง คณิตศาสตร์

ผู้วิจัยนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ไปทดลองใช้กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนกรรณสูตศึกษาลัย จังหวัดสุพรรณบุรี จำนวน 30 คน จากนั้นนำแบบสอบอัตนัยประยุกต์มาตรวจให้คะแนนตามวิธีการตรวจให้คะแนน 3 วิธี คือ วิธีตรวจให้คะแนนรูปกรีกสองชั้น วิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox โดยผู้ประเมินที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือ มีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 1 คน ตรวจให้คะแนนแบบสอบดังกล่าว จากนั้นนำคะแนนที่ได้ไปหาค่าสถิติพื้นฐานโดยใช้สถิติบรรยาย คือ ค่าเฉลี่ย (M) ส่วนเบี่ยงเบนมาตรฐาน (SD) คะแนนสูงสุด (Maximum) คะแนนต่ำสุด (Minimum) ของแต่ละเหตุการณ์ และตรวจสอบคุณภาพของแบบสอบ โดยวิเคราะห์คุณภาพของแบบสอบทั้งรายข้อและทั้งฉบับตามทฤษฎีการทดสอบแบบดั้งเดิม ได้แก่ ค่าความยาก (p) อำนาจจำแนก (r) และความเที่ยง รวมทั้งความชัดเจนของข้อคำถาม ความเหมาะสมของข้อคำถาม ความเหมาะสมของภาษาที่ใช้ และระยะเวลาที่ใช้ในการทำแบบสอบ โดยมีเกณฑ์ในการแปลความหมายของคุณภาพด้านต่าง ๆ คือ แบบสอบในแต่ละข้อควรมีค่าความยาก อยู่ระหว่าง 0.2 ถึง 0.8 ค่าอำนาจจำแนกมีค่าตั้งแต่ 0.2 ขึ้นไป และค่าความเที่ยงตามเกณฑ์มาตรฐานที่กำหนดความเที่ยงตั้งแต่ 0.7 ขึ้นไป (Nunnally, 1987) พร้อมทั้งปรับปรุงภาษาที่ใช้ให้มีความเหมาะสม จากนั้นนำข้อสอบมารวบรวมเป็นแบบสอบวัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ (ศิริชัย กาญจนวาสี, 2556; กมลวรรณ ตังธนกานนท์, 2557; โชติกา ภาชีผล, 2559) ดังตาราง 10 และ 11

ตาราง 10 เกณฑ์การแปลความหมายของค่าความยาก (p)

ค่าความยาก (p)	ความหมาย
0.81 – 1.00	ง่ายมาก (ควรปรับปรุงหรือตัดทิ้ง)
0.60 – 0.80	ค่อนข้างง่าย
0.40 – 0.59	ปานกลาง
0.20 – 0.39	ค่อนข้างยาก
0 – 0.19	ยากมาก (ควรปรับปรุงหรือตัดทิ้ง)
0.60 – 1.00	ดีมาก
0.40 – 0.59	ดี
0.20 – 0.39	พอใช้ได้
0.10 – 0.19	ค่อนข้างต่ำ ควรปรับปรุง
0 – 0.09	ต่ำมาก ควรปรับปรุง

ค่าความยาก ค่าอำนาจจำแนก และค่าความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์โดยใช้วิธีตรวจให้คะแนนที่ต่างกัน จากผลการทดลองใช้แบบสอบ เมื่อพิจารณาค่าความยากในแต่ละวิธีตรวจให้คะแนน พบว่า ค่าความยากของแบบสอบทั้งฉบับทั้ง 3 วิธีตรวจให้คะแนน อยู่ในช่วง 0.37 – 0.59 โดยวิธีตรวจให้คะแนน Knox ยากกว่าวิธีตรวจให้คะแนนรูปริกสองชั้น และวิธีตรวจให้คะแนนวิเคราะห์ย่อย ตามลำดับ แสดงว่าเกณฑ์การตรวจให้คะแนนทั้ง 3 วิธี มีความยากในระดับปานกลางค่อนข้างไปทางยาก และเกณฑ์การตรวจให้คะแนน Knox ยากที่สุด เมื่อพิจารณาค่าความยากรายเหตุการณ์ พบว่า ค่าความยากของแบบสอบทั้ง 3 เหตุการณ์ อยู่ในช่วง 0.34 – 0.62 โดยเหตุการณ์ที่ 2 ยากกว่าเหตุการณ์ที่ 3 และเหตุการณ์ที่ 1 ตามลำดับ แสดงว่าแบบสอบทั้ง 3 เหตุการณ์ มีความยากในระดับปานกลางค่อนข้างไปทางยาก และเหตุการณ์ที่ 2 ยากที่สุด

เมื่อพิจารณาค่าอำนาจจำแนกในแต่ละวิธีตรวจให้คะแนน พบว่า พบว่า ค่าอำนาจจำแนกของแบบสอบทั้งฉบับทั้ง 3 วิธีตรวจให้คะแนน อยู่ในช่วง 0.31 – 0.37 โดยวิธีตรวจให้คะแนนแบบรูปริกสองชั้น มีอำนาจจำแนกสูงกว่าวิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox ตามลำดับ แสดงว่า เกณฑ์การตรวจให้คะแนนทั้ง 3 วิธี มีความสามารถในการจำแนกผู้สอบได้ และวิธีตรวจให้คะแนนแบบรูปริกสองชั้น มีความสามารถในการจำแนกผู้สอบได้สูงสุด เมื่อพิจารณาค่าอำนาจจำแนกรายเหตุการณ์ พบว่า ค่าอำนาจจำแนกของแบบสอบทั้ง 3

เหตุการณ์ อยู่ในช่วง 0.29 – 0.41 โดยส่วนใหญ่เหตุการณ์ที่ 3 มีอำนาจจำแนกสูงสุด แสดงว่าแบบสอบทั้ง 3 เหตุการณ์มีอำนาจจำแนกอยู่ในระดับพอใช้ได้ และเหตุการณ์ที่ 3 มีความสามารถในการจำแนกผู้สอบได้ดีที่สุด

เมื่อพิจารณาค่าความเที่ยงในแต่ละวิธีตรวจให้คะแนน พบว่า ค่าความเที่ยงของแบบสอบทั้งฉบับทั้ง 3 วิธีตรวจให้คะแนน อยู่ในช่วง 0.836 – 0.934 โดยวิธีตรวจให้คะแนนรูปภิกสองมีค่าความเที่ยงสูงกว่าวิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox ตามลำดับ แสดงว่า เกณฑ์การตรวจให้คะแนนทั้ง 3 วิธี มีความเที่ยงอยู่ในระดับสูง และวิธีตรวจให้คะแนนรูปภิกสองชั้นมีความเที่ยงสูงสุด เมื่อพิจารณาค่าความเที่ยงรายเหตุการณ์ พบว่า ค่าความเที่ยงของแบบสอบทั้ง 3 เหตุการณ์ อยู่ในช่วง 0.463 – 0.878 โดยเหตุการณ์ที่ 3 มีค่าความเที่ยงสูงกว่าเหตุการณ์ที่ 2 และเหตุการณ์ที่ 1 ตามลำดับ แสดงว่า แบบสอบทั้ง 3 เหตุการณ์ค่าความเที่ยงอยู่ในระดับสูง ตามเกณฑ์ความเที่ยงตั้งแต่ 0.7 ขึ้นไปที่ Nunnally (1987) เสนอไว้ และเหตุการณ์ที่ 3 มีค่าความเที่ยงสูงสุด ผลการวิเคราะห์เป็นดังตาราง 11

ตาราง 11 ค่าความยาก ค่าอำนาจจำแนก และค่าความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์โดยใช้วิธีการตรวจให้คะแนนที่ต่างกัน จากกลุ่มทดลองใช้ (try out)

เหตุการณ์	วิธีตรวจให้คะแนน รูปภิกสองชั้น			วิธีตรวจให้คะแนน วิเคราะห์ย่อย			วิธีตรวจให้คะแนน Knox		
	p	r	r _{tt}	p	r	r _{tt}	p	r	r _{tt}
	เหตุการณ์ 1	0.49	0.41	0.742	0.62	0.30	0.677	0.43	0.29
เหตุการณ์ 2	0.42	0.32	0.813	0.54	0.31	0.786	0.34	0.29	0.648
เหตุการณ์ 3	0.48	0.38	0.878	0.58	0.36	0.855	0.35	0.35	0.726
ทั้งฉบับ	0.46	0.37	0.934	0.59	0.33	0.919	0.37	0.31	0.836

ผลการวิเคราะห์ค่าสถิติพื้นฐานของวิธีตรวจให้คะแนนในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนนรูปภิกสองชั้น (double

layer scoring rubric) ที่ได้จากกลุ่มทดลอง นักเรียนชั้นมัธยมศึกษาปีที่ 3 พบว่า เมื่อพิจารณา คะแนนรวม ทั้ง 3 เหตุการณ์ จากคะแนนเต็ม 36 คะแนน คะแนนเฉลี่ย 19.97 คะแนน ส่วน เบี่ยงเบนมาตรฐาน 9.32 คะแนน สรุปได้ว่า คะแนนเฉลี่ยทั้งฉบับของแบบสอบอัตนัยประยุกต์วัด ทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนนรูบริกสองชั้นทำได้ใน ระดับสูงกว่าครึ่งหนึ่งของแบบสอบ

ผลการวิเคราะห์ค่าสถิติพื้นฐานของวิธีตรวจให้คะแนนในแบบสอบอัตนัยประยุกต์วัด ทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ที่ได้จากกลุ่มทดลอง นักเรียนชั้นมัธยมศึกษาปีที่ 3 พบว่า เมื่อพิจารณาคะแนนรวมทั้ง 3 เหตุการณ์ จากคะแนนเต็ม 72 คะแนน คะแนนเฉลี่ย 43.03 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 10.90 คะแนน สรุปได้ว่า คะแนนเฉลี่ยทั้งฉบับของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนนวิเคราะห์ย่อยทำได้ในระดับสูงกว่า ครึ่งหนึ่งของแบบสอบ

ผลการวิเคราะห์ค่าสถิติพื้นฐานของวิธีตรวจให้คะแนนในแบบสอบอัตนัยประยุกต์วัด ทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนน Knox ที่ได้จากกลุ่มทดลอง นักเรียนชั้นมัธยมศึกษาปีที่ 3 พบว่า เมื่อพิจารณาคะแนนรวมทั้ง 3 เหตุการณ์ จากคะแนนเต็ม 30 คะแนน คะแนนเฉลี่ย 11.00 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 4.88 คะแนน สรุปได้ว่า คะแนน เฉลี่ยทั้งฉบับของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยใช้วิธี ตรวจให้คะแนน Knox ทำได้ในระดับต่ำกว่าครึ่งหนึ่งของแบบสอบ ดังตาราง 12

ตาราง 12 ค่าสถิติพื้นฐานของคะแนนรายข้อของแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิเคราะห์ตามทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนที่ต่างกัน
จากผลการทดลองใช้แบบสอบ

เหตุการณ์ (O)	ข้อคำถาม (I)	วิธีการตรวจให้คะแนนตามแบบ Knox			วิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อย			วิธีการให้คะแนนแบบรูปภิกของขึ้น					
		Max	Min	M	SD	Max	Min	M	SD	Max	Min	M	SD
1	1.1 การนิยามปัญหา	1	0	0.70	0.47	4	1	2.27	0.87	2	0	0.90	0.61
	1.2 การตั้งสมมติฐาน	1	0	0.70	0.47	4	1	2.33	0.80	2	0	1.00	0.64
	1.3 การรวบรวมข้อมูล	1	0	0.73	0.45	4	1	2.77	0.77	2	0	1.27	0.69
	1.4 การวิเคราะห์ข้อมูล	2	0	1.17	0.87	6	1	3.60	1.43	3	0	1.37	1.00
	1.5 การสร้างข้อสรุป	2	0	0.97	0.85	6	1	3.73	1.44	3	0	1.37	1.00
	รวม	7	1	4.27	1.84	23	10	14.70	3.65	11	1	5.90	2.83
2	2.1 การนิยามปัญหา	2	0	0.70	0.54	4	1	2.60	0.93	2	0	1.23	0.73
	2.2 การตั้งสมมติฐาน	1	0	0.70	0.47	4	1	2.50	0.94	2	0	1.33	0.66
	2.3 การรวบรวมข้อมูล	2	0	0.70	0.54	4	1	2.67	1.12	2	0	1.33	0.80
	2.4 การวิเคราะห์ข้อมูล	2	0	0.70	0.65	6	1	2.73	1.14	3	0	1.77	1.10
	2.5 การสร้างข้อสรุป	1	0	0.53	0.51	5	0	2.40	1.16	3	0	1.53	1.04
	รวม	6	0	3.33	1.75	21	7	12.90	3.91	12	2	7.20	3.35
3	3.1 การนิยามปัญหา	1	0	0.70	0.47	4	1	2.50	0.90	2	0	1.07	0.83
	3.2 การตั้งสมมติฐาน	1	0	0.63	0.49	4	1	2.80	0.92	2	0	1.10	0.80
	3.3 การรวบรวมข้อมูล	2	0	0.67	0.55	4	1	2.97	0.93	2	0	1.20	0.71
	3.4 การวิเคราะห์ข้อมูล	2	0	0.83	0.65	6	1	3.63	1.10	3	0	1.77	1.01
	3.5 การสร้างข้อสรุป	2	0	0.57	0.68	6	2	3.53	1.20	3	0	1.73	1.05
	รวม	7	0	3.40	1.98	23	8	15.43	4.04	12	1	6.87	3.65
	รวมทั้งฉบับ	20	3	11.00	4.88	59	26	43.03	10.90	33	6	19.97	9.32

จากการสร้างแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนระดับชั้นมัธยมศึกษาตอนต้น เนื้อหาวิชาคณิตศาสตร์ ตามหลักสูตรแกนกลาง การศึกษาขั้นพื้นฐาน พ.ศ. 2551 (ฉบับปรับปรุง พ.ศ. 2561) ตามสาระทางคณิตศาสตร์ จำนวน 3 สาระ ได้แก่ จำนวนและพีชคณิต (เรื่อง อัตราส่วน และระบบสมการ) การวัดและเรขาคณิต (เรื่อง ทฤษฎีบทพีทาโกรัส และพื้นที่ผิวและปริมาตร) สถิติและความน่าจะเป็น (เรื่อง สถิติ) โดยเลือก เหตุการณ์ที่จะนำมาสร้างแบบสอบอัตนัยประยุกต์และข้อคำถามจากเหตุการณ์จริงที่คัดเลือก เนื้อหาที่สนใจและมีความเหมาะสมกับเนื้อหาในระดับชั้นของนักเรียนจากแหล่งข้อมูลต่าง ๆ เช่น หนังสือพิมพ์ ข่าว อินเทอร์เน็ต เป็นต้น ปรับเนื้อหาให้เหมาะสมโดยการตัดทอนหรือปรับข้อความ แต่ยังคงความหมายใจความเดิม จากนั้นตรวจสอบคุณภาพของแบบสอบอัตนัยประยุกต์วัดทักษะ การคิดวิจารณ์ญาณทางคณิตศาสตร์และวิธีการตรวจให้คะแนน 3 วิธี คือ วิธีตรวจให้คะแนนรูปกริ สองชั้น (double layer scoring rubric) วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) และ วิธีตรวจให้คะแนนตาม Knox ผ่านการนำไปทดลองใช้กับกลุ่มทดลองและนำผลที่ได้มาปรับปรุง พัฒนาแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ รวมถึงวิธีการตรวจ ให้คะแนนให้มีประสิทธิภาพยิ่งขึ้น และเพื่อนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์และวิธีการตรวจให้คะแนนไปใช้อย่างมีประสิทธิภาพ ผู้วิจัยได้จัดทำ คู่มือการใช้แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับ นักเรียนชั้นมัธยมศึกษาตอนต้น โดยมีการชี้แจงการใช้คู่มือ การอธิบายความสำคัญของแบบสอบ อัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ แนวปฏิบัติและการนำแบบสอบไป ใช้ การตรวจให้คะแนนและการบันทึกคะแนน รวมถึงตัวอย่างแบบสอบอัตนัยประยุกต์วัดทักษะ การคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 วิธีการตรวจให้คะแนน และแบบบันทึกคะแนน ซึ่งจะปรากฏอยู่ในภาคผนวก

ขั้นตอนที่ 2 การเก็บรวบรวมข้อมูล

2.1 การกำหนดประชากรและตัวอย่างวิจัย

2.1.1 ประชากร

นักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปีที่ 3 ในโรงเรียนสังกัด สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา สุพรรณบุรี จำนวน 33,751 คน (ข้อมูลจากระบบ สารสนเทศสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาสุพรรณบุรี ณ วันที่ 31 พฤษภาคม 2564)

2.1.2 ตัวอย่างวิจัย

ตัวอย่างวิจัยที่ใช้ในครั้งนี้ คือ นักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2564 ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา สุพรรณบุรี กำหนดขนาดตัวอย่างวิจัยด้วยการใช้ WebPower package ในโปรแกรม R สำหรับการทดสอบ Multi-way ANOVA (3-way ANOVA) ระหว่างจำนวนพาเซท (ผู้สอบ, ผู้ประเมิน และเหตุการณ์) และวิธีการตรวจให้คะแนน (วิธีการตรวจให้คะแนนรูปภิกสองชั้น วิธีการตรวจให้คะแนนวิเคราะห์ย่อย และวิธีการตรวจให้คะแนน Knox) ด้วยสถิติที่ใช้ทดสอบ คือ F-test กำหนดขนาดอิทธิพลขนาดใหญ่ เท่ากับ 0.40 ความคลาดเคลื่อนร้อยละ 5 ระหว่างจะได้ขนาดตัวอย่าง 80 คน เพื่อป้องกันการขาดหายของตัวอย่างในขณะเก็บรวบรวมข้อมูล ผู้วิจัยจึงกำหนดขนาดตัวอย่างวิจัยเป็น 90 คน

2.2 การสุ่มตัวอย่างวิจัย

เกณฑ์การพิจารณาโรงเรียนของตัวอย่างวิจัย คือ โรงเรียนที่มีการจัดการเรียนการสอนระดับชั้นมัธยมศึกษาลดความสามารถของผู้เรียน และเป็นโรงเรียนที่มีความต้องการพัฒนาทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ รวมถึงผู้บริหารและครูมีความสนใจและให้ความสำคัญร่วมมือด้านการบริหารจัดการ อำนวยความสะดวก และอนุญาตให้ผู้วิจัยดำเนินการทดลองได้ และเพื่อให้ได้ตัวอย่างที่เป็นตัวแทนของคุณลักษณะที่ต้องการวัด ผู้วิจัยกำหนดขนาดของตัวอย่างโดยเลือกโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา สุพรรณบุรี จนได้จำนวนตัวอย่างวิจัยทั้งสิ้น 90 คน ผ่านการสุ่มตัวอย่างแบบหลายขั้นตอน (multi-stage random sampling) ดังนี้

ขั้นที่ 1 เลือกโรงเรียนจากสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 9 สุพรรณบุรี โดยวิธีการจับฉลาก จำนวน 3 โรงเรียน จากจำนวนโรงเรียนทั้งหมด 32 โรงเรียน คือ โรงเรียนสระยายโสมวิทยา โรงเรียนบางลี่วิทยา และโรงเรียนกาญจนาภิเษกวิทยาลัย สุพรรณบุรี ซึ่งเป็นตัวแทนของโรงเรียนขนาดเล็ก โรงเรียนขนาดกลาง และโรงเรียนขนาดใหญ่ ตามลำดับ

ขั้นที่ 2 สุ่มห้องเรียนจากโรงเรียนตัวอย่าง โรงเรียนละ 1 ห้องเรียน โดยเป็นนักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปีที่ 3 ดังนี้ สุ่มห้องเรียนจากโรงเรียนสระยายโสมวิทยา จำนวน 2 ห้องเรียน จากทั้งหมด 4 ห้องเรียน โรงเรียนบางลี่วิทยา จำนวน 2 ห้องเรียน จากทั้งหมด 7 ห้องเรียน และโรงเรียนกาญจนาภิเษกวิทยาลัย สุพรรณบุรี จำนวน 2 ห้องเรียน จากทั้งหมด 12 ห้องเรียน

ชั้นที่ 3 สุ่มตัวอย่างวิจัยจากแต่ละห้องเรียน ดังนี้ สุ่มตัวอย่างวิจัยจากห้องเรียนในโรงเรียนสระยายโสมวิทยา ห้องละ 11 คน สุ่มตัวอย่างวิจัยจากห้องเรียนในโรงเรียนบางลี่วิทยา ห้องละ 15 คน สุ่มตัวอย่างวิจัยจากห้องเรียนในโรงเรียนกาญจนาภิเษกวิทยาลัย สุพรรณบุรี ห้องละ 19 คน จนได้ตัวอย่างวิจัย 90 คน ดังตาราง 13

ตาราง 13 จำนวนนักเรียนที่ใช้เป็นตัวอย่างวิจัยในการเก็บข้อมูล

โรงเรียน	ห้องเรียน	ตัวอย่างวิจัย (คน)
โรงเรียนสระยายโสมวิทยา	3/1 และ 3/3	22
โรงเรียนบางลี่วิทยา	3/2 และ 3/7	30
โรงเรียนกาญจนาภิเษกวิทยาลัย สุพรรณบุรี	3/3 และ 3/5	38
รวม		90

2.3 เครื่องมือที่ใช้ในงานวิจัย

2.3.1 แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีจำนวน 3 เหตุการณ์ โดยแต่ละเหตุการณ์มีข้อคำถามจำนวน 5 ข้อ การสร้างข้อคำถามจะสอดคล้องกับขั้นตอนการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยในข้อคำถามที่ 1 ของทุกเหตุการณ์ จะวัดความสามารถในการนิยามปัญหา ข้อคำถามที่ 2 ของทุกเหตุการณ์จะวัดความสามารถในการตั้งสมมติฐาน ข้อคำถามที่ 3 ของทุกเหตุการณ์จะวัดความสามารถในการรวบรวมข้อมูล ข้อคำถามที่ 4 ของทุกเหตุการณ์จะวัดความสามารถในการวิเคราะห์ข้อมูล และข้อคำถามที่ 5 ของทุกเหตุการณ์จะวัดความสามารถในการสร้างข้อสรุป

2.3.2 วิธีการตรวจให้คะแนน งานวิจัยในครั้งนี้ใช้วิธีการตรวจให้คะแนน 3 วิธี คือ วิธีตรวจให้คะแนนรูปกริกสองชั้น (double layer scoring rubric) วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) และวิธีตรวจให้คะแนนตาม Knox มีรายละเอียด ดังนี้

วิธีตรวจให้คะแนนรูปกริกสองชั้น ในชั้นที่ 1 ให้คะแนนตามกระบวนการคิดวิจารณ์ญาณ 5 ขั้นตอน และในแต่ละขั้นตอนมีเกณฑ์การให้คะแนน ดังนี้ ข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 ข้อละ 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (4 คะแนน) และการเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) จากนั้นแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้คะแนนเต็ม ข้อละ 2 คะแนน ส่วนข้อคำถามที่ 4 และ ข้อคำถามที่ 5 ข้อละ 3 องค์ประกอบ คือ ความ

ถูกต้องของเนื้อหา (4 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) และการสื่อความหมายทางคณิตศาสตร์ (4 คะแนน) จากนั้นแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้ คะแนนเต็ม ข้อละ 3 คะแนน รวมคะแนนเต็ม แต่ละเหตุการณ์ 12 คะแนน รวมทั้งฉบับ คิดเป็น 36 คะแนน

วิธีตรวจให้คะแนนวิเคราะห์ย่อย มีเกณฑ์การให้คะแนน ดังนี้ ข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 ข้อละ 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) ส่วนในข้อคำถามที่ 4 และข้อคำถามที่ 5 ข้อละ 3 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) และการสื่อความหมายทางคณิตศาสตร์ (2 คะแนน) รวมคะแนนเต็มของแต่ละเหตุการณ์ 24 คะแนน รวมทั้งฉบับคิดเป็น 72 คะแนน

วิธีตรวจให้คะแนน Knox มีเกณฑ์การให้คะแนนแต่ละข้อคำถามเป็น 3 ระดับ ได้แก่ 2 คะแนน คือ ผู้ตอบคำถามได้ถูกต้องและครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ ตามแนวคำตอบ 1 คะแนน คือ ผู้ตอบคำถามผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ และ 0 คะแนน คือ ผู้ตอบคำถามไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม รวมคะแนนเต็มของแต่ละเหตุการณ์ 10 คะแนน รวมทั้งฉบับคิดเป็น 30 คะแนน

2.4 การดำเนินการเก็บข้อมูลจากการออกแบบการวัดโดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง

ผู้วิจัยออกแบบสถานการณ์โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง ผ่านการออกแบบการวัดแบบไขว้ (cross design) 2 ฟาเซท ($p \times r \times o$) คือ ให้ผู้เรียนทุกคนทำแบบสอบทุกเหตุการณ์ โดยผู้ประเมินแต่ละคนตรวจแบบสอบทุกเหตุการณ์ โดยมีขั้นตอนการดำเนินการดังนี้

2.4.1 ขั้นตอนการดำเนินการประสานงาน

ผู้วิจัยทำหนังสือจากงานหลักสูตร คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ขอความร่วมมือในการเก็บข้อมูลจากผู้อำนวยการโรงเรียนที่เป็นตัวอย่างวิจัย ผู้วิจัยติดต่อประสานกับครูประจำวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 ที่เป็นห้องตัวอย่างวิจัย เพื่อติดต่อเรื่องวันและเวลาที่ผู้วิจัยจะดำเนินการสอบและจะดำเนินการเก็บข้อมูลกับนักเรียนของโรงเรียนตัวอย่างวิจัย จำนวน 1 ครั้ง ซึ่งเวลาในการสอบนักเรียนชั้นมัธยมศึกษาปีที่ 3 เป็นเวลา 1

ชั่วโมง ในทุกโรงเรียนตัวอย่างวิจัย โดยใช้ผู้คุมสอบจำนวน 2 คน ในทุกโรงเรียนตัวอย่างวิจัย ประกอบด้วยผู้วิจัยและครูประจำภาคคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 ที่เป็นห้องตัวอย่างวิจัยนั้น

2.4.2 ขั้นตอนการดำเนินการสอบ

1) ผู้คุมสอบแจกแบบสอบอัตนัยประยุกต์ไว้บนโต๊ะสอบทุกโต๊ะ โต๊ะละ 1 ชุด เมื่ออนุญาตให้นักเรียนเข้าประจำที่สอบแล้ว ผู้วิจัยแนะนำตนเองและขอความร่วมมือจากนักเรียนในการทำแบบสอบ

2) อนุญาตให้นักเรียนเปิดซองข้อสอบพร้อมกัน

3) ผู้วิจัยให้นักเรียนทุกคนอ่านคำแนะนำในการทำแบบสอบอัตนัยประยุกต์พร้อมทั้งให้เวลาเขียนชื่อ นามสกุล ชั้น และเลขที่ ที่ข้อคำถามแผ่นแรกให้เรียบร้อย

4) ผู้คุมสอบอธิบายย้ำถึงวิธีการทำแบบสอบให้นักเรียนฟัง โดยให้นักเรียนทำแบบสอบทีละข้อคำถาม และต้องทำข้อคำถามนั้นให้เสร็จ จึงจะสามารถเปลี่ยนไปทำข้อคำถามต่อไปได้ สำหรับเวลาในการทำแบบสอบ คือ 1 ชั่วโมง โดยให้นักเรียนบริหารจัดการเวลาสอบเอง และเน้นย้ำว่าเมื่อนักเรียนทำข้อคำถามในแต่ละหน้าเสร็จเรียบร้อยแล้ว ให้ตรวจสอบความเรียบร้อยก่อนแล้วจึงส่งข้อสอบโดยคว่ำกระดาษลงและวางไว้ได้ช่องใส่แบบสอบ และบอกให้นักเรียนเริ่มทำแบบสอบได้ ผู้วิจัยคุมสอบบันทึกเวลาเริ่มต้นในการทำข้อสอบและเวลาสิ้นสุดในการทำแบบสอบอัตนัยประยุกต์

2.4.3 ขั้นตอนการตรวจให้คะแนนตามการออกแบบการวัด 2 ฟาเซท

1) ผู้วิจัยรวบรวมแบบสอบจากการเก็บข้อมูลในตัวอย่างวิจัย แล้วดำเนินการส่งต่อให้ผู้ประเมิน จำนวน 2 คน ที่ผ่านการคัดเลือกและฝึกตรวจให้คะแนนจนมีผลการประเมินได้มาตรฐาน

2) ผู้ประเมินแต่ละคนตรวจแบบสอบอัตนัยประยุกต์ ด้วยวิธีการตรวจให้คะแนน 3 วิธี คือ วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) วิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนนตาม Knox โดยดำเนินการตรวจทีละเหตุการณ์ จนครบ 3 เหตุการณ์ โดยบันทึกคะแนนสอบผ่านแบบบันทึกคะแนนตามที่ปรากฏในภาคผนวก ๑ และภาคผนวก ๒

3) ผู้วิจัยนำผลการบันทึกคะแนนแบบสอบอัตนัยประยุกต์ที่ตรวจให้คะแนนทั้ง 3 วิธีข้างต้นมาดำเนินการวิเคราะห์ข้อมูล

ขั้นตอนที่ 3 การวิเคราะห์ข้อมูล

3.1 การวิเคราะห์ค่าสถิติพื้นฐานของผลการตรวจให้คะแนนแบบสอบอัตนัย ประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

นำคะแนนที่ได้จากการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ด้วยวิธีการตรวจให้คะแนน 3 วิธี คือ วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) และวิธีตรวจให้คะแนนตาม Knox จากผู้ประเมิน คือ ผู้ที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน มาวิเคราะห์ค่าสถิติพื้นฐานโดยใช้สถิติบรรยาย คือ ค่าเฉลี่ย (M) ส่วนเบี่ยงเบนมาตรฐาน (SD) คะแนนสูงสุด (Maximum) คะแนนต่ำสุด (Minimum)

3.2 การวิเคราะห์คุณภาพของการตรวจให้คะแนนด้วยวิธีตรวจให้คะแนนรูปริกสองชั้น ในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

นำคะแนนที่ได้จากการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ด้วยวิธีตรวจให้คะแนนแบบรูปริกสองชั้น (double layer scoring rubric) จากผู้ประเมินจำนวน 1 คน ที่ทำการตรวจให้คะแนนแบบสอบทุกฉบับ และตรวจแบบสอบทุกฉบับจำนวน 2 ครั้ง ห่างกัน 4 สัปดาห์ มาวิเคราะห์หาความเที่ยงในการตรวจให้คะแนนจากผู้ตรวจคนเดียว (intra rater reliability) ด้วยค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation) และนำคะแนนที่ได้จากการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ด้วยวิธีการให้คะแนนแบบรูปริกสองชั้นจากผู้ประเมิน คือ ผู้ที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน ที่ทำการตรวจให้คะแนนแบบสอบทุกฉบับอย่างเป็นอิสระ เพื่อนำมาวิเคราะห์ความเที่ยงของการตรวจให้คะแนนระหว่างผู้ประเมิน (inter rater reliability) ด้วยค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation) ควบคู่กับค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) ที่วิเคราะห์จากโมเดลอิทธิพลผสมแบบสองทาง (two-way mixed-effects model) ภายใต้ผู้ประเมิน 2 คน (โดยใช้สัญลักษณ์ คือ $ICC_{3,2}$) ที่ระดับความเชื่อมั่น 95% กำหนดเกณฑ์การพิจารณาระดับความสอดคล้องตามแนวทางของ Portney and Watkins (2015) อ้างถึงในบุษยารัตน์ จันทร์ประเสริฐ และคณะ (2560) ดังตาราง 14

ตาราง 14 เกณฑ์การแปลความหมายของค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (ICC)

ค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (ICC)	ความหมาย
0.90 – 1.00	ผู้ประเมินมีความสอดคล้องในการประเมินระดับดีมาก
0.75 – 0.89	ผู้ประเมินมีความสอดคล้องในการประเมินระดับดี
0.50 – 0.74	ผู้ประเมินมีความสอดคล้องในการประเมินระดับพอใช้
0 – 0.49	ผู้ประเมินมีความสอดคล้องในการประเมินระดับต่ำ

3.3 การวิเคราะห์ความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด

วิจารณ์ญาณทางคณิตศาสตร์ ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกัน ภายใต้จำนวนเหตุการณ์ จำนวน ข้อคำถาม และจำนวนผู้ประเมินที่ต่างกัน โดยใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (G-Theory)

นำคะแนนที่ได้จากการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกัน จากผู้ประเมิน คือ ผู้ที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือ มีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน มาคำนวณเป็นคะแนน T-Score ซึ่งเป็นการวิเคราะห์ค่าความเที่ยง (ค่าสัมประสิทธิ์การสรุปอ้างอิง: G-Coefficient) ของแบบสอบอัตนัยประยุกต์ที่ใช้วิธีตรวจให้คะแนนรูปกรสองชั้น โดยเปรียบเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox ภายใต้จำนวนผู้ประเมิน และจำนวนเหตุการณ์ที่ต่างกัน เมื่อใช้ผู้ประเมินจำนวน 1 และ 2 คน ภายใต้จำนวนเหตุการณ์ที่ใช้ 2, 3, 4, 5 และ 6 เหตุการณ์ แต่ละเหตุการณ์จะมี 5 ข้อคำถาม ซึ่งออกแบบการวัดแบบไขว้ (cross design) 2 ฟาเซท ($p \times r \times o$) คือ นักเรียนทุกคนทำแบบสอบทุกเหตุการณ์ โดยผู้ประเมินแต่ละคนตรวจแบบสอบทุกเหตุการณ์ ซึ่งใช้โปรแกรมสำเร็จรูป EduG และเปรียบเทียบค่าความเที่ยง (ค่าสัมประสิทธิ์การสรุปอ้างอิง: G-Coefficient) ด้วยวิธีการของพิทแมน (Pittman) โดยสูตร (Feldt, L. S. et al., 1987 อ้างถึงในมาฆพันธ์ อานาคิล และกมลวรรณ ตังธนากานนท์, 2557)

			$t = \frac{(r_{tt_1} - r_{tt_2})\sqrt{(N-2)}}{\sqrt{4(1-r_{tt_1})(1-r_{tt_2})(1-r_{x_1x_2}^2)}}$	เมื่อ $df = n - 2$
เมื่อ	t	แทน	การแจกแจงแบบที	
	r_{tt_1}	แทน	ค่าความเที่ยงของแบบสอบเมื่อตรวจให้คะแนนด้วยวิธีตรวจให้คะแนนแบบที่ 1	
	r_{tt_2}	แทน	ค่าความเที่ยงของแบบสอบเมื่อตรวจให้คะแนนด้วยวิธีตรวจให้คะแนนแบบที่ 2	
	N	แทน	จำนวนกลุ่มตัวอย่าง	
	$r_{x_1x_2}$	แทน	สหสัมพันธ์ระหว่างคะแนนจากวิธีตรวจให้คะแนนแบบที่ 1 กับ 2	

บทที่ 4 ผลการวิเคราะห์ข้อมูล

งานวิจัย เรื่อง ความเที่ยงภายในได้เงื่อนไขการทดสอบที่ต่างกันของแบบสอบอัตนัย ประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง มี วัตถุประสงค์ (1) เพื่อพัฒนาแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์และวิธีตรวจให้คะแนนรูปริกสองชั้น สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 (2) เพื่อตรวจสอบคุณภาพวิธีตรวจให้คะแนนรูปริกสองชั้นสำหรับแบบสอบอัตนัยประยุกต์ วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ตาม สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ (3) เพื่อเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนที่แตกต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง ซึ่งประกอบด้วย 2 ฟาเซท คือ จำนวนผู้ประเมินและจำนวนเหตุการณ์ โดยมีตัวอย่างวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2564 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาสุพรรณบุรี จำนวน 90 คน

ในบทนี้ผู้วิจัยเสนอผลการวิเคราะห์ข้อมูลออกเป็น 3 ตอน ดังนี้

ตอนที่ 1 ผลการพัฒนาแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์และวิธีตรวจให้คะแนนแบบรูปริกสองชั้น สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบสอบอัตนัยประยุกต์ (Modified Essay Question test: MEQ) เป็นแบบสอบความเรียงรูปแบบหนึ่งที่ได้รับการพัฒนาจากแบบสอบอัตนัยแบบดั้งเดิมที่มักวัดความรู้ที่อยู่ในระดับท่องจำหรือเข้าใจพื้นฐานเท่านั้น และมีความเป็นปรนัยในการตรวจให้คะแนนค่อนข้างต่ำ โดยแบบสอบอัตนัยประยุกต์นี้จะบรรยายเหตุการณ์ที่เกิดขึ้นจริงตามลำดับของเหตุการณ์นั้น ๆ แต่จะไม่นำเสนอเหตุการณ์ทั้งหมดเพียงครั้งเดียวเพื่อเป็นการเพิ่มข้อมูลที่อาจจะเกี่ยวข้องหรือไม่เกี่ยวข้องกับคำถาม ผู้สอบจะต้องนำข้อมูลมาพิจารณาเอง ในแต่ละข้อคำถามจะเป็นอิสระจากกันและผู้สอบไม่มีโอกาสกลับไปแก้ไขคำตอบของข้อ ที่ผ่านมาได้ โดยผู้สอบจะต้องใช้ข้อมูลที่กำหนดมาให้เท่านั้นตอบคำถาม ดังนั้นจึงทำให้แบบสอบอัตนัยประยุกต์แตกต่างกับแบบสอบอัตนัยแบบดั้งเดิมและเป็นที่ยอมรับในการนำมาวัดความรู้ขั้นสูง

มากกว่าแบบสอบถามประเภทเลือกตอบและประเภทอื่น ๆ เพราะมีค่าความตรง ความเที่ยง และค่าอำนาจจำแนกสูง

ผู้วิจัยจึงได้พัฒนาแบบสอบถามอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ ซึ่งเป็นการวัดกระบวนการคิดขั้นสูง จำนวน 3 เหตุการณ์ แต่ละเหตุการณ์จะมี 5 ข้อคำถาม รวมทั้งสิ้น ฉบับละ 15 ข้อคำถาม แต่ละข้อคำถามวัดตามกระบวนการคิดวิจารณ์ญาณทางคณิตศาสตร์ ดังนี้ ข้อคำถามที่ 1 ของทุกเหตุการณ์จะวัดความสามารถในการนิยามปัญหา ข้อคำถามที่ 2 วัดความสามารถในการตั้งสมมติฐาน ข้อคำถามที่ 3 วัดความสามารถในการรวบรวมข้อมูล ข้อคำถามที่ 4 วัดความสามารถในการวิเคราะห์ข้อมูล และข้อคำถามที่ 5 วัดความสามารถในการสร้างข้อสรุป ดังภาพที่ 10

ภาพที่ 10 ตัวอย่างแบบสอบถามอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

เวลาสำหรับทำเหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ 1 ข้อคำถามที่ 1	ชื่อ-สกุล ชั้น เลขที่
<p>สถานการณ์ เด็กชายตะวันอยากเก็บสะสมเงินในสกุลอื่น ๆ บ้างนอกจากสกุลเงินบาท ได้แก่ สกุลเงินสกุลหยวน สกุลดอลลาร์สหรัฐ โดยให้แต่ละสกุลเงินที่จะแลกเปลี่ยนมีมูลค่าเท่ากัน เด็กชายตะวันจึงคิดว่าคงต้องเตรียมเงินในสกุลเงินบาทไป 3 ส่วน เพื่อที่จะนำไปแลกเปลี่ยนเป็นสกุลเงินต่าง ๆ ตามที่ต้องการ</p>		
<p>ข้อคำถามที่ 1 คำถาม : จากสถานการณ์ที่กำหนดให้ นักเรียนจะสามารถระบุปัญหาที่ตะวันคิดไว้ว่าอย่างไร (ตอบแบบละเอียดโดยเขียนบริบทสถานการณ์ในโจทย์ประกอบให้ครบถ้วน) คำตอบ</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p>		
<p>ข้อควรระวัง ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>		

ภาพที่ 10 (ต่อ)

เวลาสำหรับทำเหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ 1 ข้อคำถามที่ 2	ชื่อ-สกุล ชั้น เลขที่
<p>สถานการณ์ เด็กชายตะวันระบุปัญหาว่า การนำสกุลเงินบาทไปแลกเปลี่ยนเป็นสกุลเงินอื่น ๆ ตามที่ต้องการในมูลค่าเท่า ๆ กัน จะต้องใช้เงินบาทเป็นจำนวนเท่ากันหรือไม่</p>		
<p>ข้อคำถามที่ 2 คำถาม : จากปัญหา นักเรียนคิดว่าตะวันสามารถตั้งสมมติฐานได้อย่างไร (ตอบแบบละเอียดโดยเขียนบริบทสถานการณ์ในโจทย์ประกอบให้ครบถ้วน) คำตอบ</p>		
<p>ข้อควรระวัง ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>		

ภาพที่ 10 (ต่อ)

เวลาสำหรับทำเหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ 1 ข้อคำถามที่ 3	ชื่อ-สกุล ชั้น เลขที่																									
<p>สถานการณ์ จากปัญหา เด็กชายตะวันตั้งคำถามว่า การจะแลกเงินในสกุลต่าง ๆ ให้มีมูลค่าเท่า ๆ กัน ต้องเตรียมเงินบาทไปแลกในปริมาณที่ไม่เท่ากัน</p>																											
<p>ข้อคำถามที่ 3 คำถาม : หากเด็กชายตะวันค้นหาข้อมูลมาได้ดังนี้</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>บาท</th> <th>เยน</th> <th>ดอลลาร์สหรัฐ</th> <th>หยวน</th> </tr> </thead> <tbody> <tr> <th>บาท</th> <td>50</td> <td>175</td> <td>-</td> <td>-</td> </tr> <tr> <th>เยน</th> <td>-</td> <td>350</td> <td>-</td> <td>20</td> </tr> <tr> <th>ดอลลาร์สหรัฐ</th> <td>-</td> <td>-</td> <td>7.5</td> <td>50</td> </tr> <tr> <th>หยวน</th> <td>-</td> <td>-</td> <td>-</td> <td>75</td> </tr> </tbody> </table> <p>ถ้านักเรียนเป็นเด็กชายตะวัน นักเรียนจะทำการคัดเลือกและรวบรวมข้อมูลที่จำเป็นและเพียงพอเพื่อนำไปสู่การวิเคราะห์ข้อมูล สำหรับการทดสอบสมมติฐานที่ตั้งไว้อย่างไร</p> <p>คำตอบ</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p>				บาท	เยน	ดอลลาร์สหรัฐ	หยวน	บาท	50	175	-	-	เยน	-	350	-	20	ดอลลาร์สหรัฐ	-	-	7.5	50	หยวน	-	-	-	75
	บาท	เยน	ดอลลาร์สหรัฐ	หยวน																							
บาท	50	175	-	-																							
เยน	-	350	-	20																							
ดอลลาร์สหรัฐ	-	-	7.5	50																							
หยวน	-	-	-	75																							
<p>ข้อควรระวัง ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>																											

ภาพที่ 10 (ต่อ)

เวลาสำหรับทำเหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ 1 ข้อคำถามที่ 4	ชื่อ-สกุล ชั้น เลขที่
<p>สถานการณ์ เด็กชายตะวันทำการรวบรวมข้อมูลเพื่อนำไปสู่การทดสอบสมมติฐาน ได้ข้อมูลดังนี้</p> <p>เงิน 350 เยน แลกได้ 20 หยวน</p> <p>เงิน 50 หยวน แลกได้ 7.5 ดอลลาร์สหรัฐ</p> <p>เงิน 50 บาท แลกได้ 175 เยน</p>		
<p>ข้อคำถามที่ 4</p> <p>คำถาม : นักเรียนคิดว่า เด็กชายตะวันควรแสดงผลการวิเคราะห์ข้อมูล เพื่อเปรียบเทียบอัตราค่าเงินสำหรับการแลกเปลี่ยนอย่างไรให้เห็นภาพชัดเจน</p> <p>คำตอบ</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p>		
<p>ข้อควรระวัง ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>		

ภาพที่ 10 (ต่อ)

เวลาสำหรับทำเหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ 1 ข้อคำถามที่ 5	ชื่อ-สกุล ชั้น เลขที่
<p>สถานการณ์ จากความคิดของเด็กชายตะวันที่อยากเก็บสะสมเงินในสกุลอื่น ๆ บ้างนอกจากสกุลเงินบาท เขาจึงมีข้อสงสัยว่า ‘ถ้าต้องการนำสกุลเงินบาทไปแลกเปลี่ยนเป็นสกุลเงินอื่น ๆ ตามที่ต้องการในมูลค่าเท่า ๆ กัน จะต้องใช้เงินบาทเป็นจำนวนเท่ากันหรือไม่’ และจากการที่เด็กชายตะวันได้บันทึกผลการวิเคราะห์ข้อมูลดังนี้</p> <p style="text-align: center;">บาท : เยน : หยวน : ดอลลาร์สหรัฐ = 100 : 350 : 20 : 3</p>		
<p>ข้อคำถามที่ 5</p> <p>คำถาม : ถ้าต้องการแลกเปลี่ยนสกุลต่าง ๆ ได้แก่ สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ สกุลละ 210 หน่วย ต้องเตรียมเงินสกุลบาทไปเป็นจำนวนเท่าไร และผลการวิเคราะห์ข้อมูลนี้สอดคล้องกับสมมติฐานที่ตั้งไว้ข้างต้นหรือไม่</p> <p>คำตอบ</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p>		
<p>ข้อควรระวัง ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์ แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>		

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับวิธีตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric) เป็นเกณฑ์การตรวจให้คะแนนที่แยกองค์ประกอบในการประเมินแต่ละด้าน มีการอธิบายลักษณะของงานในแต่ละระดับความสามารถ ประกอบด้วยรายการย่อย ๆ 2 ชั้น โดยชั้นที่ 1 เป็นเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) ที่อธิบายลักษณะของงานในข้อรายการย่อย ๆ ลดหลั่นตามระดับความสามารถ ส่วนชั้นที่ 2 เป็นการแปลงคะแนนผลรวมของคะแนนที่ได้ในระดับที่ 1 ออกมาให้เป็นไปตามเกณฑ์ ที่กำหนดไว้สำหรับวิธีการกำหนดเกณฑ์ในชั้นที่ 2 ในงานวิจัยที่ผ่านมายังไม่มีวิธีการกำหนดหลักเกณฑ์ที่แน่ชัด สำหรับงานวิจัยฉบับนี้ผู้วิจัยจึงดำเนินการกำหนดหลักเกณฑ์ จากการพิจารณาคะแนนที่ได้ในชั้นที่ 1 โดยนำมาแปลงคะแนนในชั้นที่ 2 ให้มีคะแนนเต็มแต่ละข้อคำถามในทุกเหตุการณ์ให้สอดคล้อง และเทียบเคียงกับเกณฑ์การตรวจให้คะแนนในวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox

ผู้วิจัยจึงได้พัฒนาได้วิธีตรวจให้คะแนนรูบริกสองชั้นสำหรับแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยในแต่ละข้อคำถามวัดตามกระบวนการคิดวิจารณ์ญาณ 5 ขั้นตอน และในแต่ละขั้นตอนมีเกณฑ์การให้คะแนน ดังนี้ ชั้นที่ 1 สำหรับข้อคำถามที่ 1 ถึงข้อ 3 วัด 2 องค์ประกอบ คือ (1) ความถูกต้องของเนื้อหา (4 คะแนน) แบ่งออกเป็น (1.1) ระบุข้อมูลที่สถานการณ์กำหนด (2 คะแนน) (1.2) อธิบายเงื่อนไขของสถานการณ์ (2 คะแนน) และ (2) การเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) แบ่งออกเป็น (2.1) กำหนดขั้นของการวัด (2 คะแนน) (2.2) บริบทขั้นของการวัดอยู่ภายใต้ข้อมูลในสถานการณ์ (2 คะแนน) ซึ่งในแต่ละข้อรายการย่อยใช้มาตรฐานค่า (Rating scale) 3 ระดับ ได้แก่ 2 คะแนน คือ ถูกต้องทั้งหมด, 1 คะแนน คือ ถูกต้องบางส่วน และ 0 คะแนน คือ ไม่ถูกต้องทั้งหมด หรือไม่มีการตอบคำถาม จากนั้นนำผลรวมของคะแนนในระดับชั้นที่ 1 มาแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้ คือ คะแนนเต็มข้อละ 2 คะแนน สำหรับชั้นที่ 1 ของข้อคำถามที่ 4 และข้อคำถามที่ 5 วัดข้อละ 3 องค์ประกอบ คือ (1) ความถูกต้องของเนื้อหา (4 คะแนน) (2) การเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) และ (3) การสื่อความหมายทางคณิตศาสตร์ (4 คะแนน) โดยแบ่งออกเป็น (3.1) เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ (2 คะแนน) (3.2) การคำนวณ (2 คะแนน) ซึ่งในแต่ละข้อรายการย่อยใช้มาตรฐานค่า (Rating scale) 3 ระดับ (2, 1, 0) จากนั้นนำผลรวมของคะแนนในระดับชั้นที่ 1 มาแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่

กำหนดไว้ คือ คะแนนเต็ม ข้อละ 3 คะแนน คะแนนเต็มเหตุการณ์ละ 12 คะแนน รวมทั้งฉบับ 36
คะแนน ดังภาพที่ 11

ภาพที่ 11 ตัวอย่างวิธีให้คะแนนรูปกรีกสองชั้น (Double layer scoring rubric) ในแบบสอบอัตรณ์ัย
ประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 1 นิยามปัญหา	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบทของ สถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์ กำหนด คือ สกุลเงินบาทเป็นสกุลเงินตั้งต้น ที่นำไปแลก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
		1.2 อธิบายเงื่อนไขของ สถานการณ์ คือ ต้องการให้ แต่ละสกุลเงินที่จะแลกมีมูลค่า เท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
	2. การเข้าใจใน สิ่งที่ต้องการวัด - กำหนดปัญหา ภายใต้บริบทของ สถานการณ์	2.1 กำหนดปัญหา คือ ปัญหาที่เกี่ยวข้องกับการเตรียม เงินบาทไปแลกเป็นสกุลเงิน อื่น ๆ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
		2.2 บริบทของการกำหนด ปัญหาอยู่ภายใต้ข้อมูลใน สถานการณ์ คือ สกุลเงินที่นำไปแลก ได้แก่ สกุลเยน สกุลเงิน และ สกุลดอลลาร์สหรัฐ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
คะแนนรวม = คะแนน			คิดเป็น = คะแนน		

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูปกรีก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-8	2
4-6	1
0-3	0

ภาพที่ 11 (ต่อ)

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 2 ตั้งสมมติฐาน	1. ความถูกต้องของเนื้อหา - ระบุข้อมูลและอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนด คือ ต้องเตรียมสกุลเงินบาทไปแลกในปริมาณที่ไม่เท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ คือ สกุลเงินที่ต้องการแลกมีมูลค่าเท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดสมมติฐานภายใต้บริบทของสถานการณ์	2.1 กำหนดสมมติฐาน คือ คาดคะเนปริมาณเงินบาทที่นำไปแลกได้ว่าต้องเตรียมไปในปริมาณที่ไม่เท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
		2.2 บริบทของการกำหนดสมมติฐานอยู่ภายใต้ข้อมูลที่สถานการณ์ คือ สมมติฐานที่เกี่ยวข้องกับการกำหนดปริมาณเงินบาทไปแลกเป็นสกุลเงินอื่น ๆ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
คะแนนรวม = คะแนน			คิดเป็น = คะแนน		

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-8	2
4-6	1
0-3	0

ภาพที่ 11 (ต่อ)

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ชั้นที่ 3 รวบรวม ข้อมูล	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบทของ สถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์ กำหนด คือ แนวทางการรวบรวมข้อมูลที่ เกี่ยวข้องกับอัตราการแลกเปลี่ยน สกุลเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ คือ เหตุผลในการใช้แนวทางเก็บ รวบรวมข้อมูล	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
	2. การเข้าใจใน สิ่งที่ต้องการวัด - กำหนดแนวทางการ รวบรวม ข้อมูลภายใต้ บริบทของ สถานการณ์	2.1 กำหนดแนวทางการรวบรวม ข้อมูล คือ แนวทางการรวบรวมข้อมูลที่ เกี่ยวข้องกับอัตราการแลกเปลี่ยน เงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
		2.2 บริบทของแนวทางการเก็บ รวบรวมข้อมูลภายใต้ข้อมูลใน สถานการณ์ คือ การแลกเปลี่ยน เงินตราระหว่างสกุลเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบ คำถาม
คะแนนรวม = คะแนน			คิดเป็น = คะแนน		

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-8	2
4-6	1
0-3	0

ภาพที่ 11 (ต่อ)

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 4 วิเคราะห์ ข้อมูล	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบทของ สถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์ กำหนด คือ อัตราแลกเปลี่ยนแต่ละคู่ สกุลเงินตามสถานการณ์กำหนด	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ คือ เปรียบเทียบอัตราแลกเปลี่ยนค่าเงิน ครบทุกสกุลเงิน (สกุลบาท สกุลเงิน สกุลหยวน และสกุลดอลลาร์สหรัฐ)	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจใน สิ่งที่ต้องการวัด - วิเคราะห์ข้อมูล ภายใต้บริบทของ สถานการณ์	2.1 วิเคราะห์ข้อมูล คือ เปรียบเทียบอัตราส่วนค่าเงิน แลกเปลี่ยนทุกสกุลเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการวิเคราะห์ข้อมูล อยู่ภายใต้ข้อมูลในสถานการณ์ คือ สกุลเงินที่นำมาเปรียบเทียบ ประกอบด้วย สกุลเงิน บาท สกุลเงิน สกุลหยวน และสกุลดอลลาร์สหรัฐ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
	3. การสื่อ ความหมายทาง คณิตศาสตร์ - เปลี่ยนประโยค ภาษาเป็น ประโยคทาง คณิตศาสตร์ และ แสดงการคำนวณ	3.1 เปลี่ยนประโยคภาษาเป็น ประโยคทางคณิตศาสตร์ คือ เขียน สัญลักษณ์ อัตราส่วนในการ เปรียบเทียบ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
		3.2 แสดงการคำนวณ คือ ทำเป็นอัตราส่วนอย่างต่ำ คือ บาท : เยน : หยวน : ดอลลาร์สหรัฐ = 100 : 350 : 20 : 3 (หรืออัตราส่วน รูปแบบอื่นที่ไม่ใช่เศษส่วนอย่างต่ำ)	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
คะแนนรวม =			คะแนน	คิดเป็น =	คะแนน

ระดับชั้นที่ 2 (คะแนนเต็ม 3 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0

ภาพที่ 11 (ต่อ)

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 5 สร้าง ข้อสรุป	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบทของ สถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนด คือ เตรียมเงินสกุลบาทไปแลกเงินสกุลต่าง ๆ ได้แก่ สกุลเยน สกุลหยวน และสกุลดอลลาร์ สหรัฐ สกุลละ 210 หน่วย	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ คือ อัตราแลกเปลี่ยนเงินสกุลต่าง ๆ มาจาก ผลการทดสอบสมมติฐานของสถานการณ์ ดังนี้ บาท : เยน : หยวน : ดอลลาร์สหรัฐ เท่ากับ 100 : 350 : 20 : 3	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่ง ที่ต้องการวัด - กำหนดข้อสรุป การเตรียมเงินสกุล บาทไปแลก และ สรุปผลสมมติฐาน สอดคล้องกับ บริบทของ สถานการณ์	2.1 เขียนข้อสรุปของการแลกเปลี่ยน และ ผลสมมติฐาน คือ 1) เตรียมเงินสกุลบาทไปเป็นจำนวน 8,110 บาท และ 2) สรุปสมมติฐานว่า การนำสกุลเงินบาทไป แลกเป็นเงินสกุลอื่น ๆ ตามที่ต้องการใน มูลค่าเท่า ๆ กัน จะต้องเตรียมเงินบาทเป็น จำนวนไม่เท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการสร้างข้อสรุปอยู่ภายใต้ ข้อมูลในสถานการณ์ คือ สกุลเงินที่นำมาเปรียบเทียบกับสกุลเงินบาท ประกอบด้วย สกุลเยน สกุลหยวน และสกุล ดอลลาร์สหรัฐ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
	3. การสื่อ ความหมาย ทางคณิตศาสตร์ - เปลี่ยนประโยค ภาษาเป็นประโยค ทางคณิตศาสตร์ และ คำนวณหา คำตอบ	3.1 เปลี่ยนประโยคภาษาเป็นประโยคทาง คณิตศาสตร์ คือ เขียนสัดส่วนในแต่ละคู่การแลกเปลี่ยน สกุลเงิน (บาท-เยน, บาท-หยวน และ บาท-ดอลลาร์สหรัฐ)	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
		3.2 คำนวณหาคำตอบ คือ คำนวณอัตรา การแลกเปลี่ยนแต่ละคู่สกุลบาทกับสกุลเงิน อื่น ๆ ได้อย่างถูกต้อง คือ เตรียมเงินบาทไป แลกเป็นเงินสกุลเยน 60 บาท, เตรียมเงิน บาทไปแลกเป็นเงินสกุลหยวน 1,050 บาท และเตรียมเงินบาทไปแลกเป็นเงินสกุล ดอลลาร์สหรัฐ 7,000 บาท	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้อง หรือ ไม่มีการตอบคำถาม
คะแนนรวม = คะแนน			คิดเป็น = คะแนน		

ระดับชั้นที่ 2 (คะแนนเต็ม 3 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0

ตอนที่ 2 ผลการตรวจสอบคุณภาพของวิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

2.1 ความตรงเชิงเนื้อหาของวิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ผู้วิจัยพัฒนาขึ้น

เมื่อผู้วิจัยพัฒนาวิธีการตรวจให้คะแนนแบบรูปริกสองชั้นแล้ว ผู้วิจัยนำไปให้อาจารย์ที่ปรึกษาวิทยานิพนธ์พิจารณาเพื่อปรับปรุงแก้ไข จากนั้นนำไปให้ผู้ทรงคุณวุฒิ จำนวน 3 คน ตรวจสอบความตรงเชิงเนื้อหา โดยประเมินว่า สถานการณ์มีความสอดคล้องกับมาตรฐานการเรียนรู้ ตัวชี้วัด และสาระการเรียนรู้ทางคณิตศาสตร์หรือไม่ แนวคำตอบมีความสอดคล้องกับข้อคำถามและถูกต้องหรือไม่ และเกณฑ์การให้คะแนนมีความสอดคล้องกับแนวคำตอบหรือไม่ โดยกำหนดระดับความสอดคล้องของข้อคำถามกับกระบวนการคิดวิจารณ์ญาณ ดังนี้ 1 = สอดคล้อง 0 = ไม่แน่ใจว่าสอดคล้อง และ -1 = ไม่สอดคล้อง จากนั้นนำมาหาค่าดัชนีความสอดคล้อง (Item-Objective Congruence: IOC) โดยกำหนดเกณฑ์ความสอดคล้องของผู้ทรงคุณวุฒิ ≥ 0.5 สำหรับข้อคำถามใดที่มีค่าระดับความสอดคล้องของผู้ทรงคุณวุฒิ และสำหรับข้อคำถามใดที่มีค่าดัชนีความสอดคล้อง < 0.5 ผู้วิจัยจะนำมาปรับปรุงตามข้อเสนอแนะของผู้ทรงคุณวุฒิต่อไป

ผลการตรวจสอบความตรงเชิงเนื้อหาของแนวคำตอบกับวิธีการตรวจให้คะแนนรูปริกสองชั้น พบว่า วิธีการตรวจให้คะแนนรูปริกสองชั้นที่ผู้วิจัยพัฒนาขึ้นเพื่อใช้ตรวจให้คะแนนข้อคำถามของเหตุการณ์ในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ทุกข้อมีระดับความสอดคล้องกับข้อคำถามและแนวคำตอบเมื่อเทียบกับเกณฑ์ ≥ 0.5 แสดงว่า ข้อคำถามมีความสอดคล้องกับเนื้อหา และขั้นของกระบวนการคิดวิจารณ์ญาณ และเมื่อพิจารณาคำแนะนำที่ได้จากผู้เชี่ยวชาญได้ข้อมูล โดยสรุปว่า ทุกเหตุการณ์ในความหมายของเกณฑ์ให้คะแนน 0 คะแนน ควรปรับภาษา จาก 'ไม่ถูกต้องทั้งหมด' เป็น "ไม่ถูกต้อง" โดยเหตุการณ์ที่ 1 ชั้น 4) การวิเคราะห์ข้อมูล ในด้านองค์ประกอบของความถูกต้องของเนื้อหา ควรมีการปรับคำอธิบายของเกณฑ์ให้กระชับและครอบคลุมมากขึ้น เหตุการณ์ที่ 2 ชั้น 2) การตั้งสมมติฐานในด้านองค์ประกอบของการเข้าใจสิ่งที่ต้องการวัด ควรปรับเนื้อหาในแนวคำตอบให้ใช้ความรู้ในระดับชั้นมัธยมศึกษาตอนต้น เพื่อให้สอดคล้องกับระดับความสามารถของผู้สอบ และชั้น 4) การวิเคราะห์ข้อมูลในด้านองค์ประกอบของการสื่อความหมายทางคณิตศาสตร์ ควรเพิ่มกรณีแนวคำตอบทุกกรณีที่เป็นไปได้ และเหตุการณ์ที่ 3 ชั้น 4) การวิเคราะห์ข้อมูล ในด้าน

องค์ประกอบของการสื่อความหมายทางคณิตศาสตร์ ควรปรับคำอธิบายของข้อรายการย่อยให้ สอดคล้องกับการสื่อความหมายทางคณิตศาสตร์ ดังตาราง 15



ตาราง 15 ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถามขั้นตอนการวัดทักษะการคิดวิเคราะห์โดยวิธีตรวจสอบ (double layer scoring rubric)

เหตุการณ์ (O)	ข้อคำถาม (I)	ความตรงเชิงเนื้อหาของวิธีตรวจสอบแบบนรูบริการสองชั้น					
		ความถูกต้องของเนื้อหา	การเข้าใจในสิ่งที่ต้องการวัด	การสื่อความหมายทาง คณิตศาสตร์	IOC	IOC	ข้อเสนอแนะ
1	1.1 การนิยาม ปัญหา	0.8	0.8	-	-	-	ไม่ถูกต้องประเภทอื่น
	1.2 การ ตั้งสมมติฐาน	0.8	0.8	-	-	-	ไม่ถูกต้องประเภทอื่น
	1.3 การรวบรวม ข้อมูล	0.8	0.8	0.8	0.8	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง ทั้งหมด' เป็น 'ไม่ถูกต้อง'
	1.4 การวิเคราะห์ ข้อมูล	0.8	0.8	0.8	0.8	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้ ครอบคลุมทุกประเด็น
	1.5 การสร้าง ข้อสรุป	0.8	0.8	0.8	0.8	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง ทั้งหมด' เป็น 'ไม่ถูกต้อง'

ตารางที่ 15 (ต่อ)

		ความตรงเชิงเนื้อหาของวิธีตรวจให้คะแนนรูปrikสองชั้น			
เหตุการณ์ (O)	ข้อคำถาม (I)	ความถูกต้องของเนื้อหา	การเข้าใจในสิ่งที่ต้องการวัด	การสื่อความหมายทาง คณิตศาสตร์	
		IOC	ข้อเสนอแนะ	IOC	ข้อเสนอแนะ
2.1	การนิยาม ปัญหา	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง ทั้งหมด' เป็น "ไม่ถูกต้อง"	-	ไม่ถูกต้องประกอบนี้
		0.8	ปรับเนื้อหาในแนวคำตอบให้ ใช้ความรู้ในระดับชั้น มัธยมศึกษาตอนต้น	-	ไม่ถูกต้องประกอบนี้
2.3	การรวบรวม ข้อมูล	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง ทั้งหมด' เป็น "ไม่ ถูกต้อง"	0.8	ไม่ถูกต้องประกอบนี้
		0.8	ปรับภาษา จาก 'ไม่ถูกต้อง ทั้งหมด' เป็น "ไม่ถูกต้อง"	0.8	ปรับภาษาคำอธิบายเกณฑ์ให้ ครอบคลุมทุกประเด็น
2.5	การสร้าง ข้อสรุป	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง ทั้งหมด' เป็น "ไม่ถูกต้อง"	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง ทั้งหมด' เป็น "ไม่ถูกต้อง"

ตารางที่ 15 (ต่อ)

เหตุการณ์ (O)	ข้อคำถาม (I)	ความตรงเชิงเนื้อหาของวิธีตรวจให้คะแนนรัฐบริกของชั้น			
		ความถูกต้องของเนื้อหา	การเข้าใจในสิ่งที่ต้องการวัด	การสื่อความหมายทาง คณิตศาสตร์	ข้อเสนอนะ
		IOC	IOC	IOC	IOC
3	3.1 การนิยาม ปัญหา	0.8	0.8	-	ไม่ถูกต้องประกอบนี้
	3.2 การ ตั้งสมมติฐาน	0.8	0.8	-	ไม่ถูกต้องประกอบนี้
	3.3 การรวบรวม ข้อมูล	0.8	0.8	-	ไม่ถูกต้องประกอบนี้
	3.4 การวิเคราะห์ ข้อมูล	0.8	0.8	0.8	ปรับคำอธิบายของข้อรายการ ขอยให้สอดคล้องกับการสื่อ ความทางคณิตศาสตร์
	3.5 การสร้าง ข้อสรุป	0.8	0.8	0.8	ปรับภาษา จาก 'ไม่ถูกต้อง ทั้งหมด' เป็น 'ไม่ถูกต้อง'

2.2 ค่าสถิติเบื้องต้นของคะแนนที่ใช้วิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบ
 อัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

ผู้วิจัยนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์
 สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่เก็บข้อมูลจากตัวอย่างวิจัยจำนวน 90 คน มาให้ผู้ตรวจให้
 คะแนน คือ ผู้ที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะ
 ชำนาญการ (ศศ.2) ขึ้นไป จำนวน 2 คน โดยก่อนทำการตรวจให้คะแนน ผู้วิจัยได้อธิบายเกณฑ์
 การตรวจให้คะแนนรูปริกสองชั้นในการตรวจให้คะแนนกับผู้ประเมินก่อนการตรวจให้คะแนน
 โดยที่ผู้ตรวจให้คะแนนทั้งสองคนจะตรวจแบบสอบทุกฉบับอย่างเป็นอิสระต่อกัน และตรวจแบบ
 สอบทุกฉบับจำนวน 2 ครั้ง ห่างกัน 4 สัปดาห์ จากนั้นนำคะแนนที่ได้จากผู้ประเมินมาวิเคราะห์
 หาค่าสถิติพื้นฐานโดยใช้สถิติบรรยาย คือ ค่าเฉลี่ย (M) ส่วนเบี่ยงเบนมาตรฐาน (SD) คะแนน
 สูงสุด (Maximum) และคะแนนต่ำสุด (Minimum) ของแต่ละเหตุการณ์ และทั้งฉบับ

ผลการวิเคราะห์ค่าสถิติพื้นฐานแบบสอบอัตนัยประยุกต์วัดทักษะการคิด
 วิจารณ์ญาณทางคณิตศาสตร์ที่ตรวจให้คะแนนด้วยวิธีตรวจให้คะแนนรูปริกสองชั้น จำนวน 3
 เหตุการณ์ เหตุการณ์ละ 12 คะแนน รวมคะแนนทั้งฉบับ 36 คะแนน เมื่อพิจารณาคะแนนที่ได้จาก
 ผู้ประเมินคนที่ 1 พบว่า การตรวจให้คะแนนครั้งที่ 1 มีค่าเฉลี่ยของคะแนนทั้งฉบับ คือ 23.94
 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 7.17 คะแนน การตรวจให้คะแนน ครั้งที่ 2 มีค่าเฉลี่ยของคะแนน
 ทั้งฉบับ คือ 26.00 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 8.40 คะแนน เมื่อพิจารณาคะแนนที่ได้จากผู้
 ประเมินคนที่ 2 พบว่า การตรวจให้คะแนนครั้งที่ 1 มีค่าเฉลี่ยของคะแนนทั้งฉบับ คือ 26.03
 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 8.53 คะแนน การตรวจให้คะแนนครั้งที่ 2 มีค่าเฉลี่ยของคะแนน
 ทั้งฉบับ คือ 25.03 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 7.67 คะแนน ดังตาราง 16

ตาราง 16 ค่าสถิติพื้นฐานของคะแนนรายข้อของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิเคราะห์ตามทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนนรูปภิกสองชั้น (double layer scoring rubric) จากผู้ประเมินคนที่ 1 และผู้ประเมินคนที่ 2 ที่ตรวจให้คะแนน 2 ครั้ง

ผู้ประเมิน	แบบสอบ	ตรวจครั้งที่ 1					ตรวจครั้งที่ 2				
		Max	Min	M	SD	Max	Min	M	SD		
1	เหตุการณ์ที่ 1	12	2	9.14	2.43	12	2	9.64	2.45		
	เหตุการณ์ที่ 2	12	1	7.52	2.98	12	1	8.11	3.40		
	เหตุการณ์ที่ 3	12	2	7.28	2.73	12	1	8.24	3.15		
	รวมทั้งฉบับ	36	6	23.94	7.17	36	5	26.00	8.40		
2	เหตุการณ์ที่ 1	12	1	9.53	2.58	12	2	9.14	2.42		
	เหตุการณ์ที่ 2	12	1	8.17	3.43	12	1	7.78	3.17		
	เหตุการณ์ที่ 3	12	2	8.33	3.14	12	2	8.11	2.95		
	รวมทั้งฉบับ	36	5	26.03	8.53	36	7	25.03	7.67		

2.3 ความเที่ยงในการตรวจให้คะแนนจากผู้ประเมินคนเดียว (intra rater reliability)

ผู้วิจัยนำคะแนนที่ได้จากการตรวจแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิเคราะห์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ทำการตรวจให้คะแนนแบบ สอบ 90 ฉบับ ด้วยวิธีตรวจให้คะแนนรูปริกสองชั้นจากผู้ประเมิน คือ ผู้ที่ปฏิบัติการสอน คณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือ มีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน ที่ทำการตรวจให้คะแนนแบบสอบทุกฉบับ และตรวจแบบสอบทุกฉบับจำนวน 2 ครั้ง ห่างกัน 4 สัปดาห์ มาหาความสัมพันธ์โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation)

ผลการตรวจให้คะแนนจากผู้ประเมินคนที่ 1 ที่ตรวจให้คะแนน 2 ครั้ง พบว่า ค่าสัมประสิทธิ์ของคะแนนที่ได้จากการตรวจของผู้ประเมินคนเดียว โดยตรวจแบบสอบทุกฉบับ จำนวน 2 ครั้ง ห่างกันครั้งละ 4 สัปดาห์ ด้วยวิธีตรวจให้คะแนนรูปริกสองชั้น มีนัยสำคัญทางสถิติที่ ระดับ 0.1 ทุกเหตุการณ์ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน อยู่ในช่วง 0.875 ถึง 0.935 และเมื่อพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ในแต่ละเหตุการณ์ พบว่า เหตุการณ์ที่ 2 มีค่าสัมประสิทธิ์ของคะแนนมากที่สุด คือ 0.922 รองลงมา คือ เหตุการณ์ที่ 1 และ 3 ตามลำดับ และเมื่อพิจารณาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานของผู้ตรวจคนเดียวที่ทำการตรวจให้คะแนนทั้ง 2 ครั้ง มีค่าใกล้เคียงกันทุกเหตุการณ์ จากข้อมูลดังกล่าวสรุปได้ว่า แบบสอบอัตนัยประยุกต์วัดทักษะ การคิดวิเคราะห์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้ คะแนนรูปริกสองชั้นที่ผู้วิจัยพัฒนาขึ้น มีความเป็นปรนัยในการตรวจให้คะแนน ไม่ว่าจะผู้ประเมินจะ ตรวจให้คะแนนกี่ครั้ง คะแนนที่ได้ก็จะมีค่าใกล้เคียงกัน ดังตาราง 17

ตาราง 17 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ประเมินคนที่ 1 ที่ตรวจให้คะแนน 2 ครั้ง

เหตุการณ์ (O)	ผู้ประเมินคนที่ 1				r_{xy}
	ตรวจครั้งที่ 1		ตรวจครั้งที่ 2		
	M_1	SD_1	M_2	SD_2	
เหตุการณ์ที่ 1	9.14	2.43	9.64	2.45	0.902**
เหตุการณ์ที่ 2	7.52	2.98	8.11	3.40	0.922**
เหตุการณ์ที่ 3	7.28	2.73	8.24	3.15	0.875**
ทั้งหมด	23.94	7.17	26.00	8.40	0.935**

** $p < .1$

ผลการตรวจให้คะแนนจากผู้ประเมินคนที่ 2 ที่ตรวจให้คะแนน 2 ครั้ง พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ประเมินคนเดียว โดยตรวจแบบสอบทุกฉบับ จำนวน 2 ครั้ง ห่างกันครั้งละ 4 สัปดาห์ ด้วยวิธีตรวจให้คะแนนรูปริกสองชั้น มีนัยสำคัญทางสถิติที่ระดับ 0.1 ทุกเหตุการณ์ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน อยู่ในช่วง 0.914 ถึง 0.966 และเมื่อพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ในแต่ละเหตุการณ์ พบว่า เหตุการณ์ที่ 2 มีค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนมากที่สุด คือ 0.950 รองลงมา คือ เหตุการณ์ที่ 3 และ 1 ตามลำดับ และเมื่อพิจารณาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานของผู้ตรวจคนเดียวที่ทำการตรวจให้คะแนนทั้ง 2 ครั้ง มีค่าใกล้เคียงกันทุกเหตุการณ์ จากข้อมูลดังกล่าวสรุปได้ว่า แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนรูปริกสองชั้นที่ผู้วิจัยพัฒนาขึ้น มีความเป็นปรนัยในการตรวจให้คะแนน ไม่ว่าจะผู้ประเมินจะตรวจให้คะแนนกี่ครั้ง คะแนนที่ได้ก็จะมีค่าใกล้เคียงกัน ดังตาราง 18

ตาราง 18 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ประเมินคนที่ 2 ที่ตรวจให้คะแนน 2 ครั้ง

เหตุการณ์ (O)	ผู้ประเมินคนที่ 2				r_{xy}
	ตรวจครั้งที่ 1		ตรวจครั้งที่ 2		
	M_1	SD_1	M_2	SD_2	
เหตุการณ์ที่ 1	9.53	2.58	9.14	2.42	0.914**
เหตุการณ์ที่ 2	8.17	3.43	7.78	3.17	0.950**
เหตุการณ์ที่ 3	8.33	3.14	8.11	2.95	0.937**
ทั้งหมด	26.03	8.53	25.03	7.67	0.966**

**p < .1

2.4 ความเที่ยงในการตรวจให้คะแนนจากผู้ประเมิน 2 คน (inter rater reliability)

ผู้วิจัยนำคะแนนที่ได้จากการตรวจแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิเคราะห์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ทำการตรวจให้คะแนนแบบ สอบ 90 ฉบับ ด้วยวิธีตรวจให้คะแนนรูปริกสองชั้นจากผู้ประเมิน คือ ผู้ที่ปฏิบัติการสอน คณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือ มีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน ที่ทำการตรวจให้คะแนนแบบสอบทุกฉบับอย่างเป็นอิสระมาหาความสัมพันธ์โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation)

ผลการตรวจให้คะแนนจากผู้ประเมิน 2 คน พบว่า ค่าสัมประสิทธิ์ของคะแนนที่ได้จากการตรวจของผู้ประเมิน 2 คนอย่างเป็นอิสระ โดยตรวจแบบสอบทุกฉบับด้วยวิธีตรวจให้คะแนนรูปริกสองชั้น มีนัยสำคัญทางสถิติที่ระดับ 0.1 ทุกเหตุการณ์ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันในแต่ละเหตุการณ์อยู่ในช่วง 0.797 ถึง 0.844 และค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันทั้งหมดมีค่าเท่ากับ 0.877 ซึ่งให้ผลสอดคล้องกับค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (intra-class correlation: ICC) ที่ระดับความเชื่อมั่น 95% ในแต่ละเหตุการณ์อยู่ในช่วง 0.789 ถึง 0.836 และค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้นทั้งหมดมีค่าเท่ากับ 0.864 แสดงว่า ทุกเหตุการณ์มีความสอดคล้องในการตรวจให้คะแนนระหว่างผู้ประเมินอยู่ในระดับดี และเมื่อพิจารณา ค่าสัมประสิทธิ์สหสัมพันธ์และค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้นในแต่ละเหตุการณ์ พบว่า เหตุการณ์ที่ 2 มีค่าสัมประสิทธิ์สหสัมพันธ์และค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้นของคะแนนมากที่สุดและให้ผลสอดคล้องกัน คือ 0.844 และ 0.836 ตามลำดับ รองลงมา คือ เหตุการณ์ที่ 1 และ 3 ตามลำดับ และเมื่อพิจารณาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานของผู้ประเมิน

ทั้ง 2 คน มีค่าใกล้เคียงกันทุกเหตุการณ์ จากข้อมูลดังกล่าวสรุปได้ว่า แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนรูปกรีกสองชั้น ที่ผู้วิจัยพัฒนาขึ้นมีความเป็นปรนัยในการตรวจให้คะแนน ไม่ว่าจะผู้ใดตรวจให้คะแนน คะแนนที่ได้ก็จะมีค่าใกล้เคียงกัน ดังตาราง 19

ตาราง 19 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ประเมิน 2 คน

เหตุการณ์ (O)	ผู้ประเมินคนที่ 1		ผู้ประเมินคนที่ 2		r_{xy}	ICC _{3,2}
	M_1	SD_1	M_2	SD_2		
เหตุการณ์ที่ 1	9.14	2.43	9.53	2.58	0.809**	0.808
เหตุการณ์ที่ 2	7.52	2.98	8.17	3.43	0.844**	0.836
เหตุการณ์ที่ 3	7.28	2.73	8.33	3.14	0.797**	0.789
ทั้งฉบับ	23.94	7.17	26.03	8.53	0.877**	0.864

**p < .1

ตอนที่ 3 ผลการเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณ ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง

1.1 ผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีการตรวจให้คะแนน 3 วิธี

ผู้วิจัยเก็บข้อมูลโดยให้ตัวอย่างวิจัย จำนวน 90 คน ทำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 3 เหตุการณ์ เหตุการณ์ละ 5 ข้อคำถาม โดยมีเกณฑ์ในการเลือกผู้ประเมิน คือ ผู้ที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน ตรวจให้คะแนนแต่ละเหตุการณ์ โดยใช้วิธีตรวจให้คะแนนรูปกรีกสองชั้น เทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox จากนั้นนำคะแนน ที่ได้มาหาความ

เที่ยงของแบบสอบด้วยค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ตามทฤษฎีการสรุปอ้างอิง (G-Theory)

การวิเคราะห์ความแปรปรวนขององค์ประกอบจากแหล่งความแปรปรวนต่าง ๆ ที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกัน ได้ออกแบบการวัดแบบไขว้ (cross design) 2 องค์ประกอบ คือ $p \times r \times o$ เมื่อ p แทน ผู้สอบ r แทน ผู้ประเมิน และ o แทน เหตุการณ์ นั่นคือ นักเรียนทุกคนทำแบบสอบอัตนัยประยุกต์ ใช้ผู้ตรวจให้คะแนนตรวจแบบสอบทุกเหตุการณ์ โดยใช้โปรแกรมสำเร็จรูป EduG

ผลการวิเคราะห์ความแปรปรวนขององค์ประกอบจากแหล่งความแปรปรวน 7 แหล่ง ที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนรูปภิกสองชั้น (double layer scoring rubric) พบว่า ความแปรปรวนของคะแนนจริง (σ^2_p) มีค่าน้อย คิดเป็นร้อยละ 20.7 นอกนั้นเป็นความแปรปรวนของ ความคลาดเคลื่อนจากแหล่งอื่น ๆ จำนวน 6 แหล่ง ได้แก่ ความแปรปรวนของคะแนนเฉลี่ยระหว่างผู้ประเมิน (σ^2_r) คิดเป็นร้อยละ 0.2 ความแปรปรวนของคะแนนเฉลี่ยระหว่างเหตุการณ์ (σ^2_o) คิดเป็นร้อยละ 20.6 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบกับผู้ประเมิน (σ^2_{pr}) คิดเป็นร้อยละ 2.1 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบกับเหตุการณ์ (σ^2_{po}) คิดเป็นร้อยละ 17.8 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้ประเมินกับเหตุการณ์ (σ^2_{ro}) คิดเป็นร้อยละ 0.8 และความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบ, ผู้ประเมิน และเหตุการณ์ (σ^2_{pro}) คิดเป็นร้อยละ 37.8 แสดงให้เห็นว่า มีความคลาดเคลื่อนที่เหลือที่ไม่สามารถวัดได้ค่อนข้างมาก นอกจากนี้จะเห็นได้ว่าความแปรปรวนของคะแนนเฉลี่ยระหว่างเหตุการณ์ (σ^2_o) มีค่ามากกว่าความแปรปรวนของคะแนนเฉลี่ยระหว่างผู้ประเมิน (σ^2_r) ดังนั้น หากต้องการลดความคลาดเคลื่อนของการวัดในสถานการณ์นี้ ควรให้ความสำคัญกับจำนวนเหตุการณ์ที่ใช้ ดังตาราง 20

ตาราง 20 ผลการวิเคราะห์ความแปรปรวน G-Study for $p \times r \times o$ ของแบบสอบถามดัชนีประยุทธิ์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนรูปกรสองชั้น

Source of variation	G-Study สำหรับวิธีตรวจให้คะแนนรูปกรสองชั้น				
	df	SS	MS	Estimated Error Component	Percentage of Error Variance
P	89	523.3274	5.8801	0.1676	20.7
R	1	3.2726	3.2726	0.0016	0.2
O	14	437.9052	31.2789	0.1673	20.6
PR	89	50.0607	0.5625	0.0171	2.1
PO	1246	741.5615	0.5952	0.1442	17.8
RO	14	12.3941	0.8853	0.0064	0.8
PRO,e	1246	382.2726	0.3068	0.3068	37.8
Total	2699	2150.7941	-	0.8110	100%

ผลการวิเคราะห์ความแปรปรวนขององค์ประกอบจากแหล่งความแปรปรวน 7 แหล่ง ที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบถามดัชนีประยุทธิ์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) พบว่า ความแปรปรวนของคะแนนจริง (σ^2_p) มีค่าน้อย คิดเป็นร้อยละ 16.8 นอกนั้นเป็นความแปรปรวนของความคลาดเคลื่อนจากแหล่งอื่น ๆ จำนวน 6 แหล่ง ได้แก่ ความแปรปรวนของคะแนนเฉลี่ยระหว่างผู้ประเมิน (σ^2_r) คิดเป็นร้อยละ 0.4 ความแปรปรวนของคะแนนเฉลี่ยระหว่างเหตุการณ์ (σ^2_o) คิดเป็นร้อยละ 19.0 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบกับผู้ประเมิน (σ^2_{pr}) คิดเป็นร้อยละ 3.1 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบกับเหตุการณ์ (σ^2_{po}) คิดเป็นร้อยละ 12.6 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้ประเมินกับเหตุการณ์ (σ^2_{ro}) คิดเป็นร้อยละ 1.5 และความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบ, ผู้ประเมิน และเหตุการณ์ (σ^2_{pro}) คิดเป็นร้อยละ 46.6 แสดงให้เห็นว่า มีความคลาดเคลื่อนที่เหลือที่ไม่สามารถวัดได้ค่อนข้างมาก นอกจากนี้จะเห็นได้ว่าความแปรปรวนของคะแนนเฉลี่ยระหว่างเหตุการณ์ (σ^2_o) มีค่ามากกว่าความแปรปรวนของคะแนนเฉลี่ยระหว่างผู้ประเมิน

(σ^2_R) ดังนั้น หากต้องการลดความคลาดเคลื่อนของการวัดในสถานการณ์นี้ ควรให้ความสำคัญกับจำนวนเหตุการณ์ที่ใช้ ดังตาราง 21

ตาราง 21 ผลการวิเคราะห์ความแปรปรวน G-Study for $p \times r \times o$ ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนวิเคราะห์ย่อย

G-Study สำหรับวิธีตรวจให้คะแนนวิเคราะห์ย่อย					
Source of variation	df	SS	MS	Estimated Error Component	Percentage of Error Variance
P	89	1294.3263	14.5430	0.3927	16.8
R	1	16.8033	16.8033	0.0085	0.4
O	14	1191.9652	85.1404	0.4458	19.0
PR	89	193.2300	2.1711	0.0720	3.1
PO	1246	2095.9682	1.6822	0.2953	12.6
RO	14	60.3800	4.3129	0.0358	1.5
PRO,e	1246	1360.0867	1.0916	1.0916	46.6
Total	2699	6212.7596	-	2.3417	100%

ผลการวิเคราะห์ความแปรปรวนขององค์ประกอบจากแหล่งความแปรปรวน 7 แหล่ง ที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนน Knox พบว่า ความแปรปรวนของคะแนนจริง (σ^2_p) มีค่าน้อย คิดเป็นร้อยละ 7.6 นอกนั้นเป็นความแปรปรวนของความคลาดเคลื่อนจากแหล่งอื่น ๆ จำนวน 6 แหล่ง ได้แก่ ความแปรปรวนของคะแนนเฉลี่ยระหว่างผู้ประเมิน (σ^2_R) คิดเป็นร้อยละ 1.9 ความแปรปรวนของคะแนนเฉลี่ยระหว่างเหตุการณ์ (σ^2_O) คิดเป็นร้อยละ 8.0 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบกับผู้ประเมิน (σ^2_{PR}) คิดเป็นร้อยละ 0.4 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบกับเหตุการณ์ (σ^2_{PO}) คิดเป็นร้อยละ 54.0 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้ประเมินกับเหตุการณ์ (σ^2_{RO}) คิดเป็นร้อยละ 0.3 และความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบ, ผู้ประเมิน และเหตุการณ์ (σ^2_{PRO}) คิดเป็นร้อยละ 27.7 จะเห็นได้ว่าความแปรปรวนของคะแนนเฉลี่ยระหว่างเหตุการณ์ (σ^2_O) มีค่า

มากกว่าความแปรปรวนของคะแนนเฉลี่ยระหว่างผู้ประเมิน (σ^2_R) ดังนั้น หากต้องการลดความคลาดเคลื่อนของการวัดในสถานการณ์นี้ ควรให้ความสำคัญกับจำนวนเหตุการณ์ที่ใช้ ดังตาราง 22

ตาราง 22 ผลการวิเคราะห์ความแปรปรวน G-Study for $p \times r \times o$ ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนน Knox

Source of variation	G-Study สำหรับวิธีตรวจให้คะแนน Knox				
	df	SS	MS	Estimated Error Component	Percentage of Error Variance
P	89	116.1574	1.3051	0.0268	7.6
R	1	9.1293	9.1293	0.0066	1.9
O	14	79.6407	5.6886	0.0283	8.0
PR	89	10.6374	0.1195	0.0015	0.4
PO	1246	596.6926	0.4789	0.1906	54.0
RO	14	2.8985	0.2070	0.0012	0.3
PRO,e	1246	121.8348	0.0978	0.0978	27.7
Total	2699	936.9907	-	0.3528	100%

3.2 ผลการศึกษาเพื่อตัดสินใจการสรุปอ้างอิง (D-Study) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปกรีกสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox

การวิเคราะห์เป็นการนำข้อมูลจาก G-Study มาสรุปอ้างอิงคุณภาพของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปกรีกสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox เพื่อลด ความคลาดเคลื่อนของการวัด โดยการตรวจสอบคุณภาพด้านความเที่ยงด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบภายใต้การออกแบบการวัด $p \times r \times o$ เมื่อใช้ผู้ประเมินจำนวน 1 และ 2 คน ภายใต้จำนวนเหตุการณ์ที่ใช้ 2, 3, 4, 5 และ 6 เหตุการณ์ เหตุการณ์ละ 5 ข้อคำถาม โดยใช้วิธีตรวจให้คะแนนรูปกรีกสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์

ย่อยและวิธีตรวจให้คะแนน Knox ผู้วิจัยเลือกค่าความเที่ยงตั้งแต่ 0.7 ขึ้นไป ตามเกณฑ์ที่ Nunnally (1978) เสนอไว้ ผลการวิเคราะห์ดังนี้

สำหรับวิธีตรวจให้คะแนนรูปภิกสองชั้นสำหรับแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ พบว่า เมื่อใช้ผู้ประเมิน 1 คน ควรมีจำนวน 2 เหตุการณ์ จึงจะได้ค่าสัมประสิทธิ์ การสรุปอ้างอิงตัดสินใจเชิงสัมพัทธ์ (relative coefficient) เท่ากับ 0.73 และค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) เท่ากับ 0.67 เมื่อใช้ผู้ประเมิน 2 คน ควรมีจำนวน 2 เหตุการณ์ จึงจะได้ค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมพัทธ์ เท่ากับ 0.81 และค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมบูรณ์ เท่ากับ 0.73 ดังตาราง 23

ตาราง 23 ผลการศึกษา D (D-Study of p x r x o design) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนรูปภิกสองชั้น

		D-Study สำหรับวิธีตรวจให้คะแนนรูปภิกสองชั้น									
		$n_r=1$					$n_r=2$				
Effect		$n_o=2$	$n_o=3$	$n_o=4$	$n_o=5$	$n_o=6$	$n_o=2$	$n_o=3$	$n_o=4$	$n_o=5$	$n_o=6$
Error	σ^2_δ	0.06	0.05	0.04	0.04	0.03	0.04	0.03	0.02	0.02	0.02
Variance	σ^2_Δ	0.08	0.06	0.05	0.04	0.04	0.06	0.04	0.03	0.03	0.02
G-	ρ^2_δ	0.73	0.78	0.81	0.83	0.84	0.81	0.86	0.88	0.89	0.90
Coefficient	ρ^2_Δ	0.67	0.74	0.77	0.79	0.81	0.73	0.81	0.84	0.86	0.87

เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยสำหรับแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ พบว่า เมื่อใช้ผู้ประเมิน 1 คน ควรมีจำนวน 3 เหตุการณ์ จึงจะได้ค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมพัทธ์ เท่ากับ 0.70 และค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมบูรณ์ เท่ากับ 0.66 เมื่อใช้ผู้ประเมิน 2 คน ควรมีจำนวน 2 เหตุการณ์ จึงจะได้ค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมพัทธ์ เท่ากับ 0.77 และค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมบูรณ์ เท่ากับ 0.70 ดังตาราง 24

ตาราง 24 ผลการศึกษา D (D-Study of $p \times r \times o$ design) ของแบบสอบถามระดับมัธยมศึกษาการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนนวิเคราะห์ย่อย

D-Study สำหรับวิธีตรวจให้คะแนนวิเคราะห์ย่อย											
Effect		$n_r=1$					$n_r=2$				
		$n_o=2$	$n_o=3$	$n_o=4$	$n_o=5$	$n_o=6$	$n_o=2$	$n_o=3$	$n_o=4$	$n_o=5$	$n_o=6$
Error	σ^2_δ	0.21	0.16	0.14	0.13	0.12	0.12	0.09	0.08	0.07	0.06
Variance	σ^2_Δ	0.27	0.20	0.17	0.16	0.14	0.17	0.13	0.11	0.09	0.08
G-	ρ^2_δ	0.65	0.70	0.74	0.75	0.77	0.77	0.81	0.83	0.85	0.86
Coefficient	ρ^2_Δ	0.60	0.66	0.69	0.72	0.73	0.70	0.76	0.79	0.81	0.82
t											

เมื่อเทียบกับวิธีตรวจให้คะแนน Knox สำหรับแบบสอบถามระดับมัธยมศึกษาการคิดวิจารณ์ทางคณิตศาสตร์ พบว่า เมื่อใช้ผู้ประเมิน 1 คน ควรมีจำนวน 6 เหตุการณ์ จึงจะได้ค่าสัมประสิทธิ์ การสรุปอ้างอิงตัดสินใจเชิงสัมพัทธ์ เท่ากับ 0.71 และค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมบูรณ์ เท่ากับ 0.59 เมื่อใช้ผู้ประเมิน 2 คน ควรมีจำนวน 5 เหตุการณ์ จึงจะได้ค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมพัทธ์ เท่ากับ 0.72 และค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิงสัมบูรณ์ เท่ากับ 0.65 ดังตาราง 25

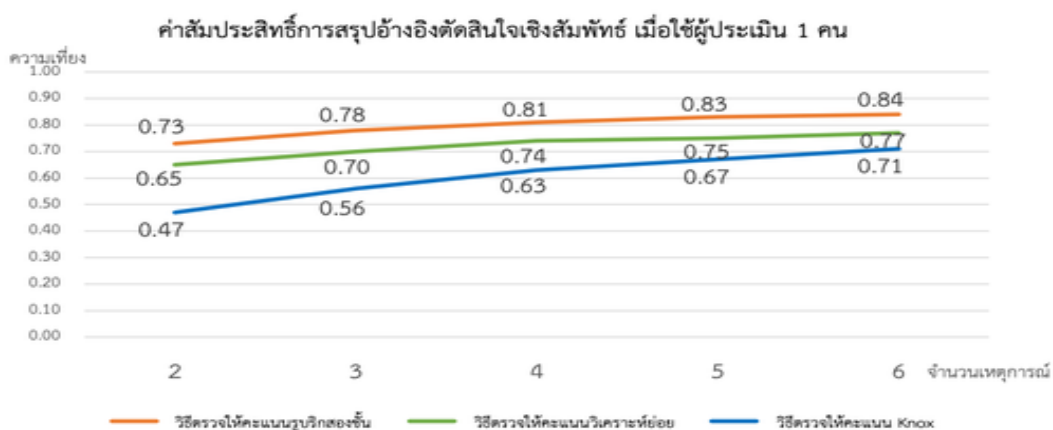
ตาราง 25 ผลการศึกษา D (D-Study of $p \times r \times o$ design) ของแบบสอบถามระดับมัธยมศึกษาการคิดวิเคราะห์ การคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีตรวจให้คะแนน Knox

D-Study สำหรับวิธีตรวจให้คะแนน Knox											
Effect		$n_r=1$					$n_r=2$				
		$n_o=2$	$n_o=3$	$n_o=4$	$n_o=5$	$n_o=6$	$n_o=2$	$n_o=3$	$n_o=4$	$n_o=5$	$n_o=6$
Error	σ^2_δ	0.03	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.01	0.01
Variance	σ^2_Δ	0.04	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.01	0.01
G-	ρ^2_δ	0.47	0.56	0.63	0.67	0.71	0.52	0.62	0.68	0.72	0.75
Coefficient	ρ^2_Δ	0.40	0.48	0.53	0.56	0.59	0.47	0.55	0.61	0.65	0.67

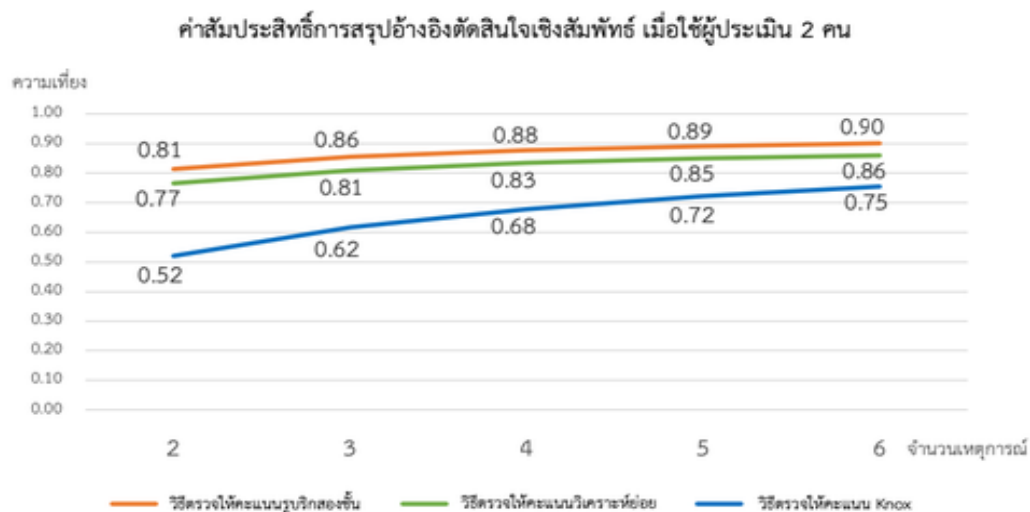
สรุปได้ว่า แบบสอบถามระดับมัธยมศึกษาการคิดวิเคราะห์ การคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปกรอกสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ในระดับความเที่ยงที่ผู้วิจัยยอมรับตั้งแต่ 0.7 ขึ้นไป พิจารณากรณีจำนวนผู้ประเมินเท่ากัน หากเพิ่มจำนวนเหตุการณ์จะทำให้วิธีตรวจให้คะแนนรูปกรอกสองชั้นมีค่าความเที่ยงของแบบสอบถามสูงขึ้น และเมื่อพิจารณาจำนวนผู้ประเมินต่างกัน หากเพิ่มจำนวนผู้ประเมิน จะทำวิธีตรวจให้คะแนนรูปกรอกสองชั้นมีค่าความเที่ยงของแบบสอบถามเช่นกัน ทั้งนี้ความเที่ยงของแบบสอบถามที่ใช้วิธีตรวจให้คะแนนรูปกรอกสองชั้นยังคงมีค่าความเที่ยงสูงสุดเมื่อใช้จำนวนผู้ประเมินและจำนวนเหตุการณ์เท่ากัน โดยเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox

เพื่อแสดงให้เห็นการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมพัทธ์ (relative coefficient) ของแบบสอบถามระดับมัธยมศึกษาการคิดวิเคราะห์ การคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปกรอกสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ 2 องค์ประกอบ คือ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 1 - 2 คน และจำนวนเหตุการณ์ 2 - 6 เหตุการณ์ ตามรายละเอียดข้างต้น ผู้วิจัยนำเสนอการเปรียบเทียบเป็นกราฟเส้นตรง ดังภาพที่ 12 และภาพที่ 13

ภาพที่ 12 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมพัทธ์ (relative coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปกรงสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนน วิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 1 คน และจำนวนเหตุการณ์ 2 – 6 เหตุการณ์

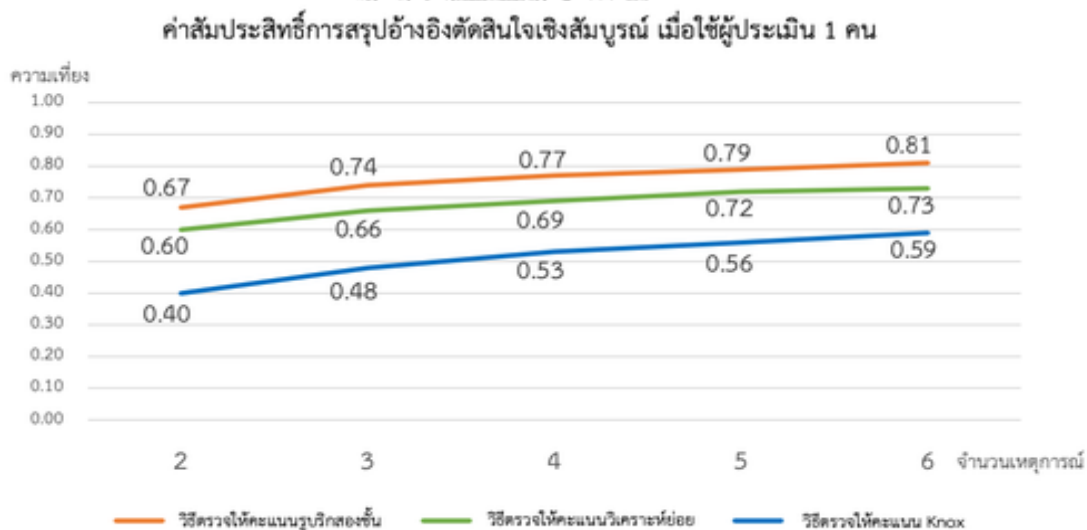


ภาพที่ 13 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมพัทธ์ (relative Coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปกรงสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนน วิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 2 คน และจำนวนเหตุการณ์ 2 – 6 เหตุการณ์

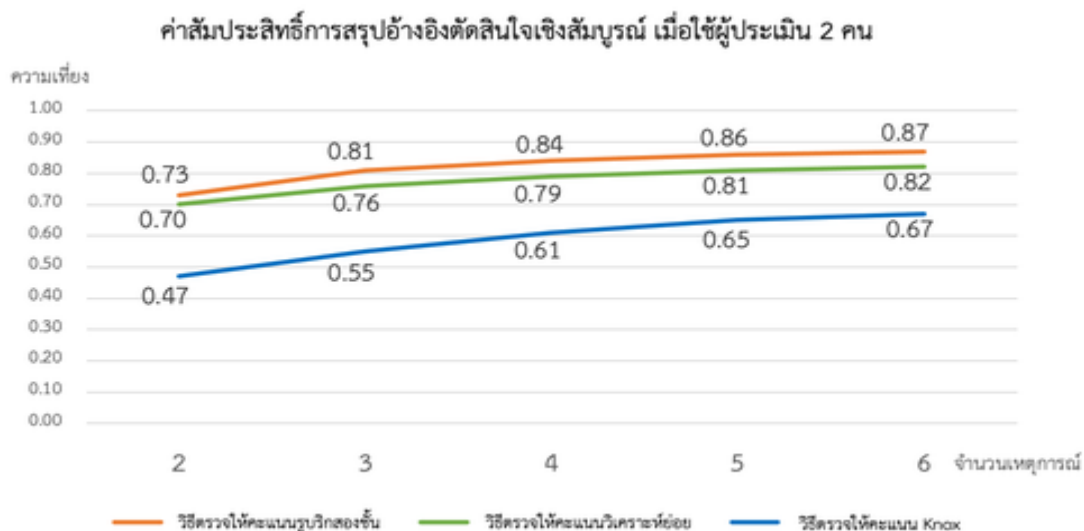


เพื่อแสดงให้เห็นการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์ (absolute Coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปริกสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ 2 องค์ประกอบ คือ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 1 - 2 คน และจำนวนเหตุการณ์ 2 - 6 เหตุการณ์ ตามรายละเอียดข้างต้น ผู้วิจัยนำเสนอการเปรียบเทียบเป็นกราฟเส้นตรง ดังภาพที่ 14 และภาพที่ 15

ภาพที่ 14 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปริกสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนน วิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 1 คน และจำนวนเหตุการณ์ 2 - 6 เหตุการณ์



ภาพที่ 15 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์ (absolute Coefficient) ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีตรวจให้คะแนนรูปrikสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบ $p \times r \times o$ เมื่อใช้จำนวนผู้ประเมิน 2 คน และจำนวนเหตุการณ์ 2 – 6 เหตุการณ์



3.3 ผลการเปรียบเทียบค่าความเที่ยงด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนรูปrikสองชั้น โดยเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ด้วยวิธีการของพิตแมน (Pittman)

จากการตรวจสอบคุณภาพด้านความเที่ยงด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับแบบสอบอิงกลุ่มที่มีความเที่ยงตั้งแต่ 0.7 ขึ้นไปผ่านการพิจารณาค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมพัทธ์ (relative coefficient) พบว่า สำหรับวิธีตรวจให้คะแนนรูปrikสองชั้น ภายใต้ผู้ประเมิน 1 คน ต้องใช้แบบสอบจำนวน 2 เหตุการณ์ ค่าความเที่ยงของแบบสอบ คือ 0.73 และภายใต้ผู้ประเมิน 2 คน ต้องใช้แบบสอบจำนวน 2 เหตุการณ์ ค่าความเที่ยงของแบบสอบ คือ 0.81 ดังตาราง 23 สำหรับวิธีตรวจให้คะแนนวิเคราะห์ย่อย ภายใต้ผู้ประเมิน 1 คน ต้องใช้แบบสอบจำนวน 3 เหตุการณ์ ค่าความเที่ยงของแบบสอบ คือ 0.70 และภายใต้ผู้ประเมิน 2 คน ต้องใช้แบบสอบจำนวน 2 เหตุการณ์ ค่าความเที่ยงของแบบสอบ คือ 0.77 ดังตาราง 4.10 และสำหรับวิธีตรวจให้คะแนน Knox ภายใต้ผู้ประเมิน 1 คน ต้องใช้แบบสอบจำนวน 6 เหตุการณ์ ค่าความเที่ยงของแบบสอบ คือ 0.71 และภายใต้ผู้ประเมิน 2 คน ต้องใช้แบบสอบจำนวน 5 เหตุการณ์ ค่าความ

เที่ยงของแบบสอบ คือ 0.72 ดังตาราง 25 ดังนั้นสรุปได้ว่า สำหรับทุกวิธีตรวจให้คะแนนแบบสอบ
 อัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่มีความเที่ยงตั้งแต่ 0.7 ขึ้นไป เมื่อ
 พิจารณาภายใต้ผู้ประเมิน 1 คน จะใช้แบบสอบ 6 เหตุการณ์ พบว่า วิธีตรวจให้คะแนนรูปrikสอง
 ชั้น ให้ค่าความเที่ยงสูงสุด คือ 0.84 รองลงมา คือ วิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้
 คะแนน Knox ให้ค่าความเที่ยง คือ 0.77 และ 0.71 ตามลำดับ ดังตาราง 23 ถึงตาราง 25 เมื่อ
 พิจารณาภายใต้ผู้ประเมิน 2 คน จะใช้แบบสอบ 5 เหตุการณ์ พบว่า วิธีตรวจให้คะแนนรูปrikสอง
 ชั้น ให้ค่าความเที่ยงสูงสุด คือ 0.89 รองลงมา คือ วิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้
 คะแนน Knox ให้ค่าความเที่ยง คือ 0.85 และ 0.72 ตามลำดับ ดังตาราง 23 ถึงตาราง 25 จากนั้น
 นำความเที่ยงของแบบสอบ มาเปรียบเทียบความแตกต่างรายคู่ ผ่านการทดสอบด้วยสถิติที่
 (t-test) ด้วยวิธีของพิตแมน (Pittman)

สำหรับผลการเปรียบเทียบความแตกต่างของความเที่ยงด้วยค่าสัมประสิทธิ์การสรูป
 อัจฉิงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่ใช้วิธีตรวจให้
 คะแนนรูปrikสองชั้น โดยเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox
 พบว่า เมื่อใช้ผู้ประเมิน 1 คน ภายใต้แบบสอบ 6 เหตุการณ์ วิธีตรวจให้คะแนนรูปrikสองชั้นมี
 ความเที่ยงสูงกว่าวิธีตรวจให้คะแนนวิเคราะห์ย่อยอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.1 ($t =$
 $2.080, p = 0.0404$) และวิธีตรวจให้คะแนนรูปrikสองชั้นมีความเที่ยงสูงกว่าวิธีตรวจให้คะแนน
 Knox อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.1 ($t = 3.063, p = 0.0029$) เมื่อใช้ผู้ประเมิน 2 คน
 ภายใต้ แบบสอบ 5 เหตุการณ์ วิธีตรวจให้คะแนนรูปrikสองชั้นมีความเที่ยงสูงกว่าวิธีตรวจให้
 คะแนนวิเคราะห์ย่อยอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.1 ($t = 1.776, p = 0.0792$) และวิธีตรวจ
 ให้คะแนนรูปrikสองชั้นมีความเที่ยงสูงกว่าวิธีตรวจให้คะแนน Knox อย่างมีนัยสำคัญที่ระดับ 0.1
 ($t = 4.830, p < .00001$) ดังตาราง 26

ตาราง 26 ผลการเปรียบเทียบความแตกต่างของความเที่ยงของของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนนรูปริกสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้ผู้ประเมิน 1 และ 2 คน

จำนวนผู้ประเมิน	การเปรียบเทียบ	$r_{x_1x_2}$	t	p
1 คน	วิธีรูปริกสองชั้น - วิธีวิเคราะห์	0.568	2.080**	0.0404
	วิธีรูปริกสองชั้น - วิธี Knox			
2 คน	วิธีรูปริกสองชั้น - วิธีวิเคราะห์	0.568	1.776**	0.0792
	วิธีรูปริกสองชั้น - วิธี Knox			

** p <.1

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาแบบสอบถามวัดทักษะการคิด
วิจารณ์ทางคณิตศาสตร์และตรวจสอบคุณภาพวิธีการตรวจให้คะแนนแบบรูปrikสองชั้นใน
แบบสอบถามวัดทักษะการคิด โดยเปรียบเทียบความเที่ยงของแบบสอบถามวัดทักษะการคิด
วิจารณ์ทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนรูปrikสองชั้น โดยเทียบกับวิธีตรวจให้
คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox แบบสอบถามวัดทักษะการคิดที่มี
คะแนนในแต่ละวิธีนั้นศึกษาอยู่ภายใต้จำนวนเหตุการณ์ที่ต่างกัน โดยการประยุกต์ใช้ทฤษฎีการ
สรุปอ้างอิง

ประชากรที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ที่ศึกษาอยู่ใน
โรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 9 สุพรรณบุรี จำนวน 33,751 คน

ตัวอย่างวิจัยที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปี
ที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2564 ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา
สุพรรณบุรี จำนวน 90 คน ขนาดตัวอย่างกำหนดด้วย WebPower package ในโปรแกรม R
สำหรับการทดสอบ Multi-way ANOVA (3-way ANOVA) ระหว่างจำนวนพาเซท (ผู้สอบ,
ผู้ประเมิน และเหตุการณ์) และวิธีการตรวจให้คะแนน (วิธีตรวจให้คะแนนรูปrikสองชั้น, วิธีตรวจให้
คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox) ด้วยสถิติที่ใช้ทดสอบ คือ F-test กำหนด
ขนาดอิทธิพลขนาดใหญ่ เท่ากับ 0.40 ความคลาดเคลื่อนร้อยละ 5 ระหว่างจะได้ ขนาดตัวอย่าง
80 คน แต่ป้องกันการขาดหายของตัวอย่าง ในขณะที่เก็บรวบรวมข้อมูล ผู้วิจัยจึงกำหนดขนาด
ตัวอย่างวิจัยเป็น 90 คน โดยมีเกณฑ์การพิจารณาโรงเรียนของตัวอย่างวิจัยจากการสุ่มตัวอย่าง
แบบหลายขั้นตอน (multi-stage random sampling) ของขนาดโรงเรียนขนาดเล็ก โรงเรียนขนาด
กลาง และโรงเรียนขนาดใหญ่ จำนวน 3 โรงเรียน จากจำนวนโรงเรียนทั้งหมด 32 โรงเรียน เพื่อได้
ตัวอย่างวิจัยที่ครอบคลุมทุกตัวทุกขนาดของโรงเรียน

เครื่องมือที่ใช้ในการวิจัย คือ 1) แบบสอบถามวัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์
ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีจำนวน 3 เหตุการณ์ โดยแต่ละเหตุการณ์
มีข้อความถาม จำนวน 5 ข้อ การสร้างข้อความจะสอดคล้องกับขั้นตอนการคิดวิจารณ์ทาง
คณิตศาสตร์ โดยในข้อความที่ 1 ของทุกเหตุการณ์ จะวัดความสามารถในการนิยามปัญหา ข้อ
คำถามที่ 2 ของทุกเหตุการณ์จะวัดความสามารถในการตั้งสมมติฐาน ข้อคำถามที่ 3 ของทุก

เหตุการณ์จะวัดความสามารถในการรวบรวมข้อมูล ข้อคำถามที่ 4 ของทุกเหตุการณ์จะวัดความสามารถในการวิเคราะห์ข้อมูล และข้อคำถามที่ 5 ของทุกเหตุการณ์จะวัดความสามารถในการสร้างข้อสรุป 2) วิธีการตรวจให้คะแนน งานวิจัยในครั้งนี้ใช้วิธีการตรวจให้คะแนน 3 วิธี คือ (1) วิธีการให้คะแนนแบบรูปกรีกสองชั้น ในชั้นที่ 1 มีเกณฑ์การให้คะแนน ดังนี้ ข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 ข้อละ 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (4 คะแนน) และการเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) จากนั้นแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้คะแนนเต็มข้อละ 2 คะแนน ส่วนข้อคำถามที่ 4 และข้อคำถามที่ 5 ข้อละ 3 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (4 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) และการสื่อความหมายทางคณิตศาสตร์ (4 คะแนน) จากนั้นแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้คะแนนเต็ม ข้อละ 3 คะแนน (2) วิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อย มีเกณฑ์การให้คะแนน ดังนี้ ข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 ข้อละ 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) ส่วนในข้อคำถามที่ 4 และข้อคำถามที่ 5 ข้อละ 3 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) และการสื่อความหมายทางคณิตศาสตร์ (2 คะแนน) และ (3) วิธีการตรวจให้คะแนนตามแบบ Knox มีเกณฑ์การให้คะแนนแต่ละข้อคำถามเป็น 3 ระดับ ได้แก่ 2 คะแนน 1 คะแนน และ 0 คะแนน

การวิเคราะห์ข้อมูลโดยวิเคราะห์ค่าสถิติพื้นฐานของผลการตรวจให้คะแนนแบบรูปกรีกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ โดยนำคะแนนที่ได้จากการตรวจให้คะแนนแบบสอบนี้ จำนวน 3 เหตุการณ์ จากผู้ประเมิน คือ ผู้ที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน มาวิเคราะห์ค่าสถิติพื้นฐานโดยใช้สถิติพื้นฐาน คือ ค่าเฉลี่ย (M) ส่วนเบี่ยงเบนมาตรฐาน (SD) คะแนนสูงสุด (Maximum) คะแนนต่ำสุด (Minimum) และวิเคราะห์คุณภาพของวิธีการตรวจให้คะแนนแบบรูปกรีกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ จากผู้ประเมิน คือ ผู้ที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน ที่ทำการตรวจให้คะแนนแบบสอบทุกฉบับ และตรวจแบบสอบ ทุกฉบับจำนวน 2 ครั้ง ห่างกัน 4 สัปดาห์ มาวิเคราะห์หาความเที่ยงในการตรวจให้คะแนนภายในผู้ประเมินคนเดียวด้วยสถิติค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation) และความเที่ยงในการตรวจให้คะแนนระหว่างผู้ประเมิน 2 คน ด้วยสถิติค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment

correlation) โดยพิจารณาคู่กับค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) ที่วิเคราะห์จากโมเดลอิทธิพลผสมแบบสองทาง (two-way mixed-effects model) และวิเคราะห์ความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ ที่มีการออกแบบวิธีการวัด 2 ฟาเซท คือ จำนวนผู้ประเมินและจำนวนเหตุการณ์ที่ต่างกัน โดยใช้ทฤษฎีการสรุปอ้างอิง โดยนำคะแนนที่ได้จากการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์มาวิเคราะห์ค่าความเที่ยง ภายใต้จำนวนผู้ประเมิน 2 คน และจำนวนเหตุการณ์ที่ใช้ต่างกัน คือ 2, 3, 4, 5 และ 6 เหตุการณ์ เหตุการณ์ละ 5 ข้อคำถาม มาหาผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่จำนวนผู้ประเมินและจำนวนเหตุการณ์ที่ต่างกัน ภายใต้การออกแบบการวัด 2 ฟาเซท คือ $p \times r \times o$ เมื่อ p แทน ผู้สอบ r แทน ผู้ประเมิน และ o แทน เหตุการณ์ นั่นคือ ผู้ประเมินตรวจให้คะแนนแบบสอบทุกเหตุการณ์ของผู้สอบทุกคน โดยใช้โปรแกรม EduG ในการวิเคราะห์ค่าความเที่ยง และเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนต่างกัน โดยมีจำนวนเหตุการณ์เท่ากัน คือ 3 เหตุการณ์ ภายใต้ผู้ประเมินจำนวน 2 คน ด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมพัทธ์ (relative coefficient)

สรุปผลการวิจัย

การสรุปผลการวิจัยแบ่งออกเป็น 3 ตอน เพื่อตอบวัตถุประสงค์การวิจัย คือ (1) ผลการพัฒนาแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์และวิธีตรวจให้คะแนนรูปริกสองชั้น สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 (2) ผลการตรวจสอบคุณภาพของวิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ (3) ผลการเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง แต่ละตอนมีรายละเอียดดังนี้

1. ผลการพัฒนาแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทาง คณิตศาสตร์และวิธีตรวจให้คะแนนแบบรูปริกสองชั้น สำหรับนักเรียนชั้นมัธยมศึกษา ปีที่ 3

แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่พัฒนาขึ้น มีจำนวน 3 เหตุการณ์ วัดตามสาระทางคณิตศาสตร์ จำนวน 2 สาระ ได้แก่ สาระจำนวนและพีชคณิต 1 เรื่อง คือ อัตราส่วน สาระการวัดและเรขาคณิต 2 เรื่อง คือ ทฤษฎีบทพีทาโกรัส และพื้นที่ผิวและปริมาตร โดยแต่ละเหตุการณ์จะวัดตามขั้นตอนของกระบวนการคิดวิจารณ์ทางคณิตศาสตร์ นั่นคือ ข้อคำถามที่ 1 ของทุกเหตุการณ์จะวัดความสามารถในการนิยามปัญหา ข้อคำถามที่ 2 วัดความสามารถในการตั้งสมมติฐาน ข้อคำถามที่ 3 วัดความสามารถในการรวบรวมข้อมูล ข้อคำถามที่ 4 วัดความสามารถในการวิเคราะห์ข้อมูล และข้อคำถามที่ 5 วัดความสามารถในการสร้างข้อสรุป รวมเหตุการณ์ละ 5 ข้อคำถาม รวมทั้งสิ้น ฉบับละ 15 ข้อคำถาม สำหรับวิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่พัฒนาขึ้น มีเกณฑ์การให้คะแนนที่สอดคล้องกับแต่ละข้อคำถามที่วัดตามกระบวนการคิดวิจารณ์ 5 ขั้นตอน นั่นคือ ข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 สำหรับชั้นที่ 1 วัดข้อละ 2 องค์กรประกอบ ดังนี้ องค์กรประกอบที่ 1 ความถูกต้องของเนื้อหา (4 คะแนน) ซึ่งแบ่งออกเป็น 2 ข้อรายการย่อย คือ ระบุข้อมูลที่สถานการณ์กำหนด (2 คะแนน) และอธิบายเงื่อนไขของสถานการณ์ (2 คะแนน) และองค์กรประกอบที่ 2 การเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) ซึ่งแบ่งออกเป็น 2 ข้อรายการย่อย คือ กำหนดตามขั้นของการวัดตามกระบวนการคิดวิจารณ์ทางคณิตศาสตร์ (กำหนดปัญหา/ ตั้งสมมติฐาน/ การรวบรวมข้อมูล) (2 คะแนน) และกำหนดบริบทของข้อมูลภายใต้สถานการณ์ที่กำหนด (2 คะแนน) จากนั้นแปลงคะแนนที่ได้จากชั้นที่ 1 เป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้ คือ ก่อนแปลงคะแนน มีคะแนนอยู่ในช่วง 7-8 คะแนน หลังแปลงคะแนนในชั้นที่ 2 จะได้คะแนน 2 คะแนน, ก่อนแปลงคะแนน มีคะแนนอยู่ในช่วง 4-6 คะแนน หลังแปลงคะแนนในชั้นที่ 2 จะได้คะแนน 1 คะแนน และก่อนแปลงคะแนน มีคะแนนอยู่ในช่วง 0-3 คะแนน หลังแปลงคะแนนในชั้นที่ 2 จะได้คะแนน 0 คะแนน ส่วนข้อคำถามที่ 4 และข้อคำถามที่ 5 สำหรับชั้นที่ 1 วัดข้อละ 3 องค์กรประกอบ ดังนี้ องค์กรประกอบที่ 1 ความถูกต้องของเนื้อหา (4 คะแนน) องค์กรประกอบที่ 2 การเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) สอดคล้องเช่นเดียวกับข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 และองค์กรประกอบที่ 3 การสื่อความหมายทางคณิตศาสตร์ (4 คะแนน) ซึ่งแบ่งออกเป็น 2 ข้อรายการย่อย คือ การเปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ และการคำนวณหาคำตอบ จากนั้นแปลงคะแนนที่ได้จากชั้นที่ 1 เป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้ คือ ก่อนแปลงคะแนน มีคะแนนอยู่ในช่วง 10-12

คะแนน หลังแปลงคะแนนในชั้นที่ 2 จะได้คะแนน 3 คะแนน, ก่อนแปลงคะแนน มีคะแนนอยู่ในช่วง 7-9 คะแนน หลังแปลงคะแนนในชั้นที่ 2 จะได้คะแนน 2 คะแนน, ก่อนแปลงคะแนน มีคะแนนอยู่ในช่วง 4-6 คะแนน หลังแปลงคะแนนในชั้นที่ 2 จะได้คะแนน 1 คะแนน และก่อนแปลงคะแนน มีคะแนนอยู่ในช่วง 0-3 คะแนน หลังแปลงคะแนนในชั้นที่ 2 จะได้คะแนน 0 คะแนน ดังนั้นแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ตรวจให้คะแนน ด้วยวิธีแบบรูปริกสองชั้น แต่ละเหตุการณ์จะมีคะแนนเต็ม 12 คะแนน รวมทั้งนับคิดเป็น 36 คะแนน และจากการสอบ วัดกับกลุ่มทดลองใช้ จำนวน 30 คน พบว่า แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่พัฒนาขึ้น ทั้งฉบับมีค่าความเที่ยงสูง (0.934) แบบสอบมีความยากง่ายอยู่ในระดับระดับปานกลางค่อนข้างไปทางยาก (0.46) และแบบสอบ มีความสามารถในการจำแนกผู้สอบได้ (0.37)

2. ผลการตรวจสอบคุณภาพของวิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

2.1 ผลการตรวจสอบความตรงเชิงเนื้อหาของวิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ผู้วิจัยพัฒนาขึ้น โดยนำไปให้ผู้ทรงคุณวุฒิ จำนวน 3 คน เพื่อตรวจสอบความสอดคล้องของมาตรฐานการเรียนรู้ ตัวชี้วัด และสาระการเรียนรู้ทางคณิตศาสตร์ ความสอดคล้องของแนวคำตอบกับข้อคำถาม และความสอดคล้องของเกณฑ์การให้คะแนนกับแนวคำตอบ พบว่า วิธีการตรวจให้คะแนนแบบรูปริกสองชั้นที่ผู้วิจัยพัฒนาขึ้นเพื่อใช้ตรวจให้คะแนนข้อคำถามของเหตุการณ์ในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ ทุกข้อมีระดับความสอดคล้องกับข้อคำถามและแนวคำตอบ รวมถึงมีความสอดคล้องกับเนื้อหา และขั้นของกระบวนการคิดวิจารณ์ญาณ และเมื่อพิจารณาคำแนะนำที่ได้จากผู้เชี่ยวชาญได้ข้อมูลโดยสรุปว่า ในทุกข้อรายการย่อยได้รับคำแนะนำให้ปรับแก้คำว่า “ไม่ถูกต้องทั้งหมด” จากเกณฑ์การให้ 0 คะแนน เป็น “ไม่ถูกต้อง” เนื่องจากไปซ้ำซ้อนกับเกณฑ์การให้ 1 คะแนน (ถูกต้องบางส่วน)

2.2 ผลการวิเคราะห์ค่าสถิติเบื้องต้นของคะแนนที่ใช้วิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ ที่ได้จากผู้ตรวจให้คะแนนจำนวน 2 คน โดยทำการตรวจให้คะแนน 2 ครั้ง ห่างกัน 4 สัปดาห์ โดยภาพรวมพบว่า ค่าเฉลี่ยของคะแนนในแต่ละเหตุการณ์และค่าเฉลี่ยรวมทั้งฉบับ จากผู้ประเมินคนที่ 1 และผู้ประเมินคนที่ 2 สูงกว่าร้อยละ 50 ของคะแนนเต็ม โดยค่าเฉลี่ยในแต่ละเหตุการณ์ของการตรวจ

ครั้งที่ 1 และครั้งที่ 2 ของผู้ประเมินคนที่ 1 พบว่า มีค่าใกล้เคียงกัน และคะแนนที่ได้จากการตรวจครั้งที่ 2 จะสูงกว่าคะแนนที่ได้จากการตรวจครั้งที่ 1 ส่วนค่าเฉลี่ยในแต่ละเหตุการณ์ของการตรวจครั้งที่ 1 และครั้งที่ 2 ของผู้ประเมินคนที่ 2 พบว่า มีค่าใกล้เคียงกัน โดยคะแนนที่ได้จากการตรวจครั้งที่ 1 จะสูงกว่าคะแนนที่ได้จากการตรวจครั้งที่ 2 จากข้อมูลดังกล่าวสรุปได้ว่า ไม่ว่าจะผู้ประเมินจะตรวจแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ด้วยของวิธีการตรวจให้คะแนนแบบรูปริกสองชั้นก็ครั้งที่ 1 ก็จะได้คะแนนใกล้เคียงกัน

2.3 ความเที่ยงในการตรวจให้คะแนนจากผู้ประเมินคนเดียว จากผู้ที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน ที่ทำการตรวจให้คะแนนแบบสอบ 90 ฉบับ และตรวจแบบสอบทุกฉบับจำนวน 2 ครั้ง ห่างกัน 4 สัปดาห์ ด้วยวิธีการตรวจให้คะแนนแบบรูปริกสองชั้น พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนมีนัยสำคัญทางสถิติที่ระดับ 0.1 ทุกเหตุการณ์ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์อยู่ในระดับสูง และเมื่อพิจารณาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานของผู้ประเมินที่ทำการตรวจให้คะแนนทั้ง 2 ครั้ง มีค่าใกล้เคียงกันทุกเหตุการณ์ จากข้อมูลดังกล่าวสรุปได้ว่า แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ตรวจให้คะแนนแบบรูปริกสองชั้นที่พัฒนาขึ้น มีความเป็นปรนัยในการตรวจให้คะแนน ไม่ว่าจะผู้ประเมินจะตรวจให้คะแนนทั้งหมดกี่ครั้ง คะแนนที่ได้ก็จะมีค่าใกล้เคียงกัน

2.4 ความเที่ยงในการตรวจให้คะแนนจากผู้ประเมิน 2 คน จากผู้ที่ปฏิบัติการสอนคณิตศาสตร์ที่ประสบการณ์สอนมากกว่า 10 ปี หรือมีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน ที่ทำการตรวจให้คะแนนแบบสอบ 90 ฉบับอย่างเป็นอิสระ พบว่า พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนอยู่ในระดับสูงและมีนัยสำคัญทางสถิติที่ระดับ 0.1 ซึ่งสอดคล้องกับค่าสัมประสิทธิ์ภายในชั้นจากโมเดลอิทธิพลผสมแบบสองทางที่ระดับความเชื่อมั่น 95% ทุกเหตุการณ์ แสดงว่า ทุกเหตุการณ์มีความสอดคล้องในการตรวจให้คะแนนระหว่างผู้ประเมิน โดยมีค่าสัมประสิทธิ์สหสัมพันธ์อยู่ในระดับสูง และเมื่อพิจารณาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานของ ผู้ประเมินทั้ง 2 คน มีค่าใกล้เคียงกันทุกเหตุการณ์ จากข้อมูลดังกล่าวสรุปได้ว่า แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนนรูปริกสองชั้นมีความเป็นปรนัยในการตรวจให้คะแนน ไม่ว่าจะให้ผู้ประเมินคนใดตรวจให้คะแนนคะแนนที่ได้ก็จะมีค่าใกล้เคียงกัน

3. ผลการเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีการตรวจให้ คะแนนที่แตกต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง

3.1 ผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยใช้วิธีการตรวจให้คะแนน 3 วิธี คือ วิธีตรวจให้คะแนนรูปริกสองชั้น วิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox ที่มีวิธีการออกแบบ 2 ฟาเซท ได้แก่ จำนวนผู้ประเมินและจำนวนเหตุการณ์ที่ต่างกัน พบว่า ทุกวิธีตรวจให้คะแนนความแปรปรวนของจำนวนเหตุการณ์มีผลต่อสัมประสิทธิ์การสรุปอ้างอิงมากที่สุด แสดงว่า แบบสอบอัตนัยประยุกต์ควรมีจำนวนเหตุการณ์ให้เหมาะสมกับความรู้ของผู้สอบและสามารถใช้ได้จริงในการจัดสอบ เนื่องจากจำนวนข้อสอบอัตนัยที่เยอะเกินไป ทำให้ผู้สอบเกิดความเมื่อยล้า และไม่สามารถวัดความรู้ได้อย่างแท้จริง สอดคล้องกับสุโรยา หมั่นหมัด (2549) ที่กล่าวไว้ว่า แบบสอบอัตนัยควรมีข้อคำถาม 5-35 ข้อ จึงจะสามารถวัดพฤติกรรมที่ต้องการวัดได้เพียงพอ

3.2 ผลการศึกษาเพื่อตัดสินใจการสรุปอ้างอิง (D-Study) จำนวนเหตุการณ์ที่ใช้ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนรูปริกสองชั้น เมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้การออกแบบให้ผู้ประเมินตรวจให้คะแนนแบบสอบทุกเหตุการณ์ของผู้สอบทุกคน หรือ ($p \times r \times o$) เมื่อใช้ผู้ประเมินจำนวน 1 และ 2 คน ภายใต้จำนวนเหตุการณ์ที่ใช้ 2, 3, 4, 5 และ 6 เหตุการณ์ โดยแต่ละเหตุการณ์มี 5 ข้อคำถาม ในระดับความเที่ยงที่ผู้วิจัยยอมรับตั้งแต่ 0.7 ขึ้นไป ตามเกณฑ์ที่ Nunnally (1978) เสนอไว้ เมื่อพิจารณาจำนวนผู้ประเมินเท่ากัน พบว่า เพิ่มจำนวนเหตุการณ์จะทำให้วิธีตรวจให้คะแนนรูปริกสองชั้นมีค่าความเที่ยงของแบบสอบสูงขึ้น เมื่อพิจารณาจำนวนผู้ประเมินต่างกัน พบว่า เพิ่มจำนวนผู้ประเมินจะทำให้วิธีตรวจให้คะแนนรูปริกสองชั้นมีค่าความเที่ยงของแบบสอบสูงขึ้นเช่นกัน ทั้งนี้ความเที่ยงของแบบสอบที่ใช้วิธีตรวจให้คะแนนรูปริกสองชั้นยังคงมีค่าความเที่ยงสูงสุดเมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้จำนวนผู้ประเมินและจำนวนเหตุการณ์เท่ากัน

3.3 ผลการเปรียบเทียบความแตกต่างของความเที่ยงด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ โดยใช้วิธีการตรวจให้คะแนนที่แตกต่างกัน คือ วิธีตรวจให้คะแนนรูปริกสองชั้น วิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox เมื่อพิจารณาความเที่ยงของแบบสอบตั้งแต่ 0.7 ขึ้นไป สำหรับวิธีตรวจให้คะแนนรูปริกสองชั้น พบว่า เมื่อใช้ผู้ประเมิน 1 คน ควรใช้แบบสอบอัตนัยประยุกต์จำนวน

2 เหตุการณ์ เมื่อใช้ผู้ประเมิน 2 คน ควรใช้แบบสอบถามอัตโนมัติประยุกต์จำนวน 2 เหตุการณ์ สำหรับวิธีตรวจให้คะแนนวิเคราะห์ย่อย พบว่า เมื่อใช้ผู้ประเมิน 1 คน ควรใช้แบบสอบถามอัตโนมัติประยุกต์จำนวน 3 เหตุการณ์ เมื่อใช้ผู้ประเมิน 2 คน ควรใช้แบบสอบถามอัตโนมัติประยุกต์จำนวน 2 เหตุการณ์ และสำหรับวิธีตรวจให้คะแนน Knox พบว่า เมื่อใช้ผู้ประเมิน 1 คน ควรใช้แบบสอบถามอัตโนมัติประยุกต์จำนวน 3 เหตุการณ์ เมื่อใช้ผู้ประเมิน 2 คน ควรใช้แบบสอบถามอัตโนมัติประยุกต์จำนวน 2 เหตุการณ์ จะเห็นได้ว่าแบบสอบถามอัตโนมัติวัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ไม่ว่าจะใช้วิธีตรวจให้คะแนนใด พบว่า เมื่อใช้ผู้ประเมิน 1 คน แบบสอบถามประยุกต์ควรมีจำนวน 6 เหตุการณ์ เมื่อใช้ผู้ประเมิน 2 คน แบบสอบถามประยุกต์ควรมีจำนวน 5 เหตุการณ์ จึงจะได้ความเที่ยงของแบบสอบในระดับที่ยอมรับได้ 0.7 ขึ้นไป โดยที่วิธีตรวจให้คะแนนรูปrikสองชั้นจะให้ค่าความเที่ยงสูงกว่าวิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox ตามลำดับ และเมื่อพิจารณาผลการเปรียบเทียบความแตกต่างของความเที่ยงของแบบสอบถามประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนนรูปrikสองชั้น โดยเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox พบว่า วิธีตรวจให้คะแนนรูปrikสองชั้นมีความเที่ยงสูงกว่าวิธีตรวจให้คะแนนวิเคราะห์ย่อย อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.1 และวิธีตรวจให้คะแนนรูปrikสองชั้นมีความเที่ยงสูงกว่าวิธีตรวจให้คะแนน Knox อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.1

อภิปรายผลการวิจัย

จากผลการวิเคราะห์และสรุปผลการวิจัย ผู้วิจัยนำเสนอประเด็นการอภิปรายที่น่าสนใจ ดังนี้

1. การพัฒนาแบบสอบถามอัตโนมัติวัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์และวิธีตรวจให้คะแนนแบบรูปrikสองชั้น สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

แบบสอบถามอัตโนมัติวัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่พัฒนาขึ้นมีจำนวน 3 เหตุการณ์ วัดตามสาระทางคณิตศาสตร์ จำนวน 2 สาระ ได้แก่ สาระจำนวนและพีชคณิต 1 เรื่อง คือ อัตราส่วน สาระการวัดและเรขาคณิต 2 เรื่อง คือ ทฤษฎีบทพีทาโกรัส และพื้นที่ผิวและปริมาตร โดยแต่ละเหตุการณ์จะวัดตามขั้นตอนของกระบวนการคิดวิจารณ์ทางคณิตศาสตร์ซึ่งเป็นการคิดขั้นสูงที่บูรณาการความรู้ทางคณิตศาสตร์ สอดคล้องกับแนวคิดของ Dressel and Mayhew (1957) โดยผู้วิจัยสังเคราะห์จากแนวคิดและงานวิจัยที่เกี่ยวข้องกับการคิด

วิจารณ์ญาณทางคณิตศาสตร์ ดังนี้ ข้อคำถามที่ 1 ของทุกเหตุการณ์จะวัดความสามารถในการนิยามปัญหา ข้อคำถามที่ 2 วัดความสามารถในการตั้งสมมติฐาน ข้อคำถามที่ 3 วัดความสามารถในการรวบรวมข้อมูล ข้อคำถามที่ 4 วัดความสามารถในการวิเคราะห์ข้อมูล และข้อคำถามที่ 5 วัดความสามารถในการสร้างข้อสรุปรวมเหตุการณ์ละ 5 ข้อคำถาม รวมทั้งสิ้น ฉบับละ 15 ข้อคำถาม ทั้งนี้การสร้างเหตุการณ์ที่วัดการคิดวิจารณ์ญาณทางคณิตศาสตร์ไม่ง่าย ต้องมีบริบทที่ใกล้เคียงกับชีวิตจริงของผู้สอบ เพื่อลดความได้เปรียบเสียเปรียบ และองค์ประกอบด้านเหตุการณ์เป็นตัวแปรหนึ่งในพาเซทที่ผู้วิจัยพิจารณาเพื่อหาข้อสรุปของจำนวนเหตุการณ์ที่เหมาะสมต่อการนำไปสร้างในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ แต่เนื้อหาหรือเรื่องที่ต้องการวัดในแต่ละเหตุการณ์ที่ผู้วิจัยสร้างขึ้นวัดต่างเรื่องต่างสาระกัน อาจทำให้เกิดความไม่แน่ชัดของข้อสรุปของจำนวนเหตุการณ์ที่ใช้วัดว่าวัดในเนื้อหาเรื่องใด หรือสามารถวัดได้ทุกเนื้อหาสาระทางคณิตศาสตร์

สำหรับวิธีตรวจให้คะแนนรูปริกสองชั้นเป็นวิธีการประเมินที่มีรูปแบบการกำหนดคำอธิบายที่ชัดเจนซึ่งสอดคล้องกับงานวิจัยของณัฐธณี ศิริโชติ และกมลวรรณ ดงธนากานนท์ (2559) โดยมีเกณฑ์การประเมิน 2 ชั้น ดังนี้ ระดับชั้นที่ 1 เป็นเกณฑ์การให้คะแนนตามคำอธิบายขององค์ประกอบย่อยลดหลั่นตามระดับความสามารถ ส่วนระดับชั้นที่ 2 เป็นการแปลงผลรวมของคะแนนที่ได้จากระดับชั้นที่ 1 ให้ออกมาเป็นช่วงคะแนนที่กำหนด ในงานวิจัยที่ผ่านมายังไม่มีรายละเอียดหรือการกำหนดเกณฑ์ที่แน่ชัดว่าควรกำหนดเกณฑ์การให้คะแนนหลังแปลงคะแนนไว้ในช่วงใดสำหรับชั้นที่ 2 ของวิธีตรวจให้คะแนนรูปริกสองชั้น สำหรับงานวิจัยฉบับนี้ผู้วิจัยจึงพิจารณาจากคะแนนที่ได้ในชั้นที่ 1 โดยแปลงคะแนนในชั้นที่ 2 ให้มีช่วงคะแนนที่สอดคล้องกับเกณฑ์การตรวจให้คะแนนในวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ดังนี้ สำหรับชั้นที่ 2 ของข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 กำหนดเกณฑ์การให้คะแนนหลังแปลงคะแนนในชั้นที่ 2 เป็น 3 ระดับ (2, 1, 0) ส่วนข้อคำถามที่ 4 และข้อคำถามที่ 5 กำหนดเกณฑ์การให้คะแนนหลังแปลงคะแนนในชั้นที่ 2 เป็น 4 ระดับ (3, 2, 1, 0) ทั้งนี้แม้ว่าหลังการแปลงคะแนนในชั้นที่ 2 ของวิธีตรวจให้คะแนนรูปริกสองชั้น คะแนนที่ได้จะมีความละเอียดน้อยลง แต่สารสนเทศที่ได้จากชั้นที่ 1 จะบ่งชี้ถึงความสามารถของผู้สอบได้ว่าควรพัฒนาหรือปรับปรุงในเรื่องใดผ่านการพิจารณาคะแนนที่ได้จากคำอธิบายแต่ละข้อรายการย่อย

เมื่อพิจารณาค่าความเที่ยงที่ได้จากแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์โดยใช้วิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในกลุ่มทดลองใช้พบว่า ค่าความเที่ยงที่ได้มีค่าสูง ซึ่งสอดคล้องกับ Hamzah M. S. G. B. et al. (2014) ที่กล่าวไว้

ว่า วิธีการตรวจให้คะแนนแบบรูปริกสองชั้นมีความเป็นปรนัยมากกว่าวิธีการตรวจให้คะแนนวิเคราะห์ย่อยซึ่งเป็นวิธีการตรวจให้คะแนนที่ละเอียดสำหรับแบบสอบอัตนัยประยุกต์และแบบสอบความเรียงในงานวิจัยที่ผ่านมา เนื่องด้วยชั้นที่ 1 ของวิธีการตรวจให้คะแนนรูปริกสองชั้น จะแยกคำอธิบายตามข้อรายการย่อย ในแต่ละระดับคะแนนที่ได้ (2, 1, 0) จะสามารถแยกผู้สอบที่มีความสามารถไม่เท่ากันได้จากคำอธิบายในแต่ละระดับความสามารถ 2 คะแนน หมายถึง ทำได้ถูกต้องทั้งหมดตามคำอธิบายตามข้อรายการย่อย 1 คะแนน หมายถึง ทำถูกต้องบางส่วนตามคำอธิบายตามข้อรายการย่อย และ 0 คะแนน หมายถึง ทำไม่ถูกต้องทั้งหมดตามคำอธิบายตามข้อรายการย่อย หรือไม่มีการตอบคำถาม เกณฑ์การให้คะแนนที่ชัดเจนของวิธีการตรวจให้คะแนนรูปริกสองชั้น ทำให้ง่ายต่อการบันทึกคะแนนตามคำอธิบายในองค์ประกอบย่อยและง่ายต่อการอ่านผลการประเมินซึ่งคะแนนที่ได้จากการตรวจให้คะแนนรูปริกสองชั้น นอกจากจะได้ผลคะแนนที่สามารถตัดสินผลสัมฤทธิ์ของผู้เรียนได้แล้ว ยังสะท้อนจุดแข็งหรือจุดอ่อนในแต่ละองค์ประกอบของเนื้อหาที่เรียนซึ่งสอดคล้องกับ Abdullah N. et al. (2015) ที่กล่าวไว้ว่า ผลการประเมินที่ได้จากการตรวจให้คะแนนแบบรูปริกสองชั้น นอกจากจะเป็นค่าเฉลี่ยของการประเมินในด้านนั้น ๆ แล้วยังทำให้ทราบข้อมูลย้อนกลับ เพื่อนำไปแก้ไขปรับปรุงได้ตรงประเด็น และพัฒนาต่อยอดในด้านนั้น ๆ ได้ดียิ่งขึ้น ในขณะที่วิธีการตรวจให้คะแนนวิเคราะห์ย่อยเกิดจากการนำคำอธิบายตามข้อรายการย่อยในชั้นที่ 1 ของวิธีการตรวจให้คะแนนรูปริกสองชั้น มารวมเป็นคำอธิบายในแต่ละระดับคะแนนที่ได้ (2, 1, 0) ซึ่งจะเกิดปัญหาเมื่อผู้สอบทำได้ไม่ครบทุกองค์ประกอบในคำอธิบายคะแนนในแต่ละระดับ ทำให้ผู้สอบที่มีระดับความสามารถไม่เท่ากันแต่ได้รับคะแนนในช่วงเดียวกัน ซึ่งการตัดสินผลคะแนนจะขึ้นอยู่กับผู้ประเมินเป็นสำคัญแทนที่เกณฑ์การตรวจให้คะแนนจะเป็นตัวตัดสินความสามารถของผู้สอบ ทั้งนี้จะเห็นได้ว่าวิธีการตรวจให้คะแนนรูปริกสองชั้นจะสามารถแยกความรู้ความสามารถที่แท้จริงของผู้สอบได้อย่างชัดเจน

2. การตรวจสอบคุณภาพวิธีการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

ผลการตรวจให้คะแนนแบบรูปริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ พบว่า วิธีการตรวจให้คะแนนแบบรูปริกสองชั้นมีความสอดคล้องของเกณฑ์การให้คะแนนกับแนวคำตอบ เนื่องจากผู้วิจัยกำหนดประเด็นหลักที่ต้องการประเมินในวิธีการตรวจให้คะแนนแบบรูปริกสองชั้น ให้สอดคล้องกับขั้นของการวัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ และสอดคล้องกับมาตรฐานการเรียนรู้ ตัวชี้วัด สาระการเรียนรู้ทางคณิตศาสตร์

เพื่อให้การตรวจให้คะแนนแบบรูบริกสองชั้นใน แบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจัยวรรณกรรมทางคณิตศาสตร์วัดได้ครอบคลุมทุกประเด็นในทุกเหตุการณ์สอดคล้องกับ So J. Y. et al. (2018) ที่พัฒนาวิธีการตรวจให้คะแนนแบบรูบริกในการทดสอบทางคลินิกของนักศึกษา แพทย์ ซึ่งกล่าวไว้ว่า การกำหนดสิ่งที่ต้องการประเมินหรือองค์ประกอบในการสร้างเกณฑ์การให้ คะแนนรูบริก ควรกำหนดจากจุดประสงค์ในการประเมิน เพื่อวัดได้ครอบคลุมที่ทุกประเด็นที่ ต้องการ

ผลการตรวจให้คะแนนแบบรูบริกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจัยวรรณกรรมทางคณิตศาสตร์จากความเที่ยงในการตรวจให้คะแนนจากผู้ประเมินคนเดียว และ ความเที่ยงในการตรวจให้คะแนนจากผู้ประเมิน 2 คน พบว่า แบบสอบอัตนัยประยุกต์วัดทักษะ การคิดวิจัยวรรณกรรมทางคณิตศาสตร์ที่ตรวจให้คะแนนแบบรูบริกสองชั้นที่พัฒนาขึ้น มีความเที่ยง ในการตรวจให้คะแนนจากผู้ประเมินคนเดียว และความเที่ยงในการตรวจให้คะแนนจากผู้ประเมิน 2 คนทั้งรายเหตุการณ์และทั้งฉบับมีความเที่ยงอยู่ในระดับสูง เนื่องจากผู้วิจัยคัดเลือกผู้ประเมินที่ มีความเชี่ยวชาญและประสบการณ์การจัดการเรียนการสอนเกี่ยวกับทักษะการคิดขั้นสูงทาง คณิตศาสตร์ และก่อนการตรวจให้คะแนน ผู้วิจัยได้อธิบายการตรวจให้คะแนนแบบรูบริกสองชั้น ตามคำอธิบายทุกข้อรายการย่อยในแต่ละเหตุการณ์เพื่อสร้างความเข้าใจที่ตรงกัน เนื่องจากผู้ ประเมินยังไม่เคยมีประสบการณ์ในการใช้วิธีการตรวจให้คะแนนแบบรูบริกสองชั้นที่มีเกณฑ์การให้ คะแนนหลายระดับมาก่อนและให้ผู้ประเมินฝึกใช้เกณฑ์การตรวจให้คะแนนแบบรูบริกสองชั้นใน การตรวจให้คะแนนจากตัวอย่างแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจัยวรรณกรรมทาง คณิตศาสตร์ที่ผู้เชี่ยวชาญได้ให้คะแนนไว้ก่อนแล้วเพื่อเป็นมาตรฐานในการตรวจสอบคุณภาพใน การตรวจให้คะแนนของผู้ประเมินและลดความคลาดเคลื่อนในการตรวจให้คะแนน มากที่สุด ซึ่งสอดคล้องกับ Jonsson and Svingby (2017), Rezaei and Lovorn (2010) อ้างถึงใน Emily Saxton et al. (2012) ที่กล่าวไว้ว่า การให้ผู้ประเมินฝึกตรวจให้คะแนนโดยใช้แบบสอบที่ เป็นหลักเทียบ จะเป็นเครื่องมือที่สำคัญในการเพิ่มความเที่ยงของการตรวจให้คะแนน อีกทั้งยัง สอดคล้องกับ Sandlund and Greer (2020) ซึ่งกล่าวไว้ว่า การฝึกตรวจให้คะแนนรูบริกโดยใช้ ตัวอย่างงานเดียวกัน จะช่วยให้ผู้ประเมินส่วนใหญ่ประเมินงานเดียวกัน ได้คะแนนใกล้เคียงกัน และช่วยเพิ่มความเที่ยงระหว่างผู้ประเมิน โดยความเที่ยงที่กล่าวมานั้น มีความเที่ยงในรายฉบับสูง กว่าความเที่ยงในแต่ละเหตุการณ์ ดังนั้นจะเห็นได้ว่า แบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจัยวรรณกรรมทางคณิตศาสตร์มีความปรนัยในการตรวจให้คะแนน ไม่ว่าจะผู้ประเมินจะตรวจให้ คะแนนกี่ครั้ง หรือไม่ว่าจะมีผู้ใดตรวจให้คะแนน คะแนนที่ได้ก็จะมีค่าใกล้เคียงกัน

3. การเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ใช้วิธีการตรวจให้ คะแนนที่แตกต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง

ผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อสัมประสิทธิ์การสรุป
อ้างอิงของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้
คะแนนรูปริกสองชั้น โดยเทียบกับวิธีการตรวจให้คะแนนวิเคราะห์ย่อยและวิธีการตรวจให้คะแนน Knox
ภายใต้การออกแบบ 2 ฟาเซท ได้แก่ จำนวนผู้ประเมินและจำนวนเหตุการณ์ที่ต่างกัน พบว่า ทั้ง
3 วิธีการตรวจให้คะแนนความแปรปรวนของคะแนนเฉลี่ยระหว่างเหตุการณ์มีผลต่อสัมประสิทธิ์
การสรุปอ้างอิงมากที่สุด เมื่อเทียบกับความแปรปรวนของคะแนนเฉลี่ยระหว่างผู้ประเมินที่มีค่า
ค่อนข้างน้อย แสดงว่า แบบสอบอัตนัยประยุกต์ควรมีจำนวนข้อสอบให้เหมาะสมกับความถี่ของผู้สอบ
และสามารถใช้ได้จริงในช่วงเวลาที่จัดการสอบ เนื่องจากจำนวนข้อสอบอัตนัยที่เยอะเกินไป อาจ
ทำให้ผู้สอบเกิดความเมื่อยล้าในการสอบ จึงทำให้ไม่สามารถวัดความรู้ได้อย่างแท้จริง ในทาง
กลับกันถ้าจำนวนข้อสอบที่น้อยเกินไป ก็อาจทำให้วัดได้ไม่ครอบคลุมตัวชี้วัดที่ตั้งไว้ สอดคล้องกับ
สุโรยา หมั่นหมัด (2549) ที่กล่าวไว้ว่า แบบสอบอัตนัยควรมีข้อคำถาม 5-35 ข้อ จึงจะสามารถวัด
พฤติกรรมที่ต้องการวัดได้เพียงพอ คะแนนที่ได้จากการเดาจึงลดน้อยลง อีกทั้งผู้ตรวจแต่ละคนมี
การตรวจให้คะแนนที่สอดคล้องกัน เนื่องจากผู้ตรวจทุกคนได้ใช้เกณฑ์การตรวจที่ผู้วิจัยสร้างขึ้น
ผ่านการกำหนดเกณฑ์ที่ชัดเจน และฝึกอบรมการตรวจให้คะแนน ซึ่งแสดงถึงความเป็นปรนัยของ
เกณฑ์การตรวจให้คะแนน

ผลการศึกษาเพื่อตัดสินใจการสรุปอ้างอิงเป็นการศึกษาเพื่อการตัดสินใจเลือกใช้
แบบสอบสำหรับการวัดผลแบบอิงกลุ่มที่สามารถนำไปปฏิบัติใช้ได้จริงในการสอบและมีความ
เที่ยงในระดับที่ต้องการ ซึ่งในงานวิจัยครั้งนี้ ผู้วิจัยเลือกค่าความเที่ยงตั้งแต่ 0.7 ขึ้นไป ตามเกณฑ์
ที่ Nunnally (1978) เสนอไว้ว่า ผ่านการพิจารณาจากค่าสัมประสิทธิ์การสรุปอ้างอิงตัดสินใจเชิง
สัมพัทธ์ พบว่า เมื่อเพิ่มจำนวนผู้ประเมินหรือเพิ่มจำนวนเหตุการณ์ จะทำให้ค่าความเที่ยง
ของแบบสอบอัตนัยประยุกต์สูงขึ้น ซึ่งสอดคล้องกับงานวิจัย มาฆพันธ์ อำนาคิล และกมลวรรณ
ตั้งธนกานนท์ (2557) และ สุภชิต ผดุงผล และกมลวรรณ ตั้งธนกานนท์ (2559) ทั้งนี้ความเที่ยง
ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนน
รูปริกสองชั้นยังคงมีค่าความเที่ยงสูงสุด เมื่อเทียบกับวิธีการตรวจให้คะแนนวิเคราะห์ย่อยและวิธีการตรวจ
ให้คะแนน Knox ภายใต้จำนวนผู้ประเมินและจำนวนเหตุการณ์เท่ากัน

ผลการเปรียบเทียบความแตกต่างของความเที่ยงด้วยค่าสัมประสิทธิ์การสรุปอ้างอิง
ของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนน

รูปริกสองชั้น โดยเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox สำหรับการวัดผลแบบอิงกลุ่ม พบว่า เมื่อใช้ผู้ประเมิน 1 คน แบบสอบอัตนัยประยุกต์ควรมีจำนวน 6 เหตุการณ์ เมื่อใช้ผู้ประเมิน 2 คน แบบสอบอัตนัยประยุกต์ควรมีจำนวน 5 เหตุการณ์ จึงจะได้แบบสอบที่มีความเที่ยงในระดับที่ยอมรับได้ 0.7 ขึ้นไป โดยที่วิธีตรวจให้คะแนนรูปริกสองชั้นมีความเที่ยงสูงกว่าวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.1 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ว่า แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนนที่ต่างกันในระดับความเที่ยงที่ยอมรับ 0.7 ขึ้นไป วิธีตรวจให้คะแนนรูปริกสองชั้นจะมีความเที่ยงสูงสุดเมื่อเทียบกับวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox ภายใต้จำนวนผู้ประเมินและจำนวนเหตุการณ์ที่ต่างกัน เนื่องจากวิธีตรวจให้คะแนนแบบรูปริกสองชั้น เป็นวิธีประเมินทักษะที่เป็นพฤติกรรมเดี่ยว ๆ หรือลักษณะนั้น ๆ ตามระดับคุณภาพของการปฏิบัติงาน ทำให้ครูผู้ที่ใช้แบบประเมินวิธีตรวจให้คะแนนแบบรูปริกสองชั้นเข้าใจง่ายและสามารถนำผลประเมินมาเป็นข้อมูลที่เป็นประโยชน์ต่อการพัฒนาผู้เรียนได้ดีกว่าวิธีการตรวจให้คะแนนรูปริกทั่วไป ซึ่งสอดคล้องกับงานวิจัยของณัฐณี ศิริโชติ และกมลวรรณ ตังธนากานนท์ (2559) ที่เป็นการศึกษาเปรียบเทียบการประเมินตนเองในทักษะการปฏิบัติทดลองความสามารถทางวิทยาศาสตร์ของผู้เรียนผ่านเครื่องมือเกณฑ์การให้คะแนนแบบรูปริกสองชั้นและเกณฑ์การให้คะแนนรูปริกทั่วไป พบว่า ผู้เรียนมีความเข้าใจต่อข้อรายการย่อยและเกณฑ์การประเมินซึ่งส่งผลต่อการประเมินตนเองของวิธีการให้คะแนนแบบรูปริกสองชั้นได้ดีกว่าการให้คะแนนรูปริกทั่วไป ในขณะที่วิธีตรวจให้คะแนนวิเคราะห์ย่อย แบ่งเกณฑ์การให้คะแนนตามองค์ประกอบ 3 องค์ประกอบ คือ ความถูกต้องของเนื้อหา การเข้าใจสิ่งที่ต้องการวัด และการสื่อความหมายทางคณิตศาสตร์ ส่วนวิธีตรวจให้คะแนน Knox ให้คะแนนแบบมาตรฐานประมาณค่า 3 ระดับ (2, 1, 0) ซึ่งทั้งวิธีตรวจให้คะแนนวิเคราะห์ย่อยและวิธีตรวจให้คะแนน Knox จะพิจารณาเพียงความครบถ้วนของคำตอบเท่านั้น ไม่ให้สารสนเทศย้อนกลับว่าผู้เรียนควรปรับปรุงส่วนใดได้ละเอียดเท่าวิธีตรวจให้คะแนนรูปริกสองชั้น

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลวิจัยไปใช้

การนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนนรูปริกสองชั้นที่พัฒนาขึ้นในการวิจัยครั้งนี้ไปใช้ ผู้วิจัยเสนอแนะไว้ 3 แนวทาง ดังนี้

แนวทางที่ 1 สำหรับการนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ไปใช้

1.1 ครูผู้ใช้แบบสอบควรรีศึกษาคู่มือการใช้แบบสอบอัตโนมัติประยุกต์วัดทักษะการคิด วิจรรณญาณทางคณิตศาสตร์ก่อนนำไปใช้ แบบสอบอัตโนมัติประยุกต์ฉบับนี้เป็นแบบสอบที่มีความ ยากง่ายอยู่ในระดับปานกลางค่อนข้างไปทางยากมีคุณภาพด้านความเที่ยงอยู่ในเกณฑ์ที่เหมาะสม จึงเหมาะกับการนำไปใช้ส่งเสริมทักษะการคิดวิจรรณญาณทางคณิตศาสตร์ซึ่งเป็นทักษะการคิด ขั้นสูงสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น

1.2 แบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจรรณญาณทางคณิตศาสตร์ที่ผู้วิจัย สร้างขึ้นเหมาะสำหรับครูที่ต้องการใช้ในการวัดผลแบบอิงกลุ่ม ยังไม่สามารถนำไปใช้ในการวัดผล แบบอิงเกณฑ์ได้ เนื่องด้วยแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจรรณญาณทางคณิตศาสตร์ ที่ใช้วิธีตรวจให้คะแนน Knox ยังมีความเที่ยงของแบบสอบแบบอิงเกณฑ์ที่ผ่านการพิจารณาจาก ค่าสัมประสิทธิ์การสุรุปอ้างอิงการตัดสินใจเชิงสัมบูรณ์ไม่ถึงในระดับที่ยอมรับได้

แนวทางที่ 2 สำหรับการนำวิธีตรวจให้คะแนนรูปrikสองชั้นไปใช้

2.1 เนื่องจากวิธีตรวจให้คะแนนรูปrikสองชั้นที่ผู้วิจัยพัฒนาขึ้นเป็นวิธีตรวจให้ คะแนนที่มีความละเอียด เนื่องจากในแต่ละข้อคำถาม จะประเมินตามข้อรายการย่อยซึ่งแบ่งตาม องค์ประกอบ คือ ด้านความถูกต้องของเนื้อหา ด้านการเข้าใจสิ่งที่ต้องการวัด และด้านการสื่อ ความหมายทางคณิตศาสตร์ ด้วยเหตุนี้วิธีตรวจให้คะแนนรูปrikสองชั้นจึงพิจารณาขั้นตอนของ กระบวนการคิดซึ่งไม่ใช่การพิจารณาเพียงแค่ความครบถ้วนสมบูรณ์ของคำตอบ ทำให้ได้ สารสนเทศย้อนกลับไปสู่ตัวผู้เรียน สามารถชี้ให้เห็นถึงจุดเด่นที่ควรพัฒนาต่อยอด และจุดบกพร่อง ที่ควรแก้ไขปรับปรุง ดังนั้นครูที่ต้องการใช้วิธีตรวจให้คะแนนรูปrikสองชั้นควรรีศึกษาเกณฑ์การ ตรวจให้คะแนนรูปrikสองชั้น และฝึกตรวจประเมินแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิด วิจรรณญาณทางคณิตศาสตร์ด้วยวิธีตรวจให้คะแนนรูปrikสองชั้น

2.2 วิธีตรวจให้คะแนนรูปrikสองชั้นเป็นวิธีที่ให้สารสนเทศย้อนกลับที่ละเอียด ชี้ให้เห็นถึงจุดเด่นที่ควรพัฒนาต่อยอด และจุดบกพร่องที่ควรแก้ไขปรับปรุงจากการทดสอบทั้งใน แ่งของผู้เรียนและผู้สอน ทั้งนี้ครูที่ต้องการนำวิธีตรวจให้คะแนนรูปrikสองชั้นไปใช้ในการตรวจให้ คะแนนแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจรรณญาณทางคณิตศาสตร์ควรพิจารณา สารสนเทศที่ได้จากทั้งชั้นสองชั้น โดยชั้นที่ 1 เหมาะสำหรับการใช้งานจริงของครู เพื่อนำ สารสนเทศที่ได้จากผู้สอบไปปรับปรุงหรือพัฒนาต่อยอดในตัวผู้เรียนต่อ สำหรับชั้นที่ 2 เหมาะสำหรับการรายงานผลการสอบในภาพรวมของผู้เรียน

แนวทางที่ 3 สำหรับการนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่มีความเที่ยงในระดับที่ยอมรับได้จากการออกแบบการวัดแบบไขว้ ($p \times r \times o$) ไปใช้

3.1 กรณีที่ครูต้องการใช้วิธีตรวจให้คะแนนรูปกรอกสองชั้นตรวจให้คะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับการวัดผลแบบอิงกลุ่มที่มีค่าความเที่ยงของแบบสอบตั้งแต่ 0.7 ขึ้นไป สำหรับผู้ประเมิน 1 คน ควรใช้แบบสอบอัตนัยประยุกต์จำนวน 2 เหตุการณ์ และสำหรับผู้ประเมิน 2 คน ควรใช้แบบสอบอัตนัยประยุกต์จำนวน 2 เหตุการณ์ กรณีที่ครูต้องการใช้วิธีตรวจให้คะแนนวิเคราะห์ย่อยตรวจให้คะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่มีค่าความเที่ยงตั้งแต่ 0.7 ขึ้นไป สำหรับผู้ประเมิน 1 คน ควรใช้แบบสอบอัตนัยประยุกต์จำนวน 3 เหตุการณ์ และสำหรับผู้ประเมิน 2 คน ควรใช้แบบสอบอัตนัยประยุกต์จำนวน 2 เหตุการณ์ กรณีที่ครูต้องการใช้วิธีตรวจให้คะแนน Knox ตรวจให้คะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่มีค่าความเที่ยงตั้งแต่ 0.7 ขึ้นไป สำหรับผู้ประเมิน 1 คน ควรใช้แบบสอบอัตนัยประยุกต์จำนวน 6 เหตุการณ์ และสำหรับผู้ประเมิน 2 คน ควรใช้แบบสอบอัตนัยประยุกต์จำนวน 5 เหตุการณ์

3.2 กรณีที่มีข้อจำกัดของความพร้อมในการหาผู้ประเมินที่มีคุณสมบัติตามที่กำหนดในการตรวจให้คะแนนด้วยวิธีตรวจให้คะแนนรูปกรอกสองชั้นในแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่มีความเที่ยงของแบบสอบในระดับที่ยอมรับได้ 0.7 ขึ้นไป เมื่อใช้ผู้ประเมิน 2 คน ใช้แบบสอบ 2 เหตุการณ์ มีความเที่ยงของแบบสอบแบบอิงกลุ่มที่พิจารณาผ่านค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมพัทธ์อยู่ที่ 0.81 และความเที่ยงของแบบสอบแบบอิงเกณฑ์ที่พิจารณาผ่านค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจ เชิงสัมบูรณ์อยู่ที่ 0.73 สามารถลดจำนวนผู้ประเมินลงแต่ต้องใช้แบบสอบที่มีจำนวนเหตุการณ์ที่เพิ่มขึ้น ซึ่งยังคงมีเวลาที่เหมาะสมในการจัดการสอบ คือ ผู้ประเมิน 1 คน ใช้แบบสอบ 3 เหตุการณ์ ใช้เวลาในการสอบ 1 ชั่วโมง มีความเที่ยงของแบบสอบแบบอิงกลุ่มที่พิจารณาผ่านค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมพัทธ์อยู่ที่ 0.78 และความเที่ยงของแบบสอบแบบอิงเกณฑ์ที่พิจารณาผ่านค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจ เชิงสัมบูรณ์อยู่ที่ 0.74 ซึ่งจะเห็นได้ว่าค่าความเที่ยงของแบบสอบอัตนัยประยุกต์ยังคงผ่านเกณฑ์ความเที่ยงของแบบสอบในระดับที่ยอมรับ 0.7 ขึ้นไป และเป็นแบบสอบที่สามารถวัดผลได้ทั้งแบบอิงกลุ่มและอิงเกณฑ์

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

สำหรับการแนะนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนนรูปrikสองชั้นที่พัฒนาขึ้นในการทำวิจัยครั้งต่อไป ผู้วิจัยเสนอแนะไว้ 3 แนวทาง ดังนี้

แนวทางที่ 1 สำหรับการแนะนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ในการทำวิจัยครั้งต่อไป

1.1 เนื่องจากแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่พัฒนาขึ้นสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น ดังนั้นควรมีแบบสอบอัตนัยประยุกต์วัดทักษะการคิดขั้นสูงในด้านอื่น ๆ อาทิ ทักษะการคิดแก้ปัญหา Problem solving path 'MasterMind Matrix' ซึ่งเป็นเส้นทางการคิดแก้ปัญหาในวิธีการคิดแนวใหม่ที่ออกแบบการคิดผ่านแผนผังการคิดแบบผสม เพื่อช่วยในการคิดแก้ปัญหาที่มีลักษณะยากและท้าทายในชีวิตจริงที่ละขั้นตอน และขยายไปถึงการใช้แบบสอบอัตนัยประยุกต์สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย เพื่อนำผลการวิจัยไปใช้ในการพัฒนาและปรับปรุงคุณภาพการจัดการเรียนการสอนในหลักสูตรฐานสมรรถนะต่อไป

1.2 การออกแบบแต่ละเหตุการณ์ต้องมีความเท่าเทียมกันในการวัด ทั้งความยากง่ายของแต่ละเหตุการณ์ผ่านการวัดในเรื่องเดียวกันทุกเหตุการณ์ อาทิ วัดการคิดวิจารณ์ผ่านเรื่องอัตราส่วน ทุกเหตุการณ์ที่สร้างขึ้นในระบบสอบอัตนัยประยุกต์ก็ควรเป็นเรื่องอัตราส่วน เพื่อนำไปสู่การหาข้อสรุปของจำนวนเหตุการณ์ที่ใช้ในระบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ที่ใช้วิธีตรวจให้คะแนน 3 วิธี คือ วิธีตรวจให้คะแนนรูปrikสองชั้น วิธีตรวจให้คะแนนวิเคราะห์ย่อย และวิธีตรวจให้คะแนน Knox จนได้แบบสอบที่มีความเที่ยงในระดับที่ยอมรับได้ เพื่อให้แบบสอบอัตนัยประยุกต์ฉบับนี้สามารถนำไปใช้ได้ทั้งการวัดผลแบบอิงกลุ่มและอิงเกณฑ์

แนวทางที่ 2 สำหรับการแนะนำวิธีตรวจให้คะแนนรูปrikสองชั้นในการทำวิจัยครั้งต่อไป

2.1 ควรพัฒนาวิธีตรวจให้คะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ในรูปแบบอื่น อาทิ วิธีตรวจให้คะแนนรูปrikแบบแอนโนเทต ที่เป็นการประเมินพฤติกรรมหรือลักษณะนั้น ๆ โดยภาพรวม แล้วอธิบายจุดเด่นและจุดด้อยของสิ่งที่ประเมินในแต่ละองค์ประกอบ ซึ่งทั้งจุดเด่นและจุดด้อยนี้จะเป็นสารสนเทศเพื่อการวินิจฉัยและให้ข้อมูลย้อนกลับในการพัฒนาผู้เรียนได้เช่นกัน

2.2 ควรศึกษาวิธีการกำหนดเกณฑ์สำหรับขั้นที่ 2 ของวิธีตรวจให้คะแนนรูปrikสองขั้นว่าควรใช้วิธีใดในการกำหนดช่วงหลังการแปลงคะแนน ช่วงคะแนนที่ได้ควรมีที่ระดับจึงจะเหมาะสมและในแต่ละระดับคะแนน ที่ได้ (3, 2, 1, 0) ให้คำอธิบายรายละเอียดผลการวัดการคิดวิจารณ์ญาณทางคณิตศาสตร์ของผู้สอบได้อย่างไร

แนวทางที่ 3 สำหรับการแนะนำแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่มีความเที่ยงในระดับที่ยอมรับได้จากการออกแบบการวัดแบบไขว้ ($p \times r \times o$) ในการทำวิจัยครั้งต่อไป

3.1 ควรศึกษาองค์ประกอบอื่น ๆ ที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง เช่น จำนวนข้อคำถามย่อย หรือศึกษาเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของการวัดที่ใช้รูปแบบความสัมพันธ์แบบ Cross และ Nested ในแง่จำนวนเหตุการณ์และจำนวนผู้ประเมิน

3.2 สำหรับการสร้างแบบสอบอัตโนมัติประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ผ่านการออกแบบการวัดแบบไขว้ ($p \times r \times o$) ที่ใช้จำนวนผู้ประเมินน้อย แต่ใช้จำนวนเหตุการณ์ในแบบสอบมากขึ้นเพื่อให้ได้แบบสอบที่มีความเที่ยงในระดับที่ยอมรับได้ 0.7 ขึ้นไป ต้องพิจารณาถึงขอบเขตเรื่องที่จะนำมาสร้างในแต่ละเหตุการณ์ของแบบสอบ เนื่องจากเรื่องที่มีเนื้อหาที่มีความลึก อาจต้องขยายเวลาที่ใช้ในการทำแต่ละเหตุการณ์ ซึ่งอาจทำให้มีตัวแปรแทรกซ้อนที่ส่งผลต่อผู้สอบมากขึ้น อาทิ ความเมื่อยล้าในการทำแบบสอบ

บรรณานุกรม

ภาษาไทย

กติกิกร กมลรัตน์สมบัติ และกมลวรรณ ตังธนากานนท์. (2559). *ผลของการให้ข้อมูลย้อนกลับจากแบบสอบอัตนัยประยุกต์ที่มีต่อพัฒนาการความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษา* [วิทยานิพนธ์ปริญญาามหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย].

Chulalongkorn University Intellectual Repository (CUIR).

<http://cuir.car.chula.ac.th/handle/123456789/50593>

กนกวรรณ ศรีรักษา. (2554). *Modified Essay Question (MEQ)*. ขอนแก่นเวชสาร.

กมลวรรณ ตังธนากานนท์. (2563). *การวัดและประเมินทักษะการปฏิบัติ* (พิมพ์ครั้งที่ 3). สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

กลุ่มนิเทศ ติดตาม และประเมินผลการจัดการศึกษา สพพ.นครปฐม เขต 2. (2564, 11 กุมภาพันธ์).

หลักสูตฐานสมรรถนะ. <http://chaisri-nites.hi-supervisory5.net/home/smrrthna>

กิตติพันธ์ วิบูลศิลป์ และจินดิษฐ์ ละออบปักษิณ. (2560). *ผลการจัดกิจกรรมการเรียนรู้คณิตศาสตร์ตามแนวคิดห้องเรียนย้อนกลับทางร่วมกับการเรียนรู้เชิงรุกที่มีต่อความสามารถในการแก้ปัญหาคณิตศาสตร์และความสามารถในการคิดอย่างมีวิจารณญาณของนักเรียนชั้นมัธยมศึกษาปีที่ 5* [วิทยานิพนธ์ปริญญาามหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย].

Chulalongkorn University Intellectual Repository (CUIR).

<http://cuir.car.chula.ac.th/handle/123456789/58427>

จิรายุ เถาว์โท และคณะ. (2559). การศึกษาค่าความเชื่อมั่นของคะแนนแบบทดสอบอัตนัยคณิตศาสตร์ของ นักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีจำนวนผู้ตรวจและรูปแบบการตรวจให้คะแนนต่างกัน โดยใช้ทฤษฎีการสรุปอ้างอิง. *วารสารหาดใหญ่วิชาการ*, 14(1), 1-14.

<https://so01.tcithaijo.org/index.php/HatyaiAcademicJournal/article/view/82192/65334>

ชนิสรา สงวนไ้ และศิริชัย กาญจนวาสิ. (2558). *การเปรียบเทียบความเที่ยงของแบบสอบวัดความสามารถใน การแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของ ผลการวัด* [วิทยานิพนธ์ปริญญาามหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย].

Chulalongkorn University Intellectual Repository (CUIR).

<http://cuir.car.chula.ac.th/handle/123456789/51081>

ชูศรี วงศ์รัตน์. (2560). *เทคนิคการสร้างเครื่องมือวิจัย: แนวทางการนำไปใช้อย่างมืออาชีพ*. อมรการพิมพ์.

เชิดศักดิ์ ไอรณณิรัตน์. (2558). การสร้างข้อสอบอัตนัยประยุกต์. *เวชบันทึกศิริราช*,

8(1), 48-49. <https://he02.tci-thaijo.org/index.php/simedbull/article/view/81482>

- โชติกา ภาชีผล. (2559). *การวัดและประเมินผลการเรียนรู้ Learning Measurement and Evaluation*. โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ฐิติมา ฐิติรุ่งเรือง. (2544). *การพัฒนาแบบทดสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาวิชาแนวคิดพื้นฐานและหลักการพยาบาล* [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยทักษิณ]. TSU Knowledge bank (TSUKB).
<http://kb.tsu.ac.th/jspui/bitstream/123456789/1324/1/69745.pdf>
- ณัฐณี ศิริโชติ และกมลวรรณ ตั้งชนกานนท์. (2559). *ปฏิสัมพันธ์ระหว่างวิธีการประเมินตนเองและความสามารถทางวิทยาศาสตร์ที่มีต่อการพัฒนาการทักษะปฏิบัติการทดลองของนักเรียน* [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย]. Chulalongkorn University Intellectual Repository (CUIR).
<http://cuir.car.chula.ac.th/bitstream/123456789/50594/1/5783825127.pdf>
- ณัฐภรณ์ หลาวทอง. (2561). *การสร้างเครื่องมือการวิจัยทางการศึกษา* (พิมพ์ครั้งที่ 2). สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ดร.ณิ อภัยกวี, ประกิตติยา ทักษิณ และกมลวรรณ ตั้งชนกานนท์. (2562). การศึกษาผลการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ใช้โมเดลหลายองค์ประกอบของรรัสซ และทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด. *วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม*, 26(1), 110-124.
- ดวงมณี หล้าคำดง. (2544). *การสร้างแบบทดสอบเอ็ม อี คิว กลุ่มสร้างเสริมประสบการณ์ชีวิต เรื่องพลังงานและสารเคมี สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ดำเนินการสอบโดยใช้คอมพิวเตอร์* [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยมหาสารคาม].
- ทวี สระน้ำคำ และสุรี รอดโพธิ์ทอง. (2551). *ผลของวิธีการสอนแบบนิรนัยและวิธีสอนแบบอุปนัยที่มีแบบฝึกหัดหลังเรียนต่างกัน โดยใช้บทเรียนบทเว็บในวิชาฟิสิกส์ที่มีต่อการคิดวิจารณ์ญาณของนักเรียนชั้นมัธยมศึกษาปีที่ 4* [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย]. Chulalongkorn University Intellectual Repository (CUIR).
<http://cuir.car.chula.ac.th/bitstream/123456789/59036/1/Tawee%20Sranamkam.pdf>
- นงรัก ไทธานี. (2542). *การสร้างแบบทดสอบเอ็ม อี คิว ที่ดำเนินการสอบโดยใช้คอมพิวเตอร์เพื่อวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์* [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยมหาสารคาม].
- นภา หลิมรัตน์. (2551, ตุลาคม). *การวัดผลและประเมินผลการศึกษา*. PSU Medical Education Resources. <http://teachingresources.psu.ac.th/document/2551/limrat/1.pdf>
- บุษยรัตน์ จันทร์ประเสริฐ, ณัฐภรณ์ หลาวทอง, สัจวง ังดกระโทก. (2560). *ความสอดคล้องในแนว*

เดียวกันระหว่างข้อสอบในการประเมินระดับชาติ กับข้อสอบในการประเมินระดับชั้นเรียน กลุ่ม
สาระการเรียนรู้วิทยาศาสตร์: การประยุกต์ใช้โมเดลหลายองค์ประกอบของราล์ช และทฤษฎีการ
สรุปอ้างอิงความน่าเชื่อถือของผลการวัด [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์
มหาวิทยาลัย]. Chulalongkorn University Intellectual Repository (CUIR).

<http://cuir.car.chula.ac.th/handle/123456789/59575>

พรพิมล ปิลันธนาภรณ์ และชิตชนก เชิงเขาว์. (2561). การพัฒนาแบบทดสอบอัตนัยประยุกต์วัด
ความสามารถในการตัดสินใจทางคลินิกของนักศึกษาแพทย์ มหาวิทยาลัยสงขลานครินทร์.
วารสารมหาวิทยาลัยราชภัฏยะลา, 15(3), 282-289.

พวงแก้ว ปุณยกนก. (2531). แบบสอบอัตนัยประยุกต์ (เอ็ม อี คิว) เพื่อใช้วัดทักษะการแก้ปัญหา.

https://doi.nrct.go.th/ListDoi/listDetail?Resolve_DOI=10.14457/CU.res.1988.1

มาฆพันธ์ อำนาคิล และกมลวรรณ ตั้งชนกานนท์. (2557). การเปรียบเทียบความเที่ยงของแบบสอบ
อัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ที่ใช้วิธีการตรวจให้คะแนน
ที่แตกต่าง ภายใต้จำนวนเหตุการณ์ที่ต่างกัน: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความ
น่าเชื่อถือของผลการวัด [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย].

Chulalongkorn University Intellectual Repository (CUIR).

<http://cuir.car.chula.ac.th/bitstream/123456789/46285/1/5683866827.pdf>

ระพีพัฒน์ แก้วอ่ำ. (2559). การใช้คำถามปลายเปิดในการสอนคณิตศาสตร์. *วารสารศรีนครินทรวิ
โรฒวิจัยและพัฒนา (สาขามนุษยศาสตร์และสังคมศาสตร์*, 8(15), 206-211.

<http://ejournals.swu.ac.th/index.php/swurd/article/viewFile/7458/6852>

ลักขมี จันทราช. (2544). การเปรียบเทียบคุณภาพแบบสอบอัตนัยประยุกต์ (MEQ) และแบบทดสอบ
เลือกตอบ (MCQ) เพื่อวัดความสามารถในการแก้ปัญหา [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต
มหาวิทยาลัยทักษิณ]. TSU Knowledge bank (TSUKB).

วรนิพิฏ์ พันธุ์หนองหว่า และอัมพร ม้าคอง. (2559). ผลการจัดกิจกรรมการเรียนรู้คณิตศาสตร์โดย
ใช้กระบวนการแก้ปัญหาที่มีโครงสร้างไม่สมบูรณ์ร่วมกับกลวิธีการเสริมต่อความคิดที่มีต่อ
ความสามารถในการคิดอย่างมีวิจารณญาณ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 [วิทยานิพนธ์
ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย].

วีรวัฒน์ ไทยท่า และทรงชัย อักขรคิด. (2564). ความสามารถในการตั้งปัญหาทางคณิตศาสตร์ของ
นักศึกษาครู Mathematical Problem Posing Abilities of Preservice Teachers. *วารสาร
คณิตศาสตร์ Mathematical Journal*, 66(704), 26-45.

ศิริชัย กาญจนวาสี. (2556). ทฤษฎีการทดสอบแนวใหม่ *Modern Test Theories* (พิมพ์ครั้งที่ 4).

สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2563). *ทฤษฎีการทดสอบแนวใหม่ Modern Test Theories* (พิมพ์ครั้งที่ 5).

สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2564). *เกี่ยวกับ PISA* (2564, 11 มิถุนายน).

<https://pisathailand.ipst.ac.th/about-pisa/>

สมคิด เทียบพิสุทธิ. (2550). *การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบความเรียงประยุกต์วิชา*

คณิตศาสตร์ที่มีวิธีการตรวจและจำนวนผู้ตรวจต่างกัน [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต

มหาวิทยาลัยทักษิณ]. TSU Knowledge bank (TSUKB).

<http://kb.tsu.ac.th/jspui/bitstream/123456789/681/1/114500.pdf>

สายหมอก ขุนศักดิ์ดา. (2543). *การสร้างแบบทดสอบวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์*

เรื่องฟังก์ชันตรีโกณมิติ และกำหนดการเชิงเส้น ตามแนวคิดเอ็มอีคิว ในระดับชั้นมัธยมศึกษาปีที่ 5 [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยขอนแก่น]. Khon Kaen University library.

<https://opac.kku.ac.th/Catalog/Bibitem.aspx?BibID=b00193157>

สำนักงานราชบัณฑิตยสภา. (2558). *พจนานุกรมศัพท์ศึกษาศาสตร์ร่วมสมัย*

ฉบับราชบัณฑิตยสภา. สำนักงานราชบัณฑิตยสภา.

สำนักงานเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ. (2563). *การจัดการเรียนรู้ฐานสมรรถนะ*

เชิงรุก. กลุ่มมาตรฐานการศึกษา สำนักมาตรฐานการศึกษาและพัฒนาการเรียนรู้.

สำนักงานเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ. (2563). *การจัดการเรียนรู้ฐานสมรรถนะ*

เชิงรุก. กลุ่มมาตรฐานการศึกษา สำนักมาตรฐานการศึกษาและพัฒนาการเรียนรู้.

สิริภักดิ์ ศิริโพ. (2558). *การคิดวิจารณ์ญาณ: มุมมองการศึกษาและการบริหารจัดการ*. ศาสตราจารย์

คิด รวมบทความเรื่องการคิดและการสอนคิด. มหาวิทยาลัยธุรกิจบัณฑิตย์.

สุจิตรา ปทุมลังการ. (2552, มกราคม). *ความรู้เกี่ยวกับหลักสูตรฐานสมรรถนะและการประเมินผล*.

<http://www.atsn.ac.th/images/Upload/file/CBCApplications.pdf>

สุพัฒตา ภูสอด้สี. (2553). *การสร้างแบบทดสอบเอ็มอีคิว วัดความสามารถในการแก้ปัญหาทาง*

คณิตศาสตร์ เรื่องเรขาคณิตวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 4 [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยสารคาม].

สุภชิต ผดุงผล และกมลวรรณ ตั้งธนานนท์. (2559). *การเปรียบเทียบความเที่ยงของแบบสอบอัตนัย*

ประยุกต์สำหรับวัดความสามารถในการใช้วิธีการทางวิทยาศาสตร์ในวิชาฟิสิกส์ ภายใต้จำนวน เหตุการณ์และจำนวนผู้ตรวจที่ต่างกัน: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของ ผลการวัด [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย]. Chulalongkorn

University Intellectual Repository (CUIR).

<http://cuir.car.chula.ac.th/handle/123456789/55715>

สุไรยา หมั่นหมัด. (2549). การพัฒนาแบบทดสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยทักษิณ]. TSU Knowledge bank (TSUKB).

<http://kb.tsu.ac.th/jspui/bitstream/123456789/463/1/102970.pdf>

สุไรยา หมั่นหมัด. (2549). การพัฒนาแบบสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยทักษิณ]. TSU Knowledge bank (TSUKB).

<http://kb.tsu.ac.th/jspui/bitstream/123456789/463/1/102970.pdf>

อนุสรณ์ หงส์ขุนทด. (2563, 23 สิงหาคม). ทำไมต้องหลักสูตรฐานสมรรถนะ.

<http://krukob.com/web/1-159/>

อัมพร ม้าคนอง. (2544). ทักษะและกระบวนการทางคณิตศาสตร์ : การพัฒนาเพื่อพัฒนาการ. สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

อำมาลา สารชาติ. (2548). การพัฒนาแบบสอบ เอ็ม อี คิว สำหรับวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยราชภัฏอุบลราชธานี]. Thai Library Integrated System (ThaiLIS).

ภาษาอังกฤษ

Abatihun, A. S. (2013). Function of Training Background and Teaching Experience with Particular Reference to Final Examinations. *American Journal of Educational Research*, 1(8), 300-306. <https://doi.org/10.12691/education-1-8-6>

Abdullah, N., Idris, N., Hamzah, M. S. G. B., Sembak, S. (2015). Planing and implementation of School-Based Assessment (SBA) among teachers. *Social and Behavior Sciences*, 211, 247-254.

Albert, O. (2003). *Developing and using classroom assessment* (3 rd ed.). Upper Saddle River, NJ: Merrill.

Badger, E., & Thomas, B. (1992). Open-ended questions in reading. *Practical Assessment, Res. And Evaluation*, 3(4), 1-13. <https://doi.org/10.7275/fryf-z044>

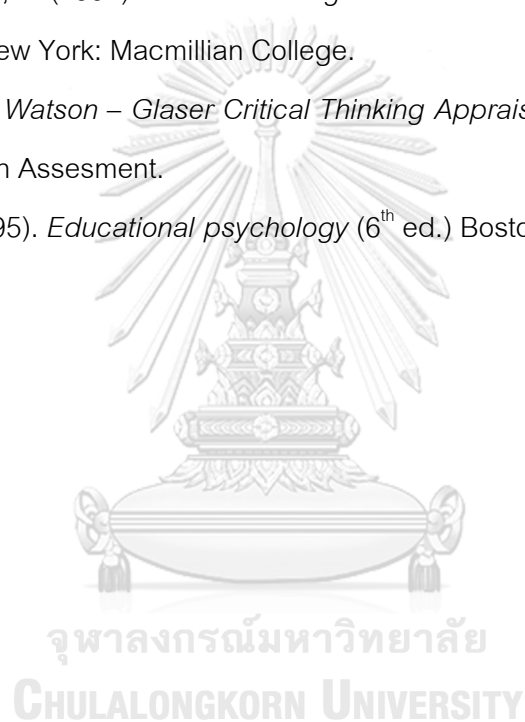
Brennan, R. L., Xiaohong, G., & Dean, A. C. (1995). Genelizability Analysis of Work Keys Listening and Writing Tests. *Educational and Psychological Measurement*, 55(2), 157-176. <https://doi.org/10.1177/0013164495055002001>

- Carl, W. S., Stephen, R. H., James, W. M., Melissa, B. W., Renee, E. L., Martha, R., Timothy, T. B., Melvin, D. L., & Kinnard, P. W. (1999). Using Generalizability Theory to Estimate the Reliability Writing Scores Derived from Holistic and Analytical Scoring Methods. *Educational and Psychological Measurement*, 59(3), 492-506.
<https://in.booksc.eu/ireader/44276528>
- Charles, D. K., & Richard, L.A. (1990). *Classroom measurement and evaluation*. F.E. Peacock. Chulalongkorn University Intellectual Repository (CUIR).
<http://cuir.car.chula.ac.th/handle/123456789/55144>
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Nageswari, R. (1972). Psychometrics: The Dependability of Behavior Measurements. *Science*, 178(4067), 1275.
<https://doi.org/10.1126/science.178.4067.1275>
- Cronbach, L. J., Nageswari, R., & Gleser, G. C. (1963). Theory of Generalizability: A liberalization of Reliability Theory. *British Journal of Mathematical and Statistical Psychology*, 16, 137-173. <https://doi.org/10.1111/j.2044-8317.1963.tb00206.x>
- Decaroli, J. (1973). What Research Say to the Classroom Teacher: Critical Thinking. *Social Education*, 37(1), 67-68.
- Dressel, P. L., & Mayhew, L. B. (1957). *General education explorations in evaluation* (2nd ed.). American Council on Education.
- Ebel, R. L., & Frisbie, D. A. (1972). *Essentials of Educational Measurement*. Englewood: Prentice-Hall.
- Ennis, R. H. (1985) Critical thinking and specificity. *Educational Researcher*, 18(3), 4-10.
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. *Teaching for Thinking*, 9-26.
- Facione, P. A. (1992). *Critical Thinking: What It is and Why It Counts*.
https://www.researchgate.net/publication/251303244_Critical_Thinking_What_It_Is_and_Why_It_Counts
- Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical Inference for Coefficient Alpha. *Applied Psychological Measurement*, 11(1), 93-103.
<https://doi.org/10.1177/014662168701100107>
- Feletti, G. I. (1980). Reliability and validity studies on modified essay questions. *Academic Medicine*, 55(11), 933-941. <https://doi.org/10.1097/00001888-198011000-00006>

- Gagne, R. M. (1985). *The conditions of learning* (4th ed.). New York, NY : Holt, Rinehart & Winston.
- Glaser, G. C. (1941). An Experiment in the Development of Critical Thinking. *Contributions to Education (Columbia University)*, 843.
<http://www.worldcat.org/oclc/5993128>
- Gleser, G. C., Cronbach, L. J., & Nageswari, R.. (1965). Generalizability of Scores Influenced by Multiple Sources of Variance. *Psychometrika*, 30, 395-418.
<https://doi.org/10.1007/BF02289531>
- Halpern, D. F. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of critical thinker. *New directions for teaching and learning*, 99(80), (pp. 69-74). <https://doi.org/10.1002/tl.8005>
- Hamzah, M. S. G. B., Idris, N., Abdullah, S. K., Abdullah, N., Muhammad, M. M. (2014). Development of the Double Layer Rubric for the Study on the Implementation of School-Based Assessment Among Teachers. *US-China Education Review*, 5(4), 245-256.
https://pdfs.semanticscholar.org/a63c/65fc0d7064c5ec51e60949e8424f244781b4.pdf?_ga=2.222926363.1275499131.1626344558-596869108.1626344558
- Hodgkin, & Knox. (1980). How to Use Modified Essay Question. *Medical Teacher*, 2, 20-24.
- Hoyt, C. (1941). Test Reliability Obtained by Analysis of Variance. *Psychometrika*, 6, 153-160. <https://doi.org/10.1007/BF02289270>
- Jonsson, A. (2010). The use of transparency in the interactive examination for student teachers. *Assessment in Education: Principles, Policy & Practice*, 17(2), 185-199.
- Linquist. (1953). *Design and Analysis of Experiments in Education and Psychology*. Houghton Mifflin.
- Lisa, R. H., Alison, M. H., Bridgette, L. K., & Sean, P. L. (2020). Using generalizability theory to evaluate the comparative reliability of developmental measures in neurogenetic syndrome and low-risk populations. *Journal of Neurodevelopmental Disorders*, 12(1), 1-15. <https://doi.org/10.1186/s11689-020-09318-1>
- Medley, & Mitzel. (1963). Measuring Classroom Behavior by Systematic Observation. In N.L. Gange. *Handbook of Research on Teaching*. Chicago: Rand McNally.
- Mehtap, K., & Erdinc, C. (2020). Prospective mathematics teachers' critical thinking

- processes about scientific research: Newspaper article example. *Turkish Journal of Education*, 9(1), 22-45. <https://dergipark.org.tr/en/download/article-file/948207>
- Nagandla, K., Das Gupta, E., Motilal, T., Teng, C. L., Gangadaran, S. (2018). Performance of medical students in final professional examination: Can in-course continuous assessments predict students at risk?. *The National Medical Journal of India*, 31(5), 293-295. <https://doi.org/10.4103/0970-258X.261197>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill Inc.
- Penny, J. A., Johnson, R. L., & Gordon, B. (2000). The effect of rating augmentation on inter-rater reliability: An empirical study of a holistic rubric. *Assessing Writing*, 7(2), 143-164. [https://doi.org/10.1016/S1075-2935\(00\)00012-X](https://doi.org/10.1016/S1075-2935(00)00012-X)
- Penny, J. A., Johnson, R. L., & Gordon, B. (2000b). Using rating augmentation to expand the scale of an analytic rubric: *Journal of Experimental Education*, 68(3), 269-287. <http://www.jstor.org/stable/20152633>
- Quoc, C. T., Christian, U. K., Richard, J. S., Jason, L., & Oleg, N. M. (2020). Applying Generalizability Theory to Differentiate Between Trait and State in the Five Facet Mindfulness Questionnaire (FFMQ). *Mindfulness*, 11, 953-963. <https://doi.org/10.1007/s12671-020-01324-7>
- Rajaratnam, N., Cronbach, L. J., & Gleser, G. C. (1965). Generalizability of Stratified Parallel Tests. *Psychometrika*, 30, 39-56. <https://doi.org/10.1007/BF02289746>
- Richard, J. S., Olga, Z., Klaus, B., Susanne, S., & Julian, P. M. (2019). Assessment of University Students' Critical Thinking: Next Generation Performance Assessment. *International Journal of Testing*, 19, 337-362. <https://doi.org/10.1080/15305058.2018.1543309>
- Sanchez, W. B. (2013). Open-Ended Questions and The Process Standards. *The Mathematics Teacher*, 107(3), 206-211.
- Siti, A. A., & Arief, A. (2020). Mathematical Critical Thinking Ability in Solving Open-Ended Questions Viewed from Students' Curiosity. *Unnes Journal of Mathematics Education*, 9(1), 36-42. <https://journal.unnes.ac.id/sju/index.php/ujme/article/view/38099/15685>
- So, J. Y., Sang, Y. L., Sun, J. I., & Bee, S. K. (2018). Holistic rubric vs analytic rubric for measuring clinical performance levels in medical student. *BMC Medical Education*, 18(124), 1-6.

- Sudweeks, R. R., Reeve, S., & Bradshaw, W. S. (2004). A Comparison of Generalizability Theory and Many-Facet Rasch Measurement in an Analysis of College of Sophomore Writing. *Assessment Writing*, 9(3), 239-261. <https://eric.ed.gov/?id=EJ803625>
- Truong, Q.C., Krageloh, C.U., Siegert, R.J., Landon, J. & Medvedev, O. N. (2020). Applying Generalizability Theory to Differentiate Between Trait and State in the Five Facet Mindfulness Questionnaire (FFMQ). *Mindfulness*, 11, 953-963. <https://doi.org/10.1007/s12671-020-01324-7>
- Warnick, B., & Inch, E. (1994). *Critical thinking and communication: The use of reason in argument*. New York: Macmillian College.
- Watson, G. (1980). *Watson – Glaser Critical Thinking Appraisal – UK Edition Practice Test*. Pearson Assesment.
- Woolfolk, A. E. (1995). *Educational psychology* (6th ed.) Boston, MA: Allyn & Bacon.





ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY



ภาคผนวก ก
รายชื่อผู้ทรงคุณวุฒิ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

รายชื่อผู้ทรงคุณวุฒิตรวจสอบเครื่องมือและเกณฑ์การให้คะแนน

ผู้ทรงคุณวุฒิทางการสอนคณิตศาสตร์

ผู้ช่วยศาสตราจารย์ ดร.จنگล ทำสวน
คณิตศาสตร์

อาจารย์ประจำ ภาควิชาการศึกษา

ครูมานัส ทิพย์สัมฤทธิ์กุล

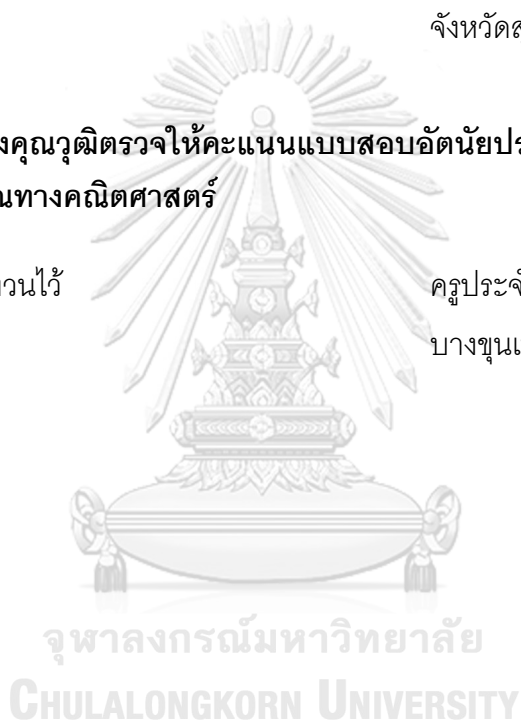
จุฬาลงกรณ์มหาวิทยาลัย

ครูเชี่ยวชาญ สาขาหลักสูตรและการ
สอน(คณิตศาสตร์) โรงเรียนสงวนหญิง
จังหวัดสุพรรณบุรี

รายชื่อผู้ทรงคุณวุฒิตรวจให้คะแนนแบบสอบอัตโนมัติวัดทักษะการคิด วิจารณ์ญาณทางคณิตศาสตร์

ครูชนิสรา สงวนไฉ่

ครูประจำโรงเรียนรัตนโกสินทร์สมโภช
บางขุนเทียน





ภาคผนวก ข

คู่มือการใช้แบบสอบอัตรานัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

คู่มือการใช้



แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์
ทางคณิตศาสตร์

สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3



นางสาวกัจจิรา จงสุกใส

รศ.ดร.กมลวรรณ ตั้งธนานนท์
(อาจารย์ที่ปรึกษา)

เอกสารฉบับนี้เป็นส่วนหนึ่งของ
วิทยานิพนธ์ปริญญาโทมหาบัณฑิต
สาขาวิชาวิธีวิทยาการพัฒน
นวัตกรรมการศึกษ
ภาควิชาวิจัยและจิตวิทยา
การศึกษ คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2564



คำนำ

คู่มือการใช้แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 เล่มนี้ จัดทำขึ้นเพื่อให้ครูใช้เป็นเครื่องมือและแนวทางในการวัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ซึ่งเป็นหนึ่งในทักษะการคิดขั้นสูงของหลักสูตรฐานสมรรถนะ โดยมีการกำหนดขอบเขตเนื้อหาวิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จากสาระทางคณิตศาสตร์ ได้แก่ สาระจำนวนและพีชคณิต (อัตราส่วน) และสาระการวัดและเรขาคณิต (ทฤษฎีบทพีทาโกรัส และพื้นที่ผิวและปริมาตร) เนื่องจากสาระการวัดและเรขาคณิตอาศัยความรู้พื้นฐานจากสาระจำนวนและพีชคณิต และประยุกต์เข้ากับสถานการณ์ในชีวิตจริงเพื่อส่งเสริมการคิดวิจารณ์ญาณทางคณิตศาสตร์ของนักเรียน โดยข้อคำถามในแต่ละเหตุการณ์จะวัดตามขั้นของการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ได้จากการสังเคราะห์เอกสารและงานวิจัยที่เกี่ยวข้อง 5 ขั้นตอน ดังนี้ ข้อคำถามที่ 1 ของทุกเหตุการณ์จะวัดความสามารถในการนิยามปัญหา ข้อคำถามที่ 2 ของทุกเหตุการณ์จะวัดความสามารถในการตั้งสมมติฐาน ข้อคำถามที่ 3 ของทุกเหตุการณ์จะวัดความสามารถในการรวบรวมข้อมูล ข้อคำถามที่ 4 ของทุกเหตุการณ์จะวัดความสามารถในการวิเคราะห์ข้อมูล และข้อคำถามที่ 5 ของทุกเหตุการณ์ จะวัดความสามารถในการสร้างข้อสรุป

สำหรับวิธีตรวจให้คะแนนของแบบสอบอัตนัยประยุกต์ฉบับนี้ คือ วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) ที่มีการตรวจให้คะแนน 2 ชั้น โดยชั้นที่ 1 ประเมินตามข้อรายการย่อยในแต่ละองค์ประกอบ ดังนี้ ข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 ข้อละ 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (4 คะแนน) และการเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) จากนั้นแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้คะแนนเต็มข้อละ 2 คะแนน ส่วนข้อคำถามที่ 4 และข้อคำถามที่ 5 ข้อละ 3 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (4 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) และการสื่อความหมายทางคณิตศาสตร์ (4 คะแนน) จากนั้นแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้คะแนนเต็ม ข้อละ 3 คะแนน ซึ่งวิธีตรวจให้คะแนนรูปริกสองชั้นเป็นวิธีที่ละเอียดเหมาะแก่การนำไปใช้และสามารถให้สารสนเทศย้อนกลับแก่ผู้สอบเพื่อนำไปปรับปรุงและพัฒนาต่อไปได้

สารบัญ

คำนำ	หน้า
สารบัญ	135
คำชี้แจงในคู่มือการใช้	136
ส่วนที่ 1 : ความสำคัญของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์	138
ส่วนที่ 2 : แนวปฏิบัติและการนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ไปใช้สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3	141
ส่วนที่ 3 : การตรวจให้คะแนนและการบันทึกคะแนน	144
บรรณานุกรม	154
ภาคผนวก ค แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3	158
ภาคผนวก ง แนวคำตอบและวิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric)	178
ภาคผนวก จ แบบบันทึกคะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์	214

คำชี้แจงในการใช้คู่มือ

จุดมุ่งหมายของคู่มือ

คู่มือนี้มีจุดมุ่งหมายเพื่อนำเสนอแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 และแนวปฏิบัติในการตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric)

ส่วนประกอบของคู่มือ

คู่มือนี้มีเนื้อหาแยกเป็น 3 ส่วน ดังนี้ ส่วนที่ 1 ความสำคัญของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์และวิธีตรวจให้คะแนนรูบริกสองชั้น ส่วนที่ 2 แนวปฏิบัติและนำเสนอแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ไปใช้ และส่วนที่ 3 แนวปฏิบัติในการตรวจให้คะแนนแบบสอบและการบันทึกคะแนน

คำนิยามศัพท์

แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ตามหลักสูตรฐานสมรรถนะ หมายถึง แบบสอบประเภทความเรียงที่ผู้วิจัยสร้างขึ้นโดยเหตุการณ์จำลองเป็นกรณีศึกษาตามเนื้อหาวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 มีการเสนอเหตุการณ์จำลองตามลำดับเหตุการณ์ และเสนอแยกออกเป็นข้อ ๆ ที่เป็นอิสระจากกัน แทรกคำถามตามวัตถุประสงค์ของการวัดเป็นระยะ ๆ แต่ละเหตุการณ์ประกอบด้วย ข้อคำถามจำนวน 5 ข้อ ตามขั้นตอนการคิดวิจารณ์เพื่อฝึกการจำแนกข้อเท็จจริงจากแหล่งข้อมูลที่กำหนดให้จนสามารถนำมาวิเคราะห์ และสร้างข้อสรุปของปัญหาในเหตุการณ์นั้นได้ สำหรับการทำให้แบบทดสอบผู้สอบต้องนำข้อมูลจากเหตุการณ์ที่เสนอมาใช้ตัดสินใจเพื่อตอบคำถามปลายเปิดเอง และต้องทำข้อสอบให้เสร็จที่ละหน้า เนื่องจากผู้สอบไม่สามารถย้อนกลับไปแก้ไขคำตอบในข้อที่ทำไปแล้วได้

การคิดวิจารณ์ทางคณิตศาสตร์ หมายถึง กระบวนการคิดทางคณิตศาสตร์ของนักเรียนในระดับชั้นมัธยมศึกษาตอนต้นผ่านแบบสอบอัตนัยประยุกต์อาศัยเหตุผลในการพิจารณาข้อมูลและหลักฐานที่มีอยู่อย่างรอบคอบ ตรวจสอบความน่าเชื่อถือของข้อมูล คุณภาพสัมพันธ์ของข้อมูลที่เกี่ยวข้องกับปัญหาหรือสถานการณ์ เพื่อสร้างกระบวนการจัดระบบข้อมูลในการแก้ปัญหาหรือตั้งสมมติฐาน นำมาสู่การสร้างข้อสรุปที่มีหลักฐานมาสนับสนุนความคิด ตั้งอยู่บนตรรกะอย่างมีเหตุและผล จนได้คำตอบที่เหมาะสม ซึ่งกระบวนการคิดวิจารณ์ทางคณิตศาสตร์ มี 5 ขั้นตอน ได้แก่ 1) ขั้น

นิยามปัญหา 2) ขั้นตั้งสมมติฐาน 3) ขั้นรวบรวมข้อมูล 4) ขั้นวิเคราะห์ข้อมูล และ 5) ขั้นสร้างข้อสรุป ซึ่งแต่ละขั้นตอน มีความหมายดังนี้

1) ขั้นนิยามปัญหา หมายถึง การกำหนดประเด็นปัญหาจากการพิจารณาข้อมูลหรือข้อโต้แย้ง ภายใต้ข้อตกลงเบื้องต้นในสถานการณ์ที่กำหนดให้ โดยปราศจากการให้ข้อมูลอื่นที่นอกเหนือจากบริบทของสถานการณ์หรือความคิดเห็นส่วนตัวเพิ่มเข้าไป

2) ขั้นตั้งสมมติฐาน หมายถึง การสร้างแนวทางคำตอบที่เป็นไปได้มากที่สุดหรือคาดคะเนแนวทางในการหาคำตอบ โดยอาศัยข้อมูลในสถานการณ์ที่มีอยู่เดิมประกอบการตัดสินใจอย่างสมเหตุสมผล

3) ขั้นรวบรวมข้อมูล หมายถึง การคัดเลือกข้อมูลที่มีความสัมพันธ์กับประเด็นปัญหา ข้อโต้แย้ง หรือข้อมูลที่คลุมเครือ โดยดึงข้อมูลหรือความรู้จากประสบการณ์เดิมมาใช้ เพื่อแสวงหาข้อมูลที่ต้องการและชัดเจนมากยิ่งขึ้น

4) ขั้นวิเคราะห์ข้อมูล หมายถึง การนำข้อมูลที่ผ่านการรวบรวมจากการพิจารณาความน่าเชื่อถือและ ความเพียงพอของข้อมูล มาเชื่อมโยงหาความสัมพันธ์โดยใช้ความรู้ทางคณิตศาสตร์ เพื่อนำไปสู่การสร้างข้อสรุป

5) ขั้นสร้างข้อสรุป หมายถึง การนำผลการวิเคราะห์ข้อมูล และหลักฐานที่มีความหนักแน่นน่าเชื่อถือมาใช้ในการตัดสินใจสรุปผลการทดสอบสมมติฐานอย่างเหมาะสมกับสถานการณ์ โดยตั้งอยู่บนตรรกะเหตุและผลทางคณิตศาสตร์

วิธีการตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric) หมายถึง วิธีการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์กับเกณฑ์การให้คะแนนแบบรูบริก 2 ชั้น โดยแบ่งออกเป็น 2 ระดับชั้น ชั้นที่ 1 อธิบายตามองค์ประกอบเช่นเดียวกับวิธีการตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) และในแต่ละองค์ประกอบจะประเมินตามข้อรายการย่อย โดยใช้มาตราประมาณค่า (rating scale) 3 ระดับ ได้แก่ 2 คะแนน คือ ถูกต้องทั้งหมด 1 คะแนน คือ ถูกต้องบางส่วน และ 0 คะแนน คือ ไม่ถูกต้องทั้งหมด หรือไม่มีการตอบคำถาม จากนั้นนำผลรวมของคะแนนในระดับชั้นที่ 1 มาแปลงเป็นคะแนนในระดับชั้นที่ 2 ตามเกณฑ์ที่กำหนดไว้

ส่วนที่ 1
ความสำคัญของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์และ
วิธีตรวจให้คะแนนรูปrikสองชั้น

การคิดวิจารณ์ญาณ (critical thinking) เป็นหนึ่งในสมรรถนะการคิดขั้นสูง (Higher Order Thinking: HOT) ตามหลักสูตรฐานสมรรถนะที่ทางสำนักงานเลขาธิการสภาการศึกษา (2563) ได้วางแนวทางไว้ โดยการคิดวิจารณ์ญาณเป็นกระบวนการคิดที่ต้องใช้เหตุผลในการพิจารณาอย่างรอบคอบผ่านการศึกษาค้นคว้าข้อมูล หลักฐาน จนสามารถแยกแยะข้อเท็จจริงหรือข้อคิดเห็นได้ ตลอดจนพิจารณาความน่าเชื่อถือของข้อมูล ตัดสินใจอย่างรอบคอบรอบด้าน เพื่อให้เห็นว่าสิ่งใดควรเชื่อหรือควรทำ และเพราะเหตุใด จนนำไปสู่แนวทางในการแก้ไขปัญหา (อัมพร ม้าคะนอง, 2554; สิริภักตร์ ศิริโท, 2558) ดังนั้นการศึกษาจำเป็นต้องมีเครื่องมือที่ดีในการวัดทักษะการคิดวิจารณ์ญาณของผู้เรียนให้เป็นที่ไปตามเป้าหมายที่ตั้งไว้ เพื่อส่งเสริมให้ผู้เรียนเกิดทักษะการเรียนรู้และการคิดขั้นสูง จนสามารถนำไปประยุกต์ใช้ในชีวิตจริงได้

ขั้นตอนในการคิดวิจารณ์ญาณทางคณิตศาสตร์

ทักษะการคิดขั้นสูงส่วนใหญ่จะนำทักษะการแก้ปัญหามาใช้ เนื่องจากทักษะนี้สามารถประยุกต์ใช้ได้หลากหลายวิชา ไม่ว่าจะเป็นวิทยาศาสตร์ หรือคณิตศาสตร์ แต่กลับพบว่ายังไม่เคยมีการนำทักษะการคิดวิจารณ์ญาณมาพัฒนากระบวนการคิดของผู้เรียนในรายวิชาคณิตศาสตร์ในประเทศไทยทั้งที่ทักษะการคิดวิจารณ์ญาณนี้เป็นกระบวนการคิดพิจารณาไตร่ตรองข้อมูล และหลักฐานที่มีอยู่อย่างรอบคอบ ประยุกต์ใช้ทักษะการคิดขั้นสูงในหลากหลายด้าน และเป็นทักษะพื้นฐานที่สำคัญในการนำไปต่อยอดสู่สมรรถนะอื่น ๆ ได้ จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง พบว่า มีนักการศึกษาหลายคนได้เสนอขั้นตอนการคิดวิจารณ์ญาณทางคณิตศาสตร์ไว้ อย่างหลากหลาย เมื่อทำการสังเคราะห์แนวคิดและขั้นตอนในการคิดวิจารณ์ญาณจากนักการศึกษา แล้วพบว่า มีนักการศึกษาที่เสนอขั้นตอนในการคิดวิจารณ์ญาณทางคณิตศาสตร์ที่ตรงกัน ดังนี้ (Dressel & Mayhew, 1957; Decaroli, 1973; Gagne, 1985; Ennis, 1987; Woolfolk, 1995; Facione, P. A., 1992; Watson, G. 1980; วรณีพิฏ พันธ์หนองหว้า, 2559) 1. ขั้นตอนิยามปัญหา 2. ขั้นตอนตั้งสมมติฐาน 3. ขั้นตอนรวบรวมข้อมูล 4. ขั้นตอนวิเคราะห์ข้อมูล และ 5. ขั้นตอนสร้างข้อสรุป โดยแต่ละขั้นตอนมีรายละเอียดดังนี้

- 1) **ขั้นนิยามปัญหา** หมายถึง การกำหนดประเด็นปัญหาจากการพิจารณาข้อมูลหรือข้อโต้แย้ง ภายใต้ข้อตกลงเบื้องต้นในสถานการณ์ที่กำหนดให้ โดยปราศจากการให้ข้อมูลอื่นที่นอกเหนือจากบริบทของสถานการณ์หรือความคิดเห็นส่วนตัวเพิ่มเข้าไป
- 2) **ขั้นตั้งสมมติฐาน** หมายถึง การสร้างแนวทางคำตอบที่เป็นไปได้มากที่สุดหรือคาดคะเนแนวทาง ในการหาคำตอบ โดยอาศัยข้อมูลในสถานการณ์ที่มีอยู่เดิมประกอบการตัดสินใจอย่างสมเหตุสมผล
- 3) **ขั้นรวบรวมข้อมูล** หมายถึง การคัดเลือกข้อมูลที่มีความสัมพันธ์กับประเด็นปัญหา ข้อโต้แย้ง หรือข้อมูลที่คลุมเครือ โดยดึงข้อมูลหรือความรู้จากประสบการณ์เดิมมาใช้ เพื่อแสวงหาข้อมูลที่ถูกต้องและชัดเจนมากยิ่งขึ้น
- 4) **ขั้นวิเคราะห์ข้อมูล** หมายถึง การนำข้อมูลที่ผ่านการรวบรวมจากการพิจารณาความน่าเชื่อถือและ ความเพียงพอของข้อมูล มาเชื่อมโยงหาความสัมพันธ์โดยใช้ความรู้ทางคณิตศาสตร์ เพื่อนำไปสู่การสร้างข้อสรุป
- 5) **ขั้นสร้างข้อสรุป** หมายถึง การนำผลการวิเคราะห์ข้อมูล และหลักฐานที่มีความหนักแน่นน่าเชื่อถือ มาใช้ในการตัดสินใจสรุปผลการทดสอบสมมติฐานอย่างเหมาะสมกับสถานการณ์ โดยตั้งอยู่บนตรรกะเหตุและผลทางคณิตศาสตร์

สำหรับการคิดวิจารณ์ญาณทางคณิตศาสตร์เป็นทักษะการคิดขั้นสูง โดยส่วนใหญ่เครื่องมือในการประเมินการคิดขั้นสูงนิยมใช้แบบสอบความเรียงหรือแบบสอบอัตนัย (essay question) ที่มักจะมีกาให้บทความยาวมาในครั้งเดียว ไม่มีตัวเลือก ผู้สร้างแบบสอบจะเขียนข้อความคำถามแล้วให้ผู้สอบเขียนตอบด้วยตัวเอง แม้ว่าข้อสอบอัตนัยจะวัดความคิดขั้นสูงแต่ก็ยังไม่บรรลุวัตถุประสงค์เนื่องจากผู้สร้างข้อสอบตั้งเป้าหมายในการตรวจให้คะแนนได้ง่ายเป็นสำคัญ ทำให้แบบสอบอัตนัยนี้มักวัดความรู้ที่อยู่ในระดับท่องจำหรือเข้าใจพื้นฐานเท่านั้น และมีความเป็นปรนัยในการตรวจให้คะแนนค่อนข้างต่ำ จากข้อจำกัดของแบบสอบอัตนัยแบบดั้งเดิมจึงได้รับการพัฒนาเป็นแบบสอบอัตนัยประยุกต์ (Modified Essay Question test: MEQ)

แบบสอบอัตนัยประยุกต์ หรือแบบสอบเอ็มอีคิว (MEQ) เป็นแบบสอบความเรียงหรือแบบสอบอัตนัย (essay question) ที่มีลักษณะคำถามเป็นปลายเปิด (open-ended question) โดยบรรยายเหตุการณ์ที่เกิดขึ้นจริงตามลำดับเหตุการณ์ที่ละขั้นตอน จะไม่นำเสนอเหตุการณ์ทั้งหมดเพียงครั้งเดียวต่อเนื่องเหมือนแบบสอบความเรียง ผู้ตอบต้องอาศัยข้อมูลจากเหตุการณ์นั้นมาหาคำตอบ (นภา หลิมรัตน์, 2551) ซึ่งแบบสอบอัตนัยประยุกต์ถูกสร้างขึ้นเพื่อมาใช้แทนแบบสอบอัตนัยแบบดั้งเดิมที่มีความเป็นปรนัยในการตรวจให้คะแนนค่อนข้างต่ำ และวัดความรู้ที่อยู่ในระดับ

ท่องจำหรือเข้าใจพื้นฐานเท่านั้น แบบสอบอัตนัยประยุกต์พัฒนาขึ้นโดย Hodgkin และ Knox คณะกรรมการตรวจสอบแห่งราชวิทยาลัยเวชปฏิบัติทั่วไป สหราชอาณาจักร ในปี ค.ศ. 1965 เพื่อเป็นเครื่องมือวัดทักษะการแก้ปัญหา (problem solving skill) ซึ่งจัดเป็นทักษะการคิดขั้นสูง (higher order cognitive thinking) ของแพทย์ที่จบการศึกษาในระดับปริญญาตรีที่ต้องการสอบเข้าเป็นสมาชิกของสมาคม ต่อมา Feletti (1980) แห่งมหาวิทยาลัยนิวคาสเซิล ได้ปรับปรุงแบบสอบอัตนัยประยุกต์โดยเพิ่มการกำหนดเวลาในการทำข้อสอบแต่ละข้อ และผู้สอบต้องทำข้อสอบทุกข้อให้เสร็จภายในเวลาที่กำหนด นอกจากนี้ Feletti ยังสร้างคำตอบเพื่อใช้ในการให้คะแนนและกำหนดคะแนนผ่านหรือระดับความสามารถขั้นต่ำที่ผ่านด้วย (Feletti, 1980; กนกวรรณ ศรีรักษา, 2554; สุโรยา หมั่นหมัด, 2549 อ้างถึงใน มาฆพันธ์ อำนาคิล และกมลวรรณ ตังธนกานนท์; 2557 และสุภชิต ผดุงผล และกมลวรรณ ตังธนกานนท์, 2559; เชิดศักดิ์ ไอรณรัตน์, 2558)

หลักการสร้างแบบสอบอัตนัยประยุกต์

สำหรับขั้นตอนการสร้างแบบสอบอัตนัยประยุกต์โดยภาพรวม มีขั้นตอนในการสร้างและพัฒนาเหมือนกับแบบสอบโดยทั่วไป แตกต่างกันที่รายละเอียดบางขั้นตอน จากการศึกษางานวิจัย สุโรยา หมั่นหมัด (2549) มีขั้นตอนในการสร้างแบบสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ดังนี้

1. กำหนดจุดมุ่งหมายในการพัฒนาแบบสอบ
2. ศึกษาทฤษฎี หลักสูตร และเอกสารงานวิจัยที่เกี่ยวข้อง เพื่อวิเคราะห์ขอบเขตของเนื้อหา ในการพัฒนา โดยศึกษาตามหลักสูตรการศึกษาขั้นพื้นฐาน มาตรฐานการเรียนรู้ และตัวชี้วัดในกลุ่มสาระวิทยาศาสตร์ พร้อมทั้งศึกษากระบวนการแก้ปัญหาทางวิทยาศาสตร์
3. กำหนดกรอบปัญหา ตามขอบเขตของเนื้อหา และวิธีการแก้ปัญหาทางวิทยาศาสตร์
4. สร้างเครื่องมือตามกรอบปัญหา โดยใช้แบบสอบอัตนัยประยุกต์เป็นเครื่องมือในการวัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์
5. ตรวจสอบคุณภาพขั้นต้นของแบบสอบอัตนัยประยุกต์และเกณฑ์การตรวจให้คะแนน โดยให้ ผู้เชี่ยวชาญที่มีประสบการณ์ทางการสอนวิทยาศาสตร์อย่างน้อย 5 ปี และผู้เชี่ยวชาญที่มีประสบการณ์ทางการวัดผลทางการศึกษา เป็นผู้ตรวจสอบความตรงของสถานการณ์และข้อคำถาม
6. ทดลองใช้แบบทดสอบ (try out) และตรวจสอบคุณภาพหลังการทดลอง โดยหาคุณภาพในด้านความยากง่าย ค่าอำนาจจำแนก ความตรง และความเที่ยงของแบบทดสอบ

7. จัดทำคู่มือการใช้แบบทดสอบ และจัดพิมพ์แบบทดสอบเป็นรูปเล่ม

นอกจากนี้ พวงแก้ว ปุณยกนก (2531) อ้างถึงใน สุภชิต ผดุงผล และกมลวรรณ ตั้งธนาภานนท์ (2559) เสนอขั้นตอนการสร้างโมเดลคำตอบ โดยนำแบบสอบที่สร้างขึ้นไปให้ผู้ทรงคุณวุฒิ หรือผู้เชี่ยวชาญตรวจสอบความถูกต้องและทดลองทำข้อสอบ เพื่อนำคำตอบมาสร้างเป็นโมเดลคำตอบทั้งหมด โดยจะเลือกคำตอบที่ผู้ทรงคุณวุฒิเห็นตรงกันเกินครึ่งหนึ่งมาเป็นคำตอบในการให้คะแนน ส่วนคำตอบที่ผู้เชี่ยวชาญมีความเห็นไม่ตรงกันจะนำมาอภิปรายเพื่อแก้ไขปรับปรุงหรือตัดทิ้งไป และหาเวลาที่เหมาะสมในการทำข้อสอบแต่ละข้อ

จากแนวคิดของนักวิชาการที่กล่าวไว้ข้างต้น การสร้างแบบสอบอัตโนมัติจึงเป็นขั้นตอนที่ส่งผลต่อความน่าเชื่อถือของแบบสอบหากมีขั้นตอนในการสร้างที่ถูกต้อง ผู้วิจัยจึงได้กำหนดขั้นตอนการสร้างแบบสอบอัตโนมัติ ดังนี้ 1) กำหนดจุดมุ่งหมาย 2) ศึกษาเอกสารที่เกี่ยวข้อง 3) กำหนดขอบเขตของแบบสอบ 4) เลือกเหตุการณ์ที่จะนำมาสร้างสถานการณ์และข้อคำถาม 5) กำหนดรูปแบบของแบบสอบอัตโนมัติในการวัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ 6) สร้างโมเดลคำตอบ เกณฑ์การให้คะแนน การตรวจให้คะแนน และการบันทึกคะแนน 7) ตรวจสอบความตรง 8) ทดลองใช้ (try out) 9) คัดเลือกข้อสอบที่อยู่ในเกณฑ์ และปรับปรุงแก้ไข 10) นำแบบสอบอัตโนมัติที่พัฒนาแล้วไปใช้จริง 11) วิเคราะห์ข้อสอบ และหาคุณภาพ 12) สร้างเกณฑ์ และ 13) สร้างคู่มือการใช้ และจัดทำรูปเล่ม

ส่วนที่ 2
แนวปฏิบัติและนำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ไปใช้
สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

รูปแบบของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์
สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

สำหรับรูปแบบของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ แบ่งแบบสอบออกเป็น 3 ส่วน ดังนี้ ส่วนที่ 1 คือ ส่วนด้านบนของกระดาษ ซึ่งแบ่งออกเป็น 3 ช่องย่อย ช่องซ้าย คือ เวลาที่ใช้ในการทำเหตุการณ์นี้ ช่องกลาง คือ เลขที่ของเหตุการณ์ในแบบสอบและข้อความคำถามในเหตุการณ์นั้น ช่องขวา คือ ส่วนที่นักเรียนกรอกชื่อ นามสกุล ชั้น และเลขที่ ส่วนที่ 2 คือ ส่วนที่อยู่ถัดลงมาจากส่วนที่ 1 เป็นส่วนที่ใช้แสดงสถานการณ์ ข้อคำถาม และวันที่ไว้สำหรับให้ตอบคำถาม และส่วนที่ 3 คือ ส่วนที่แสดงคำเตือนในการทำแบบสอบอัตนัยประยุกต์ ดังตัวอย่างต่อไปนี้

ตัวอย่างแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

เวลาสำหรับทำ เหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ ... ข้อคำถามที่ ...	ชื่อ-สกุล ชั้น เลขที่
สถานการณ์		
ข้อคำถามข้อที่ ...		
คำถาม :		
คำตอบ		
.....		
.....		
.....		
ข้อควรระวัง ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก		

การจัดพิมพ์แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

1. พิมพ์เหตุการณ์ลงในกระดาษขนาด A4 โดยในแต่ละเหตุการณ์จะพิมพ์ข้อความถามละ 1 แผ่น และเพื่อไม่ให้ผู้สอบเกิดความสับสนในแต่ละเหตุการณ์ ควรพิมพ์ชื่อเหตุการณ์คั่นระหว่างเหตุการณ์หนึ่ง ๆ
2. พิมพ์คำแนะนำในการทำแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ โดยมีรายละเอียดดังนี้

คำแนะนำในการทำแบบสอบอัตนัยประยุกต์ วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

1. ข้อสอบนี้มีทั้งหมด 3 เหตุการณ์ โดยแต่ละเหตุการณ์จะมีข้อความถาม จำนวน 5 ข้อ คำถาม รวมข้อความถามทั้งสิ้น จำนวน 15 ข้อ รวม 15 หน้า
2. มีเวลาในการทำแบบสอบทั้งหมด 1 ชั่วโมง โดยแบ่งเวลาในการทำแต่ละเหตุการณ์ 20 นาที ซึ่งอยู่รวมบนซ้ายของแบบสอบอัตนัยประยุกต์ ผู้สอบต้องจัดสรรเวลาในการทำแบบสอบอัตนัยประยุกต์เอง และทำข้อสอบทุกข้อให้เสร็จทันในเวลาที่กำหนด
3. นักเรียนต้องกรอกชื่อ นามสกุล ชั้น และเลขที่ ที่มุมบนขวาของกระดาษคำถามทุกแผ่น
4. เริ่มทำข้อสอบตั้งแต่เหตุการณ์ที่ 1 ข้อคำถามที่ 1 เมื่อทำเสร็จให้นักเรียนตรวจสอบให้เรียบร้อย แล้วส่งข้อสอบโดยคว่ำกระดาษลง วางไว้ใต้ช่องใส่แบบสอบอัตนัยประยุกต์ และไม่ต้องนำข้อสอบกลับมาแก้ไขหรือเปิดดูข้อคำถามถัดไป

3. การเรียงหน้าของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

เรียงหน้าของแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ โดยเริ่มต้นที่คำแนะนำในการทำแบบสอบอัตนัยประยุกต์ จากนั้นจึงเรียงเหตุการณ์ของแบบสอบจากเหตุการณ์ที่ 1 จนถึงเหตุการณ์ที่ 3

4. การบรรจุแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

บรรจุแบบสอบอัตนัยประยุกต์ลงในซองสีน้ำตาลที่มีขนาดของซองเท่ากับกระดาษ A5 เนื่องจากต้องการอำนวยความสะดวกให้กับนักเรียนเวลาดึงแบบสอบออกจากซองข้อสอบ และพิมพ์ชื่อแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ พร้อมทั้งเลขที่ตามลำดับที่หน้าซองของแบบสอบทุกซอง

ขั้นตอนการดำเนินการสอบ

1. ผู้คุมสอบแจกแบบสอบอัตนัยประยุกต์ไว้บนโต๊ะสอบทุกโต๊ะ โต๊ะละ 1 ชุด เมื่ออนุญาตให้นักเรียนเข้าประจำที่สอบและนั่งเรียบร้อยแล้ว ผู้คุมสอบอนุญาตให้นักเรียนเปิดซองข้อสอบได้

2. ผู้คุมสอบให้นักเรียนทุกคนอ่านคำแนะนำในการทำแบบสอบอัตนัยประยุกต์ พร้อมทั้งให้เวลาเขียนชื่อ นามสกุล และเลขที่ ที่ข้อความถามแผ่นแรกให้เรียบร้อย

3. ผู้คุมสอบอธิบายย้ำถึงวิธีการทำแบบสอบให้นักเรียนฟัง โดยให้นักเรียนทำแบบสอบที่ละข้อคำถามและต้องทำข้อคำถามนั้นให้เสร็จ จึงจะสามารถเปลี่ยนไปทำข้อคำถามต่อไปได้ สำหรับเวลาที่ใช้ในการทำแบบสอบ คือ เหตุการณ์ละ 20 นาที รวมทั้งฉบับ 1 ชั่วโมง โดยให้นักเรียนบริหารจัดการเวลาสอบเองและเน้นย้ำว่า เมื่อนักเรียนทำข้อคำถามในแต่ละหน้าเสร็จเรียบร้อยแล้ว ให้ตรวจสอบความเรียบร้อยก่อนแล้วจึงส่งข้อสอบ โดยคว่ำกระดาษลงและวางไว้ในช่องใส่แบบสอบอัตนัยประยุกต์ และบอกให้นักเรียนเริ่มทำแบบสอบ อัตนัยประยุกต์ ผู้คุมสอบบันทึกเวลาในการทำข้อสอบและเวลาสิ้นสุดในการทำแบบสอบอัตนัยประยุกต์

สำหรับแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จะปรากฏอยู่ในภาคผนวก ค

ส่วนที่ 3

การตรวจให้คะแนนและการบันทึกคะแนน

วิธีตรวจให้คะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่พัฒนาขึ้น คือ วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) โดยแบ่งออกเป็น 2 ระดับชั้น ชั้นที่ 1 ประเมินตามข้อรายการย่อย ดังนี้

สำหรับข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 วัด 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา และการเข้าใจในสิ่งที่ต้องการวัด สำหรับข้อคำถามที่ 4 และข้อคำถามที่ 5 วัด 3 องค์ประกอบ คือ ความถูกต้องของเนื้อหา การเข้าใจในสิ่งที่ต้องการวัด และการสื่อความหมายทาง โดยในแต่ละข้อรายการย่อยใช้มาตราประมาณค่า (rating scale) 3 ระดับ ได้แก่ 2 คะแนน คือ ถูกต้องทั้งหมด, 1 คะแนน คือ ถูกต้องบางส่วน และ 0 คะแนน คือ ไม่ถูกต้องทั้งหมด หรือไม่มีการตอบคำถาม จากนั้นนำผลรวมของคะแนนในระดับชั้นที่ 1 มาแปลงเป็นคะแนนในชั้นที่ 2 ตามเกณฑ์ที่กำหนด ซึ่งวิธีตรวจให้คะแนนรูปริกสองชั้นแต่ละชุดจะใช้ตรวจข้อคำถามแต่ละข้อเท่านั้น ตัวอย่างเช่น ข้อคำถามที่ 1 ถึงข้อคำถามที่ 5 ของเหตุการณ์ที่ 1 โดยใช้วิธีตรวจให้คะแนนรูปริกสองชั้น ดังนี้

วิธีตรวจให้คะแนนแบบรูปริกสองชั้นของข้อคำถามที่ 1 ในเหตุการณ์ที่ 1

ระดับชั้นที่ 1

จุฬาลงกรณ์มหาวิทยาลัย

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย		1	0
ชั้นที่ 1 นิยาม ปัญหา	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบท ของสถานการณ์	1.1 ระบุข้อมูลที่ สถานการณ์กำหนดได้ ถูกต้อง คือ สกุลเงินบาทเป็นสกุลเงิน ตั้งต้นที่นำไปแลก	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		1.2 อธิบายเงื่อนไขของ สถานการณ์ได้ถูกต้อง คือ แลกเป็นเงินสกุล อื่น ๆ ในมูลค่าเท่ากัน	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			ถูกต้องทั้งหมด	ถูกต้อง บางส่วน	ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ คำถาม
			ถูกต้องทั้งหมด	ถูกต้อง บางส่วน	ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ

					คำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดปัญหาภายใต้บริบทของสถานการณ์	2.1 กำหนดปัญหาได้ถูกต้อง คือ ปัญหาที่เกี่ยวข้องกับการเตรียมเงินบาทไปแลกเป็นสกุลเงินอื่น ๆ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการกำหนดปัญหาอยู่ภายใต้ข้อมูลในสถานการณ์ คือ สกุลเงินที่นำไปแลกได้แก่ สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
คะแนนรวม = คะแนน คิดเป็น = คะแนน					

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-8	2
4-6	1
0-3	0

วิธีตรวจให้คะแนนแบบรูปริกสองชั้นของข้อคำถามที่ 2 ในเหตุการณ์ที่ 1

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0	
<p>ขั้นที่ 2 ตั้งสมมติฐาน</p>	<p>1. ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ ต้องเตรียมสกุลเงินบาทไปแลกเปลี่ยนในปริมาณที่ไม่เท่ากัน</p>	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม	
		<p>1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ สกุลเงินที่ต้องการแลกเปลี่ยนมีมูลค่าเท่ากัน</p>	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม	
	<p>2. การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- กำหนดสมมติฐานภายใต้บริบทของสถานการณ์</p>	<p>2.1 กำหนดสมมติฐานได้ถูกต้อง คือ คาดคะเนปริมาณเงินบาทที่นำไปแลกเปลี่ยนได้ว่าต้องเตรียมไป ในปริมาณที่ไม่เท่ากัน</p>	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม	
		<p>2.2 บริบทของการกำหนดสมมติฐานอยู่ภายใต้ข้อมูลในสถานการณ์ คือ สมมติฐานที่เกี่ยวข้องกับการกำหนดปริมาณเงินบาทไปแลกเปลี่ยนสกุลเงินอื่น ๆ</p>	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มี การตอบคำถาม	
	<p>คะแนนรวม = คะแนน คิดเป็น = คะแนน</p>					

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูปริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-8	2
4-6	1
0-3	0

วิธีตรวจให้คะแนนแบบรูปริกสองชั้นของข้อคำถามที่ 3 ในเหตุการณ์ที่ 1

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 3 รวบรวม ข้อมูล	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบท ของสถานการณ์	1.1 ระบุข้อมูลที่ สถานการณ์กำหนดได้ ถูกต้อง คือ แนวทางการรวบรวม ข้อมูลที่เกี่ยวข้องกับอัตรา การแลกเปลี่ยนสกุลเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ คำถาม
		1.2 อธิบายเงื่อนไขของ สถานการณ์ได้ถูกต้อง คือ เหตุผลในการใช้แนวทางการ เก็บรวบรวมข้อมูล	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ คำถาม
	2. การเข้าใจใน สิ่งที่ต้องการวัด - กำหนด แนวทางการรวบรวม ข้อมูลภายใต้ บริบทของ สถานการณ์	2.1 กำหนดแนวทางการ รวบรวมข้อมูลได้ถูกต้อง คือ แนวทางการรวบรวมข้อมูลที่ เกี่ยวข้องกับอัตรา การแลกเปลี่ยนเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ คำถาม
		2.2 บริบทของแนวทางการ เก็บรวบรวมข้อมูลอยู่ ภายใต้ข้อมูลใน สถานการณ์ คือ การ แลกเปลี่ยนเงินตราระหว่าง	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		สกุลเงิน			ไม่มีการตอบ คำถาม
คะแนนรวม =			คะแนน	คิดเป็น =	คะแนน

ระดับขั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-8	2
4-6	1
0-3	0

วิธีตรวจให้คะแนนแบบรูบริกสองชั้นของข้อคำถามที่ 4 ในเหตุการณ์ที่ 1

ระดับขั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 4 วิเคราะห์ ข้อมูล	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบท ของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์ กำหนดได้ถูกต้อง คือ อัตราแลกเปลี่ยนแต่ละคู่ สกุลเงินตามที่สถานการณ์ กำหนด	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
		1.2 อธิบายเงื่อนไขของ สถานการณ์ได้ถูกต้อง คือ เปรียบเทียบ อัตรา แลกเปลี่ยนค่าเงินครบ ทุกสกุลเงิน (สกุลบาท สกุลเยน สกุลเงิน และสกุล ดอลลาร์สหรัฐ)	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
	2. การเข้าใจใน สิ่งที่ต้องการวัด - วิเคราะห์ข้อมูล ภายใต้บริบท ของสถานการณ์	2.1 วิเคราะห์ข้อมูลได้ ถูกต้อง คือ เปรียบเทียบ อัตราส่วนค่าเงินแลกเปลี่ยนทุก สกุลเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
		2.2 บริบทของการวิเคราะห์ ข้อมูลอยู่ภายใต้ข้อมูลใน สถานการณ์ คือ	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		สกุลเงินที่นำมาเปรียบเทียบ ประกอบด้วย สกุลเงินบาท สกุลเยน สกุลเงินหยวน และสกุลดอลลาร์สหรัฐ	ถูกต้องทั้งหมด	ถูกต้องบางส่วน	ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
	3. การสื่อ ความหมายทาง คณิตศาสตร์ - เปลี่ยนประโยค ภาษา เป็น ประโยคทาง คณิตศาสตร์ และคำนวณได้ ถูกต้อง	3.1 เปลี่ยนประโยคภาษา เป็น ประโยคทาง คณิตศาสตร์ได้ถูกต้อง คือ เขียนสัญลักษณ์อัตราส่วนใน การเปรียบเทียบ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
		3.2 คำนวนได้ถูกต้อง คือ ทำเป็นอัตราส่วนอย่างต่ำ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
คะแนนรวม = คะแนน			คิดเป็น = คะแนน		

ระดับขั้นที่ 2 (คะแนนเต็ม 3 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0

วิธีตรวจให้คะแนนแบบรูปริกสองชั้นของข้อคำถามที่ 5 ในเหตุการณ์ที่ 1

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 5 สร้าง ข้อสรุป	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ เตรียมเงินสกุลบาทไปแลกเงินสกุลต่าง ๆ ได้แก่ สกุลเยน สกุลเงินหยวน และสกุลดอลลาร์สหรัฐ สกุลละ 210 หน่วย	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ อัตราการแลกเปลี่ยนเงินสกุลต่าง ๆ มาจากผลการทดสอบสมมติฐานของสถานการณ์ ดังนี้ บาท : เยน : หยวน : ดอลลาร์สหรัฐ เท่ากับ 100 : 350 : 20 : 3	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดข้อสรุปการเตรียมเงินสกุลบาทไปแลก และสรุปผลสมมติฐานสอดคล้องกับบริบทของสถานการณ์	2.1 เขียนข้อสรุปของการแลกเงิน และผลสมมติฐาน คือ 1) เตรียมเงินสกุลบาทไปเป็นจำนวน 8,110 บาท และ 2) สรุปสมมติฐานว่า การนำสกุลเงินบาทไปแลกเป็นเงินสกุลอื่น ๆ ตามที่ต้องการในมูลค่าเท่า ๆ กัน จะต้องเตรียมเงินบาทเป็นจำนวนไม่เท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการสร้างข้อสรุปอยู่ภายใต้ข้อมูลในสถานการณ์ คือ สกุลเงินที่นำมาเปรียบเทียบกับสกุลเงินบาทประกอบด้วย สกุลเงินเยน สกุลเงินหยวน และสกุลดอลลาร์สหรัฐ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		3.1 เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ได้ถูกต้อง คือ เขียนสัดส่วนในแต่ละคู่การแลกเปลี่ยนสกุลเงิน (บาท-เยน, บาท-หยวน และบาท-ดอลลาร์สหรัฐ)	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	3.2 คำนวนได้ถูกต้อง คือ คำนวน	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
	ทางคณิตศาสตร์ และคำนวณได้ ถูกต้อง	อัตรา การแลกเปลี่ยนแต่ละคู่สกุลบาทกับ สกุลเงินอื่น ๆ ได้อย่างถูกต้อง คือ เตรียม เงินบาทไปแลกเปลี่ยนเงินสกุลเยน 60 บาท, เตรียมเงินบาทไปแลกเปลี่ยน เป็นเงินสกุลหยวน 1,050 บาท และเตรียมเงินบาทไปแลกเปลี่ยน เงินสกุลดอลลาร์สหรัฐ 7,000 บาท	ถูกต้องทั้งหมด	ถูกต้องบางส่วน	ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
คะแนนรวม = คะแนน			คิดเป็น = คะแนน		

ระดับขั้นที่ 2 (คะแนนเต็ม 3 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0

ขั้นตอนในการตรวจให้คะแนนแบบสอบ

1. ให้ตรวจคำตอบที่ละเอียดของนักเรียนทุกคนและควรตรวจให้เสร็จทีละคน นั่นคือ ถ้าจะตรวจคำตอบ ข้อคำถามที่ 1 ของเหตุการณ์ที่ 1 ก็ให้ตรวจข้อคำถามที่ 1 ของเหตุการณ์ที่ 1 ของนักเรียนทุกคนจบครบ แล้วจึงเริ่มตรวจข้อคำถามที่ 2 ของเหตุการณ์ที่ 1 ครูไม่ควรตรวจข้อสอบให้เสร็จเป็นคน ๆ

2. ให้มีความคงเส้นคงวาในการตรวจ เกณฑ์ในการให้คะแนนจะเหมือนกัน สำหรับทุก ๆ คำตอบ ครูบางคนอาจจะให้คะแนนค่อนข้างน้อย สำหรับคำตอบแผ่นแรก แต่จะให้คะแนนมากขึ้น สำหรับคำตอบแผ่นหลัง ทั้ง ๆ ที่คำตอบเหล่านั้นเป็นไปในแนวทางเดียวกัน

3. เวลาตรวจไม่ควรดูชื่อนักเรียน เพื่อป้องกันไม่ให้เกิดอคติต่อการให้คะแนน และไม่ควรนำลายมือของผู้ตอบมามีส่วนในการให้คะแนน

4. พยายามตรวจข้อคำถามใดข้อคำถามหนึ่งของนักเรียนทุกครั้งให้เสร็จ ไม่ควรหยุดพักจนกว่าจะตรวจข้อคำถามนั้นเสร็จแล้ว

5. พยายามเขียนข้อแนะนำและแก้ไขความผิดในกระดาษคำตอบ เพื่อให้ผู้ตอบรู้ว่าตนเองบกพร่องตรงไหน

สำหรับแนวคำตอบวิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) ของแบบสอบอัตรัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จะปรากฏในภาคผนวก ง

การบันทึกคะแนน

เมื่อผู้ตรวจให้คะแนนจะต้องบันทึกคะแนนลงในแบบบันทึกคะแนน โดยคะแนนที่บันทึกจะแบ่งออกเป็นดังนี้

สำหรับข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 วัด 2 องค์ประกอบ คือ 1) ความถูกต้องของเนื้อหา (4 คะแนน) แบ่งออกเป็น 1.1) ระบุข้อมูลที่สถานการณ์กำหนด (2 คะแนน) 1.2) อธิบายเงื่อนไขของสถานการณ์ (2 คะแนน) และ 2) การเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) แบ่งออกเป็น 2.1) กำหนดขั้นของการวัด (2 คะแนน) 2.2) บริบทขั้นของการวัดอยู่ภายใต้ข้อมูลในสถานการณ์ (2 คะแนน) โดยในแต่ละข้อรายการย่อยใช้มาตราประมาณค่า (rating scale) 3 ระดับ ได้แก่ 2 คะแนน คือ ถูกต้องทั้งหมด, 1 คะแนน คือ ถูกต้องบางส่วน และ 0 คะแนน คือ ไม่ถูกต้องทั้งหมด หรือไม่มีการตอบคำถาม จากนั้นนำผลรวมของคะแนนในระดับขั้นที่ 1 มาแปลงเป็นคะแนนในขั้นที่ 2 คะแนนเต็มข้อละ 2 คะแนน

สำหรับข้อคำถามที่ 4 และข้อคำถามที่ 5 วัด 3 องค์ประกอบ คือ 1) ความถูกต้องของเนื้อหา (4 คะแนน) 2) การเข้าใจในสิ่งที่ต้องการวัด (4 คะแนน) และ 3) การสื่อความหมายทางคณิตศาสตร์ (4 คะแนน) แบ่งออกเป็น 3.1) เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ (2 คะแนน) 3.2) การคำนวณ (2 คะแนน) โดยในแต่ละข้อรายการย่อยใช้มาตราประมาณค่า (rating scale) 3 ระดับ ได้แก่ 2 คะแนน คือ ถูกต้องทั้งหมด, 1 คะแนน คือ ถูกต้องบางส่วน และ 0 คะแนน คือ ไม่ถูกต้องทั้งหมด หรือไม่มีการตอบคำถาม จากนั้นนำผลรวมของคะแนนในระดับชั้นที่ 1 มาแปลงเป็นคะแนนในชั้นที่ 2 คะแนนเต็ม ข้อละ 3 คะแนน คะแนนเต็มเหตุการณ์ละ 12 คะแนน รวมทั้งฉบับ 36 คะแนน สำหรับแบบบันทึกคะแนนจะปรากฏอยู่ในภาคผนวก จ



บรรณานุกรม

ภาษาไทย

- กติกิร กมลรัตน์นะสมบัติ และกมลวรรณ ตั้งธนกานนท์. (2559). *ผลของการให้ข้อมูลย้อนกลับจากแบบสอบอัตโนมัติประยุกต์ที่มีต่อพัฒนาการความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษา [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย]*. Chulalongkorn University Intellectual Repository (CUIR).
<http://cuir.car.chula.ac.th/handle/123456789/50593>
- กนกวรรณ ศรีรักษา. (2554). *Modified Essay Question (MEQ)*. ขอนแก่นเวชสาร.
- กมลวรรณ ตั้งธนกานนท์. (2563). *การวัดและประเมินทักษะการปฏิบัติ (พิมพ์ครั้งที่ 3)*. สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- กลุ่มนิเทศ ติดตาม และประเมินผลการจัดการศึกษา สพป.นครปฐม เขต 2. (2564, 11 กุมภาพันธ์). *หลักฐตฐฐานสมรรถนะ*. <http://chaisri-nites-hi-supervisory5.net/home/smrrthna>
- กิตติพันธ์ วิบูลศิลป์ และจินดิษฐ์ ละออบปักษิน. (2560). *ผลการจัดกิจกรรมการเรียนรู้คณิตศาสตร์ตามแนวคิดห้องเรียนย้อนกลับทางร่วมกับการเรียนรู้เชิงรุกที่มีต่อความสามารถในการแก้ปัญหาคณิตศาสตร์และความสามารถในการคิดอย่างมีวิจารณญาณของนักเรียนชั้นมัธยมศึกษาปีที่ 5 [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย]*. Chulalongkorn University Intellectual Repository (CUIR).
<http://cuir.car.chula.ac.th/handle/123456789/58427>
- จิรายุ เถาว์โท และคณะ. (2559). *การศึกษาค่าความเชื่อมั่นของคะแนนแบบทดสอบอัตโนมัติคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีจำนวนผู้ตรวจและรูปแบบการตรวจให้คะแนนต่างกัน โดยใช้ทฤษฎีการสรุปอ้างอิง. วารสารหาดใหญ่วิชาการ, 14(1), 1-14.*
<https://so01.tci-thaijo.org/index.php/HatyaiAcademicJournal/article/view/82192/65334>
- ชนิสรา สงวนไ้ว และศิริชัย กาญจนวาสี. (2558). *การเปรียบเทียบความเที่ยงของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของ ผลการวัด [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย]*. Chulalongkorn University Intellectual Repository (CUIR).
<http://cuir.car.chula.ac.th/handle/123456789/51081>

- เชิดศักดิ์ ไอรอมณีรัตน์. (2558). การสร้างข้อสอบอัตนัยประยุกต์. *เวชบันทึกศิริราช*, 8(1), 48-49. <https://he02.tci-thaijo.org/index.php/simedbull/article/view/81482>
- โชติกา ภาชีผล. (2559). *การวัดและประเมินผลการเรียนรู้ Learning Measurement and Evaluation*. โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- จิตติมา จิตรุ่งเรือง. (2544). *การพัฒนาแบบทดสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาวิชาแนวคิดพื้นฐานและหลักการพยาบาล* [วิทยานิพนธ์ปริญญาามหาบัณฑิต มหาวิทยาลัยทักษิณ]. TSU Knowledge bank (TSUKB). <http://kb.tsu.ac.th/jspui/bitstream/123456789/1324/1/69745.pdf>
- ณัฐฐณี ศิริโชติ และกมลวรรณ ตั้งธนานนท์. (2559). *ปฏิสัมพันธ์ระหว่างวิธีการประเมินตนเองและความสามารถทางวิทยาศาสตร์ที่มีต่อพัฒนาการทักษะปฏิบัติการทดลองของนักเรียน* [วิทยานิพนธ์ปริญญาามหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย]. Chulalongkorn University Intellectual Repository (CUIR). <http://cuir.car.chula.ac.th/bitstream/123456789/50594/1/5783825127.pdf>
- ณัฐฐภรณ์ หลาวทอง. (2561). *การสร้างเครื่องมือการวิจัยทางการศึกษา* (พิมพ์ครั้งที่ 2). สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ดรุณี อภัยกาวิ, ประกิตติยา ทักษิณ และกมลวรรณ ตั้งธนานนท์. (2562). การศึกษาผลการตรวจให้คะแนนแบบสอบอัตนัยประยุกต์ใช้โมเดลหลายองค์ประกอบของราล์ช และทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด. *วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม*, 26(1), 110-124.
- นงรัศม์ ไทธานี. (2542). *การสร้างแบบสอบเอ็ม อี คิว ที่ดำเนินการสอบโดยใช้คอมพิวเตอร์เพื่อวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์* [วิทยานิพนธ์ปริญญาามหาบัณฑิต มหาวิทยาลัยมหาสารคาม].
- นภา หลิมรัตน์. (2551, ตุลาคม). *การวัดผลและประเมินผลการศึกษา*. PSU Medical Education Resources. <http://teachingresources.psu.ac.th/document/2551/limrat/1.pdf>
- พรพิมล ปิณฑนาภรณ์ และชิตชนก เจริญเชาว์. (2561). *การพัฒนาแบบทดสอบอัตนัยประยุกต์วัดความสามารถในการตัดสินใจทางคลินิกของนักศึกษาแพทย์* มหาวิทยาลัยสงขลานครินทร์. *วารสารมหาวิทยาลัยราชภัฏยะลา*, 15(3), 282-289.
- พวงแก้ว ปุณยณก. (2531). *แบบสอบอัตนัยประยุกต์ (เอ็ม อี คิว) เพื่อใช้วัดทักษะการแก้ปัญหา*. https://doi.nrct.go.th/ListDoi/listDetail?Resolve_DOI=10.14457/CU.res.1988.1

- มาฆพันธ์ อ๋านาคิล และกมลวรรณ ตังธนทานนท์. (2557). การเปรียบเทียบความเที่ยงของแบบ
สอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ที่ใช้วิธีการตรวจให้
คะแนนที่แตกต่าง ภายใต้จำนวนเหตุการณ์ที่ต่างกัน: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง
ความน่าเชื่อถือของผลการวัด [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์
มหาวิทยาลัย]. Chulalongkorn University Intellectual Repository (CUIR).
<http://cuir.car.chula.ac.th/bitstream/123456789/46285/1/5683866827.pdf>
- ลักษมี จันทราช. (2544). การเปรียบเทียบคุณภาพแบบสอบอัตนัยประยุกต์ (MEQ) และ
แบบทดสอบเลือกตอบ (MCQ) เพื่อวัดความสามารถในการแก้ปัญหา [วิทยานิพนธ์
ปริญญาโทมหาบัณฑิต มหาวิทยาลัยทักษิณ]. TSU Knowledge bank (TSUKB).
- วรนิพิฏ์ พันธุ์หนองหว้า และอัมพร ม้าคนอง. (2559). ผลการจัดกิจกรรมการเรียนรู้คณิตศาสตร์
โดยใช้กระบวนการแก้ปัญหาที่มีโครงสร้างไม่สมบูรณ์ร่วมกับกลวิธีการเสริมต่อความคิดที่มี
ต่อความสามารถในการคิดอย่างมีวิจารณญาณ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3
[วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย].
- สำนักงานเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ. (2563). การจัดการเรียนรู้ฐานสมรรถนะ
เชิงรุก. กลุ่มมาตรฐานการศึกษา สำนักงานมาตรฐานการศึกษาและพัฒนาการเรียนรู้.
- สำนักงานเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ. (2563). การจัดการเรียนรู้ฐานสมรรถนะ
เชิงรุก. กลุ่มมาตรฐานการศึกษา สำนักงานมาตรฐานการศึกษาและพัฒนาการเรียนรู้.
- สิริภักดิ์ ศิริโท. (2558). การคิดวิจารณ์: มุมมองการศึกษาและการบริหารจัดการ. ศาสตร์
การคิด รวบรวมความเรียงการคิดและการสอนคิด. มหาวิทยาลัยธุรกิจบัณฑิต.
- สุพัฒตา ภูสอดสี. (2553). การสร้างแบบทดสอบเอ็มอีคิว วัดความสามารถในการแก้ปัญหาทาง
คณิตศาสตร์ เรื่องเรขาคณิตวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 4 [วิทยานิพนธ์ปริญญา
มหาบัณฑิต มหาวิทยาลัยสารคาม].
- สุภชิต ผดุงผล และกมลวรรณ ตังธนทานนท์. (2559). การเปรียบเทียบความเที่ยงของแบบสอบ
อัตนัยประยุกต์สำหรับวัดความสามารถในการใช้วิธีการทางวิทยาศาสตร์ในวิชาฟิสิกส์
ภายใต้จำนวนเหตุการณ์และจำนวนผู้ตรวจที่ต่างกัน: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง
ความน่าเชื่อถือของผลการวัด [วิทยานิพนธ์ปริญญาโทมหาบัณฑิต จุฬาลงกรณ์
มหาวิทยาลัย]. Chulalongkorn University Intellectual Repository (CUIR).
<http://cuir.car.chula.ac.th/handle/123456789/55715>

- สุโรยา หมั่นหมัด. (2549). การพัฒนาแบบทดสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 [วิทยานิพนธ์ปริญญา มหาบัณฑิต มหาวิทยาลัยทักษิณ]. TSU Knowledge bank (TSUKB).
<http://kb.tsu.ac.th/jspui/bitstream/123456789/463/1/102970.pdf>
- สุโรยา หมั่นหมัด. (2549). การพัฒนาแบบสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 [วิทยานิพนธ์ปริญญา มหาบัณฑิต มหาวิทยาลัยทักษิณ]. TSU Knowledge bank (TSUKB).
<http://kb.tsu.ac.th/jspui/bitstream/123456789/463/1/102970.pdf>
- อัมพร ม้าคนอง. (2544). ทักษะและกระบวนการทางคณิตศาสตร์ : การพัฒนาเพื่อพัฒนาการ. สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- อำมาลา สารชาติ. (2548). การพัฒนาแบบสอบ เอ็ม อี คิว สำหรับวัดความสามารถในการแก้ โจทย์ปัญหาทางคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 [วิทยานิพนธ์ปริญญา มหาบัณฑิต มหาวิทยาลัยราชภัฏอุบลราชธานี]. Thai Library Integrated System (ThaiLIS).

ภาษาอังกฤษ

- Decaroli, J. (1973). What Research Say to the Classroom Teacher: Critical Thinking. *Social Education*, 37(1), 67-68.
- Dressel, P. L., & Mayhew, L. B. (1957). *General education explorations in evaluation* (2nd ed.). American Council on Education.
- Ennis, R. H. (1985) Critical thinking and specificity. *Educational Researcher*, 18(3), 4-10.
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. *Teaching for Thinking*, 9-26.
- Facione, P. A. (1992). *Critical Thinking: What It is and Why It Counts*.
https://www.researchgate.net/publication/251303244_Critical_Thinking_What_It_Is_and_Why_It_Counts
- Feletti, G. I. (1980). Reliability and validity studies on modified essay questions. *Academic Medicine*, 55(11), 933-941. <https://doi.org/10.1097/00001888-198011000-00006>

Gagne, R. M. (1985). *The conditions of learning* (4th ed.). New York, NY : Holt, Rinehart & Winston.

Watson, G. (1980). *Watson – Glaser Critical Thinking Appraisal – UK Edition Practice Test*. Pearson Assesment.

Woolfolk, A. E. (1995). *Educational psychology* (6th ed.) Boston, MA: Allyn & Bacon.





ภาคผนวก ค

แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์

สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

แบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์
สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

คำชี้แจง

1. ข้อสอบนี้มีทั้งหมด 3 เหตุการณ์ โดยแต่ละเหตุการณ์จะมีข้อคำถาม จำนวน 5 ข้อคำถาม รวมข้อคำถามทั้งสิ้น จำนวน 15 ข้อ รวม 15 หน้า
2. มีเวลาในการทำแบบสอบทั้งหมด 1 ชั่วโมง โดยแบ่งเวลาในการทำแต่ละเหตุการณ์ 20 นาที ซึ่งอยู่มุมบนซ้ายของแบบสอบอัตนัยประยุกต์ ผู้สอบต้องจัดสรรเวลาในการทำแบบสอบอัตนัยประยุกต์เอง และทำข้อสอบทุกข้อให้เสร็จทันในเวลาที่กำหนด
3. นักเรียนต้องกรอกชื่อ นามสกุล ชั้น และเลขที่ ที่มุมบนขวาของกระดาษคำถามทุกแผ่น
4. เริ่มทำข้อสอบตั้งแต่เหตุการณ์ที่ 1 ข้อคำถามที่ 1 เมื่อทำเสร็จให้นักเรียนตรวจสอบให้เรียบร้อย แล้วส่งข้อสอบโดยคว่ำกระดาษลง วางไว้ใต้ช่องใส่แบบสอบอัตนัยประยุกต์ และไม่อนุญาตให้กลับมาแก้ไขหรือเปิดดูข้อคำถามอีกต่อไป





<p>เวลาสำหรับทำเหตุการณ์นี้ 20 นาที</p>	<p>เหตุการณ์ที่ 1 ข้อคำถามที่ 3</p>	<p>ชื่อ-สกุล ชั้น เลขที่</p>																																																	
<p><u>สถานการณ์</u> จากปัญหา เด็กชายตะวันตั้งคำถามว่า การจะแลกเงินในสกุลต่าง ๆ ให้มีมูลค่าเท่า ๆ กัน ต้องเตรียมเงินบาทไปแลกในปริมาณที่ไม่เท่ากัน</p>																																																			
<p>ข้อคำถามที่ 3 คำถาม : หากเด็กชายตะวันค้นหาข้อมูลมาได้ดังนี้</p> <table border="1" data-bbox="316 907 1370 1429"> <thead> <tr> <th></th> <th>บาท</th> <th>ปอนด์</th> <th>เยน</th> <th>ดอลลาร์ สหรัฐ</th> <th>วอน</th> <th>หยวน</th> </tr> </thead> <tbody> <tr> <th>บาท</th> <td>50</td> <td>1.13</td> <td>175</td> <td>-</td> <td>1800</td> <td>-</td> </tr> <tr> <th>ปอนด์</th> <td>-</td> <td>5</td> <td>770</td> <td>-</td> <td>7950</td> <td>-</td> </tr> <tr> <th>เยน</th> <td>-</td> <td>-</td> <td>350</td> <td>-</td> <td>3620</td> <td>20</td> </tr> <tr> <th>ดอลลาร์ สหรัฐ</th> <td>-</td> <td>5.3</td> <td>-</td> <td>7.5</td> <td>-</td> <td>50</td> </tr> <tr> <th>วอน</th> <td>-</td> <td>-</td> <td>48</td> <td>0.44</td> <td>500</td> <td>-</td> </tr> <tr> <th>หยวน</th> <td>-</td> <td>8.23</td> <td>-</td> <td>-</td> <td>-</td> <td>75</td> </tr> </tbody> </table> <p>ถ้านักเรียนเป็นเด็กชายตะวัน นักเรียนจะทำการคัดเลือกและรวบรวมข้อมูลที่จำเป็นและ เพียงพอเพื่อนำไปสู่การวิเคราะห์ข้อมูล สำหรับการทดสอบสมมติฐานที่ตั้งไว้อย่างไร</p> <p>คำตอบ</p> <p>.....</p> <p>.....</p> <p>.....</p>				บาท	ปอนด์	เยน	ดอลลาร์ สหรัฐ	วอน	หยวน	บาท	50	1.13	175	-	1800	-	ปอนด์	-	5	770	-	7950	-	เยน	-	-	350	-	3620	20	ดอลลาร์ สหรัฐ	-	5.3	-	7.5	-	50	วอน	-	-	48	0.44	500	-	หยวน	-	8.23	-	-	-	75
	บาท	ปอนด์	เยน	ดอลลาร์ สหรัฐ	วอน	หยวน																																													
บาท	50	1.13	175	-	1800	-																																													
ปอนด์	-	5	770	-	7950	-																																													
เยน	-	-	350	-	3620	20																																													
ดอลลาร์ สหรัฐ	-	5.3	-	7.5	-	50																																													
วอน	-	-	48	0.44	500	-																																													
หยวน	-	8.23	-	-	-	75																																													
<p><u>ข้อควรระวัง</u> ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>																																																			
<p>เวลาสำหรับทำเหตุการณ์นี้</p>	<p>เหตุการณ์ที่ 1</p>	<p>ชื่อ-สกุล</p>																																																	

เวลาสำหรับทำ เหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ 1 ข้อคำถามที่ 5	ชื่อ-สกุล ชั้น เลขที่
<p><u>สถานการณ์</u> จากความคิดของเด็กชายตะวันที่ยากเก็บสะสมเงินในสกุลอื่น ๆ บ้างนอกจากสกุลเงินบาท เขาจึงมีข้อสงสัยว่า 'ถ้าต้องการนำสกุลเงินบาทไปแลกเปลี่ยนเป็นสกุลเงินอื่น ๆ ตามที่ต้องการในมูลค่าเท่า ๆ กัน จะต้องใช้เงินบาทเป็นจำนวนเท่ากันหรือไม่' และจากการที่เด็กชายตะวันได้บันทึกผลการวิเคราะห์ข้อมูลดังนี้</p> <p style="text-align: center;">บาท : เยน : หยวน : ดอลลาร์สหรัฐ = 100 : 350 : 20 : 3</p>		
<p>ข้อคำถามที่ 5</p> <p>คำถาม : ถ้าต้องการแลกเปลี่ยนสกุลต่าง ๆ ได้แก่ สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ สกุลละ 210 หน่วย</p> <p>ต้องเตรียมเงินสกุลบาทไปเป็นจำนวนเท่าไร และผลการวิเคราะห์ข้อมูลนี้สอดคล้องกับสมมติฐานที่ตั้งไว้ข้างต้นหรือไม่</p> <p>คำตอบ</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p>		
<p><u>ข้อควรระวัง</u> ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>		



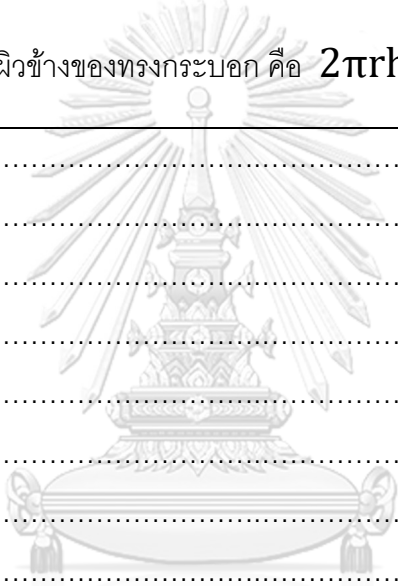
เหตุการณ์ที่ 2

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

เวลาสำหรับทำ เหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ 2 ข้อคำถามที่ 5	ชื่อ-สกุล ชั้น เลขที่
<p>สถานการณ์ จากความคิดของเด็กหญิงมิตตี้สะในตอนที่ว่า ‘สามารถใช้เครื่องมือในการวัดและทฤษฎีบททางคณิตศาสตร์มาช่วยคำนวณสำหรับการตรวจสอบได้ว่าค่ากล่าวอ้างนั้นเป็นจริงหรือไม่’ และจากการที่เด็กหญิงมิตตี้สะนำข้อมูลที่รวบรวมมาได้คำนวณหาความยาวเส้นทแยงมุมของพื้นที่ได้ เท่ากับ 1.30173 กิโลเมตร</p>		
<p>ข้อคำถามที่ 5</p> <p>คำถาม : ถ้านักเรียนเป็นเด็กหญิงมิตตี้สะจะสรุปผลการทดสอบสมมติฐานได้ว่าอย่างไร และจะสรุปว่าค่ากล่าวอ้างในพินัยกรรมบอกไว้เพียงว่าที่ดินนี้มีเส้นทแยงมุมยาว 1.3 กิโลเมตร เป็นจริงหรือไม่อย่างไร</p> <p>คำตอบ</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p>		
<p><u>ข้อควรระวัง</u> ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>		



เวลาสำหรับทำ เหตุการณ์นี้ 20 นาที	เหตุการณ์ที่ 3 ข้อคำถามที่ 4	ชื่อ-สกุล ชั้น เลขที่
<p>สถานการณ์ จากสมมติฐานที่ทากิคุงตั้งไว้ว่าสูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอก มีความสัมพันธ์</p> <p>กับปริซึม และจะต้องใช้วิธีการหา/คำนวณ/สูตร เดียวกันกับปริซึม ดังนั้น ทากิคุงจึงทำการรวบรวม ข้อมูลเพื่อนำ</p> <p>ไปสู่การทดสอบสมมติฐาน ได้ข้อมูลดังนี้</p> <p>เมื่อเปรียบเทียบลักษณะของทรงกระบอก และปริซึม พบว่า ทรงกระบอกนั้นมีหน้าตัดที่ อยู่ฝั่งซ้าย</p> <p>กันเป็นรูปเดียวกัน นั่นคือ รูปวงกลม แต่ว่าพื้นที่ผิวข้างนั้นมีความโค้ง ต้องทำการคลี่ออกมาแล้วจึง พบว่าได้เป็นรูปคลี่ด้านของของทรงกระบอกเป็นรูปสี่เหลี่ยมมุมฉาก และจากสูตรการหาปริมาตร และพื้นที่ผิวของข้างปริซึมที่เรียนเสร็จเรียบร้อยแล้ว กล่าวไว้ว่า</p> $\text{สูตรการหาปริมาตรของปริซึม} = \text{พื้นที่ฐาน} \times \text{ความสูง}$ $\text{สูตรการหาพื้นที่ผิวข้างของปริซึม} = \text{ความยาวรอบฐาน} \times \text{ความสูง}$		
<p>ข้อคำถามที่ 4</p> <p>คำถาม : นักเรียนคิดว่า ทากิคุงเมื่อนำข้อมูลที่รวบรวมมาได้มาทำการจัดระบบ จะได้ข้อสรุปถึง สูตรการหาปริมาตรของทรงและบอ และสูตรการหาพื้นที่ผิวข้างของทรงกระบอกได้อย่างไร จงอธิบาย (ทั้งสูตรการคำนวณ และที่มาพอสังเขป)</p> <p>คำตอบ</p> <p>.....</p> <p>.....</p> <p>.....</p>		
<p><u>ข้อควรระวัง</u> ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>		
เวลาสำหรับทำ	เหตุการณ์ที่ 3	ชื่อ-สกุล

เหตุการณ์นี้ 20 นาที	ข้อคำถามที่ 5 ชั้น เลขที่
<p>สถานการณ์ จากข้อสังเกตของทากิคุงในตอนแรกเมื่อสังเกตเห็นรูปทรงของทรงกระบอกที่มีความคิดว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกนั้นน่าจะมีความสัมพันธ์กับของปริซึม' และจากการที่ทากิคุงนำข้อมูลที่รวบรวมมาจัดระบบได้ดังนี้</p> <p style="text-align: center;">สูตรปริมาตรของทรงกระบอก คือ $\pi r^2 h$</p> <p style="text-align: center;">สูตรพื้นที่ผิวข้างของทรงกระบอก คือ $2\pi r h$</p>		
<div style="text-align: center;">  <p>จุฬาลงกรณ์มหาวิทยาลัย CHULALONGKORN UNIVERSITY</p> </div>		
<p><u>ข้อควรระวัง</u> ขอให้ตรวจสอบเวลาที่ใช้ในการทำ และตรวจสอบให้แน่ใจว่าได้ทำข้อคำถามนี้ครบถ้วนสมบูรณ์แล้วจึงทำข้อถัดไป เพราะไม่อนุญาตให้กลับมาแก้ไขอีก</p>		



ภาคผนวก ง

แนวคำตอบ วิธีตรวจให้คะแนนรูปกรสองชั้น (double layer scoring rubric)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

เหตุการณ์ที่ 1

แนวคำตอบข้อคำถามที่ 1

1. เด็กชายตะวันระบุปัญหาว่า การนำสกุลเงินบาทไปแลกเปลี่ยนสกุลเงินอื่น ๆ ตามที่ต้องการในปริมาณเท่า ๆ กัน ในแต่ละสกุลเงินจะต้องเตรียมเงินบาทเป็นจำนวนเท่ากันหรือไม่
2. เด็กชายตะวันระบุปัญหาว่า การนำสกุลเงินบาทไปแลกเปลี่ยนสกุลเงินอื่น ๆ ตามที่ต้องการในปริมาณเท่า ๆ กัน ตะวันจะต้องเตรียมเงินบาทมากที่สุดในสกุลเงินใด
3. เด็กชายตะวันระบุปัญหาว่า การนำเงินสกุลบาทไปแลกเปลี่ยนเงินสกุลอื่น ๆ ตามที่ต้องการในปริมาณเท่า ๆ กัน ตะวันจะต้องเตรียมเงินบาทน้อยที่สุดในสกุลเงินใด

วิธีตรวจให้คะแนนรูปกรีกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 1

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 1 นิยาม ปัญหา	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ สกุลเงินบาทเป็นสกุลเงินตั้งต้นที่นำไปแลกเปลี่ยน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ แลกเป็นเงินสกุลอื่น ๆ ในมูลค่าเท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดปัญหาภายใต้บริบทของสถานการณ์	2.1 กำหนดปัญหาได้ถูกต้อง คือ ปัญหาที่เกี่ยวข้องกับการเตรียมเงินบาทไปแลกเปลี่ยนสกุลเงินอื่น ๆ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการกำหนดปัญหาอยู่	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		ภายใต้อัตราข้อมูลใน สถานการณ์ คือ สกุลเงินที่นำไปแลกได้แก่ สกุลเยน สกุลหยวน และ สกุลดอลลาร์สหรัฐ	ถูกต้องทั้งหมด	ถูกต้อง บางส่วน	ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ คำถาม
คะแนนรวม =			คะแนน	คิดเป็น =	คะแนน

ระดับขั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูปริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-8	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 2

เด็กชายตะวันตกสามารถตั้งสมมติฐานไว้ว่า การจะแลกเงินในสกุลต่าง ๆ ให้มีมูลค่าเท่า ๆ กัน ต้องเตรียมเงินบาทไปแลกในปริมาณที่ไม่เท่ากัน

วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 2

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 2 ตั้งสมมติฐาน	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ ต้องเตรียมสกุลเงินบาทไปแลกในปริมาณที่ไม่เท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ สกุลเงินที่ต้องการแลกเปลี่ยนมีมูลค่าเท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดสมมติฐานภายใต้บริบทของสถานการณ์	2.1 กำหนดสมมติฐานได้ถูกต้อง คือ คาดคะเนปริมาณเงินบาทที่นำไปแลกได้ว่าต้องเตรียมไป ในปริมาณที่ไม่เท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการกำหนดสมมติฐานอยู่ภายใต้ข้อมูลในสถานการณ์ คือ สมมติฐานที่เกี่ยวข้องกับการกำหนดปริมาณเงินบาทไปแลกเปลี่ยนสกุลเงินอื่น ๆ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
คะแนนรวม = คะแนน คิดเป็น = คะแนน					

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-8	2
4-6	1
0-3	0



แนวคำตอบข้อคำถามที่ 3

1. ดูจากตารางแสดงอัตราการแลกเปลี่ยนเงินระหว่างประเทศ เพื่อเปรียบเทียบอัตราแลกเปลี่ยนในแต่ละสกุลเงิน
2. หากไม่มีตารางแสดงอัตราการแลกเปลี่ยนเงินระหว่างประเทศ จะต้องใช้ข้อมูลที่เกี่ยวข้องเพื่อเปรียบเทียบและสรุปผล

วิธีตรวจให้คะแนนรูปกรีกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 3

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 3 รวบรวม ข้อมูล	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้องคือ แนวทางการรวบรวมข้อมูลที่เกี่ยวข้องกับอัตราการแลกเปลี่ยนสกุลเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด <u>หรือ</u> ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้องคือ เหตุผลในการใช้แนวทางการเก็บรวบรวมข้อมูล	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด <u>หรือ</u> ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดแนวทางการรวบรวมข้อมูลภายใต้บริบทของสถานการณ์	2.1 กำหนดแนวทางการรวบรวมข้อมูลได้ถูกต้องคือ แนวทางการรวบรวมข้อมูลที่เกี่ยวข้องกับอัตราการแลกเปลี่ยนเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด <u>หรือ</u> ไม่มีการตอบคำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		2.2 บริบทของแนวทางการเก็บรวบรวมข้อมูลอยู่ภายใต้ข้อมูลในสถานการณ์ คือ การแลกเปลี่ยนเงินตราระหว่างสกุลเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ คำถาม
คะแนนรวม = คะแนน			คิดเป็น = คะแนน		

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-8	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 4

บาท : เยน : หยวน : ดอลลาร์สหรัฐ = 100 : 350 : 20 : 3 หรืออัตราส่วนใด ๆ ที่เท่ากับอัตราส่วนข้างต้น

วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 4

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 4 วิเคราะห์ ข้อมูล	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบท ของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์ กำหนดได้ถูกต้อง คือ อัตราแลกเปลี่ยนแต่ละคู่ สกุลเงินตามที่สถานการณ์ กำหนด	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
		1.2 อธิบายเงื่อนไขของ สถานการณ์ได้ถูกต้อง คือ เปรียบเทียบอัตรา แลกเปลี่ยนค่าเงินครบ ทุกสกุลเงิน (สกุลบาท สกุลเยน สกุลหยวน และสกุล ดอลลาร์สหรัฐ)	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
	2. การเข้าใจใน สิ่งที่ต้องการวัด - วิเคราะห์ข้อมูล ภายใต้บริบท ของสถานการณ์	2.1 วิเคราะห์ข้อมูลได้ ถูกต้อง คือ เปรียบเทียบ อัตราส่วนค่าเงินแลกเปลี่ยนทุก สกุลเงิน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
		2.2 บริบทของการวิเคราะห์ ข้อมูลอยู่ภายใต้ข้อมูลใน สถานการณ์ คือ สกุลเงินที่นำมาเปรียบเทียบ ประกอบด้วย สกุลเงินบาท สกุล เยน สกุลหยวน และสกุล ดอลลาร์สหรัฐ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
	3. การสื่อ ความหมาย ทาง คณิตศาสตร์	3.1 เปลี่ยนประโยคภาษา เป็นประโยคทาง คณิตศาสตร์ได้ถูกต้อง คือ เขียนสัญลักษณ์อัตราส่วนใน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
	- เปลี่ยนประโยค ภาษาเป็น ประโยคทาง คณิตศาสตร์ และคำนวณได้ ถูกต้อง	การเปรียบเทียบ			ไม่มีการตอบ คำถาม
		3.2 คำนวนได้ถูกต้อง คือ ทำเป็นอัตราส่วนอย่างต่ำ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
		คะแนนรวม =	คะแนน	คิดเป็น =	คะแนน

ระดับขั้นที่ 2 (คะแนนเต็ม 3 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 5

คำตอบที่หนึ่ง การคำนวณเตรียมเงินสกุลบาทไปแลกเงินสกุลต่าง ๆ ได้แก่ สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ สกุลละ 210 หน่วย ได้ดังนี้

จากสถานการณ์กำหนดให้ บาท : เยน : หยวน : ดอลลาร์สหรัฐ = 100 : 350 : 20
: 3

กำหนดให้เงินบาทที่ต้องเตรียมไป X บาท

พิจารณาอัตราการแลกเปลี่ยนระหว่างเงินสกุลบาท และสกุลเยน คือ
บาท : เยน = 100 : 350 เมื่อต้องการแลกเปลี่ยนสกุลเงินเยน 210 หน่วย
จะได้ว่า $X : 210 = 100 : 350$

$X : 210 = 2 : 7$ (ทำเป็นอัตราส่วนอย่างต่ำ)

$X : 210 = 2 \times 30 : 7 \times 30$ (นำค่าคงที่ 30 มาคูณ เพื่อให้อัตราส่วนเท่ากัน)

$X : 210 = 60 : 210$

จะได้ $X = 60$ นั่นคือ ต้องเตรียมเงินบาทไปแลกเปลี่ยนเงินสกุลเยน 60 บาท

พิจารณาอัตราการแลกเปลี่ยนระหว่างเงินสกุลบาท และสกุลหยวน คือ
บาท : หยวน = 100 : 20 เมื่อต้องการแลกเปลี่ยนสกุลเงินหยวน 210 หน่วย
จะได้ว่า $X : 210 = 100 : 20$

$X : 210 = 5 : 1$ (ทำเป็นอัตราส่วนอย่างต่ำ)

$X : 210 = 5 \times 210 : 1 \times 210$ (นำค่าคงที่ 210 มาคูณ เพื่อให้อัตราส่วนเท่ากัน)

$X : 210 = 1,050 : 210$

จะได้ $X = 1,050$ นั่นคือ ต้องเตรียมเงินบาทไปแลกเปลี่ยนเงินสกุลหยวน 1,050 บาท

พิจารณาอัตราการแลกเปลี่ยนระหว่างเงินสกุลบาท และสกุลดอลลาร์สหรัฐ คือ
บาท : หยวน = 100 : 3 เมื่อต้องการแลกเปลี่ยนสกุลเงินดอลลาร์สหรัฐ 210 หน่วย
จะได้ว่า $X : 210 = 100 : 3$

$X : 210 = 100 \times 70 : 3 \times 70$ (นำค่าคงที่ 70 มาคูณ เพื่อให้อัตราส่วนเท่ากัน)

$X : 210 = 7,000 : 210$

จะได้ $X = 7,000$ นั่นคือ ต้องเตรียมเงินบาทไปแลกเปลี่ยนเงินสกุลดอลลาร์สหรัฐ 7,000 บาท

ต้องเตรียมเงินสกุลบาททั้งหมด $60 + 1,050 + 7,000$ เท่ากับ 8,110 บาท จึงจะได้เงินแต่ละสกุลตามเงื่อนไข #

คำตอบที่สอง การนำสกุลเงินบาทไปแลกเปลี่ยนเงินสกุลอื่น ๆ ตามที่ต้องการในมูลค่าเท่า ๆ

กัน จะต้องเตรียมเงินบาทเป็นจำนวนไม่เท่ากัน

วิธีตรวจให้คะแนนรูปกรีกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 5

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 5 สร้าง ข้อสรุป	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ เตรียมเงินสกุลบาทไปแลกเปลี่ยนสกุลต่าง ๆ ได้แก่ สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ สกุลละ 210 หน่วย	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ อัตราการแลกเปลี่ยนเงินสกุลต่าง ๆ มาจากผลการทดสอบสมมติฐานของสถานการณ์ ดังนี้ บาท : เยน : หยวน : ดอลลาร์สหรัฐ เท่ากับ 100 : 350 : 20 : 3	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดข้อสรุปการเตรียมเงินสกุลบาทไปแลกเปลี่ยน และสรุปผลสมมติฐานสอดคล้องกับบริบทของสถานการณ์	2.1 เขียนข้อสรุปของการแลกเปลี่ยน และผลสมมติฐาน คือ 1) เตรียมเงินสกุลบาทไปเป็นจำนวน 8,110 บาท และ 2) สรุปสมมติฐานว่า การนำสกุลเงินบาทไปแลกเปลี่ยนเงินสกุลอื่น ๆ ตามที่ต้องการในมูลค่าเท่า ๆ กัน จะต้องเตรียมเงินบาทเป็นจำนวนไม่เท่ากัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการสร้างข้อสรุปอยู่ภายใต้ข้อมูลในสถานการณ์ คือ สกุลเงินที่นำมาเปรียบเทียบกับสกุลเงินบาทประกอบด้วย สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
3. การสื่อความหมายทางคณิตศาสตร์ - เปลี่ยนประโยค	3.1 เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ได้ถูกต้อง คือ เขียนสัดส่วนในแต่ละคู่การแลกเปลี่ยนสกุลเงิน (บาท-เยน, บาท-หยวน และ บาท-ดอลลาร์สหรัฐ)	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม	

	ภาษาเป็นประโยคทางคณิตศาสตร์และจำนวนได้ถูกต้อง	3.2 จำนวนได้ถูกต้อง คือ จำนวนอัตรา การแลกเปลี่ยนแต่ละคู่สกุลบาทกับสกุลเงินอื่น ๆ ได้อย่างถูกต้อง คือ เตรียมเงินบาทไปแลกเป็นเงินสกุลเยน 60 บาท, เตรียมเงินบาทไปแลกเป็นเงินสกุลหยวน 1,050 บาท และเตรียมเงินบาทไปแลกเป็นเงินสกุลดอลลาร์สหรัฐ 7,000 บาท	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
คะแนนรวม = คะแนน คิดเป็น = คะแนน					

ระดับขั้นที่ 2 (คะแนนเต็ม 3 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0

เหตุการณ์ที่ 2

แนวคำตอบข้อคำถามที่ 1

เด็กหญิงมีนิสัยระบุนปัญหาว่า จะต้องใช้เครื่องมือในการวัดประเภทและทฤษฎีบททางคณิตศาสตร์ใดบ้างมาช่วยคำนวณสำหรับการตรวจสอบค่ากล่าวเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉากในพินัยกรรม

วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 1

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 1 นิยาม ปัญหา	1. ความถูกต้องของเนื้อหา - ระบุนข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุนข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ ที่ดินรูปสี่เหลี่ยมมุมฉาก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ การหาเส้นทแยงมุมของรูปที่ดินในพินัยกรรม	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดปัญหาภายใต้บริบทของสถานการณ์	2.1 กำหนดปัญหาได้ถูกต้อง คือ เครื่องมือในการวัด และทฤษฎีบททางคณิตศาสตร์ใดบ้าง สำหรับการตรวจสอบค่ากล่าวข้าง	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการกำหนดปัญหาอยู่ภายใต้ข้อมูลในสถานการณ์ คือ ค่ากล่าวข้างในพินัยกรรม	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
คะแนนรวม = คะแนน คิดเป็น = คะแนน					

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-9	2
4-6	1
0-3	0



แนวคำตอบข้อคำถามที่ 2

- เด็กหญิงมีตีสี่จะสามารถตั้งสมมติฐานไว้ว่า การจะตรวจสอบความยาวของเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉากในค่ากล่าวอ้างของพินัยกรรมนั้น จะต้องใช้เครื่องมือวัดประเภทวัดความยาว และทฤษฎีบททางคณิตศาสตร์เกี่ยวกับสามเหลี่ยม (ทฤษฎีบทพีทาโกรัส) หรือ
- เด็กหญิงมีตีสี่จะสามารถตั้งสมมติฐานไว้ว่า การจะตรวจสอบความยาวของเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉากในค่ากล่าวอ้างของพินัยกรรมนั้น จะต้องใช้เครื่องมือวัดประเภทวัดขนาดของมุม และทฤษฎีบททางคณิตศาสตร์เกี่ยวกับสามเหลี่ยม (ตรีโกณมิติ)

วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 2

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 2 ตั้งสมมติฐาน	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ ที่ดินในพินัยกรรมเป็นรูปสี่เหลี่ยมมุมฉาก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ การตรวจสอบความยาวของเส้นทแยงมุมของที่ดินในพินัยกรรม	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดสมมติฐานภายใต้บริบทของสถานการณ์	2.1 กำหนดสมมติฐานได้ถูกต้อง คือ 1) ต้องใช้เครื่องมือวัดประเภทวัดความยาว/วัดขนาดของมุม และ 2) ทฤษฎีบททางคณิตศาสตร์เกี่ยวกับสามเหลี่ยม (ทฤษฎีบทพีทาโกรัส /	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		ตรีโกณมิติ)			
		2.2 บริบทของการกำหนด สมมติฐานอยู่ภายใต้ข้อมูล ในสถานการณ์ คือ ค่ากล่าวอ้างในพินัยกรรม	<input type="checkbox"/> ถูกต้อง ทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด <u>หรือ</u> ไม่มีการตอบ คำถาม
คะแนนรวม = คะแนน			คิดเป็น = คะแนน		

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-9	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 3

1. เดินวัดระยะทางตามด้านกว้าง และด้านยาวของพื้นที่ โดยใช้เครื่องมือวัดความยาวที่มีเพื่อวัดระยะทาง และ
2. อธิบายทฤษฎีบทพีทาโกรัสว่ากล่าวไว้ว่าอย่างไร เพื่อสนับสนุนค่ากล่าวอ้างในพินัยกรรมเกี่ยวกับเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉาก

วิธีตรวจให้คะแนนรูปกรีกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 3

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 3 รวบรวม ข้อมูล	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ การตรวจสอบค่ากล่าวอ้างในพินัยกรรมเกี่ยวกับเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉาก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด <u>หรือ</u> ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ การอธิบายทฤษฎีบทเกี่ยวกับรูปสามเหลี่ยมมุมฉาก (พีทาโกรัส) เพื่อสนับสนุนค่ากล่าวอ้างในพินัยกรรม และนำไปสู่การทดสอบสมมติฐาน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด <u>หรือ</u> ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดแนวทางการรวบรวมข้อมูลภายใต้บริบทของสถานการณ์	2.1 กำหนดแนวทางการรวบรวมข้อมูลได้ถูกต้อง คือ ใช้เครื่องมือวัดความยาวที่มีเพื่อวัดระยะทางตามด้านกว้าง ด้านยาวของ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด <u>หรือ</u> ไม่มีการตอบคำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		พื้นที่			
		2.2 บริบทของแนวทางการเก็บรวบรวมข้อมูลอยู่ภายใต้ข้อมูลในสถานการณ์คือ เครื่องนับก้าว เครื่องแปลงจำนวนก้าวเป็นระยะทาง	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
คะแนนรวม =			คะแนน	คิดเป็น =	คะแนน

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูปริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-9	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 4

นำความยาวด้านยาว และความยาวด้านกว้างจากข้อมูลในสถานการณ์มาใช้กับทฤษฎีบทพีทาโกรัส เพื่อตรวจสอบว่า ตามทฤษฎีแล้วความยาวของเส้นทแยงมุมของพื้นที่ต้องมีค่าประมาณเท่าใด ดังนี้

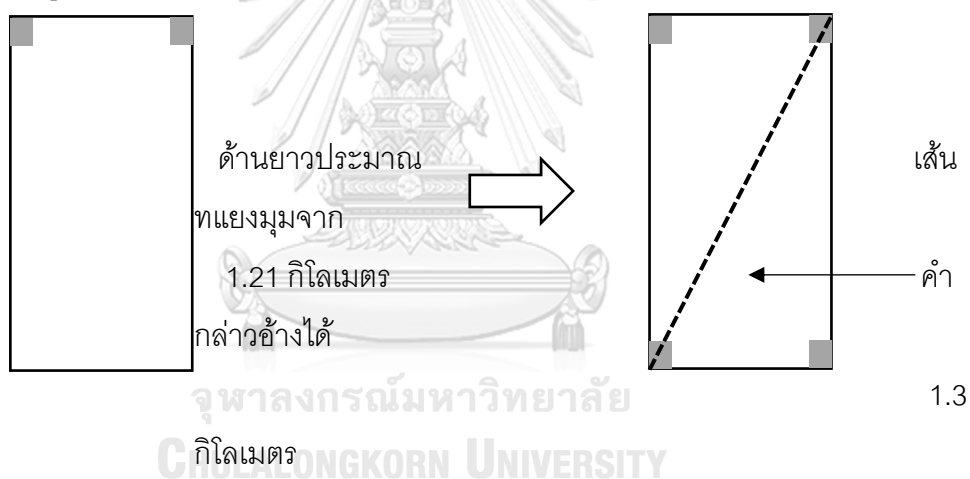
จากทฤษฎีบทพีทาโกรัสกล่าวถึงรูปสามเหลี่ยมมุมฉากไว้ว่า “ผลบวกของกำลังสองของด้านประกอบมุมฉาก เท่ากับกำลังสองของด้านตรงข้ามมุมฉาก”

จะได้สมการความสัมพันธ์จากทฤษฎีบทพีทาโกรัสได้ดังนี้

$$\text{“ด้านประกอบมุมฉาก}^2 + \text{ด้านประกอบมุมฉาก}^2 = \text{ด้านตรงข้ามมุมฉาก}^2\text{”} \quad \text{---}$$

(สมการที่ 1)

จากการรวบรวมข้อมูลขนาดมรดกที่ดินรูปสี่เหลี่ยมมุมฉากของเด็กหญิงมิตตี้สะ พบว่าด้านยาวของพื้นที่ยาวประมาณ 1.21 กิโลเมตร และด้านกว้างของพื้นที่ยาวประมาณ 0.48 กิโลเมตร ดังรูป



ด้านกว้างประมาณ 0.48 กิโลเมตร

พบว่าเมื่อลากเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉาก จะเกิดรูปสามเหลี่ยมมุมฉาก 2 รูป ซึ่งสามารถใช้ความสัมพันธ์ที่ได้จากทฤษฎีบทพีทาโกรัสได้ และเมื่อนำข้อมูลที่ได้จากการรวบรวมของเด็กหญิงมิตสึฮะไปแทนค่าในสมการที่ 1 เพื่อหาความยาวของเส้นทแยงมุมมรดกที่ดินรูปสี่เหลี่ยมมุมฉากในคำกล่าวอ้างในพินัยกรรมนี้ จะได้

$$0.48^2 + 1.21^2 = \text{ด้านตรงข้ามมุมฉาก}^2$$

$$0.2304 + 1.4641 = \text{ด้านตรงข้ามมุมฉาก}^2$$

$$1.6945 = \text{ด้านตรงข้ามมุมฉาก}^2$$

$$1.3017 \approx \text{ด้านตรงข้ามมุมฉาก}$$

เนื่องจากความยาวของด้านตรงข้ามมุมฉากที่ได้จากรูปสามเหลี่ยมมุมฉากในทฤษฎีบทพีทาโกรัส คือ ความยาวของเส้นทแยงมุมในรูปสี่เหลี่ยมมุมฉากของมรดกที่ดินในพินัยกรรม ดังนั้นที่ดินผืนนี้มีเส้นทแยงมุมประมาณ 1.3017 กิโลเมตร #

วิธีตรวจให้คะแนนรูปกรีกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 4

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 4 จัดระบบ ข้อมูล	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบท ของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์ กำหนดได้ถูกต้อง คือ ใช้ข้อมูลของความยาว ด้านกว้างและด้านยาว ของที่ได้จากการ รวบรวมข้อมูลใน สถานการณ์	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
		1.2 อธิบายเงื่อนไขของ สถานการณ์ได้ถูกต้อง คือ ใช้ทฤษฎีบทพีทาโกรัส ที่กล่าวถึงรูป สามเหลี่ยมมุมฉากช่วย ตรวจสอบค่ากล่าวอ้าง	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
	2. การเข้าใจใน สิ่งที่ต้องการวัด - วิเคราะห์ข้อมูล ภายใต้บริบท ของสถานการณ์	2.1 วิเคราะห์ข้อมูลได้ ถูกต้อง คือ บอกได้ว่าความ ยาวด้านตรงข้ามมุมฉากใน ทฤษฎีบทพีทาโกรัส คือ ความ ยาวของเส้นทแยงมุมของ รูปสี่เหลี่ยมมุมฉาก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
		2.2 บริบทของการวิเคราะห์ ข้อมูลอยู่ภายใต้ข้อมูลใน สถานการณ์ คือ มรดกที่ดิน เป็นรูปสี่เหลี่ยมมุมฉาก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
	3. การสื่อ ความหมาย ทาง คณิตศาสตร์ - เปลี่ยนประโยค ภาษาเป็น ประโยคทาง	3.1 เปลี่ยนประโยคภาษา เป็นประโยคทางคณิตศาสตร์ ได้ถูกต้อง คือ เขียน สมการความสัมพันธ์ จากทฤษฎีบทพีทา โกรัส	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
	คณิตศาสตร์ และคำนวณได้ ถูกต้อง	3.2 คำนวณได้ถูกต้อง คือ นำข้อมูลของความยาว ด้านกว้างและด้านยาว ของที่ได้จากการ รวบรวมข้อมูลใน สถานการณ์ไปแทนค่า ในความสัมพันธ์จาก ทฤษฎีบทพีทาโกรัสได้ ความยาวเส้นทแยงมุม ของพื้นที่นั้นมี ค่าประมาณ 1.3017 กิโลเมตร	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบ คำถาม
คะแนนรวม = คะแนน			คิดเป็น = คะแนน		

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 5

คำตอบที่หนึ่ง เครื่องมือที่ใช้วัดความยาว ได้แก่ เครื่องนับก้าว เครื่องแปลงจำนวนก้าวเป็นระยะทางเพื่อวัดระยะทาง และทฤษฎีบททางคณิตศาสตร์ที่ใช้เกี่ยวกับรูปสามเหลี่ยมมุมฉาก คือ ทฤษฎีบทพีทาโกรัส เพื่อสนับสนุนค่ากล่าวอ้างในพินัยกรรมเกี่ยวกับเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉาก

ดังนั้นสามารถสรุปผลการทดสอบสมมติฐานได้ว่า “การจะตรวจสอบค่ากล่าวอ้างในพินัยกรรมเกี่ยวกับเส้นทแยงมุมของมรดกที่ดินรูปสี่เหลี่ยมมุมฉากนั้น สามารถใช้เครื่องมือวัดความยาว และทฤษฎีบทพีทาโกรัสตรวจสอบความยาวเส้นทแยงมุมพื้นที่ได้จริง”

คำตอบที่สอง สามารถนำข้อมูลความยาวของเส้นทแยงมุมของพื้นที่ที่ได้จากการคำนวณโดยใช้หลัก การประมาณเพื่อสรุปค่ากล่าวอ้างความยาวของเส้นทแยงมุมของที่ดินในพินัยกรรมให้เป็นจำนวนเต็มได้ว่า “เส้นทแยงมุมของที่ดินในค่ากล่าวอ้างตามพินัยกรรมเท่ากับ 1.3 กิโลเมตร เป็นความจริง” เนื่องจากค่าที่คำนวณได้จากการวัดถือว่าสอดคล้องกับทฤษฎีบทพีทาโกรัส (คลาดเคลื่อนน้อยมาก) #

วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 5

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 5 สร้าง ข้อสรุป	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบทของ สถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ ถูกต้อง คือ ใช้เครื่องมือที่ใช้วัดความยาวและ ทฤษฎีทางคณิตศาสตร์ช่วยในการหาความ ยาวของเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุม ฉาก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ ถูกต้อง คือ 1) เครื่องมือที่ใช้วัดความยาว ได้แก่ เครื่องนับก้าว เครื่องแปลงจำนวนก้าว เป็นระยะทาง เพื่อวัดระยะทาง และ 2) ทฤษฎีทางคณิตศาสตร์ที่ใช้เกี่ยวกับรูป สามเหลี่ยมมุมฉาก คือ ทฤษฎีบทพีทาโกรัส เพื่อสนับสนุนความยาวเส้นทแยงมุมของที่ดิน รูปสี่เหลี่ยมมุมฉากภายใต้เงื่อนไขของ สถานการณ์ (คำกล่าวอ้างในพินัยกรรม)	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่ง ที่ต้องการวัด - กำหนดข้อสรุปผล การทดสอบ สมมติฐานและคำ กล่าวอ้างใน พินัยกรรม สอดคล้องกับบริบท ของสถานการณ์	2.1 เขียนข้อสรุปของสรุปผลการทดสอบ สมมติฐานและคำกล่าวอ้างในพินัยกรรม คือ 1) สรุปผลการทดสอบสมมติฐานได้ว่า “การจะตรวจสอบคำกล่าวอ้างในพินัยกรรม เกี่ยวกับเส้นทแยงมุมของมรดกที่ดินรูป สี่เหลี่ยมมุมฉากนั้น สามารถใช้เครื่องมือวัด ความยาว และทฤษฎีบทพีทาโกรัสตรวจสอบ ความยาว เส้นทแยงมุมพื้นที่ได้จริง” และ 2) สรุปคำกล่าวอ้างความยาวของเส้นทแยงมุม ของที่ดินในพินัยกรรมได้ว่า “เส้นทแยงมุมของ ที่ดินในคำกล่าวอ้างตามพินัยกรรมเป็นความ จริง”	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการสร้างข้อสรุปอยู่ภายใต้ ข้อมูลในสถานการณ์ คือ เส้นทแยงมุมของที่ดินในคำกล่าวอ้างตาม พินัยกรรม เท่ากับ 1.3 กิโลเมตร	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	3. การสื่อ ความหมาย ทางคณิตศาสตร์ - เปลี่ยนประโยค ภาษาเป็นประโยค ทางคณิตศาสตร์ และ คำนวณ ได้ ถูกต้อง	3.1 เปลี่ยนประโยคภาษาเป็นประโยคทาง คณิตศาสตร์ได้ถูกต้อง คือ การประมาณ (\approx) ค่าความยาวของเส้นทแยง มุมที่ได้จากการคำนวณในสถานการณ์ของ มรดกที่ดินให้เป็นจำนวนเต็ม	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		3.2 คำนวณได้ถูกต้อง คือ ผลการประมาณค่าความยาวของเส้นทแยงมุม ของพื้นที่ 1.30173 \approx 13 กิโลเมตร	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
คะแนนรวม = คะแนน คิดเป็น = คะแนน					

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0



เหตุการณ์ที่ 3

แนวคำตอบข้อคำถามที่ 1

หากิคุงระบุปัญหาว่า ต้องใช้ความรู้ทางคณิตศาสตร์เรื่องใด เพื่อหาความสัมพันธ์ระหว่าง ปริมาตร และพื้นที่ผิวข้างของปริซึมกับทรงกระบอก และความรู้ที่นั้นกล่าวไว้ว่าอย่างไรบ้าง

วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 1

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ชั้นที่ 1 นิยาม ปัญหา	1. ความถูกต้อง ของเนื้อหา - ระบุข้อมูล และ อธิบายเงื่อนไข ภายใต้บริบทของ สถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์ กำหนดได้ถูกต้อง คือ ความรู้ทางคณิตศาสตร์ ที่ใช้หาปริมาตรและพื้นที่ผิว ของทรงกระบอกมี ความสัมพันธ์กับปริมาตรและ พื้นที่ผิวข้างของปริซึม	<input type="checkbox"/> ถูกต้อง ทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ ไม่มี การตอบ คำถาม
		1.2 อธิบายเงื่อนไขของ สถานการณ์ได้ถูกต้อง คือ อธิบายความรู้ทาง คณิตศาสตร์ที่ นำมาใช้	<input type="checkbox"/> ถูกต้อง ทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ คำถาม
	2. การเข้าใจใน สิ่งที่ต้องการวัด - กำหนดปัญหา ภายใต้บริบทของ สถานการณ์	2.1 กำหนดปัญหาได้ ถูกต้อง คือ 1) ระบุปัญหาว่า ต้องใช้ ความรู้ทาง คณิตศาสตร์เรื่องใด และ 2) ความรู้ทาง คณิตศาสตร์ที่ใช้ นั้น กล่าวอย่างไร	<input type="checkbox"/> ถูกต้อง ทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ คำถาม
		2.2 บริบทของการกำหนด ปัญหาอยู่ภายใต้ข้อมูลใน สถานการณ์ คือ การหาปริมาตรและ	<input type="checkbox"/> ถูกต้อง ทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		พื้นที่ผิวข้างของ ทรงกระบอกมี ความสัมพันธ์กับ ปริมาตรและพื้นที่ผิวข้างของ ปริซึม			ไม่มีการตอบ คำถาม
คะแนนรวม =			คะแนน	คิดเป็น =	คะแนน

ระดับขั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูปริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-9	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 2

1. ทาकिคุงสามารถตั้งสมมติฐานไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม และจะต้องใช้วิธีการหาเดียวกันกับสูตรการหาปริมาตรและพื้นที่ผิวข้างของปริซึม
2. ทาकिคุงสามารถตั้งสมมติฐานไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม และจะต้องใช้การคำนวณเดียวกันกับสูตรการหาปริมาตรและพื้นที่ผิวข้างของปริซึม
3. ทาकिคุงสามารถตั้งสมมติฐานไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม และจะต้องมีสูตรเดียวกันกับสูตรการหาปริมาตรและพื้นที่ผิวข้างของปริซึม

วิธีตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 2

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 2 ตั้งสมมติฐาน	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ การตั้งสมมติฐานหาสูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดสมมติฐานภายใต้บริบทของสถานการณ์	2.1 กำหนดสมมติฐานได้ถูกต้อง คือ 1) สูตรการหาปริมาตร และ 2) พื้นที่ผิวข้างของทรงกระบอกใช้วิธีการหา/คำนวณ/สูตรเดียวกันกับปริซึม	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		2.2 บริบทของการกำหนด สมมติฐานอยู่ภายใต้ข้อมูล ในสถานการณ์ คือ สูตรการหาปริมาตรและ พื้นที่ผิวข้างของปริซึม	<input type="checkbox"/> ถูกต้อง ทั้งหมด	<input type="checkbox"/> ถูกต้อง บางส่วน	<input type="checkbox"/> ไม่ถูกต้อง ทั้งหมด หรือ ไม่มีการตอบ คำถาม
คะแนนรวม =			คิดเป็น = คะแนน		

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-9	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 3

1. เปรียบเทียบลักษณะทางกายภาพของทรงกระบอก ว่ามีความสัมพันธ์กับปริซึมอย่างไร เพื่อรวบรวมข้อมูลความเหมือนและความต่าง นำไปสู่การทดสอบสมมติฐาน และ
2. ค้นหาสูตรการคำนวณปริมาตร และพื้นที่ผิวข้างของปริซึม ว่ากล่าวไว้ว่าอย่างไร เพื่อนำมาสร้างข้อสรุปสูตรการหาปริมาตรและพื้นที่ผิวของทรงกระบอก

วิธีตรวจให้คะแนนรูปริกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 3

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 3 รวบรวม ข้อมูล	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ ค้นหาสูตรการคำนวณปริมาตร และพื้นที่ผิวข้างของปริซึม	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมดหรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ อธิบายสูตรการหาปริมาตร และพื้นที่ผิวข้างของปริซึม	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดแนวทางการรวบรวมข้อมูลภายใต้บริบทของสถานการณ์	2.1 กำหนดแนวทางการรวบรวมข้อมูลได้ถูกต้อง คือ เปรียบเทียบลักษณะทางกายภาพของทรงกระบอกว่ามีความสัมพันธ์กับปริซึมอย่างไร	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของแนวทางการเก็บรวบรวมข้อมูลอยู่ภายใต้ข้อมูลในสถานการณ์ คือ เพื่อรวบรวมข้อมูลความเหมือนและความต่างของปริซึมที่เรียนไป	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		แล้ว นำไปสู่การ ทดสอบสมมติฐาน			
		คะแนนรวม =	คะแนน	คิดเป็น =	คะแนน

ระดับขั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
7-9	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 4

จากสมมติฐานที่ตั้งไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม

คำตอบที่หนึ่ง การหาสูตรการคำนวณปริมาตรของทรงกระบอก

จากเปรียบเทียบลักษณะภายนอกของทรงกระบอก และปริซึม พบว่า ทรงกระบอกนั้นมีหน้าตัดที่อยู่ฝั่งข้ามกันเป็นรูปเดียวกัน นั่นคือ รูปวงกลม และความสูงของทรงกระบอกและปริซึมในขนาดเท่ากัน จะใช้ความสูงเดียวกัน (มีค่าเท่ากับ h)

$$\begin{aligned} \text{จาก สูตรการคำนวณปริมาตรของปริซึม} &= (\text{พื้นที่ฐาน}) \times \text{ความสูง} \\ \text{ทำให้ได้ว่า สูตรการคำนวณปริมาตรของทรงกระบอก} &= (\text{พื้นที่ฐาน}) \times \text{ความสูง} \\ &= (\text{พื้นที่รูปวงกลม}) \times \\ &\quad \text{ความสูง} \\ &= (\pi r^2)h \\ &= \pi r^2 h \end{aligned}$$

#

คำตอบที่สอง การหาสูตรการคำนวณพื้นที่ผิวข้างของทรงกระบอก

จากรูปคลี่ด้านข้างของทรงกระบอกเป็นรูปสี่เหลี่ยมมุมฉาก พบว่า ความยาวด้านหนึ่งของรูปคลี่จะเท่ากับความยาวเส้นรอบวงของวงกลม (มีค่าเท่ากับ $2\pi r$) และ **พาลงกรณ์มหาวิทยาลัย**

ความยาวอีกด้านหนึ่งจะเท่ากับความสูงของทรงกระบอก (มีค่าเท่ากับ h)

$$\begin{aligned} \text{จาก สูตรการคำนวณพื้นที่ผิวข้างของของปริซึม} &= (\text{ความยาวรอบฐาน}) \\ &\quad \times \text{ความสูง} \\ \text{ดังนั้น สูตรการคำนวณพื้นที่ผิวข้างของทรงกระบอก} &= (\text{ความยาวรอบฐาน}) \\ &\quad \times \text{ความสูง} \\ &= (2\pi r)h \\ &= 2\pi r h \end{aligned}$$

#

วิธีตรวจให้คะแนนรูปrikสองชั้น (double layer scoring rubric) ข้อคำถามที่ 4

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 4 จัดระบบ ข้อมูล	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ 1) ทรงกระบอกมีหน้าตัดเป็นรูปเดียวกันกับปริซึม คือ รูปวงกลม และ 2) รูปคลี่ด้านข้างของทรงกระบอกเป็นรูปสี่เหลี่ยมมุมฉาก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ 1) สำหรับการหาสูตรปริมาตรของทรงกระบอก จากการเปรียบเทียบลักษณะภายนอกของทรงกระบอกและปริซึม และ 2) สำหรับการหาสูตรพื้นที่ผิวข้างของทรงกระบอก จากรูปคลี่ด้านข้างของทรงกระบอก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2. การเข้าใจในสิ่งที่ต้องการวัด - จัดระบบข้อมูลภายใต้บริบทของสถานการณ์	2.1 จัดระบบข้อมูลได้ถูกต้อง คือ 1) สำหรับการหาสูตรปริมาตรของทรงกระบอก สามารถดูจากความสัมพันธ์ของสูตรการหาปริมาตรปริซึมกับลักษณะภายนอกของทรงกระบอก - ปริซึม และ 2) สำหรับการหาสูตรพื้นที่ผิวข้างของทรงกระบอก สามารถดูจากความสัมพันธ์ของสูตรการหาพื้นที่ผิวข้างของปริซึมกับรูปคลี่ของทรงกระบอก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน
		2.2 บริบทของการจัดระบบข้อมูลอยู่ภายใต้ข้อมูลในสถานการณ์ คือ จากสมมติฐานที่ตั้งไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอก มีความสัมพันธ์กับ	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		ปริซึม			

วิธีตรวจให้คะแนนรูปกรีกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 4 (ต่อ)

ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
<p>ขั้นที่ 4</p> <p>จัดระบบ</p> <p>ข้อมูล</p>	<p>3. การสื่อ</p> <p>ความหมาย</p> <p>ทาง</p> <p>คณิตศาสตร์</p> <p>- เปลี่ยนประโยค</p> <p>ภาษาเป็น</p> <p>ประโยคทาง</p> <p>คณิตศาสตร์</p> <p>และคำนวณได้</p> <p>ถูกต้อง</p>	<p>3.1 เปลี่ยนประโยคภาษาเป็น</p> <p>ประโยคทางคณิตศาสตร์ได้</p> <p>ถูกต้อง คือ</p> <p>1) สำหรับการหาสูตรปริมาตรของ</p> <p>ทรงกระบอก สามารถระบุได้ว่า</p> <p>พื้นที่ฐานของทรงกระบอก คือ พื้นที่</p> <p>รูปวงกลม (πr^2) และความสูงของ</p> <p>ทรงกระบอก คือ ความสูงของปริซึม</p> <p>(h)</p> <p>และ</p> <p>2) สำหรับการหาสูตรพื้นที่ผิวข้าง</p> <p>ของทรงกระบอก สามารถระบุได้ว่า</p> <p>ความยาวด้านหนึ่งของรูป</p> <p>คลี่จะเท่ากับความยาวเส้น</p> <p>รอบวงของวงกลม ($2\pi r$)</p> <p>และความยาวอีกด้านหนึ่ง</p> <p>จะเท่ากับ ความสูงของ</p> <p>ทรงกระบอก (h)</p>	<p><input type="checkbox"/></p> <p>ถูกต้อง</p> <p>ทั้งหมด</p>	<p><input type="checkbox"/></p> <p>ถูกต้องบางส่วน</p>	<p><input type="checkbox"/></p> <p>ไม่ถูกต้อง</p> <p>ทั้งหมด หรือ</p> <p>ไม่มีการตอบ</p> <p>คำถาม</p>
		<p>3.2 คำนวณได้ถูกต้อง คือ</p> <p>1) สำหรับการหาสูตรปริมาตรของ</p> <p>ทรงกระบอก = (พื้นที่ฐาน) x</p> <p>ความสูง = $\pi r^2 h$</p> <p>และ</p> <p>2) สำหรับการหาสูตรพื้นที่ผิวข้าง</p> <p>ของทรงกระบอก</p> <p>= (ความยาวรอบฐาน) x</p> <p>ความสูง</p>	<p><input type="checkbox"/></p> <p>ถูกต้อง</p> <p>ทั้งหมด</p>	<p><input type="checkbox"/></p> <p>ถูกต้องบางส่วน</p>	<p><input type="checkbox"/></p> <p>ไม่ถูกต้อง</p> <p>ทั้งหมด หรือ</p> <p>ไม่มีการตอบ</p> <p>คำถาม</p>

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
		= $2\pi rh$			
	คะแนนรวม = คะแนน คิดเป็น = คะแนน				

ระดับชั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0

แนวคำตอบข้อคำถามที่ 5

คำตอบที่หนึ่ง จากความสัมพันธ์ของปริซึมและทรงกระบอก สามารถสรุปผลการทดสอบสมมติฐานได้ว่า “สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกจะต้องใช้วิธีการหา/คำนวณ/สูตร เดียวกันกับปริซึม” #

คำตอบที่สอง ถ้า ‘ทรงกระบอกที่มีความยาวรัศมีที่ฐาน 7 เซนติเมตร สูง 10 เซนติเมตร’ และจากข้อมูลสูตรปริมาตรและพื้นที่ผิวข้างของทรงกระบอกในสถานการณ์ที่ว่า

“สูตรปริมาตรของทรงกระบอก คือ $\pi r^2 h$

สูตรพื้นที่ผิวข้างของทรงกระบอก คือ $2\pi r h$ ”

ดังนั้น สามารถนำความยาวรัศมีที่ฐาน (r) 7 เซนติเมตร สูง (h) 10 เซนติเมตร ไปแทนในสูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกได้ดังนี้

$$\begin{aligned} \text{ทรงกระบอกนี้มีปริมาตร} &= \pi r^2 h && \text{เมื่อกำหนดให้ } \pi = \frac{22}{7} \\ &= \pi (7)^2 (10) && \text{เมื่อกำหนดให้ } \pi = \frac{22}{7} \\ &= \frac{22}{7} (7)^2 (10) \\ &= 1,540 && \text{ลูกบาศก์เซนติเมตร} \end{aligned}$$

ดังนั้น ปริมาตรของทรงกระบอกนี้ เท่ากับ 1,540 ลูกบาศก์เซนติเมตร #

$$\begin{aligned} \text{ทรงกระบอกนี้มีพื้นที่ผิวข้าง} &= 2\pi r h && \text{เมื่อกำหนดให้ } \pi = \frac{22}{7} \\ &= 2\pi (7)(10) && \text{เมื่อกำหนดให้ } \pi = \frac{22}{7} \\ &= 2\left(\frac{22}{7}\right)(7)(10) \\ &= 440 && \text{ตารางเซนติเมตร} \end{aligned}$$

ดังนั้น พื้นที่ผิวของทรงกระบอกนี้ เท่ากับ 440 ตารางเซนติเมตร #

วิธีตรวจให้คะแนนรูปกรีกสองชั้น (double layer scoring rubric) ข้อคำถามที่ 5

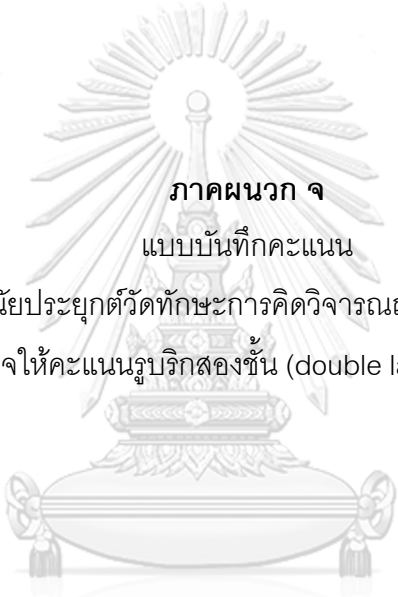
ระดับชั้นที่ 1

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
ขั้นที่ 5 สร้าง ข้อสรุป	1. ความถูกต้องของเนื้อหา - ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	1.1 ระบุข้อมูลที่สถานการณ์กำหนดได้ถูกต้อง คือ 1) สูตรปริมาตรของทรงกระบอกคือ $\pi r^2 h$ และ 2) สูตรพื้นที่ผิวข้างของทรงกระบอก คือ $2\pi r h$ และ 3) กำหนดให้ $\pi = \frac{22}{7}$	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		1.2 อธิบายเงื่อนไขของสถานการณ์ได้ถูกต้อง คือ สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกและปริซึม นั้นมีความสัมพันธ์กัน	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
	2. การเข้าใจในสิ่งที่ต้องการวัด - กำหนดข้อสรุปผลการทดสอบสมมติฐาน และสามารถนำข้อมูลที่กำหนดให้ไปแทนค่าในตัวแปรของสูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอก	2.1 เขียนข้อสรุปผลการทดสอบสมมติฐาน และสามารถนำข้อมูลที่กำหนดให้ไปแทนค่าในตัวแปรของสูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอก คือ 1) ผลการทดสอบสมมติฐานสรุปได้ว่า “สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกจะต้องใช้วิธีการหา/คำนวณ/สูตร เดียวกันกับปริซึม” และ 2) สามารถนำข้อมูลทรงกระบอกที่ให้ ความยาวรัศมีที่ฐาน (r) และความสูง (h) ไปแทนในสูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอก	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		2.2 บริบทของการสร้างข้อสรุปอยู่ภายใต้ข้อมูลในสถานการณ์ คือ ทรงกระบอกมีความยาวรัศมีที่ฐาน 7 เซนติเมตร และสูง 10 เซนติเมตร	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
3. การสื่อ	3.1 เปลี่ยนประโยคภาษาเป็นประโยคทาง	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

ขั้นตอน	องค์ประกอบ	ข้อรายการย่อย	2	1	0
	<p>ความหมาย ทางคณิตศาสตร์</p> <p>- เปลี่ยนประโยค ภาษาเป็นประโยค ทางคณิตศาสตร์ และคำนวณได้ ถูกต้อง</p>	<p>คณิตศาสตร์ได้ถูกต้อง คือ</p> <p>1) ปริมาตรของทรงกระบอก = $\pi r^2 h$ $= \frac{22}{7} (7)^2 (10)$ ลูกบาศก์ เซนติเมตร <u>และ</u></p> <p>2) พื้นที่ผิวข้างของทรงกระบอก = $2\pi rh$ $= 2(\frac{22}{7})(7)(10)$ ตารางเซนติเมตร</p>	ถูกต้องทั้งหมด	ถูกต้องบางส่วน	ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		<p>3.2 คำนวณได้ถูกต้อง คือ</p> <p>1) ปริมาตรของทรงกระบอกนี้ เท่ากับ 1,540 ลูกบาศก์เซนติเมตร <u>และ</u> 2) พื้นที่ผิวข้างของ ทรงกระบอกนี้ เท่ากับ 440 ตาราง เซนติเมตร</p>	<input type="checkbox"/> ถูกต้องทั้งหมด	<input type="checkbox"/> ถูกต้องบางส่วน	<input type="checkbox"/> ไม่ถูกต้องทั้งหมด หรือ ไม่มีการตอบคำถาม
		คะแนนรวม =	คะแนน	คิดเป็น =	คะแนน

ระดับขั้นที่ 2 (คะแนนเต็ม 2 คะแนน)

เกณฑ์การให้คะแนนแบบรูบริก	
ก่อนแปลงระดับคะแนน	หลังแปลงระดับคะแนน
10-12	3
7-9	2
4-6	1
0-3	0



ภาคผนวก จ

แบบบันทึกคะแนน

แบบสอบอัตรานัยประยุกต์วัดทักษะการคิดวิจารณ์ญาณทางคณิตศาสตร์
โดยใช้วิธีตรวจให้คะแนนรูปกรสองชั้น (double layer scoring rubric)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

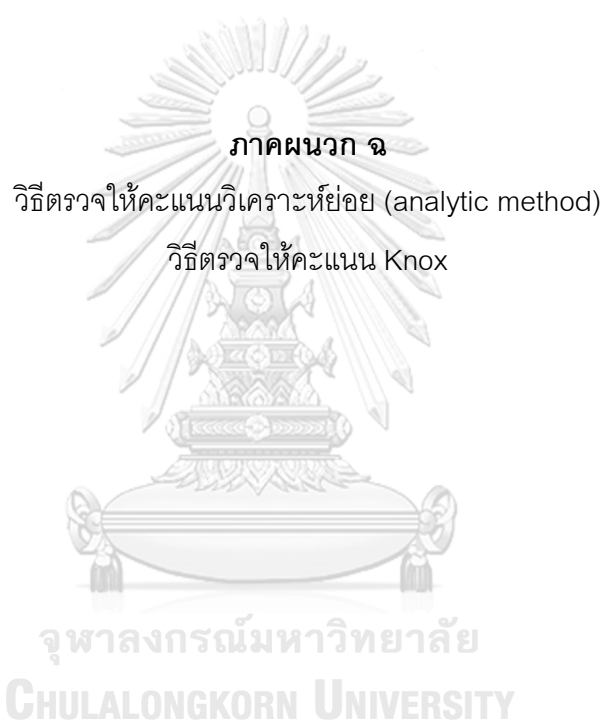
แบบบันทึกคะแนน

แบบสอบอัตรานัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์ โดยใช้วิธีตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric)

คำชี้แจง 1. เกณฑ์ตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric) ในแต่ละเหตุการณ์เป็นดังนี้ สำหรับชั้นที่ 1 ของข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 แบ่งออกเป็น 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) และการเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) จากนั้นแบ่งลงคะแนนในชั้นที่ 2 (คะแนนเต็มข้อละ 2 คะแนน) ส่วนข้อคำถามที่ 4 และข้อคำถามที่ 5 แบ่งออกเป็น 3 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) และการสื่อความหมายทางคณิตศาสตร์ จากนั้นแบ่งลงคะแนนในชั้นที่ 2 (คะแนนเต็มข้อละ 3 คะแนน) รวมคะแนนเต็มแต่ละเหตุการณ์ 12 คะแนน

2. ให้ผู้ตรวจใช้เกณฑ์การตรวจให้คะแนนรูบริกสองชั้น (double layer scoring rubric) พิจารณาคำตอบของผู้เรียนแล้วตัดสินคุณภาพตามคุณภาพของคำตอบโดยใช้คะแนนลงในช่องตาราง

ลำดับที่	ข้อคำถามที่ 1		ข้อคำถามที่ 2		ข้อคำถามที่ 3		ข้อคำถามที่ 4		ข้อคำถามที่ 5		รวม (12 คะแนน)
	ชั้นที่ 1	ชั้นที่ 2	ชั้นที่ 1	ชั้นที่ 2	ชั้นที่ 1	ชั้นที่ 2	ชั้นที่ 1	ชั้นที่ 2	ชั้นที่ 1	ชั้นที่ 2	
	การเข้าใจสิ่งที่ต้องการวัด	การเข้าใจสิ่งที่ต้องการวัด	การเข้าใจสิ่งที่ต้องการวัด	การเข้าใจสิ่งที่ต้องการวัด	การเข้าใจสิ่งที่ต้องการวัด	การเข้าใจสิ่งที่ต้องการวัด	การเข้าใจสิ่งที่ต้องการวัด	การสื่อความหมายทางคณิตศาสตร์	การสื่อความหมายทางคณิตศาสตร์	การเข้าใจสิ่งที่ต้องการวัด	
1											
2											
3											
4											
5											



ภาคผนวก จ

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method)

วิธีตรวจให้คะแนน Knox

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

เหตุการณ์ที่ 1

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 1

องค์ประกอบ	2	1	0
<p>ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สกุลเงินบาทเป็นสกุลเงินตั้งต้นที่นำไปแลก และ</p> <p>2) แลกเป็นเงินสกุลอื่น ๆ ในมูลค่าเท่ากัน</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สกุลเงินบาทเป็นสกุลเงินตั้งต้นที่นำไปแลก หรือ</p> <p>2) แลกเป็นเงินสกุลอื่น ๆ ในมูลค่าเท่ากัน</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- กำหนดปัญหาภายใต้บริบทของสถานการณ์</p>	<p>กำหนดปัญหาได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>ปัญหาเกี่ยวข้องกับการเตรียมเงินบาทไปแลกเป็นสกุลเงินอื่น ๆ ภายใต้บริบทของสถานการณ์ (สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ)</p>	<p>กำหนดปัญหาผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>ปัญหาเกี่ยวข้องกับการเตรียมเงินบาทไปแลกเป็นสกุลเงินอื่น ๆ นอกเหนือจากบริบทของสถานการณ์</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 1

2	1	0
อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ตามแนวคำตอบ คือ สกุลเงินบาทเป็นสกุลเงินตั้ง ต้นที่นำไปแลก และ แลกเปลี่ยนเงินสกุลอื่น ๆ ใน มูลค่าเท่ากัน	อธิบายผิดบางส่วน หรือไม่ ครบถ้วนสมบูรณ์ตามแนว คำตอบ คือ สกุลเงินบาทเป็นสกุลเงินตั้งต้น ที่นำไปแลก หรือ แลกเปลี่ยนเงินสกุลอื่น ๆ ในมูลค่า เท่ากัน	ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 2

องค์ประกอบ	2	1	0
ความถูกต้องของ เนื้อหา - ระบุข้อมูล และอธิบาย เงื่อนไขภายใต้บริบทของ สถานการณ์	ระบุข้อมูลภายใต้ บริบทของ สถานการณ์ได้ ถูกต้อง และ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ 1) ต้องเตรียมสกุลเงิน บาทไปแลกในปริมาณ ที่ไม่เท่ากัน และ 2) สกุลเงินที่ต้องการ แลกเปลี่ยนมูลค่าเท่ากัน	ระบุข้อมูลภายใต้ บริบทของ สถานการณ์ผิด บางส่วน หรือไม่ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ 1) ต้องเตรียม สกุลเงินบาทไปแลก ใน ปริมาณ ที่ ไม่ เท่ากัน หรือ 2) สกุลเงินที่ต้องการ แลกเปลี่ยนมูลค่าเท่ากัน	ไม่ถูกต้องตาม แนวคำตอบ หรือ ไม่มีการตอบ คำถาม
การเข้าใจในสิ่งที่ ต้องการวัด - กำหนดสมมติฐาน ภายใต้บริบทของ สถานการณ์	กำหนดสมมติฐานได้ ถูกต้อง และ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ คาดคะเนปริมาณเงิน บาทที่นำไปแลกได้ว่า	กำหนดสมมติฐาน ผิดบางส่วน หรือไม่ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ ปัญหาเกี่ยวข้องกับ	ไม่ถูกต้องตาม แนวคำตอบ หรือ ไม่มีการตอบ คำถาม

องค์ประกอบ	2	1	0
	ต้องเตรียมไป ใน ปริมาณที่ไม่เท่ากัน เพื่อนำไปแลกเปลี่ยนสกุลเงินอื่น ๆ ภายใต้บริบท <u>ของสถานการณ์</u> (สกุลเงิน สกกุลหยวน และ สกกุลดอลลาร์สหรัฐ)	การเตรียมเงินบาทไป แลกเป็นสกุลเงิน อื่น ๆ <u>นอกเหนือจาก</u> <u>บริบท</u> ของ สถานการณ์	

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 2

2	1	0
อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ตามแนวคำตอบ คือ การจะแลกเปลี่ยนในสกุลต่าง ๆ ให้มีมูลค่าเท่า ๆ กัน ต้องเตรียมเงินบาทไปแลกเปลี่ยนใน ปริมาณที่ไม่เท่ากัน	อธิบายผิดบางส่วน หรือไม่ ครบถ้วนสมบูรณ์ ตามแนว คำตอบ คือ เตรียมเงินบาทไป แลกใน ปริมาณที่ไม่เท่ากัน	ไม่ถูกต้องตามแนวคำตอบ <u>หรือ</u> ไม่มีการตอบคำถาม

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 3

องค์ประกอบ	2	1	0
ความถูกต้องของ <u>เนื้อหา</u> - ระบุข้อมูล และอธิบาย เงื่อนไขภายใต้บริบทของ สถานการณ์	ระบุข้อมูลภายใต้ <u>บริบท</u> ของ สถานการณ์ ได้ ถูกต้อง และ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ 1) แนวทางการ รวบรวม ข้อมูล ที่ เกี่ยวข้องกับอัตราการใช้	ระบุข้อมูลภายใต้ <u>บริบท</u> ของ สถานการณ์ ผิด บางส่วน หรือไม่ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ 1) แนวทางการ รวบรวม ข้อมูล ที่ เกี่ยวข้องกับอัตราการใช้	ไม่ถูกต้องตาม แนวคำตอบ <u>หรือ</u> ไม่มีการตอบ คำถาม

องค์ประกอบ	2	1	0
	แลกเปลี่ยนสกุลเงิน <u>และ</u> 2) อธิบายเหตุผลในการใช้แนวทางได้อย่างถูกต้อง	แลกเปลี่ยนสกุลเงิน <u>หรือ</u> 2) อธิบายเหตุผลในการใช้แนวทางได้อย่างถูกต้อง	
การเข้าใจในสิ่งที่ต้องการวัด - กำหนดแนวทางการรวบรวมข้อมูลภายใต้บริบทของสถานการณ์	กำหนดแนวทางการรวบรวมข้อมูลได้ถูกต้อง <u>และ</u> ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ การรวบรวมข้อมูลที่เกี่ยวข้องกับอัตราการแลกเปลี่ยนระหว่างสกุลเงินภายใต้บริบทของสถานการณ์	กำหนดแนวทางการรวบรวมข้อมูลผิดพลาดบางส่วน <u>หรือ</u> ไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ การรวบรวมข้อมูลที่เกี่ยวข้องกับอัตราการแลกเปลี่ยนระหว่างสกุลเงินนอกเหนือจากบริบทของสถานการณ์	ไม่ถูกต้องตามแนวคำตอบ <u>หรือ</u> ไม่มีการตอบคำถาม

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 3

2	1	0
อธิบายได้ถูกต้อง ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ ดูจากตารางแสดงอัตราการแลกเปลี่ยนเงินระหว่างประเทศ เพื่อเปรียบเทียบอัตราการแลกเปลี่ยนในแต่ละสกุลเงิน <u>หรือ</u> หากไม่มีตารางแสดงอัตราการแลกเปลี่ยนเงินระหว่าง	อธิบายผิดพลาดบางส่วน <u>หรือ</u> ไม่ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ ดูจากตารางแสดงอัตราการแลกเปลี่ยนเงินระหว่างประเทศ <u>หรือ</u> หากไม่มีตารางแสดงอัตราการแลกเปลี่ยนเงินระหว่างประเทศ จะต้องใช้ข้อมูล	ไม่ถูกต้องตามแนวคำตอบ <u>หรือ</u> ไม่มีการตอบคำถาม

2	1	0
ประเทศ จะต้องใช้ข้อมูลที่เกี่ยวข้องเพื่อเปรียบเทียบและสรุปผลอัตราการแลกเปลี่ยน		

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 4

องค์ประกอบ	2	1	0
<p>ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) อัตราการแลกเปลี่ยนแต่ละคู่สกุลเงินตามที่สถานการณ์กำหนด และ</p> <p>2) เปรียบเทียบอัตราแลกเปลี่ยนค่าเงินครบทุกสกุลเงิน (สกุลบาท สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ)</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ผิดบางส่วน หรือ ไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) อัตราการแลกเปลี่ยนแต่ละคู่สกุลเงินตามที่สถานการณ์กำหนด หรือ</p> <p>2) เปรียบเทียบอัตราแลกเปลี่ยนค่าเงินครบ 4 สกุลเงิน ภายใต้บริบทของสถานการณ์ไม่ถูกต้อง</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- วิเคราะห์ข้อมูลภายใต้บริบทของสถานการณ์</p>	<p>วิเคราะห์ข้อมูลได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ เปรียบเทียบอัตราค่าเงินแลกเปลี่ยนทุกสกุลเงินภายใต้บริบทของ<u>สถานการณ์</u></p>	<p>วิเคราะห์ข้อมูลผิดบางส่วน หรือ ไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ เปรียบเทียบอัตราค่าเงินแลกเปลี่ยนแต่มีสกุลใดสกุลหนึ่งนอกเหนือจาก<u>บริบทของสถานการณ์</u></p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>
<p>การสื่อความหมายทาง</p>	<p>เปลี่ยนประโยคภาษาเป็นประโยคทาง</p>	<p>เปลี่ยนประโยคภาษาเป็นประโยคทาง</p>	<p>ไม่ถูกต้องตามแนวคำตอบ</p>

องค์ประกอบ	2	1	0
คณิตศาสตร์ - เปลี่ยนประโยค ภาษาเป็นประโยค ทางคณิตศาสตร์ และคำนวณได้ ถูกต้อง	คณิตศาสตร์ และ คำนวณได้ถูกต้อง คือ 1) เขียนสัญลักษณ์ อัตราส่วนในการ เปรียบเทียบ <u>และ</u> 2) คำนวณเป็นอัตราส่วน อย่างต่ำได้ถูกต้อง	คณิตศาสตร์ และ คำนวณผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ 1) เขียนสัญลักษณ์ อัตราส่วนในการ เปรียบเทียบ <u>หรือ</u> 2) คำนวณเป็นอัตราส่วน อย่างต่ำไม่ถูกต้อง	หรือ ไม่มีการตอบ คำถาม

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 4

2	1	0
อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ตามแนวคำตอบ คือ บาท : เยน : หยวน : ดอลลาร์ สหรัฐ = 100 : 350 : 20 : 3	อธิบายผิดบางส่วน หรือไม่ ครบถ้วนสมบูรณ์ ตามแนว คำตอบ คือ บาท : เยน : หยวน : ดอลลาร์ สหรัฐ = 5,000 : 17,500 : 1,000 : 15	ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 5

องค์ประกอบ	2	1	0
<p>ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูลและอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ได้ถูกต้องและครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) เตรียมเงินสกุลบาทไปแลกเงินสกุลต่าง ๆ ได้แก่ สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ สกุลละ 210 หน่วย และ</p> <p>2) อัตราการแลกเปลี่ยนเงินสกุลต่าง ๆ มาจากผลการทดสอบสมมติฐานของสถานการณ์ ดังนี้ บาท : เยน : หยวน : ดอลลาร์สหรัฐ เท่ากับ 100 : 350 : 20 : 3</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ ผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) เตรียมเงินสกุลบาทไปแลกเงินสกุลต่าง ๆ ได้แก่ สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ สกุลละ 210 หน่วย หรือ</p> <p>2) อัตราการแลกเปลี่ยนเงินสกุลต่าง ๆ มาจากผลการทดสอบสมมติฐานของสถานการณ์ ดังนี้ บาท : เยน : หยวน : ดอลลาร์สหรัฐ เท่ากับ 100 : 350 : 20 : 3</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม</p>
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- กำหนดข้อสรุปการเตรียมเงินสกุลบาทไปแลกและสรุปผลสมมติฐาน</p>	<p>เขียนข้อสรุปของการแลกเปลี่ยนและผลการทดสอบสมมติฐานได้ถูกต้องและครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) เตรียมเงินสกุลบาทไปเป็นจำนวน 8,110 บาท และ</p> <p>2) สรุปสมมติฐานว่า การนำสกุลเงินบาทไปแลกเป็นเงินสกุลอื่น ๆ (สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ) ตามที่ต้องการใน</p>	<p>เขียนข้อสรุปของการแลกเปลี่ยนและผลการทดสอบสมมติฐาน ผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ 1) เตรียมเงินสกุลบาทไปเป็นจำนวน 8,110 บาท หรือ</p> <p>2) สรุปสมมติฐานว่า การนำสกุลเงินบาทไปแลกเป็นเงินสกุลอื่น ๆ (สกุลเยน สกุลหยวน และสกุลดอลลาร์สหรัฐ) ตามที่</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม</p>

องค์ประกอบ	2	1	0
	มูลค่าเท่า ๆ กัน จะต้องเตรียมเงินบาทเป็นจำนวนไม่เท่ากัน	ต้องการในมูลค่าเท่า ๆ กัน จะต้องเตรียมเงินบาทเป็นจำนวนไม่เท่ากัน	
<p>การสื่อ ความหมาย ทาง คณิตศาสตร์</p> <p>- เปลี่ยน ประโยคภาษา เป็นประโยค ทาง คณิตศาสตร์ และคำนวณได้ ถูกต้อง</p>	<p>เปลี่ยนประโยคภาษาเป็น ประโยคทางคณิตศาสตร์ ในเนื้อหาที่เกี่ยวข้อง ภายใต้บริบทของ สถานการณ์ ได้ถูกต้อง และครบถ้วนสมบูรณ์ตาม แนวคำตอบ คือ</p> <p>1) คำนวณสัดส่วนในแต่ละ การแลกเปลี่ยนสกุลเงินทุกคู่ ได้อย่างถูกต้อง (บาท-เยน, บาท-หยวน และบาท- ดอลลาร์สหรัฐ) และ</p> <p>2) คำนวณ อัตราการ แลกเปลี่ยนแต่ละคู่สกุลบาท กับสกุลเงินอื่น ๆ ได้อย่าง ถูกต้อง คือ เตรียมเงินบาทไป แลกเปลี่ยนเงินสกุลเยน 60 บาท , เตรียมเงินบาทไปแลกเปลี่ยน เงินสกุลหยวน 1,050 บาท และเตรียมเงินบาทไปแลกเปลี่ยน เป็นเงินสกุลดอลลาร์สหรัฐ 7,000 บาท</p>	<p>เปลี่ยนประโยคภาษาเป็น ประโยคทางคณิตศาสตร์ ในเนื้อหาที่เกี่ยวข้อง ภายใต้บริบทของ สถานการณ์ผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ</p> <p>1) คำนวณสัดส่วนในแต่ละ การแลกเปลี่ยนสกุลเงินทุกคู่ ได้อย่างถูกต้อง (บาท-เยน, บาท-หยวน และบาท- ดอลลาร์สหรัฐ) หรือ</p> <p>2) คำนวณ อัตราการ แลกเปลี่ยนแต่ละคู่สกุลบาท กับสกุลเงินอื่น ๆ ได้อย่าง ถูกต้อง คือ เตรียมเงินบาทไป แลกเปลี่ยนเงินสกุลเยน 60 บาท , เตรียมเงินบาทไปแลกเปลี่ยน เงินสกุลหยวน 1,050 บาท และเตรียมเงินบาทไปแลกเปลี่ยน เป็นเงินสกุลดอลลาร์สหรัฐ 7,000 บาท</p>	<p>ไม่ ถูก ตั้ ง ตาม แ น ว คำตอบ <u>หรือ</u> ไม่มีการตอบ คำถาม</p>

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 5

2	1	0
<p>อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ตามแนวคำตอบ คือ เตรียมเงินสดบาทไป เป็นจำนวน 8,110 บาท และ สรุปสมมติฐานว่า การนำสกุลเงินบาทไปแลกเป็นเงินสดอื่น ๆ ตามที่ต้องการในมูลค่าเท่า ๆ กัน จะต้องเตรียมเงินบาทเป็นจำนวนไม่เท่ากัน</p>	<p>อธิบายผิดบางส่วน หรือไม่ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ เตรียมเงินสดบาทไปเป็นจำนวน 8,110 บาท หรือ สรุปสมมติฐานว่า การนำสกุลเงินบาทไปแลกเป็นเงินสดอื่น ๆ ตามที่ต้องการในมูลค่าเท่า ๆ กัน จะต้องเตรียมเงินบาทเป็นจำนวนไม่เท่ากัน</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม</p>

เหตุการณ์ที่ 2

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 1

องค์ประกอบ	2	1	0
<p>ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ที่ดินรูปสี่เหลี่ยมมุมฉาก และ</p> <p>2) การหาเส้นทแยงมุมของรูปที่ดินในพินัยกรรม</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ที่ดินรูปสี่เหลี่ยมมุมฉาก หรือ</p> <p>2) การหาเส้นทแยงมุมของรูปที่ดินในพินัยกรรม</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- กำหนดปัญหาภายใต้บริบทของสถานการณ์</p>	<p>กำหนดปัญหาได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>เครื่องมือในการวัดและทฤษฎีบททางคณิตศาสตร์ใดบ้างสำหรับการตรวจสอบภายใต้บริบทของสถานการณ์ (คำกล่าวอ้างในพินัยกรรม)</p>	<p>กำหนดปัญหาผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>เครื่องมือในการวัดและทฤษฎีบททางคณิตศาสตร์ใดบ้างสำหรับการตรวจสอบนอกเหนือจากบริบทของสถานการณ์</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 1

2	1	0
<p>อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ภายใต้บริบทของ สถานการณ์ ตามแนว คำตอบ คือ</p> <p>ระบุปัญหาว่า จะต้องใช้ เครื่องมือในการวัดและ</p> <p>ทฤษฎีบททางคณิตศาสตร์</p> <p>ใดบ้างสำหรับการตรวจสอบ</p> <p>ค่ากล่าวอ้างเส้นทแยงมุมของ</p> <p>ที่ดินรูปสี่เหลี่ยมมุมฉากใน</p> <p>พินัยกรรม</p>	<p>อธิบายผิดบางส่วน หรือไม่ ครบถ้วน สมบูรณ์ภายใต้ บริบทของสถานการณ์ตาม แนวคำตอบ คือ</p> <p>ระบุปัญหาว่า จะต้องใช้ เครื่องมือในการวัดหรือ</p> <p>ทฤษฎีบททางคณิตศาสตร์</p> <p>ใดบ้างสำหรับ การ</p> <p>ตรวจสอบค่ากล่าวอ้างเส้น</p> <p>ทแยงมุมของที่ดินรูปสี่เหลี่ยม</p> <p>มุมฉากในพินัยกรรม</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 2

องค์ประกอบ	2	1	0
<p>ความถูกต้องของ เนื้อหา</p> <p>- ระบุข้อมูล และอธิบาย เงื่อนไขภายใต้บริบทของ สถานการณ์</p>	<p>ระบุข้อมูลภายใต้ บริบทของ สถานการณ์ ได้ ถูกต้อง และ ครบถ้วน สมบูรณ์ ตามแนวคำตอบ คือ</p> <p>1) การตรวจสอบความ ยาวของเส้นทแยงมุม ของที่ดินในพินัยกรรม และ</p> <p>2) ที่ดินในพินัยกรรม เป็นรูปสี่เหลี่ยมมุม ฉาก</p>	<p>ระบุข้อมูลภายใต้ บริบทของ สถานการณ์ ผิด บ้างส่วน หรือไม่ ครบถ้วน สมบูรณ์ ตามแนวคำตอบ คือ</p> <p>1) การตรวจสอบ ความยาวของเส้น ทแยงมุมของที่ดินใน พินัยกรรม หรือ</p> <p>2) ที่ดินในพินัยกรรม เป็นรูปสี่เหลี่ยมมุม ฉาก</p>	<p>ไม่ถูกต้องตาม แนวคำตอบ หรือ ไม่มีการตอบ คำถาม</p>

องค์ประกอบ	2	1	0
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- กำหนดสมมติฐานภายใต้บริบทของสถานการณ์</p>	<p>กำหนดสมมติฐานได้ถูกต้องและครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ต้องใช้เครื่องมือวัดประเภทวัดความยาว/วัดขนาดของมุม และ</p> <p>2) ทฤษฎีบททางคณิตศาสตร์เกี่ยวกับสามเหลี่ยม (ทฤษฎีบทพีทาโกรัส/ตรีโกณมิติ) ภายใต้บริบทของสถานการณ์ (คำกล่าวอ้างในพินัยกรรม)</p>	<p>กำหนดสมมติฐานผิดบางส่วนหรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ต้องใช้เครื่องมือวัดประเภทวัดความยาว/วัดขนาดของมุม หรือ</p> <p>2) ทฤษฎีบททางคณิตศาสตร์เกี่ยวกับสามเหลี่ยม (ทฤษฎีบทพีทาโกรัส/ตรีโกณมิติ) นอกเหนือจากบริบทของสถานการณ์</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 2

2	1	0
<p>อธิบายได้ถูกต้อง ครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ</p> <p>การจะตรวจวัดความยาวของเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉากในคำกล่าวอ้างของพินัยกรรมนั้น จะต้องใช้เครื่องมือวัดประเภทวัดความยาว/วัดขนาดของมุม</p>	<p>อธิบายผิดบางส่วนหรือไม่ครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ</p> <p>การจะตรวจวัดความยาวของเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉากในคำกล่าวอ้างของพินัยกรรมนั้น จะต้องใช้เครื่องมือวัดประเภทวัดความยาว/วัดขนาดของมุม</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>

2	1	0
และ ทฤษฎีบททางคณิตศาสตร์ เกี่ยวกับสามเหลี่ยม (ทฤษฎี บทพีทาโกรัส/ ตรีโกณมิติ)	หรือ ทฤษฎีบททางคณิตศาสตร์ เกี่ยวกับสามเหลี่ยม (ทฤษฎีบท พีทาโกรัส/ตรีโกณมิติ)	

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 3

องค์ประกอบ	2	1	0
ความถูกต้องของ เนื้อหา - ระบุข้อมูล และอธิบาย เงื่อนไขภายใต้บริบทของ สถานการณ์	ระบุข้อมูลภายใต้ บริบทของ สถานการณ์ได้ ถูกต้อง และ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ การอธิบายทฤษฎีบท พีทาโกรัส เพื่อ สนับสนุนค่ากล่าวอ้าง ในพินัยกรรมเกี่ยวกับ เส้นทแยงมุมของที่ดิน รูปสี่เหลี่ยมมุมฉาก เพื่อนำไปสู่การ ทดสอบสมมติฐาน	ระบุข้อมูลภายใต้ บริบทของ สถานการณ์ผิด บางส่วน หรือไม่ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ การอธิบายทฤษฎีบท พีทาโกรัส เพื่อนำไปสู่ การทดสอบ สมมติฐาน	ไม่ถูกต้องตาม แนวคำตอบ <u>หรือ</u> ไม่มีการตอบ คำถาม
การเข้าใจในสิ่งที่ ต้องการวัด - กำหนดแนวทางการ รวบรวมข้อมูลภายใต้ บริบทของสถานการณ์	กำหนดแนวทางการ รวบรวมข้อมูลได้ ถูกต้อง และ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ ใช้เครื่องมือวัดความ ยาวที่มี เพื่อวัด ระยะทางตามด้าน	กำหนดแนวทางการ รวบรวมข้อมูลผิด บางส่วน หรือไม่ ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ ใช้เครื่องมือวัดความ ยาวที่มี เพื่อวัด ระยะทางตามด้าน	ไม่ถูกต้องตาม แนวคำตอบ <u>หรือ</u> ไม่มีการตอบ คำถาม

องค์ประกอบ	2	1	0
	กว้าง ด้านยาว ของพื้นที่ภายใต้บริบท ของสถานการณ์ (เครื่องนับก้าว เครื่อง แปลงจำนวนก้าวเป็น ระยะทาง)	กว้าง ด้านยาว ของพื้นที่ นอกเหนือ จากบริบทของ สถานการณ์	

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 3

2	1	0
อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ภายใต้บริบทของ สถานการณ์ ตามแนว คำตอบ คือ ใช้เครื่องมือวัดความยาวที่มี เพื่อวัดระยะทางตามด้าน กว้าง ด้านยาวของพื้นที่ และ ทฤษฎีบทพีทาโกรัส มา สนับสนุนค่ากล่าวอ้างใน พินัยกรรมเกี่ยวกับเส้นทแยง มุมของที่ดินรูปสี่เหลี่ยมมุม ฉาก เพื่อนำไปสู่การทดสอบ สมมติฐาน	อธิบายผิดบางส่วน หรือไม่ ครบถ้วนสมบูรณ์ภายใต้ บริบทของสถานการณ์ตาม แนวคำตอบ คือ ใช้เครื่องมือวัดความยาวที่มี เพื่อวัดระยะทางตามด้านกว้าง ด้านยาวของพื้นที่ หรือ อธิบายทฤษฎีบทพีทาโกรัสมา สนับสนุนค่ากล่าวอ้างใน พินัยกรรมเกี่ยวกับเส้นทแยงมุม ของที่ดินรูปสี่เหลี่ยมมุมฉาก เพื่อนำไปสู่การทดสอบ สมมติฐาน	ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 4

องค์ประกอบ	2	1	0
ความถูกต้องของ เนื้อหา - ระบุข้อมูล และ	ระบุข้อมูลภายใต้บริบท ของสถานการณ์ ได้ ถูกต้อง และครบถ้วน สมบูรณ์ตามแนวคำตอบ	ระบุข้อมูลภายใต้บริบท ของสถานการณ์ ผิด บางส่วน หรือ ไม่ ครบถ้วนสมบูรณ์ตาม	ไม่ถูกต้องตาม แนวคำตอบ หรือ ไม่มีการตอบ

องค์ประกอบ	2	1	0
อธิบายเงื่อนไขภายใต้บริบทของสถานการณ์	คือ 1) ใช้ข้อมูลของความยาวด้านกว้างและด้านยาวของที่ได้จากการรวบรวมข้อมูลในสถานการณ์ และ 2) ใช้ทฤษฎีบทพีทาโกรัสที่กล่าวถึงรูปสามเหลี่ยมมุมฉาก ช่วยตรวจสอบค่ากล่าวอ้าง	แนวคำตอบ คือ 1) ใช้ข้อมูลของความยาวด้านกว้างและด้านยาวของที่ได้จากการรวบรวมข้อมูลในสถานการณ์ หรือ 2) ใช้ทฤษฎีบทพีทาโกรัสที่กล่าวถึง รูปสามเหลี่ยมมุมฉาก ช่วยตรวจสอบค่ากล่าวอ้าง	คำถาม
การเข้าใจในสิ่งที่ต้องการวัด - วิเคราะห์ข้อมูลภายใต้บริบทของสถานการณ์	วิเคราะห์ข้อมูลได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบคือ บอกได้ว่าความยาวด้านตรงข้ามมุมฉากในทฤษฎีบทพีทาโกรัส คือ ความยาวของเส้นทแยงมุมของรูปสี่เหลี่ยมมุมฉากภายใต้บริบทของสถานการณ์ (มรดกที่ดินเป็นรูปสี่เหลี่ยมมุมฉาก)	วิเคราะห์ข้อมูลผิดบางส่วน หรือ ไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ บอกได้ว่าความยาวด้านตรงข้ามมุมฉากในทฤษฎีบทพีทาโกรัส คือ ความยาวของเส้นทแยงมุมของรูปสี่เหลี่ยมมุมฉากนอกเหนือบริบทของสถานการณ์	ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม
การสื่อความหมายทางคณิตศาสตร์ - เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ และคำนวณได้	เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ และคำนวณได้ถูกต้อง คือ 1) เขียนสมการความสัมพันธ์จากทฤษฎีบทพีทาโกรัสได้ และ 2) นำข้อมูลของความยาว	เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ และคำนวณผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ 1) เขียนสมการความสัมพันธ์จากทฤษฎี	ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม

องค์ประกอบ	2	1	0
ถูกต้อง	ด้านกว้างและด้านยาวของที่ได้จากการรวบรวมข้อมูลในสถานการณ์ไปแทนค่าในความสัมพันธ์จากทฤษฎีบทพีทาโกรัส ได้ความยาวเส้นทแยงมุมของพื้นที่นั้นมีค่าประมาณ 1.3017 กิโลเมตร”	บทพีทาโกรัสได้ <u>หรือ</u> 2) นำข้อมูลของความยาวด้านกว้างและด้านยาวของที่ได้จากการรวบรวมข้อมูลในสถานการณ์ไปแทนค่าในความสัมพันธ์จากทฤษฎีบทพีทาโกรัส ได้ความยาวเส้นทแยงมุมของพื้นที่นั้นมีค่าประมาณ 1.3017 กิโลเมตร”	

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 4

2	1	0
อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ 1) เขียนสมการความสัมพันธ์จากทฤษฎีบทพีทาโกรัสได้ เพื่อช่วยในการตรวจสอบค่ากล่าวอ้าง <u>และ</u> 2) นำข้อมูลของความยาวด้านกว้างและด้านยาวของที่ได้จากการรวบรวมข้อมูลในสถานการณ์ไปแทนค่าในสมการความสัมพันธ์ทฤษฎีบทพีทาโกรัสได้ <u>และ</u>	อธิบายผิดบางส่วน <u>หรือไม่</u> ครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ 1) เขียนสมการความสัมพันธ์จากทฤษฎีบทพีทาโกรัสได้เพื่อช่วยในการตรวจสอบค่ากล่าวอ้าง <u>หรือ</u> 2) นำข้อมูลของความยาวด้านกว้างและด้านยาวของที่ได้จากการรวบรวมข้อมูลในสถานการณ์ไปแทนค่าในสมการความสัมพันธ์ทฤษฎีบทพีทาโกรัสได้ <u>หรือ</u>	ไม่ถูกต้องตามแนวคำตอบ <u>หรือ</u> ไม่มีการตอบคำถาม

2	1	0
3) คำนวณความยาวด้านตรงข้ามมุมฉากได้จากทฤษฎีบทพีทาโกรัสและสรุปค่ากล่าวอ้างความยาวของเส้นทแยงมุมของที่ดินได้ว่า “ความยาวด้านตรงข้ามมุมฉากได้จากทฤษฎีบทพีทาโกรัส คือ ความยาวของเส้นทแยงมุมของมรดกที่ดิน และมีค่าประมาณ 1.3017 กิโลเมตร”	3) คำนวณความยาวด้านตรงข้ามมุมฉากได้จากทฤษฎีบทพีทาโกรัสและสรุปค่ากล่าวอ้างความยาวของเส้นทแยงมุมของที่ดินได้ว่า “ความยาวด้านตรงข้ามมุมฉากได้จากทฤษฎีบทพีทาโกรัส คือ ความยาวของเส้นทแยงมุมของมรดกที่ดิน และมีค่าประมาณ 1.3017 กิโลเมตร”	

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 5

องค์ประกอบ	2	1	0
<p>ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูลและอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ได้ถูกต้องและครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ใช้เครื่องมือที่ใช้วัดความยาว ได้แก่ เครื่องนับก้าว เครื่องแปลงจำนวนก้าวเป็นระยะทาง เพื่อวัดระยะทาง และ</p> <p>2) ทฤษฎีบททางคณิตศาสตร์ช่วยในการหาความยาวของเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉาก คือ ทฤษฎีบทพีทาโกรัส เพื่อสนับสนุนความยาวเส้นทแยงมุมของ</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ผิดบางส่วนหรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ใช้เครื่องมือที่ใช้วัดความยาว ได้แก่ เครื่องนับก้าว เครื่องแปลงจำนวนก้าวเป็นระยะทาง เพื่อวัดระยะทาง หรือ</p> <p>2) ทฤษฎีบททางคณิตศาสตร์ช่วยในการหาความยาวของเส้นทแยงมุมของที่ดินรูปสี่เหลี่ยมมุมฉาก คือ ทฤษฎีบทพีทาโกรัส เพื่อสนับสนุนความยาวเส้นทแยงมุมของที่ดิน</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีคำตอบคำถาม</p>

องค์ประกอบ	2	1	0
	ที่ดินรูปสี่เหลี่ยมมุมฉาก ภายใต้เงื่อนไขของ สถานการณ์ (คำกล่าวอ้างใน พินัยกรรม)	นอกเหนือเงื่อนไขของ สถานการณ์	
<p>การเข้าใจใน สิ่งที่ต้องการ วัด</p> <p>- กำหนดข้อ สรุปผลการ ทดสอบ สมมติฐานและ คำกล่าวอ้างใน พินัยกรรม</p>	<p>เขียนข้อสรุปของสรุปผล การทดสอบสมมติฐานและ คำกล่าวอ้างในพินัยกรรม คือ 1) สรุปผลการทดสอบ สมมติฐานได้ว่า “การจะ ตรวจสอบคำกล่าวอ้างใน พินัยกรรมเกี่ยวกับเส้นทแยง มุมของมรดกที่ดินรูปสี่เหลี่ยม มุมฉากนั้น สามารถใช้ เครื่องมือวัดความยาว และ ทฤษฎีบท พิทาโกรัส ตรวจสอบความยาวเส้น ทแยงมุมพื้นที่ได้จริง” และ</p> <p>2) สรุปคำกล่าวอ้างความ ยาวของ เส้นทแยง มุมของที่ดินในพินัยกรรมได้ ว่า “เส้นทแยงมุมของที่ดินใน คำกล่าวอ้างตามพินัยกรรม เท่ากับ 1.3 กิโลเมตร เป็น ความจริง”</p>	<p>เขียนข้อสรุปของสรุปผล การทดสอบสมมติฐานและ คำกล่าวอ้างในพินัยกรรม ผิดบางส่วน หรือไม่ครบถ้วน สมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สรุปผลการทดสอบ สมมติฐานได้ว่า “การจะ ตรวจสอบคำกล่าวอ้างใน พินัยกรรมเกี่ยวกับเส้นทแยง มุมของมรดกที่ดินรูปสี่เหลี่ยม มุมฉากนั้น สามารถใช้ เครื่องมือวัดความยาว และ ทฤษฎีบทพิทาโกรัสตรวจสอบ ความยาวเส้นทแยงมุมพื้นที่ได้ จริง” หรือ</p> <p>2) สรุปคำกล่าวอ้างความยาว ของเส้นทแยงมุมของที่ดินใน พินัยกรรมได้ว่า “เส้นทแยงมุม ของที่ดินในคำกล่าวอ้างตาม พินัยกรรม เท่ากับ 1.3 กิโลเมตร เป็นความจริง”</p>	<p>ไม่ถูกต้อง ตามแนว คำตอบ หรือ ไม่มี การ ตอบคำถาม</p>
<p>การสื่อ ความหมาย</p>	<p>เปลี่ยนประโยคภาษาเป็น ประโยคทางคณิตศาสตร์</p>	<p>เปลี่ยนประโยคภาษาเป็น ประโยคทางคณิตศาสตร์ใน</p>	<p>ไม่ถูกต้อง ตามแนว</p>

องค์ประกอบ	2	1	0
ทาง คณิตศาสตร์ - เปลี่ยน ประโยคภาษา เป็นประโยค ทาง คณิตศาสตร์ และคำนวณได้ ถูกต้อง	<p>ในเนื้อหาที่เกี่ยวข้อง ภายใต้บริบทของ สถานการณ์ ได้ถูกต้อง และครบถ้วนสมบูรณ์ตาม แนวคำตอบ คือ</p> <p>1) ใช้หลักการประมาณ (\approx) ค่าความยาวของเส้นทแยง มุมที่ได้จากการคำนวณใน สถานการณ์ของมรดกที่ดินให้ เป็นจำนวนเต็ม และ</p> <p>2) ผลการประมาณค่าความ ยาวของเส้นทแยงมุมของ พื้นที่ $1.30173 \approx 13$ กิโลเมตร</p>	<p>เนื้อหาที่เกี่ยวข้องภายใต้ บริบทของสถานการณ์ผิด บางส่วน หรือไม่ครบถ้วน สมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ใช้หลักการประมาณ (\approx) ค่าความยาวของเส้นทแยงมุม ที่ได้จากการคำนวณใน สถานการณ์ของมรดกที่ดินให้ เป็น จำนวนเต็ม หรือ</p> <p>2) ผลการประมาณค่าความ ยาวของเส้นทแยงมุมของพื้นที่ $1.30173 \approx 13$ กิโลเมตร</p>	ค ำ ต อ บ หรือ ไม่ มี ก า ร ต อ บ ค ำ ถ าม

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 5

2	1	0
<p>อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ภายใต้บริบทของ สถานการณ์ ตามแนว คำตอบ คือ</p> <p>ผลการทดสอบสมมติฐาน สรุปได้ว่า “การจะตรวจสอบ ค่ากล่าวอ้างในพินัยกรรม เกี่ยวกับเส้นทแยงมุมของ มรดกที่ดินรูปสี่เหลี่ยมมุมฉาก นั้น สามารถใช้เครื่องมือวัด ความยาว และทฤษฎีบทพีทา โกรัสตรวจสอบความยาวเส้น ทแยงมุมพื้นที่ได้จริง” และ</p>	<p>อธิบายผิดบางส่วน หรือไม่ ครบถ้วนสมบูรณ์ภายใต้ บริบทของสถานการณ์ตาม แนวคำตอบ คือ</p> <p>ผลการทดสอบสมมติฐานสรุป ได้ว่า “การจะตรวจสอบค่า กล่าวอ้างในพินัยกรรมเกี่ยวกับ เส้นทแยงมุมของมรดกที่ดินรูป สี่เหลี่ยมมุมฉากนั้น สามารถใช้ เครื่องมือวัดความยาว และ ทฤษฎีบทพีทาโกรัสตรวจสอบ ความยาวเส้นทแยงมุมพื้นที่ได้ จริง”</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม</p>

2	1	0
<p>ค่ากล่าวอ้างความยาวของเส้นทแยงมุมของที่ดินในพินัยกรรมสรุปได้ว่า “เส้นทแยงมุมของที่ดินในค่ากล่าวอ้างตามพินัยกรรม เท่ากับ 1.3 กิโลเมตร เป็นความจริง”</p>	<p>หรือ ค่ากล่าวอ้างความยาวของเส้นทแยงมุมของที่ดินในพินัยกรรมสรุปได้ว่า “เส้นทแยงมุมของที่ดินในค่ากล่าวอ้างตามพินัยกรรม เท่ากับ 1.3 กิโลเมตร เป็นความจริง”</p>	



เหตุการณ์ที่ 3

วิธีตรวจวิเคราะห์แบบวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 1

องค์ประกอบ	2	1	0
<p>ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ความรู้ทางคณิตศาสตร์ที่ใช้หาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริมาตรและพื้นที่ผิวข้างของปริซึม และ</p> <p>2) อธิบายความรู้ทางคณิตศาสตร์ที่นำมาใช้</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>ความรู้ทางคณิตศาสตร์ที่ใช้หาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกสัมพันธ์กับปริมาตรและพื้นที่ผิวข้างของปริซึม</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- กำหนดปัญหภายใต้บริบทของสถานการณ์</p>	<p>กำหนดปัญหาได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>ต้องใช้ความรู้ทางคณิตศาสตร์เรื่องใดภายใต้บริบทของสถานการณ์ (การหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริมาตรและพื้นที่ผิวข้างของปริซึม)</p>	<p>กำหนดปัญหาผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>ต้องใช้ความรู้ทางคณิตศาสตร์เรื่องใดนอกเหนือจากบริบทของสถานการณ์</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 1

2	1	0
<p>อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ภายใต้บริบทของ สถานการณ์ ตาม แนว คำตอบ คือ</p> <p>ระบุปัญหาว่าต้องใช้ความรู้ ทางคณิตศาสตร์เรื่องใด เพื่อ หาความสัมพันธ์ระหว่าง ปริมาตรและพื้นที่ผิวข้างของ ทรงกระบอกกับปริซึม และ ความรู้ทางคณิตศาสตร์ที่ใช้นั้นกล่าวไว้ว่าอย่างไรบ้าง</p>	<p>อธิบายผิดบางส่วน หรือไม่ ครบถ้วน สมบูรณ์ ภายใต้ บริบทของสถานการณ์ตาม แนวคำตอบ คือ</p> <p>ระบุปัญหาว่า ต้องใช้ความรู้ ทางคณิตศาสตร์เรื่องใด เพื่อหา ความสัมพันธ์ระหว่างปริมาตร และ พื้นที่ผิวข้างของ ทรงกระบอกกับปริซึม</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 2

องค์ประกอบ	2	1	0
<p>ความถูกต้องของ เนื้อหา</p> <p>- ระบุข้อมูล และอธิบาย เงื่อนไขภายใต้บริบทของ สถานการณ์</p>	<p>ระบุข้อมูลภายใต้ บริบทของ สถานการณ์ได้ ถูกต้อง และ ครบถ้วน สมบูรณ์ ตามแนวคำตอบ คือ</p> <p>1) สูตรการหาปริมาตร และ 2) พื้นที่ผิวข้าง ของทรงกระบอก มี ความสัมพันธ์กับปริซึม</p>	<p>ระบุข้อมูลภายใต้ บริบทของ สถานการณ์ผิด บางส่วน หรือไม่ ครบถ้วน สมบูรณ์ ตามแนวคำตอบ คือ</p> <p>1) สูตรการหา ปริมาตร หรือ</p> <p>2) พื้นที่ผิวข้างของ ทรงกระบอก มี ความสัมพันธ์กับ ปริซึม</p>	<p>ไม่ถูกต้องตาม แนวคำตอบ หรือ</p> <p>ไม่มีการตอบ คำถาม</p>

องค์ประกอบ	2	1	0
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- กำหนดสมมติฐานภายใต้บริบทของสถานการณ์</p>	<p>กำหนดสมมติฐานได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สูตรการหาปริมาตร และ 2) พื้นที่ผิวข้างของทรงกระบอกใช้วิธีการหา/คำนวณ/สูตรเดียวกันกับปริซึมภายใต้บริบทของสถานการณ์ (สูตรการหาปริมาตรและพื้นที่ผิวข้างของปริซึม)</p>	<p>กำหนดปัญหาผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สูตรการหาปริมาตร หรือ</p> <p>2) พื้นที่ผิวข้างของทรงกระบอกใช้วิธีการหา/คำนวณ/สูตรเดียวกันกับปริซึม นอกเหนือบริบทจากสถานการณ์</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 2

2	1	0
<p>อธิบายได้ถูกต้อง ครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ</p> <p>ตั้งสมมติฐานไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม</p> <p>และ</p> <p>จะต้องใช้วิธีการหา/คำนวณ/สูตรเดียวกันกับสูตรการหาปริมาตรและพื้นที่ผิวข้างของปริซึม</p>	<p>อธิบายผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ</p> <p>ตั้งสมมติฐานไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม</p> <p>หรือ</p> <p>สูตรการหาปริมาตรและพื้นที่ผิวของทรงกระบอกใช้วิธีการหา/คำนวณ/สูตรเดียวกันกับสูตรการหาปริมาตรและพื้นที่ผิวข้างของปริซึม</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีการตอบคำถาม</p>

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 3

องค์ประกอบ	2	1	0
<p>ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูล และอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ ค้นหาสูตรการคำนวณ ปริมาตร และพื้นที่ผิวข้างของปริซึม ว่ากล่าวไว้ว่าอย่างไร</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ผิดบางส่วน หรือไม่ ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ ค้นหาสูตรการคำนวณปริมาตร และพื้นที่ผิวข้างของปริซึม</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือไม่มีคำตอบ</p>
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- กำหนดแนวทางการรวบรวมข้อมูลภายใต้บริบทของสถานการณ์</p>	<p>กำหนดแนวทางการรวบรวมข้อมูลได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ เปรียบเทียบลักษณะทางกายภาพของทรงกระบอก ว่ามีความสัมพันธ์กับปริซึมอย่างไร ภายใต้บริบทของสถานการณ์ (เพื่อรวบรวมข้อมูลความเหมือนและความต่างของปริซึมที่เรียนไปแล้ว นำไปสู่การทดสอบสมมติฐาน)</p>	<p>กำหนดแนวทางการรวบรวมข้อมูลผิดบางส่วน หรือไม่ ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ เปรียบเทียบลักษณะทางกายภาพของทรงกระบอกกับปริซึมหรือนอกเหนือจากบริบทของสถานการณ์</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือไม่มีคำตอบ</p>

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 3

2	1	0
<p>อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ</p> <p>เปรียบเทียบลักษณะทางกายภาพของทรงกระบอกที่มีความสัมพันธ์กับปริซึมอย่างไร และ</p> <p>ค้นหาสูตรการคำนวณปริมาตร และพื้นที่ผิวข้างของปริซึม ว่ากล่าวได้ว่าอย่างไร</p>	<p>อธิบายผิดบางส่วน หรือไม่ ครบถ้วน สมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ</p> <p>เปรียบเทียบลักษณะทางกายภาพของทรงกระบอกที่มีความสัมพันธ์กับปริซึมอย่างไร หรือ</p> <p>ค้นหาสูตรการคำนวณปริมาตร และพื้นที่ผิวข้างของปริซึม ว่ากล่าวได้ว่าอย่างไร</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม</p>

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 4

องค์ประกอบ	2	1	0
<p>ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูลและอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ได้ถูกต้อง และ ครบถ้วน สมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สำหรับการหาสูตรปริมาตรของทรงกระบอก จากการเปรียบเทียบลักษณะภายนอกของทรงกระบอกและปริซึม สามารถบอกได้ว่า ทรงกระบอกมีหน้าตัดเป็นรูปเดียวกันกับปริซึม คือ รูปวงกลม และ</p> <p>2) สำหรับการหาสูตรพื้นที่ผิวข้างของทรงกระบอก จากรูปคลี่ด้านข้างของทรงกระบอกเป็นรูป</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ผิดบางส่วน หรือไม่ ครบถ้วน สมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สำหรับการหาสูตรปริมาตรของทรงกระบอก จากการเปรียบเทียบลักษณะภายนอกของทรงกระบอกและปริซึม สามารถบอกได้ว่า ทรงกระบอกมีหน้าตัดเป็นรูปเดียวกันกับปริซึม คือ รูปวงกลม หรือ 2) สำหรับการหาสูตรพื้นที่ผิวข้างของทรงกระบอก จากรูปคลี่ด้านข้างของทรงกระบอกเป็นรูป</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม</p>

องค์ประกอบ	2	1	0
	สี่เหลี่ยมมุมฉาก	สี่เหลี่ยมมุมฉาก	
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- จัดระบบข้อมูลภายใต้บริบทของสถานการณ์</p>	<p>จัดระบบข้อมูลได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สำหรับการหาสูตรปริมาตรของทรงกระบอก สามารถสร้างสูตรการหาปริมาตรทรงกระบอกได้จากความสัมพันธ์ของสูตรการหาปริมาตรปริซึม และลักษณะภายนอกของทรงกระบอก กับปริซึม และ</p> <p>2) สำหรับการหาสูตรพื้นที่ผิวข้างของทรงกระบอก สามารถสร้างสูตรการหาพื้นที่ผิวข้างของทรงกระบอกได้จากความสัมพันธ์ของสูตรการหาพื้นที่ผิวข้างของปริซึมกับรูปคลี่ของทรงกระบอก ภายใต้บริบทของสถานการณ์ (สมมติฐานที่ตั้งไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม)</p>	<p>จัดระบบข้อมูลผิดบางส่วนหรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สำหรับการหาสูตรปริมาตรของทรงกระบอก สามารถสร้างสูตรการหาปริมาตรทรงกระบอกได้จากความสัมพันธ์ของสูตรการหาปริมาตรปริซึม และลักษณะภายนอกของทรงกระบอก กับปริซึม หรือ</p> <p>2) สำหรับการหาสูตรพื้นที่ผิวข้างของทรงกระบอก สามารถสร้างสูตรการหาพื้นที่ผิวข้างของทรงกระบอกได้จากความสัมพันธ์ของสูตรการหาพื้นที่ผิวข้างของปริซึมกับรูปคลี่ของทรงกระบอก นอกเหนือบริบทของสถานการณ์</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือไม่มีคำตอบคำถาม</p>
<p>การสื่อความหมายทางคณิตศาสตร์</p> <p>- เปลี่ยนประโยคภาษา</p>	<p>เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ และคำนวณได้ถูกต้อง คือ</p> <p>1) สำหรับการหาสูตรปริมาตรของทรงกระบอก สามารถระบุได้ว่าพื้นที่ฐานของทรงกระบอก คือพื้นที่รูปวงกลม (πr^2) และความ</p>	<p>เปลี่ยนประโยคภาษาเป็นประโยคทางคณิตศาสตร์ และคำนวณผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) สำหรับการหาสูตรปริมาตรของทรงกระบอก สามารถระบุ</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือไม่มี</p>

องค์ประกอบ	2	1	0
เป็นประโยคทางคณิตศาสตร์และคำนวณได้ถูกต้อง	<p>สูงของทรงกระบอก คือ ความสูงของปริซึม (h) ดังนั้น สูตรปริมาตรของทรงกระบอก = (พื้นที่ฐาน) x ความสูง = $\pi r^2 h$ และ</p> <p>2) สำหรับการหาสูตรพื้นที่ผิวข้างของทรงกระบอก สามารถระบุได้ว่า ความยาวด้านหนึ่งของรูปคลี่จะเท่ากับความยาวเส้นรอบวงของวงกลม ($2\pi r$) และความยาวอีกด้านหนึ่งจะเท่ากับความสูงของทรงกระบอก (h)</p> <p>ดังนั้น สูตรพื้นที่ผิวข้างของทรงกระบอก</p> $= (\text{ความยาวรอบฐาน}) \times \text{ความสูง}$ $= 2\pi r h$	<p>ได้ว่า พื้นที่ฐานของทรงกระบอก คือ พื้นที่รูปวงกลม (πr^2) และความสูงของทรงกระบอก คือ ความสูงของปริซึม (h) ดังนั้น สูตรปริมาตรของทรงกระบอก = (พื้นที่ฐาน) x ความสูง = $\pi r^2 h$ หรือ</p> <p>2) สำหรับการหาสูตรพื้นที่ผิวข้างของทรงกระบอก สามารถระบุได้ว่า ความยาวด้านหนึ่งของรูปคลี่จะเท่ากับความยาวเส้นรอบวงของวงกลม ($2\pi r$) และความยาวอีกด้านหนึ่งจะเท่ากับความสูงของทรงกระบอก (h) ดังนั้น สูตรพื้นที่ผิวข้างของทรงกระบอก = (ความยาวรอบฐาน) x ความสูง = $2\pi r h$</p>	การตอบคำถาม

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 4

2	1	0
<p>อธิบายได้ถูกต้อง ครบถ้วนสมบูรณ์ ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ</p> <p>จากสมมติฐานที่ตั้งไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม</p> <p>1) สำหรับการหาสูตรปริมาตรของทรงกระบอก สามารถสร้างสูตรการหา</p>	<p>อธิบายผิดบางส่วน หรือไม่ครบถ้วน สมบูรณ์ ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ คือ</p> <p>จากสมมติฐานที่ตั้งไว้ว่า สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกมีความสัมพันธ์กับปริซึม</p> <p>1) สำหรับการหาสูตรปริมาตรของทรงกระบอก สามารถสร้างสูตรการหา</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ ไม่มีการตอบคำถาม</p>

2	1	0
<p>ปริมาตรทรงกระบอกได้จาก ความสัมพันธ์ของสูตรการหาปริมาตร ปริซึม และลักษณะภายนอกของ ทรงกระบอก กับ ปริซึม <u>พบว่า</u> ทรงกระบอกมีหน้าตัดเป็นรูปเดียวกันกับ ปริซึม คือ รูปวงกลม จึง สามารถระบุได้ว่า พื้นที่ฐานของทรงกระบอก คือ พื้นที่ รูปวงกลม (πr^2) และความสูงของ ทรงกระบอก คือ ความสูงของปริซึม (h) ดังนั้น สูตรปริมาตรของทรงกระบอก</p> $= (\text{พื้นที่ฐาน}) \times \text{ความสูง} = \pi r^2 h$ <p><u>และ</u></p> <p>2) สำหรับการหาสูตรพื้นที่ผิวข้างของ ทรงกระบอก สามารถสร้างสูตรการหา พื้นที่ผิวข้างทรงกระบอกได้จาก ความสัมพันธ์ของสูตรการหาพื้นที่ผิว ข้างของปริซึมกับรูปคลี่ของทรงกระบอก <u>พบว่า</u> รูปคลี่ด้านข้างของทรงกระบอก เป็นรูปสี่เหลี่ยมมุมฉาก จึง สามารถระบุได้ว่า ความยาวด้านหนึ่งของรูปคลี่จะ เท่ากับความยาวเส้นรอบวงของวงกลม ($2\pi r$) และความยาวอีกด้านหนึ่งจะ เท่ากับความสูงของทรงกระบอก (h) ดังนั้น สูตรพื้นที่ผิวข้างของทรงกระบอก</p> $= (\text{ความยาวรอบฐาน}) \times \text{ความสูง} = 2\pi r h$	<p>ปริมาตรทรงกระบอกได้จาก ความสัมพันธ์ของสูตรการหาปริมาตร ปริซึม และลักษณะภายนอกของ ทรงกระบอก กับ ปริซึม <u>พบว่า</u> ทรงกระบอกมีหน้าตัดเป็นรูปเดียวกันกับ ปริซึม คือ รูปวงกลม จึง สามารถระบุได้ว่า พื้นที่ฐานของทรงกระบอก คือ พื้นที่ รูปวงกลม (πr^2) และความสูงของ ทรงกระบอก คือ ความสูงของปริซึม (h) ดังนั้น สูตรปริมาตรของทรงกระบอก</p> $= (\text{พื้นที่ฐาน}) \times \text{ความสูง} = \pi r^2 h$ <p><u>หรือ</u></p> <p>2) สำหรับการหาสูตรพื้นที่ผิวข้างของ ทรงกระบอก สามารถสร้างสูตรการหา พื้นที่ผิวข้างทรงกระบอกได้จาก ความสัมพันธ์ของสูตรการหาพื้นที่ผิว ข้างของปริซึมกับรูปคลี่ของทรงกระบอก <u>พบว่า</u> รูปคลี่ด้านข้างของทรงกระบอก เป็นรูปสี่เหลี่ยมมุมฉาก จึง สามารถระบุได้ว่า ความยาวด้านหนึ่งของรูปคลี่จะ เท่ากับความยาวเส้นรอบวงของวงกลม ($2\pi r$) และความยาวอีกด้านหนึ่งจะ เท่ากับความสูงของทรงกระบอก (h) ดังนั้น สูตรพื้นที่ผิวข้างของทรงกระบอก</p> $= (\text{ความยาวรอบฐาน}) \times \text{ความสูง} = 2\pi r h$	

วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method) ข้อคำถามที่ 5

องค์ประกอบ	2	1	0
<p>ความถูกต้องของเนื้อหา</p> <p>- ระบุข้อมูลและอธิบายเงื่อนไขภายใต้บริบทของสถานการณ์</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกและปริซึมนั้นมีความสัมพันธ์กันโดย</p> <p>1) สูตรปริมาตรของทรงกระบอก คือ $\pi r^2 h$ และ</p> <p>2) สูตรพื้นที่ผิวข้างของทรงกระบอก คือ $2\pi r h$ และ</p> <p>3) กำหนดให้ $\pi = \frac{22}{7}$</p>	<p>ระบุข้อมูลภายใต้บริบทของสถานการณ์ ผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกและปริซึมนั้นมีความสัมพันธ์กันโดย</p> <p>1) สูตรปริมาตรของทรงกระบอก คือ $\pi r^2 h$</p> <p>หรือ</p> <p>2) สูตรพื้นที่ผิวข้างของทรงกระบอก คือ $2\pi r h$</p> <p>หรือ</p> <p>3) กำหนดให้ $\pi = \frac{22}{7}$</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีคำตอบคำถาม</p>
<p>การเข้าใจในสิ่งที่ต้องการวัด</p> <p>- กำหนดข้อสรุปผลการทดสอบสมมติฐาน และสามารถนำข้อมูลที่กำหนดให้ไปแทนค่าในตัวแปรของสูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ผลการทดสอบสมมติฐานสรุปได้ว่า “สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกจะต้องใช้วิธีการ</p>	<p>เขียนข้อสรุปผลการทดสอบสมมติฐาน และสามารถนำข้อมูลที่กำหนดให้ไปแทนค่าในตัวแปรของสูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกได้ถูกต้อง และครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ</p> <p>1) ผลการทดสอบสมมติฐานสรุปได้ว่า “สูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอกจะต้องใช้วิธีการ</p>	<p>เขียนข้อสรุปผลการทดสอบสมมติฐาน และสามารถนำข้อมูลที่กำหนดให้ไปแทนค่าในตัวแปรของสูตรการหาปริมาตรและพื้นที่ผิวข้างของทรงกระบอก ผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ตามแนวคำตอบ คือ 1) ผลการทดสอบสมมติฐานสรุปได้ว่า “สูตรการหาปริมาตรและพื้นที่ผิว</p>	<p>ไม่ถูกต้องตามแนวคำตอบ หรือ</p> <p>ไม่มีคำตอบคำถาม</p>

องค์ประกอบ	2	1	0
<p>แปรของสูตร การหาปริมาตร และพื้นที่ผิว ข้างของ ทรงกระบอก</p>	<p>หา/คำนวณ/สูตร เดียวกันกับ ปริซึม” และ 2) ส าม า ร ถ น ำ ข ้อ มู ล ทรงกระบอกที่ให้ความยาวรัศมี ที่ฐาน (r) 7 เซนติเมตร และ ความสูง (h) 10 เซนติเมตร ไป แทนในสูตรการหาปริมาตรและ พื้นที่ผิวข้างของทรงกระบอก</p>	<p>ข้างของทรงกระบอกจะต้อง ใช้วิธีการหา/คำนวณ/สูตร เดียวกันกับปริซึม” หรือ 2) ส าม า ร ถ น ำ ข ้อ มู ล ทรงกระบอกที่ให้ความยาว รัศมีที่ฐาน (r) 7 เซนติเมตร และ ความ สูง (h) 10 เซนติเมตร ไปแทนในสูตรการ หาปริมาตรและพื้นที่ผิวข้าง ของทรงกระบอก</p>	
<p>ก า ร ส ี อ ความหมาย ท ำ ง คณิตศาสตร์ - เ ป ลี่ ย น ประโยคภาษา เป็นประโยค ท ำ ง คณิตศาสตร์ และคำนวณได้ ถูกต้อง</p>	<p>เปลี่ยนประโยคภาษาเป็น ประโยคทางคณิตศาสตร์ใน เนื้อหาที่เกี่ยวข้องภายใต้ บริบทของสถานการณ์ ได้ ถูกต้อง และครบถ้วน สมบูรณ์ตามแนวคำตอบ คือ 1) สูตรปริมาตรของ ทรงกระบอก คือ $\pi r^2 h = \frac{22}{7} (7)^2 (10)$ $= 1,540$ ลูกบาศก์เซนติเมตร และ 2) สูตรพื้นที่ผิวข้างของ ทรงกระบอก คือ $2\pi rh =$ $2\left(\frac{22}{7}\right)(7)(10)$ $= 440$ ตารางเซนติเมตร</p>	<p>เปลี่ยนประโยคภาษาเป็น ประโยคทางคณิตศาสตร์ ในเนื้อหาที่เกี่ยวข้อง ภายใต้ บริบทของ สถานการณ์ผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ ตามแนวคำตอบ คือ 1) สูตรปริมาตรของทรงกระบอก คือ $\pi r^2 h =$ $\frac{22}{7} (7)^2 (10)$ $= 1,540$ ลูกบาศก์เซนติเมตร หรือ 2) สูตรพื้นที่ผิวข้างของ ทรงกระบอก คือ $2\pi rh =$ $2\left(\frac{22}{7}\right)(7)(10)$ $= 440$ ตารางเซนติเมตร</p>	<p>ไม่ถูกต้อง ตามแนว คำตอบ หรือ ไม่มี การ ตอบคำถาม</p>

วิธีตรวจให้คะแนน Knox ข้อคำถามที่ 5

2	1	0
<p>อธิบายได้ถูกต้อง ครบถ้วน สมบูรณ์ภายใต้บริบทของ สถานการณ์ ตาม แนว คำตอบ คือ</p> <p>ผลการทดสอบสมมติฐานสรุป ได้ว่า “สูตรการหาปริมาตร และ พื้นที่ผิวข้างของ ทรงกระบอกจะต้องใช้วิธีการ หา/คำนวณ/สูตร เดียวกันกับ ปริซึม”</p> <p><u>และ</u></p> <p>เนื่องจากสูตรการหาปริมาตร และ พื้นที่ผิวข้างของ ทรงกระบอกและปริซึมนั้น มีความสัมพันธ์กัน เมื่อ กำหนดให้ $\pi = \frac{22}{7}$ และ ทรงกระบอกมีความยาวรัศมี ที่ฐาน (r) 7 เซนติเมตร และ ความสูง (h) 10 เซนติเมตร ไปแทนในสูตรการหาปริมาตร และ พื้นที่ผิวข้างของ ทรงกระบอก จะได้ว่า สูตร ปริมาตรของทรงกระบอก คือ $\pi r^2 h = \frac{22}{7} (7)^2 (10) = 1,540$ ลูกบาศก์เซนติเมตร <u>และ</u> สูตรพื้นที่ผิวข้างของ ทรงกระบอก คือ $2\pi r h =$</p>	<p>อธิบายผิดบางส่วน หรือไม่ ครบถ้วน สมบูรณ์ภายใต้ บริบทของสถานการณ์ตาม แนวคำตอบ คือ</p> <p>ผลการทดสอบสมมติฐานสรุป ได้ว่า “สูตรการหาปริมาตรและ พื้นที่ผิวข้างของทรงกระบอก จะต้องใช้วิธีการหา/คำนวณ/ สูตร เดียวกันกับปริซึม”</p> <p><u>หรือ</u></p> <p>เนื่องจากสูตรการหาปริมาตร และพื้นที่ผิวข้างของ ทรงกระบอกและปริซึมนั้น มีความสัมพันธ์กัน เมื่อ กำหนดให้ $\pi = \frac{22}{7}$ และ ทรงกระบอกมีความยาวรัศมีที่ ฐาน (r) 7 เซนติเมตร และ ความสูง (h) 10 เซนติเมตร ไป แทนในสูตรการหาปริมาตรและ พื้นที่ผิวข้างของทรงกระบอกจะ ได้ว่า สูตรปริมาตรของ ทรงกระบอก คือ $\pi r^2 h = \frac{22}{7} (7)^2 (10) = 1,540$ ลูกบาศก์เซนติเมตร <u>หรือ</u> สูตรพื้นที่ผิวข้างของ ทรงกระบอก คือ $2\pi r h = 2(\frac{22}{7})(7)(10)$</p>	<p>ไม่ถูกต้องตามแนวคำตอบ <u>หรือ</u> ไม่มีการตอบคำถาม</p>

2	1	0
$2\binom{22}{7}(10)$ = 440 ตารางเซนติเมตร	= 440 ตารางเซนติเมตร	



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาคผนวก ข

แบบบันทึกคะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

โดยใช้วิธีตรวจให้คะแนนวิเคราะห์ย่อย (analytic method)

แบบบันทึกคะแนนแบบสอบอัตนัยประยุกต์วัดทักษะการคิดวิจารณ์ทางคณิตศาสตร์

โดยใช้วิธีตรวจให้คะแนน Knox

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

แบบบันทึกคะแนน

แบบสอบอรรถกถาการตีความทางคณิตศาสตร์ โดยใช้วิธีตรวจวิเคราะห์ย่อย (analytic method)

คำชี้แจง 1. เกณฑ์ตรวจวิเคราะห์ย่อย (analytic method) ในแต่ละเหตุการณ์เป็นดังนี้ สำหรับข้อคำถามที่ 1 ถึงข้อคำถามที่ 3 แบ่งออกเป็น 2 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) และการเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) ส่วนข้อคำถามที่ 4 และข้อคำถามที่ 5 แบ่งออกเป็น 3 องค์ประกอบ คือ ความถูกต้องของเนื้อหา (2 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) และการสื่อความหมายทางคณิตศาสตร์ รวมคะแนนเต็มแต่ละเหตุการณ์ 24 คะแนน

2. ให้ผู้ตรวจใช้เกณฑ์การตรวจวิเคราะห์ย่อย (analytic method) พิจารณาคำตอบของผู้เรียนแล้วตัดสินคุณภาพตามคุณภาพของคำตอบโดยใช้คะแนนลงในตาราง

ลำดับที่	ข้อคำถามที่ 1		ข้อคำถามที่ 2		ข้อคำถามที่ 3		ข้อคำถามที่ 4		ข้อคำถามที่ 5			รวม (24 คะแนน)
	ความถูกต้อง ของเนื้อหา	การเข้าใจสิ่ง ที่ต้องการวัด	ความถูกต้อง ของเนื้อหา	การเข้าใจสิ่ง ที่ต้องการวัด	ความถูกต้อง ของเนื้อหา	การเข้าใจสิ่ง ที่ต้องการวัด	การสื่อ ความหมาย ทาง คณิตศาสตร์	การสื่อ ความหมาย ทาง คณิตศาสตร์	ความ ถูกต้อง ของ เนื้อหา	การเข้าใจสิ่ง ที่ต้องการวัด	การสื่อ ความหมาย ทาง คณิตศาสตร์	
1												
2												
3												
4												
5												

แบบฉบับที่คะแนน
แบบสอบอัตนัยประยุกต์ทักษะการคิดวิเคราะห์โดยใช้วิธีตรวจให้คะแนน Knox

คำชี้แจง 1. เกณฑ์ตรวจให้คะแนน Knox ของแต่ละข้อคำถามในทุกเหตุการณ์เป็น 3 ระดับ ได้แก่ 2 คะแนน คือ ผู้ตอบข้อคำถามได้ถูกต้องและครบถ้วน สมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ 1 คะแนน คือ ผู้ตอบข้อคำถามผิดบางส่วน หรือไม่ครบถ้วนสมบูรณ์ภายใต้บริบทของสถานการณ์ตามแนวคำตอบ และ 0 คะแนน คือ ผู้ตอบข้อคำถามไม่ถูกต้องตามแนวคำตอบ หรือไม่มีการตอบคำถาม รวมคะแนนเต็มแต่ละเหตุการณ์ 10 คะแนน

2. ให้ผู้ตรวจใช้เกณฑ์การตรวจให้คะแนนวิเคราะห์ (analytic method) พิจารณาคำตอบของผู้เรียนแล้วตัดสินคุณภาพตามคุณภาพของคำตอบโดยใส่คะแนนลงในตาราง

ลำดับที่	ข้อคำถามที่ 1 (2 คะแนน)	ข้อคำถามที่ 2 (2 คะแนน)	ข้อคำถามที่ 3 (2 คะแนน)	ข้อคำถามที่ 4 (2 คะแนน)	ข้อคำถามที่ 5 (2 คะแนน)	รวม (10 คะแนน)
1						
2						
3						
4						
5						

ประวัติผู้เขียน

ชื่อ-สกุล	ภักจิรา จงสุกใส
วัน เดือน ปี เกิด	21 มกราคม 2540
วุฒิการศึกษา	สำเร็จการศึกษาคณะครุศาสตร์บัณฑิต (เกียรตินิยมอันดับ 1) สาขาวิชามัธยมศึกษา (วิทยาศาสตร์) วิชาเอกคณิตศาสตร์ (เดี่ยว) ภาควิชาหลักสูตรและการสอน คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2562 ศึกษาต่อในระดับปริญญาโท สาขาวิชาวิธีวิทยาการพัฒน นวัตกรรมทางการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2563



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY