

คุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2
ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน
โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด



เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวิจัยและพัฒนาศักยภาพมนุษย์
กรกฎาคม 2561

คุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2
ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน
โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด



เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวิจัยและพัฒนาศักยภาพมนุษย์

กรกฎาคม 2561

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

คุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2
ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน
โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด



เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวิจัยและพัฒนาศักยภาพมนุษย์
กรกฎาคม 2561

พานุกพงศ์ อึ้งสืบเชื้อ. (2561). คุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียน
ชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกันโดยประยุกต์
ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด. ปรินญานิพนธ์ กศ.ม.(การวิจัย
และพัฒนาศักยภาพมนุษย์). กรุงเทพฯ: บัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒ.
อาจารย์ที่ปรึกษาปรินญานิพนธ์: อาจารย์ ดร.อัญชลี สุขในสิทธิ์, อาจารย์ดร.พินดา ศกุนตนาค.

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อสร้างและตรวจสอบคุณภาพของแบบสอบการรู้เรื่อง
คณิตศาสตร์ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกันโดยการประยุกต์ใช้ทฤษฎี
การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน
400 คน ของโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) ภาคใต้ที่กำลัง
ศึกษาในปีการศึกษา 2560 ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการวิจัย
คือ แบบสอบการรู้เรื่องคณิตศาสตร์ มีลักษณะเป็นข้อสอบสถานการณ์ตามกระบวนการทาง
คณิตศาสตร์ จำนวน 36 ข้อ สถิติที่ใช้ในการวิเคราะห์ข้อมูล สถิติทดสอบ t-test ตามสูตรของเวิร์ส,
ไนส์ และอาร์แมน และการวิเคราะห์หาค่าสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) โดยใช้ทฤษฎี
การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

ผลการวิจัย พบว่า

แบบสอบการรู้เรื่องคณิตศาสตร์มีความเที่ยงตรงเชิงเนื้อหา ความยากง่ายและอำนาจจำแนก
เหมาะสม และมีความเชื่อมั่นอยู่ในระดับสูง เมื่อตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนน
แบบภาพรวม พบว่า เมื่อเพิ่มความยาวแบบสอบจาก 12 ข้อ เป็น 24 ข้อ และ 36 มีค่าสัมประสิทธิ์
การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ 0.868, 0.929 และ 0.952 ตามลำดับ และ
มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ 0.842, 0.914 และ 0.941 และ
ตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับ
การตัดสินใจเชิงสัมพัทธ์เพิ่มจาก 0.889, 0.941 และ 0.960 และ ค่าสัมประสิทธิ์การสรุปอ้างอิง
สำหรับการตัดสินใจเชิงสัมบูรณ์ 0.856, 0.923 และ 0.947 และเมื่อเปรียบเทียบค่าสัมประสิทธิ์
การสรุปอ้างอิงของแบบสอบการรู้เรื่องคณิตศาสตร์ที่ตรวจให้คะแนนรูปแบบเดียวกัน เมื่อระดับ
ความยาวแบบสอบจำนวน 12 ข้อ 24 ข้อ และ 36 ข้อ พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิง
มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

QUALITY OF MATHEMATICS LITERACY TEST FOR MATHAYOMSUKSA 2 STUDENTS
THAT REGRADING DIFFERENT TEST LENGTHS AND SCORING RUBRICS
BY APPLYING GENERALIZABILITY THEORY



Presented in Partial Fulfillment of the Requirements for the
Master of Education degree in Research and Development on Human Potentials
at Srinakharinwirot University
July 2018

Panupong Uengsuebchua. (2018). *Quality of Mathematical Literacy Tests for Mathayomsuksa Two Students Regarding Different Test Lengths and Scoring Rubrics by Applying Generalizability Theory*. Master's thesis (Research and Development on Human Potential). Bangkok: Graduate School. Srinakarinwirot University. Advisor Committee: Dr. Anchalee Suknaisith, Dr. Panida Sakuntanak.

The purpose of this research was to develop and verify the quality of mathematical literacy test of different lengths and using a scoring rubric to apply generalizability theory. The sample used in this research included four hundred students in Grade Eight from southern schools under the Office of the Basic Education Commission (OBEC) in the 2017 academic year and obtained by multi-stage sampling. The instrument was a thirty-six item mathematical literacy test, designed as a situational test based on mathematical processes. The data were analyzed by basics statistics, hypothesis test statistic by Z – test's Wert, Neidt and Ahmann, and G Coefficient.

The results of the study were as follows: the findings indicated that the mathematical literacy test had proper content validity, difficulties, discrimination and high reliability. When considering the generalizability coefficient it was found that the mathematical literacy test quality when the test length was increased from twelve to twenty-four to thirty-six items. The G-coefficient relative decision was on a holistic scoring rubric, which increased from 0.868, 0.929 and 0.952 respectively. The G-coefficient absolute decision also increased from 0.842, 0.914 and 0.941 respectively. The G-coefficient relative decision was on an analytic scoring rubric, which increased from 0.889, 0.941 and 0.960 respectively. The G-coefficient absolute decision also increased from 0.856, 0.923 and 0.947 respectively. In terms of the generalizability coefficient, it was found that the math literacy test, examined by both holistic and analytic rubrics, had when there was a different test length, from twelve to twenty-four to thirty-six items. In comparison, the generalizability coefficient was significantly different at a level of 0.05.

ปริญญานิพนธ์

เรื่อง

คุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2

ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน

โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

ของ

ภานุพงศ์ อังสืบเชื้อ

ได้รับการอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวิจัยและพัฒนาศักยภาพมนุษย์

ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย
(ผู้ช่วยศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล)

วันที่ เดือน กรกฎาคม พ.ศ. 2561

อาจารย์ที่ปรึกษาปริญญานิพนธ์

คณะกรรมการสอบปากเปล่า

.....ที่ปรึกษาหลัก

..... ประธาน

(อาจารย์ ดร.อัญชลี สุขในสิทธิ์)

(รองศาสตราจารย์ ดร.ชัยวิชิต เขียวระนะ)

.....ที่ปรึกษาร่วม

.....กรรมการ

(อาจารย์ ดร.พนิดา ศกุนตนาค)

(อาจารย์ ดร.อัญชลี สุขในสิทธิ์)

.....กรรมการ

(อาจารย์ ดร.พนิดา ศกุนตนาค)

.....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.สุรัชย์ มีชาญ)

ประกาศคุณูปการ

ปริญญานิพนธ์นี้สำเร็จได้ด้วยความกรุณาอย่างสูงจากคณะกรรมการควบคุมปริญญานิพนธ์ อาจารย์ ดร.อัญชลี สุขในสิทธิ์ ประธานควบคุมปริญญานิพนธ์ และอาจารย์ ดร.พนิดา ศกุนตนาค กรรมการควบคุมปริญญานิพนธ์ที่ได้ให้คำปรึกษา ข้อชี้แนะในการจัดทำงานวิจัยนี้ทุกขั้นตอน และแก้ไขข้อบกพร่องต่างๆ ด้วยความเอาใจใส่ ตลอดจนให้กำลังใจเสมอมา ผู้วิจัยขอกราบขอบพระคุณในความกรุณาของท่านทั้งสองด้วยความเคารพอย่างสูง

ขอกราบขอบพระคุณรองศาสตราจารย์ ดร.ชัยวิชิต เชียรชนะ และผู้ช่วยศาสตราจารย์ ดร.สุรัช มีชาญ ที่ให้ความกรุณาเป็นกรรมการ ในการสอบปริญญานิพนธ์และให้ข้อเสนอแนะที่เป็นประโยชน์เพิ่มเติมทำให้ปริญญานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอกราบขอบพระคุณคณาจารย์ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒทุกท่านที่ประสิทธิ์ประสาทวิชาความรู้แก่ผู้วิจัย พร้อมทั้งให้คำปรึกษาและเอาใจใส่อย่างดียิ่งตลอดมา

ขอขอบพระคุณผู้เชี่ยวชาญทุกท่านที่ให้ความช่วยเหลือตรวจสอบคุณภาพของเครื่องมือในการวิจัย ขอขอบพระคุณผู้บริหาร คณาจารย์โรงเรียนกลุ่มตัวอย่างที่เป็นส่วนสำคัญส่วนหนึ่ง สำหรับการวิจัย และขอขอบคุณนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานภาคใต้ ภาคเรียนที่ 2 ปีการศึกษา 2560 จำนวน 400 คนที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูลงานวิจัยอย่างยิ่ง

ท้ายสุดผู้วิจัยขอกราบขอบพระคุณบิดา มารดา และขอบคุณเพื่อนๆ ทุกคนที่ให้กำลังใจที่ดีเยี่ยมตลอดระยะเวลาที่ศึกษาและทำวิจัย จนทำให้ผู้วิจัยประสบความสำเร็จในการศึกษา

ภานุพงศ์ อึ้งสืบเชื้อ

สารบัญ

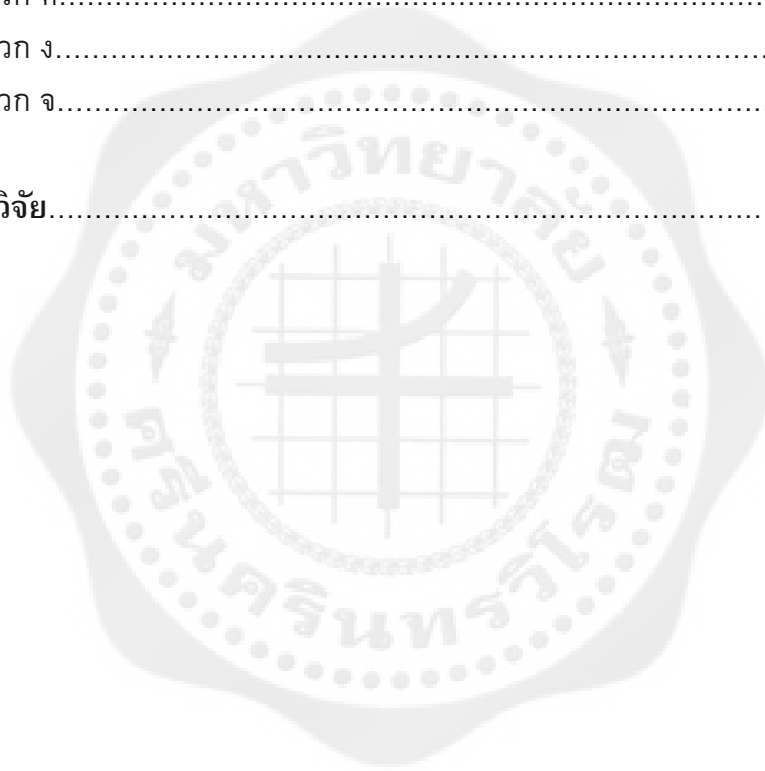
บทที่	หน้า
1 ภูมิหลัง	1
ความมุ่งหมายในการวิจัย	5
ความสำคัญในการวิจัย	5
ขอบเขตของการวิจัย	6
ประชากรที่ใช้ในการวิจัย	6
กลุ่มตัวอย่างที่ใช้ในการวิจัย	6
ตัวแปรที่ศึกษา	6
นิยามศัพท์เฉพาะ	6
กรอบแนวคิดในการวิจัย	9
สมมติฐานการวิจัย	9
2 เอกสารงานวิจัยที่เกี่ยวข้อง	10
แนวคิดและทฤษฎีที่เกี่ยวกับการรู้เรื่องคณิตศาสตร์	11
นิยามการรู้เรื่องคณิตศาสตร์และระดับการรู้เรื่องคณิตศาสตร์	12
การจัดกรอบโครงสร้างการประเมินผลการรู้เรื่องคณิตศาสตร์	15
กระบวนการคิดเป็นคณิตศาสตร์	19
สมรรถนะทางคณิตศาสตร์	21
การประเมินผลด้านการรู้เรื่องคณิตศาสตร์	23
งานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ	37
แนวคิดและทฤษฎีที่เกี่ยวกับกฎเกณฑ์การให้คะแนน	44
ความหมายและความสำคัญของกฎเกณฑ์การให้คะแนน	44
องค์ประกอบของกฎเกณฑ์การให้คะแนน	49
ลักษณะของกฎเกณฑ์การให้คะแนน	50
ประเภทของกฎเกณฑ์การให้คะแนน	51
การสร้างกฎเกณฑ์การให้คะแนน	54
งานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ	57

สารบัญ (ต่อ)

บทที่	หน้า
2 (ต่อ)	
แนวคิดและทฤษฎีที่เกี่ยวกับทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือ ของผลการวัด.....	62
ความเป็นมาของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด.....	62
แนวคิดสำคัญและข้อตกลงเบื้องต้นของทฤษฎีการสรุปอ้างอิง ของความน่าเชื่อถือของผลการวัด.....	63
แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวน สำหรับการออกแบบการวัดที่มีหนึ่งฟาเซท.....	68
ความแปรปรวนของคะแนนความคลาดเคลื่อนสัมบูรณ์และสัมพัทธ์.....	70
ค่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด.....	71
งานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ.....	73
3 วิธีดำเนินการวิจัย.....	81
การกำหนดการเลือกประชากรและกลุ่มตัวอย่าง.....	81
การสร้างเครื่องมือที่ใช้ในการวิจัย.....	84
การเก็บรวบรวมข้อมูล.....	92
การจัดกระทำและการวิเคราะห์ข้อมูล.....	92
4 ผลการวิเคราะห์ข้อมูล.....	97
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล.....	97
การนำเสนอผลการวิเคราะห์ข้อมูล.....	99
ผลการวิเคราะห์ข้อมูล.....	99
5 สรุปผล อภิปรายและข้อเสนอแนะ.....	111
สังเขปจุดมุ่งหมายและวิธีการดำเนินการวิจัย.....	111
สรุปผลการวิจัย.....	112
อภิปรายผลการวิจัย.....	113
ข้อเสนอแนะในการนำผลการวิจัยไปใช้.....	116
ข้อเสนอแนะในการวิจัยครั้งต่อไป.....	116

สารบัญ (ต่อ)

บทที่	หน้า
บรรณานุกรม.....	117
ภาคผนวก.....	123
ภาคผนวก ก.....	124
ภาคผนวก ข.....	126
ภาคผนวก ค.....	131
ภาคผนวก ง.....	136
ภาคผนวก จ.....	154
ประวัติย่อผู้วิจัย.....	161



บัญชีตาราง

ตาราง	หน้า
1 ANOVA สำหรับการวัดที่มีสององค์ประกอบแบบไขว้ $p \times i$	70
2 เปรียบเทียบสูตรสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ และเชิงสัมพัทธ์.....	73
3 ประชากรนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียน สังกัดสำนักงาน คณะกรรมการการศึกษา ขั้นพื้นฐานภาคใต้ จำแนกตามจังหวัด.....	81
4 ขนาดกลุ่มตัวอย่างนักเรียน จำแนกตามจังหวัด.....	82
5 กรอบโครงสร้างแบบสอบการรู้เรื่องคณิตศาสตร์.....	83
6 กฎเกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics).....	84
7 กฎเกณฑ์การให้คะแนนแบบแยกส่วน (Analytic Scoring Rubrics).....	86
8 ค่าความยากง่าย ค่าอำนาจจำแนก และความเชื่อมั่นแบบสอบการรู้เรื่อง คณิตศาสตร์.....	100
9 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบการรู้เรื่องคณิตศาสตร์ จำแนกตาม กระบวนการรู้เรื่องคณิตศาสตร์.....	101
10 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบการรู้เรื่องคณิตศาสตร์ เรื่องปริภูมิและรูปทรงสามมิติ จำแนกตามกฎเกณฑ์การให้คะแนน.....	102
11 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบการรู้เรื่องคณิตศาสตร์ เรื่องการเปลี่ยนแปลงและความสัมพันธ์ จำแนกตามกฎเกณฑ์การให้คะแนน...	103
12 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบการรู้เรื่องคณิตศาสตร์ เรื่องปริมาณ จำแนกตามกฎเกณฑ์การให้คะแนน.....	104
13 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบการรู้เรื่องคณิตศาสตร์ เรื่องความไม่แน่นอน จำแนกตามกฎเกณฑ์การให้คะแนน.....	105
14 ผลการวิเคราะห์ความแปรปรวนจากแหล่งความแปรปรวนต่างๆที่ตรวจให้ คะแนนตามกฎเกณฑ์การให้คะแนนแบบภาพรวม.....	106
15 ผลการศึกษา D ที่มีกฎเกณฑ์ให้คะแนนแบบภาพรวม.....	107
16 ผลการวิเคราะห์ความแปรปรวนจากแหล่งความแปรปรวนต่างๆ ที่ตรวจให้ คะแนนตามกฎเกณฑ์การให้คะแนนแบบแยกส่วน.....	108
17 ผลการศึกษา D ที่มีกฎเกณฑ์ให้คะแนนแบบแยกส่วน.....	109
18 ผลการทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบสอบการรู้เรื่อง คณิตศาสตร์ที่มีความยาวแบบสอบ และกฎเกณฑ์การให้คะแนนต่างกัน.....	110

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 กรอบแนวคิดในการวิจัย.....	9
2 ความสัมพันธ์ระหว่างคำถามกับระดับความสามารถการรู้เรื่องคณิตศาสตร์.....	14
3 องค์ประกอบและกรอบโครงสร้างการประเมินคณิตศาสตร์.....	15
4 กระบวนการคิดเป็นคณิตศาสตร์.....	20
5 แผนภูมิสรุปสมรรถนะทางคณิตศาสตร์.....	23
6 กรอบการประเมินการรู้เรื่องคณิตศาสตร์ของ PISA 2015.....	24
7 แหล่งของความแปรปรวนและองค์ประกอบของความแปรปรวน (One Facet, p X i Design).....	69
8 แหล่งความคลาดเคลื่อนแบบสุ่มและความคลาดเคลื่อนแบบสัมพัทธ์ สำหรับ p X i Design.....	71
9 ขั้นตอนดำเนินการสร้างแบบสอบการรู้เรื่องคณิตศาสตร์.....	90
10 กราฟเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบการรู้เรื่อง คณิตศาสตร์ที่มีความยาวแบบสอบต่างกัน.....	110

บทที่ 1

บทนำ

ภูมิหลัง

ในปัจจุบันเป็นยุคแห่งกระแสโลกาภิวัตน์โลกมีความเปลี่ยนแปลงที่เกิดขึ้นอย่างรวดเร็วทั้งด้านวิชาการ และความก้าวหน้าทางเทคโนโลยีสารสนเทศ ทำให้ประเทศชาติไม่สามารถอยู่โดยลำพัง ต้องอาศัยการร่วมมือและพึ่งพาอาศัยกัน สังคมโลกในยุคปัจจุบันเต็มไปด้วยข้อมูลข่าวสาร ทำให้ต้องคิด วิเคราะห์ แยกแยะ และมีการตัดสินใจอย่างรวดเร็ว เพื่อให้ทันต่อสถานการณ์ในสังคมที่มีความซับซ้อนมากขึ้น จึงเป็นแรงผลักดันสำคัญที่ทำให้ต้องมีการปฏิรูปการศึกษา ซึ่งคุณภาพของการศึกษาเป็นตัวบ่งชี้สำคัญประการหนึ่งสำหรับความพร้อมในการก้าวเข้าสู่ศตวรรษที่ 21 อย่างมีประสิทธิภาพของการแข่งขันในเวทีโลก (สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน. 2555: 3) วิชาคณิตศาสตร์มีบทบาทสำคัญยิ่งต่อการพัฒนาความคิดของมนุษย์ ทำให้มนุษย์มีความคิดสร้างสรรค์ คิดอย่างมีเหตุมีผล เป็นระบบ มีระเบียบแบบแผน สามารถวิเคราะห์ปัญหาหรือสถานการณ์ได้อย่างถี่ถ้วนรอบคอบ ช่วยคาดการณ์ วางแผน ตัดสินใจ แก้ปัญหา และนำไปใช้ในชีวิตรประจำวันได้อย่างเหมาะสม (วิมลรัตน์ ศรีสุข. 2553: 74 - 75) คณิตศาสตร์มีประโยชน์ต่อการดำรงชีวิตและช่วยพัฒนาชีวิตให้ดีขึ้น ช่วยพัฒนาคนให้เป็นมนุษย์ที่สมบูรณ์ มีความสมดุลทั้งทางร่างกาย จิตใจ สติปัญญา และอารมณ์ สามารถคิดเป็น ทำเป็น แก้ปัญหาเป็น และสามารถอยู่ร่วมกับผู้อื่นได้อย่างมีความสุข (ภัทรา สุวรรณบัตร. 2552: 66)

จากอดีตจนถึงปัจจุบันพบว่า ผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ของนักเรียนไทยอยู่ในระดับต่ำมาตลอดระยะเวลา 30 ปี โดยเฉพาะนักเรียนระดับมัธยมศึกษามีคะแนนการทดสอบระดับชาติ (National Test : NT) ต่ำกว่าร้อยละ 50 (รัตนา ตั้งศิริชัยพงษ์. 2553: 6) ในปัจจุบันประเทศไทยได้เข้าร่วมโครงการประเมินผลนักเรียนร่วมกับนานาชาติ (Program for International Student Assessment : PISA) จุดที่ PISA พิจารณาเป็นหลัก คือ ต้องการทราบว่า “สิ่งสำคัญที่เกี่ยวกับคณิตศาสตร์ที่คนทั่วไปต้องรู้และทำได้คืออะไร” หรือคำถามที่ว่า “เยาวชนอายุ 15 ปี มีความสามารถพื้นฐานทางคณิตศาสตร์เพียงพอหรือไม่อย่างไร” เพื่อตอบคำถามเหล่านี้ PISA ได้กำหนดการประเมิน “การรู้เรื่องคณิตศาสตร์” เพื่อบ่งบอกภาพสมรรถนะของบุคคลในด้านวิธีการคิดโดยผ่านกฎเกณฑ์และวิธีการทางคณิตศาสตร์ สามารถตีความและใช้คณิตศาสตร์ภายใต้บริบทต่างๆที่หลากหลาย ตลอดระยะเวลาที่ประเทศไทยเข้าร่วมโครงการ PISA แนวโน้มทิศทางผลการประเมินของ PISA เมื่อเทียบในช่วง PISA 2009 กับ PISA 2012 พบว่า คะแนนของนักเรียนไทยวิชาคณิตศาสตร์เพิ่มขึ้นน้อยที่สุด โดยเฉพาะอย่างยิ่งผลการประเมินการรู้เรื่องคณิตศาสตร์ PISA 2015 พบว่า เด็กไทยมีคะแนนเฉลี่ยต่ำกว่าคะแนนเฉลี่ย

OECD (คะแนน 433) ผลการประเมินจัดอยู่ในกลุ่มต่ำ โดยอยู่ลำดับที่ 32 จาก 41 ประเทศ และเมื่อวิเคราะห์แบ่งตามพื้นที่ภูมิศาสตร์ พบว่า นักเรียนภาคใต้และภาคอีสานตอนล่าง เป็นกลุ่มที่มีคะแนนต่ำสุด (โครงการ PISA ประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2557: 14-15, 183,184, 181-182) ซึ่งสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ได้ร่วมกับสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) และสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) จัดทำแนวข้อสอบ PISA Liked ตามกรอบแนวคิดในการวัดสมรรถนะผู้เรียนแบบเดียวกับ PISA เพื่อเป็นการลดสัดส่วนของจำนวนนักเรียนในระดับต่ำให้เหลือน้อยลงมากที่สุดเท่าที่จะทำได้ และเป็นการยกระดับคุณภาพการศึกษา โดยข้อสอบ PISA Liked เน้นการคิดวิเคราะห์เพื่อการนำไปใช้ได้จริง เพื่อวัดสมรรถนะของผู้เรียนในแต่ละช่วงชั้นว่ามีความรู้ ความเข้าใจ และรู้จักการนำไปใช้ และได้มีการพัฒนาการจัดทำแนวข้อสอบ PISA Liked มาตั้งแต่ปี 2555 โดยรูปแบบข้อสอบเป็นโจทย์คำถามเชื่อมโยงเนื้อหาสาระในแต่ละวิชาเข้ากับเรื่องราวในชีวิตประจำวัน (ASTV ผู้จัดการออนไลน์. 8 กันยายน 2557) สอดคล้องกับจากรุณันท์ ขวัญแน่น (2558: 68) ที่ศึกษาการพัฒนารูปแบบการประเมินการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนระดับการศึกษาขั้นพื้นฐาน โดยรูปแบบการประเมินเป็นกรอบแนวคิด หรือแบบแผนในการประเมินที่แสดงให้เห็นถึงความสัมพันธ์อย่างเป็นระบบระหว่างองค์ประกอบหรือสิ่งต่างๆ ที่เกี่ยวข้องกับการประเมินการรู้เรื่องคณิตศาสตร์ รูปแบบการประเมินนี้ จะช่วยให้เกิดความเข้าใจเกี่ยวกับการประเมินได้อย่างเป็นระบบ และสามารถนำไปเป็นแนวทางสู่การปฏิบัติจากความสำคัญดังกล่าวข้างต้นจำเป็นอย่างยิ่งที่ต้องมีการวัดและประเมินผลการรู้เรื่องคณิตศาสตร์ เพื่อพัฒนาผู้เรียนเป็นสำคัญ และผลการประเมินจะเป็นสารสนเทศสำหรับการกำกับติดตามสนับสนุนและพัฒนการเรียนรู้ นับว่าเป็นสิ่งที่จำเป็นอย่างยิ่งต่อเป้าหมายการเรียนรู้ตั้งแต่ระดับบุคคล ระดับสถานศึกษา ระดับเขตพื้นที่การศึกษา และระดับประเทศ ตลอดจนถึงความเป็นสากลระดับนานาชาติ

เนื่องจากในบริบทการศึกษาประเทศไทยนั้นรูปแบบการประเมินผลการเรียนรู้ของผู้เรียน เพื่อวัดผลสัมฤทธิ์ทางการเรียนหรือความสามารถของผู้เรียนนั้นมีการพัฒนาและปรับเปลี่ยนวิธีการวัดและการประเมินให้มีความหลากหลาย และสอดคล้องกับเนื้อหาสาระของแต่ละกลุ่มสาระการเรียนรู้ การวัดและประเมินที่ทำให้ทราบว่าคะแนนที่ได้จากการสอบเกิดจากความรู้ที่แท้จริง วิธีการตรวจให้คะแนนจึงมีบทบาทสำคัญเพื่อให้คะแนนที่ได้จากแบบสอบมีความน่าเชื่อถือ การตรวจให้คะแนนแบบสองค่า (binary) เช่น การตรวจให้คะแนนแบบ 0 หรือ 1 ถูกหรือผิด เห็นด้วยหรือไม่เห็นด้วย เป็นวิธีการตรวจให้คะแนนที่ใช้กันอย่างแพร่หลายในปัจจุบัน เนื่องจากมีความเป็นปรนัยในการตรวจให้คะแนน ใช้เวลาตรวจน้อย แต่ยังมีข้อจำกัดที่สำคัญ คือ เปิดโอกาสให้มีการตอบถูกโดยการเดา ปัญหาการเดาสุ่มจึงเป็นแหล่งความคลาดเคลื่อนของการวัด นอกจากนี้โมเดลการวัดของทฤษฎีการทดสอบแบบดั้งเดิม ถือว่าเป็นความคลาดเคลื่อนของผลการวัดเป็นคะแนนรวมความคลาดเคลื่อนทุกแหล่งเข้าด้วยกัน และ

เป็นความคลาดเคลื่อนรวมอันเดียวกันที่ไม่สามารถแบ่งแยกได้ (Unique error) จึงเป็นข้อจำกัดทางทฤษฎีที่ไม่สามารถศึกษารายละเอียดของแหล่งความคลาดเคลื่อนของการวัดในสถานการณ์หรือเงื่อนไขการวัดต่าง ๆ ได้ (ศิริชัย กาญจนวาสี. 2555: 11) ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (G-theory) เป็นทฤษฎีที่ Cronbach, Gleser and Rajaratnam เป็นผู้เสนอในปี 1963 และต่อมาในปี 1972 ได้มีการพัฒนาทฤษฎีดังกล่าวให้มีความชัดเจนและสามารถนำไปประยุกต์ใช้ได้จริงในทางปฏิบัติ เป็นการเพิ่มความแม่นยำ ในการแปลความหมายของผลการวัด โดยมีจุดประสงค์ที่ต้องการสรุปอ้างอิงคะแนนจากการทดสอบ (Observed Score) ไปสู่เอกภพของคะแนน (Universe Score) (Hussain Alkharusi. 2012: 188; อ้างอิงจาก Crocker; Algina. 1986; Shavelson; Webb. 1991) ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดเป็นส่วนขยายของทฤษฎีการทดสอบแบบดั้งเดิม เนื่องจากคะแนนที่สังเกตได้มีความคลาดเคลื่อนจากคะแนนที่แท้จริง และแหล่งความคลาดเคลื่อนนั้นมีโอกาสเกิดขึ้นพร้อม ๆ กันในกระบวนการวัดผล (สุรศักดิ์ อมรรัตนศักดิ์. 2544: 206) ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดยอมรับว่าความคลาดเคลื่อนเกิดขึ้นจากหลายแหล่ง ดังนั้นจึงมีความพยายามที่จะหาวิธีการในการจำแนกความคลาดเคลื่อนของผลการวัด โดยการประมาณค่าความแปรปรวนจากแหล่งต่าง ๆ เพื่อบรรยายถึงปริมาณความแปรปรวนที่ได้มาจากแหล่งต่าง ๆ อย่างชัดเจน

จากการศึกษาความเชื่อมั่นของแบบสอบหรือแบบวัดตามทฤษฎีการทดสอบแบบดั้งเดิมที่ยังใช้กันอยู่นั้น เป็นการประมาณค่าความเชื่อมั่นของแบบสอบที่ใช้กับกลุ่มบุคคลภายใต้เงื่อนไขของการทดสอบที่เฉพาะ ได้แก่ การทดสอบที่ผู้สอบทุกคนต้องทำข้อสอบเหมือนกันทุกข้อ ผลการตอบได้รับการให้คะแนนโดยผู้ตรวจคนเดียว และแบบสอบถามที่ใช้ตั้งอยู่บนข้อตกลงเบื้องต้นของความเป็นคู่ขนานระหว่างแบบสอบ ถ้าสถานการณ์ของการทดสอบแตกต่างไปจากที่กล่าวมา เทคนิคการประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบดั้งเดิมจะไม่สามารถนำมาประยุกต์ใช้ได้ เช่น กรณีของแบบสอบที่มีรูปแบบความยาวและจำนวนครั้งของการทดสอบแตกต่างกัน เป็นต้น ในการศึกษาคุณภาพของแบบสอบองค์ประกอบที่มีผลต่อความเชื่อมั่นของแบบสอบอีกประการหนึ่ง คือ ความยาวของแบบสอบ (ยศอารีย์ รวยธนพานิช. 2534: 3; อ้างอิงจาก Adams.1964: 97) การเพิ่มจำนวนข้อสอบเพื่อให้ความเชื่อมั่นสูงขึ้น จำนวนข้อที่เพิ่มขึ้นจะต้องมีคุณภาพเหมือนเดิม ภายใต้สถานการณ์ของการใช้แบบสอบหรือเงื่อนไขของการทดสอบที่ความยาวของแบบสอบจะมีผลต่อความน่าเชื่อถือมากน้อยเพียงใด เมื่อกำหนดการสรุปอ้างอิงสัมประสิทธิ์ความน่าเชื่อถือของคะแนนแบบสอบที่มีความยาวของข้อสอบตามเงื่อนไขต่างกัน ในการวิจัยครั้งนี้ผู้วิจัยกำหนดระดับความยาว 12 ข้อ 24 ข้อ และ 36 ข้อ เมื่อพิจารณาค่าสัมประสิทธิ์ความน่าเชื่อถือของแบบสอบ (G-Coefficient) หรือ ค่าความเชื่อมั่น (Reliability) และค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนนสำหรับการตัดสินใจทั้งเชิงสัมพัทธ์และเชิงสัมบูรณ์ ว่าควรให้แบบทดสอบมีความยาวแบบสอบจำนวน

ที่ข้อนี้จะเหมาะสม ทฤษฎีนี้จึงมีความสอดคล้องเหมาะสมสำหรับกฎเกณฑ์การให้คะแนน กระบวนการรู้เรื่องคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้น โดยมีการตรวจให้คะแนน 2 รูปแบบ คือ 1)แบบแยกองค์ประกอบ (Analytic Scoring Rubrics) เกณฑ์แบบนี้ใช้เวลาตรวจค่อนข้างนาน สร้างความเมื่อยล้าหากชิ้นงานที่ต้องตรวจมีจำนวนมาก แต่ข้อดีคือ มีสารสนเทศที่ครบถ้วนส่งกลับไปยังนักเรียน และ2)แบบรวมองค์ประกอบ (Holistic Scoring Rubrics) แม้ว่ารูปแบบนี้ตรวจง่าย ใช้เวลาน้อยและมีความเชื่อมั่นสูง แต่ก็ยากที่จะอธิบายเหตุผลของคะแนนที่นักเรียนได้รับ (บุษวรรษ์ แสนปลื้ม. 2556: 2 – 3; อ้างอิงจาก ส.วาสนา ประवालพฤกษ์. 2533: 39 - 42; ชัยฤทธิ์ ศิลาเดช. 2544: 29; บุญเชิด ภิญญอนันตพงษ์. 2544: 90; สำนักทดสอบทางการศึกษากกรมวิชาการ. 2554: 35 - 38) จะเห็นได้ว่าวิธีการตรวจให้คะแนนทั้ง 2 รูปแบบดังที่กล่าวข้างต้นมีข้อดีและข้อด้อยแตกต่างกัน ดังนั้นผู้วิจัยมีความสนใจสร้างรูปแบบกฎเกณฑ์การให้คะแนนและวิธีการตรวจให้คะแนนแบบสอบการรู้เรื่องคณิตศาสตร์ทั้งสองรูปแบบ เพื่อวัดความสามารถ กระบวนการเรียนรู้ตามคำนิยามการรู้เรื่องคณิตศาสตร์

ด้วยเหตุผลดังกล่าวข้างต้น ผู้วิจัยจึงสนใจสร้างแบบสอบการรู้เรื่องคณิตศาสตร์ เมื่อความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุป อ้างอิงความน่าเชื่อถือของผลการวัด ผู้วิจัยศึกษาความแปรปรวนขององค์ประกอบหลักได้แก่ ผู้สอบ ข้อสอบที่มีการตรวจให้คะแนนด้วยกฎเกณฑ์การให้คะแนนต่างกัน พิจารณาคูณลักษณะ คำสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด เพื่อเป็นแนวทางให้กับผู้ที่เกี่ยวข้อง ในการเลือกใช้ความยาวแบบสอบและกฎเกณฑ์การให้คะแนนที่มีความเหมาะสม เพื่อประหยัดเวลา แรงงานในการตรวจแบบทดสอบและสอดคล้องกับสภาพจริงในทางปฏิบัติ

คำถามของการวิจัย

1. แบบสอบการรู้เรื่องคณิตศาสตร์ควรมีความยาวแบบสอบกี่ข้อ และกฎเกณฑ์การให้คะแนนข้อสอบการรู้เรื่องคณิตศาสตร์ควรมีลักษณะอย่างไร
2. แบบสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาว และกฎเกณฑ์การให้คะแนนต่างกันมีคุณภาพอย่างไร
3. แบบสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาว และกฎเกณฑ์การให้คะแนนต่างกันมีค่าสัมประสิทธิ์การสรุปอ้างอิงแตกต่างกันหรือไม่

ความมุ่งหมายของการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยได้ตั้งความมุ่งหมายไว้ดังนี้

1. เพื่อสร้างและตรวจสอบคุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน
2. เพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน

ความสำคัญของการวิจัย

การศึกษาวิจัยครั้งนี้ เพื่อให้ได้เครื่องมือสำหรับวัดและประเมินการรู้เรื่องคณิตศาสตร์ (Mathematics Literacy) โดยให้ความสำคัญทั้งเชิงทฤษฎีและการปฏิบัติ

1. ในทางทฤษฎี เพื่อให้ได้เครื่องมือการวัดประเมินมีความน่าเชื่อถือและความเที่ยงตรง ในการศึกษาครั้งนี้แสดงข้อมูลเชิงประจักษ์ทางด้านความยาวแบบสอบและกฎเกณฑ์การให้คะแนนที่ส่งผลต่อความเชื่อมั่นของผลการวัด นอกจากนี้ยังทำให้ทราบขนาดความแปรปรวนของแหล่งต่างๆ ที่มีอิทธิพลต่อความน่าเชื่อถือของคะแนน เป็นประโยชน์ในแง่ของการตรวจสอบและเป็นหลักฐานที่แสดงถึงแหล่งความคลาดเคลื่อน โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (G-theory) เพื่อการอธิบายและแปลผลการวัดได้อย่างเที่ยงตรง และเป็นธรรม

2. ในทางปฏิบัติ เพื่อพัฒนาการวัดและประเมินผลรายวิชาคณิตศาสตร์ ซึ่งส่วนใหญ่เครื่องมือวัดเป็นแบบสอบเลือกตอบที่มีวิธีการให้คะแนนแบบ 0,1 ผู้วิจัยจึงได้สร้างแบบสอบที่มีกฎเกณฑ์การให้คะแนน (Scoring rubrics) ซึ่งเป็นรูปแบบการวัดประเมินที่มีการกำหนดประเด็นไว้อย่างชัดเจน ครูผู้สอนสามารถวินิจฉัยคำตอบของนักเรียนเปรียบเทียบกับเกณฑ์การประเมินในแต่ละด้าน ซึ่งแต่ละด้านจะมีคำอธิบายคุณภาพในแต่ละระดับ เพื่อให้ครูผู้สอนสามารถประเมินได้อย่างชัดเจน และตัดสินผลการเรียนของนักเรียนได้ถูกต้อง มีความเที่ยงตรงยุติธรรม กฎเกณฑ์การให้คะแนนจึงเป็นเทคนิควิธีการประเมินที่เหมาะสม และยังเป็นแนวทางสำหรับครูผู้สอนรายวิชาอื่นๆ ได้อีกด้วย

ขอบเขตของการวิจัย

ประชากรที่ใช้ในการวิจัย

ประชากรในการวิจัยครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) ภาคใต้ปีการศึกษา 2560 จำนวน 645 โรงเรียน มีประชากรจำนวนทั้งสิ้น 64,253 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) ภาคใต้ปีการศึกษา 2560 จำนวน 400 คน โดยการสุ่มแบบหลายขั้นตอน (Multi - stage random sampling)

ตัวแปรที่ศึกษา

1. ตัวแปรอิสระมี 2 ตัวแปร ได้แก่

1.1 ความยาวของแบบสอบ มี 3 ระดับ คือ 12 ข้อ 24 ข้อ และ 36 ข้อ

1.2 รูปแบบกฎเกณฑ์การให้คะแนน (Scoring rubrics) มี 2 รูปแบบ คือ การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics) และการให้คะแนนแบบแยกส่วน (Analytic Scoring Rubrics)

2. ตัวแปรตาม คือ ค่าสัมประสิทธิ์การสรุปอ้างอิง (G- Coefficient) ตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

นิยามศัพท์เฉพาะ

1. การรู้เรื่องคณิตศาสตร์ (Mathematics Literacy) หมายถึง ความสามารถของผู้สอบเกี่ยวกับ กระบวนการทางคณิตศาสตร์ แบ่งออกเป็น 3 กระบวนการ คือ 1)การคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ 2)การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผลทางคณิตศาสตร์ และ 3)การตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์

1.1 การคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ หมายถึง ความสามารถของผู้สอบในการวิเคราะห์ สร้างแนวทางและนำไปแก้ปัญหา สามารถแปลงปัญหาจากสถานการณ์ในชีวิตจริงให้อยู่ในขอบเขตคณิตศาสตร์ และกำหนดโครงสร้างทางคณิตศาสตร์ การแสดงเครื่องหมายแทน มีลักษณะจำเพาะกับปัญหาในโลกชีวิตจริง สามารถให้เหตุผล รู้จักการตั้งสมมติฐาน และพิจารณาข้อจำกัดได้อย่างสมเหตุสมผล ประกอบด้วยกระบวนการระบุประเด็นทางคณิตศาสตร์ของปัญหาในชีวิตจริง ทำสถานการณ์หรือปัญหาให้อยู่ในรูปอย่างง่าย เพื่อให้การวิเคราะห์ทางคณิตศาสตร์ง่ายขึ้น และแปลงปัญหาให้อยู่ในรูปของคณิตศาสตร์

1.2 การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผลทางคณิตศาสตร์ หมายถึง ความสามารถของผู้สอบในการประยุกต์ใช้แนวคิดหลักทางคณิตศาสตร์ ข้อเท็จจริง วิธีดำเนินการ และเหตุผลทางคณิตศาสตร์ในการแก้ปัญหาเชิงคณิตศาสตร์เพื่อให้ได้ข้อสรุปทางคณิตศาสตร์ เป็นกระบวนการที่แต่ละคนแสดงวิธีดำเนินการทางคณิตศาสตร์ที่จำเป็นเพื่อให้ได้มาซึ่งผลลัพธ์ และค้นหาวิธีแก้ปัญหาทางคณิตศาสตร์ มีการสร้างแบบจำลองของสถานการณ์ปัญหา สร้างกฎเกณฑ์ ระบุความเชื่อมโยงระหว่างองค์ความรู้ทางคณิตศาสตร์ และสร้างข้อโต้แย้งทางคณิตศาสตร์ ประกอบด้วยกระบวนการคิดและนำกลยุทธ์ในการหาวิธีแก้ปัญหาทางคณิตศาสตร์ไปใช้ ใช้เครื่องมือทางคณิตศาสตร์เพื่อช่วยให้วิธีการแก้ปัญหาถูกต้องหรือเหมาะสม และนำกฎเกณฑ์ ขั้นตอนวิธี และโครงสร้างคณิตศาสตร์ไปใช้แก้ปัญหา

1.3 การตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์ หมายถึง ความสามารถของผู้สอบในการสะท้อนวิธีแก้ปัญหา ผลลัพธ์ หรือข้อสรุปทางคณิตศาสตร์แล้ว ตีความออกมาในบริบทของปัญหาโลกชีวิตจริง รวมถึงการแปลความหมายของวิธีแก้ปัญหาหรือการให้เหตุผลทางคณิตศาสตร์ย้อนกลับไปที่บริบทของปัญหา และตัดสินใจว่าผลลัพธ์ที่ได้เป็นเหตุเป็นผลและเข้ากันได้กับบริบทของปัญหาหรือไม่ มีการสร้างและสื่อสารคำอธิบายหรือข้อโต้แย้งในบริบทของปัญหา และการสะท้อนทั้งกระบวนการสร้างแบบจำลองและผลที่ได้ ประกอบด้วยกระบวนการตีความผลลัพธ์ทางคณิตศาสตร์กลับไปสู่บริบทชีวิตความเป็นจริง ประเมินความเป็นเหตุเป็นผลเป็นผลของวิธีการแก้ปัญหาทางคณิตศาสตร์ในบริบทของปัญหาชีวิตจริง และอธิบายความสมเหตุสมผลของผลลัพธ์หรือข้อสรุปทางคณิตศาสตร์กับบริบทของปัญหา

2. กฎเกณฑ์การให้คะแนน (Scoring rubrics) หมายถึง แนวทางในการตรวจให้คะแนนแบบสอบการรู้เรื่องคณิตศาสตร์ มีลักษณะการตรวจให้คะแนน 2 วิธี คือ การให้คะแนนแบบภาพรวมและการให้คะแนนแบบแยกส่วน

2.1 การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics) หมายถึง การพิจารณาให้คะแนนของคำตอบในภาพรวม โดยที่กำหนดคะแนนเต็มสำหรับคำตอบที่มีความถูกต้องครบถ้วน คะแนนบางส่วนสำหรับคำตอบที่ถูกต้อง แต่มีข้อผิดพลาดบางส่วน

2.2 การให้คะแนนแบบแยกส่วน (Analytic Scoring Rubrics) หมายถึง การพิจารณาให้คะแนนของคำตอบในแต่ละประเด็นย่อย และในแต่ละประเด็นมีการกำหนดระดับคะแนนแตกต่างกัน ขึ้นอยู่กับน้ำหนักความสำคัญของคำตอบ

3. ความยาวแบบสอบ หมายถึง จำนวนข้อสอบของแบบสอบการรู้เรื่องคณิตศาสตร์ ตามกรอบโครงสร้างการรู้เรื่องคณิตศาสตร์ครอบคลุม 4 สาระ แต่ละสาระประกอบด้วยสถานการณ์ 1 - 3 สถานการณ์ แต่ละสถานการณ์แบ่งตามกระบวนการคณิตศาสตร์ 3 กระบวนการ ได้แบบสอบการรู้เรื่องคณิตศาสตร์ ที่มีความยาวแบบสอบการรู้เรื่องคณิตศาสตร์ 12 ข้อ 24 ข้อ และ 36 ข้อ

4. ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หมายถึง การวิเคราะห์ความน่าเชื่อถือของคะแนนที่สังเกตได้ จากแบบสอบที่ทำการวัด ในเงื่อนไขหรือสถานการณ์ต่างๆ ของการทดสอบไปยังคะแนนจริง โดยถือว่าคะแนนความคลาดเคลื่อนของการวัดสามารถเกิดจากหลายแหล่งที่สามารถประมาณค่าแยกกันภายใต้การวิเคราะห์ครั้งเดียวกันได้ โดยการวิจัยครั้งนี้ต้องการสรุปอ้างอิงความน่าเชื่อถือของคะแนนที่สังเกตได้จากแบบสอบการรู้เรื่องคณิตศาสตร์ ไปยังคะแนนจริง ประกอบด้วยขั้นตอนของการศึกษา 2 ขั้นตอน คือ

ขั้นตอนที่ 1 การศึกษา G-Study หมายถึง ขั้นตอนการกำหนดเป้าหมายของการสรุปอ้างอิงความน่าเชื่อถือของคะแนนที่ได้ไปยังเงื่อนไขรูปแบบการวัดที่ต้องการศึกษา ในการวิจัยครั้งนี้กำหนดเป็น One - Facet เพื่อการเก็บรวบรวมข้อมูลวิเคราะห์ความแปรปรวนของแหล่งความคลาดเคลื่อนที่สนใจ 1 แหล่ง คือ ความยาวแบบสอบ

ขั้นตอนที่ 2 การศึกษา D-Study หมายถึง ขั้นตอนการตัดสินใจของการสรุปอ้างอิงความน่าเชื่อถือของคะแนนที่ได้จากเครื่องมือไปยังเงื่อนไขเป้าหมายของการวัด เพื่อนำเครื่องมือ ไปใช้วัดในเงื่อนไขที่ทำให้ได้ความน่าเชื่อถือระดับที่พึงประสงค์ นำไปสู่การคำนวณค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) สำหรับการตัดสินใจเชิงสัมพัทธ์ และสำหรับการตัดสินใจเชิงสมบูรณ์

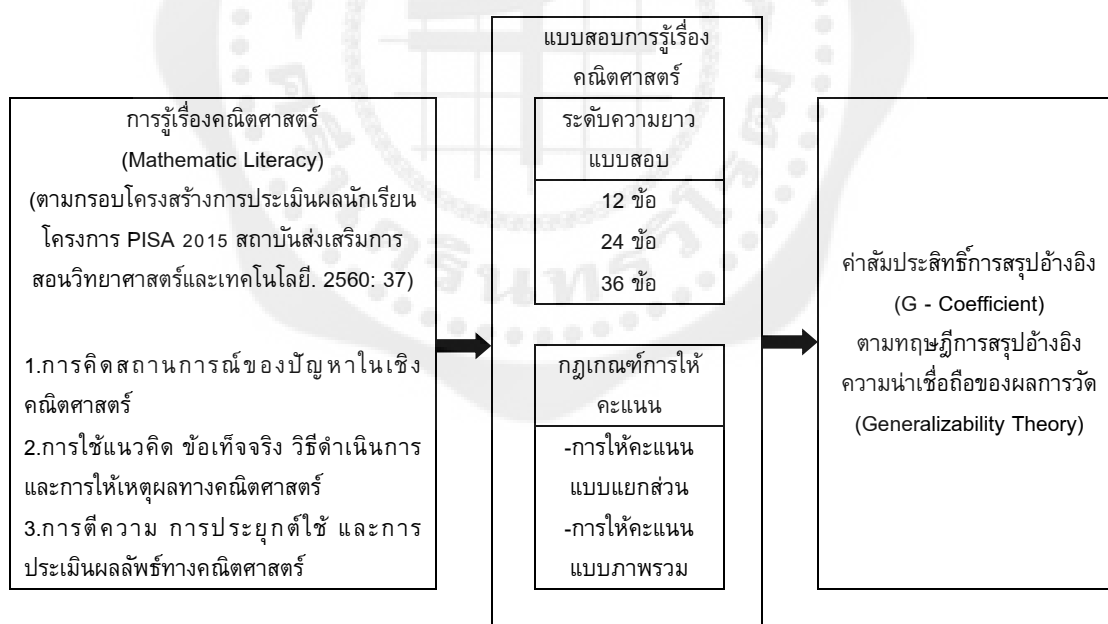
5. เอกภพ (Universe) หมายถึง เงื่อนไขการวัดองค์ประกอบทั้งหมดของแบบสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน

6. ฟาเซท (Facet) หมายถึง กลุ่มเงื่อนไขของการวัดที่ศึกษา ในการวิจัยครั้งนี้มี 1 ฟาเซท ซึ่งมีกลุ่มเงื่อนไขของการวัดของคุณภาพแบบสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาวของแบบสอบจำนวน 12 ข้อ 24 ข้อ และ 36 ข้อ โดยการตรวจให้คะแนน 2 รูปแบบ คือ กฎเกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics) และกฎเกณฑ์การให้คะแนนแบบแยกส่วน (Analytic Scoring Rubrics)

7. คุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์ หมายถึง ค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) ซึ่งเป็นค่าความเชื่อมั่นของคะแนนที่ได้จากแบบสอบภายใต้เงื่อนไขหรือสถานการณ์ที่กำหนด แสดงถึงระดับค่าความแม่นยำและความน่าเชื่อถือของการสรุปอ้างอิงคะแนนที่ได้จากแบบสอบไปยังคะแนนเอกภพของผู้สอบ

กรอบแนวคิดในการวิจัย

การรู้เรื่องคณิตศาสตร์ (Mathematic Literacy) เป็นการวัดความสามารถกระบวนการทางคณิตศาสตร์ ประกอบด้วย การคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผลทางคณิตศาสตร์ และการตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2557: 22) ผู้วิจัยสนใจสร้างเครื่องมือวัดและประเมินการรู้เรื่องคณิตศาสตร์ เนื่องจากการวัดและประเมินผลการรู้เรื่องคณิตศาสตร์เป็นเรื่องใหม่สำหรับประเทศไทย ผู้วิจัยจึงได้ศึกษาค้นคว้าการสร้างแบบสอบที่มีความเชื่อมั่นสูง โดยมีเงื่อนไขการวัดหนึ่งเงื่อนไข คือ ระดับความยาว 12, 24 และ 36 ข้อ ที่มีกฎเกณฑ์การให้คะแนน 2 รูปแบบ คือ กฎเกณฑ์การให้คะแนนแบบภาพรวม และกฎเกณฑ์การให้คะแนนแบบแยกส่วน โดยพิจารณาค่าสัมประสิทธิ์การสรุปอ้างอิงจากการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ผลการวิเคราะห์ที่ได้สามารถนำไปใช้สำหรับวางแผนเพื่อตัดสินใจว่าควรเลือกใช้เงื่อนไขใดจึงเหมาะสม ผู้วิจัยจึงได้ใช้กรอบแนวคิดนี้พัฒนาวิธีการวัด เพื่อประหยัดเวลา แรงงานในการตรวจแบบสอบ เพื่อให้ได้เครื่องมือการวัดผลทางการศึกษาที่มีคุณภาพ ดังแสดงในภาพประกอบ 1



ภาพประกอบ 1 กรอบแนวคิดในการวิจัย

สมมติฐานการวิจัย

แบบสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบต่างกันมีค่าสัมประสิทธิ์การสรุปอ้างอิงแตกต่างกัน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยเพื่อศึกษาคุณภาพแบบสอบถามการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบถามและกฎเกณฑ์การให้คะแนนต่างกัน โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ผู้วิจัยได้ศึกษาเอกสารรายงานการวิจัยและบทความทางวิชาการ วารสารวิชาการทั้งในและต่างประเทศประมวลเป็นองค์ความรู้ ดังนั้นเพื่อให้เกิดความเข้าใจในกระบวนการวิจัยครั้งนี้ ผู้วิจัยนำเสนอออกเป็น 3 ตอน ตามลำดับดังนี้

ตอนที่ 1 แนวคิดและทฤษฎีที่เกี่ยวกับการรู้เรื่องคณิตศาสตร์

- 1.1 นิยามการรู้เรื่องคณิตศาสตร์และระดับการรู้เรื่องคณิตศาสตร์
- 1.2 การจัดกรอบโครงสร้างการประเมินผลการรู้เรื่องคณิตศาสตร์
- 1.3 กระบวนการคิดเป็นคณิตศาสตร์
- 1.4 สมรรถนะทางคณิตศาสตร์
- 1.5 การประเมินผลด้านการรู้เรื่องคณิตศาสตร์
- 1.6 งานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ

ตอนที่ 2 แนวคิดและทฤษฎีเกี่ยวกับกฎเกณฑ์การให้คะแนน

- 2.1 ความหมายและความสำคัญของกฎเกณฑ์การให้คะแนน
- 2.2 องค์ประกอบของกฎเกณฑ์การให้คะแนน
- 2.3 ลักษณะของกฎเกณฑ์การให้คะแนน
- 2.4 ประเภทของกฎเกณฑ์การให้คะแนน
- 2.5 การสร้างกฎเกณฑ์การให้คะแนน
- 2.6 งานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ

ตอนที่ 3 แนวคิดและทฤษฎีที่เกี่ยวกับการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

- 3.1 ความเป็นมาของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด
- 3.2 แนวคิดสำคัญและข้อตกลงเบื้องต้น
- 3.3 แหล่งความแปรปรวนและองค์ประกอบของความแปรปรวน
- 3.4 ความแปรปรวนของคะแนนความคลาดเคลื่อนสัมบูรณ์และสัมพัทธ์
- 3.5 ค่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด
- 3.6 งานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ

ตอนที่ 1 แนวคิดและทฤษฎีที่เกี่ยวกับการรู้เรื่องคณิตศาสตร์ (Mathematics Literacy)

การรู้เรื่องคณิตศาสตร์ เป็นสิ่งจำเป็นอย่างหนึ่งในการดำเนินชีวิตปัจจุบัน การที่ PISA ใช้คำว่า “การรู้เรื่องคณิตศาสตร์” (Mathematics Literacy) เพื่อเน้นให้ชัดเจนว่า การประเมินความรู้และทักษะคณิตศาสตร์ของ OECD/PISA ไม่ได้เน้นความรู้เนื้อหาวิชา คณิตศาสตร์ที่เรียนตามหลักสูตรในโรงเรียน แต่เน้นการนำคณิตศาสตร์ที่เคยได้เรียนมา เอามาใช้ในสถานการณ์ของชีวิตจริง ซึ่งนักเรียนจะต้องสามารถขยายความรู้จากที่เคยได้เรียนมาจาก โรงเรียน นำมาประยุกต์กับสถานการณ์จริงในบริบทต่างๆ ที่หลากหลาย จุดประสงค์ของการ ประเมินการรู้เรื่องคณิตศาสตร์ของ OECD/PISA ก็เพื่อจะศึกษาว่าเยาวชนกลุ่มอายุ 15 ปี จะสามารถเป็นประชาชนที่รับรู้สาระ มีข้อมูลข่าวสาร และเป็นผู้บริโภคที่ฉลาดเพียงใด และสามารถนำสิ่งที่ได้ศึกษาเล่าเรียนในโรงเรียนไปใช้ในสถานการณ์ที่จะต้องพบเจอในชีวิตจริงได้หรือไม่อย่างไร

จุดมุ่งหมายหลักของการประเมินผลของ PISA ก็เพื่อต้องการพัฒนาตัวชี้วัดว่าระบบ การศึกษาของประเทศที่ร่วมโครงการสามารถให้การศึกษาเพื่อเตรียมตัวเยาวชนอายุ 15 ปี ให้พร้อมที่จะมีบทบาทหรือมีส่วนร่วมสร้างสรรค์ และดำเนินชีวิตอย่างมีคุณภาพในสังคมได้มากน้อย เพียงใด การประเมินของ PISA มีจุดหมายที่มองไปในอนาคตมากกว่าการจำกัดอยู่ที่การวัดและ ประเมินผลตามหลักสูตรที่นักเรียนได้เรียนในปัจจุบัน และการประเมินผลก็มุ่งความชัดเจนที่จะ หาคำตอบว่านักเรียนสามารถนำสิ่งที่ได้ การศึกษาเล่าเรียนในโรงเรียนไปใช้ในสถานการณ์ที่ นักเรียนมีโอกาสที่จะต้องพบเจอในชีวิตจริงได้หรือไม่อย่างไร PISA ได้ให้ความสำคัญกับปัญหา ในชีวิตจริง ในสถานการณ์จริงในโลก (คำว่า “โลก” ในที่นี้หมายถึง สถานการณ์ของธรรมชาติ สังคม และวัฒนธรรมที่บุคคลนั้นๆ อาศัยอยู่)

การที่ PISA ให้ความสำคัญกับ ปัญหาในชีวิตจริง เพราะว่าประชาชนทุกวันนี้ต้อง เผชิญหน้ากับกิจวัตรประจำวันที่เกี่ยวข้องกับคณิตศาสตร์ เป็นต้นว่าปริมาณ รูปทรง มิติ ความน่าจะเป็น และแนวคิดทางคณิตศาสตร์ต่างๆ อีกมากมาย PISA จึงให้ความสำคัญที่ ความต้องการให้นักเรียนเผชิญหน้ากับปัญหาทางคณิตศาสตร์ที่มีอยู่ในแวดวงของการดำเนิน ชีวิต ซึ่งต้องการให้นักเรียนระบุสถานการณ์ที่สำคัญของปัญหา กระตุ้นให้หาข้อมูล สืบวจ ตรวจสอบ และนำไปสู่การแก้ปัญหา ในกระบวนการนี้ต้องการทักษะหลายอย่างเป็นต้นว่า ทักษะการคิดและการใช้เหตุผล ทักษะการโต้แย้ง การสื่อสาร ทักษะการสร้างตัวแบบ การตั้ง ปัญหาและการแก้ปัญหา การนำเสนอ การใช้สัญลักษณ์ การดำเนินการ ในกระบวนการเหล่านี้ นักเรียนต้องใช้ทักษะต่างๆ ที่หลากหลายมารวมกันหรือใช้ทักษะหลายอย่าง ที่ทับซ้อนหรือคาบ เกี่ยวกัน เพราะกำลังคนในปัจจุบันถูกคาดหวังให้เป็นกำลังงานที่มีความคิดและสมรรถนะสูง ซึ่งจะส่งผลกระทบต่องานที่ทำในหน้าที่และสำหรับทุกคนไม่ว่าจะทำงานระดับใดจะถูก

คาดหวังว่าจะไม่ใช่เฉพาะร่างกายทำงานช้าๆ อย่างเดิมเท่านั้น แต่จะต้องพบกับความเปลี่ยนแปลงทางเทคโนโลยีและต้องสามารถปรับเปลี่ยนตัวเองให้สามารถจัดการกับเทคโนโลยีเครื่องจักรกล ต้องสามารถจัดการกับข้อมูลข่าวสารที่ถาโถมเข้ามาตลอดเวลา แนวโน้มของทุกๆ อาชีพ ชี้บอกว่า “บุคคลต้องมีความสามารถที่จะเข้าใจ สื่อสาร ใช้และอธิบายแนวคิด และวิธีการที่ยึดถือการคิดแบบคณิตศาสตร์เป็นหลัก”

ในแง่ภาษา แต่เดิมคำว่า “Literacy” ใช้ในความหมายของการรู้เรื่องในด้านภาษาของมนุษย์ เพราะการรู้เรื่องทางภาษา ซึ่งได้แก่ การรู้ที่จะอ่าน เขียน ฟัง และพูด เป็นเครื่องมือที่สำคัญที่สุดที่ทำให้มนุษย์สามารถปฏิบัติกิจกรรมต่างๆ ร่วมกันได้ ในทำนองเดียวกันถ้าคิดถึงคณิตศาสตร์ในฐานะเป็นภาษาอย่างหนึ่ง นักเรียนก็ต้องเรียนรู้คณิตศาสตร์ (แนวคิด เนื้อหาสาระ) และต้องเรียนรู้การที่จะใช้แนวคิดนั้นๆ ในการแก้ปัญหาใหม่ ในสถานการณ์ใหม่ๆ ที่เรียกว่า “สถานการณ์ทางสังคมได้ด้วย” (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2554: 87-88)

ปกติคนเราจะต้องพบกับสถานการณ์ต่างๆ เช่น การจับจ่ายใช้สอย การเดินทาง การทำอาหาร การจัดระเบียบการเงินของตน การประเมินสถานการณ์ การตัดสินใจ ประเด็นปัญหาทางสังคมการเมือง ฯลฯ ซึ่งความรู้คณิตศาสตร์สามารถเข้ามาช่วยทำให้การมองประเด็น การตั้งปัญหา หรือการแก้ปัญหา มีความชัดเจนยิ่งขึ้น การใช้คณิตศาสตร์ดังกล่าวนั้น แม้จะต้องมีรากฐานมาจากทักษะคณิตศาสตร์ในชั้นเรียน แต่ก็จำเป็นต้องมีความสามารถในการใช้ทักษะนั้นๆ ในสถานการณ์อื่นๆ นอกเหนือไปจากสถานการณ์ของปัญหาคณิตศาสตร์ล้วนๆ หรือแบบฝึกคณิตศาสตร์ที่เรียนในโรงเรียนที่นักเรียนจะสามารถคิดอยู่ในวงจำกัดของเนื้อหาวิชา โดยไม่ต้องคำนึงถึงความเป็นจริงมากนัก แต่การใช้คณิตศาสตร์ในชีวิตจริงนักเรียนต้องรู้จักสถานการณ์ หรือสิ่งแวดล้อมของปัญหาต้องเลือกตัดสินใจว่าจะใช้ความรู้คณิตศาสตร์อย่างไร (คัดลอกจากเว็บไซต์ สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี) <http://pisathailand.ipst.ac.th/>

1.1 นิยามการรู้เรื่องคณิตศาสตร์

PISA ให้นิยาม “การรู้เรื่องคณิตศาสตร์” ไว้ว่า “...คือสมรรถนะของบุคคลในการที่จะบ่งบอกและเข้าใจบทบาทของคณิตศาสตร์ที่มีในโลกเพื่อที่สามารถตัดสินใจบนพื้นฐานความรู้ที่เข้มแข็ง และเพื่อใช้และผูกพันกับคณิตศาสตร์ที่จะตอบสนองความจำเป็นต่อชีวิตของแต่ละบุคคล ในอันที่จะเป็นพลเมืองที่มีความคิด มีความหวังใฝ่ และสร้างสรรค์สังคม...” (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2554: 88 - 89)

การขยายความคำที่ใช้ในนิยามดังกล่าวจะช่วยอธิบายให้ชัดเจน ดังนี้

คำว่า “การรู้เรื่องคณิตศาสตร์” เน้นย้ำถึงการใช้คณิตศาสตร์ตอบสนองสถานการณ์ต่างๆ ในโลก อย่างการคิดวิเคราะห์ มองหาความสัมพันธ์และแปรผันได้หลากหลาย และชัดเจนว่าคนที่แสดงสมรรถนะอย่างนั้นได้จะต้องมีความรู้คณิตศาสตร์เป็นพื้นฐานอยู่แล้ว อย่างไรก็ตามความรู้ที่ว่านั้นไม่จำกัดอยู่เพียงนิยาม ข้อเท็จจริง และวิธีการแก้โจทย์คณิตศาสตร์ รวมทั้งไม่ใช้การฝึกทักษะทางคณิตศาสตร์เรื่องใดเรื่องหนึ่ง เป็นต้นว่าทักษะการแก้สมการ การพิสูจน์เอกลักษณ์ ฯลฯ หากแต่เกี่ยวข้องกับการรวมเอาความรู้ต่างๆ ดังกล่าวแล้วที่มีอยู่มาสร้างเป็นแนวคิดตอบสนองสถานการณ์ที่ก่อกวนจากภายนอก

โลก ตามคำนิยาม หมายถึง โลกธรรมชาติ สังคม วัฒนธรรม ที่บุคคลคนนั้นอาศัยอยู่

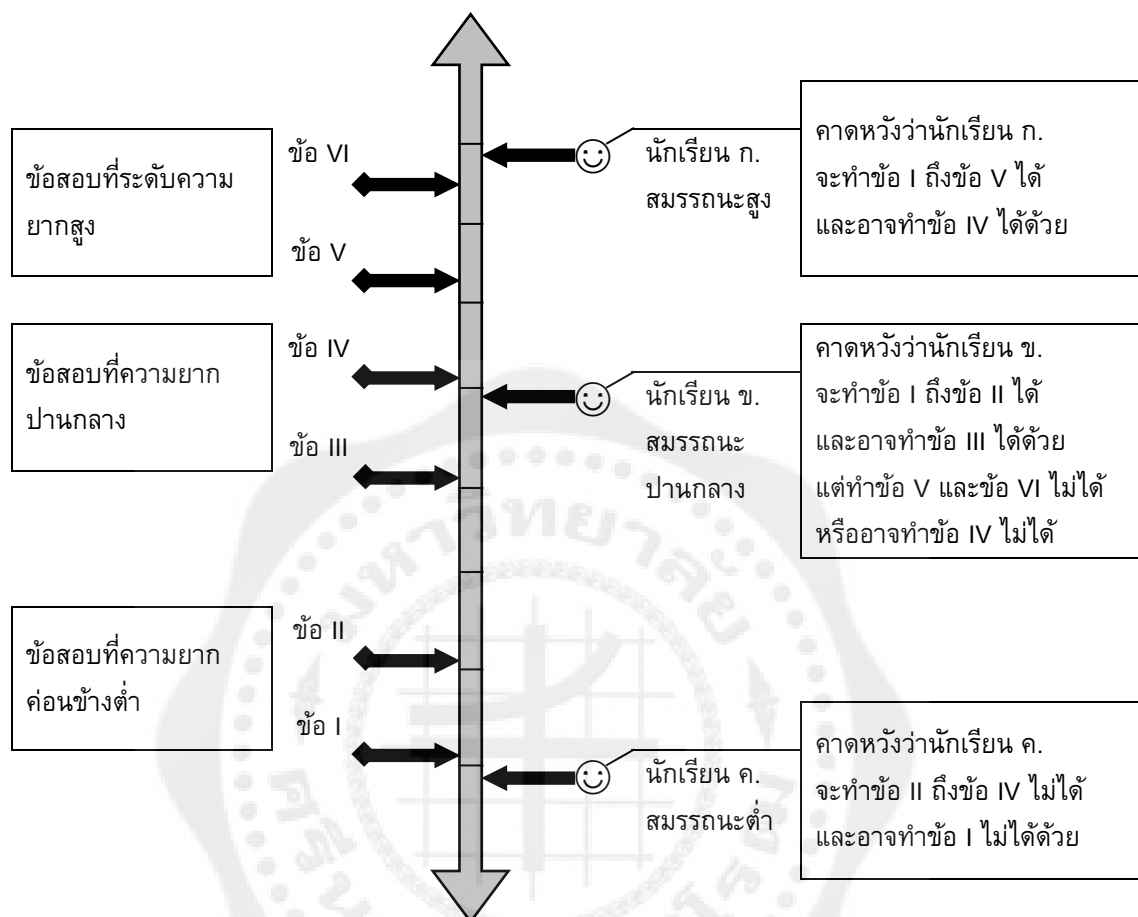
ใช้และผูกพัน หมายถึง ความครอบคลุมถึงการใช้คณิตศาสตร์ การแก้ปัญหาทางคณิตศาสตร์ และยังใช้ในความหมายที่กว้างขึ้น ครอบคลุมถึงการที่บุคคลหนึ่งจะเกี่ยวข้องกับคณิตศาสตร์ ไม่ว่าจะเป็นทางการสื่อสาร การสัมพันธ์เชื่อมโยงกับคณิตศาสตร์ การประมาณการ หรือแม้กระทั่ง ความชอบและความสนุกที่ได้ทำหรือเกี่ยวข้องกับคณิตศาสตร์

ต่อชีวิตของแต่ละบุคคล หมายถึง ชีวิตส่วนตัว การงานอาชีพ และชีวิตด้านสังคมกับเพื่อนกับญาติพี่น้อง และชีวิตในฐานะพลเมืองของชุมชนและประเทศชาติ

นิยามระดับการรู้เรื่องคณิตศาสตร์

ข้อสอบใน PISA 2012 ก็เหมือนการสอบครั้งก่อนๆ คือ ออกแบบให้สามารถบอสมรรถนะทางคณิตศาสตร์ของนักเรียนตามมาตรฐานที่กำหนดไว้ โดยคำถามแต่ละข้อที่ระบุระดับความยากและคะแนนไว้มีความสัมพันธ์กับตำแหน่งบนมาตรฐาน ผลการสอบของนักเรียนจะสัมพันธ์กับระดับสมรรถนะคณิตศาสตร์ของนักเรียนที่ระบุไว้บนมาตรฐานวัด PISA 2012 กำหนดคะแนนของนักเรียนและความยากของภารกิจออกเป็นระดับต่างๆ 6 ระดับ ตั้งแต่ระดับต่ำสุด (ระดับ 1) จนถึงระดับสูงสุด (ระดับ 6) ดังแสดง ในภาพประกอบ 2 (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2557: 33)

มาตรวัดการรู้เรื่องคณิตศาสตร์



ภาพประกอบ 2 ความสัมพันธ์ระหว่างคำถามกับระดับความสามารถการรู้เรื่องคณิตศาสตร์
ของนักเรียน

ที่มา: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2557: 34)

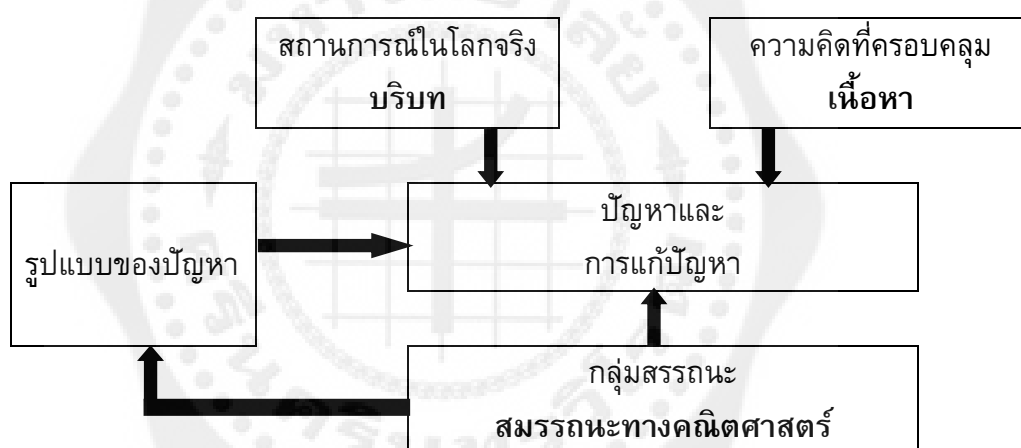
1.2 การจัดกรอบโครงสร้างการประเมินผลความรู้เรื่องคณิตศาสตร์

กรอบการประเมินผลของ OECD/PISA เน้นที่การประเมินว่านักเรียนอายุ 15 ปี รู้เรื่องคณิตศาสตร์มากน้อยเพียงใด นั่นคือสามารถนำฐานความรู้คณิตศาสตร์มาใช้และเผชิญหน้ากับปัญหาในโลกจริงได้เพียงใด ขอบเขตของคณิตศาสตร์ครอบคลุมองค์ประกอบ 3 ด้านด้วยกัน ได้แก่

ด้านแรก คือ สถานการณ์หรือบริบทที่ปัญหานั้นตั้งอยู่ (situation on context)

ด้านที่สอง คือ เนื้อหาคณิตศาสตร์ (mathematical content) ที่ต้องนำมาใช้ในการแก้ปัญหา

และด้านที่สาม คือ สมรรถนะทางคณิตศาสตร์ของนักเรียนที่ควรได้รับการกระตุ้นให้สามารถเชื่อมต่อกับโลกจริงที่ปัญหานั้นๆ เกิดขึ้นโดยใช้คณิตศาสตร์ และให้สามารถแก้ปัญหาได้โดยใช้คณิตศาสตร์นั้นๆ



ภาพประกอบ 3 องค์ประกอบและกรอบโครงสร้างการประเมินคณิตศาสตร์

ที่มา: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2554: 89)

สถานการณ์หรือบริบท ซึ่งปัญหาทางคณิตศาสตร์ตั้งอยู่สถานการณ์หรือบริบทที่ใกล้กับตัวเด็กมากที่สุด คือ บริบทส่วนตัว ถัดมาคือบริบทในโรงเรียน ในการงานอาชีพ บริบทในชุมชนหรือสังคมสาธารณะ ที่ห่างออกไปคือบริบททางวิทยาศาสตร์ บางกรณีปัญหาอาจเป็นบริบทในวิชาคณิตศาสตร์ด้วยกัน กรณีนี้จัดเป็นบริบทภายในคณิตศาสตร์ (Intra-mathematics)

สาระเนื้อหาที่ครอบคลุม

แนวคิดทางคณิตศาสตร์ได้ถูกสร้างขึ้นมาเพื่อเป็นเครื่องมืออันที่เข้าใจ จัดระเบียบ และวิเคราะห์ปรากฏการณ์ต่าง ๆ ในธรรมชาติ ในสังคมและในการคิดจินตนาการ ต่าง ๆ ในโรงเรียน หลักสูตรคณิตศาสตร์จะถูกจัดเป็นสาขาวิชา (เลข พีชคณิต เรขาคณิต ฯลฯ) ที่สะท้อนถึงที่มา แนวคิดที่ยึดถือมาและเป็นฐานของการจัดการแผนการเรียนการสอน อย่างไรก็ตามในโลกของความเป็นจริง ปรากฏการณ์ที่เกี่ยวข้องกับคณิตศาสตร์ไม่ได้จัดระเบียบมาเป็นหมวดหมู่หรือแยกสาขาวิชามาให้ และก็ไม่ค่อยมีปรากฏการณ์ใดที่สามารถใช้ความรู้จากสาขาวิชา เดียวโดดๆ มาแก้ปัญหาได้ หากแต่ต้องการใช้พื้นฐานความรู้ที่กว้างขวางครอบคลุมหลายด้าน กว่าที่ใช้อยู่ในห้องเรียน

เนื่องจากระดับของการรู้เรื่องคณิตศาสตร์จะดูจากคนนั้นสามารถใช้ความรู้และ ทักษะทางคณิตศาสตร์มาแก้ปัญหาในโลกของความเป็นจริงตามสถานการณ์หรือบริบทที่ แตกต่างหลากหลายได้ดีเพียงใด ดังนั้นในการประเมินใช้ปรากฏการณ์เป็นตัวตั้งในการนำไปสู่ แนวคิดโครงสร้างหรือความคิดหลักการทางคณิตศาสตร์ วิธีนี้จึงประกันได้ว่าจะตรงกับ จุดมุ่งหมายในนิยามของการประเมิน ซึ่งไม่เหมือนกับการประเมินผลคณิตศาสตร์ที่พบเห็นใน หลักสูตรทั่วไป (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2554: 90)

โครงสร้างการประเมินคณิตศาสตร์ที่ครอบคลุม 4 เรื่องต่อไปนี้ เป็นแนวคิดที่ ครอบคลุมเนื้อหาสาระที่นักเรียนได้เรียนมาแล้วตามหลักสูตรคณิตศาสตร์ในโรงเรียน ได้แก่ ปริภูมิและรูปทรงสามมิติ, การเปลี่ยนแปลงและความสัมพันธ์, ปริมาณ และความไม่แน่นอน แต่ละเรื่องมีลักษณะและรายละเอียดดังนี้

1) ปริภูมิและรูปทรงสามมิติ (Space and Shape)

เรื่องของแบบรูป (Pattern) มีอยู่ทุกหนทุกแห่งในโลก แม้แต่ในการพูด ดนตรี การก่อสร้าง ศิลปะ ฯลฯ รูปทรงเป็นแบบรูปที่เห็นได้ทั่วไป เป็นต้นว่ารูปร่างของบ้าน โรงเรียน อาคาร สะพาน ถนน ดอกไม้ ฯลฯ จึงเป็นเรื่องสำคัญที่นักเรียนจะเข้าใจลักษณะประจำตัวของ วัตถุและตำแหน่งเปรียบเทียบกับ ต้องเข้าใจว่าจะมองวัตถุ สิ่งของนั้นๆอย่างไรและทำไม ซึ่งเป็น เครื่องหมายแสดงถึงความเข้าใจความสัมพันธ์ระหว่างรูปร่างกับภาพ เช่น ระหว่างภาพเมืองกับ รูปร่างของตัวเมือง และแสดงถึงความเข้าใจว่าจะนำเสนอรูปร่างสามมิติออกมาในรูปของสองมิติ ได้อย่างไร สามารถบอกว่าเงาเกิดขึ้นได้อย่างไร และแปลงเงาออกมาได้เรื่องของรูปทรงแม้จะดู เป็นเรขาคณิต แต่ก็นิยามออกไปไกลกว่าสาระของวิชาเรขาคณิต ทั้งในเนื้อหาความหมายและ วิธีการขยายกว้างไปถึงเรื่องของทฤษฎีการมองเห็นรู้ว่ารูปทรงเหมือนหรือแตกต่างกัน วิเคราะห์ความแตกต่างขององค์ประกอบรู้จักในมิติต่างๆ (สถาบันส่งเสริมการสอนวิทยาศาสตร์ และเทคโนโลยี. 2554: 91)

ลักษณะสำคัญของการรู้เรื่องปริภูมิและรูปทรงสามมิติ ได้แก่

1. รู้จักรูปทรงและแบบรูปของ (patterns) ของรูปทรง
2. สามารถพรรณนา เข้ารหัส และถอดรหัสของสาระที่มองจากภาพได้
3. เข้าใจการเปลี่ยนแปลงในรูปทรง
4. ระบุความเหมือนและความแตกต่าง
5. ระบุตำแหน่งเปรียบเทียบ
6. แปลความรูปสองมิติและสามมิติ บอกความสัมพันธ์ระหว่างทั้งสองอย่างได้
7. มองเห็นแนวทางผ่านปริภูมิได้

2) การเปลี่ยนแปลงและความสัมพันธ์ (Change and Relationships)

ปรากฏการณ์ทุกอย่างในโลกล้วนเกิดการเปลี่ยนแปลงมากมายมหาศาล แสดงให้เห็นถึงความสัมพันธ์ทั้งชั่วคราวและถาวรระหว่างการเปลี่ยนแปลงต่างๆ ตัวอย่างเช่น มีการเปลี่ยนแปลงของสิ่งมีชีวิตขณะเจริญเติบโต การหมุนเวียนของฤดูกาล การขึ้นลงของกระแสน้ำ การเปลี่ยนแปลงของอวกาศ การขึ้นลงของหุ้นการว่างงานของคน การเปลี่ยนแปลงบางกระบวนการสามารถบอกได้หรือสร้างเป็นตัวแทนได้โดยตรงโดยใช้ฟังก์ชันทางคณิตศาสตร์ ความสัมพันธ์ทางคณิตศาสตร์ส่วนมากเป็นรูปของสมการหรือสมการ แต่ความสัมพันธ์ในธรรมชาติอื่นๆ ก็อาจเกิดขึ้นได้เช่นกัน ความสัมพันธ์หลายอย่างโดยเฉพาะในเรื่องธรรมชาติไม่สามารถใช้คณิตศาสตร์ได้ก็เกิดขึ้นได้เช่นกัน

การนำเสนอการเปลี่ยนแปลงและความสัมพันธ์อาจแทนได้หลายวิธีด้วยกัน ได้แก่ เชิงตัวเลข (เช่น ในตาราง) เชิงสัญลักษณ์ กราฟ เชิงพีชคณิต และเรขาคณิต การแปลความระหว่างการแทนเหล่านั้นเป็นสิ่งสำคัญนักเรียนควรจะต้องรู้เรื่องของการเปลี่ยนแปลงเชิงเส้นตรง (กระบวนการบวก) การเปลี่ยนแปลงเชิงเอ็กซ์โพเนนเชียล (คูณ) และการเปลี่ยนแปลงเชิงพีรีอดิก (Periodic) กระบวนการโลจิสติก (Logistic) นักเรียนควรต้องมองเห็นความสัมพันธ์ระหว่างแบบของการเปลี่ยนแปลงต่างๆ เหล่านี้ เช่น ความแตกต่างหลักระหว่างกระบวนการเปลี่ยนแปลงเชิงเส้นตรงกับเชิงเอ็กซ์โพเนนเชียลรู้ว่าการเปลี่ยนแปลงในกระบวนการโลจิสติกเกิดขึ้นอย่างไรและเพราะเหตุใด ทั้งในแบบต่อเนื่องและไม่ต่อเนื่อง (Discrete) (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2554: 91 - 92)

3) ปริมาณ (Quantity)

จุดหลักของปริมาณ ก็นับถึงความเข้าใจเรื่องขนาดเปรียบเทียบ การรู้จักแบบรูปของจำนวน และการใช้จำนวนแทนปริมาณและแทนวัตถุต่างๆ ในโลกจริง (การนับและการวัด) นอกจากนี้ปริมาณยังเกี่ยวข้องกับกระบวนการและความเข้าใจเรื่องจำนวนในเรื่องต่างๆ อย่างหลากหลาย กลยุทธ์ที่ใช้ในการคิดเกี่ยวกับปริมาณ ได้แก่ การใช้เหตุผลเชิงปริมาณ (Quantitative reasoning) ซึ่งประกอบด้วยความรู้สึกเชิงจำนวน (Number sense) การแทนจำนวนวิธีต่างๆ ความเข้าใจวิธีการรู้ค่าของจำนวน การคิดเลขอย่างฉลาด การคิดเลขในใจ และการประมาณค่า การใช้จำนวนในชีวิตจริงที่พบบ่อยที่สุด มักเป็นการวัดขนาด เช่น ความยาว พื้นที่ ความสูง มวลความดันอากาศ ค่าของเงิน รวมทั้งการวัดต่างๆ ทุกชนิด ความเข้าใจวิธีการหมายถึงความสามารถที่จะทำงานคณิตศาสตร์ที่เกี่ยวข้องกับการเปรียบเทียบ อัตราส่วนและร้อยละ ความรู้สึกเชิงจำนวนบอกถึงขนาดเปรียบเทียบ การใช้ตัวเลขแทนด้วยวิธีการต่างๆ แบบรูปสมมูลของจำนวนและการใช้สิ่งเหล่านี้บอกถึงคุณลักษณะของสิ่งที่ปรากฏในโลกธรรมชาติ ปริมาณยังหมายรวมถึง ความรู้สึกเชิงปริมาณและการประมาณค่านั้น คือ ต้องมีความรู้กว้างเกี่ยวกับ ปริมาณ (การวัด) เป็นต้นว่า รู้ว่ารถยนต์กำลังวิ่งด้วยความเร็วเฉลี่ย 50 หรือ 100 หรือ 200 กิโลเมตรต่อชั่วโมงดีสูงประมาณเท่าไร แม่น้ำกว้างสักเท่าไร เป็นต้น

การใช้ความรู้สึกเชิงตัวเลขที่เหมาะสมจะทำให้นักเรียนสามารถแก้ปัญหาที่ตรงไปตรงมา ที่กลับกันหรือที่ต้องการเรื่องของสัดส่วนมาใช้สามารถประมาณอัตราของการเปลี่ยนแปลง และบอกเหตุผลในการเลือกใช้ข้อมูลและระดับของความถูกต้องที่ต้องการสำหรับเรื่องหนึ่งๆ สามารถเลือกวิธีการและลำดับขั้นตอนเพื่อแสดงว่าทำไมจึงเป็นเช่นนั้นและมีกรณีใดบ้างที่ทำไม่สำเร็จ สามารถสร้างตัวแบบของวิธีการที่ใช้ สำหรับแก้ปัญหาที่ใช้ข้อมูลจากที่มีอยู่ในโลก (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2554: 93 - 94)

4) ความไม่แน่นอน (Uncertainty)

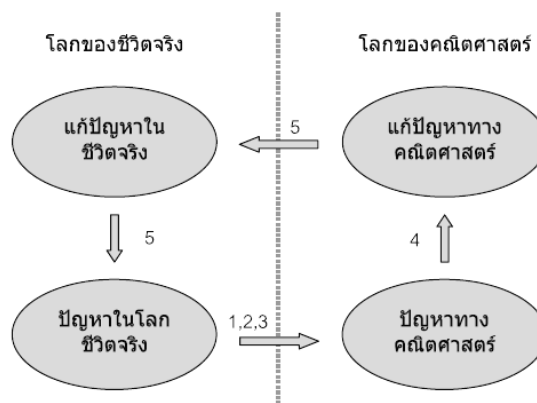
ในวงการวิทยาศาสตร์และเทคโนโลยีมักเกี่ยวข้องกับความไม่แน่นอนอยู่มากในเรื่องของความรู้วิทยาศาสตร์มีบ่อยครั้งที่พบว่าไม่ถูกหรือ ไม่ใช้อย่างที่รู้และเข้าใจกันอยู่ ณ ขณะนั้น นั่นหมายความว่ามีความไม่แน่นอนเกิดขึ้นตลอดเวลา แม้กระทั่งการพยากรณ์ทางวิทยาศาสตร์ เช่น การพยากรณ์อากาศ ความไม่แน่นอนยังเกิดขึ้นตลอดเวลาในชีวิตจริง เช่น ผลการเลือกตั้ง การพยากรณ์อากาศที่ไม่เที่ยงตรง การล้มละลายทางเศรษฐกิจ การเงิน การพยากรณ์ต่างๆ ที่ผิดพลาด แสดงให้เห็นถึงความไม่แน่นอนของโลก ความไม่แน่นอนมักจะเกี่ยวข้องกับสองเรื่อง คือ ข้อมูลและโอกาส ซึ่งปรากฏการณ์ทั้งสองเรื่องเป็นของสถิติและความน่าจะเป็น ประเทศสมาชิก OECD จึงมีข้อเสนอแนะสำหรับหลักสูตรคณิตศาสตร์ในโรงเรียนที่ให้ความสำคัญกับเรื่องของสถิติและความน่าจะเป็นให้เป็นจุดเด่นมากขึ้นกว่าในอดีต ด้วยเหตุผลที่ว่าโลกปัจจุบันในยุคของ “สังคมข้อมูลข่าวสาร” ข้อมูลข่าวสารที่หลั่งไหลเข้ามาและแม้ว่าจะอ้าง

ว่าเป็นข้อมูลที่ถูกต้องตรวจสอบได้ก็จริง แต่ในชีวิตจริงเราก็ต้องเผชิญกับความไม่แน่นอนหลายอย่าง เช่น ผลการเลือกตั้งที่ไม่คาดคิด คณิตศาสตร์ที่เข้ามามีบทบาทในส่วนนี้คือ การเก็บข้อมูล การวิเคราะห์ข้อมูล การเสนอข้อมูลความน่าจะเป็น และการอ้างอิง (สถิติ) สถิตินับเป็นการเพิ่มเนื้อหให้กับการศึกษาคณิตศาสตร์ที่มีความสำคัญและเฉพาะตัว นั่นคือการใช้เหตุผลจากข้อมูลที่ประจักษ์ชัด ซึ่งควรจะเป็นส่วนหนึ่งของวิธีการคิดการใช้เหตุผลของพลเมืองที่มีสติปัญญา (เพราะว่าโดยทั่วไปคนส่วนหนึ่งซึ่งมีจำนวนไม่น้อยมักตัดสินใจหรือตัดสินใจปัญหาต่างๆ ด้วยการใช้เฉพาะความเห็นที่ไม่มีข้อมูลหรือการใช้เหตุผลทางสถิติเป็นเครื่องมือ และมีไม่น้อยที่ปัญหาสำคัญของชาติถูกตัดสินบนพื้นฐานของความเห็นแทนที่จะเป็นพื้นฐานของการใช้เหตุผลและสถิติจากข้อมูลที่ประจักษ์) (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2554: 95) องค์ประกอบหลักของกระบวนการคิดเชิงสถิติ ประกอบด้วย การมีความแปรผันทั่วไป ในกระบวนการต้องการข้อมูลในกระบวนการ ออกแบบการสร้างข้อมูลโดยคิดถึงความแปรผันไว้ด้วย การทำข้อมูลให้เป็นเชิงปริมาณ และการอธิบายความแปรผัน

คำว่า “ข้อมูล” ไม่ใช่เป็นเพียงตัวเลข หากเป็นตัวเลขที่มีบริบท ข้อมูลได้มาจากการวัดและแทนด้วยจำนวน การวัดทำให้จำนวนบางจำนวนมีความหมายมีสาระ แต่บางจำนวนก็ไม่มี ความหมายหรือสาระการออกแบบการสุ่มตัวอย่างเป็นหัวข้อหลักของสถิติ ทำให้การวิเคราะห์ข้อมูลเป็นที่เข้าใจขึ้นและเป็นตัวแทนของประชากรเรื่องการสุ่มตัวอย่างเป็นเรื่องที่นักเรียนอายุ 15 ปี ควรต้องรู้ ความไม่แน่นอนที่เกิดขึ้นไม่ว่าจะเป็นชนิดเดียวหรือเป็นแบบรูปซ้ำกัน มักได้จากการสุ่มตัวอย่างทั้งสิ้น (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2554: 95) แนวคิดคณิตศาสตร์ที่เกี่ยวข้องกับเรื่องนี้ได้แก่ การหาข้อมูล การวิเคราะห์และการนำเสนอข้อมูล หรือทำให้ข้อมูลเป็นที่ประจักษ์ชัด ความน่าจะเป็น และการอ้างอิง

1.3 กระบวนการคิดเป็นคณิตศาสตร์

เนื่องจาก PISA ต้องการตรวจสอบความสามารถของนักเรียนในด้านการแก้ปัญหา ซึ่งครอบคลุมการวิเคราะห์ การใช้เหตุผล และการสื่อสารแนวคิดทางคณิตศาสตร์อย่างมีประสิทธิภาพ ในการแก้ปัญหานั้น นักเรียนจะต้องใช้กระบวนการความรู้ และทักษะคณิตศาสตร์ ที่ได้เรียนมาในโรงเรียนและจากประสบการณ์ชีวิต สำหรับในการประเมินผลของ PISA จะเรียกกระบวนการพื้นฐานที่นักเรียนใช้แก้ปัญหาที่ปรากฏในชีวิตจริงว่ากระบวนการคิดให้เป็นคณิตศาสตร์ (Mathematisation) ซึ่งมีวงจรดังนี้



ภาพประกอบ 4 กระบวนการคิดเป็นคณิตศาสตร์

ที่มา : สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2554: 96)

ทฤษฎีพื้นฐานกระบวนการคิดเป็นคณิตศาสตร์ PISA กำหนดเป็น 5 ขั้นตอน ดังภาพประกอบ 4 สามารถสรุปกระบวนการได้ดังนี้ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2554: 99 - 100)

ขั้นที่ 1 เริ่มด้วยปัญหาที่อยู่ในโลกจริง ต้องแปลปัญหาจากความเป็นจริงไปเป็นปัญหาคณิตศาสตร์ สามารถระบุคณิตศาสตร์ที่ใช้ได้กับปัญหานั้นๆ แสดงปัญหาในรูปที่แตกต่างไป รวมทั้งจัดแนวความคิดคณิตศาสตร์ที่เกี่ยวข้องและระบุข้อตกลงเบื้องต้นที่สอดคล้องเหมาะสม

ขั้นที่ 2 จัดให้อยู่ในรูปแบบตามแนวความคิดคณิตศาสตร์ เข้าใจความสัมพันธ์ระหว่างภาษาของปัญหาในโลกจริงกับภาษา สัญลักษณ์ สูตรทำให้เข้าใจปัญหานั้นในทางคณิตศาสตร์ มองหารูปแบบ ความสัมพันธ์และแบบรูปคณิตศาสตร์ และมองลักษณะปัญหาในเชิงคณิตศาสตร์

ขั้นที่ 3 ค่อยๆ ตัดข้อเท็จจริงที่เป็นปัญหาในโลกจริงออกไปก่อนโดยนำคณิตศาสตร์เข้ามาเชื่อมโยงกับปัญหา เช่น การสร้างข้อตกลงเบื้องต้น การทำให้เป็นโจทย์คณิตศาสตร์ การลงข้อสรุป และแปลปัญหาให้เป็นโจทย์ปัญหาคณิตศาสตร์

ขั้นที่ 4 แก้ปัญหาคณิตศาสตร์ ขั้นตอนนี้รวมถึงการใช้และการแสดงแทนเปลี่ยนแปลงไปมา การใช้สัญลักษณ์ กฎ ภาษาเฉพาะทาง และการทำโจทย์คณิตศาสตร์ ใช้ ปรับตัวแบบทางคณิตศาสตร์ ผสมผสานและบูรณาการตัวแบบ ให้ความเห็น สนับสนุน โต้แย้ง และสรุปการแก้โจทย์

ขั้นที่ 5 แปลผลจากการแก้ปัญหาคณิตศาสตร์ให้กลับเป็นปัญหาในสถานการณ์ของโลกจริง รวมถึงการระบุข้อจำกัดของการแก้ปัญหานั้นๆ ด้วย เข้าใจว่าคณิตศาสตร์ทำได้แค่ไหน และมีข้อจำกัดอย่างไร คิดสะท้อนถึงข้ออภิปราย โต้แย้ง และหาคำอธิบายถึงความใช้ได้ของการแก้โจทย์ สื่อสารทั้งกระบวนการคิดและผลที่ได้ วิพากษ์ตัวแบบและข้อจำกัด

1.4 สมรรถนะทางคณิตศาสตร์ (Mathematical Competencies)

ในการคิดเชิงคณิตศาสตร์ตามที่กล่าวมานั้น นอกจากต้องการความรู้เนื้อหา คณิตศาสตร์ และกระบวนการทางคณิตศาสตร์แล้ว สิ่งสำคัญที่ทำให้สามารถคิดในเชิง คณิตศาสตร์ได้ คือ สมรรถนะทางคณิตศาสตร์ ซึ่งแต่ละบุคคลสามารถมีสมรรถนะได้ในระดับที่ แตกต่างกันได้ สมรรถนะทางคณิตศาสตร์อาจมีมากมายหลายอย่าง แต่ในการวัดและประเมินผล ของ OECD/PISA ได้ตัดสินใจเลือกใช้ 8 สมรรถนะ ได้แก่

1. การคิดและการใช้เหตุผล (Thinking and Reasoning) สมรรถนะนี้เกี่ยวข้องกับ ความสามารถในการตั้งคำถาม รู้คำตอบทางคณิตศาสตร์ บอกความแตกต่างของประโยค (statements) (เช่น นิยามทฤษฎี conjecture สมมติฐาน ตัวอย่าง ฯลฯ) และความเข้าใจและการ ใช้ข้อจำกัดของคณิตศาสตร์

2. การสร้างข้อโต้แย้ง (Argumentation) เกี่ยวข้องกับการรู้จักการพิสูจน์ทาง คณิตศาสตร์ (และรู้ว่าการพิสูจน์แตกต่างจากการใช้เหตุผลอย่างไร) สามารถติดตามและประเมิน การโต้แย้งทางคณิตศาสตร์แบบต่างๆ มีความรู้สึกถึงความจริง (รู้ว่าอะไรเกิดขึ้นได้/ไม่ได้ และ ทำไม) และสามารถสร้างและแสดงการโต้แย้งทางคณิตศาสตร์

3. การสื่อสาร (Communication) เกี่ยวข้องกับการแสดงออกของตนเอง ความสามารถที่ทำให้ผู้อื่นเข้าใจตน โดยวิธีการต่างๆ บนพื้นฐานของคณิตศาสตร์ ทั้งในรูปของ การพูดและการเขียน และสามารถเข้าใจการพูดและ การเขียนของผู้อื่นด้วยเช่นกัน

4. การสร้างตัวแบบ (Modeling) เกี่ยวข้องกับการวางโครงสร้างของ สถานการณ์ที่จะต้องนำมาสร้างเป็นตัวแบบ (Model) การแปลความเป็นจริงให้เข้าสู่โครงสร้าง ทางคณิตศาสตร์ การประเมินความน่าเชื่อถือของตัวแบบ วิเคราะห์ วิจารณ์ตัวแบบและผลที่ เกิดขึ้น การสื่อสารแนวคิดของตัวแบบและผล (รวมทั้งข้อจำกัด) การติดตามและควบคุม กระบวนการของการสร้างตัวแบบ

5. การตั้งและการแก้ปัญหา (Problem posing and solving) เป็นสมรรถนะที่ เกี่ยวข้องกับการตั้งคำถามการสร้างเป็นปัญหาคณิตศาสตร์ และการนิยามปัญหาคณิตศาสตร์ แบบต่างๆ (เช่น การแก้ การประยุกต์คำถามเปิด คำถามปิด) และการแก้ปัญหาคณิตศาสตร์ แบบต่างๆ โดยวิธีการที่หลากหลาย

6. การแสดงเครื่องหมายแทน (Representation) สมรรถนะด้านนี้เกี่ยวข้องกับการ แปลรหัส (decoding) และการเข้ารหัส (encoding) การแปลความ การตีความ และการบอก ความแตกต่างของการแสดงเครื่องหมายของคณิตศาสตร์แบบต่างๆ และความสัมพันธ์ระหว่าง การแสดงเครื่องหมายแทนแบบต่างๆ การเลือกและการเปลี่ยนระหว่างรูปแบบต่างๆ ของการ แสดงเครื่องหมายแทน (Representation) ให้สอดคล้องกับสถานการณ์และจุดประสงค์

7. การใช้สัญลักษณ์ ภาษา และการดำเนินการ (Using symbolic, language and operation) เกี่ยวข้องกับการแปลรหัส การตีความสัญลักษณ์ ภาษาคณิตศาสตร์ และความเข้าใจการเชื่อมโยงของภาษาคณิตศาสตร์กับภาษาธรรมชาติ การแปลความจากภาษาธรรมชาติไปเป็นสัญลักษณ์/ภาษาคณิตศาสตร์สามารถจัดการกับประโยคหรือพจน์ที่มีสัญลักษณ์และสูตรความสามารถในการใช้ตัวแปร การแก้สมการ และการคำนวณ

8. ใช้ตัวช่วยและเครื่องมือ (Using aids and tools) สมรรถนะนี้เกี่ยวข้องกับการรับรู้ และความสามารถในการใช้ตัวช่วยและเครื่องมือ (รวมทั้งเครื่องมือภาคเทคโนโลยีสารสนเทศ) ที่สามารถช่วยกิจกรรมทางคณิตศาสตร์ นอกจากนี้ยังหมายถึงความรู้ถึงข้อจำกัดของเครื่องมืออื่นๆ ด้วย

อย่างไรก็ตาม PISA ไม่ได้ต้องการที่จะสร้างข้อสอบเพื่อวัดสมรรถนะต่างๆ เฉพาะแต่ละสมรรถนะโดยตรงเพราะสมรรถนะของคนไม่ใช่สิ่งที่จะแยกออกมาวัดได้โดดๆ แต่ในการแสดงความสามารถอย่างใดอย่างหนึ่งอาจมีหลายสมรรถนะซ้อนกันอยู่ ดังนั้น PISA จึงไม่ได้วัดสมรรถนะโดดๆ แต่ในการตอบข้อสอบ นักเรียนจำเป็นต้องมีและสามารถใช้สมรรถนะดังกล่าว แต่อาจจะใช้หลายสมรรถนะหรือเป็นกลุ่มของสมรรถนะในการแก้ปัญหา รวมไปถึงเป็นสามกลุ่ม คือ

กลุ่มการทำใหม่ (Reproduction Cluster) กลุ่มสมรรถนะนี้ หมายถึง การทำคณิตศาสตร์ตามแบบตัวอย่างที่เคยฝึกฝนมาแล้ว โดยใช้กระบวนการความรู้ และทักษะทางคณิตศาสตร์ทั่วไปที่มักใช้อยู่ในการสอบคณิตศาสตร์ตามมาตรฐานโรงเรียน ซึ่งมักเป็นการแก้โจทย์ การคำนวณแบบเดิมๆ แต่ก็ต้องใช้สมรรถนะทางคณิตศาสตร์ทุกสมรรถนะ ข้างต้นในการแก้ปัญหาโจทย์

กลุ่มการเชื่อมโยง (Connection Cluster) กลุ่มการเชื่อมโยงต่อยอดมาจากกลุ่มการทำใหม่ โดยประยุกต์ต่อไปใช้แก้ปัญหาในสถานการณ์ที่ไม่เหมือนเดิม ไม่ได้พบบ่อยๆ เป็นประจำ แต่เนื้อหาของปัญหายังคงเกี่ยวข้องกับสมรรถนะกลุ่มแรกอยู่บ้างบางส่วน

กลุ่มการสะท้อนและสื่อสาร (Reflection and Communication Cluster) สมรรถนะกลุ่มการสะท้อนและสื่อสารทางคณิตศาสตร์นี้ เรื่องของการคิดไตร่ตรอง สะท้อนกลับที่นักเรียนต้องใช้ในการแก้ปัญหาเข้ามารวมอยู่ด้วย จึงเกี่ยวข้องกับความสามารถในการวางแผน กลยุทธ์การแก้ปัญหา และใช้กลยุทธ์นั้นในการแก้ปัญหตามสถานการณ์ของปัญหานั้นซึ่งมักมีองค์ประกอบที่เพิ่มมากขึ้นหรือซับซ้อนขึ้นหรือมีความหมายใหม่ (หรือไม่คุ้นเคย) มากขึ้นกว่าในกลุ่มการเชื่อมโยง

การจำแนกกลุ่มสมรรถนะทางคณิตศาสตร์ มีลักษณะจำเพาะต่างกัน สรุปได้ดังแผนภูมิต่อไปนี้



ภาพประกอบ 5 แผนภูมิสรุปสมรรถนะทางคณิตศาสตร์

ที่มา: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2554: 100 - 102)

1.5 การประเมินการรู้เรื่องคณิตศาสตร์

กรอบการประเมินผลของ OECD/PISA เน้นที่การประเมินว่า นักเรียนอายุ 15 ปี รู้เรื่องคณิตศาสตร์มากน้อยเพียงใด สามารถนำฐานความรู้คณิตศาสตร์มาใช้ และเผชิญหน้ากับปัญหาในโลกจริงได้เพียงใดขอบเขตของคณิตศาสตร์ครอบคลุมองค์ประกอบ 3 ด้าน ได้แก่ กระบวนการทางคณิตศาสตร์ (process) ที่อธิบายสิ่งที่แต่ละคนทำเพื่อเชื่อมโยงบริบทของปัญหากับคณิตศาสตร์ แล้วนำไปสู่การแก้ปัญหา เนื้อหาคณิตศาสตร์ (content) ที่ต้องนำมาใช้ในการแก้ปัญหา สถานการณ์หรือบริบท (contexts) ที่ปัญหานั้นตั้งอยู่ (โครงการ PISA ประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2557: 2)

กรอบการประเมินการรู้เรื่องคณิตศาสตร์ของ PISA 2015 มีดังนี้



ภาพประกอบ 6: กรอบการประเมินการรู้เรื่องคณิตศาสตร์ของ PISA 2015

ที่มา: สุชาดา ปัทมวิภาต.(2557. พฤษภาคม – มิถุนายน) นิตยสารสสวท. 42(188): 36

ภารกิจการประเมินการรู้เรื่องทางคณิตศาสตร์ของ PISA จึงให้ความชัดเจนที่ความต้องการให้นักเรียนเผชิญหน้ากับปัญหาทางคณิตศาสตร์ที่มีอยู่ในแวดวงของการดำเนินชีวิต ซึ่งต้องการให้นักเรียนระบุสถานการณ์ที่สำคัญของปัญหา กระตุ้นให้หาข้อมูล สืบรวจตรวจสอบ และนำไปสู่การแก้ปัญหาในกระบวนการนี้ต้องการทักษะหลายอย่าง เป็นต้นว่า ทักษะการคิด และการใช้เหตุผล ทักษะการโต้แย้ง การสื่อสาร ทักษะการสร้างตัวแบบ การตั้งปัญหาและการแก้ปัญหา การนำเสนอ การใช้สัญลักษณ์ การดำเนินการ ในกระบวนการเหล่านี้นักเรียนต้องใช้ทักษะต่างๆ ที่หลากหลายมารวมกัน หรือใช้ทักษะหลายอย่างที่ทับซ้อนหรือคาบเกี่ยวกัน

ดังนั้นการที่ PISA เลือกใช้คำว่า “การรู้เรื่องคณิตศาสตร์” แทนคำว่า “ความรู้คณิตศาสตร์” ก็เพื่อเน้นความชัดเจนของความรู้คณิตศาสตร์ที่นำมาใช้ในสถานการณ์ต่างๆ ทั้งนี้โดยถือข้อตกลงเบื้องต้นว่าการที่คนหนึ่งจะใช้คณิตศาสตร์ได้ คนนั้นจะต้องมีความรู้พื้นฐานและทักษะทางคณิตศาสตร์มากพออยู่แล้ว ซึ่งนั่นก็หมายถึงสิ่งที่นักเรียนได้เรียนไปขณะอยู่ในโรงเรียน เจตคติและความรู้สึกที่เกี่ยวข้องกับคณิตศาสตร์ เช่น ความมั่นใจ ความอยากรู้อยากเห็น ความสนใจความรู้สึกว่าตรงปัญหาหรือตรงกับประเด็น และความอยากรู้อยากเข้าใจสิ่งต่างๆ รอบตัว แม้จะไม่ถือว่าเป็นเรื่องคณิตศาสตร์ แต่ก็ถือว่ามีส่วนสำคัญในการทำให้รู้เรื่องคณิตศาสตร์ เพราะโดยความเป็นจริงแล้วการรู้เรื่องคณิตศาสตร์จะไม่เกิดขึ้น หากบุคคลขาดเจตคติและความรู้สึกต่อคณิตศาสตร์ และมีหลักฐานเป็นที่ยอมรับว่ามีความเกี่ยวข้องสัมพันธ์กันระหว่างการรู้เรื่องทางคณิตศาสตร์กับเจตคติและความรู้สึกต่อคณิตศาสตร์ในการประเมินผลของ PISA จะไม่มีการวัดด้านนี้ใดๆ โดยตรง แต่จะมีการหยิบยกมาพิจารณาในบางองค์ประกอบของการประเมิน (การวัดผลประเมินผลเพื่อคุณภาพการเรียนรู้และตัวอย่างข้อสอบจากโครงการประเมินผลนักเรียนนานาชาติ PISA. 2550: 99)

แบบทดสอบ PISA ประเมินว่านักเรียนสามารถทำกระบวนการต่อไปนี้ได้ดีเพียงใด

การคิดสถานการณ์ของ ปัญหาในเชิง คณิตศาสตร์	- ระบุประเด็นทางคณิตศาสตร์ของปัญหาในชีวิตจริง - ทำสถานการณ์หรือปัญหาให้อยู่ในรูปอย่างง่าย เพื่อให้การวิเคราะห์ทางคณิตศาสตร์ง่ายขึ้น - แปลงปัญหาให้อยู่ในรูปของคณิตศาสตร์
การใช้หลักการ/ กระบวนการ คณิตศาสตร์ในการ แก้ปัญหา	- คิดและนำกลยุทธ์ในการหาวิธีการแก้ปัญหาทางคณิตศาสตร์ไปใช้ - ใช้เครื่องมือทางคณิตศาสตร์ เพื่อช่วยหาวิธีการแก้ปัญหาที่ถูกต้องหรือเหมาะสม - นำกฎเกณฑ์ ขั้นตอนวิธี และโครงสร้างทางคณิตศาสตร์ไปใช้แก้ปัญหา
การตีความ และ ประเมินผลลัพธ์ทาง คณิตศาสตร์	- ตีความผลลัพธ์ทางคณิตศาสตร์กลับไปสู่บริบทชีวิตจริง - ประเมินความเป็นเหตุเป็นผลของวิธีการแก้ปัญหาทางคณิตศาสตร์ในบริบทของปัญหาชีวิตจริง - อธิบายความสมเหตุสมผลของผลลัพธ์หรือข้อสรุปทางคณิตศาสตร์กับบริบทของปัญหา

สำหรับสัดส่วนของข้อสอบคณิตศาสตร์ของ PISA 2015 จะวัดกระบวนการการคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ ประมาณ 25% วัดการใช้หลักการหรือกระบวนการทางคณิตศาสตร์ในการแก้ปัญหาประมาณ 50% และวัดการตีความและประเมินผลลัพธ์ทางคณิตศาสตร์ ประมาณ 25% (สุชาติ บัณฑิตวิภาต. 2557: 35 - 37)

กระบวนการทางคณิตศาสตร์ตามคำนิยามของการรู้เรื่องคณิตศาสตร์ คือ ความสามารถของแต่ละบุคคลในการคิด การใช้ และการตีความคณิตศาสตร์ สามคำนี้มีประโยชน์ และมีความสำคัญต่อการจัดการกระบวนการทางคณิตศาสตร์ ซึ่งสามารถอธิบายได้ว่า แต่ละคน สามารถเชื่อมโยงบริบทของปัญหากับคณิตศาสตร์ และแก้ปัญหาได้อย่างไร สำหรับ PISA 2012 กระบวนการทางคณิตศาสตร์แบ่งออกเป็น 3 กระบวนการ ดังนี้ คือ การคิดสถานการณ์ของปัญหา ในเชิงคณิตศาสตร์ การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผลทางคณิตศาสตร์ และการตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์ (สถาบันส่งเสริมการสอน วิทยาศาสตร์และเทคโนโลยี. 2557: 3)

การคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์

นิยามของคำว่า “การคิด” ในการรู้เรื่องคณิตศาสตร์ หมายถึง ความสามารถของแต่ละบุคคลในการรู้และบอกโอกาสในการใช้คณิตศาสตร์ แล้วกำหนดโครงสร้างทางคณิตศาสตร์ ให้กับปัญหาที่พบในสถานการณ์สำหรับกระบวนการของ การคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ คือ การที่บุคคลตัดสินใจได้ว่าส่วนใดที่เขาสามารถดึงคณิตศาสตร์ที่จำเป็นไปใช้ในการวิเคราะห์ สร้างแนวทาง และนำไปแก้ปัญหาโดยบุคคลเหล่านี้สามารถแปลงปัญหาจากสถานการณ์ในชีวิตจริงให้อยู่ในขอบเขตคณิตศาสตร์ และกำหนดโครงสร้างทางคณิตศาสตร์ การแสดงเครื่องหมายแทน และลักษณะจำเพาะให้กับปัญหาในโลกชีวิตจริงซึ่งสามารถให้เหตุผล ตั้งสมมติฐาน และพิจารณาข้อจำกัดได้อย่างสมเหตุสมผล กระบวนการนี้ ประกอบด้วยกิจกรรมต่าง ๆ ดังต่อไปนี้ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2557: 4-5)

การระบุประเด็นทางคณิตศาสตร์ของปัญหาที่ตั้งอยู่ในบริบทโลกชีวิตจริง และการระบุตัวแปรที่สำคัญ

การรู้โครงสร้างทางคณิตศาสตร์ (รวมถึง กฎเกณฑ์ ความสัมพันธ์ และแบบรูป) ของปัญหาหรือสถานการณ์

การทำสถานการณ์หรือปัญหาให้อยู่ในรูปอย่างง่าย เพื่อให้การวิเคราะห์ทางคณิตศาสตร์ง่ายขึ้น

การระบุข้อจำกัดและสมมติฐานที่อยู่เบื้องหลังแบบจำลองทางคณิตศาสตร์และจากการทำให้อยู่ในรูปอย่างง่ายที่รวบรวมได้จากบริบท

การนำเสนอสถานการณ์ในเชิงคณิตศาสตร์ โดยใช้ตัวแปร สัญลักษณ์ แผนภาพ และแบบจำลองมาตรฐานที่เหมาะสม

การนำเสนอปัญหาในหลากหลายวิธี รวมถึงการจัดการกับปัญหาให้สอดคล้องกับแนวคิดทางคณิตศาสตร์และการสร้างสมมติฐานที่เหมาะสม

การรู้เข้าใจและการอธิบายความสัมพันธ์ระหว่างภาษาที่เฉพาะกับบริบทของปัญหากับภาษาที่เป็นสัญลักษณ์และภาษาอย่างเป็นทางการที่จำเป็นต้องใช้ในการแสดงเชิงคณิตศาสตร์

การแปลงปัญหาให้อยู่ในรูปของภาษาทางคณิตศาสตร์หรือใช้การแทน

การรู้แง่มุมต่าง ๆ ของปัญหา ซึ่งสอดคล้องกับปัญหาที่รู้หรือแนวคิดหลักทางคณิตศาสตร์ที่รู้จักข้อเท็จจริง หรือวิธีดำเนินการ

การใช้เทคโนโลยีเพื่อแสดงความสัมพันธ์ภายในปัญหาที่อยู่ในสถานการณ์เชิงคณิตศาสตร์ (เช่น ตารางโปรแกรมทำงาน หรือรายการที่มีให้บนเครื่องคำนวณเชิงกราฟ)

ตัวอย่างข้อสอบจาก PISA 2012 การคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์

(สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.2557: 56 - 58)

การปีนภูเขาฟูจิ

ภูเขาฟูจิ เป็นภูเขาไฟสงบแล้วที่มีชื่อเสียงโด่งดังในญี่ปุ่น



คำถามที่ 1: การปีนภูเขาฟูจิ

ภูเขาฟูจิเปิดให้คนขึ้นไปปีนเขาเฉพาะช่วงตั้งแต่ 1 กรกฎาคม ถึง 27 สิงหาคมของแต่ละปีในช่วงเวลาดังกล่าวมีคนขึ้นไปปีนภูเขาฟูจิประมาณ 200,000 คน โดยเฉลี่ยมีคนขึ้นไปปีนภูเขาฟูจิวันละประมาณกี่คน

1. 340

คะแนนเต็ม

2. 710

ข้อ 3. 3,400

3. 3,400

ไม่ได้คะแนน

4. 7,100

คำตอบอื่น ๆ

5. 7,400

ลักษณะเฉพาะของข้อสอบ

เนื้อหาทางคณิตศาสตร์: ปริมาณ

บริบท : สังคม

กระบวนการ : คิดในเชิงคณิตศาสตร์

รูปแบบของข้อสอบ : เลือกตอบ

เจตนาของคำถาม : ระบุอัตราเฉลี่ยรายวันจาก

จำนวนรวมและช่วงเวลาเฉพาะที่กำหนดให้

คำถามที่ 2: การปีนภูเขาฟูจิ

เส้นทางเดินโกเทมบะเป็นทางเดินขึ้นภูเขาฟูจิ มีระยะทางประมาณ 9 กิโลเมตร (กม.) การเดินขึ้นภูเขาไปและกลับคิดเป็นระยะทาง 18 กม. โดยต้องกลับมาถึงก่อนเวลา 20.00 น. โตชิคาว่า เขาสามารถเดินขึ้นภูเขาได้ในอัตราเร็วเฉลี่ย 1.5 กิโลเมตรต่อชั่วโมง และเดินลงด้วยอัตราเร็วเป็นสองเท่า ที่อัตราเร็วนี้รวมเวลาพักทานอาหารและหยุดพักแล้ว

จงใช้อัตราเร็วเฉลี่ยที่โตชิประมาณไว้ เพื่อหาว่าเขาควรเริ่มออกเดินอย่างช้าที่สุดในเวลาเท่าใดเพื่อจะกลับมาให้ทันเวลา 20.00 น.

.....

.....

.....

.....

การให้คะแนน การปีนภูเขาฟูจิ

คะแนนเต็ม

รหัส 1: 11 (โมงเช้า)
[ใส่หรือไม่ใส่โมงเช้าก็ได้ หรือ
การเขียนเวลาด้วยวิธีอื่นที่มี
ความหมายเหมือนกัน
ตัวอย่างเช่น 11.00 น.]

ไม่ได้คะแนน

รหัส 0: คำตอบอื่น ๆ

ลักษณะเฉพาะของข้อสอบ

เนื้อหาทางคณิตศาสตร์: การเปลี่ยนแปลงความสัมพันธ์

บริบท : สังคม

กระบวนการ : คิดในเชิงคณิตศาสตร์

รูปแบบของข้อสอบ : สร้างคำตอบแบบอิสระ

เจตนาของคำถาม : คำนวณเวลาเริ่มต้นในการ

เดินทาง เมื่อกำหนดอัตราเร็วที่แตกต่างกันสองอัตรา

ระยะทางที่เดินทางทั้งหมดและเวลาสิ้นสุด

การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผลทางคณิตศาสตร์

นิยามของคำว่า “การใช้” ในการรู้เรื่องคณิตศาสตร์ หมายถึง ความสามารถของแต่ละบุคคลในการประยุกต์ใช้แนวคิดหลักทางคณิตศาสตร์ ข้อเท็จจริง วิธีดำเนินการ และเหตุผลทางคณิตศาสตร์ในการแก้ปัญหาเชิงคณิตศาสตร์เพื่อให้ได้ข้อสรุปทางคณิตศาสตร์ ในกระบวนการการใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการและการให้เหตุผลทางคณิตศาสตร์ เพื่อแก้ปัญหาซึ่งเป็นกระบวนการที่แต่ละคนแสดงวิธีดำเนินการทางคณิตศาสตร์ที่จำเป็นเพื่อให้ได้มาซึ่งผลลัพธ์ และค้นหาวิธีแก้ปัญหาทางคณิตศาสตร์ (เช่น แสดงการคำนวณเลขคณิต การแก้สมการ การอนุมานเชิงตรรกศาสตร์จากสมมติฐานทางคณิตศาสตร์ การใช้เชิงสัญลักษณ์การสกัดข้อมูลทางคณิตศาสตร์จากตารางและกราฟ การแสดงแทนและการจัดการกับรูปร่างและรูปทรงและการวิเคราะห์ข้อมูล) นักเรียนพยายามสร้างแบบจำลองของสถานการณ์ปัญหา สร้างกฎเกณฑ์ ระบุความเชื่อมโยงระหว่างองค์ความรู้ทางคณิตศาสตร์ และสร้างข้อโต้แย้งทางคณิตศาสตร์ กระบวนการนี้ประกอบด้วยกิจกรรมต่าง ๆ ดังต่อไปนี้ (ตัวอย่างข้อสอบคณิตศาสตร์ PISA 2012 2557: 5)

การคิดและนำกลยุทธ์ในการหาวิธีแก้ปัญหาทางคณิตศาสตร์ไปใช้

การใช้เครื่องมือทางคณิตศาสตร์ รวมทั้งเทคโนโลยีเพื่อช่วยหาวิธีแก้ปัญหาที่ถูกต้องหรือเหมาะสม

การนำข้อเท็จจริง กฎเกณฑ์ ขั้นตอนวิธี และโครงสร้างทางคณิตศาสตร์มาใช้ในการแก้ปัญหา

การจัดการด้านจำนวน ข้อมูลและข้อสนเทศเกี่ยวกับกราฟและสถิติ นิพจน์พีชคณิต และสมการ และการแสดงแทนทางเรขาคณิต

การเขียนแผนภาพ กราฟ และโครงสร้างทางคณิตศาสตร์และการสกัดข้อมูลทางคณิตศาสตร์จากสิ่งเหล่านั้น

การใช้และการสลับที่ระหว่างการแสดงแทนต่าง ๆ ในกระบวนการแก้ปัญหา

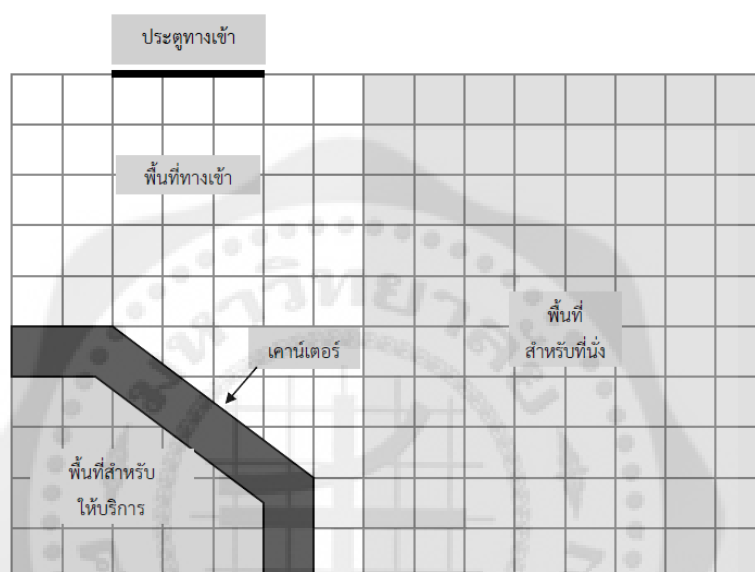
การสร้างข้อสรุปทั่วไปบนพื้นฐานของผลลัพธ์ที่เกิดจากการนำวิธีการทางคณิตศาสตร์ไปใช้ในการแก้ปัญหา

การสะท้อนข้อโต้แย้งทางคณิตศาสตร์ การอธิบายและการแสดงเหตุผลต่อผลลัพธ์ทางคณิตศาสตร์

ตัวอย่างข้อสอบจาก PISA 2012 การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผลทางคณิตศาสตร์ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.2557: 24 - 26)

ร้านไอศกรีม

แบบแปลนพื้นร้านไอศกรีมของมาลีเป็นดังนี้ เธอกำลังจะปรับปรุงร้านพื้นที่สำหรับให้บริการถูกล้อมด้วยเคาน์เตอร์



หมายเหตุ : รูปสี่เหลี่ยมจัตุรัสแต่ละรูปในช่องตาราง แทน 0.5 เมตร x 0.5 เมตร

คำถามที่ 1: ร้านไอศกรีม

มาลีต้องการทำขอบเคาน์เตอร์ใหม่ตามแนวขอบด้านนอก ความยาวขอบทั้งหมดที่เธอต้องทำเป็นเท่าใด จงแสดงวิธีทำ

.....

.....

.....

.....

.....

.....

.....

การให้คะแนน ร้านไอศกรีม

คะแนนเต็ม

รหัส 2 : คำตอบอยู่ในช่วงตั้งแต่ 4.45 ถึง 4.55 (ตอบในหน่วยเมตร ใส่หรือไม่ใส่หน่วยก็ได้)

ได้คะแนนบางส่วน

รหัส 1 : คำตอบที่แสดงให้เห็นว่าวิธีทำบางส่วนถูกต้อง (เช่น ใช้ทฤษฎีบทพีทาโกรัส หรือการอ่านมาตราส่วน) แต่มีข้อผิดพลาด เช่น ใช้มาตราส่วนไม่ถูกต้อง หรือมีการคำนวณผิด

- ตั้งแต่ 8.9 ถึง 9.1 ม. [ไม่ได้ใช้มาตราส่วน]
- 2.5 ม. (หรือ 5 หน่วย) [ใช้พีทาโกรัสคำนวณหาด้านตรงข้ามมุมฉากได้ 5 หน่วย (2.5 เมตร) แต่ไม่ได้ รวมด้านประกอบมุมฉากอีกสองด้าน]

ไม่ได้คะแนน

รหัส 0 : คำตอบอื่น ๆ

ลักษณะเฉพาะของข้อสอบ

เนื้อหาทางคณิตศาสตร์: ปริภูมิและรูปทรง

บริบท : การงานอาชีพ

กระบวนการ : ใช้หลักการทางคณิตศาสตร์

รูปแบบของข้อสอบ : สร้างคำตอบแบบอิสระ

เจตนาของคำถาม : ใช้ทฤษฎีบทพีทาโกรัสหรือใช้การวัดเพื่อหาความยาวของด้านตรงข้ามมุมฉากและแปลงหน่วยตามมาตราส่วนในภาพ

คำถามที่ 2: ร้านไอศกรีม

มาลีต้องการปูพื้นใหม่ในร้านด้วย พื้นที่ว่างของพื้นที่ทั้งหมดในร้านเป็นเท่าใด หากไม่นับรวมพื้นที่ส่วนให้บริการและเคาน์เตอร์ จงแสดงวิธีทำ

.....

.....

.....

.....

การให้คะแนน ร้านไอศกรีม

คะแนนเต็ม

รหัส 2 : 31.5 [ใส่หรือไม่ใส่หน่วยก็ได้]

ได้คะแนนบางส่วน

รหัส 1 : แสดงให้เห็นอย่างชัดเจนว่ามีการใช้ตารางในการคำนวณพื้นที่ แต่ใช้มาตราส่วนไม่ถูกต้องหรือมีข้อผิดพลาดเกี่ยวกับเลขคณิต

- 126 [คำตอบแสดงการคำนวณพื้นที่ถูกต้อง แต่ไม่ได้ใช้มาตราส่วนเพื่อให้ได้ค่าที่แท้จริง]
- $7.5 \times 5 (=37.5) - 3 \times 2.5 (=7.5) - \frac{1}{2} \times 2 \times 1.5 (=1.5) = 28.5$ ม.2 [ใช้การลบออก แทนที่จะใช้การบวกพื้นที่รูปสามเหลี่ยม เมื่อแบ่งพื้นที่ทั้งหมดเป็นพื้นที่ย่อยๆ]

ไม่ได้คะแนน

รหัส 0 : คำตอบอื่น ๆ

ลักษณะเฉพาะของข้อสอบ

เนื้อหาทางคณิตศาสตร์: ปริภูมิและรูปทรง

บริบท : การงานอาชีพ

กระบวนการ : ใช้หลักการทางคณิตศาสตร์

รูปแบบของข้อสอบ : สร้างคำตอบแบบอิสระ

เจตนาของคำถาม : คำนวณพื้นที่ของรูปหลายเหลี่ยม

การตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์

นิยามของคำว่า “ตีความ” ในการรู้เรื่องคณิตศาสตร์ มุ่งเน้นไปที่ความสามารถของแต่ละบุคคลในการสะท้อนวิธีแก้ปัญหา ผลลัพธ์ หรือข้อสรุปทางคณิตศาสตร์แล้วตีความออกมาในบริบทของปัญหาโลกชีวิตจริงซึ่งรวมถึงการแปลความหมายของวิธีแก้ปัญหาหรือการให้เหตุผลทางคณิตศาสตร์ย้อนกลับไปที่บริบทของปัญหา และตัดสินใจว่าผลลัพธ์ที่ได้เป็นเหตุเป็นผลและเข้ากันได้กับบริบทของปัญหาหรือไม่ บุคคลที่ใช้กระบวนการนี้อาจสร้างและสื่อสารคำอธิบายหรือข้อโต้แย้งในบริบทของปัญหา และการสะท้อนทั้งกระบวนการสร้างแบบจำลองและผลที่ได้ กระบวนการประเภทนี้รวม “ตีความ” และ “ประเมิน” ไว้ด้วยกัน กระบวนการนี้ประกอบด้วยกิจกรรมต่าง ๆ ดังต่อไปนี้ (โครงการ PISA ประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2557: 6)

การตีความผลลัพธ์ทางคณิตศาสตร์กลับไปบริบทโลกชีวิตจริง

การประเมินความเป็นเหตุเป็นผลของวิธีแก้ปัญหาทางคณิตศาสตร์ในบริบทของปัญหาโลกชีวิตจริง

ความเข้าใจว่าสถานการณ์ในชีวิตจริงส่งผลกระทบต่อผลลัพธ์และการคิดคำนวณตามกระบวนการหรือตัวแบบทางคณิตศาสตร์อย่างไร เพื่อตัดสินใจว่าจะปรับปรุงหรือนำผลไปใช้อย่างไร

อธิบายได้ว่าเพราะเหตุใดผลลัพธ์หรือข้อสรุปทางคณิตศาสตร์จึงเหมาะสมหรือไม่เหมาะสมกับบริบทของปัญหา

ความเข้าใจขอบเขตและข้อจำกัดของแนวคิดคณิตศาสตร์และวิธีแก้ปัญหาทางคณิตศาสตร์ และ

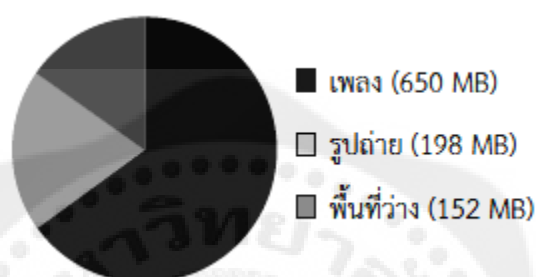
วิจารณ์และระบุข้อจำกัดของแบบจำลองที่ใช้แก้ปัญหา

ตัวอย่างข้อสอบจาก PISA 2012 การตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2557: 15 - 17)

แฟลชไดรฟ์

แฟลชไดรฟ์ เป็นอุปกรณ์จัดเก็บข้อมูลคอมพิวเตอร์แบบพกพาที่มีขนาดเล็ก
อนันต์มีแฟลชไดรฟ์อันหนึ่งซึ่งเก็บเพลงและรูปถ่ายไว้ แฟลชไดรฟ์มีความจุ 1 GB (1,000 MB)
กราฟข้างล่างแสดงสถานะดิสก์ปัจจุบันของแฟลชไดรฟ์ของอนันต์

สถานะดิสก์ของแฟลชไดรฟ์



คำถามที่ 1: แฟลชไดรฟ์

อนันต์ต้องการย้ายอัลบั้มรูปถ่ายขนาด 350 MB ลงในแฟลชไดรฟ์ของเขา แต่พื้นที่ว่างในแฟลชไดรฟ์มีไม่เพียงพอ ซึ่งเขาไม่ต้องการลบรูปถ่ายใด ๆ ที่มีอยู่ออก แต่เขายินดีที่จะลบอัลบั้มเพลงสองอัลบั้มออกแฟลชไดรฟ์ของอนันต์ เก็บอัลบั้มเพลงขนาดต่าง ๆ ต่อไปนี้

อัลบั้ม	ขนาด
อัลบั้ม 1	100 MB
อัลบั้ม 2	75 MB
อัลบั้ม 3	80 MB
อัลบั้ม 4	55 MB
อัลบั้ม 5	60 MB
อัลบั้ม 6	80 MB
อัลบั้ม 7	75 MB
อัลบั้ม 8	125 MB

ถ้าลบอัลบั้มเพลงอย่างมากที่สุดสองอัลบั้มจะทำให้แฟลชไดรฟ์ของอนันต์มีพื้นที่ว่างเพียงพอที่จะเพิ่มอัลบั้มรูปถ่ายใช่หรือไม่ จงเขียนวงกลมล้อมรอบคำว่า “ใช่” หรือ “ไม่ใช่” และแสดงวิธีการคำนวณเพื่อสนับสนุนคำตอบของนักเรียน

คำตอบ : ใช่ / ไม่ใช่

.....
.....
.....

การให้คะแนน แฟลชไดรฟ์

คะแนนเต็ม

- รหัส 1 : ใช่ (โดยตรงหรือโดยนัย) และ ระบุชื่อสองอัลบั้ม (หรือขนาด) ซึ่งใช้พื้นที่ 198 MB หรือมากกว่า
- เขาต้องลบออก 198 MB (350 - 152) ดังนั้น เขาอาจลบอัลบั้มเพลงใด ๆ สองอัลบั้ม ที่รวมกันแล้วมากกว่า 198 MB ตัวอย่างเช่น อัลบั้ม 1 และ อัลบั้ม 8
 - ใช่ เขาอาจจะลบอัลบั้ม 7 และ อัลบั้ม 8 ซึ่งทำให้ได้พื้นที่ว่าง $152 + 75 + 125 = 352$ MB

ไม่ได้คะแนน

รหัส 0 : คำตอบอื่น ๆ

ลักษณะเฉพาะของข้อสอบ

เนื้อหาทางคณิตศาสตร์: ปริมาณ

บริบท : ส่วนตัว

กระบวนการ : ดีความ

รูปแบบของข้อสอบ : สร้างคำตอบแบบอิสระ

เจตนาของคำถาม : คำนวณและเปรียบเทียบค่าที่ได้ให้เป็นไปตามเกณฑ์ที่กำหนด

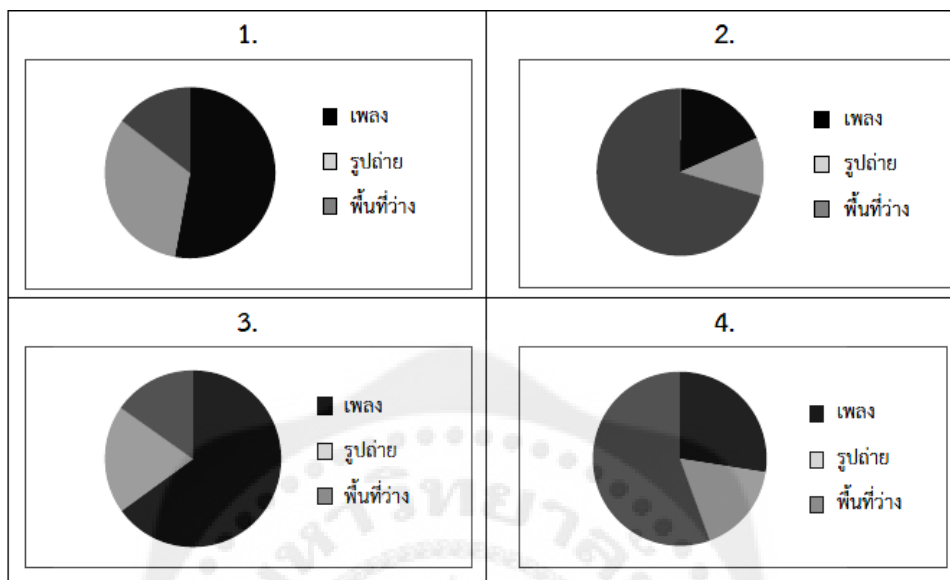
คำถามที่ 2: แฟลชไดรฟ์

ในช่วงหลายสัปดาห์ต่อมา อนันต์ลบรูปถ่ายและเพลงบางส่วนออก แต่ได้เพิ่มไฟล์รูปถ่ายและเพลงใหม่เข้าไปด้วย สถานะดิสก์ใหม่แสดงในตารางข้างล่าง:

เพลง	550 MB
รูปถ่าย	338 MB
พื้นที่ว่าง	112 MB

พี่ชายของอนันต์ให้แฟลชไดรฟ์อันใหม่กับเขา ซึ่งมีความจุ 2 GB (2,000 MB) ซึ่งเป็นพื้นที่ว่างทั้งหมด อนันต์จึงย้ายสิ่งที่เก็บไว้ในแฟลชไดรฟ์อันเก่าลงในอันใหม่

กราฟต่อไปนี้ แสดงสถานะติสก์ของแฟลชไดรฟ์อันใหม่
จงเขียนวงกลมล้อมรอบ 1, 2, 3 หรือ 4



การให้คะแนน แฟลชไดรฟ์

คะแนนเต็ม

ข้อ 4. [รูป 4]

ไม่ได้คะแนน

คำตอบอื่น ๆ

ลักษณะเฉพาะของข้อสอบ

เนื้อหาทางคณิตศาสตร์ : ความไม่แน่นอนและข้อมูล

บริบท : ส่วนตัว

กระบวนการ : ตีความ

รูปแบบของข้อสอบ : เลือกตอบ

เจตนาของคำถาม : เข้าใจความสัมพันธ์ระหว่าง

ภาษาของโจทย์และสัญลักษณ์ กับภาษาที่เป็นทางการที่จำเป็นต้องใช้ในการนำเสนอทางคณิตศาสตร์

การที่จะทราบได้ว่านักเรียนสามารถเข้าไปมีส่วนร่วมในแต่ละกระบวนการเหล่านี้ได้อย่างมีประสิทธิภาพเพียงใดนั้นเป็นสิ่งสำคัญต่อการจัดทำนโยบายทางการศึกษาในปัจจุบัน ในกระบวนการการคิดในเชิงคณิตศาสตร์ ซึ่งให้เห็นว่า นักเรียนสามารถรู้และบอกโอกาสที่จะใช้คณิตศาสตร์ในสถานการณ์ของปัญหา แล้วให้โครงสร้างทางคณิตศาสตร์ที่จำเป็นต้องใช้ในการแปลงสถานการณ์ของปัญหาให้อยู่ในรูปทางคณิตศาสตร์อย่างมีประสิทธิภาพ ส่วนกระบวนการใช้หลักคณิตศาสตร์ ซึ่งให้เห็นว่านักเรียนสามารถลงมือคำนวณ ดำเนินการ และประยุกต์แนวคิดหลัก และข้อเท็จจริงที่นำไปสู่การแก้ปัญหาเชิงคณิตศาสตร์กับปัญหาที่ถูกเปลี่ยนให้เป็นปัญหาเชิงคณิตศาสตร์ได้ดีเพียงใด และสำหรับกระบวนการตีความ ซึ่งให้เห็นว่านักเรียนสามารถสะท้อนข้อสรุปและวิธีการแก้ปัญหาทางคณิตศาสตร์ ดีความผลที่ได้ไปสู่อุปสรรคปัญหาในโลกชีวิตจริง และระบุได้ว่าผลลัพธ์หรือข้อสรุปเป็นเหตุเป็นผลหรือไม่ การที่นักเรียนจะนำคณิตศาสตร์มาใช้กับปัญหาหรือสถานการณ์ขึ้นอยู่กับทักษะที่อยู่ภายในสามกระบวนการนี้ และความรู้ความเข้าใจเกี่ยวกับประสิทธิภาพของแต่ละกระบวนการจะช่วยให้สามารถตัดสินใจหรือวางแผนเกี่ยวกับการจัดการเรียนการสอนในห้องเรียนหรือวางนโยบายทางการศึกษาต่อไป

1.6 งานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ

จารุพันธ์ ขวัญแนน; เอี่ยมพร หลินเจริญ; และคณะ. (2558) ศึกษาการพัฒนา รูปแบบการประเมินการรู้คณิตศาสตร์ การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนารูปแบบการประเมิน การรู้คณิตศาสตร์ สำหรับนักเรียนระดับการศึกษาขั้นพื้นฐาน วิธีดำเนินการวิจัย แบ่งออกเป็น 4 ขั้นตอน คือ 1) สร้างองค์ประกอบและตัวบ่งชี้การประเมินการรู้คณิตศาสตร์ สำหรับนักเรียนระดับการศึกษาขั้นพื้นฐาน 2) สร้างและตรวจสอบคุณภาพของรูปแบบการประเมินการรู้คณิตศาสตร์ สำหรับนักเรียนระดับการศึกษา ขั้นพื้นฐาน 3) ทดลองใช้เครื่องมือประเมินการรู้คณิตศาสตร์สำหรับนักเรียนระดับการศึกษาขั้นพื้นฐาน 4) ประเมินรูปแบบการประเมินการรู้คณิตศาสตร์สำหรับนักเรียนระดับการศึกษาขั้นพื้นฐาน ผลการวิจัย พบว่า

1. องค์ประกอบและตัวบ่งชี้การประเมินการรู้คณิตศาสตร์ของนักเรียนระดับ การศึกษาขั้นพื้นฐาน ประกอบด้วย 4 องค์ประกอบหลัก 21 ตัวบ่งชี้ ได้แก่ องค์ประกอบที่ 1 การแก้ปัญหาคณิตศาสตร์ในสถานการณ์โลกจริงมี 6 ตัวบ่งชี้ องค์ประกอบที่ 2 ความรู้ในเนื้อหา ทางคณิตศาสตร์ มี 4 ตัวบ่งชี้ องค์ประกอบที่ 3 สมรรถนะทางคณิตศาสตร์ในการแก้ปัญหา มี 8 ตัวบ่งชี้ และองค์ประกอบที่ 4 กลุ่มของสมรรถนะคณิตศาสตร์ในการแก้ปัญหา มี 3 ตัวบ่งชี้

2. รูปแบบการประเมินการรู้คณิตศาสตร์ มีความเหมาะสมและมีความเป็นไปได้ใน การนำไปใช้อยู่ในระดับมากขึ้นไปทุกองค์ประกอบ รูปแบบมีลักษณะเป็นโครงสร้างที่สัมพันธ์กัน 8 องค์ประกอบ ประกอบด้วย 1) วัตถุประสงค์การประเมิน 2) สิ่งที่มีประเมิน 3) วิธีการประเมิน 4) ผู้มีส่วนเกี่ยวข้องกับการประเมิน 5) เวลา/สถานที่ประเมิน 6) การจัดเก็บข้อมูลการประเมิน 7) เกณฑ์การตัดสินผลการประเมินและ 8) การเขียนรายงานการประเมิน

3. ผลการทดลองใช้เครื่องมือในการประเมินการรู้คณิตศาสตร์ของนักเรียนที่สร้างขึ้นพบว่า เครื่องมือมีค่าความยากอยู่ระหว่าง 0.50-0.79 และค่าอำนาจจำแนกอยู่ระหว่าง 0.21-0.26 มีค่าความเชื่อมั่น 0.96 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันมีความตรงเชิงโครงสร้างตามองค์ประกอบและตัวบ่งชี้การประเมินการรู้คณิตศาสตร์ทุกองค์ประกอบ ค่าไค-สแควร์มีค่าเท่ากับ 183.23 และมีค่าความน่าจะเป็นที่เหตุการณ์จะเกิดขึ้น ถ้าสมมติฐานหลักเป็นจริงเท่ากับ 0.05 ($p = 0.05$) ที่องศาอิสระเท่ากับ 154 ($df = 154$) ค่าดัชนีวัดระดับความกลมกลืน มีค่าเท่ากับ 0.95 ($GFI = 0.95$) และดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว มีค่าเท่ากับ 0.92 ($AGFI = 0.92$) ค่าดัชนีรากกำลังสองของเศษเหลือมีค่าเท่ากับ 0.01 ($RMR = 0.01$) และค่าดัชนีกำลังสองของความคลาดเคลื่อน เท่ากับ 0.02 ($RMSEA=0.02$)

4. รูปแบบการประเมินการรู้คณิตศาสตร์ สำหรับนักเรียนระดับการศึกษาขั้นพื้นฐานพบว่า มีความเป็นประโยชน์ ความเป็นไปได้ ความเหมาะสมและความถูกต้อง ต่อการนำไปใช้ในการประเมินการรู้คณิตศาสตร์ สำหรับนักเรียนระดับการศึกษาขั้นพื้นฐานในระดับมากที่สุด

ฉวีวรรณ แก้วไทรชะ; และสุพจน์ ไชยสังข์. (2557) ศึกษาการวิเคราะห์การสอบพิชชาและโอเน็ตของสถาบันทดสอบการศึกษาแห่งชาติ เพื่อูปการเรียนการสอนวิชาคณิตศาสตร์ การวิจัยครั้งนี้เป็นการวิจัยเชิงสำรวจเพื่อวิเคราะห์ความสัมพันธ์ระหว่างข้อสอบ O-NET และข้อสอบ PISA ซึ่งข้อสอบ O-NET เป็นข้อสอบประเมินตามมาตรฐานของผลลัพธ์การเรียนรู้ที่กำหนดในหลักสูตร แต่ข้อสอบ PISA เป็นข้อสอบประเมินการรู้เรื่องคณิตศาสตร์ โดยกำหนดเป็นนิยามว่า “ความสามารถของแต่ละบุคคลและการเข้าใจบทบาทของคณิตศาสตร์ในโลก การตัดสินใจที่ดีและการใช้คณิตศาสตร์ตามความต้องการของชีวิตของแต่ละบุคคล ในฐานะพลเมืองที่มีความรู้มีความตระหนักและมีการสะท้อนสิ่งที่เป็ประโยชน์” (Organization for Economic Co-operation and Development [OECD], 2003) จากการสอบถามครูคณิตศาสตร์ พบว่าข้อสอบที่ใช้สถานการณ์หรือบริบทนักเรียนใช้เวลาในการอ่านและทำความเข้าใจโจทย์ปัญหา มากกว่าโจทย์ปัญหาที่มีบริบทภายในคณิตศาสตร์จากผลการวิเคราะห์ความสัมพันธ์ระหว่างสถานการณ์และบริบทของโจทย์ปัญหาต่อไปนี้ ผู้วิจัยได้นำบริบทที่กำหนดไว้ในกรอบความคิดของข้อสอบ PISA เป็นเกณฑ์การเปรียบเทียบ 4 กลุ่มใหญ่ ได้แก่ (1) บริบทส่วนตัว เน้นคณิตศาสตร์ที่เกี่ยวข้องกับกิจกรรมของบุคคลวันต่อวัน (2) บริบทการงานอาชีพ เน้นคณิตศาสตร์ที่เกี่ยวข้องกับการทำงานในชีวิตจริงที่ไม่ต้องใช้ทักษะจนถึงงานที่ต้องใช้ความเชี่ยวชาญระดับสูง (3) บริบททางสังคม เน้นคณิตศาสตร์ที่เกี่ยวข้องกับกิจกรรมที่เกิดในชุมชนระดับท้องถิ่น ระดับชาติหรือระดับโลก และ(4) บริบททางวิทยาศาสตร์ เน้นคณิตศาสตร์ที่เกี่ยวข้องกับกิจกรรมทางวิทยาศาสตร์ เทคโนโลยี และภายในคณิตศาสตร์ จากการวิเคราะห์ข้อมูลพบว่า

(1) ด้านพฤติกรรมการรู้เรื่องคณิตศาสตร์ พบว่า ครูคณิตศาสตร์จากโรงเรียนแบบฉบับมีความต้องการให้นักเรียนสามารถปฏิบัติได้ทุกพฤติกรรมในระดับมาก (3.51 ถึง 4.51)

(2) ด้านปัจจัยที่สำคัญที่มีต่อความสำเร็จของการจัดการเรียนการสอน พบว่า

(2.1) ปัจจัยที่เกี่ยวข้องกับนักเรียน ได้แก่ ปัจจัยด้านพฤติกรรมผู้เรียน เช่น การมีระเบียบวินัยในการเรียน การให้ความร่วมมือในกิจกรรมทั้งในและนอกห้องเรียน การมีมนุษยสัมพันธ์ที่ดีกับครูและเพื่อนนักเรียน การเรียนพิเศษนอกโรงเรียน

(2.2) ปัจจัยที่เกี่ยวข้องกับครู ได้แก่ วุฒิการศึกษา ซึ่งทุกโรงเรียนมีครูสำเร็จการศึกษาด้านคณิตศาสตร์ที่สามารถพัฒนานักเรียนได้ แต่มีจำนวนที่แตกต่างกันระหว่างวุฒิปริญญาโทและวุฒิปริญญาตรีบางโรงเรียนมีวุฒิปริญญาโทมากกว่าบางโรงเรียนมีจำนวนพอๆกัน บางโรงเรียนมีแต่ปริญญาตรีอย่างเดียว นอกจากนั้นโรงเรียนแบบฉบับที่มีครูสำเร็จการศึกษาด้านคณิตศาสตร์จะมีความมั่นใจในการจัดการเรียนการสอนที่พัฒนาทักษะและกระบวนการทางคณิตศาสตร์ได้อย่างถูกต้องมีประสิทธิภาพ

สำหรับการวิจัยครั้งนี้ได้กำหนดเนื้อหาสาระคณิตศาสตร์ของ O-NET และ PISA ให้เทียบเคียงกัน เพื่อให้สามารถเสนอแนวคิดที่เป็นประโยชน์ต่อการปฏิรูปการเรียนการสอนคณิตศาสตร์ในประเทศไทยได้สาระที่วิเคราะห์ ได้แก่ จำนวนและการดำเนินการ เรขาคณิตและการวัด พีชคณิต การวิเคราะห์ข้อมูลและความน่าจะเป็นและ สาระอื่น ๆ ที่ไม่ได้ระบุไว้ในหลักสูตรการศึกษาขั้นพื้นฐานของประเทศไทยสำหรับนักเรียนอายุ 15 ปี ซึ่งกำลังเรียนอยู่ในช่วงชั้นที่ 3 (มัธยมศึกษาตอนต้น) การพัฒนาข้อสอบเทียบเคียงกับข้อสอบ PISA Chang, G.C. (2013, p. 25 – 26) กล่าวว่านักวิชาการในประเทศไทยยังมีข้อถกเถียงในเรื่องการสอบ O-NET ว่า จะเน้นเฉพาะการวัดความรู้เชิงวิชาการ (academic knowledge) อย่างเดียว หรือจะวัดทักษะการคิดระดับสูงด้วย (higher-order skills) เช่น ความสามารถในการให้เหตุผล (the ability to reason) การวิเคราะห์และสังเคราะห์สารสนเทศ (analyze and synthesize information) ซึ่งขณะนี้การประเมินระดับชาติ (และนานาชาติ) ได้ใช้งานวิจัยที่สัมพันธ์กับการพัฒนานักเรียนเกี่ยวกับทักษะการรู้หนังสือและทักษะเชิงตัวเลข (literacy and numeracy skills) อย่างไรก็ตามการพัฒนาข้อสอบที่เทียบเคียงกับ PISA โดยไม่ส่งผลกระทบต่อการประเมินระดับชาติ จะช่วยให้เกิดประโยชน์ต่อการพัฒนาการเรียนการสอนคณิตศาสตร์ โดยใช้กรอบมาตรฐานของ PISA ซึ่งปัจจุบันข้อสอบนี้ได้จัดทำเป็นระบบออนไลน์ และเปิดโอกาสให้นักเรียนสามารถเข้าทดสอบได้ จากเว็บไซต์ของ CTB/McGraw-Hill เนื่องจากการประเมิน PISA ใช้ประเมินนักเรียนอายุ 15 ปีเป็นเกณฑ์ (อ้างถึง Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment, 2009)

อารียาดี วิจายา และคณะ (Ariyadi Wijaya; et al. 2014) ศึกษาเรื่องความยากของการแก้ปัญหาตามบริบทบนพื้นฐานการรู้เรื่องคณิตศาสตร์ โดยการวิเคราะห์ข้อผิดพลาดของนักเรียน จุดมุ่งหมายการศึกษาการวิจัยครั้งนี้ เพื่อศึกษาความยากของการแก้ปัญหาตามบริบทบนพื้นฐานการรู้เรื่องคณิตศาสตร์ กลุ่มตัวอย่างในการวิจัยครั้งนี้เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 9 และปีที่ 10 จำนวน 362 คน ได้คัดเลือกจากโรงเรียน 11 แห่งที่ตั้งอยู่ในเขตชนบทและเมืองจาการ์ตา ประเทศอินโดนีเซีย โดยใช้โปรแกรมสำหรับการประเมินผลนักเรียนนานาชาติ (PISA) จำแนกนักเรียนเป็นสามกลุ่มสุรณะ ได้แก่ กลุ่มการทำใหม่ กลุ่มการเชื่อมโยง และกลุ่มการสะท้อน ตามกระบวนการรู้เรื่องคณิตศาสตร์ 1.การทำความเข้าใจ 2.การแปลงปัญหาให้อยู่ในรูปคณิตศาสตร์ 3.วิธีดำเนินการทางคณิตศาสตร์ และ 4.การตีความและการประยุกต์ใช้ในสถานการณ์จริง สถิติที่ใช้ในการวิเคราะห์ข้อมูล คือ ไคสแควร์ ผลการวิจัย พบว่าข้อผิดพลาดมากที่สุดในสองกระบวนการแรก คือ กระบวนการทำความเข้าใจ คิดเป็นร้อยละ 38 ความเข้าใจที่คลาดเคลื่อนข้างต้นส่งผลให้กระบวนการแปลงปัญหาให้อยู่ในรูปคณิตศาสตร์เพิ่มขึ้น คิดเป็นร้อยละ 42 ข้อผิดพลาดที่น้อย คือ กระบวนการวิธีดำเนินการทางคณิตศาสตร์ และกระบวนการตีความและการประยุกต์ใช้ในสถานการณ์จริง คิดเป็นร้อยละ 17 และร้อยละ 3 ของข้อผิดพลาดทั้งหมดตามลำดับ นอกจากนี้ยังพบนักเรียนที่มีผลสัมฤทธิ์ต่ำมีจำนวนมากกว่านักเรียนที่มีผลสัมฤทธิ์สูง การค้นพบนี้ชี้ให้เห็นว่านักเรียนที่มีผลสัมฤทธิ์ต่ำอาจติดอยู่ในขั้นเริ่มต้นของกระบวนการสร้างแบบจำลอง ไม่สามารถมาถึงขั้นตอนวิธีการดำเนินการทางคณิตศาสตร์เมื่อต้องแก้ปัญหาตามบริบท

ดันแคน มาร์กูเร และมาโมลาลูว่า อเมเลีย โมโคเเนา (Duncan Mhakure; & Mamolahluwa Amelia Mokoena. 2011) ศึกษาเปรียบเทียบหลักสูตรที่มีรูปแบบการสอน FET ตามหลักสูตรการรู้เรื่องคณิตศาสตร์กับหลักสูตรแกนกลางกลุ่มสาระคณิตศาสตร์ เป็นการเปรียบเทียบมุมมองแนวคิดของทั้งสองหลักสูตรของครูที่สอนรายวิชาคณิตศาสตร์ สังกัดหน่วยงานคณิตศาสตร์วิทยาศาสตร์และเทคโนโลยี ผลการวิจัยพบว่า ทั้งสองหลักสูตรมีความคล้ายกัน แต่มีมาตรฐานการประเมินที่แตกต่างกัน ซึ่งปัจจัยที่เป็นอุปสรรคต่อการจัดการเรียนการสอน เช่น การขาดครูผู้สอนรายวิชาคณิตศาสตร์ที่มีคุณสมบัติเหมาะสม การขาดสนับสนุนจากผู้ปกครอง ทศนคติเชิงลบที่มีต่อรายวิชาคณิตศาสตร์ และขาดการสนับสนุนจากหน่วยงานที่เกี่ยวข้อง วิธีการดำเนินการวิจัย ผู้วิจัยออกแบบการวิจัยทั้งเชิงปริมาณและเชิงคุณภาพ กลุ่มตัวอย่างเป็นครูผู้สอนรายวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษาตอนต้น จำนวน 842 คน จาก 160 โรงเรียน ในสี่เมืองของประเทศแอฟริกาใต้ เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบถามประกอบด้วยคำถามแบบปิดและแบบปลายเปิด คำถามที่ปิดให้การตอบตรงตามความเป็นจริง คำถามที่ถูกปิดนี้มีโครงสร้างแบบ Likert มีห้าระดับ "ไม่เห็นด้วยอย่างยิ่ง", "ไม่เห็นด้วย", "ค่อนข้างไม่เห็นด้วย", "ค่อนข้างเห็นด้วย", "เห็นด้วย" และ "เห็นด้วยอย่างยิ่ง" ด้านล่างแต่ละคำถามปลายเปิดมีช่องว่างสำหรับครูในการแสดงความรู้สึกในการเขียนเกี่ยวกับ

ปัญหาใด ๆ ที่เกิดจากคำถามเกี่ยวกับการรู้เรื่องคณิตศาสตร์กับหลักสูตรคณิตศาสตร์ สรุปผลการวิจัย พบว่า ทั้งสองหลักสูตรมีความคล้ายกัน โดย PISA (2003) การรู้เรื่องคณิตศาสตร์เป็นผลมาจากทั้งสองบริบทและเนื้อหา ในแง่ของความแตกต่างระหว่างสองหลักสูตรเนื้อหาหลักสูตรการรู้เรื่องคณิตศาสตร์ คือ การขับเคลื่อนด้วยบริบทที่ได้จากประสบการณ์ในชีวิตประจำวันของผู้เรียนในขณะที่ในทางตรงกันข้ามบริบทหลักสูตรคณิตศาสตร์ถูกขับเคลื่อนโดยเนื้อหาที่ได้จากการดำเนินงานทางคณิตศาสตร์ การศึกษาายังแสดงให้เห็นว่ามีปัญหาเกี่ยวกับการจัดการเรียนการสอนคณิตศาสตร์และการรู้เรื่องคณิตศาสตร์ในโรงเรียน ปัญหาบางส่วนค่อนข้างคุ้นเคยของระบบการศึกษาในประเทศแอฟริกาใต้ ตามที่ AMESA (2003) ระบุว่า ปัญหาการขาดคุณสมบัติและประสบการณ์ครูผู้สอนรายวิชาคณิตศาสตร์ ขาดเอกสารการเรียนการสอนที่เกี่ยวข้อง และขาดการพัฒนาวิชาชีพ ปัญหามุมมองของประสบการณ์ในการจัดการเรียนการสอนที่โรงเรียน ไม่มีการนำเทคโนโลยีมาใช้ประกอบการจัดการเรียนการสอนเพื่อพัฒนาตนเองเป็นครูมืออาชีพ

ฮัมซา เวนการ์ต และคณะ (Hamsa Venkat; et al. 2009) ศึกษาการประเมินการรู้เรื่องคณิตศาสตร์เกี่ยวกับหลักการใช้เหตุผลและการแก้ปัญหาโจทย์คณิตศาสตร์ การวิจัยครั้งนี้เป็นการพิจารณาการประเมินกระบวนการรู้เรื่องคณิตศาสตร์สำหรับหลักการใช้เหตุผลและการแก้ปัญหาโจทย์คณิตศาสตร์ ซึ่งถูกระบุไว้ว่าเป็นองค์ประกอบที่สำคัญของการรู้เรื่องคณิตศาสตร์ โดยดำเนินการวิเคราะห์ผ่านโครงสร้างภายในสมรรถนะการประเมินการรู้เรื่องคณิตศาสตร์ มีการนำเข้าข้อมูลผ่านข้อคำถามที่ถูกถามภายในกรอบการประเมินการรู้เรื่องคณิตศาสตร์ ข้อบกพร่องบางส่วนเหล่านี้มีตัวอย่างให้เห็นในผ่านคำถามที่ปรากฏในการสอบคณิตศาสตร์ 2008 ในรายงานการวิจัยครั้งนี้ได้มีการสรุปผลกระทบของโครงสร้างทฤษฎีนี้สำหรับการพัฒนาความสามารถหลักการใช้เหตุผลและการแก้ปัญหาโจทย์คณิตศาสตร์ ที่สอดคล้องกับจุดมุ่งหมายของหลักสูตร บทความนี้อ้างอิงการประเมินทฤษฎีในคำชี้แจงของหลักสูตรการรู้เรื่องคณิตศาสตร์ (Department of Education (DOE), 2007) ผลการวิจัย พบว่า จากการเผยแพร่หลักฐานการประเมินการขับเคลื่อนการสอน (Clarke, 1996) แสดงให้เห็นว่าการประเมินผลทฤษฎีการรู้เรื่องคณิตศาสตร์ถูกออกแบบมาเพื่อจัดการเรียนการสอน แสดงให้เห็นถึงความสำคัญของการแก้ปัญหาอย่างแท้จริง ซึ่งเอกสารที่เกี่ยวข้องกับหลักสูตรชี้ให้เห็นข้อด้อยของรูปแบบการประเมินการรู้เรื่องคณิตศาสตร์ที่เน้นการคำนวณในชีวิตประจำวัน (DOE. 2003: 9)

ฮุสไยอิน ฮัสนู และไกลเรย์ เบอร์เบอร์เลอกู (Huseyin Husnu; & Giray Berberoglu. 2009) ศึกษาการตัดสินใจและการวิเคราะห์การทำหน้าที่ต่างกันของแบบสอบการรู้เรื่องคณิตศาสตร์ PISA 2003 โดยเปรียบเทียบลักษณะของมนุษย์ในกลุ่มภาษาและวัฒนธรรมที่แตกต่างกัน เป็นประเด็นที่สำคัญของการประเมินผลการศึกษา โดยเพิ่มความสนใจการศึกษาเปรียบเทียบระหว่างประเทศ เพื่อความเป็นธรรมของผลลัพธ์ของกลุ่มที่มีภาษาและวัฒนธรรมแตกต่างกัน ในการศึกษาครั้งนี้ผู้วิจัยจำลองข้อมูลขึ้นเพื่อศึกษาการทำหน้าที่ต่างกันของข้อสอบ (DIF) โดยใช้ชุดข้อมูลสำหรับการประเมินการรู้เรื่องคณิตศาสตร์ (PISA) 2003 โดยวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ DIF ด้วยวิธีแมนเทล-เฮนส์เซล Mantel-Haenszel (MH) และทฤษฎีการตอบสนองข้อสอบ ด้วยวิธีความน่าจะเป็นอัตราส่วน (IRT-LR) ผลการตัดสินใจ (1) ทักษะทางปัญญา (2) ข้อผิดพลาดทางด้านภาษาและ (3) การใช้ปริมาณคำ ทั้งสามเป็นแหล่งความคลาดเคลื่อนที่สำคัญการทำหน้าที่ต่างกันของข้อสอบสำหรับการประเมินการรู้เรื่องคณิตศาสตร์ (PISA) 2003 ผลการวิจัยจากการเปรียบเทียบข้ามวัฒนธรรม การวิเคราะห์เนื้อหาอย่างเป็นระบบอาจช่วยคาดการณ์การทำหน้าที่ต่างกันของข้อสอบได้

จอห์น โอ แอนเดอร์สัน และคณะ (John O. Anderson; et al. 2007) ศึกษาการใช้ชุดข้อมูลขนาดใหญ่สำหรับการประเมินผล เพื่อการวิจัยด้านการรู้เรื่องวิทยาศาสตร์และการรู้เรื่องคณิตศาสตร์ โดยใช้โปรแกรมสำหรับการประเมินผลนักเรียนนานาชาติ (PISA) การประเมินผลสัมฤทธิ์ของนักเรียนในวงกว้างถือเป็นกรอบแนวคิดเกี่ยวกับผลสัมฤทธิ์ทางการเรียนในความสัมพันธ์กับการเรียนรู้บางอย่างเช่น ความเป็นอยู่ ทัศนคติ การรับรู้ และภูมิหลังของนักเรียน ผลการวิจัยการประเมินโครงการประเมินผลนักเรียนนานาชาติ (PISA) เป็นตัวอย่างที่สามารถอธิบายและพูดถึงผลลัพธ์ที่ได้จากการประเมินผล PISA 2000 และ PISA 2003 สะท้อนให้เห็นข้อมูลเชิงลึกจากการทดสอบ สำหรับผลและแนวโน้มการรู้เรื่องคณิตศาสตร์ PISA 2003 พบว่า เนื้อหาเรื่องพื้นที่และรูปทรง มีระดับช่วงคะแนน 544 – 607 คิดเป็นร้อยละ 30 ของกลุ่มตัวอย่าง OECD นักเรียนมีความสามารถอธิบายถึงวิธีการแก้ปัญหาได้ มีการให้เหตุผลยกตัวอย่างประกอบภายใต้บริบทที่มีความแตกต่าง มีการเชื่อมโยงและดำเนินการตามกระบวนการ (OECD. 2005: 55) ซึ่งพบว่า เพศชายมีคะแนนการรู้เรื่องคณิตศาสตร์สูงกว่าเพศหญิง ยกเว้นประเทศไอซ์แลนด์ที่ปรากฏเด่นชัดว่าเพศหญิงมีคะแนนการรู้เรื่องคณิตศาสตร์สูงกว่าเพศชาย แม้จะพบความแตกต่างระหว่างเพศแต่ก็ไม่มี ความแตกต่างทางนัยสำคัญทางสถิติ นอกจากนี้ความแตกต่างทางเพศทั้งสี่เนื้อหา พบว่า เรื่องการเปลี่ยนแปลงและความสัมพันธ์มีระดับความแตกต่างเพศชายน้อยกว่าเพศหญิง และเรื่องพื้นที่และรูปทรงสามมิติ ความแตกต่างเพศชายกับเพศหญิงต่ำสุด

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องการรู้เรื่องคณิตศาสตร์เป็นเป้าหมายของ PISA สำหรับด้านคณิตศาสตร์ คือ การสร้างตัวชี้วัดที่แสดงถึงว่าประเทศชาติหนึ่งๆ ได้เตรียมตัวนักเรียนวัยอายุ 15 ปี ให้มีความพร้อมเพียงใดในด้านการใช้ความรู้คณิตศาสตร์เพื่อให้เป็นพลเมืองที่มีความคิด และมีสติปัญญาพร้อมที่จะมีส่วนเป็นผู้สร้างสังคม เพื่อบรรลุเป้าหมายนี้ PISA จึงพัฒนาการประเมินผลที่เน้นความชัดเจนกับการตรวจสอบว่านักเรียนสามารถใช้ความรู้ที่เรียนมาได้อย่างมีประสิทธิภาพหรือไม่เพียงใด PISA เน้นที่การใช้กระบวนการทางคณิตศาสตร์ ความรู้ และความเข้าใจคณิตศาสตร์ในการแก้ปัญหาที่เกิดขึ้นในชีวิตจริง โดยนักเรียนต้องคิด ต้องประยุกต์ใช้ความรู้ในการแก้ปัญหาด้วยตนเอง แนวโน้มของการรู้เรื่องคณิตศาสตร์ตามความหมายของประเทศสมาชิก OECD/PISA ยอมรับและนำมาใช้ในการประเมินความพร้อมของเยาวชน ในการเป็นประชาชนที่มีคุณภาพและมีสมรรถนะทางเศรษฐกิจในอนาคต อาจจะไม่เหมือนสิ่งที่เรากำลังพร่ำสอนกันอยู่ในโรงเรียน จุดเน้นของ OECD คือ คณิตศาสตร์ในชีวิตจริง ซึ่งมีใช้หมายความว่าเพียงการนำความรู้คณิตศาสตร์มาใช้เล็กน้อยๆ เช่น นำมาคิดหาต้นทุน กำไรแต่ยังหมายรวมไปถึงการใช้คณิตศาสตร์สองทาง คือ เมื่อรู้ว่าคุณคณิตศาสตร์และนำไปใช้ได้ยังไม่พอ แต่ต้องรู้จักโลกของความจริง สามารถนำปัญหาจริงที่พบในโลกมาคิดในเชิงของคณิตศาสตร์ สามารถแก้ปัญหาในเชิงของคณิตศาสตร์ โดยใช้ความรู้คณิตศาสตร์และแปลงการแก้ปัญหาคณิตศาสตร์ไปตอบปัญหาในโลกของความเป็นจริงอีกต่อหนึ่ง ดังนั้นคณิตศาสตร์ที่เป็นจุดเน้นของเขา คือ คณิตศาสตร์ในโลกของความจริงการสอนคณิตศาสตร์จึงไม่จำกัดอยู่เฉพาะหัวข้อคณิตศาสตร์แบบดั้งเดิม ดังที่กำลังเน้นกันอยู่ในหลักสูตรในโรงเรียนเท่านั้น

ตอนที่ 2 แนวคิดและทฤษฎีที่เกี่ยวกับกฎเกณฑ์การให้คะแนน (Scoring Rubric)

การวัดผลและประเมินผลเป็นกระบวนการตรวจสอบคุณภาพของผู้เรียนด้วยวิธีการที่หลากหลาย เช่นการประเมินตามสภาพจริง การใช้แบบทดสอบ ซึ่งเครื่องมือทางการวัดผลแต่ละวิธีก็มีจุดเด่นและข้อจำกัด ที่แตกต่างกัน ดังนั้นการเลือกใช้เครื่องมือหรือวิธีในการวัดผลและประเมินผลในแต่ละครั้ง ควรคำนึงถึงความเหมาะสม และจากการศึกษาเอกสารที่เกี่ยวข้องการสร้างเครื่องมือในการวัดผลและประเมินผลวิชาคณิตศาสตร์ ตามหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 กลุ่มสาระการเรียนรู้คณิตศาสตร์ และคู่มือวัดผลประเมินผลคณิตศาสตร์ของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) ผู้วิจัยเลือกใช้วิธีในการวัดกระบวนการทางคณิตศาสตร์ตามคำนิยามของการรู้เรื่องคณิตศาสตร์ โดยใช้แบบทดสอบที่มีกฎเกณฑ์การให้คะแนน (Scoring Rubric)

Rubric มาจากภาษาละตินว่า “Rubrica Terra” ซึ่งเป็นคำโบราณที่ใช้ในทางศาสนา หมายถึง การทำเครื่องหมายสีแดงไว้บนสิ่งสำคัญ ส่วนในทางการศึกษานั้น “Rubric” หมายถึง สมรรถภาพหรือรายการของคุณลักษณะที่บรรยายถึงความสามารถในการแสดงออกของนักเรียนเพื่อใช้เป็นแนวทางในการให้คะแนนที่ชัดเจนคำว่า “Rubric” มีความหมายเช่นเดียวกับคำว่า “กฎ หรือกติกา” (Rule) ส่วนคำว่า “Rubric Assessment” หมายถึง แนวทางในการให้คะแนน (Scoring Guide) ที่สามารถแยกแยะความสำเร็จในการเรียนหรือคุณภาพ การปฏิบัติของนักเรียนได้อย่างชัดเจนจากระดับดีมากจนไปถึงระดับที่ต้องปรับปรุงแก้ไข ดังนั้น “Rubric” จึงเป็นเครื่องมือในการให้คะแนน ซึ่งประกอบด้วยเกณฑ์ด้านต่างๆที่ใช้พิจารณาชิ้นงานหรือการปฏิบัติงาน และระดับคุณภาพของคะแนนของเกณฑ์แต่ละด้าน (สุรศักดิ์ อมรรัตนศักดิ์. 2554: 254 - 255)

2.1 ความหมายของกฎเกณฑ์การให้คะแนน (Scoring rubrics)

สเทนมาร์ค (ซูลีพร ดัดงาม. 2552: 35; อ้างอิงจาก Stenmark.1991: 24) ได้ให้ความหมายของเกณฑ์การให้คะแนน (Scoring Rubric) ที่ใช้สำหรับคำถามแบบปลายเปิด (Open-Ended Question) ซึ่งได้มีการปรับปรุงมาจาก California Assessment Program Grade 12 Generalized Rubric ไว้ว่าเป็นเกณฑ์การให้คะแนนเฉพาะ (Specific Rubric) เป็นแนวทางในการให้คะแนนเฉพาะของคำถามแบบปลายเปิดของแต่ละข้อ สำหรับเกณฑ์การให้คะแนนทั่วไป (General Rubric) เป็นตัวแบบของเกณฑ์การให้คะแนน เพื่อใช้พิจารณาเกณฑ์โดยเฉพาะรายละเอียดจะเปลี่ยนแปลงสำหรับระดับคะแนนที่ต่างกัน แนวคิดพื้นฐานสามารถประยุกต์ใช้ได้สำหรับทุกระดับ

กู๊ดริช (อรรถนีย์ ชูช่วยสุวรรณ . 2552: 22; อ้างอิงจาก Goodrich. 1997) กล่าวว่า “รูบรีค” (Rubric) เป็นเครื่องมือการให้คะแนนซึ่งจะแสดงเกณฑ์สำหรับชิ้นงานจะเป็นการเชื่อมต่อกัน ในการแบ่งลำดับของคุณลักษณะแต่ละเกณฑ์ จากดีมากจนถึงต้องปรับปรุงแก้ไข

โรนีส (อรรถนีย์ ชูช่วยสุวรรณ . 2552: 23; อ้างอิงจาก Ronis. 2000) กล่าวว่า รูบรีคเป็นแนวทางการให้คะแนนผลงาน หรือการปฏิบัติของนักเรียนรูบรีคจะเป็นตัวแทนคุณลักษณะที่สังเกตได้จากผลงาน กระบวนการ หรือการปฏิบัติ รูบรีคยังช่วยให้ระบบการให้คะแนนมีความง่ายในการเรียนรู้และนำไปใช้ และปรับปรุงผลงานการปฏิบัติ

เกรนท์ วิกินส์ (ราตรี นันทสคุณธ์. 2553: 71; อ้างอิงจาก Grant Wiggins. 2000) ได้ให้ความหมายว่า Rubric Score หมายถึง แนวทางในการให้คะแนนของการประเมินแบบอัตโนมัติ ซึ่งเป็นการกำหนดกฎเกณฑ์ของการประเมินกระบวนการปฏิบัติงาน มีการอธิบายคุณลักษณะของกระบวนการปฏิบัติงาน ณ จุดต่างๆ รวมทั้งสามารถระบุคุณภาพของการปฏิบัติงาน ณ จุดนั้นๆได้อีกด้วย

นอแมน กรอร์น (ราตรี นันทสคุณธ์. 2553: 71; อ้างอิงจาก Norman; & Gronlund.1988) ได้ให้ความหมายว่า Rubric Score หมายถึง เป็นชุดของการอธิบายคุณลักษณะของระดับคะแนนต่างๆของกระบวนการปฏิบัติงานที่ใช้สำหรับตัดสินการปฏิบัติงานหรือชิ้นงาน

เสาวนีย์ เกรียร์ (2540: 159) กล่าวว่า รูบรีคเป็นเครื่องมือในการให้คะแนนที่ประกอบด้วยประเด็นต่างๆ ที่จะใช้พิจารณางานหนึ่ง และคำอธิบายระดับคุณภาพของแต่ละประเด็นประเมิน ซึ่งอาจเรียงลำดับตั้งแต่ดีเลิศ ไปจนถึงต้องปรับปรุง หรือให้เป็นระดับตัวเลขตั้งแต่มากที่สุด ประเด็นประเมินอาจกำหนดเพิ่มเติมได้หลายข้อ คำอธิบายระดับคุณภาพควรอธิบายให้ชัดเจน กระชับที่สุด เป็นคำอธิบายที่สามารถบอกได้ว่า ทำไม่ต้องดีเลิศ ดี ต้องปรับปรุง

บุญเชิด ภิญโญอนันตพงษ์ (2544: 89) กล่าวว่า กฎเกณฑ์การให้คะแนนคือชุดของแนวทางให้คะแนนผลการปฏิบัติเรื่องใดเรื่องหนึ่งสำหรับใช้ประเมินคุณภาพของการปฏิบัติของผู้เรียน แนวทางในการให้คะแนนอาจทำให้รูปของมาตราประมาณค่าหรือแบบตรวจสอบรายการ โดยปกติกฎเกณฑ์การให้คะแนนชุดหนึ่งสำหรับประเมินจุดหมายความการเรียนรู้ข้อหนึ่งหรือส่วนหนึ่งของการปฏิบัติ ในกรณีที่การปฏิบัติมีความสลับซับซ้อนมากขึ้น จำเป็นต้องทำการประเมินหลายๆ จุดหมายความการเรียนรู้ หรือการปฏิบัติหลายๆอย่าง ดังนั้นจึงต้องใช้กฎเกณฑ์การให้คะแนนหลายชุดในการประเมินการปฏิบัติจากหลายๆ จุดหมายความการเรียนรู้ กฎเกณฑ์การให้คะแนนแต่ละชุดประกอบด้วยตัวเลขที่สะท้อนระดับคุณภาพของการปฏิบัติ เช่น 1 ถึง 4 เมื่อ 4 หมายถึงคุณภาพระดับสูงสุด, 3 หมายถึง คุณภาพระดับสูง, 2 หมายถึง คุณภาพระดับพอใช้, และ 1 หมายถึงคุณภาพยังไม่เป็นที่น่าพอใจ

สมศักดิ์ ภู่วิภาดาพรรณ (2544: 137 - 152) ให้ความหมายของ รุบริกส์ (Rubrics) ว่าเป็นเครื่องมือในการให้คะแนน (Scoring Tool) ที่มีการระบุเกณฑ์ (Criteria) ประเมินชิ้นงาน และคุณภาพ (Quality) ของชิ้นงานในแต่ละเกณฑ์ เป็นการประเมินศักยภาพของผู้เรียนโดยการลงมือปฏิบัติ นั้น ไม่มีคำเฉลยหรือคำตอบถูกที่แน่ชัดลงไปเหมือนแบบทดสอบเลือกตอบ การประเมินผลงานแต่ละชิ้นของผู้เรียนที่ได้ลงมือปฏิบัติจึงมีความจำเป็นที่จะต้องประเมินคุณภาพของงานอย่างเป็นปรนัย ซึ่งเป็นการยากที่จะทำได้ และได้ค้นพบการสร้างเกณฑ์การให้คะแนน หรือ rubric ขึ้นมา

องอาจ นัยพัฒน์ (2558: 25) กล่าวว่า กฎเกณฑ์การให้คะแนน (Scoring rubrics) คือแนวทางสำหรับใช้ประกอบการตัดสินใจในการให้คะแนน (Scoring guidelines) ซึ่งบ่งชี้ผลการเรียนรู้ของผู้เรียนได้อย่างเที่ยงตรงคงเส้นคงวาและยุติธรรม โดยมีการกำหนดเกณฑ์การให้คะแนน (Scoring criteria) ผลการเรียนรู้ของผู้เรียนที่นำมาใช้เพื่อให้มั่นใจในความเที่ยงตรง ความคงเส้นคงวาและความยุติธรรมของการวัดประเมิน และเป็นเครื่องมือสำหรับให้คะแนน (Scoring tool) ที่ระบุความคาดหวังอย่างเจาะจงสำหรับงานหรือหน้าที่ที่มอบหมายให้ผู้เรียนปฏิบัติ

จากการศึกษาความหมายของกฎเกณฑ์การให้คะแนน (Scoring rubrics) ข้างต้น ผู้วิจัยสรุปได้ว่ากฎเกณฑ์การให้คะแนน (Scoring rubrics) หมายถึง เครื่องมือที่มีแนวทางการให้คะแนนเรื่องใดเรื่องหนึ่งที่มี ประกอบด้วยประเด็นต่างๆ และคำอธิบายระดับคุณภาพแต่ประเด็นที่ต้องการประเมิน ซึ่งคำอธิบายระดับคุณภาพควรอธิบายให้ชัดเจน กระชับที่สุด เพื่อให้มีความเป็นปรนัย

ความสำคัญของกฎเกณฑ์การให้คะแนน (Scoring rubrics)

ก๊อริช (อรรถนีย์ ชูช่วยสุวรรณ. 2552: 22 – 23; อ้างอิงจาก Goodrich.1997) ได้กล่าวถึงความสำคัญของรูบริกไว้ดังนี้

1. รูบริกสามารถจะพัฒนาการปฏิบัติงานของนักเรียนคล้ายกับเป็นการกำกับติดตาม โดยทำให้ความคาดหวังของครูมีความชัดเจน และเป็นการแสดงว่านักเรียนจะพบกับสิ่งที่คาดหวังได้อย่างไร ผลที่ได้จะเป็นการบอกการปรับปรุงแก้ไขในคุณลักษณะของงานนักเรียน และในการเรียนรู้
2. รูบริกสามารถช่วยตัดสินเกี่ยวกับคุณลักษณะงานของนักเรียนด้วยตนเองและงานอื่นๆ เมื่อรูบริกนำมาใช้ประเมินตนเองและกลุ่มจะทำให้เพิ่มความสามารถในการแก้ปัญหาขึ้น
3. รูบริกจะช่วยลดเวลาของครูที่ใช้ในการประเมินงานของนักเรียนได้
4. รูบริกใช้ง่าย และมีคำอธิบายที่ชัดเจน

โรนีส (อรรถนีย์ ชูช่วยสุวรรณ. 2552: 23; อ้างอิงจาก Ronis. 2000) ได้กล่าวว่า
 รูบริกจะเป็นส่วนสำคัญที่จะช่วยให้ครูตอบคำถามต่อไปนี้

1. เป้าหมายการปฏิบัติของนักเรียนคืออะไร
2. อะไรเป็นสิ่งที่นักเรียนต้องรู้อย่างแท้จริง โดยขึ้นอยู่กับเกณฑ์มาตรฐาน
3. อะไรจะเป็นหลักฐานประกอบความเข้าใจของนักเรียน
4. ผลงานจะต้องมีลักษณะอย่างไรบ้าง
5. อะไรเป็นสิ่งที่ผิดพลาดแสดงถึงคะแนนที่น้อยลง

องอาจ นัยวัฒน์ (2558: 25) ความสำคัญของกฎเกณฑ์การให้คะแนน (Scoring rubrics) ช่วยนิยามความคาดหวังของอาจารย์ (ผู้ทำการวัดประเมิน) ชัดเจนยิ่งขึ้น ช่วยนิยามผลการเรียนรู้ที่ซับซ้อน/เป็นนามธรรมมีความชัดเจนขึ้น ช่วยให้ได้สารสนเทศป้อนกลับ (Feedback) ไปยังผู้เรียนได้ทันเหตุการณ์ ทำให้มีแรงจูงใจและผลสัมฤทธิ์ของผู้เรียนเพิ่มสูงขึ้น โดยอาศัยตัวแบบผลผลิตและการปฏิบัติงานที่มีคุณภาพเป็นสิ่งเสริมแรง ช่วยให้ผลการวัดประเมินการเรียนรู้ของผู้เรียนด้านเดียวกันมีความเที่ยงตรงและความเชื่อมั่น ช่วยให้การสื่อสารกับผู้มีส่วนได้ส่วนเสียกับการวัดประเมินเข้าใจตรงกัน ให้แนวทางใน “การปรับการเรียน” ของผู้เรียน และ “เปลี่ยนการสอน” ของอาจารย์ ก่อให้เกิดการกระตุ้นการคิดอย่างมีวิจารณญาณ (Critical thinking) และส่งเสริมทักษะการวัดประเมินการเรียนรู้ของตนเองตลอดชีวิต เอื้ออำนวย การตรวจให้คะแนนดำเนินไปอย่างมีประสิทธิภาพ

ฉัตรศิริ ปิยะพิมลสิทธิ์ (2544) ได้กล่าวว่า รูบริกมีความสำคัญต่อการประเมิน ศักยภาพของผู้เรียนโดยให้ลงมือปฏิบัตินั้น ไม่มีคำเฉลยหรือคำตอบถูกที่แน่ชัดลงไปเหมือนแบบทดสอบเลือกตอบ การประเมินผลงานแต่ละชิ้นของผู้เรียนที่ได้ลงมือปฏิบัติจึงมีความจำเป็นที่จะต้องประเมินคุณภาพของงานอย่างเป็นปรนัย ซึ่งเป็นการยากที่จะทำได้ และได้ค้นพบการสร้างเกณฑ์การให้คะแนนหรือ rubric ขึ้นมาเพื่อกำหนดแนวทางในการตัดสินอย่างยุติธรรม และปราศจากความลำเอียง Rubric จะต้องมีความชัดเจนในเกณฑ์การให้คะแนนอย่างพอเพียงถึงขนาดที่ผู้ประเมิน 2 คนสามารถใช้ Rubric เดียวกันประเมินชิ้นงานของผู้เรียนชิ้นเดียวกันแล้วให้คะแนนได้ตรงกัน ระดับของความสอดคล้องในการให้คะแนนของผู้ประเมิน 2 คนที่ประเมินอย่างเป็นอิสระจากกันจะเรียกว่า ความเชื่อมั่น (Reliability) ของการประเมิน

สมศักดิ์ ภู่วิภาดาพรรณ (2544: 137 – 152) ซึ่งมีความสำคัญดังนี้

1. เพื่อกำหนดแนวทางในการตัดสินอย่างยุติธรรม และปราศจากความลำเอียง Rubric จะต้องมีความชัดเจนในเกณฑ์การให้คะแนนอย่างพอเพียงถึงขนาดที่ผู้ประเมิน 2 คนสามารถใช้ Rubric เดียวกันประเมินชิ้นงานของผู้เรียนชิ้นเดียวกันแล้วให้คะแนนได้ตรงกัน ระดับของความสอดคล้องในการให้คะแนนของผู้ประเมิน 2 คนที่ประเมินอย่างเป็นอิสระจากกันจะเรียกว่า “ความเชื่อมั่น (Reliability)” ของการประเมิน

2. เป็นเครื่องมือที่สามารถใช้ได้ทั้งการสอนและการประเมิน เราสามารถใช้ Rubric เพื่อพัฒนาหรือปรับปรุงการปฏิบัติงานของนักเรียนได้ และช่วยให้ครูสามารถตั้งความคาดหวังกับการปฏิบัติงานของนักเรียนได้อย่างชัดเจน นอกจากนี้ยังสามารถให้นักเรียนเห็นได้อย่างชัดเจนว่าทำอะไรจึงจะปฏิบัติงานได้ตามความคาดหวังที่ตั้งไว้

3. เป็นเครื่องมือที่มีประโยชน์ในการช่วยเหลือนักเรียนให้เป็นผู้ที่สามารถตัดสินคุณภาพชิ้นงานอย่างมีเหตุผล ทั้งงานของตนเองและผู้อื่น นักเรียนจะรู้ข้อผิดพลาดของตนเองและผู้อื่น การทำเช่นนี้บ่อยๆ ช่วยให้นักเรียนเกิดความรับผิดชอบในงานของตนเองมากยิ่งขึ้น

4. เป็นเครื่องมือที่ช่วยลดเวลาที่ครูใช้ในการประเมินผลงานของนักเรียนลงได้ เพราะโดยปกติครูมักประเมินผลงานของนักเรียนทีละชิ้น ถ้าใช้ Rubric ในการประเมินงานนักเรียนสามารถประเมินงานของตนเองและของเพื่อนๆ ได้ นอกจากนี้ Rubric ยังช่วยให้นักเรียนได้ข้อมูลย้อนกลับเกี่ยวกับจุดเด่น และสิ่งที่ควรปรับปรุงแก้ไขในชิ้นงานของตนเองได้อีกด้วย

5. Rubric มีลักษณะยืดหยุ่นที่สามารถทำให้ครูสอนนักเรียนที่มีความหลากหลายแตกต่างกันไปได้อย่างดี

6. Rubric ใช้ได้ง่ายและอธิบายได้ง่ายเช่นกัน การใช้ Rubric จะช่วยให้นักเรียนทราบว่านักเรียนได้เรียนรู้อะไร และเมื่อมีการประชุมผู้ปกครอง ครูอาจใช้ Rubric อธิบายผู้ปกครองเข้าใจง่าย โดยผู้ปกครองจะทราบได้ว่าบุตรหลานของตนต้องทำอะไรบ้างจึงจะประสบผลสำเร็จในการเรียน

2.2 องค์ประกอบของกฎเกณฑ์การให้คะแนน (Scoring rubrics)

องอาจ นัยวัฒน์ (2558: 25) กำหนดองค์ประกอบกฎเกณฑ์การให้คะแนน (Scoring rubrics) ดังนี้

1. การอธิบายงานหรือภารกิจ (Tasks) ที่มอบหมายให้ผู้เรียนปฏิบัติ (เพื่อจะได้ทราบว่ามีผลการเรียนรู้ผ่านมาตรฐานคุณภาพตามที่หลักสูตรระบุไว้หรือไม่)
2. ระดับผลการปฏิบัติงาน (Performance levels) (ระดับมาตรฐานของผลการปฏิบัติงาน หรือสมรรถนะจากการเรียนรู้ที่ใช้ในการปฏิบัติงาน)
3. เกณฑ์ (หรือมิติ) คุณภาพ (Criteria or dimensions of quality) (องค์ประกอบคุณภาพในด้านต่างๆ ของงานหรือภารกิจที่มอบหมายให้ปฏิบัติ)
4. ตัวบ่งชี้สมรรถนะการปฏิบัติงาน (Performance descriptors)
5. ข้อความบรรยายระดับคุณภาพตามมาตรฐานต่างๆ ของผลการปฏิบัติงานตามเกณฑ์ใดเกณฑ์หนึ่งที่ต้องการวัดประเมิน

สุรศักดิ์ อมรรัตนศักดิ์ (2554: 256) ได้มีการกำหนดเกณฑ์การประเมินต้องประกอบด้วย 3 องค์ประกอบ คือ องค์ประกอบที่หนึ่ง เกณฑ์หรือประเด็นที่จะประเมิน (Criteria) เป็นการพิจารณาว่าภาระงานหรือชิ้นงานนั้นๆ ประกอบด้วยคุณภาพกี่ด้าน อะไรบ้าง องค์ประกอบที่สอง ระดับความสามารถหรือระดับคุณภาพ (Performance Level) เป็นการกำหนดจำนวนระดับของเกณฑ์ส่วนมากจะประกอบด้วย 3 - 6 ระดับ และองค์ประกอบที่สาม การบรรยายคุณภาพของแต่ละระดับความสามารถ (Quality Description) เป็นการเขียนคำอธิบายความสามารถให้เห็นถึงความแตกต่างอย่างชัดเจนในแต่ละระดับ ซึ่งจะช่วยให้ง่ายต่อการตรวจคะแนน การกำหนดเกณฑ์การให้คะแนนจะต้องประกอบด้วย 3 องค์ประกอบ คือ

1. ประเด็นที่จะประเมิน (Criteria) คือ สิ่งที่จะสะท้อนผลการเรียนรู้หลักๆ หรือมาตรฐานการเรียนรู้ที่เป็นเป้าหมายของแต่ละหน่วย/ภาระงาน
2. ระดับความสามารถ (Performance Levels) ส่วนใหญ่จะกำหนดเป็นเลขคี่มากกว่าเลขคู่ ทั้งนี้เพื่อป้องกันการให้คะแนนที่ตกอยู่ตรงกลาง ทำให้จำแนกความสามารถได้ยาก และแต่ละระดับอาจกำหนดเป็นตัวเลขหรือคำแสดงคุณภาพต่างๆ เช่น ดีมาก ดี พอใช้ ยังต้องปรับปรุง เป็นต้น
3. คำอธิบายคุณภาพของแต่ละระดับความสามารถ (Quality Descriptors) ว่าคุณภาพของความสามารถแต่ละระดับที่คาดหวังนั้นเป็นอย่างไร คำบรรยายเหล่านี้จะต้องมีความชัดเจนในการใช้ภาษาที่กระชับรัดกุม เข้าใจง่ายและเห็นความแตกต่างระหว่างระดับความชัดเจน

ฉัตรศิริ ปิยะพิมลสิทธิ์ (2544) Scoring rubric มีหลายองค์ประกอบ ในแต่ละองค์ประกอบก็มีประโยชน์ มีความสำคัญ องค์ประกอบมีดังนี้

1. จะมีอย่างน้อย 1 คุณลักษณะหรือ 1 มิติที่เป็นพื้นฐานในการตัดสินผู้เรียน
2. การนิยามและการยกตัวอย่างจะต้องมีความชัดเจนในแต่ละคุณลักษณะหรือมิติ
3. มาตรการให้คะแนนจะต้องเป็นอัตราส่วนกันในแต่ละคุณลักษณะหรือมิติ
4. จะต้องมีความมาตรฐานที่เด่นชัดในแต่ละระดับของการให้คะแนน

ในแต่ละระดับการให้คะแนนจะต้องมีความชัดเจนในการนิยาม และความกว้างของระดับคะแนน ไม่ควรเกิน 6 ถึง 7 ระดับ ถ้ามีระดับของการให้คะแนนกว้างมากเกินไปจะมีความลำบากในการตัดสิน ความแตกต่างในแต่ละระดับ เช่น ความกว้างคะแนนเป็น 100 ทำให้ยากที่จะอธิบายว่าคะแนน 81 มีคุณภาพแตกต่างจาก 80 หรือ 82 อย่างไร และจะทำให้ความสอดคล้องของการประเมินด้วยผู้ประเมินหลายคนลดลงไป การจะกำหนดความกว้างของการให้คะแนนเป็นเท่าไรนั้น จะต้องมีความเหมาะสมและ มีความชัดเจนในการนิยามที่ครอบคลุมตั้งแต่แย่ที่สุด (poor) จนถึงดีเลิศที่สุด (excellent)

2.3 ลักษณะของกฎเกณฑ์การให้คะแนน (Scoring rubrics)

รูบรีค Rubric เป็นชุดคะแนนที่ใช้เป็นแนวทาง สำหรับการประเมินผลงานของผู้เรียน ลักษณะของรูบรีคที่ดี มีดังนี้ (สุรศักดิ์ อมรรัตนศักดิ์. 2554: 256 – 257; อ้างอิงจาก Wiggings. 1998)

1. มีความเกี่ยวข้องกับจุดประสงค์หรือเป้าหมายทั่วไป (General goals) คือเกี่ยวข้องกับงานที่ทำ
2. จำแนกการปฏิบัติได้อย่างเที่ยงตรง (Performance validity)
3. ในแต่ละ Rubric จะไม่มีการรวมเกณฑ์การให้คะแนน
4. วิเคราะห์งานได้อย่างละเอียด
5. ภาษาที่ใช้อธิบายคุณลักษณะงาน จำแนกคุณภาพของงานได้อย่างถูกต้อง
6. สามารถตัดสินงานได้อย่างถูกต้อง
7. อธิบายระดับของคะแนนได้อย่างชัดเจนและมีความเที่ยงตรงการให้คะแนน
8. ตัดสินการให้คะแนนจากผลงานปฏิบัติมากกว่ากระบวนการรูปแบบเนื้อหาหรือความตั้งใจในการทำงาน

นอกจากนี้ สรุตศักดิ์ อมรรัตนศักดิ์. (2554: 257; อ้างอิงจาก Wiggings. 1998) ยังได้กล่าวถึงคุณลักษณะของ Rubric เพิ่มเติม ดังนี้

1. คะแนนต้องมีลักษณะต่อเนื่อง (Continuous) กล่าวคือ ให้คะแนนเป็นจำนวนเต็ม เช่น ให้คะแนนเป็น 5, 4, 3, 2, และ 1 คะแนน แต่ละคะแนนมีความห่างเท่ากัน
2. มีความสอดคล้องกัน (Parallel) คะแนนแต่ละระดับแสดงถึงความลดหลั่นของคุณภาพงาน
3. มีความเกี่ยวเนื่องกัน (Coherent) ในแต่ละระดับของการให้คะแนน
4. น้ำหนักการให้คะแนนในแต่ละระดับมีความเหมาะสม มีเหตุผล (not arbitrary) น้ำหนักของการให้คะแนนในแต่ละระดับสามารถอ้างอิงไปยังระดับอื่นๆได้
5. มีความเที่ยง (Valid) คะแนนในแต่ละระดับ แสดงถึงคุณภาพของการปฏิบัติ เป็นสิ่งสะท้อนถึงคุณภาพของงาน ไม่ได้เน้นถึงปริมาณ แต่เป็นเกณฑ์ตามสภาพจริง (Authentic criteria)
6. เชื่อถือได้ (Reliable) กล่าวคือ มีความคงเส้นคงวาในการให้คะแนน ถึงแม้จะให้ใครเป็นผู้ประเมิน และจะประเมินช่วงใดก็ตาม

2.4 ประเภทของกฎเกณฑ์การให้คะแนน (Scoring rubrics)

องอาจ นัยวัฒน์ (2558: 26 - 27) ได้แบ่งประเภทกฎเกณฑ์การให้คะแนน (Scoring rubrics) เป็น 2 ประเภท คือ ประเภทที่ 1 แบบองค์รวมหรือแบบวิเคราะห์ : ตัดสินโดยรวมหรือรายด้าน ได้แก่ กฎเกณฑ์แบบองค์รวม (Holistic rubrics) เป็นการบรรยายคุณภาพของผลการปฏิบัติงาน (สิ่งที่ต้องการวัดประเมิน) แบบกว้างๆ โดยรวม 1 ค่า (A single overall or global scoring) และกฎเกณฑ์แบบวิเคราะห์ (Analytic rubrics) เป็นการบรรยายคุณภาพของผลการปฏิบัติงาน (สิ่งที่ต้องการวัดประเมิน) แบบแบ่งแยกส่วน ในแต่ละองค์ประกอบจำเป็น (Component) หรือลักษณะสำคัญ (Feature) ที่ใช้เป็นเกณฑ์การวัดประเมิน (Separate scoring for each of the criteria of the assessment scale) แบบทั่วไปหรือแบบเจาะจง : บรรยายคุณภาพทุกงานหรือเฉพาะบางงาน ได้แก่ กฎเกณฑ์แบบทั่วไป (Generic rubrics) เป็นการบรรยายคุณภาพของผลการปฏิบัติงาน (สิ่งที่ต้องการวัดประเมิน) โดยใช้คำศัพท์หรือระบุข้อกำหนดในรูปทั่วไป ดังนั้นจึงสามารถนำไปใช้ให้คะแนนผลการปฏิบัติงานที่แตกต่างกันได้หลายงาน และกฎเกณฑ์แบบเจาะจง (Task-specific rubrics) เป็นการบรรยายคุณภาพของผลการปฏิบัติงาน (สิ่งที่ต้องการวัดประเมิน) โดยใช้คำศัพท์หรือระบุข้อกำหนดในรูปเจาะจงเฉพาะงาน (Particular task)

ดังนั้นจึงนำไปใช้ให้คะแนนผลการปฏิบัติงานเฉพาะงานใดงานหนึ่งเท่านั้น ประเภทที่ 2 เป็นกฎเกณฑ์ที่สามารถจำแนกและบรรยายได้ทั้ง 2 มิติ มิติที่หนึ่ง การตัดสินโดยรวมบนมาตรเดียว (One scale) หรือแยกส่วนบนมาตรรายด้าน (Several scale) มิติที่สอง การบรรยายคุณภาพของงานที่ปฏิบัติในรูปทั่วไป (General) หรือเจาะจงเฉพาะ (Specific)

มิติที่ 2 \ มิติที่ 1	แบบองค์รวม	แบบวิเคราะห์
แบบทั่วไป	แบบทั่วไป/องค์รวม (Generic/holistic rubrics)	แบบทั่วไป/วิเคราะห์ (Generic/analytic rubrics)
แบบเจาะจง	แบบเจาะจง/องค์รวม (Generic/analytic rubrics)	แบบเจาะจง/วิเคราะห์ (Task-specific/analytic rubrics)

ที่มา: องอาจ นัยพัฒน์. 2558 เอกสารประกอบการบรรยายมาตรฐานผลการเรียนรู้ตามกรอบ TQF: HEd และกฎเกณฑ์ (แนวทาง) การให้คะแนน: แนวคิดพื้นฐานและวิธีการสร้างเพื่อวัดประเมินผลการเรียนรู้ตามกรอบ TQF: HEd

นิทโก (ปวีณา ปิอาทิตย์. 2545: 19 – 20; อ้างอิงจาก Nitko. 2000) กล่าวถึง เกณฑ์การให้คะแนนว่าเป็นกฎเกณฑ์หรือหลักที่ใช้ในการประเมินคุณภาพการปฏิบัติงานของนักเรียนโดยกฎเกณฑ์นี้อาจอยู่ในรูปของมาตรประมาณค่า หรือแบบตรวจสอบรายการ มาตรวัดจะประกอบด้วย ตัวเลขหลายๆตัวที่ใช้บ่อยคือ 0 ถึง 3 หรือ 1 ถึง 4 จะเป็นการบอกถึงระดับคุณภาพของการปฏิบัติ ตัวเลขที่ใช้จะแสดงถึงลำดับของระดับคุณภาพเช่น ตัวเลข 4 หมายถึงระดับสูงสุด (ดีมาก) ตัวเลข 3 หมายถึง ระดับรองลงมา (ดี) ลดหลั่นจนถึงระดับต่ำสุด จากที่กล่าวมาจะพบว่า การให้คะแนนการปฏิบัติของนักเรียนแต่ละคนคุณภาพของการปฏิบัติที่ระดับต่างๆจะบรรยายด้วยเกณฑ์การให้คะแนน

นอกจากนี้เขาได้แบ่งเกณฑ์การให้คะแนน เป็น 2 แบบ คือ เกณฑ์ทั่วไป (general scoring rubric) และเกณฑ์เฉพาะเจาะจง (specific scoring rubric) โดยมีรายละเอียดดังนี้

1. เกณฑ์ทั่วไป (general scoring rubric) เป็นเกณฑ์การให้คะแนนที่ไม่ได้มีการมุ่งเน้นในงานนั้น โดยจะเป็นกรอบทั่วไปในการพัฒนาเกณฑ์ที่เฉพาะเจาะจง

2. เกณฑ์เฉพาะเจาะจง (specific scoring rubric) เป็นเกณฑ์การให้คะแนนที่พัฒนาจากเกณฑ์ทั่วไปในการให้คะแนน แต่ละส่วนของงานนั้น ในการใช้เกณฑ์เฉพาะเจาะจงนี้ในการประเมินนักเรียน ต้องแน่ใจว่าเกณฑ์นี้เป็นส่วนหนึ่งในกรอบการให้คะแนนทั่วไป

นอกจากนี้เขายังได้แบ่งประเภทของเกณฑ์การให้คะแนน เป็น 3 ประเภทดังนี้

1. เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) เกณฑ์ประเภทนี้จะป็นมาตรวัด (scale) หรือคะแนนที่แสดงผลงานหรือกระบวนการโดยรวมทั้งหมด ซึ่งจะไม่แยกคะแนนออกเป็นแต่ละองค์ประกอบหรือรายการ วิธีการนี้ใช้ง่ายและประหยัดเวลา

2. เกณฑ์การให้คะแนนแบบแยกส่วน (analytic scoring rubric) เกณฑ์นี้จะใช้มาตรวัดหรือคะแนนที่แยกลักษณะของผลงานหรือกระบวนการออกเป็นแต่ละองค์ประกอบหรือรายการโดยคะแนนที่ได้จากแต่ละองค์ประกอบหรือรายการ จะต้องรวมคะแนนเป็นคะแนนรวมทั้งหมดของผลงาน วิธีนี้จะใช้เวลามาก ในรายละเอียดของแต่ละระดับจะมีประโยชน์เมื่อต้องการวินิจฉัยหรือช่วยนักเรียนให้เข้าใจถึงสิ่งที่คาดหวังจากการประเมินที่ได้จากการประเมิน จะใช้เวลามากแต่ก็มีประโยชน์อย่างยิ่งที่จะช่วยในการเรียนรู้ของนักเรียน

3. เกณฑ์การให้คะแนนแบบผสมผสาน (annotated holistic scoring rubric) ผู้ประเมินจะใช้มาตรวัดรวม ในการประเมินก่อนซึ่งมาตรวัด หรือคำอธิบายจะแสดงลักษณะของจุดเด่น (strengths) และจุดด้อย (weakness) ที่สนับสนุน มาตรวัดรวม วิธีการนี้เกิดจากการรวมเกณฑ์แบบภาพรวม และแบบแยกองค์ประกอบไว้ด้วยกัน เมื่อประเมินการปฏิบัติโดยภาพรวม จะต้องเลือกหนึ่งหรือสองประเด็นในการอธิบายหรือใช้มาตรวัดแบบแยกองค์ประกอบ ข้อดีของวิธีการนี้จะทำให้มีความรวดเร็วในการให้คะแนนและเป็นการสนับสนุนว่าในแต่ละระดับนั้นจะให้ข้อมูลย้อนกลับแก่นักเรียนได้

จากการแบ่งประเภทของกฎเกณฑ์การให้คะแนน (Scoring rubrics) ดังกล่าวนมาข้างต้น ได้มีนักวิชาการให้คำจำกัดความแตกต่างกันไปสรุปได้ว่า เกณฑ์การให้คะแนน (scoring rubric) โดยทั่วไปจะเสนอ เป็น 2 รูปแบบ คือ 1) เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) เกณฑ์แบบนี้ใช้เวลาตรวจก่อนข้างนาน สร้างความเมื่อยล้าหากชิ้นงานที่ต้องตรวจมีจำนวนมาก แต่ข้อดีคือมีสารสนเทศที่ครบถ้วนส่งกลับไปยังนักเรียน เพื่อพัฒนาการเรียนให้ดียิ่งขึ้นเป็นเกณฑ์การให้คะแนนที่แยกเป็นประเด็นย่อยๆ ตามองค์ประกอบ และกำหนดน้ำหนักคะแนนในแต่ละส่วน 2) เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) แม้ว่ารูปแบบนี้ตรวจง่าย ใช้เวลาน้อยและมีความเชื่อมั่นสูงแต่ก็ยากที่จะอธิบายเหตุผลของคะแนน เป็นเกณฑ์การให้คะแนนที่พิจารณาโดยรวมไม่มีการแยกประเด็นย่อยๆ การให้คะแนนขึ้นอยู่กับคุณภาพของผลงาน

2.5 การสร้างกฎเกณฑ์การให้คะแนน (Scoring rubrics)

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2551 ได้กำหนดขั้นตอนการสร้างเกณฑ์การให้คะแนน (Scoring Rubric) มีขั้นตอนดังนี้

- 1) ศึกษาเอกสารที่เกี่ยวข้องกับการสร้างเกณฑ์การให้คะแนน
- 2) ศึกษาเนื้อหาสาระการเรียนรู้/จุดประสงค์การเรียนรู้/มาตรฐานการเรียนรู้ของวิชาคณิตศาสตร์
- 3) สร้างเกณฑ์การให้คะแนน
- 4) ให้ผู้เชี่ยวชาญตรวจสอบความเที่ยงตรงเชิงเนื้อหาของเกณฑ์ การให้คะแนน
- 5) ทดลองใช้ เพื่อปรับปรุงแก้ไข
- 6) นำเกณฑ์ไปใช้จริง และหาค่าความเชื่อมั่น
- 7) สร้างคู่มือการใช้เกณฑ์

ซึ่งมีเทคนิคการสร้างเกณฑ์การให้คะแนน (Scoring Rubric) ในแต่ละระดับคุณภาพควรมีคำอธิบายที่ชัดเจน โดยมีการกำหนดคุณลักษณะ ดังนี้

- 1) ต้องมีอย่างน้อยหนึ่งคุณลักษณะหรือหนึ่งมิติเป็นพื้นฐานการตัดสินผู้เรียน
 - 2) การนิยามและการยกตัวอย่างต้องมีความชัดเจนในแต่ละคุณลักษณะหรือมิติ
 - 3) เกณฑ์/ประเด็นที่ประเมิน ต้องเป็นประเด็นสำคัญของงานเท่านั้น
 - 4) ต้องมีความสอดคล้องกันระหว่างเกณฑ์/ประเด็นที่ประเมิน ระดับคะแนน กับจุดมุ่งหมายของการประเมิน
 - 5) คำอธิบายคุณภาพงานในแต่ละระดับ ต้องสามารถสังเกตได้ ประเมินได้
 - 6) ระบบการให้ระดับคะแนนต้องมีความเป็นไปได้ คำอธิบายต้องแยกจากกัน
- อย่างชัดเจนทุกฝ่ายควรมีส่วนร่วมในการกำหนดเกณฑ์การให้คะแนน

ตัวอย่างเกณฑ์การประเมินภาพรวม เกณฑ์การประเมินปัญหาคณิตศาสตร์ แบบปลายเปิด (สำนักทดสอบทางการศึกษา. 2546: 19; อ้างอิงจากรายงานของ California CAP วิชาคณิตศาสตร์ 1989)

คำตอบแสดงถึงมีศักยภาพ	
6 ดีเยี่ยม	ได้คำตอบที่สมบูรณ์ ชัดเจน การดำเนินการคำตอบมีความสัมพันธ์เกี่ยวเนื่องกันอย่างชัดเจนแจ่มแจ้ง การอธิบายสละสลวย มีแผนภูมิแสดงให้เห็นชัดเจนและเข้าใจง่าย อธิบายคำตอบอย่างมีประสิทธิภาพ แสดงถึงความเข้าใจแนวคิดทางคณิตศาสตร์และกระบวนการแก้ปัญหา ซึ่งให้เห็นถึงองค์ประกอบที่สำคัญของปัญหาครบถ้วนอาจแสดงตัวอย่างประกอบที่ช่วยสนับสนุนข้อโต้แย้งอย่างชัดเจน
5 มีความสามารถ	คำตอบสมบูรณ์ อธิบายเหตุผลชัดเจน อาจแสดงแผนภูมิประกอบคำตอบ อธิบายคำตอบอย่างมีประสิทธิภาพ แสดงถึงความเข้าใจแนวคิดและกระบวนการแก้ปัญหา ซึ่งให้เห็นถึงองค์ประกอบที่สำคัญของปัญหาครบถ้วนเสนอหลักฐานสนับสนุนข้อโต้แย้ง
คำถามในระดับพอใจ	
4 มีข้อผิดพลาดเล็กน้อยแต่ยังใช้ได้	คำตอบสมบูรณ์ในระดับพอใจโดยอธิบายคำตอบไม่ชัดเจนนัก การนำเสนอข้อโต้แย้งไม่สมบูรณ์ แผนภูมิประกอบอาจไม่สมบูรณ์หรือไม่ชัดเจน แต่แสดงความเข้าใจ ในพื้นฐานของแนวคิดทางคณิตศาสตร์และใช้แนวคิดนั้นอย่างมีประสิทธิภาพ
3 มีข้อผิดพลาดมากอยู่ในระดับค่อนข้างพอใช้	เริ่มแก้ปัญหา ได้อย่างเหมาะสม แต่ไม่สามารถแก้ปัญหาได้อย่างครบถ้วนหรือข้ามบางตอนที่สำคัญของปัญหา อาจไม่สามารถแสดงความเข้าใจในแนวคิดทางคณิตศาสตร์และกระบวนการแก้ปัญหาได้อย่างครบถ้วน อาจคำนวณผิดพลาดในประเด็นที่สำคัญ หรือใช้ศัพท์ทางคณิตศาสตร์ผิดพลาด คำตอบอาจแสดงถึงความไม่เหมาะสมในการใช้ยุทธศาสตร์ในการแก้ปัญหา
คำตอบในระดับยังไม่พอใจ	
2 เริ่มแก้ปัญหาแต่ยังไม่สำเร็จ	คำอธิบายที่เขียนไม่สามารถเข้าใจได้ แผนภูมิประกอบไม่ชัดเจน แสดงให้เห็นถึงความไม่เข้าใจในปัญหาอาจแสดงการคำนวณไม่ถูกต้องเป็นส่วนใหญ่
1 ไม่สามารถเริ่มแก้ปัญหาได้อย่างมีประสิทธิภาพ	คำตอบไม่สะท้อนให้เห็นถึงความเข้าใจในปัญหา อธิบายปัญหาไม่ถูก ไม่สามารถใช้ข้อมูลที่กำหนดให้ได้อย่างเหมาะสม
0 ไม่แสดงความพยายามในการหาคำตอบ/ไม่ตอบ	

ตัวอย่างเกณฑ์การประเมินแยกประเด็น (สำนักทดสอบทางการศึกษา. 2546: 26; อ้างอิงจากรายงานของ California CAP วิชาคณิตศาสตร์ 1989)

ความเข้าใจในคณิตศาสตร์	
ระดับ	คำอธิบายคุณภาพ
6	แสดงความเข้าใจในแนวคิดที่อยู่ภายใต้หลักคณิตศาสตร์ที่เกี่ยวข้องในขั้นสูง สามารถใช้หลักการสูตรเทคนิคและเครื่องมืออย่างเจาะจงในระดับมีความเชี่ยวชาญ สามารถจับประเด็นสำคัญของปัญหาได้ โดยประยุกต์หลักการทางคณิตศาสตร์อย่างเหมาะสม และใช้เครื่องมืออย่างมีประสิทธิภาพในการแก้ปัญหา ผลงานแสดงให้เห็นว่านักเรียนมีความสามารถในการดำเนินการไว้อย่างมีประสิทธิภาพ เฉพาะและสามารถสรุปได้จากรูปแบบ
5	แสดงความเข้าใจในแนวคิดที่อยู่ภายใต้หลักคณิตศาสตร์ที่เกี่ยวข้องได้ครบถ้วน ใช้หลักการสูตร เทคนิคและเครื่องมืออย่างเหมาะสมและครอบคลุม สามารถจับประเด็นสำคัญของปัญหาได้โดยประยุกต์หลักการทางคณิตศาสตร์อย่างเหมาะสมในการแก้ปัญหา
4	แสดงความเข้าใจในแนวคิดที่อยู่ภายใต้หลักทางคณิตศาสตร์ที่เกี่ยวข้องได้อย่างเหมาะสม ผลงานแสดงให้เห็นถึงความสามารถในการควบคุมการใช้ระบบการคำนวณ สูตร แนวคิดทางคณิตศาสตร์ เป็นต้น แต่มีข้อบกพร่องเกี่ยวกับความสามารถในการใช้เทคนิคเฉพาะที่ช่วยในการแก้ปัญหาย่อยบ้าง ที่เกี่ยวกับรูปแบบ สูตร หรือการลงสรุป อาจมีจุดอ่อนหรือขาดแนวคิดที่จำเป็นต้องใช้ในการแก้ปัญหา แต่ความบกพร่องนี้เป็นเพียงข้อบกพร่องเล็กน้อย
3	แสดงความเข้าใจในแนวคิดที่อยู่ภายใต้หลักคณิตศาสตร์ที่เกี่ยวข้องได้ค่อนข้างจำกัด หลักการทั่วไปหรือสูตรทางคณิตศาสตร์ที่เกี่ยวข้องมักจะนำมาใช้อย่างไม่เข้าใจและไม่ถูกต้อง ผลงานของผู้เรียนอาจแสดงความเหมาะสมในการแก้ปัญหา แต่มักไม่ใช้การสรุปอย่างผิวเผินหรือตื้นๆ และ ใช้กลยุทธ์เฉพาะในการแก้ปัญหาให้ยุ่งยากโดยไม่จำเป็น ทำให้มีความผิดพลาดเกิดขึ้น
2	แสดงถึงความไม่เข้าใจในแนวคิดที่อยู่ภายใต้หลักคณิตศาสตร์ที่เกี่ยวข้องกับปัญหา มีความบกพร่องในการใช้หลักการทั่วไป สูตร ระบบทางคณิตศาสตร์ การคำนวณ หรือข้อมูลที่กำหนดให้หรือนำมาใช้น้อยมาก
1	คำตอบไม่พอเพียงที่ใช้ตัดสินผู้เรียนด้านความรู้ความสามารถทางคณิตศาสตร์ที่เกี่ยวข้องกับปัญหาได้ (โดยเฉพาะอย่างยิ่งความผิดพลาดหรือความบกพร่องในการแก้ปัญหา)

2.6 งานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ

ฐลีพร ดัดงาม (2552) ศึกษาการพัฒนาเกณฑ์การตรวจให้คะแนนการเขียนเรียงความของนักเรียนช่วงชั้นที่ 2 จุดมุ่งหมายเพื่อพัฒนาเกณฑ์การตรวจให้คะแนนการเขียนเรียงความกลุ่มตัวอย่างที่ใช้ในการร่วมพิจารณาเกณฑ์การตรวจให้คะแนน ได้แก่ ผู้เชี่ยวชาญด้านภาษาไทย จำนวน 30 คน และผู้เชี่ยวชาญด้านการวัดผลการศึกษา จำนวน 2 คน และกลุ่มตัวอย่างที่ใช้ในการทดลองเกณฑ์การตรวจให้คะแนนเรียงความเป็นนักเรียนช่วงชั้นที่ 2 โรงเรียนอัสสัมชัญแผนกประถมที่ได้มาจากการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) จำนวน 60 คน เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ เกณฑ์การตรวจให้คะแนนการเขียนแบบแยกองค์ประกอบ สถิติที่ใช้ในการวิจัย ได้แก่ สถิติพื้นฐาน ค่าความเชื่อมั่นของคะแนนที่ได้จากการตรวจให้คะแนนการเขียนเรียงความ และค่าความเชื่อมั่นของผู้ตรวจโดยมีผู้ตรวจจำนวนสองและสามคน ผลการวิจัย พบว่า เกณฑ์การตรวจให้คะแนนการเขียนแบบแยกองค์ประกอบ มีความสอดคล้องระหว่างประเด็นการตรวจกับเกณฑ์การตรวจให้คะแนน (IOC) อยู่ระหว่าง 0.75-1.00 ค่าความเชื่อมั่นของคะแนนที่ได้จากการตรวจให้คะแนนการเขียนเรียงความ (α – coefficient) 0.962 ค่าสัมประสิทธิ์ความเชื่อมั่นของเกณฑ์การตรวจให้คะแนนการเขียนเรียงความ r_{tt} – Hoyt's ANNOVA Procedure 0.962 ค่าเฉลี่ย ($\bar{R}$) ค่าความเชื่อมั่นกรณีผู้ตรวจ 2 คน มีค่า 0.915 และกรณีผู้ตรวจ 3 คน มีค่า 0.917

อรจรรย์ ชูช่วยสุวรรณ (2552) ศึกษาการพัฒนาแบบทดสอบวัดความสามารถในการเชื่อมโยงทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5 จุดมุ่งหมายเพื่อสร้างแบบทดสอบวัดความสามารถในการเชื่อมโยงทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5 และเพื่อหาคุณภาพของแบบทดสอบกลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนที่กำลังศึกษาอยู่ในภาคเรียนที่ 2 ปีการศึกษา 2551 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษากรุงเทพมหานคร เขต 1 จำนวน 374 คน ซึ่งได้มาโดยการสุ่มแบบสองขั้นตอน (Two-Stage Random Sampling) เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดความสามารถในการเชื่อมโยงทางคณิตศาสตร์ที่วัดเฉพาะความสามารถในการเชื่อมโยงภายในวิชา ที่ผู้วิจัยสร้างขึ้นมีลักษณะเป็นแบบทดสอบอัตนัย โดยใช้เกณฑ์การตรวจให้คะแนนแบบภาพรวมที่ผู้วิจัยสร้างขึ้นปรับปรุงมาจากเกณฑ์การให้คะแนนของกรมวิชาการ ผลการวิจัย พบว่า แบบทดสอบวัดความสามารถในการเชื่อมโยงทางคณิตศาสตร์ ที่วัดเฉพาะความสามารถในการเชื่อมโยงภายในวิชา มีดัชนีความยากตั้งแต่ 0.43 ถึง 0.61 และมีดัชนีอำนาจจำแนกตั้งแต่ 0.55 ถึง 0.71 มีความเที่ยงตรงเชิงโครงสร้าง ตั้งแต่ 0.817 ถึง 0.903 มีค่าความเชื่อมั่นของแบบทดสอบ 0.821 คุณภาพด้านความเชื่อมั่นของผู้ตรวจให้คะแนนระหว่างผู้ตรวจ 2 คน และผู้ตรวจ 3 คน มีดัชนีความสอดคล้อง 0.929 และ 0.934 ตามลำดับ และจากศึกษาความสามารถในการเชื่อมโยงภายในวิชาคณิตศาสตร์ของนักเรียน พบว่า นักเรียนมีความสามารถในการเชื่อมโยงภายในวิชาคณิตศาสตร์อยู่ในระดับผ่านเกณฑ์ คิดเป็นร้อยละ 71.39

บุษวรรษ์ แสนปลื้ม (2556) ศึกษาการใช้วิธีการตรวจคุณลักษณะ และสัดส่วนจำนวน ผู้ตรวจให้คะแนนที่มีผลต่อความเที่ยงตรงของการวัดความสามารถในการเขียนของนักเรียนชั้น ประถมศึกษาปีที่ 3 การวิจัยในครั้งนี้ มีจุดมุ่งหมายเพื่อสร้างและพัฒนาคุณภาพของเครื่องมือวัด ความสามารถในการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 3 เพื่อศึกษาคุณลักษณะของผู้ตรวจ ให้คะแนน คือ ความเข้มงวด/ใจดี ความลำเอียงด้านเพศ และการทำหน้าที่ต่างกันของผู้ให้ คะแนนเมื่อเวลาผ่านไป และเพื่อตรวจสอบความเที่ยงตรงของความสามารถในการเขียนของ นักเรียน เครื่องมือที่ใช้ในการวิจัย คือ วิธีการให้คะแนนการเขียนเรียงความ 3 รูปแบบ และ ข้อสอบการเขียนจดหมายและวิธีการตรวจให้คะแนนการเขียนจดหมาย 3 รูปแบบ กลุ่มตัวอย่าง ที่ใช้ในการวิจัย คือ นักเรียนชั้นประถมศึกษาปีที่ 3 จำนวน 50 คน ที่ได้มาจากการสุ่มแบบหลาย ขั้นตอน และนักศึกษาคณะครุศาสตร์ ที่เรียนวิชาเอกภาษาไทยชั้นปีที่ 4 มหาวิทยาลัยราชภัฏ 3 แห่ง ที่ได้มาจากการสุ่มอย่างง่าย จำนวน 130 คน ผลการวิจัยสรุปได้ว่า

1. เครื่องมือวัดความสามารถในการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 3 มี คุณภาพตามทฤษฎีมาตรฐานเดิม ทฤษฎีการสรุปอ้างอิงและทฤษฎีการตอบสนองข้อสอบ

2. คุณลักษณะของผู้ตรวจให้คะแนนการวัดความสามารถในการเขียนของนักเรียน ชั้นประถมศึกษาปีที่ 3 พบว่า 1) ความเข้มงวด/ใจดี 1.1) แบบแยกองค์ประกอบ เข้มงวด 17 คน เป็นกลาง 89 คน ใจดี 9 คน 1.2) แบบรวมองค์ประกอบ เข้มงวด 10 คน เป็นกลาง 83 คน ใจดี 15 คน 1.3) แบบผสมผสาน เข้มงวด 17 คน เป็นกลาง 79 คน ใจดี 16 คน 2) การระบุ คุณลักษณะระบุความเข้มงวด/ใจดี มีความสอดคล้องกัน ความลำเอียงด้านเพศ 2.1) แบบแยก องค์ประกอบ เข้าข้างเพศชาย 8 คนเป็นกลาง 92 คน เพศหญิง 15 คน 2.2) แบบรวม องค์ประกอบ เข้าข้างเพศชาย 5 คน เป็นกลาง 97 คนเพศหญิง 8 คน 2.3) แบบผสมผสาน เข้าข้างเพศชาย 10 คน เป็นกลาง 91 คน เพศหญิง 11 คน การระบุคุณลักษณะความลำเอียง ด้านเพศมีความสอดคล้องกัน 3) การทำหน้าที่ต่างกันของผู้ตรวจให้คะแนนเมื่อเวลาผ่านไป 3.1) แบบแยกองค์ประกอบไม่คงที่ 83 คนคงที่ 33 คน 3.2) แบบรวมองค์ประกอบไม่คงที่ 41 คน คงที่ 68 คน 3.3) แบบผสมผสานไม่คงที่ 71 คน คงที่ 41 คน การระบุคุณลักษณะการทำหน้าที่ ต่างกันของผู้ตรวจให้คะแนนเมื่อเวลาผ่านไปไม่สอดคล้องกัน

3. ผลการหาความเที่ยงตรงของการวัดความสามารถในการเขียนของนักเรียนชั้น ประถมศึกษาปีที่ 3 มีความเที่ยงตรงเชิงโครงสร้างแบบลักษณะหลากหลาย และมีความ เที่ยงตรงตามสภาพ โดยที่ 1) เมื่อผู้ตรวจมีคุณลักษณะลำเอียง การใช้วิธีการตรวจให้คะแนน แบบแยกองค์ประกอบในสัดส่วน จำนวนผู้ตรวจให้คะแนน 3 คน ผู้ตรวจให้คะแนนมีความ ลำเอียงด้านเพศเข้าข้างเพศชาย 1 คน และลำเอียงเข้าข้างเพศหญิง 2 คน มีความเที่ยงตรงตาม สภาพสูงที่สุด และ 2) เมื่อผู้ตรวจมีคุณลักษณะเป็นกลาง การใช้วิธีการตรวจให้คะแนนแบบแยก องค์ประกอบ ในสัดส่วนจำนวนผู้ตรวจให้คะแนน 3 คน โดยผู้ตรวจให้คะแนนเป็นกลาง ไม่เข้มงวด/ใจดี มีความความเที่ยงตรงตามสภาพสูงที่สุด

ปวีณา ปีอาทิตย์ (2545) ศึกษาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนที่ต่างกัน การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) เปรียบเทียบความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนจากผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และผู้ประเมินจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ และ 2) เพื่อหาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กลุ่มตัวอย่างประกอบด้วย ผลงานการเขียนวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 จำนวน 117 ผลงาน และผู้ประเมินจำนวน 8 คน ที่ใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ การวิจัยนี้คำนวณค่าความเที่ยงจากค่าสัมประสิทธิ์การสรุปอ้างอิงด้วยโปรแกรมสำเร็จรูป GENOVA ผลการวิจัยพบว่า

1. ภายใต้งานจำนวน 117 ชิ้น ผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.71484 และเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.81410

2. ถ้าต้องการใช้ผู้ประเมิน 1 คน และกำหนดให้ค่าสัมประสิทธิ์การสรุปอ้างอิงของเกณฑ์การให้คะแนนแบบภาพรวมเป็น 0.2, 0.4, 0.6 และ 0.8 ใช้จำนวนชิ้นงานอย่างน้อย 4, 8, 16 และ 28 ชิ้นตามลำดับ ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2 และ 0.4 ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.6 ใช้จำนวนชิ้นงานอย่างน้อย 8 ชิ้น และเมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.8 ใช้จำนวนชิ้นงานอย่างน้อย 12 ชิ้น ทั้งนี้ขึ้นอยู่กับผู้ปฏิบัติว่าจะเลือกใช้เกณฑ์การให้คะแนนให้เหมาะสมกับการเรียนการสอน

ชูซานเน เลน และคณะ (ปวีณา ปีอาทิตย์, 2545 : 30; อ้างอิงจาก Suzanne Lane; et al. 1996) ศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิงและความตรงของการประเมินผลงานของนักเรียนในวิชาคณิตศาสตร์ที่พัฒนาจากเครื่องมือประเมินความรู้แบบ QCAI (QUASAR Cognitive Assessment Instrument) เป็นชุดของข้อสอบแบบปลายเปิด โดยเน้นวิธีการตรวจให้คะแนนแบบประเมินภาพรวม ใช้ทฤษฎีการสรุปอ้างอิงตรวจสอบความสอดคล้องภายในข้อสอบ และภายในผู้ตรวจ โดยมีแบบแผน 3 รูปแบบ คือ 1) $P \times T$ ผู้สอบทุกคนทำข้อสอบทุกข้อ 2) $P \times T \times R$ ผู้สอบทุกคนทำข้อสอบทุกข้อ และผู้ตรวจแต่ละคนตรวจข้อสอบทุกข้อ 3) $(P:S) \times T$ ผู้สอบที่แบ่งในโรงเรียนทำข้อสอบทุกข้อ ผลการศึกษา พบว่า ความคลาดเคลื่อนเนื่องจากผู้ตรวจมีน้อยมาก เนื่องจากการให้รายละเอียดเกี่ยวกับเกณฑ์การให้คะแนนแบบรูปรีด และตัวอย่างการตอบของนักเรียน รวมทั้งการฝึกฝนวิธีการตรวจให้คะแนนกับผู้ตรวจถูกต้องเพียงพอ มีความแตกต่างที่ต้องคำนึงในการทำข้อสอบของนักเรียนในแต่ละข้อ ค่าความเที่ยงที่ได้จากการศึกษาในขั้นตอนการตัดสินใจการสรุปอ้างอิง พบว่ามีค่าสัมประสิทธิ์การสรุปอ้างอิงอยู่ระหว่าง 0.80 – 0.97 โดยขึ้นอยู่กับชุดของข้อสอบและระดับชั้น

อังเดร จอห์นสัน และกันนิวรา สเวนบี (Anders Jonsson; & Gunilla Svingby. 2007: 130 - 144) ศึกษาผลการใช้เกณฑ์การให้คะแนนด้านความเชื่อมั่น ความเที่ยงตรง และประโยชน์ของการใช้เกณฑ์การให้คะแนน ในการประเมินผลการปฏิบัติงานได้รับการเสนอ อาทิ เช่น ความสอดคล้องที่เพิ่มขึ้นของการให้คะแนน ความเป็นไปได้ความสะดวกในการตัดสินใจ ถูกต้องของสมรรถนะที่ซับซ้อนและการส่งเสริมการเรียนรู้ บทความนี้ได้มีการค้นคว้าหลักฐาน ข้อความกล่าวอ้างเหล่านี้สามารถพบได้ในงานวิจัยต่างๆ ฐานข้อมูลหลายคนค้นหาการวิจัยเชิง ประจักษ์เกี่ยวกับรูปрик รวมทั้งหมดจำนวน 75 บทความ สำหรับการสังเคราะห์นี้ สรุป ผลการวิจัย (1) เกณฑ์การให้คะแนนมีความน่าเชื่อถือของการประเมินผลการปฏิบัติงานสามารถ เพิ่มโดยใช้รูปริกโดยเฉพาะอย่างยิ่งถ้าพวกเขามีการวิเคราะห์หัวข้อที่เฉพาะเจาะจง และเสริม ด้วยตัวอย่างและ / หรือการฝึกอบรมผู้ประเมิน (2) รูปริกไม่อำนวยความสะดวกการตัดสินใจ ถูกต้องของการประเมินผลการปฏิบัติงาน อย่างไรก็ตามการประเมินที่ถูกต้องจะได้รับการ อำนวยความสะดวกโดยใช้กรอบที่ครอบคลุมมากขึ้นของความถูกต้องเที่ยงตรง (3) ดูเหมือน รูปริกจะมีศักยภาพในการเรียนรู้การส่งเสริมและ / หรือปรับปรุงการสอนเหตุผลหลักสำหรับการ ศักยภาพนี้อยู่ในข้อเท็จจริงที่ว่ารูปริกทำให้ความคาดหวังและเกณฑ์อย่างชัดเจนซึ่งจะอำนวยความสะดวกให้ข้อเสนอแนะและการประเมินตนเอง

คราเทเลน กรีนเบิร์ก (Kathleen P. Greenberg. 2015: 211 - 217) ศึกษาการใช้ รูปริกในการประเมินตามสภาพจริง เพื่อช่วยให้ผู้เรียนปรับปรุงทักษะกระบวนการทาง วิทยาศาสตร์ ด้านทักษะการเขียน ดังนั้นวัตถุประสงค์ของการวิจัยนี้คือ ประการแรกการใช้รูปริก เพื่อคะแนนคำตอบของนักเรียน ประการที่สองไม่ใช้รูปริกเพื่อให้คะแนน แต่เพื่อนร่วมชั้นจะ เขียนผลการวัดอีกครั้ง เพื่อให้ให้นักเรียนจะได้รับโอกาสในการแก้ไขตนเอง คำถามเหล่านี้ได้รับการ แยกกับกลุ่มที่แตกต่างกันของนักเรียน ขั้นตอนและผลของแต่ละการจัดการอธิบายไว้ ด้านล่างในส่วนชื่อ “การใช้รูปริกกับการไม่ใช้รูปริก” และ “การเขียนใหม่หลังจากการวัดผลของ เพื่อนร่วมชั้น” การอภิปรายทั่วไปของทั้งสองชุดของผลการทดลอง ดังต่อไปนี้ การจัดการที่ 1 : การใช้รูปริกกับการไม่ใช้รูปริก กลุ่มตัวอย่างนักเรียนจำนวน 78 คน ที่ใช้รูปริก และนักเรียน จำนวน 68 คนที่ไม่ใช้รูปริก ผลการวิเคราะห์ข้อมูลการทดสอบความเป็นอิสระ t-test พบว่า นักเรียนที่ใช้รูปริกในการประเมินมีค่าเฉลี่ยสูงกว่านักเรียนที่ไม่ใช้รูปริกในการประเมินอย่างมี นัยสำคัญทางสถิติ ค่า $t = 3.30$, $p = .001$, $d = 0.610$ การจัดการ 2 : การเขียนใหม่หลังจากการ วัดผลของเพื่อนร่วมชั้น วัตถุประสงค์ของการศึกษารั้งนี้เพื่อตรวจสอบลายลักษณ์อักษรจะช่วยให้ นักเรียนปรับปรุงคุณภาพของเอกสารของตัวเองจากแผ่นคะแนนจะสะท้อนกลับไปยังนักเรียน โดยที่เพื่อนร่วมชั้นให้คะแนนรายงานเท่านั้น “ความคิดเห็นของพวกเขา” เป็นสิ่งที่พวกเขาได้ เรียนรู้จากกระบวนการของการจัดลำดับของคนอื่น ผลการวิเคราะห์ข้อมูล ค่า t-test จากการ เขียนใหม่ค่าเฉลี่ย 70.07 ส่วนเบี่ยงเบนมาตรฐาน 13.33 ต้นฉบับค่าเฉลี่ย 65.16 ส่วนเบี่ยงเบน มาตรฐาน 16.27 ซึ่งการเขียนใหม่สูงกว่าคะแนนต้นฉบับอย่างมีนัยสำคัญ $t = 3.023$, $p = .001$,

$d = .33$ โดยการใช้การวิเคราะห์การวัดซ้ำแบบแปรปรวน ANOVA เพื่อดำเนินการตรวจสอบผลกระทบหลักของเวลา $F = 23.38$, $p = .001$, Wilks' $\lambda = .209$, partial $\eta^2 = .791$, และ ส่วน, $F = 9.497$, $p = .004$, Wilks' $\lambda = .816$, partial $\eta^2 = .184$, แต่ไม่แตกต่างอย่างมีนัยสำคัญ $F = 0.591$, $p = .735$, Wilks' $\lambda = .912$, partial $\eta^2 = .184$, แสดงให้เห็นว่ามีการเปลี่ยนแปลงที่ส่งผลให้ในคะแนนมีการเปลี่ยนแปลงในคะแนนจากเดิมอย่างมีนัยสำคัญมากขึ้น

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกฎเกณฑ์การให้คะแนน (Scoring rubrics) ที่ได้กล่าวไว้ข้างต้น พบว่า ประเภทของเกณฑ์การตรวจให้คะแนนส่งผลให้เกิดความคลาดเคลื่อนต่อการให้คะแนนสำหรับผู้สอบ ดังนั้นผู้วิจัยจึงนำกฎเกณฑ์การให้คะแนน ทั้ง 2 รูปแบบ คือ การตรวจให้คะแนนแบบแยกองค์ประกอบ (Analytic Scoring Rubrics) และการตรวจให้คะแนนแบบรวมองค์ประกอบ (Holistic Scoring Rubrics) จากความหมายของกฎเกณฑ์การให้คะแนน ประเภทของกฎเกณฑ์การให้คะแนน หรือเกณฑ์การตรวจให้คะแนน (Rubrics) เป็นชุดของแนวทางในการตรวจให้คะแนนผลการปฏิบัติเรื่องใดเรื่องหนึ่งสำหรับใช้ประเมินคุณภาพการปฏิบัติงานของผู้เรียน แนวทางในการตรวจให้คะแนนนั้นๆ อาจทำในรูปของมาตราส่วนประเมินค่า หรือแบบตรวจสอบรายการ ประกอบด้วยประเด็นต่างๆ (Criteria) ที่จะใช้พิจารณางานหนึ่งๆ และคำอธิบายระดับคุณภาพของแต่ละประเด็นประเมิน ซึ่งอาจเรียงลำดับตั้งแต่ดีเลิศไปจนถึงต้องปรับปรุง หรือให้เป็นระดับตัวเลขตั้งแต่มากที่สุด (เช่น 4) ไปจนถึงน้อยที่สุด (เช่น 0) ประเด็นประเมินอาจกำหนดเพิ่มเติมได้หลายข้อ คำอธิบายระดับคุณภาพควรอธิบายให้ชัดเจนที่สุด กระชับที่สุดเป็นคำอธิบายที่สามารถบอกได้ว่าทำไมจึงต้องดีเลิศ ดี ต้องปรับปรุง (บุญเชิด ภิญญอนันตพงษ์. 2544: 90) ประเภทของกฎเกณฑ์การให้คะแนนแบ่งออกเป็น วิธีการตรวจให้คะแนนเฉพาะ (Specific Rubric) เป็นแนวทางในการตรวจให้คะแนนเฉพาะของคำถามแบบปลายเปิดในแต่ละข้อสำหรับวิธีการตรวจให้คะแนนทั่วไป (General Rubric) จะเป็นตัวแบบของวิธีการตรวจให้คะแนน เพื่อใช้พิจารณาเกณฑ์เฉพาะรายละเอียดจะเปลี่ยนแปลงสำหรับระดับคะแนนที่แตกต่างกัน ซึ่งวิธีการตรวจให้คะแนนจะเน้นการแสดงความสามารถของนักเรียนในด้านความเข้าใจในความคิดรวบยอด ความรู้เกี่ยวกับกระบวนการหรือขั้นตอน ทักษะในการแก้ปัญหา การใช้เทคโนโลยีใหม่ๆ ในการประมวลผลข้อมูล และการสื่อความหมายกับผู้อื่น (ชัยฤทธิ์ ศิลาเดช. 2544: 29) กล่าวโดยสรุป วิธีการตรวจให้คะแนน หรือเกณฑ์การตรวจให้คะแนน (Rubrics) คือ ชุดของแนวทางในการตรวจให้คะแนนผลการปฏิบัติเรื่องใดเรื่องหนึ่ง สำหรับใช้ประเมินคุณภาพการปฏิบัติงานของผู้เรียน ประกอบด้วยประเด็นต่างๆ (Criteria) ที่จะใช้พิจารณางานหนึ่งๆ และคำอธิบายระดับคุณภาพของแต่ละประเด็นประเมิน ซึ่งเกณฑ์การตรวจให้คะแนนจะเน้นการแสดงความสามารถของนักเรียนในด้านความเข้าใจในความคิดรวบยอด ความรู้เกี่ยวกับกระบวนการหรือขั้นตอน ทักษะในการแก้ปัญหาการใช้เทคโนโลยีใหม่ๆ ในการประมวลผลข้อมูล และการสื่อความหมายกับผู้อื่น

ตอนที่ 3 แนวคิดและทฤษฎีที่เกี่ยวกับทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory)

ศิริชัย กาญจนวาสี. (2555: 11) การศึกษาความเที่ยงของแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิมที่ยังใช้กันอยู่นั้น เป็นการประมาณค่าความเที่ยงของแบบสอบที่ใช้กับกลุ่มบุคคล ภายใต้เงื่อนไขของการทดสอบที่เฉพาะ ได้แก่ การทดสอบที่ผู้สอบทุกคนต้องทำข้อสอบเหมือนกันทุกข้อ ผลการตอบได้รับการให้คะแนนโดยผู้ตรวจคนเดียว และแบบสอบถามที่ใช้ตั้งอยู่บนข้อตกลงเบื้องต้นของความเป็นคู่ขนานระหว่างแบบสอบ แต่ถ้าสถานการณ์ของการทดสอบแตกต่างกันไป เทคนิคการประมาณค่าความเที่ยงตามทฤษฎีการทดสอบแบบดั้งเดิมไม่สามารถนำมาประยุกต์ใช้ได้ เช่น กรณีของการทดสอบที่มีผู้ตรวจหลายคนให้คะแนนข้อสอบ ความเรียงของผู้สอบทุกคนหรือกรณีของการทดสอบที่ใช้ผู้ตรวจหลายคนคะแนนผลการตอบแบบสอบที่มีความยาวต่างกัน หรือกรณีที่ผู้ตรวจหลายคนให้คะแนนผลการตอบแบบสอบที่มีรูปแบบความยาวและจำนวนครั้งของการทดสอบแตกต่างกัน เป็นต้น นอกจากนี้โมเดลการวัดของทฤษฎีการทดสอบแบบดั้งเดิม ถือว่าเป็นความคาดเคลื่อนหรือความเคลื่อนของการวัดเป็นคะแนนรวมความคาดเคลื่อนทุกแหล่งเข้าด้วยกัน และเป็นความคาดเคลื่อนรวมอันเดียวกันที่ไม่สามารถแบ่งแยกได้ (Unique error) จึงเป็นข้อจำกัดทางทฤษฎีที่ไม่สามารถศึกษารายละเอียดของแหล่งความคาดเคลื่อนของการวัดในสถานการณ์หรือเงื่อนไขการวัดต่างๆ ได้

3.1 ความเป็นมาทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

ศิริชัย กาญจนวาสี. (2555: 11 - 12) นักทฤษฎีการทดสอบหลายท่านพยายามศึกษาเชิงจำแนกแหล่งความคาดเคลื่อนของผลการวัด โดยใช้เทคนิคการวิเคราะห์ความแปรปรวน (ANOVA) การนำเทคนิคการวิเคราะห์ความแปรปรวน ซึ่งเป็นวิธีการทางสถิติมาใช้ประมาณค่าความเที่ยงของแบบสอบ และความคาดเคลื่อนของผลการวัด วิธีที่เป็นที่รู้จักกันดี ได้แก่ วิธีการวิเคราะห์ของฮอยท์ (Hoyt.1994) ในเวลาต่อมาครอนบาคและคณะ (Cronbach, Gleser; & Rajaratnam.1963) ได้เสนอแนวคิดทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดหรือ G-Theory ขึ้นเป็นครั้งแรกเพื่อประมาณค่าความเที่ยงทั่วไปของแบบสอบในสถานการณ์ของการทดสอบหรือเงื่อนไขการวัดแบบต่างๆ นอกจากนี้ราชารัตนัม ครอนบาค และเกลสเซอร์ (Rajaratnam, Cronbach; & Gleser.1965,1972) ได้พัฒนาทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดให้มีรูปแบบของโมเดล จึงทำให้แนวคิดและการประมาณค่าความเที่ยงตามทฤษฎีการทดสอบแบบเดิมได้ขยายกว้างไกลออกไปและมีความครอบคลุมเงื่อนไขการวัดหรือแหล่งความคาดเคลื่อนต่างๆ ได้ด้วย ซึ่งจะเป็นประโยชน์ต่อการพัฒนาแบบสอบหรือแบบวัดให้มีระดับความเที่ยงที่ต้องการ ภายใต้สถานการณ์การทดสอบที่ต้องการนำแบบสอบหรือแบบวัดไปใช้ พร้อมทั้งได้มีการประยุกต์ใช้ในการวิเคราะห์ความตรงทั่วไปได้อีกด้วย

3.2 แนวคิดสำคัญและข้อตกลงเบื้องต้นของทฤษฎีการสรุปอ้างอิงของความน่าเชื่อถือของผลการวัด

นอกจากนี้ครอนบาค และคณะ (ศิริชัย กาญจนวาสี. 2555: 12 อ้างอิงจาก Cronbach; et al.1972) ได้เสนอทฤษฎีสำหรับการวิเคราะห์ความน่าเชื่อถือของผลการวัดพฤติกรรม (dependability of behavioral measurement) สำหรับสถานการณ์ของการวัดผลลักษณะต่างๆ ทั้งด้านจิตวิทยาและการศึกษาซึ่งต่อมารู้จักกันในชื่อของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability theory) หรือ G theory ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดเป็นทฤษฎีที่ขยายแนวคิดจากทฤษฎีการทดสอบแบบดั้งเดิม (classical test theory) ซึ่งเป็นทฤษฎีที่ให้ความสำคัญกับการตรวจสอบและระบุถึงปริมาณของแหล่งความคลาดเคลื่อนจากกระบวนการวัดที่ส่งผลให้เกิดความไม่คงที่ของคะแนนสังเกตได้ (inconsistencies in observed scores) (สุวิมล กฤษณกุล. 2558: 6 อ้างอิงจาก Brennan. 2011: 2)

จากโมเดลการวัดตามแนวคิดทฤษฎีการทดสอบแบบดั้งเดิม คือ

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

The diagram shows the equation $\sigma_X^2 = \sigma_T^2 + \sigma_E^2$ with lines connecting σ_T^2 to σ_S^2 and σ_E^2 . Below σ_S^2 is the label 'Systematic Error variance' and below σ_e^2 is the label 'Random Error variance'.

โมเดลการวัดดังกล่าวมีข้อควรพิจารณา 2 ประการ คือ

1. คะแนนจริงและคะแนนคลาดเคลื่อนแบบสุ่มเป็นตัวแปรที่ไม่สามารถสังเกตได้โดยตรง หากต้องการใช้โมเดลนี้จำเป็นต้องมีข้อตกลงเบื้องต้นเพิ่มเติม คือ

1.1 หากระบุค่าคะแนนจริงเหมือนกับเป็นค่าคาดหวังของคะแนนสังเกตได้ จะเป็นเงื่อนไขให้ค่าคาดหวังของคะแนนคลาดเคลื่อนเป็นศูนย์

1.2 หากระบุค่าคาดหวังของคะแนนคลาดเคลื่อนเหมือนกับมีค่าเป็นศูนย์ จะเป็นเงื่อนไขให้คะแนนจริงเป็นค่าคาดหวังของคะแนนสังเกตได้

ข้อสังเกตจากข้อตกลงเบื้องต้นเพิ่มเติม 1.1 และ 1.2 คือ ผลที่เกิดขึ้นจากข้อตกลงเบื้องต้นเพิ่มเติมทั้ง 2 ข้อได้ผลเหมือนกัน คะแนนจริงและคะแนนคลาดเคลื่อนเกี่ยวข้องกัน

2. คะแนนจริงเป็นคะแนนที่ไม่สามารถทราบค่าที่แท้จริงได้

การศึกษาความน่าเชื่อถือของผลการวัดตามแนว G-Theory อยู่บนพื้นฐานข้อตกลงเบื้องต้น ดังนี้

ศิริชัย กาญจนวาสี (2555: 14) ได้อธิบายว่าทฤษฎีการทดสอบแบบดั้งเดิมมุ่งที่จะอธิบายความไม่คงเส้นคงวาของการวัดด้วยการพิจารณาจากแหล่งความคลาดเคลื่อนเพียงแหล่งเดียวสำหรับการทดสอบในแต่ละครั้ง เช่น

1. ความคลาดเคลื่อนมาจากการทิ้งช่วงระยะเวลาของการสอบซ้ำจะประมาณค่าความเชื่อมั่นด้วยวิธีทดสอบซ้ำ (test-retest method)

2. ความคลาดเคลื่อนมาจากชุดของแบบทดสอบคู่ขนานจะประมาณค่าความเชื่อมั่นด้วยวิธีแบบทดสอบสมมูล (equivalence forms method) และ

3. ความคลาดเคลื่อนมาจากความยาวของแบบทดสอบจะประมาณค่าความเชื่อมั่นด้วยวิธีแบ่งครึ่งข้อสอบและปรับเป็นทั้งฉบับด้วยสูตรของสเปียร์แมน (Spearman formula) เป็นต้น ในขณะที่ G theory ได้พิจารณาถึงแหล่งความคลาดเคลื่อนพหุ (multiple sources of error) ที่สามารถวิเคราะห์ความคลาดเคลื่อนจากหลายแหล่งพร้อมกันไปโดยใช้เทคนิคการวิเคราะห์ความแปรปรวน (ANOVA)

ศิริชัย กาญจนวาสี (2555: 13 - 15) กล่าวว่า การศึกษาความน่าเชื่อถือของผลการวัดตามแนว G theory อยู่บนพื้นฐานของข้อตกลงเบื้องต้น ดังนี้

1. คุณลักษณะที่มุ่งวัดของบุคคลไม่ว่าจะเป็นความรู้ ทักษะ หรือคุณลักษณะอื่นๆ ซึ่งเป็นเป้าหมายของการวัด เป็นค่าที่อยู่ในสภาวะคงที่ (steady state)

2. ผู้สอบคนเดียวได้คะแนนต่างกันจากการวัดในแต่ละสถานการณ์หรือเงื่อนไขของการวัด เนื่องมาจากความคลาดเคลื่อนที่เป็นระบบอย่างน้อย 1 แหล่งโดยองค์ประกอบด้านวุฒิภาวะและการเรียนรู้ระหว่างการวัดไม่เป็นแหล่งความคลาดเคลื่อนของคะแนนที่ได้จากการวัด

3. เมื่อพิจารณาผู้สอบทั้งกลุ่ม ความแปรปรวนของคะแนนสังเกตได้ประกอบด้วยความแปรปรวนของคะแนนจริง ซึ่งเป็นความแตกต่างที่แท้จริงระหว่างบุคคล ความแปรปรวนของคะแนนความคลาดเคลื่อนที่เป็นระบบอย่างน้อย 1 แหล่งและความแปรปรวนของคะแนนความคลาดเคลื่อนสุ่ม

หลักการของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory)

สำหรับหลักการของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ประกอบด้วยความคลาดเคลื่อนมาตรฐานของการวัดตามแนวคิดของลอร์ดแนวคิดสำคัญ และข้อตกลงเบื้องต้น และแหล่งความคลาดเคลื่อนที่สามารถเกิดขึ้นในกระบวนการวัด ดังรายละเอียด ความคลาดเคลื่อนมาตรฐานของการวัดตามแนวคิดของลอร์ด (Lord's Standard Error of Measurement) การวิเคราะห์ตาม G theory สามารถวิเคราะห์ความคลาดเคลื่อนจากหลายแหล่งพร้อมกันไปนั้น เบรนนัน กล่าวว่า (สุวิมล กฤษฎะหาสน์. 2558: 7; อ้างอิงจาก Brennan. 2011: 4 - 5) หลักการของการวิเคราะห์ของ G theory ประยุกต์ใช้หลักการของการประมาณค่าความคลาดเคลื่อนมาตรฐานของการวัดตามแนวคิดของลอร์ด ค.ศ. 1955 และ 1957

ลอร์ดได้เสนอแนวคิดในการประมาณค่าความคลาดเคลื่อนมาตรฐานของการวัดด้วยการพิจารณาจากความคลาดเคลื่อนของคะแนนสังเกตเฉลี่ยของผู้สอบโดยที่คะแนนสังเกตนั้น เกิดจากการที่ผู้สอบทำข้อสอบ จำนวน k ข้อที่สุ่มแบบเป็นไปได้อันทั้งหมดจากเอกภพของข้อสอบ (universe of items)

$$\sigma(E^*) = \frac{\sqrt{\tau_p(1-\tau_p)}}{k}$$

หรือ $\hat{\sigma}(E^*) = \frac{\sqrt{X_p(1-X_p)}}{k-1}$ นำไปใช้ในทางปฏิบัติ

เมื่อ $\hat{\sigma}(E^*)$ แทน ความคลาดเคลื่อนมาตรฐานของการวัดตามแนวคิดของลอร์ด

$\hat{\sigma}(E^*)$ แทน ค่าที่ได้จากการประมาณค่าความคลาดเคลื่อนมาตรฐานของการวัดตามแนวคิดของลอร์ด

τ_p แทน คะแนนจริงของผู้สอบที่ได้จากคะแนนเฉลี่ยของผู้สอบ

k แทน จำนวนข้อสอบที่สุ่มแบบเป็นไปได้อันทั้งหมดจากเอกภพของข้อสอบ (universe of items)

X_p แทน คะแนนสังเกตเฉลี่ยของผู้สอบ

จะเห็นว่าแนวคิดความคลาดเคลื่อนมาตรฐานของการวัดของลอร์ดเป็นที่มาของ ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด สามารถสรุปที่มาของแนวคิดได้ 2 ประเด็น คือ

1. ลอร์ดใช้กระบวนการสุ่มเพื่อประมาณค่าความแปรปรวนของความคลาดเคลื่อน มากกว่าการใช้ความเป็นคู่ขนานของแบบทดสอบตามแนวคิดของสเปียร์แมนดังสูตร

$$\sigma(E) = \sigma(X)\sqrt{1 - \rho^2(X, T)}$$

เมื่อ $\sigma(E)$ แทน ความคลาดเคลื่อนมาตรฐานของการวัด

$\sigma(X)$ แทน ค่าเบี่ยงเบนมาตรฐาน ของคะแนนสังเกตได้

$\rho^2(X, T)$ แทน ความเชื่อมั่น (reliability) หรือร้อยละของความสามารถของ การอธิบายหรือทำนายคะแนนจริงจากคะแนนสังเกตได้

2. ลอร์ดใช้การออกแบบการวัดที่คำนึงถึงความแปรปรวนภายในบุคคล

จากสาระสำคัญดังกล่าวแสดงให้เห็นว่าทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของ ผลการวัดใช้แนวคิดของลอร์ดที่เน้นการประมาณค่าความแปรปรวนของความคลาดเคลื่อน ส่วนทฤษฎีทดสอบแบบดั้งเดิมเน้นที่ความเป็นคู่ขนานของแบบทดสอบ (parallelism assumption) ดังนั้น G theory จึงมีการออกแบบการวัดที่ต่างจากทฤษฎีทดสอบแบบดั้งเดิม กล่าวคือการออกแบบการวัดตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ประกอบด้วย 1) ระบุแหล่งความคลาดเคลื่อนที่เป็นระบบอย่างน้อย 1 แหล่ง ที่มีเรียกชื่อเฉพาะ ว่า เงื่อนไขของการวัด (facet) จำแนกเป็น เงื่อนไขของการวัด 1 เงื่อนไข (one facet design) และมากกว่า 1 เงื่อนไข (multifaceted designs) 2) นำเงื่อนไขการวัดที่ระบุไว้มาออกแบบการ วัด ให้สอดคล้องกับวัตถุประสงค์ของการศึกษาเชิงสรุปอ้างอิง (generalizability study) โดยที่ การศึกษาเชิงสรุปอ้างอิงเป็นการศึกษาและวิเคราะห์ความแปรปรวนของคะแนนจริงและคะแนน คลาดเคลื่อนจำแนกเป็นแบบไขว้(crossed design) แบบแฝงเต็มรูป (fully nested design) และ แบบแฝงบางส่วน (partially nested design) ดังนั้นผลการวิเคราะห์ที่ได้จากการออกแบบการวัด ขั้นต้น คือ ผลการศึกษาเชิงสรุปอ้างอิง ซึ่งผู้วิจัยสามารถนำไปใช้สำหรับวางแผนเพื่อ ตัดสินใจว่าสถานการณ์ของการวัดในลักษณะใดควรนำไปใช้ เพื่อให้เกิดความน่าเชื่อถือของผล การวัด ที่เรียกว่า การศึกษาเชิงตัดสินใจ (decision study)

คำศัพท์ที่เกี่ยวข้องกับ G-Theory

ศิริชัย กาญจนวาสี (2555: 15) G-Theory ได้จำแนกคำศัพท์ประชากรและเอกภพให้มีความหมายแตกต่างกัน ดังนี้

1. ประชากร หมายถึง สิ่งที่มีวัดทั้งหมด ในสถานการณ์ของการทดสอบทั่วไป สิ่งที่มีวัดมักได้แก่บุคคลหรือผู้ทำการสอบ

2. เอกภพ หมายถึง เงื่อนไขของการวัดที่สนใจทั้งหมด กลุ่มเงื่อนไขของการวัดเรียกว่า “ฟาเซท” (Facet) ซึ่งเป็นองค์ประกอบที่คาดว่าจะมีกระทบต่อความคาดเคลื่อนของผลการวัด เช่น ความยาวของข้อสอบ รูปแบบของข้อสอบ จำนวนครั้งของการทดสอบ จำนวนผู้ตรวจให้คะแนน เป็นต้น สำหรับเงื่อนไขของการวัดเป็นระดับฟาเซท (องค์ประกอบ) ของการวัด เช่น ฟาเซทของจำนวนผู้ตรวจ อาจกำหนดจำนวนระดับเป็น 1,2,3 คน ฟาเซทของความยาวของแบบสอบ อาจกำหนดระดับความยาวของแบบสอบเป็น 10,30,50 ข้อ เป็นต้น หลังจากการออกแบบจำนวนฟาเซท (องค์ประกอบ) และจำนวนเงื่อนไข (ระดับการวัด) ของแต่ละองค์ประกอบแล้ว การวัดที่ครอบคลุมเงื่อนไขที่สนใจทั้งหมดนั้นเรียกว่า เอกภพของการสรุปอ้างอิง (Universe of Generalization) ซึ่งเป็นเงื่อนไขของการวัดทั้งหมดที่เป็นเป้าหมายของการสรุปอ้างอิงคุณภาพของแบบสอบ เมื่อทำการเก็บรวบรวมข้อมูลการทดสอบของประชากรผู้ทำการสอบภายใต้เงื่อนไขการวัดที่สนใจทั้งหมด เรียกว่า เอกภพของค่าที่ได้จากการสังเกตทั้งหมด (Universe of Admissible Observation)

3. ฟาเซท ที่ต้องการศึกษาอาจเป็นองค์ประกอบสุ่ม (random) หรือองค์ประกอบเจาะจง (fix) ถ้าเงื่อนไขการวัดถูกเลือกมาอย่างเจาะจงจากองค์ประกอบที่ศึกษา แสดงว่าผู้ศึกษาทำการสรุปอ้างอิงความเที่ยงของแบบสอบ ไปยังองค์ประกอบเฉพาะระดับของเงื่อนไขที่เลือกศึกษาเท่านั้น แต่ถ้าเงื่อนไขการวัดได้รับการสุ่มเพื่อเป็นตัวแทนของคิประกอบที่ศึกษาแสดงว่าผู้ศึกษาสามารถทำการสรุปอ้างอิงความเที่ยงของแบบสอบ ไปยังระดับต่างขององค์ประกอบที่ศึกษาได้

4. การศึกษา G และการศึกษา D ทฤษฎีการสรุปอ้างอิงทางการทดสอบประกอบด้วยขั้นตอนการศึกษาที่สำคัญ 2 ขั้นตอน ได้แก่ การศึกษาสรุปอ้างอิงหรือการศึกษา G (Generalizability Study or G-Study) กับการศึกษาเชิงตัดสินใจหรือการศึกษา D (Decision Study or D-Study)

4.1 การศึกษา G (G-Study) เป็นการสรุปอ้างอิงผลจากการศึกษาตัวอย่างของการวัดตามเงื่อนไขที่สนใจ บรรยายความแปรปรวนของความคาดเคลื่อนจากแหล่งความคาดเคลื่อนต่างๆ เพื่อสรุปอ้างอิงไปยังเอกภพของการวัด

4.2 การศึกษา D (D-Study) เป็นการใช้ข้อมูลจากการศึกษา G ที่สอดคล้องกับจุดประสงค์เฉพาะของการตัดสินใจเลือกใช้แบบทดสอบในสถานการณ์ต่างๆของการวัด

จุดประสงค์ของการศึกษา G ต้องการประมาณค่าความแปรปรวนของคะแนนจริงและความแปรปรวนของคะแนนความคาดเคลื่อนจากแหล่งความคาดเคลื่อนต่างๆที่สนใจ และใช้เป็นข้อมูลสำหรับการวางแผนเพื่อตัดสินใจ ในการศึกษา D เกี่ยวกับค่าความเที่ยงของแบบสอบในสถานการณ์ของการวัดต่างๆ โดยกำหนดสถานการณ์หรือเงื่อนไขการวัดที่สนใจ แล้วคำนวณค่าความเที่ยง, เปรียบเทียบค่าความเที่ยง และตัดสินใจเลือกสถานการณ์ (เงื่อนไขการวัด) ที่เหมาะสม ซึ่งเมื่อทำการทดสอบแล้วจะได้ค่าความเที่ยงถึงระดับที่ต้องการ ดังนั้นการออกแบบ G-Study จึงควรครอบคลุมเงื่อนไขของการวัด ที่ต้องการตัดสินใจนำแบบสอบไปใช้ใน D-Study (ศิริชัย กาญจนวาสี. 2555: 16)

3.3 แหล่งความแปรปรวนและองค์ประกอบความแปรปรวนสำหรับการออกแบบการวัดกรณีออกแบบการวัดแบบไขว้ (Crossed design) กรณีการวัดที่มีหนึ่งองค์ประกอบแบบไขว้ (One – facet, $p \times i$ Design) คะแนนสังเกตของบุคคลในข้อสอบข้อหนึ่งได้รับอิทธิพลมาจากส่วนต่าง ๆ ดังนี้ (Shavelson; & Webb. 1991: 19)

$$\begin{aligned}
 X_{pi} &= \mu && \text{[ค่าคงที่หรือค่าเฉลี่ยทั้งหมด]} \\
 &+ \mu_p - \mu && \text{[อิทธิพลของบุคคล]} \\
 &+ \mu_i - \mu && \text{[อิทธิพลของข้อสอบ]} \\
 &+ X_{pi} - \mu_p - \mu_i - \mu && \text{[อิทธิพลของส่วนที่เหลือ]}
 \end{aligned}$$

เมื่อ μ คือ ค่าเฉลี่ยของคะแนนทั้งหมดจากประชากรในเอกภพ (Grand mean)
 μ_p คือ คะแนนเอกภพของแต่ละคน
 μ_i คือ คะแนนเฉลี่ยรายข้อ

จะพบว่า อิทธิพลอื่นที่ต่างไปจากค่าเฉลี่ยทั้งหมด (Grand mean) มีการกระจายเกิดขึ้น แต่ค่าเฉลี่ยทั้งหมดมีค่าคงที่ จึงมีค่าความแปรปรวนเป็นศูนย์ การแจกแจงที่มีค่าเฉลี่ยเป็นศูนย์ และมีความแปรปรวน (σ^2) เรียกว่า “องค์ประกอบความแปรปรวน (Variance component)” ประกอบด้วยส่วนแรกเป็นอิทธิพลของบุคคล ค่าเฉลี่ยของความแตกต่างของคะแนนระหว่างบุคคลเป็นศูนย์ คือ

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

ความแปรปรวนของอิทธิพลของบุคคล ใช้สัญลักษณ์เป็น (σ_p^2) เรียกว่าองค์ประกอบความแปรปรวนของบุคคล หรือ ความแปรปรวนของคะแนนเอกภพ เขียนสมการได้เป็น

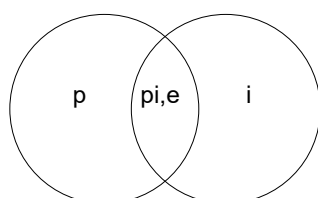
$$\sigma_p^2 = E_p (\mu_p - \mu)^2$$

ค่าเฉลี่ยและองค์ประกอบความแปรปรวนของข้อสอบ จะมีค่าเฉลี่ยรายข้อเป็นศูนย์ และองค์ประกอบความแปรปรวนของข้อสอบใช้สัญลักษณ์เป็น (σ_i^2) อิทธิพลส่วนสุดท้าย คือส่วนที่เหลือ มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนเป็น $(\sigma_{pi,e}^2)$ จะได้ว่าความแปรปรวนของคะแนนสังเกต เป็นผลรวมของ 3 องค์ประกอบความแปรปรวนข้างต้น ดังนี้

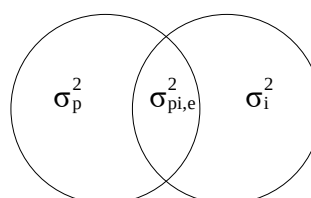
$$\sigma_{X_{pi}}^2 = \sigma_p^2 + \sigma_i^2 + \sigma_{pi,e}^2$$

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

ความแปรปรวนของคะแนนการวัดหนึ่งองค์ประกอบแบบไขว้ (One - facet crossed design) ขึ้นอยู่กับ 4 แหล่ง คือ 1) ความแตกต่างระหว่างบุคคลหรือสิ่งที่วัด 2) ความแตกต่างระหว่างความยากของข้อสอบ 3) ปฏิสัมพันธ์ระหว่างบุคคลกับข้อสอบ และ 4) ความคลาดเคลื่อนที่เกิดขึ้นอย่างสุ่ม แต่ความแปรปรวนจากแหล่งที่ 3 และ ที่ 4 ไม่สามารถแยกออกจากกันได้ เรียกว่าเป็นส่วนที่เหลือ จึงเป็นแหล่งความแปรปรวนร่วมกัน ดังภาพประกอบ 7



แหล่งของความแปรปรวน
(Source of variation)



องค์ประกอบของความแปรปรวน
(Variance component)

ภาพประกอบ 7 แหล่งของความแปรปรวนและองค์ประกอบของความแปรปรวน
(One Facet, p X i Design)

3.4 ความแปรปรวนของคะแนนความคาดเคลื่อนสัมบูรณ์และสัมพัทธ์ (Absolute and Relative Error Variance)

ตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) คะแนนจริง (True Score ; T_p) ของผู้สอบ คือ ค่าเฉลี่ยของคะแนนจากการสอบซ้ำๆ ด้วยแบบสอบคู่ขนาน ความแปรปรวนของคะแนนจริงจึงเป็นความแปรปรวนของค่าเฉลี่ยของการสอบซ้ำนั้น และความแปรปรวนของคะแนนที่สังเกตได้เป็นผลรวมของความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนความคาดเคลื่อน ดังนี้

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

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

การประมาณองค์ประกอบความแปรปรวน สามารถนำการวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) มาประยุกต์ใช้กับการศึกษา G (G-Study) (น้ำผึ้ง อินทะเนตร. 2554: 51; อ้างอิงจาก Shavelson; & Webb.1991: 28-29, 33; Brennan. 2001: 26-27) ดังนี้ ในกรณีการวัดที่มีองค์ประกอบเดียวแบบไขว้ $p \times i$

ตาราง 1 ANOVA สำหรับการวัดที่มีสององค์ประกอบแบบไขว้ $p \times i$

แหล่งความแปรปรวน	Sum of Squares	df	Mean Squares	Expected Mean Squares
Person (p)	SS_p	$n_p - 1$	$MS_p = SS_p / df_p$	$E(MS_p) = \sigma_{pi,e}^2 + n_i \sigma_p^2$
Item (i)	SS_i	$n_i - 1$	$MS_i = SS_i / df_i$	$E(MS_i) = \sigma_{pi,e}^2 + n_p \sigma_i^2$
pi,e	$SS_{pi,e}$	$(n_p - 1)(n_i - 1)$	$MS_{pi,e} = SS_{pi,e} / df_{pi,e}$	$E(MS_{pi,e}) = \sigma_{pi,e}^2$

ที่มา: ชาร์เวลสันและเวบบ์ (Shavelson; & Webb. 1991: 28)

จากตาราง ในการประมาณค่า ค่าเฉลี่ยกำลังสองที่คาดหวัง (Expected mean squares: $E(MS)$) จะใช้ค่าเฉลี่ยกำลังสองของค่าสังเกต (Observed mean square: MS) เป็นตัวประมาณค่า และการประมาณค่า σ^2 ใช้ $\hat{\sigma}^2$ เป็นตัวประมาณค่า จะได้

$$MS_p = \hat{\sigma}_{pi,e}^2 + n_i \hat{\sigma}_p^2$$

$$MS_i = \hat{\sigma}_{pi,e}^2 + n_p \hat{\sigma}_i^2$$

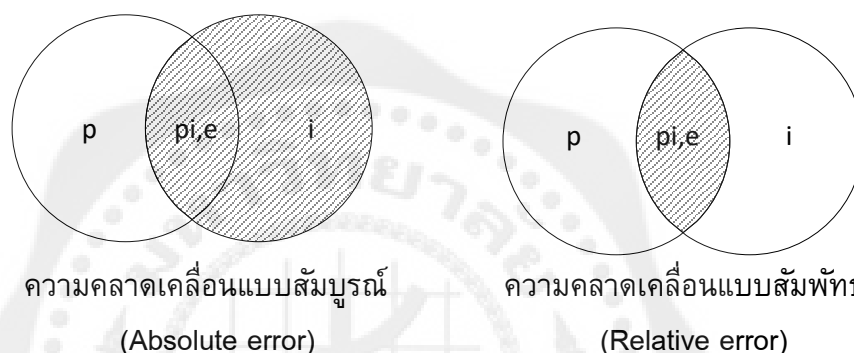
$$MS_{pi,e} = \hat{\sigma}_{pi,e}^2$$

เมื่อจัดสมการใหม่ ได้เป็น

$$\begin{aligned}\hat{\sigma}_{pi,e}^2 &= MS_{pi,e} \\ \hat{\sigma}_i^2 &= (MS_i - \hat{\sigma}_{pi,e}^2) / n_p = (MS_i - MS_{pi,e}) / n_p \\ \hat{\sigma}_p^2 &= (MS_p - \hat{\sigma}_{pi,e}^2) / n_i = (MS_p - MS_{pi,e}) / n_i\end{aligned}$$

แหล่งความคลาดเคลื่อนแบบสัมบูรณ์และแบบสัมพัทธ์ จะแตกต่างกันไปตามรูปแบบการวัดที่ได้ออกแบบไว้ มีรายละเอียดดังต่อไปนี้ (Shavelson; & Webb. 1991: 86)

กรณีการวัดที่มีองค์ประกอบเดียว แบบ p X i Design (One-facet crossed design)



ภาพประกอบ 8 แหล่งความคลาดเคลื่อนแบบสัมบูรณ์และความคลาดเคลื่อนแบบสัมพัทธ์ สำหรับ p X i Design

3.5 ค่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Coefficient)

ทฤษฎีการสรุปอ้างอิง มุ่งให้ความสำคัญกับองค์ประกอบความแปรปรวนและความคลาดเคลื่อนของการวัดแล้วยังสามารถประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง ซึ่งคล้ายคลึงกับค่าสัมประสิทธิ์ความเชื่อมั่นในทฤษฎีการทดสอบแบบมาตรฐานเดิม (Shavelson; & Webb. 1991: 92) ที่เป็นอัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของของคะแนนที่สังเกตได้ตามทฤษฎีการทดสอบแบบดั้งเดิม สัมประสิทธิ์ความเที่ยง (Reliability) ของแบบสอบ เป็นสัดส่วนระหว่างความแปรปรวนของคะแนนจริงกับคะแนนที่สังเกตได้ ดังนี้

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

ค่าสัมประสิทธิ์การสรุปอ้างอิงเป็นอัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภาพและความแปรปรวนของคะแนนสังเกตที่คาดหวัง สามารถประมาณค่าได้จากกำลังสองของค่าสหสัมพันธ์ระหว่างคะแนนเอกภาพกับคะแนนสังเกต (Cronbach; et al. 1972) ดังนั้นสำหรับ G-theory สัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient หรือ ρ_i^2) เป็นสัดส่วนระหว่างความแปรปรวนของคะแนนเอกภาพกับความแปรปรวนของค่าความคาดหมายของคะแนนที่สังเกตได้ดังนี้

$$G - coefficient = \frac{\sigma_p^2}{\sigma_p^2 + \text{ErrorVariance}}$$

ศิริชัย กาญจนวาสี (2555: 18) เนื่องจากความแปรปรวนของคะแนนความคลาดเคลื่อนจากองค์ประกอบต่างๆของการวัดมี 2 ประเภท จึงทำให้สัมประสิทธิ์การสรุปอ้างอิงมี 2 ประเภท ได้แก่ 1. สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (ρ_{Abs}^2) เมื่อคะแนนความคลาดเคลื่อนเป็นความแปรปรวนของคะแนนความคาดเคลื่อนสัมบูรณ์ สัมประสิทธิ์นี้บ่งบอกความเที่ยงของแบบทดสอบ ในสถานการณ์ของการตัดสินใจที่ขึ้นกับคะแนนผู้สอบตามลำพัง ไม่มีการเปรียบเทียบภายในกลุ่มหรือระหว่างกลุ่ม เช่น ความเที่ยงของแบบสอบอิงเกณฑ์ ซึ่งใช้แสดงความน่าเชื่อถือสำหรับการตรวจสอบความสามารถของผู้สอบเทียบกับเกณฑ์กำหนดไว้ล่วงหน้า เป็นต้น การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงสัมบูรณ์เทียบเท่ากับค่าความเชื่อมั่นในทฤษฎีการทดสอบแบบมาตรฐานเดิม (Shavelson; & Webb. 1991: 93 อ้างอิงจาก Brennan; & Kane.1977. An index of dependability for mastery test. Journal of Educational Measurement.14: 277 - 289) เรียกสัมประสิทธิ์นี้ว่า “ดัชนีความเชื่อถือได้ (Index of dependability)” และ ใช้สัญลักษณ์ ϕ (phi) หรือ ρ_{Abs}^2 2. สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (ρ_{Rel}^2) เมื่อคะแนนความคาดเคลื่อนเป็นความแปรปรวนของคะแนนความคาดเคลื่อนสัมพัทธ์ สัมประสิทธิ์นี้บ่งบอกความเที่ยงของแบบทดสอบ ในสถานการณ์ของการตัดสินใจที่มีการเปรียบเทียบคะแนนระหว่างผู้สอบ เช่น ความเที่ยงของแบบสอบอิงกลุ่ม ซึ่งแสดงความน่าเชื่อถือสำหรับการเปรียบเทียบคะแนนระหว่างผู้สอบด้วยการเข้าโค้งการแจกแจงคะแนน และตัดเกรด หรือการสอบแข่งขันหาผู้สอบที่ได้คะแนนสูงเมื่อเทียบกับกลุ่ม เพื่อเข้าศึกษาต่อในสถานศึกษาต่างๆ เป็นต้น มีสูตรดังนี้

ตาราง 2 เปรียบเทียบสูตรสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์และ
สูตรสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์

สัมประสิทธิ์การสรุปอ้างอิง	Shavelson; & Webb. 1991: 93	Brennan. 2001: 35
สำหรับการตัดสินใจ เชิงสัมบูรณ์ (ρ_{Abs}^2)	$\phi = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_{Abs}^2}$	$\phi = \frac{\sigma_{(p)}^2}{\sigma_{(p)}^2 + \sigma_{(\Delta)}^2}$
สำหรับการตัดสินใจ เชิงสัมพัทธ์ (ρ_{Rel}^2)	$E \rho_{Rel}^2 = \frac{\sigma_p^2}{E \sigma^2(X_{pi})} = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_{Rel}^2}$	$E \rho^2 = \frac{\sigma_{(p)}^2}{\sigma_{(p)}^2 + \sigma_{(\delta)}^2}$

3.6 งานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ

อรอุมา อัยวรรณ (2548) ศึกษาการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบอัตนัยการแก้โจทย์ปัญหาคณิตศาสตร์ที่มีความยาวของแบบทดสอบ วิธีการตรวจ และจำนวนผู้ตรวจแตกต่างกัน การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบอัตนัยการแก้โจทย์ปัญหาคณิตศาสตร์ที่มีความยาวของแบบทดสอบแตกต่างกัน คือ 3,6, และ 9 ข้อ การตรวจแตกต่างกัน คือ การตรวจโดยวิธีรวมและตรวจโดยวิธีการวิเคราะห์ และจำนวนผู้ตรวจแตกต่างกัน คือ 2,3, และ 5 คน กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 สังกัดเขตพื้นที่การศึกษากาฬสินธุ์ เขต 3 ปีการศึกษา 2546 จำนวน 60 คน ซึ่งได้มาโดยการสุ่มหลายขั้นตอน เครื่องมือที่ใช้ในการวิจัยในครั้งนี้ คือ แบบทดสอบอัตนัยการแก้โจทย์ปัญหาคณิตศาสตร์ จำนวน 1 ฉบับ จำนวน 9 ข้อ ผลการวิจัยปรากฏ ดังนี้ 1. แบบทดสอบอัตนัยการแก้โจทย์ปัญหาคณิตศาสตร์ที่มีความยาวของแบบทดสอบต่างกันเมื่อวิธีการตรวจเหมือนกันหรือจำนวนผู้ตรวจเท่ากัน มีค่าสัมประสิทธิ์การสรุปอ้างอิงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 2.แบบทดสอบอัตนัยการแก้โจทย์ปัญหาคณิตศาสตร์ที่มีวิธีการตรวจแตกต่างกัน เมื่อความยาวของแบบทดสอบเท่ากันหรือจำนวนผู้ตรวจเท่ากัน มีค่าสัมประสิทธิ์การสรุปอ้างอิงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 และ 3.แบบทดสอบอัตนัยการแก้โจทย์ปัญหาคณิตศาสตร์ที่มีจำนวนผู้ตรวจแตกต่างกัน เมื่อความยาวของแบบทดสอบเท่ากัน หรือวิธีการตรวจเหมือนกันมีค่าสัมประสิทธิ์การสรุปอ้างอิงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

น้ำผึ้ง อินทเนตร (2554) ศึกษาคุณลักษณะของคะแนนแบบทดสอบปลายเปิดวิชาคณิตศาสตร์เมื่อจำนวนผู้ตรวจและรูปแบบการตรวจให้คะแนนต่างกัน โดยใช้โมเดลการสรุปอ้างอิงและโมเดลหลายองค์ประกอบของราล์ซ การวิจัยในครั้งนี้เป็นการศึกษาคุณลักษณะของคะแนนแบบทดสอบปลายเปิดวิชาคณิตศาสตร์โดยใช้โมเดลการสรุปอ้างอิงและโมเดลหลายองค์ประกอบของราล์ซ ภายใต้เงื่อนไขจำนวนผู้ตรวจต่างกันสามลักษณะ คือ 2 คน 3 คน และ 4 คน และรูปแบบการตรวจให้คะแนนต่างกันสามลักษณะคือ รูปแบบที่ 1 ผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน รูปแบบที่ 2 ผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบบางคน และรูปแบบที่ 3 ผู้ตรวจตรวจข้อสอบบางข้อของผู้สอบทุกคน คุณลักษณะของคะแนนพิจารณาจากขนาดขององค์ประกอบความแปรปรวน ค่าความเชื่อมั่น และค่าความเที่ยงตรงตามสภาพเครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ แบบทดสอบปลายเปิดวิชาคณิตศาสตร์ระดับมัธยมศึกษาตอนต้นตามหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 จำนวน 12 ข้อ กลุ่มตัวอย่างที่ใช้คือนักเรียนระดับชั้น มัธยมศึกษาปีที่ 4 จากโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาจังหวัดน่าน ปีการศึกษา 2552 จำนวน 180 คน ผลการวิจัยสรุปได้ว่า 1. เมื่อวิเคราะห์ด้วยโมเดลการสรุปอ้างอิง พบว่า เมื่อใช้รูปแบบการตรวจให้คะแนนเดียวกันในทุกเงื่อนไขจำนวนผู้ตรวจ ความแปรปรวนขององค์ประกอบเดียวกันมีค่าใกล้เคียงกัน และค่าสัมประสิทธิ์การสรุปอ้างอิงในรูปแบบการตรวจที่ 2 มีค่าสูงสุด รองลงมาคือรูปแบบการตรวจที่ 1 และรูปแบบการตรวจที่ 3 มีค่าต่ำสุด ค่าสัมประสิทธิ์การสรุปอ้างอิงในรูปแบบการตรวจที่ 1 มีค่าสูงขึ้น เมื่อจำนวนผู้ตรวจเพิ่มขึ้น คะแนนในทุกเงื่อนไขที่ต่างกันมีความเที่ยงตรงตามสภาพสูงและแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ 2. เมื่อวิเคราะห์ด้วยโมเดลหลายองค์ประกอบของราล์ซ พบว่า ในทุกเงื่อนไขจำนวนผู้ตรวจและรูปแบบการตรวจให้คะแนน ความแปรปรวนของผู้สอบมีค่าสูงที่สุด รองลงมาคือ ความแปรปรวนของข้อสอบ และความแปรปรวนของผู้ตรวจมีค่าต่ำที่สุด และค่าความเชื่อมั่นแยกส่วนของผู้สอบของรูปแบบการตรวจที่ 1 มีค่าสูงสุดในทุกเงื่อนไขจำนวนผู้ตรวจ ค่าความเชื่อมั่นแยกส่วนของผู้สอบของรูปแบบการตรวจที่ 1 มีค่าสูงขึ้น เมื่อจำนวนผู้ตรวจเพิ่มขึ้นคะแนนความสามารถในทุกเงื่อนไขที่ต่างกันมีความเที่ยงตรงตามสภาพสูงและแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

สาวิตรี จุ้ยทอง (2554) ศึกษาการตรวจสอบคุณภาพรูปแบบของชุดเครื่องมือการประเมินผลการเรียนรู้ตามสภาพจริงกลุ่มสาระการเรียนรู้คณิตศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อสร้างชุดเครื่องมือการประเมินผลการเรียนรู้ตามสภาพจริงกลุ่มสาระการเรียนรู้คณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 5 ฉบับ และเพื่อตรวจสอบคุณภาพรูปแบบของชุดเครื่องมือการประเมินผลการเรียนรู้ตามสภาพจริง จำนวน 3 รูปแบบ โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่

1 ภาคเรียนที่ 1 ปีการศึกษา 2553 ของโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน จำนวน 43 โรงเรียน มีประชากรทั้งสิ้น 16,793 คน กลุ่มตัวอย่างประกอบด้วยนักเรียนจำนวน 7 โรงเรียน รวม 10 ห้องเรียน จำนวนนักเรียน 413 คน ซึ่งได้มาโดยวิธีการสุ่มแบบหลายขั้นตอน (Multi - Stage Sampling) โดยทำการสร้างชุดเครื่องมือการประเมินผลการเรียนรู้ตามสภาพจริงให้สอดคล้องกับแผนการจัดการเรียนรู้ที่ผู้วิจัยสร้างขึ้น ซึ่งประกอบด้วยเครื่องมือการประเมินจำนวน 5 ฉบับ ได้แก่ แบบทดสอบวัดความสามารถจริง (Authentic Test) แบบสังเกตพฤติกรรมการแสดงออกในการเรียนรู้ (Behavior – Observation) แบบประเมินแฟ้มสะสมงาน (Portfolio) แบบวัดการรับรู้ความสามารถของตนเอง (Self – Efficacy) และแบบวัดการกำกับตนเอง (Self - Regulation) การตรวจสอบคุณภาพของเครื่องมือ ใช้สถิติวิเคราะห์ประกอบด้วยสถิติบรรยาย ค่าสัมประสิทธิ์แอลฟาของครอนบัค (Cronbach's Alpha Coefficient: α) สหสัมพันธ์แบบไบซีเรียล (Biserial Correlation) และตรวจสอบคุณภาพรูปแบบของชุดเครื่องมือการประเมินผลการเรียนรู้ตามสภาพจริงทั้ง 3 รูปแบบ โดยหาค่าสัมประสิทธิ์การสรุปอ้างอิง (G – Coefficient) ผลการวิจัยพบว่า

1. ชุดเครื่องมือการประเมินผลการเรียนรู้ตามสภาพจริง กลุ่มสาระการเรียนรู้คณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ที่ผู้วิจัยได้สร้างขึ้น จำนวน 5 ฉบับ ได้แก่ แบบทดสอบวัดความสามารถจริง (Authentic Test) แบบสังเกตพฤติกรรมการแสดงออกในการเรียนรู้ (Behavior – Observation) แบบประเมินแฟ้มสะสมงาน (Portfolio) แบบวัดการรับรู้ความสามารถของตนเอง (Self – Efficacy) และ แบบวัดการกำกับตนเอง (Self - Regulation) โดยแบบทดสอบวัดความสามารถจริง (Authentic Test) มีค่าความเที่ยงตรงตามเนื้อหา (Content Validity) ตั้งแต่ 0.80 – 1.00 ค่าความยากง่าย (Difficulty) ตั้งแต่ 0.49 – 0.61 ค่าอำนาจจำแนก (Discrimination) ตั้งแต่ 0.32 – 0.52 ค่าความเชื่อมั่น (Reliability) ของแบบทดสอบ เท่ากับ 0.89 และค่าดัชนีความสอดคล้องระหว่างผู้ประเมิน (Rater Agreement Index : RAI) เท่ากับ 0.76

2. การศึกษาคุณภาพรูปแบบของชุดเครื่องมือการประเมินผลการเรียนรู้ตามสภาพจริง กลุ่มสาระการเรียนรู้คณิตศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 3 รูปแบบ โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) ซึ่งเป็นการศึกษาคุณภาพด้านความเชื่อมั่น (Reliability) โดยศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) พบว่า ชุดเครื่องมือการประเมินผลการเรียนรู้ตามสภาพจริง รูปแบบที่ 1 ซึ่งประกอบด้วยแบบประเมินจำนวน 3 ฉบับ มีค่าสัมประสิทธิ์การสรุปอ้างอิง เท่ากับ 0.89 รูปแบบที่ 2 ซึ่งประกอบด้วยแบบประเมินจำนวน 4 ฉบับ มีค่าสัมประสิทธิ์การสรุปอ้างอิง เท่ากับ 0.92 และรูปแบบที่ 3 ซึ่งประกอบด้วยแบบประเมินจำนวน 5 ฉบับ มีค่าสัมประสิทธิ์การสรุปอ้างอิง เท่ากับ 0.93

อกนิษฐ์ บุญสะอาด (2557) ศึกษาการพัฒนาแบบประเมินความคิดรวบยอดวิชาฟิสิกส์โดยใช้แผนผังมโนทัศน์ : การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด การศึกษาครั้งนี้มีความมุ่งหมายเพื่อพัฒนาและตรวจสอบคุณภาพของแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์วิชาฟิสิกส์ เรื่องงานและพลังงาน ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 และเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์แบบการสร้างผังคำและคำเชื่อม และแบบการเลือกคำเชื่อมที่มีลำดับของการสร้างแผนผังมโนทัศน์ที่แตกต่างกัน และจำนวนครั้งในการวัดที่แตกต่างกันโดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 แผนการเรียนวิทยาศาสตร์ คณิตศาสตร์ ปีการศึกษา 2555 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 16 ขนาดใหญ่พิเศษ จำนวน 1 โรงเรียน ที่ได้มาจากการสุ่มสองขั้นตอน (Two stage Sampling) เครื่องมือที่ใช้ในการวิจัย คือ แบบประเมินความคิดรวบยอดวิชาฟิสิกส์โดยใช้แผนผังมโนทัศน์แบบการสร้างผังคำและคำเชื่อม และแบบการเลือกคำเชื่อม สถิติที่ใช้ในการวิเคราะห์ข้อมูล สถิติทดสอบ t-test ตามสูตรของพิทแมน และการวิเคราะห์หาค่าสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) ผลการวิจัยพบว่า 1.ค่าความเที่ยงตรงเชิงเนื้อหา (Content validity) แบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์แบบการสร้างผังคำและคำเชื่อม มีค่าดัชนีค่าความสอดคล้องจากการพิจารณา อยู่ระหว่าง 0.94 ถึง 1.00 ค่าความยาก (Difficulty) มีค่าตั้งแต่ 0.24 ถึง 0.88 ค่าอำนาจจำแนกมีค่าตั้งแต่ 0.21 ถึง 0.75 ค่าความเชื่อมั่นของแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์แบบการสร้างผังคำและคำเชื่อม ชุดที่ 1,2,3 และ 4 มีค่าเท่ากับ 0.66 0.68 , 0.72 และ 0.71 ตามลำดับ 2.ค่าความเที่ยงตรงเชิงเนื้อหา (Content validity) แบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์แบบการเลือกคำเชื่อม มีค่าดัชนีค่าความสอดคล้องจากการพิจารณา อยู่ระหว่าง 0.94 ถึง 1.00 ค่าความยาก (Difficulty) มีค่าตั้งแต่ 0.20 ถึง 0.88 ค่าอำนาจจำแนกมีค่าตั้งแต่ 0.21 ถึง 0.75 ค่าความเชื่อมั่นของแบบประเมินความคิดรวบยอด โดยใช้แผนผังมโนทัศน์แบบการเลือกคำเชื่อม ชุดที่ 1, 2, 3 และ 4 มีค่าเท่ากับ 0.72 0.83 , 0.69 และ 0.67 ตามลำดับ 3.การเปรียบเทียบคุณภาพความเชื่อมั่น โดยเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์ พบว่า 3.1 ค่าสัมประสิทธิ์การสรุปอ้างอิงของลำดับของการสร้างแผนผังมโนทัศน์ที่เริ่ม จากแบบการสร้างผังคำและคำเชื่อม ตามด้วยแบบการเลือกคำเชื่อม (CS) ไปสู่ลำดับของการสร้างแผนผังมโนทัศน์ที่เริ่มจากแบบการเลือกคำเชื่อม ตามด้วยแบบการสร้างผังคำและคำเชื่อม(SC)(CS → SC) ซึ่งมีค่าเท่ากับ 0.96 ส่วนค่าสัมประสิทธิ์การสรุปอ้างอิงของลำดับของการสร้างแผนผังมโนทัศน์ที่เริ่มจากแบบการเลือกคำเชื่อม ตามด้วยแบบการสร้างผังคำและคำเชื่อม (SC) ไปสู่ลำดับของการสร้างแผนผังมโนทัศน์ที่เริ่มจากแบบการสร้างผังคำและคำเชื่อม ตามด้วยแบบการเลือกคำเชื่อม (CS) (SC → CS) มีค่าเท่ากับ 0.54 และเมื่อทำการทดสอบความแตกต่างของค่าสัมประสิทธิ์การสรุป

อ้างอิง พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงมีค่าแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 3.2 ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์แบบการเลือกคำเชื่อม (SS) จากการวัด 1 ครั้งและ 2 ครั้ง มีค่าเท่ากับ 0.52 และ 0.69 ตามลำดับ เมื่อทำการทดสอบความแตกต่างของค่าสัมประสิทธิ์การสรุปอ้างอิง พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงมีค่าแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 3.3 ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์แบบการสร้างผังคำและคำเชื่อม (CC) จากการวัด 1 ครั้ง และ 2 ครั้ง มีค่าเท่ากับ 0.40 และ 0.58 เมื่อทำการทดสอบความแตกต่างของค่าสัมประสิทธิ์การสรุปอ้างอิง พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงมีค่าแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

คาล ชวาร์ส และคณะ (Carl W. Swartz; et al.1999) ศึกษาการใช้ทฤษฎีความน่าเชื่อถือของผลการวัด ในการประเมินค่าความเชื่อมั่นของวิธีการให้คะแนนแยกองค์ประกอบและแบบองค์รวม วัดคุณสมบัติของการวิจัยนี้คือการใช้ GT ในการตรวจสอบความน่าเชื่อถือของคะแนนที่ได้มาจากการประเมินแบบแยกองค์ประกอบและแบบองค์รวม กลุ่มตัวอย่างในการศึกษาครั้งนี้เป็นโรงเรียนมัธยมภาคใต้และภาคกลางอลาบามา ทั้งชนบท / ในชนานเมือง เป็นนักเรียน 251 คนที่อยู่ในช่วงอายุ 11-15 ปี ($M = 12.25$, $SD = 1.06$) นักเรียนเหล่านี้เป็น 55% หญิง, 52% เป็นคนผิวขาว 48% เป็นแอฟริกันอเมริกันและ 61% เครื่องมือที่ใช้ คือ TOWL เป็นมาตรฐาน towl-2 ตามกระบวนการ นักเรียนได้รับเวลา 15 นาที ในการเขียนการตอบ คะแนนที่ได้ตามขั้นตอนการให้คะแนนมาตรฐาน มีประเด็นดังนี้ ใจความสำคัญ บริบท คำศัพท์ การสร้างประโยค การสะกด และรูปแบบการเขียน ขั้นตอนการวิเคราะห์ข้อมูลจำนวนผู้ตรวจ 4 คน ในการศึกษา generalizability เป็นสมาชิกของคณะศูนย์การศึกษาทางคลินิกของการพัฒนาและการเรียนรู้ (CDL) ที่มหาวิทยาลัยนอร์ทแคโรไลนาแซเปิลฮิลล์ (UNC-CH) ผู้ประเมินสำเร็จการศึกษาระดับปริญญาเอก (จิตวิทยา, การศึกษา, การพูดและการได้ยิน) และหนึ่งในผู้ประเมินสำเร็จการศึกษาปริญญาโทในด้านการศึกษา (การศึกษาพิเศษ) ผู้ประเมินแต่ละคนมีประสบการณ์การจัดการเรียนการสอนทางคลินิกและประสบการณ์การวิจัยอย่างกว้างขวางการทำงานกับเด็กและวัยรุ่น การออกแบบการวัดเอกภพหนึ่งฟาเซท แบบ $P \times R$ Design มีแหล่งความแปรปรวน 3 แหล่ง ได้แก่ (ก) บุคคล (P) คือ นักเรียนที่สอบ (ข) ผู้ประเมิน (R) และ (ค) การปฏิสัมพันธ์จำนวนผู้ประเมินกับผู้สอบ (PR)

โดยการพิจารณาแหล่งที่มาของความแปรปรวน โดยใช้การวิเคราะห์ความแปรปรวน (ANOVA) แบ่งความแปรปรวนเนื่องที่มาจากแต่ละแหล่ง โดยคำนวณค่าสัมประสิทธิ์ G – Coefficients ดังสูตรต่อไปนี้

การตัดสินใจเชิงสัมพัทธ์ (relative decisions) $\sigma_p^2 / (\sigma_p^2 + \sigma_{rp}^2) / n_r$

การตัดสินใจเชิงสัมบูรณ์ (absolute decisions) $\sigma_p^2 / (\sigma_p^2 + [\sigma_r^2 / n_r + \sigma_{rp}^2 / n_r])$

การประมาณค่าความแปรปรวนแหล่งที่มาของแต่ละคนถูกนำมาใช้ในการดำเนินการ การศึกษา D –Study เพื่อตรวจสอบผลกระทบของจำนวนผู้ประเมินและประเภทของการ ตัดสินใจของค่าสัมประสิทธิ์ generalizability ผลการวิเคราะห์ข้อมูลโปรแกรมประโยชน์ของ ทฤษฎี G คือการใช้การประมาณการความแปรปรวนที่มีผลจากการศึกษาเชิงสรุปอ้างอิงและ ดำเนินการศึกษาเชิงตัดสินใจ เพื่อประเมินความน่าเชื่อถือของคะแนนเมื่อแหล่งที่มาของ ข้อผิดพลาดในการวัดจะจัดการอย่างเป็นระบบ เพราะคะแนนที่ได้สามารถนำมาใช้ในการ ตัดสินใจเชิงสัมพัทธ์ลำดับของนักเรียนจากสูงไปต่ำและยังสามารถใช้ในการทำให้การตัดสินใจ เชิงสัมบูรณ์ว่าคะแนนของนักเรียนถึงเกณฑ์มาตรฐานหรือไม่ การศึกษา D-Study เกี่ยวข้องไปสู่ การตัดสินใจเชิงสัมบูรณ์เป็นสิ่งที่สำคัญอย่างยิ่งเพราะห้องเรียนเดียวครูการศึกษาพิเศษหรือ นักจิตวิทยาโรงเรียนอาจจะใช้คะแนนจากการทดสอบในการพัฒนาการวางแผนการศึกษาและ / หรือการตัดสินใจอย่างมีนัยสำคัญที่ส่งผลกระทบต่อนักเรียน ค่าสัมประสิทธิ์ generalizability (สี่ผู้ประเมิน 0.98 กับหนึ่งในผู้ประเมิน .94) ความน่าเชื่อถือของคะแนนที่เพิ่มขึ้นในลักษณะ เดียวกับการเพิ่มจำนวนข้อของแบบสอบมักจะเพิ่มความน่าเชื่อถือของคะแนน

ออกसान่า นาอูเมนโก (Oksana Naumenko. 2015) ศึกษาการปรับปรุง ประสิทธิภาพการประเมินผลคะแนนการปฏิบัติที่การตรวจสอบ : หลักสูตรการเรียนการสอน เกี่ยวกับทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด เกณฑ์การให้คะแนนเป็น สมมติฐานช่วงการวัดการสอบภาษาต่างประเทศการเขียนบนพื้นฐานของมาตรฐานที่แตกต่าง จากจาก Common European Framework of Reference ภาษา (CEFR; สภายุโรป 2009) สำหรับวัตถุประสงค์ของตัวอย่างที่เกณฑ์การให้คะแนน คือ 1) ช่วงภาษาศาสตร์ทั่วไป 2) ช่วงคำศัพท์ 3) ความถูกต้องตามหลักไวยากรณ์ 4) การควบคุมคำศัพท์ 5) การควบคุมภาพ ฉาย และ 6) การเชื่อมโยงกัน วิเคราะห์ส่วนประกอบเกณฑ์ (หรือองค์ประกอบ) การ เปรียบเทียบพฤติกรรมรายละเอียดกลุ่มตัวอย่างผู้อ่าน ตั้งสมมติฐานในการศึกษาครั้งนี้มี ผู้เชี่ยวชาญด้านการเขียน 12 ภาษา เนื้อหาที่ใช้ในการประเมินการตอบสนองของผู้เข้าสอบ เป็นลายลักษณ์อักษร ผลการศึกษาครั้งนี้ได้รับการการจัดระเบียบเป็นส่วนที่สำคัญโดย คำถามการวิจัย (เช่น RQS A, Ba, Bb, BC) การออกแบบการทดลอง $[(p \times r):t] \times i$ กับผู้ ประเมินสามคนซ้อนกันภายในสี่สิบสองของผู้ประเมินทั้งหมด โดยมีการคำนวณที่ถูกสร้างขึ้น ด้วยมือและการตรวจสอบผ่านชุดโปรแกรม GENOVA (Crick; & Brennan, 1986). ที่ออกแบบ มาเพื่อดำเนินการศึกษา

ปีเตอร์ สโปเรน และคณะ (Pieter Spooren; et al. 2014: 88 - 94) ศึกษาการประเมิน ความเที่ยงตรงและความเชื่อมั่นของการสแกนนักเรียนอย่างรวดเร็วสำหรับการประเมินผลการ เรียนการสอนจากการวิเคราะห์ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด วัตถุประสงค์เพื่อสังเกตคะแนนในหลายมิติของการจัดการเรียนการสอน (Apodaca; & Grad, 2005, Burdsal; & Harrison, 2008, Ginns; & Barrie, 2009; Spooren; & Mortelmans, 2009) โดยขั้นตอน

การตรวจสอบของเครื่องมือดังกล่าวได้จากการตอบแบบสอบถาม SET จำนวน 37 ข้อ สำหรับการประเมินผลการเรียนการสอนของนักเรียน (Mortelmans; & Spoooren. 2009) สมมติฐานพื้นฐานสำหรับแบบสอบถาม พบว่า ควรจะเป็น (ก) มีความเที่ยงตรงและความเชื่อมั่น (ข) สั้นได้ใจความ และ (ค) คำถามเกี่ยวกับมิติที่สำคัญของการเรียนการสอน นอกจากนี้แบบสอบถามควรทำงานได้ดีกับหลักสูตรและสาขาวิชาที่มีความแตกต่างกัน ดังนั้นขั้นตอนการสุ่มตัวอย่างต้องการที่จะครอบคลุมมากที่สุดเท่าที่เป็นไปได้ที่สาขาวิชาและระดับการศึกษา นอกจากนี้เพื่อให้เครื่องมือที่ใช้ในการทดสอบบนความมั่นคงของข้อมูลที่จะเก็บมากกว่าสองครั้งหรือโอกาสที่แตกต่างกัน กลุ่มตัวอย่างการบริหารจัดการในชั้นเรียนในช่วงฤดูใบไม้ร่วง ปีการศึกษา 2012-2013 ที่มหาวิทยาลัยแอนด์เวิร์ธ (ขนาดกลางมหาวิทยาลัยเบลเยียมที่มีประมาณ 13,000 คน) นักเรียนจาก 6 คณะ (วิทยาศาสตร์สังคมและวิทยาศาสตร์ทางการแพทย์, กฎหมาย, วรรณคดีและปรัชญา เศรษฐศาสตร์และเภสัชกรรมทางการแพทย์และสัตวแพทยศาสตร์) โดยประเมิน 16 หลักสูตร สองครั้ง (ในช่วงเวลา 3-4 สัปดาห์) เพื่อให้นักเรียนกรอกแบบสอบถามสองครั้ง เพื่อให้ข้อเสนอแนะเกี่ยวกับหลักสูตรที่พวกเขาได้ปฏิบัติตามในช่วงปีการศึกษา 2011 - 2012 จำนวนนักเรียนที่ได้รับการประเมินหลักสูตร (ช่วงเวลาแรกนักเรียนจำนวน 1,139 คน) และ (ช่วงเวลาที่สองนักเรียนจำนวน 2,941 คน) เพื่อให้มีความน่าเชื่อถือและความมั่นคงในระหว่างขั้นตอนการทดสอบการตรวจสอบ ผลการวิเคราะห์ข้อมูล การวิเคราะห์หาความสัมพันธ์ ณ ช่วงเวลาแรกระหว่างความสัมพันธ์ระหว่างช่วงเวลาที่สองของทุกหลักสูตรอย่างมีนัยสำคัญทางสถิติ ในระดับปานกลางไปจนถึงขนาดใหญ่อยู่ในช่วงระหว่าง 0.25 และ 0.59 (Cohen, 1992) ความสัมพันธ์ระหว่างข้ออยู่ระหว่าง 0.35 และ 0.59 ซึ่งมีข้อบ่งชี้แรกของความเที่ยงตรงมาบรรจบกันของข้อในแบบสอบถาม (Ginns; & Barrie. 2009) อย่างไรก็ตามความสัมพันธ์ระหว่างข้อที่แสดงให้เห็นความหลากหลายมากขึ้น (กับความสัมพันธ์ระหว่าง 0.34 และ 0.92) นี่คือการค้นพบที่น่าสนใจในขณะที่แสดงให้เห็นว่าเครื่องมือที่จะช่วยให้นักเรียนทราบ (จุดแข็งและ / หรือจุดอ่อน) ของตนเอง การวิเคราะห์ความน่าเชื่อถือ (G-Study) ประเมินความน่าเชื่อถือของแบบสอบถาม SET โดยการออกแบบการวัดแบบไขว้ (P (บุคคล) $\times$ I (ข้อ) $\times$ T (เวลา)) ในการวิเคราะห์ข้อคำถามกับเวลาใช้การสุ่ม เนื่องจากผลของการวัดแบบสอบถาม SET มักใช้ในการเปรียบเทียบคุณภาพการเรียนการสอน (ระหว่างครูผู้สอนหรือภายในของแต่ละบุคคลพื้นฐานระหว่างประสิทธิภาพก่อนหน้าและปัจจุบัน) โดยไม่ต้องใช้ (การตัดสินใจเชิงสัมบูรณ์) คะแนนตัดสินใจเชิงสัมพัทธ์เป็นที่ต้องการ สำหรับการวิเคราะห์โปรแกรม SPSS G1.sps ของ (Mushquash; & O'Connor. 2006) สำหรับการวิเคราะห์ทฤษฎี generalizability องค์ประกอบความแปรปรวนบุคคลมีขนาดใหญ่ที่สุดประมาณค่าองค์ประกอบ ($p = 0.33$) เป็นรูปแบบในคำตอบของนักเรียน (คะแนนเฉลี่ย) เป็นความแปรปรวนที่พึงประสงค์ (และไม่ควรได้รับการพิจารณาถึงความผิดพลาด) สะท้อนให้เห็นถึงความแตกต่างระหว่างนักเรียนแต่ละคน ส่วนประกอบความแปรปรวนสำหรับข้อ (I) กับเวลา (T) มีการปฏิสัมพันธ์ระหว่างข้อและเวลา อยู่ที่ประมาณใกล้เคียงกับ 0 ซึ่งแสดงให้เห็นว่ามีเกือบจะไม่มี

เปลี่ยนแปลงในองศาของข้อที่แตกต่างกันในแบบสอบถามวัดคุณภาพของหลักสูตร พบว่า คะแนนยังคงเหนียวแน่นมากกว่าสองครั้ง ส่วนที่เหลืออีกองค์ประกอบความแปรปรวนเป็นตัวชี้วัดของความผิดพลาดความแปรปรวน มีการทับซ้อนเล็กน้อยระหว่างบุคคลและเวลา อธิบายความแปรปรวนประมาณ 21% เป็นตัวชี้วัดของข้อผิดพลาดทั้งหมดความแปรปรวนในการออกแบบ (เชิงสัมบูรณ์) และความแปรปรวนข้อผิดพลาดทั้งหมดส่วนประกอบที่เกี่ยวข้องกับทั้งผู้ตอบแบบสอบถามและเป็นหนึ่งในแง่มุมอื่นๆ (เชิงสัมพัทธ์) สะท้อนให้เห็นถึงความแปรปรวนข้อผิดพลาดเล็กน้อย ในรูปแบบเชิงสัมพัทธ์และเชิงสัมบูรณ์ ค่าสัมประสิทธิ์ $G - \text{Coefficient}$ ระบุว่าคะแนนของ SET มีความเชื่อมั่นได้ทั่วทุกข้อเกินเกณฑ์มาตรฐานสำหรับความน่าเชื่อถือ (โดยปกติจะอยู่ระหว่าง 0.70 และ 0.80) ผลการศึกษา D-Study แสดงให้เห็นว่าความน่าเชื่อถือของแบบสอบถามเป็นที่ยอมรับเมื่อใช้เพียง ครั้งเดียว โดยมีค่าสัมประสิทธิ์ $G - \text{Coefficient}$ 0.75

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) กล่าวโดยสรุปได้ว่า ทฤษฎีการสรุปอ้างอิงขยายจากทฤษฎีการทดสอบแบบดั้งเดิมเป็นการประมาณค่าขนาดของความแปรปรวนจากหลายๆ แหล่งภายใต้บริบทการวัดจากการวิเคราะห์เพียงครั้งเดียว โดยมีการประมาณค่าสัมประสิทธิ์ความเชื่อมั่น หรือที่เรียกว่า “สัมประสิทธิ์การสรุปอ้างอิง (Generalizability coefficient)” เป็นการศึกษาองค์ประกอบความแปรปรวนของแต่ละแหล่งของความคลาดเคลื่อนที่ส่งผลการวัด นำการวัดที่ได้ไปสู่การตัดสินใจ 2 ลักษณะ ได้แก่ การตัดสินใจเชิงสมบูรณ์ (Absolute decision) และการตัดสินใจเชิงสัมพัทธ์ (Relative decision) โดยมีรูปแบบการวิเคราะห์ที่แตกต่างกันอย่างชัดเจน โดยการศึกษา G-Study เป็นการประมาณค่าน้ำหนักของแหล่งความคลาดเคลื่อนของการวัดที่เป็นไปได้ทั้งหมด ส่วนการศึกษา D-Study เป็นการนำข้อมูลจากการศึกษา G-Study นำมาออกแบบการวัดที่ทำให้มีความคลาดเคลื่อนน้อยที่สุดตามเป้าหมายที่ต้องการ การประมาณค่าองค์ประกอบความแปรปรวนในการศึกษา G-Study ใช้ในการออกแบบกระบวนการวัดที่มีผลทางด้านความคุ้มค่าของเวลาและค่าใช้จ่ายในการทดสอบ และสามารถนำไปใช้ได้หลายรูปแบบของการวัดทั้งรูปแบบการวัดแบบไขว้ (Crossed) แบบแฝง (Nested) หรือแบบแฝงบางส่วน (Partially nested) ผู้วิจัยจึงสนใจศึกษาคุณภาพแบบสอบการรู้เรื่องคณิตศาสตร์ที่มีจำนวนข้อและกฎเกณฑ์การให้คะแนนต่างกัน โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ซึ่งผู้วิจัยได้ศึกษาองค์ประกอบความแปรปรวนในแต่ละแหล่ง พิจารณาคูณลักษณะของคะแนนจากค่าสัมประสิทธิ์การสรุปอ้างอิง โดยทำการประมาณค่าและเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง เมื่อความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน เพื่อตรวจสอบคุณลักษณะของคะแนน ได้แก่ ค่าสัมประสิทธิ์การสรุปอ้างอิงของความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน ซึ่งเป็นแนวทางในการนำไปใช้ปฏิบัติได้จริงได้อย่างเหมาะสมต่อไป

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยเพื่อสร้างและตรวจสอบคุณภาพของแบบสอบถามรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบถามและกฎเกณฑ์การให้คะแนนต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. การกำหนดประชากรและกลุ่มตัวอย่าง
2. การสร้างเครื่องมือที่ใช้ในการวิจัย
3. การเก็บรวบรวมข้อมูล
4. การจัดกระทำและการวิเคราะห์ข้อมูล

การกำหนดประชากรและการเลือกกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย

ประชากรในการวิจัยครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานภาคใต้ ปีการศึกษา 2560 จำนวน 645 โรงเรียน มีประชากรทั้งสิ้นจำนวน 64,253 คน ดังตาราง 3

ตาราง 3 ประชากรนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำแนกตามจังหวัด

จังหวัด	จำนวนโรงเรียน	จำนวนนักเรียน
1. กระบี่	36	4,099
2. ชุมพร	58	4,891
3. ตรัง	40	4,985
4. นครศรีธรรมราช	121	11,945
5. นราธิวาส	48	3,321
6. ปัตตานี	30	2,159
7. พังงา	30	2,363
8. พัทลุง	39	3,809
9. ภูเก็ต	21	3,118
10. ยะลา	26	1,987
11. ระนอง	30	1,832
12. สงขลา	50	8,683
13. สตูล	33	2,527
14. สุราษฎร์ธานี	83	8,624
รวม	645	64,253

ที่มา: ศูนย์เทคโนโลยีสารสนเทศและการสื่อสาร (ประมวลผล ณ วันที่ 10 มกราคม 2560)

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานภาคใต้ ภาคเรียนที่ 2 ปีการศึกษา 2560 จำนวน 400 คน ผู้วิจัยกำหนดขนาดตัวอย่างจากการเปิดตารางสำเร็จรูปสำหรับศึกษาค่าเฉลี่ยของประชากรที่ระดับความเชื่อมั่น 95% เมื่อยอมให้ค่าความคลาดเคลื่อนของการประมาณค่าสัดส่วนที่เกิดขึ้นในระดับ $\pm 5\%$ ของส่วนเบี่ยงเบนมาตรฐาน (ศิริชัย กาญจนวาสี และคณะ. 2551 : 52) และใช้การสุ่มตัวอย่างแบบหลายขั้นตอน (Multi stage random sampling) ดังนี้

ขั้นตอนที่ 1 สุ่มจังหวัดในภาคใต้ จำนวน 5 จังหวัดจาก 14 จังหวัด โดยใช้วิธีจับสลากขึ้นมาทีละจังหวัด ได้แก่ จังหวัดกระบี่ ตรัง นครศรีธรรมราช พังงา และ ภูเก็ต

ขั้นตอนที่ 2 สุ่มโรงเรียนแต่ละจังหวัด จาก 5 จังหวัด โดยใช้วิธีจับสลากขึ้นมาทีละโรงเรียน จังหวัดละ 4 โรงเรียน รวมจำนวนโรงเรียนทั้งหมด 20 โรงเรียน

ขั้นตอนที่ 3 สุ่มกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 2 ในแต่ละโรงเรียน โรงเรียนละ 20 คน

ตาราง 4 ขนาดกลุ่มตัวอย่างนักเรียน จำแนกตามจังหวัด

จังหวัด	โรงเรียน	จำนวนนักเรียน ชั้นม.2 ทั้งหมด	ขนาดกลุ่ม ตัวอย่างนักเรียน ชั้นม.2
1. กระบี่	คลองท่อมราษฎร์รังสรรค์	368	20
	อำมาตย์พานิชนุกูล	473	20
	อ่าวลึกประชาสรรค์	446	20
	เมืองกระบี่	380	20
2. ตรัง	สภาราชินี 1	403	20
	วิเชียรมาตุ	438	20
	รัฐานุประดิษฐ์นุสรณ์	143	20
	วังวิเศษ	175	20
3. นครศรีธรรมราช	เบญจมาภูกิต	481	20
	กัลยาณิศรีธรรมราช	585	20
	สตรีทุ่งสง	382	20
	นบพิตรวิทยา	117	20
4. พังงา	ตะกั่วป่า(เสนานุกูล)	342	20
	สตรีพังงา	133	20
	ดีบุกพังงาวิทยายน	376	20
	ท้ายเหมืองวิทยา	249	20

ตาราง 4 (ต่อ)

จังหวัด	ชื่อโรงเรียน	จำนวนนักเรียน ชั้นม.2 ทั้งหมด	ขนาดกลุ่ม ตัวอย่างนักเรียน ชั้นม.2
5. ภูเก็ต	ภูเก็ตวิทยาลัย	519	20
	สตรีภูเก็ต	494	20
	เฉลิมพระเกียรติสมเด็จพระศรีนครินทร์ภูเก็ต	441	20
	เมืองกลาง	415	20

การสร้างเครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ แบบสอบการรู้เรื่องคณิตศาสตร์ ผู้วิจัยสร้างแบบสอบครอบคลุม 4 สาระ ได้แก่ ปริภูมิและรูปทรงสามมิติ การเปลี่ยนแปลงและความสัมพันธ์ ปริมาณและความไม่แน่นอน แต่ละเนื้อหา มี 3 สถานการณ์ แต่ละสถานการณ์มี 3 ข้อ ตามกระบวนการการรู้เรื่องคณิตศาสตร์ รวม 36 ข้อ การทดสอบใช้ระยะเวลา 120 นาที ดังตาราง 5

ตาราง 5 กรอบโครงสร้างแบบสอบการรู้เรื่องคณิตศาสตร์

เนื้อหา	สถานการณ์	กระบวนการทางคณิตศาสตร์			รวม (ข้อ)
		การคิด สถานการณ์ของ ปัญหาในเชิง คณิตศาสตร์	การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และ การให้เหตุผลทาง คณิตศาสตร์	การตีความการ ประยุกต์ใช้ และ การประเมิน ผลลัพธ์ทาง คณิตศาสตร์	
1. ปริภูมิและรูปทรง สามมิติ (Space and Shape)	1.1 ค่าใช้จ่ายการตัดหญ้า	1	1	1	3
	1.2 บรรจุนมกล่องลงในลัง กระดาษ	1	1	1	3
	1.3 ปูกระเบื้องพื้นห้องนอน	1	1	1	3
2. การเปลี่ยนแปลง และความสัมพันธ์ (Change and Relationships)	2.1 การทำไร่นาสวนผสมตาม หลักการเกษตรทฤษฎีใหม่ของ พระบาทสมเด็จพระเจ้าอยู่หัวฯ	1	1	1	3
	2.2 อัตราการเต้นของหัวใจ	1	1	1	3
	2.3 การโดยสารเครื่องบิน	1	1	1	3

ตาราง 5 (ต่อ)

เนื้อหา	สถานการณ์	กระบวนการทางคณิตศาสตร์			รวม (ข้อ)
		การคิด สถานการณ์ของ ปัญหาในเชิง คณิตศาสตร์	การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผล ทางคณิตศาสตร์	การตีความการ ประยุกต์ใช้ และ การประเมิน ผลลัพธ์ทาง คณิตศาสตร์	
3.ปริมาณ (Quantity)	3.1 น้ำแครอท	1	1	1	3
	3.2 เครื่องเล่น MP3	1	1	1	3
	3.3 จำนวนสัตว์แดรี่ฟาร์ม	1	1	1	3
4.ความไม่แน่นอน (Uncertainty)	4.1 การคิดค่าไฟฟ้า	1	1	1	3
	4.2 ราคาน้ำมัน	1	1	1	3
	4.3 ราคาขางพารา	1	1	1	3
รวม		12	12	12	36

แบบสอบการรู้เรื่องคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้นมีวิธีการตรวจให้คะแนน 2 รูปแบบ คือ กฎเกณฑ์การให้คะแนนแบบภาพรวม และกฎเกณฑ์การให้คะแนนแบบแยกส่วน

กฎเกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics)

การพิจารณาให้คะแนนของคำตอบในภาพรวม โดยที่กำหนดคะแนนเต็มสำหรับคำตอบที่มีความถูกต้องครบถ้วน คะแนนบางส่วนสำหรับคำตอบที่ถูกต้อง แต่มีข้อผิดพลาดบางส่วน

ตาราง 6 กฎเกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics)

กระบวนการทาง คณิตศาสตร์	3 คะแนน	2 คะแนน	1 คะแนน
1) การคิดสถานการณ์ของ ปัญหาในเชิงคณิตศาสตร์	สามารถวิเคราะห์สถานการณ์ และระบุได้ว่าโจทย์ต้องการ อะไรมีสิ่งใดกำหนด ให้บ้าง แล้วทำให้อยู่ในรูปร่างง่าย ตามแนวคิดหลักการ คณิตศาสตร์แล้วแปลงปัญหา โดยใช้ภาษาและ สัญลักษณ์ เครื่องหมาย หรือ หน่วยคณิตศาสตร์ ในการสื่อ ความหมาย มีการนำเสนอได้ อย่างถูกต้องและเหมาะสมกับ สถานการณ์	สามารถวิเคราะห์สถานการณ์ และระบุได้ว่าโจทย์ต้องการ อะไรมีสิ่งใดกำหนด ให้บ้าง แล้วทำให้อยู่ในรูปร่างง่าย ตามแนวคิดหลักการ คณิตศาสตร์ แต่ไม่สามารถ แปลงปัญหาโดยใช้ภาษา และสัญลักษณ์ เครื่องหมาย หรือหน่วยคณิตศาสตร์ ในการสื่อความหมาย ในการ นำเสนอ	สามารถวิเคราะห์ สถานการณ์และระบุได้ว่า โจทย์ต้องการอะไรมีสิ่งใด กำหนด ให้บ้าง แต่ไม่ สามารถทำให้อยู่ในรูปร่าง ง่ายตามแนวคิดหลักการ คณิตศาสตร์ และไม่สามารถ แปลงปัญหาโดยใช้ภาษา และสัญลักษณ์ เครื่องหมาย หรือหน่วยคณิตศาสตร์ ใน การสื่อความหมาย ในการ นำเสนอ

ตาราง 6 (ต่อ)

กระบวนการทาง คณิตศาสตร์	3 คะแนน	2 คะแนน	1 คะแนน
2) การใช้หลักการ/ กระบวนการทาง คณิตศาสตร์ในการ แก้ปัญหา	สามารถคิดวิธีการแก้ปัญหา มี การนำเครื่องมือทาง คณิตศาสตร์ เช่น กราฟ สถิติ นิพจน์พีชคณิตและสมการ และ เรขาคณิต มาใช้ในการนำเสนอ รวมทั้งมีการนำกฎเกณฑ์ ขั้นตอนวิธีการแก้ปัญหา และ โครงสร้างทางคณิตศาสตร์ เช่น การแกสมการ การอนุมานเชิง ตรรกศาสตร์ มาใช้ในการ แก้ปัญหา เพื่อให้ได้ผลลัพธ์ได้ อย่างถูกต้อง	สามารถคิดวิธีการแก้ปัญหา มี การนำเครื่องมือทาง คณิตศาสตร์ เช่น กราฟ สถิติ นิพจน์พีชคณิตและสมการ และ เรขาคณิต มาใช้ในการนำเสนอ รวมทั้งมีการนำกฎเกณฑ์ ขั้นตอนวิธีการแก้ปัญหาและ โครงสร้างทางคณิตศาสตร์ เช่น การแกสมการ การอนุมานเชิง ตรรกศาสตร์ มาใช้ในการ แก้ปัญหา แต่ผลลัพธ์ไม่ถูกต้อง	สามารถคิดวิธีการแก้ปัญหา มีการนำเครื่องมือทาง คณิตศาสตร์ เช่น กราฟ สถิติ นิพจน์พีชคณิตและ สมการ และเรขาคณิต มา ใช้ในการนำเสนอ แต่ไม่มี การนำกฎเกณฑ์ ขั้นตอน วิธีการแก้ปัญหา และ โครงสร้างทางคณิตศาสตร์ เช่น การแกสมการ การ อนุมานเชิงตรรกศาสตร์ มา ใช้ในการแก้ปัญหาเพื่อให้ ได้ผลลัพธ์
3) การตีความ และการ ประเมินผลลัพธ์ทาง คณิตศาสตร์	สามารถตีความผลลัพธ์ และ สามารถประเมินผลลัพธ์ที่ได้ โดยมีการแสดงความเป็นเหตุ เป็นผลของวิธีการแก้ปัญหา นำความรู้หลักการทาง คณิตศาสตร์มาอธิบายความ สมเหตุสมผลของผลลัพธ์ หรือข้อสรุปทางคณิตศาสตร์ ที่ได้ว่ามีความสอดคล้องและ เหมาะสมกับบริบทของปัญหา	สามารถตีความ ผลลัพธ์ และ สามารถประเมินผลลัพธ์ที่ได้ โดยมีการแสดงความเป็นเหตุ เป็นผลของวิธีการแก้ปัญหา นำความรู้หลักการ ทาง คณิตศาสตร์ มาอธิบายความ สมเหตุสมผลของผลลัพธ์ หรือข้อสรุปทางคณิตศาสตร์ ที่ได้ แต่ไม่มีความสอดคล้อง และเหมาะสมกับบริบทของ ปัญหา	สามารถตีความผลลัพธ์ และสามารถ ประเมิน ผลลัพธ์ที่ได้ แต่ไม่มีการ แสดงความเป็นเหตุเป็นผล ของวิธีการแก้ปัญหา ไม่มี การนำความรู้ตามหลักการ ทางคณิตศาสตร์ มาอธิบาย ความสมเหตุสมผลของ ผลลัพธ์หรือข้อสรุปทาง คณิตศาสตร์ที่ได้ไม่มีความ สอดคล้องและเหมาะสมกับ บริบทของปัญหา

กฎเกณฑ์การให้คะแนนแบบแยกส่วน (Analytic Scoring Rubrics)

การพิจารณาให้คะแนนของคำตอบในแต่ละประเด็นย่อย และในแต่ละประเด็นมีการกำหนดระดับคะแนนแตกต่างกัน ขึ้นอยู่กับน้ำหนักความสำคัญของคำตอบ

ตาราง 7 กฎเกณฑ์การให้คะแนนแบบแยกส่วน (Analytic Scoring Rubrics)

กระบวนการทาง คณิตศาสตร์	3 คะแนน	2 คะแนน	1 คะแนน
1) การคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์			
1.1) ระบุประเด็นทางคณิตศาสตร์ของปัญหาในชีวิตจริง	สามารถวิเคราะห์สถานการณ์และระบุได้ว่าโจทย์ต้องการอะไร มีสิ่งใดกำหนดให้บ้าง	สามารถวิเคราะห์สถานการณ์และระบุได้ว่าโจทย์ต้องการอะไร มีสิ่งใดกำหนดให้บ้าง แต่ระบุไม่ครบถ้วน	ไม่สามารถวิเคราะห์สถานการณ์และระบุไม่ได้ว่าโจทย์ต้องการอะไร มีสิ่งใดกำหนดให้บ้าง
1.2) ทำสถานการณ์หรือปัญหาให้อยู่ในรูปอย่างง่าย เพื่อให้การวิเคราะห์ทางคณิตศาสตร์ง่ายขึ้น	สามารถวิเคราะห์สถานการณ์แล้วทำให้อยู่ในรูปอย่างง่ายที่รวบรวมได้จากบริบทตามแนวคิดหลักการคณิตศาสตร์บนพื้นฐานข้อเท็จจริง	สามารถวิเคราะห์สถานการณ์แล้วทำให้อยู่ในรูปอย่างง่ายที่รวบรวมได้จากบริบท ไม่เป็นไปตามแนวคิดหลักการคณิตศาสตร์บนพื้นฐานข้อเท็จจริง	ไม่สามารถวิเคราะห์สถานการณ์แล้วทำให้อยู่ในรูปอย่างง่ายที่รวบรวมได้จากบริบท ไม่มีแนวคิดตามหลักการคณิตศาสตร์บนพื้นฐานข้อเท็จจริง
1.3) แปลงปัญหาให้อยู่ในรูปของคณิตศาสตร์	แปลงปัญหาโดยใช้ภาษาและสัญลักษณ์ เครื่องหมาย หรือหน่วยทางคณิตศาสตร์ในการสื่อความหมาย ในการนำเสนอได้อย่างถูกต้องและเหมาะสมกับสถานการณ์	แปลงปัญหาโดยใช้ภาษาและสัญลักษณ์ เครื่องหมาย หรือหน่วยทางคณิตศาสตร์ในการสื่อความหมาย แต่มีการนำเสนอไม่ถูกต้องและไม่เหมาะสมกับสถานการณ์	ไม่มีการแปลงปัญหา ไม่มีการใช้ภาษาและสัญลักษณ์ เครื่องหมาย หรือหน่วยที่ใช้ทางคณิตศาสตร์ในการสื่อความหมาย และการนำเสนอ

ตาราง 7 (ต่อ)

กระบวนการทาง คณิตศาสตร์	3 คะแนน	2 คะแนน	1 คะแนน
2) การใช้หลักการ/กระบวนการทางคณิตศาสตร์ในการแก้ปัญหา			
2.1) คิดและนำกลยุทธ์ในการ หาวิธีแก้ปัญหาทาง คณิตศาสตร์ไปใช้	สามารถคิดวิธีการแก้ปัญหา โดยอาศัยหลักการหรือ วิธีการทางคณิตศาสตร์ มาใช้ในการแก้ปัญหา เพื่อให้ ได้ผลลัพธ์ที่ถูกต้อง	สามารถคิดวิธีการแก้ปัญหา โดยอาศัยหลักการหรือ วิธีการทางคณิตศาสตร์ มาใช้ในการแก้ปัญหา เพื่อให้ ได้ผลลัพธ์ แต่ผลลัพธ์ที่ได้ไม่ ถูกต้อง	ไม่สามารถคิดวิธีการ แก้ปัญหาโดยอาศัยหลักการ หรือวิธีการทางคณิตศาสตร์ มาใช้ในการแก้ปัญหา เพื่อให้ ได้ผลลัพธ์
2.2) ใช้เครื่องมือทาง คณิตศาสตร์ เพื่อช่วยให้ วิธีการแก้ปัญหาถูกต้องหรือ เหมาะสม	มีการนำเครื่องมือทาง คณิตศาสตร์ เช่น กราฟ สถิติ นิพจน์พีชคณิตและสมการ และเรขาคณิต มาใช้ในการ นำเสนอได้อย่างถูกต้องและ เหมาะสมกับสถานการณ์	มีการนำเครื่องมือทาง คณิตศาสตร์ เช่น กราฟ สถิติ นิพจน์พีชคณิตและสมการ และเรขาคณิต มาใช้ในการ นำเสนอ แต่ไม่ถูกต้องและไม่ เหมาะสมกับสถานการณ์	ไม่มีการนำเครื่องมือทาง คณิตศาสตร์ เช่น กราฟ สถิติ นิพจน์พีชคณิตและสมการ และเรขาคณิต มาใช้ในการ นำเสนอ
2.3) นำกฎเกณฑ์ ขั้นตอนวิธี และโครงสร้างคณิตศาสตร์ไป ใช้แก้ปัญหา	มีการนำกฎเกณฑ์ ขั้นตอน วิธีการแก้ปัญหา และ โครงสร้างทางคณิตศาสตร์ เช่น การแก้สมการ การ อนุมานเชิงตรรกศาสตร์ มา ใช้ในการแก้ปัญหา เพื่อให้ได้ ผลลัพธ์ได้อย่างถูกต้อง	มีการนำกฎเกณฑ์ ขั้นตอน วิธีการแก้ปัญหา และ โครงสร้างทางคณิตศาสตร์ เช่น การแก้สมการ การ อนุมานเชิงตรรกศาสตร์ มา ใช้ในการแก้ปัญหา แต่ ผลลัพธ์ที่ได้ไม่ถูกต้อง	ไม่มีการนำกฎเกณฑ์ขั้นตอน วิธีการแก้ปัญหา และ โครงสร้างทางคณิตศาสตร์ เช่น การแก้สมการ การ อนุมานเชิงตรรกศาสตร์ มา ใช้ในการแก้ปัญหา เพื่อให้ได้ ผลลัพธ์

ตาราง 7 (ต่อ)

กระบวนการรู้เรื่อง คณิตศาสตร์	3 คะแนน	2 คะแนน	1 คะแนน
3) การตีความ และการประเมิน ผลลัพธ์ทางคณิตศาสตร์			
3.1) ตีความผลลัพธ์ทางคณิตศาสตร์ กลับไปสู่บริบทชีวิตความเป็นจริง	สามารถตีความผลลัพธ์โดยสะท้อนวิธีการแก้ปัญหาของสถานการณ์ ภายใต้การคิดคำนวณบนพื้นฐานคณิตศาสตร์ เพื่อหาผลลัพธ์หรือข้อสรุปที่ถูกต้อง	สามารถตีความผลลัพธ์โดยสะท้อนวิธีการแก้ปัญหาของสถานการณ์ ภายใต้การคิดคำนวณบนพื้นฐานคณิตศาสตร์ ผลลัพธ์หรือข้อสรุปที่ได้ไม่ถูกต้อง	ไม่สามารถตีความผลลัพธ์โดยสะท้อนวิธีการแก้ปัญหาของสถานการณ์ ภายใต้การคิดคำนวณบนพื้นฐานคณิตศาสตร์ เพื่อหาผลลัพธ์หรือข้อสรุป
3.2) ประเมินความเป็นเหตุเป็นผลของวิธีการแก้ปัญหาทางคณิตศาสตร์ ในบริบทของปัญหาชีวิตจริง	สามารถประเมินผลลัพธ์ที่ได้ โดยมีการแสดงความเป็นเหตุเป็นผลของวิธีการแก้ปัญหาทางคณิตศาสตร์ และตัดสินได้ว่าผลลัพธ์ได้สามารถเข้ากันได้กับสถานการณ์	สามารถประเมินผลลัพธ์ที่ได้ โดยมีการแสดงความเป็นเหตุเป็นผลของวิธีการแก้ปัญหาทางคณิตศาสตร์ แต่ไม่สามารถตัดสินได้ว่าผลลัพธ์ที่ได้สามารถเข้ากันได้กับสถานการณ์	ไม่สามารถประเมินผลลัพธ์ที่ได้ ไม่มีการแสดงความเป็นเหตุเป็นผลของวิธีการแก้ปัญหาทางคณิตศาสตร์ และไม่สามารถตัดสินได้ว่าผลลัพธ์ที่ได้สามารถเข้ากันได้กับสถานการณ์
3.3) อธิบายความสมเหตุสมผลของผลลัพธ์หรือข้อสรุปทางคณิตศาสตร์กับบริบทของปัญหา	นำความรู้หลักการทางคณิตศาสตร์ มาอธิบายความสมเหตุสมผลของผลลัพธ์หรือข้อสรุปทางคณิตศาสตร์ ที่ได้ว่ามีความสอดคล้องและเหมาะสมกับบริบทของปัญหา	นำความรู้หลักการทางคณิตศาสตร์ มาอธิบายความสมเหตุสมผลของผลลัพธ์หรือข้อสรุปทางคณิตศาสตร์ ที่ได้ แต่ไม่มีความสอดคล้องและไม่เหมาะสมกับบริบทของปัญหา	ไม่มีการนำความรู้หลักการทางคณิตศาสตร์มาอธิบายความสมเหตุสมผลของผลลัพธ์หรือข้อสรุปทางคณิตศาสตร์ที่ได้

ตัวอย่างเครื่องมือที่ใช้ในการวิจัย

สถานการณ์ : พีระมิด



ชาวอียิปต์ได้สร้างพีระมิดที่มีฐานเป็นรูปด้านเท่ามุมเท่า มีสันยาวเท่ากันทุกเส้น มีสูงเอียงทุกเส้นยาวเท่ากันและส่วนสูงตั้งฉากกับฐาน ABCD เป็นรูปสี่เหลี่ยมจัตุรัส เป็นฐานที่รองรับน้ำหนักของพีระมิด พีระมิดทุกด้านยาวเท่ากัน 12 เมตร โดย E เป็นจุดยอดของพีระมิดวัดจากฐานสูง 6 เมตร

สูตรการคำนวณหาปริมาตรของพีระมิด

$$\frac{1}{3} \times \text{พื้นที่ฐาน} \times \text{สูง}$$

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากสถานการณ์ข้างต้น นักเรียนวาดภาพพีระมิด พร้อมทั้งระบุขนาดตามที่โจทย์กำหนดให้

สำหรับวาดพีระมิด	แสดงวิธีทำ
	
	
	
	พีระมิดมีปริมาตร

ข้อที่ 3 : พีระมิด คือ วัตถุทรงสามมิติซึ่งฐานเป็นรูปหลายเหลี่ยม มีหน้าทุกหน้า (ผิวข้าง) เป็นรูปสามเหลี่ยมโดยมีจุดยอดรวมกันที่จุดหนึ่ง ทั้งนี้ยอดไม่อยู่ในระนาบเดียวกับฐาน



มาโนชเดินผ่านเต็นท์ที่ลักษณะแบบภาพ
ดังกล่าวส่วนหลังคาเต็นท์ดังกล่าวจัดเป็น
พีระมิดหรือไม่ เพราะเหตุใด
ตอบ ☐ เป็นพีระมิด / ☐ ไม่เป็นพีระมิด

เพราะ

.....

ขั้นตอนการสร้างแบบสอบการรู้เรื่องคณิตศาสตร์ (Mathematics Literacy)



ภาพประกอบ 9 ขั้นตอนดำเนินการสร้างแบบสอบการรู้เรื่องคณิตศาสตร์ (Mathematics Literacy)

ในการสร้างแบบสอบวัดการรู้เรื่องคณิตศาสตร์ จากการศึกษาเอกสารหลักสูตรและการวิเคราะห์หลักสูตร เพื่อที่จะระบุจุดมุ่งหมายในการชี้แนวทางการเลือกตัวแทนของเนื้อหาที่จะออกข้อสอบว่าออกข้อสอบอย่างไรจึงจะเป็นตัวแทนที่ดีของเนื้อหา วิธีการนี้จะเป็นหลักประกันได้ว่าแบบทดสอบนั้นจะมีความเที่ยงตรงด้านเนื้อหา (Content validity) มีการดำเนินการสร้างแบบทดสอบ ดังนี้

1. ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการสร้างแบบสอบวัดผลการรู้เรื่องคณิตศาสตร์ (Mathematics Literacy) เพื่อเป็นแนวทางการสร้างแบบทดสอบ
2. ศึกษาหลักสูตรคู่มือครูและแบบเรียนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 เพื่อวิเคราะห์เนื้อหา จุดประสงค์การเรียนรู้ สาระและทักษะการคิดคำนวณเกี่ยวกับการรู้เรื่องคณิตศาสตร์ (Mathematics Literacy)
3. สร้างตารางวิเคราะห์เนื้อหา โดยกำหนดเนื้อหา จุดประสงค์เชิงพฤติกรรมให้สอดคล้องกับพฤติกรรมที่ต้องการวัด
4. สร้างแบบสอบตามรูปแบบแบบสอบการรู้เรื่องคณิตศาสตร์ มีกรอบโครงสร้างการรู้เรื่องคณิตศาสตร์ครอบคลุม 4 สาระ แต่ละสาระประกอบด้วยสถานการณ์ 5 สถานการณ์ แต่ละสถานการณ์แบ่งตามกระบวนการคณิตศาสตร์ 3 กระบวนการ ได้แบบสอบการรู้เรื่องคณิตศาสตร์ จำนวน 60 ข้อ
5. ตรวจสอบคุณภาพเครื่องมือ
 - 5.1 นำเครื่องมือที่สร้างขึ้นให้ผู้เชี่ยวชาญ จำนวน 5 คน เป็นผู้เชี่ยวชาญด้านการวัดและประเมินผล 2 ท่าน และผู้เชี่ยวชาญด้านเนื้อหาคณิตศาสตร์ 3 ท่าน (ดูจากภาคผนวก ก) ตรวจสอบคุณภาพเบื้องต้น ด้านความเที่ยงตรงด้านเนื้อหา (Content Validity) เพื่อให้ผู้เชี่ยวชาญประเมินว่าข้อสอบแต่ละข้อนั้น สามารถวัดได้ตรงกับเนื้อหาหรือไม่ โดยวิธีการหาค่าดัชนีความสอดคล้องเป็นรายข้อ (Index of Congruency: IOC)
 - 5.2 นำแบบสอบไปทดลองใช้กับนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 80 คน
 - 5.3 นำแบบสอบมาตรวจให้คะแนนตามกฎเกณฑ์การให้คะแนนแบบภาพรวมและกฎเกณฑ์การให้คะแนนแบบแยกส่วน
 - 5.4 นำผลการวิเคราะห์คุณภาพแบบสอบ เพื่อหาค่าความยากง่ายโดยพิจารณาคัดเลือกข้อสอบที่มีค่าความยากอยู่ระหว่าง 0.20 ถึง 0.80 ที่ตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนน ทั้ง 2 รูปแบบ
 - 5.5 หาค่าอำนาจจำแนกของข้อสอบของแบบทดสอบชนิดอัตนัย ใช้การแบ่งกลุ่มผู้ตอบกลุ่มสูงกลุ่มต่ำเทคนิค 25 %
 - 5.6 หาค่าความเชื่อมั่นสัมประสิทธิ์แอลฟา (α – Coefficient) ของครอนบาค
6. คัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์จำนวน 36 ข้อ แล้วจัดพิมพ์เพื่อนำไปใช้กับนักเรียนกลุ่มตัวอย่าง

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในครั้งนี้ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลด้วยตนเองในเดือนพฤศจิกายนถึงเดือนธันวาคม พ.ศ.2560 โดยมีการดำเนินการเก็บรวบรวมข้อมูล ดังนี้

1. ทำหนังสือจากบัณฑิตวิทยาลัย เพื่อขอความอนุเคราะห์ผู้อำนวยการโรงเรียนที่เป็นกลุ่มตัวอย่าง
2. ประสานงานและชี้แจงกับผู้ผู้อำนวยการโรงเรียนและครูผู้สอนรายวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนที่เป็นกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลาและสถานที่สอบ
3. ดำเนินการสอบ โดยนำแบบสอบถามการรู้เรื่องคณิตศาสตร์ ไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างของแต่ละโรงเรียน ก่อนให้นักเรียนลงมือทำแบบสอบ ผู้วิจัยได้ชี้แจงอธิบายให้นักเรียนเข้าใจเกี่ยวกับวิธีการทำแบบสอบ
4. นำแบบสอบทั้งหมดมาตรวจตามกฎเกณฑ์การให้คะแนนที่ผู้วิจัยสร้างขึ้น
5. นำผลการตรวจคะแนนที่ได้มาทำการวิเคราะห์ข้อมูลเพื่อตอบความมุ่งหมายของการวิจัย

การจัดกระทำข้อมูลและการวิเคราะห์ข้อมูล

1. คัดเลือกข้อสอบมีคุณสมบัติที่เหมาะสม จำนวน 36 ข้อ ไปเก็บข้อมูลกับกลุ่มตัวอย่าง แล้วนำผลที่ได้มาตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวมและกฎเกณฑ์การให้คะแนนแบบแยกส่วน ซึ่งการวิจัยครั้งนี้ออกแบบการวัดเป็นแบบ One - facet design โดยมีรูปแบบการวัดเป็น $p \times i$ เมื่อ p แทน ผู้สอบ และ i แทน จำนวนข้อสอบ โดยผู้ตรวจดำเนินการตรวจข้อสอบทุกข้อของผู้สอบทุกคน ($p \times i$ design) ซึ่งมีรายละเอียดดังนี้

ความยาวแบบสอบ 36 ข้อ ที่มีกฎเกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics) และกฎเกณฑ์การให้คะแนนแบบแยกส่วน (Analytic Scoring Rubrics) ของผู้สอบทุกคน

P	i1		i2		...		i36	
	Holistic	Analytic	Holistic	Analytic	Holistic	Analytic	Holistic	Analytic
1								
2								
...								
400								

การจัดกระทำข้อมูลตามเงื่อนไขและระดับของเงื่อนไขที่ศึกษา

รูปแบบกฎเกณฑ์การให้คะแนน (Scoring rubrics)	เงื่อนไขที่ศึกษา (ความยาวแบบสอบ)		
กฎเกณฑ์การให้คะแนนแบบภาพรวม (Holistic Scoring Rubrics)	12 ข้อ	24 ข้อ	36 ข้อ
กฎเกณฑ์การให้คะแนนแบบแยกส่วน (Analytic Scoring Rubrics)	12 ข้อ	24 ข้อ	36 ข้อ

3. การหาคุณภาพแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด โดยตรวจสอบคุณภาพด้านความเชื่อมั่นค่าจากสัมประสิทธิ์การสรุปอ้างอิง และเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบที่มีความยาวต่างกัน เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนรูปแบบเดียวกัน

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ผู้วิจัยได้ใช้สถิติพื้นฐานและสถิติที่ใช้ตรวจสอบคุณภาพเครื่องมือ โดยมีรายละเอียด ดังนี้

1. สถิติพื้นฐาน

1.1 ค่าเฉลี่ย ($\bar{X}$)

1.2 ส่วนเบี่ยงเบนมาตรฐาน (SD)

2. สถิติที่ใช้ตรวจสอบหาคุณภาพเครื่องมือ

2.1 ตรวจสอบความสอดคล้องของผู้เชี่ยวชาญ โดยพิจารณาหาค่าดัชนีความสอดคล้อง (Index of Congruency: IOC) โดยใช้สูตรของโรวินีลลีและแฮมเบิลตัน (ล้วน สายยศ; และอังคณา สายยศ. 2539: 249; อ้างอิงจาก Rowineli; & Hambleton.1977)

$$IOC = \frac{\sum R}{N}$$

เมื่อ	IOC	แทน ดัชนีความสอดคล้องมีค่าอยู่ระหว่าง -1 ถึง +1
	$\sum R$	แทน ผลรวมของการพิจารณาของผู้เชี่ยวชาญทั้งหมด
	N	แทน จำนวนผู้เชี่ยวชาญทั้งหมด

2.2 ค่าความยากง่าย (Difficulty index) ของแบบทดสอบชนิดอัตนัย ใช้ค่าดัชนีค่าความยากง่ายตามสูตรของวิทย์นีย์และซาร์เบอร์ (สมศักดิ์ สินธุระเวชญ์. 2530: 179; อ้างอิงจาก Whitney; & Sabers. 1970: 214 - 215)

$$p_D = \frac{S_U + S_L - (2n X_{\min})}{2n (X_{\max} - X_{\min})}$$

เมื่อ	p_D	แทน ค่าดัชนีค่าความยากง่าย
	S_U	แทน ผลรวมคะแนนกลุ่มเก่ง
	S_L	แทน ผลรวมของคะแนนกลุ่มอ่อน
	n	แทน จำนวนผู้เข้าสอบของกลุ่มเก่งหรือกลุ่มอ่อน
	$X_{\max}$	แทน คะแนนที่นักเรียนทำได้สูงสุด
	$X_{\min}$	แทน คะแนนที่นักเรียนทำได้ต่ำสุด

2.3 ค่าอำนาจจำแนก (Discriminant index) ของแบบทดสอบชนิดอัตนัย ใช้การแบ่งกลุ่มผู้ตอบ กลุ่มสูงกลุ่มต่ำเทคนิค 25 % ตามสูตรของวิทย์นีย์และซาร์เบอร์ (ล้วน สายยศ; และ อังคณา สายยศ. 2539: 201 อ้างอิงจาก Whitney; & Sabers. 1970: 214 - 215)

$$D = \frac{S_U - S_L}{n(X_{\max} - X_{\min})}$$

เมื่อ	D	แทน ดัชนีค่าอำนาจจำแนก
	S_U	แทน ผลรวมคะแนนกลุ่มเก่ง
	S_L	แทน ผลรวมของคะแนนกลุ่มอ่อน
	n	แทน จำนวนผู้เข้าสอบของกลุ่มเก่งหรือกลุ่มอ่อน
	$X_{\max}$	แทน คะแนนที่นักเรียนทำได้สูงสุด
	$X_{\min}$	แทน คะแนนที่นักเรียนทำได้ต่ำสุด

2.4 ความเชื่อมั่น (Reliability) โดยวิธีสัมประสิทธิ์แอลฟา (α - Coefficient) ของครอนบาค (Cronbach alpha procedure) (ล้วน สายยศ; และอังคณา สายยศ. 2539: 218)

$$\alpha = \frac{k}{k-1} \left\{ 1 - \frac{\sum \sigma_i^2}{\sigma^2} \right\}$$

เมื่อ α แทน สัมประสิทธิ์ความเชื่อมั่น
 k แทน จำนวนข้อสอบ
 σ_i^2 แทน คะแนนความแปรปรวนเป็นรายข้อ

2.5 ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (ρ_{Abs}^2)
 คำนวณจากสูตรของชเวลสัน; และแวนบี (Shavelson; & Wabb. 1991: 96)

$$\rho_{Abs}^2 = \frac{\sigma_p^2}{n_i} + \frac{\sigma_i^2}{n_i} + \frac{\sigma_{pi}^2}{n_i} + \frac{\sigma_{pi,e}^2}{n_i}$$

เมื่อ ρ_{Abs}^2 แทน สัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์
 σ_p^2 แทน ความแปรปรวนของคะแนนเอกภพของผู้สอบ
 σ_i^2 แทน ความแปรปรวนของคะแนนเฉลี่ยข้อสอบ
 σ_{pi}^2 แทน ความแปรปรวนของคะแนนปฏิสัมพันธ์ผู้สอบกับข้อสอบ
 $\sigma_{pi,e}^2$ แทน ความแปรปรวนของคะแนนคลาดเคลื่อนหรือส่วนที่เหลือ

2.6 ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (ρ_{Rel}^2)
 คำนวณจากสูตรของชเวลสัน; และแวนบี (Shavelson; & Wabb. 1991: 96)

$$\rho_{Rel}^2 = \frac{\sigma_{pi}^2}{n_i} + \frac{\sigma_{pi,e}^2}{n_i}$$

เมื่อ ρ_{Rel}^2 แทน สัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์
 σ_{pi}^2 แทน ความแปรปรวนของคะแนนปฏิสัมพันธ์ผู้สอบกับข้อสอบ
 $\sigma_{pi,e}^2$ แทน ความแปรปรวนของคะแนนคลาดเคลื่อนหรือส่วนที่เหลือ

3. สถิติที่ใช้ในการทดสอบสมมุติฐาน

ทดสอบความแตกต่างค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสอบการรู้เรื่องคณิตศาสตร์ที่ตรวจโดยใช้กฎเกณฑ์การให้คะแนนรูปแบบเดียวกัน แต่มีความยาวแบบสอบต่างกัน โดยแปลงค่าสัมประสิทธิ์การสรุปอ้างอิงเป็นคะแนนมาตรฐาน (Fisher's Z) แล้วทดสอบด้วยค่าไค-สแควร์ (Chi - Square) ของเวอร์ส; ไนส์; และอาร์แมน (Wert; Neidt; & Ahmann. 1954: 296)

$$Z = \frac{Z_1 + Z_2}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}$$

เมื่อ Z แทน คะแนนมาตรฐานของโค้งปกติ
 Z_1 Z_2 แทน คะแนนมาตรฐาน (Fisher's Z) ที่ได้จากค่าความเชื่อมั่นที่ 1 และ
 ความเชื่อมั่นที่ 2 ตามลำดับ
 N_1 N_2 แทน ขนาดกลุ่มตัวอย่างในกลุ่มที่ 1 และ 2 ตามลำดับ

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้ เป็นการวิจัยเพื่อศึกษาคุณภาพของแบบสอบถามรู้เรื่องคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบถามและกฎเกณฑ์การให้คะแนนต่างกัน โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) เพื่อให้เกิดความเข้าใจตรงกันในการนำเสนอ และการแปลผลการวิเคราะห์ข้อมูล ผู้วิจัยจึงได้ กำหนดสัญลักษณ์ที่ใช้ในการวิเคราะห์ข้อมูลดังนี้

สัญลักษณ์และอักษรย่อในการวิเคราะห์ข้อมูล

การนำเสนอผลการวิเคราะห์ครั้งนี้ ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อ ดังนี้

n	แทน	จำนวนนักเรียนในกลุ่มตัวอย่าง
$\bar{X}$	แทน	ค่าเฉลี่ย (Mean)
SD	แทน	ส่วนเบี่ยงเบนมาตรฐาน
H	แทน	เกณฑ์การให้คะแนนแบบภาพรวม
A	แทน	เกณฑ์การให้คะแนนแบบแยกส่วน
p	แทน	ค่าความยากง่าย
r	แทน	ค่าอำนาจจำแนก
α	แทน	ค่าสัมประสิทธิ์ความเชื่อมั่น
df	แทน	องศาความเป็นอิสระ (degrees of freedom)
SS	แทน	ผลบวกกำลังสอง (Sum of Square)
MS	แทน	ค่าเฉลี่ยผลบวกกำลังสอง (Mean Square)
ρ^2_{Abs}	แทน	ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์
ρ^2_{Rel}	แทน	ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์
σ^2_p	แทน	ความแปรปรวนของคะแนนเอกภพ
σ^2_{Abs}	แทน	ความแปรปรวนของคะแนนความคลาดเคลื่อนแบบสัมบูรณ์
σ^2_{Rel}	แทน	ความแปรปรวนของคะแนนความคลาดเคลื่อนแบบสัมพัทธ์
Z_r	แทน	ค่าความเชื่อมั่นที่แปลงเป็นคะแนนมาตรฐานซี โดยวิธีการฟิชเชอร์ซี (Fisher's Z tranformation)
Z	แทน	การทดสอบความแตกต่างของคะแนนมาตรฐานซี (Z – test)
*p < .05	แทน	มีนัยสำคัญทางสถิติที่ระดับ .05

การนำเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาครั้งนี้ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับดังนี้

ตอนที่ 1 การตรวจสอบคุณภาพเบื้องต้นของแบบสอบถามการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน

1.1 ผลการตรวจสอบความสอดคล้องของผู้เชี่ยวชาญ (IOC)

1.2 ผลการทดลองใช้ (try out) วิเคราะห์ค่าความยากง่าย ค่าอำนาจจำแนก และความเชื่อมั่น

1.3 ค่าสถิติพื้นฐานคะแนนจากแบบสอบถามการรู้เรื่องคณิตศาสตร์

ตอนที่ 2 คุณภาพของแบบสอบถามการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน จากการคำนวณค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบถามการรู้เรื่องคณิตศาสตร์

2.1 ผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบถามการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบต่างกันที่ใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม

2.2 การศึกษาเชิงตัดสินใจการสรุปอ้างอิงที่ได้จากแบบสอบถามการรู้เรื่องคณิตศาสตร์ที่มีความยาว แบบสอบต่างกันและตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม

2.3 ผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบถามการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบต่างกันที่ใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน

2.4 การศึกษาเพื่อการตัดสินใจการสรุปอ้างอิงที่ได้จากแบบสอบถามการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบต่างกันและตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน

2.5 การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบถามการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบต่างกัน

ผลการวิเคราะห์ข้อมูล

ผู้วิจัยขอเสนอผลการวิเคราะห์ข้อมูล มีรายละเอียดดังต่อไปนี้

ตอนที่ 1 การตรวจสอบคุณภาพเบื้องต้นของแบบสอบถามการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน

1.1 ผลการตรวจสอบความสอดคล้องของผู้เชี่ยวชาญ (IOC)

จากการศึกษาทฤษฎีและเอกสารและงานวิจัยที่เกี่ยวข้องกับการรู้เรื่องคณิตศาสตร์ นำมาเป็นกรอบแนวคิดในการสร้างข้อคำถามวัดกระบวนการรู้เรื่องคณิตศาสตร์ 1 ฉบับ จำนวน 36 ข้อ ผู้วิจัยได้นำแบบสอบที่สร้างขึ้นตรวจสอบคุณภาพเบื้องต้น ด้านความเที่ยงตรงเชิงเนื้อหาเกี่ยวกับความเหมาะสมของข้อคำถามและความสอดคล้องกับนิยามศัพท์โดยผู้เชี่ยวชาญ 5 ท่าน พิจารณาข้อคำถามที่มีความเที่ยงตรงเชิงเนื้อหาเหมาะสม จะต้องมีความสอดคล้อง (Index of Item-Objective Congruence : IOC) มีค่าตั้งแต่ 0.50 ขึ้นไป ผลการตรวจสอบ พบว่า แบบสอบถามการรู้เรื่องคณิตศาสตร์ มีค่าดัชนีความสอดคล้อง ตั้งแต่ 0.60 – 1.00 ข้อสอบที่ผ่านเกณฑ์ทั้งหมดจำนวน 60 ข้อ (ดูจากภาคผนวก ข)

1.2 ผลการทดลองใช้ (try out) วิเคราะห์ค่าความยากง่าย ค่าอำนาจจำแนก และความเชื่อมั่น

ผู้วิจัยนำแบบทดสอบไปให้นักเรียนที่ไม่ใช่กลุ่มตัวอย่าง โดยทดสอบกับนักเรียน
ชั้นมัธยมศึกษาปีที่ 2 จำนวน 80 คน ปรากฏผลดังแสดงในตาราง 8

ตาราง 8 ค่าความยากง่าย อำนาจจำแนก และความเชื่อมั่นแบบสอบการรู้เรื่องคณิตศาสตร์

เนื้อหา	กฎเกณฑ์ การให้ คะแนน	ก่อนคัดเลือก				หลังคัดเลือก			
		จำนวน ข้อ	p	r	α	จำนวน ข้อ	p	r	α
1. ปริภูมิและรูปทรง สามมิติ	H	15	0.44 - 0.73	0.12 - 0.47	0.877	9	0.39 - 0.68	0.21 - 0.47	0.848
	A		0.39 - 0.78	0.07 - 0.50	0.911		0.49 - 0.78	0.20 - 0.50	0.875
2. การเปลี่ยนแปลง และความสัมพันธ์	H	15	0.26 - 0.71	0.13 - 0.47	0.847	9	0.33 - 0.71	0.21 - 0.47	0.859
	A		0.38 - 0.80	0.05 - 0.45	0.859		0.38 - 0.78	0.22 - 0.45	0.850
3. ปริมาณ	H	15	0.06 - 0.70	0.06 - 0.48	0.856	9	0.23 - 0.70	0.22 - 0.48	0.842
	A		0.06 - 0.72	0.06 - 0.47	0.904		0.28 - 0.73	0.21 - 0.47	0.919
4. ความไม่แน่นอน	H	15	0.39 - 0.65	0.15 - 0.41	0.804	9	0.39 - 0.61	0.22 - 0.41	0.834
	A		0.33 - 0.69	0.11 - 0.44	0.912		0.33 - 0.63	0.20 - 0.44	0.901
ภาพรวม	H	60	0.06 - 0.73	0.06 - 0.48	0.965	36	0.23 - 0.71	0.21 - 0.48	0.952
	A		0.06 - 0.80	0.05 - 0.50	0.969		0.28 - 0.78	0.20 - 0.50	0.960

จากตาราง 8 พบว่า จากแบบสอบการรู้เรื่องคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้นทั้งหมด
60 ข้อ มีข้อคำถามที่ผ่านเกณฑ์การคัดเลือก จำนวน 36 ข้อ เมื่อตรวจให้คะแนน
โดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวมพบว่าแบบสอบมีค่าความยากง่ายอยู่ระหว่าง
0.23 - 0.71 ค่าอำนาจจำแนกอยู่ระหว่าง 0.21 - 0.48 และมีค่าความเชื่อมั่นเท่ากับ 0.952
เมื่อตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วนพบว่าแบบสอบ
มีค่าความยากง่ายอยู่ระหว่าง 0.28 - 0.78 ค่าอำนาจจำแนกอยู่ระหว่าง 0.20 - 0.50 และ
มีค่าความเชื่อมั่นเท่ากับ 0.960

1.3 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบการรู้เรื่องคณิตศาสตร์
ปรากฏผลดังแสดงในตาราง 9 - 13

ตาราง 9 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบการรู้เรื่องคณิตศาสตร์ จำแนกตาม
กระบวนการรู้เรื่องคณิตศาสตร์

เนื้อหาสาระ	กฎเกณฑ์ การให้ คะแนน	คะแนน เต็ม	การคิด สถานการณ์ของ ปัญหาในเชิง คณิตศาสตร์		การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และ การให้เหตุผลทาง คณิตศาสตร์		การตีความการ ประยุกต์ใช้ และ การประเมินผลลัพธ์ ทางคณิตศาสตร์	
			$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
1.ปริภูมิและรูปทรง สามมิติ	H	9	5.62	1.44	5.64	2.04	4.56	1.73
	A	27	20.10	3.62	19.83	5.21	17.49	5.24
2.การเปลี่ยนแปลง และความสัมพันธ์	H	9	5.98	2.37	5.76	1.90	4.68	1.79
	A	27	21.05	5.43	20.36	4.86	17.43	5.43
3.ปริมาณ	H	9	5.61	1.96	4.77	1.61	4.98	1.95
	A	27	20.08	4.74	18.07	4.84	18.45	5.37
4.ความไม่แน่นอน	H	9	4.03	0.96	5.22	2.00	4.76	1.82
	A	27	15.28	3.63	18.05	5.67	17.11	5.25

จากตาราง 9 เมื่อพิจารณาคะแนนที่ได้จากแบบสอบการรู้เรื่องคณิตศาสตร์ตาม
กระบวนการรู้เรื่องคณิตศาสตร์ เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวมและแบบ
แยกส่วน พบว่า กระบวนการคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ มีค่าเฉลี่ยสูงสุด
รองลงมาคือกระบวนการใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และกระบวนการให้เหตุผลทาง
คณิตศาสตร์ และการตีความการประยุกต์ใช้ และการประเมิน ผลลัพธ์ทางคณิตศาสตร์
ตามลำดับ

ตาราง 10 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบถามรู้เรื่องคณิตศาสตร์ เรื่องปริภูมิและรูปทรงสามมิติ จำแนกตามกฎเกณฑ์การให้คะแนน

สถานการณ์	กฎเกณฑ์การให้คะแนน	จำนวนนักเรียนทั้งหมด	คะแนนเต็ม	ข้อที่ 1 การคิด สถานการณ์ ของปัญหาใน เชิง คณิตศาสตร์		ข้อที่ 2 การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และ การให้เหตุผลทาง คณิตศาสตร์		ข้อที่ 3 การตีความ การประยุกต์ใช้ และการ ประเมินผลลัพธ์ ทาง คณิตศาสตร์	
				$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
สถานการณ์ที่ 1 : ค่าใช้จ่าย การตัดหญ้า	H	400	3	1.75	0.45	2.01	0.86	1.62	0.72
	A	400	9	6.63	0.95	7.14	1.83	6.34	1.87
สถานการณ์ที่ 2 : บรรจุนมกล่องลง ในลังกระดาษ	H	400	3	2.12	0.71	1.80	0.76	1.72	0.83
	A	400	9	7.45	1.46	6.39	1.96	5.99	2.30
สถานการณ์ที่ 3 : ปูกระเบื้องพื้น ห้องนอน	H	400	3	1.75	0.69	1.46	0.66	1.35	0.61
	A	400	9	6.48	1.91	5.53	1.89	5.48	1.80

จากตาราง 10 คะแนนที่ได้จากแบบสอบถามรู้เรื่องคณิตศาสตร์ตามกระบวนการรู้เรื่องคณิตศาสตร์ เรื่องปริภูมิและรูปทรงสามมิติ เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวมและแบบแยกส่วน พบว่า กระบวนการคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ มีค่าเฉลี่ยสูงสุด รองลงมาคือกระบวนการใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และกระบวนการให้เหตุผลทางคณิตศาสตร์ และการตีความการประยุกต์ใช้ และการประเมิน ผลลัพธ์ทางคณิตศาสตร์ ตามลำดับ

ตาราง 11 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบถามรู้เรื่องคณิตศาสตร์ เรื่องการเปลี่ยนแปลงและความสัมพันธ์ จำแนกตามกฎเกณฑ์การให้คะแนน

สถานการณ์	กฎเกณฑ์ การให้ คะแนน	จำนวน นักเรียน ทั้งหมด	คะแนน เต็ม	ข้อที่ 1 การคิด สถานการณ์ ของปัญหาใน เชิง คณิตศาสตร์		ข้อที่ 2 การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และ การให้เหตุผลทาง คณิตศาสตร์		ข้อที่ 3 การตีความ การประยุกต์ใช้ และการ ประเมินผลลัพธ์ ทางคณิตศาสตร์	
				$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
สถานการณ์ที่ 1 : การทำไร่นาสวน ผสมตามหลักการ เกษตรทฤษฎีใหม่	H	400	3	2.18	0.75	1.97	0.93	1.83	0.90
	A	400	9	7.55	1.59	6.99	2.01	6.50	2.19
สถานการณ์ที่ 2 : อัตราการเต้นของ หัวใจ	H	400	3	2.30	0.75	1.85	0.73	1.62	0.74
	A	400	9	7.77	1.51	6.51	1.95	6.08	1.95
สถานการณ์ที่ 3 : การโดยสาร เครื่องบิน	H	400	3	1.69	0.65	1.50	0.62	1.49	0.74
	A	400	9	6.48	1.73	5.48	1.98	5.46	2.14

จากตาราง 11 คะแนนที่ได้จากแบบสอบถามรู้เรื่องคณิตศาสตร์ตามกระบวนการรู้เรื่องคณิตศาสตร์ เรื่องการเปลี่ยนแปลงและความสัมพันธ์ เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวมและแบบแยกส่วน พบว่า กระบวนการคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ มีค่าเฉลี่ยสูงสุด รองลงมาคือกระบวนการใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และกระบวนการให้เหตุผลทางคณิตศาสตร์ และการตีความการประยุกต์ใช้ และการประเมิน ผลลัพธ์ทางคณิตศาสตร์ ตามลำดับ

ตาราง 12 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบถามรู้เรื่องคณิตศาสตร์ เรื่องปริมาณ
จำแนกตามกฎเกณฑ์การให้คะแนน

สถานการณ์	กฎเกณฑ์ การให้ คะแนน	จำนวน นักเรียน ทั้งหมด	คะแนน เต็ม	ข้อที่ 1 การคิด สถานการณ์ ของปัญหาใน เชิง คณิตศาสตร์		ข้อที่ 2 การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผล ทางคณิตศาสตร์		ข้อที่ 3 การตีความ การประยุกต์ใช้ และการ ประเมินผลลัพธ์ ทาง คณิตศาสตร์	
				$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
สถานการณ์ที่ 1 : น้ำแครอท	H	400	3	2.17	0.73	1.67	0.64	1.78	0.82
	A	400	9	7.56	1.35	6.26	1.72	6.36	2.09
สถานการณ์ที่ 2 : เครื่องเล่น MP3	H	400	3	1.73	0.63	1.46	0.54	1.58	0.73
	A	400	9	6.66	1.57	5.52	1.79	5.90	2.10
สถานการณ์ที่ 3 : จำนวนสัตว์ของ เดซี่ฮัทฟาร์ม	H	400	3	1.91	0.74	1.59	0.75	1.48	0.74
	A	400	9	6.96	1.74	5.89	2.03	5.60	2.07

จากตาราง 12 คะแนนที่ได้จากแบบสอบถามรู้เรื่องคณิตศาสตร์ตามกระบวนการรู้เรื่องคณิตศาสตร์ เรื่องปริมาณ เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวมและแบบแยกส่วน พบว่า กระบวนการคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ มีค่าเฉลี่ยสูงสุดรองลงมาก็คือกระบวนการใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และกระบวนการให้เหตุผลทางคณิตศาสตร์ และการตีความการประยุกต์ใช้ และการประเมิน ผลลัพธ์ทางคณิตศาสตร์ตามลำดับ

ตาราง 13 ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบสอบถามการรู้เรื่องคณิตศาสตร์ เรื่องความไม่แน่นอน จำแนกตามกฎเกณฑ์การให้คะแนน

สถานการณ์	กฎเกณฑ์ การให้ คะแนน	จำนวน นักเรียน ทั้งหมด	คะแนน เต็ม	ข้อที่ 1 การคิด สถานการณ์ ของปัญหาใน เชิง คณิตศาสตร์		ข้อที่ 2 การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และ การให้เหตุผลทาง คณิตศาสตร์		ข้อที่ 3 การตีความ การประยุกต์ใช้ และการ ประเมินผลลัพธ์ ทางคณิตศาสตร์	
				$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
สถานการณ์ที่ 1 : การคิดค่าไฟฟ้า	H	400	3	1.77	0.52	1.11	0.35	1.15	0.37
	A	400	9	6.64	1.30	4.29	1.46	4.36	1.53
สถานการณ์ที่ 2 : ราคาน้ำมัน	H	400	3	1.96	0.72	1.68	0.75	1.58	0.78
	A	400	9	6.94	1.64	5.81	2.22	5.30	2.37
สถานการณ์ที่ 3 : ราคายางพารา	H	400	3	1.93	0.73	1.42	0.66	1.41	0.70
	A	400	9	6.81	1.72	5.30	1.96	5.00	2.11

จากตาราง 13 คะแนนที่ได้จากแบบสอบถามการรู้เรื่องคณิตศาสตร์ตามกระบวนการรู้เรื่องคณิตศาสตร์ เรื่องความไม่แน่นอน เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม และแบบแยกส่วน พบว่า กระบวนการคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ มีค่าเฉลี่ยสูงสุด รองลงมาคือกระบวนการใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และกระบวนการให้เหตุผลทางคณิตศาสตร์ และการตีความการประยุกต์ใช้ และการประเมิน ผลลัพธ์ทางคณิตศาสตร์ตามลำดับ

ตอนที่ 2 คุณภาพแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน โดยศึกษาคุณภาพด้านความเชื่อมั่น (Reliability) จากค่าสัมประสิทธิ์การสรุปอ้างอิง (G - Coefficient) ตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) โดยออกแบบการวัดแบบ One - facet design โดยมีรูปแบบการวัดเป็น $p \times i$ เมื่อ p แทน ผู้สอบ และ i แทน จำนวนข้อสอบ ปรากฏผลดังนี้

2.1 ผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงที่ได้จากแบบสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบต่างกันและตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม รายละเอียดแสดงในตาราง 14

ตาราง 14 ผลการวิเคราะห์ความแปรปรวนจากแหล่งความแปรปรวนต่างๆ ที่ตรวจให้คะแนนตามกฎเกณฑ์การให้คะแนนแบบภาพรวม (G-Study of $p \times i$ Design)

EFFECT	df	SS	MS	ESTIMATED VARIANCE COMPONENT	% OF TOTAL VARIANCE
Person (p)	399	2654.933	6.653	0.176	30.82
Item (i)	35	1049.248	29.979	0.074	12.96
Residual (pi,e)	13965	44486.752	0.321	0.321	56.22
TOTAL	14399	8190.933		0.571	100

เมื่อ p แทน ผู้สอบ
 i แทน จำนวนข้อสอบ
 pi แทน ปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบ
 pi, e แทน ความคลาดเคลื่อนหรือส่วนที่เหลือที่ไม่สามารถระบุแหล่งได้

จากตาราง 14 พบว่า องค์ประกอบความแปรปรวนของบุคคล (σ_p^2) เป็นค่าความแปรปรวนของคะแนนจริง (Universe score) คิดเป็น 30.82% ของความแปรปรวนทั้งหมด องค์ประกอบของความแปรปรวนของข้อสอบ (σ_i^2) แสดงถึงความคลาดเคลื่อนของข้อสอบ คิดเป็น 12.96% ของความแปรปรวนทั้งหมด สำหรับองค์ประกอบของความแปรปรวนส่วนที่เหลือ ($\sigma_{pi,e}^2$) จึงเป็นความแปรปรวนของความคลาดเคลื่อนที่เหลือจาก 2 แหล่งดังกล่าว คิดเป็น 56.22 % ของความแปรปรวนทั้งหมด

2.2 การศึกษาเชิงตัดสินใจการสรุปอ้างอิง (D-STUDY)

การศึกษาเชิงตัดสินใจการสรุปอ้างอิงเพื่อลดความคลาดเคลื่อนและเพิ่มความแม่นยำของการสรุปอ้างอิงคุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์ โดยมีเงื่อนไขในการศึกษา คือ จำนวนข้อสอบ 12 ข้อ 24 ข้อ และ 36 ข้อ ที่มีกฎเกณฑ์ให้คะแนนแบบภาพรวม โดยผลการวิเคราะห์ดังแสดงในตาราง 15

ตาราง 15 ผลการศึกษา D ที่มีกฎเกณฑ์ให้คะแนนแบบภาพรวม (D-Study of p X i Design)
เมื่อ p แทน ผู้สอบ i แทน จำนวนข้อสอบ

Estimated Variance Components in D-Study					
EFFECT	G-Study		$n_i' = 12$	$n_i' = 24$	$n_i' = 36$
Person (p)	$\hat{\sigma}_p^2$	= 0.176	$\hat{\sigma}_p^2$	= 0.176	0.176
Item (i)	$\hat{\sigma}_i^2$	= 0.074	$\hat{\sigma}_i^2$	= 0.006	0.003
RESIDUAL (pir,e)	$\hat{\sigma}_{pi,e}^2$	= 0.321	$\hat{\sigma}_{pi,e}^2$	= 0.027	0.013
Error Variance	Relative error variance		$\hat{\sigma}_\delta^2$	= 0.027	0.013
	Absolute error variance		$\hat{\sigma}_\Delta^2$	= 0.033	0.016
G-Coefficient	Relative decision		ρ_δ^2	= 0.868	0.929
	Absolute decision		ρ_Δ^2	= 0.842	0.914

จากตาราง 15 ผลการศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิง (G - Coefficient) ในการศึกษา D (D-Study) สำหรับแบบสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบต่างกัน และตรวจให้คะแนนโดยใช้กฎเกณฑ์ให้คะแนนแบบภาพรวม พบว่า การเพิ่มจำนวนความยาวแบบสอบจาก 12 ข้อ เป็น 24 ข้อ ค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนนสำหรับการตัดสินใจเชิงสัมพัทธ์ เพิ่มขึ้นจาก 0.868 เป็น 0.929 (เพิ่มขึ้น 0.061) และค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนนสำหรับการตัดสินใจเชิงสัมบูรณ์ เพิ่มขึ้นจาก 0.842 เป็น 0.914 (เพิ่มขึ้น 0.072) และเมื่อเพิ่มจำนวนความยาวแบบสอบจาก 24 ข้อ เป็น 36 ข้อ ค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนนสำหรับการตัดสินใจเชิงสัมพัทธ์ เพิ่มขึ้นจาก 0.929 เป็น 0.952 (เพิ่มขึ้น 0.023) และค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนนสำหรับการตัดสินใจเชิงสัมบูรณ์ เพิ่มขึ้นจาก 0.914 เป็น 0.941 (เพิ่มขึ้น 0.027) จากการพิจารณาผลการวิเคราะห์ข้อมูลค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีกฎเกณฑ์ให้คะแนนแบบภาพรวมพบว่า ความยาวแบบสอบ 12 ข้อ 24 ข้อ และ 36 ข้อ มีค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมพัทธ์กับการตัดสินใจเชิงสัมบูรณ์ค่อนข้างสูง เมื่อจำนวนข้อเพิ่มมากขึ้นค่าความเชื่อมั่นยิ่งเพิ่มขึ้น

2.3 ผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงที่ได้จากแบบสอบถามการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบถามต่างกันและตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน รายละเอียดแสดงในตาราง 16

ตาราง 16 ผลการวิเคราะห์ความแปรปรวนจากแหล่งความแปรปรวนต่างๆ ที่ตรวจให้คะแนนตามกฎเกณฑ์การให้คะแนนแบบแยกส่วน (G–Study of p X i Design)

EFFECT	df	SS	MS	ESTIMATED VARIANCE COMPONENT	% OF TOTAL VARIANCE
Person (p)	399	20570.034	51.554	1.375	33.18
Item (i)	35	10047.004	287.057	0.713	17.21
Residual (pi,e)	13965	28709.441	2.0558	2.056	49.61
TOTAL	14399	59326.479		4.144	100

เมื่อ p แทน ผู้สอบ
i แทน จำนวนข้อสอบ
pi แทน ปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบ
pi, e แทน ความคลาดเคลื่อนหรือส่วนที่เหลือที่ไม่สามารถระบุแหล่งได้

จากตาราง 16 พบว่า องค์ประกอบความแปรปรวนของบุคคล (σ_p^2) เป็นค่าความแปรปรวนของคะแนนจริง (Universe score) คิดเป็น 33.18% ของความแปรปรวนทั้งหมด องค์ประกอบของความแปรปรวนของข้อสอบ (σ_i^2) แสดงถึงความคลาดเคลื่อนของข้อสอบ คิดเป็น 17.21% ของความแปรปรวนทั้งหมด สำหรับองค์ประกอบของความแปรปรวนส่วนที่เหลือ ($\sigma_{pi,e}^2$) จึงเป็นความแปรปรวนของความคลาดเคลื่อนที่เหลือจาก 2 แหล่งดังกล่าว คิดเป็น 49.61 % ของความแปรปรวนทั้งหมด

2.4 การศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง (D-STUDY)

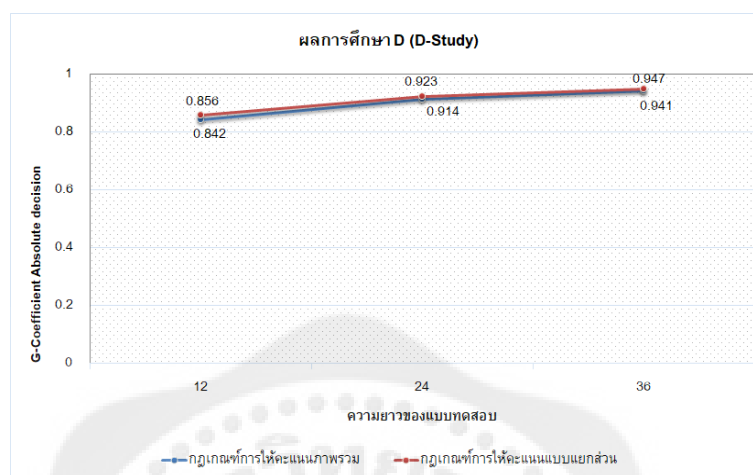
การศึกษากการตัดสินใจการสรุปอ้างอิงเพื่อลดความคลาดเคลื่อนและเพิ่มความแม่นยำของการสรุปอ้างอิงคุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์ โดยมีเงื่อนไขในการศึกษา คือ จำนวนข้อสอบ 12 ข้อ 24 ข้อ และ 36 ข้อ ที่มีกฎเกณฑ์ให้คะแนนแบบแยกส่วน โดยผลการวิเคราะห์ดังแสดงในตาราง 17

ตาราง 17 ผลการศึกษา D ที่มีกฎเกณฑ์ให้คะแนนแบบแยกส่วน (D-Study of p X i Design)
เมื่อ p แทน ผู้สอบ i แทน จำนวนข้อสอบ

Estimated Variance Components in D-Study					
EFFECT	G-Study		$n_i' = 12$	$n_i' = 24$	$n_i' = 36$
Person (p)	$\hat{\sigma}_p^2 = 1.375$	$\hat{\sigma}_p^2 =$	1.375	1.375	1.375
Item (i)	$\hat{\sigma}_i^2 = 0.713$	$\hat{\sigma}_i^2 =$	0.059	0.030	0.020
RESIDUAL (pir,e)	$\hat{\sigma}_{pi,e}^2 = 2.056$	$\hat{\sigma}_{pi,e}^2 =$	0.171	0.086	0.057
Error Variance	Relative error variance	$\hat{\sigma}_\delta^2 =$	0.171	0.086	0.057
	Absolute error variance	$\hat{\sigma}_\Delta^2 =$	0.231	0.115	0.077
G-Coefficient	Relative decision	$\rho_\delta^2 =$	0.889	0.941	0.960
	Absolute decision	$\rho_\Delta^2 =$	0.856	0.923	0.947

จากตาราง 17 ผลการศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิง (G - Coefficient) ในการศึกษา D (D-Study) สำหรับแบบสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาวแบบสอบต่างกัน และตรวจให้คะแนนโดยใช้กฎเกณฑ์ให้คะแนนแบบแยกส่วน พบว่า การเพิ่มจำนวนความยาวแบบสอบจาก 12 ข้อ เป็น 24 ข้อ ค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนนสำหรับการตัดสินใจเชิงสัมพัทธ์ เพิ่มขึ้นจาก 0.889 เป็น 0.941 (เพิ่มขึ้น 0.052) และค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนนสำหรับการตัดสินใจเชิงสัมบูรณ์ เพิ่มขึ้นจาก 0.856 เป็น 0.923 (เพิ่มขึ้น 0.067) และเมื่อเพิ่มจำนวนความยาวแบบสอบจาก 24 ข้อ เป็น 36 ข้อ ค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนนสำหรับการตัดสินใจเชิงสัมพัทธ์ เพิ่มขึ้นจาก 0.941 เป็น 0.960 (เพิ่มขึ้น 0.019) และค่าสัมประสิทธิ์ความน่าเชื่อถือของคะแนนสำหรับการตัดสินใจเชิงสัมบูรณ์ เพิ่มขึ้นจาก 0.923 เป็น 0.947 (เพิ่มขึ้น 0.024) จากการพิจารณาผลการวิเคราะห์ข้อมูลค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีกฎเกณฑ์ให้คะแนนแบบแยกส่วน พบว่า ความยาวแบบสอบ 12 ข้อ 24 ข้อ และ 36 ข้อ มีค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมพัทธ์กับการตัดสินใจเชิงสัมบูรณ์ค่อนข้างสูง เมื่อจำนวนข้อเพิ่มมากขึ้นค่าความเชื่อมั่นยิ่งเพิ่มขึ้น

2.5 การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบต่างกัน



ภาพประกอบ 10 เปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสอบการรู้เรื่องคณิตศาสตร์

ผู้วิจัยได้ทดสอบความแตกต่างค่าสัมประสิทธิ์การสรุปอ้างอิงโดยแปลงค่าเป็นคะแนนมาตรฐาน (Fisher's Z) แล้วทำการทดสอบซี (Z-test) ปรากฏผลแสดงตาราง 18

ตาราง 18 ผลการทดสอบความแตกต่างค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบการรู้เรื่องคณิตศาสตร์ที่ตรวจใช้กฎเกณฑ์การให้คะแนนรูปแบบเดียวกัน ที่มีความยาวแบบสอบต่างกัน

รูปแบบกฎเกณฑ์การให้คะแนน	ความยาวแบบสอบ (ข้อ)	Z_r	เปรียบเทียบความยาวแบบสอบ (ข้อ)	Z
กฎเกณฑ์การให้คะแนนแบบภาพรวม	12	1.128	12 กับ 24	4.55*
	24	1.551	24 กับ 36	2.75*
	36	1.747	36 กับ 12	7.31*
กฎเกณฑ์การให้คะแนนแบบแยกส่วน	12	1.278	12 กับ 24	4.66*
	24	1.609	24 กับ 36	2.72*
	36	1.802	36 กับ 12	7.38*

* $p < 0.05$

จากตาราง 18 พบว่า แบบสอบการรู้เรื่องคณิตศาสตร์ที่ตรวจโดยใช้กฎเกณฑ์การให้คะแนนรูปแบบเดียวกัน แต่มีระดับความยาวแบบสอบต่างกัน มีค่าสัมประสิทธิ์การสรุปอ้างอิงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

ในบทนี้ผู้วิจัยนำเสนอสาระสำคัญเกี่ยวกับความมุ่งหมายของการวิจัย วิธีดำเนินการวิจัย สรุปผลการวิจัย อภิปรายผลการวิจัย และข้อเสนอแนะ ซึ่งมีรายละเอียดโดยสรุปดังนี้

สังเขปความมุ่งหมายและวิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีความมุ่งหมาย 2 ประการ คือ 1) เพื่อสร้างและตรวจสอบคุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน และ 2) เพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบและกฎเกณฑ์การให้คะแนนต่างกัน กลุ่มตัวอย่างที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานภาคใต้ ปีการศึกษา 2560 จำนวน 400 คน ซึ่งได้มาจากการสุ่มแบบหลายขั้นตอน (Multi stage random sampling) ตัวแปรที่ศึกษาประกอบด้วยตัวแปรอิสระ 2 ตัวแปร คือ ความยาวแบบสอบจำนวน 12 ข้อ 24 ข้อ และ 36 ข้อ และกฎเกณฑ์การให้คะแนน 2 รูปแบบ คือ กฎเกณฑ์การให้คะแนนแบบภาพรวมกับกฎเกณฑ์การให้คะแนนแบบแยกส่วน ตัวแปรตาม คือ ค่าสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) ตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

การดำเนินการเก็บรวบรวมข้อมูล นักเรียนทุกคนจะทำแบบสอบการรู้เรื่องคณิตศาสตร์ จำนวน 36 ข้อ ผู้วิจัยใช้ระยะเวลาในการเก็บรวบรวมข้อมูลเป็นระยะเวลา 1 เดือน แล้วนำแบบสอบที่ได้มาตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวมและกฎเกณฑ์การให้คะแนนแบบแยกส่วน ซึ่งการวิจัยครั้งนี้ออกแบบการวัดเป็นแบบ One - facet design โดยมีรูปแบบการวัดเป็น $p \times i$ เมื่อ p แทน ผู้สอบ และ i แทน จำนวนข้อสอบ โดยผู้ตรวจดำเนินการตรวจข้อสอบทุกข้อของผู้สอบทุกคน ($p \times i$ design) หลังจากนั้นจึงนำคะแนนที่ได้มาวิเคราะห์คุณภาพของแบบสอบด้านความง่าย ยาก อำนาจจำแนก และความเชื่อมั่นโดยใช้สัมประสิทธิ์แอลฟาของครอนบาค รวมทั้งวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) ตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) เมื่อพิจารณาจากเงื่อนไขหรือพาเซทในกรณีความยาวแบบสอบแตกต่างกัน โดยใช้กฎเกณฑ์การให้คะแนนรูปแบบเดียวกัน และเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงด้วยการทดสอบซี (Z-test)

สรุปผลการวิจัย

1. ผลการตรวจสอบคุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์ด้านความยากง่าย อำนาจจำแนก และความเชื่อมั่นโดยใช้สัมประสิทธิ์แอลฟาของครอนบาค พบว่า ในภาพรวมแบบสอบการรู้เรื่องคณิตศาสตร์ที่ตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม มีค่าความยากง่ายระหว่าง 0.23 – 0.71 ค่าอำนาจจำแนกระหว่าง 0.21 – 0.48 และค่าความเชื่อมั่นเท่ากับ 0.952 เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน พบว่า แบบสอบมีค่าความยากง่ายระหว่าง 0.28 – 0.78 ค่าอำนาจจำแนกระหว่าง 0.20 – 0.50 และค่าความเชื่อมั่นเท่ากับ 0.960

เมื่อพิจารณาคุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์ จำแนกตามเนื้อหาสาระ พบว่า 1) แบบสอบเรื่องปริภูมิและรูปทรงสามมิติ เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม มีค่าความยากง่ายระหว่าง 0.39 – 0.68 ค่าอำนาจจำแนกระหว่าง 0.21 – 0.47 และค่าความเชื่อมั่นเท่ากับ 0.848 เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน มีค่าความยากง่ายระหว่าง 0.49 – 0.78 ค่าอำนาจจำแนกระหว่าง 0.20 – 0.50 และค่าความเชื่อมั่นเท่ากับ 0.875 2) แบบสอบเรื่องการเปลี่ยนแปลงและความสัมพันธ์ เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม มีค่าความยากง่ายระหว่าง 0.33 – 0.71 ค่าอำนาจจำแนกระหว่าง 0.21 – 0.47 และค่าความเชื่อมั่นเท่ากับ 0.859 เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน มีค่าความยากง่ายระหว่าง 0.38 – 0.78 ค่าอำนาจจำแนกระหว่าง 0.22 – 0.45 และค่าความเชื่อมั่นเท่ากับ 0.850 3) แบบสอบเรื่องปริมาณ พบว่า เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม มีค่าความยากง่ายระหว่าง 0.23 – 0.70 ค่าอำนาจจำแนกระหว่าง 0.22 – 0.48 และค่าความเชื่อมั่นเท่ากับ 0.842 เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน มีค่าความยากง่ายระหว่าง 0.28 – 0.73 ค่าอำนาจจำแนกระหว่าง 0.21 – 0.47 และค่าความเชื่อมั่นเท่ากับ 0.919 และ 4) แบบสอบเรื่องความไม่แน่นอน เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม มีค่าความยากง่ายระหว่าง 0.39 – 0.61 ค่าอำนาจจำแนกระหว่าง 0.22 – 0.41 และค่าความเชื่อมั่นเท่ากับ 0.834 เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน มีค่าความยากง่ายระหว่าง 0.33 – 0.63 ค่าอำนาจจำแนกระหว่าง 0.20 – 0.44 และค่าความเชื่อมั่นเท่ากับ 0.901

2. ผลการตรวจสอบคุณภาพของแบบสอบการรู้เรื่องคณิตศาสตร์จากค่าสัมประสิทธิ์การสรุปอ้างอิง (G- Coefficient) เมื่อตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนแบบภาพรวม พบว่า เมื่อเพิ่มความยาวแบบสอบจาก 12 ข้อ เป็น 24 ข้อ และ 36 ข้อ มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์เท่ากับ 0.868, 0.929 และ 0.952 ตามลำดับ และมีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์เท่ากับ 0.842, 0.914 และ 0.941 ตามลำดับ เมื่อตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนแบบแยกส่วน พบว่า เมื่อเพิ่มความยาวแบบสอบจาก 12 ข้อ เป็น 24 ข้อ และ 36 ข้อ มีค่าสัมประสิทธิ์การสรุป

อ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์เท่ากับ 0.889, 0.941 และ 0.960 ตามลำดับ และมีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์เท่ากับ 0.856, 0.923 และ 0.947 ตามลำดับ เมื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบการรู้เรื่องคณิตศาสตร์ที่ตรวจโดยใช้กฎเกณฑ์การให้คะแนนรูปแบบเดียวกัน เมื่อแบบสอบมีความยาวต่างกัน พบว่าค่าสัมประสิทธิ์การสรุปอ้างอิงมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

อภิปรายผล

1. แบบสอบการรู้เรื่องคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้นเป็นไปตามกรอบโครงสร้างการประเมินการรู้เรื่องคณิตศาสตร์ของโครงการประเมินผลนักเรียนร่วมกับนานาชาติ หรือ PISA ซึ่งครอบคลุม 4 สาระ ได้แก่ ปริภูมิและรูปร่างสามมิติ (Space and Shape) การเปลี่ยนแปลงและความสัมพันธ์ (Change and Relationships) ปริมาณ (Quantity) และความไม่แน่นอน (Uncertainty) และครอบคลุมกระบวนการคณิตศาสตร์ 3 กระบวนการ ได้แก่ 1) การคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ 2) การใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผลทางคณิตศาสตร์ และ 3) การตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์ จุดเด่นของแบบสอบนี้คือ ข้อสอบมีลักษณะเป็นสถานการณ์ที่ไม่ได้เน้นความรู้เนื้อหาวิชาคณิตศาสตร์ที่เรียนตามหลักสูตรในโรงเรียน แต่เน้นการนำคณิตศาสตร์ที่ได้เรียนมาใช้ในสถานการณ์ของชีวิตจริง โดยนักเรียนจะต้องขยายความรู้จากที่เรียนมาประยุกต์ใช้กับสถานการณ์จริงในบริบทต่าง ๆ ที่หลากหลาย (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2557: 1) ซึ่งแตกต่างจากแบบสอบคณิตศาสตร์ที่ใช้กันอย่างแพร่หลายในประเทศไทยที่ส่วนใหญ่เน้นถามเนื้อหาสาระโดยตรงตามหลักสูตรเท่านั้น นอกจากนี้แบบสอบการรู้เรื่องคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้นมีลักษณะเป็นข้อสอบอัตนัยทั้งหมด ซึ่งเป็นลักษณะข้อสอบที่ช่วยลดปัญหาเกี่ยวกับการเดาคำตอบ สามารถวัดความรู้ความสามารถระดับสูงได้ดีกว่าข้อสอบประเภทอื่น และส่งเสริมผู้ตอบแสดงความคิดเห็นอย่างเต็มที่ (ศิริชัย กาญจนวาสี. 2556: 219) อย่างไรก็ตาม ข้อสอบอัตนัยมีข้อจำกัดบางประการ เช่น เกณฑ์การให้คะแนนไม่แน่นอน การตรวจให้คะแนนยาก และเสียเวลามาก (บุญเชิด ภิญโญนนตพงษ์. 2526: 121) ผู้วิจัยจึงสร้างกฎเกณฑ์การให้คะแนนเพื่อใช้เป็นแนวทางในการตรวจข้อสอบ โดยมีการกำหนดประเด็นและคำอธิบายคุณภาพของคำตอบในแต่ละระดับ เพื่อสะท้อนความสามารถของผู้สอบให้ได้มากที่สุด โดยคำอธิบายคุณภาพจะต้องมีความชัดเจนในการใช้ภาษาที่กระชับรัดกุม เข้าใจง่าย และเห็นความแตกต่างระหว่างระดับอย่างชัดเจน (สุรศักดิ์ อมรรัตนศักดิ์. 2554: 256) และเนื่องจากกฎเกณฑ์การให้คะแนนนั้นมีหลายรูปแบบ ในการวิจัยครั้งนี้ผู้วิจัยจึงเลือกใช้ทั้งกฎเกณฑ์การให้คะแนนแบบภาพรวมและกฎเกณฑ์การให้คะแนนแบบแยกส่วน เพื่อเป็นแนวทางสำหรับใช้ประกอบการตัดสินใจในการให้คะแนนได้อย่างเที่ยงตรง คงเส้นคงวา และยุติธรรม

2. เมื่อพิจารณาคุณภาพของแบบสอบถามการรู้เรื่องคณิตศาสตร์ โดยพิจารณาจากค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด พบว่า เมื่อเพิ่มความยาวแบบสอบถามจาก 12 ข้อ เป็น 24 ข้อ และ 36 ข้อ ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (ρ^2_{Rel}) และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (ρ^2_{Abs}) เพิ่มขึ้น และมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เมื่อตรวจโดยใช้กฎเกณฑ์การให้คะแนนรูปแบบเดียวกัน สอดคล้องกับวัชรียา ศรีวีสรัตน์ (2558) พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัดและค่าความแปรปรวนของความคลาดเคลื่อนของแบบทดสอบเรียงความวิชาภาษาไทย เมื่อจำนวนผู้ตรวจคงที่ ถ้าเพิ่มจำนวนข้อสอบ ค่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัดเพิ่มขึ้น และค่าความแปรปรวนของความคลาดเคลื่อนลดลงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 อย่างไรก็ตาม เป็นที่น่าสังเกตว่าเมื่อเพิ่มความยาวแบบสอบถามจาก 24 ข้อ เป็น 36 ข้อ ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (ρ^2_{Rel}) และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (ρ^2_{Abs}) มีค่าเพิ่มขึ้นเพียงเล็กน้อย ซึ่งการพิจารณาเลือกความยาวแบบสอบถามที่มีความเหมาะสมนั้น บราวน์ (Brown 1996: 246) ได้กล่าวว่า ในการพิจารณาความน่าเชื่อถือของผลการวัด จำนวนข้อเป็นเพียงส่วนหนึ่งของข้อมูลหลายๆ ส่วนของการทดสอบเพื่อประกอบการตัดสินใจในการนำไปใช้และพัฒนาแบบสอบ การเพิ่มจำนวนข้อสอบนั้นย่อมส่งผลให้ความเชื่อมั่นสูงขึ้น จำนวนข้อที่เพิ่มขึ้นจะต้องมีคุณภาพเหมือนเดิม ภายใต้สถานการณ์ของการใช้แบบสอบหรือเงื่อนไขของการทดสอบที่ความยาวของแบบสอบมีผลต่อความน่าเชื่อถือของผลการวัด ทั้งนี้เพื่อเป็นการประหยัดเวลาในการทดสอบและการตรวจแบบสอบ การพิจารณาความยาวแบบสอบขึ้นอยู่กับวัตถุประสงค์ที่ต้องการศึกษาและครอบคลุมเนื้อหาที่ต้องการวัด นอกจากนี้เมื่อตรวจโดยใช้รูปแบบกฎเกณฑ์การให้คะแนนที่ต่างกัน พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (ρ^2_{Rel}) และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (ρ^2_{Abs}) มีความแตกต่างกันเพียงเล็กน้อย สอดคล้องกับแคมป์เบลล์ และฟิสต์ (ศิริชัย กาญจนวาสี. 2544: 97; อ้างอิงจาก Campbell; & Fiske. 1959) ที่กล่าวไว้ว่า คะแนนจากการวัดลักษณะเดียวกัน ถึงแม้จะใช้การวัดต่างกันย่อมมีความสัมพันธ์กันสูง การตัดสินใจเลือกกฎเกณฑ์การให้คะแนนสามารถเลือกใช้กฎเกณฑ์การให้คะแนนได้ทั้ง 2 รูปแบบ หรือถ้าต้องการเลือกรูปแบบกฎเกณฑ์การให้คะแนนรูปแบบใดรูปแบบหนึ่ง การพิจารณาเลือกประเภทกฎเกณฑ์การให้คะแนนจะต้องคำนึงถึงสอดคล้องกับจุดมุ่งหมายและลักษณะของผลผลิต หรือผลการปฏิบัติงานที่ต้องการวัดประเมิน และนำไปใช้ได้จริงในทางปฏิบัติ (องอาจ นัยวัฒน์. 2558)

3. ผลการศึกษาระดับการรู้เรื่องคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำแนกตามกระบวนการรู้เรื่องคณิตศาสตร์ พบว่า กระบวนการรู้เรื่องคณิตศาสตร์ที่นักเรียนมีค่าเฉลี่ยสูงที่สุด คือ กระบวนการคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ รองลงมา คือ กระบวนการใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผลทางคณิตศาสตร์ และกระบวนการตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์ ตามลำดับ เนื่องจากกระบวนการคิดสถานการณ์ของปัญหาในเชิงคณิตศาสตร์ เป็นขั้นตอนในการวิเคราะห์โจทย์ที่กำหนดให้ ต้องบอกให้ได้ว่าจุดไหนคือจุดสำคัญ จุดไหนที่สามารถมองข้ามไปได้ ต้องรู้จักคำ ภาพ หรือจุดเด่นอื่นๆ ของสถานการณ์ และทำให้อยู่ในรูปของประโยคสัญลักษณ์ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2560: 38) เป็นกระบวนการที่นักเรียนต้องมีความสามารถในการวิเคราะห์สถานการณ์นำไปสู่การหาคำตอบ สอดคล้องกับอัมพร ม้าคนอง (2553: 173) ที่กล่าวว่า องค์ประกอบที่สำคัญขององค์ประกอบหนึ่งของการสามารถในการแก้โจทย์ปัญหาคณิตศาสตร์ คือ การสร้างโจทย์หรือประเด็นปัญหา ซึ่งเป็นความสามารถในการเชื่อมโยงข้อมูลที่มีอยู่เพื่อหาความสัมพันธ์ ซึ่งจะนำไปสู่การสร้างโจทย์ สัญลักษณ์ เพื่อแก้ปัญหานั้นๆ ขั้นตอนนี้จึงไม่มีความซับซ้อนมากนัก แตกต่างจากกระบวนการใช้แนวคิด ข้อเท็จจริง วิธีดำเนินการ และการให้เหตุผลทางคณิตศาสตร์ที่มีความซับซ้อนเพิ่มขึ้น นักเรียนจะต้องแสดงออกมาให้อยู่ในรูปของคณิตศาสตร์ เช่น การคำนวณตัวเลข นักเรียนต้องรู้จักเครื่องมือคณิตศาสตร์ที่นำมาใช้ให้มีความเหมาะสมกับสถานการณ์และนำไปสู่การแก้ปัญหานั้นๆ นักเรียนต้องดำเนินการคำนวณ รู้จักการแทนค่าในสูตร มีการแก้สมการหรือใช้ความรู้ตามหลักเกณฑ์ การสร้างกราฟ แผนภูมิต่างๆ มีการสกัดข้อมูลแล้วนำเสนอในรูปของคณิตศาสตร์ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2560: 39) เป็นกระบวนการคิดคำนวณแก้โจทย์ปัญหา กระบวนการนี้จะต้องใช้ระยะเวลาค่อนข้างนานในการคิดคำนวณเพื่อหาคำตอบที่ถูกต้อง นักเรียนต้องมีความสามารถในการแก้โจทย์ปัญหา สอดคล้องกับฐิติยา วงศ์วิทยานุกูล (2555: 171) พบว่า ความสามารถในการคิดคำนวณมีอิทธิพลทางตรงต่อความสามารถในการแก้โจทย์ปัญหาคณิตศาสตร์ และมีความสัมพันธ์ทางบวกกับความสามารถในการแก้โจทย์ปัญหาคณิตศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 สำหรับกระบวนการตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์ เป็นกระบวนการที่ค่อนข้างมีความซับซ้อนมากกว่ากระบวนการอื่นๆ เนื่องจากนักเรียนต้องตีความผลลัพธ์ทางคณิตศาสตร์กลับไปทบทวนโลกชีวิตจริง ประเมินความเป็นเหตุเป็นผลของวิธีแก้ปัญหาวางคณิตศาสตร์ในบริบทของปัญหาโลกชีวิตจริง เข้าใจว่าสถานการณ์ในชีวิตจริงส่งผลกระทบต่อผลลัพธ์และการคิดคำนวณตามกระบวนการทางคณิตศาสตร์อย่างไร รวมทั้งสามารถอธิบายได้ว่า เพราะเหตุใดผลลัพธ์หรือข้อสรุปทางคณิตศาสตร์จึงเหมาะสมหรือไม่เหมาะสมกับบริบทของปัญหา (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2560: 40) จึงเป็นกระบวนการที่ค่อนข้างยากสำหรับนักเรียนที่ไม่มีความคุ้นเคยกับการทำแบบสอบลักษณะนี้

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. จากผลการวิจัยพบว่า เมื่อแบบสอบการรู้เรื่องคณิตศาสตร์มีจำนวนข้อมากขึ้น ค่าความเชื่อมั่นของแบบสอบก็จะเพิ่มขึ้น แต่เมื่อพิจารณาค่าสัมประสิทธิ์การสรุปอ้างอิง พบว่า เมื่อเพิ่มความยาวแบบสอบจาก 12 ข้อ เป็น 24 ข้อ และ 36 ข้อ ค่าสัมประสิทธิ์การสรุปอ้างอิง มีค่าเพิ่มขึ้น แต่มีความแตกต่างกันเพียงเล็กน้อย ดังนั้น เพื่อความสะดวกในการนำไปใช้ ครูผู้สอนอาจเลือกใช้แบบสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาว 12 ข้อ ก็น่าจะเพียงพอ เนื่องจากมีค่าสัมประสิทธิ์การสรุปอ้างอิงอยู่ในระดับสูง และเพื่อลดความเหนื่อยล้าจากการทำข้อสอบของนักเรียน ประหยัดเวลาและค่าใช้จ่ายในการทดสอบ รวมทั้งลดภาระงานในการตรวจให้คะแนน

2. แบบสอบการรู้เรื่องคณิตศาสตร์เป็นข้อสอบสถานการณ์ที่ไม่ได้เน้นความรู้เนื้อหาวิชาคณิตศาสตร์ที่เรียนตามหลักสูตรในโรงเรียน แต่เน้นการนำคณิตศาสตร์ที่ได้เรียนมาใช้ในสถานการณ์ของชีวิตจริง โดยนักเรียนจะต้องขยายความรู้จากที่เรียนมาประยุกต์ใช้กับสถานการณ์จริงในบริบทต่าง ๆ ที่หลากหลาย ซึ่งนักเรียนไทยอาจจะยังไม่คุ้นเคยกับแบบสอบในลักษณะนี้ ดังนั้น ครูผู้สอนควรบูรณาการวัดประเมินโดยใช้แบบสอบลักษณะนี้ร่วมกับการจัดการเรียนรู้ในชั้นเรียน รวมทั้งปรับกระบวนการจัดการเรียนรู้ให้นักเรียนได้ฝึกคิด สร้างสถานการณ์ในชั้นเรียนให้นักเรียนได้ลองเรียนรู้ฝึกฝน ตั้งคำถามและให้รู้จักนำไปใช้ เพื่อเตรียมความพร้อมให้นักเรียนเผชิญหน้ากับปัญหาทางคณิตศาสตร์ที่เกิดขึ้นในการดำเนินชีวิตประจำวันในบริบทที่แตกต่างกัน รวมทั้งเป็นการเตรียมความพร้อมให้นักเรียนคุ้นชินกับข้อสอบลักษณะนี้ที่เริ่มนำมาใช้ในการทดสอบระดับชาติ (O-NET) มากขึ้น

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. การวิจัยครั้งนี้ผู้วิจัยตรวจให้คะแนนเพียงคนเดียว ซึ่งต้องใช้ระยะเวลาในการตรวจนานมาก ดังนั้น ในการศึกษาร้อยต่อไปควรมีการใช้จำนวนผู้ตรวจเป็นเงื่อนไขในการศึกษา (Facet) เพื่อศึกษาว่าควรเพิ่มจำนวนผู้ตรวจเป็นเท่าใดจึงจะทำให้ความเชื่อมั่นอยู่ในระดับที่เหมาะสม และเชื่อถือได้ เพื่อช่วยลดระยะเวลาในการตรวจแบบสอบ

2. การวิเคราะห์ข้อสอบรายข้อในการวิจัยครั้งนี้ใช้ทฤษฎีการทดสอบแบบดั้งเดิมเท่านั้น ดังนั้นการศึกษาร้อยต่อไปควรมีการศึกษาคุณภาพของแบบทดสอบโดยประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า เพื่อให้ได้ข้อมูลสารสนเทศที่สะท้อนความแม่นยำของการประมาณค่าความสามารถ (θ) ของผู้สอบ และช่วยเพิ่มประสิทธิภาพในการวัดและประเมินผลให้มีความถูกต้องแม่นยำมากยิ่งขึ้น



บรรณานุกรม

- โครงการประชุมอบรมเชิงปฏิบัติการเพื่อพัฒนาวิทยากรหลักระดับเขตพื้นที่การศึกษา เรื่องการประเมินในชั้นเรียน (Classroom Assessment). การสร้างเกณฑ์การประเมิน Rubric Sampler. สำนักทดสอบทางการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ มีนาคม - เมษายน 2550.
- โครงการ PISA ประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.(2554). ผลการประเมิน PISA 2009 การอ่าน คณิตศาสตร์ และวิทยาศาสตร์. กรุงเทพฯ: อรุณการพิมพ์.
- โครงการ PISA ประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.(2557). ผลการประเมิน PISA 2012 คณิตศาสตร์ การอ่าน และวิทยาศาสตร์ นักเรียนรู้อะไร และทำอะไรได้บ้าง. กรุงเทพฯ: อรุณการพิมพ์.
- โครงการ PISA ประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.(2560). สรุปผลการวิจัย PISA 2015. กรุงเทพฯ: อรุณการพิมพ์.
- จารุพันธ์ ขวัญแน่น. (2015). การพัฒนารูปแบบการประเมินการรู้คิดศาสตร์. *Journal of Community Development Research (Humanities and Social Sciences)*. 2015: 8.
- ฉัตรศิริ ปิยะพิมลสิทธิ์. (2544). เกณฑ์การให้คะแนน (Scoring Rubrics). สืบค้นจาก <http://www.watpon.com>.
- ฉวีวรรณ แก้วไทรฮะ; และสุพจน์ ไชยสังข์. (2557). การวิเคราะห์การสอบพิชชาและโอเน็ตของสถาบันทดสอบการศึกษาแห่งชาติ เพื่อปฏิรูปการเรียนการสอนวิชาคณิตศาสตร์. สถาบันทดสอบการศึกษาแห่งชาติ (องค์การมหาชน).
- ชลีพร ดัดงาม. (2552). การพัฒนาเกณฑ์การตรวจให้คะแนนการเขียนเรียงความของนักเรียน ช่วงชั้นที่ 2. ปริญญาโท กศ.ม (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- จิตติยา วงศ์วิทยานุกูล. (2555). “ปัจจัยเชิงสาเหตุที่ส่งผลต่อความสามารถในการแก้โจทย์ปัญหา คณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 6”. *การวัดผลการศึกษา* 17(1): 163 - 171.
- น้ำผึ้ง อินทะเนตร. (2554). การศึกษาคูณลักษณะของคะแนนแบบทดสอบปลายเปิดวิชา คณิตศาสตร์เมื่อจำนวนผู้ตรวจและรูปแบบการตรวจให้คะแนนต่างกัน โดยใช้โมเดลการสรุปอ้างอิงและโมเดลหลายองค์ประกอบของราล์ฟ. ปริญญาโท กศ.ด.(การทดสอบและวัดผล การศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- บุญเชิด ภิญโญอนันตพงษ์. (2536). การวัดและประเมินผลการศึกษาทฤษฎีและการประยุกต์. กรุงเทพฯ: อักษรเจริญทัศน์.

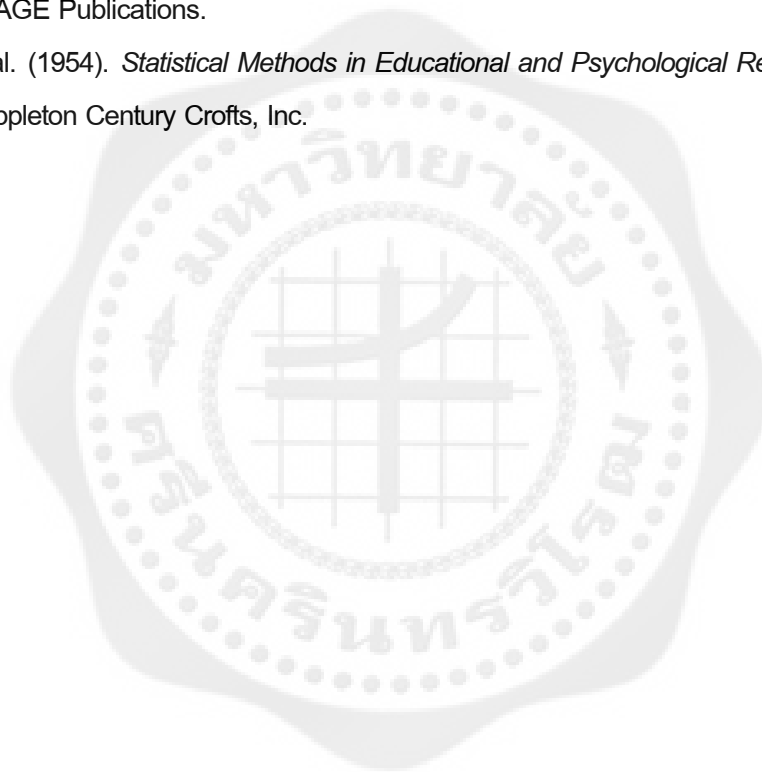
- บุญเชิด ภิญโญนนท์พงษ์. (2544). "คุณภาพของเครื่องมือวัด" หน่วยที่ 3. ประมวลสาระชุดวิชาการ พัฒนาเครื่องมือสำหรับการประเมินการศึกษา. นนทบุรี: มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- บุษวรรษ์ แสนปลื้ม. (2556). การใช้วิธีการตรวจคุณลักษณะ และสัดส่วนจำนวนผู้ตรวจให้คะแนนที่มีผลต่อความเที่ยงตรงของการวัดความสามารถในการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 3. ปรินญานิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ปวีณา ปิอาทิตย์. (2545). การศึกษาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนที่ต่างกัน. ปรินญานิพนธ์ กศ.ม. (การวัดและประเมินผลการศึกษา) บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ภัทรา สุวรรณบัตร. (2552). วิธีการเรียนรู้สู่ความสำเร็จในการเรียนคณิตศาสตร์, วารสารวิชาการ, 12 (2): 66.
- มหาวิทยาลัยสุโขทัยธรรมาธิราช สาขาวิชาศึกษาศาสตร์. (2545). การพัฒนาเครื่องมือสำหรับการประเมินการศึกษา (Development of Education Evaluation Instruments). สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- ยศอารีย์ รวยชนพานิช. (2534). การเปรียบเทียบความเชื่อมั่นของแบบทดสอบปรนัยเลือกตอบ ที่มีช่วงความยากและความยาวของแบบทดสอบต่างกัน ในวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- รัตนา ตั้งศิริชัยพงษ์. (2553). รูปแบบการสอนวิชาคณิตศาสตร์เพิ่มเติมที่ประยุกต์ใช้กลยุทธ์ การพัฒนาตนเองด้วยสัญญาการเรียนในการพัฒนาสัมฤทธิ์ผลทางการเรียนคณิตศาสตร์ ของนักเรียนด้อยสัมฤทธิ์ชั้นมัธยมศึกษาปีที่ 5 โรงเรียนท่าบ่อจังหวัดหนองคาย. วารสารวิชาการ, 14(4): 6.
- ราตรี นันทสุคนธ์. (2553). หลักการวัดและประเมินผลการศึกษา. กรุงเทพฯ: บริษัทจุดทอง จำกัด.
- ล้วน สายยศ; และอังคณา สายยศ. (2539). เทคนิคการวัดผลการเรียนรู้. กรุงเทพฯ: ชมรมเด็ก.
- วัชรียา ศรีวลีรัตน์. (2558). การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผล การวัดของข้อสอบเรียงความวิชาภาษาไทย ชั้นประถมศึกษาปีที่ 6 ที่มีจำนวนข้อสอบ และจำนวนผู้ตรวจต่างกัน. ปรินญานิพนธ์ ค.บ. (วิจัยและประเมินผลการศึกษา). กาญจนบุรี: บัณฑิตวิทยาลัย มหาวิทยาลัยราชภัฏกาญจนบุรี. ถ่ายเอกสาร.
- วิมลรัตน์ ศรีสุข. (2553). การพัฒนากระบวนการเรียนการสอนโดยการบูรณาการรูปแบบการสร้าง มโนทัศน์กับรูปแบบการแปลงเพื่อเสริมสร้างความรู้ทางคณิตศาสตร์และความสามารถ ทางความคิดอุปนัยของนักเรียนมัธยมศึกษา, วารสารวิชาการ, 13(2): 74 - 75.
- ศิริชัย กาญจนวาสี. (2544). ทฤษฎีการทดสอบแบบดั้งเดิม. พิมพ์ครั้งที่ 4 กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

- ศิริชัย กาญจนวาสี; ทวีวัฒน์ ปิตยานนท์; และดิเรก ศรีสุขโข. (2551). การเลือกใช้สถิติที่เหมาะสมสำหรับการวิจัย. พิมพ์ครั้งที่ 5. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2555). ทฤษฎีการทดสอบแนวใหม่ (MODERN TEST THEORIES). พิมพ์ครั้งที่ 4 (ฉบับปรับปรุงเพิ่มเติม) กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สาวิตรี จ้อยทอง. (2554). การตรวจสอบคุณภาพรูปแบบของชุดเครื่องมือการประเมินผลการเรียนรู้ตามสภาพจริงกลุ่มสาระการเรียนรู้คณิตศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 โดยการใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory). ปรินซ์นิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สมศักดิ์ ภู่วิภาดาธรรม. (2544). การยึดผู้เรียนเป็นศูนย์กลางและการประเมินตามสภาพจริง. เชียงใหม่: The Knowledge Center.
- สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน. (2555). มัธยมศึกษายุคใหม่สู่มาตรฐานสากล 2561. กรุงเทพฯ: ผู้พิมพ์.
- สำนักทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ. (2546). แนวทางการประเมินผลด้วยทางเลือกใหม่ตามหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 กลุ่มสาระการเรียนรู้คณิตศาสตร์. กรุงเทพฯ: สำนักทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ ผู้พิมพ์.
- สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2551). ตัวอย่างการจัดรายวิชาและคำอธิบายรายวิชาตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551.
- สุชาติ ปัทมวิภาต. (2557). การประเมินการรู้เรื่องคณิตศาสตร์ของ PISA 2015. นิตยสาร สสวท. 42(188).
- สุนีย์ คล้ายนิล. (2547). ความรู้และทักษะของเยาวชนไทยสำหรับโลกวันพรุ่งนี้. กรุงเทพฯ: เซเว่นพรินติ้ง กรุ๊ปจำกัด.
- สุนีย์ คล้ายนิล; และคณะ. (2550). การวัดผลประเมินผลเพื่อการเรียนรู้และตัวอย่างข้อสอบจากโครงการประเมินนักเรียนนานาชาติ PISA กรุงเทพฯ: เซเว่น พรินติ้ง กรุ๊ปจำกัด.
- สุรศักดิ์ อมรัตน์ศักดิ์. (2554). การประเมินผลการศึกษา (Educational Evaluation). พิมพ์ครั้งที่ 2 กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยรามคำแหง.
- สุวิมล กฤษณฤาสน์. (2558). ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด. สารานุกรมศึกษาศาสตร์ (Encyclopedia of Education) (50):6-16.
- เสาวนีย์ เกียรติ. (2540). คู่มือการอบรมการใช้แฟ้มสะสมงานนักเรียน. กรุงเทพฯ: เนติกุลการพิมพ์.
- ศูนย์เทคโนโลยีสารสนเทศและการสื่อสาร สำนักงานปลัดกระทรวงศึกษาธิการ. สืบค้นจาก <http://www.mis.moe.go.th>.

- อกนิษฐ์ บุญสะอาด. (2557). การพัฒนาแบบประเมินความคิดรวบยอดวิชาฟิสิกส์โดยใช้แผนผังมโนทัศน์: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด. ปรินุญานินพนธ์ กศ.ม. (การวิจัยและพัฒนาศึกษากายภาพมนุษย์). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- องอาจ นัยพัฒน์. (2552). การออกแบบการวิจัยวิธีการเชิงปริมาณ เชิงคุณภาพและผสมผสานวิธีการ. กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- องอาจ นัยพัฒน์. (2558). เอกสารประกอบการบรรยายมาตรฐานผลการเรียนรู้ตามกรอบ TQF: HE และกฎเกณฑ์ (แนวทาง) การให้คะแนน: แนวคิดพื้นฐานและวิธีการสร้างเพื่อวัดประเมินผลการเรียนรู้ตามกรอบ TQF: HEd. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- อรอุมา อุทัยวรรณ. (2548). การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบอัตนัย การแก้โจทย์ปัญหาคณิตศาสตร์ ที่มีความยาวของแบบทดสอบ วิธีการตรวจและจำนวนผู้ตรวจแตกต่างกัน. ปรินุญานินพนธ์ กศ.ม. (การวัดผลการศึกษา). มหาสารคาม: บัณฑิตวิทยาลัย มหาวิทยาลัยมหาสารคาม. ถ่ายเอกสาร.
- อรรถนีย์ ชูช่วยสุวรรณ. (2552). การพัฒนาแบบทดสอบวัดความสามารถในการเชื่อมโยงทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5. ปรินุญานินพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- อัมพร ม้าคนอง. (2553). ทักษะและกระบวนการทางคณิตศาสตร์: การพัฒนาเพื่อพัฒนาการ. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.
- ASTV ผู้จัดการออนไลน์. สอบ O-Net ปี 59 ใช้แนวข้อสอบ PISA Liked เน้นคิดวิเคราะห์. 8 กันยายน 2557 สืบค้นจาก <http://www.manager.co.th>.
- Anders Jonsson; & Gunilla Svingby. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review* 2 (2007). p.130 - 144.
- Ariyadi Wijaya; et al. (2014). Difficulties in solving context-based PISA mathematics tasks: An analysis of students' errors. *The Mathematics Enthusiast TME*. 11(3): 555.

- Brown, J.D.; & Ross, J.A. (1996). *Decision dependability of item types, sections, tests, and the overall TOEFL test battery*. In Milanovic, M. & Sville, N. (Eds), *Performance testing, cognition and assessment* (p. 231-265). Cambridge: Cambridge University Press.
- Camilli, G.; & Shepard, L.A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage Publications. (Vol. 4).
- Carl W. Swartz; et al. (1999). Using Generalizability Theory to estimate the reliability of writing scores derived from holistic and analytic scoring methods. *Educational and Psychological Measurement*, 59(3): 492 - 506.
- Duncan Mhakure; & Mamolahluwa Amelia Mokoena. (2011). A Comparative Study of the FET Phase Mathematical Literacy and Mathematics Curriculum. *US-China Education Review* (2011). p.309 - 323 *US-China Education Review*, ISSN 1548 - 6613.
- Feldt; & Leonard S. (1980). *A Test of Hypothesis that's Cronbach's alpha Reliability Coefficient the Same for Two Test Administered to The Same Sample*. *Psychometrika*. 45(1): 101 - 102.
- Hamsa Venkat; et al. (2009). *Critiquing the Mathematical Literacy Assessment Taxonomy: Where is the Reasoning and the Problem Solving?* *Pythagoras*, (70): 43 - 56.
- Hussain Alkharusi. (2012). Generalizability Theory: An Analysis of Variance Approach to Measurement Problems in Educational Assessment. *Journal of Studies in Education*. 2(1).
- Huseyin Husnu; & Giray Berberoglu. (2009). Judgmental and Statistical DIF Analyses of the PISA 2003 Mathematics Literacy Items. *International Journal of Testing*. (9): 108 - 121.
- John O. Anderson; et al. (2007). Using Large-scale Assessment Datasets for Research in Science and Mathematics Education: Programme for International Student Assessment (PISA). *International Journal of Science and Mathematics Education December 2007*, 5(4): 591 - 614.
- Kathleen P. Greenberg. (2015). *Rubric Use in Formative Assessment: A Detailed Behavioral Rubric Helps Students Improve Their Scientific Writing Skills*. *Society for The Teaching of Psychology* 2015, 42(3): 211 - 217.
- Oksana Naumenko. (2015). Improving Performance Assessment Score Validation Practices: An Instructional Module on Generalizability Theory. *Working Papers on Language and Diversity in Education August 2015*, 1(1).
- Pieter Spooren. (2014). Assessing the validity and reliability of a quick scan for student's evaluation of teaching. Results from G Theory. *Studies in Educational Evaluation* (43): 88 - 94.

- Ralph Bloch; & Geoff Norman. (2015). *G String IV Version 6.1.1 User Manual Revised*.
- Richard J. Shavelson; & Noreen M. Webb. (2005). *Generalizability Theory*. In Encyclopedia of Statistics in Behavioral Science.
- Robert L. Brennan. (2000). Performance Assessments Form the Perspective of Generalizability Theory. *Applied Psychological Measurement*, 24(4): 339 - 353.
- Thomas Eckes. (2008). Rater types in writing performance assessments: A classification approach to rater variability. *Language Testing* 25 (2) 155 - 185. SAGE Publications.
- Shavelson, R.j; & Webb, N.M. (1991). *Generalizability theory: A primer volume 1*. California: SAGE Publications.
- Wert; et al. (1954). *Statistical Methods in Educational and Psychological Research*. New York: Appleton Century Crofts, Inc.





ภาคผนวก



รายชื่อผู้เชี่ยวชาญในการตรวจสอบคุณภาพเครื่องมือ

ผู้ช่วยศาสตราจารย์ ดร.เรืองเดช ศิริกิจ

อาจารย์ประจำภาควิชาสำนักทดสอบ
การศึกษาและจิตวิทยา มหาวิทยาลัย
ศรีนครินทรวิโรฒ ประสานมิตร

อาจารย์ ดร.สุรินทร์ บุญพัฒนาภรณ์

อาจารย์โรงเรียนสาธิต มหาวิทยาลัย
ศรีนครินทรวิโรฒ ประสานมิตร (ฝ่ายมัธยม)

อาจารย์ ดร.นอร์ ไซยพรพัฒนา

อาจารย์โรงเรียนประชาบำรุง

อาจารย์ กฤษณา ตันรัตนพิทักษ์

อาจารย์โรงเรียนกุฎีวิทยาลัย

อาจารย์ ถนอมเกียรติ งานสกุล

อาจารย์โรงเรียนสตรีภูเก็ต





ภาคผนวก ข

ผลการตรวจสอบความสอดคล้องของผู้เชี่ยวชาญ

ผลการตรวจสอบความสอดคล้องของผู้เชี่ยวชาญจำนวน 5 คน

ตารางผลการตรวจสอบความสอดคล้อง เรื่องปฏิภูมิและรูปทรงสามมิติ

เรื่อง ปฏิภูมิและรูปทรง สามมิติ	ข้อ	ผู้เชี่ยวชาญคนที่					รวม	IOC	ผลการ ประเมิน
		1	2	3	4	5			
สถานการณ์ที่ 1 ค่าใช้จ่ายการตัดหญ้า	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 2 บรรจุหมกล่องลงในลัง กระดาษ	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	0	+1	+1	+1	+1	4	0.80	✓
สถานการณ์ที่ 3 พีระมิด	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	0	+1	4	0.80	✓
สถานการณ์ที่ 4 บรรจุน้ำดื่ม	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 5 การก่อกองน้ำด้วยอิฐ	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	0	+1	+1	0	+1	3	0.60	✓

ตารางผลการตรวจสอบความสอดคล้อง เรื่องการเปลี่ยนแปลงและความสัมพันธ์

เรื่อง การเปลี่ยนแปลง และความสัมพันธ์	ข้อ	ผู้เชี่ยวชาญคนที่					รวม	IOC	ผลการ ประเมิน
		1	2	3	4	5			
สถานการณ์ที่ 1 การทำไร่นาสวนผสม ตามหลักการเกษตร ทฤษฎีใหม่	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 2 อัตราการเต้นของหัวใจ	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 3 เวลา	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 4 อัตราแลกเปลี่ยนเงิน	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	0	+1	+1	+1	+1	4	0.80	✓
สถานการณ์ที่ 5 การโดยสารเครื่องบิน	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓

ตารางผลการตรวจสอบความสอดคล้อง เรื่องปริมาณ

เรื่อง ปริมาณ	ข้อ	ผู้เชี่ยวชาญคนที่					รวม	IOC	ผลการ ประเมิน
		1	2	3	4	5			
สถานการณ์ที่ 1 ดินสอ	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	0	+1	+1	+1	+1	4	0.80	✓
สถานการณ์ที่ 2 มูลค่าการส่งออก ผลิตภัณฑ์ยางของ ประเทศไทย	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 3 เครื่องเล่น MP3	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 4 การทำเค้ก	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 5 จำนวนสัตว์ของแดรี่ฮัท ฟาร์ม	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓

ตารางผลการตรวจสอบความสอดคล้อง เรื่องความไม่แน่นอน

เรื่อง ความไม่แน่นอน	ข้อ	ผู้เชี่ยวชาญคนที่					รวม	IOC	ผลการประเมิน
		1	2	3	4	5			
สถานการณ์ที่ 1 การพยากรณ์อากาศ	1	0	+1	+1	+1	0	3	0.60	✓
	2	+1	+1	+1	+1	0	4	0.80	✓
	3	+1	+1	+1	+1	0	4	0.80	✓
สถานการณ์ที่ 2 ทอยลูกเต๋า	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 3 แนวโน้มประชากรไทย	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	0	+1	+1	+1	+1	4	0.80	✓
	3	0	+1	+1	+1	+1	4	0.80	✓
สถานการณ์ที่ 4 ราคาน้ำมัน	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓
สถานการณ์ที่ 5 ราคายางพารา	1	+1	+1	+1	+1	+1	5	1.00	✓
	2	+1	+1	+1	+1	+1	5	1.00	✓
	3	+1	+1	+1	+1	+1	5	1.00	✓



ภาคผนวก ค

**ผลการหาคุณภาพค่าความยากง่ายและอำนาจจำแนก
จากการทดลองใช้ (try out)**

ผลการหาคุณภาพค่าความยากง่ายและอำนาจจำแนก

ตารางค่าความยากง่ายและอำนาจจำแนก เรื่องปริภูมิและรูปทรงสามมิติ

เรื่องปริภูมิและรูปทรงสามมิติ	ข้อที่	เกณฑ์การให้คะแนน	ค่าความยาก	ค่าอำนาจจำแนก	สรุปผล
สถานการณ์ที่ 1 ค่าใช้จ่ายการตัดหญ้า	1	H	0.53	0.47	ใช้ได้
		A	0.50	0.50	ใช้ได้
	2	H	0.47	0.42	ใช้ได้
		A	0.39	0.39	ใช้ได้
	3	H	0.68	0.21	ใช้ได้
		A	0.65	0.20	ใช้ได้
สถานการณ์ที่ 2 บรรจุนมกล่องลงในลังกระดาษ	1	H	0.50	0.39	ใช้ได้
		A	0.39	0.39	ใช้ได้
	2	H	0.66	0.21	ใช้ได้
		A	0.57	0.20	ใช้ได้
	3	H	0.50	0.30	ใช้ได้
		A	0.53	0.31	ใช้ได้
สถานการณ์ที่ 3 พีระมิด	1	H	0.64	0.25	ใช้ได้
		A	0.54	0.24	ใช้ได้
	2	H	0.44	0.22	ใช้ได้
		A	0.48	0.26	ใช้ได้
	3	H	0.55	0.32	ใช้ได้
		A	0.50	0.31	ใช้ได้
สถานการณ์ที่ 4 : บรรจุน้ำดื่ม	1	H	0.73	0.12	ใช้ไม่ได้
		A	0.67	0.07	ใช้ไม่ได้
	2	H	0.54	0.29	ใช้ได้
		A	0.53	0.31	ใช้ได้
	3	H	0.57	0.28	ใช้ได้
		A	0.72	0.28	ใช้ได้
สถานการณ์ที่ 5 ปูกระเบื้องพื้นห้องนอน	1	H	0.70	0.26	ใช้ได้
		A	0.78	0.22	ใช้ได้
	2	H	0.68	0.32	ใช้ได้
		A	0.67	0.33	ใช้ได้
	3	H	0.49	0.31	ใช้ได้
		A	0.56	0.33	ใช้ได้

ตารางค่าความยากง่ายและอำนาจจำแนก เรื่องการเปลี่ยนแปลงและความสัมพันธ์

เรื่องการเปลี่ยนแปลงและ ความสัมพันธ์	ข้อที่	เกณฑ์การให้ คะแนน	ค่าความยาก	ค่าอำนาจจำแนก	สรุปผล
สถานการณ์ที่ 1 การทำไร่นาสวนผสมตาม หลักการเกษตรทฤษฎีใหม่ ของพระบาทสมเด็จพระเจ้าอยู่หัวฯ	1	H	0.53	0.27	ใช้ได้
		A	0.75	0.25	ใช้ได้
	2	H	0.46	0.26	ใช้ได้
		A	0.48	0.25	ใช้ได้
	3	H	0.33	0.20	ใช้ได้
		A	0.38	0.22	ใช้ได้
สถานการณ์ที่ 2 อัตราการเดินของหัวใจ	1	H	0.57	0.30	ใช้ได้
		A	0.60	0.30	ใช้ได้
	2	H	0.55	0.25	ใช้ได้
		A	0.52	0.22	ใช้ได้
	3	H	0.53	0.21	ใช้ได้
		A	0.75	0.25	ใช้ได้
สถานการณ์ที่ 3 เวลา	1	H	0.58	0.10	ใช้ไม่ได้
		A	0.50	0.30	ใช้ได้
	2	H	0.57	0.13	ใช้ไม่ได้
		A	0.50	0.05	ใช้ไม่ได้
	3	H	0.28	0.13	ใช้ไม่ได้
		A	0.40	0.20	ใช้ได้
สถานการณ์ที่ 4 อัตราแลกเปลี่ยนเงิน	1	H	0.69	0.15	ใช้ไม่ได้
		A	0.80	0.20	ใช้ได้
	2	H	0.26	0.24	ใช้ได้
		A	0.28	0.23	ใช้ได้
	3	H	0.39	0.29	ใช้ได้
		A	0.35	0.25	ใช้ได้
สถานการณ์ที่ 5 การโดยสารเครื่องบิน	1	H	0.71	0.23	ใช้ได้
		A	0.78	0.23	ใช้ได้
	2	H	0.47	0.47	ใช้ได้
		A	0.45	0.45	ใช้ได้
	3	H	0.34	0.34	ใช้ได้
		A	0.45	0.45	ใช้ได้

ตารางค่าความยากง่ายและอำนาจจำแนก เรื่องปริมาณ

เรื่องปริมาณ	ข้อที่	เกณฑ์การให้คะแนน	ค่าความยาก	ค่าอำนาจจำแนก	สรุปผล
สถานการณ์ที่ 1 ดินสอ	1	H	0.54	0.25	ใช้ได้
		A	0.56	0.44	ใช้ได้
	2	H	0.29	0.29	ใช้ได้
		A	0.31	0.31	ใช้ได้
	3	H	0.28	0.22	ใช้ได้
		A	0.31	0.25	ใช้ได้
สถานการณ์ที่ 2 น้ำแครอท	1	H	0.65	0.31	ใช้ได้
		A	0.69	0.25	ใช้ได้
	2	H	0.61	0.26	ใช้ได้
		A	0.58	0.21	ใช้ได้
	3	H	0.64	0.34	ใช้ได้
		A	0.65	0.31	ใช้ได้
สถานการณ์ที่ 3 เครื่องเล่น MP3	1	H	0.61	0.21	ใช้ได้
		A	0.72	0.28	ใช้ได้
	2	H	0.23	0.23	ใช้ได้
		A	0.28	0.28	ใช้ได้
	3	H	0.53	0.38	ใช้ได้
		A	0.56	0.38	ใช้ได้
สถานการณ์ที่ 4 การทำเค้ก	1	H	0.70	0.30	ใช้ได้
		A	0.66	0.34	ใช้ได้
	2	H	0.06	0.06	ใช้ไม่ได้
		A	0.06	0.06	ใช้ไม่ได้
	3	H	0.16	0.16	ใช้ไม่ได้
		A	0.16	0.16	ใช้ไม่ได้
สถานการณ์ที่ 5 จำนวนสัตว์ของแดรี่ฮัท ฟาร์ม	1	H	0.63	0.38	ใช้ได้
		A	0.63	0.38	ใช้ได้
	2	H	0.48	0.48	ใช้ได้
		A	0.47	0.47	ใช้ได้
	3	H	0.34	0.34	ใช้ได้
		A	0.41	0.41	ใช้ได้

ตารางค่าความยากง่ายและอำนาจจำแนก เรื่องความไม่แน่นอน

เรื่องความไม่แน่นอน	ข้อที่	เกณฑ์การให้คะแนน	ค่าความยาก	ค่าอำนาจจำแนก	สรุปผล
สถานการณ์ที่ 1 การคิดค่าไฟฟ้า	1	H	0.47	0.31	ใช้ได้
		A	0.33	0.22	ใช้ได้
	2	H	0.61	0.24	ใช้ได้
		A	0.50	0.22	ใช้ได้
	3	H	0.56	0.22	ใช้ได้
		A	0.61	0.28	ใช้ได้
สถานการณ์ที่ 2 ทอยลูกเต๋า	1	H	0.62	0.22	ใช้ได้
		A	0.69	0.25	ใช้ได้
	2	H	0.50	0.22	ใช้ได้
		A	0.44	0.17	ใช้ไม่ได้
	3	H	0.61	0.24	ใช้ได้
		A	0.57	0.17	ใช้ไม่ได้
สถานการณ์ที่ 3 แนวโน้มประชากรไทย	1	H	0.63	0.15	ใช้ไม่ได้
		A	0.56	0.11	ใช้ไม่ได้
	2	H	0.46	0.31	ใช้ได้
		A	0.61	0.28	ใช้ได้
	3	H	0.54	0.24	ใช้ได้
		A	0.44	0.17	ใช้ไม่ได้
สถานการณ์ที่ 4 ราคาน้ำมัน	1	H	0.61	0.22	ใช้ได้
		A	0.54	0.20	ใช้ได้
	2	H	0.59	0.30	ใช้ได้
		A	0.50	0.22	ใช้ได้
	3	H	0.59	0.41	ใช้ได้
		A	0.39	0.22	ใช้ได้
สถานการณ์ที่ 5 ราคายางพารา	1	H	0.65	0.24	ใช้ได้
		A	0.63	0.22	ใช้ได้
	2	H	0.54	0.38	ใช้ได้
		A	0.56	0.39	ใช้ได้
	3	H	0.39	0.34	ใช้ได้
		A	0.50	0.44	ใช้ได้



ภาคผนวก ง
เครื่องมือที่ใช้ในการวิจัย

แบบสอบการรู้เรื่องคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2

สอบวัน.....ที่.....เดือน.....พ.ศ.เวลา

ชื่อ – นามสกุล

โรงเรียน

คำชี้แจง

1. ข้อสอบทั้งหมดมีจำนวน 4 เรื่อง ประกอบด้วย
 - 1.1 เรื่องปริภูมิและรูปทรงสามมิติ จำนวน 3 สถานการณ์
 - 1.2 เรื่องการเปลี่ยนแปลงและความสัมพันธ์ จำนวน 3 สถานการณ์
 - 1.3 เรื่องปริมาณ จำนวน 3 สถานการณ์
 - 1.4 เรื่องความไม่แน่นอน จำนวน 3 สถานการณ์
2. ตรวจสอบความเรียบร้อยแบบสอบก่อนลงมือทำ หากมีหน้าไม่ครบหรือแบบสอบมีปัญหาให้รีบแจ้งกรรมการคุมสอบ
3. เวลาในการทดสอบทั้งสิ้น **120 นาที** เริ่มทำแบบสอบได้หลังจากกรรมการคุมสอบให้สัญญาณ
4. ไม่อนุญาตให้ออกจากห้องสอบก่อนเวลา **30 นาที** หลังจากกรรมการคุมสอบให้สัญญาณ
5. เขียนคำตอบในแบบสอบด้วยลายมือตัวบรรจง โดยใช้ปากกาสีน้ำเงินหรือสีดำเท่านั้น ไม่อนุญาตให้ใช้ดินสอ ยกเว้นการทดเลข
6. ในขณะที่เข้าห้องสอบ ห้ามใช้เครื่องมือสื่อสารและเครื่องคำนวณทุกชนิด มิฉะนั้นจะถือว่าทุจริตในการสอบ

เรื่องปริภูมิและรูปทรงสามมิติ

สถานการณ์ที่ 1 : ค่าใช้จ่ายการตัดหญ้า



วิชัยต้องการตัดหญ้าสนามหญ้า
รูปสี่เหลี่ยมผืนผ้าที่มีขนาดกว้าง 30 วา ยาว
50 วา ผู้รับเหมาตัดหญ้ามัคค่าตัดหญ้า
ตารางเมตรละ 0.75 บาท วิชัยจะต้องจ่ายค่า
ตัดหญ้ากี่บาท

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากสถานการณ์ข้างต้นให้นักเรียนวาดภาพจำลองสนามหญ้าตามขนาดที่กำหนด และ
คำนวณหาค่าตัดหญ้าที่วิชัยต้องจ่าย

วาดภาพจำลองสนามหญ้า	วิธีทำ
	
	
	
	
	
	
	
สนามหญ้ามียี่พื้นที่ขนาด ตารางเมตร	วิชัยจะต้องจ่ายค่าตัดหญ้า บาท

ข้อที่ 3 : ผู้รับเหมาตัดหญ้ารายใหม่คิดราคาเหมาค่าตัดหญ้าสนามนี้ทั้งหมด 5,000 บาท ถ้า
นักเรียนเป็นวิชัยจะเลือกผู้รับเหมาตัดหญ้ารายใหม่หรือรายเก่า เพราะเหตุใดนักเรียนจึงเลือก
ผู้รับเหมาคนนั้น (พร้อมทั้งแสดงราคาเปรียบเทียบของผู้รับเหมาตัดหญ้าทั้งสอง
ประกอบการตัดสินใจ)

ตอบ

เพราะ.....

.....

.....

.....

.....

เรื่องปริภูมิและรูปทรงสามมิติ

สถานการณ์ที่ 2 : บรรจุนมกล่องลงในลังกระดาษ



นมกล่องทรงสี่เหลี่ยมมุม
ฉาก
กว้าง 4 เซนติเมตร
ยาว 6 เซนติเมตร
สูง 10 เซนติเมตร



ลังกระดาษแข็งทรงสี่เหลี่ยมมุมฉาก
กว้าง 12 เซนติเมตร
ยาว 30 เซนติเมตร
สูง 20 เซนติเมตร

สมชายนำนมกล่องบรรจุลงในลังกระดาษสี่เหลี่ยมมุมฉาก จัดวางในแนวตั้งจะบรรจุได้กี่กล่อง

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากสถานการณ์ข้างต้นนักเรียนมีวิธีการใดตรวจสอบว่าวางกล่องนมลงในลังกระดาษ
ได้พอดี พร้อมทั้งคำนวณหาปริมาตรของนมกล่องและลังกระดาษ และหาจำนวนนมกล่องที่
บรรจุลงในลังกระดาษ

วิธีการหาคำตอบ

แสดงวิธีทำ

ปริมาตรของนมกล่อง

ปริมาตรของลังกระดาษ



.....
.....
.....



.....
.....
.....

บรรจุนมกล่องได้ จำนวน กล่อง

ข้อที่ 3 : ถ้าหากเปลี่ยนจากนมกล่องเป็นนมขวดที่มีขนาด 200 มิลลิลิตร บรรจุในลังกระดาษ
ขนาดเท่าเดิมจัดวางในแนวตั้ง นักเรียนคิดว่าจำนวนนมขวดที่บรรจุลังกระดาษจะมีปริมาณ
เพิ่มขึ้น น้อยลงหรือเท่าเดิม เพราะเหตุใด

ตอบ ☐ เพิ่มขึ้น / ☐ เท่าเดิม / ☐ น้อยลง

เพราะ

.....
.....

เรื่องปริภูมิและรูปทรงสามมิติ

สถานการณ์ที่ 3 : ปูกระเบื้องพื้นห้องนอน



ปรีชาต้องการปูกระเบื้องพื้น
ห้องนอนขนาดกว้าง 3 เมตร ยาว 3
เมตร ปรีชาจะต้องใช้กระเบื้องอย่าง
น้อยกี่กล่องในการปูพื้นห้องนอน

*กระเบื้องปูพื้น ขนาด: 40 x 40

เซนติเมตร

บรรจุ 6 แผ่น / กล่อง

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากสถานการณ์ข้างต้น ปรีชาจะต้องใช้กระเบื้องอย่างน้อยกี่กล่องในการปูพื้นห้องนอน
นักเรียน มีวิธีการหาคำตอบอย่างไร เพื่อให้ได้ผลลัพธ์ที่ถูกต้อง

วิธีการหาคำตอบ

แสดงวิธีทำ

ตอบ ปรีชาจะต้องใช้กระเบื้องอย่างน้อย จำนวน กล่อง ในการปูพื้นห้องนอน

ข้อที่ 3 : ผู้รับเหมาแนะนำว่ากระเบื้องปูพื้น ขนาด 20x100 เซนติเมตร เหมาะสำหรับการปูพื้น
ห้องนอน ซึ่งภายในบรรจุกระเบื้อง 5 แผ่น/กล่อง เพื่อลดต้นทุนปริมาณกระเบื้องที่ใช้ ถ้า
นักเรียนเป็นปรีชาจะเชื่อคำแนะนำของผู้รับเหมาหรือไม่ เพราะเหตุใด (แสดงวิธีคิดคำนวณ
เพื่อสนับสนุนคำตอบ)

ตอบ ☐ เชื่อ / ☐ ไม่เชื่อ

เพราะ

เรื่องการเปลี่ยนแปลงและความสัมพันธ์

สถานการณ์ที่ 1 : การทำไร่นาสวนผสมตามหลักการเกษตรทฤษฎีใหม่ของพระบาทสมเด็จพระเจ้าอยู่หัวฯ

การทำไร่นาสวนผสมตามหลักการเกษตรทฤษฎีใหม่ของพระบาทสมเด็จพระเจ้าอยู่หัว แบ่งพื้นที่ออกเป็น 4 ส่วน ตามอัตราส่วน 30 : 30 : 30 : 10 หมายถึง พื้นที่ส่วนที่หนึ่งประมาณ 30 % ขุดสระเก็บกักน้ำ เพื่อใช้เก็บกักน้ำฝนในฤดูฝนและใช้เสริมการปลูกพืชในฤดูแล้ง พื้นที่ส่วนที่สองประมาณ 30 % ปลูกข้าวในฤดูฝน เพื่อใช้เป็นอาหารประจำวันในครัวเรือน พื้นที่ส่วนที่สามประมาณ 30 % ให้ปลูกไม้ผล ไม้ยืนต้นพืชผัก พืชไร่ พืชสมุนไพร ฯลฯ และพื้นที่ส่วนที่สี่ประมาณ 10 % ใช้เป็นที่อยู่อาศัย เลี้ยงสัตว์ และโรงเรือนอื่นๆ เป็นต้น

ที่มา : มูลนิธิชัยพัฒนา

นายแสนจัดสรรที่ดินเพื่อทำการเกษตรทฤษฎีใหม่ ดังนี้ บ้านของนายแสนอยู่ทางทิศใต้ของนาข้าว มีสระเก็บกักน้ำอยู่ทางทิศเหนือของแปลงผัก เล้าไก่อยู่ทางทิศตะวันตกเฉียงเหนือของบ้าน และนาข้าวอยู่ทางทิศตะวันออกของสระเก็บกักน้ำ อยากทราบว่าบ้านอยู่ทางทิศใดของสระเก็บกักน้ำ และนาข้าวอยู่ทางทิศใดของแปลงผัก

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

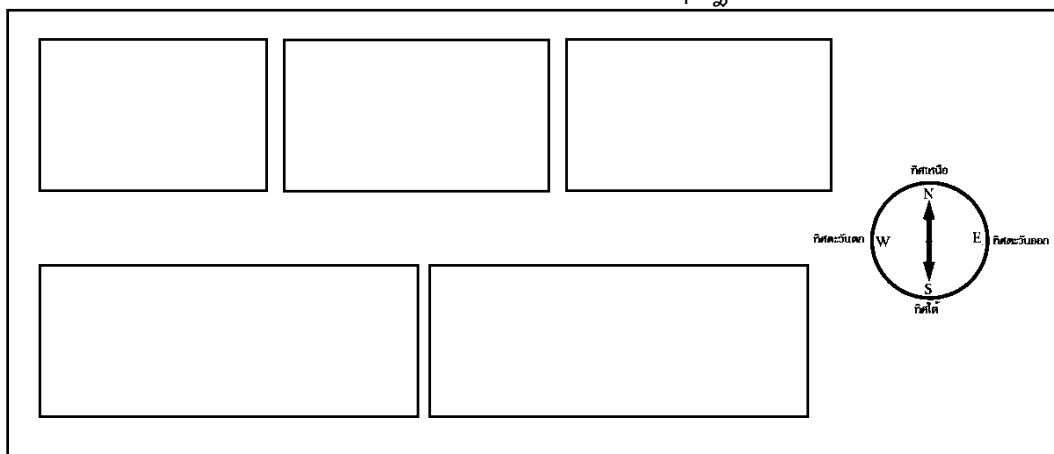
สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากคำถามสถานการณ์ข้างต้น นายแสนจัดสรรที่ดินเพื่อทำการเกษตรทฤษฎีใหม่ โดยพิจารณาตัวเลือกที่กำหนดให้ต่อไปนี้ แล้วเติมลงในช่องว่างของแผนผังด้านล่างให้ถูกต้องสอดคล้องตามคำบรรยายข้างต้น

1.บ้าน	2.นาข้าว	3.แปลงผัก	4.สระเก็บกักน้ำ	5.เล้าไก่
--------	----------	-----------	-----------------	-----------

แผนผังการจัดสรรที่ดินเพื่อทำการเกษตรทฤษฎีใหม่ของนายแสน



โปรดระบุทิศของสถานที่ต่อไปนี้ให้สอดคล้องกับข้อมูลในคำถาม

บ้านอยู่ทางทิศ.....ของสระเก็บกักน้ำ และนาข้าวอยู่ทางทิศ.....ของแปลงผัก

ข้อที่ 3 : ถ้านายแสนและนายเสก มีพื้นที่ทำกิน จำนวน 15 ไร่ ต้องการแบ่งพื้นที่ออกเป็น 4 ส่วน ตามหลักการเกษตรทฤษฎีใหม่ จะจัดสรรที่ดินแต่ละส่วนได้กี่ไร่ และอะไรเป็นปัจจัยหลักที่สามารถปรับเปลี่ยนและยืดหยุ่นได้ตามความเหมาะสมของพื้นที่ (พร้อมอธิบายเหตุผลสนับสนุนคำตอบ)

บ้านนายแสน จ.น่าน ฝนตกปานกลาง ห่างไกลชลประทาน		บ้านนายเสก จ.ชุมพร มีฝนตกตลอดปี	
การจัดสรรที่ดิน	จำนวนพื้นที่ (ระบุหน่วย)	การจัดสรรที่ดิน	จำนวนพื้นที่ (ระบุหน่วย)
ทำนา		ทำนา	
สระเก็บกักน้ำ		สระเก็บกักน้ำ	
ปลูกพืชผัก ผลไม้		ปลูกพืชผัก ผลไม้	
ที่อยู่อาศัย		ที่อยู่อาศัย	

เพราะ.....

.....

.....

.....

เรื่องการเปลี่ยนแปลงและความสัมพันธ์

สถานการณ์ที่ 2 : อัตราการเต้นของหัวใจ

อัตราการเต้นของหัวใจ (Heart Rate) คือ จังหวะการบีบตัวของหัวใจ เพื่อส่งเลือดแดงที่มีออกซิเจนไปยังส่วนต่างๆของร่างกาย อัตราการเต้นของหัวใจนับเป็นครั้งต่อวินาที (bpm) ย่อมาจากคำว่า “beat per minute” โดยทาง American College of Sports Medicine ได้กำหนดอัตราการเต้นของหัวใจสูงสุดของคนหนึ่งคน ไว้ดังนี้

อัตราการเต้นของหัวใจสูงสุดต่อนาทีแบบแยกชายหญิง

ชาย $\text{Max HR} = 214 - (0.8 \times \text{อายุ})$

หญิง $\text{Max HR} = 209 - (0.7 \times \text{อายุ})$

ตารางแสดงระดับความหนักของการออกกำลังกายต่ออัตราการเต้นของหัวใจ

Zone		เปอร์เซ็นต์ของค่า Max HR	ผลลัพธ์ที่ได้
ZONE 1 เบามาก		50% – 60% ของ ค่า Max HR (ระยะเวลาประมาณ 20-40 นาที)	ช่วยให้เลือดไหลเวียนดี และเอาไ้ฟืนฟูหลังจากออกกำลังกายหนักๆมา
ZONE 2 เบา		60% – 70% ของ ค่า Max HR (ระยะเวลาประมาณ 40-80 นาที)	พัฒนาได้นอดกลั้น สามารถเผาผลาญไขมันและคาร์โบไฮเดรตได้ดีมาก
ZONE 3 ปานกลาง		70% – 80% ของ ค่า Max HR (ระยะเวลาประมาณ 10-40 นาที)	เพิ่มความสามารถทางเอโรบิค ผักเพื่อชลอความล้าที่เกิดจากกรดแลคติก
ZONE 4 หนัก		80% – 90% ของ ค่า Max HR (ระยะเวลาประมาณ 2-10 นาที)	ออกกำลังกายหนักๆเพื่อเพิ่มความสามารถสูงสุดรวมทั้งความอด
ZONE 5 หนักมาก		90% – 100% ของ ค่า Max HR (ไม่ควรเกิน 5 นาที)	เอาไ้ใช้ในการพัฒนาศักยภาพสูงสุด ซึ่งจะทำแค่ช่วงสั้นๆไม่กี่นาทีเท่านั้น

ที่มา : <http://www.healthiie.com>.

ผู้ชายอายุ 25 ปีและผู้หญิงมีอายุ 30 ปี จะมีอัตราการเต้นของหัวใจสูงสุดต่อนาทีเท่ากันหรือไม่ พร้อมทั้งแสดงอัตราการเต้นของหัวใจสูงสุดต่อนาทีแบบแยกชายหญิง

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากสถานการณ์ข้างต้น ผู้ชายอายุ 25 ปีและผู้หญิงมีอายุ 30 ปี จะมีอัตราการเต้นของหัวใจสูงสุดตอนาที พร้อมทั้งแสดงอัตราการเต้นของหัวใจสูงสุดตอนาทีแบบแยกชายหญิง นักเรียนมีวิธีการหาคำตอบอย่างไร เพื่อให้ได้ผลลัพธ์ที่ถูกต้อง

วิธีการหาคำตอบ

แสดงวิธีทำ

.....

.....

.....

.....

ตอบ

ข้อที่ 3 : นายพุ่ม อายุ 30 ปี ต้องการออกกำลังกายเพื่อลดน้ำหนัก ระยะเวลา 30 นาที ระดับความหนักของการออกกำลังกายที่ส่งผลต่ออัตราการเต้นของหัวใจ ควรอยู่ระดับโซนไหน ถ้าเกินโซนที่กำหนดจะส่งผลอย่างไร

ตอบ ความหนักการออกกำลังกายที่ส่งผลต่ออัตราการเต้นของหัวใจของนายพุ่มควรอยู่ระดับโซน

เพราะ

.....

.....

.....

.....

.....

เรื่องการเปลี่ยนแปลงและความสัมพันธ์

สถานการณ์ที่ 3 : การโดยสารเครื่องบิน



เครื่องบินโดยสารของสายการบินแห่งหนึ่งบินจากสนามบินดอนเมืองไปสนามบินภูเก็ต ด้วยอัตราความเร็วเฉลี่ย 900 กิโลเมตร ต่อชั่วโมงในสภาพอากาศนิ่ง (ไม่มีกระแสลม) ขณะที่เครื่องบินออกจากสนามบินดอนเมืองมีกระแสลมพัดจากกรุงเทพฯ ไปภูเก็ต ด้วยอัตราเร็ว 20 กิโลเมตรต่อชั่วโมง สนามบินดอนเมืองห่างจากสนามบินภูเก็ต 860 กิโลเมตร เครื่องบินลำนี้มีกำหนดเดินทางออกจากสนามบินดอนเมือง เวลา 10.30 น. เครื่องบินลำนี้ถึงสนามบินภูเก็ตเวลาที่โมง

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากสถานการณ์ข้างต้น เครื่องบินออกจากสนามบินดอนเมืองเวลา 10.30 น. เครื่องบินถึงสนามบินภูเก็ตเวลาเท่าใด นักเรียนมีวิธีการหาคำตอบอย่างไร เพื่อให้ได้ผลลัพธ์ที่ถูกต้อง

วิธีการหาคำตอบ

แสดงวิธีทำ.....

ตอบ เครื่องบินลำนี้ถึงสนามบินภูเก็ตเวลา (พร้อมทั้งระบุหน่วย)

ข้อที่ 3 : เมื่อเวลา 20:45 น. เครื่องบินลำนี้บินกลับสนามบินดอนเมือง ในขณะที่กระแสลมยังมีทิศทางเช่นเดิมและอัตราเร็วเท่าเดิม เครื่องบินมาถึงสนามบินดอนเมืองใช้ระยะเวลาในการเดินทางมากกว่าหรือน้อยกว่าการเดินทางไปภูเก็ต ปัจจัยใดที่ทำให้การเดินทางของเครื่องบินลำนี้มีระยะเวลาในการเดินทางเป็นเช่นนั้น (แสดงวิธีคิดคำนวณเพื่อสนับสนุนคำตอบ)

ตอบ ระยะเวลาเดินทางกลับของเครื่องบินลำนี้ ☐ มากกว่า / ☐ น้อยกว่า การเดินทางจากสนามบินดอนเมืองไปภูเก็ต

เพราะ

เรื่องปริมาณ

สถานการณ์ที่ 1 : น้ำแครอท



=



น้ำแครอท 1 กล่องใหญ่ เท่ากับ น้ำแครอท
3 กล่องเล็ก



=



น้ำแครอท 1 กล่องเล็ก สามารถเทได้ 2
แก้ว

ปิยวรรณมีน้ำแครอท 1 กล่องใหญ่ กับน้ำแครอท 3 กล่องเล็ก ปิยวรรณจะเทน้ำแครอท
ได้กี่แก้ว

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากสถานการณ์ข้างต้น ปิยวรรณมีน้ำแครอท 1 กล่องใหญ่ กับน้ำแครอท 3 กล่องเล็ก
ปิยวรรณจะเทน้ำแครอทได้กี่แก้ว นักเรียนมีวิธีการหาคำตอบอย่างไร เพื่อให้ได้ผลลัพธ์ที่ถูกต้อง
วิธีการหาคำตอบ

แสดงวิธีทำ

.....

.....

.....

.....

ตอบ

ข้อที่ 3 : ปิยวรรณดื่มน้ำแครอทวันละ 2 แก้ว ภายในระยะเวลาหนึ่งสัปดาห์ปิยวรรณดื่มน้ำ
แครอทหมดหรือไม่ เพราะเหตุใด (แสดงวิธีคิดคำนวณเพื่อสนับสนุนคำตอบ)

ตอบ ☐ ดื่มหมด ☐ ดื่มไม่หมด

เพราะ.....

.....

.....

เรื่องปริมาณ

สถานการณ์ที่ 2 : เครื่องเล่น MP3

ร้านมิวสิคฟอรั่ย์ จัดมหกรรมลดสินค้าทั้งร้าน เมื่อซื้อสินค้าจำนวนสองชิ้นขึ้นไป ร้านมิวสิคฟอรั่ย์ ลดราคา 30% จากราคาขายของสินค้าเหล่านี้ ชิมัมพร มีเงินอยู่ 2,000 บาท สามารถซื้อสินค้าชนิดใดได้บ้าง

เครื่องเล่น MP3	หูฟัง	ลำโพง
		
ราคา 1,290 บาท	ราคา 590 บาท	ราคา 890 บาท

หมายเหตุ : ราคาสินค้าที่ปรากฏเป็นราคาเต็ม

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากสถานการณ์ข้างต้น ชิมัมพร มีเงินอยู่ 2,000 บาท สามารถซื้อสินค้าชนิดใดได้บ้าง นักเรียนมีวิธีการหาคำตอบอย่างไร เพื่อให้ได้ผลลัพธ์ที่ถูกต้อง

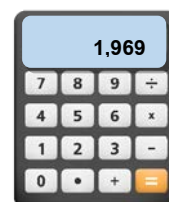
วิธีการหาคำตอบ

แสดงวิธีทำ

.....

ชิมัมพรสามารถซื้อสินค้า

ข้อที่ 3 : ชิมัมพรคำนวณราคาเต็มของเครื่องเล่น MP3 หูฟังและลำโพงด้วยเครื่องคิดเลขได้จำนวน 1,969 บาท คำตอบที่ได้ไม่ถูกต้องหรือไม่ นักเรียนคิดว่า ชิมัมพรคำนวณผิดพลาด เพราะอะไร และราคาที่ถูกต้องกี่บาท



คำตอบที่ได้จากการคำนวณเครื่องเลข ☐ ถูกต้อง / ☐ ไม่ถูกต้อง

เพราะ

.....

ราคาสินค้าทั้งหมด คิดเป็นจำนวนเงิน บาท

เรื่องปริมาณ

สถานการณ์ที่ 3 : จำนวนสัตว์ของแดรี่ฮัท ฟาร์ม



แดรี่ฮัท ฟาร์ม จ.พังงา สถานที่ท่องเที่ยวแบบฟาร์มสัตว์แห่งแรกของภาคใต้ รวบรวมความน่ารัก ของสัตว์น้อยหลายนชนิดมาไว้ในที่เดียวกัน อาทิเช่น ลามะ แกะ ม้า และวัว โดยมีจำนวนลามะ $\frac{1}{3}$ ของวัว จำนวนแกะ $\frac{1}{2}$ ของม้า จำนวนม้า $\frac{1}{2}$ ของวัว และมีจำนวนวัวทั้งหมด 12 ตัว อยากทราบว่าแดรี่ฮัท ฟาร์ม มีสัตว์แต่ละชนิดกี่ตัว

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากสถานการณ์ข้างต้น ให้นักเรียนนำเสนอข้อมูลจำนวนสัตว์แต่ละชนิด พร้อมทั้งแสดงวิธีหาจำนวนสัตว์แต่ละชนิดของแดรี่ฮัท ฟาร์ม

วิธีการนำเสนอข้อมูล

นำเสนอข้อมูลจำนวนสัตว์แต่ละชนิด	จงหาจำนวนสัตว์แต่ละชนิดของฟาร์ม วิธีทำ
	
	
	
	

ข้อที่ 3 : ตลอดระยะเวลา 3 ปีที่ผ่านมา แดรี่ฮัท ฟาร์ม ได้เพิ่มจำนวนสัตว์ทุกชนิด ปีละ 2 ตัว ยกเว้นลามะที่เพิ่มมาปีละ 3 ตัว ถ้าผ่านไป 10 ปี จำนวนลามะจะมีจำนวนมากกว่าสัตว์ชนิดอื่น ถูกต้องหรือไม่ เพราะเหตุใด (แสดงวิธีคิดคำนวณเพื่อสนับสนุนคำตอบ)

ตอบ ☐ ถูกต้อง / ☐ ไม่ถูกต้อง

เพราะ

.....

.....

เรื่องความไม่แน่นอน

สถานการณ์ที่ 1 : การคิดค่าไฟฟ้า

อัตราค่าไฟฟ้าประเทศไทย ประเภทบ้านพักอาศัยและปริมาณการใช้ไฟฟ้าเกินกว่า 150 หน่วยต่อเดือน มีอัตราดังต่อไปนี้ (ไม่รวมภาษีมูลค่าเพิ่ม และไม่รวมค่า FT)

พลังงานไฟฟ้า	หน่วยที่	ค่าพลังงานไฟฟ้า
ต่ำกว่า 35 หน่วย (กิโลวัตต์ต่อชั่วโมง)	ไม่มีการใช้ไฟฟ้า	83.18 บาท
35 หน่วย (กิโลวัตต์ต่อชั่วโมง)	หน่วยที่ 1 - 35	85.21 บาท
115 หน่วยต่อไป	หน่วยที่ 36 - 150	1.12 บาท
250 หน่วยต่อไป	หน่วยที่ 151 – 400	2.13 บาท
เกินกว่า 400 หน่วย	หน่วยที่ 401 เป็นต้นไป	2.42 บาท

ที่มา : <http://www.mea.or.th> การไฟฟ้านครหลวง

ในเดือนสิงหาคมบ้านหลังหนึ่ง ใช้ไฟฟ้าไป 500 หน่วย อยากทราบว่าบ้านหลังนี้มีค่าไฟฟ้าที่ต้องจ่ายกี่บาท (ไม่รวมภาษีมูลค่าเพิ่ม และไม่รวมค่า FT)

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากข้อมูลสถานการณ์ข้างต้น ในเดือนสิงหาคมบ้านหลังหนึ่ง ใช้ไฟฟ้าไป 500 หน่วย อยากทราบว่าบ้านหลังนี้มีค่าไฟฟ้าที่ต้องจ่ายกี่บาท (ไม่รวมภาษีมูลค่าเพิ่ม และไม่รวมค่า FT) นักเรียนมีวิธีการหาคำตอบ เพื่อให้ได้ผลลัพธ์ที่ถูกต้อง

วิธีการหาคำตอบ

แสดงวิธีการหาคำตอบ

.....
.....

ตอบ

ข้อที่ 3 : ต่อมาบ้านหลังนี้ได้ต่อเติมบ้าน โดยแบ่งเป็น 2 ส่วน ทำการติดตั้งระบบไฟฟ้าใหม่ บ้านหลังเดิม ยังคงใช้ไฟฟ้า 500 หน่วยต่อเดือน ส่วนที่ต่อเติมแบ่งเป็นสองห้องติดกัน โดยห้องแรกใช้ไฟฟ้า 300 หน่วย และห้องที่สองใช้ไฟฟ้า 200 หน่วยต่อเดือน อยากทราบว่าบ้านหลังใหม่นี้ค่าไฟมากกว่ากัน และมากกว่าเท่าไร (ไม่รวมภาษีมูลค่าเพิ่ม และไม่รวมค่า FT) (แสดงวิธีคิดคำนวณเพื่อสนับสนุนคำตอบ)

ตอบ ☐ บ้านหลังเดิม ☐ บ้านส่วนที่ต่อเติม

เพราะ

เรื่องความไม่แน่นอน

สถานการณ์ที่ 2 : ราคาน้ำมัน

วันที่	เบนซิน 91	เบนซิน 95	ดีเซล
9 ม.ค. 2560	27.68	27.95	29.69
23 ม.ค. 2560	27.08	27.35	29.39
6 ก.พ. 2560	27.38	27.65	29.39
20 ก.พ. 2560	27.98	28.25	29.79
6 มี.ค. 2560	27.58	27.85	29.59
20 มี.ค. 2560	27.08	27.35	29.59
3 เม.ย. 2560	26.78	27.05	29.09
17 เม.ย. 2560	26.28	26.55	28.49
1 พ.ค. 2560	26.28	26.55	28.09
15 พ.ค. 2560	26.28	26.55	27.69

ราคาขายปลีกมาตรฐาน ในเขตกทม. และปริมณฑล ประจำปี พ.ศ. 2560 (หน่วย : บาท/ลิตร)

ที่มา : <http://www.ptplc.com>

จากข้อมูลราคาน้ำมันข้างต้น อยากทราบว่าวันที่เท่าไรที่ราคาน้ำมันทั้ง 3 ชนิดราคาสูงสุด และสูงกว่าวันที่ราคาต่ำสุดอยู่ที่บาท (ตอบแยกราคาน้ำมันแต่ละชนิด)

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากข้อมูลสถานการณ์ข้างต้น นักเรียนคิดว่าจากข้อมูลราคาน้ำมันข้างต้น ควรนำเสนอ

ข้อมูลด้วยรูปแบบใด เหมาะสมที่สุด

นำเสนอข้อมูลด้วยรูปแบบ

แสดงวิธีการนำเสนอข้อมูล

อยากทราบว่าวันที่เท่าไรที่ราคาน้ำมันทั้ง 3 ชนิดราคาสูงที่สุด

แสดงวิธีทำ

.....

.....

.....

.....

ตอบ วันที่ เป็นวันที่ราคาน้ำมันทั้ง 3 ชนิดราคาสูงที่สุด

ข้อที่ 3 : สมชายอาศัยอยู่ที่จังหวัดนครศรีธรรมราช ซึ่งราคาน้ำมันมีราคาสูงกว่าราคาขายปลีกมาตรฐาน สมหมายอาศัยอยู่ที่จังหวัดระยอง ราคาน้ำมันเท่ากับราคาขายปลีกมาตรฐาน ปัจจัยใดที่ส่งผลให้ราคาขายปลีกน้ำมันจำหน่ายภาคอื่นๆของประเทศไทย มีราคาสูงกว่าหรือต่ำกว่าราคาขายปลีกมาตรฐาน เพราะเหตุใด

เพราะ

.....

.....

.....

.....

เรื่องความไม่แน่นอน

สถานการณ์ที่ 3 : ราคายางพารา

วัน/เดือน/ปี	ราคาประมูล ณ ตลาดกลางยางพารา อ.หาดใหญ่ (บาท)	
	ยางแผ่นดิบ	น้ำยางสด (ณ โรงงาน)
28 เมษายน 2560	67.80	63.00
27 เมษายน 2560	67.30	62.00
26 เมษายน 2560	67.30	61.00
25 เมษายน 2560	66.80	60.50
24 เมษายน 2560	65.90	60.00

ที่มา : <http://www.rubber.co.th> การยางแห่งประเทศไทย

หมายเหตุ : ราคาที่ปรากฏเป็นราคาต่อกิโลกรัม

พวงศัศกัฒมียางแผ่นดิบ 7 กิโลกรัม และน้ำยางสด 5 กิโลกรัม ถ้านักเรียนเป็นพวงศัศกัฒมี

จะขายยางแผ่นดิบและน้ำยางสดวันไหน ได้ราคาดีที่สุด และพวงศัศกัฒมีขายได้กี่บาท

ข้อที่ 1 : จากสถานการณ์ข้างต้น ให้นักเรียนตอบคำถามต่อไปนี้

สิ่งที่โจทย์กำหนดให้ :

สิ่งที่โจทย์ถาม :

สร้างประโยคสัญลักษณ์ :

ข้อที่ 2 : จากข้อมูลสถานการณ์ข้างต้น พวงศัศกัฒมีจะขายยางแผ่นดิบและน้ำยางสดวันไหนได้ราคาดีที่สุด และพวงศัศกัฒมีขายได้กี่บาท นักเรียนมีวิธีการหาคำตอบอย่างไร เพื่อให้ได้ผลลัพธ์ที่ถูกต้อง

วิธีการหาคำตอบ

แสดงวิธีทำ

.....

.....

ตอบ วันที่ ขายยางแผ่นดิบ บาท และขายราคาน้ำยางสด บาท

ข้อที่ 3 : จากข้อมูลสถานการณ์ข้างต้น นักเรียนสามารถคาดการณ์แนวโน้มราคาขายได้หรือไม่ เพราะอะไร

ตอบ คาดการณ์แนวโน้มราคาขาย ☐ ได้ ☐ ไม่ได้

เพราะ

.....

.....

.....

ภาคผนวก จ

การวิเคราะห์ข้อมูลด้วย โปรแกรม G String IV Version 6.1.1

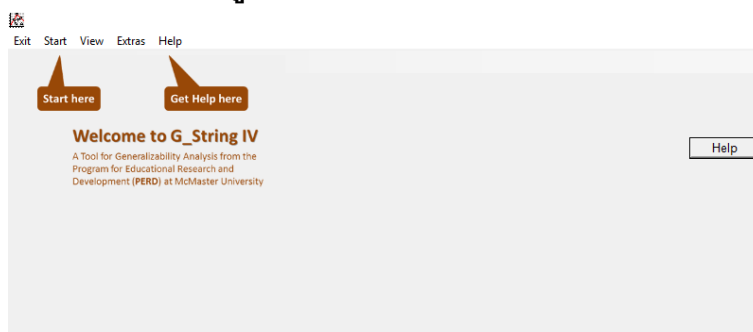
- การออกแบบเพื่อวิเคราะห์ Generalizability Theory
- โปรแกรม G String IV
- ผลการวิเคราะห์ข้อมูล

การออกแบบเพื่อวิเคราะห์ Generalizability Theory

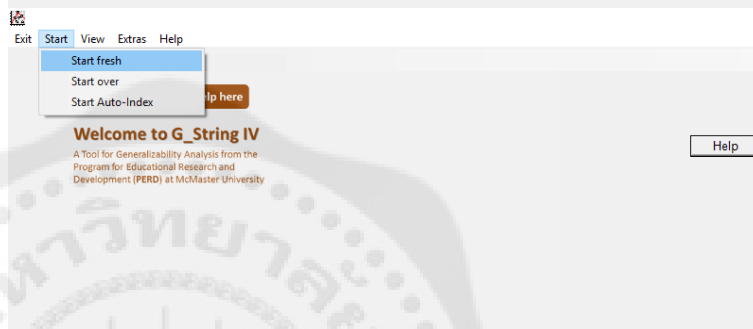
การพัฒนาและตรวจสอบคุณภาพแบบสอบการรู้เรื่องคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบและกฎการให้คะแนนต่างกันโดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) ในการศึกษาคุณภาพคุณภาพแบบสอบการรู้เรื่องคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีความยาวแบบสอบและกฎการให้คะแนนต่างกันครั้งนี้ ประกอบด้วยนักเรียนจำนวน 400 คน ทำข้อสอบจำนวน 36 ข้อ และกฎเกณฑ์การให้คะแนน 2 แบบ จึงออกแบบการวิเคราะห์เป็นแบบไขว้ทั้งหมด นักเรียนทุกคนทำข้อสอบทุกข้อ ข้อสอบทุกข้อถูกตรวจด้วยกฎเกณฑ์การให้คะแนนสองวิธี $p \times i$ เมื่อ p แทน ผู้สอบ และ i แทน จำนวนข้อสอบ โดยใช้โปรแกรมสำเร็จรูป G String IV Version 6.1.1 วิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิง (G – Coefficient) ของคะแนนแบบสอบการรู้เรื่องคณิตศาสตร์ แล้วเปรียบเทียบค่าสัมประสิทธิ์ความน่าเชื่อถือสำหรับการตัดสินใจเชิงสัมพัทธ์และค่าสัมประสิทธิ์ความน่าเชื่อถือสำหรับการตัดสินใจเชิงสมบูรณ์ของความยาวแต่ละระดับ และกฎเกณฑ์การให้คะแนนทั้งสองแบบ

ขั้นตอนการวิเคราะห์ข้อมูลด้วยโปรแกรม G String IV

1. เปิดโปรแกรม G String IV เลือก Start



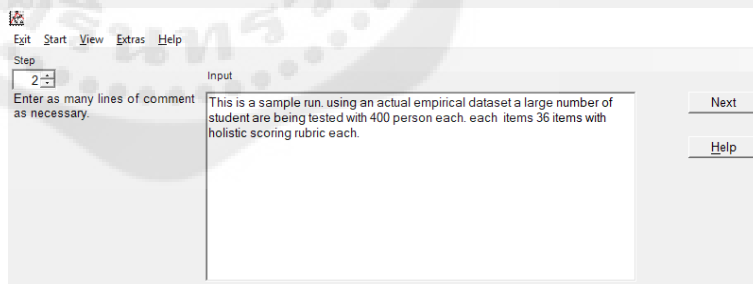
2. เลือกเมนู Start > Start fresh



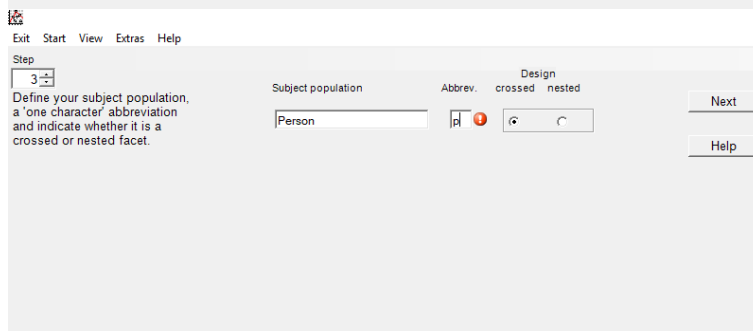
3. ในช่อง Input ให้ตั้งชื่อเรื่อง เลือก Next



4. ในช่อง Input ให้เขียนคำอธิบายการวิเคราะห์ เลือก Next



5. ในช่อง Subject population ให้ตั้งชื่อของกลุ่มตัวอย่าง และเลือกออกแบบการวัดแบบไขว้ทั้งหมด เลือก Next



6. ใส่จำนวนฟาเซทที่ต้องการใช้วิเคราะห์การวิเคราะห์ครั้งนี้ใช้ฟาเซท จำนวน 1 ฟาเซท
เลือก Next

Exit Start View Extras Help

Step 4

How many facets are there (excluding "subjects")?

Number of facets: 1

Next Help

7. กำหนดชื่อแต่ละฟาเซทที่ต้องการวิเคราะห์ รวมทั้งเลือกการออกแบบ crossed
เลือก Next

Exit Start View Extras Help

Step 5

For each facet give a descriptive name, a 'one character' abbreviation and indicate whether it is a crossed or nested facet.

Facet name: Items

Abbrev.: i

Design: crossed nested

Next Help

8. ตรวจสอบข้อมูลของแต่ละฟาเซท
เลือก Next

Exit Start View Extras Help

Step 6

Drag-and-Drop the nested facets to the appropriate 'parents'

From: To:

Next Help

9. เลือกตัวแปรจากข้อมูลที่บันทึกไว้ เลือก Next

Exit Start View Extras Help

Step 7

Mark the item for which the data file switches lines (records).

☒ p

☐ i

Next Help

10. ใส่จำนวนกลุ่มตัวอย่างที่ใช้ศึกษา
การวิจัยครั้งนี้ใช้กลุ่มตัวอย่างจำนวน 400 คน
เลือก Next

Exit Start View Extras Help

Step 8

Enter sample sizes for "Person" within the appropriate cell.

p 400

Next Help

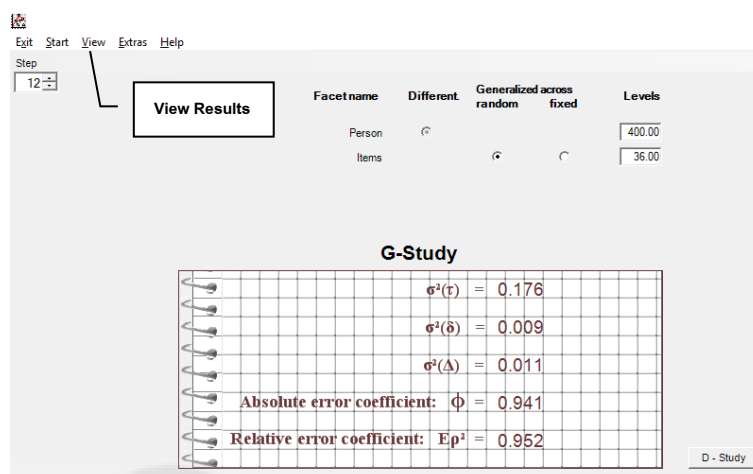
11. ใส่จำนวนข้อที่ใช้
สำหรับการวิเคราะห์
ฟาเซท การวิจัยครั้งนี้ใช้
จำนวนข้อทั้งหมด 36 ข้อ
เลือก Next

12. นำเข้าข้อมูลที่จัดเตรียม
ไว้ในโปรแกรม Notepad
หรือข้อมูลที่เป็นไฟล์ .txt
เลือก Next

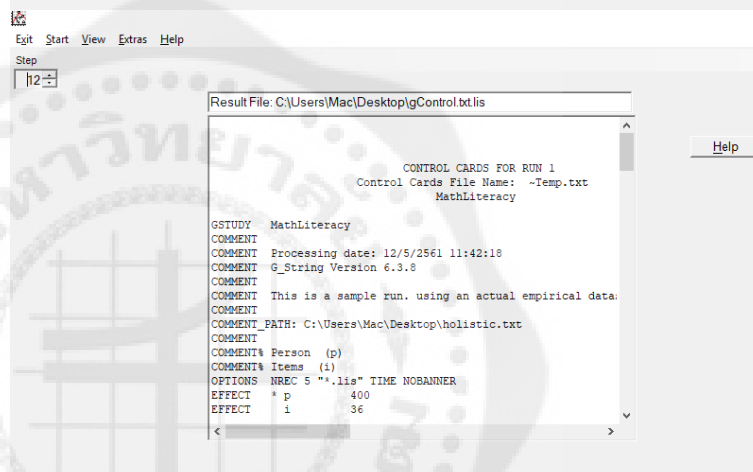
13. ถ้าหากต้องการศึกษา
ค่าอื่นๆ ศึกษาจากคู่มือเบร
นเนน เลือก Next

14. แสดงผลการวิเคราะห์
G-Study เลือก D-Study
เพื่อหาผลการวิเคราะห์
ค่า D-Study ในแต่ละ
เงื่อนไขที่ออกแบบไว้

15. เลือก View > View Results เพื่อต้องการดูข้อมูลคำสั่งการทำงาน และผลการวิเคราะห์ข้อมูล



16. ภาพแสดงคำสั่งการทำงานของโปรแกรม



17. ตาราง ANOVA แสดงค่าผลการวิเคราะห์ความแปรปรวนของแต่ละฟาเซทที่มีกฎเกณฑ์การให้คะแนนแบบภาพรวม

ANOVA TABLE FOR RUN 1
MathLiteracy

Effect	df	T	SS	MS	VC
p	399	2654.93293	2654.93293	6.65397	0.17591
i	35	1049.24801	1049.24801	29.97851	0.07414
pi	13965	8190.93266	4486.75172	0.32129	0.32129
Mean		0.00000			
Total	14399		8190.93265		

Grand Mean: 0.00000

18. ตาราง ANOVA แสดงค่าผลการวิเคราะห์ความแปรปรวนของแต่ละฟาเซทที่มีกฎเกณฑ์การให้คะแนนแบบแยกส่วน

ANOVA TABLE FOR RUN 1
MathLiteracy

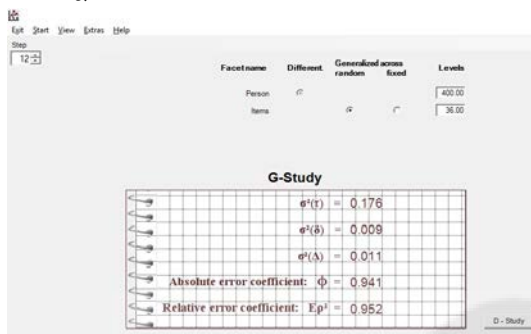
Effect	df	T	SS	MS	VC
p	399	20570.03445	20570.03444	51.55397	1.37495
i	35	10047.00390	10047.00389	287.05725	0.71250
pi	13965	59326.47890	28709.44056	2.05581	2.05581
Mean		0.00001			
Total	14399		59326.47889		

Grand Mean: 0.00002

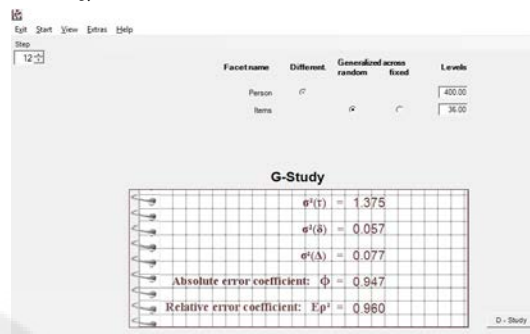
ผลการวิเคราะห์การตรวจสอบคุณภาพรูปแบบของชุดเครื่องมือ

ผลการศึกษาเชิงสรุปร่างอิง (Generalizability Study or G - Study)

กฎเกณฑ์การให้คะแนนแบบภาพรวม



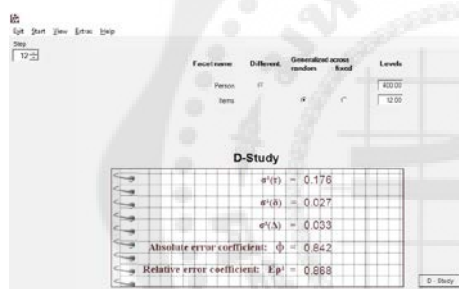
กฎเกณฑ์การให้คะแนนแบบแยกส่วน



ผลการศึกษาเชิงตัดสินใจ (Decision Study or D - Study)

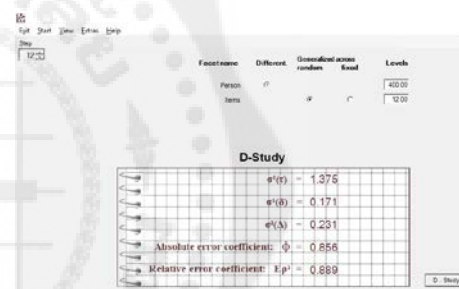
กฎเกณฑ์การให้คะแนนแบบภาพรวม

ความยาวแบบสอบ12ข้อ

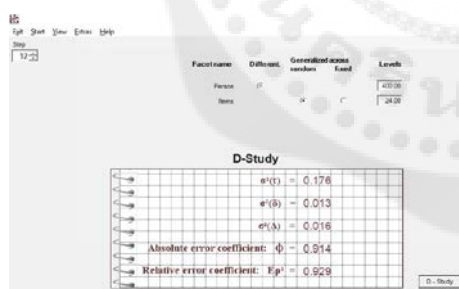


กฎเกณฑ์การให้คะแนนแบบแยกส่วน

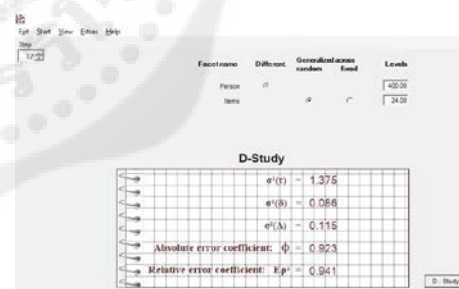
ความยาวแบบสอบ12ข้อ



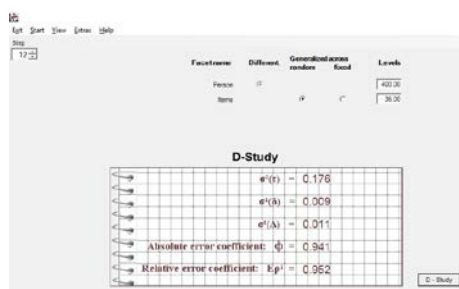
ความยาวแบบสอบ24ข้อ



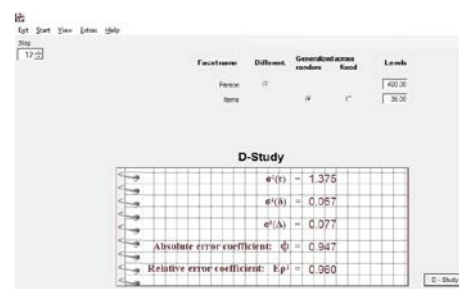
ความยาวแบบสอบ24ข้อ



ความยาวแบบสอบ36ข้อ



ความยาวแบบสอบ36ข้อ





ประวัติย่อผู้วิจัย

ชื่อ ชื่อสกุล นายภานุพงศ์ อึ้งสืบเชื้อ
 วันเดือนปีเกิด 7 มกราคม 2531
 สถานที่อยู่ปัจจุบัน 5 ถนนปฏิพัทธ์ ตำบลตลาดเหนือ อำเภอเมือง
 จังหวัดภูเก็ต 83000

ประวัติการศึกษา

พ.ศ. 2549 มัธยมศึกษาตอนปลาย
 จาก โรงเรียนสตรีภูเก็ต
 พ.ศ. 2555 ครุศาสตรบัณฑิต (ค.บ.) สาขาวิชาการศึกษาปฐมวัย
 จาก มหาวิทยาลัยราชภัฏภูเก็ต
 พ.ศ. 2561 การศึกษามหาบัณฑิต (กศ.ม.) สาขาวิชาการวิจัย
 และพัฒนาศักยภาพมนุษย์ แขนงวิชาการทดสอบ
 และวัดผลการศึกษา
 จาก มหาวิทยาลัยศรีนครินทรวิโรฒ