



ใบรับรองวิทยานิพนธ์
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

ศึกษาศาสตรมหาบัณฑิต (วิจัยและประเมินผลการศึกษา)

ปริญญา

วิจัยและประเมินผลการศึกษา	การศึกษา
สาขา	ภาควิชา
เรื่อง ความตรงสัมพันธ์ของค่าประมาณความสามารถจากการสอบแบบปรับเหมาะรูปพีระมิด ข้างตัด	
Relative Validity of the Ability Estimates From the Truncated Pyramid Adaptive Testing	
นามผู้วิจัย นางมนัสนันท์ มากสวาสดี	
ได้พิจารณาเห็นชอบโดย	
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	(รองศาสตราจารย์สุจินต์ เลี้ยงจรรยารัตน์, ศษ.ด.)
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	(รองศาสตราจารย์ภาวิณี ศรีสุขวัฒนานันท์, ศษ.ด.)
หัวหน้าภาควิชา	(ผู้ช่วยศาสตราจารย์สุดารัตน์ สารสว่าง, Ph.D.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์กัญญา ชีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ เดือน พ.ศ.

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วิทยานิพนธ์

เรื่อง

ความตรงสัมพันธ์ของค่าประมาณความสามารถจากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด

Relative Validity of the Ability Estimates from the Truncated Pyramid Adaptive Testing

โดย

นางมนัสนันท์ มากสวัสดิ์

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

เพื่อความสมบูรณ์แห่งปริญญาศึกษาศาสตรมหาบัณฑิต (วิจัยและประเมินผลการศึกษา)

พ.ศ. 2553

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

มนัสนันท์ มากสวาสดี 2553: ความตรงสัมพันธ์ของค่าประมาณความสามารถจากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด ปริญญศึกษาศาสตรมหาบัณฑิต (วิจัยและประเมินผลการศึกษา) สาขาวิจัยและประเมินผลการศึกษา ภาควิชาการศึกษา อาจารย์ที่ปรึกษา วิทยานิพนธ์หลัก: รองศาสตราจารย์สุจินต์ เลี้ยงจรรยาวัณ, ศษ.ด. 113 หน้า

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาความตรงสัมพันธ์ของค่าประมาณความสามารถจากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดที่มีวิธีการประมาณค่าความสามารถ 3 วิธี พลวิจัย เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนแผนวิทยาศาสตร์-คณิตศาสตร์ ภาคเรียนที่ 2 ปีการศึกษา 2552 โรงเรียนเศรษฐบุตรบำเพ็ญ สังกัดสำนักงานการศึกษาขั้นพื้นฐานกรุงเทพมหานครเขต 2 จำนวน 127 คน เครื่องมือที่ใช้ในการวิจัย คือแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ผู้วิจัยได้ดำเนินการ 2 ขั้นตอน คือ 1) ขั้นการสร้างและพัฒนาแบบสอบ โดยเขียนข้อสอบและตรวจสอบความตรงตามเนื้อเรื่องจากการพิจารณาโดยผู้เชี่ยวชาญและทดสอบครั้งที่ 1 กับกลุ่มนักเรียนจำนวน 150 คน วิเคราะห์คุณภาพข้อสอบตามทฤษฎีการสอบแบบดั้งเดิม ทดสอบครั้งที่ 2 กับกลุ่มนักเรียนจำนวน 554 คน ได้ค่าสัมประสิทธิ์ความเที่ยงแอลฟาของแบบสอบเท่ากับ 0.84 และคำนวณค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ 2) ขั้นการวิจัยเพื่อทดสอบสมมติฐาน ให้พลวิจัยทำการสอบ 2 วิธี คือ ตอบคำถามด้วยการเขียนตอบแบบดั้งเดิมทั้ง 54 ข้อ และตอบคำถามผ่านจอคอมพิวเตอร์แบบปรับเหมาะรูปพีระมิดข้างตัด โดยมีสมมติฐานของงานวิจัยว่าความสัมพันธ์ระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด 3 วิธี คือ 1) ประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก 2) ประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ และ 3) ประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม มีความสัมพันธ์กัน

ผลการวิจัย พบว่า ความสัมพันธ์ระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด 3 วิธี กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

Manassanan Maksavad 2010: Relative Validity of the Ability Estimates from the Truncated Pyramid Adaptive Testing. Master of Education (Educational Research and Evaluation), Major Field: Educational Research and Evaluation, Department of Education. Thesis Advisor: Associate Professor Sujin Liengjaroonrat, Ph.D. 113 pages.

The purpose of the research was to study the relative validity the estimating ability in integrated science process skill of the truncated pyramid adaptive test based on the estimating ability in three methods. The subjects of the study consisted of 127 students of Matayomsuksa 4, Science-Mathematics program the second semester of the academic year 2009, Setthabutbumpen School, under the office of Bangkok Education Service Area Zone 2. The research instrument was integrated science process skill test, which composed 2 steps: inventing and developing the test, and doing research in order to test hypotheses. The content validity was proved by the experts, and the result was first tested with the group of 150 students, analyzing the quality of the test by conventional test theory. The result was, then, second tested with the group of 554 students, calculate the item parameter which yielded Cronbach Alpha coefficient reliability 0.84. The doing research in order to test hypotheses step, the researcher had the subjects do the test in 2 ways: writing the answers by the conventional test of 54 items and answering through computers of the truncated pyramid adaptive test. The hypotheses of the research was that the correlation between the estimating ability in the truncated pyramid adaptive test by the estimating ability of average difficulty of the correct answers, the estimating ability of average difficulty of all the items done and the estimating ability of the level of difficulty of the last item done with the estimating ability of conventional test significantly related at the level of .01 according to Pearson's correlate coefficient which was the relative validity.

The result of the research was the correlation between the estimating ability of the truncated pyramid adaptive test in three methods with the estimating ability of conventional test was significantly related at the level of .01.

Student's signature

Thesis Advisor's signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยความกรุณาอย่างสูงยิ่งของ รศ.ดร.สุจินต์ เลี้ยงจรรยารัตน์ ประธานกรรมการที่ปรึกษาวิทยานิพนธ์หลัก รศ.ดร.ภาวิณี ศรีสุขวัฒนานันท์ กรรมการที่ปรึกษาวิทยานิพนธ์ร่วม ที่กรุณาให้คำแนะนำ ให้คำปรึกษา ตลอดจนให้ความช่วยเหลือ แก้ไข แนะนำแนวทาง เพื่อให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอขอบพระคุณ อาจารย์ ดร.วสันต์ ทองไทย ประธานสาขาวิจัยและประเมินผลการศึกษา ที่ได้ให้คำแนะนำ ปรึกษา รศ. วิจัย พาณิชย์สวอย ประธานกรรมการสอบ ผศ.ดร.ชุติมา หทัยรัตนา ผู้ทรงคุณวุฒิภายนอก ที่ให้แนวคิดและข้อเสนอแนะแก่ผู้วิจัย ทำให้วิทยานิพนธ์มีความสมบูรณ์ยิ่งขึ้น ผู้วิจัยขอกราบขอบพระคุณทุกท่านไว้เป็นอย่างสูง

ขอขอบพระคุณผู้เชี่ยวชาญทั้ง 11 ท่าน คือ ดร.โชคชัย ยืนยง ดร.ปัฐมาภรณ์ พิมพ์ทอง ดร.ประสงค์ เมธิพิณิตกุล อาจารย์वासนา กิรติจำเริญ อาจารย์อรุณี ชูทัย อาจารย์วรรณา ลาวัณ อาจารย์อาชวีณี ไชยสุนทร อาจารย์นงพะงา เรืองฤทธิ์ชาญกุล อาจารย์พัชรินทร์ สุภสมบัติ อาจารย์สยามไท โสภณ และอาจารย์ปราชญา ปานกลีบ ที่ให้ความอนุเคราะห์ในการตรวจสอบเครื่องมือที่ใช้ในการวิจัยและให้คำแนะนำอันเป็นประโยชน์ในการทำวิจัยให้มีความสมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ ผู้บริหาร คณะครู โรงเรียนเศรษฐบุตรบำเพ็ญ ที่ให้ความอนุเคราะห์ในการตรวจสอบและวิเคราะห์ข้อมูลทางคอมพิวเตอร์

ขอขอบพระคุณ ผู้บริหาร คณะครู โรงเรียนสตรีเศรษฐบุตรบำเพ็ญ โรงเรียนนวมินทราชินูทิศ สตรีวิทยา 2 โรงเรียนนวมินทราชินูทิศ เบญจมราชาลัย โรงเรียนฤทธิยะวรรณาลัย 2 และโรงเรียนพรตพิทยพยัต ที่ให้ความร่วมมือในการเก็บข้อมูลการวิจัยเป็นอย่างดี

ขอขอบคุณคณาจารย์ทุกท่านที่ประสิทธิ์ประสาทความรู้ และให้คำปรึกษาแก่ผู้วิจัย ขอขอบคุณ นางเฉลียว มากสวาสดี และ นายวิทยา มากสวาสดี ที่ให้ความช่วยเหลือในการศึกษา และเป็นกำลังใจสำคัญในการทำวิจัยด้วยดีตลอดมา

ท้ายนี้ผู้ค้นคว้าอันพึงมีจากวิทยานิพนธ์ฉบับนี้ ขอมอบเป็นเครื่องบูชาพระคุณบิดา มารดา ครูอาจารย์และผู้มีพระคุณทุกท่าน

มนัสนันท์ มากสวาสดี

เมษายน 2553

สารบัญ

หน้า

สารบัญตาราง	(3)
สารบัญภาพ	(5)
บทที่ 1 บทนำ	1
ความสำคัญของปัญหา	1
วัตถุประสงค์ของการวิจัย	3
ขอบเขตของการวิจัย	4
ข้อตกลงเบื้องต้น	4
นิยามศัพท์	5
ประโยชน์ที่จะได้รับ	6
บทที่ 2 การตรวจเอกสาร	7
ทฤษฎีการตอบสนองข้อสอบ	7
การสอบแบบดั้งเดิมและการสอบแบบปรับเหมาะ	15
ประเภทของการสอบแบบปรับเหมาะ	18
วิธีการประมาณค่าความสามารถด้วยการสอบแบบปรับเหมาะ	
รูปพีระมิด	28
คุณภาพของข้อสอบและแบบสอบตามทฤษฎีการสอบแบบดั้งเดิม	31
งานวิจัยที่เกี่ยวข้อง	36
สมมติฐานงานวิจัย	41
บทที่ 3 วิธีดำเนินการวิจัย	42
พลวิจัย	42
เครื่องมือที่ใช้ในการวิจัย	42
การสร้างและพัฒนาเครื่องมือที่ใช้ในการวิจัย	43
การออกแบบการสอบแบบปรับเหมาะและการสอบแบบดั้งเดิม	47
การเก็บรวบรวมข้อมูล	52

สารบัญ (ต่อ)

	หน้า
การวิเคราะห์ข้อมูล	53
บทที่ 4 ผลการวิจัยและข้อวิจารณ์	55
ผลการวิจัย	55
ข้อวิจารณ์	61
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	65
สรุปผลการวิจัย	65
ข้อเสนอแนะ	68
เอกสารและสิ่งอ้างอิง	70
ภาคผนวก	73
ภาคผนวก ก รายชื่อผู้เชี่ยวชาญในการตรวจสอบเครื่องมือ	74
ภาคผนวก ข ผลการตรวจสอบความตรงตามเนื้อเรื่อง	77
ภาคผนวก ค ตัวอย่างการสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้น บูรณาการแบบปรับเหมาะ	83
ภาคผนวก ง แบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ	86
ประวัติการศึกษาและการทำงาน	113

สารบัญตาราง

ตารางที่		หน้า
1	ความแตกต่างระหว่างการสอบแบบดั้งเดิมและการสอบแบบปรับเหมาะ	17
2	ผลการตรวจสอบความตรงตามเนื้อเรื่องจากการพิจารณาโดยผู้เชี่ยวชาญ	55
3	ค่าสถิติของแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ 2 ชุด	56
4	ค่าพารามิเตอร์ของแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้น บูรณาการ จากข้อสอบจำนวน 76 ข้อ	57
5	ค่าความสัมพันธ์ระหว่างค่าประมาณความสามารถที่ได้จากการสอบ แบบปรับเหมาะรูปพีระมิดข้างตัดกับค่าประมาณความสามารถจาก การสอบแบบดั้งเดิม	60
ตารางผนวกที่		
1	ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะ กระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ (ทักษะการตั้งสมมติฐาน)	78
2	ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะ กระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ (ทักษะการกำหนดนิยาม เชิงปฏิบัติการ)	79

สารบัญตาราง (ต่อ)

ตารางผนวกที่		หน้า
3	ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะ กระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ (ทักษะการกำหนดและ ควบคุมตัวแปร)	80
4	ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะ กระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ (ทักษะการทดลอง)	81
5	ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะ กระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ (ทักษะการตีความหมายข้อมูล และลงข้อสรุป)	82

สารบัญภาพ

ภาพที่		หน้า
1	เส้นโค้งลักษณะข้อสอบของโมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์	11
2	แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดแบบขนาดขั้นคงที่	20
3	แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดแบบขนาด ขั้นแปรผัน	21
4	แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด แบบวิธีสกัดการสะท้อนกลับ	22
5	แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด แบบวิธีรักษาการสะท้อนกลับ	23
6	แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดแบบมีหลายข้อ ในแต่ละชั้น	24
7	แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดชนิดให้น้ำหนัก หลายตัวเลือกเพื่อแยกทาง	25
8	แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดวิธี สกัดการสะท้อนกลับ ชนิด 12 ชั้น	27
9	การประมาณค่าความสามารถตามค่าความยากของข้อสอบที่ต่อจากข้อสุดท้าย ที่ทำ	29

สารบัญญภาพ (ต่อ)

ภาพที่		หน้า
10	การประมาณค่าความสามารถตามค่าความยากของข้อสอบที่ต่อจากข้อสุดท้าย ที่ทำ	30
11	ลำดับขั้นตอนการสร้างและพัฒนาแบบสอบวัดทักษะกระบวนการ ทางวิทยาศาสตร์ชั้นบูรณาการ	43
12	แบบจำลองโครงสร้างแบบสอบปรับเหมาะรูปพีระมิดข้างตัดที่แสดงตาม ค่าความยาก	49
13	แบบจำลองโครงสร้างแบบสอบปรับเหมาะรูปพีระมิดข้างตัดตามทักษะ ที่ต้องการวัด	50
14	ผลการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดของพลวิจัยที่พิจารณา จากคะแนนรายข้อและค่าประมาณความสามารถทั้ง 3 วิธี	58
15	ผลการสอบแบบดั้งเดิมของพลวิจัยที่พิจารณาจากค่าประมาณความสามารถ จากโปรแกรม BILOG และคะแนนรายข้อ	59

บทที่ 1

บทนำ

ความสำคัญของปัญหา

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) หรือ ทฤษฎีลักษณะแฝง (Latent Trait Theory) เป็นทฤษฎีที่ใช้อธิบายความสัมพันธ์ระหว่างคุณลักษณะ (Latent) ของนักเรียน หรือความสามารถ (Ability) ของนักเรียน ซึ่งเป็นสิ่งที่สังเกตไม่ได้ กับ พฤติกรรมการตอบข้อสอบของนักเรียนซึ่งเป็นสิ่งที่สังเกตได้ โดยความสัมพันธ์นี้อธิบายได้ด้วยโมเดลทางคณิตศาสตร์ (Hambleton and Swaminathan, 1985: 9) ที่มีการนำค่าประจำข้อสอบแต่ละข้อหรือเรียกว่า ค่าพารามิเตอร์ของข้อสอบ (Item Parameter) ได้แก่ ค่าอำนาจจำแนก ค่าความยาก และค่าสัมประสิทธิ์การเดา มารวมกันประมาณค่าความสามารถ หรือเรียกว่า ค่าพารามิเตอร์ของนักเรียน (Examinee Parameter) ซึ่งอาจใช้ค่าพารามิเตอร์ของข้อสอบทุกค่าหรือบางค่าก็ได้ขึ้นอยู่กับโมเดลทางคณิตศาสตร์ที่เลือกใช้ในการวิเคราะห์ สำหรับค่าพารามิเตอร์ที่ได้จากทฤษฎีนี้จะมีค่าคงที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ (Hambleton and Swaminathan, 1985: 11) ทำให้สามารถประมาณค่าความสามารถที่แท้จริงและเปรียบเทียบความสามารถของนักเรียนได้ แม้ว่าจะทำแบบสอบคนละชุดก็ตาม ทฤษฎีนี้จึงมีคุณค่าในการพัฒนาแบบสอบ เพื่อใช้ในการสอบวัดความสามารถ ซึ่งจะทำให้การวัดผลมีประสิทธิภาพยิ่งขึ้น

ผู้ริเริ่มให้แนวคิดและหลักการเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ (IRT) คือ Ferguson และ Lawley ในปี ค.ศ.1942 และ ค.ศ.1943 โดยมีแนวคิดว่า ผลการสอบของนักเรียนแต่ละคนในการทำแบบสอบใด ๆ ก็ตามจะถูกหรือผิดมากน้อยเพียงใดขึ้นอยู่กับความสัมพันธ์เชิงคณิตศาสตร์ระหว่างคุณลักษณะแฝง (Latent Trait) หรือความสามารถ (Ability) ของนักเรียนกับคุณลักษณะของแบบสอบ หลังจากนั้นความสนใจเกี่ยวกับทฤษฎีการตอบสนองข้อสอบ (IRT) ได้หยุดลงชั่วคราว จนกระทั่งปลาย ค.ศ.1960 ทฤษฎีนี้จึงได้รับความสนใจอีกครั้ง เนื่องจากพบจุดอ่อนของทฤษฎีการสอบแบบดั้งเดิมที่ใช้กันมานาน เช่น ความลำเอียงของข้อสอบ ค่าพารามิเตอร์ที่เปลี่ยนแปลงตามกลุ่มผู้สอบ รวมทั้งมีการเริ่มต้นใช้การสอบแบบเทเลอร์ (Tailored Test) ในการวัดผล (Hambleton and Swaminathan, 1985: XV)

การใช้ข้อสอบจากคลังข้อสอบมาสร้างเป็นแบบสอบซึ่งแบบสอบแต่ละชุดออกแบบมาให้เหมาะสมกับความสามารถของนักเรียนแต่ละคน เรียกว่า การสอบแบบเทเลอร์ (Tailored Test) คำนี้ใช้ครั้งแรกโดย W.W. Turnbull (Lord, 1980 อ้างใน ศิริชัย กาญจนวาสิ, 2538: 2) ต่อมา D.J. Weiss ได้เสนอให้ใช้คำว่า การสอบแบบปรับเหมาะ (Adaptive Test) (Weiss, 1974 อ้างใน ศิริชัย กาญจนวาสิ, 2538: 2) ซึ่ง Weiss และ Betz ได้ชี้ว่าแบบสอบวัดเชาว์ปัญญารายบุคคลของ Binet เป็นแบบสอบที่ประยุกต์แนวคิดของการสอบที่ปรับให้เหมาะสมกับผู้สอบเป็นครั้งแรก เนื่องจากมีการจัดเรียงข้อสอบตามลำดับความยากของข้อสอบ โดยใช้อายุเป็นเกณฑ์สำหรับทดสอบเป็นรายบุคคล และกำหนดข้อสอบเริ่มต้นที่เหมาะสมกับอายุจริงของผู้สอบ ถ้าตอบถูกก็จะเลือกข้อที่เหมาะสมกับระดับอายุที่สูงขึ้นมาให้ตอบ แต่ถ้าตอบผิดก็จะเลือกข้อที่เหมาะสมกับระดับอายุที่ต่ำกว่ามาให้ตอบ

โดยทั่วไปการสอบแบบปรับเหมาะ (Adaptive Test) จะมีการปรับระดับความยากของข้อสอบในแต่ละข้อให้เหมาะกับระดับความสามารถของนักเรียน นักเรียนที่มีความสามารถสูงจะได้รับข้อสอบที่ยาก ส่วนนักเรียนที่มีความสามารถต่ำจะได้รับข้อสอบที่ง่าย หลังจากการตอบข้อสอบจึงประมาณค่าความสามารถของนักเรียนตามแบบแผนการตอบข้อสอบแต่ละข้อที่ได้รับต่างกัน เนื่องจากนักเรียนที่มีความสามารถสูงทำข้อสอบที่ยากกว่าถูกต้องย่อมได้คะแนนหรือได้ค่าประมาณความสามารถที่สูงกว่านักเรียนที่ทำข้อสอบข้อที่ง่ายถูก ดังนั้นจึงต้องประมาณค่าความสามารถของนักเรียนตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ซึ่งเป็นทฤษฎีที่สามารถอธิบายความสัมพันธ์ระหว่างระดับความสามารถของนักเรียนกับผลการสอบ (ศิริชัย กาญจนวาสิ, 2550: 53) การสอบแบบปรับเหมาะจะใช้ข้อสอบจากคลังข้อสอบซึ่งมีการรายงานค่าพารามิเตอร์ของข้อสอบแต่ละข้อ เพื่อนำมาใช้ในการพิจารณาคัดเลือกข้อสอบที่มีค่าความยากเหมาะสมกับระดับความสามารถของนักเรียน และใช้จำนวนข้อสอบน้อยกว่าการสอบแบบดั้งเดิม กล่าวคือนักเรียนที่มีความสามารถต่ำไม่ต้องทำข้อที่ยาก และนักเรียนที่มีความสามารถสูงไม่ต้องทำข้อที่ง่าย การสอบแบบนี้ก็สามารถประมาณค่าความสามารถได้อย่างแม่นยำ

ในวงการศึกษาก็ได้มีการนำคอมพิวเตอร์มาใช้ในการเรียนการสอนอย่างแพร่หลาย ถ้าหากโรงเรียนได้พัฒนาคอมพิวเตอร์ให้เป็นคลังข้อสอบขนาดใหญ่อย่างเป็นระบบแล้ว จึงควรมีการวัดผลด้วยคอมพิวเตอร์ที่อาศัยหลักการสอบแบบปรับเหมาะ ซึ่งจะช่วยให้เกิดประโยชน์ต่อนักเรียนและครูในการแปลความหมายของผลการวัดที่ได้ แต่เนื่องจากการสอบแบบปรับเหมาะมีหลายรูปแบบ จึงต้องพิจารณาวิธีการประมาณค่าความสามารถที่เหมาะสมกับรูปแบบของการสอบ ซึ่ง

การสอบแบบปรับเหมาะรูปพีระมิดเป็นรูปแบบแรกของการสอบแบบปรับเหมาะที่มีการจัดวางลำดับของข้อสอบเป็นรูปพีระมิด ในการสอบนักเรียนจะตอบข้อสอบโดยเริ่มตอบข้อที่อยู่บนยอดพีระมิดซึ่งมีความยากปานกลางก่อน ถ้าตอบถูกขั้นต่อไปจะแยกไปตอบข้อที่ยากขึ้น แต่ถ้าตอบผิดขั้นต่อไปจะตอบข้อที่ง่ายกว่า กระบวนการสอบจะเป็นเช่นนี้ไปจนถึงขั้นสุดท้ายซึ่งเป็นฐานของพีระมิด (สุพัฒน์ สุกมลสันต์, 2539: 43) ในการสอบรูปแบบนี้มีผู้เสนอวิธีการประมาณค่าความสามารถ 6 วิธี (Weiss, 1974: 30-33) คือ (1) ประมาณค่าความสามารถตามค่าความยากเฉลี่ยของข้อที่ทำถูก (2) ประมาณค่าความสามารถตามค่าความยากเฉลี่ยของข้อที่ได้ทำ (3) ประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (4) ประมาณค่าความสามารถตามค่าความยากของข้อสอบที่ถัดจากข้อสุดท้ายที่ได้ตอบ (5) ประมาณค่าความสามารถตามผลการตอบข้อสอบข้อสุดท้าย และ (6) ประมาณค่าความสามารถจากข้อสอบทุกข้อตามวิธีของแฮนทซ์เซน แต่ละวิธีมีข้อจำกัดแตกต่างกันในเรื่องของจำนวนขั้นที่ใช้ในการสอบ ความไวในการเดาคำตอบ ความเที่ยงและค่าสหสัมพันธ์ระหว่างการสอบแบบปรับเหมาะกับรูปแบบอื่น การประมาณค่าความสามารถโดยวิธีต่าง ๆ นั้น อาจใช้ได้ดีในสถานการณ์ที่ต่างกัน ดังนั้นการจะกล่าว่วิธีใดเหมาะสมกับสถานการณ์ใดจึงต้องอาศัยผลจากการวิจัย

วิธีการประมาณค่าความสามารถจากการสอบแบบปรับเหมาะรูปพีระมิดดังกล่าวที่ไม่ซับซ้อนเหมาะให้ครูนำไปใช้ประกอบการสอบ 3 วิธี คือ (1) วิธีประมาณค่าความสามารถตามค่าความยากเฉลี่ยของข้อที่ทำถูก (2) วิธีประมาณค่าความสามารถตามค่าความยากเฉลี่ยของข้อที่ได้ทำ และ (3) วิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ ซึ่งแต่ละวิธีมีข้อจำกัดแตกต่างกัน ผู้วิจัยจึงสนใจศึกษาเพื่ออธิบายว่าวิธีทั้ง 3 มีความตรงในการอธิบายความสามารถที่แท้จริงให้มีความน่าเชื่อถือมากน้อยเพียงใด ถ้ามีความน่าเชื่อถือมากจะนำเสนอให้ครูนำทฤษฎีการตอบสนองข้อสอบไปประยุกต์ใช้กับการสอบแบบปรับเหมาะ เพื่อใช้อธิบายความสามารถของนักเรียน

วัตถุประสงค์ของการวิจัย

เพื่อศึกษาความตรงสัมพัทธ์ของค่าประมาณความสามารถ 3 วิธี จากการสอบแบบปรับเหมาะรูปพีระมิดข้างต้น

ขอบเขตของการวิจัย

1. แบบสอบวัดความสามารถที่ใช้ในการวิจัยครั้งนี้เป็นแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ซึ่งประกอบด้วยทักษะการตั้งสมมติฐาน ทักษะการกำหนดนิยามเชิงปฏิบัติการ ทักษะการกำหนดและควบคุมตัวแปร ทักษะการทดลอง และทักษะการตีความหมายข้อมูลและลงข้อสรุป ที่ผู้วิจัยสร้างขึ้นเป็นแบบปรนัย 4 ตัวเลือก มีจำนวนทั้งหมด 54 ข้อ

2. การออกแบบการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด ชนิด 12 ชั้น ได้นำข้อคำถามจากการสอบแบบดั้งเดิมมาเสนอผ่านจอกอมพิวเตอร์ ซึ่งมีการกำหนดเส้นทางการตอบเป็นรูปพีระมิด จำนวน 12 ชั้น โดยเริ่มตอบข้อที่ 1 ซึ่งมีค่าความยากปานกลาง และในชั้นที่ 1 ถึง 5 มีจำนวนข้อเท่ากับจำนวนชั้น และในชั้นที่ 6 ถึง 12 ได้ตัดข้อที่มีค่าความยากสูง ๆ และต่ำ ๆ ออก

3. พลวิจัย เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนแผนวิทยาศาสตร์-คณิตศาสตร์ ภาคเรียนที่ 2 ปีการศึกษา 2552 โรงเรียนเศรษฐบุตรบำเพ็ญ สังกัดสำนักงานการศึกษาขั้นพื้นฐาน กรุงเทพมหานครเขต 2 จำนวน 127 คน

4. วิธีการประมาณค่าความสามารถที่นำมาหาความตรงสัมพัทธ์ คือ

4.1 วิธีการประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก

4.2 วิธีการประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (ทั้งข้อที่ตอบถูกและตอบผิด)

4.3 วิธีการประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (ทั้งข้อที่ตอบถูกและตอบผิด)

ข้อตกลงเบื้องต้น

การหาค่าประมาณความสามารถจากการสอบแบบดั้งเดิมและค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) ได้ใช้โปรแกรม BILOG ตามวิธีการของเบล์

นิยามศัพท์

ความตรงสัมพัทธ์ของค่าประมาณความสามารถ หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ เพียร์สันระหว่างค่าประมาณความสามารถจากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดโดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบกับค่าประมาณความสามารถจากการสอบแบบดั้งเดิมซึ่งใช้เป็นเกณฑ์อ้างอิง

ค่าประมาณความสามารถจากการสอบแบบดั้งเดิม (θ_T) หมายถึง ระดับความสามารถของนักเรียนที่ได้จากการตอบคำถามโดยวิธีเขียนตอบ ซึ่งได้จากการนำค่าพารามิเตอร์ของข้อสอบแต่ละข้อที่นักเรียนได้ตอบมาคำนวณตามทฤษฎีการตอบสนองข้อสอบด้วยโปรแกรม BILOG

ค่าประมาณความสามารถจากการสอบแบบปรับเหมาะโดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) หมายถึง ระดับความสามารถของนักเรียนที่ได้จากการตอบคำถามผ่านจอคอมพิวเตอร์ ซึ่งคำนวณจากสัดส่วนระหว่างผลรวมของค่าความยากของข้อที่นักเรียนตอบถูกกับจำนวนข้อที่นักเรียนตอบถูก

ค่าประมาณความสามารถจากการสอบแบบปรับเหมาะโดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) หมายถึง ระดับความสามารถของนักเรียนที่ได้จากการตอบคำถามผ่านจอคอมพิวเตอร์ ซึ่งคำนวณจากสัดส่วนระหว่างผลรวมของค่าความยากของข้อที่นักเรียนได้ตอบกับจำนวนข้อที่นักเรียนได้ตอบ

ค่าประมาณความสามารถจากการสอบแบบปรับเหมาะโดยวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) หมายถึง ระดับความสามารถของนักเรียนที่ได้จากการตอบคำถามผ่านจอคอมพิวเตอร์ซึ่งได้จากค่าความยากของข้อที่นักเรียนได้ตอบเป็นข้อสุดท้าย

ประโยชน์ที่จะได้รับ

ครู นักวัดผล และผู้ที่สนใจได้แนวทางในการเลือกวิธีวัดและประเมินผลที่เหมาะสมกับระดับ
ความสามารถของนักเรียน



บทที่ 2

การตรวจเอกสาร

ในการวิจัยครั้งนี้ผู้วิจัยได้ศึกษาค้นคว้าทฤษฎี เอกสาร และงานวิจัยที่เกี่ยวข้องมีรายละเอียดตามลำดับต่อไปนี้

1. ทฤษฎีการตอบสนองข้อสอบ
2. การสอบแบบดั้งเดิมและการสอบแบบปรับเหมาะ
3. ประเภทของการสอบแบบปรับเหมาะ
4. วิธีการประมาณค่าความสามารถด้วยการสอบแบบปรับเหมาะรูปพีระมิด
5. คุณภาพของข้อสอบและแบบสอบตามทฤษฎีการสอบแบบดั้งเดิม
6. งานวิจัยที่เกี่ยวข้อง

ทฤษฎีการตอบสนองข้อสอบ

ความเป็นมาของทฤษฎีการตอบสนองข้อสอบ

ผู้ริเริ่มนำทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) มาใช้ในวงการวัดและประเมินผลการศึกษา คือ Ferguson และ Lawley ในปี ค.ศ.1942 และ ค.ศ.1943 ซึ่งอาศัยหลักการที่ว่าผลการสอบของแต่ละคนในการทำแบบสอบใด ๆ ขึ้นอยู่กับลักษณะแฝง (Latent Trait) ของบุคคลหรือความสามารถ (Ability) ของบุคคล โดยเสนอบทความเกี่ยวกับปัญหาการสร้างและการเลือกข้อสอบซึ่งแสดงแนวคิดในรูปของโมเดลนอร์มอลโอจีฟ (Normal Ogive Model) ที่เป็นโมเดลหนึ่งในทฤษฎีการตอบสนองข้อสอบ (IRT) ต่อมาในปี ในปี ค.ศ.1950 ราสช์ (Rasch) นักคณิตศาสตร์ชาวเดนมาร์กได้เสนอโมเดลราสช์ (Rasch Model หรือ 1-Parameter Logistic Model) แบบ 1 พารามิเตอร์ โดยมีแนวคิดว่าความยากของข้อสอบเป็นสิ่งเดียวที่มีอิทธิพลต่อการตอบสนองข้อสอบ และในปี ค.ศ.1952 Lord ได้พิมพ์ผลงานวิจัยเพื่อทำปริญาเอก ซึ่งเขาได้แสดงว่าทฤษฎีการตอบสนองข้อสอบ (IRT) เป็น โมเดล หรือ เป็นทฤษฎีอย่างชัดเจน และเรียกว่า ทฤษฎีโค้งคุณลักษณะของข้อสอบ (ICC) ดังนั้น Lord จึงได้รับการยกย่องว่าเป็นบิดา หรือ ผู้พบทฤษฎีการตอบสนองข้อสอบ (IRT) รวมทั้งได้ขยายแนวคิดทฤษฎีการตอบสนองข้อสอบ โดย

เสนอ Normal Ogive Function แบบ 2 พารามิเตอร์ที่มีการเพิ่มพารามิเตอร์อำนาจจำแนก ในปี ค.ศ.1968 เบิร์นบอม (Birbaum) ซึ่งเป็นนักสถิติได้พัฒนาโมเดลนอร์มอลโอจิว (Normal Ogive Model) ให้เป็นโมเดลโลจิสติก (Logistic Model) ซึ่งทำให้เป็นฟังก์ชันที่สามารถคำนวณได้ง่ายขึ้น ในปี ค.ศ.1974 ลอร์ด (Lord) ได้เสนอแนวคิดโมเดลโลจิสติกแบบ 3 พารามิเตอร์ โดยเพิ่มพารามิเตอร์โอกาสในการเดาข้อสอบ ในปี ค.ศ.1970 Urry ได้เปลี่ยนชื่อ ทฤษฎีโค้งลักษณะของข้อสอบ (Item Characteristic Curve Theory) เป็น ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) (Hambleton and Swaminatan, 1985: 8)

แนวคิดของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างลักษณะแฝง (Latent Trait) ของบุคคล หรือความสามารถ (Ability) ของบุคคลกับคำตอบของบุคคลนั้นภายใต้คุณลักษณะของข้อสอบนั้น ซึ่งอธิบายได้ด้วยค่าประจำข้อสอบแต่ละข้อ หรือค่าพารามิเตอร์ของข้อสอบ ได้แก่ อำนาจจำแนก (a) ความยาก (b) และโอกาสในการเดา (c) ทั้งนี้ การสอบส่วนใหญ่ต้องการสรุปเกี่ยวกับระดับความสามารถหรือทักษะของบุคคล โดยการจะสรุปได้ต้องทราบว่าความสามารถของบุคคลจะเป็นตัวกำหนดคำตอบแต่ละข้อให้มีลักษณะอย่างไร ซึ่งทฤษฎีนี้ได้อธิบายในรูปของฟังก์ชันทางคณิตศาสตร์ที่แสดงว่าระดับความสามารถกับคำตอบของบุคคลนั้นสัมพันธ์กันอย่างไร ความสัมพันธ์นี้ได้มาจากฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) หรือเรียกว่า ฟังก์ชันลักษณะข้อสอบ (Item Characteristic Function) (Hambleton and Swaminathan, 1984: 25) เมื่อนำความสัมพันธ์นี้มาเขียนกราฟจะได้โค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) ซึ่งฟังก์ชันลักษณะข้อสอบมีหลายลักษณะขึ้นอยู่กับรูปแบบการวิเคราะห์ที่กำหนดและลักษณะของข้อสอบที่เกี่ยวข้อง เช่น โมเดลของ Rasch เรียกว่า 1-Parameter Logistic Model โมเดลของ Lord เรียกว่า 3-Parameter Logistic Model

โมเดลของทฤษฎีการตอบสนองข้อสอบ

โมเดลทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างผลการตอบข้อสอบกับความสามารถที่อยู่ภายใน (θ) ซึ่งเป็นคุณลักษณะ (Trait) หรือ ความสามารถ (Ability) เช่น ความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์ที่กำลังจะถูกวัด มีหลายโมเดลด้วยกัน แต่โมเดลที่นิยมใช้กันคือ โมเดลโลจิสติก (Logistic Model) 1 2 และ 3 พารามิเตอร์ (ศิริชัย กาญจนวาสี, 2550: 54)

ในที่นี้ขอนำเสนอโมเดลที่เกี่ยวข้องกับ ค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) และ ค่าสัมประสิทธิ์การเดา (c) คือ โมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์

สำหรับข้อกระทงที่เป็นแบบ Dichotomous คือ การให้คะแนนแบบตอบถูกให้ 1 คะแนน ตอบผิดให้ 0 คะแนน ฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) สามารถแสดงได้ด้วยความน่าจะเป็น P หรือ $P(\theta)$ ของคำตอบข้อนั้นที่ถูกต้อง ซึ่งน่าจะเป็นว่า โอกาสในการตอบถูก ($P(\theta)$) จะเพิ่มมากขึ้น เมื่อความสามารถของนักเรียนสูงขึ้น จากข้อตกลงนี้ความน่าจะเป็น $P(\theta)$ ดังกล่าวสามารถแสดงได้โดย ฟังก์ชันโลจิสต์ 3 พารามิเตอร์ (Three Parameter Logistic Function) ดังนี้ (Hambleton and Swaminathan, 1991: 17)

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad i = 1, 2, 3, \dots, n$$

เมื่อ $P_i(\theta)$ คือ โอกาสที่นักเรียนซึ่งมีความสามารถในระดับ θ จะทำข้อสอบข้อที่ i ได้ถูกต้อง

θ คือ ระดับความสามารถที่แท้จริงของนักเรียน

a คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b คือ ค่าความยากของข้อสอบข้อที่ i

c คือ ค่าสัมประสิทธิ์การเดาของข้อสอบข้อที่ i

e คือ ค่าคงตัวทางคณิตศาสตร์ มีค่า = 2.71828

D คือ ตัวประกอบของสเกล มีค่าเท่ากับ 1.7

ในการวิจัยครั้งนี้ได้นำโมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์ของทฤษฎีการตอบสนองข้อสอบ (IRT) มาใช้ เนื่องจากเป็นโมเดลที่ให้ค่าพารามิเตอร์ของข้อสอบที่เหมาะสมในการนำมาพิจารณาการคัดเลือกข้อสอบสำหรับการสอบแบบปรับเหมาะ

ค่าพารามิเตอร์ของทฤษฎีการตอบสนองข้อสอบ

ค่าพารามิเตอร์ในการตอบข้อคำถามแบ่งออกเป็น 2 ชนิด คือ ค่าพารามิเตอร์ของข้อสอบ (Item Parameter) ได้แก่ ค่าอำนาจจำแนก (a) ค่าความยาก (b) ค่าสัมประสิทธิ์การเดา (c) และค่าพารามิเตอร์ของนักเรียน (Examinee Parameter) ได้แก่ ระดับความสามารถของนักเรียน (θ) ซึ่งค่าพารามิเตอร์ของโมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์ มีรายละเอียด ดังนี้ (Warm, 1978: 52-53 อ้างใน วนิตา มุขดาร์, 2544: 21-22)

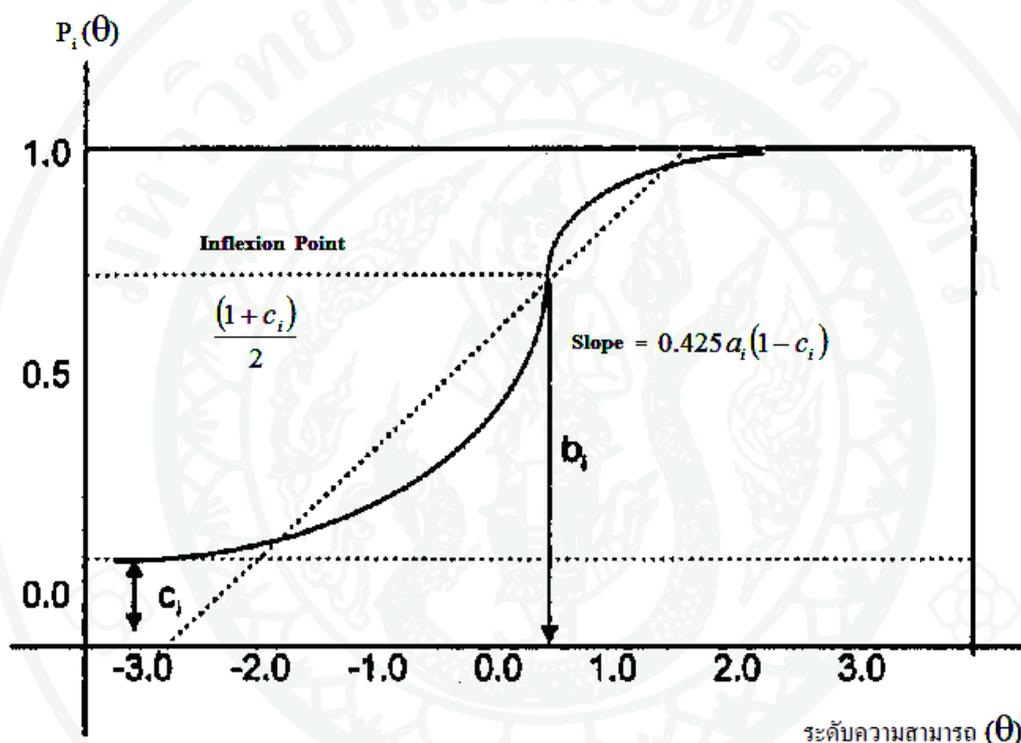
1. ค่าความยากของข้อสอบ (b – Parameter หรือ Item Difficulty) แทนด้วย b_i เป็นค่าของระดับความสามารถ θ ระดับหนึ่งที่ทำให้เส้นโค้งลักษณะข้อสอบมีการเปลี่ยนโค้ง และเมื่อไม่มีการเดา ($c_i = 0$) นักเรียนที่มี $\theta = b_i$ มีโอกาสทำข้อสอบนั้น ๆ ได้ถูกต้อง เท่ากับ 0.5 แต่ถ้ามีการเดา ($c_i \neq 0$) โอกาสของการทำข้อสอบได้ถูกต้อง เท่ากับ $\frac{(1+c_i)}{2}$ ในทางทฤษฎีมีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติมีค่าตั้งแต่ -3.00 ถึง $+3.00$ โดยค่าเป็นลบแสดงว่าข้อสอบง่ายและค่าเป็นบวกแสดงว่าข้อสอบยาก

2. ค่าอำนาจจำแนกของข้อสอบ (a- Parameter หรือ Discriminating Power) แทนด้วย a_i มีค่าเป็นสัดส่วนโดยตรงกับความชัน (Slope) ของเส้นโค้งลักษณะข้อสอบ ณ จุดเปลี่ยนโค้ง (Inflexion Point) หรือ จุดบนเส้น ICC ณ ที่โค้งมีความชัน (Slope) มากที่สุด ซึ่งความชัน ณ จุดเปลี่ยนโค้ง เท่ากับ $0.425a_i(1-c_i)$ จุดนี้เป็นจุดที่แสดงความสามารถในการจำแนกนักเรียนที่มีความสามารถต่ำและสูงออกจากกันได้ชัดเจน ถ้าค่าอำนาจจำแนกของข้อสอบมีค่ามาก โค้งลักษณะข้อสอบ (ICC) จะยังมีความชันมาก ในทางทฤษฎี ค่า a_i ที่เป็นไปได้อยู่ระหว่าง 0 ถึง $+\infty$ แต่ในทางปฏิบัติ ค่า a_i ของข้อสอบโดยทั่วไปมีค่าตั้งแต่ 0 ถึง $+2.00$ ถ้าค่าเป็น 0 แสดงว่าข้อสอบนั้นไม่มีค่าอำนาจจำแนกหรือมีค่าอำนาจจำแนกต่ำ แต่ถ้ามีค่า $+2.00$ แสดงว่าข้อสอบนั้นมีค่าอำนาจจำแนกสูง

3. ค่าสัมประสิทธิ์การเดา (c-Parameter หรือ Pseudo-Guessing Parameter หรือ Pseudo-Chance Score Level) แทนด้วย c_i เป็นค่ากำกับเส้นโค้งที่ต่ำที่สุด (Lower Assymtote) ของโค้งลักษณะข้อสอบที่แสดงถึงค่าโอกาสที่นักเรียนที่มีความสามารถต่ำมาก ($\theta = -\infty$) จะทำข้อสอบถูกปกติค่า c_i มีค่าตั้งแต่ 0.00 ถึง 0.40 ข้อสอบที่มีค่า c_i ตั้งแต่ 0.3 ขึ้นไปเป็นข้อสอบที่ไม่ดี

โดยทั่วไปแล้วค่า c_i จะมีค่าประมาณ $\frac{1}{A} - 0.05$ เมื่อ A เป็นจำนวนตัวเลือกในข้อสอบ ดังนั้น ข้อสอบที่มี 4 ตัวเลือก จะมีค่า c_i ประมาณ 0.20

4. ค่าความสามารถของนักเรียน (Examinee Ability) หมายถึง ระดับความสามารถของนักเรียน (θ) ที่คำนวณจากคะแนนจริง ในทางทฤษฎีมีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ผลการวิเคราะห์ส่วนใหญ่ให้ค่าตั้งแต่ -3.00 ถึง $+3.00$ โดยค่าที่เป็นลบแสดงว่ามีความสามารถต่ำ และค่าที่เป็นบวกแสดงว่ามีความสามารถสูง



ภาพที่ 1 เส้นโค้งลักษณะข้อสอบของโมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์

เนื่องจากการสอบแบบปรับเหมาะเป็นการสอบที่ต้องปรับระดับความยากของข้อสอบให้เหมาะกับระดับความสามารถของนักเรียน ค่าความยากจึงเป็นค่าที่ผู้สอนต้องนำมาพิจารณาในการคัดเลือกข้อสอบให้ตรงกับความสามารถของนักเรียน โดยในการวิจัยครั้งนี้เลือกข้อสอบที่มีค่าความยาก -3.00 ถึง $+3.00$ ซึ่งเป็นค่าที่ใช้ในทางปฏิบัติและสามารถตรวจสอบนักเรียนที่มีระดับความสามารถแตกต่างกันได้ และเลือกข้อสอบที่มีค่าอำนาจจำแนกสูง ๆ สำหรับค่าสัมประสิทธิ์การเดาเลือกใช้ข้อสอบที่มีค่านี้นี้ต่ำกว่า 0.30 เพื่อป้องกันไม่ให้นักเรียนสามารถเดาคำตอบที่ถูกต้องได้

นอกจากนี้แล้วการเลือกข้อสอบที่มีค่าอำนาจจำแนกสูงและค่าสัมประสิทธิ์การเดาคำจะทำให้สามารถจำแนกคนที่มีความสามารถสูงออกจากคนที่มีความสามารถต่ำ ส่งผลให้ค่าประมาณความสามารถของนักเรียนที่ได้เป็นค่าความสามารถที่แท้จริงไม่มีผลมาจากข้อสอบที่ขาดคุณสมบัติในการจำแนกระดับความสามารถของนักเรียนหรือเป็นข้อสอบที่สามารถเดาคำตอบได้

วิธีการประมาณค่าพารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ

การประมาณค่าพารามิเตอร์ต่าง ๆ ตามทฤษฎีการตอบสนองข้อสอบ (IRT) มีด้วยกันหลายวิธี แต่วิธีที่น่าสนใจ ได้แก่ วิธีแมกซิมัมไลค์ลิฮูด (Maximum Likelihood Estimation) และวิธีของเบย์ (Bayesian Estimation) การเลือกใช้วิธีการประมาณค่าเป็นสิ่งสำคัญในการนำทฤษฎีการตอบสนองข้อสอบ (IRT) ไปประยุกต์ใช้ เพราะวิธีการประมาณค่าพารามิเตอร์ที่มีประสิทธิภาพจะทำให้สามารถสรุปอ้างอิงความสามารถของนักเรียนได้อย่างถูกต้อง แต่ถ้าวิธีการประมาณค่าพารามิเตอร์ขาดประสิทธิภาพก็จะทำให้การสรุปอ้างอิงขาดความแม่นยำ

วิธีแมกซิมัมไลค์ลิฮูด (Maximum Likelihood Estimation) เป็นวิธีการทางสถิติที่นำมาใช้ประมาณค่าพารามิเตอร์ ในกรณีที่ประสิทธิภาพและลักษณะของการกระจายค่าความสามารถเข้าใกล้โค้งปกติ (Hambleton and Swaminathan, 1985: 83) แต่วิธีนี้มีปัญหาบางประการ สรุปได้ดังนี้

1. ไม่สามารถประมาณค่าพารามิเตอร์ของข้อสอบที่นักเรียนทุกคนตอบถูกหรือตอบผิด
2. การประมาณค่าอำนาจจำแนก (a) ของข้อสอบอาจจะได้ค่าประมาณที่มีค่าไม่คงที่ ซึ่งจะมีผลทำให้เกิดความคลาดเคลื่อนสูงมากในการประมาณค่าความสามารถของนักเรียน
3. ไม่สามารถประมาณค่าความสามารถของนักเรียนที่ตอบถูกหมดทุกข้อหรือตอบผิดหมดทุกข้อ
4. บางครั้งค่าพารามิเตอร์ หรือค่าที่ได้จากการประมาณไม่ตกอยู่ในขอบเขตของพารามิเตอร์ กล่าวคือ อาจมีค่าใดค่าหนึ่งอยู่นอกขอบเขตที่ยอมรับได้ ในกรณีเช่นนี้ต้องมีการกำหนดขอบเขตจำกัดของค่าประมาณไว้ เพื่อให้ค่าประมาณที่ได้ไม่สูงหรือต่ำเกินไป แต่การ

ปฏิบัตินี้เป็นจุดอ่อนของกระบวนการประมาณค่าด้วยวิธีแมกซิมัมไลค์ลิสต์ โดยเฉพาะในโมเดล 2 และ 3 พารามิเตอร์ ซึ่งทำให้เกิดปัญหาเกี่ยวกับความตรง (Validity) ของค่าประมาณที่ได้

5. ค่าพารามิเตอร์ที่ได้จะมีค่าใกล้เคียงค่าจริงเมื่อนักเรียนที่เข้าสอบและข้อสอบที่ใช้มีจำนวนมาก แต่ถ้านักเรียนที่เข้าสอบและจำนวนข้อสอบมีจำนวนน้อย ประสิทธิภาพของการประมาณค่าจะมีความแปรปรวนน้อยและลักษณะของการกระจายจะเข้าใกล้โค้งปกติ เมื่อประมาณค่าความสามารถของนักเรียน หรือ ค่าพารามิเตอร์เพียงอย่างเดียวอย่างหนึ่งเท่านั้น Swaminathan และ Gifford (1985: 349)

สำหรับวิธีของเบส์เป็นวิธีประมาณค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) โดย Swaminathan และ Gifford (1985: 349-360) ได้พัฒนาขึ้นเพื่อแก้ปัญหาเกี่ยวกับวิธีประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัมไลค์ลิสต์ดังกล่าว

ในการวิจัยครั้งนี้ได้ประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เนื่องจากในการสร้างแบบสอบปรับเหมาะได้คำนึงถึงค่าพารามิเตอร์ทั้ง 3 ค่า จึงใช้โมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) ดังนั้นจึงไม่เหมาะสมในการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัมไลค์ลิสต์ เนื่องจากการใช้วิธีนี้ต้องมีนักเรียนที่เข้าสอบและข้อสอบจำนวนมาก รวมทั้งอาจเกิดปัญหาเกี่ยวกับความตรงของค่าประมาณที่ได้

โปรแกรมคอมพิวเตอร์สำหรับประมาณค่าพารามิเตอร์

ในอดีตมีโปรแกรมคอมพิวเตอร์ที่ใช้ในการประมาณค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) ไม่เป็นที่แพร่หลาย จนกระทั่งในช่วงปี ค.ศ.1970 โปรแกรม BICAL และ LOGIST เป็นที่รู้จักและใช้กันอย่างกว้างขวาง และมีการพัฒนาโปรแกรมอื่น ๆ อีก ในปี ค.ศ. 1980 มีโปรแกรมการประมาณค่าใหม่ ๆ มากมาย ซึ่ง BILOG เป็นโปรแกรมที่มีชื่อเสียงโดดเด่นและเหมาะสมในการประมาณค่าพารามิเตอร์แบบ 1, 2 และ 3 พารามิเตอร์ ซึ่งโปรแกรม BILOG นี้สามารถใช้ได้กับคอมพิวเตอร์ขนาดใหญ่และคอมพิวเตอร์ขนาดเล็ก ในขณะที่โปรแกรมอื่นใช้ได้เฉพาะคอมพิวเตอร์ขนาดเล็ก (Hambleton and Swaminathan, 1991: 46 - 47) นอกจากนี้แล้ว Liane N. Patsula และ Marc E. Gessaroli (1995) ได้เปรียบเทียบระหว่างโปรแกรม LOGIST กับ โปรแกรม BILOG ซึ่งเป็นโปรแกรมที่ใช้กันอย่างแพร่หลายว่า ในการประมาณค่าพารามิเตอร์โมเดล 3 พารามิเตอร์ให้ได้ค่าพารามิเตอร์ที่คงที่ โปรแกรม LOGIST จะต้องใช้กลุ่มตัวอย่างขนาด

ใหญ่ และจำนวนข้อสอบมาก (กลุ่มตัวอย่าง 1,000 คน หรือมากกว่า และข้อสอบ 40 ข้อ หรือมากกว่า) ในขณะที่โปรแกรม BILOG ซึ่งเป็นโปรแกรมประมาณค่าพารามิเตอร์ที่ใหม่กว่าและโดยทั่วไปต้องการขนาดกลุ่มตัวอย่างและจำนวนข้อสอบที่น้อยกว่าโปรแกรม LOGIST เช่นเดียวกับข้อเสนอแนะของ Hulin, Lissak และ Gragrow (1982 อ้างใน อุทุมพร จามรمان, 2537: 164) ที่ว่าในการประมาณค่าพารามิเตอร์โมเดล 3 พารามิเตอร์ ควรมีขนาดกลุ่มตัวอย่างอย่างน้อย 500 คน และ จำนวนข้อสอบอย่างน้อย 30 ข้อ ด้วยเหตุนี้โปรแกรม BILOG จึงถูกใช้เป็นมาตรฐานในปัจจุบัน

ในการวิจัยครั้งนี้ได้เลือกใช้โปรแกรม BILOG เนื่องจากใช้ขนาดกลุ่มตัวอย่างและความยาวของแบบสอบไม่มากก็สามารถประมาณค่าพารามิเตอร์โมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) ได้อย่างแม่นยำ

ลักษณะเด่นของทฤษฎีการตอบสนองข้อสอบ

จากการศึกษาทฤษฎีการตอบสนองข้อสอบ (IRT) พบว่าทฤษฎีนี้จะมีลักษณะเด่นกว่าทฤษฎีการสอบแบบดั้งเดิม ดังนี้

1. ความเป็นอิสระจากกลุ่มนักเรียน (Sample Free) กล่าวคือ ไม่ว่าจะนำข้อสอบไปสอบกับนักเรียนกลุ่มใดก็ยังคงลักษณะข้อสอบก็เหมือนเดิม

2. ความไม่แปรเปลี่ยนของพารามิเตอร์ (Invariant of Item Parameter) หมายถึง การประมาณค่าพารามิเตอร์ เช่น ระดับความยาก อำนาจจำแนก และโอกาสในการเดาข้อสอบจากนักเรียนที่มีความสามารถระดับใดหรือกลุ่มใดก็ตามจะให้ค่าพารามิเตอร์ที่คงเดิม กล่าวคือไม่ว่ากลุ่มนักเรียนที่เข้าสอบจะมีระดับความสามารถสูงหรือต่ำ จะอยู่ในภาคเหนือหรือภาคใต้ก็จะให้ค่าพารามิเตอร์ที่ไม่ต่างกัน

3. ความเป็นอิสระของข้อสอบ (Item Free) ในการประมาณค่าความสามารถของนักเรียนจะใช้ข้อสอบชุดใดก็ได้จำนวนเท่าใดก็ได้ ซึ่งบางครั้งจะใช้เพียง 3-5 ข้อ ก็สามารถประมาณค่าความสามารถที่แท้จริงของนักเรียนได้ โดยข้อสอบต้องคัดเลือกจากคลังข้อสอบขนาดใหญ่ และแต่ละข้อระบุคุณลักษณะที่ต้องการวัดได้ จึงทำให้ความสามารถของนักเรียนไม่เปลี่ยนแปลงไปตามลักษณะการสอบ ความสามารถของนักเรียนจึงคงที่ไม่ว่าจะวัดเมื่อใดหรือวัดอย่างไร

จากลักษณะเด่นของทฤษฎีการตอบสนองข้อสอบ (IRT) สามารถนำมาประยุกต์ใช้ในการสอบแบบปรับเหมาะที่มีการปรับระดับความยากของข้อสอบตามระดับความสามารถของนักเรียน และค่าความยากของข้อสอบซึ่งเป็นค่าพารามิเตอร์ชนิดหนึ่ง ได้จากการวิเคราะห์ตามทฤษฎีนี้จะเป็นคุณลักษณะเฉพาะของข้อสอบข้อนั้น โดยไม่ขึ้นกับกลุ่มนักเรียนที่เข้าสอบ ทำให้สามารถประมาณค่าความสามารถที่แท้จริงของนักเรียนจากการทำข้อสอบได้ แม้ว่าจะทำข้อสอบเพียงไม่กี่ข้อก็ตาม รวมทั้งสามารถสร้างคลังข้อสอบที่มีการรายงานค่าพารามิเตอร์ของข้อสอบแต่ละข้อที่ได้จากการวิเคราะห์ตามทฤษฎีนี้

การสอบแบบดั้งเดิมและการสอบแบบปรับเหมาะ

การสอบแบบดั้งเดิมเป็นวิธีการสอบที่กำหนดให้นักเรียนทุกคนได้รับแบบสอบชุดเดียวกัน จำนวนข้อสอบเท่ากัน ซึ่งมีข้อจำนวนข้อสอบครอบคลุมองค์ประกอบของคุณลักษณะหรือเนื้อหาที่ต้องการวัดและครอบคลุมความสามารถของนักเรียนทั้งกลุ่ม ดังนั้นข้อสอบจึงมีค่าความยากของข้อสอบกว้าง และทำให้แบบสอบแต่ละฉบับมีจำนวนข้อก่อนข้างมาก นิยมให้มีค่าความยากอยู่ระหว่าง 0.20 - 0.80 เพื่อรองรับนักเรียนที่มีระดับความสามารถสูง ปานกลาง หรือต่ำ แต่อย่างไรก็ตามเมื่อนักเรียนมีความสามารถต่ำมากหรือสูงมาก ข้อสอบชุดดังกล่าวก็ไม่สามารถให้ผลการสอบที่สะท้อนความสามารถได้อย่างมีประสิทธิภาพ นอกจากนี้นักเรียนที่มีระดับความสามารถสูงจะเสียเวลาในการทำข้อสอบที่ง่ายเกินไป ขณะที่นักเรียนที่มีระดับความสามารถต่ำต้องทำข้อสอบที่ยากเกินไปอาจเกิดความท้อแท้ได้

การสอบแบบปรับเหมาะเป็นการจัดวางตำแหน่งของข้อสอบตามความยาก แล้วกำหนดให้นักเรียนทำตามความสามารถของตน กล่าวคือนักเรียนจะได้รับข้อสอบที่ยากขึ้น ๆ เมื่อเขาตอบได้ในทางตรงกันข้ามนักเรียนจะได้รับข้อสอบที่ง่ายขึ้น ๆ เมื่อเขาทำไม่ได้ ซึ่งการสอบปรับเหมาะนี้จะให้นักเรียนทำข้อสอบน้อยกว่าการสอบแบบดั้งเดิมและใช้เวลาสอบน้อยลงด้วย นอกจากนี้จากการศึกษาของ Weiss (Weiss and Kingsbury, 1984: 365 อ้างใน วสันต์ ทองไทย, 2539: 2) พบว่าการวัดความสามารถของนักเรียนในตัวแปรเดียวกันด้วยการสอบแบบปรับเหมาะมีความตรงและความเที่ยงเท่ากับหรือสูงกว่าการสอบแบบดั้งเดิม ถึงแม้ว่าจำนวนข้อสอบที่ทำจะลดลง 50 เปอร์เซ็นต์ และการสอบแบบนี้ยังช่วยผ่อนคลายความเครียดในการสอบได้ เช่น กรณีของนักเรียนที่มีความสามารถต่ำไม่จำเป็นต้องเครียดกับการทำข้อสอบที่ยากเกินความสามารถ ในขณะที่นักเรียนที่มีระดับความสามารถสูงก็ไม่จำเป็นต้องทำข้อสอบที่ง่ายเกินไปจนอาจมีผลให้ขาดความระมัดระวังในการทำข้อสอบ เป็นต้น ทั้งนี้ข้อสอบข้อที่ง่ายเกินไปหรือยากเกินไปเมื่อเทียบกับ

ระดับความสามารถของนักเรียนย่อมให้สารสนเทศ (Information) น้อย ซึ่งการคัดเลือกข้อสอบที่มีระดับความยากพอเหมาะกับความสามารถของนักเรียนตามทฤษฎีการตอบสนองข้อสอบ (IRT) จะให้สารสนเทศที่สำคัญเกี่ยวกับความสามารถที่แท้จริงของนักเรียน โดยการสอบแบบปรับเหมาะจะมีการคัดเลือกข้อสอบบนพื้นฐานของผลการตอบข้อสอบข้อที่ผ่านมา เมื่อนักเรียนทำข้อสอบข้อเริ่มต้นหรือชุดแรก (ขึ้นอยู่กับการออกแบบ) จากคลังข้อสอบแล้วจะมีการประเมินความสามารถของนักเรียน เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีค่าความยากเหมาะสมที่จะใช้วัดระดับความสามารถของนักเรียน แล้วประมาณค่าความสามารถของนักเรียนอีกครั้ง แล้วจึงเลือกข้อสอบที่เหมาะสมข้อต่อไป โดยอาศัยหลักที่ว่าถ้าตอบข้อสอบถูก ข้อถัดไปจะยากขึ้น แต่ถ้าตอบผิด ข้อถัดไปจะง่ายกว่าเดิม การตอบข้อสอบจะเป็นเช่นนี้จนกว่าการสอบจะสิ้นสุด (Green *et al*, 1984 อ้างใน สุพัฒน์ สุกมลสันต์, 2542: 279) ดังนั้นนักเรียนแต่ละคนจึงทำข้อสอบแตกต่างกัน และจำนวนข้อของการสอบไม่จำเป็นต้องเท่ากัน ทั้งนี้ความแตกต่างระหว่างการสอบแบบดั้งเดิมและการสอบแบบปรับเหมาะ แสดงได้ดังตารางที่ 1 (Lewis and Gary, 2003: 51; Tian *et al*, 2007)

ตารางที่ 1 ความแตกต่างระหว่างการสอบแบบดั้งเดิมและการสอบแบบปรับเหมาะ

ประเด็น	การสอบแบบดั้งเดิม	การสอบแบบปรับเหมาะ
ส่วนประกอบของแบบสอบ	กำหนดให้นักเรียนแต่ละคน ตอบข้อสอบเหมือนกัน ทั้ง ฉบับ	กำหนดให้นักเรียนแต่ละคน ได้ตอบข้อสอบที่แตกต่างกัน
จำนวนข้อสอบ	นักเรียนได้รับแบบสอบที่มี ความยาวเท่ากันและยาวกว่า การสอบแบบปรับเหมาะ	นักเรียนได้รับแบบสอบที่มี ความยาวต่างกันและสั้นกว่า การสอบแบบดั้งเดิม
ระดับความยากของข้อสอบ	ขึ้นอยู่กับนักเรียนส่วนมากที่ เข้าสอบ	ขึ้นอยู่กับนักเรียนแต่ละคน
เวลาในการทำแบบสอบ	กำหนดเวลาที่แน่นอนให้ นักเรียนทุกคนได้รับเท่ากัน	ไม่กำหนดเวลาขึ้นอยู่กับ ความสามารถของนักเรียนแต่ละ คน
การจัดเตรียมชุดการสอบ	ต้องใช้เวลามากในการสำเนา และจัดชุดแบบสอบ	ใช้เวลาเพียงสั้น ๆ
ระยะเวลารับผลการสอบ	ต้องรอเวลา	ทันที
ประสิทธิภาพของแบบสอบ	ความเที่ยงและความตรงใน การประมาณค่าความสามารถ มีความแม่นยำความน้อยกว่า การสอบแบบปรับเหมาะ	ความเที่ยงและความตรงใน การประมาณค่าความสามารถ มีความแม่นยำมากกว่าการ สอบแบบดั้งเดิม
ลำดับการทำข้อสอบ	สามารถเลือกทำข้อสอบข้อใด ก่อนก็ได้ และ เปลี่ยนแปลง คำตอบได้	กำหนดลำดับการตอบข้อสอบ ที่แน่นอน และ ไม่สามารถ ย้อนกลับไปตอบข้อสอบข้อที่ ผ่านมาแล้ว
การประมาณค่าความสามารถ	นับจำนวนข้อที่ตอบถูก	นำค่าพารามิเตอร์ของข้อสอบ ไปประมาณค่าความสามารถ ด้วย

ในการวิจัยครั้งนี้ได้จัดชุดการสอบแบบดั้งเดิมและดำเนินการสอบเหมือนกับการสอบโดยทั่วไป ประมาณค่าความสามารถด้วยโปรแกรม BILOG ซึ่งค่าความยากของข้อสอบในการสอบแบบปรับเหมาะได้จากวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) เนื่องจากทฤษฎีการสอบแบบดั้งเดิมให้ค่าความยากที่ขึ้นอยู่กับความสามารถของกลุ่มนักเรียนที่เข้าสอบ ทำให้ทฤษฎีการสอบแบบดั้งเดิมมีคุณสมบัติไม่เพียงพอในการนำมากำหนดค่าความยากของข้อสอบและการประมาณค่าความสามารถที่แท้จริงของนักเรียน

ประเภทของการสอบแบบปรับเหมาะ

การสอบแบบปรับเหมาะแบ่งออกเป็น 2 ประเภทใหญ่ ๆ ดังนี้ (Weiss, 1974: 78-110 cited in Hambleton and Swaminathan, 1985: 297)

1. กลวิธีสองขั้นตอน (Two-Stage Strategies) เป็นการสอบที่นักเรียนต้องทำแบบสอบ 2 ชุด โดยชุดแรกเป็นแบบสอบแยกทาง (Routing Test) ซึ่งนักเรียนทุกคนต้องสอบเหมือนกัน เพื่อแยกกลุ่มนักเรียนตามระดับความสามารถ ส่วนชุดที่สองเป็นแบบสอบหลัก (Main Test) ซึ่งแบ่งเป็นแบบสอบชุดย่อย ๆ หลายชุดตามระดับความยาก นักเรียนที่ได้รับการประเมินจากแบบสอบแยกกลุ่มว่ามีระดับความสามารถในระดับต่ำจะได้รับแบบสอบหลักที่มีค่าความยากในระดับต่ำ ส่วนนักเรียนที่ได้รับการประเมินจากแบบสอบแยกกลุ่มว่ามีระดับความสามารถในระดับสูงจะได้รับแบบสอบหลักที่มีค่าความยากในระดับสูง (Lord, 1971: 85 อ้างใน สุพัฒน์ สุกมลสันต์, 2539: 41)

2. กลวิธีหลายขั้นตอน (Multi-Stage Strategies) เป็นการสอบที่นักเรียนต้องตอบข้อสอบเพื่อแยกทางหลายครั้ง ซึ่งต่างจากกลวิธีสองขั้นตอนที่นักเรียนต้องตัดสินใจแยกทางเพียงครั้งเดียวจากการตอบแบบสอบแยกทาง (Routing Test) ไปยังแบบสอบหลัก (Main Test) สำหรับรูปแบบของกลวิธีหลายขั้นตอนแบ่งออกเป็น 2 รูปแบบ ดังนี้

2.1 แบบแยกทางคงที่ (Fixed Branching Model) เป็นการสอบแบบปรับเหมาะที่มีหลายขั้นตอน ในแต่ละตอนอาจจะมีข้อสอบเพียงข้อเดียวหรือหลายข้อก็ได้ โดยมีแบบหรือเส้นทางในการตอบข้อสอบที่แน่นอน แบ่งเป็นรูปแบบย่อยได้ ดังนี้

2.1.1 แบบสอบรูปพีระมิด เป็นรูปแบบแรกของแบบสอบปรับเหมาะ และแบ่งออกเป็นรูปแบบย่อย ๆ 5 รูปแบบ คือ รูปพีระมิดขนาดขั้นคงที่ (Constant Step Size Pyramid) รูปพีระมิดขั้นแปรผัน (Variable Step Size Pyramid) รูปพีระมิดข้างตัด (Truncated Pyramid) รูปพีระมิดชนิดมีหลายข้อในแต่ละขั้น (Multiple – Item Pyramid) และ รูปพีระมิดชนิดให้น้ำหนักหลายตัวเลือกเพื่อแยกทาง (Pyramid Using Differential Response Option Branching)

2.1.2 แบบสอบเฟล็กซิเลเวล (Flexilevel Test) เป็นแบบสอบที่ดัดแปลงมาจากรูปพีระมิด โดยมีข้อสอบเพียงข้อเดียวในแต่ละระดับความยาก ซึ่งต่างกับรูปพีระมิดที่มีข้อสอบมากกว่าหนึ่งข้อในแต่ละระดับความยาก จึงทำให้การแยกทางของนักเรียนมีความแตกต่างกัน

2.1.3 แบบสอบปรับระดับขั้น (Stradaptive Test) เป็นแบบสอบที่นำข้อสอบมาจัดกลุ่ม (Stratified) ตามค่าความยากแล้วแบ่งออกเป็นระดับขั้น (Stratum) ในแต่ละระดับขั้นจะประกอบด้วยข้อสอบที่มีค่าความยากใกล้เคียงและจัดเรียงลำดับขั้นตามค่าความยากของข้อสอบที่สูงขึ้น

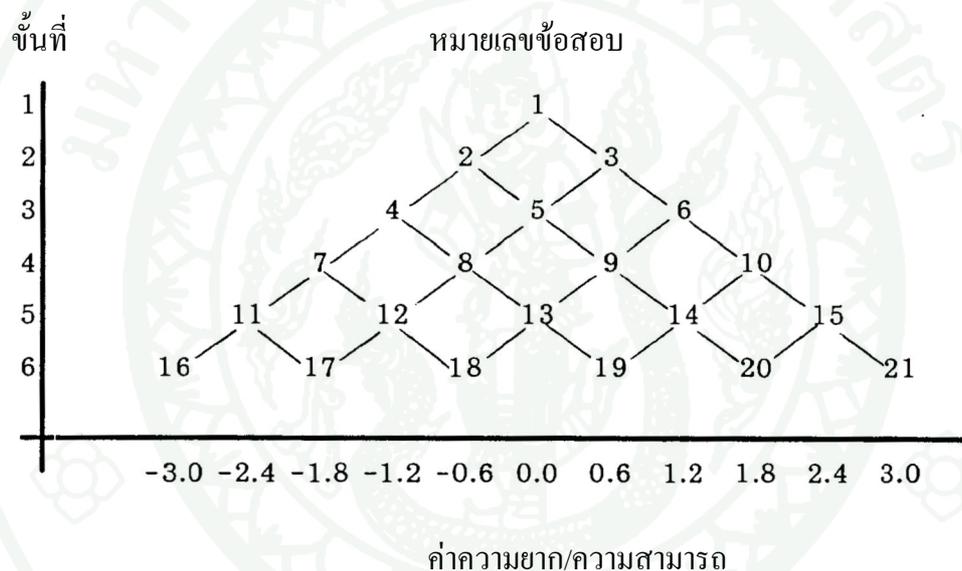
2.2 แบบแยกทางแปรผัน (Variable–Branching Model) เป็นรูปแบบที่ไม่มีการกำหนดข้อสอบและเส้นทางในการตอบข้อสอบไว้ล่วงหน้า โดยดำเนินการสอบด้วยเครื่องคอมพิวเตอร์ ซึ่งเครื่องคอมพิวเตอร์จะคัดเลือกข้อสอบจากกลุ่มข้อสอบที่มีค่าความยากของข้อสอบใกล้เคียงกับระดับความสามารถของนักเรียน โดยใช้ข้อมูลจากผลการตอบในแต่ละข้อ การทดสอบจะยุติเมื่อการประมาณค่าความสามารถของนักเรียนมีค่าคงที่ และให้ค่าความคลาดเคลื่อนมาตรฐานเท่ากับเกณฑ์ที่ตั้งไว้ สำหรับรูปแบบแยกทางแปรผันที่แพร่หลาย ได้แก่ กลวิธีของเบส์ (Bayesian Strategies) และกลวิธีประมาณค่าความเป็นไปได้สูงสุด (Maximum Likelihood Strategies)

ในการศึกษาครั้งนี้ ผู้วิจัยได้ใช้การสอบแบบดั้งเดิมและการสอบแบบปรับเหมาะรูปพีระมิด ซึ่งเป็นรูปแบบหนึ่งของการสอบแบบปรับเหมาะรูปแบบแยกทางคงที่ (Fixed Branching Model) ที่ใช้แบบสอบเพียงชุดเดียวและมีแบบหรือเส้นทางในการตอบข้อสอบที่แน่นอน เหมาะสมในการนำมาหาค่าประมาณความสามารถด้วยวิธีที่แตกต่างกัน ดังนั้นผู้วิจัยจึงขอเสนอรายละเอียดเกี่ยวกับแบบสอบรูปพีระมิดเท่านั้น

การสอบแบบปรับเหมาะรูปพีระมิด

การสอบแบบปรับเหมาะรูปพีระมิด สามารถจัดรูปแบบของแบบสอบได้โดยอาศัยโครงสร้างของแบบจำลองรูปพีระมิด ซึ่งแบ่งย่อยได้เป็น 5 แบบ ดังนี้

1. รูปพีระมิดขนาดขั้นคงที่ (Constant Step Size Pyramid) ได้แก่ การสอบแบบปรับเหมาะที่จัดเป็นรูปสามเหลี่ยม โดยมีข้อสอบในแต่ละชั้นเท่ากับลำดับที่ของชั้น กล่าวคือ ชั้นที่ 1 จะมีข้อสอบ 1 ข้อ ชั้นที่ 2 จะมีข้อสอบ 2 ข้อ ในชั้นที่ 6 จะมีข้อสอบ 6 ข้อ ดังนั้นแบบสอบปรับเหมาะรูปพีระมิดขนาดขั้นคงที่ จะมีข้อสอบทั้งหมด 21 ข้อ ดังภาพ

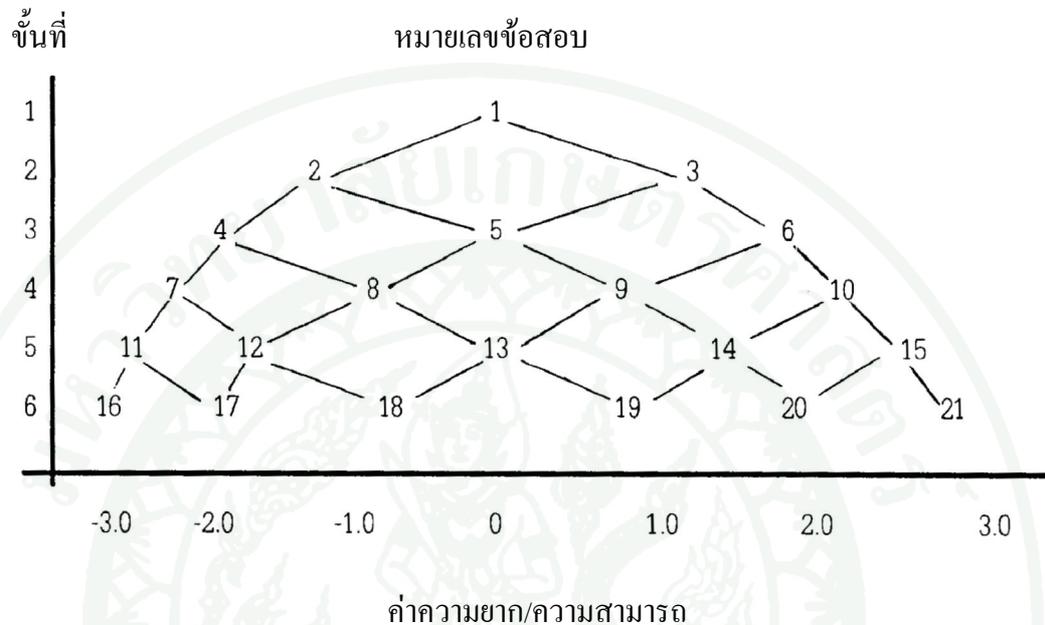


ภาพที่ 2 แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดแบบขนาดขั้นคงที่

จากภาพ แกนนอนแสดงถึงค่าความยากของข้อสอบ (b) ที่สัมพันธ์กับข้อสอบในแบบสอบปรับเหมาะรูปพีระมิด โดยมีค่าความยากอยู่ระหว่าง -3.0 ถึง 3.0 ข้อสอบที่อยู่ในแนวตั้งแนวเดียวกัน จะมีค่าความยากเท่ากัน

2. รูปพีระมิดขั้นแปรผัน (Variable Step Size Pyramid) โดยที่ Paterson และ Lord (1971 cited in Wiess, 1974: 18) มีความเห็นว่า รูปพีระมิดแบบขนาดขั้นคงที่ ขาดประสิทธิภาพเกี่ยวกับความไวในการกำหนดเส้นทางการตอบข้อสอบ จึงกำหนดช่วงห่างของค่าความยากของข้อที่อยู่ติดกันภายในชั้นเดียวกันให้มีค่าไม่เท่ากัน โดยให้ข้อที่อยู่กลาง ๆ มีช่วงห่างของค่าความยากมาก

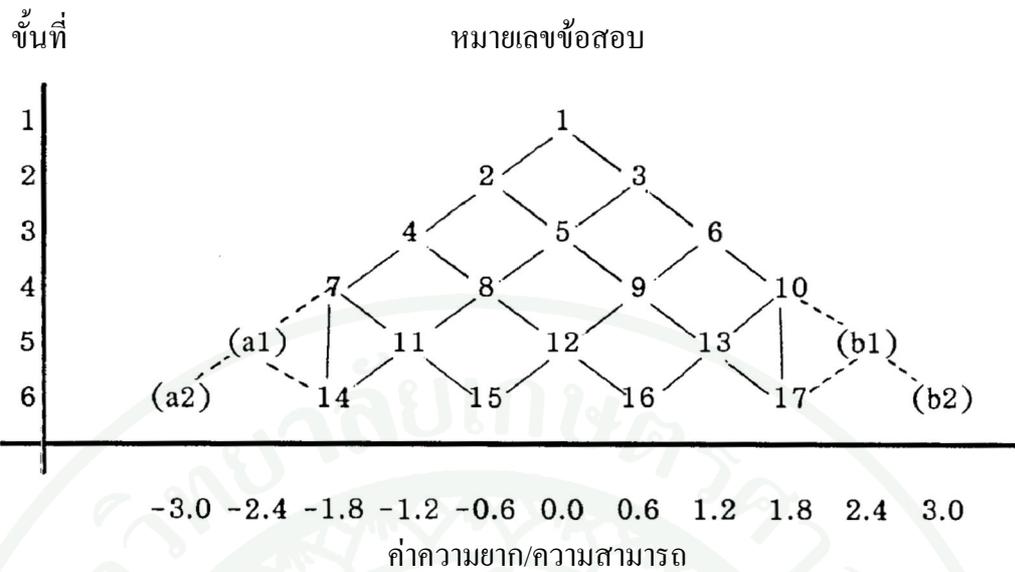
แล้วค่อย ๆ ลดลงในข้อที่ก่อนไปทางง่ายหรือยาก Paterson ได้กำหนดให้การเพิ่มหรือลดนั้น จะเพิ่มหรือลดไปยังตำแหน่งกึ่งกลางระหว่าง 2 ชั้นที่มาก่อน หรือ ระหว่างข้อในชั้นที่มาก่อนกับค่าความยากสูงสุดหรือต่ำสุด ดังภาพ



ภาพที่ 3 แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดแบบขนาดชั้นแปรผัน

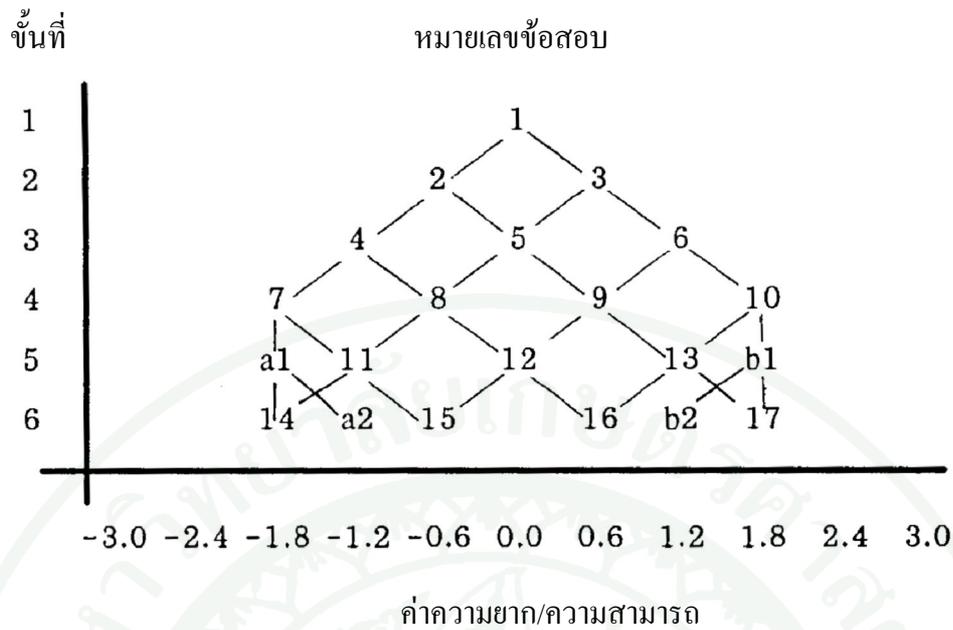
จากภาพ ในชั้นที่ 1 มีข้อสอบเพียงข้อเดียว คือ ข้อ 1 ซึ่งมีความยากเท่ากับ 0.00 ในข้อที่ 2 มีความยากอยู่กึ่งกลางระหว่าง -3.00 (ความยากต่ำสุด) กับ 0.00 (ความยากของข้อที่ 1) จึงมีความยากเท่ากับ -1.50 ในชั้นที่ 3 มีข้อสอบ 3 ข้อ คือ ข้อ 4 มีความยากอยู่กึ่งกลางระหว่าง -3.00 (ความยากต่ำสุด) กับ -1.50 (ความยากของข้อที่ 2) จึงมีความยากเท่ากับ -2.25 ข้อ 5 มีความยากเท่ากับ 0.00 ข้อ 6 มีความยากอยู่กึ่งกลางระหว่าง 1.50 (ความยากของข้อที่ 3) กับ 3.00 (ความยากสูงสุด) จึงมีความยากเท่ากับ 2.25 สำหรับในชั้นอื่น ๆ ก็ในทำนองเดียวกัน

3. รูปพีระมิดข้างตัด (Truncated Pyramid) เป็นรูปแบบที่ Mussio (1972 cited in Weiss, 1974: 22) เป็นผู้เสนอรูปแบบนี้ขึ้นมา เพื่อให้ลดจำนวนข้อสอบลงจากรูปพีระมิดแบบขนาดชั้นคงที่ จึงใช้วิธีสกัดการสะท้อนกลับ หรือ วิธีรักษาการสะท้อนกลับ (Reflecting or Retaining Barrier) ดังภาพ



ภาพที่ 4 แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด แบบวิธีสกัดการสะท้อนกลับ

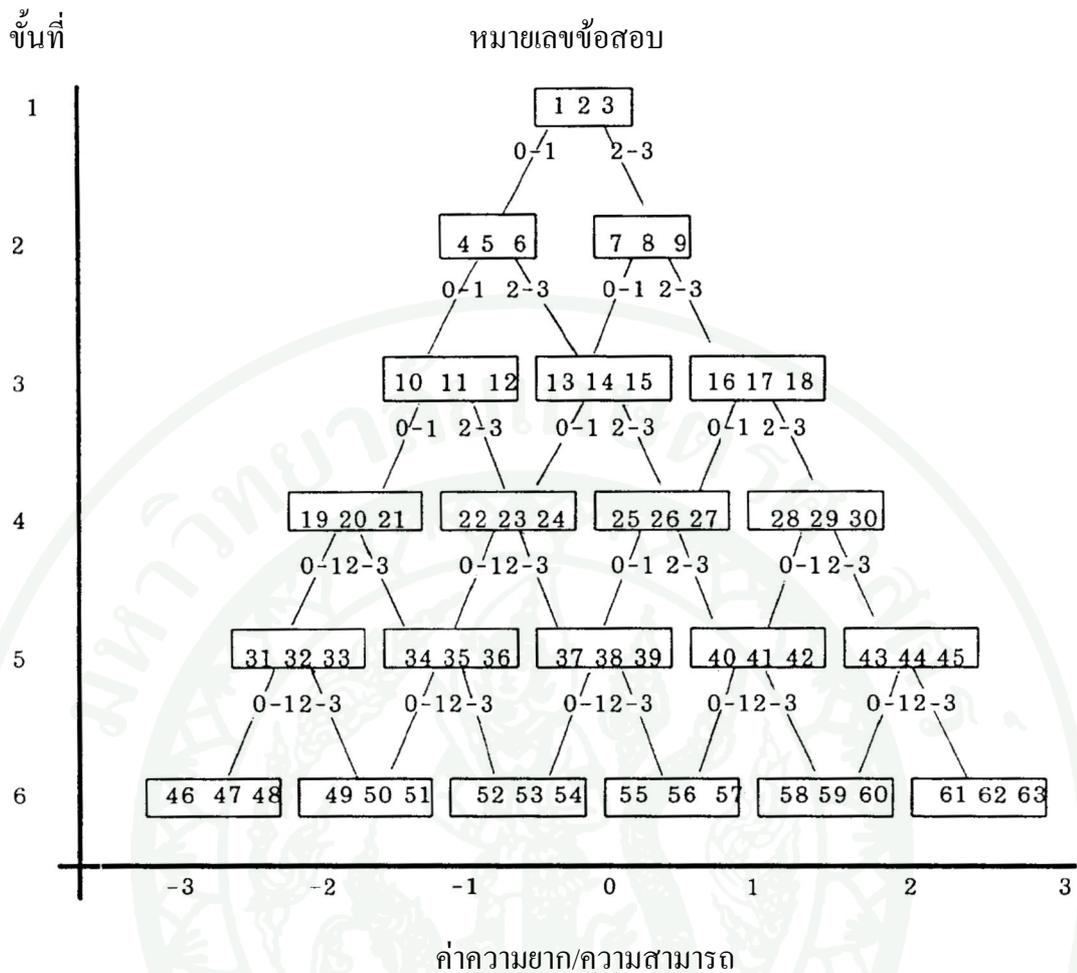
จากภาพ การตอบในชั้นที่ 1 ถึงชั้นที่ 3 มีเส้นทางการตอบเช่นเดียวกับรูปพีระมิดขนาดชั้นลงที่ จนถึงจุดสกัดในชั้นที่ 4 ถ้าตอบข้อสอบข้อที่ 7 ถูก ก็จะแยกไปตอบข้อที่ 11 แต่ถ้าตอบข้อที่ 7 ผิด ก็จะลงไปตอบข้อที่ 14 ในทำนองเดียวกัน ถ้าตอบข้อ 10 ผิด ก็จะแยกไปตอบข้อ 13 แต่ถ้าตอบข้อ 10 ถูก ก็จะลงไปตอบข้อ 17 เลย



ภาพที่ 5 แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด แบบวิธีรักษาการสะท้อนกลับ

จากภาพ การตอบข้อสอบในชั้นที่ 1 ถึง ชั้นที่ 3 มีเส้นทางการตอบเช่นเดียวกับรูปพีระมิดขนาดชั้นคงที่ จนถึงจุดสกัดในชั้นที่ 4 ถ้าตอบข้อสอบข้อที่ 7 ถูก ก็จะแยกไปตอบข้อที่ 11 แต่ถ้าตอบข้อที่ 7 ผิด ก็ลงไปตอบข้อ a1 ซึ่งมีความยากเท่ากับข้อ 7 และถ้าตอบข้อ a1 ผิดอีกก็จะลงไปตอบข้อ 14 แต่ถ้าตอบข้อ a1 ถูกก็จะแยกไปตอบข้อที่ a2 ซึ่งมีความยากเท่ากับข้อ 11

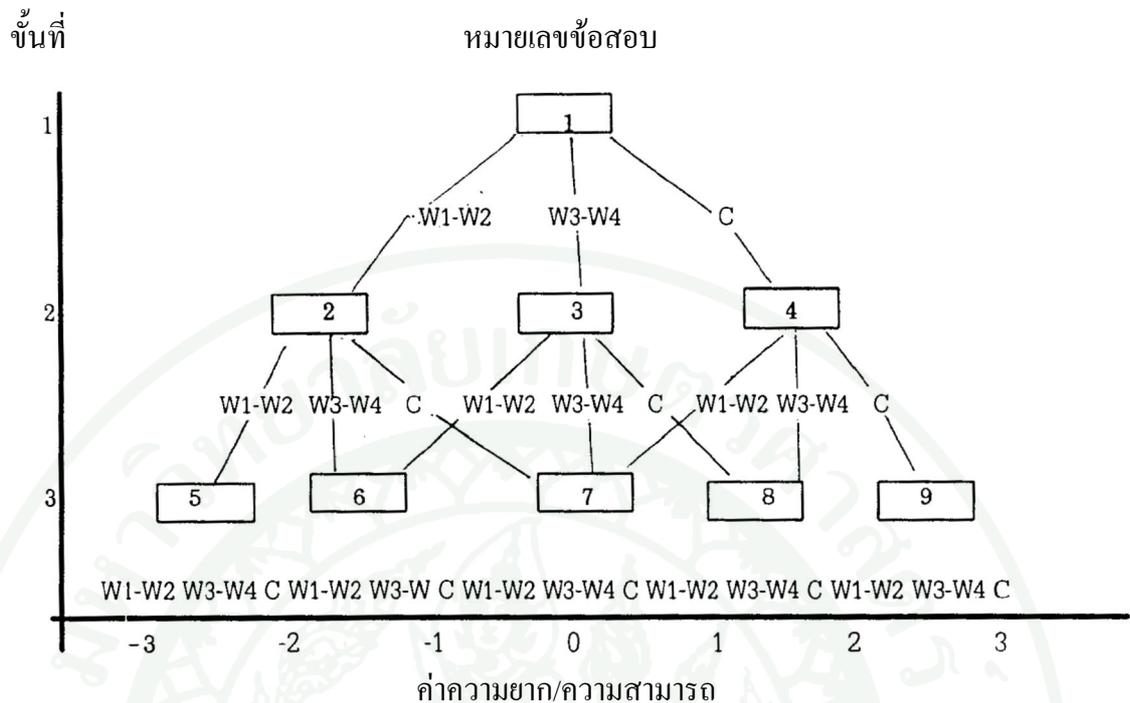
4. รูปพีระมิดแบบมีหลายข้อในแต่ละชั้น (Multiple-Item Pyramid) เป็นรูปแบบที่ Krathworth และ Huyser (1956 cited in Weiss, 1974: 25) เป็นผู้เสนอรูปแบบนี้ขึ้นมา โดยให้มีจำนวนข้อในแต่ละชั้นมากกว่าหนึ่งข้อ ดังภาพ



ภาพที่ 6 แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดแบบมีหลายข้อในแต่ละชั้น

จากภาพ ในชั้นที่ 1 ถ้านักเรียนตอบข้อสอบผิดทุกข้อ หรือ ตอบถูก 1 ข้อ จะแยกไปทำข้อสอบข้อที่ 4 5 และ 6 แต่ถ้าตอบถูก 2 ถึง 3 ข้อ ก็จะแยกไปทำข้อสอบข้อที่ 7 8 และ 9 สำหรับในชั้นอื่น ๆ ก็ทำนองเดียวกัน

5. รูปพีระมิดชนิดใ้ให้นักหลายตัวเลือกเพื่อแยกทาง (Pyramid Using Differential Response Option Branching) เป็นรูปแบบที่ Bayroff และ Seeley เสนอขึ้น (1967 cited in Weiss, 1974: 28) ซึ่งมีรูปแบบคล้ายกับแบบที่ใช้ข้อสอบหลาย ๆ ข้อในแต่ละชั้น กล่าวคือแทนที่จะใช้ข้อสอบหลายข้อในแต่ละชั้น ก็ใช้ตัวเลือกของข้อสอบแทน รูปแบบนี้เหมาะสำหรับข้อสอบเลือกตอบ ที่สามารถกำหนดน้ำหนักคะแนนของตัวเลือกได้ ดังภาพ



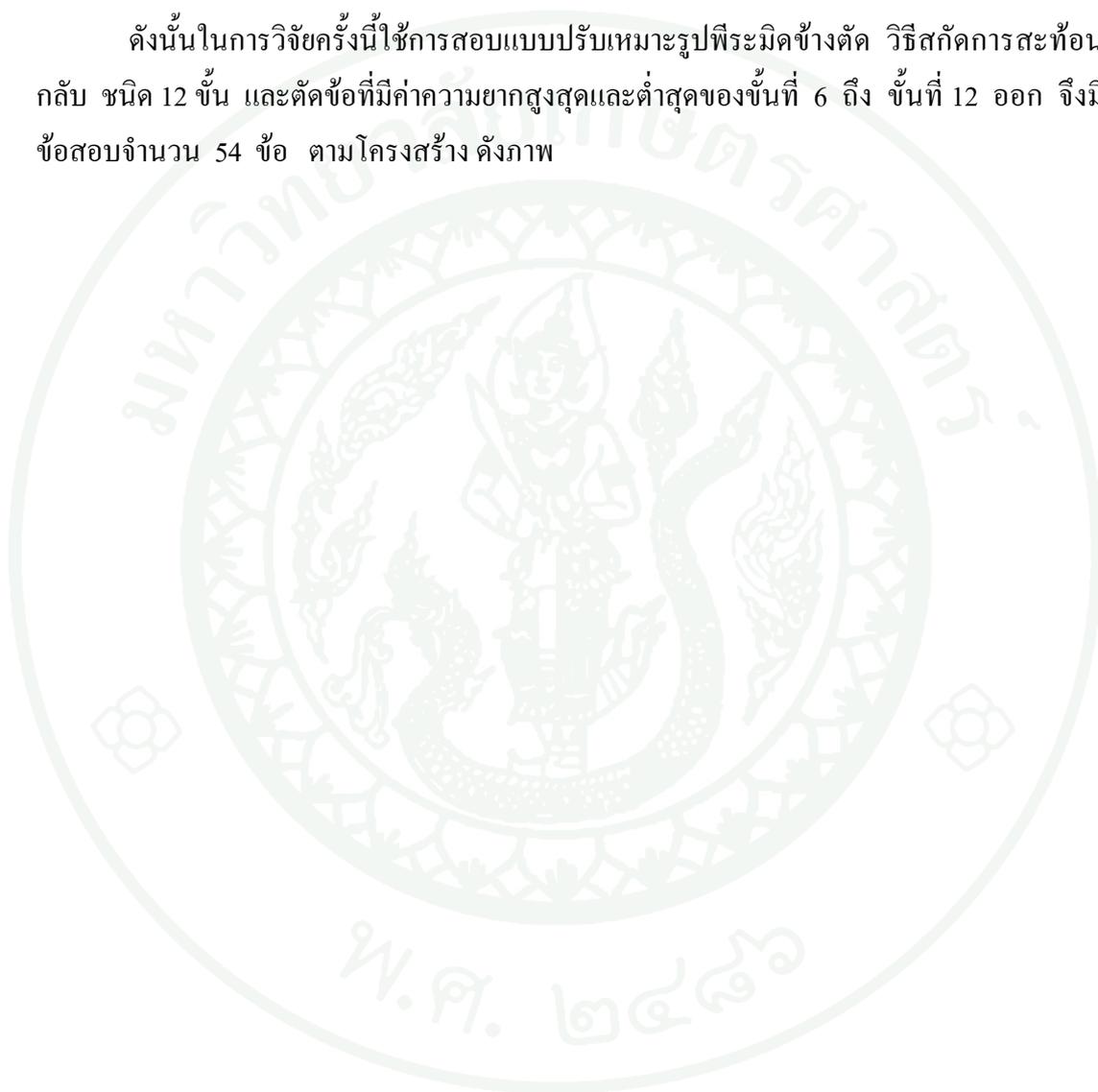
ภาพที่ 7 แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดชนิดให้น้ำหนักหลายตัวเลือก เพื่อแยกทาง

จากภาพ ในชั้นที่ 1 เมื่อนักเรียนตอบข้อสอบหมายเลข 1 ถูก หรือ ตัวเลือก C จะถูกแยกทางให้ไปตอบข้อสอบหมายเลข 4 ที่อยู่ในชั้นที่ 2 ซึ่งเป็นข้อสอบที่มีความยากมากที่สุด ในชั้นที่ 2 แต่ถ้านักเรียนตอบตัวเลือกที่มีระดับความผิดมาก คือ ตัวเลือก W_1 หรือ W_2 จะถูกแยกทางให้ไปตอบข้อสอบหมายเลข 2 ที่อยู่ในชั้นที่ 2 ซึ่งมีความยากน้อยกว่าข้อสอบหมายเลข 1 ส่วนนักเรียนที่เลือกตอบตัวเลือก W_3 หรือ W_4 ซึ่งมีระดับความผิดน้อยกว่า ตัวเลือก W_1 หรือ W_2 แต่ยังไม่ใช่คำตอบที่ถูกต้อง จะต้องแยกทางไปตอบข้อสอบหมายเลข 3 ที่อยู่ในชั้นที่ 2 ซึ่งมีความยากเท่ากับข้อสอบหมายเลข 1 ที่อยู่ในชั้นที่ 1 เพราะฉะนั้นนักเรียนที่จะเลือกตอบตัวเลือก W_3 หรือ W_4 ส่วนใหญ่จะมีความสามารถในระดับเฉลี่ยของกลุ่มคนทั่วไป ส่วนนักเรียนที่มีระดับความสามารถสูงหรือต่ำกว่านี้จะถูกแยกทางให้ไปตอบข้อสอบที่มีระดับความยากสูงหรือต่ำกว่า การสอบจะดำเนินการเช่นนี้จนถึงขั้นสุดท้ายของแบบสอบ

การวิจัยครั้งนี้ได้นำแบบจำลองโครงสร้างการสอบแบบปรับเหมาะประเภทกลวิธีหลายขั้นตอนรูปพีระมิดข้างต้นแบบวิธีสกัดการสะท้อนกลับ เพราะเป็นการสอบที่มีการประมาณค่าความสามารถของนักเรียนทีละข้อ จึงทำให้มีความแม่นยำในการประมาณค่าความสามารถของ

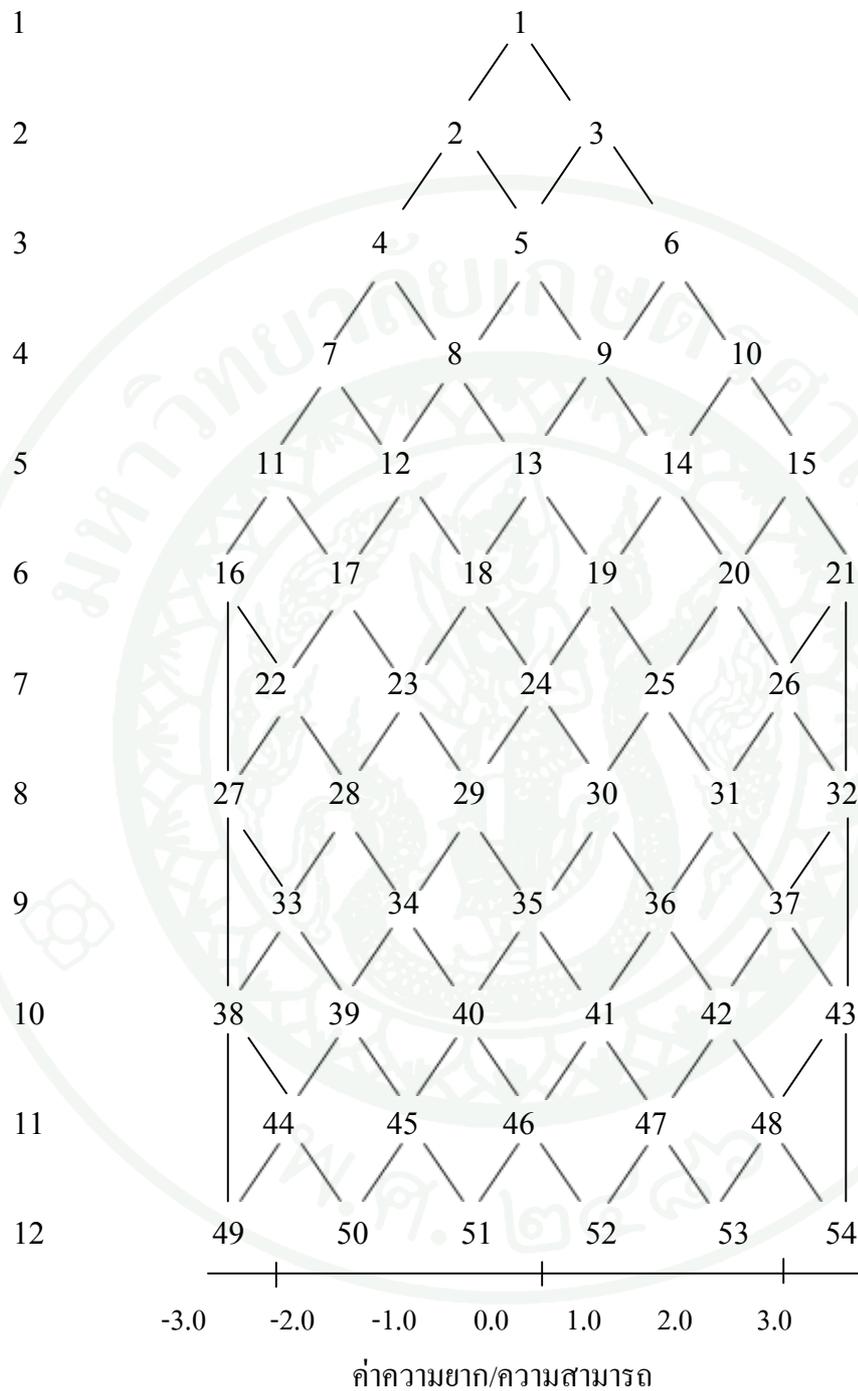
นักเรียนมากกว่าการสอบแบบปรับเหมาะประเภทลวิธีสองขั้นตอน (Hambleton and Swaminathan, 1985: 296 -298) นอกจากนี้ผลการศึกษาของ วสันต์ ทองไทย (2539) พบว่า การสอบแบบปรับเหมาะรูปพีระมิดขนาดชั้นคงที่และรูปพีระมิดข้างตัดแบบวิธีสกัดการสะท้อนกลับ ให้ผลการประมาณค่าความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์ไม่แตกต่างกัน

ดังนั้นในการวิจัยครั้งนี้ใช้การสอบแบบปรับเหมาะรูปพีระมิดข้างตัด วิธีสกัดการสะท้อนกลับ ชนิด 12 ชั้น และตัดข้อที่มีค่าความยากสูงสุดและต่ำสุดของชั้นที่ 6 ถึง ชั้นที่ 12 ออก จึงมีข้อสอบจำนวน 54 ข้อ ตามโครงสร้าง ดังภาพ



ชั้นที่

หมายเลขข้อสอบ



ภาพที่ 8 แบบจำลองโครงสร้างการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดวิธีสกัดการสะท้อนกลับ
ชนิด 12 ชั้น

วิธีการประมาณค่าความสามารถด้วยการสอบแบบปรับเหมาะรูปพีระมิด

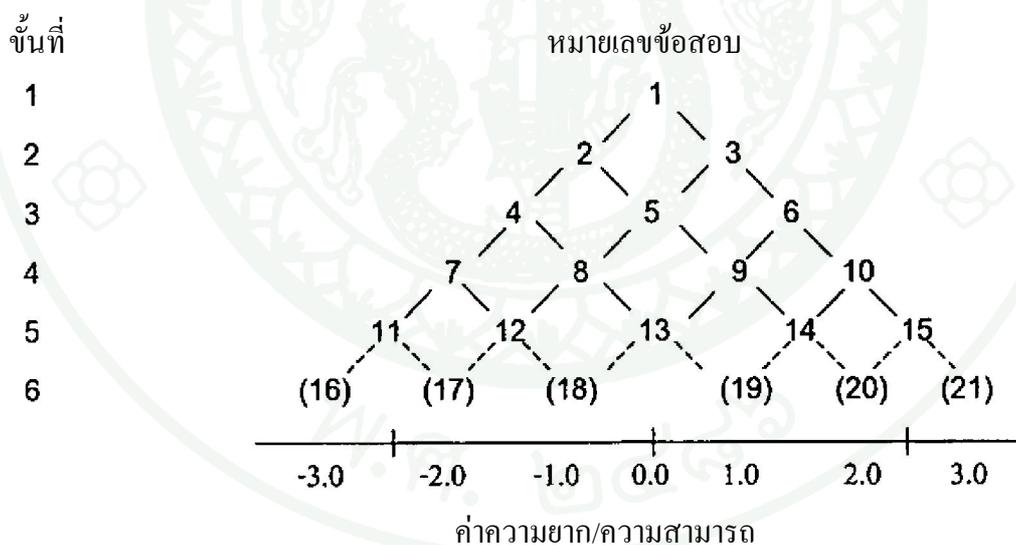
ในการสอบแบบปรับเหมาะ นักเรียนแต่ละคนจะได้ทำข้อสอบที่มีความแตกต่างกันทั้งจำนวนข้อสอบ ค่าพารามิเตอร์ของข้อสอบ และเวลาในการดำเนินการสอบ จึงไม่เหมาะสมในการประมาณค่าความสามารถของนักเรียนจากการนับจำนวนข้อสอบที่ตอบ ดังนั้นจึงมีการนำทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้ในการประมาณค่าความสามารถของนักเรียน โดยคำนึงถึงความยากของข้อสอบที่นักเรียนแต่ละคนตอบ เนื่องจากอยู่บนมาตราวัดเดียวกันกับค่าความสามารถ ซึ่งวิธีการประมาณค่าความสามารถของนักเรียนด้วยการสอบแบบปรับเหมาะรูปพีระมิด มีรายละเอียดดังนี้

1. วิธีการประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (Average Difficulty of All Item Answered Correctly) (Weiss, 1974: 33) โดยการนำค่าความยากของข้อสอบทุกข้อที่นักเรียนทำถูกมาหาค่าเฉลี่ย ค่าเฉลี่ยที่ได้นี้ถือเป็นค่าความสามารถของนักเรียน ซึ่งวิธีนี้จะมี ความหมายในการบอกระดับความสามารถของนักเรียนได้มากพอกับวิธีการประมาณค่าความสามารถตามผลการตอบข้อสุดท้าย แต่ให้ข้อมูลของค่าความสามารถมากกว่า นอกจากนี้วิธีนี้ และวิธีการประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบจะมีความคงที่ของค่าความสามารถ อย่างไรก็ตามการประมาณค่าความสามารถด้วยวิธีนี้จะไม่มีความไวต่อการเดาคำตอบของนักเรียน ซึ่งอาจมีผลให้ค่าความเที่ยงในการวัดลดลง

2. วิธีการประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (Average Difficulty) Lord (1970 cited in Weiss, 1974: 33) ได้เสนอให้นำค่าความยากของข้อสอบทุกข้อที่นักเรียนได้ทำ โดยไม่คำนึงถึงว่าจะตอบถูกหรือไม่มาหาค่าเฉลี่ย ค่าเฉลี่ยที่ได้นี้ถือเป็นค่าความสามารถของนักเรียน วิธีการประมาณค่าความสามารถโดยวิธีนี้จะให้ข้อมูลของค่าความสามารถเช่นเดียวกับวิธีการประมาณค่าความสามารถจากจำนวนข้อสอบที่ตอบถูกในการสอบรูปพีระมิดข้างตัด แต่วิธีนี้มีข้อจำกัดในการวัดระดับความสามารถของนักเรียนได้ไม่กว้างมากนัก เนื่องจากค่าความสามารถที่ได้จากวิธีนี้จะรวมค่าความยากของข้อสอบที่มีค่าความยากใกล้เคียงกับความสามารถเฉลี่ยของนักเรียนแล้วนำมาเป็นค่าความสามารถ ค่าที่ได้จึงไม่กระจายไปครอบคลุมถึงระดับความสามารถของนักเรียนที่สูงและต่ำมาก ๆ

3. วิธีการประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (Difficulty of the Final Item) (Weiss, 1974: 31) โดยนำค่าความยากของข้อสุดท้ายที่นักเรียนตอบไม่ว่าจะตอบถูกหรือผิด ความยากนี้ถือเป็นค่าความสามารถของนักเรียน ซึ่งเป็นการประมาณค่าความสามารถในการสอบที่ปฏิบัติได้สูงสุด (Maximum Performance) แต่ค่าความสามารถที่ได้จะมีความน่าเชื่อถือน้อยลงเมื่อมีการตอบข้อสอบโดยการเดาและถ้าจำนวนขั้นของพีระมิดมีน้อยค่าความสามารถที่ได้จะน้อยเกินไป มีผลให้ค่าสหสัมพันธ์ระหว่างการสอบแบบปรับเหมาะรูปพีระมิดกับแบบสอบอื่นมีค่าต่ำ

4. วิธีการประมาณค่าความสามารถตามค่าความยากของข้อสอบที่ต่อจากข้อสุดท้ายที่ทำ (Difficulty of (N+1)th Item) (Weiss, 1974: 31) โดยถ้าขั้นสุดท้ายที่ทำถูกค่าความยากในขั้นต่อไปจะเพิ่มขึ้น ดังนั้นค่าความสามารถก็จะเท่ากับค่าความยากในขั้นต่อไปที่เพิ่มขึ้นหรือถ้าขั้นสุดท้ายทำผิด ค่าความยากในขั้นต่อไปก็จะลดลง ดังนั้นค่าความสามารถจะเท่ากับค่าความยากในขั้นต่อไปที่ลดลง การประมาณค่าความสามารถตามค่าความยากโดยวิธีนี้จึงมีการสมมติข้อสอบต่อจากขั้นที่ N ในแบบสอบจริง ตามแผนภาพ



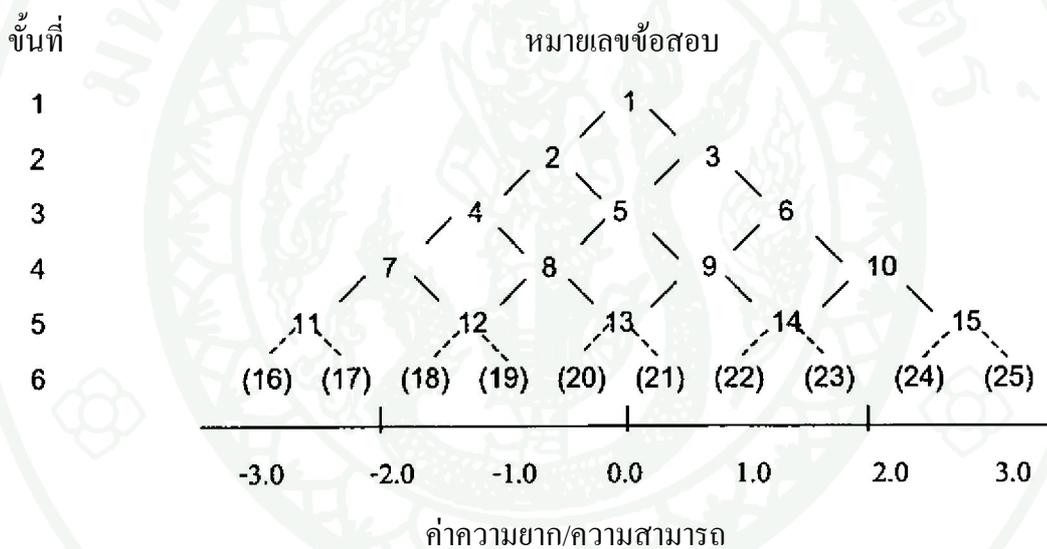
ภาพที่ 9 การประมาณค่าความสามารถตามค่าความยากของข้อสอบที่ต่อจากข้อสุดท้ายที่ทำ

จากภาพ ถ้านักเรียนทำข้อสอบหมายเลขข้อ 12 ถูก ก็จะประมาณค่าความสามารถได้เท่ากับ -1.0 ซึ่งเป็นค่าความยากของหมายเลขข้อ 18 แต่ถ้านักเรียนทำข้อสอบหมายเลขข้อ 12 ผิดก็จะประมาณค่าความสามารถได้เท่ากับ -2.0 ซึ่งเป็นค่าความยากของหมายเลขข้อ 17 โดย

หมายเลขข้อ 18 และ 17 จะอยู่ในชั้นที่ 6 ซึ่งเป็นชั้นที่สมมติขึ้น โดยการประมาณค่าความสามารถด้วยวิธีนี้จะให้ช่วงของค่าความสามารถแคบกว่าแบบสอบรูปพีระมิดมีจำนวนชั้นน้อย

5. วิธีการประมาณค่าความสามารถตามผลการตอบข้อสุดท้าย (Terminal Right-Wrong)

(Weiss, 1974: 31) จากการประมาณค่าความสามารถตามวิธีในข้อ 1 อาจทำให้ค่าความสามารถของนักเรียนที่ทำข้อสอบง่ายกว่าแล้วตอบถูกมีความสามารถเท่ากับนักเรียนที่ทำข้อสอบยากกว่าแล้วตอบผิด ทั้ง ๆ ที่นักเรียนทั้งสองคนน่าจะมีความสามารถที่แตกต่างกัน การประมาณค่าความสามารถตามผลการตอบข้อสุดท้ายจึงมีการสมมติข้อสอบต่อจากชั้นที่ N ในแบบสอบจริงซึ่งมีทิศทางต่างจากการประมาณค่าความสามารถตามค่าความยากของข้อสอบที่ต่อจากข้อสุดท้ายที่ทำตามแผนภาพ



ภาพที่ 10 การประมาณค่าความสามารถตามค่าความยากของข้อสอบที่ต่อจากข้อสุดท้ายที่ทำ

จากภาพ จะเห็นได้ว่าข้อสอบแต่ละข้อในชั้นสุดท้ายหรือชั้นที่ 5 นั้นสามารถแยกทางไปยังข้อสอบที่สมมติขึ้นในชั้นที่ 6 อีก 12 ข้อ โดยต่างจากการประมาณค่าความสามารถวิธีที่ 2 ซึ่งมีข้อสอบสมมติเพียง 6 ข้อ สำหรับวิธีนี้ถ้านักเรียนทำข้อสอบหมายเลขข้อ 12 ถูกก็จะประมาณค่าความสามารถได้เท่ากับ -1.2 ซึ่งเป็นค่าความยากของหมายเลขข้อ 19 แต่ถ้านักเรียนทำข้อสอบหมายเลข 12 ผิด ก็จะประมาณค่าความสามารถได้เท่ากับ -1.8 ซึ่งเป็นค่าความยากของหมายเลขข้อ 18

6. วิธีการประมาณค่าความสามารถจากข้อสอบทุกข้อตามวิธีของแฮนทซ์เซน (All Item Score) Hansen (1969 cited in Weiss, 1974: 34) ได้เสนอวิธีการให้คะแนนข้อสอบทุกข้อ ซึ่งมีข้อตกลงว่าค่าความยากของข้อสอบและระดับความสามารถของนักเรียนถือเป็นเรื่องเดียวกัน โดยข้อสอบที่ตอบถูกจะให้ 2 คะแนน ข้อสอบที่ง่ายกว่าข้อที่ตอบถูกในชั้นเดียวกันให้อีกข้อละ 2 คะแนน ข้อสอบที่ยากกว่าข้อที่ตอบถูกให้ 1 คะแนน ส่วนข้อที่มีค่าความยากมากกว่าข้อที่ตอบถูกในชั้นเดียวกันที่เหลือจะให้ 0 คะแนน สำหรับข้อที่ตอบผิดจะให้ 0 คะแนน ข้อสอบที่มีค่าความยากน้อยกว่าและอยู่ติดกับข้อสอบที่ตอบผิดที่อยู่ในชั้นเดียวกันให้ 1 คะแนน และข้อสอบที่มีความยากน้อยกว่าข้อที่ตอบผิดที่เหลือทุกข้อให้ข้อละ 2 คะแนน ซึ่งผลการศึกษาของ Larkin และ Weiss พบว่า การประมาณค่าความสามารถด้วยวิธีนี้มีความสัมพันธ์กับการประมาณค่าตามค่าความยากเฉลี่ยของข้อสอบที่ได้ทำสูงถึง .99

จากวิธีการประมาณค่าความสามารถข้างต้น พบว่าแต่ละวิธีมีข้อจำกัดแตกต่างกันในเรื่องของจำนวนชั้นของรูปพีระมิดที่ใช้ในการสอบต้องมีจำนวนชั้นไม่น้อยเกินไป เพื่อให้ระดับความสามารถที่ได้มีค่าไม่น้อยและมีช่วงความกว้างของค่าประมาณความสามารถไม่แคบเกินไป ความไวในการเดาคำตอบที่อาจตอบถูกโดยการเดามากกว่าตอบถูกจากความสามารถที่แท้จริง ทำให้ผลการประมาณค่าความสามารถขาดความแม่นยำ รวมทั้งความตรงตามเกณฑ์สัมพัทธ์ ความเที่ยงและค่าสหสัมพันธ์ระหว่างการสอบแบบปรับเหมาะ กับรูปแบบอื่นที่ควรมีค่าสูง ซึ่งการประมาณค่าความสามารถโดยวิธีต่าง ๆ นั้น อาจใช้ได้ดีในสถานการณ์ที่ต่างกัน ในการวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาวิธีการประมาณค่าความสามารถตามวิธีที่ 1 2 และ 3 ซึ่งเป็นวิธีที่ไม่ต้องใช้วิธีการทางเทคนิคซับซ้อนเหมาะสำหรับเสนอแนะให้ผู้สอนนำไปใช้ในการสอบภายในโรงเรียน โดยหลีกเลี่ยงข้อจำกัดที่อาจเกิดขึ้น เช่น การใช้จำนวนชั้นของรูปพีระมิดจำนวน 12 ชั้น การคัดเลือกข้อสอบที่มีค่าสัมประสิทธิ์การเดาไม่เกิน 0.3

คุณภาพของข้อสอบและแบบสอบตามทฤษฎีการสอบแบบดั้งเดิม

ทฤษฎีการสอบแบบดั้งเดิม (Classical Test Theory) เป็นทฤษฎีที่มุ่งตรวจสอบความสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนที่แท้จริง (ศิริชัย กาญจนวาสี, 2550: 35) และวิเคราะห์คุณภาพของข้อสอบแต่ละข้อโดยพิจารณาด้านความยาก อำนาจจำแนก ประสิทธิภาพตัวลอง รวมทั้งตรวจสอบคุณภาพของแบบสอบด้านความตรงและความเที่ยง มีรายละเอียด ดังนี้

คุณภาพของข้อสอบ

ค่าความยาก (p หรือ Level of Difficulty of the Items) คือ สัดส่วนของจำนวนนักเรียนที่ตอบถูกในแต่ละข้อกับจำนวนนักเรียนทั้งหมด ซึ่งค่าความยากนี้จะมีค่าระหว่าง 0 ถึง 1 ถ้าค่าความยากต่ำแสดงว่าข้อสอบข้อนั้นค่อนข้างยาก เนื่องจากมีสัดส่วนของนักเรียนที่ตอบถูกน้อย ถ้าค่าความยากสูงแสดงว่าข้อสอบข้อนั้นค่อนข้างง่าย เนื่องจากมีสัดส่วนของนักเรียนที่ตอบถูกมาก โดยทั่วไปข้อสอบที่มีค่าความยากระหว่าง 0.2 – 0.8 ถือว่าเป็นข้อสอบที่มีความยากง่ายพอเหมาะ (ศิริชัย กาญจนวาสี, 2552: 225)

ค่าอำนาจจำแนก (r หรือ Discrimination Power of the Items) คือ ผลต่างระหว่างสัดส่วนจำนวนนักเรียนที่ตอบถูกในกลุ่มเก่ง กับสัดส่วนจำนวนนักเรียนที่ตอบถูกในกลุ่มอ่อน เช่น กลุ่มเก่ง 10 คน ตอบถูก 9 คน แต่กลุ่มอ่อน 10 คน ตอบถูกเพียง 2 คน ดังนั้นค่าอำนาจจำแนกมีค่าเท่ากับ 0.7 ซึ่งค่านี้จะบ่งชี้ถึงคุณภาพของข้อสอบว่าสามารถจำแนกนักเรียนในกลุ่มสูงออกจากนักเรียนในกลุ่มต่ำ หรือนักเรียนเก่งจากนักเรียนอ่อนได้ดีเพียงใด ถ้าค่าอำนาจจำแนกสูงแสดงว่าข้อสอบข้อนั้นมีคุณภาพในการจำแนกนักเรียนเก่งออกจากนักเรียนอ่อนได้ดี ถ้าค่าอำนาจจำแนกต่ำแสดงว่าข้อสอบข้อนั้นมีคุณภาพค่อนข้างต่ำในการจำแนกนักเรียนเก่งออกจากนักเรียนอ่อน ค่าอำนาจจำแนกนั้นมีค่าอยู่ระหว่าง -1 ถึง +1 แต่อำนาจจำแนกที่ดีจะต้องมีค่าเป็นบวกและควรมีค่าตั้งแต่ 0.2 ขึ้นไป

ประสิทธิภาพตัวลวง (r_w หรือ Item Distractor) คือ ความสามารถของตัวเลือกที่ผิดที่ทำให้ให้นักเรียนในกลุ่มต่ำเลือกตอบมากกว่านักเรียนในกลุ่มสูง และควรเป็นตัวเลือกที่มีนักเรียนเลือกอย่างน้อยร้อยละ 5

คุณภาพของแบบสอบ

ความตรง (Validity) คือ คุณสมบัติของแบบสอบที่สามารถวัดได้ในสิ่งที่ต้องการวัดตามวัตถุประสงค์และสถานการณ์ที่กำหนดไว้ ถ้าแบบสอบมีความตรงจะทำให้ผลที่ได้จากการวัดมีความถูกต้อง สามารถนำคะแนนที่ได้ไปแปลความหมายถึงสิ่งที่ต้องการวัดได้อย่างเหมาะสม สามารถจำแนกความตรงได้เป็น 3 ประเภทหลัก ๆ ได้แก่ (ศิริชัย กาญจนวาสี, 2552: 225)

1. ความตรงตามเนื้อเรื่อง (Content Validity) คือ คุณสมบัติของเครื่องมือที่ใช้วัดความสอดคล้องของเนื้อหาสาระของข้อสอบกับวัตถุประสงค์ของการสอบ การตรวจสอบความตรงตามเนื้อเรื่อง ทำได้โดยการวิเคราะห์เนื้อหาของแบบสอบทั้งฉบับโดยอาศัยผู้เชี่ยวชาญหรือผู้ทรงคุณวุฒิเป็นผู้พิจารณาเนื้อหาสาระของข้อสอบแต่ละข้อว่าสามารถวัดได้ตรงตามวัตถุประสงค์ของการสอบหรือไม่

2. ความตรงตามเกณฑ์สัมพัทธ์ (Criterion-Related Validity) คือ คุณสมบัติของเครื่องมือที่ใช้วัดความสอดคล้องสัมพันธ์กันระหว่างคะแนนที่ได้จากแบบสอบที่สร้างขึ้นกับเกณฑ์ภายนอกที่สามารถใช้วัดลักษณะที่ต้องการวัดนั้นได้ เกณฑ์ภายนอกอาจเป็นสมรรถนะ ผลการปฏิบัติหรือพฤติกรรม การตรวจสอบความตรงตามเกณฑ์สัมพัทธ์ ทำได้โดยการคำนวณค่าสถิติสหสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบกับค่าหรือคะแนนที่ได้จากเกณฑ์ภายนอก

3. ความตรงเชิงทฤษฎี (Construct Validity) คือ คุณสมบัติของแบบสอบที่สามารถวัดหรืออธิบายพฤติกรรมหรือสิ่งที่ต้องการวัดได้ตรงตามทฤษฎี ส่วนใหญ่ความตรงตามทฤษฎีนั้น จะใช้พิจารณาแบบสอบที่วัดนามธรรมและวัดโดยตรงได้ยาก เช่น สถิติปัญญา ความคิดสร้างสรรค์ การตรวจสอบความตรงเชิงทฤษฎีอาจจะใช้เทคนิคของการวิเคราะห์ตัวประกอบ (Factor Analysis)

ความเที่ยง (Reliability) คือ คุณสมบัติของแบบสอบที่วัดได้สม่ำเสมอคงเส้นคงวา วัดกี่ครั้งก็ได้ผลเหมือนเดิม หรือใกล้เคียงกับของเดิมมาก การหาค่าความเที่ยงของแบบสอบอาจทำได้หลายวิธี แต่การวัดความคงที่ภายในจะเป็นการหาค่าความเที่ยงของแบบสอบโดยใช้การสอบเพียงครั้งเดียว ซึ่งสามารถคำนวณหาค่าความเที่ยงได้หลายวิธี ดังนี้

1. วิธีแบ่งครึ่ง (Split-Half Method) โดยนำผลการสอบมาแบ่งเป็นข้อมูล 2 ชุด อาจแบ่งเป็นข้อคู่-ข้อคี่ แบ่งเป็นครึ่งฉบับแรก ครึ่งฉบับหลัง มาหาค่าสัมประสิทธิ์สหสัมพันธ์จะได้สัมประสิทธิ์ความเที่ยงของแบบสอบครึ่งฉบับ แล้วจึงนำไปปรับเป็นสัมประสิทธิ์สหสัมพันธ์ของแบบสอบทั้งฉบับ โดยใช้สูตรของสเปียร์แมน บราวน์ (Spearman Brown) ดังนี้

$$r_{tt} = \frac{2r_{xy}}{1 + r_{xy}}$$

เมื่อ r_{tt} คือ ความเที่ยงของแบบทดสอบทั้งฉบับ

r_{xy} คือ ความเที่ยงของแบบทดสอบครึ่งฉบับ

2. วิธีของ Kuder – Richarson โดยนำผลการสอบมาคำนวณค่าสัมประสิทธิ์จากสูตรของคูเดอร์และริชาร์ดสัน สูตรที่ใช้มี 2 สูตร คือสูตร KR-20 กับสูตร KR-21

สูตร KR – 20

$$r_{tt} = \frac{k}{k-1} \frac{(S_t^2 - \sum pq)}{S_t^2}$$

สูตร KR – 21

$$r_{tt} = \frac{k}{k-1} \left[1 - \frac{\bar{x}(k-\bar{x})}{kS_t^2} \right]$$

เมื่อ k คือ จำนวนข้อของแบบทดสอบ

S_t^2 คือ ความแปรปรวนของคะแนนรวม

p คือ สัดส่วนของผู้ตอบข้อสอบถูกต้องในแต่ละข้อ

q คือ สัดส่วนของผู้ตอบข้อสอบผิดในแต่ละข้อ ($1 - p$)

\bar{x} คือ ค่าเฉลี่ยเลขคณิตของคะแนนสอบ

สูตร KR - 20 และ KR - 21 นี้ใช้ได้เฉพาะการหาความเที่ยงของแบบสอบที่ให้คะแนนแต่ละข้อเป็นแบบตอบผิดได้ 0 ตอบถูกได้ 1 เท่านั้น สูตร KR - 21 ใช้ในกรณีข้อสอบทุกข้อมีค่าความยากเท่ากัน อย่างไรก็ตาม บุญเรียง ขจรศิลป์ (2543: 166) อธิบายว่า ค่าที่คำนวณจากสูตร KR - 21 มักจะน้อยกว่าค่าที่คำนวณได้จากสูตร KR - 20 ทั้งนี้ เพราะ โดยทั่วไปแล้วความยากง่ายของข้อคำถามแต่ละข้อมักไม่เท่ากัน ซึ่งขัดกับข้อตกลงของสูตร KR - 21 ดังนั้นสูตร KR - 21 จึงเหมาะสำหรับการคำนวณความเที่ยงเชิงความสอดคล้องภายในอย่างหายๆ เท่านั้น เพราะมีความสะดวกในการคำนวณมากกว่าสูตร KR - 20

3. วิธีของ Cronbach โดยใช้สูตรของครอนบัก ซึ่งปรับมาจากสูตร KR - 20 สำหรับหาความเที่ยงของแบบสอบที่ให้คะแนนแตกต่างกันไปในแต่ละข้อ โดยไม่จำเป็นต้องเป็นแบบตอบผิดได้ 0 ตอบถูกได้ 1 สูตรการคำนวณเป็นดังนี้

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum S_i^2}{S_t^2} \right)$$

เมื่อ α คือ ค่าความเที่ยง หรือ สัมประสิทธิ์แอลฟา

k คือ จำนวนข้อสอบ

S_i^2 คือ ความแปรปรวนของคะแนนแต่ละข้อ

S_t^2 คือ ความแปรปรวนของคะแนนทั้งฉบับ

ค่าความเที่ยงโดยใช้สูตรสัมประสิทธิ์แอลฟา สามารถหาได้โดยใช้ผลการสอบจากแบบสอบชุดเดียวที่นำไปสอบกับนักเรียนกลุ่มเดียว และวิธีนี้มีการใช้กันอย่างกว้างขวางโดยไม่จำกัดเฉพาะแบบสอบที่ให้คะแนนแบบ 0 กับ 1 ซึ่งในกรณีที่แบบสอบให้คะแนนแบบ 0 กับ 1 การคำนวณค่าความเที่ยงด้วยวิธีนี้จะให้ค่าความเที่ยงเท่ากับการคำนวณด้วยสูตร KR - 20 (บุญเรียง ขจรศิลป์, 2543: 168)

ในการวิจัยครั้งนี้ผู้วิจัยวางแผนในการนำข้อสอบไปทดลองใช้ครั้งที่ 1 แล้วนำผลที่ได้มาวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการสอบแบบดั้งเดิมด้วยโปรแกรม EVANA ใช้เทคนิค 27 %

ในการจำแนกนักเรียนเป็นกลุ่มเก่งและกลุ่มอ่อน เนื่องจากมีนักเรียนที่ใช้ในการสอบจำนวน 150 คน (บุญเรียง ขจรศิลป์, 2543: 114) และใช้เกณฑ์ในการคัดเลือกข้อสอบที่มีค่าความยากระหว่าง 0.2 – 0.8 และมีค่าอำนาจจำแนกตั้งแต่ 0.2 ขึ้นไป และรายงานค่าความเที่ยงด้วยสูตร KR – 20 เมื่อคำนวณด้วยโปรแกรม EVANA และ สูตรสัมประสิทธิ์แอลฟา เมื่อคำนวณด้วยโปรแกรม SPSS Version 12

งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการสอบแบบปรับเหมาะสมส่วนใหญ่จะเป็นการสร้างแบบสอบปรับเหมาะสม เพื่อใช้ทดสอบในวิชาต่าง ๆ สำหรับงานวิจัยที่เกี่ยวข้องกับการประมาณค่าความสามารถ ผู้วิจัยได้รวบรวมไว้ ดังนี้

งานวิจัยภายในประเทศ

อรินทร์ น่วมถนอม (2537) ได้ศึกษาความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบสอบดั้งเดิม แบบสอบสองขั้นตอน และแบบสอบเฟล็กซิเลเวล โดยศึกษากับกลุ่มตัวอย่างกับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ผลการวิจัยพบว่า ค่าความเที่ยงตรงของการประมาณค่าความสามารถระหว่างแบบสอบดั้งเดิม กับแบบสอบสองขั้นตอน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนค่าความเที่ยงตรงของการประมาณค่าความสามารถระหว่างแบบสอบดั้งเดิมกับแบบสอบเฟล็กซิเลเวล และแบบสอบสองขั้นตอนกับแบบสอบเฟล็กซิเลเวลมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของนักเรียนระหว่างแบบสอบดั้งเดิมกับแบบสอบสองขั้นตอน และแบบสอบดั้งเดิมกับแบบสอบเฟล็กซิเลเวลมีค่าสูงกว่า 1 ในกลุ่มนักเรียนที่มีความสามารถระดับต่ำมาก ระดับปานกลาง และระดับสูงมาก แต่มีค่าต่ำกว่า 1 ในกลุ่มนักเรียนที่มีระดับความสามารถต่ำและระดับสูง

ชูใจ ทรงเมฆ (2538) ได้ศึกษาการเปรียบเทียบผลการประมาณค่าความสามารถด้วยวิธีการทดสอบแบบปรับเหมาะสม และวิธีการทดสอบแบบดั้งเดิม ผลการวิจัยพบว่าค่าประมาณความสามารถจากการสอบแบบปรับเหมาะสม และการสอบแบบดั้งเดิม ไม่มีนัยสำคัญของความแตกต่างทางสถิติที่ระดับ .05 จึงสรุปได้ว่าการใช้วิธีการสอบแบบดั้งเดิมและแบบปรับเหมาะสมให้ผลการประมาณค่าความสามารถในด้านผลสัมฤทธิ์ทางการเรียนที่ไม่แตกต่างกัน

วสันต์ ทองไทย (2539) ได้ศึกษาการเปรียบเทียบผลการประมาณค่าความสามารถด้วยวิธีการทดสอบแบบปรับเหมาะรูปปิรามิดขนาดชั้นคงที่และรูปปิรามิดข้างตัด: การทดสอบทักษะกระบวนการทางวิทยาศาสตร์ พบว่า การทดสอบทั้ง 2 วิธี ให้ค่าประมาณความสามารถที่ไม่แตกต่างกัน นอกจากนี้ได้ทำการทดสอบความแตกต่างของค่าสารสนเทศและค่าความเที่ยงของการประมาณค่าความสามารถของการทดสอบแบบรูปปิรามิดขนาดชั้นคงที่ และรูปปิรามิดข้างตัดโดยใช้ค่าดัชนีร้อยละของความแตกต่าง พบว่า การสอบทั้งสองวิธีมีค่าสารสนเทศ และค่าความเที่ยงของการประมาณค่าความสามารถไม่แตกต่างกัน ดังนั้นจึงสรุปได้ว่าการสอบแบบปรับเหมาะรูปปิรามิดขนาดชั้นคงที่และรูปปิรามิดข้างตัดให้ผลการประมาณค่าความสามารถในด้านทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 6 ไม่แตกต่างกัน

อัศวิน คงเพชรศักดิ์ (2547) ได้ศึกษาความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของนักเรียนโดยใช้แบบสอบปรับเหมาะรูปพีระมิดในวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 พบว่าค่าความเที่ยงตรงของการประมาณค่าความสามารถของแบบสอบปรับเหมาะรูปพีระมิดที่ใช้ขนาดชั้นคงที่จำนวน 5 ชั้น 15 ชั้น และ ที่มีหลายข้อในแต่ละชั้นจำนวน 5 ชั้น มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .10 ยกเว้นค่าความเที่ยงของการประมาณค่าความสามารถระหว่างรูปแบบที่ใช้ขนาดชั้นคงที่ จำนวน 15 ชั้น กับรูปแบบที่มีหลายข้อในแต่ละชั้นจำนวน 5 ชั้น ที่ตรวจให้คะแนนโดยพิจารณาค่าความยากตามข้อสอบที่ถัดจากข้อสุดท้ายที่ได้ตอบ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และการประมาณค่าความสามารถของรูปแบบที่ใช้ขนาดชั้นคงที่จำนวน 15 ชั้น มีค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถสูงสุดในกลุ่มนักเรียนทุกระดับความสามารถ สำหรับรูปแบบที่ใช้ขนาดชั้นคงที่จำนวน 5 ชั้นจะมีค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของแบบสอบสูงกว่ารูปแบบที่มีข้อสอบหลายข้อในแต่ละชั้นจำนวน 5 ชั้น ในกลุ่มนักเรียนที่มีความสามารถระดับปานกลาง และระดับต่ำ แต่ในกลุ่มนักเรียนที่มีความสามารถระดับต่ำมาก ระดับสูงและระดับสูงมาก ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของรูปแบบที่ใช้ขนาดคงที่จำนวน 5 ชั้น มีค่าต่ำกว่ารูปแบบที่มีข้อสอบหลายข้อในแต่ละชั้นจำนวน 5 ชั้น

งานวิจัยภายในประเทศข้างต้น ได้ศึกษาเกี่ยวกับความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถที่มีรูปแบบของแบบสอบปรับเหมาะแตกต่างกัน รวมทั้งศึกษาเกี่ยวกับการเปรียบเทียบผลการประมาณค่าความสามารถระหว่างแบบสอบรูปแบบต่าง ๆ แต่ยังไม่มียานวิจัยที่ศึกษาเกี่ยวกับความตรงสัมพัทธ์ของค่าประมาณความสามารถด้านทักษะกระบวนการ

ทางวิทยาศาสตร์ชั้นบูรณาการจากการสอบแบบปรับเหมาะรูปพีระมิดที่มีวิธีประมาณค่าความสามารถแตกต่างกัน

งานวิจัยต่างประเทศ

Lakin and Weiss (1975 อ้างใน วสันต์ ทองไทย, 2539: 37-38) ได้ศึกษาถึงความสัมพันธ์ระหว่างการสอบด้วยรูปพีระมิด กับวิธีการประมาณค่าความสามารถ และการสอบแบบดั้งเดิม โดยศึกษาจากการสอบรูปพีระมิด 3 ฉบับ เป็นชนิด 15 ชั้น และใช้วิธีการคงที่ทั้ง 3 ฉบับ ซึ่งค่าความแตกต่างระหว่างค่าความยากของข้อสอบในชั้นที่อยู่ติดกันคงที่ และใช้กฎการแยกทาง แต่ละชุดมีจำนวนข้อสอบ 120 ข้อ ใช้วิธีการประมาณค่าความสามารถ 6 วิธี คือ 1) ให้ตามจำนวนข้อที่ถูก 2) ให้ตามค่าความยากเฉลี่ยของข้อสอบที่ทำ 3) ให้ค่าความยากเฉลี่ยของข้อสอบที่ตอบถูก 4) ให้ตามค่าความยากของข้อสอบข้อสุดท้ายที่ทำ 5) ให้ตามค่าความยากของข้อสอบทุกข้อกับการสอบแบบดั้งเดิม 1 ฉบับ ซึ่งมีข้อสอบจำนวน 40 ข้อ และได้ปรับค่าการเดาแล้วแบบสอบทุกฉบับเป็นแบบที่เลือกตอบชนิด 5 ตัวเลือก ถ้ามุ่งเกี่ยวกับคำศัพท์ที่อยู่ในระดับมหาวิทยาลัย ซึ่งคัดเลือกมาจากกลุ่มข้อสอบจำนวน 369 ข้อ และได้ดำเนินการสอบโดยใช้เครื่องคอมพิวเตอร์ โดยแบ่งการสอบออกเป็น 2 ระยะ คือ ระยะแรก แบ่งกลุ่มตัวอย่างออกเป็น 8 กลุ่ม โดยกลุ่มที่ 1 และ กลุ่มที่ 2 ซึ่งมีจำนวนกลุ่มละ 125 คน ทำการสอบรูปพีระมิดฉบับที่ 1 และการสอบแบบดั้งเดิม ส่วนกลุ่มที่ 3 ซึ่งมีจำนวน 142 คน ทำการสอบรูปพีระมิดฉบับที่ 3 และการสอบปรับระดับชั้น ระยะที่สอง ทิ้งช่วงให้ห่างจากระยะแรก เป็นเวลา 7 สัปดาห์ สุ่มกลุ่มตัวอย่างจากกลุ่มที่ 1 และ กลุ่มที่ 2 จำนวน 112 คน ให้ทำการสอบรูปพีระมิดฉบับที่ 3 และแบบสอบสองชั้นตอน ผลการศึกษาปรากฏว่าคะแนนที่ได้จากการสอบรูปพีระมิดกับการสอบแบบดั้งเดิม อันเนื่องมาจากการแบ่งกลุ่มในการดำเนินการสอบก่อนหรือหลังไม่แตกต่างกัน

Bizot *et al.* (1994) ได้ศึกษาผลการใช้รูปแบบการสอบและค่าพารามิเตอร์จากการสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) โดยมีจุดมุ่งหมายเพื่อประเมินโมเดลตามทฤษฎีการตอบสนองข้อสอบ (IRT) ระหว่างโมเดลแบบสองพารามิเตอร์กับสามพารามิเตอร์ ซึ่งใช้สำหรับประมาณค่าความสามารถเริ่มต้นการสอบและเกณฑ์การยุติการสอบ โดยใช้ข้อสอบคำศัพท์จากบทสนทนาจำนวน 80 ข้อ กับกลุ่มตัวอย่าง 3 กลุ่ม คือ กลุ่มนักเรียนชั้นมัธยมศึกษาตอนปลาย กลุ่มนักเรียนชั้นมัธยมศึกษาตอนต้น และกลุ่มนักเรียนชั้นมัธยมศึกษาตอนต้นและตอนปลายจากกลุ่มแรกและกลุ่มที่สองเพื่อทำการสอบซ้ำ วิธีการประมาณค่าความสามารถเริ่มแรกนั้นใช้ 2 วิธี คือ การประมาณค่าความสามารถแบบสุ่มและการประมาณค่าความสามารถที่ได้จากการคำนวณด้วยโมเดล

แบบสองพารามิเตอร์กับสามพารามิเตอร์ ทำการเก็บข้อมูลกับกลุ่มตัวอย่างทั้งหมด 2,697 คน ผลการศึกษาพบว่า รูปแบบการสอบและค่าพารามิเตอร์จากการสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) แตกต่างกัน โดยโมเดลตามทฤษฎีการตอบสนองข้อสอบ (IRT) แบบสามพารามิเตอร์มีความเหมาะสมที่สุดทั้งในเชิงทฤษฎีและเชิงปฏิบัติการ

Bryson (n.d. อ้างใน อศวิน กงเพชรศักดิ์, 2547: 47) ได้ศึกษาเปรียบเทียบค่าสหสัมพันธ์ระหว่างแบบสอบพีระมิดชนิด 5 ชั้น จำนวน 2 ฉบับ และแบบสอบดั้งเดิมที่มีข้อสอบ 5 ข้อ จำนวน 2 ฉบับ กับคะแนนผลการสอบแบบสอบฉบับหนึ่งที่มีข้อสอบจำนวน 100 ข้อ แบบสอบดั้งเดิมสอบโดยการเขียนตอบ ส่วนแบบสอบพีระมิดสอบด้วยคอมพิวเตอร์ แบบสอบพีระมิดฉบับการเลือกตอบให้นักเรียนใช้ค่าอำนาจจำแนกเป็นเกณฑ์ ส่วนอีกฉบับใช้การพยากรณ์คะแนนสูงสุดเป็นเกณฑ์ นำแบบสอบทั้งสองฉบับนี้ และแบบสอบดั้งเดิมไปทดลองกับกลุ่มตัวอย่าง พบว่าแบบสอบดั้งเดิมฉบับหนึ่งมีค่าสหสัมพันธ์กับคะแนนรวมสูงกว่าแบบสอบพีระมิดทั้งสองฉบับ และแบบสอบพีระมิดฉบับหนึ่งมีค่าสหสัมพันธ์กับคะแนนรวมต่ำกว่าแบบสอบดั้งเดิมทั้งสองฉบับ

Wen, Chang, and Hau (2001) ได้ศึกษาจำนวนชั้นที่เหมาะสมของวิธีการสอบแบบปรับเหมาะที่ดำเนินการด้วยคอมพิวเตอร์ตามระดับชั้นของค่าอำนาจจำแนก จากข้อสังเกตว่า ถ้าจำนวนชั้นน้อยที่สุดเท่ากับ 1 การเลือกข้อสอบจากค่าความยาก (b) ให้ใกล้เคียงกับค่าประมาณความสามารถ ให้ประสิทธิภาพน้อยกว่าวิธีสารสนเทศสูงสุด แต่ถ้าจำนวนชั้นมากเกินไปอาจไม่มีจำนวนข้อสอบที่มีความยากระดับต่าง ๆ เพียงพอในแต่ละชั้น การศึกษาทำจากข้อสอบ 200 400 และ 800 ข้อ ความยาวแบบสอบเท่ากับ 48 ข้อ จำนวนชั้นของข้อสอบ คือ 1 2 3 4 6 8 12 16 24 และ 48 ชั้น ผลการศึกษาพบว่า ประสิทธิภาพการประมาณค่าความสามารถของผู้สอบและสมมูลการใช้ข้อสอบจะไม่เพิ่มขึ้นหรือลดลงตามจำนวนชั้น จำนวนชั้นที่ใช้ขึ้นอยู่กับวิธีการประยุกต์ โครงสร้างของข้อสอบ ความยาวแบบสอบ และเงื่อนไขในการสอบอื่นๆ

Widiatmo (2004: 1336-A) ศึกษาการออกแบบแบบทดสอบที่ดำเนินการสอบด้วยคอมพิวเตอร์ในการจำลองสถานการณ์และประเมินความสามารถทางสติปัญญาด้านภาษาของนักเรียน พบว่า การออกแบบเพื่อสร้างแบบสอบวัดความสามารถทางสติปัญญาที่ดำเนินการสอบด้วยคอมพิวเตอร์โดยใช้ข้อสอบทั้งฉบับครบทุกข้อกับการลดจำนวนข้อสอบต่างก็มีค่าความลำเอียงของข้อสอบและค่าความคลาดเคลื่อนมาตรฐานในการวัดอยู่ในระดับต่ำ แต่มีประสิทธิภาพเชิงสัมพันธ์สูงกว่าแบบทดสอบแบบเดิมที่ใช้กระดาษ - ดินสอ (Paper - Pencil Test) และแบบทดสอบ

ที่ดำเนินการสอบด้วยคอมพิวเตอร์ทุกฉบับมีประสิทธิภาพในการวัดและมีการใช้กลุ่มข้อสอบสูงกว่าการสอบด้วยกระดาษ - ดินสอ (Paper - Pencil Test) อย่างน้อย 40 - 60%

งานวิจัยต่างประเทศข้างต้น ได้ศึกษาเกี่ยวกับประสิทธิภาพของการสอบแบบปรับเหมาะ ซึ่งใช้วิธีประมาณค่าความสามารถ โมเดลตามทฤษฎีการตอบสนองข้อสอบ (IRT) และรูปแบบการสอบที่แตกต่างกัน โดยแบบสอบปรับเหมาะได้สร้างจากโปรแกรมทางคอมพิวเตอร์ซึ่งสามารถเลือกข้อสอบแต่ละข้อตามเกณฑ์ต่าง ๆ ที่ผู้สร้างกำหนดขึ้นและยุติการสอบอย่างอัตโนมัติเมื่อการประมาณค่าความสามารถที่ได้มีความคลาดเคลื่อนต่ำ ซึ่งต้องใช้ค่าใช้จ่ายสูงในการสร้างโปรแกรมทางคอมพิวเตอร์ให้มีลักษณะดังกล่าว

จากการศึกษาเอกสารงานวิจัยทั้งในและต่างประเทศ ผู้วิจัยสนใจสร้างแบบสอบปรับเหมาะรูปพีระมิดข้างตัดด้วยโปรแกรมสำเร็จรูปทางคอมพิวเตอร์ที่แพร่หลายและศึกษาความสัมพันธ์ของการประมาณค่าความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการจากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดที่มีวิธีการประมาณค่าความสามารถต่างกัน 3 วิธี คือ วิธีการประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก ค่าเฉลี่ยความยากของข้อที่ได้ทำ และค่าความยากของข้อสุดท้ายที่ได้ตอบ เมื่อใช้ค่าความสามารถที่ได้จากการสอบแบบดั้งเดิมซึ่งประมาณค่าความสามารถด้วยโปรแกรม BILOG เป็นเกณฑ์อ้างอิง

ในการวิจัยครั้งนี้ผู้วิจัยได้สร้างและพัฒนาคุณภาพแบบสอบตามทฤษฎีการสอบดั้งเดิมโดยคัดเลือกข้อสอบที่มีค่าความยากระหว่าง 0.2 – 0.8 และมีค่าอำนาจจำแนกตั้งแต่ 0.2 ขึ้นไป พัฒนาคุณภาพแบบสอบ ประเมินค่าพารามิเตอร์ของข้อสอบและค่าความสามารถของนักเรียนด้วยโมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) ด้วยโปรแกรม BILOG โดยเลือกข้อสอบที่มีค่าความยาก -3.00 ถึง +3.00 ค่าอำนาจจำแนกสูง ๆ และค่าสัมประสิทธิ์การเดามีค่าต่ำกว่า 0.30 สำหรับนำมาออกแบบการสอบแบบปรับเหมาะ และทดสอบสมมติฐานโดยจัดการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดและแบบดั้งเดิมเพื่อศึกษาความสัมพันธ์ของค่าประมาณความสามารถจากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดโดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ โดยหาค่าความสัมพันธ์กับค่าประมาณความสามารถจากการสอบแบบดั้งเดิม ซึ่งใช้เป็นเกณฑ์ในการเปรียบเทียบ

สมมติฐานการวิจัย

ความสัมพันธ์ระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดโดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) มีความสัมพันธ์กัน



บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ ผู้วิจัยมีวัตถุประสงค์เพื่อศึกษาความตรงสัมพัทธ์ของค่าประมาณความสามารถ 3 วิธี จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด ผู้วิจัยได้ใช้ระเบียบวิธีวิจัยเชิงประจักษ์และดำเนินการตามหัวข้อต่าง ๆ ดังนี้

1. พลวิจัย
2. เครื่องมือที่ใช้ในการวิจัย
3. การสร้างและพัฒนาแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ
4. การออกแบบการสอบแบบปรับเหมาะและการสอบแบบดั้งเดิม
5. การเก็บรวบรวมข้อมูล
6. การวิเคราะห์ข้อมูล

พลวิจัย

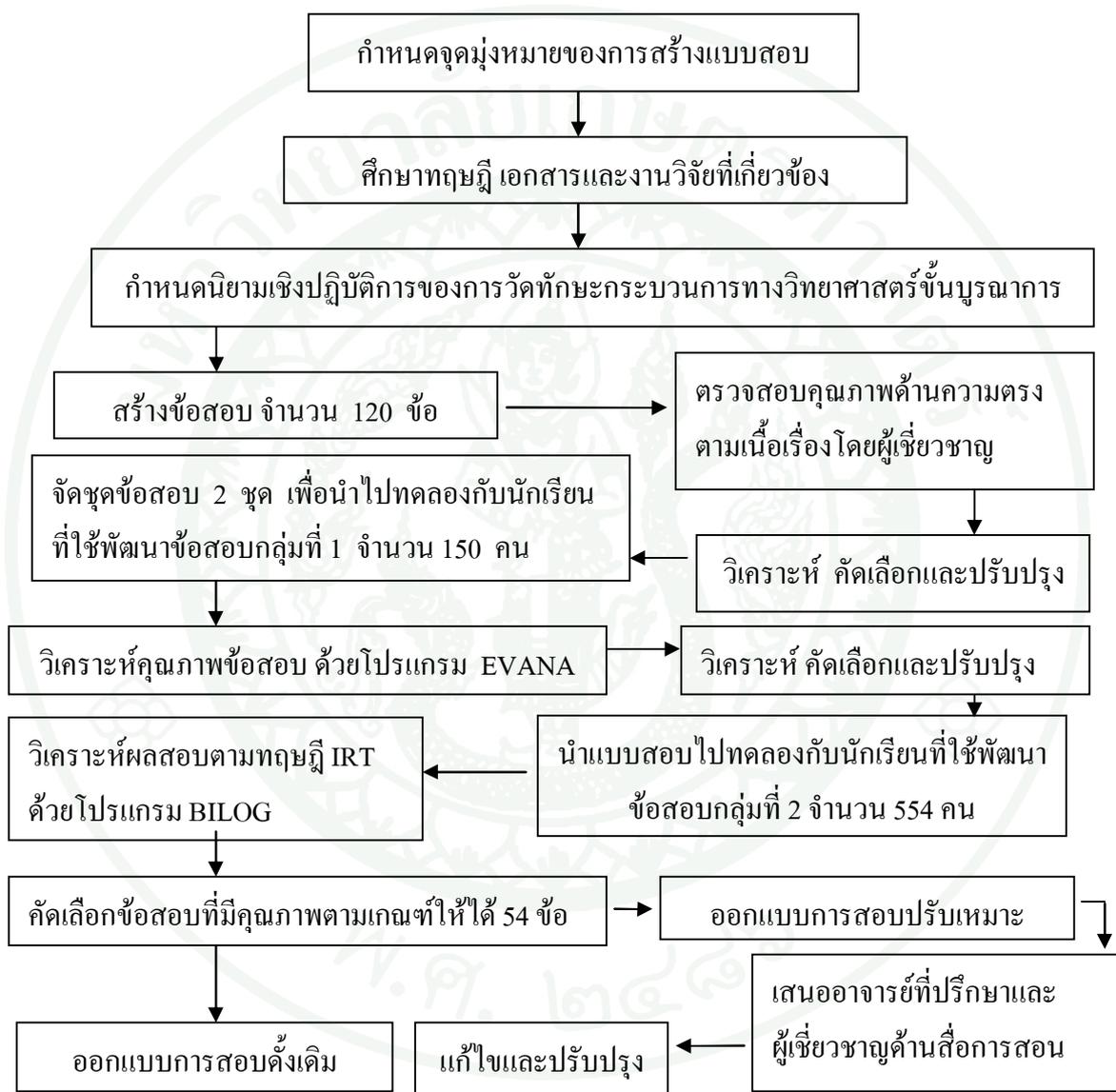
พลวิจัย ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนแผนวิทยาศาสตร์-คณิตศาสตร์ ภาคเรียนที่ 2 ปีการศึกษา 2552 โรงเรียนเศรษฐบุตรบำเพ็ญ สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กรุงเทพมหานคร เขต 2 จำนวน 127 คน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ แบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ผู้วิจัยเป็นผู้สร้างขึ้น เป็นแบบสอบเลือกตอบ 4 ตัวเลือก โดยใช้เนื้อหาตามสาระการเรียนรู้วิทยาศาสตร์

การสร้างและพัฒนาเครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ แบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ วิธีการดำเนินการสร้างและพัฒนาแบบสอบเป็นไปตามขั้นตอน ดังแสดงในภาพ



ภาพที่ 11 ลำดับขั้นตอนการสร้างและพัฒนาแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

รายละเอียดขั้นตอนการสร้างและพัฒนาแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์
ขั้นบูรณาการ

1. กำหนดจุดมุ่งหมายในการสร้างแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้น
บูรณาการ

2. กำหนดนิยามเชิงปฏิบัติการของความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์
ขั้นบูรณาการทั้ง 5 ทักษะ

3. สร้างข้อสอบแบบเลือกตอบ 4 ตัวเลือก ตามกรอบนิยามเชิงปฏิบัติการที่กำหนดไว้
ทักษะละ 24 ข้อ โดยให้แต่ละทักษะมีระดับความยาก - ง่ายแตกต่างกันรวมทั้งสิ้น 120 ข้อ

4. เสนออาจารย์ที่ปรึกษาวิทยานิพนธ์และผู้เชี่ยวชาญที่มีวุฒิการศึกษาไม่ต่ำกว่าปริญญาโท
โดยแบ่งเป็นอาจารย์สอนในคณะศึกษาศาสตร์ มหาวิทยาลัยของรัฐ จำนวน 2 ท่าน นักวิชาการ
จากสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี จำนวน 2 ท่าน และ อาจารย์ผู้สอนใน
กลุ่มสาระการเรียนรู้วิทยาศาสตร์ระดับมัธยมศึกษา จำนวน 6 ท่าน เพื่อตรวจสอบความตรงตาม
เนื้อเรื่อง (Content Validity) ความถูกต้องของเฉลย จากดัชนีความสอดคล้อง (IOC) และคัดเลือก
ข้อที่มีค่า IOC ตั้งแต่ 0.7 ขึ้นไป จากสูตร

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

เมื่อ IOC คือ ดัชนีความสอดคล้องของข้อคำถามกับทักษะกระบวนการที่ต้องการวัด

$\sum R$ คือ ผลรวมของคะแนนความคิดเห็นของผู้เชี่ยวชาญทั้งหมด

N คือ จำนวนผู้เชี่ยวชาญ

5. วิเคราะห์รายข้อ คัดเลือกข้อที่อยู่ในเกณฑ์และปรับปรุงแบบสอบตามข้อเสนอแนะของ
ผู้เชี่ยวชาญ ผลการคัดเลือกจากค่าดัชนีความสอดคล้อง (IOC) ที่มีค่าตั้งแต่ 0.7 ขึ้นไปได้ข้อสอบ
จำนวน 89 ข้อ แบ่งตามทักษะต่าง ๆ ได้ ดังนี้

ทักษะการตั้งสมมติฐาน	จำนวน	15	ข้อ
ทักษะการกำหนดนิยามเชิงปฏิบัติการ	จำนวน	17	ข้อ
ทักษะการกำหนดและควบคุมตัวแปร	จำนวน	18	ข้อ
ทักษะการทดลอง	จำนวน	17	ข้อ
ทักษะการตีความหมายข้อมูลและลงข้อสรุป	จำนวน	22	ข้อ

6. จัดแบบสอบเป็น 2 ชุด โดยให้แต่ละชุดวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการครบทั้ง 5 ทักษะ โดยชุดที่ 1 มีจำนวนข้อสอบ 44 ข้อ และ ชุดที่ 2 จำนวน 45 ข้อ

7. นำแบบสอบไปทดสอบกับนักเรียนที่ใช้ในการพัฒนาข้อสอบกลุ่มที่ 1 จำนวน 150 คน ซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 ที่เรียนแผนวิทยาศาสตร์-คณิตศาสตร์ โรงเรียนเศรษฐบุตรบำเพ็ญ โดยนักเรียนแต่ละคนจะได้แบบสอบ 1 ชุดและผู้วิจัยเปิดโอกาสให้นักเรียนถามเพื่อความชัดเจนของภาษา แล้วตรวจให้คะแนนตามเกณฑ์ดังนี้ ตอบถูกได้ 1 ตอบผิดได้ 0 คะแนน แล้วหาค่าความยาก (p) ค่าอำนาจจำแนก (r) และพิจารณาประสิทธิภาพตัวลง (r_w) ของข้อสอบรายข้อตามทฤษฎีการสอบดั้งเดิม ด้วยโปรแกรม EVANA หลังจากนั้นนำผลมาปรับปรุงข้อคำถามตัวเลือก และ ตัวลง

8. วิเคราะห์ คัดเลือกและปรับปรุงแก้ไขข้อสอบ โดยคัดเลือกข้อที่มีค่าความยาก (p) 0.2 - 0.8 และค่าอำนาจจำแนก (r) มากกว่า 0.2 ได้ข้อสอบจำนวน 76 ข้อ จากข้อสอบ 2 ชุด มีรายละเอียด ดังนี้

ชุดที่ 1 จำนวน 38 ข้อ

ทักษะการตั้งสมมติฐาน	จำนวน	4	ข้อ
ค่าความยาก (p)	0.53	ถึง	0.79
ค่าอำนาจจำแนก (r)	0.26	ถึง	0.42
ทักษะการกำหนดนิยามเชิงปฏิบัติการ	จำนวน	8	ข้อ
ค่าความยาก (p)	0.26	ถึง	0.74
ค่าอำนาจจำแนก (r)	0.21	ถึง	0.58

ทักษะการกำหนดและควบคุมตัวแปร	จำนวน	8	ข้อ
ค่าความยาก (p)	0.34	ถึง	0.84
ค่าอำนาจจำแนก (r)	0.26	ถึง	0.74

ทักษะการทดลอง	จำนวน	6	ข้อ
ค่าความยาก (p)	0.26	ถึง	0.82
ค่าอำนาจจำแนก (r)	0.32	ถึง	0.58

ทักษะการตีความหมายข้อมูลและลงข้อสรุป	จำนวน	12	ข้อ
ค่าความยาก (p)	0.26	ถึง	0.79
ค่าอำนาจจำแนก (r)	0.26	ถึง	0.63

ชุดที่ 2 จำนวน 38 ข้อ

ทักษะการตั้งสมมติฐาน	จำนวน	8	ข้อ
ค่าความยาก (p)	0.31	ถึง	0.78
ค่าอำนาจจำแนก (r)	0.22	ถึง	0.50

ทักษะการกำหนดนิยามเชิงปฏิบัติการ	จำนวน	5	ข้อ
ค่าความยาก (p)	0.25	ถึง	0.75
ค่าอำนาจจำแนก (r)	0.33	ถึง	0.56

ทักษะการกำหนดและควบคุมตัวแปร	จำนวน	6	ข้อ
ค่าความยาก (p)	0.22	ถึง	0.83
ค่าอำนาจจำแนก (r)	0.28	ถึง	0.56

ทักษะการทดลอง	จำนวน	10	ข้อ
ค่าความยาก (p)	0.39	ถึง	0.75
ค่าอำนาจจำแนก (r)	0.22	ถึง	0.61

ทักษะการตีความหมายข้อมูลและลงข้อสรุป	จำนวน 9	ข้อ
ค่าความยาก (p)	0.53	ถึง 0.75
ค่าอำนาจจำแนก (r)	0.28	ถึง 0.50

9. นำข้อสอบที่ปรับปรุงแล้วทั้ง 2 ชุด ไปทดสอบกับนักเรียนที่ใช้ในการพัฒนาข้อสอบ กลุ่มที่ 2 จำนวน 554 คน ซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 ที่เรียนแผนวิทยาศาสตร์ - คณิตศาสตร์ ในโรงเรียนสตรีศรีนครปฐมเพื่อญ โรงเรียนนวมินทราชินูทิศ สตรีวิทยา 2 โรงเรียนฤทธิยะวรรณาลัย 2 โรงเรียนนวมินทราชินูทิศ เบญจมราชาลัย และโรงเรียนพรตพิทยพยัต สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กรุงเทพมหานคร เขต 2 โดยเว้นระยะการทำข้อสอบแต่ละชุดประมาณ 1 สัปดาห์

10. นำผลการตอบมาวิเคราะห์หาค่าพารามิเตอร์การตอบสนองข้อสอบ (IRT) โมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์ ด้วยโปรแกรม BILOG เพื่อหาค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าสัมประสิทธิ์การเดา (c) ของข้อสอบแต่ละข้อ เพื่อเก็บเข้าคลังข้อสอบ

การออกแบบการสอบแบบปรับเหมาะและการสอบแบบดั้งเดิม

1. คัดเลือกข้อสอบที่มีค่าความยากประมาณ -3.00 ถึง +3.00 ค่าอำนาจจำแนกสูง ๆ และค่าสัมประสิทธิ์การเดาไม่เกิน 0.30 ที่ได้จากการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) ด้วยโปรแกรม BILOG จำนวน 54 ข้อ โดยพิจารณาจากแบบจำลองโครงสร้างแบบสอบปรับเหมาะรูปพีระมิดข้างตัด ดังในภาพที่ 8 โดยให้มีค่าความยากของข้อสอบในแนวตั้งมีค่าใกล้เคียงกัน ได้ผลการคัดเลือก ดังนี้

ทักษะการตั้งสมมติฐาน	จำนวน 8	ข้อ
ค่าอำนาจจำแนก (a)	0.49	ถึง 0.91
ค่าความยาก (b)	-0.52	ถึง 1.44
ค่าการเดา (c)	0.00	ถึง 0.19

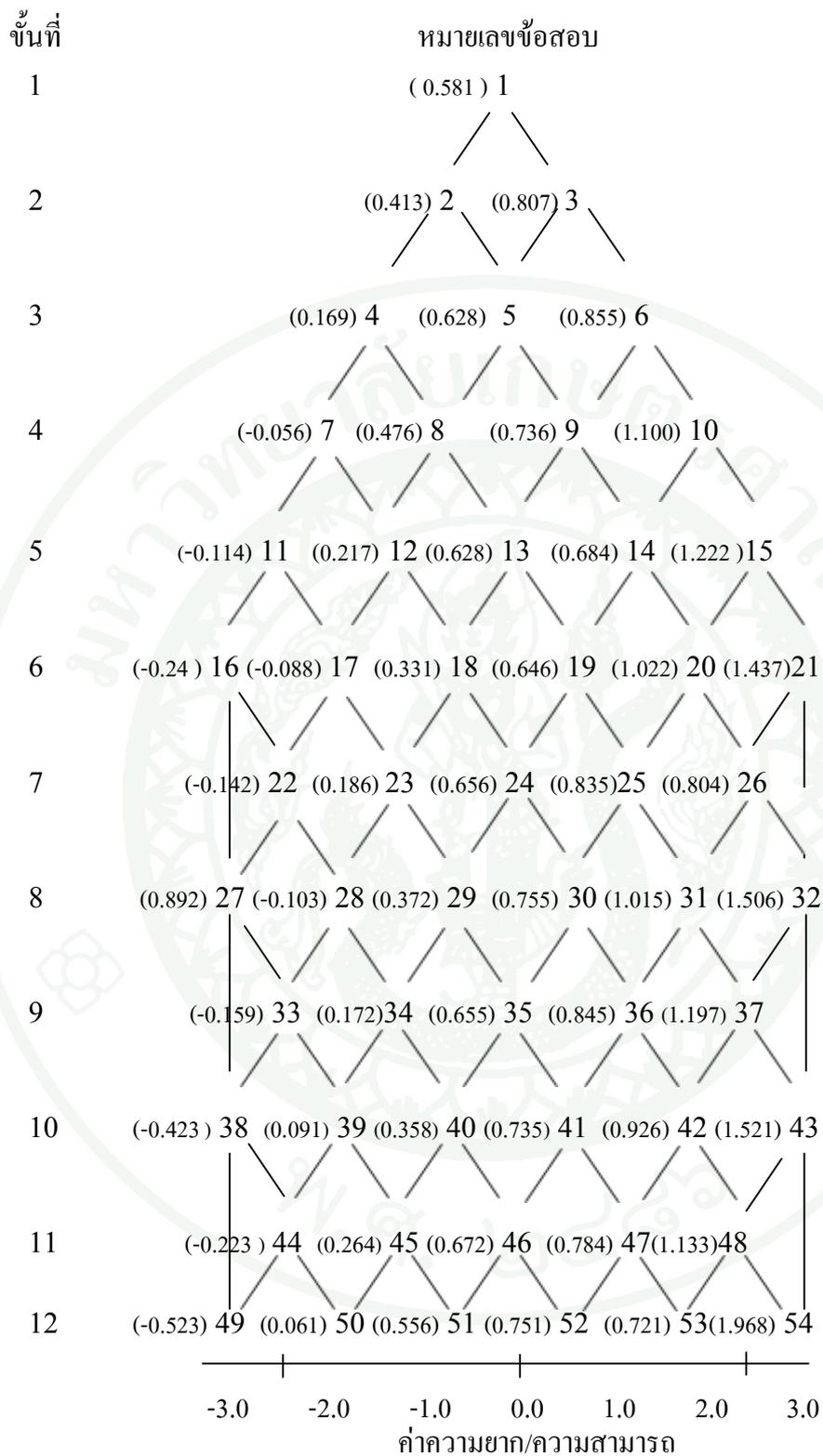
ทักษะการกำหนดนิยามเชิงปฏิบัติการ	จำนวน	9	ข้อ
ค่าอำนาจจำแนก (a)		0.35	ถึง 2.19
ค่าความยาก (b)		-0.32	ถึง 1.52
ค่าการเดา (c)		0.00	ถึง 0.24

ทักษะการกำหนดและควบคุมตัวแปร	จำนวน	12	ข้อ
ค่าอำนาจจำแนก (a)		0.36	ถึง 2.60
ค่าความยาก (b)		-0.42	ถึง 1.23
ค่าการเดา (c)		0.001	ถึง 0.075

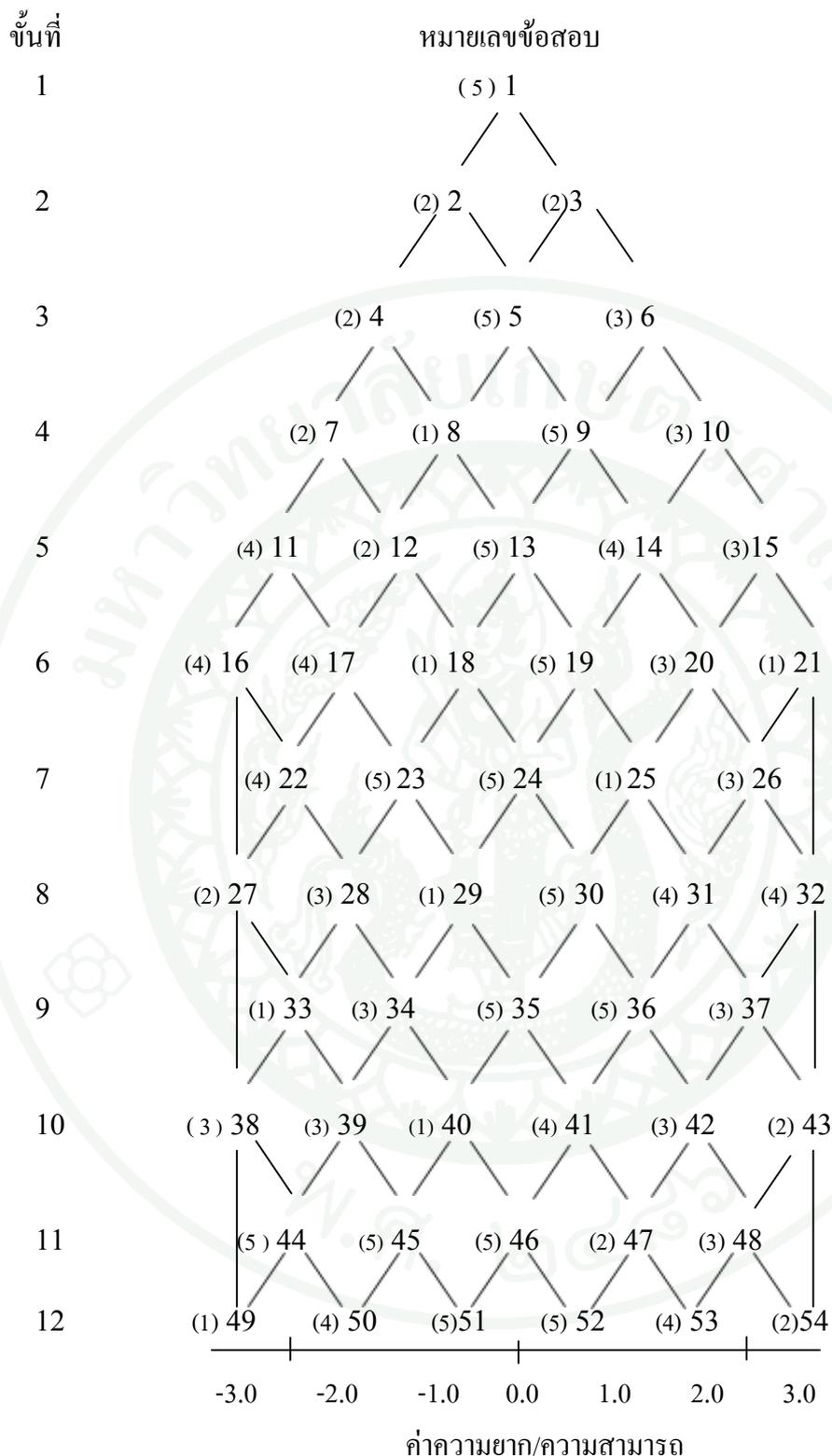
ทักษะการทดลอง	จำนวน	10	ข้อ
ค่าอำนาจจำแนก (a)		0.41	ถึง 2.58
ค่าความยาก (b)		-0.24	ถึง 1.51
ค่าการเดา (c)		0.001	ถึง 0.260

ทักษะการตีความหมายข้อมูลและลงข้อสรุป	จำนวน	15	ข้อ
ค่าอำนาจจำแนก (a)		0.20	ถึง 3.39
ค่าความยาก (b)		-0.22	ถึง 0.85
ค่าการเดา (c)		0.00	ถึง 0.25

2. วางแผนผังเพื่อสร้างแบบสอบปรับเหมาะรูปพีระมิดข้างตัด จำนวน 12 ชั้น จากการพิจารณาระดับความยากของข้อสอบในแนวคิดให้มีค่าใกล้เคียงกันจะได้แบบสอบประกอบด้วยข้อสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการทั้ง 5 ทักษะ จำนวน 54 ข้อ ดังแสดงในภาพที่ 12 โดยแต่ละข้อมีค่าความยากตามตัวเลขในวงเล็บด้านซ้าย และภาพที่ 13 แสดงทักษะที่ต้องการวัดในข้อนั้น ๆ



ภาพที่ 12 แบบจำลองโครงสร้างแบบสอบปรับเหมาะรูปพีระมิดข้างตัดที่แสดงตามค่าความยาก



ภาพที่ 13 แบบจำลองโครงสร้างแบบสอบปรับเหมาะรูปพีระมิดข้างตัดตามทักษะที่ต้องการวัด

- หมายเหตุ
- 1 แทน ทักษะการตั้งสมมติฐาน
 - 2 แทน ทักษะการกำหนดนิยามเชิงปฏิบัติการ
 - 3 แทน ทักษะการกำหนดและควบคุมตัวแปร
 - 4 แทน ทักษะการทดลอง
 - 5 แทน ทักษะการตีความหมายข้อมูลและลงข้อสรุป

3. จัดชุดการสอบแบบปรับเหมาะ โดยนำข้อสอบบนที่กลงเครื่องคอมพิวเตอร์ โดยใช้โปรแกรม Adobe Captivate Version 2 ซึ่งจัดเรียงข้อสอบตามหลักของการสอบแบบปรับเหมาะรูปแบบพีระมิดข้างตัด ตามภาพที่ 12 เพื่อดำเนินการสอบโดยใช้เครื่องคอมพิวเตอร์

4. นำโปรแกรม Adobe Captivate Version 2 ที่บันทึกข้อสอบไปให้ผู้เชี่ยวชาญด้านสื่อการสอนตรวจสอบการจัดรูปแบบของข้อความ ความเหมาะสมของขนาดและสีของตัวอักษร การดำเนินการของโปรแกรมในการคัดเลือกข้อสอบที่มีค่าความยากเพิ่มขึ้นหรือลดลงตามผลการตอบของนักเรียน ตามภาพที่ 12 ซึ่งผู้เชี่ยวชาญได้ให้ข้อเสนอแนะว่าควรจัดระยะห่างของคำถามและตัวเลือกแต่ละข้อให้เหมาะสม พื้นหลังไม่ควรเป็นสีขาว ตัวอักษรควรเป็นตัวหนาและตัวอักษรของแต่ละตัวเลือกควรใช้คนละสี ซึ่งได้นำข้อเสนอแนะของผู้เชี่ยวชาญด้านสื่อการสอนมาปรับปรุงแก้ไข

5. จัดชุดการสอบแบบดั้งเดิม โดยใช้ข้อสอบเดียวกับการสอบแบบปรับเหมาะเพื่อดำเนินการสอบแบบดั้งเดิม โดยเรียงลำดับตามทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ และแต่ละทักษะจะเรียงจากข้อสอบที่มีค่าระดับความยากจากง่ายไปยาก

6. นำแบบสอบทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการในข้อข้างต้นจัดพิมพ์เป็นฉบับ เพื่อดำเนินการสอบแบบดั้งเดิมโดยวิธีการเขียนตอบ

การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล โดยแบ่งเป็น 2 ตอน คือการเก็บรวบรวมข้อมูลเพื่อพัฒนาแบบสอบ และการเก็บรวบรวมข้อมูลเพื่อทดสอบสมมติฐาน ดังนี้

ตอนที่ 1 การเก็บรวบรวมข้อมูลเพื่อพัฒนาแบบสอบ ผู้วิจัยดำเนินการตามลำดับ ดังนี้

1. ผู้วิจัยนำหนังสือจากโครงการปริญญาโท สาขาวิจัยและประเมินผลการศึกษาภาคพิเศษ มหาวิทยาลัยเกษตรศาสตร์ ไปติดต่อโรงเรียนสตรีศรีอยุธยา ๒ โรงเรียนนวมินทราชินูทิศ สตรีวิทยา 2 โรงเรียนนวมินทราชินูทิศ เบญจมราชาลัย โรงเรียนฤทธิยะวรรณาลัย 2 และโรงเรียนพรตพิทยพยัต สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กรุงเทพมหานคร เขต 2 เพื่อขอความอนุเคราะห์ในการเก็บรวบรวมข้อมูลจากนักเรียนในโรงเรียน สำหรับใช้เป็นข้อมูลในการพัฒนาแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ โดยชี้แจงกับอาจารย์ผู้สอนวิชาวิทยาศาสตร์เกี่ยวกับความสำคัญ วัตถุประสงค์ของการวิจัยในครั้งนี้ และประโยชน์ของนักเรียนที่ทำแบบสอบและของครูที่จะได้รับ รวมทั้งนัดหมายวัน เวลาในการสอบ

2. นำแบบสอบทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการทั้ง 76 ข้อ ไปทดสอบกับนักเรียนในข้อ 1 ซึ่งผู้วิจัยได้จัดเป็น 2 ชุด ชุดละ 38 ข้อ สำหรับใช้สอบ 2 ช่วง แต่ละช่วงห่างกันประมาณ 1 สัปดาห์ เพื่อให้ให้นักเรียนไม่ล้าในการทำข้อสอบจนเกินไปและทำแบบสอบอย่างเต็มความสามารถ

3. วิเคราะห์ผลการสอบที่ได้ตามทฤษฎีการตอบสนองข้อสอบ ด้วยโปรแกรม BILOG เพื่อหาค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าสัมประสิทธิ์การเดา (c) แล้วเก็บในคลังข้อสอบไว้สำหรับจัดชุดการสอบแบบดั้งเดิมและแบบปรับเหมาะสมรูปพีระมิดข้างตัด

ตอนที่ 2 การเก็บรวบรวมข้อมูลเพื่อทดสอบสมมติฐาน ผู้วิจัยดำเนินการตามลำดับ ดังนี้

1. ติดต่อประสานงานกับอาจารย์ผู้ควบคุมดูแลห้องคอมพิวเตอร์ เพื่อบันทึกแบบสอบปรับเหมาะสมรูปพีระมิดข้างตัดในเครื่องคอมพิวเตอร์ ทดสอบประสิทธิภาพการทำงานของเครื่องคอมพิวเตอร์ และกำหนดวันเวลาในการสอบ

2. ทำความเข้าใจกับพลวิชัยถึงความสำคัญของการสอบ ระเบียบและวิธีการสอบ

3. จัดการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดกับพลวิชัย โดยให้พลวิชัยเข้าสอบในครั้งแรกจำนวน 26 คน ตามจำนวนเครื่องคอมพิวเตอร์ และตอบคำถามผ่านจอคอมพิวเตอร์ เมื่อมีพลวิชัยสอบเสร็จจึงเรียกพลวิชัยคนต่อไปเข้าทดสอบ

4. นำผลการสอบแบบปรับเหมาะมาหาค่าประมาณความสามารถ โดยประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก ค่าเฉลี่ยความยากของข้อที่ได้ทำ และค่าความยากของข้อสุดท้ายที่ได้ตอบ

5. หลังการสอบในข้อ 4 เว้นระยะห่าง 2 สัปดาห์ จัดการสอบแบบดั้งเดิมกับพลวิชัยพร้อมกันทุกคนโดยวิธีเขียนตอบ

6. นำผลการสอบแบบดั้งเดิมมาประมาณค่าความสามารถของนักเรียน ด้วยโปรแกรม BILOG โดยนำค่าพารามิเตอร์ของข้อสอบที่ได้จากตอนที่ 1 ข้อที่ 3 ซึ่งเป็นค่าที่ได้จากผลการตอบของนักเรียนที่ใช้พัฒนาข้อสอบกลุ่มที่ 2 เป็นค่าอ้างอิงเพื่อใช้ประมาณค่าความสามารถจากการสอบแบบดั้งเดิม เช่นเดียวกับการประมาณค่าความสามารถจากการสอบแบบปรับเหมาะทั้ง 3 วิธี ซึ่งใช้ค่าความยากที่ได้จากผลการตอบของนักเรียนที่ใช้พัฒนาข้อสอบกลุ่มที่ 2

การวิเคราะห์ข้อมูล

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยทำการวิเคราะห์ข้อมูลโดยวิเคราะห์ความตรงสัมพันธ์ของค่าประมาณความสามารถของนักเรียน โดยหาค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด โดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_r) ซึ่งคำนวณจากสมการ (บุญเรียง ขจรศิลป์, 2548: 159)

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

เมื่อ r คือ ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน

n คือ จำนวนพลวิจัย

x คือ ค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูป
พรีมิตข้างตัดแต่ละวิธี (θ_{A_1} , θ_{A_2} และ θ_{A_3})

y คือ ค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T)

โดยทดสอบความสัมพันธ์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ดังนั้นหากค่า r มีค่ามากกว่า α ที่ได้จากการคำนวณจะยอมรับ H_0 กล่าวคือค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพรีมิตข้างตัดวิธีที่ทดสอบนั้น ไม่มีความสัมพันธ์กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) แต่ถ้าค่า r มีค่าน้อยกว่า α ที่ได้จากการคำนวณจะปฏิเสธ H_0 กล่าวคือค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพรีมิตข้างตัดวิธีที่ทดสอบนั้นมีความสัมพันธ์กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T)

บทที่ 4

ผลการวิจัยและข้อวิจารณ์

ผลการวิจัย

การวิจัยครั้งนี้ผู้วิจัยมีวัตถุประสงค์เพื่อศึกษาความตรงสัมพัทธ์ของค่าประมาณความสามารถ 3 วิธี จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด โดยแบ่งการวิจัยเป็น 2 ตอน คือ ตอนที่ 1 การสร้างและพัฒนาแบบสอบ และ ตอนที่ 2 การวิจัยเพื่อทดสอบสมมติฐาน โดยการเปรียบเทียบค่าประมาณความสามารถจากการสอบทั้ง 2 วิธี ผลการวิเคราะห์ มีดังนี้

ตอนที่ 1 การสร้างและพัฒนาแบบสอบ

1. ผลการตรวจสอบความตรงตามเนื้อเรื่อง (Content Validity) จากการพิจารณาคำดัชนีความสอดคล้อง (IOC) ของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ สรุปได้ดังตารางที่ 2

ตารางที่ 2 ผลการตรวจสอบความตรงตามเนื้อเรื่องจากการพิจารณาโดยผู้เชี่ยวชาญ

ทักษะ	จำนวนข้อสอบ (ข้อ)				
	ทั้งหมด	คัดเลือก	ไม่ต้องแก้ไข	ได้แก้ไข	ตัดทิ้ง
การตั้งสมมติฐาน	24	15	4	11	9
การกำหนดนิยามเชิงปฏิบัติการ	24	17	12	5	7
การกำหนดและควบคุมตัวแปร	24	18	14	4	6
การทดลอง	24	17	7	10	7
การตีความหมายข้อมูลและ ลงข้อสรุป	24	22	12	10	2

ผู้วิจัยได้สร้างข้อสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ทักษะละ 24 ข้อ และให้ผู้เชี่ยวชาญแต่ละท่านได้ตรวจสอบความตรงตามเนื้อเรื่อง พบว่ามีข้อสอบที่ได้คัดเลือก

ผ่านเกณฑ์ค่า IOC ที่ได้กำหนดไว้ โดยไม่ต้องแก้ไข บางข้อได้แก้ไขตามข้อเสนอแนะของผู้เชี่ยวชาญ สำหรับข้อที่มีค่า IOC ไม่เหมาะสมผู้วิจัยได้ตัดทิ้ง

2. ผลการพัฒนาคุณภาพข้อสอบรายข้อตามทฤษฎีการสอบแบบดั้งเดิม ได้ค่าความยาก (p) และ ค่าอำนาจจำแนก (r) จากผลการตอบของนักเรียนที่ใช้พัฒนาข้อสอบกลุ่มที่ 1 ซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนเศรษฐบุตรบำเพ็ญ จำนวน 150 คน ด้วยข้อสอบ 2 ชุด รวมทั้งสิ้น 89 ข้อ วิเคราะห์ด้วยโปรแกรม EVANA สรุปผลได้ดังตารางที่ 3

ตารางที่ 3 ค่าสถิติของแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ 2 ชุด

ชุด ที่	จำนวน ข้อ (ข้อ)	จำนวน นักเรียน(คน)	\bar{X}	S.D.	S.E.	KR-20	p	r
1	44	75	26.47	5.61	2.87	0.74	0.08 - 1.00	-0.05 - 0.74
2	45	75	25.86	5.46	2.97	0.70	0.11 - 0.92	-0.06 - 0.61

จากตาราง พบว่าแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการทั้ง 2 ชุด มีค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานใกล้เคียงกัน ค่าความเที่ยงอยู่ในเกณฑ์ค่อนข้างสูง ค่าความคลาดเคลื่อนมาตรฐานมีค่าต่ำ ค่าความยากกระจายอยู่ในช่วง 0.08 ถึง 1.00 และค่าอำนาจจำแนกกระจายอยู่ในช่วง -0.05 ถึง 0.74

3. ผลการพัฒนาคุณภาพแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการตามทฤษฎีการตอบสนองข้อสอบ (IRT) โมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์ ด้วยโปรแกรม BILOG ได้ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าสัมประสิทธิ์การเดา (c) จากผลการตอบของนักเรียนที่ใช้พัฒนาข้อสอบกลุ่มที่ 2 ซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนสตรีเศรษฐบุตรบำเพ็ญ โรงเรียนนวมินทราชินูทิศ สตรีวิทยา 2 โรงเรียนนวมินทราชินูทิศ เบญจมราชาลัย โรงเรียนฤทธิยะวรรณาลัย 2 และโรงเรียนพรตพิทยพยัต จำนวน 554 คน โดยนักเรียนได้ทำข้อสอบจำนวน 76 ข้อจากแบบสอบ 2 ชุด ชุดละ 38 ข้อ ดังนั้นก่อนการวิเคราะห์จึงนำคำตอบของนักเรียนจากการสอบทั้ง 2 ชุดรวมเป็น 1 ชุด เพื่อให้การวิเคราะห์อยู่บนพื้นฐานเดียวกัน สรุปได้ดังตารางที่ 4

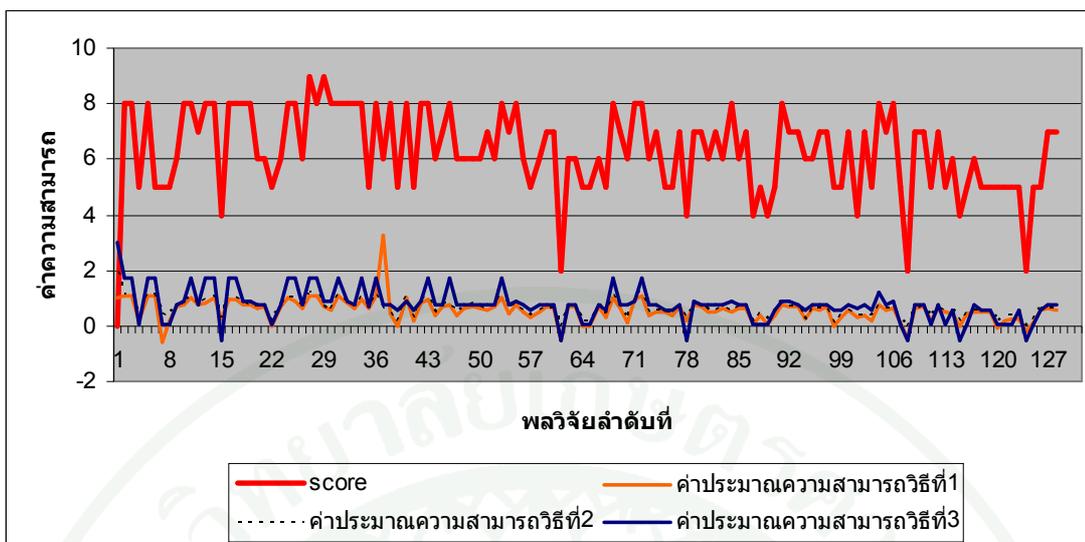
ตารางที่ 4 ค่าพารามิเตอร์ของแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ จากข้อสอบจำนวน 76 ข้อ

ประเภทของ ค่าพารามิเตอร์	ค่าพารามิเตอร์ของแบบสอบ	\bar{X}	S.D.
ค่าอำนาจจำแนก (a)	0.03 - 3.74	1.00	0.88
ค่าความยาก (b)	-0.06 - 2.15	0.60	0.59
ค่าสัมประสิทธิ์การเดา (c)	0.00 - 0.30	0.06	0.10

จากตาราง พบว่าในคลังข้อสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการจำนวน 76 ข้อ มีค่าอำนาจจำแนกอยู่ในช่วงกว้าง ส่วนใหญ่อยู่ในเกณฑ์ที่กำหนด ค่าความยากมีการกระจายค่อนข้างน้อยและส่วนใหญ่มีค่าค่อนข้างต่ำ สำหรับค่าสัมประสิทธิ์การเดามีค่าไม่เกิน 0.3 ตามที่กำหนด มีค่าความเที่ยงเท่ากับ 0.84 เมื่อวิเคราะห์ตามเกณฑ์ที่กำหนดและรูปแบบของแบบสอบที่เหมาะสม ได้ข้อสอบจำนวน 54 ข้อ จากคลังข้อสอบจำนวน 76 ข้อ เพื่อใช้ในจัดแบบสอบสำหรับการสอบแบบดั้งเดิมและแบบที่เหมาะสม

ตอนที่ 2 การวิจัยเพื่อทดสอบสมมติฐาน

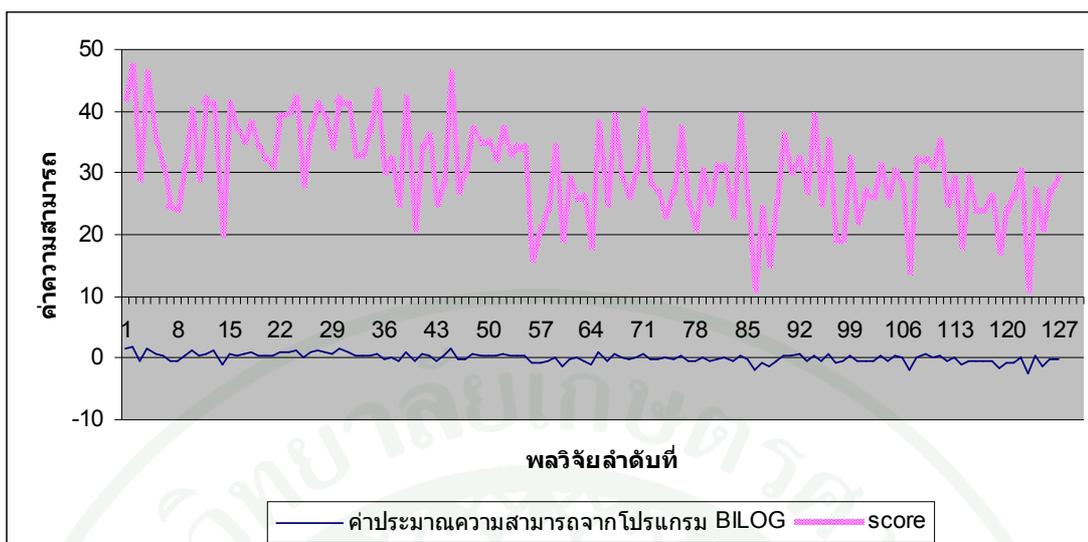
1. ผลการหาค่าประมาณความสามารถโดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) รวมทั้งคะแนนรายข้อที่ได้ตรวจตามเกณฑ์ตอบถูกได้ 1 ตอบผิดได้ 0 ที่ได้จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด แสดงได้ดังภาพที่ 14



ภาพที่ 14 ผลการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดของพลวิจัยที่พิจารณาจากคะแนนรายชื่อ และค่าประมาณความสามารถทั้ง 3 วิธี

จากภาพ แสดงให้เห็นว่าจากการสอบแบบปรับเหมาะ พลวิจัยส่วนใหญ่มีคะแนนรายชื่อ และค่าประมาณความสามารถทั้ง 3 วิธีไปในทิศทางเดียวกัน ทั้งนี้ผู้วิจัยไม่ได้กำหนดเวลาในการสอบจึงมีพลวิจัยที่ใช้เวลานานที่สุด คือ 16 นาที และพลวิจัยได้ทำข้อสอบจำนวนไม่เท่ากัน โดยมีพลวิจัยได้ทำข้อสอบจำนวน 10 11 และ 12 ข้อ เป็นจำนวน 10 13 และ 104 คนตามลำดับ ซึ่งไม่มีพลวิจัยคนใดได้ทำข้อสอบจำนวน 9 ข้อ

2. ผลการหาค่าประมาณความสามารถ(θ_r) ด้วยโปรแกรม BILOG และคะแนนรายชื่อ ที่ได้ตรวจตามเกณฑ์ตอบถูกได้ 1 ตอบผิดได้ 0 จากการสอบแบบดั้งเดิม แสดงได้ดังภาพที่ 15



ภาพที่ 15 ผลการสอบแบบดั้งเดิมของผลวิจัยที่พิจารณาจากค่าประมาณความสามารถจากโปรแกรม BILOG และคะแนนรายข้อ

จากภาพ แสดงให้เห็นว่าจากการสอบแบบดั้งเดิม ผลวิจัยส่วนใหญ่มีคะแนนรายข้อและค่าประมาณความสามารถจากโปรแกรม BILOG ไปในทิศทางเดียวกัน แต่ยังมีคะแนนรายข้อของผลวิจัยบางคนที่ได้คะแนนน้อยกว่า ในขณะที่ค่าประมาณความสามารถจากโปรแกรม BILOG มีค่ามากกว่า ทั้งนี้ผู้วิจัยไม่ได้กำหนดเวลาในการสอบจึงมีผลวิจัยที่ใช้เวลาในการทำข้อสอบ 38 นาที ถึง 3 ชั่วโมง 12 นาที

3. ผลการหาค่าความสัมพันธ์ระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด โดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) โดยนำค่าประมาณความสามารถของผลวิจัยที่ได้จากการสอบทั้ง 2 วิธีไปคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันด้วยโปรแกรม SPSS Version 12 คำสั่ง Correlation ที่ระดับนัยสำคัญทางสถิติ .01 ได้ผลดังตารางที่ 5

ตารางที่ 5 ค่าความสัมพันธ์ระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะ
รูปพีระมิดข้างตัดกับค่าประมาณความสามารถจากการสอบแบบดั้งเดิม

ค่าความสัมพันธ์ระหว่าง	ค่าสัมประสิทธิ์ สหสัมพันธ์	p
$\theta_{A_1} - \theta_T$	0.690 ***	.000
$\theta_{A_2} - \theta_T$	0.713 ***	.000
$\theta_{A_3} - \theta_T$	0.748 ***	.000

***มีนัยสำคัญทางสถิติที่ระดับ .01 ($p < .01$)

จากตาราง แสดงให้เห็นว่าเมื่อใช้ค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) เป็นเกณฑ์อ้างอิงเพื่อหาค่าความตรงสัมพันธ์ด้วยค่าสัมประสิทธิ์สหสัมพันธ์กับค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดที่มีการประมาณค่าความสามารถต่างกัน 3 วิธี มีค่าเท่ากับ 0.690, 0.713 และ 0.748 ตามลำดับ โดยค่าสัมประสิทธิ์สหสัมพันธ์ของค่าประมาณความสามารถตามค่าเฉลี่ยความยากของข้อสุดท้ายที่ได้ตอบมีค่าสูงสุด (θ_{A_3}) รองลงมาคือค่าประมาณความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และค่าประมาณความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) มีค่าต่ำที่สุด ซึ่งความสัมพันธ์ระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะโดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) มีความสัมพันธ์กันที่ระดับนัยสำคัญทางสถิติ .01

ดังนั้นสรุปได้ว่า ค่าประมาณความสามารถจากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัดโดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) มีความตรงสัมพันธ์ เมื่อใช้ค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) เป็นเกณฑ์อ้างอิง

ข้อวิจารณ์

ผู้วิจัยขอเสนอข้อวิจารณ์ โดยแบ่งเป็น 2 ตอน ดังนี้

ตอนที่ 1 การสร้างและพัฒนาแบบสอบ

1. ผู้วิจัยได้สร้างและพัฒนาแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการที่มีคุณภาพตรงตามวัตถุประสงค์ของการวัด โดยผู้วิจัยได้ใช้เวลาในการเขียนข้อสอบเป็นเวลานาน เนื่องจากต้องรวบรวมและคัดเลือกข้อมูลที่น่ามาสร้างสถานการณ์คำถามที่หลากหลายเหมาะสมกับนักเรียนและทักษะที่ต้องการวัด หลังจากเขียนข้อสอบเสร็จแล้วผู้วิจัยได้รับความอนุเคราะห์จากเพื่อนครูที่ทำการสอนในกลุ่มสาระการเรียนรู้วิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 4 และ 5 เพื่อตรวจสอบอย่างไม่เป็นทางการในด้านความชัดเจนของภาษา ความถูกต้องของคำถาม และความเหมาะสมของตัวลง จนผู้วิจัยมีความมั่นใจในระดับหนึ่งจึงเสนออาจารย์ที่ปรึกษาและผู้เชี่ยวชาญตรวจสอบความตรงตามเนื้อเรื่อง ทำให้ได้ข้อมูลที่เป็นประโยชน์ในการพัฒนาข้อสอบให้มีคุณภาพมากขึ้น และจากการตรวจสอบของผู้เชี่ยวชาญทำให้คัดเลือกข้อสอบได้จำนวน 89 ข้อ จากข้อสอบทั้งสิ้น 120 ข้อ โดยผู้วิจัยได้ตัดข้อสอบออกไปจำนวน 31 ข้อ แทนการปรับปรุง เนื่องจากเวลาอันมีจำกัด เพื่อทำให้ได้ข้อสอบที่มีหลักฐาน (ค่า IOC) แสดงคุณภาพของข้อสอบ นอกจากนี้ผู้เชี่ยวชาญได้ให้ข้อเสนอแนะว่าข้อสอบบางข้อที่แสดงสถานการณ์ของปัญหาสามารถใช้ตั้งคำถามเพื่อวัดความสามารถของนักเรียนได้หลายทักษะซึ่งจะทำให้ข้อสอบมีคุณค่ามากขึ้น แต่การสอบแบบปรับเหมาะข้อสอบแต่ละข้อต้องเป็นอิสระจากกันทำให้ผู้วิจัยเจาะจงถามเพียง 1 คำถาม จาก 1 สถานการณ์ อย่างไรก็ตามผู้วิจัยได้เขียนข้อสอบเกินจำนวนที่ต้องการจริง ทำให้สามารถคัดเลือกข้อสอบตามจำนวนที่ต้องการได้ สำหรับข้อสอบที่ตัดออกเมื่อผู้วิจัยมีโอกาสจะนำสถานการณ์ของคำถามไปปรับปรุงและพัฒนาเพื่อใช้ในการสร้างแบบสอบที่ต้องการวัดหลายทักษะ จาก 1 สถานการณ์ หรือเก็บไว้ในคลังข้อสอบเพื่อใช้ในการสอบต่อไป

2. ผู้วิจัยได้พัฒนาข้อสอบจากการทดสอบกับนักเรียน 2 กลุ่ม ทำให้ได้ข้อมูลเพื่อปรับปรุงและพัฒนาแบบสอบ ซึ่งผลการสอบของนักเรียนกลุ่มที่ 1 วิเคราะห์ตามทฤษฎีการสอบดั้งเดิมได้ข้อมูลด้านความยาก อำนาจจำแนก และประสิทธิภาพตัวลง โดยข้อสอบบางข้อถึงแม้จะมีค่าอำนาจจำแนกและค่าความยากอยู่ในเกณฑ์ที่ต้องการ แต่เมื่อพิจารณาประสิทธิภาพตัวลงจะพบว่ามีความยากบางตัวที่ควรปรับปรุง เนื่องจากไม่สามารถลงนักเรียนในกลุ่มอ่อนให้ตอบได้ หรือตัวลงบางตัวที่มีนักเรียนในกลุ่มเก่งเลือกเป็นจำนวนมาก ผู้วิจัยจึงได้ตรวจสอบรายชื่อของนักเรียน

ในกลุ่มอ่อนที่ไม่เลือกและกลุ่มเก่งที่เลือกตัวลวงนั้น เพื่อสอบถามถึงเหตุผลในการเลือกตอบคำถาม ทำให้ได้ข้อมูลมาปรับปรุงตัวลวงให้มีประสิทธิภาพมากขึ้น สำหรับข้อที่มีค่าความยากและค่าอำนาจจำแนกไม่ตรงตามเกณฑ์ ผู้วิจัยได้ตัดออกจำนวน 13 ข้อ โดยข้อสอบที่ตัดออกผู้วิจัยได้เก็บไว้เพื่อปรับปรุงและพัฒนาใส่คลังข้อสอบ สำหรับใช้ในการสอบแบบปรับเหมาะในอนาคต ส่วนผลการสอบของนักเรียนกลุ่มที่ 2 วิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) ได้ค่าพารามิเตอร์ของข้อสอบทั้ง 3 ค่า คือ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าสัมประสิทธิ์การเดา (c) ซึ่งได้มาจากกลุ่มนักเรียนที่มีจำนวนมากและประมาณค่าพารามิเตอร์ด้วยโปรแกรมทางคอมพิวเตอร์ที่มีมาตรฐาน จึงได้ค่าพารามิเตอร์ของข้อสอบที่น่าเชื่อถือสำหรับเก็บในคลังข้อสอบเพื่อใช้ในการทดสอบสมมติฐาน

จากกระบวนการสร้างและพัฒนาข้อสอบข้างต้นทำให้ผู้วิจัยได้ข้อสอบที่มีคุณภาพสำหรับนำไปใช้ในการทดสอบสมมติฐาน เพื่อให้ได้ข้อมูลที่นำไปสู่ข้อสรุปที่ถูกต้อง อย่างไรก็ตามการสอบครั้งนี้ผู้วิจัยมุ่งวัดความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ซึ่งประกอบด้วยทักษะย่อย 5 ทักษะ คือ ทักษะการตั้งสมมติฐาน ทักษะการกำหนดนิยามเชิงปฏิบัติการ ทักษะการกำหนดและควบคุมตัวแปร ทักษะการทดลอง และทักษะการตีความหมายข้อมูลและลงข้อสรุป ซึ่งการออกแบบการสอบแบบปรับเหมาะครั้งนี้ได้พิจารณาจากค่าความยากของข้อสอบ ดังนั้นจึงพบว่าข้อสอบที่วัดทักษะย่อยบางทักษะมีค่าความยากที่ไม่กระจาย ทำให้มีข้อจำกัดในการออกแบบจัดวางข้อสอบตามการสอบแบบปรับเหมาะที่ต้องการวัดทักษะย่อยได้ครบทุกทักษะทำให้การสอบครั้งนี้นักเรียนบางคนขาดการวัดในบางทักษะ (ขึ้นอยู่กับเส้นทางการตอบของนักเรียน) จึงควรมีการสร้างคลังข้อสอบที่มีข้อสอบแต่ละทักษะจำนวนมากเพื่อให้มีค่าความยากกระจาย และจัดเก็บข้อสอบโดยแบ่งตามทักษะนั้น ๆ เพื่อใช้ในการสร้างการสอบแบบปรับเหมาะที่วัดทักษะแต่ละทักษะเท่านั้น ทั้งนี้การสร้างและพัฒนาแบบสอบเป็นกระบวนการที่ละเอียดอ่อน มีการทำงานหลายขั้นตอน โดยปกติแล้วนักทดสอบจะพัฒนาระบบคลังข้อสอบอย่างต่อเนื่อง นักวิจัยจึงเริ่มสร้างและเก็บข้อสอบไว้ เพื่อเป็นการเริ่มต้นสร้างคลังข้อสอบควบคู่ไปกับการวิจัยครั้งนี้ด้วย

ตอนที่ 2 การวิจัยเพื่อทดสอบสมมติฐาน

1. ในการวิจัยครั้งนี้ได้ให้พลวิจัยทำการสอบ 2 วิธี คือ ตอบคำถามด้วยการเขียนตอบแบบดั้งเดิมทั้ง 54 ข้อ และตอบคำถามผ่านจอกอมพิวเตอร์แบบปรับเหมาะรูปพีระมิดข้างตัด โดยให้ใช้เวลาสอบอย่างเต็มที่ ซึ่งการสอบแบบดั้งเดิมพลวิจัยใช้เวลาในการทำข้อสอบ 38 นาที ถึง 3 ชั่วโมง 12 นาที ส่วนการสอบแบบปรับเหมาะพลวิจัยใช้เวลานานที่สุด 16 นาที จึงเห็นได้ว่า การสอบแบบปรับเหมาะใช้เวลาน้อยกว่าการสอบแบบดั้งเดิมประมาณ 2.5 ถึง 12 เท่า เนื่องจากการสอบแบบปรับเหมาะพลวิจัยได้ทำข้อสอบเพียง 1 ใน 5 ของการสอบแบบดั้งเดิม

2. ค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะ โดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) มีความสัมพันธ์กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) ซึ่งประมาณค่าด้วยโปรแกรม BILOG ดังนั้นการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด ชนิด 12 ชั้นสามารถนำมาประมาณค่าความสามารถแทนการสอบแบบดั้งเดิมได้ ถึงแม้ว่าการสอบแบบปรับเหมาะจะใช้ข้อสอบเพียง 1 ใน 5 ของการสอบแบบดั้งเดิม ซึ่งใกล้เคียงกับงานวิจัยของ เบย์รอฟฟ์ และซีลีย์ (สุพรรณ สุกลมลสันต์, 76: 2539) ที่พบว่า การสอบแบบปรับเหมาะที่มีจำนวนข้อสอบประมาณ 0.5 เท่าของการสอบแบบดั้งเดิมจะได้ค่าสัมประสิทธิ์สหสัมพันธ์กับแบบสอบชุดหนึ่งที่ใช้เป็นเกณฑ์

3. การประมาณค่าความสามารถด้วยทฤษฎีการตอบสนองข้อสอบ (IRT) สามารถบอกระดับความสามารถของนักเรียนที่มีความแตกต่างกันได้ดีกว่าการประมาณค่าความสามารถซึ่งพิจารณาจากจำนวนข้อที่นักเรียนทำถูก เนื่องจากนักเรียนที่ได้จำนวนข้อที่ทำถูกเท่ากันจะมีค่าประมาณความสามารถตามทฤษฎีการตอบสนองข้อสอบ (IRT) ต่างกัน ทั้งนี้เป็นเพราะว่าการประมาณค่าความสามารถของนักเรียนตามทฤษฎีการตอบสนองข้อสอบ (IRT) พิจารณาจากแบบแผนการตอบข้อสอบของนักเรียนแต่ละคน ดังนั้นถ้านักเรียนตอบข้อสอบแต่ละข้อที่มีค่าความยากต่างกัน นักเรียนก็ควรได้ค่าประมาณความสามารถต่างกัน ถึงแม้ว่าจำนวนข้อที่นักเรียนตอบถูกจะเท่ากันก็ตาม ซึ่งเป็นการยืนยันความเชื่อพื้นฐานของทฤษฎีการตอบสนองข้อสอบ (IRT) ที่ว่า พฤติกรรมการตอบสนองต่อข้อสอบของนักเรียน เป็นสิ่งที่สังเกตได้ว่าถูกหรือผิดจะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่อยู่ภายในตัวนักเรียน ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้

โดยตรง จึงสามารถประมาณค่าความสามารถนักเรียนได้อย่างอิสระโดยไม่ขึ้นอยู่กับกลุ่มนักเรียนที่เข้าสอบ (ศิริชัย กาญจนวาสี, 2550: 53)

4. วิธีการประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) Weiss (1974: 33) ได้เสนอว่าการประมาณค่าความสามารถด้วยวิธีนี้จะมีผลไวต่อการเดาคำตอบของนักเรียน ซึ่งอาจมีผลให้ค่าความเที่ยงในการวัดลดลง ซึ่งการวิจัยครั้งนี้ได้วิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) โมเดลโลจิสติก (Logistic Model) 3 พารามิเตอร์ เพื่อคัดเลือกข้อสอบที่มีค่าสัมประสิทธิ์การเดาไม่เกิน 0.3 ทำให้นักเรียนมีโอกาสในการทำข้อสอบถูกโดยการเดาน้อย ดังนั้นค่าประมาณความสามารถที่ได้จากวิธีนี้จึงมีความสัมพันธ์กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิมซึ่งประมาณค่าด้วยโปรแกรม BILOG

5. วิธีการประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) ซึ่ง Lord (cited in Weiss, 1974: 33) ได้เสนอให้นำค่าความยากของข้อสอบทุกข้อที่นักเรียนได้ทำมาประมาณค่าความสามารถโดยไม่คำนึงถึงว่าจะตอบถูกหรือไม่ ทำให้ค่าประมาณความสามารถของนักเรียนแต่ละคนมีค่าไม่แตกต่างกันมากนัก แต่ถ้ามีจำนวนชั้นของพืระมิตมากขึ้นก็จะทำให้ค่าประมาณความสามารถที่ได้มีค่าแตกต่างกันมากขึ้น โดยการวิจัยครั้งนี้ได้สร้างแบบสอบปรับเหมาะ ชนิด 12 ชั้น ทำให้ค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพืระมิตข้างตัดกับการสอบแบบดั้งเดิมมีความสัมพันธ์กัน

6. วิธีการประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) เป็นวิธีที่ง่ายที่สุดในการประมาณค่าความสามารถ เนื่องจากพิจารณาจากค่าความยากของข้อสุดท้ายที่ได้ตอบเพียงค่าเดียว ทั้งนี้ Weiss (1974: 31) ได้เสนอว่าถ้าจำนวนชั้นของพืระมิตมีน้อยวิธีนี้จะให้ค่าประมาณความสามารถที่มีค่าต่ำ ทำให้ค่าสหสัมพันธ์ระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปพืระมิตกับการสอบแบบอื่นมีค่าต่ำด้วย แต่การวิจัยครั้งนี้ได้จัดชุดการสอบแบบปรับเหมาะรูปพืระมิตข้างตัด ชนิด 12 ชั้น จึงทำให้ค่าประมาณความสามารถจากการสอบแบบปรับเหมาะรูปพืระมิตข้างตัดกับการสอบแบบดั้งเดิมมีความสัมพันธ์กัน

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

สรุปผลการวิจัย

สาระสำคัญของการวิจัยเรื่องความตรงสัมพันธ์ของค่าประมาณความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการจากการสอบแบบปรับเหมาะรูปพีระมิด สรุปได้ดังนี้

วัตถุประสงค์ของการวิจัย

เพื่อศึกษาความตรงสัมพันธ์ของค่าประมาณความสามารถ 3 วิธี จากการสอบแบบปรับเหมาะรูปพีระมิดข้างตัด

ระเบียบวิธีวิจัย

การวิจัยครั้งนี้ใช้ระเบียบวิธีวิจัยเชิงประจักษ์ (Empirical Reserch) เพื่อนำข้อมูลมาใช้ในการทดสอบสมมติฐาน

พลวิจัย

พลวิจัย เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนแผนวิทยาศาสตร์ - คณิตศาสตร์ ภาคเรียนที่ 2 ปีการศึกษา 2552 โรงเรียนเศรษฐบุทรบาเพ็ญ สังกัดสำนักงานการศึกษาขั้นพื้นฐาน กรุงเทพมหานครเขต 2 จำนวน 127 คน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ซึ่งประกอบด้วยทักษะการตั้งสมมติฐาน ทักษะการกำหนดนิยามเชิงปฏิบัติการ ทักษะ

การกำหนดและควบคุมตัวแปร ทักษะการทดลอง และทักษะการตีความหมายข้อมูลและลงข้อสรุป เป็นแบบปรนัย 4 ตัวเลือก สำหรับใช้ในการสอบ 2 วิธี คือ

1. การสอบแบบปรับเหมาะรูปร่างพีระมิดข้างตัดสร้างขึ้นตามแบบจำลองรูปร่างพีระมิดข้างตัดที่กำหนดแนวทางในการตอบ ชนิด 12 ชั้น โดยใช้ข้อคำถามจากการสอบดั้งเดิมมาออกแบบการสอบปรับเหมาะรูปร่างพีระมิดข้างตัดนำเสนอด้วยคอมพิวเตอร์จากโปรแกรม Adobe Captivate Version 2 โดยให้พลวิจัยเข้าสอบในครั้งแรกจำนวน 26 คน ตามจำนวนเครื่องคอมพิวเตอร์ และตอบคำถามผ่านจอคอมพิวเตอร์ เมื่อมีพลวิจัยสอบเสร็จจึงเรียกพลวิจัยคนต่อไปเข้าทดสอบ ซึ่งพลวิจัยจะต้องตอบคำถาม จำนวน 9 – 12 ข้อ ขึ้นอยู่กับความสามารถของตนและข้อสอบแต่ละข้ออาจไม่เหมือนกัน

2. การสอบแบบดั้งเดิมที่พลวิจัยตอบโดยการเขียนตอบพร้อมกันทุกคน ซึ่งมีข้อคำถามเดียวกับการสอบแบบปรับเหมาะ จำนวน 54 ข้อ แบ่งตามทักษะได้ ดังนี้

ทักษะการตั้งสมมติฐาน	จำนวน 8 ข้อ
ทักษะการกำหนดนิยามเชิงปฏิบัติการ	จำนวน 9 ข้อ
ทักษะการกำหนดและควบคุมตัวแปร	จำนวน 12 ข้อ
ทักษะการทดลอง	จำนวน 10 ข้อ
ทักษะการตีความหมายข้อมูลและลงข้อสรุป	จำนวน 15 ข้อ

การวิเคราะห์ข้อมูล

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยทำการวิเคราะห์ข้อมูล โดยวิเคราะห์ความตรงสัมพันธ์ของค่าประมาณความสามารถของนักเรียน โดยหาค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะรูปร่างพีระมิดข้างตัด โดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) ซึ่งคำนวณจากสมการ (บุญเรียง ขจรศิลป์, 2548: 159)

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

เมื่อ r คือ ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน

n คือ จำนวนพลวิจัย

x คือ ค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะสมรูป
พีระมิดข้างตัดแต่ละวิธี (θ_{A_1} , θ_{A_2} และ θ_{A_3})

y คือ ค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T)

โดยทดสอบความสัมพันธ์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ดังนั้นหากค่า r มีค่ามากกว่า α ที่ได้จากการคำนวณจะยอมรับ H_0 กล่าวคือค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะสมรูปพีระมิดข้างตัดวิธีที่ทดสอบนั้นไม่มีความสัมพันธ์กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) แต่ถ้าค่า r มีค่าน้อยกว่า α ที่ได้จากการคำนวณจะปฏิเสธ H_0 กล่าวคือค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะสมรูปพีระมิดข้างตัดวิธีที่ทดสอบนั้นมีความสัมพันธ์กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T)

ผลการวิจัย

ความสัมพันธ์ระหว่างค่าประมาณความสามารถที่ได้จากการสอบแบบปรับเหมาะสม โดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ได้ตอบ (θ_{A_3}) กับค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ดังนั้นสรุปได้ว่า ค่าประมาณความสามารถจากการสอบแบบปรับเหมาะสมรูปพีระมิดข้างตัดโดยวิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ทำถูก (θ_{A_1}) วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของข้อที่ได้ทำ (θ_{A_2}) และวิธีประมาณค่าความสามารถตามค่า

ความยากของข้อสุดท้ายที่ได้ตอบ (θ_{4_3}) มีความตรงสัมพันธ์ เมื่อใช้ค่าประมาณความสามารถที่ได้จากการสอบแบบดั้งเดิม (θ_T) เป็นเกณฑ์อ้างอิง

ข้อเสนอแนะ

ข้อเสนอแนะสำหรับการนำผลการวิจัยไปใช้

1. ครูควรสร้างข้อสอบให้มีจำนวนมาก และวิเคราะห์ตามทฤษฎีการสอบแบบดั้งเดิม เพื่อปรับปรุงตัวเลือกและประสิทธิภาพตัวลวงของข้อสอบ แล้วจึงนำไปทดสอบกับนักเรียนเพื่อวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) ให้ได้ข้อสอบที่มีคุณภาพ

2. ครูควรเก็บรวบรวมข้อสอบที่มีคุณภาพไว้ในคลังข้อสอบให้มีจำนวนมากพอ โดยแบ่งเป็นหมวดหมู่ เป็นเรื่องย่อย ๆ เช่น หมวดทักษะกระบวนการทางวิทยาศาสตร์ เมื่อครูต้องการนำไปใช้ในการสอบแบบปรับเหมาะก็สามารถนำข้อสอบไปใช้ได้สะดวก ซึ่งการสอบแบบปรับเหมาะจากข้อสอบที่มีคุณภาพจะทำให้สามารถประมาณค่าความสามารถของนักเรียนได้อย่างแม่นยำโดยที่นักเรียนไม่เกิดความเครียดจากการสอบ

3. ครูที่มีข้อสอบในคลังจำนวนมากถ้านำมาสร้างแบบสอบแบบปรับเหมาะจะทำให้สามารถศึกษาพัฒนาการด้านความสามารถของนักเรียนที่ได้จากการวัด โดยเปิดโอกาสให้นักเรียนทำแบบสอบหลายครั้ง เช่น การสอบก่อนเรียนและหลังเรียน การสอบประจำปี เนื่องจากการสอบแต่ละครั้งนักเรียนจะได้รับข้อสอบที่วัดในเรื่องเดียวกันแต่มีระดับความยากของข้อสอบแตกต่างกันตามระดับความสามารถของคนที่ได้รับการพัฒนามากขึ้น

4. ครูสามารถเลือกใช้วิธีการประมาณค่าวิธีใดวิธีหนึ่งใน 3 วิธีที่ผู้วิจัยได้นำเสนอก็ได้ หรือครูจะศึกษาเรียนรู้ข้อได้เปรียบเสียเปรียบจากวิธีใดวิธีหนึ่งก็จะเป็นประโยชน์ต่อการวัดผลของครูมากยิ่งขึ้น

5. โรงเรียนที่มีความพร้อมในการจัดการเรียนรู้แบบ E-Learning ที่สามารถออนไลน์ในระบบ Internet ควรนำหลักการสอบแบบปรับเหมาะไปใช้ เพื่อให้นักเรียนสามารถทดสอบความรู้ได้ด้วยตนเอง ซึ่งทำให้ผู้สอนไม่ต้องใช้เวลาในคาบเรียนสำหรับทดสอบนักเรียน เนื่องจากการ

สอบลักษณะนี้ต้องเตรียมเครื่องคอมพิวเตอร์สำหรับการสอบให้เพียงพอกับจำนวนนักเรียน หากเครื่องคอมพิวเตอร์มีน้อยเกินไปผู้สอนต้องใช้เวลาในการดำเนินการสอบมากขึ้น

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ควรมีการวิจัยเกี่ยวกับการสอบแบบปรับเหมาะในรูปแบบอื่น ๆ เช่น กลวิธีของเบย์แบบปรับระดับชั้น กลวิธีการประมาณค่าความเป็นไปได้สูงสุด
2. ควรมีการวิจัยเกี่ยวกับวิธีประมาณค่าความสามารถจากการสอบแบบปรับเหมาะในวิชาอื่น ๆ เพื่อหาความสอดคล้องของค่าประมาณความสามารถที่ได้ด้วยวิธีใดวิธีหนึ่งกับค่าประมาณความสามารถจากการสอบแบบดั้งเดิมหรือคิดค้นวิธีการประมาณค่าที่มีความแม่นยำมากยิ่งขึ้น

เอกสารและสิ่งอ้างอิง

ชูใจ ทรงเมฆ. 2538. การเปรียบเทียบผลการประมาณค่าความสามารถด้วยวิธีการทดสอบแบบ
ปรับเหมาะและวิธีการทดสอบแบบดั้งเดิม. วิทยานิพนธ์ศึกษาศาสตรมหาบัณฑิต สาขาวิจัย
และประเมินผลการศึกษา, มหาวิทยาลัยเกษตรศาสตร์.

บุญเรียง ขจรศิลป์. 2543. วิธีวิจัยทางการศึกษา. พิมพ์ครั้งที่ 5. กรุงเทพมหานคร:
เอส.พี.เอ็น.การพิมพ์.

_____. 2548. การวิเคราะห์และแปลความหมายข้อมูลในการวิจัยโดยใช้โปรแกรมสำเร็จรูป
SPSS for Window Version 10-12. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: เอส. พี.เอ็น.การ
พิมพ์.

วนิดา มุขดาร์. 2544. การเปรียบเทียบการประมาณค่าความสามารถวิชาภาษาไทย ของนักศึกษา
การศึกษานอกโรงเรียน จังหวัดสมุทรปราการ ตามทฤษฎีการตอบสนองข้อสอบระหว่าง
โมเดลหนึ่งพารามิเตอร์ สองพารามิเตอร์ และสามพารามิเตอร์ เมื่อขนาดกลุ่มตัวอย่าง
แตกต่างกัน. วิทยานิพนธ์ศึกษาศาสตรมหาบัณฑิต สาขาวิชาการวัดผลศึกษา, มหาวิทยาลัย
ศรีนครินทรวิโรฒ.

วสันต์ ทองไทย. 2539. การเปรียบเทียบผลการประมาณค่าความสามารถด้วยวิธีการทดสอบแบบ
ปรับเหมาะรูปพีระมิตขนาดชั้นคงที่ และ รูปพีระมิตข้างตัด : การทดสอบทักษะ
กระบวนการทางวิทยาศาสตร์. วิทยานิพนธ์ศึกษาศาสตรมหาบัณฑิต สาขาวิจัยและ
ประเมินผลการศึกษา, มหาวิทยาลัยเกษตรศาสตร์.

ศิริชัย กาญจนวาตี. 2538. การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ. กรุงเทพ-
มหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

_____. 2550. ทฤษฎีการทดสอบแนวใหม่ (MODERN TEST THEORIES). พิมพ์ครั้งที่
3. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. 2552. ทฤษฎีการทดสอบแบบดั้งเดิม (CLASSICAL TEST THEORY).
พิมพ์ครั้งที่ 5. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สุพัฒน์ สุขมลสันต์. 2539. ธนาการข้อทดสอบและการทดสอบปรับเปลี่ยนด้วยคอมพิวเตอร์.
กรุงเทพมหานคร: วิทยพัฒน์.

_____. 2542. การวิเคราะห์ข้อสอบและตัดเกรดด้วยคอมพิวเตอร์. กรุงเทพมหานคร:
วิทยพัฒน์.

อัศวิน คงเพชรศักดิ์. 2547. การศึกษาความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณค่า
ความสามารถของผู้สอบโดยใช้แบบทดสอบปรับเหมาะรูปปิรามิดในวิชาคณิตศาสตร์ ชั้น
ประถมศึกษาปีที่ 6. วิทยานิพนธ์ศึกษาศาสตร์มหาบัณฑิต สาขาวิชาการวัดและประเมินผล
การศึกษา, มหาวิทยาลัยรามคำแหง.

อทุมพร จามรมาน. 2537. ทฤษฎีการวัดทางจิตวิทยา. กรุงเทพมหานคร: ฟีนีქซ์พับลิชชิง.

อรินทร์ น่วมถนอม. 2537. การศึกษาความเที่ยงตรงและประสิทธิภาพสัมพัทธ์ของการประมาณ
ค่าความสามารถในวิชาคณิตศาสตร์ระหว่างแบบทดสอบดั้งเดิม แบบทดสอบสองขั้นตอน
และแบบทดสอบเฟล็กซิเบิล. วิทยานิพนธ์ศึกษาศาสตร์มหาบัณฑิต สาขาวิชาการวัดผล
ศึกษา, มหาวิทยาลัยศรีนครินทรวิโรฒ.

Bizot *et al.* 1994. **The Practical Impact of IRT Models and Parameters When Converting a
Test to Adaptive Format** (Online). [http://eric.ed.gov/ERICWebPortal/
contentdelivery/servlet/ERICServlet?accno=ED397061](http://eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED397061). , June 25, 2009.

Weiss, D.J. **Strategies of Adaptive Ability Measurement** (Online).
<http://www.psych.umn.edu/psylabs/catcentral/pdf%20files/we74-05.pdf>, April 15, 2010.

Hambleton, R.K. and H. Swaminathan. 1984. **Item Response Theory Principles and
Applications**. U.S.A. : Kluwer Academic Publishers Group.

Hambleton, R.K. and H. Swaminathan. 1985. **Item Response Theory**. Boston : Kluwer Nijhoff Publishing.

_____. 1991. **Fundamentals of Item Response Theory**. Newbury Park : Sage.

Lewis R. Aiken and Gary Groth-Marnat. 2003. **Psychological Testing and Assessment**. U.S.A. : Pearson Education Group.

Liane N. Patsula and Marc E. Gessaroli. 1995. **A Comparison of Item Parameter Estimates and ICCs Produced with TESTGRAF and BILOG under Different Test Lengths and Sample Sizes** (Online). http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/14/da.pdf , July 15, 2009.

Swaminathan, H., and Gifford, J.A.. 1985. Bayesian estimation in the two-parameter logistic model. **Psychometrika**. 50 (3) : 349-364: September, 1985.

Tian Jian-quan *et al.* 2007. "An Introduction to the Computerized Adaptive Testing." **US-China Education Review** (Online). <http://www.teacher.org.cn/doc/ucedu200701/ucedu20070114.pdf> , June 25, 2009.

Wen, Chang, and Hau. 2001. **Optimum Number of Strata in the a-Stratified Computerized Adaptive Testing Design** (Online). http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1a/22/17.pdf , June 25, 2009.

Widiatmo and Heru Sundari. "A Simulation and Evaluation of Computerized Adaptive Testing Designs for the Verbal Battery of the Cognitive Abilities Test (CogAT)," **Dissertation Abstracts International**. 65(01) : 1336-A : October, 2004.



ภาคผนวก



รายชื่อผู้เชี่ยวชาญในการตรวจสอบเครื่องมือ

1. ดร.โชคชัย ยืนยง
อาจารย์ผู้สอนประจำคณะศึกษาศาสตร์
มหาวิทยาลัยขอนแก่น
2. ดร.ปัฐมาภรณ์ พิมพ์ทอง
อาจารย์ผู้สอนประจำคณะศึกษาศาสตร์
มหาวิทยาลัยขอนแก่น
3. ดร.ประสงค์ เมธิพิณิตกุล
นักวิชาการสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี
4. อาจารย์वासนา กิรติจำเริญ
นักวิชาการสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี
5. อาจารย์อรุณี ชูทัย
อาจารย์ผู้สอนในกลุ่มสาระการเรียนรู้วิทยาศาสตร์ระดับมัธยมศึกษา
โรงเรียนเศรษฐบุทรบำเพ็ญ
6. อาจารย์วรรณภา ลาวัง
อาจารย์ผู้สอนในกลุ่มสาระการเรียนรู้วิทยาศาสตร์ระดับมัธยมศึกษา
โรงเรียนเศรษฐบุทรบำเพ็ญ
7. อาจารย์อาชวินี ไชยสุนทร
อาจารย์ผู้สอนในกลุ่มสาระการเรียนรู้วิทยาศาสตร์ระดับมัธยมศึกษา
โรงเรียนเศรษฐบุทรบำเพ็ญ
8. อาจารย์นงพะงา เรืองฤทธิ์ชาญกุล
อาจารย์ผู้สอนในกลุ่มสาระการเรียนรู้วิทยาศาสตร์ระดับมัธยมศึกษา
โรงเรียนเศรษฐบุทรบำเพ็ญ

9. อาจารย์พัชรินทร์ ศุภสมบัติ
อาจารย์ผู้สอนในกลุ่มสาระการเรียนรู้วิทยาศาสตร์ระดับมัธยมศึกษา
โรงเรียนสตรีศรีบุญคุรบำเพ็ญ

10. อาจารย์สยามไท โสภณ
อาจารย์ผู้สอนในกลุ่มสาระการเรียนรู้วิทยาศาสตร์ระดับมัธยมศึกษา
โรงเรียนบางกะปิ

11. อาจารย์ปราชญา ปานกลีบ
ผู้เชี่ยวชาญด้านสื่อการสอน
โรงเรียนศรีบุญคุรบำเพ็ญ



ภาคผนวก ข
ผลการตรวจสอบความตรงตามเนื้อเรื่อง

ตารางผนวกที่ 1 ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะกระบวนการทาง
วิทยาศาสตร์ชั้นบูรณาการ (ทักษะการตั้งสมมติฐาน)

ข้อที่	คะแนนความคิดเห็นของผู้เชี่ยวชาญ					ค่า IOC	สรุปผล
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5		
1	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
2	+1	+1	0	0	+1	0.60	ตัดทิ้ง
3	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
4	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
5	+1	0	0	+1	+1	0.60	ตัดทิ้ง
6	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
7	0	0	0	-1	+1	0.40	ตัดทิ้ง
8	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
9	+1	+1	0	0	+1	0.60	ตัดทิ้ง
10	0	+1	+1	0	+1	0.60	ตัดทิ้ง
11	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
12	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
13	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
14	0	+1	0	+1	+1	0.60	ตัดทิ้ง
15	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
16	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
17	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
18	+1	0	0	+1	+1	0.60	ตัดทิ้ง
19	+1	0	0	+1	+1	0.60	ตัดทิ้ง
20	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
21	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
22	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
23	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
24	+1	0	0	+1	+1	0.60	ตัดทิ้ง

ตารางผนวกที่ 2 ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะกระบวนการทาง
วิทยาศาสตร์ขั้นบูรณาการ (ทักษะการกำหนดนิยามเชิงปฏิบัติการ)

ข้อที่	คะแนนความคิดเห็นของผู้เชี่ยวชาญ					ค่า IOC	สรุปผล
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5		
1	+1	0	+1	0	+1	0.60	ตัดทิ้ง
2	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
3	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
4	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
5	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
6	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
7	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
8	+1	0	+1	0	+1	0.60	ตัดทิ้ง
9	+1	0	0	0	+1	0.40	ตัดทิ้ง
10	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
11	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
12	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
13	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
14	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
15	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
16	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
17	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
18	+1	0	0	+1	+1	0.60	ตัดทิ้ง
19	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
20	-1	0	+1	-1	+1	0.00	ตัดทิ้ง
21	0	0	+1	0	+1	0.40	ตัดทิ้ง
22	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
23	+1	0	+1	0	+1	0.60	ตัดทิ้ง
24	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้

ตารางผนวกที่ 3 ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะกระบวนการทาง
วิทยาศาสตร์ชั้นบูรณาการ (ทักษะการกำหนดและควบคุมตัวแปร)

ข้อที่	คะแนนความคิดเห็นของผู้เชี่ยวชาญ					ค่า IOC	สรุปผล
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5		
1	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
2	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
3	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
4	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
5	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
6	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
7	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
8	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
9	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
10	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
11	+1	-1	0	+1	+1	0.40	ตัดทิ้ง
12	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
13	0	+1	0	+1	+1	0.60	ตัดทิ้ง
14	0	0	+1	+1	+1	0.60	ตัดทิ้ง
15	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
16	0	+1	0	+1	+1	0.60	ตัดทิ้ง
17	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
18	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
19	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
20	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
21	0	0	0	+1	0	0.20	ตัดทิ้ง
22	+1	-1	0	+1	+1	0.40	ตัดทิ้ง
23	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
24	0	+1	+1	+1	+1	0.80	คัดเลือกไว้

ตารางผนวกที่ 4 ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะกระบวนการทาง
วิทยาศาสตร์ชั้นบูรณาการ (ทักษะการทดลอง)

ข้อที่	คะแนนความคิดเห็นของผู้เชี่ยวชาญ					ค่า IOC	สรุปผล
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5		
1	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
2	+1	0	0	+1	+1	0.60	ตัดทิ้ง
3	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
4	0	+1	+1	+1	+1	0.80	คัดเลือกรับ
5	+1	+1	0	+1	0	0.60	ตัดทิ้ง
6	+1	+1	-1	+1	+1	0.60	ตัดทิ้ง
7	0	+1	+1	+1	+1	0.80	คัดเลือกรับ
8	+1	+1	0	+1	0	0.60	ตัดทิ้ง
9	0	+1	+1	+1	+1	0.80	คัดเลือกรับ
10	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
11	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
12	+1	0	0	+1	+1	0.60	ตัดทิ้ง
13	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
14	0	+1	+1	+1	+1	0.80	คัดเลือกรับ
15	0	+1	+1	+1	+1	0.80	คัดเลือกรับ
16	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
17	0	0	0	+1	+1	0.40	ตัดทิ้ง
18	0	+1	+1	+1	+1	0.80	คัดเลือกรับ
19	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
20	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
21	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
22	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
23	+1	+1	+1	+1	+1	1.00	คัดเลือกรับ
24	+1	0	0	+1	+1	0.60	ตัดทิ้ง

ตารางผนวกที่ 5 ผลการตรวจสอบความตรงตามเนื้อเรื่องของแบบสอบวัดทักษะกระบวนการทาง
วิทยาศาสตร์ขั้นบูรณาการ (ทักษะการตีความหมายข้อมูลและลงข้อสรุป)

ข้อที่	คะแนนความคิดเห็นของผู้เชี่ยวชาญ					ค่า IOC	สรุปผล
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5		
1	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
2	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
3	+1	0	0	+1	+1	0.60	ตัดทิ้ง
4	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
5	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
6	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
7	0	+1	+1	+1	+1	0.80	คัดเลือกไว้
8	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
9	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
10	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
11	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
12	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
13	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
14	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
15	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
16	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
17	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
18	+1	0	0	+1	+1	0.60	ตัดทิ้ง
19	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
20	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
21	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
22	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
23	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้
24	+1	+1	+1	+1	+1	1.00	คัดเลือกไว้



การแสดงข้อความของการสอบแบบปรับเหมาะจาก
โปรแกรม ADOBE CAPTIVATE VERSION 2
ผ่านจอคอมพิวเตอร์

ตารางแสดงปริมาณความร้อนของถ่านอัดแท่ง 500 g. ที่ได้จากหญาค่าผสมกับ
ซีลีอในอัตราส่วนที่ต่างกัน จากตารางข้างต้น นักเรียนควรสรุปผลการทดลอง
ได้อย่างไร

หญาค่า : ซีลีอ	ปริมาณความร้อน (cal.)
1 : 4	250
2 : 3	180
2.5 : 2.5	130
3 : 2	110
4 : 1	80

- 1) ปริมาณหญาค่าแปรผันตรงกับปริมาณความร้อน
- 2) ปริมาณความร้อนขึ้นอยู่กับชนิดของวัตถุที่ใช้ในการเผาไหม้
- 3) เมื่อใช้หญาค่า 150 g. และ ซีลีอ 100 g. จะได้ปริมาณความร้อน 110 cal.
- 4) อัตราส่วนของหญาค่า : ซีลีอ เท่ากับ 1 : 4 จะได้ปริมาณความร้อนมากที่สุด

Question 1 of 54

Submit

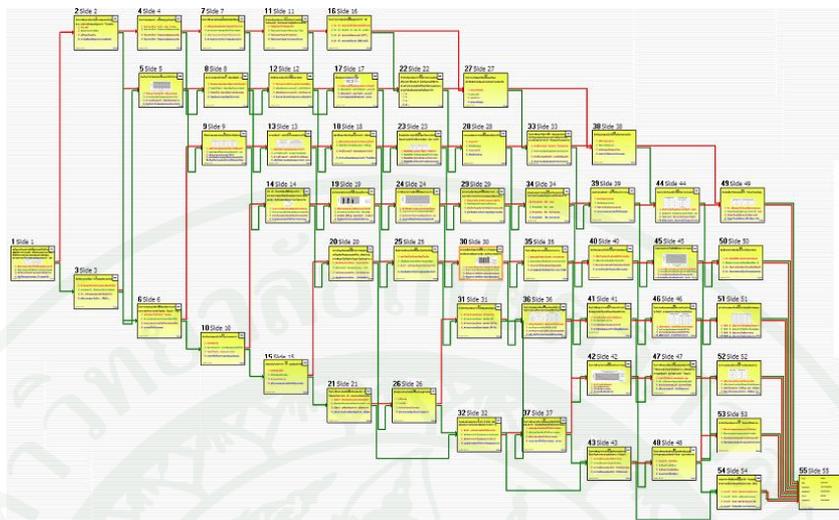
ถ้านักเรียนต้องการทดสอบสมมติฐานที่ว่า " สารละลายชนิดต่าง ๆ เมื่อมีความ
เข้มข้นต่างกันจะมีค่าดัชนีหักเหแสงต่างกันด้วย " นักเรียนควรเตรียมสารละลาย
ตามข้อใด

- 1) สารละลายกลูโคส, สารละลายซูโครส และสารละลายเกลือแกงที่มีความ
เข้มข้นชนิดละ 25%
- 2) สารละลายกลูโคส เข้มข้น 25 % , สารละลายซูโครส เข้มข้น 50 %
และสารละลายเกลือแกง เข้มข้น 75 %
- 3) สารละลายกลูโคส เข้มข้น 25 % , 50 % , 75 % และ 100 % ที่มีปริมาตร
100, 200, 300 และ 400 cm³. ตามลำดับ
- 4) สารละลายกลูโคส เข้มข้น 25 % , 50 % , 75 % , 100 % และสารละลาย
ซูโครส, สารละลายเกลือแกง ที่มีความเข้มข้นเท่ากับสารละลายกลูโคส

Question 31 of 54

Submit

การแสดงผลเส้นทางการดำเนินการสอบแบบปรับเหมาะตามคำตอบของนักเรียนในแต่ละข้อ



การรายงานผลการสอบแบบปรับเหมาะ ซึ่งผู้สอนสามารถกดปุ่ม Review Quiz เพื่อกลับไปดูคำตอบของนักเรียนได้

จำนวนข้อที่ตอบถูก

0

สิ้นสุดการสอบ
ดูจากที่นั่งได้ค่ะ

Continue Review Quiz



ภาคผนวก ง
แบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

แบบสอบทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

คำชี้แจง

1. ข้อสอบเป็นแบบปรนัย 4 ตัวเลือก จำนวน 54 ข้อ 19 หน้า
2. เขียนชื่อ – นามสกุล ด้วยปากกาในกระดาษคำตอบ
3. เขียนรหัสด้วยปากกา โดย ตัวเลข 2 หลักแรกเป็นเลขที่และหลักที่ 3 เป็นห้องของนักเรียน เช่น
เลขที่ 1 ห้อง ม.4/1 รหัส 011 เลขที่ 30 ห้อง ม.4/3 รหัส 303
4. ในการตอบให้ใช้ปากกาทำเครื่องหมาย X ลงในกระดาษคำตอบลงในข้อที่ถูกต้องที่สุด
5. หากนักเรียนต้องการเปลี่ยนคำตอบให้ลบคำตอบเดิมออก หรือ ทำดังนี้

ข้อ	1	2	3	4
1.	X		X	

6. ห้ามนำแบบสอบและกระดาษคำตอบออกจากห้องสอบ
7. นักเรียนควรคิดให้รอบคอบก่อนที่จะตอบ เพราะการเดาไม่ช่วยให้คะแนนดีขึ้นเลย

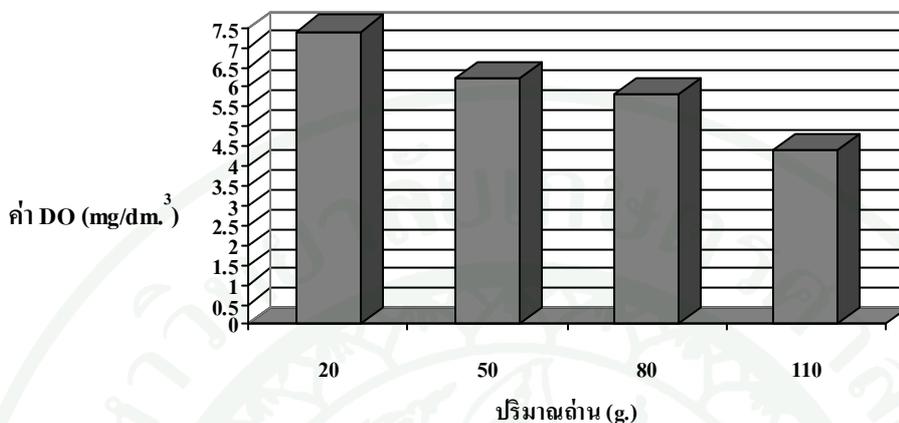
1. นักปฐุสัต์ว้ได้ทดลองนำไ้ช่นไทยที่อยู๋ในวัยเจริญพันธุ์ แต่มีอายุต่างกันมารีดน้ำเชื้อแล้วผสมพันธุ์กับไ้ไทยตัวเมียเพื่อหาเปอร์เซ็นต์การผสมติดได้ผลการทดลอง ดังตาราง

อายุของไ้ช่นไทย (เดือน)	เปอร์เซ็นต์การผสมติดของน้ำเชื้อสด
8	98
12	86
30	68

สมมติฐานที่นักปฐุสัต์ว้กำลังพิสูจน์คือข้อใด

- น้ำเชื้อของไ้ไทยที่มีอายุน้อยจะผสมติดได้ดีกว่าไ้ไทยที่มีอายุมาก
 - ไ้ช่นไทยยังมีอายุมากเท่าไร จะมีปริมาณของน้ำเชื้อสดน้อยลงเท่านั้น
 - ไ้ช่นไทยที่มีอายุ 8 เดือน จะมีเปอร์เซ็นต์การผสมติดของน้ำเชื้อสด เท่ากับ 98
 - ไ้ช่นไทยที่มีอายุ 30 เดือน จะมีเปอร์เซ็นต์การผสมติดของน้ำเชื้อสดสูงกว่า 50
2. ในการศึกษาวิธีการกำจัดแมลงวันทอง นักเรียนกลุ่มหนึ่งพบว่าสารสกัดจากพืชบางชนิดสามารถทำให้แมลงวันทองตายได้ พวกเขาจึงใช้สารสกัดจากพืชชนิดเดียวกัน ปริมาณเท่ากัน มาผสมกับน้ำในปริมาณที่ต่างกัน แล้วนำไปวางในที่ต่าง ๆ ซึ่งพบแมลงวันทอง นักเรียนกลุ่มนี้กำลังพิสูจน์สมมติฐานในข้อใด
- ถ้าปริมาณน้ำมีผลทำให้แมลงวันทองตาย ดังนั้นปริมาณน้ำที่มากขึ้น จะทำให้แมลงวันทองตายมากขึ้น
 - ถ้าปริมาณพืชที่นำมาสกัดมีผลทำให้แมลงวันทองตาย ดังนั้นปริมาณพืชที่มากขึ้น จะทำให้แมลงวันทองตายมากขึ้น
 - ถ้าความเข้มข้นของสารละลายจากสารสกัดจากพืชมีผลต่อแมลงวันทองขนาดใหญ่ ดังนั้นความเข้มข้นของสารละลายจากสารสกัดจากพืชที่มากขึ้น จะทำให้แมลงวันทองขนาดใหญ่ตายมากขึ้น
 - ถ้าความเข้มข้นของสารละลายจากสารสกัดจากพืชมีผลทำให้แมลงวันทองตาย ดังนั้นความเข้มข้นของสารละลายจากสารสกัดจากพืชที่มากขึ้น จะทำให้แมลงวันทองตายมากขึ้น

3. ศึกษาเกี่ยวกับคุณภาพน้ำเพื่อทำให้น้ำในตู้ปลาเน่าเสียช้าลงโดยใส่ถ่านในตู้ปลาที่มีสภาพแวดล้อมเหมือนกันและมีค่าการละลายในน้ำของออกซิเจน (DO) เริ่มต้น 2 mg./dm.^3 ได้ผลการทดลองดังแผนภูมิ



ข้อใดเป็นสมมติฐานของสุชา

1. ปริมาณถ่านและค่าการละลายในน้ำของออกซิเจนมีผลต่อคุณภาพน้ำ
 2. เมื่อใส่ถ่านในตู้เลี้ยงปลา 110 g. ค่าการละลายในน้ำของออกซิเจน จะเพิ่ม 2 mg./dm.^3
 3. ถ้าปริมาณน้ำมีผลต่อคุณภาพน้ำในตู้เลี้ยงปลา ดังนั้นน้ำในตู้เลี้ยงปลาที่มีปริมาณถ่านมากจะมีคุณภาพน้ำต่ำกว่าน้ำในตู้เลี้ยงปลาที่มีปริมาณถ่านน้อย
 4. ถ้าถ่านมีผลต่อคุณภาพน้ำในตู้เลี้ยงปลา ดังนั้นน้ำในตู้เลี้ยงปลาที่มีปริมาณถ่านมากจะมีค่าการละลายในน้ำของออกซิเจนสูงกว่าน้ำในตู้เลี้ยงปลาที่มีปริมาณถ่านน้อย
4. นักเรียนคนหนึ่งตั้งวัตถุประสงค์ของโครงการว่าเพื่อเปรียบเทียบปริมาณเอ็นไซม์เซลลูเลสจากเชื้อราต่างสายพันธุ์ นักเรียนคนนี้ควรตั้งสมมติฐานการศึกษาว่าอย่างไร
1. เชื้อราแต่ละสายพันธุ์มีปริมาณเอ็นไซม์เซลลูเลสต่างกัน
 2. เชื้อราสายพันธุ์เดียวกันจะให้ปริมาณเอ็นไซม์แต่ละชนิดเท่ากัน
 3. อัตราส่วนของเอ็นไซม์แต่ละชนิดที่พบขึ้นอยู่กับสายพันธุ์ของเชื้อรา
 4. การเจริญเติบโตของเชื้อราแต่ละสายพันธุ์มีผลต่อปริมาณเอ็นไซม์เซลลูเลส

5. ชาวบ้านที่อาศัยบริเวณที่มีการทำนาทุ่งพบว่ามีการแพร่กระจายของเกลือจากนาทุ่ง ทำให้เกิดผลกระทบกับสิ่งแวดล้อม จึงได้เลือกพืชทะเล 4 ชนิด คือ ผักบุ้งทะเล, ถั่วทะเล, ปอทะเล และ หูกวาง มาทดสอบประสิทธิภาพการดูดซับเกลือ นักเรียนคิดว่าชาวบ้านกำลังพิสูจน์สมมติฐานตามข้อใด
1. พืชทะเลทั้ง 4 มีความเหมาะสมในการนำมาปลูกบริเวณที่มีการทำนาทุ่ง
 2. พืชที่ปลูกบริเวณชายฝั่งทะเลสามารถเจริญเติบโตได้ดีกว่าพืชที่ปลูกในดินเลนจากก้นบ่อทุ่ง
 3. พืชที่เจริญเติบโตได้ดีบริเวณชายฝั่งทะเลสามารถลดการแพร่กระจายของเกลือที่เกิดจากการทำนาทุ่ง
 4. ประสิทธิภาพในการดูดซับความเค็มของพืชทะเลแต่ละชนิดขึ้นอยู่กับชนิดของเกลือที่แพร่กระจายจากนาทุ่ง
6. ชาวนากลุ่มหนึ่งได้นำก้อนเชื้อสำหรับเพาะเห็ดที่ผลิตจากวัสดุต่าง ๆ คือ ฟางข้าว ขุยมะพร้าว ขี้เลื่อยและยางพารา มาใช้ในการเพาะเห็ดฟางแล้วนับระยะเวลาตั้งแต่เริ่มเพาะจนเห็ดให้ดอก ชาวนากลุ่มนี้กำลังทดสอบสมมติฐานตามข้อใด
1. ลักษณะของก้อนเชื้อต่างกันมีผลทำให้อายุการให้ดอกเห็ดต่างกัน
 2. วัสดุที่ใช้ทำก้อนเชื้อที่เหมาะสมมีผลทำให้จำนวนของดอกเห็ดมากขึ้น
 3. วัสดุที่ใช้ทำก้อนเชื้อต่างกันมีผลต่อระยะเวลาตั้งแต่เริ่มเพาะจนเห็ดให้ดอก
 4. วัสดุที่เหมาะสมที่สุดในการทำก้อนเชื้อคือฟางข้าว รองลงมาคือขี้เลื่อย ขางพารา และขุยมะพร้าวตามลำดับ
7. สับปะรดเป็นผลไม้ที่มีเอนไซม์โบรมีเลน ซึ่งสามารถย่อยเคซีนในน้ำนมได้ ถ้านักโภชนาการต้องการศึกษาปริมาณเอนไซม์โบรมีเลนในส่วนต่าง ๆ ของผลสับปะรด ควรตั้งสมมติฐานอย่างไร
1. เนื้อสับปะรดจะมีปริมาณเอนไซม์โบรมีเลนมากกว่าเปลือกและแกนสับปะรด
 2. ยิ่งน้ำหนักของสับปะรดมากเท่าใด ยิ่งทำให้มีปริมาณเอนไซม์โบรมีเลนมากขึ้นเท่านั้น
 3. เอนไซม์โบรมีเลนที่อยู่ในเนื้อสับปะรดจะย่อยเคซีนในน้ำนมได้ดีกว่าเอนไซม์โบรมีเลนที่อยู่ในเปลือกและแกนสับปะรด
 4. ประสิทธิภาพในการย่อยเคซีนในน้ำนมของเอนไซม์โบรมีเลนที่ได้จากเนื้อสับปะรดจะดีกว่าที่ได้จากเปลือกและแกนสับปะรด

8. ในการศึกษาสมบัติของพืชท้องถิ่น นักวิจัยสนใจการทำน้ำกรดจากพืชท้องถิ่นเพื่อใช้เป็นวัตถุดิบในการผลิตยางพารา โดยใช้เปลือกสับปะรด เปลือกเงาะ ใบมะขามและใบชะมวง มาต้มแล้วหมักเพื่อวัดค่า pH นักวิจัยกลุ่มนี้ควรตั้งสมมติฐานการทดลองอย่างไร
1. เมื่อนำพืชท้องถิ่นแต่ละชนิดที่มีค่า pH ต่างกันมาผลิตยางพาราจะได้ยางพาราที่มีสมบัติต่างกัน
 2. พืชท้องถิ่นแต่ละชนิดสามารถนำมาใช้เป็นวัตถุดิบในการทำยางพาราได้ต่างกัน
 3. หลังการหมักเปลือกสับปะรด เปลือกเงาะ ใบมะขามและใบชะมวงจะให้ของเหลวที่มีค่า pH ต่างกัน
 4. เมื่อนำเปลือกสับปะรด เปลือกเงาะ ใบมะขามและใบชะมวงมาใช้เป็นวัตถุดิบในการผลิตยางพาราจะได้ยางพาราที่มีสมบัติต่างกัน
9. ในการทดสอบเพื่อเปรียบเทียบประสิทธิภาพของฉนวนกันความร้อนใต้หลังคาโดยสร้างบ้านจำลอง 2 หลัง ที่มีลักษณะเหมือนกัน แต่หลังที่ 1 ใช้ฉนวนกันความร้อนใต้หลังคา และหลังที่ 2 ไม่ใช้ฉนวนกันความร้อน ตั้งบ้านทั้งสองหลังให้ได้รับความร้อนจากหลอดไฟหลอดเดียวกัน ซึ่งมีระยะห่างจากหลังคาบ้านแต่ละหลังกับหลอดไฟและระยะเวลาที่บ้านได้รับความร้อนจากหลอดไฟเท่ากัน นักเรียนคิดว่าในการทดสอบข้างต้นสามารถเปรียบเทียบประสิทธิภาพของฉนวนกันความร้อนได้จากสิ่งใด
1. อุณหภูมิของฉนวน
 2. อุณหภูมิภายในบ้าน
 3. แสงสว่างภายในบ้าน
 4. อุณหภูมิของหลังคาภายนอกบ้าน
10. ในการศึกษาชนิดของปุ๋ยพืชสดที่มีผลต่อการเพิ่มคุณภาพดิน โดยชนิดของปุ๋ยพืชสดเป็นตัวแปรต้น และการเพิ่มคุณภาพดินเป็นตัวแปรตาม นักเรียนสามารถกำหนดนิยามเชิงปฏิบัติการตัวแปรตามอย่างไร
1. เปรียบเทียบลักษณะของต้นโพธิ์หลังจากใส่ปุ๋ยพืชสดกับปุ๋ยเคมี
 2. ความเหมาะสมของดินในการปลูกพืช โดยนับจำนวนครั้งในการใส่ปุ๋ยพืชสด
 3. ปริมาณแร่ธาตุในดินที่ต้นโพธิ์ดูดซึมไปใช้วัดโดยการชั่งน้ำหนักของดินที่หายไป
 4. ความสามารถในการย่อยสลายสารอินทรีย์ในปุ๋ยพืชสด โดยนับระยะเวลาที่ปุ๋ยพืชสดใช้ในการย่อยสลาย

11. ในการทดลองนำเมล็ดธัญพืชชนิดต่าง ๆ มาทำเป็นถุงเก็บความร้อนที่ใช้ในการบรรเทาอาการปวดเมื่อย วิธีใดเหมาะสมที่สุดในการวัดระดับความร้อนของถุงเก็บความร้อน
1. วัดเวลาที่ทำให้น้ำแข็ง 1 ก้อน ละลายขณะวางบนถุงเก็บความร้อน
 2. วัดเวลาที่ทำให้อุณหภูมิของถุงเก็บความร้อนเพิ่มขึ้น 5°C ขณะวางกลางแดด
 3. วัดเวลาที่ทำให้อุณหภูมิของถุงเก็บความร้อนขณะอยู่ในไมโครเวฟเพิ่มขึ้นถึง 100°C
 4. วัดเวลาที่ทำให้อุณหภูมิของถุงเก็บความร้อนลดลงจนถึงอุณหภูมิห้องหลังจากนำออกนอกไมโครเวฟ
12. นักศึกษากลุ่มหนึ่งได้ศึกษาชนิดของกระดาษที่นำมาผสมกับกาวและไส้มันสำปะหลังแล้ว ทำให้กระดาษประดิษฐ์มีการลอยน้ำได้ดีที่สุด นักศึกษากลุ่มนี้ควรตรวจสอบการลอยน้ำของกระดาษอย่างไร
1. วัดระยะทางที่กระดาษสามารถเคลื่อนที่ไปได้
 2. เมื่อปล่อยกระดาษลงน้ำแล้ววัดอัตราการเคลื่อนที่ของกระดาษ
 3. เมื่อปล่อยกระดาษลงน้ำแล้ววัดความสูงของกระดาษที่อยู่เหนือน้ำ
 4. จับเวลาตั้งแต่ผิวของกระดาษสัมผัสน้ำจนผิวของกระดาษทั้งหมดจมน้ำ
13. ในการศึกษาเกี่ยวกับการดองผลไม้ชนิดต่าง ๆ มีการตั้งสมมติฐานว่า “ถ้าชนิดของผลไม้มีผลต่อความเป็นกรดของผลไม้ดอง ดังนั้นมะม่วงจะให้ความเป็นกรดมากที่สุด รองลงไปคือ มะดัน มะกอก และฝรั่ง ตามลำดับ” วิธีใดเหมาะสมที่สุดในการวัดความเป็นกรดของผลไม้ดอง
1. ใช้ pH มิเตอร์
 2. สอบถามจากผู้ชิม
 3. การเปลี่ยนสีของกระดาษลิตมัส
 4. เปรียบเทียบกับรสชาติของมะนาว

14. ในการศึกษาเพื่อทดสอบสมมติฐานที่ว่า “วัสดุจากธรรมชาติชนิดต่าง ๆ มีสมบัติในการดูดซับน้ำมันได้ต่างกัน” โดยนำวัสดุจากธรรมชาติที่ต้องการศึกษา คือ แกลบ จี้เลื่อย และธูปฤาษี อย่างละ 500 g. ใส่ในถุงตาข่ายแล้วนำไปแช่ในบีกเกอร์แต่ละใบที่บรรจุน้ำมันดีเซลปริมาตร 1 dm^3 ทิ้งไว้ 30 นาที แล้วนำไปชั่ง เพื่อเปรียบเทียบความสามารถในการดูดซับน้ำมัน นักเรียนควรนิยามคำว่า “ความสามารถในการดูดซับน้ำมัน” อย่างไรจึงจะสามารถวัดได้ตรงกัน

1. เวลาที่ใช้ในการดูดซับน้ำมันปริมาตร 1 dm^3
2. น้ำหนักของวัสดุจากธรรมชาติที่เพิ่มขึ้นหลังการทดลอง
3. ปริมาตรของน้ำมันที่วัสดุจากธรรมชาติสามารถกักเก็บได้
4. จำนวนชนิดของน้ำมันพืชที่วัสดุจากธรรมชาติสามารถดูดซับได้

15. ในโครงการเรื่อง “การลดปริมาณโลหะหนักในน้ำทิ้งด้วยดินเหนียวบดละเอียด” ควรให้นิยามเชิงปฏิบัติการคำที่ขีดเส้นใต้ได้อย่างไร

1. มวลของดินเหนียวบดละเอียดที่เพิ่มขึ้น หลังการปล่อยให้ น้ำทิ้ง ไหลผ่านดินเหนียวบดละเอียด
2. มวลของน้ำที่ลดลงหลังการปล่อยให้ น้ำ ไหลผ่านดินเหนียวบดละเอียด
3. น้ำหนักของของแข็งที่อยู่ในน้ำหลังการปล่อยให้ น้ำ ไหลผ่านดินเหนียวบดละเอียด
4. ปริมาณของ Cu^{2+} , Pb^{2+} , Fe^{2+} และ Fe^{3+} ในน้ำที่ลดลงหลังการไหลผ่านดินเหนียวบดละเอียด

16. เกษตรกรได้คิดแก้ปัญหาหน้าในบ่อปลาที่มีความเป็นกรดเบสไม่เหมาะสม เนื่องจากเป็นบ่อที่สร้างด้วยปูนซีเมนต์ โดยทดลองใช้ใบหิ่เหล็ก ใบมะเฟือง และใบมะยม มาทดสอบผลการปรับสภาพน้ำให้เป็นกลาง เกษตรกรควรจะให้คำนิยามเชิงปฏิบัติการคำว่า “การปรับสภาพน้ำ” อย่างไร

1. การทำให้น้ำเปลี่ยนสภาพจากเบสเป็นกลาง
2. การทำให้น้ำมีค่าความเป็นกรด-เบสเหมาะสม
3. การทำให้ค่า pH ซึ่งวัดจาก pH มิเตอร์ มีค่าใกล้เคียง 7
4. การทำให้สมบัติทางกายภาพของน้ำมีคุณภาพมากขึ้น

17. ในการศึกษาการเกิดปฏิกิริยาเคมีระหว่างโลหะกับสารละลายชนิดต่าง ๆ โดยนำโลหะแต่ละชนิด ที่มีขนาด 1×1 นิ้ว ใส่งในสารละลายที่ต้องการทดสอบแล้วสังเกตตะกอนและแก๊สที่เกิดขึ้น รวมทั้งสีและกลิ่นที่เปลี่ยนแปลง ในการศึกษาครั้งนี้นักเรียนสามารถให้คำนิยามเชิงปฏิบัติการ คำว่า “ การเกิดปฏิกิริยาเคมี ” ได้อย่างไร
1. การเปลี่ยนแปลงที่มีสารใหม่เกิดขึ้น
 2. การเปลี่ยนสารตั้งต้นให้เป็นผลิตภัณฑ์
 3. การเปลี่ยนแปลงที่ทำให้เกิดตะกอนหรือแก๊ส หรือทำให้สีและกลิ่นเปลี่ยนไป
 4. การเปลี่ยนแปลงที่ทำให้โลหะและสารละลายกลายเป็นสารอื่นที่มีสมบัติแตกต่างจากเดิม
18. นักเรียนกลุ่มหนึ่งได้ร่วมกันรณรงค์ในโครงการกำจัดยุงลายจึงนำเสนอโครงการ เรื่อง “ การกำจัดยุงโดยใช้พืชสมุนไพร ” พบว่าการใช้สะเดาสามารถทำลายลูกน้ำของยุงได้ผลดีที่สุด จากโครงการข้างต้น ข้อใดเป็นตัวแปรต้น
1. ปริมาณยุงลาย
 2. วิธีการใช้สะเดา
 3. ชนิดของพืชสมุนไพร
 4. ผลการใช้สะเดาทำลายลูกน้ำ
19. พรเทพต้องการทดสอบสมมติฐานที่ว่า “ ถ้าเครื่องบินกระดาษมีพื้นที่ปีกมากขึ้นจะทำให้มีระยะการกระจัดมากขึ้น ” ในการทดสอบสมมติฐานนี้ ตัวแปรต้น คือข้อใด
1. ระยะการกระจัด
 2. พื้นที่ปีกของเครื่องบินกระดาษ
 3. ระยะเวลาที่ลอยตัวอยู่ในอากาศ
 4. พื้นที่กระดาษที่ใช้ทำเครื่องบิน
20. แม่ค้าขายส้มตำสนใจที่จะทำการทดลองตามสมมติฐานว่า “ วิธีการเก็บรักษามะละกอที่ต่างกัน อาจทำให้ชะลอการสุกของมะละกอได้ ” จึงแบ่งมะละกอดิบออกเป็น 4 กลุ่ม โดยแต่ละกลุ่มต้องควบคุมสิ่งต่าง ๆ ให้เหมือนกัน ยกเว้น ข้อใด
1. อายุของมะละกอ
 2. พันธุ์ของมะละกอ
 3. วิธีการเก็บรักษามะละกอ
 4. จำนวนผลมะละกอที่ใช้ในแต่ละกลุ่ม

21. จากประสบการณ์ของนักเรียนคนหนึ่ง พบว่าการเพาะเมล็ดพืชบริเวณที่มีแสงสว่างและไม่มีแสงสว่าง เมล็ดพืชก็สามารถงอกได้ แต่นักเรียนสนใจจะศึกษาว่าปริมาณแสงมีผลต่อระยะเวลาในการงอกของเมล็ดพืชหรือไม่ จึงทดลองโดยใช้เมล็ดถั่วเขียว การทดลองนี้ควรจะกำหนดตัวแปรต่าง ๆ อย่างไร

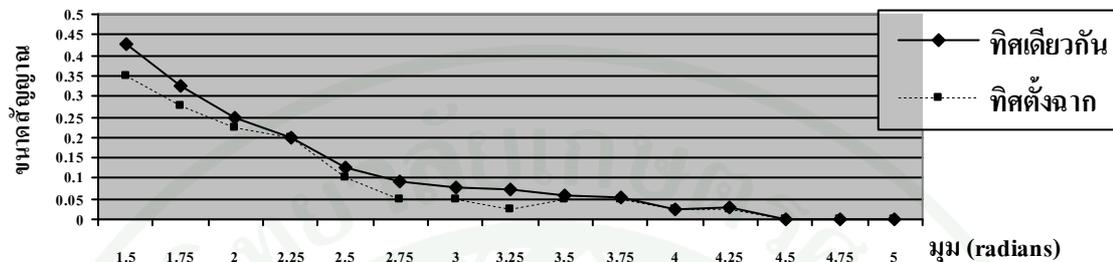
1. ตัวแปรต้น คือ แสง
ตัวแปรตาม คือ เวลาในการงอกของเมล็ดถั่วเขียว
2. ตัวแปรต้น คือ แสง
ตัวแปรตาม คือ อัตราการงอกของเมล็ดถั่วเขียว
3. ตัวแปรต้น คือ สภาพแวดล้อมในการเพาะเมล็ดถั่วเขียว
ตัวแปรตาม คือ อัตราการงอกของเมล็ดถั่วเขียว
4. ตัวแปรต้น คือ สภาพแวดล้อมในการเพาะเมล็ดถั่วเขียว
ตัวแปรตาม คือ อัตราเร็วในการงอกของเมล็ดถั่วเขียว

22. ในการทดลองผลิตไข่เค็มเพื่อหาเวลาที่เหมาะสมในการหมักไข่เค็ม โดยนำไข่มาพอกดิน แล้วแบ่งเป็น 2 กลุ่ม คือ กลุ่มที่แช่กรดอะซิติกและไม่แช่กรดอะซิติก นำไปเรียงในกล่อง เพื่อรอเวลานำไข่ไปต้ม จากการทดลองข้างต้น ข้อใดเป็นตัวแปรต้น

1. การแช่ไข่ในกรดอะซิติก
2. ชนิดของไข่ที่ใช้ทำไข่เค็ม
3. ระยะเวลาการแช่ไข่ในกรดอะซิติก
4. ความเข้มข้นของกรดอะซิติกที่ใช้แช่ไข่เค็ม

23. ในการศึกษาความสัมพันธ์ระหว่างขนาดสัญญาณที่ได้จากลำโพงประดิษฐ์กับมุมต่าง ๆ ของระนาบการสั่น เมื่อวางแม่เหล็กของลำโพงให้มีลักษณะต่างกันได้ผลดังกราฟ ข้อใดเป็น ตัวแปรต้น

กราฟเปรียบเทียบขนาดสัญญาณที่วัดที่มุมต่าง ๆ เมื่อวางแม่เหล็ก
มีทิศสัมพันธ์กับแรงโน้มถ่วงของโลก



1. ขนาดสัญญาณ
2. การวางตัวของแม่เหล็ก
3. ประสิทธิภาพของลำโพง
4. ความเข้มของสนามแม่เหล็ก

24. กระเทียมมีคุณสมบัติในการยับยั้งการเจริญเติบโตของแบคทีเรีย เพื่อนำผลการศึกษาไปใช้ประโยชน์ในด้านต่าง ๆ จึงมีการทดลองนำน้ำมันหอมระเหยจากกระเทียมพันธุ์เดียวกัน ปริมาตร 2 cm^3 ใส่จานเพาะเลี้ยงจำนวนเท่ากับชนิดของแบคทีเรียที่ต้องการทดลองในปริมาณที่เท่ากัน บันทึกการเจริญเติบโตของแบคทีเรียหลังจากทิ้งไว้ 1, 2, 3 และ 4 สัปดาห์ การทดลองครั้งนี้มี ตัวแปรต้น และ ตัวแปรตาม คืออะไร

1. พันธุ์ของกระเทียม , ระยะเวลาในการทดลอง
2. ปริมาณน้ำมันหอมระเหย , การเจริญเติบโตของแบคทีเรีย
3. ชนิดของแบคทีเรีย , ประสิทธิภาพการยับยั้งการเจริญเติบโตของแบคทีเรีย
4. พันธุ์ของกระเทียม , ประสิทธิภาพการยับยั้งการเจริญเติบโตของแบคทีเรีย

25. นักเรียนกลุ่มหนึ่งได้รายงานผลการทดลองว่า “ เวลาที่ใช้ในการงอกของเห็ดจากแปลงใบสับปะรดมากกว่าแปลงผักตบ และมวลรวมของดอกเห็ดจากแปลงสับปะรดจะน้อยกว่าแปลงผักตบ ” การทดลองของนักเรียนกลุ่มนี้มีสิ่งใดเป็น ตัวแปรตาม

1. มวลของใบสับปะรดและมวลของผักตบ
2. ขนาดและจำนวนของดอกเห็ดที่ได้หลังการทดลอง
3. วัสดุที่ใช้ทำแปลงเพาะเห็ดและเวลาที่ใช้ในการเพาะเห็ด
4. เวลาที่ใช้ในการงอกของเห็ดและมวลรวมของดอกเห็ด

26. ในการศึกษาอิทธิพลของคลื่นเสียงต่อการงอกของเมล็ดข้าวโพด พบว่าที่ความถี่เสียง 1,000 เฮิร์ต ระดับความเข้มเสียง 80 เดซิเบล เปอร์เซนต์การงอกของเมล็ดข้าวโพดมีมากที่สุด คือ 84 % และเปอร์เซนต์การงอกของเมล็ดข้าวโพดที่ความถี่เสียง 400 เฮิร์ต ระดับความเข้มเสียง 90 เดซิเบล มีน้อยที่สุด คือ 28 % การทดลองครั้งนี้ต้องควบคุมตัวแปรใด
1. แหล่งกำเนิดเสียง
 2. ระดับความถี่เสียง
 3. ระดับความเข้มเสียง
 4. ระยะเวลาที่ให้เสียง
27. ในการศึกษาปริมาณของเกลือที่ใช้ในการดองจะทำให้ผลิตภัณฑ์ที่ได้จากการดองมีความเป็นกรด-เบสต่างกันหรือไม่ ตัวแปรควบคุม ในการศึกษาครั้งนี้คือข้อใด
1. ชนิดของวัตถุดิบที่ใช้ในการดอง, ระยะเวลาที่ใช้ในการดอง
 2. ปริมาณวัตถุดิบที่ใช้ในการดอง, ปริมาณเกลือที่ใช้ในการดอง
 3. ปริมาณเกลือที่ใช้ในการดอง, ขนาดภาชนะที่ใช้ในการดอง
 4. ระยะเวลาที่ใช้ในการดอง, ความเป็นกรด-เบสของผลิตภัณฑ์
28. กลุ่มแม่บ้านประจำตำบลต้องการศึกษาว่า “ สารสกัดที่ได้จากฟ้าทะลายโจร บอระเพ็ด สะเดา และว่านหางจระเข้ สามารถป้องกันการเกิดหนองของปลาเค็มได้หรือไม่” นักเรียนคิดว่าตัวแปรใดที่มีผลต่อการทดลองน้อยที่สุด
1. ชนิดของพืช
 2. ชนิดของปลาเค็ม
 3. ช่วงเวลาในการเก็บพืช
 4. ปริมาณของสารสกัดที่ได้จากพืช
29. ปัจจุบันกล้วยไม้เป็นพืชเศรษฐกิจที่สร้างรายได้เข้าประเทศอย่างมาก โดยกล้วยไม้ต้องการความชื้นในการเจริญเติบโต แต่ถ้าขาดความชื้นหรือมีความชื้นมากเกินไปจะทำให้กล้วยไม้ตาย ดังนั้นจึงมีการศึกษาเกี่ยวกับเรื่องนี้ นักเรียนคิดว่าตัวแปรต้นในการศึกษานี้คือข้อใด
1. ปริมาณความชื้น
 2. รายได้เข้าประเทศ
 3. การเจริญเติบโตของกล้วยไม้
 4. สภาพแวดล้อมที่เหมาะสมต่อการปลูกกล้วยไม้

30. กิ่งทราบว่ายางเป็ญฟูนลยอยูบบนผวน้ำได้เพราะน้ำมีแรงดึงผวน เขาจึงทดลองนำของเหลวชนิดต่าง ๆ มาทดสอบว่าของเหลวต่างชนิดกันจะมีแรงดึงผวนเหมือนหรือต่างจากน้ำอย่างไร กิ่งต้องวางแผนการทดลองอย่างไร

1. นำน้ำสะอาดใส่ในบีกเกอร์ปริมาตร 50, 100, 150 และ 200 cm³. แล้วโรยเป็ญฟูนจำนวน 1 ซ้อน เบอร์ 2 ลงในบีกเกอร์แต่ละใบพร้อมกัน
2. นำน้ำสะอาด น้ำสบู่ น้ำผงซักฟอก และน้ำแชมพูปริมาตร 150 cm³. ใส่ในแก้วน้ำทั้ง 4 ใบ ซึ่งมีขนาดเท่ากันแล้วโรยเป็ญฟูนจำนวน 1 ซ้อน เบอร์ 2 ลงในแก้วพร้อมกัน
3. นำน้ำสะอาดที่มีอุณหภูมิ 10°C, 30°C, 50°C และ 70°C ใส่ในแก้วน้ำทั้ง 4 ใบ ใบละ 150 cm³. แล้วโรยเป็ญฟูนจำนวน 1 ซ้อน เบอร์ 2 ลงในแก้วพร้อมกัน
4. นำน้ำสะอาดปริมาตร 200 cm³. ใส่ในบีกเกอร์จำนวน 4 ใบ แล้วใช้ซ้อนเบอร์ 2 ตักเป็ญฟูน เป็ญมัน เป็ญข้าวโพดและเป็ญข้าวเหนียวอย่างละ 1 ซ้อน โรยลงในบีกเกอร์แต่ละใบพร้อมกัน

31. ถ้านักเรียนต้องการใช้สารละลายชนิดหนึ่งปริมาตร 50 cm³ นักเรียนจะมีวิธีการตวงสารละลายชนิดนี้โดยใช้กระบอกตวงตามลำดับขั้นตอนต่อไปนี้ได้อย่างไร

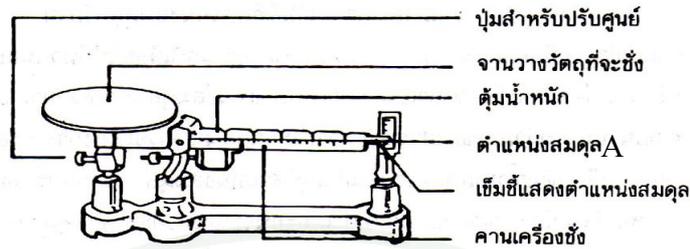
- ก. ใช้หลอดหยดปรับปริมาตรให้ส่วนโค้งเว้าต่ำสุดของสารละลายอยู่ตรงกับขีดบอกปริมาตร
- ข. ถ่ายเทสารละลายจากบีกเกอร์ลงในกระบอกตวงจนโค้งเว้าอยู่ต่ำกว่าขีดบอกปริมาตร
- ค. เทสารละลายจากขวดบรรจุสารเคมีลงในบีกเกอร์
- ง. อ่านปริมาตรของของเหลวในกระบอกตวงให้ได้ 50 cm³

1. ข ก ง ค
2. ค ข ก ง
3. ค ก ข ง
4. ค ข ง ก

32. ถ้านักเรียนต้องการดมกลิ่นสารเคมีที่บรรจุในบีกเกอร์ นักเรียนควรปฏิบัติตามข้อใด

1. ใช้มือโบกให้ไอของสารบริเวณปากบีกเกอร์ผ่านเข้าจมูกช้า ๆ
2. ใช้แท่งแก้วจุ่มสารขึ้นมาดม โดยให้ปลายแท่งแก้วอยู่ต่ำกว่าจมูก
3. ใช้แท่งแก้วแตะสารแล้วยกขึ้นมาเหนือจมูก สายแท่งแก้วไปมาช้า ๆ
4. ใช้แท่งแก้วจุ่มสารมาแตะบริเวณหลังมือที่อยู่ห่างจากจมูกพอสมควรแล้วสูดหายใจเข้าช้า ๆ

33. ขั้นตอนแรกของการใช้เครื่องชั่งดังภาพ คือข้อใด



เครื่องชั่งชนิดจานเดียว

1. กดคานเครื่องชั่งลงก่อนวางวัตถุที่ต้องการชั่ง
2. เลื่อนดัมน้ำหนักไปยังตำแหน่งที่ใกล้เคียงน้ำหนักของวัตถุ
3. เลื่อนดัมน้ำหนักไปตำแหน่งที่ทำให้เข็มชี้แสดงตำแหน่งสมดุลอยู่ที่ตำแหน่ง A
4. ตรวจสอบเข็มชี้แสดงตำแหน่งสมดุลให้อยู่ตรงกลางเมื่อดัมน้ำหนักอยู่ในตำแหน่ง 0 กรัม

34. ชัยสังเกตเห็นว่าส้มเขียวหวานที่ชาวสวนเก็บมานั้นมีสีต่างกันจึงต้องการศึกษาว่า “สีเปลือกส้มมีความสัมพันธ์กับปริมาณวิตามินซีที่มีอยู่ในผลส้มหรือไม่” ชัยควรคัดเลือกผลส้มที่ใช้ในการทดลองตามข้อใด

1. นำส้มที่มีสีต่างกันตามสายพันธุ์มาทดสอบอย่างน้อย 4 ชนิด
2. ซื้อมะม่วงจากตลาดที่มีเปลือกสีส้ม สีเขียว สีส้มมากกว่าสีเขียว และสีเขียวกว่าสีส้ม
3. นำส้มจากต้นเดียวกันแต่มีเปลือกสีส้ม สีเขียว สีส้มมากกว่าสีเขียว และสีเขียวกว่าสีส้ม
4. นำส้มจากสวนเดียวกันแต่คนละต้นที่มีเปลือกสีส้ม สีเขียว สีส้มมากกว่าสีเขียว และสีเขียวกว่าสีส้ม

35. ในการศึกษาการผลิตไวน์มะพร้าว น้ำหอมติดอยู่บนต้นโดยเปรียบเทียบปริมาณแอลกอฮอล์ ปริมาณน้ำตาล ปริมาณกรดและค่า pH ของไวน์มะพร้าว น้ำหอมที่ผลิตขึ้น ณ สถานที่ที่แตกต่างกัน คือ น้ำมะพร้าวใส่ในขวดรูปชมพู่ น้ำมะพร้าวที่อยู่ในผลนอกต้นและน้ำมะพร้าวที่อยู่ในผลมะพร้าวติดอยู่บนต้น ขั้นตอนใดจำเป็นในการทดลองนี้

1. การเติมปริมาณหัวเชื้อที่ต่างกัน
2. การคัดเลือกขนาดมะพร้าวที่ต่างกัน
3. การคัดเลือกพันธุ์มะพร้าวที่เหมือนกัน
4. การคัดเลือกมะพร้าวที่มีเปลือกหนาเท่ากัน

36. คำกำลังทดสอบสมมติฐานที่ว่า “ ถ้าความยาวของเชือกมีผลต่อการแกว่งของลูกตุ้ม ดังนั้นเส้นเชือกยาวจะใช้เวลาในการแกว่งมากกว่าเส้นเชือกสั้น ” โดยจัดอุปกรณ์การทดลองดังภาพ และกำหนดให้ θ แทน มุมของเชือกที่กระทำกับแนวตั้ง 30° และ l แทน ความยาวของเชือก 20, 30, 40, 50, 60, 70 และ 80 cm. คำควรบันทึกข้อมูลใดลงในตารางบันทึกผลการทดลอง



1. มวลของลูกตุ้มที่ใช้ในการแกว่งแต่ละครั้ง
 2. ระยะทางที่ลูกตุ้มเคลื่อนที่จากปลายข้างหนึ่งไปยังปลายอีกข้างหนึ่ง
 3. จำนวนรอบที่ลูกตุ้มเคลื่อนจากจุดเริ่มต้นแล้วกลับมาในตำแหน่งเดิม
 4. เวลาตั้งแต่ลูกตุ้มเคลื่อนที่จากจุดเริ่มต้นแล้วกลับมาในตำแหน่งเดิมครบ 20 รอบ
37. ถ้านักเรียนต้องการทำให้สารที่มีสถานะของเหลว 2 ชนิด ผสมกันในหลอดทดลองโดยการเขย่าสาร นักเรียนจะมีวิธีการเขย่าสารอย่างไร
1. ปิดจุกหลอดทดลองแล้วใช้มือจับหลอดทดลองยกขึ้น-ลงหลาย ๆ ครั้ง
 2. ปิดจุกหลอดทดลองแล้วคว่ำพลิกหลอดทดลองไปทางซ้าย-ขวาหลาย ๆ ครั้ง
 3. จับหลอดทดลองบริเวณส่วนบนแล้วหมุนหลอดทดลองไปในทิศทางตามเข็มนาฬิกา
 4. จับหลอดทดลองด้านบนแล้วกระแทกหลอดทดลองส่วนล่างกับฝ่ามืออีกข้างหนึ่งเบาๆ
38. ถ้านักเรียนต้องการทดสอบสมมติฐานที่ว่า “ สารละลายชนิดต่าง ๆ เมื่อมีความเข้มข้นต่างกัน จะมีค่าดัชนีหักเหแสงต่างกันด้วย ” นักเรียนควรเตรียมสารละลายตามข้อใด
1. สารละลายกลูโคส, สารละลายซูโครส และสารละลายเกลือแกงที่มีความเข้มข้นชนิดละ 25%
 2. สารละลายกลูโคส เข้มข้น 25 %, สารละลายซูโครส เข้มข้น 50 % และสารละลายเกลือแกง เข้มข้น 75 %
 3. สารละลายกลูโคส เข้มข้น 25 %, 50 %, 75 % และ 100 % ที่มีปริมาตร 100, 200, 300 และ 400 cm^3 . ตามลำดับ
 4. สารละลายกลูโคส เข้มข้น 25 %, 50 %, 75 %, 100 % และสารละลายซูโครส, สารละลายเกลือแกง ที่มีความเข้มข้นเท่ากับสารละลายกลูโคส

39. ใบของต้นขอบชะนางทั้ง 2 ชนิด คือ ขอบชะนางแดงและขอบชะนางขาว มีสรรพคุณแก้ ริดสีดวงทวารและกำจัดพยาธิตานซาง ถ้าต้องการทดสอบเพื่อเปรียบเทียบน้ำหนักของสารสกัด ที่ได้จากใบขอบชะนางทั้ง 2 ชนิด เมื่อใช้ตัวทำละลายชนิดต่าง ๆ โดยสกัดด้วยเครื่องมือชนิดหนึ่ง ขั้นตอนใดไม่จำเป็นในการทดลอง

1. ชั่งน้ำหนักสารสกัดที่ได้จากใบขอบชะนางทั้ง 2 ชนิด
2. สกัดสารจากใบขอบชะนางแดงและชะนางขาวที่มีน้ำหนักเท่ากัน
3. สกัดสารจากใบขอบชะนางทั้ง 2 ชนิดด้วยเฮกเซน คลอโรฟอร์มและเมทานอล
4. หยอดสารสกัดที่ได้จากใบขอบชะนางทั้ง 2 ชนิด ลงในจานเพาะเลี้ยงที่มีพยาธิตานซาง

40. ต้นกระพังโหมเป็นวัสดุที่ชาวบ้านสนใจนำมาทำเครื่องจักสานแทนไม้ไผ่และหวาย จึงมีการ ทดลองนำต้นกระพังโหมความยาว 20 cm. ถ่วงด้วยลูกตุ้มน้ำหนักขนาด 500 g. ได้ผลดังตาราง

จำนวนลูกตุ้มที่ใช้ถ่วงน้ำหนัก (อัน)	ความยาวของลำต้นกระพังโหมหลัง การถ่วงด้วยลูกตุ้ม (cm.)
1	20
2	20
3	20
4	30
5	ต้นกระพังโหมขาด

จากการทดลองควรสรุปเกี่ยวกับต้นกระพังโหมอย่างไร

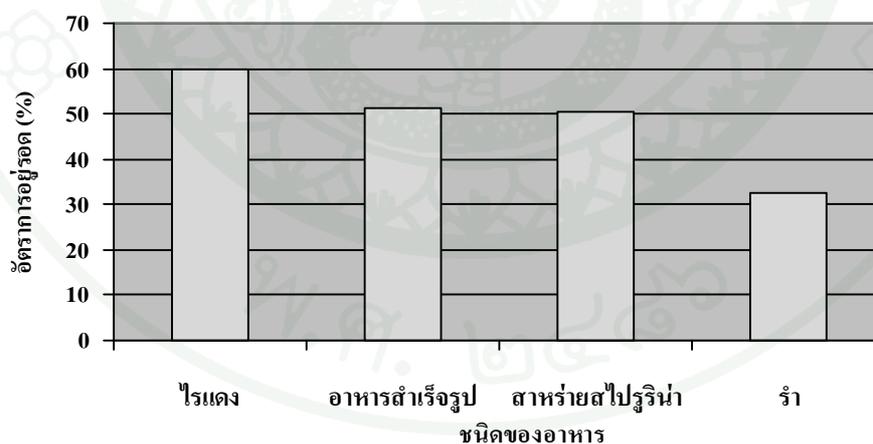
1. ลูกตุ้มน้ำหนักสามารถทำให้ลำต้นกระพังโหมมีความยาวเพิ่มขึ้น 10 เซนติเมตร
2. ต้นกระพังโหมสามารถรับน้ำหนักได้มากที่สุดประมาณ 2.0 กิโลกรัม
3. ไม่ควรนำต้นกระพังโหมมาใช้ทำเครื่องจักสานแทนไม้ไผ่และหวาย
4. สามารถนำต้นกระพังโหมมาทำเครื่องจักสานเพื่อบรรจุลูกตุ้ม 3 อัน

41. ปัจจุบันมีการปนเปื้อนของโลหะหนักที่เป็นพิษหลายชนิดในสิ่งแวดล้อม เช่น ตะกั่ว แคดเมียม นักวิจัยจึงสังเคราะห์สารเคมีที่เรียกว่า Sodalite ให้มีคุณสมบัติในการดูดซับโลหะไอออนเนื่องจากโครงสร้างเป็นรูพรุนและผลึกแข็งอย่างเป็นระเบียบ และได้นำ Sodalite ไปทดสอบการดูดซับโลหะ Pb^{2+} และ Cd^{2+} ที่ปนเปื้อนในน้ำ ได้ผลดังตาราง

โลหะไอออน	ค่าเฉลี่ยร้อยละของการดูดซับ
Pb^{2+}	99
Cd^{2+}	85

จากผลการทดลอง ข้อสรุปใดถูกต้อง

1. Sodalite ดูดซับ Pb^{2+} ได้ดีกว่า Cd^{2+}
 2. Sodalite ให้มีคุณสมบัติในการดูดซับโลหะไอออนได้ดี
 3. Sodalite มีประสิทธิภาพในการกำจัดโลหะหนักที่เป็นพิษหลายชนิดในสิ่งแวดล้อม
 4. Sodalite มีโครงสร้างเป็นรูพรุนและผลึกแข็งอย่างเป็นระเบียบตรงตามที่นักวิจัยต้องการ
42. กลุ่มเกษตรกรได้ศึกษาอัตราการอยู่รอดของกึ่งฝอยหลังการเลี้ยงด้วยอาหารชนิดต่าง ๆ ได้ผลดังกราฟ



จากกราฟข้างต้น ข้อใดถูกต้อง

1. การเลี้ยงกึ่งฝอยด้วยสาหร่ายสไปรูลีนาจะทำให้กึ่งฝอยมีโปรตีนมากขึ้น
2. การเลี้ยงกึ่งฝอยด้วยอาหารสำเร็จรูป กึ่งฝอยจะเจริญเติบโตได้ดีกว่าการเลี้ยงด้วยรำ
3. การเลี้ยงกึ่งฝอยด้วยไรแดงจะทำให้กึ่งฝอยมีอัตราการเจริญเติบโตเร็วกว่าการเลี้ยงด้วยรำ
4. การเลี้ยงกึ่งฝอยด้วยรำจะทำให้กึ่งฝอยมีอัตราการตายสูงกว่าการเลี้ยงด้วยสาหร่ายสไปรูลีนา อาหารสำเร็จรูปและไรแดง

43. ชาวบ้านกลุ่มหนึ่งต้องการลดปริมาณขยะจากใบอ้อย จึงทดลองนำไปผลิตถ่านโดยใช้ใบอ้อย ผสมกับแกลบในปริมาณที่ต่างกัน แล้ววัดพลังงานความร้อนจากการเผาถ่าน ได้ผลดังตาราง

วิธีที่	ใบอ้อย (g.)	แกลบ (g.)	พลังงานความร้อน (kcal)
1	100	300	30.76
2	200	200	21.84
3	300	300	23.99
4	100	200	33.61
5	200	300	20.72
6	300	100	30.24

ชาวบ้านกลุ่มนี้ควรเลือกใช้วิธีใดในการผลิตถ่าน

1. วิธีที่ 4 เนื่องจากได้ค่าพลังงานความร้อนมากที่สุด
2. วิธีที่ 1 เนื่องจากใช้ใบอ้อยน้อยกว่าแกลบแต่ได้ค่าพลังงานความร้อนค่อนข้างมาก
3. วิธีที่ 6 เนื่องจากใช้ปริมาณใบอ้อยมากกว่าปริมาณ แกลบ แต่ได้ค่าพลังงานความร้อนมาก
4. วิธีที่ 3 เนื่องจากใช้อัตราส่วนระหว่างใบอ้อยและแกลบเท่ากับวิธีที่ 2 แต่ได้ค่าพลังงานความร้อนมากกว่า

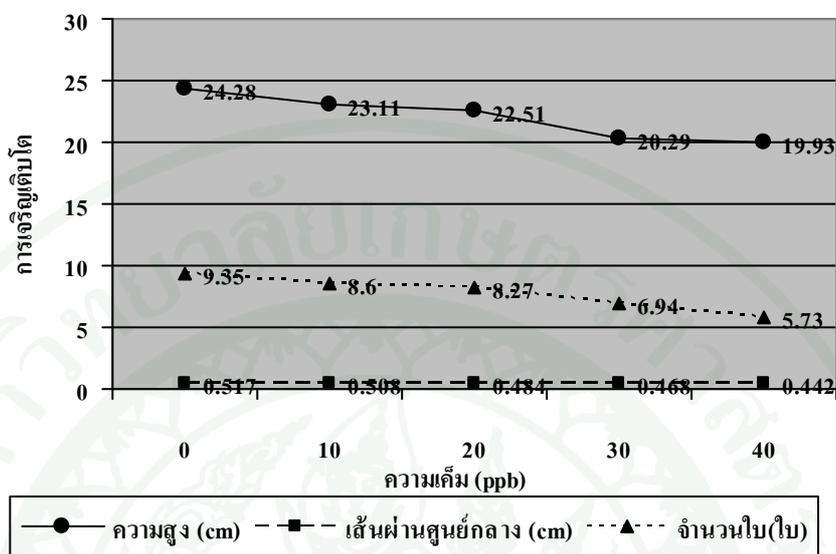
44. เครื่องเหียงสารเป็นอุปกรณ์ที่ใช้ในห้องปฏิบัติการทางเคมี เพื่อช่วยลดระยะเวลาที่ใช้ในการตกตะกอนของสารซึ่งต้องบรรจุสารลงในหลอดทดลองก่อนนำเข้าเครื่องเหียง โดยการใช้ในแต่ละครั้งสามารถใส่หลอดทดลองได้หลายหลอด ดังนั้นเพื่อเพิ่มประสิทธิภาพของเครื่องเหียง จึงได้มีการศึกษาเวลาที่ใช้ในการตกตะกอนเมื่อใส่จำนวนหลอดทดลองในเครื่องเหียงสารแตกต่างกัน ได้ผลดังตาราง

จำนวนหลอดทดลอง	ครั้งที่	เวลาที่ใช้ในการตกตะกอน (วินาที)					
		1	2	3	4	5	เฉลี่ย
2		32	29	30	31	28	30
4		40	37	42	38	43	40
6		45	44	43	46	47	45

ข้อสรุปที่ถูกต้อง

1. ปริมาณตะกอนที่ได้ขึ้นอยู่กับเวลาที่ใช้ในการตกตะกอน
2. ปริมาณสารที่ใช้ในการตกตะกอนมีผลต่อระยะเวลาที่ใช้ในการตกตะกอน
3. เครื่องเหียงสารขาดประสิทธิภาพ เพราะจากผลการทดลองใช้เวลาในการตกตะกอนไม่คงที่
4. เมื่อใส่หลอดทดลอง 2 หลอดใช้เวลาในการตกตะกอนเร็วกว่าใส่หลอดทดลอง 4 และ 6 หลอด ตามลำดับ

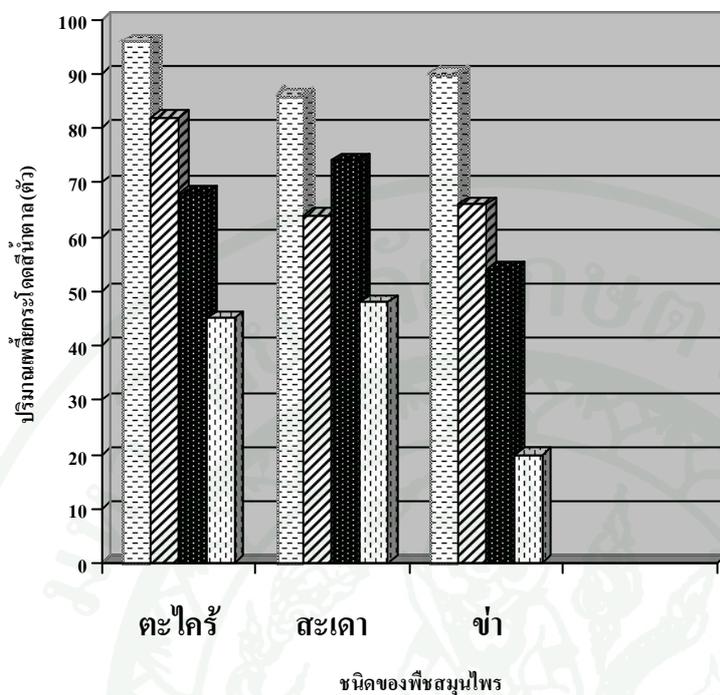
45. ต้นพังกาหัวสุมดอกแดงพบได้ในป่าชายเลนซึ่งมีระดับความเค็มแตกต่างกันในแต่ละพื้นที่ จึงมีการศึกษาความเค็มของน้ำที่มีผลต่ออัตราการเจริญเติบโตของต้นพังกาหัวสุมดอกแดงได้ผลดังกราฟ



ข้อสรุปของการศึกษาข้อใดถูกต้อง

1. ไม่สามารถสรุปได้ เนื่องจากผลของการเจริญเติบโตมีแนวโน้มไม่ชัดเจน
2. ความเค็มของน้ำมีอิทธิพลต่อการเจริญเติบโตของต้นพังกาหัวสุมดอกแดง
3. ต้นพังกาหัวสุมดอกแดงสามารถนำเกลือที่อยู่ในน้ำไปใช้ในการเจริญเติบโต
4. การเจริญเติบโตของต้นพังกาหัวสุมดอกแดงแปรผันตรงกับระดับความเค็ม

46. จากแผนภูมิแท่งเปรียบเทียบปริมาณเพ็ลยกระโดดสีน้ำตาลในนาข้าวจำลองที่เหลืออยู่หลังการฉีดพ่นด้วยสารสกัดจากพืชสมุนไพรในปริมาณต่าง ๆ ต่อหน้า 2 dm.^3 ข้อใดถูกต้อง



50 g./หน้า 2 dm.

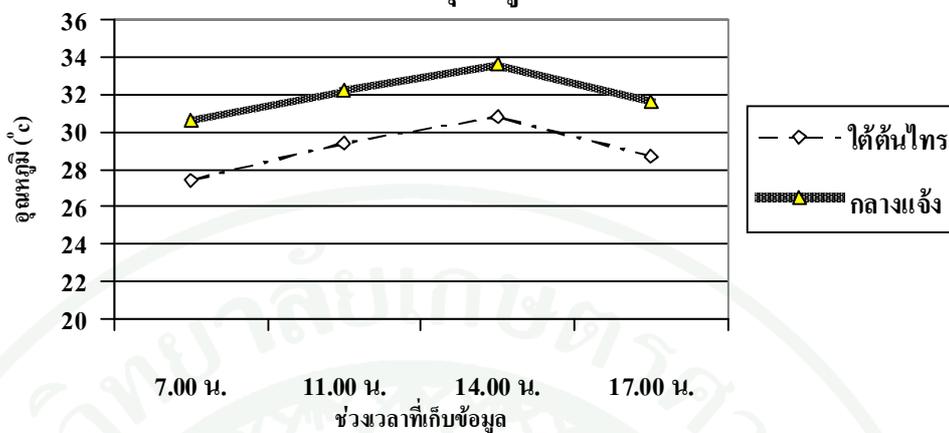
100 g./หน้า 2 dm.

150 g./หน้า 2 dm.

200 g./หน้า 2 dm.

1. ข่าสกัด 200 g. ผสมกับน้ำ 2 dm.^3 เหมาะสมที่สุดในการกำจัดเพ็ลยกระโดดสีน้ำตาล
2. ปริมาณของสารสกัดจากตะไคร้ที่เหมาะสมที่สุดในการกำจัดเพ็ลยกระโดดสีน้ำตาลคือ 50 g. ต่อ 2 dm.^3
3. ความสามารถในการกำจัดเพ็ลยกระโดดสีน้ำตาลขึ้นอยู่กับความเข้มข้นของสารสกัดจากพืชสมุนไพรเท่านั้น
4. พืชสมุนไพรต่างชนิดกันมีความสามารถในการกำจัดเพ็ลยกระโดดสีน้ำตาลได้เท่ากันถ้าใช้ความเข้มข้นของสารสกัดเท่ากัน

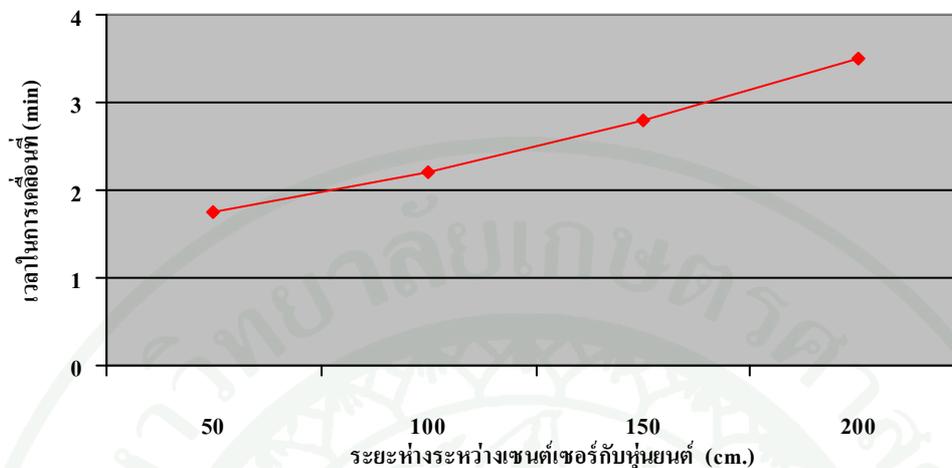
47. ในการศึกษาความแตกต่างระหว่างอุณหภูมิใต้ต้นไม้กับบริเวณกลางแจ้งที่อยู่ในบริเวณเดียวกันโดยการเปรียบเทียบความแตกต่างของอุณหภูมิของทั้งสองบริเวณ ได้ผลดังกราฟ



การศึกษานี้ควรได้ข้อสรุปอย่างไร

1. อุณหภูมิใต้ต้นไม้ต่ำกว่าบริเวณกลางแจ้งประมาณ 3 °C
2. อุณหภูมิบนพื้นดินช่วงเวลากลางวันจะสูงกว่าช่วงเวลาเช้าและเย็น
3. ช่วงเวลาที่เก็บข้อมูลไม่มีผลต่ออุณหภูมิใต้ต้นไม้และบริเวณกลางแจ้ง
4. การปลูกต้นไม้ใต้ต้นไม้ เวลา 7.00 น. จะทำให้อุณหภูมิของอากาศลดต่ำกว่าปลูกต้นไม้บริเวณกลางแจ้ง 3 °C

48. การศึกษาความสัมพันธ์ของระยะห่างระหว่างเซนต์เซอร์กับเวลาที่ใช้ในการเคลื่อนที่เป็นเส้นตรงของหุ่นยนต์ ได้ผลดังกราฟ ข้อใดถูกต้อง



1. ประสิทธิภาพของเซนต์เซอร์มีผลต่อการเคลื่อนที่ของหุ่นยนต์
2. ระยะห่างระหว่างเซนต์เซอร์มีค่าน้อยลงจะทำให้หุ่นยนต์ใช้เวลาในการเคลื่อนที่ลดลง
3. ยิ่งให้เวลาในการเคลื่อนที่มาก หุ่นยนต์จะเคลื่อนที่ได้ระยะทางมากกว่าการให้เวลาน้อย
4. หุ่นยนต์ที่อยู่ใกล้เซนต์เซอร์จะมีอายุการใช้งานน้อยกว่าหุ่นยนต์ที่อยู่ไกลจากเซนต์เซอร์

49. ในการผลิตแอลกอฮอล์จากฟางข้าวนั้น จะได้น้ำตาลออกมาหลังการหมักด้วยกรดซัลฟิวริก ชาวบ้านได้บันทึกปริมาณน้ำตาลที่ได้จากการหมักด้วยกรดซัลฟิวริกที่มีความเข้มข้นต่างกัน ที่อุณหภูมิต่างกัน ดังนี้

ความเข้มข้นของกรดซัลฟิวริก (%)	ปริมาณน้ำตาล (mg.)	
	อุณหภูมิ 50 °C	อุณหภูมิ 100 °C
1.0	2.5	5.9
2.5	12.9	20.1
5	35.3	50.5
10	72.1	96.3
20	105.6	173.8

จากตารางข้างต้นสามารถสรุปผลการทดลองได้อย่างไร

1. ปริมาณน้ำตาลที่ได้จากการหมักฟางข้าวขึ้นอยู่กับความเข้มข้นของกรดซัลฟิวริกเท่านั้น
2. ความเข้มข้นของกรดซัลฟิวริก 20% ที่อุณหภูมิสูงจะได้ปริมาณน้ำตาลมากกว่าที่อุณหภูมิต่ำ
3. ความเข้มข้นของกรดซัลฟิวริก 20% ที่อุณหภูมิ 50 °C จะได้ปริมาณน้ำตาล 105.6 mg.
4. ปริมาณน้ำตาลที่ได้จากการหมักฟางข้าวขึ้นอยู่กับความเข้มข้นของกรดซัลฟิวริกและอุณหภูมิที่ใช้ในการหมัก

50. การผลิตน้ำหมักชีวภาพนอกจากใช้เศษอาหารเป็นวัตถุดิบแล้วต้องเติมกากน้ำตาลเพื่อใช้ในการหมัก หลังจากหมักแล้วสามารถนำน้ำหมักชีวภาพไปใช้ในการบำบัดน้ำเสีย จึงมีการศึกษาความเข้มข้นของกากน้ำตาลที่มีผลต่อค่า pH ของน้ำหมักชีวภาพเมื่อใช้ระยะเวลาในการหมักต่างกัน ได้ผลดังตาราง

ความเข้มข้นของกากน้ำตาล (mg. ต่อเศษอาหาร 1 kg.)	ค่า pH ของน้ำหมักชีวภาพ			
	วันเริ่มต้น	วันที่ 10	วันที่ 20	วันที่ 30
15	7	3	3	4
25	7	3	5	6

ข้อใดสรุปผลการทดลองได้ถูกต้อง

- ค่า pH ของน้ำหมักชีวภาพขึ้นอยู่กับความเข้มข้นของกากน้ำตาลเพียงอย่างเดียว
 - เมื่อใช้ความเข้มข้นของกากน้ำตาล 15 mg. ต่อเศษอาหาร 1 kg. ระยะเวลาในการหมักไม่มีผลต่อค่า pH
 - การใช้กากน้ำตาลความเข้มข้นต่างกันมีผลต่อค่า pH ของน้ำหมักชีวภาพเมื่อเก็บไว้เป็นเวลานานกว่า 10 วัน
 - เมื่อใช้ความเข้มข้นของกากน้ำตาล 25 mg. ต่อเศษอาหาร 1 kg. ค่า pH ของน้ำหมักชีวภาพจะลดลงตามระยะเวลาในการหมัก
51. फिल्मกรองแสงรถยนต์ที่มีประสิทธิภาพต้องสามารถดูดกลืนแสงได้ดี และมีแสงผ่านหลังจากผ่านฟิล์มกรองแสงน้อย ดังนั้นจึงมีการทดสอบประสิทธิภาพของฟิล์มกรองแสงเบอร์ต่าง ๆ ได้ผลดังตาราง

เบอร์ของฟิล์มกรองแสง	% ของแสงที่ผ่านฟิล์ม	% ของแสงที่ถูกดูดกลืน
30	75.24	24.76
40	64.86	35.14
50	58.97	41.01
60	55.67	44.33

จากผลการทดลอง ข้อใดถูกต้อง

- ฟิล์มกรองแสงเบอร์ 50 มีประสิทธิภาพดีกว่าเบอร์ 40
- เบอร์ของฟิล์มกรองแสงถูกกำหนดตามเปอร์เซ็นต์ของแสงที่ผ่านฟิล์ม
- เปอร์เซ็นต์ของแสงที่ผ่านฟิล์มแปรผันตรงกับเปอร์เซ็นต์ของแสงที่ถูกดูดกลืน
- ประสิทธิภาพของฟิล์มกรองแสงไม่ขึ้นกับปริมาณแสงที่ผ่านฟิล์มและที่ถูกดูดกลืน

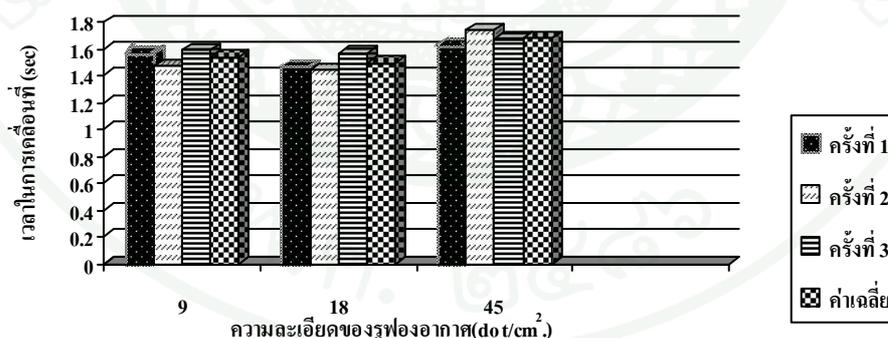
52. ตารางแสดงปริมาณความร้อนของถ่านอัดแท่ง 500 g. ที่ได้จากหญ้าคาผสมกับขี้เลื่อยในอัตราส่วนที่ต่างกัน

หญ้าคา : ขี้เลื่อย	ปริมาณความร้อน (cal.)
1 : 4	250
2 : 3	180
2.5 : 2.5	130
3 : 2	110
4 : 1	80

จากตารางข้างต้น นักเรียนควรสรุปผลการทดลองได้อย่างไร

1. ปริมาณหญ้าคาแปรผันตรงกับปริมาณความร้อน
2. ปริมาณความร้อนขึ้นอยู่กับชนิดของวัตถุดิบที่ใช้ในการเผาไหม้
3. เมื่อใช้หญ้าคา 150 g. และ ขี้เลื่อย 100 g. จะได้ปริมาณความร้อน 110 cal.
4. อัตราส่วนของหญ้าคา : ขี้เลื่อย เท่ากับ 1 : 4 จะได้ปริมาณความร้อนมากที่สุด

53. การเคลื่อนที่ของวัตถุต่าง ๆ ในของเหลวจะเกิดแรงเสียดทานขึ้น ดังนั้นหากลดแรงเสียดทานนี้ลงได้จะช่วยเพิ่มประสิทธิภาพในการเคลื่อนที่และลดการสูญเสียพลังงาน จึงมีการศึกษาเกี่ยวกับความละเอียดของรูปฟองอากาศเมื่อใช้ฟิล์มอากาศห่อหุ้มวัตถุที่เคลื่อนที่ จัปเวลาที่ใช้ในการเคลื่อนที่ได้ผลดังแผนภูมิ ข้อใดถูกต้อง



1. ความละเอียดของรูปฟองอากาศแปรผันตรงกับระยะเวลาในการเคลื่อนที่
2. ยิ่งความละเอียดของรูปฟองอากาศมากขึ้นระยะเวลาในการเคลื่อนที่จะน้อยลง
3. ฟิล์มอากาศที่มีความละเอียดของรูปฟองอากาศเท่ากับ 18 dot/cm². ช่วยลดแรงเสียดทานในการเคลื่อนที่ได้ดีที่สุด
4. ฟิล์มอากาศที่มีความละเอียดของรูปฟองอากาศ เท่ากับ 45 dot/cm². ช่วยลดแรงเสียดทานในการเคลื่อนที่ได้ดีที่สุด

54. ในการทดลองเพื่อเปรียบเทียบวิธีการวิเคราะห์หาปริมาณกรดแอสซิติคในน้ำส้มสายชูตัวอย่างที่พบในท้องตลาดระหว่างวิธี A กับ วิธี B ได้ผลดังตาราง

ตัวอย่าง	ปริมาณที่ระบุในฉลาก % (มวล/ปริมาตร)	ปริมาณกรดแอสซิติค	
		วิธี A % (มวล/ปริมาตร)	วิธี B % (มวล/ปริมาตร)
ทิพรส	5	4.75	5.11
ภูเขาทอง	5	4.71	5.11
อสร.	5	4.81	5.16
ฉลากทอง	5	4.65	5.20

จากตารางบันทึกผลการวิเคราะห์ข้างต้น ข้อสรุปในข้อใดถูกต้องที่สุด

1. บริษัทที่ผลิตน้ำส้มสายชูใส่กรดแอสซิติคไม่เท่ากัน
2. การวิเคราะห์ด้วยวิธี B จะได้ปริมาณกรดแอสซิติคสูงกว่าวิธี A
3. การวิเคราะห์หาปริมาณกรดแอสซิติคด้วยวิธี B จะให้ผลการวิเคราะห์ดีกว่าวิธี A
4. อสร.มีปริมาณกรดแอสซิติคสูงที่สุด รองลงมาคือ ทิพรส ภูเขาทอง และฉลากทอง ตามลำดับ

ประวัติการศึกษาและการทำงาน

ชื่อ-นามสกุล	นางมนัสนันท์ มากสวัสดิ์
วัน เดือน ปี ที่เกิด	วันที่ 28 เดือนมีนาคม พ.ศ.2523
สถานที่เกิด	กรุงเทพมหานคร
ประวัติการศึกษา	ศึกษาศาสตรบัณฑิต (การสอนวิทยาศาสตร์) มหาวิทยาลัยเกษตรศาสตร์
ตำแหน่งปัจจุบัน	ครูผู้ช่วย
สถานที่ทำงานปัจจุบัน	โรงเรียนเศรษฐบุทรบำรุงพิเศษ

