

การแสดงผลฐานความเที่ยงตรงและความเชื่อมั่นของแบบวัดทักษะกระบวนการ
ทางวิทยาศาสตร์ ชั้นบูรณาการ

ปริญญานิพนธ์
ของ
นพวรรณ ศรีเกตุ

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา
พฤษภาคม 2550

การแสดงความเห็นเกี่ยวกับความเที่ยงตรงและความเชื่อมั่นของแบบวัดทักษะกระบวนการ
ทางวิทยาศาสตร์ ชั้นประถมศึกษา

ปริญญานิพนธ์
ของ
นพวรรณ ศรีเกตุ

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา

พฤษภาคม 2550

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

ปริญญานิพนธ์
เรื่อง
การแสดงผลฐานความเที่ยงตรงและความเชื่อมั่นของแบบวัดทักษะกระบวนการ
ทางวิทยาศาสตร์ ชั้นบูรณาการ

ของ
นพวรรณ ศรีเกตุ

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.เพ็ญศิริ จีระเดชากุล)
วันที่.....เดือน พฤษภาคม พ.ศ. 2550

คณะกรรมการควบคุมปริญญานิพนธ์

คณะกรรมการสอบปากเปล่า

.....ประธาน
(รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์)

.....ประธาน
(ผู้ช่วยศาสตราจารย์ ดร.องอาจ นัยพัฒน์)

.....กรรมการ
(รองศาสตราจารย์นิภา ศรีไพโรจน์)

.....กรรมการ
(รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์)

.....กรรมการ
(รองศาสตราจารย์นิภา ศรีไพโรจน์)

.....กรรมการ
(อาจารย์ชวลิต รวยอาจิณ)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลงได้ด้วยความกรุณาเอาใจใส่ให้คำปรึกษาเป็นอย่างดีของ คณะกรรมการที่ปรึกษาปริญญานิพนธ์ ได้แก่ รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์ ในฐานะประธานที่ปรึกษาปริญญานิพนธ์ และรองศาสตราจารย์ นิภา ศรีไพโรจน์ ในฐานะกรรมการที่ปรึกษาปริญญานิพนธ์ ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.องอาจ นัยพัฒน์ และอาจารย์ชวลิต รวยอาจิน ที่กรุณาเป็นกรรมการสอบปริญญานิพนธ์ และได้ให้คำแนะนำที่ทำให้ปริญญานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น และที่สำคัญยิ่งผู้วิจัยขอกราบระลึกถึงพระคุณอาจารย์ภาควิชาการวัดผลและวิจัย การศึกษาทุกท่านที่ได้อบรมสั่งสอนให้วิชาความรู้ ซึ่งเป็นประโยชน์ต่อการทำปริญญานิพนธ์ฉบับนี้จน สำเร็จลุล่วง

ขอกราบขอบพระคุณผู้เชี่ยวชาญทั้ง 5 ท่าน คือ รองศาสตราจารย์ ละเอียด รักษ์เฝ้า รองศาสตราจารย์ ชูศรี วงศ์รัตนะ อาจารย์พรวณี จินตมาศ อาจารย์สุรศักดิ์ บานแย้ม และ อาจารย์มะนิดดา เกตุสุข ที่ได้กรุณาตรวจสอบและให้ข้อเสนอแนะในการแก้ไขปรับปรุงเครื่องมือที่ใช้ และเก็บข้อมูลในการทำวิจัยครั้งนี้เป็นอย่างดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณผู้อำนวยการ ผู้บริหาร คณะครู และนักเรียนโรงเรียนที่เป็นกลุ่ม ตัวอย่างและให้ความร่วมมือในการเก็บรวบรวมข้อมูลเพื่อใช้ในการวิจัยเป็นอย่างดี

ขอกราบขอบพระคุณนายวีรชัย ศรีเกตุ และนางศรีอรุณ ศรีเกตุ ซึ่งเป็นบิดาและมารดา และครอบครัวที่ให้กำลังใจ ความห่วงใย ทำให้ลูกประสบความสำเร็จในการศึกษา รวมทั้ง ขอขอบพระคุณอาจารย์ลัดดา ศรีเลิศสมบัติ อาจารย์อโนทัย สมานทรัพย์ และอาจารย์วีรชัย ศรี เกตุ ที่ให้ความอนุเคราะห์ตรวจสอบบทดสอบนักเรียน ขอขอบคุณอาจารย์บุญยิ่ง เข้มพงษ์ ที่ให้ คำปรึกษา เพื่อน ๆ พี่ ๆ น้อง ๆ ทุกคนที่เป็นกำลังใจ และคอยช่วยเหลือผู้วิจัยมาโดยตลอด

คุณค่าทั้งหมดที่เกิดขึ้นจากปริญญานิพนธ์เล่มนี้ ผู้วิจัยขอน้อมรำลึกและบูชาพระคุณแก่ บุพการีของผู้วิจัย และบูรพาจารย์ทุกท่านที่อยู่เบื้องหลังในการวางรากฐานการศึกษาให้กับผู้วิจัย ตั้งแต่อดีตจนถึงปัจจุบัน

นพวรรณ ศรีเกตุ

สารบัญ

บทที่		หน้า
1	บทนำ	1
	ภูมิหลัง.....	1
	ความมุ่งหมายของการวิจัย.....	4
	ความสำคัญของการวิจัย.....	4
	ขอบเขตของการวิจัย.....	4
	ประชากรที่ใช้ในการวิจัย.....	4
	กลุ่มตัวอย่างที่ใช้ในการวิจัย.....	5
	ตัวแปรที่ใช้ในการวิจัย.....	5
	นิยามศัพท์เฉพาะ.....	5
2	เอกสารและงานวิจัยที่เกี่ยวข้อง	8
	แบบทดสอบอัตนัย หรือแบบทดสอบความเรียง.....	9
	นิยามของแบบทดสอบความเรียง.....	9
	ลักษณะและประเภทของแบบทดสอบความเรียง.....	11
	การสร้างแบบทดสอบความเรียง.....	16
	องค์ประกอบที่ส่งผลต่อคะแนนในแบบทดสอบความเรียง.....	19
	กฎเกณฑ์การให้คะแนน.....	20
	ความหมายของกฎเกณฑ์การให้คะแนน.....	20
	ประเภทของกฎเกณฑ์การให้คะแนน.....	21
	การสร้างกฎเกณฑ์การให้คะแนน.....	22
	ทฤษฎีค่าความเชื่อมั่น.....	27
	การประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม.....	27
	การประมาณค่าความเชื่อมั่นที่เกี่ยวกับการตรวจให้คะแนน.....	29
	ดัชนีความสอดคล้องของผู้ประเมิน.....	31
	การประมาณค่าความเชื่อมั่นตามโครงสร้างทฤษฎีการสรุปอ้างอิง.....	

สารบัญ (ต่อ)

บทที่

หน้า

2 (ต่อ)

หลักฐานแสดงความเที่ยงตรง.....	43
ความหมายของความเที่ยงตรง.....	43
วิธีการแสดงหลักฐานความเที่ยงตรง.....	44
ทักษะกระบวนการทางวิทยาศาสตร์.....	49
ความหมายทักษะกระบวนการทางวิทยาศาสตร์.....	49
ประเภทของทักษะกระบวนการทางวิทยาศาสตร์.....	50
หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544.....	54
งานวิจัยที่เกี่ยวข้อง.....	56
งานวิจัยที่เกี่ยวข้องกับความเที่ยงตรงและความเชื่อมั่น.....	56
งานวิจัยที่เกี่ยวข้องกับทักษะกระบวนการทางวิทยาศาสตร์.....	60

3 วิธีดำเนินการวิจัย.....64

การกำหนดประชากรและการสุ่มกลุ่มตัวอย่าง.....	64
เครื่องมือที่ใช้ในการวิจัย.....	65
ตัวอย่างเครื่องมือที่ใช้ในการวิจัย.....	70
วิธีดำเนินการเก็บข้อมูล.....	73
การจัดกระทำและการวิเคราะห์ข้อมูล.....	73
สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	74

4 ผลการวิเคราะห์ข้อมูล.....78

5 สรุปผล อภิปราย และข้อเสนอแนะ.....92

ความมุ่งหมายของการวิจัย.....	92
ประชากรที่ใช้ในการวิจัย.....	92
กลุ่มตัวอย่างที่ใช้ในการวิจัย.....	92

สารบัญ (ต่อ)

บทที่

หน้า

5 (ต่อ)

เครื่องมือที่ใช้ในการวิจัย.....	93
วิธีการดำเนินการเก็บข้อมูล.....	93
สรุปผลการวิเคราะห์ข้อมูล.....	94
อภิปรายผล.....	97
ข้อเสนอแนะ.....	99
 บรรณานุกรม.....	 100
 ภาคผนวก.....	 108
ภาคผนวก ก คุณภาพเครื่องมือ.....	109
ภาคผนวก ข คู่มือการตรวจให้คะแนน.....	122
ภาคผนวก ค รายนามผู้เชี่ยวชาญ.....	147
 ประวัติย่อผู้วิจัย.....	 149

บัญชีตาราง

ตาราง	หน้า
1 การเปรียบเทียบลักษณะข้อสอบปรนัย และข้อสอบแบบความเรียง.....	11
2 ข้อดี และข้อจำกัดของแบบความเรียง	13
3 เกณฑ์การให้คะแนนแบบรวมองค์ประกอบ.....	23
4 เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ.....	24
5 อธิบายคุณภาพ “การพุดอภิปราย”.....	25
6 การกำหนดคะแนนการตัดสินระดับคุณภาพ.....	25
7 การประมาณค่าความแปรปรวนขึ้นมาใหม่ของแบบแผน	42
8 จำนวนนักเรียนที่ใช้เป็นกลุ่มตัวอย่างในการเก็บรวบรวมข้อมูล.....	65
9 ผลการวิเคราะห์เหตุการณ์ที่บ่งชี้ทักษะกระบวนการทางวิทยาศาสตร์.....	68
10 ค่าสถิติพื้นฐานของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ.....	80
11 ค่าสถิติที่ได้จากการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับแรกด้วยโปรแกรม ลิสเรลจำแนกเป็นรายตอน (N = 185).....	86
12 ค่าสถิติจากการทดสอบความเหมาะสมของโมเดลกับข้อมูลเชิงประจักษ์ ของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการจำแนกเป็น รายตอนด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน อันดับที่สอง.....	90
13 ค่าความเชื่อมั่นด้วยการใช้ดัชนีความสอดคล้องของผู้ประเมิน ของแบบวัด กระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ.....	90
14 แสดงการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบวัดทักษะ กระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่มีผู้ตรวจให้คะแนนจำนวน 3 คน....	91
15 ดัชนีความสอดคล้องของการประเมินจากผู้เชี่ยวชาญในการตรวจสอบคุณภาพ ของเครื่องมือ (IOC).....	110
16 ค่าความยาก (p) ค่าอำนาจจำแนก (r_{xy}) ของแบบวัดทักษะกระบวนการทาง วิทยาศาสตร์ขั้นบูรณาการ.....	113
17 ค่าความยาก (p) ค่าอำนาจจำแนก (r_{xy}) ของแบบวัดทักษะกระบวนการทาง วิทยาศาสตร์ขั้นบูรณาการ ของการทดสอบครั้งที่ 2.....	116

บัญชีตาราง (ต่อ)

ตาราง	หน้า
18 ค่าความยาก (p) ค่าอำนาจจำแนก (r_{xy}) ของแบบวัดทักษะกระบวนการทาง วิทยาศาสตร์ขั้นบูรณาการ ของการทดสอบกับกลุ่มตัวอย่าง.....	118
19 ค่าสถิติที่ได้จากการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับแรก โดยใช้โปรแกรม ลิสเรลจำแนกเป็นรายตอน (N = 185)	120

บัญชีภาพประกอบ

ภาพประกอบ

หน้า

- 1 แผนภูมิ Venn แสดงองค์ประกอบความแปรปรวนของแบบแผน $p \times i \times r$
เมื่อองค์ประกอบที่ศึกษาทั้งหมดเป็นองค์ประกอบสุ่ม.....39
- 2 ลำดับขั้นการสร้างแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ.....66
- 3 โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับแรก ของโมเดลแบบวัด
ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการจำแนกเป็นตอน.....84
- 4 โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สองของแบบวัด
ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ จำแนกเป็นรายตอน.....89

การแสดงผลฐานความเที่ยงตรงและความเชื่อมั่นของแบบวัดทักษะกระบวนการ
ทางวิทยาศาสตร์ ชั้นประถมศึกษา

บทคัดย่อ
ของ
นพวรรณ ศรีเกตุ

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา
พฤษภาคม 2550

นพวรรณ ศรีเกตุ. (2550). การแสดงหลักฐานความเที่ยงตรงและความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ ชั้นบูรณาการ. ปรินซ์ตันนิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม : รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ , รองศาสตราจารย์ นิภา ศรีไพโรจน์

การวิจัยครั้งนี้ มีจุดมุ่งหมายสำคัญเพื่อแสดงหลักฐานความเที่ยงตรง และความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ตามแนวของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2526 : 1 – 16) ศึกษาจากกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนชิโนรสวิทยาลัย โรงเรียนสตรีวัดระฆัง โรงเรียนสวนอนันต์ และโรงเรียน มหรรณพาราม ภาคเรียนที่ 2 ปีการศึกษา 2549 จำนวน 185 คน ซึ่งเลือกมาโดยการสุ่มแบบสองขั้นตอน

แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ซึ่งวัดสาระเกี่ยวกับวิชาวิทยาศาสตร์พื้นฐาน ด้านฟิสิกส์ เคมี และชีววิทยา แบ่งออกเป็น 5 ตอน ๆ ละ 6 ข้อ รวมทั้งหมดจำนวน 30 ข้อ เป็นแบบวัดชนิดปลายเปิด

ผลการวิจัยพบว่า แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการที่สร้างขึ้นมีคุณภาพใช้ได้ โดยมีหลักฐานแสดงความเที่ยงตรงตามโครงสร้างคำนวณด้วยสูตรของ ไฮเซน และ บอร์นสเตรดต์ (Heise and Bohrnstedt) มีค่า .70 ซึ่งค่าอยู่ในเกณฑ์สูง และมีหลักฐานความเที่ยงตรงเชิงโครงสร้างภายในฉบับด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันทั้งอันดับแรกและอันดับสอง ไม่เดลกับข้อมูลเชิงประจักษ์มีความเหมาะสมเจาะพอดี ซึ่งทักษะการตั้งสมมติฐานมีหลักฐานค่าความเชื่อมั่นอยู่ระหว่าง .88 - .93 หลักฐานความเที่ยงตรงอยู่ระหว่าง .78 - .87 ทักษะการกำหนดนิยามเชิงปฏิบัติการมีหลักฐานค่าความเชื่อมั่นอยู่ระหว่าง .86 - .96 หลักฐานความเที่ยงตรงอยู่ระหว่าง .73 - .93 ทักษะการกำหนดและควบคุมตัวแปรมีหลักฐานค่าความเชื่อมั่นอยู่ระหว่าง .83 - .91 หลักฐานความเที่ยงตรงอยู่ระหว่าง .68 - .83 ทักษะการทดลองมีหลักฐานค่าความเชื่อมั่นอยู่ระหว่าง .89 - .94 หลักฐานความเที่ยงตรงอยู่ระหว่าง .79 - .89 ทักษะการตีความหมายข้อมูลและลงข้อสรุปมีหลักฐานค่าความเชื่อมั่นอยู่ระหว่าง .89 - .95 หลักฐานความเที่ยงตรงอยู่ระหว่าง .79 - .90

หลักฐานแสดงความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการด้วยการใช้ดัชนีความสอดคล้องของผู้ประเมิน 3 คน มีค่า 0.93 และมีหลักฐานแสดงความเชื่อมั่นด้วยการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง มีค่า 0.68 ซึ่งค่าความเชื่อมั่นที่ได้อยู่ในเกณฑ์สูง

VALIDITY AND RELIABILITY EVIDENCE OF INTEGRATED
SCIENCE PROCESS SKILLS TEST

AN ABSTRACT
BY
NOPHAWAN SRIKHET

Presented in Partial Fulfillment of the Requirements for the
Master of Education Degree in Educational Measurement
at Srinakharinwirot University

April 2007

Nophawan Srikhet (2550). *The Validity and Reliability Evidence of Integrated Science Process Skill Test*. Master thesis, M.Ed. (Educational Measurement). Bangkok: Graduate School, Srinakharinwirot University. Advisor Committee: Assoc. Prof. Dr. Boonchird Pinyoanuntapong, Assoc. Prof. Nipa Sripairot.

The main purpose of this research was to develop the integrated science process skill test according to the guidelines of The Institute for the Promotion of Teaching Science and Technology (1983: 1 - 16); and the specific purpose was to show the validity and reliability evidence of the test. The samples of this research were 285 Matthayomsuksa 5 students of Chinorot Wittayalai School, Satri Wat Rakhang School, Suan Anun School, and Mahunapharam School in 2nd semester of 2006. They were selected by two-stage random sampling.

The test comprised 30 open-ended questions with the contents of Basic Science, Physics, Chemistry, and Biology subjects. It was separated into 5 parts with 6 questions each.

The results of this research revealed that the quality of the integrated science process skill test was fairly good. There was an evidence of its construct validity calculated by Heise & Bohrnstedt's formula at the high level of .70. There was also an evidence of construct validity within the test by using confirmed component analysis in both first and second orders. The model fitted well the empirical data. The reliability evidence of hypothesis formulating skill was between .88 - .93 while the validity evidence was between .78 - .87. The reliability evidence of operational defining skill was between .86 - .96 while the validity evidence was between .73 - .93. The reliability evidence of variable identifying and controlling skill was between .83 - .91 while the validity evidence was between .68 - .83. The reliability evidence of experimenting skill was between .89 - .94 while the validity evidence was between .79 - .89. The reliability evidence of data interpreting and summarizing skill was between .89 - .95 while the validity evidence was between .79 - .90.

The reliability evidence of the integrated science process skill test by using index of congruence of 3 raters was 0.93, and the reliability evidence with the use of generalization coefficient estimate was at the high level of 0.68.

บทที่ 1

บทนำ

ภูมิหลัง

การศึกษา เป็นกระบวนการที่มุ่งพัฒนาคน ให้มีความเจริญงอกงาม ตามทิศทางที่สังคมปรารถนา โดยพยายามเปลี่ยนแปลงพฤติกรรมของผู้เรียน ให้มีคุณสมบัติตรงตามความต้องการของสังคม ดังนั้นนักการศึกษาจึงพยายามสร้างหรือปลูกฝังคุณลักษณะต่าง ๆ ที่จำเป็นที่ผู้เรียนยังไม่มีหรือมีอยู่น้อยให้เกิดขึ้น ในขณะเดียวกันก็มุ่งจัดหรือเปลี่ยนแปลงพฤติกรรมที่ไม่พึงประสงค์ให้หมดไป การจัดการศึกษาเพื่อให้บรรลุเป้าหมายดังกล่าวจำเป็นต้องอาศัยองค์ประกอบหลายประการ องค์ประกอบหนึ่งก็คือ การวัดผล

พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 หมวด 4 การจัดการศึกษา มาตรา 26 ระบุไว้ว่า ให้สถานศึกษาจัดการประเมินผู้เรียน โดยพิจารณาจากพัฒนาการของผู้เรียน ความประพฤติ การสังเกตพฤติกรรมการเรียน การร่วมกิจกรรม และทดสอบควบคู่กันไปในกระบวนการเรียนการสอน ให้มีวิธีการวัดประเมินที่หลากหลาย (พ.ร.บ.การศึกษาแห่งชาติ พ.ศ.2542. : 12 – 14) ให้มีความเหมาะสมกับวัย และลักษณะความสามารถของเด็ก ดังนั้นในการวัดประเมินนักเรียน จึงจำเป็นต้องใช้เครื่องมือชนิดต่าง ๆ ทั้งการวัดแบบไม่เป็นทางการ เช่น การสังเกต การบันทึก การสนทนากับเด็ก และการวัดประเมินแบบทางการ ได้แก่ แบบทดสอบชนิดต่าง ๆ คือ แบบทดสอบถูกผิด จับคู่ เติมคำ เลือกตอบ และแบบทดสอบอัตนัย หรือ แบบทดสอบความเรียง (บุญเชิด ภิญโญนนตพงษ์. 2544 : 62)

วิทยาศาสตร์เป็นเรื่องของการเรียนรู้เกี่ยวกับธรรมชาติ โดยเฉพาะมนุษย์ใช้กระบวนการสังเกต สำรวจตรวจสอบ และการทดลองเกี่ยวกับปรากฏการณ์ธรรมชาติ และนำผลมาจัดระบบหลักการแนวคิดและทฤษฎี ดังนั้นการเรียนการสอนวิทยาศาสตร์จึงมุ่งเน้นให้ผู้เรียนได้เป็นผู้เรียนรู้และค้นพบด้วยตนเองมากที่สุด นั่นคือให้ได้ทั้งกระบวนการและองค์ความรู้ ตั้งแต่เริ่มแรกก่อนเข้าเรียนเมื่ออยู่ในสถานศึกษา และเมื่อออกจากสถานศึกษาไปประกอบอาชีพแล้ว การเรียนรู้วิทยาศาสตร์เป็นการพัฒนาผู้เรียนให้ได้รับทั้งความรู้ กระบวนการและเจตคติ ผู้เรียนทุกคนควรได้รับการกระตุ้นส่งเสริมให้สนใจและกระตือรือร้นที่จะเรียนรู้วิทยาศาสตร์ มีความสงสัย เกิดคำถามในสิ่งต่าง ๆ ที่เกี่ยวกับโลกธรรมชาติรอบตัว มีความมุ่งมั่นและมีความสุขที่จะศึกษา ค้นคว้า สืบเสาะหาความรู้เพื่อรวบรวมข้อมูล วิเคราะห์ผล นำไปสู่คำตอบของคำถาม สามารถตัดสินใจด้วยการใช้ข้อมูลอย่างมีเหตุผล สามารถสื่อคำถาม คำตอบ ข้อมูลและสิ่งที่ค้นพบจากการเรียนรู้ให้ผู้อื่นเข้าใจได้ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. 2546 : 1 – 2)

การวัดทักษะกระบวนการทางวิทยาศาสตร์ ยังมีปัญหาในทางปฏิบัติ ยังไม่เป็นรูปธรรมที่สามารถตรวจสอบได้ ผู้สอนส่วนใหญ่ขาดความรู้ความเข้าใจเกี่ยวกับการวัดทักษะกระบวนการทางวิทยาศาสตร์ (สุวัฒน์ นิยมคำ. 2531 : 2) ขาดความเข้าใจในการเลือกวิธีการวัดประเมินและวิธีสร้างกฎเกณฑ์การให้คะแนน (Scoring Rubric) ที่เป็นปรนัยให้มีความเชื่อมั่นสูง และลดความคลาดเคลื่อนจากการวัดในแหล่งต่าง ๆ ให้เหลือน้อยที่สุด โดยมากครูผู้สอนจะนิยมวัดผลสัมฤทธิ์ด้านความรู้ความจำ โดยใช้แบบทดสอบที่มีตัวเลือกกำหนดให้ตอบ ที่สะดวกทั้งวิธีการตรวจให้คะแนน และสามารถนำไปวิเคราะห์ หาคุณภาพของแบบทดสอบ แต่วิธีการดังกล่าวนี้ เราอาจไม่ทราบถึงพัฒนาการที่แท้จริง เกี่ยวกับวิธีการแสวงหาความรู้ของนักเรียน และทักษะพื้นฐานความรู้ทางวิทยาศาสตร์ที่จำเป็น ส่งผลให้ผลสัมฤทธิ์ทางการเรียนในวิชาวิทยาศาสตร์ตกต่ำ

ข้อสอบอัตนัยหรือแบบทดสอบความเรียงเป็นข้อสอบที่มีคุณค่ามากชนิดหนึ่ง สามารถวัดพฤติกรรมด้านความรู้ความคิดระดับสูงได้ดี (บุญเชิด ภิญญอนันตพงษ์. 2525 : 46) ข้อสอบอัตนัยจะเขียนคำถาม หรือโจทย์ โดยกำหนดเป็นสถานการณ์ หรือปัญหาในรูปใดรูปหนึ่งอย่างกว้าง ๆ หรือเฉพาะเจาะจง เพื่อให้ผู้ตอบได้แสดงความรู้ ความเข้าใจ ความคิดเห็น ทักษะคิด ได้อย่างไม่จำกัด ข้อสอบอัตนัยมีจำนวนข้อไม่มาก การตอบข้อสอบอัตนัยจึงมีลักษณะจัดระเบียบคำตอบในเวลาที่กำหนด ใช้สำนวนภาษาและแบบฉบับของตนเองเขียนตอบ เขียนคำตอบให้ครอบคลุมอย่างสมบูรณ์ และระมัดระวัง ผู้ตรวจให้คะแนนต้องเป็นผู้มีความรู้ในเนื้อหาวิชา การให้คะแนนต้องอาศัยทักษะและความพยายามในการอ่าน ผู้ตรวจต้องทำใจให้บริสุทธิ์ โดยยึดคุณธรรมอันสูงส่งในการให้คะแนน (บุญเชิด ภิญญอนันตพงษ์. 2547 : 59) และมีความเป็นอัตนัย คือ ให้การตัดสินตามความรู้สึกส่วนตัว ซึ่งการตัดสินของแต่ละคนอาจไม่ตรงกัน หรืออาจจะตรงกันข้ามก็ได้ ทั้งนี้เพราะบุคคลย่อมแตกต่างกันในด้านต่าง ๆ การนำไปใช้ต้องพยายามวัดพฤติกรรมที่เหมาะสมแล้วยังต้องทำให้การตรวจให้คะแนนมีความเชื่อถือได้มากที่สุด ปัญหาใหญ่ของคะแนนแบบทดสอบความเรียงที่จะให้ความเชื่อมั่นสูงขึ้นอยู่กับผู้ตรวจ ดังนั้น หัวใจสำคัญของการให้คะแนนจึงต้องมีพัฒนาการให้คะแนนด้วยวิธีที่จะให้คะแนนได้อย่างมีความเชื่อถือได้จากการวัดนั้น ๆ ปัญหาในการให้คะแนนไม่สอดคล้องกันมีสาเหตุสำคัญมาจาก 3 กรณี คือ 1) กรณีที่มีผู้ตรวจต่างกัน ตรวจให้คะแนนข้อเดียวกัน 2) กรณีที่ผู้ตรวจคนเดียวกันให้คะแนนข้อเดียวกันในเวลาต่างกัน และ 3) การตรวจคำตอบที่มีอิสระในการตอบ (Hopkins and Antes. 1990 : 236 ; Coffman. 1971 : 276 – 277) การลดความคลาดเคลื่อนดังกล่าวจำเป็นต้องมีวิธีการตรวจที่รัดกุมซึ่งมีวิธีการตรวจให้คะแนน 2 วิธีที่เหมาะสมกับการตรวจแบบทดสอบความเรียงที่สุด คือ การตรวจด้วยวิธีวิเคราะห์ย่อย (Analytic Method) หรือวิธีให้ตามประเด็น (Point Method) การตรวจแบบนี้ขึ้นอยู่กับเตรียมรายละเอียดคำแนะนำในการตรวจด้วยประเด็นย่อย ๆ อย่างละเอียด และวิธีประเมินภาพรวม (Global Method) หรือเรียกว่า

วิธีประเมินรวม (Holistic Method) ที่จะให้คะแนนในภาพรวมของคำตอบที่จะขึ้นอยู่กับมาตรฐานของการตัดสินใจของผู้ตรวจเอง (Hopkins and Antes. 1990 : 237) และวิธีการตรวจก็เป็นปัจจัยหนึ่งที่มีอิทธิพลต่อความเชื่อมั่นของผู้ตรวจ (Rater Reliability) ที่จะมีความสอดคล้องกันได้สูง นักวัดผลการศึกษาและนักวิจัยเสนอว่า การชี้แจงให้ผู้ตรวจมีความรู้ความเข้าใจในเรื่องที่จะตรวจ รู้จักใช้แบบแผนหรือคู่มือการตรวจที่เหมาะสม สามารถทำให้การตรวจมีความเชื่อมั่นสูงได้ โดยทั่วไปการตรวจโดยวิธีวิเคราะห์ย่อย (Analytic Method) จะให้ค่าความเชื่อมั่นสูงกว่าวิธีประเมินรวม (Holistic Method) นอกจากนี้การตรวจทีละข้อจะดีกว่าการตรวจทีละคน (วิญญาณ วิศาลาภรณ์. 2533 : 79) เพราะเป็นการลดความแปรปรวนในการตรวจจากการเปลี่ยนข้อบ่งชี้ ๆ นอกจากนี้สิ่งสำคัญของการตรวจทั้งสองวิธี คือ คำตอบเหล่านั้นต้องไม่คำนึงว่าเป็นของใคร เพื่อหลีกเลี่ยงความลำเอียงที่เกี่ยวข้องกับความประทับใจของครูที่มีต่อนักเรียน และการตรวจทีละมาก ๆ ก็อาจทำให้เกิดความเหน็ดเหนื่อยขึ้นได้ สิ่งนี้เป็นเหตุให้การให้คะแนนไม่แน่นอนขึ้นได้ ดังนั้นการให้คะแนนจะเป็นมาตรฐานได้นั้นต้องยึดถึงมาตรฐานคะแนน และเกณฑ์การตรวจเดียวกันทุกแผ่นคำตอบเสมอ (บุญเชิด ภิญญอนันตพงษ์. 2547 : 60)

ดังนั้นในการสร้างแบบทดสอบความเรียง จึงต้องพัฒนาวิธีการตรวจให้คะแนนควบคู่กันไปด้วย และต้องตรวจสอบความเชื่อมั่นของการตรวจให้คะแนน ตัวประกอบที่มีผลต่อค่าความเชื่อมั่นของแบบประเมินผลทักษะกระบวนการทางวิทยาศาสตร์ ที่มีลักษณะเป็นคำถามปลายเปิด ประกอบด้วย ผู้ตรวจให้คะแนน วิธีการตรวจ จำนวนแบบทดสอบ จำนวนผู้ตอบ และทฤษฎีที่ประมาณค่าความเชื่อมั่น โดยเฉพาะความไม่แน่นอนของผู้ตรวจให้คะแนน (Rater Unreliability) นับเป็นปัญหาที่สำคัญ ความไม่แน่นอนของผู้ตรวจให้คะแนนจะส่งผลให้คะแนนแปรเปลี่ยนไปตามผู้ตรวจให้คะแนนแต่ละคน แม้แต่ผู้ตรวจให้คะแนนที่เป็นคนเดียวกัน ก็จะทำให้คะแนนแปรเปลี่ยนไปตามระยะเวลาที่ทำการตรวจ (บุญเชิด ภิญญอนันตพงษ์. 2533 : 47)

ด้วยเหตุผลดังกล่าวข้างต้น ผู้วิจัยจึงสนใจที่จะศึกษาว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ มีความเที่ยงตรงเชิงโครงสร้างเป็นอย่างไร โดยพิจารณาจากหลักฐานความเที่ยงตรงที่คำนวณด้วยสูตรของ ไฮสไน์ และบอร์นสเตรดต์ (Heise and Bohrnstedt) และการวิเคราะห์องค์ประกอบเชิงยืนยัน และพิจารณาค่าความเชื่อมั่น ด้วยค่าดัชนีความสอดคล้องของผู้ประเมิน (Rater Agreement Indexes : RAI) และคำนวณโดยใช้ทฤษฎีการสรุปอ้างอิง โดยใช้เนื้อหาวิชาวิทยาศาสตร์พื้นฐาน ระดับชั้นมัธยมศึกษาตอนปลาย

ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีจุดมุ่งหมายสำคัญเพื่อแสดงหลักฐานความเที่ยงตรงและความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ โดยมีจุดมุ่งหมายเฉพาะของการวิจัยดังนี้

1. เพื่อแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ 2 วิธี คือ

1.1 ความเที่ยงตรงเชิงโครงสร้างด้วยสูตร $\rho_{T_k S_k}$ ของไฮเซน และบอร์นสเตรดต์ (Heise and Bohrnstedt. 1970 : 109 , 111 – 113)

1.2 ความเที่ยงตรงเชิงโครงสร้าง โดยใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) (บุญเชิด ภิญโญนนตพงษ์. 2547 : 193)

2. เพื่อแสดงหลักฐานความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ 2 วิธี คือ

2.1 ค่าดัชนีความสอดคล้องของผู้ประเมิน (Rater Agreement Indexes : RAI) ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

2.2 ค่าความเชื่อมั่นของการตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่คำนวณโดยใช้ทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

ความสำคัญของการวิจัย

การวิจัยครั้งนี้ เป็นการศึกษาคุณภาพด้านความเที่ยงตรง และความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ได้แก่ ทักษะการตั้งสมมติฐาน ทักษะการกำหนดนิยามเชิงปฏิบัติการ ทักษะการกำหนดและควบคุมตัวแปร ทักษะการทดลอง และทักษะการตีความหมายข้อมูลและลงข้อสรุป พร้อมเกณฑ์การให้คะแนน ผลที่ได้จากการศึกษา จะเป็นแนวทางในการสร้างแบบทดสอบที่มีคุณภาพ และนำไปใช้ ในการวัดทักษะกระบวนการทางวิทยาศาสตร์ ซึ่งจะนำไปดำเนินการปรับปรุง / แก้ไข และพัฒนาให้เกิดพฤติกรรมด้านทักษะกระบวนการทางวิทยาศาสตร์ในตัวผู้เรียนต่อไป

ขอบเขตของการวิจัย

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนมัธยมศึกษาสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กลุ่มที่ 3 เขตพื้นที่การศึกษา 3 กรุงเทพมหานคร ปีการศึกษา 2549 มีทั้งสิ้น 11 โรงเรียน จำนวนนักเรียนทั้งหมด 9,317 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 (ทุกแผนการเรียน) โรงเรียนมัธยมศึกษาสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กลุ่มที่ 3 เขตพื้นที่การศึกษา 3 กรุงเทพมหานคร ปีการศึกษา 2549 จำนวน 285 คน ซึ่งได้มาโดยการสุ่มแบบสองขั้นตอน (Two – Stage Random Sampling) โดยใช้เนื้อหาวิชาฟิสิกส์พื้นฐาน เคมีพื้นฐาน และชีววิทยาพื้นฐาน ในการวิจัย

ตัวแปรที่ใช้ในการวิจัย

การวิจัยครั้งนี้ ใช้ตัวแปรที่เกี่ยวข้องกับการพัฒนาแบบทดสอบ ได้แก่

1. การแสดงหลักฐานความเที่ยงตรง
 - 1.1 ใช้ $\rho_{T_k S_k}$ ของไฮเซน และบอร์นสเตรดต์ (Heise and Bohrnstedt.)
 - 1.2 ใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis)
2. การแสดงหลักฐานความเชื่อมั่น
 - 2.1 ใช้ค่าดัชนีความสอดคล้องของผู้ประเมิน (Rater Agreement Indexes : RAI)
 - 2.2 ใช้ทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

นิยามศัพท์เฉพาะ

1. **ทักษะกระบวนการทางวิทยาศาสตร์** (Science Process Skills) หมายถึง การปฏิบัติ และฝึกฝนความนึกคิดอย่างเป็นระบบ เป็นความสามารถ ในการแสวงหาความรู้ทางวิทยาศาสตร์ แบ่งออกเป็น 2 ประเภท คือ ทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน (Basic Science Process Skills) และทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ (Integrated Science Process Skill) ซึ่งผู้วิจัยทำการวิจัยเฉพาะในทักษะกระบวนการทาง วิทยาศาสตร์ขั้นบูรณาการ

2. **ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ** (Integrated Science Process Skills) หมายถึง วิธีการแสวงหาความรู้ทางวิทยาศาสตร์ เป็นการปฏิบัติและฝึกฝนความคิดอย่างเป็นระบบ โดยอาศัยการผสมผสานจากทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน ประกอบด้วย 5 ทักษะ ได้แก่ ทักษะการตั้งสมมติฐาน ทักษะการกำหนดและควบคุมตัวแปร ทักษะการกำหนดนิยามเชิงปฏิบัติการ ทักษะการตีความหมายข้อมูลและลงข้อสรุป และทักษะการทดลอง มีรายละเอียดดังนี้

2.1 **ทักษะการตั้งสมมติฐาน** (Formulating Hypothesis) หมายถึง ความสามารถในการ

การให้คำอธิบายซึ่งเป็นคำตอบล่วงหน้าก่อนที่จะดำเนินการทดลองเพื่อตรวจสอบความถูกต้องเป็นจริงในเรื่องนั้น ๆ โดยอาศัยทักษะกระบวนการขั้นพื้นฐานที่สำคัญ ได้แก่ การสังเกต การพยากรณ์ ประกอบกับประสบการณ์ กฎ หลักการ และทฤษฎีที่เกี่ยวข้อง

2.2 ทักษะการกำหนดและควบคุมตัวแปร (Identifying and Controlling) หมายถึง ความสามารถในการบ่งชี้และกำหนดตัวแปรต้น ตัวแปรตาม และควบคุมตัวแปรที่เป็นสาเหตุอื่น ๆ ในปรากฏการณ์หนึ่ง ๆ ที่ต้องการศึกษา โดยอาศัยทักษะกระบวนการขั้นพื้นฐานที่สำคัญ ได้แก่ การสังเกต การพยากรณ์ และการลงความเห็นจากข้อมูล เป็นต้น

2.3 ทักษะการกำหนดนิยามเชิงปฏิบัติการ (Defining Operationally) หมายถึง ความสามารถในการกำหนดความหมาย และขอบเขตของคำหรือตัวแปรต่าง ๆ ในการปฏิบัติการทดลองให้สามารถทดสอบหรือวัดได้ โดยอาศัยทักษะกระบวนการขั้นพื้นฐานที่สำคัญ ได้แก่ การสังเกต การวัด เป็นต้น

2.4 ทักษะการตีความหมายข้อมูลและลงข้อสรุป (Interpreting Data and Making Conclusion) หมายถึง ความสามารถในการแปลความหมายหรือบรรยายลักษณะและคุณสมบัติของข้อมูล พร้อมทั้งสรุปความสัมพันธ์ของตัวแปร ที่อยู่ในข้อมูลที่ต้องการศึกษา โดยอาศัยทักษะกระบวนการขั้นพื้นฐานที่สำคัญ ได้แก่ การสังเกต การวัด การจัดกระทำและสื่อความหมายข้อมูล และการลงความเห็นจากข้อมูล เป็นต้น

2.5 ทักษะการทดลอง (Experimentation) หมายถึง ความสามารถในการดำเนินการตรวจสอบสมมติฐานโดยการทดลอง ซึ่งเริ่มตั้งแต่การออกแบบการทดลอง การปฏิบัติการทดลองตามขั้นตอนที่ออกแบบไว้ ตลอดจนการใช้วัสดุอุปกรณ์ได้อย่างถูกต้อง โดยอาศัยทักษะกระบวนการขั้นพื้นฐานและทักษะกระบวนการขั้นบูรณาการทุกทักษะ

3. แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ หมายถึง ชุดข้อคำถามที่ให้ผู้ตอบมีอิสระในการเขียนคำตอบแสดงหลักวิชา ความรู้ ข้อเท็จจริงต่าง ๆ ที่สะสมความรู้จากการเรียนมาใช้ในการตอบโดยแสดงให้เห็นความสามารถในด้านทักษะการตั้งสมมติฐาน ทักษะการกำหนดและควบคุมตัวแปร ทักษะการกำหนดนิยามเชิงปฏิบัติการ ทักษะการตีความหมายข้อมูลและลงข้อสรุป และทักษะการทดลอง

4. การแสดงหลักฐานความเที่ยงตรง หมายถึง การแสดงข้อมูลจากแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ว่าสามารถวัดได้ตรงตามที่ต้องการวัดด้วยวิธีการวิเคราะห์หาความเที่ยงตรงเชิงโครงสร้างสองวิธีการดังนี้

4.1 ความเที่ยงตรงเชิงโครงสร้างด้วยสูตร $\rho_{T_k S_k}$ ของไฮเซน และบอร์นสเตดต์

(Heise and Bohrnstedt. 1970 : 116)

4.2 ความเที่ยงตรงเชิงโครงสร้างโดยใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน

(Confirmatory Factor Analysis)

5. **การแสดงผลหลักฐานความเชื่อมั่น** หมายถึง การแสดงข้อมูลจากแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ว่าสามารถวัดได้คงที่แน่นอน ด้วยวิธีการสองวิธีดังนี้

5.1 ค่าดัชนีความสอดคล้องของผู้ประเมิน (Rater Agreement Indexes : RAI) โดยใช้ผู้ตรวจ 3 คน

5.2 ค่าความเชื่อมั่นของการตรวจให้คะแนนคำนวณโดยใช้ทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

6. **ผู้ตรวจ** หมายถึง ครูที่สอนวิชาวิทยาศาสตร์ในระดับมัธยมศึกษาตอนปลาย มีประสบการณ์การสอน 5 ปีขึ้นไป จำนวน 3 คน เป็นผู้พิจารณาคำตอบและให้คะแนน

7. **กฎเกณฑ์การให้คะแนน** (Scoring Rubric) หมายถึง แนวทางที่กำหนดขึ้นเพื่อชี้บอกระดับของทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการในแต่ละด้าน ซึ่งข้อมูลได้มาจากการพิจารณาผลของผู้ตรวจให้คะแนนเพื่อให้เกิดความเข้าใจตรงกัน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการวิจัยในครั้งนี้ จำแนกตามหัวข้อสำคัญดังต่อไปนี้

1. แบบทดสอบอัตนัย หรือแบบทดสอบความเรียง
 - 1.1 นิยามของแบบทดสอบความเรียง
 - 1.2 ลักษณะและประเภทของแบบทดสอบความเรียง
 - 1.3 การสร้างแบบทดสอบความเรียง
 - 1.4 องค์ประกอบที่ส่งผลต่อคะแนนในแบบทดสอบความเรียง
2. กฎเกณฑ์การให้คะแนน
 - 2.1 ความหมายของกฎเกณฑ์การให้คะแนน
 - 2.2 ประเภทของกฎเกณฑ์การให้คะแนน
 - 2.3 การสร้างกฎเกณฑ์การให้คะแนน
3. ทฤษฎีค่าความเชื่อมั่น
 - 3.1 การประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม
 - 3.2 การประมาณค่าความเชื่อมั่นที่เกี่ยวกับการตรวจให้คะแนน
 - 3.3 ดัชนีความสอดคล้องของผู้ประเมิน
 - 3.4 การประมาณค่าความเชื่อมั่นตามโครงสร้างทฤษฎีการสรุปอ้างอิง
4. หลักฐานแสดงความเที่ยงตรง
 - 4.1 ความหมายของความเที่ยงตรง
 - 4.2 วิธีการแสดงหลักฐานความเที่ยงตรง
5. ทักษะกระบวนการทางวิทยาศาสตร์
 - 5.1 ความหมายทักษะกระบวนการทางวิทยาศาสตร์
 - 5.2 ประเภทของทักษะกระบวนการทางวิทยาศาสตร์
6. หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544
7. งานวิจัยที่เกี่ยวข้อง
 - 7.1 งานวิจัยที่เกี่ยวข้องกับความเที่ยงตรงและความเชื่อมั่น
 - 7.2 งานวิจัยที่เกี่ยวข้องกับทักษะกระบวนการทางวิทยาศาสตร์

1. แบบทดสอบแบบอัตนัย หรือ แบบทดสอบความเรียง (Essay Type Test)

1.1 นิยามของแบบทดสอบความเรียง

คอฟแมน (นิศารัตน์ คงสวัสดิ์. 2544 : 3 ; อ้างอิงจาก Coffman. 1971 : 271) ได้นิยามการทดสอบแบบความเรียง (Essay Examination) ว่าหมายถึง การสอบที่ใช้ข้อสอบแบบความเรียง ตั้งแต่หนึ่งข้อขึ้นไปกับกลุ่มนักเรียนภายใต้สถานการณ์อย่างหนึ่ง เพื่อเก็บข้อมูลในการประเมินผลคำถามแบบความเรียงแตกต่างกับคำถามแบบตอบสั้น ๆ ตรงที่ต้องใช้ผู้เชี่ยวชาญตัดสิน ไม่ใช่การตรวจอย่างง่าย โดยใช้ตัวเฉลยแบบเดียวกับแบบทดสอบปรนัย แตกต่างจากรายงานที่ทำส่งครูในการเรียนวิชาต่าง ๆ ตรงที่การดำเนินการสอบจะต้องดำเนินให้เป็นมาตรฐานเดียวกัน และต้องการข้อสอบที่เป็นตัวแทนผลการสอบที่ต้องการ

กลอนแลนด์ (ชาญวิทย จรัสสุทธิศิร. 2545 : 31 ; อ้างอิงจาก Gronlund. 1982 : 71) สรุปความหมายของแบบทดสอบความเรียงไว้ว่า แม้ว่าจะกล่าวว่าแบบทดสอบแบบความเรียงจะไม่มีประสิทธิภาพในการวัดผลการเรียนรู้ แต่ข้อสอบแบบนี้จะให้อิสระในการตอบซึ่งเป็นสิ่งที่จำเป็นในการวัดผลที่ซับซ้อน รวมไปถึงการวัดความสามารถในการสร้าง การจัดระเบียบ การรวมเป็นหนึ่งเดียว การแสดงออกและพฤติกรรมที่คล้ายคลึงกันซึ่งก่อให้เกิดการสร้างและการสังเคราะห์ความคิด

ฮอบกินส์ และแอนเตส (Hopkins and Antes. 1990 : 231) ให้ความหมายของแบบทดสอบความเรียงไว้ว่า เป็นข้อสอบประเภทข้อสอบขยายความ (Constructed – response Items) ที่สมบูรณ์ที่สุด เป็นข้อสอบที่ต้องการเห็นพัฒนาการในการตอบคำถามของนักเรียน

เกย์ (นิศารัตน์ คงสวัสดิ์. 2544 : 7 ; อ้างอิงจาก Gay. 1991 : 224) กล่าวว่าแบบทดสอบความเรียง เป็นแบบทดสอบหนึ่งที่ประกอบด้วยหัวข้อที่ต้องการให้นักเรียนเขียนตอบ ซึ่งโดยปกติแล้วจะเป็นคำตอบที่ยาว ซึ่งแบบทดสอบความเรียงไม่เหมาะสมสำหรับวัดผลระดับความรู้ความจำ แต่เหมาะสำหรับวัดผลในระดับสูง เพราะแบบทดสอบแบบนี้จะต้องมีการสังเคราะห์ และการประเมินค่า

คูบิสซ์ และบอริช (Kubiszyn and Borich. 1999 : 112) กล่าวว่า แบบทดสอบความเรียงเป็นแบบทดสอบสำหรับให้นักเรียนนำเสนอ มากกว่าเลือกตัวเลือกที่ถูกต้อง แบบทดสอบชนิดนี้ต้องให้นักเรียนเรียบเรียงคำตอบที่กว้างกว่าเดิมเพื่อตอบคำถามที่ไม่ใช่คำถามเดียว

วอร์ด (Ward. 1999 : 129) ให้ความหมายของแบบทดสอบความเรียงว่า เป็นแบบทดสอบที่วัดทักษะขั้นสูงที่ยากที่จะสร้างด้วยแบบทดสอบแบบปรนัยได้ แบบทดสอบประเภทนี้ต้องการให้นักเรียนแสดงความรู้ กระบวนการของความรู้ การรวบรวมข้อมูล และการนำเสนอความรู้ นั้นในรูปของเหตุผลที่ปะติดปะต่อกันได้

โกวิท ประวาลพุกษ์ และสมศักดิ์ สินธุระเวชญ์ (2525 : 101) ให้นิยามแบบทดสอบความเรียงว่า เป็นข้อสอบที่ให้เด็กได้มีอิสระในการตอบเต็มที่ ให้โอกาสเด็กได้แสดงว่าเขารู้อะไร สามารถวัดสมรรถภาพด้านบูรณาการความคิด ความสามารถในการจัดรวบรวมความคิดที่สมบูรณ์ (Organization Ability) การสังเคราะห์ และการประเมินค่าอย่างมีประสิทธิภาพ เช่น ความรู้ในการแก้ปัญหาใหม่ ๆ และเป็นต้นคิดหรือค้นหาวิธีการใหม่ ๆ เพื่อแก้ปัญหาต่าง ๆ ได้

วิญญา วิศาลาภรณ์ (2533 : 60) ให้ความหมายของแบบทดสอบความเรียง ว่า เป็นข้อสอบที่ให้เสรีภาพเป็นอย่างมากแก่นักเรียนในการตอบ นักเรียนมีอิสระในการที่จะเสนอความคิดด้วยภาษาของตนเอง นักเรียนต้องใช้ความสามารถอธิบายปัญหาต่าง ๆ ได้ ดังนั้น ข้อสอบแบบความเรียงหรือข้อสอบอัตนัยจึงสามารถใช้วัดความรู้ความคิดขั้นสูงที่ซับซ้อนได้ ใช้วัดทัศนคติได้ นักเรียนต้องมีความสามารถในการจัดระเบียบความรู้ความคิด แสดงความริเริ่มและรู้จักสังเคราะห์

บุญชม ศรีสะอาด (2534 : 80) กล่าวถึงแบบทดสอบความเรียงว่า เป็นข้อสอบที่ให้ผู้สอบเขียนตอบตามความคิดของตนเอง แม้ว่าโดยทั่วไปจะได้รับการนำไปใช้น้อย แต่ก็เป็นเครื่องมือที่จำเป็นในการวัดผลโดยเฉพาะในการวัดเกี่ยวกับความคิดเห็น เจตคติ ความสามารถในการคิดสังเคราะห์ (Synthesis) ความสามารถในการอธิบายให้คนอื่นเข้าใจ ความสามารถในการบูรณาการความรู้ การวัดในลักษณะเหล่านี้ข้อสอบแบบอัตนัยจะวัดได้ดีมาก

ล้วน สายยศ และอังคณา สายยศ (2539 : 86) ได้กล่าวถึงแบบทดสอบความเรียง ว่า มีจุดประสงค์วัดความสามารถในการบรรยาย อธิบายและแสดงเหตุผลตามความคิดเห็นของตน อาจจำกัดความยาวหรือให้เขียนได้ตามสบายก็ได้

จากการศึกษานิยามของแบบทดสอบความเรียงจากนักการศึกษาหลาย ๆ ท่านที่กล่าวมานี้ ผู้วิจัยสรุปได้ว่าแบบทดสอบความเรียงเป็นแบบทดสอบที่ให้ผู้ตอบมีอิสระในการตอบอย่างเต็มที่ตามความคิดของตนเอง ผู้ตอบได้แสดงความสามารถทางการใช้ภาษา การใช้เหตุผล ซึ่งแบบทดสอบแบบนี้เหมาะสำหรับใช้วัดประเมินความสามารถทางความคิดในระดับความเข้าใจ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า

1.2 ลักษณะและประเภทของแบบทดสอบความเรียง

1.2.1 การเปรียบเทียบลักษณะของแบบทดสอบความเรียงกับแบบทดสอบอื่น ๆ

กลอนลันด์ (Gronlund. 1977 : 73) ได้เปรียบเทียบสรุปลักษณะระหว่างข้อสอบแบบปรนัย และข้อสอบแบบความเรียง ไว้ดังต่อไปนี้

ตาราง 1 การเปรียบเทียบลักษณะข้อสอบปรนัย และข้อสอบแบบความเรียง

	ข้อสอบแบบปรนัย	ข้อสอบแบบความเรียง
การวัดผลการเรียนรู้	ดีสำหรับวัดผลการเรียนรู้ขั้นความรู้ความเข้าใจ การนำไปใช้ และการวิเคราะห์ แต่ไม่เพียงพอต่อการวัดการสังเคราะห์ และการประมาณค่า	ไม่มีประสิทธิภาพในการวัดผลการเรียนรู้ แต่ดีในการวัดความเข้าใจ การนำไปใช้ และการวิเคราะห์ และดีที่สุดสำหรับการวัดผลการสังเคราะห์และประเมินค่า
การสุ่มตัวอย่างเนื้อหา	การใช้ข้อสอบจำนวนหลายข้อจะมีความครอบคลุมกว้าง ซึ่งทำให้การสุ่มตัวอย่างเนื้อหาได้ครอบคลุม	การใช้ข้อสอบจำนวนไม่มากข้อ ส่งผลให้มีขอบเขตจำกัด ทำให้การสุ่มตัวอย่างเนื้อหาไม่ครอบคลุม
การเตรียมการสร้าง	การเตรียมคำถามที่ดีทำได้ยาก และใช้เวลานาน	การเตรียมคำถามที่ดีทำได้ยาก แต่ง่ายกว่าคำถามแบบปรนัย
การให้คะแนน	มีความเป็นปรนัย ง่าย และมีความเชื่อมั่นสูง	มีความเป็นอัตนัย ยาก และให้ความเชื่อมั่นต่ำ
ปัจจัยที่ส่งผลต่อคะแนนที่ได้	ใช้ความสามารถในการอ่านและการเดา	ใช้ความสามารถในการเขียนและการลง
ผลกระทบต่อการเรียนที่เป็นไปได้	กระตุ้นให้นักเรียนจำ ดีความ และวิเคราะห์ความคิดของผู้อื่น	กระตุ้นให้นักเรียนจัดระเบียบ รวบรวม และแสดงความคิดเห็นของตนเอง

เวียสมา และคนอื่น ๆ (Wiersma and others. 1985 : 168) ได้เปรียบเทียบลักษณะที่แตกต่างระหว่างแบบทดสอบความเรียงกับแบบทดสอบปรนัย ไว้ดังนี้

1. แม้ว่าการให้คะแนนแบบทดสอบความเรียงจะมีความเป็นปรนัยน้อยกว่าแบบทดสอบปรนัย แต่ยังมีจุดสำคัญต่าง ๆ ที่ดีกว่าแบบทดสอบปรนัย
2. ข้อสอบปรนัยถูกฟอร์มด้วยส่วนคำถาม และส่วนคำตอบตามรูปแบบของแบบทดสอบที่ดี แต่แบบทดสอบความเรียงจะมีส่วนคำถามที่มีตั้งแต่ขนาดสั้นบรรทัดเดียว จนถึงยาวเป็นหน้า ๆ

3. การตอบแบบทดสอบความเรียงต่อข้อ ใช้เวลานานกว่าการตอบแบบทดสอบปรนัย

4. คำตอบของแบบทดสอบความเรียง ต้องอาศัยทักษะในการเขียนมากกว่าการตอบคำถามแบบทดสอบปรนัยที่ไม่ต้องอาศัยทักษะกระบวนการเขียนเลย

1.2.2 ข้อดี ข้อเสียของแบบทดสอบความเรียง

ฮอบกินส์ และแอนเตส (Hopkins and Antes. 1990 : 232 – 233) ได้สรุปจุดแข็งจุดอ่อนของแบบทดสอบความเรียง ไว้ว่าการใช้แบบทดสอบความเรียงสอบนักเรียนนั้นทำให้นักเรียนมีอิสระในการตอบเต็มที่ ความมีอิสระมากหรือน้อยขึ้นอยู่กับลักษณะข้อสอบที่สร้างมาว่าต้องการการตอบแบบใด ทำให้สามารถทดสอบพฤติกรรมการเรียนรู้ของนักเรียนในระดับที่สูงขึ้น เช่น การวิเคราะห์ การสังเคราะห์ และการประเมินค่า โดยการให้นักเรียนนำความรู้และหลักการต่าง ๆ ไปประยุกต์ใช้ , เปรียบเทียบ , อธิบาย และอธิบายความเหมือนความต่างของความรู้และหลักการ ซึ่งทำให้นักเรียนได้แสดงความคิดเห็นโดยไม่มีการตัดสินว่าความคิดเห็นนั้นถูกหรือผิด ไม่เหมือนข้อสอบปรนัย ซึ่งให้คะแนนตามความถูกผิด ปัจจัยสำคัญในการตอบข้อสอบแบบความเรียง คือความสามารถของนักเรียนในการสนับสนุนความคิดเห็นของตน ความจำเป็นในการจัดระเบียบและการแสดงความคิดเห็น ช่วยเปิดโอกาสให้นักเรียนมีพัฒนาการทางทักษะในการเขียน นอกจากนี้การให้อิสระในการตอบเต็มที่ ทำให้นักเรียนวิเคราะห์หรือประเมินการเขียน และจัดองค์ประกอบต่าง ๆ ได้ และข้อสอบแบบความเรียงที่ดี จะใช้เวลาในการออกข้อสอบน้อยกว่าข้อสอบแบบปรนัย เพราะมีหัวข้อให้เขียนเพียงไม่กี่หัวข้อ ไม่เหมือนข้อสอบแบบปรนัยที่ต้องคิดและเขียนตัวเลือกมากมาย

ส่วนจุดอ่อนของแบบทดสอบความเรียงจะมีผลต่อการตรวจคำตอบของนักเรียนที่ใช้เวลานานมาก ยากที่จะให้คะแนน ถ้าไม่มีเกณฑ์ในการให้คะแนน ผลจากการวัดก็จะไม่มีประโยชน์ คำตอบที่ได้มักเป็นส่วนหนึ่งของเนื้อหาวิชาเท่านั้น ซึ่งอาจน้อยกว่าสิ่งที่ต้องการวัด และถ้านักเรียนไม่สามารถเขียนแสดงความคิดเห็นออกมาได้ ก็อาจจะแสดงผลสำเร็จที่ไม่ดี นักเรียนคนที่มีความสำเร็จในการเรียนรู้เท่ากัน แต่มีระดับของทักษะการแสดงออกทางความคิดต่างกัน ก็อาจจะได้คะแนนที่ต่างกันมากในการสอบด้วยแบบทดสอบความเรียง

คูบิสซ์ และบอริช (Kubiszyn and Borich. 1999 : 109) ได้รวบรวมข้อดี และข้อเสียของข้อสอบแบบความเรียงไว้ 4 อย่าง ดังนี้

ข้อดี

1. มีประสิทธิภาพมากที่สุดในการประเมินผลการเรียนรู้ที่ซับซ้อน
2. การสร้างข้อสอบง่าย
3. เน้นย้ำในด้านทักษะการสื่อสาร หมายความว่าครูต้องใช้เวลาในการสอนทักษะ

การสื่อสารที่ตรงกับหลักสูตร รวมถึงคำศัพท์พิเศษและรูปแบบในการเขียน

4. ไม่มีการเดา

ข้อเสีย

1. ยากต่อการให้คะแนน
2. คะแนนที่ได้ไม่มีความแน่นอนเชื่อถือได้
3. จำกัดเนื้อหาจากการสอนทั้งหมด
4. การลวง เนื่องจากว่ายิ่งคำตอบของนักเรียนยาวเท่าไร ก็มีแนวโน้มที่จะได้

คะแนนมากกว่าคำตอบที่สั้นมากเท่านั้น โดยไม่ได้คำนึงถึงเนื้อหา

บุญเชิด ภิญโญอนันตพงษ์ (2547: 65) ได้สรุปข้อดี และข้อจำกัดของแบบความเรียง ดังนี้

ตาราง 2 ข้อดี และข้อจำกัดของแบบความเรียง

ข้อดี	ข้อจำกัด
1. วัดกระบวนการความคิด เช่น การจัดระเบียบ และโครงสร้าง, การคัดเลือกความคิดที่สำคัญ ๆ ที่เกี่ยวข้อง ความสามารถในการเขียน ได้ดี.	1. การให้คะแนนไม่แน่นอน คะแนนที่ได้ขึ้นอยู่กับผู้ตรวจ เช่น อารมณ์ ทัศนคติของผู้ตรวจลายมือของผู้ตอบด้วย
2. วัดความคิดริเริ่มสร้างสรรค์ ความสามารถประเมินค่าได้ดี	2. ขาดความเที่ยงตรงทางเนื้อหา เพราะข้อสอบได้น้อยจึงไม่ครอบคลุม
3. วัดทัศนคติ ข้อคิดเห็นต่าง ๆ ได้ดี	3. ตรวจข้อสอบยากและเสียเวลามาก
4. ผู้ตอบมีอิสระเขียนแสดงความคิดเห็นได้เต็มที่	

1.2.3 ประเภทของแบบทดสอบความเรียง

กลอนลันด์ (Gronlund. 1977 : 74 – 75) ได้แบ่งรูปแบบคำถามของแบบทดสอบความเรียงไว้ 2 ประเภท คือ

1. คำถามที่ต้องการกำหนดขอบเขตของคำตอบ (Restricted – Response Questions) คำถามนี้จะมีการจำกัดคำตอบที่นักเรียนต้องตอบลงในข้อสอบ โดยทั่วไปแล้วขอบเขตของเนื้อหาที่ใช้พิจารณาจะถูกกำหนดขึ้นอย่างแคบ ๆ ด้วยปัญหาและรูปแบบเฉพาะของคำตอบก็จะระบุไว้ด้วยเช่นกัน (โดยคำว่า “เขียนเป็นข้อ ๆ” “ให้คำนิยาม” และ “ให้เหตุผล” เป็นต้น) ในบางกรณีการตอบอาจจะถูกจำกัดมากขึ้นไปอีก ด้วยการใช้ส่วนเกริ่นนำหรือการใช้คำสั่งพิเศษ

คำถามรูปแบบจำกัดคำตอบนี้มีทั้งข้อดีและข้อเสีย คำถามแบบนี้สามารถเตรียมได้ง่ายกว่า สัมพันธ์กับผลการเรียนรู้เฉพาะโดยตรงมากกว่า และให้คะแนนง่ายกว่า อย่างไรก็ตามคำถามแบบนี้ก็ให้โอกาสนักเรียนน้อยในการที่จะแสดงความสามารถของพวกเขาในด้านการจัดระเบียบ การรวบรวม และการพัฒนารูปแบบในการตอบแบบใหม่ ๆ หากต้องการวัดผลในระดับการสังเคราะห์ และประเมินค่า ก็ควรใช้คำถามแบบไม่จำกัดคำตอบ จะเหมาะสมกว่า

2. คำถามแบบไม่จำกัดคำตอบ (Extended – Response Questions) คำถามนี้จะให้อิสระในการตอบกับนักเรียน ไม่จำกัดในการกำหนดรูปแบบและขอบเขตการตอบ แม้ว่าจะมีการจำกัดในทางปฏิบัติบ้าง เช่น การจำกัดด้านเวลา หรือนำกระดาษในการตอบ นักเรียนมีอิสระในการแสดงทักษะในการวิเคราะห์และประเมินผล ดังนั้นปริมาณโครงสร้างจึงหลากหลายกันไปในแต่ละข้อโดยขึ้นอยู่กับผลการเรียนรู้ที่ต้องการวัด

คำถามที่ไม่จำกัดคำตอบนี้จะทำให้เกิดการรวบรวมความคิดในเชิงสร้างสรรค์ การประเมินเนื้อหาโดยรวมและวิธีการแก้ปัญหาอย่างกว้าง ๆ สิ่งเหล่านี้ เป็นผลการเรียนรู้ที่สำคัญ ซึ่งไม่สามารถวัดได้ด้วยรูปแบบข้อสอบประเภทอื่น ปัญหาใหญ่ที่สุดของข้อสอบลักษณะนี้ คือ การประเมินคำตอบด้วยความเชื่อมั่นที่สุด เพื่อให้ได้ผลการเรียนรู้ที่มีประโยชน์ สิ่งเหล่านี้เป็นงานที่ยากและใช้เวลานาน แต่ผลที่ได้ก็นับว่าคุ้มค่า

ฮอปกินส์ และแอนเตส (Hopkins and Antes. 1990 : 231 – 232) ได้เสนอรูปแบบของแบบทดสอบความเรียงอีกครั้งว่า ข้อสอบประเภทนี้ต้องการคำตอบที่เป็นประโยคหลายประโยคต่อเนื่องกัน ซึ่งแสดงความสมเหตุสมผล ผู้ตรวจความถูกต้องและคุณภาพของข้อสอบแบบนี้ต้องเป็นผู้ที่มีความรู้ในวิชาที่สอบ ซึ่งแบ่งออกเป็น 2 ลักษณะ คือ

1. แบบที่มีขอบเขตกว้างแต่ให้เวลาจำกัด (Extended – response) ข้อสอบลักษณะนี้จะทดสอบทักษะของผู้เขียน ในการเลือกว่าจะเขียนอะไร และจะเขียนมากน้อยแค่ไหนในแต่ละส่วน รวมทั้งความสามารถในการจัดระเบียบ และแสดงความคิดในเชิงตรรกะ ข้อสอบเขียนบรรยาย แบบขยายความนี้ เป็นข้อสอบประเภทปลายเปิด (Open – ended) และไม่ได้จำกัดความคิดของนักเรียน ข้อสอบแบบนี้มีประโยชน์ในการประเมินการเขียนโดยเฉพาะศิลปะเกี่ยวกับการใช้ภาษา

2. แบบจำกัดคำตอบ (Limited – response) จะมีขอบเขตแคบหรือกว้างก็ได้ แต่นักเรียนควรจะรู้อย่างแน่ชัดว่าขีดจำกัดนั้นคืออะไร ข้อสอบแบบนี้จะให้ขีดจำกัดที่ระบุได้อย่างชัดเจน เพื่อสร้างขอบเขตในการตอบให้นักเรียน กล่าวคือนักเรียนต้องตอบในรอบเนื้อหาที่จำกัด ข้อสอบแบบนี้มีประโยชน์ต่อครูผู้สอนที่ต้องการทดสอบผลสัมฤทธิ์ทางการเรียนซึ่งต่างจากแบบทดสอบความเรียงแบบที่มีขอบเขตกว้างที่ต้องการวัดด้านทักษะการเขียน ผู้ออกข้อสอบควรพิจารณาถึงจุดมุ่งหมายในการวัดผลก่อนตัดสินใจเลือกว่าจะใช้รูปแบบใด

บุญเชิด ภิญญอนันตพงษ์ (2547 : 60 – 61) ได้เสนอประเภทของแบบทดสอบความเรียง ไว้ 2 แบบ คือ

1. แบบไม่จำกัดคำตอบ หรือแบบขยายความ ข้อสอบแบบนี้จะถามความรู้ความสามารถต่าง ๆ โดยให้ผู้ตอบแสดงความคิดเห็นอย่างอิสระ สามารถวัดสมรรถภาพด้านความคิดริเริ่มสร้างสรรค์ ทักษะคิด การประเมินค่าได้อย่างกว้างขวาง ปริมาณคำตอบจึงขึ้นอยู่กับคำถามและความรู้ที่ผู้สอบว่ามีมากน้อยเพียงใด การกำหนดเวลาให้เขียนตอบจึงต้องกำหนดให้เหมาะสม ข้อสอบแบบนี้ส่งเสริมให้ผู้เรียนรู้จักการรวบรวมความคิดต่าง ๆ การประเมินคุณค่าของสิ่งเหล่านั้น และการใช้วิธีการต่าง ๆ ในการแก้ปัญหา ซึ่งข้อสอบแบบนี้ ไม่สามารถจะนำมาใช้วัดได้

2. แบบจำกัดคำตอบ ข้อสอบแบบนี้จะถามแบบจำเพาะเจาะจง และต้องการคำตอบเฉพาะเรื่อง ซึ่งผู้ตอบต้องจัดเรียงความคิดให้เป็นระเบียบ เพื่อให้ตรงประเด็นของคำถามเพียงสั้น ๆ ดังนั้นจึงต้องระมัดระวังเรื่องคำสั่งของโจทย์ ขอบเขตของเนื้อหา เวลาที่ให้นักเรียนเขียนตอบ

1.2.4 ลักษณะของแบบทดสอบความเรียง

บุญเชิด ภิญญอนันตพงษ์ (2547 : 59) กล่าวถึงลักษณะของแบบทดสอบความเรียงว่า ข้อสอบความเรียงจะเขียนคำถาม หรือโจทย์ โดยกำหนดเป็นสถานการณ์หรือปัญหาในรูปแบบหนึ่งอย่างกว้าง ๆ หรือเฉพาะเจาะจงเพื่อให้ผู้ตอบได้แสดงความรู้ ความเข้าใจ ความคิดเห็น ทักษะคิดได้อย่างไม่จำกัด ดังนั้นโดยทั่วไปแบบทดสอบความเรียงจึงมีจำนวนข้อไม่มากนัก ส่วนมากมีจำนวนข้อไม่เกิน 10 ข้อ ลึก แล กว้าง แตกต่างกันไป การเตรียมนักเรียนเพื่อตอบคำถามเป็นสิ่งจำเป็นและสำคัญมาก ซึ่งทำได้โดยการฝึกหัดให้นักเรียนได้รู้และเข้าใจขอบเขตของคำในคำถาม และให้รู้จักแบ่งเวลาเพื่อจะไม่ให้ใช้เวลาทำข้อใดข้อหนึ่งมากเกินไป จะช่วยให้การวัดและการตรวจให้คะแนนยุติธรรมและแม่นยำยิ่งขึ้น

คำตอบของข้อสอบอัตนัยจะมีลักษณะและปริมาณไม่แน่นอน เนื่องจากผู้ตอบต้องเขียนตอบ โดยเรียบเรียงสำนวนจากความคิด ความเข้าใจอย่างอิสระ ดังนั้นคำตอบจึงอาจมีเพียงประโยคเดียวจนถึงหลาย ๆ หน้ากระดาษ การตอบแบบทดสอบความเรียงจึงมีลักษณะดังนี้ 1) จัดระเบียบคำตอบภายในเวลาที่กำหนดให้ 2) ใช้สำนวนภาษา และแบบฉบับของตนเองเขียนตอบ 3) เขียนตอบไม่มากจนเกินไป 4) เขียนคำตอบให้ครอบคลุมคำถามอย่างสมบูรณ์และระมัดระวัง

พวงรัตน์ ทวีรัตน์ (2530 : 108) กล่าวว่าลักษณะของแบบทดสอบแบบความเรียงนี้มี 2 ส่วน คือ ส่วนที่เป็นคำถาม ซึ่งอาจกำหนดในรูปแบบสถานการณ์หรือคำถามในรูปแบบใดรูปแบบหนึ่งอย่างกว้าง ๆ หรือบางครั้งก็เฉพาะเจาะจงเพื่อให้ผู้ตอบได้แสดงความรู้ ความเข้าใจ ความคิดเห็น หรือทักษะคิดอย่างอิสระ ลักษณะการถามจะมีความตื่นลึกแตกต่างกันแล้วแต่ผู้ออกข้อสอบต้องการ แต่โดยทั่วไปมักเป็นที่เข้าใจว่า คำถามของแบบทดสอบความเรียงจะเป็นคำถามที่แคบลึกในประเด็น

โดยประเด็นหนึ่ง ผู้ตอบต้องมีความรู้ในเรื่องนั้นเป็นอย่างดีจึงจะสามารถเขียนตอบได้ การเตรียมตัวเพื่อตอบข้อสอบอัตนัยจึงเป็นสิ่งจำเป็นมาก โดยปกติแบบทดสอบความเรียงจะมีจำนวนคำถามไม่มากนัก และคำถามแต่ละข้ออาจแบ่งเป็นคำถามย่อย ๆ ได้อีก และบางครั้งก็ระบุการให้คะแนนไว้ด้วย เพื่อช่วยให้การตรวจมีความเชื่อถือได้มากยิ่งขึ้น ส่วนที่เป็นคำตอบ ซึ่งมีลักษณะและปริมาณไม่แน่นอน และโดยทั่วไปมักจะอยู่ในรูปของการเขียนบรรยาย โดยผู้ตอบจะเขียนเรียบเรียงสำนวนจากความคิดของตนเอง

1.3 การสร้างแบบทดสอบความเรียง

ได้มีนักวัดผลการศึกษาทั้งในและต่างประเทศนำเสนอขั้นตอนในการสร้างแบบทดสอบความเรียงไว้หลาย ๆ ท่าน ดังนี้

กลอนลันด์ (Gronlund. 1977 : 75 – 76) เสนอกฎการตั้งคำถามแบบทดสอบความเรียงเพื่อให้ข้อสอบมีคุณภาพสูงไว้ดังนี้

1. ให้ใช้แบบทดสอบความเรียงเพื่อวัดผลการเรียนที่ซับซ้อนเท่านั้น
2. เชื่อมโยงคำถามกับผลการเรียนรู้ที่ต้องการวัดโดยตรงที่สุดเท่าที่จะเป็นไปได้
3. สร้างคำถามที่นักเรียนเข้าใจได้อย่างชัดเจน
4. ไม่มีตัวเลือกของคำถามให้นักเรียนเลือกตอบ ยกเว้นว่าผลการเรียนรู้ต้องการให้เป็นไปเช่นนั้น

5. ให้เวลาในการตอบอย่างเพียงพอ และระบุเวลาไว้ในแต่ละคำถามด้วย

ฮอปกินส์ และแอนเตส (Hopkins and Antes. 1990 : 233 – 234) เสนอข้อแนะนำในการสร้างและพัฒนาแบบทดสอบความเรียงไว้ดังนี้

1. ใช้แบบทดสอบความเรียงเมื่อต้องการคำตอบของนักเรียนที่เป็นการอธิบาย การเปรียบเทียบความเหมือน การเปรียบเทียบความแตกต่าง การบรรยาย การวิเคราะห์ การสังเคราะห์ หรือประเมินผล
2. ไม่ใช้แบบทดสอบแบบความเรียงในการวัดความจำ
3. ใช้ถ้อยคำในการตั้งคำถามที่ชัดเจนและกระชับ
4. ตั้งคำถามด้วยสถานการณ์ใหม่ ๆ ที่นักเรียนไม่คุ้นเคย ไม่เช่นนั้นนักเรียนก็จะจำสิ่งที่เรียนมาแล้วนั้นมาตอบแบบเดิม ๆ
5. กำหนดขอบเขตในการตอบที่ต้องการไว้ในข้อสอบด้วย และขีดจำกัดในการตอบควรกำหนดตามเนื้อหาที่นักเรียนได้เรียน นอกจากนี้ความยากของคำตอบที่ต้องการก็อาจจะกำหนดโดยเวลา จำนวนข้อ หรือปริมาณที่ว่างในกระดาษคำตอบสำหรับเขียนตอบ
6. บอกให้นักเรียนรู้ว่าจะจัดการกับคำศัพท์เฉพาะหรือศัพท์เทคนิคในเนื้อหาที่จะสอบ

อย่างไร นักเรียนควรรู้ว่าจะต้องให้คำจำกัดความหรืออธิบายคำเหล่านี้ในการตอบด้วยหรือไม่

7. เขียนข้อสอบเป็นข้อ ๆ เพื่อวัดด้านต่าง ๆ ที่ได้ศึกษามา และไม่ควรออกข้อสอบที่ให้ผู้สอบนำข้อมูลทั่วไปมาตอบ

8. ถ้อยคำที่ใช้ในข้อสอบหรือคำสั่งจะต้องมีความหมายชัดเจน

9. ให้นักเรียนทุกคนตอบข้อสอบด้วยชุดข้อสอบชุดเดียวกันและเหมือนกัน

10. เลือกใช้แบบทดสอบแบบความเรียงแบบขยายความ เพื่อวัดทักษะการเขียน การสื่อสารหรือการจัดองค์ประกอบเท่านั้น

เวอร์ด (Ward. 1999 : 130) ให้คำแนะนำสำหรับการเขียนแบบทดสอบความเรียงไว้ 5 รูปแบบ ดังนี้

1. ตั้งคำถามเกี่ยวกับความคิดรวบยอดที่สำคัญ
2. ให้นักเรียนแสดงความรู้ของเขา ไม่ใช่รายงานข้อมูลที่มีอยู่แล้ว
3. ให้นักเรียนแสดงความสามารถของตนอย่างสมบูรณ์ที่สุดเท่าที่จะทำได้
4. คำถามต้องสั้นและชัดเจนมากกว่าที่จะเป็นคำถามทั่วไป
5. แจ้งหัวข้อที่จะให้คะแนนให้ผู้สอบทราบก่อนทำข้อสอบทุกครั้ง

บุญเชิด ภิญญอนันตพงษ์ (2525 : 34 – 41) ได้ให้คำแนะนำในการสร้างข้อสอบอัตนัยให้มีคุณภาพ ต้องคำนึงถึงจุดอ่อน 2 ประการ คือ ทำอย่างไรจึงจะสร้างคำถามให้สามารถวัดในสิ่งที่ต้องการวัดได้ และทำอย่างไรจึงจะทำให้การตรวจข้อสอบอัตนัยมีความเชื่อถือได้มากที่สุด ดังนั้นในการสร้างข้อสอบจะต้องเขียนข้อคำถาม และกำหนดแนวการตอบและเกณฑ์สำหรับตัดสินให้คะแนนอย่างละเอียดด้วย สาเหตุหนึ่งที่ทำให้คะแนนในการตรวจเชื่อถือไม่ค่อยได้ เพราะการเฉลยส่วนนี้ไป ดังนั้นการสร้างข้อสอบอัตนัยจึงควรคำนึงถึงข้อแนะนำต่อไปนี้

1. วิเคราะห์จุดประสงค์ของวิชาที่จะออกข้อสอบให้ละเอียดครอบคลุม และเลือกรูปแบบคำถามให้เหมาะสมกับจุดประสงค์นั้น ๆ
2. เขียนคำถามให้ชัดเจนและตรงจุด ระบุให้แน่ชัดว่าต้องการให้ผู้ตอบทำอะไร
3. เขียนคำชี้แจงให้ละเอียดและชัดเจน
4. ควรเขียนข้อสอบแบบจำกัดคำตอบ มากกว่าข้อสอบแบบไม่จำกัดคำตอบเพื่อให้การตรวจมีความเชื่อมั่นสูงขึ้น
5. หลีกเลี่ยงคำถามประเภทวัดความจำอย่างเดียว เพราะการวัดความจำสามารถใช้ข้อสอบปรนัยอื่น ๆ วัดได้ดีกว่า
6. ควรเขียนข้อสอบให้มากข้อ และหลายรูปแบบ เพื่อให้ครอบคลุมเนื้อหา และสิ่งที่ต้องการจะวัด ซึ่งช่วยให้การวัดมีความเที่ยงตรงตามเนื้อหา (Content Validity) สูงขึ้น

7. ควรกำหนดเวลาที่ใช้ตอบแต่ละข้ออย่างคร่าว ๆ การคาดคะเนเวลาที่จะใช้เขียนตอบจะเป็นการช่วยให้ผู้ตอบสามารถกำหนดขอบเขตในการตอบได้เหมาะสม และสามารถทำข้ออื่น ๆ ได้ทัน

8. ควรเฉลยคำตอบไว้ทั้งคำตอบที่ถูกต้องและแนวคำตอบอื่น ๆ ที่ควรจะเป็นเพื่อสามารถให้คะแนนได้เชื่อมั่นยิ่งขึ้นและยังเป็นการตรวจสอบดูว่าโจทย์ชัดเจนพอที่จะให้ตอบตามคำตอบหรือไม่ อีกทั้งเป็นการคาดคะเนคะแนนและเวลาให้ตอบที่เหมาะสมอีกด้วย

9. ถ้าข้อสอบมีมากข้อ และยากไม่เท่ากัน ควรจัดเรียงลำดับจากง่ายหายาก เพื่อช่วยให้ผู้ตอบเกิดกำลังใจในการตอบ การให้ข้อยากไว้ตอนต้นอาจทำให้เด็กตกใจหรือขวัญเสีย ซึ่งจะกระทบกระเทือนคะแนนข้ออื่น ๆ ได้ ดังนั้นควรเขียนคำชี้แจงให้นักเรียนเลือกตอบข้อใดก่อนก็ได้ เพื่อให้นักเรียนเลือกทำข้อง่ายตามความรู้สึกของนักเรียนเอง

10. ถ้าเป็นไปได้ควรออกข้อสอบให้มีความยากใกล้เคียงกัน

11. ให้ผู้อื่นที่มีความรู้ในสาขาวิชาที่ออกข้อสอบ อ่าน วิพากษ์วิจารณ์ เพื่อตรวจสอบข้อบกพร่อง เช่น เขียนคำถามครอบคลุมสิ่งสำคัญที่ต้องการวัดหรือไม่ คำถามยากพอเหมาะหรือไม่

12. หลีกเลี่ยงวิธีการสอบโดยให้เปิดตำราสอบ เพราะการสอบโดยให้เปิดตำราดูได้นี้ เด็กที่เรียนเก่งอยู่แล้วมีความชำนาญในการใช้หนังสือ มีทักษะการอ่าน มีความคล่องแคล่ว มีความพยายาม และมีความอดทนสูงกว่าเด็กที่เรียนอ่อน ดังนั้นเด็กเก่งจึงมีแนวโน้มที่จะได้คะแนนสูงกว่าปกติ ส่วนเด็กอ่อนจะมีแนวโน้มที่จะได้คะแนนต่ำกว่าเดิม เพราะแม้แต่ค้นหาคำตอบจนไม่สามารถจะตอบได้ทันตามกำหนดเวลา

13. ไม่ควรออกข้อสอบให้เลือกตอบเพียงบางข้อ เพราะนักเรียนอาจจะเลือกตอบข้อต่างกัน ซึ่งจะทำให้เกิดความแตกต่างในตัวข้อสอบเอง คะแนนของแต่ละคนจึงขึ้นอยู่กับข้อที่เลือกทำ

ล้วน สายยศ และอังคณา สายยศ (2539 : 86 – 88) ได้เสนอหลักการสร้างข้อคำถามแบบความเรียงไว้ 5 ข้อ ดังนี้

1. ดูจุดประสงค์ของการสอบก่อน แล้วจึงเขียนข้อคำถาม
2. ควรใช้คำถามที่มีความกระชับ ชัด ด้วยหลักการถามและหลักภาษา ทั้งนี้เพื่อไม่ให้ผู้สอบเข้าใจผิด หรือมัวตีความหมายของคำถามที่เป็นไปได้หลายแง่หลายมุม
3. คำถามหนึ่ง ๆ ควรเป็นเรื่องเดียว ทั้งนี้เพื่อให้ผู้ตอบสามารถตอบตรงเป้าหมายที่ผู้ถามต้องการ
4. คำถามควรคำนึงถึงเวลาที่จะให้ผู้ตอบทำการตอบ
5. คำถามทุกคำถามผู้สอบควรทำเฉลยไว้ และวางแผนการให้คะแนนแต่ละส่วนว่าเป็นเท่าไรในเวลาที่จำกัดไว้ คำตอบที่สมบูรณ์ที่สุดควรเป็นอย่างไร เพื่อใช้เปรียบเทียบ

1.4 องค์ประกอบที่ส่งผลต่อคะแนนในแบบทดสอบความเรียง

เวอร์ต (Ward. 1999 : 162 – 164) ได้แบ่งปัญหาในการให้คะแนนโดยผู้ตรวจออกเป็น 2 ด้าน คือ ด้านความคลาดเคลื่อนทางเนื้อหา และด้านความคลาดเคลื่อนจากผู้ตรวจ ความคลาดเคลื่อนนี้จะต้องทำให้เกิดขึ้นน้อยที่สุด วิธีที่ช่วยได้ก็คือ การเขียนกำหนดเกณฑ์การตรวจให้ชัดเจนที่สุด

ความคลาดเคลื่อนทางเนื้อหา แบ่งเป็น 2 ชนิด คือ

1. ความคลาดเคลื่อนจากความใจดี (Generosity Error) เป็นผลที่ผู้ตรวจจะให้คะแนนสูงซึ่งจะพบในหลาย ๆ สถานการณ์ เช่น โดยทั่วไปครูมากกว่า 90% มักจะให้คะแนนลูกศิษย์ของตนเองสูง

2. ความคลาดเคลื่อนจากความลำเอียง (Halo Effect) เป็นพฤติกรรมส่วนบุคคลและตัวแปรที่เป็นความประทับใจส่วนตัวต่อนักเรียน หรือต่อลายมือที่สวยงาม จะเป็นผลให้คะแนนที่ได้สูงกว่า เช่น ซาร่าเป็นนักเรียนที่ดี (หรือที่น่ารัก , เป็นคนใจดี) ก็จะทำให้คะแนนของเธอสูงขึ้นในทุก ๆ หัวข้อ ทั้ง ๆ ที่เธออาจจะไม่มีความสามารถในด้านนั้น ๆ ก็ได้

ความคลาดเคลื่อนจากผู้ตรวจ มี 2 ลักษณะ คือ

1. ความคลาดเคลื่อนภายในผู้ตรวจ (Intrajudge Errors) เป็นความคลาดเคลื่อนที่เกิดจากการใช้มาตรฐานของผู้ตรวจเองที่อาจเกิดจากความเหนื่อยล้า ความเบื่อ หรือความคุ้นเคยจากตัวกระดาษคำตอบขณะทำการตรวจให้คะแนน ผู้ตรวจต่างคนก็จะมีมาตรฐานต่างกัน

2. ความคลาดเคลื่อนจากภายนอกผู้ตรวจ (Interjudge Errors) เป็นความคลาดเคลื่อนจากความแตกต่างระหว่างผู้ตรวจต่างคนกัน เช่น ผู้ตรวจคนที่ 1 อาจจะให้คะแนนสูงกับนักเรียนทุกคน และอีกคนอาจจะให้คะแนนต่ำกับนักเรียนทุกคน

บุญเชิด ภิญโญนนตพงษ์ (2547 : 63 – 64) กล่าวว่า การตรวจให้คะแนนของแบบทดสอบความเรียง มีสิ่งที่ทำให้เกิดความคลาดเคลื่อนไปจากความเป็นจริง ซึ่งพอสรุปได้ 4 ประการ คือ

1. ผลมาจากสิ่งอื่น ความคลาดเคลื่อนของการให้คะแนนอาจเป็นผลมาจากสิ่งอื่น เช่น เมื่อผู้ตรวจอ่านคำตอบของคำถามข้อแรกซึ่งนักเรียนผู้นี้ตอบได้ดีมาก และได้คะแนนสูงสุด แต่ในข้อถัด ๆ มา คำตอบของนักเรียนคนนี้ไม่ดีเท่าที่ควร คะแนนที่ได้รับก็ยังคงได้คะแนนสูงอยู่ก็ ทั้งนี้ เพราะอิทธิพลของข้อแรกส่งผลมาให้คะแนนข้อถัดมาคลาดเคลื่อนจากที่ควรจะได้จริงไป นอกจากนั้น ความรู้จำมักคุ้นก็มีผลต่อการตรวจให้คะแนน

2. ความใจดี มีครูจำนวนมากที่มักจะให้คะแนนสูงกว่าที่ควรจะเป็นเสมอ หรือเรียกว่า

มักจะให้คะแนนเพื่ออยู่เสมอ เช่น จะให้เกรด 4 และ 3 ค่อนข้างหมด ทั้ง ๆ ที่คำตอบของนักเรียนจะอยู่ในระดับปานกลางก็ตาม ความคลาดเคลื่อนชนิดนี้เป็นเพราะความใจดีของครู เมื่อเห็นคำตอบนั้นถูกบ้างก็จะให้คะแนนสูงมากทั้ง ๆ ที่ความเป็นจริงแล้วถ้าให้ครูคนอื่น ๆ ตรวจคำตอบนั้นจะอยู่ในระดับปานกลางเท่านั้น ความคลาดเคลื่อนชนิดนี้มีลักษณะคล้ายกับชนิดแรกนั่นเอง

3. การปรับโทษ การให้คะแนนลดหรือการตัดทอนคะแนนของนักเรียนซึ่งไม่สามารถทำคำตอบได้ตามที่คาดหวังไว้ ถือว่าเป็นการปรับโทษ ซึ่งเป็นสิ่งที่ไม่ยุติธรรมสำหรับนักเรียน ทั้งนี้เพราะความสามารถที่แท้จริงของนักเรียนอาจจะอยู่ในระดับที่ใช้ได้ แต่ทว่าครูได้ตั้งมาตรฐานและคาดหวังความสามารถของนักเรียนไว้สูงจนเกินไป เมื่อนักเรียนทำไม่ได้ดังที่ตนตั้งไว้ จึงทำให้เกิดความขัดแย้งว่านักเรียนมีความสามารถต่ำจึงให้คะแนนน้อย

4. องค์ประกอบแทรกซ้อน อิทธิพลที่มีผลต่อการให้คะแนนข้อสอบอัตนัยอาจจะเกิดมาจากองค์ประกอบแทรกซ้อนอื่น ๆ ซึ่งเป็นการยากต่อการควบคุม ทั้งนี้เพราะเป็นผลมาจากตัวนักเรียนเอง เช่น ลายมือ ตัวสะกดการันต์ หลักไวยากรณ์ การเรียบเรียงถ้อยคำ สิ่งเหล่านี้ย่อมมีอิทธิพลต่อการให้คะแนนอยู่เสมอ องค์ประกอบเหล่านี้ครูจำเป็นต้องแจ้งให้นักเรียนทราบว่าการตรวจให้คะแนนจะยึดถือหลักเกณฑ์ดังกล่าวด้วย เพื่อให้ นักเรียนพยายามและระมัดระวังในการตอบมากขึ้น

2. กฎเกณฑ์การให้คะแนน (Scoring Rubric)

2.1 ความหมายของกฎเกณฑ์การให้คะแนน

สเตนมาร์ค (Stenmark. 1991 : 24) ได้อธิบายถึงกฎเกณฑ์การให้คะแนนที่ใช้สำหรับคำถามแบบปลายเปิด (Open – ended Questions) ซึ่งปรับปรุงมาจาก California Assessment Program Grade 12 Generalized Rubric ระบุว่า กฎเกณฑ์การให้คะแนนเฉพาะ (Specific Rubric) เป็นแนวทางในการให้คะแนนเฉพาะของคำถามแบบปลายเปิดในแต่ละข้อ สำหรับกฎเกณฑ์การให้คะแนนทั่วไป (General Rubric) จะเป็นตัวแบบของกฎเกณฑ์การให้คะแนนเพื่อใช้พิจารณา กฎเกณฑ์เฉพาะ รายละเอียดจะเปลี่ยนแปลงสำหรับระดับคะแนนที่แตกต่างกัน แนวคิดพื้นฐานสามารถประยุกต์ใช้ได้

คอทนี (Courtney. 1996 : 77 – 79) อธิบายว่า เป็นคำแนะนำที่เกี่ยวกับการให้คะแนน ถ้าระบบการให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนที่มี 5 ระดับ นักเรียนจะได้ 5 คะแนนเต็มก็ต่อเมื่อสามารถทำได้ครบทุกอย่างที่ระบุไว้ กฎเกณฑ์การให้คะแนนเน้นการแสดงความสามารถของนักเรียนในด้านความเข้าใจในความคิดรวบยอด ความรู้เกี่ยวกับกระบวนการหรือขั้นตอน ทักษะในการแก้ปัญหา การใช้เทคโนโลยีใหม่ ๆ ในการประมวลผลข้อมูล และการสื่อความหมายกับผู้อื่น

เสาวนีย์ เกรียร์ (2540 : 159) กล่าวว่า เป็นเครื่องมือในการให้คะแนนที่ประกอบด้วย ประเด็นต่าง ๆ (Criteria) ที่จะใช้พิจารณางานหนึ่ง ๆ และคำอธิบายระดับคุณภาพของแต่ละประเด็น ประเมิน ซึ่งอาจเรียงลำดับตั้งแต่ดีเลิศไปจนถึงต้องปรับปรุง หรือให้เป็นระดับตัวเลขตั้งแต่มากที่สุด (เช่น 4) ไปจนถึงน้อยที่สุด (เช่น 0) ประเด็นประเมินอาจกำหนดเพิ่มเติมได้หลายข้อ คำอธิบายระดับคุณภาพควรอธิบายให้ชัดเจนที่สุด กระชับที่สุด เป็นคำอธิบายที่สามารถบอกได้ว่าทำไมจึงต้อง ดีเลิศ ดี ต้องปรับปรุง

บุญเชิด ภิญญอนันตพงษ์ (2544 : 90) ให้ความหมายของกฎเกณฑ์การให้คะแนนว่าเป็นชุดของแนวทางในการให้คะแนนผลการปฏิบัติเรื่องใดเรื่องหนึ่ง สำหรับใช้ประเมินคุณภาพการปฏิบัติงานของผู้เรียน แนวทางในการให้คะแนนนั้น อาจทำในรูปของมาตราประเมินค่า หรือแบบตรวจสอบรายการ

จากผลการศึกษาดังกล่าว ผู้วิจัยให้ความหมายของกฎเกณฑ์การให้คะแนนว่าหมายถึงแนวทางที่กำหนดขึ้นเพื่อชี้บอกระดับของพฤติกรรม คุณภาพของกระบวนการทำงานและผลลัพธ์เพื่อสิ่งที่ต้องการ ซึ่งจะต้องมีการกำหนดเป็นมาตรวัด และรายการคุณลักษณะที่บรรยายถึงความสามารถในการแสดงออก ของแต่ละจุดในมาตรวัดอย่างชัดเจน ซึ่งจะทำให้ผู้ตรวจให้คะแนนมีความเข้าใจตรงกัน และกฎเกณฑ์การให้คะแนนจะช่วยส่งเสริมการเรียนรู้ของนักเรียน ทำให้เป้าหมายการแสดงผลงานของนักเรียนมีความชัดเจนขึ้น ทำให้นำไปสู่การบรรลุจุดประสงค์หรือสมรรถภาพที่สำคัญของมาตรฐานการศึกษาได้

2.2 ประเภทของกฎเกณฑ์การให้คะแนน

ชัยฤทธิ์ ศิลาเดช (2540 : 29) กล่าวว่า การให้คะแนนของรูปรีด เป็นการตอบคำถามว่านักเรียนทำอะไรได้สำเร็จ มีระดับความสำเร็จในขั้นต่าง ๆ กัน หรือมีผลงานอย่างไร และได้แบ่งการให้คะแนนแบบรูปรีดไว้ 2 แบบ คือ

1. แบบรวมองค์ประกอบ (Holistic Rubrics) คือการให้คะแนนงานชิ้นใดชิ้นหนึ่ง โดยดูภาพรวมของชิ้นงานว่ามีความเข้าใจในความคิดรวบยอด การสื่อความหมาย กระบวนการที่ใช้ และผลงานเป็นอย่างไร แล้วเขียนอธิบายคุณภาพของงาน หรือความสำเร็จของงานเป็นขั้น ๆ โดยอาจจะแบ่งระดับคุณภาพตั้งแต่ 0 – 4 หรือ 0 – 6

2. แบบแยกองค์ประกอบ (Analytic Rubrics) เพื่อให้การมองคุณภาพของงานหรือความสามารถของนักเรียนได้อย่างชัดเจน จึงได้มีการแยกองค์ประกอบของการให้คะแนน และอธิบายคุณภาพของงานในแต่ละองค์ประกอบเป็นระดับโดยทั่วไปแล้วจะมีการแยกองค์ประกอบของงานเป็น 4 ด้าน คือ

2.1 ความเข้าใจในความคิดรวบยอด ข้อเท็จจริง เป็นการแสดงให้เห็นว่านักเรียน

เข้าใจในความคิดรวบยอด หลักการในการแก้ปัญหาที่ถามอย่างกระจ่างชัด

2.2 การสื่อความหมาย สื่อสาร คือความสามารถในการอธิบาย นำเสนอ การบรรยายเหตุผล แนวคิดให้ผู้อื่นเข้าใจได้ดี มีความคิดสร้างสรรค์

2.3 การใช้กระบวนการและยุทธวิธี สามารถเลือกยุทธวิธี กระบวนการนำไปสู่ความสำเร็จอย่างมีประสิทธิภาพ

2.4 ผลสำเร็จของงาน ความถูกต้องแม่นยำในผลสำเร็จของงาน หรืออธิบายที่มา และตรวจสอบผลงาน

บุญเชิด ภิญโญนันตพงษ์ (2544 : 90) กล่าวว่า การพิจารณากฎเกณฑ์การให้คะแนน ถ้าพิจารณาตามลักษณะของงานหรือกิจกรรม แบ่งออกเป็น 2 รูปแบบ คือ กฎเกณฑ์การให้คะแนนแบบทั่วไป (General Rubric) สำหรับใช้ประเมินงานหรือกิจกรรมหลาย ๆ ชั้นที่อยู่ในกรอบเนื้อเรื่องเดียวกัน และกฎเกณฑ์การให้คะแนนแบบเฉพาะ (Specific Rubric) เป็นแนวทางในการให้คะแนนทั่วไป ที่เขียนให้เฉพาะเจาะจงกับการปฏิบัติงานหรือกิจกรรมแต่ละชิ้น ถ้าพิจารณาตามเกณฑ์หรือองค์ประกอบของการให้คะแนน ก็ยังแบ่งออกได้เป็น 2 ลักษณะ คือ เกณฑ์รวม (Holistic) เป็นการพิจารณาผลงานของผู้เรียนในภาพรวมว่ามีคุณภาพในระดับใด และเกณฑ์ย่อย (Analytic) เป็นการพิจารณาผลงานของผู้เรียนแยกเป็นรายด้าน

สำหรับกฎเกณฑ์การให้คะแนนที่ใช้ในการตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ผู้วิจัยได้เลือกใช้รูปแบบเฉพาะ (Specific Rubric) กำหนดระดับคุณภาพเป็น 4 ระดับ ในแต่ละทักษะ

2.3 การสร้างกฎเกณฑ์การให้คะแนน

การสร้างกฎเกณฑ์การให้คะแนนมีขั้นตอนในการดำเนินการ 7 ขั้นตอน (โครงการพัฒนาทรัพยากรมนุษย์. ม.ป.ป.) ดังนี้

ขั้นที่ 1 วิเคราะห์ผลการเรียนรู้ของนักเรียน

วิเคราะห์ผลการเรียนรู้ของนักเรียนในแต่ละจุดประสงค์ หรือแต่ละหัวเรื่อง (Theme) เมื่อนักเรียนได้ปฏิบัติกิจกรรมการเรียนรู้แล้วจะเกิดผลการเรียนรู้อะไรบ้าง ผลการเรียนรู้ของนักเรียนไม่จำเป็นจะต้องมีครบทุกประเภท ทุกครั้งที่นักเรียนได้ปฏิบัติกิจกรรมการเรียนรู้ บางครั้งในการวางแผนการสอน ครูคาดหวังในผลการเรียนรู้ของนักเรียนเพียงกระบวนการอย่างเดียว บางครั้งครูคาดหวังทั้งกระบวนการและผลงาน

ขั้นที่ 2 กำหนดประเด็นที่จะต้องประเมิน

การกำหนดประเด็นที่ต้องการประเมินอาจให้นักเรียน ร่วมกันเสนอความคิด ในการกำหนดประเด็นประเมินผลการเรียนรู้แต่ละอย่างของเขาเอง แต่อย่างไรก็ตามเนื่องจากในการวางแผน

แผนการสอน ครูผู้สอนได้มีการคาดหวังความสำเร็จจากการเรียนรู้ของนักเรียนเอาไว้ล่วงหน้าแล้ว จึงควรกำหนดรายการประเมินที่สำคัญ ๆ ของความสำเร็จจากการเรียนแต่ละด้านเอาไว้ล่วงหน้าด้วย เพื่อจะได้แจ้งให้นักเรียนทราบและเพื่อการตรวจสอบผลงานของนักเรียนเอง

ขั้นที่ 3 คัดเลือกประเด็นประเมินผลที่สำคัญ

การกำหนดประเด็นที่จะต้องประเมิน เราจะพบว่า ในผลการเรียนรู้หนึ่งอย่าง จะมีประเด็นที่ต้องการประเมินมาก โดยเฉพาะถ้าเปิดโอกาสให้นักเรียนมีส่วนร่วมในการเสนอประเด็นประเมินด้วยแล้ว จะมีมุมมองในการประเมินที่หลากหลาย หรือรวมประเด็นที่สำคัญ ๆ

ขั้นที่ 4 เลือกรูปแบบในการสร้างเกณฑ์

การสร้างเกณฑ์การประเมินผลการเรียนรู้ของนักเรียน เพื่อพิจารณาตัดสินให้คะแนน (Scoring Rubric) มีรูปแบบในการสร้าง 2 รูปแบบ

รูปแบบที่ 1 การให้คะแนนแบบรวมองค์ประกอบ (Holistic Scoring Rubric)

การสร้างเกณฑ์การประเมินแบบรวมองค์ประกอบ หมายถึงการให้คะแนนผลการเรียนรู้ โดยรวมทุกประเด็นที่กำหนดเพื่อการประเมิน แล้วเขียนอธิบายคุณภาพของผลการเรียนรู้แต่ละระดับ

ตาราง 3 เกณฑ์การให้คะแนนแบบรวมองค์ประกอบ

ระดับคุณภาพ	คำอธิบาย
ระดับ 3	หมายถึง ตอบคำถามจากเรื่องที่อ่านถูกทุกข้อ บอกเนื้อหาสาระถูกต้องได้ใจความ ต่อเนื่อง บอกข้อคิดได้ตรงประเด็นสมบูรณ์ เสนอความคิดเห็นด้วยเหตุผลและประโยชน์
ระดับ 2	หมายถึง ตอบคำถามจากเรื่องที่อ่านผิดไม่เกิน 3 ข้อ จาก 5 ข้อ บอกเนื้อหาสาระได้ถูกต้องแต่รวน บอกข้อคิดได้ตรงประเด็นแต่ไม่ต่อเนื่อง เสนอความคิดเห็นด้วยเหตุผล
ระดับ 1	หมายถึง ตอบคำถามจากเรื่องที่อ่านผิดมากกว่า 3 ข้อ ใน 5 ข้อ บอกเนื้อหาสาระได้บ้าง บอกข้อคิดได้บ้าง แต่รวน เสนอความคิดเห็นแต่ไม่แสดงเหตุผล

รูปแบบที่ 2 การให้คะแนนแบบแยกองค์ประกอบ (Analytic Scoring Rubric)

การให้คะแนนแบบแยกองค์ประกอบ หมายถึง การให้คะแนนโดยการแยกองค์ประกอบของสิ่งที่จะประเมิน เพื่อให้มองเห็นคุณภาพของงาน หรือความสามารถของนักเรียนได้อย่างชัดเจน ผลการประเมินจะบ่งบอกถึงจุดเด่น จุดด้อยของแต่ละประเด็นประเมินได้ชัดเจน การสร้างเกณฑ์การ

ประเมินในรูปแบบนี้จะต้องเขียนคำอธิบายคุณภาพของงานในแต่ละองค์ประกอบ และแต่ละระดับขององค์ประกอบให้ชัดเจน

ตาราง 4 เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

ประเด็นประเมิน	คำอธิบายระดับคุณภาพ		
	3 (ดี)	2 (พอใช้)	1 (ควรปรับปรุง)
1. การตอบคำถามจากเรื่องที่อ่าน	ตอบคำถามถูกต้องทุกข้อ	ตอบคำถามผิดไม่เกิน 3 ข้อ จาก 5 ข้อ	ตอบคำถามผิดมากกว่า 3 ข้อ จาก 5 ข้อ
2. การบอกความสำคัญของเรื่องที่อ่าน	บอกเนื้อหาสาระถูกต้องได้ในความต่อเนื่อง	บอกเนื้อหาสาระถูกต้อง แต่วกวน	บอกเนื้อหาสาระได้บ้าง
3. การบอกข้อคิดจากเรื่องที่อ่าน	บอกข้อคิดได้ตรงประเด็น สมบูรณ์	บอกข้อคิดได้ตรงประเด็น แต่ไม่ต่อเนื่อง	บอกข้อคิดได้บ้างแต่วกวน
4. การเสนอความคิดเห็นจากเรื่องที่อ่าน	เสนอความคิดเห็นด้วยเหตุผลและเป็นประโยชน์	เสนอความคิดเห็นด้วยเหตุผล แต่มีประโยชน์น้อย	เสนอความคิดเห็นแต่ไม่แสดงเหตุผล

ขั้นที่ 5 กำหนดค่าระดับคุณภาพในการประเมิน

ค่าระดับคุณภาพ คือ ตัวเลขที่บ่งบอกถึงคะแนนการประเมินผลการเรียนรู้ของนักเรียน ซึ่งผู้สอนเป็นผู้กำหนด หรืออาจจะให้นักเรียนมีส่วนร่วมในการกำหนดค่าระดับคุณภาพก็ได้ ตัวเลขค่าระดับคุณภาพอาจจะมีระดับ 0-1-2-3 หรือ 0-3-4-5-6 หรือ 0-10-15-20 ในกรณีที่กำหนดค่าระดับต่ำสุดที่เลข 0 นั้น หมายถึงนักเรียนไม่มีผลการเรียนรู้ หรือไม่มีผลงานเข้ารับการประเมิน หรือไม่ได้ดำเนินการตามรายการประเมินนั้น และการกำหนดค่าระดับจาก 0-10 แสดงว่าผู้กำหนดค่าระดับคุณภาพ พิจารณาแล้วเห็นว่าคุณภาพของงานหรือสิ่งที่จะประเมินควรมีคะแนนสูงสุดคือ 10

ขั้นที่ 6 บรรยายคุณภาพการประเมินแต่ละระดับ

เมื่อได้เลือกรูปแบบในการกำหนดเกณฑ์การให้คะแนน และกำหนดตัวเลขค่าระดับแล้ว จะต้องเขียนคำอธิบายขอบข่ายการพิจารณาตัดสินให้คะแนนแต่ละระดับคุณภาพให้ชัดเจน ผลงาน

หรือกระบวนการในลักษณะไหน อย่างไร จึงจะมีคุณภาพระดับ 4 ลักษณะไหนอย่างไรจึงมีคุณภาพระดับ 3 ระดับ 2 และระดับ 1 คำบรรยายคุณภาพแต่ละระดับจะเกิดประโยชน์สูงสุด

ตาราง 5 อธิบายคุณภาพ “การพูดอภิปราย”

ระดับคุณภาพ	คำอธิบาย
4 ดีมาก	หมายถึง มีความคิดริเริ่มสร้างสรรค์ มีการนำเสนอข้อเท็จจริงและแสดงเหตุผลอย่างเหมาะสม มีการยกตัวอย่างเพื่อสนับสนุนประเด็นอภิปราย การออกเสียงถูกต้องชัดเจน มีมารยาทในการพูด และรักษาเวลาได้
3 ดี	หมายถึง มีความคิดริเริ่มสร้างสรรค์ มีการนำเสนอข้อเท็จจริงและแสดงเหตุผลอย่างเหมาะสม แต่ขาดตัวอย่างเพื่อสนับสนุนประเด็นอภิปราย ส่วนการออกเสียงถูกต้องชัดเจน มีมารยาทในการพูด และรักษาเวลาได้ดี
2 ใช้ได้	หมายถึง ไม่มีความคิดริเริ่มสร้างสรรค์ มีการนำเสนอข้อเท็จจริงและข้อคิดเห็น แต่ขาดตัวอย่างเพื่อสนับสนุนประเด็นอภิปราย มีการออกเสียงถูกต้องชัดเจน แต่ไม่รักษาเวลาและมารยาทในการพูด
1 ใช้ไม่ได้	หมายถึง มีการนำเสนอข้อเท็จจริงและข้อคิดเห็น แต่ขาดตัวอย่างเพื่อสนับสนุนประเด็นอภิปราย การออกเสียงไม่ถูกต้องชัดเจน ขาดการรักษาเวลาและมารยาทในการพูด
0	หมายถึง ไม่มีการอภิปราย

ขั้นที่ 7 กำหนดคะแนนการตัดสินระดับคุณภาพ

การกำหนดคะแนนการตัดสินระดับคุณภาพเป็นการกำหนดช่วงคะแนนจากการประเมินผลการเรียนรู้

ตาราง 6 การกำหนดคะแนนการตัดสินระดับคุณภาพ

คะแนน	ระดับคุณภาพ
9 – 12	ดี
5 – 8	พอใช้
1 – 4	ควรปรับปรุง

บุญเชิด ภิญญอนันตพงษ์ (2544 : 94) กล่าวว่า การสร้างกฎเกณฑ์การให้คะแนนการประเมินผลภาคปฏิบัติต้องมีส่วนประกอบในการดำเนินการที่สำคัญ 3 ขั้นตอน คือ

1. กำหนดการปฏิบัติให้ชัดเจน

- 1.1 ระบุชนิดของการปฏิบัติว่าเป็นทักษะวิธีการทำหรือผลงานสำเร็จที่สร้างขึ้น
- 1.2 กำหนดจุดสนใจของการประเมินว่าสนใจเป็นรายบุคคลหรือรายกลุ่ม
- 1.3 กำหนดกฎเกณฑ์การปฏิบัติที่สะท้อนคุณลักษณะที่สำคัญของการปฏิบัติที่ประสบผลสำเร็จ

2. ออกแบบกิจกรรมสำหรับใช้ประเมิน

- 2.1 ระบุลักษณะของงานหรือกิจกรรมว่าเป็นผลงานหรือการปฏิบัติตามที่มอบหมายให้ทำ หรือเป็นการแสดงออกที่เกิดขึ้นเองตามธรรมชาติ
- 2.2 กำหนดรายการของงานหรือกิจกรรม ได้แก่ นิยามจุดหมายของการปฏิบัติ เงื่อนไขสถานการณ์ของการปฏิบัติ และเกณฑ์มาตรฐาน
- 2.3 กำหนดจำนวนชิ้นงานหรือกิจกรรมให้ทำว่าจะใช้เพียงงานเดียวหรือหลายงาน จึงจะเพียงพอในสภาพการณ์นั้น

3. ให้คะแนนและจัดบันทึกผล ดังนี้

- 3.1 ระบุรูปแบบการให้คะแนนว่าใช้เกณฑ์ในการให้คะแนนแบบใด เช่น แบบทั่วไป หรือแบบเฉพาะเนื้อหา และแบบเกณฑ์รวมหรือแบบเกณฑ์ย่อย
 - 3.2 กำหนดวิธีการจัดบันทึกว่าเป็นแบบตรวจสอบรายการ มาตรฐานประเมินค่าการปฏิบัติ การจัดบันทึกพฤติกรรมตามช่วงเวลา หรือการจัดจำธรรมชาติ
 - 3.3 ระบุผู้ประเมินว่าจะให้ใครเป็นผู้สังเกตและประเมิน ผลการปฏิบัติจะเป็นประโยชน์แก่ครูผู้สอน ครูที่เคยสอน ตัวผู้เรียนเอง หรือเพื่อน ๆ ในกลุ่ม
- การวิจัยในครั้งนี้ ผู้วิจัยได้ใช้วิธีผสมผสานจากหลายแนวทางที่กล่าวมาข้างต้น มาสร้างเป็นกฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ โดยเลือกใช้รูปแบบเกณฑ์เฉพาะ (Specific Rubric) จำนวน 5 ทักษะ ตามทักษะย่อยของทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

3. ทฤษฎีค่าความเชื่อมั่น (Reliability Theory)

3.1 การประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory)

ทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory) อธิบายความแม่นยำในการวัดโดยใช้ค่าความเชื่อมั่น ยึดข้อตกลงที่เป็นคุณสมบัติคู่ขนานหรือความเท่าเทียมเป็นสำคัญ ซึ่งเป็นไปตามสมมติฐานของสเปียร์แมน (Spearman, 1910) ที่ได้กล่าวถึงข้อตกลงเกี่ยวกับแบบทดสอบคู่ขนานเป็นครั้งแรกว่า เป็นแบบทดสอบที่วัดคุณลักษณะเดียวกัน

คะแนนที่ได้จากการสอบ (X) ประกอบด้วยคะแนนจริง (T) รวมกับคะแนนความคลาดเคลื่อน (E) จากข้อตกลงของสเปียร์แมน สามารถเขียนสมการได้ดังนี้

$$X = T + E$$

เมื่อ X แทน คะแนนสังเกต หรือคะแนนที่ได้จากการสอบ (Observed Scores)

T แทน คะแนนจริง (True Score)

E แทน คะแนนความคลาดเคลื่อน (Error Score)

คะแนนจริง (True Score) หมายถึง คะแนนที่ผู้เข้าสอบได้รับจากการทดสอบด้วยเครื่องมือวัดใด ๆ หรือหมายถึงคะแนนเฉลี่ยของผู้เข้าสอบทำได้จากการสอบหลาย ๆ ครั้งโดยมีข้อตกลงว่า ต้องไม่มีอิทธิพลจากการฝึกฝน ความเมื่อยล้า และการเรียนรู้ในแบบทดสอบซ้ำ คะแนนความคลาดเคลื่อน (Error Score) หมายถึง คะแนนที่ผู้เข้าสอบได้หรือไม่ได้เนื่องจากการผิดพลาดที่เกิดจากการวัด ซึ่งมี 2 ชนิด คือ ความคลาดเคลื่อนอย่างสุ่ม (Random Error) เป็นความคลาดเคลื่อนที่เกิดขึ้นโดยบังเอิญ เช่น เด็กบางคนป่วยไม่สามารถทำข้อสอบได้เต็มที่ ทำให้ได้คะแนนลดลง ถ้าความคลาดเคลื่อนอย่างสุ่มเกิดขึ้นในการวัดแล้ว ไม่มีความสัมพันธ์กับคะแนนจริง นั่นคือสหสัมพันธ์ระหว่างคะแนนจริงกับคะแนนความคลาดเคลื่อนมีค่าเป็น 0 ($r_{TE} = 0$) (Mehrens and Lehmann, 1973 : 91) อีกชนิดหนึ่งเรียกว่าความคลาดเคลื่อนที่เป็นระบบ (Systematic Error) เป็นความคลาดเคลื่อนที่เกิดขึ้นกับนักเรียนทุก ๆ คน เช่น แบบทดสอบที่ใช้พิมพ์โจทย์ตกไป 2 ข้อทุกฉบับ ซึ่งไม่มีผลกระทบต่อค่าความเชื่อมั่น เพราะคะแนนที่คลาดเคลื่อนเกิดขึ้นกับตัวนักเรียนแต่ละคนเหมือนกัน (บุญเชิด ภิญญอนันตพงษ์, 2524 : 270 – 271)

นอกจากนั้นมีข้อตกลงเพิ่มเติมเกี่ยวกับคะแนนความคลาดเคลื่อนว่า มีค่าเฉลี่ยเป็น 0 ($\sum E = 0$) มีความแปรปรวนเท่ากัน เป็นอิสระต่อกัน และเป็นอิสระต่อคะแนนจริง ภายใต้เงื่อนไขนี้แบบทดสอบแต่ละฉบับมีค่าเฉลี่ยเท่ากัน ความแปรปรวนเท่ากัน และค่าสหสัมพันธ์ระหว่างแบบทดสอบคู่ขนานดังกล่าวมีค่าเท่ากัน เท่ากับอัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับ

ความแปรปรวนของคะแนนสังเกต และมีค่าเท่ากับกำลังสองของค่าสหสัมพันธ์ระหว่างคะแนนจริงกับคะแนนสังเกต

ดังนั้นความแปรปรวนของคะแนนสังเกต (σ_X^2) จึงเป็นผลรวมของความแปรปรวนของคะแนนจริง (σ_T^2) กับความแปรปรวนของคะแนนความคลาดเคลื่อน (σ_E^2) เขียนเป็นสมการได้ดังนี้

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

สำหรับสัมประสิทธิ์ความเชื่อมั่น ก็คือ อัตราส่วนระหว่างความแปรปรวนของคะแนนจริง (σ_T^2) กับความแปรปรวนของคะแนนสังเกต (σ_X^2) เขียนเป็นสมการได้ดังนี้

$$\rho_{xx} = \rho_{XT}^2 = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2}$$

ในระยะเวลาที่ใกล้เคียงกันนั้นบราวน์ (Brown) ได้พัฒนาทฤษฎีความเชื่อมั่นโดยเริ่มด้วยการนิยามแบบทดสอบคู่ขนาน ขณะที่สเปียร์แมนเริ่มศึกษาจากคะแนนสังเกต ประกอบด้วยคะแนนจริงรวมกับคะแนนความคลาดเคลื่อน และจากแนวคิดเกี่ยวกับความเชื่อมั่นของบราวน์ไปสอดคล้องกับทฤษฎีของสเปียร์แมน (ไพรัตน์ วงษ์นาม. 2533 : 18) จากข้อตกลงเกี่ยวกับคุณสมบัติคู่ขนาน นำไปสู่สูตรการพัฒนาค่าความเชื่อมั่นหลายสูตร รวมทั้งสูตรความสอดคล้องภายใน (Internal Consistency) คือสูตรสเปียร์แมน - บราวน์ที่ประมาณค่าความเชื่อมั่นจากการสอบครั้งเดียว โดยคำนึงจากค่าสหสัมพันธ์ของแบบทดสอบที่แบ่งเป็น 2 ส่วน ซึ่งอาจแบ่งเป็นข้อคู่ - ข้อคี่ หรือ ครึ่งแรก - ครึ่งหลัง ค่าสหสัมพันธ์เมื่อปรับด้วยสูตรแล้ว ค่าที่ได้เป็นค่าประมาณของค่าสหสัมพันธ์ระหว่างแบบทดสอบฉบับนั้นกับแบบทดสอบฉบับอื่นที่มีความคู่ขนานกัน

สูตรหาความเชื่อมั่นแบบสอดคล้องภายในอีกสองสูตร คือ KR - 20 และ KR - 21 ใช้ค่าสถิติจากการสอบครั้งเดียว เป็นค่าประมาณสหสัมพันธ์ระหว่างแบบทดสอบคู่ขนานสองฉบับ ถ้าข้อสอบในแบบทดสอบฉบับนั้นวัดองค์ประกอบเดียว สูตร KR - 20 เป็นสูตรที่ให้ค่าสหสัมพันธ์เหมาะสม การใช้สูตร KR - 20 ต้องปรับเป็นสูตร KR - 21 ถ้าความแปรปรวนของข้อสอบแต่ละข้อมีค่าเท่ากัน สูตรทั้งสองนี้ใช้กับการให้คะแนนแบบ 0 หรือ 1 การให้คะแนนแบบนี้ข้อสอบทั้งฉบับจะวัดองค์ประกอบเดียวก็ต่อเมื่อข้อสอบมีค่าเฉลี่ยเท่ากัน และความแปรปรวนเท่ากัน หมายความว่าสูตรทั้งสองยึดข้อตกลงเบื้องต้นเกี่ยวกับคุณสมบัติคู่ขนาน (Cronbach and others. 1963 : 139 - 140)

โดยสรุปทฤษฎีความเชื่อมั่นแบบมาตรฐานเดิม พัฒนาจากข้อตกลงของความเท่าเทียมหรือความเป็นคู่ขนานของการวัด จากนั้นความเชื่อมั่นนี้จึงถือว่าเป็นค่าสัมประสิทธิ์การวัดที่มีความเท่าเทียมหรือมีความคู่ขนานกัน จากนั้นการหาค่าความเชื่อมั่นต้องมีการทดสอบเพื่อให้ได้ค่าจาก

การทดสอบอย่างน้อยสองชุด แต่ละชุดใช้เครื่องวัดที่คู่ขนานกัน คือมีเนื้อหาอย่างเดียวกัน ค่าเฉลี่ยเท่ากัน ความแปรปรวนเท่ากันและค่าสหสัมพันธ์เท่ากัน แล้วคำนวณค่าสหสัมพันธ์ระหว่างค่าที่ได้จากการทดสอบทั้งสองชุด ผลที่ได้เป็นค่าประมาณความเชื่อมั่น กรณีที่ต้องประมาณค่าความเชื่อมั่นจากการสอบครั้งเดียว ต้องแบ่งแบบทดสอบเป็นส่วนย่อย อาจแบ่งเป็นสองส่วน เช่น ข้อคู่ - ข้อคี่ หรือ ครึ่งแรก - ครึ่งหลัง หรือแบ่งย่อยตามจำนวนข้อ โดยยึดข้อตกลงส่วนย่อยของแต่ละส่วน ต้องมีความเท่าเทียมกัน

แต่ถ้าสถานการณ์ของแบบทดสอบแตกต่างกันไป เทคนิคการประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม ไม่สามารถนำมาประยุกต์ใช้ได้เช่น กรณีของการทดสอบที่มีผู้ตรวจให้คะแนนหลายคน ตรวจให้คะแนนข้อสอบหลายข้อ ของผู้เข้าสอบทุกคนและคะแนนผลการตอบแบบทดสอบที่มีค่าคะแนนต่างกันเป็นต้น เพราะรูปแบบการวัดของทฤษฎีการทดสอบแบบมาตรฐานเดิมถือว่าคะแนนความคลาดเคลื่อน หรือความคลาดเคลื่อนของการวัดเป็นคะแนนรวมความคลาดเคลื่อนทุกแห่งเข้าด้วยกัน และเป็นความคลาดเคลื่อนรวมอันเดียวกันที่ไม่สามารถระบุว่าเป็นอะไรได้ จึงเป็นข้อจำกัดทางทฤษฎีที่ไม่สามารถศึกษารายละเอียดของแหล่งความคลาดเคลื่อนจากการวัดต่าง ๆ ได้

3.2 การประมาณค่าความเชื่อมั่นที่เกี่ยวข้องกับการตรวจให้คะแนน (Reliability of Rating)

ฮอบกินส์ และแอนเตส (Hopkins and Antes. 1990 : 550) อธิบายถึงความเชื่อมั่นของผู้ตรวจให้คะแนนที่มีตั้งแต่สองคนขึ้นไป จะมีความเห็นสอดคล้องกันเกี่ยวกับจุดเน้น หรือส่วนสำคัญต่าง ๆ ของคำตอบที่ผู้ตรวจให้คะแนนตามความเห็นของผู้ตรวจ ความเชื่อมั่นของผู้ตรวจสูงหมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์สูงตามไปด้วย ผู้ตรวจให้คะแนนที่ดีควรเป็นผู้ที่สอนวิชานั้นจริง ๆ

เจมส์ , เดมารี และวอล์ฟ (James, Damaree and Wolf. 1993) ได้กล่าวว่า ความเชื่อมั่นที่เกี่ยวข้องกับการตรวจให้คะแนนที่ได้ถือเป็นดัชนีบ่งบอกความสอดคล้อง (Index of Agreement) ไม่ใช่ความคงที่ (Consistency) ได้พัฒนาสัมประสิทธิ์ความเชื่อมั่นระหว่างผู้ตรวจ (Inter rater reliability Coefficient , r_{wg}) ซึ่งจะประเมินความสอดคล้องระหว่างผู้ตรวจ โดยกำหนดให้ r_{wg} เป็นการลดสัดส่วนความแปรปรวนที่ผิดพลาดสำหรับการให้คะแนนที่ได้มาเพื่อเทียบกับการจัดแบ่งที่แสดงถึงการตอบแบบสุ่ม ตัวอย่างเช่น การแบ่งในกรณีที่แต่ละคะแนนในเกณฑ์ของอันดับคะแนนมีความเป็นไปได้เท่า ๆ กันในการนำไปใช้ กล่าวคือ

$$r_{wg} = 1 - \left(\frac{S_X^2}{\sigma_E^2} \right)$$

โดยที่ S_X^2 เท่ากับความแปรปรวนของผลการตรวจให้คะแนน และ σ_E^2 เท่ากับความแปรปรวนของประชากรในการจัดแบ่งอันดับคะแนนที่ได้จากการสุ่ม เช่น การตรวจให้คะแนนที่เท่ากันของคำตอบที่เป็นไปได้ในแต่ละประเด็น ค่าของ r_{wg} มีค่าตั้งแต่ 0.0 จนถึง 1.0 เมื่อความแปรปรวนของคะแนนที่ได้มาเป็นแบบสุ่ม (เช่น $S_X^2 = \sigma_E^2$) ค่า $r_{wg} = 0$ ซึ่งให้เห็นว่า ไม่มีความสอดคล้องกันเลย หรือหากผลการตรวจให้คะแนนเหมือนกัน (เช่น $S_X^2 = 0$) ค่า $r_{wg} = 1$ ซึ่งให้เห็นว่ามีความสอดคล้องโดยสมบูรณ์ระหว่างผู้ตรวจทั้งหมด

ค่า r_{wg} ที่ได้ควรจะเป็นดัชนีของความเห็นพ้อง (Index of Agreement) ไม่ใช่ค่าความสอดคล้อง (Consistency) เจมส์ (James. 1993) อธิบายว่า การให้คะแนนบางประเภทอาจมีระดับความสอดคล้องสูง เช่น ผู้ตรวจ 2 คน หรือมากกว่านั้นให้คะแนนออกมาคล้ายคลึงกัน เมื่อให้คะแนนตัวอย่างหลาย ๆ ตัวอย่างบนมาตรฐานเดียวกัน แต่ค่าการให้คะแนนที่แท้จริงที่ให้ผู้ตรวจคนเดียวอาจจะต่างกันมาก ให้พิจารณาถึงกรณีที่ผู้ตรวจคนที่ 1 ให้คะแนน 4 ทุกครั้ง ผู้ตรวจคนที่ 2 ให้คะแนน 2 ทุกครั้ง และทุกครั้งที่ผู้ตรวจคนที่ 1 ให้คะแนน 3 ผู้ตรวจคนที่ 2 ให้คะแนน 1 ค่าสหสัมพันธ์ระหว่างความคงที่ระหว่างผู้ตรวจ 2 คน นั้นจะสูงมาก แต่ระดับความสอดคล้องระหว่างผู้ตรวจนี้จะต่ำ เพราะว่าพวกเขาจะไม่ให้ค่าที่เป็นคะแนนซ้ำกันเหมือนเดิม ข้อดีของสัมประสิทธิ์ r_{wg} ก็คือสัมประสิทธิ์ที่ได้มาจะแปรผันได้ตามค่าสมมติในการให้คะแนนเช่นเดียวกับการจัดอันดับคะแนนจากการตรวจให้คะแนน

บุญเชิด ภิญญอนันตพงษ์ (2521 : 302) ได้กล่าวถึงการประมาณค่าความเชื่อมั่นในการให้คะแนน (Reliability of Rating) หรือค่าความเชื่อมั่นของผู้ตรวจให้คะแนน (Reliability of Rator) เมื่อผู้ให้คะแนนมีส่วนเกี่ยวข้องกับคะแนนที่ตรวจ เช่น การให้คะแนนการปฏิบัติหรือการให้คะแนนข้อสอบอัตนัย โดยใช้ผู้ตรวจให้คะแนนมากกว่าหนึ่งคนขึ้นไป ลักษณะเช่นนี้มักจะมีความคลาดเคลื่อนเกิดขึ้นได้เสมอ เช่น ครูวิทยาศาสตร์ 10 คน ตรวจข้อสอบอัตนัย 5 ข้อ จะพบว่าครูแต่ละคนอาจให้คะแนนแตกต่างกันไป เราจึงต้องตรวจวัดดูว่าครูแต่ละคนให้คะแนนสอดคล้องตามกันหรือไม่เพียงใด ดังนั้นการประมาณค่าความเชื่อมั่นในการตรวจให้คะแนนของแต่ละคนจึงเป็นการคำนวณหาความสัมพันธ์ระหว่างผู้ให้คะแนนแต่ละคนจะให้คะแนนสอดคล้องตรงกันมากน้อยเพียงใดนั่นเอง ซึ่งจะเป็นสหสัมพันธ์ภายในกลุ่ม (Intracorrelation) การตรวจให้คะแนนอาจจำแนกเป็นสองลักษณะตามผู้ให้คะแนนคือ

1. การตรวจให้คะแนนแบบสมบูรณ์ หมายถึง การที่ผู้ตรวจแต่ละคนให้คะแนนข้อสอบของนักเรียนแต่ละคนครบหมด เช่น นักเรียน 3 คน ครู 2 คน นักเรียนแต่ละคนจะได้รับคะแนนจากครูทั้งสองคน
2. การตรวจให้คะแนนแบบไม่สมบูรณ์ หมายถึง ผู้ตรวจบางคนไม่ตรวจให้คะแนน

นักเรียนบางคน เช่น ศาสตราจารย์ทางฟิสิกส์ 8 คน ตรวจให้คะแนนความสามารถทางการวิจัยนิสิตปริญญาโท 28 คน ศาสตราจารย์แต่ละคนขอร้องว่าเขาจะให้คะแนนเฉพาะนิสิตที่เลือกวิจัยในหัวข้อที่ตนรู้เท่านั้น ดังนั้นนิสิตบางคนจะได้รับคะแนนจากผู้ตรวจไม่ครบทุกคน

3.3 ดัชนีความสอดคล้องของผู้ประเมิน (Rater Agreement Indexes : RAI)

เบอร์รี สตอค และคนอื่น ๆ (Burry – Stock and others. 1996) ได้ศึกษาดัชนีความสอดคล้องของผู้ประเมิน (Rater Agreement Indexes : RAI) ได้ผลการศึกษาที่สามารถสรุปวิธีการคำนวณได้จากผลการสังเกตพฤติกรรมในระดับช่วงคะแนน (I) ที่มีข้อดัชนีจากค่ามากที่สุด คือ 1 ไปหาค่าน้อยที่สุด คือ 0 สรุปแนวคิดจากการตรวจให้คะแนนของผู้ประเมินหลาย ๆ คน ได้ 6 วิธี ดังนี้

1. RAI สำหรับข้อสอบ 1 ข้อ ผู้ประเมิน 2 คน สำหรับ 1 เนื้อหา เป็นสูตรคำนวณจากผลการให้คะแนนของผู้ประเมิน 2 คน ให้คะแนนพฤติกรรมเดียวกัน ที่มีพิสัยของดัชนีจากค่ามากที่สุด (I) ถ้าผู้ประเมิน 2 คน ให้คะแนนเหมือนกัน ค่า RAI = 1 ถ้าผู้ประเมิน 2 คน ให้คะแนนตรงกันข้าม ค่า RAI = 0

$$RAI = 1 - \frac{|R_1 - R_2|}{I - 1}$$

เมื่อ R_1 แทน ผลการตรวจของผู้ตรวจให้คะแนนคนที่ 1
 R_2 แทน ผลการตรวจของผู้ตรวจให้คะแนนคนที่ 2
 I แทน จำนวนช่วงคะแนน (0 , 1 , 2 , 3 , 4 , 5)

2. RAI ที่ใช้ผลรวมของคะแนนที่มีผู้ตรวจให้คะแนน 2 คน 1 เนื้อหา และ K พฤติกรรม ใช้คะแนนที่ได้จากการสังเกตสำหรับ k พฤติกรรมของผู้ตรวจให้คะแนน 2 คน สูตรนี้สามารถแสดงค่า RAI สำหรับพฤติกรรม k

$$RAI = 1 - \frac{\sum_{k=1}^K |R_{1k} - R_{2k}|}{k(I - 1)}$$

เมื่อ R_{1k} แทน คะแนนที่ได้จากการตรวจของผู้ตรวจคนที่ 1
 R_{2k} แทน คะแนนที่ได้จากการตรวจของผู้ตรวจคนที่ 2
 I แทน จำนวนช่วงคะแนน (0 , 1 , 2 , 3 , 4 , 5)

3. RAI ที่ใช้ผลรวมของคะแนนที่มีผู้ตรวจให้คะแนน 2 คน n เนื้อหา k พฤติกรรม สำหรับผลการให้คะแนนจากผู้ตรวจให้คะแนน 2 คน ที่ให้คะแนนในวิชาที่ n ใน k พฤติกรรมที่เหมือนกันและต่อเนื่องกัน

$$RAI = 1 - \frac{\sum_{k=1}^K \sum_{n=1}^n |R_{1kn} - R_{2kn}|}{kn(I - 1)}$$

เมื่อ R_{1kn} แทน คะแนนที่ได้จากการตรวจของผู้ตรวจคนที่ 1 สำหรับพฤติกรรมที่ k ในวิชาที่ n

R_{2k} แทน คะแนนที่ได้จากการตรวจของผู้ตรวจคนที่ 2 สำหรับพฤติกรรมที่ k ในวิชาที่ n

I แทน จำนวนช่วงคะแนน (0, 1, 2, 3, 4, 5)

สำหรับการตรวจที่มีจำนวนผู้ให้คะแนนมีมากกว่า 2 คน หรือมีผู้ให้คะแนน M คน ให้ $\bar{R}$ แสดงค่าเฉลี่ยจากผู้ให้คะแนน M คน

4. RAI สำหรับข้อสอบข้อเดียว ผู้ตรวจให้คะแนนหลายคน 1 วิชา มีที่มาของคะแนนจากผู้ตรวจให้คะแนน M คน ที่ให้คะแนนที่เหมือนกันใน 1 พฤติกรรม ที่ต่อเนื่องกันที่ I ระดับ

$$RAI = 1 - \frac{\sum_{m=1}^M |R_m - \bar{R}|}{(M-1)(I-1)}$$

$$\text{โดยที่ } \bar{R} = \frac{1}{M} \sum_{m=1}^M R_m$$

เมื่อ R แทน ผลการตรวจของผู้ตรวจให้คะแนน

M แทน จำนวนผู้ตรวจให้คะแนนทั้งหมด

I แทน จำนวนช่วงคะแนน (0, 1, 2, 3, 4, 5)

5. RAI สำหรับผู้ตรวจให้คะแนน M คน K พฤติกรรม 1 วิชา ซึ่งหาค่าเฉลี่ยของพฤติกรรมที่ k ของ RAIs ของผู้ตรวจให้คะแนน M คน ใน 1 พฤติกรรม (นั่นคือ RAI ของผู้ตรวจให้คะแนน M คน ใน 1 พฤติกรรม สามารถหาค่าเฉลี่ยของผู้ตรวจให้คะแนน 2 คน) RAIs ของผู้ตรวจให้คะแนน 2 คน หรือแบบคู่ แสดงได้ถึงข้อเปรียบเทียบกับผู้ตรวจให้คะแนนคนเดียวกับค่าเฉลี่ยกับผู้ให้คะแนนอื่น ๆ

$$RAI = 1 - \frac{\sum_{k=1}^K \sum_{m=1}^M |R_{mk} - \bar{R}_k|}{k(M-1)(I-1)}$$

$$\text{โดยที่ } \bar{R}_k = \frac{1}{M} \sum_{m=1}^M R_{mk}$$

เมื่อ R แทน ผลการตรวจของผู้ตรวจให้คะแนน

M แทน จำนวนผู้ตรวจให้คะแนนทั้งหมด

K แทน จำนวนพฤติกรรมย่อย

I แทน จำนวนช่วงคะแนน (0 , 1 , 2 , 3 , 4 , 5)

6. RAI สำหรับผู้ตรวจให้คะแนน M คน K พฤติกรรม N วิชา ให้ R_{mnk} เป็นคะแนนของผู้ให้คะแนน M คน ในการตรวจ k พฤติกรรม จากจำนวน N^{th} วิชา

จากข้อมูลการศึกษาค้นคว้า ผู้วิจัยได้นำสูตร 6 มาใช้ในการวิเคราะห์ข้อมูล ดังมีรายละเอียดของสูตรดังนี้

$$RAI = 1 - \frac{\sum \sum \sum |R_{mnk} - \bar{R}_{nk}|}{KN(M-1)(I-1)}$$

$$\text{โดยที่ } \bar{R}_{nk} = \frac{1}{M} \sum R_{mk}$$

เมื่อ R แทน ผลการตรวจของผู้ตรวจให้คะแนน

M แทน จำนวนผู้ตรวจให้คะแนนทั้งหมด

N แทน จำนวนนักเรียนในกลุ่มตัวอย่าง

K แทน จำนวนพฤติกรรมย่อย

I แทน จำนวนช่วงคะแนน (0 , 1 , 2 , 3 , 4 , 5)

3.4 การประมาณค่าความเชื่อมั่นตามโครงสร้างทฤษฎีการสรุปอ้างอิง

(Generalizability Theory)

ครอนบาค และคณะ (Cronbach and other. 1972) ได้เสนอทฤษฎีสำหรับวิเคราะห์ความน่าเชื่อถือของผลการวัดพฤติกรรมสำหรับสถานการณ์ของการวัดผลลักษณะต่าง ๆ โดยทฤษฎีนี้ได้ขยายแนวคิดของความเที่ยง (Reliability) ตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ให้สามารถสรุปผลความเที่ยงในสถานการณ์หรือเงื่อนไขการทดสอบลักษณะต่าง ๆ ได้ โดยทฤษฎีการสรุปอ้างอิงได้ให้แนวคิดในการแยกส่วนความคลาดเคลื่อน (Error) จากหลายแหล่ง (Multiple error sources) ประกอบด้วยแหล่งความคลาดเคลื่อนอย่างมีระบบ (Systematic source) และความคลาดเคลื่อนสุ่ม (Random source) มิใช่เป็นความคลาดเคลื่อนรวมเพียงแหล่งเดียวเหมือนอย่างทฤษฎีการทดสอบแบบดั้งเดิม

ทฤษฎีการสรุปอ้างอิงได้เสนอวิธีวิเคราะห์ค่าสัมพัทธ์ของความแปรปรวนของคะแนนความคลาดเคลื่อนอย่างเป็นระบบจากแหล่งต่าง ๆ อันเป็นสถานการณ์หรือเงื่อนไขของการวัด เช่น ความยาวของแบบทดสอบ ชุดของแบบทดสอบ จำนวนครั้งของการทดสอบ จำนวนผู้ตรวจ เป็นต้น รวมทั้งปฏิสัมพันธ์ระหว่างผู้สอบกับเงื่อนไขของการวัด เช่น ปฏิสัมพันธ์ระหว่างผู้สอบกับชุดของแบบสอบ ผู้สอบกับครั้งของการทดสอบ เป็นต้น จึงทำให้ทราบและสามารถควบคุมแหล่งความคลาดเคลื่อนได้ตรงประเด็น ยังผลให้สัมประสิทธิ์ความน่าเชื่อถือของการวัดสูงขึ้น ทฤษฎีการสรุป

อ้างอิงจึงเป็นทั้งวิธีประเมินความน่าเชื่อถือหรือความเที่ยงของเครื่องมือวัดผล และกลยุทธ์ของการออกแบบการวัดให้ได้ผลของการวัดที่มีความน่าเชื่อถือ หรือความเที่ยงสูงถึงระดับที่ต้องการ เพื่อนำผลไปใช้เป็นสารสนเทศสำหรับการตัดสินใจได้อย่างมีประสิทธิภาพ

3.4.1 โครงสร้างของทฤษฎีสรุปอ้างอิง

ความเข้าใจเกี่ยวกับทฤษฎีการสรุปอ้างอิงทางการทดสอบเบื้องต้นนั้นจำเป็นต้องทำความเข้าใจและมีความรู้ในเรื่องพื้นฐานทางโครงสร้างของทฤษฎีนี้ในเรื่องต่าง ๆ ดังนี้ (Brennan. 1983 : 2 – 6)

1. องค์ประกอบ (Facet) หมายถึง ชุดของเงื่อนไขการวัดที่คล้ายคลึงกัน (Similar Conditions of Measurement) เช่น องค์ประกอบของข้อสอบ (Item Facet) องค์ประกอบของผู้ตรวจข้อสอบ (Rater Facet) เป็นต้น

2. เงื่อนไขการวัด (Condition) หมายถึง ระดับขององค์ประกอบที่ทำให้ได้ค่าสังเกตแต่ละค่าขึ้นมาในการวัดครั้งหนึ่ง ๆ อาจกล่าวได้ว่า เป็นขนาดของกลุ่มตัวอย่างขององค์ประกอบนั้น เช่น ข้อสอบแต่ละข้อเป็นเงื่อนไขการวัดหนึ่ง ๆ ขององค์ประกอบแบบทดสอบ

3. รูปแบบความสัมพันธ์ระหว่างองค์ประกอบ หมายถึง องค์ประกอบที่บ่งชี้แหล่งความแปรปรวนของรูปแบบการวัด ในแบบจำลองการวิเคราะห์ความแปรปรวน รูปแบบความสัมพันธ์นี้ จัดแบ่งออกเป็น 3 ลักษณะ คือ

ความสัมพันธ์แบบไขว้ (Crossed) หมายถึง ความสัมพันธ์แต่ละลักษณะที่แต่ละระดับของสิ่งที่ถูกวัด ถูกวัดภายใต้เงื่อนไขเดียวกันทั้งหมด สัญลักษณ์คือ “ X ” อ่านว่า “Crossed with” เช่น กำหนดรูปแบบของความสัมพันธ์ระหว่างองค์ประกอบเป็น $p \times i \times r$ หมายถึง นักเรียน (p) ทำข้อสอบ (i) ทุกข้อ และผู้ตรวจ (r) ตรวจข้อสอบของนักเรียนทุกคน ทุก ๆ ข้อ

ความสัมพันธ์แบบแฝง (Nested) หมายถึง ความสัมพันธ์ในลักษณะที่แต่ละระดับของสิ่งที่ถูกวัดภายใต้เงื่อนไขที่แตกต่างกัน สัญลักษณ์คือ “ : ” อ่านว่า “Nested with” เช่น รูปแบบความสัมพันธ์ระหว่างองค์ประกอบเป็น $i : r$ หมายถึง ผู้ตรวจ (r) ตรวจข้อสอบ (i) ต่างข้อกัน

ความสัมพันธ์แบบผสม (Confounded) หมายถึง ความสัมพันธ์ที่มีทั้งความสัมพันธ์แบบไขว้และความสัมพันธ์แบบแฝงปนกันอยู่ เช่น $p \times (i : t)$ หมายถึง นักเรียน (p) ทำข้อสอบ (i) บางข้อที่อยู่ในแบบทดสอบ (t)

4. เอกภพ (Universe) หมายถึง เงื่อนไขของการวัดทั้งหมดของแต่ละองค์ประกอบ

เช่น จำนวนข้อสอบทุกข้อ จำนวนผู้ตรวจทั้งหมด เอกภพมีความหมายคล้ายกับคำว่า ประชากร (Population) แต่คำว่าเอกภพมักใช้กับองค์ประกอบที่เกี่ยวข้องกับการวัด ส่วนคำว่าประชากรมักใช้กับองค์ประกอบที่เป็นสิ่งที่ถูกวัด

คะแนนเอกภพ (Universe Scores) หมายถึง คะแนนเงื่อนไขการวัดทั้งหมดของแต่ละองค์ประกอบซึ่งคล้ายกับคะแนนจริง (Ideal Scores) ของสิ่งที่ถูกวัดในทฤษฎีการวัดแบบมาตรฐานเดิม

5. เอกภพของการสังเกตที่ยอมรับได้ (Universe of Admissible Observation) เป็นกลุ่มเงื่อนไขของการวัดที่สามารถวัดหรือสังเกตได้ในแต่ละองค์ประกอบ เช่น ถ้าต้องการวัดผลการเรียนของนักเรียนโดยครูต้องการข้อสอบที่สามารถนำไปใช้ได้ และต้องการผู้ตรวจที่มีคุณภาพ ครูจึงได้กำหนดเงื่อนไขของการวัด เป็นองค์ประกอบของข้อสอบ และองค์ประกอบของผู้ตรวจให้คะแนน เรียกว่า เอกภพของการสังเกตที่ยอมรับได้ของครูคนนี้ ประกอบด้วย องค์ประกอบของข้อสอบและองค์ประกอบของผู้ตรวจให้คะแนน

6. เอกภพของการสรุปอ้างอิง (Universe of Generalization) หมายถึง เงื่อนไขของการวัดทั้งหมดขององค์ประกอบที่ผู้วิจัยต้องการสรุปอ้างอิงผลการวัด จากกลุ่มตัวอย่างของเงื่อนไขการวัดเหล่านี้ไปยังกลุ่มเงื่อนไขของการวัดทั้งหมดขององค์ประกอบนั้น ๆ ดังนั้น เงื่อนไขขององค์ประกอบในเอกภพของการสรุปอ้างอิงจึงเป็นสับเซต หรือเซตย่อย ๆ ในเอกภพของการสังเกตที่ยอมรับได้ เอกภพของการสรุปอ้างอิงนี้ ผู้วิจัยต้องระบุประเด็นต่าง ๆ ให้ชัดเจน ซึ่งนำไปใช้ในการออกแบบการวัดในการวิเคราะห์เพื่อการตัดสินใจ (D Study)

7. การศึกษาเพื่อการสรุปอ้างอิง (Generalizability Study : G Study) หมายถึง การวิเคราะห์การสรุปอ้างอิง เพื่อประมาณค่าความแปรปรวนจากแหล่งต่าง ๆ ของรูปแบบการวัดโดยเริ่มจากการกำหนดองค์ประกอบต่าง ๆ กำหนดความสัมพันธ์ขององค์ประกอบ และกำหนดเอกภพของการสังเกตที่ยอมรับได้ และมีการวิเคราะห์ความแปรปรวน การประมาณค่าความแปรปรวนจากแหล่งต่าง ๆ

8. การศึกษาเพื่อการตัดสินใจ (Decision Study : D Study) หมายถึง การศึกษาเพื่อประเมินวิธีการวัดที่เหมาะสมเพื่อการตัดสินใจหรือลงข้อสรุป เป็นขั้นสุดท้ายของการวิเคราะห์การสรุปอ้างอิง โดยเริ่มจากผู้วิจัยต้องระบุสิ่งที่ต้องการวัด เอกภพของการอ้างอิง จำนวนเงื่อนไขของการวัด หรือขนาดของกลุ่มตัวอย่างของแต่ละองค์ประกอบ และรูปแบบของการวัด แล้วประมาณค่าความแปรปรวนจากการศึกษา ค่าสัมประสิทธิ์การสรุปอ้างอิง คือ ความเชื่อมั่นของแบบทดสอบ และสุดท้ายปรับปรุงรูปแบบการวัดเพื่อหาวิธีที่เหมาะสมเพื่อให้เกิดประโยชน์สูงสุดต่อไป

ขนาดของกลุ่มตัวอย่างในการศึกษาเพื่อการตัดสินใจ ไม่จำเป็นต้องมีขนาดเหมือนกับขนาดของกลุ่มตัวอย่างที่ใช้ในการศึกษาเพื่อการสรุปอ้างอิง แต่ขึ้นอยู่กับการนำไปใช้เพื่อการตัดสินใจเกี่ยวกับสิ่งที่จะเกิดขึ้นจริง ๆ ซึ่งจะถูกกำหนดขึ้นโดยมีขนาดของกลุ่มตัวอย่างในการศึกษาเพื่อการสรุปอ้างอิงเป็นพื้นฐาน

9. ความแปรปรวนของความคลาดเคลื่อนแบบสัมบูรณ์ (Absolute Error

Variance : $\sigma^2(\Delta)$ หมายถึง ความแปรปรวนของผลต่างระหว่างคะแนนการสังเกตกับคะแนนเอกภพ ซึ่งมีค่าเท่ากับผลรวมของความแปรปรวนที่ประมาณได้ทั้งหมด ยกเว้นความแปรปรวนของแหล่งบุคคล

10. ความแปรปรวนของความคลาดเคลื่อนแบบสัมพัทธ์ (Relative Error

Variance : $\sigma^2(\delta)$ หมายถึง ความแปรปรวนของผลต่างระหว่างส่วนเบี่ยงเบนของคะแนนสังเกตและส่วนเบี่ยงเบนของคะแนนเอกภพ (ความแตกต่างของคะแนนเอกภพจากคะแนนเฉลี่ยของประชากรของคะแนนเอกภพ)

11. สัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient : $E\rho^2$) หมายถึง

อัตราส่วนของความแปรปรวนของคะแนนเอกภพและความแปรปรวนของคะแนนสังเกตที่คาดหวัง ซึ่งสัมประสิทธิ์การสรุปอ้างอิงที่ประมาณค่าได้จากกำลังสองของค่าสหสัมพันธ์ระหว่างคะแนนเอกภพและคะแนนสังเกต

สัมประสิทธิ์การสรุปอ้างอิงสำหรับการวัดที่มีการตัดสินใจในเชิงเปรียบเทียบหรือการวัดแบบอิงกลุ่ม ความคลาดเคลื่อนในการวัดเป็นความคลาดเคลื่อนแบบสัมพัทธ์ ดังนั้น ค่าสัมประสิทธิ์การสรุปอ้างอิงประมาณได้จาก อัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพและผลบวกของความแปรปรวนคะแนนเอกภพกับความแปรปรวนของความคลาดเคลื่อนสัมพัทธ์

สัมประสิทธิ์การสรุปอ้างอิงสำหรับการวัดที่มีการตัดสินใจแบบสมบูรณ์ หรือการวัดแบบอิงเกณฑ์ ความคลาดเคลื่อนในการวัดเป็นความคลาดเคลื่อนแบบสมบูรณ์ ค่าสัมประสิทธิ์การสรุปอ้างอิงประมาณค่าได้จาก อัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพและผลบวกของความแปรปรวนของคะแนนเอกภพ กับความแปรปรวนของความคลาดเคลื่อนแบบสมบูรณ์

3.4.2 หลักพื้นฐานของทฤษฎีการสรุปอ้างอิง

1. ข้อตกลงเบื้องต้นของทฤษฎีการสรุปอ้างอิง อาศัยข้อตกลงของแบบการวิเคราะห์ความแปรปรวนที่สัมพันธ์กับรูปแบบของการวัดที่ต้องการ โดยทั่วไปมีข้อตกลงดังนี้

ต้องระบุเอกภพที่ต้องการสรุปอ้างอิงให้ชัดเจน จนสามารถบอกได้ว่ามีเงื่อนไขใดบ้างที่เป็นสมาชิกของเอกภพนั้น

เงื่อนไขการวัดเป็นอิสระต่อกัน กล่าวคือ คะแนนของนักเรียนที่ทำข้อสอบข้อ 1 ถูกหรือผิดไม่ขึ้นอยู่กับการตอบข้ออื่น

คะแนนการสังเกต (X) สมมติให้เป็นค่าการวัดในมาตราช่วง (Interval Scale)

2. ความหมายของค่าสัมประสิทธิ์การสรุปอ้างอิง

ค่าสัมประสิทธิ์การสรุปอ้างอิง เป็นดัชนีที่สามารถอธิบาย ความแม่นยำของการวัด เช่นเดียวกับค่าสัมประสิทธิ์ความเชื่อมั่นแบบมาตรฐานเดิม สามารถใช้คำนวณความเชื่อมั่นของคะแนนเอกภพหรือใช้ในสมการถดถอย ในการประมาณค่าคะแนนเอกภพและใช้ในการปรับแก้ค่าสหสัมพันธ์ที่ลดลงอันเนื่องมาจากความคลาดเคลื่อน (Correcting Correlation for Attenuation) (Cronbach and others. 1972 : 156)

ค่าสัมประสิทธิ์การสรุปอ้างอิง เป็นค่ากำลังสองของสหสัมพันธ์ระหว่าง คะแนนเอกภพกับคะแนนสังเกต (Brennan and Kane. 1979 : 40)

ค่าสัมประสิทธิ์การสรุปอ้างอิง อธิบายได้ในรูปอัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพกับคะแนนสังเกต (Cronbach and others. 1972)

3. ข้อแตกต่างระหว่างค่าความเชื่อมั่นกับค่าสัมประสิทธิ์การสรุปอ้างอิง ถึงแม้ว่าค่าสัมประสิทธิ์การสรุปอ้างอิง มีความหมายเช่นเดียวกันค่าความเชื่อมั่นของทฤษฎีการวัดแบบมาตรฐานเดิม แต่มีประเด็นที่แตกต่างกันไปดังนี้

การวัดแต่ละครั้งมีค่าสัมประสิทธิ์การสรุปอ้างอิงได้มากกว่า 1 ค่า

การสรุปอ้างอิงไปยังเอกภพใด ต้องระบุและ อธิบายเอกภพนั้นให้ชัดเจน และต้องรู้เงื่อนไขนั้นมาศึกษาด้วย

ค่าสัมประสิทธิ์การสรุปอ้างอิง สามารถบอกถึงความเป็นเอกพันธ์ของเอกภพได้ ถ้าข้อสอบที่นำมาศึกษา เป็นตัวอย่างสุ่มจากเอกภพข้อสอบที่มีความเป็นเอกพันธ์ ก็สามารถหาค่าคะแนนสังเกตแทนคะแนนเอกภพได้อย่างมั่นใจ

4. การวิเคราะห์สัมประสิทธิ์การสรุปอ้างอิง แนวคิดในการประมาณค่าพารามิเตอร์ของทฤษฎีการสรุปอ้างอิง ครอนบัค และคนอื่น ๆ (Cronbach and others. 1972) เบรนนอน (Brennan. 1983) เสนอวิธีการประมาณค่าพารามิเตอร์เป็น 2 ขั้นตอน ดังนี้

ขั้นที่ 1 การศึกษาเพื่อการสรุปอ้างอิง (G Study) เป็นขั้นแรกของการวิเคราะห์ข้อมูลเพื่อที่จะประมาณค่าความแปรปรวนจากแหล่งต่าง ๆ ซึ่งมีกระบวนการดังนี้

1.1 กำหนดสิ่งที่จะวัด (Object of Measurement) เช่น นักเรียนมัธยมศึกษาปีที่ 4 เป็นต้น

1.2 กำหนดองค์ประกอบ (Facets) ของการวัด เช่น องค์ประกอบข้อสอบ องค์

ประกอบผู้ตรวจ เป็นต้น

1.3 กำหนดเอกภาพของการสังเกตที่ยอมรับได้ เช่น การวิจัยครั้งนี้ เอกภาพของการสังเกตที่ยอมรับได้ คือ องค์ประกอบข้อสอบ และองค์ประกอบของผู้ตรวจให้คะแนน

1.4 กำหนดความสัมพันธ์ระหว่างองค์ประกอบ เช่น ถ้าต้องการศึกษาแบบให้ผู้สอบทุกคนทำข้อสอบทุกข้อ ก็จะเป็นแบบไขว้ (Crossed) คือ $p \times i$ (p คือ นักเรียน และ i คือ ข้อสอบ) แต่ถ้าให้นักเรียนทุกคนทำข้อสอบบางข้อจะเป็นแบบ (Nested) คือ $p : i$

1.5 กำหนดชนิดขององค์ประกอบว่าเป็นชนิดใดดังต่อไปนี้

15.1 องค์ประกอบสุ่มแท้จริง (Purely Random Facets) ได้แก่ องค์ประกอบที่ทุกเงื่อนไขการวัดได้มาจากการสุ่มอย่างง่ายจากเอกภาพที่มีขนาดไม่จำกัด (∞)

15.2 องค์ประกอบสุ่มจำกัด (Finite Random Facets) ได้แก่ องค์ประกอบที่ทุกเงื่อนไขการวัดได้มาจากการสุ่มอย่างง่ายจากเอกภาพที่มีขนาดจำกัด

15.3 องค์ประกอบกำหนด (Finite Random Facets) ได้แก่ องค์ประกอบที่ประกอบด้วย เงื่อนไขการวัดจำนวนจำกัดในเอกภาพ และเงื่อนไขการวัดเหล่านี้ถูกเลือกมาศึกษาทั้งหมด

1.6 กำหนดรูปแบบ (Model) การวัดว่าเป็นรูปแบบใดในต่อไปนี้

16.1 รูปแบบกำหนด (Fixed Model) เป็นรูปแบบที่ประกอบด้วย องค์ประกอบกำหนดทั้งหมด

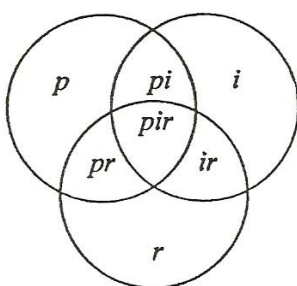
16.2 รูปแบบสุ่ม (Random Model) เป็นรูปแบบที่ประกอบด้วย องค์ประกอบสุ่มแท้จริงและหรือองค์ประกอบสุ่มจำกัด

16.3 รูปแบบผสม (Mixed Model) เป็นรูปแบบที่ประกอบด้วย องค์ประกอบสุ่มแท้จริงและ / หรือสุ่มจำกัด และองค์ประกอบกำหนดผสมกัน

การกำหนดสิ่งที่วัด องค์ประกอบ เอกภาพของการสังเกตที่ยอมรับได้และความสัมพันธ์ระหว่างองค์ประกอบว่าเป็นแบบ “Crossed” หรือ “Nested” แล้วก็ใช้ ANOVA ทำการวิเคราะห์ให้สอดคล้องกับแบบแผนของความสัมพันธ์ขององค์ประกอบเพื่อประมาณค่า Mean Square แล้วประมาณค่าความแปรปรวนของแหล่งต่าง ๆ เราเรียกค่าประมาณความแปรปรวนของสิ่งที่ถูกวัดว่า ค่าประมาณความแปรปรวนของคะแนนเอกภาพ (Estimated Universe Score Variance) ซึ่งสอดคล้องกับค่าประมาณความแปรปรวนของคะแนนจริง (True Score Variance) ของทฤษฎีวัดมาตรฐานเดิม เราเรียกค่าประมาณความแปรปรวนขององค์ประกอบอื่นว่า ค่าประมาณความแปรปรวนของความคลาดเคลื่อน (Estimated Error Score Variance) การวิเคราะห์ขั้น G

Study ใช้ค่าขนาดของกลุ่มตัวอย่างที่ศึกษาจริง ๆ แต่ในการวิเคราะห์ขั้น D Study สามารถกำหนดขนาดของกลุ่มตัวอย่างแตกต่างไปจากที่วิเคราะห์ใน G Study ได้

1.7 การเก็บรวบรวมและการวิเคราะห์ข้อมูล เช่น ในรูปแบบสุ่ม (Random Model) โดยการสุ่มกลุ่มตัวอย่างมาศึกษา คือ ผู้สอบ n_p คน ข้อสอบ n_i ข้อ และผู้ตรวจ n_r คน ซึ่งเป็นการศึกษา G Study แบบ $p \times i \times r$ คือ ผู้สอบทุกคนทำข้อสอบทุกข้อและผู้ตรวจแต่ละคนจะตรวจข้อสอบทุกข้อ แล้วคำนวณค่าความแปรปรวนขององค์ประกอบต่าง ๆ วิธีการดังกล่าวนี้จะได้แหล่งความแปรปรวน 7 แหล่ง (Brennan. 1983 : 27 – 28) เขียนแทนด้วยแผนภูมิ Venn ได้ดังภาพประกอบ 1 ดังนี้



ภาพประกอบ 1 แผนภูมิ Venn แสดงองค์ประกอบความแปรปรวนของแบบแผน $p \times i \times r$ เมื่อองค์ประกอบที่ศึกษาทั้งหมดเป็นองค์ประกอบสุ่ม

จากภาพประกอบจะพบว่ามีอิทธิพลหลัก (Main Effects) 3 แหล่ง ได้จากอิทธิพลจากผู้สอบ (p) อิทธิพลจากข้อสอบ (i) ส่วนผลปฏิสัมพันธ์ (Interaction Effects) 4 แหล่ง คือ ปฏิสัมพันธ์ระหว่างผู้สอบและข้อสอบ (pi) ปฏิสัมพันธ์ระหว่างผู้สอบและผู้ตรวจ (pr) ปฏิสัมพันธ์ระหว่างข้อสอบและผู้ตรวจ (ir) และปฏิสัมพันธ์ระหว่างผู้สอบ ข้อสอบ และผู้ตรวจ (pir)

ค่าความแปรปรวนของแหล่งต่าง ๆ แหล่งนี้คำนวณจาก Mean Square โดยใช้ ANOVA แบบ Factorial Design $p \times i \times r$ ในการประมาณค่าความแปรปรวนจาก Mean Square นั้นใช้สูตรเฉพาะอย่างที่สุดคล้องกับรูปแบบ เช่น สูตรในการประมาณค่าความแปรปรวนของแหล่งต่าง ๆ เมื่อองค์ประกอบที่ศึกษาเป็นองค์ประกอบสุ่มทั้งหมด และกำหนดรูปแบบความสัมพันธ์ระหว่างองค์ประกอบเป็นแบบ Crossed ใช้สูตรของ (Brennan. 1983 : 130) ดังนี้

$$\begin{aligned}
 EMS_p &= \sigma^2(pir) + n_i \sigma_{(pr)}^2 + n_i \sigma_{(pi)}^2 + n_i n_r \sigma_{(p)}^2 \\
 EMS_i &= \sigma^2(pir) + n_p \sigma_{(ir)}^2 + n_r \sigma_{(pi)}^2 + n_p n_r \sigma_{(i)}^2 \\
 EMS_r &= \sigma^2(pir) + n_r \sigma_{(pr)}^2 + n_p \sigma_{(ir)}^2 + n_p n_i \sigma_{(r)}^2 \\
 EMS_{pi} &= \sigma^2(pir) + n_r \sigma_{(pi)}^2 \\
 EMS_{pr} &= \sigma^2(pir) + n_i \sigma_{(pr)}^2
 \end{aligned}$$

$$EMS_{ir} = \sigma^2(pir) + n_p \sigma_{(ir)}^2$$

$$EMS_{(pir)} = \sigma^2(pir)$$

โครงสร้างของสมการนี้สัมพันธ์กับแผนภูมิ Venn ข้างต้น กล่าวคือในวงกลม p ประกอบด้วยพจน์ (Term) ต่าง ๆ 4 พจน์ มี p, pi, pr และ pir เป็นเทอมทางขวามือของสูตรในการหา EMS_p หรือในพื้นที่ pi ซึ่งเป็นส่วนร่วมระหว่างวงกลม p และ i ประกอบด้วย 2 พจน์ คือ pi และ pir ก็คือพจน์ที่ปรากฏอยู่ทางขวามือของสูตรในการหา EMS_{pi} นั่นเอง ดังนั้นจึงสามารถสร้างสูตรหาค่า EMS โดยอาศัยแผนภูมิ Venn โดยใช้ n_p เป็นตัวคูณควบถ้าองค์ประกอบนั้นไม่รวม p และคูณ n_i ถ้าองค์ประกอบนั้นไม่มี i รวมอยู่ด้วย และ n_r ถ้าองค์ประกอบนั้นไม่มี r รวมอยู่ด้วย

การคำนวณเมื่อองค์ประกอบที่ศึกษาเป็นแบบผสม คือ มีทั้งองค์ประกอบกำหนดและองค์ประกอบสุ่ม และรูปแบบความสัมพันธ์ระหว่างองค์ประกอบเป็นแบบ Crossed สมมติให้ผู้ตรวจ (r) เป็นองค์ประกอบกำหนด สูตรในการคำนวณรวมทั้งแหล่งความแปรปรวนที่ต้องการประมาณค่าเป็นดังนี้ (Brennan. 1983 : 71 – 72)

$$EMS_p = n_r \sigma_{(pi)}^2 + n_i n_r \sigma_{(p)}^2$$

$$EMS_i = n_r \sigma_{(pi)}^2 + n_p n_r \sigma_{(i)}^2$$

$$EMS_{pi} = n_r \sigma_{(pi)}^2$$

เนื่องจากองค์ประกอบของผู้ตรวจ (r) เป็นองค์ประกอบที่กำหนด จึงหายไปจากแผนภูมิเดิมเหลือเพียงวงกลม p และ r เท่านั้น

ค่าความแปรปรวนที่ได้จาก G Study เป็นค่าประมาณความแปรปรวนที่แท้จริงของเอกภพของเงื่อนไขการวัดที่มีขนาด 1 หน่วย กล่าวคือ เป็นความแปรปรวนของข้อสอบ 1 ข้อ ผู้ตรวจ 1 คน ของนักเรียน 1 คน เช่น $\sigma_{(p)}^2$ เป็นค่าประมาณความแปรปรวนของ $\sigma_{(p)}^2$ หมายความว่า จากประชากรที่ทำการวัดนั้น จะได้คะแนนของนักเรียน (p) แต่ละคนจากการสอบข้อสอบทุก ๆ ข้อ (N_i) และได้รับการตรวจจากผู้ตรวจทุกคน (N_r) ในเอกภพของการสังเกตคะแนนเฉลี่ยของนักเรียนแต่ละคน คือคะแนนเอกภพ (μ_p) ดังนั้นความแปรปรวนของ (μ_p) จากการสอบวัดทุก ๆ คนในประชากรคือ $\sigma_{(\mu_p)}^2$ หรืออาจเขียนให้ง่ายว่า $\sigma_{(p)}^2$ เป็นความแปรปรวนของคะแนนเอกภพ ตรงกับความแปรปรวนของคะแนนจริงตามทฤษฎีการวัดมาตรฐานเดิม

ขั้นที่ 2 การศึกษาเพื่อการตัดสินใจสรุปอ้างอิง (D Study) เป็นขั้นการนำผลที่ได้จากการประมาณความแปรปรวนในการศึกษาเพื่อการสรุปอ้างอิง (G Study) มาเลือกวิธีวัดที่เหมาะสมที่สุดภายใต้การตัดสินใจ มีขั้นตอนดังนี้

2.1 กำหนดเอกภพของการสรุปอ้างอิง (Universe of Generalization) เป็นเป้าหมายสำคัญของ D Study ได้แก่การกำหนดลักษณะเฉพาะของเอกภพของการอ้างอิงที่ผู้ตัดสินใจต้องการอ้างอิงถึง อาจจะประกอบด้วยเงื่อนไขทั้งหมดในเอกภพของการสังเกตที่ยอมรับได้ หรืออาจเป็นเซตย่อย (Subset) ของเอกภพของการสังเกตที่ยอมรับได้ (Brennan. 1983 : 3)

2.2 กำหนดรูปแบบ (Model) หมายถึง จะใช้รูปแบบใด เช่น รูปแบบสุ่ม (Random Model) หรือรูปแบบกำหนด (Fixed Model) หรือรูปแบบผสม (Mixed Model) ซึ่งโดยหลักการแล้วควรใช้โมเดลแบบสุ่มดีที่สุด เพราะจะทำให้การสรุปอ้างอิงได้กว้างขวางกว่าโมเดลแบบผสมและโมเดลแบบกำหนดตามลำดับ

2.3 ขนาดของกลุ่มตัวอย่างของ D Study (D Study Sample Sizes) หมายถึง จำนวนเงื่อนไขขององค์ประกอบใน D Study ไม่จำเป็นต้องเหมือนกัน และสามารถกำหนดได้แตกต่างจากใน G Study โดยใช้สัญลักษณ์ (') แทนขนาดของกลุ่มตัวอย่างใน D Study เช่น n'_i และ n'_r แทนจำนวนข้อสอบและจำนวนผู้ตรวจ

2.4 กำหนดโครงสร้างของแบบที่ศึกษา (D Study design Structure) โครงสร้างของแบบที่ศึกษาหรือความสัมพันธ์ขององค์ประกอบที่ศึกษา อาจจะเหมือนกันหรือแตกต่างกันจาก G Study ก็ได้

2.5 การประมาณค่าความแปรปรวนในชั้น D Study (Estimate D Study Variance Components) ในชั้นนี้จะมีการประมาณค่าความแปรปรวนชั้น มาใหม่ ขึ้นอยู่กับการตัดสินใจของนักวิจัยโดยอาศัยผลการประมาณค่าในชั้น G Study เป็นฐานและให้สอดคล้องกับแบบแผน และขนาดกลุ่มตัวอย่างที่ต้องการตัดสินใจ เช่นในแบบแผน $p \times I \times R$ ขนาดตัวอย่างข้อสอบ = n'_i และขนาดของผู้ตรวจ = n'_r ดังแสดงในตาราง 7 (Brennan. 1983 : 58)

ตาราง 7 การประมาณค่าความแปรปรวนขึ้นมาใหม่ของแบบแผน $p \times I \times R$

ความแปรปรวนใหม่	ความแปรปรวนเดิม / ขนาดของกลุ่มตัวอย่าง
$\sigma^2_{(P)}$	$\sigma^2_{(P)}$
$\sigma^2_{(I)}$	$\sigma^2_{(i)} / n'_i$
$\sigma^2_{(R)}$	$\sigma^2_{(r)} / n'_r$
$\sigma^2_{(PI)}$	$\sigma^2_{(pi)} / n'_i$
$\sigma^2_{(PR)}$	$\sigma^2_{(pr)} / n'_r$
$\sigma^2_{(IR)}$	$\sigma^2_{(ir)} / n'_i n'_r$
$\sigma^2_{(PIR)}$	$\sigma^2_{(pir)} / n'_i n'_r$

2.6 การประมาณค่าความแปรปรวนของความคลาดเคลื่อน นอกจากการประมาณค่าองค์ประกอบความแปรปรวน (Variance Components) ตามรูปแบบและขนาดของกลุ่มตัวอย่างที่ต้องการแล้ว ยังจะต้องประมาณค่าความแปรปรวนของความคลาดเคลื่อนซึ่งแบ่งเป็น 2 ชนิด คือ

2.6.1 ความแปรปรวนคลาดเคลื่อนสัมบูรณ์ (Absolute Error Variance) ใช้กับสัญลักษณ์ว่า $\sigma^2(\Delta)$ เป็นความแปรปรวนของความแตกต่างระหว่างคะแนนที่ได้จากการสังเกตกับคะแนนที่ได้จากเอกภพของผู้เข้าสอบ คำนวณจากผลบวกของค่าองค์ประกอบความแปรปรวนทั้งหมด ยกเว้นความแปรปรวนของคะแนนเอกภพ จากแบบแผน $p \times I \times R$ จะได้

$$\sigma^2(\Delta) = \sigma^2(I) + \sigma^2(R) + \sigma^2(pI) + \sigma^2(pR) + \sigma^2(IR) + \sigma^2(pIR)$$

2.6.2 ความแปรปรวนคลาดเคลื่อนสัมพัทธ์ (Relative Error Variance) ใช้สัญลักษณ์ว่า $\sigma^2(\delta)$ หมายถึง ค่าความแปรปรวนของผลต่างระหว่างส่วนเบี่ยงเบนคะแนนที่สังเกตได้ และส่วนเบี่ยงเบนคะแนนเอกภพ คำนวณจากผลบวกของค่าองค์ประกอบความแปรปรวนระหว่างสิ่งที่ถูกวัดกับองค์ประกอบอื่น ๆ จากแบบแผน $p \times I \times R$ จะได้

$$\sigma^2(\delta) = \sigma^2(pI) + \sigma^2(pR) + \sigma^2(pIR)$$

2.7 การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง เป็นขั้นสุดท้ายของ D Study คือการคำนวณค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) ซึ่งเป็นดัชนีที่บ่งถึงความเชื่อถือได้ของการวัด (Dependability of Measurement) ซึ่งใช้สูตรคำนวณได้จากอัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพและความแปรปรวนของคะแนนที่สังเกตที่คาดหวัง

ในการวัดแบบอิงกลุ่ม จะได้สูตรที่ใช้ในการคำนวณสัมประสิทธิ์การสรุปอ้างอิง ดังนี้
(Brennan. 1983 : 17)

$$E\rho^2 = \frac{\sigma^2(\rho)}{\sigma^2(\rho) + \sigma^2(\Delta)}$$

หรือ

$$E\rho^2 = \frac{\sigma^2(\rho)}{\sigma^2(\rho) + \sigma^2(\delta)}$$

เมื่อ $E\rho^2$ คือ ค่าสัมประสิทธิ์การสรุปอ้างอิงซึ่งมีความหมายเช่นเดียวกับความเชื่อมั่น

4. หลักฐานแสดงความเที่ยงตรง

4.1 ความหมายของความเที่ยงตรง

ความเที่ยงตรงของแบบทดสอบเป็นอีกคุณสมบัติหนึ่ง ที่สำคัญที่สุดของเครื่องมือวัดทุกชนิด ซึ่งนักวัดผลการศึกษาได้ให้ความหมายไว้ดังนี้

อนาสตาซี (Anastasi. 1968: 29) ได้ให้ความหมายของความเที่ยงตรงของแบบทดสอบว่าเป็นระดับที่แบบทดสอบวัดได้จริงในสิ่งที่ต้องการจะวัด

สแตนเลย์ และฮอบกินส์ (Stanley and Hopkins. 1972: 101) ได้กล่าวว่า ความเที่ยงตรงของการวัด หมายถึง การวัดนั้นตรงตามหน้าที่ที่จะวัดได้ดีเพียงใด หรือระดับที่การวัดสามารถบรรลุจุดมุ่งหมายบางอย่าง

ทักแมน (Tuckman. 1975: 114) กล่าวว่า ความเที่ยงตรงของแบบทดสอบ หมายถึง แบบทดสอบฉบับหนึ่งวัดในสิ่งที่เราต้องการให้วัด หรือตรงตามจุดมุ่งหมายที่ต้องการวัดหรือไม่

ลัวัน สายยศ และ อังคณา สายยศ (2539: 259) ได้กล่าวไว้ว่า ความเที่ยงตรงเชิงโครงสร้าง หมายถึง คุณภาพของเครื่องมือที่สามารถวัดได้ตรงตามทฤษฎีต่าง ๆ ของโครงสร้างนั้น หรือวัดได้ครอบคลุมตามลักษณะโครงสร้างของแบบทดสอบมาตรฐาน

บุญเชิด ภิญญอนันตพงษ์ (2547: 172-173) สรุปได้ว่า เครื่องมือวัดที่มีความเที่ยงตรง หมายถึง เครื่องมือวัดนั้นสามารถวัดในสิ่งที่ผู้วัดต้องการวัดได้อย่างถูกต้อง ดังนั้น ในแนวคิดเดิมจึงถือว่า ความเที่ยงตรง เป็นคุณภาพของเครื่องมือวัด ซึ่งเป็นสมบัติที่สำคัญของเครื่องมือทุกชนิด ไม่ว่าจะเป็นแบบทดสอบผลสัมฤทธิ์ แบบทดสอบความถนัด มาตรวัดทัศนคติ จริยธรรม และบุคลิกภาพ เครื่องมือวัดทุกฉบับจะต้องมีคุณภาพด้านความเที่ยงตรง จึงจะเชื่อได้ว่าเป็นเครื่องมือวัดที่ดีและผลที่ได้จากการวัดจะถูกต้องตรงตามที่ต้องการ การวัดใด ๆ จะไร้ประโยชน์เมื่อเครื่องมือที่ใช้วัดนั้นขาด

ความเที่ยงตรง และความเที่ยงตรงของเครื่องมือวัด แบ่งเป็นหลายชนิด ขึ้นอยู่กับตัวแปร เกณฑ์ที่ใช้เทียบ

ส่วนแนวคิดปัจจุบัน ได้ให้ความหมายของความเที่ยงตรงตามมาตรฐานทางเทคนิค สำหรับประเมินคุณภาพของเครื่องมือวัด ซึ่งคณะกรรมการร่วมประกอบด้วยสมาชิกของสมาคมคือ สมาคมวิจัยการศึกษาอเมริกัน สมาคมจิตวิทยาอเมริกัน และสภาการวัดผลการศึกษาแห่งชาติของสหรัฐอเมริกา ได้เสนอไว้ในหนังสือ “Standards for Educational and Psychological Testing” ที่พิมพ์ในปี 1985 ให้ความหมายความเที่ยงตรงไว้ว่า ความเที่ยงตรงเป็นลักษณะสำคัญที่สุดที่ใช้ในการประเมินเครื่องมือวัด ซึ่งเป็นเรื่องเกี่ยวกับการลงความเห็น (Inference) จากคะแนนที่วัดได้อย่างเหมาะสม อย่างมีความหมาย และให้ประโยชน์ ส่วนการตรวจสอบความเที่ยงตรง (Validation) เป็นกระบวนการในการรวบรวมพยานหลักฐาน (Evidence) เพื่อสนับสนุนการลงความเห็นดังกล่าว และสามารถเก็บรวบรวมหลักฐานได้หลายวิธี ความเที่ยงตรงจึงขึ้นอยู่กับปริมาณ และชนิดของหลักฐานที่สามารถสนับสนุนการลงความเห็นของผู้วัดจากข้อมูลที่วัดได้

4.2 วิธีการแสดงหลักฐานความเที่ยงตรง

บุญเชิด ภิญญอนันตพงษ์ (2547 : 178 – 195) กล่าวว่า เนื่องจากหลักฐานความเที่ยงตรงของค่าวัดจากเครื่องมือวัดเป็นความสัมพันธ์ หรือความสอดคล้องระหว่างค่าวัดของเครื่องมือวัดนั้น กับสิ่งที่ต้องการวัดหรือตัวเกณฑ์ ดังนั้น การแสดงหลักฐานความเที่ยงตรง จึงเป็นการหาความสัมพันธ์หรือความสอดคล้องระหว่างค่าวัดของตัวแปรทั้งสอง วิธีการแสดงหลักฐานความเที่ยงตรงจึงขึ้นอยู่กับชนิดของค่าวัดที่ได้จากตัวแปรทั้งสอง ดังนี้

1. วิธีการแสดงหลักฐานความเที่ยงตรงตามเนื้อหา เป็นการแสดงหรือหาว่าเครื่องมือวัดนั้น สามารถวัดได้ตรงและครอบคลุมเนื้อหาวิชา การแสดงหลักฐานความเที่ยงตรงมากน้อยเพียงใด โดยการเทียบกับตารางวิเคราะห์หลักสูตร หรือตารางกำหนดข้อสอบ (Test Blueprint) ซึ่งกำหนดตัวอย่างหัวข้อเนื้อหาสาระวิชาและพฤติกรรมจากเนื้อหาสาระวิชาทั้งหมด และถือว่าเป็นตัวแทนที่ดีแล้ว

การแสดงหลักฐานความเที่ยงตรงตามเนื้อหาของเครื่องมือวัด สามารถพิจารณา จากความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ของวิชา (ซึ่งจะครอบคลุมทั้งเนื้อหาและพฤติกรรมที่ต้องการวัด) โดยคำนวณจากดัชนีความสอดคล้องของระหว่างข้อสอบกับจุดประสงค์ ค่า IOC ที่มีค่า .5 ขึ้นไป แสดงว่า ข้อสอบวัด หรือเป็นตัวแทนจุดประสงค์ของวิชา

2. วิธีการแสดงหลักฐานความเที่ยงตรงตามเกณฑ์ เป็นการหาค่าความสัมพันธ์ของค่าที่วัดได้จากเครื่องมือวัดที่ต้องการกับค่าที่วัดได้จากเกณฑ์ ซึ่งสามารถคำนวณได้หลายวิธีขึ้นอยู่กับค่าที่วัดได้ ซึ่งอาจใช้สูตร Pearson Product Moment (เมื่อข้อมูลเป็นคะแนนทั้ง 2 ชุด) หรือ Spearman

Rank Order (เมื่อข้อมูลเป็นการจัดอันดับ) ซึ่งแยกตามเกณฑ์เป็นการแสดงหลักฐานความเที่ยงตรงตามสภาพ และการแสดงหลักฐานความเที่ยงตรงตามพยากรณ์

3. วิธีการแสดงหลักฐานความเที่ยงตรงตามโครงสร้าง เป็นการแสดงหลักฐานความเที่ยงตรงว่าเครื่องมือวัดนั้นสามารถวัดขอบเขตความหมาย หรือคุณลักษณะประจำตามโครงสร้างทางทฤษฎีที่สมมุติขึ้นนั้นได้เพียงใด การแสดงหลักฐานความเที่ยงตรงตามโครงสร้างสามารถทำได้หลายวิธีการ ดังนี้

3.1 วิธีพิจารณาเทียบกับโครงสร้างที่กำหนด เครื่องมือวัดผลการเรียนที่เขียนข้อสอบวัดตามตารางลักษณะเฉพาะ หรือตารางวิเคราะห์หลักสูตร สามารถแสดงหลักฐานความเที่ยงตรงตามโครงสร้างได้โดยให้ผู้เชี่ยวชาญวิเคราะห์

3.2 วิธีเปรียบเทียบจากกลุ่มที่ต่างกัน การศึกษาว่าเครื่องมือวัดโครงสร้างของสิ่งที่จะวัดได้โดยใช้กลุ่มตัวอย่างที่ต่างกันสองกลุ่มที่รู้แน่ชัดว่า กลุ่มหนึ่งมีคุณลักษณะในสิ่งที่ต้องการวัด ส่วนอีกกลุ่มหนึ่งไม่มีคุณลักษณะในสิ่งนั้น แล้วเปรียบเทียบคะแนนที่ได้จากทั้งสองกลุ่ม แล้วใช้ t - test ทดสอบ ก็สามารถสรุปว่า เครื่องมือนั้นมีหลักฐานแสดงความเที่ยงตรงตามโครงสร้างสูง

3.3 วิธีเทียบกับเครื่องมือมาตรฐานที่วัดคุณลักษณะเดียวกัน ค่าสหสัมพันธ์ของเครื่องมือวัด กับเครื่องมือมาตรฐานที่วัดคุณลักษณะเดียวกันสามารถบ่งชี้ หลักฐานความเที่ยงตรงตามโครงสร้างได้

3.4 วิธีการวิเคราะห์หลายลักษณะหลายวิธี (Multitrait-Multimethods) การแสดงหลักฐานความเที่ยงตรงตามโครงสร้างต่างจากการแสดงหลักฐานความเที่ยงตรงตามเกณฑ์ที่ต้องใช้หลักฐานต่าง ๆ มากกว่า โดยอาศัยสมมติฐานที่ว่า ถ้าเครื่องมือวัดกับเกณฑ์มีลักษณะร่วมกันจะมีค่าสหสัมพันธ์กันสูง และถ้าเครื่องมือวัดกับเกณฑ์มีลักษณะต่างกันจะมีค่าสหสัมพันธ์กันต่ำ นำมาวิเคราะห์พร้อมกัน ซึ่งแคมเบลล์และฟิสค์ (Campbell and Fiske) ได้พัฒนาแนวคิดนี้ให้เหมาะสมเรียกว่า การวิเคราะห์หลายลักษณะ หลายวิธีการ โดยใช้วิธีการเทียบความเที่ยงตรงร่วม (Convergent Validity) กับความเที่ยงตรงแยก (Divergent Validity) ซึ่งความเที่ยงตรงร่วมควรมีค่าสูงกว่า และความเที่ยงตรงแยกควรมีค่าต่ำกว่า แต่ค่าความเที่ยงของเครื่องมือวัดแต่ละวิธีควรมีค่าสูงสุด

3.5 วิธีการหาค่าความสอดคล้องภายในเครื่องมือวัด การแสดงหลักฐานความเที่ยงตรงตามโครงสร้างดังกล่าวไปแล้ว จะอาศัยสหสัมพันธ์ระหว่างเครื่องมือวัดกับเกณฑ์ภายนอกที่ยอมรับ สำหรับวิธีนี้จะอาศัยความสอดคล้องภายในเครื่องมือวัด โดยไม่ใช้เกณฑ์ภายนอก ซึ่งสามารถพิจารณาจากดัชนีต่าง ๆ ดังนี้

3.5.1 พิจารณาจากดัชนีอำนาจจำแนกรายข้อเพราะข้อสอบที่มีค่าอำนาจ

จำแนกสูง เป็นข้อสอบที่วัดในทิศทางเดียวกันกับส่วนรวม ถือว่ามีหลักฐานความเที่ยงตรงตามโครงสร้างภายในสูง

3.5.2 พิจารณาจากระดับความสัมพันธ์ระหว่างคะแนนส่วนย่อยภายในเครื่องมือวัดกับคะแนน

3.5.3 พิจารณาจากค่าความเชื่อมั่นของเครื่องมือวัดที่ หาด้วยสูตรความสอดคล้องภายใน เช่น สูตร KR-20 หรือสูตร แอลฟา ของครอนบัค (Cronbach's Coefficient-Alpha) ดังนั้น เครื่องมือวัดได้มีค่าความเชื่อมั่นสูง ก็สามารถสรุปว่า มีหลักฐานความเที่ยงตรงตามโครงสร้างภายในสูงได้

3.6 วิธีการวิเคราะห์องค์ประกอบ (Factor Analysis) การแสดงหลักฐานความเที่ยงตรงตามโครงสร้างที่ตรงประเด็นมากที่สุดคือ วิเคราะห์องค์ประกอบ เพราะวิธีการทางสถิติที่สามารถตรวจชี้ ลักษณะประจำทางจิตวิทยา เนื่องจากตัวแปรต่าง ๆ เมื่อนำมาหาค่าสัมประสิทธิ์สหสัมพันธ์ จะพบว่า มีตัวแปรบางคู่มีความ สัมพันธ์กันสูงหรือบางที่ก็พบว่า มีกลุ่มตัวแปรบางกลุ่มมีความสัมพันธ์ซึ่งกันและกันสูง นั่นแสดงว่าตัวแปรเหล่านั้นวัดบางสิ่งบางอย่างที่เป็นองค์ประกอบร่วมกัน การวิเคราะห์องค์ประกอบ เป็นการจัดสมรรถภาพ หรือคุณลักษณะต่าง ๆ ทางจิตวิทยาที่วัดได้ให้เป็นหมวดหมู่ตามโครงสร้าง ซึ่งค่าน้ำหนักองค์ประกอบแรกก่อนหมุนแกน จะเป็นค่าที่แสดงหลักฐานความเที่ยงตรงตามโครงสร้างได้

การวิเคราะห์ข้อมูลทางสถิติที่มีหลักการเชิงวิชาการ มีวิธีการวิเคราะห์ซับซ้อน เป็นวิธีการที่มีอำนาจสูง และเป็นประโยชน์อย่างยิ่งต่อการวิจัยทางสังคมศาสตร์ และพฤติกรรมศาสตร์จนได้รับการยกย่องว่าเป็นวิธีการที่เยี่ยมยอดทางการวิเคราะห์ข้อมูลทางสถิติทั้งปวง คือ การวิเคราะห์องค์ประกอบ (เพ็ญนภา กุลนภาดล. 2547: 62; อ้างอิงจาก Kerlinger. 1973: 659) ซึ่งการวิเคราะห์องค์ประกอบ (Factor Analysis) เป็นชื่อทั่วไปที่ใช้เรียกววิธีการวิเคราะห์ข้อมูลที่มีวิธีการ และเป้าหมายการวิเคราะห์ต่างกัน คือ การวิเคราะห์ส่วนประกอบ การวิเคราะห์องค์ประกอบร่วม การวิเคราะห์องค์ประกอบเชิงสำรวจ และการวิเคราะห์องค์ประกอบเชิงยืนยัน วิธีการวิเคราะห์องค์ประกอบเหล่านี้ไม่ว่าวิธีใดวิธีหนึ่งต่างก็เป็นวิธีการที่มีประโยชน์ต่อนักวิจัยทั้งสิ้น

นางลักษณ์ วิรัชชัย (2542: 122) กล่าวถึง การวิเคราะห์องค์ประกอบเป็นวิธีการอธิบายข้อมูลให้ง่ายขึ้นด้วยการลดจำนวนตัวแปร (variable reduction) โดยการพยายามหาโครงสร้างตัวประกอบจำนวนน้อย ๆ ที่จะแทนตัวแปรจำนวนมาก ๆ ทั้งนี้เนื่องจาก การวิจัยทางสังคมศาสตร์ และการวิจัยทางพฤติกรรมศาสตร์นั้น เรามุ่งเก็บรวบรวมข้อมูลซึ่งเป็นลักษณะภายในที่ไม่สามารถสังเกตได้โดยตรงหรืออาจเรียกว่าตัวแปรแฝง และต้องศึกษาคุณลักษณะดังกล่าวนี้จากพฤติกรรมการแสดงออกของบุคคล โดยการวัดหรือการสังเกตพฤติกรรมเหล่านั้น แทนคุณลักษณะที่ต้องการศึกษา

ในทางปฏิบัติ นักวิจัยจะเก็บรวบรวมข้อมูลได้เป็นตัวแปรสังเกตได้หลายตัว และใช้การวิเคราะห์องค์ประกอบมาวิเคราะห์ข้อมูล เพื่อให้ได้องค์ประกอบอันเป็นคุณลักษณะที่ผู้วิจัยต้องการศึกษา กล่าวได้ว่า วิธีการวิเคราะห์องค์ประกอบเป็นการวิเคราะห์ข้อมูลทางสถิติ ที่ช่วยให้ นักวิจัยสร้างองค์ประกอบจากตัวแปรหลายๆ ตัวแปรโดยรวมกลุ่มตัวแปรที่เกี่ยวข้องสัมพันธ์กันเป็นองค์ประกอบเดียวกัน และแต่ละองค์ประกอบ คือ ตัวแปรแฝงอันเป็นคุณลักษณะที่ผู้วิจัยต้องการศึกษา

โดย การวิเคราะห์องค์ประกอบมีวัตถุประสงค์ 2 ประการ คือ

1. เพื่อสำรวจระบุ หรือศึกษาว่า องค์ประกอบร่วมที่สามารถอธิบายความสัมพันธ์ร่วมกันระหว่างตัวแปรต่าง ๆ โดยที่จำนวนองค์ประกอบร่วมที่หาได้ จะมีจำนวนน้อยกว่าจำนวนตัวแปร ซึ่งเป็นการลดจำนวนตัวแปรลง และได้องค์ประกอบหลัก ๆ เรียกว่า การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis Model) การวิเคราะห์ในลักษณะนี้ทำให้ง่ายต่อการทำความเข้าใจลักษณะของข้อมูล และสะดวกในการแปลความหมาย

2. เพื่อต้องการทดสอบสมมุติฐานเกี่ยวกับแบบแผน โครงสร้างขององค์ประกอบว่า องค์ประกอบแต่ละตัว ประกอบด้วยตัวแปรอะไรบ้าง และตัวแปรแต่ละตัวควรมีน้ำหนักหรืออัตราความสัมพันธ์กับองค์ประกอบมากน้อยเพียงใด ตรงกับที่คาดคะเนไว้หรือไม่ หรือสรุปได้ว่า เพื่อต้องการทดสอบว่าองค์ประกอบอย่างนี้ ตรงกับโมเดลหรือตรงกับทฤษฎีที่มีอยู่หรือไม่ โมเดลนี้เรียกว่า Confirmatory Factor Analysis Model

แซทฟีลด์ และโคลลิน (เพ็ญญา กุลนาคดล. 2547: 63; อ้างอิงจาก Chatfield and Collins. 1980: 89) กล่าวว่า ในปัจจุบันนักวิจัยเริ่มใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis Model: CFA) แทนการวิเคราะห์องค์ประกอบเชิงสำรวจมากขึ้น สาเหตุที่เป็นเช่นนี้เนื่องจาก องค์ประกอบเชิงสำรวจมีรูปแบบวิเคราะห์หลากหลาย และได้ผลการวิเคราะห์ที่ไม่สอดคล้องกัน นอกจากนี้การวิเคราะห์องค์ประกอบเชิงสำรวจยังมีข้อตกลงเบื้องต้นที่เข้มงวดและไม่ตรงตามความเป็นจริง เช่น ข้อตกลงเบื้องต้นที่ว่า ตัวแปรสังเกตได้ทุกตัวเป็นผลมาจากองค์ประกอบร่วมทุกตัว ส่วนที่เป็นความคลาดเคลื่อนของตัวแปรไม่สัมพันธ์กัน รวมทั้งสเกลองค์ประกอบที่สร้างขึ้นแปลความหมายได้ยาก เพราะในบางครั้งสเกลองค์ประกอบเกิดจากการสุ่มตัวแปรที่ไม่น่าจะมีองค์ประกอบร่วมกัน จุดอ่อนของการวิเคราะห์องค์ประกอบเชิงสำรวจนี้ ทำให้ Long (1983: 12) กล่าวว่า สำหรับนักวิจัยหลายคนใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงสำรวจเป็น GIGO model (Garbage In and Garbage Out model) และเสนอว่า ถ้าทำได้นักวิจัยไม่ควรใช้การวิเคราะห์องค์ประกอบเชิงสำรวจเลย

เทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยัน เป็นการวิเคราะห์องค์ประกอบที่มีการปรับปรุงจุดอ่อนของการวิเคราะห์องค์ประกอบเชิงสำรวจได้เกือบทั้งหมด โดยข้อตกลงเบื้องต้นของการ

วิเคราะห์องค์ประกอบเชิงยืนยันมีความสมเหตุสมผลตรงตามความเป็นจริงมากกว่าการวิเคราะห์องค์ประกอบเชิงสำรวจ นักวิจัยต้องมีทฤษฎีสนับสนุนในการกำหนดเงื่อนไขบังคับ ซึ่งใช้ในการวิเคราะห์หาค่าน้ำหนักองค์ประกอบ และเมื่อได้ผลการวิเคราะห์แล้วยังมีการตรวจสอบความกลมกลืนของโมเดลกับข้อมูลเชิงประจักษ์อีกด้วย รวมทั้งยังมีการตรวจสอบโครงสร้างของโมเดลว่ามีความแตกต่างกันระหว่างกลุ่มตัวอย่างหลาย ๆ กลุ่มหรือไม่ (นงลักษณ์ วิรัชชัย. 2542: 122)

วัตถุประสงค์ของการใช้การวิเคราะห์องค์ประกอบเชิงยืนยันมี 3 ประการ (นงลักษณ์ วิรัชชัย. 2542: 122) คือ

1. นักวิจัยใช้เทคนิคการวิเคราะห์องค์ประกอบนี้ เพื่อตรวจสอบทฤษฎีที่ใช้เป็นพื้นฐานในการวิเคราะห์องค์ประกอบ
2. ใช้เพื่อสำรวจและระบุองค์ประกอบ
3. ใช้เป็นเครื่องมือในการสร้างตัวแปรใหม่ แต่เทคนิคนี้สามารถใช้วิเคราะห์ข้อมูลโดยมีข้อตกลงเบื้องต้น น้อยกว่าเทคนิคการวิเคราะห์องค์ประกอบเชิงสำรวจ เช่น ส่วนที่เป็นความคลาดเคลื่อนอาจสัมพันธ์กันได้

สุทธิพงศ์ บุญผดุง (2541: 10 – 13) ได้ศึกษาสูตรการหาค่าความเที่ยงตรงเชิงโครงสร้างของ ไฮเซนและบอร์นสเตรดต์ (Heise and Bohrnstedt. 1970: 109, 111 – 113) ซึ่งไฮเซนและบอร์นสเตรดต์ ได้เสนอวิธีหาค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ โดยใช้ทฤษฎีตัวประกอบร่วม (Common Factor Theory) ในรูปของการหาสหสัมพันธ์ (Correlation) ระหว่างคะแนนจริงของคุณลักษณะที่ได้ T_k กับคะแนนองค์ประกอบ S_k เมื่อคะแนนจริงของคุณลักษณะ (Traits) สัมพันธ์กันแบบออร์โธโกนอล (Orthogonal : $\rho_{T_1 T_2} = 0$) นั่นคือ แต่ละคุณลักษณะ (T_1, T_2) ไม่มีความสัมพันธ์กัน และการหาน้ำหนักองค์ประกอบของคำถามข้อที่ i จากองค์ประกอบที่ k สามารถหาได้จาก การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis) จากการหมุนแกนแบบมูมจาก ด้วยวิธีแวนริแมกซ์ (Varimax) และสกัดองค์ประกอบด้วยวิธี PC (Principal Component Analysis) โดยใช้คอมพิวเตอร์โปรแกรม SPSS

สูตรสำหรับหาค่าความเที่ยงตรง ที่สุทธิพงศ์ บุญผดุง (2541: 10 -13) ได้ศึกษาจากสูตรของไฮเซนและบอร์นสเตรดต์ (Heise and Bohrnstedt. 1970: 109, 111 – 113) มีดังนี้

$$\rho_{T_k S_k} = \frac{\sum_{i=1}^n S_i f_{ik}}{\sqrt{\sum_{i=1}^n \sum_{j=1}^n \text{Cov}(x_i, x_j)}}$$

$\rho_{T_k S_k}$ แทน ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ

S_i แทน ส่วนเบี่ยงเบนมาตรฐานของข้อคำถามข้อที่ i

f_{ik} แทน น้ำหนักองค์ประกอบของคำถามข้อที่ i จากองค์ประกอบที่ k

$Cov(x_i, x_j)$ แทน ค่าความแปรปรวนร่วมระหว่างคำถามข้อที่ i กับคำถามข้อที่ j

จากวิธีดังกล่าวผู้วิจัยเลือกใช้การคำนวณค่าความเที่ยงตรงด้วยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยัน และการคำนวณด้วยสูตรไฮเซน และ บอร์นสเตรดต์ (Heise and Bohrnstedt. 1970 : 109, 111 – 113)

5. ทักษะกระบวนการทางวิทยาศาสตร์

5.1 ความหมายทักษะกระบวนการทางวิทยาศาสตร์

คลอฟเฟอร์ (Klopfer. 1971 : 568 – 573) ให้ความหมายทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า ทักษะกระบวนการทางวิทยาศาสตร์เป็นกระบวนการที่ใช้ในการแสวงหาความรู้ทางวิทยาศาสตร์

ปีเตอร์สัน (Peterson. 1978 : 153) ได้ให้ความหมายทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า เป็นปฏิบัติการสืบเสาะแสวงหาความรู้ทางวิทยาศาสตร์ที่ประกอบด้วย การสังเกต การตั้งคำถาม การทดลอง การเปรียบเทียบ การสรุปหาพิพจน์ การสรุปหลักเกณฑ์ การสื่อความหมายและการนำไปใช้

แอนเดอร์สัน (Anderson. 1978 : 15) ได้ให้ความหมายทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า เป็นวิธีการที่นักวิทยาศาสตร์ใช้ในการเสาะแสวงหาความรู้ ความหมายที่สำคัญของกระบวนการในการเสาะแสวงหาความรู้ ทำให้เกิดพัฒนาการทางด้านสติปัญญา

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2524 : 1) ให้ความหมายทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า เป็นพฤติกรรมที่เกิดจากการปฏิบัติและการฝึกฝนความคิดอย่างมีระบบ ซึ่งก่อให้เกิดความงอกงามทางสติปัญญา

วรรณทิพา รอดแรงคำ และจิต นวนแก้ว (2533 : V) ให้ความหมายทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า เป็นทักษะทางสติปัญญาที่นักวิทยาศาสตร์และผู้ที่มีแนวโน้มวิธีการทางวิทยาศาสตร์มาแก้ปัญหา ใช้ในการศึกษาค้นคว้า สืบเสาะหาความรู้และแก้ปัญหาต่าง ๆ

สุนีย์ คล้ายนิล และคนอื่น ๆ (2535 : 10) ได้อธิบายว่าทักษะกระบวนการทางวิทยาศาสตร์เป็นกระบวนการหรือกิจกรรมที่ใช้ปฏิบัติ จนเกิดความชำนาญหรือเกิดทักษะในกิจกรรมที่ใช้ปฏิบัติจนเกิดความชำนาญ หรือเกิดทักษะในกิจกรรมนั้น ๆ ในทางวิทยาศาสตร์

วิษชุดา งามอักษร (2541 : 39) ได้สรุปความหมายของทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า ทักษะกระบวนการทางวิทยาศาสตร์ หมายถึงความสามารถในการสืบเสาะหา

ความรู้ โดยผ่านการปฏิบัติ และฝึกฝนความคิดอย่างมีระบบ จนเกิดเป็นความคล่องแคล่ว และชำนาญ

จากการศึกษานิยามความหมายทักษะกระบวนการทางวิทยาศาสตร์ จากนักการศึกษาหลาย ๆ ท่านที่กล่าวมานี้ ผู้วิจัยสรุปได้ว่าทักษะกระบวนการทางวิทยาศาสตร์เป็นปฏิบัติการสืบเสาะหาความรู้ทางวิทยาศาสตร์ โดยใช้พฤติกรรมที่เกิดจากการการฝึกฝน ความคิดอย่างมีระบบ และฝึกปฏิบัติจนเกิดความชำนาญ เพื่อใช้ในการศึกษาค้นคว้า แก้ปัญหา และสืบเสาะหาความรู้

5.2 ประเภทของทักษะกระบวนการทางวิทยาศาสตร์

สมาคมเพื่อความก้าวหน้าทางวิทยาศาสตร์แห่งอเมริกา (นงนุช สหัสดี. 2545 : 13 อ้างอิงจาก The American Association for the Advancement of Science. 1970 : 33 – 176) ได้แบ่งทักษะกระบวนการทางวิทยาศาสตร์เป็น 13 กระบวนการด้วยกัน โดยแบ่งออกเป็น 2 ประเภทใหญ่ ๆ ดังนี้คือ

1. ทักษะเบื้องต้น จัดไว้ 8 ทักษะ ดังนี้

1.1 การสังเกต (Observing) หมายถึง การใช้ประสาทสัมผัสซึ่งได้แก่ จมูก ตา หู ลิ้น และผิวหนัง เพื่อเก็บรวบรวมข้อมูล คุณลักษณะ และรายละเอียดของสิ่งของหรือปรากฏการณ์อย่างใดอย่างหนึ่งทั้งที่เป็นเชิงปริมาณและคุณภาพ

1.2 การวัด (Measurement) หมายถึง การใช้เครื่องมือต่าง ๆ เพื่อเก็บรวบรวมข้อมูล รวมทั้งการกะประมาณค่าที่ควรจะได้วัดได้

1.3 การใช้จำนวนเลข (Using Number) หมายถึง การนำตัวเลขมากำหนดคุณลักษณะต่าง ๆ เช่น ความกว้าง ความยาว ความสูง พื้นที่ ปริมาตร หรือจำนวนของต่าง ๆ รวมทั้งการคำนวณเบื้องต้น เช่น การหาค่าเฉลี่ยหรืออัตราส่วน

1.4 การจัดจำพวก (Classifying) หมายถึง การจำแนกสิ่งของหรือเหตุการณ์ออกเป็นประเภทต่าง ๆ โดยพิจารณาจากคุณสมบัติที่เหมือนกัน สัมพันธ์กัน หรือต่างกันของสิ่งของหรือเหตุการณ์นั้น ๆ ซึ่งอาจมีวิธีแบ่งได้หลายวิธี ทั้งนี้ขึ้นอยู่กับเกณฑ์ที่ใช้

1.5 การสื่อความหมาย (Communicating) หมายถึง การพูดหรือการแสดงสัญลักษณ์ต่าง ๆ เช่น แผนภูมิ สมการ กราฟ หรือตัวอักษร เป็นต้น เพื่อให้บุคคลอื่นเข้าใจหรือรับทราบความคิด ความรู้สึกต่าง ๆ ได้ตามต้องการ

1.6 การใช้ความสัมพันธ์เกี่ยวกับมิติกับเวลา (Using Space – time Relationship) หมายถึง การนำเอาความสัมพันธ์ระหว่างมิติกับเวลา หรือมิติกับมิติ หรือเวลากับเวลา มาอธิบายสิ่งใดสิ่งหนึ่งหรือสถานการณ์ใดสถานการณ์หนึ่ง ในที่นี้มีมิติหมายถึงคุณสมบัติเกี่ยวกับความกว้าง ความยาว ความหนา รูปร่าง สมมาตร หรือตำแหน่งที่อยู่ของสิ่งใดสิ่งหนึ่ง

ความสัมพันธ์ระหว่างมิติกับมิติ เช่น ความสัมพันธ์ระหว่างจังหวะการแกว่งของลูกตุ้มนาฬิกา กับจังหวะการเดินของชีพจร

ความสัมพันธ์ระหว่างเวลากับเวลา เช่น การหาตำแหน่งของวัตถุที่เคลื่อนที่เมื่อเวลาเปลี่ยนไป

1.7 การสรุปอ้างอิง (Infering) หมายถึง การอธิบายปรากฏการณ์หรือข้อเท็จจริงต่าง ๆ โดยอาศัยข้อมูลที่เกิดขึ้นได้ประกอบกับประสบการณ์เดิม

1.8 การทำนาย (Predicting) หมายถึง การคาดคะเนสิ่งที่จะเกิดขึ้นในอนาคต น่าจะเป็นอย่างไร โดยอาศัยหลักฐานส่วนใหญ่ที่ได้จากการสังเกต หรือการวัดประกอบกับการสรุปอ้างอิง

2. ทักษะเชิงซ้อน จัดไว้ 5 ทักษะ ดังนี้

2.1 การให้นิยามเชิงปฏิบัติการ (Defining Operationally) หมายถึง การให้ความหมายของสิ่งใดสิ่งหนึ่ง ในรูปที่สังเกต วัด หรือนำมาปฏิบัติการได้และบอกด้วยว่าในสถานการณ์หนึ่ง ๆ จะมีวิธีสังเกต หรือวิธีวัดสิ่งนั้นได้อย่างไร

2.2 การกำหนดและควบคุมตัวแปร (Identifying, Controlling and Manipulating Variable) การกำหนดตัวแปร หมายถึง การแยกตัวแปรต่าง ๆ ออกเป็นตัวแปรต้น ตัวแปรตาม และตัวแปรอื่น ๆ ที่ต้องควบคุม

การควบคุมตัวแปร หมายถึง การพยายามทำให้สรุปได้ว่าผลการทดลอง (ตัวแปรตาม) เป็นผลจากตัวแปรต้น โดยการควบคุมตัวแปรอื่น ๆ ที่อาจมีผลต่อตัวแปรตาม

2.3 การสร้างสมมติฐาน (Formulating Hypothesis) หมายถึง การคาดการณ์ว่าตัวแปรต่าง ๆ จะมีความสัมพันธ์กันอย่างไร เป็นการลงสรุปของคำอธิบาย โดยอาศัยการสังเกตหรือการสรุปอ้างอิงเป็นพื้นฐาน

2.4 การประมวลผลและการตีความหมายข้อมูล (Data Processing and Interpreting) การประมวลผลข้อมูล หมายถึง การรวบรวมข้อมูลให้อยู่ในรูปของตาราง ข้อความ หรือข้อความกึ่งตาราง หรือกราฟ และการคำนวณค่าสถิติพื้นฐานจากข้อมูล

การตีความหมายข้อมูล หมายถึง การบอกความสัมพันธ์ของตัวแปรต่าง ๆ จากข้อมูลที่ประมวลผลมาแล้ว หรือการให้ความหมายของข้อมูลเชิงปริมาณเป็นเชิงคุณภาพ

2.5 การออกแบบการทดลอง (Designing and Investigation) หมายถึง การกำหนดโครงการทดลองเพื่อเก็บรวบรวมข้อมูลมาทดสอบสมมติฐาน โดยคำนึงถึง นิยามปฏิบัติการของตัวแปรต่าง ๆ ที่เกี่ยวข้อง การควบคุมตัวแปรต่าง ๆ เครื่องมือและวิธีการที่จะใช้ในการเก็บรวบรวมข้อมูล

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2526 : 1 – 16) ได้กล่าวถึงทักษะกระบวนการทางวิทยาศาสตร์ว่ามี 13 ทักษะ ดังต่อไปนี้

1. การสังเกต (Observation) หมายถึง การใช้ประสาทสัมผัสอย่างใดอย่างหนึ่งหรือหลายอย่างรวมกัน ได้แก่ หู ตา จมูก ลิ้น และผิวหนัง เข้าไปสัมผัสโดยตรงกับวัตถุหรือเหตุการณ์ โดยมีจุดประสงค์ที่จะหาข้อมูล ซึ่งเป็นรายละเอียดของสิ่งนั้น ๆ โดยไม่ใส่ความคิดเห็นของผู้สังเกตลงไป

2. การวัด (Measurement) หมายถึง การเลือกและการใช้เครื่องมือทำการวัดหาปริมาณของสิ่งต่าง ๆ ออกมาเป็นตัวเลขที่แน่นอนได้อย่างเหมาะสมและถูกต้อง โดยมีหน่วยกำกับเสมอ

3. การจำแนกประเภท (Classification) หมายถึง การแบ่งพวกหรือเรียงลำดับวัตถุหรือสิ่งที่อยู่ในปรากฏการณ์โดยมีเกณฑ์ดังกล่าว อาจจะใช้ความเหมือน ความแตกต่างหรือความสัมพันธ์อย่างใดอย่างหนึ่งก็ได้

4. การหาความสัมพันธ์ระหว่างสเปสกับสเปสและสเปสกับเวลา (Space / Space Relationship and Space / time Relationship) หมายถึง ที่ว่างที่วัตถุนั้นครองที่ ซึ่งจะมีรูปร่างลักษณะเช่นเดียวกับวัตถุนั้น โดยทั่วไปแล้วสเปสของวัตถุจะมี 3 มิติ คือ ความกว้าง ความยาว ความสูง

ความสัมพันธ์ระหว่างสเปสกับสเปสของวัตถุ ได้แก่ ความสัมพันธ์ระหว่างตำแหน่งที่อยู่ของวัตถุหนึ่งกับอีกวัตถุหนึ่ง

ความสัมพันธ์ระหว่างสเปสของวัตถุกับเวลา ได้แก่ ความสัมพันธ์ระหว่างการเปลี่ยนแปลงตำแหน่งที่อยู่ของวัตถุกับเวลา หรือความสัมพันธ์ระหว่างสเปสของวัตถุที่เปลี่ยนไปกับเวลา

5. การคำนวณ (Using Number) หมายถึง การนับจำนวนของวัตถุและการนับตัวเลขแสดงจำนวนที่นับได้ มาคิดคำนวณโดยการบวก ลบ คูณ หาร หรือหาค่าเฉลี่ย

6. การจัดกระทำและสื่อความหมายข้อมูล (Organizing Data and Communion) หมายถึง การนำข้อมูลที่ได้จากการสังเกต การวัด การทดลอง และจากแหล่งอื่น ๆ มาจัดกระทำเสียใหม่ โดยการหาความถี่ เรียงลำดับ จัดแยกประเภท หรือคำนวณหาค่าใหม่ เพื่อให้ผู้อื่นเข้าใจความหมายของข้อมูลชุดนั้นดีขึ้น โดยอาจเสนอในรูปของตาราง แผนภูมิ แผนภาพ ไดอะแกรม วงจร กราฟ สมการ เขียน บรรยาย เป็นต้น

7. การลงความคิดเห็นจากข้อมูล (Inferring) หมายถึง การเพิ่มความคิดเห็นให้กับข้อมูลที่ได้จากการสังเกตอย่างมีเหตุผล โดยอาศัยความรู้หรือประสบการณ์เดิมมาช่วย

8. การพยากรณ์ (Prediction) หมายถึง การสรุปคำตอบล่วงหน้าก่อนจะทดลอง โดยอาศัยปรากฏการณ์ที่เกิดขึ้น ๆ หลักการ กฎ หรือทฤษฎีที่มีอยู่แล้วในเรื่องนั้น ๆ มาช่วยในการสรุป

การพยากรณ์ข้อมูลเกี่ยวกับตัวเลข ได้แก่ ข้อมูลที่เป็นตาราง หรือกราฟ ทำได้ 2 แบบ คือ การพยากรณ์ภายในขอบเขตของข้อมูลที่มีอยู่กับการพยากรณ์ภายนอกขอบเขตของข้อมูลที่มีอยู่

9. การตั้งสมมติฐาน (Formulating Hypothesis) หมายถึง การคิดหาคำตอบล่วงหน้าจะทำการทดลอง โดยอาศัยหลักการสังเกต ความรู้ ประสบการณ์เดิมเป็นพื้นฐานคำตอบที่คิดหาล่วงหน้านี้อย่างไม่ทราบหรือยังไม่เป็นหลักการ กฎหรือทฤษฎีมาก่อน สมมติฐานหรือคำตอบที่คิดไว้ล่วงหน้ามักกล่าวไว้เป็นข้อความที่บอกความสัมพันธ์ระหว่างตัวแปรต้นกับตัวแปรตาม สมมติฐานที่ตั้งไว้อาจถูกหรือผิดก็ได้ ซึ่งจะทราบได้ภายหลังการทดลองหาคำตอบเพื่อสนับสนุนหรือคัดค้านสมมติฐานที่ตั้งไว้

10. การกำหนดนิยามเชิงปฏิบัติการ (Defining Operationally) หมายถึง การกำหนดความหมายและขอบเขตของคำต่าง ๆ ให้เข้าใจตรงกันและสามารถสังเกตหรือวัดได้

11. การกำหนดและควบคุมตัวแปร (Identifying and Controlling Variables) หมายถึง การชี้แจงตัวแปรที่ต้องควบคุมในสมมติฐานหนึ่ง ๆ

ตัวแปรต้น คือ สิ่งที่เป็นสาเหตุที่ทำให้เกิดผลต่าง ๆ หรือสิ่งที่เราต้องการทดลองดูว่าเป็นสาเหตุที่ก่อให้เกิดผลเช่นนั้นจริงหรือไม่

ตัวแปรตาม คือ สิ่งที่เป็นผลเนื่องมาจากตัวแปรต้น เมื่อตัวแปรต้นหรือสิ่งที่เป็สาเหตุเปลี่ยนแปลงไป ตัวแปรตามหรือสิ่งที่เป็ผลจะเปลี่ยนแปลงไปด้วย

ตัวแปรควบคุม คือ สิ่งอื่น ๆ นอกเหนือจากตัวแปรต้นที่มีผลต่อการทดลองด้วย ซึ่งจะต้องควบคุมให้เหมือน ๆ กัน มิเช่นนั้นอาจทำให้ผลการทดลองคลาดเคลื่อน

12. การทดลอง (Experimenting) หมายถึง กระบวนการปฏิบัติการเพื่อหาคำตอบหรือทดสอบสมมติฐานที่ตั้งไว้ใน การทดลอง จะประกอบด้วยกิจกรรม 3 ขั้นตอน คือ

12.1 การออกแบบการทดลอง หมายถึง การวางแผนการทดลองก่อนลงมือทดลองจริง เพื่อกำหนดวิธีการทดลอง อุปกรณ์หรือสารเคมีที่จะต้องใช้ในการทดลอง

12.2 การปฏิบัติการทดลอง หมายถึง การลงมือปฏิบัติการทดลองจริง

12.3 การบันทึกผลการทดลอง หมายถึง การจดบันทึกข้อมูลที่ได้จากการทดลอง ซึ่งอาจเป็นผลการสังเกต การวัด และอื่น ๆ

13. การตีความหมายข้อมูลและลงข้อสรุป (Interpreting Data and Conclusion)

หมายถึง การแปลความหมายหรือการบรรยายลักษณะและสมบัติของข้อมูลที่มีอยู่ การตีความหมาย ข้อมูลในบางครั้ง อาจต้องใช้ทักษะกระบวนการทางวิทยาศาสตร์อื่น ๆ ด้วย เช่น ทักษะการสังเกต ทักษะการคำนวณ เป็นต้น

การลงข้อสรุป หมายถึง การสรุปความสัมพันธ์ของข้อมูลทั้งหมด ความสามารถที่แสดงว่าเกิดทักษะแล้ว คือ แปลความหมาย หรือบรรยายลักษณะและสมบัติของข้อมูลที่มีอยู่ได้ บอกความสัมพันธ์ของข้อมูลที่มีอยู่ได้

ทักษะกระบวนการทางวิทยาศาสตร์ทั้ง 13 ทักษะที่กล่าวมา แบ่งเป็น 2 ประเภทใหญ่ คือ ทักษะขั้นพื้นฐาน (Basic Science Process Skills) และทักษะขั้นบูรณาการ (Integrated Science Process Skills) ซึ่งผู้วิจัยได้นำเอาพฤติกรรมที่แสดงว่าเกิดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการในด้านต่าง ๆ มาเป็นแนวทางการออกแบบ สร้างแบบวัดประเมินทักษะกระบวนการทางวิทยาศาสตร์

6. หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544

หลักสูตรการศึกษาขั้นพื้นฐาน (กระทรวงศึกษาธิการ. 2544 : 9 - 16) เป็นหลักสูตรที่กำหนดมาตรฐานการเรียนรู้ในการพัฒนาผู้เรียนตั้งแต่ชั้นประถมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 6 สำหรับผู้เรียนทุกคน ทุกกลุ่มเป้าหมาย สามารถปรับใช้ได้กับการจัดการศึกษาทุกรูปแบบ ทั้งในระบบ นอกระบบ และการศึกษาตามอัธยาศัย

หลักสูตรการศึกษาขั้นพื้นฐานที่สถานศึกษานำไปใช้จัดการเรียนรู้ในสถานศึกษานั้น กำหนดโครงสร้างที่เป็นสาระการเรียนรู้ จำนวนเวลา อย่างกว้าง ๆ มาตรฐานการเรียนรู้ที่แสดงคุณภาพผู้เรียนเมื่อเรียนจบ 12 ปี และเมื่อจบการเรียนรู้แต่ละช่วงชั้นของสาระการเรียนรู้แต่ละกลุ่ม สถานศึกษาต้องนำโครงสร้างดังกล่าวนี้ไปจัดทำเป็นหลักสูตรสถานศึกษา โดยคำนึงถึงสภาพปัญหา ความพร้อม เอกลักษณ์ ภูมิปัญญาท้องถิ่น และคุณลักษณะอันพึงประสงค์ ทั้งนี้สถานศึกษาต้องจัดทำรายวิชาในแต่ละกลุ่มให้ครบถ้วนตามมาตรฐานที่กำหนด ซึ่งสำหรับวิชาวิทยาศาสตร์ในช่วงชั้นที่ 4 (ม.4 - 6) นักเรียนทุกคนจะต้องเรียนให้ครบ 8 สาระ 13 มาตรฐาน ได้แก่

สาระที่ 1 : สิ่งมีชีวิตกับกระบวนการดำรงชีวิต

มาตรฐาน ว 1.1 : เข้าใจหน่วยพื้นฐานของสิ่งมีชีวิต ความสัมพันธ์ของโครงสร้างและหน้าที่ของระบบต่าง ๆ ของสิ่งมีชีวิตที่สัมพันธ์กัน มีกระบวนการสืบเสาะหาความรู้ สื่อสารสิ่งที่เรียนรู้ และนำความรู้ไปใช้ในการดำรงชีวิตของตนเอง และดูแลสิ่งมีชีวิต

มาตรฐาน ว 1.2 : เข้าใจกระบวนการและความสำคัญของการถ่ายทอดลักษณะทางพันธุกรรม วิวัฒนาการของสิ่งมีชีวิต ความหลากหลายทางชีวภาพ การใช้เทคโนโลยีชีวภาพที่มีผลต่อ

มนุษย์และสิ่งแวดล้อม มีกระบวนการสืบเสาะหาความรู้ และจิตวิทยาศาสตร์ สื่อสารสิ่งที่เรียนรู้และนำความรู้ไปใช้ประโยชน์

สาระที่ 2 : ชีวิตกับสิ่งแวดล้อม

มาตรฐาน ว 2.1 : เข้าใจสิ่งแวดล้อมในท้องถิ่น ความสัมพันธ์ระหว่างสิ่งแวดล้อมกับสิ่งมีชีวิต ความสัมพันธ์ระหว่างสิ่งมีชีวิตต่าง ๆ ในระบบนิเวศ มีกระบวนการสืบเสาะหาความรู้และจิตวิทยาศาสตร์ สื่อสารสิ่งที่เรียนรู้และนำความรู้ไปใช้ประโยชน์

มาตรฐาน ว 2.2 : เข้าใจความสำคัญของทรัพยากรธรรมชาติ การใช้ทรัพยากรธรรมชาติในระดับท้องถิ่น ประเทศ และโลก นำความรู้ไปใช้ในการจัดการทรัพยากรธรรมชาติและสิ่งแวดล้อมในท้องถิ่นอย่างยั่งยืน

สาระที่ 3 : สารและสมบัติของสาร

มาตรฐาน ว 3.1 : เข้าใจสมบัติของสาร ความสัมพันธ์ระหว่างสมบัติของสารกับโครงสร้างและแรงยึดเหนี่ยวระหว่างอนุภาค มีกระบวนการสืบเสาะหาความรู้และจิตวิทยาศาสตร์ สื่อสารสิ่งที่เรียนรู้และนำความรู้ไปใช้ประโยชน์

มาตรฐาน ว 3.2 : เข้าใจหลักการและธรรมชาติของการเปลี่ยนแปลงสถานะของสาร การเกิดสารละลาย การเกิดปฏิกิริยาเคมี มีกระบวนการสืบเสาะหาความรู้ และจิตวิทยาศาสตร์ สื่อสารสิ่งที่เรียนรู้และนำความรู้ไปใช้ประโยชน์

สาระที่ 4 : แรงและการเคลื่อนที่

มาตรฐาน ว 4.1 : เข้าใจธรรมชาติของแรงแม่เหล็กไฟฟ้า แรงโน้มถ่วง และแรงนิวเคลียร์ มีกระบวนการสืบเสาะหาความรู้ สื่อสารสิ่งที่เรียนรู้และนำความรู้ไปใช้ประโยชน์อย่างถูกต้องและมีคุณธรรม

มาตรฐาน ว 4.2 : เข้าใจลักษณะการเคลื่อนที่แบบต่าง ๆ ของวัตถุในธรรมชาติ มีกระบวนการสืบเสาะหาความรู้ และจิตวิทยาศาสตร์ สื่อสารสิ่งที่เรียนรู้และนำความรู้ไปใช้ประโยชน์

สาระที่ 5 : พลังงาน

มาตรฐาน ว 5.1 : เข้าใจความสัมพันธ์ระหว่างพลังงานกับการดำรงชีวิต การเปลี่ยนรูปพลังงาน ปฏิสัมพันธ์ระหว่างสารและพลังงาน ผลของการใช้พลังงานต่อชีวิต และสิ่งแวดล้อม มีกระบวนการสืบเสาะหาความรู้ และจิตวิทยาศาสตร์ สื่อสารสิ่งที่เรียนรู้และนำความรู้ไปใช้ประโยชน์

สาระที่ 6 : กระบวนการเปลี่ยนแปลงของโลก

มาตรฐาน ว 6.1 : เข้าใจกระบวนการต่าง ๆ ที่เกิดขึ้นบนผิวโลกและภายในโลก ความสัมพันธ์ของกระบวนการต่าง ๆ ที่มีผลต่อการเปลี่ยนแปลงภูมิอากาศ ภูมิประเทศและสิ่งแวดล้อม

ของโลก มีกระบวนการสืบเสาะหาความรู้และจิตวิทยาศาสตร์ สื่อสารสิ่งที่เรียนรู้ และนำความรู้ไปใช้ประโยชน์

สาระที่ 7 : ดาราศาสตร์ และอวกาศ

มาตรฐาน ว 7.1 : เข้าใจวิวัฒนาการของระบบสุริยะและกาแล็กซี ปฏิสัมพันธ์ภายในระบบสุริยะ และผลต่อสิ่งมีชีวิตบนโลก มีกระบวนการสืบเสาะหาความรู้และจิตวิทยาศาสตร์ สื่อสารสิ่งที่เรียนรู้ และนำความรู้ไปใช้ประโยชน์

มาตรฐาน ว 7.2 : เข้าใจความสำคัญของเทคโนโลยีอวกาศที่นำมาใช้ในการสำรวจอวกาศและทรัพยากรธรรมชาติ ด้านการเกษตรและการสื่อสาร สื่อสารสิ่งที่เรียนรู้และนำความรู้ไปใช้ประโยชน์อย่างมีคุณธรรมต่อสิ่งมีชีวิตและสิ่งแวดล้อม

สาระที่ 8 : ธรรมชาติของวิทยาศาสตร์และเทคโนโลยี

มาตรฐาน ว 8.1 : ใช้กระบวนการทางวิทยาศาสตร์และจิตวิทยาศาสตร์ ในการสืบเสาะหาความรู้ การแก้ปัญหา รู้ว่าปรากฏการณ์ทางธรรมชาติที่เกิดขึ้นส่วนใหญ่มีรูปแบบที่แน่นอนสามารถอธิบาย และตรวจสอบได้ภายใต้ข้อมูล และเครื่องมือที่มีอยู่ในช่วงเวลานั้น ๆ เข้าใจว่าวิทยาศาสตร์ เทคโนโลยี สังคมและสิ่งแวดล้อมมีความเกี่ยวข้องสัมพันธ์กัน

7. งานวิจัยที่เกี่ยวข้อง

7.1 งานวิจัยที่เกี่ยวข้องกับความเที่ยงตรงและความเชื่อมั่น

อิบราฮิม (Ibrahim. 1984 : 499 – A) ได้ประยุกต์ทฤษฎีการสรุปอ้างอิง เพื่อประมาณค่าความแปรปรวนที่มีต่อการประเมินวัตถุประสงค์ทางการศึกษา (The Rating of Evaluational Goals) โดยสุ่มกลุ่มตัวอย่างครู 80 คน และนักเรียน 80 คน ในประเทศชูดาน ประเมินวัตถุประสงค์ทางการศึกษา 2 ชนิด คือ วัตถุประสงค์ที่สำคัญจริง ๆ และวัตถุประสงค์ตามที่คาดหวังองค์ประกอบในการศึกษามีผู้ประเมิน กลุ่มผู้ประเมิน จำนวนครั้งของการประเมิน ถิ่นที่อยู่ของผู้ประเมิน ชนิดของวัตถุประสงค์ สถานที่ทำงานของผู้ประเมิน และเพศของผู้ประเมิน พบว่าองค์ประกอบที่มีผลต่อการประเมินมากที่สุด ได้แก่ ผู้ประเมินและกลุ่มผู้ประเมิน ส่วนสถานที่ทำงานของผู้ประเมินมีผลเล็กน้อยเท่านั้น ส่วนองค์ประกอบอื่นที่เหลือไม่มีผลต่อการประเมินเลย

คุนนอน (Kunnan. 1992 : 30 – 47) ได้ทำการศึกษาวิธีวิเคราะห์ 3 ชนิด โดยแบ่งข้อสอบเป็น 3 ฉบับ คือ การฟัง การอ่าน และศัพท์ ไวยากรณ์ ฉบับละ 30 ข้อเท่า ๆ กันทุกฉบับเป็นแบบทดสอบชนิดเลือกตอบ และแบ่งกลุ่มตัวอย่างออกเป็น 4 กลุ่มตามความสามารถในการสอบ และการคำนวณหาความเชื่อมั่นโดยทฤษฎีสรุปอ้างอิง โดยใช้โปรแกรมสำเร็จรูป GENOVA ซึ่งออกแบบแผนเป็น $p \times (i : t)$ คือ ข้อสอบที่แบ่งอยู่ในแบบทดสอบซึ่งเป็นโดเมนผสม และการ

คำนวณความเที่ยงตรงโดยใช้วิเคราะห์กลุ่มใช้โปรแกรมสำเร็จรูป SPSS – X โดยแบ่งแบบทดสอบแต่ละฉบับออกเป็น 2 ฉบับ ผลของการวิจัยพบว่า ความเชื่อมั่นของแบบทดสอบแตกต่างกันทุก ๆ กลุ่ม และค่าจะสูงในกลุ่มที่มีความสามารถต่ำได้ดีกว่ากลุ่มที่มีความสามารถสูง

เฟอร์นันเดซ และเมธัว (Fernandez and Meteo. 1993 : 425 – 435) ได้ศึกษาความเที่ยงตรงเชิงโครงสร้าง โดยการวิเคราะห์องค์ประกอบของแบบทดสอบการประเมินวิชาการ (The Academic Setting Evaluation Questionnaire : ASEQ) ที่ได้สร้างขึ้น กลุ่มตัวอย่างที่ใช้เป็นนิสิตระดับมหาวิทยาลัยในประเทศสเปน จำนวน 748 คน ผลการศึกษาพบว่า แบบทดสอบ ASEQ จะวัดใน 3 องค์ประกอบ และมีความเที่ยงตรงตามองค์ประกอบดังนี้ 1) ด้านความพอใจกับสภาวะในการทำงาน มีน้ำหนักองค์ประกอบตั้งแต่ .32 - .71 2) ด้านสภาวะสิ่งแวดล้อมมีน้ำหนักองค์ประกอบตั้งแต่ .64 - .80 3) ด้านความสัมพันธ์กับเพื่อน มีน้ำหนักองค์ประกอบตั้งแต่ .31 - .62

สเตค (Stake. 1994 : 56 – 72) ได้ศึกษาความเที่ยงตรงของแบบวัดมโนภาพแห่งตนด้านการปฏิบัติ (The Performance Self – Concept Scale : PSES) ที่ได้สร้างไว้ในปี 1977 โดยการวิเคราะห์องค์ประกอบจากกลุ่มตัวอย่างที่เป็นผู้ใหญ่จำนวน 365 คน เป็นผู้ชาย 96 คน ผู้หญิง 269 คน ผลจากการวิเคราะห์องค์ประกอบพบว่า แบบวัดมโนภาพแห่งตน ด้านการปฏิบัติวัดมโนภาพแห่งตนใน 6 ด้าน ซึ่งในแต่ละด้านมีความเที่ยงตรงตามองค์ประกอบดังนี้ 1) ความชื่นชมมีน้ำหนักองค์ประกอบตั้งแต่ .62 - .79 2) ความมีศีลธรรม มีน้ำหนักองค์ประกอบตั้งแต่ .59 - .79 3) ความสามารถ มีน้ำหนักองค์ประกอบตั้งแต่ .48 - .72 4) ความฉลาด มีน้ำหนักองค์ประกอบตั้งแต่ .42 - .82 5) ความมีอำนาจ มีน้ำหนักองค์ประกอบตั้งแต่ .56 - .71 6) ความอ่อนแอ มีน้ำหนักองค์ประกอบตั้งแต่ .51 - .67

ทอมสัน (Thompson. 1996 : 197 – 208) ได้ศึกษาความเที่ยงตรงเชิงโครงสร้างโดยการวิเคราะห์องค์ประกอบ ผลจากการวิเคราะห์องค์ประกอบพบว่า เป็นความเที่ยงตรงเชิงโครงสร้างอย่างมีเหตุผล ซึ่งได้นำมาอธิบายสิ่งที่ปรากฏออกมาจากเกณฑ์ที่กำหนดได้อย่างมีเหตุผล การฝึกฝนทักษะต่าง ๆ นั้นก็เป็นส่วนหนึ่งในการพิจารณาการตรวจสอบความเที่ยงตรงเชิงโครงสร้าง เป็นส่วนประกอบหลักของการวิเคราะห์และรับรองความเที่ยงตรง โดยใช้โปรแกรมคอมพิวเตอร์ช่วยในการวิเคราะห์องค์ประกอบ

ไพร์ตัน วังษ์นาม (2533. 105 – 123) ได้ศึกษาเปรียบเทียบผลของวิธีตรวจ การชี้แจง และการรู้ผลการเรียนของผู้ตอบที่มีต่อสัมประสิทธิ์การสรุปอ้างอิง (Generalization Coefficient) ของการสอบด้วยแบบทดสอบความเรียง ที่วัดความสามารถในการแสดงความคิดเห็นเกี่ยวกับข่าวและเหตุการณ์ ของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 30 คน ข้อสอบความเรียง 5 ข้อ และครูผู้ตรวจให้คะแนนจำนวน 20 คน สุ่มแบ่งผู้ตรวจให้คะแนนเป็น 2 กลุ่ม กลุ่มละ 10 คน สำหรับ

การตรวจ 2 วิธี คือวิธีประเมินรวมและวิธีวิเคราะห์ ในแต่ละวิธีการทำการตรวจ 3 ครั้ง คือ ครั้งแรกตรวจโดยใช้ประสบการณ์เดิมของผู้ตรวจให้คะแนน ครั้งที่สองตรวจตามเกณฑ์ที่ได้รับการชี้แจงจากผู้วิจัยโดยไม่รู้ผลการเรียนของผู้ตอบ ครั้งที่สาม ตรวจตามเกณฑ์ที่ได้รับการชี้แจงจากผู้วิจัยเหมือนครั้งที่สอง แต่รู้ผลการเรียนของผู้ตอบด้วย พบว่า เมื่ออ้างอิงสรุปไปยังเอกภาพของข้อสอบและผู้ตรวจให้คะแนนพร้อมกัน ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบความเรียงที่ตรวจโดยวิธีประเมินรวม มีค่าอยู่ระหว่าง 0.3328 – 0.4782 และวิธีวิเคราะห์มีค่าอยู่ระหว่าง 0.3348 – 0.5895 และใช้ข้อสอบ 5 ข้อ ผู้ตรวจให้คะแนนอย่างน้อย 5 คน ใช้วิธีการตรวจวิธีวิเคราะห์ตามที่ได้รับการชี้แจง และไม่รู้ผลการเรียนของผู้ตอบ จะให้ค่าสัมประสิทธิ์การสรุปอ้างอิงไม่น้อยกว่า 0.50 เมื่ออ้างอิงสรุปไปยังเอกภาพของผู้ตรวจให้คะแนนอย่างเดียว ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบความเรียงที่ตรวจโดยวิธีประเมินรวม มีค่าอยู่ระหว่าง 0.743 – 0.6865 และวิธีวิเคราะห์มีค่าอยู่ระหว่าง 0.5985 – 0.7761 และใช้ข้อสอบ 5 ข้อ ผู้ตรวจให้คะแนนอย่างน้อย 5 คน ใช้วิธีการตรวจวิธีวิเคราะห์ตามที่ได้รับการชี้แจง และไม่รู้ผลการเรียนของผู้ตอบ จะให้ค่าสัมประสิทธิ์การสรุปอ้างอิงไม่น้อยกว่า 0.50 เมื่ออ้างอิงสรุปไปยังเอกภาพของผู้ตรวจให้คะแนนอย่างเดียว พบว่า เฉพาะกรณีที่ผู้ตรวจให้คะแนนได้รับการชี้แจงและไม่รู้ผลการเรียนของผู้ตอบเท่านั้นที่ให้ค่าสัมประสิทธิ์การสรุปอ้างอิงสูงกว่าการตรวจโดยใช้ประสบการณ์เดิม ถ้ารู้ผลการเรียนของผู้ตอบค่าสัมประสิทธิ์การสรุปอ้างอิงจะไม่แตกต่างกัน

ชาญ วชิรเดช (2538) ศึกษารูปแบบของแบบทดสอบการแก้ปัญหา ที่สอบด้วยไมโครคอมพิวเตอร์ 3 รูปแบบ คือ แบบทดสอบความสามารถในการแก้ปัญหาที่อาศัยหลักการเพียงอย่างเดียว ประเมิน 4 ตัวเลือก แบบทดสอบความสามารถในการแก้ปัญหาที่อาศัยหลักการและข้อเท็จจริง ประเมิน 4 ตัวเลือก แบบทดสอบความสามารถในการแก้ปัญหาที่อาศัยหลักการเพียงอย่างเดียว ประเมินเลือกตอบหลายตัวเลือก แบบทดสอบแต่ละรูปแบบประกอบด้วยเงื่อนไขสถานการณ์ของโจทย์ปัญหา 6 เงื่อนไข แต่ละเงื่อนไขประกอบด้วยคำถามย่อย 5 ข้อ เป็นข้อสอบทั้งสิ้น 30 ข้อ พบว่า ความแปรปรวน ของผลรวมระหว่างนักเรียนทุกคน ทำข้อสอบทุกข้อ ที่แฝงอยู่ในเงื่อนไขสถานการณ์ของโจทย์ที่กำหนด มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบการแก้ปัญหาที่สอบด้วยไมโครคอมพิวเตอร์ทั้งสามรูปแบบ สัมประสิทธิ์การสรุปอ้างอิง ของแบบทดสอบทั้งสามรูปแบบแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบที่กำหนดเงื่อนไขสถานการณ์ของโจทย์ปัญหาที่กำหนด ตั้งแต่ 1 เงื่อนไขโดยใช้จำนวนข้อสอบ 1 – 10 ข้อ แบบทดสอบความสามารถในการแก้ปัญหาที่อาศัยหลักการและข้อเท็จจริงชนิดประเมิน 4 ตัวเลือก จะมีค่าสัมประสิทธิ์การสรุปอ้างอิงสูงกว่ารูปแบบอื่นๆ เงื่อนไข และค่าสัมประสิทธิ์การสรุปอ้างอิงของ

แบบทดสอบทั้งสามรูปแบบมีค่าสัมประสิทธิ์การสุปร้องประมาณใกล้เคียงกับค่าพารามิเตอร์ของความเชื่อมั่นที่แท้จริง

มิ่ง เทพครเมือง (2539: บทคัดย่อ) ได้ศึกษาเกี่ยวกับการตรวจสอบหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบการคิดเอกนัยด้านสัญลักษณ์ 6 ฉบับ ตามทฤษฎีของกิลฟอร์ด กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2538 ของโรงเรียนในสังกัดกรมสามัญศึกษาจังหวัดเชียงราย จำนวน 586 คน ผลการศึกษาพบว่า แบบทดสอบวัดความสามารถในการคิดเอกนัยทางสัญลักษณ์ทั้ง 6 ฉบับ มีค่าความเชื่อมั่นอยู่ระหว่าง .742 ถึง .8144 และค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบที่ได้จากการวิเคราะห์องค์ประกอบเชิงยืนยันแบบความน่าจะเป็นสูงสุดพบว่า แบบทดสอบแต่ละฉบับมีความเที่ยงตรงในการวัดองค์ประกอบการคิดเอกนัยทางสัญลักษณ์แบบหน่วย แบบกลุ่ม แบบความสัมพันธ์ แบบระบบ แบบการแปลงรูป และแบบการประยุกต์โดยมีค่าน้ำหนักองค์ประกอบที่มีนัยสำคัญทางสถิติที่ระดับ .05 แบบทดสอบทั้ง 6 ฉบับมีความเที่ยงตรงในการวัดองค์ประกอบการคิดเอกนัยทางสัญลักษณ์ร่วมกัน โดยมีค่าน้ำหนักองค์ประกอบที่มีนัยสำคัญทางสถิติที่ระดับ .05

นพดล กองศิลป์ (2540: บทคัดย่อ) ได้ศึกษาเกี่ยวกับการแสดงหลักฐานของความเที่ยงตรงเชิงโครงสร้างและค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ ของแบบทดสอบวัดความรู้ความเข้าใจทางภาษาแบบประยุกต์ตามโครงสร้างทางสติปัญญาของกิลฟอร์ด โดยใช้วิธีการวิเคราะห์เชิงสำรวจและวิธีวิเคราะห์องค์ประกอบเชิงยืนยันด้วยวิธีความน่าจะเป็นสูงสุด ตามวิธีการของโจเรสกอกและซอร์บอม (Joreskog and Sorbom: 1989) และคำนวณค่าความเชื่อมั่นด้วยสูตรของเฟลด์ต – ราจู (Feldt – Raju) กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 900 คน ผลการศึกษาความเที่ยงตรงเชิงโครงสร้าง ของแบบทดสอบวัดความรู้ความเข้าใจทางภาษาประยุกต์ทั้ง 5 ฉบับ คือ การให้ทางเลือก การทดสอบเครื่องมือเครื่องใช้ การตั้งคำถามให้ตรงกับเนื้อเรื่อง การมองเห็นความบกพร่อง และการมองเห็นปัญหา โดยใช้การวิเคราะห์องค์ประกอบเชิงสำรวจ พบว่า แบบทดสอบแต่ละฉบับวัดองค์ประกอบร่วมกัน 1 องค์ประกอบโดยมีน้ำหนักร่วมกันอยู่ระหว่าง .548 ถึง .713 ความเที่ยงตรงเชิงโครงสร้างที่ได้จากการวิเคราะห์ องค์ประกอบเชิงยืนยันด้วยวิธีความน่าจะเป็นสูงสุด พบว่าแบบทดสอบทั้ง 5 ฉบับ มีความเที่ยงตรงในการวัดองค์ประกอบความรู้ความเข้าใจทางภาษาแบบประยุกต์ โดยมีน้ำหนักองค์ประกอบที่มีนัยสำคัญทางสถิติที่ระดับ .01 และค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์มีค่าอยู่ระหว่าง .784 ถึง .887

สนอง ตรงเที่ยง (2540: 61–63) ได้ศึกษาความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบวัดการประเมินค่าทางสัญลักษณ์แบบการแปลงรูป และแบบทดสอบวัดการประเมินค่าทางภาษาแบบการแปลงรูปตามทฤษฎีโครงสร้างทางปัญญาของกิลฟอร์ด กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษา

ปีที่ 1 ของโรงเรียนสังกัดกรมสามัญศึกษาในจังหวัดนนทบุรี จำนวน 610 คน ผลการศึกษาพบว่า ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบจากการวิเคราะห์เชิงสำรวจพบว่า แบบทดสอบทั้งสองฉบับวัดองค์ประกอบร่วมกัน คือ องค์ประกอบการประเมินค่าทางสัญลักษณ์แบบการแปลงรูป และองค์ประกอบการประเมินค่าทางภาษาแบบการแปลงรูป สำหรับการวิเคราะห์เชิงยืนยันด้วยวิธีความน่าจะเป็นสูงสุด พบว่า แบบทดสอบฉบับถอดรหัส ฉบับการสลับที่ตัวอักษร และฉบับการตัดสินความถูกต้องของนิพจน์ วัดองค์ประกอบด้านการประเมินค่าทางสัญลักษณ์แบบการแปลงรูปอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และแบบทดสอบฉบับตัดสินการปรับใช้ประโยชน์ของสิ่งของ และฉบับการเปรียบเทียบประโยคที่ใช้บรรยายภาพ วัดองค์ประกอบด้านการประเมินค่าทางภาษาแบบการแปลงรูปอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แบบทดสอบการประเมินค่าทางสัญลักษณ์แบบการแปลงรูปจำนวน 3 ฉบับ มีค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ คือ .740 .749 และ .763 ตามลำดับ แบบทดสอบการประเมินค่าทางภาษาแบบการแปลงรูป จำนวน 2 ฉบับ มีค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ คือ .717 และ .742 ตามลำดับ

อุษณีย์ บัวศิริพันธุ์ (2543) ศึกษาเปรียบเทียบ ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจ จำนวนผู้ตรวจให้คะแนน และประสบการณ์ของผู้ตรวจให้คะแนน แตกต่างกัน กรณีผลการศึกษาเกี่ยวกับวิธีการตรวจให้คะแนน การตรวจแบบประเมินรวมโดยใช้เกณฑ์การให้คะแนนของผู้ตรวจให้คะแนนเอง เมื่อมีจำนวนผู้ตรวจให้คะแนน 3, 4, 5 และ 6 คน มีคะแนนเฉลี่ยอยู่ระหว่าง 8.79 – 9.05 โดยการตรวจให้คะแนนเมื่อมีจำนวนผู้ตรวจตั้งแต่ 4 คน คะแนนเฉลี่ยมีแนวโน้มสูงที่สุด รองลงมา คือ มีจำนวนผู้ตรวจให้คะแนน 5 คน และเมื่อมีจำนวนผู้ตรวจให้คะแนน 6 คน คะแนนเฉลี่ยมีแนวโน้มต่ำที่สุด ส่วนการตรวจให้คะแนนแบบประเมินรวมโดยใช้เกณฑ์การให้คะแนนแบบรูปคี่ เมื่อมีจำนวนผู้ตรวจให้คะแนน 3, 4, 5 และ 6 มีคะแนนเฉลี่ยอยู่ระหว่าง 9.82 – 10.05 โดยการตรวจให้คะแนนเมื่อมีจำนวนผู้ตรวจให้คะแนน 4 คน คะแนนเฉลี่ยมีแนวโน้มสูงที่สุด รองลงมาคือ จำนวนผู้ตรวจให้คะแนน 3 และ 6 คน และเมื่อมีจำนวนผู้ตรวจให้คะแนน 5 คน คะแนนเฉลี่ยมีแนวโน้มต่ำที่สุด

7.2 งานวิจัยที่เกี่ยวข้องกับทักษะกระบวนการทางวิทยาศาสตร์

โรบินสัน (Robinson. 1973 : 1522 – A) ได้พัฒนาแบบทดสอบวัดทักษะด้านการควบคุมตัวแปร (Controlling Variable) และการแปลความหมายจากข้อมูล (Interpreting Data) แบบทดสอบสองฉบับนี้ เป็นแบบปรนัยชนิดเลือกตอบหรือจากการได้ทำการวิเคราะห์เป็นรายข้อ ได้นำข้อสอบที่มีค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป ค่าสัมประสิทธิ์สหสัมพันธ์กับคะแนนที่ได้จากการสอบเป็นรายบุคคลได้ค่าสหสัมพันธ์ .66 แบบทดสอบอีกฉบับหนึ่ง

พีเรซ (Perez. 1978 : 3496 – A) สร้างขึ้นสำหรับใช้ในประเทศฟิลิปปินส์ ลักษณะของแบบทดสอบประกอบด้วย ทักษะการสังเกต 11 ข้อ การเปรียบเทียบ 6 ข้อ การหาจำนวน 8 ข้อ การจำแนกประเภท 11 ข้อ การลงข้อสรุป 7 ข้อ การทำนาย 5 ข้อ และการทดลอง 11 ข้อ เมื่อคำนวณค่าความเชื่อมั่นทั้งฉบับด้วยสูตร KR – 20 ได้เท่ากับ .87 ค่าความยากง่ายเฉลี่ยเท่ากับ .39 และข้อสอบมีค่าอำนาจจำแนกที่มีประสิทธิภาพในการนำไปใช้มีจำนวน 48 ข้อ เป็นแบบทดสอบที่พัฒนาขึ้นให้มีความสัมพันธ์กับเกณฑ์ภายนอก คือ The Individual Competency Measures อยู่ในระดับสูงและแบบทดสอบอีกฉบับหนึ่งมีความเชื่อมั่นสูง สามารถนำไปใช้กับ เกณฑ์การทำวิจัยเกี่ยวกับทักษะกระบวนการทางวิทยาศาสตร์

เพททัส และฮาเลย์ (Pettus and Haley. 1980 : 273 – 277) ได้ศึกษาปัจจัยที่เกี่ยวข้องกับระดับการปฏิบัติด้านทักษะกระบวนการทางวิทยาศาสตร์ ของนักเรียนเกรด 9 ถึง 12 ในรัฐเวอร์จิเนีย ผลการศึกษาพบว่า นักเรียนมีทักษะกระบวนการทางวิทยาศาสตร์เพิ่มขึ้นตามระดับชั้นเรียนที่เพิ่มขึ้น และนักเรียนหญิงมีทักษะกระบวนการทางวิทยาศาสตร์สูงกว่านักเรียนชาย

รูบิน (Rubin. 1989 : 3469 – A) ได้ศึกษายุทธศาสตร์การสอนทำต้นแบบเป็นระบบเพื่อพัฒนาทักษะกระบวนการทางวิทยาศาสตร์แบบบูรณาการ และความสามารถในการใช้เหตุผลในเชิงพุทธิพิสัยแบบเป็นทางการ กลุ่มตัวอย่างเป็นนักเรียนระดับกลางในเมือง นักเรียนกลุ่มที่ 1 ได้รับการสอนทักษะกระบวนการทางวิทยาศาสตร์ในช่วงเวลา 3 เดือน จากครูซึ่งได้รับการฝึกกลยุทธ์การสอนแบบเชิงระบบทดลอง กลุ่มที่ 2 ได้รับการสอนจากครูที่ฝึกกลยุทธ์การควบคุม วงจรการเรียนรู้ และกระบวนการทางวิทยาศาสตร์ กลุ่มที่ 3 เป็นกลุ่มควบคุมได้รับการสอนวิทยาศาสตร์ตามแบบเดิม ผลการวิจัยพบว่านักเรียนในกลุ่มที่ 1 และกลุ่มที่ 2 มีผลสัมฤทธิ์ทางการเรียนกระบวนการทางวิทยาศาสตร์แตกต่างกัน อย่างมีนัยสำคัญทางสถิติ นักเรียนกลุ่มที่ 2 และ 3 มีผลสัมฤทธิ์ทางการเรียนทักษะกระบวนการทางวิทยาศาสตร์แตกต่างกัน อย่างมีนัยสำคัญทางสถิติ โดยสรุปกลยุทธ์การสอนการทำต้นแบบเป็นระบบ จัดเป็นวิธีการที่มีประสิทธิภาพในการพัฒนาความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์แบบบูรณาการของนักเรียน

รอย และรอยเชาดีฮูรี (ธิดา สอนงนารถ. 2542 : 33 อ้างอิงจาก Royh and Roychoudhury. 1993 : 127 – 152) ได้ศึกษาการพัฒนาทักษะกระบวนการทางวิทยาศาสตร์ในสิ่งแวดล้อมจริง กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษา จำนวน 157 คน ซึ่งแบ่งเป็นกลุ่มทดลอง 3 กลุ่ม กลุ่มแรกได้รับการสอนโดยวิธีชี้แนะแนวทางในวิชาฟิสิกส์ กลุ่มที่สองได้รับการสอนวิชาฟิสิกส์ในระดับสูง และกลุ่มที่สาม ได้รับการสอนในวิชาวิทยาศาสตร์ทั่วไป โดยนักเรียนทั้ง 3 กลุ่มจะเป็นศูนย์กลางในการเรียน โดยจะได้รับการฝึกให้ปฏิบัติการทดลองในห้องปฏิบัติการอย่างอิสระเหมือนกัน ผลการวิจัยพบว่า กลุ่มทดลองทั้ง 3 กลุ่ม มีพัฒนาทักษะกระบวนการทางวิทยาศาสตร์สูงขึ้น เช่น มี

ทักษะในตัวแปรต่าง ๆ เหมือนกัน คือทักษะการแปรความหมายข้อมูล ทักษะการตั้งสมมติฐาน ทักษะการให้คำนิยาม และทักษะการทดลอง

ปรีชา เจตินัย (2531 : 63 – 64) ได้เปรียบเทียบทักษะกระบวนการทางวิทยาศาสตร์ และเจตคติเชิงวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีรูปแบบการจัดชั้นเรียนและแผนการเรียนแตกต่างกันในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา เขตการศึกษา 11 ปีการศึกษา 2530 ในส่วนที่เกี่ยวกับทักษะกระบวนการทางวิทยาศาสตร์ พบว่า

1. นักเรียนที่มีแผนการเรียนแตกต่างกัน มีทักษะกระบวนการทางวิทยาศาสตร์แตกต่างกันอย่างมีนัยสำคัญทางสถิติ
2. นักเรียนที่มีรูปแบบการจัดชั้นเรียนต่างกัน มีทักษะกระบวนการทางวิทยาศาสตร์แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

วิวิรรณ อังคนุรักษ์พันธุ์ (2531 : บทคัดย่อ) ได้สร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนสาธิต สังกัดทบวงมหาวิทยาลัยของรัฐ จำนวน 1,095 คน ประกอบด้วยข้อคำถามที่วัดองค์ประกอบที่สำคัญของทักษะกระบวนการทางวิทยาศาสตร์ 13 ทักษะ คือ การสังเกต การวัด การจำแนกประเภท การหาความสัมพันธ์ระหว่างสเปกกับสเปส และสเปกกับเวลา การคำนวณ การจัดกระทำและสื่อความหมายข้อมูล การลงความคิดเห็นจากข้อมูล การพยากรณ์ การทดลอง การตั้งสมมติฐาน การกำหนดนิยามเชิงปฏิบัติการ การกำหนดและควบคุมตัวแปร การตีความหมายข้อมูลและลงข้อสรุป ผลการศึกษาพบว่า ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบมีค่าอยู่ในช่วง .21 ถึง .78 และ .20 ถึง .85 ตามลำดับ คำนวณค่าความเชื่อมั่นด้วยสูตร KR – 20 ได้ค่าความเชื่อมั่นของแบบทดสอบแต่ละทักษะมีค่าอยู่ในช่วง .2290 ถึง .5769 และรวมทั้งฉบับมีค่าเท่ากับ .8383 มีนัยสำคัญทางสถิติที่ระดับ .01 ค่าความเที่ยงตรงเชิงโครงสร้างมีค่าน้ำหนักองค์ประกอบอยู่ในช่วง .1972 ถึง .6202 มีนัยสำคัญทางสถิติที่ระดับ .01 และเกณฑ์ปกติของแบบทดสอบอยู่ในช่วง T15 ถึง T 87

มนวิภา อ่อนศรี (2541 : บทคัดย่อ) ศึกษาความสัมพันธ์ของทักษะกระบวนการทางวิทยาศาสตร์กับความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ กลุ่มตัวอย่างเป็นนักเรียนมัธยมศึกษา ปีที่ 3 โรงเรียนสังกัดกรมสามัญศึกษาศูนย์กลาง กลุ่มที่ 5 ปีการศึกษา 2540 จำนวน 552 คน โดยวิธีสุ่มแบบสองขั้นตอน เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลประกอบด้วยแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ด้านการสังเกต การวัด การจำแนกประเภท การหาความสัมพันธ์ระหว่างสเปกกับสเปส และสเปกกับเวลา การคำนวณ การจัดกระทำและสื่อความหมายข้อมูล การลงความเห็นจากข้อมูล การพยากรณ์ การตั้งสมมติฐาน การกำหนดนิยามเชิงปฏิบัติการ การกำหนดและควบคุมตัวแปร การทดลอง การตีความหมายข้อมูลและลงข้อสรุป

และแบบทดสอบวัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ผลการวิจัยพบว่า ทักษะกระบวนการทั้ง 13 ทักษะกับความสามารถในการแก้ปัญหาทางวิทยาศาสตร์มีความสัมพันธ์ทางบวกกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และทักษะกระบวนการทางวิทยาศาสตร์ด้านการจำแนกประเภทส่งผลต่อความสามารถในการแก้ปัญหาทางวิทยาศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนทักษะกระบวนการทางวิทยาศาสตร์ด้านการสังเกต การจัดกระทำและสื่อความหมายข้อมูล และด้านการตีความหมายข้อมูลและลงข้อสรุป ส่วนผลต่อความสามารถในการแก้ปัญหาทางวิทยาศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนทักษะกระบวนการทางวิทยาศาสตร์ด้านอื่น ๆ ส่งผลต่อความสามารถในการแก้ปัญหาทางวิทยาศาสตร์อย่างไม่มีนัยสำคัญทางสถิติ

ในการวัดประเมินผู้เรียนต้องมีเครื่องมือวัดที่มีคุณภาพ ซึ่งมีดัชนีบ่งชี้คุณภาพของเครื่องมือวัดที่สำคัญ ได้แก่ ดัชนีความง่าย ดัชนีอำนาจจำแนก ดัชนีความเที่ยงตรงและดัชนีความเชื่อมั่นของเครื่องมือวัดประเมิน สำหรับงานวิจัยครั้งนี้ ผู้วิจัยใช้วิธีการแสดงหลักฐานความเที่ยงตรงและความเชื่อมั่นของแบบวัดประเมินทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

บทที่ 3

วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. การกำหนดประชากรและการสุ่มกลุ่มตัวอย่าง
2. เครื่องมือที่ใช้ในการวิจัย
3. ตัวอย่างเครื่องมือที่ใช้ในการวิจัย
4. วิธีการดำเนินการเก็บรวบรวมข้อมูล
5. การจัดกระทำและการวิเคราะห์ข้อมูล
6. สถิติที่ใช้ในการวิเคราะห์ข้อมูล

การกำหนดประชากรและการสุ่มกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 ภาคเรียนที่ 1 โรงเรียนมัธยมศึกษาสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กลุ่มที่ 3 เขตพื้นที่การศึกษา 3 กรุงเทพมหานคร ปีการศึกษา 2549 มีทั้งสิ้น 11 โรงเรียน จำนวนนักเรียนทั้งหมด 9,317 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 (ทุกแผนการเรียน) ภาคเรียนที่ 1 โรงเรียนมัธยมศึกษาสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กลุ่มที่ 3 เขตพื้นที่การศึกษา 3 กรุงเทพมหานคร ปีการศึกษา 2549 จำนวน 4 โรงเรียน จำนวนนักเรียน 285 คน ซึ่งได้มาโดยการสุ่มแบบสองขั้นตอน (Two – Stage Random Sampling) มีขั้นตอนการสุ่มดังนี้

ขั้นที่ 1 ประมาณค่าขนาดของกลุ่มตัวอย่าง โดยใช้สูตรการกำหนดขนาดกลุ่มตัวอย่าง เพื่อประมาณค่าเฉลี่ยของประชากรที่ระดับความเชื่อมั่น .05 (มยุรี ศรีชัย. 2538 : 104) ได้ขนาดกลุ่มตัวอย่างจำนวน 285 คน

ขั้นที่ 2 สุ่มโรงเรียนโดยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) ใช้ขนาดโรงเรียนเป็นชั้น (Strata) ทั้งหมด 2 ขนาด ได้โรงเรียนขนาดละ 2 โรงเรียน รวมทั้งหมด 4 โรงเรียน

ขั้นที่ 3 สุ่มนักเรียนในแต่ละโรงเรียนตามสัดส่วนจำนวนนักเรียน ได้นักเรียน 285 คน ดังแสดงในตาราง 8

ตาราง 8 จำนวนนักเรียนที่ใช้เป็นกลุ่มตัวอย่างในการเก็บรวบรวมข้อมูล

ขนาดโรงเรียน	โรงเรียน	ประชากร (ม.5)	กลุ่มตัวอย่าง
ใหญ่พิเศษ	ชินรสวิทยาลัย	436	127
	สตรีวัดระฆัง	335	97
ใหญ่	สวนอนันต์	98	29
	มหารณพาราม	111	32
รวม		980	285

เครื่องมือที่ใช้ในการวิจัย

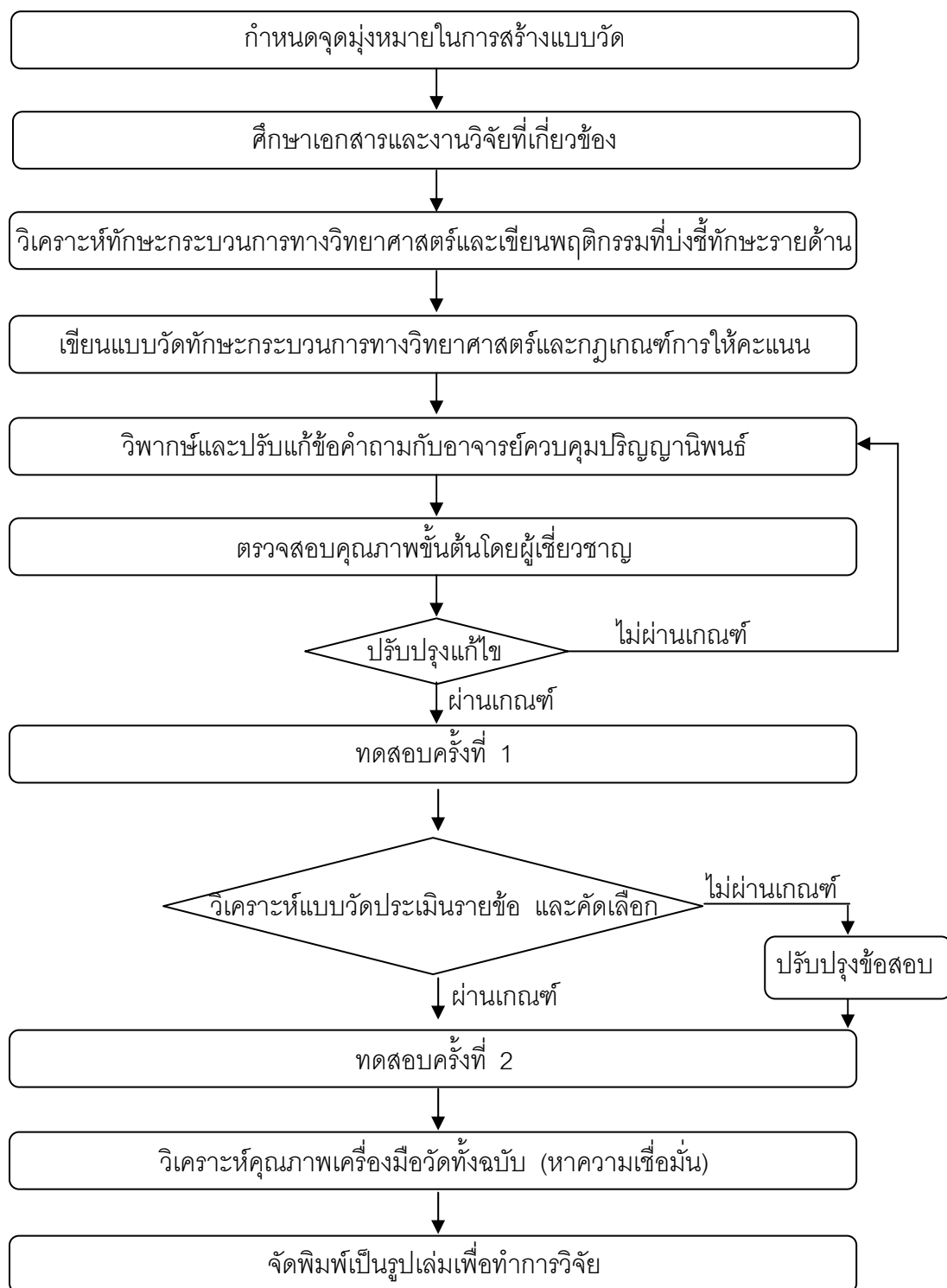
เครื่องมือที่ใช้ในการวิจัย เป็นแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ มีลักษณะเป็นคำถามปลายเปิดที่เกี่ยวกับฟิสิกส์พื้นฐาน เคมีพื้นฐาน และชีววิทยาพื้นฐาน ประกอบด้วย 5 ทักษะ คือ

- | | |
|---|-------------|
| 1. ทักษะการตั้งสมมติฐาน | จำนวน 6 ข้อ |
| 2. ทักษะการกำหนดนิยามเชิงปฏิบัติการ | จำนวน 6 ข้อ |
| 3. ทักษะการกำหนดและควบคุมตัวแปร | จำนวน 6 ข้อ |
| 4. ทักษะการทดลอง | จำนวน 6 ข้อ |
| 5. ทักษะการตีความหมายข้อมูลและลงข้อสรุป | จำนวน 6 ข้อ |

รวมเป็นแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ 1 ฉบับ จำนวน 5 ทักษะ จำนวนข้อสอบรวม 30 ข้อ

การสร้างเครื่องมือที่ใช้ในการวิจัย

ในการสร้างแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้



ภาพประกอบ 2 ลำดับขั้นตอนการสร้างแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

จากภาพประกอบแสดงลำดับขั้นตอนการสร้างแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้น
บูรณาการ มีรายละเอียดการดำเนินการสร้างตามขั้นตอนดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบวัด

1.1 เพื่อสร้างแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ 5 ทักษะ
ลักษณะเป็นคำถามปลายเปิด สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย

1.2 เพื่อนำคะแนนจากแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการไป
ใช้ในการศึกษาความเชื่อมั่นและความเที่ยงตรงที่มีสูตรต่างกัน

2. ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับความหมาย ประเภทของทักษะ
กระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ และการประเมินผลทักษะกระบวนการทางวิทยาศาสตร์

2.2 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับความหมาย วิธีการสร้าง และศึกษา
อิทธิพลที่มีผลต่อการวัดผลโดยใช้แบบทดสอบความเรียง

2.3 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องเกี่ยวกับทฤษฎีค่าความเชื่อมั่น ได้แก่ การ
ประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม การประมาณค่าความเชื่อมั่นที่
เกี่ยวกับการตรวจให้คะแนน การประมาณค่าความเชื่อมั่นตามโครงสร้างทฤษฎีการสรุปอ้างอิง การ
หาค่าความเชื่อมั่นของผู้ตรวจให้คะแนนที่มีจำนวนมากกว่า 2 คน ตามวิธีของดัชนีความสอดคล้อง
ของผู้ตรวจ (RAI)

2.4 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องเกี่ยวกับทฤษฎีค่าความเที่ยงตรง ได้แก่
ความหมายของความเที่ยงตรง ชนิดของหลักฐานแสดงความเที่ยงตรง วิธีการแสดงความเที่ยงตรง
และองค์ประกอบที่มีอิทธิพลต่อความเที่ยงตรง

3. วิเคราะห์ทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการและเขียนพฤติกรรมที่บ่งชี้
ทักษะรายด้าน โดยพฤติกรรมที่บ่งชี้ทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการย่อยทั้ง 5
ทักษะนั้นชาญวิทย์ จรัสสุทธิอิสร์ (2545 ; 82) ได้เสนอผลการวิเคราะห์การวัดทักษะกระบวนการทาง
วิทยาศาสตร์แต่ละทักษะดังแสดงในตารางที่ 9

ตาราง 9 ผลการวิเคราะห์พฤติกรรมที่บ่งชี้ทักษะกระบวนการทางวิทยาศาสตร์

ทักษะ	พฤติกรรมที่บ่งชี้
การตั้งสมมติฐาน	คาดคะเนคำตอบก่อนการทดลองได้สอดคล้องกับตัวแปรต้นและตัวแปรตามโดยอาศัยความรู้และประสบการณ์เดิม
การกำหนดนิยามเชิงปฏิบัติการ	การกำหนดความหมายและขอบเขตของคำหรือตัวแปรต่าง ๆ ในปฏิบัติการทดลองให้สามารถทดสอบหรือวัดได้
การกำหนดและควบคุมตัวแปร	การบ่งชี้และกำหนดตัวแปรต้น ตัวแปรตาม และควบคุมตัวแปรที่เป็นสาเหตุอื่น ๆ ในปรากฏการณ์หนึ่ง ๆ ที่ต้องการศึกษา
การทดลอง	การออกแบบการทดลอง ระบุวัสดุอุปกรณ์ที่ใช้ในการทดลอง อธิบายขั้นตอนการทดลอง และเสนอแนะการบันทึกผลการทดลอง
การตีความหมายและลงข้อสรุป	การแปลความหมายหรือบรรยายลักษณะและคุณสมบัติของข้อมูล พร้อมทั้งสรุปความสัมพันธ์ของตัวแปรที่อยู่ในข้อมูลที่ต้องการศึกษา

4. เขียนแบบวัดทักษะกระบวนการทางวิทยาศาสตร์และกฎเกณฑ์การให้คะแนน โดยใช้เนื้อหาวิชา ฟิสิกส์พื้นฐาน เคมีพื้นฐาน และชีววิทยาพื้นฐาน ซึ่งเขียนตามพฤติกรรมที่บ่งชี้ในตารางที่ 9 จำนวน 60 ข้อ เพื่อทดสอบและหาคุณภาพรายข้อ และคัดเลือกไว้ 30 ข้อ พร้อมทั้งจัดทำคู่มือการตรวจให้คะแนนแยกเป็นรายข้อเพื่อให้ผู้ตรวจใช้ในการให้คะแนนแต่ละข้อ

5. วิพากษ์และปรับแก้ข้อคำถามกับอาจารย์ควบคุมปริญญาโท

6. ตรวจสอบคุณภาพขั้นต้นโดยผู้เชี่ยวชาญ นำข้อคำถามที่ผู้วิจัยสร้างขึ้นไปตรวจสอบคุณภาพข้างต้น โดยพิจารณาความเที่ยงตรงเชิงเนื้อหา (Content Validity) และความเป็นปรนัย (Objectivity) ของข้อคำถามและคู่มือการตรวจให้คะแนน โดยผู้เชี่ยวชาญ 5 ท่านเป็นผู้พิจารณาว่าข้อคำถามนั้นวัดได้ตรงตามนิยามหรือไม่ พร้อมทั้งปรับปรุงภาษาให้เหมาะสมกับระดับของนักเรียน

7. ปรับปรุงแก้ไข ทำการปรับปรุงแก้ไขข้อคำถามตามคำแนะนำของผู้เชี่ยวชาญ โดยนำผลการพิจารณาแบบวัดประเมินทักษะกระบวนการทางวิทยาศาสตร์ของผู้เชี่ยวชาญ มาวิเคราะห์เพื่อหาความเที่ยงตรงเชิงเนื้อหาโดยใช้ดัชนีความสอดคล้อง (IOC) โดยรวมความเห็นในช่องเห็นด้วย (+1) ไม่แน่ใจ (0) และไม่เห็นด้วย (-1) ตามวิธีของ ไรวินลลีและแฮมเบิลตัน

(บุญเชิด ภิญญอนันตพงษ์. 2526) คัดเลือกข้อที่มีค่า IOC ตั้งแต่ 0.5 ขึ้นไป คัดเลือกไว้ 59 ข้อ ซึ่งถือว่าข้อคำถามนั้นสามารถวัดได้ตามนิยามที่กำหนด ปรากฏผลการคำนวณค่า IOC มีค่าระหว่าง 0.6 ถึง 1.0 ดังแสดงในภาคผนวก

8. ทดสอบครั้งที่ 1 นำแบบวัดประเมินที่ผ่านการตรวจสอบคุณภาพในขั้นต้นแล้วไป ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ ในการหาคุณภาพของเครื่องมือ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนนวมินทราชินูทิศ สตวิทย์วิทยา พุทธมณฑล จำนวน 90 คน โดยการให้นักเรียนทำข้อสอบทีละ ข้อเพื่อตรวจสอบเวลาที่เหมาะสมในการทำข้อสอบ โดยพิจารณาจากจำนวนนักเรียนที่ทำเสร็จ เรียบร้อยประมาณร้อยละ 80 ของนักเรียนที่เข้าสอบ

9. วิเคราะห์แบบทดสอบรายข้อ และคัดเลือก นำผลการทดสอบมาตรวจให้คะแนนโดย ผู้วิจัยเป็นผู้ดำเนินการเอง นำผลการตรวจให้คะแนนมาวิเคราะห์รายข้อเพื่อหาค่าความยาก โดยใช้ สัดส่วนจำนวนผู้ตอบ เพื่อคัดเลือกข้อสอบที่ดี หาค่าอำนาจจำแนกของข้อคำถามรายข้อ โดยใช้สูตร สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson Product Moment Correlation) โดยคัดเลือกไว้ตอน ละ 6 ข้อ จำนวน 30 ข้อ ได้ค่าความยากอยู่ระหว่าง 0.30 ถึง 0.78 และหาค่าอำนาจจำแนกเป็น รายข้อมีค่าอยู่ระหว่าง 0.21 ถึง 0.82 ได้ข้อสอบจำนวน 30 ข้อ ซึ่งค่าความเชื่อมั่นโดยใช้ สัมประสิทธิ์แอลฟา (Coefficient-Alpha) ของครอนบัค (Cronbach) มีค่าอยู่ระหว่าง .90 ถึง .91 และมีค่าความเชื่อมั่นรวมทั้งฉบับเท่ากับ .91

10. ทดสอบครั้งที่ 2 โดยนำแบบวัดประเมินทักษะกระบวนการทางวิทยาศาสตร์ที่ ปรับปรุงแก้ไขแล้วไปทดสอบกับกลุ่มตัวอย่างที่ใช้ ในการหาคุณภาพของเครื่องมือ คือ นักเรียนชั้น มัธยมศึกษาปีที่ 5 โรงเรียนนวมินทราชินูทิศ สตวิทย์วิทยา พุทธมณฑล จำนวน 90 คน

11. วิเคราะห์คุณภาพเครื่องมือวัดทั้งฉบับ โดยนำผลการทดสอบมาตรวจให้คะแนนตาม เกณฑ์การให้คะแนนที่กำหนดไว้ และวิเคราะห์หาคุณภาพของแบบวัดประเมินด้านความเชื่อมั่น ซึ่งมี ค่าความยากอยู่ระหว่าง 0.31 ถึง 0.78 ค่าอำนาจจำแนกเป็นรายข้อมีค่าอยู่ระหว่าง 0.23 ถึง 0.82 และมีค่าความเชื่อมั่นเท่ากับ .91

12. จัดพิมพ์เป็นรูปเล่มเพื่อทำการวิจัย เพื่อนำไปเก็บข้อมูล โดยนำแบบวัดประเมิน ทักษะกระบวนการทางวิทยาศาสตร์ที่มีคุณภาพ ไปเก็บข้อมูลเพื่อทำการวิจัยต่อไป

ตัวอย่างเครื่องมือที่ใช้ในการวิจัย

แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

ข้อ 1 ทักษะการตั้งสมมติฐาน

จงศึกษาข้อมูลจากตารางต่อไปนี้

การเคลื่อนที่ของแหล่งกำเนิดเสียง	ระดับเสียงที่ได้ยิน
1. เคลื่อนที่เข้าหาผู้ฟัง	สูง
2. เคลื่อนที่ออกห่างจากผู้ฟัง	ต่ำ

ข้อมูลในการทดลองจากตารางนี้ ควรตั้งสมมติฐานว่าอย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ถ้าแหล่งกำเนิดเสียงเคลื่อนที่เข้าหาผู้ฟังแล้ว จะได้ยินระดับเสียงจะสูงกว่าเคลื่อนที่ออกห่างจากผู้ฟัง
1	-ระดับเสียงสูง – ต่ำที่ได้ยิน ขึ้นอยู่กับระยะห่างระหว่างแหล่งกำเนิดเสียง กับผู้ฟัง -ระยะทางมีผลต่อความถี่
0	ไม่ตอบ หรือตอบไม่ถูก

ข้อ 2 ทักษะการกำหนดนิยามเชิงปฏิบัติการ

ในการทดลองเกี่ยวกับการเคลื่อนที่ของวัตถุในแนวตั้งภายใต้แรงโน้มถ่วงของโลก โดยติดลูกเทนนิสกับแถบกระดาษ แล้วปล่อยให้ลูกเทนนิสตกลงสู่พื้น พร้อมกับดึงแถบกระดาษผ่านเครื่องเคาะสัญญาณเวลา คำว่า “ภายใต้แรงโน้มถ่วงของโลก” จะกำหนดนิยามว่าอย่างไร

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	-ภายใต้แรงโน้มถ่วงของโลก หมายถึง การเคลื่อนที่ที่อยู่ในอำนาจแรงดึงดูดของโลก -ภายใต้แรงโน้มถ่วงของโลก หมายถึง แรงที่โลกกระทำกับวัตถุในทิศทางจากที่สูงลงสู่ที่ต่ำ
1	-ภายใต้แรงโน้มถ่วงของโลก หมายถึง การเคลื่อนที่จากที่สูงลงสู่ที่ต่ำ -ภายใต้แรงโน้มถ่วงของโลก หมายถึง แรงที่โลกกระทำต่อวัตถุ -ภายใต้แรงโน้มถ่วงของโลก หมายถึง การตกในแนวตั้ง , การเคลื่อนที่ที่วิ่งเข้าหาโลก , ตกลงมาโดยอิสระ
0	ไม่ตอบ หรือตอบไม่ถูก

ข้อ 3 ทักษะการกำหนดและควบคุมตัวแปร

ผู้ติemannำสารละลายกลูโคสใส่ในถุงเยื่อเซลโลเฟน แล้วนำถุงนี้ไปแช่ในน้ำกลั่น ทิ้งไว้ 20 นาที นำน้ำกลั่นที่ใช้แช่ถุง 5 cc ใส่ในหลอดทดลองขนาดกลางไปหยดด้วยสารละลายเบเนดิกต์ 3 หยด ต่อจากนั้นนำหลอดทดลองไปแช่ในน้ำเดือด 5 นาที สารละลายในหลอดแก้วจะเปลี่ยนเป็นสีส้มเหลือง จงระบุตัวแปรต้นและตัวแปรตาม

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ตัวแปรต้น คือ กลูโคสที่ผ่านเซลโลเฟนออกมา และตัวแปรตาม คือ การเปลี่ยนสีของสารละลาย
1	ตอบถูกบางส่วนของคำตอบ เช่น -ตัวแปรต้น คือ กลูโคสที่ผ่านเซลโลเฟนออกมา หรือ -ตัวแปรตาม คือ การเปลี่ยนสีของสารละลาย
0	ไม่ตอบ หรือตอบไม่ถูก

ข้อ 4 ทักษะการทดลอง

จงออกแบบการทดลองเพื่อยืนยันว่า วัตถุที่มีน้ำหนักต่างกัน ตกจากที่สูงเท่ากัน จะถึงพื้นพร้อมกัน

.....

.....

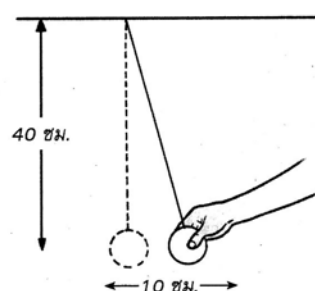
.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ทั้งวัตถุ 2 ชนิด ขนาดต่างกัน พร้อม ๆ กัน จากที่สูงเท่ากันตกในสุญญากาศ
1	ตอบถูกบางส่วน of คำตอบ เช่น -ทั้งวัตถุ 2 ชนิด ขนาดต่างกัน พร้อม ๆ กัน จากที่สูงเท่ากัน
0	ไม่ตอบ หรือตอบไม่ถูก

ข้อ 5 ทักษะการตีความหมายข้อมูลและลงข้อสรุป

1. แขนงลูกตุ้มลูกหนึ่งด้วยความยาวต่างกัน 2 ครั้ง แล้วดึงลูกตุ้มให้เกิดการแกว่ง โดยดึงให้ห่างจากแนวตั้งต่างกัน 2 ครั้ง จับเวลาเมื่อลูกตุ้มแกว่งครบรอบ 50 รอบ ได้ผลดังตาราง



ระยะทางที่ดึงลูกตุ้มให้ห่างจากแนวตั้ง (cm)	ลูกตุ้มยาว 60 cm		ลูกตุ้มยาว 40 cm	
	เวลาที่ลูกตุ้มแกว่งครบ 50 รอบ (s)	ความถี่ (Hz)	เวลาที่ลูกตุ้มแกว่งครบ 50 รอบ (s)	ความถี่ (Hz)
8 cm	80	0.625	65	0.769
10 cm	80	0.625	65	0.769

จากข้อมูลข้างต้น ลงข้อสรุปได้ว่าอย่างไร

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- เชือกผูกลูกตุ้มยาวจะมีความถี่น้อยกว่าเชือกผูกลูกตุ้มสั้น และระยะห่างจากแนวตั้ง ไม่มีผลต่อความถี่ - ความยาวของเชือกผูกลูกตุ้มมีผลต่อเวลาที่ใช้ในการแกว่งลูกตุ้มและความถี่ในการแกว่ง
1	ตอบถูกบางส่วนของคำตอบ เช่น - เชือกผูกลูกตุ้มยาวจะมีความถี่น้อยกว่าเชือกผูกลูกตุ้มสั้น หรือ - ระยะห่างจากแนวตั้ง ไม่มีผลต่อความถี่ หรือ - สิ่งที่มีผลต่อการเปลี่ยนแปลงความถี่ในการแกว่งคือความยาวเชือก
0	ไม่ตอบ หรือตอบไม่ถูก

วิธีการดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลในการวิจัย ดังนี้

1. นำหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อขอความอนุเคราะห์กับโรงเรียนที่เป็นกลุ่มตัวอย่าง
2. ติดต่อประสานงาน ขออนุญาตผู้บริหารและครูผู้สอนของโรงเรียนที่ใช้เป็นกลุ่มตัวอย่างเพื่อกำหนดวัน เวลา สถานที่ วิธีดำเนินการสอบ และจัดประชุมชี้แจงครูที่ได้รับการคัดเลือกให้มาตรวจให้คะแนนจำนวน 3 คน ในครั้งนี้เพื่อให้เกิดความเข้าใจอันดี
3. นำแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ผ่านการวิเคราะห์คุณภาพรายข้อแล้วไปทดสอบกับกลุ่มตัวอย่าง จำนวน 285 คน ซึ่งผู้ดำเนินการสอบชี้แจงให้กลุ่มตัวอย่าง และครูผู้ควบคุมการสอบให้ทราบถึงจุดมุ่งหมายของการวิจัย และความสำคัญของการวิจัยนี้ก่อนดำเนินการสอบ โดยนำไปเก็บรวบรวมข้อมูลระหว่างเดือนมกราคม พ.ศ.2550

การจัดกระทำและการวิเคราะห์ข้อมูล

1. นำกระดาษคำตอบของผู้เข้าสอบทั้งหมดมาถ่ายเอกสารจำนวน 3 ชุด จัดให้ผู้ตรวจจำนวน 3 คน ได้ตรวจให้คะแนนตามแนวทางในการตอบและกฎเกณฑ์การให้คะแนนที่ผู้วิจัยได้สร้างขึ้น ในคู่มือการตรวจให้คะแนน
2. นำผลที่ได้จากการตรวจให้คะแนนมาวิเคราะห์ข้อมูล ดังนี้
 - 2.1 คำนวณค่าสถิติพื้นฐานของแบบวัดประเมินจำแนกตามผู้ตรวจให้คะแนน

2.2 คำนวณค่าความเชื่อมั่นของแบบวัดประเมินโดยใช้

2.2.1 การหาค่าดัชนีความสอดคล้องของผู้ประเมิน โดยใช้ผู้ตรวจจำนวน 3 ท่าน ซึ่งมีประสบการณ์ในการสอนวิทยาศาสตร์ 5 ปีขึ้นไป

2.2.2 การหาค่าความเชื่อมั่นของการตรวจให้คะแนน โดยผู้ตรวจ 3 ท่าน แต่ละท่านตรวจข้อสอบทุกข้อ ของนักเรียนทุกคน

2.3 คำนวณค่าความเที่ยงตรงของแบบวัดประเมินโดยใช้

2.3.1 สูตร $\rho_{T_k S_k}$ ของ ไชยน์ และบอร์นสเตดท์

2.3.2 การวิเคราะห์องค์ประกอบเชิงยืนยัน ซึ่งผู้วิจัยแบ่งการวิเคราะห์ออกเป็น 2 อันดับ คือ การวิเคราะห์อันดับแรก (First Order) เป็นการวิเคราะห์เพื่อแสดงหลักฐานว่าข้อคำถามที่วัดในแต่ละทักษะมีความเที่ยงตรงมากน้อยเพียงใด และการวิเคราะห์อันดับที่สอง (Second Order) เป็นการวิเคราะห์เพื่อแสดงหลักฐานทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการมีความเที่ยงตรงมากน้อยเพียงใด

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. สถิติพื้นฐานของแบบทดสอบ ได้แก่ คะแนนเฉลี่ย ความเบี่ยงเบนมาตรฐาน และสัมประสิทธิ์ความแปรผัน (Coefficient of Variation)
2. สถิติที่ใช้ในการกำหนดขนาดกลุ่มตัวอย่าง เพื่อกะประมาณค่าเฉลี่ยของประชากรด้วยความเชื่อมั่น 95 เปอร์เซนต์ ใช้สูตรของมยูรี ศรีชัย (มยูรี ศรีชัย. 2538 : 104)

$$n = \frac{\sum_{g=1}^K \frac{N_g^2 \sigma_g^2}{W_g}}{\frac{N^2 e^2}{4} + \sum_{g=1}^K N_g \sigma_g^2}$$

เมื่อ	n	แทน	ขนาดกลุ่มตัวอย่างทั้งหมด
	N	แทน	ขนาดประชากร
	e	แทน	ความคลาดเคลื่อนที่ยอมรับให้เกิดได้
	W_g	แทน	สัดส่วนของประชากรในชั้นนั้น
	N_g	แทน	ขนาดของประชากรในชั้นนั้น
	σ_g	แทน	ส่วนเบี่ยงเบนมาตรฐานของชั้นนั้น

3. สถิติที่ใช้ในการหาคุณภาพเครื่องมือ

3.1 หาค่าความยากของแบบวัดประเมินทักษะกระบวนการทางวิทยาศาสตร์ชั้น

บูรณาการ โดยใช้การวิเคราะห์โดยใช้สัดส่วนจำนวนผู้ตอบ (บุญเชิด ภิญโญนนต์พงษ์. 2526: 89)

$$P = \frac{R}{N}$$

เมื่อ	P	แทน	ค่าความยาก
	R	แทน	จำนวนคนที่ตอบข้อสอบข้อนั้นถูก
	N	แทน	จำนวนคนที่ทำข้อสอบข้อนั้นทั้งหมด

3.2 หาค่าความเที่ยงตรงเชิงพินิจ (IOC) ของแบบวัดประเมินทักษะกระบวนการทาง

วิทยาศาสตร์ชั้นบูรณาการ จากค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (บุญเชิด ภิญโญนนต์พงษ์. 2547 : 179)

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

เมื่อ	IOC	แทน	ดัชนีความสอดคล้อง
	$\sum R$	แทน	ผลรวมของคะแนนความคิดเห็นของผู้เชี่ยวชาญทั้งหมด
	N	แทน	จำนวนผู้เชี่ยวชาญ

3.3 ค่าอำนาจจำแนกรายของข้อสอบรายข้อ โดยใช้สูตร Pearson Product Moment

Correlation : r_{xy} (บุญเชิด ภิญโญนนต์พงษ์. 2545 : 84) สูตรที่ใช้คือ

$$r_{xy} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

เมื่อ	r_{xy}	แทน	สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน หรือ ดัชนีอำนาจจำแนกของข้อสอบรายข้อ
	N	แทน	จำนวนกลุ่มตัวอย่าง
	$\sum X$	แทน	ผลรวมของคะแนนรายข้อ
	$\sum Y$	แทน	ผลรวมของคะแนนรวมที่ไม่รวมข้อที่นำมาคำนวณ
	$\sum X$	แทน	ผลรวมของผลคูณของคะแนน X และคะแนน Y
	$\sum X^2$	แทน	ผลรวมกำลังสองของคะแนน X
	$\sum Y^2$	แทน	ผลรวมกำลังสองของคะแนน Y

3.4 การแสดงหลักฐานความเชื่อมั่น โดยใช้การวิเคราะห์ตามแบบคะแนนจริงสมมูล โดยใช้สูตรสัมประสิทธิ์แอลฟา (Coefficient – Alpha) ของครอนบัค (Cronbach) (บุญเชิด ภิญโญนนันทพงษ์. 2547: 220)

$$\alpha = \frac{K}{K-1} \left[1 - \frac{\sum S_i^2}{S_x^2} \right]$$

เมื่อ	α	แทน	ค่าความเชื่อมั่นของเครื่องมือวัด
	K	แทน	จำนวนส่วนย่อยหรือจำนวนข้อคำถามของเครื่องมือวัด
	S_i^2	แทน	ค่าความแปรปรวนของคะแนนแต่ละส่วนย่อยหรือแต่ละข้อ
	S_x^2	แทน	ค่าความแปรปรวนของคะแนนรวมทั้งฉบับของเครื่องมือวัด

4. สถิติที่ใช้ในการวิจัย

4.1 การแสดงหลักฐานความเที่ยงตรง

4.1.1 ค่าความเที่ยงตรงเชิงโครงสร้างของแบบวัดประเมิน โดยใช้สูตรของ Heise and Bohrnstedt. (Heise and Bohrnstedt. 1970 : 116) สูตรที่ใช้คือ

$$\rho_{T_k S_k} = \frac{\sum_{i=1}^n S_i f_{ik}}{\sqrt{\sum_{i=1}^n \sum_{j=1}^n \text{Cov}(x_i, x_j)}}$$

โดย	$\rho_{T_k S_k}$	แทน	ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ
	S_i	แทน	ส่วนเบี่ยงเบนมาตรฐานของข้อคำถามข้อที่ i
	f_{ik}	แทน	น้ำหนักองค์ประกอบของคำถามข้อที่ i จากองค์ประกอบที่ k

$\text{Cov}(x_i, x_j)$ แทน ค่าความแปรปรวนร่วมระหว่างคำถามข้อที่ i กับคำถามข้อที่ j

4.1.2 ค่าความเที่ยงตรงเชิงโครงสร้างของแบบวัดประเมิน โดยใช้ Confirmatory Factor Analysis ซึ่งใช้โปรแกรมลิสเรล (LISREL 8.54)

4.2 การแสดงหลักฐานความเชื่อมั่น

4.2.1 ค่าความเชื่อมั่น (Reliability) ของผู้ตรวจแบบวัดประเมินทักษะ

กระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ คือ ผู้ตรวจให้คะแนน M คน K พฤติกรรม N วิชา โดยใช้สูตร RAI ของ Burry – Stock (Burry – Stock and other. 1996 : 259) สูตรที่ใช้คือ

$$RAI = 1 - \frac{\sum_{K=1}^K \sum_{N=1}^N \sum_{M=1}^M |R_{mnk} - \bar{R}_{nk}|}{KN(M-1)(I-1)}$$

โดยที่ $\bar{R}_{nk} = \frac{1}{M} \sum_{M=1}^M R_{mnk}$

เมื่อ	RAI	แทน	ดัชนีความสอดคล้องของผู้ประเมิน
	R	แทน	ผลการตรวจของผู้ตรวจให้คะแนน
	M	แทน	จำนวนผู้ตรวจให้คะแนนทั้งหมด
	N	แทน	จำนวนนักเรียนในกลุ่มตัวอย่าง
	K	แทน	จำนวนพฤติกรรมย่อย
	I	แทน	จำนวนช่วงคะแนน (5, 4, 3, 2, 1, 0)

4.2.2 การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง ($E\rho^2$) โดยใช้สูตรของ

Brennan. (Brennan. 1983 : 60) สูตรที่ใช้คือ

$$E\rho^2 = \frac{\sigma^2(\rho)}{\sigma^2(\rho) + \sigma^2(\delta)}$$

เมื่อ	$E\rho^2$	แทน	ค่าสัมประสิทธิ์การสรุปอ้างอิง
	$\sigma^2(\rho)$	แทน	ความแปรปรวนของคะแนนเอกภาพ
	$\sigma^2(\delta)$	แทน	ความแปรปรวนของคะแนนความคลาดเคลื่อน

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การเสนอผลการวิเคราะห์ข้อมูล ผู้วิจัยดำเนินการตามลำดับดังนี้

สัญลักษณ์และอักษรย่อในการวิเคราะห์ข้อมูล

ในการนำเสนอผลการวิเคราะห์ข้อมูลในการวิจัยในครั้งนี้ เพื่อให้เกิดความเข้าใจเกี่ยวกับความหมายในการนำเสนอผลการวิเคราะห์ที่ตรงกัน ผู้วิจัยได้กำหนดสัญลักษณ์ และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังต่อไปนี้

$\bar{X}$	แทน	ค่าเฉลี่ย
SD	แทน	ค่าความเบี่ยงเบนมาตรฐาน
CV	แทน	สัมประสิทธิ์การกระจาย
P	แทน	ค่าความยากง่าย
r	แทน	ค่าอำนาจจำแนก
IOC	แทน	ค่าความเที่ยงตรงเชิงพินิจ
S_i	แทน	ส่วนเบี่ยงเบนมาตรฐานรายข้อ
f_{ik}	แทน	ค่าน้ำหนักองค์ประกอบของคำถามข้อที่ i จากองค์ประกอบที่ k
$Cov(x_i, x_j)$	แทน	ค่าความแปรปรวนร่วมกันระหว่างคำถามข้อที่ i กับคำถามข้อที่ k
χ^2	แทน	ค่าจาก χ^2 - distribution
RAI	แทน	ดัชนีความสอดคล้องของผู้ประเมิน
$\bar{R}_{nk}$	แทน	ค่าเฉลี่ยของความเชื่อมั่นของผู้ตรวจ
df	แทน	องศาความเป็นอิสระ
SS	แทน	ผลรวมของกำลังสองของค่าเบี่ยงเบน
MS	แทน	ค่าเฉลี่ยกำลังสองของค่าเบี่ยงเบน
σ^2	แทน	ค่าประมาณความแปรปรวน
$\sigma^2(\rho)$	แทน	ค่าความแปรปรวนของคะแนนเอกภพ
$\sigma^2(\delta)$	แทน	ค่าความแปรปรวนของคะแนนความคลาดเคลื่อน
$E\rho^2$	แทน	ค่าสัมประสิทธิ์การสรุปอ้างอิง

การนำเสนอผลการวิเคราะห์ข้อมูล

การศึกษาค้นคว้าในครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับขั้นตอนดังนี้

ตอนที่ 1 ค่าสถิติพื้นฐานของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

ตอนที่ 2 หลักฐานการแสดงความคุณภาพเครื่องมือ

2.1 การแสดงหลักฐานความเที่ยงตรงเชิงพินิจ

2.2 การแสดงหลักฐานคุณภาพรายข้อเบื้องต้น (การทดสอบครั้งที่ 1)

2.2.1 การแสดงหลักฐานค่าความยาก

2.2.2 การแสดงหลักฐานค่าอำนาจจำแนก

2.3 การแสดงหลักฐานคุณภาพเครื่องมือ (การทดสอบครั้งที่ 2)

2.2.1 การแสดงหลักฐานค่าความยาก

2.2.2 การแสดงหลักฐานค่าอำนาจจำแนก

2.2.3 การแสดงหลักฐานค่าความเชื่อมั่น

ตอนที่ 3 การแสดงหลักฐานคุณภาพเครื่องมือที่ศึกษาจากกลุ่มตัวอย่าง

3.1 การแสดงหลักฐานความเที่ยงตรงของแบบทดสอบ

3.1.1 วิเคราะห์ด้วยสูตรของของไฮเซน และบอร์นสเตดต์ (Heise and

Bohrnstedt)

3.1.2 วิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) โดยใช้

โปรแกรมสำเร็จลิสเรล (LISREL 8.54)

3.2 การแสดงหลักฐานความเชื่อมั่น

3.2.1 วิเคราะห์ความสอดคล้องของผู้ประเมิน (Rater Agreement Indexes : RAI)

3.2.2 วิเคราะห์ค่าความเชื่อมั่นของการตรวจให้คะแนนที่คำนวณโดยใช้ทฤษฎีการ

สรุปอ้างอิง (Generalizability Theory)

ผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ค่าสถิติพื้นฐานของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

การวิเคราะห์ข้อมูลในขั้นตอนนี้ ผู้วิจัยได้นำคะแนนจากแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการไปให้ผู้ตรวจ 3 คน ตรวจให้คะแนนผู้เข้าสอบทุกคนและตรวจทุกทักษะโดยใช้กฎเกณฑ์การให้คะแนน (Scoring Rubric) ที่ผู้วิจัยสร้างขึ้นเป็นคู่มือในการตรวจ แล้วนำผลมาวิเคราะห์หาค่าสถิติพื้นฐานของแบบวัด จำแนกผลตามผู้ตรวจ ปากฏผลดังตาราง 10

ตาราง 10 ค่าสถิติพื้นฐานของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ จำนวน 5
ทักษะ คะแนนเต็ม 60 คะแนน

ผู้ตรวจ	$\bar{X}$	S.D.	C.V.	r_{tt}	SE_M
คนที่ 1	32.57	13.68	2.38	.98	2.07
คนที่ 2	34.37	14.99	2.29	.99	1.46
คนที่ 3	34.85	15.47	2.25	.99	1.46

ผลการวิเคราะห์พบว่า สถิติพื้นฐานของคะแนนแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ มีคะแนนเฉลี่ยจำแนกตามผู้ตรวจคนที่ 1 ถึง 3 มีค่าอยู่ระหว่าง 32.57 ถึง 34.85 โดยคะแนนเฉลี่ยของผู้ตรวจคนที่ 3 มีค่ามากที่สุด คือ 34.85 และรองลงมาคือคะแนนเฉลี่ยของผู้ตรวจคนที่ 2 มีค่า 34.37 และคะแนนเฉลี่ยของผู้ตรวจคนที่ 1 มีค่าน้อยที่สุด คือ 32.57 เมื่อพิจารณาค่าส่วนเบี่ยงเบนมาตรฐาน พบว่ามีค่าอยู่ระหว่าง 13.68 ถึง 15.47 และค่าสัมประสิทธิ์ความแปรผัน (Coefficient of Variation) มีค่าอยู่ระหว่าง 2.25 ถึง 2.38

ส่วนค่าความเชื่อมั่นของแบบวัดทั้งฉบับ มีค่าอยู่ระหว่าง 0.98 ถึง 0.99 โดยคะแนนของผู้ตรวจคนที่ 3 มีค่าความเชื่อมั่นสูงสุด รองลงมาคือ คะแนนของผู้ตรวจคนที่ 2 และคะแนนของผู้ตรวจคนที่ 1 มีค่าความเชื่อมั่นต่ำสุด แสดงว่า แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ตรวจให้คะแนนโดยผู้ตรวจ 3 คน มีหลักฐานค่าความเชื่อมั่นอยู่ในเกณฑ์ดี

ตอนที่ 2 หลักฐานการแสดงคุณภาพเครื่องมือ

2.1 การแสดงหลักฐานความเที่ยงตรงเชิงพินิจ

การดำเนินการตอนนี้ เป็นการประเมินแบบทดสอบเบื้องต้น ผู้วิจัยได้นำแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการที่สร้างขึ้น แบ่งออกเป็น 5 ตอน จำนวน 60 ข้อ พร้อมทั้งคำนิยาม นำไปให้ผู้เชี่ยวชาญทางด้านวัดผลจำนวน 3 คน และทางด้านวิทยาศาสตร์จำนวน 2 คน เป็นผู้พิจารณาข้อสอบเป็นรายข้อว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการวัดได้ตรงตามคำนิยามหรือไม่ พร้อมทั้งให้ปรับปรุงแก้ไขภาษา ให้เหมาะสมกับกลุ่มนักเรียนที่ศึกษา นำผลการพิจารณาแบบวัดของผู้เชี่ยวชาญ มาวิเคราะห์เพื่อแสดงหลักฐานความเที่ยงตรงเชิงพินิจ (IOC) โดยใช้ดัชนีความสอดคล้อง แล้วคัดเลือกเฉพาะข้อที่มีค่าดัชนีความสอดคล้องตั้งแต่ .5 ขึ้นไป ผลการวิเคราะห์พบว่า แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการที่คัดเลือกแล้วมีค่าดัชนีความสอดคล้องอยู่ในช่วง .60 ถึง 1.00 จำนวน 59 ข้อ แสดงว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการมีหลักฐานความเที่ยงตรงเชิงพินิจ

2.2 การแสดงหลักฐานคุณภาพรายข้อเบื้องต้น (การทดสอบครั้งที่ 1)

2.2.1 การแสดงหลักฐานค่าความยาก

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำผลการทดสอบของแบบวัดจำนวน 59 ข้อ ที่ทดสอบกับนักเรียนจำนวน 90 คน นำมาตรวจให้คะแนน แล้วนำมาวิเคราะห์ค่าความยากรายข้อ ปรากฏผลว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้น บูรณาการ จำนวน 59 ข้อ เป็นข้อสอบที่ยาก ($p = .20 - .39$) จำนวน 18 ข้อ ค่อนข้างง่าย ($p = .51 - .60$) จำนวน 14 ข้อ ง่าย ($p = .61 - .80$) และง่ายมาก ($p = .81 - 1.00$) อย่างละ 10 ข้อ ค่อนข้างยาก (.40 - .49) จำนวน 6 ข้อ และง่ายพอเหมาะ ($p = .50$) จำนวน 1 ข้อ แสดงว่าข้อสอบส่วนใหญ่ยาก รองลงมาเป็นข้อสอบที่ค่อนข้างง่าย ที่เหลือนอกนั้นเป็นข้อสอบที่ง่าย ง่ายมาก และค่อนข้างยาก

2.2.2 การแสดงหลักฐานค่าอำนาจจำแนก

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำผลการทดสอบครั้งที่ 1 ของแบบวัดจำนวน 59 ข้อ ที่ทดสอบกับนักเรียนจำนวน 90 คน นำมาตรวจให้คะแนน แล้วนำมาวิเคราะห์ค่าอำนาจจำแนก รายข้อ ปรากฏผลว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้น บูรณาการมีจำนวนข้อที่ตัดออก ($r \leq .19$) ทั้งหมด 12 ข้อ มีจำนวนข้อที่คัดเลือกไว้ทั้งหมด 47 ข้อ โดยข้อที่จำแนกได้บ้าง ($.20 \leq r \leq .39$) จำนวน 25 ข้อ ข้อที่จำแนกได้ปานกลาง ($.40 \leq r \leq .59$) จำนวน 15 ข้อ ข้อที่จำแนกได้ดี ($.60 \leq r \leq .79$) จำนวน 6 ข้อ และจำแนกได้ดีมาก ($.80 \leq r$) จำนวน 1 ข้อ แสดงว่าข้อสอบส่วนใหญ่สามารถจำแนกได้บ้าง รองลงมาเป็นข้อสอบที่จำแนกได้ปานกลาง และที่เหลือเป็นข้อสอบที่จำแนกได้ดีและจำแนกได้ดีมาก

2.3 การแสดงหลักฐานคุณภาพเครื่องมือ (การทดสอบครั้งที่ 2)

2.3.1 การแสดงหลักฐานค่าความยาก

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำผลการทดสอบครั้งที่ 2 ของแบบวัดจำนวน 30 ข้อ ที่ได้ปรับปรุงข้อคำถามให้เหมาะสม นำไปทดสอบกับนักเรียนจำนวน 90 คน นำมาตรวจให้คะแนน แล้วนำมาวิเคราะห์ค่าความยากรายข้อ ปรากฏผลว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้น บูรณาการ จำนวน 30 ข้อ ที่ทดสอบกับนักเรียนจำนวน 90 คน มีจำนวนข้อที่คัดเลือกไว้จำนวน 30 ข้อ โดยข้อที่ยาก ($p = .20 - .39$) มีจำนวน 8 ข้อ ค่อนข้างยาก (.40 - .49) จำนวน 5 ข้อ ยากง่ายพอเหมาะ (.50) จำนวน 1 ข้อ ค่อนข้างง่าย (.51 - .60) จำนวน 8 ข้อ และข้อที่ง่าย (.61 - .80) จำนวน 8 ข้อ แสดงว่าความยากของข้อสอบมีระดับความยากที่กระจายตั้งแต่ยาก ค่อนข้างยาก ยากง่ายพอเหมาะ ค่อนข้างง่าย และง่าย

2.3.2 การแสดงหลักฐานค่าอำนาจจำแนก

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำผลการทดสอบครั้งที่ 2 ของแบบวัดจำนวน 30 ข้อ ที่ได้ปรับปรุงข้อคำถามให้เหมาะสม นำไปทดสอบกับนักเรียนจำนวน 90 คน นำมาตรวจให้คะแนน แล้วนำมาวิเคราะห์ค่าอำนาจจำแนก ปรากฏผลว่า แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้น บูรณาการ จำนวน 30 ข้อ มีจำนวนข้อที่คัดเลือกไว้จำนวน 30 ข้อ โดยข้อที่จำแนกได้บ้าง ($.20 \leq r \leq .39$) มีจำนวน 10 ข้อ ข้อที่จำแนกได้ปานกลาง ($.40 \leq r \leq .59$) มีจำนวน 13 ข้อ ข้อที่จำแนกได้ดี ($.60 \leq r \leq .79$) มีจำนวน 5 ข้อ และข้อที่จำแนกได้ดีมาก ($.80 \leq r$) มีจำนวน 2 ข้อ แสดงว่าอำนาจจำแนกของข้อสอบมีลักษณะกระจาย โดยข้อสอบส่วนมากมีอำนาจจำแนกปานกลาง

2.3.3 การแสดงหลักฐานค่าความเชื่อมั่น

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำข้อสอบมาหาค่าความเชื่อมั่นรวมทั้งฉบับโดยใช้สัมประสิทธิ์แอลฟา (α) ผลการวิเคราะห์พบว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้น บูรณาการเมื่อทำการทดสอบครั้งที่ 1 มีค่าความเชื่อมั่น .91 เมื่อทำการทดสอบครั้งที่ 2 มีค่าความเชื่อมั่น .91 ซึ่งหลักฐานค่าความเชื่อมั่นของคะแนนอยู่ในเกณฑ์สูง

ตอนที่ 3 การแสดงหลักฐานคุณภาพเครื่องมือที่ศึกษาจากกลุ่มตัวอย่าง

3.1 การแสดงหลักฐานความเที่ยงตรงของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

การศึกษาครั้งนี้ ผู้วิจัยได้คัดเลือกข้อสอบและนำไปทดสอบกับกลุ่มตัวอย่างจำนวน 185 คน โดยใช้วิธีการวิเคราะห์ด้วยสูตรของไฮสน์ และบอร์นสเตรดต์ และวิธีวิเคราะห์องค์ประกอบเชิงยืนยัน ตามลำดับดังต่อไปนี้

3.1.1 การแสดงหลักฐานความเที่ยงตรงด้วยสูตรของไฮสน์ และบอร์นสเตรดต์

การวิเคราะห์ในตอนนี้ ผู้วิจัยได้นำคะแนนจากการตรวจให้คะแนนของผู้ตรวจแต่ละคนมาหาค่าความแปรปรวนร่วมระหว่างข้อคำถาม ($Cov(x_i, x_j)$) , ส่วนเบี่ยงเบนมาตรฐานรายข้อ (S_j) , วิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis) ด้วยโปรแกรม SPSS จากการหมุนแกนแบบมูมจาก ด้วยวิธีแวนริแมกซ์ (Varimax) และสกัดองค์ประกอบด้วยวิธี PC (Principal Component Analysis) หาค่าน้ำหนักองค์ประกอบรายข้อ (f_{ik}) นำมาคำนวณหาค่าความเที่ยงตรง (Construct validity) ด้วยสูตรของไฮสน์ และบอร์นสเตรดต์ (Heise and Bohrnstedt. 1970 : 109 , 111 – 113) ผลปรากฏว่าค่าผลรวมของคะแนนจากการตรวจให้คะแนนของผู้ตรวจ แต่ละคนมาหาค่าความแปรปรวนร่วมระหว่างข้อคำถาม ($\sum \sum Cov(x_i, x_j)$) มีค่า 161.65 ค่าผลรวมระหว่างผลคูณของส่วนเบี่ยงเบนมาตรฐานรายข้อกับค่าน้ำหนักองค์ประกอบรายข้อ

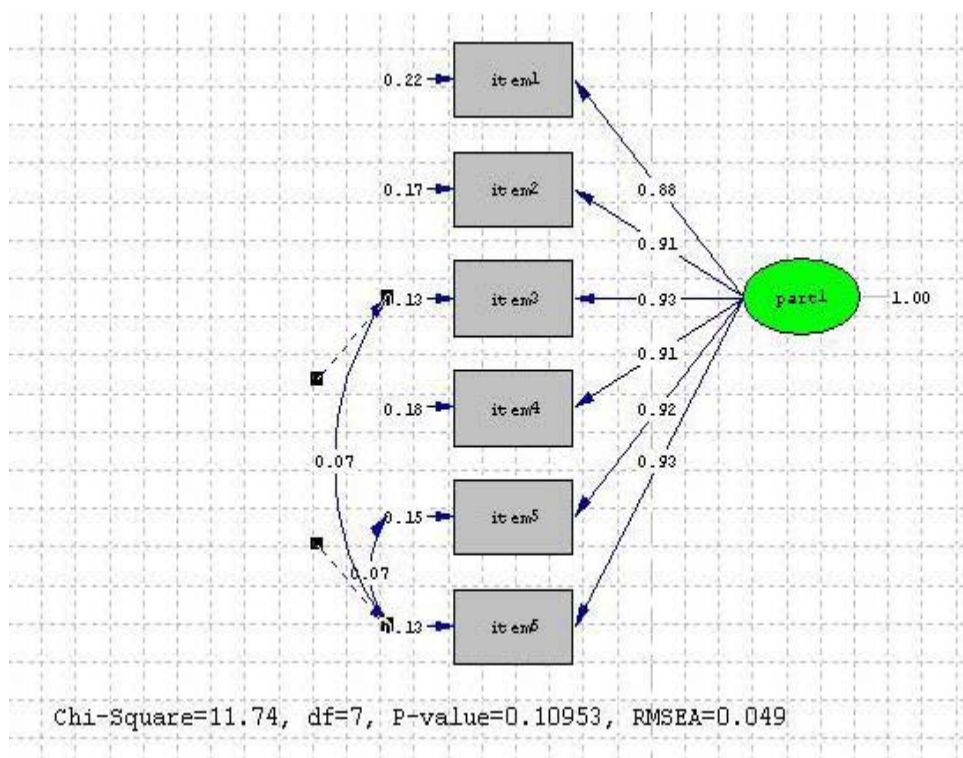
$(\sum S_{if_{ik}})$ มีค่า 8.89 และหลักฐานค่าความเที่ยงตรงของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการมีค่า 0.70 ซึ่งผลการวิเคราะห์เป็นหลักฐานแสดงว่าแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ มีความเที่ยงตรงอยู่ในเกณฑ์สูง

3.1.2 การแสดงหลักฐานความเที่ยงตรง ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis)

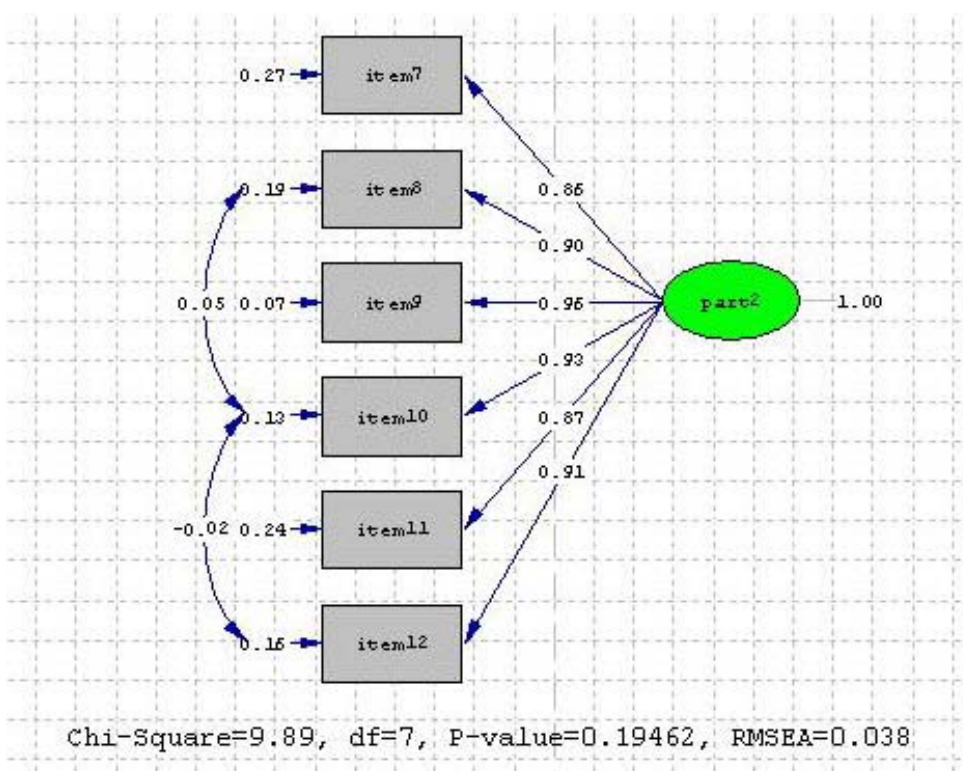
ผู้วิจัยแบ่งการวิเคราะห์ออกเป็น 2 อันดับ คือ การวิเคราะห์อันดับแรก (First Order) เป็นการวิเคราะห์เพื่อแสดงหลักฐานว่าข้อคำถามที่วัดในแต่ละทักษะมีความเที่ยงตรงมากน้อยเพียงใด และการวิเคราะห์อันดับที่สอง (Second Order) เป็นการวิเคราะห์เพื่อแสดงหลักฐานทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการมีความเที่ยงตรงมากน้อยเพียงใด ดังต่อไปนี้

3.1.2 การวิเคราะห์องค์ประกอบเชิงยืนยัน อันดับแรก

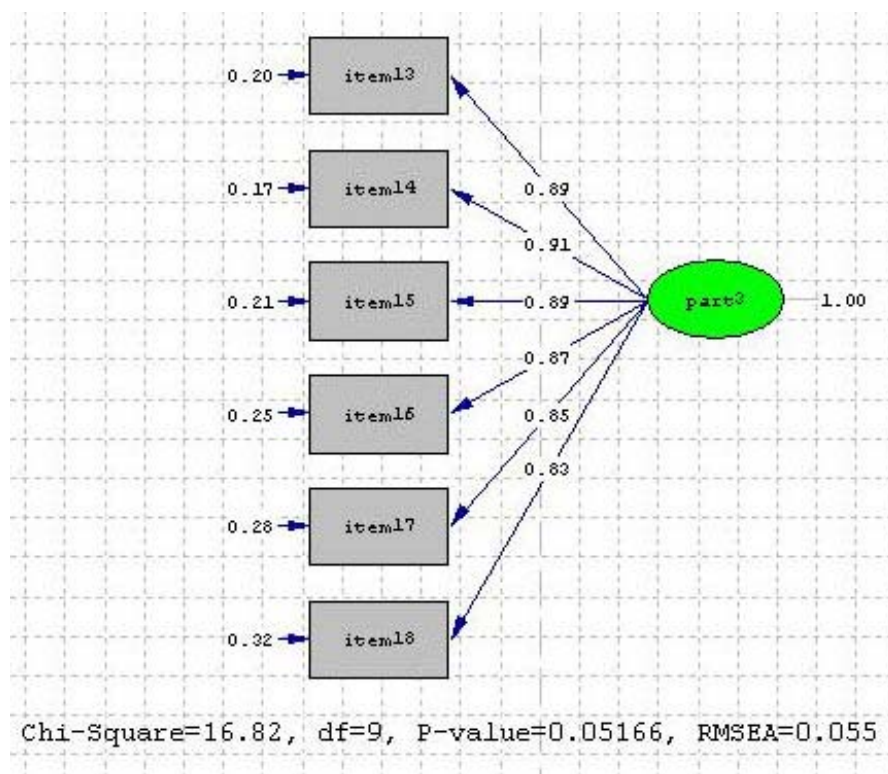
การวิเคราะห์ตอนนี้ เป็นการวิเคราะห์อันดับแรก (First Order) ผู้วิจัยได้นำคะแนนมาคำนวณค่าสหสัมพันธ์รายข้อในแต่ละตอน แล้วนำมาสร้างเมตริกซ์สหสัมพันธ์คะแนนดิบ (Correlation Matrix) มาวิเคราะห์เพื่อแสดงความเที่ยงตรง โดยใช้การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับแรกด้วยโปรแกรมลิสเรล (LISREL 8.54) ผู้วิจัยได้นำคะแนนมาแบ่งเป็น 5 ส่วนย่อย จำแนกเป็นรายตอนในแต่ละองค์ประกอบ แล้วนำคะแนน 5 ส่วนย่อยของแต่ละองค์ประกอบมาทดสอบความเหมาะสมพอดีของโมเดลกับข้อมูลเชิงประจักษ์ ปรากฏว่าโมเดลของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ไม่เหมาะสมพอดีกับข้อมูลเชิงประจักษ์ ดังนั้น ผู้วิจัยจึงได้ปรับโมเดลใหม่ โดยพิจารณาค่าดัชนีปรับแต่งโมเดล (modification indices) ตามวิธีของ จอเรสกอกและซอร์บอม (Joreskog and Sorbom. 1989 : 21) ปรากฏผลดังแสดงในภาพประกอบ 3



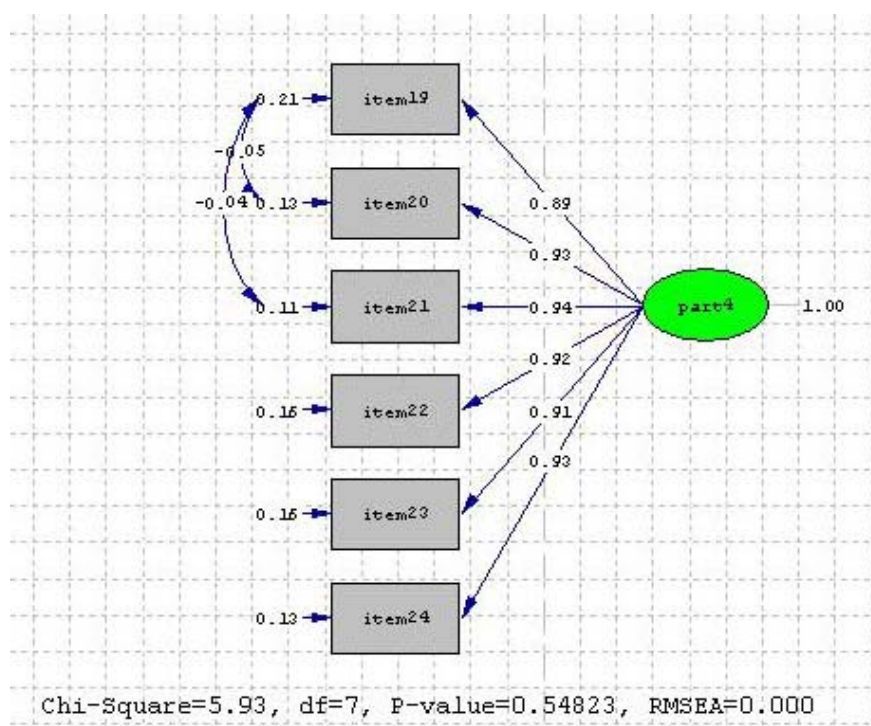
ทักษะการตั้งสมมติฐาน



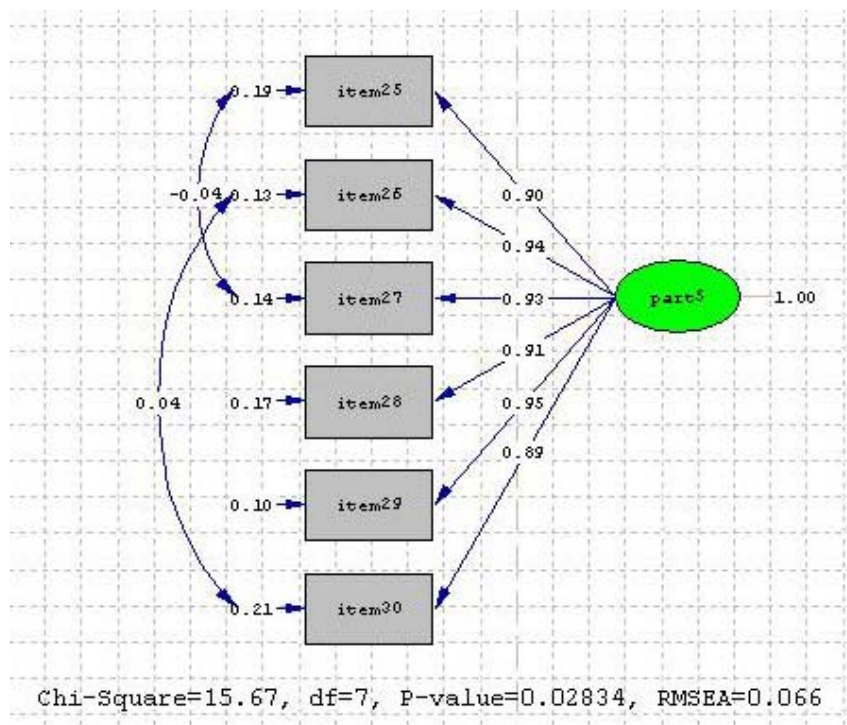
ทักษะการกำหนดนิยามเชิงปฏิบัติการ



ทักษะการกำหนดและควบคุมตัวแปร



ทักษะการทดลอง



ทักษะการตีความหมายข้อมูลและลงข้อสรุป

ภาพประกอบ 3 โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับแรก ของโมเดลแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการจำแนกเป็นตอน

การตรวจสอบความสอดคล้องของข้อมูลที่ได้จากการวัด และหาค่าความเชื่อมั่น (R^2) คำนวณหาค่าขององค์ประกอบ (b) ของคะแนนแต่ละข้อคำถาม รวมทั้งดัชนีแสดงความเหมาะสมพอดีแต่ละตอนกับข้อมูลเชิงประจักษ์ ปรากฏผลดังตาราง 11

ตาราง 11 ค่าสถิติที่ได้จากการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับแรกด้วยโปรแกรมลิสเรลจำแนกเป็นรายตอน (N = 185)

แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ	b	R^2	χ^2	p	GFI	RMSEA
การตั้งสมมติฐาน	.88 - .93	.78 - .87	11.74	.11	.99	.04
การกำหนดนิยามเชิงปฏิบัติการ	.86 - .96	.73 - .93	9.89	.20	.95	.03
การกำหนดและควบคุมตัวแปร	.83 - .91	.68 - .83	16.82	.05	.98	.05
การทดลอง	.89 - .94	.79 - .89	5.93	.55	.99	.00
การตีความหมายข้อมูลและลงข้อสรุป	.89 - .95	.79 - .90	15.67	.03	.98	.06

ผลการวิเคราะห์ตามตาราง 11 พบว่า หลักฐานค่าความเชื่อมั่นของการวัด พิจารณาจากค่าขอบเขตอันดับต่ำของความเชื่อมั่น (Squared Multiple Correlation : R^2) รายข้อจำแนกเป็นรายตอน มีหลักฐานค่าความเชื่อมั่นอยู่ระหว่าง .68 - .93 ซึ่งค่าสัมประสิทธิ์ความเชื่อมั่น สำหรับการวัดทักษะการตั้งสมมติฐาน พบว่าแบบวัดข้อ 3 และ ข้อ 6 สามารถแสดงค่าความเที่ยงตรงร่วมที่เกิดจากความสัมพันธ์ของข้อคำถามที่วัดคุณลักษณะเดียวกันได้มากที่สุด การวัดทักษะการกำหนดนิยามเชิงปฏิบัติการ พบว่าแบบวัดข้อ 9 สามารถแสดงค่าความเที่ยงตรงร่วมที่เกิดจากความสัมพันธ์ของข้อคำถามที่วัดคุณลักษณะเดียวกันได้มากที่สุด แบบวัดข้อ 10 สามารถแสดงค่าความเที่ยงตรงร่วมที่เกิดจากความสัมพันธ์ของข้อคำถามที่วัดคุณลักษณะเดียวกันได้รองลงมา การวัดทักษะการกำหนดและควบคุมตัวแปร พบว่าแบบวัดข้อ 14 สามารถแสดงค่าความเที่ยงตรงร่วมที่เกิดจากความสัมพันธ์ของข้อคำถามที่วัดคุณลักษณะเดียวกันได้มากที่สุด แบบวัดข้อ 13 สามารถแสดงค่าความเที่ยงตรงร่วมที่เกิดจากความสัมพันธ์ของข้อคำถามที่วัดคุณลักษณะเดียวกันได้รองลงมา การวัดทักษะการทดลอง พบว่าแบบวัดข้อ 21 สามารถแสดงค่าความเที่ยงตรงร่วมที่เกิดจากความสัมพันธ์ของข้อคำถามที่วัดคุณลักษณะเดียวกันได้มากที่สุด แบบวัดข้อ 20 และ ข้อ 24 สามารถแสดงค่าความเที่ยงตรงร่วมที่เกิดจากความสัมพันธ์ของข้อคำถามที่วัดคุณลักษณะเดียวกันได้รองลงมา การวัดทักษะการตีความหมายข้อมูลและลงข้อสรุป พบว่าแบบวัดข้อ 29 สามารถแสดงค่าความเที่ยงตรงร่วมที่เกิดจากความสัมพันธ์ของข้อคำถามที่วัดคุณลักษณะเดียวกันได้มากที่สุด แบบวัดข้อ 26 สามารถแสดงค่าความเที่ยงตรงร่วมที่เกิดจากความสัมพันธ์ของข้อคำถามที่วัดคุณลักษณะเดียวกันได้รองลงมา

สำหรับค่าน้ำหนักองค์ประกอบ (b) รายข้อจำแนกเป็นรายตอน มีค่าอยู่ระหว่าง .83 - .96 โดย แบบวัดข้อ 3 และ ข้อ 6 สามารถวัดทักษะการตั้งสมมติฐานได้ดีที่สุด แบบวัดข้อ 9 สามารถวัดทักษะการกำหนดนิยามเชิงปฏิบัติการได้ดีที่สุด แบบวัดข้อ 10 สามารถวัดทักษะการกำหนดนิยามเชิงปฏิบัติการได้รองลงมา แบบวัดข้อ 14 สามารถวัดทักษะการกำหนดและควบคุมตัวแปรได้ดีที่สุด แบบวัดข้อ 13 สามารถวัดทักษะการกำหนดและควบคุมตัวแปรได้รองลงมา แบบวัดข้อ 20 และ 24 สามารถวัดทักษะการทดลองได้ดีที่สุด แบบวัดข้อ 29 สามารถวัดทักษะการตีความหมายข้อมูลและลงข้อสรุปได้ดีที่สุด แบบวัดข้อ 26 สามารถวัดทักษะการตีความหมายข้อมูลและลงข้อสรุปได้รองลงมา เมื่อพิจารณาค่าน้ำหนักองค์ประกอบโดยรวมแล้วพบว่าวัดองค์ประกอบเดียวกัน

เนื่องจากผลลัพธ์ค่าน้ำหนักองค์ประกอบของแบบวัดทั้ง 30 ข้อ ที่จำแนกเป็นรายตอน ๆ ละ 6 ข้อ สำหรับการทดสอบค่าไค – สแควร์นั้น ได้ปรับค่าพารามิเตอร์โดยเปลี่ยนให้เป็นพารามิเตอร์

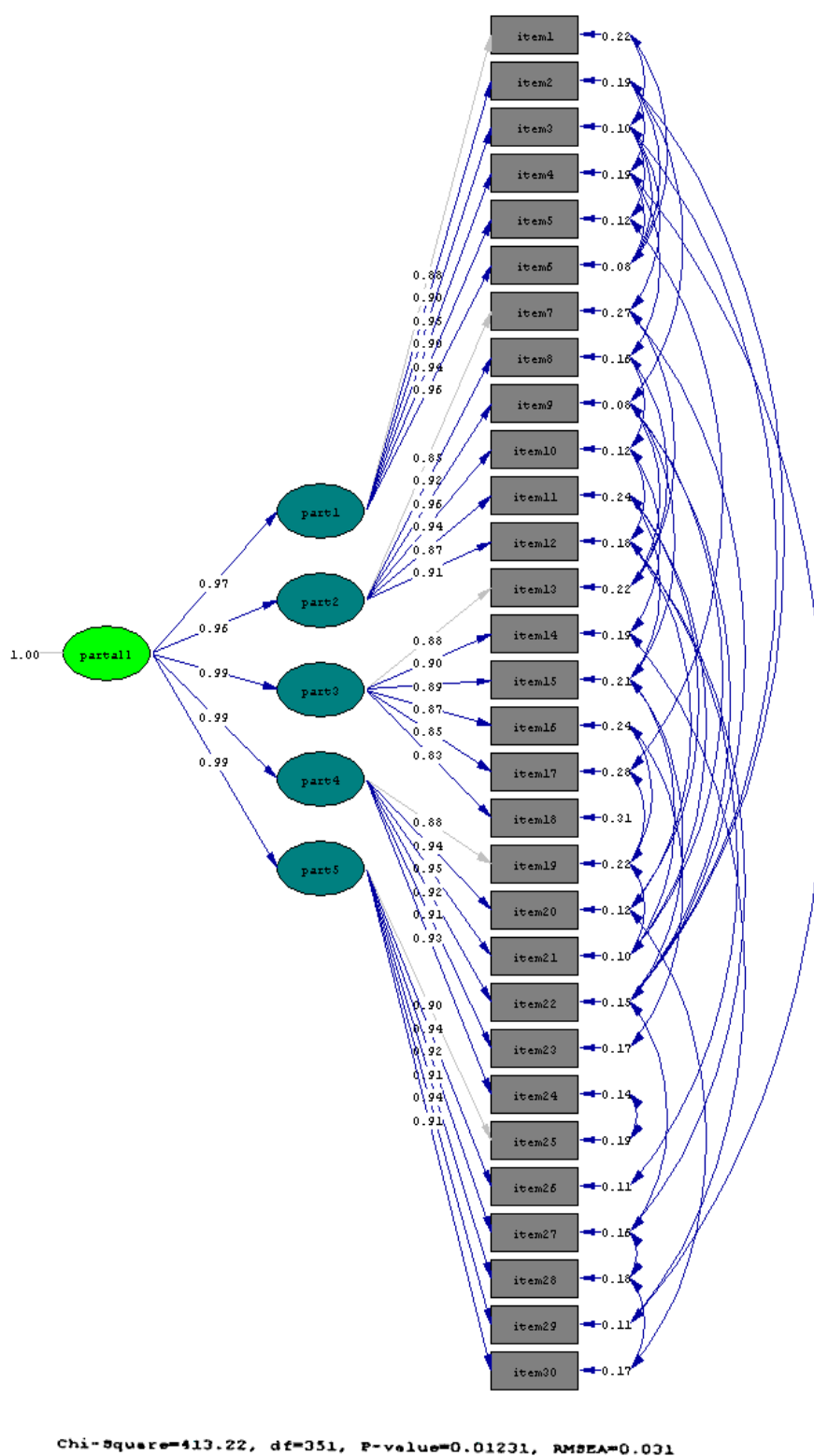
อิสระ เพื่อให้ค่าไค – สแควร์ไม่มีนัยสำคัญทางสถิติ หมายความว่าโมเดลกับข้อมูลเชิงประจักษ์มีความเหมาะสมจะพอดี

หลักฐานการทดสอบความเหมาะสมจะพอดีของโมเดลกับข้อมูลเชิงประจักษ์จำแนกเป็นรายตอน ค่าไค – สแควร์ (Chi – Square : χ^2) มีค่าอยู่ระหว่าง 5.93 ถึง 15.67 ซึ่งมีค่าแตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ (ตามเกณฑ์พิจารณาความเหมาะสมจะพอดี ค่า p มีค่ามากกว่า .01 ซึ่งมีค่าแตกต่างจากศูนย์อย่างไม่มีนัยสำคัญ) เมื่อพิจารณาค่าค่าสถิติ GFI ของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการจำแนกรายตอน 5 ตอน มีค่า GFI อยู่ระหว่าง .95 ถึง .99 (ตามเกณฑ์พิจารณาความเหมาะสมจะพอดี ค่า GFI ควรมีค่ามากกว่า .90) เมื่อพิจารณาค่าสถิติ RMSEA ของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการจำแนกรายตอน 5 ตอน มีค่า RMSEA อยู่ระหว่าง .00 ถึง .06 (ตามเกณฑ์พิจารณาความเหมาะสมจะพอดี ค่า RMSEA ควรมีค่าต่ำกว่า .05) แสดงว่าโมเดลของการวัดกับข้อมูลเชิงประจักษ์มีความเหมาะสมจะพอดีอยู่ในระดับดี ผลการทดสอบดังกล่าวเป็นหลักฐานแสดงว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการมีความเที่ยงตรงในการวัด

3.1.3 การวิเคราะห์องค์ประกอบเชิงยืนยัน อันดับที่สอง

การทดสอบความเหมาะสมจะพอดีของโมเดลกับข้อมูลเชิงประจักษ์ ของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการจำแนกเป็นรายตอน

การวิเคราะห์ตอนนี้ ผู้วิจัยได้นำคะแนนมาแบ่งเป็น 5 ส่วนย่อยตามตอนของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ แล้วนำคะแนน 5 ส่วนย่อยของแต่ละองค์ประกอบมาทดสอบความเหมาะสมจะพอดีของโมเดลกับข้อมูลเชิงประจักษ์ด้วยการวิเคราะห์องค์ประกอบอันดับที่สอง (Second Order) ปรากฏว่า โมเดลของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ไม่เหมาะสมจะพอดีกับข้อมูลเชิงประจักษ์ ดังนั้น ผู้วิจัยจึงได้ดำเนินการปรับโมเดลใหม่ โดยพิจารณาค่าดัชนีปรับแต่งโมเดล (modification indices) ตามวิธีของโจเรสกอกและซอร์บอม (Joreskog and Sorbom. 1989 : 21) โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สองของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ จำแนกเป็นรายตอน ดังแสดงในภาพประกอบ 4



ภาพประกอบ 4 โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สองของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ จำแนกเป็นรายตอน

ค่าสถิติจากการทดสอบความเหมาะสมของโมเดลกับข้อมูลเชิงประจักษ์ ของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการจำแนกเป็นรายตอน ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สอง ปรากฏผลดังตาราง 12

ตาราง 12 ค่าสถิติจากการทดสอบความเหมาะสมของโมเดลกับข้อมูลเชิงประจักษ์ ของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการจำแนกเป็นรายตอน ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน อันดับที่สอง

	χ^2	Probability level (p)	GFI	RMSEA
ค่าสถิติ	413.22	.012	.87	.03

ผลการวิเคราะห์ตามตาราง 13 พบว่า ค่าไค – สแควร์ (χ^2) มีค่า 413.22 ($p = .01$) เมื่อพิจารณาค่าสถิติ GFI มีค่า .87 และค่าสถิติ RMSEA มีค่า .03 เป็นไปตามเกณฑ์ของความเหมาะสมพอดี จึงพิจารณาได้ว่า แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการโมเดลเหมาะสมพอดีกับข้อมูลเชิงประจักษ์

3.2 การแสดงหลักฐานความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

3.2.1 การแสดงหลักฐานความเชื่อมั่น ด้วยการใช้ดัชนีความสอดคล้องของผู้ประเมิน (Rater Agreement Indexes : RAI)

ผู้วิจัยได้คำนวณค่าความเชื่อมั่นของผู้ตรวจ โดยใช้ดัชนีความสอดคล้องของผู้ประเมิน ของ Burry – Stock (Burry – Stock and other. 1996 : 259) ปรากฏผลดังตาราง 13

ตาราง 13 ค่าความเชื่อมั่นด้วยการใช้ดัชนีความสอดคล้องของผู้ประเมิน ของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

จำนวนผู้ตรวจ	ผลรวมคะแนนผู้ตรวจคนที่ 1	ผลรวมคะแนนผู้ตรวจคนที่ 2	ผลรวมคะแนนผู้ตรวจคนที่ 3	$\sum \sum \sum R_{mnk} - \bar{R}_{nk} $	RAI
3 คน	10312	11345	11611	2485.90	0.93

ผลการวิเคราะห์จากตาราง 13 พบว่า ค่าความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ เมื่อใช้ผู้ตรวจ 3 คน มีค่าเท่ากับ 0.93 ซึ่งผลการวิเคราะห์เป็นหลักฐานแสดงว่าแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ มีความเชื่อมั่นอยู่ในเกณฑ์สูง

3.2.2 การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่มีผู้ตรวจให้คะแนนจำนวน 3 คน

ผู้วิจัยได้คำนวณโดยให้คะแนนที่ผู้ตรวจทุกคน ตรวจให้คะแนนแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการทุกทักษะของผู้เข้าสอบทุกคน ปรากฏผลดังตาราง 14

ตาราง 14 แสดงการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่มีผู้ตรวจให้คะแนนจำนวน 3 คน

แหล่งความแปรปรวน	df	SS	MS	$\sigma^2(\rho)$	$\sigma^2(\delta)$	$E\rho^2$
p	284	263470.86	927.71			
r	2	3295.47	1647.74	3.24	1.54	0.68
e	18	2629.86	4.63			

ผลการวิเคราะห์ตามตาราง 14 พบว่า การตรวจให้คะแนนแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่มีผู้ตรวจให้คะแนนจำนวน 3 คน ผู้เข้าสอบจำนวน 185 คน มีค่าสัมประสิทธิ์การสรุปอ้างอิง เท่ากับ 0.68 แสดงว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ มีหลักฐานค่าความเชื่อมั่นอยู่ในเกณฑ์สูง

บทที่ 5

สรุปผล อภิปราย และข้อเสนอแนะ

ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีจุดมุ่งหมายสำคัญเพื่อแสดงหลักฐานความเที่ยงตรงและความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ โดยมีจุดมุ่งหมายเฉพาะของการวิจัยดังนี้

1. เพื่อแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ 2 วิธี คือ

1.1 ความเที่ยงตรงเชิงโครงสร้างด้วยสูตร $\rho_{T_k S_k}$ ของไฮเซน และบอร์นสเตรดต์ (Heise and Bohrnstedt. 1970 : 109 , 111 – 113)

1.2 ความเที่ยงตรงเชิงโครงสร้าง โดยใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) (บุญเชิด ภิญโญนนตพงษ์. 2547 : 193)

2. เพื่อแสดงหลักฐานความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ 2 วิธี คือ

2.1 ค่าดัชนีความสอดคล้องของผู้ประเมิน (Rater Agreement Indexes : RAI) ทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

2.2 ค่าความเชื่อมั่นของการตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการที่คำนวณโดยใช้ทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนมัธยมศึกษาสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กลุ่มที่ 3 เขตพื้นที่การศึกษา 3 กรุงเทพมหานคร ปีการศึกษา 2549 มีทั้งสิ้น 11 โรงเรียน จำนวนนักเรียนทั้งหมด 9,317 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 (ทุกแผนการเรียน) โรงเรียนมัธยมศึกษาสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กลุ่มที่ 3 เขตพื้นที่การศึกษา 3 กรุงเทพมหานคร ปีการศึกษา 2549 จำนวน 285 คน ซึ่งได้มาโดยการสุ่มแบบสองขั้นตอน (Two – Stage Random Sampling) โดยใช้เนื้อหาวิชาฟิสิกส์พื้นฐาน เคมีพื้นฐาน และชีววิทยาพื้นฐาน ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย เป็นแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ มีลักษณะเป็นคำถามปลายเปิดที่เกี่ยวกับฟิสิกส์พื้นฐาน เคมีพื้นฐาน และชีววิทยาพื้นฐาน ประกอบด้วย 5 ทักษะ คือ

- | | |
|---|-------------|
| 1. ทักษะการตั้งสมมติฐาน | จำนวน 6 ข้อ |
| 2. ทักษะการกำหนดนิยามเชิงปฏิบัติการ | จำนวน 6 ข้อ |
| 3. ทักษะการกำหนดและควบคุมตัวแปร | จำนวน 6 ข้อ |
| 4. ทักษะการทดลอง | จำนวน 6 ข้อ |
| 5. ทักษะการตีความหมายข้อมูลและลงข้อสรุป | จำนวน 6 ข้อ |

รวมเป็นแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ 1 ฉบับ จำนวน 5 ทักษะ จำนวนข้อสอบรวม 30 ข้อ ใช้เวลา 150 นาที

วิธีการดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลในการวิจัย ดังนี้

- นำหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อขอความอนุเคราะห์กับโรงเรียนที่เป็นกลุ่มตัวอย่าง
- ติดต่อประสานงาน ขออนุญาตผู้บริหารและครูผู้สอนของโรงเรียนที่ใช้เป็นกลุ่มตัวอย่างเพื่อกำหนดวัน เวลา สถานที่ วิธีดำเนินการสอบ และจัดประชุมชี้แจงครูที่ได้รับการคัดเลือกให้มาตรวจให้คะแนนจำนวน 3 คน ในครั้งนี้เพื่อให้เกิดความเข้าใจอันดี
- นำแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ผ่านการวิเคราะห์คุณภาพรายข้อแล้วไปทดสอบกับกลุ่มตัวอย่าง จำนวน 285 คน ซึ่งผู้ดำเนินการสอบชี้แจงให้กลุ่มตัวอย่าง และครูผู้ควบคุมการสอบให้ทราบถึงจุดมุ่งหมายของการวิจัย และความสำคัญของการวิจัยก่อนดำเนินการสอบ โดยนำไปเก็บรวบรวมข้อมูลระหว่างเดือนมกราคม พ.ศ.2550
- นำกระดาษคำตอบของผู้เข้าสอบทั้งหมดมาถ่ายเอกสารจำนวน 3 ชุด จัดให้ผู้ตรวจจำนวน 3 คน ได้ตรวจให้คะแนนตามแนวทางในการตอบและกฎเกณฑ์การให้คะแนนที่ผู้วิจัยได้สร้างขึ้น ในคู่มือการตรวจให้คะแนน
- นำผลที่ได้จากการตรวจให้คะแนนมาวิเคราะห์ข้อมูล ดังนี้
 - คำนวณค่าสถิติพื้นฐานของแบบวัดประเมินจำแนกตามผู้ตรวจให้คะแนน
 - คำนวณค่าความเที่ยงตรงของแบบวัดประเมินโดยใช้
 - สูตร $\rho_{T_k S_k}$ ของ ไฮซัน และบอร์นสเตดท์

5.2.2 การวิเคราะห์องค์ประกอบเชิงยืนยัน ซึ่งผู้วิจัยแบ่งการวิเคราะห์ออกเป็น 2 อันดับ คือ การวิเคราะห์อันดับแรก (First Order) เป็นการวิเคราะห์เพื่อแสดงหลักฐานว่าข้อคำถามที่วัดในแต่ละทักษะมีความเที่ยงตรงมากน้อยเพียงใด และการวิเคราะห์อันดับที่สอง (Second Order) เป็นการวิเคราะห์เพื่อแสดงหลักฐานทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการมีความเที่ยงตรงมากน้อยเพียงใด

5.3 คำนวณค่าความเชื่อมั่นของแบบวัดประเมินโดยใช้

5.3.1 การหาค่าดัชนีความสอดคล้องของผู้ประเมินโดยใช้ผู้ตรวจจำนวน 3 ท่าน ซึ่งมีประสบการณ์ในการสอนวิทยาศาสตร์ 5 ปีขึ้นไป

5.3.2 การหาค่าความเชื่อมั่นของการตรวจให้คะแนน โดยผู้ตรวจ 3 ท่าน แต่ละท่านตรวจข้อสอบทุกข้อ ของนักเรียนทุกคน

สรุปผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้ มีจุดมุ่งหมายสำคัญเพื่อพัฒนาแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ตามแนวของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท. 2526 : 1 – 16) โดยแบบวัดมีลักษณะเป็นคำถามปลายเปิดที่เกี่ยวกับฟิสิกส์พื้นฐาน เคมีพื้นฐาน และชีววิทยาพื้นฐาน ซึ่งมีจุดมุ่งที่จะแสดงหลักฐานความเที่ยงตรง และความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ สรุปได้ดังนี้

1. ค่าสถิติพื้นฐานที่ได้จากการทดสอบกับกลุ่มตัวอย่าง เป็นดังนี้ คะแนนเฉลี่ย ($\bar{X}$) ของผู้ตรวจคนที่ 3 มีค่ามากที่สุด รองลงมาคือคะแนนเฉลี่ยของผู้ตรวจคนที่ 2 และคะแนนเฉลี่ยของผู้ตรวจคนที่ 1 มีค่าต่ำสุด มีค่า 34.85 , 34.37 และ 32.57 ตามลำดับ เมื่อพิจารณาค่าส่วนเบี่ยงเบนมาตรฐานของผู้ตรวจทั้ง 3 คน มีค่าอยู่ระหว่าง 13.68 ถึง 15.47

ส่วนค่าความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ของผู้ตรวจทั้ง 3 คน พบว่า คะแนนของผู้ตรวจคนที่ 3 มีค่าความเชื่อมั่นสูงสุด รองลงมาคือคะแนนของผู้ตรวจคนที่ 2 และคะแนนของผู้ตรวจคนที่ 1 มีค่าต่ำสุด มีค่า .98 , .99 และ .99 ตามลำดับ แสดงว่า แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ตรวจโดยผู้ตรวจ 3 คน มีหลักฐานค่าความเชื่อมั่นอยู่ในเกณฑ์สูง

2. แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่สร้างขึ้น นำมาวิเคราะห์คุณภาพเครื่องมือ มีค่าความยากง่ายข้ออยู่ระหว่าง 0.31 ถึง 0.78 ส่วนค่าอำนาจจำแนกรายข้อมีค่าอยู่ระหว่าง 0.23 ถึง 0.82 โดยค่าอำนาจจำแนกรายข้อของแบบวัดส่วนใหญ่สามารถจำแนกได้ปานกลาง รองลงมาเป็นข้อสอบที่จำแนกได้บ้าง และที่เหลือเป็นข้อสอบที่จำแนกได้ดีและจำแนกได้ดีมาก

ส่วนหลักฐานค่าความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ นั้นเมื่อทำการทดสอบครั้งที่ 1 มีค่าความเชื่อมั่นรวมทั้งฉบับเท่ากับ .91 และเมื่อทำการทดสอบครั้งที่ 2 มีค่าความเชื่อมั่นรวมทั้งฉบับเท่ากับ .91 แสดงว่า แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการทั้งฉบับมีหลักฐานค่าความเชื่อมั่นอยู่ในเกณฑ์สูง

3. หลักฐานแสดงความเที่ยงตรงของแบบทดสอบ

การแสดงผลหลักฐานความเที่ยงตรง คำนวณด้วยสูตรของไฮสน์ และบอร์นสเตรดต์ พบว่าหลักฐานค่าความเที่ยงตรงของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการมีค่า .70 ซึ่งผลการวิเคราะห์เป็นหลักฐานแสดงว่าแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ มีความเที่ยงตรงอยู่ในเกณฑ์สูง

การแสดงผลหลักฐานความเที่ยงตรงเชิงโครงสร้างภายในฉบับของแบบทดสอบ ด้วย การวิเคราะห์องค์ประกอบเชิงยืนยัน โดยใช้โปรแกรมสำเร็จลิสเรล

ผู้วิจัยแบ่งการวิเคราะห์องค์ประกอบแบ่งการวิเคราะห์ออกเป็น 2 อันดับ คือ การวิเคราะห์อันดับแรก (First Order) เป็นการวิเคราะห์เพื่อแสดงหลักฐานว่าข้อคำถามที่วัดในแต่ละทักษะมีความเที่ยงตรงมากน้อยเพียงใด และการวิเคราะห์อันดับที่สอง (Second Order) เป็นการวิเคราะห์เพื่อแสดงหลักฐานทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการมีค่าความเที่ยงตรงมากน้อยเพียงใด ซึ่งในแต่ละอันดับจะจำแนกเป็นรายตอน

3.2.1 การวิเคราะห์อันดับแรก (First Order) ผู้วิจัยดำเนินการโดยนำคะแนนที่ได้จากผู้ตรวจ 3 คน มาวิเคราะห์องค์ประกอบเชิงยืนยัน ด้วยการตรวจสอบความสอดคล้องของข้อมูลที่ได้จากการวัดด้วยแบบทดสอบในแต่ละตอน และหาค่าน้ำหนักองค์ประกอบ (b) ของแต่ละข้อคำถาม รวมทั้งดัชนีแสดงความเหมาะสมของโมเดลกับข้อมูลเชิงประจักษ์ จำแนกเป็นรายตอน

ผลการวิเคราะห์ข้อมูลพบว่า ค่าดัชนีวัดความเหมาะสมของโมเดล ได้แก่ ค่า ไค - สแควร์ (χ^2) ค่าดัชนีวัดความเหมาะสมของโมเดล (GFI) และค่าความคลาดเคลื่อนเฉลี่ย (RMSEA) ของข้อมูลจากแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการทุกตอน โดยทุกตอนมีค่า ไค - สแควร์ แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ แสดงว่า โมเดลกับข้อมูลเชิงประจักษ์มีความเหมาะสมพอดี

ค่าน้ำหนักองค์ประกอบ (b) ซึ่งเป็นค่าความเที่ยงตรงของข้อคำถามของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ รายข้อจำแนกเป็นรายตอน ตอนที่ 1 วัดทักษะการตั้งสมมติฐาน มีค่าตั้งแต่ .88 ถึง .93 โดยแบบวัดข้อ 3 และ ข้อ 6 สามารถวัดได้ดีที่สุด ตอนที่ 2 วัดทักษะการกำหนดนิยามเชิงปฏิบัติการ มีค่าตั้งแต่ .86 ถึง .96 โดยแบบวัดข้อ 9 สามารถวัดได้ดีที่สุด ตอนที่ 3 วัดทักษะการกำหนดและควบคุมตัวแปร มีค่าตั้งแต่ .83 ถึง .91 โดยแบบวัดข้อ

14 สามารถวัดได้ดีที่สุด ตอนที่ 4 วัดทักษะการทดลอง มีค่าตั้งแต่ .89 ถึง .94 โดยแบบวัดข้อ 20 และ 24 สามารถวัดได้ดีที่สุด และตอนที่ 5 วัดทักษะการตีความหมายข้อมูลและลงข้อสรุป มีค่าตั้งแต่ .89 ถึง .95 โดยแบบวัดข้อ 29 สามารถวัดได้ดีที่สุด เมื่อพิจารณาโดยรวมแล้ว ตอนที่ 2 มีค่าโดยรวมสูงสุด รองลงมาคือตอนที่ 5 , ตอนที่ 4 , ตอนที่ 1 และตอนที่ 3 มีค่าต่ำสุด

แสดงว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ รายข้อจำแนกเป็นรายตอน มีความเที่ยงตรงในการวัด

3.2.2 การวิเคราะห์ห้ลำดับที่สอง (Second Order) เป็นการวิเคราะห์เพื่อแสดงหลักฐานว่า คุณลักษณะที่วัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ มีค่าความเที่ยงตรงมากน้อยเพียงใด ผู้วิจัยดำเนินการโดยนำคะแนนแต่ละตอนที่วัดองค์ประกอบทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ มาวิเคราะห์ห้องค์ประกอบเชิงยืนยัน ด้วยการตรวจสอบความเหมาะสมของโมเดลกับข้อมูลเชิงประจักษ์ และหาค่าน้ำหนักองค์ประกอบ ซึ่งเป็นค่าความเที่ยงตรงซึ่งจำแนกเป็นรายตอน

ผลการวิเคราะห์ข้อมูลพบว่า ค่าไค – สแควร์ (χ^2) มีค่า 413.22 ($p = .012$) เป็นไปตามเกณฑ์ของความเหมาะสมพอดี จึงพิจารณาได้ว่า แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ โมเดลกับข้อมูลเชิงประจักษ์มีความเหมาะสมพอดี

ส่วนค่าดัชนีวัดความเหมาะสมพอดี (GFI) มีค่า .87 มีค่าเข้าใกล้ 1.00 ส่วนค่าความคลาดเคลื่อนเฉลี่ย (RMSEA) มีค่า .03 มีค่าเข้าใกล้ศูนย์ แสดงว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ผู้วิจัยสร้างขึ้น โมเดลกับข้อมูลเชิงประจักษ์มีความเหมาะสมพอดี

จากการวิเคราะห์องค์ประกอบเชิงยืนยันทั้งสองอันดับดังกล่าว เป็นหลักฐานแสดงว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ผู้วิจัยสร้างขึ้น มีความเที่ยงตรงในการวัด

4. การแสดงหลักฐานค่าความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ นำมาตรวจด้วยผู้ตรวจ 3 คน พบว่าค่าความเชื่อมั่นด้วยการใช้ดัชนีความสอดคล้องของผู้ประเมิน 3 คน มีค่า 0.93 ซึ่งมีค่าเข้าใกล้ 1.00 และวิธีวิเคราะห์ด้วยการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง ที่มีผู้ตรวจ 3 คน มีค่า 0.68 แสดงว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการมีหลักฐานค่าความเชื่อมั่นอยู่ในเกณฑ์สูง

อภิปรายผล

การวิจัยครั้งนี้ มีจุดมุ่งหมายสำคัญเพื่อพัฒนาแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ตามแนวของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท. 2526 : 1 – 16) โดยแบบวัดมีลักษณะเป็นคำถามปลายเปิดที่เกี่ยวกับฟิสิกส์พื้นฐาน เคมีพื้นฐาน และชีววิทยาพื้นฐาน ซึ่งมีจุดมุ่งที่จะแสดงหลักฐานความเที่ยงตรง และความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ผลการวิจัยสามารถนำมาอภิปรายได้ดังนี้

1. แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการซึ่งมีลักษณะข้อคำถามเป็นแบบปลายเปิด มีค่าความยากอยู่ในช่วง 0.30 ถึง 0.78 และมีค่าอำนาจจำแนกตั้งแต่ 0.21 ถึง 0.82 พบว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ มีคุณภาพทั้งค่าความยากและอำนาจจำแนก ในระดับที่ใช้ประเมินผลความสามารถของผู้เข้าสอบได้ เช่นเดียวกับแบบทดสอบปรนัยเลือกตอบ แต่ให้ความโดดเด่นที่สามารถวัดความคิดของผู้เข้าสอบได้ตรงกว่า เพราะไม่ต้องมีตัวเลือกให้ผู้เข้าสอบ ซึ่งสอดคล้องกับคำกล่าวของกรอนลันด์ (Gronlund. 1965 : 108) ที่กล่าวว่าแบบทดสอบอัตนัยใช้ได้ดีกับการวัดความเข้าใจ ทักษะการคิด ความสามารถในการเขียน การแก้ปัญหา โดยเฉพาะการวัดความคิดริเริ่มสร้างสรรค์ และเหมาะสมอย่างยิ่งกับคำถามที่เน้นให้นักเรียนใช้ความสามารถในการหาคำตอบเอง

2. การแสดงหลักฐานความเที่ยงตรงของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ที่คำนวณด้วยสูตรของไฮเซน และบอร์นสเตรดต์ (Heise and Bohrnstedt) พบว่า การตรวจให้คะแนนโดยผู้ตรวจ 3 คน มีค่า .70 ซึ่งมีค่าพอใช้ ส่วนการแสดงผลหลักฐานความเที่ยงตรงเชิงโครงสร้างภายในฉบับ ผู้วิจัยใช้วิธีวิเคราะห์องค์ประกอบเชิงยืนยัน โดยทำการวิเคราะห์ทั้งอันดับแรก (First Order) และอันดับที่สอง (Second Order) ซึ่งผลการวิจัยพบว่า โดยรวมโมเดลกับข้อมูลเชิงประจักษ์มีความเหมาะสมเจาะพอดี เมื่อพิจารณารายละเอียดลงไปสู่การตรวจให้คะแนน พบว่า การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับแรก โมเดลกับข้อมูลเชิงประจักษ์มีความเหมาะสมเจาะพอดีกันทุกตอน โดยในส่วนของ การวัดทักษะการตั้งสมมติฐาน มีข้อคำถามทั้งหมด 6 ข้อ มีข้อคำถามที่มีความเกี่ยวข้องกันคือข้อ 3 กับข้อ 6 และ ข้อ 5 กับข้อ 6 การวัดทักษะการกำหนดนิยามเชิงปฏิบัติการ มีข้อคำถามทั้งหมด 6 ข้อ มีข้อคำถามที่มีความเกี่ยวข้องกันคือข้อ 8 กับข้อ 10 และ ข้อ 10 กับข้อ 12 การวัดทักษะการทดลอง มีข้อคำถามทั้งหมด 6 ข้อ มีข้อคำถามที่มีความเกี่ยวข้องกันคือข้อ 19 กับข้อ 20 และ ข้อ 19 กับข้อ 21 การวัดทักษะการตีความหมายข้อมูลและลงข้อสรุป มีข้อคำถามทั้งหมด 6 ข้อ มีข้อคำถามที่มีความเกี่ยวข้องกันคือข้อ 25 กับข้อ 27 และ ข้อ 26 กับข้อ 30 ทั้งนี้เป็นเพราะว่า คะแนนในแต่ละชุดมีจากข้อคำถามในแต่ละตอน ซึ่งผู้วิจัยพิจารณาลักษณะของข้อคำถามเป็นหลัก เนื่องจากในแต่ละตอนของคำถาม ประกอบด้วยข้อคำถามครบตาม

จุดประสงค์ ผลการวิเคราะห์จึงให้ผลสอดคล้องกันและเป็นพยานหลักฐานสนับสนุนความเที่ยงตรงของแบบวัดที่ผู้วิจัยสร้างขึ้น

ส่วนการทดสอบคุณลักษณะที่วัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ มีค่าความเที่ยงตรงมากน้อยเพียงใดนั้น พบว่าความเหมาะสมจะพอดีของโมเดลกับข้อมูลเชิงประจักษ์เป็นไปตามเกณฑ์ของความเหมาะสมจะพอดี มีความเที่ยงตรงในการวัด ข้อคำถามแต่ละข้อในทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการนั้นส่วนมากมีความเกี่ยวข้องกัน ซึ่งสามารถสรุปได้ว่าโมเดลกับข้อมูลเชิงประจักษ์มีความเหมาะสมจะพอดี เพราะในการวิเคราะห์แบบจำลองโครงสร้างนั้นกำหนดข้อตกลงเบื้องต้นว่าตัวแปรต่างๆ ในรูปแบบจำลองทางทฤษฎีมีการนำความคลาดเคลื่อนในการวัดมาร่วมวิเคราะห์ด้วยและยอมให้ความคลาดเคลื่อนมีความสัมพันธ์กันได้ (นงลักษณ์ วิรัชชัย. 2542 : 122) และในแต่ละลักษณะประกอบด้วยข้อคำถามจากแต่ละตอน

3. การแสดงหลักฐานความเชื่อมั่นของแบบวัด พบว่าคะแนนที่วิเคราะห์ด้วยดัชนีความสอดคล้องของผู้ประเมิน (RAI) จำนวน 3 คน มีค่าเท่ากับ 0.93 ซึ่งพบว่าผู้ตรวจจำนวน 3 คนมีค่าความเชื่อมั่นอยู่ในเกณฑ์สูง ซึ่งอุษณีย์ บัวศิริพันธุ์ (2543 : 55 – 56) กล่าวว่า ประสิทธิภาพของผู้ตรวจที่มีอายุการทำงานมากกว่า 5 ปี มีความเชี่ยวชาญในเนื้อหาวิชา และผ่านการตรวจให้คะแนนแบบทดสอบความเรียงมามาก และมีการกำหนดเกณฑ์การให้คะแนนที่เป็นมาตรฐาน จะส่งผลต่อคะแนนที่ได้จากการตรวจดีกว่า ส่วนการแสดงหลักฐานความเชื่อมั่นโดยใช้การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของการตรวจให้คะแนนที่กำหนดให้ผู้ตรวจทุกคนตรวจแบบวัดทุกทักษะของผู้เข้าสอบทุกคน มีค่าเท่ากับ 0.68 กรอนลันด์ (Gronlund. 1976 : 117) กล่าวว่า ถ้าความสามารถของผู้เข้าสอบในกลุ่มแตกต่างกันมาก การกระจายของคะแนนจะมากทำให้ความเชื่อมั่นสูง ถ้าความสามารถของกลุ่มผู้เข้าสอบใกล้เคียงกันมาก การกระจายของคะแนนจะมีค่าน้อยทำให้ความเชื่อมั่นต่ำ สอดคล้องกับ คันทิงแฮม (ชาญวิทย จรัสสุทธิศิร. 2545 : 100 อ้างอิงจาก Cunningham. 1986 : 112 – 118) ที่พบว่า ความสามารถผู้เข้าสอบที่หลากหลายจะส่งผลให้ค่าความเชื่อมั่นสูงขึ้นเมื่อความแปรปรวนของคะแนนเพิ่มขึ้น จะเห็นได้ว่าคะแนนจากการตรวจให้คะแนนของนักเรียนแต่ละคนจะมีค่าใกล้เคียงกัน ทำให้การกระจายของคะแนนมีค่าน้อย ส่งผลต่อค่าความเชื่อมั่นของแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

สรุปจากผลดังกล่าว จึงเป็นหลักฐานแสดงว่า แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ผู้วิจัยสร้างขึ้น มีคุณภาพในการวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการของนักเรียนระดับมัธยมศึกษาตอนปลาย

ข้อเสนอแนะ

ผลการวิจัยครั้งนี้ สามารถสรุปเป็นข้อเสนอแนะในการนำไปใช้ และในการวิจัย ได้ดังนี้

1. ข้อเสนอแนะในการนำไปใช้

1.1 จากผลการวิจัยพบว่าแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการมีหลักฐานแสดงความเที่ยงตรงและความเชื่อมั่นสูง ดังนั้นจึงควรนำไปใช้ในการวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

1.2 จากการดำเนินการวิจัย มีข้อสังเกตว่าคุณสมบัติของผู้ตรวจให้คะแนนควรกำหนดให้ชัดเจน ต้องเป็นผู้สอนในรายวิชานั้นๆ ควรมีความเชี่ยวชาญและมีประสบการณ์ในการสอนมากพอสมควร

1.3 จากการดำเนินการวิจัย ควรนำแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการไปใช้กับผู้เรียน และดำเนินการปรับปรุง / แก้ไข และพัฒนาให้เกิดพฤติกรรมด้านทักษะกระบวนการทางวิทยาศาสตร์ในตัวผู้เรียนต่อไป

2. ข้อเสนอแนะในการวิจัย

ควรมีการศึกษาโดยการออกแบบเครื่องมือที่เน้นการวัดการปฏิบัติ หรือจำลองสถานการณ์ในการปฏิบัติให้เหมือนกับสถานการณ์จริง เพื่อให้เกิดความสมบูรณ์ในการประเมินพฤติกรรมอันจะแสดงถึงความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์ของผู้เรียนได้อย่างถูกต้อง

บรรณานุกรม

บรรณานุกรม

- กระทรวงศึกษาธิการ. (2544). *หลักสูตรการศึกษาขั้นพื้นฐาน*. พิมพ์ครั้งที่ 2. กรุงเทพฯ: โรงพิมพ์
องค์การรับส่งสินค้าและพัสดุภัณฑ์.
- โกวิท ประวาลพุกษ์ และ สมศักดิ์ สิ้นธุระเวชญ์. (2525). *การประเมินในชั้นเรียน*. กรุงเทพฯ:
วัฒนาพานิช.
- โครงการพัฒนาทรัพยากรมนุษย์. (ม.ป.ป.) *การประเมินผลโดยใช้แฟ้มสะสมงานดีเด่น*. กระป๋:
โครงการโครงการพัฒนาทรัพยากรมนุษย์กระทรวงศึกษาธิการ.
- ชัยฤทธิ์ ศิลาเดช. (2540). *การพัฒนาแฟ้มสะสมงานในการประเมินผลการเรียนวิชาภาษาอังกฤษ
ระดับชั้นมัธยมศึกษาปีที่ 3*. วิทยานิพนธ์ ค.ด. (การวัดผลการศึกษา). กรุงเทพฯ:
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ชาญ วิชรเดช. (2538). *สัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบการแก้ปัญหาที่สอบด้วย
ไมโครคอมพิวเตอร์*. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ:
บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ชาญวิทย์ จรัสสุทธิศิร. (2545). *การพัฒนาเกณฑ์การให้คะแนนทักษะกระบวนการทาง
วิทยาศาสตร์ชั้นบูรณาการ*. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ:
บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ธิดา สอนงนารถ. (2542). *การสร้างแบบฝึกทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ
สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2*. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา).
กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- นงนุช สหัสดี. (2545). *การศึกษาความสัมพันธ์ระหว่างทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณา
การกับผลสัมฤทธิ์ทางการเรียนวิชาฟิสิกส์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5*.
วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัย
ศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- นงลักษณ์ วิรัชชัย. (2542). *โมเดลลิเรล : สถิติวิเคราะห์สำหรับการวิจัย*. พิมพ์ครั้งที่ 3. กรุงเทพฯ:
จุฬาลงกรณ์มหาวิทยาลัย.
- นพดล กองศิลป์. (2540). *การแสดงผลหลักฐานความเที่ยงตรงเชิงโครงสร้างและความเชื่อมั่นแบบ
คะแนนจริงสัมพันธ์ของแบบทดสอบวัดความรู้ความเข้าใจทางภาษาแบบประยุกต์ตาม
โครงสร้างทางสติปัญญาของกิลฟอร์ด*. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา).
กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

- นิศารัตน์ คงสวัสดิ์. (2544). ผลการวิเคราะห์ค่าความเชื่อมั่นของผู้ตรวจแบบทดสอบความเรียง ที่มีจำนวนผู้ตรวจ และวิธีการตรวจต่างกัน. ปรินุญานินพนธ์ กศ.ม. (การวัดผลการศึกษา).
กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- บุญชม ศรีสะอาด. (2534, มกราคม – เมษายน). การวิเคราะห์ข้อสอบแบบอื่นนอกเหนือจากแบบเลือกตอบ. วารสารวัดผลการศึกษา. 12(36) : 79 – 84.
- บุญเขต ภิญญอนันตพงษ์. (2521). การวัดและประเมินผลการศึกษา : ทฤษฎีและการประยุกต์.
กรุงเทพฯ: คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- (2524). ทฤษฎีการทดสอบ. กรุงเทพฯ: คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- (2525). เอกสารประกอบการสอนวิชา วัดผล 301 การสร้างแบบทดสอบ 1. กรุงเทพฯ: คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- (2526). การทดสอบแบบอิงเกณฑ์ : แนวคิดและปฏิบัติ. กรุงเทพฯ: โอเดียนสโตร์.
- (2533). Congeneric part Reliability. วารสารวัดผลการศึกษา. 12(34).
- (2544). การประเมินการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ : แนวคิดและวิธีการ. กรุงเทพฯ: สำนักงานคณะกรรมการการศึกษาแห่งชาติ
- (2545). คุณภาพเครื่องมือวัด. ประมวลสาระชุดวิชา การพัฒนาเครื่องมือสำหรับการประเมินการศึกษา หน่วยที่ 3. หน้า 65 – 153. กรุงเทพฯ: มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- (2547). การวัดประเมินการเรียนรู้. กรุงเทพฯ: คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ปรีชา เจริญ. (2531). การเปรียบเทียบทักษะกระบวนการทางวิทยาศาสตร์และเจตคติทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีรูปแบบการจัดชั้นเรียนและแผนการเรียนแตกต่างกันในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา เขตการศึกษา 11 ปีการศึกษา 2530. ปรินุญานินพนธ์ กศ.ม. (การมัธยมศึกษา). มหาสารคาม: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ มหาสารคาม. ถ่ายเอกสาร.
- พวงรัตน์ ทวีรัตน์. (2530). การสร้างและพัฒนาแบบทดสอบวัดผลสัมฤทธิ์. สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ.
- เพ็ญภา กุลนาคดล. (2547). การศึกษาและพัฒนาการปฏิบัติหน้าที่ของสมาชิกครอบครัวในครอบครัวของผู้ป่วยโรคซึมเศร้า. ปรินุญานินพนธ์ กศ.ด. (จิตวิทยาการให้คำปรึกษา).
กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

- ไพรัตน์ วงษ์นาม. (2533). *สัมประสิทธิ์การอ้างอิงสรุปของแบบทดสอบความเรียง*. วิทยานิพนธ์ ค.ด. (การวิจัยทางการศึกษา) กรุงเทพฯ: บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- มนวิภา อ่อนศรี. (2541). *การศึกษาความสัมพันธ์ระหว่างทักษะกระบวนการทางวิทยาศาสตร์กับความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3*. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- มยุรี ศรีชัย. (2538). *เทคนิคการสุ่มกลุ่มตัวอย่าง*. กรุงเทพฯ : ห้างหุ้นส่วนจำกัด วี.เจ.พรินติ้ง.
- มิ่ง เทพครเมือง. (2539) . *การตรวจสอบหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบการคิดเอกนัยด้านสัญลักษณ์ 6 ฉบับ ตามทฤษฎีของกิลฟอร์ด*. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- รวีวรรณ อังคนุรักษ์พันธุ์. (2531). *การสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์*. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ล้วน สายยศ; และอังคณา สายยศ. (2539). *เทคนิคการวัดผลการเรียนรู้*. กรุงเทพฯ: สุวีริยาสาส์น.
- วรรณทิพา รอดแรงคำ; และ จิต นวนแก้ว. (2533). *กิจกรรมทักษะกระบวนการทางวิทยาศาสตร์สำหรับนักเรียน*. กรุงเทพฯ: สถาบันส่งเสริมพัฒนาคุณภาพวิชาการ.
- วิญญา วิศาลาภรณ์. (2533). *การสร้างแบบทดสอบเพื่อการวิจัยและพัฒนาการเรียนการสอน*. กรุงเทพฯ: อักษราพิพัฒน์.
- วิชชุดา งามอักษร. (2541). *การศึกษาลักษณะพฤติกรรมการเรียนวิชาวิทยาศาสตร์ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ และความสามารถในการคิดอย่างมีเหตุผลของนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยการสอนแบบ เอส เอส ซี เอส กับการสอนตามคู่มือครู*. วิทยานิพนธ์ กศ.ม. (การมัธยมศึกษา) กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่าย เอกสาร.
- สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2524). *ทักษะกระบวนการทางวิทยาศาสตร์*. กรุงเทพฯ: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.
- . (2526). *ทักษะกระบวนการทางวิทยาศาสตร์*. กรุงเทพฯ: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.

- (2546). *การจัดสาระการเรียนรู้กลุ่มวิทยาศาสตร์ หลักสูตรการศึกษาขั้นพื้นฐาน*. กรุงเทพฯ: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.
- สนอง ตรงเที่ยง. (2540). *การแสดงผลหลักฐานของความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ วัดการประเมินค่าทางสัญลักษณ์แบบการแปลงรูป และแบบทดสอบวัดการประเมินค่าทางภาษาแบบการแปลงรูปตามทฤษฎีโครงสร้างทางสถิติปัญญาของกิลฟอร์ด*. ปริญญาโท กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สุทธิพงศ์ บุญผดุง. (2541). *การสร้างแบบทดสอบวัดลักษณะความรับผิดชอบสำหรับนักเรียน ชั้นมัธยมศึกษาปีที่ 3 โดยใช้การแสดงผลหลักฐานความเที่ยงตรง ความไม่เที่ยงตรง และความเชื่อมั่น*. ปริญญาโท กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สุนีย์ คล้ายนิล และคนอื่น ๆ. (2535). *เอกสารประกอบการอบรมครูผู้สอนกลุ่มสร้างเสริมประสบการณ์ชีวิต เนื้อหาวิทยาศาสตร์*. กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สุวัฒน์ นิยมคำ. (2531). *ทฤษฎีและทางปฏิบัติในการสอนวิทยาศาสตร์แบบสืบเสาะหาความรู้ เล่ม 1 – 2*. กรุงเทพฯ: เจเนอรัลบุ๊คเซนเตอร์.
- สำนักงานคณะกรรมการการศึกษาแห่งชาติ. (2542). *พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542*. สำนักนายกรัฐมนตรี.
- เสาวนีย์ เกรียร์. (2540). *คู่มือการอบรมการใช้แฟ้มสะสมผลงานนักเรียน*. กรุงเทพฯ: เนติกุลการพิมพ์.
- อุษณีย์ บัวศิริพันธุ์. (2543). *การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจ จำนวนผู้ตรวจ และประสบการณ์ของผู้ตรวจแตกต่างกัน*. ปริญญาโท กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- Anastasi, Anne. (1968). *Psychological Testing*. New York: Macmillan Company.
- Anderson, C.R. (1978, August). The Effective of A Simulation Learning Game in Teaching Consumer Credit to Senior Approach to Instruction, *Dissertation Abstract International*. 31(2): 127 – 132.
- Brennan, R.L. and Kane, M.T. (1979, June). Generalizability Theory : A review. *New Directions for Testing and Measurement*. 4(2) : 33 – 51

- (1983). *Element of Generalizability Theory*. Iowa : The American College Tesing Program.
- Burphy – Stock, Judith A.; Laurie, Cecelia, Chissom, Brad S. (1996, April). Rater Agreement Indexes For Performance Assessment, *Education and Psychological measurement*. 2 (56) : 251 – 261.
- Cronbach, L.J., Rajaratnam, N. And Gleser, G.C. (1963, August). Theory of Generalizability : A Liberalization of Reliability theory, *The British Journal of Statistical Psychology*. 16(1) :137 – 163
- Cronbach and other. (1972). *The Dependability of Behavioral Measurement : Theory of Generalizability for Score and Profiles*. New York : John Wiley & Son, Inc.
- Courtney, E.W. (1996). Alternative Assessment Techniques at the Elementary and Secondary Levels, *เอกสารการอบรมเชิงปฏิบัติการ. คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์*.
- Fernandez, J. and Meteo, M. (1993, April). “The Development and Factorial Validation Of The Academic Setting Evaluation Questionnaire” *Educational and Psychological Measurement*. 53(8) : 425 – 435.
- Gronlund, Norman E. (1965). *Measurement and Evaluation in Teaching*. New York : Macmillan.
- (1976). *Measurement and Evaluation in Teaching*. 3rd ed. New York : Macmillan.
- (1977). *Constructing Achievement Test*. 2nd ed. Prentice – Hall, Inc.
- Heise, D.R. and Bohrnstedt, G.W. (1970). Validity Invalidity and Reliability, *Sociological Methodology*. San Francisco : Jossey Bass.
- Hopkins, C.D., and Antes, R.L. (1990). *Classroom Measurement and Evaluation*. 3rd ed. Itasca, IL: F.E. Peacock.
- Ibrahim, A.M. (1984, August). The dependability of need assessment data : An application of Generalizability Theory of the rating of the educational goals, *Dissertation Abstracts International*. 45(2) : 499 – A
- James, L.R. Demaree, R.G., Wolf, G. (1993, October). r_{wg} : An assessment of within – group interrater Agreement, *Journal of applied psychology*. 78(2) : 306 – 309.

- Joreskog, K.G. and Sorbom, D. (1989). *LISREL 7 : User's Reference Guide*. Chicago : Scientific Software, Inc
- Klopfer, Leopold E. (1971). Evaluation of Learning in Science, *Handbook on Formative and Summative Evaluation of Student Learning*. New York: Benjamin S. Bloom, et al. McGraw – Hill, Co.
- Kubiszyn, T. And Borich, G. (1999). *Education Testing and Measurement : Classroom application and practice*. 6th Ed. Haper Collins College Publishers.
- Kunnan, A.J. (1992, November). An investigation of criterion – referenced test using G – theory, and factor And cluster analysis, *Language testing*. 9(1) : 30 – 49
- Long, J.S. (1983). *Confirmatory Factor Analysis*. Beverly Hills, California : Sage Publications, Inc.
- Perez, Carolina V. (1978, December). The Development and Evaluation of a Test of Science Process for Use in The Philippines, *Dissertation Abstracts International*. 39(6) : 3496 – A.
- Peterson, Kenneth D. (1978, March). Science Inquire Training for High School Students, *Journal of Research in Science Teaching*. 15(3) : 153.
- Pettus, A.M. and C.D. Haley. (1980, April). Identifying Factors Related to Science Process Skill Performance Levels. *School Science and Mathematics*. 30(2): 273 – 276
- Robinson, Richard Wayne. (1973, September). The Development of Item Which Access the Process of Controlling Variable and Interpreting Data, *Dissertation Abstracts International*. 35(3) : 1522 – A.
- Rubin, Rochelle Leventhal. (1989, April). Using a Systematic Modeling Teaching Strategy to Promote the Development of Integrated Science Process Skills and Formal Cognitive Reasoning Ability (Reasoning). *Dissertation Abstracts International*. 50(11) : 3469 – A.
- Spearman, C. (1910, March). Coefficient of Correlation Calculated from Faulty Data, *British Journal of Psychology*. 3(2) : 271 – 295.
- Stake, J.E. (1994, May) "Development and Validation of The Six – Factor Self – Concept Scale for Adults" *Education and Psychological Measurement*. 54(11) : 56 – 72.

- Stanley, Julian C. & Hopkins, Kenneth D. (1972). *Educational and Psychological Evaluation*. 5th ed. Englewood Cliffs: Prentice-Hall.
- Stenmark, J.K. (1991). *Mathematics Assessment : myths, models, good question, and Practical Suggestion*. University of California : The National Council of Teachers of Mathematics, INC.
- Thompson, Bruce. (1996, July) "Factor Analytic Evidence For The Construct Validity of Scores : A Historical Overview And Some Guidelines" *Educational and Psychological Evaluation*. 56(2) : 197 – 208.
- Tuckman, Bruce W. (1975). *Measurement of Educational Outcomes : Fundamentals of Testing*. New York: Harcourt Brace Jovanovich Inc.
- Ward, Annie W. (1999). *Assessment in Classroom*. p. 412. Wadsworth Publishing Company.
- Wiersma and others. (1985). *Educational Measurement and testing*. Boston: Allyn and Bacon, Inc.

ภาคผนวก

ภาคผนวก ก
คุณภาพของเครื่องมือ

ตาราง 15 ดัชนีความสอดคล้องของการประเมินจากผู้เชี่ยวชาญในการตรวจสอบคุณภาพของ
เครื่องมือ (IOC)

ทักษะ	ข้อ	ผู้เชี่ยวชาญท่านที่					รวม	IOC	แปลผล
		1	2	3	4	5			
การตั้งสมมติฐาน	1.1	0	0	1	1	1	3	0.60	ใช้ได้
	1.2	1	1	1	1	1	5	1.00	ใช้ได้
	1.3	1	0	1	1	1	4	0.80	ใช้ได้
	1.4	1	0	1	1	1	4	0.80	ใช้ได้
	1.5	1	0	1	1	1	4	0.80	ใช้ได้
	1.6	1	1	1	1	1	5	1.00	ใช้ได้
	1.7	1	1	1	1	1	5	1.00	ใช้ได้
	1.8	1	0	1	1	1	4	0.80	ใช้ได้
	1.9	1	1	1	1	1	5	1.00	ใช้ได้
	1.10	0	1	1	1	1	4	0.80	ใช้ได้
	1.11	1	0	1	1	1	4	0.80	ใช้ได้
	1.12	0	1	1	1	1	4	0.80	ใช้ได้
การกำหนดนิยามเชิงปฏิบัติการ	2.1	1	1	1	1	1	5	1.00	ใช้ได้
	2.2	1	1	1	1	1	5	1.00	ใช้ได้
	2.3	1	1	1	1	1	5	1.00	ใช้ได้
	2.4	1	1	1	1	1	5	1.00	ใช้ได้
	2.5	1	1	1	1	1	5	1.00	ใช้ได้
	2.6	1	1	1	1	1	5	1.00	ใช้ได้
	2.7	1	1	1	1	1	5	1.00	ใช้ได้
	2.8	1	1	1	1	1	5	1.00	ใช้ได้
	2.9	1	1	1	1	1	5	1.00	ใช้ได้
	2.10	1	1	1	1	1	5	1.00	ใช้ได้
	2.11	1	1	1	1	1	5	1.00	ใช้ได้
	2.12	1	1	1	1	1	5	1.00	ใช้ได้

ตาราง 15 (ต่อ)

ทักษะ	ข้อ	ผู้เชี่ยวชาญท่านที่					รวม	IOC	แปลผล
		1	2	3	4	5			
การกำหนดและควบคุมตัวแปร	3.1	1	1	1	1	1	5	1.00	ใช้ได้
	3.2	1	1	1	1	1	5	1.00	ใช้ได้
	3.3	1	1	1	1	1	5	1.00	ใช้ได้
	3.4	1	1	1	1	1	5	1.00	ใช้ได้
	3.5	1	1	1	1	1	5	1.00	ใช้ได้
	3.6	1	1	1	1	1	5	1.00	ใช้ได้
	3.7	1	1	1	1	1	5	1.00	ใช้ได้
	3.8	1	1	1	1	1	5	1.00	ใช้ได้
	3.9	0	1	1	1	1	4	0.80	ใช้ได้
	3.10	1	1	1	1	1	5	1.00	ใช้ได้
	3.11	-1	1	1	1	1	4	0.80	ใช้ได้
	3.12	1	1	1	1	1	5	1.00	ใช้ได้
การทดลอง	4.1	1	1	1	1	1	5	1.00	ใช้ได้
	4.2	1	1	1	1	1	5	1.00	ใช้ได้
	4.3	1	1	1	1	1	5	1.00	ใช้ได้
	4.4	1	1	1	1	1	5	1.00	ใช้ได้
	4.5	1	1	1	1	1	5	1.00	ใช้ได้
	4.6	1	1	1	1	1	5	1.00	ใช้ได้
	4.7	1	1	1	1	1	5	1.00	ใช้ได้
	4.8	1	1	1	1	1	5	1.00	ใช้ได้
	4.9	1	1	1	1	1	5	1.00	ใช้ได้
	4.10	1	1	1	1	1	5	1.00	ใช้ได้
	4.11	1	1	1	1	1	5	1.00	ใช้ได้
	4.12	1	1	1	1	1	5	1.00	ใช้ได้

ตาราง 15 (ต่อ)

ทักษะ	ข้อ	ผู้เชี่ยวชาญท่านที่					รวม	IOC	แปลผล
		1	2	3	4	5			
การตีความหมายข้อมูล และลงข้อสรุป	5.1	1	1	1	1	1	5	1.00	ใช้ได้
	5.2	1	1	1	1	1	5	1.00	ใช้ได้
	5.3	1	1	1	1	1	5	1.00	ใช้ได้
	5.4	1	1	1	1	1	5	1.00	ใช้ได้
	5.5	1	1	1	1	1	5	1.00	ใช้ได้
	5.6	1	1	1	1	1	5	1.00	ใช้ได้
	5.7	1	1	1	1	1	5	1.00	ใช้ได้
	5.8	1	1	1	1	1	5	1.00	ใช้ได้
	5.9	1	1	1	1	1	5	1.00	ใช้ได้
	5.10	-1	-1	-1	-1	-1	-5	-1.00	ใช้ไม่ได้
	5.11	1	1	1	1	1	5	1.00	ใช้ได้
	5.12	1	1	1	1	1	5	1.00	ใช้ได้

ตาราง 16 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของแบบวัดทักษะกระบวนการทาง
วิทยาศาสตร์ขั้นบูรณาการ

ทักษะ	ข้อ	ค่าความยาก (p)	ค่าอำนาจจำแนก (r)	ผลการคัดเลือก
การตั้งสมมติฐาน	1.1	0.27	0.18	ใช้ไม่ได้
	1.2	0.58	0.00	ใช้ไม่ได้
	1.3	0.58	0.31	ใช้ได้
	1.4	0.38	0.62	ใช้ได้
	1.5	0.54	0.26	คัดออก
	1.6	0.56	0.72	ใช้ได้
	1.7	0.84	0.19	ใช้ไม่ได้
	1.8	0.37	0.52	ใช้ได้
	1.9	0.68	0.57	ใช้ได้
	1.10	0.10	0.30	ใช้ไม่ได้
	1.11	0.56	0.61	ใช้ได้
	1.12	0.52	0.44	คัดออก
การกำหนดนิยามเชิง ปฏิบัติการ	2.1	0.51	0.44	ใช้ได้
	2.2	0.23	0.43	คัดออก
	2.3	0.53	0.55	ใช้ได้
	2.4	0.28	0.16	ใช้ไม่ได้
	2.5	0.92	0.23	ใช้ไม่ได้
	2.6	0.43	0.53	ใช้ได้
	2.7	0.40	0.40	ใช้ได้
	2.8	0.57	0.26	คัดออก
	2.9	0.27	0.25	คัดออก
	2.10	0.33	0.58	ใช้ได้
	2.11	0.28	0.27	คัดออก
	2.12	0.31	0.59	ใช้ได้

ตาราง 16 (ต่อ)

ทักษะ	ข้อ	ค่าความยาก (p)	ค่าอำนาจจำแนก (r)	ผลการคัดเลือก
การกำหนดและควบคุม ตัวแปร	3.1	0.32	0.25	คัดออก
	3.2	0.68	0.59	ใช้ได้
	3.3	0.86	0.29	ใช้ไม่ได้
	3.4	0.78	0.48	ใช้ได้
	3.5	0.90	0.39	ใช้ไม่ได้
	3.6	0.37	0.29	ใช้ได้
	3.7	0.95	0.25	ใช้ไม่ได้
	3.8	0.78	0.26	ใช้ได้
	3.9	0.83	0.17	ใช้ไม่ได้
	3.10	0.30	0.27	ใช้ได้
	3.11	0.88	0.02	ใช้ไม่ได้
	3.12	0.48	0.25	ใช้ได้
การทดลอง	4.1	0.82	0.04	ใช้ไม่ได้
	4.2	0.12	0.17	ใช้ไม่ได้
	4.3	0.54	0.46	ใช้ได้
	4.4	0.51	0.21	ใช้ได้
	4.5	0.70	0.70	ใช้ได้
	4.6	0.02	0.20	ใช้ไม่ได้
	4.7	0.95	0.28	ใช้ไม่ได้
	4.8	0.46	0.01	ใช้ไม่ได้
	4.9	0.51	0.63	ใช้ได้
	4.10	0.07	0.26	ใช้ไม่ได้
	4.11	0.38	0.32	ใช้ได้
	4.12	0.52	0.82	ใช้ได้

ตาราง 16 (ต่อ)

ทักษะ	ข้อ	ค่าความยาก (p)	ค่าอำนาจจำแนก (r)	ผลการคัดเลือก
การตีความหมายข้อมูล และลงข้อสรุป	5.1	0.5	0.35	ใช้ได้
	5.2	0.84	0.40	ใช้ไม่ได้
	5.3	0.51	0.11	ใช้ไม่ได้
	5.4	0.73	0.37	ใช้ได้
	5.5	0.77	-0.09	ใช้ไม่ได้
	5.6	0.43	0.40	ใช้ได้
	5.7	0.40	0.51	ใช้ได้
	5.8	0.63	0.31	คัดออก
	5.9	0.62	0.23	ใช้ได้
	5.11	0.37	0.18	ใช้ไม่ได้
	5.12	0.65	0.78	ใช้ได้
ค่าความเชื่อมั่น	.91			

ตาราง 17 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของแบบวัดทักษะกระบวนการทาง
วิทยาศาสตร์ขั้นบูรณาการ ของการทดสอบครั้งที่ 2

ทักษะ	ข้อ	ค่าความยาก (p)	ค่าอำนาจจำแนก (r)
การตั้งสมมติฐาน	1	0.58	0.29
	2	0.37	0.60
	3	0.56	0.67
	4	0.37	0.54
	5	0.68	0.55
	6	0.56	0.63
การกำหนดนิยามเชิง ปฏิบัติการณ์	7	0.51	0.41
	8	0.53	0.58
	9	0.43	0.53
	10	0.41	0.38
	11	0.33	0.58
	12	0.31	0.60
การกำหนดและควบคุม ตัวแปร	13	0.8	0.57
	14	0.78	0.50
	15	0.38	0.27
	16	0.78	0.26
	17	0.31	0.29
	18	0.48	0.28
การทดลอง	19	0.53	0.49
	20	0.50	0.23
	21	0.70	0.69
	22	0.52	0.59
	23	0.38	0.33
	24	0.52	0.82

ตาราง 17 (ต่อ)

ทักษะ	ข้อ	ค่าความยาก (p)	ค่าอำนาจจำแนก (r)
การตีความหมายข้อมูล และลงข้อสรุป	25	0.43	0.23
	26	0.72	0.41
	27	0.42	0.42
	28	0.38	0.54
	29	0.62	0.26
	30	0.65	0.77
ค่าความเชื่อมั่น	.91		

ตาราง 18 ค่าความยาก (p) ค่าอำนาจจำแนก (r) ของแบบวัดทักษะกระบวนการทาง
วิทยาศาสตร์ขั้นบูรณาการ ของการทดสอบกับกลุ่มตัวอย่าง

ทักษะ	ข้อ	ค่าความยาก (p)	ค่าอำนาจจำแนก (r)
การตั้งสมมติฐาน	1	0.61	0.90
	2	0.63	0.90
	3	0.64	0.94
	4	0.63	0.91
	5	0.61	0.92
	6	0.63	0.93
การกำหนดนิยามเชิง ปฏิบัติการณ์	7	0.64	0.85
	8	0.63	0.92
	9	0.64	0.96
	10	0.64	0.93
	11	0.67	0.90
	12	0.61	0.92
การกำหนดและควบคุม ตัวแปร	13	0.63	0.92
	14	0.64	0.92
	15	0.65	0.90
	16	0.66	0.89
	17	0.60	0.87
	18	0.62	0.84
การทดลอง	19	0.63	0.91
	20	0.61	0.92
	21	0.64	0.94
	22	0.65	0.91
	23	0.63	0.90
	24	0.62	0.92

ตาราง 18 (ต่อ)

ทักษะ	ข้อ	ค่าความยาก (p)	ค่าอำนาจจำแนก (r)
การตีความหมายข้อมูล และลงข้อสรุป	25	0.65	0.90
	26	0.64	0.92
	27	0.67	0.91
	28	0.64	0.91
	29	0.61	0.94
	30	0.61	0.91
ค่าความเชื่อมั่น	.99		

ตาราง 19 ค่าสถิติที่ได้จากการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับแรก โดยใช้โปรแกรมลิสเรล
จำแนกเป็นรายตอน (N = 185)

ตอนที่	ข้อที่	b	SE	t	R^2	χ^2	df	P – value
1	1	.88	.05	18.90	.78	11.74	7	.11
	2	.91	.05	19.96	.83			
	3	.93	.05	20.81	.87			
	4	.91	.05	19.74	.82			
	5	.92	.05	20.30	.85			
	6	.93	.05	20.71	.87			
2	7	.86	.05	18.02	.73	9.89	7	.20
	8	.90	.05	19.54	.81			
	9	.96	.04	22.11	.93			
	10	.93	.05	20.76	.87			
	11	.87	.05	18.53	.76			
	12	.91	.05	20.05	.82			
3	13	.89	.05	19.13	.80	16.82	9	.05
	14	.91	.05	19.80	.83			
	15	.89	.05	18.93	.79			
	16	.87	.05	18.29	.75			
	17	.85	.05	17.55	.72			
	18	.83	.05	16.94	.68			
4	19	.89	.05	19.08	.79	5.93	7	.55
	20	.93	.05	20.82	.87			
	21	.94	.04	21.26	.89			
	22	.92	.05	20.24	.84			
	23	.91	.05	20.13	.84			
	24	.93	.05	20.86	.87			

ตาราง 19 (ต่อ)

ตอนที่	ข้อที่	b	SE	t	R^2	χ^2	df	P – value
5	25	.90	.05	19.52	.81	15.67	7	.03
	26	.94	.05	20.90	.87			
	27	.93	.05	20.53	.86			
	28	.91	.05	19.97	.83			
	29	.95	.04	21.50	.90			
	30	.89	.05	19.20	.79			

ภาคผนวก ข
คู่มือการตรวจให้คะแนนแบบวัดทักษะกระบวนการทางวิทยาศาสตร์
ชั้นบูรณาการ

คู่มือการตรวจให้คะแนน แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

คำชี้แจง

1. คู่มือการตรวจให้คะแนนแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ประกอบด้วย 2 องค์ประกอบด้วยกัน คือ
 - 1.1 แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ
 - 1.2 กฎเกณฑ์การให้คะแนน (Scoring Rubric)
2. การตรวจให้คะแนนแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ กำหนดให้ตรวจความถูกต้องของคำตอบในแต่ละข้อโดยใช้แนวทางในการตอบประกอบการพิจารณาให้คะแนน และนำผลการตรวจ มาจัดระดับคุณภาพของผู้เข้าสอบ แยกเป็นรายทักษะ โดยใช้กฎเกณฑ์การให้คะแนน (Scoring Rubric) ประกอบการพิจารณา กำหนดระดับคุณภาพเป็น 3 ระดับ ทุกทักษะกระบวนการย่อย

แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

คำชี้แจง

1. แบบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการฉบับนี้ มีคำถามทั้งหมด 30 ข้อ ใช้เวลาในการทำแบบวัด 150 นาที
2. แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการฉบับนี้ ประกอบด้วย ทักษะกระบวนการย่อยจำนวน 6 ทักษะ คือ ทักษะการตั้งสมมติฐาน ทักษะการกำหนดนิยามเชิงปฏิบัติการ ทักษะการกำหนดและควบคุมตัวแปร ทักษะการทดลอง ทักษะการตีความหมายข้อมูล และลงข้อสรุป
3. แบบทดสอบทุกข้อเป็นแบบอัตนัย โดยที่แต่ละข้อกำหนดสถานการณ์ให้นักเรียนศึกษา ก่อนแล้วจึงเขียนตอบลงในที่ว่างที่เว้นไว้ในแต่ละข้อ
4. อย่าเปิดแบบทดสอบจนกว่าจะได้รับสัญญาณให้ลงมือทำ
5. เมื่อได้รับสัญญาณเตือนหมดเวลาในการทำข้อสอบ ให้นักเรียนหยุดการทำข้อสอบในทันที

ข้อ 1 ทักษะการตั้งสมมติฐาน

1.1 จงศึกษาข้อมูลจากตารางต่อไปนี้

การเคลื่อนที่ของแหล่งกำเนิดเสียง	ระดับเสียงที่ได้ยิน
1. เคลื่อนที่เข้าหาผู้ฟัง	สูง
2. เคลื่อนที่ออกห่างจากผู้ฟัง	ต่ำ

ข้อมูลในการทดลองจากตารางนี้ ควรตั้งสมมติฐานว่าอย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

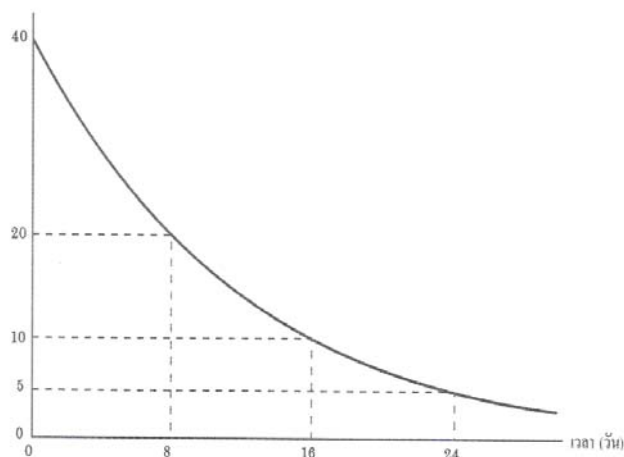
ระดับคุณภาพ	แนวทางในการตอบ
2	ถ้าแหล่งกำเนิดเสียงเคลื่อนที่เข้าหาผู้ฟังแล้ว จะได้ยินระดับเสียงจะสูงกว่าเคลื่อนที่ออกห่างจากผู้ฟัง
1	- ระดับเสียงสูง – ต่ำที่ได้ยิน ขึ้นอยู่กับระยะห่างระหว่างแหล่งกำเนิดเสียง กับผู้ฟัง - ระยะทางมีผลต่อความถี่
0	ไม่ตอบ หรือตอบไม่ถูก

1.2 จงศึกษาข้อมูลจากกราฟต่อไปนี้

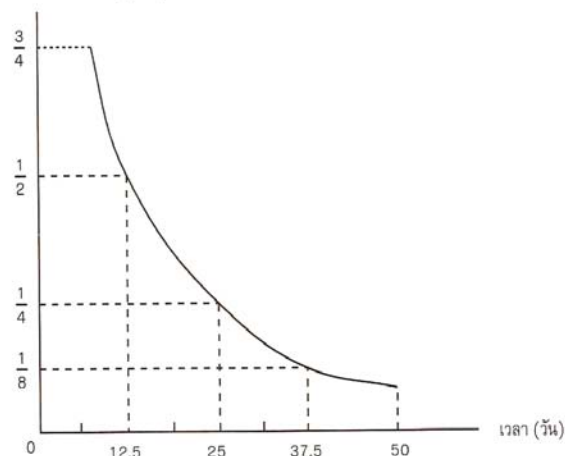
กราฟเปรียบเทียบอัตราการสลายตัวของธาตุกัมมันตรังสี

ธาตุไอโอดีน – 131 และธาตุแบเรียม – 140

จำนวนนิวเคลียส



ปริมาณแบเรียม-140 (กรัม)



ไอโอดีน – 131 มีครึ่งชีวิต 8 วัน

แบเรียม – 140 มีครึ่งชีวิต 12.5 วัน

จากกราฟ ควรตั้งสมมติฐานอย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- ถ้าธาตุแบเรียม – 140 มีครึ่งชีวิตมากกว่าธาตุไอโอดีน – 131 แล้ว ธาตุแบเรียม – 140 จะมีอัตราการสลายตัวช้ากว่าธาตุ ไอโอดีน – 131 - ถ้าธาตุไอโอดีน – 131 และ ธาตุแบเรียม – 140 มีปริมาณเท่ากันแล้ว จะใช้เวลาในการสลายตัวต่างกัน เมื่อมีครึ่งชีวิตต่างกัน
1	ถ้าธาตุมีครึ่งชีวิตต่างกันแล้ว จะมีอัตราการสลายตัวต่างกัน
0	ไม่ตอบ หรือตอบไม่ถูก

1.3 ผู้ติมานำสารละลายกลูโคสใส่ในถุงเยื่อเซลโลเฟน แล้วนำถุงนี้ไปแช่ในน้ำกลั่น ทั้งไว้ 20 นาที นำน้ำกลั่นที่ใส่แช่ถุง 5 cc ใส่ในหลอดทดลองขนาดกลาง นำไปหยดด้วยสารละลายเบเนดิกต์ 3 หยด ต่อกันนั้นนำหลอดทดลองไปแช่ในน้ำเดือด 5 นาที สารละลายในหลอดแก้วจะเปลี่ยนเป็นสีส้มเหลือง นักเรียนจะตั้งสมมติฐานว่าอย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- ถ้าสารละลายกลูโคสสามารถแพร่ผ่านถุงเยื่อเซลโลเฟนออกมาได้แล้ว น้ำที่ใส่แช่ถุงก็จะมีกลูโคสปนอยู่ทำให้เกิดการเปลี่ยนสีเมื่อทำปฏิกิริยากับเบเนดิกต์ หรือ - ถ้าสารละลายกลูโคส มีขนาดโมเลกุลเล็กกว่าช่องว่างระหว่างโมเลกุลของเยื่อเซลโลเฟนแล้วกลูโคสจะผ่านออกมาปนอยู่ในน้ำที่แช่ถุงได้ - น้ำตาลกลูโคสจะมีขนาดโมเลกุลเล็กจนผ่านเยื่อเซลโลเฟนได้ - สารละลายที่มีความเข้มข้นสูงจะแพร่ไปสู่สารละลายที่มีความเข้มข้นต่ำกว่า
1	- ถ้าสารละลายเปลี่ยนสีแล้ว แสดงว่าน้ำที่แช่ถุงมีกลูโคส - เบเนดิกต์เป็นสารที่ใช้ทดสอบน้ำตาลเมื่อนำมาทำปฏิกิริยากับกลูโคสจึงเกิดปฏิกิริยา
0	ไม่ตอบ หรือตอบไม่ถูก

1.4 ในการทดลองหาลักษณะเด่น – ด้อย ของต้นถั่ว ทำได้โดย นำถั่ว 2 ต้น ที่ลักษณะต่างกันมาผสมพันธุ์กันแล้วนำเมล็ดพันธุ์ (F_1) ที่ได้มาปลูก ปรากฏว่าลูกพันธุ์ที่ได้มีลักษณะเหมือนต้นแม่ ทุกต้น นำรุ่นลูกนี้ผสมกันเอง แล้วนำเมล็ดพันธุ์ (F_2) ที่ได้มาปลูก ปรากฏว่า รุ่นหลาน (F_2) มีลักษณะเหมือนต้นแม้อยู่ละ 75 เหมือนต้นพ่ออยู่ละ 25 นักเรียนจะตั้งสมมติฐานว่าอย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- ถ้าลูกในกลุ่ม F_2 ส่วนมากเหมือนต้นใด แล้วต้นนั้นจะเป็นลักษณะเด่น หรือ - ถ้าเป็นลักษณะเด่นแล้วจะพบลักษณะนั้นในกลุ่ม F_2 มากกว่ามีเป็นลักษณะด้อย - ลักษณะทางพันธุกรรมมีผลต่อการแสดงออกของฟีโนไทป์ ดังนั้นแม่จะมีลักษณะยีนส์เด่นพันธุ์แท้ พ่อมีลักษณะด้อยพันธุ์แท้ ลูกเป็นพันธุ์ทาง
1	- ถ้ารุ่น F_2 ส่วนใหญ่เหมือนต้นแม่แล้ว ต้นแม่เป็นลักษณะเด่น - ลักษณะเด่น – ด้อย จะปรากฏในรุ่นหลาน
0	ไม่ตอบ หรือตอบไม่ถูก

1.5 ในการทดลองครั้งหนึ่งได้ผลการทดลองดังตาราง นักเรียนจะตั้งสมมติฐานว่าอย่างไร

สารทดสอบ	ผลที่ได้จากการทดสอบกับ สารละลายเบเนดิกต์	ผลที่ได้จากการทดสอบกับ สารละลายไอโอดีน
1. แป้งมัน	ไม่เปลี่ยนแปลง	เปลี่ยนเป็นสีน้ำเงิน
2. กลัวย่น้ำว่า	เปลี่ยนเป็นสีเหลือง	เปลี่ยนเป็นสีน้ำเงิน
3. น้ำผึ้ง	เปลี่ยนเป็นสีเหลือง	ไม่เปลี่ยนแปลง
4. วุ้นเส้น	ไม่เปลี่ยนแปลง	เปลี่ยนเป็นสีน้ำเงิน

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- ถ้าสารทดสอบมีแบ่งเป็นส่วนประกอบแล้ว เมื่อถูกกับสารละลายไอโอดีนจะเปลี่ยนเป็นสีน้ำเงิน และถ้าสารทดสอบมีน้ำตาลเป็นส่วนประกอบแล้ว เมื่อถูกกับสารละลายเบเนดิกต์จะเปลี่ยนเป็นสีเหลือง - แบ่งมัน และอุ่นเส้นเป็นแบ่ง น้ำผึ้งมีน้ำตาล กลัวย่น้ำว่ามีทั้งน้ำตาลและแบ่ง
1	ตอบถูกบางส่วนของคำตอบ เช่น - ถ้าสารทดสอบมีแบ่งเป็นส่วนประกอบแล้ว เมื่อถูกกับสารละลายไอโอดีนจะเปลี่ยนเป็นสีน้ำเงิน หรือ - ถ้าสารทดสอบมีน้ำตาลเป็นส่วนประกอบแล้ว เมื่อถูกกับสารละลายเบเนดิกต์จะเปลี่ยนเป็นสีเหลือง
0	ไม่ตอบ หรือตอบไม่ถูก

1.6 ในการทดลองเพื่อหาสมบัติบางประการของพลาสติก ผู้ทดลองนำขวดน้ำ (PVC) พวงกุญแจ (ยูเรียฟอร์มัลดีไฮด์) และหลอดกาแฟ (พอลิเอทิลีน) มาเผาแล้วใช้กระดาษลิตมัส ชุบน้ำให้ขึ้นมา อังควัน ผลการทดลองได้ดังตาราง

ตัวอย่างพลาสติก	การเปลี่ยนสีของกระดาษลิตมัส
1. ขวดน้ำ (PVC)	น้ำเงินเปลี่ยนเป็นแดง
2. พวงกุญแจ (ยูเรียฟอร์มัลดีไฮด์)	แดงเปลี่ยนเป็นน้ำเงิน
3. หลอดกาแฟ (พอลิเอทิลีน)	ไม่เกิดการเปลี่ยนแปลง

นักเรียนจะตั้งสมมติฐานว่าอย่างไร

.....

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- ถ้าให้ควันท่เกิดจากการเผาพลาสติกถูกกับน้ำแล้ว จะมีฤทธิ์ความเป็นกรดเบสต่างกัน - พลาสติกชนิดต่างกัน ทำให้มีผลต่อการเปลี่ยนสีของกระดาษลิตมัสต่างกัน
1	ตอบถูกบางส่วนของคำตอบ เช่น - ควันท่เกิดจากการเผาพลาสติกต่างชนิดกัน จะมีฤทธิ์เป็นกรดหรือเบสต่างกัน - ขวดน้ำ (PVC) มีสมบัติเป็นกรด พวงกุญแจ (ยูเรียฟอรัลดีไฮด์) มีสมบัติเป็นเบส หลอดกาแฟ (พอลิเอทิลีน) ไม่มีคุณสมบัติเป็นกรด – เบส (เป็นกลาง)
0	ไม่ตอบ หรือตอบไม่ถูก

ข้อ 2 ทักษะการกำหนดนิยามเชิงปฏิบัติการ

2.1 ในการทดลองเกี่ยวกับการเคลื่อนที่ของวัตถุในแนวโค้งภายใต้แรงโน้มถ่วงของโลก โดยติดลูกเทนนิสกับแถบกระดาษ แล้วปล่อยให้ลูกเทนนิสตกลงสู่พื้น พร้อมกับดึงแถบกระดาษผ่านเครื่องเคาะสัญญาณเวลา คำว่า “ภายใต้แรงโน้มถ่วงของโลก” จะกำหนดนิยามว่าอย่างไร ให้สามารถใช้เป็นแนวทางในการสังเกตและวัดได้ตรงกัน

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- ภายใต้แรงโน้มถ่วงของโลก หมายถึง การเคลื่อนที่ที่อยู่ในอำนาจแรงดึงดูดของโลก - ภายใต้แรงโน้มถ่วงของโลก หมายถึง แรงที่โลกกระทำกับวัตถุในทิศทางจากที่สูงลงสู่ที่ต่ำ
1	- ภายใต้แรงโน้มถ่วงของโลก หมายถึง การเคลื่อนที่จากที่สูงลงสู่ที่ต่ำ - ภายใต้แรงโน้มถ่วงของโลก หมายถึง แรงที่โลกกระทำต่อวัตถุ - ภายใต้แรงโน้มถ่วงของโลก หมายถึง การตกในแนวตั้ง , การเคลื่อนที่ที่วิ่งเข้าหาโลก , ตกลงมาโดยอิสระ
0	ไม่ตอบ หรือตอบไม่ถูก

2.2 จงศึกษาข้อมูลจากตารางต่อไปนี้

การเคลื่อนที่ของแหล่งกำเนิดเสียง	ระดับเสียงที่ได้ยิน
1. เคลื่อนที่เข้าหาผู้ฟัง	สูง
2. เคลื่อนที่ออกห่างจากผู้ฟัง	ต่ำ

คำว่า “ระดับเสียง” จะกำหนดนิยามว่าอย่างไร ให้สามารถใช้เป็นแนวทางในการสังเกตและวัดได้ตรงกัน

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- ระดับเสียง หมายถึง ความถี่ของคลื่นเสียง หรือ - ระดับเสียง หมายถึง ช่วงกว้าง (Amplitude) ของคลื่นเสียง
1	- ระดับเสียง หมายถึง ความแหลม – ทุ้ม หรือ - ระดับเสียง หมายถึง ความดัง – ค่อย ของเสียง - ระดับเสียง หมายถึง ความสูงต่ำของเสียงที่คนสามารถได้ยิน - ระดับเสียง หมายถึง ความหนักเบาของเสียง - ระดับเสียง หมายถึง ความเข้มของเสียง
0	ไม่ตอบ หรือตอบไม่ถูก

2.3 จิตินำสารละลายกลูโคสใส่ในถุงเยื่อเซลล์โพลีเฟน แล้วนำถุงนี้ไปแช่ในน้ำกลั่น ทั้งไว้ 20 นาที นำน้ำกลั่นที่ใส่แฉ่ง 5 cc ใส่ในหลอดทดลองขนาดกลาง นำไปหยดด้วยสารละลายเบเนดิกต์ 3 หยด ต่อกจากนั้นนำหลอดทดลองไปแช่ในน้ำเดือด 5 นาที สารละลายในหลอดแก้วจะเปลี่ยนเป็นสีส้มเหลือง คำว่า “สารละลาย” จะกำหนดนิยามอย่างไร ให้สามารถใช้เป็นแนวทางในการสังเกตและวัดได้ตรงกัน

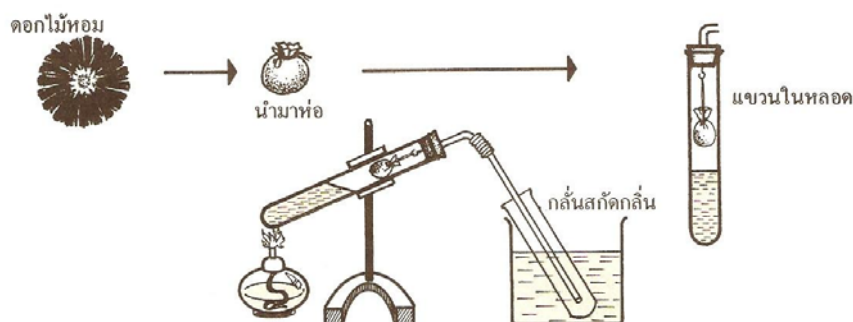
.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- สารละลาย หมายถึง ของเหลวที่มีธาตุ หรือสารประกอบตั้งแต่ 1 ชนิดขึ้นไป ละลายอยู่ - สารละลาย หมายถึง สารที่เกิดจากการรวมตัวกันของสาร 2 ชนิดขึ้นไปแล้ว มองเห็นเป็นเนื้อเดียวกัน - สารละลาย หมายถึง ของเหลวที่มีทั้งตัวทำละลายและตัวถูกละลายและมองเห็นเป็นเนื้อเดียวกัน
1	- สารละลาย หมายถึง น้ำที่มีกลูโคสละลายอยู่ - สารละลาย หมายถึง สารที่ผสมกันแล้วไม่มีตะกอน - สารละลาย หมายถึง ของเหลวที่มีสารหลายชนิดรวมเข้าด้วยกัน
0	ไม่ตอบ หรือตอบไม่ถูก

2.4 จงพิจารณาภาพการสกัดน้ำมันหอมระเหยต่อไปนี้



คำว่า “น้ำมันหอมระเหย” จะกำหนดนิยามอย่างไร ให้สามารถใช้เป็นแนวทางในการสังเกตและวัดได้ตรงกัน

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	น้ำมันหอมระเหย หมายถึง สารที่เป็นน้ำมัน ซึ่งอยู่ในพืช สามารถระเหยได้ มีกลิ่นหอม และสามารถสกัดได้ด้วยความร้อน
1	ตอบถูกบางส่วนของคำตอบ เช่น - น้ำมันหอมระเหย หมายถึง น้ำมันที่มีกลิ่นหอม ซึ่งสามารถสกัดได้จากพืช - น้ำมันหอมระเหย หมายถึง น้ำมันที่มีกลิ่นหอม และสามารถระเหยได้ - น้ำมันหอมระเหย หมายถึง น้ำมันที่ได้จากการกลั่นดอกไม้
0	ไม่ตอบ หรือตอบไม่ถูก

2.5 ในอุตสาหกรรมการกลั่นน้ำมันปิโตรเลียม โรงกลั่นจะให้ความร้อนกับน้ำมันดิบด้วยอุณหภูมิต่าง ๆ กัน จนสามารถกลั่นแยกสารต่าง ๆ ออกมาได้ในแต่ละช่วงของจุดเดือด คำว่า “น้ำมันดิบ” จะกำหนดนิยามอย่างไร ให้สามารถใช้เป็นแนวทางในการสังเกตและวัดได้ตรงกัน

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	น้ำมันดิบ หมายถึง ปิโตรเลียมเหลวที่สูบออกมาจากพื้นดิน มีสีดำหรือน้ำตาล ประกอบด้วย คาร์บอน ไฮโดรเจน ซัลเฟอร์ ออกซิเจน ไนโตรเจน และโลหะอื่น ๆ
1	- น้ำมันดิบ หมายถึง น้ำมันที่ขุดขึ้นมา แล้วยังไม่ได้นำมากลั่นแยกส่วน - น้ำมันดิบ หมายถึง น้ำมันที่เกิดขึ้นจากการทับถมกันของซากสิ่งมีชีวิต ภายใต้ความกดดันสูง และอุณหภูมิได้เปลือกโลก ซึ่งน้ำมันนี้สามารถนำไปกลั่นลำดับส่วนในหอกลั่นจะได้ผลิตภัณฑ์เกิดขึ้นมากมาย - น้ำมันดิบ หมายถึง สารที่มีสารประกอบไฮโดรคาร์บอนหลาย ๆ ชนิดปะปนกันอยู่
0	ไม่ตอบ หรือตอบไม่ถูก

2.6 ในการทดลองเพื่อให้ทราบว่าปฏิกิริยาเคมีใดเป็นชนิดคายความร้อน ได้ผลการทดลองดังตาราง

สารที่ใช้	อุณหภูมิ
1. มะนาว และดินสอพอง	เพิ่มขึ้น
2. กรด และน้ำ	เพิ่มขึ้น
3. การสลายตัวของผงฟู	ลดลง
4. การสังเคราะห์ด้วยแสงของพืช	ลดลง

คำว่า “ปฏิกิริยาเคมี” จะกำหนดนิยามอย่างไร ให้สามารถใช้เป็นแนวทางในการสังเกตและวัดได้ตรงกัน

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ปฏิกิริยาเคมี หมายถึง การเปลี่ยนแปลงของสารที่จะได้สารใหม่ที่มีสมบัติเปลี่ยนไปจากเดิม ทำให้สารใหม่มักกลับเป็นสารเดิมได้ยาก ขณะเปลี่ยนแปลงจะมีฟองแก๊ส ตะกอน ความร้อนเกิดขึ้น หรือสีของสารละลายเกิดการเปลี่ยนแปลง
1	ตอบถูกบางส่วนของคำตอบ เช่น - ปฏิกิริยาเคมี หมายถึง การเปลี่ยนแปลงของสารที่จะได้สารใหม่ที่มีสมบัติเปลี่ยนไปจากเดิม
0	ไม่ตอบ หรือตอบไม่ถูก

ข้อ 3 ทักษะการกำหนดและควบคุมตัวแปร

3.1 สาครปล่อยกระแสไฟฟ้าจากแบตเตอรี่เข้าสู่ขดลวดที่วางอยู่ในสนามแม่เหล็ก ทำให้ขดลวดหมุน พบว่าถ้าเขาใช้แบตเตอรี่ที่มีความต่างศักย์ต่างกัน ขดลวดจะหมุนด้วยความเร็วต่างกัน จงระบุตัวแปรตาม และตัวแปรควบคุม

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ตัวแปรตาม คือ ความเร็วในการหมุน และตัวแปรควบคุม คือ อำนาจสนามแม่เหล็ก หรือความเข้มสนามแม่เหล็ก หรือ จำนวนรอบของขดลวด
1	ตอบถูกบางส่วนของคำตอบ เช่น - ตัวแปรตาม คือ ความเร็วในการหมุน หรือ - ตัวแปรควบคุม คือ อำนาจสนามแม่เหล็ก หรือความเข้มสนามแม่เหล็ก หรือ จำนวนรอบของขดลวด
0	ไม่ตอบ หรือตอบไม่ถูก

3.2 ถ้านำกระดิ่งไฟฟ้าใส่ไว้ในครอบแก้วที่ปิดสนิท และกระดิ่งไม่แตะกับส่วนใดของครอบแก้ว เมื่อเปิดสวิตช์ จะได้ยินเสียงกระดิ่งดัง แต่เมื่อสูบลมอากาศในครอบแก้วออกจนหมด จะไม่ได้ยินเสียงกระดิ่ง ทั้ง ๆ ที่ยังมองเห็นกระดิ่งสั่นอยู่ จงระบุตัวแปรต้นและตัวแปรตาม

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ตัวแปรต้น คือ อากาศในครอบแก้ว และตัวแปรตาม คือ เสียงที่ได้ยิน
1	ตอบถูกบางส่วนของคำตอบ เช่น - ตัวแปรต้น คือ อากาศในครอบแก้ว หรือ - ตัวแปรตาม คือ เสียงที่ได้ยิน
0	ไม่ตอบ หรือตอบไม่ถูก

3.3 สารละลายกลูโคสใสในถุงเยื่อเซลโลเฟน แล้วนำถุงนี้ไปแช่ในน้ำกลั่น ทั้งไว้ 20 นาที นำน้ำกลั่นที่ใช้แช่ถุง 5 cc ใส่ในหลอดทดลองขนาดกลาง นำไปหยดด้วยสารละลาย เบเนดิกต์ 3 หยด ต่อจากนั้นนำหลอดทดลองไปแช่ในน้ำเดือด 5 นาที สารละลายในหลอดแก้วจะเปลี่ยนเป็นสี ส้มเหลือง จงระบุตัวแปรต้นและตัวแปรตาม

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ตัวแปรต้น คือ กลูโคสที่ผ่านเซลโลเฟนออกมา และตัวแปรตาม คือ การเปลี่ยนสีของสารละลาย
1	ตอบถูกบางส่วนของคำตอบ เช่น - ตัวแปรต้น คือ กลูโคสที่ผ่านเซลโลเฟนออกมา หรือ - ตัวแปรตาม คือ การเปลี่ยนสีของสารละลาย
0	ไม่ตอบ หรือตอบไม่ถูก

3.4 ในการทดลองเกี่ยวกับการคายน้ำของปากใบ โดยนำต้นไม้ที่จะทำการทดลองไปวางไว้ กลางแจ้ง และภายในห้อง โดยใช้เวลาเท่า ๆ กัน ผลปรากฏว่า ต้นไม้ขณะอยู่กลางแจ้งจะมีการคาย น้ำมากกว่าต้นไม้ที่วางไว้ในห้อง จงระบุตัวแปรต้น และตัวแปรตาม

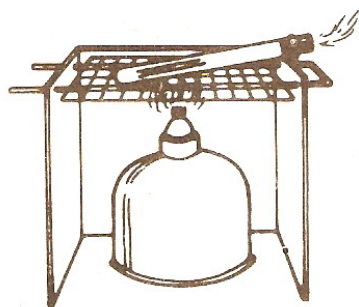
.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ตัวแปรต้น คือ สภาพที่ทำการทดลอง หรือ สถานที่ทำการทดลอง หรือ ความเข้มของแสงสว่าง และตัวแปรตาม คือ อัตราการคายน้ำ
1	ตอบถูกบางส่วนของคำตอบ เช่น - ตัวแปรต้น คือ สภาพที่ทำการทดลอง หรือ ความเข้มของแสงสว่าง หรือ สถานที่ทำการทดลอง หรือ - ตัวแปรตาม คือ อัตราการคายน้ำ
0	ไม่ตอบ หรือตอบไม่ถูก

3.5 การทดลองเลียนแบบการทำถ่านไม้



ทดลองใช้ก้านไม้ขีดไฟ 3 ก้าน ใส่ลงในหลอดทดลองทนไฟขนาดกลางปิดด้วยจุกยางที่
 เสาะเป็นร่องเล็กๆ ไว้ 1 ร่อง เพื่อให้อากาศออกได้ นำหลอดนี้ไปเผาบนตะแกรงด้วยตะเกียง
 แอลกอฮอล์ นาน 5 นาที ได้ของแข็งสีดำ เพราะ ต่อจากนั้นใช้ก้านไม้ขีดไฟ 3 ก้าน เผาบน
 ตะเกียงแอลกอฮอล์ใช้เวลาไม่เกิน 5 นาที ได้ขี้เถ้า จมระบุตัวแปรต้น และตัวแปรตาม

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ตัวแปรต้น คือ ลักษณะการเผาไม้ขีด หรือ วิธีการเผาไม้ขีด หรือ สถานการณ์ที่ อากาศเข้าไปทำปฏิกิริยาขณะเผาไหม้ได้หรือไม่ได้ และตัวแปรตาม คือ ขี้เถ้า และ ถ่าน หรือ ผลลัพธ์ที่ได้
1	ตอบถูกบางส่วนของคำตอบ เช่น - ตัวแปรต้น คือ ลักษณะการเผาไม้ขีด หรือ วิธีการเผาไม้ขีด หรือ สถานการณ์ที่ อากาศเข้าไปทำปฏิกิริยาขณะเผาไหม้ได้หรือไม่ได้ หรือ - ตัวแปรตาม คือ ขี้เถ้า และถ่าน หรือ ผลลัพธ์ที่ได้
0	ไม่ตอบ หรือตอบไม่ถูก

3.6 ในการทดลองเพื่อให้ทราบว่าปฏิกิริยาเคมีใดเป็นชนิดคายความร้อน ได้ผลการทดลองดังตาราง

สารที่ใช้	อุณหภูมิ
1. มะนาว และดินสอพอง	เพิ่มขึ้น
2. กรด และน้ำ	เพิ่มขึ้น
3. การสลายตัวของผงฟู	ลดลง
4. การสังเคราะห์ด้วยแสงของพืช	ลดลง

จงระบุตัวแปรต้น และตัวแปรตาม

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ตัวแปรต้น คือ ชนิดของปฏิกิริยาเคมีที่เกิดขึ้น และตัวแปรตาม คือ อุณหภูมิที่เปลี่ยนแปลงไป
1	ตอบถูกบางส่วนของคำตอบ เช่น - ตัวแปรต้น คือ ชนิดของปฏิกิริยาเคมีที่เกิดขึ้น หรือ - ตัวแปรตาม คือ อุณหภูมิที่เปลี่ยนแปลงไป
0	ไม่ตอบ หรือตอบไม่ถูก

ข้อ 4 ทักษะการทดลอง

4.1 จงออกแบบการทดลองเพื่อยืนยันว่า วัตถุที่มีน้ำหนักต่างกัน ตกจากที่สูงเท่ากัน จะถึงพื้นพร้อมกัน

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ทั้งวัตถุ 2 ชนิด ขนาดต่างกัน พร้อม ๆ กัน จากที่สูงเท่ากันตกในสุญญากาศ
1	ตอบถูกบางส่วนของคำตอบ เช่น - ทั้งวัตถุ 2 ชนิด ขนาดต่างกัน พร้อม ๆ กัน จากที่สูงเท่ากัน
0	ไม่ตอบ หรือตอบไม่ถูก

4.2 ถ้าต้องการมองเห็นเส้นแรงแม่เหล็ก นักเรียนจะทำการทดลองอย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	เอากระดาษวางทับแท่งแม่เหล็ก แล้วใช้ผงตะไบเหล็กโรยบนกระดาษ
1	- โรยผงเหล็กบนกระดาษ แล้วยกกระดาษไปวางบนแท่งแม่เหล็ก หรือ - วางแท่งแม่เหล็กแล้วโรยด้วยผลตะไบเหล็ก
0	ไม่ตอบ หรือตอบไม่ถูก

4.3 ถ้าใส่พวกเหล็กออกไซด์ลงไปในวันที่ใช้ปลูกต้นดอกไฮเดรนเยียแล้วจะมีผลทำให้ดอกไฮเดรนเยียเปลี่ยนสีไปจากเดิมได้ จงออกแบบการทดลองเพื่อพิสูจน์สมมติฐานนี้

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	1. ปลุกต้นไฮเดรนเยียรกระถางละ 1 ต้น 2 กระถาง 2. กระถางที่ 1 ฝังตะปู หรือเหล็กออกไซด์ ลงไปในดิน 3. สังเกตเปรียบเทียบสีของดอกระหว่างกระถางที่ 1 และ 2
1	ตอบถูกบางส่วนของคำตอบ เช่น - ปลุกต้นไฮเดรนเยียรกระถางละ 1 ต้น 2 กระถาง และกระถางที่ 1 ฝังตะปู หรือเหล็กออกไซด์ ลงไปในดิน หรือ - ใส่เหล็กออกไซด์ลงไปในดินที่ปลูก แล้วสังเกตดูการเปลี่ยนแปลง
0	ไม่ตอบ หรือตอบไม่ถูก

4.4 การที่จะนำแบ่งมันไปใช้ ในการทดสอบกับสารละลายเบเนดิกต์กับสารละลายไอโอดีน จะต้องจัดเตรียมแบ่งเพื่อทำการทดลองนี้อย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ใส่ผงแบ่งมันลงในน้ำแล้วเขย่าให้เป็นของเหลว
1	ตอบถูกบางส่วนของคำตอบ เช่น - ใส่ผงแบ่งมันลงในน้ำ
0	ไม่ตอบ หรือตอบไม่ถูก

4.5 จงออกแบบการทดลองเพื่อแสดงเกี่ยวกับแรงยึดเหนี่ยวระหว่างอิเล็กตรอนกับนิวเคลียส โดยกำหนดอุปกรณ์ให้คือ ถ้วยพลาสติกใส 1 ใบ , แท่งแม่เหล็ก 1 อัน , ลูกกลมเหล็กขนาดเส้นผ่านศูนย์กลาง 3 มิลลิเมตร 100 ลูก

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ใส่ลูกกลมเหล็กลงในถ้วยพลาสติกใส แล้วจุ่มปลายด้านหนึ่งของแท่งแม่เหล็กลงในถ้วยที่บรรจุลูกเหล็ก ยกแท่งแม่เหล็กขึ้นจากถ้วยแล้วแกะลูกเหล็กออกจากแท่งแม่เหล็ก เปรียบเทียบแรงที่ใช้ในการแกะลูกนอกสุดจนถึงลูกในสุด
1	ตอบถูกบางส่วน of คำตอบ เช่น - ใช้แท่งแม่เหล็กดูดลูกเหล็ก แล้วเปรียบเทียบแรงที่ใช้ในการแกะลูกนอกสุดจนถึงลูกในสุด
0	ไม่ตอบ หรือตอบไม่ถูก

4.6 จงออกแบบการทดลองเพื่อแสดงให้เห็นว่า พื้นที่ผิวมีผลต่ออัตราการเกิดปฏิกิริยา โดยกำหนดสารให้คือ เหล็กเม็ดใหญ่ และน้ำกลั่น

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	1. นำเกลือเม็ดใหญ่ขนาดเท่ากันมา 2 เม็ด 2. เม็ดที่ 1 ตำให้ละเอียด 3. ใส่น้ำลงในแก้ว 2 ใบเท่า ๆ กัน แล้วนำเกลือจากข้อ 2 ใส่ลงในแก้วใบ 1 และนำเกลืออีกเม็ดหนึ่งจากข้อ 1 ใส่ลงในแก้วใบ 2 พร้อม ๆ กัน 4. คนน้ำทั้ง 2 แก้ว จับเวลาเปรียบเทียบว่าเกลือในแก้วใดละลายก่อนกัน
1	ตอบถูกบางส่วน of คำตอบ เช่น - นำเกลือ 2 เม็ดขนาดต่าง ๆ กัน ใส่น้ำลงในแก้ว 2 ใบเท่า ๆ กัน คนน้ำทั้ง 2 แก้ว จับเวลาเปรียบเทียบว่าเกลือในแก้วใดละลายก่อนกัน
0	ไม่ตอบ หรือตอบไม่ถูก

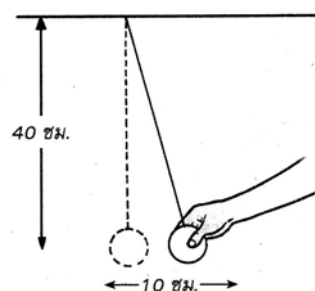
ข้อ 5 ทักษะการตีความหมายข้อมูลและลงข้อสรุป

5.1 แขนงลูกตุ้มลูกหนึ่งด้วยความยาวต่างกัน

2 ครั้ง แล้วดึงลูกตุ้มให้เกิดการแกว่ง โดยดึงให้ห่างจากแนว

ตั้ง ต่างกัน 2 ครั้ง จับเวลาเมื่อลูกตุ้มแกว่งครบรอบ 50

รอบ ได้ผลดังตาราง



ระยะทางที่ดึงลูกตุ้มให้ห่างจาก แนวตั้ง (cm)	ลูกตุ้มยาว 60 cm		ลูกตุ้มยาว 40 cm	
	เวลาที่ลูกตุ้มแกว่ง ครบ 50 รอบ (s)	ความถี่ (Hz)	เวลาที่ลูกตุ้มแกว่ง ครบ 50 รอบ (s)	ความถี่ (Hz)
8 cm	80	0.625	65	0.769
10 cm	80	0.625	65	0.769

จากข้อมูลข้างต้น ลงข้อสรุปได้ว่าอย่างไร

.....

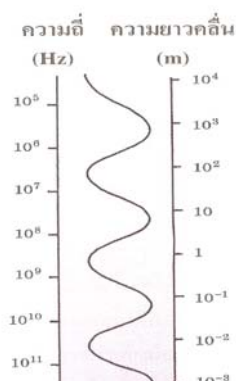
.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- เชือกผูกลูกตุ้มยาวจะมีความถี่น้อยกว่าเชือกผูกลูกตุ้มสั้น และระยะห่างจากแนวตั้ง ไม่มีผลต่อความถี่ - ความยาวของเชือกผูกลูกตุ้มมีผลต่อเวลาที่ใช้ในการแกว่งลูกตุ้มและความถี่ในการแกว่ง
1	ตอบถูกบางส่วนของคำตอบ เช่น - เชือกผูกลูกตุ้มยาวจะมีความถี่น้อยกว่าเชือกผูกลูกตุ้มสั้น หรือ - ระยะห่างจากแนวตั้ง ไม่มีผลต่อความถี่ หรือ - สิ่งที่มีผลต่อการเปลี่ยนแปลงความถี่ในการแกว่งคือความยาวเชือก
0	ไม่ตอบ หรือตอบไม่ถูก

5.2



จากข้อมูลข้างต้นแสดงเกี่ยวกับความถี่และความยาวคลื่น ลงข้อสรุปได้ว่าอย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ความถี่สูง ความยาวคลื่นน้อย และความถี่น้อย ความยาวคลื่นมาก หรือ ความถี่แปรผกผันกับความยาวคลื่น
1	ตอบถูกบางส่วนของคำตอบ เช่น - ความถี่สูง ความยาวคลื่นน้อย หรือ - ความถี่น้อย ความยาวคลื่นมาก
0	ไม่ตอบ หรือตอบไม่ถูก

5.3 ในการทดลองใช้ยาปฏิชีวนะยับยั้งการเจริญเติบโตของแบคทีเรียชนิดหนึ่ง โดยการป้ายเชื้อโรคลงบนแผ่นกระดาษ 3 แผ่นในปริมาณเท่ากัน แล้วหยดยาปฏิชีวนะ 3 ชนิดลงบนกระดาษแต่ละแผ่น ปรากฏผลดังตาราง

ชื่อยาปฏิชีวนะ	ปริมาณยาบนแผ่นกระดาษ (μg)	ขนาดเส้นผ่านศูนย์กลางบริเวณที่แบคทีเรียถูกทำลาย (mm)
1. เตตราซัยคลิน	30	20 – 29
2. สเตรปโตมัยซิน	10	15 – 23
3. แอมพิซิลลิน	10	28 – 36

จากผลการทดลอง จะสรุปผลได้อย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ยาแอมพิซิลลิน ใช้ทำลายแบคทีเรียได้ดีกว่ายาสเตรปโตมัยซิน และยาสเตรปโตมัยซินใช้ทำลายแบคทีเรียได้ดีกว่ายาเตตราซัยคลิน
1	ตอบถูกบางส่วนของคำตอบ เช่น - ยาแอมพิซิลลิน ใช้ทำลายแบคทีเรียได้ดีที่สุด หรือ - ยาเตตราซัยคลิน ใช้ทำลายแบคทีเรียได้น้อยที่สุด
0	ไม่ตอบ หรือตอบไม่ถูก

5.4 ในการทดลองเลือกใช้ตัวทำละลายเพื่อสกัดสารเคมีจากพืช (ขิง) ได้ผลดังตาราง

ของเหลวที่ใช้เป็นตัวทำละลาย	ลักษณะและกลิ่นของสารที่ได้จากการสกัด
1. เอทิลแอลกอฮอล์	สารละลายมีสีเหลือง มีกลิ่นขิงปนกลิ่นแอลกอฮอล์
2. น้ำ	สารละลายไม่มีสี มีกลิ่นขิง

จากผลการทดลอง จะสรุปผลได้อย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	แอทิลแอลกอฮอล์ สกัดสารเคมีจากขิงได้ดีกว่าน้ำ
1	- แอทิลแอลกอฮอล์ และน้ำ สามารถสกัดสารเคมีจากขิงได้ หรือ - เอทานอล และน้ำสามารถสกัดสารเคมีจากพืชได้ต่างกัน เอทานอลจะได้สารละลายที่มีกลิ่น แต่น้ำจะให้สารละลายมีกลิ่นแต่ไม่มีสี
0	ไม่ตอบ หรือตอบไม่ถูก

5.5 ในการทดลองเพื่อเปรียบเทียบปริมาณกรดไขมันไม่อิ่มตัวในน้ำมันพืช และน้ำมันสัตว์ พบว่า จำนวนหยดของทิงเจอร์ไอโอดีนที่ใช้ในน้ำมันสัตว์ น้อยกว่าในน้ำมันพืช ซึ่งจำนวนหยดของทิงเจอร์ไอโอดีนเป็นสัดส่วนโดยตรงกับจำนวนพันธะคู่ในกรดไขมันไม่อิ่มตัว นักเรียนจะสรุปผลการทดลองนี้ว่าอย่างไร

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	ในน้ำมันสัตว์มีกรดไขมันไม่อิ่มตัวน้อยกว่าในน้ำมันพืช
1	ในน้ำมันสัตว์มีพันธะคู่น้อยกว่าในน้ำมันพืช
0	ไม่ตอบ หรือตอบไม่ถูก

5.6 จาการ่างจสรุปความสัมพันธ์ของไอโซโทปไฮโดรเจนทั้ง 3 ชนิด

	โปรตอน	นิวตรอน	อิเล็กตรอน
${}^1_1\text{H}$	1	0	1
${}^2_1\text{H}$	1	1	1
${}^3_1\text{H}$	1	2	1

.....

.....

.....

กฎเกณฑ์การให้คะแนน

ระดับคุณภาพ	แนวทางในการตอบ
2	- ไฮโดรเจนทั้งสามชนิดมี โปรตอน และอิเล็กตรอนเท่ากัน แต่มีนิวตรอนต่างกัน หรือ - H-3 มีนิวตรอนมากกว่า H-2 , H-2 มีนิวตรอนมากกว่า H-1 และ ทั้ง 3 มีโปรตอนและอิเล็กตรอนเท่ากัน
1	ไฮโดรเจนทั้งสามไอโซโทป แต่ละตัวต่างกันที่จำนวนนิวตรอน
0	ไม่ตอบ หรือตอบไม่ถูก

ภาคผนวก ค
รายนามผู้เชี่ยวชาญ

รายนามผู้เชี่ยวชาญในการตรวจสอบคุณภาพเครื่องมือ

ผู้เชี่ยวชาญด้านการวัดผลและประเมินผล

- | | |
|--------------------------|--|
| 1. รศ. ละเอียด รักษ์เฝ้า | คณะศึกษาศาสตร์ ภาควิชาการวัดผลและวิจัยการศึกษา
มหาวิทยาลัยศรีนครินทรวิโรฒ |
| 2. รศ.ชูศรี วงศ์รัตนะ | คณะศึกษาศาสตร์ ภาควิชาการวัดผลและวิจัยการศึกษา
มหาวิทยาลัยศรีนครินทรวิโรฒ |
| 3. อาจารย์พรณี จินตมาศ | หัวหน้างานวัดผลและประเมินผล โรงเรียนราชสีมาวิทยาลัย |

ผู้เชี่ยวชาญด้านการจัดการเรียนการสอนวิชาวิทยาศาสตร์

- | | |
|-----------------------------|--|
| 1. อาจารย์สุรศักดิ์ บานแย้ม | ครูชำนาญการ กลุ่มสาระการเรียนรู้วิทยาศาสตร์
โรงเรียนสวนอนันต์ |
| 2. อาจารย์มะนิดา เกตุสุข | ครูชำนาญการ กลุ่มสาระการเรียนรู้วิทยาศาสตร์
โรงเรียนสวนอนันต์ |

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ – สกุล	นางสาวนพวรรณ ศรีเกตุ
วันเดือนปีเกิด	6 สิงหาคม 2522
สถานที่เกิด	กรุงเทพมหานคร
สถานที่อยู่ปัจจุบัน	48/5 หมู่ 5 ต.บางขุนกอง อ.บางกรวย จ.นนทบุรี 11130
ตำแหน่งหน้าที่การงานปัจจุบัน	ครูผู้ช่วย
สถานที่ทำงานปัจจุบัน	โรงเรียนสวนอนันต์
ประวัติการศึกษา	
พ.ศ. 2537	มัธยมศึกษาตอนต้น โรงเรียนสตรีวัดระฆัง
พ.ศ. 2540	มัธยมศึกษาตอนปลาย โรงเรียนสตรีวัดระฆัง
พ.ศ. 2544	วิทยาศาสตร์บัณฑิต (วท.บ.) วิชาเอกฟิสิกส์ มหาวิทยาลัยบูรพา
พ.ศ. 2545	ประกาศนียบัตรบัณฑิตวิชาชีพครู สถาบันราชภัฏบ้านสมเด็จเจ้าพระยา
พ.ศ. 2550	การศึกษามหาบัณฑิต (กศ.ม.) สาขาวิชาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ