

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาค่าความสอดคล้องของคะแนนสอบ ค่าความเที่ยง และค่าความตรงของแบบสอบ เมื่อตรวจให้คะแนนด้วยวิธีการที่ต่างกัน 2 วิธี คือ วิธีการให้น้ำหนักคะแนนของกิบบอนส์และคณะ กับวิธีของลอร์ด ในแบบสอบเลือกตอบวิชาวิทยาศาสตร์ โดยมีขั้นตอนในการวิจัยตามลำดับดังนี้

ประชากรและกลุ่มตัวอย่าง

ประชากร

ประชากรในการวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2533 ของโรงเรียนสังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ เขตจังหวัดปทุมธานี จำนวน 4,559 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างในการวิจัยครั้งนี้ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2533 ที่สุ่มมาจากประชากรโดยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) ผู้วิจัยดำเนินการสุ่มตัวอย่างตามขั้นตอนต่อไปนี้

1. สํารวจจำนวนโรงเรียน และจำนวนนักเรียนในแต่ละอำเภอมีดังนี้

1.1 อำเภอเมือง	จำนวนโรงเรียน 3 โรงเรียน	จำนวนนักเรียน 1,160 คน
1.2 อำเภอลองหลวง	จำนวนโรงเรียน 2 โรงเรียน	จำนวนนักเรียน 793 คน
1.3 อำเภอธัญบุรี	จำนวนโรงเรียน 3 โรงเรียน	จำนวนนักเรียน 1,617 คน
1.4 อำเภอลาดหลุมแก้ว	จำนวนโรงเรียน 2 โรงเรียน	จำนวนนักเรียน 266 คน
1.5 อำเภอลำลูกกา	จำนวนโรงเรียน 2 โรงเรียน	จำนวนนักเรียน 359 คน
1.6 อำเภอสามโคก	จำนวนโรงเรียน 2 โรงเรียน	จำนวนนักเรียน 230 คน
1.7 อำเภอหนองเสือ	จำนวนโรงเรียน 1 โรงเรียน	จำนวนนักเรียน 134 คน
รวม	จำนวนโรงเรียน 15 โรงเรียน	จำนวนนักเรียน 4,559 คน

## 2. คำนวณหาขนาดที่พอดีของกลุ่มตัวอย่างที่ใช้ในการวิจัยจากสูตร

$$n = \frac{N}{1 + Ne^2} \quad (\text{Taro Yamane, 1969})$$

เมื่อ  $n$  คือ ขนาดของกลุ่มตัวอย่าง

$N$  คือ ขนาดของประชากร

$e$  คือ ความคลาดเคลื่อนที่ยอมรับได้ (ผู้วิจัยยอมให้เกิดความคลาดเคลื่อนได้ 3 เปอร์เซ็นต์)

จากจำนวนประชากรทั้งหมด 4,559 คน สามารถคำนวณหาขนาดที่พอดีของกลุ่มตัวอย่างได้ 893 คน แต่เนื่องจากการประมาณค่าพารามิเตอร์ของข้อสอบ โดยการวิเคราะห์ข้อมูลด้วยโปรแกรม Logist 5 Version 2.5 ค่าพารามิเตอร์จะมีความคงที่มากขึ้น เมื่อใช้ผู้สอบจำนวนมากประมาณ 1,000 คน และมีข้อสอบจำนวนมาก เช่น ประมาณ 40 ข้อ (Wood and Other, 1976 อ้างถึงใน สุนัดาน์ สุกมลสันต์, 2528) ดังนั้นผู้วิจัยจึงสุ่มกลุ่มตัวอย่างแบบแบ่งชั้นได้กลุ่มตัวอย่างจำนวน 969 คน ตามตารางที่ 1 ซึ่งใกล้เคียงกับจำนวนที่ต้องใช้ในการวิเคราะห์ด้วยโปรแกรม Logist 5 Version 2.5 ดังที่กล่าวมาแล้ว

4. หลังจากได้คำนวณกลุ่มตัวอย่างในแต่ละอำเภอแล้ว จึงทำการสุ่มโรงเรียนตามสัดส่วน (ประมาณ 2:1) และสุ่มห้องเรียนตามสัดส่วน (ประมาณ 4:1) ของจำนวนห้องเรียนในชั้นมัธยมศึกษาปีที่ 3 จากโรงเรียนที่สุ่มมาได้ในครั้งแรก และใช้นักเรียนทั้งหมดในห้องเรียนเป็นกลุ่มตัวอย่าง รายละเอียดของจำนวนโรงเรียน และจำนวนห้องเรียน และจำนวนกลุ่มตัวอย่างแสดงใน ตารางที่ 1

ตารางที่ 1 แสดงกลุ่มตัวอย่างจริงที่ใช้ในการวิจัย จำแนกตามอำเภอและโรงเรียน

อำเภอ	โรงเรียน	จำนวนห้องที่มี	จำนวนห้องที่ล่ม	จำนวนนักเรียน
เมือง	คณะราษฎร์บำรุง	10	3	141
	ปทุมวิไล	10	3	138
คลองหลวง	ธรรมศาสตร์คลองหลวง	11	3	140
	วิทยาคม			
ัญบุรี	ัญบุรี	12	3	140
	ัญรัตน	12	3	146
ลาดหลุมแก้ว	บัวแก้วเกษร	7	2	87
ลำลูกกา	ลำลูกกา	7	2	99
สามโคก	ชัยสิทธิ์वास	4	1	40
หนองเสือ	หนองเสือวิทยาคม	4	1	38
รวม	9	77	21	969

#### เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ว.305 สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 เรื่องบรรยากาศ โลก ดวงดาว และอวกาศ ที่ผู้วิจัยสร้างขึ้น ตามหลักสูตรชั้นมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง 2533) โดยมีขั้นตอนในการดำเนินการสร้างและปรับปรุงดังนี้

1. ศึกษาหลักสูตรและวัตถุประสงค์ของบทเรียน จากหนังสือแบบเรียนวิชาวิทยาศาสตร์ ว. 305 และคู่มือครูของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี กระทรวงศึกษาธิการ
2. สร้างตารางวิเคราะห์เนื้อหาและพฤติกรรม (Two-way Specification Table) วิชาวิทยาศาสตร์ ว. 305 เรื่อง บรรยากาศ โลก ดวงดาว และอวกาศ โดยตารางนี้จะแสดงเนื้อหาและพฤติกรรมที่ต้องการวัด
3. นำตารางในข้อ 2 พร้อมทั้งจุดประสงค์เชิงพฤติกรรมและเวลาที่ใช้ในการสอนในแต่ละเนื้อหา ไปให้อาจารย์ผู้สอนวิชาวิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 ซึ่งผ่านการอบรมการสอนวิชาวิทยาศาสตร์ของสถาบันการสอนวิทยาศาสตร์และเทคโนโลยี จำนวน 5 ท่าน ให้นำนักความสำคัญแล้วนำมาปรับเป็นตารางโครงสร้างเฉลี่ยของน้ำหนักความสำคัญในแต่ละเนื้อหา
4. กำหนดจำนวนข้อกระทงในแต่ละเนื้อหาวิชา ตามน้ำหนักในตารางโครงสร้างเฉลี่ย ดังแสดงในตารางที่ 2

ศูนย์วิจัยทรัพยากรชีวภาพ  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2 จำนวนข้อกระทงของแต่ละเนื้อหาย่อยในแต่ละจุดประสงค์

เนื้อหา	จุดประสงค์	ความรู้ ความจำ	ความ เข้าใจ	นำไปใช้	วิเคราะห์	สังเคราะห์	ประเมินผล	รวม
1.	ส่วนประกอบของอากาศ	-	1	-	-	-	-	1
2.	สมบัติของอากาศ	3	9	3	-	1	-	16
3.	ลม	1	6	-	-	1	-	8
4.	อุตุนิยมวิทยากับชีวิตประจำวัน	-	1	-	-	-	1	2
5.	วัตถุในท้องฟ้ามีอะไรบ้าง	5	1	-	-	-	-	6
6.	มองท้องฟ้าด้วยกล้องโทรทรรศน์	8	4	2	-	-	-	14
7.	สู่อวกาศ	2	5	2	1	1	-	11
8.	ประโยชน์และความก้าวหน้า ของการสำรวจอวกาศ	-	1	-	-	-	-	1
9.	ประเทศไทยกับการสำรวจ อวกาศ	-	1	-	-	-	-	1
รวม		19	29	7	1	3	1	60

5. เขียนข้อกระทงตามจำนวนข้อในตารางที่ 2 โดยข้อกระทงที่เขียนขึ้นเป็นแบบสอบเลือกตอบ 5 ตัวเลือก จำนวน 60 ข้อ พร้อมทั้งคำเฉลยนำไปให้ผู้ทรงคุณวุฒิ โดยเป็นอาจารย์ผู้สอนวิชาวิทยาศาสตร์ ในชั้นมัธยมศึกษาตอนต้นมาไม่น้อยกว่า 5 ปี และผ่านการอบรมการสอนวิชาวิทยาศาสตร์ 5 จากสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยีมาแล้ว จำนวน 10 ท่าน เพื่อตรวจสอบความตรงเชิงเนื้อหา และตัดสินว่าแบบสอบที่ผู้วิจัยสร้างขึ้นตรงตามจุดประสงค์เชิงพฤติกรรมที่กำหนดไว้หรือไม่ จากนั้นนำไปให้ผู้ทรงคุณวุฒิ ซึ่งเคยเป็นผู้ผ่านการศึกษาวิชาการสร้างแบบทดสอบ และเคยผ่านการอบรมทางด้านการวัดผล การศึกษามาแล้ว จำนวน 3 ท่าน ตรวจสอบความตรงตามลำดับขั้นของความรู้ที่ต้องการวัดของข้อกระทง โดยถือเอาผลของการตัดสินของผู้ทรงคุณวุฒิที่มีความเห็นตรงกันอย่างน้อย 2 ใน 3 ท่าน ถือเป็นการวัดตามลำดับขั้นนั้น ๆ ของข้อสอบ (รายละเอียดแสดงในภาคผนวก) แล้วจัดทำเป็นแบบสอบจริง

6. ทำแบบสอบที่ได้จัดทำแล้วไปทดลองใช้ครั้งที่ 1 กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนมัธยมวัดดาวคอง จำนวน 50 คน โดยการทดสอบครั้งนี้มีวัตถุประสงค์เพื่อปรับปรุงภาษาที่ใช้ในข้อกระทง และเพื่อวิเคราะห์ข้อสอบรายข้อ (Item Analysis) โดยใช้เกณฑ์ของจอห์นสัน (Johnson) คัดเลือกข้อกระทงที่เหมาะสม ซึ่งใช้เกณฑ์ดังนี้

6.1 ข้อกระทงที่มีค่าความยาก ( $p$ ) ของคำตอบที่ถูกตั้งแต่ 0.20 ถึง .80

6.2 ข้อกระทงที่มีค่าอำนาจจำแนก ( $r$ ) ของคำตอบที่ถูก ตั้งแต่ 0.2 ขึ้นไป

โดยในการทดลองสอบครั้งที่ 1 พบว่ามีข้อสอบที่มีคุณสมบัติตามเกณฑ์จำนวน 52 ข้อ และมีข้อสอบที่มีคุณสมบัติต่ำกว่าเกณฑ์ 8 ข้อ ได้แก่ ข้อที่ 1, 3, 12, 18, 26, 28, 34, 38 แต่เนื่องจากข้อสอบที่ผู้วิจัยสร้างขึ้นนั้น สร้างจากการกำหนดจำนวนข้อตามน้ำหนักในตารางโครงสร้างเฉลี่ยที่กำหนดโดยครูผู้สอนซึ่งมีประสบการณ์ในการสอนมาแล้วดังกล่าวข้างต้น และเพื่อมิให้น้ำหนักของข้อสอบนั้นเปลี่ยนแปลง ผู้วิจัยจึงได้นำข้อสอบที่มีคุณสมบัติต่ำกว่าเกณฑ์ดังกล่าวมาปรับปรุงข้อคำถาม และตัวเลือกขึ้นใหม่ โดยไม่ได้ตัดข้อกระทงออกแต่อย่างใด

7. นำแบบสอบที่ได้ปรับปรุงแล้วในข้อที่ 6 ไปทดลองใช้ครั้งที่ 2 กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนทวีธาภิเศก จำนวน 104 คน เพื่อซักซ้อมดำเนินการสอบ ศึกษาเวลาที่ใช้ในการสอบ และปัญหาในด้านต่าง ๆ เพื่อความสมบูรณ์ของการสอบจริง

ปรากฏว่าแบบสอบฉบับนี้ใช้เวลาในการตอบประมาณ 60 นาที แล้วจึงนำมาวิเคราะห์หาค่าคุณภาพของข้อสอบเป็นรายข้อ ปรากฏว่า แบบสอบฉบับนี้มีค่าความยากอยู่ระหว่าง .23 ถึง .75 และมีค่าอำนาจจำแนกอยู่ระหว่าง .21 ถึง .51 และคำนวณค่าความเที่ยงจากสูตร คูเดอร์ริชาร์ดสัน 20 (KR - 20) ปรากฏว่าได้ค่าความเที่ยงเป็น .8826 จึงถือว่าแบบสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ฉบับนี้ มีค่าความเที่ยงอยู่ในระดับที่เชื่อถือได้ รายละเอียดของแบบสอบ แสดงในตารางที่ 3

ตารางที่ 3 เนื้อหาและพฤติกรรมที่วัดในแบบสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ ว. 305

เนื้อหา	พฤติกรรม	ความรู้ ความจำ ข้อที่	ความ เข้าใจ ข้อที่	การนำ ไปใช้ ข้อที่	การ วิเคราะห์ ข้อที่	การ สังเคราะห์ ข้อที่	การประ เมินผล ข้อที่	รวม
<b>บทที่ 13 บรรยากาศ</b>								
13.1	ส่วนประกอบของอากาศ		1					1
13.2	สมบัติของอากาศ							
13.2.1	ความหนาแน่น ของอากาศ	2	3					3
13.2.2	ความดันของ อากาศ		5,6,7			4		4
13.2.3	อุณหภูมิของอากาศ		8,10,11	9				4
13.2.4	ความชื้นของ อากาศ	12,13	14,15	16,17				6

ตารางที่ 3 (ต่อ)

เนื้อหา	พฤติกรรม	ความรู้	ความ	การนำ	การ	การ	การประ	รวม
		ความจำ ข้อที่	เข้าใจ ข้อที่	ไปใช้ ข้อที่	วิเคราะห์ ข้อที่	สังเคราะห์ ข้อที่	เมินผล ข้อที่	
13.3	ลม							
13.3.1	สภาพอากาศ เหนือพื้นดินและ พื้นน้ำ	-	18, 19					2
13.3.2	การเกิดลม	20	21, 22, 24, 25			23		6
13.4	อุตุนิยมวิทยากับชีวิตประจำวัน		27				26	2
<b>บทที่ 14 โลก ดวงดาวและอวกาศ</b>								
14.1	วัตถุในท้องฟ้ามีอะไรบ้าง							
14.1.1	ดาว	28, 29 30						3
14.1.2	วัตถุในระบบสุริยะ	31, 32	33					3
14.2	มองท้องฟ้าด้วยกล้อง โทรทรรศน์							
14.2.1	เลนส์นูนทำให้เกิด ภาพได้อย่างไร	34, 35	36, 39	41				8
14.2.2	กล้องโทรทรรศน์ หักเหแสงมีหลักการ อย่างไร	42, 43 44		45				3



ตารางที่ 3 (ต่อ)

เนื้อหา	ความรู้ ความจำ ข้อที่	ความ เข้าใจ ข้อที่	การนำ ไปใช้ ข้อที่	การ วิเคราะห์ ข้อที่	การ สังเคราะห์ ข้อที่	การประ เมินผล ข้อที่	รวม
14.2.3 กล้องโทรทรรศน์ ประเภทสะท้อนแสง	46		47				2
14.3 ลู่อากาศ							
14.3.1 แรงแม่เหล็กของ โลก	48		50		51	49	4
14.3.2 การเดินทางออกสู่ อวกาศ	52	53, 55 56, 57 58	54				7
14.4 ประโยชน์ของความ ก้าวหน้าของการสำรวจ อวกาศ		59					1
14.5 ประเทศไทยกับการสำรวจ อวกาศ		60					1
รวม	19	29	7	1	3	2	60

8. นำแบบสอบที่ได้ทดลองใช้ในข้อที่ 7 ซึ่งได้ผ่านการวิเคราะห์คุณภาพดังกล่าวมาแล้ว มาจัดทำเป็นแบบสอบจริงที่ใช้ในการวิจัยเป็น 2 ฉบับ โดยแบบสอบทั้ง 2 ฉบับนั้น ใช้ข้อกระทงและตัวเลือกชุดเดียวกัน แต่ต่างกันที่วิธีการตอบ กล่าวคือ

ฉบับที่ 1 เป็นแบบสอบที่นักเรียนสามารถเลือกตอบได้ตั้งแต่ 1 ถึง หลายตัวเลือก ตามวิธีการตรวจให้คะแนนของกิบบอนส์และคณะ

ฉบับที่ 2 เป็นแบบสอบฉบับที่นักเรียนเลือกตอบได้เพียงตัวเลือกเดียว เพื่อนำไปวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบตามทฤษฎีตอบสนองข้อสอบ (IRT) และนำไปคำนวณหาน้ำหนักคะแนนตามวิธีของลอร์ด

หลังจากนั้นจึงนำแบบสอบทั้ง 2 ฉบับไปใช้ทดลองจริงต่อไป

#### การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล โดยนำหนังสือรับรองจากบัณฑิตวิทยาลัยไปติดต่อขอความร่วมมือในการวิจัย จากโรงเรียนมัธยมศึกษาที่เป็นกลุ่มตัวอย่างของแต่ละอำเภอ โดยมีวิธีดำเนินการเก็บรวบรวมข้อมูลตามลำดับขั้นดังนี้

#### 1. ขั้นการเก็บรวบรวมจากกลุ่มตัวอย่างจริง

1.1 ติดต่อกับฝ่ายวิชาการของโรงเรียนและอาจารย์ผู้สอน เพื่อชี้แจงวัตถุประสงค์ของการวิจัย พร้อมทั้งนัดหมายเรื่อง วัน เวลา และสถานที่สอบ

1.2 ผู้วิจัยนำแบบสอบฉบับที่ 1 ไปทดสอบกับนักเรียนในแต่ละโรงเรียนด้วยตนเอง ตามวัน และเวลาที่กำหนดไว้ หลังจากนั้นเป็นเวลาประมาณ 2 สัปดาห์ ผู้วิจัยจึงได้นำแบบสอบฉบับที่ 2 ไปทดสอบกับนักเรียนกลุ่มเดิม โดยในการทดสอบครั้งหลังนี้ ผู้วิจัยไม่ได้แจ้งให้นักเรียนทราบล่วงหน้า และอธิบายให้นักเรียนฟังว่า ผลจากการทดสอบจะนำไปใช้ในการวิจัยเท่านั้น ไม่มีผลต่อการเรียนการสอนของนักเรียนแต่อย่างใด จึงคาดว่านักเรียนไม่ได้ไปศึกษาเพิ่มเติมมา ส่วนการบริหารแบบสอบ ผู้วิจัยขอความร่วมมือจากอาจารย์ผู้สอนให้ช่วยในการคุมสอบ

## 2. ชั้นการบันทึกข้อมูล

หลังจากได้ดำเนินการสอบแล้ว ผู้วิจัยนำเอากระดาษคำตอบมาตรวจสอบความเรียบร้อย และตรวจให้น้ำหนักคะแนน โดยมีวิธีการตรวจให้เป็น 2 แบบดังนี้

ฉบับที่ 1 การตรวจให้คะแนนโดยวิธีของกิบบอนส์และคณะ มีวิธีการดังนี้  
ถ้านักเรียนทำเครื่องหมาย (X) ตัวเลือกเดียวแล้วถูกต้องจะได้ 4 คะแนน แต่ถ้าผิดจะได้ -1 คะแนน

ถ้านักเรียนทำเครื่องหมาย (X) 2 ตัวเลือก แล้วมีตัวเลือกที่ถูกรวมอยู่ด้วย จะได้ 3 คะแนน แต่ถ้าไม่มีตัวเลือกที่ถูกรวมอยู่จะได้ -2 คะแนน

ถ้านักเรียนทำเครื่องหมาย (X) 3 ตัวเลือก แล้วมีตัวเลือกที่ถูกรวมอยู่จะได้ 2 คะแนน แต่ถ้าไม่มีตัวเลือกที่ถูกรวมอยู่จะได้ -3 คะแนน

ถ้านักเรียนทำเครื่องหมาย (X) 4 ตัวเลือก แล้วมีตัวเลือกที่ถูกรวมอยู่ด้วย จะได้ 1 คะแนน แต่ถ้าไม่มีตัวเลือกที่ถูกรวมอยู่จะได้ -4 คะแนน

ถ้านักเรียนเว้นว่างข้อใด หรือตอบทุกตัวเลือก จะได้ 0 คะแนน

ฉบับที่ 2 การตรวจให้น้ำหนักคะแนนตามวิธีของลอร์ด มีลำดับขั้นดังนี้ คือ ถ้าตอบถูกต้องจะได้ 1 คะแนน ถ้าตอบผิดหรือไม่ตอบจะได้ 0 คะแนน หลังจากนั้นจึงได้เตรียมข้อมูลลงเทป เพื่อนำไปวิเคราะห์หาค่าพารามิเตอร์อำนาจจำแนก (a) ค่าการเดา (c) จากการวิเคราะห์ด้วยโปรแกรม Logist 5 Version 2.5 วิเคราะห์หาค่าความยาก (p) จากทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory) แล้วนำค่าเหล่านั้นมาคำนวณหาค่าน้ำหนักคะแนนรายชื่อตามสูตรของลอร์ดต่อไป

3. ในการศึกษาเรื่องความตรงตามสภาพของแบบสอบ ผู้วิจัยใช้วิธีหาค่าสหสัมพันธ์แบบลำดับที่ (Rank Correlation) โดยหาความสัมพันธ์ระหว่างการจัดลำดับที่ของคะแนนสอบที่ได้จากวิธีการให้น้ำหนักคะแนนของกิบบอนส์และคณะ และวิธีของลอร์ด กับการจัดลำดับที่ความสามารถของนักเรียนโดยครูผู้สอน ซึ่งใช้เป็นตัวเกณฑ์ในการศึกษา

จากการศึกษาระเบียบกระทรวงศึกษาธิการว่าด้วย การประเมินผลการเรียนตาม หลักสูตรมัธยมศึกษาในปัจจุบัน กำหนดให้ครูผู้สอนเป็นผู้ประเมินผลการเรียนของนักเรียนเป็น ระยะ ๆ ตลอดภาคเรียน เพื่อศึกษาผลการเรียน เพื่อการจัดสอนซ่อมเสริมและเพื่อจัดสอบ ซ่อมให้กับนักเรียนที่ไม่สามารถสอบผ่านจุดประสงค์ใดจุดประสงค์หนึ่ง และนำผลการประเมิน ย่อยทั้งหมดไปประเมินผลรวมในปลายภาคอีกครั้ง และจากสภาพการเรียนการสอนใน วิทยาศาสตร์มัธยมศึกษาตอนต้น ที่จัดทำหลักสูตร โดย สวท. นั้น ผู้เรียนจะต้องเรียนภาค ทฤษฎี ควบคู่ไปกับการเรียนภาคปฏิบัติในห้องปฏิบัติการ (Lab) ซึ่งครูผู้สอนก็ต้องทำหน้าที่ สอนทั้งในภาคทฤษฎีและภาคปฏิบัติเช่นเดียวกัน จึงเป็นเหตุให้ครูผู้สอนได้มีโอกาสนั่งใกล้ชิด กับนักเรียน และสามารถสังเกตพฤติกรรมที่คาดว่าจะส่งผลไปยังผลสัมฤทธิ์ทางการเรียนได้ พฤติกรรมเหล่านั้น ได้แก่ ความตั้งใจและกระตือรือร้นในการเรียน การมาเรียนทันเวลา และสม่ำเสมอ การทำงานที่ครูมอบหมายได้เสร็จสิ้น การมีส่วนร่วมในกิจกรรมการเรียน เหล่านี้เป็นต้น ด้วยเหตุนี้ผู้วิจัย จึงได้ให้ครูผู้สอนใช้คุณสมบัติดังกล่าว อันได้แก่ ภาระหน้าที่ในการประเมินผลการเรียนทั้งการประเมินผลย่อยและการประเมินผลรวม การ สังเกตพฤติกรรมที่คาดว่าจะมีผลต่อผลสัมฤทธิ์ทางการเรียน และใช้คะแนนผลสัมฤทธิ์ทางการ เรียนในภาคเรียน เดิมมาประกอบการพิจารณาจัดลำดับที่ความสามารถด้านผลสัมฤทธิ์ของ นักเรียน ซึ่งจากเหตุผลดังกล่าวมาแล้วในข้างต้น จึงน่าจะกล่าวได้ว่า การจัดลำดับที่ความ สามารถของนักเรียน โดยครูผู้สอน น่าจะใช้เป็นเกณฑ์ในการหาความตรงตามสภาพที่เชื่อถือ ได้วิธีหนึ่ง

#### การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลสำหรับการวิจัยในครั้งนี้ ผู้วิจัยได้ทำการวิเคราะห์ด้วย เครื่องคอมพิวเตอร์ ของสถาบันการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย โดยใช้ โปรแกรม SPSS\* และโปรแกรม Logist 5 Version 2.5 และคำนวณค่าน้ำหนัก คะแนนของ ลอร์ด (Lord) ขึ้นตอนต่าง ๆ ในการวิเคราะห์ข้อมูลมีดังนี้

## 1. การวิเคราะห์เพื่อตรวจสอบคุณสมบัติของแบบสอบถามโมเดลโลจิสติก

### 1.1 จำนวนข้อสอบที่ตัดออก

เนื่องจากการวิเคราะห์ข้อสอบด้วยโปรแกรม Logist อาจประสบปัญหาในการประมาณค่าพารามิเตอร์ไม่ได้ เนื่องจากสาเหตุสำคัญ 2 ประการคือ (Winkersky อ้างถึงใน อาคุณ ช่างแก้วมณี, 2533: 65-66)

1.1.1 เงื่อนไข Non-Positive นั่นคือ ในการประมาณค่าด้วยโปรแกรม Logist จะต้องมีการประมาณซ้ำ ๆ กันหลายครั้ง ซึ่งใน Step ที่ 2 และ 3 จะมีค่าอนุพันธ์อันดับ 2 (Second Derivative) ปรากฏอยู่ในขบวนการประมาณค่า และค่าอนุพันธ์อันดับ 2 จะต้องมีค่าเป็นลบ

1.1.2 เงื่อนไข Zero-Determinant นั่นคือ ในการประมาณค่าพารามิเตอร์ด้วยโปรแกรม Logist จะต้องมีการหารด้วย Determinant ของ Matrix ของผลการตอบ  $n$  ข้อ ของคน  $N$  คน ถ้า Determinant มีค่าเป็นศูนย์ จะไม่สามารถประมาณค่าได้

ด้วยเหตุทั้ง 2 ประการ จึงจำเป็นต้องตัดข้อสอบที่มีปัญหาออก ทั้งนี้เนื่องจากอาจมีผลกระทบต่อค่าความตรงของค่าพารามิเตอร์ และอาจมีผลกระทบต่อค่าพารามิเตอร์อื่น ๆ ซึ่งในการวิจัยครั้งนี้ พบว่า ข้อสอบที่ไม่สามารถประมาณค่าได้ เนื่องจากสาเหตุสำคัญ 2 ประการข้างต้น จำนวน 5 ข้อ ได้แก่ ข้อที่ 6, 18, 26 และ 34, 49 จึงเหลือข้อสอบในการคำนวณค่าคะแนนทั้งสิ้น จำนวน 55 ข้อ

ผู้วิจัยได้นำข้อสอบที่เหลือไปตรวจสอบความตรงตามเนื้อหาอีกครั้ง พบว่า ข้อสอบที่ตัดออกนั้นกระจายไปตามเนื้อหาต่าง ๆ ข้อสอบที่เหลืออยู่ยังคงครอบคลุมถึงเนื้อหา และพฤติกรรมที่ต้องการวัดได้ ดังแสดงรายละเอียดในตารางที่ 4

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4 เนื้อหาและพฤติกรรมที่วัดในแบบสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ ว.305 หลังจากผ่านการวิเคราะห์ด้วยโปรแกรม Logist

เนื้อหา	พฤติกรรม	ความรู้ ความจำ ข้อที่	ความ เข้าใจ ข้อที่	การนำ ไปใช้ ข้อที่	การ วิเคราะห์ ข้อที่	การ สังเคราะห์ ข้อที่	การประ เมินผล ข้อที่	รวม
<b>บทที่ 13 บรรยากาศ</b>								
13.1	ส่วนประกอบของอากาศ		1					1
13.2	สมบัติของอากาศ							
13.2.1	ความหนาแน่น ของอากาศ	2	3					3
13.2.2	ความดันของ อากาศ		5, 7			4		3
13.2.3	อุณหภูมิของอากาศ		8, 10, 11	9				4
13.2.4	ความชื้นของ อากาศ	12, 13	14, 15	16, 17				6
13.3	ลม							
13.3.1	สภาพอากาศ เหนือพื้นดินและ พื้นน้ำ	-	19					1
13.3.2	การเกิดลม	20	21, 22, 24, 25			23		6
13.4	อุตุนิยมวิทยากับชีวิตประจำวัน		27					1

ตารางที่ 4 (ต่อ)

เนื้อหา	พหุคูณ ความรู้ ความจำ ข้อที่	ความ เข้าใจ ข้อที่	การนำ ไปใช้ ข้อที่	การ วิเคราะห์ ข้อที่	การ สังเคราะห์ ข้อที่	การประ เมินผล ข้อที่	รวม
<b>บทที่ 14 โลก ดวงดาวและอวกาศ</b>							
14.1 วัตถุในท้องฟ้ามีอะไรบ้าง							
14.1.1 ดาว	28, 29 30						3
14.1.2 วัตถุในระบบสุริยะ	31, 32	33					3
14.2 มองท้องฟ้าด้วยกล้อง โทรทรรศน์							
14.2.1 เลนส์ทำให้เกิด ภาพได้อย่างไร	35 37, 38	36, 39 40	41				7
14.2.2 กล้องโทรทรรศน์ หักเหแสงมีหลักการ อย่างไร	42, 43 44		45				3
14.2.3 กล้องโทรทรรศน์ ประเภทสะท้อนแสง	46		47				2
14.3 สู่อวกาศ							
14.3.1 แรงโน้มถ่วงของ โลก	48		50		51		3
14.3.2 การเดินทางออกสู่อวกาศ	52	53, 55 56, 57 58	54				7

ตารางที่ 4 (ต่อ)

เนื้อหา	เหตุการณ์	ความรู้ ความจำ ข้อที่	ความ เข้าใจ ข้อที่	การนำ ไปใช้ ข้อที่	การ วิเคราะห์ ข้อที่	การ สังเคราะห์ ข้อที่	การประ เมินผล ข้อที่	รวม
14.4 ประโยชน์ของความ ก้าวหน้าของการสำรวจ อวกาศ			59					1
14.5 ประเทศไทยกับการสำรวจ อวกาศ			60					1
รวม		18	27	7	1	3	-	55

ศูนย์วิจัยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



## 1.2 การวัดเพียงคุณลักษณะเดียวของแบบสอบ

ตรวจสอบได้โดยเ้าคะแนนผลการตอบแบบสอบ มาทำการวิเคราะห์ตัวประกอบสำคัญ (Principle Components) แล้วหมุนแกนด้วยวิธีแวนริแมกซ์ (Varimax) โดยในการตรวจสอบคุณสมบัติของการวัดเพียงคุณลักษณะเดียวนั้นพิจารณาจากค่าไอเกน (Eigen Value) ถ้าหากค่าไอเกนของตัวประกอบที่ 1 ของแบบสอบมีค่าสูงกว่า ค่าไอเกนของตัวประกอบตัวที่ 2 อย่างมาก และค่าไอเกนของตัวประกอบที่สองมีค่าสูงกว่าค่าไอเกนของตัวประกอบตัวถัด ๆ ไปเพียงเล็กน้อย สามารถกล่าวได้ว่า แบบสอบนั้นวัดเพียงคุณลักษณะเดียว (Christofersson (1975), Indow and Samejima (1962), Muten (1977) อ้างถึงใน Lord, 1980) ซึ่งการตรวจสอบนี้ก็เพื่อเป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีสองตอข้อสอบ (IRT) ที่กำหนดว่า แบบสอบที่จะใช้วิเคราะห์ด้วยโมเดลต่าง ๆ ในทฤษฎีนี้ จะต้องมึคุณสมบัติในการวัดเพียงคุณลักษณะเดียว และมีความเป็นอิสระในการตอบสนองต่อข้อสอบ ซึ่งในการตรวจสอบความเป็นอิสระในการตอบสนองข้อสอบนั้นให้พิจารณาว่า ถ้าแบบสอบมีคุณสมบัติการวัดเพียงคุณลักษณะเดียวแล้ว ก็จะมีคุณสมบัติในการตอบสนองข้อสอบด้วย (Hambleton and Swaminatan, 1985 ; Lord, 1980) ดังนั้น ผู้วิจัยจึงทำการทดสอบคุณสมบัติดังกล่าว เพื่อให้แน่ใจว่าแบบสอบแบบนี้เป็นไปตามข้อตกลงเบื้องต้นของการใช้ Three-Parameter Logistic Model ในทฤษฎีสองตอข้อสอบ (IRT) ดังแสดงค่าไอเกน (Eigen Value) ในตารางที่ 5

ศูนย์วิจัยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย  
จุฬาลงกรณ์มหาวิทยาลัย

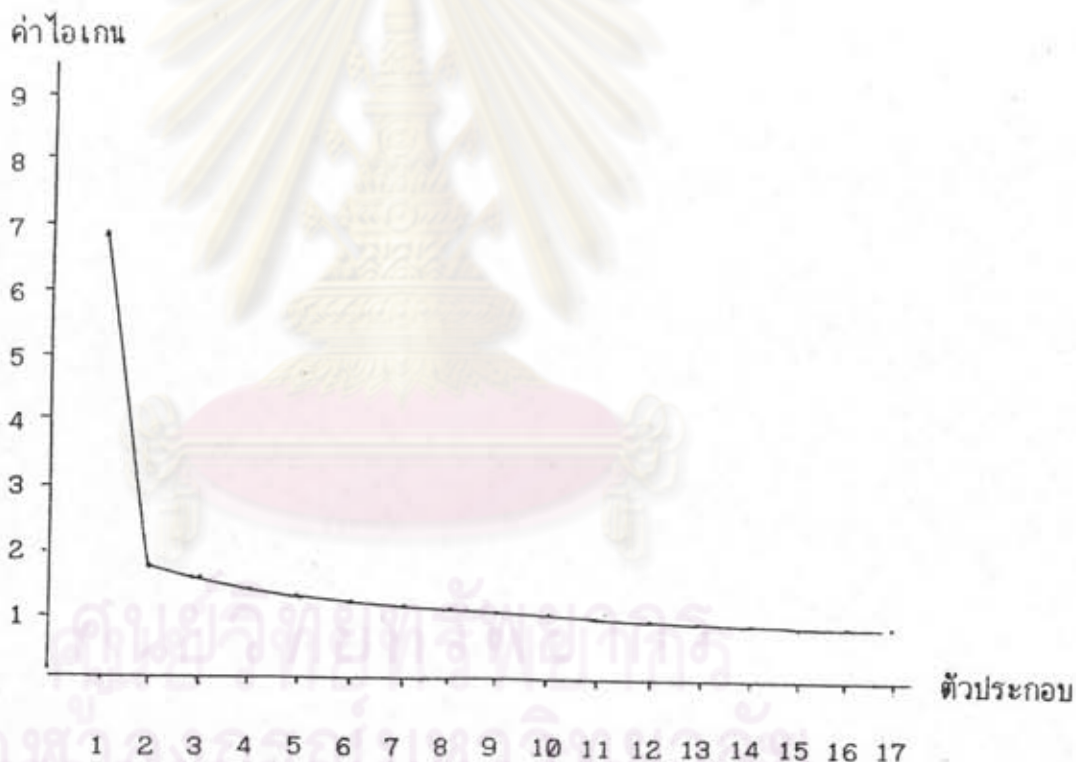
ตารางที่ 5 แสดงค่าไอเกน (Eigen Value) ของแต่ละตัวประกอบของแบบสอบถามที่มี  
ค่าไอเกนเกิน 1.00

ตัวประกอบที่	ค่าไอเกน
1	6.86794
2	1.85284
3	1.65367
4	1.45663
5	1.40601
6	1.35828
7	1.30499
8	1.24817
9	1.22210
10	1.20495
11	1.16745
12	1.14046
13	1.08842
14	1.07740
15	1.05699
16	1.03259
17	1.01797

จากตารางที่ 5 พบว่า แบบสลับฉบับนี้ เมื่อทำการวิเคราะห์ตัวประกอบแล้วจะได้ตัวประกอบที่มีค่าไอเกน เกิน 1.00 อยู่ 17 ตัวประกอบ โดยตัวประกอบตัวที่ 1 มีค่าไอเกนสูงสุด 6.86794 ตัวประกอบตัวที่ 2 มีค่าไอเกน 1.85284 และตัวประกอบตัวต่อมามีค่าไอเกนลดลงเรื่อย ๆ ตามลำดับ โดยตัวประกอบตัวที่มีค่าไอเกนต่ำสุด คือ ตัวประกอบที่ 17 มีค่าไอเกน 1.01797

ผู้วิจัยได้นำค่าไอเกนมาเขียนเป็นกราฟเพื่อแสดงความแตกต่างของแต่ละตัวประกอบ

ภาพที่ 2 แสดงค่าไอเกนของตัวประกอบที่มีเกิน 1.00 ของแบบสลับ



พบว่า ค่าไอเกนของตัวประกอบที่ 1 มีค่าสูงกว่าตัวประกอบที่ 2 อย่างมาก โดยจะสังเกตเห็นจากค่าไอเกนของตัวประกอบที่ 1 มาถึงตัวประกอบที่ 2 ลดลงอย่างรวดเร็ว แต่จากตัวประกอบที่ 2 มาถึงตัวประกอบที่ 3 และอื่น ๆ ตามลำดับ ค่าไอเกนจะลดลง

เพียงเล็กน้อย ลักษณะดังกล่าวนี้ (Christoffersson 1975, Indow and Somejima (1962), Muten (1977) อ้างถึงใน Lord, 1980) กล่าวว่า เป็นแบบสอบที่วัดคุณลักษณะเดียว ดังนั้น แบบสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ (ว.305) ฉบับนี้น่าจะมีคุณสมบัติการวัดเพียงลักษณะเดียว และยังมีผลในคุณสมบัติของความเป็นอิสระในการตอบข้อสอบอีกด้วย (Lord, 1980; Hambleton and Swaminatan, 1985)

2. การคำนวณหาน้ำหนักคะแนนสอบรายข้อ (Optimal Scoring Weight) โดยสูตรของลอร์ด (Lord) มีวิธีการดังนี้

2.1 นำคะแนนจากแบบสอบฉบับที่ 2 มาคำนวณด้วยโปรแกรม SPSS\* เพื่อหาค่าความยาก (p) ตามทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory) ด้วยสูตร

$$P = \frac{\text{จำนวนนักเรียนที่ตอบถูก}}{\text{จำนวนนักเรียนทั้งหมด}}$$

เมื่อ P คือ ค่าความยาก (Difficulty)

2.2 นำคะแนนจากแบบสอบฉบับที่ 2 มาคำนวณด้วยโปรแกรม Logist 5 Version 2.5 เพื่อประมาณค่าพารามิเตอร์ของข้อสอบ คือ ค่าอำนาจจำแนกของข้อสอบ ( $a_i$ ) ค่าการเดาของข้อสอบ ( $c_i$ ) ตาม Three-Parameter Logistic Model จากสูตร

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

ในเมื่อ  $p_i(\theta)$  คือ ความน่าจะเป็นที่ผู้สอบที่มีระดับความสามารถ ( $\theta$ ) จะตอบข้อสอบข้อที่  $i$  ได้ถูกต้อง

( $\theta$ ) คือ ระดับความสามารถที่แท้จริงของผู้สอบ

$a_i$  คือ ค่าอำนาจจำแนกของข้อสอบข้อที่  $i$

$b_i$  คือ ค่าความยากของข้อสอบข้อที่  $i$

$c_i$  คือ ค่าการเดาของข้อสอบข้อที่  $i$

$D$  คือ ค่า Scaling Factor มีค่า = 1.7

$e$  คือ ค่าคงที่มีค่าเท่ากับ 2.71828

2.3 นำค่าความยาก ( $p$ ) ค่าอำนาจจำแนก ( $a_i$ ) และค่าการเดาของข้อสอบ ( $c_i$ ) ในแต่ละข้อกระทง มาคำนวณหาน้ำหนักคะแนนสอบรายข้อ (Optimal Scoring Weight) จากสูตรของลอร์ด (Lord, 1980)

$$W_i = \frac{a_i \cdot p_i - c_i}{1 - c_i \cdot p_i}$$

ในเมื่อ  $W_i$  คือ ค่าคะแนน

$a_i$  คือ ค่าอำนาจจำแนกข้อสอบข้อที่  $i$

$p_i$  คือ ค่าความยากที่วิเคราะห์จากทฤษฎีมาตรฐานดั้งเดิม

$c_i$  คือ ค่าการเดาข้อสอบข้อที่  $i$

### 3. การวิเคราะห์ข้อมูลในการวิจัย

น้ำหนักคะแนนสอบที่ได้จากการตรวจให้นำหนักคะแนนทั้ง 2 วิธี คือวิธีของกิบบอนส์ และคณะ และวิธีการให้คะแนนของลอร์ด ในข้อ 2.3 มาวิเคราะห์หาค่าสถิติต่าง ๆ ดังนี้

3.1 คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนสอบระหว่างวิธีการให้น้ำหนักคะแนนกิบบอนส์และคณะ กับวิธีการของลอร์ด โดยใช้สูตร (Guilford, 1976)

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

เมื่อ  $r_{xy}$  คือ ความสัมพันธ์ระหว่างวิธีการให้น้ำหนักคะแนนของกิบบอนส์และ  
คะแนนกับวิธีของลอร์ด

$x$  คือ คะแนนสอบที่ได้จากการให้น้ำหนักคะแนนโดยวิธีของกิบบอนส์  
และคะแนน

$y$  คือ คะแนนสอบที่ได้จากการให้คะแนนโดยวิธีของลอร์ด

$N$  คือ จำนวนนักเรียน

3.2 ทดสอบสัมประสิทธิ์สหสัมพันธ์ระหว่างวิธีการให้คะแนน โดยใช้สถิติทดสอบ  
ที่ (t-test) จากสูตร (Ferguson, 1981)

$$t = \frac{r_{xy} \sqrt{N-2}}{\sqrt{(1-r_{xy}^2)}} \quad , \quad df = N-2$$

เมื่อ  $t$  แทน สถิติทดสอบที่

$r_{xy}$  แทน สัมประสิทธิ์สหสัมพันธ์ระหว่างวิธีการให้น้ำหนักคะแนน  
ของกิบบอนส์และคะแนน กับวิธีของลอร์ด

$N$  แทน จำนวนกลุ่มตัวอย่าง

3.3 คำนวณค่าความเที่ยง (Reliability) ของแบบสอบในการให้  
น้ำหนักคะแนนตามวิธีของกิบบอนส์และคะแนน กับวิธีของลอร์ด โดยใช้สูตรแอลฟาของครอนบัค  
(Cronbach' Alpha) (Mehren and Lehman, 1984)

$$\alpha = \frac{n}{n-1} \left[ \frac{1 - \sum s^2_i}{s^2_x} \right]$$

- เมื่อ  $\alpha$  แทน ค่าความเที่ยงของแบบสอบ  
 $n$  แทน จำนวนข้อกระทง  
 $\sum s^2_i$  แทน ผลรวมความแปรปรวนของการกระจายของคะแนนในข้อกระทง  
 $s^2_x$  แทน ความแปรปรวนของคะแนนสอบทั้งหมด

3.4 เปรียบเทียบค่าความแตกต่างของค่าความเที่ยงของแบบสอบระหว่างวิธีการให้น้ำหนักคะแนนต่างกัน ระหว่างวิธีของกิบบอนส์และคณะ กับวิธีของลอร์ด โดยวิธีการของพิตแมน (Pitman) ที่เสนอโดย Feldt (Feldt, 1987: อ้างถึงใน นิลิษฐ์ ตันเทวนิช, 2532)

$$t = \frac{(r_{\text{tt1}} - r_{\text{tt2}}) \sqrt{(N-2)}}{\sqrt{4(1-r_{\text{tt1}})(1-r_{\text{tt2}})(1-r^2_{x \times 2})}}, \quad (df = n - 2)$$

- เมื่อ  $t$  คือ การแจกแจงแบบที  
 $r_{\text{tt1}}$  คือ ค่าความเที่ยงของแบบสอบเมื่อให้น้ำหนักคะแนน โดยวิธีของกิบบอนส์และคณะ  
 $r_{\text{tt2}}$  คือ ค่าความเที่ยงของแบบสอบเมื่อให้น้ำหนักคะแนน โดยวิธีของลอร์ด  
 $N$  คือ จำนวนกลุ่มตัวอย่าง  
 $r_{x \times 2}$  คือ สหสัมพันธ์ระหว่างวิธีการให้น้ำหนักคะแนน ตามวิธีของกิบบอนส์และคณะ กับวิธีของลอร์ด

3.5 หาความตรงตามสภาพ (Concurrent Validity) ของแบบสอบ โดยหาความสัมพันธ์ระหว่างผลการจัดลำดับความสามารถ (Rank order) ของนักเรียน โดยครูผู้สอน กับลำดับคะแนนที่ได้จากการตอบแบบสอบทั้งสองของนักเรียน โดยใช้สูตรสหสัมพันธ์แบบลำดับที่ (Rank Correlation) จากสูตร

$$r = 1 - \frac{6 \sum_{i=1}^N D^2}{N(N^2 - 1)}$$

เมื่อ D คือ ผลต่างของอันดับความสามารถ  
N คือ จำนวนนักเรียน

3.6 ทดสอบนัยสำคัญทางสถิติของค่าสัมประสิทธิ์ความตรงตามสภาพของแบบสอบ ซึ่งได้จากวิธีการให้น้ำหนักคะแนนของกิบบอนส์และคณะ กับวิธีการของลอร์ด ด้วยสถิติทดสอบที (t-test) จากสูตร (Ferguson, 1981)

$$t = \frac{r \sqrt{N - 2}}{\sqrt{1 - r^2}}$$

เมื่อ t แทน สถิติทดสอบที  
r แทน ค่าสัมประสิทธิ์ความตรงของแบบสอบ  
N แทน จำนวนนักเรียนในแต่ละห้องเรียน

3.7 เปรียบเทียบค่าความแตกต่างของค่าความตรงตามสภาพของแบบสอบระหว่างวิธีการให้น้ำหนักคะแนนตามวิธีของกิบบอนส์และคณะ กับวิธีของลอร์ด โดยใช้วิธีการทดสอบของวิลคอกซอน (The Wilcoxon Matched-Pairs Signed Rank test) มีวิธีการดังนี้คือ (Siegel, 1956)



- 3.7.1 หาความแตกต่างระหว่างค่าความตรง 2 กลุ่ม โดยคิดเครื่องหมายด้วย
- 3.7.2 จัดลำดับที่ความแตกต่างโดยคิดเฉพาะค่าตัวเลข และเมื่อได้ลำดับที่แล้วให้ติดเครื่องหมายไว้
- 3.7.3 หาค่า  $T$  โดยการรวมลำดับที่ของเครื่องหมายที่มีจำนวนน้อยกว่า
- 3.7.4 นับจำนวน  $N$  ซึ่งเท่ากับจำนวนลำดับที่ทั้งหมด แล้วนำค่า  $N$  ไปเปิดตารางของ Wilcoxon เพื่อหาค่าวิกฤติของ  $T$
- การตัดสินใจ ถ้าค่า  $T$  ที่นับได้มีค่าเท่ากับหรือน้อยกว่าค่า  $T$  จากตารางตามระดับนัยสำคัญที่ตั้ง จะไม่ยอมรับสมมุติฐานศูนย์



ศูนย์วิจัยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย  
จุฬาลงกรณ์มหาวิทยาลัย