



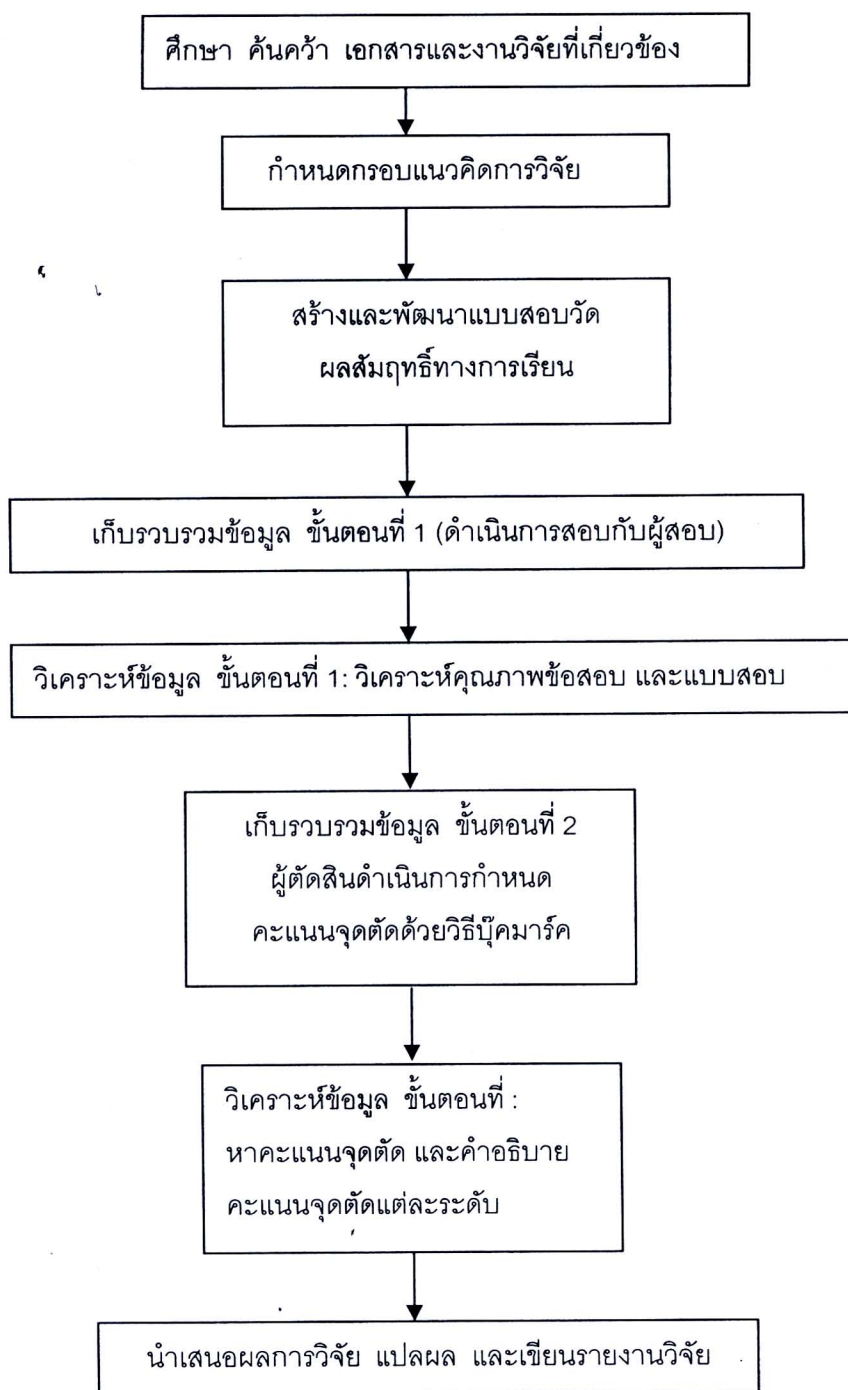
บทที่ 3 วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อสร้างแบบสอบผลสัมฤทธิ์ทางการเรียนวิชา การวัดและประเมินผลในชั้นเรียน ระดับปริญญาตรี มหาวิทยาลัยบูรพา เพื่อตรวจสอบความยาก อำนาจจำแนก ความตรงและความเที่ยงของแบบสอบผลสัมฤทธิ์ทางการเรียนวิชา การวัดและประเมินผลในชั้นเรียนที่สร้างขึ้น และเพื่อหาคะแนนจุดตัดของแบบสอบที่สร้างขึ้นโดยการกำหนดมาตรฐานด้วยวิธีบูคมาร์ค ซึ่งมีรายละเอียดของการดำเนินการวิจัยดังนี้

ขั้นตอนการดำเนินการวิจัย

1. ศึกษาเอกสาร งานวิจัยที่เกี่ยวข้องกับการกำหนดมาตรฐาน เช่น ความหมาย วิธีการกำหนดมาตรฐานทั้งแบบดั้งเดิม และแบบใหม่ วิธีการพัฒนาแบบสอบวัดผลสัมฤทธิ์ทางการเรียน
2. กำหนดกรอบแนวคิดการวิจัย
3. กำหนดประชากร และกลุ่มตัวอย่าง แบ่งออกเป็น 2 กลุ่มคือ
 - 3.1 ผู้ตัดสิน คือ อาจารย์ที่สอนรายวิชา การวัดและประเมินผลในชั้นเรียน
 - 3.2 ผู้สอบ คือ นิสิตชั้นปีที่ 3 คณะศึกษาศาสตร์ ที่ลงเรียนรายวิชา 400204 การวัดและประเมินในชั้นเรียน ปีการศึกษา 2553
4. สร้างและพัฒนาแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผลการศึกษา แบ่งเป็น 2 ฉบับคือ ฉบับที่ 1 แบ่งเป็น 2 ตอน คือแบบเลือกตอบ 4 ตัวเลือก จำนวน 60 ข้อ และแบบเขียนตอบ จำนวน 1 ข้อ และฉบับที่ 2 แบ่งเป็น 2 ตอน คือแบบเลือกตอบ 4 ตัวเลือก จำนวน 50 ข้อ และแบบเขียนตอบ จำนวน 1 ข้อ
5. ดำเนินการเก็บรวบรวมข้อมูล ระยะที่ 1 ดำเนินการสอบกับผู้สอบ คือ นิสิตคณะศึกษาศาสตร์ที่ลงทะเบียนเรียนวิชา 400204 การวัดและประเมินผลในชั้นเรียน ปีการศึกษา 2553
6. ดำเนินการวิเคราะห์ข้อมูล ระยะที่ 1 โดยการวิเคราะห์ข้อสอบเพื่อหาค่าความยาก อำนาจจำแนก ของข้อสอบโดยใช้โปรแกรม SPSS และ EXCEL โดยใช้สูตรในทฤษฎีการทดสอบแบบดั้งเดิม
7. นำค่าความยากที่ได้จากข้อสอบเลือกหลายคำตอบ และข้อสอบอัตนัยมาคำนวณหาค่าความสามารถของผู้สอบ (θ) ในแต่ละข้อ โดยใช้โปรแกรม Xcalibre และโปรแกรม Multilog เพื่อเป็นข้อมูลในคู่มือจัดเรียงข้อสอบ

8. ดำเนินการเก็บรวบรวมข้อมูล ระยะที่ 2 กำหนดมาตรฐานด้วยวิธีบูคมาร์คโดยกลุ่มผู้ตัดสินที่เป็นอาจารย์สอนวิชาการวัดและประเมินผลในชั้นเรียน
 9. ดำเนินการวิเคราะห์ข้อมูล หาคะแนนจุดตัด 7 ระดับพร้อมคำอธิบายแต่ละระดับ
 10. นำเสนอผลการวิจัย แปลผล และเขียนรายงานการวิจัย
- ขั้นตอนการดำเนินงานวิจัยทั้งหมดข้างต้น เสนอเป็นภาพที่ 3-1 ได้ดังนี้



ภาพที่ 3-1 ขั้นตอนการดำเนินงานวิจัย

การวิจัยเรื่อง การพัฒนาแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผลในชั้นเรียน โดยกำหนดคะแนนมาตรฐานด้วยวิธีปฏิกิริยา เป็นการวิจัยประยุกต์ (Applied Research) แบ่งการศึกษาเป็น 2 ขั้นตอนดังนี้

ขั้นตอนที่ 1 การสร้างและพัฒนาแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินในชั้นเรียน สำหรับนิสิตปริญญาตรี คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา

ขั้นตอนที่ 2 การหาคะแนนจุดตัดด้วยการกำหนดมาตรฐานด้วยวิธีปฏิกิริยา

ขั้นตอนที่ 1 การสร้างและพัฒนาแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินในชั้นเรียน สำหรับนิสิตปริญญาตรี คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา
วัตถุประสงค์ : 1) เพื่อสร้างแบบสอบผลสัมฤทธิ์ทางการเรียนวิชา การวัดและประเมินผลในชั้น เรียนระดับปริญญาตรี มหาวิทยาลัยบูรพา

2) เพื่อตรวจสอบความยาก อำนาจจำแนก ความตรง และความเที่ยงของแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผลในชั้นเรียนที่สร้างขึ้น

การดำเนินการในขั้นตอนนี้ มีวิธีดำเนินการวิจัยดังนี้

1. กำหนดประชากรและกลุ่มตัวอย่าง

1.1 ประชากร นิสิตคณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา ที่ลงทะเบียนเรียนวิชา 400204 การวัดและประเมินผลในชั้นเรียน ปีการศึกษา 2553 จำนวน 990 คน
(<http://reg.buu.ac.th/registrar/home.asp>)

1.2 กลุ่มตัวอย่าง นิสิตคณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา ที่ลงทะเบียนเรียนวิชา 400204 การวัดและประเมินผลในชั้นเรียน ภาคเรียนที่ 2 ปีการศึกษา 2553 โดยวิธีการสุ่มแบบยกกกลุ่ม (Cluster Sampling) จำนวน 667 คน

สูตรหาขนาดกลุ่มตัวอย่าง กรณีต้องการศึกษาค่าสัดส่วน ผู้วิจัยใช้สูตรของ โคชแรน (Cochran, 1963 อ้างถึงใน ศิริชัย กาญจนวาสี, 2545)

$$n = \frac{NZ^2P(1-P)}{NE^2 + Z^2P(1-P)}$$

เมื่อ

n = ขนาดของกลุ่มตัวอย่าง

N = ขนาดของประชากร

Z = ค่าที่กำหนดจากค่าความเชื่อมั่นที่ผู้วิจัยต้องการจะใช้
เพื่อการสรุปผล ซึ่งระบุเป็นร้อยละ โดยทั่วไปแล้ว เช่น

ถ้าความเชื่อมั่น 99% จะเท่ากับ 2.58 หรือประมาณ 3

P = ค่าสัดส่วนของลักษณะที่ต้องการศึกษา

$P(1 - P)$ = ค่าความแปรปรวนของลักษณะที่ต้องการศึกษา
มีค่าสูงสุดเป็น 0.25

E = ค่าความคลาดเคลื่อนที่ผู้วิจัยยอมรับในการสรุปผล

ในการวิจัยนี้ผู้วิจัยกำหนดระดับความเชื่อมั่น 99% ความคลาดเคลื่อน 5% และขนาดประชากร 990 คน

$$n = \frac{990(3)^2(0.25)}{990(0.05)^2 + (3)^2(0.25)}$$
$$= 471.42$$
$$\approx 471$$

ดังนั้น จะเห็นว่ากลุ่มตัวอย่างขั้นต่ำคือ 471 คน

ตารางที่ 3-1 รายชื่อสาขาวิชาและจำนวนกลุ่มตัวอย่างนิสิตที่ลงทะเบียนเรียนรายวิชาการวัดและประเมินผลในชั้นเรียน ที่ใช้เก็บข้อมูลเพื่อทดลองแบบสอบครั้งที่ 1 ครั้งที่ 2 และครั้งที่ 3

สาขาวิชา	กลุ่มเรียนที่	ประชากร	กลุ่มตัวอย่าง
		จำนวนนิสิต	จำนวนนิสิต
การทดลองแบบสอบครั้งที่ 1			
1) เทคโนโลยีการศึกษา	1	301	5
รวม		301	5
การทดลองแบบสอบครั้งที่ 2			
1) การสอนสังคม	1	26	26
2) การสอนภาษาอังกฤษ	1	21	21
รวม		47	47

ตารางที่ 3-1 (ต่อ)

สาขาวิชา	กลุ่มเรียนที่	ประชากร	กลุ่มตัวอย่าง
		จำนวนนิสิต	จำนวนนิสิต
การทดลองแบบสอบครั้งที่ 3			
1) การสอนพลศึกษา	1	43	-
2) การสอนพลศึกษา	2	37	37
3) เทคโนโลยีการศึกษา	1	42	42
4) เทคโนโลยีการศึกษา	2	36	36
5) เทคโนโลยีการศึกษา	3	52	52
6) เทคโนโลยีการศึกษา	4	50	50
7) เทคโนโลยีการศึกษา	5	42	42
8) เทคโนโลยีการศึกษา	6	50	50
9) การสอนภาษาอังกฤษ	1	36	36
10) สังคมศึกษา ศาสนาและวัฒนธรรม	1	47	47
11) การสอนภาษาจีน	1	40	40
12) การสอนภาษาจีน	2	46	-
13) การสอนฟิสิกส์	1	36	-
14) การสอนปฐมวัย	1	47	47
15) การสอนปฐมวัย	2	33	33
16) การสอนปฐมวัย	3	35	35
17) เทคโนโลยีอุตสาหกรรม	1	46	46
18) เทคโนโลยีอุตสาหกรรม	2	52	-
19) เทคโนโลยีอุตสาหกรรม	3	48	-
20) การสอนศิลปะและนาฏยศิลป์	1	25	25
21) การสอนภาษาญี่ปุ่น	1	49	49
22) การสอนเคมี	1	40	-
23) การสอนภาษาไทย	1	58	-
รวม		990	667

2. ศึกษาคำอธิบายรายวิชา และเนื้อหาในหนังสือเกี่ยวกับการวัดและประเมินการศึกษา

3. วิเคราะห์วัตถุประสงค์การเรียนรู้เป็นรายข้อ และเนื้อหาวิชา การวัดและประเมินในชั้นเรียน เพื่อสร้างตารางโครงสร้างเนื้อหา จากนั้นนำตารางโครงสร้างเนื้อหาไปให้อาจารย์ที่สอนในรายวิชาดังกล่าว โดยเป็นอาจารย์มหาวิทยาลัย ที่มีประสบการณ์การสอนในวิชาดังกล่าวไม่ต่ำกว่า 3 ปี จำนวน 3 คน กำหนดน้ำหนักเนื้อหา และพิจารณาความครอบคลุมของเนื้อหา ตลอดจนความเหมาะสมของวัตถุประสงค์การเรียนรู้

4. สร้างข้อสอบแบบเลือกคำตอบ 4 ตัวเลือก และข้อสอบอัตนัย ให้สอดคล้องกับตารางโครงสร้างกำหนดน้ำหนักเนื้อหา ตามระดับพุทธิพิสัยของแนวคิดของ บลุ่ม คือ ระดับความรู้ ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า

5. ตรวจสอบความตรงเชิงเนื้อหาของแบบสอบ โดยผู้วิจัยนำแบบสอบมาให้ผู้เชี่ยวชาญจำนวน 3 คน ที่มีประสบการณ์ในการสอนมาไม่ต่ำกว่า 3 ปี พิจารณาดัดสินความสอดคล้องระหว่างข้อสอบและจุดประสงค์การเรียนรู้

6. การนำแบบสอบไปทดลองใช้เพื่อตรวจสอบคุณภาพของแบบสอบ ผู้วิจัยดำเนินการทดลองใช้แบบสอบ 3 ครั้ง ดังนี้

6.1 การทดลองใช้แบบสอบครั้งที่ 1 เพื่อตรวจสอบด้านภาษาเกี่ยวกับคำชี้แจงในการสอบ และข้อคำถาม ว่านักเรียนมีความเข้าใจตรงกันหรือไม่ มีภาษาตรงจุดไหนมีความกำกวมเพื่อนำไปปรับปรุงแก้ไข ตลอดจนเพื่อนำข้อมูลที่ได้มากำหนดเวลาในการสอบที่มีความเหมาะสมสำหรับการดำเนินงานมีดังนี้

(1) ผู้วิจัยสุ่มตัวอย่างอย่างง่ายที่มีลักษณะคล้ายประชากรได้นิสิตจำนวน 10 คน โดยเป็นนิสิตสาขาเทคโนโลยีการศึกษา ภาคเรียนที่ 1 ปีการศึกษา 2553

(2) นำนิสิตมาทำแบบสอบโดยจับเวลาในการทำข้อสอบของนิสิตแต่ละคน เพื่อนำเวลาที่นักเรียนแต่ละคนทำมาหาค่าเฉลี่ยในการทำแบบสอบ

(3) เมื่อนิสิตทำแบบสอบเสร็จ ผู้วิจัยสัมภาษณ์นิสิตแต่ละคนเกี่ยวกับภาษาที่ใช้ในแบบทดสอบว่ามีคำสั่ง และข้อคำถามชัดเจน เข้าใจตรงกันหรือไม่

(4) ปรับปรุงภาษาที่ใช้ในแบบทดสอบ

6.2 การทดลองใช้แบบสอบครั้งที่ 2 เพื่อนำผลการสอบไปคำนวณเพื่อหาคุณภาพข้อสอบ และแบบสอบ สำหรับการดำเนินงานมีดังนี้

(1) ผู้วิจัยสุ่มตัวอย่างแบบยกกกลุ่ม (Cluster Random Sampling) 1 กลุ่มเรียน ได้นิสิตจำนวน 47 คน เป็นนิสิตสาขาการสอนภาษาอังกฤษ และการสอนสุขศึกษา พลศึกษา ภาคเรียนที่ 1 ปีการศึกษา 2553

(2) นำนิสิตมาทำแบบสอบ

(3) คำนวณหาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ความโด่ง ของคะแนนสอบ และหาคุณภาพข้อสอบคือ ค่าความยาก ค่าอำนาจจำแนก ด้วยทฤษฎีการทดสอบแบบดั้งเดิม และหาความเที่ยงแบบสอดคล้องภายในวิธีคูเดอร์-ริชาร์ดสัน (KR-20) ด้วยโปรแกรม Excel และ SPSS

6.3 การนำแบบสอบไปใช้จริง เป็นการนำแบบสอบที่ได้รับการปรับปรุงในการทดลองใช้ครั้งที่ 2 นำมาใช้จริง สำหรับการดำเนินงานมีดังนี้

(1) ผู้วิจัยสุ่มตัวอย่างแบบยกกกลุ่มได้นิสิตที่ลงเรียนวิชา 400204 การวัดและประเมินในชั้นเรียนจำนวน 23 กลุ่มเรียน รวมนิสิตทั้งสิ้น 667 คน

(2) นำนิสิตมาทำแบบสอบ

(3) คำนวณหาค่าคุณภาพของข้อสอบ และแบบสอบ หาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่ามัธยฐาน ฐานนิยม พิสัย ความเบ้ ความโด่ง ของคะแนนสอบของกลุ่มตัวอย่างนักเรียน สำหรับการหาคุณภาพของข้อสอบคือ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c) และการหาคุณภาพแบบสอบคือ ฟังก์ชันสารสนเทศของแบบสอบ (test information) หรือ ความเที่ยงของแบบสอบนั่นเอง โดยใช้โมเดลการตอบสนองข้อสอบ (IRT) แบบ 3 พารามิเตอร์ สามารถวิเคราะห์โดยใช้โปรแกรม Xcalibre และ Multilog และหาความตรงตามเกณฑ์ แบบความตรงตามสภาพ (Concurrent Validity) โดยการหาสัมประสิทธิ์สหสัมพันธ์ระหว่าง คะแนนสอบจากแบบสอบที่ผู้วิจัยสร้างขึ้น กับเกรดเฉลี่ยรวมของนิสิต (GPAX) วิเคราะห์ด้วยโปรแกรม SPSS

สำหรับเกณฑ์ที่ใช้คัดเลือกข้อสอบที่มีคุณภาพดีคือ ค่าความยากอยู่ระหว่าง -2.50 ถึง +2.50 ค่าอำนาจจำแนกอยู่ระหว่าง +0.50 ถึง +2.50 และค่าการเดาไม่เกิน 0.30 (ศิริชัย กาญจนวาสี, 2545)

ขั้นตอนที่ 2 การหาคะแนนจุดตัดโดยการกำหนดมาตรฐานด้วยวิธีบูคมาร์ค

วัตถุประสงค์ เพื่อหาคะแนนจุดตัดของแบบสอบที่สร้างขึ้นด้วยการกำหนดมาตรฐาน ด้วยวิธีบูคมาร์ค

การดำเนินงานการวิจัยมีขั้นตอนดังนี้

1. การกำหนดประชากรและกลุ่มตัวอย่าง

1.1 ประชากร อาจารย์มหาวิทยาลัยบูรพาที่ทำการสอนวิชาการวัดและประเมินผล การศึกษาระดับปริญญาตรีที่เป็นอาจารย์ประจำ และอาจารย์พิเศษ จำนวน 9 คน

1.2 กลุ่มตัวอย่าง อาจารย์มหาวิทยาลัยบูรพาที่ทำการสอนวิชาการวัดและ ประเมินผลการศึกษา ระดับปริญญาตรีที่เป็นอาจารย์ประจำ และอาจารย์พิเศษ จำนวน 6 คน โดยการสุ่มอย่างง่าย มีคุณสมบัติคือ สอนในวิชาการวัดและประเมินผลการศึกษา และจบ การศึกษาในระดับปริญญาตรี ปริญญาโท หรือปริญญาเอกในสาขาการวัดและประเมินผล การศึกษาหรือสาขาที่เกี่ยวข้อง

2. ผู้วิจัยสรุปขั้นตอนการกำหนดมาตรฐานบุคคมาค์จากการสังเคราะห์เอกสารงานวิจัย ที่เกี่ยวข้องดังนี้ (Buckendahl, 2002; Wang, 2003; Beretvas, 2004; Lewis, Green, Mitzel, Baum & Patz, 1998; Kiplinger, 1997; Cizek, Bunch & Koons, 2004; Ferrara, Johnson & Chen, 2005) ดังนี้

2.1 นำผลการสอบของนักเรียนมาหาค่าพารามิเตอร์ความยาก และอำนาจจำแนก โดยการวิเคราะห์ตามทฤษฎี IRTค่าพารามิเตอร์เหล่านี้จะถูกนำมาแปลงเข้าสู่ตรหาค่า θ จากนั้น นำค่า θ ที่คำนวณได้ไปจัดทำคู่มือจัดเรียงข้อสอบต่อไป

2.2 สร้างคู่มือการจัดเรียงข้อสอบ (Ordered Item Booklet: OIB) คู่มือการจัดเรียง ข้อสอบดังกล่าว มีการจัดเรียงข้อสอบตามค่าความยาก โดยจัดเรียงข้อสอบจากข้อสอบง่ายที่สุด ไปยังข้อสอบยากที่สุด และมีข้อมูลความสามารถของผู้สอบ (θ) ประกอบการพิจารณาของผู้ ตัดสิน OLB มีรูปแบบการจัดเรียงข้อสอบ 1 ข้อ ต่อ 1 หน้า

2.3 ผู้ตัดสินทั้ง 6 คน ร่วมสร้างนิยามความสามารถของนักเรียนใน 1) ระดับดีเยี่ยม 2) ระดับดีมาก 3) ระดับดี 4) ระดับดีพอใช้ 5) ระดับพอใช้ 6) ระดับอ่อน 7) ระดับอ่อนมาก และ 8) ระดับตก ว่านักเรียนแต่ละระดับควรมีความรอบรู้อย่างไร เพื่อสร้างให้เกิดความเข้าใจตรงกันในการพิจารณาตัดสินคะแนนจุดตัดในแต่ละระดับ

2.4 ผู้ตัดสินพิจารณากำหนดมาตรฐาน โดยดำเนินการพิจารณา 3 รอบ

2.5 การตัดสินรอบที่ 1 เป้าหมายหลัก คือเพื่อให้ผู้ตัดสินมีความคุ้นเคยกับคู่มือการจัดเรียงข้อสอบ ในรอบนี้จะทำงานเป็นกลุ่มย่อย ผู้ตัดสินจะทำงานอิสระจากกันนั้นคือผู้ตัดสินแต่ละคนทำการกำหนดคะแนนจุดตัดที่ระดับพอใจเป็นจุดแรก โดยให้ผู้ตัดสินแต่ละคนพิจารณา ข้อสอบทีละข้อในคู่มือจัดเรียงข้อสอบซึ่งพิจารณาตั้งแต่หน้าแรกซึ่งเป็นข้อสอบที่ง่ายที่สุด โดยตอบ

คำถาม 2 คำถามในแต่ละข้อต่อไปนี้ 1) นิสิตต้องมีความรู้ความสามารถอะไรเพื่อให้ตอบข้อสอบข้อนี้ถูก? 2) อะไรที่ทำให้ข้อสอบข้อนี้ยากกว่าข้อสอบข้อที่แล้วในคู่มือจัดเรียงข้อสอบ?

การตอบคำถามเหล่านี้เป็นการเตรียมตัวผู้ตัดสินใจในการทำการตัดสินใจคะแนนจุดตัด โดยผู้ตัดสินใจทำการกำหนดบุคมาาร์ค (คั่นหนังสือ) บนหน้าที่ผู้ตัดสินใจคิดว่า “ผู้ที่คาบเส้นระดับอ่อนมากจะมีโอกาสตอบข้อสอบข้อนี้ถูก 67%” หรือมีผู้คาบเส้นระดับพอใช้ 100 คน จะมีผู้ตอบถูก 67 คน โดยผู้ตัดสินใจฝึกให้เข้าใจว่านิสิตที่คาบเส้นในระดับพอใช้ควรมีโอกาสตอบถูกอย่างน้อย 67% ในข้อสอบหน้าถัดไป และมีโอกาสตอบถูกสูงกว่า 67% ในหน้าก่อนหน้าที่คั่นหนังสือ

หลังจากดำเนินการกำหนดคะแนนจุดตัดที่ระดับอ่อนมากแล้ว ผู้ตัดสินใจดำเนินการหาคะแนนจุดตัดระดับอ่อน โดยใช้กระบวนการคล้ายกับที่กล่าวมา ในระดับนี้ผู้ตัดสินใจกำหนดบุคมาาร์ค (คั่นหนังสือ) บนหน้าที่คิดว่า “ผู้ที่คาบเส้นระดับอ่อนมีโอกาสดตอบข้อสอบข้อนี้ถูก 67% หรือ 2/3” จากนั้นผู้ตัดสินใจดำเนินการหาคะแนนจุดตัดระดับพอใช้ ดีพอใช้ ระดับดีมาก ระดับดีเยี่ยม ตามลำดับเช่นเดียวกับข้างต้น ผลที่ได้จากการหาคะแนนจุดตัดจะได้ระดับคะแนนจุดตัด 7 ระดับ คือ 1) ระดับดีเยี่ยม 2) ระดับดีมาก 3) ระดับดี 4) ระดับดีพอใช้ 5) ระดับพอใช้ 6) ระดับอ่อน 7) ระดับอ่อนมาก

2.6 การตัดสินใจรอบที่ 2 จัดให้มีการอภิปรายในกลุ่มย่อย ก่อนการอภิปรายผู้ตัดสินใจได้รับข้อมูลเลขที่หน้าในคู่มือจัดเรียงข้อสอบที่ผู้ตัดสินใจคนอื่นกำหนดบุคมาาร์ค (คั่นหนังสือ) เมื่อผู้ตัดสินใจพิจารณาข้อมูลนั้นแล้ว ผู้ตัดสินใจร่วมกันอภิปรายถึงเหตุผล และความเหมาะสมของคะแนนจุดตัดที่ผู้ตัดสินใจแต่ละคนตัดสินใจ โดยเฉพาะให้ผู้ตัดสินใจคนที่ให้คะแนนจุดตัดสูงสุด และต่ำสุด อภิปรายเหตุผลที่ใช้คะแนนจุดตัดตรงนี้ ซึ่งการอภิปรายอาจอภิปรายในประเด็นตัวข้อสอบเกี่ยวกับระดับทักษะ และเนื้อหาทางวิชาการที่ควรจะรอบรู้สำหรับนิสิตในแต่ละคะแนนจุดตัด เหตุผลในการอภิปรายก็เพื่อให้ผู้ตัดสินใจได้รับฟังเหตุผลของผู้ตัดสินใจคนอื่น การอภิปรายในกลุ่มย่อยนั้น ผู้ตัดสินใจแสดงความคิดเห็นและประสบการณ์ที่มีความหลากหลาย หลังจากนั้นเปิดโอกาสให้ผู้ตัดสินใจพิจารณาคู่มือจัดเรียงข้อสอบอีกครั้งหนึ่งโดยนำความคิดเห็นจากผู้ตัดสินใจคนอื่นมาพิจารณาประกอบด้วย เพื่อหาคะแนนจุดตัดในรอบที่ 2 การพิจารณาคะแนนจุดตัดดำเนินการพิจารณาทีละระดับครบทั้ง 7 ระดับ เมื่อได้เลขที่หน้าของผู้ตัดสินใจแต่ละคนกำหนดขึ้นมาแล้ว จากนั้นเลขที่หน้าจะถูกแปลงเป็นค่าความสามารถ (θ) (เลขที่หน้าที่คั่นคู่มือจัดเรียงข้อสอบไว้นี้ตรงกับค่า θ เท่าไร นำค่า θ เหล่านี้มาหาค่าเฉลี่ยเพื่อนำมาประมาณความสามารถขั้นต่ำของนิสิต)

2.7 การตัดสินใจรอบที่ 3 นำผู้ตัดสินใจทั้ง 2 กลุ่มย่อยมารวมกันเป็นกลุ่มใหญ่เพียงกลุ่มเดียว ผู้อำนวยการความสะดวกนำข้อมูลผลกระทบ (Impact Data) คือ เปอร์เซนต์ของนิสิตที่อยู่ในแต่

ละระดับ ค่าเฉลี่ยคะแนนจุดตัดจากกลุ่มย่อยทั้งสองกลุ่ม มาแสดงให้ผู้ตัดสินพิจารณา จากนั้น ผู้ตัดสินอภิปรายร่วมกันโดยพิจารณาข้อสอบที่อยู่ระหว่างค่าเฉลี่ยคะแนนจุดตัดทั้งสองกลุ่ม แล้ว เปิดโอกาสให้ผู้ตัดสินทำการกำหนดคะแนนจุดตัดอีกครั้งอย่างอิสระจากกัน โดยการกำหนดหน้าที่ คั่นหนังสือในคู่มือจัดเรียงข้อสอบ เลขที่หน้าที่ของผู้ตัดสินแต่ละคนที่กำหนดขึ้นมานั้นถูกแปลงเป็น ค่าความสามารถ (θ) (เลขที่หน้าที่คั่นคู่มือจัดเรียงข้อสอบไว้นี้ตรงกับค่า θ เท่าไร แล้วนำค่า θ เหล่านี้มาหาค่าเฉลี่ยเพื่อนำมาประมาณความสามารถขั้นต่ำของนิสิต

2.8 เมื่อได้มาตรฐานที่มีคะแนนจุดตัดทั้ง 7 ระดับแล้ว จากนั้นผู้ตัดสินระดมพลังสมองเพื่อเขียนบรรยายระดับการปฏิบัติของนิสิตในประเด็นความรู้ ทักษะ และความสามารถของนิสิต ณ ระดับคะแนนจุดตัดของระดับนั้น รวมทั้งสิ้น 8 ระดับ

2.9 ผู้วิจัยคำนวณหาความตรงของคะแนนจุดตัด โดยหาสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมนระหว่าง ระดับคะแนนที่ได้จากการกำหนดคะแนนมาตรฐานด้วยวิธีบูคมาร์ค กับ ระดับคะแนนเฉลี่ยรวมของนิสิต (GPAX) วิเคราะห์ด้วยโปรแกรม SPSS

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยนี้คือ แบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผลในชั้นเรียน

แบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผลในชั้นเรียน เป็นแบบสอบที่ผู้วิจัยได้ดำเนินการสร้างและพัฒนาขึ้นมาเอง แบบสอบเป็นแบบเลือกตอบ 4 ตัวเลือก และแบบอัตนัย จำนวน 2 ฉบับ คือ ฉบับที่ 1 เป็นแบบสอบกลางภาค ฉบับที่ 2 แบบสอบปลายภาค

สำหรับขั้นตอนการสร้างและพัฒนาแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผลในชั้นเรียน มีลำดับขั้นดังนี้

1) วิเคราะห์คำอธิบายรายวิชา เพื่อสร้างตารางโครงสร้างเนื้อหา จากนั้นนำตารางโครงสร้างเนื้อหาไปให้อาจารย์ที่สอนวิชาการวัดและประเมินผลการศึกษา ที่มีประสบการณ์การสอนไม่ต่ำกว่า 3 ปี จำนวน 3 คน กำหนดน้ำหนักเนื้อหา และพิจารณาความครอบคลุมของเนื้อหา ตลอดจนความเหมาะสมของวัตถุประสงค์การเรียนรู้ ตัวอย่างแบบประเมินความครอบคลุมของวัตถุประสงค์การเรียนรู้ในขอบเขตเนื้อหา แบบประเมินการกำหนดน้ำหนักความสำคัญของแต่ละเนื้อหา ดังต่อไปนี้

ตัวอย่างแบบประเมินความครอบคลุมของวัตถุประสงค์การเรียนรู้ในขอบเขตเนื้อหา

คำชี้แจง ขอให้ท่านพิจารณาว่าวัตถุประสงค์เหล่านี้วัดได้ครอบคลุมในขอบเขตเนื้อหานั้นหรือไม่ โดยพิจารณาตามเกณฑ์ดังต่อไปนี้

- 1 หมายถึง แน่ใจว่าวัตถุประสงค์นี้วัดได้ครอบคลุมขอบเขตเนื้อหานั้นจริง
- 0 หมายถึง ไม่แนใจว่าวัตถุประสงค์นี้วัดได้ครอบคลุมขอบเขตเนื้อหานั้นจริง
- 1 หมายถึง แน่ใจว่าวัตถุประสงค์นี้วัดได้ไม่ครอบคลุมขอบเขตเนื้อหานั้นจริง

วัตถุประสงค์	คะแนนการตัดสิน			ข้อเสนอแนะ
	- 1	0	1	
คำศัพท์เกี่ยวกับการวัดและประเมินผล 1) นักเรียนสามารถบอกความหมายของการวัดและการประเมินผลได้ถูกต้อง เหตุการณ์ต่อไปนี้เป็นกรวัด (Measurement) ก. ขวัญเป็นเด็กอ้วน ข. แพนเค้กมาโรงเรียนสาย ค. พิงก์รับประทานอาหาร 20 นาที ง. ริต้าเป็นเด็กมีความรับผิดชอบสูง				

สำหรับตารางโครงสร้างเนื้อหา ผู้เชี่ยวชาญได้กำหนดน้ำหนักความสำคัญของเนื้อหาของแต่ละเนื้อหา ปรากฏดังตารางที่ 3-2

ตารางที่ 3-2 ตารางโครงสร้างกำหนดน้ำหนักความสำคัญของเนื้อหาและวัตถุประสงค์การเรียนรู้
ของวิชาการวัดผลและประเมินผลในชั้นเรียน

เนื้อหา	น้ำหนัก ความ สำคัญ	จำนวน ข้อ	เลขที่ข้อ
แบบสอบฉบับที่ 1			
1. หลักการวัดและประเมินในชั้นเรียน	18	(11)	
1.1 คำศัพท์เกี่ยวกับการวัดและประเมินผล			
1) นิสิตสามารถบอกความหมายของการวัด และการประเมินผลได้ถูกต้อง		6	2, 3, 4, 7 9, 41
1.2 ประเภทการประเมินผล			
1) นิสิตสามารถจำแนกสถานการณ์ที่กำหนด ว่าเป็นการประเมินผลประเภทใดได้ถูกต้อง		4	5, 6, 43, 44
1.3 การวัดประเมินการเรียนรู้แนวใหม่		1	1
1) นิสิตสามารถบอกลักษณะการวัดและ ประเมินผลแนวใหม่ได้ถูกต้อง			
2. แนวทางการวัดและประเมินผล การเรียนรู้	3	(4)	
1) นิสิตสามารถอธิบายแนวทางการวัด และประเมินผลการเรียนรู้ พ.ศ. 2551 ได้		4	57, 58, 59 60
3. การวางแผนประเมินในชั้นเรียน	16	(10)	
1) นิสิตสามารถเขียนวัตถุประสงค์การเรียนรู้ได้ ถูกต้อง		7	13, 14, 15, 21, 22, 27, 42
2) นิสิตสามารถสร้างตารางแผนงานประเมิน ได้ถูกต้อง		3	10, 11, 12

ตารางที่ 3-2 (ต่อ)

เนื้อหา	น้ำหนัก ความ สำคัญ	จำนวน ข้อ	เลขที่ข้อ
4. แบบสอบวัดผลสัมฤทธิ์ทางการเรียน	19	(12)	
4.1) ประเภทของแบบสอบวัดผลสัมฤทธิ์ ทางการเรียน	1	8	
1) นิสิตสามารถจำแนกประเภทของแบบสอบ วัดผลสัมฤทธิ์ทางการเรียน			
4.2) ขั้นตอนการสร้างแบบสอบวัดผลสัมฤทธิ์ ทางการเรียนได้ถูกต้อง	1	20	
1) นิสิตสามารถบอกขั้นตอนการสร้างแบบสอบ วัดผลสัมฤทธิ์ทางการเรียนได้ถูกต้อง			
4.3) การสร้างแบบสอบวัดผลสัมฤทธิ์ทางการ เรียนแบบต่างๆ			
1) นิสิตสามารถสร้างข้อสอบวัดระดับความรู้ 6 ระดับได้ถูกต้อง		10	16, 17, 18, 23, 24, 25, 45, 46, 48, 47
5. การประเมินทางจิตพิสัย	18	(11)	19, 26, 28, 29,
1) เมื่อครูกำหนดสถานการณ์มาให้ นิสิต สามารถ ระบุเครื่องมือการประเมินที่เหมาะสมได้		9	30, 38, 49, 50, 55
2) เมื่อกำหนดข้อคำถามมา นิสิตสามารถระบุ รูปแบบแบบวัดทางจิตวิทยาได้		3	33, 34, 35
6. การประเมินตามสภาพจริง และการ ประเมินการปฏิบัติ	18	(11)	
6.1 ความหมายและลักษณะการประเมินตาม สภาพจริง			
1) นิสิตสามารถอธิบายความหมายและ ลักษณะ การประเมินตามสภาพจริงได้ถูกต้อง		2	31, 52

ตารางที่ 3-2 (ต่อ)

เนื้อหา	น้ำหนัก ความ สำคัญ	จำนวน ข้อ	เลขที่ข้อ
6.2 ความหมายการประเมินการปฏิบัติ และ วิธีประเมินการปฏิบัติ			
1) นิสิตสามารถอธิบายความหมายและวิธีประเมิน การปฏิบัติได้ถูกต้อง		2	32, 53
6.3 รูปแบบงานในการประเมินการปฏิบัติ			
1) เมื่อครูกำหนดสถานการณ์มาให้ นิสิตสามารถชี้ ได้ว่าเป็นการประเมินการปฏิบัติรูปแบบใด		5	36, 37, 39, 40, 56
6.4 มโนทัศน์คะแนนรูบริก			
1) นิสิตสามารถอธิบายแนวคิดคะแนนรูบริกได้ ถูกต้อง		1	51, 54
2) นิสิตสามารถสร้างสร้างเกณฑ์ให้คะแนนให้ได้ ถูกต้อง		1	1 (อัตร้อย)
รวม	100	61	

ตารางที่ 3-3 ตารางโครงสร้างกำหนดน้ำหนักความสำคัญของเนื้อหาและวัตถุประสงค์การเรียนรู้
ของวิชาการวัดผลและประเมินผลในชั้นเรียน

เนื้อหา	น้ำหนัก ความ สำคัญ	จำนวน ข้อ	เลขที่ข้อ
แบบสอบฉบับที่ 2			
1) สถิติเบื้องต้นเกี่ยวกับการวัดและ ประเมินผล	21	(11)	
1.1) การวัดแนวโน้มเข้าสู่ส่วนกลาง การวัด การกระจาย			

ตารางที่ 3-3 (ต่อ)

เนื้อหา	น้ำหนัก ความ สำคัญ	จำนวน ข้อ	เลขที่ข้อ
1) เมื่อครูกำหนดข้อมูลมาให้ นิสิตสามารถ คำนวณหาแนวโน้มเข้าสู่ส่วนกลาง และการวัด การกระจายได้		3	1, 2, 7
1.2) ตำแหน่งเปอร์เซ็นต์ไทล์			
1) เมื่อครูกำหนดข้อมูลมาให้ นิสิตสามารถ คำนวณหาตำแหน่งเปอร์เซ็นต์ไทล์และแปล ความหมายได้ถูกต้อง		2	3, 4
1.3) คะแนนมาตรฐาน			
1) เมื่อกำหนดข้อมูลมา นิสิตสามารถคำนวณหา คะแนนมาตรฐาน Z และ คะแนนมาตรฐาน T ได้ ถูกต้อง		3	5, 6, 43
1.4) ระดับการวัด			
1) นิสิตสามารถบอกได้ว่าตัวแปรที่กำหนดให้ เป็น ข้อมูลระดับใดได้ถูกต้อง	36	1	32
1.5) โค้งการแจกแจงข้อมูล			
1) นิสิตสามารถแปลความหมายของโค้งการแจก แจงข้อมูลได้ถูกต้อง		1	31
1.6) สหสัมพันธ์			
1) เมื่อกำหนดข้อมูลมาให้ นิสิตสามารถคำนวณ สหสัมพันธ์แบบเพียร์สันได้ถูกต้อง		1	37
2) การวิเคราะห์คุณภาพข้อสอบรายข้อ			
2.1) การวิเคราะห์ข้อสอบแบบอิงกลุ่ม	37	(19)	
1) เมื่อกำหนดข้อมูลมาให้ นิสิตสามารถ คำนวณหาค่าความยาก อำนาจจำแนก และแปล ความหมายได้ถูกต้อง		11	11, 12, 13, 14, 15, 16 45, 46, 47 10, 49

ตารางที่ 3-3 (ต่อ)

เนื้อหา	น้ำหนัก ความ สำคัญ	จำนวน ข้อ	เลขที่ข้อ
2.2) การวิเคราะห์ข้อสอบแบบอิงเกณฑ์			
1) เมื่อกำหนดข้อมูลมาให้ นิสิตสามารถ คำนวณหาค่าอำนาจจำแนกปี (B-Index) และ ดัชนีความไว และแปลความหมายได้ถูกต้อง		5	9, 41, 42, 48, 50
2.3) การวิเคราะห์ข้อสอบแบบอัตนัย			
1) เมื่อกำหนดข้อมูลคะแนนสอบของข้อสอบ อัตนัยมาให้ นิสิตสามารถคำนวณค่าความยาก และอำนาจจำแนกได้ถูกต้อง		3	33, 34, 35
3) การวิเคราะห์คุณภาพแบบสอบ		(19)	
3.1) ความตรง			
1) นิสิตสามารถชี้ได้ว่าดัชนีคุณภาพของแบบสอบ และข้อสอบคือดัชนีอะไร		1	8
2) นิสิตสามารถคำนวณหาค่าดัชนีความ สอดคล้องระหว่างข้อสอบและวัตถุประสงค์ได้ ถูกต้อง		1	21
3) นิสิตสามารถระบุได้ว่าแบบสอบที่กำหนดให้ ควรใช้ความตรงประเภทใดได้ถูกต้อง		4	22, 26, 36, 40
3.2) ความเที่ยง			
1) เมื่อกำหนดลักษณะแบบสอบมาให้ นิสิต สามารถระบุได้ว่าควรใช้ความเที่ยงแบบใด		6	17, 20, 23 24, 25, 28
2) นิสิตสามารถหาความคลาดเคลื่อนมาตรฐาน ได้ถูกต้อง		1	18

ตารางที่ 3-3 (ต่อ)

เนื้อหา	น้ำหนัก ความ สำคัญ	จำนวน ข้อ	เลขที่ข้อ
3) นิสิตสามารถหาค่าความเที่ยงของแบบสอบได้ ถูกต้อง		4	19, 29, 30, 38
4) นิสิตสามารถระบุข้อดีของความเที่ยงแต่ละ แบบได้ถูกต้อง		1	39
4) การรายงานผลการเรียน	6	(3)	
1) นิสิตสามารถกำหนดเกรดแบบอิงกลุ่มและ อิงเกณฑ์ได้ถูกต้อง		1	27, 1 (อัตรัย)
2) นิสิตสามารถคำนวณหาคะแนนพัฒนาการได้ ถูกต้อง		1	44
รวม	100	51	

2) สร้างข้อสอบแบบเลือกคำตอบ 4 ตัวเลือกและข้อสอบอัตรัย จำนวน 2 ฉบับ คือ ฉบับที่ 1 แบบสอบกลางภาคเรียน ฉบับที่ 2 แบบสอบปลายภาคเรียน

3) ตรวจสอบความตรงเชิงเนื้อหาของแบบสอบ โดยผู้วิจยนำแบบสอบ ทั้ง 2 ฉบับมาให้ผู้เชี่ยวชาญจำนวน 3 คน ที่มีประสบการณ์ในการสอนมาไม่ต่ำกว่า 3 ปี พิจารณาตัดสินความสอดคล้องระหว่างข้อสอบและจุดประสงค์การเรียนรู้ โดยพิจารณาตามเกณฑ์ดังนี้

- +1 หมายถึง แน่ใจว่าข้อสอบข้อนี้สอดคล้องกับวัตถุประสงค์การเรียนรู้
- 0 หมายถึง ไม่แน่ใจว่าข้อสอบข้อนี้สอดคล้องกับวัตถุประสงค์การเรียนรู้
- 1 หมายถึง แน่ใจว่าข้อสอบข้อนี้ไม่สอดคล้องกับวัตถุประสงค์การเรียนรู้

เมื่อผู้เชี่ยวชาญพิจารณาตัดสินความสอดคล้องระหว่างข้อสอบและวัตถุประสงค์การเรียนรู้เรียบร้อยแล้ว นำผลประเมินมาหาค่าดัชนีความสอดคล้องตามวิธีของโรบินลลีและแฮมเบิลตัน (Rovinelli & Hambleton, 1977 cited in Hambleton, 1978) แบบประเมินความสอดคล้องระหว่างข้อสอบกับจุดประสงค์การเรียนรู้ มีตัวอย่างดังต่อไปนี้

แบบประเมินความสอดคล้องระหว่างข้อสอบกับจุดประสงค์การเรียนรู้

คำชี้แจง ขอให้ท่านพิจารณาว่าวัตถุประสงค์เหล่านี้วัดได้ครอบคลุมในขอบเขตเนื้อหาหรือไม่ โดยพิจารณาตามเกณฑ์ดังต่อไปนี้

- 1 หมายถึง แน่ใจว่าข้อสอบเหล่านี้วัดได้สอดคล้องกับจุดประสงค์นั้นจริง
- 0 หมายถึง ไม่แน่ใจว่าข้อสอบเหล่านี้วัดได้สอดคล้องกับจุดประสงค์นั้นจริง
- 1 หมายถึง แน่ใจว่าข้อสอบเหล่านี้วัดได้ไม่สอดคล้องกับจุดประสงค์นั้นจริง

วัตถุประสงค์	คะแนนการตัดสิน			ข้อเสนอแนะ
	- 1	0	1	
1.1 คำศัพท์เกี่ยวกับการวัดและประเมินผล 1) นิสิตสามารถบอกความหมายของการวัดและการประเมินผลได้ถูกต้อง				

หลังจากนั้นผู้วิจัยนำคะแนนความคิดเห็นของผู้เชี่ยวชาญมาคำนวณหาค่าดัชนี IOC รายละเอียดดังตารางที่ 3-4 และ 3-5

ตารางที่ 3-4 ผลการตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์การเรียนรู้ของแบบสอบกลางภาคเรียน

ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC
1	1.0	19	1.0	37	1.0	55	1.0	73	1.0
2	0.6	20	1.0	38	1.0	56	0.6	74	1.0
3	0.6	21	1.0	39	1.0	57	1.0	75	1.0
4	1.0	22	1.0	40	1.0	58	0.6	76	1.0

ตารางที่ 3-4 (ต่อ)

ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC
5	1.0	23	1.0	41	0.6	59	1.0	77	1.0
6	1.0	24	1.0	42	1.0	60	1.0	78	1.0
7	1.0	25	1.0	43	1.0	61	1.0	79	1.0
8	0.6	26	1.0	44	1.0	62	1.0	80	1.0
9	1.0	27	1.0	45	1.0	63	1.0	81	1.0
10	0.6	28	1.0	46	1.0	64	0.6	82	1.0
11	0.6	29	1.0	47	1.0	65	1.0	83	1.0
12	1.0	30	0.6	48	1.0	66	1.0	84	1.0
13	1.0	31	1.0	49	1.0	67	1.0	85	1.0
14	1.0	32	1.0	50	1.0	68	1.0	86	1.0
15	1.0	33	1.0	51	0.6	69	1.0	87	1.0
16	0.6	34	1.0	52	1.0	70	1.0	88	1.0
17	0.6	35	1.0	53	1.0	71	1.0	89	1.0
18	1.0	36	1.0	54	1.0	72	1.0	90	1.0

* ข้อสอบที่มีค่า IOC ต่ำกว่า 0.5

จากตารางที่ 3-4 ค่าดัชนีความสอดคล้อง (IOC) อยู่ในช่วง 0.6 – 1.0 ข้อสอบทั้งหมด 90 ข้อ มีค่าดัชนีความสอดคล้องสูงกว่า 0.5 แสดงว่าผู้เชี่ยวชาญมีความคิดเห็นและพิจารณาตัดสินแล้วว่าข้อสอบทั้งหมดนั้นวัดได้ตรงตามวัตถุประสงค์

ตารางที่ 3-5 ผลการตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์การเรียนรู้ของแบบ
สอบกลางภาคเรียน

ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC
1	1.0	15	1.0	29	1.0	43	1.0	57	1.0
2	1.0	16	1.0	30	1.0	44	0.6	58	1.0
3	0.6	17	1.0	31	1.0	45	1.0	59	1.0

ตารางที่ 3-5 (ต่อ)

ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC	ข้อที่	IOC
4	1.0	18	0.6	32	1.0	46	1.0	60	1.0
5	1.0	19	1.0	33	0.6	47	1.0	61	1.0
6	1.0	20	1.0	34	1.0	48	1.0	62	1.0
7	1.0	21	1.0	35	1.0	49	1.0	63	1.0
8	0.6	22	1.0	36	1.0	50	1.0	64	1.0
9	1.0	23	1.0	37	1.0	51	1.0	65	1.0
10	1.0	24	1.0	38	1.0	52	0.6	66	1.0
11	1.0	25	1.0	39	1.0	53	0.6	67	1.0
12	1.0	26	0.6	40	1.0	54	1.0	68	1.0
13	1.0	27	1.0	41	1.0	55	1.0	69	1.0
14	1.0	28	1.0	42	1.0	56	1.0	70	1.0

* ข้อสอบที่มีค่า IOC ต่ำกว่า 0.5

จากตารางที่ 3-5 ค่าดัชนีความสอดคล้อง (IOC) อยู่ในช่วง 0.6 – 1.0 ข้อสอบทั้งหมด 70 ข้อ มีค่าดัชนีความสอดคล้องสูงกว่า 0.5 แสดงว่าผู้เชี่ยวชาญมีความคิดเห็นและพิจารณาตัดสินแล้วว่าข้อสอบทั้งหมดนั้นวัดได้ตรงตามวัตถุประสงค์

4) การนำแบบสอบไปทดลองใช้เพื่อตรวจสอบคุณภาพของแบบสอบ ผู้วิจัยดำเนินการทดลองใช้แบบสอบ 3 ครั้ง ดังนี้

4.1 การทดลองใช้แบบสอบครั้งที่ 1 เพื่อตรวจสอบด้านภาษาเกี่ยวกับคำชี้แจงในการสอบ และข้อคำถาม ว่านักเรียนมีความเข้าใจตรงกันหรือไม่ มีภาษาตรงจุดไหนมีความกำกวมเพื่อนำไปปรับปรุงแก้ไข ตลอดจนเพื่อนำข้อมูลที่ได้มากำหนดเวลาในการสอบที่มีความเหมาะสมสำหรับการดำเนินงานมีดังนี้

- (1) ผู้วิจัยสุ่มตัวอย่างอย่างง่ายนิสิต จำนวน 5 คน
- (2) เมื่อนักเรียนทำแบบสอบเสร็จ ผู้วิจัยสัมภาษณ์นิสิตแต่ละคนเกี่ยวกับภาษาที่ใช้ในแบบทดสอบว่ามีคำสั่ง และข้อคำถามชัดเจน เข้าใจตรงกันหรือไม่
- (3) ปรับปรุงภาษาที่ใช้ในแบบทดสอบ

ผลการทดลองใช้แบบสอบครั้งที่ 1 ปรากฏว่า นิสิตมีความเข้าใจในคำชี้แจง และ
ข้อคำถามแต่ละข้อดี

4.2 การทดลองใช้แบบสอบครั้งที่ 2 นำผลการสอบไปคำนวณเพื่อหาคุณภาพ
ข้อสอบ และแบบสอบ นั่นคือ ผู้วิจัยคำนวณหาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ความ
โด่ง ของคะแนนสอบ และหาคุณภาพข้อสอบคือ ค่าความยาก ค่าอำนาจจำแนก และหาคุณภาพ
แบบสอบคือ ความเที่ยงแบบสอดคล้องภายในวิธีคูเดอร์-ริชาร์ดสัน (KR-20) วิเคราะห์โดยใช้
โปรแกรม SPSS และ EXCEL

ผลการทดลองใช้ครั้งที่ 2 ปรากฏค่าสถิติ ดังต่อไปนี้
(1) ค่าสถิติพื้นฐาน มีดังต่อไปนี้

ตารางที่ 3-6 ค่าสถิติพื้นฐานของแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผล
การศึกษาฉบับที่ 1 แบบสอบกลางภาคเรียน ฉบับที่ 2 แบบสอบปลายภาคเรียน



ค่าสถิติพื้นฐาน	ฉบับที่ 1		ฉบับที่ 2	
	แบบสอบกลางภาคเรียน		แบบสอบปลายภาคเรียน	
	แบบ เลือกตอบ	แบบอัตนัย	แบบ เลือกตอบ	แบบอัตนัย
คะแนนเต็ม	90	5	70	5
คะแนนเฉลี่ย	46.98	2.74	29.91	1.81
ส่วนเบี่ยงเบน มาตรฐาน	9.20	1.421	7.40	1.624
ความเบ้	-.20	.095	.69	.196
ความโด่ง	-.45	-1.023	-.25	-1.650
คะแนนสูงสุด	64	0	18	0
คะแนนต่ำสุด	26	5	49	4

จากตารางที่ 3-6 พบว่า แบบสอบฉบับที่ 1 นิสิตส่วนใหญ่ทำคะแนนได้เกินครึ่งของ
คะแนนเต็ม แต่ฉบับที่ 2 นิสิตส่วนใหญ่ได้คะแนนต่ำกว่าครึ่งของคะแนนเต็ม และเมื่อพิจารณาค่า
ความเบ้ พบว่า ฉบับที่ 1 ติดค่าลบ หรือโค้งเบ้ซ้าย แสดงว่าแบบสอบค่อนข้างง่าย ฉบับที่ 2 มีค่า
มากกว่า 0 หรือโค้งเบ้ขวา แสดงว่า แบบสอบค่อนข้างยาก

(2) ค่าความยาก ค่าอำนาจจำแนก และค่าความเที่ยง

เมื่อพิจารณาค่าความยาก อำนาจจำแนก และค่าความเที่ยง ของแบบสอบทั้ง 4 ฉบับมีรายละเอียดดังนี้

ตารางที่ 3-7 ค่าความยาก และอำนาจจำแนกของแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการ
วัดและประเมินผลการศึกษาฉบับที่ 1 แบบสอบกลางภาคเรียน ฉบับที่ 2 แบบสอบ
ปลายภาคเรียน

ลักษณะข้อสอบ	จำนวน ข้อ	ค่าอำนาจจำแนก		ค่าความยาก		ความ เที่ยง
		$\bar{X}$	S.D.	$\bar{X}$	S.D.	
ฉบับที่ 1						
ตอนที่ 1 แบบเลือกตอบ	90	0.175	0.213	0.513	0.243	0.795
ตอนที่ 2 แบบอัตนัย	1	0.470		0.548		-
ฉบับที่ 2						
ตอนที่ 1 แบบเลือกตอบ	70	0.175	0.149	0.412	0.206	0.749
ตอนที่ 2 แบบอัตนัย	1	0.591		0.366		-

จากตารางที่ 3-7 แสดงว่า ข้อสอบแบบเลือกตอบทั้ง 2 ฉบับ มีค่าความยากง่ายเฉลี่ยปานกลาง แต่ค่าอำนาจจำแนกเฉลี่ยทั้งสองฉบับต่ำ ส่วนข้อสอบอัตนัยฉบับที่ 1 มีความยากง่ายปานกลาง ฉบับที่ 2 ค่อนข้างยาก ส่วนอำนาจจำแนกของข้อสอบอัตนัยทั้งสองฉบับดีมาก สามารถจำแนกเด็กเก่ง และเด็กอ่อนออกจากกันได้ สำหรับค่าความเที่ยงของแบบสอบทั้งสองฉบับอยู่ในระดับปานกลาง คือ 0.795 และ 0,749 ตามลำดับ

สำหรับเกณฑ์ที่ใช้คัดเลือกข้อสอบที่มีคุณภาพดีคือ ค่าความยากอยู่ระหว่าง 0.20 – 0.80 และค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป (ศิริชัย กาญจนวาสี, 2544) ส่วนข้อสอบที่ไม่เป็นไปตามเกณฑ์ดังกล่าว ผู้วิจัยได้ดำเนินการปรับปรุงข้อสอบ เช่น ข้อสอบที่ยากมาก ผู้วิจัยจะปรับปรุงข้อสอบข้อนั้นให้ง่ายขึ้นโดยการลดความซับซ้อนของโจทย์ เป็นต้น และข้อสอบบางส่วนผู้วิจัยดำเนินการตัดทิ้ง ในข้อสอบที่ไม่เป็นไปตามเกณฑ์ ดังตารางที่ 3-8

ตารางที่ 3-8 ข้อสอบที่ไม่มีคุณภาพตามเกณฑ์ที่กำหนด ต้องได้รับการปรับปรุงข้อสอบ

ลักษณะข้อสอบ	เลขที่ข้อแต่ละฉบับ	ข้อสอบที่ต้องปรับปรุง/ตัดทิ้ง
ฉบับที่ 1		
ตอนที่ 1 แบบเลือกตอบ	1-90	ข้อที่ 1, 4, 6, 8, 9, 10, 15, 16, 17, 21, 24, 25, 29, 30, 32, 43, 44, 46, 48, 50, 51, 54, 56, 62, 64, 65, 67, 69, 70, 72, 73, 74, 78, 79, 83, 84, 85, 87, 88, 89, 90
ตอนที่ 2 แบบอัตนัย	1	ไม่มีปรับปรุง
ฉบับที่ 2		
ตอนที่ 1 แบบเลือกตอบ	1-70	ข้อที่ 1, 2, 5, 6, 8, 17, 19, 23, 35, 43, 44, 49, 65, 67, 68, 69, 70
ตอนที่ 2 แบบอัตนัย	1	ไม่มีปรับปรุง

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลแบ่งออกเป็น 2 ระยะคือ ระยะที่ 1 ดำเนินการสอบกับนิสิต และระยะที่ 2 ดำเนินการเก็บข้อมูลกับกลุ่มผู้ตัดสิน มีขั้นตอนการดำเนินงานเก็บรวบรวมข้อมูลแต่ละระยะดังนี้

การเก็บข้อมูลระยะที่ 1 (ดำเนินการเก็บข้อมูลกับนิสิต)

1. ผู้วิจัยติดต่อประสานงานการสอบกลางภาคกับอาจารย์ประจำวิชาในแต่ละกลุ่มเพื่อนัดวัน เวลา สถานที่ในการสอบเพื่อเก็บข้อมูล
2. ผู้วิจัยดำเนินการสอบนิสิตคณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา ที่ลงทะเบียนเรียนวิชาการวัดและประเมินผลการศึกษาในภาคเรียนที่ 1 ปีการศึกษา 2552 เพื่อพัฒนาแบบสอบวิชาการวัดและประเมินผลในชั้นเรียน โดยเป็นการทดลองใช้ครั้งที่ 1 จำนวน 5 คน เพื่อตรวจสอบความเป็นปรนัยของข้อสอบ และทดลองใช้ครั้งที่ 2 กับนิสิตจำนวน 47 คน และดำเนินการเก็บข้อมูลจริงกับนิสิตคณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา ที่ลงทะเบียนเรียนวิชาการวัดและประเมินผลการศึกษาในภาคเรียนที่ 2 ปีการศึกษา 2552 จำนวน 667 คน วิธีการเก็บข้อมูลคือ ดำเนินการเก็บข้อมูลโดยการจัดสอบกลางภาคเรียน และปลายภาคเรียนพร้อมกันทุกกลุ่มเรียน

3. นำผลการสอบจากการทดลองใช้แบบสอบถามวิเคราะห์ข้อสอบเพื่อปรับปรุงข้อสอบ และนำผลการสอบจากการเก็บข้อมูลจริงมาหาค่าพารามิเตอร์ของข้อสอบ

การเก็บข้อมูลระยะที่ 2 (ดำเนินการเก็บข้อมูลกับผู้ตัดสิน)

1. ผู้วิจัยจัดทำคู่มือการกำหนดมาตรฐานด้วยวิธีบูคมาร์ค เพื่อให้ผู้ตัดสินที่ประกอบด้วยอาจารย์ประจำ และอาจารย์พิเศษที่สอนในรายวิชาการวัดและประเมินในชั้นเรียน มหาวิทยาลัยบูรพา ศึกษา

2. ผู้วิจัยส่งหนังสือขอเชิญอาจารย์ผู้สอน เพื่อเป็นผู้ตัดสินในการกำหนดคะแนนจุดตัด ในแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินในชั้นเรียน พร้อมคู่มือการกำหนด มาตรฐานด้วยวิธีบูคมาร์คเพื่อให้ศึกษา และนัด วัน เวลา สถานที่ เพื่อเข้าร่วมประชุมกลุ่ม กำหนด มาตรฐานร่วมกัน

3. ผู้วิจัยจัดเตรียม แบบฟอร์มสำหรับกำหนดคะแนนจุดตัด คู่มือจัดเรียงข้อสอบ เพื่อใช้ เป็นข้อมูลสารสนเทศสำหรับผู้ตัดสินใช้ในการกำหนดคะแนนจุดตัด

4. ดำเนินการกำหนดมาตรฐานด้วยวิธีบูคมาร์ค ตามวัน เวลา และสถานที่ที่กำหนด

5. ในการประชุมกลุ่มครั้งนี้มีอาจารย์เข้าร่วมประชุมทั้งหมด 6 คน เริ่มด้วยผู้วิจัยบรรยาย วิธีการกำหนดมาตรฐานด้วยวิธีบูคมาร์ค พร้อมลงฝึกปฏิบัติ ชว่งบายดำเนินการเก็บข้อมูลจริงโดยให้ผู้ ตัดสินกำหนดมาตรฐานด้วยวิธีบูคมาร์ค จากนั้น ระดมสมองเขียนบรรยายลักษณะนิสิตในแต่ละระดับ

สำหรับวิธีการกำหนดมาตรฐานด้วยวิธีบูคมาร์ค (Bookmark) มีดังนี้

ผู้วิจัยสรุปขั้นตอนการกำหนดมาตรฐานบูคมาร์คจากการสังเคราะห์เอกสารงานวิจัยที่ เกี่ยวข้องดังนี้ (Buckendahl, 2002; Wang, 2003; Beretvas, 2004; Lewis, Green, Mitzel, Baum & Patz, 1998; Kiplinger, 1997; Lin, 2005; Dawber & Lewis, 2005; Wisconsin Department of Public Instruction, 2004; Cizek, Bunch & Koons, 2004; Ferrara, Johnson & Chen, 2005)

1. สร้างคู่มือการจัดเรียงข้อสอบ (Ordered Item Booklet: OIB) คู่มือการจัดเรียง ข้อสอบดังกล่าว มีการจัดเรียงข้อสอบตามค่าความยาก โดยจัดเรียงข้อสอบจากข้อสอบง่ายที่สุด ไปยังข้อสอบยากที่สุด มีรูปแบบการจัดเรียงข้อสอบ 1 ข้อ ต่อ 1 หน้า

2. ผู้ตัดสินทั้ง 6 คน ร่วมสร้างนิยามความสามารถของนักเรียนใน 1) ระดับดีเยี่ยม 2) ระดับดีมาก 3) ระดับดี 4) ระดับดีพอใช้ 5) ระดับพอใช้ 6) ระดับอ่อน 7) ระดับอ่อนมาก และ 8) ระดับตก ว่านักเรียนแต่ละระดับควรมีความรอบรู้อย่างไร เพื่อสร้างให้เกิดความเข้าใจตรงกันใน การพิจารณาตัดสินคะแนนจุดตัดในแต่ละระดับ

3. ผู้ตัดสินพิจารณากำหนดมาตรฐาน โดยดำเนินการพิจารณา 3 รอบ ดังนี้

รอบที่ 1 เป้าหมายหลัก คือเพื่อให้ผู้ตัดสินมีความคุ้นเคยกับคู่มือการจัดเรียงข้อสอบ ในรอบนี้จะทำงานเป็นกลุ่มย่อย ผู้ตัดสินจะทำงานอิสระจากกันนั่นคือผู้ตัดสินแต่ละคนทำการ กำหนดคะแนนจุดตัดที่ระดับพอใจเป็นจุดแรก โดยให้ผู้ตัดสินแต่ละคนพิจารณาข้อสอบที่ละข้อใน คู่มือจัดเรียงข้อสอบซึ่งพิจารณาตั้งแต่หน้าแรกซึ่งเป็นข้อสอบที่ง่ายที่สุด โดยตอบคำถาม 2 คำถาม ในแต่ละข้อต่อไปนี้

- (1) นิสิตต้องมีความรู้ความสามารถอะไร เพื่อให้ตอบข้อสอบข้อนี้ถูก ?
- (2) อะไรที่ทำให้ข้อสอบข้อนี้ยากกว่าข้อสอบข้อที่แล้วในคู่มือจัดเรียงข้อสอบ ?

การตอบคำถามเหล่านี้เป็นการเตรียมตัวผู้ตัดสินในการทำการตัดสินหาคะแนนจุดตัด โดยผู้ตัดสินทำการกำหนดนิคมาร์ค (คั่นหนังสือ) บนหน้าที่ผู้ตัดสินคิดว่า “ผู้ที่คาบเส้นระดับอ่อนมากจะมีโอกาสตอบข้อสอบข้อนี้ถูก 67%” หรือมีผู้คาบเส้นระดับพอใช้ 100 คน จะมีผู้ตอบถูก 67 คน โดยผู้ตัดสินถูกฝึกให้เข้าใจว่านิสิตที่คาบเส้นในระดับพอใช้ควรมีโอกาสตอบถูกอย่างน้อย 67% ในข้อสอบหน้าถัดไป และมีโอกาสตอบถูกสูงกว่า 67% ในหน้าก่อนหน้าที่คั่นหนังสือ

หลังจากดำเนินการกำหนดคะแนนจุดตัดที่ระดับอ่อนมากแล้ว ผู้ตัดสินดำเนินการหา คะแนนจุดตัดระดับอ่อน โดยใช้กระบวนการคล้ายกับที่กล่าวมา ในระดับนี้ผู้ตัดสินกำหนด นิคมาร์ค (คั่นหนังสือ) บนหน้าที่คิดว่า “ผู้ที่คาบเส้นระดับอ่อนมีโอกาสดตอบข้อสอบข้อนี้ถูก 67%” จากนั้นผู้ตัดสินดำเนินการหาคะแนนจุดตัดระดับพอใช้ ดีพอใช้ ระดับดีมาก ระดับดีเยี่ยม ตามลำดับเช่นเดียวกับข้างต้น ผลที่ได้จากการหาคะแนนจุดตัดจะได้ระดับคะแนนจุดตัด 7 ระดับ สำหรับวิชาการวัดและประเมินในชั้นเรียน คือ 1) ระดับดีเยี่ยม 2) ระดับดีมาก 3) ระดับดี 4) ระดับ ดีพอใช้ 5) ระดับพอใช้ 6) ระดับอ่อน 7) ระดับอ่อนมาก

รอบที่ 2 จัดให้มีการอภิปรายในกลุ่ม ก่อนการอภิปรายผู้ตัดสินได้รับข้อมูลเลขที่หน้า ในคู่มือจัดเรียงข้อสอบที่ผู้ตัดสินคนอื่นกำหนดนิคมาร์ค (คั่นหนังสือ) เมื่อผู้ตัดสินพิจารณาข้อมูล นั้นแล้ว ผู้ตัดสินร่วมกันอภิปรายถึงเหตุผล และความเหมาะสมของคะแนนจุดตัดที่ผู้ตัดสินแต่ละ คนตัดสินไว้ โดยเฉพาะให้ผู้ตัดสินคนที่ให้คะแนนจุดตัดสูงสุด และต่ำสุดอภิปรายเหตุผลที่ใช้ คะแนนจุดตัดตรงนี้ ซึ่งการอภิปรายอาจอภิปรายในประเด็นตัวข้อสอบเกี่ยวกับระดับทักษะ และ เนื้อหาทางวิชาการที่ควรจะต้องรู้สำหรับนักเรียนในแต่ละคะแนนจุดตัด เหตุผลในการอภิปรายก็ เพื่อให้ผู้ตัดสินได้รับฟังเหตุผลของผู้ตัดสินคนอื่น การอภิปรายในกลุ่มย่อยนั้น ผู้ตัดสินแสดงความ คิดเห็นและประสบการณ์ที่มีความหลากหลาย หลังจากนั้นเปิดโอกาสให้ผู้ตัดสินพิจารณาคู่มือ จัดเรียงข้อสอบอีกครั้งหนึ่งโดยนำความคิดเห็นจากผู้ตัดสินคนอื่นมาพิจารณาประกอบด้วย เพื่อหา คะแนนจุดตัดในรอบที่ 2 การพิจารณาคะแนนจุดตัดดำเนินการพิจารณาทีละระดับครบทั้ง 7

ระดับ เมื่อได้เลขที่หน้าที่ของผู้ตัดสินแต่ละคนกำหนดขึ้นมาแล้ว นำเลขที่หน้ามาหาค่าเฉลี่ยเพื่อหาคะแนนจุดตัด

รอบที่ 3 ผู้อำนวยการความสะดวกนำข้อมูลเลขที่หน้าในคู่มือจัดเรียงข้อสอบที่ผู้ตัดสินแต่ละคนกำหนดบัพมาร์ค (คั่นหนังสือ)ในรอบที่ 2 มาแสดงให้ผู้ตัดสินพิจารณา จากนั้นผู้ตัดสินอภิปรายร่วมกันโดยพิจารณาข้อสอบที่อยู่ระหว่างค่าเฉลี่ยคะแนนจุดตัดทั้งสองกลุ่ม แล้วเปิดโอกาสให้ผู้ตัดสินทำการกำหนดคะแนนจุดตัดอีกครั้งอย่างอิสระจากกัน โดยการกำหนดหน้าที่คั่นหนังสือในคู่มือจัดเรียงข้อสอบ เลขที่หน้าที่ของผู้ตัดสินแต่ละคนที่กำหนดขึ้นมานั้นถูกแปลงเป็นค่าความสามารถ (θ) (เลขที่หน้าที่คั่นคู่มือจัดเรียงข้อสอบไว้นี้ตรงกับค่า θ เท่าไร แล้วนำค่า θ เหล่านี้มาหาค่าเฉลี่ยเพื่อนำมาประมาณความสามารถขั้นต่ำของนักเรียน)

4. เมื่อได้มาตรฐานที่มีคะแนนจุดตัดทั้ง 7 ระดับแล้ว จากนั้นผู้ตัดสินระดมพลังสมองเพื่อเขียนบรรยายระดับการปฏิบัติของนักเรียนในประเด็นความรู้ ทักษะ และความสามารถของนิสิต ณ ระดับคะแนนจุดตัดของระดับนั้น รวมทั้งสิ้น 8 ระดับ

การวิเคราะห์ข้อมูล

1. วิเคราะห์ข้อสอบโดยทฤษฎีการทดสอบแบบดั้งเดิมในการทดลองใช้ครั้งที่ 2 แบบเลือกตอบ

ความยากของข้อสอบ (Item Difficulty Index: p)

$$p = \frac{\text{จำนวนผู้สอบที่ตอบข้อสอบข้อนั้นถูก}}{\text{จำนวนผู้สอบทั้งหมด}} = \frac{R}{N}$$

อำนาจจำแนกของข้อสอบ (Item Discriminating Power Index: D หรือ r)

$$r = \frac{R_H - R_L}{N_H \text{ or } N_L} \quad \text{หรือ} \quad r = P_H - P_L$$

เมื่อ R_H = จำนวนผู้สอบในกลุ่มสูงที่ตอบข้อสอบข้อนั้นถูก

R_L = จำนวนผู้สอบในกลุ่มต่ำที่ตอบข้อสอบข้อนั้นถูก

N_H = จำนวนผู้สอบทั้งหมดในกลุ่มสูง

N_L = จำนวนผู้สอบทั้งหมดในกลุ่มต่ำ

แบบอัดนัย

$$\text{ความยาก} = \frac{P_H + P_L}{2}$$

$$\text{อำนาจจำแนก} = p_H - p_L$$

p_H แทน สัดส่วนของคะแนนที่ผู้สอบในกลุ่มสูงตอบถูก

p_L แทน สัดส่วนของคะแนนที่ผู้สอบในกลุ่มต่ำตอบถูก

2. การวิเคราะห์ค่าความสามารถ

เมื่อได้ผลการสอบของนักเรียน นำมาหาค่าพารามิเตอร์ความยากโดยการวิเคราะห์ตามทฤษฎี IRT ค่าพารามิเตอร์เหล่านี้จะถูกนำมาแปลงเข้าสู่สูตรหาค่า θ เพื่อใช้เป็นข้อมูลจัดทำคู่มือจัดเรียงข้อสอบ

สำหรับสูตรการหาค่า θ ในงานวิจัยนี้ มาจาก Wright and Stone (1979 cited in Cizek & Bunch, 2007) ได้เสนอสูตรดั้งเดิมแบบโมเดล Rasch สำหรับข้อสอบแบบเลือกตอบ ที่มีระบบการให้คะแนนเป็น 1 0 โดยกำหนดความน่าจะเป็นในการตอบถูก, $P(X=1)$ เมื่อกำหนดค่าความยาก คือ β_j และความสามารถของผู้สอบ (θ_i) ดังนี้

$$P(x = 1/\theta_i, \beta_j) = \exp(\theta_i - \beta_j) / [1 + \exp(\theta_i - \beta_j)] \quad (1)$$

เมื่อกำหนด β_j = ค่าความยากของข้อสอบ

θ_i = ความสามารถของผู้สอบ

P = ความน่าจะเป็นของการตอบถูก

$\exp$ = natural logarithm e (2.71828...)

เมื่อ แทนค่า $p = 2/3$ แล้วแก้สมการหา θ_i ดังนี้

$$\exp(\theta_i - \beta_j) / [1 + \exp(\theta_i - \beta_j)] = 2/3 \quad (2)$$

$$\exp(\theta_i - \beta_j) = 2/3 * [1 + \exp(\theta_i - \beta_j)] \quad (3)$$

$$\exp(\theta_i - \beta_j) = 2/3 + 2/3 * \exp(\theta_i - \beta_j) \quad (4)$$

$$\exp(\theta_i - \beta_j) - 2/3 * \exp(\theta_i - \beta_j) = 2/3 \quad (5)$$

$$1/3 * \exp((\theta_i - \beta_j)) = 2/3 \quad (6)$$

$$\exp((\theta_i - \beta_j)) = 2/3 \div 1/3 \quad (7)$$

$$\exp((\theta_i - \beta_j)) = 2 \quad (8)$$

$$\theta_i - \beta_j = .693 \quad (9)$$

$$\theta_i = \beta_j + .693 \quad (10)$$

สำหรับสูตรการหาค่า θ จากข้อสอบแบบให้คะแนนหลายค่า เช่น ข้อสอบอัตนัย มาจาก Wright and Masters (1982 cited in Cizek & Bunch, 2007) ได้เสนอสูตรตามกรอบแนวคิดของ Rasch โมเดล Partial-Credit Model (PCM) โดยยกตัวอย่างข้อสอบที่ให้คะแนนแบบ 6 ค่า (0, 1, 2, 3, 4, 5) ซึ่งการคำนวณจากสูตรข้างล่างนี้ ผู้วิจัยใช้โปรแกรม EXCEL ช่วยในการคำนวณ

$$\text{กำหนด } \pi_{nix} = \frac{\exp \sum \theta_n - \delta_{ij}}{\sum \exp \sum (\theta_n - \delta_{ij})} \quad (11)$$

เมื่อกำหนด π_{nix} = likelihood ของผู้สอบแต่ละคน

θ_n = ความสามารถของผู้สอบ

δ_{ij} = ค่าความยากของข้อสอบข้อที่ i ณ คะแนน j

i = ข้อสอบ

j = คะแนน

เมื่อคะแนนเท่ากับ 0 ค่า $\delta_{io} \equiv 0$ แทนค่าในสมการ 11 ได้ดังนี้

$$\sum (\theta_n - \delta_{ij}) = 0 \text{ และ } \exp \sum (\theta_n - \delta_{ij}) = 1 \quad (12)$$

$$\begin{aligned} \text{ขั้นที่ 1 } \sum (\theta_n - \delta_{ij}) &= \sum (\theta_n - \delta_{io}) + \theta_n - \delta_{i1} \\ &= 0 + \theta_n - \delta_{i1} \\ &= \theta_n - \delta_{i1} \end{aligned} \quad (13)$$

ขั้น 2-4 ทำคล้ายขั้น 1 ได้สมการดังนี้

$$\text{ขั้นที่ 2 } \sum (\theta_n - \delta_{ij}) = 2\theta_n - \delta_{i1} - \delta_{i2} \quad (14)$$

$$\text{ขั้นที่ 3 } \sum (\theta_n - \delta_{ij}) = 3\theta_n - \delta_{i1} - \delta_{i2} - \delta_{i3} \quad (15)$$

$$\text{ขั้นที่ 4 } \sum (\theta_n - \delta_{ij}) = 4\theta_n - \delta_{i1} - \delta_{i2} - \delta_{i3} - \delta_{i4} \quad (16)$$

$$\text{ขั้นที่ 5 } \sum (\theta_n - \delta_{ij}) = 4\theta_n - \delta_{i1} - \delta_{i2} - \delta_{i3} - \delta_{i4} - \delta_{i5} \quad (17)$$

นำค่า exponential มาคูณค่า summation ของสมการข้างต้นได้ดังนี้

$$\text{ขั้นที่ 1 คะแนนเท่ากับ 0 } \exp(0) \quad (18)$$

$$\text{ขั้นที่ 2 คะแนนเท่ากับ 1 } \exp(\theta_n - \delta_{i1}) \quad (19)$$

$$\text{ขั้นที่ 3 คะแนนเท่ากับ 2 } \exp(2\theta_n - \delta_{i1} - \delta_{i2}) \quad (20)$$

$$\text{ขั้นที่ 4 คะแนนเท่ากับ 3 } \exp(3\theta_n - \delta_{i1} - \delta_{i2} - \delta_{i3}) \quad (21)$$

$$\text{ขั้นที่ 5 คะแนนเท่ากับ 4 } \exp(4\theta_n - \delta_{i1} - \delta_{i2} - \delta_{i3} - \delta_{i4}) \quad (22)$$

$$\text{ขั้นที่ 6 คะแนนเท่ากับ 5 } \exp(5\theta_n - \delta_{i1} - \delta_{i2} - \delta_{i3} - \delta_{i4} - \delta_{i5}) \quad (23)$$