

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ สำหรับแบบสอบคัดเลือกวิชาภาษาไทยและวิชาภาษาอังกฤษด้วยวิธี MH สำหรับผู้สอบจำแนกตามเพศภูมิสำเนา ประสบการณ์ในการสอบ และสังกัดของสถานศึกษา ซึ่งมีวิธีดำเนินการวิจัยตามขั้นตอนดังต่อไปนี้

ผลวิจัย

ผลวิจัยที่ใช้ในการวิจัยครั้งนี้ คือ ผู้สอบคัดเลือกวิชาภาษาไทยและภาษาอังกฤษระดับบัณฑิตศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย โดยศูนย์ทดสอบทางการศึกษา ได้มีการจัดดำเนินการสอบรหัส 383 ประกอบด้วยผู้สอบวิชาภาษาไทยจำนวน 506 คน และผู้สอบวิชาภาษาอังกฤษจำนวน 501 คน

เครื่องมือ

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ประกอบด้วย

1. แบบสอบวิชาภาษาไทย (ศทท 383) สร้างโดยคณะกรรมการพัฒนาแบบสอบภาษาไทยของศูนย์ทดสอบทางการศึกษา ซึ่งนำมาใช้เฉพาะส่วนที่เป็นข้อสอบปรนัยชนิด 4 ตัวเลือก จำนวน 95 ข้อ

2. แบบสอบวิชาภาษาอังกฤษ (ศทอ 383) สร้างโดยคณะกรรมการพัฒนาแบบสอบภาษาอังกฤษของศูนย์ทดสอบทางการศึกษา ซึ่งนำมาใช้เฉพาะส่วนที่เป็นข้อสอบปรนัยชนิด 5 ตัวเลือก จำนวน 80 ข้อ

แบบสอบคัดเลือกทั้ง 2 ฉบับ ศูนย์ทดสอบทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์ มหาวิทยาลัยได้ทำการวิเคราะห์คุณภาพของข้อสอบไว้แล้ว ซึ่งผลการวิเคราะห์วิชาภาษาไทยและวิชาภาษาอังกฤษมีค่าความเที่ยงแบบความสอดคล้องภายใน (KR-20) เท่ากับ 0.70 และ 0.77 ตามลำดับ

ข้อมูลและการเก็บรวบรวมข้อมูล

ข้อมูลสำหรับการวิจัยมีดังต่อไปนี้

1. ข้อมูลในการแบ่งกลุ่มย่อยได้แก่ตัวแปรเพศ ภูมิสำเนา ประสบการณ์ในการสอบ และสังกัดของสถานศึกษาของผู้สอบ รวบรวมจากหลักฐานใบสมัครของผู้สอบ
 2. คะแนนผลการสอบรายข้อ วิชาภาษาไทยและวิชาภาษาอังกฤษ รวบรวมจากแผ่นบันทึกแม่เหล็ก (diskette) ซึ่งบันทึกผลการสอบจากกระดาษคำตอบของผู้สอบ
- การเก็บรวบรวมข้อมูลมีวิธีดำเนินการดังนี้
1. ผู้วิจัยทำหนังสือขออนุญาตจากบัณฑิตวิทยาลัย เพื่อติดต่อขอข้อมูลหรือผลการสอบจากศูนย์ทดสอบทางการศึกษา
 2. ผู้วิจัยคัดลอกข้อมูลที่ต้องการ ได้แก่ข้อมูลตัวแปรที่ใช้ในการแบ่งกลุ่มย่อย และข้อมูลผลการสอบวิชาภาษาไทยและภาษาอังกฤษ
 3. ผู้วิจัยตรวจทานความถูกต้อง และความสมบูรณ์ของข้อมูลของแบบสอบทั้งวิชาภาษาไทยและวิชาภาษาอังกฤษ

เนื่องจากข้อจำกัดของโปรแกรมการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ MH_{DF} ของ Fidalgo A. (1995) สามารถวิเคราะห์ข้อสอบได้เพียง 75 ข้อ ผู้วิจัยจึงทำการสุ่มข้อสอบของแต่ละฉบับออกให้เหลือเท่ากับ 75 ข้อ เพื่อให้โปรแกรมดำเนินการวิเคราะห์ได้โดยทำการสุ่มข้อสอบออกดังนี้

วิชาภาษาไทยมีข้อสอบปรนัยทั้งหมด 95 ข้อ ซึ่งประกอบด้วยข้อสอบตามโครงสร้างของเนื้อหาที่ต้องการวัดใน 2 องค์ประกอบ คือ การใช้ภาษาไทย (T₁) และเหตุผลเชิงภาษา (T₂) จากลักษณะเนื้อหา 5 ด้าน ได้แก่ความสามารถในด้านต่อไปนี้

| | | | |
|-----------------|-------------|----|-----|
| 1. หลักภาษา | จำนวนข้อสอบ | 20 | ข้อ |
| 2. การใช้ภาษา 1 | จำนวนข้อสอบ | 23 | ข้อ |
| 3. การใช้ภาษา 2 | จำนวนข้อสอบ | 15 | ข้อ |
| 4. วรรณคดี | จำนวนข้อสอบ | 19 | ข้อ |
| 5. ครรภวิทยา | จำนวนข้อสอบ | 18 | ข้อ |

ผู้วิจัยทำการสุ่มข้อสอบออกทั้งหมดจำนวน 20 ข้อ ด้วยการสุ่มอย่างง่ายตามสัดส่วนโครงสร้างเนื้อหาของแบบสอบแต่ละด้านทั้ง 5 ด้าน ผลการสุ่มได้ข้อสอบเพื่อตัดออกดังแสดงรายละเอียดในตารางที่ 3

ตารางที่ 3 จำนวนข้อสอบที่ตัดออกจากแบบสอบวิชาภาษาไทย จำนวนตามโครงสร้างเนื้อหา

| ลักษณะเนื้อหา | จำนวนข้อที่ตัด | ข้อที่ |
|-----------------|----------------|--------------------|
| 1. หลักภาษา | 4 | 3, 10, 12, 19 |
| 2. การใช้ภาษา 1 | 5 | 22, 26, 31, 34, 40 |
| 3. การใช้ภาษา 2 | 3 | 84, 92, 95 |
| 4. วรรณคดี | 4 | 62, 70, 74, 78 |
| 5. ตรรกวิทยา | 4 | 46, 49, 51, 61 |
| รวมทั้งฉบับ | 20 | |

ดังนั้นการวิจัยครั้งนี้ จึงใช้ข้อสอบวิชาภาษาไทยจำนวน 75 ข้อ จากนั้นผู้วิจัยจัดการเรียงลำดับข้อสอบใหม่แล้วแยกตามโครงสร้างของเนื้อหาที่ต้องการวัดได้ 5 ด้าน รายละเอียดดังตารางที่ 4

ตารางที่ 4 ลักษณะของข้อสอบวิชาภาษาไทยจำนวนตามโครงสร้างเนื้อหา

| ลักษณะเนื้อหา | จำนวนข้อ (%) | ข้อที่ |
|-----------------|--------------|-----------|
| 1. หลักภาษา | 16 (21.33) | 1. - 16. |
| 2. การใช้ภาษา 1 | 18 (24.00) | 17. - 34. |
| 3. การใช้ภาษา 2 | 12 (16.00) | 64. - 75. |
| 4. วรรณคดี | 15 (20.00) | 49. - 63. |
| 5. ตรรกวิทยา | 14 (18.67) | 35. - 48. |
| รวมทั้งฉบับ | 75 | |

สำหรับวิชาภาษาอังกฤษจากเดิมมีข้อสอบปรนัยทั้งหมด 80 ข้อ ซึ่งประกอบด้วยข้อสอบตามโครงสร้างของเนื้อหาที่ต้องการวัดใน 2 องค์ประกอบ คือ ความสามารถในการอ่าน (Reading) และการบูรณาการ (Integrate) จากลักษณะเนื้อหา 7 ด้าน ได้แก่ความสามารถในด้านต่อไปนี้

| | | |
|----------------------|-------------|--------|
| 1. การจับใจความสำคัญ | จำนวนข้อสอบ | 6 ข้อ |
| 2. การให้รายละเอียด | จำนวนข้อสอบ | 17 ข้อ |
| 3. การถอดความ | จำนวนข้อสอบ | 7 ข้อ |

| | | |
|--------------------------|-------------|--------|
| 4. การอ้างอิง | จำนวนข้อสอบ | 9 ข้อ |
| 5. การสรุปข้อความ | จำนวนข้อสอบ | 8 ข้อ |
| 6. การบอกความหมายคำศัพท์ | จำนวนข้อสอบ | 13 ข้อ |
| 7. การใช้ภาษา | จำนวนข้อสอบ | 20 ข้อ |

ผู้วิจัยทำการสุ่มข้อสอบออกจำนวน 5 ข้อ โดยใช้วิธีการสุ่มแบบหลายขั้นตอน (multi-stage) โดยดำเนินการดังนี้

1. ใช้วิธีการสุ่มอย่างง่าย ทำการสุ่มโครงสร้างของเนื้อหาของแบบสอบจากเดิม 7 ด้าน ได้โครงสร้างของเนื้อหาที่ต้องการสุ่มออก 5 ด้าน ได้แก่ความสามารถในด้านการให้รายละเอียด การอ้างอิง การสรุปข้อความ การบอกความหมายคำศัพท์ และความสามารถในด้านการใช้ภาษา

2. ใช้วิธีการสุ่มอย่างง่ายสุ่มข้อสอบของแต่ละด้านออกด้านละ 1 ข้อ ข้อสอบที่สุ่มออก ได้แก่ข้อที่ 20, 26, 50, 60, และข้อที่ 80 รวม 5 ข้อ

ดังนั้นในการวิจัยครั้งนี้ ข้อมูลการตอบของแบบสอบวิชาภาษาอังกฤษที่นำไปใช้ในการวิจัยเท่ากับ 75 ข้อ จากนั้นผู้วิจัยจัดการเรียงลำดับข้อสอบใหม่แล้วแยกตามโครงสร้างของเนื้อหาที่ต้องการวัดได้ 7 ด้าน รายละเอียดดังตารางที่ 5

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5 ลักษณะของข้อสอบวิชาภาษาอังกฤษจำแนกตามโครงสร้างเนื้อหา

| ลักษณะเนื้อหา | จำนวนข้อ (%) | ข้อที่ |
|--------------------------|-----------------|--|
| 1. การจับใจความสำคัญ | 6 (8.00) | 1,12,20,29,39,48 |
| 2. การให้รายละเอียด | 16 (21.33) | 2,3,11,15,16,18,21,27,31, 35, 37,40,41,42,49,50 |
| 3. การถอดความ | 7 (9.33) | 5,6,30,32,43,51,52 |
| 4. การอ้างอิง | 8 (10.67) | 23,24,34,36,45,46,53,55 |
| 5. การสรุปใจความ | 7 (9.33) | 4,13,25,28,38,44,54 |
| 6. การบอกความหมายคำศัพท์ | 12 (16.00) | 7,8,9,10,14,17,19,22,26, 33,47,56 |
| 7. การใช้ภาษา | 19 (25.33) | 57-75 |
| รวมทั้งฉบับ | 75 | |

การวิเคราะห์ข้อมูล

ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลตามลำดับต่อไปนี้

1. คำนวณค่าสถิติพื้นฐาน เพื่อบรรยายลักษณะการแจกแจงของคะแนนของแบบสอบวิชาภาษาไทยและภาษาอังกฤษ
2. ใช้โปรแกรมวิเคราะห์ข้อสอบ CTIA/Item (สุทัศน์ สุกมลสันต์,2535) ซึ่งเป็นโปรแกรมวิเคราะห์ข้อสอบตามทฤษฎีแบบดั้งเดิม(Classical Test Theory) โดยดำเนินการ ดังนี้

2.1 วิเคราะห์คุณภาพของข้อสอบวิชาภาษาไทยและภาษาอังกฤษ โดยการวิเคราะห์ค่าความยาก (p) และค่าอำนาจจำแนก(r) ของข้อสอบรายข้อจากสูตร

$$p = \frac{P_H + P_L}{2}$$

และ $r = P_H - P_L$

- เมื่อ p = ค่าความยากของข้อสอบ
 r = ค่าอำนาจจำแนกของข้อสอบ
 P_H = สัดส่วนของคนที่ตอบถูกในกลุ่มสูง
 P_L = สัดส่วนของคนที่ตอบถูกในกลุ่มต่ำ

2.2 วิเคราะห์คุณภาพของแบบสอบวิชาภาษาไทยและภาษาอังกฤษ โดยการคำนวณค่าความเที่ยงเชิงความสอดคล้องภายในของแบบสอบแต่ละฉบับ ด้วยสูตร KR-20 ของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson) ดังนี้

$$r_{tt} = \left[\frac{n}{n-1} \right] \left[1 - \frac{\sum pq}{S_x^2} \right]$$

- เมื่อ r_{tt} = ค่าความเที่ยง
 n = จำนวนข้อสอบ
 $\sum pq$ = ผลรวมของความแปรปรวนรายข้อ
 S_x^2 = ความแปรปรวนของคะแนนรวม
 p = สัดส่วนของผู้ที่ตอบถูก
 q = สัดส่วนของผู้ที่ตอบผิด (1-p)

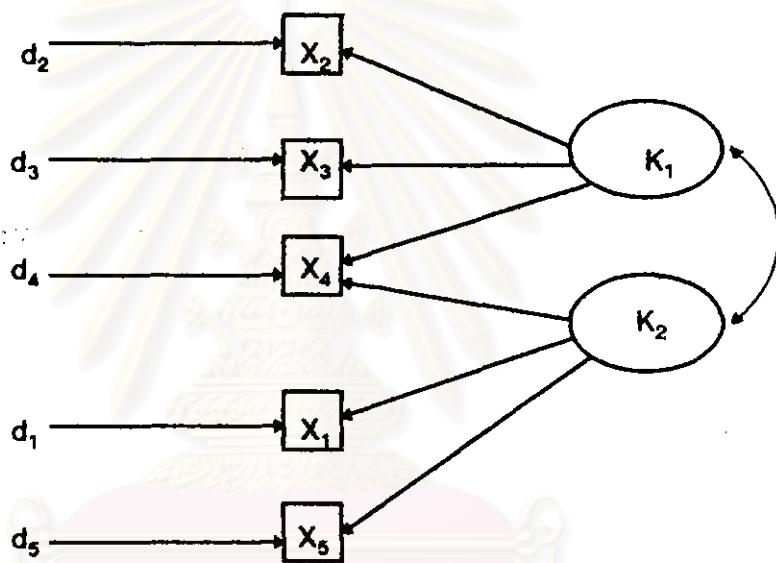
3. ทดสอบความตรงเชิงโครงสร้างของแบบสอบวิชาภาษาไทยและภาษาอังกฤษ ด้วยวิธีวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) โดยดำเนินการดังนี้

3.1 คำนวณค่าเฉลี่ย (\bar{X}) และส่วนเบี่ยงเบนมาตรฐาน (S.D.) ของคะแนนเป็นรายด้านของผู้สอบแต่ละคน และคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันของข้อสอบแต่ละด้าน โดยใช้โปรแกรมคอมพิวเตอร์ SPSS/PC

3.2 วิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) ของแบบสอบวิชาภาษาไทยและภาษาอังกฤษ โดยใช้โปรแกรมคอมพิวเตอร์ LISREL 8.10 ซึ่งมีโมเดลองค์ประกอบของแบบสอบดังนี้

3.2.1 โมเดลองค์ประกอบของแบบสอบวิชาภาษาไทย ลักษณะของโมเดลเชิงโครงสร้างของแบบสอบวิชาภาษาไทยในการวิเคราะห์ยืนยันความตรงเชิงโครงสร้างหรือทฤษฎีนั้นแสดงได้ดังแผนภาพที่ 1 ดังนี้

แผนภาพที่ 1 โมเดลวิเคราะห์องค์ประกอบของแบบสอบวิชาภาษาไทย



เมื่อ X_1 = คะแนนหลักภาษา

X_2 = คะแนนการใช้ภาษา 1

X_3 = คะแนนการใช้ภาษา 2

X_4 = คะแนนวรรณคดี

X_5 = คะแนนตรรกวิทยา

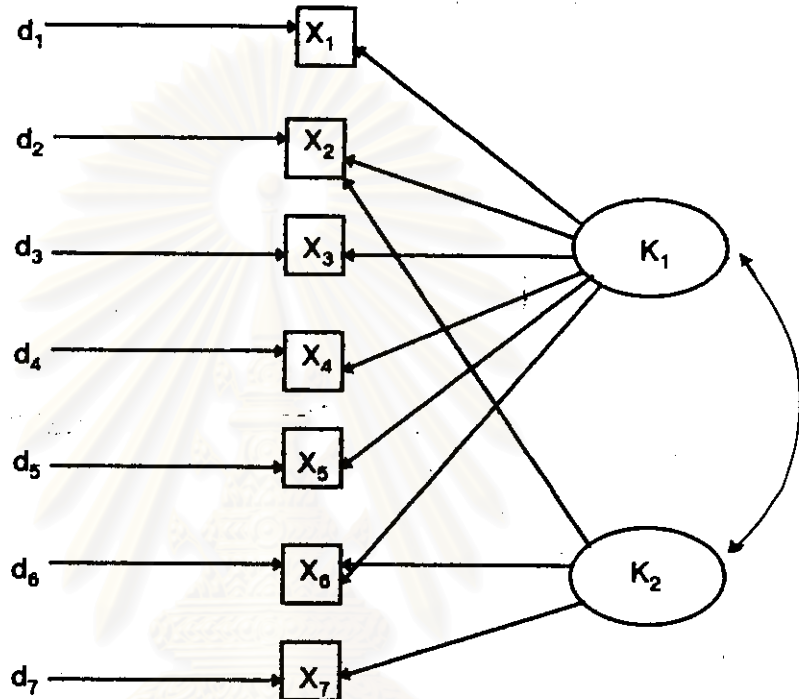
$K_1 = T_1$ = การใช้ภาษาไทย

$K_2 = T_2$ = เหตุผลเชิงภาษา

d_1 ถึง d_5 = ความคลาดเคลื่อนของคะแนนแต่ละด้าน

3.2.2 โมเดลองค์ประกอบของแบบสอบวิชาภาษาอังกฤษ ลักษณะของโมเดลเชิงโครงสร้างของแบบสอบวิชาภาษาอังกฤษในการวิเคราะห์ยืนยันความตรงเชิงโครงสร้างหรือทฤษฎีนั้นแสดงได้ดังแผนภาพที่ 2 ดังนี้

แผนภาพที่ 2 โมเดลวิเคราะห์องค์ประกอบของแบบสอบวิชาภาษาอังกฤษ



- เมื่อ X_1 = คะแนนการจับใจความสำคัญ
 X_2 = คะแนนการให้รายละเอียด
 X_3 = คะแนนการถอดความ
 X_4 = คะแนนการอ้างอิง
 X_5 = คะแนนการสรุปอิงความ
 X_6 = คะแนนการบอกความหมายคำศัพท์
 X_7 = คะแนนการใช้ภาษา
 K_1 = reading = ความสามารถในการอ่าน
 K_2 = integrate = ความสามารถในการบูรณาการ
 d_1 ถึง d_7 = ความคลาดเคลื่อนคะแนนแต่ละด้าน

4. ทำการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (uniform DIF)

จำแนกตามกลุ่มตัวแปร เพศ ภูมิภาค ประสิทธิภาพในการสอบ และสังกัดของสถานศึกษา โดยใช้โปรแกรมสำเร็จรูป MH_{DF} ของ Fidalgo A. (1995) ซึ่งดำเนินการดังนี้

4.1 ตรวจสอบผลการสอบวิชาภาษาไทยและภาษาอังกฤษรายข้อ โดยการให้คะแนนแบบ 2 ค่า (dichotomous score) ตอบถูกให้ 1 ตอบผิดให้ 0

4.2 กำหนดชนิดของการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ . และกำหนดเกณฑ์ในการจับคู่ชั้นคะแนน (Score Category) ซึ่งการวิจัยครั้งนี้ผู้วิจัยใช้เกณฑ์ในการจับคู่ชั้นคะแนนเป็นคะแนนที่ผู้สอบสามารถทำได้ (Total Score) จำนวนมากที่สุดโดยการเปรียบเทียบคู่ชั้นคะแนนที่ห่างกันช่วงละ 1 คะแนน จากคะแนนรวมต่ำสุดถึงคะแนนรวมสูงสุดที่ผู้สอบทั้งหมดทำได้

4.3 ค่าความน่าจะเป็นในรูปแบบสัดส่วนจากความถี่ของการตอบข้อสอบถูกและผิดระหว่างกลุ่มผู้สอบในกลุ่มอ้างอิง (R) และกลุ่มเปรียบเทียบ (F) สำหรับช่วงคะแนน j ดังตารางที่ 6

ตารางที่ 6 ความถี่ของผู้สอบที่ตอบถูกและผิดในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ สำหรับช่วงคะแนน j

| กลุ่มผู้สอบ | คะแนนของข้อสอบที่ต้องการวิเคราะห์ DIF | | |
|-------------|---------------------------------------|-----------|----------|
| | ตอบถูก(1) | ตอบผิด(0) | รวม |
| R | A_j | B_j | N_{Rj} |
| F | C_j | D_j | N_{Fj} |
| รวม | m_{1j} | m_{0j} | N_j |

จากนั้นคำนวณสัดส่วนของการตอบของผู้สอบในทุกช่วงคะแนน จากสูตร ดังนี้

$$\alpha_{MH} = \frac{\sum A_j D_j / N_j}{\sum B_j C_j / N_j}$$

เมื่อ α_{MH} = สัดส่วนของการตอบข้อสอบถูกและผิดระหว่างกลุ่มในแต่ละข้อในทุกช่วงคะแนน j

A_j = เป็นความถี่ที่สังเกตได้ในการตอบถูกในช่วงคะแนน j ของกลุ่ม R

B_j = เป็นความถี่ที่สังเกตได้ในการตอบผิดในช่วงคะแนน j ของกลุ่ม R

C_j = เป็นความถี่ที่สังเกตได้ในการตอบถูกในช่วงคะแนน j ของกลุ่ม F

D_j = เป็นความถี่ที่สังเกตได้ในการตอบผิดในช่วงคะแนน j ของกลุ่ม F

m_{1j} = จำนวนผู้ตอบถูกทั้งหมด ในช่วงคะแนน j

m_{0j} = จำนวนผู้ตอบผิดทั้งหมด ในช่วงคะแนน j

N_j = จำนวนผู้สอบทั้งหมด

4.4 ทดสอบนัยสำคัญของค่าสถิติไค-สแควร์ เพื่อทดสอบค่า αMH ที่คำนวณได้ว่าแตกต่างจาก 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 หรือไม่ ที่ระดับชั้นความเป็นอิสระเท่ากับ 1

4.5 แปลงค่า αMH ให้เป็นค่าเคลต้า หรือ MH_{DIF} โดยใช้สูตรดังนี้

$$MH_{DIF} = -2.35 \ln(\alpha MH)$$

จากนั้นทดสอบนัยสำคัญของค่าสถิติไค-สแควร์เพื่อทดสอบค่า MH_{DIF} ที่คำนวณได้ว่าแตกต่างจาก 0 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 หรือไม่ที่ระดับชั้นความเป็นอิสระเท่ากับ 1 เพื่อนำไปพิจารณาข้อสอบ ซึ่งมีเกณฑ์ในการพิจารณาค่าของ MH_{DIF} ดังนี้

1). ค่า MH_{DIF} เท่ากับ 0 หรือไม่แตกต่างจาก 0 อย่างมีนัยสำคัญทางสถิติ แสดงว่าข้อสอบนั้นทำหน้าที่ไม่แตกต่างกันระหว่างกลุ่ม (no DIF)

2). ค่า MH_{DIF} แตกต่างจาก 0 อย่างมีนัยสำคัญทางสถิติ และมีค่าเป็นบวก (positive) แสดงว่าข้อสอบนั้นทำหน้าที่ต่างกันระหว่างกลุ่ม โดยจะเข้าข้าง (favor) กลุ่มเปรียบเทียบ (F)

3). ค่า MH_{DIF} แตกต่างจาก 0 อย่างมีนัยสำคัญทางสถิติ และมีค่าเป็นลบ (negative) แสดงว่าข้อสอบนั้นทำหน้าที่ต่างกันระหว่างกลุ่ม โดยจะเข้าข้าง (favor) กลุ่มอ้างอิง (R)

5. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูป (nonuniform DIF) โดยใช้โปรแกรมคอมพิวเตอร์ MH_{DIF} ซึ่งมีวิธีดำเนินการเตรียมข้อมูลโดยการแปลงคะแนนเช่นเดียวกับ ข้อ 4.1 จากนั้นกำหนดชนิดของการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบให้เป็นแบบอเนกรูป ซึ่งมีหลักในการคำนวณคล้ายกับการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป แต่จะต่างกันตรงที่การคัดเลือกผู้สอบเข้ากลุ่ม โดยจะพิจารณาช่วงระดับความสามารถของผู้สอบหรือคะแนนเป็นเกณฑ์เพิ่มเติมด้วย กล่าวคือจะใช้ค่าเฉลี่ยของคะแนนรวมทั้งหมดเป็นเกณฑ์ในการคัดเลือกผู้สอบ สำหรับผู้สอบที่จัดให้เป็นกลุ่มที่มีความสามารถสูงได้แก่ผู้สอบที่ได้คะแนนรวมสูงกว่าค่าเฉลี่ย ส่วนผู้สอบที่จัดให้เป็นกลุ่มที่มีความสามารถต่ำได้แก่ผู้สอบที่ได้คะแนนรวมน้อยกว่าหรือเท่ากับค่าเฉลี่ย จากนั้นจะคำนวณค่าความน่าจะเป็นในรูปของสัดส่วนของการตอบข้อสอบถูกและผิดระหว่างกลุ่มอ้างอิง (R) และกลุ่มเปรียบเทียบ (F) โดยใช้หลักการคำนวณเช่นเดียวกับข้อ 4.3 -4.5

6. สร้างตารางเปรียบเทียบจำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้งแบบเอกรูปและแบบอเนกรูปในด้านค่าความยาก ค่าอำนาจจำแนก และลักษณะเนื้อหาของข้อสอบรายข้อ ในแบบสอบวิชาภาษาไทยและภาษาอังกฤษ

7. วิเคราะห์คุณภาพของแบบสอบหลังการตัดข้อสอบที่ทำหน้าที่ต่างกันแบบเอกรูป หรือ แบบอนเอกรูป จำแนกตามตัวแปรต่างๆ โดยดำเนินการดังนี้

7.1 คำนวณค่าความเที่ยงเชิงความสอดคล้องภายในของแบบสอบแต่ละฉบับหลัง การข้อสอบที่ทำหน้าที่ต่างกันโดยใช้สูตร KR-20 ของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson) ดังนี้

$$r_{tt} = \left[\frac{n}{n-1} \right] \left[1 - \frac{\sum pq}{S_x^2} \right]$$

เมื่อ r_{tt} = ค่าความเที่ยง

n = จำนวนข้อสอบ

$\sum pq$ = ผลรวมของความแปรปรวนรายข้อ

S_x^2 = ความแปรปรวนของคะแนนรวม

p = สัดส่วนของผู้ที่ตอบถูก

q = สัดส่วนของผู้ที่ตอบผิด ($1-p$)

7.2 วิเคราะห์ความตรงเชิงโครงสร้างหรือความตรงตามทฤษฎีของแบบสอบหลัง การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธีวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) โดยดำเนินการเช่นเดียวกับข้อ 3.1 และข้อ 3.2 แต่ตัดข้อสอบที่พบ DIF จากการวิเคราะห์ การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูปหรือแบบอนเอกรูป เมื่อจำแนกผู้สอบตามตัวแปร เพศ สภาพสังคม ประสบการณ์ในการสอบและสังกัดของสถานศึกษาออกจากแบบสอบวิชา ภาษาไทยและภาษาอังกฤษ โดยพิจารณาจากค่าสถิติไค-สแควร์ (chi-square) ดัชนี วัดระดับความกลมกลืน (GFI) ดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) และดัชนีวัด กำลังสองเฉลี่ยของเศษ (RMR)

8. เปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยง ก่อนและหลังการตัดข้อสอบ ที่ทำหน้าที่ต่างกันออกจากแบบสอบในทุกเงื่อนไขของการวิเคราะห์โดยดำเนินการดังนี้

8.1 ทำการปรับขยายจำนวนข้อสอบให้เท่ากับ 75 ข้อ ในทุกกรณีที่เกิดข้อสอบที่ ทำหน้าที่ต่างกันออกจากแบบสอบเพื่อให้จำนวนข้อสอบเท่ากัน และสามารถนำมาเปรียบเทียบค่าความเที่ยงกันได้ จากนั้นคำนวณค่าความเที่ยงของแบบสอบฉบับใหม่โดยใช้สูตรของ สเปียร์แมน : Spearman C. (Allen J.M. and Yen.W., 1979) ดังนี้

$$r'_n = \frac{N (r_n)}{1 + (N - 1) r_n}$$

- เมื่อ r'_n = ค่าความเที่ยงของแบบสอบฉบับใหม่
 r_n = ค่าความเที่ยงของแบบสอบฉบับเดิม
 N = จำนวนเท่าของข้อสอบที่เพิ่มขึ้น

8.2 ทดสอบนัยสำคัญความแตกต่างของค่าความเที่ยงฉบับใหม่ที่ได้จากการปรับขยายจำนวนข้อสอบหลังจากตัดข้อสอบที่ทำหน้าที่ต่างกันออกจากแบบสอบในทุกเงื่อนไขของการวิเคราะห์ กับค่าความเที่ยงของแบบสอบฉบับเดิม โดยใช้สูตรดังนี้ (พิศิษฐ์ คัณฑวณิช, 2532 ; Felde, 1980)

$$t = \frac{(r_n - r'_n) \sqrt{(N-2)}}{\sqrt{4(1-r_n)(1-r'_n)(1-r_{t12})}}$$

- เมื่อ t = ค่าสถิติทดสอบที่ (t-test)
 r_n = ค่าความเที่ยงของแบบสอบฉบับเดิม (ก่อนตัดข้อสอบที่ DIF)
 r'_n = ค่าความเที่ยงของแบบสอบฉบับใหม่ (ปรับขยายข้อสอบหลังตัดข้อสอบที่พบ DIF)
 r_{t12} = ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบฉบับเดิมกับฉบับใหม่
 N = จำนวนกลุ่มตัวอย่าง

9. เปรียบเทียบคุณภาพด้านความตรงตามทฤษฎีของโมเดลแบบสอบวิชาภาษาไทยและภาษาอังกฤษ โดยเปรียบเทียบจากค่าไค-สแควร์สัมพัทธ์ (Relative chi-square: χ^2/df) ของแบบสอบทั้งกรณีก่อนและหลังการตัดข้อสอบที่พบ DIF ออกจากแบบสอบ