บทที่ 3

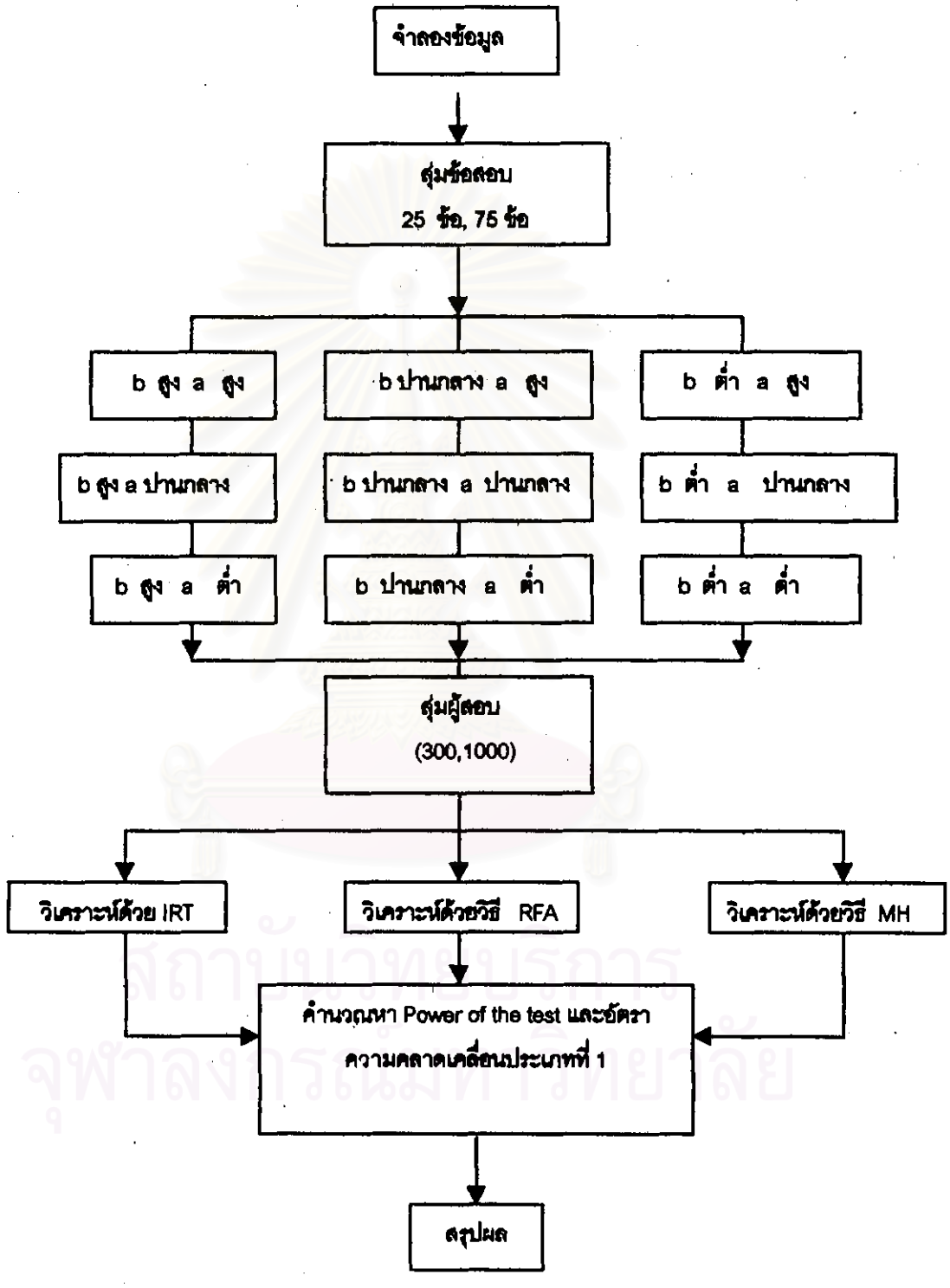
วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ใช้ระเบียบวิธีการวิจัยเชิงทดลองโดยมีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ 3 วิธีคือ วิธีการวิเคราะห์องค์ประกอบจำกัด (RFA) วิธีแมนเทล-แฮนส์เซล (MH) และวิธีการตอบสนองข้อสอบ (IRT) แบบ 2 พารามิเตอร์ เมื่อมีการจัดกลุ่มค่าความยากของข้อสอบออกเป็น 3 ระดับคือ ข้อสอบที่มีค่าความยากสูง ปานกลาง และต่ำ กลุ่มค่าอำนาจจำแนกแบ่งออกเป็น 3 ระดับคือ ข้อสอบที่มีค่าอำนาจจำแนกสูง ปานกลาง และต่ำ ขนาดกลุ่มตัวอย่างแบ่งออกเป็น 2 กลุ่ม คือกลุ่มตัวอย่างขนาดเล็ก (300 คน) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) ความยาวของแบบสอบแบ่งออกเป็น 2 ขนาดคือแบบสอบสั้น (25 ข้อ) และแบบสอบยาว (75 ข้อ) ขนาดความลำเอียงของข้อสอบแบ่งออกเป็น 2 ขนาดคือ ลำเอียงสูงและลำเอียงต่ำ โดยจะนำเสนอเนื้อหาเป็นตอน ๆ ดังนี้

1. การจำลองข้อมูล
2. การสุ่มข้อสอบ
3. การสุ่มผู้สอบ
4. การวิเคราะห์ข้อมูลด้วยวิธีการวิเคราะห์องค์ประกอบจำกัด (RFA)
5. การวิเคราะห์ข้อมูลด้วยวิธีแมนเทล-แฮนส์เซล (MH)
6. การวิเคราะห์ข้อมูลด้วยวิธีการตอบสนองข้อสอบ (IRT) แบบ 2 พารามิเตอร์
7. การคำนวณอัตราความคลาดเคลื่อนประเภทที่ 1 และคำนวณหาค่าอำนาจการตรวจสอบ (Power of the test) ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบของแต่ละวิธี
8. เกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบทั้ง 3 วิธี

การดำเนินการวิจัยทั้ง 8 ขั้นตอนนี้สามารถเขียนเป็นแผนภาพแสดงขั้นตอนการดำเนินการวิจัยทั้งหมดได้ดังแผนภาพที่ 1

แผนภาพที่ 1 ขั้นตอนการดำเนินการวิจัย (ดัดแปลงมาจาก รัชรินทร์ มุคดา,2540)



จากแผนภาพที่ 1 ขั้นตอนแรกเป็นการจำลองข้อมูล ด้วยโปรแกรม SPSS for Windows ขั้นตอนที่ 2 เป็นการสุ่มข้อสอบเพื่อให้ได้แบบสอบที่จะศึกษา 2 ขนาดคือแบบสอบสั้น 25 ข้อ และแบบสอบยาว 75 ข้อ แต่ละแบบสอบมีค่าความยาก ค่าอำนาจจำแนกและขนาดความลำเอียงตามเงื่อนไขที่กำหนด ขั้นตอนที่ 3 เป็นการสุ่มขนาดของกลุ่มตัวอย่าง ซึ่งในการวิจัยครั้งนี้กำหนดกลุ่มตัวอย่างออกเป็น 2 ขนาดคือ กลุ่มตัวอย่างขนาดเล็ก (300 คน) และกลุ่มตัวอย่างขนาดใหญ่ (1,000 คน) แต่ละขนาดแบ่งเป็น 2 กลุ่มย่อยคือ กลุ่มอ้างอิง และกลุ่มเปรียบเทียบโดยมีอัตราส่วนระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็น 1:1 ขั้นตอนที่ 4 เป็นการวิเคราะห์ข้อมูลในแต่ละเงื่อนไขที่ได้จากขั้นตอนที่ 2 และขั้นตอนที่ 3 ด้วยวิธี RFA วิธี MH วิธี IRT โดยแต่ละเงื่อนไขจะวิเคราะห์ซ้ำทั้งหมด 10 รอบ (10 replications) ขั้นตอนที่ 5 คำนวณอัตราความคลาดเคลื่อนประเภทที่ 1 และคำนวณหาค่าอำนาจการตรวจสอบ (Power of the test) ของทั้ง 3 วิธีในทุกเงื่อนไข ขั้นตอนที่ 6 เป็นการสรุปผลการวิจัยทั้งหมด รายละเอียดในแต่ละขั้นตอนมีดังนี้

1. การจำลองข้อมูล

การศึกษาครั้งนี้ผู้วิจัยยึดสมการที่ใช้ในการจำลองข้อมูลจากการศึกษาของ Oort (Oort, 1998) เป็นสมการในการจำลองข้อมูล ซึ่งการศึกษาครั้งนี้ผู้วิจัยจำลองข้อมูลด้วยโปรแกรม SPSS for Windows โดยจำลองการให้คะแนนเป็นแบบ 2 ค่า (dichotomous score) ตอบถูกได้ 1 ตอบผิดได้ 0

โดย ผลการตอบข้อสอบ (X_{ij}) ของผู้สอบคนที่ j ในข้อสอบข้อที่ i จะตอบถูกหรือได้ 1 เมื่อ

$$U_{ij} > [1 - \{\exp(a_i(t_i + (m_i / a_i)))/(1 + (\exp(a_i(t_i + (m_i / a_i)))))\}] \quad \text{..... (35)}$$

ในทางตรงกันข้ามผลการตอบข้อสอบ (X_{ij}) ของผู้สอบคนที่ j ในข้อสอบข้อที่ i จะตอบผิดหรือได้ 0 เมื่อ

$$U_{ij} < [1 - \{\exp(a_i(t_i + (m_i / a_i)))/(1 + (\exp(a_i(t_i + (m_i / a_i)))))\}] \quad \text{.....(36)}$$

เมื่อ a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

t_j คือ ค่าความสามารถของผู้สอบคนที่ j

$m_i = -b_i a_i$

b_i คือ ค่าความยากของข้อสอบข้อที่ i

U_{ij} ต้องมีการกระจายอยู่ระหว่าง 0 ถึง 1 ซึ่งแปลงมาจากค่า e_{ij} ที่มีการกระจายเป็นโค้งปกติ ค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 (Oort, 1998)

สมการ (35) และสมการ (36) มีที่มาดังนี้

Zelen และ Severo (1976, Equation 26.2.18 cited in Oort, 1998:85) กล่าวว่า ผลการตอบข้อสอบ (X_{ij}) ของผู้สอบคนที่ j ในข้อสอบข้อที่ i จะตอบถูกหรือได้ 1 เมื่อ

$$\text{ถ้า} \quad U_{ij} > \left|1 - \Pr ob\left(X_{ij} = 1 \middle| T = t_j\right)\right| \quad \text{.....(37)}$$

ในทางตรงกันข้ามผลการตอบข้อสอบ (X_{ij}) ของผู้สอบคนที่ j ในข้อสอบข้อที่ i จะตอบผิดหรือได้ 0 เมื่อ

$$U_{ij} < \left|1 - \Pr ob\left(X_{ij} = 1 \middle| T = t_j\right)\right| \quad \text{.....(38)}$$

เมื่อ

$$\Pr ob\left(X_{ij} = 1 \middle| T = t_j\right) = \Psi\left(\delta_T Y_{ij}\right) \quad \text{.....(39)}$$

โดยที่ $\delta_T Y_{ij}$ เป็นค่าความคาดหวังของคะแนนของคนที่ j ที่ทำข้อสอบข้อที่ i

$$\delta_T Y_{ij} = \delta\left(Y_{ij} \middle| T = t_j\right) = m_i + c_i t_i \quad \text{.....(40)}$$

และ Ψ เป็นฟังก์ชันโลจิสติก (logistic function)

$$\Psi(x) = \exp(x)/(1 + \exp(x)) \quad \text{.....(41)}$$

เมื่อกำหนดให้ a_i คือ c_i และ $b_i = \left(-m_i / c_i\right)$ (Oort, 1996:85)

$$a_i = c_i \quad \text{.....(42)}$$

$$b_i = (-m_i / c_i) \quad \text{.....(43)}$$

หรือ

$$b_i = (-m_i / a_i) \text{.....(44)}$$

แทนค่าตัวแปรในสมการ (42) และ สมการ (43) ลงในสมการ (40) จะได้สมการที่เปลี่ยนไปคือ

$$\delta_T Y_{ij} = \delta(Y_{ij} | T = t_i) = a_i(t_i - b_i) \text{..... (45)}$$

นำสมการที่ (45) ไปแทนค่าในสมการที่ (39) จะได้สมการเป็น

$$\text{Prob}(X_{ij} = 1 | T = t_j) = \Psi(a_i(t_i - b_i)) \quad \text{.....(46)}$$

เมื่อ Ψ เป็นฟังก์ชันโลจิสติก (logistic function) สมการที่ได้คือ

$$\Psi(a_i(t_i - b_i)) = \exp(a_i(t_i - b_i))/(1 + \exp(a_i(t_i - b_i))) \text{.....(47)}$$

แทนค่าสมการที่ (44) ลงในสมการที่ (47) จะได้สมการ

$$\Psi(a_i(t_i - (-m_i / a_i))) = \exp(a_i(t_i - (-m_i / a_i)))/(1 + \exp(a_i(t_i - (-m_i / a_i)))) \text{.....(48)}$$

จากความสัมพันธ์ของสมการ (39) จะได้สมการใหม่เป็น

$$\text{Prob}(X_{ij} = 1 | T = t_j) = \exp(a_i(t_i - (-m_i / a_i)))/(1 + \exp(a_i(t_i - (-m_i / a_i)))) \text{.....(49)}$$

นำสมการ (49) แทนค่าในสมการ (37) จะได้สมการ (35) และถ้านำสมการที่ (49) แทนค่าในสมการ (38) จะได้สมการ (36) ซึ่งเป็นสมการที่ใช้ในการจำลองข้อมูลผลการตอบข้อสอบ (X_{ij}) ของกลุ่มตัวอย่างทั้งสองกลุ่มที่ใช้ในการศึกษาครั้งนี้ด้วยโปรแกรม SPSS for Windows

จากสมการ (35) และสมการ (36) จะเห็นได้ว่า ค่าพารามิเตอร์ของข้อสอบที่ใช้ในสมการดังกล่าวมีทั้งหมด 4 ตัว คือ a_i, m_i, t_j, และ U_{ij} การจำลองค่าพารามิเตอร์เหล่านี้เพื่อให้ได้ผล

การตอบข้อสอบ (X_{ij}) ผู้วิจัยได้ดำเนินการดังนี้ ในการจำลองค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a_i) และค่าพารามิเตอร์ความยากของข้อสอบ (b_i) ผู้วิจัยยึดผลการศึกษาของ คณิต ไข่มุกด์ (2534) เป็นเกณฑ์ในการแบ่งกลุ่มค่าอำนาจจำแนกของข้อสอบ (a_i) และค่าความยากของข้อสอบ (b_i) เพื่อให้ได้ลักษณะของแบบสอบแต่ละฉบับตรงตามเงื่อนไขที่ต้องการศึกษา ซึ่งมีรายละเอียดดังนี้

1. ค่าอำนาจจำแนกของข้อสอบ (a_i)

1.1 กลุ่มข้อสอบที่มีค่าอำนาจจำแนกสูง กำหนดค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.675 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.32 ซึ่งค่าพิสัยที่เป็นผลลัพธ์ของค่าอำนาจจำแนกของข้อสอบที่จำลองได้อยู่ระหว่าง 0.85 ถึง 2.50

1.2 กลุ่มข้อสอบที่มีค่าอำนาจจำแนกปานกลาง กำหนดค่าอำนาจจำแนกเฉลี่ยเท่ากับ 0.7 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.055 ซึ่งค่าพิสัยที่เป็นผลลัพธ์ของค่าอำนาจจำแนกของข้อสอบที่จำลองได้อยู่ระหว่าง 0.55 ถึง 0.85

1.3 กลุ่มข้อสอบที่มีค่าอำนาจจำแนกต่ำ กำหนดค่าอำนาจจำแนกเฉลี่ยเท่ากับ 0.525 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.006 ซึ่งค่าพิสัยที่เป็นผลลัพธ์ของค่าอำนาจจำแนกของข้อสอบที่จำลองได้อยู่ระหว่าง 0.50 ถึง 0.55

2. ค่าความยากของข้อสอบ (b_i)

2.1 กลุ่มข้อสอบที่มีความยากสูง กำหนดค่าความยากเฉลี่ยเท่ากับ 1.835 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.29 ซึ่งค่าพิสัยที่เป็นผลลัพธ์ของค่าอำนาจจำแนกของข้อสอบที่จำลองได้อยู่ระหว่าง 0.67 ถึง 2.50

2.2 กลุ่มข้อสอบที่มีความยากปานกลาง กำหนดค่าความยากเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.32 ซึ่งค่าพิสัยที่เป็นผลลัพธ์ของค่าอำนาจจำแนกของข้อสอบที่จำลองได้อยู่ระหว่าง -0.67 ถึง 0.67

2.3 กลุ่มข้อสอบที่มีความยากต่ำ กำหนดค่าความยากเฉลี่ยเท่ากับ -1.835 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.33 ซึ่งค่าพิสัยที่เป็นผลลัพธ์ของค่าอำนาจจำแนกของข้อสอบที่จำลองได้อยู่ระหว่าง –2.50 ถึง –0.67

รายละเอียดของค่าอำนาจจำแนกเฉลี่ยและค่าความยากเฉลี่ยของแบบสอบที่ศึกษาในแต่ละเงื่อนไขแสดงได้ในตารางที่ 4

ตารางที่ 4 ค่าอำนาจจำแนกและค่าความยากของแบบสอบที่ศึกษาในแต่ละเงื่อนไข

เงื่อนไขที่	ขนาดกลุ่มสอบ	ลักษณะแบบสอบที่ศึกษา			
		ความยาวของแบบสอบ	ขนาดความคล้ายเสียง	ค่าความยาก	ค่าอำนาจจำแนก
1	150	25	ข้อสอบที่ไม่มีความคล้ายเสียง 15 ข้อ (ข้อ 1-15) ข้อสอบที่มีความคล้ายเสียงต่ำ 5 ข้อ (ข้อ16-20) ข้อสอบที่มีความคล้ายเสียงสูง 5 ข้อ (ข้อ 21-25)	สูง ($0.67 < b_i < 2.5$, $\bar{b} = 1.835$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
2					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
3					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
4				ปานกลาง ($-0.67 \le b_i \le 0.67$, $\bar{b} = 0$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
5					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
6					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
7				ต่ำ ($-2.50 < b_i < -.67$, $\bar{b} = -1.835$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
8					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
9					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
10	150	75	ข้อสอบที่ไม่มีความคล้ายเสียง 60 ข้อ (ข้อ 1-60) ข้อสอบที่มีความคล้ายเสียงต่ำ 5 ข้อ (ข้อ 61-65) และข้อสอบที่มีความคล้ายเสียงสูง 10 ข้อ (ข้อ 66-75)	สูง ($0.67 < b_i < 2.5$, $\bar{b} = 1.835$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
11					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
12					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
13				ปานกลาง ($-0.67 \le b_i \le 0.67$, $\bar{b} = 0$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
14					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
15					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
16				ต่ำ ($-2.50 < b_i < -.67$, $\bar{b} = -1.835$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
17					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
18					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
19	500	25	ข้อสอบที่ไม่มีความคล้ายเสียง 15 ข้อ (ข้อ 1-15) ข้อสอบที่มีความคล้ายเสียงต่ำ 5 ข้อ (ข้อ16-20) ข้อสอบที่มีความคล้ายเสียงสูง 5 ข้อ (ข้อ 21-25)	สูง ($0.67 < b_i < 2.5$, $\bar{b} = 1.835$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
20					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
21					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
22				ปานกลาง ($-0.67 \le b_i \le 0.67$, $\bar{b} = 0$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
23					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
24					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
25				ต่ำ ($-2.50 < b_i < -.67$, $\bar{b} = -1.835$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
26					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
27					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)

ตารางที่ 4 (ต่อ)

เงื่อนไขที่	ขนาดผู้สอบ	ลักษณะแบบสอบที่ศึกษา			
		ความยาวของแบบสอบ	ขนาดความลำเอียง	ค่าความยาก	ค่าอำนาจจำแนก
28	500	75	ข้อสอบที่ไม่มีความลำเอียง 60 ข้อ (ข้อ 1-60) ข้อสอบที่มีความลำเอียงต่ำ 5 ข้อ (ข้อ 61-65) และ ข้อสอบที่มีความลำเอียงสูง 10 ข้อ (ข้อ 66-75)	สูง ($0.67 < b_i < 2.5$, $\bar{b} = 1.835$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
29					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
30					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
31				ปานกลาง ($-0.67 \le b_i \le 0.67$, $\bar{b} = 0$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
32					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
33					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)
34				ต่ำ ($-2.50 < b_i < -.67$, $\bar{b} = -1.835$)	สูง ($0.85 \le a_i \le 2.5$, $\bar{a} = 1.675$)
35					ปานกลาง ($0.55 < a_i < 0.85$, $\bar{a} = 0.7$)
39					ต่ำ ($0.50 \le a_i \le 0.55$, $\bar{a} = 0.525$)

รายละเอียดขั้นตอนการจำลองค่าพารามิเตอร์แต่ละตัวมีดังนี้

1. การจำลองค่าอำนาจจำแนกของข้อสอบ (a_i) เนื่องจากการศึกษาครั้งนี้แบ่งค่าอำนาจจำแนกของข้อสอบออกเป็น 3 ระดับ คือ อำนาจจำแนกสูง ปานกลาง และต่ำ ในแต่ละระดับจะกำหนดพิสัยการกระจายของค่าอำนาจจำแนกของข้อสอบไว้ ฉะนั้นก่อนจะจำลองค่าอำนาจจำแนกของข้อสอบในแต่ละระดับ ผู้วิจัยต้องนำค่าอำนาจจำแนกของข้อสอบในแต่ละระดับมาหาค่าเฉลี่ยผลลัพธ์ได้ดังตารางที่ 4 แล้วนำค่าเฉลี่ยดังกล่าวไปจำลองค่าอำนาจจำแนกของข้อสอบในแต่ละเงื่อนไข ตัวอย่างเช่น ในเงื่อนไขที่ 2 ค่าเฉลี่ยอำนาจจำแนกปานกลางเท่ากับ 0.70 ผู้วิจัยนำค่า 0.7 นี้ไปจำลองค่าอำนาจจำแนกของเงื่อนไขที่ 2 ได้ดังนี้

1.1 เข้าสู่โปรแกรม SPSS for Windows

1.2 เติมตัวเลขในแต่ละแถว (low) ของคอลัมน์ (column) ที่ 1 ให้เท่ากับจำนวนข้อสอบที่ต้องการศึกษา

1.3 ใช้คำสั่ง Transform

1.4 เลือกคำสั่ง Compute ซึ่งเป็นคำสั่งย่อยในคำสั่ง Transform

1.5 ตั้งชื่อตัวแปรใหม่ในช่อง Target Variable เช่น การจำลองค่าอำนาจจำแนกปานกลาง ผู้วิจัยกำหนดชื่อตัวแปรใหม่เป็น *a* medium

1.6 เลือกฟังก์ชั่น RV.NORMAL(mean,stddev) ผู้วิจัยต้องการค่าอำนาจจำแนก (a_i) ในเงื่อนไขที่ 2 ที่มีการแจกแจงปกติ โดยมีค่าเฉลี่ยเท่ากับ 0.7 ค่าส่วนเบี่ยงเบน

มาตรฐานเท่ากับ 0.055 ฟังชั่นจะเปลี่ยนเป็น RV.NORMAL(0.7,0.055) แล้วคลิก OK ตัวอย่างค่าอำนาจจำแนก (a_i) ที่จำลองได้แสดงในภาคผนวก ก

1.7 จัดเก็บแฟ้มข้อมูลค่า (a_i)

2. การจำลองค่าความยากของข้อสอบ (b_i) เนื่องจากการศึกษาครั้งนี้แบ่งค่าความยากของข้อสอบออกเป็น 3 ระดับ คือ ความยากสูง ปานกลาง และต่ำ ในแต่ละระดับจะกำหนดพิสัยการกระจายของค่าความยากไว้ ฉะนั้นก่อนจะจำลองค่าความยากของข้อสอบในแต่ละระดับ ผู้วิจัยต้องนำค่าความยากของข้อสอบในแต่ละระดับมาหาค่าเฉลี่ย ผลลัพธ์ได้ดังตารางที่ 4 แล้วนำค่าเฉลี่ยดังกล่าวไปจำลองค่าความยากของข้อสอบในแต่ละเงื่อนไข ขั้นตอนการจำลองค่าความยากของข้อสอบปฏิบัติเช่นเดียวกับการจำลองค่าอำนาจจำแนกของข้อสอบ ตัวอย่างค่าความยากของข้อสอบ (b_i) ที่จำลองขึ้นแสดงในภาคผนวก ข

3. การจำลองค่าพารามิเตอร์ m_i ($m_i = -b_i a_i$ เมื่อ a_i คือค่าอำนาจจำแนกของข้อสอบ และ b_i คือค่าความยากของข้อสอบ)

3.1 เข้าสู่โปรแกรม SPSS for Windows

3.2 รวมแฟ้มข้อมูลค่าพารามิเตอร์ a_i และค่าพารามิเตอร์ b_i โดยใช้เมนูบาร์ Data ใช้คำสั่ง merge files ซึ่งเป็นเมนูย่อยในเมนูบาร์ Data เพื่อรวมแฟ้มข้อมูลค่าพารามิเตอร์ที่ใช้ทั้งหมด

3.3 เลือกคำสั่ง Transform

3.4 เลือกคำสั่ง Compute ซึ่งเป็นคำสั่งย่อยในคำสั่ง Transform

3.5 ตั้งชื่อตัวแปรใหม่ เช่น ในการจำลองผู้วิจัยกำหนดให้เป็นค่า (m_1) ในช่อง Target Variable

3.6 เขียนสมการ $m_1 = -b_i a_i$ ลงในช่องคำสั่ง numeric expression ซึ่งเป็นคำสั่งที่กำหนดผลลัพธ์ให้กับ Target Variable

3.7 คลิก OK ปฏิบัติตามข้อ 3.6, 3.7 ไปจนกระทั่งได้ค่า m_i เท่ากับจำนวนข้อสอบในแต่ละขนาดที่ต้องการศึกษา ตัวอย่างค่า m_i ที่จำลองขึ้นแสดงในภาคผนวก ค

3.8 จัดเก็บแฟ้มข้อมูลค่า (m_i)

4. การจำลองค่า t_j เนื่องจากค่า t_j เป็นค่าความสามารถของผู้สอบแต่ละคนที่ตอบข้อสอบในเงื่อนไขที่ศึกษา มีการกระจายเป็นโค้งปกติที่มีค่าเฉลี่ยเท่ากับ 0 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ผู้วิจัยดำเนินการจำลองค่า t_j ดังนี้

4.1 เข้าสู่โปรแกรม SPSS for Windows

4.2 เติมตัวเลขในแต่ละแถว (low) ของคอลัมน์ (column) ที่ 1 ให้เท่ากับจำนวนกลุ่มตัวอย่างที่ต้องการศึกษา

4.3 ใช้คำสั่ง Transform

4.4 เลือกคำสั่ง Compute ซึ่งเป็นคำสั่งย่อยในคำสั่ง Transform

4.5 ตั้งชื่อตัวแปรใหม่ ในการจำลองผู้วิจัยกำหนดให้เป็นค่า (t) ในช่อง Target variable

4.6 เลือกฟังก์ชั่น RV.NORMAL(mean,stddev) ผู้วิจัยต้องการได้ (t_j) ที่มีการแจกแจงเป็นโค้งปกติ ที่มีค่าเฉลี่ย เท่ากับ 0 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ฟังก์ชั่นก็จะเปลี่ยนเป็น RV.NORMAL(0,1) แล้วคลิก OK (ตัวอย่างค่า t_j ที่จำลองขึ้นแสดงในภาคผนวก ง)

4.7 จัดเก็บแฟ้มข้อมูลค่า (t_j)

5. การจำลองค่าพารามิเตอร์ U_g เนื่องจากการกระจายของ U_g ต้องมีพิสัยการกระจายอยู่ระหว่าง 0 ถึง 1 ซึ่งค่า U_g นี้แปลงมาจากค่า e_g ที่มีการกระจายเป็นโค้งปกติ มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 (Oort, 1998) ผู้วิจัยดำเนินการจำลองดังนี้

5.1 เข้าสู่โปรแกรม SPSS for Windows

5.2 เติมตัวเลขในแต่ละแถว (low) ของคอลัมน์ (column) ที่ 1 ให้เท่ากับจำนวนกลุ่มตัวอย่างที่ต้องการศึกษา

5.3 ใช้คำสั่ง Transform

5.4 เลือกคำสั่ง Compute ซึ่งเป็นคำสั่งย่อยในคำสั่ง Transform

5.5 ตั้งชื่อตัวแปรใหม่ ในการจำลองผู้วิจัยกำหนดให้เป็นค่า (e_1) ไปเรื่อย ๆ จนถึง (e_{10}) ในช่อง Target Variable เท่ากับจำนวนข้อสอบที่จะศึกษา

5.6 เลือกฟังก์ชั่น RV.NORMAL(mean,stddev) โดยค่า mean และ ค่า stddev คือค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของข้อมูลที่ต้องการจำลองนั่นเอง ในการจำลองค่า (e_g) ผู้วิจัยต้องการได้ (e_j) ที่มีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ 0 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ฟังก์ชั่นที่ใช้จะเปลี่ยนเป็น RV.NORMAL(0,1) แล้วคลิก OK

5.7 ปฏิบัติตามข้อ 5.6 ไปจนกระทั่งได้ค่า (e_j) เท่ากับจำนวนข้อสอบในแต่ละขนาดที่ต้องการศึกษา

5.8 นำค่า (e_{ij}) ที่มีค่าสูงที่สุด (สมมุติว่าการจำลองครั้งนี้ค่า (e_{ij}) สูงที่สุดคือ 3.98) ไปหารค่า (e_{ij}) ที่จำลองขึ้นทั้งหมด เพื่อให้ได้ค่า U_{ij} ตามต้องการโดยระบุชื่อตัวแปรใหม่ให้เป็น U_1 ในช่อง Target Variable

5.9 เลือกฟังก์ชั่น ABS(numexpr) เติมฟังชั่นของตัวแปรในวงเล็บ เช่น ผู้วิจัยต้องการแปลงค่า (e_{ij}) ให้เป็นค่า U_1 ก็จะได้เป็น ABS(e_{ij} / 3.98) ผลลัพธ์ที่ได้ก็คือ U_1 นั่นเอง ปฏิบัติเช่นนี้จนกระทั่งได้ค่า U_{ij} เท่ากับจำนวนข้อสอบในแต่ละขนาดที่ต้องการศึกษา ตัวอย่างข้อมูลค่า U_{ij} ที่จำลองขึ้นแสดงในภาคผนวก จ

5.10 จัดเก็บแฟ้มข้อมูลค่า U_{ij}

6. การจำลองข้อสอบให้มีความลำเอียง จากการศึกษาของ Oort พบว่าการกำหนดให้ข้อสอบมีความลำเอียงสูงหรือต่ำนั้น สามารถกำหนดได้โดยนำขนาดของความลำเอียงไปลบออกจากค่า m_i ($m_i = -b_i a_i$) เฉพาะข้อที่ผู้วิจัยต้องการกำหนดให้เป็นข้อสอบมีความลำเอียง (Oort, 1998) การศึกษาครั้งนี้ในแบบสอบสั้นมีข้อสอบ 25 ข้อ แบบสอบชุดนี้ผู้วิจัยกำหนดให้มีข้อสอบที่ไม่ลำเอียง 15 ข้อ และข้อสอบที่มีความลำเอียงทั้งหมด 10 ข้อ เป็นข้อสอบที่มีความลำเอียงสูง 5 ข้อ ข้อสอบมีความลำเอียงต่ำ 5 ข้อ โดยผู้วิจัยกำหนดให้ข้อสอบข้อที่ 1-15 เป็นข้อสอบที่ไม่มีความลำเอียง ข้อสอบข้อที่ 16-20 เป็นข้อสอบที่มีความลำเอียงต่ำ (0.3SD) และข้อสอบข้อที่ 21-25 เป็นข้อสอบที่มีความลำเอียงสูง (0.8SD) เฉพาะกรณีที่ข้อสอบมีความลำเอียงต่ำ (ข้อที่ 16-20) ผู้วิจัยนำค่า 0.3 ไปลบออกจากค่า m_i ของข้อสอบข้อที่ 16 (m_{16}) ถึงข้อที่ 20 (m_{20}) จากนั้นนำค่า m_i ใหม่นี้ไปแทนค่าในสมการ (35) เพื่อจำลองเป็นผลการตอบข้อสอบข้อที่ 16 ถึงข้อที่ 20 ของกลุ่มอ้างอิง ส่วนค่า m_i เดิมจะนำไปแทนค่าในสมการ (35) เพื่อจำลองเป็นผลการตอบข้อสอบข้อที่ 16 ถึงข้อที่ 20 ของกลุ่มเปรียบเทียบ ในทำนองเดียวกัน ข้อสอบข้อที่ 21-25 ที่ถูกกำหนดให้เป็นข้อสอบที่มีความลำเอียงสูง (0.8 SD) ก็ปฏิบัติทำนองเดียวกันคือ ผู้วิจัยนำค่า 0.8 ไปลบออกจากค่า m_i ของข้อสอบข้อที่ 21 (m_{21}) ถึงข้อที่ 25 (m_{25}) จากนั้นนำค่า m_i ใหม่นี้ไปจำลองเป็นผลการตอบข้อสอบข้อที่ 21 ถึงข้อที่ 25 ของกลุ่มอ้างอิง ส่วนค่า m_i เดิมจะนำไปจำลองเป็นผลการตอบข้อสอบข้อที่ 21 ถึงข้อที่ 25 ของกลุ่มเปรียบเทียบ ส่วนขั้นตอนการจำลองข้อสอบให้มีความลำเอียงในแบบสอบยาว (75 ข้อ) มีวิธีการการจำลองเหมือนกับการจำลองข้อมูลในแบบสอบสั้น (25 ข้อ) จะต่างกันเฉพาะข้อสอบข้อที่ถูกกำหนดให้เป็นข้อสอบที่มีความลำเอียงนั้นจะต่างกัน กล่าวคือ ผู้วิจัยจะกำหนดให้แบบสอบยาว 75 ข้อ มีข้อสอบที่ไม่ลำเอียง 60 ข้อ และข้อสอบที่ลำเอียงทั้งหมด 15 ข้อ เป็นข้อสอบที่ลำเอียงสูง 10 ข้อ ลำเอียงต่ำ5 ข้อ โดยผู้วิจัยกำหนดให้ข้อสอบข้อที่ 1-60 เป็นข้อสอบที่ไม่มีความลำเอียง ข้อสอบข้อที่ 61-65 เป็นข้อสอบที่ลำเอียงต่ำ (0.3SD) และข้อสอบข้อที่ 66-75 เป็นข้อสอบที่ลำเอียงสูง (0.8SD)

7. การจำลองข้อมูลผลการตอบข้อสอบ (X_{ij}) ในแต่ละเงื่อนไข เมื่อผู้วิจัยจำลองค่าพารามิเตอร์ที่ใช้ในสมการที่ (35) ครบหมดทุกตัว จากนั้นนำค่าพารามิเตอร์ที่จำลองขึ้นไปแทนค่าลงในสมการที่ (35) ขั้นตอนการจำลองผลการตอบข้อสอบ (X_{ij}) มีดังนี้

7.1 เข้าสู่โปรแกรม SPSS for Windows

7.2 รวมแฟ้มข้อมูลค่าพารามิเตอร์ที่ใช้แทนค่าในสมการที่ 35 โดยใช้เมนูบาร์ Data ใช้คำสั่ง merge files ซึ่งเป็นเมนูย่อยในเมนูบาร์ Dataเพื่อรวมแฟ้มข้อมูลค่าพารามิเตอร์ที่ใช้ทั้งหมด

7.3 เลือกคำสั่ง Transform

7.4 เลือกคำสั่ง Compute ซึ่งเป็นคำสั่งย่อยในคำสั่ง Transform

7.5 ตั้งชื่อตัวแปรใหม่ในช่อง Target variable เช่น การจำลองผลการตอบข้อสอบข้อที่ 1 ผู้วิจัยตั้งชื่อตัวแปรเป็น X_1 ในช่อง Target Variable

7.6 เขียนสมการที่ (35) ลงในช่องคำสั่ง numeric expression ซึ่งเป็นคำสั่งที่กำหนดผลลัพธ์ให้กับ Target Variable

7.7 แล้วคลิก OK

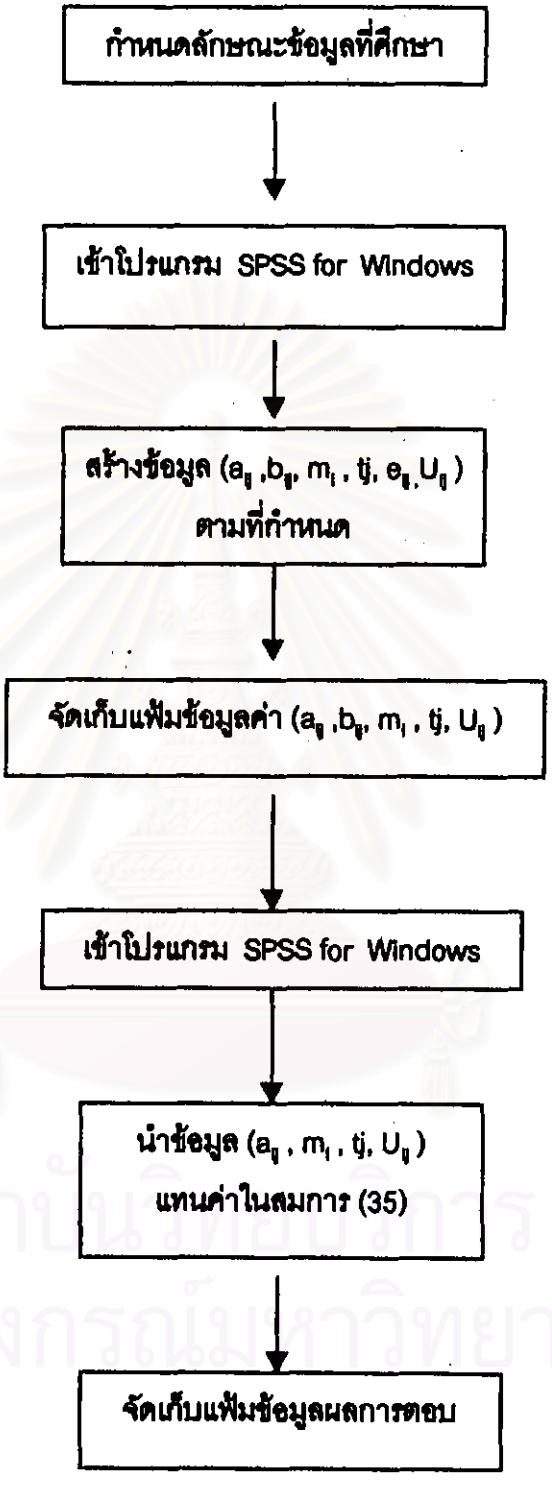
7.8 ปฏิบัติตามข้อ 7.5, 7.6, 7.7 ไปเรื่อย ๆ จนได้ข้อมูลครบตามต้องการ โดยเปลี่ยนค่า i จาก 1 ไปจนถึง 25 ในแบบสอบ 25 ข้อ และจาก 1 ไปจนถึง 75 ในแบบสอบ 75 ข้อ ตัวอย่างค่า X_{ij} ที่จำลองขึ้นแสดงในภาคผนวก ฉ

7.9 จัดเก็บแฟ้มข้อมูลค่า X_{ij}

ขั้นตอนในการจำลองข้อมูลทั้งหมดสามารถเขียนเป็นแผนภาพการจำลองข้อมูลตั้งแต่ขั้นตอนแรกจนกระทั่งได้ผลการตอบข้อสอบ (X_{ij}) ของกลุ่มตัวอย่างแต่ละกลุ่มที่ใช้ในการศึกษาครั้งนี้ดังแผนภาพที่ 2

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แผนภาพที่ 2_ ขั้นตอนการจำลองข้อมูลด้วยโปรแกรม SPSS for Windows



2. การสุ่มข้อสอบ

เป็นการเตรียมข้อมูลให้มีลักษณะตามเงื่อนไขที่กำหนดด้วยวิธีสุ่มอย่างง่ายแบบไม่ใส่กลับคืนจากข้อมูลที่ได้จำลองไว้ ในขั้นตอนนี้เป็นการสุ่มข้อสอบเพื่อให้ได้แบบสอบที่มีความยาวตามที่ต้องการศึกษาโดยสุ่มข้อสอบให้มีจำนวน 25 ข้อและ 75 ข้อ ตามเงื่อนไขของค่าความยากของข้อสอบ (b) ค่าอำนาจจำแนกของข้อสอบ (a) และความลำเอียงของข้อสอบตามที่กำหนดขึ้น

3. การสุ่มกลุ่มตัวอย่าง

การสุ่มกลุ่มตัวอย่าง เป็นการสุ่มอย่างง่ายแบบไม่ใส่กลับคืน ข้อมูลที่จำลองขึ้นมีจำนวนกลุ่มตัวอย่างระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบในอัตราส่วนที่เท่ากัน คือ 1:1 ทุกๆ ขนาดของกลุ่มตัวอย่าง โดยจำนวนกลุ่มตัวอย่างที่ศึกษาในครั้งนี้มีทั้งหมด 2 ขนาด คือ 300 และ 1,000 คน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5 ลักษณะแบบสอบที่ศึกษาในแต่ละวิธีตรวจสอบของกลุ่มอ้างอิง

วิธีตรวจสอบ	กลุ่มที่ศึกษา	ขนาดผู้สอบ	ลักษณะแบบสอบที่ศึกษา				
			ความยาวแบบสอบ	ขนาดความลำเอียง	ค่าความยาก	ค่าอำนาจจำแนก	replication
IRT, RFA, MH	กลุ่มอ้างอิง	150	25	ข้อสอบที่ไม่มีความลำเอียง 15 ข้อ (ข้อ 1-15)ข้อสอบที่มีความลำเอียงต่ำ 5 ข้อ (ข้อ16-20) ข้อสอบที่มีความลำเอียงสูง 5 ข้อ (ข้อ 21-25)	สูง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ปานกลาง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ต่ำ	สูง	10
						ปานกลาง	10
						ต่ำ	10
			75	ข้อสอบที่ไม่มีความลำเอียง 60 ข้อ (ข้อ 1-60) ข้อสอบที่มีความลำเอียงต่ำ 5 ข้อ (ข้อ 61-65) และข้อสอบที่มีความลำเอียงสูง 10 ข้อ (ข้อ 66-75)	สูง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ปานกลาง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ต่ำ	สูง	10
						ปานกลาง	10
						ต่ำ	10
		500	25	ข้อสอบที่ไม่มีความลำเอียง 15 ข้อ (ข้อ 1-15)ข้อสอบที่มีความลำเอียงต่ำ 5 ข้อ (ข้อ16-20) ข้อสอบที่มีความลำเอียงสูง 5 ข้อ (ข้อ 21-25)	สูง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ปานกลาง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ต่ำ	สูง	10
						ปานกลาง	10
						ต่ำ	10
			75	ข้อสอบที่ไม่มีความลำเอียง 60 ข้อ (ข้อ 1-60) ข้อสอบที่มีความลำเอียงต่ำ 5 ข้อ (ข้อ 61-65) และข้อสอบที่มีความลำเอียงสูง 10 ข้อ (ข้อ 66-75)	สูง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ปานกลาง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ต่ำ	สูง	10
						ปานกลาง	10
						ต่ำ	10

ตารางที่ 6 ลักษณะแบบสอบที่ศึกษาในแต่ละวิธีตรวจสอบของกลุ่มเปรียบเทียบ

วิธีตรวจสอบ	กลุ่มที่ศึกษา	ขนาดผู้สอบ	ลักษณะแบบสอบที่ศึกษา				
			ความยาวแบบสอบ	ขนาดความคลาดเคลื่อน	ค่าความยาก	ค่าอำนาจจำแนก	replication
IRT, RFA, MH	กลุ่มเปรียบเทียบ	150	25 ข้อ	-	สูง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ปานกลาง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ต่ำ	สูง	10
						ปานกลาง	10
						ต่ำ	10
			75 ข้อ	-	สูง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ปานกลาง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ต่ำ	สูง	10
						ปานกลาง	10
						ต่ำ	10
		500	25 ข้อ	-	สูง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ปานกลาง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ต่ำ	สูง	10
						ปานกลาง	10
						ต่ำ	10
			75 ข้อ	-	สูง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ปานกลาง	สูง	10
						ปานกลาง	10
						ต่ำ	10
					ต่ำ	สูง	10
						ปานกลาง	10
						ต่ำ	10

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4. การวิเคราะห์ด้วยวิธีการวิเคราะห์องค์ประกอบจำกัด (RFA)

การวิเคราะห์ข้อสอบทำหน้าที่ต่างกันด้วยวิธี RFA ผู้วิจัยใช้โปรแกรม LISREL มีขั้นตอนการวิเคราะห์ 2 ขั้นตอนคือ 1) เตรียมข้อมูลสำหรับการวิเคราะห์องค์ประกอบจำกัด ให้อยู่ในรูปเมตริกซ์สหสัมพันธ์โดยใช้โปรแกรมย่อย PRELIS ซึ่งเป็นโปรแกรมย่อยที่อยู่ในโปรแกรม LISREL ข้อมูลนำเข้าในโปรแกรม PRELIS คือ ผลการตอบข้อสอบ (X_{ij}) ของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบซึ่งจัดเก็บอยู่ในแฟ้มเดียวกัน ในการจัดเก็บผลการตอบของผู้สอบทั้งสองกลุ่มนี้ต้องมีรหัส (code) บอกว่าผลการตอบข้อสอบเป็นของผู้สอบกลุ่มอ้างอิงหรือกลุ่มเปรียบเทียบ ผลลัพธ์ที่ได้จากการวิเคราะห์คือ เมตริกซ์สหสัมพันธ์ 2) นำเมตริกซ์สหสัมพันธ์นี้ไปตรวจสอบหาข้อสอบที่ทำหน้าที่ต่างกันด้วยโปรแกรม LISREL เกณฑ์การพิจารณาข้อสอบทำหน้าที่ต่างกันคือ จะพิจารณาที่ค่าดัชนี MI และดัชนี EPC ว่าแตกต่างจาก 0 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 หรือไม่ (Oort ,1988) ตัดข้อสอบที่มีค่า MI และดัชนี EPC ที่ต่างจาก 0 ออกจากแบบสอบชุดนั้นทีละข้อ โดยตัดข้อสอบที่มีค่า MI สูงที่สุดออกก่อน นำข้อสอบที่เหลือมาวิเคราะห์ซ้ำไปเรื่อย ๆ จนกว่าไม่มีข้อสอบที่มีค่า MI และดัชนี EPC ที่แตกต่างจาก 0 เหลืออยู่ในแบบสอบชุดนั้น

5. การวิเคราะห์ด้วยวิธีแมนเทล-แฮนส์เซล (MH)

การวิเคราะห์ข้อสอบทำหน้าที่ต่างกันด้วยวิธี MH นั้นใช้โปรแกรม MH ที่พัฒนาโดย Fidago A. (1995) ในการวิเคราะห์ผู้วิจัยเลือกใช้เทคนิคการวิเคราะห์แบบ 2 ขั้นตอนทั้งนี้เพราะจากงานวิจัยของ Clauser และคณะ (1993) พบว่าผลการตรวจสอบด้วยเทคนิค 2 ขั้นตอน พบข้อสอบทำหน้าที่ต่างกันสูงกว่าเทคนิค 1 ขั้นตอนโดยไม่ทำให้อัตราความคลาดเคลื่อนประเภทที่ 1 สูงขึ้น

ในการวิเคราะห์ด้วยโปรแกรม MH นั้นแฟ้มข้อมูลที่ใช้ป้อนเข้าคือ ผลการตอบข้อสอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่จัดเก็บอยู่ในแฟ้มเดียวกัน ซึ่งในการจัดเก็บผลการตอบของผู้สอบสองกลุ่มต้องมีรหัส (code) บอกว่าผลการตอบข้อสอบเป็นของผู้สอบกลุ่มอ้างอิงหรือกลุ่มเปรียบเทียบ ผลลัพธ์ที่ได้จากการวิเคราะห์คือ ค่าดัชนี α_{MH} และดัชนี MH_{DIF} และทำการทดสอบนัยสำคัญด้วยค่าสถิติไคสแควร์

6. การวิเคราะห์ข้อมูลด้วยวิธีการตอบสนองข้อสอบ (IRT)

การศึกษาการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบครั้งนี้ เป็นการเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธี IRT แบบ 2 พารามิเตอร์ วิธี MH และวิธี RFA ในทุกเงื่อนไขที่ผู้วิจัยกำหนดขึ้น ซึ่งในส่วนของวิธีการตอบสนองข้อสอบ (IRT) แบบ 2 พารามิเตอร์นั้นมีขั้นตอนการวิเคราะห์ดังนี้

เริ่มต้นด้วยการวิเคราะห์องค์ประกอบของแบบสอบด้วยโปรแกรม SPSS for Windows เพื่อตรวจสอบความเป็นเอกมิติ (Unidimentionality) ของแบบสอบ ซึ่งเป็นข้อตกลงเบื้องต้นของการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีการตอบสนองข้อสอบ (IRT) จากนั้นนำแฟ้มข้อมูลผลการตอบข้อสอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบมาคำนวณหาค่าความแปรปรวนร่วมของข้อมูลด้วยโปรแกรม BILOG 3.04 โดยแยกแฟ้มข้อมูลวิเคราะห์ระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ นำแฟ้มข้อมูลที่ได้จากการประมาณค่าจากโปรแกรม BILOG 3.04 มาเทียบมาตรของค่าพารามิเตอร์ของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบโดยโปรแกรม EQUATE 2.0 การวิเคราะห์จะเลือกใช้แบบ interactive mode ซึ่งมีลักษณะเป็นคำถามทีละคำถามให้ผู้ใช้ตอบเกี่ยวกับความต้องการของผู้ใช้ ผลลัพธ์ที่ได้จากการวิเคราะห์จะได้ ค่า A (slope) และค่า K (intercept) ของสมการเชิงเส้นที่ใช้เทียบมาตรค่าพารามิเตอร์ของกลุ่มอ้างอิงและค่าพารามิเตอร์ของกลุ่มเปรียบเทียบ จากนั้นคำนวณหาพื้นที่ความแตกต่างระหว่างโค้งคุณลักษณะข้อสอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบโดยใช้โปรแกรม AREA แบบ 2 พารามิเตอร์ แล้วทดสอบค่าสถิติ Z ที่ระดับ 0.05

ขั้นตอนการวิเคราะห์มีรายละเอียดดังนี้

6.1 นำแฟ้มข้อมูลผลการตอบข้อสอบของทั้งสองกลุ่มมาคำนวณเมตริกซ์ค่าพารามิเตอร์และค่าความแปรปรวนร่วมของข้อมูลผลการตอบด้วยโปรแกรม BILOG 3.04 ในการวิเคราะห์ด้วยโปแกรมนี้ จะต้องมีแฟ้มข้อมูลนำเข้าคือ 1)แฟ้มข้อมูลซึ่งจะเป็นคำสั่งให้โปรแกรมดำเนินการวิเคราะห์ตามที่ผู้ใช้กำหนด 2) แฟ้มข้อมูลผลการตอบข้อสอบในแต่ละข้อของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่ใช้ในการวิเคราะห์ โดยบรรทัดแรกเป็นคำตอบที่ถูกต้องผู้วิจัยกำหนดให้เป็น 1 และบรรทัดที่สองเป็น omit ผู้วิจัยกำหนดให้เป็นเลข 9 และบรรทัดที่สามเป็นข้อมูลประกอบด้วยเลขประจำตัวและผลการตอบข้อสอบของผู้สอบ ในการเตรียมข้อมูลในแฟ้มดังกล่าวนี้ต้องเว้นช่องว่างระหว่างเลขประจำตัวกับผลการตอบข้อสอบข้อที่ 1 อย่างน้อย 1 คอลัมน์ ผลลัพธ์ที่ได้จากการวิเคราะห์ในแต่ละเงื่อนไขในขั้นตอนนี้ได้แก่ แฟ้มข้อมูลค่าความแปรปรวนร่วมของกลุ่มอ้างอิงและแฟ้มข้อมูลความแปรปรวนร่วมของกลุ่มเปรียบเทียบ

6.2 ในการเปรียบเทียบโค้งลักษณะข้อสอบของผู้สอบสองกลุ่มต้องแปลงมาตรของค่าพารามิเตอร์ที่แสดงลักษณะข้อสอบที่ได้จากผู้สอบ 2 กลุ่มที่แยกกันวิเคราะห์ให้อยู่บนมาตรเดียวกันก่อนนำไปเปรียบเทียบกัน ซึ่งเราสามารถปฏิบัติได้โดยใช้โปรแกรม EQUATE

มาตรของค่าพารามิเตอร์ของข้อสอบของผู้สอบทั้งสองกลุ่มจะเทียบให้มีมาตรเดียวกันได้ด้วยสูตรดังนี้

$$a_i^*2 = (1/A)a_i2 \quad \text{..... (50)}$$

$$b_i^*2 = Ab_i2 + K \quad \text{..... (51)}$$

เมื่อ * คือค่าพารามิเตอร์ที่เทียบแล้ว

i คือ ข้อสอบข้อที่ i

2 คือกลุ่มที่ต้องการจะนำไปเทียบมาตร

A, K คือค่าคงที่ซึ่งจะทำให้ค่าเฉลี่ยและความแปรปรวนของค่า b ที่แปลงแล้วของกลุ่มที่ 2 มีค่าเท่ากับค่าเฉลี่ยและความแปรปรวนของกลุ่มแรกโดย A เป็นค่าความชันและ K เป็นค่า intercept

การวิเคราะห์ด้วยโปรแกรม EQUATE ผู้วิจัยเลือกใช้แบบ intertive mode ซึ่งมีลักษณะเป็นคำถามทีละคำถามให้ผู้ใช้ตอบเกี่ยวกับข้อมูลตามความต้องการของผู้ใช้ ข้อมูลนำเข้าเป็นเมตริกซ์ความแปรปรวนร่วมที่ได้จากการวิเคราะห์ด้วยโปรแกรม BILOG ในการวิเคราะห์แต่ละเงื่อนไขจะใช้เมตริกซ์ 2 เมตริกซ์ คือเมตริกซ์ของกลุ่มอ้างอิงและเมตริกซ์ของกลุ่มเปรียบเทียบซึ่งแยกกันอยู่คนละแฟ้ม ผลลัพธ์ที่ได้จากการวิเคราะห์ด้วยโปรแกรม EQUATE คือค่า A (slope) และค่า K (intercept) ของสมการเชิงเส้นที่ใช้เทียบมาตรค่าพารามิเตอร์กลุ่มเปรียบเทียบให้มีมาตรเดียวกับค่าพารามิเตอร์กลุ่มอ้างอิง

6.3 การคำนวณพื้นที่ความแตกต่างระหว่างโค้งลักษณะข้อสอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ ใช้โปรแกรมการคำนวณพื้นที่ความแตกต่างระหว่างโค้งคุณลักษณะข้อสอบของ Raju แบบ 2 พารามิเตอร์ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

ในการวิเคราะห์ด้วยโปรแกรม AREA ของ Raju แยกขั้นตอนการวิเคราะห์ออกเป็น 2 ขั้นตอนคือ

6.3.1 วิเคราะห์ด้วยโปรแกรมย่อย AREAB ซึ่งจะคำนวณพื้นที่แบบมีเครื่องหมาย (SA) และพื้นที่ไม่มีเครื่องหมาย (UN) ในการวิเคราะห์ด้วยโปรแกรมย่อย AREAB ข้อมูลที่ต้องนำ

เข้าคือ 1) แฟ้มข้อมูลที่มีค่า A และค่า K ที่ได้จากการวิเคราะห์ด้วยโปรแกรม EQUATE 2) แฟ้มข้อมูลความแปรปรวนร่วมของกลุ่มอ้างอิงที่ได้จากการวิเคราะห์ด้วยโปรแกรม BILOG 3) แฟ้มข้อมูลความแปรปรวนร่วมของกลุ่มเปรียบเทียบที่ได้จากโปรแกรม BILOG ซึ่งแฟ้มข้อมูลค่าความแปรปรวนร่วมทั้ง 2 แฟ้มที่นำเข้านี้ไม่ต้องดัดแปลงเมตริกซ์เหมือนเมตริกซ์ที่ใช้นำเข้าในโปรแกรม EQUATE ผลลัพธ์ที่ได้จากโปรแกรมย่อย AREAB นี้จะนำไปเป็นข้อมูลนำเข้าในโปรแกรม AREASEB ต่อไป

6.3.2 วิเคราะห์ด้วยโปรแกรมย่อย AREASEB เพื่อทดสอบนัยสำคัญด้วยสถิติ Z ที่ระดับ 0.05 ขั้นตอนนี้ใช้ข้อมูลนำเข้าคือ 1) ผลลัพธ์ที่ได้จากโปรแกรมย่อย AREAB 2) แฟ้มข้อมูลความแปรปรวนร่วมของกลุ่มอ้างอิงที่ได้จากโปรแกรม BILOG 3) แฟ้มข้อมูลความแปรปรวนร่วมของกลุ่มเปรียบเทียบที่ได้จากโปรแกรม BILOG ผลลัพธ์ที่ได้จากการวิเคราะห์ด้วยโปรแกรมย่อย AREASEB นี้ได้แก่ พื้นที่แบบมีเครื่องหมายและพื้นที่แบบไม่มีเครื่องหมาย ผลการทดสอบนัยสำคัญด้วยสถิติ Z และสถิติ χ^2 ของ LORD

7. คำนวณหาคำนวณหาค่าอำนาจการตรวจสอบ (Power of the test) และอัตราความคลาดเคลื่อนประเภทที่ 1 (Type I error rate) ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแต่ละวิธี

หลังจากวิเคราะห์ด้วยวิธีทั้งสามแล้ว นำผลการวิเคราะห์ที่ได้มาคำนวณหาอัตราความคลาดเคลื่อนประเภทที่ 1 (Type I error rate) อัตราความคลาดเคลื่อนประเภทที่ 2 และค่าอำนาจการตรวจสอบ (Power of the test) ของแต่ละวิธี เพื่อเปรียบเทียบประสิทธิภาพของแต่ละวิธีซึ่งสามารถคำนวณได้ดังนี้

อัตราความคลาดเคลื่อนประเภทที่ 1

$$= \frac{\text{จำนวนข้อสอบที่ระบุผิดพลาดว่าทำหน้าที่ต่างกันทั้งที่ความจริงทำหน้าที่ไม่ต่างกัน}}{\text{จำนวนข้อสอบที่ทำหน้าที่ไม่ต่างกันทั้งหมดในแบบสอบ}}$$

อัตราความคลาดเคลื่อนประเภทที่ 2

$$= \frac{\text{จำนวนข้อสอบที่ระบุผิดพลาดว่าทำหน้าที่ไม่ต่างกันทั้งที่ความจริงข้อสอบทำหน้าที่ต่างกัน}}{\text{จำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้งหมดในแบบสอบ}}$$

อำนาจการตรวจสอบ (Power of tests)

$$= \frac{\text{จำนวนข้อสอบที่ตรวจพบว่าทำหน้าที่ต่างกันได้ถูกต้อง}}{\text{จำนวนข้อสอบที่ทำหน้าที่ต่างกันทั้งหมดในแบบสอบ}}$$

8. เกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธี RFA วิธี MH และวิธี IRT

ผู้วิจัยใช้ข้อสอบที่มีความลำเอียงเป็นเกณฑ์ในการเปรียบเทียบเพื่อคำนวณหาค่าอำนาจการตรวจสอบ และใช้ข้อสอบข้อที่ไม่มีความลำเอียงเป็นเกณฑ์ในการเปรียบเทียบเพื่อคำนวณหาค่าอัตราความคลาดเคลื่อนประเภทที่ 1 ในการตรวจสอบด้วยวิธีการวิเคราะห์องค์ประกอบจำกัด (RFA) วิธีแมนเทล-แฮนส์เซล (MH) และวิธีการตอบสนองข้อสอบ (IRT) ทั้งนี้เพราะผู้วิจัยทราบว่าข้อสอบข้อใดเป็นข้อสอบที่มีความลำเอียงและข้อสอบข้อใดบ้างที่ไม่มีความลำเอียง นอกจากนี้ผู้วิจัยกำหนดเกณฑ์ที่ใช้เปรียบเทียบประสิทธิภาพในการทำหน้าที่ต่างกันของข้อสอบของทั้ง 3 วิธี คือ วิธีใดมีอำนาจการตรวจสอบ (Power of the test) สูงและมีความคลาดเคลื่อนประเภทที่ 1 ต่ำแสดงว่า วิธีนั้นมีประสิทธิภาพสูงสุดในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย