บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

บทนี้ว่าด้วยเอกสาร บทความและงานวิจัยที่เกี่ยวข้องกับการเทียบมาตราซึ่งรายละเอียดต่างๆจะครอบคลุมเกี่ยวกับ นิยามของการเทียบมาตรา รูปแบบการเทียบมาตราต่างๆ ได้แก่ การเทียบมาตรารูปแบบอิควิเปอร์เซ็นไทล์ (Equipercentile Equating) การเทียบมาตรา เชิงเส้นตรง (Linear Equating) การเทียบมาตรารูปแบบการวิเคราะห์ตัวประกอบ (Equating Using The Confirmatory Factor Analysis Model) การเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบ (Item Response Theory Model) การประเมินคุณภาพของการเทียบมาตรา และงานวิจัยที่เกี่ยวกับการเทียบมาตรา

นิยามของการเทียบมาตรา

กัลลิคเสน (Gulliksen, 1950 : 298-304 อ้างจากภาวิณี ศรีสุขวัฒนานันท์, 2529 : 17) ได้ให้ความหมายของการเทียบมาตราไว้ว่า เป็นวิธีการทำคะแนนที่ได้จากแบบสอบสองชุดในวิชาเดียวกัน ให้เป็นคะแนนสมมูล (Equivalent Scores) ที่เปรียบเทียบกันได้โดยตรง โดยเสนอวิธีการให้กลุ่มผู้สอบกลุ่มเดียวทำแบบสอบสองชุด แล้วแปลงคะแนนแต่ละชุดให้เป็นคะแนนมาตรฐาน แล้วนำคะแนนแปลงมาเทียบกันโดยตรง

ฟลานาแกน (Flanagan, 1951 : 747-748) ได้ให้ความหมายของวิธีการเทียบมาตราไว้ว่า เป็นวิธีการทำคะแนนที่ได้จากแบบสอบต่างชุด ให้มีคุณลักษณะที่เปรียบเทียบกันได้ "คุณลักษณะที่เปรียบเทียบกันได้" หมายความว่า เมื่อกำหนดประชากรให้การแจกแจงของคะแนนจริงจากแบบสอบทั้งสองชุดที่ได้จากกลุ่มตัวอย่างซึ่งเลือกมาขนาดใหญ่ใด ๆ มีลักษณะเหมือนกันแล้ว คะแนนดิบของแบบสอบทั้งสองชุด จึงจะสามารถเปรียบเทียบกันได้ หรือถ้าความเที่ยงของแบบสองชุดนั้นเท่ากันในประชากรแล้ว ก็สามารถเปรียบเทียบการแจกแจงของค่าที่ได้เช่นกันความหมายดังกล่าวเป็นนิยามเชิงทฤษฎี ในทางปฏิบัติได้นิยามไว้ว่า ในประชากรที่กำหนดถ้าคะแนนจากแบบสอบสองชุดใด ๆ มีค่าเฉลี่ยเท่ากัน หรือเกือบเท่ากันในทุกๆ กลุ่มตัวอย่างขนาดใหญ่ใด ๆ แล้วการเปรียบเทียบจะทำได้

เลวิน (Levin, 1955 quoted in Holland and Rubin, 1982 : 10) ได้ให้ข้อตกลงเบื้องต้นของการเทียบมาตราไว้ว่า แบบสอบชุดใด ๆ จะเทียบมาตรากันได้ก็ต่อเมื่อแบบสอบทั้งสองนั้นคู่ขนานกัน (Parallel) ในด้านโครงสร้าง (Structure) เวลา(Timing) รูปแบบ (Format) ชนิดของข้อกระทง (Item type) และเนื้อหาวิชา (Subject Matter)

แองกอฟฟ์ (Angoff, 1971 : 562) ได้ให้ความหมายของการเปรียบเทียบมาตราไว้ว่า เป็นวิธีการแปลงระบบหน่วยคะแนนของแบบสอบชุดหนึ่ง ไปสู่ระบบหน่วยคะแนนของแบบสอบอีกชุดหนึ่ง คะแนนที่ผ่านการแปลงแล้ว จะให้ความหมายของการสมมูลกันโดยตรง

ลอร์ด (Lord, 1980 : 195) ได้ให้ความหมายของการเทียบมาตราไว้ว่า เป็นวิธีการแปลงคะแนนจากแบบสอบต่างชุด ที่มีความยากต่างกันจากความสามารถเดียวกัน

สงบ ลักษณะ (2522 : 22) ได้กล่าวว่า คะแนนจากแบบสอบ 2 ฉบับวัดสิ่งเดียวกัน แต่ไม่จำเป็นต้องเป็นข้อสอบคู่ขนาน จะถือว่าเทียบเท่ากันได้ถ้าคะแนนจากแบบสอบทั้ง 2 ฉบับนั้นมาจากคะแนนจริง(True Score) หรือความสามารถแท้ (True Ability) ที่เท่ากัน

ชูศักดิ์ ขัมภลิขิต (2527 : 2) ได้สรุปความหมายของการเทียบมาตราไว้ว่าเป็นกระบวนการที่เกี่ยวโยงกับกิจกรรม 2 ประการ คือ

1) กระบวนการที่ทำให้แบบสอบ 2 ฉบับใด ๆ มีความทัดเทียมกัน หรือเท่ากันในเชิงโครงสร้าง

2) การใช้วิธีการทางสถิติเพื่อปรับ (Adjust) คะแนนที่ได้จากแบบสอบแต่ละฉบับให้อยู่ในมาตราเดียวกัน และเทียบกันได้

จากความหมายดังกล่าว สรุปได้ว่า การเทียบมาตราคือ กระบวนการทางสถิติเพื่อที่จะปรับคะแนนที่ได้จากแบบสอบต่างชุดกันที่มีโครงสร้างเดียวกันให้สามารถเปรียบเทียบกันได้

รูปแบบการเทียบมาตรา

รูปแบบการเทียบมาตราพิจารณาลักษณะรูปแบบของทฤษฎีการวัดแล้วสามารถจำแนกได้ 2 รูปแบบ คือ การเทียบมาตรารูปแบบดั้งเดิม (Traditional Model) ซึ่งแบ่งย่อยเป็น การเทียบมาตรารูปแบบอิควิเปอร์เซ็นไทล์ (Equipercentile Equating) การเทียบมาตราเชิงเส้นตรง (Linear Equating) การเทียบมาตรารูปแบบการวิเคราะห์ตัวประกอบ (Equating Using The Confirmatory Factor Analysis Model) และการเทียบมาตรารูปแบบทฤษฎีการตอบข้อสอบ

(Item Response Theory Model) หลักการเทียบมาตราแต่ละรูปแบบมีรายละเอียดดังต่อไปนี้

1. การเทียบมาตรารูปแบบดั้งเดิม (Traditional Model)

1.1 การเทียบมาตรารูปแบบอิควิเปอร์เซ็นไทล์(Equipercentile Equating)

1.1.1 การเทียบมาตรารูปแบบอิควิเปอร์เซ็นไทล์โดยไม่ใช้แบบสอบร่วม

มโนทัศน์ของการเทียบมาตราวิธีอิควิเปอร์เซนไทล์ เริ่มจากการแจกแจงของคะแนนจากแบบสอบฉบับ X และฉบับ Y ที่มีลักษณะคล้ายกันหรือแตกต่างกันบ้างเล็กน้อย เช่น ต่างกันที่ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน การเทียบมาตราสมมูลทำได้โดยใช้คะแนน ณ ตำแหน่งเปอร์เซนไทล์เดียวกันของคะแนน 2 ฉบับนั้น ผลการเทียบคะแนนแสดงด้วยกราฟ ขั้นตอนการแปลงคะแนนมีดังนี้ คือ เลือกกลุ่มผู้สอบที่มีความสามารถกระจาย มีทั้งเก่งปานกลาง และอ่อนแล้วแบ่งกลุ่มย่อยโดยสุ่ม 2 กลุ่มให้กลุ่มหนึ่งทำแบบสอบ X และอีกกลุ่มหนึ่งทำแบบสอบ Y ผลการสอบนำมาทำการแจกแจงคะแนน X และ Y คำนวณหาจุดกลางเปอร์เซนไทล์ของแต่ละการแจกแจง อ่านและทำเครื่องหมายสำหรับค่า X และ Y จากการแจกแจงคะแนนที่สมนัยกันบนกระดาษกราฟ ทำจุดบนกระดาษประมาณ 30 จุด แล้วลากเส้นเชื่อมจุดจะเกิดเป็นเส้นกราฟ ทำการปรับเส้นให้เรียบเส้นกราฟนี้จะใช้อ่านค่า X ที่สมนัยกับ Y หรืออ่านค่า Y ที่สมนัยกับ X ก็ได้ จากนั้นทำตารางสำเร็จเพื่ออ่านค่าคะแนนแปลงโดยปกติการเทียบมาตราวิธีอิลวิเปอร์เซนไทล์ให้ภาพของการแปลงที่สะท้อนถึงระดับความยากง่ายของแบบสอบสองชุด ถ้าแบบสองชุดมีความยากใกล้เคียงกันเส้นกราฟจะมีลักษณะใกล้เคียงเส้นตรง แต่ถ้าแบบสอบมีความยากต่างกันเส้นกราฟจะเป็นเส้นโค้ง (Curvilinear) คะแนนสมมูลที่เกิดขึ้น จะถูกยืดหรือหยุ่นคะแนนดิบเพื่อให้คงรักษาคะแนนให้เหมือนชุดก่อนตามต้องการ (Angoff 1984 : 97-101)

ขั้นตอนในการเทียบมาตราวิธีอิควิเปอร์เซนไทล์ สรุปเป็นขั้นตอนได้ 2 ขั้นตอน (Two-Stage) คือ

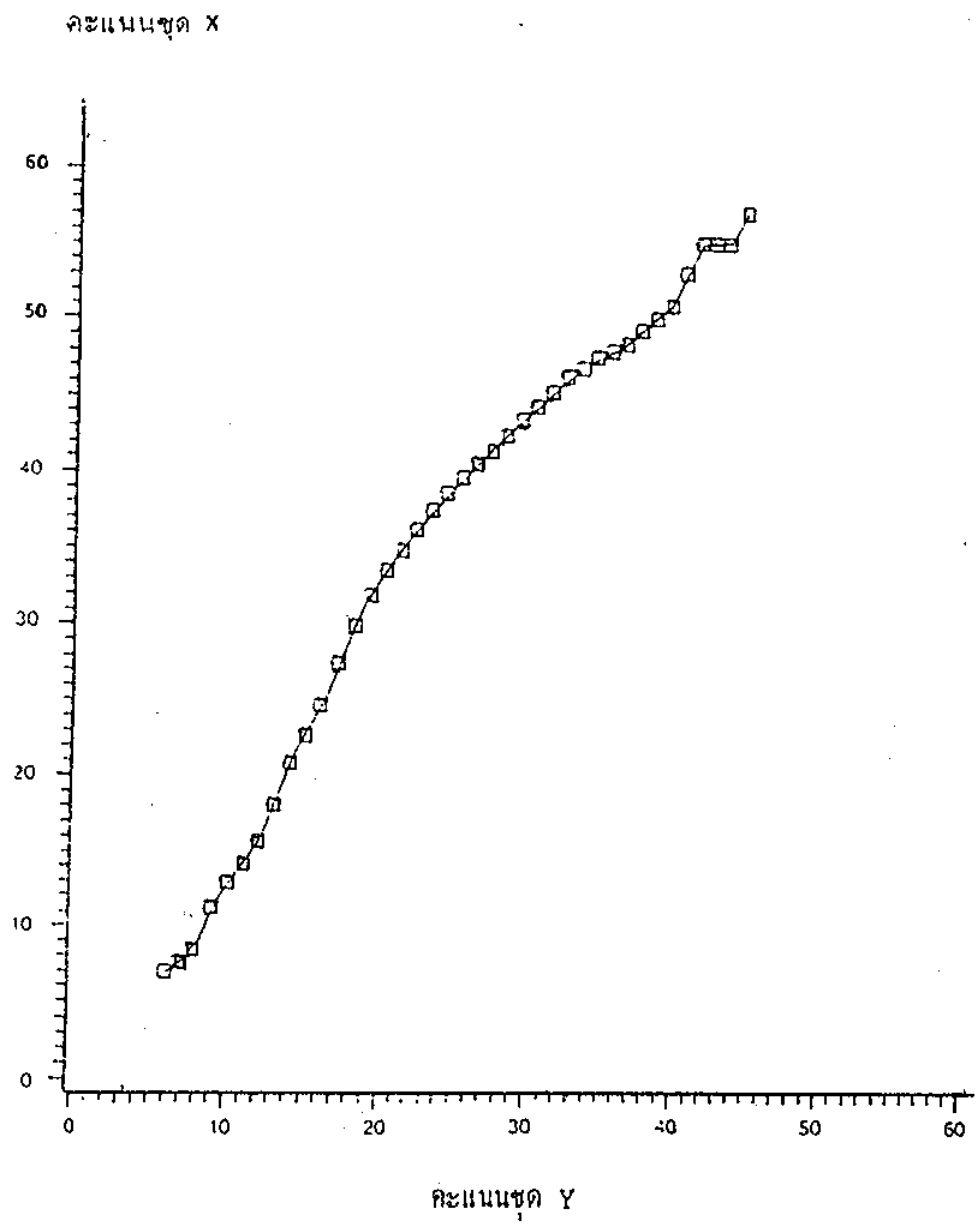
1.1.1.1 การกระจายความถี่สะสมที่มีความสัมพันธ์เป็นตารางหรือเป็นกราฟสำหรับสองแบบสอบที่นำมาเทียบมาตรา คือ

ก. นำคะแนนของกลุ่มผู้สอบที่ มีความสามารถกระจายทั้ง เก่ง ปานกลาง และอ่อนซึ่งถูกแบ่งเป็นกลุ่มย่อยสองกลุ่มโดยการสุ่ม ให้กลุ่มหนึ่งทำแบบสอบ X และอีกกลุ่มทำแบบสอบ Y มาทำการแจกแจงคะแนน X และ Y

ข. คำนวณหาจุดกลางเปอร์เซนไทล์ของแต่ละการแจกแจง

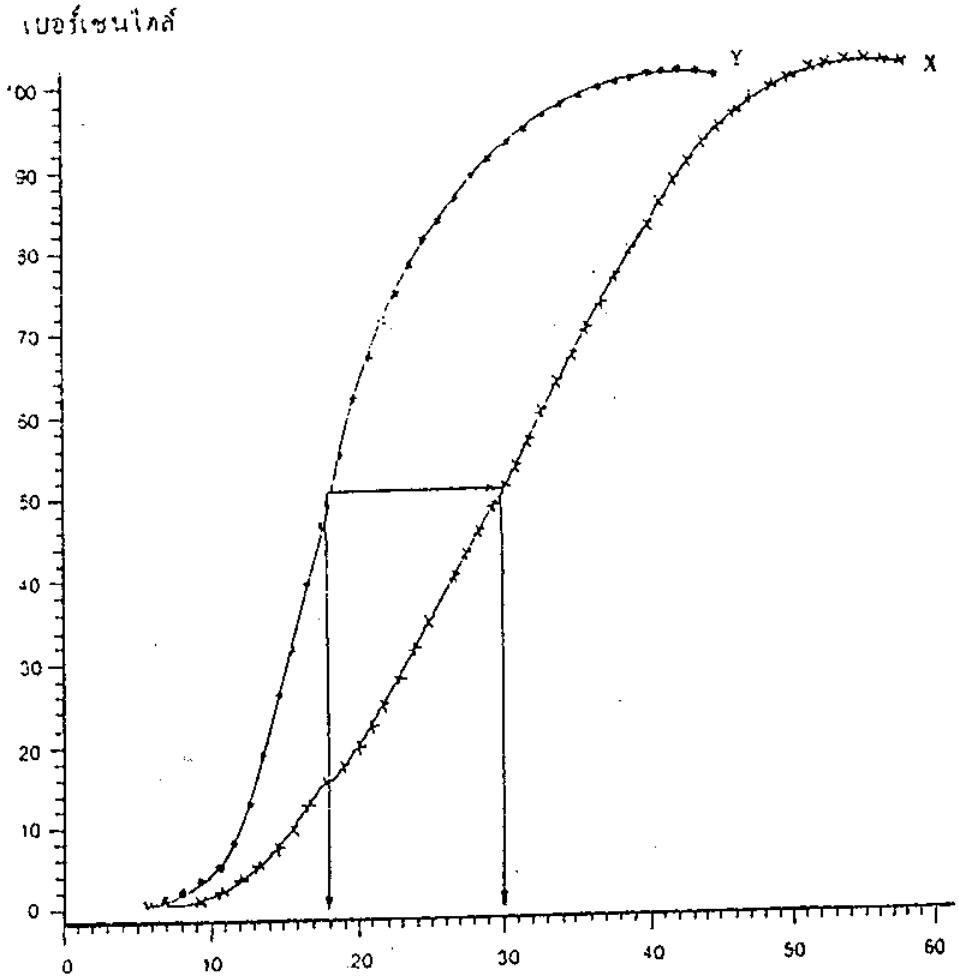
ค. อ่านและทำเครื่องหมายสำหรับค่าคะแนนของแบบสอบฉบับ X และฉบับ Y

ของการแจกแจงที่สมมูลกันบนกระดาษกราฟ ซึ่งแองกอฟฟ์แนะนำให้ใช้กระดาษ Arithmatic Graph Paper (โดยแกนนอนเป็นคะแนนดิบ แกนตั้งเป็นตำแหน่งเปอร์เซนไทล์) ดังภาพที่ 1 ประมาณ 30 จุด และลากเส้นเชื่อมเกิดเป็นเส้นกราฟ



ภาพที่ 1: กระบวนการเทียบมาตราวิธีอิควิเปอร์เซนไทล์

1.1.1.2 เทียบมาตราคะแนนจากรายเอียดในข้อ ค ขั้นตอนที่หนึ่ง นำมาพล็อตกราฟใหม่ลงบนกระดาษกราฟ โดยแกนนอนเป็นคะแนนฉบับ Y แกนตั้งเป็นคะแนนฉบับ X ทำการปรับเส้นกราฟให้เรียบ เส้นกราฟนี้จะใช้อ่านค่า X ที่สมมูลกันกับ Y จากนั้นสร้างตารางสำเร็จเพื่ออ่านค่าคะแนนแปลงจากกราฟ ดังภาพที่ 2



ภาพที่ 2 : การแปลงคะแนนของการเทียบมาตราวิธีอิควิเปอร์เซนไทล์

1.1.2 การเทียบมาตราวิธีอิควิเปอร์เซนไทล์โดยใช้แบบสอบร่วม

ขั้นตอนการเทียบมาตราวิธีอิควิเปอร์เซนไทล์ โดยใช้แบบสอบร่วม สำหรับกลุ่มที่ไม่ได้สุ่มซึ่งที่มีความสามารถไม่แตกต่างกันมาก โดยที่กลุ่ม α สอบฉบับ X และฉบับ U กลุ่ม β สอบฉบับ Y และฉบับ U ขั้นตอนในการเทียบมาตรา คือ ประมาณค่า ความถี่ในแบบสอบฉบับ X และ Y สำหรับกลุ่ม $t(\alpha+\beta)$ มีขั้นตอนดังนี้

1.1.2.1. รวมคะแนนของฉบับ U ที่สอบโดยกลุ่ม α , β และกลุ่ม t

1.1.2.2. หาสัดส่วนความถี่ $\frac{f_i t}{f_i \alpha}$ และ $\frac{f_i t}{f_i \beta}$ ที่ทุกช่วงคะแนน i

1.1.2.3. กระจายความถี่ของฉบับ U ในแต่ละช่วงคะแนนของฉบับ X และฉบับ Y

1.1.2.4. คูณความถี่ของฉบับ Uในแต่ละช่วงคะแนนของฉบับ Y ด้วยสัดส่วน $\frac{f_i t}{f_i \alpha}$

1.1.2.5. คูณความถี่ของฉบับ U ในแต่ละช่วงคะแนนของฉบับ Y ด้วยสัดส่วน $\frac{f_i t}{f_i \beta}$

1.1.2.6. หาตำแหน่งเปอร์เซนไทล์ของความถี่ที่ประมาณค่าใหม่ของทั้งสองฉบับ

1.1.2.7. ดำเนินการเทียบตามขั้นตอนที่สองของการเทียบโดยวิธีอิควิเปอร์เซนไทล์ดังกล่าวแล้วข้างต้น

1.2 การเทียบมาตรารูปแบบเชิงเส้นตรง (Linear Equating)

การเทียบมาตราเชิงเส้นตรงมีนิยามว่า คะแนนของแบบสอบ 2 ชุด จะเท่าเทียมกัน ถ้าต่างก็ตรงกับคะแนนมาตรฐานค่าเดียวกัน (Angoff, 1971 : 564 ; Petersen, et al. 1982 : 73) เนื่องจากการเทียบมาตราเป็นการศึกษาเชิงประจักษ์ เพื่อกำหนดคะแนนแปลงที่ได้จากแบบสอบชุดหนึ่งไปสู่แบบสอบอีกชุดหนึ่ง จึงมีองค์ประกอบหลายประการเข้ามาเกี่ยวข้อง ได้แก่ การออกแบบเพื่อเก็บรวบรวมข้อมูล และการจัดกระทำข้อมูลในทางสถิติเพื่อการแปลงคะแนน แองกอฟฟ์ (Angoff, 1984 : 93-121) ได้เสนอแบบแผนการรวบรวมข้อมูลและการจัดกระทำข้อมูลในทางสถิติของการเทียบมาตราเชิงเส้นตรงไว้ 6 รูปแบบ รูปแบบที่ 1-2 เป็นการเทียบมาตราโดยไม่ได้ใช้แบบสอบร่วม รูปแบบที่ 3-6 เป็นการเทียบมาตราโดยใช้แบบสอบร่วม ซึ่งแต่ละรูปแบบจะมีวิธีการประมาณค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐาน ของแบบสอบแตกต่างกันไป ตามเงื่อนไขของรูปแบบการรวบรวมข้อมูล แต่ทุกรูปแบบจะตัดสินคะแนนสมมูล จากค่าคะแนนมาตรฐานค่าเดียวกัน คือ

$$\frac{Y-M_y}{S_y}=\frac{X-M_x}{S_x}$$

เมื่อ X , Y คือ คะแนนจากแบบสอบฉบับ X , Y

M_x , M_y คือ ค่าเฉลี่ยจากคะแนนแบบสอบฉบับ X , Y

S_x , S_y คือ ค่าเบี่ยงเบนมาตรฐานของคะแนนแบบสอบฉบับ X , Y

ในแต่ละรูปแบบการรวบรวมข้อมูลของแองกอฟฟ์ ซึ่งเสนอรายละเอียดแยกระหว่างกรณีที่แบบสอบเทียบมาตรามีความเที่ยงเท่ากัน และเมื่อแบบสอบเทียบมาตรามีความเที่ยงไม่เท่ากันจึงขอเสนอแบบแผนการรวบรวมข้อมูลของแองกอฟฟ์ 6 รูปแบบ ดังตารางที่ 1 - 4 และรายละเอียดดังต่อไปนี้

ตารางที่ 1 Equivalent - Groups Design

กลุ่มตัวอย่าง	แบบสอบ	
	X	Y
P_1	√	
P_2		√

ตารางที่ 2 Counterbalanced Random - Groups Design

กลุ่มตัวอย่าง	แบบสอบ			
	X		Y	
	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 1	ครั้งที่ 2
P_1	√			√
P_2		√	√	

ตารางที่ 3 Anchor - Test - Random - Groups Design

กลุ่มตัวอย่าง	แบบสอบ		
	X	Y	U
P_1	√		√
P_2		√	√

ตารางที่ 4 Anchor - Test Nonequivalent - Groups Design

กลุ่มตัวอย่าง	แบบสอบ		
	X	Y	U
P_1	√		√
Q_1		√	√

1.2.1 Random Group - One Test Administered To Each Group

รูปแบบนี้ผู้สอบมาจากการสุ่มจากประชากรกลุ่มเดียวกัน แบ่งเป็น 2 กลุ่ม คือ กลุ่ม α สอบฉบับ X กลุ่ม β สอบฉบับ Y

1.2.1.1. การเทียบมาตราเชิงเส้นตรงสำหรับแบบสอบที่มีความเที่ยงเท่ากัน

คำนวณค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐานของคะแนนแบบฉบับ X โดย กลุ่ม α และแบบสอบฉบับ Y โดยกลุ่ม β จัดให้คะแนนมาตรฐานของทั้ง 2 กลุ่มเท่ากัน

$$\frac{Y - M_{y\beta}}{S_{y\beta}} = \frac{X - M_{x\alpha}}{S_{x\alpha}} \qquad \text{.....[1]}$$

โดยที่ Y คือ คะแนนสอบที่ได้จากฉบับ Y

$M_{y\beta}$ คือ ค่าเฉลี่ยของคะแนนฉบับ Y จากการสอบโดยกลุ่ม β

$S_{y\beta}$ คือ ค่าเบี่ยงเบนมาตรฐานของคะแนนฉบับ Y จากการสอบโดยกลุ่ม β

X คือ คะแนนสอบที่ได้จากฉบับ X

$M_{x\alpha}$ คือ ค่าเฉลี่ยของคะแนนฉบับ X จากการสอบโดยกลุ่ม α

$S_{x\alpha}$ คือ ค่าเบี่ยงเบนมาตรฐานของคะแนนฉบับ X จากการสอบโดยกลุ่ม α

เมื่อจัดเทอมในสมการ [1] เสียใหม่ จะได้สมการเชิงเส้นตรง คือ

$$Y = \frac{S_{y\beta}}{S_{y\alpha}} X + M_{y\beta} - \frac{S_{y\beta}}{S_{x\alpha}} M_{x\alpha} \qquad \text{.....[2]}$$

ซึ่งจะอยู่ในรูป $Y = AX + B$

เมื่อ $A = \frac{S_{y\beta}}{S_{x\alpha}}$

$B = M_{y\beta} - A_{x\alpha}$

1.2.1.2 การเทียบมาตราเชิงเส้นตรงสำหรับแบบสอบที่มีความเที่ยงไม่เท่ากัน

ถ้าแบบสอบฉบับ X และฉบับ Y มีความเที่ยงแตกต่างกันอย่างมีนัยสำคัญแล้ว จะไม่มีหนทางที่จะพัฒนาการแปลงคะแนนให้มีลักษณะเทียบเคียงกันได้เลย อย่างไรก็ตาม สถานการณ์ในทางปฏิบัติที่เรียกว่าการเทียบมาตรานั้นเราสามารถที่จะนำเอาความเที่ยงของแบบสอบทั้งสองฉบับรวมเข้าไปด้วย ดังสมการ

$$\frac{Y - M_{y\beta}}{S_{y\beta}} = \frac{X - M_{x\alpha}}{S_{x\alpha}} \quad \text{.....[3]}$$

เมื่อ $S_{y\beta} = S_y \sqrt{r_{yy}}$

$S_{x\alpha} = S_x \sqrt{r_{xx}}$

เมื่อ Y คือ คะแนนสอบที่ได้จากฉบับ Y

$M_{y\beta}$ คือ ค่าเฉลี่ยของคะแนนฉบับ Y จากการสอบโดยกลุ่ม β

S_y คือ ค่าเบี่ยงเบนมาตรฐานของคะแนนสอบ (Observe Score) ของฉบับ Y

$S_{y\beta}$ คือ ค่าเบี่ยงเบนมาตรฐานของคะแนนจริง (True Score) ของฉบับ Y

r_y คือ ความเที่ยงของแบบสอบฉบับ Y

X คือ คะแนนสอบที่ได้จากฉบับ X

$M_{x\alpha}$ คือ ค่าเฉลี่ยของคะแนนฉบับ X จากการสอบโดยกลุ่ม α

S_x คือ ค่าเบี่ยงเบนมาตรฐานของคะแนนสอบ (Observe Score) ของฉบับ X

$S_{\bar{x}x}$ คือ ค่าเบี่ยงเบนมาตรฐานของคะแนนจริง (True Score) ของฉบับ X

r_{xx} คือ ความเที่ยงของแบบสอบฉบับ X

สมการสำหรับแปลงคะแนนจากสเกลฉบับ X ไปสู่ฉบับ Y จะเป็น

$$Y = \frac{S_{y\beta}}{S_{x\alpha}} X + M_{y\beta} - \frac{S_{y\beta}}{S_{x\alpha}} M_{x\alpha} \quad \text{.....[4]}$$

ซึ่งอยู่ในรูป $Y = AX + B$

เมื่อ $A = \frac{S_{y\beta}}{S_{x\alpha}}$

$B = M_{y\beta} - AM_{x\alpha}$

1.2.2 Random Groups ,Both Forms Administered To Each Group, C ounterbalanced

ลักษณะเช่นเดียวกับรูปแบบที่ 1 เลือกกลุ่มตัวอย่างขนาดใหญ่แล้วแบ่งแบบสุ่มออกเป็น 2 กลุ่ม โดยกลุ่มหนึ่งเป็นกลุ่ม α สอบฉบับ X แล้วตามด้วยฉบับ Y อีกกลุ่มหนึ่งเป็นกลุ่ม β สอบฉบับ Y แล้วตามด้วยฉบับ X ในการที่จะป้องกันความคลาดเคลื่อนที่จะเกิดขึ้นในการสอบโดยเฉพาะอย่างยิ่ง เมื่อสอบแบบฉบับ X และฉบับ Y พร้อมกัน ควรจะเย็บแบบสอบทั้ง 2 ฉบับติดกัน

1.2.2.1. วิธีการเทียบมาตราเชิงเส้นตรงกรณีแบบสอบมีความเที่ยงเท่ากัน

วิธีการนี้โดยมีข้อตกลงว่า ผลกระทบที่เกิดจากการฝึกที่มีต่อแบบสอบฉบับ Y อันเป็นผลมาจากการสอบฉบับ X ก่อน และผลกระทบที่มีต่อฉบับ X อันเป็นผลมาจากการสอบฉบับ Y ก่อน ต่างก็เป็นสัดส่วนของค่าเบี่ยงเบนมาตรฐานของแบบสอบทั้ง 2 ฉบับ

$$\frac{K_x}{S_x} = \frac{K_y}{S_y} = H$$

การประมาณค่าที่ดีที่สุดของ H คือ การหาค่าเฉลี่ยของความแตกต่างระหว่างค่าเฉลี่ยของฉบับ X และค่าเฉลี่ยของฉบับ Y เมื่อแต่ละฉบับอยู่ในหน่วยของค่าเบี่ยงเบนมาตรฐาน

$$H = \frac{1}{2}\left[\frac{M_{\alpha\beta} - M_{x\alpha}}{S_x} + \frac{M_{y\alpha} - M_{y\beta}}{S_Y}\right] \quad \text{.....[5]}$$

ซึ่งสมการนี้ก็คือ ต้นแบบของสมการสำหรับเทียบเชิงเส้นตรง สมการ [1]

$$\frac{(Y - M_y)}{S_y} = \frac{(X - M_x)}{S_x}$$

สมการที่ใช้สำหรับแทนค่าในสมการ [1] คือ

$$M_x = \frac{1}{2}(M_{x\alpha} + M_{x\beta} - K_x) \quad \text{.....[6]}$$

$$M_y = \frac{1}{2}(M_{y\alpha} + M_{y\beta} - K_y) \quad \text{.....[7]}$$

$$S_x^2 = \frac{1}{2}(S_{x\alpha}^2 + S_{x\beta}^2) \quad \text{..... [8]}$$

$$S_y^2 = \frac{1}{2}(S_{y\alpha}^2 + S_{y\beta}^2) \quad \text{.....[9]}$$

เมื่อสมการ [6] ถึง [9] แทนค่าลงในสมการ [1] จะได้สมการเชิงเส้นในรูปของ $Y = AX + B$ เมื่อ

$$A = \sqrt{\frac{S_y^2 + S_y^2}{S_x^2 + S_x^2}}$$

$$B = \frac{1}{2}(M_{y\alpha} + M_{y\beta}) - A(M_{x\alpha} + M_{x\beta})$$

1.2.2.2 การเทียบมาตรฐานเชิงเส้นตรงสำหรับแบบสอบที่มีความเที่ยงไม่เท่ากัน

เมื่อฉบับ X และฉบับ Y มีความเที่ยงไม่เท่ากัน ค่าเฉลี่ยและความแปรปรวนในกลุ่มไม่ควรจะคำนวณจากคะแนนที่ได้จากการสอบในสมการ[8] และ [9] แต่ควรจะคำนวณจากคะแนนจริง โดยค่าที่สมนัยกับสมการ [8] และ [9] คือสมการ [10] และ [11]

$$S_x^2 = \frac{1}{2}(S_{x\alpha}^2 r_{xx\alpha} + S_x^2 r_{xx\beta}) \quad \text{.....[10]}$$

$$S_y^2 = \frac{1}{2}(S_{y\alpha}^2 r_{yy\alpha} + S_y^2 r_{yy\beta}) \quad \text{.....[11]}$$

1.2.3 Random Groups - One Test Administered To Each Group Common Equating Test Administrated To Both Groups

1.2.3.1 การเทียบมาตราเชิงเส้นตรงเมื่อแบบสอบมีความเที่ยงเท่ากัน

คะแนนสอบที่ได้ส่วนใหญ่ขึ้นกับลักษณะ 2 ประการ คือ ความสามารถรายบุคคลและลักษณะของแบบสอบ ในการที่จะเปรียบเทียบผลการสอบของแต่ละบุคคลซึ่งสอบแบบสอบที่ต่างกัน จึงจำเป็นที่จะต้องปรับคะแนนเสียก่อน กล่าวคือ การเทียบมาตรานั้นจะเป็นเพียงการเทียบความแตกต่างของคะแนนอันเป็นผลมาจากความแตกต่างของแต่ละบุคคล หรือของกลุ่ม ถ้ากลุ่ม α และ β ไม่ได้สุ่มมาจากประชากรเดียวกัน ความแตกต่างระหว่างกลุ่มอาจจะเป็นองค์ประกอบที่มีนัยสำคัญในการแปรเปลี่ยนค่าของ A และ B สมการ [2] ,[3] และจะเป็นผลทำให้เกิดอคติ (Bias) ในวิธีการเทียบมาตรา แม้ว่ากลุ่มจะถูกเลือกมาโดยวิธีการสุ่ม ความแตกต่างเล็กน้อยระหว่างกลุ่มอาจเกิดขึ้นได้ ถ้าไม่ป้องกันก็อาจเกิดอคติในสมการแปลงคะแนน จะเป็นผลต่อการเปรียบเทียบภายหลังในการที่ควบคุมให้บังเกิดผล การเทียบมาตราจึงจำเป็นต้องใช้คะแนนของแบบฉบับ U ซึ่งเป็นข้อสอบที่เพิ่มขึ้น หรือเป็นข้อสอบร่วมกันระหว่างฉบับ X และฉบับ Y แบบสอบร่วมนี้ใช้สำหรับปรับความแตกต่างที่อาจจะพบระหว่างกลุ่ม α และ β ในการบริหารแบบสอบ กลุ่ม α จะได้รับแบบสอบฉบับ X และฉบับ U และกลุ่ม β จะได้รับแบบสอบฉบับ Y และฉบับ U ดังสมการ

$$\hat{\mu}_x = M_{x\alpha} + b_{xu\alpha}\left(\hat{\mu}_u - M_{u\alpha}\right) \quad \text{.....[12]}$$

$$\hat{\mu}_y = M_{y\beta} + b_{yu\beta}\left(\hat{\mu}_u - M_{u\beta}\right) \quad \text{.....[13]}$$

$$\hat{\sigma}_x^2 = s_{x\alpha}^2 + b_x^2\,\alpha\left(\hat{\sigma}_u^2 - s_{u\alpha}^2\right) \quad \text{.....[14]}$$

$$\hat{\sigma}_x^2 = s_{y\beta}^2 + b_y^2\,\beta\left(\hat{\sigma}_u^2 - s_{u\beta}^2\right) \quad \text{.....[15]}$$

เมื่อ $\hat{\mu}_u = M_u$

$\hat{\sigma}_u^2 = s_u^2$

และ $t = \alpha + \beta$

ซึ่งการประมาณค่าเหล่านี้ใช้กับสมการ [1] $Y = AX + B$

เมื่อ $A = \dfrac{\hat{\sigma}_y}{\hat{\sigma}_x}$

และ $B = \hat{\mu}_y - A\hat{\mu}_x$

การเทียบมาตราโดยวิธีนี้มีประโยชน์มากสามารถยืดหยุ่นและปรับให้เข้ากับสถานการณ์อื่นๆได้ เช่น ฉบับ U อาจจะสอบรวม หรือแยกจากฉบับ X และฉบับ Y ฉบับ U อาจจะเป็นส่วนหนึ่งร่วมกับฉบับ X แต่แยกจากฉบับ Y และฉบับ U อาจจะร่วมเข้าไปภายในฉบับ X และฉบับ Y โดยกระจายทั่วไปตลอดทั้งฉบับ อย่างไรก็ตามแบบสอบร่วมควรจะมีความยาวและความเที่ยงเพียงพอที่จะให้ข้อมูลไปปรับความแตกต่างระหว่างกลุ่มตามที่คาดหวัง ซึ่งแองกอฟฟ์ (Angoff 1971 : 578) ได้เสนอกฎเกณฑ์เกี่ยวกับจำนวนข้อของแบบสอบร่วม คือ ไม่ควรน้อยกว่า 20 ข้อ หรือ 20 % ของจำนวนข้อในแต่ละแบบฉบับแล้วแต่จำนวนใดมากกว่าให้ใช้จำนวนนั้น และเพื่อที่จะหลีกเลี่ยงเกี่ยวกับแบบสอบที่ใช้วัดความเร็ว ควรกระจายฉบับ U กับกลุ่ม α และ β คือจะต้องเป็นการสอบเหมือนกันทั้งสองกลุ่มในกรณีการแยกสอบคนละเวลา

การเทียบมาตรารูปแบบที่ 3 นี้ยังมีวิธีการที่ยืดหยุ่นได้ คือสามารถที่จะเทียบแบบสอบที่มากกว่า 2 ฉบับขึ้นไป อาจจะเป็น 3 ฉบับ 4 ฉบับ หรือมากว่านั้น เช่น กรณี 3 ฉบับ คือ ฉบับ X ,Y และ Z แบบสอบทั้ง 3 ฉบับนี้สอบโดยกลุ่ม α ,β และ γ ตามลำดับ ค่าเฉลี่ยและความแปรปรวนสำหรับกลุ่มรวม (α ,β และ γ) อาจจะประมาณได้โดยใช้ข้อตกลงเกี่ยวกับการเทียบมาตราของแบบสอบ 2 ฉบับ

1.2.3.2 การเทียบมาตราเชิงเส้นตรงเมื่อแบบสอบมีความเที่ยงไม่เท่ากัน

ในกรณีของกลุ่มสุ่ม แบบสอบฉบับ X และฉบับ Y มีความเที่ยงไม่เท่ากัน ซึ่งเหมาะกับการแปลงโดยคะแนนจริงมากกว่าคะแนนที่สอบได้ ข้อตกลงที่เพิ่มขึ้นอีกคือ ฉบับ U จะต้องมีลักษณะคู่ขนานกันทางโครงสร้างกับทั้งฉบับ X และฉบับ Y ภายใต้เงื่อนไขนี้เมื่อฉบับ U แยกออกจากฉบับ X และ Y ความชัน (Slope) และจุดตัดแกน (Intercept) ของสมการ $Y = AX + B$ คือ

$$A = \frac{b_{yu\beta}}{b_{xy\beta}}$$

$$B = \hat{\mu}_y - A\hat{\mu}_x$$

เมื่อ $b_x\,\alpha$ และ $b_y\,\beta$ คือ สัมประสิทธิ์กระถดถอย (Regression Coefficients) ที่พบในกลุ่ม α และ β สำหรับการทำนาย X จาก U และ Y จาก U และ $\hat{\mu}_x$, $\hat{\mu}_y$ คำนวณจากสมการที่ [12] ,[13]

นอกจากนี้เพิ่มข้อตกลงเบื้องต้นอีกว่า ฉบับ U ต้องเป็นฉบับคู่ขนานกับแบบสอบ X และ Y เมื่อฉบับ U รวมอยู่ในส่วนของ X และ Y แล้ว

PRINCE OF SONGKLA UNIVERSITY
PATTANI, THAILAND

$$A = \frac{b_{xu\alpha}\hat{\sigma}_y^2}{b_{yu\beta}\hat{\sigma}_x^2}$$

และ $$B = \hat{\mu}_y - A\hat{\mu}_x$$

ส่วนค่าของ $\hat{\mu}_x, \hat{\mu}_y, \hat{\sigma}_x^2$ และ $\hat{\sigma}_y^2$ คำนวณได้จากสมการที่ [12] - [15]

1.2.4 Nonrandom Group - One Test To Each Group, Common Equating Test Administered To Both Groups

สถานการณ์ในการสอบบางครั้งถือเป็นความลับที่ไม่เปิดเผย ไม่สามารถที่จะแนะนำแบบสอบฉบับใหม่ก่อนที่สอบได้ และคำสั่งของโปรแกรมการสอบไม่อนุญาตให้อธิบายนอกเหนือไปกว่าที่สอบ ภายใต้สถานการณ์ เหล่านี้การเทียบมาตราจากการสุ่ม 3 วิธีไม่สามารถนำมาใช้ได้และข้อมูลที่ใช้สำหรับการเทียบมาตราต้องมาจากผลการสอบของกลุ่มที่สอบจริงๆแต่ละฉบับ ซึ่งการสอบต่างฉบับกันและเวลาสอบต่างกัน ผู้สอบอาจจะไม่ได้มาจากกลุ่มประชากรเดียวกัน ตัวอย่างเช่น มีการจัดสอบฉบับใหม่ในเดือนกันยายน 2537 ซึ่งผู้สอบอาจมีลักษณะเหมือนกันหลายๆประการ แต่ยังไม่แน่ใจว่ากลุ่มผู้สอบทั้งสองฉบับมาจากประชากรเดียวกัน ดังนั้นการเลือกกลุ่ม α และ β จึงควรระมัดระวังเพื่อที่จะลดความแตกต่างให้น้อยลง ใช้วิธีการ คือ

1.2.4.1 การเทียบมาตราเชิงเส้นตรงสำหรับกลุ่มที่มีความสามารถแตกต่างกันน้อย วิธีการที่จะเสนอต่อไปนี้เป็นวิธีการหาความเป็นไปได้สูงสุด (Maximum Likelihood) กับวิธีการที่ใช้การปรับด้วยคะแนนจริง (True Score Adaptation) โดยกลุ่ม α สอบฉบับ X และ กลุ่ม β สอบฉบับ Y และทั้งสองกลุ่มสอบฉบับ U ซึ่งเป็นแบบสอบร่วม และใช้ปรับความแตกต่างที่อาจจะเกิดขึ้นระหว่างกลุ่ม การประมาณค่าเฉลี่ยและความแปรปรวนของทั้งฉบับ X และฉบับ Y กระทำโดยการรวมกลุ่ม α และกลุ่ม β เป็นกลุ่ม t การประยุกต์ใช้กับสมการที่ [1] ซึ่งเป็นสมการเชิงเส้นที่เป็นความสัมพันธ์เชิงเส้นของคะแนนดิบฉบับ X และฉบับ Y สมการที่จะเสนอต่อไปจะขึ้นกับข้อตกลงเบื้องต้น 3 ประการ ของทฤษฎี Univariate Selection Theory นั้นคือ

ก. จุดตัดแกนของคะแนนฉบับ Y จากฉบับ U โดยกลุ่ม t เหมือนกับกลุ่ม α

$$M_{xt} - b_{xut}M_{ut} = M_{x\alpha} - b_{xu\alpha}M_{u\alpha} \qquad [16]$$

ACC. No ... 3711Eb
DATE RECEIVED .....
CALL No. ... T420 ...

ข. สัมประสิทธิ์กระถดถอย (Regression Coefficient) ของฉบับ X จากฉบับ U สำหรับกลุ่ม t เหมือนกับกลุ่ม α

$$b_{xut} = b_{xu\alpha} \qquad \text{.....[17]}$$

ค. ความแปรปรวนของความคลาดเคลื่อนในการประมาณค่าของ X จากฉบับ U ของกลุ่ม t เหมือนกับกลุ่ม α

$$S_{xt}^2(1-r_{xut}^2) = S_{x\alpha}^2(1-r_{xu\alpha}^2) \qquad \text{.....[18]}$$

แทนค่าสมการ [17] ลงใน [16] เพื่อหาค่า M_x

$$\hat{M}_{xt} = M_{x\alpha} + b_{xu\alpha}(M_{ut} - M_{u\alpha}) \qquad \text{.....[19]}$$

(สัญลักษณ์ ($\hat{}$) ใช้ในการประมาณค่า)
แทนค่าลงในสมการ [18] โดยที่ b_x S_u แทน r_x S_x และ $b_x \alpha S_{u\alpha}$
แทนค่า $r_{xu\alpha}S_{x\alpha}$ เพื่อหาค่า S_x

$$\hat{S}_{xt}^2 = S_{x\alpha}^2 + b_{xu\alpha}^2(S_{ut}^2 - S_{u\alpha}^2) \qquad \text{.....[20]}$$

ด้วยข้อตกลงเบื้องต้นลักษณะเดียวกันของฉบับ X และฉบับ Y นำมาหาความสัมพันธ์ระหว่างฉบับ Y และฉบับ U กลุ่ม β และกลุ่ม t ได้ผลลักษณะเดียวกับสมการ [19] , [20] คือ

$$\hat{M}_{yt} = M_{y\beta} + b_{yu\beta}(M_{ut} - M_{u\beta}) \qquad \text{.....[21]}$$

และ
$$S_{yt}^2 = S_{y\beta}^2 + b_{yu\beta}^2(S_{ut}^2 - S_{u\beta}^2) \qquad \text{.....[22]}$$

สัญลักษณ์ $\hat{M}_x, S_x^2, \hat{M}_y$ และ $\hat{S}_y^2$ ที่ใช้ในสมการ [19] ถึง [22] นำมาใช้แทน
สัญลักษณ์ μ_x, σ_x^2, μ_y และ σ_y^2 ตามลำดับ เพราะว่าในที่นี้เป็นการประมาณค่ากลุ่มรวม ไม่ใช่การประมาณค่าสำหรับประชากร

สมการ [19] ถึง [22] แทนค่าในสมการ [1] คือ

$$\frac{(Y - M_y)}{S_y} = \frac{(X - M_x)}{Sx}$$

ซึ่งจะได้สมการสำหรับแปลงคะแนน $Y = AX + B$ โดยที่

$$A = \frac{\hat{S}_{yt}}{\hat{S}_{xt}} \qquad \text{.....[23]}$$

$$B = \hat{M}_{yt} - A\hat{M}_{xt} \qquad \text{.....[24]}$$

เป็นที่สังเกตได้ว่า วิธีการคำนวณสำหรับการประมาณค่าในสมการ [19] -[22] เหมือนกับสมการ [12] -[15] ตามลำดับ แม้ว่าสมการทั้ง 2 ชุดจะแตกต่างกัน

ลักษณะเช่นเดียวกันนี้ สามารถที่จะยืดหยุ่นใช้กับสถานการณ์อื่นได้อีก เช่นฉบับ U อาจจะสอบรวมหรือแยกจากฉบับ X และฉบับ Y หรือเป็นส่วนหนึ่งของฉบับ X และเป็นส่วนหนึ่งของฉบับ Y หรืออาจแยกจากฉบับ X แต่เป็นส่วนหนึ่งของฉบับ Y หรืออาจจะกระจายข้อสอบของฉบับ U ปะปนกับฉบับ X และฉบับ Y ก็ได้ แต่ต้องสามารถหาคะแนนรวมของแต่ละฉบับได้ถูกต้อง

วิธีการประมาณค่าความเป็นไปได้สูงสุด (Maximum Likelihood) นี้ไม่จำกัดเฉพาะการเทียบแบบสอบ 2 ฉบับเท่านั้น แต่สามารถที่จะเทียบแบบสอบตั้งแต่ 3 ฉบับขึ้นไปได้ ในการบริหารการสอบแต่ละฉบับอาจจะสอบแยกกลุ่มกัน แต่ฉบับ U ซึ่งเป็นแบบสอบร่วมต้องสอบทุกกลุ่มแล้วการประมาณค่าเฉลี่ยและความแปรปรวนของแต่ละฉบับที่จะเทียบกันก็ใช้กลุ่มรวมสอบฉบับ U ซึ่งสมการ [19] - [22] สามารถใช้ได้หลายๆครั้งเท่ากับจำนวนแบบสอบที่จะนำมาเทียบกัน และอีกประการหนึ่ง ฉบับ U ก็ไม่จำกัดการวัดตัวแปรเดียว แต่อาจจะใช้ได้หลายๆ ตัวแปร

1.2.4.2 การเทียบมาตราเชิงเส้นตรงสำหรับกลุ่มตัวอย่างมีความสามารถต่างกัน

ก. แบบสอบที่มีความเที่ยงเท่ากัน

เมื่อกลุ่ม α และ β มีความสามรถแตกต่างกันมาก ข้อตกลงที่เป็นพื้นฐานของทฤษฎีการเลือกนั้นไม่มีความเหมาะสม กับสถานการณ์นี้ควรใช้กับสถานการณ์เบื้องต้นแบบอื่นต่อภายใต้ข้อจำกัดที่ว่า ฉบับ U มีลักษณะคู่ขนานกันทั้งฉบับ X และฉบับ Y นั่นคือ

(1) จุดตัดแกนของเส้นกระถดถอย (Regression Coefficient) ที่เกี่ยวกับคะแนนจริงของฉบับ X และฉบับ U โดยที่กลุ่ม ε เหมือนกับกลุ่ม α

$$M_{xt} - \frac{M_{xt}}{S_{ut}} M_{ut} = M_{x} - \frac{S_{x\alpha}}{S_{u}} M_{u\alpha} \qquad \text{.....[25]}$$

เมื่อ $S_x = S_x \sqrt{r_x}$

และ $S_u = \sqrt{S_u \sqrt{r_u}}$

(2) ความชันของเส้นสัมพันธ์สำหรับกลุ่ม t เหมือนกับกลุ่ม α

$$\frac{S_{xt}}{S_{ut}} = \frac{S_{x\alpha}}{S_{u\alpha}} \qquad \text{.....[26]}$$

(3) ความแปรปรวนของความคลาดเคลื่อนของการวัดจากฉบับ X กลุ่ม t เหมือนกับกลุ่ม α

$$S_{xt}^2(1 - r_{xxt}) = S_{x\alpha}^2(1 - r_{xx\alpha}) \qquad \text{.....[27]}$$

จากสมการ [25] ,[26] ,[27] จะได้

$$\hat{M}_{xt} = M_{x\alpha} + \frac{S_{xx}}{S_u}(M_{ut} - M_{ux}) \qquad \text{.....[28]}$$

$$\hat{S}_{xt}^2 = S_{x\alpha}^2 + \frac{S_{x\alpha}^2}{S_{u\alpha}}(S_{ut}^2 - S_{u\alpha}^2) \qquad \text{.....[29]}$$

ด้วยข้อตกลงเบื้องต้นแบบเดียวกันสำหรับฉบับ X และฉบับ Y จะได้

$$\hat{M}_{yt} = M_{y\beta} + \frac{S_{yy}}{S_u}(S_{ut}^2 - S_{u\beta}^2) \qquad \text{.....[30]}$$

$$\hat{S}_{yt}^2 = S_{y\beta}^2 + \frac{S_{y\beta}^2}{S_{u\beta}^2}(S_{ut}^2 - S_{u\beta}^2) \qquad \text{.....[31]}$$

สมการที่แสดงความสัมพันธ์ระหว่างคะแนนฉบับ X และฉบับ Y คือ $Y = AX + B$

เมื่อ $$A = \frac{\hat{S}_{yt}}{\hat{S}xt} \qquad \text{.....[32]}$$

$$B = \hat{M}_{yt} - A\hat{M}_{xt} \qquad \text{.....[33]}$$

เพื่อเป็นการง่ายในการคำนวณ แองกอฟฟ์ (Angoff) ได้เสนอข้อมูลที่ได้จากการทดลองเทียบมาตราโดยใช้การประมาณค่าอัตราส่วนของค่าเบี่ยงเบนมาตรฐานของคะแนนจริง นั่นคือ $\frac{\tilde{S}_x}{\tilde{S}_u} = n$ ซึ่งเป็นอัตราส่วนของความยาวของแบบสอบฉบับ X กับฉบับ U เมื่อฉบับ U

เป็นข้อสอบร่วมในฉบับ X และฉบับ U คู่ขนานกับฉบับ X

$$n = \frac{S_x}{r_{xu} S_u} = \frac{1}{b_{ux}}$$

เมื่อฉบับ U เป็นการสอบแยกจากฉบับ X แล้ว

$$n = \frac{\left(Sx^2 + Sux\right)}{\left(Su^2 + Sux\right)}$$

ซึ่งเมื่อใช้กับฉบับ Y และฉบับ U ก็ปรับสมการในลักษณะเดียวกัน

ข. แบบสอบที่มีความเที่ยงไม่เท่ากัน

เมื่อฉบับ U สอบแยกจากฉบับ X และฉบับ Y จะได้ $Y = AX + B$

เมื่อ

$$A = \frac{b_{y\beta} r_{yu\alpha}}{b_{xu\alpha} r_{uy\beta}} \qquad \text{.....[34]}$$

$$B = M_{y\beta} - A M_{x\alpha} + \frac{b_{yu\beta}}{r_{uu\beta}}\left(M_{u\alpha} - M_{u\beta}\right) \qquad \text{.....[35]}$$

เมื่อฉบับ U สอบรวมกับฉบับ X และฉบับ Y จะได้

$$A = \frac{b_{ux\alpha}}{b_{uy\beta}} \qquad \text{.....[36]}$$

$$B = M_{y\beta} - A M_{x\alpha} + \left[\frac{M_{u\alpha} - M_{u\beta}}{b_{uy\beta}}\right] \qquad \text{.....[37]}$$

1.2.5 Other Methods Involving Score Data

1.2.5.1 แบบสอบฉบับ X และฉบับ Yเทียบไปสู่แบบสอบร่วม

วิธีการนี้เกี่ยวข้องกับการบริหารแบบสอบร่วมฉบับ U ซึ่งอาจจะสอบฉบับ U ก่อนแล้วตามด้วยฉบับ X และฉบับ Y หรืออาจจะสอบฉบับ X ก่อนแล้วตามด้วยฉบับ U ถ้าหากฉบับ X และฉบับ Y เป็นแบบสอบที่มีลักษณะคู่ขนานกันก็มีเหตุผลที่จะเชื่อถือได้ว่า

แต่ละฉบับมีผลต่อฉบับ U เหมือนกัน เมื่อฉบับ U สอบภายหลัง หรือฉบับ U มีผลต่อฉบับ X และฉบับ Y เหมือนกัน เมื่อฉบับ U สอบก่อนฉบับ X และฉบับ Y การเทียบมาตราโดยวิธีนี้ คือ เทียบจากฉบับ X ไปสู่ฉบับ U และเทียบจากฉบับ Y ไปสู่ฉบับ U ซึ่งคะแนนที่เทียบแล้วของทั้งสองฉบับที่ระดับคะแนนฉบับ U เดียวกัน ถือว่าเท่าเทียมกัน ดังนั้น

ถ้า $X = A_{xu}U + B_{xu}$

เมื่อ $A_{xu} = \frac{S_{x\alpha}}{S_{y\beta}}$ ..... [38]

$B_{xu} = M_{x\alpha} - A_{xu}M_{u\alpha}$ ..... [39]

และ ถ้า $Y = A_{yu}U + B_{yu}$

เมื่อ $A_{yu} = \frac{S_{y\beta}}{S_{u\beta}}$ ..... [40]

$B_{yu} = M_{y\beta} - A_{yu}M_{u\beta}$ ..... [41]

แล้วจะได้ $Y = A_{yx}X + B_{yx}$

เมื่อ $A_{yx} = \frac{S_{yu}}{S_{xu}}$ ..... [42]

$B_{yx} = B_{yu} - A_{yu}B_{xu}$ ..... [43]

ในการที่ดูถึงความเหมาะสมของสมการแปลง $Y = A_{yx}X + B_{yx}$ ฉบับ U ต้องมีลักษณะคู่ขนานกับฉบับ X และ Y โดยที่มีอิสระในการเลือกกลุ่ม α และ β แต่ถ้าฉบับ U ไม่มีลักษณะคู่ขนานกับฉบับ X และ Y แล้ว กลุ่ม α และ β ต้องมาจากสุ่มประชากรกลุ่มเดียวกัน

1.2.5.2 แบบสอบร่วมฉบับ U เป็นตัวทำนาย (Predict) ฉบับ X และ Y

วิธีการบริหารการสอบคือ กลุ่ม α สอบฉบับ X กลุ่ม β สอบฉบับ Y และทั้งสองกลุ่มสอบฉบับ U ซึ่งวิธีการในการกำหนดคะแนนที่เทียบเคียงกันทั้ง 2 ฉบับ โดยการทำนายของคะแนนฉบับ U เดียวกัน ดังนั้น

ถ้า $\hat{X} = b_{xu}U + D_{xu}$

เมื่อ $b_{xu} = r_{xu\alpha}\frac{S_{x\alpha}}{S_{y\alpha}}$

$D_{xu} = M_{x\alpha} - b_{xu}M_{u\alpha}$

และถ้า $\hat{Y} = b_{yu}U + D_{yu}$

เมื่อ $b_{yu} = r_{yu\beta}\frac{S_{y\beta}}{S_{u\beta}}$

$$D_{yu} = M_{y\beta} - b_{yu} M_{u\beta}$$

แล้วจะได้ $Y = AX + B$

เมื่อ $$A = \frac{b_{yu}}{b_{xu}}$$ .....[44]

$$B = \frac{D_{yu}}{AD_{xu}}$$ .....[45]

1.2.5.3 แบบสอบฉบับ X และฉบับ Y เป็นตัวทำนายแบบสอบร่วม กลุ่ม α สอบฉบับ X กลุ่ม β สอบฉบับ Y และทั้งสองกลุ่มสอบฉบับ U

ถ้า $U = b_{xu} X + D_{xu}$

เมื่อ $$b_{xu} = r_{xu\alpha} \frac{S_{y\alpha}}{S_{x\alpha}}$$

$$D_{xu} = M_{u\alpha} - b_{ux} M_{x\alpha}$$

และถ้า $U = b_{uy} y + D_{uy}$

เมื่อ $$b_{uy} = r_{uy\beta} \frac{S_{u\beta}}{S_{y\beta}}$$

$$D_{uy} = M_{u\beta} - b_{uy\beta} M_{y\beta}$$

แล้วจะได้ $Y = AX + B$

เมื่อ $$A = \frac{b_{ux}}{b_{uy}}$$ .....[45]

$$B = \frac{D_{ux} - D_{uy}}{b_{uy}}$$ .....[46]

1.2.6 การเทียบมาตราที่ขึ้นกับลักษณะของข้อสอบ

1.2.6.1 วิธีการของเธอร์สโตน (Thurstone's Absolute Scaling Method) เธอร์สโตน (Therstone, 1925 อ้างจาก วรเนตร พิชิตรเกริกพล, 2537 : 34) ได้เสนอวิธีการโดยมีสถานการณ์ คือ กลุ่ม α สอบฉบับ X กลุ่ม β สอบฉบับ Y แบบสอบฉบับ X และแบบสอบฉบับ Y มีข้อสอบชุดหนึ่งซึ่งมีลักษณะร่วมกันคือ ค่าความยาก (p) ซึ่งต่างจากค่า p ทั่วๆไป คือ แปลงให้มีลักษณะเช่นเดียวกับการกระจายปกติ ข้อตกลงเบื้องต้นเช่นเดียวกับวิธีการอื่นๆ คือ ฉบับ X

และฉบับ Y มีลักษณะคู่ขนานกัน ดังนั้นจึงสามารถแปลงเป็นสเกลร่วมได้ และข้อตกลงอีกอย่างหนึ่งคือ กลุ่ม α และกลุ่ม β ลักษณะการกระจายเป็นแบบปกติ

1.2.6.2. วิธีการเทียบมาตราของแฟน (Swwine - Fan Method of Equating)

วิธีการนี้คล้ายกับของเธอร์สโตน ซึ่งแฟน (Fan) จัดดำเนินการโดยแบบสอบมีกลุ่มข้อสอบร่วมกันเหมือนกันทั้งสองฉบับ (ฉบับ X และ ฉบับ Y) กลุ่ม α สอบฉบับ X กลุ่ม β สอบฉบับ Y จะมีการนำค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานจากคะแนนดิบของข้อสอบที่ทำถูกมาใช้ในการคำนวณ

1.3 การเทียบมาตรารูปแบบการวิเคราะห์ตัวประกอบ (Equating Using The Confirmatory Factor Analysis Model)

การเทียบมาตราโดยวิธีนี้ใช้ความเป็นไปได้ของการวิเคราะห์ตัวประกอบสูงสุด (Maximum Likelihood Factor Analysis) ในการประมาณค่าพารามิเตอร์ โดยการเทียบมาตราของแบบสอบทั้ง 2 ฉบับ ที่ใช้แบบสอบร่วม ดังนี้

ให้กลุ่ม α สอบแบบสอบฉบับ X และ U_1

ให้กลุ่ม β สอบแบบสอบฉบับ Y และ U_2

ถึงแม้ว่า U_1 จะมีโครงสร้างที่คู่ขนานกับ U_2 แต่ก็ยังแตกต่างกันในหน่วยของการวัด จุดเริ่มต้นของสเกลและความแปรปรวนที่ไม่เท่ากัน ฉะนั้นความจำเป็นและเงื่อนไขสำหรับที่จะทำให้ U_1 และ U_2 เท่าเทียมกันขึ้นอยู่กับสององค์ประกอบร่วม คือ รูปแบบของสององค์ประกอบที่เหมาะสมกับกลุ่มตัวอย่างทั้งสอง การเทียบมาตรารูปแบบการวิเคราะห์ตัวประกอบ มีขั้นตอนดังนี้

1.3.1 รวบรวมข้อมูลที่ได้จากแบบสอบทั้ง 2 ฉบับ

1.3.2 นำข้อมูลไปวิเคราะห์หาองค์ประกอบสำคัญ เพื่อหาน้ำหนักองค์ประกอบและค่าความร่วมกัน

1.3.3 ปรับคะแนนให้อยู่ในสเกลเดียวกันของแบบสอบทั้ง 2 ฉบับ โดยใช้สูตรของร็อค (Rock, 1982 : 252 อ้างจาก เรวดี อินทสระ, 2530 : 40) ดังนี้

$$f = \frac{X_i - V_i}{\Lambda_i} \qquad \text{.....[47]}$$

เมื่อ f คือ ค่าองค์ประกอบร่วม (Common Factor) ในแต่ละข้อ

x_i คือ ค่าความน่าจะเป็นในการตอบถูกของข้อสอบ ข้อที่ i

v_i คือ ค่าความร่วมกันในแฟกเตอร์ของข้อสอบข้อที่ i

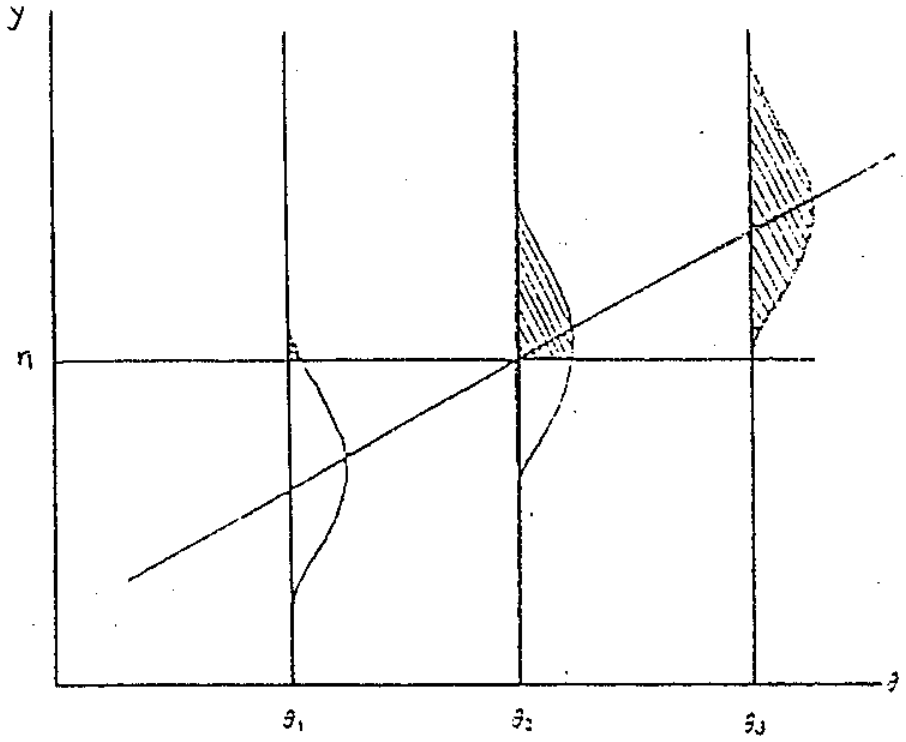
Λ_i คือ ค่าน้ำหนัก (Loading) ของข้อสอบข้อที่ i

2. การเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบ (Item Response Theory Model)

ก่อนที่จะกล่าวถึงการเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบ (Item Response Theory Model) นั้น ผู้วิจัยขอกล่าวถึงหลักการของทฤษฎีการตอบสนองข้อสอบก่อนเพื่อให้เข้าใจในขั้นต่อไป

2.1 หลักการของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างคุณลักษณะภายใน หรือความสามารถที่มีอยู่ภายในตัวบุคคลกับพฤติกรรมการตอบข้อสอบของบุคคลนั้นว่ามีโอกาสตอบข้อสอบถูกมากน้อยเพียงไร ทฤษฎีนี้มีพื้นฐานความเชื่อว่า พฤติกรรมการตอบสนองข้อสอบของผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายใน หรือความสามารถที่อยู่ภายในตัวบุคคล ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง ทฤษฎีนี้พยายามที่จะอธิบายความสัมพันธ์ระหว่างลักษณะภายในตัวบุคคลกับพฤติกรรมที่บุคคลตอบสนองต่อข้อสอบ การอธิบายความสัมพันธ์ดังกล่าวแสดงออกมาในรูปแบบทางคณิตศาสตร์ โดยให้คะแนนที่ได้รับจากการตอบข้อสอบ (y) แทนพฤติกรรมการตอบสนองต่อข้อสอบ ให้ θ แทนลักษณะหรือความสามารถในตัวบุคคล และ r_i เป็นเกณฑ์ในการทำข้อสอบนั้นว่าถูกหรือผิด เป็นเกณฑ์ที่บอกว่า y แค่ไหนจึงจะทำข้อสอบข้อ i ได้ถูก ดังนั้นถ้า $y > r_i$ แสดงว่าผู้สอบทำข้อสอบข้อ i ได้ถูก และถ้า $y < r_i$ แสดงว่าทำข้อสอบข้อ i ผิด ความสัมพันธ์ระหว่างความสามารถ (θ) กับพฤติกรรมการตอบสนองข้อสอบ (y) แสดงได้ดังภาพที่ 4



ภาพที่ 3 : ความสัมพันธ์ระหว่างความสามารถกับพฤติกรรมการตอบสนองข้อสอบ

จากภาพแสดงว่า ถ้านำโอกาสที่จะตอบถูก (พื้นที่ส่วนที่แรเงา) ในระดับความสามารถต่าง ๆ มาเขียนกราฟใหม่ จะได้โค้งลักษณะของข้อสอบ (ICC) เป็นรูปต่าง ๆ ทั้งนี้ขึ้นอยู่กับฟังก์ชันทางคณิตศาสตร์ และจำนวนพารามิเตอร์ที่จะใช้อธิบายฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างระดับความสามารถกับพฤติกรรมการตอบสนองข้อสอบเรียกว่า ฟังก์ชันการตอบสนองข้อสอบ (Item Characteristic Function) ซึ่งสามารถเขียนฟังก์ชันได้ดังนี้

$$P_i(\theta) = \text{Prob}(U_i = 1|\theta) \quad \text{เมื่อ} \quad U_i = 0 \, , \, 1$$

จากฟังก์ชันข้างต้นนี้หมายถึงโอกาสที่ผู้สอบซึ่งมีความสามารถ θ จะตอบคำถามข้อ i ได้ถูกต้อง

2.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

2.2.1. มิติของลาเท้นท์สเปส (Dimensionality of Latent Space)

มิติของลาเท้นท์สเปส หมายถึง ลักษณะหรือความสามารถที่เป็นตัวกำหนด

พฤติกรรมการตอบสนองข้อสอบข้อใดข้อหนึ่งอาจมีได้หลายลักษณะ ลักษณะทั้งหมดนั้นรวมเรียกว่า ลาเท้นท์สเปส (Latent Space) จำนวนลักษณะทั้งหมดในลาเท้นท์สเปสก็คือ มิติของลาเท้นท์สเปส ในทฤษฎีการตอบสนองข้อสอบโดยทั่วไปมักจะถือว่า ลักษณะหรือความสามารถที่เป็นตัวกำหนดพฤติกรรมการตอบสนองข้อสอบมีเพียงลักษณะเดียว และยอมรับเป็นข้อตกลงเบื้องต้นที่สำคัญประการหนึ่งของรูปแบบหลายแบบในทฤษฎีนี้ เรียกว่า Unidimensionality Assumption นั่นคือ ข้อสอบในแบบสอบมีลักษณะเป็นเอกพันธ์ (Homogeneous) การถือว่าลาเท้นท์สเปสที่สนใจมีเพียงมิติเดียวจะช่วยให้รูปแบบของทฤษฎีมีความซับซ้อนน้อยลง และง่ายแก่การแปลความหมายของคะแนนจากแบบสอบ วิธีการตรวจสอบว่าข้อสอบชุดนี้วัดความสามารถหรือคุณลักษณะเดียวกันมีหลายวิธี ได้แก่ การวิเคราะห์ตัวประกอบ (Factor Analysis) และการเลือกข้อสอบที่มีค่าอำนาจจำแนกสูง เนื่องจากข้อสอบที่มีค่าอำนาจจำแนกสูงจะมีสหสัมพันธ์กับคะแนนของแบบสอบทั้งหมดสูง ยูริ (Urry, 1981 อ้างจาก จิราพร เนียมสุวรรณ, 2533 : 29) เสนอไว้ว่า ควรเลือกข้อสอบที่มีค่าอำนาจจำแนกตั้งแต่ 0.80 ขึ้นไป

2.2.2. ความเป็นอิสระในการตอบข้อสอบ(Local independence)

ความเป็นอิสระในการตอบข้อสอบ หมายถึง การตอบข้อสอบข้อต่าง ๆ ในแบบสอบของผู้สอบนั้นมีความเป็นอิสระซึ่งข้อตกลงเกี่ยวกับความเป็นอิสระในการตอบสนองข้อสอบนั้นมี 2 ลักษณะ คือ

2.2.2.1. ความเป็นอิสระทางสถิติ (Statistically indenpendence) กล่าวคือ ข้อสอบแต่ละข้อเป็นอิสระไม่เกี่ยวข้องกัน ดังนั้นคำตอบในแต่ละข้อของแต่ละคนเป็นอิสระต่อกัน แต่รวมกันแล้วจะวัดคุณลักษณะเดียวเท่านั้น

2.2.2.2. ความเป็นอิสระจากตำแหน่ง (Uncorrelated Independence) กล่าวคือ ข้อสอบแต่ละข้อจะปรากฏอยู่ในตำแหน่งใดของข้อสอบก็ได้ จะไม่มีผลต่อการตอบข้อสอบ ลอร์ด (Lord, 1980 : 19) กล่าวว่า เมื่อแบบสอบมีคุณสมบัติของการวัดเพียงมิติเดียวแล้วก็จะมีความเป็นอิสระในการตอบข้อสอบ ดังนั้นการตรวจสอบความเป็นอิสระในการตอบข้อสอบก็ใช้การวิเคราะห์ตัวประกอบได้เหมือนกัน

2.2.2.3. โค้งลักษณะของข้อสอบ (Item Characteristi Curve : ICC) โค้งลักษณะของข้อสอบ คือ ฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบข้อหนึ่ง ได้ถูกต้องกับระดับความสามารถที่วัดด้วยแบบสอบที่มีข้อสอบข้อนั้นอยู่ กล่าวอีกนัยหนึ่ง โค้งลักษณะของข้อสอบเป็นฟังก์ชันดีกรีที่ไม่ใช่เส้นตรง (Non Linear) ของคะแนนข้อสอบบนความสามารถหรือคุณลักษณะที่วัดด้วยแบบสอบนั้น

โค้งลักษณะของข้อสอบมีหลายรูปแบบขึ้นอยู่กับความเชื่อเกี่ยวกับความสัมพันธ์ระหว่างระดับความสามารถกับความน่าจะเป็นในการตอบถูก ทำให้มีรูปแบบที่ใช้อธิบายความสัมพันธ์ดังกล่าวในทฤษฎีการตอบสนองข้อสอบหลายรูปแบบด้วยกันในที่นี้จะกล่าวถึงรูปแบบ โลจีสติก (Logistic Model) เท่านั้นผู้ที่มีบทบาทสำคัญในการพัฒนารูปแบบนี้คือ เบอร์นบอม(Birnbaum) (Hambleton and Swaminathan, 1985 : 36) ซึ่งแสดงความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบถูกกับระดับความสามารถในรูปของฟังก์ชันการแจกแจงสะสมแบบโลจีสต์ (Logistic Cumulation Distibution Function) มีเส้นโค้งลักษณะของข้อสอบเป็นรูปตัว S เช่นเดียวกับรูปแบบปกติสะสม (Normal Ogive Model) และเมื่อปรับค่าตัวแปรด้วย Scaling Factor ซึ่งมีค่าเท่ากับ 1.7 แล้วโค้งลักษณะของข้อสอบของรูปแบบทั้งสองมีความแตกต่างกันไม่เกิน .01 สำหรับทุก ๆ ค่าของความสามารถ แต่ในแง่ของการคำนวณ Logistic Model มีความง่ายและสะดวกกว่ามาก จึงเป็นที่นิยมในการนำไปใช้ปฏิบัติงานจริง ซึ่งแบ่งออกเป็น 4 รูปแบบตามจำนวนพารามิเตอร์ของโค้งลักษณะของข้อสอบ (สุพัฒน์ สุกมลสันต์,2538 : 62)

2.3 รูปแบบโลจีสติก

2.3.1 รูปแบบโลจีสติกที่มีพารามิเตอร์หนึ่งตัว (One - Parameter Logistic Model)

รูปแบบนี้เบิร์นบอม (Birnbaum) พัฒนาขึ้นในปี ค.ศ. 1968 ซึ่งบังเอิญตรงกับแบบราสช์ (Rasch, 1980 : 2) โดยราสช์นำมาใช้ในปี ค.ศ. 1960 เป็นรูปแบบที่อธิบายคุณลักษณะของข้อสอบด้วยค่าพารามิเตอร์เพียงตัวเดียว คือ ระดับความยาก (β) รูปแบบนี้เชื่อว่าความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบได้ถูกหรือไม่นั้น ขึ้นอยู่กับระดับความสามารถของผู้สอบ (θ) และระดับความยากของข้อสอบ ดังนั้นจึงถือว่าค่าการเดา (c) เป็นศูนย์ และค่าอำนาจจำแนก (α หรือแทนด้วย a) ของข้อสอบจะคงที่ทั้งฉบับเขียนเป็นฟังก์ชันได้ดังนี้

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}} \quad ; \quad i = 1, 2, \ldots, n \quad \ldots\ldots\ldots\ldots [48]$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้ตอบที่มีระดับความสามารถ θ ตอบข้อสอบข้อ i ถูกต้อง

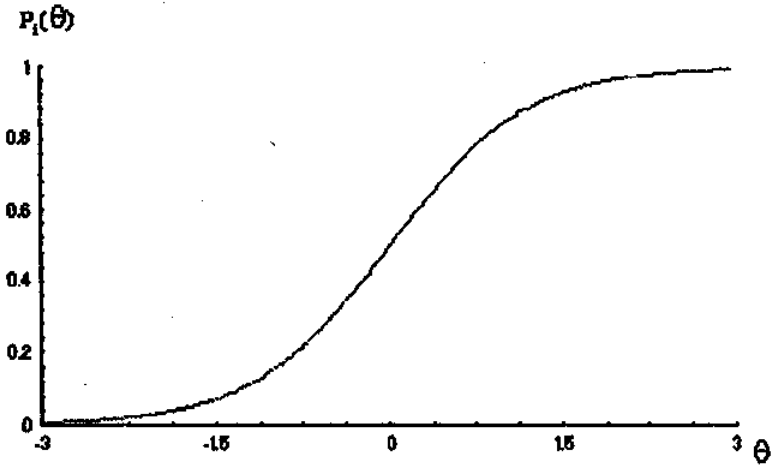
θ คือ ระดับความสามารถของผู้สอบที่ประมาณได้จากคะแนนรวมของการตอบแบบสอบโดยการปรับให้เป็นคะแนนที่มีค่าเฉลี่ยเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 โดยมีพิสัย

การกระจายอยู่ระหว่าง - ∞ ถึง +∞ แต่ในทางปฏิบัติ ระดับความสามารถจะมีค่าอยู่ระหว่าง -3 ถึง +3 ซึ่งระดับความสามารถ -3 หมายถึงผู้สอบมีระดับความสามารถต่ำมาก และระดับความสามารถ +3 หมายถึงผู้สอบมีความสามารถสูงมาก

b_i คือ ค่าความยากของข้อสอบข้อ i ที่แสดงถึงระดับความสามารถที่จุดโค้งลักษณะของข้อสอบมีความชันมากที่สุด โดยทั่วไปมีพิสัยของค่าความยากของข้อสอบอยู่ระหว่าง -2 ถึง +2 ซึ่งค่าที่ใกล้ -2 หมายถึงข้อสอบที่ง่ายมาก และค่าที่ใกล้ +2 หมายถึงข้อสอบที่ยากมาก

D คือ Scaling Factor มีค่าเท่ากับ 1.7

e คือ ค่าคงที่มีค่าเท่ากับ 2.71828...



ภาพที่ 4 : โค้งลักษณะของข้อสอบที่มีหนึ่งพารามิเตอร์

2.3.2. รูปแบบโลจิสติกที่มีพารามิเตอร์สองตัว (Two - Parameter Logistic Model)

เป็นรูปแบบที่เบิร์นบอม (Birnbaum 1957 , 1958a , 1958b , 1968) ได้เสนอไว้แทนรูปแบบโค้งปกติสะสม (Normal Ogive Model) เนื่องมาจากรูปแบบโลจิสติก เป็นรูปแบบเชิงคณิตศาสตร์ที่นำมาใช้ในทางปฏิบัติได้สะดวกกว่า รูปแบบนี้ถือว่าไม่มีการเดา ดังสมการ

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad \text{หรือ} \quad \text{.....[49]}$$

$$P_i(\theta) = \left[1 + e^{-Da_i(\theta - b_i)}\right]^{-1} \quad i = 1, 2, \ldots, n \quad \ldots\ldots\ldots\ldots[50]$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้ตอบที่มีระดับความสามารถ θ ตอบข้อสอบข้อ i ถูกต้อง

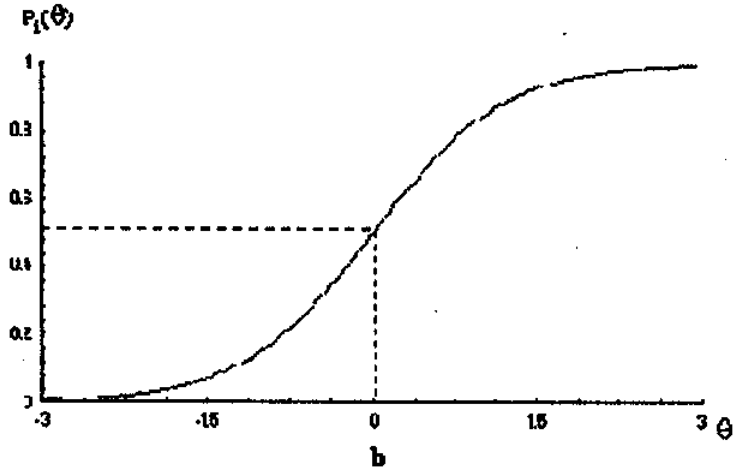
θ คือ ระดับความสามารถของผู้สอบที่ประมาณได้จากคะแนนรวมของการตอบแบบสอบโดยการปรับให้เป็นคะแนนที่มีค่าเฉลี่ยเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 โดยมีพิสัยการกระจายอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติระดับความสามารถจะมีค่าอยู่ระหว่าง -3 ถึง +3 ซึ่งระดับความสามารถ -3 หมายถึงผู้สอบมีระดับความสามารถต่ำมาก และระดับความสามารถ +3 หมายถึงผู้สอบมีความสามารถสูงมาก

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อ i ที่เป็นสัดส่วนโดยตรงกับความชันของโค้งลักษณะของข้อสอบ ณ จุดเปลี่ยนโค้ง หรือที่จุด $\theta = b_i$ โดยทั่วไปพิสัยของค่าอำนาจจำแนกของข้อสอบอยู่ระหว่าง 0 ถึง 2

b_i คือ ค่าความยากของข้อสอบข้อ i ที่แสดงถึงระดับความสามารถที่จุดโค้งลักษณะของข้อสอบมีความชันมากที่สุด โดยทั่วไปมีพิสัยของค่าความยากของข้อสอบอยู่ระหว่าง -2 ถึง +2 ซึ่งค่าที่ใกล้ -2 หมายถึงข้อสอบที่ง่ายมาก และค่าที่ใกล้ +2 หมายถึงข้อสอบที่ยากมาก

D คือ Scaling Factor มีค่าเท่ากับ 1.7

e คือ ค่าคงที่มีค่าเท่ากับ 2.71828...



ภาพที่ 5 : โค้งลักษณะของข้อสอบที่มีสองพารามิเตอร์

2.3.3. รูปแบบโลจีสติกที่มีสามพารามิเตอร์ (Three - Parameter Logist Model)
เป็นรูปแบบที่มีพารามิเตอร์อธิบายโค้งลักษณะของข้อสอบสามตัว ซึ่งมีรูปสมการดังนี้

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad ; (i = 1,2,...,n) \quad .....[51]$$

หรือ $$P_i(\theta) = c_i + (1 - c_i)\left[1 + e^{-Da_i(\theta - b_i)}\right]^{-1} \quad ; (i = 1,2,...,n) \quad .....[52]$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้ตอบที่มีระดับความสามารถ θ ตอบข้อสอบข้อ i ถูกต้อง

θ คือ ระดับความสามารถของผู้สอบที่ประมาณได้จากคะแนนรวมของการตอบแบบสอบโดยการปรับให้เป็นคะแนนที่มีค่าเฉลี่ยเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 โดยมีพิสัยการกระจายอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติระดับความสามารถจะมีค่าอยู่ระหว่าง -3 ถึง +3 ซึ่งระดับความสามารถ -3 หมายถึงผู้สอบมีระดับความสามารถต่ำมาก และระดับความสามารถ +3 หมายถึงผู้สอบมีความสามารถสูงมาก

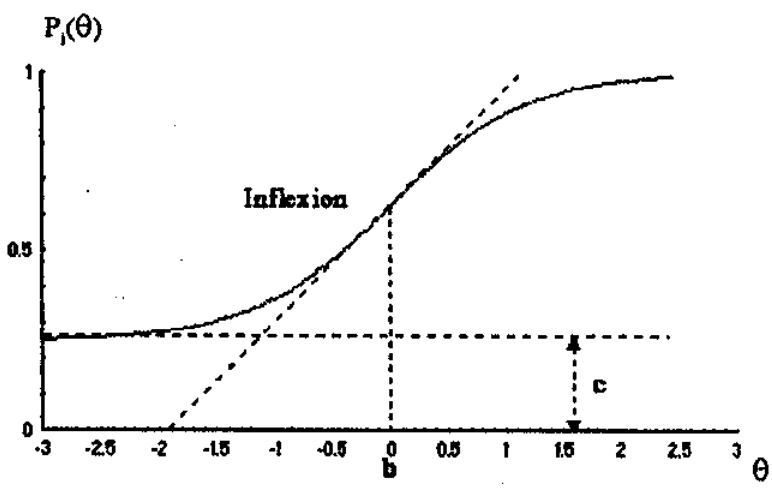
a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อ i ที่เป็นสัดส่วนโดยตรงกับความชันของโค้งลักษณะของข้อสอบ ณ จุดเปลี่ยนโค้ง หรือที่จุด $\theta = b_i$ โดยทั่วไปพิสัยของค่าอำนาจจำแนกของข้อสอบอยู่ระหว่าง 0 ถึง 2

b_i คือ ค่าความยากของข้อสอบข้อ i ที่แสดงถึงระดับความสามารถที่จุดโค้งลักษณะของข้อสอบมีความชันมากที่สุด โดยทั่วไปมีพิสัยของค่าความยากของข้อสอบอยู่ระหว่าง -2 ถึง +2 ซึ่งค่าที่ใกล้ -2 หมายถึงข้อสอบที่ง่ายมาก และค่าที่ใกล้ +2 หมายถึงข้อสอบที่ยากมาก

c_i คือ ค่าการเดาของข้อสอบข้อ i ซึ่งเป็นค่าความน่าจะเป็นที่ผู้สอบมีความสามารถต่ำมากจะมีความน่าจะเป็นในการตอบข้อสอบข้อ i ได้ถูกหรือเป็นค่ากำกับโค้งที่ต่ำสุด (Lower Assymptote) ของโค้งลักษณะข้อสอบมีค่า 0 ถึง 1

D คือ Scaling Factor มีค่าเท่ากับ 1.7

e คือ ค่าคงที่มีค่าเท่ากับ 2.71828...



ภาพที่ 6 : โค้งลักษณะของข้อสอบที่มีสามพารามิเตอร์

ลักษณะเฉพาะของโค้งของข้อสอบเมื่อระดับความสามารถของผู้สอบ (θ) เปลี่ยนแปลงไปจะมีรูปแบบแตกต่างกัน รูปแบบต่าง ๆ เหล่านี้เปลี่ยนแปลงไปตามฟังก์ชันการตอบสนองรายข้อ (Item Response Function) หรือที่นิยมเรียกกันทั่วไปว่า โค้งลักษณะเฉพาะของข้อสอบ รูปแบบนี้มีลักษณะเฉพาะดังต่อไปนี้ (Warm, 1978 : 42)

2.3.3.1 เส้นกำกับบน (Upper Assymptote) จะมีค่าสูงที่สุดได้เกือบเท่ากับ 1.0 แต่ไม่เท่ากับ 1.0 เพราะความน่าจะเป็นในการตอบข้อสอบได้ถูกมีค่าไม่เท่ากับ 1.0 หรือ 100% เพราะไม่มีอะไรแน่นอนที่สุด

2.3.3.2 เมื่อผู้สอบมีความสามารถมากขึ้น ความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องมีมากขึ้น เส้นโค้งลักษณะเฉพาะของข้อสอบจะสูงขึ้นเรื่อย ๆ เมื่อเคลื่อนจากซ้ายไปขวา และจุดที่โค้งมีความชันมากที่สุด คือ จุดที่ดีที่สุดในการแบ่งผู้สอบที่มีความสามารถต่างกัน

2.3.3.3 เส้นกำกับล่าง (Lower Assymptote) จะมีค่าได้น้อยที่สุดเกือบ 0.0 แต่ไม่เท่ากับ 0.0 เพราะการสอบย่อมมีการเดาบ้างไม่มากก็น้อย จะไม่มีการเดาเลยย่อมเป็นไปไม่ได้

2.3.4. รูปแบบโลจิสติกที่มีพารามิเตอร์สี่ตัว (Four - Parameter Logistic Model)

รูปแบบนี้เป็นรูปแบบล่าสุดที่ แมกโดแนลด์ บาร์ตัน และลอร์ด (Mc Donald, Barton and Lord) คิดใช้ขึ้นในปีค.ศ. 1967 และ 1981 (Hambleton and Swaminathan 1985 : 48) เพื่อนำค่าพารามิเตอร์เกี่ยวกับความคลาดเคลื่อนอื่น ๆ ในการสอบมาคำนวณในรูปแบบด้วย เช่น ความสะเพร่า ความรอบรู้มากกว่าเนื้อหาที่สอบ (จนคิดว่าข้อสอบที่ถูกต้องยังไม่ถูกต้องมากเพียงพอ) เป็นต้น รูปแบบนี้เขียนเป็นฟังก์ชันได้ดังนี้

$$P_i(\theta) = c_i + (\gamma_i - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \; ; i = 1, 2, \ldots, n \quad .....[53]$$

เมื่อ γ คือ ความสะเพร่าและความคลาดเคลื่อนอื่น ๆ มีค่าต่ำกว่า 1 เล็กน้อย

แต่อย่างไรก็ตามแนวคิดของรูปแบบดังกล่าวนี้ ในปัจจุบันยังไม่มีการนำมาใช้ในทางปฏิบัติ

2.4. การเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบ

การเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบ มีทั้งการเทียบมาตราโดยใช้คะแนนจริงและการเทียบมาตราโดยใช้คะแนนสังเกต ทั้งสองวิธีนี้มีทั้งข้อได้เปรียบและเสียเปรียบกล่าวคือ การเทียบมาตราด้วยคะแนนจริงไม่สามารถอธิบายคะแนนที่อยู่ต่ำกว่าระดับการเดาได้ โดยจะให้ความหมายของคะแนนสมมูลเฉพาะคะแนนที่อยู่เหนือค่าเฉลี่ยการเดา (Lord, 1980 : 199) แม้จะเป็นการเทียบมาตราโดยใช้คะแนนจริงแต่ยังเป็นคะแนนที่ประมาณได้จากสูตรในการคำนวณ ดังนั้นจึงยังคงมีความคลาดเคลื่อนอยู่ ส่วนการเทียบมาตราโดยใช้คะแนนสังเกตนั้นเป็นการเทียบคะแนนโดยประมาณ ซึ่งอธิบายคะแนนสมมูลจาก X และ Y ได้ครอบคลุมพิสัยที่สังเกตได้ ลอร์ด (Lord,1980 :198-203) ได้แบ่งรูปแบบการเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบออกเป็นหลายวิธี ดังนี้

2.4.1. การเทียบมาตราโดยใช้คะแนนจริง (True-Score Equating)

พิจารณาจากความสัมพันธ์เชิงคณิตศาสตร์ระหว่างความสามารถและจำนวนข้อที่ทำถูก ซึ่งจะได้คะแนนจริงจากแบบสอบ 2 ฉบับ ดังนี้

$$\xi = \xi(\theta) = \sum_{i=1}^{m} P_i(\theta_x) \quad \text{.....[54]}$$

$$\eta = \eta(\theta) = \sum_{j=1}^{n} P_j(\theta_y) \quad \text{.....[55]}$$

เมื่อ ξ คือ คะแนนจริงของผู้สอบที่ได้จากแบบสอบฉบับ X

η คือ คะแนนจริงของผู้สอบที่ได้จากแบบสอบฉบับ Y

m คือ จำนวนข้อของแบบสอบฉบับ X

n คือ จำนวนข้อของแบบสอบฉบับ Y

θ_x คือ ค่าความสามารถของผู้สอบที่ได้จากแบบสอบฉบับ X

θ_y คือ ค่าความสามารถของผู้สอบที่ได้จากแบบสอบฉบับ Y

ในทางปฏิบัติ เราจะหาค่า $P_i(\theta)$ และ $P_j(\theta)$ ได้จากการประมาณค่าพารามิเตอร์ของข้อสอบโดยเลือกว่าจะใช้ชนิดหนึ่งพารามิเตอร์ สองพารามิเตอร์ หรือสามพารามิเตอร์ จากนั้นนำคะแนนจริงของผลการสอบจากแบบสอบฉบับ X และ Y มาหาความสัมพันธ์กันโดยใช้ค่าความสามารถ (θ) ที่ระดับเดียวกัน

2.4.2. การเทียบมาตราโดยใช้คะแนนจริงด้วยแบบสอบร่วม (True-Score Equating with An Anchor Test)

การใช้แบบสอบร่วมทำได้ 2 กรณี คือ ใช้แบบสอบร่วมรวมเข้าเป็นชุดเดียวกับแบบสอบที่ต้องการเทียบเรียกว่า แบบสอบร่วมภายใน (Internal Anchor Test) ส่วนในกรณีที่จัดแยกเป็นชุดแบบสอบต่างหากจากแบบสอบที่ต้องการเทียบเรียกว่า แบบสอบร่วมภายนอก (External Anchor Test) แล้วหาค่าความสามารถ (θ) ของกลุ่มตัวอย่างที่ทำแบบสอบร่วมไปเทียบมาตราก้บแบบสอบฉบับ X และ Y

2.4.3. การเทียบมาตราโดยใช้คะแนนดิบด้วยแบบสอบร่วม (Raw- Score Equating with an Anchor)

ปัญหาของการเทียบมาตราโดยใช้คะแนนจริงคือ ไม่สามารถทราบคะแนนจริงของแต่ละคนได้ นอกจากใช้วิธีประมาณผลการสอบโดยสมการ

$$\xi = \sum_{i=1}^{n} P_i(\theta)$$

เมื่อ $p_i(\theta)$ คือ ความน่าจะเป็นที่ทำข้อสอบข้อที่ i ได้ถูกที่ระดับความสามารถ θ

n คือ จำนวนข้อสอบทั้งหมด

ซึ่งค่าที่ได้เป็นค่าประมาณเท่านั้น ยังไม่มีคุณสมบัติเป็นคะแนนจริง ดังนั้นการเทียบมาตราโดยใช้คะแนนจริงจึงคล้ายกับการเทียบมาตราโดยใช้คะแนนดิบชนิดใหม่นั่นเอง สำหรับการเทียบมาตราโดยใช้คะแนนดิบในข้อนี้อาศัยข้อมูลจากการตอบแบบสอบร่วม (Anchor Test)

ในการเทียบมาตราโดยวิธีนี้จะเริ่มด้วยการประมาณลักษณะการแจกแจงของความสามารถของกลุ่มผู้สอบรวม $\gamma(\theta)$ ซึ่งหมายถึง ผู้สอบทั้งหมดที่ทำแบบสอบร่วม การแจกแจง(θ) ในกลุ่มเป็นการประมาณการแจกแจงของ θ

การประมาณการแจกแจงของคะแนนดิบ X สำหรับกลุ่ม $\phi_x(x)$ ได้จากสมการ

$$\hat{\phi}_x(x) = \frac{1}{N}\sum_{a=1}^{N} \hat{\phi}_x\left(x|\hat{\theta}_a\right) \qquad \text{.....[56]}$$

เมื่อ a =1,,2,3,...,N คือ ผู้สอบแต่ละระดับ และ $\phi_x\left(x|\hat{\theta}_x\right)$ เป็น

การแจกแจงของคะแนนดิบ x สำหรับผู้ที่มีความสามารถ θ

ถ้า $\gamma(\theta)$ มีลักษณะต่อเนื่อง (Continuous) สามารถประมาณด้วยสมการ

$$\hat{\phi}_x(x) = \int_{-\alpha}^{\alpha} \phi_x(x|\theta_a)\gamma(\theta)d\theta \quad \text{.....[57]}$$

เมื่อ $\phi_x(x|\theta_a)$ ได้มาจากการประมาณค่าพารามิเตอร์โดยใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Model) ในแบบสอบฉบับ X และ Y

ในการศึกษานี้ ผู้วิจัยได้เลือกวิธีการเทียบมาตราโดยใช้ทฤษฎีการตอบสนองข้อสอบ ที่ใช้คะแนนจริงของแบบสอบ (True Score Equating) มาเทียบมาตราใช้ค่าความสามารถที่ θ เดียวกัน เนื่องจากในทางปฏิบัติแล้วการเทียบมาตราโดยใช้แบบสอบร่วมของแต่ละคนหรือแต่ละโรงเรียน มีความเป็นไปได้น้อยถึงแม้ว่าการเทียบมาตราโดยใช้แบบสอบร่วมจะช่วยลดความคลาดเคลื่อนได้ ผู้วิจัยจึงเชื่อว่าการเทียบมาตราโดยใช้ทฤษฎีการตอบสนองข้อสอบที่ใช้คะแนนจริงจากแบบสอบแต่ละฉบับมีความเหมาะสมมากกว่าการใช้แบบสอบร่วม

การประเมินคุณภาพของการเทียบมาตรา

การเทียบมาตรารูปแบบใดก็ตามจะมีคุณภาพดีที่สุดเมื่อคะแนนที่ได้จากแบบสอบที่ต้องการเป็นไปตามเงื่อนไขต่าง ๆ ที่กำหนดไว้ของแต่ละรูปแบบ แต่ในสถานการณ์จริงมีข้อจำกัดทำให้ไม่สามารถได้ข้อมูลตามเงื่อนไข ดังนั้นจึงจำเป็นจะต้องตรวจสอบความคลาดเคลื่อนมาตรฐานของการเทียบมาตราและความเพียงพอ (Adequacy) ของการเทียบมาตราหรือการประเมินคุณภาพของรูปแบบที่แตกต่างกัน ดังนี้

1. ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา

ความคลาดเคลื่อนมาตรฐานของการเทียบมาตราเป็นผลจากการใช้รูปแบบการเทียบมาตราต่างกัน เกิดจากการแปรผันของการสุ่ม (Sample Fluctuation) ของค่าประมาณที่ใช้ในรูปแบบการเทียบมาตรา สำหรับการเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบนั้นเกิดขึ้นจากการประมาณความสัมพันธ์ของการเทียบ η ไปยัง ξ ทั้งนี้เนื่องจากการประมาณคะแนนจริงของผู้สอบได้มาจากค่าประมาณประชากรรายข้อ ค่าเหล่านี้คือ ที่มาของความคลาดเคลื่อน

เชิงสุ่มของการเทียบมาตรา (Lord, 1982 : 2 อ้างจาก ภาวิณี ศรีสุขวัฒนานันท์,2529 : 98)

2. ความเพียงพอของการเทียบมาตรา

วิธีการเทียบมาตราของแต่ละวิธี ประกอบด้วยรูปแบบการเทียบมาตรา (Model) ซึ่งมีข้อกำหนดที่ว่าด้วยสมมติฐานเบื้องต้น (Assumption) ของแต่ละรูปแบบ และประกอบด้วยการออกแบบ (Designs) เพื่อจัดเก็บข้อมูลให้เป็นไปตามข้อตกลงต่างๆ ถ้าหากทุกอย่างเป็นไปตามเงื่อนไขเชื่อได้ว่าผลของการเทียบมาตราจะมีความถูกต้อง (Accurate) และความคมชัด (Precise) ทางทฤษฎี แต่ในความเป็นจริงการสอบมักไม่เป็นตามอุดมการณ์เพราะมีหลายสิ่งที่อยู่นอกเหนือการควบคุม ดังนั้นการเทียบมาตราจึงมีการตรวจสอบความเพียงพอของรูปแบบในการเทียบมาตรา ซึ่งมีผู้เสนอไว้หลายวิธี แต่ในที่นี้ผู้วิจัยได้ใช้รูปแบบการประเมินความเพียงพอของการเทียบมาตราซึ่งเป็นการประเมินคุณภาพโดยใช้เกณฑ์ภายนอกตามแนวของ ภาวิณี ศรีสุขวัฒนานันท์ (2529,102) โดยใช้ดัชนีความแตกต่างมาตรฐาน (Index of standard Discripancy) ซึ่งดัดแปลงจากสูตรของโคเลนได้ใช้ข้อมูลคะแนนจากผู้สอบเองเป็นเกณฑ์ในการหาความแตกต่าง ข้อมูลเหล่านี้ได้มาจากการออกแบบด้วยการใช้กลุ่มสอบทานผล ซึ่งผู้สอบในกลุ่มตัวอย่างนี้ได้รับการสอบด้วยแบบสอบทั้งสองชุด ผู้วิจัยได้เลือกวิธีการประเมินความเพียงพอในการเทียบมาตราจากการวิเคราะห์กลุ่มสอบทานผล เนื่องจากเป็นวิธีที่ใช้คะแนนของตนเองเป็นเกณฑ์จึงมีความอิสระไม่ขึ้นกับกระบวนการแปลงคะแนนอื่น ๆ และใช้แนวคิดของปีเตอร์เสนและคณะคือ ใช้ค่าความแปรปรวนเป็นตัวถ่วงน้ำหนักเพื่อให้ค่าที่ได้มีหน่วยเป็นมาตรฐาน ค่าดัชนีความแตกต่าง คือ ค่าความคลาดเคลื่อนรวมในการเทียบมาตราของวิธีการเทียบมาตราที่ใช้นั้นซึ่งคำนวณจากค่าเฉลี่ยยกกำลังสองของความแตกต่างระหว่างคะแนนแปลงกับคะแนนเกณฑ์และถ่วงน้ำหนักด้วยความแปรปรวนของคะแนนเกณฑ์

$$C = \frac{\Sigma\left(X_i - X_i^*\right)^2}{nS_x^2}$$

เมื่อ C คือ ดัชนีความแตกต่าง

X_i คือ คะแนนจากการสอบชุด X ของผู้สอบคนที่ i

X_i^* คือ คะแนนจากแบบสอบฉบับ X ที่ได้จากการนำคะแนนจากฉบับ Y ไปแปลงจากตารางเทียบมาตราของคนที่ I

n คือ จำนวนในกลุ่มสอบทานผล

S_x^2 คือ ค่าความแปรปรวนของคะแนนจากแบบสอบฉบับ X

เกณฑ์ในการเทียบมาตราตามระดับการยอมรับ จะใช้เกณฑ์เดียวกับ ปีเตอร์เสนและคณะ (Petersen, et al. 1982 : 93-94 อ้างจาก ภาวิณี ศรีสุขวัฒนานันท์,2529 : 103) โดยได้กำหนดคุณภาพในการเทียบมาตราตามระดับการยอมรับ ดังนี้

ระดับน่าพอใจอย่างยิ่ง เมื่อ $C \leq (0.05\ S.D_{IX})^2$

ระดับน่าพอใจ เมื่อ $(.05\ S.D_{IX})^2 < C \leq (0.10\ S.D_{IX})^2$

ระดับปานกลาง เมื่อ $(0.10\ S.D_{IX})^2 < C \leq (0.15\ S.D_{IX})^2$

ระดับไม่น่าพอใจ เมื่อ $(0.15\ S.D_{IX})^2 < C \leq (0.20\ S.D_{IX})^2$

ระดับไม่น่าพอใจอย่างยิ่ง เมื่อ $(0.20\ S.D_{IX})^2 < C$

งานวิจัยที่เกี่ยวกับการเทียบมาตรา

1. งานวิจัยที่เกี่ยวกับการเทียบมาตราวิธีอิควิเปอร์เซนไทล์

การเทียบมาตราวิธีอิควิเปอร์เซนไทล์เป็นการเทียบมาตราตามทฤษฎีการทดสอบดั้งเดิม โดยยึดหลักการแปลงคะแนนที่ได้จากแบบสอบ 2 ฉบับให้เป็นตำแหน่งที่เปอร์เซนไทล์เดียวกัน คะแนนของแบบสอบ 2 ฉบับนั้นจึงเท่าเทียมกัน การเทียบมาตราโดยวิธีนี้มีข้อดีกว่าการเทียบมาตราเชิงเส้นตรง เนื่องจากให้ผลที่เพียงพอที่จะนำมาใช้ในการเทียบมาตรา เมื่อแบบสอบมีค่าความยากแตกต่างกัน (Kolen, 1981: 1-11) ซึ่งสอดคล้องกับผลการวิจัยของสไลด์และลินน์ (Slinde and Linn, 1977 :23-31) ที่พบว่าการเทียบมาตราวิธีอิควิเปอร์เซนไทล์ให้ผลการเทียบมาตราดีกว่าการเทียบเชิงเส้นตรงในกรณีที่แบบสอบมีค่าความยากต่างกัน ปีเตอร์เสน มาร์โค และสตีเวอร์ท (Petersen Marco and Stewart,1982 : 71-135) ได้ศึกษาความเพียงพอของการเทียบมาตราวิธีอิควิเปอร์เซนไทล์กับวิธีเชิงเส้นตรง โดยใช้กลุ่มตัวอย่างต่างๆกัน คือกลุ่มสุ่มอย่างง่าย กลุ่มที่มีความสามารถใกล้เคียงกันและกลุ่มที่มีความสามารถไม่ใกล้เคียงกัน พบว่า การเทียบมาตราวิธีอิควิเปอร์เซนไทล์ให้ผลการเทียบที่ดีกว่าวิธีเชิงเส้นตรง เมื่อแบบสอบทั้ง 2 ฉบับ มีความสัมพันธ์เชิงเส้นโค้งอันเนื่องมาจากมีความยากแตกต่างกันและจากกลุ่มตัวอย่างที่ไม่คล้ายคลึงกัน และพบว่าถ้ากลุ่มตัวอย่างมีความสามารถใกล้เคียงกันแล้ว วิธีอิควิเปอร์เซนไทล์จะให้ผลการเทียบมาตราที่มีความคลาดเคลื่อนรวมมากกว่าวิธีเชิงเส้นตรง แต่โคเลนและวิทนีย์ (Kolen and Whitney,1982 : 229-293) ได้ผลการวิจัยที่แตกต่างจากนี้ โดยโคเลนและวิทนีย์ได้เปรียบเทียบความเพียงพอของการเทียบมาตราวิธีอิควิเปอร์เซนไทล์ และวิธีเชิงเส้นตรง โดยใช้แบบสอบมาตรฐานวัดผลสัมฤทธิ์ทางการเรียน 5 วิชา ที่มีเนื้อหาแตกต่างกัน ใช้กลุ่มตัวอย่างขนาดเล็ก ประมาณกลุ่มละ 200 คน

ผลการวิจัยพบว่า วิธีอิควิเปอร์เซนไทล์ให้ผลการเทียบมาตราที่มีความคงเส้นคงวามากกว่าวิธีเชิงเส้นตรง เนื่องจากผลการศึกษาที่แตกต่างกันโคเลนและวิทนีย์จึงได้สรุปว่า วิธีการเทียบมาตราที่มีความพึงพอขึ้นกับองค์ประกอบหลายอย่าง ได้แก่ คุณลักษณะต่างๆของแบบสอบ วิธีที่ใช้ในการเทียบและคุณลักษณะของกลุ่มตัวอย่างที่ใช้การวิจัย เป็นต้น

สำหรับการเทียบมาตราวิธีอิควิเปอร์เซนไทล์ในประเทศไทยนั้น เริ่มจาก ภาวิณี ศรีสุขวัฒนานันท์ (2529 : 155-170) ได้เปรียบเทียบผลของการใช้วิธีการเทียบมาตราวิธีอิควิเปอร์เซนไทล์ วิธีเชิงเส้นตรงและทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ ในแบบสอบวัดผลสัมฤทธิ์และแบบสอบคัดเลือก ใช้กลุ่มตัวอย่างกลุ่มละ 1,500 คน ผลการวิจัยพบว่า การเทียบมาตราวิธีอิควิเปอร์เซนไทล์เป็นวิธีที่ให้ความเพียงพอมากที่สุดเมื่อเป็นแบบสอบคัดเลือก แต่ถ้าเป็นแบบสอบวัดผลสัมฤทธิ์ทางการเรียนแล้ว การเทียบมาตราวิธีอิควิเปอร์เซนไทล์จะให้ผลการเทียบรองลงมา ผู้วิจัยจึงอภิปรายว่า วิธีการเทียบมาตราที่ให้ความเพียงพอต่างกันในกรณีแบบสอบคัดเลือก และแบบสอบวัดผลสัมฤทธิ์ทางการเรียนอาจเนื่องมาจากปัจจัยด้านแบบสอบและปัจจัยด้านประชากรผู้สอบ ต่อมาวรเนตร พิชิตเกริกพล (2537 : 94-100) ได้เปรียบเทียบคุณภาพของการเทียบมาตราระหว่างแบบอิควิเปอร์เซนไทล์กับเชิงเส้นตรง โดยศึกษากรณีของลักษณะแบบสอบที่มีความยากใกล้เคียงกันและแตกต่างกัน และกลุ่มตัวอย่างที่มีความสามารถใกล้เคียงกันและแตกต่างกัน ใช้กลุ่มตัวอย่าง 900 คน พบว่า การเทียบมาตราวิธีอิควิเปอร์เซนไทล์ให้คุณภาพดีกว่าวิธีเชิงเส้นตรงเมื่อแบบสอบมีความยากใกล้เคียงกันกับกลุ่มตัวอย่างที่มีความสามารถแตกต่างกันและในกรณีที่แบบสอบมีความยากแตกต่างกันกับกลุ่มตัวอย่างที่มีความสามารถใกล้เคียงกันซึ่งสอดคล้องกับการวิจัยของโคเลน สไลด์และลินน์ดังที่กล่าวมาแล้ว

สรุปงานที่วิจัยเกี่ยวกับการเทียบมาตราวิธีอิควิเปอร์เซนไทล์

1) เมื่อแบบสอบมีค่าความยากแตกต่างกัน การเทียบมาตราวิธีอิควิเปอร์เซนไทล์ให้ผลการเทียบที่ดีกว่าการเทียบมาตราเชิงเส้นตรง

2) เมื่อเป็นแบบสอบคัดเลือก การเทียบมาตราวิธีอิควิเปอร์เซนไทล์เป็นวิธีที่ให้ความเพียงพอมากมากกว่าการเทียบมาตราเชิงเส้นตรง

3) ในกรณีที่แบบสอบมีความยากใกล้เคียงกันการเทียบมาตราวิธีอิควิเปอร์เซนไทล์ให้คุณภาพดีกว่าวิธีเชิงเส้นตรงเมื่อเทียบกับกลุ่มตัวอย่างที่มีความสามารถแตกต่างกัน

2 งานวิจัยที่เกี่ยวกับการเทียบมาตราเชิงเส้นตรง

การเทียบมาตราเชิงเส้นตรงเป็นการเทียบมาตราตามทฤษฎีดั้งเดิมโดยยึดหลักว่า คะแนนของแบบสอบ 2 ฉบับ จะเท่าเทียมกัน ถ้าต่างก็ตรงกับคะแนนมาตรฐานเดียวกัน มีผู้วิจัยเกี่ยวกับการเทียบมาตราเชิงเส้นตรงไว้ ดังนี้ โคเลน (Kolen, 1981 : 1-11) ได้ศึกษาการเทียบมาตราเชิงเส้นตรงโดยใช้แบบสอบวัดผลสัมฤทธิ์ที่มีค่าความยากเท่าเทียมกันและแตกต่างกัน ใช้กลุ่มตัวอย่างขนาดใหญ่ แล้วให้นักเรียนแต่ละคนได้ทำแบบสอบคนละ 1 ฉบับ โดยการสุ่ม พบว่า การเทียบมาตราเชิงเส้นตรงจะให้ผลการเทียบมาตรามีความคงเส้นคงวามากที่สุดเมื่อแบบสอบมีค่าความยากแตกต่างกัน ต่อมาโคเลนและวิทนีย์ (Kolen and Whitney,1982 : 279-293) ได้เปรียบเทียบความถูกต้องของการเทียบมาตรา 4 วิธี คือ วิธีอิควิเปอร์เซนไทล์ เชิงเส้นตรง วิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งและสามพารามิเตอร์ โดยใช้แบบสอบมาตรฐานวัดผลสัมฤทธิ์ทางการเรียน โดยใช้กลุ่มตัวอย่างขนาดเล็ก กลุ่มละประมาณ 200 คน พบว่า การเทียบมาตราเชิงเส้นตรงให้ผลการเทียบมาตราดีกว่าการเทียบมาตราอิควิเปอร์เซนไทล์ ซึ่งผลการวิจัยนี้ขัดแย้งกับการศึกษาของโคเลนดังที่กล่าวมาแล้ว ในปีเดียวกันปีเตอร์เสน มาร์โค และสตีเวอร์ท (Petersen Marco and Stewart,1982 : 71-135) ได้ศึกษาความเพียงพอของการเทียบมาตราเชิงเส้นตรงหลายๆ วิธี โดยใช้แบบสอบ TSWE (Test of Standard Written English) ใช้กลุ่มตัวอย่างขนาดต่างกันคือ กลุ่มสุ่มอย่างง่าย กลุ่มที่มีความสามารถคล้ายคลึงกันและกลุ่มที่มีความสามารถไม่คล้ายคลึงกัน ผลการวิจัยพบว่า การเทียบมาตราที่ใช้เทคนิคพอทพอฟ (Potthoff) ให้ผลการเทียบมาตราที่ไม่น่าพอใจในทุกสถานการณ์ สำหรับกลุ่มตัวอย่างที่มีความสามารถไม่คล้ายคลึงกันการเทียบมาตราเชิงเส้นตรงที่ใช้เทคนิคของทัคเกอร์ 1 (Tucker 1) ให้ผลการเทียบมาตราที่มีความคลาดเคลื่อนมากกว่าวิธีเชิงเส้นตรงอื่นๆ สำหรับการศึกษาเกี่ยวกับการเทียบมาตราเชิงเส้นตรงในประเทศไทยนั้น ภาวิณี ศรีสุขวัฒนานันท์ (2529 : 155 -170) ได้เปรียบเทียบผลของการใช้วิธีการเทียบมาตรา วิธีอิควิเปอร์เซนไทล์ วิธีเชิงเส้นตรงและโดยใช้ทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ในแบบสอบวัดผลสัมฤทธิ์และแบบสอบคัดเลือก ผลการวิจัยพบว่า การเทียบมาตราเชิงเส้นตรงเป็นวิธีที่มีความเพียงพอมากที่สุดในกรณีของแบบสอบวัดผลสัมฤทธิ์ ส่วนวรเนตร พิชิตเกริกพล (2537: 94-100) ได้เปรียบเทียบคุณภาพของการเทียบมาตราระหว่างแบบอิควิเปอร์เซนไทล์กับเชิงเส้นตรง โดยศึกษากรณีของลักษณะแบบสอบที่มีความยากใกล้เคียงกันและแตกต่างกัน และกลุ่มตัวอย่างที่มีความสามารถใกล้เคียงกันและแตกต่างกัน พบว่า ในกรณีที่แบบสอบมีความยากใกล้เคียงกันและกลุ่มตัวอย่างมีความสามารถใกล้เคียงกันแล้ววิธีเชิงเส้นตรงจะให้ผลการเทียบมาตราที่ดีกว่าวิธีอิควิเปอร์เซนไทล์ ซึ่งสอดคล้องกับผลการศึกษาของโคเลนที่กล่าวมาแล้ว

สรุปงานวิจัยที่เกี่ยวกับการเทียบมาตราเชิงเส้นตรง

1) การเทียบมาตราเชิงเส้นตรงให้ผลการเทียบมาตราคงเส้นคงวาน้อย เมื่อแบบสอบมีความยากแตกต่างกัน

2) เมื่อกลุ่มตัวอย่างมีขนาดเล็กและความสามารถใกล้เคียงกัน การเทียบมาตราเชิงเส้นตรงจะให้ความเพียงพอของการเทียบมาตรามากกว่าวิธีอิควิเปอร์เซนไทล์

3) การเทียบมาตราเชิงเส้นตรงจะให้ความเพียงพอของการเทียบมาตรามากที่สุดในกรณีที่เป็นแบบสอบวัดผลสัมฤทธิ์

3. งานวิจัยที่เกี่ยวกับการเทียบมาตราวิธีการใช้เทคนิคการวิเคราะห์องค์ประกอบ

เป็นวิธีการเทียบมาตราที่เสนอโดยร็อค (Rock,1982 :251 อ้างจาก เรวดี อินทสระ, 2530 : 2) มีผู้ศึกษาน้อยมาก เนื่องจากมีความคลาดเคลื่อนมาตรฐานของการเทียบมาตราสูงและมีความเพียงพอของการเทียบมาตราอยู่ในระดับปานกลางเท่านั้น (เรวดี อินทสระ,2530 : 64-69) โดยผู้วิจัยได้เปรียบเทียบความคลาดเคลื่อนของการเทียบมาตราระหว่างวิธีการใช้ทฤษฎีการตอบสนองข้อสอบกับวิธีการใช้เทคนิคการวิเคราะห์ตัวประกอบ ใช้แบบสอบวัดผลสัมฤทธิ์ทางการเรียนกับกลุ่มตัวอย่างจำนวน 2,823 คน ความคลาดเคลื่อนมาตรฐานของการเทียบมาตราวิธีการใช้เทคนิคการวิเคราะห์ตัวประกอบ มีค่าสูงกว่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตราวิธีการใช้ทฤษฎีการตอบสนองข้อสอบ ส่วนความเพียงพอของการเทียบมาตราอยู่ในระดับปานกลางเท่านั้น

4. งานวิจัยที่เกี่ยวกับการเทียบมาตราวิธีการใช้ทฤษฎีการตอบสนองข้อสอบ

การเทียบมาตราวิธีการใช้ทฤษฎีการตอบสนองข้อสอบเป็นการเทียบมาตราโดยใช้ค่าความสามารถเดียวกันของผู้สอบ มีผู้ศึกษาไว้ดังนี้ มาร์โค ปีเตอร์เสนและสตีเวอร์ท (Marco Petersen and Stewart,1978 อ้างจาก เรวดี อินทสระ,2530 : 22) ได้ประเมินและเปรียบเทียบวิธีการเทียบมาตราเชิงเส้นตรง วิธีอิควิเปอร์เซนไทล์ และวิธีการใช้โค้งลักษณะข้อสอบ ผลปรากฏว่า เมื่อแบบสอบมีค่าความยากแตกต่างกันการเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบให้ผลดีที่สุด ต่อมาโคเลน (Kolen,1981 : 1-11) ได้ทำการเปรียบเทียบผลการเทียบมาตราระหว่างวิธีเชิงเส้นตรง วิธีอิควิเปอร์เซนไทล์ และวิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งและสามพารามิเตอร์ พบว่า การเทียบมาตราวิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ ให้ผลแน่นอนที่สุด โคเลนได้อภิปรายสรุปว่า

วิธีที่ให้ผลการเทียบมาตราที่มีประสิทธิภาพน้อยที่สุดคือ วิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์ อาจเนื่องมาจากการไม่ได้ระบุการเดา ซึ่งเป็นส่วนหนึ่งของการกำหนดค่าความสามารถ ดังนั้นข้อมูลที่มาจากข้อสอบที่มีความยากต่างกัน การเทียบมาตราอาจจะได้รับผลกระทบและให้ประสิทธิภาพที่ต่ำลง โคเลนและวิทนีย์ (Kolen and Whitney,1982 279-293) ได้เปรียบเทียบความถูกต้องของการเทียบมาตรา 4 วิธี คือ วิธีอิควิเปอร์เซนไทล์ วิธีเชิงเส้นตรง และวิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งและสามพารามิเตอร์ พบว่าวิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ ให้ผลยังไม่เป็นที่ยอมรับ ส่วนวิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์ให้ผลคงที่มากกว่า ซึ่งผลการวิจัยนี้ขัดแย้งกับผลการวิจัยของโคเลน (Kolen,1981) ต่อมา ลอร์ดและวินเกอร์สกี้ (Lord and Wingersky,1984 : 153-161 อ้างจาก เรวดี อินทสระ, 530 : 28) ได้เปรียบเทียบการเทียบมาตราวิธีใช้ทฤษฎีการตอบสนองข้อสอบที่ใช้คะแนนจริงกับวิธีอิควิเปอร์เซนไทล์ ซึ่งผลการวิจัยไม่สามารถสรุปได้ว่าวิธีการเทียบมาตราทั้ง 2 วิธี วิธีใดให้ผลดีกว่ากัน และได้ทำการวิเคราะห์ซ้ำอีกครั้งผลก็คล้ายคลึงกันคือ ไม่สามารถสรุปได้ สำหรับการเทียบมาตราโดยใช้ทฤษฎีการตอบสนองข้อสอบในประเทศไทยนั้นเริ่มจาก ภาวิณี ศรีสุขวัฒนานันท์ (2529 : 155-170) โดยเปรียบเทียบผลของการใช้การเทียบมาตราระหว่างวิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ วิธีอิควิเปอร์เซนไทล์ และวิธีเชิงเส้นตรง พบว่ากรณีแบบสอบคัดเลือก รูปแบบการเทียบมาตราที่ให้ผลดีคือ วิธีอิควิเปอร์เซนไทล์ วิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ และวิธีเชิงเส้นตรงตามลำดับ ต่อมา เรวดี อินทสระ (2530 : 64-69)ได้เปรียบเทียบผลการเทียบมาตราระหว่างวิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์กับวิธีการใช้เทคนิคการวิเคราะห์องค์ประกอบ ผลการวิจัยพบว่า ความคลาดเคลื่อนมาตรฐานของการเทียบมาตราของการใช้ทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ มีค่าน้อยกว่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตราการใช้เทคนิคการวิเคราะห์องค์ประกอบ และมีความเพียงพอของการเทียบมาตราอยู่ในระดับที่น่าพอใจ ส่วนความเพียงพอของการเทียบมาตราการใช้เทคนิคการวิเคราะห์องค์ประกอบอยู่ในระดับปานกลางเท่านั้น

สรุปผลการวิจัยเกี่ยวกับการเทียบมาตราวิธีการใช้ทฤษฎีการตอบสนองข้อสอบ

1) เมื่อแบบสอบมีค่าความยากแตกต่างกันการเทียบมาตราวิธีการใช้ทฤษฎีการตอบสนองข้อสอบให้ผลดีกว่าวิธีเชิงเส้นตรง และวิธีอิควิเปอร์เซนไทล์

2) การเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งและสามพารามิเตอร์ไม่สามารถสรุปได้ว่าการเทียบมาตรารูปแบบใดที่ให้ผลการเทียบมาตราที่ดีกว่ากัน

3) การเทียบมาตราระหว่างวิธีการใช้ทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ ให้ผลดีกว่าการเทียบมาตราวิธีการใช้เทคนิคการวิเคราะห์องค์ประกอบ