หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

วรรณคดีที่เกี่ยวข้อง

ในการนำเสนอวรรณคดีที่เกี่ยวข้อง ผู้วิจัยได้นำเสนอโดยแยกกล่าวเป็นตอน ๆ ดังนี้

ตอนที่ 1 พัฒนาการของระบบมาตราและการเทียบมาตรา

ตอนที่ 2 ทฤษฎีที่เกี่ยวข้องกับการเทียบมาตรา

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการเทียบมาตรา

ตอนที่ 1 พัฒนาการของระบบมาตราและการเทียบมาตรา

การเทียบมาตราได้มีการพัฒนาแนวความคิดและเทคนิควิธีมานาน ดังที่ไรท์และสโตน (Wright & Stone 1979) ได้เสนอบทความเกี่ยวกับพัฒนาการของระบบมาตราในหนังสือ BEST TEST DESIGN หน้า vii - x พอสรุปความได้ดังนี้

เธอร์สโตน (Thurstone 1925) ได้ค้นคิดวิธีการแปลงสัดส่วนจำนวนคนที่ตอบแบบสอบได้ถูกในกลุ่มอายุหนึ่งให้เป็นหน่วยความเบี่ยงเบนปกติ (normal deviate) และใช้ค่าเหล่านี้เป็นมาตราพื้นฐาน ส่วนค่ามาตราของกลุ่มอายุอื่นๆจะกำหนดขึ้นจากการสร้างข้อตกลงว่าด้วยความสัมพันธ์เชิงเส้นตรงระหว่างมาตราต่างๆ ของข้อกระทงที่ได้จากแบบสอบหลายๆฟอร์มโดยผู้ตอบมาจากหลายๆกลุ่มอายุ สำหรับค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานจัดเป็นค่าพารามิเตอร์ที่จะใช้ในการแปลงคะแนนให้อยู่บนมาตราเดียวกัน (Common scale) มาตราของเธอร์สโตนประสบปัญหาของความไม่แน่นอน (Invariance to sampling) จะแปรผันตามการแจกแจงความสามารถของกลุ่มตัวอย่าง ต่อมาเขาได้ปรับปรุงแก้ไขวิธีการ โดยส่วนมากจะอาศัยเทคนิคตำแหน่งเปอร์เซนไทล์ และคะแนนมาตรฐาน ซึ่งภายหลังโลวิงเกอร์ (Lovinger 1947) ได้พิสูจน์ให้เห็นว่าวิธีการนั้นเป็นวิธีที่อ่อนยังมีคุณภาพไม่เข้าขั้น

โลวิงเกอร์ (Lovinger 1974: 46) ได้แสดงให้เห็นว่า การวัดที่ถูกต้องมีเกณฑ์ที่จำเป็นคือ ความเป็นเอกพันธ์ของแบบสอบ และสเกลที่เพิ่มขึ้นทางเดียว นอกจาก

นั้น วิธีการสร้างมาตราที่ยอมรับได้จะต้องให้ผลที่เป็นอิสระไม่ผูกพันกับหน่วยเริ่มต้นที่ใช้สร้าง หรือกลุ่มผู้รับการทดสอบกลุ่มแรก

ต่อมากัลลิกเซน (Gulliksen 1950) ได้ย้ำถึงความเข้มงวดของนิยามการเทียบมาตรา โดยเสนอวิธีการให้กลุ่มผู้สอบกลุ่มเดียวทำแบบสอบสองฟอร์ม ให้ตรวจสอบความเป็นคู่ขนานของแบบสอบเสียก่อน โดยทดสอบด้วยสถิติของวิคส์(Wilks) หรือในกรณีที่แบบสอบทั้งสองต่างมีอำนาจทำนายแบบสอบอีกฉบับหนึ่ง อาจพิจารณาจากค่าสหสัมพันธ์ระหว่างแบบสอบหนึ่งกับเกณฑ์ ($r_{yc} = r_{xc}$) วิธีนี้จะบอกได้ว่าแบบสอบทั้งสองฟอร์มนั้นเพียงพอที่จะนำมาเทียบกันหรือไม่

กัลลิกเซน ได้อาศัยความคิดของ เคลเลย์ (Kelley 1947) ในการใช้แบบสอบร่วม (Anchor test) เพื่อสร้างมาตรสำหรับแบบสอบใหม่ นอกจากนี้ได้พัฒนาความคิดต่อจาก เลยาร์ด ทัคเคอร์ (Ledyard Tucker) ในการเทียบมาตราจากแบบสอบหลายๆฟอร์ม เมื่อกลุ่มผู้รับการทดสอบหลายกลุ่ม โดยใช้แบบสอบร่วม และสมการถดถอยเชิงเส้นตรงในการหาค่าประมาณที่ได้จากคะแนนแบบสอบโดยตรง

ทัคเคอร์ (Tucker 1953) ได้เสนอความคิดในที่ประชุมของสถาบันบริการทดสอบทางการศึกษา (Educational Testing Service) เกี่ยวกับแบบสอบในอุดมการของเขาว่า แบบสอบควรมีคุณภาพพอที่จะกำหนดคะแนนแทนตำแหน่งบนความต่อเนื่องลักษณะใดลักษณะหนึ่ง เพื่อให้ความแตกต่างที่มีลักษณะเป็นแบบฉบับนั้นแทนคะแนนสองตำแหน่งที่มีความสอดคล้องกับความแตกต่างของผลการสอบทุกกรณีไม่ว่าจะเป็นคะแนนระดับใด ทัคเคอร์ได้เสนอให้ประเมินความเป็นเอกพันธ์ของแบบสอบด้วย

เลวิน (Levine 1955 cited in Holland & Rubin 1982: 10) ได้ให้ข้อตกลงเบื้องต้นของการเทียบมาตราไว้ว่า "แบบสอบสองฉบับใดๆจะเทียบมาตราคะแนนกันได้ก็ต่อเมื่อแบบสอบทั้งสองนั้นคู่ขนานกัน (parallel) ในด้านต่อไปนี้

1. โครงสร้าง (Structure)
2. เวลา (Timing)
3. รูปแบบ (Format)
4. ชนิดของข้อกระทง (Item type)
5. เนื้อหาวิชา (Subject matter)

แองกอฟฟ (Angoff 1960: 815 cited in Wright & Stone 1979) ได้เขียนบทความลงในสารานุกรมการวัดผล โดยกล่าวว่า การสร้างมาตราของแบบสอบมักได้จากข้อมูลที่มาจากการสอบ ทำให้ระบบที่สร้างขึ้นมีความหมายตราบเท่าที่กลุ่มผู้สอบมีลักษณะคล้ายกลุ่มแรกเท่านั้น แต่ถ้ากาลต่อมาปรากฏว่า คำนิยามของกลุ่มตัวอย่างเปลี่ยนไป จุดมุ่งหมายการศึกษาเปลี่ยนจุดเน้นไป ปัญหาของระบบมาตราของแบบสอบนั้นจะเกิดขึ้น แองกอฟฟกล่าว่า เป็นปัญหาที่รุนแรงมากและย้ำว่า แบบสอบใด ๆ จะยังคงมีคุณค่าต่อไปถ้ามีหน่วยมาตราที่คงที่ ไม่แปรผันไปกับเวลา มีลักษณะเป็นระบบของหน่วย สามารถเก็บเป็นข้อมูลสะสมไว้เพื่อเปรียบเทียบได้

แองกอฟฟ (Angoff 1971) ได้ให้ความหมายของการเทียบมาตราว่า "เป็นวิธีการเปลี่ยนระบบหน่วยคะแนนของแบบสอบฟอร์มหนึ่งไปสู่ระบบหน่วยคะแนนของแบบสอบอีกฟอร์มหนึ่ง เพื่อทำให้ได้คะแนนที่แปลงจากฟอร์มทั้งสองมีความเท่าเทียมกันโดยตรงหลังจากผ่านกระการเปลี่ยนแปลง"

ส่วนอีกด้านหนึ่งของการวัดผล ได้มีการพัฒนาทฤษฎีคุณลักษณะแฝง (Latent trait theory) ตั้งแต่ต้นศตวรรษที่ 19 โดยผู้ที่สร้างโมเดลนี้ครั้งแรกคือ เวเบอร์และเฟคเนอร์ (Weber & Fechner) เขาได้สร้างโมเดลง่ายๆโดยโมเดลขึ้นกับโอกาสในการตอบข้อกระทง ซึ่งต่อมาลอเลย์และฟินเนย์ (Lawley & Finney 1940) ได้ใช้ นอร์มัล โอจัยฝ์ (normal ogive) เป็นโมเดลของการตอบข้อกระทง และต่อมามีผู้คิดค้นให้เกิดการนำไปใช้ได้แก่ ลอร์ด (Lord 1952) เบิร์นบอม (Birnbaum ,1968) เบเคอร์ (Baker 1968) ราสช (Rasch 1968) แอนเดอร์เซน (Andersen ,1973) ตลอดจน ไรท์ (Wright 1977) ทฤษฎีนี้อาศัยแนวความคิด รูปแบบความน่าจะเป็น (probabilistic model)โดยวิเคราะห์ค่าพารามิเตอร์ของตัวข้อสอบและประมาณค่าพารามิเตอร์ความสามารถของคนในแต่ละระดับ ค่าพารามิเตอร์ที่ได้มีคุณสมบัติของความไม่แปรเปลี่ยน

ลอร์ด (Lord 1977) ได้เสนอความรู้ในการเทียบมาตรา โดยเน้นทฤษฎีการสนองตอบข้อสอบ ซึ่งเขาเสนอให้ใช้ค่าพารามิเตอร์ของข้อกระทงและค่าความสามารถของคนหาคะแนนจริงของแบบสอบ แล้วเทียบคะแนนจริงของแบบสอบทั้งสองฟอร์มในระดับความสามารถที่เท่ากัน

ในระหว่างปี 1978 - 1980 สถาบันบริการทดสอบทางการศึกษา ได้ตั้งทีมงานขึ้นมาเรียกว่า Program statistics Research Project (PSRP) เพื่อทำการวิจัยและศึกษาเกี่ยวกับการเทียบมาตรา ในแง่ของเทคนิควิธีที่เกี่ยวกับการใช้สถิติเพื่อการเปรียบเทียบคะแนนและศึกษาค้นคว้าทฤษฎีความตรงของแบบสอบ (Test Validity) ในด้านเทคนิควิธีทางสถิติ PSRP จัดประชุมสัมนาเกี่ยวกับการเทียบมาตรา เมื่อเดือน เมษายน 1980 และได้จัดพิมพ์หนังสือซึ่งรวบรวมข้อเขียน ข้อวิพากวิจารณ์เกี่ยวกับการเทียบมาตรา รวบรวมโดย ฮอลแลนด์และรูบิน (Holland & Rubin 1982)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 2 ทฤษฎีที่เกี่ยวข้องกับการเทียบมาตรา

<u>มโนทัศน์ของการเทียบมาตรา</u> (Braun & Holland 1982 : 10-18)

1. สมมุติให้ ประชากร P แทนประชากรของผู้สอบซึ่งอาจสอบแบบสอบฟอร์ม X หรือแบบสอบฟอร์ม Y

2. กำหนดให้ $F(x)$ และ $G(y)$ แทนฟังก์ชันการกระจาย(distribution function) ของคะแนน ฟอร์ม X และฟอร์ม Y ตามลำดับ

3. สมมุติให้ ฟังก์ชันการเทียบมาตรา e_x แทนฟังก์ชันของการเทียบมาตราจากแบบสอบฟอร์ม Y ไปสู่แบบสอบฟอร์ม X

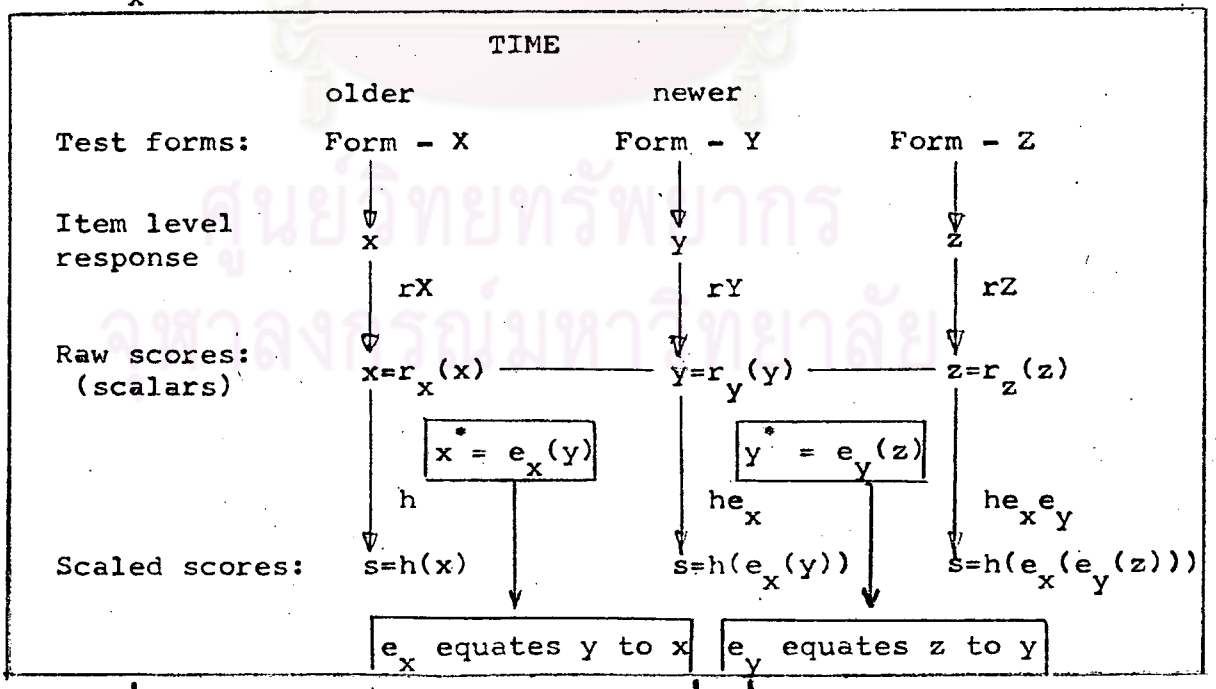
ฉะนั้น $X^* = e_x(y)$

4. ให้ $F^*(x)$ เป็นฟังก์ชันการกระจายของ X^*

$$\therefore \quad F^*(x) = G(e_x^{-1}(x))$$

<u>คำนิยาม</u>

"ถ้า $F = F^*$ แล้ว แบบสอบฟอร์มXและฟอร์ม Y สามารถเทียบโดยฟังก์ชัน e_x ในประชากร P "



ภาพที่ 1 สถานการณ์ของการสอบและระบบอย่างง่ายของการเทียบมาตรา

(Braun & Holland 1982 : 11)

จากภาพที่ 1 อธิบายความหมายได้ดังต่อไปนี้

"ระดับการสนองตอบข้อกระทง" (Item level response) เป็นเวคเตอร์ของคะแนนรายข้อ โดยส่วนใหญ่แล้วจะใช้ 1 แทนการตอบถูก 0 แทนการตอบผิด และเลขอื่น ๆ แทนข้อที่เว้นว่าง (omitted) ฉะนั้นแบบสอบฟอร์ม X จะมีเวคเตอร์ของการสนองตอบข้อกระทง x แทนของแต่ละคนที่สอบแบบสอบฟอร์ม X

"คะแนนดิบ" (Raw score) เป็นปริมาณตัวเลขที่ได้จากการตอบข้อกระทงของแต่ละคนซึ่งคำนวณจากฟังก์ชั่น r โดยทั่วไป ฟังก์ชั่น r ใช้วิธีใดวิธีหนึ่ง คือ

r = จำนวนข้อถูก หรือ

r = จำนวนข้อถูก - คะแนนการเดา (แบบสอบชนิดเลือกตอบ)

ในที่นี้คะแนนดิบใช้สัญญลักษณ์ x ตรงกับคะแนนของแบบสอบฟอร์ม X y ตรงกับคะแนนของแบบสอบฟอร์ม Y และ z ตรงกับคะแนนของแบบสอบฟอร์ม Z

"ฟังก์ชั่นการเทียบมาตรา" (Equating function) คือกฎเกณฑ์หรือฟังก์ชั่นที่กำหนดในการเทียบคะแนนจากแบบสอบฉบับใหม่ (Y) ให้อยู่ในมาตราเดียวกับแบบสอบฉบับเดิม (X) ใช้สัญญลักษณ์ e เช่น

e_x เป็นฟังก์ชั่นที่เทียบคะแนนจากฟอร์ม y ไปสู่ฟอร์ม X

e_y เป็นฟังก์ชั่นที่เทียบคะแนนจากฟอร์ม Z ไปสู่ฟอร์ม Y

คะแนน y ใดๆเมื่อใช้ฟังก์ชั่น e_x เปลี่ยนคะแนนแล้วมักจะใช้สัญญลักษณ์ x^* แทน เพื่อเป็นการง่ายต่อความเข้าใจว่า $x^* = e_x(y)$ คือ คะแนน x แต่มาจากแบบสอบฉบับอื่น

"สเกลฟังก์ชั่น" (The Scaling Function) บางครั้งการรายงานคะแนนการสอบนิยมรายงานคะแนนในรูปแบบของคะแนนมาตรฐาน (Scaled score) ที่ง่ายต่อการเข้าใจและแปลความ (เช่น อีทีเอส รายงาน SAT ด้วยคะแนน 200-800) ฉะนั้นการเทียบมาตราจำเป็นต้องเทียบในรูปของคะแนนมาตรฐานแต่มีข้อจำกัดว่าสเกลฟังก์ชั่นที่ใช้จะต้องเป็นสเกลฟังก์ชั่นเดียวกัน ในภาพที่ 1 สเกลฟังก์ชั่นคือ s

"ฟังก์ชั่นการแปลงคะแนน" (The Conversion Function) คือฟังก์ชั่นที่แปลงคะแนนดิบเป็นคะแนนมาตรฐาน ฉะนั้นสำหรับแบบสอบฟอร์ม X สเกลฟังก์ชั่น

010743

และฟังก์ชั่นการแปลงคะแนนคือฟังก์ชั่นเดียวกัน

$$\text{Test X}: \quad s = h(x)$$

$$\text{Test Y}: \quad s = h(e_x(y))$$

$$\text{Test Z}: \quad s = h(e_x(e_y(z)))$$

<u>ฟังก์ชั่นการเทียบมาตราตามทฤษฎีคลาสิคอล</u>

ทฤษฎีคลาสิคอล (Classical Test Theory) มีวิธีการเทียบมาตราที่สำคัญ 2 วิธี คือ

1. การเทียบโดยใช้เปอร์เซนไทล์
2. การเทียบเชิงเส้นตรง

1. <u>การเทียบโดยใช้เปอร์เซนไทล์</u> (Equipercentile equating)

มีหลักการเกิดขึ้นจากความต้องการพื้นฐาน คือ $F = F^*$ โดยมีฟังก์ชั่นการเทียบมาตราด้วยวิธีนี้ คือ $e_x(y) = F^{-1}(G(y))$

ลักษณะที่สำคัญ

1. ในทางทฤษฎี $F(x)$ และ $G(y)$ มาจากกลุ่มตัวอย่างที่ได้จากการสุ่ม (P_1 และ P_2) จาก P

2. ในทางปฏิบัติ สถาบันบริการทดสอบทางการศึกษา (Educational Testing Service) ทำการสร้างแบบสอบฟอร์ม Y และแบบสอบฟอร์ม X แล้วสุ่มสลับกันลงในกล่องแต่ละสนามสอบ ฉะนั้นครั้งหนึ่งผู้สอบจะสอบฟอร์ม Y และอีกครั้งหนึ่งจะสอบฟอร์ม X

3. การเทียบโดยใช้เปอร์เซนไทล์จะเทียบ ฟอร์ม X และฟอร์ม Y โดยผู้สอบจากประชากร $\underline{P}$

4. ทั้ง F และ G เป็นลักษณะต่อเนื่อง (Continuous) จาก $F^{-1}(G(y))$

5. ความยาวของแบบสอบฟอร์ม X และฟอร์ม Y ไม่ควรต่างกันมากนัก

2. <u>การเทียบเชิงเส้นตรง</u>

จุดอ่อนของฟังก์ชั่นการเทียบโดยใช้เปอร์เซนไทล์ คือ การประมาณค่า $F(x)$ และ $G(y)$ ฉะนั้นโมเดลที่ง่ายและอาศัยการประมาณค่าพารามิเตอร์ที่เป็นไปได้ในแง่

ของการปฏิบัติ คือ โมเดลเชิงเส้นตรง (linear model)

$$e_x(y) = a\,y + b \qquad (1)$$

จากโมเดลของฟังก์ชั่นการเทียบเชิงเส้นตรงนี้ มีพารามิเตอร์ที่ต้องประมาณค่า 2 ตัว คือ a และ b

ในกรณีของสมการ (1) ซึ่งตรงกันกับฟังก์ชั่นการเทียบโดยใช้เปอร์เซนไทล์ ถ้า

$$F^{-1}(G(y)) = a\,y + b \qquad (2)$$

หรือเช่นเดียวกันถ้า

$$G(y) = F(ay + b) \qquad (3)$$

สมการที่ (3) ตรงกับข้อตกลงที่ให้ F และ G อยู่ในรูปของ

$$F(x) = H((x-u_x)/\sigma_x) \qquad (4)$$

และ

$$G(y) = H((y-u_y)/\sigma_y) \qquad (5)$$

สำหรับฟังก์ชั่นการกระจายบางอย่าง H ในสมการ (4) และ (5) u_x , u_y , σ_x^2 , σ_y^2 คือค่าเฉลี่ยและความแปรปรวนของแบบสอบฟอร์ม X และฟอร์ม Y ถ้าประยุกต์สมการ (4) และ (5) กับสมการ (2) เราจะได้รูปฟอร์มของ a และ b ในสมการ (2) คือ

$$a = \sigma_x/\sigma_y \qquad (6)$$

$$b = u_x - (\sigma_x/\sigma_y)u_y \qquad (7)$$

นำไปใช้กับสมการ (1) เราจะได้ฟังก์ชั่นการเทียบเชิงเส้นตรง คือ

$$x^* = e_x(y) = u_x + (\sigma_x/\sigma_y)(y - u_y)$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

โมเดลของการเทียบมาตรา

การเทียบมาตรา (Equating) เป็นการปรับสเกลคะแนนจากการสอบหลายๆ ครั้งให้อยู่ในสเกลเดียวเพื่อประโยชน์ในการเปรียบเทียบ ซึ่งการเทียบมาตราได้มีการพัฒนากันมานานจากอดีตจนถึงปัจจุบัน โดยถ้าใช้เกณฑ์ทฤษฎีแบบสอบดั้งเดิมและทฤษฎีใหม่พอจะแยกการเทียบมาตราเป็น 2 โมเดล คือ

1. โมเดลการเทียบมาตราตามทฤษฎีคลาสิคอล

    1.1 โมเดลการเทียบโดยใช้คะแนนสอบ (Observed score equating model)

    1.2 โมเดลการเทียบโดยใช้คะแนนจริง (True score equating model) ของเลวิน (Levine 1955)

2. โมเดลการเทียบมาตราตามทฤษฎีสนองตอบข้อสอบ

    2.1 โมเดลการเทียบโดยใช้คะแนนจริง (True score equating model) ของลอร์ด (Lord 1981)

    2.2 โมเดลการเทียบโดยใช้คะแนนดิบ (Raw score equating model)

ในการเสนอรายละเอียดแต่ละโมเดลนั้น เฉพาะโมเดลการเทียบมาตราตามทฤษฎีคลาสิคอลมีหลายรูปแบบ เพื่อความต่อเนื่องในเนื้อหา จะเสนอโมเดลที่ 1.1 และ 1.2 ติดต่อกันในแต่ละรูปแบบ โดยใช้ความเที่ยงของแบบสอบเป็นสิ่งที่กำหนด กล่าวคือ โมเดลการเทียบโดยใช้คะแนนสอบจะอยู่ในกรณีของแบบสอบที่มีความเที่ยงเท่ากันทั้งสองฟอร์ม แต่โมเดลการเทียบโดยใช้คะแนนจริงจะอยู่ในกรณีของแบบสอบที่มีความเที่ยงไม่เท่ากัน

1. <u>โมเดลการเทียบมาตราตามทฤษฎีคลาสิคอล</u>

จากทฤษฎีที่กล่าวแล้วข้างต้น ในการเทียบมาตราจากแบบสอบฟอร์มหนึ่งไปสู่อีกฟอร์มหนึ่ง จะมีองค์ประกอบหลายประการเข้ามาเกี่ยวข้องในการเทียบคะแนนที่เทียบเคียงกัน เช่น กลุ่มผู้สอบได้มาจากการสุ่มหรือไม่ได้สุ่ม ระดับความสามารถของกลุ่มผู้สอบแตกต่างกันหรือใกล้เคียงกัน ลักษณะการกระจายของคะแนน เป็นต้น ซึ่งองค์ประกอบต่างๆเหล่านี้ไม่สามารถที่จะใช้โมเดลเดียวกันในการเทียบมาตราได้ ดังนั้นจึงมีการคิดค้นและพัฒนารูปแบบในการเทียบมาตราต่างๆมากมาย แองกอฟฟ์ (Angoff 1971: 569-586) ได้รวบรวมและเสนอโมเดลการเทียบมาตราไว้ 6 รูปแบบ ด้วยกัน โดยที่แต่ละรูปแบบจะมีทั้งโมเดลที่เทียบโดยใช้คะแนนสอบ (Observe score equating model) และโมเดลการเทียบโดยใช้คะแนนจริง (True score

equating model) ของเลวิน (Levine 1955) ซึ่งมีรายละเอียด ดังต่อไปนี้

<u>รูปแบบที่ 1</u> (Random group-one test administered to each group)

รูปแบบนี้ผู้สอบมาจากการสุ่มจากประชากรกลุ่มเดียวกัน แบ่งเป็นสองกลุ่ม คือ กลุ่ม α สอบฟอร์ม X กลุ่ม β สอบฟอร์ม Y

1. <u>การเทียบเชิงเส้นตรงสำหรับแบบสอบที่มีความเที่ยงเท่ากัน</u>

ขั้นแรกคำนวณค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของคะแนนแบบสอบฟอร์ม X โดยกลุ่ม α และแบบสอบฟอร์ม Y โดยกลุ่ม β จัดให้คะแนนมาตรฐานของทั้งสองกลุ่มเท่ากัน

$$\frac{Y - M_{Y\beta}}{s_{Y\beta}} = \frac{X - M_{X\alpha}}{s_{X\alpha}} \qquad (1)$$

โดยที่ Y คือคะแนนสอบที่ได้จากฟอร์ม Y

$M_{Y\beta}$ คือค่าเฉลี่ยของคะแนนฟอร์ม Y จากการสอบโดยกลุ่ม β

$s_{Y\beta}$ คือค่าเบี่ยงเบนมาตรฐานของคะแนนฟอร์ม Y จากการสอบโดยกลุ่ม β

X คือคะแนนสอบที่ได้จากฟอร์ม X

$M_{X\alpha}$ คือค่าเฉลี่ยของคะแนนฟอร์ม X จากการสอบโดยกลุ่ม α

$s_{X\alpha}$ คือค่าเบี่ยงเบนมาตรฐานของคะแนนฟอร์ม X จากการสอบโดยกลุ่ม α

เมื่อจัดเทอมในสมการ (1) เสียใหม่ จะได้สมการเชิงเส้นตรง คือ

$$Y = \frac{s_{Y\beta}}{s_{X\alpha}} X + M_{Y\beta} - \frac{s_{Y\beta}}{s_{X\alpha}} M_{X\alpha} \qquad (2)$$

ซึ่งจะอยู่ในรูป $Y = aX + b$

เมื่อ $a = s_{Y\beta} / s_{X\alpha}$

$b = M_{Y\beta} - aM_{X\alpha}$

ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา

เช่นเดียวกับวิธีการทางสถิติอื่น ๆ การเทียบมาตราจะมีความคลาดเคลื่อนแบบสุ่มซึ่งเกิดจากการแกว่ง (fluctuation) ของค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของคะแนนทั้งฟอร์ม X และฟอร์ม Y ซึ่งเป็นผลมาจากการสร้างแบบสอบ

ลอร์ด (Lord 1950) ได้อธิบายความคลาดเคลื่อนมาตรฐานในลักษณะของค่าเบี่ยงเบนมาตรฐานของคะแนนแปลงในสเกล y ที่ตรงกับคะแนน x ที่กำหนด ซึ่งคะแนนที่แปลงแล้ว y^* ได้มาจากกลุ่มตัวอย่าง α และ β ที่เป็นอิสระต่อกัน ความคลาดเคลื่อนมาตรฐานของการเทียบมาตราของรูปแบบที่ 1 มีสมการคือ

$$S.E._{y^*}^2 = \frac{2 S_y^2}{N_t} (Z_x^2 + 2) \qquad (3)$$

โดยที่ $S.E._{y^*}^2$ คือความแปรปรวนของความคลาดเคลื่อนของคะแนนที่แปลงแล้ว

N_t คือจำนวนผู้สอบทั้งสองกลุ่ม ($Nt = N_\alpha + N_\beta$)

Z_x คือคะแนนมาตรฐาน ได้จาก $Z_x = \frac{(x-M_x)}{S_x}$

<u>การเทียบโดยใช้เปอร์เซนไทล์สำหรับแบบสอบที่มีความเที่ยงเท่ากัน</u>

จากการกระจายของคะแนนทั้งสองฟอร์ม ฟอร์ม x สอบโดยกลุ่ม α ฟอร์ม y สอบโดยกลุ่ม β คำนวณจุดกลางตำแหน่งเปอร์เซนไทล์สำหรับคะแนนแต่ละชุด พล็อตลงบนกระดาษกราฟ Arithmetic probability paper โดยแกนนอนเป็นคะแนนดิบ แกนตั้งเป็นตำแหน่งเปอร์เซนไทล์ พล็อตตำแหน่งเปอร์เซนไทล์ทั้งสองชุดลงบนกระดาษกราฟแผ่นเดียวกัน อ่านคะแนนทั้งสองฟอร์มที่มีตำแหน่งเปอร์เซนไทล์เดียวกันประมาณ 30 จุด แล้วนำมาพล็อตใหม่ลงบนกระดาษกราฟธรรมดาโดยแกนนอนเป็นคะแนนฟอร์ม x แกนตั้งเป็นคะแนนฟอร์ม y สร้างตารางเทียบคะแนนโดยอ่านจากกราฟที่พล็อตใหม่

2. <u>การเทียบเชิงเส้นตรงสำหรับแบบสอบที่มีความเที่ยงไม่เท่ากัน</u>

ถ้าแบบสอบฟอร์ม x และฟอร์ม y มีความเที่ยงแตกต่างกันอย่างมีนัยสำคัญแล้ว จะไม่มีหนทางที่จะพัฒนาการแปลงคะแนนให้มีลักษณะเทียบเคียงกันได้เลย อย่างไรก็ตาม สถานการณ์ในทางปฏิบัติที่เรียกว่า การเทียบมาตรานั้น เราสามารถที่จะนำเอาความเที่ยงของแบบสอบทั้งสองฟอร์มรวมเข้าไปด้วย ดังสมการ

$$\frac{y - M_{y\beta}}{S_{\tilde{y}_\beta}} = \frac{x - M_{x\alpha}}{S_{\tilde{x}_\alpha}} \qquad (4)$$

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

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

โดยที่ y คือคะแนนสอบที่ได้จากฟอร์ม Y

$M_{Y\beta}$ คือค่าเฉลี่ยของคะแนนฟอร์ม Y จากการสอบโดยกลุ่ม β

s_Y คือค่าเบี่ยงเบนมาตรฐานของคะแนนสอบ (observe score) ของฟอร์ม Y

$s_{\tilde{Y}\beta}$ คือค่าเบี่ยงเบนมาตรฐานของคะแนนจริง (True score) ของฟอร์ม Y

r_{xy} คือความเที่ยงของแบบสอบฟอร์ม Y

x คือคะแนนสอบที่ได้จากฟอร์ม X

$M_{X\alpha}$ คือค่าเฉลี่ยของคะแนนฟอร์ม X จากการสอบโดยกลุ่ม α

s_X คือค่าเบี่ยงเบนมาตรฐานของคะแนนสอบของฟอร์ม X

$s_{\tilde{X}\alpha}$ คือค่าเบี่ยงเบนมาตรฐานของคะแนนจริงของฟอร์ม X

r_{XX} คือความเที่ยงของแบบสอบฟอร์ม X

สมการสำหรับแปลงคะแนนจากสเกลฟอร์ม X ไปสู่สเกลฟอร์ม Y จะเป็น

$$Y = \frac{s_{\tilde{Y}\beta}}{s_{\tilde{X}\alpha}} X + M_{Y\beta} - \frac{s_{\tilde{Y}\beta}}{s_{\tilde{X}\alpha}} M_{X\alpha} \qquad (5)$$

ซึ่งอยู่ในรูป $Y = AX + B$

เมื่อ $A = s_{\tilde{Y}\beta} / s_{\tilde{X}\alpha}$

$B = M_{Y\beta} - AM_{X\alpha}$

<u>รูปแบบที่</u> 2 Random groups , both forms administered to each group, counterbalanced

ลักษณะเช่นเดียวกับรูปแบบที่ 1 เลือกกลุ่มตัวอย่างขนาดใหญ่แล้วแบ่งแบบสุ่มออกเป็น 2 กลุ่ม โดยกลุ่มหนึ่งเป็น α สอบฟอร์ม X แล้วตามด้วยฟอร์ม Y อีกกลุ่มหนึ่งเป็นกลุ่ม β สอบฟอร์ม Y แล้วตามด้วยฟอร์ม X ในการที่จะป้องกันความคลาดเคลื่อนที่จะเกิดขึ้นในการสอบโดยเฉพาะอย่างยิ่ง เมื่อสอบแบบสอบฟอร์ม X และฟอร์ม Y พร้อมกัน ควรจะเย็บแบบสอบทั้งสองฟอร์มติดกันโดยครึ่งหนึ่งให้ฟอร์ม X อยู่ด้านบนและอีกครึ่งหนึ่งให้ฟอร์ม Y อยู่ด้านบน และในการสอบควรจะแจกสลับกันเรียงตามผู้สอบ คือ XY , YX ,XY , YX , ...

1. <u>แบบสอบความเที่ยงเท่ากัน</u>

ก. การเทียบเชิงเส้นตรง

ลอร์ด (Lord 1950) ได้เสนอวิธีการนี้โดยมีข้อตกลงว่า ผลกระทบที่เกิดจาก

การฝึกที่มีต่อแบบสอบฟอร์ม Y อันเป็นผลมาจากการสอบฟอร์ม X ก่อน และผลกระทบที่มีต่อฟอร์ม X อันเป็นผลมาจากการสอบฟอร์ม Y ก่อน ต่างก็เป็นสัดส่วนของความเบี่ยงเบนมาตรฐานของแบบสอบทั้งสองฉบับ

$$\frac{K_x}{s_x} = \frac{K_y}{s_y} = H$$

การประมาณค่าที่ดีที่สุดของ H คือการหาค่าเฉลี่ยของความแตกต่างระหว่างค่าเฉลี่ยของฟอร์ม X และค่าเฉลี่ยของฟอร์ม Y เมื่อแต่ละฟอร์มอยู่ในหน่วยของค่าเบี่ยงเบนมาตรฐาน

$$H = \tfrac{1}{2}\left(\frac{M_{x\beta} - M_{x\alpha}}{s_x} + \frac{M_{y\alpha} - M_{y\beta}}{s_y}\right) \qquad (6)$$

ซึ่งสมการนี้ก็คือต้นแบบของสมการสำหรับเทียบเชิงเส้นตรง สมการ (1)

$$(y - M_y) / s_y = (x - M_x) / s_x$$

รูปฟอร์มที่ใช้สำหรับแทนค่าในสมการ (1) คือ

$$M_x = \tfrac{1}{2}(M_{x\alpha} + M_{x\beta} - K_x) \qquad (7)$$

$$M_y = \tfrac{1}{2}(M_{y\alpha} + M_{y\beta} - K_y) \qquad (8)$$

$$s_x^2 = \tfrac{1}{2}(s_{x\alpha}^2 + s_{x\beta}^2) \qquad (9)$$

$$s_y^2 = \tfrac{1}{2}(s_{y\alpha}^2 + s_{y\beta}^2) \qquad (10)$$

เมื่อสมการ (7) ถึง (10) แทนลงในสมการ (1) จะได้สมการเชิงเส้นในรูปของ $y = Ax + B$ โดยที่

$$A = \sqrt{\frac{s_y^2 + s_y^2}{s_x^2 + s_x^2}}$$

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

ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา

ลอร์ด (Lord 1950) ได้ตั้งข้อตกลงเบื้องต้นไว้ว่า ค่าเบี่ยงเบนมาตรฐานของแต่ละฟอร์มและความสัมพันธ์ระหว่างฟอร์มเหมือนกันในประชากรแล้ว ความแปรปรวนของความคลาดเคลื่อนของคะแนนแปลงในสเกล y คือ

$$S.E._{y^*}^2 = S_y^2(1-r_{xy})\frac{Z_x^2(1+r_{xy})+2}{Nt} \qquad (11)$$

เมื่อ $Z_x = \frac{x-M_x}{s_x}$

การเทียบมาตราด้วยรูปแบบนี้มีความถูกต้องสูง โดยอาจจะสังเกตจากการเปรียบเทียบ ความแปรปรวนของความคลาดเคลื่อนในสมการ (11) กับสมการ (3) ตัวอย่างเช่น แบบสอบ 2 ฟอร์ม มีค่าสหสัมพันธ์ .80 ความแปรปรวนของความคลาดเคลื่อนของคะแนนแปลง y ที่จุด $Z_x=0$ ของรูปแบบที่ 3 มีค่า $\frac{1}{10}$ ของขนาดของความคลาดเคลื่อนของรูปแบบที่ 1 นั่นคืออาจจะกล่าวได้ว่า รูปแบบที่ 1 จะได้รับความถูกต้องเหมือนรูปแบบที่ 2 จะต้องใช้ผู้สอบจำนวน 10 เท่าของจำนวนเดิม แสดงว่าการใช้รูปแบบที่ 2 มีข้อดีในด้านของความประหยัดใช้ผู้สอบน้อยแต่มีความถูกต้องมากกว่า

ข. การเทียบเชิงเส้นโค้ง (Curvilinear analog)

การเทียบเชิงเส้นโค้งหรือการเทียบโดยใช้เปอร์เซนไทล์ของรูปแบบนี้ ใช้วิธีการเช่นเดียวกับรูปแบบที่ 1 ซึ่งข้อมูลของฟอร์ม X และฟอร์ม Y จะรวมกันเสียก่อนสำหรับแบบสอบแต่ละฟอร์ม นั่นคือรวมกลุ่มคะแนนของผู้สอบกลุ่ม α และ β สำหรับฟอร์มX และรวมคะแนนของผู้สอบกลุ่ม α และ β สำหรับฟอร์ม Y แล้วจึงใช้วิธีเทียบโดยใช้เปอร์เซนไทล์ ดังได้อธิบายไว้ในรูปแบบที่ 1 แล้ว

2. <u>การเทียบเชิงเส้นตรงสำหรับแบบสอบที่มีความเที่ยงไม่เท่ากัน</u>

เมื่อฟอร์ม X และฟอร์มx มีความเที่ยงไม่เท่ากัน ค่าเฉลี่ยของความแปรปรวนในกลุ่มไม่ควรจะคำนวณจากคะแนนที่ได้จากการสอบ ในสมการ (9) และ (10) แต่ควรจะคำนวณจากคะแนนจริง โดยค่าที่สมนัยกับสมการ (9) และ (10) คือสมการ (12) และสมการ (13)

$$S^2_{\tilde{x}} = \tfrac{1}{2}(S^2_{x\alpha} r_{xx\alpha} + S^2_{x\beta} r_{xx\beta}) \qquad (12)$$

$$S^2_{\tilde{y}} = \tfrac{1}{2}(S^2_{y\alpha} r_{yy\alpha} + S^2_{y\beta} r_{yy\beta}) \qquad (13)$$

รูปแบบที่ 3 Random groups - one test administered to each group Common equating test administered to both groups

1. การเทียบเชิงเส้นตรงเมื่อแบบสอบมีความเที่ยงเท่ากัน

คะแนนสอบที่ได้ส่วนใหญ่ขึ้นกับลักษณะ 2 ประการคือ ความสามารถรายบุคคล และลักษณะของแบบสอบ ในการที่จะเปรียบเทียบผลการสอบของแต่ละบุคคล ซึ่งสอบแบบสอบที่แตกต่างกัน จึงจำเป็นที่จะต้องปรับคะแนนเสียก่อน กล่าวคือ การเทียบมาตรานั้นจะเป็นเพียงการเทียบความแตกต่างของคะแนนอันเป็นผลมาจากความแตกต่างของแต่ละบุคคลหรือของกลุ่ม ถ้ากลุ่ม α และ β ไม่ได้สุ่มมาจากประชากรเดียวกัน ความแตกต่างระหว่างกลุ่มอาจจะเป็นองค์ประกอบที่มีนัยสำคัญมากในการแปรเปลี่ยนของค่า A และ B (สมการ (2),(4)) และจะเป็นผลทำให้เกิดอคติ (bias) ในวิธีการเทียบมาตรา แม้ว่ากลุ่มจะถูกเลือกมาโดยวิธีการสุ่ม ความแตกต่างเล็กน้อยระหว่างกลุ่มอาจจะเกิดขึ้นได้ ซึ่งถ้าไม่ป้องกันก็จะเกิดอคติในสมการแปลงคะแนนซึ่งจะเป็นผลต่อการเปรียบเทียบภายหลัง

ในการที่จะควบคุมให้มันเกิดผล การเทียบมาตราจำเป็นต้องใช้คะแนนของแบบสอบร่วม ฟอร์ม U ซึ่งเป็นข้อกระทงที่เพิ่มขึ้นหรือเป็นข้อร่วมกันระหว่างฟอร์ม X และฟอร์ม Y แบบสอบร่วมนี้ใช้สำหรับปรับความแตกต่างที่อาจจะพบระหว่างกลุ่ม α และ β ในการบริหารแบบสอบ กลุ่ม α จะได้รับแบบสอบฟอร์ม X และฟอร์ม U และกลุ่ม β จะได้รับแบบสอบฟอร์ม Y และฟอร์ม U

ลอร์ด (Lord 1955a) ได้พัฒนาสมการสำหรับรูปแบบนี้โดยใช้การประมาณค่าความเป็นไปได้สูงสุด (maximum likelihood estimates) ประมาณค่าเฉลี่ยและความแปรปรวนของฟอร์ม X และฟอร์ม Y โดยสมการเหล่านี้คือ

$$\hat{\mu}_x = M_{x\alpha} + b_{xu\alpha}(\mu_u - M_{u\alpha}) \quad (14)$$

$$\hat{\mu}_y = M_{y\beta} + b_{yu\beta}(\mu_u - M_{u\beta}) \quad (15)$$

$$\hat{\sigma}^2_x = S^2_{x\alpha} + b^2_{xu\alpha}(\hat{\sigma}^2_u - S^2_{u\alpha}) \quad (16)$$

$$\hat{\sigma}^2_y = S^2_{y\beta} + b^2_{yu\beta}(\hat{\sigma}^2_u - S^2_{u\beta}) \quad (17)$$

เมื่อ $\mu_u = M_{ut}$ และ $\hat{\sigma}^2_u = S^2_{ut}$ และ $t = \alpha + \beta$
ซึ่งการประมาณค่าเหล่านี้ใช้กับสมการ (1) $Y = AX + B$
โดยที่ $A = \hat{\sigma}_y / \hat{\sigma}_x$ และ $B = \hat{\mu}_y - A\hat{\mu}_x$

สำหรับรูปแบบนี้ ลอร์ด (Lord 1950) ได้เสนอสูตรในการหาความแปรปรวนของความคลาดเคลื่อน คือ

$$S.E._{y^*}^2 = 2\hat{\sigma}^2_y(1-\hat{r}^2)\frac{(1+\hat{r}^2)Z^2_x + 2}{N_t} \quad (18)$$

โดยถือว่า $\hat{r} = \frac{b_{xu\alpha}\hat{\sigma}_u}{\hat{\sigma}_x} = \frac{b_{yu\beta}\hat{\sigma}_u}{\hat{\sigma}_y}$

จากสูตรนี้จะเห็นได้ว่า เมื่อ $\hat{r} = 0$ ความแปรปรวนของความคลาดเคลื่อนจะเหมือนกับการเทียบเชิงเส้นตรงในรูปแบบที่ 1 จากสมการ (14) ถึง (17) ถ้ากลุ่ม α และกลุ่ม β มีค่าเฉลี่ยในฟอร์ม U เท่ากันแล้ว ค่าในวงเล็บของสมการ (14),(15) จะพบว่าเป็นศูนย์ นั่นคือการปรับค่าความแตกต่างระหว่างกลุ่มก็ไม่จำเป็น และการประมาณค่าเฉลี่ยประชากรของคะแนน ฟอร์ม X และฟอร์ม Y ที่ดีที่สุดคือค่าเฉลี่ยที่สังเกตได้จริงๆของกลุ่ม α และกลุ่ม β

เป็นที่น่าสังเกตว่า สมการเหล่านี้ไม่มีข้อจำกัดเกี่ยวกับธรรมชาติของฟอร์ม U เลย การจะใช้ประโยชน์ของฟอร์ม U ในการเทียบมาตรานั้น จะขึ้นกับความสัมพันธ์ของฟอร์ม U กับฟอร์ม X และฟอร์ม Y ตัวอย่างเช่น ถ้า $r_{xu} = 0$ และ $r_{yu} = 0$ (เนื่องจาก X และ Y คู่ขนานกัน) จะเห็นได้ว่า คะแนนหรือค่าที่สังเกตได้ของฟอร์ม U ไม่เกี่ยวข้องกับการวัดโดยฟอร์ม X และฟอร์ม Y เลย ดังนั้นจึงไม่มีประโยชน์ในการปรับคะแนนโดยใช้คะแนนจากฟอร์ม U เลย

การเทียบมาตราด้วยวิธีนี้มีประโยชน์มาก สามารถยืดหยุ่นและปรับให้เข้ากับสถานการณ์อื่น ๆ ได้ เช่น ฟอร์ม U อาจจะสอบรวมหรือแยกจากฟอร์มX และฟอร์ม Y ฟอร์ม U อาจจะเป็นส่วนหนึ่งร่วมกับฟอร์มX แต่แยกจากฟอร์ม Y ช่วงเวลาการสอบฟอร์มU อาจจะแยกจากการสอบฟอร์มX และฟอร์ม Y และฟอร์ม U อาจจะรวมเข้าไปภายในฟอร์ม X และ ฟอร์ม Y โดยกระจายทั่วไป ตลอดทั้งฉบับ อย่างไรก็ตาม แบบสอบร่วมควรจะมีความยาวและความเที่ยงเพียงพอที่จะให้ข้อมูลไปปรับความแตกต่างระหว่างกลุ่มตามที่คาดหวัง ซึ่งแองกอฟฟ (angoff 1971: 578) ได้เสนอกฎเกณฑ์เกี่ยวกับจำนวนข้อของแบบสอบร่วม คือ ไม่ควรน้อยกว่า 20 ข้อ หรือ 20 % ของจำนวนข้อในแต่ละฟอร์มแล้วแต่จำนวนใดจะมากกว่าให้ใช้จำนวนนั้น และเพื่อที่จะหลีกเลี่ยงเกี่ยวกับแบบสอบที่ใช้วัดความเร็ว ควรกระจายฟอร์ม U ปะปนไปในทั้งฟอร์มx และ Y หลักเกณฑ์ทั่วๆไปเกี่ยวกับการใช้ฟอร์ม U กับกลุ่ม α และกลุ่ม β คือจะต้องเป็นการสอบเหมือนกันทั้งสองกลุ่ม ในกรณีของการแยกสอบคนละเวลา ตัวอย่างเช่น จัดสอบภายหลังฟอร์ม X และฟอร์ม Y หรือ จัดสอบก่อนฟอร์ม X และฟอร์ม Y เพราะจะได้มีผลเท่ากันในทางปฏิบัติและข้อกระทงในฟอร์ม Uไม่ควรซ้ำกันกับข้อกระทงของฟอร์ม X และฟอร์ม Y

การเทียบมาตราด้วยรูปแบบที่ 3 นี้ ยังมีวิธีการที่ยืดหยุ่นได้ คือ สามารถเทียบแบบสอบที่มากกว่า 2 ฟอร์มขึ้นไป อาจจะเป็น 3 ฟอร์ม 4 ฟอร์ม หรือมากกว่านั้น เช่น กรณี 3 ฟอร์ม คือฟอร์ม X,Y และ Z แบบสอบทั้งสามฟอร์มนี้สอบโดยกลุ่ม α, β และ γ ตามลำดับ ค่าเฉลี่ยและความแปรปรวนสำหรับกลุ่มรวม $(\alpha+\beta+\gamma)$ จะประมาณได้โดยใช้ข้อตกลงเกี่ยวกับการเทียบมาตราของแบบสอบสองฟอร์ม

แบบสอบร่วม ฟอร์ม U ไม่จำเป็นที่จะต้องเป็นตัวแปรเดียวเสมอไป อาจจะเป็น 2,3,4, ... หรือเท่ากับคะแนนที่ใช้ปรับระหว่างกลุ่มตัวอย่าง เช่น ฟอร์ม X และฟอร์ม Y เป็นแบบสอบที่ประกอบด้วย ข้อกระทงที่เป็นทั้งภาษาและคณิตศาสตร์ด้วย คะแนนของข้อกระทงทั้งสองชนิดอาจจะเป็นคะแนนรวมกันคือ ฟอร์ม U จะประยุกต์ใช้วิธีการเดียวกับที่กล่าวมาแล้ว และอีกกรณีหนึ่งคือแยกเป็นคะแนนฟอร์ม V และฟอร์ม M แล้วใช้การรวมแบบ (multiple combination) สมการที่ใช้สำหรับ

ประมาณค่าเฉลี่ยและความแปรปรวนของกลุ่ม t คือ

$$\hat{\mu}_x = M_{x\alpha} + b_{xv\cdot m\alpha}(\hat{\mu}_v - M_{v\alpha}) + b_{xm\cdot v\alpha}(\hat{\mu}_m - M_{m\alpha}) \quad (19)$$

$$\hat{\mu}_y = M_{y\beta} + b_{yv\cdot m\beta}(\hat{\mu}_v - M_{v\beta}) + b_{ym\cdot v\beta}(U_m - M_{m\beta}) \quad (20)$$

$$\hat{\sigma}_x = S^2_{x\alpha} + b^2_{xv\cdot m\alpha}(\hat{\sigma}^2_v - S^2_{v\alpha}) + b^2_{xm\cdot v\alpha}(\hat{\sigma}^2_m - S^2_{m\alpha}) + 2b_{xv\cdot m\alpha}b_{xm\cdot v\alpha}(\hat{\sigma}_{vm} - S_{vm\alpha}) \quad (21)$$

$$\hat{\sigma}_y = S^2_{y\beta} + b^2_{yv\cdot m\beta}(\hat{\sigma}^2_v - S^2_{v\beta}) + b^2_{ym\cdot v\beta}(\hat{\sigma}^2_m - S^2_{m\beta}) + 2b_{yv\cdot m\beta}b_{ym\cdot v\beta}(\hat{\sigma}_{vm} - S_{vm\beta}) \quad (22)$$

เมื่อ $b_{xv\cdot m\alpha}$ คือสัมประสิทธิ์กระถดถอยสำหรับการทำนาย X จาก V โดยที่ M เป็นค่าคงที่ และ $S_{vm\alpha} = r_{vm\alpha}\cdot S_{v\alpha}\cdot S_{m\alpha}$,

$\hat{\mu}_v = M_{vt}$, $\hat{\mu}_m = M_{mt}$, $\hat{\sigma}^2_v = S^2_{vt}$,

$\hat{\sigma}^2_m = S^2_{mt}$, $\hat{\sigma}_{vm} = S_{vmt}$

ซึ่งอาจกล่าวได้ว่า การประมาณค่าพารามิเตอร์ของประชากรสำหรับตัวแปรร่วม V และ M ได้มาโดยตรงจากค่าสถิติที่สังเกตได้สำหรับกลุ่มรวม t ($t = \alpha + \beta$)

อีกกรณีหนึ่งซึ่งยืดหยุ่นได้คือ แบบสอบร่วมฟอร์ม U ที่สอบโดยกลุ่ม α ไม่จำเป็นต้องเหมือนกันกับฟอร์ม U ที่สอบโดยกลุ่ม β อาจจะเป็นแบบสอบร่วมครึ่งหนึ่ง (quasi-common test) หรืออาจจะเป็นแบบสอบสองฟอร์มที่แตกต่างกัน (ฟอร์ม U และ W) แต่ทั้งสองฟอร์มต้องวัดในเรื่องเดียวกัน กระบวนการในการเทียบมาตราคือ แปลงแบบสอบร่วมฟอร์ม W ให้อยู่ในสเกลเดียวกับฟอร์ม U ก่อน หรือทั้งสองฟอร์มอาจจะแปลงไปเป็นสเกลใหม่เลยก็ได้แล้วจึงใช้วิธีการที่กล่าวมาแล้วตอนต้นปรับความแตกต่างระหว่างกลุ่ม

2. <u>การเทียบเชิงเส้นตรงเมื่อแบบสอบมีความเที่ยงไม่เท่ากัน</u>

เลวิน (Levine 1955) ได้เสนอไว้ในกรณีของกลุ่มสุ่ม แบบสอบฟอร์ม X และฟอร์ม Y มีความเที่ยงไม่เท่ากัน ซึ่งเหมาะสมกับการแปลงโดยคะแนนจริงมากกว่าคะแนนที่สอบได้ ข้อตกลงที่เพิ่มเติมขึ้นอีกคือ ฟอร์ม U จะต้องมีลักษณะคู่ขนานกันทางโครงสร้างกับทั้งฟอร์ม X และฟอร์ม Y

ภายใต้เงื่อนไขนี้ เมื่อฟอร์ม U แยกออกมาจากฟอร์ม X และ Y ความชัน (slope) และจุดตัดแกน (intercept) ของสมการ $y = AX + B$ คือ

$$A = b_{yu\beta} / b_{xu\alpha}$$

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

เมื่อ $b_{xu\alpha}$ และ $b_{yu\beta}$ คือสัมประสิทธิกระถดถอย (regression coefficient) ที่พบในกลุ่ม α และ β สำหรับการทำนาย X จาก U และ Y จาก U

เมื่อเป็นลักษณะแบบสอบร่วมภายใน (internal anchor test) คือ ฟอร์ม U รวมเข้าไปเป็นส่วนหนึ่งของฟอร์ม X และฟอร์ม Y

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

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

ค่าของ $\hat{\mu}_x$, $\hat{\mu}_y$, σ^2_x , σ^2_y คำนวณเช่นเดียวกับสมการ (14) ถึง (17)

<u>รูปแบบที่</u> 4 Nonrandom group - one test to each group , Common equating test administered to both groups

สถานการณ์ในการสอบบางครั้งถือเป็นความลับที่ไม่เปิดเผย ไม่สามารถที่จะแนะนำแบบสอบฟอร์มใหม่ก่อนที่จะสอบได้ และคำสั่งของโปรแกรมการสอบไม่อนุญาตให้อธิบายนอกเหนือไปกว่าที่สอบ ภายใต้สถานการณ์เหล่านี้ การเทียบมาตราโดยการสุ่มจาก 3 วิธีแรกไม่สามารถนำมาใช้ได้ และข้อมูลที่ใช้สำหรับเทียบมาตราต้องมาจากผลการสอบของกลุ่มที่สอบจริงๆแต่ละฟอร์ม ซึ่งการสอบต่างฟอร์มกันและเวลาสอบต่างกันผู้สอบอาจจะไม่ได้มาจากกลุ่มประชากรเดียวกัน ตัวอย่างเช่น มีการจัดสอบฟอร์มใหม่ในเดือนกันยายน 2527 และต้องการเทียบกับคะแนนของแบบสอบฟอร์มเก่าที่จัดสอบในเดือนกันยายน 2526 ซึ่งผู้สอบอาจจะมีลักษณะเหมือนกันหลายๆประการ แต่ยังไม่แน่ใจว่า กลุ่มผู้สอบทั้งสองฟอร์มมาจากประชากรเดียวกัน ดังนั้นเมื่อมีความระมัดระวังเกิดขึ้นดังตัวอย่าง การเลือกกลุ่ม α และ β เพื่อที่จะลดความแตกต่างให้น้อยลงใช้วิธีการ คือ

1. การเทียบเชิงเส้นตรงสำหรับกลุ่มที่มีความสามารถแตกต่างกันน้อย

วิธีการที่จะเสนอต่อไปนี้เป็นการรวมวิธีการที่เสนอโดยลอร์ด (Lord 1955a) ที่ใช้วิธีหาความเป็นไปได้สูงสุด (Maximum likelihood method) กับวิธีการที่เสนอโดย เลวิน (Levine 1955) ที่ใช้การปรับด้วยคะแนนจริง (true-score adaptation) โดยมีรูปแบบดังต่อไปนี้

กลุ่ม α สอบฟอร์ม X กลุ่ม β สอบฟอร์ม Y และทั้งสองกลุ่มสอบฟอร์ม V ซึ่งเป็นแบบสอบร่วม และใช้ปรับความแตกต่างที่อาจจะเกิดขึ้นระหว่างกลุ่ม การประมาณค่าเฉลี่ยและความแปรปรวนของทั้งฟอร์ม X และฟอร์ม Y กระทำโดยรวมกลุ่ม α และ β เป็นกลุ่ม t และประยุกต์ใช้กับสมการที่ 1 ซึ่งเป็นสมการเชิงเส้นที่เป็นความสัมพันธ์ของคะแนนดิบฟอร์ม X และฟอร์ม Y (Gulliksen 1950 chap 19 and angoff 1961a) สมการที่จะเสนอต่อไปจะขึ้นกับข้อตกลงเบื้องต้น 3 ประการของทฤษฎี Univariate selection theory นั่นคือ

1. จุดตัดแกนของคะแนนฟอร์ม X กับฟอร์ม V โดยกลุ่ม t เหมือนกับกลุ่ม α

$$M_{xt} - b_{xut}M_{ut} = M_{x\alpha} - b_{xu\alpha}M_{u\alpha} \qquad (23)$$

2. สัมประสิทธิ์กระถดถอย (regression coefficient) ของ X กับ V สำหรับกลุ่ม t เหมือนกับกลุ่ม α

$$b_{xut} = b_{xu\alpha} \qquad (24)$$

3. ความแปรปรวนของความคลาดเคลื่อนในการประมาณค่าของ X จากกลุ่ม t เหมือนกับกลุ่ม α

$$S^2_{xt}(1 - r^2_{xut}) = S^2_x (1 - r^2_{xu}) \qquad (25)$$

แทนค่าสมการ (24) ลงใน (23) เพื่อหาค่า $\hat{M}_{xt}$

$$\hat{M}_{xt} = M_{x\alpha} + b_{xu\alpha}(M_{ut} - M_{u\alpha}) \qquad (26)$$

(สัญญลักษณ์ ($\wedge$) ใช้แสดงค่าประมาณ)

แทนค่าลงในสมการ (25) โดยที่ $b_{xut}S_{ut}$ แทน $r_{xut}S_{xt}$ และ $b_{xu\alpha}S_{u\alpha}$ แทน $r_{xu\alpha}S_{x\alpha}$

ด้วยข้อตกลงเบื้องต้นลักษณะเดียวกันของฟอร์ม X และฟอร์ม Y นำมาหาความสัมพันธ์ระหว่าง ฟอร์ม Y และฟอร์ม U กลุ่ม β และกลุ่ม t ได้ผลลักษณะเช่นเดียวกับสมการ (26),(27) คือ

$$\hat{M}_{yt} = M_{y\beta} + b_{yu\beta}(M_{ut} - M_{u\beta}) \qquad (28)$$

และ

$$\hat{S}^2_{yt} = S^2_{y\beta} + b^2_{yu\beta}(S^2_{ut} - S^2_{u\beta}) \qquad (29)$$

(สัญญลักษณ์ $\hat{M}_{xt}$, $\hat{S}^2_{xt}$, $\hat{M}_{yt}$ และ $\hat{S}^2_{yt}$ ที่ใช้ในสมการ (26) ถึง (29) แทนสัญญลักษณ์ $\hat{\mu}_x$, $\hat{\sigma}^2_x$, $\hat{\mu}_y$ และ $\hat{\sigma}^2_y$ ตามลำดับ เพราะว่าในที่นี้เป็นการประมาณค่ากลุ่มรวม ไม่ใช่ประมาณค่าสำหรับประชากร)

สมการ (26) ถึง (29) แทนค่าในสมการ (1) คือ

$$(y - M_y) / s_y = (x - M_x) / S_x$$

ซึ่งจะได้สมการสำหรับแปลงคะแนน $y = Ax + B$ โดยที่

$$A = \hat{S}_{yt} / \hat{S}_{xt}$$

$$B = \hat{M}_{yt} - A\hat{M}_{xt}$$

เป็นที่สังเกตได้ว่า วิธีการคำนวณสำหรับการประมาณค่าในสมการ (26)-(29) เหมือนกับสมการ (14)-(17) ตามลำดับ แม้ว่าสมการทั้ง 2 ชุดจะแตกต่างกัน

ลักษณะเช่นเดียวกันนี้ สามารถที่จะยืดหยุ่นใช้กับสถานการณ์อื่นได้อีก เช่น ฟอร์ม U อาจจะสอบรวมหรือแยกจากฟอร์ม X และฟอร์ม Y หรือเป็นส่วนหนึ่งของฟอร์ม X และเป็นส่วนหนึ่งของฟอร์ม Y หรืออาจแยกจากฟอร์ม X แต่เป็นส่วนหนึ่งของฟอร์ม Y ก็ได้ เกี่ยวกับเวลาในการสอบ อาจจะแยกสอบฟอร์ม U จากฟอร์ม X และฟอร์ม Y หรืออาจจะกระจายข้อกระทงของฟอร์ม U ปะปนกับฟอร์ม X และฟอร์ม Y ก็ได้ แต่ต้องสามารถหาคะแนนรวมของแต่ละฟอร์มได้ถูกต้อง

วิธีการประมาณค่าความเป็นไปได้สูงสุด (Maximum likelihood) นี้ ไม่จำกัดเฉพาะการเทียบแบบสอบ 2 ฟอร์มเท่านั้น แต่สามารถเทียบแบบสอบตั้งแต่ 3 ฟอร์มขึ้นไปได้ ซึ่งในการบริหารแบบสอบแต่ละฟอร์มอาจจะสอบแยกกลุ่มกัน แต่ฟอร์ม U ซึ่งเป็นแบบสอบร่วมต้องสอบทุกกลุ่ม แล้วการประมาณค่าเฉลี่ย และความแปรปรวนของแต่

ละฟอร์มที่จะเทียบกันก็ใช้กลุ่มรวมสอบฟอร์ม U ซึ่งสมการ (26)-(29) สามารถใช้หลายๆครั้งเท่ากับจำนวนแบบสอบที่จะมาเทียบกัน และอีกประการหนึ่ง ฟอร์ม U ก็ไม่จำกัดเฉพาะการวัดตัวแปรเดียว แต่อาจจะใช้ได้หลายๆตัวแปรซึ่งมีลักษณะเดียวกับสมการ (19) ถึง (22) แบบสอบร่วมอาจจะเป็นลักษณะครึ่งหนึ่ง (quasi-common) ซึ่งอาจจะกล่าวได้ว่าเป็นแบบสอบที่แตกต่างกัน 2 ฟอร์ม ฟอร์มหนึ่งสอบกับกลุ่ม α และอีกฟอร์มหนึ่งสอบกับกลุ่ม β มีข้อจำกัดอยู่ว่าทั้งสองฟอร์มจะต้องวัดเรื่องเดียวกัน คะแนนอยู่บนสเกลเดียวกัน

2. การเทียบเชิงเส้นตรงสำหรับกลุ่มที่มีความสามารถแตกต่างกัน

2.1 แบบสอบที่มีความเที่ยงเท่าเทียมกัน

เลวิน (Levine 1955) ได้เสนอว่าเมื่อกลุ่ม α และ β มีความสามารถแตกต่างกันมาก ข้อตกลงเบื้องต้นที่เป็นพื้นฐานของทฤษฎีการเลือก (Classical selection theory) นั้นไม่มีความเหมาะสมกับสถานการณ์นี้ ควรใช้กับข้อตกลงเบื้องต้นแบบอื่นแต่ภายใต้ข้อจำกัดที่ว่า ฟอร์ม U มีลักษณะคู่ขนานกับทั้งฟอร์ม X และฟอร์ม Y นั่นคือ

1. จุดตัดแกนของเส้นกระถดถอย (regression line) ที่เกี่ยวกับคะแนนจริงของฟอร์ม X กับฟอร์ม U โดยกลุ่ม t เหมือนกับกลุ่ม α

$$M_{xt} - \frac{S_{\tilde{x}t}}{S_{\tilde{u}t}} M_{ut} = M_{x\alpha} - \frac{S_{\tilde{x}\alpha}}{S_{\tilde{u}\alpha}} M_{u\alpha} \qquad (30)$$

โดยที่ $S_{\tilde{x}} = S_x \sqrt{r_{xx}}$ และ $S_{\tilde{u}} = S_u \sqrt{r_{uu}}$

2. ความชันของเส้นความสัมพันธ์สำหรับกลุ่ม t เหมือนกับกลุ่ม α

$$\frac{S_{\tilde{x}t}}{S_{\tilde{u}t}} = \frac{S_{\tilde{x}\alpha}}{S_{\tilde{u}\alpha}} \qquad (31)$$

3. ความแปรปรวนของความคลาดเคลื่อนของการวัดจากฟอร์ม X กลุ่ม t เหมือนกับกลุ่ม α

$$S^2_{xt}(1 - r_{xxt}) = S^2_{x\alpha}(1 - r_{xx\alpha}) \qquad (32)$$

จากสมการ (30),(31),(32) จะได้

$$M_{xt} = M_{x\alpha} + \frac{S_{\tilde{x}\alpha}}{S_{u\alpha}}(M_{ut} - M_{u\alpha}) \qquad (33)$$

$$\hat{S}^2_{xt} = S^2_{x\alpha} + \frac{S^2_{\tilde{x}\alpha}}{S^2_{\tilde{u}\alpha}}(S^2_{ut} - S^2_{u\alpha}) \qquad (34)$$

ด้วยข้อตกลงเบื้องต้นแบบเดียวกันสำหรับฟอร์ม X และฟอร์ม Y จะได้

$$M_{yt} = M_{y\beta} + \frac{S_{\tilde{y}\beta}}{S_{\tilde{u}\beta}}(M_{ut} - M_{u\beta}) \qquad (35)$$

$$S^2_{yt} = S^2_{y\beta} + \frac{S^2_{\tilde{y}\alpha}}{S^2_{\tilde{u}\alpha}}(S^2_{ut} - S^2_{u\beta}) \qquad (36)$$

สมการที่แสดงความสัมพันธ์ระหว่างคะแนนของฟอร์ม X และฟอร์ม Y คือ

$$Y = AX + B$$

โดยที่ $A = \hat{S}_{yt} / \hat{S}_{xt}$ $\quad B = \hat{M}_{yt} - A\hat{M}_{xt}$

เพื่อเป็นการง่ายในการคำนวณ แองกอฟฟ์ (Angoff,1953) ได้แสดงวิธีการที่ได้จากการทดลองของเขาโดยใช้การประมาณค่าอัตราส่วนของความเบี่ยงเบนมาตรฐานของคะแนนจริง นั่นคือ $S_{\tilde{x}} / S_{\tilde{u}} = U$ ซึ่งเป็นอัตราส่วนของความยาวแบบสอบฟอร์ม X กับฟอร์ม U เมื่อฟอร์ม U เป็นการสอบรวมในฟอร์ม X

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

เมื่อฟอร์ม U เป็นการสอบแยกจากฟอร์ม X

$$n = (S^2_x + S_{ux}) / (S^2_u + S_{ux})$$

ซึ่งเมื่อใช้กับฟอร์ม Y ก็จะได้สมการเช่นเดียวกัน

2.2 <u>แบบสอบที่มีความเที่ยงไม่เท่ากัน</u>

เมื่อฟอร์ม U สอบแยกจากฟอร์ม X และฟอร์ม Y จะได้

$$y = Ax + B$$

เมื่อ

$$A = \frac{b_{yu\beta}r_{uu\alpha}}{b_{xu\alpha}r_{uu\beta}}$$

$$B = M_{y\beta} - AM_{x\alpha} + \frac{b_{yu\beta}(M_{u\alpha} - M_{u\beta})}{r_{uu\beta}}$$

เมื่อฟอร์ม u สอบรวมกับฟอร์ม x และฟอร์ม Y จะได้

$$A = \frac{b_{ux\alpha}}{b_{uy\beta}}$$

$$B = M_{y\beta} - AM_{x\alpha} + \frac{(M_{u\alpha} - M_{u\beta})}{b_{uy\beta}}$$

3. การเทียบโดยใช้เปอร์เซนไทล์สำหรับกลุ่มที่ความสามารถไม่แตกต่างกันมาก

ลอร์ด (Lord 1957) และเลวิน (Levine 1958) ได้เสนอวิธีการเทียบโดยใช้ตำแหน่งเปอร์เซนไทล์ซึ่งมาจากพื้นฐานข้อตกลงเบื้องต้นในสมการที่ (23),(24) และ (25) โดยที่กลุ่ม α สอบฟอร์ม x และฟอร์ม u กลุ่ม β สอบฟอร์ม Y และฟอร์ม u ขั้นตอนของกระบวนการเทียบมาตราคือ ประมาณค่าความถี่ในแบบสอบฟอร์ม x และ Y สำหรับกลุ่มรวม t ($\alpha+\beta$) โดยมีวิธีการดังนี้

1. รวมคะแนนของฟอร์ม u ที่สอบโดยกลุ่ม α, β และกลุ่ม t
2. หาสัดส่วนความถี่ $f_{it}/f_{i\alpha}$ และ $f_{it}/f_{i\beta}$ ที่ทุกช่วงคะแนน i
3. กระจายความถี่ของฟอร์ม u ที่ทุกช่วงคะแนนของฟอร์ม x และ Y
4. คูณความถี่ของฟอร์ม u ในแต่ละช่วงคะแนนของฟอร์ม x ด้วยสัดส่วน $f_{it} / f_{i\alpha}$
5. คูณความถี่ของฟอร์ม u ในแต่ละช่วงคะแนนของฟอร์ม Y ด้วยสัดส่วน $f_{it} / f_{i\beta}$
6. หาตำแหน่งเปอร์เซนไทล์ของความถี่ที่ประมาณค่าใหม่ของทั้งสองชุด
7. ขั้นตอนต่อไปดำเนินวิธีการเทียบโดยเปอร์เซนไทล์เช่นเดียวกับกลุ่มสุ่มที่กล่าวแล้วข้างต้น

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

รูปแบบที่ 5 Other methods involving score data

1. แบบสอบฟอร์ม X และฟอร์ม Y เทียบไปสู่แบบสอบร่วม

ก. การเทียบเชิงเส้นตรง

วิธีการนี้เกี่ยวข้องกับการบริหารแบบสอบร่วมฟอร์ม U ซึ่งอาจจะสอบฟอร์ม U ก่อนแล้วตามด้วยฟอร์ม X และฟอร์ม Y หรืออาจจะสอบฟอร์ม X และฟอร์ม Y ก่อนแล้วตามด้วยฟอร์ม U ถ้าหากฟอร์ม X และฟอร์ม Y เป็นแบบสอบที่มีลักษณะคู่ขนานกันก็มีเหตุผลที่จะถือได้ว่า แต่ละฟอร์มมีผลต่อฟอร์ม U เหมือนกันเมื่อฟอร์ม U สอบภายหลัง หรือฟอร์ม U มีผลต่อฟอร์ม X และฟอร์ม Y เหมือนกันเมื่อฟอร์ม U สอบก่อนฟอร์ม X และฟอร์ม Y การเทียบมาตราโดยวิธีการนี้ก็คือ เทียบจากฟอร์ม X ไปสู่ฟอร์ม U และเทียบจากฟอร์ม Y ไปสู่ฟอร์ม U ซึ่งคะแนนที่เทียบแล้วของทั้งสองฟอร์มที่ระดับคะแนนของฟอร์ม U เดียวกันถือว่าเท่าเทียมกัน ดังนั้น

ถ้า $x = A_{xu}U + B_{xu}$ เมื่อ $A_{xu} = S_{x\alpha} / S_{u\alpha}$

$B_{xu} = M_{x\alpha} - A_{xu}M_{u\alpha}$

และ ถ้า $y = A_{yu}U + B_{yu}$ เมื่อ $A_{yu} = S_{y\beta} / S_{u\beta}$

$B_{yu} = M_{y\beta} - A_{yu} / A_{xu\beta}$

แล้วจะได้ $y = A_{yx}X + B_{yx}$ เมื่อ $A_{yx} = A_{yu} / A_{xu}$

$B_{yx} = B_{yu} - A_{yx}B_{xu}$

ในการที่จะประกันถึงความเหมาะสมของสมการแปลง $y = A_{yx}X + B_{yx}$ ฟอร์ม U ต้องมีลักษณะคู่ขนานกับฟอร์ม X และ Y โดยที่มีอิสระในการเลือกกลุ่ม α และ β แต่ถ้าฟอร์ม U ไม่มีลักษณะคู่ขนานกับฟอร์ม X และ Y แล้วกลุ่ม α และ β ต้องมาจากการสุ่มจากประชากรเดียวกัน

ภายใต้เงื่อนไขของการสุ่มตัวอย่าง ลอร์ด (Lord 1950) ได้เสนอความแปรปรวนของความคลาดเคลื่อนของการเทียบมาตราของวิธีนี้ คือ

$$S.E._{y\cdot}^2 = 4S_y^2 (1-r)\frac{Z_x^2(1 + r) + 2}{N_t} \qquad (37)$$

เมื่อ $r = r_{xu\alpha} = r_{yu\beta}$

ข. การเทียบเชิงเส้นโค้ง

การเทียบเชิงเส้นโค้งก็มีลักษณะเช่นเดียวกับที่ได้อธิบายไว้ก่อนแล้ว แต่มีข้อแตกต่างกันคือ เทียบระหว่างแบบสอบร่วมฟอร์ม U กับฟอร์ม X โดยใช้ตำแหน่งเปอร์เซนไทล์เดียวกัน และใช้วิธีเดียวกันเทียบระหว่างฟอร์ม U กับฟอร์ม Y แล้วจึงนำคะแนนฟอร์ม X และฟอร์ม Y ที่มีตำแหน่งเปอร์เซนไทล์บนสเกลของฟอร์ม U มาพล็อตกราฟและสร้างตารางแปลงคะแนน

2. <u>แบบสอบร่วมฟอร์ม U เป็นตัวทำนาย (predict) ฟอร์ม X และฟอร์ม Y</u>

ก. การเทียบเชิงเส้นตรง

วิธีการนี้การบริหารแบบสอบก็เช่นเดียวกัน คือ กลุ่ม α สอบฟอร์ม X กลุ่ม β สอบฟอร์ม Y และทั้งสองกลุ่มสอบฟอร์ม U ซึ่งวิธีการในการกำหนดคะแนนที่เทียบเคียงกันทั้ง 2 ฟอร์ม โดยการทำนายของคะแนนฟอร์ม U เดียวกัน ดังนั้น

ถ้า $\hat{X} = b_{xu}U + D_{xu}$ เมื่อ $b_{xu} = r_{xy\alpha}(S_{x\alpha} / S_{u\alpha})$

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

และถ้า $\hat{Y} = b_{yu}U + D_{yu}$ เมื่อ $b_{yu} = r_{yu\beta}(S_{y\beta} / S_{u\beta})$

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

แล้วจะได้ $y = Ax + B$ เมื่อ $A = b_{yu} / b_{xu}$ (38)

$B = D_{yu} / AD_{xu}$ (39)

ข. การเทียบเชิงเส้นโค้ง

ในการเทียบเชิงเส้นโค้งโดยใช้ตำแหน่งเปอร์เซนไทล์ ใช้วิธีการดังนี้ คือ ที่ทุก ๆ ช่วงคะแนนของฟอร์ม U หาความถี่ของฟอร์ม X ที่กระจายทุกช่วงคะแนนแล้วหาค่าเฉลี่ยของคะแนน X ขั้นที่สองทำเช่นเดียวกันคือกระจายคะแนนฟอร์ม Y แล้วหาความถี่และค่าเฉลี่ยของคะแนน Y ขั้นที่สาม นำค่าเฉลี่ยของทั้งสองสเกลมาพล็อตกราฟ ซึ่งค่าเฉลี่ย $M_{x.u}$ กับ $M_{y.u}$ จะเป็นตัวกำหนดความสัมพันธ์ระหว่างคะแนนฟอร์ม X กับคะแนนฟอร์ม Y

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

I 15536488

3. แบบสอบฟอร์ม X และฟอร์ม Y เป็นตัวทำนายแบบสอบร่วม

ก. การเทียบเชิงเส้นตรง

กลุ่ม α สอบฟอร์ม X กลุ่ม β สอบฟอร์ม Y แล้วทั้งสองกลุ่มสอบฟอร์ม U

ถ้า $\hat{U} = b_{ux}X + D_{ux}$ เมื่อ $b_{ux} = r_{ux\alpha}(S_{u\alpha} / S_{x\alpha})$

$D_{ux} = M_{u\alpha}$

และถ้า $\hat{U} = b_{uy}Y + D_{uy}$ เมื่อ $b_{uy} = r_{uy\beta}(S_{u\beta} / S_{y\beta})$

$D_{uy} = M_{u\beta}$

แล้วจะได้ $y = Ax + B$ เมื่อ $A = b_{ux} / b_{uy}$ (40)

$B = (D_{ux} - D_{uy}) / b_{uy}$ (41)

ข. การเทียบเชิงเส้นโค้ง

ใช้ตำแหน่งเปอร์เซนไทล์ที่เท่ากันเทียบโดยมีขั้นตอนดังนี้ ที่ทุกๆช่วงคะแนนของฟอร์ม X หาความถี่และค่าเฉลี่ยของฟอร์ม U เมื่อกำหนดให้ X คงที่ และที่ทุกๆช่วงคะแนนของฟอร์ม Y หาความถี่และค่าเฉลี่ยของฟอร์ม U เมื่อกำหนดให้ Y คงที่ หาความถี่สะสมและตำแหน่งเปอร์เซนไทล์ของทั้งสองสเกลแล้วจึงพล็อตกราฟ ค่าของ X และ Y อ่านจากกราฟที่มีค่า U เท่ากันเป็นตัวกำหนดความสัมพันธ์

รูปแบบที่ 6 การเทียบมาตราที่ขึ้นกับลักษณะข้อกระทง

เธอร์สโตน (Thurstone 1925)และ แฟน (Fan 1957) ได้เสนอวิธีการนี้ โดยมีสถานการณ์คือ กลุ่ม α สอบฟอร์ม X กลุ่ม β สอบฟอร์ม Y แบบสอบฟอร์ม X และแบบสอบฟอร์ม Y มีข้อสอบชุดหนึ่งซึ่งมีลักษณะร่วมกันคือ ค่าความยาก (P) ซึ่งต่างจากค่า P ทั่วๆไป คือ แปลงให้มีลักษณะเช่นเดียวกับการกระจายแบบปกติ (z') วิธีนี้เช่นเดียวกับวิธีอื่นๆ โดยมีข้อตกลงเบื้องต้นว่า ฟอร์ม X และฟอร์ม Y มีลักษณะคู่ขนานกัน ดังนั้นจึงสามารถแปลงเป็นสเกลร่วมได้ และข้อตกลงอีกประการคือ กลุ่ม α และ β มีลักษณะการกระจายเป็นแบบปกติ

จุดประสงค์ของวิธีนี้คือหาความสัมพันธ์ระหว่างความยากของข้อกระทงที่ร่วมกันของทั้งสองกลุ่มที่จะนำไปสู่สมการแปลงคะแนนดิบของฟอร์ม X ไปสู่ฟอร์ม Y นั่นคือ

$$Y = Ax + B$$

เมื่อ $A = S_{Y\alpha} / S_{X\alpha} \quad (= S_{Y\beta} / S_{X\beta})$

$B = M_{Y\alpha} - AM_{X\alpha} \quad (= M_{Y\beta} - AM_{X\beta})$

ถ้าการกระจายของคะแนนความสามารถเป็นโค้งปกติ ภายในกลุ่ม α และ β แล้ว ความสัมพันธ์ระหว่างข้อกระทงและความสามารถ คือ

$$Z_{i\alpha} = \frac{Y_i - M_{Y\alpha}}{S_{Y\alpha}} \qquad (42)$$

และ

$$Z_{i\beta} = \frac{Y_i - M_{Y\beta}}{S_{Y\beta}} \qquad (43)$$

เมื่อ $M_{Y\alpha}$, $M_{Y\beta}$ คือค่าเฉลี่ยของคะแนนความสามารถของกลุ่ม α และกลุ่ม β

$S_{Y\alpha}$, $S_{Y\beta}$ คือค่าเบี่ยงเบนมาตรฐานของคะแนนกลุ่ม α และกลุ่ม β

$Z_{i\alpha}$, $Z_{i\beta}$ คือค่าคะแนนมาตรฐานของตำแหน่งข้อกระทงที่ i ของสเกลสำหรับกลุ่ม α และกลุ่ม β

ให้สมการที่ (42) และ (43) เท่ากัน

$$Z_{i\beta} = (S_{Y\alpha} / S_{Y\beta}) Z_{i\alpha} + \frac{M_{Y\alpha} - M_{Y\beta}}{S_{Y\beta}} \qquad (44)$$

ถ้าพล็อตความเบี่ยงเบนปกติระหว่าง $Z'_{i\beta}$ กับ $Z'_{i\alpha}$ ความสัมพันธ์ระหว่างทั้งสองสามารถอธิบายได้ด้วย

$$Z'_{i\beta} = (S'_{\beta} / S'_{\alpha}) Z'_{i\alpha} + M'_{\beta} - (S'_{\beta} / S'_{\alpha}) M'_{\alpha} \qquad (45)$$

เมื่อ M'_{α} , M'_{β} คือค่าเฉลี่ยของความเบี่ยงเบนปกติของทั้งสองกลุ่ม

S'_{α} , S'_{β} คือค่าเบี่ยงเบนมาตรฐานของความเบี่ยงเบนปกติของทั้งสองกลุ่ม

สมมุติให้ สมการ (44) และ (45) เป็นการอธิบายความสัมพันธ์เหมือนกัน จึงสรุปได้ว่าความชันเท่ากัน และจุดตัดแกนเท่ากัน คือ

$$S_{Y\alpha} / S_{Y\beta} = S'_{\beta} / S'_{\alpha} \qquad (46)$$

$$(M_{Y\alpha} - M_{Y\beta}) / S_{Y\beta} = M'_{\beta} - (S'_{\beta} / S'_{\alpha}) M'_{\alpha} \qquad (47)$$

จากสมการ (46),(47) ค่าประมาณ $\hat{S}_{Y\alpha}$ และ $\hat{M}_{Y\alpha}$ ซึ่งจะใช้สำหรับคำนวณค่าความชัน คือ $A = \hat{S}_{Y\alpha} / S_{X\alpha}$

และจุดตัดแกน คือ $B = \hat{M}_{Y\alpha} - AM_{X\alpha}$

สำหรับสมการแปลงคะแนนจากฟอร์ม X ไปสู่ฟอร์ม Y หาได้จาก

$$\hat{S}_{Y\alpha} = S_{Y\beta}(S'_{\beta} / S'_{\alpha}) \text{ และ}$$

$$\hat{M}_{Y\alpha} = S_{Y\beta}(M'_{\beta} - (S'_{\beta} / S'_{\alpha})M'_{\alpha}) + M_{Y\beta}$$

2. โมเดลการเทียบมาตราโดยใช้ทฤษฎีสนองตอบข้อสอบ

2.1 วิธีการเทียบโดยใช้คะแนนจริง (True - score equating)

จากทฤษฎีคะแนนทราบว่าหากใช้แบบสอบที่มีระดับความยากต่างกัน ลักษณะการกระจายของคะแนนดิบจากแบบสอบสองฉบับที่ได้จากการสอบจากกลุ่มประชากรที่มีความสามารถหรือคุณลักษณะระดับเดียวกันจะมีลักษณะแตกต่างกัน หากเราสามารถหาค่าคะแนนจริงของผู้เข้าสอบที่ทำแบบสอบที่มีความยากต่างกันได้ จะพบว่าความสัมพันธ์ระหว่างคะแนนจริงของผู้เข้าสอบจากแบบสอบฉบับดังกล่าว จะมีลักษณะไม่เป็นเส้นตรง (nonlinear) ดังนั้นถ้าสามารถหาคะแนนจริงได้จากผลการสอบทั้งสองฉบับของบุคคลที่มีความสามารถหรือคุณลักษณะที่เท่ากัน (same θ) ก็จะสามารถทำการเทียบโดยคะแนนจริงได้

จากความสัมพันธ์เชิงคณิตศาสตร์ (Mathematical model) ระหว่างความสามารถ (θ) และคะแนนจริงของแบบสอบสองฉบับใด ๆ ในที่นี้กำหนดให้เป็นแบบสอบฟอร์ม X และฟอร์ม Y ความสัมพันธ์ดังกล่าว คือ

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

$$\eta \equiv \eta(\theta) \equiv \sum_{j=1}^{n_y} P_j(\theta)$$

ในเมื่อ ξ คือคะแนนจริงจากแบบสอบฟอร์ม X

η คือคะแนนจริงจากแบบสอบฟอร์ม Y

n_x คือจำนวนข้อสอบของแบบสอบฟอร์ม X

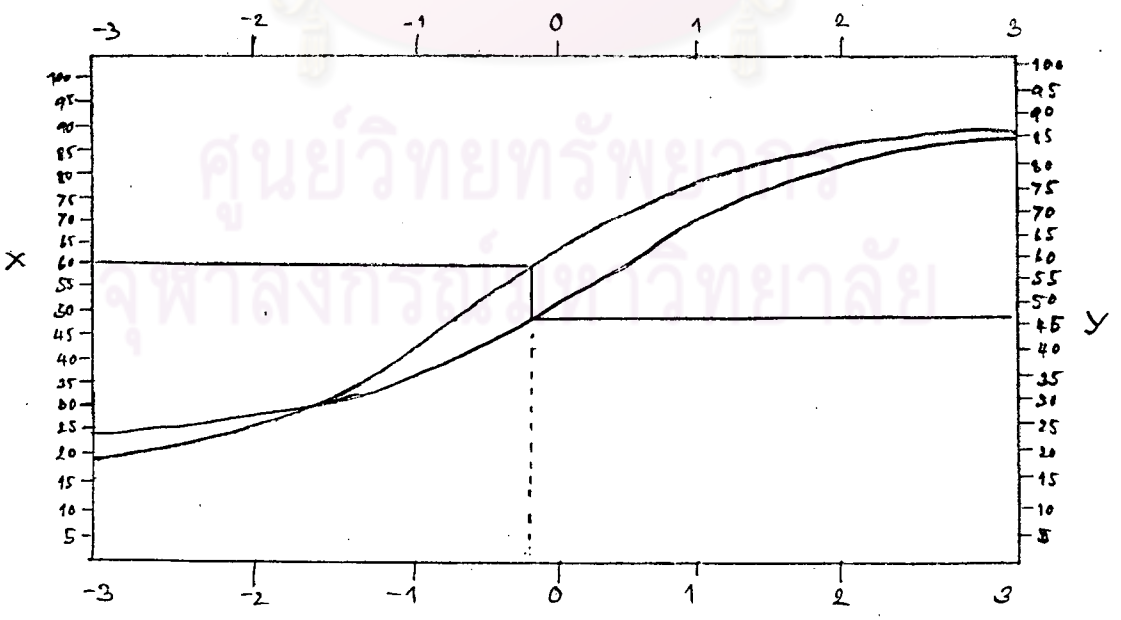
n_y คือจำนวนข้อสอบของแบบสอบฟอร์ม Y

สมการ (1) และ (2) เป็นสมการพาราเมตริก (parametric) ระหว่างคะแนนจริงของแบบสอบฟอร์ม X และฟอร์ม Y ดังนั้นหากแทนค่า (θ) ใดๆลงไปก็จะสามารถคำนวณค่าของ ξ และ η ได้ และในทางตรงกันข้ามก็สามารถประมาณค่า (θ) เมื่อทราบคะแนนจริงได้

ในทางปฏิบัติเราไม่สามารถหาค่า $P_i(\theta)$ และ $P_j(\theta)$ ได้ แต่สามารถประมาณค่าทั้งสองนี้ได้จากการใช้ค่าพารามิเตอร์ของข้อสอบแต่ละข้อ คือค่า a , b และ c จากสูตร

$$P_i(\theta) = C_i + (1 - C_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

จากนั้นก็สามารถนำคะแนนจริงของผลการสอบแบบสอบฟอร์ม X และฟอร์ม Y มาหาความสัมพันธ์กันได้โดยใช้ค่า θ ที่ระดับเดียวกันเป็นหลัก ดังภาพที่ 2



ภาพที่ 2 การเทียบมาตราโดยใช้คะแนนจริงตามทฤษฎีสนองตอบข้อสอบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2.2 วิธีการเทียบโดยใช้คะแนนดิบ (Raw score equating)

เนื่องจากปัญหาในทางปฏิบัติจริงไม่สามารถที่จะนำเอาลักษณะของการเทียบโดยใช้คะแนนจริงมาใช้ได้เพราะ ไม่สามารถทราบคะแนนจริงของผู้สอบแต่ละคนได้ นอกจากใช้วิธีประมาณจากผลการสอบโดยสมการ

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

อย่างไรก็ตาม คะแนนที่ได้ก็เป็นเพียงค่าประมาณเท่านั้นยังไม่มีคุณสมบัติเป็นคะแนนจริง ดังนั้นการเทียบมาตราโดยคะแนนจริงประมาณจึงเสมือนกับการเทียบมาตราโดยคะแนนดิบชนิดใหม่เท่านั้นเอง

สำหรับการเทียบมาตราโดยคะแนนดิบตามทฤษฎีสนองตอบข้อสอบนี้เป็นการอาศัยข้อมูลจากการตอบแบบสอบร่วม (Anchor test) และคุณสมบัติของข้อสอบแต่ละข้อมาประมาณลักษณะการกระจายของคะแนนจากฟอร์ม X และฟอร์ม Y ในแต่ละระดับความสามารถ (θ) และนำมาเทียบกันโดยยึดถือการกระจายร่วม (joint distribution) ของผลการตอบแบบสอบร่วม

การดำเนินการเทียบมาตราโดยวิธีนี้จะเริ่มต้นด้วยการประมาณลักษณะการกระจาย $r(\theta)$ ของความสามารถของกลุ่มรวม ซึ่งหมายถึงผู้สอบทั้งหมดที่ทำแบบสอบร่วม การกระจายของ (θ) ในกลุ่มเป็นการประมาณการกระจายของ $r(\theta)$

การประมาณการกระจายของคะแนนดิบจากแบบสอบ ในการกระจายของ θ ประมาณได้ด้วยสมการ

$$\hat{\phi}_x(x) = \frac{1}{N}\sum_{a=1}^{N} \hat{\phi}_x(x/\hat{\theta}_a)$$

โดยที่ $a = 1,2 \ldots\ldots N$ คือผู้สอบในแต่ละระดับ θ ของกลุ่มรวม

ถ้า $r(\theta)$ มีลักษณะต่อเนื่อง (Continuous) ก็สามารถประมาณได้ด้วยสมการ

$$\hat{\phi}_x(x) = \int_{-\infty}^{\infty} \phi_x(x/\theta) r(\theta)\, d\theta$$

โดยที่ $\phi_x(x/\hat{\theta}_a)$ ได้มาจากการประมาณโดยใช้ค่าพารามิเตอร์ของข้อกระทง (Item parameter) คือ (a,b,c) ของข้อสอบในแบบสอบฟอร์ม X

ในทำนองเดียวกัน สมการต่างๆสามารถนำมาประยุกต์ใช้กับแบบสอบฟอร์ม Y ได้เช่นเดียวกัน

เนื่องจากแบบสอบฟอร์ม X และฟอร์ม Y มีการกระจายของคะแนนเป็นอิสระจากกันเมื่อระดับความสามารถ (θ) คงที่ การเทียบมาตราโดยกรณีนี้จึงต้องอาศัยการกระจายร่วม (joint distribution) ของคะแนนจากแบบสอบฟอร์ม X และฟอร์ม Y เป็นตัวเชื่อมในการเทียบมาตรา ซึ่งการกระจายร่วมของความสามารถประมาณได้จากสมการ

$$\phi(x,y) = \frac{1}{N}\sum_{a=1}^{N}\phi_x(x/\theta_a)\phi_y(Y/\theta_a)$$

หรือ

$$\phi(x,y) = \int_{-\infty}^{\infty}\hat{\phi}_x(x/\theta)\hat{\phi}_y(y/\theta)\hat{r}(\theta)\,d\theta$$

จากนั้นก็สามารถอาศัยความสัมพันธ์ของการกระจายร่วมทำการเทียบมาตราโดยเทียบคะแนนดิบจากแบบสอบฟอร์ม X ไปสู่ฟอร์ม Y โดยวิธีการเทียบที่ใช้เปอร์เซนไทล์

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการเทียบมาตรา

จากการที่ผู้วิจัยได้ศึกษาผลงานวิจัยต่างๆที่ผ่านมา พบว่าในประเทศไทยมีผลงานวิจัยเกี่ยวกับการเทียบมาตรานอยมาก แต่อย่างไรก็ตามในต่างประเทศได้มีผลงานวิจัยเกี่ยวกับการเทียบมาตราไว้หลายชิ้น ซึ่งส่วนใหญ่แล้วเป็นการประเมินความถูกต้องของวิธีการเทียบมาตรา โดยพอจะแยกเป็น 2 พวกใหญ่ ๆ คือ พวกแรกศึกษาความถูกต้องของการเทียบมาตราในโครงการเดี่ยวๆ โดยทดสอบความคล้ายคลึงกันของผลที่ได้จากกลุ่มที่แตกต่างกัน ซึ่งกลุ่มอาจจะแตกต่างกันในลักษณะของความสามารถ (ability) สถานะทางเศรษฐกิจและสังคม (socio-economic status) และอื่นๆ การศึกษาแบบนี้ขึ้นกับพื้นฐานที่ว่าถ้าข้อตกลงของการเทียบแบบสอบที่ไม่คู่ขนานกันใช้ได้แล้ว การเทียบมาตราที่ขึ้นกับกลุ่มที่แตกต่างกันควรจะเหมือนกัน ซึ่งถ้าไม่เหมือนกันก็ควรจะมาจากความคลาดเคลื่อนในการสุ่ม (Sampling error) (Kolen 1981: 2) พวกที่สองได้ศึกษาเกี่ยวกับการเปรียบเทียบวิธีการเทียบมาตราวิธีต่างๆ ซึ่งผู้วิจัยจะนำเสนอแยกเป็น 2 ตอนดังต่อไปนี้

<u>การเทียบมาตราเมื่อใช้กลุ่มตัวอย่างแตกต่างกัน</u>

ในต่างประเทศได้มีการศึกษาหลายๆครั้ง ทดสอบผลการใช้ราสช์โมเดลกับกลุ่มตัวอย่างที่แตกต่างกันโดยใช้แบบสอบเป็นแบบ ข้อคู่-ข้อคี่ , ง่าย-ยาก กลุ่มหนึ่งที่ศึกษาได้แก่ เคอรี่ ,บาส์ชอว์ และ เรนท์ (Curry, Bashaw & Rentz 1978)ทินส์เลย์และเดวิส (Tinsley & Dawis 1973) ไวย์เลย์ และเดวิส (Whilely & Dawis 1974) และ ไรท์ (Wright 1968) ได้สรุปว่าพารามิเตอร์ความสามารถคงที่ (Invariance) ไม่แปรผันกับข้อกระทงชุดย่อยๆ แต่การศึกษาชิ้นอื่นๆ ได้แก่งานของ ลอยด์และฮูเวอร์ (Loyd & Hoover 1980) ไสลด์และลินน์ (Slinde & Linn 1978) และ สไลด์และลินน์ (Slinde & Linn 1979)โดยใช้ราสช์โมเดลกับแบบสอบที่มีความยากแตกต่างกัน พบว่าขึ้นกับความสามารถของกลุ่ม กัสตาฟสัน (Gustafsson 1979b) และสไลด์กับลินน์ (Slinde & Linn 1979b) ได้ตั้งสมมติฐานว่า ผลของการเดาอาจจะมีส่วนทำให้ผลการเทียบมาตราโดยใช้ราสช์โมเดลแตกต่างกัน กัสตาฟสัน (Gustafsson 1979a) ได้แสดงให้เห็นว่าการเดาของผู้สอบอาจจะเป็นผลให้ได้ความสัมพันธ์ทางลบระหว่างระดับความยากและอำนาจจำแนก เนื่องจากสไลด์และลินน์พบหลักฐานว่า เมื่อ

ความสัมพันธ์ทางลบเกิดขึ้น ผลของการเดาอาจจะมีองค์ประกอบในผลงานวิจัยของเขา (อ้างจากงานของโคเลน (Kolen 1981 : 2)

เบียนชินีและโลเรต (Bianchini & Loret 1972) ได้ศึกษาเกี่ยวกับแบบสอบร่วม โดยแยกเป็น 2 กรณีคือ ศึกษานอร์ม (norm) และการเทียบมาตรา ในส่วนของการเทียบมาตราได้ออกแบบที่จะหาคะแนนดิบที่เทียบเคียงกันได้ของแบบสอบการอ่านทั้งหมดซึ่งเป็นแบบสอบที่วัดความเข้าใจและคำศัพท์ 7 ชุด (batteries) การศึกษานี้ได้ทำการเทียบมาตราโดยเทียบแบบสอบชุดอื่นๆกับ MAT ซึ่งเป็นแบบสอบร่วม โดยใช้กลุ่มตัวอย่าง 13,500 คน และในปีต่อมาได้เทียบแบบสอบชุด GMT กับแบบสอบร่วมโดยใช้กลุ่มตัวอย่าง 14,400 คน วิธีการเทียบมาตราใช้การเทียบเชิงเส้นตรงและการเทียบโดยใช้เปอร์เซนไทล์ การประเมินผลใช้การประมาณค่าความคลาดเคลื่อนในการเทียบมาตรา (error of equating) ผลการศึกษาพบว่า การเทียบมาตราที่ใช้เพศเป็นตัวแปรอิสระได้ผลไม่แตกต่างกัน ส่วนการเทียบระหว่าง ไอคิว (I.Q.) 3 กลุ่ม ได้ผลแตกต่างกันบ้างเล็กน้อย ผลการเทียบที่แตกต่างกันเห็นได้ชัดเจนคือ การเทียบจากกลุ่มที่มีการแข่งขัน (race) กลุ่มที่มีช่วงคะแนนสูงจะมีความคงที่ของคะแนนดีกว่ากลุ่มอื่นๆส่วนด้านวิธีการพบว่า การเทียบโดยใช้เปอร์เซนไทล์ได้ผลเป็นที่น่าพอใจกว่าการเทียบเชิงเส้นตรง

<u>การเปรียบเทียบวิธีการเทียบมาตรา</u>

ลอร์ด (Lord 1977) มาร์โค (Marco 1977) วูดส์และวิลเลย์ (Wood & Wiley 1977,1978) ได้เปรียบเทียบวิธีการเทียบมาตราจากวิธีดั้งเดิม(Conventional) กับวิธีที่ใช้ทฤษฎีสนองตอบข้อสอบ (Item Response Theory) ได้ผลแตกต่างกัน ซึ่งผลการวิจัยนี้แตกต่างจากการวิเคราะห์ซ้ำของ เรนซ์และบาส์ชอร์ (Rentz & Bashaw 1977) โดยใช้รูปแบบที่ใช้แบบสอบร่วมซึ่งเขาสรุปว่าผลการเทียบโดยใช้ราสช์โมเดลวิเคราะห์กับการเทียบโดยใช้เปอร์เซนไทล์ให้ผลเหมือนกัน (อ้างจากโคเลน (Kolen 1981 : 2))

สไลด์และลินน์ (Slinde & Linn 1977: 23-31) ได้สรุปและวิจารณ์การเทียบเชิงเส้นตรงและการเทียบโดยใช้เปอร์เซนไทล์ ซึ่งเน้นด้านการเทียบแบบสอบผลสัมฤทธิ์ที่มีระดับความยากแตกต่างกันโดยใช้ข้อมูลจากการศึกษาแบบสอบร่วม ซึ่งเขา

อภิปรายว่า วิธีการเทียบโดยใช้เปอร์เซนไทล์ดีกว่าการเทียบเชิงเส้นตรง แต่ทั้งสองวิธีนี้ยังไม่มีความเหมาะสมกับการเทียบมาตราตามแนวตั้ง (Vertical equating)

ในปีต่อมา สไลด์และลินน์(Slinde & Linn 1978:23-34) ได้ศึกษาปัญหาเกี่ยวกับการใช้ราสช์โมเดลในการเทียบมาตราตามแนวตั้ง โดยใช้คะแนนจากการสอบเข้ามหาวิทยาลัยอิลินอย วิชาคณิตศาสตร์ 50 ข้อ กลุ่มตัวอย่างจำนวน 1,365 คน ผู้วิจัยได้คัดเลือกข้อสอบที่มีค่าความยากต่ำกว่า .20 และสูงกว่า .80 ออกเหลือจำนวน 36 ข้อ แล้วแบ่งเป็น 2 ชุด คือข้อสอบง่าย 18 ข้อ และข้อสอบยาก 18 ข้อ ส่วนกลุ่มตัวอย่าง ผู้วิจัยได้คัดเลือกออก 58 คน เนื่องจากคะแนนเป็น 0 หรือ 18 ของแบบสอบทั้งสองฉบับ แล้วแบ่งเป็น 3 กลุ่ม คือกลุ่มต่ำจำนวน 390 คน ซึ่งสอบได้ตั้งแต่ 1-10 คะแนน กลุ่มปานกลางจำนวน 411 คน สอบได้คะแนน 11-13 และกลุ่มสูงจำนวน 506 คน สอบได้คะแนนตั้งแต่ 14 ขึ้นไป ผู้วิจัยได้ประมาณค่าพารามิเตอร์ความสามารถโดยใช้ล็อก(Log ability) แล้วนำไปพล็อตกราฟ ผลการศึกษาพบว่าเมื่อทดสอบความเหมาะสมของข้อมูลกับราสช์โมเดลด้วยสถิติไคสแควร์ระหว่างกลุ่มพบว่า ไม่มีนัยสำคัญที่ .01 และเมื่อพล็อตกราฟแล้วพบว่า การเทียบมาตราด้วยราสช์โมเดลตามแนวตั้ง(Vertical equating) ได้ผลไม่เป็นที่น่าพอใจ

มาร์โค ปีเตอร์เซน และสจ๊วต (Marco , Petersen & Stewart 1979) ได้ประเมินและเปรียบเทียบวิธีการเทียบมาตราระหว่างวิธีการเทียบเชิงเส้นตรง , วิธีการเทียบโดยใช้เปอร์เซนไทล์ และวิธีการเทียบโดยใช้โค้งลักษณะข้อสอบ ใช้แบบสอบชุด Scholastic Aptitude Test แบ่งแบบสอบแล้วเทียบเข้าสู่กันด้วยแบบสอบร่วมที่มีความยากเทียบเคียงกัน ผลก็คือเมื่อแบบสอบที่มีความยากแตกต่างกัน วิธีการเทียบโดยใช้โค้งลักษณะข้อสอบให้ผลดีที่สุด และวิธีเทียบเชิงเส้นตรงได้ผลแย่ที่สุด(Kolen 1981 : 2)

คาร์แกนและสตอค(Kagan & Stock 1980 : 34-37) ได้ศึกษาคะแนนที่เท่าเทียมกันระหว่างแบบสอบชุด Miller Analogies Test และ The Graduate Record Examination โดยใช้การแปลงคะแนนแบบเชิงเส้นตรง (linear method) และวิธีรีเกรสชั่น (Regression method)โดยกลุ่มประชากรจากมหาวิทยาลัยอริโซนา ปี ค.ศ. 1970-1976 จำนวน 18,017 คน ใช้กลุ่มตัวอย่างที่สอบทั้งสองฉบับจำนวน

154 คน เป็นชาย 86 คน หญิง 68 คน ผลการศึกษาพบว่า สหสัมพันธ์ของคะแนน GREV ของกลุ่มผู้หญิงสูงกว่ากลุ่มผู้ชายอย่างมีนัยสำคัญ ค่าเฉลี่ยคะแนนปฏิบัติ GREQ กลุ่มผู้ชายสูงกว่ากลุ่มผู้หญิงอย่างมีนัยสำคัญ ค่าความเบี่ยงเบนมาตรฐานของคะแนนที่แปลงเชิงเส้นตรงมีค่าใกล้เคียงกับคะแนนชุดเดิม แต่จากวิธีรีเกรสชั่น ได้ค่าเบี่ยงเบนมาตรฐานของคะแนนที่แปลงน้อยกว่าคะแนนชุดเดิม

โคเลน (Kolen 1981 : 1-12) ได้เปรียบเทียบผลการเทียบมาตราระหว่าง 3 วิธี คือ การเทียบเชิงเส้นตรง การเทียบโดยใช้เปอร์เซนไทล์ และการเทียบโดยใช้โค้งลักษณะข้อสอบ ใช้ข้อมูลจากโครงการของ IOWA Test of Educational Development (ITED) ค.ศ. 1978 ศึกษาทั้ง 2 แบบ คือ เทียบจากฟอร์มที่มีความยากเทียบเคียงกันและเทียบจากฟอร์มที่มีความยากต่างกัน กลุ่มตัวอย่างเป็นนักเรียนเกรด 9 และเกรด 10 จำนวน 10,728 คน จากโรงเรียนใน IOWA 34 แห่ง การศึกษาครั้งนี้ใช้เกณฑ์ในการเปรียบเทียบผลคือ cross-validation criterion ซึ่งวัดความใกล้เคียงของการกระจายของคะแนนที่แปลงแล้วกับการกระจายของคะแนนปฏิบัติ ผลการศึกษาพบว่า การเทียบโดยใช้เปอร์เซนไทล์เหมาะสมที่สุดกับแบบสอบที่มีความยากแตกต่างกัน ส่วนการเทียบโดยใช้โค้งลักษณะข้อสอบพบว่ากรณีใช้พารามิเตอร์ตัวเดียวเป็นวิธีการที่ไม่ถูกต้องสำหรับแบบสอบที่มีความยากแตกต่างกัน ซึ่งเนื่องจากอาจจะมีการเดาเกิดขึ้นมาก ส่วนกรณีใช้พารามิเตอร์ 3 ตัว ก็มีปัญหาเกิดขึ้นอย่างหลีกเลี่ยงไม่ได้คือ การประมาณค่าที่ไม่ถูกต้องของ Lower asymtote parameter

โคเลนและไวท์เนย์ (Kolen & Whitney 1982 : 279-293) ได้เปรียบเทียบความถูกต้องของวิธีการเทียบมาตรา 4 วิธี คือ การเทียบเชิงเส้นตรง การเทียบโดยใช้เปอร์เซนไทล์ การเทียบโดยใช้โค้งลักษณะข้อสอบซึ่งใช้พารามิเตอร์ตัวเดียวและพารามิเตอร์สามตัว ใช้แบบสอบของ General Education Development ซึ่งมี 12 ฟอร์ม ให้ฟอร์มหนึ่งเป็นแบบสอบร่วม ที่เหลืออีก 11 ฟอร์มเทียบเข้าสู่แบบสอบร่วม ใช้กลุ่มตัวอย่างเพียง ฟอร์มละ 200 คน การเปรียบเทียบผลใช้การวิเคราะห์แบบ cross-validation analysis ผลการวิจัยพบว่า วิธีการเทียบโดยใช้เปอร์เซนไทล์และการใช้โค้งลักษณะข้อสอบพารามิเตอร์สามตัว ได้ผลยังไม่เป็นที่ยอมรับ ส่วนวิธีเทียบเชิงเส้นตรงและวิธีใช้พารามิเตอร์ตัวเดียว ได้ผลคงที่มากกว่า ในการประเมินผล ผลของการ

ศึกษานี้ร่วมกับการเปรียบเทียบอื่น ๆ สรุปได้ว่า ความถูกต้องของวิธีการเทียบมาตราขึ้นกับองค์ประกอบหลายอย่าง เช่น ลักษณะของแบบสอบ รูปแบบในการเทียบมาตรา และขนาดของกลุ่มตัวอย่าง จากผลการวิจัยครั้งนี้ได้มีข้อเสนอว่า ถ้าขนาดของกลุ่มตัวอย่างเล็กในการเทียบตามแนวระนาบ (horizontal equating) ควรจะใช้วิธีการเทียบเชิงเส้นตรง

โฮม (Holmes 1982: 139-147) ได้ศึกษาการเทียบมาตราตามแนวตั้งโดยใช้ราสช์โมเดล ซึ่งมีจุดประสงค์แยกเป็น 3 ประการคือ ประการแรกเป็นการเลือกข้อกระทงที่เหมาะสมกับข้อมูล ประการที่สองประเมินข้อกระทงที่ได้จากการเลือกครั้งแรกโดยทดสอบว่าอยู่ในมิติเดียวกัน (Unidimensionality) ประการที่สามเป็นการเทียบมาตราตามแนวตั้ง โฮมได้ใช้ข้อมูลจากกลุ่มตัวอย่างเกรด 3 จำนวน 5,310 คน เกรด 4 จำนวน 3,393 คน ตอบแบบสอบชุด Comprehensive Tests of Basic Skills (1973) ฟอร์ม S จำนวน 183 ข้อ โฮมได้ใช้นักเรียนเกรด 3 และเกรด 4 อย่างละ 1000 คน เป็นกลุ่มตรวจสอบ cross validation และที่เหลือเป็นกลุ่มที่ใช้เทียบมาตรา

จากจุดประสงค์แรกเลือกข้อกระทงที่เหมาะสมได้จำนวน 32 ข้อ และประการที่สอง ประเมินข้อกระทงได้ดัชนีค่าอำนาจจำแนกระหว่าง .80 และ 1.20 จุดประสงค์ที่สามแบ่ง 32 ข้อ เป็น 3 ฟอร์ม ฟอร์มง่าย 12 ข้อ มีค่าระดับความยาก -0.48 ถึง 0.48 ฟอร์มปานกลาง 8 ข้อ 0.25 ถึง 0.73 และฟอร์มยาก 12 ข้อ 0.53 ถึง 1.27 ผลการเทียบมาตราโดยใช้ราสช์โมเดล ปรากฏว่าได้ค่าเฉลี่ยของคะแนนที่แปลงแล้วยังไม่เป็นที่น่าพอใจ

ปีเตอร์เซน มาร์โค และสจ๊วต (Petersen, Marco & Steward 1982: 71-135) ได้ศึกษาความถูกต้องของรูปแบบการเทียบมาตราโดยใช้วิธีเทียบเชิงเส้นตรงและวิธีเทียบโดยใช้เปอร์เซนไทล์ เขาได้เทียบคะแนนจากแบบสอบ 2 ฉบับ คือ Scholatic Aptitude Test(SAT) และ the Test of Standard Written English (TSWE) กลุ่มตัวอย่างแบ่งเป็น 3 ประเภทคือ กลุ่มสุ่ม (Random sample) กลุ่มที่เหมือนกัน (Similar sample) เป็นกลุ่มที่มีค่าเฉลี่ยของความสามารถทางภาษาเท่ากัน กลุ่มที่สามเป็นกลุ่มที่ไม่เหมือนกัน (Dissimilar sample) เป็นกลุ่มที่มีค่าเฉลี่ยของความสามารถทางภาษาแตกต่างกัน แต่ละกลุ่มแบ่งเป็น 3 กลุ่มย่อย กลุ่มละ 1,577 คน ในการศึกษาครั้งนี้คณะผู้วิจัยได้ออกแบบการเทียบเป็น 11 แบบ ซึ่งแต่ละแบบจะแตกต่างกันโดยขึ้นกับลักษณะหลายประการ คือ

เทียบภายในแบบสอบเดียวกันหรือแบบสอบต่างฟอร์มกัน, ลักษณะแบบสอบร่วม, ลักษณะเนื้อหาของแบบสอบ, ความยากของแบบสอบ และความเหมือนกันหรือแตกต่างกันของกลุ่มตัวอย่าง เกณฑ์ที่ใช้ในการวัดความถูกต้องของการเทียบมาตราคือ "I.C.C. equipercentile criterion" ผลการศึกษาพบว่า เมื่อใช้กับกลุ่มตัวอย่างเล็กๆ การเทียบโดยใช้เปอร์เซนไทล์จะมีความคลาดเคลื่อนมากกว่าการเทียบเชิงเส้นตรง การเทียบเชิงเส้นตรงใช้ได้ผลไม่ดีเมื่อเทียบจากแบบสอบที่มีความยากแตกต่างกัน การเทียบโดยใช้เปอร์เซนไทล์ดีกว่าการเทียบเชิงเส้นตรงในกรณีที่ความสัมพันธ์ไม่เป็นเส้นตรง และจากกลุ่มตัวอย่างที่ไม่เหมือนกัน แบบสอบร่วมและแบบสอบที่ใช้เทียบที่มีระดับความยากแตกต่างกันมากจะมีความคลาดเคลื่อนมากกว่าที่มีระดับความยากแตกต่างกันน้อย แบบสอบสองฟอร์มที่แตกต่างกันมากจะให้ความคลาดเคลื่อนในการเทียบมาตรา (equating error)มากกว่าแบบสอบที่แตกต่างกันน้อย

ปีเตอร์เซนและคณะ(Petersen and others 1983 : 137-156) ได้วิเคราะห์ซ้ำกับเมื่อปี 1979 โดยศึกษาวิธีการเทียบมาตรา 3 วิธี คือการเทียบเชิงเส้นตรง การเทียบโดยใช้เปอร์เซนไทล์ และการเทียบโดยใช้ทฤษฎีสนองตอบข้อสอบ โดยใช้คะแนนจากแบบสอบวัดความถนัดทางการเรียน ผลปรากฏว่าสอดคล้องกับผลการศึกษาครั้งก่อน นั่นคือการใช้วิธีการเทียบตามทฤษฎีสนองตอบข้อสอบได้ผลมีความคงที่มากที่สุด

ลอร์ดและวิงเกอร์สกี้ (Lord & Wingersky 1984 : 153-161) ได้เปรียบเทียบการเทียบมาตราระหว่างวิธีการเทียบโดยใช้คะแนนจริงซึ่งวิเคราะห์ตามทฤษฎีสนองตอบข้อสอบ(Item Response Theory) และวิธีการเทียบคะแนนที่สอบได้โดยใช้เปอร์เซนไทล์ เครื่องมือที่ใช้ในการวิจัยคือแบบสอบชุด Scholastic Aptitude Test (SAT) มี 6 ฟอร์ม แต่ละฟอร์มมี 85 ข้อ การเทียบมาตราจะเป็นลักษณะลูกโซ่คือเทียบจากแบบสอบฟอร์มที่หนึ่งไปจนถึงฟอร์มที่หก โดยที่ฟอร์มหนึ่งและฟอร์มหกเป็นแบบสอบฉบับเดียวกัน ระหว่างฟอร์มมีแบบสอบร่วม จำนวน 40 ข้อ เป็นตัวปรับให้คะแนนอยู่ในสเกลเดียวกัน กลุ่มตัวอย่างได้จากการสุ่มจากประชากรเดียวกัน จำนวน 2,670 คน แยกสอบฟอร์มละสองกลุ่ม แต่ละกลุ่มสอบแบบสอบร่วมต่างกัน ผลการศึกษาได้ข้อสรุปว่า ยังไม่มีการตัดสินใจทางทฤษฎีที่ชัดเจนเกี่ยวกับการประยุกต์ใช้ทฤษฎีสนองตอบข้อสอบ วาวิธีการเทียบมาตราทั้งสองวิธี แบบใดดีกว่ากัน และได้มีการวิเคราะห์ซ้ำอีกโดยไม่ใช้แบบสอบร่วม ก็ได้ผลคล้ายคลึงกัน

การเทียบมาตราในประเทศไทย

ตามหลักสูตร พุทธศักราช 2503 การสอบระดับมัธยมศึกษาปีที่ 5 มี 2 ประเภท คือสอบปลายปีการศึกษา ประมาณเดือนมีนาคมของทุกปี และสอบเทียบ ม.ศ. 5 เดือนสิงหาคม สำหรับผู้ที่ไม่ได้เรียนตามระบบโรงเรียน การสอบทั้งสองประเภททางกระทรวงศึกษาธิการเป็นหน่วยงานที่รับผิดชอบในการสร้างแบบสอบให้ใช้เหมือนกันทั่วประเทศ ซึ่งทุกปีได้มีการสร้างแบบสอบฟอร์มใหม่ๆขึ้น โดยยึดเนื้อหาเดิมและกำหนดจุดตัดสำหรับได้ตกคือ 50 % แบบสอบแต่ละฟอร์มขึ้นอยู่กับกรรมการสร้างข้อสอบซึ่งการสร้างแบบสอบฟอร์มใหม่ๆมีการเปลี่ยนแปลงหลายอย่าง เช่น ความยาวของแบบสอบ ชนิดของแบบสอบ และเวลาที่ใช้ในการสอบ ดังนั้นจึงปรากฏว่าแบบสอบแต่ละปีนั้นไม่มีความเท่าเทียมกันในเชิงสถิติ และก็ไม่ปรากฏว่ามีการเทียบมาตราคะแนนของฟอร์มใหม่กับฟอร์มเก่าๆเลย คะแนนของนักเรียนที่สอบต่างฉบับกันจึงไม่สามารถเปรียบเทียบกันได้ และผลการสอบปรากฏว่าจำนวนนักเรียนที่สอบผ่านแต่ละปีแตกต่างกันมาก จึงเป็นที่สงสัยว่าความสามารถของนักเรียนแต่ละปีแตกต่างกันอย่างไร ซึ่งการเปรียบเทียบความสามารถของนักเรียนที่สอบคนละครั้งและจากแบบสอบต่างฟอร์มกันก็มีเทคนิคการเทียบมาตราเท่านั้นที่สามารถปรับให้คะแนนที่ได้จากการสอบคนละครั้งให้อยู่ในสเกลเดียวกัน

เยาวดี รางชัยกุล (2518) เป็นบุคคลแรกที่ได้เล็งเห็นถึงความสำคัญของประเด็นดังกล่าว จึงได้ทำการวิจัยเกี่ยวกับการเทียบมาตราโดยมีจุดมุ่งหมายคือสำรวจชอบเขตของแบบสอบฟอร์มต่างๆในโปรแกรมการสอบมัธยมศึกษาปีที่ 5ในประเทศไทยว่ามีความเท่าเทียมกันเพียงใดและเสนอแนะวิธีที่ดีที่สุดสำหรับการเทียบแบบสอบต่างฟอร์ม โดยข้อมูลที่ใช้ในการศึกษาได้มาจากกระทรวงศึกษาธิการซึ่งเป็นผลการสอบของนักเรียนชั้นมัธยมศึกษาปีที่ 5 สายวิทยาศาสตร์ จำนวน 500 คน จากการสอบ 4 ครั้งคือ เดือนมีนาคมและสิงหาคม 2516 เดือนมีนาคมและสิงหาคม 2517 คะแนนที่นำมาเทียบกันจากการสอบ 4 ครั้งคือวิชาภาษาไทย ภาษาอังกฤษ สังคมศึกษา คณิตศาสตร์ และวิทยาศาสตร์ เทคนิคการเทียบมาตราใช้วิธีการเทียบเชิงเส้นตรง รูปแบบที่ 1 ที่เสนอโดยแองกอฟฟ์ (Angoff 1971) ผลการศึกษาพบว่าเมื่อใช้สเกลร่วม (Common scale) ที่มีช่วงคะแนน 1-175 ค่าเฉลี่ย 100 ค่าเบี่ยงเบนมาตรฐาน 25 คะแนนที่เทียบกับสเกลร่วมทั้ง 4 ฟอร์ม มีคะแนนไม่เท่าเทียมกันทุกๆวิชา กล่าวคือ นักเรียนคนหนึ่งสอบได้คะแนนรวมของวิชาภาษาไทย ฟอร์มที่หนึ่ง (มีนาคม 2516) สูงกว่าคะแนนรวมของวิชาภาษาไทย ฟอร์มที่สอง (มีนาคม 2517) และยังพบว่าคะแนนจุดตัด

(cut-off score) ของแต่ละฟอร์มไม่สามารถเปรียบเทียบกันได้ เช่น ฟอร์มที่หนึ่งคะแนนจุดตัดเมื่อเทียบกับสเกลร่วมคือ 88 แต่ฟอร์มที่สองคะแนนจุดตัดคือ 78 ส่วนการศึกษาในรายวิชาทั้ง 5 วิชา พบว่าแต่ละวิชามีความแปรเปลี่ยนของคะแนนแตกต่างกัน วิชาภาษาไทยมีความแปรเปลี่ยนน้อยที่สุด วิชาคณิตศาสตร์มีความแปรเปลี่ยนของคะแนนมากที่สุด ผลการศึกษาได้เสนอแนะการเทียบมาตรากับกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 5 ว่าควรใช้กับรูปแบบที่ 4 ที่แองกอฟฟ์เสนอไว้เมื่อปี ค.ศ. 1971

จากงานวิจัยต่างๆที่กล่าวมาแล้วพอสรุปเป็นประเด็นสำคัญๆ ดังต่อไปนี้

1. การเทียบมาตราโดยใช้ทฤษฎีสนองตอบข้อสอบ (Item Response Theory) ซึ่งเป็นทฤษฎีวัดผลใหม่มีผลการเทียบดีกว่าใช้วิธีการเทียบตามทฤษฎีคลาสิคอล (Classical Test Theory) ซึ่งเป็นทฤษฎีวัดผลเก่าทั้งสองแบบคือ การเทียบตามแนวระนาบ (Horizontal Equating) และการเทียบตามแนวตั้ง (Vertical Equating)

2. ในด้านทฤษฎีวัดผลใหม่คือวิธีการเทียบโดยใช้โค้งลักษณะข้อสอบ พารามิเตอร์ 3 ตัว ให้ผลการเทียบดีกว่าชนิดพารามิเตอร์ตัวเดียวหรือราสชโมเดล (Rasch model) เพราะพารามิเตอร์ตัวเดียวมีข้อตกลงเบื้องต้นของโมเดลนี้เกี่ยวกับการเดาว่า ค่าการเดาของผู้สอบทุกคนเท่ากันหมด แต่ในทางปฏิบัติแล้ว การเดาของทุกคนไม่เท่ากัน ดังนั้นการเทียบมาตราด้วยการใช้โมเดลที่มีพารามิเตอร์ตัวเดียวจึงมีความคลาดเคลื่อนสูงกว่าโมเดลที่มีพารามิเตอร์สามตัว ซึ่งโมเดลนี้วิเคราะห์ค่าการเดาออกมาด้วย

3. การเทียบมาตราโดยใช้ทฤษฎีคลาสิคอล คือวิธีการเทียบชิงเส้นตรงและการเทียบโดยใช้เปอร์เซนไทล์นั้น ในกรณีของแบบสอบที่มีความยากแตกต่างกันและกลุ่มตัวอย่างมีความสามารถแตกต่างกันมากๆแล้วการเทียบโดยใช้เปอร์เซนไทล์ดีกว่าการเทียบเชิงเส้นตรง

4. ในส่วนของกลุ่มตัวอย่าง ถ้ากลุ่มตัวอย่างที่นำมาเทียบมาตรามีจำนวนน้อยแล้ว การเทียบโดยใช้โค้งลักษณะข้อสอบจะมีความคลาดเคลื่อนสูงกว่าการเทียบแบบดั้งเดิม เพราะว่าในการประมาณค่าพารามิเตอร์ของทฤษฎีวัดผลใหม่ต้องการผู้สอบจำนวนมาก ถ้าน้อยแล้วค่าพารามิเตอร์ที่ได้จะมีความคงที่น้อยลง