บทที่ 7
ความเชื่อมั่นและความเที่ยงตรง

ความเชื่อมั่นและความเที่ยงตรงเป็นคุณภาพที่สำคัญของแบบทดสอบ เนื่องจากผู้เขียนได้อธิบายความหมายของความเชื่อมั่นและความเที่ยงตรงในบทที่ 6 แล้ว จึงจะไม่อธิบายซ้ำอีกในบทนี้ เนื้อหาของบทนี้จะกล่าวถึงรายละเอียดของเรื่องต่อไปนี้

เนื้อหา

7.1 ทฤษฎีของความเชื่อมั่น

7.2 วิธีหาค่าความเชื่อมั่นของแบบทดสอบ

- 7.2.1 การวัดความมั่นคงในการตอบ
- 7.2.2 การวัดความเป็นคู่ขนาน
- 7.2.3 การวัดความคงที่ภายใน

7.3 ความคลาดเคลื่อนมาตรฐานของคะแนน

7.4 องค์ประกอบที่มีอิทธิพลต่อความเชื่อมั่นของแบบทดสอบ

7.5 ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์

7.6 ความเที่ยงตรง

- 7.6.1 ความเที่ยงตรงตามเนื้อหา
- 7.6.2 ความเที่ยงตรงตามโครงสร้าง
- 7.6.3 ความเที่ยงตามเกณฑ์

7.7 องค์ประกอบที่มีอิทธิพลต่อความเที่ยงตรงของแบบทดสอบ

7.8 ความสัมพันธ์ระหว่างความเชื่อมั่นกับความเที่ยงตรง

วัตถุประสงค์

เมื่อผู้อ่านได้ศึกษารายละเอียดของเนื้อหาบทที่ 7 แล้วควรจะมีความสามารถดังนี้

1. เข้าใจความคิดเชิงทฤษฎีของความเชื่อมั่นของแบบทดสอบว่า ความเชื่อมั่นของแบบทดสอบคือสัดส่วนระหว่างความแปรปรวนของคะแนนจริง กับความแปรปรวนของคะแนนดิบ
2. คำนวณค่าความเชื่อมั่นของแบบทดสอบได้
3. คำนวณค่าความคลาดเคลื่อนมาตรฐานของคะแนนได้
4. บอกองค์ประกอบที่มีอิทธิพลต่อความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบได้

5. วิเคราะห์ความสัมพันธ์ระหว่างประเภทของแบบทดสอบกับชนิดของความเที่ยงตรงได้

6. วิเคราะห์ความสัมพันธ์ระหว่างความเชื่อมั่นและความเที่ยงตรงได้

7.1 ทฤษฎีของความเชื่อมั่น (Theory of Reliability)

ทฤษฎีของความเชื่อมั่นสามารถอธิบายอย่างง่าย ๆ โดยเริ่มจากความรู้ที่ว่าคะแนนดิบ (observed score) ประกอบไปด้วยคะแนนจริง (true score) รวมกับความคลาดเคลื่อนของคะแนน (error score) ซึ่งอาจเขียนเป็นสูตรได้ดังนี้

$$X = T + E \quad \ldots\ldots(7-1)$$

เมื่อ X = คะแนนดิบ

T = คะแนนจริง

E = ความคลาดเคลื่อนของคะแนน

เนื่องจากความคลาดเคลื่อนของคะแนนเกิดขึ้นแบบสุ่ม ซึ่งโดยทฤษฎีแล้วจะมีทั้งความคลาดเคลื่อนในทางบวกและทางลบ และค่าเฉลี่ยของความคลาดเคลื่อนเหล่านี้จะมีค่าเป็นศูนย์ ดังนั้นถ้าความคลาดเคลื่อนของคะแนนเกิดขึ้นแบบสุ่ม มันจะไม่มีความสัมพันธ์กับคะแนนจริง และจะไม่มีความสัมพันธ์ซึ่งกันและกันด้วย จากข้อตกลงเบื้องต้นนี้ เราสามารถเขียนความแปรปรวนของแบบทดสอบ ให้อยู่ในรูปของสมการข้างล่างนี้

$$S_x^2 = S_t^2 + S_e^2 \quad \ldots\ldots(7-2)$$

เมื่อ S_x^2 = ความแปรปรวนของคะแนนดิบ

S_t^2 = ความแปรปรวนของคะแนนจริง

S_e^2 = ความแปรปรวนของความคลาดเคลื่อนของคะแนน

โดยทฤษฎีแล้ว ความเชื่อมั่น (r_{xx}) สามารถอธิบายได้ในรูปของสัดส่วนของความแปรปรวนของคะแนนจริง กับความแปรปรวนของคะแนนดิบ ดังนี้

$$r_{xx} = \frac{S_t^2}{S_x^2} \quad \ldots\ldots(7-3)$$

แต่จากสมการ (7–2) นั้น $S_t^2 = S_x^2 - S_e^2$

แทนค่า S_t^2 ในสมการ (7–3) จะได้

$$r_{xx} = \frac{S_x^2 - S_e^2}{S_x^2}$$ หรือ

$$r_{xx} = 1 - \frac{S_e^2}{S_x^2} \quad \ldots\ldots(7-4)$$

สูตร (7 – 4) คือ สูตรที่เราใช้อธิบายความหมายของค่าความเชื่อมั่นจากสูตรนี้อาจสรุปได้ ดังนี้

1. r_{xx} จะมีค่าเป็นศูนย์ ($r_{xx} = .00$) เมื่อ $S_e^2 = S_x^2$

2. r_{xx} จะมีค่าเป็น 1.00 เมื่อ $S_e^2 = 0$

3. ถ้าคะแนนของแบบทดสอบมีความคงเส้นคงวาเป็นที่เชื่อถือได้แล้ว S_e^2 จะมีค่าน้อยกว่า S_x^2 เสมอ

4. ความเชื่อมั่นของแบบทดสอบ (r_{xx}) ไม่เคยมีค่าติดลบ ทั้งนี้เพราะ S_e^2 ไม่เคยมีค่ามากกว่า S_x^2

7.2 วิธีหาค่าความเชื่อมั่นของแบบทดสอบ

มีวิธีการหลายอย่างที่ใช้ในการหาค่าความเชื่อมั่นของแบบทดสอบ แต่ที่นิยมใช้กันมากมีอยู่ 3 แบบ คือ

7.2.1 การวัดความมั่นคงของการสอบ (Measures of stability) ได้แก่วิธีสอบซ้ำ (test retest method)

7.2.2 การวัดความเป็นคู่ขนาน (Measures of equivalence)

7.2.3 การวัดความคงที่ภายใน (Measures of internal consistency)

1) วิธีหาแบบแบ่งเป็นข้อคู่–ข้อคี่ (Split-half method)

2) วิธีหาแบบ Kuder-Richardson

3) วิธีหาแบบ Coefficient alpha

7.2.1 การวัดความมั่นคงในการตอบ

การหาความเชื่อมั่นของข้อสอบด้วยวิธีนี้มีชื่อเรียกอีกอย่างหนึ่งว่าการสอบซ้ำ (test-retest method) ซึ่งการหาความเชื่อมั่นด้วยวิธีนี้เป็นการหาความเชื่อมั่นโดยการนำข้อสอบฉบับหนึ่งไปสอบกับนักเรียนกลุ่มหนึ่งซ้ำกัน 2 ครั้งในเวลาที่ใกล้เคียงกัน ดังนั้น สิ่งสำคัญในการหาค่าความเชื่อมั่นแบบนี้ก็คือ **"ช่วงเวลาของการสอบ 2 ครั้ง"** ถ้าระยะเวลาระหว่างการสอบครั้งแรกกับครั้งหลังกระชั้นกันเกินไป เช่น สอบ 2 ครั้ง ภายในวันเดียวกัน ผู้สอบก็อาจจะยังจำคำตอบที่สอบครั้งแรกได้ หรือถ้าปล่อยเวลาให้ล่วงเลยไปเป็นเดือนหรือเป็นปีแล้วจึงสอบครั้งที่ 2 ก็จะทำให้นักเรียนเกิดการเรียนรู้มากขึ้น ดังนั้นระยะเวลาระหว่างการทดสอบทั้ง 2 ครั้ง จึงมีผลต่อค่าความเชื่อมั่นมาก เวลาที่พอเหมาะ คือหลังจากการสอบครั้งแรกไปแล้วควรเว้นระยะเวลาประมาณ 3–7 วัน จึงค่อยสอบซ้ำ

สูตรสำหรับหาค่าความเชื่อมั่นแบบสอบซ้ำ คือ

$$r = \frac{n\Sigma xy - \Sigma x\,\Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2]\,[n\Sigma y^2 - (\Sigma y)^2]}} \quad \text{.....(7–5)}$$

เมื่อ r = ค่าความเชื่อมั่นของแบบทดสอบ

n = จำนวนผู้เข้าสอบ

Σx = ผลรวมของคะแนนที่ได้จากการสอบครั้งแรก

Σx^2 = คะแนนจากการสอบครั้งแรกแต่ละตัวยกกำลังสองแล้วนำมารวมกัน

$(\Sigma x)^2$ = คะแนนจากสอบครั้งแรกทั้งหมดรวมกัน แล้วยกกำลังสอง

Σy = ผลรวมของคะแนนจากการสอบครั้งหลัง

Σy^2 = คะแนนจากการสอบครั้งหลังแต่ละตัวยกกำลังสองแล้วนำมารวมกัน

$(\Sigma y)^2$ = คะแนนจากการสอบครั้งหลังทั้งหมดรวมกัน แล้วยกกำลังสอง

ตัวอย่าง : ข้อสอบฉบับหนึ่งมี 20 ข้อ ทำการทดสอบซ้ำ (retest) กับนักเรียน 10 คน ได้คะแนนดังรายการข้างล่าง ให้หาค่าความเชื่อมั่นของแบบทดสอบนี้

โดยกำหนดให้ X แทน คะแนนจากการสอบครั้งแรก

Y แทน คะแนนจากการสอบครั้งหลัง

คนที่	X	X^2	Y	Y^2	XY
1	18	324	18	324	324
2	9	81	12	144	108
3	10	100	10	100	100
4	12	144	10	100	120
5	7	49	8	64	56
6	5	25	6	36	30
7	15	225	16	256	240
8	10	100	14	196	140
9	8	64	9	81	72
10	3	9	6	36	18
รวม (Σ)	97	1121	109	1337	1208

$n = 10, \quad \Sigma x = 97, \quad \Sigma y = 109, \quad \Sigma xy = 1208, \quad \Sigma x^2 = 1121, \quad \Sigma y^2 = 1337$

$$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

$$= \frac{10(1208) - (97)(109)}{\sqrt{[(10)(1121) - (97)^2][(10)(1337) - (109)^2]}}$$

$$= \frac{12080 - 10573}{\sqrt{(11210 - 9409)(13370 - 11881)}}$$

$$= \frac{1507}{\sqrt{(1801)(1489)}}$$

$$= \frac{1507}{\sqrt{2681689}}$$

$$= 0.92$$

∴ แบบทดสอบฉบับนี้มีความเชื่อมั่นประมาณ .92 จัดได้ว่าเป็นความเชื่อมั่นที่สูงมาก แสดงว่าคะแนนที่ได้จากการทดสอบแบบทดสอบฉบับนี้เป็นที่เชื่อถือได้

อนึ่ง ในการทดสอบครั้งที่ 2 นั้น ถ้ามีนักเรียนขาดสอบไปบ้าง (ซึ่งเป็นสิ่งที่เกิดขึ้นได้) เวลาจะนำคะแนนที่ได้จากการทดสอบ 2 ครั้ง มาคำนวณหาค่าความเชื่อมั่น ก็ให้ดึงกระดาษคำตอบของคนที่ขาดสอบในการสอบครั้งที่สองออกจากกระดาษคำตอบของการสอบครั้งแรก เพื่อให้จำนวนนักเรียนในการสอบทั้ง 2 ครั้งเท่ากัน และเป็นกลุ่มเดียวกัน

7.2.2 การวัดความเป็นคู่ขนาน

การหาความเชื่อมั่นแบบนี้หาได้โดยการนำข้อสอบ 2 ฉบับ ซึ่งเป็นคู่ขนานกันไปสอบกับนักเรียนกลุ่มหนึ่งแล้วหาความสัมพันธ์ของคะแนน 2 ชุดที่ได้นั้น ค่าสหสัมพันธ์ที่ได้บ่งบอกถึงการวัดความเป็นคู่ขนานของข้อสอบ

จึงอาจกล่าวได้ว่า ค่าความเชื่อมั่นของแบบทดสอบที่หาด้วยวิธีนี้ไม่สามารถจะบ่งบอกถึงความคงเส้นคงวาในการตอบคำถามของนักเรียน แต่จะบ่งบอกว่าข้อสอบมีข้อคำถามที่เป็นตัวแทนของสิ่งที่จะวัดได้พอเพียงหรือไม่ ตัวอย่างเช่น ในการสอบข้อสอบผลสัมฤทธิ์ ในข้อสอบฉบับหนึ่งถ้าจะต้องออกข้อสอบวัดกันจริง ๆ แล้วจะต้องถามถึง 1,000 ข้อ แต่เราก็ทำไม่ได้จึงต้องสุ่มตัวอย่างมาเพียงพอให้เป็นตัวแทน วิธีการที่จะให้ข้อคำถามเป็นตัวแทนที่แท้จริงก็คือ สร้างข้อสอบในวิชาเดียวกันมา 2 ชุด โดยใช้เนื้อหาเดียวกัน มีระดับความยากง่ายเท่ากัน มีอำนาจจำแนกเหมือนกัน แล้วนำผลจากการสอบข้อสอบ 2 ชุด นั้นมาหาค่าสหพันธ์ ถ้าผลที่ได้มีความสัมพันธ์กันสูง ก็หมายความว่าข้อสอบ 2 ชุดนั้น วัดเนื้อหาเดียวกันจริง สำหรับสูตรสำหรับหาความเชื่อมั่นก็ใช้สูตรเดียวกันกับการหาแบบการสอบซ้ำ คือ

$$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

เมื่อ X = คะแนนที่ได้จากการสอบข้อสอบฉบับ ก

Y = คะแนนที่ได้จากการสอบข้อสอบฉบับ ข

วิธีการคาดคะเนความเชื่อมั่นโดยวิธีคู่ขนานนี้ มีปัญหาเกี่ยวกับการกำหนดเวลาที่ทำการสอบข้อสอบ 2 ชุด แบบเดียวกับการหาแบบการสอบซ้ำ อย่างไรก็ดีการหาความเชื่อมั่นโดยใช้ข้อสอบคู่ขนานก็มีข้อจำกัด เนื่องจากข้อสอบคู่ขนานสร้างยาก ส่วนใหญ่มักจะใช้ในข้อสอบมาตรฐาน ซึ่งตามปกติแล้วข้อสอบมาตรฐานวิชาหนึ่งจะต้องมีหลาย ๆ ชุด (form) การหาความเชื่อมั่นแบบใช้ข้อสอบคู่ขนานนี้ก็เช่นเดียวกับแบบสอบซ้ำ คือ ถ้าช่วงเวลาที่สอบ 2 form เว้นห่างกันมาก ก็จะทำให้ความเชื่อมั่นของข้อสอบต่ำกว่าที่ควรจะเป็น

7.2.3 การวัดความคงที่ภายใน

1) แบบแบ่งเป็นข้อคู่ — ข้อคี่

การหาความเชื่อมั่นของแบบทดสอบนี้ทำได้โดย การนำเอาข้อสอบฉบับที่ต้องการหาความเชื่อมั่น ไปทำการทดสอบกับนักเรียน 1 ครั้ง แล้วนำกระดาษคำตอบมาตรวจให้คะแนน ซึ่งในการตรวจนี้จะต้องใช้ key สำหรับตรวจ 2 ชุด คือ key ของข้อคู่ และ key ของข้อคี่ วิธีการตรวจแบบนี้จะทำให้สามารถแยกคะแนนของนักเรียนแต่ละคนออกเป็น 2 ชุด คือ คะแนนของข้อคู่ และคะแนนของข้อคี่ แล้วนำคะแนน 2 ชุดที่ได้มาหาค่าสหสัมพันธ์ โดยใช้สูตร (7—5) ดังนี้

$$r_{½½} = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

เมื่อ $r_{½½}$ = ค่าความเชื่อมั่นของข้อสอบครึ่งฉบับ

X = คะแนนของข้อคู่

Y = คะแนนของคี่

ค่าสหสัมพันธ์ที่ได้นี้จะบ่งบอกความเป็นคู่ขนานของข้อสอบกลุ่มข้อคู่ และกลุ่มข้อคี่ เนื่องจากค่าความเชื่อมั่นที่คำนวณได้เป็นค่าความเชื่อมั่นของข้อสอบเพียงครึ่งฉบับ จึงจำเป็นต้องนำค่าความเชื่อมั่นนั้นมาขยายให้เป็นค่าความเชื่อมั่นของข้อสอบทั้งฉบับ โดยใช้สูตรของ Spearman - Brown ดังนี้

$$\text{ความเชื่อมั่นของข้อสอบทั้งฉบับ } (r_{xx}) = \frac{2 \times \text{ความเชื่อมั่นของข้อสอบครึ่งฉบับ}}{1 + \text{ความเชื่อมั่นของข้อสอบครึ่งฉบับ}}$$

หรืออาจเขียนเป็นสูตรได้ดังนี้ $r_{xx} = \frac{2r_{½½}}{1+r_{½½}}$ .....(7–6)

ตัวอย่าง : ถ้าค่าความเชื่อมั่นของข้อสอบฉบับหนึ่ง ซึ่งหาโดยวิธีการแบ่งเป็นข้อคู่–ข้อคี่ เท่ากับ .60 อยากทราบว่าความเชื่อมั่นของข้อสอบทั้งฉบับมีค่าเท่าใด

$$r_{xx} = \frac{2 \times (.60)}{1 + (.60} = \frac{1.20}{1.60} = 0.75$$

ฉะนั้นข้อสอบฉบับนี้มีความเชื่อมั่น .75

สูตร (7–6) ใช้เมื่อต้องการหาความเชื่อมั่นของแบบทดสอบเมื่อเราขยายความยาวของข้อสอบให้มีความยาวเป็น 2 เท่าของข้อสอบฉบับเดิม ดังนั้นถ้าเราต้องการขยายความยาวของข้อสอบให้มีความยาวเป็นจำนวน K เท่าของข้อสอบฉบับเดิมแล้วเราสามารถหาค่าความเชื่อมั่นของข้อสอบฉบับใหม่ได้จากสูตร

$$r_{xx} = \frac{kr}{1 + (k-1)r} \quad .....(7–7)$$

เมื่อ r_{xx} = ความเชื่อมั่นของข้อสอบฉบับใหม่ที่มีความยาวเป็น k เท่าของฉบับเดิม

r = ความเชื่อมั่นของข้อสอบฉบับเดิม

K = จำนวนเท่าของข้อสอบที่ขยายให้ยาวขึ้น

ตัวอย่าง : ข้อสอบฉบับหนึ่งมี 50 ข้อ มีความเชื่อมั่น .60 ถ้าขยายความยาวของข้อสอบเป็น 150 ข้อ ข้อสอบฉบับใหม่จะมีความเชื่อมั่นเท่าใด

$$r_{xx} = \frac{3\,(.60)}{1 + (3-1)\,(.60)} = \frac{1.8}{2.2} = 0.818$$

หลักการหาค่าความเชื่อมั่นของแบบทดสอบ โดยวิธีแบ่งเป็นข้อคู่ กับข้อคี่ นี้ใช้หลักการเดียวกันกับแบบคู่ขนาน คือ ค่าความเชื่อมั่นจะเป็นตัวบ่งบอกว่า คำถามเป็นตัวแทนของสิ่งที่จะวัดหรือไม่ ถ้าข้อสอบข้อคู่ กับข้อคี่ แต่ละคู่วัดเนื้อหาเดียวกัน และมีความยากง่ายปาน ๆ กันแล้ว ค่า r_{xx} ที่หาได้จะมีค่าสูง

การหาค่าความเชื่อมั่น ของข้อสอบแบบแบ่งเป็นข้อคู่–ข้อคี่ นั้นดีกว่าแบบคู่ขนาน ตรงที่ไม่ต้องออกข้อสอบ 2 ฉบับ และเมื่อนำข้อสอบไปทดสอบกับนักเรียน 1 ครั้ง ก็สามารถนำมาหาความเชื่อมั่นได้ วิธีนี้จึงเป็นที่นิยมใช้มาก เพราะสะดวกกว่าแบบสอบซ้ำหรือแบบคู่ขนาน แต่การหาความเชื่อมั่นแบบนี้ ก็ยังไม่ค่อยสะดวกมากนัก เพราะต้องมาทำ key 2 ชุด ตรวจแยกคะแนนของนักเรียนแต่ละคน และ**วิธีการนี้ไม่เหมาะสำหรับนำมาใช้หาความเชื่อมั่นของข้อสอบแบบจำกัดเวลามาก ๆ** (speeded tests)

2) แบบคูเดอร์ — ริชาร์ดสัน (Kuder - Richardson)

การหาค่าความเชื่อมั่นของข้อสอบแบบนี้ เป็นวิธีการหาที่ง่ายและสะดวกที่สุด วิธีนี้ดีตรงที่สอบเพียงครั้งเดียว ก็สามารถนำคะแนนมาหาความเชื่อมั่นได้โดยไม่ต้องเอาข้อสอบมาแบ่งครึ่ง สูตรการคำนวณหาค่าความเชื่อมั่นแบบคูเดอร์ – ริชาร์ดสัน มี 2 สูตร คือ K–R20 และ K–R21 ดังนี้

$$\text{K–R20: } r_{xx} = \frac{K}{K-1}\left[1-\frac{\Sigma pq}{S_x^2}\right] \quad \text{.....(7–8)}$$

$$\text{K–R21: } r_{xx} = \frac{K}{K-1}\left[1-\frac{\bar{X}(K-\bar{X})}{KS_x^2}\right] \quad \text{.....(7–9)}$$

เมื่อ k = จำนวนข้อของแบบทดสอบ

p = สัดส่วนของนักเรียนที่ตอบข้อสอบแต่ละข้อถูก (เช่น ถ้าข้อสอบข้อ 1 มีคนตอบถูก 40 คน จากนักเรียนทั้งหมด 50 คน p จะมีค่าเท่ากับ 40/50 = .80)

q = สัดส่วนของนักเรียนที่ตอบข้อสอบแต่ละข้อผิด (q = 1-p)

Σpq = ผลรวมทั้งหมดของ pq

$\bar{X}$ = คะแนนเฉลี่ยของกลุ่ม

S_x^2 = ความแปรปรวนของคะแนน $= \frac{N\Sigma x^2-(\Sigma x)^2}{N(N-1)}$

ข้อแตกต่างระหว่าง K–R20 และ K–R21 คือ การหาความเชื่อมั่นแบบ K–R 21 ยึดหลักว่า ข้อสอบแต่ละข้อต้องมีความยาก–ง่ายเท่ากัน (มีค่า p เท่ากันทุกข้อ) **ถ้าคำถามแต่ละข้อมีความ**

ยาก—ง่ายต่างกันแล้ว ค่าความเชื่อมั่นที่คำนวณได้จากสูตรจะมีค่าต่ำกว่าความเป็นจริงเล็กน้อย

สูตร K—R 21 เป็นที่นิยมใช้ของครูทั่วไปมากกว่าสูตร K—R 20 ทั้งนี้เนื่องจากวิธีการคิดคำนวณง่ายกว่า เพราะเราเพียงแต่คำนวณหา ค่าเฉลี่ยของกลุ่ม $\bar{X}$ และค่าความแปรปรวนของคะแนน (ซึ่งเป็นสิ่งที่ครูต้องหาอยู่แล้ว หลังจากตรวจข้อสอบเสร็จ) ก็สามารถหาค่าความเชื่อมั่นของข้อสอบได้โดยการใช้สูตร K-R 20 นั้น ครูจะหาได้ก็ต่อเมื่อครูหาค่า p และ q ของแต่ละข้อมาคูณกัน แล้วยังต้องหาผลบวกของผลคูณเหล่านั้น ซึ่งเป็นการเสียเวลามากกว่า ครูจึงไม่ค่อยนิยมหาความเชื่อมั่นโดยใช้สูตร K-R 20

ตัวอย่าง การหาความเชื่อมั่นของแบบทดสอบโดยใช้สูตร K—R 20

ข้อสอบฉบับหนึ่งมี 20 ข้อ หาค่า p และ q ได้ดังตารางข้างล่าง จงหาความเชื่อมั่นของข้อสอบฉบับนี้ ถ้าข้อสอบฉบับนี้มีความแปรปรวนของคะแนนเท่ากับ 25

ข้อที่	p	q	pq	ข้อที่	p	q	pq
1	.68	.32	.22	11	.81	.19	.15
2	.37	.63	.24	12	.80	.20	.16
3	.72	.28	.20	13	.53	.47	.25
4	.52	.48	.25	14	.36	.64	.23
5	.19	.81	.15	15	.25	.75	.19
6	.51	.49	.25	16	.16	.84	.13
7	.74	.26	.19	17	.51	.49	.25
8	.21	.79	.17	18	.63	.37	.23
9	.23	.77	.18	19	.71	.29	.21
10	.50	.50	.25	20	.18	.82	.15

$$\Sigma pq = 4.05$$

$$S_x^2 = 25$$

$$K = 20$$

แทนค่าในสูตร

$$r_{xx} = \frac{20}{20-1}\left[1-\frac{4.05}{25}\right]$$

$$= \frac{20}{19}(1-.16)$$

$$= 0.88$$

ฉะนั้นแบบทดสอบฉบับนี้มีค่าความเชื่อมั่นเท่ากับ .88

ตัวอย่าง แบบทดสอบฉบับหนึ่งมี 70 ข้อ นำไปทดสอบกับนักเรียนกลุ่มหนึ่ง ได้ค่าคะแนนเฉลี่ย 47.91 ความแปรปรวนของคะแนนเท่ากับ 29.04 อยากทราบว่าข้อสอบฉบับนี้มีความเชื่อมั่นเท่าไร

$$r_{xx} = \frac{K}{K-1}\left[1-\frac{\bar{X}(K-\bar{X})}{KS_x^2}\right]$$

$$= \frac{70}{69}\left[1-\frac{47.91(70-47.91)}{70(29.04)}\right]$$

$$= \frac{70}{69}\left[1-\frac{1058.33}{2032.80}\right]$$

$$= .49$$

3) แบบโคเอฟฟิเชี่ยนแอลฟา (Coefficient alpha)

การหาความเชื่อมั่นของข้อสอบแบบนี้ เป็นวิธีการที่**เหมาะสำหรับหาความเชื่อมั่นของข้อสอบแบบอัตนัย** ซึ่งสามารถทำได้โดยการนำข้อสอบที่ต้องการหาค่าความเชื่อมั่นไปทำการทดสอบกับนักเรียนเพียงครั้งเดียว แล้วนำมาตรวจให้คะแนน แล้วจึงนำคะแนนที่นักเรียนแต่ละคนได้มาแทนค่าในสูตรหาความเชื่อมั่น เป็นดังนี้

$$\alpha = \frac{K}{K-1}\left[1-\frac{\Sigma S_i^2}{S_x^2}\right] \quad \text{.....(7-10)}$$

เมื่อ α = ค่าความเชื่อมั่นของแบบทดสอบ

K = จำนวนข้อของข้อสอบ

S_i^2 = ค่าความแปรปรวนของข้อสอบ 1 ข้อ ดังนั้น ถ้ามีข้อสอบ 5 ข้อ ก็จะต้องคำนวณหาค่า S_i^2 5 ครั้ง

S_x^2 = ค่าความแปรปรวนของข้อสอบทั้งฉบับ

ซึ่ง S_i^2 หามาได้จากสูตรข้างล่างนี้

$$S_i^2 = \frac{N\Sigma x_i^2 - \Sigma(x_i)^2}{N(N-1)} \text{ หรือ } = \frac{\Sigma fx^2 - \frac{(\Sigma fx)^2}{N}}{N-1}$$

เมื่อ N = จำนวนผู้เข้าสอบ

X_i = คะแนนที่นักเรียนแต่ละคนทำให้ได้จากการตอบข้อสอบที่ i

$\sum$ = จำนวนความถี่

ตัวอย่าง : ข้อสอบฉบับหนึ่งมี 5 ข้อ ค่าความแปรปรวนของข้อสอบทั้งฉบับเท่ากับ 64 ค่าความแปรปรวนของข้อสอบแต่ละข้อเท่ากับ 8, 10, 6, 12, 15 อยากทราบว่าข้อสอบฉบับนี้มีค่าความเชื่อมั่นเท่าใด

$$\alpha = \frac{5}{4}\left(1-\frac{(8+10+6+12+15)}{64}\right)$$

$$= \frac{5}{4}(1-.797)$$

$$= .25$$

7.3 ความคลาดเคลื่อนมาตรฐานของคะแนน

ถ้าเราสามารถให้นักเรียนคนหนึ่ง ทำข้อสอบฉบับหนึ่งซ้ำกันหลาย ๆ ครั้ง คะแนนที่เขาจะได้ในแต่ละครั้งนั้น จะแปรปรวนไปไม่เหมือนเดิม ความแปรปรวนของคะแนนนั้นมีความสัมพันธ์โดยตรงกับความเชื่อมั่นของการทดสอบ ถ้าข้อสอบมีความเชื่อมั่นต่ำ คะแนนของนักเรียนก็จะมีความแปรปรวนมาก ถ้าข้อสอบมีความเชื่อมั่นสูงคะแนนก็มีความแปรปรวนน้อย ในการสอบนักเรียนหลาย ๆ ครั้ง นี้ทำให้สามารถคาดคะเนความแปรปรวนของคะแนนได้ ตัวที่ใช้ในการคาดคะเนความแปรปรวนนี้เราเรียกว่า "ความคลาดเคลื่อนมาตรฐานของการวัด" ซึ่งเขียนย่อ ๆ ได้เป็น $SE_{(meas)}$ ซึ่ง $SE_{(meas)} = S\sqrt{1 - r_{xx}}$

ตัวอย่าง ข้อสอบฉบับหนึ่งมีความเบี่ยงเบนมาตรฐาน 8 มีความเชื่อมั่น .25 อยากทราบว่าข้อสอบฉบับนี้มีความคลาดเคลื่อนมาตรฐานของการวัดเท่าไร

$$SE_{(meas)} = 8\sqrt{1 - .25}$$

$$= 6.9$$

ตามปกติแล้วในคู่มือการทดสอบมักจะต้องบอกความคลาดเคลื่อนมาตรฐานในการวัดไว้ ซึ่งเราจำเป็นจะต้องใช้ค่านี้ประกอบการแปลความหมายคะแนนที่ได้จากการสอบ เช่น ด.ช.วินัยสอบภาษาไทยได้ 97 คะแนน และการสอบครั้งนี้ มีความคลาดเคลื่อนมาตรฐาน 5 หน่วย ความคลาดเคลื่อนมาตรฐานนี้ จะเป็นตัวกำหนดขอบข่ายของคะแนนจริงที่ ด.ช.วินัยควรจะสอบได้ ถ้าเรากำหนดขอบข่ายของคะแนนภายในช่วง 1 ความคลาดเคลื่อนมาตรฐานแล้ว คะแนนจริงที่

ได้นั้นจะเชื่อถือได้เพียง 68% แต่ถ้ากำหนดขอบข่ายของคะแนนจริงให้ตกอยู่ภายใต้ 2 ความคลาดเคลื่อนมาตรฐานแล้ว คะแนนที่ได้นั้นจะเชื่อถือได้ถึง 95% แต่ถ้าจะให้เชื่อถือได้ถึง 99% ก็ต้องให้คะแนนตกอยู่ในช่วงของ 3 ความคลาดเคลื่อนมาตรฐาน ดังนี้

จำนวนของความคลาดเคลื่อนมาตรฐาน	ระดับของความมั่นใจ	หน่วยของคะแนนที่คลาดเคลื่อนไปจากคะแนน 97	ขอบเขตของคะแนนจริง
1	68%	5	92-102
2	95%	10	87-107
3	99%	15	82-112

ช่วงของคะแนนจริงของการสอบครั้งหนึ่ง จะแคบหรือกว้างขึ้นอยู่กับ ค่าความคลาดเคลื่อนมาตรฐานของการวัด ถ้าความคลาดเคลื่อนมาตรฐานมีมาก ช่วงของคะแนนจริงก็ห่างมาก ถ้าความคลาดเคลื่อนมาตรฐานมีน้อย ช่วงของคะแนนจริงก็แคบตามลงมาด้วย หรือถ้าคะแนนไม่มีความคลาดเคลื่อนเลย คะแนนจริงก็คือ คะแนนที่สอบได้นั่นเอง จากตัวอย่างข้างบน ถ้าวินัยสอบได้คะแนน 97 คะแนน เรามีความมั่นใจถึง 99% ที่จะเชื่อว่าคะแนนจริงของเขาจะตกอยู่ระหว่างคะแนน 82-112

ความสัมพันธ์ระหว่างความเชื่อมั่นกับ ความคลาดเคลื่อนมาตรฐาน แสดงไว้ในตาราง 7.1 ตารางนี้แสดงค่าความคลาดเคลื่อนมาตรฐาน ซึ่งจะแปรเปลี่ยนไปตามค่าความเชื่อมั่นของข้อสอบและความเบี่ยงเบนมาตรฐาน เราอาจจะสรุปได้ว่าในขณะที่ค่าความเบี่ยงเบนมาตรฐานคงที่และถ้าค่าความเชื่อมั่นสูงแล้ว ความคลาดเคลื่อนมาตรฐานในการวัดจะลดลง ดังนั้น ค่าความเชื่อมั่นยิ่งสูงเพียงใด คะแนนที่ได้จากการทดสอบจะยิ่งมีความคลาดเคลื่อนลดน้อยลง และถ้าข้อสอบมีความเชื่อมั่นต่ำลง คะแนนที่ได้จากการทดสอบจะยิ่งมีความคลาดเคลื่อนมากขึ้น

ตาราง 7.1 ความสัมพันธ์ระหว่างความคลาดเคลื่อนของคะแนนกับ ความเชื่อมั่นของแบบทดสอบ

S	ค่าความเชื่อมั่น (r_{xx})					
	95	.90	.85	.80	.75	.70
30	6.7	9.5	11.6	13.4	15.0	16.4
28	6.3	8.9	10.8	12.5	14.0	15.3
26	5.8	8.2	10.1	11.6	13.0	14.2
24	5.4	7.6	9.3	10.7	12.0	13.1
22	4.9	7.0	8.5	9.8	11.0	12.0
20	4.5	6.3	7.7	8.9	10.0	11.0
18	4.0	5.7	7.0	8.0	9.0	9.9
16	3.6	5.1	6.2	7.2	8.0	8.8
14	3.1	4.4	5.4	6.3	7.0	7.7
12	2.7	3.8	4.6	5.4	6.0	6.6
10	2.2	3.2	3.9	4.5	5.0	5.5
8	1.8	2.5	3.1	3.6	4.0	4.4
6	1.3	1.9	2.3	2.7	3.0	3.3
4	.9	1.3	1.5	1.8	2.0	2.2
2	.4	.6	.8	.9	1.0	1.1

ตารางนี้หามาได้จากสูตร : $SE_{(meas)} = S\sqrt{1 - r_{xx}}$

S = ความเบี่ยงเบนมาตรฐาน

r_{xx} = ความเชื่อมั่นของข้อสอบ

ถ้าคู่มือการทดสอบฉบับหนึ่งไม่ได้รายงานค่าความคลาดเคลื่อนมาตรฐานในการวัดไว้ เราก็อาจจะประมาณค่า ความคลาดเคลื่อนมาตรฐานได้จากตารางข้างต้น

7.4 องค์ประกอบที่มีอิทธิพลต่อความเชื่อมั่นของแบบทดสอบ

มีองค์ประกอบจำนวนมากที่มีผลทำให้ความเชื่อมั่นของแบบทดสอบต่ำหรือสูงซึ่งองค์ประกอบเหล่านี้ จะต้องนำมาใช้พิจารณาประกอบกับการแปลผลต่อค่าความเชื่อมั่นหรือในการเปรียบเทียบค่าความเชื่อมั่นระหว่างแบบทดสอบ 2 หรือมากกว่า 2 ฉบับ เช่น ในการหาความเชื่อมั่นของข้อสอบแบบการสอบซ้ำนั้น ช่วงระยะเวลาที่ใช้ในการทดสอบมีอิทธิพลต่อค่าความเชื่อมั่นของแบบทดสอบเป็นอันมาก เพราะยิ่งสอบซ้ำในเวลาที่ใกล้เคียงกับการสอบครั้งแรกเท่าใด ค่าความเชื่อมั่นของข้อสอบก็จะยิ่งสูงมากขึ้นเท่านั้น ดังนี้เป็นต้น

การพิจารณาองค์ประกอบซึ่งมีอิทธิพลต่อความเชื่อมั่นของแบบทดสอบนี้ไม่เพียงแต่จะช่วยในการแปลความหมายของค่าความเชื่อมั่นของแบบทดสอบมาตรฐานให้ถูกต้องยิ่งขึ้น ยังจะช่วยให้ครูสามารถสร้างข้อสอบให้มีความเชื่อมั่นได้สูง ๆ อีกด้วย

สิ่งที่มีอิทธิพลต่อความเชื่อมั่นของแบบทดสอบที่สำคัญ มี 4 องค์ประกอบใหญ่ ๆ คือ

1. ความยาวของแบบทดสอบ

โดยปกติทั่ว ๆ ไปแล้ว ความยาวของแบบทดสอบมีความสำคัญมาก ข้อสอบยิ่งมีความยาวมากเท่าไร ความเชื่อมั่นของแบบทดสอบก็จะยิ่งสูงขึ้น ทั้งนี้เนื่องจากข้อสอบยิ่งมากเท่าไร ก็จะสามารถเป็นตัวแทนของสิ่งที่จะวัดได้มากเท่านั้น และโอกาสที่ผู้เข้าสอบจะได้คะแนนจากการเดาก็จะลดน้อยลงไป สมมติว่าการวัดความสามารถในการสะกดคำ แล้วเราให้นักเรียนสะกดคำเพียงคำเดียว ผลที่ได้ออกมาก็จะขาดความเชื่อมั่น เด็กอาจจะเป็นคนที่มีความสามารถในการสะกดคำสูงหรือต่ำก็ได้ ทั้งนี้ขึ้นอยู่กับความยากง่ายของคำที่นำมาให้สะกด เพราะถ้าเราเลือกคำยากนักเรียนส่วนใหญ่อาจจะสะกดผิดหมด แต่ถ้าเป็นคำง่าย ๆ นักเรียนส่วนใหญ่ก็จะสะกดถูกหมด ดังนั้น การใช้คำเพียงคำเดียวก็ไม่สามารถที่จะวัดความสามารถในการสะกดคำของนักเรียน จึงอาจกล่าวได้ว่าการเพิ่มจำนวนข้อคำถามที่ต้องการจะวัดให้มากขึ้น ก็จะเป็นการช่วยเพิ่มความคงเส้นคงวาของการวัดให้มากยิ่งขึ้นด้วย

ความยาวของแบบทดสอบเป็นส่วนหนึ่ง ซึ่งจะช่วยลดคะแนนที่ได้จากการเดา ตัวอย่างเช่น ในข้อสอบแบบถูกผิด 10 ข้อ นักเรียนคนหนึ่งอาจจะมีความรู้เพียง 7 ข้อ แต่อีก 3 ข้อนั้นต้องเดา เขาอาจจะเดาถูกทั้ง 3 ข้อ ซึ่งเขาก็จะได้คะแนนเต็ม หรือเขาอาจจะเดาผิดทั้ง 3 ข้อ ซึ่งเขาก็จะได้คะแนนเพียง 7 คะแนน อย่างไรก็ดี ถ้านักเรียนคนนี้ทำข้อสอบแบบถูก ผิด 100 ข้อ โอกาสที่เขาจะเดาถูก ก็จะลดน้อยลงไปเนื่องจากมีโอกาสที่จะเดาผิดมากยิ่งขึ้น และคะแนนที่ได้ก็จะสามารถบ่งถึงความรู้ที่แท้จริงของเขา

มีตัวอย่างที่แสดงให้เห็นว่า ความยาวของแบบทดสอบมีแนวโน้มที่จะเพิ่มค่าความเชื่อมั่น ของแบบทดสอบโดยดูจากการหาความเชื่อมั่นด้วยวิธี Split - half เช่น ข้อสอบฉบับหนึ่งถ้าหาความเชื่อมั่นด้วยวิธี Split - half จะได้ค่าความเชื่อมั่น = .60 แต่เมื่อมาใช้สูตร Spearman -

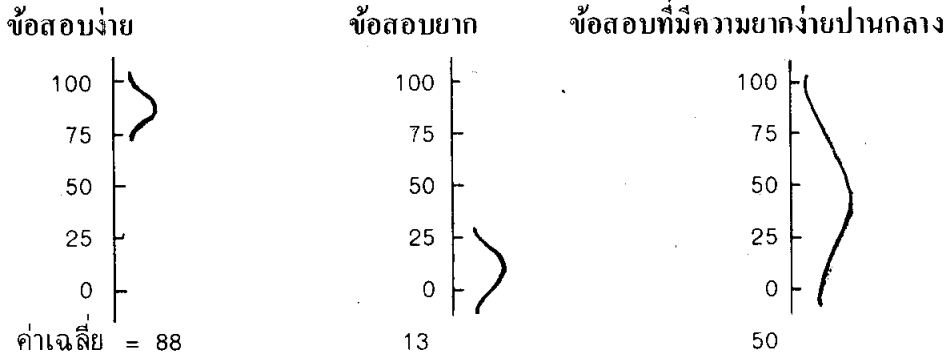
Brown ขยายเป็นความเชื่อมั่นของข้อสอบทั้งฉบับแล้วจะได้ = .75 จากตัวอย่างนี้ ก็แสดงให้เห็นอย่างชัดเจนว่า ถ้าจำนวนข้อสอบเพิ่มมากขึ้น ความเชื่อมั่นก็จะเพิ่มมากขึ้นตาม แต่มิได้หมายความว่าจะต้องเพิ่มเป็นสัดส่วนตามจำนวนข้อสอบที่เพิ่มขึ้น

สิ่งสำคัญที่จะต้องระวังอีกประการหนึ่งในการเพิ่มจำนวนข้อในแบบทดสอบก็คือ **ข้อสอบที่เพิ่มขึ้นนั้น จะต้องมีคุณภาพใกล้เคียงกับข้อสอบเดิม ทั้งในด้านเนื้อหา ความยากง่าย และอำนาจจำแนก** จึงจะมีผลต่อการเพิ่มค่าความเชื่อมั่นดังกล่าว

ในการสร้างข้อสอบเพื่อใช้ในชั้นเรียน สิ่งที่ครูจะต้องตระหนักถึงมากที่สุดก็คือ จะต้องสอบนักเรียนมาก ๆ ข้อ เพื่อช่วยให้ข้อสอบมีความเชื่อมั่นสูง ๆ ซึ่งจะต้องกินเวลาในการทดสอบมาก แต่ในขณะเดียวกัน เวลาที่กระทรวงกำหนดในการทดสอบและอายุของนักเรียน ก็เป็นอุปสรรคต่อการจะสอบข้อสอบคราวละมาก ๆ ข้อ ดังนั้นวิธีการแก้ไขอุปสรรคด้านนี้ที่ดีที่สุดก็คือ การสอบบ่อย ๆ สอบหลาย ๆ ครั้ง นั่นเอง

2. ความยากง่ายของแบบทดสอบ

ข้อสอบที่ง่ายเกินไปหรือยากเกินไป จะทำให้ข้อสอบมีความเชื่อมั่นต่ำ เพราะข้อสอบที่ง่ายเกินไป หรือยากเกินไปนั้น จะทำให้คะแนนที่ได้จากการทดสอบมีการกระจายแคบ เนื่องจากตัวข้อสอบง่ายเกินไป ผู้เข้าสอบส่วนใหญ่ก็จะทำข้อสอบได้มาก คะแนนก็จะมาออกันตรงคะแนนสูง แต่ถ้าข้อสอบยากเกินไปผู้สอบส่วนใหญ่ ก็จะทำข้อสอบไม่ค่อยได้ เหมือน ๆ กัน คะแนนก็จะมาออกันอยู่ตรงคะแนนต่ำ ๆ ความแตกต่างระหว่างกลุ่มจะน้อย (ดูรูป)



	ข้อสอบง่าย	ข้อสอบยาก	ข้อสอบที่มีความยากง่ายปานกลาง
ค่าเฉลี่ย	= 88	13	50
ช่วงห่างของคะแนน	= 25	25	100
ค่าความเชื่อมั่น	= .50	.50	.90

จากรูปเป็นการเปรียบเทียบการกระจายของคะแนนตามทฤษฎี จากการสอบข้อสอบ 100 ข้อ ความเชื่อมั่นหาได้จากการใช้สูตร Kuder - Richardson โดยการสมมติค่าความเบี่ยงเบนมาตรฐานขึ้นมาอย่างสมเหตุสมผล

จากความรู้เรื่องนี้ เราสามารถนำไปใช้ในการสร้างข้อสอบให้มีความเชื่อมั่นสูง ๆ ได้ โดยพยายาม สร้างข้อสอบให้มีคะแนนเฉลี่ยของผู้สอบเท่ากับ 50% ของจำนวนข้อ และให้

คะแนนกระจายอยู่ระหว่างคะแนนศูนย์ถึงคะแนนเต็ม

3. ความเป็นปรนัยของแบบทดสอบ

ความเป็นปรนัยของแบบทดสอบ หมายถึง คุณลักษณะสำคัญ 3 ประการของแบบทดสอบ คือ

1. ผู้สอบอ่านโจทก์แล้วต้องเข้าใจคำถามตรงกัน
2. ไม่ว่าจะให้ใครตรวจคำตอบ จะต้องได้คะแนนเท่ากัน
3. สามารถนำคะแนนมาแปลความหมายออกมาในรูปเดียวกัน

ข้อสอบมาตรฐานส่วนมาก ทั้งข้อสอบความถนัด และผลสัมฤทธิ์มักจะต้องมีความเป็นปรนัยสูง เพราะความเป็นปรนัยของข้อสอบจะเป็นตัวช่วยควบคุมความแปรผันของคะแนน ซึ่งจะมีผลทำให้ความเชื่อมั่นของข้อสอบสูงด้วย

4. ความเป็นเอกพันธ์ของกลุ่ม

ถ้ากลุ่มผู้เข้าสอบมีความเป็นเอกพันธ์ต่อกันมากเพียงใด ข้อสอบก็จะมีความเชื่อมั่นน้อยลงเท่านั้น ทั้งนี้เนื่องจากถ้านักเรียนมีความสามารถใกล้เคียงกันแล้ว คะแนนที่ได้จากการทดสอบก็จะเกาะกลุ่มใกล้เคียงกันมาก ความแปรปรวนของคะแนนก็จะน้อย ซึ่งจะส่งผลทำให้ความเชื่อมั่นของข้อสอบต่ำลงไปด้วย ถ้าผู้เข้าสอบมีความสามารถแตกต่างกันมาก ๆ จะส่งผลให้ค่าความเชื่อมั่นของข้อสอบสูงตามไปด้วย เรื่องนี้สามารถอธิบายให้เข้าใจได้อย่างชัดเจนโดยการศึกษาดูสูตร 7-4 ดังนี้

$$r_{xx} = 1-\frac{S_e^2}{S_x^2}$$

ถ้า S_x^2 (ความแปรปรวนของคะแนน) มีค่ามากก็จะทำให้ S_e^2/S_x^2 มีค่าน้อยลง ซึ่งจะมีผลทำให้ r_{xx} มีค่ามาก แต่ถ้ากลุ่มมีความเป็นเอกพันธ์กันมาก S_x^2 ก็จะมีค่าน้อย ซึ่งจะทำให้ S_e^2/S_x^2 มีค่ามาก จึงมีผลทำให้ r_{xx} มีค่าน้อยลง

7.5 ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์

เนื่องจากการวัดแบบอิงกลุ่ม มีจุดประสงค์ที่จะเปรียบเทียบความสามารถของผู้เรียนกับเพื่อนคนอื่น ๆ ส่วนการวัดแบบอิงเกณฑ์ต้องการเปรียบเทียบความสามารถของผู้เรียนกับเกณฑ์มาตรฐาน ดังนั้น วิธีการหาค่าความเชื่อมั่นของการวัดแบบอิงกลุ่มบางวิธีอาจจะไม่เหมาะสมที่จะนำมาใช้กับการวัดแบบอิงเกณฑ์ ได้มีนักวัดผลเสนอวิธีการหาค่าความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์ไว้หลายวิธี เนื่องจากยังเป็นที่อภิปรายโต้เถียงกันอยู่ ยังหาข้อสรุปไม่ได้ว่าวิธีใด

เหมาะสมที่สุด ดังนั้น ผู้เขียนจะยกตัวอย่างให้นักศึกษาดูพอเป็นสังเขป คือ

1. Livingston (1972) ได้เสนอวิธีการหาค่าความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์ โดยเขาได้ปรับปรุงมาจาก หลักการหาค่าความเชื่อมั่นแบบอิงกลุ่ม ดังนี้

$$r_{cc} = \frac{r_{xx}\delta_x^2 + (\bar{X}-C)^2}{\delta_x^2 + (\bar{X}-C)^2} \qquad \text{.....(7-11)}$$

$$= \frac{r_{xx}S_x^2 + (\bar{X}-C)^2}{S_x^2 + (\bar{X}-C)^2}$$

เมื่อ r_{cc} = ค่าความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์

r_{xx} = ค่าความเชื่อมั่นของข้อสอบที่คำนวณหาค่าแบบอิงกลุ่ม

$\bar{X}$ = ค่าเฉลี่ยของกลุ่ม

C = เกณฑ์มาตรฐานที่กำหนดไว้สำหรับตัดสินว่าใครมีความรู้มากพอที่จะให้ผ่านหรือให้สอบได้

S_x^2 = ความแปรปรวนของคะแนน

ตัวอย่าง แบบทดสอบฉบับหนึ่งมี 70 ข้อ นำไปทดสอบกับนักเรียนจำนวน 50 คน ได้ค่าสถิติดังนี้

คะแนนเฉลี่ย	= 48
ความแปรปรวนของคะแนน	= 29
ความเชื่อมั่น	= .49

ถ้ากำหนดว่าผู้ที่สอบได้คะแนนตั้งแต่ 50 คะแนนขึ้นไปเป็นผู้ที่มีความรอบรู้ในเนื้อหาวิชานั้นจริง อยากทราบว่าค่าความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์เป็นเท่าใด

จากสูตร

$$r_{cc} = \frac{r_{xx}S_x^2 + (\bar{X}-C)^2}{S_x^2 + (\bar{X}-C)^2}$$

$$= \frac{(.49)\ (29) + (48 - 50)^2}{29 + (48 - 50)^2}$$

$$= \frac{14.21 + 4}{29 + 4}$$

$$= .55$$

ฉนั้นความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์เท่ากับ .55

จากสูตร (7-11)

- ถ้า r_{xx} เพิ่มขึ้น r_{cc} จะเพิ่มขึ้น
- ถ้า r_{xx} = 1.00 r_{cc} จะ = 1.00 ด้วย
- ถ้า $\bar{X}$ = C แล้ว r_{cc} จะเท่ากับ r_{xx}

ดังนั้น ค่าความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์ที่คำนวณหาโดยสูตรของ Livingston จะมีค่ามากกว่า หรือเท่ากับค่าความเชื่อมั่นของข้อสอบฉบับเดียวกัน แต่คำนวณหา โดยใช้สูตรของการวัดแบบอิงกลุ่ม

2. Harris (1972) เสนอวิธีการหาค่าความเชื่อมั่นของแบบทดสอบ โดยการแบ่งกลุ่มผู้เข้าสอบออกเป็น 2 กลุ่ม เช่นเดียวกับการแบบของ Livingston คือ เป็นพวกผ่านเกณฑ์กับไม่ผ่านเกณฑ์ หรือพวกสัมฤทธิ์ผล - ไม่สัมฤทธิ์ผล ฯลฯ สูตรคำนวณหาค่าความเชื่อมั่นนั้น Harris เสนอให้ใช้การวิเคราะห์ความแปรปรวนแบบวันเวย์ (One way ANOVA)* ดังนี้

$$\mu_c^2 = \frac{SS_B}{SS_B + SS_W} \quad \text{.....(7-12)}$$

เมื่อ μ_c^2 = ค่าความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์.

SS_B = ผลบวกกำลังสองระหว่างกลุ่ม (ซึ่งมีเพียง 2 กลุ่ม คือ พวกผ่านเกณฑ์กับ ไม่ผ่านเกณฑ์) (Sum Square Between)

SS_W = ผลบวกกำลังสองภายในกลุ่ม (Sum Square Within)

ตัวอย่าง นำแบบทดสอบวิชาคณิตศาสตร์ ชั้น ม.1 จำนวน 20 ข้อ ไปทดสอบกับนักเรียน 15 คน ได้คะแนนดังข้างล่างนี้ ให้หาความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์ โดยที่กำหนดว่านักเรียนที่สอบผ่านจะต้องได้คะแนนตั้งแต่ 14 คะแนน จากคะแนนเต็ม 20 คะแนน (เกณฑ์ = ทำได้ 70 % ขึ้นไป)

*สำหรับผู้ที่ไม่มีความรู้หรือประสบการณ์เกี่ยวกับเรื่อง ANOVA ให้อ่านตำรา MR 311 : วิธีเชิงปริมาณทางการศึกษา ซึ่งเขียนโดยผู้เขียนคนเดียวกันเพิ่มเติม)

คะแนนของกลุ่มผ่านเกณฑ์			คะแนนของกลุ่มไม่ผ่านเกณฑ์		
นร.คนที่	X	X^2	นร.คนที่	X	X^2
1	15	225	1	5	25
2	18	324	2	8	64
3	16	256	3	10	100
4	17	289	4	12	144
5	20	400	5	6	36
6	14	196			
7	16	256			
8	16	256			
9	19	361			
10	17	289			
n_j 10	–	–	5	–	–
Σx_{ij}	168	–		41	
$(\Sigma x_{ij})^2$	28224	–		1681	–
$\frac{(\Sigma x_{ij})^2}{n_j}$	2822.4	–		336.2	–
Σx_{ij}^2	–	2852		–	369

วิธีการหาค่าสถิติต่าง ๆ

1. **กลุ่มที่สอบผ่านเกณฑ์** ($n_j = 10$)

$$\Sigma x_{ij} = 15 + 18 + 16 + 17 + 20 + 14 + 16 + 16 + 19 + 17 = 168$$

$$(\Sigma x_{ij})^2 = (168)^2 = 28224$$

$$\frac{(\Sigma x_{ij})^2}{n_j} = \frac{28224}{10} = 2822.4$$

$$\Sigma x_{ij}^2 = 225 + 324 + 256 + 289 + 400 + 196 + 256 + 256 + 361 + 289 = 2852$$

2. กลุ่มที่สอบไม่ผ่านเกณฑ์ $(n_j = 5)$

$$\Sigma x_{ij} = 5 + 8 + 10 + 12 + 6 = 41$$

$$(\Sigma x_{ij})^2 = (41)^2 = 1681$$

$$\frac{(\Sigma x_{ij})^2}{n_j} = \frac{1681}{5} = 336.2$$

$$\Sigma x_{ij}^2 = 25 + 64 + 100 + 144 + 36 = 369$$

การคำนวณหาค่าความเชื่อมั่นแบบอิงเกณฑ์ตามวิธีของ Harris

$$\mu_C^2 = \frac{SS_B}{SS_B + SS_w}$$

$$SS_B = \Sigma\frac{(\Sigma x_{ij})^2}{n_j} - \frac{(\Sigma\Sigma x_{ij})^2}{N}$$

$$= (2822.4 + 336.2) - \frac{(168 + 41)^2}{15}$$

$$= 3158.6 - \frac{43681}{15}$$

$$= 3158.6 - 2912.07$$

$$= 246.53$$

$$SS_w = \Sigma\Sigma x_{ij}^2 - \Sigma\frac{(\Sigma x_{ij})^2}{n_j}$$

$$= 3221 - 3158.6$$

$$= 62.4$$

แทนค่า SS_B กับ SS_W ในสูตร

$$\mu_C^2 = \frac{SS_B}{SS_B + SS_W}$$

$$= \frac{246.53}{246.53 + 62.40}$$

$$= \frac{246.53}{308.93}$$

$$= 0.798$$

ฉนั้นความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์เท่ากับ 0.798

7.6 ความเที่ยงตรง (Validity)

ความเที่ยงตรงเป็นคุณสมบัติที่สำคัญอีกประการหนึ่งของแบบทดสอบ แบบทดสอบที่มีความเที่ยงตรงสูง คือ แบบทดสอบที่มีสามารถวัดได้ตรงตามวัตถุประสงค์ที่ผู้ออกข้อสอบต้องการจะวัด

ความเที่ยงตรงของแบบทดสอบแบ่งออกเป็นชนิดใหญ่ ๆ ที่สำคัญ 3 ชนิด คือ

7.6.1 ความเที่ยงตรงตามเนื้อหา (Content Validity)

7.6.2 ความเที่ยงตรงตามโครงสร้าง (Construct Validity)

7.6.3 ความเที่ยงตรงตามเกณฑ์ (Criterion - Related Validity)

1) ความเที่ยงตรงตามสภาพ (Concurrent Validity)

2) ความเที่ยงตรงตามพยากรณ์ (Predictive Validity)

7.6.1 ความเที่ยงตรงตามเนื้อหา

การหาความเที่ยงตรงตามเนื้อหาของแบบทดสอบใด ๆ มีจุดประสงค์เพื่อจะศึกษาดูว่า แบบทดสอบนั้นสามารถวัดเนื้อหาวิชาตามที่กำหนดไว้ในหลักสูตรได้มากน้อยเพียงใด ดังนั้น โดยความหมายทั่ว ๆ ไปแล้ว ความเที่ยงตรงตามเนื้อหาจะหมายถึง **ความสามารถของแบบทดสอบฉบับหนึ่งที่จะมีข้อคำถามเป็นตัวแทนที่แท้จริงของเนื้อหาตามที่หลักสูตรกำหนดไว้**

ความเที่ยงตรงตามเนื้อหา เกี่ยวข้องเป็นอันดับแรกในการทดสอบผลสัมฤทธิ์ทางการเรียน ข้อสอบของครูจะมีความเที่ยงตรงตามเนื้อหาสูงมาก ถ้าครูออกข้อสอบสอดคล้องกับน้ำหนักที่กำหนดไว้ในตารางวิเคราะห์หลักสูตร ดังนั้น ข้อสอบที่มีความเที่ยงตรงตามเนื้อหา คือ ข้อสอบที่สร้างขึ้นโดยใช้ตารางวิเคราะห์หลักสูตรภาคเนื้อหาเป็นเกณฑ์ นั่นเอง

7.6.2 ความเที่ยงตรงตามโครงสร้าง

แบบทดสอบที่มีความเที่ยงตรงตามโครงสร้าง หมายถึง แบบทดสอบที่สามารถวัดคุณลักษณะบางประการทางจิตวิทยา ได้ครบถ้วนตามที่ต้องการ

ลักษณะทั่วไปทางจิตวิทยาที่เกี่ยวข้องกับโครงสร้างได้แก่ เชาว์ปัญญา ทัศนคติ การคิดอย่างวิพากษ์วิจารณ์ ความเข้าใจในการอ่าน ทักษะ ความถนัดทางคณิต ความสามารถในการวิเคราะห์ สังเคราะห์ หรือประเมินค่า แต่ละโครงสร้างจะต้องเกี่ยวกับทฤษฎี ซึ่งอธิบายและพยากรณ์พฤติกรรมของมนุษย์

ความเที่ยงตรงตามโครงสร้างจึงหมายถึงความสามารถของแบบทดสอบที่จะให้ผลที่สามารถนำมาแปลความหมายในลักษณะที่เกี่ยวข้องกับโครงสร้างทางจิตวิทยาได้

ตัวอย่างเช่น สมมติว่าจะตรวจสอบความเที่ยงตรงตามโครงสร้างของข้อสอบวัดเชาว์ปัญญาฉบับหนึ่ง นอกจากจะต้องรู้ว่า "เชาว์ปัญญา" คืออะไรแล้วยังต้องคาดคะเนว่า

1. คะแนนที่ได้จากแบบทดสอบนี้จะต้องเพิ่มขึ้นตามระดับอายุ (เราเชื่อว่าปัญญาของคนจะต้องเพิ่มขึ้นเมื่ออายุเพิ่มขึ้น จนถึง 16 ปี)

2. คะแนนจะต้องสามารถใช้ในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนที่เรียนอยู่ในโรงเรียน

3. คะแนนจะต้องมีความสัมพันธ์ในทางบวกกับอันดับของเชาว์ปัญญาซึ่งครูเป็นผู้จัดอันดับ

4. คะแนนจะต้องมีความสัมพันธ์ในทางบวกกับคะแนนของข้อสอบวัดเชาว์-ปัญญาฉบับอื่น ๆ

5. คะแนนจะต้องจำแนกความสามารถของผู้เข้าสอบได้อย่างเด่นชัด เช่น มีความถนัดทางด้านใด

6. คะแนนจะต้องไม่ได้มาจากการสอนของครูโดยตรง (เด็กจะทำข้อสอบได้เองจากประสบการณ์)

สำหรับแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนนั้น ความเที่ยงตรงตามโครงสร้างหมายถึง ความสามารถของข้อสอบที่จะถามวัดได้สอดคล้องกับพฤติกรรม ที่กำหนดไว้ในตารางวิเคราะห์หลักสูตร นั่นเอง

วิธีการคาดคะเนความเที่ยงตรงตามโครงสร้าง

จากหลักเกณฑ์ข้างบน จะเห็นว่าไม่มีวิธีใดวิธีหนึ่งโดยเฉพาะ ที่จะช่วยสร้างความเที่ยงตรงตามโครงสร้าง เพราะความเที่ยงตรงตามโครงสร้างจะขึ้นกับข้อมูลหลายอย่าง การหาความเที่ยงตรงตามโครงสร้างอาจทำได้ ดังนี้

1. วิเคราะห์ข้อสอบแต่ละข้อ เพื่อดูองค์ประกอบต่าง ๆ ของข้อสอบแต่ละข้อ ว่าวัดองค์ประกอบทางด้านใด ดังนั้น ข้อสอบวิทยาศาสตร์ ก็จะต้องสามารถวัดความสามารถด้านความรู้ ความเข้าใจ และความสามารถด้านอื่น ๆ หรือข้อสอบวัดเหตุผลทางเลขคณิต ก็ต้องมีข้อคำถามที่สามารถวัดความมีเหตุผลจริง ๆ

2. เปรียบเทียบคะแนนแต่ละกลุ่ม ในบางกรณีก็สามารถที่จะพยากรณ์ความแตกต่างของคะแนนแต่ละกลุ่มที่จะสอบได้ เช่น ความสามารถจะต้องเพิ่มขึ้นตามระดับอายุ (จากวัยเด็ก-วัยรุ่น) เด็กชายต้องได้คะแนนทางความเข้าใจในจักรกลสูงกว่าเด็กหญิง

3. เปรียบเทียบคะแนนก่อนและหลังของการกระทำบางอย่าง คะแนนจากการทดสอบบางอย่างอาจจะขึ้นลงเปลี่ยนแปลงได้ด้วยการฝึกฝน เช่น เรารู้ว่าความกระวนกระวายใจ มีผลต่อการทำข้อสอบ ดังนั้น เราก็คาดได้ว่าหลังจากก่อให้ผู้สอบเกิดความกระวนกระวายใจแล้ว คะแนนที่ได้จากการทดสอบต้องเปลี่ยนแปลงจากเดิม

4. หาค่าสหสัมพันธ์กับแบบทดสอบชนิดอื่น ซึ่งวัดองค์ประกอบเดียวกัน หรือวัดสิ่งเดียวกัน เช่น ข้อสอบความถนัดทางการเรียนฉบับหนึ่ง ก็ควรมีความสัมพันธ์สูงกับข้อสอบความถนัดทางการเรียนฉบับอื่น ๆ แต่มีความสัมพันธ์ต่ำกับข้อสอบความถนัดทางดนตรี

5. ตรวจสอบจากตารางวิเคราะห์หลักสูตรภาคพฤติกรรม

ความเที่ยงตรงทางโครงสร้างมีความสำคัญมากสำหรับข้อสอบทุกชนิดทั้ง achievement test, aptitude test และ personal - social development test

7.6.3 ความเที่ยงตรงตามเกณฑ์

การหาความเที่ยงตรงตามเกณฑ์ของแบบทดสอบ มีจุดประสงค์จะดูว่าคะแนนที่ได้จากการทดสอบแบบทดสอบฉบับนั้น สามารถนำมาใช้พยากรณ์ผลการเรียนภายภาคหน้า (Predictive Validity) หรือนำมาคาดคะเนผลการเรียนในปัจจุบันได้ดีเพียงใด (Concurrent Validity) ความเที่ยงตรงตามเกณฑ์มีอยู่ 2 ชนิด คือ **ความเที่ยงตรงตามพยากรณ์** (Predictive Validity) กับ**ความเที่ยงตรงตามสภาพ** (Concurrent Validity) การหาความเที่ยงตรงทั้ง 2 ชนิดนี้ เป็นการหาความสัมพันธ์ระหว่างคะแนนของแบบทดสอบที่จะหาความเที่ยงตรงกับตัวเกณฑ์ (Criteria) ในการหาความเที่ยงตรงตามพยากรณ์นั้น ตัวเกณฑ์ที่จะใช้เปรียบเทียบคือ คะแนนที่ได้จากผลการเรียน ซึ่งจะต้องติดตามไปในอนาคต ดังนั้น **ข้อสอบที่ใช้ในการสอบคัดเลือกจึงต้องมีความเที่ยงตรงตามพยากรณ์สูง** ส่วนการหาความเที่ยงตรงตามสภาพ ตัวเกณฑ์จะได้จากการที่ครูจัดอันดับความสามารถที่นักเรียนแสดงออกในห้องเรียนให้ครูสังเกตเห็นได้ ถ้าข้อสอบมีความเที่ยงตรงตามสภาพสูง ก็แสดงว่าคะแนนที่ได้จากแบบทดสอบฉบับนั้นสอดคล้องเป็นไปในทางเดียวกันกับอันดับที่ครูจัด นั่นคือ ถ้าเขาได้คะแนนจากแบบทดสอบสูงก็จะต้องได้อันดับดีด้วย ในปัจจุบันนี้ความเที่ยงตรงทั้ง 2 ชนิดนี้ มักเรียกรวมเป็นชื่อเดียวกันว่า Criterion - Related Validity ทั้งนี้เนื่องจากความเที่ยงตรงทั้ง 2 ชนิดนี้ มีหลักการและความหมายเป็นแบบเดียวกัน แต่จะต่างกันตรงระยะเวลาที่จะได้ตัวเกณฑ์มาเปรียบเทียบ

ในการศึกษาเกี่ยวกับความเที่ยงตรงตามเกณฑ์นั้นเราจะต้องพิจารณาหา**ตัวเกณฑ์**อย่างรอบคอบ สมมุติว่าเราต้องการจะทราบว่าคะแนนที่ได้จากแบบทดสอบความถนัดสามารถนำมาใช้ในการพยากรณ์ "ความสำเร็จในการเรียน" มากน้อยเพียงใด ในที่นี้ความสำเร็จในการเรียนเป็น**ตัวเกณฑ์** เราจะวัดความสำเร็จในการเรียนได้อย่างไร โดยปกติแล้ว นักการศึกษามักจะใช้เกรดเฉลี่ย (GPA) แทนความสำเร็จในการเรียน นั่นคือถ้านักเรียนได้เกรดเฉลี่ยสูง แสดงว่าเขาประสบความสำเร็จในการเรียนสูง แต่ถ้าเขาได้เกรดเฉลี่ยต่ำก็แสดงว่าเขาประสบความสำเร็จในการเรียนน้อย หรือไม่ประสบความสำเร็จในการเรียนเลย อย่างไรก็ดีในการใช้เกรดเฉลี่ยเพียงอย่างเดียวแทนความหมายของความสำเร็จในการเรียนนั้นนักการศึกษาส่วนมากตระหนักว่าไม่เป็นการถูกต้องมากนัก น่าที่จะใช้สัดส่วนระหว่างจำนวนผู้ที่เรียนจบกับผู้ที่เรียนไม่จบประกอบ

ด้วย ตัวอย่างที่คล้ายคลึงกับตัวอย่างที่กล่าวไปแล้วข้างต้นคือ ถ้าเราพยายามจะพยากรณ์ความสำเร็จในการทำงาน ในกรณีนี้ อันดับของสมรรถภาพในการทำงานของพนักงาน ซึ่งพิจารณาโดยหัวหน้าหน่วยงานมักจะถูกนำไปใช้เป็นตัวเกณฑ์ในการวัด ซึ่งอันดับที่กล่าวนี้อาจจะเป็นตัวเกณฑ์ที่ไม่เหมาะสมพอ ถ้าคะแนนจากแบบทดสอบไม่มีความสัมพันธ์กับอันดับที่ให้โดยหัวหน้าหน่วยงาน เราจะตอบไม่ได้โดยแน่ใจเลยว่าการที่ได้ค่าสัมประสิทธิ์สหสัมพันธ์ต่ำนั้นเกิดจากแบบทดสอบไม่ใช่ตัวพยากรณ์ที่ดีของความสำเร็จในการทำงาน หรืออันดับที่ให้โดยหัวหน้าหน่วยงานไม่ถูกต้อง หรือทั้งสองอย่าง ดังนั้นตัวเกณฑ์ที่ดีนั้นควรจะต้องเป็นตัวเกณฑ์ที่เชื่อถือได้ เชื่อมั่นได้และปราศจากอคติใด ๆ ทั้งสิ้น

วิธีหาค่าความเที่ยงตรงตามเกณฑ์

1) การหาค่าความเที่ยงตรงตามเกณฑ์นั้น หาได้โดยการคำนวณหาค่าสหสัมพันธ์ Correlation Coefficient) ของคะแนน 2 ชุด แบบเดียวกันกับการหาค่าความเชื่อมั่นของแบบทดสอบ แต่ต่างกันตรงที่การหาความเชื่อมั่นนั้น คะแนน 2 ชุด เกิดจากแบบทดสอบฉบับเดียวกัน ส่วนการหาค่าความเที่ยงตรงนั้น คะแนน 2 ชุด ได้มาจากแบบทดสอบต่างฉบับกัน คือแบบทดสอบที่ต้องการใช้พยากรณ์กับแบบทดสอบที่ใช้เป็นตัวเกณฑ์ในการเปรียบเทียบ ดังตัวอย่าง

สมมุติว่าสุพจน์ ซึ่งเป็นนักเรียนฝึกหัดครูปีที่ 3 ต้องการจะรู้ว่าคะแนนที่ได้จากการทดสอบความถนัดทางการเรียนจะสามารถพยากรณ์ผลสัมฤทธิ์ในการเรียนเลขคณิตชั้น ป.6 ของนักเรียนเขาได้ดีเพียงไร เนื่องจากนักเรียนทุกคนที่เข้ามาเรียนในชั้น ป.6 นี้ต้องผ่านการสอบความถนัดทางการเรียนก่อนเป็นอันดับแรก ปัญหาที่สำคัญที่สุดของสุพจน์ ก็คือขาดเกณฑ์ (Criterion) ที่ดีพอสำหรับความสำเร็จในการเรียนเลขคณิต เขาตัดสินใจใช้คะแนนจากการสอบไล่ซึ่งส่วนใหญ่ถามเกี่ยวกับวิชาเลขคณิตระดับ ป.6 เป็นเกณฑ์ ปัญหาอยู่ที่ว่านักเรียนที่ได้คะแนนความถนัดทางการเรียนสูงจะได้คะแนนจากการสอบไล่สูงด้วยหรือไม่ หรือถ้าเขาได้คะแนนจากการสอบถามความถนัดทางการเรียนต่ำ เขาจะได้คะแนนในการสอบไล่ต่ำไปด้วยหรือไม่

จากตัวอย่างข้างบน สุพจน์พบว่าส่วนใหญ่คะแนนที่ได้จากการทดสอบความถนัดกับคะแนนผลสัมฤทธิ์ขึ้นลงตามกัน ซึ่งก็เป็นแต่เพียงการพิจารณาขั้นต้นเท่านั้น ส่วนวิธีการจะหาค่าความเที่ยงตรงตามเกณฑ์ ที่ถูกต้องตามหลักแล้ว จะต้องนำคะแนนที่ได้จากการทดสอบ 2 ชุดนั้น มาหาค่าสหสัมพันธ์ (correlation coefficient) กัน ดังนี้

คะแนนที่ได้จากการสอบ 2 ครั้ง นำมาลงตารางได้ ดังนี้

ตาราง 7.2 แสดงคะแนนและอันดับที่ของนักเรียนที่ได้จากการสอบ

นักเรียนคนที่	คะแนนความถนัด	คะแนนเลขคณิต ป.6	อันดับของคะแนนความถนัด	อันดับของคะแนนเลขคณิต	ความแตกต่างของอันดับ(D)	D^2
1	119	77	1	3	−2	4
2	118	76	2	4	−2	4
3	116	72	3	6	−3	9
4	115	67	4	8	−4	16
5	112	82	5	1	4	16
6	109	63	6	10	−4	16
7	108	60	7	12	−5	25
8	106	78	8	2	6	36
9	105	69	9	7	2	4
10	104	49	10	18	−8	64
11	102	48	11	19	−8	64
12	100	58	12	14	−2	4
13	98	56	13	16	−3	9
14	97	57	14	15	−1	1
15	95	74	15	5	10	100
16	94	62	16	11	5	25
17	93	46	17	20	3	9
18	91	65	18	9	9	81
19	90	59	19	13	6	36
20	89	54	20	17	3	9
						$\Sigma D^2 = 532$

วิธีการคำนวณหา RANK - ORDER CORRELATION

1. เรียงคะแนนความถนัดจากมาก ⟶ น้อย แล้วกรอกลงคะแนนคณิตศาสตร์แต่ละคนทำได้ลงไปคู่กัน

2. จัดอันดับของคะแนนจาก 1 ⟶ N (จำนวนคนในกลุ่ม)

3. หาค่าความแตกต่าง (D) ของอันดับโดยเอาคะแนนความถนัดตั้ง ลบด้วยคะแนนคณิตศาสตร์

4. ยกกำลังสองของผลต่างของอันดับ (D^2)

5. หาผลบวกของกำลังสองของผลต่าง (ΣD^2)

6. คำนวณหาค่าความสัมพันธ์จากสูตร

$$\rho \text{ (rho)} = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)}$$

Σ = ผลบวกของ

D = ความแตกต่างของอันดับ

N = จำนวนคนในกลุ่ม

ρ = เป็นตัวอักษรกรีกใช้แทนค่า rank - order correlation coefficient ซึ่งมีความหมายเช่นเดียวกับค่า r

ฉะนั้น

$$\rho = 1 - \frac{6X532}{20(20^2 - 1)}$$

$$= 1 - \frac{3192}{7980}$$

$$= 1 - .40 = .60$$

วิธีการหาค่าสัมประสิทธิ์สหสัมพันธ์แบบนี้ เป็นวิธีการที่ง่ายเป็นการหาโดยจัดอันดับของคะแนน แต่ถ้าจำนวนกลุ่มตัวอย่างมาก ๆ ก็มีวิธีการหาแบบอื่นที่สะดวกกว่านี้ เช่น หาจากคะแนนดิบโดยตรงเลย โดยใช้สูตร (7-5) ดังนี้

$$r_{xy} = \frac{n\Sigma xy - \Sigma x \, \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2] \; [(n\Sigma y^2 - (\Sigma y)^2]}}$$

เมื่อ r_{xy} = ค่าความเที่ยงตรงของแบบทดสอบ

ความหมายของสัมประสิทธิ์สหสัมพันธ์

1.00 = มีความสัมพันธ์กันอย่างดีในทางบวก

.00 = ไม่มีความสัมพันธ์กันเลย

– 1.00 = มีความสัมพันธ์กันอย่างดียิ่งในทางลบ แต่ค่านี้ไม่นิยมใช้ในทางวัดผล เพราะขัดแย้งกับความรู้สึกของคนทั่ว ๆ ไป

ค่าความเที่ยงตรงของแบบทดสอบที่สุพจน์สร้างขึ้น มีค่า .60 ซึ่งนับว่าคะแนนระหว่างการสอบ 2 ฉบับ มีความสัมพันธ์กันค่อนข้างสูง ซึ่งถ้าพิจารณาดูค่าประสิทธิภาพในการพยากรณ์จากตาราง Expectancy Table จะเห็นว่า พวกที่ได้คะแนนสูงเป็นกลุ่มที่ 1 ในวิชาความถนัดจะมีโอกาสได้คะแนนสูงเป็นกลุ่มที่ 1 ในวิชาคณิตศาสตร์ ด้วย

2) สร้างตาราง 2 ทาง สำหรับพยากรณ์ ผลจากการสอบจะสามารถใช้ในการพยากรณ์ผลงานในอนาคต หรือคาดคะเนผลงานในปัจจุบันได้ดีเพียงไรนั้น เราอาจจะทำด้วยวิธีง่าย ๆ โดยการขีดรอยคะแนน (tally) ข้อมูลโดยตรงลงในตาราง 2 มิติได้ ซึ่งเราอาจเรียกตารางนั้นว่า Expectancy table ดังนี้

คะแนนความถนัด	คะแนนคณิตศาสตร์ ป.6				รวม
	45 - 54	55 - 64	65 - 74	75 - 84	
สูงกว่าเฉลี่ย (เกิน 100 คะแนน)			2 //	3 ///	5
เฉลี่ย	2 //	5 /////	2 //	1 /	10
ต่ำกว่าเฉลี่ย (ต่ำกว่า 95 คะแนน)	2 //	2 //	1 /		5
รวม	4	7	5	4	20

จากตารางจะเห็นว่าพวกที่สอบได้คะแนนสูงในวิชาความถนัด ไม่มีใครได้คะแนนต่ำในวิชาเลข และพวกที่ได้คะแนนต่ำในวิชาความถนัดก็ไม่มีใครได้คะแนนสูงในวิชาเลขเลย จึงสรุปได้ว่าข้อสอบความถนัดนี้มีความเที่ยงตรงตามเกณฑ์อยู่ในเกณฑ์ดี

7.7 องค์ประกอบที่มีอิทธิพลต่อความเที่ยงตรงของแบบทดสอบ

องค์ประกอบเป็นจำนวนมากที่มีผลทำให้ แบบทดสอบขาดความเที่ยงตรง แต่ก็มีองค์ประกอบหลายตัวที่เราสามารถหลีกเลี่ยงได้ เช่น ไม่วัดความรู้ทางสังคมวิทยาด้วยข้อสอบภาษาอังกฤษ หรือไม่วัดทักษะในการแก้ปัญหาโจทย์คณิตศาสตร์ชั้น ป.3 ด้วยเนื้อหาที่ใช้เรียนในชั้น ป.6 หรือเขียนคำสั่งชี้แจงสับสน ซึ่งสิ่งเหล่านี้ครูจะพยายามหาทางหลีกเลี่ยงไม่ว่าจะเป็นการสร้างข้อสอบที่ใช้ในห้องเรียน หรือข้อสอบมาตรฐานก็ตาม

1. องค์ประกอบที่เกิดขึ้นจากตัวแบบทดสอบเอง

องค์ประกอบที่มีผลทำให้ผลที่ได้จากการทดสอบขาดความเที่ยงตรง มีดังนี้

1) คำชี้แจงไม่ชัดเจน เช่นไม่บอกวิธีการตอบ หรือบอกไม่ชัดเจน เปิดโอกาสให้เดาได้หรือไม่ สิ่งเหล่านี้มีผลทำให้ข้อสอบขาดความเที่ยงตรง ทั้งสิ้น

2) ใช้ภาษาและศัพท์ที่สูงหรือยากเกินไป เพราะเราไม่ได้วัดความสามารถในการอ่าน

3) ความยากง่ายไม่พอเหมาะ ข้อสอบที่ยากเกินไปหรือง่ายเกินไปจะไม่มีอำนาจในการจำแนกเด็ก ซึ่งจะมีผลทำให้ความเที่ยงตรงของข้อสอบต่ำลงไปด้วย

4) ข้อคำถามไม่ดี คือ ข้อสอบที่สร้างขึ้นมาอย่างไม่มีหลักเกณฑ์ไม่สามารถวัดในสิ่งที่ต้องการจะวัดได้

5) คำถามคลุมเคลือ ซึ่งบางทีนั้นคำถามที่คลุมเคลือ หรือกำกวมนี้อาจมีผลทำให้เด็กเก่งสามารถเดาคำตอบได้ถูกมากกว่าเด็กอ่อน

6) ข้อคำถามไม่วัดพฤติกรรมที่ต้องการจะวัด เช่น ต้องการวัดความสามารถทางด้านความเข้าใจหรือทักษะ แต่กลับออกคำถามวัดแต่ความรู้ในเนื้อเรื่อง

7) ข้อสอบสั้นเกินไป แบบทดสอบแต่ละฉบับเป็นแต่เพียงกลุ่มตัวอย่างของคำถามทั้งหมด ถ้าเราออกข้อสอบน้อยเกินไป ข้อสอบก็ไม่สามารถเป็นตัวแทนของคำถามที่แท้จริงได้

8) การเรียงข้อสอบที่ไม่ถูกต้อง ตามปกติแล้วข้อสอบควรจะเรียงจากข้อง่ายไปหาข้อยาก ถ้าข้อสอบเรียงข้อยาก ๆ ไว้ต้น ๆ แล้วก็จะเป็นผลเสียทางด้านแรงจูงใจ

9) เรียงคำตอบถูกไว้อย่างมีระบบ เช่น กก ขข หรือ กขคงจ กขคงจ ซึ่งจะทำให้ผู้ตอบเดาได้

ถ้าครูพยายามขจัดสิ่งต่าง ๆ ที่กล่าวมาแล้ว ก็อาจจะช่วยให้ข้อสอบมีความเที่ยงตรงสูงขึ้น

2. องค์ประกอบที่เกี่ยวกับวิธีดำเนินการสอบและการให้คะแนน

การไม่ค่อยเคร่งครัดเวลาในการเก็บข้อสอบ การอธิบายเพิ่มเติมเมื่อมีข้อสงสัย การทำเสียงดังในระหว่างการสอบ รวมทั้งการให้คะแนนข้อสอบแบบอัตนัย การไม่ปฏิบัติตาม

คำสั่งชี้แจง ที่กำหนดไว้สิ่งเหล่านี้มีผลทำให้ข้อสอบมีความเที่ยงตรงต่ำทั้งสิ้น

3. องค์ประกอบที่เกี่ยวกับการตอบของนักเรียน

แรงจูงใจมีผลทำให้ความสามารถในการตอบของนักเรียนแตกต่างกัน เด็กบางคนอาจจะเกิดความกลัวสภาพการณ์ ในการสอบ จึงไม่สามารถจะทำข้อสอบได้อย่างปกติ หรือการเดาตอบ ก็มีผลทำให้ข้อสอบขาดความเที่ยงตรง เช่น ในข้อสอบแบบถูกผิด นักเรียนคนหนึ่งอาจจะเดาถูกตลอด หรืออีกคนหนึ่ง อาจจะเดาข้อสอบผิดพลาดตลอด

4. องค์ประกอบที่เกี่ยวกับความเชื่อมั่นของตัวพยากรณ์กับเกณฑ์

7.8 ความสัมพันธ์ระหว่างความเชื่อมั่นกับความเที่ยงตรง

ความเที่ยงตรงของแบบทดสอบ จะมีค่าไม่เกินรากที่สองของค่าความเชื่อมั่น นั่นคือ

$$r_{xy} = \sqrt{r_{xx}} \quad \text{.....(7–14)}$$

เมื่อ r_{xy} = ค่าความเที่ยงตรงสูงสุดของแบบทดสอบ

r_{xx} = ค่าความเชื่อมั่นของแบบทดสอบ

ตัวอย่าง : แบบทดสอบฉบับหนึ่งมีค่าความเชื่อมั่น .79 อยากทราบว่าแบบทดสอบฉบับนี้จะมีค่าความเที่ยงตรงสูงสุด ไม่เกินเท่าใด

$$r_{xy} = \sqrt{.79}$$
$$= 0.89$$

สรุปสาระสำคัญ

สาระสำคัญของเนื้อหาในบทนี้สรุปได้เป็นข้อ ๆ ดังนี้

1. ความเชื่อมั่นเป็นปริมาณของความคงเส้นคงวาระหว่างการวัดในสิ่งเดียวกัน 2 ครั้ง

2. วิธีการคาดคะเนความเชื่อมั่นของแบบทดสอบมีหลายวิธี เช่น การวัดความมั่นคงในการตอบ การวัดความเป็นคู่ขนาน และการวัดความคงที่ภายใน

3. การวัดความมั่นคงในการตอบ สามารถกำหนดได้โดย การนำเอาแบบทดสอบฉบับหนึ่งไปทดสอบกับคนกลุ่มเดียวกัน 2 ครั้ง แล้วคำนวณหาความสัมพันธ์ของคะแนน 2 ชุดนั้น

การเปลี่ยนแปลงใด ๆ ของคะแนนจากการสอบครั้งแรกกับการทดสอบครั้งหลัง ถือว่าเป็นความคลาดเคลื่อน

4. การวัดความเป็นคู่ขนาน สามารถกำหนดได้โดยการทดสอบคนกลุ่มหนึ่งด้วยแบบทดสอบ 2 ฉบับ ที่เป็นคู่ขนานกันในวันเดียวกัน แล้วหาความสัมพันธ์ของคะแนน 2 ชุดนั้น

5. วิธีการวัดความคงที่ภายในแบบทดสอบทุกวิธีทำได้โดยการดำเนินการสอบแบบทดสอบฉบับหนึ่งเพียงครั้งเดียว

6. โดยทั่ว ๆ ไปแล้ว แบบทดสอบที่ยาวกว่าจะมีความเชื่อมั่นมากกว่า

7. วิธีการวัดความคงที่ภายในไม่ควรนำมาใช้หาความเชื่อมั่นของแบบทดสอบที่กำหนดเวลาให้ทำในระยะเวลาที่จำกัดมาก ๆ (speeded tests)

8. ความเชื่อมั่นของข้อสอบจะสูงขึ้นถ้ากลุ่มที่เข้าสอบเป็นกลุ่มที่มีความแตกต่างกันอย่างมากในด้านสติปัญญา (heterogeneous group)

9. การคาดคะเนความเชื่อมั่นของแบบทดสอบโดยทั่ว ๆ ไปยึดหลักความแปรปรวนของคะแนนจริง แต่การวัดแบบอิงเกณฑ์ไม่สนใจในเรื่องความแปรปรวนของคะแนน ดังนั้นวิธีการหาความเชื่อมั่นของแบบทดสอบโดยทั่ว ๆ ไปอาจจะไม่เหมาะสมที่จะนำมาใช้กับการหาความเชื่อมั่นของการวัดแบบอิงเกณฑ์

10. แบบทดสอบที่มีความเชื่อมั่นสูงไม่จำเป็นต้องเป็นแบบทดสอบที่มีความเที่ยงตรงสูงตามไปด้วย

11. นักการศึกษามักจะจำแนกความเที่ยงตรงของแบบทดสอบออกเป็น 3 ประเภทใหญ่ ๆ คือ ความเที่ยงตรงตามเนื้อหา ความเที่ยงตรงตามโครงสร้าง และความเที่ยงตรงตามเกณฑ์

12. ความเที่ยงตรงตามเนื้อหาเกี่ยวข้องโดยตรงกับเรื่องที่เนื้อหาของแบบทดสอบนั้นเป็นตัวแทนที่แท้จริงของมวลประสบการณ์ที่ต้องการทดสอบทั้งหมดหรือไม่

13. ความเที่ยงตรงทางโครงสร้างเกี่ยวข้องโดยตรงกับการวัดพฤติกรรมทางจิตวิทยาที่เกี่ยวข้องกับการเรียนการสอน

14. ความเที่ยงตรงตามเกณฑ์เป็นวิธีการของการศึกษาความสัมพันธ์ระหว่างคะแนนที่ได้จากแบบทดสอบกับเกณฑ์ภายนอก

15. เกณฑ์ที่ดีจะต้องมีความเชื่อถือได้ มีความเชื่อมั่นและไม่มีอคติ

16. องค์ประกอบสำคัญที่มีผลต่อความเที่ยงตรงของแบบทดสอบคือ ความเชื่อมั่นของตัวพยากรณ์กับเกณฑ์ที่ใช้วัดความแตกต่างในด้านความสามารถทางสติปัญญาของผู้เข้าสอบ และคุณลักษณะภายในของแบบทดสอบ

คำถามท้ายบทที่ 7

1. ให้อธิบายความหมายและความสัมพันธ์ของความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบ
2. ท่านจะมีวิธีการปรับปรุงความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบได้โดยวิธีใดบ้าง จงอธิบาย
3. เรามีวิธีการตรวจสอบความเที่ยงตรงตามเนื้อหาของแบบทดสอบได้อย่างไร
4. ถ้ามีผู้กล่าวว่าเนื่องจากแบบทดสอบที่เขาสร้างขึ้นมามีความเชื่อมั่นประมาณ .95 ดังนั้น แบบทดสอบของเขาก็จะต้องมีความเที่ยงตรงสูงตามไปด้วย ท่านเห็นด้วยกับคำกล่าวนี้หรือไม่ เพราะเหตุใด
5. ให้เปรียบเทียบข้อดี-ข้อเสีย ระหว่างวิธีการหาความเชื่อมั่นโดยการวัดความมั่นคงในการตอบกับการวัดความคงที่ภายใน

บรรณานุกรม

Ivens, S.H., **An Investigation of Item** Analysis, Reliability and Validity in Relation to Criterion Referenced Tests, Dissertation, 1970.

Magnusson, D., **Test Theory,** Addison - Wesley, London, 1967.

Mehrens, W.A., and I.J. Lehmann, **Measurement and Evaluation in Education and Psychology,** Holt, Rinehart and Winston, 1978.

Popham, W.J., **Criterion - referenced measurement,** Englewood Cliffs, New Jersey : Prentice - Hall, 1978.