

บทที่ 7

การตรวจสอบคุณภาพของเครื่องมือทั้งฉบับ

เนื้อหา

เนื้อหาที่จะกล่าวถึงในบทที่ 7 มีรายละเอียด ดังนี้

- 7.1 วิธีหาความเชื่อมั่น (reliability) ของเครื่องมือ
- 7.2 องค์ประกอบที่มีอิทธิพลต่อความเชื่อมั่นของเครื่องมือ
- 7.3 วิธีหาความเที่ยงตรง (validity) ของเครื่องมือ
- 7.4 องค์ประกอบที่มีอิทธิพลต่อความเที่ยงตรงของเครื่องมือ
- 7.5 ความสัมพันธ์ระหว่างความเชื่อมั่น กับความเที่ยงตรง

7.1 วิธีหาความเชื่อมั่น (reliability) ของเครื่องมือ

มีวิธีการหลายอย่างที่ใช้ในการหาค่าความเชื่อมั่นของเครื่องมือที่ใช้ในการประเมินพัฒนาการระดับปฐมวัย แต่ที่นิยมใช้กันมากมีอยู่ 3 แบบ คือ

1. การวัดความมั่นคงในการตอบ (Measures of stability) ได้แก่วิธีการสอบซ้ำ (test retest method)
2. การวัดความเป็นคู่ขนาน (Measures of equivalence)
3. การวัดความคงที่ภายใน (Measures of internal consistency)
 - 1) วิธีหาแบบแบ่งเป็นข้อคู่ - ข้อคี่ (Split - half method)
 - 2) วิธีหาแบบ Kuder-Richardson
 - 3) วิธีหาแบบ Coefficient alpha

1. การวัดความมั่นคงในการตอบ การหาความเชื่อมั่นของข้อสอบด้วยวิธีนี้มีชื่อเรียกอีกอย่างหนึ่งว่า การสอบซ้ำ (test retest method) ซึ่งการหาความเชื่อมั่นด้วยวิธีนี้เป็นการหาความเชื่อมั่นโดยการนำแบบสอบถามหรือแบบทดสอบฉบับหนึ่งไปสอบกับนักเรียนกลุ่มหนึ่งซ้ำกัน 2 ครั้ง ในเวลาที่ใกล้เคียงกัน ดังนั้นสิ่งสำคัญในการหาค่าความเชื่อมั่นแบบนี้ก็คือ “ช่วงเวลาของการสอบ 2 ครั้ง” ถ้าระยะเวลาระหว่างการสอบครั้งแรกกับครั้งหลังกระชั้นชิดเกินไป เช่น สอบ 2 ครั้ง ภายในวันเดียวกัน ผู้สอบก็อาจจะยังจำคำตอบที่สอบครั้งแรกได้ หรือถ้าปล่อยเวลา

ให้ล่องเลยไปเป็นเดือนหรือเป็นปีแล้วจึงสอบครั้งที่ 2 ก็จะทำให้นักเรียนเกิดการเรียนรู้มากขึ้น
ตั้งนั้นระยะเวลาระหว่างการทดสอบทั้ง 2 ครั้ง จึงมีผลต่อค่าความเชื่อมั่นมาก เวลาที่เหมาะสมคือ
หลังจากการสอบครั้งแรกไปแล้วควรเว้นระยะเวลาประมาณ 3-7 วัน จึงค่อยสอบซ้ำ

สูตรสำหรับหาค่าความเชื่อมั่นแบบสอบซ้ำ คือ

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

เมื่อ r_{xx} = ค่าความเชื่อมั่นของแบบทดสอบ

n = จำนวนผู้เข้าสอบ

$\sum x$ = ผลรวมของคะแนนที่ได้จากการสอบครั้งแรก

$\sum x^2$ = คะแนนจากการสอบครั้งแรกแต่ละตัวยกกำลังสองแล้วนำมารวมกัน

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

$\sum y$ = ผลรวมของคะแนนที่ได้จากการสอบครั้งหลัง

$\sum y^2$ = คะแนนจากการสอบครั้งหลังแต่ละตัวยกกำลังสองแล้วนำมารวมกัน

$(\sum y)^2$ = คะแนนจากสอบครั้งหลังทั้งหมดรวมกัน แล้วยกกำลังสอง

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

คนที่	คะแนนจากการสอบ ครั้งแรก (x)	(x) ²	คะแนนจากการสอบ ครั้งหลัง (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
Σ(รวม)	97	1121	109	1337	1208

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

$$\begin{aligned}
 r_{xx} &= \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}} \\
 &= \frac{10(1208) - (97)(109)}{\sqrt{[(10)(1121) - (97)^2][(10)(1337) - (109)^2]}} \\
 &= \frac{1507}{\sqrt{2681689}} \\
 &= \frac{1507}{1637.586} \\
 &= 0.920
 \end{aligned}$$

∴ แบบทดสอบฉบับนี้มีความเชื่อมั่น 0.92 จัดได้ว่าเป็นความเชื่อมั่นที่สูงมาก แสดงว่าคะแนนที่ได้จากการทดสอบฉบับนี้เป็นที่เชื่อถือได้

อนึ่งในการทดสอบครั้งที่ 2 นั้น ถ้ามีนักเรียนขาดสอบไปบ้าง (ซึ่งเป็นสิ่งที่เกิดขึ้นได้) เวลาจะนำคะแนนที่ได้จากการทดสอบ 2 ครั้ง มาคำนวณหาค่าความเชื่อมั่น ก็ให้ดึงกระดาษคำตอบของคนที่ขาดสอบในการสอบครั้งที่สองออกจากกระดาษคำตอบของการสอบครั้งแรก เพื่อให้จำนวนคนจากการสอบ 2 ครั้งเท่ากัน และเป็นกลุ่มเดียวกัน

2. การวัดความเป็นคู่ขนาน การหาความเชื่อมั่นแบบนี้หาได้โดยการนำแบบสอบถามหรือแบบทดสอบ 2 ฉบับ ซึ่งเป็นคู่ขนานกันไปสอบกับนักเรียนกลุ่มหนึ่งแล้วหาความสัมพันธ์ของคะแนน 2 ชุดที่ได้นั้น ค่าสหพันธ์ที่ได้บ่งบอกถึงการวัดความเป็นคู่ขนานของข้อสอบ

จึงอาจกล่าวได้ว่า ค่าความเชื่อมั่นของเครื่องมือที่หาด้วยวิธีนี้ไม่สามารถจะบ่งบอกถึงความคงเส้นคงวาในการตอบคำถามของผู้ตอบ แต่จะบ่งบอกว่าข้อสอบมีข้อคำถามที่เป็นตัวแทนของสิ่งที่จะวัดได้พอเพียงหรือไม่ ตัวอย่างเช่น ในการสอบข้อสอบผลสัมฤทธิ์ในข้อสอบฉบับหนึ่ง ถ้าจะต้องออกข้อสอบวัดกันจริง ๆ แล้วจะต้องถามถึง 1,000 ข้อ แต่เราก็ทำไม่ได้จึงต้องสุ่มตัวอย่างมาเพียงพอให้เป็นตัวแทน วิธีการที่จะให้ข้อคำถามเป็นตัวแทนที่แท้จริงก็คือ สร้างข้อสอบในวิชาเดียวกันมา 2 ชุด โดยใช้เนื้อหาเดียวกัน มีระดับความยากง่ายเท่ากัน มีอำนาจจำแนกเหมือนกัน แล้วนำผลจากการสอบข้อสอบ 2 ชุดนั้นมาหาค่าสหพันธ์ ถ้าผลที่ได้มีความสัมพันธ์กันสูง ก็หมายความว่าข้อสอบ 2 ชุดนั้นวัดเนื้อหาเดียวกันจริง สำหรับสูตรการหาความเชื่อมั่นที่ใช้สูตรเดียวกันกับการหาแบบการสอบซ้ำ คือ

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

เมื่อ x = คะแนนที่ได้จากการสอบข้อสอบฉบับ ก

y = คะแนนที่ได้จากการสอบข้อสอบฉบับ ข

วิธีการคาดคะเนความเชื่อมั่นโดยวิธีคู่ขนานนี้ มีปัญหาเกี่ยวกับการกำหนดเวลาที่ทำ การสอบข้อสอบ 2 ชุด แบบเดียวกับการหาแบบการสอบซ้ำ อย่างไรก็ตามการหาความเชื่อมั่นโดยใช้ข้อสอบคู่ขนานก็มีข้อจำกัด เนื่องจากข้อสอบคู่ขนานสร้างยาก ส่วนใหญ่มักจะใช้ในข้อสอบมาตรฐาน ซึ่งตามปกติแล้วข้อสอบมาตรฐานวิชาหนึ่งจะต้องมีหลาย ๆ ชุด (form) การหาความเชื่อมั่นแบบใช้ข้อสอบคู่ขนานนี้ก็เช่นเดียวกับแบบสอบซ้ำ คือ ถ้าช่วงเวลาที่สอบ 2 forms เว้นห่างกันมากก็จะทำให้ความเชื่อมั่นของข้อสอบต่ำกว่าที่ควรจะเป็น

3. การวัดความคงที่ภายใน

1) แบบแบ่งเป็นข้อคู่ - ข้อคี่

การหาความเชื่อมั่นของแบบสอบถามหรือแบบทดสอบนี้ทำได้โดยการนำเอาข้อสอบฉบับที่ต้องการหาความเชื่อมั่นไปทำการทดสอบกับผู้ตอบ 1 ครั้ง แล้วนำกระดาษคำตอบมาตรวจให้คะแนนซึ่งในการตรวจนี้จะต้องใช้ key สำหรับตรวจ 2 ชุด คือ key ของข้อคู่ และ key ของข้อคี่ วิธีการตรวจแบบนี้จะทำให้สามารถแยกคะแนนของผู้ตอบแต่ละคนออกเป็น 2 ชุด คือคะแนนของข้อคู่ และคะแนนของข้อคี่ แล้วนำคะแนน 2 ชุดที่ได้มาหาค่าสหสัมพันธ์ โดยใช้สูตรดังนี้

$$r_{1/2} = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

เมื่อ $r_{1/2}$ = ค่าความเชื่อมั่นของข้อสอบครึ่งฉบับ

x = คะแนนของข้อคู่

y = คะแนนของข้อคี่

ค่าสหสัมพันธ์ที่ได้นี้จะบ่งบอกความเป็นคู่ขนานของข้อสอบกลุ่มข้อคู่ และกลุ่มข้อคี่ เนื่องจากค่าความเชื่อมั่นที่คำนวณได้เป็นค่าความเชื่อมั่นของข้อสอบเพียงครึ่งฉบับจึงจำเป็นต้องนำค่าความเชื่อมั่นนั้นมาขยายให้เป็นค่าความเชื่อมั่นของข้อสอบทั้งฉบับ โดยใช้สูตรของ Spearman - Brown ดังนี้

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

$$r_{xx} = \frac{2r_{1/2}}{1 + r_{1/2}}$$

ตัวอย่าง : ถ้าค่าความเชื่อมั่นของข้อสอบฉบับหนึ่ง ซึ่งหาโดยวิธีการแบ่งเป็นข้อคู่-ข้อคี่เท่ากับ .60 อยากทราบว่าความเชื่อมั่นของข้อสอบทั้งฉบับมีค่าเท่าใด

$$\begin{aligned} r_{xx} &= \frac{2 \times (.60)}{1 + (.60)} \\ &= \frac{1.20}{1.60} \\ &= .75 \end{aligned}$$

ฉะนั้นข้อสอบฉบับนี้มีความเชื่อมั่น .75

สูตร r_{xx} ที่กล่าวมาแล้วใช้เมื่อต้องการหาความเชื่อมั่นของแบบทดสอบเมื่อเราขยายความยาวของข้อสอบให้มีความยาวเป็น 2 เท่าของข้อสอบฉบับเดิม ดังนั้นถ้าเราต้องการขยายความยาวของข้อสอบให้มีความยาวเป็นจำนวน n เท่าของข้อสอบฉบับเดิมแล้วเราสามารถหาค่าความเชื่อมั่นของข้อสอบฉบับใหม่ได้จากสูตร

$$\begin{aligned} r_{xx} &= \frac{nr}{1 + (n - 1)r} \\ \text{เมื่อ } r_{xx} &= \text{ความเชื่อมั่นของข้อสอบฉบับใหม่ที่มีความยาวเป็น } n \text{ เท่า} \\ &\quad \text{ของฉบับเดิม} \\ r &= \text{ความเชื่อมั่นของข้อสอบฉบับเดิม} \\ n &= \text{จำนวนเท่าของข้อสอบที่ขยายให้ยาวขึ้น} \end{aligned}$$

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

$$\begin{aligned} r_{xx} &= \frac{3(.60)}{1 + (3 - 1)(.60)} \\ &= \frac{1.8}{2.2} = .818 \end{aligned}$$

หลักการหาค่าความเชื่อมั่นของแบบทดสอบโดยวิธีแบ่งเป็นข้อคู่กับข้อคี่นี้ใช้หลักการเดียวกันกับแบบคู่ขนานคือ ค่าความเชื่อมั่นจะเป็นตัวบ่งบอกว่าค่าถามเป็นตัวแทนของสิ่งที่วัดหรือไม่ ถ้าข้อสอบข้อคู่กับข้อคี่แต่ละคู่วัดเนื้อหาเดียวกัน และมีความยากง่ายปาน ๆ กันแล้วค่า r_{xx} ที่หาได้จะมีค่าสูง

การหาค่าความเชื่อมั่นของข้อสอบแบบแบ่งเป็นข้อคู่-ข้อคี่ นั้นดีกว่าแบบคู่ขนานตรงที่ไม่ต้องออกข้อสอบ 2 ฉบับ และเมื่อนำข้อสอบไปทดสอบกับนักเรียน 1 ครั้ง ก็สามารถนำมาหาความเชื่อมั่นได้วิธีนี้จึงเป็นที่นิยมใช้มาก เพราะสะดวกกว่าแบบสอบซ้ำหรือแบบคู่ขนาน แต่การหาความเชื่อมั่นแบบนี้ก็ยังไม่ค่อยสะดวกมากนัก เพราะต้องมาทำ key 2 ชุด ตรวจแยกคะแนนของนักเรียนแต่ละคน และวิธีการนี้ไม่เหมาะสำหรับนำมาใช้หาความเชื่อมั่นของข้อสอบแบบจำกัดเวลามาก ๆ (speed tests)

2) วิธีหาแบบคูเดอร์-ริชาร์ดสัน (Kuder-Richardson)

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

$$K-R20 : r_{xx} = \frac{n}{n-1} \left[1 - \frac{\sum pq}{\frac{s_x^2}{n}} \right]$$

$$K-R21 : r_{xx} = \frac{n}{n-1} \left[1 - \frac{\bar{x}(n-\bar{x})}{ns_x^2} \right]$$

da n = จำนวนข้อของแบบทดสอบ

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

cpq = ผลรวมทั้งหมดของ pq

$\bar{x}$ = คะแนนเฉลี่ยของกลุ่ม

s_x^2 = ความแปรปรวนของคะแนน

ข้อแตกต่างระหว่าง K-R20 และ K-R21 คือ การหาความเชื่อมั่นแบบ K-R21 ยึดหลักว่าข้อสอบแต่ละข้อต้องมีความยาก-ง่ายเท่ากัน (มีค่า p เท่ากันทุกข้อ) ถ้าคำถามแต่ละข้อมีความยาก-ง่ายต่างกันแล้ว ค่าความเชื่อมั่นที่คำนวณได้จากสูตรจะมีค่าต่ำกว่าความเป็นจริงเล็กน้อย

ตัวอย่าง การหาความเชื่อมั่นของแบบทดสอบโดยใช้สูตร K-R20 และ K-R21

ตัวอย่าง 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

$$cpq = 4.05$$

$$s_x^2 = 25$$

$$n = 20$$

แทนค่าในสูตร

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

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

$$= .88$$

ฉะนั้นแบบทดสอบฉบับนี้มีค่าความเชื่อมั่นเท่ากับ .88

ตัวอย่าง แบบสอบถามฉบับหนึ่งมี 70 ข้อ นำไปทดสอบกับนักเรียนกลุ่มหนึ่งได้ค่าคะแนนเฉลี่ย 47.91 ความแปรปรวนของคะแนนเท่ากับ 29.04 อยากทราบว่าแบบสอบถามฉบับนี้มีความเชื่อมั่นเท่าไร

$$\begin{aligned} r_{xx} &= \frac{n}{n-1} \left[1 - \frac{\bar{x}^2 (n-1)}{ns_{xx}^2} \right] \\ &= \frac{70}{69} \left[1 - \frac{47.91^2 (70-1)}{70(29.04)} \right] \\ &= .49 \end{aligned}$$

3) วิธีหาแบบโคเอฟฟิเชียนแอลฟา (Coefficient alpha)

การหาค่าความเชื่อมั่นของเครื่องมือวัดแบบนี้ เป็นวิธีการหาที่เหมาะสมสำหรับหาความเชื่อมั่นของข้อสอบแบบอัตนัยและแบบสอบถาม ซึ่งสามารถทำได้โดยการนำเครื่องมือที่ต้องการหาค่าความเชื่อมั่นไปทำการทดสอบกับกลุ่มตัวอย่างเพียงครั้งเดียวแล้วนำมาตรวจให้คะแนนแล้วจึงนำคะแนนที่แต่ละคนได้มาแทนค่าในสูตรหาความเชื่อมั่น ดังนี้

$$\alpha = \frac{k}{k-1} \left[1 - \frac{\sum s_i^2}{S_x^2} \right]$$

เมื่อ α = ค่าความเชื่อมั่นของแบบทดสอบ หรือแบบสอบถาม

k = จำนวนของข้อสอบ หรือข้อคำถาม

s_i^2 = ค่าความแปรปรวนของข้อสอบ 1 ข้อ ดังนั้นถ้ามีข้อสอบ 5 ข้อ ก็จะต้องคำนวณหาค่า 5 ครั้ง

S_x^2 = ค่าความแปรปรวนของข้อสอบทั้งฉบับ

ซึ่ง s_i^2 หาได้จากสูตรข้างล่างนี้

$$s_i^2 = \frac{n \sum x_i^2 - (\sum x_i)^2}{n(n-1)}$$

เมื่อ n = จำนวนผู้เข้าสอบ

x_i = คะแนนที่นักเรียนแต่ละคนทำได้จากการตอบข้อสอบที่ i

ตัวอย่าง : แบบสอบถามฉบับหนึ่งมี 5 ข้อ ค่าความแปรปรวนของแบบสอบถามทั้งฉบับเท่ากับ 64 ค่าความแปรปรวนของแบบสอบถามแต่ละข้อเท่ากับ 8, 10, 6, 12, 15 อยากทราบว่าแบบสอบถามฉบับนี้มีค่าความเชื่อมั่นเท่าใด

$$\begin{aligned}\alpha &= \frac{5}{4} \left(1 - \frac{(8+10+6+12+15)}{64} \right) \\ &= \frac{5}{4} (1 - .797) \\ &= .25\end{aligned}$$

ตัวอย่าง : ในการให้คะแนนพัฒนาการทางสติปัญญา ทางอารมณ์ และทางสังคมของนักเรียนชั้นอนุบาล 1 จำนวน 10 คน ได้คะแนนดังข้างล่าง จงหาความเชื่อมั่นของแบบประเมินฉบับนี้โดยวิธี Coefficient alpha

กำหนดให้ x_1 = คะแนนพัฒนาการทางสติปัญญา
 x_2 = คะแนนพัฒนาการทางอารมณ์
 x_3 = คะแนนพัฒนาการทางสังคม
 x_T = คะแนนรวมของนักเรียน

คนที่ \ ด้านที่	x_1	x_2	x_3	x_T	x_1^2	x_2^2	x_3^2	x_T^2
1	10	8	10	28	100	64	100	784
2	5	9	9	23	25	81	81	529
3	6	3	—	9	36	9	—	81
4	8	8	7	23	64	64	49	529
5	1	5	6	12	1	25	36	144
6	10	10	10	30	100	100	100	900
7	2	4	3	9	4	16	9	81
8	4	5	5	14	16	25	25	196
9	7	7	6	20	49	49	36	400
10	10	10	8	28	100	100	64	784
$\Sigma(\text{รวม})$	63	69	64	196	495	533	500	4428

จากสูตร

$$r_{\alpha} = \frac{k}{k-1} \left[1 - \frac{\sum S_i^2}{S_x^2} \right], \text{ เมื่อ } k = \text{จำนวนข้อคำถาม}$$

$$S^2 = \frac{N \sum x^2 - (\sum x)^2}{N(N-1)}, \text{ เมื่อ } N = \text{จำนวนนักเรียน}$$

$$S_1^2 = \frac{N \sum x_1^2 - (\sum x_1)^2}{N(N-1)}$$

$$= \frac{10(495) - (63)^2}{(10)(9)}$$

$$= \frac{4950 - 3969}{90}$$

$$= \frac{981}{90} = 10.9$$

$$S_2^2 = \frac{N \sum x_2^2 - (\sum x_2)^2}{N(N-1)}$$

$$= \frac{10(533) - (69)^2}{(10)(9)}$$

$$= \frac{5330 - 4761}{90}$$

$$= \frac{569}{90} = 6.32$$

$$S_3^2 = \frac{N \sum x_3^2 - (\sum x_3)^2}{N(N-1)}$$

$$= \frac{10(500) - (64)^2}{(10)(9)}$$

$$\begin{aligned}
&= \frac{5000 - 4096}{90} \\
&= \frac{904}{90} = 10.04 \\
S_x^2 &= \frac{N \sum x_T^2 - (\sum x_T)^2}{N(N-1)} \\
&= \frac{10(4428) - (196)^2}{(10)(9)} \\
&= \frac{44280 - 38416}{90} \\
&= \frac{5864}{90} = 65.16 \\
r_\alpha &= \frac{3}{3-1} \left[1 - \frac{(10.9 + 6.32 + 10.04)}{65.16} \right] \\
&= \frac{3}{2} \left(1 - \frac{27.26}{65.16} \right) \\
&= \frac{3}{2} (1 - 0.42) \\
&= .87 = \text{ค่าความเชื่อมั่นของเครื่องมือ}
\end{aligned}$$

7.2 องค์ประกอบที่มีอิทธิพลต่อความเชื่อมั่นของเครื่องมือ

มีองค์ประกอบจำนวนมากที่มีผลทำให้ความเชื่อมั่นของเครื่องมือต่ำหรือสูง ซึ่งองค์ประกอบเหล่านี้จะต้องนำมาใช้พิจารณาประกอบกับการแปลผลต่อค่าความเชื่อมั่น หรือในการเปรียบเทียบค่าความเชื่อมั่นระหว่างเครื่องมือ 2 หรือมากกว่า 2 ฉบับ เช่น ในการหาความเชื่อมั่นของเครื่องมือแบบการสอบซ้ำนั้น ช่วงระยะเวลาที่ใช้ในการทดสอบมีอิทธิพลต่อค่าความเชื่อมั่นของเครื่องมือเป็นอันมาก เพราะยิ่งถามซ้ำในเวลาใกล้เคียงกับการถามครั้งแรกเท่าใด ค่าความเชื่อมั่นของเครื่องมือก็จะยิ่งสูงมากขึ้นเท่านั้น ดังนี้ เป็นต้น

สิ่งที่มีอิทธิพลต่อความเชื่อมั่นของเครื่องมือที่สำคัญ มี 4 องค์ประกอบใหญ่ ๆ คือ

1. ความยาวของเครื่องมือ

โดยปกติทั่ว ๆ ไปแล้ว ความยาวของเครื่องมือมีความสำคัญมาก ถ้าแบบทดสอบหรือแบบสอบถามยังมีความยาวมากเท่าไร ความเชื่อมั่นของแบบทดสอบหรือแบบสอบถามก็จะยิ่งสูงขึ้น 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 ข้อนั้น ต้องเดา เขาอาจจะเดาถูกทั้ง 3 ข้อ ซึ่งเขาก็จะได้คะแนนเต็ม หรือเขาอาจจะเดาผิดทั้ง 3 ข้อ ซึ่งเขาก็จะได้คะแนนเพียง 7 คะแนน อย่างไรก็ตาม ถ้านักเรียนคนนั้นทำข้อสอบแบบถูก ผิด 100 ข้อ โอกาสที่เขาจะเดาถูก ก็จะลดน้อยลงไปเนื่องจากมีโอกาสดูผิดมากยิ่งขึ้น และคะแนนที่ได้ก็จะสามารถบ่งถึงความรู้ที่แท้จริงของเขา

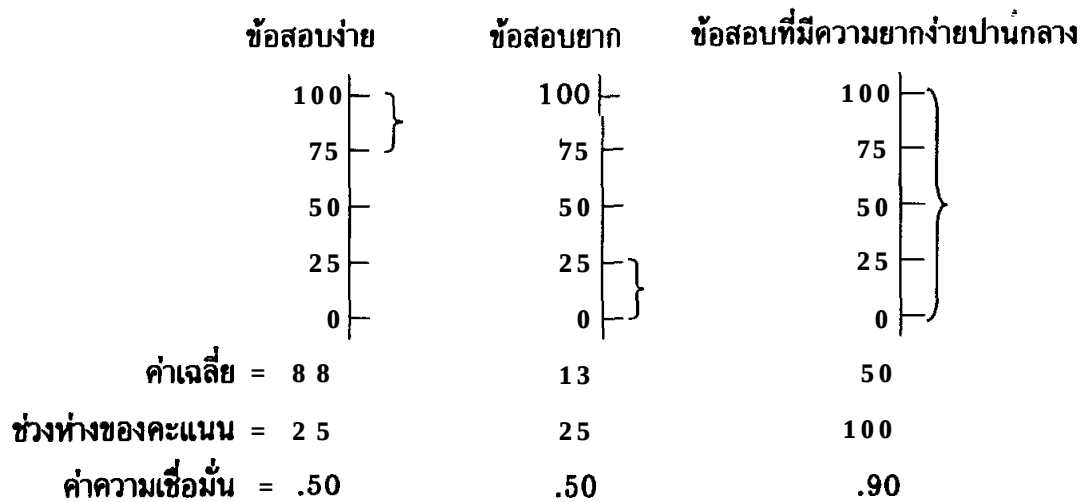
มีตัวอย่างที่แสดงให้เห็นว่า ความยาวของแบบทดสอบมีแนวโน้มที่จะเพิ่มค่าความเชื่อมั่นของแบบทดสอบโดยดูจากการหาความเชื่อมั่นด้วยวิธี Split - half เช่น ข้อสอบฉบับหนึ่ง ถ้าหาความเชื่อมั่นด้วยวิธี Split - half จะได้ค่าความเชื่อมั่น = .60 แต่เมื่อมาใช้สูตร Spearman-Brown ขยายเป็นความเชื่อมั่นของข้อสอบทั้งฉบับแล้วจะได้ = .75 จากตัวอย่างนี้ ก็แสดงให้เห็นอย่างชัดเจนว่า ถ้าจำนวนข้อสอบเพิ่มมากขึ้น ความเชื่อมั่นก็จะเพิ่มมากขึ้นตาม แต่ไม่ได้หมายความว่า จะต้องเพิ่มเป็นสัดส่วนตามจำนวนข้อสอบที่เพิ่มขึ้น

สิ่งสำคัญที่จะต้องระวังอีกประการหนึ่งในการเพิ่มจำนวนข้อในแบบทดสอบ หรือแบบสอบถามก็คือ ข้อสอบที่เพิ่มขึ้นนั้นจะต้องมีคุณภาพใกล้เคียงกับข้อคำถามเดิม ทั้งในด้านเนื้อหา ความยากง่ายและอำนาจจำแนก จึงจะมีผลต่อการเพิ่มค่าความเชื่อมั่นดังกล่าว

ในการสร้างข้อสอบเพื่อใช้ในชั้นเรียน สิ่งที่เราจะต้องตระหนักถึงมากที่สุดก็คือ 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 คะแนนก็จะมาออกกันตรงคะแนนสูง แต่ถ้าข้อสอบยากเกินไป ผู้สอบส่วนใหญ่ก็จะทำข้อสอบไม่ค่อยได้ เหมือน ๆ กัน คะแนนก็จะมาออกกันอยู่ตรงคะแนนต่ำ ๆ ความแตกต่างระหว่างกลุ่มจะน้อย (ดูรูป)



จากรูปเป็นการเปรียบเทียบการกระจายของคะแนนตามทฤษฎี จากการสอบข้อสอบ 100 ข้อ ความเชื่อมั่นหาได้จากการใช้สูตร Kuder - Richardson โดยการสมมติค่าความเบี่ยงเบนมาตรฐานขึ้นมาอย่างสมเหตุสมผล

จากความรู้เรื่องนี้ เราสามารถนำไปใช้ในการสร้างข้อสอบให้มีความเชื่อมั่นสูง ๆ ได้ โดยพยายามสร้างข้อสอบให้มีคะแนนเฉลี่ยของผู้เข้าสอบเท่ากับ 50% ของจำนวนข้อ และให้คะแนนกระจายอยู่ระหว่างคะแนนศูนย์ถึงคะแนนเต็ม

3. ความเป็นปรนัยของแบบทดสอบ

ความเป็นปรนัยของแบบทดสอบ หมายถึง คุณลักษณะสำคัญ 3 ประการของแบบทดสอบ คือ

1. ผู้สอบอ่านใจทบทวนแล้วต้องเข้าใจคำถามตรงกัน
2. ไม่ว่าจะให้ใครตรวจคำตอบ จะต้องได้คะแนนเท่ากัน
3. สามารถนำคะแนนมาแปลความหมายออกมาในรูปเดียวกัน

ข้อสอบมาตรฐานส่วนมาก ทั้งข้อสอบความถนัด และผลสัมฤทธิ์ก็จะต้องมีความเป็นปรนัยสูง เพราะความเป็นปรนัยของข้อสอบจะเป็นตัวช่วยควบคุมความแปรผันของคะแนนซึ่งจะมีผลทำให้ความเชื่อมั่นของข้อสอบสูงด้วย

4. ความเป็นเอกพันธ์ของกลุ่ม

ถ้ากลุ่มผู้เข้าสอบมีความเป็นเอกพันธ์ต่อกันมากเพียงใด ข้อสอบก็就会有ความเชื่อมั่นน้อยลงเท่านั้น ทั้งนี้เนื่องจากถ้านักเรียนมีความสามารถใกล้เคียงกันแล้ว คะแนนที่ได้จากการทดสอบก็จะเกาะกลุ่มใกล้เคียงกันมาก ความแปรปรวนของคะแนนก็จะน้อย ซึ่งจะส่งผลทำให้ความเชื่อมั่นของข้อสอบต่ำลงไปด้วย ถ้าผู้เข้าสอบมีความสามารถแตกต่างกันมาก ๆ จะส่งผลให้ค่าความเชื่อมั่นของข้อสอบสูงตามไปด้วย เรื่องนี้สามารถอธิบายให้เข้าใจได้อย่างชัดเจนโดยการศึกษาสูตรดังนี้

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

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 มีค่ามากจึงมีผลทำให้ r_{xx} มีค่าน้อยลง

7.3 วิธีหาความเที่ยงตรง (validity) ของเครื่องมือ

เครื่องมือวัดทางการศึกษาควรมีความเที่ยงตรง กล่าวคือเครื่องมือนั้นต้องสามารถวัดในสิ่งที่ต้องการวัดได้ และวัดได้ตรงจุดที่ต้องการวัดให้มากที่สุดเท่าที่จะเป็นไปได้

ความเที่ยงตรงของเครื่องมือแบ่งออกเป็น 3 ประเภท คือ

1. ความเที่ยงตรงตามเนื้อหา (Content validity)
2. ความเที่ยงตรงตามโครงสร้าง (Construct validity)
3. ความเที่ยงตรงตามเกณฑ์ (Criterion-related validity)

วิธีหาความเที่ยงตรงแต่ละประเภท มีวิธีการดังนี้

1. ความเที่ยงตรงตามเนื้อหา (Content validity)

เครื่องมือวัดที่มีความเที่ยงตรงตามเนื้อหา คือเครื่องมือที่สามารถวัดได้ตรงและครอบคลุมเนื้อหาสาระที่ต้องการจะวัดได้ครบถ้วนสมบูรณ์ เช่น “แบบทดสอบวิชาคณิตศาสตร์ ชั้น อ.3” คำถามต้องวัดเนื้อหาวิชาคณิตศาสตร์ชั้น อ.3 และครอบคลุมเนื้อหาที่เรียนทั้งหมดหรือ “แบบวัดเจตคติที่มีต่ออาชีพครู” คำถามที่ใช้ถามก็ต้องครอบคลุมเรื่องต่างๆ ที่เกี่ยวข้องกับอาชีพครู เป็นต้น เครื่องมือวัดสำหรับเด็กปฐมวัยควรมีความเที่ยงตรงตามเนื้อหาสูงกว่าแบบอื่นๆ

การหาความเที่ยงตรงตามเนื้อหาของเครื่องมือมีหลายวิธี ดังนี้

1. ใช้ผู้เชี่ยวชาญตรวจสอบ
2. ใช้วิธีการสอบก่อน-หลัง (Pretest - posttest Design)
3. ใช้วิธีการหาค่าสัมประสิทธิ์สหสัมพันธ์

วิธีการหาความเที่ยงตรงตามเนื้อหาแต่ละวิธีมีรายละเอียด ดังนี้

1. การใช้ผู้เชี่ยวชาญตรวจสอบ

การใช้ผู้เชี่ยวชาญตรวจสอบความเที่ยงตรงของเครื่องมือวัดทางการศึกษา เป็นวิธีที่นิยมใช้กันมากที่สุด การใช้ผู้เชี่ยวชาญตรวจสอบ ทำได้ 3 วิธี คือ

1) ใช้ดุลยพินิจของผู้เชี่ยวชาญด้านเนื้อหาที่เกี่ยวข้องกับแบบวัด พิจารณาข้อคำถามอย่างคร่าวๆ ว่า ข้อคำถามนั้นๆ มีความเที่ยงตรงตามเนื้อหาหรือไม่ ซึ่งผลการพิจารณาของผู้เชี่ยวชาญโดยวิธีนี้จะเป็นอัตนัยค่อนข้างสูง

2) ใช้ตารางวิเคราะห์เนื้อหา หรือตารางวิเคราะห์หลักสูตร วิธีนี้ทำได้โดยการพิจารณาสัดส่วนน้ำหนักของจำนวนข้อคำถามกับเนื้อหาที่นำมาสร้างข้อคำถาม

ส่วนใหญ่วิธีการหาความเที่ยงตรงโดยวิธีนี้จะเกี่ยวข้องกับแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน โดยที่ผู้สร้างแบบทดสอบจะต้องพยายามเขียนข้อคำถามให้ครอบคลุมเนื้อหาตามที่กำหนดไว้ในตารางวิเคราะห์เนื้อหาเพื่อการเขียนข้อคำถาม

ฉะนั้นในการคัดเลือกข้อคำถามที่มีคุณภาพดีมาใช้ทางการศึกษา จึงไม่ควรพิจารณาจากเฉพาะค่าสถิติรายข้อเท่านั้น แต่ควรต้องพิจารณาจากสัดส่วนจำนวนข้อคำถามที่เหมาะสมกับเนื้อหาแต่ละเนื้อหาที่จะถามด้วย

3) ใช้วิธีการทางสถิติ วิเคราะห์หาความเที่ยงตรงออกมาในเชิงปริมาณ หรือตัวเลข

วิธีหาค่าดัชนีความเที่ยงตรงตามเนื้อหาโดยใช้สถิติที่นิยมใช้กัน มี 3 วิธี คือ

1) ใช้ดัชนีความสอดคล้องระหว่างข้อคำถาม และลักษณะเฉพาะของกลุ่มพฤติกรรม วิธีนี้ได้โดยการนำเอาข้อคำถามที่สร้างขึ้น และลักษณะเฉพาะของกลุ่มพฤติกรรมไปให้ผู้เชี่ยวชาญ ประมาณ 10 คน พิจารณาลงความเห็น โดยมีเกณฑ์การให้คะแนน ดังนี้

ให้คะแนน +1 เมื่อผู้เชี่ยวชาญแน่ใจว่าข้อคำถามนั้นเป็นตัวแทนของเนื้อหาหรือลักษณะเฉพาะของกลุ่มพฤติกรรมนั้น

ให้คะแนน 0 เมื่อผู้เชี่ยวชาญไม่แน่ใจว่าข้อคำถามนั้นเป็นตัวแทนของเนื้อหาหรือลักษณะเฉพาะของกลุ่มพฤติกรรมหรือไม่

ให้คะแนน -1 เมื่อแน่ใจว่าข้อคำถามนั้นไม่เป็นตัวแทนของเนื้อหาหรือลักษณะเฉพาะของกลุ่มพฤติกรรมนั้น

วิธีหาค่าดัชนีความเที่ยงตรง มีสูตรดังนี้

$$IOC = \frac{\sum R}{N}$$

เมื่อ IOC = ดัชนีความสอดคล้องระหว่างข้อคำถามกับลักษณะเฉพาะของกลุ่มพฤติกรรม (Index of objective congruency)

$\sum R$ = ผลรวมของคะแนนความคิดเห็นในแต่ละข้อของผู้เชี่ยวชาญทั้งหมด

N = จำนวนผู้เชี่ยวชาญทั้งหมด

เกณฑ์การพิจารณา

ถ้า IOC มากกว่าหรือเท่ากับ 0.5 แสดงว่าคำถามข้อนั้นใช้ได้ (เป็นตัวแทนของลักษณะเฉพาะของกลุ่มพฤติกรรมนั้น)

ถ้า IOC น้อยกว่าหรือเท่ากับ 0.5 แสดงว่าคำถามข้อนั้นไม่ดี ควรตัดทิ้งไป หรือต้องทำการปรับปรุงใหม่ให้ดีขึ้น

2) ใช้ดัชนีความเหมาะสมระหว่างข้อคำถาม และลักษณะเฉพาะของกลุ่มพฤติกรรม วิธีนี้ทำได้โดยการนำลักษณะเฉพาะของกลุ่มพฤติกรรม ไปให้ผู้เชี่ยวชาญเฉพาะแต่ละคนพิจารณาว่าข้อคำถามนั้นเหมาะสมที่จะเป็นตัวแทนลักษณะเฉพาะกลุ่มพฤติกรรมหรือไม่ โดยกำหนดเกณฑ์การให้คะแนน ดังนี้

ให้คะแนน 5	เมื่อข้อคำถามนั้นเหมาะสมมากที่สุด
ให้คะแนน 4	เมื่อข้อคำถามนั้นเหมาะสมมาก
ให้คะแนน 3	เมื่อข้อคำถามนั้นเหมาะสมปานกลาง
ให้คะแนน 2	เมื่อข้อคำถามนั้นเหมาะสมน้อย
ให้คะแนน 1	เมื่อข้อคำถามนั้นเหมาะสมน้อยที่สุด

แล้วนำข้อคำถามแต่ละข้อที่ผู้เชี่ยวชาญได้พิจารณาให้คะแนนมาคำนวณหาสิ่งต่อไปนี้

1) หาค่าเฉลี่ย ($\bar{X}$) จากสูตร $\bar{X} = \frac{\sum x}{N}$

2) หาค่าความเบี่ยงเบนมาตรฐาน (S) จากสูตร

$$S = \sqrt{\frac{N\sum x^2 - (\sum x)^2}{N(N-1)}}$$

เกณฑ์การพิจารณา

ข้อคำถามที่ดีคือข้อคำถามที่มีค่า $\bar{X} \geq 3.5$, $S \leq 1.00$

3) ใช้ดัชนีการจับคู่ระหว่างข้อคำถามและลักษณะเฉพาะของกลุ่มพฤติกรรม วิธีนี้ทำได้โดยการนำลักษณะเฉพาะของกลุ่มพฤติกรรมและข้อคำถามที่เป็นตัวแทนของพฤติกรรมที่สร้างขึ้น ไปให้ผู้เชี่ยวชาญพิจารณาตัดสินว่าข้อคำถามใดเป็นตัวแทนของพฤติกรรมใด ถ้าข้อคำถามใดเป็นตัวแทนของกลุ่มพฤติกรรมใดก็ให้จับคู่กัน แล้วหาค่าดัชนีของการจับคู่ จากสูตร

$$M = \left(\frac{f}{N} \right) \times 100$$

เมื่อ	M	=	ดัชนีการจับคู่
	f	=	จำนวนความถี่ของผู้เชี่ยวชาญที่พิจารณาตัดสินว่าข้อ คำถามนั้นจับคู่ได้กับลักษณะเฉพาะของกลุ่มพฤติกรรมนั้น
	N	=	จำนวนผู้เชี่ยวชาญทั้งหมด

เกณฑ์การตัดสิน

ข้อคำถามที่มีความเที่ยงตรงเชิงเนื้อหา จะต้องมีค่า M มากกว่า 80%

2. ใช้วิธีสอนก่อน - หลัง

การหาความเที่ยงตรงตามเนื้อหาของเครื่องมือวิธีนี้ทำได้โดยทำแบบทดสอบไปทดสอบกับนักเรียนก่อนทำการสอนเนื้อหานั้น แล้วให้ตรวจกระดาษคำตอบและเก็บคะแนนไว้ หลังจากนั้นก็ให้ทำการสอนเนื้อหานั้นกับนักเรียนกลุ่มเดิม แล้วนำแบบทดสอบฉบับเดิมไปทดสอบกับนักเรียนกลุ่มเดิมซ้ำอีกครั้งหนึ่ง แล้วทำการทดสอบดูว่า คะแนนที่ได้จากการทดสอบครั้งหลังเพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติหรือไม่ ถ้าใช่ก็แสดงว่า เนื้อหาของแบบทดสอบครอบคลุมเนื้อหาที่ต้องการวัด นั่นคือแบบทดสอบนั้นมีความเที่ยงตรงตามเนื้อหาจริง

3. ใช้วิธีหาค่าสัมประสิทธิ์สหสัมพันธ์

การหาค่าความเที่ยงตรงตามเนื้อหาของแบบทดสอบหรือแบบสอบถามโดยใช้วิธีหาค่าสัมประสิทธิ์สหสัมพันธ์สามารถทำได้โดยกำหนดให้ผู้เชี่ยวชาญ 2 กลุ่ม สร้างแบบทดสอบ หรือแบบสอบถาม เกี่ยวกับเนื้อหาที่ต้องการวัดแบบเดียวกัน แล้วนำแบบทดสอบหรือแบบสอบถามทั้ง 2 ฉบับที่เป็นคู่ขนานกันไปทดลองกลุ่มตัวอย่างกลุ่มเดียวกัน แล้วนำคะแนนที่ได้จากการทดลองเครื่องมือที่เป็นคู่ขนานกัน มาหาค่าสัมประสิทธิ์สหสัมพันธ์ โดยใช้สูตร

$$r_{xy} = \frac{N\sum xy - \sum x \sum y}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}}$$

เมื่อ x = คะแนนที่ได้จากการทดลองใช้เครื่องมือฉบับที่ 1

y = คะแนนที่ได้จากการทดลองใช้เครื่องมือฉบับที่ 2 (ซึ่งเป็นคู่ขนานกับแบบที่ 1)

ถ้าค่า r_{xy} มีค่าตั้งแต่ 0.50 ถึง 1.00 ก็อาจสรุปได้ว่าเครื่องมือที่สร้างขึ้นมีความเที่ยงตรงตามเนื้อหา

2. ความเที่ยงตรงตามโครงสร้าง (Construct validity)

เครื่องมือที่มีความเที่ยงตรงตามโครงสร้าง หมายถึง เครื่องมือที่สามารถวัดได้ตามลักษณะ หรือตามทฤษฎีต่าง ๆ ของโครงสร้างนั้น เช่น

- ลักษณะความคิดสร้างสรรค์
- ความมีวินัยในตัวเอง
- ความซื่อสัตย์
- พฤติกรรมด้านพุทธิพิสัย

ฯลฯ

เราสามารถสร้างเครื่องมือวัดลักษณะของสิ่งเหล่านี้ได้ โดยนิยามคุณลักษณะที่ต้องการวัดให้ชัดเจนว่า คุณลักษณะเหล่านั้นครอบคลุมพฤติกรรมอะไรบ้าง แล้วจึงสร้างข้อคำถามให้ครอบคลุมลักษณะย่อย ๆ เหล่านั้น

การหาความเที่ยงตรงตามโครงสร้างของเครื่องมือวัดมีหลายวิธี ในที่นี้จะเสนอวิธีที่นิยมใช้กันมาก 4 วิธี คือ

- 1) หาความสัมพันธ์ระหว่างคะแนนรวมทั้งฉบับกับคะแนนเฉพาะข้อนั้น ๆ
- 2) หาความสัมพันธ์ระหว่างเครื่องมือที่ได้มาตรฐานกับเครื่องมือที่ต้องการ

หาความเที่ยงตรง

- 3) เปรียบเทียบกลุ่มที่แตกต่างกัน (Group differences)
- 4) การวิเคราะห์องค์ประกอบ (Factor analysis)

วิธีการหาความเที่ยงตรงตามโครงสร้าง แต่ละวิธีมีรายละเอียด ดังนี้

- 1) หาความสัมพันธ์ระหว่างคะแนนรวมทั้งฉบับกับคะแนนเฉพาะข้อนั้น ๆ

วิธีนี้ใช้หาความเที่ยงตรงตามโครงสร้างของเครื่องมือวัดโดยถือคะแนนรวมเป็นคะแนนที่แทนโครงสร้างรวมของลักษณะที่ต้องวัด อนาสตาซี (Anastasi, Anne, 1988 : 155) เรียกวิธีการหาความเที่ยงตรงวิธีนี้ว่า Internal consistency

ถ้าข้อคำถามใดมีค่าสัมประสิทธิ์สหสัมพันธ์กับคะแนนรวมสูง แสดงว่าข้อคำถามนั้นวัดสิ่งที่เป็นโครงสร้างของคุณลักษณะนั้น

สูตรที่ใช้หาความเที่ยงตรงตามโครงสร้าง ใช้สูตรสหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation) ดังนี้

$$r_{xy} = \frac{N\sum xy - \sum x \sum y}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}}$$

เมื่อ r_{xy} = ค่าความเที่ยงตรงตามโครงสร้าง

x = คะแนนรายข้อของแต่ละคน

y = คะแนนรวมของแต่ละคน

N = จำนวนคนที่น่าเครื่องมือไปทดลองใช้

ค่าความเที่ยงตรงตามโครงสร้างควรมีค่า r ตั้งแต่ 0.50 ถึง 1.00

2) หาความสัมพันธ์ระหว่างเครื่องมือที่ได้มาตรฐานกับเครื่องมือที่ต้องหา

ความเที่ยงตรง

การหาความเที่ยงตรงโดยวิธีนี้ขึ้นอยู่กับคุณภาพของเครื่องมือที่เป็นมาตรฐานด้วย โดยที่เครื่องมือมาตรฐานจะต้องวัดคุณลักษณะหรือโครงสร้างที่สอดคล้องกับสิ่งที่ต้องการวัด

สูตรการคำนวณ ใช้สูตร

$$r_{xy} = \frac{N\sum xy - \sum x \sum y}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}}$$

เมื่อ r_{xy} = ค่าความเที่ยงตรงตามโครงสร้าง

x = คะแนนรวมของแต่ละคนจากเครื่องมือที่สร้างขึ้น

y = คะแนนรวมของแต่ละคนจากเครื่องมือมาตรฐาน

N = จำนวนคนที่น่าเครื่องมือไปทดลอง

3) เปรียบเทียบกลุ่มที่แตกต่างกัน

วิธีหาความเที่ยงตรงตามโครงสร้างของเครื่องมือโดยการเปรียบเทียบกลุ่มที่แตกต่างกัน ทำได้โดยการนำเครื่องมือที่สร้างขึ้นไปทดลองใช้กับกลุ่มตัวอย่าง 2 กลุ่มที่ต่างกันอย่างชัดเจน เช่น กลุ่มเก่ง-กลุ่มอ่อน กลุ่มมีความพร้อม-กลุ่มไม่มีความพร้อม แล้วทดสอบความแตกต่างของคะแนนเฉลี่ยของทั้ง 2 กลุ่ม โดยใช้การทดสอบแบบที (t-test) จากสูตร t-test ตามเงื่อนไข ดังนี้

กรณีที 1 : ถ้า $\sigma_1^2 = \sigma_2^2$ ใช้สูตรที่ 1

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \dots\dots\dots(1)$$

$$df = n_1 + n_2 - 2$$

เมื่อ $S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$

$$S_1^2 = \frac{n_1 x_1^2 - (\sum x_1)^2}{n_1 (n_1 - 1)},$$

$$S_2^2 = \frac{n_2 x_2^2 - (\sum x_2)^2}{n_2 (n_2 - 1)}$$

กรณีที 2 : ถ้า $\sigma_1^2 \neq \sigma_2^2$ ใช้สูตรที่ 2

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \dots\dots\dots(2)$$

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} \right)^2}{\frac{\left(\frac{S_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2} \right)^2}{n_2 - 1}}$$

เมื่อ σ^2 = ความแปรปรวนของประชากร
 s^2 = ความแปรปรวนของกลุ่มตัวอย่าง
 n_1 = จำนวนกลุ่มตัวอย่างกลุ่มที่ 1
 n_2 = จำนวนกลุ่มตัวอย่างกลุ่มที่ 2

ในกรณีที่เราไม่ทราบว่า $\sigma_1^2 = \sigma_2^2$ หรือไม่ แต่ถ้าเราควบคุมให้ $n_1 = n_2$ แล้วเราก็สามารถใช้สูตรที่ 1 (เช่นเดียวกับเมื่อทราบว่า $\sigma_1^2 = \sigma_2^2$) ได้โดยไม่ผิดข้อตกลงเบื้องต้น

ถ้าทดสอบความแตกต่างของคะแนนเฉลี่ยของคำตอบของทั้ง 2 กลุ่ม แล้วพบว่าคะแนนเฉลี่ยแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ก็สามารถสรุปได้ว่าเครื่องมือที่มีความเที่ยงตรงตามโครงสร้าง

4) การวิเคราะห์องค์ประกอบ

เป็นวิธีการทางสถิติขั้นสูงสำหรับตรวจสอบองค์ประกอบหรือคุณลักษณะเฉพาะของเครื่องมือ ผลที่ได้จากการวิเคราะห์องค์ประกอบจะบ่งบอกให้ทราบว่าเครื่องมือเหล่านั้นมีลักษณะหรือโครงสร้างร่วมกันหรือไม่

3. ความเที่ยงตรงตามเกณฑ์ (Criterion - relates validity)

การหาความเที่ยงตรงตามเกณฑ์ของแบบทดสอบ มีจุดประสงค์จะดูว่าคะแนนที่ได้จากการทดสอบแบบทดสอบฉบับนั้น สามารถนำมาใช้พยากรณ์ผลการเรียนภายภาคหน้า (Predictive validity) หรือนำมาคาดคะเนผลการเรียนในปัจจุบันได้ดีเพียงใด (Concurrent Validity) ความเที่ยงตรงตามเกณฑ์มีอยู่ 2 ชนิด คือ ความเที่ยงตรงตามพยากรณ์ (Predictive Validity) กับความเที่ยงตรงตามสภาพ (Concurrent Validity) การหาความเที่ยงตรงทั้ง 2 ชนิดนี้ เป็นการหาความสัมพันธ์ระหว่างคะแนนของแบบทดสอบที่จะหาความเที่ยงตรงกับตัวเกณฑ์ (Criteria) ในการหาความเที่ยงตรงตามพยากรณ์นั้น ตัวเกณฑ์ที่จะใช้เปรียบเทียบคือ คะแนนที่ได้จากผลการเรียน ซึ่งจะต้องติดตามไปในอนาคต ดังนั้น ข้อสอบที่ใช้ในการสอบคัดเลือกจึงต้องมีความเที่ยงตรงตามพยากรณ์สูง ส่วนการหาความเที่ยงตรงตามสภาพ ตัวเกณฑ์จะได้อาจจากการที่ครูจัดอันดับความสามารถที่นักเรียนแสดงออกในห้องเรียนให้ครูสังเกตเห็นได้ ถ้าข้อสอบมีความเที่ยงตรงตามสภาพสูงก็แสดงว่าคะแนนที่ได้จากแบบทดสอบฉบับนั้นสอดคล้องเป็นไปในทางเดียวกันกับอันดับที่ครูจัด นั่นคือ ถ้าเขาได้คะแนนจากแบบทดสอบสูงก็ต้องได้อันดับดีด้วย ในปัจจุบันนี้ความเที่ยงตรงทั้ง 2 ชนิดนี้ มักเรียกรวมเป็นชื่อเดียวกันว่า Criterion - Related Validity ทั้งนี้เนื่องจากความเที่ยงตรงทั้ง 2 ชนิดนี้ มีหลักการและความหมายเป็นแบบเดียวกันแต่จะต่างกันตรงระยะเวลาที่จะได้ตัวเกณฑ์มาเปรียบเทียบ

ในการศึกษาเกี่ยวกับความเที่ยงตรงตามเกณฑ์นั้นเราจะต้องพิจารณาหาตัวเกณฑ์อย่างรอบคอบ สมมติว่าเราต้องการจะทราบว่าจะเกณฑ์ที่ได้จากแบบทดสอบความถนัดสามารถนำมาใช้ในการพยากรณ์ “ความสำเร็จในการเรียน” มากน้อยเพียงใด ในที่นี้ความสำเร็จในการเรียนเป็นตัวเกณฑ์ เราจะวัดความสำเร็จในการเรียนได้อย่างไร โดยปกติแล้ว นักการศึกษามักจะใช้เกรดเฉลี่ย (GPA) แทนความสำเร็จในการเรียน นั่นคือถ้านักเรียนได้เกรดเฉลี่ยสูง แสดงว่าเขาประสบความสำเร็จในการเรียนสูง แต่ถ้าเขาได้เกรดเฉลี่ยต่ำก็แสดงว่าเขาประสบความสำเร็จในการเรียนน้อย หรือไม่ประสบความสำเร็จในการเรียนเลย อย่างไรก็ตามการใช้เกรดเฉลี่ยเพียงอย่างเดียวแทนความหมายของความสำเร็จในการเรียนนั้น นักการศึกษาส่วนมากตระหนักว่าไม่เป็นการถูกต้องมากนัก น่าที่จะใช้สัดส่วนระหว่างจำนวนผู้ที่เรียนจบกับผู้ที่เรียนไม่จบประกอบด้วย ตัวอย่างที่คล้ายคลึงกับที่กล่าวไปแล้วข้างต้น คือ ถ้าเราพยายามจะพยากรณ์ความสำเร็จในการทำงาน ในกรณีนี้ อันดับของสมรรถภาพในการทำงานของพนักงาน ซึ่งพิจารณาโดยหัวหน้าหน่วยงานมักจะถูกนำไปใช้เป็นตัวเกณฑ์ในการวัด ซึ่งอันดับที่กล่าวนี้อาจจะเป็นตัวเกณฑ์ที่ไม่เหมาะสมพอ ถ้าคะแนนจากแบบทดสอบไม่มีความสัมพันธ์กับอันดับที่ให้โดยหัวหน้าหน่วยงาน เราจะตอบไม่ได้โดยแน่ใจเลยว่าเกณฑ์ที่ได้ค่าสัมประสิทธิ์สหสัมพันธ์ต่ำนั้นเกิดจากแบบทดสอบไม่ใช่ตัวพยากรณ์ที่ดีของความสำเร็จในการทำงาน 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

สมมติว่าสุพจน์ ซึ่งเป็นนักเรียนชั้นอนุบาล 3 ต้องการจะรู้ว่าคะแนนที่ได้จากการทดสอบความถนัดทางการเรียนจะสามารถพยากรณ์ผลสัมฤทธิ์ในการเรียนเลขคณิต ชั้นอนุบาล 3 ของนักเรียนเขาได้ดีเพียงไร เนื่องจากนักเรียนทุกคนที่เข้ามาเรียนในชั้นอนุบาล 3 นี้ต้องผ่านการสอบความถนัดทางการเรียนก่อนเป็นอันดับแรก ปัญหาที่สำคัญที่สุดของสุพจน์ ก็คือขาดเกณฑ์ (Criterion) ที่ดีพอสำหรับความสำเร็จในการเรียนเลขคณิต เขาตัดสินใจใช้คะแนนจากการสอบไล่ ซึ่งส่วนใหญ่ถามเกี่ยวกับวิชาเลขคณิตระดับอนุบาล 3 เป็นเกณฑ์ ปัญหาอยู่ที่ว่านักเรียนที่ได้คะแนนความถนัดทางการเรียนสูงจะได้คะแนนจากการสอบไล่สูงด้วยหรือไม่ หรือถ้าเขาได้คะแนนจากการสอบถามความถนัดทางการเรียนต่ำ เขาจะได้คะแนนในการสอบไล่ต่ำไปด้วยหรือไม่

จากตัวอย่าง สุ่มพบว่าส่วนใหญ่คะแนนที่ได้จากการทดสอบความถนัด กับคะแนนผลสัมฤทธิ์ขึ้นลงตามกัน ซึ่งก็เป็นแต่เพียงการพิจารณาขั้นต้นเท่านั้น ส่วนวิธีการจะหาค่าความเที่ยงตรงตามเกณฑ์ที่ถูกต้องตามหลักแล้ว จะต้องนำคะแนนที่ได้จากการทดสอบ 2 ชุดนั้น มาหาค่าสหสัมพันธ์ (correlation coefficient) กัน ดังนี้

คะแนนที่ได้จากการสอบ 2 ครั้ง นำมาลงตารางได้ดังนี้

นักเรียน คนที่	คะแนน ความถนัด	คะแนน เลขคณิต อ.3	อันดับของ คะแนน ความถนัด	อันดับของ คะแนน เลขคณิต	ความแตกต่าง ของอันดับ (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 $\longrightarrow$ น้อย แล้วรอกลงคะแนน
คณิตศาสตร์แต่ละคนทำได้ลงไปคู่กัน
2. จัดอันดับของคะแนนจาก 1 $\longrightarrow$ N (จำนวนคนในกลุ่ม)
3. หาค่าความแตกต่าง (D) ของอันดับโดยเอาคะแนนความถนัดตั้ง ลบด้วย
คะแนนคณิตศาสตร์
4. ยกกำลังสองของผลต่างของอันดับ (D^2)
5. หาผลบวกของกำลังสองของผลต่าง ($\sum D^2$)
6. คำนวณหาค่าความสัมพันธ์จากสูตร

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

Σ = ผลบวกของ

D = ความแตกต่างของอันดับ

N = จำนวนคนในกลุ่ม

P = เป็นตัวอักษรกรีกใช้แทนค่า rank - order correlation coefficient ซึ่งมีความหมายเช่นเดียวกับค่า r

ฉะนั้น

$$\begin{aligned} \rho &= 1 - \frac{6 \times 532}{20(20^2 - 1)} \\ &= 1 - \frac{3192}{7980} \\ &= 1 - .40 = .60 \end{aligned}$$

วิธีหาค่าสัมประสิทธิ์สหสัมพันธ์แบบนี้ เป็นวิธีการที่ง่ายเป็นการหาโดยจัดอันดับ
ของคะแนน แต่ถ้าจำนวนกลุ่มตัวอย่างมาก ๆ ก็มีวิธีการหาแบบอื่นที่สะดวกกว่านี้ เช่น หาจากคะแนน
ดิบโดยตรงเลย โดยใช้สูตร

$$r_{xy} = \frac{N\sum xy - \sum x \sum y}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}}$$

เมื่อ r_{xy} = ค่าความเที่ยงตรงของแบบทดสอบ

ความหมายของสัมประสิทธิ์สหสัมพันธ์

1.00 = มีความสัมพันธ์กันอย่างดีในทางบวก

.00 = ไม่มีความสัมพันธ์กันเลย

-1.00 = มีความสัมพันธ์กันอย่างดียิ่งในทางลบ แต่ค่านี้ไม่นิยมใช้ในทางวัดผลเพราะขัดแย้งกับความรู้สึกของคนทั่วไป

ค่าความเที่ยงตรงของแบบทดสอบที่สุพจน์สร้างขึ้น มีค่า .60 ซึ่งนับว่า

คะแนนระหว่างการสอบ 2 ฉบับ มีความสัมพันธ์กันค่อนข้างสูง ซึ่งถ้าพิจารณาดูค่าประสิทธิภาพในการพยากรณ์จากตาราง Expectancy Table จะเห็นว่า พวกที่ได้คะแนนสูงเป็นกลุ่มที่ 1 ในวิชาความถนัดจะมีโอกาสได้คะแนนสูงเป็นกลุ่มที่ 1 ในวิชาคณิตศาสตร์ ด้วย

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

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

จากตารางจะเห็นว่าพวกที่สอบได้คะแนนสูงในวิชาความถนัด ไม่มีใครได้คะแนนต่ำในวิชาเลข และพวกที่ได้คะแนนต่ำในวิชาความถนัดก็ไม่มีใครได้คะแนนสูงในวิชาเลขเลย จึงสรุปได้ว่าข้อสอบความถนัดนี้มีความเที่ยงตรงตามเกณฑ์อยู่ในเกณฑ์ดี

7.4 องค์ประกอบที่มีอิทธิพลต่อความเที่ยงตรงของเครื่องมือ

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

1. องค์ประกอบที่เกิดขึ้นจากตัวแบบทดสอบเอง

องค์ประกอบที่มีผลทำให้ผลที่ได้จากการทดสอบขาดความเที่ยงตรง มีดังนี้

- 1) คำชี้แจงไม่ชัดเจน เช่น ไม่บอกวิธีการตอบ หรือบอกไม่ชัดเจน เปิดโอกาสให้เดาได้หรือไม่ สิ่งเหล่านี้มีผลทำให้ข้อสอบขาดความเที่ยงตรง ทั้งสิ้น
 - 2) ใช้ภาษาและศัพท์ที่สูงหรือยากเกินไป เพราะเราไม่ได้วัดความสามารถในการอ่าน
 - 3) ความยากง่ายไม่พอเหมาะ ข้อสอบที่ยากเกินไปหรือง่ายเกินไปจะไม่มีอำนาจในการจำแนกเด็ก ซึ่งจะมีผลทำให้ความเที่ยงตรงของข้อสอบต่ำลงไปด้วย
 - 4) ข้อคำถามไม่ดี คือ ข้อสอบที่สร้างขึ้นมาอย่างไม่มีหลักเกณฑ์ไม่สามารถวัดในสิ่งที่ต้องการจะวัดได้
 - 5) คำถามคลุมเครือ ซึ่งบางทีนั้นคำถามที่คลุมเครือ หรือกำกวมนี้อาจมีผลทำให้เด็กเก่งสามารถเดาคำตอบได้ถูกมากกว่าเด็กอ่อน
 - 6) ข้อคำถามไม่วัดพฤติกรรมที่ต้องการจะวัด เช่น ต้องการวัดความสามารถทางด้านความเข้าใจหรือทักษะ แต่กลับออกคำถามวัดแต่ความรู้ในเรื่อง
 - 7) ข้อสอบนั้นสั้นเกินไป แบบทดสอบแต่ละฉบับเป็นแต่เพียงกลุ่มตัวอย่างของคำถามทั้งหมด ถ้าเราออกข้อสอบน้อยเกินไป ข้อสอบก็ไม่สามารถเป็นตัวแทนของคำถามที่แท้จริงได้
 - 8) การเรียงข้อสอบที่ไม่ถูกต้อง ตามปกติแล้วข้อสอบควรจะเรียงจากข้อง่ายไปหาข้อยาก ถ้าข้อสอบเรียงข้อยาก ๆ ไว้ต้น ๆ แล้วก็จะเป็ผลเสียทางด้านแรงจูงใจ
 - 9) เรียงคำตอบถูกไว้อย่างมีระบบ เช่น กก ขข หรือ กขคจ กขคจ ซึ่งจะช่วยให้ผู้ตอบเดาได้
- ถ้าครูพยายามจัดสิ่งต่าง ๆ ที่กล่าวมาแล้ว ก็อาจจะช่วยให้ข้อสอบมีความเที่ยงตรงสูงขึ้น

2. องค์ประกอบที่เกี่ยวกับวิธีดำเนินการสอบและการให้คะแนน

การไม่คอยเคร่งครัดเวลาในการเก็บข้อสอบ การอธิบายเพิ่มเติมเมื่อมีข้อสงสัย การทำเสียงดังในระหว่างการสอบ รวมทั้งการให้คะแนนข้อสอบแบบอัตนัย การไม่ปฏิบัติตามคำสั่งชี้แจง ที่กำหนดไว้สิ่งเหล่านี้มีผลทำให้ข้อสอบมีความเที่ยงตรงต่ำทั้งสิ้น

3. องค์ประกอบที่เกี่ยวกับการตอบของนักเรียน

แรงจูงใจมีผลทำให้ความสามารถในการตอบของนักเรียนแตกต่างกัน เด็กบางคน อาจเกิดความกลัวสภาพการณ์ในการสอบ จึงไม่สามารถจะทำข้อสอบได้อย่างปกติ หรือการเดาตอบก็มีผลทำให้ข้อสอบขาดความเที่ยงตรง เช่น ในข้อสอบแบบถูกผิด นักเรียนคนหนึ่งอาจจะเดาถูกตลอด หรืออีกคนหนึ่ง อาจจะเดาข้อสอบผิดพลาดตลอด

4. องค์ประกอบที่เกี่ยวกับความเชื่อมั่นของตัวพยากรณ์กับเกณฑ์

7.5 ความสัมพันธ์ระหว่างความเชื่อมั่นกับความเที่ยงตรง

ความเที่ยงตรงของแบบทดสอบ จะมีค่าไม่เกินรากที่สองของค่าความเชื่อมั่น นั่นคือ

$$\begin{aligned} r_{xy} &= \sqrt{r_{xx}} \\ \text{เมื่อ } r_{xy} &= \text{ค่าความเที่ยงตรงสูงสุดของแบบทดสอบ} \\ r_{xx} &= \text{ค่าความเชื่อมั่นของแบบทดสอบ} \end{aligned}$$

ตัวอย่าง : แบบทดสอบฉบับหนึ่งมีค่าความเชื่อมั่น .79 อยากทราบว่าแบบทดสอบฉบับนี้จะมีค่าความเที่ยงตรงสูงสุด ไม่เกินเท่าใด

$$\begin{aligned} r_{xy} &= \sqrt{.79} \\ &= 0.89 \end{aligned}$$

คำถามท้ายบทที่ 7

1. จงอธิบายความหมายและความสัมพันธ์ของความเชื่อมั่นและความเที่ยงตรงของเครื่องมือที่ใช้ในการประเมินเด็กปฐมวัย
2. ท่านคิดว่าความเที่ยงตรงแบบใดจำเป็นมากที่สุดสำหรับเครื่องมือที่ใช้ในการประเมินระดับปฐมวัย และให้อธิบายวิธืหาความเที่ยงตรงดังกล่าวอย่างละเอียด
3. ความเชื่อมั่นของเครื่องมือที่ใช้ในการประเมินมีวิธีการหาที่วิธี จงอธิบายพร้อมยกตัวอย่างประกอบ
4. ท่านจะมีวิธีปรับปรุงค่าความเชื่อมั่นและความเที่ยงตรงของเครื่องมือที่ใช้ในการประเมินได้ด้วยวิธีใดบ้างจงอธิบาย