

วิธีการหาคุณภาพเครื่องมือวัดและประเมินผลการเรียนรู้

การสร้างเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล ครูผู้สอนต้องหาคุณภาพของเครื่องมือเพื่อเป็นการยืนยันว่า เครื่องมือดังกล่าวมีคุณภาพซึ่งการหาคุณภาพของเครื่องมือสามารถจำแนกเป็น 2 ลักษณะคือ

1. การหาคุณภาพของเครื่องมือทั้งฉบับ

การวิเคราะห์ข้อสอบทั้งฉบับ เป็นการตรวจสอบคุณภาพของเครื่องมือวัด เกี่ยวกับความเที่ยงตรง (Validity) และความเชื่อมั่น (Reliability)

(1) ความเที่ยงตรง หมายถึง ความสามารถของเครื่องมือวัด ที่สามารถวัดได้ในสิ่งที่ต้องการวัด เป็น ความสอดคล้องระหว่างผลการวัด กับสิ่งที่ต้องการวัด ความตรงที่ใช้ในการทดสอบจำแนกเป็น 3 ชนิด ได้แก่ ความตรง ตามเนื้อหา ความตรงตามโครงสร้าง และ ความตรงตามเกณฑ์ที่เกี่ยวข้อง โดยแบบทดสอบวัดผลสัมฤทธิ์ จะเกี่ยวข้อง กับความตรงตามเนื้อหา มากกว่าความตรงชนิดอื่น ๆ

การหาค่าความเที่ยงตรงเชิงเนื้อหา (Content Validity) เป็นการหาค่าความเที่ยงตรงที่ให้ผู้เชี่ยวชาญ พิจารณาว่าข้อสอบ หรือ ข้อคำถามแต่ละข้อ วัดได้ตรงตามสิ่งที่ต้องการวัดเนื้อหาหรือวัตถุประสงค์การเรียนรู้มากน้อย เพียงใด โดยใช้เกณฑ์การประเมิน ดังนี้

ให้คะแนน +1 หมายถึง แน่ใจว่าข้อสอบวัดจุดประสงค์/เนื้อหานั้น

ให้คะแนน 0 หมายถึง ไม่แน่ใจว่าข้อสอบวัดจุดประสงค์/เนื้อหานั้น

ให้คะแนน -1 หมายถึง แน่ใจว่าข้อสอบไม่วัดจุดประสงค์/เนื้อหานั้น แล้วนำข้อมูลที่ได้จากการพิจารณาของผู้เชี่ยวชาญ หาค่าความสอดคล้องระหว่างข้อคำถามแต่ละข้อ กับจุดประสงค์หรือเนื้อหา (Index of Item-Objective Congruence หรือ IOC) จากสูตร

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

เมื่อ $\sum R$ แทน ผลรวมของคะแนนการพิจารณาของผู้เชี่ยวชาญ
N แทน จำนวนผู้เชี่ยวชาญ

เกณฑ์การตัดสินค่า IOC ถ้ามีค่า 0.50 ขึ้นไป แสดงว่า ข้อคพถามนั้นวัดได้ตรงจุด ประสงค์ หรือตรงตาม เนื้อหานั้น แสดงว่า ข้อคกถามข้อนั้น ใช้ได้

(2) ความเชื่อมั่น ความเชื่อมั่น หมายถึง ความคงที่ของคะแนนที่วัดได้
แต่ละครั้ง วิธีการหาค่าความ เชื่อมั่นของแบบทดสอบทำได้หลายวิธี คือ

1. วิธีสอบซ้ำ
2. วิธีแบบทดสอบคู่ขนาน
3. วิธีหาความสอดคล้องภายใน แบ่งเป็น
 - 3.1 วิธีแบ่งครึ่งแบบทดสอบ
 - 3.2 วิธีหาจากสูตรคูเดอร์และริชาร์ดสัน
 - 3.3 วิธีหาจากสูตรสัมประสิทธิ์แอลฟา

จากสูตรของสเปียร์แมน บราวน์ (Spearman Brown) ดังนี้

$$R_{tt} = \frac{2r_{mm}}{1+r_{mm}}$$

เมื่อ R_{tt} แทน ความเที่ยงแบบแบบทดสอบทั้งฉบับ

R_{mm} แทน สัมประสิทธิ์สหสัมพันธ์ของแบบทดสอบครึ่งฉบับ

วิธีหาจากสูตรของคูเดอร์และริชาร์ดสัน

การหาความเที่ยงโดยวิธีนี้ เป็นการหาความสัมพันธ์ของคะแนนจากการใช้แบบทดสอบ
ฉบับเดียวและสอบเพียงครั้งเดียวโดยนำผลการสอบมาคำนวณค่าสัมประสิทธิ์ ใช้สูตร
ของคูเดอร์และริชาร์ดสันซึ่งเป็นการหาความเที่ยงของ แบบทดสอบที่มีระบบการให้
คะแนนแบบ 0,1 (ผิด 0, ถูก 1) สูตรที่ใช้มี 2 สูตร คือ สูตร KR - 20 กับสูตร KR - 21
สูตร KR -20 ในกรณีที่ค่าความยากง่ายของข้อสอบแต่ละข้อไม่เท่ากัน

$$R_{KR-20} = \frac{K}{K-1} \left[1 - \frac{\sum pq}{S^2} \right]$$

- เมื่อ R_{tt} แทน ความเที่ยงของแบบทดสอบ
 K แทน จำนวนข้อสอบ
 P แทน ความยากง่ายของข้อสอบแต่ละข้อ (สัดส่วนที่ตอบถูก)
 q แทน สัดส่วนที่ตอบผิด ($1-p$)
 S^2 แทน ความแปรปรวนของคะแนนรวมของแบบทดสอบ

$$S^2 = \frac{N \sum X^2 - (\sum X)^2}{N^2}$$

สูตร KR-21 ในกรณีที่ค่าความยากง่ายของข้อสอบทุกข้อเท่ากันหรือไม่แตกต่างกันมาก

$$R_{KR-21} = \frac{K}{K-1} \left[\frac{1 - \bar{X}(K - \bar{X})}{KS^2} \right]$$

- เมื่อ R_t แทน ความเที่ยงของแบบทดสอบ
 K แทน จำนวนข้อสอบ
 X แทน ค่าเฉลี่ยของคะแนนรวมของแบบทดสอบทั้งฉบับ
 S^2 แทน ความแปรปรวนของคะแนนรวมของแบบทดสอบ

สูตร KR - 20 และ KR - 21 นี้ใช้ได้เฉพาะการหาความเที่ยงของแบบทดสอบที่ให้คะแนนแต่ละข้อ เป็นแบบ 0 กับ 1 เท่านั้น สูตร KR-21 ใช้ในกรณีข้อสอบทุกข้อมีค่าความยากเท่ากันซึ่งในทางปฏิบัติต้องพิจารณาเงื่อนไขที่เป็นจริงด้วย

วิธีหาจากสูตรสัมประสิทธิ์ แอลฟา การหาค่าความเที่ยงโดยใช้สูตรของครอนบัค (Cronbach) นี้ปรับมาจากสูตร KR – 20 ใช้หาความเที่ยงของเครื่องมือวัดที่ให้คะแนนแตกต่างกันไปในแต่ละข้อได้ โดยไม่จำเป็นต้องเป็นระบบการให้ คะแนนแบบ ๑ กับ ๐ สูตรการคำนวณเป็นดังนี้

การหาค่าความเที่ยงโดยใช้สูตรสัมประสิทธิ์แอลฟา สามารถหาได้โดยใช้ผลการสอบจากแบบทดสอบฉบับเดียว นำไปสอบกับบุคคลกลุ่มเดียว และนำไปใช้กันได้อย่างกว้างขวาง โดยไม่จำกัดเฉพาะแบบทดสอบที่ให้คะแนนแบบ 1 กับ 0

2. การวิเคราะห์หาคุณภาพของข้อสอบรายข้อ

การวิเคราะห์ ข้อสอบรายข้อ เป็นการตรวจสอบคุณภาพของข้อสอบแต่ละข้อโดยพิจารณาจากสมบัติที่สำคัญ 3 ประการ ได้แก่ ความยาก อำนาจจำแนกแลประสิทธิภาพของตัวลง

1.1 ความยากของข้อสอบ. ความยากของข้อสอบ หมายถึง สัดส่วนของจำนวนผู้ที่ตอบข้อสอบได้ถูกต้องต่อจำนวนผู้ที่ตอบข้อสอบทั้งหมด หรือหมายถึงจำนวน

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

เมื่อ α แทน ความเที่ยงของแบบทดสอบ

S^2 แทน ความแปรปรวนของข้อสอบแต่ละข้อ

S^2 แทน ความแปรปรวนของคะแนนรวมของแบบทดสอบ

K แทน จำนวนข้อสอบทั้งหมด

ร้อยละของผู้ตอบข้อสอบนั้น ๆ ถูก ตัวอย่างเช่น ค่า $p = 0.30$ แสดงว่า จำนวนผู้ตอบ 100 คนมีผู้ที่ตอบข้อนั้น ๆ ถูก 30 คน ค่าความยากง่ายจะมีค่าระหว่าง 0 ถึง 1.00 สามารถหาได้จาก

สูตร

$$P = \frac{R_H + R_L}{N_H + N_L}$$

P คือ ความยากง่าย

R_H คือ จำนวนนักเรียนที่ตอบถูกในกลุ่มคะแนนสูง

R_L คือ จำนวนนักเรียนที่ตอบถูกในกลุ่มคะแนนต่ำ

N_H คือ จำนวนนักเรียนทั้งหมดในกลุ่มคะแนนสูง

N_L คือ จำนวนนักเรียนทั้งหมดในกลุ่มคะแนนต่ำ

ในการพิจารณาค่าความยากง่ายนั้น ถ้าข้อสอบมีค่าความยากง่ายสูง เช่น $p = 0.95$ แสดงว่า มีผู้ตอบ ถูกจำนวนมาก จึงถือว่าเป็นข้อสอบที่ง่าย แต่ในทางกลับกัน ถ้าข้อสอบมีผู้ตอบถูกน้อย เช่น $p = 0.15$ แสดงว่าเป็นข้อสอบที่ยาก ข้อสอบที่ดีจะมีระดับความยากง่าย เท่ากับ 0.5 ซึ่งจะทำให้เกิดค่าอำนาจการจำแนกสูงสุดและมี ความเชื่อมั่นสูง อย่างไรก็ตามในการสอบวัดความรู้ผลการเรียน โดยทั่วไป มักนิยมให้มีข้อสอบที่มีระดับความยากง่ายใน ระดับต่าง ๆ ปะปนกันไป โดยจัดให้มีข้อสอบมีค่าความยากง่ายพอเหมาะ (p มีค่าใกล้เคียง 0.5) เป็นส่วนใหญ่ รวมทั้งให้มีข้อสอบที่ค่อนข้างยาก และค่อนข้างง่ายอีกจำนวนหนึ่ง แต่ถ้าเป็นการสอบแข่งขันเพื่อคัดเลือกผู้ที่มีความรู้ความสามารถควรมีสัดส่วนของข้อสอบที่ยากสูงขึ้น ทั้งนี้ ข้อสอบที่ดีควรมีค่าความยากง่ายระหว่าง 0.20 – 0.80 ใน ข้อสอบประเภท 4 ตัวเลือก ส่วนข้อสอบประเภทถูก – ผิด ค่าความยากง่าย ควรอยู่ระหว่าง 0.60 - 0.70

เกณฑ์การแปลความหมายค่าความยากง่าย (p) ของข้อสอบ (ล้วน สายยศ และ อังคณา สายยศ, ๒๕๔๓)

ความยากง่ายของข้อสอบ (p)	ความหมาย
๐.๘๑ - ๑.๐๐	ง่ายมาก (ควรปรับปรุงหรือตัดทิ้ง)
๐.๖๐ - ๐.๘๐	ค่อนข้างง่าย (ดี)
๐.๔๐ - ๐.๕๙	ยากพอเหมาะ (ดีมาก)
๐.๒๐ - ๐.๓๙	ค่อนข้างยาก (ดี)
๐ - ๐.๑๙	ยากมาก (ควรปรับปรุงหรือตัดทิ้ง)

1.2 อำนาจจำแนก (r) หมายถึง ความสามารถของข้อสอบในการจำแนกหรือแยกให้เห็นความแตกต่างระหว่างผู้สอบที่มีผลสัมฤทธิ์ต่างกัน เพื่อที่จะใช้พยากรณ์หรือบ่งชี้ความแตกต่างที่เห็นชัดในด้านความสามารถ เช่น จำแนกคนเก่งกับคนอ่อนจากกันได้ โดยถือว่าคนเก่งควรทำข้อสอบข้อนั้นได้ ส่วนผู้ที่ย่อมนไม่ควรทำข้อสอบข้อนั้นได้ อำนาจจำแนกของข้อสอบ จะมีค่าตั้งแต่ - 1 ถึง + 1 ค่าอำนาจจำแนกที่ดี ควรมีค่าตั้งแต่ 0.20 ขึ้นไป

$$r = \frac{R_H - R_L}{N_H \text{ or } N_L}$$

r คือ ค่าอำนาจจำแนก

R_H คือ จำนวนนักเรียนที่ตอบถูกในกลุ่มคะแนนสูง

R_L คือ จำนวนนักเรียนที่ตอบถูกในกลุ่มคะแนนต่ำ

N_H คือ จำนวนนักเรียนทั้งหมดในกลุ่มคะแนนสูง

N_L คือ จำนวนนักเรียนทั้งหมดในกลุ่มคะแนนต่ำ

กรณีที่ค่า r ติดลบ แสดงว่า ข้อสอบข้อนั้นจำแนกกลับ คนเก่งทพไม่ได้ แต่คนอ่อนทำได้ ถือว่าเป็น ข้อสอบที่ไม่ดีควรตัดทิ้ง นอกจากนี้ อาจารย์ผู้สอนควรตรวจสอบการจัดการเรียนสอนของตน ว่าเพราะเหตุใดผู้ที่เรียน เก่งจึงไม่เข้าใจ ในเรื่องที่สอน

เกณฑ์การแปลความหมายค่าอำนาจจำแนก (r) ของข้อสอบ

อำนาจจำแนกของข้อสอบ (r)	ความหมาย
๐.๖๐ - ๑.๐๐	อำนาจจำแนกดีมาก
๐.๔๐ - ๐.๕๙	อำนาจจำแนกดี
๐.๒๐ - ๐.๓๙	อำนาจจำแนกพอใช้
๐.๑๐ - ๐.๑๙	อำนาจจำแนกต่ำ (ควรปรับปรุงหรือตัดทิ้ง)
-๐.๐๐ - ๐.๐๙	อำนาจจำแนกต่ำมาก (ควรปรับปรุงหรือตัดทิ้ง)

ขั้นตอนการหาค่าความยากง่ายและอำนาจจำแนก

1. ตรวจให้คะแนนข้อสอบ แล้วเรียงกระดาษคำตอบจากคะแนนมากไปหาน้อย
2. แบ่งกระดาษคำตอบออกเป็น 2 กลุ่ม กลุ่มแรกเรียกว่า กลุ่มสูง(PH) โดยนับจากคะแนนสูงลงมา ประมาณ 27% ของกระดาษคำตอบทั้งหมด และกลุ่มหลังเรียกว่ากลุ่มต่ำ (PL) โดยนับจากคะแนนต่ำขึ้นไปประมาณ 27% ของกระดาษคำตอบทั้งหมด
3. หาจำนวนคนที่ตอบถูกในแต่ละข้อของกลุ่มสูงและกลุ่มต่ำ
4. หาค่าความยาก (P) และค่าอำนาจจำแนก (r) ตามสูตร