

บทที่ 8

การตรวจสอบคุณภาพของเครื่องมือเป็นรายข้อ

เนื้อหา

เนื้อหาที่จะกล่าวถึงในบทที่ 8 มีรายละเอียดดังนี้

8.1 การวิเคราะห์ข้อคำถามของแบบทดสอบอิงกลุ่ม

8.2 การวิเคราะห์ข้อคำถามของแบบทดสอบอิงเกณฑ์

8.3 การวิเคราะห์ข้อคำถามของแบบสอบถาม

วิธีการตรวจสอบคุณภาพของเครื่องมือเป็นรายข้อมีรายละเอียดดังนี้

8.1 การวิเคราะห์ข้อคำถามของแบบทดสอบอิงกลุ่ม

1. การแปลความหมายข้อมูลที่ได้จากการวิเคราะห์ข้อสอบรายข้อ
 2. การวิเคราะห์ข้อสอบโดยใช้สูตรคำนวณ
 3. การหาค่าอำนาจจำแนกข้อสอบเป็นรายข้อโดยการใช้สูตรสหสัมพันธ์
1. การแปลความหมายข้อมูลที่ได้จากการวิเคราะห์ข้อสอบรายข้อ

ขบวนการวิเคราะห์ข้อสอบเกี่ยวข้องโดยตรงกับการกำหนดระดับความยาก (p) และค่าอำนาจจำแนก (r) ของข้อสอบแต่ละข้อ จึงมีความจำเป็นที่ผู้อ่านจะต้องเข้าใจความหมายของคำว่า “ระดับความยาก” และ “อำนาจจำแนก” ของข้อสอบเสียก่อนจึงจะทำให้สามารถแปลความหมายข้อมูลที่ได้จากการวิเคราะห์ข้อสอบได้อย่างถูกต้อง คำวิเคราะห์ของตัวถูกนี้เราถือว่าเป็นตัวเลขที่แสดงถึงคุณภาพของข้อสอบข้อนั้น

ข้อมูลที่ได้จากการวิเคราะห์ข้อสอบ แยกได้เป็น 2 พวก คือ ข้อมูลของตัวเลือกที่เป็นตัวถูก กับข้อมูลของตัวเลือกที่เป็นตัวลวง ดังนั้นในการแปลความหมายข้อมูลที่ได้จากการวิเคราะห์ข้อสอบจึงต้องแยกอธิบายเป็น 2 พวก ดังนี้

ก. การแปลความหมายค่าวิเคราะห์ของตัวถูก

1. ระดับความยาก (Level of difficulty : p)

ระดับความยากของข้อคำถามใด (ใช้แทนด้วยอักษร p) หมายถึงจำนวนเปอร์เซ็นต์หรือสัดส่วน (percentage or proportion) ของนักเรียนที่เลือกตอบตัวเลือกนั้น เมื่อพิจารณาเปรียบเทียบกับจำนวนนักเรียนที่เข้าสอบทั้งหมด ดังนั้น

$$P = \frac{\text{จำนวนคนที่เลือกตอบตัวเลือกนั้น}}{\text{จำนวนคำตอบทั้งหมด}}$$

ถ้าตัวเลือกที่นำมาวิเคราะห์เป็นตัวเลือก ค่า p ก็จะมีค่าตั้งแต่เปอร์เซ็นต์ หรือ สัดส่วนที่ผู้เข้าสอบกลุ่มนั้นทำข้อสอบถูก ดังนั้น

$$P = \frac{\text{จำนวนคนที่ตอบข้อสอบข้อนั้นถูก}}{\text{จำนวนคนตอบทั้งหมด}}$$

ลักษณะความยากของข้อสอบแต่ละข้อ หรือแต่ละตัวเลือกมี ดังนี้

- 1) p (ระดับความยาก) มีค่าตั้งแต่ .00 ถึง 1.00
- 2) ถ้านักเรียนเลือกตอบด้วยตัวเลือกใดมาก แสดงว่าข้อหรือตัวเลือกนั้นง่าย ค่าของ p จะสูง
- 3) ถ้านักเรียนเลือกตอบตัวเลือกได้น้อย แสดงว่าข้อหรือตัวเลือกนั้นยาก ค่าของ p จะต่ำ
- 4) ถ้านักเรียนทำข้อสอบข้อใดถูก 100% (ไม่ผิดเลย) แสดงว่าข้อสอบข้อนั้นง่ายมาก ($p = 1.00$)
- 5) ถ้าข้อสอบข้อใดไม่มีคนตอบถูกเลย หมายความว่าข้อสอบข้อนั้นยากมาก ($p = .00$)
- 6) ข้อสอบที่มีค่า $p = .50$ ถือว่าข้อสอบข้อนั้นยากพอเหมาะ หมายความว่า เป็นข้อสอบที่มี 50% หรือครึ่งหนึ่งของผู้เข้าสอบตอบถูก อย่างไรก็ตามถ้าข้อสอบข้อ 1 และข้อ 5 ต่างก็มีค่า $p = .50$ ไม่ได้หมายความว่าผู้ที่ตอบข้อ 1 ถูก จะต้องตอบข้อ 5 ถูกด้วย ผู้ที่ตอบข้อ 5 ถูก อาจจะเป็นนักเรียนคนอื่น แต่จะมีจำนวนเท่ากับครึ่งหนึ่งของผู้เข้าสอบเช่นกัน
- 7) การแปลความหมายระดับความยากของข้อสอบ
 - p .81 - 1.00 เป็นข้อสอบที่ง่ายมาก
 - p .61 - .80 เป็นข้อสอบที่ค่อนข้างง่าย
 - p .41 - .60 เป็นข้อสอบที่ง่ายพอเหมาะ
 - p .21 - .40 เป็นข้อสอบที่ค่อนข้างยาก
 - p .00 - .20 เป็นข้อสอบที่ยากมาก

2. อำนาจจำแนก (Discrimination index : r)

อำนาจจำแนก (ใช้แทนด้วยอักษร r) ของข้อคำถามใด หมายถึง ประสิทธิภาพของข้อคำถามในการแบ่งนักเรียนออกเป็นสองกลุ่ม คือกลุ่มเก่งที่มีความรู้ กับกลุ่มอ่อนที่มีความรู้น้อย ค่า r คือ ค่าสหสัมพันธ์ (correlation) ระหว่างคะแนนที่ได้จากการตอบข้อสอบข้อนั้น กับคะแนนรวมทั้งฉบับ ของผู้เข้าสอบคนเดียวกันนั้น วิธีการหาค่าอำนาจจำแนกมีหลายวิธีซึ่งจะได้อธิบายให้ละเอียดในเนื้อหาตอนต่อไป ค่าอำนาจจำแนกของข้อสอบมีความหมายดังนี้

- 1) ค่าอำนาจจำแนก หรือค่า r มีค่าตั้งแต่ -1.00 ถึง +1.00
- 2) ข้อสอบข้อใดที่นักเรียนกลุ่มเก่งทำถูกหมดทุกคนและนักเรียนกลุ่มอ่อนทำผิดหมดทุกคน r จะมีค่าเท่ากับ 1.00 จัดเป็นข้อสอบที่ดีเยี่ยมเป็นที่ต้องการยิ่ง
- 3) ข้อสอบข้อใดที่นักเรียนกลุ่มเก่งทำผิดหมด แต่นักเรียนกลุ่มอ่อนกลับทำถูกหมด r จะมีค่าเท่ากับ -1.00 แสดงว่าข้อสอบข้อนี้เป็นข้อสอบที่ไม่ดี เพราะเป็นข้อสอบที่ลวงเด็กเก่ง ซึ่งผิดจุดประสงค์ เพราะในการทดสอบใดนั้น ครูผู้ออกข้อสอบหวังว่านักเรียนที่เก่งจะต้องทำข้อสอบได้ แต่นักเรียนอ่อนจะทำข้อสอบไม่ได้
- 4) ข้อสอบข้อใดมีจำนวนนักเรียนกลุ่มเก่งและจำนวนนักเรียนกลุ่มอ่อนทำถูกเท่า ๆ กัน ข้อสอบข้อนั้นจะมีค่า r เท่ากับ .00 แสดงว่าข้อสอบข้อนี้เป็นข้อสอบที่ไม่สามารถแยกนักเรียนเก่งออกจากนักเรียนอ่อนได้ นั่นคือ เราไม่สามารถจะบอกได้เลยว่าคนที่ทำข้อสอบนี้ถูก เป็นเด็กเก่งและมีความรู้ในเรื่องที่ถามจริง ๆ หรือตอบถูกเพราะการเดา

เกณฑ์ที่ใช้พิจารณาค่าอำนาจจำแนกของข้อสอบมีดังนี้

ดัชนีอำนาจจำแนก (ของตัวถูก)	คุณภาพของข้อสอบ
.40 ขึ้นไป	ดีมาก
.30 ถึง .39	ดีพอสมควร อาจต้องปรับปรุงบ้าง
.20 ถึง .29	พอใช้ได้แต่ต้องปรับปรุง
ต่ำกว่า .20	ใช้ไม่ได้ ต้องพิจารณาปรับปรุงใหม่หรือตัดทิ้งไปเลย

ข. การแปลความหมายค่าวิเคราะห์ของตัวลง

1. ค่า p จะเป็นตัวอย่างบ่งบอกให้ทราบว่า มีนักเรียนเลือกตอบตัวลงเป็นสัดส่วนเท่าใด ถ้าค่าต่ำมาก แสดงว่าตัวลงนั้นไม่ค่อยมีใครเลือก เห็นสมควรจะต้องปรับปรุงเพราะไม่เกิดประโยชน์ใด ๆ ที่จะเขียนไว้เฉย ๆ โดยไม่มีใครเลือกตอบ

2. ค่า r มีหลักในการพิจารณา ดังนี้

- ถ้าค่า r ติดลบ แสดงว่าตัวลงนั้นไม่จำเป็นต้องแก้ไข
- ถ้าค่า r เป็นศูนย์ แสดงว่าตัวลงนั้นไม่จำเป็นต้องแก้ไข
- ถ้าค่า r เป็นบวก แสดงว่าตัวลงนั้นเป็นตัวลงที่ดี

2. การวิเคราะห์ข้อสอบโดยใช้สูตรคำนวณ

ในการวิเคราะห์ข้อสอบนั้น เมื่อครูสร้างข้อสอบและนำไปทดสอบกับนักเรียนแล้ว ครูจะต้องนำเอาผลการสอบนั้นมาตรวจให้คะแนน แล้วมานั่งนับเป็นรายข้อว่า ข้อใดนักเรียนทำถูกกี่คน คิดเป็นเปอร์เซ็นต์ จากนั้นก็ใช้วิธีการทางสถิติคำนวณหาค่าต่าง ๆ ที่เราต้องการออกมาอย่างไรก็ดีการวิเคราะห์แบบนี้นอกจากครูจะต้องมีความรู้ และความชำนาญในวิชาสถิติแล้ว ยังเป็นการเสียเวลามาก จึงมีนักวัดผลคิดวิธีการวิเคราะห์ข้อสอบแบบง่าย ๆ ขึ้นมาเรียกว่า “เทคนิค 27%” นั่นคือ ครูไม่ต้องเอาคำตอบทั้งหมด (100%) มาวิเคราะห์เอามาแต่เพียง 27% ของเด็กเก่ง ๆ ที่ได้คะแนนมาก ๆ กับ 27% ของเด็กอ่อน ๆ ที่ได้คะแนนน้อย ๆ ก็พอ เหลือตรงกลางเป็นเด็กที่ได้คะแนนกลาง ๆ อีก 46% ไม่ต้องนำมาคิดการที่ใช้ 27% ไม่ใช่ 25% หรือ 30% ก็เพราะว่ามีผู้ทำวิจัยแล้วพบว่า การตัดหัวตัดหางมาคิดวิเคราะห์อย่างละ 27% นี้จะให้ผลได้เหมือนกันกับนำผลการสอบของนักเรียนทั้งหมด (100%) มาวิเคราะห์เช่นกัน

การใช้เทคนิค 27% มาเป็นหลักในการหาค่าความยากง่าย และค่าอำนาจจำแนกนั้นข้อสอบที่จะนำมาวิเคราะห์โดยวิธีนี้จะต้องเป็นข้อสอบที่ให้คะแนนตอบถูก 1 และตอบผิด 0 เท่านั้น ข้อสอบประเภทที่ให้คะแนนแตกต่างกันไปจากนี้ไม่สามารถนำมาวิเคราะห์ได้ด้วยวิธีนี้

ขั้นตอนการดำเนินการวิเคราะห์มีดังนี้

1. ตรวจให้คะแนนกระดาษคำตอบ
2. เรียงกระดาษคำตอบของนักเรียนจากคะแนนสูงสุดมาหาคะแนนต่ำสุดกระดาษคำตอบที่มีคะแนนเท่ากันให้รวมอยู่ในลำดับเดียวกันโดยจะให้แผ่นใดอยู่ก่อนหลังก็ได้
3. คำนวณว่า 27% ของกลุ่มที่เข้าสอบทั้งหมดคิดเป็นจำนวนเท่าใด โดยเอา 0.27 คูณจำนวนทั้งหมด เช่น มีผู้เข้าสอบ 80 คน 27% จะเท่ากับ $80 \times 0.27 = 21.6$ หรือ 22 คน

(ถ้าตัวเลขหลังจุดทศนิยมสูงกว่า .5 ให้ปัดเป็นเลขจำนวนเต็ม 1 ถ้าน้อยกว่าหรือเท่ากับ .5 ให้ปัดทิ้งไป เช่น 20.4 ให้คิดเป็น 20 คน) ถ้ามีผู้สอบ 120 คน 27% จะเท่ากับ $120 \times 0.27 = 32.4 = 32$ คน

ในการวิเคราะห์ข้อสอบโดยใช้ตารางสำเร็จของ Chung - Teh Fan นั้น ผู้ทำการวิเคราะห์ต้องระลึกไว้ในใจว่า วิธีนี้เหมาะสำหรับการวิเคราะห์ข้อสอบที่มีคนสอบจำนวนมาก (เกิน 30 คนขึ้นไป) ถ้ามีคนสอบเพียง 20 คน ไม่นิยมนำมาวิเคราะห์ด้วยวิธีนี้ เพราะผลที่ได้จะคลาดเคลื่อนไปควรวิเคราะห์ด้วยวิธีอื่น (เช่น ใช้สูตรคำนวณอย่างง่าย ๆ) แต่ถ้ามีผู้เข้าสอบหลายพันคน ก็ไม่จำเป็นต้องนำกระดาษคำตอบทั้งหมดมาวิเคราะห์ เพราะเป็นการเสียเวลามากเกินไป เราอาจจะสุ่มตัวอย่างมาเพียงบางส่วน แล้วจึงหา 27% ของกลุ่มที่เราสุ่มมา เพื่อเลือกจำนวนกระดาษคำตอบเป็นกลุ่มสูงและกลุ่มต่ำต่อไป

4. แบ่งกระดาษคำตอบของผู้เข้าสอบออกเป็นกลุ่มสูงและกลุ่มต่ำ (ในกรณีที่มีคนสอบทั้งหมด 80 คน 27% ของกลุ่มสูงมีจำนวนเท่ากับ 22 คน) การนับจำนวนของกลุ่มสูงให้นับจากผู้ที่ได้คะแนนสูงสุดลงไป ส่วนการนับจำนวนของกลุ่มต่ำที่ให้นับจากผู้ที่ได้คะแนนต่ำสุดขึ้นไป สำหรับกลุ่มกลางที่เหลือให้ตัดทิ้งไปไม่ต้องนำมาคำนวณ ตัวอย่างเช่น ถ้ามีกระดาษคำตอบที่จะเอามาวิเคราะห์ทั้งหมด 80 คน ก็ให้นับจากคนที่ได้คะแนนสูงสุดลงมา 22 คน เป็นพวกกลุ่มสูงและนับจากคนที่ได้คะแนนต่ำสุดขึ้นไป 22 คน เป็นพวกกลุ่มต่ำ

5. นำกระดาษคำตอบของกลุ่มสูงไปทำรอยคะแนน (tally) เพื่อให้ทราบว่าในข้อคำถามแต่ละข้อนั้นคนกลุ่มสูง 22 คน ได้เลือกตอบตัวเลือกใดในข้อสอบแต่ละข้อจาก ก. ถึง จ. (ถ้ามี 5 ตัวเลือก) เพื่อจะได้หาผลรวมของแต่ละตัวเลือก (H) และจะได้นำไปคิดเป็นค่าสัดส่วนของกลุ่มสูง (P_H)

6. นำกระดาษของคนกลุ่มต่ำไปทำรอยคะแนนโดยวิธีการแบบเดียวกับที่อธิบายไว้ในข้อ 5 เพื่อจะได้หาผลรวมของแต่ละตัวเลือกของคำถามแต่ละข้อ (L) และนำไปคิดหาค่าสัดส่วนของกลุ่มต่ำ (P_L) การขีดรอยคะแนนของกลุ่มต่ำให้แยกทำในกระดาษคนละแผ่นกับของกลุ่มสูง

ตัวอย่าง : ตารางวิเคราะห์ข้อสอบวิชาภาษาไทย ชั้น อ.3 จำนวนข้อสอบ 50 ข้อ
ทำการทดสอบกับนักเรียน 80 คน แสดงไว้ในตาราง 1

ตาราง 1 ตาราง tally ของกลุ่มสูง วิชาภาษาไทย ชั้น อ.3 กลุ่มตัวอย่าง 80 คน

ข้อที่ งานที่	1					2					...	50				
	ก	ข	ค	ง	จ	ก	ข	ค	ง	จ	...	ก	ข	ค	ง	จ
1	/									/		/				
2	/									/		/				
3		/				/				/		/				
4		/														
20		/								/				/		
21		/								/				/		
22	/									/		/				
รวม (H)	3	15	3	1		4	3	2	1	12	...	10		2	7	3
P_H	14	68	.14	.04	.00	18	14	09	04	55	...	45	.00	09	32	14

การแปลความหมายตาราง tally ของกลุ่มสูง

- ตัวเลือกที่มีเครื่องหมายวงกลมล้อมรอบคือตัวถูก จากตารางข้างบน ข้อ 1 ตัวถูกคือ ข. ข้อ 2 ตัวถูกคือ จ.
- จากช่องรวม (H) แสดงให้เห็นว่า ข้อ 1 ก. มีคนเลือกตอบ 3 คน ข้อ ข. มีคนเลือกตอบ 15 คน ข้อ 1 ค. มีคนเลือกตอบ 3 คน
- จากช่อง P_H ค่า P_H ของข้อ 1 ก. = $.14 = \frac{3}{22}$

ตาราง 2 ตาราง tally ของกลุ่มต่ำ วิชาภาษาไทย ชั้น อ.3 กลุ่มตัวอย่าง 80 คน

ข้อที่ เนที่	1					2					50				
	ก	ข	ค	ง	จ	ก	ข	ค	ง	จ	ก	ข	ค	ง	จ
1	/								/							
2	/							/						/		
3		/							/			/				
4		/								/					/	
20	/					/										/
21	/						/								/	
22	/							/				/				
รวม (H)	6	7	4	2	3	2	5	4	4	7	10	-	3	6	3
P _L	.27	.32	.18	.09	.14	.09	.23	.18	.18	.3245	.00	.14	.27	.14

การแปลความหมายตาราง tally ของกลุ่มต่ำ แปลได้เช่นเดียวกับของกลุ่มสูง

7. คำนวณค่า P_H และ P_L ของแต่ละตัวเลือกจากสูตร

$$P_H = \frac{\text{จำนวนคนที่ตอบตัวเลือกนั้น}}{\text{จำนวนคนทั้งหมดในกลุ่มสูง}}$$

$$P_L = \frac{\text{จำนวนคนที่ตอบตัวเลือกนั้น}}{\text{จำนวนคนทั้งหมดในกลุ่มต่ำ}}$$

หมายเหตุ

จำนวนคนทั้งหมดในกลุ่มสูง ย่อมเท่ากับจำนวนคนทั้งหมดในกลุ่มต่ำเสมอ
เพราะคิดมาจาก 27% ของคนสอบทั้งหมด

จากตัวอย่างตารางวิเคราะห์ข้อสอบวิชาภาษาไทย ชั้น อ.3 ค่า P_H และ P_L
ของข้อสอบข้อ 1 หาได้ดังนี้

ข้อ	ตัวเลือก	P_H	P_L
1	ก	$3/22 = .14$	$6/22 = .27$
	ข	$15/22 = .68$	$7/22 = .32$
	ค	$3/22 = .14$	$4/22 = .18$
	ง	$1/22 = .04$	$2/22 = .09$
	จ	$0/22 = .00$	$3/22 = .14$

8. สำหรับตัวเลือกที่เป็นคำตอบถูกให้ใช้สูตรคำนวณข้างล่างนี้

$$P = \frac{H + L}{n_H + n_L} \quad (\text{หรือ} = \frac{P_H + P_L}{2})$$

$$r = \frac{H - L}{n_H} \quad \text{หรือ} = \frac{H - L}{n_L} \quad (\text{หรือ} = P_H - P_L)$$

เมื่อ p = ระดับความยากง่ายของข้อสอบ

r = อำนาจจำแนกของข้อสอบ

H = จำนวนคนที่เลือกตอบตัวเลือกใด ๆ ของกลุ่มสูง ซึ่งในที่นี้คือ
ช่องรวมตามแนวขี้นนั่นเอง

L = จำนวนคนที่เลือกตอบตัวเลือกใด ๆ ของกลุ่มต่ำ ซึ่งในที่นี้คือ
ช่องรวมตามแนวขี้น

n_H = จำนวนคนทั้งหมดของกลุ่มสูง (27% ของคนสอบทั้งหมด)

n_L = จำนวนคนทั้งหมดของกลุ่มต่ำ (27% ของคนสอบทั้งหมด) ซึ่ง

$n_H = n_L$ เสมอ

ตัวอย่าง วิธีคำนวณหาค่า p และ r

จากตาราง tally ของกลุ่มสูง และกลุ่มต่ำวิชาภาษาไทย ชั้น อ.3 ข้อ 1 (คำตอบถูก คือตัวเลือก ข.) คำนวณหาค่า p และ r ได้ดังนี้

ตัวเลือก	ค่า H	ค่า L	P	r
น	3	6	.20	.14
ข	15	7	.50	.36
ค	3	4	.16	.05
ง	1	2	.07	.05
จ	0	3	.07	.14

$$\begin{aligned} \text{ค่า } p \text{ ของตัวถูก (ตัวเลือก ข.)} &= \frac{H + L}{n_H + n_L} \\ &= \frac{15 + 7}{22 + 22} = \frac{22}{44} = .5 \end{aligned}$$

$$\begin{aligned} \text{ค่า } r \text{ ของตัวถูก} &= \frac{H - L}{n_H} \\ &= \frac{15 - 7}{22} = \frac{8}{22} = .36 \end{aligned}$$

9. สำหรับตัวเลือกที่เป็นตัวลวง ในการหาค่าอำนาจจำแนก (r) ของตัวลวงใช้สูตรข้างล่างนี้

$$r = \frac{L - H}{n_H} \text{ หรือ } = \frac{L - H}{n_L} \text{ หรือ } = P_L - P_H$$

สำหรับสูตรการหาค่า p ของตัวลวง ใช้สูตรเดียวกันกับการหาค่า p ของตัว

ถูกคือ

$$P = \frac{H + L}{n_H + n_L}$$

ตัวอย่าง การหาค่า p และ r ของตัวลงของข้อสอบข้อ 1

ตัวเลือก	$r = \frac{L - H}{n_H}$	$P = \frac{H + L}{n_H + n_L}$
ก	$\frac{6 - 3}{22} = .14$	$\frac{6 + 3}{22 + 22} = .20$
ค	$\frac{4 - 3}{22} = .05$	$\frac{4 + 3}{22 + 22} = .16$
ง	$\frac{2 - 1}{22} = .05$	$\frac{2 + 1}{22 + 22} = .07$
จ	$\frac{3 - 0}{22} = .14$	$\frac{3 + 0}{22 + 22} = .07$

10. การแปลความหมายผลการวิเคราะห์

ข้อ	ตัวเลือก	P_H	P_L	p	r	ความหมาย
1	ก	.14	.27	.20	.13	ดี
	ข	.68	.32	.50	.36	ดี
	ค	.14	.18	.16	.04	ดี
	ง	.04	.09	.06	.05	ดี
	จ	.00	.14	.07	.14	ดี
2	ก	.18	.09	.13	-.09	ไม่ดีต้องปรับปรุง
	ข	.14	.23	.18	.09	ดี
	ค	.09	.18	.13	.09	ดี
	ง	.04	.18	.10	.14	ดี
	จ	.55	.32	.43	.23	ดี
50	ก	.45	.45	.45	.00	ไม่ดีต้องปรับปรุง
	ข	.00	.00	.00	.00	ไม่ดีต้องปรับปรุง
	ค	.09	.14	.11	.05	ดี
	ง	.32	.27	.29	-.05	ไม่ดีต้องปรับปรุง
	จ	.14	.14	.14	.00	ไม่ดีต้องปรับปรุง

3. การหาค่าอำนาจจำแนกของข้อสอบรายข้อโดยการใช้สูตรสหสัมพันธ์

(correlation technique)

วิธีที่นิยมหากันมากมีอยู่ 2 วิธีคือ

1. Point biserial correlation (r_{pbis})

เป็นการหาค่าสหสัมพันธ์ระหว่างคะแนนที่นักเรียนทำได้ในข้อสอบแต่ละข้อ (0 หรือ 1) กับคะแนนรวม การหาค่าอำนาจจำแนกด้วยวิธีนี้จะทำได้ต่อเมื่อ 1) คะแนนของข้อสอบฉบับนั้นมีการกระจายแบบโค้งปกติ (Normal distribution) และ 2) ข้อสอบแต่ละข้อจะต้องให้คะแนนแบบ “ตอบถูกให้คะแนน 1” “ตอบผิดให้คะแนน 0” เท่านั้น อำนาจจำแนกของข้อสอบที่หาด้วยวิธี Point biserial correlation นี้จะมีค่าสูงสุดเมื่อ $p = q = .50$ แต่ถ้าข้อสอบง่ายเกินไปหรือยากเกินไปค่าอำนาจจำแนกจะมีค่าน้อยไปกว่าความเป็นจริง สูตรการหาค่าอำนาจจำแนกของข้อสอบแต่ละข้อด้วยวิธี Point biserial correlation มีดังนี้

$$r_{pbis} = \frac{M_p - M_q}{s_y} \times \sqrt{pq}$$

- เมื่อ
- M_p = ค่าเฉลี่ยของคะแนนรวมของนักเรียนกลุ่มที่ตอบข้อสอบข้อนั้นถูก
 - M_q = ค่าเฉลี่ยของคะแนนรวมของนักเรียนกลุ่มที่ตอบข้อสอบข้อนั้นผิด
 - s_y = ความเบี่ยงเบนมาตรฐานของคะแนนรวมทั้งหมด
 - p = สัดส่วนของนักเรียนที่ตอบข้อสอบข้อนั้นถูก (= ค่าความยากของข้อสอบ)
 - q = สัดส่วนของนักเรียนที่ตอบข้อสอบข้อนั้นผิดซึ่งมีค่าเท่ากับ $1 - p$

ตาราง 3 ตัวอย่างการหา r_{pbis} ของข้อสอบข้อ 1 (Magnusson, 1967)

คะแนนรวม (X)	จำนวนคนที่ตอบข้อ 1 ผิด	จำนวนคนที่ตอบข้อ 1 ถูก	รวมความถี่
20	-	1	1
19	-	-	-
18	-	1	1
17	-	3	3
16	-	4	4
15	-	6	6
14	-	6	6
13	-	7	7
12	4	4	8
11	4	6	10
10	8	5	13
9	10	3	13
8	9	2	11
7	3	2	5
6	4	-	4
5	5	-	5
4	-	-	-
3	3	-	3
2	-	-	-
1	-	-	-
N	50	50	100
Σ	413	639	1052
\bar{X} (หรือ M)	8.26 = M_q	12.78 = M_p	10.52 = M_y

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

$$M_q = \frac{(12 \times 4) + (11 \times 4) + (10 \times 8) + \dots + (5 \times 5) + (4 \times 0) + (3 \times 3)}{50}$$

$$= \frac{413}{50} = 8.26$$

$$M_p = \frac{(20 \times 1) + (19 \times 0) + (18 \times 1) + (17 \times 3) + \dots + (8 \times 2) + (7 \times 2)}{50}$$

$$= \frac{639}{50} = 12.78$$

$$p = \frac{50}{100} = .50$$

$$q = \frac{50}{100} = .50$$

$$\begin{aligned} \therefore r_{pbis} &= \frac{M_p - M_q}{s_y} \times \sqrt{pq} \\ &= \frac{12.78 - 8.26}{3.49} \times \sqrt{(.5)(.5)} \\ &= 0.648 \end{aligned}$$

ดังนั้น ค่าอำนาจจำแนกของข้อสอบข้อ 1 เท่ากับ 0.648

2. Biserial correlation (r_{bis})

การหาค่าอำนาจจำแนกแบบ Biserial Correlation นี้เป็นการหาความสัมพันธ์ระหว่างผลการสอบของนักเรียนที่ได้จากการตอบข้อสอบ 1 ข้อ กับคะแนนรวมทั้งฉบับ และผลที่ได้จากการตอบข้อสอบแต่ละข้อไม่จำเป็นจะต้องเป็นคะแนน 0 - 1 (นั่นคือ ตอบผิดให้คะแนน 0 ตอบถูกให้คะแนน 1) เสมอไป การให้คะแนนข้อสอบแต่ละข้ออยู่ในระบบ 2 พวก เช่น ผ่าน - ไม่ผ่าน สอบได้ - สอบตก ถูก - ผิด หรือชอบ - ไม่ชอบ เป็นต้น และการวิเคราะห์ข้อสอบแบบนี้จะต้อง

ตั้งอยู่บนข้อตกลงเบื้องต้นว่า ทั้งคะแนนรวมและคะแนนที่ได้จากการตอบข้อสอบ 1 ข้อมีการกระจาย เป็นแบบโค้งปกติ เนื่องจากค่า biserial correlation เป็นอิสระจากค่าความยากง่ายของแบบทดสอบ (นั่นคือ ข้อสอบไม่จำเป็นต้องมีความยากง่ายปานกลาง) ดังนั้น ข้อสอบที่ค่อนข้างยาก เช่น ข้อสอบ ที่ใช้ในการสอบคัดเลือก หรือข้อสอบที่ค่อนข้างง่าย เช่น ข้อสอบ mastery test จึงควรหาค่าอำนาจ จำแนกของข้อสอบด้วยวิธีนี้ การวิเคราะห์ข้อสอบด้วยวิธี biserial correlation นี้มีสูตรคำนวณ ดังนี้

$$r_{bis} = \frac{M_p - M_q}{s_y} \times \frac{pq}{Y}$$

- da M_p = ค่าเฉลี่ยของคะแนนรวมของนักเรียนกลุ่มที่ตอบข้อสอบข้อนั้นถูก
 M_q = ค่าเฉลี่ยของคะแนนรวมของนักเรียนกลุ่มที่ตอบข้อสอบข้อนั้นผิด
 s_y = ความเบี่ยงเบนมาตรฐานของคะแนนรวมทั้งหมด
 P = สัดส่วนของนักเรียนที่ตอบข้อสอบข้อนั้นถูก
 q = สัดส่วนของนักเรียนที่ตอบข้อสอบข้อนั้นผิด ($q = 1 - p$)
 Y = ค่า ordinate ของโค้งปกติตรงจุดที่แบ่ง p และ $1 - p$ (หรือ q) ซึ่งหาได้จากตารางสำเร็จในภาคผนวก

ตัวอย่าง วิธีการหาค่าอำนาจจำแนกแบบ biserial correlation จากตารางที่กล่าวมาแล้วข้างต้น

$$\begin{aligned} M_p &= 12.78 & p &= .5 \\ M_q &= 8.26 & q &= 1-p = .5 \\ S &= 3.49 & Y &= .3989 \text{ (คือจุดที่แบ่งพื้นที่ใต้โค้งเป็นข้างละ .5)} \end{aligned}$$

$$\begin{aligned} r_{bis} &= \frac{M_p - M_q}{s_y} \times \frac{pq}{Y} \\ \dots r_{bis} &= \frac{12.78 - 8.26}{3.49} \times \frac{(.5)(.5)}{.3989} \\ &= .812 \end{aligned}$$

จากตัวอย่างข้างบนนี้แสดงให้เห็นว่าค่า r_{bis} จะมากกว่า r_{pbis} โดยปกติแล้วจะมากกว่ากันอยู่ประมาณ 25% ถึงแม้ว่าข้อสอบข้อนั้นจะมีค่า $p = .50$ ก็ตาม

8.2 การวิเคราะห์ข้อคำถามของแบบทดสอบอิงเกณฑ์

เหตุผลสำคัญประการหนึ่งของการสร้างข้อสอบแบบอิงเกณฑ์คือ ความต้องการที่จะกำหนดเป้าหมายของการสอนให้กระจ่างชัดเจนที่สุด จนกระทั่งผู้เรียนส่วนใหญ่ (หรือทั้งหมด) สัมฤทธิ์ผลในการเรียน สอดคล้องกับพฤติกรรมที่กำหนดไว้ จากการที่ผู้เรียนทุกคนสัมฤทธิ์ผลตามเกณฑ์ที่กำหนดไว้จะเป็นสาเหตุทำให้ผู้เรียนทุกคนสอบได้คะแนนเต็มจึงทำให้ไม่มีความแปรปรวนของคะแนน ดังนั้นถ้าการสอนมีประสิทธิภาพ การหาอำนาจจำแนกของข้อสอบโดยวิธีสหสัมพันธ์ (correlation) จึงไม่สามารถทำได้เนื่องจากความแปรปรวนของคะแนน (S^2) มีค่าเป็นศูนย์

อนึ่งในการประเมินผลแบบอิงเกณฑ์ โดยทฤษฎีแล้วมีหลักอยู่ว่าก่อนที่เด็กนักเรียนจะเริ่มเรียนบทเรียนใดบทเรียนหนึ่ง เด็กจะต้องไม่มีความรู้ในหน่วยการเรียนรู้หรือเนื้อหานั้น ๆ มาก่อนเลย (ถ้านักเรียนมีความรู้ในหน่วยการเรียนนั้นแล้ว) และถ้าการสอนมีประสิทธิภาพแล้ว หลังจากครูทำการสอนจบหน่วยการเรียนรู้ นักเรียนจะต้องมีความรู้ถึงเกณฑ์ที่กำหนดไว้ หรือรู้เรื่องที่ครูสอนไปเป็นอย่างดี จะเห็นได้ว่าการประเมินผลแบบอิงเกณฑ์นั้นมุ่งที่จะให้นักเรียนมีความรู้เท่าเทียมกัน ซึ่งจากจุดประสงค์ที่กล่าวมาแล้วข้างต้น มีผลทำให้คะแนนที่ได้จากการทดสอบหลังจากที่เรียนจบบทเรียนไปแล้ว มีความแปรปรวนน้อยมาก หรือไม่มีการแปรปรวนเลย ($S^2 = 0$) ดังนั้นการวิเคราะห์ข้อสอบของการวัดแบบอิงเกณฑ์จึงต้องใช้วิธีการที่แตกต่างกันออกไปจากการวัดแบบอิงกลุ่ม เนื่องจากการวัดแบบอิงเกณฑ์นั้นยังเป็นเรื่องใหม่ในวงการวัดผล จึงยังไม่สามารถจะหาข้อยุติได้ว่า ควรจะใช้การวิเคราะห์ข้อสอบแบบใดจึงจะดี และมีประสิทธิภาพมากที่สุด ได้มีนักทฤษฎีทางการวัดผลการศึกษาชาวอเมริกันหลายท่าน เสนอวิธีการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ได้หลายวิธี แต่ผู้เขียนจะเสนอวิธีวิเคราะห์ข้อสอบอิงเกณฑ์แบบง่าย ๆ ซึ่งครูสามารถนำไปใช้ได้ 3 วิธีคือ

1. ให้ผู้เชี่ยวชาญวิเคราะห์เนื้อหา
2. ทำ Item - Response Chart
3. หาคความไวของข้อสอบ (S)

1. การวิเคราะห์ข้อสอบอิงเกณฑ์โดยการใช้ผู้เชี่ยวชาญวิเคราะห์เนื้อหา

วิธีนี้ทำได้โดยให้ผู้เชี่ยวชาญทางเนื้อหาวิชาทำการตัดสินคุณภาพของข้อสอบแต่ละข้อ โดยกำหนดให้มีผู้ตัดสินหลาย ๆ คนทำการวิเคราะห์และตรวจสอบดูว่าข้อสอบแต่ละข้อถามวัดเนื้อหาและพฤติกรรมสอดคล้องกับที่กำหนดไว้ในตารางวิเคราะห์หลักสูตรหรือไม่ ถ้าผู้เชี่ยวชาญเห็นว่าข้อสอบข้อใด ถามวัดไม่ตรงกับที่กำหนดไว้ในตารางวิเคราะห์หลักสูตรก็ให้เขียนอธิบายไว้ว่า ทำไมจึงคิดว่าข้อสอบข้อนั้นวัดไม่สอดคล้องกับตารางวิเคราะห์หลักสูตร

ถ้าผู้เชี่ยวชาญตั้งแต่ 2 คนขึ้นไป มีความเห็นว่าข้อสอบข้อใดวัดไม่สอดคล้องกับที่กำหนดไว้ในตารางวิเคราะห์หลักสูตรและเพราะเหตุใด ถ้าเหตุผลของผู้เชี่ยวชาญทั้ง 2 คน ตรงกันหรือใกล้เคียงกันก็แสดงว่าข้อสอบข้อนั้นเป็นข้อสอบที่ไม่ดี ควรจะนำไปปรับปรุงแก้ไขหรือตัดทิ้งไปเลย

2. การวิเคราะห์ข้อสอบอิงเกณฑ์โดยวิธีทำ Item - Response Chart

การวิเคราะห์ข้อสอบด้วยวิธีนี้ยึดหลักการดังที่กล่าวไว้ข้างต้นว่า ก่อนเรียนนักเรียนจะต้องไม่มีความรู้ในเรื่องนั้น และหลังการเรียนบทเรียนไปแล้ว นักเรียนควรจะรู้นั้นคือ เราหวังว่าถ้านักเรียนทำข้อสอบก่อนการสอนข้อสอบควรจะยาก และถ้าเขาทำข้อสอบหลังจากการสอนจบสิ้นแล้วข้อสอบควรจะง่าย ดังนั้นในการวิเคราะห์ข้อสอบด้วย วิธีนี้จึงทำได้โดยการนำข้อสอบ ที่ถามเกี่ยวกับเนื้อหาที่กำลังจะสอนไปทดสอบกับนักเรียนก่อนทำการสอนแล้วนำมาตรวจให้คะแนน และบันทึกผลการสอบของนักเรียนแล้วก็นำแบบทดสอบฉบับเดิมนั้นไปทำการทดสอบกับนักเรียนกลุ่มเดิมอีกที และนำผลการสอบมาตรวจให้คะแนน แล้วบันทึกผลการสอบลงในตารางเดียวกันกับการบันทึกครั้งแรก ดังตัวอย่าง

ตาราง 4 Item - Response Chart ของข้อสอบ 5 ข้อแสดงคำตอบ ถูก (+) และคำตอบผิด (-) ของนักเรียนทุกคนเมื่อทำการทดสอบก่อนการสอน และหลังการสอน

ข้อสอบข้อที่	1		2		3		4		5	
	ก	ล	ก	ล	ก	ล	ก	ล	ก	ล
สมศักดิ์		+	+	+	-	-	+	-	-	+
สิริ		+	+	+	-	-	+	-	+	+
วินิจ	-	+	+	+	-	-	+	-	-	+
สุนีย์	-	+	+	+	-	-	+	-	-	+
จิตรี		+	+	+	-	-	+	-	+	+
วรรณภา		+	+	+	-	-	+	-	-	-

จากตาราง 4 แปลความหมายได้ดังนี้

ข้อสอบข้อ 1 ก่อนสอบไม่มีใครทำข้อสอบถูกเลยแต่หลังจากการสอนแล้วนักเรียนทุกคนทำข้อสอบถูกหมด ข้อสอบข้อนี้จัดได้ว่าเป็นข้อสอบอุดมคติ แต่ก็คงหาได้ยากในทางปฏิบัติ

ข้อสอบข้อ 2 ก่อนสอนนักเรียนทุกคนทำข้อสอบถูกต้องแต่หลังจากสอนแล้วทุกคนก็ทำข้อสอบถูกต้องเหมือนเดิม แสดงว่าข้อสอบข้อนี้ง่ายมาก ที่จริงแล้วถ้าก่อนสอนเด็กทำข้อสอบถูกต้องแล้ว ครูก็ไม่ควรสอนจุดประสงค์ข้อนั้น เพราะนักเรียนควรรู้หมดแล้ว ทำให้ไม่เกิดผลอะไรขึ้นมา เป็นการเสียเวลาโดยใช้เหตุ ข้อสอบข้อนี้จึงไม่ควรใช้

ข้อสอบข้อ 3 ทั้งก่อนและหลังการสอนไม่มีใครทำข้อสอบนี้ถูกเลย แสดงว่าข้อสอบยากมาก หรืออาจจะเป็นเพราะว่าจุดประสงค์ในการสอนและการสอบไม่เหมือนกันจะต้องแก้ไขจุดประสงค์เสียใหม่

ข้อสอบข้อ 4 ข้อสอบข้อนี้ก่อนสอนนักเรียนทำถูกต้อง แต่หลังจากสอนแล้วกลับทำไม่ได้เลย แสดงว่าข้อสอบอาจจะกำกวม เด็กอ่านโจทย์แล้วไม่เข้าใจตรงกันหรืออาจจะเป็นการสอนไม่ถูกต้อง เด็กอาจจะเกิดความเข้าใจผิดในเนื้อหาที่เรียน ทั้ง ๆ ที่ก่อนสอนเด็กเข้าใจถูกต้องอยู่แล้ว ข้อสอบข้อนี้จึงไม่ควรใช้

ข้อสอบข้อ 5 ข้อสอบข้อนี้ก่อนสอนมีนักเรียนบางคนทำถูก และบางคนที่ทำผิด แต่หลังจากสอนแล้วมีนักเรียนที่ทำข้อสอบผิดกลับทำถูก และคนที่ทำถูกอยู่ก็ยังคงทำถูกเหมือนเดิม ข้อสอบแบบนี้จัดว่าเป็นข้อสอบที่มีประสิทธิภาพที่เราสามารถหาได้ในทางปฏิบัติ

3. การวิเคราะห์ข้อสอบอิงเกณฑ์โดยการหาความไวของข้อสอบ

แบ่งการวิเคราะห์เป็น 2 อย่าง คือ

ก. การวิเคราะห์ข้อสอบเป็นรายข้อ

ข. การวิเคราะห์ข้อสอบรายตัวเลือก

ก. การวิเคราะห์ข้อสอบเป็นรายข้อ

ในการวิเคราะห์ข้อสอบเป็นรายข้อมีขั้นตอนในการดำเนินงาน ดังนี้

1. นำแบบทดสอบที่ต้องการจะตรวจสอบคุณภาพไปทดสอบกับนักเรียนที่จะเรียนวิชาของท่านก่อนทำการสอนแล้วนำมาตรวจให้คะแนน

2. นำกระดาษคำตอบทั้งหมดของนักเรียนมาขีดรอยคะแนนลงในตารางเพื่อสำรวจดูว่าข้อไหนมีคนทำถูกกี่คน แล้วหาค่าความยากง่ายของข้อสอบแต่ละข้อจากสูตร

$$p = \frac{\text{จำนวนคนตอบถูกในข้อสอบข้อใด ๆ}}{\text{จำนวนนักเรียนทั้งหมด}}$$

ค่า p ที่คำนวณได้จากการสอบครั้งแรกเรียกว่า P_{pre} (ดูตัวอย่างตาราง 5)

3. หลังจากที่ครูสอนจบบทเรียนแล้ว นำข้อสอบฉบับเดียวกันกับที่กล่าวไว้ในข้อ 1 ไปทดสอบกับนักเรียนกลุ่มเดิมอีก แล้วนำมาตรวจหาค่า p ด้วยวิธีการที่กล่าวมาแล้วในข้อ 2 p ที่ได้ในข้อ 3 นี้มีชื่อเรียกว่า P_{post} (ดูตัวอย่างตาราง 6)

4. หาค่าความไวของข้อสอบจากสูตร

$$S = \frac{P_{pre} - P_{post}}{P_{pre}}$$

เมื่อ S = ค่าความไวของข้อสอบ (Sensitive Index)

S ควรมีค่าตั้งแต่ 0.30 ขึ้นไป จึงจะถือว่าเป็นข้อสอบที่ดี (ดูตัวอย่างการหา S จากตาราง 7)

ตาราง 5 ตารางคำนวณค่า P_{pre} ของข้อสอบวิชาเลขคณิต (30 ข้อ) ชั้น a.3 จำนวนนักเรียนที่สอบ 50 คน (สอบก่อนการสอน)

คนที่ ไอ้ที่	1	2	3	4	5	6	18	19	50	รวม	ค่า P_{pre}
1											-	$0/50 = .00$
2	/	/	/	/	/	/		/	/	/	50	$50/50 = 1.00$
3		/		/		/					25	$25/50 = .50$
4												$0/50 = .00$
5												
6												
7												
27												
26												
29												
30												

การแปลความหมายตาราง 5

- ข้อสอบข้อ 1 ไม่มีนักเรียนตอบถูกเลย ค่า $P_{pre} = .00$
- ข้อสอบข้อ 2 นักเรียนทุกคนตอบถูก ค่า $P_{pre} = 1.00$
- ข้อสอบข้อ 3 นักเรียนตอบถูกเพียง 25 คน ค่า $P_{pre} = .50$
- ข้อสอบข้อ 4 ไม่มีนักเรียนตอบถูกเลย ค่า $P_{pre} = .00$

ตาราง 6 ตารางคำนวณค่า P_{post} ของข้อสอบวิชาเลขคณิต (30 ข้อ) ชั้น อ.3 จำนวนนักเรียนที่สอบ 50 คน (สอบก่อนการสอน)

คนที่ ข้อที่	1	2	3	4	5	6	18	19	50	รวม	ค่า P_{post}
1	/	/	/	/	/	/		/	/	1	50	$0/50 = 1.00$
2											-	$0/50 = .00$
3	/	/	/	/	/	/		/	/	/	50	$0/50 = 1.00$
4											-	$0/50 = .00$
5												
6												
7												
27												
28												
29												
30												

การแปลความหมายตาราง 5

- ข้อสอบข้อ 1 ไม่มีนักเรียนตอบถูกเลย ค่า $P_{post} = .00$
- ข้อสอบข้อ 2 นักเรียนทุกคนตอบถูก ค่า $P_{post} = 1.00$
- ข้อสอบข้อ 3 นักเรียนตอบถูกเพียง 25 คน ค่า $P_{post} = .50$
- ข้อสอบข้อ 4 ไม่มีนักเรียนตอบถูกเลย ค่า $P_{post} = .00$

ตาราง 7 ตารางคำนวณค่า S (ความไวของข้อสอบ) ของวิชาเลขคณิต ชั้น อ.3

ข้อสอบ ข้อที่	P	P _{pre}	P _{post} - P _{pre}	S	ความหมาย
1	1.00	.00	1.00 - .00	1.00	ดี เป็นข้อสอบอุดมคติ
2	.00	1.00	.00 - 1.00	-1.00	ไม่ดี ต้องตัดทิ้ง
3	1.00	.50	1.00 - .50	.50	พอใช้ได้ แต่ต้องเปลี่ยนแปลงจุด ประสงค์ให้ซับซ้อนขึ้นกว่าเดิม
4	.00	.00	.00 - .00	.00	ไม่ดี ต้องตัดทิ้ง
5	.60	.05	.60 - .05	.55	ไม่ดี ต้องตัดทิ้ง
6	.90	.10	.90 - .10	.80	ดี เก็บไว้ใช้ต่อไป
7	1.00	1.00	1.00 - 1.00	.00	ไม่ดี ต้องตัดทิ้ง
29					
30					

การแปลความหมายค่า S จากตาราง

ข้อสอบข้อ 1 S มีค่าเท่ากับ 1.00 จัดได้ว่าเป็นข้อสอบที่ดี เป็นไปตามอุดมคติจริง ๆ เนื่องจากตอนสอบครั้งแรกไม่มีนักเรียนตอบถูกเลย ($p = .00$) แต่หลังจากผ่านการเรียนไปแล้วนักเรียนทุกคนทำข้อสอบถูก ($p = 1.00$) จึงเป็นข้อสอบที่มีความไวในการวัดผลมากที่สุด ข้อสอบข้อนี้ต้องเก็บเอาไว้ใช้อีก

ข้อสอบข้อ 2 S มีค่าเท่ากับ -1.00 จัดได้ว่าเป็นข้อสอบที่ไม่ดีต้องตัดทิ้งไป ทั้งนี้เนื่องจากตอนสอบครั้งแรกนักเรียนทุกคนตอบข้อสอบถูกหมดทุกข้อ แสดงว่ามีความรู้ในเรื่องนั้นเป็นอย่างดี แต่พอนักเรียนเรียนเนื้อหาผ่านไป แล้ว มาทำข้อสอบอีกกลับทำข้อสอบผิดหมดทุกคน แสดงว่าคำถามอาจจะกำกวม หรืออาจจะเป็นเพราะการสอนไม่ถูกต้อง เด็กอาจจะเกิดความเข้าใจผิดในเนื้อหาที่เรียน ทั้ง ๆ ที่ก่อนสอนเด็กเข้าใจถูกต้องอยู่แล้ว

ข้อสอบข้อ 3 ก่อนเรียนนักเรียนครึ่งหนึ่งรู้เนื้อหานั้นอยู่แล้ว ($p = .50$) แต่หลังจากเรียนไปแล้วนักเรียนทุกคนทำข้อสอบถูกหมด ($p = 1.00$) จึงเป็นข้อสอบที่พอใช้ได้ แต่อาจจะเป็นข้อสอบที่ต้องปรับปรุงจุดประสงค์เชิงพฤติกรรมที่จะวัดให้ซับซ้อนขึ้นได้โดยที่ยังถามเนื้อหาเดิมอยู่

ข้อสอบข้อ 4 และข้อ 7 เป็นข้อสอบที่ไม่ดีทั้งคู่เพราะ S มีค่าเท่ากับ .00

ข้อสอบข้อ 5 เป็นข้อสอบที่ไม่ดีต้องตัดทิ้งไปเนื่องจากหลังจากทำการสอนแล้วนักเรียนเพียง 60% ที่ตอบข้อสอบข้อนี้ถูก ($p = .60$) อย่างไรก็ตามข้อสอบข้อนี้เป็นข้อสอบที่เคยวิเคราะห์และทดลองใช้กับเด็กกลุ่มอื่นมาแล้ว ปรากฏว่าเป็นข้อสอบที่มีความไวในการทดสอบสูง (S มีค่าใกล้ 1.00) แต่พอมาใช้กับเด็กกลุ่มนี้ S กลับมีค่าลดน้อยลง แสดงว่านักเรียนบางคนไม่เกิดการเรียนรู้ในบทเรียนนั้น ๆ ครูจะต้องจัดการสอนซ่อมเสริมให้กับนักเรียนกลุ่มนี้ใหม่เป็นพิเศษ เพื่อให้เขาเกิดการเรียนรู้ขึ้นมาไม่ใช่ตัดข้อสอบทิ้งไป

ข้อสอบข้อ 6 S มีค่าเท่ากับ .80 จัดเป็นข้อสอบที่ดี เพราะก่อนสอนมีนักเรียนเพียง 10% ที่ตอบข้อสอบข้อนี้ถูก และหลังจากการสอนแล้วมีนักเรียนถึง 90% ($P = .90$) ตอบข้อสอบข้อนี้ถูก

กล่าวโดยสรุปแล้วในการพิจารณาคัดเลือกข้อสอบ โดยพิจารณาผลจากการวิเคราะห์ข้อสอบด้วยวิธีนี้การหาค่าความไวของข้อสอบ (S) นั้น ควรถือหลักดังนี้

1. พิจารณาจากค่าความยากง่ายของข้อสอบก่อนการสอน (P_{pre})

ค่า P_{pre}	ความหมาย
.41 ขึ้นไป	เป็นข้อสอบที่ง่ายเกินไป ควรตัดทิ้ง
.21 - .40	เป็นข้อสอบที่พอใช้ได้ แต่ต้องปรับปรุงจุดประสงค์เชิงพฤติกรรมให้ซับซ้อนขึ้น
ต่ำกว่า .21	เป็นข้อสอบที่ดี

2. พิจารณาจากค่าความยากง่ายของข้อสอบหลังการสอน (P_{post})

ค่า P_{post}	ความหมาย
.80 ขึ้นไป	เป็นข้อสอบที่ดี
.70 - .79	เป็นข้อสอบที่พอใช้ได้ แต่ครูอาจจะต้องปรับปรุงตัวคำถามหรือตัวเลือกบ้างเล็กน้อย
ต่ำกว่า .69	เป็นข้อสอบที่ไม่ดี ตัดทิ้งไป

3. พิจารณาจากค่าความไวของข้อสอบ (S)

ค่า S	ความหมาย
1.00	เป็นข้อสอบที่ดี เป็นไปตามทฤษฎี
.80 - .99	เป็นข้อสอบที่ดี หาได้ในเชิงปฏิบัติ
.30 - .79	เป็นข้อสอบที่พอใช้ได้
.00 - .29	เป็นข้อสอบที่ไม่ดี ควรตัดทิ้ง
-1.00 ถึง .00	เป็นข้อสอบที่ไม่ดีอย่างยิ่ง ต้องตัดทิ้ง

ข. การวิเคราะห์รายตัวเลือก

การวิเคราะห์รายตัวเลือกมีประโยชน์สำหรับการปรับปรุงข้อสอบ การวิเคราะห์นี้ถือหลักว่า ถ้านักเรียนเรียนรูบทเรียนนั้น ๆ แล้วเมื่อนำข้อสอบไปทดสอบหลังจากที่ครูสอนไปแล้วไม่ควรจะมีนักเรียนคนใดเลือกตอบตัวลงเลย ในการหาค่าของตัวเลือกที่เป็นตัวลง ต้องใช้สูตรที่แปลกไปจากของตัวเลือกที่เป็นตัวถูก ดังนี้

$$S \text{ ของตัวถูก} = P_{\text{post}} - P_{\text{pre}}$$

$$S \text{ ของตัวลง} = P_{\text{pre}} - P_{\text{post}}$$

ขั้นตอนในการวิเคราะห์รายตัวเลือกมีดังนี้

1. นำผลของการทดสอบก่อนทำการสอนทั้งหมดมาขีดรอยคะแนน (tally) เพื่อสำรวจดูว่าข้อไหนมีใครเลือกตัวเลือกใดบ้าง ดูตัวอย่างตาราง tally ตารางที่ 1 หรือ 2 แล้วคำนวณหาค่า P_{pre} ของแต่ละตัวเลือกจากสูตร

$$P_{\text{pre}} = \frac{\text{จำนวนคนที่เลือกตอบในแต่ละตัวเลือก}}{\text{จำนวนคนทั้งหมด}}$$

2. นำผลของการทดสอบหลังจากที่ทำการสอนไปแล้วทั้งหมดมาขีดรอยคะแนน (tally) เพื่อสำรวจดูว่าข้อไหนมีใครเลือกตัวเลือกใดบ้าง แล้วคำนวณหาค่า P_{post} ของแต่ละตัวเลือกจากสูตร

$$P_{\text{post}} = \frac{\text{จำนวนคนที่เลือกตอบในแต่ละตัวเลือก}}{\text{จำนวนคนทั้งหมด}}$$

3. นำผลที่ได้จากข้อ 1 และข้อ 2 มาเขียนตอบลงในตาราง 8 เพื่อคำนวณ

หาค่า S

ตาราง 8 ตารางวิเคราะห์ข้อสอบรายตัวเลือก

ข้อ	ตัวเลือก	ผลการสอบครั้งแรก		ผลการสอบหลังจากสอนไปแล้ว		S
		จำนวนคนที่ตอบ	P _{pre}	จำนวนคนที่ตอบ	P _{post}	
1	ก					
	ข					
	ค					
	ง					
	จ					
2	ก					
	ข					
	ค					
	ง					
	จ					

การพิจารณาตัวเลือก

1. ตัวเลือกที่เป็นตัวถูก ใช้หลักการพิจารณาแบบเดียวกับหลักการพิจารณาเลือกข้อสอบเมื่อวิเคราะห์แบบรายข้อ นั่นคือ S ควรมียค่าตั้งแต่ .30 ขึ้นไป

2. สำหรับตัวเลือกที่เป็นตัวลวง มีหลักในการพิจารณาดังนี้

- ตัวลวงใดที่นักเรียนไม่เลือกตอบเลยในการสอบหลังจากที่สอนไปแล้ว ถือว่าเป็นตัวลวงที่ใช้ได้

- ตัวลวงใดมีค่า S ตีตก แสดงว่าเป็นตัวลวงที่ไม่ดี ต้องปรับปรุงแก้ไข

- ตัวลวงใดที่มีค่า S เป็นบวก ถือว่าเป็นตัวลวงที่ใช้ได้

ในการวิเคราะห์ข้อสอบแบบอิงเกณฑ์นั้น ถ้าจะนำผลจากการวิเคราะห์มาปรับปรุงการสอนและการสอบให้มีประสิทธิภาพแล้ว ควรจะดำเนินการวิเคราะห์ทั้ง 3 วิธี ที่กล่าวมาแล้วข้างต้น อย่างไรก็ตามถ้าครูไม่มีเวลาเพียงพอ อาจจะเลือกเอาวิธีใดวิธีหนึ่งมาใช้ก็ได้

8.3 การวิเคราะห์ข้อคำถามของแบบสอบถาม

การตรวจสอบคุณภาพข้อคำถามรายชื่อเป็นวิธีการตรวจสอบอำนาจในการจำแนกของข้อคำถามของแบบสอบถามโดยใช้การทดสอบแบบที (t-test) ทดสอบความแตกต่างของผู้ตอบที่ให้คะแนนการประเมินสูง กับผู้ตอบที่ให้คะแนนการประเมินต่ำ ถ้าคนสองกลุ่มนี้ตอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติก็แสดงว่าข้อคำถามนั้นมีอำนาจในการจำแนก (discrimination power) ถือว่าเป็นข้อคำถามที่มีความไวพอที่จะใช้วัดความคิดเห็นเกี่ยวกับประสิทธิภาพในการสอนได้

วิธีวิเคราะห์หาอำนาจจำแนกของข้อคำถามมีขั้นตอน ดังนี้

1) ให้คะแนนตามค่าน้ำหนักในแต่ละข้อคำถาม แล้วหาคะแนนรวมของผู้ตอบแบบสอบถามแต่ละคน

2) เรียงคะแนนของแต่ละคนได้จากสูงไปหาต่ำ หรือต่ำไปหาสูง

3) หากจำนวน 25% (หรือ 27% หรือ 33%) ของผู้ตอบทั้งหมดในตัวอย่างนี้ จะใช้เฉพาะ 25% ของคนกลุ่มที่ได้คะแนนสูง และ 25% ของคนกลุ่มที่ได้คะแนนต่ำมาทำการวิเคราะห์ที่เหลือนตรงกลางไม่นำมาคิด

4) นำแบบสอบถามของกลุ่มสูง 25% และกลุ่มต่ำ 25% มา tally เป็นรายชื่อ

5) หาผลรวมของคะแนนยกกำลังสอง ($\sum fx^2$) ของกลุ่มสูงและกลุ่มต่ำ

6) คำนวณหาค่าเฉลี่ยของคำตอบของกลุ่มสูง (\bar{X}_H) ของคำถามแต่ละข้อ และค่าเฉลี่ยของคำตอบของกลุ่มต่ำ (\bar{X}_L) ของคำถามแต่ละข้อ โดยใช้สูตร

$$\bar{X}_H = \frac{\sum fx_H}{n_H} \quad \text{และ} \quad \bar{X}_L = \frac{\sum fx_L}{n_L}$$

7) คำนวณหาความแปรปรวนของคำตอบแต่ละข้อของคะแนนกลุ่มสูง S_H^2 จากสูตร

$$S_H^2 = \left[\sum fx^2 - \frac{(\sum fx)^2}{n_H} \right] / n_H - 1$$

8) คำนวณหาความแปรปรวนของคำตอบแต่ละข้อของคะแนนกลุ่มต่ำ S_L^2 จากสูตร

$$S_L^2 = \left[\sum fx^2 - \frac{(\sum fx)^2}{n_L} \right] / n_L - 1$$

เมื่อ n_H = จำนวน 25% ของกลุ่มสูง

n_L = จำนวน 25% ของกลุ่มต่ำ = n_H

9) ทำการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของกลุ่มสูง และกลุ่มต่ำของคำถามแต่ละข้อ จากสูตร

$$t = \frac{\bar{X}_H - \bar{X}_L}{\sqrt{\frac{S_H^2}{n_H} + \frac{S_L^2}{n_L}}}, \quad df = n_H + n_L - 2$$

10) กำหนดระดับความเชื่อมั่นว่า จะใช้ระดับความเชื่อมั่น 95% ($\alpha = 0.05$) หรือ 99% ($\alpha = 0.01$)

11) เปิดตารางหาค่าที่ใช้การทดสอบแบบทางเดียว (One-tailed test)

12) ถ้าค่าที่คำนวณได้ในข้อ 9 มีค่ามากกว่าค่าที่ได้จากการเปิดตารางในข้อ 11 ก็แสดงว่าข้อคำถามข้อนั้นมีอำนาจในการจำแนก ถือว่าเป็นข้อคำถามที่มีความไวพอที่จะใช้วัดความคิดเห็นเกี่ยวกับประสิทธิภาพในการสอนได้ แต่ถ้าค่าที่คำนวณได้ในข้อ 9 มีค่าน้อยกว่าค่าที่ในข้อ 11 แสดงว่าข้อคำถามนั้นมีอำนาจในการจำแนกต่ำ ไม่ควรนำมาใช้

ตัวอย่าง นำแบบสอบถามเกี่ยวกับ “ประสิทธิภาพด้านการบริการและประชาสัมพันธ์ของห้องสมุดโรงเรียนอนุบาล” ไปทดลองกับกลุ่มตัวอย่างจำนวน 84 คน แล้วนำผลมาวิเคราะห์คุณภาพของข้อคำถามแต่ละข้อ อยากทราบว่าคำถามแต่ละข้อมีอำนาจจำแนกหรือไม่

วิธีการ

1. หา 25% ของผู้เข้าตอบ 84 คนมีค่าเท่ากับ 21 คน
2. หาคะแนนรวมของผู้ตอบแบบสอบถามแต่ละคน
3. เรียงคะแนนรวมจากมากไปหาน้อย
4. นำแบบสอบถามของกลุ่มสูง 21 คน (นับจากคะแนนรวมมากที่สุดลงมา) กับกลุ่มต่ำ 21 คน (นับจากคะแนนรวมต่ำสุดขึ้นไป) มา tally เป็นรายข้อ ดังนี้

คำถามข้อที่ 1

ระดับ ความคิดเห็น	กลุ่มสูง					ระดับ ความคิดเห็น	กลุ่มต่ำ				
	x	f	fx	x ²	fx ²		x	f	fx	x ²	fx ²
มากที่สุด	5	5	25	25	125	มากที่สุด	5	1	5	25	25
มาก	4	13	52	16	208	มาก	4	6	24	16	96
ปานกลาง	3	3	9	9	27	ปานกลาง	3	13	39	9	117
น้อย	2	0	0	4	0	น้อย	2	1	2	4	4
น้อยที่สุด	1	0	0	1	0	น้อยที่สุด	1	0	0	1	0

1. คำนวณหาค่าเฉลี่ยของกลุ่มสูงและกลุ่มต่ำ

$$\bar{X}_H = \frac{86}{21} = 4.09 \quad \bar{X}_L = \frac{70}{21} = 3.34$$

2. คำนวณหาค่าความแปรปรวนของกลุ่มสูงและกลุ่มต่ำ

$$S_H^2 = \left[\sum fx^2 - \frac{(\sum fx)^2}{n_H} \right] / n_H - 1$$

$$= \left[360 - \frac{(86)^2}{21} \right] / 21 - 1$$

$$= 0.39$$

$$S_L^2 = \left[\sum fx^2 - \frac{(\sum fx)^2}{n_L} \right] / n_L - 1$$

$$= \left[224 - \frac{(70)^2}{21} \right] / 21 - 1$$

$$= 0.43$$

3. ทดสอบความแตกต่างระหว่างกลุ่มสูงและกลุ่มต่ำ

$$t = \frac{\bar{X}_H - \bar{X}_L}{\sqrt{\frac{S_H^2}{n_H} + \frac{S_L^2}{n_L}}}, \quad df = n_H + n_L - 2$$
$$= \frac{4.09 - 3.34}{\sqrt{\frac{0.39}{21} + \frac{0.43}{21}}}$$
$$= 3.75$$

4. เปิดค่าตาราง t ที่ระดับความเชื่อมั่น 0.05, $df = 40$ ได้ค่า $t = 1.68$

5. เนื่องจากค่า t ที่คำนวณได้ มีค่ามากกว่าค่า t ที่เปิดตาราง ฉะนั้นข้อคำถามนี้มีอำนาจจำแนก นำมาใช้ได้

คำถามท้ายบทที่ 8

1. การวิเคราะห์ข้อสอบมีประโยชน์อย่างไร จงอธิบาย
2. การวิเคราะห์ข้อสอบแบบอิงกลุ่มยึดหลักการใดที่แตกต่างจากการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ จงอธิบาย
3. ท่านจะตรวจสอบคุณภาพของข้อคำถามของแบบสอบถามได้โดยวิธีใด จงอธิบาย
4. จงอธิบายความหมายของคำต่อไปนี้
 - 1) ความยากง่ายของข้อสอบ
 - 2) อำนาจจำแนกของข้อสอบ
 - 3) ความไวของข้อสอบ