

บทที่ 9

การวิเคราะห์ข้อสอบ (Item Analysis)

การวิเคราะห์ข้อสอบคืออะไร

การวิเคราะห์ข้อสอบเป็นขบวนการตรวจสอบคุณภาพของข้อคำถามแต่ละข้อว่ามีคุณลักษณะตรงตามวัตถุประสงค์ที่ผู้ออกข้อสอบต้องการข้อละเท่าใด จุดมุ่งหมายใหญ่ของการวิเคราะห์ข้อสอบก็เพื่อจะทราบคุณภาพของคำถามอยู่ 2 ประการ คือ ข้อสอบข้อนั้น (หรือตัวเลือกนั้น) มีระดับความยากเท่าใด และมีอำนาจจำแนกเท่าใด การวิเคราะห์ข้อสอบนี้จะกระทำได้ก็ต่อเมื่อได้นำเอาข้อสอบนั้นไปทดสอบกับนักเรียนมาแล้วเท่านั้น เพราะจะต้องใช้ผลจากการที่นักเรียนตอบข้อสอบมาเป็นหลักในการวิเคราะห์

ระดับความยาก (Level of difficulty : p)

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

$$p = \frac{\text{จำนวนคนที่เลือกตอบตัวเลือกนั้น}}{\text{จำนวนคนตอบทั้งหมด}}$$

ถ้าตัวเลือกที่นำมาวิเคราะห์เป็นตัวเลือก ค่า p ก็จะมีหมายถึงเปอร์เซ็นต์หรือสัดส่วนที่ผู้เข้าสอบกลุ่มนั้นทำข้อสอบถูก ดังนั้น

$$p = \frac{\text{จำนวนคนที่ตอบข้อสอบข้อนั้นถูก}}{\text{จำนวนคำตอบทั้งหมด}}$$

ลักษณะความยากของข้อสอบแต่ละข้อ หรือแต่ละตัวเลือก มีดังนี้

1. p (ระดับความยาก) มีค่าตั้งแต่ .00 ถึง 1.00
2. ถ้านักเรียนเลือกตอบตัวเลือกใดมาก แสดงว่าข้อหรือตัวเลือกนั้นง่ายค่าของ p จะสูง
3. ถ้านักเรียนเลือกตอบตัวเลือกได้น้อย แสดงว่าข้อหรือตัวเลือกนั้นยากค่าของ p จะต่ำ
4. ถ้านักเรียนทำข้อสอบข้อใดถูก 100 % (ไม่ผิดเลย) แสดงว่าข้อสอบข้อนั้นง่ายมาก (p = 1.00)
5. ถ้าข้อสอบข้อใดไม่มีคนตอบถูกเลย หมายความว่าข้อสอบข้อนั้นยากมาก (p = .00)

6. ข้อสอบที่มีค่า $p = .50$ ถือว่าข้อสอบข้อนั้นยากพอเหมาะ หมายความว่า เป็นข้อสอบที่มี 50% หรือครึ่งหนึ่งของผู้เข้าสอบตอบถูก อย่างไรก็ตาม ถ้าข้อสอบข้อ 1 และข้อ 5 ต่างก็มีค่า $p = .50$ ไม่ได้หมายความว่า ผู้ที่ตอบข้อ 1 ถูก จะต้องตอบข้อ 5 ถูกด้วย ผู้ที่ตอบข้อ 5 ถูก อาจจะเป็นคนอื่น แต่จะมีจำนวนเท่ากับครึ่งหนึ่งของผู้เข้าสอบเช่นกัน

7. การแปลความหมายระดับความยากของข้อสอบ

- $p .81 - 1.00$ เป็นข้อสอบที่ง่ายมาก
- $p .61 - .80$ เป็นข้อสอบที่ค่อนข้างง่าย
- $p .41 - .50$ เป็นข้อสอบที่ง่ายพอเหมาะ
- $p .21 - .40$ เป็นข้อสอบที่ค่อนข้างยาก
- $p .00 - .20$ เป็นข้อสอบที่ยากมาก

อำนาจจำแนก (Discrimination index : r)

อำนาจจำแนก (ใช้แทนด้วยอักษร r) ของข้อคำถามใด หมายถึงประสิทธิภาพของข้อคำถามในการแบ่งนักเรียนออกเป็นสองกลุ่ม คือกลุ่มเก่งที่มีความรู้ดี กับกลุ่มอ่อนที่มีความรู้ต่ำ ค่า r จะหาได้จากการหาค่าสหสัมพันธ์ (correlation) ระหว่างคะแนนที่ได้จากข้อสอบนั้นกับคะแนนรวมทั้งหมด ซึ่งวิธีหาค่าสหสัมพันธ์และเรียกชื่อต่างกันหลายวิธี เช่น แบบ Point - biserial correlation (เขียนย่อ ๆ ว่า r_{pbis}) หรือแบบ Biserial correlation (เขียนย่อ ๆ ว่า r_{bis}) หรือ Tetrachoric correlation (เขียนย่อ ๆ ว่า r_{tet}) หรือ Phi coefficient (r_{phi}) เนื่องจากการหาค่าอำนาจจำแนกของข้อสอบทั้ง 4 วิธีนี้ มีข้อตกลงเบื้องต้นและวิธีการหาค่าที่แตกต่างกัน และผู้ที่ทำการวิเคราะห์ข้อสอบจะต้องใช้ความรู้ทางสถิติเข้ามาช่วยจึงจะหาค่าออกมาได้ ผู้เขียนจึงจะไม่แสดงวิธีการหาค่า r ทั้ง 4 แบบนี้ไว้ในหนังสือเล่มนี้แต่จะเสนอวิธีการหาค่า r แบบง่าย ๆ เพื่อให้ครูทั่วไปสามารถทำการวิเคราะห์ข้อสอบได้ด้วยตนเอง ส่วนผู้ที่มีความประสงค์อย่างแรงกล้าที่จะศึกษาวิธีการหาค่าอำนาจจำแนกด้วยวิธีทั้ง 4 แบบที่กล่าวมาแล้วข้างต้น ก็อาจจะหาอ่านได้ในหนังสือ "การสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน" ซึ่งผู้เขียนได้เขียนไว้สำหรับใช้เป็นตำราเรียนในวิชา MR 311 (วิธีการสร้างแบบทดสอบ 1)

การแปลความหมายค่าอำนาจจำแนก

1. ค่าอำนาจจำแนกหรือค่า r มีค่าตั้งแต่ -1.00 ถึง $+1.00$
2. ข้อสอบข้อใดที่นักเรียนกลุ่มเก่งทำถูกหมดทุกคน และนักเรียนกลุ่มอ่อนทำผิดหมดทุกคน r จะมีค่าเท่ากับ $+1.00$ ซึ่งจัดเป็นข้อสอบที่ดีเยี่ยม เป็นที่ต้องการอย่างยิ่ง

3. ข้อสอบข้อใดที่นักเรียนกลุ่มเก่งทำผิดหมด แต่นักเรียนกลุ่มอ่อนกลับทำถูกหมดทุกคน ค่า r จะเท่ากับ -1.00 ข้อสอบข้อนี้จะเป็นข้อสอบที่ไม่ดี เพราะเป็นข้อสอบที่ลวงนักเรียนเก่ง ซึ่งผิดจุดประสงค์ เพราะในการทดสอบใด ๆ นั้น เราหวังว่านักเรียนเก่งจะทำข้อสอบได้ นักเรียนอ่อนจะทำข้อสอบไม่ได้

4. ข้อสอบข้อใดที่นักเรียนกลุ่มเก่งกับกลุ่มอ่อนทำถูกได้เท่า ๆ กัน จะมีค่า r เท่ากับ $.00$ แสดงว่าข้อสอบข้อนี้เป็นข้อสอบที่ไม่สามารถจะแยกนักเรียนเก่งออกจากนักเรียนอ่อนได้ นั่นคือ เราไม่สามารถจะบอกได้เลยว่าผู้ที่ทำข้อสอบข้อนี้ถูกเป็นเด็กเก่ง มีความรู้ในเรื่องที่ตามจริง ๆ หรือตอบถูกเพราะการเดา

Ebel (1972) ได้ให้ข้อเสนอแนะเกณฑ์ที่ใช้พิจารณาคุณค่าอำนาจจำแนกของข้อสอบ ดังนี้

ดัชนีอำนาจจำแนก	คุณภาพของข้อสอบ
.40 ขึ้นไป	ดีมาก
.30 ถึง .39	ดีพอสมควร อาจต้องปรับปรุงบ้าง
.20 ถึง .29	พอใช้ได้ แต่ต้องปรับปรุง
ต่ำกว่า .20	ใช้ไม่ได้ ต้องพิจารณาปรับปรุงใหม่หรือตัดทิ้งไปเลย

ประโยชน์ของการวิเคราะห์ข้อสอบ

1. ทำให้รู้คุณภาพของข้อสอบ ซึ่งจะช่วยให้ครูสามารถปรับปรุงข้อสอบได้ถูกต้อง ผลจากการวิเคราะห์ข้อสอบจะบอกได้ว่า จุดอ่อนอยู่ตรงไหน ภาษาที่ใช้กำกวมหรือไม่

2. ช่วยเป็นแนวทางให้ครูเขียนข้อสอบเพิ่มขึ้น ผลจากการวิเคราะห์ข้อสอบจะแสดงให้เห็นอย่างชัดเจนว่า องค์ประกอบของข้อสอบส่วนไหนน่าจะเป็นตัวทำให้ข้อสอบมีคุณภาพดี หรือด้วยตัวเลขที่ได้จากการวิเคราะห์จะช่วยทำให้ครูระมัดระวังเรื่องภาษาที่ใช้ในการเขียนตัวเลือกทั้งที่เป็นตัวถูกและตัวลวง

3. ช่วยทันแรงในการออกข้อสอบระยะยาว โดยการเก็บ คัดเลือกข้อสอบที่วิเคราะห์แล้วไว้ใช้ต่อไป

4. เป็นจุดเริ่มต้นของการสร้างข้อสอบคู่ขนาน (parallel test)

5. เป็นจุดเริ่มต้นของการสร้างธนาคารข้อสอบ (Item Bank)

6. เป็นจุดเริ่มต้นของการสร้างข้อสอบมาตรฐาน เพื่อใช้ในกลุ่มโรงเรียนที่ใกล้เคียงกัน หรืออาจจะปรับปรุงเป็นข้อสอบมาตรฐานระดับท้องถิ่น หรือระดับภาคการศึกษาได้

หลักการวิเคราะห์ข้อสอบ

โดยทฤษฎีแล้ว ในการวิเคราะห์ข้อสอบนั้น เมื่อครูสร้างข้อสอบและนำไปทดสอบกับนักเรียนแล้ว ครูจะต้องนำเอาผลการสอบนั้นมาตรวจให้คะแนน แล้วก็มาโยงนับเป็นรายชื่อว่า ข้อใด มีนักเรียนทำถูกกี่คน คิดเป็นกี่เปอร์เซ็นต์ จากนั้นก็ใช้วิธีการทางสถิติคำนวณหาค่าต่าง ๆ ที่เราต้องการออกมา อย่างไรก็ตามวิธีการวิเคราะห์แบบนี้นอกจากครูจะต้องมีความรู้ความชำนาญในวิชาสถิติแล้ว ยังเป็นการเสียเวลามาก จึงมีนักวิจัยคิดค้นวิธีการวิเคราะห์ข้อสอบแบบง่าย ๆ ขึ้นมา เรียกว่า "เทคนิค 27%" คือเราไม่ต้องเอาคำตอบทั้งหมด (100%) มาวิเคราะห์ เอาแค่เพียง 27% ของเด็กเก่ง ๆ ที่ได้คะแนนมาก ๆ กับ 27% ของเด็กอ่อน ๆ ที่ได้คะแนนน้อย ๆ ก็พอ เหลือตรงกลางเป็นเด็กที่ได้คะแนนกลาง ๆ อีก 46% ไม่ต้องนำมาคิด การที่ใช้ 27% ไม่ใช่ 25% หรือ 30% ก็เพราะว่ามีผู้ทำวิจัยแล้วพบว่า การตัดหัวตัดหางมาคิดวิเคราะห์อย่างละ 27% นี้ จะให้ผลได้เหมือนกับนำผลการสอบของนักเรียนทั้งหมด 100% มาวิเคราะห์เช่นกัน

การใช้เทคนิค 27% มาเป็นหลักในการหาค่าความยากง่าย และค่าอำนาจจำแนกนั้น สามารถทำได้โดยวิธีดังต่อไปนี้

1. โดยใช้ตารางสำเร็จของ Chung - Teh Fan
2. โดยวิธีใช้สูตรคำนวณอย่างง่าย ๆ

การวิเคราะห์ข้อสอบโดยใช้ตารางสำเร็จของ Chung - Teh Fan

Chung - Teh Fan ได้สร้างตารางสำเร็จสำหรับการวิเคราะห์ข้อสอบมาจากสูตรการคำนวณหา Product Moment Correlation ซึ่งจำแนกคนออกเป็นสองกลุ่มโดยใช้เทคนิค 27% ข้อสอบที่จะนำมาวิเคราะห์โดยวิธีนี้จะต้องเป็นข้อสอบที่ให้คะแนนตอบถูก 1 และตอบผิด 0 เท่านั้น ถ้าข้อสอบประเภทที่ให้คะแนนแตกต่างกันไปจากนี้ ไม่สามารถนำมาวิเคราะห์ได้ด้วยวิธีนี้

ขั้นตอนการดำเนินการวิเคราะห์ มีดังนี้

1. ตรวจให้คะแนนกระดาษคำตอบ
2. เรียงกระดาษคำตอบของนักเรียนจากคะแนนสูงสุดมาหาคะแนนต่ำสุด กระดาษคำตอบที่มีคะแนนเท่ากันให้รวมอยู่ในลำดับเดียวกัน โดยจะให้แผ่นใดอยู่ก่อนหลังก็ได้
3. คำนวณว่า 27% ของกลุ่มที่เข้าสอบทั้งหมด คิดเป็นจำนวนเท่าใดโดยเอา .27 คูณจำนวนทั้งหมด เช่นมีผู้เข้าสอบ 80 คน 27% จะเท่ากับ $80 \times .27 = 21.6$ หรือ 22 คน (ถ้าตัวเลขหลัง

จุดคณนิยมสูงกว่า .5 ให้ปิดเป็นเลขจำนวนเต็ม 1 ถ้าน้อยกว่าหรือเท่ากับ .5 ให้ปิดทิ้งไป เช่น 20.4 ให้คิดเป็น 20 คน) ถ้ามีผู้สอบ 120 คน 27% จะเท่ากับ $120 \times .27 = 32.4 = 32$ คน

ในการวิเคราะห์ข้อสอบโดยใช้ตารางสำเร็จของ Chung-Teh Fan นั้นผู้ทำการวิเคราะห์ต้องระลึกไว้ในใจว่า วิธีนี้เหมาะสำหรับการวิเคราะห์ที่มีคนสอบจำนวนมาก น่าจะเกิน 30 คนขึ้นไป ถ้ามีคนสอบเพียง 20 คน ก็ไม่นิยมนำมาวิเคราะห์ด้วยวิธีนี้ เพราะผลที่ได้จะคลาดเคลื่อนไปควรวิเคราะห์ด้วยวิธีอื่น แต่ถ้ามีผู้สอบหลายพันคนก็ไม่จำเป็นต้องนำกระดาษคำตอบทั้งหมดมาวิเคราะห์ เพราะเป็นการเสียเวลามากเกินไปเราอาจจะสุ่มตัวอย่างมาเพียงบางส่วน แล้วจึงหา 27% ของกลุ่มที่เราสุ่มมา เพื่อเลือกจำนวนกระดาษคำตอบเป็นกลุ่มสูงและกลุ่มต่ำต่อไป

4. แบ่งกระดาษคำตอบของผู้เข้าสอบออกเป็นกลุ่มสูงและกลุ่มต่ำ (ในกรณีที่มีคนสอบทั้งหมด 80 คน 27% ของกลุ่มสูงก็คือ 22 คน) การนับกลุ่มสูงให้นับจากผู้ที่ได้คะแนนสูงสุดลงไป ส่วนการนับกลุ่มต่ำให้นับจากผู้ที่ได้คะแนนต่ำสุดขึ้นไป สำหรับกลุ่มกลางที่เหลือ ให้ตัดทิ้งไปไม่ต้องนำมาคำนวณ ถ้ามีกระดาษคำตอบที่จะเอามาวิเคราะห์ทั้งหมด 80 คน ก็ให้นับจากคนที่ได้คะแนนสูงสุดลงมา 22 คน เป็นพวกกลุ่มสูง และนับจากคนที่ได้คะแนนต่ำสุดขึ้นไป 22 คน เป็นพวกกลุ่มต่ำ

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

6. นำกระดาษคำตอบของคนกลุ่มต่ำไปทำรอยคะแนนเช่นเดียวกับของกลุ่มสูง เพื่อจะได้หาผลรวมของแต่ละตัวเลือกของแต่ละข้อคำถาม (L) และนำไปคิดค่าสัดส่วนของกลุ่มต่ำ (P_L) การขีดรอยคะแนนของกลุ่มต่ำให้แยกทำในกระดาษคนละแผ่นกับกลุ่มสูง

7. คำนวณหาค่า 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)

8. นำค่า P_H และ P_L ของแต่ละตัวเลือกไปเปิดเทียบหาค่า p และ r จากตารางวิเคราะห์ข้อสอบของ Chung-Teh Fan

ตัวอย่างตารางวิเคราะห์ข้อสอบเมื่อมีคนสอบ 80 คน และมีข้อสอบ 50 ข้อ

ตาราง 9.1 ตารางรอยคะแนนของกลุ่มสูง

ข้อที่	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

ตาราง 9.2 ตารางรอยคะแนนของกลุ่มคำ

ข้อที่ คนตี	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 (L)	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	.32		.45	.00	.14	.27	.14

เมื่อได้ค่า P_H และ P_L แล้ว นำมาเขียนใส่ไว้ในที่เดียวกัน เพื่อให้สะดวกในการเปิดค่า P และค่า r ดังนี้

ข้อ	ตัวเลือก	P_H	P_L	p	r	ความหมาย
1	ก	.14	.27	.20	.19	ดี
	ข	.68	.32	.50	.36	ดี
	ค	.14	.18	.16	.07	ดี
	ง	.04	.09	.06	.19	ดี
	จ	.00	.14	.06	.46	ดี
2	ก	.18	.09	.13	-.17	ไม่ดีต้องปรับปรุง
	ข	.14	.23	.18	.14	ดี
	ค	.09	.18	.13	.17	ดี
	ง	.04	.18	.10	.33	ดี
	จ	.55	.32	.43	.24	ดี
...						
50	ก	.45	.45	.45	.00	ไม่ดีต้องปรับปรุง
	ข	.00	.00	.00	.00	ไม่ดีต้องปรับปรุง
	ค	.09	.14	.11	.10	ดี
	ง	.32	.27	.29	-.06	ไม่ดีต้องปรับปรุง
	จ	.14	.14	.14	.00	ไม่ดีต้องปรับปรุง

จากตัวอย่างดังกล่าว การหาค่า P_H และ P_L ของข้อ 1 หาได้ดังนี้

ข้อ 1 ก. $P_H = 3/22 = .14$ $P_L = 6/22 = .27$

ข. $P_H = 15/22 = .68$ $P_L = 7/22 = .32$

ค. $P_H = 3/22 = .14$ $P_L = 4/22 = .18$

ง. $P_H = 1/22 = .04$ $P_L = 2/22 = .09$

จ. $P_H = 0/22 = .00$ $P_L = 3/22 = .14$

ข้อเสนอแนะในการใช้หนังสือตารางวิเคราะห์ข้อสอบของ Chung - Teh Fan

1. หนังสือเล่มนี้จะบอกค่าความยากง่าย (p) ค่าอำนาจจำแนก (r) และค่าความยากง่ายมาตรฐาน หรือเคลต้า (Δ) โดยหามาจากค่า P_H และ P_L
2. ค่า P_L จะอยู่แนวนอนของหน้ากระดาษเรียงตามแนวนอน ตั้งแต่ค่า $P_L = .01$ ถึง $P_L = .95$ ค่า P_H จะอยู่ทางสดมภ์ซ้ายมือสุดและขวามือสุดซึ่งจะมีค่าสูงสุดเท่ากับ .99 ค่า P_L ที่น้อยกว่า .50 จะมีอยู่ทั้งสองหน้ากระดาษ ค่า P_L ที่มากกว่า .50 จะมีอยู่เพียงหน้าเดียวของกระดาษ
3. หนังสือนี้จะเปิดอ่านค่าได้ก็ต่อเมื่อ P_L น้อยกว่า P_H ถ้า P_L มากกว่า P_H จะต้องสลับค่า P_H ให้เป็นค่า P_L และเปลี่ยนค่า P_L เป็น P_H ไว้ในใจเสียก่อน (แต่ไม่ต้องไปแก้ตัวเลขในตารางที่เราทำรอยคะแนน) แล้วจึงเปิดอ่านค่าในตาราง
4. ในการเปิดอ่านค่า p และ r จากตารางนั้น ให้เปิดหาค่า P_L ก่อนที่แนวนอนจนพบว่าอยู่หน้าใดแล้วจึงดูค่า P_H ที่สดมภ์ขวาสุด หรือซ้ายสุดจนพบ ลากเส้นจากจุด P_L และ P_H ในแนวดิ่งและแนวนอนตามลำดับมาตัดกันที่ใด จุดนั้นก็จะเป็นค่า p และ r ของตัวเลือกนั้น

ตัวอย่าง $P_H = .55$, $P_L = .32$

ขั้นที่ 1 เปิดหาค่า $P_L = .32$ ก่อน พบว่าอยู่หน้า 235 และหน้า 236

ขั้นที่ 2 ดูค่า $P_H = .55$ พบว่าอยู่หน้า 235

ขั้นที่ 3 จากสดมภ์ $P_L = .32$ และแนว $P_H = .55$ ไล่มาตัดกัน

จะพบค่าตัวเลข 3 ตัวคือ .43, .24, 13.7 ซึ่งเป็นค่าของ p , r , Δ ตามลำดับ

ในการวิเคราะห์ข้อสอบนั้นค่าที่เรานำมาใช้ในการพิจารณา คือค่า p และค่า r

เท่านั้น

ในการคัดเลือกข้อสอบนั้น เราจะคัดเลือกโดยพิจารณาจากค่า p และค่า r ของตัวเลือกที่เป็นตัวถูก สำหรับค่า r หรือค่าอำนาจจำแนกของตัวเลือกที่เป็นตัวลวงนั้น เรานำมาใช้เป็นแนวทางในการปรับปรุงข้อสอบให้มีคุณภาพดีขึ้น

4.1 การหาค่า p และ r ของตัวถูก

- 1) ถ้า P_H มากกว่า P_L เปิดตามวิธีที่กล่าวมาแล้วข้างต้นแล้วอ่านค่า p และ r ตามตาราง เช่น

P_H	P_L	p	r
.68	.32	.50	.36

ตัวเลือกที่เป็นตัวถูกนี้ กลุ่มเก่งเลือกตอบมากกว่ากลุ่มอ่อน (∴) P_H มากกว่า P_L แสดงว่าเป็นตัวเลือกที่ดี

- 2) ถ้า P_H น้อยกว่า P_L จะต้องสลับค่า P_H กับ P_L ในใจเสียก่อน แล้วจึงเปิดตารางหาค่า p และ r ตามวิธีข้างต้น ในกรณีนี้ เมื่อได้ค่า r แล้ว ให้ใส่เครื่องหมายลบ หน้าค่าของ r แสดงถึงอำนาจจำแนกที่ติดลบ

ตัวอย่าง

P_H	P_L	p	r
.32	.68	.50	-.36

ตัวถูกนี้เด็กเก่งเลือกตอบน้อยกว่าเด็กอ่อน แสดงว่าเป็นข้อสอบที่ไม่ดีควรตัดทิ้งไป

4.2 การหาค่า p และ r ของตัวหลง

- 1) ถ้า P_H น้อยกว่า P_L จะต้องสลับค่า P_H เป็น P_L ในใจแล้วเปิดหาค่า p และ r ตามปกติ ค่า r ที่ได้จะเป็นค่าบวก เพราะสำหรับตัวหลงนั้นเราต้องการให้คนอ่อนเลือกตอบมากกว่าคนเก่ง

ตัวอย่าง

P_H	P_L	p	r
.14	.27	.20	.19

ตัวดวงนี้ดี เพราะกลุ่มเก่งเลือกตอบตัวดวงนี้น้อยกว่ากลุ่มอ่อน นั่นคือ คนเก่งจริงรู้ว่าตัวเลือกตัวนี้ไม่ใช่ตัวเลือกที่ถูก จึงไม่เลือก ส่วนคนอ่อนเลือกมากกว่าเพราะไม่รู้จริง ค่าอำนาจจำแนก (r) ที่วิเคราะห์ได้จึงมีค่าเป็นบวก

2) ถ้า P_H มากกว่า P_L ให้เปิดตารางตามวิธีในข้อ 4 แต่จะต้องใส่เครื่องหมายลบที่หน้า r เช่น

P_H	P_L	p	r
.27	.14	.20	-.19

ตัวดวงนี้ไม่ดี เพราะกลุ่มเก่งเลือกตอบตัวดวงมากกว่ากลุ่มอ่อนค่าอำนาจจำแนกจึงเป็นลบ แสดงว่าตัวเลือกนี้จะต้องมีอะไรคลุมเครือไม่แจ่มชัดเพราะที่จริงแล้วกลุ่มเก่งจะต้องเลือกตอบตัวเลือกที่ไม่ถูกน้อยกว่ากลุ่มอ่อน

5. ถ้าค่า P_H หรือ P_L น้อยมาก และไม่มีในตาราง ให้ใช้ค่าต่ำสุดที่ใกล้เคียงนั้นเปิดหาค่า p และ r แล้วใส่เครื่องหมายดอกจัน (*) บนตัว p และตัว r เพื่อแสดงว่าเป็นค่าประมาณ (ถ้า P_H หรือ P_L ที่เป็น .00 ให้ปัดเป็น .01)

6. ถ้าค่า P_H หรือ P_L สูงมาก และไม่มีในตาราง ให้ใช้ค่าสูงสุดที่ใกล้เคียงในตารางนั้นเปิดหาค่า p และ r แล้วให้ใส่เครื่องหมายดอกจัน (*) บนตัว p และตัว r เพื่อแสดงว่าเป็นค่าประมาณ (สำหรับ P_H หรือ P_L ที่มีค่าเป็น 1.00 ให้ลดลงเป็น .99)

การวิเคราะห์ข้อสอบโดยการใช้สูตรคำนวณอย่างง่าย ๆ

1. สำหรับตัวเลือกที่เป็นค่าตอบถูก ให้ใช้สูตรคำนวณข้างล่างนี้ได้เลย

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

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

- เมื่อ
- p = ระดับความยากง่ายของข้อสอบ
 - r = อำนาจจำแนกของข้อสอบ
 - H = จำนวนคนที่เลือกตอบตัวเลือกใด ๆ ของกลุ่มสูง
ซึ่งในที่นี้ก็คือช่องรวมตามแนวยืนนั่นเอง
 - L = จำนวนคนที่เลือกตอบตัวเลือกใด ๆ ของกลุ่มต่ำซึ่ง
ก็คือช่องรวมตามแนวยืนนั่นเอง
 - n_H = จำนวนคนทั้งหมดของกลุ่มสูง (27% ของคนสอบ)
 - n_L = จำนวนคนทั้งหมดของกลุ่มต่ำ (27% ของคนสอบ)

วิธีคำนวณหาค่า H และ L นั้น ทำแบบเดียวกับขั้นตอนการดำเนินการวิเคราะห์ข้อสอบ โดยใช้ตารางสำเร็จของ Chung-Teh Fan ชั้นที่ 1 ถึง 6 ก็จะได้ค่า H และ L จากตาราง 9.1 และ 9.2

ตัวอย่าง การหาค่า p และ r ของข้อสอบข้อ 1 จากตาราง 9.1 และ 9.2 ซึ่งค่าตอบถูก คือ ข.

ค่า H ของตัวเลือก ก. = 3,	ค่า L ของตัวเลือก ก. = 6
" " ข. = 15,	" " ข. = 7
" " ค. = 3,	" " ค. = 4
" " ง. = 1,	" " ง. = 2
" " จ. = 0,	" " จ. = 3

ดังนั้นค่า p ของตัวเลือก ข. = $\frac{15 + 7}{22 + 22} = \frac{22}{44} = .50$

ค่า r ของตัวเลือก ข.
$$= \frac{15 - 7}{22} = \frac{8}{22} = .36$$

2. สำหรับตัวเลือกที่เป็นตัวลง เนื่องจากเหตุผลที่ได้อธิบายไว้ในตอนแรกแล้วว่าในการเขียนข้อสอบนั้น เราต้องการให้คนอ่อนตอบผิด คนเก่งมีความรู้จริงจึงจะตอบถูกดังนั้น โดยปกติแล้วคนอ่อนจะต้องเลือกตอบตัวลงมากกว่าคนเก่ง ค่า L จึงมักจะมากกว่าค่า H เสมอ (ถ้าเป็นตัวลงที่ดี) สูตรการหาค่าอำนาจจำแนก (r) จึงต้องแตกต่างไปจากการหาค่า r ของตัวถูก ดังนี้

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

สำหรับสูตรการหาค่า P ของตัวลงนั้นเหมือนกับสูตรการหา P ของตัวถูก ตัวอย่าง การหาค่า P และ r ของตัวลงของข้อสอบข้อ 1

ตัวเลือก ก. $r = \frac{6 - 3}{22} = .14, P = \frac{6 + 3}{22 + 22} = .20$

" ค. $r = \frac{4 - 3}{22} = .05, P = \frac{4 + 3}{22 + 22} = .16$

" ง. $r = \frac{2 - 1}{22} = .05, P = \frac{2 + 1}{22 + 22} = .07$

" จ. $r = \frac{3 - 0}{22} = .14, P = \frac{3 + 0}{22 + 22} = .07$

การแปลความหมายค่าวิเคราะห์ตัวเลือกที่เป็นตัวลง

1. ค่า P จะเป็นตัวบ่งบอกให้เราทราบว่ามีคนเลือกตัวลงเป็นสัดส่วนเท่าไร ถ้าค่า P ต่ำมาก แสดงว่าตัวลงข้อนั้นไม่ค่อยมีใครเลือก เห็นควรจะต้องปรับปรุง เพราะไม่เกิดประโยชน์ใดๆ ที่จะเขียนไว้เฉยๆ โดยไม่มีคนเลือกตอบ

2. ค่า r มีหลักในการพิจารณาดังนี้

ถ้าค่า r ติดลบ แสดงว่าตัวเลือกนั้นไม่ดี ต้องแก้ไข

ถ้าค่า r เป็นศูนย์ แสดงว่าตัวเลือกนั้นไม่ดี ต้องแก้ไข

ถ้าค่า r เป็นบวก แสดงว่าเป็นตัวเลือกที่ดี

แนวคิดในการเลือกข้อสอบ

โดยทั่ว ๆ ไปแล้ว ในการคัดเลือกข้อสอบวัดผลสัมฤทธิ์นั้น ควรจะคัดเลือกข้อสอบที่มีค่าอำนาจจำแนก (r) ตั้งแต่ .2 ขึ้นไป จนถึง + 1.00 ส่วนค่าความยากง่าย (P) มีหลักการพิจารณา ดังนี้

1. ในการคัดเลือกข้อสอบเพื่อใช้ในการสอบประจำภาคเรียน หรือสอบไล่ปลายปี (หรือเมื่อจบกระบวนการวิชาใดกระบวนการวิชาหนึ่ง) นั้น ควรเลือกข้อสอบที่มีความยากง่ายปานกลาง ($P = .41$ ถึง $.60$) หรืออาจจะเลือกข้อสอบที่ค่อนข้างง่ายหรือค่อนข้างยากปนอยู่บ้างก็ได้ แต่ไม่ควรเลือกข้อสอบที่มีค่า p ใกล้ 1.00 หรือ .00 เพราะข้อสอบที่ง่ายเกินไป นักเรียนทุกคนจะทำได้หมด แต่เข้าเป็นข้อสอบที่ยากเกินไปนักเรียนทุกคนก็ทำไม่ได้ ดังนั้นข้อสอบที่ยากเกินไปหรือง่ายเกินไปจึงไม่ควรนำมาใช้ในการทดสอบ เมื่อเรียนจบกระบวนการวิชา เพราะข้อสอบประเภทนี้ไม่สามารถจำแนกได้ว่านักเรียนคนไหนเก่งกว่ากัน

2. ในการคัดเลือกข้อสอบเพื่อใช้ในการสอบคัดเลือกนั้น ไม่ควรเลือกข้อสอบที่ค่อนข้างง่ายมาใช้ ข้อสอบควรจะเป็นข้อสอบที่ค่อนข้างยาก เพราะเราต้องการคัดเลือกเด็กเก่งเข้ามาเรียน ค่า p ควรจะมีค่าตั้งแต่ .05 ถึง .60 แต่ไม่ควรจะเกิน .60 เพราะจะเป็นข้อสอบที่ค่อนข้างง่าย

การวิเคราะห์ข้อสอบรายข้อแบบง่าย ๆ

1. แบบขงมือ

วิธีนี้เป็นวิธีวิเคราะห์ข้อสอบอย่างง่าย ๆ เหมาะสำหรับใช้วิเคราะห์ข้อสอบภายในห้องเรียนที่มีนักเรียนไม่มากนัก วิธีนี้จะช่วยให้ครูทราบว่าข้อสอบที่ตนเขียนขึ้นมีข้อบกพร่องตรงไหน เมื่อครูตรวจข้อสอบและให้คะแนนแล้วก็ส่งคืนให้แก่นักเรียน เพื่อให้นักเรียนตรวจสอบดูว่าข้อใดที่นักเรียนทำผิด ในการวิเคราะห์ข้อสอบแบบนี้ให้ครูถามนักเรียนว่า ข้อไหนใครทำผิดบ้าง โดยเริ่มจากข้อที่ 1 จนถึงข้อสุดท้าย ใครทำข้อไหนผิดให้ยกมือขึ้น แล้วให้ครูจดจำนวนที่เด็กทำผิดแต่ละข้อไว้บนกระดาษค่า ข้อสอบข้อใดที่เด็กส่วนใหญ่ทำผิด หรือนักเรียนทั้งห้องทำผิด จัดเป็นข้อสอบที่ค่อนข้างยาก ไม่ควรนำมาใช้อีก หรือข้อสอบข้อใดไม่มีคนทำผิดเลย หรือทำผิดเพียง 1 - 2 คน ถือว่าเป็นข้อสอบที่ไม่ดีอีกเช่นกัน ไม่ควรนำมาใช้อีก

2. แบบใช้สูตรคำนวณ

การใช้สูตรคำนวณนี้ให้ครูเรียงกระดาษคำตอบของนักเรียนจากคะแนนสูงไปหาคะแนนต่ำสุด แล้วหาว่า 27% ของผู้เข้าสอบทั้งหมดมีกี่คน จากนั้น ก็แบ่งกระดาษคำตอบออกเป็นสอง

พวก คือพวกกลุ่มสูงและพวกกลุ่มต่ำ ตัวอย่างเช่น มีคนสอบ 50 คน 27% ของ 50 เท่ากับ 13 คน (โดยประมาณ) จากนั้นก็แบ่งกระดาษคำตอบเป็น 2 พวกพวกหนึ่งคือพวกกลุ่มสูง 13 คน (โดยนับจากคนที่ได้คะแนนสูงสุดลงมา) และพวกกลุ่มต่ำ 13 คน (โดยนับจากคนที่ได้คะแนนต่ำสุดขึ้นไป) แล้วมาหาว่าข้อสอบแต่ละข้อนั้นกลุ่มสูงทำถูกกี่คน กลุ่มต่ำทำถูกกี่คน จากนั้นจึงนำค่าที่นับได้มาใช้ในการหาค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) จากสูตร ดังนี้

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

$$r = \frac{H - L}{n_H}$$

เมื่อ p = ระดับความยากง่ายของข้อสอบ

r = อำนาจจำแนกของข้อสอบ

H = จำนวนคนในกลุ่มสูงที่ทำข้อสอบแต่ละข้อถูก

L = จำนวนคนในกลุ่มต่ำที่ทำข้อสอบแต่ละข้อถูก

n_H = จำนวนคนทั้งหมดของกลุ่มสูง (27% ของคนสอบ)

n_L = จำนวนคนทั้งหมดของกลุ่มต่ำ (27% ของคนสอบ)

ซึ่ง n_L จะเท่ากับ n_H เสมอ

การวิเคราะห์ข้อสอบแบบอิงเกณฑ์

เหตุผลสำคัญประการหนึ่งของการสร้างข้อสอบแบบอิงเกณฑ์ คือต้องการจะกำหนดเป้าหมายของการสอนให้กระจ่างชัดเจนที่สุด จนกระทั่งผู้เรียนส่วนใหญ่ (หรือทั้งหมด) สัมฤทธิ์ผลในการเรียน สอดคล้องกับพฤติกรรมที่กำหนดไว้ จากการศึกษาที่ผู้เรียนทุกคนสัมฤทธิ์ผลตามที่กำหนดไว้จะเป็นสาเหตุทำให้ไม่มีความแปรปรวนของคะแนน ดังนั้นถ้าการสอนมีประสิทธิภาพจริงแล้ว การหาอำนาจจำแนกของข้อสอบโดยวิธีสหสัมพันธ์ (Correlation) จึงไม่สามารถทำได้ เนื่องจากความแปรปรวนของคะแนน (S^2) มีค่าเป็นศูนย์

อนึ่ง ในการประเมินผลแบบอิงเกณฑ์นั้น โดยทฤษฎีแล้วมีหลักอยู่ว่า ก่อนที่เด็กนักเรียนจะเริ่มเรียนบทเรียนใดบทเรียนหนึ่ง เด็กควรจะไม่มีความรู้ในหน่วยการเรียนนั้น (ถ้านักเรียนรู้ในหน่วยการเรียนนั้นแล้ว ครูก็ไม่จำเป็นต้องสอนอีก เพราะจะทำให้เสียเวลาทั้งของครูและของนักเรียนเอง) ถ้าการสอนมีประสิทธิภาพ หลังจากที่ครูทำการสอนไปแล้วนักเรียนควรจะได้คะแนนเต็มทุก

คนจะเห็นได้ว่าการประเมินผลแบบอิงเกณฑ์นั้นมุ่งที่จะให้นักเรียนมีความรู้เท่าเทียมกัน ซึ่งจะทำให้คะแนนที่ได้จากการทดสอบ หลังจากที่ยื่นจบบทเรียนไปแล้ว มีความแปรปรวนน้อยมาก หรือไม่มีความแปรปรวนเลย ดังนั้นการวิเคราะห์ข้อสอบของการวัดแบบอิงเกณฑ์จึงต้องใช้วิธีการที่แตกต่างออกไปจากการวัดแบบอิงกลุ่มเนื่องจากการวัดแบบอิงเกณฑ์ยังเป็นเรื่องใหม่ในวงการวัดผล จึงยังไม่สามารถจะหาข้อยุติได้ว่าควรจะใช้การวิเคราะห์ข้อสอบแบบใดจึงจะดี และถูกต้องมากที่สุด ได้มีผู้เสนอวิธีการวิเคราะห์ข้อสอบประเภทอิงเกณฑ์ไว้หลายวิธี แต่ผู้เขียนจะเลือกยกตัวอย่างมาแสดงให้ดูคร่าว ๆ เพียง 2 วิธีเท่านั้น ดังนี้

วิธีที่ 1 ใช้ผู้เชี่ยวชาญทางเนื้อหาวิชาทำการตัดสินคุณภาพของข้อสอบแต่ละข้อโดยกำหนดให้ผู้ตัดสินหลาย ๆ คน ทำการวิเคราะห์และตรวจสอบดูว่าข้อสอบแต่ละข้อถามวัดเนื้อหาและพฤติกรรมสอดคล้องกับที่กำหนดไว้ในตารางวิเคราะห์หลักสูตรหรือไม่ ถ้าผู้เชี่ยวชาญเห็นว่าข้อสอบข้อใดถามวัดไม่ตรงกับที่กำหนดไว้ในตารางวิเคราะห์ ก็ให้เขียนอธิบายไว้ว่าทำไมถึงคิดว่าข้อสอบนั้นไม่สอดคล้องกับตารางวิเคราะห์หลักสูตร

ถ้าผู้เชี่ยวชาญตั้งแต่ 2 คนขึ้นไป มีความเห็นว่าข้อสอบข้อใดไม่สอดคล้องกับตารางวิเคราะห์หลักสูตร และเพราะเหตุใดสาเหตุของผู้เชี่ยวชาญทั้ง 2 คน ตรงกันหรือใกล้เคียงกัน ก็แสดงว่าข้อสอบข้อนั้นเป็นข้อสอบที่ไม่ดี ควรจะนำมาปรับปรุงแก้ไข หรือทิ้งไปเลย

วิธีที่ 2 หาความไวของข้อสอบ โดยวิธี Pre-test Post-test ดังนี้

1. นำแบบทดสอบไปทดสอบกับนักเรียนจำนวนหนึ่งก่อนทำการสอน และนักเรียนกลุ่มนี้จะต้องยังไม่เคยเรียนเรื่องที่น่ามาออกข้อสอบมาก่อน (ซึ่งในทางปฏิบัติจริงอาจทำได้ยาก) แล้วนำมาตรวจให้คะแนน

2. นำกระดาษคำตอบทั้งหมดของนักเรียนมาขีดรอยคะแนนลงในตารางเพื่อดูว่าข้อไหนมีคนทำถูกกี่คน แล้วเรหาค่าความยากง่ายของข้อสอบ แต่ละข้อจากสูตร

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

ค่า p ที่ได้เราเรียกว่า P_{pre}

3. หลังจากที่ได้ครูสอนจบบทเรียนแล้ว นำข้อสอบฉบับเดียวกันนี้ไปสอบกับนักเรียนกลุ่มเดิมอีก แล้วนำมาตรวจหาค่า p ด้วยวิธีที่กล่าวมาแล้วในข้อ 2 p ที่ได้นี้เรียกว่า P_{post}

4. หาค่าความไวของข้อสอบ จากสูตร

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

เมื่อ S คือค่าความไวของข้อสอบ ตัวอย่างการหาค่า S แสดงไว้ในตาราง 9.3 - 9.5

ตาราง 9.3 ตารางคำนวณค่า P_{pre} ของข้อสอบวิชาสังคมศึกษา (15 ข้อ) ชั้น ม.1
จำนวนนักเรียนที่สอบ 10 คน สอบวันที่ 11 พฤศจิกายน 2524

ข้อที่	คนที่										รวม ความดี	ค่า P_{pre}	
	1	2	3	4	5	6	7	8	9	10			
1												—	0/10 = .00
2	/	/	/	/	/	/	/	/	/	/		10	10/10 = 1.00
3	/	/		/		/	/					5	5/10 = .50
4												—	0/10 = .00
5													
6													
7													
.													
.													
.													
15													

การแปลความหมายตาราง 9.3

- ข้อสอบข้อ 1 ไม่มีนักเรียนตอบถูกเลย ค่า $P_{pre} = .00$
- ข้อสอบข้อ 2 นักเรียนทุกคนตอบถูก ค่า $P_{pre} = 1.00$
- ข้อสอบข้อ 3 นักเรียนตอบถูกเพียง 5 คน ค่า $P_{pre} = .50$
- ข้อสอบข้อ 4 ไม่มีนักเรียนตอบถูกเลย ค่า $P_{pre} = .00$

ตาราง 9.4 ตารางคำนวณค่า P_{post} ของข้อสอบวิชาสังคมศึกษา (15 ข้อ) ชั้น ม.1
จำนวนนักเรียนที่สอบ 10 คน สอบวันที่ 15 ธันวาคม 2524

ข้อที่ \ กนที่	1	2	3	4	5	6	7	8	9	10	รวม ความถี่	ค่า P_{post}
1	/	/	/	/	/	/	/	/	/	/	10	10/10 = 1.00
2											—	0/10 = .00
3	/	/	/	/	/	/	/	/	/	/	10	10/10 = 1.00
4											—	0/10 = .00
5												
6												
.												
.												
.												
15												

การแปลความหมายตาราง 9.4

ข้อสอบข้อ 1 นักเรียนทุกคนตอบถูก	ค่า $P_{post} = 1.00$
ข้อสอบข้อ 2 ไม่มีนักเรียนตอบถูกเลย	ค่า $P_{post} = .00$
ข้อสอบข้อ 3 นักเรียนทุกคนตอบถูก	ค่า $P_{post} = 1.00$
ข้อสอบข้อ 4 ไม่มีนักเรียนตอบถูกเลย	ค่า $P_{post} = .00$

ตาราง 9.5 ตารางคำนวณค่า S (ความไวของข้อสอบ) ของวิชาสังคมศึกษา ชั้น ม.1

ข้อสอบ ข้อที่	P _{post}	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 แต่ต้องเปลี่ยนแปลง จุดประสงค์ให้ซับซ้อนขึ้นกว่าเดิม
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 ต้องตัดทิ้งไป
.					
.					
.					
15					

การแปลความหมายค่า S จากตาราง 9.5

ข้อสอบข้อ 1 S มีค่าเท่ากับ 1.00 จัดได้ว่าเป็นข้อสอบที่ดี เป็นไปตามอุดมคติจริงๆ เนื่องจากก่อนสอบครั้งแรกไม่มีนักเรียนตอบถูกเลย ($p = .00$) แต่หลังจากผ่านการเรียนไปแล้ว นักเรียนทุกคนทำข้อสอบถูก ($p = 1.00$) จึงเป็นข้อสอบที่มีความไวในการวัดผลมากที่สุด ข้อสอบข้อนี้ต้องเก็บเอาไว้ใช้อีก

ข้อสอบข้อ 2 S มีค่าเท่ากับ -1.00 จัดได้ว่าเป็นข้อสอบที่ไม่ดีต้องตัดทิ้งไป ทั้งนี้เนื่องจากก่อนสอบครั้งแรกนักเรียนทุกคนตอบข้อสอบถูกหมดทุกข้อ แสดงว่ามีความรู้ในเรื่องนั้นเป็นอย่างดี แต่พอนักเรียนเรียนเนื้อหาอันนั้นไปแล้ว มาทำข้อสอบอีกกลับทำข้อสอบผิดหมดทุกคน แสดงว่าคำถามอาจจะกำกวม หรืออาจจะเป็นการสอนไม่ถูกต้อง เด็กอาจจะเกิดความเข้าใจผิดในเนื้อหาที่เรียน ทั้ง ๆ ที่ก่อนสอนเด็กเข้าใจถูกต้องอยู่แล้ว

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

จะเป็นข้อสอบที่ต้องปรับปรุงจุดประสงค์เชิงพฤติกรรม ที่จะวัดให้ซับซ้อนขึ้นได้ โดยที่ยังตามเนื้อหาเดิมอยู่

ข้อสอบข้อ 4 และข้อ 7 เป็นข้อสอบที่ไม่ดีทั้งคู่เพราะ S มีค่าเท่ากับ .00

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

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

กล่าวโดยสรุปแล้วในการพิจารณาคัดเลือกข้อสอบ โดยพิจารณาผลจากการวิเคราะห์ข้อสอบด้วยวิธีนี้การหาค่าความไวของข้อสอบ (S) ควรถือหลักดังนี้

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

ธนาคารข้อสอบ (Item Bank)

คุณประโยชน์ที่สำคัญยิ่งที่ได้จากการวิเคราะห์ข้อสอบก็คือ การสร้างธนาคารข้อสอบ (Item Bank) ธนาคารนี้คือที่รวมข้อคำถามดี ๆ ที่เราคัดเลือกแล้วจากการวิเคราะห์

วิธีการสร้างธนาคารข้อสอบนั้นอาจทำได้โดยหากระดาษแข็งมาตัดเป็น "บัตรข้อสอบ" (Item card) ให้มีขนาดกว้างประมาณ 4 - 6 นิ้วยาวประมาณ 6 - 8 นิ้ว ควรเตรียมบัตรนี้ไว้มาก ๆ เพราะจะต้องใช้บัตร 1 แผ่น สำหรับคำถาม 1 ข้อ ถ้าใช้กระดาษหลาย ๆ สีก็ยิ่งดี เพราะจะได้สังเกตเห็นได้ง่ายว่าสีใดเป็นของวิชาใด แบบฟอร์มของบัตรข้อสอบนี้ อาจดัดแปลงให้เหมาะสมกับความต้องการของเราได้ แต่ควรจะมีสิ่งต่าง ๆ ข้างล่างนี้อยู่ด้วย ดังนี้

ในด้านหน้าของบัตร ควรประกอบไปด้วย

1. เลขที่ ควรใส่เลขที่ของคำถามที่เราคัดเลือกแล้ว โดยเริ่มเรียงตั้งแต่ 1 เป็นต้นไป ตามลำดับ เพื่อสะดวกในการค้นหา
2. วิชา - ชั้น ควรระบุให้ชัดเจนว่าวิชาใดใช้กับชั้นใด
3. ตัวข้อสอบ พิมพ์ข้อสอบลงไป พร้อมกับกาเครื่องหมายดอกจันทรีไว้ที่ตัวถูก
4. ความมุ่งหมาย ควรระบุไว้ด้วยว่าข้อสอบข้อนี้ใช้สำหรับวัดพฤติกรรมอะไร

ตัวอย่าง ด้านหน้าของบัตรข้อสอบ

เลขที่ 5 วิชาวิทยาศาสตร์ ชั้น ม.1

อุปกรณ์ชนิดใดที่เปลี่ยนพลังงานกลเป็นพลังงานไฟฟ้า

ก. พัดลม

ข. ตู้เย็น

ค. มอเตอร์

• ง. ไดนาโม

จ. แบตเตอรี่

ด้านหลังของบัตรควรประกอบไปด้วย

- ใช้ครั้งที่เท่าไร เมื่อไร ควรบันทึกไว้ว่าข้อสอบนี้เคยใช้มากี่ครั้งแล้วและใช้เพื่อสอบประจำภาค หรือสอบคัดเลือก ฯลฯ
- ค่าวิเคราะห์ ให้บ่งค่าวิเคราะห์ p และ r ของการใช้ข้อสอบแต่ละครั้ง ทั้งตัวลงและตัวถูก

ตัวอย่างด้านหลังของบัตรข้อสอบ

ใช้ครั้งที่ วันที่ ชั้น จำนวนนักเรียน เพื่อ	1	2	3	4
	20 สิงหาคม 24			
	ม.1			
	50			
	ทดสอบย่อย			
ตัวเลือก	p	r	p r	p r
ก	.20	.19		
ข	.16	.07		
ค	.06	.19		
ง	.50	.36		
จ	.06	.46		

การคัดเลือกข้อสอบเพื่อเก็บไว้ในธนาคารข้อสอบ

- โดยปกติแล้วถ้าเราต้องการเก็บข้อสอบเพื่อไว้ใช้ในการทดสอบเมื่อสิ้นภาคเรียน หรือเมื่อสอนจบกระบวนการแล้วนั้น ควรจะคัดเลือกเก็บข้อสอบที่มีค่า r ตั้งแต่ .20 ขึ้นไป ส่วน p นั้น ควรมีค่าอยู่ระหว่าง .20 ถึง .80 เอาไว้ในธนาคาร
- เนื่องจากในบางครั้งเราอาจจะต้องการเลือกข้อคำถามจากธนาคารเอาไปใช้ในการสอบคัดเลือก ซึ่งมีจุดมุ่งหมายที่จะคัดเอาเฉพาะคนเก่งจริง ๆ ดังนั้นข้อสอบที่จัดว่าเป็นข้อสอบที่ดีสำ

หรับเก็บไว้ในธนาคารเพื่อนำมาใช้ในการสอบคัดเลือกในคราวต่อไป ควรจะมีค่า p ตั้งแต่ .05 ถึง .06 ส่วนค่า r ก็ควรมีค่าตั้งแต่ .20 หรือใกล้เคียงกับ .20 ขึ้นไป

3. ในบางครั้งเราต้องการคัดเลือกข้อสอบเพื่อนำไปทดสอบหลังจากที่สอบจบหน่วยเรียน (หรือบทเรียน) ย่อย ๆ แล้ว การวัดแบบนี้เป็นการวัดว่านักเรียนมีความรู้ในสิ่งที่ครูสอนไปหรือไม่ ถ้าการสอนได้ผลนักเรียนจะต้องเปลี่ยนพฤติกรรมจาก "ไม่รู้" เป็น "รู้" นั่นก็คือนักเรียนทุกคนควรจะได้คะแนนเต็มหรือใกล้เคียงเต็ม การวัดแบบนี้เป็นการวัดแบบอิงเกณฑ์นั่นเอง ฉะนั้นเราก็ควรมีข้อสอบที่ค่อนข้างจะง่าย ซึ่งจัดเนื้อหา พฤติกรรมตามที่ต้องการเก็บไว้ในธนาคารข้อสอบด้วย เพื่อจะได้ประหยัดเวลาของครูในการออกข้อสอบใหม่ ดังนั้นเราจึงควรคัดเลือกข้อสอบที่มีค่า p ระหว่าง .70 ถึง .95 เก็บไว้ในธนาคารข้อสอบด้วย ข้อสอบประเภทนี้ค่า r จะต่ำ เนื่องจากเป็นข้อสอบที่ไม่มีประสิทธิภาพในการจำแนกเด็กเก่งออกจากเด็กอ่อน ซึ่งก็ตรงกับจุดมุ่งหมายของการวัดแบบอิงเกณฑ์ ที่ไม่ต้องการแบ่งเด็กออกเป็นกลุ่ม เก่ง - อ่อน แต่ต้องการจะทดสอบดูแต่เพียงว่าเขารู้เรื่องที่สอนไปหรือไม่

4. จึงอาจกล่าวโดยสรุปได้ว่า ในการคัดเลือกข้อสอบเพื่อเก็บไว้ในธนาคารข้อสอบนั้น ควรจะคัดเลือกข้อสอบที่มีค่า p ตั้งแต่ .05 ถึง .95 ส่วนค่า r นั้น พิจารณาโดยยึดหลักดังนี้

4.1 ถ้า p มีค่าระหว่าง .05 - .70 ค่า r ควรมีค่าตั้งแต่ .20 ขึ้นไป

4.2 ถ้า p มีค่ามากกว่า .70 r อาจมีค่าต่ำกว่า .20 ได้ แต่ต้องมากกว่า .00

การใช้หลักการคัดเลือกข้อสอบดังที่กล่าวมาแล้วข้างต้น จะทำให้ครูมีข้อสอบสำรองเก็บไว้ในธนาคารเพื่อการนำมาใช้ในจุดประสงค์ต่าง ๆ กัน