

การวิเคราะห์ขอมูลและการอภิปรายผลการวิจัย

๑. การวิเคราะห์แบบสอบที่นำไปทดลองสอบ

นำคะแนนของนักเรียนแต่ละคนที่ได้จากการทดลองสอบมาเรียงจากมากไปหาน้อย เพื่อวิเคราะห์รายข้อ (Item analysis) จัดแบ่ง กลุ่มตัวอย่างประชากรที่ทดลองสอบออกเป็น ๘ กลุ่มตามจำนวนแบบสอบแต่ละชุดดังนี้

- ๑.๑ กลุ่มโรงเรียนเทศบาลวัดธาตุทองที่ทดลองทำแบบสอบ O II A
- ๑.๒ กลุ่มโรงเรียนเทศบาลวัดธาตุทองที่ทดลองทำแบบสอบ O II B
- ๑.๓ กลุ่มโรงเรียนเทศบาลวัดธาตุทองที่ทดลองทำแบบสอบ JO III A
- ๑.๔ กลุ่มโรงเรียนเทศบาลวัดธาตุทองที่ทดลองทำแบบสอบ JO III B
- ๑.๕ กลุ่มโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตร ศาสตร์ ที่ทดลองทำแบบสอบ JIIA
- ๑.๖ กลุ่มโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตร ศาสตร์ ที่ทดลองทำแบบสอบ JIIB
- ๑.๗ กลุ่มโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตร ศาสตร์ ที่ทดลองทำแบบสอบ JOIIIA
- ๑.๘ กลุ่มโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตร ศาสตร์ ที่ทดลองทำแบบสอบ JOIIIB

จัดแบ่งนักเรียนเป็นแต่ละกลุ่มโดยใช้เทคนิค ๒๗ เปอร์เซ็นต์ ของจำนวนผู้เข้าสอบ เป็นกลุ่มสูงและกลุ่มต่ำ หากจำนวนคนในกลุ่มสูงและกลุ่มต่ำ ที่ตอบข้อสอบแต่ละข้อได้ถูกต้อง (U,L) แล้วหาควรระดับความยาก และค่าอำนาจจำแนกตามตารางวิเคราะห์ข้อสอบ ของ จุง-เต ฟาน^๑ (Chung-Teh Fan) ซึ่งอาจอธิบายความหมายของตัวเลข ที่แสดงระดับความยาก และค่าอำนาจจำแนกได้ดังนี้

- ๑.๑ ความหมายของตัวเลขที่แสดงระดับความยาก (Level of Difficulty)
หรือ p

^๑จุง-เต ฟาน เรื่องเดิม.

ค่าตั้งแต่	หมายความว่า
๐๐ - ๐๕	ยากเกินไป
๑๐ - ๑๕	ยากมาก
๒๐ - ๒๕	ยาก
๓๐ - ๓๕	ค่อนข้างยาก
๔๐ - ๕๐	ปานกลาง
๕๑ - ๕๕	ค่อนข้างง่าย
๖๐ - ๖๕	ง่าย
๗๐ - ๗๕	ง่ายมาก
๘๐ - ๘๕	ง่ายเกินไป

๑.๒ ความหมายของตัวเลขที่แสดงค่าอำนาจจำแนก (Biserial Correlation หรือ r มีค่าตั้งแต่ $-.๒๐$ ถึง $.๐๐$ และ ๑.๐๐

ค่าตั้งแต่	หมายความว่า
.๐๑ - .๐๕	อำนาจจำแนกต่ำมาก
.๑๐ - .๑๕	อำนาจจำแนกต่ำ
.๒๐ - .๒๕	อำนาจจำแนกพอใช้ได้
.๓๐ - .๓๕	อำนาจจำแนกดีพอใช้
.๔๐ - .๔๕	อำนาจจำแนกดี
.๕๐ ขึ้นไป	อำนาจจำแนกดีมาก
.๐๐	อำนาจจำแนกไม่บอกอะไรเลย
ค่าติดลบ	อำนาจจำแนกกลับกัน คือ เด็กอ่อน ทำถูกมากกว่าเด็กเก่ง ^๒

ข้อสอบที่มีค่า r ต่ำกว่า .๒๐ ถือว่าเป็นข้อสอบที่ไม่ดี ควรตัดออกหรือปรับปรุง ข้อสอบที่มีค่า p ตั้งแต่ ๘๐ ขึ้นไปและค่า r ตั้งแต่ .๒๐ ขึ้นไป เป็นข้อง่ายเกินไป ข้อสอบที่มีค่า p ต่ำกว่า ๒๐ ลงมาและค่า r ตั้งแต่ .๒๐ ขึ้นไป เป็นข้อยากเกินไป ข้อสอบข้อใดที่มีค่า p สูงเกิน .๕๐ ขึ้นไป จัดว่าเป็นข้อง่าย คือ มีเด็กทั้งหมดทำถูกเกิน ๕๐ เปอร์เซ็นต์ ข้อใดที่มีค่า p ต่ำกว่า .๕๐ จัดว่าเป็นข้อยาก ยิ่งค่า p น้อยขนาด .๑๐ หรือ .๐๕ ควรตัดข้อสอบข้อนั้นทิ้ง

ผลการวิเคราะห์แบบสอบที่นำไปทดลอง สอบอาจสรุปได้ดังนี้
ตารางที่ ๑ ช่วงค่าอำนาจจำแนก และความยากของแบบสอบคู่ขนานแต่ละชุดในการทดลองสอบ

โรงเรียนวัดธาตุทอง				
แบบสอบ	O II A	O II B	JO III A	JO III B
ช่วงความยากอยู่ระหว่าง	.๒๒ - .๕๕	.๒๔ - .๔๘	.๓๕ - .๔๗	.๑๓ - .๔๗
ค่าอำนาจจำแนกอยู่ระหว่าง	.๐๐ - .๔๓	.๑๔ - .๔๓	.๓๕ - .๗๒	.๒๕ - .๖๗
โรงเรียนสาธิตเกษตร				
แบบสอบ	J II A	J II B	JO III A	JO III B
ช่วงความยากอยู่ระหว่าง	.๔๗ - .๕๑	.๒๕ - .๔๔	.๒๗ - .๗๘	.๓๒ - .๔๑
ค่าอำนาจจำแนกอยู่ระหว่าง	.๕๕ - .๘๗	.๒๓ - .๗๕	.๑๒ - .๕๐	.๑๖ - .๗๗

จากตารางที่ ๑ จะเห็นว่าแบบสอบ O IIA ที่นำไปทดลองสอบ มีระดับความยากอยู่ระหว่าง .๒๒ - .๕๕ จึงเป็นข้อสอบที่ค่อนข้างง่าย ส่วนค่าอำนาจจำแนกอยู่ระหว่าง .๐๐ - .๔๓ ซึ่งแสดงว่าข้อสอบบางข้อไม่สามารถจำแนกนักเรียนที่เรียนดีและเรียนไม่ดีออกจากกันได้

แบบสอบ O IIB มีระดับความยากอยู่ระหว่าง .๒๔ - .๔๘ จึงเป็นข้อสอบที่ค่อนข้างง่าย ค่าอำนาจจำแนกอยู่ระหว่าง .๑๔ - .๔๓ ซึ่งต่ำกว่ามาตรฐานเพียงเล็กน้อย

แบบสอบ JIIA และ JIIB มีระดับความยากอยู่ระหว่าง .๔๗ - .๕๑ และ .๒๕ - .๔๔ ตามลำดับ แสดงว่ามีข้อสอบบางข้อยากเกินไป ค่าอำนาจจำแนกของแบบสอบ JIIA อยู่ระหว่าง .๔๕ - .๔๗ และแบบสอบ JIIB มีค่าอำนาจจำแนกอยู่ระหว่าง .๒๓ - .๓๕ นับว่าแบบสอบ J II A และ J II B เป็นแบบสอบที่ดี มีค่าอำนาจจำแนกสูง

แบบสอบ JO III A ที่นำไปทดลองสอบกับนักเรียนโรงเรียนวัดธาตุทอง มีระดับความยากอยู่ระหว่าง .๓๕ - .๔๗ แสดงว่ามีข้อสอบบางข้อยากเกินไป แต่ค่าอำนาจจำแนกอยู่ในเกณฑ์คืออยู่ระหว่าง .๓๕ - .๓๒ จึงเป็นข้อสอบที่ดี มีค่าอำนาจจำแนกสูง

แบบสอบ JO IIIB ที่นำไปทดลองสอบกับนักเรียนโรงเรียนวัดธาตุทอง มีระดับความยากอยู่ระหว่าง .๑๓ - .๔๗ แสดงว่ามีข้อสอบบางข้อยากเกินไปและบางข้อง่ายเกินไป ควรปรับปรุง ค่าอำนาจจำแนกอยู่ระหว่าง .๒๕ - .๒๗ นับว่าอยู่ในระดับมาตรฐาน มีค่าอำนาจจำแนกดี

แบบสอบ JO IIIA ที่นำไปทดลองสอบกับนักเรียนโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ มีระดับความยากอยู่ระหว่าง .๒๗ - .๓๕ นับว่าเป็นข้อสอบที่มีระดับความยากพอเหมาะ ค่าอำนาจจำแนกอยู่ระหว่าง .๑๒ - .๕๐ แสดงว่ามีข้อสอบบางข้อมีอำนาจจำแนกต่ำ

แบบสอบ JO IIIB ที่นำไปทดลองสอบกับนักเรียนโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ มีระดับความยากอยู่ระหว่าง .๓๒ - .๔๑ จึงมีระดับความยากพอเหมาะ ค่าอำนาจจำแนกอยู่ระหว่าง .๑๖ - .๓๗ แสดงว่ามีข้อสอบบางข้อมีค่าอำนาจจำแนกต่ำกว่าระดับมาตรฐาน

ตารางแสดงช่วงค่าอำนาจจำแนก และความยากของแบบสอบคู่ขนานแต่ละชุดในการทดลองสอบไม่ให้รายละเอียดว่ามีจำนวนข้อสอบที่มีความยากและค่าอำนาจจำแนกอยู่ในระดับมาตรฐานจำนวนเท่าใดในแต่ละแบบสอบ จึงจำเป็นต้องพิจารณาจากแผนภาพประกอบ แต่ก่อนอื่นจะขออธิบายความหมายของแผนภาพในกราฟดังนี้

ความหมายของแผนภาพ

เส้นประตงระดับ ๕๐ จะเป็นเส้นแบ่งคะแนนเป็น ๒ ซีก ซีกขวาเป็นข้อสอบประเภทายซีกซ้ายเป็นข้อสอบประเภทยาก

เส้นประตงระดับ ๒๐ และเส้นประตงระดับ ๘๐ จะเป็นเส้นแบ่งคะแนนที่อยู่ในระดับมาตรฐาน คะแนนที่อยู่ซีกซ้ายของเส้นประตงระดับ ๒๐ จะเป็นข้อสอบที่ยากเกินไป และคะแนนที่อยู่ซีกขวาของเส้นประตงระดับ ๘๐ เป็นข้อสอบที่ง่ายเกินไป

เส้นประตงแนวนอน มีระดับอำนาจจำแนก ๑.๒๐ ขึ้นไป จะแบ่งคำถามออกเป็น ๒ ข้าง บน - ล่าง ข้างบนเป็นข้อที่มีอำนาจจำแนกสูง ข้างล่างเป็นข้อที่มีอำนาจจำแนกต่ำ

เมื่อดูพรอมๆกัน เส้นสองเส้นนี้ แบ่งคำถามทั้งหมดออกเป็น ๔ ประเภทคือ

๑. พวกที่อยู่เหนือเส้นประตงแนวนอน ระหว่างเส้นประตงระดับ ๕๐ - ๘๐ จัดเป็นข้อสอบที่ง่ายและมีอำนาจจำแนกสูง (ดี)
๒. พวกที่อยู่เหนือเส้นประตงแนวนอน ระหว่างเส้นประตงระดับ ๒๐ - ๕๐ จัดเป็นข้อสอบที่ยากและมีอำนาจจำแนกสูง (ดี)
๓. พวกที่อยู่ใต้เส้นประตงแนวนอน ระหว่างเส้นประตงระดับ ๕๐ - ๘๐ จัดเป็นข้อสอบที่ง่ายและมีอำนาจจำแนกต่ำ (ไม่ดี)
๔. พวกที่อยู่ใต้เส้นประตงแนวนอน ระหว่างเส้นประตงระดับ ๒๐ - ๕๐ จัดเป็นข้อสอบที่ยากและมีอำนาจจำแนกต่ำ (ไม่ดี)

จากการอาทกราฟ จะชี้ให้เราเห็นธรรมชาติและคุณภาพของคำถามได้หลายประการ เช่น ข้อใดอยู่ในแนวอนระดับเดียวกัน แปลว่ามีอำนาจจำแนกปานกัน ถ้าอยู่ในแนวตั้งตรงกันก็มีระดับความยากเท่ากัน ถ้าอยู่ใกล้ซีกกันหรือเกาะกันเป็นกระจุก หรือบางที่ซ้อนทับกัน ก็แปลว่า คำถามเหล่านั้นมีระดับความยาก และอำนาจจำแนกปานๆกัน หรือเท่ากัน นับว่ามีคุณค่าในการวัดผล เท่าเทียมกัน ส่วนข้อคำถามที่อยู่ภายในกรอมจัดเป็นข้อสอบที่มีคุณภาพดี

ตารางที่ ๒ ค่าสถิติที่คำนวณของแบบสอบถาม O II

ค่าสถิติที่คำนวณ	O II	
	A	B
\bar{X}	๑๖.๑๖	๑๐.๑๖
S.D.	๒.๖๑	๒.๗๗
σ_m	๐.๒๖	๐.๒๘
F ₂₁	๐.๕๖	๐.๖๑
\bar{C}_{mean}	๑.๗๕	๑.๗๓
r_{XY}	๐.๕๗	
σ_{dm}	๐.๒๕	
Z ratio	๐.๑๖ ไม่มีนัยสำคัญที่ระดับ ๐.๐๕	

(ดูการคำนวณในภาคผนวก)

จากตารางที่ ๒ จะเห็นว่าแบบสอบถาม O IIA มีค่าเฉลี่ยเท่ากับ ๑๖.๑๖ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๒.๖๑ ความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๒๖ ค่าความเชื่อถือได้ของแบบสอบถามเท่ากับ ๐.๕๖ และความคลาดเคลื่อนของการวัดเท่ากับ ๑.๗๕

แบบสอบถาม O II B มีค่าเฉลี่ยเท่ากับ ๑๐.๑๖ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๒.๗๗ ความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๒๘ ค่าความเชื่อถือได้ของแบบสอบถามเท่ากับ ๐.๖๑ และความคลาดเคลื่อนของการวัดเท่ากับ ๑.๗๓

ค่าสหสัมพันธ์ของแบบสอบถาม O IIA และ O II B มีค่าเท่ากับ ๐.๕๗ หมายความว่าแบบสอบถามทั้งสองมีลักษณะตามกันและค่าสหสัมพันธ์อยู่ในระดับกลาง ความคลาดเคลื่อนของความแตกต่างระหว่างค่าเฉลี่ยของแบบสอบถามทั้งสองเท่ากับ ๐.๒๕ อัตราส่วนวิกฤตเท่ากับ ๐.๑๖ แสดงว่าแบบสอบถาม O II ไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ ๐.๐๕

ตารางที่ ๓ ค่าสถิติที่คำนวณของแบบทดลองสอบคุชานาน J II

ค่าสถิติที่คำนวณ	J II	
	A	B
\bar{x}	๕.๓๓	๔.๕๔
S.D.	๓.๙๕	๒.๙๔
σ_m	๐.๓๙	๐.๒๗
r_{21}	๐.๘๐	๐.๖๕
σ_{meas}	๑.๖๘	๑.๙๕
r_{XX}	๐.๙๓	
σ_{dm}	๐.๒๔	
χ^2 ratio	- ๒.๖๓ มีนัยสำคัญที่ระดับ ๐.๐๕	

(ดูการคำนวณในภาคผนวก)

จากตารางที่ ๓ จะเห็นว่า

แบบสอบ J IIA มีค่าเฉลี่ยเท่ากับ ๕.๓๓ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๓.๙๕ ความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๓๙ ค่าความเชื่อถือได้เท่ากับ ๐.๘๐ และความคลาดเคลื่อนของการวัดเท่ากับ ๑.๖๘

แบบสอบ J IIB มีค่าเฉลี่ยเท่ากับ ๔.๕๔ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๒.๙๔ ความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๒๗ ค่าความเชื่อถือได้เท่ากับ ๐.๖๕ และความคลาดเคลื่อนของการวัดเท่ากับ ๑.๙๕

ค่าสหสัมพันธ์แบบสอบ J IIA และ J IIB มีค่าเท่ากับ ๐.๙๓ หมายความว่าแบบสอบทั้งสองมีลักษณะตามกันและค่าสหสัมพันธ์อยู่ในระดับสูง ความคลาดเคลื่อนของความแตกต่างระหว่างค่าเฉลี่ยของแบบสอบทั้งสอง เท่ากับ ๐.๒๔ อัตราส่วนวิกฤตเท่ากับ - ๒.๖๓ แสดงว่า แบบสอบคุชานาน J II แตกต่างกันอย่างมีนัยสำคัญที่ระดับ ๐.๐๕

ตารางที่ ๔ ค่าสถิติที่คำนวณของแบบสอบถาม JO II

ค่าสถิติที่คำนวณ	วิชาศุทอง		สาริถเภษัทร	
	JO III A	JO III B	JO III A	JO III B
\bar{X}	๔.๐๐	๔.๘๒	๔.๓๖	๔.๖๖
S.D.	๒.๖๐	๒.๒๖	๓.๘๒	๒.๔๖
σ_m	๐.๒๖	๐.๒๒	๐.๓๓	๐.๒๔
r_{21}	๐.๕๐	๐.๓๖	๐.๘๐	๐.๖๒
σ_{meas}	๑.๘๔	๑.๘๑	๑.๙๑	๑.๘๒
r_{XY}	๐.๕๑		๐.๕๘	
σ_{dm}	๐.๒๔		๐.๒๙	
σ_{ratio}	- ๓.๕๐		- ๑.๐๘	
	มีนัยสำคัญที่ระดับ ๐.๐๕		ไม่มีนัยสำคัญที่ระดับ ๐.๐๕	

(การคำนวณในภาคผนวก)

๑. จากตารางที่ ๔ จะเห็นว่า แบบสอบ JO III ให้นำไปทดลองสอบนักเรียนโรงเรียนวิชาศุทอง มีค่าสถิติดังต่อไปนี้

แบบสอบ JO III มีค่าเฉลี่ยเท่ากับ ๔ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๒.๖๐ ค่าความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๒๖ ค่าความเชื่อถือได้เท่ากับ ๐.๕๐ และความคลาดเคลื่อนของการวัดเท่ากับ ๑.๘๔

แบบสอบ JO IIIB ค่าเฉลี่ยเท่ากับ ๔.๘๒ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๒.๒๖ ค่าความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๒๒ ค่าความเชื่อถือได้เท่ากับ ๐.๓๖ และความคลาดเคลื่อนของการวัดเท่ากับ ๑.๘๑

ค่าสหสัมพันธ์แบบสอบ JO IIIA และ JO IIIB วิชาศุทองเท่ากับ ๐.๕๑ แสดงว่าแบบสอบทั้งสองมีลักษณะตามกันและค่าสหสัมพันธ์อยู่ในระดับกลาง ความคลาดเคลื่อนของความแตกต่างระหว่างค่าเฉลี่ยของแบบสอบทั้งสองเท่ากับ ๐.๒๔ อัตราส่วนวิกฤตเท่า

กับ -๓.๔ แสดงว่าแบบสสอบคุณาน JO III (วิชาศุทง) แตกต่างกันอย่างมีนัยสำคัญที่ระดับ ๐.๐๕

๒. แบบสสอบ JO III ที่นำไปทดลองสสอบกับนักเรียนโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตร ศาสตร์มีค่าสถิติดังนี้

แบบสสอบ JO IIIA มีค่าเฉลี่ยเท่ากับ ๘.๓๒ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๓.๘๒ ความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๓๓ ค่าความเชื่อถือได้เท่ากับ ๐.๘๐ และความคลาดเคลื่อนของการวัดเท่ากับ ๑.๓๑

แบบสสอบ JO IIIB มีค่าเฉลี่ยเท่ากับ ๘.๖๖ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๒.๘๖ ความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๒๕ ค่าความเชื่อถือได้เท่ากับ ๐.๖๒ และความคลาดเคลื่อนของการวัดเท่ากับ ๑.๘๒

ค่าสหสัมพันธ์ของแบบสสอบ JO IIIA และ JO IIIB ที่นำไปทดลองสสอบกับนักเรียนโรงเรียนสาธิตเกษตรฯ เท่ากับ ๐.๕๘ หมายความว่าข้อสสอบทั้งสองมีลักษณะตามกัน และค่าสหสัมพันธ์อยู่ในระดับปานกลาง ความคลาดเคลื่อนของความแตกต่างระหว่างค่าเฉลี่ยของแบบทดลองสสอบทั้งสองเท่ากับ ๐.๒๗ อัตราส่วนวิกฤตเท่ากับ - ๑.๐๕ แสดงว่าแบบสสอบคุณาน JO XII (สาธิตเกษตร) ไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ ๐.๐๕

การแก้าค่าความเชื่อถือได้ของแบบสสอบแต่ละฉบับที่นำไปทดลองสสอบ

เนื่องจากแบบสสอบแต่ละฉบับมี ๑๕ ข้อ ค่าความเชื่อถือได้จึงมีค่าไม่สูงเท่าที่ควร เพราะจำนวนข้อของแบบสสอบมีอิทธิพลต่อค่าความเชื่อถือได้ จึงแก้าค่าควยสูตร สเปียร์แมน บราวน์ (Spearman Brown) ได้ค่าความเชื่อถือได้ของแบบสสอบแต่ละฉบับเมื่อกำหนดให้แต่ละฉบับมี ๕๐ ข้อ ดังนี้

แบบสสอบ O IIA มีค่าความเชื่อถือได้เท่ากับ	๐.๘๑
แบบสสอบ O II B มีค่าความเชื่อถือได้เท่ากับ	๐.๘๔
แบบสสอบ J IIA มีค่าความเชื่อถือได้เท่ากับ	๐.๘๓

แบบสอบ J II B	มีค่าความเชื่อถือได้เท่ากับ	๐.๘๖
วิชาศุทของ	แบบสอบ JO III A	มีค่าความเชื่อถือได้เท่ากับ ๐.๗๗
	แบบสอบ JO III B	มีค่าความเชื่อถือได้เท่ากับ ๐.๖๕
สาธิตเกษตร	แบบสอบ JO III A	มีค่าความเชื่อถือได้เท่ากับ ๐.๕๓
	แบบสอบ JO III B	มีค่าความเชื่อถือได้เท่ากับ ๐.๘๔

(ดูการคำนวณในภาคผนวก)

เมื่อได้รับการแก้ไขค่าความเชื่อถือได้โดยใช้สูตรของสเปียร์แมน บราวน์ (Spearman Brown) ทำให้ค่าความเชื่อถือสูงอยู่ในเกณฑ์ คือมีค่าความเชื่อถือได้อยู่ระหว่าง ๐.๖๕ - ๐.๕๓

การปรับปรุงแก้ไขข้อสอบตามผลวิเคราะห์

หลังจากการวิเคราะห์รายข้อแล้วได้พิจารณาเลือกข้อสอบที่ดีกว่าจากแบบสอบคู่ขนานทั้งสองฉบับคือ ฉบับ A และฉบับ B เรียงลำดับข้อสอบที่ใช้ได้จากง่ายมาหายาก โดยคำนึงถึงค่าอำนาจจำแนกด้วย ได้ค่าเฉลี่ยระดับความยากอยู่ระหว่าง ๐.๖ - ๐.๗ และค่าอำนาจจำแนกสูงกว่า ๐.๒๐ ขึ้นไป

ตารางที่ ๕ ค่าเฉลี่ยอำนาจจำแนกและค่าเฉลี่ยความยากง่ายของแบบทดสอบสองข้อที่คัดเลือกจากผลการวิเคราะห์รายข้อแบบสอบคู่ขนาน ฉบับ A และ ฉบับ B

	แบบสอบคู่ขนาน OII ^๑	แบบสอบคู่ขนาน JII ^๒	แบบสอบคู่ขนาน JO III
ค่าเฉลี่ยอำนาจจำแนก	๐.๖๔๒	๐.๗๐๖	๐.๖๑๘
ค่าเฉลี่ยความยากง่าย	๐.๖๗๖	๐.๖๔๒	๐.๖๔๐

(ดูรายละเอียดในภาคผนวก ข.)

จากตารางที่ ๕ จะเห็นว่า

แบบสอบ O II	มีค่าเฉลี่ยระดับความยากเท่ากับ	๐.๖๘
	มีค่าเฉลี่ยอำนาจจำแนกเท่ากับ	๐.๖๔
แบบสอบ J II	มีค่าเฉลี่ยระดับความยากเท่ากับ	๐.๖๔

มีค่าเฉลี่ยอำนาจจำแนก เท่ากับ ๐.๗๑

แบบสอบ JO III มีค่าเฉลี่ยระดับความยากเท่ากับ ๐.๖๔

มีค่าเฉลี่ยอำนาจจำแนกเท่ากับ ๐.๖๒

ค่าเฉลี่ยระดับความยากและค่าเฉลี่ยอำนาจจำแนกของแบบสอบทั้ง ๓ ชุดที่คัดเลือกแล้ว ผลเป็นที่น่าพอใจ จึงนำแบบสอบทั้ง ๓ ชุดนี้ไปทำการสอบกับตัวอย่างประชากรจริง

๒. การวิเคราะห์แบบสอบที่ใช้ทดสอบจริง

นำแบบสอบที่ปรับปรุงแล้วไปทดสอบกับตัวอย่างประชากรจริง แล้วนำคะแนนของนักเรียนแต่ละคนในกลุ่มที่เรียน Oxford English Course for Thailand และกลุ่มที่เรียน JILAP มาเรียงคะแนนตามลำดับจากมากไปหาน้อย คำนวณโดยใช้เครื่องคอมพิวเตอร์ชนิด IBM ๓๕๖/๑๔๕ ของสถาบันเทคโนโลยีแห่งเอเชีย (Asian Institute of Technology) วิเคราะห์รายชื่อของแบบสอบแต่ละชุดโดยใช้เทคนิค ๒๗ เมอร์เซนต์ ของจำนวนข้อสอบแต่ละกลุ่มเป็นกลุ่มสูงและกลุ่มต่ำ หาร้อยละของกลุ่มสูงและกลุ่มต่ำที่ตอบถูกแต่ละข้อ เพื่อหาระดับความยาก (p) และอำนาจจำแนก (r) ของแบบสอบทั้ง ๓ ฉบับ ปรากฏผลดังนี้

ตารางที่ ๖ แสดงช่วงความยากและค่าอำนาจจำแนกของแบบสอบที่ใช้ทดสอบจริง

แบบสอบ	O II	J II	JO III
ช่วงความยากอยู่ระหว่าง	๐.๑๘ - ๐.๘๘	๐.๔๐ - ๐.๕๓	๐.๒๘ - ๐.๘๗
ค่าอำนาจจำแนกอยู่ระหว่าง	๐.๑๔ - ๐.๕๕	๐.๔๔ - ๐.๘๑	๐.๑๗ - ๐.๘๔

จากตารางที่ ๖ จะเห็นว่าแบบสอบ O II มีบางข้อยากเกินไป (ระดับความยากต่ำกว่า .๒๐) และแบบสอบทั้ง ๓ ชุดมีบางข้อง่ายเกินไป (ระดับความยากสูงกว่า .๘๐ ทั้งสามชุด) ค่าอำนาจจำแนกของแบบสอบ O II และแบบสอบ JO III มีบางข้อค่าอำนาจจำแนกต่ำ (ต่ำกว่า .๒๐) สำหรับแบบสอบ J II ค่าอำนาจจำแนกสูงเป็นที่น่าพอใจ (มากกว่า .๒๐ ขึ้นไป)

จำนวนข้อสอบที่มีความยากและค่าอำนาจจำแนกอยู่ในระดับมาตรฐานของแบบสอบแต่ละชุดจะพิจารณาได้จากแผนภาพประกอบต่อไปนี้

ตารางที่ ๗ ค่าสถิติที่คำนวณได้ของแบบสอบในการทดสอบจริง

ค่าสถิติที่คำนวณ	O II	J II	JO III
\bar{X}	๗.๕๔	๑๐.๕๑	๘.๖๕
S.D.	๒.๕๗	๓.๑๓	๓.๐๗
σ_m	๐.๑๘	๐.๑๔	๐.๑๘
r_{21}	๐.๖๒	๐.๗๓	๐.๖๖
$\sigma_{mcas.}$	๑.๘๔	๑.๖๓	๑.๘๐

จากตารางที่ ๗ จะเห็นว่า

แบบสอบ O II มีค่าเฉลี่ยเท่ากับ ๗.๕๔ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๒.๕๗ ความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๑๘ ค่าความเชื่อถือได้ของแบบสอบโดยใช้สูตรของคูเคอร์ริชาร์ดสัน สูตรที่ ๒๑ (Kuder Richardson Formula 21) เท่ากับ ๐.๖๒ และความคลาดเคลื่อนมาตรฐานในการวัดเท่ากับ ๑.๘๔

แบบสอบ J II มีค่าเฉลี่ยเท่ากับ ๑๐.๕๑ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๓.๑๓ ความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๑๔ ค่าความเชื่อถือได้ของแบบสอบโดยใช้สูตรคูเคอร์ริชาร์ดสัน สูตรที่ ๒๑ เท่ากับ ๐.๗๓ และความคลาดเคลื่อนมาตรฐานในการวัดเท่ากับ ๑.๖๓

แบบสอบ JO III มีค่าเฉลี่ยเท่ากับ ๘.๖๕ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ ๓.๐๗ ความคลาดเคลื่อนของค่าเฉลี่ยเท่ากับ ๐.๑๘ ค่าความเชื่อถือได้ของแบบสอบโดยใช้สูตรคูเคอร์ริชาร์ดสัน สูตรที่ ๒๑ เท่ากับ ๐.๖๖ และความคลาดเคลื่อนมาตรฐานในการวัดเท่ากับ ๑.๘๐

เนื่องจากแบบสอบแต่ละฉบับมี ๑๕ ข้อ ค่าความเชื่อถือได้จึงมีค่าไม่สูงเท่าที่ควร จึงแก้ค่าความเชื่อถือได้ของแบบสอบที่นำไปทดสอบจริงแต่ละฉบับโดยใช้สูตร สเปียร์แมน บราวน์ (Spearman Brown) ได้ค่าความเชื่อถือได้ของแบบสอบแต่ละฉบับเมื่อกำหนดให้แต่ละฉบับมี ๕๐ ข้อ ดังนี้

แบบสอบ O II	มีค่าความเชื่อถือได้เท่ากับ	๐.๘๔๓
แบบสอบ J II	มีค่าความเชื่อถือได้เท่ากับ	๐.๘๘๕
แบบสอบ JO III	มีค่าความเชื่อถือได้เท่ากับ	๐.๘๖๓

(ผลการคำนวณในภาคผนวก ค.)

เมื่อได้รับการแก้ไขค่าความเชื่อโดยใช้สูตรของสเปียร์แมน บราวน์ แล้วทำให้ค่าความเชื่อถือสูงอยู่ในเกณฑ์ดีมาก คือมีค่าความเชื่อถือได้อยู่ระหว่าง ๐.๘๔๓ - ๐.๘๘๕

เนื่องจากค่าอำนาจจำแนกและความยากง่ายของแบบสอบที่นำไปใช้ทดสอบจริงบางข้อต่ำกว่าระดับมาตรฐานและบางข้อสูงกว่าระดับมาตรฐาน ในการพิจารณาแบบสอบทั้งฉบับจึงต้องพิจารณาจากค่าเฉลี่ยอำนาจจำแนกและค่าเฉลี่ยความยากง่ายของแบบสอบแต่ละฉบับเป็นเกณฑ์

ตารางที่ ๘ แสดงค่าเฉลี่ยอำนาจจำแนกและค่าเฉลี่ยความยากง่ายของแบบสอบที่นำไปใช้ทดสอบจริง

ค่าเฉลี่ย	แบบสอบ O II	แบบสอบ J II	แบบสอบ JO III
อำนาจจำแนก (\bar{r}_{bis})	๐.๕๕	๐.๖๗	๐.๕๘
ความยากง่าย (p)	๐.๕๓	๐.๗๐	๐.๖๓

(ดูรายละเอียดในภาคผนวก ข.)

จากตารางที่ ๘ จะเห็นว่า

แบบสอบ O II	ค่าเฉลี่ยระดับความยากเท่ากับ	๐.๕๓
	ค่าเฉลี่ยอำนาจจำแนกเท่ากับ	๐.๕๕
แบบสอบ J II	ค่าเฉลี่ยระดับความยากเท่ากับ	๐.๗๐
	ค่าเฉลี่ยอำนาจจำแนกเท่ากับ	๐.๖๗
แบบสอบ JO III	ค่าเฉลี่ยระดับความยากเท่ากับ	๐.๖๓
	ค่าเฉลี่ยอำนาจจำแนกเท่ากับ	๐.๕๘

ค่าเฉลี่ยระดับความยากและค่าเฉลี่ยอำนาจจำแนกของแบบสอบที่นำไปทดสอบจริงทั้ง ๓ ชุด

เป็นที่น่าพอใจ

๓. การเปรียบเทียบแบบทดสอบและแบบสอบที่ใช้ทดสอบจริง

ผลการทดลองสอบและการวิเคราะห์รายข้อนับว่ามีประโยชน์อย่างมากในการสร้างแบบสอบเพื่อใช้ทดสอบจริง เพราะถึงแม้เนื้อหาศัพท์ และโครงสร้างไวยากรณ์จะเหมือนกันทุกประการ แต่ตัวเลือกที่เปลี่ยนไปทำให้ผลการวิเคราะห์แตกต่างกัน ผลการวิเคราะห์แบบทดสอบช่วยให้ผู้วิจัยสามารถคัดเลือกข้อสอบที่ดีกว่าจากฉบับใดฉบับหนึ่ง เพราะทั้งสองฉบับเป็นแบบสอบคุณภาพ จากตารางที่ ๕ จะเห็นว่า แบบสอบที่ใช้ทดสอบจริงมีพัฒนาการจากแบบทดสอบ

จากตารางที่ ๕ จะเห็นว่า

แบบสอบ O IIA และ O IIB ที่นำไปทดลองสอบมีค่าเฉลี่ยอำนาจจำแนกแต่ละฉบับใกล้เคียงกัน แต่เมื่อคัดเลือกข้อสอบที่ดีจากทั้ง ๒ ฉบับแล้วได้แบบสอบที่มีค่าอำนาจจำแนกสูงขึ้น แต่เมื่อนำไปสอบจริงค่าอำนาจจำแนกต่ำลงกว่าแบบสอบที่คัดเลือกไว้ ทั้งนี้เพราะแบบสอบที่คัดเลือกไว้มาจากข้อที่มีคะแนนสูงในแต่ละฉบับ ค่าจึงมากกว่าแบบสอบจริง ส่วนค่าเฉลี่ยระดับความยากของแบบสอบ O IIA และ O IIB มีค่าเท่ากันจึงมีความยากง่ายพอกัน และอยู่ในระดับปานกลาง เมื่อคัดเลือกข้อสอบที่ดีจากทั้งสองฉบับแล้วได้แบบสอบที่มีระดับความยากง่ายเพิ่มขึ้นเพียงเล็กน้อย แต่เมื่อนำไปสอบจริง ค่าระดับความง่ายกลับลดน้อยลง แสดงว่าข้อสอบยากขึ้น แต่ก็ยังอยู่ในเกณฑ์คือไม่ยากหรือง่ายเกินไป ทั้งนี้อาจเป็นเพราะกลุ่มตัวอย่างประชากรที่ทดลองสอบอยู่ในกรุงเทพฯ จึงได้เปรียบกว่ากลุ่มตัวอย่างประชากรจริงในจังหวัดขอนแก่นในด้านการเรียน การสอน และระดับความรู้

แบบสอบ J IIA ที่นำไปทดลองสอบมีค่าเฉลี่ยอำนาจจำแนกสูงกว่า แบบสอบ J IIB เมื่อคัดเลือกข้อที่ดีจากแบบสอบทั้งสองฉบับ ได้แบบสอบที่มีค่าอำนาจจำแนกสูงกว่าเดิม แต่เมื่อนำไปสอบจริง ค่าอำนาจจำแนกต่ำลงกว่า แบบสอบที่คัดเลือกไว้ แต่ค่าอำนาจจำแนกยังคง เท่ากับค่าอำนาจจำแนกของแบบสอบ J IIA และยังคงอยู่ในเกณฑ์ที่มีอำนาจจำแนกค่อนข้างสูง ส่วนค่าเฉลี่ยระดับความยากของแบบสอบ J IIA น้อยกว่า J IIB แสดงว่า J IIA ยากกว่า J IIB เมื่อคัดเลือกข้อที่ดีจากแบบสอบทั้งสอง

ตารางที่ ๕ เปรียบเทียบค่าเฉลี่ยอำนาจจำแนกและระดับความยากของแบบทดสอบคุณานาน
แบบสอบที่คัดเลือกจากผลการวิเคราะห์แบบทดสอบสอบคุณานาน และแบบสอบ
ที่ใช้ทดสอบจริง

แบบทดสอบสอบคุณานาน								
ค่าเฉลี่ย	OIIA	OIIB	IIIA	JIIB	วิชาศุทอง JOIIIA	วิชาศุทอง JOIIIB	สาริตเกษทร JOIIIA	สาริตเกษทร JOIIIB
อำนาจจำแนก	๐.๕๓	๐.๕๕	๐.๖๗	๐.๖๐	๐.๕๔	๐.๔๗	๐.๖๕	๐.๕๖
ระดับความยาก	๐.๖๗	๐.๖๗	๐.๖๔	๐.๖๘	๐.๖๑	๐.๖๔	๐.๕๖	๐.๖๐
แบบสอบที่คัดเลือกจากผลการวิเคราะห์แบบทดสอบ								
ค่าเฉลี่ย	O II	J II	JO III					
อำนาจจำแนก	๐.๖๔	๐.๗๑	๐.๖๒					
ระดับความยาก	๐.๖๘	๐.๖๔	๐.๖๔					
แบบสอบที่ใช้ทดสอบจริง								
ค่าเฉลี่ย	O II	J II	JO III					
อำนาจจำแนก	๐.๕๕	๐.๖๗	๐.๕๘					
ระดับความยาก	๐.๕๓	๐.๗๐	๐.๖๓					

(กระจายละเอียดภาคผนวก ข.)

จุฬาลงกรณ์มหาวิทยาลัย



แล้วได้แบบสอบที่มีค่าเฉลี่ยความยากง่ายของแบบสอบ JO III เท่ากับแบบสอบ JO IIIA แต่เมื่อนำไปสอบจริงการระดับความยากง่ายกลับเพิ่มขึ้น แสดงว่าข้อสอบ JO III ที่นำไปทดสอบจริงง่ายขึ้น ทั้งนี้อาจเป็นเพราะว่าตัวอย่างประชากรจริงได้เรียนครบตามหลักสูตร แต่ในขณะที่ทำการทดลองสอบกลุ่มตัวอย่างประชากร ทดลองสอบยังเรียนไม่ครบหลักสูตร

แบบสอบ JO IIIA ที่นำไปทดลองสอบกับนักเรียนโรงเรียนวัดธาตุทองมีค่าเฉลี่ยอำนาจจำแนกมากกว่า แบบสอบ JO III B แสดงว่าแบบสอบ JO III A เป็นแบบสอบที่ดีกว่า แต่ทั้งสองฉบับมีค่าอำนาจจำแนกอยู่ในเกณฑ์มาตรฐาน สำหรับแบบสอบ JO III A และ B ที่นำไปทดลองสอบกับนักเรียนโรงเรียนสาธิตเกษตรฯ มีค่าเฉลี่ยอำนาจจำแนกทั้งสองฉบับสูงกว่าวัดธาตุทอง เมื่อคัดเลือกข้อที่ดีจากทั้ง ๔ ฉบับแล้ว ได้แบบสอบที่มีค่าอำนาจจำแนกสูงขึ้นเป็นที่พอใจ แต่เมื่อนำไปสอบจริงค่าอำนาจจำแนกต่ำลงเพียงเล็กน้อย จึงเป็นแบบสอบที่มีค่าอำนาจจำแนกอยู่ในเกณฑ์

แบบสอบ JO III A และ JO III B ที่นำไปทดลองสอบกับนักเรียนโรงเรียนวัดธาตุทองมีค่าเฉลี่ยระดับความยากใกล้เคียงกัน และ ค่าเฉลี่ยระดับความยากอยู่ในเกณฑ์มาตรฐาน ซึ่งสูงกว่าค่าเฉลี่ยแบบสอบ JO III A และ JO III B ที่นำไปทดลองสอบกับนักเรียนโรงเรียนสาธิตเกษตรฯ เล็กน้อย แสดงว่านักเรียนโรงเรียนสาธิตเกษตรฯ ทำแบบสอบได้คะแนนน้อยกว่านักเรียนโรงเรียนวัดธาตุทองเล็กน้อย เมื่อคัดเลือกข้อที่ดีจากแบบสอบทั้งสองฉบับแล้วปรากฏว่าได้แบบสอบที่มีค่าเฉลี่ยระดับความยากอยู่ในเกณฑ์มาตรฐาน และมีค่าเท่ากับค่าเฉลี่ยระดับความยากของแบบสอบ JO III B ที่ทำการทดลองสอบกับนักเรียนวัดธาตุทอง แต่เมื่อนำไปสอบจริงค่าเฉลี่ยระดับความยากลดลงเพียงเล็กน้อยหรืออาจกล่าวได้ว่าไม่แตกต่างกันเลย

จากตารางที่ ๕ อาจสรุปได้ว่า ค่าเฉลี่ยอำนาจจำแนกและค่าเฉลี่ยระดับความยากของแบบสอบทุกฉบับ ทุกชุด ทั้งแบบทดลองสอบ แบบสอบที่คัดเลือกแล้ว และแบบสอบจริง อยู่ในระดับมาตรฐานทั้งสิ้น เป็นที่น่าพอใจ ความแตกต่างมีน้อยมาก จึงเป็นแบบสอบที่ดีตั้งแต่เริ่มทดลองสอบจนกระทั่งสอบจริง

ตารางที่ ๑๐ เปรียบเทียบความเชื่อถือได้ของแบบทดสอบที่ทดลองสอบและแบบสอบจริงทั้งที่คำนวณจากสูตรคูเคอร์ วิซาร์คสัน สูตรที่ ๒๑ และค่าแก้ของสเปียร์แมน บราวน์ และค่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย

	Pretest		Post test
	Form A	Form B	
แบบทดสอบ O II			แบบทดสอบ O II
Kr ₂₁	๐.๕๕๖๔	๐.๖๑	๐.๖๑๓
S.E. _m	๐.๒๖๒	๐.๒๓๔	๑.๘๓๓
r _{tt}	๐.๘๐๓	๐.๘๓๔	๐.๘๔๓
แบบทดสอบ J II			แบบทดสอบ J II
Kr ₂₁	๐.๘๐	๐.๖๕	๐.๗๒๘
S.E. _m	๐.๓๔๔	๐.๒๓	๑.๖๓๓
r _{tt}	๐.๘๓	๐.๘๖๑	๐.๘๘๔
แบบทดสอบ J O III (วัชรธาตุทอง)			แบบทดสอบ JO III
Kr ₂₁	๐.๕๐	๐.๓๖	๐.๖๕๕
S.E. _m	๐.๒๕๓	๐.๒๒๔	๑.๘๐๑
r _{tt}	๐.๗๖๔	๐.๖๕๒	๐.๘๖๓
แบบทดสอบ J O III (สาริิตเกษตรา)			
Kr ₂₁	๐.๘๐	๐.๖๒	
S.E. _m	๐.๓๒๖	๐.๒๕๓	
r _{tt}	๐.๘๓	๐.๘๔๔	

จากตารางที่ ๑๐ จะเห็นว่าค่าความเชื่อถือได้ของแบบสอบจริงหลังจากใช้ค่า
 แก้ของสเปียร์แมน บราวน์ นั้นมีดังนี้ แบบสอบ O II ที่นำไปทดสอบกับตัวอย่างประชากร
 จริงมีค่าสูงกว่าแบบสอบ O IIA และ O IIB ที่นำไปทดสอบ แสดงว่ามีความเชื่อ
 ถือได้มากกว่าแบบสอบทั้งสองฉบับที่นำไปทดสอบ สำหรับแบบสอบ J II ที่นำไปทดสอบ
 กับตัวอย่างประชากรจริง มีค่าความเชื่อถือได้ต่ำกว่าแบบสอบ J IIA แต่สูงกว่าแบบ
 สอบ J IIB ที่นำไปทดสอบ และแบบสอบ JO III ที่นำไปทดสอบกับตัวอย่างประชากร
 จริงมีค่าความเชื่อถือได้สูงกว่าแบบสอบ JO IIIA , JO IIIB ที่นำไปทดสอบกับนัก
 เรียนวิชาคหัตถ์ และสูงกว่าแบบสอบ JO IIIB ที่นำไปทดสอบกับนักเรียนสาธิตแห่ง
 มหาวิทยาลัยเกษตรศาสตร์ แต่ต่ำกว่าแบบสอบ JO IIIA ที่นำไปทดสอบกับนักเรียนสาธิต
 แห่งมหาวิทยาลัยเกษตรศาสตร์

ส่วนค่าความคลาดเคลื่อนมาตรฐานของแบบสอบจริงแต่ละชุดมีค่ามากกว่าความคลาด
 เคลื่อนของแบบสอบชุดเดียวกันที่นำไปทดสอบ

ศูนย์วิจัยทรัพยากร
 จุฬาลงกรณ์มหาวิทยาลัย