



บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการเสนอการวิเคราะห์ข้อมูล ผู้วิจัยจะเสนอผลการวิเคราะห์เป็น 2 ตอน คือ ตอนที่ 1 ผลการวิเคราะห์การพัฒนาเครื่องมือ เพื่อปรับปรุงข้อกระทงรายชื่อ และทั้งฉบับ สำหรับจัดต้นฉบับใหม่

ตอนที่ 2 ผลการวิจัย เพื่อวิเคราะห์ผลแบบสอบตามทฤษฎีคลาสสิก ความเป็นมิติเดียวของแบบสอบ วิเคราะห์รายชื่อตามทฤษฎีการตอบสนองข้อกระทงเพื่อเลือกข้อกระทง สร้างคลังข้อสอบ การสุ่มข้อกระทงเพื่อหาเกณฑ์เลือกข้อกระทงที่เหมาะสมกับความสามารถ เมื่อกำหนดคะแนนโดเมนและความยาวแบบสอบ 20 และ 30 ข้อตามลำดับ การหาคะแนนจุดตัด ความเที่ยงตามแนวทฤษฎีคลาสสิก และทฤษฎีการตอบสนองข้อกระทง และความน่าจะเป็นในการจำแนกผิดพลาด

ตอนที่ 1 ผลการวิเคราะห์การพัฒนาเครื่องมือ

จากการสร้างแบบสอบ แยกเป็น 3 ฉบับทดลองกับกลุ่มตัวอย่างที่ใช้ทดลองเครื่องมือ ผลจากการสอบวิเคราะห์ด้วยคอมพิวเตอร์ โปรแกรมวิเคราะห์ข้อกระทง (Item Analysis) ปรากฏผล ดังตารางที่ 6

ตารางที่ 6 ผลการหาค่าความยาก ค่าอำนาจจำแนกและค่าความเที่ยงของแบบสอบ

แบบสอบ ฉบับที่	จุด ประสงค์	จำนวน ข้อ	จำนวน คน	พิสัย ค่าความยาก p	ค่าความ ยากเฉลี่ย	พิสัยค่าอำนาจ จำแนก r	ค่าอำนาจ จำแนกเฉลี่ย	ค่า ความเที่ยง α_{21}
1	1, 2	60	280	.239-.918	.579	.112-.590	.412	.907
2	3	40	241	.519-.979	.778	.034-.666	.461	.899
3	4	20	224	.384-.782	.586	.363-.631	.525	.821

จากผลการวิเคราะห์การสร้างเครื่องมือ ให้ผล ดังนี้

แบบสอบฉบับที่ 1 เมื่อวิเคราะห์รายข้อ ปรากฏว่ามีข้อที่ต้องปรับปรุงอยู่ 1 ข้อคือ ข้อที่ 60 เพราะมีค่าอำนาจจำแนกต่ำกว่า .2 คือ .112 แต่เนื่องจากเป็นข้อสุดท้ายของแบบสอบ นักเรียนอาจจะทำไม่ทัน จึงตอบข้อที่ 60 ได้ไม่ถี่ ผู้วิจัยจึงยังคงไว้ และเพื่อดูว่าจะมีผลต่อการวิจัยอย่างไร เมื่อนำไปวิเคราะห์ด้วยโปรแกรมโลจิสติก 5 ปรากฏว่าค่าความยากเฉลี่ยอยู่ในเกณฑ์ปานกลางเล็กน้อย ค่าอำนาจจำแนกในแต่ละข้อทั้งหมดเฉลี่ยแล้วอยู่ในเกณฑ์ดี ส่วนค่าความเที่ยงของแบบสอบอยู่ในเกณฑ์สูง

แบบสอบฉบับที่ 2 ผลการวิเคราะห์แบบสอบมีข้อที่ต้องปรับปรุงแก้ไข 1 ข้อ คือ ข้อที่ 1 เพราะมีค่าอำนาจจำแนกน้อยกว่า 0.2 (0.034) ซึ่งจำแนกได้น้อยมาก ค่าความยากของแบบสอบเฉลี่ยอยู่ในเกณฑ์ค่อนข้างง่ายมาก ค่าอำนาจจำแนกเฉลี่ยอยู่ในเกณฑ์ดี ส่วนค่าความเที่ยงของแบบสอบอยู่ในเกณฑ์สูง

แบบสอบฉบับที่ 3 ไม่ปรากฏว่ามีข้อกระทงที่ต้องปรับปรุงแก้ไข ค่าความยากเฉลี่ยอยู่ในเกณฑ์ปานกลาง ค่าอำนาจจำแนกเฉลี่ยอยู่ในเกณฑ์ดี ส่วนค่าความเที่ยงของแบบสอบอยู่ในเกณฑ์สูง

จากผลดังกล่าว แสดงว่า แบบสอบทั้ง 3 ฉบับ มีคุณภาพอยู่ในเกณฑ์ สามารถใช้ เป็นเครื่องมือในการวิจัยได้

จากนั้นนำปัญหาจากการบริหารการสอบ และผลการวิเคราะห์ข้อกระทงที่ได้รับการ ปรับปรุงในการทดลองขั้นต้น มาปรับปรุงและจัดต้นฉบับใหม่เป็น 4 ฉบับ ๗ ละ 30 ข้อ เพื่อให้ได้จำนวนข้อกระทงเหมาะสมกับเวลา 1 คาบ 50 นาทีในการวิจัย

ตอนที่ 2 ผลการวิจัย

ข้อมูลที่ได้จากกลุ่มตัวอย่างที่ใช้ในการวิจัยเป็นคะแนนที่ได้จากแบบสอบจำนวน 4 ฉบับ ๗ ละ 30 ข้อ กลุ่มตัวอย่าง 1,005 คน จากโรงเรียน 11 โรงเรียนและจำนวนห้องเรียน 22 ห้องเรียน โดยที่นักเรียนที่เป็นกลุ่มตัวอย่างจะต้องทำแบบสอบทั้ง 4 ฉบับ ผู้วิจัยได้ รวบรวมแบบสอบทั้ง 4 ฉบับ เป็นแบบสอบชุดเดียวจำนวน 120 ข้อ นำมาวิเคราะห์ด้วย คอมพิวเตอร์ โปรแกรมโลจิสติก 5 เพื่อสร้างเป็นคลังข้อสอบ แต่ก่อนจะนำไปวิเคราะห์ด้วย โปรแกรมโลจิสติก 5 ผู้วิจัยต้องตรวจหาค่าสถิติทั้ง 120 ข้อ ตามแนวทฤษฎีคลาสสิก และ เพื่อตรวจสอบว่าแบบสอบวัดความสามารถด้านเดียวหรือมิติเดียว (Unidimension) จึงมี ลำดับขั้นการวิเคราะห์ ดังนี้

1. การวิเคราะห์แบบสอบตามแนวทฤษฎีคลาสสิก โดยวิเคราะห์แบบสอบ 120 ข้อนี้ด้วย คอมพิวเตอร์ ใช้โปรแกรมสำเร็จรูป SPSSX วิเคราะห์ค่าสถิติต่าง ๆ ตามแนว คลาสสิก ทั้งฉบับที่จะนำมาเป็นคลังข้อสอบ ได้ผลดังตารางที่ 7

ตารางที่ 7 ค่าสถิติพื้นฐานของแบบสอบ 4 ฉบับ รวม 120 ข้อ

ค่าเฉลี่ย \bar{x}	ความแปรปรวน s^2	ส่วนเบี่ยงเบนมาตรฐาน s	ค่าความเที่ยง α_{21}
78.2358	528.4035	22.9870	.9605

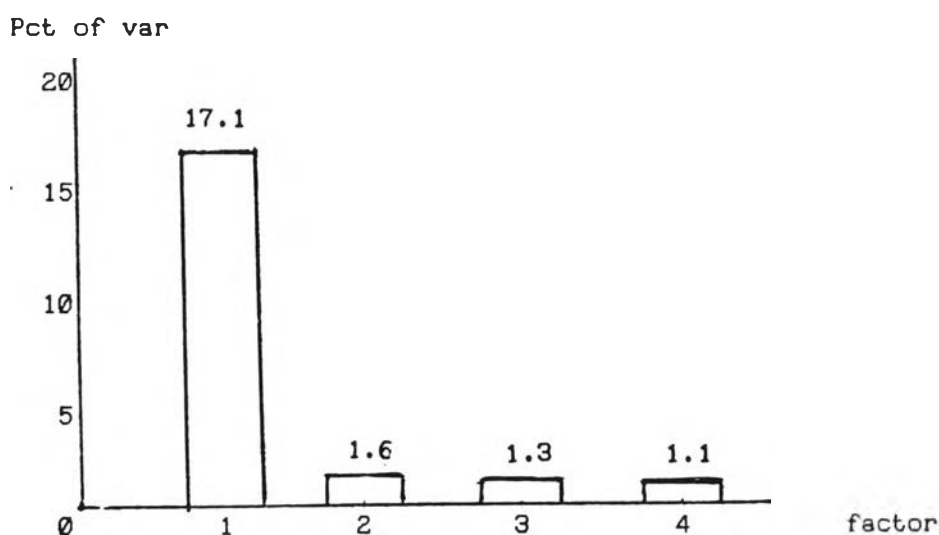
จากตารางพบว่า คะแนนที่ได้จากข้อกระทง 120 ข้อ มีค่าความแปรปรวน 528.4035 แสดงว่า คะแนนสอบมีการกระจายของคะแนนมาก ให้ค่าเฉลี่ยของคะแนนสอบมากกว่าครึ่งหนึ่งของแบบสอบ คือ 78 คะแนน แสดงว่าข้อกระทงในแบบสอบค่อนข้างง่าย และค่าความเที่ยง (α_{21}) คือ .9605 แสดงว่าแบบสอบที่จะสร้างเป็นคลังข้อสอบมีคุณภาพดี

2. การวิเคราะห์ความเป็นมิติเดียวของแบบสอบ ใช้วิธีวิเคราะห์ตัวประกอบ (Factor Analysis) โดยใช้โปรแกรมสำเร็จรูป SPSSX ให้ผล ดังตารางที่ 8

ตารางที่ 8 ผลการวิเคราะห์ตัวประกอบของข้อระทง 120 ข้อ

Factor	SS	Loadings	Pct of var	Cum Pct	Factor	SS	Loadings	Pct of var	Cum Pct
1	20.54681		17.1	17.1	21	.20845		.2	26.9
2	1.93314		1.6	18.7	22	.18790		.2	27.1
3	1.55399		1.3	20.0	23	.17471		.1	27.2
4	1.31052		1.1	21.1	24	.17510		.1	27.3
5	.95160		.8	21.9	25	.16464		.1	27.4
6	.80724		.7	22.6	26	.14905		.1	27.5
7	.58511		.5	23.1	27	.14540		.1	27.6
8	.52676		.4	23.5	28	.13296		.1	27.7
9	.48416		.4	23.9	29	.12620		.1	27.8
10	.42172		.4	24.3	30	.12240		.1	27.9
11	.40960		.3	24.6	31	.10562		.1	28.0
12	.38378		.3	24.9					
13	.32901		.3	25.2					
14	.30472		.3	25.5					
15	.28996		.2	25.7					
16	.27277		.2	25.9					
17	.25115		.2	26.1					
18	.24166		.2	26.3					
19	.24149		.2	26.5					
20	.21310		.2	26.7					

จากตารางปรากฏว่า ตัวประกอบที่ 1 มี Sum Square Loading มากที่สุด คือ 20.54681 ซึ่งมีค่ามากกว่าตัวประกอบที่ 2 อยู่มาก และตัวประกอบที่ 2 มีค่าไม่ต่าง จากตัวประกอบอื่น ๆ มากนัก ซึ่งจะเห็นได้ชัดดังภาพที่ 2



แผนภาพที่ 2 ค่าเปอร์เซ็นต์ของความแปรปรวน (Percentage of variance) ของตัวประกอบต่าง ๆ

จากตารางและแผนภาพ ผลการวิเคราะห์แสดงว่า แบบสอบทั้ง 120 ข้อ เป็นแบบสอบที่วัดความสามารถด้านเดียว หรือมิติเดียว (Unidimension) สอดคล้องกับ ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อกระทง (Lord 1980: 21)

3. การวิเคราะห์รายข้อเพื่อหาค่าพารามิเตอร์ที่เหมาะสมในการคัดเลือก ข้อกระทงเป็นคลังข้อสอบตามวิธี IRT ด้วยคอมพิวเตอร์ โปรแกรมโลจิสติก 5 ที่ใช้ พารามิเตอร์ 3 ตัว คือ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) การวิเคราะห์ข้อกระทงทั้ง 120 ข้อ พบว่า

ค่าอำนาจจำแนก (a)	มีค่าอยู่ระหว่าง	.26127	ถึง	1.56509
ค่าความยาก (b)	มีค่าอยู่ระหว่าง	-2.27647	ถึง	1.19197
ค่าการเดา (c)	มีค่าอยู่ระหว่าง	.00000	ถึง	.48015

เมื่อพิจารณาข้อที่เหมาะสมกับโมเดล โดยพิจารณาจากเกณฑ์ ดังนี้ ค่าอำนาจจำแนก (a) มากกว่า 0.3 ค่าความยาก (b) อยู่ระหว่าง -2 ถึง 2 และค่าการเดา (c) น้อยกว่า 0.3 (Urry 1977: 191) แล้ว ปรากฏว่ามีข้อที่ไม่เหมาะสมกับเกณฑ์ที่กำหนด 16 ข้อ คือ ข้อที่ 13, 25, 54, 57, 60, 69, 70, 83, 92, 97, 98, 100, 114, 119 และ 120 จึงเหลือข้อกระทงที่จะใช้เป็นคลังข้อสอบ 104 ข้อ (ดูรายละเอียดในภาคผนวก ข) แยกเป็นจำนวนข้อตามจุดประสงค์ที่ต้องการวัด ดังนี้

จุดประสงค์	จำนวนข้อ
1	38
2	17
3	32
4	17
รวม	104

4. การเลือกข้อกระทงเพื่อประมาณความสามารถ (θ_0) ที่สัมพันธ์กับคะแนนโตเมน โดยทำการสุ่มข้อกระทงอย่างง่ายในคลังข้อสอบ 104 ข้อ ให้ได้จำนวนความยาวแบบสอบ 20 และ 30 ข้อ ตามลำดับ ปรากฏผลดังตารางที่ 9

ตารางที่ 9 จำนวนและหมายเลขข้อกระทงในคลังข้อสอบ (104 ข้อ) ที่สุ่มได้จำนวน 20 และ 30 ข้อ

จุดประสงค์	ฉบับที่ 1		ฉบับที่ 2	
	ความยาวแบบสอบ 20 ข้อ		ความยาวแบบสอบ 30 ข้อ	
	จำนวนข้อ	ข้อ	จำนวนข้อ	ข้อ
1	7	3, 18, 19, 21, 22, 29, 33	10	2, 3, 5, 10, 21, 27, 30, 32, 36, 40
2	3	43, 48, 58	5	45, 46, 48, 52, 55
3	7	62, 66, 71, 77, 86, 89, 94	10	63, 65, 72, 77, 79, 80, 81, 89, 94, 95
4	3	103, 106, 112	5	102, 104, 111, 113, 117
	20		30	

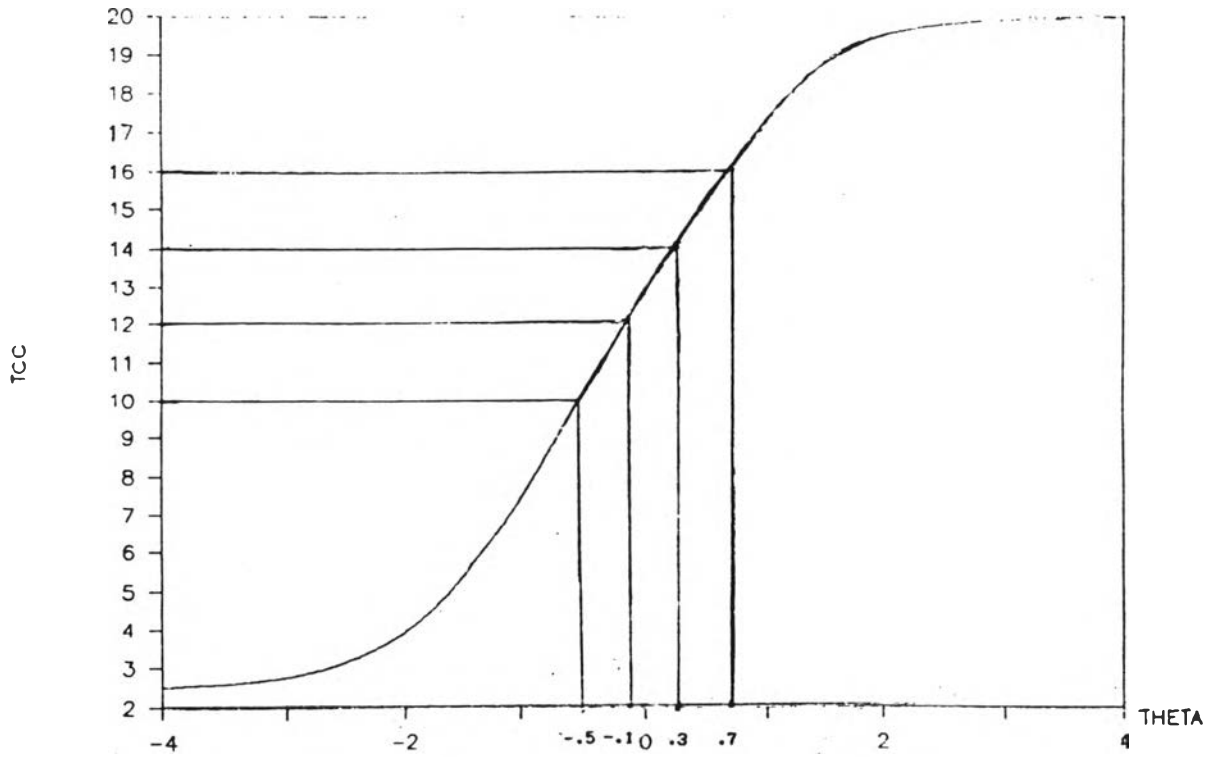
จากผลการเลือกข้อกระทงเพื่อประมาณค่าความสามารถแล้ว นำจำนวนข้อกระทงที่ได้เป็นแบบสอบความยาว 20 และ 30 ข้อ ประมาณค่าความสามารถและสร้างเกณฑ์เพื่อหาคะแนนจุดตัด ดังนี้

4.1 การประมาณค่าความสามารถเพื่อใช้เป็นเกณฑ์ในการเลือกข้อกระทงที่ใช้วัดความสามารถระดับที่ต้องการในการหาคะแนนจุดตัด โดยนำแบบสอบที่สุ่มมา 20 ข้อ และ 30 ข้อ ตามลำดับมาคำนวณค่าต่าง ๆ (ดูจากผลการคำนวณด้วยคอมพิวเตอร์ภาษาฟอร์แทรน) จากการวิเคราะห์ได้ผลดังตารางที่ 10

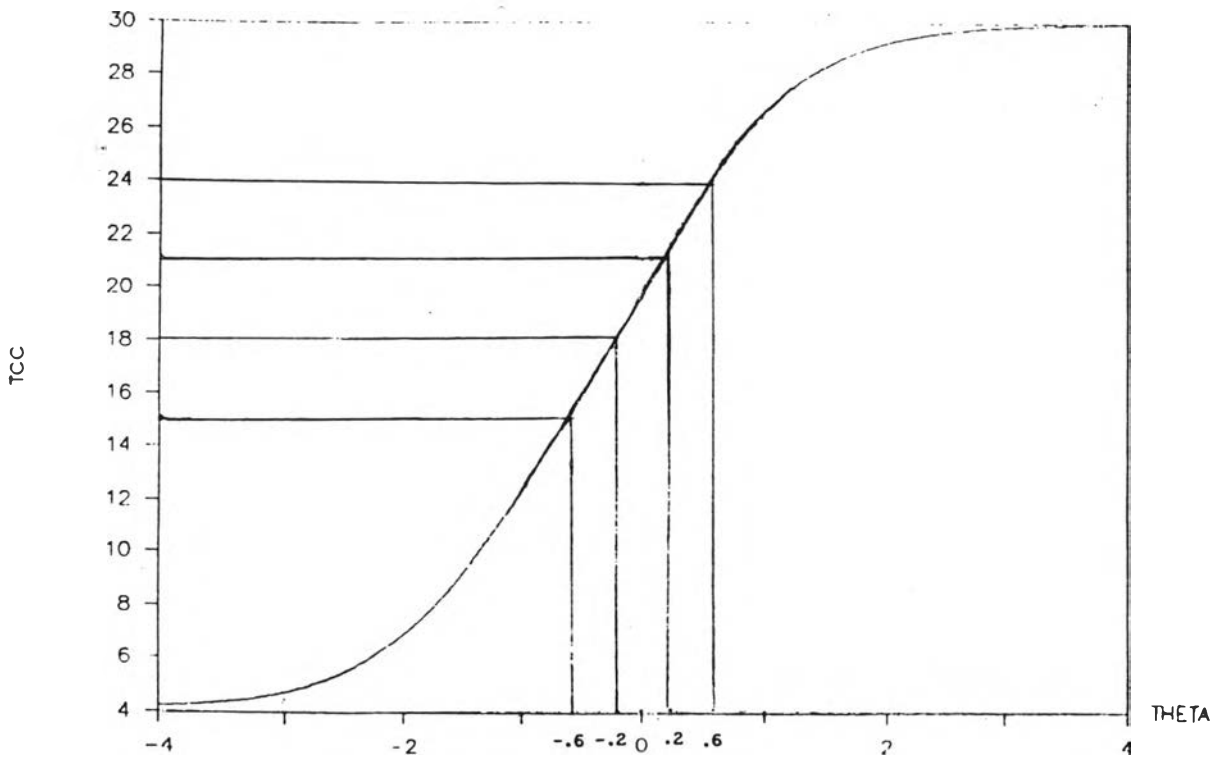
ตารางที่ 10 ระดับความสามารถ θ_0 ที่ตรงกับคะแนนโดเมน (π) ที่กำหนด สำหรับ
แบบสอบ 20 และ 30 ข้อกระทง

จำนวนข้อ	คะแนนโดเมน	ค่าความสามารถ (θ_0)
20	.5	-.5
	.6	-.1
	.7	.3
	.8	.7
30	.5	-.6
	.6	-.2
	.7	.2
	.8	.6

หรือตั้งแสดงด้วยโค้งลักษณะข้อสอบ สำหรับความยาว 20 และ 30 ข้อ เมื่อกำหนด
คะแนนโดเมน .5, .6, .7 และ .8 ได้ตั้งแผนภาพที่ 3 และแผนภาพที่ 4 ตามลำดับ



แผนภาพที่ 3 โค้งลักษณะแบบสอบของความยาว 20 ซ้อยี่สู่มได้เพื่อหาค่าความสามารถ
เมื่อกำหนดคະแนนโดเมน .5, .6, .7 และ .8



แผนภาพที่ 4 โค้งลักษณะแบบสอบของความยาว 30 ซ้อยี่สู่มได้เพื่อหาค่าความสามารถ
เมื่อกำหนดคະแนนโดเมน .5, .6, .7 และ .8

จากตารางที่ 10 แผนภาพที่ 3 และแผนภาพที่ 4 ตามลำดับ เมื่อกำหนดความยาว 20 ข้อ คะแนนโดเมน .5, .6, .7 และ .8 จะได้ค่าความสามารถที่ต้องการวัดคือ -.5, -.1, .3 และ .7 และเมื่อกำหนดความยาว 30 ข้อ คะแนนโดเมน .5, .6, .7 และ .8 จะได้ค่าความสามารถที่ต้องการ -.6, -.2, .2 และ .6

4.2 การคัดเลือกข้อกระทงที่เหมาะสมกับความสามารถ (θ_0) ที่เป็นเกณฑ์ที่ได้จากข้อ 4.1.2 โดยพิจารณาข้อกระทงในคลังข้อสอบ 104 ข้อ ที่ระดับความสามารถใกล้เคียง ๆ -.5, -.1, .3 และ .7 ในแต่ละจุดประสงค์ เมื่อความยาวแบบสอบ 20 ข้อ และเลือกข้อกระทงที่ระดับความสามารถใกล้เคียง ๆ -.6, -.2, .2 และ .6 ในแต่ละจุดประสงค์ เมื่อความยาวแบบสอบ 30 ข้อ ที่ซึ่งค่าระดับความสามารถ (θ_0) ให้โค้งแสดงลักษณะข้อกระทงสูง ๆ (Item Characteristic Curve) ซึ่งคำนวณด้วยคอมพิวเตอร์ ภาษาฟอร์แทรน หลักการเลือกข้อกระทงที่ใกล้เคียง ๆ ค่าความสามารถ (θ_0) ที่ให้โค้งแสดงสารสนเทศข้อกระทงสูง ๆ ที่ผู้วิจัยกำหนดขึ้นเองให้โค้งน้อยกว่า 20% ของความโค้งสูงสุด ทั้งนี้เพราะข้อกระทงที่ต้องการวัดที่ระดับความสามารถ (θ_0) บางค่าที่ให้โค้งแสดงสารสนเทศข้อกระทงสูง ๆ มีไม่เพียงพอ ซึ่งผลการเลือกข้อกระทงที่จะนำมาหาคะแนนจุดตัด ได้ผลดังตารางที่ 11, 12

ตารางที่ 11 ผลการเลือกและหมายเลขข้อกระทง ณ ระดับคะแนนโดเมนที่กำหนด เมื่อ
แบบสอบมีความยาว 20 ข้อ

จุด จุดประสงค์	คะแนนโดเมน .5	คะแนนโดเมน .6	คะแนนโดเมน .7	คะแนนโดเมน .8
1	4, 5, 9, 11, 14 16, 29	8, 19, 21, 28 30, 31, 39	1, 2, 3, 6, 17 24, 26	1, 2, 3, 17, 23 26, 36
2	45, 48, 59	46, 52, 56	46, 53, 55	47, 49, 50
3	62, 66, 68, 77, 78, 79, 87	66, 75, 79, 86 87, 88, 94	65, 71, 73, 75, 76, 89, 94	65, 72, 73, 76 80, 99
4	102, 105, 106	103, 106, 117	101, 103, 111	107, 113, 115, 116

จากตาราง ความยาวแบบสอบ 20 ข้อ เมื่อทำการคัดเลือกข้อกระทงที่
เหมาะสมกับความสามารถ (θ_0) พบว่า
แบบสอบฉบับคะแนนโดเมน .5, .6 มีข้อที่คัดเลือกซ้ำกัน 4 ข้อ คือ ข้อที่ 66,
79, 87 และ 106
แบบสอบฉบับคะแนนโดเมน .6, .7 มีข้อที่ซ้ำกัน 3 ข้อ คือ ข้อที่ 46, 75
และ 103
แบบสอบฉบับคะแนนโดเมน .7, .8 มีข้อที่ซ้ำกัน 8 ข้อ คือ ข้อที่ 1, 2, 3,
17, 26, 65, 73 และ 76

ตารางที่ 12 ผลการเลือกและหมายเลขข้อกระทง ณ ระดับคะแนนโดเมนที่กำหนด เมื่อ
แบบสอบมีความยาว 30 ข้อ

จุด จุดประสงค์	คะแนนโดเมน .5	คะแนนโดเมน .6	คะแนนโดเมน .7	คะแนนโดเมน .8
1	4, 5, 9, 11, 14, 15, 16, 29, 37, 38	4, 8, 19, 21, 27, 28, 29, 30, 31, 39	1, 2, 3, 6, 17, 23, 24, 26, 31, 36	1, 2, 3, 6, 17, 18, 20, 23, 26, 36
2	42, 44, 45, 48, 59	46, 48, 52, 56, 59	46, 49, 50, 53, 55	47, 49, 50, 53, 55
3	62, 66, 67, 68, 77, 78, 79, 87, 88, 91	62, 66, 68, 75, 77, 79, 86, 87, 88, 94	65, 71, 72, 73, 75, 76, 89, 94, 99	64, 65, 71, 72, 73, 76, 80, 89, 99
4	102, 104, 105, 106, 108	102, 103, 105, 106, 117	101, 103, 110, 111, 112, 118	107, 109, 113, 115, 116, 118

จากตาราง ความยาวแบบสอบ 30 ข้อ เมื่อทำการคัดเลือกข้อกระทงที่
เหมาะสมกับความสามารถ (θ_0) โดยมีจำนวนข้อกระทงในจุดประสงค์ที่ 1, 2, 3 และ 4
ตามที่คุณเชี่ยวชาญกำหนด คือ 10, 5, 10 และ 5 ข้อตามลำดับ พบว่า

แบบสอบฉบับคะแนนโดเมน .5, .6 มีข้อที่คัดเลือกซ้ำกัน 14 ข้อ คือ ข้อที่ 4,
29, 48, 59, 62, 66, 68, 77, 79, 87, 88, 102, 105 และ 106

แบบสอบฉบับคะแนนโดเมน .6, .7 มีข้อที่ซ้ำกัน 4 ข้อ คือ ข้อที่ 46, 75, 94
และ 103

แบบสอบฉบับคะแนนโดเมน .7, .8 มีข้อที่ซ้ำกัน 20 ข้อ คือ ข้อที่ 1, 2, 3,
6, 17, 23, 26, 36, 49, 50, 53, 55, 65, 71, 72, 73, 76, 89, 99 และ 118

สรุปเมื่อนำแบบสอบทั้ง 8 ฉบับมาพิจารณา ค่าอำนาจจำแนก ค่าความยาก ค่าการเดา และค่าความสามารถในแต่ละฉบับ สำหรับความยาวแบบสอบ 20 ข้อ จะได้ผลดังตารางที่ 13 และสำหรับแบบสอบความยาว 30 ข้อ จะได้ผลดังตารางที่ 14

ตารางที่ 13 พิสัยค่าอำนาจจำแนก ค่าความยาก ค่าการเดา และค่าความสามารถที่ให้โดเมน
สารสนเทศข้อกระทงสูงสุดที่ระดับคะแนนโดเมนที่กำหนดเมื่อความยาวแบบสอบ 20 ข้อ

คะแนน โดเมน	ค่าความ สามารถ	ค่าอำนาจจำแนก a	ค่าความยาก b	ค่าการเดา c	ค่าความสามารถที่ ให้ ICC สูงสุด θ^*
.5	-.5	.46058-1.17331	(-.97891)-(-.32902)	.00000-.18732	(-.85230)-(-.24333)
.6	-.1	.49519-1.17331	(-.95601)-(.00325)	.01010-.26038	(-.81338)-(-.02950)
.7	.3	.48051-1.56509	(-.22799)-(.62316)	.08427-.26755	(-.06757)-(.75443)
.8	.7	.44832-1.32215	(.07057)-(.82094)	.08427-.26855	(.28125)-(.86603)

จากตารางพบว่า แบบสอบทั้ง 4 ฉบับของความยาวแบบสอบ 20 ข้อ ให้ผลดังนี้
ค่าอำนาจจำแนกของแบบสอบฉบับคะแนนโดเมน .5 และ .6 แต่ละข้อมีค่า
ดังกล่าวใกล้เคียงกัน แต่ฉบับคะแนนโดเมน .6 ข้อที่ให้ค่าอำนาจจำแนกค่าสูงสุดสูงกว่าฉบับ
คะแนนโดเมน .5 เล็กน้อย ฉบับที่มีคะแนนโดเมน .7 มีข้อที่ให้ค่าอำนาจจำแนกสูงกว่า
ฉบับอื่น และฉบับที่มีคะแนนโดเมน .8 มีข้อที่ให้ค่าอำนาจจำแนกต่ำกว่าฉบับอื่น ๆ

ค่าความยากแบบสอบทั้ง 4 ฉบับจะยากขึ้นตามลำดับเมื่อกำหนดคะแนนโดเมนสูงขึ้น
ค่าการเดา เมื่อกำหนดคะแนนโดเมนสูงขึ้นค่าการเดาก็เพิ่มขึ้นเช่นเดียวกัน

ส่วนค่าความสามารถที่ให้โดเมนสารสนเทศข้อกระทง (ICC) สูงสุด เมื่อกำหนดคะแนน
โดเมนสูงขึ้น ค่าดังกล่าวก็เพิ่มขึ้นเช่นเดียวกัน แต่จะเห็นว่าฉบับที่มีคะแนนโดเมน .5 และ
.6 มีการคาบเกี่ยวกันมากของค่าความสามารถที่ให้ ICC สูงสุด เช่นเดียวกับแบบสอบ ฉบับ
คะแนนโดเมน .7 และ .8 ก็มีการคาบเกี่ยวกันกับค่าความสามารถที่ให้ ICC สูงสุด

ตารางที่ 14 ฟิลียค่าอำนาจจำแนก ค่าความยาก ค่าการเดา และค่าความสามารถที่ให้โด่ง
 สารสนเทศข้อกระทงสูงสุดที่ระดับคะแนนโดเมนที่กำหนดเมื่อความยาวแบบสอบ 30 ข้อ

คะแนน โดเมน	ค่าความ สามารถ	ค่าอำนาจจำแนก a	ค่าความยาก b	ค่าการเดา c	ค่าความสามารถที่ ให้ ICC สูงสุด θ^*
.5	-.6	.46058-1.17331	(-1.16437)-(-.32902)	.00000-.18732	(-1.70542)-(-.24333)
.6	-.2	.46058-1.17331	(-.98116)-(.00325)	.01010-.26938	(-.86031)-(-.02523)
.7	.2	.44832-1.56509	(-.36770)-(.68414)	.08427-.26755	(-.21693)-(.90697)
.8	.6	.34998-.156509	(.03714)-(1.19197)	.08427-.26755	(.16548)-(1.44331)

จากตารางพบว่า แบบสอบทั้ง 4 ฉบับของความยาวแบบสอบ 30 ข้อ ให้ผลดังนี้
 ค่าอำนาจจำแนกของแบบสอบฉบับคะแนนโดเมน .5 และ .6 อยู่ในช่วงเดียวกัน
 ส่วนฉบับที่มีคะแนนโดเมน .7 และ .8 มีช่วงค่าอำนาจจำแนกสูงเท่ากัน แต่ค่าอำนาจ
 จำแนกของฉบับคะแนนโดเมน .8 จะมีข้อที่มีค่าโด่งสารสนเทศข้อกระทง (ICC) ต่ำสุด
 ส่วนค่าความยากก็จะสูงขึ้นตามคะแนนโดเมนที่กำหนดสูงขึ้น
 ค่าการเดา เมื่อกำหนดคะแนนโดเมนสูงขึ้นค่าการเดาก็เพิ่มขึ้นเช่นเดียวกัน
 แต่ฉบับที่มีคะแนนโดเมน .7 และ .8 มีค่าการเดาอยู่ในช่วงเดียวกัน
 สำหรับค่าความสามารถที่ให้โด่งสารสนเทศข้อกระทงสูงสุดนั้น จะเพิ่มขึ้นเช่นเดียวกัน
 แต่จะมีการคาบเกี่ยวของค่าความสามารถมากขึ้น ซึ่งแสดงว่ามีบางข้อที่โด่งสารสนเทศข้อกระทง
 (ICC) ไม่โด่งมากที่ความสามารถที่ต้องการวัดมากนัก

5. การวิเคราะห์แบบสอบโดยใช้ทฤษฎีคลาสสิกโดยแบบสอบที่เลือกข้อกระทงมาจากคลังข้อสอบทั้ง 8 ฉบับ คือ ฉบับวัดคะแนนโดเมน .5, .6, .7, .8 ของความยาวแบบสอบ 20 ข้อ กับฉบับวัดคะแนนโดเมน .5, .6, .7, .8 ของความยาวแบบสอบ 30 ข้อ เมื่อนำมาหาค่าเฉลี่ย ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน และค่าความเที่ยง (α_{21}) ตามแนวทฤษฎีคลาสสิก จะได้ผลดังตารางที่ 15

ตารางที่ 15 ค่าเฉลี่ย ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน ค่าความเที่ยงของแบบสอบแต่ละฉบับที่วัดคะแนนโดเมน .5, .6, .7 และ .8 ตามลำดับ

คะแนนโดเมน	ความยาวแบบสอบ 20 ข้อ				ความยาวแบบสอบ 30 ข้อ			
	ค่าเฉลี่ย \bar{X}	ความแปรปรวน	ส่วนเบี่ยงเบนมาตรฐาน	ความเที่ยง α_{21}	ค่าเฉลี่ย \bar{X}	ความแปรปรวน	ส่วนเบี่ยงเบนมาตรฐาน	ความเที่ยง α_{21}
.5	13.9960	20.4760	4.5250	.8376	21.5552	41.3388	6.4295	.8837
.6	12.8090	22.7045	4.8649	.8414	19.6725	45.6746	6.8583	.8829
.7	10.9801	22.0693	4.6978	.8222	15.7990	44.4635	6.6681	.8646
.8	9.8259	19.7254	4.4414	.7913	14.6667	42.0790	6.4868	.8540

จากตารางพบว่า แบบสอบความยาว 20 ข้อ และ 30 ข้อ เมื่อวิเคราะห์ด้วยสถิติของคลาสสิก เมื่อคะแนนโดเมนสูงขึ้น ค่าเฉลี่ยมีแนวโน้มลดลง ค่าความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน และค่าความเที่ยงก็มีลักษณะเช่นเดียวกัน คือ แบบสอบที่มีคะแนนโดเมนจาก .5 เป็น .6 จะมี S, α_{21} เพิ่มขึ้น และจาก .6 เป็น .7 ค่าดังกล่าวจะลดลง สำหรับความยาวแบบสอบ 20 ข้อ แบบสอบวัดคะแนนโดเมน .7 และ .8 ที่ให้ค่าความเที่ยงของแบบสอบลดลงจะต่ำกว่าแบบสอบที่วัดคะแนนโดเมน .5 และต่ำสุดในแบบสอบฉบับวัดคะแนนโดเมน .8 ส่วนแบบสอบความยาว 30 ข้อ α_{21} ของแบบสอบแต่ละฉบับที่วัดคะแนน

โดเมนเพิ่มขึ้นจะลดลง แต่ค่าดังกล่าวนี้ไม่แตกต่างกันมาก คือ แต่ละฉบับค่าความเที่ยงต่างกันไม่เกิน .05 ส่วนเบี่ยงเบนมาตรฐานต่างกันไม่เกิน .4

6. การวิเคราะห์หาคะแนนจุดตัด ความน่าจะเป็นในการจำแนกผิดพลาด และความเที่ยงในการตัดสินใจ ณ ระดับความสามารถที่สัมพันธ์กับคะแนนโดเมน และความยาวแบบสอบที่กำหนดตามแนวทฤษฎีตอบสนองข้อกระทง ซึ่งวิธีหาคะแนนจุดตัดนี้คำนวณได้จากสูตรดังนี้

$$\text{คะแนนจุดตัด} = \sum_{i=1}^m P_i(\theta_0)$$

ค่าความน่าจะเป็นในการจำแนกผิดพลาด (P) ใช้วิธีการของโปโนเมียลโมเดล จากสูตรดังนี้

$$P = \binom{m}{x} \pi_0^x (1 - \pi_0)^{m-x}$$

ส่วนค่าความเที่ยงใช้วิธีการประยุกต์ของวอร์ม (Warm, 1980, 77) ที่สร้างความสัมพันธ์ค่าความเที่ยงกับฟังก์ชันสารสนเทศข้อสอบ $I(\theta_0)$ จากสูตร ดังนี้

$$\text{SEE} = \frac{1}{\sqrt{I(\theta_0)}}$$

$$\text{และ } r = 1 - \text{SEE}^2$$

ซึ่งผลการวิเคราะห์หาค่าต่าง ๆ ดังกล่าวได้ผลดังตารางที่ 15 และ 16 สำหรับความยาวแบบสอบ 20 ข้อ และ 30 ข้อ ตามลำดับ

ตารางที่ 16 คะแนนจุดตัด ความน่าจะเป็นในการจำแนกผิดพลาด และความเที่ยงในการจำแนกความรู้ θ
 ค่าความสามารถที่สัมพันธ์กับคะแนนโดเมนที่กำหนด ตามแนวทฤษฎีการตอบสนองข้อกระทง
 เมื่อกำหนดความยาวแบบสอบ 20 ข้อ

กำหนดคะแนน โดเมน (π_0)	ค่าความ สามารถ (θ_0)	ความสามารถของ แบบสอบในฉบับ	คะแนนจุดตัด	คะแนน โดเมน (θ_0)	ความน่าจะเป็นใน การจำแนกผิดพลาด	ความเที่ยงที่ระดับ ความสามารถ
.5	-.5	-.5	12 ^a	.6022	.1797	.9295
		-.1	14	.7117	.1904	.9151
		.3	16	.8019	.2181	.8723
		.7	17	.8684	.2360	.7894
.6	-.1	-.5	10	.5082	.1757	.9323
		-.1	13 ^a	.6305	.1814	.9377
		.3	15	.7426	.2017	.9177
		.7	17	.8303	.2360	.8680
.7	.3	-.5	7	.3677	.1819	.8768
		-.1	9	.4733	.1732	.9281
		.3	12 ^a	.5955	.1796	.9410
		.7	14	.7122	.1903	.9271
.8	.7	-.5	6	.3148	.1897	.7577
		-.1	8	.3961	.1796	.8829
		.3	10	.5052	.1760	.9325
		.7	13 ^a	.6297	.1811	.9435

^a หมายถึง คะแนนจุดตัดที่วัดได้ตรงความสามารถ θ_0 เมื่อกำหนดคะแนนโดเมน และความยาวแบบสอบที่ต้องการ

จากตารางพบว่า แบบสอบที่ใช้วิธีการเลือกข้อกระทงตามแนวทฤษฎี IRT โดยใช้วิธีประยุกต์ของวิลคอกซ์ (Wilcox 1968) จำนวน 20 ข้อ จะให้แบบสอบที่เหมาะสมใกล้เคียง ๆ ความสามารถที่สัมพันธ์กับคะแนนโดเมน .5, .6, .7 และ .8 คือ -.5, -.1, .3 และ .7 ตามลำดับ ซึ่งเมื่อนำมาหาคะแนนจุดตัด จะให้ผลในระดับความสามารถนั้น ๆ คือ ที่ระดับความสามารถ -.5, -.1, .3 และ .7 จะให้คะแนนจุดตัด คือ 12, 13, 12 และ 13 คะแนนตามลำดับ ให้ค่าความน่าจะเป็นในการจำแนกผิดพลาดที่ระดับความสามารถนั้น ๆ คือ .1797, .1814, .1796 และ .1811 ซึ่งจะเห็นว่าในแต่ละระดับความสามารถนั้น ความน่าจะเป็นในการจำแนกผิดพลาดมีค่าประมาณ .18 (18%) ส่วนค่าความเที่ยงในการจำแนกความรอบรู้ ณ ระดับความสามารถต่าง ๆ นั้นคือ .9295, .9377, .9410 และ .9435 ประมาณ .93 - .94 (93% - 94%) ซึ่งไม่แตกต่างกันมากนัก และเป็นค่าความเที่ยงในการจำแนกความรอบรู้สูงสุด

ถ้าพิจารณาแบบสอบแต่ละฉบับ และกำหนดความสามารถของแบบสอบในฉบับที่ความสามารถ -.5, -.1, .3 และ .7 จะให้คะแนนจุดตัดสูงขึ้นตามลำดับ ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบที่วัดคะแนนโดเมน .5 และ .6 จะมีค่าเพิ่มขึ้นตามความสามารถที่เพิ่มขึ้น และเมื่อกำหนดความสามารถแบบสอบในฉบับที่ความสามารถ .3 และ .7 ค่าความน่าจะเป็นในการจำแนกผิดพลาดมากกว่า .20 ส่วนแบบสอบวัดคะแนนโดเมน .6 ที่ความสามารถ -.1 จะมีค่าความน่าจะเป็นในการจำแนกผิดพลาดสูงกว่าที่วัดความสามารถ -.5 ในฉบับเดียวกัน แต่ก็น้อยมากซึ่งสูงกว่าเพียง .0057 ค่าความน่าจะเป็นในการจำแนกผิดพลาดที่ระดับความสามารถ .3 สูงกว่าระดับความสามารถ -.1 ประมาณ .0064 แต่ต่ำกว่าระดับความสามารถ -.5 และ .7 ส่วนแบบสอบวัดคะแนนโดเมน .8 ค่าความน่าจะเป็นในการจำแนกผิดพลาดที่ระดับความสามารถ .7 สูงกว่า -.1 และ .3 ประมาณ .0015 และ .0051 ตามลำดับ สำหรับค่าความเที่ยงในการจำแนกความรอบรู้ให้ค่าสูงสุดที่ระดับความสามารถ -.1 สำหรับแบบสอบที่วัดคะแนนโดเมน .7

ตารางที่ 17 คະแนจุดตัด ความน่าจะเป็นในการจำแนกผิดพลาด และความเที่ยงในการจำแนกความรู้ θ
 ค่าความสามารถที่สัมพันธ์กับคະแนโดเมนที่กำหนด ตามแนวทฤษฎีการตอบสนองข้อกระทบ
 เมื่อกำหนดความยาวแบบสอบ 30 ข้อ

กำหนดคະแน โดเมน (π_0)	ค่าความ สามารถ (θ_0)	ความสามารถของ แบบสอบในฉบับ	คະแนจุดตัด	คະแน โดเมน (π_0)	ความน่าจะเป็นใน การจำแนกผิดพลาด	ความเที่ยงที่ระดับ ความสามารถ
.5	-.6	-.6	18 [▪]	.6043	.1472	.9501
		-.2	21	.7111	.1559	.9410
		.2	24	.8001	.1795	.9139
		.6	27	.8922	.2338	.8203
.6	-.2	-.6	15	.5048	.1443	.9492
		-.2	19 [▪]	.6214	.1484	.9554
		.2	22	.7309	.1628	.9444
		.6	26	.8539	.2056	.8882
.7	.2	-.6	10	.3358	.1530	.8896
		-.2	13	.4282	.1455	.9391
		.2	16 [▪]	.5421	.1444	.9570
		.6	21	.7155	.1546	.9494
.8	.8	-.6	9	.3024	.1572	.8175
		-.2	11	.3792	.1483	.9150
		.2	15	.4842	.1423	.9525
		.6	19 [▪]	.6368	.1496	.9607

▪ หมายถึง คະแนจุดตัดที่วัดได้ตรงความสามารถ θ_0 เมื่อกำหนดคະแนโดเมน และความยาวแบบสอบที่ต้องการ

จากตารางพบว่า แบบสอบที่ใช้วิธีการเลือกข้อกระทงตามแนวทฤษฎี IRT โดยใช้วิธีประยุกต์ของวิลคอกซ์ (Wilcox, 1968) จำนวน 30 ข้อ จะให้แบบสอบที่เหมาะสมใกล้เคียง ๗ ความสามารถที่สัมพันธ์กับคะแนนโดเมน .5, .6, .7 และ .8 คือ -.6, -.2, .2 และ .6 ตามลำดับ ซึ่งเมื่อนำมาหาคะแนนจุดตัดที่ระดับความสามารถดังกล่าวจะได้คะแนนจุดตัด คือ 18, 19, 16 และ 19 คะแนนตามลำดับ ค่าความน่าจะเป็นการจำแนกผิดพลาดที่ระดับความสามารถนั้น ๆ คือ .1472, .1484, .1444 และ .1496 ตามลำดับ ซึ่งมีค่าประมาณ .15 (15%) จะเห็นว่าค่าดังกล่าวมีแนวโน้มจะสูงขึ้น นอกจากฉบับวัดคะแนนโดเมน .7 จะให้ค่าความน่าจะเป็นในการจำแนกผิดพลาดต่ำที่สุด ส่วนค่าความเที่ยงในการจำแนกความรอบรู้ ณ ระดับความสามารถนั้น ๆ คือ .9501, .9554, .9570 และ .9607 ประมาณ .95 - .96 (95% - 96%) ซึ่งเป็นค่าที่สูงสุดเมื่อเทียบกับความสามารถระดับอื่นๆ

ถ้าพิจารณาแบบสอบแต่ละฉบับพร้อมกับกำหนดความสามารถของแบบสอบในฉบับเดียวกับที่ -.6, -.2, .2 และ .6 จะให้คะแนนจุดตัดสูงขึ้นตามลำดับ ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบที่วัดคะแนนโดเมน .5 และ .6 ที่ระดับความสามารถ -.6 จะต่ำสุด และที่ความสามารถในฉบับเดียวกันที่วัดความสามารถ .8 จะให้ค่ามากกว่าที่ระดับความสามารถ .2 แต่แบบสอบฉบับวัดคะแนนโดเมน .6 เมื่อกำหนดความสามารถในฉบับเป็น -.2 จะสูงกว่า -.6 ประมาณ .0041 สำหรับแบบสอบวัดคะแนนโดเมน .7, .8 ที่ความสามารถ .2 จะให้ค่าความน่าจะเป็นในการจำแนกผิดพลาดต่ำสุด

7. การวิเคราะห์หาคะแนนจุดตัดโดยวิธีประยุกต์ของแอมเบิลตันและกรูจเตอร์ ความน่าจะเป็นในการจำแนกผิดพลาดโดยใช้ไบโนเมียลโมเดล และความเที่ยงในการจำแนกความรอบรู้ ณ ระดับคะแนนโดเมนและความยาวที่กำหนดด้วยวิธีของฮวิน โดยใช้แนวคิดทฤษฎีคลาสสิก ได้ผลดังตารางที่ 17 สำหรับความยาวแบบสอบ 20 ข้อ และตารางที่ 18 สำหรับความยาวแบบสอบ 30 ข้อ

ตารางที่ 18 คะแนนจุดตัด ความน่าจะเป็นในการจำแนกความผิดพลาด และความเที่ยงในการ
จำแนกความรอบรู้ π คะแนนโดเมนที่กำหนด โดยใช้ทฤษฎีคลาสสิก เมื่อ
ความยาวแบบสอบ 20 ข้อ

กำหนดคะแนน โดเมน (π_0)	คะแนนโดเมน ในแบบสอบ	คะแนนจุดตัด	ความน่าจะเป็นใน การจำแนกผิดพลาด	ความเที่ยงในการ จำแนกความรอบรู้ (k)
.5	.5	10 [▪]	.1762	.5509
	.6	12	.1797	.5789
	.7	14	.1916	.5418
	.8	16	.2182	.5924
.6	.5	10	.1762	.6197
	.6	12 [▪]	.1797	.6389
	.7	14	.1916	.5938
	.8	16	.2181	.6200
.7	.5	10	.1762	.4810
	.6	12	.1797	.5420
	.7	14 [▪]	.1916	.5326
	.8	16	.2182	.4785
.8	.5	10	.1762	.5376
	.6	12	.1797	.5584
	.7	14	.1916	.4040
	.8	16 [▪]	.2182	.5816

▪ หมายถึง คะแนนจุดตัดที่วัดได้ตรงคะแนนโดเมน และความยาวแบบสอบที่ต้องการวัด

จากตารางพบว่า เมื่อนำแบบสอปที่เลือกข้อกระทงตามแนวทฤษฎี IRT มาวิเคราะห์หาคะแนนจุดตัด ความเที่ยงในการจำแนกความรอบรู้โดยใช้สัมประสิทธิ์แคปปาวิธีของฮวิน (Huynh) ตามแนวทฤษฎีคลาสสิกจำนวน 20 ข้อ จะเห็นว่าความยากของแบบสอปเพิ่มขึ้น คะแนนจุดตัดจะให้ค่าที่คงที่ คือ 10, 12, 14 และ 16 คะแนนตามลำดับ เมื่อกำหนดคะแนนโดเมน .5, .6, .7 และ .8 รวมทั้งค่าความน่าจะเป็นในการจำแนกผิดพลาดก็ให้ค่าที่คงที่ และมีค่าเพิ่มขึ้นตามการกำหนดคะแนนโดเมนเพิ่มขึ้น คือ .1762, .1797, .1916 และ .2182 ตามลำดับ ส่วนค่าความเที่ยงในการจำแนกความรอบรู้จะแปรเปลี่ยนไปตามคะแนนจุดตัด คือ .5509, .6387, .5326 และ .5816 ซึ่งให้ค่าประมาณ .53 - .64 (53% - 64%)

ถ้าพิจารณาแบบสอปแต่ละฉบับตามแนวทฤษฎีคลาสสิกที่คะแนนโดเมนที่กำหนด .5, .6, .7 และ .8 สัมประสิทธิ์แคปปา (k) จะให้ค่าที่เพิ่มขึ้น ลดลงและ เพิ่มขึ้นเมื่อกำหนดคะแนนโดเมนเพิ่มขึ้น นอกจากฉบับวัดคะแนนโดเมน .7 ที่สัมประสิทธิ์แคปปาให้ค่าที่เพิ่มขึ้นและลดลงเท่านั้น

ตารางที่ 19 คະแนนจุดตัด ความน่าจะเป็นในการจำแนกความผิดพลาด และความเที่ยงในการ
จำแนกความรอบรู้ π_0 คະแนนโดเมนที่กำหนด โดยใช้ทฤษฎีคลาสสิก เมื่อ
ความยาวแบบสอบ 30 ข้อ

กำหนดคະแนน โดเมน (π_0)	คະแนนโดเมน ในแบบสอบ	คະแนนจุดตัด	ความน่าจะเป็นใน การจำแนกผิดพลาด	ความเที่ยงในการ จำแนกความรอบรู้ (k)
.5	.5	15 [▪]	.1444	.6372
	.6	18	.1474	.6488
	.7	21	.1573	.6593
	.8	24	.1795	.6596
.6	.5	15	.1444	.6398
	.6	18 [▪]	.1474	.6521
	.7	21	.1573	.6575
	.8	24	.1795	.6158
.7	.5	15	.1444	.6366
	.6	18	.1474	.5919
	.7	21 [▪]	.1573	.5735
	.8	24	.1795	.5445
.8	.5	15	.1444	.5046
	.6	18	.1474	.6189
	.7	21	.1573	.5934
	.8	24 [▪]	.1795	.5495

▪ หมายถึง คະแนนจุดตัดที่วัดได้ตรงคະแนนโดเมน และความยาวแบบสอบที่ต้องการวัด

จากตารางพบว่า เมื่อนำแบบสอบที่เลือกข้อกระทงตามแนวทฤษฎี IRT มาวิเคราะห์หาคะแนนจุดตัด ความเที่ยงในการจำแนกความรอบรู้โดยใช้สัมประสิทธิ์แคปป่า (k) วิธีของฮวิน (Huyhn) ตามแนวทฤษฎีคลาสสิก จำนวน 30 ข้อ 4 ฉบับ คะแนนจุดตัดและค่าความน่าจะเป็นในการจำแนกความผิดพลาดจะให้ค่าที่คงที่ และมีค่าเพิ่มขึ้นตามการกำหนดคะแนนโดเมน คือ .5, .6, .7 และ .8 และจะให้คะแนนจุดตัดที่ระดับคะแนนโดเมนดังกล่าว คือ 15, 18, 21 และ 24 คะแนนตามลำดับ ค่าความน่าจะเป็นในการจำแนกผิดพลาด คือ .1444, .1474, .1573 และ .1795 ตามลำดับ ส่วนค่าความเที่ยงในการจำแนกความรอบรู้จะแปรเปลี่ยนไป คือ .6372, .6521, .5735 และ .5495

ถ้าพิจารณาแบบสอบแต่ละฉบับตามแนวทฤษฎีคลาสสิก เมื่อกำหนดคะแนนโดเมน .5, .6, .7 และ .8 แล้ว พบว่า

แบบสอบฉบับคะแนนโดเมน .5 จะให้ค่า ส.ป.ส. แคปป่า เพิ่มขึ้น

แบบสอบฉบับวัดคะแนนโดเมน .6 จะให้ค่า ส.ป.ส. แคปป่า (k) สูงขึ้น และให้ค่าต่ำสุด เมื่อกำหนดคะแนนโดเมนในแบบสอบเป็น .8

แบบสอบฉบับวัดคะแนนโดเมน .7 จะให้ค่า ส.ป.ส.แคปป่า (k) ลดลง

แบบสอบฉบับวัดคะแนนโดเมน .8 จะให้ค่า ส.ป.ส. แคปป่า (k) เมื่อกำหนดคะแนนโดเมนในแบบสอบ .6 สูงกว่า .5 จะให้ค่าดังกล่าวลดลงที่คะแนนโดเมน .7 และ .8

8. การเปรียบเทียบผลการวิเคราะห์ตามแนวทฤษฎีการตอบสนองข้อกระทง และทฤษฎีคลาสสิก ได้แสดงผลการเปรียบเทียบดังตารางที่ 20 และ 21 สำหรับความยาวแบบสอบ 20 และ 30 ข้อ ตามลำดับ ดังนี้

ตารางที่ 20 การเปรียบเทียบผลการหาคะแนนจุดตัด ความน่าจะเป็นในการจำแนกผิดพลาด ความเที่ยงในการตัดสินใจ ณ ระดับคะแนนโดเมนที่ต้องการวัดที่เหมาะสมกับ ค่าความสามารถ (θ_0) ตามแนวทฤษฎีการตอบสนองข้อกระทงกับตามแนว ทฤษฎีคลาสสิก เมื่อแบบสอบมีความยาว 20 ข้อ

θ_0	คะแนนโดเมน ตามแนว IRT	คะแนนจุดตัด ตามแนว CCT	การจำแนกความผิดพลาด		ความเที่ยง	
			IRT	CCT	IRT	CCT
.5	12	10	.1797	.1762	.9295	.5509
.6	13	12	.1814	.1797	.9377	.6389
.7	12	14	.1796	.1916	.9410	.5326
.8	13	16	.1811	.2182	.9435	.5816

จากตารางที่ 20 พบว่า สำหรับแบบสอบความยาว 20 ข้อ ได้ข้อสรุปผลดังนี้
คะแนนจุดตัดที่ใช้วิธีการประยุกต์ของแฮมเบิลตันและกรูจเตอร์ (Hambleton and Gruijter, 1983) ตามแนวทฤษฎีคลาสสิก (CCT) จะได้คะแนนจุดตัดเป็นสัดส่วนโดยตรงตามคะแนนโดเมนที่กำหนดสูงขึ้น แต่คะแนนจุดตัดตามแนวทฤษฎีการตอบสนองข้อกระทง (IRT) จะให้ค่าประมาณ 60% - 65%

ค่าความน่าจะเป็นในการจำแนกผิดพลาด (P) ของคะแนนจุดตัดตามแนวทฤษฎี IRT ฉบับคะแนนโดเมน .5 และ .6 มีแนวโน้มสูงกว่าค่าความน่าจะเป็นในการจำแนกผิดพลาดตามแนวทฤษฎีคลาสสิก ซึ่งขัดแย้งกับสมมติฐาน ข้อ 1.1, 1.2 แต่ฉบับคะแนนโดเมน .7 และ .8 ค่า P ดังกล่าวของคะแนนจุดตัดตามแนว IRT จะต่ำกว่าค่า P ของคะแนนจุดตัดตามแนวทฤษฎี CCT ซึ่งสอดคล้องกับสมมติฐาน ข้อ 1.3, 1.4

ส่วนค่าความเที่ยงในการตัดสินใจของคะแนนจุดตัดประยุกต์วิธีของวอร์มตามแนวทฤษฎี IRT จะให้ค่าสูงกว่าวิธีการหาค่าความเที่ยงโดยวิธีของฮวินที่ใช้แนวทฤษฎี CCT ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ข้อ 3.1 - 3.4

ตารางที่ 21 การเปรียบเทียบผลการหาคะแนนจุดตัด ความน่าจะเป็นในการจำแนกชนิดผลผลิต ความเที่ยงในการตัดสินใจ ณ ระดับคะแนนโดเมนที่ต้องการวัดที่เหมาะสมกับ ค่าความสามารถ (θ_0) ตามแนวทฤษฎีการตอบสนองข้อกระทงกับตามแนว ทฤษฎีคลาสสิก เมื่อแบบสอบมีความยาว 30 ข้อ

θ_0	คะแนนโดเมน ตามแนว IRT	คะแนนจุดตัด ตามแนว CCT	การจำแนกความผิดพลาด		ความเที่ยง	
			IRT	CCT	IRT	CCT
.5	18	15	.1472	.1444	.9501	.6372
.6	19	18	.1484	.1474	.9554	.6521
.7	16	21	.1444	.1573	.9570	.5735
.8	19	24	.1496	.1795	.9607	.5495

จากตารางที่ 21 พบว่า สำหรับแบบสอบความยาว 30 ข้อ ได้ข้อสรุปผลดังนี้
คะแนนจุดตัดที่ใช้วิธีการประยุกต์ของแฮมเบิลตันและกรูจเตอร์ (Hambleton and Gruijter, 1977) ตามแนวทฤษฎีคลาสสิก (CCT) จะได้คะแนนจุดตัดเป็นสัดส่วนโดยตรงตามคะแนนโดเมนที่กำหนดสูงขึ้น เช่นเดียวกับแบบสอบความยาว 20 ข้อ แต่คะแนนจุดตัดตามแนวทฤษฎีการตอบสนองข้อกระทง (IRT) จะให้ค่าประมาณ 60% ยกเว้นฉบับที่กำหนดคะแนนโดเมน .7 คะแนนจุดตัดประมาณ 53 % หรือ 16 คะแนน

ค่าความน่าจะเป็นในการจำแนกชนิดผลผลิตของคะแนนจุดตัดตามแนวทฤษฎี IRT ฉบับคะแนนโดเมน .5 และ .6 มีแนวโน้มสูงกว่าของคะแนนจุดตัดตามแนวทฤษฎีคลาสสิก (CCT) ซึ่งขัดแย้งกับสมมติฐาน ข้อ 2.1, 1.2, 2.2 แต่ฉบับคะแนนโดเมน .7 และ .8 ค่า P ของคะแนนจุดตัดตามแนว IRT จะต่ำกว่าค่า P ของคะแนนจุดตัดตามแนวทฤษฎี CCT ซึ่งสอดคล้องกับสมมติฐาน ข้อ 2.3, 2.4

ค่าความเที่ยงของคะแนนจุดตัดประยุกต์โดยวิธีของวอร์มตามแนวทฤษฎี IRT จะให้ค่าสูงกว่าวิธีการหาค่าความเที่ยงโดยวิธีของอวินทุกฉบับ ซึ่งสอดคล้องกับสมมติฐานข้อ 4.1-4.4