

ไอเกนที่ 2 (Eigen Ratio : ER) สูงกว่า 3 มีแบบสอบเพียง 2 ชุด ซึ่งมีค่าดัชนี ER น้อยกว่า 3 คือ แบบสอบในผู้สอบระดับความสามารถต่ำ ระดับความยาว 20 และ 60 ข้อ ผู้วิจัยจึงได้วิเคราะห์โดยการทดสอบค่าสกรี (Scre test) เพื่อตรวจสอบว่าแบบสอบวัดคุณลักษณะเด่น (Dominant trait) เพียงลักษณะเดียว (One trait) ตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบหรือไม่

ตารางที่ 6 ผลการตรวจสอบความเป็นเอกมิติของแบบสอบ

เงื่อนไข	ค่าไอเกน		สัดส่วนของค่าไอเกน (Eigen Ratio:ER)
	องค์ประกอบที่ 1	องค์ประกอบที่ 2	
ระดับความสามารถต่ำ			
20 ข้อ	3.213	1.114	2.884
40 ข้อ	5.789	1.932	3.042
60 ข้อ	8.905	3.008	2.960
ระดับความสามารถปานกลาง			
20 ข้อ	4.063	1.335	3.043
40 ข้อ	4.768	1.300	3.667
60 ข้อ	8.517	2.480	3.434
ระดับความสามารถสูง			
20 ข้อ	4.052	1.343	3.017
40 ข้อ	4.615	1.471	3.137
60 ข้อ	8.173	2.134	3.829

แบบสอบมีความยาว 3 ระดับคือ 20, 40 และ 60 ข้อ ปรากฏว่า

แบบสอบสำหรับกลุ่มผู้สอบซึ่งมีระดับความสามารถปานกลาง มีค่าดัชนี ER สูงสุด ในระดับความยาวของแบบสอบ 40 ข้อ คือ เท่ากับ 3.667

กลุ่มผู้สอบซึ่งมีความสามารถระดับสูง มีค่าดัชนี ER ของแบบสอบสูงที่สุดในระดับ ความยาว 60 ข้อ โดยมีค่าดัชนี ER เท่ากับ 3.829

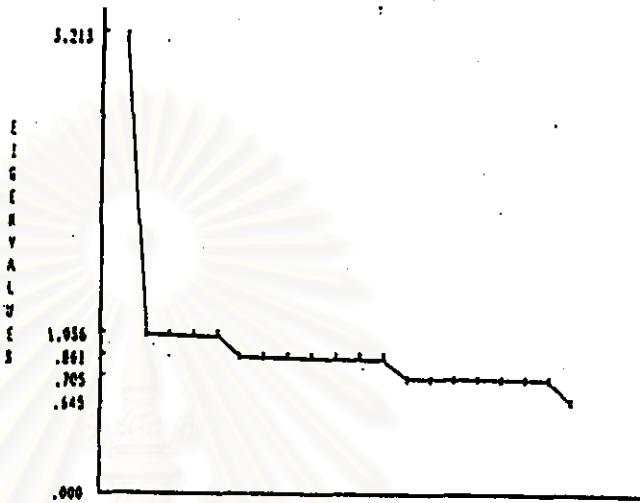
ส่วนกลุ่มผู้สอบซึ่งมีระดับความสามารถต่ำ มีค่าดัชนีของแบบสอบต่ำสุดในทุกระดับ ความยาวของแบบสอบ โดยมีค่าดัชนี ER เท่ากับ 2.884 3.042 และ 2.960 ตามลำดับ

สำหรับผลการทดสอบสกรี ปรากฏว่าแบบสอบซึ่งสุ่มมาจากเมตริกซ์คำตอบที่จำลอง ขึ้น มีองค์ประกอบเด่นเพียงองค์ประกอบเดียว ส่วนองค์ประกอบอื่นๆ ที่เหลือ มีค่าไอเกนต่ำ และเป็นค่าที่ใกล้เคียงกัน ซึ่งตรงกับที่ ลอร์ด (Lord, 1980) ได้กล่าวไว้ว่า แบบสอบที่มีความเป็น เอกมิติ ควรีผลการทดสอบสกรี (Scree Test) ที่เส้นกราฟมีความลาดชันในองค์ประกอบแรก และค่อยๆ ลาดลงในองค์ประกอบหลังๆ แสดงว่าแบบสอบในทุกเงื่อนไขที่ศึกษา มีความเป็น เอกมิติ (Unidimensionality) ตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ผลการ ทดสอบสกรีปรากฏดังแผนภูมิที่ 1 ถึง แผนภูมิที่ 3

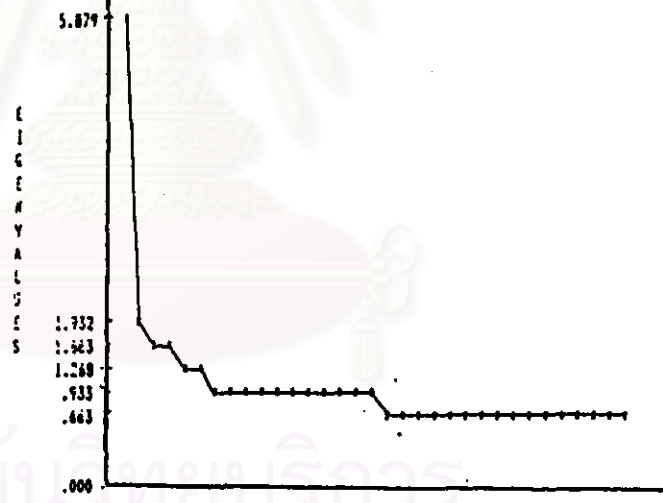
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แผนภูมิที่ 1 ผลการทดสอบค่าสกรี (Scree Test) เมื่อผู้สอบมีระดับความสามารถต่ำ

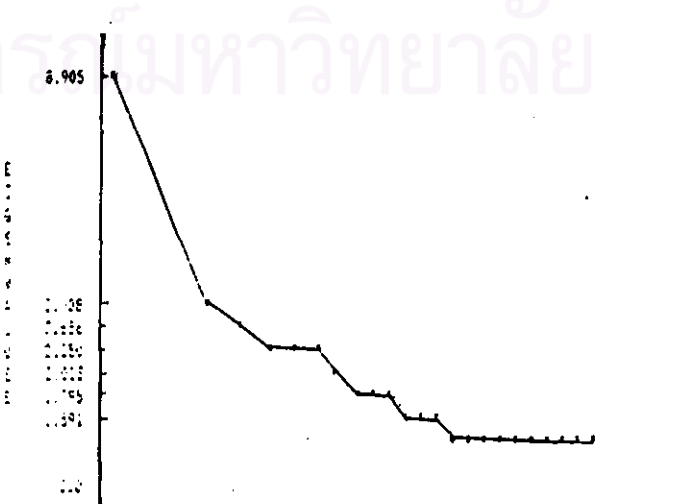
แบบสอบความยาว 20 ข้อ



แบบสอบความยาว 40 ข้อ

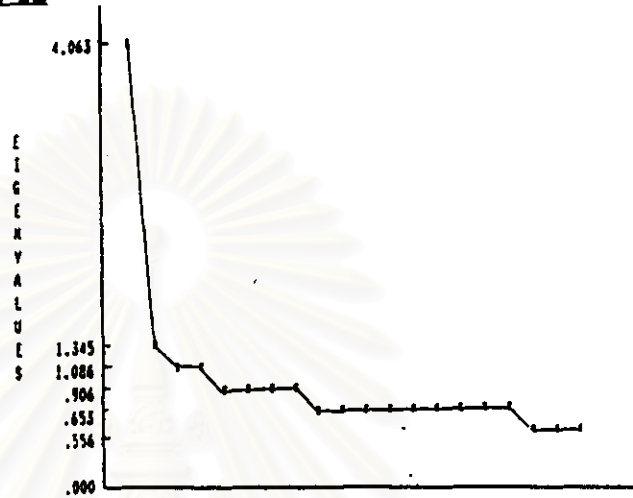


แบบสอบความยาว 60 ข้อ

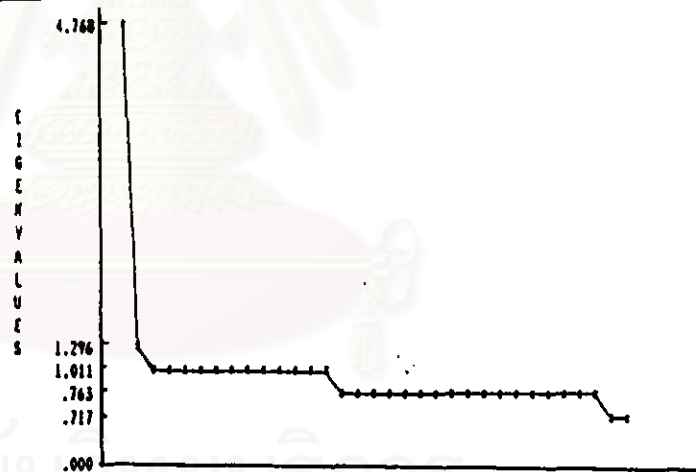


แผนภูมิที่ 2 ผลการทดสอบค่าสกรี (Scree Test)
เมื่อผู้สอบมีระดับความสามารถปานกลาง

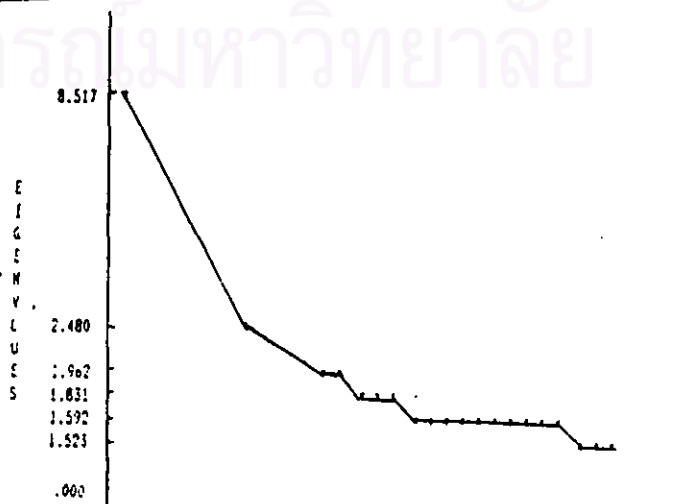
แบบสอบความยาว 20 ข้อ



แบบสอบความยาว 40 ข้อ

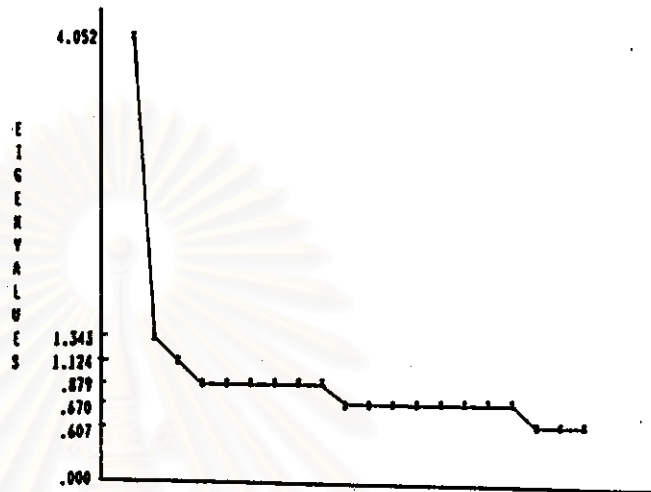


แบบสอบความยาว 60 ข้อ

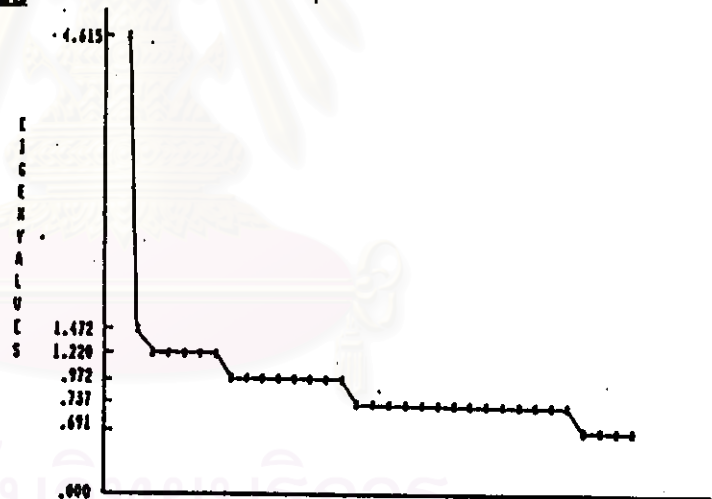


แผนภูมิที่ 3 ผลการทดสอบค่าสกรี (Scree Test)
เมื่อผู้สอบมีระดับความสามารถสูง

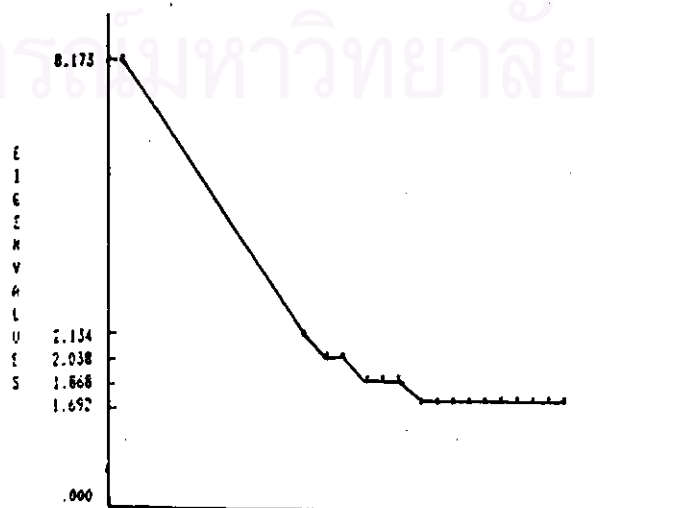
แบบสอบความยาว 20 ข้อ



แบบสอบความยาว 40 ข้อ



แบบสอบความยาว 60 ข้อ



1.2 คุณภาพของแบบสอบ

แบบสอบซึ่งเป็นเมตริกซ์คำตอบจำนวน 9 เงื่อนไข ผู้วิจัยได้วิเคราะห์คุณภาพตาม ทฤษฎีการทดสอบแบบดั้งเดิม ด้วยโปรแกรม BILOG โปรแกรมย่อย IS11 เพื่อหาค่าความยาก ค่าอำนาจจำแนก และลักษณะการแจกแจงของคะแนนของผู้สอบในกลุ่มต่าง ๆ รวมทั้งศึกษาคุณภาพของแบบสอบในด้านความเที่ยง โดยใช้สูตร KR - 20 ปรากฏผลดังตาราง

ตารางที่ 7 ค่าความยาก ค่าอำนาจจำแนก และค่าความเที่ยงเฉลี่ยของแบบสอบ

เงื่อนไข	คุณภาพของแบบสอบ			จำนวนผู้สอบ
	ค่าความยาก	ค่าอำนาจจำแนก	ค่าความเที่ยง	
ระดับความสามารถต่ำ				
20 ข้อ	0.388	0.511	0.716	1,000
40 ข้อ	0.396	0.489	0.846	1,000
60 ข้อ	0.425	0.491	0.898	1,000
ระดับความสามารถปานกลาง				
20 ข้อ	0.501	0.609	0.782	1,000
40 ข้อ	0.577	0.465	0.807	1,000
60 ข้อ	0.557	0.481	0.888	1,000
ระดับความสามารถสูง				
20 ข้อ	0.729	0.632	0.778	1,000
40 ข้อ	0.827	0.558	0.783	1,000
60 ข้อ	0.791	0.582	0.833	1,000

แบบสอบทุกระดับความยาวในกลุ่มผู้สอบความสามารถต่ำ เป็นข้อสอบที่ยาก โดยมีแบบสอบที่ยากที่สุดคือ แบบสอบชุด 20 ข้อ มีค่าความยากเท่ากับ 0.386 ส่วนค่าอำนาจจำแนกอยู่ในระดับใช้ได้

แบบสอบทุกระดับความยาวในกลุ่มผู้สอบความสามารถปานกลาง มีค่าความยากง่ายปานกลาง และมีอำนาจจำแนกในระดับดี

แบบสอบทุกระดับความยาวในกลุ่มผู้สอบความสามารถสูง เป็นข้อสอบที่ง่าย และมีอำนาจจำแนกในระดับดี

ความเที่ยงของแบบสอบในทุกระดับความสามารถของผู้สอบอยู่ในเกณฑ์ดี (0.716 - 0.898) โดยค่าความเที่ยงจะสูงขึ้นเมื่อแบบสอบมีความยาวมากขึ้น

1.3 ค่าสถิติพื้นฐานของแบบสอบ

ค่าสถิติพื้นฐานของแบบสอบ จำแนกตามระดับความสามารถของผู้สอบ ผู้วิจัยรายงานค่ามัธยฐานเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด ค่าสูงสุด ค่าฐานนิยม ความเบ้และความโด่ง ของแบบสอบในทุกเงื่อนไขที่ศึกษา ปรากฏดังตารางที่ 8 ถึงตารางที่ 10

ตารางที่ 8 ลักษณะการแจกแจงของคะแนนของแบบสอบความยาว 20 ข้อ

ค่าสถิติ	ระดับความสามารถของผู้สอบ		
	ต่ำ ($\bar{\theta} = -1.5$)	ปานกลาง ($\bar{\theta} = 0.0$)	สูง ($\bar{\theta} = 1.5$)
ค่ามัธยฐานเลขคณิต (mean)	7.72	10.01	14.58
ส่วนเบี่ยงเบนมาตรฐาน (S.D.)	3.54	3.53	3.60
ค่าต่ำสุด (minimum)	1	2	4
ค่าสูงสุด (maximum)	17	18	19
ฐานนิยม (mode)	8	10	18
ความเบ้ (skew)	0.34	0.20	-0.90
ความโด่ง (kurtosis)	-0.46	-0.60	0.15

เมื่อแบบสอบมีความยาว 20 ข้อ กลุ่มผู้สอบที่มีระดับความสามารถต่างกัน มีลักษณะการแจกแจงของคะแนนต่างกัน โดยกลุ่มผู้สอบความสามารถปานกลาง มีการกระจายของคะแนนต่ำสุด ส่วนกลุ่มผู้สอบซึ่งมีระดับความสามารถสูง มีการกระจายของคะแนนสูงสุด การแจกแจงของคะแนนของผู้สอบซึ่งมีระดับความสามารถต่ำและปานกลาง มีการแจกแจงแบบเบ้ขวา ส่วนผู้สอบซึ่งมีความสามารถสูงมีลักษณะการแจกแจงของคะแนนแบบเบ้ซ้าย

ตารางที่ 9 ลักษณะการแจกแจงของคะแนนของแบบสอบความยาว 40 ข้อ

ค่าสถิติ	ระดับความสามารถของผู้สอบ		
	ต่ำ ($\bar{\theta} = -1.5$)	ปานกลาง ($\bar{\theta} = 0.0$)	สูง ($\bar{\theta} = 1.5$)
ค่าเฉลี่ยเลขคณิต (mean)	15.86	23.07	33.09
ส่วนเบี่ยงเบนมาตรฐาน (S.D.)	6.81	5.96	4.08
ค่าต่ำสุด (minimum)	2	8	14
ค่าสูงสุด (maximum)	33	36	39
ฐานนิยม (mode)	16	22	36
ความเบ้ (skew)	0.36	-0.13	-0.97
ความโด่ง (kurtosis)	-0.53	-0.57	1.01

เมื่อแบบสอบมีความยาว 40 ข้อกลุ่มผู้สอบระดับความสามารถปานกลาง มีคะแนนกระจายมากที่สุด ลักษณะการแจกแจงของกลุ่มผู้สอบความสามารถปานกลาง และความสามารถสูงมีลักษณะเบ้ขวา ส่วนผู้สอบซึ่งมีระดับความสามารถต่ำ มีลักษณะการแจกแจงแบบเบ้ซ้าย กลุ่มผู้สอบซึ่งมีความสามารถสูงมีค่าสัมประสิทธิ์ความโด่ง (kurtosis) สูงสุด คือเท่ากับ 1.01

ตารางที่ 10 ลักษณะการแจกแจงของคะแนนของแบบสอบความยาว 60 ข้อ

ค่าสถิติ	ระดับความสามารถของผู้สอบ		
	ต่ำ ($\bar{\theta} = -1.5$)	ปานกลาง ($\bar{\theta} = 0.0$)	สูง ($\bar{\theta} = 1.5$)
ค่ามัธยฐานเลขคณิต (mean)	25.51	33.43	47.45
ส่วนเบี่ยงเบนมาตรฐาน (S.D.)	10.18	8.80	7.48
ค่าต่ำสุด (minimum)	3	14	22
ค่าสูงสุด (maximum)	43	51	59
ฐานนิยม (mode)	24	37	50
ความเบ้ (skew)	0.48	-0.10	-0.88
ความโด่ง (kurtosis)	-0.32	-0.82	0.46

เมื่อแบบสอบมีความยาว 60 ข้อ กลุ่มผู้สอบระดับความสามารถต่ำ มีคะแนนกระจายมากที่สุด (S.D. = 10.18) ส่วนกลุ่มผู้สอบซึ่งมีระดับความสามารถสูงมีการกระจายของคะแนนน้อยที่สุด (S.D. = 7.48) กลุ่มผู้สอบความสามารถปานกลางและความสามารถสูง มีลักษณะการแจกแจงของคะแนนแบบเบ้ขวา ส่วนกลุ่มผู้สอบที่มีระดับความสามารถต่ำมีลักษณะการกระจายแบบเบ้ซ้าย คะแนนของผู้สอบทุกระดับความสามารถมีความโด่ง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 2 ผลการตรวจสอบบุคคลที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนีแอลฟา (L₂ index) และดัชนีดับเบิลยูวัน (W₁ index)

2.1 จำนวนผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบ เป็นจำนวนผู้สอบในทุกเงื่อนไขที่ศึกษา ปรากฏผลดังตารางที่ 11

ตารางที่ 11 จำนวนและร้อยละของผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบ

ความยาวของแบบสอบ	จำนวนผู้สอบที่ไม่เหมาะสม					
	ดัชนี L ₂		ดัชนี W ₁ (infit)		ดัชนี W ₁ (outfit)	
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
ความสามารถต่ำ						
ความยาวของแบบสอบ						
20 ข้อ	95	9.50	185	18.50	39	3.90
40 ข้อ	202	20.20	58	5.80	51	5.10
60 ข้อ	249	24.90	79	7.90	55	5.50
ความสามารถปานกลาง						
ความยาวของแบบสอบ						
20 ข้อ	113	11.30	64	6.40	41	4.10
40 ข้อ	294	29.40	53	5.30	48	4.80
60 ข้อ	333	33.00	108	10.80	96	9.60
ความสามารถสูง						
ความยาวของแบบสอบ						
20 ข้อ	194	19.40	44	4.40	34	3.40
40 ข้อ	368	36.80	45	4.50	32	3.20
60 ข้อ	356	35.60	53	5.30	44	4.40

เมื่อผู้สอบมีระดับความสามารถต่ำ จำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบ ซึ่งได้จากการตรวจสอบด้วยดัชนีบ่งชี้ความเหมาะสมของบุคคลต่างวิธี ในแต่ละระดับความยาวของแบบสอบมีจำนวนแตกต่างกัน ดังนี้

การประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลด้วยฟังก์ชันโลจิสติก (L_2 index) มีจำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบ ในแบบสอบความยาว 20, 40 และ 60 ข้อ จำนวน 96, 202 และ 249 คน คิดเป็นร้อยละ 9.50, 20.20 และ 24.90 ตามลำดับ

การประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลจากคะแนนส่วนที่เหลือ (W_1 index) เมื่อพิจารณาค่าสถิติ Infit ในแบบสอบความยาว 20, 40 และ 60 ข้อ มีจำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบ จำนวน 185, 58 และ 79 คน คิดเป็นร้อยละ 18.50, 5.80 และ 7.90 ตามลำดับ และเมื่อพิจารณาค่าสถิติ Outfit มีจำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบจำนวน 39, 51 และ 55 คน คิดเป็นร้อยละ 3.90, 5.10 และ 5.50 ตามลำดับ

จะเห็นว่าดัชนี L_2 และดัชนี W_1 เมื่อพิจารณาค่า Outfit วิเคราะห์ผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบได้เพิ่มขึ้นเมื่อแบบสอบมีจำนวนข้อสอบมากขึ้น

เมื่อผู้สอบมีระดับความสามารถปานกลาง จำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบ ซึ่งได้จากการประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลต่างวิธี ในแต่ละระดับความยาวของแบบสอบ มีจำนวนแตกต่างกันดังนี้

การประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลด้วยฟังก์ชันโลจิสติก (L_2 index) มีจำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบ ในแบบสอบความยาว 20, 40 และ 60 ข้อ จำนวน 113, 294 และ 330 คน คิดเป็นร้อยละ 11.30, 29.40 และ 33.00 ตามลำดับ

การประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลจากคะแนนส่วนที่เหลือ (W_1 index) เมื่อพิจารณาค่าสถิติ Infit ในแบบสอบความยาว 20, 40 และ 60 ข้อ มีจำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบ จำนวน 64, 53 และ 108 คน คิดเป็นร้อยละ 6.40, 5.30 และ 10.80 ตามลำดับ และเมื่อพิจารณาค่าสถิติ Outfit มีจำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบจำนวน 41, 48 และ 96 คน คิดเป็นร้อยละ 4.10, 4.80 และ 9.60 ตามลำดับ

จะเห็นได้ว่าดัชนี L_2 และ W_1 เมื่อพิจารณาค่าสถิติ Outfit วิเคราะห์ผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบได้เพิ่มขึ้นเมื่อแบบสอบมีจำนวนข้อสอบมากขึ้น

เมื่อผู้สอบมีระดับความสามารถสูง จำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบซึ่งได้จากการประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลต่างวิธี ในแต่ละระดับความยาวของแบบสอบ มีจำนวนแตกต่างกัน ดังนี้

การประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลด้วยฟังก์ชันโลดิสติก (L_x index) มีจำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบ ในแบบสอบความยาว 20, 40 และ 60 ข้อ จำนวน 194, 368 และ 356 คน คิดเป็นร้อยละ 16.40, 36.80 และ 35.60 ตามลำดับ

การประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลจากคะแนนส่วนที่เหลือ (W_1 index) เมื่อพิจารณาค่าสถิติ Infit มีจำนวนผู้สอบซึ่งไม่เหมาะสมกับโมเดลการตอบข้อสอบ ในแบบสอบความยาว 20, 40 และ 60 ข้อ จำนวน 44, 45 และ 53 คน คิดเป็นร้อยละ 4.40, 4.50 และ 5.30 ตามลำดับ และเมื่อพิจารณาค่าสถิติ Outfit มีจำนวนผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบจำนวน 34, 32 และ 44 คน คิดเป็นร้อยละ 3.40, 3.20 และ 4.40 ตามลำดับ

2.2 คุณภาพของแบบสอบหลังการวิเคราะห์ เป็นคุณภาพของแบบสอบหลังจากนำผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบออกจากการวิเคราะห์แล้ว

2.2.1 คำสารสนเทศของแบบสอบ

คำสารสนเทศของแบบสอบเป็นคุณภาพของแบบสอบตามทฤษฎีการตอบสนองข้อสอบ ได้จากผลรวมคำสารสนเทศของข้อสอบ การเปรียบเทียบคุณภาพของแบบสอบสามารถคำนวณค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ ซึ่งผู้วิจัยได้คำนวณค่าประสิทธิภาพสัมพัทธ์ของแบบสอบในแต่ละระดับความสามารถก่อนและหลังการตรวจสอบความเหมาะสมของบุคคล กับโมเดลการตอบข้อสอบ ปรากฏผลดังตารางที่ 12 ถึง 13 และแผนภูมิที่ 4 ถึง 6

ตารางที่ 12 ค่าสารสนเทศของแบบสอบที่ได้จากแบบสอบก่อนและหลังจากการวิเคราะห์บุคคลที่เหมาะสมกับโมเดลการตอบข้อสอบ

เงื่อนไข	ค่าสารสนเทศของแบบสอบ			
	ก่อนการวิเคราะห์	หลังการวิเคราะห์		
		L_2	W_1 (Infit)	W_1 (Outfit)
ผู้สอบระดับความสามารถต่ำ				
ความยาวของแบบสอบ				
20 ข้อ	4.251	4.690	4.216	4.544
40 ข้อ	7.372	9.923	7.441	7.601
60 ข้อ	12.089	15.440	11.633	11.728
ผู้สอบระดับความสามารถปานกลาง				
ความยาวของแบบสอบ				
20 ข้อ	7.210	8.533	7.624	7.785
40 ข้อ	5.476	7.821	6.471	6.421
60 ข้อ	11.137	15.211	10.998	10.734
ผู้สอบระดับความสามารถสูง				
ความยาวของแบบสอบ				
20 ข้อ	4.739	6.248	4.507	4.538
40 ข้อ	6.506	9.104	6.632	6.730
60 ข้อ	12.201	19.856	11.417	11.541

คุณภาพของแบบสอบตามทฤษฎีการตอบสนองข้อสอบ ซึ่งวัดได้จากค่าสารสนเทศของแบบสอบ ในเงื่อนไขต่างๆ ทั้งคุณภาพของแบบสอบก่อนและหลังวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีบ่งชี้ความเหมาะสมของบุคคล เมื่อระดับความยาวของแบบสอบ ระดับความสามารถของผู้สอบ และวิธีประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลต่างกัน มีค่าแตกต่างกันดังนี้

เมื่อผู้สอบมีระดับความสามารถต่ำ แบบสอบความยาว 20,40 และ 60 ข้อ ค่าสารสนเทศของแบบสอบก่อนการวิเคราะห์แบบแผนการตอบข้อสอบ มีค่าเท่ากับ 4.251, 7.372 และ 12.089 ค่าสารสนเทศของแบบสอบซึ่งวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี L_x เท่ากับ 4.690, 9.923 และ 12.089 ค่าสารสนเทศของแบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Infit})$ เท่ากับ 4.216, 7.441 และ 11.633 ตามลำดับ ค่าสารสนเทศของแบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Outfit})$ มีค่าเท่ากับ 4.544, 7.601 และ 11.728 ตามลำดับ

เมื่อผู้สอบมีระดับความสามารถปานกลาง แบบสอบความยาว 20,40 และ 60 ข้อ ค่าสารสนเทศของแบบสอบก่อนการวิเคราะห์แบบแผนการตอบข้อสอบมีค่าเท่ากับ 7.210, 5.476 และ 11.137 ค่าสารสนเทศของแบบสอบซึ่งวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี L_x เท่ากับ 8.533, 7.821 และ 15.211 ค่าสารสนเทศของแบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Infit})$ เท่ากับ 7.624, 6.471 และ 10.998 ตามลำดับ และค่าสารสนเทศของแบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Outfit})$ เท่ากับ 7.785, 6.421 และ 10.734 ตามลำดับ

เมื่อผู้สอบมีระดับความสามารถสูง แบบสอบความยาว 20, 40 และ 60 ข้อ ค่าสารสนเทศของแบบสอบก่อนการวิเคราะห์แบบแผนการตอบข้อสอบ มีค่าเท่ากับ 4.739, 6.506 และ 12.201 ค่าสารสนเทศของแบบสอบซึ่งวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี L_x เท่ากับ 6.248, 9.104 และ 19.856 ค่าสารสนเทศของแบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Infit})$ เท่ากับ 4.507, 6.632 และ 11.417 และค่าสารสนเทศของแบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Outfit})$ มีค่าเท่ากับ 4.538, 6.730 และ 11.541 ตามลำดับ

ตารางที่ 13 ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบจากการตรวจสอบความเหมาะสมของบุคคล
ด้วยดัชนีแอลเซตและดัชนีดับเบิลยูวัน

เงื่อนไข	ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ		
	L_2	W_1 (Infit)	W_1 (Outfit)
ผู้สอบระดับความสามารถต่ำ			
ความยาวของแบบสอบ			
20 ข้อ	1.103	0.992	1.069
40 ข้อ	1.346	1.009	1.031
60 ข้อ	1.277	0.962	0.970
ผู้สอบระดับความสามารถปานกลาง			
ความยาวของแบบสอบ			
20 ข้อ	1.183	1.057	1.079
40 ข้อ	1.428	1.182	1.173
60 ข้อ	1.366	0.988	0.964
ผู้สอบระดับความสามารถสูง			
ความยาวของแบบสอบ			
20 ข้อ	1.318	0.951	0.958
40 ข้อ	1.399	1.019	1.034
60 ข้อ	1.627	0.936	0.946

ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบซึ่งได้จากการตรวจสอบความเหมาะสมของบุคคล
กับโมเดลการตอบข้อสอบ คำนวณจากอัตราส่วนของค่าสารสนเทศของแบบสอบหลังการตรวจ

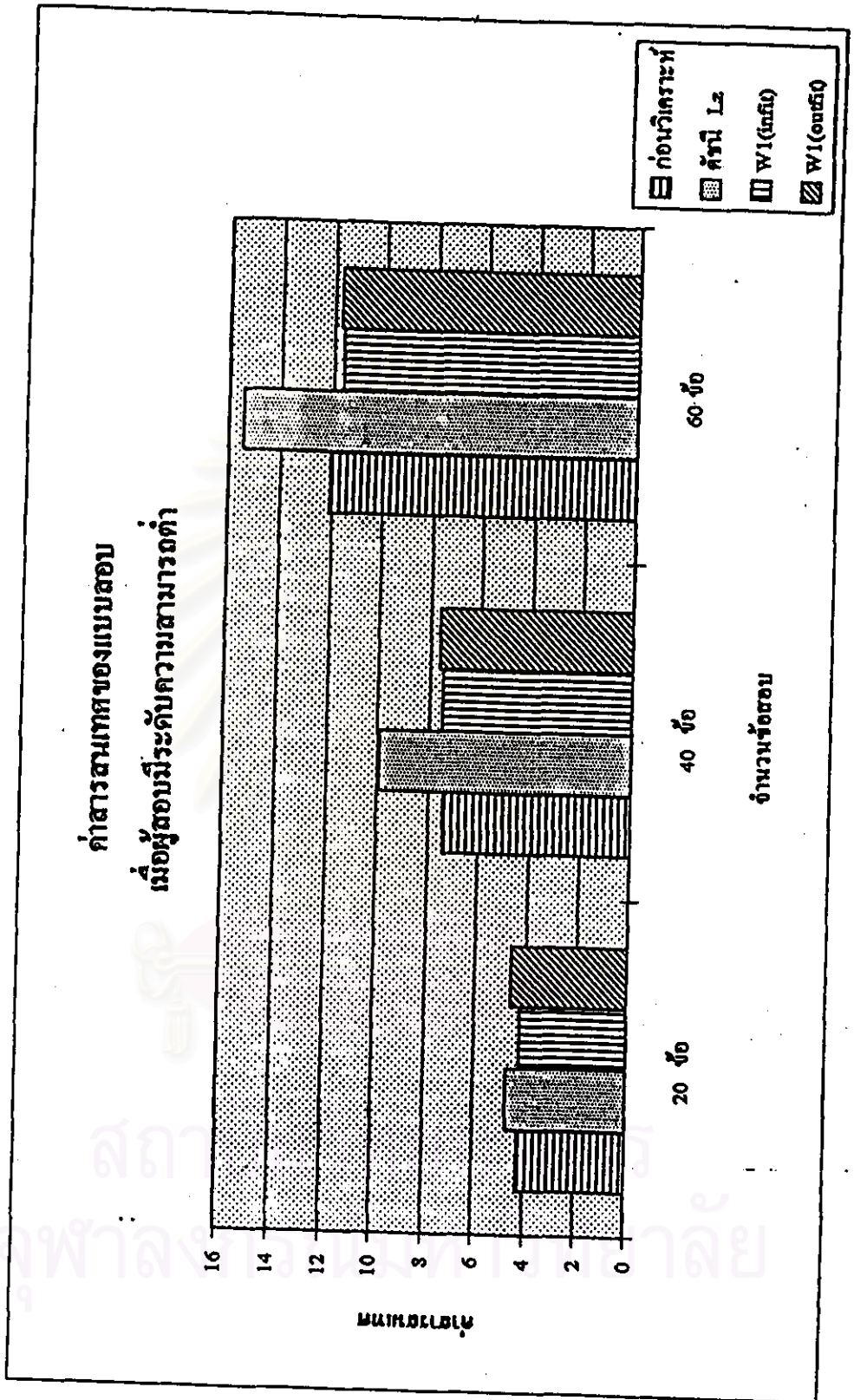
สอบความเหมาะสมของบุคคลและก่อนการวิเคราะห์ ด้วยดัชนีซึ่งมีวิธีประมาณค่าต่างกัน ในเงื่อนไขต่างๆ 9 เงื่อนไข ปรากฏผล ดังนี้

เมื่อผู้สอบมีระดับความสามารถต่ำ แบบสอบความยาว 20,40 และ 60 ข้อ ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบซึ่งวิเคราะห์แบบแผนการตอบข้อสอบดัชนี L_x เท่ากับ 1.103, 1.346, และ 1.277 ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Infit})$ เท่ากับ 0.992, 1.009 และ 0.962 ตามลำดับ ค่าประสิทธิภาพสัมพัทธ์แบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Outfit})$ มีค่าเท่ากับ 1.069, 1.031 และ 0.970 ตามลำดับ

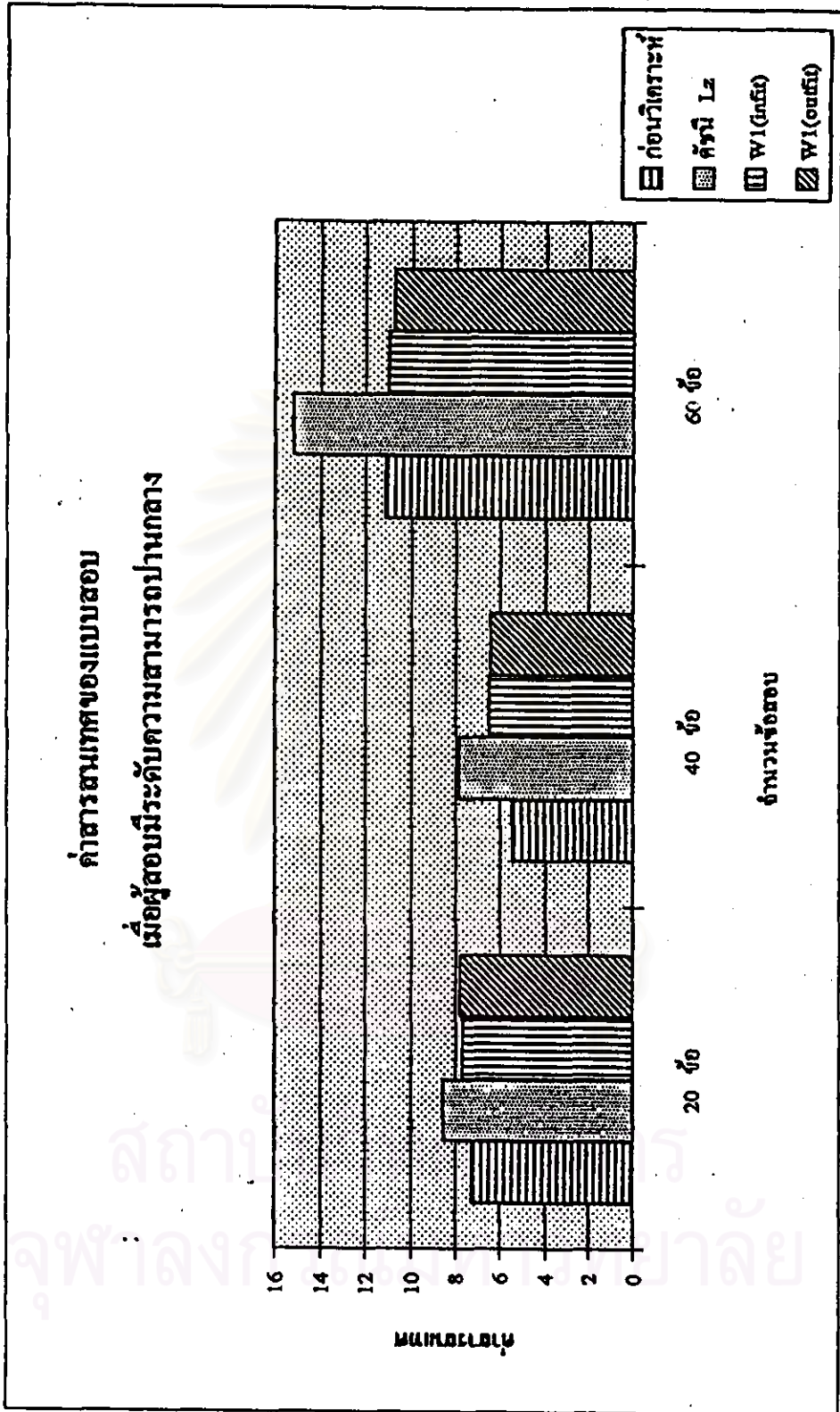
เมื่อผู้สอบมีระดับความสามารถปานกลาง แบบสอบความยาว 20,40 และ 60 ข้อ ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบซึ่งวิเคราะห์แบบแผนการตอบข้อสอบดัชนี L_x เท่ากับ 1.183, 1.428, และ 1.366 ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Infit})$ เท่ากับ 1.057, 1.182 และ 0.988 ตามลำดับ ค่าประสิทธิภาพสัมพัทธ์แบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Outfit})$ มีค่าเท่ากับ 1.079, 1.173 และ 0.964 ตามลำดับ

เมื่อผู้สอบมีระดับความสามารถสูง แบบสอบความยาว 20, 40 และ 60 ข้อ ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบซึ่งวิเคราะห์แบบแผนการตอบข้อสอบดัชนี L_x เท่ากับ 1.318, 1.399, และ 1.627 ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Infit})$ เท่ากับ 0.951, 1.019 และ 0.936 ตามลำดับ ค่าประสิทธิภาพสัมพัทธ์แบบสอบซึ่งวิเคราะห์ด้วยดัชนี $W_1(\text{Outfit})$ มีค่าเท่ากับ 0.968, 1.034 และ 0.946 ตามลำดับ

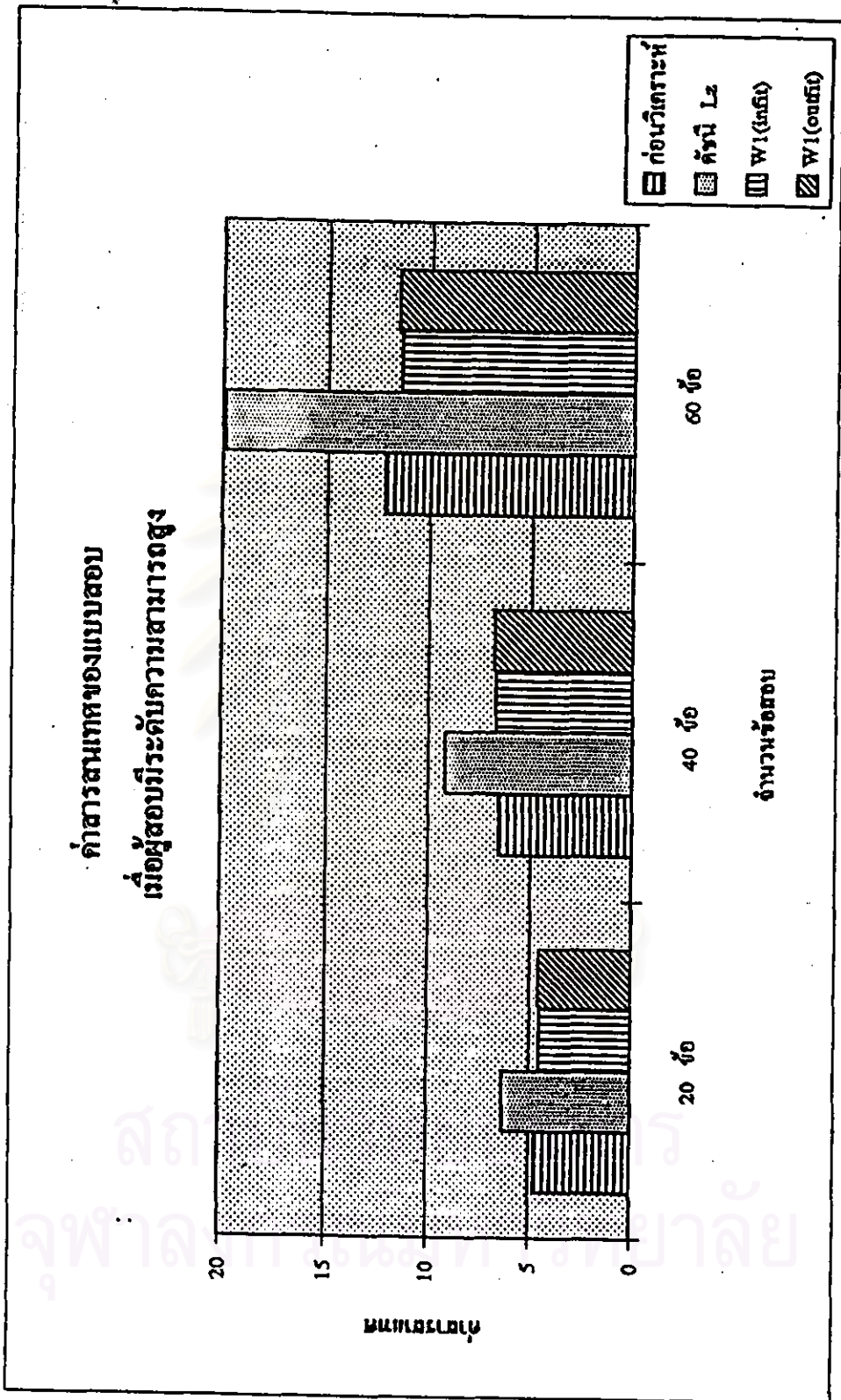
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



แผนภูมิที่ 4 ค่าสารสนเทศของแบบสอบเมื่อผู้สอบมีระดับความสามารถต่ำ



แผนภูมิที่ 5 ค่าสารสนเทศของแบบสอบเมื่อผู้สอบมีระดับความสามารถปานกลาง



แผนภูมิที่ 6 ค่าสารสนเทศของแบบสอบเมื่อผู้สอบมีระดับความสามารถสูง

2.2.2 ค่าความเที่ยงของแบบสอบ

ผู้วิจัยคำนวณค่าความเที่ยงของแบบสอบ หลังการตรวจสอบความเหมาะสมของผู้สอบ ปรากฏ ดังตารางที่ 14

ตารางที่ 14 ผลการวิเคราะห์คุณภาพของแบบสอบด้านความเที่ยงภายหลังการตรวจสอบบุคคลที่เหมาะสมกับโมเดลการตอบข้อสอบ

เงื่อนไข	ความเที่ยง		
	ดัชนี L_z	ดัชนี W_r	
		ค่าสถิติ Infit	ค่าสถิติ Outfit
ผู้สอบระดับความสามารถต่ำ			
ความยาวของแบบสอบ			
20 ข้อ	0.726	0.708	0.721
40 ข้อ	0.860	0.848	0.848
60 ข้อ	0.898	0.901	0.900
ผู้สอบระดับความสามารถปานกลาง			
ความยาวของแบบสอบ			
20 ข้อ	0.795	0.784	0.784
40 ข้อ	0.836	0.809	0.807
60 ข้อ	0.898	0.889	0.889
ผู้สอบระดับความสามารถสูง			
ความยาวของแบบสอบ			
20 ข้อ	0.808	0.770	0.771
40 ข้อ	0.836	0.783	0.780
60 ข้อ	0.909	0.880	0.879

คุณภาพของแบบสอบภายหลังการวิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนีบ่งชี้ความเหมาะสมของบุคคลที่มีวิธีประมาณค่าต่างกัน มีค่าความเที่ยงในแต่ละระดับความยาวของแบบสอบแตกต่างกัน

เมื่อผู้สอบมีระดับความสามารถต่ำ ในแบบสอบซึ่งมีความยาว 20, 40 และ 60 ข้อ แบบสอบซึ่งวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี L_2 มีค่าความเที่ยงเท่ากับ 0.726, 0.860 และ 0.898 ตามลำดับ ส่วนดัชนี W_1 เมื่อใช้ค่าสถิติ Infit เป็นเกณฑ์ในการพิจารณา มีค่าความเที่ยงเท่ากับ 0.708, 0.848 และ 0.898 ตามลำดับ และเมื่อใช้ค่าสถิติ Outfit เป็นเกณฑ์แบบสอบมีความเที่ยงเท่ากับ 0.721, 0.848 และ 0.900 ตามลำดับ

เมื่อผู้สอบมีระดับความสามารถปานกลาง ในแบบสอบซึ่งมีความยาว 20, 40 และ 60 ข้อ แบบสอบซึ่งวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี L_2 มีค่าความเที่ยงเท่ากับ 0.795, 0.838 และ 0.898 ตามลำดับ ส่วนดัชนี W_1 เมื่อใช้ค่าสถิติ Infit เป็นเกณฑ์ในการพิจารณา

มีค่าความเที่ยงเท่ากับ 0.784, 0.809 และ 0.889 ตามลำดับ และเมื่อใช้ค่าสถิติ Outfit เป็นเกณฑ์ในการพิจารณาแบบสอบมีค่าความเที่ยงเท่ากับ 0.784, 0.807 และ 0.899 ตามลำดับ

เมื่อผู้สอบมีระดับความสามารถสูง ในแบบสอบซึ่งมีความยาว 20, 40 และ 60 ข้อ แบบสอบซึ่งวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี L_2 มีค่าความเที่ยงเท่ากับ 0.808, 0.838 และ 0.909 ตามลำดับ ส่วนดัชนี W_1 เมื่อใช้ค่าสถิติ Infit เป็นเกณฑ์ในการพิจารณา มีค่าความเที่ยงเท่ากับ 0.770, 0.783 และ 0.880 ตามลำดับ และเมื่อใช้ค่าสถิติ Outfit เป็นเกณฑ์ในการพิจารณาแบบสอบมีค่าความเที่ยงเท่ากับ 0.771, 0.780 และ 0.879 ตามลำดับ

ตอนที่ 3 ผลการเปรียบเทียบผลการตรวจสอบความเหมาะสมของบุคคล

3.1 การเปรียบเทียบค่าสารสนเทศของแบบสอบ

การเปรียบเทียบค่าสารสนเทศของแบบสอบ สามารถคำนวณได้จากอัตราส่วนระหว่างค่าสารสนเทศของแบบสอบต่างชุด ณ ระดับความสามารถเดียวกัน เรียกว่าค่าประสิทธิภาพสัมพัทธ์ (Relative Efficiency : RE (Y,X) ค่า RE (Y,X) ที่มากกว่า 1.00 แสดงให้เห็นว่าแบบสอบชุด Y มีคุณภาพดีกว่าชุด X ผลการวิเคราะห์ปรากฏ ดังตารางที่ 15

ตารางที่ 15 ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบสอบจากการตรวจสอบความเหมาะสมของบุคคลด้วยดัชนีแอลซัดและดัชนีดับเบิลยูวัน

เงื่อนไข	ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ		
	L_2	W_1 (Infit)	W_1 (Outfit)
ผู้สอบระดับความสามารถต่ำ			
ความยาวของแบบสอบ			
20 ข้อ	1.103*	0.992	1.069*
40 ข้อ	1.346*	1.009*	1.031*
60 ข้อ	1.277*	0.962	0.970
ผู้สอบระดับความสามารถปานกลาง			
ความยาวของแบบสอบ			
20 ข้อ	1.183*	1.057*	1.079*
40 ข้อ	1.428*	1.182*	1.173*
60 ข้อ	1.366*	0.988	0.964
ผู้สอบระดับความสามารถสูง			
ความยาวของแบบสอบ			
20 ข้อ	1.318*	0.951	0.958
40 ข้อ	1.399*	1.019*	1.034*
60 ข้อ	1.627*	0.936	0.946

* ค่าประสิทธิภาพสัมพัทธ์สูงกว่า 1.00

ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบซึ่งได้จากการตรวจสอบความเหมาะสมของบุคคลกับโมเดลการตอบข้อสอบ ซึ่งมีวิธีประมาณค่าต่างกัน ในเงื่อนไขต่างๆ 9 เงื่อนไข ปรากฏผล ดังนี้

แบบสอปซึ่งได้จากการตรวจสอบความเหมาะสมของบุคคลกับโมเดลการตอบข้อสอบด้วยดัชนี L_2 มีค่าประสิทธิภาพสัมพัทธ์สูงกว่า 1.00 ในทุกระดับความสามารถของผู้สอบและทุกระดับความยาวของแบบสอป แสดงว่าแบบสอปหลังการวิเคราะห์มีคุณภาพสูงขึ้น

แบบสอปซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี W_1 เมื่อพิจารณาค่าสถิติ Infit แบบสอปมีค่าสารสนเทศเพิ่มขึ้น จำนวน 4 เงื่อนไข กล่าวคือ มีค่าประสิทธิภาพสัมพัทธ์สูงกว่า 1.00 เมื่อผู้สอบระดับความสามารถต่ำ ความยาว 40 ข้อ ผู้สอบระดับความสามารถปานกลาง ความยาว 20 และ 40 ข้อ และเมื่อผู้สอบมีความสามารถสูง ความยาว 40 ข้อ ในเงื่อนไขอื่นๆ แบบสอปมีค่าสารสนเทศลดลง และเมื่อพิจารณาค่าสถิติ Outfit แบบสอปมีค่าสารสนเทศเพิ่มขึ้นจำนวน 5 เงื่อนไข คือ เมื่อผู้สอบระดับความสามารถต่ำ และระดับความสามารถปานกลาง แบบสอปยาว 20 และ 40 ข้อ และ เมื่อผู้สอบมีระดับความสามารถสูง แบบสอปมีความยาว 40 ข้อ

โดยสรุปแบบสอปซึ่งวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี L_2 ให้สารสนเทศสูงที่สุดลงมาคือ ดัชนี W_1 เมื่อพิจารณาค่าสถิติ Outfit ส่วนดัชนี W_1 เมื่อพิจารณาค่าสถิติ Infit ให้สารสนเทศต่ำสุด

ทั้งนี้มิขอสังเกตว่า สารสนเทศของแบบสอปซึ่งวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี W_1 เมื่อผู้สอบมีระดับความสามารถต่ำ และระดับความสามารถสูง แบบสอปสั้น (20 ข้อ) และแบบสอปยาว (60 ข้อ) มีแนวโน้มจะให้สารสนเทศต่ำกว่าแบบสอปก่อนการวิเคราะห์แบบแผนการตอบข้อสอบ

3.2 การเปรียบเทียบค่าความเที่ยงของแบบสอป

ความเที่ยงของแบบสอปในแต่ละเงื่อนไข ก่อนและหลังการตรวจสอบบุคคลที่เหมาะสม คำนวณด้วยสูตร KR-20 และทดสอบความแตกต่างด้วยค่าสถิติพีชเชอร์-ซี ปรากฏผลดังตารางที่ 16 ถึง 21

ตารางที่ 16 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงก่อนและหลังการตรวจ
สอบผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนี L_2

เงื่อนไข	ความเที่ยง		ค่า Fisher-Z
	ก่อนการวิเคราะห์	หลังการวิเคราะห์	
ผู้สอบระดับความสามารถต่ำ			
ความยาวของแบบสอบ			
20 ข้อ	0.716	0.726	0.0671
40 ข้อ	0.846	0.860	0.2194
60 ข้อ	0.898	0.898	0.0000
ผู้สอบระดับความสามารถปานกลาง			
ความยาวของแบบสอบ			
20 ข้อ	0.782	0.795	0.1020
40 ข้อ	0.807	0.836	0.3785
60 ข้อ	0.888	0.898	0.2349
ผู้สอบระดับความสามารถสูง			
ความยาวของแบบสอบ			
20 ข้อ	0.778	0.808	0.2361
40 ข้อ	0.783	0.836	0.6624
60 ข้อ	0.883	0.909	0.7047

* $p < .05$

ความเที่ยงของแบบสอบภายหลังการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี L_2
เพื่อคัดเลือกผู้สอบที่เหมาะสม เมื่อผู้สอบมีระดับความสามารถปานกลาง และความสามารถสูง
มีค่าสูงขึ้นในทุกระดับความสามารถของผู้สอบ และทุกระดับความยาวของแบบสอบ ส่วนค่า

ความเที่ยงของแบบสอบเมื่อผู้สอบมีระดับความสามารถต่ำ แบบสอบความยาว 20 และ 40 ข้อ มีค่าความเที่ยงสูงขึ้น ส่วนความยาว 60 ข้อ มีค่าความเที่ยงเท่าเดิม เมื่อทดสอบค่านัยสำคัญทางสถิติ ปรากฏว่าความเที่ยงของแบบสอบหลังการวิเคราะห์ไม่แตกต่างจากความเที่ยงของแบบสอบก่อนการวิเคราะห์ห้อย่างมีนัยสำคัญทางสถิติที่ระดับ .05



สถาบันวิจัยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 17 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงก่อนและหลังการวิเคราะห์ ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนี W_1 โดยพิจารณาค่าสถิติ Infit

เงื่อนไข	ความเที่ยง		ค่า Fisher-Z
	ก่อนการวิเคราะห์	หลังการวิเคราะห์	
ผู้สอบระดับความสามารถต่ำ			
ความยาวของแบบสอบ			
20 ข้อ	0.716	0.708	-0.0408
40 ข้อ	0.846	0.848	0.0301
60 ข้อ	0.898	0.901	0.0801
ผู้สอบระดับความสามารถปานกลาง			
ความยาวของแบบสอบ			
20 ข้อ	0.782	0.784	0.0146
40 ข้อ	0.807	0.809	0.0215
60 ข้อ	0.888	0.889	0.0374
ผู้สอบระดับความสามารถสูง			
ความยาวของแบบสอบ			
20 ข้อ	0.778	0.770	-0.0583
40 ข้อ	0.783	0.783	0.0000
60 ข้อ	0.883	0.880	0.0690

* $p < .05$

ความเที่ยงของแบบสอบภายหลังการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี W_1 โดยใช้ค่าสถิติ Infit เป็นเกณฑ์ในการพิจารณา เมื่อผู้สอบมีระดับความสามารถปานกลาง

ทุกระดับความยาวของแบบสอบมีค่าสูงขึ้น เมื่อผู้สอบมีระดับความสามารถต่ำ แบบสอบซึ่งมีความยาว 40 และ 60 ข้อ มีค่าความเที่ยงสูงขึ้น ส่วนแบบสอบซึ่งมีความยาว 20 ข้อ มีค่าความเที่ยงลดลง เมื่อผู้สอบมีระดับความสามารถสูง ค่าความเที่ยงของแบบสอบความยาว 20 และ 60 ข้อมีค่าลดลง ส่วนแบบสอบความยาว 40 ข้อมีค่าความเที่ยงเท่าเดิม

เมื่อทดสอบนัยสำคัญทางสถิติ ปรากฏว่า ความเที่ยงของแบบสอบในทุกระดับความยาวของแบบสอบและทุกระดับความสามารถของผู้สอบ ก่อนและหลังการวิเคราะห์ผู้สอบที่เหมาะสม มีค่าความเที่ยงไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 18 ผลการเปรียบเทียบคุณภาพของแบบสอบถามด้านความเที่ยงก่อนและหลังการวิเคราะห์ ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนี W, โดยพิจารณาค่าสถิติ Outfit

เงื่อนไข	ความเที่ยง		ค่า Fisher-Z
	ก่อนการวิเคราะห์	หลังการวิเคราะห์	
ผู้สอบระดับความสามารถต่ำ			
ความยาวของแบบสอบ			
20 ข้อ	0.716	0.721	0.3003
40 ข้อ	0.846	0.848	0.0301
60 ข้อ	0.898	0.900	0.0534
ผู้สอบระดับความสามารถปานกลาง			
ความยาวของแบบสอบ			
20 ข้อ	0.782	0.784	0.0146
40 ข้อ	0.807	0.807	0.0000
60 ข้อ	0.888	0.889	0.0374
ผู้สอบระดับความสามารถสูง			
ความยาวของแบบสอบ			
20 ข้อ	0.778	0.770	0.0496
40 ข้อ	0.783	0.780	-0.0344
60 ข้อ	0.883	0.879	-0.0908

* $p < .05$

ความเที่ยงของแบบสอบภายหลังการวิเคราะห์แบบแผนการตอบข้อสอบ ด้วยดัชนี W_1 โดยใช้ค่าสถิติ $infitt$ เป็นเกณฑ์ในการพิจารณา เมื่อผู้สอบมีระดับความสามารถต่ำ ในทุกระดับความยาวของแบบสอบ มีค่าความเที่ยงสูงขึ้น เมื่อผู้สอบมีระดับความสามารถปานกลาง แบบสอบความยาว 20 และ 60 ข้อ มีค่าความเที่ยงเพิ่มขึ้นส่วนแบบสอบความยาว 40 ข้อ มีความเที่ยงเท่าเดิม เมื่อผู้สอบมีระดับความสามารถสูง แบบสอบความยาว 20 ข้อ มีค่าความเที่ยงเพิ่มขึ้น ส่วนแบบสอบซึ่งมีความยาว 40 และ 60 ข้อ มีค่าความเที่ยงลดลง

เมื่อทดสอบนัยสำคัญทางสถิติ ปรากฏว่า ความเที่ยงของแบบสอบในทุกระดับความยาวของแบบสอบ และทุกระดับความสามารถของผู้สอบ ก่อนและหลังการวิเคราะห์ผู้สอบที่เหมาะสม มีความเที่ยงไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 19 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงภายหลังการวิเคราะห์ผู้สอบ
ที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนี L_2 กับดัชนี W_1 (Infit)

เงื่อนไข	ความเที่ยง		ค่า Fisher-Z
	ดัชนี L_2	ดัชนี W_1 (Infit)	
ผู้สอบระดับความสามารถต่ำ			
ความยาวของแบบสอบ			
20 ข้อ	0.726	0.708	0.1079
40 ข้อ	0.860	0.848	0.1892
60 ข้อ	0.898	0.901	0.0801
ผู้สอบระดับความสามารถปานกลาง			
ความยาวของแบบสอบ			
20 ข้อ	0.795	0.784	0.0875
40 ข้อ	0.836	0.809	0.3570
60 ข้อ	0.898	0.889	0.2402
ผู้สอบระดับความสามารถสูง			
ความยาวของแบบสอบ			
20 ข้อ	0.808	0.770	0.2945
40 ข้อ	0.836	0.783	0.6624
60 ข้อ	0.909	0.880	0.7741

* $p < .05$

คุณภาพของแบบสอบด้านความเที่ยง ภายหลังการวิเคราะห์แบบแผนการตอบข้อ
สอบด้วยดัชนีหนึ่งชี้ ความเหมาะสมของบุคคลด้วย L_2 เมื่อผู้สอบมีระดับความสามารถต่ำ และ

ระดับความสามารถสูง ในทุกระดับความยาวของแบบสอบ มีค่าความเที่ยงของแบบสอบไม่สูง
กว่าการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี W , เมื่อใช้ค่าสถิติ $Infitt$ เป็นเกณฑ์ในการ
พิจารณา อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

คุณภาพของแบบสอบด้านความเที่ยงภายหลังการวิเคราะห์แบบแผนการตอบข้อสอบ
เมื่อผู้สอบมีระดับความสามารถปานกลาง ในทุกระดับความยาวของแบบสอบ ด้วยดัชนี L_2 และ
ดัชนี W , เมื่อใช้ค่าสถิติ $Infitt$ เป็นเกณฑ์ในการพิจารณา มีค่าความเที่ยงไม่แตกต่างกันอย่างมี
นัยสำคัญทางสถิติที่ระดับ .05



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 20 ผลการเปรียบเทียบคุณภาพของแบบสอบถามด้านความเที่ยงภายหลังการวิเคราะห์ผู้ตอบ
ที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนี L_2 กับดัชนี W_1 (Outfit)

เงื่อนไข	ความเที่ยง		ค่า Fisher-Z
	ดัชนี L_2	ดัชนี W_1 (outfit)	
ผู้สอบระดับความสามารถต่ำ			
ความยาวของแบบสอบ			
20 ข้อ	0.726	0.721	0.0291
40 ข้อ	0.860	0.848	0.1892
60 ข้อ	0.898	0.900	0.0534
ผู้สอบระดับความสามารถปานกลาง			
ความยาวของแบบสอบ			
20 ข้อ	0.795	0.784	0.0875
40 ข้อ	0.836	0.807	0.3785
60 ข้อ	0.898	0.889	0.2136
ผู้สอบระดับความสามารถสูง			
ความยาวของแบบสอบ			
20 ข้อ	0.808	0.771	0.2857
40 ข้อ	0.836	0.78	0.6968
60 ข้อ	0.909	0.879	0.7955

* $p < .05$

คุณภาพของแบบสอบถามด้านความเที่ยง ภายหลังการวิเคราะห์แบบแผนการตอบข้อ
สอบด้วยดัชนีบ่งชี้ ความเหมาะสมของบุคคลด้วย L_2 เมื่อผู้สอบมีระดับความสามารถต่ำ และ

ระดับความสามารถสูง ในทุกระดับความยาวของแบบสอบ มีค่าความเที่ยงของแบบสอบไม่สูง
กว่าการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี W_1 เมื่อใช้ค่าสถิติ $Outfit$ เป็นเกณฑ์ในการ
พิจารณา อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

คุณภาพของแบบสอบด้านความเที่ยงภายหลังการวิเคราะห์แบบแผนการตอบข้อสอบ
เมื่อผู้สอบมีระดับความสามารถปานกลาง ในทุกระดับความยาวของแบบสอบ ด้วยดัชนี L_2 และ
ดัชนี W_1 เมื่อใช้ค่าสถิติ $Outfit$ เป็นเกณฑ์ในการพิจารณา มีค่าความเที่ยงไม่แตกต่างกันอย่าง
มีนัยสำคัญทางสถิติที่ระดับ .05



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 21 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงภายหลังการตรวจสอบผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยค่าสถิติในวิธีประมาณค่าดัชนีจากคะแนนส่วนที่เหลือ

เงื่อนไข	ความเที่ยง		ค่า Fisher-Z
	ค่าสถิติ Infit	ค่าสถิติ Outfit	
ผู้สอบระดับความสามารถต่ำ			
ความยาวของแบบสอบ			
20 ข้อ	0.708	0.721	0.0787
40 ข้อ	0.848	0.848	0.0000
60 ข้อ	0.901	0.900	0.0267
ผู้สอบระดับความสามารถปานกลาง			
ความยาวของแบบสอบ			
20 ข้อ	0.784	0.784	0.0000
40 ข้อ	0.809	0.807	0.0215
60 ข้อ	0.889	0.889	0.0000
ผู้สอบระดับความสามารถสูง			
ความยาวของแบบสอบ			
20 ข้อ	0.770	0.771	0.0087
40 ข้อ	0.783	0.780	0.0344
60 ข้อ	0.880	0.879	0.0214

* $p < .05$

คุณภาพของแบบสอบด้านความเที่ยง ภายหลังการวิเคราะห์ข้อสอบด้วยดัชนีป้องกันความเหมาะสมของบุคคลด้วยดัชนี W_1 เมื่อใช้ค่าสถิติ Infit และค่าสถิติ Outfit เป็นเกณฑ์ใน

การพิจารณา ในทุกระดับความสามารถของผู้สอบและทุกระดับความยาวของแบบสอบ มีค่าความเที่ยงไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.3 การเปรียบเทียบความตรงตามทฤษฎีของแบบสอบ

ผู้วิจัยตรวจสอบความตรงตามทฤษฎีของแบบสอบ ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน ด้วยโปรแกรม LISREL 8.10 โดยผู้วิจัยได้วิเคราะห์ข้อสอบที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยค่าสถิติไค-สแควร์ จากโปรแกรม BILOG นำเฉพาะข้อสอบที่เหมาะสมมาวิเคราะห์ความตรงเชิงทฤษฎี และเนื่องจาก แบบสอบมีความยาวเกินกว่า 40 ข้อ ผู้วิจัยจึงสุ่มข้อสอบในแบบสอบความยาว 60 ข้อ ออกเงื่อนไขละ 20 ข้อ เพื่อให้สามารถวิเคราะห์ได้ ผลการวิเคราะห์ปรากฏดังตารางที่ 22 ถึง ตารางที่ 24



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 22 ผลการเปรียบเทียบคุณภาพของแบบสอบถามด้านความตรงทฤษฎี เมื่อผู้ตอบมีระดับความสามารถต่ำ

เงื่อนไข	ดัชนีของความตรง						
	χ^2	df	P	GFI	AGFI	RMR	χ^2 / df
แบบสอบถามยาว 20 ข้อ							
วิธีประมาณค่าดัชนี							
L_2	60.13	72	.840	.991	.987	.004	.835
W_1 (Infit)	79.71	114	.994	.988	.984	.004	.699
W_1 (Outfit)	68.88	85	.898	.990	.987	.004	.810
แบบสอบถามยาว 40 ข้อ							
วิธีประมาณค่าดัชนี							
L_2	515.96	523	.523	.579	.963	.005	.987
W_1 (infit)	604.28	605	.501	.967	.962	.005	.999
W_1 (Outfit)	597.35	605	.580	.968	.962	.005	.987
แบบสอบถามยาว 60 ข้อ							
วิธีประมาณค่าดัชนี							
L_2	198.73	208	.666	.976	.971	.006	.956
W_1 (Infit)	117.62	127	.713	.986	.981	.005	.926
W_1 (Outfit)	496.76	506	.607	.970	.965	.006	.981

ผลการเปรียบเทียบดัชนีความตรงจากการวิเคราะห์องค์ประกอบเชิงยืนยัน ของแบบสอบถามซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีบ่งชี้ความเหมาะสมของบุคคลซึ่งประมาณค่าที่แตกต่างกัน ปรากฏว่า เมื่อผู้ตอบมีระดับความสามารถต่ำ แบบสอบถามทุกระดับความยาว มีความตรงตามทฤษฎี เนื่องจากค่าไค - สแควร์ มีค่าแตกต่างจากศูนย์ อย่างไม่มี

นัยสำคัญทางสถิติ ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าใกล้เคียงกันและมีค่าเข้าใกล้ 1 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

เมื่อแบบสอบถามยาว 20 ข้อ ค่าไค-สแควร์สัมพันธ์ ของแบบสอบถามซึ่งได้จากการวิเคราะห์ด้วยดัชนี W_1 เมื่อพิจารณาค่าสถิติ Infit มีค่าต่างจาก 1 มากที่สุด ซึ่งตามหลักการแล้วโมเดลที่มีความตรง ควรมีค่าไค-สแควร์สัมพันธ์เข้าใกล้ 1 (Saris & Stronkhorst, 1984: อ้างถึงใน นงลักษณ์ วิรัชชัย, 2538 : 46) ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 และเมื่อพิจารณาดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) พบว่ามีค่าเท่ากับ 0.004 แสดงว่าแบบสอบถามซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบมีความตรงตามทฤษฎีไม่แตกต่างกัน ทั้งนี้ถ้าพิจารณาค่า ไค-สแควร์สัมพันธ์ โมเดลที่เหมาะสมที่สุดคือ โมเดลของแบบสอบถามซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีแอลแซด (L_1) ($\chi^2/df=0.835$)

เมื่อแบบสอบถามยาว 40 ข้อ แบบสอบถามซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีทั้ง 2 วิธี มีความสอดคล้องกับข้อมูลเชิงประจักษ์ เนื่องจาก ค่าไค-สแควร์ สัมพันธ์มีค่าใกล้เคียงกันและมีค่าเข้าใกล้ 1 ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 และเมื่อพิจารณาค่าดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) ทุกเงื่อนไขมีค่าเท่ากับ 0.005 ซึ่งมีค่าเข้าใกล้ 0 แสดงว่าแบบสอบถามซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบมีความตรงตามทฤษฎีไม่แตกต่างกัน ทั้งนี้ถ้าพิจารณาค่า ไค - สแควร์สัมพันธ์ โมเดลที่เหมาะสมที่สุดคือ โมเดลของแบบสอบถามซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี W_1 (Infit) ($\chi^2/df=0.999$)

เมื่อแบบสอบถามยาว 60 ข้อ ค่าไค-สแควร์สัมพันธ์มีค่าเข้าใกล้ 1 และมีค่าใกล้เคียงกัน ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 ค่าดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเข้าใกล้ศูนย์ แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ สรุปได้ว่าแบบสอบถามซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีทั้ง 2 วิธี มีความตรงตามทฤษฎีไม่แตกต่างกัน ทั้งนี้ถ้าพิจารณาค่า ไค - สแควร์สัมพันธ์ โมเดลที่เหมาะสมที่สุดคือ โมเดลของแบบสอบถามซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี W_1 (Outfit) ($\chi^2/df=0.981$)

ตารางที่ 23 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความตรงตามทฤษฎีเมื่อผู้สอบมีระดับความสามารถปานกลาง

เงื่อนไข	ดัชนีของความตรง						
	χ^2	df	P	GFI	AGFI	RMR	χ^2 / df
แบบสอบความยาว 20 ข้อ							
วิธีประมาณค่าดัชนี							
L_2	9.25	24	.997	.998	.995	.001	.385
W_1 (Infit)	11.19	30	.999	.998	.996	.002	.373
W_1 (Outfit)	10.77	24	.991	.998	.995	.001	.458
แบบสอบความยาว 40 ข้อ							
วิธีประมาณค่าดัชนี							
L_2	646.08	656	.602	.955	.949	.005	.985
W_1 (Infit)	689.12	696	.566	.965	.961	.005	.990
W_1 (Outfit)	705.46	696	.393	.964	.963	.005	1.01
แบบสอบความยาว 60 ข้อ							
วิธีประมาณค่าดัชนี							
L_2	516.75	525	.593	.957	.949	.005	.984
W_1 (Infit)	515.60	526	.619	.972	.965	.004	.980
W_1 (Outfit)	489.31	499	.613	.971	.963	.004	.981

ผลการเปรียบเทียบดัชนีความตรงจากการวิเคราะห์องค์ประกอบเชิงยืนยัน ของแบบสอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีบ่งชี้ความเหมาะสมของบุคคลซึ่ง

ประมาณด้วยวิธีที่แตกต่างกัน ปรากฏว่า เมื่อผู้สอบมีระดับความสามารถปานกลาง แบบสอบทุกระดับความยาว มีความตรงตามทฤษฎี เนื่องจากค่าไค-สแควร์ มีค่าแตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

เมื่อแบบสอบความยาว 20 ข้อ ปรากฏว่าแบบสอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีทั้ง 2 วิธี มีความตรงตามทฤษฎี เนื่องจากค่าไค-สแควร์ มีค่าแตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ คือ มีค่าเข้าใกล้ 1 และมีค่าใกล้เคียงกัน แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 และเมื่อพิจารณาค่าดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) ทุกเงื่อนไขมีค่าเข้าใกล้ 0 ทั้งนี้ถ้าพิจารณาค่า ไค-สแควร์สัมพัทธ์ โมเดลที่เหมาะสมที่สุดคือ โมเดลของแบบสอบที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี W_1 (Infit) ($\chi^2/df=0.468$)

เมื่อแบบสอบความยาว 40 ข้อ แบบสอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีทั้ง 2 วิธี มีความสอดคล้องกับข้อมูลเชิงประจักษ์ เนื่องจาก ค่าไค-สแควร์ สัมพัทธ์มีค่าใกล้เคียงกันและมีค่าเข้าใกล้ 1 ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 และเมื่อพิจารณาค่าดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) ทุกเงื่อนไขมีค่าเท่ากันคือ เท่ากับ .005 ซึ่งมีค่าเข้าใกล้ 0 แสดงว่าแบบสอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบมีความตรงตามทฤษฎีไม่แตกต่างกัน ทั้งนี้ถ้าพิจารณาค่า ไค-สแควร์สัมพัทธ์ โมเดลที่เหมาะสมที่สุดคือ โมเดลของแบบสอบที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี W_1 (Infit) และ W_1 (Outfit) $\chi^2/df=0.99$ และ 1.01)

เมื่อแบบสอบความยาว 60 ข้อ ค่าไค-สแควร์สัมพัทธ์มีค่าเข้าใกล้ 1 และมีค่าใกล้เคียงกัน ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 ค่าดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเข้าใกล้ศูนย์ แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ สรุปได้ว่าแบบสอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีทั้ง 2 วิธี มีความตรงตามทฤษฎีไม่แตกต่างกัน ทั้งนี้ถ้าพิจารณาค่า ไค-สแควร์สัมพัทธ์ โมเดลที่เหมาะสมที่สุดคือ โมเดลของแบบสอบที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี L_2 ($\chi^2/df=0.981$)

ตารางที่ 24 ผลการเปรียบเทียบคุณภาพของแบบสอบถามด้านความตรงทฤษฎีเมื่อผู้ตอบมีระดับ
ความสามารถสูง

เงื่อนไข	ดัชนีของความตรง						
	χ^2	df	P	GFI	AGFI	RMR	χ^2 / df
แบบสอบถามยาว 20 ข้อ							
วิธีประมาณค่าดัชนี							
L_2	85.35	96	.774	.987	.982	.003	0.889
W_1 (Infit)	78.27	82	.596	.989	.984	.003	0.955
W_1 (Oufit)	82.62	82	.460	.989	.984	.003	1.007
แบบสอบถามยาว 40 ข้อ							
วิธีประมาณค่าดัชนี							
L_2	491.56	478	.324	.960	.947	.002	1.028
W_1 (Infi)	436.08	425	.345	.973	.966	.002	1.026
W_1 (Oufi)	480.60	463	.277	.971	.965	.002	1.038
แบบสอบถามยาว 60 ข้อ							
วิธีประมาณค่าดัชนี							
L_2	324.80	333	.616	.970	.953	.002	0.975
W_1 (Infit)	499.93	509	.605	.973	.962	.002	0.982
W_1 (Oufit)	477.16	478	.502	.973	.963	.003	0.998

ผลการเปรียบเทียบดัชนีความตรงจากการวิเคราะห์องค์ประกอบเชิงยืนยัน ของแบบ
สอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีบ่งชี้ความเหมาะสมของบุคคลซึ่ง

ประมาณด้วยวิธีที่แตกต่างกัน ปรากฏว่า เมื่อผู้สอบมีระดับความสามารถสูง แบบสอบทุกระดับความยาวมีความตรงตามทฤษฎี เนื่องจากค่าไค-สแควร์ มีค่าแตกต่างจากศูนย์ อย่างไม่มีนัยสำคัญทางสถิติ ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

เมื่อแบบสอบความยาว 20 ข้อ ค่าไค-สแควร์ สัมพัทธ์ ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 และเมื่อพิจารณาดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) พบว่ามีค่าเท่ากับหรือเท่ากับ 0.003 แสดงว่าแบบสอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบมีความตรงตามทฤษฎีไม่แตกต่างกัน ทั้งนี้ถ้าพิจารณาค่าไค-สแควร์สัมพัทธ์ โมเดลที่เหมาะสมที่สุดคือ โมเดลของแบบสอบที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี $W_1(\text{Infit})$ ($\chi^2/df=0.955$)

เมื่อแบบสอบความยาว 40 ข้อ แบบสอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีทั้ง 2 วิธี มีความสอดคล้องกับข้อมูลเชิงประจักษ์ เนื่องจาก ค่าไค-สแควร์ สัมพัทธ์มีค่าใกล้เคียงกันและมีค่าเข้าใกล้ 1 ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 และเมื่อพิจารณาค่าดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) ทุกเงื่อนไขมีค่าเท่ากับหรือเท่ากับ .002 ซึ่งมีค่าเข้าใกล้ 0 แสดงว่าแบบสอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบมีความตรงตามทฤษฎีไม่แตกต่างกัน ทั้งนี้ถ้าพิจารณาค่าไค-สแควร์สัมพัทธ์ โมเดลที่เหมาะสมที่สุดคือ โมเดลของแบบสอบ ที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี $W_1(\text{Infit})$ ($\chi^2/df=1.026$)

เมื่อแบบสอบความยาว 60 ข้อ ค่าไค-สแควร์สัมพัทธ์มีค่าเข้าใกล้ 1 และมีค่าใกล้เคียงกัน ดัชนีวัดระดับความกลมกลืน GFI และ AGFI มีค่าเข้าใกล้ 1 ค่าดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเข้าใกล้ศูนย์ แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ สรุปได้ว่าแบบสอบซึ่งได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีทั้ง 2 วิธี มีความตรงตามทฤษฎีไม่แตกต่างกัน ทั้งนี้ถ้าพิจารณาค่าไค-สแควร์สัมพัทธ์ โมเดลที่เหมาะสมที่สุดคือ โมเดลของแบบสอบที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนี $W_1(\text{Outfit})$ ($\chi^2/df=0.998$)