

บทที่ 4

ผลการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาอิทธิพลที่เกิดขึ้นจากการตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ ที่ส่งผลต่อคุณภาพของแบบทดสอบวัดผลฤทธิ์ทางการเรียนระดับชาติชั้นประถมศึกษาปีที่ 6 โดยเปรียบเทียบค่าความเที่ยง ความตรงเชิงโครงสร้าง ความคงที่ของโครงสร้างองค์ประกอบ และค่าสหสัมพันธ์อันดับของผู้สอบ ของแบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ

ผู้วิจัยนำเสนอผลการวิเคราะห์ข้อมูลเป็น 6 ตอน ดังนี้

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลเบื้องต้น

ตอนที่ 2 ผลการตรวจสอบคุณภาพของแบบทดสอบก่อนการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ได้แก่ ค่าความเที่ยง ความตรงเชิงโครงสร้าง และความคงที่ของโครงสร้างองค์ประกอบ

ตอนที่ 3 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติกและวิธีชีพเสถที่ปรับใหม่

ตอนที่ 4 ผลการตรวจสอบคุณภาพของแบบทดสอบ ฉบับหลังจากตัดข้อสอบที่ทำหน้าที่ต่างกันออกจากแบบทดสอบ ได้แก่ ค่าความเที่ยง ความตรงเชิงโครงสร้าง และความคงที่ของโครงสร้างองค์ประกอบ

ตอนที่ 5 ผลการเปรียบเทียบคุณภาพของแบบทดสอบ ฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ

ตอนที่ 6 ผลการวิเคราะห์ค่าสหสัมพันธ์อันดับของผู้สอบ ฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ

เพื่อให้การนำเสนอผลการวิเคราะห์ข้อมูลมีความสะดวกและเข้าใจผลการวิเคราะห์ข้อมูลตรงกัน ผู้วิจัยจึงกำหนดสัญลักษณ์และอักษรย่อที่ใช้แทนค่าสถิติและตัวแปรต่าง ๆ ในการนำเสนอ ดังนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

n	คือ	จำนวนกลุ่มตัวอย่าง (Sample Size)
M	คือ	ค่าเฉลี่ย (Mean)
SD	คือ	ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)
p	คือ	ค่าความยากง่ายของข้อสอบ (Difficulty of the Items)
r	คือ	ค่าอำนาจจำแนกของข้อสอบ (Discrimination Power of the Items)
t	คือ	ค่าแจกแจงแบบ t (t-Distribution)
df	คือ	ค่าองศาแห่งความเป็นอิสระ (Degree of Freedom)
CFA	คือ	การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis)
Known	คือ	องค์ประกอบด้านความรู้
Process	คือ	องค์ประกอบด้านกระบวนการ
Thai Language	คือ	แบบทดสอบวิชาภาษาไทย (ตัวแปรแฝงภายนอก;K)
DIF	คือ	ข้อสอบทำหน้าที่ย่อยต่างกัน
NO DIF	คือ	ข้อสอบทำหน้าที่ย่อยไม่ต่างกัน
χ^2	คือ	ค่าสถิติไค-สแควร์ (Chi-square)
GFI	คือ	ค่าดัชนีวัดระดับความกลมกลืน (Goodness of Fit Index)
AGFI	คือ	ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (Adjust Goodness of Fit Index)
CFI	คือ	ค่าดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (Comparative Fit Index)
NFI	คือ	ค่าดัชนีวัดความสอดคล้องปกติ (Normed Fit Index)
RMSEA	คือ	ค่ารากของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนโดยประมาณ (Root Mean Squared Error of Approximation)
R^2	คือ	สัมประสิทธิ์การถดถอย (Multiple Correlation Squared)
r_s	คือ	ค่าสัมประสิทธิ์สหสัมพันธ์อันดับของสเปียร์แมน (Spearman Rank Correlation Coefficient)
B	คือ	ดัชนีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ที่ได้จากการตรวจสอบโดยใช้วิธีถดถอยโลจิสติก
β_{uni}	คือ	ดัชนีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ที่ได้จากการตรวจสอบโดยใช้วิธีชิปเทสที่ปรับใหม่
g_{byx}	คือ	ปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบกับความสามารถ (คะแนนรวม)

ผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลเบื้องต้น

ผลการวิเคราะห์ข้อมูลเบื้องต้น ของคะแนนการสอบวัดผลสัมฤทธิ์ทางการเรียน ระดับชาติ ชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2546 วิชาภาษาไทย จากจำนวนกลุ่มตัวอย่าง 2,000 คน แยกเป็นเพศชาย 1,000 คน และเพศหญิง 1,000 คน แสดงดังตารางที่ 2

ตารางที่ 2 ลักษณะการแจกแจงของคะแนนวิชาภาษาไทย

ค่าสถิติ	เพศชาย ($n = 1000$)	เพศหญิง ($n = 1000$)	รวม ($N = 2000$)
ค่าเฉลี่ย	20.16	20.54	20.35
ส่วนเบี่ยงเบนมาตรฐาน	7.75	7.99	7.87
มัธยฐาน	19.00	20.00	20.00
ฐานนิยม	26.00	26.00	26.00
ค่าต่ำสุด	4.00	4.00	4.00
ค่าสูงสุด	38.00	38.00	38.00
พิสัย	34.00	34.00	34.00
ความเบ้	0.16	0.10	0.13
ความโด่ง	-0.98	-1.01	-1.00

จากตารางที่ 2 เมื่อจำแนกกลุ่มผู้สอบตามตัวแปรเพศ พบว่าเพศชายมีค่าเฉลี่ยของคะแนนการสอบ เท่ากับ 20.16 คะแนน คะแนนมีการกระจายออกจากค่าเฉลี่ยประมาณ 7.75 คะแนน มีพิสัยของคะแนนเท่ากับ 34 คะแนน มีคะแนนต่ำสุดเท่ากับ 4 คะแนน สูงสุดเท่ากับ 38 คะแนน มีฐานนิยมของคะแนนเท่ากับ 26 คะแนน ค่าความเบ้เท่ากับ 0.16 ซึ่งเมื่อเทียบกับเส้นโค้งปกติ (Normal Curve) ที่มีค่าความเบ้เท่ากับ 0 แสดงว่าผลการสอบมีลักษณะการแจกแจงเบ้ทางบวกเล็กน้อย (Positive Skewness) ใกล้เคียงกับการแจกแจงแบบปกติมาก นั่นคือผลการสอบส่วนใหญ่ นักเรียนได้คะแนนต่ำกว่าค่าเฉลี่ยเล็กน้อย และมีค่าความโด่งเท่ากับ -0.98 เมื่อเทียบกับเส้นโค้งปกติ (Normal Curve) ที่มีค่าเท่ากับ 3.00 แสดงว่ามีการแจกแจงความถี่ในลักษณะโค้งลาด (Platy Kurtic) ที่แบนกว่าปกติ

ผู้สอบเพศหญิงมีค่าเฉลี่ยของคะแนนการสอบ เท่ากับ 20.54 คะแนน โดยคะแนนมีการกระจายออกจากค่าเฉลี่ยประมาณ 7.99 คะแนน มีพิสัยของคะแนนเท่ากับ 34 คะแนน โดยมีคะแนนต่ำสุดเท่ากับ 4 คะแนน สูงสุดเท่ากับ 38 คะแนน โดยมีฐานนิยมของคะแนนเท่ากับ 26 คะแนน ค่าความเบ้เท่ากับ 0.10 ซึ่งเมื่อเทียบกับเส้นโค้งปกติ ที่มีค่าความเบ้เท่ากับ 0 แสดงว่าผลการสอบมีลักษณะการแจกแจงเบ้ทางบวกเล็กน้อย ใกล้เคียงกับการแจกแจงแบบปกติมาก นั่นคือผลการสอบส่วนใหญ่ นักเรียนได้คะแนนต่ำกว่าค่าเฉลี่ยเล็กน้อย และมีค่าความโค้งเท่ากับ -1.01 เมื่อเทียบกับเส้นโค้งปกติ ที่มีค่าเท่ากับ 3.00 แสดงว่ามีการแจกแจงความถี่ในลักษณะโค้งลาดที่แบนกว่าปกติ โดยค่าสถิติพื้นฐานระหว่างเพศชายกับเพศหญิงส่วนใหญ่มีค่าใกล้เคียงกันมาก

ผลการตอบรวมทั้งฉบับ (ทั้งชายและหญิง) ผู้สอบมีค่าเฉลี่ยของคะแนนการสอบ เท่ากับ 20.35 คะแนน โดยคะแนนมีการกระจายออกจากค่าเฉลี่ยประมาณ 7.87 คะแนน มีพิสัยของคะแนนเท่ากับ 34 คะแนน โดยมีคะแนนต่ำสุดเท่ากับ 4 คะแนน สูงสุดเท่ากับ 38 คะแนนคือได้คะแนนเต็ม โดยมีฐานนิยมของคะแนนเท่ากับ 26 คะแนน ค่าความเบ้เท่ากับ 0.13 ซึ่งเมื่อเทียบกับเส้นโค้งปกติ ที่มีค่าความเบ้เท่ากับ 0 แสดงว่าผลการสอบมีลักษณะการแจกแจงเบ้ทางบวกเล็กน้อย ใกล้เคียงกับการแจกแจงแบบปกติ นั่นคือผลการสอบส่วนใหญ่ นักเรียนได้คะแนนต่ำกว่าค่าเฉลี่ยเล็กน้อย และมีค่าความโค้งเท่ากับ -1.00 เมื่อเทียบกับเส้นโค้งปกติ ที่มีค่าเท่ากับ 3.00 แสดงว่ามีการแจกแจงความถี่ในลักษณะโค้งลาดที่แบนกว่าปกติ

ตอนที่ 2 ผลการตรวจสอบคุณภาพของแบบทดสอบก่อนการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ

2.1 ผลการตรวจสอบคุณภาพของข้อสอบวิชาภาษาไทยเป็นรายข้อ ค่าความยาก (p) ค่าอำนาจจำแนก (r) และค่าความเที่ยงแบบสัมประสิทธิ์แอลฟาของครอนบาค วิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิม โดยใช้โปรแกรมคอมพิวเตอร์ LERTAP 5 ในการวิเคราะห์ มีรายละเอียดแสดงดังตารางที่ 3

ตารางที่ 3 ค่าความยาก ค่าอำนาจจำแนก และค่าความเที่ยงแบบสัมประสิทธิ์แอลฟาครอนบาค
ของแบบทดสอบวิชาภาษาไทย

ข้อ ที่	ค่าความยาก (<i>p</i>)	ค่าอำนาจจำแนก (<i>r</i>)	ข้อ ที่	ค่าความยาก (<i>p</i>)	ค่าอำนาจจำแนก (<i>r</i>)
1	.55	.33	21	.51	.38
2	.68	.45	22	.33	.29
3	.65	.45	23	.44	.28
4	.45	.31	24	.35	.36
5	.51	.42	25	.31	.31
6	.44	.40	26	.62	.38
7	.31	.38	27	.56	.41
8	.63	.44	28	.47	.43
9	.63	.45	29	.38	.29
10	.52	.43	30	.36	.26
11	.56	.52	31	.65	.39
12	.57	.36	32	.54	.32
13	.50	.41	33	.61	.39
14	.46	.38	34	.42	.32
15	.55	.38	35	.56	.28
16	.55	.45	36	.68	.45
17	.54	.32	37	.65	.41
18	.48	.41	38	.58	.47
19	.57	.47	39	.57	.30
20	.41	.34	40	.40	.38
ค่าเฉลี่ยทั้งฉบับ			$p = .51$ $r = .38$		
ค่าความเที่ยงแบบสัมประสิทธิ์แอลฟา				.86	

จากตารางที่ 3 พบว่าแบบทดสอบมีค่าความเที่ยงเท่ากับ .86 ซึ่งอยู่ในระดับสูง เมื่อพิจารณาความยาก (*p*) จะเห็นว่าข้อสอบส่วนใหญ่มีค่าความยากอยู่ในระดับปานกลาง โดยมี

ค่าเฉลี่ยความยากทั้งฉบับเท่ากับ .51 โดยข้อสอบที่มีความยากมากที่สุดมีค่าเท่ากับ .31 ได้แก่ ข้อสอบข้อที่ 7 และข้อที่ 25 ข้อสอบที่มีความยากต่ำสุด มีค่าเท่ากับ .68 ได้แก่ ข้อสอบข้อที่ 2 และข้อที่ 36

เมื่อแบ่งช่วงของค่าความยากตามเกณฑ์การแปลความหมาย (เสรี ชัดเข้ม, 2544, หน้า 121) พบว่าข้อสอบที่มีค่าความยากอยู่ในช่วง .20-.40 มีจำนวน 7 ข้อ คิดเป็นร้อยละ 17.50 ได้แก่ ข้อสอบข้อที่ 7, 22, 24, 25, 29, 30 และ 40 ซึ่งเป็นข้อสอบที่อยู่ในระดับค่อนข้างยาก

ข้อสอบที่มีค่าความยากอยู่ในช่วง .41-.60 มีจำนวน 24 ข้อ คิดเป็นร้อยละ 60 ได้แก่ ข้อสอบข้อที่ 1, 4, 5, 6, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 27, 28, 32, 34, 35, 38 และ 39 ซึ่งเป็นข้อสอบที่มีค่าความยากอยู่ในระดับปานกลาง

ข้อสอบที่มีค่าความยากอยู่ในช่วง .61-.80 มีจำนวน 9 ข้อ คิดเป็นร้อยละ 22.50 ได้แก่ ข้อสอบข้อที่ 2, 3, 8, 9, 26, 31, 33, 36 และ 37 ซึ่งเป็นข้อสอบที่มีค่าความยากอยู่ในระดับค่อนข้างง่าย

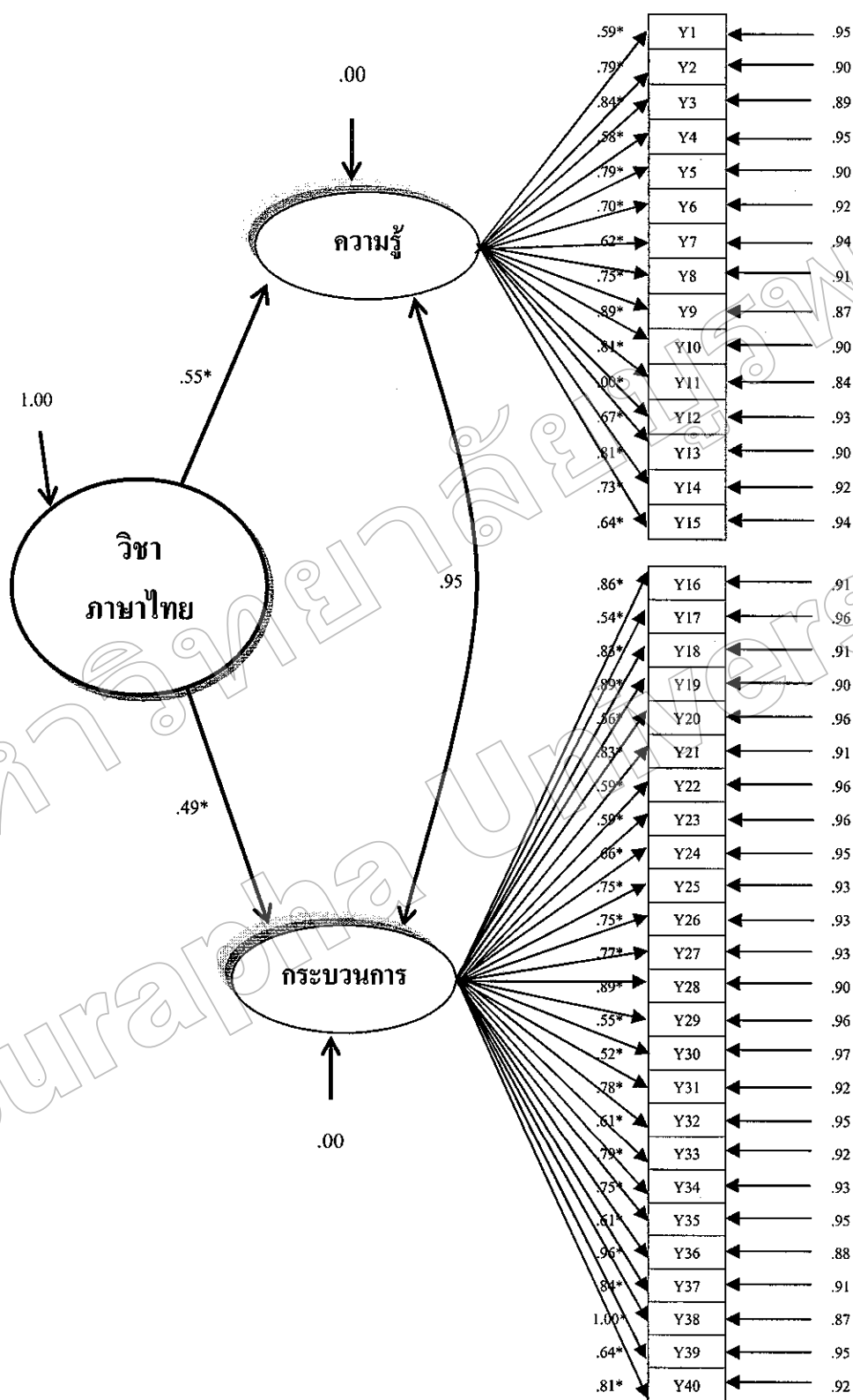
เมื่อพิจารณาค่าอำนาจจำแนก (r) จะเห็นว่าข้อสอบส่วนใหญ่มีค่าอำนาจจำแนกปานกลางค่อนข้างต่ำ โดยมีค่าเฉลี่ยของอำนาจจำแนกทั้งฉบับเท่ากับ .38 โดยข้อสอบที่มีอำนาจจำแนกสูงที่สุดมีค่าเท่ากับ .52 ได้แก่ ข้อสอบข้อที่ 11 ข้อสอบที่มีอำนาจจำแนกต่ำสุด มีค่าเท่ากับ .26 ได้แก่ ข้อสอบข้อที่ 30

เมื่อแบ่งช่วงของค่าอำนาจจำแนกของข้อสอบตามเกณฑ์การแปลความหมาย (เสรี ชัดเข้ม, 2544, หน้า 121) ปรากฏว่า ข้อสอบที่มีค่าอำนาจจำแนกอยู่ในช่วง .40 ขึ้นไป มีจำนวน 17 ข้อ คิดเป็นร้อยละ 42.50 ได้แก่ ข้อสอบข้อที่ 2, 3, 5, 6, 8, 9, 10, 11, 13, 16, 18, 19, 27, 28, 36, 37 และ 38 ซึ่งเป็นข้อสอบที่มีค่าอำนาจจำแนกอยู่ในเกณฑ์ดีมาก

ข้อสอบที่มีค่าอำนาจจำแนกอยู่ในช่วง .30-.39 มีจำนวน 18 ข้อ คิดเป็นร้อยละ 45 ได้แก่ ข้อสอบข้อที่ 1, 4, 7, 12, 14, 15, 17, 20, 21, 24, 25, 26, 31, 32, 33, 34, 39 และ 40 ซึ่งเป็นข้อสอบที่มีค่าอำนาจจำแนกอยู่ในเกณฑ์ดี

ข้อสอบที่มีค่าอำนาจจำแนกอยู่ในช่วง .20-.29 มีจำนวน 5 ข้อ คิดเป็นร้อยละ 12.50 ได้แก่ ข้อสอบข้อที่ 22, 23, 29, 30 และ 35 ซึ่งเป็นข้อสอบที่มีค่าอำนาจจำแนกอยู่ในเกณฑ์พอใช้ได้

2.2 ผลการวิเคราะห์ความตรงเชิงโครงสร้างของแบบทดสอบวิชาภาษาไทย โดยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ด้วยโปรแกรม LISREL 8.50 มีรายละเอียด ดังนี้



ภาพที่ 4 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของแบบทดสอบวิชาภาษาไทย
(N = 2000)

ตารางที่ 4 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบวิชาภาษาไทย
($N = 2000$) (ขั้นที่ 1)

ข้อ ที่	องค์ประกอบ				R ²	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
1	.59*	.05	-	-	.10	.03	-
2	.79*	.05	-	-	.18	.04	-
3	.84*	.05	-	-	.21	.04	-
4	.58*	.05	-	-	.10	.03	-
5	.79*	.05	-	-	.19	.04	-
6	.70*	.05	-	-	.15	.04	-
7	.62*	.05	-	-	.11	.03	-
8	.75*	.05	-	-	.17	.04	-
9	.89*	.05	-	-	.24	.05	-
10	.81*	.05	-	-	.20	.05	-
11	1.00*	-	-	-	.30	.07	-
12	.67*	.05	-	-	.13	.03	-
13	.81*	.05	-	-	.20	.05	-
14	.73*	.05	-	-	.16	.04	-
15	.64*	.05	-	-	.12	.03	-
16	-	-	.86*	.06	.18	-	.04
17	-	-	.54*	.05	.07	-	.02
18	-	-	.83*	.06	.17	-	.03
19	-	-	.89*	.06	.19	-	.04
20	-	-	.56*	.05	.07	-	.02
21	-	-	.83*	.06	.17	-	.04
22	-	-	.59*	.05	.09	-	.02
23	-	-	.59*	.06	.08	-	.03
24	-	-	.66*	.06	.11	-	.02

ตารางที่ 4 (ต่อ)

ข้อ ที่	องค์ประกอบ				R ²	สป.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
25	-	-	.75*	.06	.14	-	.03
26	-	-	.75*	.06	.13	-	.03
27	-	-	.77*	.06	.14	-	.02
28	-	-	.89*	.06	.19	-	.03
29	-	-	.55*	.05	.07	-	.02
30	-	-	.52*	.05	.06	-	.02
31	-	-	.78*	.06	.15	-	.03
32	-	-	.61*	.05	.09	-	.03
33	-	-	.79*	.06	.15	-	.04
34	-	-	.75*	.06	.14	-	.03
35	-	-	.61*	.05	.09	-	.02
36	-	-	.96*	.06	.22	-	.04
37	-	-	.84*	.06	.17	-	.03
38	-	-	1.00*	-	.24	-	.05
39	-	-	.64*	.06	.10	-	.02
40	-	-	.81*	.06	.16	-	.04
$\chi^2 = 580.88$ $df = 662$ $P = .99$ $GFI = .99$ $AGFI = .98$							
$CFI = 1.00$ $NFI = .95$ $RMSEA = .00$							

* $p < .05$

จากภาพที่ 4 และตารางที่ 4 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของข้อสอบวัดผลสัมฤทธิ์ทางการเรียนฉบับรวม ($N = 2000$) พบว่า โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีมาก โดยพิจารณาจากค่าสถิติไค-สแควร์เท่ากับ 580.88 มีค่าความน่าจะเป็นเท่ากับ .99 ที่องศาอิสระเท่ากับ 662 นั่นคือค่าไค-สแควร์ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ แสดงว่ายอมรับสมมติฐานหลักที่ว่าโมเดลการวิจัยสอดคล้องกับข้อมูลเชิงประจักษ์

ส่วนค่าดัชนีวัดระดับความกลมกลืน (GFI) เท่ากับ .99 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) เท่ากับ .98 ค่าดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) เท่ากับ 1.00 ค่าดัชนีวัดความสอดคล้องปกติ (NFI) เท่ากับ .95 และค่ารากของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนโดยประมาณ (RMSEA) เท่ากับ .00 ซึ่งมีค่าน้อยมากจนมีค่าเท่ากับศูนย์ แสดงว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี

เมื่อพิจารณาน้ำหนักองค์ประกอบ ในแต่ละองค์ประกอบ พบว่า

องค์ประกอบด้านความรู้ มีข้อสอบ จำนวน 15 ข้อ (ข้อที่ 1 - 15) มีค่าน้ำหนักองค์ประกอบระหว่าง .58 ถึง .89 (ข้อที่ 11 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

องค์ประกอบด้านกระบวนการ มีข้อสอบ จำนวน 25 ข้อ (ข้อที่ 16 - 40) มีค่าน้ำหนักองค์ประกอบระหว่าง .52 ถึง .96 (ข้อที่ 38 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

ตารางที่ 5 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบวิชาภาษาไทย (N = 2000) (ขั้นที่ 2)

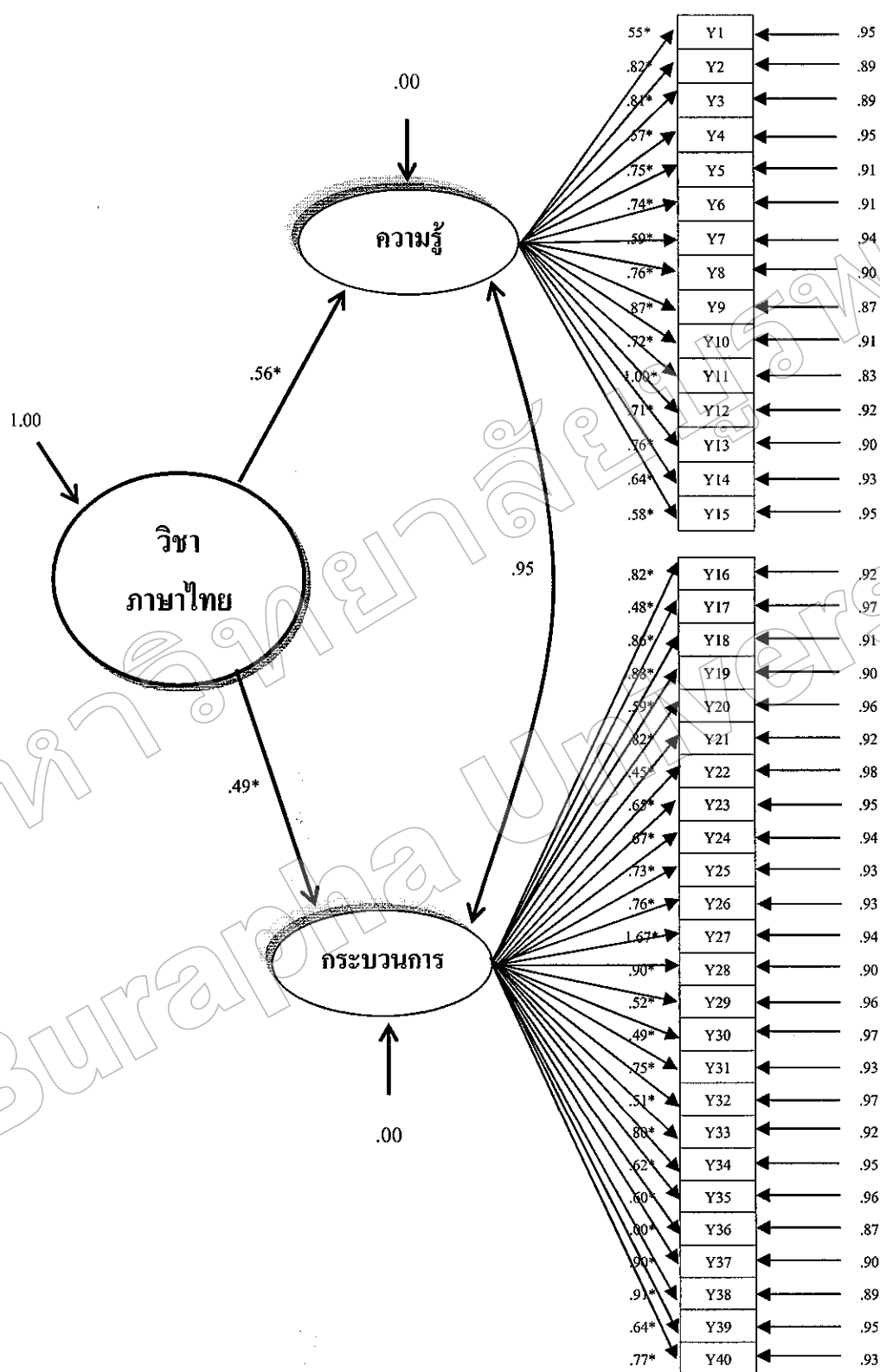
องค์ประกอบ	เมตริกซ์น้ำหนักองค์ประกอบ		R^2
	น้ำหนักองค์ประกอบ	SE	
ฉบับรวม (n=2000)			
ด้านความรู้	.55	.02	1.00
ด้านกระบวนการ	.49	.02	1.00
$\chi^2 = 580.88$ $df = 662$ $P = .99$ GFI = .99 AGFI = .98			
CFI = 1.00 NFI = .95 RMSEA = .00			

จากตารางที่ 5 พบว่า ค่าน้ำหนักองค์ประกอบด้านความรู้มีค่าเท่ากับ .55 และค่าน้ำหนักองค์ประกอบด้านกระบวนการมีค่าเท่ากับ .49 และทั้งสององค์ประกอบมีนัยสำคัญทางสถิติ โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีมาก แสดงว่าโครงสร้างของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ประกอบด้วยสององค์ประกอบ ตรงตามโครงสร้างของแบบทดสอบ

2.3 ผลการตรวจสอบคุณภาพของข้อสอบวิชาภาษาไทยในด้านความคงที่ของ
โครงสร้างองค์ประกอบ เมื่อจำแนกกลุ่มผู้สอบตามตัวแปรเพศ

2.3.1 ผลการตรวจสอบความคงที่โครงสร้างองค์ประกอบของแบบทดสอบวิชา
ภาษาไทยสำหรับเพศชาย ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ดังภาพที่ 5 และ
ตารางที่ 6 - 7

มหาวิทยาลัยบูรพา
Burapha University



ภาพที่ 5 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบวิชาภาษาไทย
สำหรับเพศชาย

ตารางที่ 6 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบวิชาภาษาไทย
สำหรับเพศชาย (ขั้นที่ 1)

ข้อ ที่	องค์ประกอบ				R ²	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
1	.55*	.07	-	-	.10	.03	-
2	.82*	.07	-	-	.22	.05	-
3	.81*	.07	-	-	.21	.05	-
4	.57*	.07	-	-	.10	.03	-
5	.75*	.07	-	-	.18	.05	-
6	.74*	.07	-	-	.17	.04	-
7	.59*	.07	-	-	.11	.03	-
8	.76*	.07	-	-	.18	.05	-
9	.87*	.07	-	-	.24	.06	-
10	.72*	.07	-	-	.16	.04	-
11	1.00*	-	-	-	.32	.08	-
12	.71*	.07	-	-	.16	.05	-
13	.76*	.07	-	-	.18	.05	-
14	.64*	.07	-	-	.13	.04	-
15	.58*	.07	-	-	.11	.03	-
16	-	-	.82*	.08	.16	-	.04
17	-	-	.48*	.08	.06	-	.02
18	-	-	.86*	.08	.18	-	.03
19	-	-	.88*	.08	.18	-	.04
20	-	-	.59*	.08	.08	-	.02
21	-	-	.82*	.08	.16	-	.03
22	-	-	.45*	.07	.05	-	.02
23	-	-	.65*	.08	.10	-	.03
24	-	-	.67*	.08	.11	-	.02

ตารางที่ 6 (ต่อ)

ข้อ ที่	องค์ประกอบ				R^2	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
25	-	-	.73*	.08	.13	-	.03
26	-	-	.76*	.08	.14	-	.03
27	-	-	.67*	.08	.11	-	.01
28	-	-	.90*	.09	.19	-	.03
29	-	-	.52*	.08	.07	-	.01
30	-	-	.49*	.08	.06	-	.02
31	-	-	.75*	.08	.14	-	.03
32	-	-	.51*	.08	.06	-	.02
33	-	-	.80*	.08	.15	-	.04
34	-	-	.62*	.08	.09	-	.02
35	-	-	.60*	.08	.09	-	.02
36	-	-	1.00*	-	.24	-	.04
37	-	-	.90*	.09	.19	-	.04
38	-	-	.91*	.08	.20	-	.04
39	-	-	.64*	.08	.10	-	.03
40	-	-	.77*	.08	.14	-	.04
$\chi^2 = 553.81$ $df = 673$ $P = 1.00$ GFI = .97 AGFI = .97							
CFI = 1.00 NFI = .90 RMSEA = .00							

* $p < .05$

จากภาพที่ 5 และตารางที่ 6 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของข้อสอบวัดผลสัมฤทธิ์ทางการเรียน สำหรับเพศชาย ($n=1000$) พบว่าโมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีมาก โดยพิจารณาจากค่าสถิติไค-สแควร์ เท่ากับ 553.81 ซึ่งมีค่าความน่าจะเป็นเท่ากับ 1.00 ที่องศาอิสระเท่ากับ 673 นั่นคือค่าไค-สแควร์ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ แสดงว่ายอมรับสมมติฐานหลักที่ว่าโมเดลการวิจัยสอดคล้องกับ

ข้อมูลเชิงประจักษ์ ส่วนค่าดัชนีวัดระดับความกลมกลืน (GFI) เท่ากับ .97 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) เท่ากับ .97 ค่าดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) เท่ากับ 1.00 ค่าดัชนีวัดความสอดคล้องปกติ (NFI) เท่ากับ .90 และค่ารากของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนโดยประมาณ (RMSEA) เท่ากับ .00 ซึ่งมีค่าน้อยมากจนมีค่าเท่ากับศูนย์ แสดงว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี

เมื่อพิจารณาค่าน้ำหนักองค์ประกอบ ในแต่ละองค์ประกอบ พบว่า

องค์ประกอบด้านความรู้ มีข้อสอบ จำนวน 15 ข้อ (ข้อที่ 1 - 15) มีค่าน้ำหนักองค์ประกอบระหว่าง .55 ถึง .87 (ข้อที่ 11 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

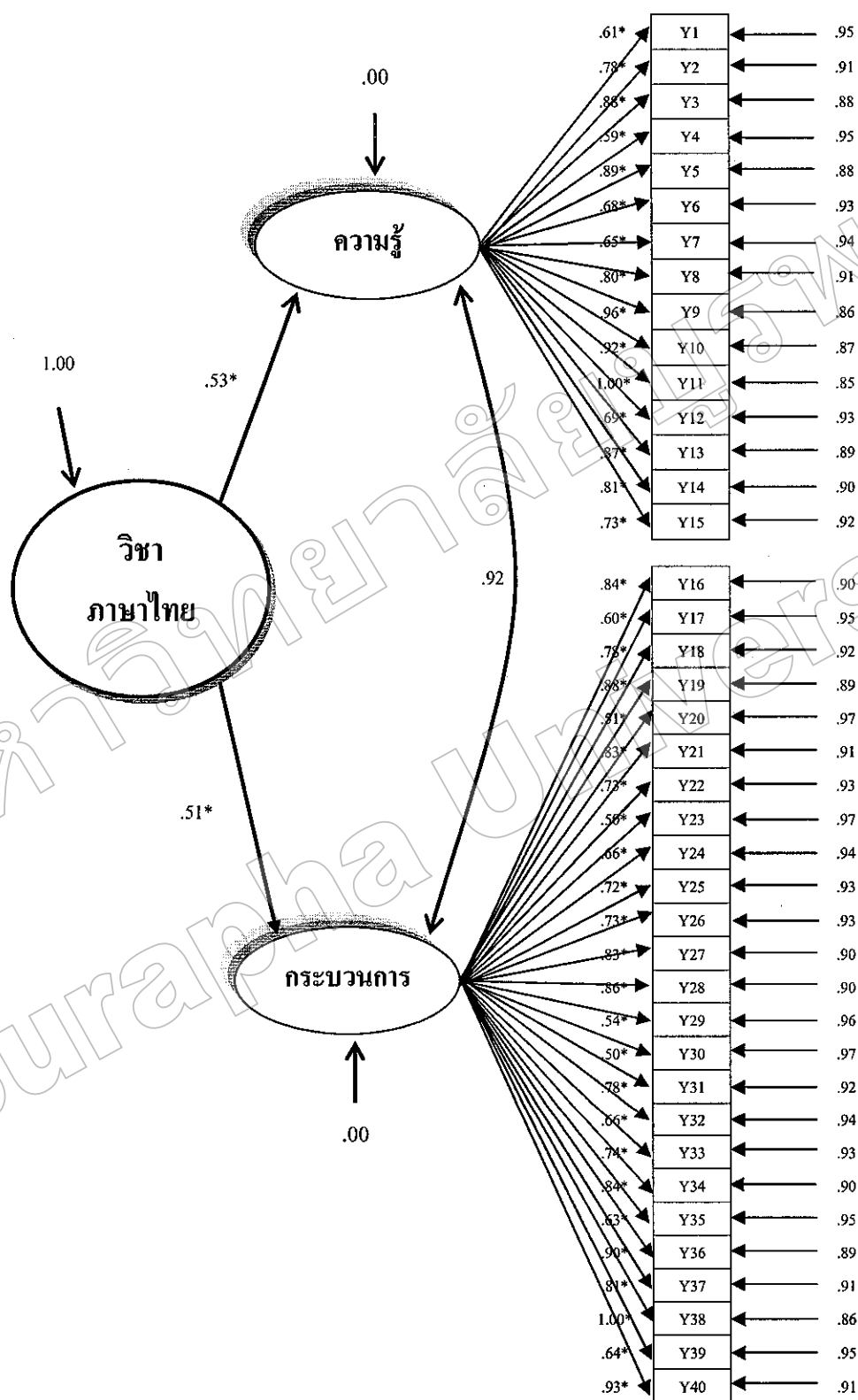
องค์ประกอบด้านกระบวนการ มีข้อสอบ จำนวน 25 ข้อ (ข้อที่ 16 - 40) มีค่าน้ำหนักองค์ประกอบระหว่าง .45 ถึง .91 (ข้อที่ 36 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

ตารางที่ 7 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบวิชาภาษาไทย สำหรับเพศชาย (ขั้นที่ 2)

องค์ประกอบ	เมทริกซ์น้ำหนักองค์ประกอบ		R^2
	น้ำหนักองค์ประกอบ	SE	
เพศชาย (n=1000)			
ด้านความรู้	.56	.03	1.00
ด้านกระบวนการ	.49	.03	1.00
$\chi^2 = 553.81$ $df = 673$ $P = 1.00$ GFI = .97 AGFI = .97			
CFI = 1.00 NFI = .90 RMSEA = .00			

จากตารางที่ 7 พบว่า ค่าน้ำหนักองค์ประกอบด้านความรู้มีค่าเท่ากับ .56 และค่าน้ำหนักองค์ประกอบด้านกระบวนการมีค่าเท่ากับ .49 และทั้งสององค์ประกอบมีนัยสำคัญทางสถิติ โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีแสดงว่าโครงสร้างของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย สำหรับเพศชาย ประกอบด้วยสององค์ประกอบ ตรงตามโครงสร้างของแบบทดสอบ

2.3.2 ผลการตรวจสอบความคงที่โครงสร้างองค์ประกอบของแบบทดสอบวิชาภาษาไทยสำหรับเพศหญิง ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สอง ดังภาพที่ 6



ภาพที่ 6 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบวิชาภาษาไทย
สำหรับเพศหญิง

ตารางที่ 8 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบวิชาภาษาไทย
สำหรับเพศหญิง (ขั้นที่ 1)

ข้อ ที่	องค์ประกอบ				R ²	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
1	.61*	.07	-	-	.11	.02	-
2	.78*	.07	-	-	.17	.04	-
3	.88*	.08	-	-	.22	.05	-
4	.59*	.07	-	-	.10	.03	-
5	.89*	.08	-	-	.23	.05	-
6	.68*	.07	-	-	.13	.03	-
7	.65*	.07	-	-	.12	.03	-
8	.80*	.07	-	-	.18	.04	-
9	.96*	.08	-	-	.26	.06	-
10	.92*	.08	-	-	.24	.05	-
11	1.00*	-	-	-	.28	.06	-
12	.69*	.07	-	-	.14	.04	-
13	.87*	.08	-	-	.21	.04	-
14	.81*	.07	-	-	.19	.04	-
15	.73*	.07	-	-	.15	.04	-
16	-	-	.84*	.08	.19	-	.04
17	-	-	.60*	.07	.09	-	.03
18	-	-	.78*	.08	.16	-	.03
19	-	-	.88*	.08	.20	-	.04
20	-	-	.51*	.07	.07	-	.01
21	-	-	.83*	.08	.18	-	.04
22	-	-	.73*	.08	.14	-	.03
23	-	-	.50*	.07	.06	-	.02
24	-	-	.66*	.07	.11	-	.03

ตารางที่ 8 (ต่อ)

ข้อ ที่	องค์ประกอบ				R ²	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
25	-	-	.72*	.08	.14	-	.03
26	-	-	.73*	.08	.14	-	.03
27	-	-	.83*	.08	.18	-	.03
28	-	-	.86*	.08	.19	-	.03
29	-	-	.54*	.07	.08	-	.01
30	-	-	.50*	.07	.07	-	.02
31	-	-	.78*	.08	.16	-	.03
32	-	-	.66*	.07	.11	-	.03
33	-	-	.74*	.08	.14	-	.03
34	-	-	.84*	.08	.18	-	.03
35	-	-	.63*	.07	.10	-	.03
36	-	-	.90*	.08	.21	-	.04
37	-	-	.81*	.08	.17	-	.04
38	-	-	1.00*	-	.26	-	.05
39	-	-	.64*	.07	.11	-	.03
40	-	-	.93*	.78	.16	-	.04
$\chi^2 = 497.11$ $df = 657$ $P = 1.00$ GFI = .98 AGFI = .97							
CFI = 1.00 NFI = .92 RMSEA = .00							

* $p < .05$

จากภาพที่ 6 และตารางที่ 8 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของข้อสอบวัดผลสัมฤทธิ์ทางการเรียน สำหรับเพศหญิง พบว่าโมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีมาก โดยพิจารณาจากค่าสถิติไค-สแควร์ เท่ากับ 497.11 ซึ่งมีค่าความน่าจะเป็นเท่ากับ 1.00 ที่องศาอิสระเท่ากับ 657 นั่นคือค่าไค-สแควร์ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ แสดงว่ายอมรับสมมติฐานหลักที่ว่าโมเดลการวิจัยสอดคล้องกับข้อมูล

เชิงประจักษ์ ส่วนค่าดัชนีวัดระดับความกลมกลืน (GFI) เท่ากับ .98 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) เท่ากับ .97 ค่าดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) เท่ากับ 1.00 ค่าดัชนีวัดความสอดคล้องปกติ (NFI) เท่ากับ .92 และค่ารากของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนโดยประมาณ (RMSEA) เท่ากับ .00 ซึ่งมีค่าน้อยมากจนมีค่าเท่ากับศูนย์ แสดงว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี

เมื่อพิจารณาน้ำหนักองค์ประกอบ ในแต่ละองค์ประกอบ พบว่า

องค์ประกอบด้านความรู้ มีข้อสอบ จำนวน 15 ข้อ (ข้อที่ 1 - 15) มีค่าน้ำหนักองค์ประกอบระหว่าง .59 ถึง .96 (ข้อที่ 11 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

องค์ประกอบด้านกระบวนการ มีข้อสอบ จำนวน 25 ข้อ (ข้อที่ 16 - 40) มีค่าน้ำหนักองค์ประกอบระหว่าง .50 ถึง .93 (ข้อที่ 38 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

ตารางที่ 9 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบวิชาภาษาไทย สำหรับเพศหญิง (ชั้นที่ 2)

องค์ประกอบ	เมทริกซ์น้ำหนักองค์ประกอบ		R ²
	น้ำหนักองค์ประกอบ	SE	
เพศหญิง (n=1000)			
ด้านความรู้	.53	.03	1.00
ด้านกระบวนการ	.51	.03	1.00
$\chi^2 = 497.11$ $df = 657$ $P = 1.00$ GFI = .98 AGFI = .97			
CFI = 1.00 NFI = .92 RMSEA = .00			

จากตารางที่ 9 พบว่า ค่าน้ำหนักองค์ประกอบด้านความรู้มีค่าเท่ากับ .53 และค่าน้ำหนักองค์ประกอบด้านกระบวนการมีค่าเท่ากับ .51 และทั้งสององค์ประกอบมีนัยสำคัญทางสถิติ โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี แสดงว่าโครงสร้างของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย สำหรับเพศหญิง ประกอบด้วยสององค์ประกอบ ตรงตามโครงสร้างของแบบทดสอบ

ตอนที่ 3 ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธีถดถอยโลจิสติกและวิธีซิปเทสท์ปรับปรุง

3.1 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก

ในแบบทดสอบวิชาภาษาไทย จำนวน 40 ข้อ รายละเอียดดังตารางที่ 10

ตารางที่ 10 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยวิธีถดถอยโลจิสติก

ข้อ ที่	กลุ่มผู้สอบ (g)		ปฏิสัมพันธ์ ระหว่างกลุ่ม (g by x)		ข้อ ที่	กลุ่มผู้สอบ (g)		ปฏิสัมพันธ์ ระหว่างกลุ่ม (g by x)	
	B	Sig	B	Sig		B	Sig	B	Sig
1	-.24	.02*	-.01	.59	18	-.06	.54	.00	.10
2	-.14	.19	-.01	.64	19	-.09	.36	-.00	.85
3	-.16	.13	-.00	.79	20	.06	.51	.01	.58
4	.11	.28	.01	.46	21	.26	.01*	.03	.06
5	-.03	.78	.01	.75	22	.18	.08	-.07	.00*
6	.16	.11	-.01	.59	23	.25	.01*	.01	.43
7	.03	.80	-.05	.00*	24	.44	.00*	-.01	.35
8	.05	.62	.04	.01*	25	.32	.00*	.00	.97
9	-.17	.11	-.01	.71	26	-.08	.42	.02	.25
10	.21	.04*	-.01	.51	27	-.19	.05*	.02	.20
11	-.03	.81	-.02	.24	28	-.18	.08	.02	.17
12	-.20	.04*	-.00	.84	29	-.01	.94	-.01	.29
13	-.16	.12	-.02	.27	30	.09	.36	-.00	.95
14	.20	.04*	-.02	.11	31	-.11	.27	.01	.43
15	-.03	.77	-.00	.88	32	-.08	.43	-.01	.72
16	.22	.03*	-.00	.85	33	-.28	.01*	.01	.67
17	.30	.00*	-.03	.03*	34	-.08	.41	-.03	.02*

ตารางที่ 10 (ต่อ)

ข้อ ที่	กลุ่มผู้สอบ (g)		ปฏิสัมพันธ์ ระหว่างกลุ่ม (g by x)		ข้อ ที่	กลุ่มผู้สอบ (g)		ปฏิสัมพันธ์ ระหว่างกลุ่ม (g by x)	
	B	Sig	B	Sig		B	Sig	B	Sig
35	-.346	.00*	.01	.55	38	-.41	.00*	-.02	.18
36	-.359	.00*	-.02	.29	39	-.25	.01*	.01	.27
37	-.046	.66	.01	.43	40	.17	.09	.00	.78
วิธีถดถอยโลจิสติกพบ DIF จำนวน 20 ข้อ									

* $p < .05$

จากตารางที่ 10 เมื่อพิจารณาค่านัยสำคัญของกลุ่มผู้สอบ (g) และปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบกับความสามารถหรือคะแนนรวม (g by x) แล้วพบว่าข้อสอบข้อใดมีนัยสำคัญที่ระดับ .05 หมายความว่าข้อสอบข้อนั้นทำหน้าที่ต่างกัน ซึ่งจากตารางพบว่ามีข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันทั้งหมด 20 ข้อ โดยแยกเป็นข้อสอบแบบอนเกรูป จำนวน 5 ข้อ ได้แก่ ข้อที่ 7, 8, 17, 22 และ 34 เป็นข้อสอบแบบเอกรูป จำนวน 15 ข้อ ได้แก่ ข้อที่ 1, 10, 12, 14, 16, 21, 23, 24, 25, 27, 33, 35, 36, 38 และ 40

3.2 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีชิปเทสที่ปรับใหม่ ในแบบทดสอบวิชาภาษาไทย จำนวน 40 ข้อ แบ่งการวิเคราะห์ออกเป็นสองกลุ่มโดยใช้คะแนนเฉลี่ย คือ 20 คะแนน เป็นเกณฑ์ในการแบ่งกลุ่ม โดยผู้ที่ได้คะแนนมากกว่าหรือเท่ากับ 20 คะแนนเป็นกลุ่มสูง และผู้ที่ได้คะแนนน้อยกว่า 20 คะแนนเป็นกลุ่มต่ำ แล้วแยกวิเคราะห์เป็นสองกลุ่ม รายละเอียดดังตารางที่ 11 - 12

ตารางที่ 11 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยวิธีชิปทดสอบที่ปรับใหม่ (กลุ่มต่ำ)

ข้อ ที่	วิธีชิปทดสอบที่ปรับใหม่		ข้อ ที่	วิธีชิปทดสอบที่ปรับใหม่	
	β_{uni}	<i>p-Value</i>		β_{uni}	<i>p-Value</i>
1	.11	.00*	21	-.05	.10
2	.10	.00*	22	-.09	.00*
3	.04	.19	23	-.05	.10
4	.03	.43	24	-.04	.18
5	-.01	.81	25	-.00	.87
6	.03	.32	26	.06	.07
7	.05	.06	27	.06	.07
8	.02	.62	28	.08	.01*
9	.06	.07	29	.04	.19
10	-.06	.06	30	-.00	.90
11	.04	.24	31	.02	.46
12	.02	.60	32	-.02	.47
13	.06	.04*	33	.08	.02*
14	-.04	.22	34	.03	.37
15	.03	.31	35	.09	.01*
16	-.01	.85	36	.15	.00*
17	-.02	.57	37	.07	.04*
18	.03	.30	38	.05	.12
19	.06	.07	39	.10	.00*
20	-.02	.56	40	-.05	.08

* $p < .05$

จากตารางที่ 11 เมื่อพิจารณาค่า *p-Value* ถ้าน้อยกว่าหรือเท่ากับ .05 แสดงว่าข้อสอบข้อนั้นมีการทำหน้าที่ต่างกัน จากตาราง (กลุ่มต่ำ) พบว่ามีข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันทั้งหมดจำนวน 10 ข้อ ได้แก่ ข้อสอบข้อที่ 1, 2, 13, 22, 28, 33, 35, 36, 37 และ 39

ตารางที่ 12 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยวิธีซิปเทสท์ปรับปรุงใหม่ (กลุ่มสูง)

ข้อ ที่	วิธีซิปเทสท์ปรับปรุงใหม่		ข้อ ที่	วิธีซิปเทสท์ปรับปรุงใหม่	
	β_{uni}	<i>p-Value</i>		β_{uni}	<i>p-Value</i>
1	.04	.17	21	-.06	.04*
2	.02	.40	22	.04	.17
3	.00	.91	23	-.03	.38
4	.01	.80	24	-.05	.13
5	.00	1.00	25	.00	1.00
6	-.06	.03*	26	.02	.53
7	.08	.01*	27	.12	.00*
8	-.02	.34	28	.07	.02*
9	.07	.01*	29	.07	.03*
10	.06	.02*	30	-.01	.78
11	.01	.78	31	.02	.39
12	-.00	.92	32	.06	.03*
13	.10	.00*	33	.04	.10
14	-.10	.65	34	.14	.00*
15	.04	.13	35	.09	.00*
16	.01	.78	36	.07	.00*
17	.04	.17	37	-.00	.85
18	.01	.66	38	.11	.00*
19	.04	.10	39	.07	.02*
20	-.01	.83	40	-.02	.54

* $p < .05$

จากตารางที่ 12 เมื่อพิจารณาค่า *p-Value* ถ้าน้อยกว่าหรือเท่ากับ .05 แสดงว่าข้อสอบข้อนั้นมีการทำหน้าที่ต่างกัน ซึ่งจากตาราง (กลุ่มสูง) พบว่ามีข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันทั้งหมดจำนวน 15 ข้อ ได้แก่ ข้อสอบข้อที่ 6, 7, 9, 10, 13, 21, 27, 28, 29, 32, 34, 35, 36, 38 และ 39

ตารางที่ 13 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยวิธีชิปทดสอบที่ปรับใหม่ทั้งกลุ่มสูง
และกลุ่มต่ำที่พบตรงกัน

ข้อ ที่	วิธีชิปทดสอบที่ปรับใหม่		ผลการ ตรวจสอบ	ข้อ ที่	วิธีชิปทดสอบที่ปรับใหม่		ผลการ ตรวจสอบ
	กลุ่มต่ำ	กลุ่มสูง			กลุ่มต่ำ	กลุ่มสูง	
1	DIF	No DIF	DIF	21	No DIF	DIF	DIF
2	DIF	No DIF	DIF	22	DIF	No DIF	DIF
3	No DIF	No DIF	No DIF	23	No DIF	No DIF	No DIF
4	No DIF	No DIF	No DIF	24	No DIF	No DIF	No DIF
5	No DIF	No DIF	No DIF	25	No DIF	No DIF	No DIF
6	No DIF	DIF	DIF	26	No DIF	No DIF	No DIF
7	No DIF	DIF	DIF	27	No DIF	DIF	DIF
8	No DIF	No DIF	No DIF	28	DIF	DIF	DIF
9	No DIF	DIF	DIF	29	No DIF	DIF	DIF
10	No DIF	DIF	DIF	30	No DIF	No DIF	No DIF
11	No DIF	No DIF	No DIF	31	No DIF	No DIF	No DIF
12	No DIF	No DIF	No DIF	32	No DIF	DIF	DIF
13	DIF	DIF	DIF	33	DIF	No DIF	DIF
14	No DIF	No DIF	No DIF	34	No DIF	DIF	DIF
15	No DIF	No DIF	No DIF	35	DIF	DIF	DIF
16	No DIF	No DIF	No DIF	36	DIF	DIF	DIF
17	No DIF	No DIF	No DIF	37	DIF	No DIF	DIF
18	No DIF	No DIF	No DIF	38	No DIF	DIF	DIF
19	No DIF	No DIF	No DIF	39	DIF	DIF	DIF
20	No DIF	No DIF	No DIF	40	No DIF	No DIF	No DIF
วิธีชิปทดสอบที่ปรับใหม่พบ DIF				จำนวน 20 ข้อ			

จากตารางที่ 13 พบว่ามีข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันได้จากการตรวจสอบด้วยวิธีชิปทดสอบที่ปรับใหม่ ที่พบในกลุ่มสูงและกลุ่มต่ำ จำนวน 20 ข้อ ได้แก่ ข้อที่ 1, 2, 6, 7, 9, 10, 13, 21, 22, 27, 28, 29, 32, 33, 34, 35, 36, 37, 38 และ 39

3.3 ผลการตัดข้อสอบที่ตรวจพบว่าทำหน้าที่ต่างกันของข้อสอบ โดยได้จากการตรวจพบตรงกันทั้งสองวิธี รายละเอียดดังตารางที่ 14

ตารางที่ 14 ผลการตัดข้อสอบทำหน้าที่ต่างกันที่ได้จากการตรวจพบตรงกันสองวิธี (วิธีถดถอยโลจิสติก กับ วิธีชิปเทสต์ปรับปรุง) ออกจากแบบทดสอบ

ข้อ ที่	วิธีการตรวจสอบ		ผลการ ตรวจสอบ	ข้อ ที่	วิธีการตรวจสอบ		ผลการ ตรวจสอบ
	ถดถอย โลจิสติก	ชิปเทสต์ ปรับปรุง			ถดถอย โลจิสติก	ชิปเทสต์ ปรับปรุง	
1	DIF	DIF	ตัดออก	21	DIF	DIF	ตัดออก
2	No DIF	DIF	คงไว้	22	DIF	DIF	ตัดออก
3	No DIF	No DIF	คงไว้	23	DIF	No DIF	คงไว้
4	No DIF	No DIF	คงไว้	24	DIF	No DIF	คงไว้
5	No DIF	No DIF	คงไว้	25	DIF	No DIF	คงไว้
6	No DIF	DIF	คงไว้	26	No DIF	No DIF	คงไว้
7	DIF	DIF	ตัดออก	27	DIF	DIF	ตัดออก
8	DIF	No DIF	คงไว้	28	No DIF	DIF	คงไว้
9	No DIF	DIF	คงไว้	29	No DIF	DIF	คงไว้
10	DIF	DIF	ตัดออก	30	No DIF	No DIF	คงไว้
11	No DIF	No DIF	คงไว้	31	No DIF	No DIF	คงไว้
12	DIF	No DIF	คงไว้	32	No DIF	DIF	คงไว้
13	No DIF	DIF	คงไว้	33	DIF	DIF	ตัดออก
14	DIF	No DIF	คงไว้	34	DIF	DIF	ตัดออก
15	No DIF	No DIF	คงไว้	35	DIF	DIF	ตัดออก
16	DIF	No DIF	คงไว้	36	DIF	DIF	ตัดออก
17	DIF	No DIF	คงไว้	37	No DIF	DIF	คงไว้
18	No DIF	No DIF	คงไว้	38	DIF	DIF	ตัดออก
19	No DIF	DIF	คงไว้	39	DIF	DIF	ตัดออก
20	No DIF	No DIF	คงไว้	40	No DIF	No DIF	คงไว้
จำนวนข้อสอบที่ตัดออกจากแบบทดสอบ				จำนวน 12 ข้อ			

จากตารางที่ 14 พบว่า จากข้อสอบทั้งหมด 40 ข้อ มีข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกัน ที่ได้จากการตรวจพบตรงกันทั้งสองวิธี จำนวน 12 ข้อ คิดเป็นร้อยละ 30 ข้อสอบที่ต้องตัดออกจากแบบทดสอบโดยไม่นำมาคิดคะแนนให้กับผู้เข้าสอบ ได้แก่ ข้อสอบข้อที่ 1, 7, 10, 21, 22, 27, 33, 34, 35, 36, 38 และ 39 คงเหลือข้อสอบที่ไม่มี DIF จำนวน 28 ข้อ

ตอนที่ 4 ผลการตรวจสอบคุณภาพของแบบทดสอบ ฉบับหลังจากตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ

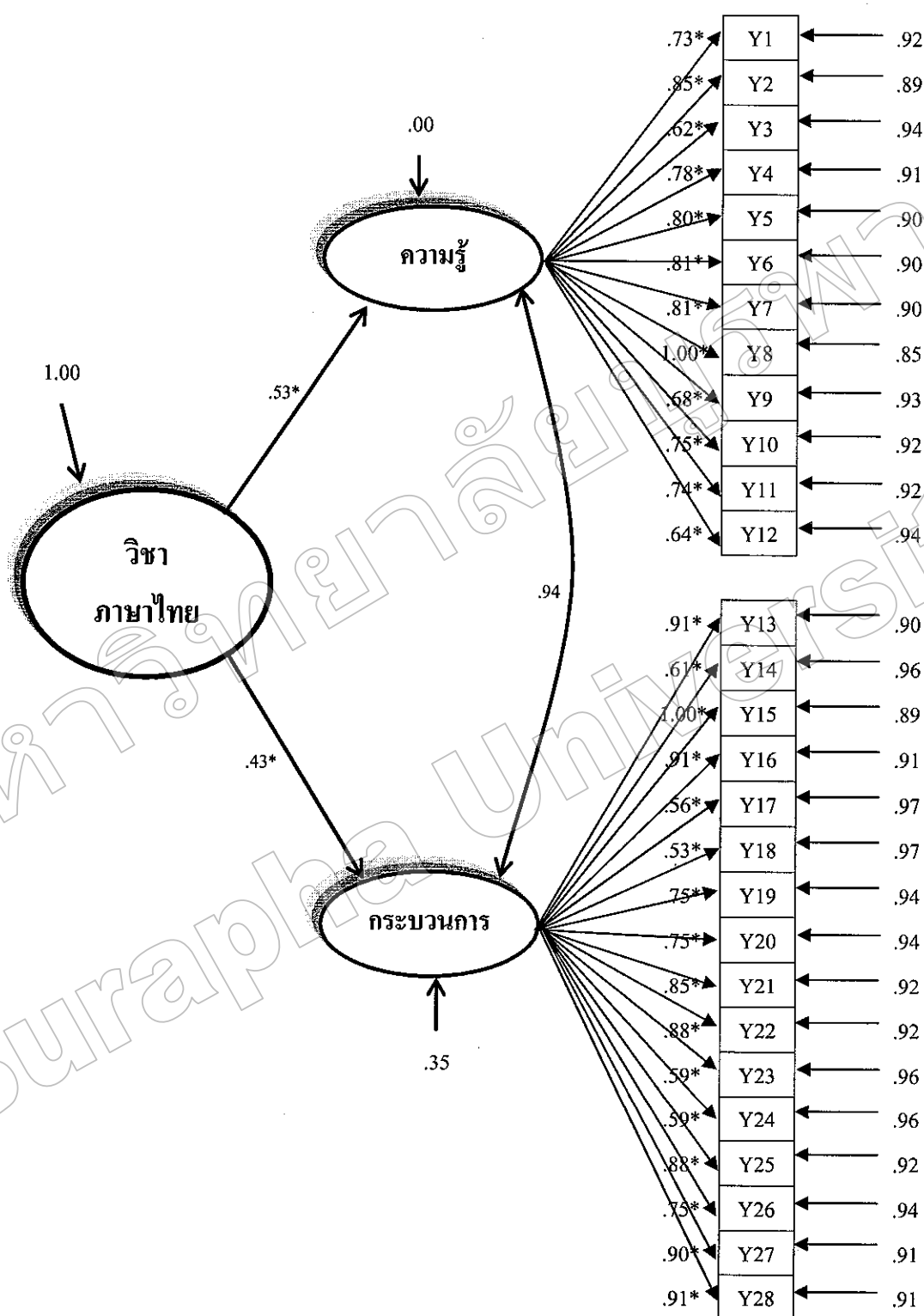
4.1 ผลการตรวจสอบคุณภาพของแบบทดสอบวิชาภาษาไทย ด้านค่าความเที่ยง ฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ แล้วปรับขยายค่าความเที่ยงให้มีสัดส่วนจำนวนข้อเท่ากับฉบับก่อนตัด โดยใช้สูตรสเปียร์แมนบราวน์ แสดงดังตารางที่ 15

ตารางที่ 15 ค่าความเที่ยงแบบสัมประสิทธิ์แอลฟาของแบบทดสอบวิชาภาษาไทยฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ

แบบทดสอบ	ค่าความเที่ยง
ฉบับหลังตัดข้อสอบที่พบ DIF ออกจากแบบทดสอบ (28 ข้อ)	.82
ปรับขยายให้มีจำนวนข้อเท่ากับ 40 ข้อ โดยใช้สูตรสเปียร์แมนบราวน์	.87

จากตารางที่ 15 พบว่าคุณภาพด้านค่าความเที่ยงของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ จำนวน 12 ข้อ เหลือข้อสอบในแบบทดสอบ จำนวน 28 ข้อ วิเคราะห์ค่าความเที่ยงแบบสัมประสิทธิ์แอลฟาของครอนบาค ได้ค่าความเที่ยงเท่ากับ .82 แล้วทำการปรับขยายเพื่อให้มีสัดส่วนจำนวนข้อสอบเท่ากับ 40 ข้อ โดยใช้สูตรสเปียร์แมนบราวน์ ได้ค่าความเที่ยงหลังปรับขยายเท่ากับ .87 ซึ่งค่าความเที่ยงที่ได้หลังจากปรับขยายแล้วสามารถนำไปเปรียบเทียบกับค่าความเที่ยงของแบบทดสอบฉบับก่อนที่จะมีการตรวจสอบ DIF ได้ เพราะมีสัดส่วนจำนวนข้อเท่ากัน

4.2 ผลการวิเคราะห์ความตรงเชิงโครงสร้างของแบบทดสอบวิชาภาษาไทยฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ โดยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ด้วยโปรแกรม LISREL 8.50 มีรายละเอียดดังนี้



ภาพที่ 7 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบฉบับหลังตัดข้อสอบ ทำหน้าที่ต่างกันออกจากแบบทดสอบ (N= 2000)

ตารางที่ 16 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบฉบับหลังตัด
ข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ ($N=2000$) (ขั้นที่ 1)

ข้อ ที่	องค์ประกอบ				R ²	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
1	.73*	.05	-	-	.15	.04	-
2	.85*	.06	-	-	.21	.06	-
3	.62*	.05	-	-	.11	.04	-
4	.78*	.05	-	-	.17	.05	-
5	.80*	.05	-	-	.18	.06	-
6	.81*	.05	-	-	.19	.06	-
7	.81*	.05	-	-	.19	.06	-
8	1.00*	-	-	-	.29	.08	-
9	.68*	.05	-	-	.13	.04	-
10	.75*	.05	-	-	.16	.05	-
11	.74*	.05	-	-	.16	.05	-
12	.64*	.05	-	-	.12	.04	-
13	-	-	.97*	.07	.20	-	.05
14	-	-	.61*	.06	.08	-	.03
15	-	-	1.00*	-	.21	-	.05
16	-	-	.91*	.07	.17	-	.04
17	-	-	.56*	.06	.06	-	.02
18	-	-	.53*	.06	.06	-	.01
19	-	-	.75*	.07	.12	-	.03
20	-	-	.75*	.07	.12	-	.04
21	-	-	.85*	.07	.15	-	.04
22	-	-	.88*	.07	.16	-	.04
23	-	-	.59*	.06	.07	-	.01
24	-	-	.59*	.06	.07	-	.03

ตารางที่ 16 (ต่อ)

ข้อ ที่	องค์ประกอบ				R^2	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
25	-	-	.88*	.07	.16	-	.05
26	-	-	.75*	.06	.12	-	.03
27	-	-	.90*	.07	.17	-	.05
28	-	-	.91*	.07	.17	-	.05
$\chi^2 = 237.18$ $df = 307$ $P = 1.00$ GFI = .99 AGFI = .99							
CFI = 1.00 NFI = .96 RMSEA = .00							

* $p < .05$

จากภาพที่ 7 และตารางที่ 16 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของข้อสอบวัดผลสัมฤทธิ์ทางการเรียนฉบับรวม ($N=2000$) พบว่าโมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีมาก โดยพิจารณาจากค่าสถิติไค-สแควร์ เท่ากับ 264.68 มีค่าความน่าจะเป็นเท่ากับ .91 ที่องศาอิสระเท่ากับ 296 นั่นคือค่าไค-สแควร์ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ แสดงว่ายอมรับสมมติฐานหลักที่ว่าโมเดลการวิจัยสอดคล้องกับข้อมูลเชิงประจักษ์ ส่วนค่าดัชนีวัดระดับความกลมกลืน (GFI) เท่ากับ .99 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) เท่ากับ .99 ค่าดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) เท่ากับ 1.00 ค่าดัชนีวัดความสอดคล้องปกติ (NFI) เท่ากับ .96 และค่ารากของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนโดยประมาณ (RMSEA) เท่ากับ .00 ซึ่งมีค่าน้อยมากจนมีค่าเท่ากับศูนย์ แสดงว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี

เมื่อพิจารณาน้ำหนักองค์ประกอบ ในแต่ละองค์ประกอบ พบว่า

องค์ประกอบด้านความรู้ มีข้อสอบ จำนวน 12 ข้อ (ข้อที่ 1 - 12) มีค่าน้ำหนักองค์ประกอบระหว่าง .62 ถึง .85 (ข้อที่ 8 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

องค์ประกอบด้านกระบวนการ มีข้อสอบ จำนวน 16 ข้อ (ข้อที่ 13 - 28) มีค่าน้ำหนักองค์ประกอบระหว่าง .53 ถึง .97 (ข้อที่ 15 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

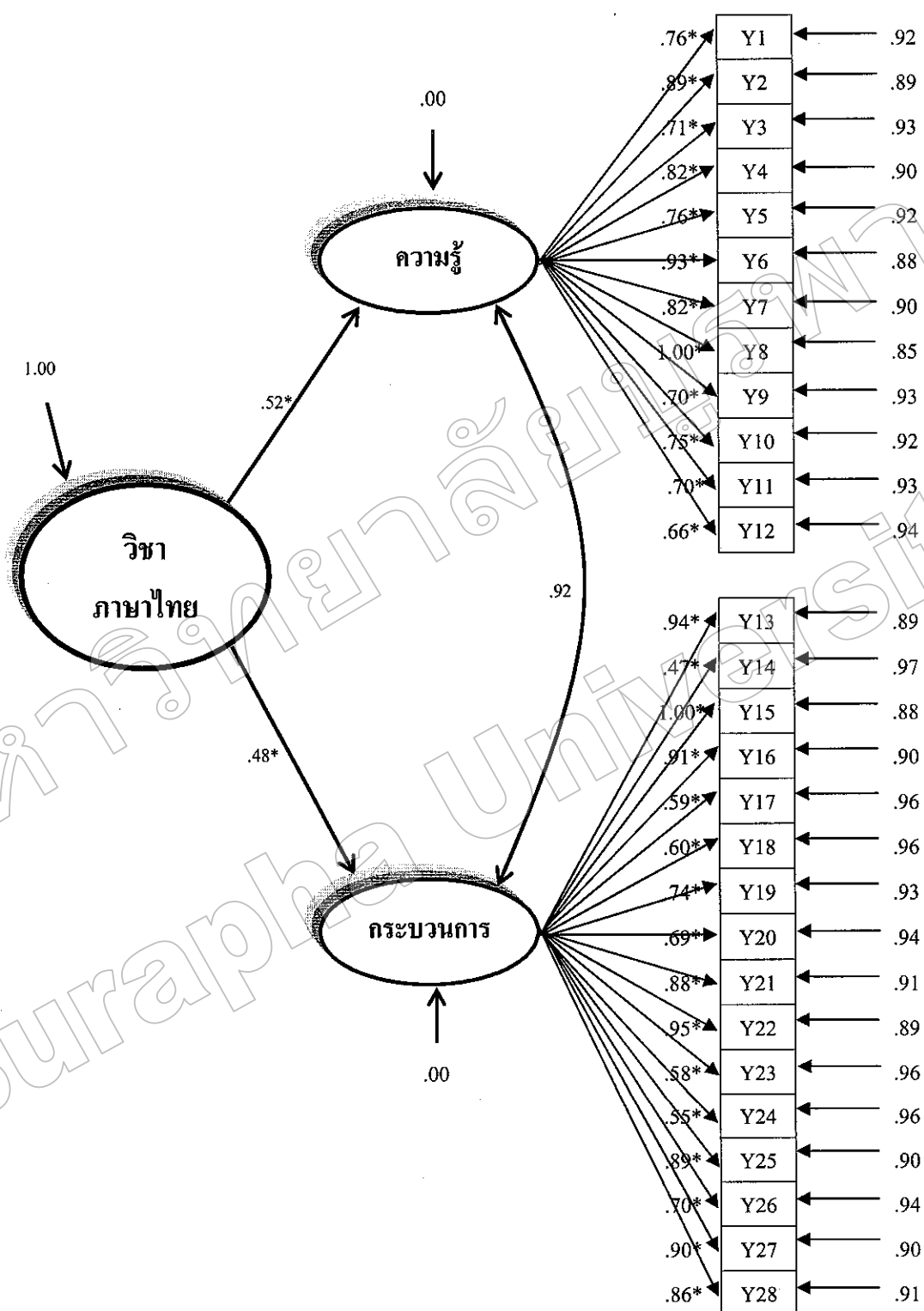
ตารางที่ 17 ผลการวิเคราะห์ห้้องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบฉบับหลังตัด
ข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ ($N=2000$) (ขั้นที่ 2)

องค์ประกอบ	เมทริกซ์น้ำหนักองค์ประกอบ		R^2
	น้ำหนักองค์ประกอบ	SE	
ฉบับรวม (N=2000)			
ด้านความรู้	.53	.02	1.00
ด้านกระบวนการ	.43	.02	.88
$\chi^2 = 237.18$	$df = 307$	$P = 1.00$	GFI = .99
CFI = 1.00	NFI = .96	RMSEA = .00	AGFI = .99

จากตารางที่ 17 พบว่า คำน้หนักองค์ประกอบด้านความรู้มีค่าเท่ากับ .53 และค่าน้ำหนักองค์ประกอบด้านกระบวนการมีค่าเท่ากับ .43 และทั้งสององค์ประกอบมีนัยสำคัญทางสถิติ โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี แสดงว่าโครงสร้างของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ประกอบด้วยสององค์ประกอบ ตรงตามโครงสร้างของแบบทดสอบ

4.3 ผลการตรวจสอบความคงที่ของโครงสร้างองค์ประกอบ ของแบบทดสอบวิชาภาษาไทยฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ โดยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ด้วยโปรแกรม LISREL 8.50 มีรายละเอียดดังนี้

4.3.1 ผลการตรวจสอบคุณภาพของโครงสร้างองค์ประกอบของข้อสอบวิชาภาษาไทยฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ สำหรับเพศชาย โดยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง แสดงดังภาพที่ 8 และตารางที่ 18-19



ภาพที่ 8 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบฉบับหลังตัดข้อสอบ ทำหน้าที่ต่างกันออกจากแบบทดสอบ สำหรับเพศชาย

ตารางที่ 18 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบฉบับหลังตัด
ข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ สำหรับเพศชาย (ขั้นที่ 1)

ข้อ ที่	องค์ประกอบ				R ²	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
1	.76*	.08	-	-	.16	.05	-
2	.89*	.08	-	-	.22	.05	-
3	.71*	.08	-	-	.14	.05	-
4	.82*	.08	-	-	.18	.06	-
5	.76*	.08	-	-	.16	.04	-
6	.93*	.08	-	-	.23	.06	-
7	.82*	.08	-	-	.18	.06	-
8	1.00*	-	-	-	.27	.08	-
9	.70*	.08	-	-	.13	.04	-
10	.75*	.08	-	-	.15	.05	-
11	.70*	.08	-	-	.13	.04	-
12	.66*	.08	-	-	.12	.04	-
13	-	-	.94*	.09	.20	-	.05
14	-	-	.47*	.08	.05	-	.02
15	-	-	1.00*	-	.23	-	.07
16	-	-	.91*	.09	.19	-	.05
17	-	-	.59*	.08	.08	-	.03
18	-	-	.60*	.08	.08	-	.03
19	-	-	.74*	.09	.13	-	.04
20	-	-	.69*	.09	.11	-	.03
21	-	-	.88*	.09	.18	-	.05
22	-	-	.95*	.09	.21	-	.05
23	-	-	.58*	.08	.08	-	.02
24	-	-	.55*	.08	.07	-	.03

ตารางที่ 18 (ต่อ)

ข้อ ที่	องค์ประกอบ				R^2	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
25	-	-	.89*	.09	.18	-	.06
26	-	-	.70*	.08	.11	-	.03
27	-	-	.90*	.09	.18	-	.05
28	-	-	.86*	.09	.17	-	.05
$\chi^2 = 240.74$ $df = 303$ $P = 1.00$ $GFI = .98$ $AGFI = .98$							
$CFI = 1.00$ $NFI = .93$ $RMSEA = .00$							

* $p < .05$

จากภาพที่ 8 และตารางที่ 18 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของข้อสอบวัดผลสัมฤทธิ์ทางการเรียนของเพศชาย ($n=1000$) พบว่าโมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีมาก โดยพิจารณาจากค่าสถิติไค-สแควร์ เท่ากับ 240.74 ซึ่งมีค่าความน่าจะเป็นเท่ากับ 1.00 ที่องศาอิสระเท่ากับ 303 นั่นคือค่าไค-สแควร์ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ แสดงว่ายอมรับสมมติฐานหลักที่ว่าโมเดลการวิจัยสอดคล้องกับข้อมูลเชิงประจักษ์ ส่วนค่าดัชนีวัดระดับความกลมกลืน (GFI) เท่ากับ .98 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) เท่ากับ .99 ค่าดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) เท่ากับ 1.00 ค่าดัชนีวัดความสอดคล้องปกติ (NFI) เท่ากับ .93 และค่ารากของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนโดยประมาณ (RMSEA) เท่ากับ .00 ซึ่งมีค่าน้อยมากจนมีค่าเท่ากับศูนย์ แสดงว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี

เมื่อพิจารณาน้ำหนักองค์ประกอบ ในแต่ละองค์ประกอบ พบว่า

องค์ประกอบด้านความรู้ มีข้อสอบ จำนวน 12 ข้อ (ข้อที่ 1 - 12) มีค่าน้ำหนักองค์ประกอบระหว่าง .66 ถึง .93 (ข้อที่ 11 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

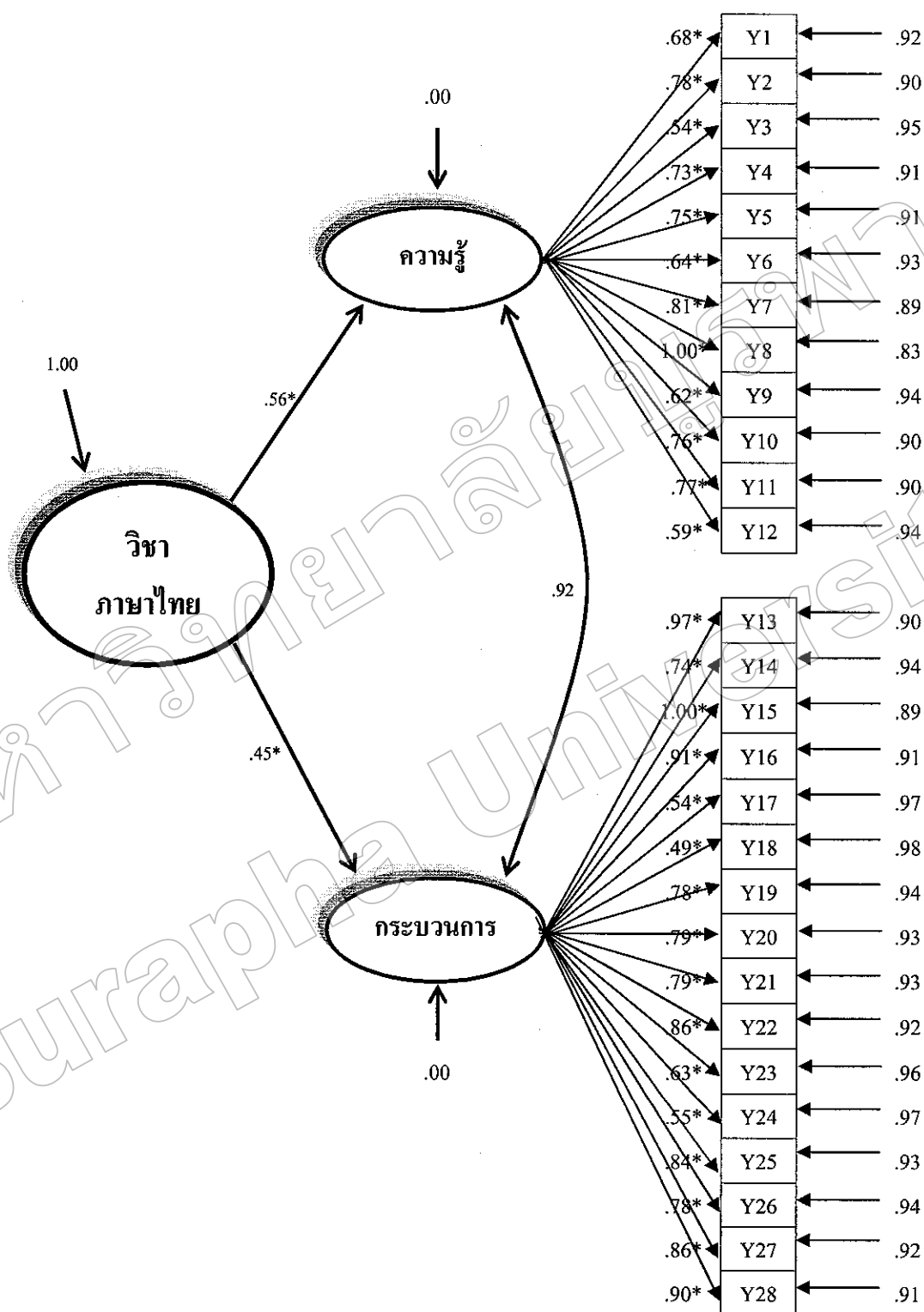
องค์ประกอบด้านกระบวนการ มีข้อสอบ จำนวน 16 ข้อ (ข้อที่ 13 - 28) มีค่าน้ำหนักองค์ประกอบระหว่าง .47 ถึง .95 (ข้อที่ 15 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

ตารางที่ 19 ผลการวิเคราะห์ห้อยค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบฉบับหลังตัด
ข้อสอบทำหน้าที่แตกต่างกันออกจากแบบทดสอบ สำหรับเพศชาย

องค์ประกอบ	เมทริกซ์น้ำหนักองค์ประกอบ		R^2
	น้ำหนักองค์ประกอบ	SE	
เพศชาย (n=1000)			
ด้านความรู้	.52	.03	1.00
ด้านกระบวนการ	.48	.03	1.00
$\chi^2 = 240.74 \quad df = 303 \quad P = 1.00 \quad GFI = .98 \quad AGFI = .98$			
CFI = 1.00 NFI = .93 RMSEA = .00			

จากตารางที่ 19 พบว่า ค่าน้ำหนักองค์ประกอบด้านความรู้มีค่าเท่ากับ .52 และ ค่าน้ำหนักองค์ประกอบด้านกระบวนการมีค่าเท่ากับ .48 และทั้งสององค์ประกอบมีนัยสำคัญทางสถิติ โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี แสดงว่าโครงสร้างของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย สำหรับเพศชาย ประกอบด้วยสององค์ประกอบ ตรงตามโครงสร้างของแบบทดสอบ

4.3.2 ผลการตรวจสอบความคงที่ของโครงสร้างองค์ประกอบ ของแบบทดสอบวิชาภาษาไทยฉบับหลังตัดข้อสอบทำหน้าที่แตกต่างกันออกจากแบบทดสอบ สำหรับเพศหญิง โดยการวิเคราะห์ห้อยค์ประกอบเชิงยืนยันอันดับสอง แสดงดังภาพที่ 9 และตารางที่ 20-21



ภาพที่ 9 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบฉบับหลังตัดข้อสอบ
ทำหน้าที่ย่างกันออกจากแบบทดสอบ สำหรับเพศหญิง

ตารางที่ 20 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบฉบับหลังตัด
ข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ สำหรับเพศหญิง (ขั้นที่ 2)

ข้อ ที่	องค์ประกอบ				R ²	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
1	.68*	.07	-	-	.15	.04	-
2	.78*	.07	-	-	.19	.06	-
3	.54*	.07	-	-	.09	.04	-
4	.73*	.07	-	-	.17	.04	-
5	.75*	.07	-	-	.18	.05	-
6	.64*	.07	-	-	.13	.04	-
7	.81*	.07	-	-	.21	.07	-
8	1.00*	-	-	-	.32	.10	-
9	.62*	.07	-	-	.12	.04	-
10	.76*	.07	-	-	.18	.06	-
11	.77*	.07	-	-	.19	.06	-
12	.59*	.07	-	-	.11	.04	-
13	-	-	.97*	.10	.19	-	.04
14	-	-	.74*	.09	.11	-	.03
15	-	-	1.00*	-	.20	-	.05
16	-	-	.91*	.09	.17	-	.04
17	-	-	.54*	.09	.06	-	.02
18	-	-	.49*	.09	.05	-	.01
19	-	-	.78*	.09	.12	-	.03
20	-	-	.79*	.09	.13	-	.05
21	-	-	.79*	.09	.13	-	.04
22	-	-	.86*	.10	.15	-	.05
23	-	-	.63*	.09	.08	-	.01
24	-	-	.55*	.09	.06	-	.02

ตารางที่ 20 (ต่อ)

ข้อ ที่	องค์ประกอบ				R^2	สปส.คะแนน	
	ความรู้		กระบวนการ			องค์ประกอบ	
	นน.องค์ประกอบ	SE	นน.องค์ประกอบ	SE		ความรู้	กระบวนการ
25	-	-	.84*	.10	.14	-	.04
26	-	-	.78*	.09	.12	-	.04
27	-	-	.86*	.10	.15	-	.04
28	-	-	.90*	.10	.16	-	.05
$\chi^2 = 184.29$		$df = 297$		$P = 1.00$	$GFI = .99$	$AGFI = .98$	
CFI = 1.00		NFI = .94		RMSEA = .00			

* $p < .05$

จากภาพที่ 9 และตารางที่ 20 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของข้อสอบวัดผลสัมฤทธิ์ทางการเรียนของเพศหญิง ($n=1000$) พบว่าโมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีมาก โดยพิจารณาจากค่าสถิติไค-สแควร์ เท่ากับ 184.29 มีค่าความน่าจะเป็นเท่ากับ 1.00 ที่องศาอิสระเท่ากับ 297 นั่นคือค่าไค-สแควร์ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ แสดงว่ายอมรับสมมติฐานหลักที่ว่าโมเดลการวิจัยสอดคล้องกับข้อมูลเชิงประจักษ์ ส่วนค่าดัชนีวัดระดับความกลมกลืน (GFI) เท่ากับ .99 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) เท่ากับ .98 ค่าดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) เท่ากับ 1.00 ค่าดัชนีวัดความสอดคล้องปกติ (NFI) เท่ากับ .94 และค่ารากของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนโดยประมาณ (RMSEA) เท่ากับ .00 ซึ่งมีค่าน้อยมากจนมีค่าเท่ากับศูนย์ แสดงว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี

เมื่อพิจารณาน้ำหนักองค์ประกอบ ในแต่ละองค์ประกอบ พบว่า

องค์ประกอบด้านความรู้ มีข้อสอบ จำนวน 12 ข้อ (ข้อที่ 1 - 12) มีค่าน้ำหนักองค์ประกอบระหว่าง .54 ถึง .81 (ข้อที่ 11 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

องค์ประกอบด้านกระบวนการ มีข้อสอบ จำนวน 16 ข้อ (ข้อที่ 13 - 28) มีค่าน้ำหนักองค์ประกอบระหว่าง .49 ถึง .97 (ข้อที่ 15 กำหนดให้เป็นค่าคงที่) โดยข้อสอบทุกข้อมีนัยสำคัญทางสถิติ

ตารางที่ 21 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ของแบบทดสอบฉบับหลังตัด
ข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ สำหรับเพศหญิง (ขั้นที่ 2)

องค์ประกอบ	เมทริกซ์น้ำหนักองค์ประกอบ		R^2
	น้ำหนักองค์ประกอบ	SE	
เพศหญิง ($n=1000$)			
ด้านความรู้	.56	.03	1.00
ด้านกระบวนการ	.45	.03	1.00
$\chi^2 = 184.29$ $df = 297$ $P = 1.00$ GFI = .99 AGFI = .98			
CFI = 1.00 NFI = .94 RMSEA = .00			

จากตารางที่ 21 พบว่า ค่าน้ำหนักองค์ประกอบด้านความรู้มีค่าเท่ากับ .56 และค่าน้ำหนักองค์ประกอบด้านกระบวนการมีค่าเท่ากับ .45 และทั้งสององค์ประกอบมีนัยสำคัญทางสถิติ โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีมาก แสดงว่าโครงสร้างของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย สำหรับเพศหญิงประกอบด้วยสององค์ประกอบ ตรงตามโครงสร้างของแบบทดสอบ

ตอนที่ 5 ผลการเปรียบเทียบคุณภาพของแบบทดสอบ ฉบับก่อนกับฉบับหลังตัด
ข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ

5.1 ผลการเปรียบเทียบค่าความเที่ยงของแบบทดสอบวิชาภาษาไทย ฉบับก่อนกับ
ฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ แสดงดังตารางที่ 22

ตารางที่ 22 ผลการเปรียบเทียบค่าความเที่ยงของแบบทดสอบวิชาภาษาไทย ฉบับก่อนกับฉบับหลัง
ตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ

ค่าความเที่ยงแบบสัมประสิทธิ์แอลฟาครอนบาค		t
ฉบับก่อน	ฉบับหลัง	
.86	.87	- 8.94*

$$t(1,999) = \pm 1.96, p < .05$$

* $p < .05$

จากตารางที่ 22 พบว่า แบบทดสอบฉบับก่อนตรวจสอบการทำหน้าที่ต่างกันของข้อสอบมีค่าความเที่ยงเท่ากับ .86 แบบทดสอบฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบมีค่าความเที่ยงเท่ากับ .87 เมื่อเปรียบเทียบค่าความเที่ยงโดยใช้สถิติทดสอบที (Fleiss, 1987 อ้างถึงใน เสรี ชัดแจ้ง, 2544) ได้ค่าสถิติทดสอบทีเท่ากับ 8.94 และค่าที่ได้มีนัยสำคัญทางสถิติ หมายความว่า เมื่อตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ ทำให้ค่าความเที่ยงของแบบทดสอบฉบับหลังมีค่าความเที่ยงแตกต่างจากแบบทดสอบฉบับก่อนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

5.2 ผลการเปรียบเทียบความตรงเชิงโครงสร้างของแบบทดสอบวิชาภาษาไทย ฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ แสดงดังตารางที่ 23

ตารางที่ 23 เปรียบเทียบค่าสถิติไค-สแควร์สอดคล้องระหว่างโมเดลแบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ ($N = 2000$)

โมเดล	χ^2	df	GFI	CFI	χ^2/df	χ^2_{diff}	ΔGFI
1. ฉบับก่อนตัด (40 ข้อ)	552.90	654	.99	1.00	.85		
2. ฉบับหลังตัด (28 ข้อ)	237.18	307	.99	1.00	.77		
ความแตกต่างระหว่าง							
โมเดล 1 กับ โมเดล 2						315.72*	.00
$\chi^2 (347, N = 2000) = 124.30, p < .05$							

* $p < .05$

จากตารางที่ 23 ผลการเปรียบเทียบค่าสถิติไค-สแควร์สอดคล้องระหว่างโมเดลของแบบทดสอบฉบับก่อนตัดข้อสอบที่พบ DIF กับโมเดลของแบบทดสอบฉบับหลังตัดข้อสอบที่พบ DIF ผลปรากฏว่า โมเดลแบบทดสอบฉบับก่อนตัดข้อสอบที่พบ DIF มีค่าสถิติค่าไค-สแควร์เท่ากับ 552.90 ที่องศาอิสระเท่ากับ 654 ส่วนโมเดลแบบทดสอบฉบับหลังตัดข้อสอบที่พบ DIF มีค่าสถิติค่าไค-สแควร์เท่ากับ 237.18 ที่องศาอิสระเท่ากับ 307 และค่าไค-สแควร์สอดคล้องเท่ากับ 315.72 ที่องศาอิสระเท่ากับ 347 ค่าไค-สแควร์สอดคล้องมีนัยสำคัญทางสถิติ $\chi^2 (347, N = 2000) = 124.30, p < .05$ แสดงว่าโมเดลแบบทดสอบฉบับก่อนตัดข้อสอบที่พบ DIF กับโมเดลแบบทดสอบฉบับหลังตัดข้อสอบที่พบ DIF มีความสอดคล้องกับข้อมูลเชิงประจักษ์แตกต่างกัน โดยโมเดลของแบบทดสอบฉบับหลังตัด DIF มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่า (พิจารณาจากค่า

องศาอิสระของโมเดล 2 น้อยกว่า) แต่ทั้งสองโมเดลประกอบด้วยสององค์ประกอบเช่นเดียวกัน เมื่อพิจารณาค่าสถิติวัดระดับความสอดคล้องของแบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบ ทำหน้าที่ต่างกันออกจากแบบทดสอบ ($N=2000$) ค่าสถิติที่ได้ไม่แตกต่างกันมากนัก ถือว่ามีค่าใกล้เคียงกัน แสดงว่าเมื่อตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบไม่ได้ทำให้โครงสร้างของแบบทดสอบเปลี่ยนแปลงไปจากเดิม

5.3 ผลการเปรียบเทียบความคงที่ของโครงสร้างองค์ประกอบ ของแบบทดสอบวิชาภาษาไทย ฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ เมื่อจำแนกกลุ่มประชากรตามตัวแปรเพศ

5.3.1 เปรียบเทียบความคงที่ของโครงสร้างองค์ประกอบ ฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ สำหรับเพศชาย

ตารางที่ 24 เปรียบเทียบค่าสถิติไค-สแควร์สอดแทรกระหว่างโมเดลแบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ สำหรับเพศชาย

โมเดล	χ^2	df	GFI	CFI	χ^2/df	χ^2_{diff}	ΔGFI
1. ฉบับก่อนตัด (40 ข้อ)	540.65	676	.97	1.00	.80		
2. ฉบับหลังตัด (28 ข้อ)	240.74	303	.98	1.00	.79		
ความแตกต่างระหว่าง							
โมเดล 1 กับ โมเดล 2						299.91*	.01
$\chi^2 (373, n=1000) = 124.30, p < .05$							

* $p < .05$

จากตารางที่ 24 ผลการเปรียบเทียบค่าสถิติไค-สแควร์สอดแทรกระหว่างโมเดลของแบบทดสอบฉบับก่อนตัดข้อสอบที่พบ DIF กับโมเดลของแบบทดสอบฉบับหลังตัดข้อสอบที่พบ DIF สำหรับเพศชาย ผลปรากฏว่า โมเดลแบบทดสอบฉบับก่อนตัดข้อสอบที่พบ DIF มีค่าสถิติไค-สแควร์เท่ากับ 540.65 ที่องศาอิสระเท่ากับ 676 ส่วนโมเดลแบบทดสอบฉบับหลังตัดข้อสอบที่พบ DIF มีค่าสถิติไค-สแควร์เท่ากับ 240.74 ที่องศาอิสระเท่ากับ 303 และค่าไค-สแควร์สอดแทรกเท่ากับ 299.91 ที่องศาอิสระเท่ากับ 373 ค่าไค-สแควร์สอดแทรกมีนัยสำคัญทางสถิติ $\chi^2 (371, n=1000) = 124.30, p < .05$ แสดงว่าโมเดลแบบทดสอบฉบับก่อนตัดข้อสอบที่พบ DIF กับโมเดลแบบทดสอบฉบับหลังตัดข้อสอบที่พบ DIF มีความสอดคล้องกับข้อมูลเชิงประจักษ์

แตกต่างกัน โดยโมเดลของแบบทดสอบฉบับหลังตัด DIF มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่า (พิจารณาจากค่าองศาอิสระของโมเดล 2 น้อยกว่า) แต่ทั้งสองโมเดลประกอบด้วยสององค์ประกอบเช่นเดียวกัน เมื่อพิจารณาค่าสถิติวัดระดับความสอดคล้องของโครงสร้างแบบทดสอบฉบับก่อนกับโครงสร้างแบบทดสอบฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบสำหรับเพศชาย ($n = 1000$) ค่าสถิติที่ได้ไม่แตกต่างกันมากนัก ถือว่ามีค่าใกล้เคียงกัน นั้นแสดงว่าเมื่อตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบไม่ได้ทำให้โครงสร้างของแบบทดสอบเปลี่ยนไปจากเดิม

5.3.2 เปรียบเทียบความคงที่ของโครงสร้างองค์ประกอบ ฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ สำหรับเพศหญิง

ตารางที่ 25 เปรียบเทียบค่าสถิติไค-สแควร์สอดแทรกระหว่างโมเดลแบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ต่างกันออกจากแบบทดสอบ สำหรับเพศหญิง

โมเดล	χ^2	df	GFI	CFI	χ^2/df	χ^2_{diff}	ΔGFI
1. ฉบับก่อนตัด (40 ข้อ)	515.19	668	.98	1.00	.77		
2. ฉบับหลังตัด (28 ข้อ)	184.29	297	.99	1.00	.62		
ความแตกต่างระหว่าง							
โมเดล 1 กับโมเดล 2						330.90*	.01
$\chi^2 (371, n=1000) = 124.30, p < .05$							

* $p < .05$

จากตารางที่ 25 ผลการเปรียบเทียบค่าสถิติไค-สแควร์สอดแทรกระหว่างโมเดลของแบบทดสอบฉบับก่อนตัดข้อสอบที่พบ DIF กับโมเดลของแบบทดสอบฉบับหลังตัดข้อสอบที่พบ DIF สำหรับเพศหญิง ผลปรากฏว่า โมเดลแบบทดสอบฉบับก่อนตัดข้อสอบที่พบ DIF มีค่าสถิติไค-สแควร์เท่ากับ 515.19 ที่องศาอิสระเท่ากับ 668 ส่วนโมเดลแบบทดสอบฉบับหลังตัดข้อสอบที่พบ DIF มีค่าสถิติไค-สแควร์เท่ากับ 184.29 ที่องศาอิสระเท่ากับ 297 และค่าไค-สแควร์สอดแทรกเท่ากับ 330.90 ที่องศาอิสระเท่ากับ 371 ค่าไค-สแควร์สอดแทรกมีนัยสำคัญทางสถิติ $\chi^2 (371, n = 1000) = 124.30, p < .05$ แสดงว่าโมเดลแบบทดสอบฉบับก่อนตัดข้อสอบที่พบ DIF กับโมเดลแบบทดสอบฉบับหลังตัดข้อสอบที่พบ DIF มีความสอดคล้องกับข้อมูลเชิงประจักษ์แตกต่างกัน โดยโมเดลของแบบทดสอบฉบับหลังตัด DIF มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่า

(พิจารณาจากค่าองศาอิสระของโมเดล 2 น้อยกว่า) แต่ทั้งสองโมเดลประกอบด้วยสององค์ประกอบเช่นเดียวกัน เมื่อพิจารณาค่าสถิติวัดระดับความสอดคล้องของโครงสร้างแบบทดสอบฉบับก่อนกับโครงสร้างแบบทดสอบฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ สำหรับเพศหญิง ($n = 1000$) ค่าสถิติที่ได้ไม่แตกต่างกันมากนัก ถือว่ามีค่าใกล้เคียงกัน แสดงว่าเมื่อตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบไม่ได้ทำให้โครงสร้างของแบบทดสอบเปลี่ยนแปลงไป

ตอนที่ 6 ผลการวิเคราะห์ค่าสหสัมพันธ์อันดับของผู้สอบ ฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ

ผลการวิเคราะห์ค่าสหสัมพันธ์อันดับของผู้สอบ ของแบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบวิชาภาษาไทย ได้ผลการวิเคราะห์แสดงดังตารางที่ 26

ตารางที่ 26 ผลการวิเคราะห์ค่าสหสัมพันธ์อันดับของผู้สอบ ฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ

ตัวแปร	ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์แมนแรงค์ (r_s)
ค่าสหสัมพันธ์อันดับของผู้สอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ	.48*

* $p < .05$

จากตารางที่ 26 ผลการวิเคราะห์ความสัมพันธ์ของลำดับที่การสอบของผู้สอบ ที่ได้จากแบบทดสอบฉบับก่อนกับลำดับที่การสอบที่ได้จากแบบทดสอบฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ พบว่าลำดับที่การสอบของผู้สอบจากแบบทดสอบฉบับก่อนกับฉบับหลังตัดข้อสอบทำหน้าที่ย่างกันออกจากแบบทดสอบ มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยมีค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมนแรงค์เท่ากับ .48 ซึ่งเป็นค่าสัมประสิทธิ์สหสัมพันธ์ในระดับปานกลาง