

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เซล (MH) กับวิธีดัดดอยโลจิสติก(LR)ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกันในแบบสอบชนิดพหุมิติ โดยใช้ข้อมูลจากผลการตอบข้อสอบของนักเรียนชั้นประถมศึกษาปีที่ 6 ในสังกัดสำนักงานการประถมศึกษากรุงเทพมหานคร ปีการศึกษา 2540 จำนวน 1,076 คน ซึ่งเครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบความสามารถทางคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยสร้างขึ้น

ผลการวิเคราะห์ข้อมูลในบทนี้ ผู้วิจัยได้นำเสนอเป็น 6 ตอน ดังนี้

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐานของผลการตอบข้อสอบ

ตอนที่ 2 การวิเคราะห์คุณภาพของแบบสอบที่ใช้ในการวิจัย

ตอนที่ 3 การวิเคราะห์สหสัมพันธ์ระหว่างตัวแปรที่ใช้ในการวิจัย

ตอนที่ 4 การวิเคราะห์องค์ประกอบของแบบสอบ

ตอนที่ 5 เปรียบเทียบการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่าง

วิธีแมนเทิล - แฮนส์เซลและวิธีดัดดอยโลจิสติกตามเกณฑ์จับคู่เปรียบเทียบ

ตอนที่ 6 เปรียบเทียบอัตราความถูกต้องและอัตราความคลาดเคลื่อนระหว่าง




วิธีแมนเทิล - แฮนส์เซลกับวิธีดัดดอยโลจิสติกในการตรวจสอบการทำหน้าที่

ต่างกันของข้อสอบ ตามเกณฑ์จับคู่เปรียบเทียบ

ในการวิเคราะห์ข้อมูลและการแปลผลการวิเคราะห์ข้อมูลในการวิจัยครั้งนี้ เพื่อให้การนำเสนอผลการวิเคราะห์ข้อมูลมีความสะดวกและเกิดความเข้าใจตรงกันเกี่ยวกับผลการวิเคราะห์ข้อมูล ผู้วิจัยจึงได้กำหนดสัญลักษณ์และความหมายที่ใช้แทนตัวแปรต่าง ๆ ดังต่อไปนี้

X_i หมายถึง ข้อสอบข้อที่ i ในแบบทดสอบความสามารถทางคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6 เมื่อ $i = 1, 2, 3, \dots, 80$

K₁ หมายถึง องค์ประกอบที่วัดความรู้ความเข้าใจทางคณิตศาสตร์

K2	หมายถึง	องค์ประกอบที่วัดความสามารถด้านการคิดคำนวณ
K3	หมายถึง	องค์ประกอบที่วัดความสามารถด้านการพิจารณา ผลลัพธ์ที่สมเหตุสมผล
K4	หมายถึง	องค์ประกอบที่วัดความสามารถด้านการแก้โจทย์ปัญหา
SUBTEST1	หมายถึง	ผลการตอบข้อสอบชุดที่ 1 ซึ่งวัดในองค์ประกอบ K1
SUBTEST2	หมายถึง	ผลการตอบข้อสอบชุดที่ 2 ซึ่งวัดในองค์ประกอบ K2
SUBTEST3	หมายถึง	ผลการตอบข้อสอบชุดที่ 3 ซึ่งวัดในองค์ประกอบ K3
SUBTEST4	หมายถึง	ผลการตอบข้อสอบชุดที่ 4 ซึ่งวัดในองค์ประกอบ K4
TOTALTEST	หมายถึง	ผลการตอบข้อสอบที่วัดรวมทั้ง 4 องค์ประกอบ
MH _{TOT}	หมายถึง	การตรวจสอบ DIF ด้วยวิธีแมนเทิล - แฮนส์เซล โดยใช้ คะแนนรวม (total test score) เป็นเกณฑ์การจับคู่เปรียบเทียบ
MH _{SUB}	หมายถึง	การตรวจสอบ DIF ด้วยวิธีแมนเทิล - แฮนส์เซล โดยใช้ คะแนนแบบสอบย่อย (subtest score) เป็นเกณฑ์การจับคู่
LR _{TOT}	หมายถึง	การตรวจสอบ DIF ด้วยวิธีถดถอยโลจิสติก โดยใช้ คะแนนรวมเป็นเกณฑ์การจับคู่เปรียบเทียบ
LR _{SUB}	หมายถึง	การตรวจสอบ DIF ด้วยวิธีถดถอยโลจิสติก โดยใช้ คะแนนแบบสอบย่อยเป็นเกณฑ์การจับคู่
LR _{MULTIPLE}	หมายถึง	การตรวจสอบ DIF ด้วยวิธีถดถอยโลจิสติก โดยใช้ คะแนนหลายแบบสอบย่อยเป็นเกณฑ์การจับคู่
GFI	หมายถึง	ดัชนีวัดระดับความกลมกลืน
AGFI	หมายถึง	ดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว
df	หมายถึง	องศาอิสระ
	หมายถึง	ตัวแปรสังเกตได้
	หมายถึง	ตัวแปรแฝง
	หมายถึง	ความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรต้นกับตัวแปรตาม หัวข้อครุศาสตร์ศึกษาศาสตร์ของอิทธิพล

ตอนที่ 1 การวิเคราะห์ข้อมูลพื้นฐานของผลการตอบข้อสอบ

ผู้วิจัยเสนอผลการวิเคราะห์ข้อมูลพื้นฐานของผลการสอบแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ซึ่งมีลักษณะการกระจายของคะแนนดังตารางที่ 13

ตารางที่ 13 ค่าสถิติพื้นฐานของแบบทดสอบ

ค่าสถิติ	SUBTEST1	SUBTEST2	SUBTEST3	SUBTEST4	TOTALTEST
คะแนนเต็ม	20.00	20.00	20.00	20.00	80.00
มัธยิมเลขคณิต(Mean)	13.12	12.60	10.18	12.23	48.09
ส่วนเบี่ยงเบนมาตรฐาน(S.D.)	4.21	4.22	3.66	4.54	14.88
สัมประสิทธิ์การกระจาย(%CV)	32.08	33.49	35.95	37.12	30.94
ต่ำสุด(Min)	2.00	2.00	2.00	1.00	14.00
สูงสุด(Max)	20.00	20.00	20.00	20.00	78.00
พิสัย(Range)	18.00	18.00	18.00	19.00	64.00
มัธยฐาน(Median)	14.00	13.00	10.00	12.00	49.00
ฐานนิยม(Mode)	16.00	14.00	10.00	12.00	51.00
ความเบ้(Sk)	-0.32	-0.20	0.13	-0.18	-0.05
ความโด่ง(Ku)	-0.74	-0.78	-0.72	-0.92	-0.94
ค่าความเที่ยง(KR-20)	0.80	0.80	0.70	0.82	0.93

จากผลการวิเคราะห์ค่าสถิติพื้นฐานของแบบสอบทั้งฉบับ ซึ่งผู้วิจัยได้แยกวิเคราะห์ตามองค์ประกอบในการวัดแต่ละด้านทั้ง 4 ด้านและวิเคราะห์รวมทั้งฉบับ เมื่อพิจารณาผลการตอบข้อสอบเปรียบเทียบระหว่างแบบสอบย่อยแต่ละด้านซึ่งมีคะแนนเต็ม 20 คะแนน ปรากฏว่า SUBTEST1 ผู้สอบมีค่าเฉลี่ยของคะแนนสอบสูงสุด ($\bar{X} = 13.12$) รองลงมา คือ SUBTEST2, SUBTEST4 และ SUBTEST3 ตามลำดับ โดยคะแนนจาก SUBTEST4 มีการกระจายออกจากค่าเฉลี่ยมากที่สุด (S.D. = 4.54) และคะแนนจาก SUBTEST3 มีการกระจายออกจากค่าเฉลี่ยน้อยที่สุด (S.D. = 3.66) ส่วนการแจกแจงคะแนนพบว่า SUBTEST1 SUBTEST2 และ SUBTEST4 มีลักษณะเบ้ซ้ายเล็กน้อย (Sk = -0.32, -0.20 และ -0.18 ตามลำดับ) ค่อนข้างแบน (Ku = -0.74, -0.78 และ -0.92 ตามลำดับ) ส่วนการแจกแจงคะแนนของ SUBTEST3 มีลักษณะเบ้ขวาเล็กน้อย (Sk = 0.13) และค่อนข้างแบน (Ku = -0.72) เมื่อพิจารณาผลการตอบข้อสอบของแบบสอบทั้งฉบับ

พบว่าจากคะแนนเต็ม 80 คะแนน ผู้สอบมีค่าเฉลี่ยของคะแนนสอบ เท่ากับ 48.09 คะแนน โดยคะแนนมีการกระจายออกจากค่าเฉลี่ยประมาณ 14.88 คะแนน มีพิสัยของคะแนนเท่ากับ 64 โดยมีคะแนนต่ำสุดเท่ากับ 14 และคะแนนสูงสุดเท่ากับ 78 ส่วนการแจกแจงคะแนนมีลักษณะเบ้ซ้ายเล็กน้อย ($S_k = -0.05$) ค่อนข้างแบน ($K_u = -0.94$) สำหรับผลการวิเคราะห์ค่าความเที่ยงของแบบสอบย่อยแต่ละด้านและวิเคราะห์รวมทั้งฉบับซึ่งคำนวณแบบ KR-20 พบว่า ค่าความเที่ยงของแบบสอบย่อยแต่ละด้านอยู่ในเกณฑ์ดี โดย SUBTEST 4 มีค่าความเที่ยงสูงสุด ($KR-20 = 0.82$) เมื่อพิจารณาค่าความเที่ยงรวมทั้งฉบับพบว่ามีค่าความเที่ยง เท่ากับ 0.93

ตารางที่ 14 เปรียบเทียบค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนระหว่างกลุ่มผู้สอบ

องค์ประกอบในการวัด	เพศหญิง	เพศชาย
ความรู้ความเข้าใจทางคณิตศาสตร์ (K1)		
ค่าเฉลี่ย (\bar{X})	13.06	13.18
(S.D.)	(4.29)	(4.16)
ความสามารถด้านการคิดคำนวณ (K2)		
ค่าเฉลี่ย (\bar{X})	12.69	12.52
(S.D.)	(4.16)	(4.28)
ความสามารถด้านการพิจารณาผลลัพธ์ที่สมเหตุสมผล(K3)		
ค่าเฉลี่ย (\bar{X})	10.30	10.08
(S.D.)	(3.61)	(4.61)
ความสามารถด้านการแก้โจทย์ปัญหา (K4)		
ค่าเฉลี่ย (\bar{X})	12.63	11.99
(S.D.)	(4.43)	(4.61)
แบบทดสอบความสามารถทางคณิตศาสตร์ชั้นป. 6 (TOTAL TEST)		
ค่าเฉลี่ย (\bar{X})	48.59	47.77
(S.D.)	(14.85)	(14.91)

จากตารางที่ 14 เมื่อนำผลการตอบข้อสอบมาวิเคราะห์เพื่อเปรียบเทียบระหว่างกลุ่มผู้สอบ โดยจำแนกตามเพศ ซึ่งเป็นเพศหญิงจำนวน 483 คน และเพศชายจำนวน 593 คน คิดเป็นร้อยละ

44.9 และ 55.1 ตามลำดับ พบว่าใน SUBTEST1 เพศชายมีคะแนนเฉลี่ยสูงกว่าเพศหญิง ($\bar{X} = 13.18$) และมีค่าส่วนเบี่ยงเบนมาตรฐานต่ำกว่าเพศหญิง (S.D. = 4.16) ส่วนใน SUBTEST2, SUBTEST3, SUBTEST4 และ TOTAL TEST พบว่าเพศหญิงมีคะแนนเฉลี่ยสูงกว่าเพศชาย ($\bar{X} = 12.69, 10.30, 12.53$ และ 48.59 ตามลำดับ) มีค่าส่วนเบี่ยงเบนมาตรฐานต่ำกว่าเพศชาย (S.D. = 4.16, 3.61, 4.43 และ 14.85 ตามลำดับ)

ตอนที่ 2 การวิเคราะห์คุณภาพของแบบสอบที่ใช้ในการวิจัย

การวิเคราะห์คุณภาพของแบบวัดความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยสร้างขึ้นวิเคราะห์โดยใช้โปรแกรมสำเร็จรูป CTIA ซึ่งผู้วิจัยจะนำเสนอผลการวิเคราะห์ในด้านค่าความยาก (p) ค่าอำนาจจำแนก (r) รายข้อ และค่าความเที่ยงแบบความสอดคล้องภายใน (KR-20) ของแบบวัด โดยนำผลการตอบข้อสอบของผู้สอบแต่ละคนจากแบบสอบทั้งฉบับ A และฉบับ B มารวมเป็นข้อมูลชุดเดียวกัน ซึ่งผลการตอบข้อสอบฉบับ A จะเป็นผลการตอบข้อสอบตั้งแต่ข้อที่ 1 ถึง ข้อที่ 40 ส่วนผลการตอบข้อสอบฉบับ B จะเป็นผลการตอบข้อสอบตั้งแต่ข้อที่ 41 ถึงข้อที่ 80 ส่วน ซึ่งมีรายละเอียดดังตารางที่ 15

ตารางที่ 15 ค่าความยาก(p) ค่าอำนาจจำแนก(r) ของแบบทดสอบ

องค์ประกอบในการวัด	แบบสอบฉบับ A			แบบสอบฉบับ B		
	ข้อ	p	r	ข้อ	p	r
1. ความรู้ความเข้าใจทางคณิตศาสตร์	1	0.785	0.452	41	0.439	0.341
	2	0.578	0.638	42	0.820	0.328
	3	0.632	0.703	43	0.376	0.579
	4	0.669	0.586	44	0.803	0.421
	5	0.744	0.369	45	0.832	0.428
	6	0.541	0.624	46	0.626	0.490
	7	0.548	0.714	47	0.634	0.631
	8	0.773	0.369	48	0.507	0.607
	9	0.665	0.676	49	0.667	0.710
	10	0.669	0.576	50	0.824	0.307

ตารางที่ 15 (ต่อ)

องค์ประกอบในการวัด	แบบสอบฉบับ A			แบบสอบฉบับ B		
	ข้อ	p	r	ข้อ	p	r
2. การคิดคำนวณ	11	0.848	0.234	61	0.705	0.452
	12	0.681	0.497	62	0.303	0.465
	13	0.610	0.652	63	0.634	0.293
	14	0.329	0.524	64	0.721	0.572
	15	0.609	0.721	65	0.648	0.634
	16	0.408	0.566	66	0.509	0.659
	17	0.784	0.455	67	0.862	0.286
	18	0.715	0.541	68	0.555	0.659
	19	0.677	0.621	69	0.635	0.524
	20	0.709	0.590	70	0.660	0.734
3. การพิจารณาผลลัพธ์ อย่างสมเหตุสมผล	21	0.511	0.521	61	0.676	0.600
	22	0.270	0.210	62	0.401	0.459
	23	0.360	0.438	63	0.564	0.666
	24	0.599	0.541	64	0.599	0.669
	25	0.580	0.272	65	0.399	0.431
	26	0.503	0.448	66	0.644	0.507
	27	0.452	0.621	67	0.562	0.679
	28	0.464	0.462	68	0.358	0.279
	29	0.889	0.259	69	0.485	0.693
	30	0.624	0.400	70	0.241	0.234
4. การแก้ไขข้อผิดพลาด	31	0.663	0.634	71	0.494	0.538
	32	0.690	0.607	72	0.664	0.572
	33	0.578	0.628	73	0.404	0.407
	34	0.481	0.617	74	0.598	0.583
	35	0.630	0.621	75	0.812	0.290

ตารางที่ 15 (ต่อ)

องค์ประกอบในการวัด	แบบสอบฉบับ A			แบบสอบฉบับ B		
	ข้อ	p	r	ข้อ	p	r
4. การแก้ปัญหา(ต่อ)	36	0.655	0.310	76	0.623	0.617
	37	0.372	0.462	77	0.716	0.548
	38	0.640	0.634	78	0.685	0.652
	39	0.679	0.721	79	0.717	0.614
	40	0.513	0.659	80	0.618	0.634
ค่าความเที่ยงของแบบสอบ (KR-20)	0.878			0.874		
ค่าความคลาดเคลื่อนในการวัด	2.734			2.721		

จากตารางที่ 15 ผลการวิเคราะห์คุณภาพของแบบวัดความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 จำนวน 80 ข้อ เมื่อพิจารณาผลการวิเคราะห์คุณภาพรายข้อในด้านค่าความยาก(p) จะเห็นว่าข้อที่มีค่าความยากสูงสุด ได้แก่ ข้อสอบข้อที่ 29 โดยมีค่าความยากเท่ากับ 0.889 และข้อสอบที่มีค่าความยากต่ำสุด ได้แก่ ข้อสอบข้อที่ 70 โดยมีค่าความยากเท่ากับ 0.241 ส่วนค่าอำนาจจำแนก(g) จะเห็นว่าข้อที่มีค่าอำนาจจำแนกสูงสุดได้แก่ ข้อสอบข้อที่ 60 โดยมีค่าเท่ากับ 0.734 และข้อสอบที่มีค่าอำนาจจำแนกต่ำสุด ได้แก่ ข้อสอบข้อที่ 22 โดยมีค่าเท่ากับ 0.210 สำหรับผลการวิเคราะห์ค่าความเที่ยงของแบบสอบพบว่า ค่าความเที่ยงของแบบสอบทั้งสองฉบับอยู่ในเกณฑ์ดีและมีค่าใกล้เคียงกัน โดยแบบสอบฉบับ A มีค่าความเที่ยงเท่ากับ 0.878 และแบบสอบฉบับ B ค่าความเที่ยงเท่ากับ 0.874 มีค่าความคลาดเคลื่อนในการวัด เท่ากับ 2.734 และ 2.721 ตามลำดับ

ตอนที่ 3 ผลการวิเคราะห์สหสัมพันธ์ระหว่างตัวแปรที่ใช้ในการวิจัย

ก่อนที่จะทำการวิเคราะห์องค์ประกอบผู้วิจัยได้ศึกษาความสัมพันธ์ระหว่างตัวแปร เพื่อพิจารณาถึงความเหมาะสมของเมทริกซ์สหสัมพันธ์ที่จะนำไปใช้ในการวิเคราะห์องค์ประกอบว่ามีความเหมาะสมหรือไม่ โดยพิจารณาความสัมพันธ์ระหว่างข้อสอบที่จะนำไปวิเคราะห์ว่ามีค่าสัมประสิทธิ์สหสัมพันธ์แตกต่างจากศูนย์หรือไม่ ถ้าหากว่าข้อสอบที่สร้างไม่มีความสัมพันธ์กัน แสดงว่าไม่มีองค์ประกอบร่วมและไม่มีประโยชน์ที่จะนำเมทริกซ์นั้นไปวิเคราะห์(เนงลักษณะ วิรัชชัย, 2539) ค่าสถิติที่จะใช้พิจารณา ได้แก่ ค่าสถิติของ Bartlett ซึ่งเป็นค่าสถิติทดสอบสมมติฐานว่า

เมทริกซ์สหสัมพันธ์นั้นเป็นเมทริกซ์เอกลักษณ์(identity matrix)หรือไม่ โดยพิจารณาที่ค่า Bartlett's test of sphericity และค่าความน่าจะเป็น นอกจากนี้ยังพิจารณาได้จากค่าดัชนีไกเซอร์-ไมเยอร์-ออลคิน(Kaiser - Meyer - Olkin measure of sampling adequacy = MSA) โดย Kim, Mueller (1978) ได้เสนอไว้ว่า ถ้ามีค่ามากกว่า 0.80 ดีมากและถ้ามีค่าน้อยกว่า 0.50 ใช้ไม่ได้

จากการวิเคราะห์ความสัมพันธ์ระหว่างข้อสอบทุกข้อในแบบสอบที่ผู้วิจัยสร้างขึ้น โดยใช้โปรแกรม SPSS for WINDOW version 7.0 พบว่าข้อสอบส่วนใหญ่มีความสัมพันธ์กันอย่างมีนัยสำคัญ($P < 0.05, 0.01, 0.001$) สำหรับค่าสถิติที่ใช้พิจารณาความเหมาะสม พบว่า ค่า Bartlett's test of sphericity มีค่าเท่ากับ 17191.911 และมีค่าความน่าจะเป็นน้อยกว่า .000 ($P < 0.000$) แสดงว่า เมทริกซ์สหสัมพันธ์ระหว่างข้อสอบแตกต่างจากเมทริกซ์เอกลักษณ์อย่างมีนัยสำคัญ ส่วนค่าดัชนีไกเซอร์-ไมเยอร์-ออลคิน มีค่าเท่ากับ 0.942 ซึ่งมีค่ามากกว่า 0.80 นั้นแสดงว่าข้อสอบในแบบสอบฉบับนี้มีความสัมพันธ์กันเหมาะสมดีมากสามารถที่จะนำไปวิเคราะห์องค์ประกอบได้

นอกจากนี้ผู้วิจัยยังได้ศึกษาความสัมพันธ์ระหว่างข้อสอบที่เมืองค์ประกอบในการวัดเดียวกันว่า มีความสัมพันธ์กันมากน้อยเพียงใด พบว่า องค์ประกอบด้านความรู้ความเข้าใจทางคณิตศาสตร์ (K1) ข้อสอบส่วนใหญ่มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, .01 และ .001 โดยมีความสัมพันธ์ของข้อสอบจำนวน 12 คู่ที่ไม่มีนัยสำคัญทางสถิติ ข้อสอบที่มีความสัมพันธ์กันสูงสุด คือ ข้อที่ 9(X9) กับข้อที่ 49(X49) มีค่าความสัมพันธ์เท่ากับ .3634 ($p < 0.001$)

องค์ประกอบด้านการคิดคำนวณ (K2) พบว่า ข้อสอบส่วนใหญ่มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, .01 และ .001 โดยมีความสัมพันธ์ของข้อสอบจำนวน 9 คู่ที่ไม่มีนัยสำคัญทางสถิติ ข้อสอบที่มีความสัมพันธ์กันสูงสุดคือ ข้อที่ 13(X13) กับข้อที่ 15(X15) มีค่าความสัมพันธ์เท่ากับ .3305 ($p < 0.001$)

องค์ประกอบด้านการพิจารณาผลลัพธ์ที่สมเหตุสมผล (K3) พบว่า ข้อสอบส่วนใหญ่มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, .01 และ .001 โดยมีความสัมพันธ์ของข้อสอบจำนวน 24 คู่ที่ไม่มีนัยสำคัญทางสถิติ ข้อสอบที่มีความสัมพันธ์กันสูงสุดคือ ข้อที่ 63(X63)กับข้อที่ 66(X66) มีค่าความสัมพันธ์เท่ากับ .3464 ($p < 0.001$) และองค์ประกอบด้านการแก้โจทย์ปัญหา(K4) พบว่า ข้อสอบส่วนใหญ่มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, .01 และ .001 โดยมีความสัมพันธ์ของข้อสอบจำนวน 7 คู่ที่ไม่มีนัยสำคัญทางสถิติ ข้อสอบที่มีความสัมพันธ์กันสูงสุดคือ ข้อที่ 78(X78)กับข้อที่ 79(X79) มีค่าความสัมพันธ์เท่ากับ .3680 ($p < 0.001$) จากการวิเคราะห์ความสัมพันธ์ดังกล่าว จะได้เมทริกซ์สหสัมพันธ์ของแต่ละองค์ประกอบที่นำไปใช้ในการวิเคราะห์องค์ประกอบเชิงยืนยัน ปากฎดังตารางที่ 16 - 19

ตารางที่ 16 เมทริกซ์สหสัมพันธ์ระหว่างตัวแปรใน SUBTEST1

ตัวแปร	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X41	X42	X43	X44	X45	X46	X47	X48	X49	X50	
X1	1.0000																				
X2	.2178**	1.0000																			
X3	.2252**	.3079**	1.0000																		
X4	.2577**	.2111**	.3112**	1.0000																	
X5	.0932*	.0732	.1272**	.1178**	1.0000																
X6	.1224**	.2438**	.2482**	.2044**	.1614**	1.0000															
X7	.1849**	.2834**	.2600**	.2628**	.1704**	.2693**	1.0000														
X8	.1763**	.1395**	.2034**	.1475**	.0389	.1112**	.1908**	1.0000													
X9	.1684**	.2315**	.2451**	.2006**	.1263**	.2302**	.2924**	.2235**	1.0000												
X10	.2096**	.1951**	.1760**	.1856**	.1450**	.2123**	.2429**	.1900**	.3003**	1.0000											
X41	.0745	.0813	.0921*	.1002*	.0843*	.1154**	.0647	-.0043	.0896*	.0723	1.0000										
X42	.1198**	.1574**	.1735**	.1378**	.0744	.0918*	.1816**	.1271**	.1328**	.1429**	.0346	1.0000									
X43	.2193**	.2054**	.2468**	.1916**	.1078**	.1768**	.1847**	.1184**	.1761**	.1998**	.1096**	.1049**	1.0000								
X44	.1393**	.1824**	.1888**	.1880**	.0901*	.1391**	.2077**	.1279**	.1519**	.1433**	.1607**	.1324**	.1919**	1.0000							
X45	.2187**	.1692**	.2699**	.1907**	.0441	.1840**	.2409**	.1599**	.2645**	.2013**	.0821*	.1675**	.1391**	.1895**	1.0000						
X46	.1483**	.1649**	.1556**	.1756**	.0628	.1019**	.1831**	.0497	.1794**	.1347**	.1175**	.1375**	.1955**	.1487**	.1560**	1.0000					
X47	.2275**	.2490**	.2160**	.2035**	.1561**	.2250**	.2366**	.1504**	.2320**	.1871**	.0227	.1654**	.2003**	.1813**	.2410**	.1467**	1.0000				
X48	.1911**	.1971**	.1964**	.2396**	.1089**	.1666**	.2284**	.1057**	.3100**	.1645**	.0580	.1617**	.1976**	.1522**	.1631**	.1460**	.2004**	1.0000			
X49	.2024**	.2794**	.3241**	.2958**	.1153**	.2598**	.3222**	.1546**	.3634**	.2413**	.0995*	.1716**	.1903**	.2106**	.2994**	.1804**	.2862**	.3458**	1.0000		
X50	.1512**	.1002*	.0529	.0959*	.0879*	.1187**	.1356**	.1175**	.1739**	.0855*	.0684	.0630	.1217**	.1336**	.0993*	.0676	.1561**	.0972*	.1717**	1.0000	
X	.7853	.5781	.6320	.6691	.7444	.5409	.5483	.7732	.6652	.6691	.4387	.8197	.3764	.8230	.8318	.6264	.6338	.5074	.6673	.8244	
S.D.	.41	.49	.48	.47	.44	.50	.50	.42	.48	.47	.50	.38	.48	.40	.37	.48	.48	.50	.47	.38	

* p < 0.01, ** p < 0.001

ตารางที่ 17 เมทริกซ์สหสัมพันธ์ระหว่างตัวแปรใน SUBTEST 2

	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X51	X52	X53	X54	X55	X56	X57	X58	X59	X60
X12	1.0000																			
X13	.0794*	1.0000																		
X14	.0933*	.1012**	1.0000																	
X15	.1740**	.1871**	.3305**	1.0000																
X16	.0574	.0444	.1797**	.2196**	1.0000															
X17	.1361**	.1263**	.2059**	.1122**	.2743**	1.0000														
X18	.0699	.1463**	.2412**	.1577**	.2736**	.1728**	1.0000													
X19	.1048**	.2390**	.2491**	.1712**	.2925**	.1738**	.2124**	1.0000												
X20	.1269**	.2118**	.2384**	.2089**	.2161**	.1153**	.3011**	.2388**	1.0000											
X51	.0549	.2010**	.1850**	.1141**	.1544**	.1217**	.1767**	.1560**	.2113**	1.0000										
X52	.0601	.1038**	.1503**	.1539**	.1722**	.0987*	.1785**	.1568**	.1446**	.1506**	1.0000									
X53	.0105	-.0149	.0325	.1340**	.0942*	.0500	-.0092	.0666	.0907*	.0144	-.0596	.0267	1.0000							
X54	.1054**	.1617**	.1993**	.1707**	.2447**	.1535**	.2032**	.2772**	.1860**	.2452**	.2392**	.1168**	.0308	1.0000						
X55	.1023**	.1691**	.2242**	.1710**	.2745**	.1797**	.1836**	.2251**	.1943**	.2232**	.1524**	.1583**	.0221	.2937**	1.0000					
X56	.0544	.1703**	.2283**	.2363**	.1920**	.2096**	.1318**	.1620**	.2076**	.2227**	.1649**	.1698**	.0952*	.1650**	.3228**	1.0000				
X57	.1145**	.2050**	.1149**	.0459	.0976*	-.0011	.1431**	.0923*	.1600**	.1579**	.2544**	.1238**	-.0478	.1828**	.1083**	.1447**	1.0000			
X58	.0676	.1256**	.1879**	.1934**	.2705**	.2071**	.1761**	.1712**	.1602**	.1962**	.1308**	.2243**	.0567	.2396**	.2948**	.2729**	.1065**	1.0000		
X59	.1133**	.1272**	.2081**	.1532**	.2461**	.1388**	.1327**	.1633**	.1729**	.1729**	.0644	.1095**	.0324	.1912**	.2012**	.2052**	.1318**	.1750**	1.0000	
X60	.1485**	.1571**	.2338**	.2230**	.2886**	.2328**	.2628**	.3153**	.3172**	.2874**	.1944**	.2268**	.0651	.2885**	.2847**	.2566**	.1892**	.3397**	.1887**	1.0000
X	.8476	.6812	.6097	.3290	.6087	.4080	.7844	.7147	.6766	.7091	.7052	.3030	.6338	.7212	.6459	.5093	.8615	.5548	.6348	.6599
S.D.	.36	.47	.49	.47	.49	.49	.41	.45	.47	.45	.46	.46	.48	.45	.48	.50	.35	.50	.48	.47

* p < 0.01, ** p < 0.001

ตารางที่ 18 เมทริกซ์สัมพันธ์ระหว่างตัวแปรใน SUBTEST 3

ตัวแปร	X21	X22	X23	X24	X25	X26	X27	X28	X29	X30	X61	X62	X63	X64	X65	X66	X67	X68	X69	X70
X21	1.0000																			
X22	.0848*	1.0000																		
X23	.1905**	-.0421	1.0000																	
X24	.1226**	.0024	.1068**	1.0000																
X25	.0152	.0180	-.0095	.0305	1.0000															
X26	.0426	-.0390	.1256**	.1089**	.0462	1.0000														
X27	.2114**	.0192	.1564**	.1969**	.0498	.0883*	1.0000													
X28	.1377**	.0925*	.0409	.0528	.0061	.1010**	.1184**	1.0000												
X29	.0583	-.0054	.0235	.1290**	.0781	.0879*	.1235**	.0784	1.0000											
X30	.0538	-.0020	.0946*	.0852*	-.0083	.0715	.1038**	.0647	.1603**	1.0000										
X61	.1247**	.0554	.1511**	.1426**	.0539	.1091**	.1781**	.1029**	.0975*	.1215**	1.0000									
X62	.0975*	-.0067	.0948*	.1109**	.0117	.1036**	.1613**	.1374**	.1129**	.1105**	.1329**	1.0000								
X63	.2090**	.0711	.1667**	.1535**	-.0304	.1455**	.2404**	.1409**	.1143**	.0947*	.2918**	.1333**	1.0000							
X64	.1358**	.0932*	.1319**	.2088**	-.0018	.1600**	.1833**	.1420**	.1948**	.1503**	.2425**	.1008**	.2244**	1.0000						
X65	.1280**	.0597	.1056**	.1621**	-.0030	.0733	.1573**	.1106**	.0451	.0214	.1790**	.0820*	.1569**	.0784	1.0000					
X66	.1156**	-.0062	.0839*	.1845**	-.0861*	.1226**	.1911**	.1581**	.1401**	.0875*	.1649**	.1165**	.1842**	.2227**	.0424	1.0000				
X67	.1640**	.0522	.1070**	.1427**	-.0070	.1454**	.2663**	.1594**	.1368**	.1342**	.2290**	.1516**	.3464**	.2824**	.1561**	.2244**	1.0000			
X68	.0396	-.0180	.0547	.0008	-.0993*	.0056	.0550	.0095	.0221	.0637	-.0047	.0466	.0227	.0458	-.0059	.0488	.0333	1.0000		
X69	.1829**	.0663	.1444**	.1939**	.0463	.1620**	.2325**	.1339**	.2000**	.1477**	.2197**	.1173**	.2420**	.3246**	.1742**	.2051**	.2830**	.0319	1.0000	
X70	.0679	.0634	.0673	.0654	.0344	-.0053	.0743	.0475	.0253	-.0023	.0511	.0411	.0566	.0487	.0210	.0463	.0455	-.0675	.1016**	1.0000
X	5112	2705	3597	5994	5799	5028	4517	4638	8894	6236	6757	4006	5641	5985	3987	6441	5623	3578	4851	2407
S.D.	.50	.44	.48	.49	.49	.50	.50	.50	.31	.48	.47	.49	.50	.49	.49	.48	.50	.48	.50	.43

* p < 0.01, ** p < 0.001

ตารางที่ 19 เมทริกซ์สัมพันธ์ระหว่างตัวแปรใน SUBTEST4

ตัวแปร	X31	X32	X33	X34	X35	X36	X37	X38	X39	X40	X71	X72	X73	X74	X75	X76	X77	X78	X79	X80
X31	1.0000																			
X32	.3328**	1.0000																		
X33	.2780**	.2728**	1.0000																	
X34	.1643**	.1761**	.2356**	1.0000																
X35	.2798**	.2640**	.2419**	.1680**	1.0000															
X36	.0986*	.0162	.0731	.1080**	.1570**	1.0000														
X37	.1340**	.1462**	.1510**	.2211**	.1114**	.0604	1.0000													
X38	.2598**	.2339**	.2302**	.1950**	.2080**	.1368**	.1758**	1.0000												
X39	.3227**	.3052**	.2275**	.2594**	.2986**	.1468**	.2277**	.2901**	1.0000											
X40	.2054**	.2224**	.2180**	.2242**	.2164**	.1460**	.1839**	.2849**	.3585**	1.0000										
X71	.1631**	.2055**	.2068**	.1744**	.1339**	.0408	.1240**	.2182**	.2213**	.1862**	1.0000									
X72	.2283**	.2535**	.2360**	.1585**	.2123**	.0463	.1204**	.2062**	.2357**	.1759**	.2438**	1.0000								
X73	.1232**	.1557**	.1018**	.1538**	.0859*	.0756	.1069**	.1320**	.1399**	.1168**	.1399**	.1136**	1.0000							
X74	.2242**	.1990**	.1738**	.1458**	.1956**	.0826*	.1763**	.1906**	.2200**	.1939**	.1330**	.1698**	.0890*	1.0000						
X75	.1150**	.0941*	.0905*	.0917*	.1789**	.1219**	.0202	.0662	.1236**	.1268**	.0565	.1463**	.0566	.1588**	1.0000					
X76	.1907**	.1657**	.1580**	.2051**	.2058**	.1332**	.1386**	.1997**	.2499**	.2504**	.1639**	.1843**	.1607**	.3074**	.1560**	1.0000				
X77	.2648**	.1826**	.1497**	.1827**	.2254**	.0975*	.1439**	.1457**	.2423**	.1483**	.1537**	.1834**	.0995*	.1801**	.1348**	.2275**	1.0000			
X78	.2862**	.2498**	.1903**	.2010**	.2387**	.0973*	.1408**	.3005**	.3357**	.2318**	.2385**	.2581**	.1062**	.2186**	.1453**	.2563**	.3308**	1.0000		
X79	.2316**	.1931**	.2351**	.1726**	.2015**	.0774	.1339**	.02806**	.2749**	.2082**	.1972**	.2199**	.0937*	.1945**	.1887**	.2421**	.2755**	.3680**	1.0000	
X80	.2765**	.2374**	.2269**	.1870**	.1425**	.0736	.1733**	.2558**	.2509**	.1984**	.2648**	.2458**	.1760**	.2169**	.1217**	.2088**	.1913**	.2986**	.2653**	1.0000
X	.6628	.6896	.5781	.4814	.6301	.6552	.3718	.6403	.6794	.5130	.4944	.6636	.4043	.5976	.8123	.6227	.7166	.6849	.7165	.6180
S.D.	.47	.46	.49	.50	.48	.48	.48	.48	.47	.50	.50	.47	.49	.49	.39	.48	.45	.46	.45	.49

* p < 0.01, ** p < 0.001

ตอนที่ 4 การวิเคราะห์องค์ประกอบของแบบทดสอบ

ในการตรวจสอบความตรงเชิงโครงสร้าง (construct validity) และยืนยันความเป็นพหุมิติ (multidimensional) ของแบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่สร้างขึ้น ผู้วิจัยใช้เทคนิคการวิเคราะห์องค์ประกอบ (factor analysis) โดยในขั้นแรกเป็นการสำรวจหรือค้นหา (exploratory factor analysis = EFA) ว่าแบบทดสอบที่สร้างขึ้นประกอบด้วยตัวแปรใดบ้าง และมีตัวแปรใดบ้างที่อยู่บนองค์ประกอบเหล่านั้น องค์ประกอบที่ได้มีความสอดคล้องกับด้านต่าง ๆ ที่กำหนดไว้ในการสร้างแบบทดสอบหรือไม่ โดยใช้โปรแกรม SPSS for WINDOW version 7.5 ในขั้นที่สองเป็นการตรวจสอบหรือยืนยัน (confirmatory factor analysis = CFA) ว่าตัวแปรมีโครงสร้างตามนิยามทางทฤษฎีหรือไม่และมีความสอดคล้องกลมกลืนกับข้อมูลการวิจัยเพียงใด โดยใช้โปรแกรม LISREL 8.10

ตารางที่ 20 ผลการวิเคราะห์องค์ประกอบเชิงสำรวจของแบบทดสอบ

องค์ประกอบที่	ตัวแปร	ค่าไอเกน	ค่าความแปรปรวนรวม
1	X59, X67, X38, X13, X78, X79, X2, X61, X40, X15	13.67	17.09
2	X17, X20, X60, X18, X3, X39, X32, X31, X19, X35	1.97	2.47
3	X48, X49, X9, X14, X58	1.67	1.96
4	X29, X30, X80, X64	1.60	1.82
5	X45, X1, X47, X10	1.39	1.73
6	X57, X51, X12	1.37	1.71
7	X26, X72, X71, X33	1.33	1.67
8	X65, X52	1.30	1.63
9	X28, X56	1.28	1.60
10	X46, X43, X44, X42, X4	1.26	1.57
11	X75, X76, X77	1.21	1.52
12	X62, X27, X37	1.21	1.51
13	X16, X21	1.19	1.50

ตารางที่ 20 (ต่อ)

ตัวประกอบที่	ตัวแปร	ค่าไอเกน	ค่าความแปรปรวนร่วม
14	X5	1.16	1.46
15	X73	1.14	1.43
16	X68, X64, X55	1.11	1.38
17	X70	1.09	1.36
18	X50	1.09	1.35
19	X41, X23	1.07	1.33
20	X36, X8	1.06	1.32
21	X11, X66	1.05	1.31
22	X53, X34	1.04	1.30
23	X25	1.01	1.26
24	X22	1.00	1.25

จากตารางที่ 20 ผู้วิจัยใช้การวิเคราะห์องค์ประกอบ โดยการสกัดตัวประกอบด้วยวิธีการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis : PC) และทำการหมุนแกนแบบ Orthogonal ด้วยวิธี Varimax พิจารณาองค์ประกอบที่มีค่าไอเกน (eigen value) สูงกว่า 1 ขึ้นไป และเลือกตัวแปรที่มีน้ำหนักองค์ประกอบ (factor loading) ตั้งแต่ 0.3 ขึ้นไป ปรากฏว่า แบบสอบชุดนี้ ประกอบด้วยองค์ประกอบ จำนวน 24 องค์ประกอบ สามารถอธิบายความแปรปรวนร่วมได้ ร้อยละ 52.53 โดยตัวแปรที่มีค่าน้ำหนักองค์ประกอบต่ำกว่า 0.3 มีจำนวน 5 ตัว คือ X6, X7, X24, X69 และ X74 นอกจากนี้หากพิจารณาองค์ประกอบที่มีตัวแปรบรรยายตั้งแต่ 3 ตัวขึ้นไป รวมด้วยพบว่า แบบสอบชุดนี้จะประกอบด้วยองค์ประกอบ จำนวน 11 องค์ประกอบและมีตัวแปร จำนวนหนึ่งที่มีค่าน้ำหนักองค์ประกอบสูงใกล้เคียงกันหลายองค์ประกอบ ดังนั้นผู้วิจัยจึงได้นำผลการวิเคราะห์มาพิจารณาร่วมกับตารางโครงสร้างของแบบสอบที่สร้าง จะเห็นว่าองค์ประกอบหลัก ที่ได้จากผลการวิเคราะห์มีตัวแปรบรรยายส่วนใหญ่ที่สอดคล้องกับองค์ประกอบในการวัดของแบบสอบ ผู้วิจัยจึงวิเคราะห์องค์ประกอบอีกครั้งเพื่อยืนยันว่าตัวแปรต่าง ๆ ในแบบสอบชุดนี้ อยู่ในองค์ประกอบที่ได้กำหนดไว้

เนื่องจากแบบสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยสร้างขึ้น มีองค์ประกอบในการวัดทั้งหมด 4 ด้าน ได้แก่ การวัดความรู้ความเข้าใจทางคณิตศาสตร์(K1)

การวัดความสามารถด้านการคิดคำนวณ(K2) การวัดความสามารถด้านการพิจารณาผลลัพธ์ที่สมเหตุสมผล(K3) และการวัดความสามารถด้านการแก้ปัญหา(K4) ดังนั้นผลการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรมลิสเรลเพื่อตรวจสอบความตรงเชิงโครงสร้างของแบบสอบถามที่ผู้วิจัยสร้างขึ้นจำนวน 80 ข้อ จึงทำการแยกโมเดลย่อย ๆ ตามองค์ประกอบในการวัดออกเป็น 4 โมเดล ได้ผลการวิเคราะห์ซึ่งมีรายละเอียดดังต่อไปนี้

4.1 การวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 1(K1)

ตารางที่ 21 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 1(K1)

ตัวแปร	น้ำหนักองค์ประกอบ b (SE)	สัมประสิทธิ์การ พยากรณ์ (R^2)	สัมประสิทธิ์คะแนน องค์ประกอบ
X1	0.43**(0.03)	0.19	0.43
X2	0.50**(0.03)	0.25	0.50
X3	0.54**(0.03)	0.30	0.54
X4	0.46**(0.03)	0.21	0.46
X5	0.27**(0.03)	0.07	0.11
X6	0.46**(0.03)	0.21	0.13
X7	0.53**(0.03)	0.28	0.15
X8	0.32**(0.03)	0.10	0.08
X9	0.48**(0.03)	0.23	0.07
X10	0.43**(0.03)	0.18	0.12
X41	0.18**(0.03)	0.03	0.13
X42	0.31**(0.03)	0.10	0.06
X43	0.41**(0.03)	0.17	0.09
X44	0.37**(0.03)	0.14	0.07
X45	0.47**(0.03)	0.22	0.13
X46	0.34**(0.03)	0.11	0.08
X47	0.49**(0.03)	0.24	0.13
X48	0.40**(0.03)	0.16	0.05

ตารางที่ 21 (ต่อ)

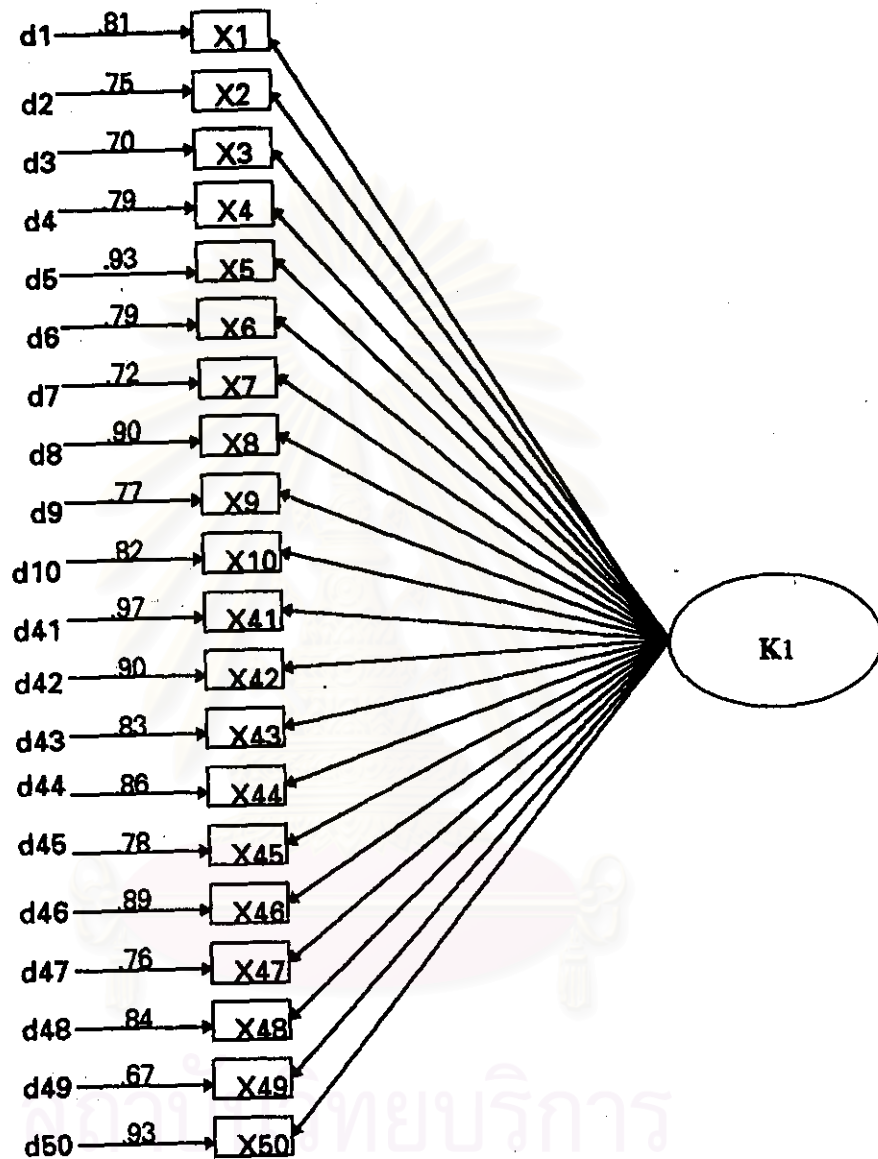
ตัวแปร	น้ำหนักองค์ประกอบ b (SE)	สัมประสิทธิ์การ พยากรณ์ (R ²)	สัมประสิทธิ์คะแนน องค์ประกอบ
X49	0.57**(.03)	0.33	0.14
X60	0.27**(.03)	0.07	0.06
Chi - square = 134.28		df = 145	P = 0.73
GFI = 0.99		AGFI = 0.98	

** p < 0.01

จากผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 1 พบว่า โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์พิจารณาได้จากค่าไคสแควร์ (Chi-square = 134.28) ซึ่งมีความน่าจะเป็นเข้าใกล้ 1 (P = 0.73) แสดงว่าค่าไคสแควร์แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ และค่าดัชนีวัดระดับความกลมกลืน (GFI = 0.99) รวมทั้งค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI = 0.98) มีค่าเข้าใกล้ 1 จึงสามารถยอมรับสมมติฐานหลักที่ว่า โมเดลการวิจัยสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์

เมื่อพิจารณาค่าน้ำหนักองค์ประกอบ พบว่า น้ำหนักองค์ประกอบของตัวแปรทั้ง 20 ตัว มีค่าเป็นบวก โดยมีขนาดตั้งแต่ 0.18 - 0.57 แม้ว่าจะมีตัวแปรบางตัวที่มีค่าน้ำหนักองค์ประกอบต่ำกว่า 0.30 แต่จะเห็นว่า มีนัยสำคัญที่ระดับ .01 ทุกค่าซึ่งแสดงให้เห็นว่าตัวแปรสังเกตได้เหล่านี้ต่างเป็นตัวแปรที่สำคัญขององค์ประกอบที่ 1 นอกจากนี้เมื่อพิจารณาจากค่าสัมประสิทธิ์การพยากรณ์ (R²) และค่าสัมประสิทธิ์คะแนนองค์ประกอบ (factor score coefficient) พบว่าให้ผลในการทำงานเองเดียวกันกับการพิจารณาน้ำหนักองค์ประกอบ จากการวิเคราะห์จะเห็นว่าตัวแปรแต่ละตัวมีเทอมความคลาดเคลื่อนซึ่งเกิดจากความสัมพันธ์ระหว่างตัวแปรนั้นกับตัวแปรอื่นในโมเดล ดังแสดงในแผนภาพที่ 6

แผนภาพที่ 6 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 1 (K1)



4.2 การวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 2 (K2)

จากผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 2 พบว่า โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ พิจารณาได้จากค่าไคสแควร์ (Chi-square = 116.52) ซึ่งมีความน่าจะเป็นเข้าใกล้ 1 ($P = 0.76$) แสดงว่าค่าไคสแควร์แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ และค่าดัชนีวัดระดับความกลมกลืน (GFI = 0.99) รวมทั้งค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI = 0.98) มีค่าเข้าใกล้ 1 ดังแสดงในตารางที่ 22 จึงสามารถยอมรับสมมติฐานหลักที่ว่าโมเดลการวิจัยสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์

ตารางที่ 22 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 2 (SUBTEST 2)

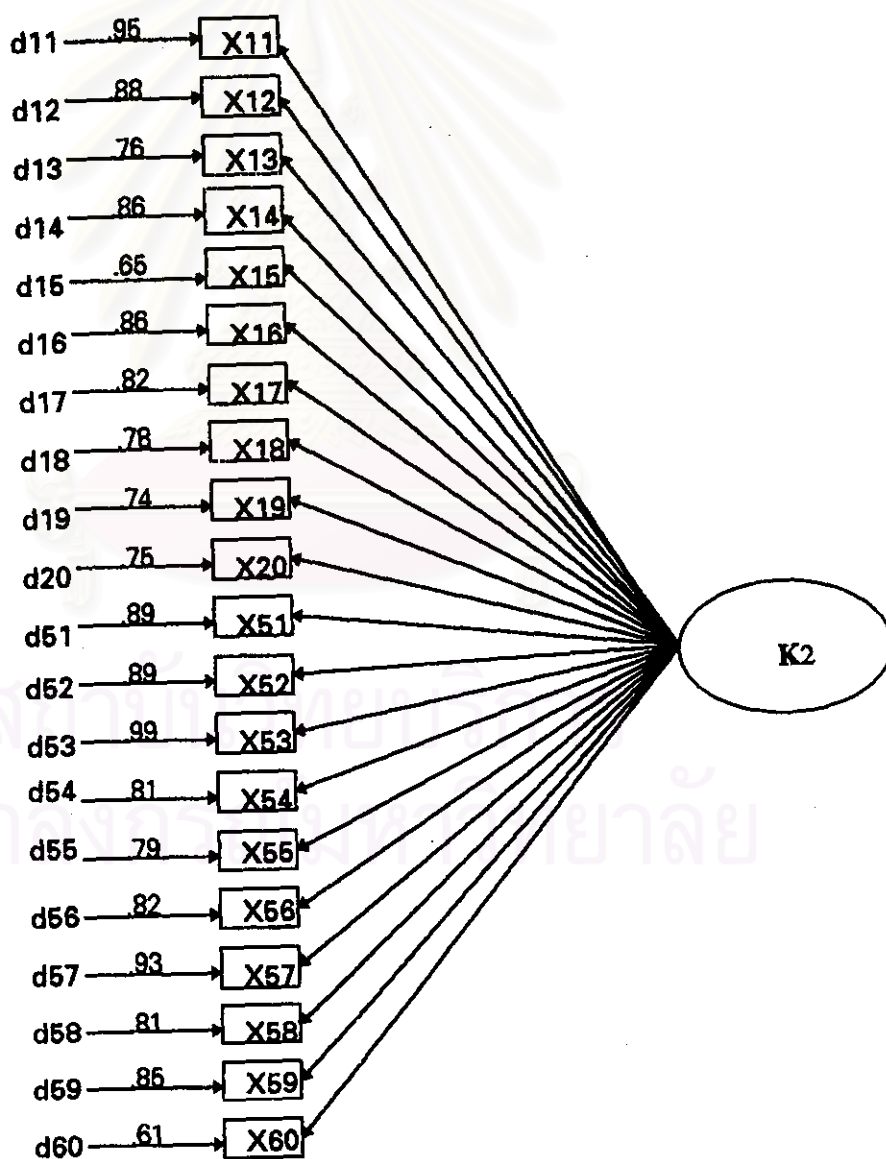
ตัวแปร	น้ำหนักองค์ประกอบ b (SE)	สัมประสิทธิ์การ พยากรณ์ (R ²)	สัมประสิทธิ์คะแนน องค์ประกอบ
X11	0.23**(0.03)	0.05	0.04
X12	0.35**(0.03)	0.12	0.08
X13	0.49**(0.03)	0.24	0.12
X14	0.37**(0.03)	0.14	0.05
X15	0.59**(0.03)	0.36	0.20
X16	0.36**(0.03)	0.13	0.06
X17	0.42**(0.03)	0.18	0.07
X18	0.46**(0.03)	0.22	0.09
X19	0.51**(0.03)	0.26	0.12
X20	0.50**(0.03)	0.25	0.14
X51	0.34**(0.03)	0.11	0.07
X52	0.33**(0.03)	0.11	0.06
X53	0.09*(0.03)	0.01	0.01
X54	0.44**(0.03)	0.20	0.07
X55	0.46**(0.03)	0.21	0.08
X56	0.43**(0.03)	0.18	0.08
X57	0.27**(0.03)	0.07	0.03
X58	0.43**(0.03)	0.19	0.06
X59	0.39**(0.03)	0.15	0.10
X60	0.62**(0.03)	0.39	0.23
Chi - square = 116.52 df = 128 P = 0.76			
GFI = 0.99 AGFI = 0.98			

* p < 0.05, ** p < 0.01

เมื่อพิจารณาน้ำหนักองค์ประกอบ พบว่า น้ำหนักองค์ประกอบของตัวแปรทั้ง 20 ตัวมีค่าเป็นบวก โดยมีขนาดตั้งแต่ 0.09 - 0.62 ถึงแม้ว่าจะมีตัวแปรบางตัวที่มีค่าน้ำหนักองค์ประกอบ

ต่ำกว่า 0.30 แต่จะเห็นว่ามีความสำคัญที่ระดับ .01 ทุกค่า ยกเว้นตัวแปร X60 ที่มีนัยสำคัญที่ระดับ .05 แสดงให้เห็นว่าตัวแปรสังเกตได้เหล่านี้ต่างเป็นตัวแปรที่สำคัญขององค์ประกอบที่ 2 นอกจากนี้เมื่อพิจารณาค่าสัมประสิทธิ์การพยากรณ์ (R^2) และค่าสัมประสิทธิ์คะแนนองค์ประกอบ (factor score coefficient) ให้ผลการวิเคราะห์ในทำนองเดียวกันกับการพิจารณาน้ำหนักองค์ประกอบ จากผลการวิเคราะห์จะเห็นว่าตัวแปรแต่ละตัวมีเทอมความคลาดเคลื่อนอยู่ด้วย ซึ่งมีรายละเอียดดังแผนภาพที่ 7

แผนภาพที่ 7 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 2 (SUBTEST 2)



4.3 การวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 3 (K3)

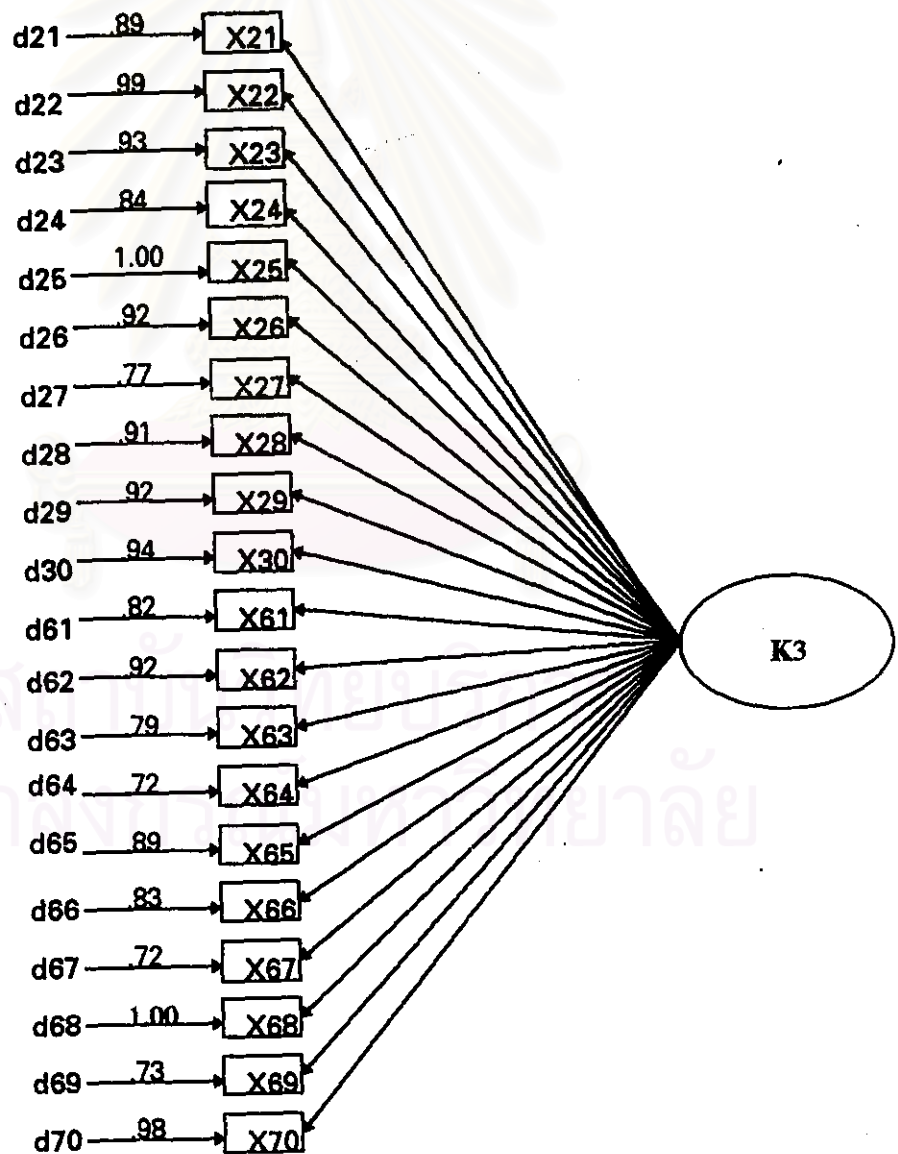
ตารางที่ 23 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 3 (SUBTEST 3)

ตัวแปร	น้ำหนักองค์ประกอบ	สัมประสิทธิ์การ	สัมประสิทธิ์คะแนน
	b (SE)	พยากรณ์ (R ²)	องค์ประกอบ
X21	0.33**(.03)	0.11	0.08
X22	0.10*(.04)	0.01	0.02
X23	0.27**(.03)	0.07	0.06
X24	0.39**(.03)	0.16	0.13
X25	0.04 (.04)	0.00	0.02
X26	0.28**(.03)	0.08	0.08
X27	0.48**(.03)	0.23	0.17
X28	0.30**(.03)	0.09	0.09
X29	0.29**(.03)	0.08	0.07
X30	0.24**(.03)	0.06	0.06
X61	0.42**(.03)	0.18	0.11
X62	0.29**(.03)	0.08	0.08
X63	0.46**(.03)	0.21	0.10
X64	0.52**(.03)	0.28	0.20
X65	0.33**(.04)	0.11	0.13
X66	0.41**(.03)	0.17	0.14
X67	0.52**(.03)	0.28	0.17
X68	0.07 (.04)	0.00	0.02
X69	0.52**(.03)	0.27	0.16
X70	0.13**(.03)	0.02	0.03
Chi - square = 140.46		df = 150	P = 0.70
GFI = 0.99		AGFI = 0.98	

** p < 0.01

จากผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 3 พบว่า โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ที่พิจารณาได้จากค่าไคสแควร์ (Chi-square = 140.46) ซึ่งมีความน่าจะเป็นเข้าใกล้ 1 ($P = 0.70$) แสดงว่าค่าไคสแควร์แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ และค่าดัชนีวัดระดับความกลมกลืน ($GFI = 0.99$) รวมทั้งค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว ($AGFI = 0.98$) มีค่าเข้าใกล้ 1 ดังแสดงในตารางที่ 23 จึงสามารถยอมรับสมมติฐานหลักที่ว่าโมเดลการวิจัยสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์

แผนภาพที่ 8 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 3 (SUBTEST 3)



จากแผนภาพที่ 8 เมื่อพิจารณาน้ำหนักองค์ประกอบ พบว่า น้ำหนักองค์ประกอบของตัวแปรทั้ง 20 ตัวมีค่าเป็นบวก มีขนาดตั้งแต่ 0.04 - 0.60 แม้ว่าจะมีตัวแปรบางตัวที่มีค่าน้ำหนักองค์ประกอบต่ำกว่า 0.30 แต่จะเห็นว่ามียุทธศาสตร์ที่ระดับ .01 ทุกค่ายกเว้นตัวแปร X25 และ X68 เท่านั้นที่ไม่มีนัยสำคัญทางสถิติ ซึ่งแสดงให้เห็นว่าตัวแปรสังเกตได้ส่วนใหญ่ต่างเป็นตัวแปรที่สำคัญขององค์ประกอบที่ 3 นอกจากนี้เมื่อพิจารณาจากค่าสัมประสิทธิ์การพยากรณ์ (R^2) และค่าสัมประสิทธิ์คะแนนองค์ประกอบ (factor score coefficient) ให้ผลในทำนองเดียวกันกับการพิจารณาน้ำหนักองค์ประกอบ จากผลการวิเคราะห์จะเห็นว่าตัวแปรแต่ละตัวมีเทอมความคลาดเคลื่อนอยู่ด้วย

4.4 การวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 4(K4)

ตารางที่ 24 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 4 (SUBTEST 4)

ตัวแปร	น้ำหนักองค์ประกอบ b (SE)	สัมประสิทธิ์การ พยากรณ์ (R^2)	สัมประสิทธิ์คะแนน องค์ประกอบ
X31	0.52**(.03)	0.27	0.10
X32	0.49**(.03)	0.24	0.10
X33	0.50**(.03)	0.26	0.16
X34	0.40**(.03)	0.16	0.08
X35	0.47**(.03)	0.22	0.11
X36	0.20**(.03)	0.04	0.04
X37	0.31**(.03)	0.10	0.06
X38	0.49**(.03)	0.24	0.11
X39	0.60**(.03)	0.36	0.18
X40	0.45**(.03)	0.20	0.06
X71	0.39**(.03)	0.15	0.06
X72	0.44**(.03)	0.20	0.09
X73	0.26**(.03)	0.07	0.04
X74	0.39**(.03)	0.16	0.06
X75	0.24**(.03)	0.06	0.03

ตารางที่ 24 (ต่อ)

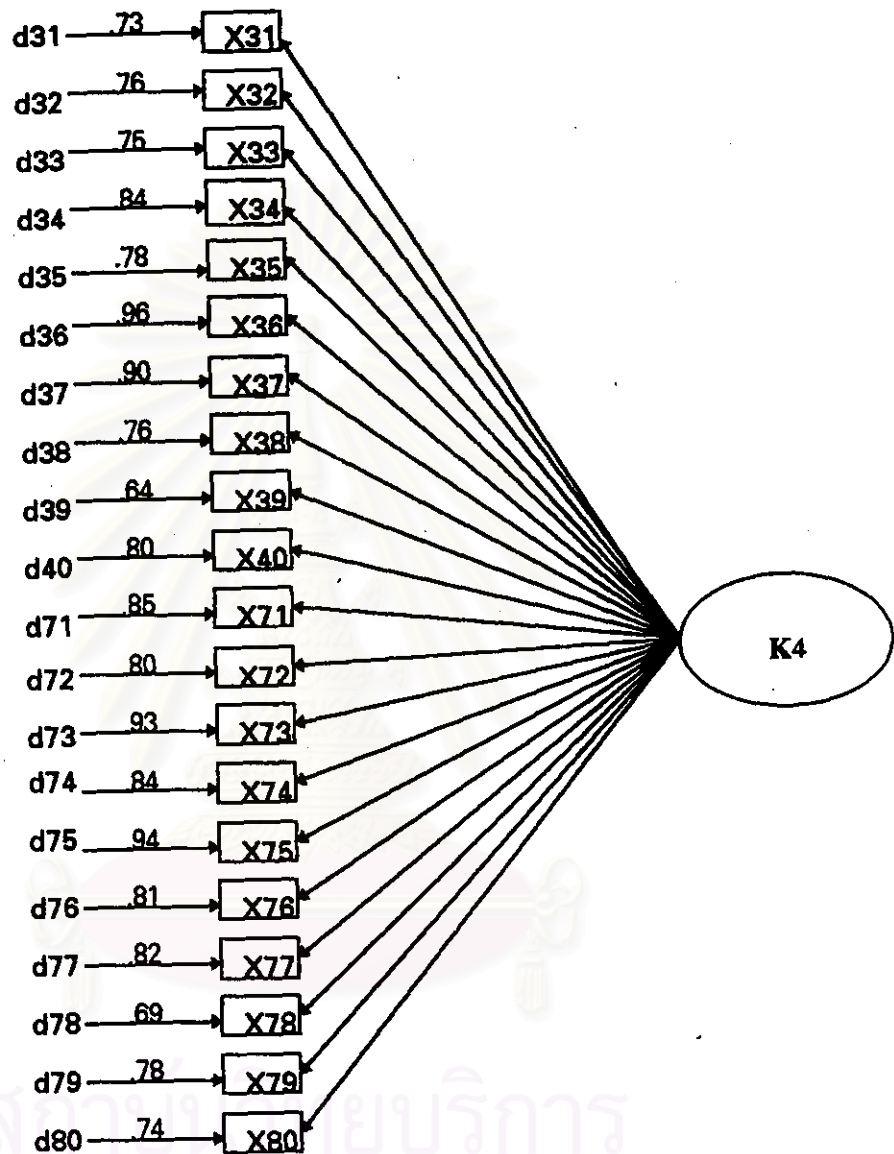
ตัวแปร	น้ำหนักองค์ประกอบ b (SE)	สัมประสิทธิ์การ พยากรณ์ (R ²)	สัมประสิทธิ์คะแนน องค์ประกอบ
X76	0.43**(0.03)	0.19	0.08
X77	0.43**(0.03)	0.18	0.09
X78	0.55**(0.03)	0.31	0.13
X79	0.47**(0.03)	0.22	0.07
X80	0.50**(0.03)	0.25	0.14
Chi - square = 130.00		df = 139	P = 0.70
GFI = 0.99		AGFI = 0.98	

** p < 0.01

จากผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 4 พบว่า โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ที่พิจารณาได้จากค่าไคสแควร์ (Chi-square = 130.00) ซึ่งมีความน่าจะเป็นเข้าใกล้ 1 (P = 0.70) แสดงว่า ค่าไคสแควร์แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ และค่าดัชนีวัดระดับความกลมกลืน (GFI = 0.99) รวมทั้งค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI = 0.98) มีค่าเข้าใกล้ 1 ดังแสดงในตารางที่ 24 จึงสามารถยอมรับสมมติฐานหลักที่ว่าโมเดลการวิจัยสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์

เมื่อพิจารณาน้ำหนักองค์ประกอบ พบว่าน้ำหนักองค์ประกอบของตัวแปรทั้ง 20 ตัว มีค่าเป็นบวก มีขนาดตั้งแต่ 0.20 - 0.60 แม้ว่าจะมีตัวแปรบางตัวที่มีค่าน้ำหนักองค์ประกอบต่ำกว่า 0.30 แต่จะเห็นว่ามีความสำคัญที่ระดับ .01 ทุกค่า ซึ่งแสดงให้เห็นว่าตัวแปรสังเกตได้เหล่านี้ต่างเป็นตัวแปรที่สำคัญขององค์ประกอบที่ 4 นอกจากนี้เมื่อพิจารณาค่าสัมประสิทธิ์การพยากรณ์ (R²) และค่าสัมประสิทธิ์คะแนนองค์ประกอบ (factor score coefficient) พบว่าให้ผลในการทำงานของเดียวกันกับการพิจารณาน้ำหนักองค์ประกอบ จากผลการวิเคราะห์จะเห็นว่าตัวแปรแต่ละตัวมีเทอมความคลาดเคลื่อนอยู่ด้วย ซึ่งมีรายละเอียดดังแผนภาพที่ 9

แผนภาพที่ 9 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบที่ 4 (SUBTEST 4)



กล่าวโดยสรุป ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของตัวแปรสังเกตได้จากโมเดลองค์ประกอบของแบบสอบถามที่ผู้วิจัยสร้างขึ้น มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ทั้ง 4 โมเดลย่อย โดยที่น้ำหนักองค์ประกอบของตัวแปรมีนัยสำคัญทางสถิติทุกค่า ยกเว้นตัวแปร X25 และ X68 ซึ่งแสดงว่า ตัวแปรสังเกตได้เหล่านี้เป็นตัวแปรที่สำคัญขององค์ประกอบที่ต้องการวัดทั้งสิ้น นอกจากนี้ยังพบเทอมความคลาดเคลื่อนของแต่ละตัวแปรที่เกิดจากความสัมพันธ์ของตัวแปรนั้นกับตัวแปรอื่นในโมเดลซึ่งการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรมลิสเรลได้นำเทอมความคลาดเคลื่อนเข้ามาวิเคราะห์ด้วย โดยยอมให้มีความสัมพันธ์กันได้ทำให้ผลจากการ

วิเคราะห์องค์ประกอบสอดคล้องกับสภาพความเป็นจริงมากยิ่งขึ้น และจากผลการวิเคราะห์จะได้ผลการตอบแบบสอบถามชนิดพหุมิติที่มีองค์ประกอบในการวัดทั้งหมด 4 ด้าน มีตัวแปรทั้งสิ้น 75 ตัว มีรายละเอียด ดังนี้

องค์ประกอบที่ 1 การวัดความรู้ความเข้าใจทางคณิตศาสตร์(K1) มีตัวแปรสังเกตได้ที่สำคัญขององค์ประกอบ ได้แก่ X1, X2, X3, X4, X5, X6, X7, X8, X9, X10, X41, X42, X43, X44, X45, X46, X47, X48, X49 และ X50 รวมตัวแปรทั้งหมด 20 ตัว

องค์ประกอบที่ 2 การวัดความสามารถด้านการคิดคำนวณ(K2) มีตัวแปรสังเกตได้ที่สำคัญขององค์ประกอบ ได้แก่ X11, X12, X13, X14, X15, X16, X17, X18, X19, X20, X52, X54, X55, X56, X57, X58, X59 และ X60 ทั้งนี้ได้ตัดตัวแปร X53 ออกไปเนื่องจากมีค่าน้ำหนักองค์ประกอบต่ำมาก ($b = 0.09$) รวมตัวแปรทั้งหมด 19 ตัว

องค์ประกอบที่ 3 การวัดความสามารถด้านการพิจารณาผลลัพธ์ที่สมเหตุสมผล (K3) มีตัวแปรสังเกตได้ที่สำคัญขององค์ประกอบ ได้แก่ X21, X23, X24, X26, X27, X28, X29, X30, X61, X62, X63, X64, X65, X66, X67 และ X69 ทั้งนี้ได้ตัดตัวแปร X25 และ X68 ออกไปเนื่องจากไม่มีนัยสำคัญทางสถิติ รวมทั้งตัดตัวแปร X22 และ X70 เนื่องจากมีค่าน้ำหนักองค์ประกอบต่ำมาก ($b = 0.10, 0.13$ ตามลำดับ) รวมตัวแปรทั้งหมด 16 ตัว

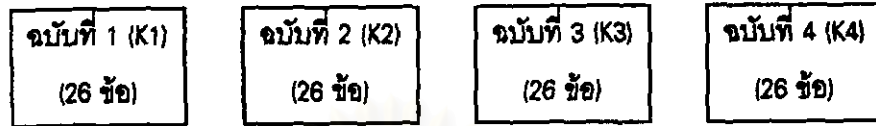
องค์ประกอบที่ 4 การวัดความสามารถด้านการแก้โจทย์ปัญหา(K4) มีตัวแปรสังเกตได้ที่สำคัญขององค์ประกอบ ได้แก่ X31, X32, X33, X34, X35, X36, X37, X38, X39, X40, X71, X72, X73, X74, X75, X76, X77, X78, X79 และ X80 รวมตัวแปรทั้งหมด 20 ตัว

ดังนั้นลักษณะของแบบทดสอบที่ใช้ในการวิจัยและผลการตอบข้อสอบที่จะนำมาใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบจึงมีลำดับชั้นต่างๆ พอที่จะสรุปได้ดังแผนภาพที่ 10

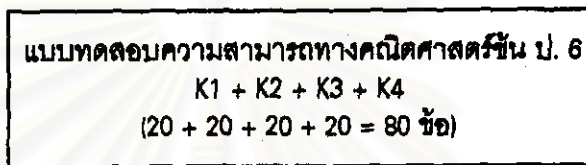
จุฬาลงกรณ์มหาวิทยาลัย

แผนภาพที่ 10 สรุปลักษณะของแบบทดสอบและผลการตอบข้อสอบที่ใช้ในการวิเคราะห์

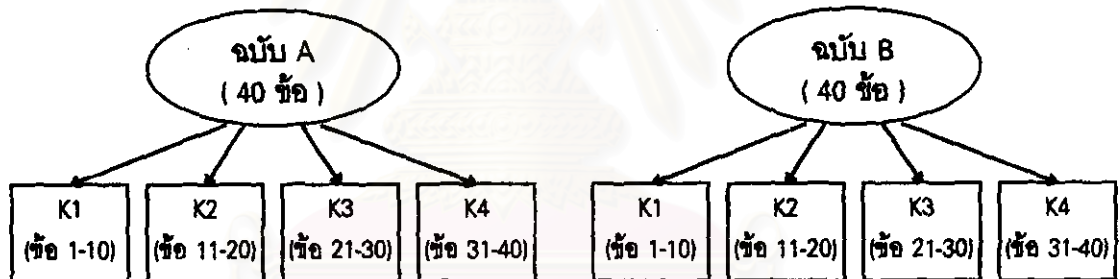
1. แบบสอบย่อยที่ใช้ Try-out



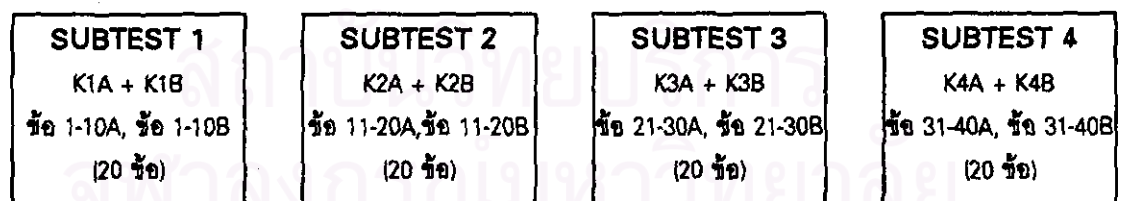
2. ข้อสอบที่ผ่านการวิเคราะห์คุณภาพข้อสอบและ Factor analysis



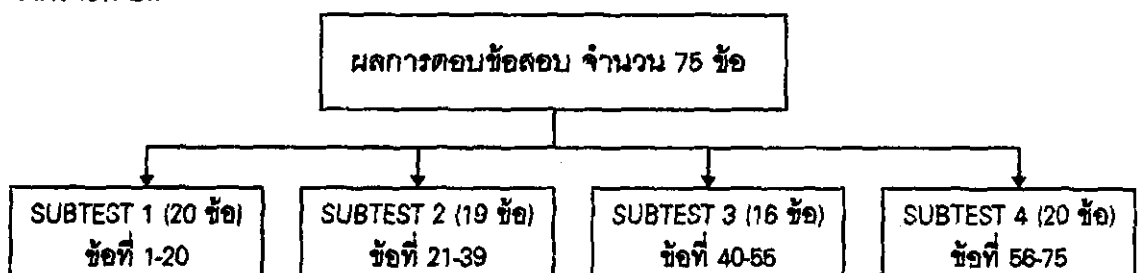
3. ลักษณะของแบบทดสอบที่นำไปใช้เก็บข้อมูลจริง



4. แยกผลการตอบข้อสอบเพื่อนำไปใช้ในการวิเคราะห์ DIF



5. ลักษณะของผลการตอบข้อสอบที่ผ่าน Item analysis และ Factor analysis พร้อมทั้งจะนำไปวิเคราะห์ DIF



ตอนที่ 5 เปรียบเทียบการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีแมนเทล-แฮนส์เซลและวิธีถดถอยโลจิสติกตามเกณฑ์จับคู่เปรียบเทียบ

5.1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีซิปเทสท์

ตารางที่ 25 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีซิปเทสท์

ข้อ	วิธี SIBTEST		ข้อ	วิธี SIBTEST	
	Beta -uni	p-value		Beta -uni	p-value
1.	-0.008	0.751	39.	-0.039	0.227
2.	0.018	0.522	40.	-0.018	0.410
3.	-0.025	0.343	41.	-0.003	0.907
4.	0.033	0.233	42.	0.001	0.975
5.	0.097	0.000*	43.	0.027	0.224
6.	0.035	0.213	44.	0.069	0.020*
7.	0.029	0.294	45.	0.018	0.513
8.	-0.048	0.049*	46.	0.022	0.446
9.	0.058	0.028*	47.	0.036	0.152
10.	0.086	0.002*	48.	-0.002	0.936
11.	-0.003	0.885	49.	-0.033	0.234
12.	0.021	0.459	50.	0.016	0.554
13.	-0.010	0.710	51.	-0.019	0.448
14.	0.048	0.076	52.	-0.018	0.519
15.	0.016	0.537	53.	-0.001	0.975
16.	0.071	0.016*	54.	-0.041	0.044*
17.	-0.006	0.800	55.	-0.007	0.813
18.	0.022	0.396	56.	-0.012	0.682
19.	-0.007	0.791	57.	0.011	0.662
20.	-0.017	0.500	58.	-0.022	0.431
21.	0.000	0.998	59.	-0.002	0.952
22.	-0.068	0.023*	60.	-0.043	0.125

ตารางที่ 25 (ต่อ)

ข้อ	วิธี SIBTEST		ข้อ	วิธี SIBTEST	
	Beta -uni	p-value		Beta -uni	p-value
23.	0.017	0.567	61.	0.019	0.496
24.	-0.102	0.001*	62.	0.077	0.009*
25.	0.023	0.433	63.	-0.016	0.587
26.	0.003	0.924	64.	-0.018	0.513
27.	-0.046	0.011*	65.	-0.006	0.814
28.	0.059	0.050	66.	-0.106	0.000*
29.	-0.056	0.036*	67.	-0.104	0.000*
30.	-0.007	0.795	68.	-0.015	0.612
31.	-0.030	0.303	69.	-0.012	0.706
32.	0.028	0.336	70.	-0.041	0.090
33.	0.012	0.679	71.	0.017	0.543
34.	-0.047	0.130	72.	-0.014	0.597
35.	0.104	0.000*	73.	-0.022	0.386
36.	0.011	0.693	74.	-0.034	0.188
37.	0.001	0.968	75.	-0.058	0.036*
38.	0.051	0.088			

* $p < .05$

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีซิบเทสท์(SIBTEST) พิจารณาจากค่า p-value หากพบว่าข้อสอบข้อใดมีนัยสำคัญ(Significant) หมายความว่า ข้อสอบข้อนั้นทำหน้าที่ต่างกัน จากตารางที่ 25 พบว่า ข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันมีทั้งหมด 16 ข้อ ได้แก่ ข้อที่ 6, 8, 9, 10, 16, 22, 24, 27, 29, 35, 44, 54, 62, 66, 67 และ 75 ซึ่งข้อสอบที่มีค่าดัชนีบ่งชี้การทำหน้าที่ต่างกันของข้อสอบ (Beta-uni) สูงสุด คือ ข้อที่ 66 (Beta-uni = -0.106) รองลงมา ได้แก่ ข้อที่ 35, 67 (Beta-uni = 0.104, -0.104) และ 24 (Beta-uni = -0.102) ตามลำดับ ส่วนข้อสอบที่มีค่าดัชนีบ่งชี้การทำหน้าที่ต่างกันของข้อสอบต่ำสุด คือ ข้อที่ 54 (Beta-uni = -0.041)

5.2 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล
ตามเกณฑ์จับคู่เปรียบเทียบ

ตารางที่ 26 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซลเมื่อใช้
คะแนนรวม(total test score) เป็นเกณฑ์จับคู่เปรียบเทียบ

ข้อ	วิธี MH		ข้อ	วิธี MH	
	MH alpha	p-value		MH alpha	p-value
1.	0.93	0.716	39.	0.84	0.208
2.	1.14	0.434	40.	0.99	0.967
3.	0.96	0.870	41.	1.01	0.981
4.	1.22	0.208	42.	0.99	0.963
5.	1.63	0.001*	43.	1.24	0.311
6.	1.17	0.335	44.	1.43	0.015*
7.	1.22	0.234	45.	1.12	0.500
8.	0.77	0.129	46.	1.03	0.871
9.	1.39	0.048*	47.	1.26	0.189
10.	1.65	0.002*	48.	0.96	0.881
11.	1.05	0.872	49.	0.83	0.251
12.	1.15	0.376	50.	1.10	0.579
13.	0.93	0.682	51.	0.95	0.826
14.	1.34	0.080	52.	0.83	0.245
15.	1.11	0.558	53.	0.96	0.813
16.	1.38	0.032*	54.	0.64	0.028*
17.	0.96	0.895	55.	0.96	0.822
18.	1.17	0.373	56.	0.89	0.473
19.	0.99	0.989	57.	1.06	0.785
20.	0.93	0.731	58.	0.87	0.375
21.	1.03	0.897	59.	0.96	0.840
22.	0.76	0.050	60.	0.80	0.151

ตารางที่ 26 (ต่อ)

ข้อ	วิธี MH		ข้อ	วิธี MH	
	MH alpha	p-value		MH alpha	p-value
23.	1.06	0.760	61.	1.09	0.603
24.	0.67	0.003*	62.	1.38	0.022*
25.	1.13	0.432	63.	0.92	0.617
26.	1.08	0.599	64.	0.93	0.702
27.	0.63	0.043*	65.	0.93	0.678
28.	0.72	0.048*	66.	0.58	0.000*
29.	0.73	0.049*	67.	0.59	0.001*
30.	0.97	0.930	68.	0.94	0.726
31.	0.81	0.154	69.	0.92	0.593
32.	1.14	0.384	70.	0.70	0.047*
33.	0.99	0.995	71.	1.14	0.426
34.	0.85	0.251	72.	1.00	0.939
35.	1.69	0.001*	73.	0.89	0.542
36.	0.99	0.977	74.	0.83	0.305
37.	1.07	0.734	75.	0.78	0.119
38.	1.21	0.201			

* $p < .05$

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทล - แฮนส์เซล(MH) เมื่อใช้คะแนนรวม (total test score) เป็นเกณฑ์จับคู่เปรียบเทียบ ซึ่งหากพิจารณาจากค่า p-value แล้วพบว่าข้อสอบข้อใดมีนัยสำคัญ(Significant) หมายความว่า ข้อสอบข้อนั้นทำหน้าที่ต่างกัน จากตารางที่ 26 พบว่า ข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันมีทั้งหมด 15 ข้อ ได้แก่ ข้อที่ 5, 9, 10, 16, 24, 27, 28, 29, 35, 44, 54, 62, 66, 67 และ 70 ข้อสอบที่มีค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องระหว่างกลุ่มสูงสุด คือ ข้อที่ 35 ($\alpha_{MH} = 1.69$) รองลงมา ได้แก่ ข้อที่ 10 ($\alpha_{MH} = 1.65$) และข้อที่ 24($\alpha_{MH} = 1.39$) ตามลำดับ ซึ่งเป็นข้อสอบที่เข้าข้างกลุ่มอ้างอิง($\alpha_{MH} > 1$) ส่วนข้อสอบที่มีค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องระหว่างกลุ่มต่ำสุด คือ ข้อที่ 66 ($\alpha_{MH} = 0.58$)

ตารางที่ 27 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เชล เมื่อใช้คะแนนแบบสอบย่อย (subtest scores) เป็นเกณฑ์จับคู่เปรียบเทียบ

แบบสอบย่อย	ข้อ	วิธี MH		ข้อ	วิธี MH	
		MH alpha	p-value		MH alpha	p-value
SUBTEST 1	1.	0.80	0.234	39.	0.83	0.171
	2.	0.93	0.651	40.	0.85	0.406
	3.	0.76	0.094	41.	0.96	0.854
	4.	1.04	0.882	42.	0.85	0.401
	5.	1.58	0.003*	43.	1.11	0.672
	6.	1.04	0.821	44.	1.33	0.049*
	7.	0.97	0.917	45.	1.09	0.622
	8.	0.66	0.014*	46.	0.94	0.694
	9.	1.25	0.186	47.	1.08	0.697
	10.	1.36	0.047*	48.	0.88	0.511
SUBTEST 2	11.	0.99	0.957	49.	0.83	0.240
	12.	1.04	0.838	50.	1.04	0.853
	13.	1.00	0.963	51.	0.98	0.966
	14.	1.26	0.152	52.	0.84	0.297
	15.	1.06	0.830	53.	0.95	0.793
	16.	1.34	0.048*	54.	0.66	0.046*
	17.	0.96	0.869	55.	0.93	0.665
	18.	1.22	0.235	56.	0.96	0.817
	19.	1.00	0.941	57.	1.12	0.527
	20.	0.88	0.489			
SUBTEST 3	21.	1.12	0.460	58.	0.89	0.480
	22.	0.82	0.186	59.	0.94	0.712
	23.	1.10	0.553	60.	0.96	0.353

ตารางที่ 27 (ต่อ)

แบบทดสอบย่อย	ข้อ	วิธี MH		ข้อ	วิธี MH	
		MH alpha	p-value		MH alpha	p-value
(SUBTEST 3)	24.	0.67	0.003*	61.	1.17	0.323
	25.	1.19	0.256	62.	1.35	0.036*
	26.	1.04	0.816	63.	0.97	0.871
	27.	0.64	0.049*	64.	1.02	0.947
	28.	0.70	0.047*	65.	1.04	0.855
SUBTEST 4	29.	0.82	0.222	66.	0.65	0.003*
	30.	1.09	0.644	67.	0.63	0.003*
	31.	0.96	0.831	68.	1.03	0.901
	32.	1.26	0.112	69.	0.98	0.929
	33.	1.02	0.971	70.	0.76	0.133
	34.	0.90	0.477	71.	1.33	0.048*
	35.	1.75	0.000*	72.	1.05	0.815
	36.	1.07	0.733	73.	1.00	0.946
	37.	1.15	0.447	74.	0.98	0.953
	38.	1.25	0.151	75.	0.85	0.335

* p < .05

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล - ฮานส์เชล(MH) เมื่อใช้คะแนนแบบทดสอบย่อย (subtest scores) เป็นเกณฑ์จับคู่เปรียบเทียบ ซึ่งพิจารณาค่า p-value จากตารางที่ 27 พบว่า ข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันมีทั้งหมด 14 ข้อ ได้แก่ ข้อที่ 5, 8, 10, 16, 24, 27, 28, 35, 44, 54, 62, 66, 67 และ 71 ข้อสอบที่มีค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องระหว่างกลุ่มสูงสุด คือ ข้อที่ 36 ($\alpha_{MH} = 1.75$) รองลงมา ได้แก่ ข้อที่ 5 ($\alpha_{MH} = 1.59$) และข้อที่ 10 ($\alpha_{MH} = 1.36$) ตามลำดับ ซึ่งเป็นข้อสอบที่เข้าข้างกลุ่มอ้างอิง ($\alpha_{MH} > 1$) ส่วนข้อสอบที่มีค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องระหว่างกลุ่มต่ำสุด คือ ข้อที่ 67 ($\alpha_{MH} = 0.63$)

5.3 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีทดสอบไจจิสติก
ตามเกณฑ์จับคู่เปรียบเทียบ

ตารางที่ 28 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีทดสอบไจจิสติกเมื่อใช้
คะแนนรวม(total test score) เป็นเกณฑ์จับคู่เปรียบเทียบ

ข้อ	ปฏิสัมพันธ์ระหว่าง				ข้อ	ปฏิสัมพันธ์ระหว่าง			
	กลุ่มผู้สอบ(ง)		กลุ่มกับความ			กลุ่มผู้สอบ(ง)		กลุ่มกับความ	
	B	sig	B	sig		B	sig	B	sig
1.	0.110	0.504	0.023	0.096	39.	0.148	0.242	0.015	0.049*
2.	-0.400	0.779	0.001	0.915	40.	0.058	0.737	-0.008	0.533
3.	0.131	0.380	0.033	0.011*	41.	-0.002	0.989	-0.003	0.813
4.	-0.165	0.263	0.009	0.467	42.	0.022	0.894	-0.014	0.288
5.	-0.462	0.002*	-0.006	0.580	43.	-0.251	0.174	-0.013	0.425
6.	-0.153	0.271	0.014	0.218	44.	-0.370	0.006*	0.001	0.892
7.	-0.0128	0.384	0.017	0.176	45.	-0.156	0.277	0.018	0.125
8.	0.303	0.049*	-0.001	0.959	46.	-0.032	0.815	-0.009	0.406
9.	-0.360	0.019*	0.008	0.522	47.	-0.096	0.066	-0.012	0.423
10.	-0.413	0.005*	0.023	0.051	48.	0.031	0.857	-0.015	0.248
11.	-0.022	0.900	-0.020	0.123	49.	0.202	0.159	-0.010	0.379
12.	-0.031	0.826	-0.001	0.982	50.	-0.086	0.549	-0.010	0.379
13.	0.057	0.097	0.015	0.209	51.	0.056	0.724	-0.026	0.047*
14.	-0.238	0.102	0.015	0.198	52.	0.117	0.430	-0.009	0.451
15.	-0.055	0.724	-0.001	0.919	53.	0.044	0.750	0.005	0.663
16.	-0.284	0.038*	0.005	0.646	54.	0.437	0.023*	-0.018	0.224
17.	0.013	0.938	0.021	0.119	55.	0.065	0.648	-0.012	0.269
18.	-0.136	0.382	-0.018	0.163	56.	0.044	0.757	0.012	0.266
19.	0.040	0.787	0.007	0.556	57.	-0.092	0.556	0.019	0.179
20.	0.122	0.429	0.013	0.314	58.	0.160	0.283	-0.003	0.782

ตารางที่ 28 (ต่อ)

ข้อ	ปฏิสัมพันธ์ระหว่าง				ข้อ	ปฏิสัมพันธ์ระหว่าง			
	กลุ่มผู้สอบ(g)		กลุ่มกับความ			กลุ่มผู้สอบ(g)		กลุ่มกับความ	
	B	sig	B	sig		B	sig	B	sig
21.	-0.081	0.535	0.001	0.978	59.	0.70	0.692	0.011	0.249
22.	0.212	0.111	0.015	0.125	60.	0.197	0.168	0.014	0.229
23.	-0.049	0.719	-0.003	0.747	61.	-0.081	0.575	0.011	0.336
24.	0.393	0.002*	-0.010	0.263	62.	-0.297	0.025*	-0.002	0.860
25.	-0.111	0.422	0.015	0.165	63.	0.058	0.675	-0.007	0.510
26.	-0.026	0.839	0.004	0.641	64.	0.052	0.720	-0.014	0.238
27.	0.451	0.041*	0.005	0.813	65.	0.054	0.712	0.006	0.610
28.	-0.236	0.055	0.007	0.446	66.	0.511	0.000*	0.007	0.531
29.	0.301	0.047*	0.019	0.148	67.	0.567	0.000*	-0.017	0.153
30.	0.031	0.838	0.030	0.023*	68.	0.059	0.652	0.021	0.024*
31.	0.122	0.387	-0.017	0.118	69.	0.078	0.571	-0.006	0.559
32.	-0.163	0.230	-0.019	0.070	70.	0.379	0.022*	-0.002	0.891
33.	0.044	0.760	-0.001	0.900	71.	-0.183	0.196	0.003	0.798
34.	0.173	0.192	-0.002	0.873	72.	0.052	0.729	0.009	0.452
35.	-0.466	0.001*	0.018	0.100	73.	0.132	0.407	0.020	0.162
36.	-0.020	0.890	-0.011	0.360	74.	0.134	0.391	0.007	0.607
37.	-0.031	0.848	0.001	0.932	75.	0.246	0.088	0.008	0.480
38.	-0.166	0.235	-0.007	0.505					

* $p < .05$

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีทดสอบไจสติก (LR เมื่อใช้คะแนนรวม (total test score) เป็นเกณฑ์จับคู่เปรียบเทียบ ซึ่งหากพิจารณาค่า Significant ของกลุ่มผู้สอบ (g) และปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบกับความสามารถ (g by x) แล้ว พบว่าข้อสอบข้อใดมีนัยสำคัญ (Significant) ที่ระดับ .05 หมายความว่าข้อสอบข้อนั้นทำหน้าที่ต่างกัน จาก

ตารางที่ 28 พบว่า ข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันมีทั้งหมด 20 ข้อ ได้แก่ ข้อที่ 3, 5, 8, 9, 10, 16, 24, 27, 29, 30, 35, 39, 44, 51, 54, 62, 66, 67, 68 และ 71 มีข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันแบบเอกรูป จำนวน 15 ข้อ ได้แก่ ข้อที่ 5, 8, 9, 10, 16, 24, 27, 29, 35, 44, 54, 62, 66, 67 และ 70 ส่วนข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันแบบอนเอกรูป มี 5 ข้อ ได้แก่ ข้อที่ 3, 30, 39, 51 และ 68

ตารางที่ 29 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีทดสอบไจสติก เมื่อใช้คะแนนแบบสอบย่อย(subtest scores) เป็นเกณฑ์จับคู่เปรียบเทียบ

แบบสอบย่อย	ข้อ	กลุ่มผู้สอบ (g)				กลุ่มผู้สอบ (g)				
		ปฏิบัติสัมพันธ์		ปฏิบัติสัมพันธ์		ปฏิบัติสัมพันธ์		ปฏิบัติสัมพันธ์		
		B	sig	B	sig	B	sig	B	sig	
SUBTEST 1	1.	0.263	0.126	0.080	0.110	39.	0.233	0.116	0.025	0.438
	2.	0.108	0.457	-0.015	0.722	40.	0.166	0.338	-0.004	0.929
	3.	0.321	0.068	0.047	0.316	41.	0.108	0.456	-0.044	0.306
	4.	-0.021	0.891	0.053	0.231	42.	0.153	0.377	-0.042	0.373
	5.	-0.398	0.007*	-0.021	0.568	43.	-0.099	0.602	0.007	0.899
	6.	-0.027	0.850	-0.007	0.858	44.	-0.295	0.034*	0.002	0.955
	7.	0.046	0.759	0.031	0.504	45.	-0.019	0.897	0.044	0.298
	8.	0.414	0.010*	-0.012	0.781	46.	0.096	0.500	-0.020	0.624
	9.	-0.187	0.232	0.065	0.167	47.	-0.079	0.636	0.018	0.729
	10.	-0.296	0.048*	0.054	0.197	48.	0.120	0.484	-0.057	0.181
SUBTEST 2	11.	-0.039	0.827	-0.047	0.301	49.	0.201	0.180	-0.061	0.130
	12.	-0.048	0.741	0.004	0.916	50.	-0.079	0.589	-0.013	0.753
	13.	0.053	0.720	0.014	0.747	51.	0.039	0.808	-0.045	0.34
	14.	-0.230	0.120	0.005	0.909	52.	0.111	0.467	-0.023	0.598
	15.	-0.064	0.683	0.030	0.547	53.	0.047	0.744	0.043	0.318
	16.	-0.288	0.039*	0.031	0.431	54.	0.436	0.027*	-0.043	0.411

ตารางที่ 29 (ต่อ)

แบบทดสอบย่อย	ข้อ	ปฏิสัมพันธ์ระหว่างกลุ่มกับความสามารถ (g by x)				ข้อ	ปฏิสัมพันธ์ระหว่างกลุ่มกับความสามารถ (g by x)			
		กลุ่มผู้สอบ (g)		ระหว่างกลุ่มกับความสามารถ			กลุ่มผู้สอบ (g)		ระหว่างกลุ่มกับความสามารถ	
		B	sig	B	sig		B	sig	B	sig
(SUBTEST2)	17.	-0.013	0.938	0.130	0.013*	55.	0.065	0.651	0.014	0.737
	18.	-0.166	0.325	-0.469	0.306	56.	0.040	0.777	0.019	0.616
	19.	0.023	0.884	0.072	0.133	57.	-0.127	0.443	0.134	0.020*
	20.	0.110	0.494	0.052	0.283					
SUBTEST 3	21.	-0.129	0.343	0.034	0.460	58.	0.113	0.459	-0.009	0.878
	22.	0.194	0.162	0.048	0.285	59.	0.040	0.770	0.027	0.539
	23.	-0.093	0.506	-0.001	0.990	60.	0.141	0.343	-0.007	0.904
	24.	0.383	0.004*	-0.067	0.116	61.	-0.154	0.302	0.047	0.410
	25.	-0.176	0.223	0.051	0.328	62.	-0.348	0.011*	0.006	0.890
	26.	-0.061	0.649	0.021	0.630	63.	0.011	0.938	-0.062	0.225
	27.	0.442	0.043*	-0.103	0.219	64.	-0.021	0.888	-0.008	0.885
	28.	-0.077	0.041*	0.020	0.641	65.	-0.013	0.933	0.056	0.322
SUBTEST 4	29.	0.204	0.191	0.036	0.405	66.	0.466	0.001*	0.018	0.607
	30.	-0.078	0.616	0.111*	0.013	67.	0.492	0.001*	-0.022	0.573
	31.	0.046	0.750	-0.003	0.931	68.	0.013	0.920	0.059	0.072
	32.	-0.240	0.087	-0.075	0.035*	69.	0.002	0.989	-0.011	0.770
	33.	-0.050	0.755	-0.009	0.822	70.	0.316	0.062	0.004	0.924
	34.	0.128	0.346	0.004	0.909	71.	-0.296	0.547	-0.009	0.810
	35.	-0.520	0.000*	0.069	0.059	72.	-0.051	0.744	0.003	0.941
	36.	-0.127	0.403	0.003	0.949	73.	0.001	0.996	0.055	0.253
	37.	-0.190	0.260	-0.019	0.714	74.	0.014	0.931	0.067	0.160
	38.	-0.258	0.078	0.001	0.974	75.	0.164	0.273	0.063	0.123

* p < .05

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก (LR) เมื่อใช้คะแนนแบบสอบย่อย (subtest scores) เป็นเกณฑ์จับคู่เปรียบเทียบ พิจารณาค่า Significant ของกลุ่มผู้สอบ (g) และปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบกับความสามารถ (g by x) จากตารางที่ 29 พบว่า ข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันมีทั้งหมด 17 ข้อ ได้แก่ ข้อที่ 5, 8, 10, 16, 17, 24, 27, 28, 30, 32, 35, 44, 54, 57, 62, 66, และ 67 มีข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันแบบเอกรูป จำนวน 13 ข้อ ได้แก่ ข้อที่ 5, 8, 10, 16, 24, 27, 28, 35, 44, 54, 62, 66 และ 67 ส่วนข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันแบบอนเอกรูป มี 4 ข้อ ได้แก่ ข้อที่ 17, 30, 32 และ 67

ตารางที่ 30 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก เมื่อใช้คะแนนหลายแบบสอบย่อย (multiple subtest scores) เป็นเกณฑ์จับคู่เปรียบเทียบ

แบบสอบย่อย	ข้อ	กลุ่มผู้สอบ (g)				ปฏิสัมพันธ์ระหว่างกลุ่มกับความสามารถ (g by x)				
		B		sig		B		sig		
		B	sig	B	sig	B	sig	B	sig	
SUBTEST 1	1.	0.287	0.098	0.083	0.097	39.	0.231	0.078	0.019	0.564
	2.	0.091	0.538	-0.009	0.836	40.	0.134	0.442	-0.012	0.973
	3.	0.295	0.059	0.046	0.325	41.	0.146	0.321	-0.050	0.246
	4.	-0.012	0.941	0.054	0.218	42.	0.170	0.328	-0.039	0.400
	5.	-0.390	0.010*	-0.024	0.521	43.	-0.114	0.555	-0.003	0.953
	6.	-0.057	0.691	-0.007	0.855	44.	-0.274	0.049*	0.001	0.976
	7.	0.014	0.926	0.034	0.457	45.	-0.011	0.940	0.044	0.302
	8.	0.403	0.013*	-0.009	0.829	46.	0.091	0.526	-0.020	0.612
	9.	-0.236	0.137	0.065	0.172	47.	-0.139	0.406	0.017	0.743
	10.	-0.302	0.045*	0.057	0.171	48.	0.112	0.518	-0.052	0.222
SUBTEST 2	11.	-0.048	0.789	-0.044	0.325	49.	0.196	0.198	-0.058	0.149
	12.	-0.032	0.828	0.006	0.869	50.	-0.073	0.620	-0.012	0.784
	13.	0.066	0.660	0.010	0.825	51.	0.006	0.971	-0.044	0.357

ตารางที่ 30 (ต่อ)

แบบทดสอบย่อย	ข้อ	ปฏิสัมพันธ์ ระหว่างกลุ่มกับ ความสามารถ (g by x)				ข้อ	ปฏิสัมพันธ์ ระหว่างกลุ่มกับ ความสามารถ (g by x)			
		กลุ่มผู้สอบ (g)		ระหว่างกลุ่มกับ ความสามารถ			กลุ่มผู้สอบ (g)		ระหว่างกลุ่มกับ ความสามารถ	
		B	sig	B	sig		B	sig	B	sig
(SUBTEST 2)	14.	-0.224	0.135	0.003	0.948	52.	0.105	0.496	-0.024	0.604
	15.	-0.105	0.513	0.026	0.613	53.	0.061	0.672	0.041	0.344
	16.	-0.289	0.040*	0.031	0.444	54.	0.402	0.043*	-0.041	0.431
	17.	-0.002	0.991	0.134	0.061	55.	0.053	0.716	0.012	0.780
	18.	-0.146	0.361	-0.048	0.295	56.	0.029	0.841	0.019	0.235
	19.	0.023	0.883	0.072	0.132	57.	-0.151	0.366	0.133	0.072
	20.	0.116	0.473	0.052	0.284					
SUBTEST 3	21.	-0.144	0.297	0.040	0.387	58.	0.141	0.361	-0.010	0.854
	22.	0.222	0.116	0.062	0.256	59.	0.025	0.854	0.033	0.461
	23.	-0.056	0.693	-0.004	0.943	60.	0.135	0.369	-0.009	0.877
	24.	0.372	0.006*	-0.067	0.116	61.	-0.121	0.422	0.043	0.445
	25.	-0.192	0.187	0.052	0.318	62.	-0.386	0.005*	0.014	0.756
	26.	-0.064	0.633	0.024	0.574	63.	-0.001	0.994	-0.060	0.240
	27.	0.459	0.041*	-0.011	0.201	64.	-0.016	0.918	-0.010	0.860
	28.	-0.277	0.062	0.023	0.599	65.	0.005	0.976	0.050	0.380
SUBTEST 4	29.	0.167	0.289	0.042	0.343	66.	0.477	0.001*	0.013	0.717
	30.	-0.076	0.229	0.110	0.015*	67.	0.515	0.001*	-0.029	0.457
	31.	0.063	0.663	-0.005	0.893	68.	-0.017	0.902	0.068	0.089
	32.	-0.254	0.071	-0.071	0.055	69.	-0.005	0.972	-0.008	0.835
	33.	-0.037	0.805	-0.007	0.860	70.	-0.295	0.087	0.006	0.879
	34.	0.123	0.371	0.011	0.736	71.	-0.286	0.055	-0.008	0.833

ตารางที่ 30 (ต่อ)

แบบทดสอบย่อย	ข้อ	ปฏิสัมพันธ์ ระหว่างกลุ่มกับ ความสามารถ (g by x)				ข้อ	ปฏิสัมพันธ์ ระหว่างกลุ่มกับ ความสามารถ (g by x)			
		กลุ่มผู้สอบ (g)					กลุ่มผู้สอบ (g)			
		B	sig	B	sig		B	sig	B	sig
(SUBTEST 4)	36.	-0.537	0.000*	0.069	0.061	72.	-0.033	0.832	0.001	0.989
	36.	-0.141	0.356	0.002	0.965	73.	-0.029	0.862	0.050	0.312
	37.	-0.141	0.409	-0.022	0.662	74.	0.046	0.778	0.065	0.164
	38.	-0.235	0.105	-0.001	0.989	75.	0.135	0.374	0.060	0.147

* $p < .05$

ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก (LR) เมื่อใช้คะแนนหลายแบบทดสอบย่อย (multiple subtest scores) เป็นเกณฑ์จับคู่เปรียบเทียบ พิจารณาว่า Significant ของกลุ่มผู้สอบ (g) และปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบกับความสามารถ (g by x) จากตารางที่ 30 พบว่า ข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันมีทั้งหมด 13 ข้อ ได้แก่ ข้อที่ 5, 8, 10, 16, 24, 27, 30, 35, 44, 54, 62, 66, และ 67 มีข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันแบบเอกรูปจำนวน 12 ข้อ ได้แก่ ข้อที่ 5, 8, 10, 16, 24, 27, 35, 44, 54, 62, 66, และ 67 ส่วนข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันแบบอนเนกรูป มี 1 ข้อ ได้แก่ ข้อที่ 30

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 31 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

วิธีที่ใช้ในการตรวจสอบ	ข้อสอบที่ทำหน้าที่ต่างกัน	จำนวน (ข้อ)	ร้อยละ
SIBTEST	5, 8, 9, 10, 16, 22, 24, 27, 29, 35, 44, 54, 62, 66, 67, 75	16	21.33
MH _{TOT}	5, 9, 10, 16, 24, 27, 28, 29, 35, 44, 54, 62, 66, 67, 70	15	20.00
MH _{SUB}	5, 8, 10, 16, 24, 27, 28, 35, 44, 54, 62, 66, 67, 71	14	18.67
LR _{TOT}	3, 5, 8, 9, 10, 16, 24, 27, 29, 30, 35, 39, 44, 51, 54, 62, 66, 67, 68, 70,	20	26.67
LR _{SUB}	5, 8, 10, 16, 17, 24, 27, 28, 30, 32, 35, 44, 54, 57, 62, 66, 67	17	22.67
LR _{MULTIPLE}	5, 8, 10, 16, 24, 27, 30, 35, 44, 54, 62, 66, 67	13	17.33

จากตารางที่ 31 เป็นการเปรียบเทียบจำนวนข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันซึ่งตรวจสอบด้วยวิธีต่าง ๆ ปรากฏว่า วิธี SIBTEST ซึ่งเป็นวิธีที่ใช้ในการตรวจสอบประสิทธิภาพของวิธีที่นำมาเปรียบเทียบ ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 16 ข้อ (21.33%) สำหรับวิธีดัดถอยโลจิสติก เมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่เปรียบเทียบ (LR_{TOT}) ตรวจพบข้อสอบทำหน้าที่ต่างกันอย่างมากที่สุด จำนวน 20 ข้อ (26.67%) รองลงมา ได้แก่ วิธีดัดถอยโลจิสติกที่ใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ (LR_{SUB}) ตรวจพบจำนวน 17 ข้อ (22.67%) วิธีแมนเทล - แอนส์เชล เมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่เปรียบเทียบ (MH_{TOT}) ตรวจพบจำนวน 15 ข้อ (20%) และวิธีแมนเทล - แอนส์เชล เมื่อใช้แบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ (MH_{SUB}) ตรวจพบจำนวน 14 ข้อ (18.67%) ตามลำดับ ส่วนวิธีที่ตรวจพบข้อสอบทำหน้าที่ต่างกันอย่างน้อยที่สุด คือ วิธีดัดถอยโลจิสติกที่ใช้คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ (LR_{MULTIPLE}) ซึ่งตรวจพบข้อสอบทำหน้าที่ต่างกันจำนวน 13 ข้อ (17.33%)

ตอนที่ 6 เปรียบเทียบอัตราความถูกต้องและอัตราความคลาดเคลื่อนระหว่างวิธีแมนเทิล-แฮนส์เซลกับวิธีถดถอยโลจิสติกในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบตามเกณฑ์จับคู่เปรียบเทียบ

6.1 อัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล - แฮนส์เซลกับวิธีถดถอยโลจิสติก เมื่อแยกตามเกณฑ์จับคู่เปรียบเทียบ

ตารางที่ 32 ผลการวิเคราะห์อัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล กับวิธีถดถอยโลจิสติก เมื่อแยกตามเกณฑ์จับคู่เปรียบเทียบ

วิธีที่ใช้ในการตรวจสอบDIF	อัตราความถูกต้อง ประเภทที่ 1	อัตราความถูกต้อง ประเภทที่ 2	อัตราความคลาดเคลื่อนประเภทที่1	อัตราความคลาดเคลื่อนประเภทที่2
MH _{TOT}	0.966	0.813	0.034	0.187
MH _{SUB}	0.966	0.760	0.034	0.250
LR _{TOT}	0.898	0.876	0.102	0.126
LR _{SUB}	0.915	0.760	0.085	0.250
LR _{MULTIPLE}	0.983	0.760	0.017	0.250

จากตารางที่ 32 เป็นผลการวิเคราะห์อัตราความถูกต้องประเภทที่ 1, 2 และอัตราความคลาดเคลื่อนประเภทที่ 1, 2 ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล (MH) กับวิธีถดถอยโลจิสติก (LR) ตามเกณฑ์จับคู่เปรียบเทียบ พบว่า

วิธี LR เมื่อใช้คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ (LR_{MULTIPLE}) มีอัตราความถูกต้องประเภทที่ 1 สูงสุด เท่ากับ 0.983 นั่นคือ มีการตรวจสอบพบว่าข้อสอบทำหน้าที่ไม่ต่างกันถูกต้องตรงกับวิธี SIBTEST ซึ่งเป็นวิธีที่ใช้ตรวจสอบคิดเป็นร้อยละ 98.3 รองลงมา คือ วิธี MH เมื่อใช้คะแนนรวมและคะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ (MH_{TOT} และ MH_{SUB}) มีอัตราความถูกต้องประเภทที่ 1 เท่ากับ 0.966 และวิธี LR ซึ่งใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ (LR_{SUB}) มีค่าเท่ากับ 0.915 ตามลำดับ ส่วนวิธีที่มีอัตราความถูกต้องประเภทที่ 1 ต่ำสุด คือ วิธี LR เมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่เปรียบเทียบ (LR_{TOT}) มีค่าเท่ากับ 0.898

สำหรับอัตราความถูกต้องประเภทที่ 2 ซึ่งหมายถึง ร้อยละของจำนวนข้อสอบที่ตรวจสอบพบว่าทำหน้าที่ต่างกันได้ถูกต้องตรงกับวิธี SIBTEST ปากกว่าวิธี LR_{TOT} เป็นวิธีที่มีอัตราความถูกต้องประเภทที่ 2 สูงสุด เท่ากับ 0.875 นั่นคือ มีการตรวจสอบพบว่าข้อสอบทำหน้าที่ต่างกันถูกต้องคิดเป็นร้อยละ 87.5 รองลงมาได้แก่ วิธี MH_{TOT} , MH_{SUB} , LR_{SUB} และ $LR_{MULTIPLE}$ มีอัตราความถูกต้องประเภทที่ 2 เท่ากับ 0.813, 0.750, 0.750 และ 0.750 ตามลำดับ

ส่วนอัตราความคลาดเคลื่อนประเภทที่ 1 ซึ่งหมายถึง การระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความจริงข้อสอบทำหน้าที่ไม่ต่างกัน พบว่าวิธี LR_{TOT} เป็นวิธีที่มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงสุด เท่ากับ 0.102 นั่นหมายถึง มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันร้อยละ 10.2 รองลงมาได้แก่ วิธี LR_{SUB} , MH_{TOT} และ MH_{SUB} มีอัตราความคลาดเคลื่อนประเภทที่ 1 เท่ากับ 0.085, 0.034 และ 0.034 ตามลำดับ ส่วนวิธีที่มีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำสุด คือ วิธี $LR_{MULTIPLE}$ เท่ากับ 0.017 นั่นคือ มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันเพียงร้อยละ 1.7

สำหรับอัตราความคลาดเคลื่อนประเภทที่ 2 ซึ่งหมายถึงการระบุผิดพลาดว่าข้อสอบทำหน้าที่ไม่ต่างกัน ทั้งที่ความจริงข้อสอบทำหน้าที่ต่างกันพบว่า วิธี MH_{SUB} , LR_{SUB} และ $LR_{MULTIPLE}$ เป็นวิธีที่มีอัตราความคลาดเคลื่อนประเภทที่ 2 สูงสุด เท่ากับ 0.250 นั่นคือ มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ไม่ต่างกันร้อยละ 25 รองลงมา ได้แก่ วิธี MH_{TOT} และ LR_{TOT} มีอัตราความคลาดเคลื่อนประเภทที่ 2 เท่ากับ 0.187 และ 0.125 ตามลำดับ ซึ่งวิธีที่มีอัตราความคลาดเคลื่อนประเภทที่ 2 ต่ำสุด คือ วิธี LR_{TOT} มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันเพียงร้อยละ 12.5

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

6.2 การทดสอบนัยสำคัญของอัตราความถูกต้องและอัตราความคลาดเคลื่อนของผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เชลกับวิธีถดถอยโลจิสติก เมื่อใช้เกณฑ์จับคู่เปรียบเทียบเหมือนกัน

ตารางที่ 33 ผลการทดสอบนัยสำคัญของการเปรียบเทียบอัตราความถูกต้อง และอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีแมนเทิล - แฮนส์เชลกับวิธีถดถอยโลจิสติกเมื่อใช้เกณฑ์จับคู่เปรียบเทียบเหมือนกัน

วิธีที่ใช้ในการตรวจสอบ DIF		Z - test		
ที่นำมาทดสอบนัยสำคัญ	ถูกต้อง 1	ถูกต้อง 2	type I error	type II error
LR _{TOT} : MH _{TOT}	- 1.659*	1.061	1.659*	- 1.061
LR _{SUB} : MH _{SUB}	- 1.308	0.000	1.308	0.000

* $p < .05$

จากตารางที่ 33 เมื่อทดสอบนัยสำคัญของการเปรียบเทียบอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีแมนเทิล-แฮนส์เชล (MH) กับวิธีถดถอยโลจิสติก (LR) เมื่อใช้เกณฑ์จับคู่คะแนนรวม (total test score) เหมือนกันทั้งสองวิธีพบว่า วิธี MH มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี LR เนื่องจากวิธี MH มีอัตราความถูกต้องประเภทที่ 1 มากกว่าและมีอัตราความคลาดเคลื่อนประเภทที่ 1 น้อยกว่าวิธี LR อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สำหรับการทดสอบนัยสำคัญของการเปรียบเทียบอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ระหว่างวิธีแมนเทิล-แฮนส์เชลกับวิธีถดถอยโลจิสติก เมื่อใช้คะแนนแบบสอบย่อย (subtest score) เป็นเกณฑ์จับคู่เหมือนกันทั้งสองวิธีพบว่า ไม่สามารถปฏิเสธสมมติฐานศูนย์ได้ ดังนั้นจึงไม่มีหลักฐานเพียงพอที่จะสรุปว่าวิธี LR มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี MH

ตารางที่ 34 ผลการทดสอบนัยสำคัญของอัตราความถูกต้อง และอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก เมื่อใช้เกณฑ์จับคู่คะแนนหลายแบบสอบย่อยเปรียบเทียบกับการใช้เกณฑ์จับคู่คะแนนรวมและคะแนนแบบสอบย่อย

วิธีที่ใช้ในการตรวจสอบ DIF ที่นำมาทดสอบนัยสำคัญ	Z - test			
	ถูกต้อง 1	ถูกต้อง 2	type I error	type II error
$LR_{MULTIPLE} : LR_{TOT}$	2.180*	- 1.953*	- 2.180*	1.953*
$LR_{MULTIPLE} : LR_{SUB}$	1.889*	0.000	- 1.889*	0.000

* $p < .05$

จากตารางที่ 34 เมื่อทดสอบนัยสำคัญของอัตราความถูกต้อง และอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก (LR) เมื่อใช้เกณฑ์จับคู่คะแนนหลายแบบสอบย่อย (multiple subtest scores) เปรียบเทียบกับการใช้เกณฑ์จับคู่คะแนนรวม (total test score) ด้วยสถิติ Z โดยพิจารณาอัตราความถูกต้องและอัตราความคลาดเคลื่อนประเภทที่ 1 เป็นหลัก พบว่าวิธี $LR_{MULTIPLE}$ มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี LR_{TOT} เนื่องจากวิธี $LR_{MULTIPLE}$ มีอัตราความถูกต้องประเภทที่ 1 มากกว่าและมีอัตราความคลาดเคลื่อนประเภทที่ 1 น้อยกว่าวิธี LR_{TOT} อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

เมื่อทดสอบนัยสำคัญของอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก เมื่อใช้เกณฑ์จับคู่คะแนนหลายแบบสอบย่อยเปรียบเทียบกับการใช้เกณฑ์จับคู่คะแนนแบบสอบย่อย (subtest score) ด้วยสถิติ Z พบว่าวิธี $LR_{MULTIPLE}$ มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี LR_{SUB} เนื่องจากวิธี $LR_{MULTIPLE}$ มีอัตราความถูกต้องประเภทที่ 1 มากกว่าและมีอัตราความคลาดเคลื่อนประเภทที่ 1 น้อยกว่าวิธี LR_{TOT} อย่างมีนัยสำคัญทางสถิติที่ระดับ .05