บทที่ 4

ผลการพัฒนาเกณฑ์และการวิเคราะห์ข้อมูล

วัตถุประสงค์ของการวิจัยเพื่อพัฒนาเกณฑ์การตัดสินความลำเอียงของดัชนี 4 ตัว คือ (1) ดัชนี SA และ (2) ดัชนี UA ซึ่งได้จากวิธีการวิเคราะห์ความลำเอียงตามทฤษฎีการตอบข้อสอบโมเดล 2 พารามิเตอร์ (3) ดัชนี α_{MH} จากวิธี MH และ (4) ดัชนี β_{SIB} จากวิธี (SIBTEST) โดยใช้ข้อมูลการตอบแบบสอบคัดเลือกบุคคลเพื่อเข้าศึกษาในระดับอุดมศึกษาของทบวงมหาวิทยาลัยประจำปีการศึกษา 2535 จำนวน 2 วิชา คือ วิชาคณิตศาสตร์ กข และวิชาภาษาอังกฤษ กข ซึ่งเป็นแบบสอบเลือกตอบ 4 ตัวเลือก ความลำเอียงที่ศึกษาคือความลำเอียงระหว่างเพศของผู้สอบ

บริบทในการศึกษาประกอบด้วยแบบสอบ 7 ความยาว คือ ความยาว 20 30 และ 40 ข้อ สำหรับวิชาคณิตศาสตร์และ 50 60 70 และ 80 ข้อ สำหรับวิชาภาษาอังกฤษ แบบสอบแต่ละความยาวมีเนื้อหาที่วัดครบตามกลุ่มเนื้อหาของข้อสอบฉบับเต็ม (คณิตศาสตร์ 41 ข้อ และภาษาอังกฤษ 100 ข้อ) ก่อนที่จะตัดข้อกระทงที่มีค่าอำนาจจำแนกต่ำกว่า .20 และค่าความยากสูงกว่า .80 หรือต่ำกว่า .20 ออกแล้ว

สำหรับขนาดผู้สอบที่ใช้ในการศึกษาประกอบด้วยขนาดผู้สอบ 6 ขนาด ได้แก่ ขนาด 100 200 400 600 800 และ 1,000 คน จากผู้เข้าสอบที่เป็นประชากรทั้งหมด 25,459 คน สำหรับวิชาคณิตศาสตร์ เป็นหญิง 13,480 คน (ร้อยละ 52.95) และชาย 11,977 คน (ร้อยละ 47.04) และสำหรับวิชาภาษาอังกฤษมีจำนวนประชากร 16,026 คน เป็นหญิง 8,876 คน (ร้อยละ 44.58) และชาย 8,878 คน (ร้อยละ 55.40)

ผลการวิเคราะห์ข้อมูลในบทนี้ผู้วิจัยได้นำเสนอเป็น 4 ตอน คือ

<u>ตอนที่ 1</u> ผลการพัฒนาเกณฑ์ตัดสินข้อสอบลำเอียงของดัชนี SA UA α_{MH} และ β_{SIB}

<u>ตอนที่ 2</u> ผลการการตัดสินข้อสอบลำเอียงทางเพศ วิชาภาษาอังกฤษ และวิชาคณิตศาสตร์ ด้วยเกณฑ์ที่พัฒนาจากข้อมูลเชิงประจักษ์

ตอนที่ 3 ความสอดคล้องในผลการตัดสินข้อสอบลำเอียงทางเพศวิชาคณิตศาสตร์และวิชาภาษาอังกฤษ เมื่อใช้การทดสอบความมีนัยสำคัญของค่าสถิติ และการใช้เกณฑ์ที่พัฒนาจากข้อมูลเชิงประจักษ์

ตอนที่ 4 ผลการเปรียบเทียบลักษณะลำเอียงของข้อสอบ ที่มีต่อเพศของผู้สอบในวิชาภาษาอังกฤษและวิชาคณิตศาสตร์

ตอนที่ 1 ผลการพัฒนาเกณฑ์ตัดสินข้อสอบลำเอียงของดัชนี SA UA α_{MH} และ β_{SIB}

1. ดัชนี SA

ตารางที่ 54 ในภาคผนวก ง เป็นตารางสรุปค่าสูงสุดและต่ำสุดของค่าเฉลี่ยรายข้อของดัชนี SA ในความยาวแบบสอบและขนาดผู้สอบต่างกัน ซึ่งพบว่าที่ความยาวแบบสอบ 40 ข้อ วิชาคณิตศาสตร์ ค่าเฉลี่ยรายข้อมีพิสัยกว้างกว่าที่ความยาวแบบสอบอื่น และมีค่าสูงสุดสูงกว่าค่าในความยาวแบบสอบอื่นด้วยเช่นกัน แต่ผลเช่นนี้ไม่พบในวิชาภาษาอังกฤษ นอกจากนั้นยังพบว่า ที่ขนาดผู้สอบ 1,000 คน ค่าเฉลี่ยรายข้อมีพิสัยกว้างกว่าขนาดอื่นทุกความยาวแบบสอบ และเป็นเช่นเดียวกันทั้งวิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ และมีค่าเฉลี่ยสูงสุดสูงกว่าด้วย

เมื่อวิเคราะห์ค่าเฉลี่ยรายข้อรวมทุกความยาวแบบสอบจำแนกตามขนาดผู้สอบ 6 ขนาด คือ 100 200 400 600 800 และ 1,000 คน ได้ผลดังตารางที่ 55 และ 56 ในภาคผนวก ง และค่าเฉลี่ยรายข้อรวมทุกขนาดผู้สอบจำแนกตามความยาวแบบสอบแสดงไว้ในตารางที่ 57 และ 58 ในภาคผนวก ง เช่นกัน ทั้งนี้ได้รวมสรุปค่าเฉลี่ยเมื่อรวมทุกข้อทุกความยาวแบบสอบและทุกขนาดผู้สอบไว้ในตารางที่ 6

ตารางที่ 6 ค่าเฉลี่ยรวมทุกข้อของดัชนี SA จากผู้สอบเพศเดียวกัน วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ เมื่อความยาวแบบสอบต่างกันในขนาดผู้สอบเดียวกัน

วิชา	ขนาดผู้สอบ											
ความยาวแบบสอบ	100		200		400		600		800		1,000	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
คณิตศาสตร์												
20 ข้อ	.39	.11	.38	.14	.38	.16	.38	.17	.42	.22	.37	.20
30 ข้อ	.35	.08	.33	.10	.34	.14	.35	.16	.35	.19	.38	.25
40 ข้อ	.89	.27	.73	.28	.64	.31	.63	.33	.60	.38	.58	.36
ภาษาอังกฤษ												
50 ข้อ	.30	.11	.27	.17	.25	.16	.24	.17	.24	.19	.53	.51
60 ข้อ	.27	.10	.25	.12	.24	.15	.23	.17	.22	.17	.22	.20
70 ข้อ	.27	.10	.24	.12	.23	.15	.23	.16	.22	.17	.22	.21
80 ข้อ	.25	.09	.24	.12	.23	.16	.22	.17	.22	.19	.22	.22

จากตารางที่ 6 จะเห็นว่า สำหรับวิชาคณิตศาสตร์เมื่อพิจารณาตามขนาดผู้สอบ ค่าเฉลี่ยรวมทุกข้อมีค่าใกล้เคียงกัน แต่เมื่อพิจารณาตามความยาวแบบสอบพบว่า ค่าเฉลี่ยที่ความยาวแบบสอบ 40 ข้อ มีค่าสูงกว่าค่าเฉลี่ยที่ความยาว 20 และ 30 ข้อ ทุกขนาดผู้สอบ และมีส่วนเบี่ยงเบนมาตรฐานสูงกว่าด้วย ในขณะที่ค่าเฉลี่ยรวมทุกข้อของวิชาภาษาอังกฤษเมื่อพิจารณาตามขนาดผู้สอบมีค่าใกล้เคียงกันเช่นกัน แต่เมื่อพิจารณาที่ความยาวแบบสอบพบว่า ค่าเฉลี่ยที่ความยาวแบบสอบ 50 ข้อ มีค่าสูงกว่าความยาวอื่นเล็กน้อย โดยเฉพาะอย่างยิ่งที่ขนาดผู้สอบ 1,000 คน แต่ความแตกต่างไม่มากเท่าที่พบในวิชาคณิตศาสตร์

เมื่อวิเคราะห์ค่าเฉลี่ยรวมทุกข้อและทุกความยาวแบบสอบจำแนกตามขนาดผู้สอบ รวมทุกข้อและทุกขนาดผู้สอบจำแนกตามความยาวแบบสอบ และรวมทุกข้อทุกขนาดผู้สอบ และทุกความยาวแบบสอบพบผลดังตารางที่ 7

ตารางที่ 7 ค่าเฉลี่ยรวมทุกข้อของดัชนี SA จากผู้สอบเพศเดียวกันวิชาคณิตศาสตร์และวิชาภาษาอังกฤษ จำแนกตามขนาดผู้สอบและความยาวแบบสอบ

วิชา	ขนาดผู้สอบ ความยาวแบบสอบ	$\bar{X}$	S.D.	S.E.	ค่าสูงสุด*
คณิตศาสตร์	ขนาดผู้สอบ				
	100 คน	.58	.31	.03	.66
	200 คน	.51	.27	.03	.59
	400 คน	.48	.28	.03	.56
	600 คน	.48	.30	.03	.56
	800 คน	.47	.32	.03	.55
	1,000 คน	.47	.33	.03	.55
	รวม	.50	.04	.02	.55
ภาษาอังกฤษ	100 คน	.270	.10	.01	.30
	200 คน	.245	.12	.01	.27
	400 คน	.235	.15	.01	.26
	600 คน	.224	.16	.01	.25
	800 คน	.220	.17	.01	.25
	1,000 คน	.293	.31	.02	.35
	รวม	.24	.02	.01	.26

ตารางที่ 7 (ต่อ)

วิชา	ขนาดผู้สอบ ความยาวแบบสอบ	$\bar{X}$	S.D.	S.E.	ค่าสูงสุด*
คณิตศาสตร์	ความยาวแบบสอบ				
	20 ข้อ	.390	.17	.02	.442
	30 ข้อ	.343	.16	.01	.369
	40 ข้อ	.670	.35	.02	.722
	รวม	.46	.18	.11	.74
ภาษาอังกฤษ	50 ข้อ	.310	.27	.02	.362
	60 ข้อ	.240	.15	.01	.265
	70 ข้อ	.230	.15	.01	.256
	80 ข้อ	.220	.15	.01	.246
	รวม	.25	.04	.02	.31

* = $\bar{X} \pm 2.58$ (S.E.)

การกำหนดเกณฑ์ตัดสินข้อสอบลำเอียงของดัชนี SA

1.1 การกำหนดเกณฑ์โดยไม่พิจารณาถึงความแตกต่างข้ามขนาดผู้สอบและความยาวแบบสอบ

ค่าเฉลี่ยรวมทุกข้อทุกความยาวผู้สอบจากตารางที่ 6 พบว่ายอดรวมทั้ง 4 แถว สำหรับวิชาคณิตศาสตร์มีค่าสูงสุด 2 ค่า คือ .55 และ .75 และค่าสูงสุดสำหรับวิชาภาษาอังกฤษมี 2 ค่าเช่นกัน คือ .26 และ .31 ผู้วิจัยได้พิจารณากำหนดเกณฑ์โดยพิจารณาค่าสูงสุดสำหรับวิชาคณิตศาสตร์คือ ค่า $|SA| > .75$ และสำหรับวิชาภาษาอังกฤษคือ ค่า $|SA| > .35$

1.2 การกำหนดเกณฑ์โดยพิจารณาความแตกต่างข้ามขนาดผู้สอบและความยาวแบบสอบ

จะเห็นได้ว่า เมื่อพิจารณาค่าเฉลี่ยตามขนาดผู้สอบ พบว่ามีค่าใกล้เคียงกัน ยกเว้นขนาดผู้สอบ 100 คน ในวิชาคณิตศาสตร์และขนาดผู้สอบ 100 และ 1,000 คน สำหรับวิชาภาษาอังกฤษ แต่ค่าเฉลี่ยสูงสุดยังมีค่าต่ำกว่าเกณฑ์ที่กำหนด (คณิตศาสตร์ > .75 ภาษาอังกฤษ > .35) แต่เมื่อพิจารณาตามขนาดผู้สอบ จะเห็นว่าค่าสูงสุดของค่าเฉลี่ย ในแต่ละความยาวแบบสอบ

สำหรับวิชาคณิตศาสตร์ เมื่อพิจารณาค่าเฉลี่ยทั้ง 2 ลักษณะแล้ว พบว่าลักษณะการกำหนดเกณฑ์ อาจเป็นดังนี้

ขนาด 100 คน	SA > .70
ขนาดอื่น ๆ	SA > .60
ความยาวแบบสอบ 20 ข้อ	SA > .50
ความยาวแบบสอบ 30 ข้อ	SA > .40
ความยาวแบบสอบ 40 ข้อ	SA > .80

เพื่อให้มีการสอดคล้องสำหรับการกำหนดเกณฑ์จากผลทั้ง 2 ลักษณะ ผู้วิจัยเสนอให้ใช้เกณฑ์ดังนี้

ความยาวแบบสอบ 20-30 ข้อ	เกณฑ์ SA > .70
ความยาวแบบสอบ 40 ข้อ	เกณฑ์ SA > .80

สำหรับวิชาภาษาอังกฤษ ยกเว้นที่ 50 ข้อ เมื่อพิจารณาค่าจากทั้ง 2 กรณีแล้ว วิชาภาษาอังกฤษอาจกำหนดค่าดังนี้

ขนาดผู้สอบ 100 คน	SA > .40	
ขนาดผู้สอบ 1,000 คน	SA > .40	และ
ขนาด 200, 400, 600 และ 800 คน	SA > .30	
ความยาวแบบสอบ 50 ข้อ	SA > .40	
ความยาวแบบสอบ 60-80 ข้อ	SA > .30	

ค่าที่ผู้วิจัยเสนอเพื่อใช้ คือ SA > .40 ในทุกขนาดและความยาวแบบสอบ ซึ่งจะครอบคลุมทั้งด้านขนาดผู้สอบและความยาวแบบสอบ

2. <u>ดัชนี UA</u>

ตารางที่ 59 ในภาคผนวก ง เป็นตารางสรุปค่าต่ำสุด และสูงสุดของค่าเฉลี่ยรายข้อของดัชนี UA เมื่อมีความยาวแบบสอบและขนาดผู้สอบต่างกัน จะเห็นว่าพิสัยของค่าเฉลี่ยสูงขึ้นทั้ง 2 วิชา เมื่อขนาดผู้สอบมากขึ้น และเมื่อพิจารณาความแตกต่างเมื่อความยาวแบบสอบต่างกันพบลักษณะคล้ายกับที่พบในกรณีของดัชนี SA คือ ในวิชาคณิตศาสตร์ ที่ความยาวแบบสอบ 40 ข้อ พิสัยของค่าเฉลี่ยรายข้อมีค่าสูงกว่าที่ความยาว 20 และ 30 ข้อ ในขณะที่วิชาภาษาอังกฤษมีพิสัยของค่าเฉลี่ยในความยาวแบบสอบต่างกันใกล้เคียงกัน ทั้งนี้รวมสรุปค่าเฉลี่ยเมื่อรวมทุกข้อทุกความยาวแบบสอบและทุกขนาดผู้สอบไว้ในตารางที่ 8

จากตารางที่ 8 จะเห็นว่า ค่าเฉลี่ยรวมทุกข้อของดัชนี UA ในวิชาคณิตศาสตร์มีลักษณะเช่นเดียวกับดัชนี SA คือ ค่าเฉลี่ยเมื่อขนาดผู้สอบต่างกันมีค่าใกล้เคียงกัน แต่เมื่อพิจารณาตามความยาวแบบสอบแล้วพบว่า ค่าเฉลี่ยรวมทุกข้อที่ขนาดแบบสอบ 40 ข้อ มีค่าสูงกว่าขนาดอื่น

ตารางที่ 8 ค่าเฉลี่ยรวมทุกข้อของดัชนี UA จากผู้สอบเพศเดียวกัน วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ เมื่อความยาวแบบสอบต่างกันในขนาดผู้สอบเดียวกัน

วิชา	ขนาดผู้สอบ											
ความยาวแบบสอบ	100		200		400		600		800		1,000	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
คณิตศาสตร์												
20 ข้อ	.61	.11	.62	.14	.63	.21	.64	.24	.72	.30	.64	.30
30 ข้อ	.51	.09	.52	.12	.56	.21	.58	.22	.58	.27	.63	.36
40 ข้อ	1.33	.22	1.17	.29	1.05	.41	1.01	.46	.98	.51	.88	.47
ภาษาอังกฤษ												
50 ข้อ	.48	.14	.47	.17	.45	.22	.44	.25	.44	.27	.43	.30
60 ข้อ	.42	.12	.43	.16	.43	.21	.42	.24	.40	.24	.41	.29
70 ข้อ	.43	.13	.42	.16	.42	.21	.42	.24	.40	.24	.41	.29
80 ข้อ	.40	.12	.40	.16	.41	.22	.40	.24	.40	.27	.41	.31

นอกจากนี้ยังพบว่า ค่าที่ได้มีแนวโน้มลดลงเมื่อขนาดผู้สอบเพิ่มขึ้น โดยจะเห็นได้ว่าที่ขนาดผู้สอบ 100 คน ค่าเฉลี่ยเท่ากับ 1.33 แต่ที่ขนาดผู้สอบ 1,000 คน ค่าเฉลี่ยลดลงเหลือ .88

สำหรับวิชาภาษาอังกฤษ พบว่าค่าเฉลี่ยที่ความยาวแบบสอบ 50 ข้อ มีค่าสูงกว่าที่ความยาวอื่นเล็กน้อย (ทศนิยมตำแหน่งที่ 2) แต่ไม่พบความแตกต่างข้ามขนาดผู้สอบ

เมื่อวิเคราะห์ค่าเฉลี่ยรวมทุกข้อ และทุกความยาวแบบสอบ จำแนกตามขนาดผู้สอบ รวมทุกข้อและทุกขนาดผู้สอบ จำแนกตามความยาวแบบสอบและรวมทุกข้อทุกขนาดผู้สอบ และทุกความยาวแบบสอบพบผลดังตารางที่ 9

ตารางที่ 9 ค่าเฉลี่ยรวมทุกข้อของดัชนี UA จากผู้สอบเพศเดียวกัน วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ จำแนกตามขนาดผู้สอบและความยาวแบบสอบ

วิชา	ขนาดผู้สอบ ความยาวแบบสอบ	$\bar{X}$	S.D.	S.E.	ค่าสูงสุด*
คณิตศาสตร์	ขนาดผู้สอบ				
	100 คน	.889	.412	.043	1.00
	200 คน	.820	.366	.039	.92
	400 คน	.802	.398	.042	.91
	600 คน	.794	.421	.044	.90
	800 คน	.797	.455	.048	.92
	1,000 คน	.774	.469	.049	.90
	รวม	.81	.04	.016	.85
ภาษาอังกฤษ	100 คน	.422	.127	.008	.44
	200 คน	.421	.162	.010	.45
	400 คน	.417	.212	.013	.45
	600 คน	.410	.238	.015	.45
	800 คน	.397	.250	.016	.44
	1,000 คน	.403	.291	.018	.45
	รวม	.41	.01	.00	.41

ตารางที่ 9 (ต่อ)

วิชา	ขนาดผู้สอบ ความยาวแบบสอบ	$\bar{X}$	S.D.	S.E.	ค่าสูงสุด*
คณิตศาสตร์	ความยาวแบบสอบ				
	20 ข้อ	.64	.22	.02	.69
	30 ข้อ	.57	.23	.02	.62
	40 ข้อ	1.07	.42	.04	1.17
	รวม**	.76	.27*	.156	1.16
ภาษาอังกฤษ	50 ข้อ	.45	.23	.01	.48
	60 ข้อ	.42	.22	.01	.45
	70 ข้อ	.42	.22	.01	.45
	80 ข้อ	.40	.23	.01	.43
	รวม	.42	.02	.01	.45

* = $\bar{X} \pm 2.58$ (S.E.)

การกำหนดเกณฑ์ตัดสินข้อสอบลำเอียงของดัชนี UA

2.1 การกำหนดเกณฑ์โดยไม่พิจารณาถึงความแตกต่างข้ามขนาดผู้สอบ และความยาวแบบสอบ

จากตารางที่ 9 ในยอดรวมทั้ง 4 แถว พบว่าค่าสูงสุดสำหรับวิชาคณิตศาสตร์มี 2 ค่า คือ .85 และ 1.16 และค่าสูงสุดสำหรับวิชาภาษาอังกฤษมี 2 ค่า เช่นกัน คือ .41 และ .45 ผู้วิจัยได้พิจารณากำหนดเกณฑ์โดยพิจารณาค่าสูงสุด สำหรับวิชาคณิตศาสตร์คือ UA > 1.20 และสำหรับวิชาภาษาอังกฤษ คือ UA > .50

2.2 การกำหนดเกณฑ์โดยพิจารณาความแตกต่างข้ามขนาดผู้สอบและความยาวแบบสอบ

สำหรับวิชาคณิตศาสตร์ เมื่อพิจารณาให้ค่าที่กำหนดครอบคลุมทั้งในด้านขนาดผู้สอบและความยาวแบบสอบแล้ว ค่าที่กำหนดควรจะเป็น 2 ลักษณะ

ความยาวแบบสอบ 20 และ 30 ข้อ UA > 1.00

ความยาวแบบสอบ 40 ข้อ UA > 1.20

จากค่าเฉลี่ยในวิชาภาษาอังกฤษ พบว่าค่าเฉลี่ยทุกความยาวแบบสอบ และทุกขนาดแบบสอบมีค่าใกล้เคียงกันมากและถ้าใช้ข้อมูลตามนี้เป็นฐาน เกณฑ์การตัดสินความลำเอียงควรเป็น UA > .50 ซึ่งเป็นเกณฑ์ในกรณีข้อ 2.1

ดังนั้นผู้วิจัยจึงได้กำหนดเกณฑ์ในบริบทนี้ ดังนี้

คณิตศาสตร์

ความยาว 20, 30 ข้อ UA > 1.00

ความยาว 40 ข้อ UA > 1.20

ภาษาอังกฤษ UA > .50

3. <u>ดัชนี α_{MH}</u>

ตารางที่ 60 และ 61 ในภาคผนวก ง แสดงค่าต่ำสุดและค่าสูงสุดของค่าเฉลี่ยรายข้อของดัชนี α_{MH} มีค่าต่ำกว่า และสูงกว่า 1.00 จำแนกตามความยาวแบบสอบและขนาดผู้สอบ ซึ่งพบว่าในกรณีที่ $\alpha_{MH} < 1.00$ ค่าต่ำสุดของค่าเฉลี่ยรายข้อมีค่าสูงขึ้นเมื่อจำนวนผู้สอบมากขึ้น ค่าที่สูงขึ้นค่อนข้างจะใกล้เคียงกันที่ขนาดผู้สอบ 800 และ 1,000 คน ยกเว้นที่วิชาภาษาอังกฤษ ความยาวแบบสอบ 60 ข้อ

เมื่อค่า $\alpha_{MH} > 1.00$ ดังตารางที่ 61 ในภาคผนวก ง พบว่าค่า α_{MH} มีอิทธิพลจากขนาดผู้สอบเล็กน้อยเช่นเดียวกับค่าดัชนี SA และ UA แต่ในทางตรงข้ามคือจะพบว่าที่ขนาดผู้สอบ 100 และ 200 พิสัยของค่าเฉลี่ยจะสูงกว่าที่ขนาดอื่น ๆ ทั้ง 2 วิชา โดยมีค่าเฉลี่ยสูงสุดสูงกว่าที่ขนาดอื่น ๆ ด้วย อย่างไรก็ตามไม่พบว่าที่ความยาวแบบสอบ 40 ข้อ ในวิชาคณิตศาสตร์มีค่าเฉลี่ยที่แตกต่างจากที่พบในความยาวแบบสอบอื่น ๆ เหมือนเช่นที่พบในดัชนี SA และ UA ทั้งนี้ได้รวมสรุปค่าเฉลี่ยเมื่อรวมทุกข้อทุกความยาวแบบสอบและทุกขนาดผู้สอบไว้ในตารางที่ 10 เมื่อ $\alpha_{MH} < 1.00$ และในตารางที่ 11 เมื่อ $\alpha_{MH} > 1.00$

จากตารางที่ 10 จะเห็นว่าค่าเฉลี่ยรวมทุกข้อของ α_{MH} ทั้งวิชาคณิตศาสตร์ และวิชาภาษาอังกฤษมีค่าเท่ากันทุกความยาวแบบสอบ ยกเว้นที่ขนาดผู้สอบ 100 คน ซึ่งมีค่าเฉลี่ยรวมทุกข้อในวิชาคณิตศาสตร์เท่ากับ .77 ในขณะที่ค่าเฉลี่ยในวิชาภาษาอังกฤษเท่ากับ .76 เท่ากันทุกความยาวแบบสอบ นอกจากค่าเฉลี่ยจะเท่ากันแล้ว ส่วนเบี่ยงเบนมาตรฐานมีค่าทั้งที่เท่ากันและใกล้เคียงกันมากเช่นกัน

อย่างไรก็ตามพบว่าค่าเฉลี่ยของ α_{MH} เมื่อ α_{MH} มีค่าน้อยกว่า 1.00 มีอิทธิพลจากขนาดผู้สอบโดยที่ค่าที่ได้จะเพิ่มขึ้นเล็กน้อยเมื่อขนาดผู้สอบเพิ่มขึ้น

และเมื่อค่า $\alpha_{MH} > 1.00$ พบผลเช่นเดียวกับที่พบ เมื่อค่า $\alpha_{MH} < 1.00$ โดยที่ค่าเฉลี่ยในแต่ละความยาวแบบสอบมีค่าใกล้เคียงกันมาก ทั้งในวิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ เช่นที่ขนาดผู้สอบ 100 คน ค่าเฉลี่ยของค่า α_{MH} มีค่าระหว่าง 1.36-1.40 ที่ขนาดผู้สอบ 200 คน มีค่าระหว่าง 1.23-1.24 เป็นต้น รายละเอียดดังตารางที่ 11

ตารางที่ 10 ค่าเฉลี่ยรวมทุกข้อของดัชนี α_{KR} จากผู้สอบเพศเดียวกัน วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ เมื่อความยาวแบบสอบต่างกัน ในขนาดผู้สอบเดียวกัน (กรณี α_{KR} < 1.00)

วิชา	ขนาดผู้สอบ											
ความยาวแบบสอบ	100		200		400		600		800		1,000	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
คณิตศาสตร์												
20 ข้อ	.77	.15	.83	.11	.88	.08	.90	.07	.91	.06	.92	.05
30 ข้อ	.77	.15	.83	.12	.88	.08	.90	.07	.91	.06	.92	.06
40 ข้อ	.77	.15	.83	.11	.88	.08	.90	.07	.91	.06	.92	.06
ภาษาอังกฤษ												
50 ข้อ	.76	.15	.83	.11	.88	.08	.90	.08	.91	.01	.92	.05
60 ข้อ	.76	.16	.83	.11	.88	.08	.90	.08	.91	.01	.92	.05
70 ข้อ	.76	.16	.83	.11	.88	.08	.90	.08	.91	.01	.92	.05
80 ข้อ	.76	.16	.83	.12	.88	.08	.90	.08	.91	.01	.92	.05

ตารางที่ 11 ค่าเฉลี่ยรวมทุกข้อของดัชนี α_{KR} จากผู้สอบเพศเดียวกัน วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษเมื่อความยาวแบบสอบต่างกัน ในขนาดผู้สอบเดียวกัน (กรณี $\alpha_{KR} > 1.00$)

วิชา	ขนาดผู้สอบ											
ความยาวแบบสอบ	100		200		400		600		800		1,000	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
คณิตศาสตร์												
20 ข้อ	1.36	.34	1.23	.21	1.15	.12	1.12	.10	1.10	.08	1.09	.07
30 ข้อ	1.37	.36	1.24	.21	1.15	.13	1.12	.10	1.10	.08	1.09	.07
40 ข้อ	1.38	.37	1.23	.20	1.15	.12	1.12	.09	1.10	.08	1.09	.07
ภาษาอังกฤษ												
50 ข้อ	1.38	.37	1.23	.20	1.15	.12	1.12	1.12	1.10	.08	1.09	.07
60 ข้อ	1.39	.39	1.23	.20	1.15	.12	1.12	1.12	1.10	.08	1.09	.07
70 ข้อ	1.39	.37	1.24	.20	1.15	.12	1.12	1.12	1.10	.08	1.09	.07
80 ข้อ	1.40	.41	1.24	.21	1.15	.12	1.12	.09	1.10	.08	1.09	.07

นอกจากนั้นเมื่อพิจารณาค่าเฉลี่ยข้ามขนาดผู้สอบ พบว่าเมื่อจำนวนผู้สอบเพิ่มขึ้นค่าเฉลี่ยของ α_{MH} จะต่ำลง

ผลการวิเคราะห์ตารางทั้งในกรณี $\alpha_{MH} < 1.00$ และ $\alpha_{MH} > 1.00$ พบว่าเมื่อ $\alpha_{MH} > 1.00$ ระยะของการเบี่ยงเบนของค่าเฉลี่ยออกจาก 1.00 จะกว้างกว่าเมื่อ $\alpha_{MH} < 1.00$ เช่นเมื่อ $\alpha_{MH} < 1.00$ ค่าเฉลี่ยของ α_{MH} เบี่ยงเบนออกจากค่า 1.00 สูงสุดคือ 1.00-.76 = .24 ในขณะที่เมื่อค่า $\alpha_{MH} > 1.00$ ค่าเฉลี่ยของ α_{MH} เบี่ยงเบนออกจาก 1.00 สูงสุดเท่ากับ .40 ดังนั้นการวิเคราะห์ค่าเฉลี่ยรวมทุกข้อและทุกความยาวแบบสอบจำแนกตามขนาดผู้สอบและรวมทุกข้อ ทุกขนาดผู้สอบจำแนกตามความยาวแบบสอบ รวมทั้งรวมทุกข้อทุกขนาดผู้สอบและทุกความยาวแบบสอบจึงใช้ผลสรุปจากค่าเฉลี่ย เมื่อ $\alpha_{MH} > 1.00$ ดังตารางที่ 12

ตารางที่ 12 ค่าเฉลี่ยรวมทุกข้อของดัชนี α_{HH} จากผู้สอบเพศเดียวกัน วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ จำแนกตามขนาดผู้สอบและความยาวแบบสอบ

วิชา	ขนาดผู้สอบ ความยาวแบบสอบ	$\bar{X}$	S.D.	S.E.	ค่าสูงสุด*
คณิตศาสตร์	ขนาดผู้สอบ				
	100 คน	1.37	.08	.01	1.40
	200 คน	1.23	.04	.00	1.23
	400 คน	1.15	.02	00	1.15
	600 คน	1.12	.02	00	1.12
	800 คน	1.10	.01	.00	1.10
	1,000 คน	1.09	.01	.00	1.09
	รวม	1.18	.01	.044	1.29
ภาษาอังกฤษ	100 คน	1.39	.08	.01	1.42
	200 คน	1.24	.04	.00	1.24
	400 คน	1.15	.02	.00	1.15
	600 คน	1.12	.02	.00	1.12
	800 คน	1.10	.01	.00	1.00
	1,000 คน	1.09	.01	.00	1.09
	รวม	1.18	.06	.024	1.24

ตารางที่ 12 (ต่อ)

วิชา	ขนาดผู้สอบ ความยาวแบบสอบ	$\bar{X}$	S.D.	S.E.	ค่าสูงสุด*
คณิตศาสตร์	ความยาวแบบสอบ				
	20 ข้อ	1.18	.10	.01	1.21
	30 ข้อ	1.18	.11	.01	1.21
	40 ข้อ	1.18	.11	.01	1.21
	รวม	1.18	.11	.00	1.18
ภาษาอังกฤษ	50 ข้อ	1.18	.11	.01	1.21
	60 ข้อ	1.18	.11	.01	1.21
	70 ข้อ	1.18	.11	.01	1.21
	80 ข้อ	1.18	.12	.01	1.21
	รวม	1.18	.11	.00	1.18

* = $\bar{X} \pm 2.58$ (S.E.)

การกำหนดเกณฑ์ตัดสินข้อสอบล่าเอียงของดัชนี α_{HH}

3.1 การกำหนดเกณฑ์โดยไม่พิจารณาถึงความแตกต่างข้ามขนาดผู้สอบและความยาวแบบสอบ

จากตารางที่ 12 ในยอดรวมทั้ง 4 แถว พบว่าค่าเฉลี่ยสูงสุดสำหรับวิชาคณิตศาสตร์ คือ 1.29 และสำหรับวิชาภาษาอังกฤษ คือ 1.24 ซึ่งเป็นค่าที่เบี่ยงเบนออกจาก 1.00 ประมาณ .30 ดังนั้น ผู้วิจัยได้พิจารณากำหนดเกณฑ์จากค่าเฉลี่ยสูงสุดคือ $\alpha_{HH} > 1.30$ ในกรณี $\alpha_{HH} > 1.00$ และ $\alpha_{HH} < .70$ กรณี $\alpha_{HH} < .1.00$ หรือกำหนดเกณฑ์ตัดสินความล่าเอียงเมื่อ $.70 > \alpha > 1.30$

3.2 การกำหนดเกณฑ์โดยพิจารณาความแตกต่างข้ามขนาดผู้สอบ และความยาวแบบสอบ

เมื่อพิจารณาค่าเฉลี่ยของดัชนี α_{MH} ตามขนาดผู้สอบและความยาวแบบสอบ ไม่พบอิทธิพลจากความยาวแบบสอบ ในขณะที่ขนาดผู้สอบมีอิทธิพลต่อค่าเฉลี่ยที่แตกต่างกัน การวิเคราะห์เพื่อกำหนดเกณฑ์ให้ครอบคลุม ทั้งค่าที่ได้จากความยาวแบบสอบและขนาดผู้สอบ ผู้วิจัยจึงขอเสนอค่าเฉลี่ยที่เป็นเกณฑ์เพื่อตัดสินข้อสอบลำเอียง โดยกำหนดค่าร่วมกันทั้งสำหรับวิชาคณิตศาสตร์และวิชาภาษาอังกฤษเนื่องจากค่าเฉลี่ยที่ได้มีค่าทั้งที่เท่ากัน และใกล้เคียงกันมาก ดังนี้

ขนาดผู้สอบ 100 คน	$.50 > \alpha_{MH} > 1.50$
ขนาดผู้สอบ 200 คน	$.60 > \alpha_{MH} > 1.40$
ขนาดผู้สอบ 400-1,000 คน	$.70 > \alpha_{MH} > 1.30$

4. ดัชนี β_{SIB}

ตารางที่ 62 ในภาคผนวก ง แสดงค่าต่ำสุดและสูงสุดของค่าเฉลี่ยรายข้อของดัชนี β_{SIB} ซึ่งพบว่ามีพิสัยแคบกว่าดัชนีอื่น ๆ มาก และพบความคล้ายคลึงกับดัชนี α_{MH} ตรงที่ค่าเฉลี่ยที่ขนาดผู้สอบ 100 และ 200 คน จะสูงกว่าขนาดอื่น ๆ และมีพิสัยกว้างกว่าขนาดอื่น ๆ ส่วนความยาวแบบสอบไม่มีอิทธิพลต่อค่าเฉลี่ยทั้งในวิชาคณิตศาสตร์และวิชาภาษาอังกฤษ ทั้งนี้ได้รวมสรุปค่าเฉลี่ยเมื่อรวมทุกข้อ ทุกความยาวแบบสอบ และทุกขนาดผู้สอบไว้ในตารางที่ 13

จากตารางที่ 13 จะเห็นว่า วิชาคณิตศาสตร์และวิชาภาษาอังกฤษมีค่าเฉลี่ยรวมของค่า β_{SIB} ใกล้เคียงกันทุกความยาวแบบสอบภายในขนาดผู้สอบเดียวกัน ในขณะที่ความยาวแบบสอบไม่มีอิทธิพลต่อค่าเฉลี่ยของค่า β_{SIB} แต่จำนวนผู้สอบมีแนวโน้มจะมีอิทธิพลต่อค่าเฉลี่ย β_{SIB} โดยที่จะเห็นได้ว่าเมื่อขนาดผู้สอบมากขึ้น ค่าเฉลี่ยที่ได้จะมีค่าลดลง ตัวอย่างเช่น วิชาคณิตศาสตร์ ในขนาดผู้สอบ 20 ข้อ ค่าเฉลี่ยเมื่อขนาดผู้สอบ 100 คน มีค่า .06 ที่ขนาด 800 และ 1,000 คน ค่าเฉลี่ยเท่ากับ .02 เช่นเดียวกับค่าที่พบในวิชาภาษาอังกฤษ ขนาดผู้สอบ 50 60 70 และ 80 ข้อ ค่าเฉลี่ยจะมีค่าเท่ากับในขนาดผู้สอบ 800 และ 1,000 คน สำหรับวิชาคณิตศาสตร์ และ 600 800 และ 1,000 คน ในวิชาภาษาอังกฤษ

ตารางที่ 13 ค่าเฉลี่ยรวมทุกข้อของดัชนี B_{SIB} จากผู้สอบเพศเดียวกัน วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษจากผู้สอบเพศเดียวกัน เมื่อความยาวแบบสอบต่างกัน ในขนาดผู้สอบเดียวกัน

วิชา	ขนาดผู้สอบ											
ความยาวแบบสอบ	100		200		400		600		800		1,000	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
คณิตศาสตร์												
20 ข้อ	.06	.01	.05	.00	.03	.00	.03	.00	.02	.00	.02	.00
30 ข้อ	.07	.01	.05	.01	.03	.00	.03	.00	.02	.00	.02	.00
40 ข้อ	.07	.01	.04	.01	.03	.00	.02	.00	.02	.00	.02	.00
ภาษาอังกฤษ												
50 ข้อ	.06	.01	.04	.00	.03	.00	.02	.00	.02	.00	.02	.00
60 ข้อ	.06	.01	.04	.00	.03	.00	.02	.00	.02	.00	.02	.00
70 ข้อ	.06	.01	.04	.00	.03	.00	.02	.00	.02	.00	.02	.00
80 ข้อ	.07	.01	.04	.00	.03	.00	.02	.00	.02	.00	.02	.00

เมื่อวิเคราะห์ค่าเฉลี่ยรวมทุกข้อและทุกความยาวแบบสอบ จำแนกตามขนาดผู้สอบ รวมทุกข้อและทุกขนาดผู้สอบจำแนกตามความยาวแบบสอบ และรวมทุกข้อทุกขนาดผู้สอบและทุกความยาวแบบสอบพบผลดังตารางที่ 14

ตารางที่ 14 ค่าเฉลี่ยรวมทุกข้อขององค์ชนี β_{SIB} จากผู้สอบเพศเดียวกัน วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ จำแนกตามขนาดผู้สอบและความยาวแบบสอบ

วิชา	ขนาดผู้สอบ ความยาวแบบสอบ	$\bar{X}$	S.D.	S.E.	ค่าสูงสุด*
คณิตศาสตร์	ขนาดผู้สอบ				
	100 คน	.07	.01	.00	.07
	200 คน	.05	.01	.00	.05
	400 คน	.03	.00	00	.03
	600 คน	.03	.00	00	.03
	800 คน	.02	.00	.00	.02
	1,000 คน	.02	.00	.00	.02
	รวม	.04	.22	.01	.06
ภาษาอังกฤษ	100 คน	.06	.01	.00	.06
	200 คน	.04	.00	.00	.04
	400 คน	.03	.00	.00	.03
	600 คน	.02	.00	.00	.02
	800 คน	.02	.00	.00	.02
	1,000 คน	.02	.00	.00	.02
	รวม	.03	.03	.01	.05

ตารางที่ 14 (ต่อ)

วิชา	ขนาดผู้สอบ ความยาวแบบสอบ	$\bar{X}$	S.D.	S.E.	ค่าสูงสุด*
คณิตศาสตร์	ความยาวแบบสอบ				
	20 ข้อ	.03	.02	.00	.03
	30 ข้อ	.04	.02	.00	.04
	40 ข้อ	.03	.02	.00	.03
	รวม**	.03	.01	.00	.03
ภาษาอังกฤษ	50 ข้อ	.03	.02	.00	.03
	60 ข้อ	.03	.02	.00	.03
	70 ข้อ	.03	.02	.00	.03
	80 ข้อ	.03	.02	.00	.03
	รวม**	.03	.00	.00	.03

* = $\bar{X} \pm 2.58$ (S.E.)

การกำหนดเกณฑ์คัดสินข้อสอบล่าเอียงของดัชนี β_{SIB}

4.1 การกำหนดเกณฑ์โดยไม่พิจารณาถึงความแตกต่างข้ามขนาดผู้สอบและความยาวแบบสอบ

จากค่าเฉลี่ยรวมทุกข้อทุกความยาวผู้สอบในตารางที่ 14 ในยอดรวมทั้ง 4 แถว พบว่าค่าสูงสุดสำหรับวิชาคณิตศาสตร์คือ .06 และสำหรับวิชาภาษาอังกฤษ คือ .05

4.2 การกำหนดเกณฑ์โดยพิจารณาความแตกต่างข้ามขนาดผู้สอบและความยาวแบบสอบ

จากตารางที่ 14 ไม่พบอิทธิพลอันเนื่องจากความยาวแบบสอบที่มีต่อค่าเฉลี่ยของดัชนี β_{SIB} แต่พบอิทธิพลจากขนาดผู้สอบเช่นเดียวกับ α_{MH} โดยที่มีค่าเฉลี่ยในขนาดผู้สอบ 100 คน สูงสุดทั้ง 2 วิชา และมีลักษณะการกระจายของค่าดัชนีต่ำ เช่นเดียวกับ α_{MH} ซึ่งต่ำกว่า SA และ UA มาก

จากค่าที่นำเสนอทั้งหมด ผู้วิจัยเสนอค่าเฉลี่ยเพื่อเป็นเกณฑ์ดังนี้

ขนาดผู้สอบ 100 คน $\beta_{SIB} > .08$

ขนาดผู้สอบ 200 คน $\beta_{SIB} > .06$

ขนาดผู้สอบ 400 คนขึ้นไป $\beta_{SIB} > .05$

เมื่อสรุปรวมเกณฑ์ของดัชนีทั้ง 4 ตัว ที่พัฒนาขึ้นจากข้อมูลเชิงประจักษ์ทั้งในกรณีใช้ค่าเฉลี่ยรวม โดยไม่คำนึงถึงบริบทด้านความยาวแบบสอบและขนาดผู้สอบ และโดยคำนึงถึงบริบทด้านนี้ด้วยได้ผลดังตารางที่ 15

ตารางที่ 15 สรุปเกณฑ์การตัดสินความลำเอียงของข้อสอบทั้งที่ไม่ได้พิจารณา (1) และพิจารณา (2) ถึงความแตกต่างข้ามขนาดผู้สอบ และความยาวแบบสอบ

ดัชนี	วิชา	เกณฑ์ (1)	เกณฑ์ (2)
SA	คณิตศาสตร์	$\|SA\| > .75$	ความยาวแบบสอบ 20-30 ข้อ SA > .70 ความยาวแบบสอบ 40 ข้อ SA > .80
	ภาษาอังกฤษ	$\|SA\| > .35$	SA > .40
UA	คณิตศาสตร์	UA > 1.20	ความยาวแบบสอบ 20-30 ข้อ UA > 1.00 ความยาวแบบสอบ 40 ข้อ UA > 1.20
	ภาษาอังกฤษ	UA > .50	UA > .50
α_{MH}	คณิตศาสตร์ ภาษาอังกฤษ	$.70 > \alpha_{MH} > 1.30$	ขนาดผู้สอบ 100 คน $.50 > \alpha_{MH} > 1.50$ ขนาดผู้สอบ 200 คน $.60 > \alpha_{MH} > 1.40$ ขนาดผู้สอบ 400-1,000 คน $.70 > \alpha_{MH} > 1.30$
β_{SIB}	คณิตศาสตร์ ภาษาอังกฤษ	$\|\beta_{SIB}\| > .06$	ขนาดผู้สอบ 100 คน $\|\beta_{SIB}\| > .08$ ขนาดผู้สอบ 200 คน $\|\beta_{SIB}\| > .06$ ขนาดผู้สอบ 400 คนขึ้นไป $\|\beta_{SIB}\| > .05$

ตอนที่ 2 **ผลการตัดสินข้อสอบลำเอียงทางเพศในวิชาคณิตศาสตร์และวิชาภาษาอังกฤษด้วยเกณฑ์ที่พัฒนาจากข้อมูลเชิงประจักษ์**

ก. ค่าสถิติที่เกี่ยวข้องของข้อมูลที่นำมาใช้ในการวิเคราะห์ความลำเอียงของข้อสอบ

ข้อมูลที่นำมาใช้ในการวิเคราะห์ความลำเอียงด้วยวิธี IRT-2 พารามิเตอร์ MH และ SIBTEST มีค่าสถิติที่เกี่ยวข้อง ดังนี้

1. คะแนนดิบ

ผลการวิเคราะห์ค่าเฉลี่ยของคะแนนดิบของผู้สอบหญิงและชายในวิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ สำหรับแบบสอบขนาด 20 30 40 50 60 70 และ 80 ข้อ จำแนกตามจำนวนผู้สอบ 100 200 400 600 800 และ 1,000 คน ตามลำดับ พบว่าผลจากวิชาคณิตศาสตร์ ผู้สอบชายได้คะแนนเฉลี่ยสูงกว่า และมีการกระจายของคะแนนสูงกว่าผู้สอบหญิงเล็กน้อย โดยที่คะแนนเฉลี่ยของผู้สอบหญิงมีค่าระหว่าง 6.29-11.89 และของผู้สอบชายมีค่าระหว่าง 7.49-13.38 ส่วนผลที่ได้จากวิชาภาษาอังกฤษแสดงว่าผู้สอบหญิงและชายมีคะแนนเฉลี่ยไม่ต่างกันมากนัก โดยที่กลุ่มผู้สอบหญิงมีคะแนนเฉลี่ยสูงกว่าและมีการกระจายของคะแนนสูงกว่า ผู้สอบชายเล็กน้อยคะแนนเฉลี่ยของกลุ่มผู้สอบหญิงมีค่าระหว่าง 23.13-36.23 และคะแนนเฉลี่ยของกลุ่มผู้สอบชายมีค่าระหว่าง 22.21-34.95

2. ค่าพารามิเตอร์ความสามารถ (θ) เมื่อพิจารณาค่าเฉลี่ยของค่าพารามิเตอร์ความสามารถ (θ) ที่ได้จากการวิเคราะห์ตามทฤษฎีการตอบข้อสอบ พบว่าค่าเฉลี่ยของค่าพารามิเตอร์ ความสามารถในวิชาคณิตศาสตร์ พบว่าส่วนมากมีค่าใกล้เคียงกัน มีค่าความสามารถของผู้สอบระหว่าง .00-.04 ทั้งหญิงและชาย ส่วนค่าความสามารถของผู้สอบหญิงและชายสำหรับวิชาภาษาอังกฤษมีค่าใกล้เคียงกันเช่นกัน โดยมีค่า θ ระหว่าง -1.873 และ 3.447 สำหรับผู้สอบหญิงและระหว่าง -1.721 และ 3.291 สำหรับผู้สอบชาย และมีค่าเฉลี่ยของ θ ระหว่าง -.02 และ .08 และระหว่าง -.10 และ .00 ตามลำดับ

3. ค่าพารามิเตอร์ความยาก (b)

ผลการวิเคราะห์ค่าพารามิเตอร์ ความยาก (b) จากการวิเคราะห์ด้วยทฤษฎีการตอบข้อสอบ พบว่าค่าพารามิเตอร์ความยากเฉลี่ยในวิชาคณิตศาสตร์จากกลุ่มผู้สอบหญิงสูงกว่าที่ได้จากกลุ่มผู้สอบชาย และพบอิทธิพลจากความยาวของแบบสอบที่มีต่อค่านี้โดยที่ในแบบสอบที่ยาวกว่า ค่าเฉลี่ยของค่าพารามิเตอร์ ความยากจะสูงกว่า นอกจากนั้นจำนวนผู้สอบก็มีอิทธิพลต่อค่าพารามิเตอร์ความยากเช่นกัน แต่น้อยกว่าขนาดแบบสอบ ค่าเฉลี่ยของค่าความยากสำหรับผู้สอบหญิงมีค่าระหว่าง0.82-3.14 และสำหรับผู้สอบชายมีค่าระหว่าง .40-2.06 ในขณะที่ในวิชาภาษาอังกฤษ สำหรับแบบสอบแต่ละความยาว และขนาดผู้สอบแต่ละขนาดมีค่าแตกต่างกันโดยค่าที่

ได้จากผู้สอบกลุ่มใหญ่กว่า จะมีค่าสูงกว่าที่ได้จากผู้สอบกลุ่มเล็กกว่า และเมื่อจำนวนข้อสอบในแบบสอบมากกว่าค่าพารามิเตอร์ ความยากก็มีค่าสูงกว่าเช่นกัน ทั้งนี้พบว่าค่าความยากเฉลี่ยของแบบสอบ สำหรับกลุ่มผู้สอบหญิงมีค่าต่ำกว่าที่ได้จากกลุ่มผู้สอบชาย ยกเว้นที่ขนาดความยาวแบบสอบ 50 ข้อ ขนาดผู้สอบ 100 คน ที่ค่าเฉลี่ยของค่าพารามิเตอร์ความยากจากกลุ่มผู้สอบชายมีค่าต่ำกว่าเล็กน้อย โดยที่ค่าเฉลี่ยของค่าความยากสำหรับผู้สอบหญิงมีค่าระหว่าง .08-.94 และสำหรับผู้สอบชายมีค่าระหว่าง .07-1.16

4. ค่าพารามิเตอร์อำนาจจำแนก (a)

ผลการวิเคราะห์ค่าเฉลี่ยของค่าพารามิเตอร์อำนาจจำแนกของข้อสอบนั้น พบว่าสำหรับวิชาคณิตศาสตร์ค่าเฉลี่ยที่ได้จากผู้สอบชายจะสูงกว่าที่ได้จากผู้สอบหญิงเล็กน้อย และค่าเฉลี่ย สำหรับขนาดแบบสอบ 40 ข้อ จะต่ำกว่าขนาด 20 และ 30 ข้อ ซึ่งมีค่าใกล้เคียงกับค่าเฉลี่ยของค่าพารามิเตอร์อำนาจจำแนกจากผู้สอบหญิงมีค่าระหว่าง .24-.70 และจากผู้สอบชายมีค่าระหว่าง .33-.70 สำหรับวิชาภาษาอังกฤษเมื่อใช้แบบสอบที่มีความยาวข้อสอบ 80 ข้อ ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ เมื่อจำนวนผู้สอบต่างกันมีค่าไม่แตกต่างกันมากนัก แต่เมื่อใช้แบบสอบขนาด 70 60 และ 50 ข้อ พบว่าค่าที่ได้จากผู้สอบจำนวนมากกว่าจะมีค่าต่ำกว่าที่ได้จากผู้สอบจำนวนน้อยกว่า และเป็นลักษณะเช่นนี้ทั้งสำหรับผู้สอบหญิงและชาย นอกจากนั้นพบว่าโดยเฉลี่ยแล้วแบบสอบแต่ละความยาว มีอำนาจจำแนกสำหรับผู้สอบหญิงและชายใกล้เคียงกัน โดยที่ค่าเฉลี่ยของค่าพารามิเตอร์อำนาจจำแนกผู้สอบหญิงมีค่าระหว่าง .45-.87 สำหรับผู้สอบชายมีค่าระหว่าง .42-.85

ข. ผลการวิเคราะห์ความลำเอียงของข้อสอบโดยใช้การทดสอบนัยสำคัญของค่าสถิติ

ผลการวิเคราะห์ความลำเอียงด้วยการทดสอบความมีนัยสำคัญของค่าสถิติตามวิธีทฤษฎีการตอบข้อสอบชนิด 2 พารามิเตอร์ วิธี MH และวิธี SIBTEST (ได้นำเสนอตัวอย่างตารางไว้ในตารางที่ 63 และในภาคผนวก จ) ตามความยาวของแบบสอบ 20 30 40 50 60 70 และ 80 ข้อ โดยจำแนกตามขนาดผู้สอบ 100 200 400 600 800 และ 1,000 คน ตามลำดับ ได้นำมาสรุปให้เห็นเฉพาะข้อสอบที่ลำเอียงในตารางที่ 64-70 ในภาคผนวก จ ผลการตัดสินความลำเอียงของข้อสอบมีรายละเอียดดังนี้

1. วิชาคณิตศาสตร์

ผลการวิเคราะห์ความลำเอียงของข้อสอบวิชาคณิตศาสตร์ด้วยการทดสอบความมีนัยสำคัญของค่าสถิติโดยภาพรวมพบว่าเมื่อขนาดผู้สอบเพิ่มขึ้น จำนวนข้อสอบลำเอียงจากวิธี IRT มีจำนวนเพิ่มขึ้น และมีจำนวนมากกว่าที่พบจากวิธี MH และวิธี SIBTEST ทั้งที่ระดับ $\alpha \leq .05$ และ $\alpha \leq .01$ ยกเว้นกรณีที่ขนาดผู้สอบ 100 และ 200 ที่ความยาวแบบสอบ 20 และ 40 จำนวนข้อสอบลำเอียงใกล้เคียงกับ วิธี MH และวิธี SIBTEST และที่ขนาดผู้สอบ 400 คน ความยาวแบบสอบ 40 ข้อ มีจำนวนน้อยข้อสอบลำอียงกว่าอีก 2 วิธี ส่วนวิธี MH และวิธี SIBTEST มีจำนวนข้อสอบลำเอียงใกล้เคียงกัน โดยที่วิธี SIBTEST มีข้อสอบลำเอียงมากกว่าเล็กน้อยในขนาดผู้สอบ 1,000 คน ดังสรุปผลเฉพาะจำนวนข้อสอบลำเอียงที่ในตารางที่ 16 โดยที่พบว่าจำนวนข้อสอบลำเอียงที่พบจากวิธี IRT มีจำนวน 0-19 และ 0-15 ข้อ วิธี MH มีจำนวน 1-6 และ 0-4 ข้อ และวิธี SIBTEST มีจำนวน 0-8 และ 0-3 ข้อ ที่ระดับ $\alpha \leq .05$ และ $\alpha \leq .01$ ตามลำดับ

2. วิชาภาษาอังกฤษ

เมื่อมีจำนวนผู้สอบ 100 และ 200 คน พบว่าจำนวนข้อสอบลำเอียงตามวิธี IRT มีจำนวนน้อยกว่าวิธี MH และวิธี SIBTEST เล็กน้อย แต่เมื่อจำนวนผู้สอบมากขึ้น จำนวนข้อสอบลำเอียงจากวิธี IRT จะมีจำนวนมากกว่าอีก 2 วิธีมากโดยเฉพาะอย่างยิ่งเมื่อจำนวนผู้สอบเพิ่มเป็น 800 และ 1,000 คน จะเห็นว่าวิธีตรวจสอบความลำเอียงของข้อสอบทั้ง 3 วิธี มีอิทธิพลจากขนาดผู้สอบโดยที่เมื่อจำนวนผู้สอบมากขึ้น จำนวนข้อสอบลำเอียงจะเพิ่มขึ้น และจะเห็นได้ชัดเมื่อกำหนดระดับนัยสำคัญที่ $\alpha \leq .05$ มากกว่าที่ $\alpha \leq .01$

เมื่อพิจารณาอิทธิพลด้านความยาวของแบบสอบต่อจำนวนข้อสอบลำเอียงไม่พบว่ามีความแตกต่างอย่างเป็นระบบ สำหรับวิธี IRT แต่วิธี MH และ SIBTEST พบข้อสอบลำเอียงจากแบบสอบที่มีความยาว 50 ข้อ น้อยกว่าขนาด 70 และ 80 เล็กน้อย ในกรณีที่ขนาดผู้สอบ 600 คนขึ้นไป อย่างไรก็ตามไม่พบอิทธิพลดังกล่าวในกรณีที่กำหนดระดับนัยสำคัญที่ $\alpha \leq .01$

สำหรับจำนวนข้อสอบลำเอียงที่พบจากวิธี IRT มีตั้งแต่ 2-44 ข้อ ที่ระดับ $\alpha \leq .05$ 0-29 ข้อที่ $\alpha \leq .01$ วิธี MH ตั้งแต่ 3-17 ข้อ และ 0-6 ข้อ ส่วนวิธี SIBTEST ตั้งแต่ 2-18 ข้อ และ 0-7 ข้อ ตามลำดับ ดังสรุปผลเฉพาะข้อสอบลำเอียงไว้ในตารางที่ 17

ตารางที่ 16 สรุปจำนวนข้อสอบลำเอียงวิชาคณิตศาสตร์จากการทดสอบค่าสถิติตามวิธี IRT-2 พารามิเตอร์ MH และ SIBTEST จำแนกตามขนาดผู้สอบและความยาวแบบสอบ

ขนาดผู้สอบ (คน)	ความยาวแบบสอบ (ข้อ)	วิธีวิเคราะห์					
		IRT		MH		SIBTEST	
		จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
100	20	2	10.00	1	5.00	2	10.00
		-	0.00	1	5.00	1	5.00
	30	2	6.67	2	6.67	1	3.33
		-	0.00	-	0.00	-	0.00
	40	-	0.00	2	5.00	3	7.50
		-	0.00	-	0.00	1	2.50
200	20	2	10.00	1	5.00	7	0.00
		-	0.00	-	0.00	-	0.00
	30	2	6.67	2	6.67	2	6.67
		-	0.00	1	3.33	-	0.00
	40	2	5.00	1	2.50	2	5.00
		-	0.00	1	2.50	-	0.00
400	20	10	50.00	2	10.00	3	15.00
		5	25.00	1	5.00	1	5.00
	30	4	14.33	5	16.67	4	13.33
		1	3.33	2	6.67	3	10.00
	40	3	7.50	6	15.00	6	15.00
		1	2.50	4	10.00	2	5.00

ตารางที่ 16 (ต่อ)

ขนาดผู้สอบ	ความยาวแบบสอบ	วิธีวิเคราะห์					
		IRT		MH		SIBTEST	
		จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
600	20	10	50.00	3	15.00	2	10.00
		6	30.00	2	10.00	-	0.00
	30	14	46.67	4	13.33	2	6.67
		5	16.67	1	3.33	-	0.00
	40	15	37.50	3	7.50	4	10.00
		5	12.50	-	0.00	1	2.50
800	20	16	80.00	5	25.00	4	20.00
		15	75.00	-	0.00	2	10.00
	30	14	46.67	5	16.67	7	23.33
		8	26.67	-	0.00	1	3.33
	40	17	42.50	6	15.00	8	20.00
		9	22.50	-	0.00	2	5.00
1,000	20	17	85.00	1	5.00	2	10.00
		11	55.00	1	5.00	1	5.00
	30	17	56.67	2	6.67	4	13.33
		11	36.67	1	3.33	1	3.33
	40	19	47.50	2	5.00	7	17.50
		13	32.50	2	5.00	2	5.00

ตารางที่ 17 สรุปจำนวนข้อสอบลำเอียงวิชาภาษาอังกฤษจากการทดสอบค่าสถิติตามวิธี IRT-2 พารามิเตอร์ MH และ SIBTEST จำแนกตามขนาดผู้สอบและความยาวแบบสอบ

ขนาดผู้สอบ	ความยาวแบบสอบ (ข้อ)	วิธีวิเคราะห์					
		IRT-2		MH		SIBTEST	
		จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
100	50	4	8.00	3	6.00	5	10.00
		-	0.00	-	0.00	-	0.00
	60	2	3.33	4	6.67	3	5.00
		-	0.00	1	1.67	3	5.00
	70	5	7.14	7	10.00	2	2.86
		-	0.00	2	2.86	1	1.43
	80	2	2.50	4	5.00	10	12.50
		-	0.00	-	0.00	3	3.75
200	50	2	4.00	4	8.00	7	14.00
		-	0.00	3	6.00	4	8.00
	60	2	3.33	5	8.33	7	11.67
		-	0.00	2	3.33	5	8.33
	70	2	2.86	4	5.71	6	8.57
		-	0.00	-	0.00	2	2.86
	80	4	5.00	5	6.25	4	5.00
		-	0.00	-	0.00	1	1.25

ตารางที่ 17 (ต่อ)

ขนาดผู้สอบ	ความยาวแบบสอบ (ข้อ)	วิธีวิเคราะห์					
		IRT-2		MH		SIBTEST	
		จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
400	50	26	52.00	6	12.00	6	12.00
		17	34.00	3	6.00	2	4.00
	60	9	15.00	7	11.67	7	11.67
		5	8.33	2	3.33	2	3.33
	70	12	17.14	6	8.57	9	12.86
		6	8.57	3	4.28	3	4.28
	80	9	11.25	7	8.75	9	11.25
		6	7.50	3	3.75	3	3.75
600	50	12	24.00	7	14.00	8	16.00
		6	12.00	4	8.00	3	6.00
	60	9	15.00	7	11.67	7	11.67
		5	8.33	3	5.00	3	5.00
	70	10	14.28	9	12.86	8	11.43
		6	8.57	3	4.28	4	5.71
	80	13	16.25	9	11.25	13	16.25
		6	7.50	5	6.25	5	6.25

ตารางที่ 17 (ต่อ)

ขนาดผู้สอบ	ความยาวแบบสอบ (ข้อ)	วิธีวิเคราะห์					
		IRT-2		MH		SIBTEST	
		จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
800	50	40	80.00	9	18.00	11	22.00
		25	50.00	4	8.00	6	12.00
	60	34	56.67	9	15.00	14	23.33
		22	36.67	5	8.33	5	8.33
	70	44	62.86	11	15.71	14	20.00
		29	41.43	6	8.57	6	8.57
	80	12	15.00	10	12.50	15	18.75
		4	5.00	4	5.00	5	6.25
1,000	50	44	88.00	14	28.00	12	24.00
		29	58.00	4	8.00	5	10.00
	60	20	33.33	14	21.67	17	28.33
		12	20.00	4	5.00	5	8.33
	70	38	54.28	17	24.28	18	25.71
		25	35.71	5	7.14	5	7.14
	80	26	32.50	17	21.25	17	21.25
		12	15.00	5	6.25	7	8.75

* ตัวเลขแถวบนได้จาก $\alpha \leq .05$ แถวล่างได้จาก $\alpha \leq .01$

ค. ผลการตัดสินข้อสอบลำเอียงเมื่อใช้เกณฑ์จากข้อมูลเชิงประจักษ์

ตารางที่ 63 ในภาคผนวก จ แสดงตัวอย่างของค่าดัชนีที่ได้จากการวิเคราะห์ความลำเอียงของข้อสอบรายข้อ และนำมาสรุปให้เห็นเฉพาะข้อสอบที่ลำเอียง ดังตารางที่ 71-77 ในภาคผนวก ฉ สำหรับผลจากดัชนี SA UA และตารางที่ 78-84 ในภาคผนวก ฉ เช่นกัน สำหรับดัชนี α_{MH} และ β_{SIB} ซึ่งผู้วิจัยได้นำมาสรุปแสดงจำนวนข้อสอบลำเอียงในแต่ละวิธีไว้ในตารางที่ 18-21 ดังนี้

1. ผลจากดัชนี SA

ตารางที่ 18 แสดงจำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้ดัชนี SA และใช้เกณฑ์ตัดสินจากข้อมูลเชิงประจักษ์ ซึ่งพบว่าร้อยละของข้อสอบลำเอียงจะเพิ่มขึ้นเมื่อขนาดตัวอย่างและจำนวนข้อสอบในแบบสอบเพิ่มขึ้นเป็นส่วนใหญ่ ร้อยละของข้อสอบลำเอียงในแบบสอบยาว 20 ข้อ เรียงตามลำดับจากขนาดผู้สอบ 100 จนถึง 1,000 คน คือ 10.00 5.00 25.00 15.00 30.00 และ 40.00 ในแบบสอบยาว 30 ข้อ มีจำนวนร้อยละ 16.67 6.67 33.33 36.67 43.33 และ 46.67 ในขณะที่ในความยาวแบบสอบ 40 ข้อ มีจำนวนร้อยละ 62.50 40.00 50.00 70.00 57.00 และ 60.00 เมื่อใช้เกณฑ์ตัดสิน $|SA| > .60$

สำหรับวิชาภาษาอังกฤษมีลักษณะเช่นเดียวกับวิชาคณิตศาสตร์ โดยพบว่ามีข้อสอบลำเอียงร้อยละ 38.00 18.00 32.00 24.00 34.00 และ 30.00 สำหรับความยาวแบบสอบ 50 ข้อ ร้อยละ 25.00 16.67 25.00 20.00 28.33 และ 13.33 สำหรับความยาวแบบสอบ 60 ข้อ ร้อยละ 22.86 21.43 28.57 28.57 41.43 และ 34.28 สำหรับความยาวแบบสอบ 70 ข้อ และร้อยละ 50.50 40.00 42.50 37.50 38.75 และ 33.75 สำหรับแบบสอบยาว 80 ข้อ

อย่างไรก็ตามเมื่อใช้เกณฑ์ $|SA| > .70$ ที่ความยาวแบบสอบ 20-30 ข้อ $|SA| > .80$ ที่ความยาวแบบสอบ 40 ข้อ และ $|SA| > .75$ สำหรับความยาวแบบสอบ 50-80 ข้อ พบว่าจำนวนข้อสอบลำเอียงในความยาวแบบสอบเพิ่มขึ้น 1 ข้อ ที่ความยาวแบบสอบ 20 ข้อ ขนาดผู้สอบ 100 และ 400 คน และที่ความยาวแบบสอบ 30 ข้อ ขนาดผู้สอบ 200 คนเพิ่มขึ้น 2 ข้อ ที่ความยาวแบบสอบ 30 ข้อ ขนาดผู้สอบ 1,000 คน ส่วนในความยาวแบบสอบ 40 มีจำนวนข้อลดลง 1 ข้อ ที่ขนาดผู้สอบ 100, 200, 400 และ 1,000 คน และลดลง 2 ข้อ ที่ขนาดผู้สอบ 800 คน รายละเอียดดังตารางที่ 18

ตารางที่ 18 จำนวนและร้อยละของข้อสอบลำเอียง จากดัชนี SA เมื่อใช้เกณฑ์จากข้อมูลเชิงประจักษ์

ขนาดผู้สอบ (คน)	คณิตศาสตร์ ความยาวแบบสอบ (ข้อ)			ภาษาอังกฤษ ความยาวแบบสอบ (ข้อ)			
	20	30	40	50	60	70	80
100	2:10.00 3:15.00	5:16.67 5:16.66	24:60.00 23:57.50	19:38.00 13:16.00	15:25.00 12:20.00	16:22.86 16:22.85	40:50.50 36:45.00
200	1:5.00 1:5.00	3:10.00 4:13.33	17:42.50 16:40.00	9:18.00 7:14.00	10:16.67 8:13.33	15:21.43 9:10.00	49:61.25 29:36.25
400	5:25.00 6:30.00	10:33.33 10:33.33	21:52.50 19:47.50	16:32.00 15:30.00	15:25.00 14:23.33	2o:28.57 16:22.85	34:42.50 33:41.25
600	3:15.00 3:15.00	11:36.67 12:40.00	26:65.00 26:65.00	12:24.00 9:18.00	12:20.00 6:10.00	20:28.57 16:22.85	30:37.50 26:32.50
800	6:30.00 6:30.00	13:43.33 13:43.33	23:57.50 21:52.50	12:24.00 12:24.00	12:20.00 14:23.33	20:28.57 22:31.42	30:37.50 27:33.75
1,000	8:40.00 8:40.00	14:46.67 16:53.33	28:70.00 27:67.50	15:30.00 12:24.00	8:13.33 7:11.66	24:34.28 18:25.71	27:33.75 17:21.25

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

2. ผลจากดัชนี UA

สรุปผลการวิเคราะห์จากดัชนี UA ในตารางที่ 19 พบว่าเมื่อใช้เกณฑ์ที่ไม่ได้พิจารณาถึงความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ ซึ่งพบข้อสอบลำเอียงทั้งที่น้อยกว่าและมากกว่าที่พบในการวิเคราะห์จากดัชนี SA อย่างไรก็ตามพบว่าในขนาดแบบสอบ 100 คน ความยาวแบบสอบ 20 และ 30 ข้อ ไม่พบว่ามีข้อสอบลำเอียง เมื่อเรียงลำดับตามขนาดผู้สอบตั้งแต่ 100 คนขึ้นไป พบข้อสอบลำเอียงในความยาวแบบสอบ 20 ข้อ ร้อยละ 0.00 5.00 35.00 35.00 25.00 และ 30.00 ในความยาวแบบสอบ 30 ข้อ พบข้อสอบลำเอียงร้อยละ 0.00 6.67 13.33 40.00 33.33 และ 40.00 และในความยาวแบบสอบ 40 ข้อ พบข้อสอบลำเอียงร้อยละ 62.50 40.00 50.00 70.00 57.50 และ 60.00 ตามลำดับ

จะเห็นว่าในกรณีของวิชาคณิตศาสตร์นี้ พบข้อสอบลำเอียงในจำนวนครึ่งหนึ่งของความยาวแบบสอบขึ้นไปถึง 10 กรณี ในความยาวแบบสอบ 40 เป็นผลจากการวิเคราะห์ด้วยดัชนี SA และ UA อย่างละ 5 กรณี ในความยาวแบบสอบ 40 ข้อ ในขณะที่ผลการวิเคราะห์ในวิชาภาษาอังกฤษพบอยู่ 6 กรณี เป็นผลจากการวิเคราะห์ด้วยดัชนี UA ทั้งหมดที่ความยาวแบบสอบ 70 และ 80 ข้อ

จำนวนข้อสอบลำเอียงในวิชาภาษาอังกฤษเรียงตามลำดับขนาดผู้สอบจาก 180 คน คิดเป็นร้อยละ 34.00 30.00 60.00 34.00 58.00 และ 64.00 สำหรับความยาวแบบสอบ 50 ข้อ ร้อยละ 23.33 23.33 25.00 28.33 46.67 และ 25.00 สำหรับความยาวแบบสอบ 60 ข้อ ร้อยละ 25.71 30.00 27.14 35.71 51.43 และ 52.87 สำหรับความยาวแบบสอบ 70 ข้อ และร้อยละ 62.50 61.25 51.25 51.20 46.25 และ 47.25 สำหรับแบบสอบยาว 80 ข้อ

เมื่อใช้เกณฑ์ที่พิจารณาถึงความแตกต่างข้ามความยาวแบบสอบ และขนาดผู้สอบสอบแล้ว พบว่าวิชาภาษาอังกฤษยังใช้เกณฑ์เดิม แต่วิชาคณิตศาสตร์จะมีข้อสอบลำเอียงเปลี่ยนแปลงเล็กน้อย โดยมีข้อสอบลำเอียงในความยาวแบบสอบ 20 และ 30 ข้อ เพิ่มขึ้น โดยมีจำนวนข้อสอบลำเอียงร้อยละ 00.00 10.00 50.00 40.00 40.00 และ 45.00 สำหรับแบบสอบยาว 20 ข้อ และร้อยละ 10.00 13.33 26.26 43.33 46.67 และ 50.00 ตามลำดับ

ตารางที่ 19 จำนวนและร้อยละของข้อสอบลำเอียง จากดัชนี UA เมื่อใช้เกณฑ์จากข้อมูลเชิงประจักษ์

ขนาดผู้สอบ (คน)	คณิตศาสตร์ ความยาวแบบสอบ (ข้อ)			ภาษาอังกฤษ ความยาวแบบสอบ (ข้อ)			
	20	30	40	50	60	70	80
100	0:0.00 0:0.00	0: 0.00 3:10.00	25:62.50 25:62.50	17:34.00 17:34.00	14:23.33 14:23.33	18:25.71 18:25.71	50:62.50 50:62.50
200	1: 5.00 2:10.00	3:10.00 4:13.33	16:40.00 16:40.00	15:30.00 15:30.00	14:23,33 14:23.33	21:30.00 21:30.00	49:61.25 49:61.25
400	7:35.00 10:50.00	4:13.33 8:26.26	20:50.00 20:50.00	30:60.00 30:60.00	15:25.00 15:25.00	19:27.14 19:27.14	41:51.25 41:52.25
600	7:35.00 8:40.00	12:40.00 13:43.33	28:70.00 28:70.00	17:34.00 17:34.00	17:28.33 17:28.33	25:35.71 25:35.71	41:51.25 41:51.25
800	5:25.00 8:40.00	10:33.33 14:46.66	23:57.50 23:57.50	29:58.00 29:58.00	28:46.67 28:46.67	36:51.43 36:51.43	37:46.25 37:46.25
1,000	6:30.00 9:45.00	12:40.00 15:50.00	24:60.00 24:60.00	32:64.00 32:64.00	15:25.00 15:25.00	37:52.87 37:52.86	38:47.50 38:47.50

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ
ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

3. ผลจากดัชนี α_{MH} และ β_{SIB}

เนื่องจากผลการตัดสินข้อสอบลำเอียงทั้ง 2 วิธี มีความใกล้เคียงกันมาก ผู้วิจัยจึงได้รายงานผลไปพร้อม ๆ กัน ดังนี้

เมื่อใช้เกณฑ์ $.70 > \alpha_{MH} > 1.30$ และ $\left|\beta_{SIB}\right| > .06$ ทั้ง 2 วิชา ได้เสนอผลสรุปไว้ในตารางที่ 20 และ 21 ซึ่งพบว่าในวิชาคณิตศาสตร์ จำนวนข้อสอบลำเอียงที่พบระหว่างการใช้ดัชนีทั้งสอง มีทั้งที่มีค่าใกล้เคียงและแตกต่างกัน แต่ต่างก็มีจำนวนลดลงเมื่อขนาดผู้สอบเพิ่มขึ้นและเมื่อความยาวแบบสอบเพิ่มขึ้น ซึ่งในกรณีหลังนี้พบในขนาดผู้สอบไม่เกิน 400 คน โดยที่ผลจากการใช้ดัชนี α_{MH} ค้นพบข้อสอบลำเอียงในความยาวแบบสอบ 20 ข้อ สำหรับขนาดผู้สอบ 6 ขนาด ระหว่างร้อยละ 0.00-30.00 ขนาดผู้สอบที่พบข้อสอบลำเอียงน้อยที่สุดคือ 1,000 คน และมากที่สุดคือขนาด 100 คน ในความยาวแบบสอบ 30 ข้อ ไม่พบข้อสอบลำเอียงที่ขนาดผู้สอบ 1,000 คนเช่นกัน และพบว่ามีข้อสอบลำเอียงมากที่สุดขนาดผู้สอบ 100 คน คิดเป็นร้อยละ 33.33 และในความยาวแบบสอบ 40 ข้อ ไม่พบข้อสอบลำเอียงที่ขนาดผู้สอบ 1,000 คน แต่พบข้อสอบลำเอียงมากที่สุดในขนาดผู้สอบ 100 คน ทั้งนี้มีร้อยละของข้อสอบลำเอียงจากผลการใช้ดัชนี α_{MH} เรียงลำดับจากขนาดผู้สอบ 100 คน คิดเป็นร้อยละ 30.00 20.00 15.00 10.00 10.00 และ 0.00 สำหรับความยาวแบบสอบ 20 ข้อ ร้อยละ 33.33 20.00 20.00 10.00 3.33 และ 0.00 สำหรับความยาวแบบสอบ 30 ข้อ และร้อยละ 40.00 22.50 22.50 5.00 5.00 และ 00.00 สำหรับความยาวแบบสอบ 40 ข้อ

ส่วนผลการใช้ดัชนี β_{SIB} สำหรับวิชาคณิตศาสตร์มีข้อสอบลำเอียงคิดเป็นร้อยละ 60.00 20.00 30.00 10.00 20.00 และ 5.00 สำหรับความยาวแบบสอบ 20 ข้อ ร้อยละ 43.33 36.67 23.33 10.00 23.33 และ 3.33 สำหรับความยาวแบบสอบ 30 ข้อ และร้อยละ 57.50 35.00 20.00 10.00 10.00 และ 7.50

สำหรับผลการใช้ดัชนี α_{MH} และ β_{SIB} ในวิชาภาษาอังกฤษ พบว่าเมื่อใช้ดัชนี α_{MH} วิเคราะห์ที่ความยาวแบบสอบ 50 ข้อ จำนวนข้อสอบลำเอียงตามขนาดผู้สอบมีค่าระหว่างร้อยละ 8.00-38.00 และระหว่าง 5.00-43.33 สำหรับความยาวแบบสอบ 60 ข้อ ระหว่าง 2.86-45.71 สำหรับความยาวแบบสอบ 60 ข้อ ระหว่าง 2.86-45.71 สำหรับความยาวแบบสอบ 70 ข้อ และระหว่าง 3.75-42.50 สำหรับแบบสอบยาว 80 ข้อ และเมื่อใช้ดัชนี β_{SIB} พบข้อสอบลำเอียงคิดเป็นร้อยละ 8.00-38.00 3.33-40.00 2.86-40.00 และ 3.75-50.00 ตามลำดับความยาวแบบสอบ 50-80 ข้อซึ่งมีค่าใกล้เคียงกับ α_{MH}

เมื่อจำแนกร้อยละของข้อสอบลำเอียงตามขนาดผู้สอบตั้งแต่ 100 คนขึ้นไป พบว่าร้อยละของข้อสอบลำเอียงเมื่อใช้ดัชนี α_{MH} มีค่าร้อยละ 38.00 34.00 16.00 10.00 8.00 และ 8.00 สำหรับความยาวแบบสอบ 50 ข้อ ร้อยละ 43.33 26.67 13.33 8.33 8.33 และ 5.00 สำหรับความยาวแบบสอบ 60 ข้อ ร้อยละ 51.42 27.14 11.43 11.43 4.28 และ 2.85 สำหรับความยาวแบบสอบ 70 ข้อ ร้อยละ 43.75 30.00 11.25 10.00 3.75 และ 5.00 สำหรับความยาวแบบสอบ 80 ข้อ

เมื่อใช้ดัชนี β_{SIB} สำหรับวิชาภาษาอังกฤษ พบข้อสอบลำเอียงร้อยละ 36.00 38.00 20.00 12.00 12.00 และ 8.00 ในความยาวแบบสอบ 50 ข้อ ร้อยละ 38.33 18.33 8.33 10.00 และ 5.00 ในความยาวแบบสอบ 60 ข้อ ร้อยละ 45.71 28.57 20.00 8.57 7.14 และ 2.86 ในความยาวแบบสอบ 70 ข้อ และ 50.00 25.00 17.50 10.00 5.00 และ 3.75 ในความยาวแบบสอบ 80 ข้อ รายละเอียดดังตารางที่ 19 และ 20

อย่างไรก็ตามเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบด้วย พบว่าที่ขนาดผู้สอบ 100 และ 200 คน ผลการใช้ดัชนี α_{MH} มีจำนวนข้อสอบลำเอียงลดลง โดยผลจากการใช้ดัชนี α_{MH} มีข้อสอบลำเอียงคิดเป็นร้อยละ 15.00 เท่ากัน เมื่อขนาดผู้สอบเท่ากับ 100 และ 200 คน ในความยาวแบบสอบ 20 ข้อ และร้อยละ 13.33 และ 6.66 ในความยาวแบบสอบ 30 ข้อ ร้อยละ 22.50 เท่ากันในความยาวแบบสอบ 40 ข้อ ร้อยละ 18.00 และ 6.25 ในความยาวแบบสอบ 50 ข้อ ร้อยละ 23.33 และ 20.00 ในความยาวแบบสอบ 60 ข้อ ร้อยละ 22.85 และ 17.14 ในความยาวแบบสอบ 70 ข้อ และร้อยละ 12.50 และ 20.00 ในความยาวแบบสอบ 80 ข้อ

สำหรับผลจากดัชนี β_{SIB} เมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบพบว่าจำนวนข้อสอบลำเอียงในขนาดผู้สอบ 100 คน ลดลงแต่ในขนาดผู้สอบ 400 คนขึ้นไป จนเพิ่มขึ้นโดยมีค่าร้อยละของข้อสอบลำเอียงจำแนกตามขนาดผู้สอบตั้งแต่ 100 คน ในแต่ละความยาวแบบสอบ ดังนี้ ร้อยละ 55.00 20.00 30.00 10.00 10.00 และ 10.00 สำหรับความยาวแบบสอบ 20 ข้อ ร้อยละ 33.33 36.67 26.66 20.00 30.00 และ 16.66 สำหรับความยาวแบบสอบ 30 ข้อ ร้อยละ 27.50 35.00 35.00 15.00 20.00 และ 15.00 สำหรับความยาวแบบสอบ 40 ข้อ ร้อยละ 28.00 38.00 26.00 18.00 14.00 และ 12.00 สำหรับความยาวแบบสอบ 50 ข้อ ร้อยละ 28.33 40.00 26.66 18.33 13.33 และ 13.33 สำหรับความยาวแบบสอบ 60 ข้อ ร้อยละ 35.71 25.71 21.42 17.41 14.28 และ 10.00 สำหรับความยาวแบบสอบ 70 ข้อ ร้อยละ 32.50 25.00 23.75 16.25 12.50 และ 11.25 สำหรับความยาวแบบสอบ 80 ข้อ

เมื่อพิจารณาในภาพรวมแล้ว พบว่าการใช้เกณฑ์โดยพิจารณาถึงความแตกต่างในขนาดผู้สอบนั้น ร้อยละของข้อสอบลำเอียงที่พบในขนาดผู้สอบต่างกัน ส่วนค่าใกล้เคียงกันมากกว่าเมื่อใช้เกณฑ์โดยไม่คำนึงถึงความแตกต่างในบริบททั้งสอง

เมื่อเปรียบเทียบผลจากดัชนีทั้ง 4 ตัว จะพบว่าในขณะที่ผลของการใช้ดัชนี SA และ UA ค่อนข้างจะมีความแตกต่างกันเป็นส่วนมากในด้านจำนวนข้อสอบลำเอียงที่พบ และข้อสอบลำเอียงจะมีจำนวนมากขึ้นเมื่อความยาวของแบบสอบเพิ่มขึ้น โดยเฉพาะอย่างยิ่งเห็นได้ชัดในความยาวแบบ 80 ข้อ และขนาดผู้สอบ 100-600 คน ผลการใช้ดัชนี α_{MH} และ β_{SIB} ในวิชาภาษาอังกฤษให้ผลการวิเคราะห์ในด้านจำนวนข้อสอบลำเอียงใกล้เคียงกันแทบทุกขนาดผู้สอบ 100 และ 200 คนเช่นกัน แต่พบว่าจำนวนข้อสอบลำเอียงจะมีความคงที่มากขึ้นที่ขนาดผู้สอบตั้งแต่ 400 คนขึ้นไป

ตารางที่ 20 จำนวนและร้อยละของข้อสอบลำเอียง จากดัชนี α_{mH} เมื่อใช้เกณฑ์จากข้อมูลเชิงประจักษ์

ขนาดผู้สอบ (คน)	คณิตศาสตร์ ความยาวแบบสอบ (ข้อ)			ภาษาอังกฤษ ความยาวแบบสอบ (ข้อ)			
	20	30	40	50	60	70	80
100	6:30.00 3:15.00	10:33.33 4:13.33	16:40.00 7:17.50	19:38.00 9:18.00	26:43.33 13:21.67	36:51.42 17:24.28	35:43.75 11:13.75
200	4:20.00 3:15.00	6:20.00 2: 6.66	9:22.50 3: 7.50	17:34.00 8:6.25	16:26.67 11:18.33	19:27.14 10:14.28	24:30.00 15:18.75
400	3:15.00 3:15.00	6:20.00 6:20.00	9:22.50 9:22.50	8:16.00 8:16.00	8:13.33 8:13.33	8:11.43 8:11.43	9:11.25 9:11.25
600	2:10.00 2:10.00	3:10.00 3:10.00	2:5.00 2:5.00	5:10.00 5:10.00	5:8.33 5:8.33	8:11.43 8:11.43	8:10.00 8:10.00
800	2::10.00 2:10.00	1:3.33 1:3.53	2:5.00 2:5.00	4:8.00 4:8.00	5:8.33 5:8.33	3:4.28 3:4.28	3:3.75 3:3.75
1,000	0:0.00 0:0.00	0:0.00 0:0.00	0:0.00 0:0.00	4:8.00 4:8.00	3:5.00 3:5.00	2:2.85 2:2.85	4:5.00 4:5.00

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตารางที่ 21 จำนวนและร้อยละของข้อสอบลำเอียง จากดัชนี B_{SIB} เมื่อใช้เกณฑ์จากข้อมูลเชิงประจักษ์

ขนาดผู้สอบ (คน)	คณิตศาสตร์ ความยาวแบบสอบ (ข้อ)			ภาษาอังกฤษ ความยาวแบบสอบ (ข้อ)			
	20	30	40	50	60	70	80
100	12:60.00 11:55.00	13:43.33 10:33.33	23:57.50 12:30.00	18:36.00 14:28.00	23:38.33 17:28.33	32:45.71 29:35.71	40:50.00 26:32.50
200	4:20.00 4:20.00	11:36.67 11:36.67	14:35.00 14:35.00	19:38.00 19:38.00	23:38.33 23:38.33	20:28.57 20:28.57	20:25.00 20:25.00
400	6:30.00 6:30.00	7:23.33 9:30.00	8:20.00 14:35.00	10:20.00 13:26.00	11:18.33 17:28.33	14:20.00 16:22.86	14:17.50 21:26.25
600	2:10.00 2:10.00	3:10.00 6:20.00	4:10.00 6:15.00	6:12.00 9:18.00	5: 8.33 11:18.33	6: 8.57 12:17.14	8:10.00 13:16.25
800	4:20.00 4:10.00	7:23.33 9:30.00	4:10.00 8:20.00	6:12.00 7:14.00	6:10.00 9:15.00	5: 7.14 12:17.14	4: 5.00 10:12.50
1,000	1: 5.00 2:10.00	1: 3.33 5:16.66	3: 7.50 6:15.00	4: 8.00 6:12.00	2: 5.00 8:13.33	2: 2.85 7:10.00	3: 3.75 9:11.25

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ
ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตอนที่ 3 ความสอดคล้องในผลการตัดสินข้อสอบลำเอียงเมื่อใช้การทดสอบความมีนัยสำคัญทางสถิติ และใช้เกณฑ์ที่พัฒนาจากข้อมูลเชิงประจักษ์

ก. ความสอดคล้องในการตัดสินข้อสอบลำเอียงเมื่อใช้การทดสอบนัยสำคัญของค่าสถิติ

1. ความสอดคล้องภายในวิธีเดียวกัน

ตารางที่ 22-24 แสดงจำนวนและร้อยละของข้อสอบลำเอียงที่สอดคล้องกันภายในวิธี IRT-2 พารามิเตอร์ MH และ SIBTEST เมื่อจำนวนผู้สอบต่างกัน พบว่า วิธี IRT-2 พารามิเตอร์ มีความสอดคล้องสูงกว่าวิธี MH และ SIBTEST เมื่อขนาดผู้สอบมากขึ้น เช่นในความยาวแบบสอบ 20 ข้อ ตั้งแต่ขนาดผู้สอบ 400 และ 600 คน ร้อยละของความสอดคล้องมีค่า 42.85 และมีค่าสูงขึ้นในขนาดผู้สอบเพิ่มขึ้น และพบผลเช่นเดียวกันในความยาวแบบสอบ 30 และ 40 ข้อ โดยมีร้อยละของความสอดคล้องตั้งแต่ร้อยละ 40.00 ระหว่างขนาดผู้สอบ 600 และ 800 คนขึ้นไป

ส่วนความสอดคล้องภายในวิธี MH และวิธี IRT-2 พารามิเตอร์ มีค่าไม่ค่อยคงที่ไม่เป็นระบบพอที่สรุปได้ อาจเป็นเพราะมีจำนวนข้อสอบลำเอียงน้อยกว่าวิธี IRT-2 พารามิเตอร์ และบางขนาดผู้สอบไม่พบว่ามีข้อสอบลำเอียง สำหรับร้อยละของความสอดคล้องภายในวิธี IRT-2 พารามิเตอร์ มีค่าระหว่าง 9.09-94.00 0.00-63.50 และ 6.25-63.00 สำหรับความยาวแบบสอบ 20 30 และ 40 ข้อ วิธี MH มีความสอดคล้องระหว่างร้อยละ 14.00-100.00 0.00-40.00 0.00-25.00 ส่วนวิธี SIBTEST มีค่าระหว่าง 20.00-40.00 0.00-25.00 และ 10.00-30.00 ตามลำดับ

ตารางที่ 22 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในวิธีวิเคราะห์เดียวกัน เมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาคณิตศาสตร์ ความยาวแบบสอบ 20 ข้อ

จำนวนผู้สอบ	IRT-2	MH	SIBTEST
100-200	-	(1,100.00)	-
100-400	(2,20.00 %)	(1,50.00)	(1,25.00)
100-600	-	-	-
100-800	(2,12.50)	(1,20.00)	(1,20.00)
100-1,000	(2,11.76)	(1,100.00)	(1,33.33)
200-400	(1,9.09)	(1,50.00)	-
200-600	(1,9.90)	-	-
200-800	(2,12.50)	(1,20.00)	-
200-1,000	(2,11.76)	(1,100.00)	-
400-600	(6,42.85)	-	-
400-800	(9,52.94)	(2,40.00)	(2,40.00)
400-1,000	(9,50.00)	(1,50.00)	-
600-800	(8,44.44)	(1,14.28)	-
600-1,000	(9,50.00)	-	-
800-1,000	(16,94.11)	(1,20.00)	-

ตารางที่ 23 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในวิธีวิเคราะห์เดียวกัน เมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาคณิตศาสตร์ ความยาวแบบสอบ 30 ข้อ

จำนวนผู้สอบ	IRT-2	MH	SIBTEST
100-200	-	(1,33.33)	-
100-400	(2,50.00)	(2,40.00)	(1,25.00)
100-600	(0, 0.00)	(0, 0.00)	(0, 0.00)
100-800	(1, 6.67)	(2,40.00)	(1,14.28)
100-1,000	(1, 5.55)	(1,33.33)	(0, 0.00)
200-400	(0, 0.00)	(1,16.67)	(0, 0.00)
200-600	(0, 0.00)	(0, 0.00)	(1,33.33)
200-800	(2,14.28)	(1,16.67)	(1,12.50)
200-1,000	(2,11.76)	(1,33.33)	(0, 0.00)
400-600	(2,12.50)	(0, 0.00)	(0, 0.00)
400-800	(2,12.50)	(3,42.86)	(2,22.22)
400-1,000	(2,10.52)	(1,16.67)	(1,14.28)
600-800	(8,40.00)	(1,12.50	(0, 0.00)
600-1,000	(9,40.90)	(0, 0.00)	(0, 0.00)
800-1,000	(12,63.15)	(1,16.67)	(1,10.00)

ตารางที่ 24 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในวิธีวิเคราะห์เดียวกัน เมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาคณิตศาสตร์ ความยาวแบบสอบ 40 ข้อ

จำนวนผู้สอบ	IRT-2	MH	SIBTEST
100-200	-	-	-
100-400	-	(1,14.28)	(1,12.50)
100-600	-	-	-
100-800	-	(1,14.28)	(1,10.00)
100-1,000	-	-	-
200-400	(1,25.00)	-	-
200-600	(1,6.25)	-	-
200-800	(2,11.76)	-	(1,11.11)
200-1,000	(2,10.52)	-	-
400-600	(2,12.50)	-	-
400-800	(2,11.11)	(4,50.00)	(2,16.67)
400-1,000	(2,10.00)	(1,14.28)	(3,30.00)
600-800	(11,52.38)	(1,11.11)	-
600-1,000	(10,45.83)	-	-
800-1,000	(14,63.63)	(1,14.28)	-

สำหรับวิชาภาษาอังกฤษซึ่งสรุปผลไว้ในตารางที่ 25-28 พบว่า ในภาพรวมแล้ว ร้อยละของความสอดคล้องเมื่อขนาดผู้สอบต่างกันมีแนวโน้มที่ลดลงจากที่พบในวิชาคณิตศาสตร์ แต่ยังพบว่าส่วนใหญ่ที่ความสอดคล้องระหว่างผู้สอบจำนวนมากขึ้น ความสอดคล้องจะสูงขึ้นเช่นกัน ทั้งนี้มีค่าร้อยละของความสอดคล้องภายในวิธี IRT-2 พารามิเตอร์ ระหว่าง 4.51-50.00 2.85-31.00 4.54-57.69 และ 6.25-22.58 สำหรับความยาวแบบสอบ 50 60 70 และ 80 ข้อ ความสอดคล้องภายในวิธี MH มีค่าระหว่าง 8.33-30.00 6.25-27.27 6.66-30.00 และ 7.69-30.00 ส่วนวิธี SIBTEST มีค่าระหว่าง 5.58-36.36 10.53-34.78 5.26-28.50 และ 4.17-28.00 ตามลำดับ รายละเอียดดังตารางที่ 25-28

ตารางที่ 25 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในวิธีวิเคราะห์เดียวกัน เมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาภาษาอังกฤษ ความยาวแบบสอบ 50 ข้อ

จำนวนผู้สอบ	IRT-2	MH	SIBTEST
100-200	-	-	-
100-400	(2, 7.14)	(1, 12.50)	(1, 10.00)
100-600	(1, 6.66)	(1, 11.11)	(1, 8.33)
100-800	(3, 7.31)	(1, 9.09)	(1, 6.67)
100-1,000	(3, 6.66)	(2, 13.33)	(2, 13.33)
200-400	(2, 7.69)	(2, 25.00)	(2, 18.18)
200-600	(2, 16.66)	(2, 22.22)	(4, 36.36)
200-800	(2, 5.00)	(1, 8.33)	(1, 5.88)
200-1,000	(2, 4.54)	(2, 12.50)	(4, 26.67)
400-600	(4, 11.36)	(3, 30.00)	(3, 27.27)
400-800	(22, 50.00)	(2, 15.38)	(3, 21.43)
400-1,000	(24, 52.17)	(3, 17.64)	(3, 20.00)
600-800	(8, 18.18)	(3, 23.07)	(4, 26.67)
600-1,000	(12, 27.27)	(4, 23.52)	(4, 25.00)
800-1,000	(32, 61.53)	(3, 15.00)	(5, 27.78)

ตารางที่ 26 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในวิธีวิเคราะห์เดียวกัน เมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาภาษาอังกฤษ ความยาวแบบสอบ 60 ข้อ

จำนวนผู้สอบ	IRT-2	MH	SIBTEST
100-200	-	-	(1, 11.11)
100-400	(1, 10.00)	(2, 22.22)	(1, 11.11)
100-600	(1, 10.00)	(1, 10.00)	(1, 11.11)
100-800	(1, 2.85)	(1, 8.33)	(1, 6.25)
100-1,000	(1, 4.76)	(2, 6.25)	(2, 11.11)
200-400	-	(1, 9.09)	(2, 16.67)
200-600	(1, 10.00)	(1, 9.09)	(3, 27.27)
200-800	(1, 2.85)	(1, 7.69)	(2, 10.53)
200-1,000	(1, 4.76)	(2, 12.50)	(3, 14.28)
400-600	(3, 20.00)	(3, 27.27)	(2, 16.67)
400-800	(5, 13.15)	(2, 14.28)	(4, 23.53)
400-1,000	(4, 31.81)	(3, 17.64)	(3, 14.78)
600-800	(5, 13.15)	(3, 23.07)	(3, 16.67)
600-1,000	(4, 16.00)	(4, 25.00)	(3, 14.28)
800-1,000	(12, 28.57)	(3, 15.78)	(8, 34.78)

ตารางที่ 27 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในวิธีวิเคราะห์เดียวกัน เมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาภาษาอังกฤษ ความยาวแบบสอบ 70 ข้อ

จำนวนผู้สอบ	IRT-2	MH	SIBTEST
100-200	(1, 16.66)	-	-
100-400	-	(2, 18.18)	-
100-600	(2, 15.38)	(1, 6.66)	-
100-800	(3, 6.52)	(2, 12.50)	-
100-1,000	(3, 7.50)	(3, 14.28)	(4, 5.26)
200-400	-	(1, 11.11)	(2, 15.38)
200-600	(1, 9.09)	(3, 30.00)	(3, 27.27)
200-800	(2, 4.54)	(1, 7.14)	(2, 11.11)
200-1,000	(2, 5.26)	(3, 16.66)	(2, 9.09)
400-600	(3, 15.78)	(2, 15.38)	(3, 21.43)
400-800	(9, 19.14)	(3, 21.42)	(3, 15.00)
400-1,000	(11, 28.20)	(3, 15.00)	(4, 17.39)
600-800	(7, 14.89)	(3, 17.64)	(3, 15.79)
600-1,000	(9, 23.07)	(5, 23.80)	(5, 23.81)
800-1,000	(30, 57.69)	(3, 16.66)	(7, 28.50)

ตารางที่ 28 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในวิธีวิเคราะห์เดียวกัน เมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาภาษาอังกฤษ ความยาวแบบสอบ 80 ข้อ

จำนวนผู้สอบ	IRT-2	MH	SIBTEST
100-200	(1, 20.00)	-	-
100-400	-	(1, 10.00)	(2, 11.76)
100-600	(1, 7.14)	(1, 8.33)	(1, 4.54)
100-800	-	(1, 7.69)	(1, 4.17)
100-1,000	(2, 7.69)	(2, 10.52)	(3, 12.50)
200-400	(1, 8.33)	(2, 20.00)	(1, 8.33)
200-600	(1, 6.25)	(2, 16.66)	(2, 13.33)
200-800	(2, 14.28)	(1, 7.14)	(2, 11.76)
200-1,000	(4, 15.38)	(2, 10.00)	(2, 10.53)
400-600	(3, 15.78)	(3, 23.07)	(4, 22.22)
400-800	(2, 10.52)	(3, 21.42)	(3, 14.28)
400-1,000	(8, 29.62)	(3, 14.28)	(3, 13.04)
600-800	(3, 13.63)	(4, 26.66)	(7, 33.33)
600-1,000	(5, 14.70)	(6, 30.00)	(6, 25.00)
800-1,000	(7, 22.58)	(3, 12.50)	(7, 28.00)

2. ความสอดคล้องระหว่างวิธีต่างกัน

ตารางที่ 29 และ 30 แสดงจำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันระหว่างวิธี IRT-2 พารามิเตอร์ วิธี MH วิธี IRT-2 พารามิเตอร์ กับวิธี SIBTEST วิธี MH กับวิธี SIBTEST รายละเอียดดังนี้

ตารางที่ 29 แสดงร้อยละของความสอดคล้องในการค้นพบข้อสอบลำเอียงระหว่างวิธี IRT และ MH วิธี IRT และ SIBTEST และวิธี MH และ SIBTEST แยกตามความยาวของแบบสอบ 20 30 40 ข้อ และขนาดผู้สอบ 6 ขนาด พบว่าเมื่อขนาดผู้สอบตั้งแต่ 600 คนขึ้นไป ร้อยละของความสอดคล้องของการค้นพบข้อสอบลำเอียงระหว่างวิธี MH และวิธี SIBTEST จะสูงกว่าวิธีอื่น ยกเว้นที่ขนาดผู้สอบ 600 คน ความยาวแบบสอบ 40 ข้อ ที่พบว่าร้อยละของความสอดคล้องระหว่างวิธี IRT และวิธี SIBTEST สูงกว่าเล็กน้อย ในส่วนที่ขนาดผู้สอบ 100 และ 400 คนนั้น วิธีที่สอดคล้องกันสูงสุดเมื่อความยาวของแบบสอบเท่ากับ 40 ข้อ คือ วิธี MH และวิธี SIBTEST เช่นกัน ส่วนในความยาวข้อสอบ 20 พบว่าวิธี IRT และวิธี SIBTEST สอดคล้องกันสูงกว่าวิธีอื่นส่วนในความยาวข้อสอบ 30 ข้อ มีทั้งระหว่าง IRT และวิธี MH และวิธี IRT และวิธี SIBTEST ที่สูงกว่าวิธีอื่น สำหรับกรณีที่มีผู้สอบขนาด 200 คน พบแต่ความสอดคล้องระหว่างวิธี MH และวิธี IRT ในความยาวแบบสอบ 30 และ 40 ข้อ

ตารางที่ 30 แสดงร้อยละของความสอดคล้องในการค้นพบข้อสอบลำเอียงต่างวิธีกันในวิชาภาษาอังกฤษ แยกตามความยาวข้อสอบและขนาดผู้สอบ ซึ่งพบว่านอกจากที่ขนาดผู้สอบ 100 คน ที่พบความสอดคล้องระหว่างวิธี IRT และวิธี MH สูงกว่าวิธีอื่น ๆ เมื่อความยาวข้อสอบเท่ากับ 50 และ 80 ข้อแล้ว ในขนาดผู้สอบอื่น ๆ ตั้งแต่ 200 คนขึ้นไปพบความสอดคล้องระหว่างวิธี MH และวิธี SIBTEST สูงกว่าวิธีอื่นแทบทุกความยาวแบบสอบ ยกเว้นที่ความยาว 50 ข้อ ขนาดผู้สอบ 200 คน ที่มีความสอดคล้องระหว่างวิธี IRT และวิธี SIBTEST สูงกว่าวิธีอื่น

ตารางที่ 29 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันระหว่างวิธีวิเคราะห์ต่างกัน จำแนกตามความยาวแบบสอบ และขนาดกลุ่มตัวอย่าง (วิชาคณิตศาสตร์)

ขนาดผู้สอบ ขนาดผู้สอบ (คน)	ความยาวแบบสอบ (ข้อ)								
	20			30			40		
	IRT-MH	IRT-SIB	MH-SIB	IRT-MH	IRT-SIB	MH-SIB	IRT-MH	IRT-SIB	MH-SIB
100	1/2 (50.00)	2/2 100.00)	1/2 (50.33)	2/2 (100.00)	1/2 (50.00)	1/2 (50.00)	- -	- -	2/3 (66.66
200	0 (00.00)	- -	- -	0/4 (00.00)	0/4 (00.00)	1/3 (33.33)	0/3 (00.00)	0/4 (00.00)	1/2 (50.00
400	2/10 (20.00)	3/10 (30.00)	1/4 (25.00)	2/7 (28.57)	2/6 (33.33)	2/7 (28.57)	0/9 (00.00)	1/8 (12.50)	3/9 (33.33
600	2/11 (18.18)	1/11 (9.09)	1/4 (25.00)	2/16 (12.50)	2/14 (14.28)	1/15 (6.66)	1/17 (5.88)	- -	2/5 (40.00
800	3/18 (16.66)	3/17 (17.64)	3/6 (50.00)	2/17 (11.76)	4/17 (23.52)	3/9 (33.33)	2/21 (9.52)	4/21 (19.04)	2/12 (16.66
1,000	1/17 (5.55)	2/17 (10.53)	1/2 (50.50)	1/18 (5.26)	2/19 (9.52)	2/4 (33.33)	- -	3/23 (11.54)	2/7 (22.22

* แถวบนคือจำนวนข้อสอบลำเอียงที่สอดคล้องกัน/จำนวนข้อสอบลำเอียงรวมทั้งหมดจาก 2 วิธี

ตารางที่ 30 จำนวนและร้อยละของข้อสอบที่ค่าเอียงสอดคล้องกันระหว่างวิธีวิเคราะห์ต่างกัน จำแนกตามความยาวแบบสอบ และขนาดกลุ่มตัวอย่าง (วิชาภาษาอังกฤษ)

ขนาดกลุ่มสอบ (คน)	ความยาวแบบสอบ (ข้อ)											
	50			60			70				80	
	IRT-MH	IRT-SIB	MH-SIB	IRT-MH	IRT-SIB	MH-SIB	IRT-MH	IRT-SIB	MH-SIB	IRT-MH	IRT-SIB	MH-SIB
100	3/4 (75.50)	3/6 (50.00)	3/5 (60.00)	1/5 (20.50)	1/4 (25.50)	3/4 (75.50)	2/10 (20.50)	2/5 (40.50)	1/8 (12.50)	2/4 (50.00)	1/11 (9.09)	2/12 (16.66)
200	1/5 (20.00)	2/7 (28.57)	2/9 (22.22)	2/5 (40.00)	2/7 (28.57)	5/7 (71.43	1/5 (20.00)	1/7 (14.28)	4/6 (66.67)	-	1/7 (12.28)	2/7 (28.57)
400	5/27 (18.52)	5/27 (18.51)	6/6 100.00)	3/13 (23.08)	3/13 (23.08)	6/8 (75.00)	5/13 (38.46)	6/15 (40.00)	6/19 (66.67)	3/13 (23.07)	4/14 (28.57	7/9 (77.77
600	5/14 (35.71)	6/14 (42.85)	6/9 (66.66)	4/12 (33.33)	5/11 (45.45)	4/10 (40.00)	5/14 (35.71)	6/12 (50.00)	7/10 (70.00)	4/18 (22.22)	7/22 (36.84)	8/14 (57.14)
800	9/40 (22.50)	10/41 (24.39)	8/12 (66.66)	8/35 (22.85)	13/35 (37.14)	9/14 (64.28)	8/47 (17.02)	10/48 (20.83)	9/16 (56.25)	6/16 (37.50)	7/20 (35.00)	9/16 (56.25
1,000	12/46 (26.08)	10/46 (21.73)	9/17 (52.94)	8/26 (30.76)	9/20 (55.00)	11/19 (57.89)	9/46 (19.56)	11/45 (24.44)	14/21 (66.66)	12/31 (38.70)	10/33 (30.30)	15/19 (78.94

* แถวบนคือจำนวนข้อสอบค่าเอียงที่สอดคล้องกัน/จำนวนข้อสอบค่าเอียงรวมทั้งหมดจาก 2 วิธี

ข. ความสอดคล้องในการตัดสินข้อสอบลำเอียงโดยใช้เกณฑ์จากข้อมูลเชิงประจักษ์

1. ความสอดคล้องภายในดัชนีเดียวกันเมื่อขนาดผู้สอบต่างกัน

1.1 วิชาคณิตศาสตร์

การตรวจสอบความสอดคล้องของข้อสอบลำเอียงระหว่างขนาดผู้สอบที่ต่างกันภายในแต่ละความยาวแบบสอบ สำหรับวิชาคณิตศาสตร์ เมื่อใช้ดัชนีทั้ง 4 ตัว นำเสนอไว้ในตารางที่ 31-33

1.1.1 ผลจากดัชนี SA และ UA

เมื่อใช้เกณฑ์ที่ไม่ได้พิจารณาถึงความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบพบว่าในความยาวแบบสอบ 20 ข้อ ความสอดคล้องในการค้นพบข้อสอบลำเอียงระหว่างขนาดผู้สอบมีค่าสูงสุดร้อยละ 62.50 ซึ่งเป็นความสอดคล้องระหว่างขนาดผู้สอบ 400 และ 1,000 คน และระหว่างขนาดผู้สอบ 800 และ 1,000 คน และมีความสอดคล้องต่ำสุดร้อยละ 12.50 ระหว่างขนาดผู้สอบ 200 และ 1,000 คน สำหรับผลจากการใช้ดัชนี SA ส่วนผลจากดัชนี UA พบความสอดคล้องสูงสุดร้อยละ 83.33 ระหว่างขนาดผู้สอบ 800 และ 1,000 คน และต่ำสุดร้อยละ 16.67 ระหว่างขนาดผู้สอบ 400 และ 600 คน

เมื่อใช้เกณฑ์ที่พิจารณาถึงความแตกต่างระหว่างความยาวแบบสอบและขนาดผู้สอบพบว่า ค่าร้อยละของความสอดคล้องไม่เปลี่ยนไปมากนัก และพบว่าในจุดที่มีความสอดคล้องสูงอยู่แล้ว คือ ความสอดคล้องระหว่างจำนวนผู้สอบ 400 และ 800 คน 400 และ 1,000 คน และ 800 และ 1,000 คน มีความสอดคล้องสูงขึ้น คิดเป็นร้อยละ 71.43 75.00 และ 75.00 ตามลำดับ ส่วนผลจาก UA นั้น พบว่าส่วนมากมีค่าร้อยละของความสอดคล้องสูงขึ้น ยกเว้นที่ความสอดคล้องระหว่างขนาดผู้สอบ 400 และ 1,000 คน และระหว่าง 800 และ 1,000 คน ที่มีค่าลดลง

เมื่อใช้ความยาวแบบสอบ 30 ข้อ และใช้เกณฑ์ที่ไม่ได้พิจารณาถึงความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ พบว่าความสอดคล้องในการค้นพบข้อสอบลำเอียงระหว่างการใช้ขนาดผู้สอบต่างกันด้วยดัชนี SA มีค่าต่ำสุดร้อยละ 9.09 ระหว่างขนาดผู้สอบ 200 และ 400 คน มีค่าสูงสุดร้อยละ 61.54 ระหว่างขนาดผู้สอบ 600 และ 800 คน ในขณะที่ดัชนี UA ให้ผลสอดคล้องต่ำกว่าเป็นส่วนมาก โดยมีความสอดคล้องต่ำสุด ระหว่างขนาดผู้สอบ 400 กับ 600 คน และ 400 กับ 1,000 คน คิดเป็นร้อยละ 14.28 และมีความสอดคล้องสูงสุดระหว่างผู้สอบขนาด 800 และ 1,000 คน คิดเป็นร้อยละ 69.23

เมื่อใช้เกณฑ์ที่พิจารณาถึงความแตกต่างข้ามความยาวแบบสอบ และขนาดผู้สอบ พบว่าร้อยละของความสอดคล้องภายในดัชนี SA ระหว่างผู้สอบขนาด 100 คน และขนาดอื่น ๆ มีทั้งเท่าเดิมและลดลง เช่นเดียวกับระหว่างขนาด 400 คน และขนาดอื่น ๆ ในขณะที่ค่าระหว่างผู้สอบขนาด 200 คน และขนาดอื่น ๆ มีค่าเพิ่มขึ้นเล็กน้อย ที่เหลือมีค่าคงเดิม ทั้งนี้มีร้อยละของความสอดคล้องระหว่าง 12.50-61.54 ส่วนผลภายในดัชนี UA พบว่าความสอดคล้องเพิ่มขึ้นทุกบริบท ยกเว้นค่าระหว่างผู้สอบ 800 และ 1,000 คน ที่มีค่าลดลง

ที่ความยาวแบบสอบ 40 ข้อ ซึ่งพบข้อสอบลำเอียงมากที่สุดนั้น พบว่ามีความสอดคล้องในการค้นพบข้อสอบลำเอียงสูงสุดระหว่างขนาดผู้สอบต่างกันเช่นกัน สำหรับความสอดคล้องจากดัชนี SA มีค่าต่ำสุดร้อยละ 45.16 ระหว่างขนาดผู้สอบ 200 และ 1,000 คน มีค่าสูงสุดร้อยละ 80.00 ระหว่างขนาดผู้สอบ 600 และ 1,000 คน ส่วนดัชนี UA ให้ผลความสอดคล้องต่ำกว่าเล็กน้อย มีค่าต่ำสุดร้อยละ 25.00 ระหว่างขนาดผู้สอบ 100 คน และ 400 คน และมีค่าสูงสุดร้อยละ 74.07 ระหว่างขนาดผู้สอบ 800 และ 1,000 คน

1.1.2 ผลจากดัชนี α_{MH} และ β_{SIB}

เมื่อพิจารณาความสอดคล้องระหว่างการค้นพบข้อสอบลำเอียงในแต่ละความยาวแบบสอบเมื่อใช้ขนาดผู้สอบต่างกัน พบว่าดัชนี α_{MH} ให้ผลความสอดคล้องระหว่างร้อยละ 0.00-66.67% ร้อยละ 0.00-33.33 และร้อยละ 0.00-21.43 สำหรับแบบสอบยาว 20 30 และ 40 ข้อ ตามลำดับ ส่วนดัชนี β_{SIB} ให้ผลความสอดคล้องระหว่างร้อยละ 0.00-33.33 ร้อยละ 0.00-26.67 และระหว่าง 0.00-25.00 ตามลำดับ เมื่อใช้เกณฑ์ที่ไม่ได้พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ แต่เมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ พบว่าผลจากดัชนี α_{MH} มีร้อยละของความสอดคล้องเท่าเดิมบ้างและลดลงบ้าง มีเพิ่มขึ้นเป็นส่วนน้อยร้อยละของความสอดคล้องในกรณีนี้มีค่าระหว่างร้อยละ 20.00-66.67 0.00-42.86 และ 00.00-22.22 สำหรับดัชนี β_{SIB} พบว่าส่วนใหญ่มีร้อยละของความสอดคล้องเท่าเดิมหรือลดลงในความยาวแบบสอบ 20 ข้อ ในขณะที่ส่วนใหญ่ความสอดคล้องจะเพิ่มขึ้นในความยาวแบบสอบ 30 และ 40 ข้อ โดยมีความสอดคล้องร้อยละ 00.00-50.00 00.00-54.54 และร้อยละ 00.00-29.41 ตามลำดับ

ตารางที่ 31 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในดัชนีเดียวกันเมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาคณิตศาสตร์ ความยาวแบบสอบ 20 ข้อ

จำนวนผู้สอบ	SA	UA	α_{MH}	β_{SIB}
100-200	-	-	(3,37.50)	(1,7.14)
	-	-	(0,20.00	(1,7.14)
100-400	(2,40.00)	-	(3,50.00)	(4,28.57)
	(2,28.57)	-	(2,50.00)	(3,21.43)
100-600	(1,25.00)	-	(0,00.00)	(0,00.00)
	(1,20.00)	-		(0,00.00)
100-800	(2,33.33)	-	(2,33.33)	(4,33.33)
	(2,28.57)	-	(1,25.00)	(3,25.00)
100-1,000	-	-	(0,00.00)	(1,8.33)
	(2,22.22)	-	-	(1,8.33)
200-400	-	-	(3,60.00)	(2,28.57)
	-	(0,0.00)	(2,50.00)	(2,25.00)
200-600	-	-	(0,00.00)	(0,0.00)
	-	(0,0.00)	(0,00.00)	(1,20.00)
200-800	(1,16.67)	-	(2,40.00)	(0,0.00)
	(1,16.67)	(1,11.11)	(1,25.00)	(0,0.00)
200-1,000	(1,12.50)	-	(0,0.00)	(1,33.33)
	(1,12.50)	(1,10.00)	-	(2,50.00)
400-600	(1,14.28)	(2,16.67)	(0,0.00)	(0,0.00)
	(1,12.50)	(6,50.00)	(0,0.00)	(0,0.00)
400-800	(4,52.14)	(3,33.33)	(2,66.67)	(2,25.00)
	(5,71.43)	(6,50.00)	(2,66.67)	(2,25.00)
400-1,000	(5,62.50)	(3,30.30)	(0,0.00)	(1,16.67)
	(6,75.00)	(6,46.15)	-	(1,14.28)

ตารางที่ 31 (ต่อ)

จำนวนผู้สอบ	SA	UA	α_{MH}	β_{SIB}
600-800	(2,28.57) (2,28.50)	(2,20.00) (5,45.45)	(0,0.00)	(0,0.00) (0,0.00)
600-1,000	(2,22.22) (2,22.22)	(4,44.44) (5,41.67)	(0,0.00) -	(0,0.00) (0,0.00)
800-1,000	(6,62.50) (6,75.00)	(5,83.33) (6,54.54)	(0,0.00) -	(0,0.00) (0,0.00)

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ
ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตารางที่ 32 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในดัชนีเดียวกันเมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาคณิตศาสตร์ ความยาวแบบสอบ 30 ข้อ

จำนวนผู้สอบ	SA	UA	α_{MH}	β_{SIB}
100-200	(1,14.28)	-	(2,12.50)	(4,20.00)
	(1,12.50)	(1,16.67)	(0,0.00)	(6,35.29)
100-400	(3,25.00)	-	(4,33.33)	(4,26.67)
	(3,25.00)	(1,10.00)	(3,42.86)	(6,40.00)
100-600	(4,33.33)	-	(2,18.18)	(0,0.00)
	(4,30.77)	(2,14.28)	(0,0.00)	(1,6.67)
100-800	(5,38.46)	-	(1,10.00)	(4,25.00)
	(5,38.46)	(2,13.33)	(0,0.00)	(3,18.75)
100-1,000	(3,18.75)	-	(0,0.00)	(1,7.69)
	(3,16,67)	(1,5.88)	-	(6,54.54)
200-400	(1,9.09)	-	(2,16.67)	(1,6.25)
	(2,16.67)	(1,9.09)	(0,0.00)	(2,11.11)
200-600	(2,18.18)	(1,33.33)	(1,10.00)	(1,7.69)
	(4,33.33)	(2,10.33)	(0,0.00)	(1,6.25)
200-800	(2,15.38)	(2,20.00)	(1,12.50)	(3,20.00)
	(3,21.43)	(3,20.00)	(0,0.00)	(6,37.50)
200-1,000	(2,14.28)	(2,16.67)	(0,0.00)	(1,9.09)
	(2,25.00)	(3,18.75)	-	(3,23.07)
400-600	(4,23.53)	(2,14.28)	(0,0.00)	(0,0.00)
	(4,22.22)	(3,16.67)	(0,0.00)	(1,7.14)
400-800	(8,53.33)	(2,16.67)	(1,16.67)	(2,18.18)
	(8,53.33)	(5,29.41)	(1,16.67)	(3,20.00)
400-1,000	(9,60.00)	(2,14.28)	(0,0.00)	(1,16.67)
	(8,44.44)	(6,35.29)	-	(2,16.67)

ตารางที่ 32 (ต่อ)

จำนวนผู้สอบ	SA	UA	α_{MH}	β_{SIB}
600-800	(8,61.54)	(5,29.41)	(0,0.00)	(1,16.67)
	(8,61.54)	(8,42.10)	(0,0.00)	(3,25.00)
600-1,000	(8,47.06)	(6,33.33)	(0,0.00)	(0,0.00)
	(8,47.06)	(9,47.36)	-	(0,0.00)
800-1,000	(10,58.82)	(9,69.23)	(0,0.00)	(1,16.67)
	(10,58.82)	(10,52.63)	-	(1,7.69)

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตารางที่ 33 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในดัชนีเดียวกันเมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาคณิตศาสตร์ ความยาวแบบสอบ 40 ข้อ

จำนวนผู้สอบ	SA	UA	α_{MH}	β_{SIB}
100-200	(12,41.38)	(11,36.67)	(3,13.64)	(7,25.00)
	(12,44.44)	(11,36.67)	(0,0.00)	(2,08.33)
100-400	(14,45.16)	(9,25.00)	(5,15.79)	(4,15.38)
	(12,60.00)	(12,36.36)	(2,14.28)	(5,23.81)
100-600	(17,51.51)	(19,55.88)	(1,5.88)	(2,8.33)
	(17,53.12)	(19,55.88)	(0,0.00)	(0,0.00)
100-800	(16,51.61)	(14,41.17)	(1,5.88)	(3,13.04)
	(14,46.66)	(14,41.17)	(1,12.50)	(2,11.11)
100-1,000	(18,52.94)	(15,44.11)	(0,0.00)	(1,4.17)
	(17,51.51)	(15,41.66)	-	(1,5.88)
200-400	(10,35.17)	(9,33.33)	(3,21.43)	(1,5.00)
	(8,47.05)	(9,33.33)	(1,09.09)	(6,16.66)
200-600	(14,48.27)	(13,41.93)	(0,0.00)	(2,13.33)
	(13,44.83)	(13,41.93)	(0,0.00)	(4,25.00)
200-800	(15,60.00)	(13,50.00)	(0,0.00)	(3,21.67)
	(12,48.00)	(13,50.00)	(0,0.00)	(5,29.41)
200-1,000	(14,45.16)	(17,73.91)	(0,0.00)	(1,6.67)
	(13,43.33)	(12,42.85)	-	(2,11.11)
400-600	(16,51.61)	(17,54.88)	(0,0.00)	(0,0.00)
	(16,51.61)	(17,54.83)	(0,0.00)	(1,05.26)
400-800	(17,62.96)	(15,53.57)	(1,10.00)	(1,9.09)
	(14,53.85)	(15,53.57)	(2,22.22)	(4,22.23)
400-1,000	(19,63.33)	(18,69.23)	(0,0.00)	(2,00.00)
	(17,58.62)	(16,57.14)	-	(2,11.11)

ตารางที่ 33 (ต่อ)

จำนวนผู้สอบ	SA	UA	α_{MH}	β_{SIB}
600-800	(20,68.96)	(21,70.00)	(0,0.00)	(0,0.00)
	(18,62.06)	(20,64.51)	(0,0.00)	(2,16.66)
600-1,000	(24,80.00)	(22,73.33)	(0,0.00)	(0,0.00)
	(23,76.67)	(22,73.33)	-	(0,0.00)
800-1,000	(22,52.94)	(20,74.07)	(0,0.00)	(0,0.00)
	(20,71.42)	(20,74.07)	-	(0,0.00)

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

1.2 วิชาภาษาอังกฤษ

การตรวจสอบความสอดคล้องของข้อสอบลำเอียงระหว่างขนาดผู้สอบต่างกันภายใน แต่ละความยาวแบบสอบสำหรับวิชาภาษาอังกฤษ สำหรับดัชนีทั้ง 4 ตัว นำเสนอในตารางที่ 33-36

1.2.1 ผลจากการใช้ดัชนี SA

ผลการวิเคราะห์ความสอดคล้องของข้อสอบลำเอียงจากดัชนี SA เมื่อใช้ขนาดผู้สอบต่างกันโดยใช้เกณฑ์ตัดสินที่ไม่ได้พิจารณาความแตกต่างข้ามความยาวแบบสอบ และขนาดผู้สอบ พบว่าที่ความยาวแบบสอบ50 ข้อ ความสอดคล้องของการตัดสินข้อสอบลำเอียงด้วยดัชนี SA ระหว่างขนาดผู้สอบต่างกันมีค่าระหว่าง 13.04-61.11 ซึ่งเป็นความสอดคล้องระหว่างขนาดผู้สอบ 200 และ 800 คน และระหว่างขนาดผู้สอบ 600 และ 800 คน โดยมีความสอดคล้องระหว่างขนาดผู้สอบ 600 และ 1,000 คน คิดเป็นร้อยละ 58.82 ระหว่างขนาด 400 และ 1,000 คน คิดเป็นร้อยละ 55.00

ที่ความยาวแบบสอบ 60 ข้อ โดยส่วนรวมพบว่าร้อยละของความสอดคล้องในการตัดสินข้อสอบลำเอียงระหว่างผู้สอบต่างขนาดกันลดลงจากความยาวแบบสอบ 60 ข้อ โดยมีค่าระหว่างร้อยละ 13.63-35.29 ซึ่งเป็นความสอดคล้องระหว่างขนาดผู้สอบ 100 และ 200 คน และระหว่างขนาดผู้สอบ 400 และ 1,000 คน ที่ความยาวแบบสอบ 70 ข้อ ความสอดคล้องในการตัดสินข้อสอบลำเอียงที่ได้จากการใช้ดัชนี SA มีค่าระหว่างร้อยละ 15.38-68.96 ซึ่งเป็นความสอดคล้องเมื่อใช้ความยาวแบบสอบ 100 และ 400 คน และระหว่างความยาวแบบสอบ 600 และ 800 คน ร้อยละของความสอดคล้องมีค่าโดยเฉลี่ยเพิ่มขึ้นจากกรณีที่ใช้ความยาวแบบสอบ 60 คน และพบผลคล้าย ๆ กับกรณีของความยาวแบบสอบ 50 ข้อ คือ เมื่อขนาดผู้สอบเพิ่มขึ้นร้อยละของความสอดคล้องระหว่างขนาดผู้สอบต่างกันจะมีค่าเพิ่มขึ้นด้วย

ความสอดคล้องของการตัดสินข้อสอบลำเอียงระหว่างขนาดผู้สอบต่างกันสำหรับความยาวแบบสอบ 80 ข้อ มีค่าระหว่างร้อยละ 22.81-54.05

เมื่อใช้เกณฑ์ตัดสินที่พิจารณาความแตกต่างข้ามความยาวแบบสอบ และขนาดผู้สอบพบว่าส่วนใหญ่ความสอดคล้องจะมีค่าลดลงหรือเท่าเดิม โดยมีร้อยละของความสอดคล้องระหว่าง 5.55-50.00 5.88-40.00 8.69-58.33 และ 15.21-48.27 ตามลำดับความยาวแบบสอบ

1.2.2 ผลจากดัชนี UA

ผลจากดัชนี UA ซึ่งมีเกณฑ์ที่เท่ากันทุกบริบทไม่ว่าจะพิจารณาจากความแตกต่างข้ามความยาวแบบสอบหรือขนาดผู้สอบหรือไม่ก็ตาม พบว่าร้อยละของความสอดคล้องระหว่างขนาดผู้สอบ โดยทั่วไปต่ำกว่าผลจากดัชนี SA ในความยาวแบบสอบ 50 ข้อ มีค่าร้อยละของความสอดคล้องระหว่าง 7.50-32.35 ซึ่งเป็นความสอดคล้องระหว่างการวิเคราะห์โดยใช้ขนาดผู้สอบ 200 และ 800 คน และระหว่างขนาดผู้สอบ 600 และ 800 คน

ในกรณีที่ใช้ความยาวแบบสอบ 60 ข้อ ความสอดคล้องระหว่างการตัดสินข้อสอบล่าเอียงมีค่าร้อยละ 19.23-45.83 ซึ่งเป็นความสอดคล้องระหว่างขนาดผู้สอบ 200 และ 600 คน และระหว่างขนาดผู้สอบ 600 และ 800 คน

ร้อยละของความสอดคล้องในความยาวแบบสอบ 70 ข้อ ส่วนใหญ่มีค่าสูงขึ้นกว่าเมื่อใช้แบบสอบยาว 60 ข้อ โดยเฉพาะในขนาดผู้สอบสูง ๆ ร้อยละของความสอดคล้องของการตัดสินมีค่าระหว่าง 18.18-66.22 ซึ่งเป็นความสอดคล้องระหว่างขนาดผู้สอบ 100 และ 200 คน และระหว่างขนาดผู้สอบ 800 และ 1,000 คน

ผลการวิเคราะห์ความสอดคล้องระหว่างขนาดผู้สอบในความยาวแบบสอบ 80 ข้อ พบว่าร้อยละของความสอดคล้องของการค้นพบข้อสอบล่าเอียงระหว่างดัชนี SA และ UA มีค่าใกล้เคียงกัน โดยที่ความสอดคล้องจากดัชนี UA มีค่าร้อยละ 35.93-55.17 ซึ่งเป็นความสอดคล้องระหว่างขนาดผู้สอบ 100 และ 800 คน และระหว่าง 200 และ 400 คน

1.2.3 ผลจากดัชนี α_{MH}

เมื่อพิจารณาถึงความสอดคล้องในการตัดสินภายในวิธีเดียวกันในขนาดผู้สอบต่างกัน พบว่าการใช้ดัชนี αMH มีความสอดคล้องระหว่างขนาดผู้สอบต่างกันร้อยละ 8.00-33.33 3.70-37.50 3.57-40.00 และ 0.00-37.50 สำหรับความยาวแบบสอบ 50 60 70 และ 80 เมื่อใช้เกณฑ์ ที่ไม่ได้พิจารษความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ แต่เมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบ และขนาดผู้สอบ พบว่าความสอดคล้องส่วนหนึ่งมีค่าเท่าเดิม เนื่องจากใช้เกณฑ์เดิมทั้งในความยาวแบบสอบ 50 60 70 และ 80 ข้อ ในขณะที่ส่วนที่มีการปรับเกณฑ์ พบว่ามีความสอดคล้องลดลงในความยาวแบบสอบ 50 ข้อ และ 80 ข้อ ส่วนใหญ่เพิ่มขึ้นในความยาวแบบสอบ 60 และ 70

1.2.4 ผลจากดัชนี β_{SIB}

เมื่อใช้ดัชนี β_{SIB} พบว่าความสอดคล้องระหว่างการตัดสินในขนาดผู้สอบต่างกันระหว่าง 9.09-42.86 7.41-40.00 6.45-33.33 และ 7.32-40.00 ตามลำดับ สำหรับแบบสอบยาว 50 60 70 และ 80 ข้อ เมื่อใช้เกณฑ์ที่ไม่ได้พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ แต่เมื่อใช้เกณฑ์โดยพิจารณาความแตกต่างในบริบทดังกล่าว พบว่าส่วนหนึ่งมีความสอดคล้องเท่าเดิม อีกส่วนหนึ่งมีทั้งเพิ่มและลดลง รายละเอียดดังตารางที่ 34-36

ตารางที่ 34 จำนวนและร้อยละของข้อสอบที่ล่าเอียงสอดคล้องกันภายในดัชนีเดียวกันเมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาภาษาอังกฤษ ความยาวแบบสอบ 50 ข้อ

	SA	UA	α_{MH}	β_{SIB}
100-200	(4,16.66)	(5,18.51)	(7,22.58)	(5,16.13)
	(2,11.11)	(5,18.51)	(2,13.33)	(4,13.79)
100-400	(9,34.61)	(9,23.68)	(6,26.08)	(5,21.74)
	(6,27.27)	(9,23.68)	(3,21.42)	(5,22.72)
100-600	(6,24.00)	(7,25.92)	(2,8.33)	(2,9.09)
	(3,15.78)	(7,25.92)	(1,7.69)	(3,15.00)
100-800	(7,24.13)	(11,31.42)	(2,8.00)	(2,9.09)
	(2,8.69)	(11,31.42)	(1,8.33)	(2,10.52)
100-1,000	(7,25.92)	(11,28.94)	(2,8.69)	(1,4.76)
	(3,13.63)	(11,28.94)	(1,8.33)	(1,5.26)
200-400	(5,25.00)	(8,38.09)	(4,19.05)	(4,14.28)
	(4,22.22)	(8,38.09)	(1,6.66)	(4,14.28)
200-600	(4,23.52)	(5,19.23)	(1,14.76)	(4,16.66)
	(2,14.28)	(5,19.23)	(0,0.00)	(4,16.66)
200-800	(3,13.04)	(3,7.50)	(3,15.00)	(4,4.00)
	(1,5.55)	(10,31.25)	(0,0.00)	(1,4.00)
200-1,000	(4,20.00)	(5,11.90)	(3,16.67)	(3,13.63)
	(4,26.66)	(8,38.09)	(1,9.09)	(3,13.63)
400-600	(9,47.36)	(9,23.68)	(2,18.18)	(3,23.08)
	(6,33.33)	(7,28.00)	(2,18.18)	(4,22.22)
400-800	(10,43.34)	(10,20.83)	(1,9.09)	(4,33.33)
	(5,22.72)	(13,43.33)	(1,9.09)	(4,25.00)
400-1,000	(11,55.00)	(11,21.56)	(3,33.33)	(3,27.27)
	(8,42.10)	(8,36.36)	(3,33.33)	(3,18.75)

ตารางที่ 34 (ต่อ)

	SA	UA	α_{MH}	β_{SIB}
600-800	(11,61.11)	(11,32.35)	(2,28.57)	(3,33.33)
	(6,40.00)	(11,32.35)	(2,28.57)	(3,23.07)
600-1000	(10,58.82)	(10,25.64)	(1,12.50)	(3,42,86)
	(7,50.00)	(6,23.07)	(1,12.50)	(3,25.00)
800-1,000	(11,52.38)	(11,22.44)	(1,14.28)	(3,42.86)
	(8,50.00)	(11,34.37)	(1,14.28)	(3,30.00)

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ
* ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตารางที่ 35 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในดัชนีเดียวกันเมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาภาษาอังกฤษ ความยาวแบบสอบ 60 ข้อ

	SA	UA	α_{MH}	β_{SIB}
100-200	(3,13.63)	(4,20.00)	(5,16.13)	(8,20.00)
	(3,17.64)	(4,16.66)	(2,13.33)	(6,17.64)
100-400	(7,30.43)	(5,20.83)	(5,22.72)	(6,21.43)
	(5,23.80)	(5,20.83)	(3,23,00)	(9,36.00)
100-600	(5,22.72)	(6,24.00)	(2,09.09)	(3,11.54)
	(1,5.88)	(6,24.00)	(2,07.69)	(4,16.66)
100-800	(4,14.81)	(7,20.00)	(2,7.14)	(3,11.54)
	(2,8.33)	(7,20.00)	(2,12.50)	(3,13.04)
100-1,000	(3,15.00)	(5,20.83)	(1,3.70)	(2,8.33)
	(2,11.76)	(4,16.00)	(1,6.66)	(4,19.04)
200-400	(3,13.63)	(8,38.09)	(4,18.18)	(5,14.28)
	(3,15.78)	(8,38.09)	(4,26.66)	(5,14.28)
200-600	(4,22.22)	(5,19.23)	(1,4.76)	(8,30.76)
	(4,40.00)	(5,19.23)	(1,6.25)	(8,30.76)
200-800	(5,23.80)	(11,35.48)	(1,5.00)	(3,10.34)
	(4,22.22)	(11,35.48)	(1,6.66)	(3,10.34)
200-1,000	(4,28.57)	(7,31.81)	(3,18.75)	(3,10.71)
	(3,25.00)	(8,38.09)	(3,27.27)	(3,10.71)
400-600	(7,35.00)	(7,29.16)	(3,23.08)	(2,15.38)
	(4,25.00)	(7,28.00)	(3,30.00)	(3,12.00)
400-800	(8,34.78)	(13,43.33)	(1,11.10)	(3,25.00)
	(7,33.33)	(13,43.33)	(1,11.10)	(5,23.80)
400-1,000	(6,35.29)	(8,36.36)	(2,22.22)	(2,20.00)
	(4,23.52)	(9,42.85)	(2,22.22)	(4,19.07)

ตารางที่ 35 (ต่อ)

	SA	UA	α_{MH}	β_{SIB}
600-800	(7,33.33)	(11,45.83)	(2,28.57)	(2.25.00)
	(5,33.33)	(11,32.35)	(2,28.57)	(4,25.00)
600-1000	(3,17.64)	(6,23.07)	(1,14.28)	(1,16.67)
	(2,18.18)	(6,23.07)	(1,14.28)	(3,18.75)
800-1,000	(4,20.00)	(11,34.37)	(1,14.28)	(2,40.00)
	(3,16.66)	(11,34.37)	(1,14.28)	(4,30.76)

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ
* ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตารางที่ 36 จำนวนและร้อยละของข้อสอบที่ล่าเอียงสอดคล้องกันภายในดัชนีเดียวกัน เมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาภาษาอังกฤษ ความยาวแบบสอบ 70 ข้อ

	SA	UA	α_{MH}	β_{SIB}
100-200	(5,19.23)	(6,18.18)	(12,29.27)	(9,23.68)
	(2, 8.69	(6,18.18)	(3.12.50)	(7,11.29)
100-400	(8,28.57)	(7,23.33)	(5,14.28)	(7,20.00)
	(7,28.00)	(7,23.33)	(6,28.57)	(7,18.42)
100-600	(5,16.12)	(10,30.33)	(5,13.89)	(3,9.37)
	3,10.34)	(10,30.33)	(2,08.69)	(5,13.88)
100-800	(6,15.38)	(10,22.72)	(2,6.06)	(2,6.45)
	(6,18.75)	(10,22.72)	(2,11.11)	6.17.14)
100-1,000	(6,17.64)	(7,10.29)	(2,2.88)	(2,6.90)
	(3,09.67)	(7,10.29)	(1,5.55)	(3,06.45)
200-400	(8,29.62)	(8,25.00)	(1,3.57)	(4,14.81)
	(4,19.04)	(8,25.00)	(1,5.88)	(5,16.12)
200-600	(10,40.00)	(12,35.29)	(2,7.14)	(2,9.09)
	(5,25.00)	(12,35.29)	(0,00.00)	(8,33.33)
200-800	(10,29.41)	(14,32.55)	(0,0.00)	(2,10.00)
	(6,24.00)	(14,32.55)	(0,00.00)	(4,14.28)
200-1,000	(12,44.44)	(15,34.88)	(1,4.17)	(2,11.11)
	(7,35.00)	(15,34.88)	(1,09.09)	(5,22.72)
400-600	(12,42.85)	(11,33.33)	(2,14.28)	(2,11.76)
	(10,38.46)	(11,33.33)	(2,14.28)	(4,16.66)
400-800	(15,44.11)	(15,37.50)	(2,22.22)	(3,21.43)
	(10,35.71)	(15,37.50)	(2,22.22)	(8,40.00)
400-1,000	(13,41.93)	(15,36.58)	(2,25.00)	(2,15.38)
	(9,36.00)	(15,36.58)	(2,25.00)	(5,27.77)

ตารางที่ 36 (ต่อ)

	SA	UA	α_{HH}	β_{SIB}
600-800	(20,68.96)	(19,45.23)	(2,22.22)	(2,25.00)
	(14,58.33)	(19,45.23)	(2,22.22)	(3,14.28)
600-1000	(17,62.96)	(16,34.78)	(1,11.11)	(1,14.28)
	(12,54.54)	(16,34.78)	(1,11.11)	(3,18.75)
800-1,000	(19,55.88)	(27,58.69)	(2,25.00)	(2,33.00)
	(12,42.85)	(27,58.69)	(1,25.00)	(3,18.75)

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ
* ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

ตารางที่ 37 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันภายในดัชนีเดียวกันเมื่อใช้ขนาดผู้สอบต่างกัน สำหรับวิชาภาษาอังกฤษ ความยาวแบบสอบ 80 ข้อ

	SA	UA	α_{MH}	β_{SIB}
100-200	(18.33.33)	(32,47.76)	(10,19.64)	(9,17.65)
	(16,32.65)	(32,47.76)	(3,13.04)	(6,15.00)
100-400	(21,32.14)	(32,54.23)	(5,12.50)	(7,15.55)
	(21,30.88)	(31,51.66)	(3,17.64)	(8,20.51)
100-600	(13,22.81)	(29,46.77)	(4,10.25)	(5,11.63)
	(12,24.00)	(29,46.77)	(2,11.76)	(5,5.34)
100-800	(16,29.09)	(23,35.93)	(1,2.70)	(2,7.32)
	(16,32.65)	(24,38.09)	(2,16.66)	(3,9.09)
100-1,000	(14,26.41)	(25,39.68)	(1,2.63)	(3,7.50)
	(7,15.21)	(25,36.68)	(1,7.14)	(4,12.90)
200-400	(18,37.50)	(21,55.17)	(6,12.50)	(4,10.81)
	(17,37.77)	(21,55.17)	(4,20.00)	(4,10.81)
200-600	(20,47.61)	(30,50.00)	(2,6.25)	(6,22.22)
	(17,44.73)	(30,50.00)	(1,4.54)	(6,22.22)
200-800	(15,31.25)	(28,48.27	(0,0.00)	(2,7.14)
	(13,30,23)	(23,48.27)	(0,0.00)	(2,7.14)
200-1,000	(19,47.50)	(28,47.45)	(1,3.45)	(4,16.00)
	(12,35.29)	(29,47.45)	(1,5.55)	(4,16.00)
400-600	(22,52.38)	(29,54.71)	(3,21.42)	(3,17.65)
	(19,47.50)	(29,54.71)	(3,21.42)	(4,13.30)
400-800	(22,51.16)	(23,41.18)	(2,20.00)	(2,14.28)
	(18,42.85)	(23,41.18)	(2,20.00)	(5,19.23)
400-1,000	(21,52.50)	(22,38.59)	(2,18.18)	(2,15.38)
	(15,42.85)	(22,38.59)	(2,18.18)	(6,25.00)

ตารางที่ 37 (ต่อ)

	SA	UA	α_{MH}	β_{SIB}
600-800	(18,41.86)	(25,47.16)	(3,37.50)	(2,20.00)
	(14,35.89)	(25,47.16)	(3,37.50)	(4,23.52)
600-1000	(20,54.05)	(25,46.29)	(1,9.09)	(3,37.50)
	(14,48.27)	(25,46.29)	(1,9.09)	(5,29.41)
800-1,000	(16,38.09)	(26,53.06)	(1,16.67)	(2,40.00
	(19,37.50)	(26,53.06)	(1,16.67)	(4,26.66)

* ตัวเลขแถวบนคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่ไม่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ
* ตัวเลขแถวล่างคือ จำนวนและร้อยละของข้อสอบลำเอียงเมื่อใช้เกณฑ์ที่พิจารณาความแตกต่างข้ามความยาวแบบสอบและขนาดผู้สอบ

2. ความสอดคล้องภายในดัชนีเดียวกัน ความยาวแบบสอบต่างกัน

2.1 วิชาคณิตศาสตร์

1) ผลจากดัชนี SA และ UA

เมื่อวิเคราะห์ความสอดคล้องของการตัดสินข้อสอบลำเอียงเมื่อใช้ความยาวแบบสอบต่างกัน ผลดังตารางที่ 38 พบว่าสำหรับผลจากดัชนี SA นั้น นอกจากที่ขนาดผู้สอบ 200 คน ซึ่งมีความสอดคล้องของการตัดสินข้อสอบลำเอียงระหว่างแบบสอบยาว 20 และ 30 ข้อ อยู่ 1 ข้อ คิดเป็นร้อยละ 50.00 ของข้อสอบลำเอียงทั้งหมดของทั้ง 2 ความยาวแล้ว พบความสอดคล้องสูงสุดระหว่างแบบสอบยาว 30 และ 40 ข้อ ที่ขนาดผู้สอบ 800 คน คิดเป็นร้อยละ 33.33 ต่ำสุดคือที่ความยาวแบบสอบ 20 และ 30 ข้อ ขนาดผู้สอบ 100 คน ที่พบว่าไม่มีข้อสอบสอดคล้องกันเลย

สำหรับผลการวิเคราะห์ความสอดคล้องของข้อสอบลำเอียงระหว่างแบบสอบที่มีความยาวต่างกัน พบว่ามีความสอดคล้องต่ำสุดร้อยละ 0.00 และสูงสุดร้อยละ 36.67

เมื่อวิเคราะห์ผลจากตารางที่ 49 เพื่อพิจารณาข้อสอบที่ได้รับการตัดสินว่าลำเอียงทุกความยาวแบบสอบพบว่า ในกรณีใช้ดัชนี SA มีอยู่ 3 ข้อ ได้แก่ ข้อ 16 ในขนาดผู้สอบ 200 800 และ 1,000 และข้อ 23 และ 26 ในขนาดผู้สอบ 800 และ 1,000 คน

ส่วนผลจากดัชนี UA พบข้อ 5 ในขนาดผู้สอบ 600 คน ข้อ 23 และ 25 ในขนาดผู้สอบ 800 และ 1,000 คน และข้อ 26 ในขนาดผู้สอบ 800 คน

ตารางที่ 38 ข้อสอบล่าเอียงและร้อยละของข้อสอบที่ล่าเอียงสอดคล้องกัน ภายในดัชนีเดียวกัน (SA และ UA) เมื่อใช้ความยาวแบบสอบต่างกัน จำแนกตามขนาดผู้สอบ (วิชาคณิตศาสตร์)

ขนาดผู้สอบ(คน)	SA			UA		
	20-30	20-40	30-40	20-30	20-40	30-40
100	- (0.00%)	15,29 (8.33%)	1,10,17,25 (16.00%)	- (0.00%)	- (0.00%)	- (0.00%)
200	16 (50.00%)	16 (5.88%)	1,16,24 (18.75%)	- (0.00%)	16 (6.25%)	1,15,24 (20.00%)
400	- (0.00%)	15,23,29,30 (18.18%)	10,16,17, 19,25,26,28 (29.27)	- (0.00%)	15,24,26,29, 30 (22.73%)	16,21,25,28 (19.04%)
600	- (0.00%)	15,25,26 (11.54%)	1,10,16,17, 22,23,24, 30 (27.59%)	5 (5.55%)	5,15,24,25, 26 (16.67%)	1,4,5,10,16, 17,21,22,23, 30,37 (36.67%)
800	16,23,26 (18.75%)	16,23,26,29 (16.00%)	1,10,11,16, 17,23,24, 26,32 (33.33%)	23,25,26 (25.00%)	5,23,26,29 (16.67%)	1,16,17,23, 24,26,41 (26.92%)
1,000	16,23,26 (15.79%)	15,16,23, 26,29,30 (20.00%)	1,11,16.17, 23,24,25, 26,28,32 (31.25%)	23,25 (12.50%)	5,23,25,29 (15.38%)	1,4,16,23, 24,25,26 (24.14%)

* ตัวเลขในวงเล็บคือ ข้อสอบที่ล่าเอียงสอดคล้องกัน

2) ผลจากดัชนี α_{MH} และ β_{SIB}

เมื่อตรวจสอบความสอดคล้องระหว่างผลการตัดสินข้อสอบลำเอียง ระหว่างความยาวแบบสอบที่ต่างกัน พบว่ามีความสอดคล้องร้อยละ 6.66 และ 25.00 ระหว่างความยาวแบบสอบ 20-30 ข้อ ในขนาดผู้สอบ 100 และ 200 คน ในขนาดอื่น ๆ ไม่พบว่ามีความสอดคล้องระหว่างความยาวแบบสอบ 20-40 ข้อ พบความสอดคล้องร้อยละ 4.76 ในขนาดผู้สอบ 100 คน และ 20.00 ในขนาด 400 คน ที่ขนาดอื่นไม่พบความสอดคล้อง และพบความสอดคล้องในขนาดผู้สอบ 100 และ 400 คน ระหว่างความยาวแบบสอบ 30-40 ข้อ โดยมีจำนวนร้อยละ 23.80 และ 15.38 ตามลำดับ ส่วนผลการตรวจสอบในกรณีใช้ดัชนี β_{SIB} พบความสอดคล้องในบางขนาดผู้สอบเช่นกัน โดยระหว่างความยาว 20 และ 30 ข้อ พบความสอดคล้องในขนาดผู้สอบ 100 200 400 และ 800 คน คิดเป็นร้อยละ 31.57 25.00 9.09 และ 20.00 ระหว่างความยาวแบบสอบ 30 และ 40 ข้อ พบความสอดคล้องที่ขนาด 100 200 400 และ 600 คิดเป็นร้อยละ 25.92 12.00 40.00 และ 20.00 และระหว่างความยาวแบบสอบ 30-40 ข้อ พบความสอดคล้องในขนาดผู้สอบ 100 200 และ 400 คน คิดเป็นร้อยละ 34.61 13.63 และ 27.27 ตามลำดับ ทั้งนี้ผลจากการใช้ดัชนี β_{SIB} ให้ผลที่สอดคล้องกันสูงกว่าผลจาก α_{MH} รายละเอียดดังตารางที่ 39

ตารางที่ 39 ข้อสอบล่าเอียงและร้อยละของข้อสอบที่ล่าเอียงสอดคล้องกัน ภายในดัชนีเดียวกัน (α_{MH} และ β_{SIB}) เมื่อใช้ความยาวแบบสอบต่างกัน จำแนกตามขนาดผู้สอบ (วิชาคณิตศาสตร์)

ขนาดผู้สอบ (คน)	α_{MH}			β_{SIB}		
	20 - 30	20 - 40	30 - 40	20 - 30	20 - 40	30 - 40
100	29 (6.66%)	29 (4.76%)	10,11,28, 29,41 (23.80%)	2,5,9,15, 29,33 (31.57%)	1,5,9,15, 33,35,36 (25.92%)	4,5,9,15, 25,31,33, 37,38 (34.61%)
200	8,39 (25.00%)	- (0.00%)	- (0.00%)	8,32,37 (25.00%)	8,37 (12.50%)	8,10,37 (13.63%)
400	-	29,39 (20.00%)	10,38 (15.38%)	30 (9.09%)	29,30,32, 39 (40.00%)	28,30,38 (27.27%)
600	-	-	-	-	37 (20.00%)	-
800	-	-	-	15,16 (20.00%)	-	-
1,000	-	-	-	-	-	-

* ในวงเล็บคือ ร้อยละของข้อสอบที่ล่าเอียงสอดคล้องกัน

2.2 วิชาภาษาอังกฤษ

2.2.1 ความสอดคล้องของการตัดสินด้วยดัชนี SA และ UA เมื่อความยาวแบบสอบต่างกัน

1) ผลจากดัชนี SA

เมื่อพิจารณาความสอดคล้องของการค้นพบข้อสอบลำเอียงในวิชาภาษาอังกฤษระหว่างแบบสอบที่ใช้ความยาวแบบสอบต่างกันพบว่าร้อยละของความสอดคล้องระหว่างแบบสอบยาว 50 และ 70 ข้อ มีค่าต่ำที่สุดคือ 6.06 ที่ขนาดผู้สอบ 1,000 คน มีความสอดคล้องสูงสุดระหว่างแบบสอบยาว 70 และ 80 ข้อ ขนาดผู้สอบ 1,000 คน คิดเป็นร้อยละ 41.67

และเมื่อพิจารณาจำนวนข้อสอบที่ได้รับการตัดสินว่าลำเอียงทุกความยาวแบบสอบ พบว่ามี 5 ข้อ ได้แก่ ข้อ 47 ในขนาดผู้สอบ 200 คน ข้อ 96 ในขนาดผู้สอบ 400 คน ข้อ 83 ในขนาดผู้สอบ 600 คน และ 1,000 คน และข้อ 98 ในขนาดผู้สอบ 800 คน รายละเอียดดังตารางที่ 40

2) ผลจากดัชนี UA

ความสอดคล้องของการตัดสิน เมื่อใช้ความยาวแบบสอบต่างกันด้วยดัชนี UA พบว่าร้อยละของความสอดคล้องต่ำสุด ได้แก่ ความสอดคล้องระห่างความยาวแบบสอบ 60 และ 70 ข้อ ขนาดผู้สอบ 100 คน มีค่าร้อยละ 10.34 และที่สอดคล้องกันสูงสุด ได้แก่ ความสอดคล้องระหว่างความยาวแบบสอบ 50 และ 70 ข้อ มีค่าร้อยละ 40.82

จำนวนข้อสอบที่ได้รับการตัดสินว่าลำเอียงทุกความยาวแบบสอบ ได้แก่ ข้อ 12 21 29 37 41 42 53 79 81 83 96 และ 98 รายละเอียดดังตารางที่ 41

ตารางที่ 40 ข้อสอบล่าเอียงและร้อยละของข้อสอบที่ล่าเอียงสอดคล้องกัน ในการตัดสินด้วยดัชนี SA เมื่อใช้ความยาวแบบสอบต่างกันจำแนกตามขนาดผู้สอบ (วิชาภาษาอังกฤษ)

ขนาดผู้สอบ(คน)	ความยาวแบบสอบ (ข้อ)					
	50-60	50-70	50-80	60-70	60-80	70-80
100	4,15,29,46,62 (17.24%)	29,41 (6.06%)	4,15,19,21,32,41,46,48,62,79,90,96 (25.53%)	25,26,46,53,62,65,75 (29.17%)	4,11,15,26,46,53,62,65,75 (19.56%)	1,26,31,37,41,59,66, (14.28%)
200	12,47,81 (18.75%)	15,42,47 (14.28%)	15,42,46,47,61,81, (17.14%)	25,33,47,83 (19.05%)	47,81,83 (7.69%)	1,15,26,35,39,42,47,53,56,79,83 (29.73%)
400	15,25,53,65,79,96 (24.00%)	12,25,47,53,87,96,98 (24.14%)	15,32,41,47,53,79,87,96,98 (21.95%)	25,26,37,53,96 (16.67%)	7,10,15,20,26,37,79,96 (19.51%)	14,26,37,39,47,59,82,87,96,98 (22.73%)
600	15,25,83 (14.28%)	15,41,42,47,79,83,96 (28.00%)	41,42,47,79,83,92,96 (20.00%)	15,37,39,46,83 (18.52%)	7,20,37,39,46,83 (16.67%)	26,35,37,39,41,42,46,47,79,82,83,87,96,98 (38.89%)

ตารางที่ 40 (ต่อ)

ขนาดผู้สอบ(คน)	ความยาวแบบสอบ (ข้อ)					
	50-60	50-70	50-80	60-70	60-80	70-80
800	15,33,71, 79,83,96, 98 (26.92%)	15,33,41, 42,71,79, 83,92,96,98 (27.78%)	41,74,92,98 (9.09%)	12,15,26,33, 37,39,71,73, 75,79,83,87, 96,98 (45.16%)	37,39,87,98 (9.30%)	20,21,37,39, 41,82,87,92, 98,100 (20.00%)
1,000	81,83 (9.52%)	15,41,42, 47,53,79, 83,96,98 (30.00%)	41,42,47, 53,83,96,98 (20.00%)	26,83 (6.67%)	25,26,83 (9.37%)	1,26,37,39, 41,42,47,48, 53,56,82,83, 87,96,98 (41.67%)

* ตัวเลขในวงเล็บคือ ร้อยละของข้อสอบที่ค่าเอียงสอดคล้องกันระหว่างความยาวแบบสอบ

ตารางที่ 41 ข้อสอบลำเอียงและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันในการตัดสินใจด้วยดัชนี UA เมื่อใช้ความยาวแบบสอบต่างกันจำแนกตามขนาดผู้สอบ (วิชาภาษาอังกฤษ)

ขนาดผู้สอบ(คน)	ความยาวแบบสอบ (ข้อ)					
	50-60	50-70	50-80	60-70	60-80	70-80
100	29,62,76,79 (14.81%)	4,29,90,96 (12.90%)	4,7,19,28,29,32,40,41,62,79,90,96 (21.42%)	25,29,46 (10.34%)	11,15,21,25,29,43,46,53,62,75,79 (20.37%)	1,4,25,26,29,31,37,46,60,66,82,90,96 (23.64%)
200	12,47,69,76 (16.00%)	7,15,42,47,81 (16.13%)	7,15,28,32,38,42,47,74,76,81,98 (20.75%)	21,26,39,47,83 (16.67%)	10,11,20,26,39,47,53,76,83 (16.67%)	1,7,14,15,24,26,35,39,41,42,47,56,59,81,83,87,90,100 (34.61%)
400	12,25,37,69,76,79,96 (17.86%)	12,15,21,37,47,71,79,81,82,87,96,98 (32.43%)	15,25,28,32,37,41,46,47,59,73,79,81,82,87,96,98 (29.09%)	12,26,37,79,96 (17.24%)	10,20,25,26 37,75,79,96 (16.67%)	15,26,35,37,39,47,72,79,81,82,87,96,98 (27.66%)

ตารางที่ 41 (ต่อ)

ขนาดผู้สอบ(คน)	ความยาวแบบสอบ (ข้อ)					
	50-60	50-70	50-80	60-70	60-80	70-80
600	12,41,45, 46,79,83, 96 (25.92%)	15,41,42, 46,92,96, 98 (23.53%)	15,25,28, 32,41,42, 45,46,47, 83,92,96,98 (28.89%)	37,39,41,43, 46,87,96 (20.00%)	7,20,37,39, 41,43,45,46, 83,87,96 (23.40%)	8,15,26,35, 37,39,41,42, 43,46,47,54, 82,87,90,92, 96,98,100 (40.42%)
800	3,12,19,21, 40,42,53,62, 63,71,75,76, 79,83,96,98 (40.00%)	4,15,21,41, 42,46,47,53, 71,75,79,81, 82,92,96,98 (33.33%)	3,21,25,28,31, 41,42,53,62, 74,76,82,92, 96,98 (30.00%)	20,21,26,37, 39,42,45,53, 56,71,73,75, 79,87,96,98 (33.33%)	3,10,20,21, 37,42,48,53, 62,76,87,95, 98 (25.00%)	8,20,21,24, 37,41,42,53, 57,60,82,87, 92,98,100 (28.86%)
1,000	11,12,19,69, 76,79,81,83, 96 (23.68%)	4,8,11,12, 15,21,34,37, 41,42,47,53, 73,79,81,83, 87,92,96,98 (40.82%)	4,12,15,21,25, 28,37,42,47, 53,69,73,74, 81,83,87,96, 96,98 (34.61%)	11,12,26,79, 81,83,96 (15.55%)	12,56,69,81, 83,96 (12.76%)	1,4,9,12,15, 20,21,24,37, 42,47,48,49, 73,81,82,83, 87,90,96,98 (21,38.89%)

* ตัวเลขในวงเล็บคือ ร้อยละของความสอดคล้องในการค้นพบข้อสอบลำเอียงระหว่างความยาวแบบสอบที่ต่างกัน

3) ผลจากดัชนี α_{MH}

ความสอดคล้องของการตัดสินด้วยดัชนี α_{MH} เมื่อความยาวแบบสอบต่างกัน พบว่ามีความสอดคล้องของการค้นพบข้อสอบลำเอียง ระหว่างความยาวแบบสอบ 50 และ 60 ข้อ ร้อยละ 0.00-37.50 ระหว่างความยาวแบบสอบ 50-70 ข้อ ร้อยละ 0.00-23.07 ระหว่างความยาวแบบสอบ 50 และ 80 ข้อ ร้อยละ 0.00-40.00 ระหว่าง 60 และ 70 ข้อ ร้อยละ 0.00-26.08 ระหว่าง 60 และ 80 ข้อ ร้อยละ 0.00-17.64 และระหว่าง 70-80 ข้อ ร้อยละ 0.00-32.00 และพบว่ามีข้อสอบที่พบว่าลำเอียงทั้งในความยาวแบบสอบ 50-80 ข้อ อยู่เพียง 2 ข้อ ได้แก่ ข้อ 11 และ 46 ในขนาดผู้สอบ 100 คน และข้อ 7 ในขนาดผู้สอบ 200 คน รายละเอียดดังตารางที่ 42

4) ผลจากดัชนี β_{SIB}

สำหรับผลจากการใช้ β_{SIB} พบความสอดคล้องระหว่างแบบสอบยาว 50-60 50-70 50-80 60-70 60-80 และ 70-80 ในขนาดผู้สอบทั้ง 6 ขนาด ระหว่างร้อยละ 0.00- 38.70 0.00-21.05 0.00-18.36 4.54-33.33 0.00-16.66 และ 0.00-23.63 ตามลำดับ รายละเอียดดังตารางที่ 42 ทั้งนี้พบว่ามีข้อสอบลำเอียงสอดคล้องกันทุกความยาวแบบสอบเพียง 1 ข้อ คือ ข้อ 16

ตารางที่ 42 ข้อสอบลำเอียงและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันในการตัดสินด้วยดัชนี α_{MH} เมื่อใช้ความยาวแบบสอบต่างกันจำแนกตามขนาดผู้สอบ (วิชาภาษาอังกฤษ)

ขนาดผู้สอบ (คน)	ความยาวแบบสอบ (ข้อ)					
	50 - 60	50 - 70	50 - 80	60 - 70	60 - 80	70 - 80
100	3,8,9,11, 46,62,65, 76 (21.62%)	3,9,11,19, 25,40,41, 46,92 (21.42%)	4,11,12, 32,40,46, 92 (15.21%)	1,3,7,9, 11,14,15, 26,33,45, 46,59 (26.08%)	1,7,11,15, 20,29,46, 59 (17.64%)	1,6,7,11, 15,27,24, 28,37,40, 46,47,59, 66,92,95 (32.00%)
200	7,15,16, 60,61,62, 81,82,90 (37.50%)	7,15,19, 29,45,46 (19.35%)	7,16,19, 61,65 (13.51%)	7,15,25, 33 (12.50%)	7,15,61 (7.89%)	1,7,19,35, 53,56,64, 84,93 (25.00%)
400	25 (6.66%)	16,25,40 (23.07%)	40 (6.25%)	25 (6.66%)	(0.00%)	17,40 (13.33%)
600	38 (11.11%)	47 (8.33%)	33,47 (18.18%)	15 (8.33%)	31 (8.33%)	1,47 (14.28%)
800	(0.00%)	40 (1,16.66%)	33,40 (40.00%)	(0.00%)	(0.00%)	40 (20.00%)
1,000	– (0.00%)	– (0.00%)	– (0.00%)	25 (25.00%)	– (0.00%)	– (0.00%)

* ในวงเล็บคือ จำนวนและร้อยละของข้อสอบลำเอียงที่สอดคล้องกัน

ตารางที่ 43 ข้อสอบล่าเอียงและร้อยละของข้อสอบที่ล่าเอียงสอดคล้องกัน ในการตัดสินด้วยดัชนี β_{SIB} เมื่อใช้ความยาวแบบสอบต่างกัน จำแนกตามขนาดผู้สอบ (วิชาภาษาอังกฤษ)

ขนาดผู้สอบ (คน)	ความยาวแบบสอบ (ข้อ)					
	50 - 60	50 - 70	50 - 80	60 - 70	60 - 80	70 - 80
100	3,11,62,76 (10.81%)	3,11,21, 25,62,82, 83,92 (21.05%)	11,21,32, 48,60,75, 79,82,92 (18.36%)	1,3,11,46, 56,62,65, 66,72 (18.36%)	1,11,12, 20,21,33, 46,56,87 (16.66%)	1,7,12,21, 46,56,59, 64,74,82, 92,95,100 (23.63%)
200	15,16,41, 42,47,60, 61,62,81, 82,90,92 (38.70%)	16,29,46, 53,62,82 (19.35%)	7,8,16,41, 61 (14.70%)	9,16,25, 32,45,62, 82 (20.00%)	12,16,41, 45,61,66 (15.78%)	16,17,19, 45,54,64 (18.75%)
400	25,29 (11.11%)	14,25,40 (15.00%)	40 (4.76%)	25 (4.54%)	10 (4.76%)	17,38,40, 59 (19.04%)
600	(22.22%)	29 (11.11%)	33 (0.00%)	15 (12.50%)	31 (0.00%)	(0.00%)
1,000	- (0.00%)	- (0.00%)	- (0.00%)	25 (33.33%)	- (0.00%)	- (0.00%)

* ในวงเล็บคือ จำนวนและร้อยละของข้อสอบล่าเอียงที่สอดคล้องกัน

ค. ความสอดคล้องในการตัดสินข้อสอบลำเอียง ระหว่างการใช้เกณฑ์ที่กำหนดและการทดสอบนัยสำคัญของค่าสถิติ

ผลการวิเคราะห์ความสอดคล้องของการตัดสินข้อสอบลำเอียง เมื่อใช้เกณฑ์ที่กำหนด (ค่าดัชนี SA UA α_{MH} และ β_{SIB}) และการทดสอบนัยสำคัญของค่า χ^2 และ Z แสดงไว้ในตารางที่ 44-46 สำหรับวิชาคณิตศาสตร์ และในตารางที่ 47-50 สำหรับวิชาภาษาอังกฤษ มีรายละเอียดดังนี้

1. วิชาคณิตศาสตร์

ผลในตารางที่ 44-46 แสดงความสอดคล้องระหว่าง SA-χ^2 ว่าสูงกว่า UA-χ^2 เล็กน้อย ในขนาดผู้สอบ 100-600 คน ในขณะที่มีทั้งสูงกว่าและต่ำกว่าในขนาดผู้สอบ 800 และ 1,000 คน สำหรับร้อยละของความสอดคล้องเรียงลำดับตามความยาวแบบสอบ 20-40 ข้อ จำแนกตามขนาดผู้สอบ 6 ขนาด ในแต่ละความยาวแบบสอบ สำหรับ SA-χ^2 มีค่าระหว่างร้อยละ 18.18-50.00 0.00-25.00 และ 0.00-33.33 ในขณะที่ร้อยละของความสอดคล้องระหว่าง UA - χ^2 มีค่าระหว่าง 0.00-50.00 0.00-60.00 และ 0.00-30.00 ตามลำดับ

นอกจากนั้นสังเกตได้ว่าร้อยละของความสอดคล้องระหว่าง SA-χ^2 ในขนาดผู้สอบ 20 ข้อ มีค่าสูงกว่าในขนาดผู้สอบ 30 และ 40 ข้อ ในขนาดผู้สอบเดียวกัน ยกเว้นที่ขนาดผู้สอบ 600 คน ซึ่งต่ำกว่า สำหรับค่าระหว่างความยาวแบบสอบ 30 และ 40 ไม่สามารถสังเกตลักษณะที่เป็นระบบได้

ความสอดคล้องของ α_{MH}-χ^2 มีค่าสูงมากทุกความยาวแบบสอบที่ขนาดผู้สอบ 400 และ 600 คน (เกินกว่า 50.00%) และมีค่าสูงกว่าค่าที่พบใน SA - χ^2 และ UA - χ^2 โดยเฉพาะอย่างยิ่งในขนาดผู้สอบ 400-800 คน ทั้งนี้ค่าร้อยละของความสอดคล้องในแต่ละความยาวแบบสอบมีค่าระหว่าง 0.00-66.67 0.00-83.33 และ 0.00-66.67 ตามลำดับ

ความสอดคล้องระหว่างผลของ α_{MH}-Z_{SIB} ในความยาวแบบสอบ 40 ข้อ มีค่าสูงกว่าค่าระหว่าง α_{MH}-χ^2_{MH} ทุกขนาดผู้สอบ แต่ในขนาด 20 และ 30 ข้อ จะมีค่าแตกต่างกันไป สามารถระบุเป็นระบบได้ ทั้งนี้จะมีค่าร้อยละของความสอดคล้องสูงสุดที่ขนาดผู้สอบ 400 600 และ 800 คน โดยมีค่าร้อยละของความสอดคล้องในแต่ละความยาวแบบสอบระหว่าง 0.00-100.00 7.69-75.00 และ 13.00-100.00 ตามลำดับ

ตารางที่ 44 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันในวิชาคณิตศาสตร์ระหว่างการใช้เกณฑ์และการทดสอบนัยสำคัญของค่าสถิติภายในวิธีเดียวกัน ความยาวแบบสอบ 20 ข้อ*

ขนาดผู้สอบ	SA - χ^2	UA - χ^2	α_{MH} - χ^2_{MH}	β_{SIB} - Z_{SIB}
100	2,2 (1, 33.33)	-,2 0.00	6,1 (1, 16.67)	12,2 (2, 16.67)
200	1,2 (1, 50.00)	1,2 (1, 50.00)	4,1 (1, 25.00)	4,- 0.00
400	5,10 (3, 25.00)	7,10 (4, 30.76)	3,2 (2, 66.67)	6,3 3, 50.00
600	3,10 (2, 18.18)	7,10 4, 30.76	2,3 (2, 66.67)	2,2 (2, 100.00)
800	6,16 (6, 37.50)	5,16 (5, 31.25)	2,5 (2, 40.00)	4,4 (4, 100.00)
1,000	8,17 (7, 47.05)	6,17 (6, 35.29)	0,1 (0.00)	1,2 1, 100.00

* แถวบนคือ จำนวนข้อสอบลำเอียงในแต่ละวิธี
แถวล่าง (ในวงเล็บ) คือ จำนวนข้อสอบที่ลำเอียงซ้ำกันและร้อยละเมื่อเทียบกับข้อสอบลำเอียงทั้งหมดที่พบ

ตารางที่ 45 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันในวิชาคณิตศาสตร์ระหว่างการใช้เกณฑ์และการทดสอบนัยสำคัญของค่าสถิติภายในวิธีเดียวกัน ความยาวแบบสอบ 30 ข้อ*

ขนาดผู้สอบ	SA - χ^2	UA - χ^2	α_{MH} - χ^2_{MH}	β_{SIB} - Z_{SIB}
100	5,2 (1, 16.67)	-,2 (0,0.00)	10,2 (2, 20.00)	13,1 (1, 7.69)
200	6,2 (0,0.00)	2,2 (4, 10.00)	6,2 (2, 33.33)	11,2 (2, 18.18)
400	10,4 (2, 16.67)	4,4 (3, 60.00)	6,5 (5, 83.33)	6,4 (4, 66.67)
600	11,14 (5, 25.00)	12,14 (8, 44.44)	3,4 (3, 75.00)	3,2 (2, 66.67)
800	13,14 (6, 28.57)	10,14 (4, 20.00)	1,5 (1, 20.00)	7,7 (6, 75.00)
1,000	14,17 (5, 19.23)	12,17 (5, 20.83)	0,2 (0, 0.00)	1,4 (1, 25.00)

* แถวบนคือ จำนวนข้อสอบลำเอียงในแต่ละวิธี
แถวล่าง (ในวงเล็บ) คือ จำนวนข้อสอบที่ลำเอียงซ้ำกันและร้อยละเมื่อเทียบกับข้อสอบลำเอียงทั้งหมดที่พบ

ตารางที่ 46 จำนวนและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันในวิชาคณิตศาสตร์ระหว่างการใช้เกณฑ์และการทดสอบนัยสำคัญของค่าสถิติภายในวิธีเดียวกัน ความยาวแบบสอบ 40 ข้อ*

ขนาดผู้สอบ	SA - χ^2	UA - χ^2	α_{MH} - χ^2_{MH}	B_{SIB} - Z_{SIB}
100	24,- (0, 0.00)	25,- (0,0.00)	16,2 (2, 12.50)	22,3 (3,13.64)
200	17,2 (2, 11.76)	16,2 (2, 12.50)	9,1 (1, 11.11)	14,2 (2, 14.28)
400	21,3 (2, 9.09)	20,3 (2, 9.52)	9,6 (6, 66.67)	8,6 (6, 75.00)
600	26,15 (8, 24.24)	28,15 (10, 30.30)	2,3 (2, 66.67)	4,4 (4, 100.00)
800	23,17 10, 33.33	23,17 (8, 25.00)	2,6 (2, 33.33)	4,8 (4, 50.00)
1,000	28,19 (12, 34.28)	24,19 (10, 30.30)	0,2 0.00	3,7 (3, 42.86)

* แถวบนคือ จำนวนข้อสอบลำเอียงในแต่ละวิธี
แถวล่าง (ในวงเล็บ) คือ จำนวนข้อสอบที่ลำเอียงซ้ำกันและร้อยละเมื่อเทียบกับข้อสอบลำเอียงทั้งหมดที่พบ

2. วิชาภาษาอังกฤษ

สำหรับวิชาอังกฤษ ความสอดคล้องในการตัดสินข้อสอบลำเอียง เมื่อพิจารณาตามขนาดผู้สอบภายในความยาวแบบสอบเดียวกัน มีค่าแตกต่างกัน โดยที่ความสอดคล้องระหว่างการใช้ดัชนี SA และการทดสอบนัยสำคัญมีร้อยละของความสอดคล้องระหว่าง 21.05-40.00 ร้อยละ 6.25-31.25 ร้อยละ 6.25-31.91 และร้อยละ 2.86-29.27 สำหรับความยาวแบบสอบ 50 60 70 และ 80 ข้อ ตามลำดับ

สำหรับความสอดคล้องระหว่างการใช้ดัชนี UA และการทดสอบค่าสถิติพบว่ามีความสอดคล้องระหว่างร้อยละ 13.33-65.22 ร้อยละ 6.67-29.63 ร้อยละ 4.54-42.86 และร้อยละ 4.00-33.33 ตามลำดับ ซึ่งพบว่าส่วนมากมีค่าแตกต่างจากผลความสอดคล้องระหว่างการใช้ดัชนี SA และการทดสอบค่าสถิติ

ผลการพิจารณาความสอดคล้องระหวางดัชนี α_{MH} และการทดสอบค่าสถิติและความสอดคล้องระหว่างดัชนี β_{SIB} และการทดสอบค่าสถิติ พบว่ามีความแตกต่างกันในแต่ละขนาดผู้สอบทั้งภายในแต่ละวิธี และระหว่างทั้ง 2 วิธีเช่นกัน โดยวิธี MH ให้ค่าความสอดคล้องเรียงลำดับ ตามความยาวแบบสอบตั้งแต่ 50 ข้อ มีค่าระหว่างร้อยละ 15.79-75.00 ร้อยละ 15.38-87.50 ร้อยละ 10.00-75.00 ตามลำดับ และร้อยละ 11.76-70.00 ซึ่งมีค่าสูงกว่าความสอดคล้องระหว่าง SA กับการทดสอบค่าสถิติเป็นส่วนใหญ่ แต่ไม่สามารถสรุปได้เมื่อเปรียบเทียบกับความสอดคล้องระหว่าง UA และการทดสอบค่าสถิติ เนื่องจากมีทั้งที่ใกล้เคียงกันและที่แตกต่างกัน ความสอดคล้องระหว่างการใช้ดัชนี α_{MH} และการทดสอบ MH - χ^2 มีค่าสูงมากที่สุดร้อยละ 60.00-87.50 ที่ขนาดผู้สอบ 400 คน รองลงมาคือที่ขนาดผู้สอบ 600 800 คน ตามลำดับ โดยมีค่าร้อยละ 10.00 - 70.00 และ 27.27 - 55.55 ตามลำดับ

ความสอดคล้องคู่สุดท้ายคือ ระหว่างการใช้ดัชนี β_{SIB} และการทดสอบค่า Z ซึ่งมีค่าทั้งที่แตกต่างและใกล้เคียงกับคู่ α_{MH} และ MH-χ^2 มีค่าร้อยละในแต่ละความยาวแบบสอบอยู่ในช่วง 7.69-75.00 5.55-54.54 7.14-75.00 และในช่วงร้อยละ 5.55-61.54 ตามลำดับความยาวของแบบสอบ

ตารางที่ 47 จำนวนข้อสอบลำเอียงและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันในวิชาภาษาอังกฤษระหว่างการใช้เกณฑ์และการทดสอบนัยสำคัญของค่าสถิติภายในวิธีเดียวกัน ความยาวแบบสอบ 50 ข้อ*

ขนาดผู้สอบ	SA - χ^2	UA - χ^2	α_{MH} - χ^2_{MH}	β_{SIB} - Z_{SIB}
100	(19,4) (4, 21.05)	17.4 (4, 23.53)	19,3 (3, 15.79)	18,5 (5, 27.78)
200	9,2 (2, 22.22)	15,2 (2, 13.33)	17,4 4 (23.53)	19,7 (7, 36.84)
400	16,26 (12, 40.00)	30,26 (20, 55.55)	8,6 (6, 75.00)	10,6 (6, 60.00)
600	12,12 (5, 26.31)	17,12 (7, 31.82)	5,7 (4, 50.00)	6,8 (6, 75.00)
800	17,40 (13, 29.54)	29,40 (25, 58.14)	4,9 (4, 44.44)	6,11 (1, 7.69)
1,000	15,44 (14, 31.11)	32,44 (30, 65.22)	4,14 (4, 28.57)	4,12 (4, 33.33)

* แถวบนคือ จำนวนข้อสอบลำเอียงในแต่ละวิธี
แถวล่าง (ในวงเล็บ) คือ จำนวนข้อสอบที่ลำเอียงซ้ำกันและร้อยละเมื่อเทียบกับข้อสอบลำเอียงทั้งหมดที่พบ

ตารางที่ 48 จำนวนข้อสอบลำเอียงและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันในวิชาภาษาอังกฤษระหว่างการใช้เกณฑ์และการทดสอบนัยสำคัญของค่าสถิติภายในวิธีเดียวกัน ความยาวแบบสอบ 60 ข้อ*

ขนาดผู้สอบ	SA - χ^2	UA - χ^2	α_{MH} - χ^2_{MH}	β_{SIB} - Z_{SIB}
100	15,2 (2,13.33)	14,2 (2,14.28)	26,4 (4, 15.38)	23,3 (3, 13.04)
200	10,2 (1, 9.09)	14,2 1, 6.67	16,2 (5, 31.25)	24,7 (7, 29.17)
400	15,9 (4, 20.00)	15,9 (5, 25.00)	8,7 (7, 87.50)	10,7 (6, 54.54)
600	12,9 (5, 31.25)	17,9 (5, 23.81)	5,7 (4, 50.00)	5,7 (4, 50.00)
800	17,34 (11, 28.20)	28,34 (16, 34.78)	5,9 (5, 55.55)	5,14 (1, 5.55)
1,000	8,20 (5, 21.74)	15,20 (7, 25.00)	3,13 (3, 23.08)	2,17 (2, 11.76)

* แถวบนคือ จำนวนข้อสอบลำเอียงในแต่ละวิธี
แถวล่าง (ในวงเล็บ) คือ จำนวนข้อสอบที่ลำเอียงซ้ำกันและร้อยละเมื่อเทียบกับข้อสอบลำเอียงทั้งหมดที่พบ

ตารางที่ 49 จำนวนข้อสอบลำเอียงและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันในวิชาภาษาอังกฤษระหว่างการใช้เกณฑ์และการทดสอบนัยสำคัญของค่าสถิติภายในวิธีเดียวกัน ความยาวแบบสอบ 70 ข้อ*

ขนาดผู้สอบ	SA - χ^2	UA - χ^2	α_{MH} - χ^2_{MH}	β_{SIB} - Z_{SIB}
100	16,5 (3, 16.67)	18,5 (3, 15.00)	32,7 (6, 18.10)	28,2 (2, 7.14)
200	15,2 (1, 6.25)	21,2 (1, 4.54)	20,4 (4, 20.00)	18,6 (5, 26.31)
400	20,12 (5, 18.52)	19,12 (6, 24.00)	8,6 (6, 75.00)	13,9 (9, 69.23)
600	20,10 (5, 20.00)	25,10 (6, 20.69)	8,9 (7, 10.00)	6,8 (6, 75.00)
800	29,44 (19, 19.54)	36,44 (24, 42.86)	3,11 (3, 27.27)	4,14 (4, 23.53)
1,000	24,38 15, 31.91	37,38 20, 36.36	2,17 (2, 11.76)	2,18 (2, 11.11)

* แถวบนคือ จำนวนข้อสอบลำเอียงในแต่ละวิธี
แถวล่าง (ในวงเล็บ) คือ จำนวนข้อสอบที่ลำเอียงซ้ำกันและร้อยละเมื่อเทียบกับข้อสอบลำเอียงทั้งหมดที่พบ

ตารางที่ 50 จำนวนข้อสอบลำเอียงและร้อยละของข้อสอบที่ลำเอียงสอดคล้องกันในวิชาภาษาอังกฤษระหว่างการเกณฑ์และการทดสอบนัยสำคัญของค่าทางสถิติภายในวิธีเดียวกัน ความยาวแบบสอบ 80 ข้อ*

ขนาดผู้สอบ	SA - χ^2	UA - χ^2	α_{MH} - χ^2_{MH}	β_{SIB} - Z_{SIB}
100	40,2 (2, 5.00)	50,2 (2, 4.00)	34,4 (4, 11.76)	40,10 (9, 21.95)
200	32,4 (1, 2.86)	49,4 (3, 6.00)	25,5 (5, 20.00)	20,4 (4, 20.00)
400	34,9 (4, 10.26)	41,9 (7, 16.28)	9,7 (6, 60.00)	12,9 (8, 61.54)
600	30,13 (7, 19.44)	41,13 (9, 20.00)	8,9 (7, 70.00)	8,13 (8, 61.54)
800	31,12 (7, 19.44)	37,12 (7, 16.67)	3,10 (3, 30.00)	4,15 (1, 5.55)
1,000	27,26 (12, 29.27)	38,26 (16, 33.33)	4,17 (4, 23.53)	3,17 (3, 17.65)

* แถวบนคือ จำนวนข้อสอบลำเอียงในแต่ละวิธี
แถวล่าง (ในวงเล็บ) คือ จำนวนข้อสอบที่ลำเอียงซ้ำกันและร้อยละเมื่อเทียบกับข้อสอบลำเอียงทั้งหมดที่พบ

ตอนที่ 4 ผลการเปรียบเทียบลักษณะของความลำเอียงของข้อสอบที่มีต่อเพศของผู้สอบในวิชาคณิตศาสตร์และวิชาภาษาอังกฤษ

จากการพิจารณาค่าดัชนีต่าง ๆ ที่ได้จากการวิเคราะห์ระหว่างผู้สอบเพศหญิงและชาย พบว่าค่าดัชนีที่ได้มีค่าแตกต่างกันไป เมื่อขนาดผู้สอบและความยาวแบบสอบแตกต่างกัน ส่วนใหญ่พบว่าข้อสอบที่ลำเอียงในความยาวแบบสอบหนึ่ง อาจไม่ใช่ข้อสอบลำเอียงในอีกความยาวหนึ่ง และภายใต้ความยาวแบบสอบเดียวกันข้อสอบที่ลำเอียงในขนาดผู้สอบหนึ่งไม่แน่ว่าจะเป็นข้อลำเอียงในอีกขนาดผู้สอบหนึ่งหรือไม่ ค่าดัชนีที่ได้แต่ละตัวมีลักษณะคล้ายกันในประเด็นที่ว่ามีค่าไม่คงที่ แต่จะเปลี่ยนไปเมื่อเปลี่ยนขนาดผู้สอบหรือความยาวแบบสอบ นอกจากนี้ความเปลี่ยนแปรมีลักษณะขึ้น ๆ ลง ๆ ไม่สามารถทำนายได้ และไม่สามารถสรุปออกมาเป็นหลักการได้ เมื่อพิจารณาถึงความลำเอียงในลักษณะด่อด้านหรือเข้าข้างกลุ่มใดกลุ่มหนึ่ง พบว่ามีข้อสอบหลายข้อที่เปลี่ยนไปโดยอาจมีทิศทางการลำเอียงเข้าข้างกลุ่มหนึ่งในสถานการณ์หนึ่ง (เช่น SA อาจมีค่าเป็นลบ) แต่อาจกลับไปเข้าข้างอีกกลุ่ม (เช่นมีค่าเป็นบวก) ในอีกสถานการณ์หนึ่งก็ได้

ในการวิจัยครั้งนี้ ดัชนีที่มีค่าแสดงถึงความลำเอียงเข้าข้างกลุ่มผู้สอบกลุ่มใดกลุ่มหนึ่งได้นั้นมีอยู่ 3 ตัว ได้แก่ ดัชนี SA จากวิธี IRT-2 พารามิเตอร์ ดัชนี α_{MH} และดัชนี β_{SIB} จากวิธี SIBTEST

ลักษณะการนำเข้าของข้อมูลของผู้วิจัยในการวิเคราะห์ค่าดัชนีทั้ง 3 ตัว ระหว่างผู้สอบเพศชายและหญิงจะเป็นดังนี้

ดัชนี SA

ในวิชาคณิตศาสตร์ ดัชนี SA ที่เป็นบวก แสดงว่าเข้าข้างเพศหญิงมากกว่าเพศชาย และวิชาภาษาอังกฤษข้อสอบที่มีค่าดัชนี SA เป็นบวกแสดงว่าเข้าข้าง (favour) ผู้สอบเพศชายมากกว่าผู้สอบเพศหญิง และ

ดัชนี α_{MH}

ในวิชาคณิตศาสตร์ข้อสอบที่มีค่าดัชนี $\alpha_{MH} > 1.00$ แสดงว่าข้อสอบนั้นเข้าข้างผู้สอบเพศชายมากกว่าเพศหญิง แต่ในวิชาภาษาอังกฤษข้อสอบที่มีค่าดัชนี $\alpha_{MH} > 1.00$ แสดงว่าข้อสอบนั้นเข้าข้างเพศหญิงมากกว่าเพศชาย

ดัชนี β_{SIB}

ในวิชาภาษาอังกฤษข้อสอบที่มีค่าดัชนี β_{SIB} เป็นบวก แสดงว่าเข้าข้างเพศหญิงมากกว่าเพศชาย และสำหรับในวิชาคณิตศาสตร์ ข้อสอบที่มีค่าดัชนี β_{SIB} เป็นบวก แสดงว่าข้อสอบนั้นเข้าข้างเพศชายมากกว่าเพศหญิง

ความหมายของ "ข้อสอบเข้าข้าง" ผู้สอบเพศใด หมายความว่า ผู้สอบเพศนั้นจะได้รับประโยชน์หรือมีโอกาสจะตอบข้อสอบข้อนั้น ๆ ได้ถูกมากกว่าอีกเพศหนึ่ง

ตารางที่ 85-90 ในภาคผนวก ช แสดงถึงความลำเอียงของข้อสอบเมื่อพิจารณาว่ากลุ่มผู้สอบกลุ่มใด จะได้รับประโยชน์จากข้อสอบที่ลำเอียงโดยจะนำเสนอผลจากการใช้ดัชนี UA α_{MH} และ β_{SIB} ซึ่งแสดงทิศทางของการลำเอียงด้วย มีรายละเอียด ดังนี้

ก. ผลจากดัชนี SA

2. วิชาคณิตศาสตร์

สำหรับผลการวิเคราะห์วิชาคณิตศาสตร์เสนอไว้ในตารางที่ 85 ในภาคผนวก ช จะเห็นว่าข้อสอบส่วนใหญ่มีความลำเอียงเข้าข้างเพศชายเป็นส่วนมาก โดยเฉพาะอย่างยิ่งในขนาดผู้สอบ 100 คน และ 200 คน แทบไม่มีข้อที่ลำเอียงเข้าข้างเพศหญิง โดยที่มีข้อสอบลำเอียงเข้าข้างเพศหญิงในความยาวแบบสอบ 20 ข้อ สำหรับขนาดผู้สอบ 100 200 400 600 และ 800 คน ร้อยละ 0.00 และในขนาดผู้สอบ 1,000 คน ร้อยละ 11.11 ในขณะที่ลำเอียงเข้าข้างผู้สอบชายร้อยละ 100.00 เท่ากัน ในขนาดผู้สอบ 100 200 400 600 และ 800 คน และร้อยละ 88.89 สำหรับขนาดผู้สอบ 1,000 คน ในความยาวแบบสอบ 30 ข้อ มีข้อสอบลำเอียงเข้าข้างผู้สอบหญิงร้อยละ 0.00 0.00 10.00 9.09 0.00 และ 7.14 และเข้าข้างผู้สอบชายร้อยละ 100.00 100.00 90.01 100.00 และ 95.86 และในความยาวแบบสอบ 40 ข้อ มีข้อสอบลำเอียงเข้าข้างผู้สอบหญิงร้อยละ 7.69 0.00 0.00 0.00 0.00 และ 3.57 ลำเอียงเข้าข้างผู้สอบชายร้อยละ 92.31 100.00 100.00 100.00 100.00 และ 96.43 ตามลำดับ

2. วิชาภาษาอังกฤษ

จากการใช้ดัชนี SA สำหรับแบบสอบวิชาภาษาอังกฤษในตารางที่ 86 ในภาคผนวก ช พบว่าข้อสอบลำเอียงในวิชาภาษาอังกฤษ ส่วนใหญ่ลำเอียงเข้าข้างผู้สอบเพศหญิงมากกว่าเพศชายโดยมีข้อสอบลำเอียงเข้าข้างเพศหญิงเมื่อใช้ขนาดผู้สอบ 100 คน เรียงลำดับจากความยาวแบบสอบ 50-80 ข้อ ร้อยละ 78.95 33.33 50.00 และ 55.00 เข้าข้างเพศชายร้อยละ 21.05 66.67 50.00 และ 45.00 ขนาดผู้สอบ 200 คน มีข้อสอบลำเอียงเข้าข้างเพศหญิงร้อยละ 55.55 70.00 80.00 และ 78.12 เข้าข้างเพศชายร้อยละ

44.44 30.00 20.00 และ 21.87 ขนาดผู้สอบ 400 คน ลำเอียงเข้าข้างเพศหญิง 87.50 66.67 80.00 และ 70.58 ลำเอียงเข้าข้างเพศชาย 12.50 33.33 20.00 และ 29.41 ขนาดผู้สอบ 600 คน ลำเอียงเข้าข้างเพศหญิงร้อยละ 100.00 66.67 85.00 และ 83.33 ลำเอียงเข้าข้างเพศชายร้อยละ 0.00 33.33 15.00 และ 16.67 ขนาดผู้สอบ 800 คน ลำเอียงเข้าข้างเพศหญิงร้อยละ 93.33 75.00 95.83 และ 85.10 ลำเอียงเข้าข้างผู้สอบชายร้อยละ 6.67 4.17 และ 14.80 ขนาดผู้สอบ 1000 คน ลำเอียงเข้าข้างเพศหญิงร้อยละ 82.36 94.11 89.66 และ 70.97 ลำเอียงเข้าข้างเพศชายร้อยละ 17.64 5.89 10.34 และ 29.03

จะสังเกตได้ว่า มีจำนวนข้อสอบลำเอียงเข้าข้างเพศชายมากว่าที่ลำเอียงเข้าข้างเพศหญิงอยู่ 1 กรณี ได้แก่ กรณีขนาดผู้สอบ 100 คน ที่ความยาวแบบสอบ 60 ข้อ อย่างไรก็ตาม พบว่า มีข้อสอบบางข้อที่ลำเอียงเข้าข้างผู้สอบหญิงในบางกรณี และลำเอียงเข้าข้างเพศชายในบางกรณี เช่น ในขนาดผู้สอบ 100 คน ข้อ 4 ลำเอียงเข้าข้างเพศหญิงในความยาวแบบสอบ 50 ข้อ และ 80 ข้อ ลำเอียงเข้าข้างเพศชายในความยาวแบบสอบ 60 ข้อ และไม่จัดว่าเป็นที่ลำเอียงในความยาวแบบสอบ 70 ข้อ ข้ออื่น ๆ ที่พบในลักษณะคล้าย ได้แก่ ข้อ 26 29 46 และ 75 ในขนาดผู้สอบ 100 คน และข้อ 47 ในขนาดผู้สอบ 200 คน เป็นต้น

ข. ผลของดัชนี α_{MH}

1. วิชาคณิตศาสตร์

เมื่อพิจารณาตารางที่ 87 ในภาคผนวก ช ซึ่งเป็นผลจากวิชาคณิตศาสตร์พบว่ามีลักษณะตรงข้าม คือ ข้อสอบลำเอียงเข้าข้างเพศชายจะมีมากกว่าที่เข้าข้างเพศหญิงในบางกรณี และเท่ากันในบางกรณีและยังพบว่ามีข้อสอบที่ลำเอียงเข้าข้างเพศชายในบางกรณี และเข้าข้างเพศหญิงในบางกรณีเช่นกัน เช่น ข้อ 29 ในขนาดผู้สอบ 100 คน และข้อ 10 ในขนาดผู้สอบ 400 คน เป็นต้น สำหรับร้อยละของข้อสอบลำเอียงเข้าข้างเพศหญิงในขนาดผู้สอบ 100 คน เรียงตามลำดับความยาวแบบสอบ 20 30 และ 40 ข้อ คือ ร้อยละ 33.33 40.00 และ 50.00 เข้าข้างผู้สอบชายร้อยละ 66.67 60.00 และ 50.00 ขนาดผู้สอบ 200 คน เข้าข้างผู้สอบหญิงร้อยละ 25.00 50.00 และ 33.33 เข้าข้างผู้สอบชายร้อยละ 66.67 60.50 และ 66.67 ขนาด 400 คน เข้าผู้สอบหญิงร้อยละ 33.33 33.33 และ 37.50 เข้าข้างผู้สอบชายร้อยละ 66.67 66.67 และ 62.50 ขนาดผู้สอบ 600 คน เข้าข้างผู้สอบหญิงร้อยละ 50.50 33.33 และ 0.00 เข้าข้างผู้สอบชาย 50.00 66.67 และ 100.00 ขนาดผู้สอบ 800 คน ไม่มีข้อสอบลำเอียงเข้าข้างผู้สอบหญิงเลยในขณะที่ลำเอียงเข้าข้างผู้สอบชายร้อยละ 100.00 ส่วนในขนาดผู้สอบ 1,000 คน ไม่มีข้อสอบข้อใดได้รับการตัดสินว่าลำเอียง

2. วิชาภาษาอังกฤษ

จากตารางที่ 88 ในภาคผนวก ช จะเห็นว่าเมื่อใช้ดัชนี α_{MH} ผลการตัดสินข้อสอบลำเอียงวิชาภาษาอังกฤษเมื่อใช้เกณฑ์เดียวกัน ข้อสอบลำเอียงส่วนมากจะลำเอียงเข้าข้างผู้สอบหญิงมากกว่าผู้สอบชายที่มีความสามารถในระดับเดียวกัน ทั้งนี้ร้อยละของข้อสอบลำเอียงเข้าข้างเพศชายในขนาดผู้สอบ 100 คน เรียงลำดับตามความยาวแบบสอบเป็นร้อยละ 42.10 38.46 44.44 และ 47.06 ข้อสอบเข้าข้างเพศหญิงมีจำนวนร้อยละ 57.90 61.54 55.56 และ 52.94 ในขนาดผู้สอบ 200 คน ข้อสอบลำเอียงเข้าข้างเพศชายมีจำนวนร้อยละ 41.18 33.33 36.82 และ 37.50 ข้อสอบเข้าข้างเพศหญิงร้อยละ 58.52 66.67 63.16 และ 62.50 ในขนาดผู้สอบ 400 คน ข้อสอบลำเอียงเข้าข้างเพศชายร้อยละ 25.00 25.00 37.50 และ 33.33 เข้าข้างเพศหญิงร้อยละ 75.00 75.00 62.50 และ 66.67 ขนาดผู้สอบ 600 คน ข้อสอบเข้าข้างผู้สอบชายร้อยละ 20.00 20.00 25.00 และ 37.50 เข้าข้างผู้สอบหญิงร้อยละ 80.00 80.00 75.00 และ 62.50 ขนาดผู้สอบ 800 คน เข้าข้างเพศชายร้อยละ 50.00 40.00 33.33 และ 33.33 เข้าข้างเพศหญิงร้อยละ 50.00 60.00 66.67 และ 66.67 และในขนาดผู้สอบ 1000 คน มีข้อสอบลำเอียงเข้าข้างเพศชายร้อยละ 25.00 33.33 50.00 และ 25.00 และลำเอียงเข้าข้างเพศหญิงร้อยละ 75.00 66.67 50.00 และ 75.00 ตามลำดับ ความยาวแบบสอบ 50 60 70 และ 80 ข้อ

เมื่อพิจารณาในรายละเอียดพบว่ามีข้อสอบหลายข้อที่บางกรณีลำเอียงเข้าข้างเพศหญิงบางกรณีเข้าข้างเพศชาย เช่น ข้อ 3 เป็นข้อสอบลำเอียงเข้าข้างเพศชายในบางกรณีความยาวแบบสอบ 50 และ 70 ข้อ แต่เข้าข้างเพศหญิงกรณีความยาว 60 ข้อ และเป็นข้อสอบไม่ลำเอียงกรณีความยาว 80 ข้อ ข้ออื่น ๆ ที่พบ เช่น ข้อ 8, 11, 45, 46 และ 92 เป็นต้น

อย่างไรก็ตามเมื่อพิจารณาในภาพรวมพบว่า ความแตกต่างระหว่างจำนวนข้อสอบลำเอียงเข้าข้างเพศหญิงและชายไม่สูงนัก

ค. ผลจากดัชนี β_{SIB}

1. วิชาคณิตศาสตร์

ข้อสอบลำเอียงในวิชาคณิตศาสตร์ส่วนมากลำเอียงเข้าข้างผู้สอบชายเช่นเดียวกับที่พบในการใช้ดัชนี α_{MH} และพบว่ามีบางข้อเช่นกันที่ลำเอียงเข้าข้างผู้สอบหญิงในขนาดผู้สอบหนึ่งและเข้าข้างผู้สอบชายในอีกขนาดหนึ่ง เช่น ข้อ 2 5 และ 33 ในขนาดผู้สอบ 100 คน

สำหรับร้อยละของข้อสอบลำเอียงเมื่อเปรียบเทียบกับจำนวนข้อสอบที่ลำเอียงทั้งหมดในแต่ละขนาดผู้สอบจำแนกตามความยาวแบบสอบ 20 30 และ 40 ข้อ พบว่าในขนาดผู้สอบ 100 คน มีข้อสอบลำเอียงเข้าข้างผู้สอบหญิงร้อยละ 25.00 69.23 และ 34.78 เข้าข้างผู้สอบชายร้อยละ 75.00 30.77 และ 65.22 ในขนาดผู้สอบ 200 คน เข้าข้างผู้สอบหญิงร้อยละ 50.00 27.27 และ 22.22 เข้าข้างผู้สอบชายร้อยละ 50.00 72.73 และ 77.78 ขนาดผู้สอบ 400 คน ลำเอียงเข้าข้างผู้สอบชายร้อยละ 16.67 42.86 และ 25.00 ลำเอียงเข้าข้างผู้สอบชายร้อยละ 83.33 57.14 และ 75.25 ขนาดผู้สอบ 600 คน ลำเอียงเข้าข้างผู้สอบหญิงร้อยละ 50.00 33.33 และ 25.00 เข้าข้างผู้สอบชายร้อยละ 50.00 66.67 และ 75.00 ขนาดผู้สอบ 800 คน ลำเอียงเข้าข้างผู้สอบหญิงร้อยละ 25.00 33.33 และ 0.00 เข้าข้างผู้สอบเพศชายร้อยละ 75.00 66.67 และ 100.00 และในขนาดผู้สอบ 1,000 คน มีข้อสอบลำเอียงเข้าข้างผู้สอบหญิงร้อยละ 100.00 ทั้งในความยาวแบบสอบ 20 และ 30 ข้อ และร้อยละ 66.67 ในความยาว 40 ข้อ ไม่มีข้อใดลำเอียงเข้าข้างผู้สอบชายในความยาวแบบสอบ 20 และ 30 ข้อ แต่ลำเอียงเข้าข้างผู้สอบชายร้อยละ 33.33 ในความยาวแบบสอบ 40 ข้อ รายละเอียดดังตารางที่ 89 ในภาคผนวก ช

2. วิชาภาษาอังกฤษ

กรณีของผลจากดัชนี B_{SIB} ในตารางที่ 90 ในภาคผนวก ช พบผลตรงกันข้ามกับที่พบในการใช้ดัชนี α_{MH} คือ ข้อสอบลำเอียงวิชาภาษาอังกฤษส่วนมากลำเอียงเข้าข้างผู้สอบชายมากกว่าหญิงและพบข้อสอบบางข้อที่บางครั้งลำเอียงเข้าข้างผู้สอบหญิง บางครั้งลำเอียงเข้าข้างผู้สอบชายเช่นกัน เช่น ข้อ 3 ข้อ 11 ข้อ 21 และข้อ 82 ในขนาดผู้สอบ 100 คน เป็นต้น สำหรับร้อยละของข้อสอบลำเอียงเข้าข้างผู้สอบชายสำหรับขนาดผู้สอบ 100 คน เรียงลำดับจากความยาวแบบสอบ 50-80 ข้อ คือ 68.75 52.38 72.41 และ 78.38 ที่เข้าข้างผู้สอบหญิงมีร้อยละ 31.25 47.62 27.59 และ 21.62 ในขนาดผู้สอบ 200 คน เข้าข้างผู้สอบชายร้อยละ 68.42 43.48 52.63 และ 63.16 เข้าข้างผู้สอบหญิงร้อยละ 31.58 56.52 47.37 และ 36.84 ในขนาดผู้สอบ 400 คน เข้าข้างผู้สอบชายร้อยละ 50.00 54.54 57.14 และ 57.14 เข้าข้างผู้สอบหญิงร้อยละ 50.00 45.45 42.82 และ 57.14 ขนาดผู้สอบ 600 คน ข้อสอบลำเอียงเข้าข้างผู้สอบชายร้อยละ 66.67 40.00 50.00 และ 62.50 เข้าข้างผู้สอบหญิงร้อยละ 33.33 60.00 50.00 และ 37.00 ขนาดผู้สอบ 800 คน เข้าข้างร้อยละ 60.00 66.67 60.00 และ 50.00 เข้าข้างผู้สอบหญิงร้อยละ 40.00 33.33 40.00 และ 50.00 และในขนาดผู้สอบ 1,000 คน มีข้อสอบลำเอียงเข้าข้างผู้สอบชายร้อยละ 50.00 เท่ากัน ในความยาวแบบสอบ 50 60 และ 70 ข้อ และร้อยละ 66.67 ในความยาวข้อสอบ 80 ข้อ ลำเอียงเข้าข้างผู้สอบหญิงร้อยละ 50.00 ในความยาวแบบสอบ 50 ข้อ และ 70 ข้อ และ 33.33 ในความยาวแบบสอบ 80 ข้อ