

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบอำนาจการตรวจสอบและอัตราความคลาดเคลื่อนประเภทที่ 1 ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่าง วิธีการวิเคราะห์องค์ประกอบจำกัด (RFA) วิธีแมนเทิล-แฮนส์เซล (MH) และวิธีการตอบสนองข้อสอบ (IRT) แบบ 2 พารามิเตอร์ เมื่อมีการจัดกลุ่มค่าความยากของข้อสอบออกเป็น 3 ระดับ คือ กลุ่มข้อสอบที่มีค่าความยากสูง ปานกลาง และต่ำ ค่าอำนาจจำแนกข้อสอบจัดเป็น 3 ระดับ คือ กลุ่มข้อสอบที่มีค่าอำนาจจำแนกสูง ปานกลาง และต่ำ กลุ่มตัวอย่างแบ่งออกเป็น 2 ขนาดคือ กลุ่มตัวอย่างขนาดเล็ก (300 คน) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) ความยาวของแบบสอบแบ่งเป็น 2 ขนาดคือ แบบสอบสั้น (25 ข้อ) และแบบสอบยาว (75 ข้อ) ความลำเอียงของข้อสอบแบ่งออกเป็น 2 ระดับคือ ลำเอียงสูง และลำเอียงต่ำ ข้อมูลที่ศึกษาได้จากการจำลองขึ้นด้วยโปรแกรม SPSS for Windows

ผลการวิเคราะห์ข้อมูลในบทนี้ได้นำเสนอเป็น 3 ตอนดังนี้

ตอนที่ 1 ผลการวิเคราะห์องค์ประกอบเพื่อหาความเป็นเอกมิติของแบบสอบในแต่ละเงื่อนไข

ตอนที่ 2 เปรียบเทียบอำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่าง วิธีการวิเคราะห์องค์ประกอบจำกัด (RFA) วิธีแมนเทิล-แฮนส์เซล (MH) และวิธีการตอบสนองข้อสอบ (IRT) แบบ 2 พารามิเตอร์

ตอนที่ 3 เปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธีการวิเคราะห์องค์ประกอบจำกัด (RFA) วิธีแมนเทิล - แฮนส์เซล (MH) และวิธีการตอบสนองข้อสอบ (IRT) แบบ 2 พารามิเตอร์

เพื่อความเข้าใจตรงกันเกี่ยวกับสัญลักษณ์ที่ใช้ในตารางผลการวิจัย ผู้วิจัยกำหนดให้สัญลักษณ์ดังกล่าวมีความหมายดังนี้

RFA คือ การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีการวิเคราะห์องค์ประกอบจำกัด

MH คือ การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล

IRT (Z-SA) คือ การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ ที่ใช้ดัชนีพื้นที่แบบคิดเครื่องหมาย และมีการทดสอบนัยสำคัญทางสถิติด้วยค่า Z ที่ระดับ .05 เป็นเกณฑ์ในการพิจารณา

IRT(Z-UN) คือ การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ ที่ใช้ดัชนีพื้นที่แบบไม่คิดเครื่องหมาย และมีการทดสอบนัยสำคัญทางสถิติด้วยค่า Z ที่ระดับ .05 เป็นเกณฑ์ในการพิจารณา

IRT(χ^2) คือ การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ ที่ใช้ดัชนีโคแอสควร์ของ Lord และมีการทดสอบนัยสำคัญทางสถิติที่ระดับ .05 เป็นเกณฑ์ในการพิจารณา

ตอนที่ 1 ผลการวิเคราะห์องค์ประกอบเพื่อตรวจสอบความเป็นเอกมิติของข้อมูลตารางที่ 7 อัตราส่วนของผลการวิเคราะห์องค์ประกอบแล้ว พบว่าแบบสอบมีความเป็นเอกมิติของกลุ่มช่างฉิ่งและกลุ่มแปรรูปเทียบในการจำลองข้อสอบ 10 ครั้ง

เดือนปีที่	ขนาดกลุ่ม	ลักษณะแบบสอบที่ใช้ศึกษา			กลุ่มช่างฉิ่ง	กลุ่มแปรรูปเทียบ
		ความยากแบบสอบ	ค่าความยาก (b)	ค่าอำนาจจำแนก (a)		
1	160 คน	25 %	สูง	สูง	0	0
2				ปานกลาง	0.40	0.40
3				ต่ำ	0.40	0.40
4			ปานกลาง	สูง	1.00	1.00
5				ปานกลาง	0.20	0.20
6				ต่ำ	0.20	0.20
7			ต่ำ	สูง	0.40	0.40
8				ปานกลาง	0.40	0.40
9				ต่ำ	0	0
10		75 %	สูง	สูง	0	0
11				ปานกลาง	0	0
12				ต่ำ	0.20	0.20
13			ปานกลาง	สูง	1.00	1.00
14				ปานกลาง	0.40	0.40
15				ต่ำ	0	0
16			ต่ำ	สูง	0.80	0.80
17				ปานกลาง	0.60	0.60
18				ต่ำ	0.60	0.60

ตารางที่ 7 (ต่อ)

เดือนไฮ ที่	ขนาดผู้ สอบ	ลักษณะแบบสอบที่ศึกษา			กลุ่มอ้างอิง	กลุ่มเปรียบเทียบ
		ความยาวแบบสอบ	จำนวนฉาก (b)	ค่าสัมประสิทธิ์ (a)		
19	500 คน	25 ข้อ	สูง	สูง	0	0
20				ปานกลาง	0	0
21				ต่ำ	0	0
22			ปานกลาง	สูง	0.80	0.80
23				ปานกลาง	0.60	0.60
24				ต่ำ	0.20	0.20
25			ต่ำ	สูง	0.80	0.80
26				ปานกลาง	0.60	0.60
27				ต่ำ	0.20	0.20
28		75 ข้อ	สูง	สูง	0	0
29				ปานกลาง	0	0
30				ต่ำ	0	0
31			ปานกลาง	สูง	1.00	1.00
32				ปานกลาง	1.00	1.00
33				ต่ำ	1.00	1.00
34			ต่ำ	สูง	1.00	1.00
35				ปานกลาง	1.00	1.00
36				ต่ำ	0.80	0.80

จากตารางที่ 7 เป็นการนำเสนออัตราส่วนผลการวิเคราะห์องค์ประกอบของแบบสอบในแต่ละเดือนไฮที่ทำการศึกษา ทั้งหมด 36 เดือนไฮ จากการจำลองข้อมูล 10 ครั้ง เพื่อทดสอบความเป็นเอกมิติของแบบสอบ ซึ่งเป็นข้อตกลงเบื้องต้นของการตรวจสอบด้วยวิธีการทดสอบองค์ประกอบ หากพบว่า เดือนไฮใดมีอัตราส่วนทั้งในกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็น 0 หมายถึง จากการจำลองข้อมูลทั้งหมด 10 ครั้งเมื่อนำข้อมูลไปวิเคราะห์องค์ประกอบไม่พบว่า แบบสอบในเดือนไฮนั้นมีค่าไอเกนขององค์ประกอบที่ 1 มากกว่าค่าไอเกนขององค์ประกอบที่ 2 เป็นสามเท่า จากการวิเคราะห์องค์ประกอบทั้งหมด 10 ครั้ง แสดงว่าเดือนไฮนั้นมีลักษณะข้อมูลไม่สอดคล้องกับข้อตกลงเบื้องต้นและไม่สามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีการทดสอบองค์ประกอบได้ ในทางตรงกันข้าม ในเดือนไฮที่มีอัตราส่วนผลการวิเคราะห์องค์ประกอบเท่ากับ 1.00 แสดงว่า เดือนไฮนั้นเมื่อนำข้อมูลไปวิเคราะห์องค์ประกอบทั้งหมด 10 ครั้ง พบว่ามีค่าไอเกนขององค์ประกอบที่ 1 มากกว่าค่าไอเกนขององค์ประกอบที่ 2 เป็นสามเท่าทั้งหมด 10 ครั้ง

ความเป็นเอกมิติของข้อมูล (Unidimensionality) หมายความว่า การตอบข้อสอบทุกข้อ ได้ถูกต้อง ต้องใช้ความสามารถหรือคุณลักษณะแฝงภายในเพียงคุณลักษณะเดียว หรือข้อสอบทุกข้อในแบบสอบวัดคุณลักษณะแฝงลักษณะเดียว หรือกลุ่มลักษณะแฝงลักษณะเดียว ซึ่งเป็นข้อตกลงเบื้องต้นในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีการตอบสนองข้อสอบ (วรมุข แหนมแสง ,2537)

ตอนที่ 2 เปรียบเทียบอำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ จากการตรวจสอบด้วยวิธีการวิเคราะห์องค์ประกอบจำกัด วิธีแมนเทิล - แอนส์เซด และวิธีการตอบสนองข้อสอบ

เป็นการนำเสนออำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ จากการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีการวิเคราะห์องค์ประกอบจำกัด (RFA) วิธีแมนเทิล-แอนส์เซด (MH) และวิธีการตอบสนองข้อสอบ (IRT) แบบ 2 พารามิเตอร์ ตามเงื่อนไขที่กำหนด ดังแสดงในตารางที่ 8-11

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 8 จำนวนการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน)

เงื่อนไข	ความยาก (b)	อำนาจจำแนก (a)	ขนาดความลำเอียง	RFA	MH	IRT (Z-SA)	IRT (Z-UN)	IRT (χ^2)
1		สูง	สูง	**	0.050	*	*	*
			ต่ำ	**	0.050	*	*	*
2	สูง	ปานกลาง	สูง	1.000	0.450	0.800	0.800	0
			ต่ำ	1.000	0.050	0	0	0
3		ต่ำ	สูง	1.000	0.500	0.800	0.800	0.600
			ต่ำ	1.000	0.350	0	0.100	0
4	ปานกลาง	สูง	สูง	1.000	0.500	0.360	0.440	0.160
			ต่ำ	1.000	0.050	0	0	0
5	กลาง	ปานกลาง	สูง	1.000	0.750	1.000	1.000	1.000
			ต่ำ	1.000	0.050	0	0	0
6	ปานกลาง	ต่ำ	สูง	1.000	0.250	1.000	1.000	1.000
			ต่ำ	1.000	0.100	0.200	0.200	0
7		สูง	สูง	1.000	0.150	0.600	0.600	0
			ต่ำ	1.000	0	0	0	0
8	ต่ำ	ปานกลาง	สูง	1.000	0.750	0.500	0.700	0
			ต่ำ	1.000	0.050	0	0	0
9		ต่ำ	สูง	1.000	0.500	*	*	*
			ต่ำ	1.000	0.500	*	*	*

* เงื่อนไขนั้นไม่เป็นไปตามข้อตกลงเบื้องต้น

** ข้อมูลไม่เหมาะในการตรวจสอบเนื่องจากไม่มี variation ระหว่างข้อมูล

จากตารางที่ 8 เป็นภาพนำเสนอผลการเปรียบเทียบอำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) รายละเอียดมีดังนี้

เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี IRT ทั้ง สามวิธี พบว่า เฉพาะกรณีข้อสอบมีความลำเอียงสูง ความยากปานกลาง อำนาจจำแนกปานกลาง และอำนาจจำแนกต่ำ วิธี IRT ทั้ง สามวิธีจะมีอำนาจการตรวจสอบสูงที่สุดเท่าเทียมกัน กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีข้อสอบมีความลำเอียงสูง ที่มีความยากปานกลาง อำนาจจำแนกปานกลางและอำนาจจำแนกต่ำ วิธี IRT (Z-UN) วิธี IRT(Z-SA) และวิธี IRT(χ^2) จะสามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ถูกต้อง 100 เปอร์เซ็นต์ แต่โดยภาพรวม กล่าวได้ว่า วิธี IRT (Z-UN) มีอำนาจการตรวจสอบสูงที่สุดรองลงมาคือวิธี IRT(Z-SA) และวิธี IRT(χ^2) ตามลำดับ ในกรณีนี้ผู้วิจัยจึงใช้ วิธี IRT (Z-UN) เป็นตัวแทนในกลุ่มวิธี IRT แบบ 2 พารามิเตอร์

สำหรับกลุ่มข้อสอบที่มีความลำเอียงสูง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีข้อสอบมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT(Z-UN) และวิธี MH ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความลำเอียงต่ำ เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีข้อสอบมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากสูง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่ข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีข้อสอบมีความยากสูงและมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ

เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีความยากสูง และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ)

และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากสูง และมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากปานกลาง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่มีข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากปานกลาง และมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ

เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีความยากปานกลาง และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากปานกลาง และมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากต่ำ เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่มีข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด ในขณะที่วิธี IRT (Z-UN) และวิธี MH มีอำนาจการตรวจสอบเท่าเทียมกัน กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากต่ำ และมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด ในขณะที่วิธี IRT (Z-UN) และวิธี MH จะสามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ถูกต้องพอ ๆ กัน

เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีความยากต่ำ และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากต่ำ และมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

เมื่อพิจารณาจากค่าความยากทั้ง 3 ระดับ พบจะสรุปได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุดทุกเงื่อนไขที่สามารถวิเคราะห์ได้ ในขณะที่วิธี MH และวิธี IRT (Z-UN) จะมีอำนาจการตรวจสอบสูงที่สุดในกรณีที่มีข้อสอบมีความยากปานกลาง

เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีอำนาจจำแนกต่ำ และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT.(Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกต่ำ และมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สรุป เมื่อพิจารณาจากค่าอำนาจจำแนกทั้ง 3 ระดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุดทุกเงื่อนไข ในขณะที่วิธี MH และวิธี IRT (Z-UN) จะมีอำนาจการตรวจสอบสูงที่สุด เฉพาะกรณีที่ข้อสอบมีค่าอำนาจจำแนกปานกลาง

กล่าวโดยสรุป เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุดทั้งในกรณีที่ข้อสอบมีความลำเอียงสูงและลำเอียงต่ำ วิธี IRT(Z-UN) จะตรวจสอบได้ดีกว่าวิธี MH เฉพาะในกรณีที่ข้อสอบมีความลำเอียงสูง แต่ถ้าข้อสอบที่มีความลำเอียงต่ำ วิธี MH จะตรวจสอบได้ดีกว่า วิธี IRT (Z-UN) ดังนั้น ข้อค้นพบนี้จึงไม่สอดคล้องกับสมมุติฐานข้อที่ 2 ที่กล่าวว่า เมื่อกลุ่มตัวอย่างมีขนาดเล็ก ค่าความยากของข้อสอบและค่าอำนาจจำแนกของข้อสอบต่างกัน ภายใต้ความยาวแบบสอบเดียวกันวิธี MH จะมีอำนาจการตรวจสอบสูงกว่าวิธี RFA และวิธี IRT แบบ 2 พารามิเตอร์ และมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าวิธี RFA และวิธี IRT แบบ 2 พารามิเตอร์ ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

ตารางที่ 9 จำนวนการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน)

เงื่อนไข	ความยาก (b)	อำนาจจำแนก (a)	ขนาดความลำเอียง	RFA	MH	IRT (Z-SA)	IRT (Z-UN)	IRT (χ^2)
10		สูง	สูง	**	0.050	*	*	*
			ต่ำ	**	0	*	*	*
11	สูง	ปานกลาง	สูง	**	0.200	*	*	*
			ต่ำ	**	0.100	*	*	*
12		ต่ำ	สูง	1.000	0.250	1.000	1.000	0.700
			ต่ำ	1.000	0.250	0.400	0.600	0
13		สูง	สูง	1.000	0.550	0.525	0.600	0.225
			ต่ำ	1.000	0.200	0.200	0.200	0.200
14	ปานกลาง	ปานกลาง	สูง	1.000	0.425	0	0.400	0
			ต่ำ	1.000	0.250	0	0	0
15		ต่ำ	สูง	1.000	0.400	*	*	*
			ต่ำ	1.000	0.200	*	*	*
16		สูง	สูง	1.000	0.275	0	0	0.050
			ต่ำ	1.000	0.050	0	0	0
17	ต่ำ	ปานกลาง	สูง	1.000	0.675	0.867	0.934	0.334
			ต่ำ	1.000	0.050	0	0	0
18		ต่ำ	สูง	1.000	0.800	0.300	0.667	0.334
			ต่ำ	1.000	0.100	0	0	0

- * เงื่อนไขนั้นไม่เป็นไปตามข้อตกลงเบื้องต้น
- ** ข้อมูลไม่เหมาะในการตรวจสอบเนื่องจากไม่มี variation ระหว่างข้อมูล

จากตารางที่ 9 เป็นกาแนะนำเสนอผลการเปรียบเทียบอำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) รายละเอียดที่พบมีดังนี้

เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี IRT ทั้ง ตามวิธี พบว่า วิธี IRT (Z-UN) มีอำนาจการตรวจสอบสูงที่สุดรองลงมาคือวิธี IRT(Z-SA) และวิธี IRT(χ^2) ตามลำดับ ในกรณีนี้ผู้วิจัยจึงใช้วิธี IRT (Z-UN) เป็นตัวแทนในกลุ่มวิธี IRT แบบ 2 พารามิเตอร์

สำหรับกลุ่มข้อสอบที่มีความลำเอียงสูง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีข้อสอบมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความลำเอียงต่ำ เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีข้อสอบมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากสูง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) พบว่า ในกรณีที่ข้อสอบมีความลำเอียงสูง วิธี RFA และวิธี IRT (Z-UN) มีอำนาจการตรวจสอบสูงที่สุดเท่ากัน ขณะเดียวกันทั้งวิธี RFA และวิธี IRT (Z-UN) ยังมีอำนาจการตรวจสอบสูงกว่าวิธี MH กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่ข้อสอบมีความยากสูง และมีความลำเอียงสูง วิธี RFA และวิธี IRT (Z-UN) จะมีอำนาจการตรวจสอบพอ ๆ กัน ขณะเดียวกันทั้งวิธี RFA และวิธี IRT (Z-UN) ยังมีอำนาจการตรวจสอบสูงกว่าวิธี MH

สำหรับกลุ่มข้อสอบที่มีความยากปานกลาง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่ข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่ข้อสอบมีความยากปานกลาง และมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT (Z-UN) ตามลำดับ

เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีความยากปานกลาง และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจ

ความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกปานกลาง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่ข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกปานกลาง และมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT (Z-UN) ตามลำดับ

แต่เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีอำนาจจำแนกปานกลาง และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกปานกลาง และมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกต่ำ เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่ข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกต่ำ และมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

แต่เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีอำนาจจำแนกต่ำ และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกต่ำ และมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สรุป เมื่อพิจารณาจากค่าอำนาจจำแนกทั้ง 3 ระดับ จะเห็นว่าเมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี RFA ยังคงมีอำนาจการตรวจสอบสูงที่สุดในทุกเงื่อนไข ในขณะที่วิธี MH จะมีอำนาจการตรวจสอบสูงที่สุดในกรณีที่ข้อสอบมีค่าอำนาจจำแนกต่ำ วิธี IRT (Z-UN) จะมีอำนาจการตรวจสอบสูงที่สุดในกรณีที่ข้อสอบมีค่าอำนาจจำแนกปานกลาง

กล่าวโดยสรุป โดยภาพรวมอาจกล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี RFA ยังคงมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ ดังนั้น ข้อค้นพบนี้จึงไม่สอดคล้องกับสมมติฐานข้อที่ 2 ที่กล่าวว่า เมื่อกลุ่มตัวอย่างมีขนาดเล็ก ค่าความยากของข้อสอบและค่าอำนาจจำแนกของข้อสอบต่างกัน ภายใต้อายุขยาแบบสอบเดียวกันวิธี MH จะมีอำนาจการตรวจสอบสูงกว่าวิธี RFA และวิธี IRT แบบ 2 พารามิเตอร์ และมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าวิธี RFA และวิธี IRT แบบ 2 พารามิเตอร์ ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ อีกเช่นกัน



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 10 จำนวนการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน)

เงื่อนไข	ความยาก (b)	อำนาจจำแนก (a)	ขนาดความต่ำเอียง	RFA	MH	IRT (Z-SA)	IRT (Z-UN)	IRT (χ^2)	
19	สูง	สูง	สูง	**	0	*	*	*	
			ต่ำ	**	0	*	*	*	
		ปานกลาง	สูง	**	0	*	*	*	
			ต่ำ	**	0	*	*	*	
		21	ต่ำ	สูง	**	0	*	*	*
				ต่ำ	**	0	*	*	*
22	สูง	ปานกลาง	สูง	1.000	0.900	0.700	0.750	0.550	
			ต่ำ	1.000	0.250	0.050	0.250	0	
23	ปานกลาง	ปานกลาง	สูง	1.000	1.000	0.467	0.733	0.600	
			ต่ำ	1.000	0.150	0.200	0.200	0	
24	ต่ำ	ปานกลาง	สูง	1.000	1.000	1.000	1.000	1.000	
			ต่ำ	1.000	0.150	0.800	0.800	0.200	
25	สูง	ปานกลาง	สูง	1.000	0.700	0.950	0.950	0.700	
			ต่ำ	1.000	0.100	0.200	0.200	0	
26	ต่ำ	ปานกลาง	สูง	1.000	1.000	1.000	1.000	1.000	
			ต่ำ	1.000	0.050	0	0	0	
27	ต่ำ	ปานกลาง	สูง	1.000	1.000	0.800	1.000	1.000	
			ต่ำ	1.000	0.150	0	0	0	

* เงื่อนไขนั้นไม่เป็นไปตามข้อตกลงเบื้องต้น

** ข้อมูลไม่เหมาะในการตรวจสอบเนื่องจากไม่มี variation ระหว่างข้อมูล

จากตารางที่ 10 เป็นการนำเสนอผลการเปรียบเทียบอำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) รายละเอียดมีดังนี้

เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี IRT ทั้ง สามวิธี พบว่า ถ้าข้อสอบมีความลำเอียงสูง เฉพาะกรณีที่มีข้อสอบมีความยากปานกลาง อำนาจจำแนกต่ำและในกรณีที่ข้อสอบมีความยากต่ำ อำนาจจำแนกปานกลางวิธี IRT ทั้ง 3 วิธีจะมีอำนาจการตรวจสอบเท่าเทียมกัน แต่โดยภาพรวมแล้วสรุปได้ว่า วิธี IRT (Z-UN) จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี IRT (Z-SA) และวิธี IRT(x^2) ตามลำดับ ในกรณีนี้ผู้วิจัยจึงใช้ วิธี IRT (Z-UN) เป็นตัวแทนในกลุ่มวิธี IRT แบบ 2 พารามิเตอร์

สำหรับกลุ่มข้อสอบที่มีความลำเอียงสูง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความลำเอียงต่ำ เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีความลำเอียงต่ำวิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากสูง พบว่า ข้อมูลที่จำลองขึ้นไม่เหมาะสำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT (Z-UN) กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีความยากสูง ข้อมูลไม่เหมาะสำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีทั้ง 5 วิธี

สำหรับกลุ่มข้อสอบที่มีความยากปานกลาง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่ข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีความยากปานกลางและมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

แต่เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีความยากปานกลาง และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจ

สอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ยืดสอบมีความยากปานกลาง และมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากต่ำ เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่ยืดสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ยืดสอบมีความยากต่ำและมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ

แต่เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีความยากต่ำ และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ยืดสอบมีความยากต่ำและมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สรุป เมื่อพิจารณาจากค่าความยากทั้ง 3 ระดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างมีขนาดใหญ่ (1000 คน) วิธี RFA ยังคงมีอำนาจการตรวจสอบสูงที่สุดในทุกเงื่อนไข วิธี IRT (Z-UN) จะมีอำนาจการตรวจสอบสูงที่สุดในกรณีที่ยืดสอบมีความยากต่ำ วิธี MH จะมีอำนาจการตรวจสอบสูงที่สุดในกรณีที่ยืดสอบมีความยากปานกลาง

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกสูง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่ยืดสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุดในขณะที่วิธี IRT (Z-UN) และวิธี MH มีอำนาจการตรวจสอบเท่าเทียมกัน กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ยืดสอบมีอำนาจจำแนกสูงและมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุดในขณะที่วิธี IRT (Z-UN) และวิธี MH จะสามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ถูกต้องพอ ๆ กัน

แต่เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีอำนาจจำแนกสูง และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างมีขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ยืดสอบมีอำนาจจำแนกสูงและมี

ความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี IRT (Z-UN) และวิธี MH ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกปานกลาง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่ข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่าเมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกปานกลาง และมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

แต่เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีอำนาจจำแนกปานกลาง และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุดในขณะที่วิธี IRT (Z-UN) และวิธี MH มีอำนาจการตรวจสอบเท่าเทียมกัน กล่าวได้ว่าเมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกปานกลางและมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุดในขณะที่วิธี IRT (Z-UN) และวิธี MH จะสามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ถูกต้องพอ ๆ กัน

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกต่ำ เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกรณีที่ข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA วิธี MH และวิธี IRT (Z-UN) มีอำนาจการตรวจสอบเท่าเทียมกัน กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกต่ำและมีความลำเอียงสูง ทั้ง 3 วิธีคือ วิธี RFA วิธี MH และวิธี IRT (Z-UN) จะมีอำนาจการตรวจสอบพอ ๆ กัน

แต่เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) ในกลุ่มข้อสอบที่มีอำนาจจำแนกต่ำ และมีความลำเอียงต่ำ พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุดในขณะที่วิธี IRT (Z-UN) และวิธี MH มีอำนาจการตรวจสอบเท่าเทียมกัน กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกต่ำ และมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุดในขณะที่วิธี IRT (Z-UN) และวิธี MH จะสามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ถูกต้องพอ ๆ กัน

เมื่อพิจารณาจากค่าอำนาจจำแนกทั้ง 3 ระดับ จะสรุปได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) วิธี RFA ยังคงมีอำนาจการตรวจสอบสูงที่สุดในทุกเงื่อนไข ในขณะที่วิธี MH และวิธี IRT ทั้ง 3 วิธี จะมีอำนาจการตรวจสอบสูงที่สุด เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกต่ำ

กล่าวโดยสรุป โดยภาพรวม วิธี RFA ยังคงมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT ทั้ง 3 วิธี แม้ว่าจะเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างจะมีขนาดใหญ่ (1000) ก็ตาม ดังนั้น ข้อค้นพบนี้จึงไม่สอดคล้องกับสมมติฐานข้อ 1 ที่กล่าวว่า เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ ค่าความยากของข้อสอบและค่าอำนาจจำแนกของข้อสอบต่างกัน ภายใต้ความยาวแบบสอบเดียวกัน วิธี RFA วิธี MH และวิธี IRT แบบ 2 พหาวมิติเตอร์ จะมีอำนาจการตรวจสอบและมีอัตราความคลาดเคลื่อนประเภทที่ 1 ไม่ต่างกันในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 11 อำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน)

เงื่อนไข	ความยาก (b)	อำนาจจำแนก (a)	ขนาดความลำเอียง	RFA	MH	IRT (Z-SA)	IRT (Z-UN)	IRT (χ^2)
28		สูง	สูง	**	0.025	*	*	*
			ต่ำ	**	0	*	*	*
		ปานกลาง	สูง	**	0.025	*	*	*
29		ปานกลาง	สูง	**	0.025	*	*	*
			ต่ำ	**	0	*	*	*
30		ต่ำ	สูง	**	0.025	*	*	*
			ต่ำ	**	0.050	*	*	*
31		สูง	สูง	1.000	1.000	0.740	0.660	0.660
			ต่ำ	1.000	0.150	0	0.040	0
32	ปานกลาง	ปานกลาง	สูง	1.000	1.000	0.220	0.400	0.520
			ต่ำ	1.000	0.350	0	0	0
33		ต่ำ	สูง	1.000	1.000	0.220	0.260	0.420
			ต่ำ	1.000	0.550	0.080	0.080	0
34		สูง	สูง	1.000	0.800	0.700	0.725	0.650
			ต่ำ	1.000	0.350	0.200	0.200	0.150
35	ต่ำ	ปานกลาง	สูง	1.000	1.000	0.800	0.920	0.900
			ต่ำ	1.000	0.550	0.040	0.200	0
36		ต่ำ	สูง	1.000	1.000	0.975	1.000	1.000
			ต่ำ	1.000	0.900	0.050	0.050	0

* เงื่อนไขนั้นไม่เป็นไปตามข้อตกลงเบื้องต้น

** ข้อมูลไม่เหมาะในการตรวจสอบเนื่องจากไม่มี variation ระหว่างข้อมูล

จากตารางที่ 11 เป็นภาพนำเสนอผลการเปรียบเทียบอำนาจการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) รายละเอียดที่พบมีดังนี้

เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี IRT ทั้ง 3 วิธี พบว่าในกรณีนี้ วิธี IRT ทั้ง 3 วิธี มีอำนาจการตรวจสอบต่างจากตารางก่อนหน้านี้ที่กล่าวมาทั้งหมด กล่าวคือ บางเงื่อนไขวิธี IRT (Z-UN) มีอำนาจการตรวจสอบสูงกว่าวิธี IRT (χ^2) และวิธี IRT (Z-SA) แต่บางเงื่อนไขวิธี IRT (χ^2) มีอำนาจการตรวจสอบสูงที่สุด บางเงื่อนไขวิธี IRT (Z-SA) มีอำนาจการตรวจสอบสูงที่สุด อย่างไรก็ตามพอจะสรุปได้ว่า ในภาพรวม วิธี IRT (Z-UN) ยังคงมีอำนาจการตรวจสอบสูงกว่าวิธี IRT (χ^2) และวิธี IRT (Z-SA)

สำหรับกลุ่มข้อสอบที่มีความลำเอียงสูง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ข้อสอบมีความลำเอียงสูง วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความลำเอียงต่ำ เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (Z-UN) พบว่า วิธี RFA มีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ข้อสอบมีความลำเอียงต่ำ วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือ วิธี MH และวิธี IRT (Z-UN) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากสูง ลักษณะข้อมูลไม่เหมาะสำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT ทั้ง 3 วิธี ดังนั้น พอจะกล่าวได้ว่าเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ข้อสอบมีความยากสูง ไม่เหมาะสำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีทั้ง 5 วิธี

สำหรับกลุ่มข้อสอบที่มีความยากปานกลาง เมื่อเปรียบเทียบอำนาจการตรวจสอบระหว่างวิธี RFA วิธี MH และวิธี IRT (χ^2) ในกรณีที่ข้อสอบมีความลำเอียงสูง พบว่า วิธี RFA และวิธี MH มีอำนาจการตรวจสอบเท่าเทียมกัน ขณะเดียวกันทั้ง วิธี RFA และวิธี MH มีอำนาจการตรวจสอบสูงกว่าวิธี IRT (χ^2) กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่ข้อสอบมีความยากปานกลางและมีความลำเอียงสูง วิธี RFA และวิธี MH จะมีอำนาจการตรวจสอบพอ ๆ กัน ขณะเดียวกันทั้ง วิธี RFA และวิธี MH ยังมีอำนาจการตรวจสอบสูงกว่าวิธี IRT (χ^2)

สรุป เมื่อพิจารณาที่อำนาจจำแนกทั้ง 3 ระดับ เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) วิธี RFA ยังคงมีอำนาจการตรวจสอบสูงที่สุดในทุกเงื่อนไข ในขณะที่วิธี MH และวิธี IRT(Z-UN) จะมีอำนาจการตรวจสอบสูงที่สุด เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกต่ำ

กล่าวโดยสรุป เมื่อแบบสอบยาวขึ้น และกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น อำนาจการตรวจสอบของทั้ง 5 วิธีจะเพิ่มขึ้นด้วย โดยภาพรวม ยังคงสรุปได้ว่า วิธี RFA จะมีอำนาจการตรวจสอบสูงที่สุด รองลงมาคือวิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ ตามลำดับ ดังนั้น ข้อค้นพบนี้จึงไม่สอดคล้องกับสมมุติฐานการวิจัยข้อที่ 1 ที่กล่าวว่า เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ ค่าความยากของข้อสอบและค่าอำนาจจำแนกของข้อสอบต่างกัน ภายใต้ความยาวแบบสอบเดียวกัน วิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ จะมีอำนาจการตรวจสอบและมีอัตราความคลาดเคลื่อนประเภทที่ 1 ไม่ต่างกันในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบอีกเช่นกัน



สถาบันวิจัยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 3 การเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบ
การทำหน้าที่ต่างกันของข้อสอบระหว่างวิธี RFA วิธี MH และวิธี IRT

ตารางที่ 12 อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบด้วยวิธี RFA วิธี MH
และวิธี IRT แบบ 2 พหามิเตอร์ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่าง
ขนาดเล็ก (300 คน)

เงื่อนไข	ความยาก (b)	อำนาจจำแนก (a)	RFA	MH	IRT (Z-SA)	IRT (Z-UN)	IRT (χ^2)
1	สูง	สูง	**	0	*	*	*
2		ปานกลาง	**	0	0.060	0.060	0
3		ต่ำ	**	0	0.23	0.200	0
4	ปาน กลาง	สูง	0	0	0.146	0.160	0
5		ปานกลาง	0	0.033	0.600	0.800	0.060
6		ต่ำ	0	0.016	0.600	0.930	0.060
7	ต่ำ	สูง	0	0	0	0	0
8		ปานกลาง	0	0.033	0	0	0
9		ต่ำ	0	0.033	*	*	*

* เงื่อนไขนั้นไม่เป็นไปตามข้อตกลงเบื้องต้น

** ข้อมูลไม่เหมาะในการตรวจสอบเนื่องจากไม่มี variation ระหว่างข้อมูล

จากตารางที่ 12 เป็นการนำเสนออัตราความคลาดเคลื่อนประเภทที่ 1 ซึ่งหมายถึง
การระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันของการ
ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พหามิเตอร์เมื่อ
เป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) รายละเอียดมีดังต่อไปนี้

เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทำหน้าที่ต่าง
กันของข้อสอบระหว่างวิธี IRT(Z-SA) วิธี IRT (Z-UN) และวิธี IRT(χ^2) พบว่า วิธี IRT (Z-UN)
มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงที่สุดรองลงมาคือ วิธี IRT(Z-SA) และวิธี IRT(χ^2) ตาม
ลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี IRT
(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกัน

สูงที่สุด รองลงมาคือ วิธี IRT(Z-SA) และวิธี IRT(χ^2) ตามลำดับ ในกรณีนี้ผู้วิจัยจึงใช้ วิธี IRT (Z-UN) เป็นตัวแทนในกลุ่มวิธี IRT แบบ 2 พารามิเตอร์

สำหรับกลุ่มข้อสอบที่มีความยากสูง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากสูง วิธี IRT(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากปานกลาง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากปานกลาง วิธี IRT (Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากต่ำ เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี MH มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี IRT(Z-UN) และวิธี RFA กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากต่ำ วิธี MH มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี IRT(Z-UN) และวิธี RFA

สรุป เมื่อพิจารณาจากค่าความยากทั้ง 3 ระดับ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี IRT(Z-UN) จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง ในกรณีที่มีข้อสอบมีความยากปานกลาง และวิธี MH จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง ในกรณีที่มีข้อสอบมีความยากต่ำ ในขณะที่วิธี RFA สามารถควบคุมอัตราตามความคลาดเคลื่อนประเภทที่ 1 ไปได้ดี

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกสูง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกสูง วิธี IRT(Z-UN) ระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกปานกลาง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกปานกลาง วิธี IRT(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกต่ำ เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกต่ำ วิธี IRT(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สรุป เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี IRT (Z-UN) จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง ในกรณีที่มีข้อสอบมีอำนาจจำแนกต่ำ วิธี MH จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง ในกรณีที่มีข้อสอบมีอำนาจจำแนกปานกลาง ในขณะที่วิธี RFA สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดี

กล่าวโดยสรุป วิธี RFA และวิธี MH สามารถควบคุมระดับอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ภายใต้ ระดับ .05 แต่วิธี IRT ทั้ง 3 วิธี ไม่สามารถควบคุมได้ กล่าวได้ว่า วิธี IRT ไม่เหมาะที่จะใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเป็นอย่างยิ่ง ในกรณีแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างมีขนาดเล็ก (300 คน) ดังนั้นสมมุติฐานการวิจัยข้อที่ 2 ที่กล่าวว่า เมื่อกลุ่มตัวอย่างมีขนาดเล็ก ค่าความยากของข้อสอบและค่าอำนาจจำแนกของข้อสอบต่างกัน ภายใต้ความยาวแบบสอบเดียวกันวิธี MH จะมีอำนาจการตรวจสอบสูงกว่าวิธี RFA และวิธี IRT แบบ 2 พารามิเตอร์ และมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าวิธี RFA และวิธี IRT แบบ 2 พารามิเตอร์ ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ จึงถูกปฏิเสธอีกเช่นกัน

ตารางที่ 13 อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทำหน้าที่ต่างกัน
ของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT เมื่อเป็นแบบสอบยาว (75 ข้อ)
และกลุ่มตัวอย่างขนาดเล็ก (300 คน)

เงื่อนไข ที่	ความยาก (b)	อำนาจจำแนก (a)	RFA	MH	IRT (Z-SA)	IRT (Z-UN)	IRT (χ^2)
10	สูง	สูง	**	0	*	*	*
11		ปานกลาง	**	0	*	*	*
12		ต่ำ	**	0	0.016	0.016	0
13	ปาน กลาง	สูง	0	0.004	0.004	0.025	0.025
14		ปานกลาง	0	0	0	0	0
15		ต่ำ	0	0	*	*	*
16	ต่ำ	สูง	0	0	0	0	0
17		ปานกลาง	0	0	0	0	0
18		ต่ำ	0	0.004	0	0	0

* เงื่อนไขนั้นไม่เป็นไปตามข้อตกลงเบื้องต้น

** ข้อมูลไม่เหมาะในการตรวจสอบเนื่องจากไม่มี variation ระหว่างข้อมูล

จากตารางที่ 13 เป็นการนำเสนออัตราความคลาดเคลื่อนประเภทที่ 1 ซึ่งหมายถึงการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) รายละเอียดมีดังต่อไปนี้

เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธี IRT(Z-SA) วิธี IRT (Z-UN) และวิธี IRT(χ^2) พบว่า วิธี IRT (Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงที่สุด ในขณะที่ วิธี IRT(Z-SA) และวิธี IRT(χ^2) มีอัตราความคลาดเคลื่อนประเภทที่ 1 เท่าเทียมกัน กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี IRT (Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงที่สุด ในขณะที่ วิธี IRT(Z-SA) และวิธี IRT(χ^2) มีการ

ระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันพอ ๆ กัน ในกรณีนี้ผู้วิจัยจึงใช้ วิธี IRT (Z-UN) เป็นตัวแทนในกลุ่มวิธี IRT แบบ 2 พารามิเตอร์

สำหรับกลุ่มข้อสอบที่มีความยากสูง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากสูง วิธี IRT(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากปานกลาง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากปานกลาง วิธี IRT(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากต่ำ เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี MH มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี IRT(Z-UN) และวิธี RFA กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีความยากต่ำ วิธี MH มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี IRT(Z-UN) และวิธี RFA ในขณะที่วิธี RFA และวิธี IRT(Z-UN) สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก

สรุป เมื่อพิจารณาจากค่าความยากทั้ง 3 ระดับ เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี RFA มีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำที่สุด ในขณะที่วิธี IRT(Z-UN) จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง ในกรณีที่มีข้อสอบมีความยากปานกลาง วิธี MH จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง ในกรณีที่มีข้อสอบมีความยากปานกลางและมีความยากต่ำ

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกสูง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกสูง วิธี IRT(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกปานกลาง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี RFA วิธี MH และวิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 อยู่ในระดับต่ำที่สุด กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีอำนาจจำแนกปานกลาง วิธี RFA วิธี MH และวิธี IRT(Z-UN) สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกต่ำ เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า วิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) เฉพาะกรณีที่มีอำนาจจำแนกต่ำ วิธี IRT(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่าวิธี MH และวิธี RFA ตามลำดับ

สรุป เมื่อพิจารณาจากค่าอำนาจจำแนกทั้ง 3 ระดับ เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดเล็ก (300 คน) วิธี RFA มีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำที่สุด ในขณะที่วิธี MH จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง ในกรณีที่ข้อสอบมีอำนาจจำแนกสูงและอำนาจจำแนกต่ำ วิธี IRT (Z- UN) จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง ในกรณีที่ข้อสอบมีอำนาจจำแนกสูง

กล่าวโดยสรุป วิธีการตรวจสอบทั้ง 5 วิธี คือวิธี RFA วิธี MH และวิธี IRT ทั้ง 3 วิธี สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดี ในกรณีที่แบบสอบเป็นแบบสอบขนาดยาว (75 ข้อ) แม้ว่ากลุ่มตัวอย่างจะมีขนาดเล็ก (300 คน) และใน 5 วิธีนี้ วิธี RFA มีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำสุด และที่เหลือ 4 วิธี มีอัตราความคลาดเคลื่อนประเภทที่ 1 พอ ๆ กัน ดังนั้น สมมติฐานการวิจัยข้อที่ 2 ที่กล่าวว่า เมื่อกลุ่มตัวอย่างมีขนาดเล็ก ค่าความยากของข้อสอบและค่าอำนาจจำแนกของข้อสอบต่างกัน ภายใต้ความยาวแบบสอบเดียวกันวิธี MH จะมีอำนาจการตรวจสอบสูงกว่าวิธี RFA และวิธี IRT แบบ 2 พารามิเตอร์ และมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าวิธี RFA และวิธี IRT แบบ 2 พารามิเตอร์ ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ จึงถูกปฏิเสธอีกเช่นกัน

ตารางที่ 14 อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน)

เงื่อนไขที่	ความยาก (b)	อำนาจจำแนก (a)	RFA	MH	IRT (Z-SA)	IRT (Z-UN)	IRT (χ^2)
19	สูง	สูง	**	0	*	*	*
20		ปานกลาง	**	0	*	*	*
21		ต่ำ	**	0	*	*	*
22	ปานกลาง	สูง	0	0.167	0.267	0	0
23		ปานกลาง	0	0.250	0.033	0.033	0.033
24		ต่ำ	0	0.250	1.000	1.000	1.000
25	ต่ำ	สูง	0	0	0.067	0.067	0
26		ปานกลาง	0	0.034	0.667	0.689	0
27		ต่ำ	0	0.083	0.133	0.200	0

- * เงื่อนไขนั้นไม่เป็นไปตามข้อตกลงเบื้องต้น
- ** ข้อมูลไม่เหมาะในการตรวจสอบเนื่องจากไม่มี variation ระหว่างข้อมูล

จากตารางที่ 14 เป็นการนำเสนออัตราความคลาดเคลื่อนประเภทที่ 1 ซึ่งหมายถึงการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) รายละเอียดมีดังต่อไปนี้

เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธี IRT(Z-SA) วิธี IRT (Z-UN) และวิธี IRT (χ^2) พบว่า วิธี IRT (Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงที่สุด รองลงมาคือวิธี IRT(Z-SA) และวิธี IRT (χ^2) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) วิธี IRT (Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงที่สุด รองลงมาคือวิธี IRT(Z-SA) และวิธี IRT (χ^2) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากสูง พบว่าข้อมูลไม่เหมาะสำหรับการวิเคราะห์ด้วย
วิธี RFA วิธี MH และวิธี IRT ทั้ง 3 วิธี

สำหรับกลุ่มข้อสอบที่มีความยากปานกลาง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อน
ประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี MH มีอัตราความคลาดเคลื่อน
ประเภทที่ 1 สูงกว่าวิธี IRT(Z-UN) และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น
(25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีความยากปานกลาง
วิธี MH มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกัน
สูงกว่า วิธี IRT(Z-UN) และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากต่ำ เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1
ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อน
ประเภทที่ 1 สูงที่สุด รองลงมาคือวิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น
(25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีความยากต่ำ วิธี IRT
(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกัน
สูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สรุป เมื่อพิจารณาจากความยากทั้ง 3 ระดับ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่าง
ขนาดใหญ่ (1000 คน) วิธี RFA สามารถควบคุมไม่ให้มีอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดี
มาก ในขณะที่วิธี MH จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง ในกรณีที่มีข้อสอบมีความยาก
ปานกลาง และวิธี IRT (Z-UN) จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงในกรณีที่มีข้อสอบมี
ความยากต่ำ

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกสูง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภท
ที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี MH มีอัตราความคลาดเคลื่อนประเภท
ที่ 1 สูงกว่าวิธี IRT(Z-UN) และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และ
กลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกสูง วิธี MH มีการระบุผิด
พลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่าวิธี IRT
(Z-UN) และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกปานกลาง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อน
ประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT (Z-UN)
มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่าวิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็น
แบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจ

จำแนกปานกลาง วิธี IRT(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างสูงกว่าวิธี MH และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีค่าอำนาจจำแนกต่ำ เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-UN) พบว่า วิธี IRT(Z-UN) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงที่สุด รองลงมาคือวิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกต่ำ วิธี IRT(Z-UN) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สรุป เมื่อพิจารณาจากค่าอำนาจจำแนกทั้ง 3 ระดับ วิธี RFA ยังคงสามารถควบคุมไม่ให้มีอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก ในขณะที่วิธี MH และวิธี IRT ทั้ง 3 มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง .ในกรณีที่มีข้อสอบมีอำนาจจำแนกต่ำ

กล่าวโดยสรุป เมื่อเป็นแบบสอบสั้น (25 ข้อ) แม้กลุ่มตัวอย่างจะมีขนาดใหญ่ (1000 คน) วิธี RFA ยังคงควบคุมระดับอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก ในขณะที่ 4 วิธีที่เหลือไม่สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ กล่าวคือ ทั้ง 4 วิธีให้ค่าอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า .05 มาก ทำให้สมมุติฐานการวิจัยข้อที่ 1 ที่กล่าวว่า เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ ค่าความยากของข้อสอบและค่าอำนาจจำแนกของข้อสอบต่างกัน ภายใต้ความยาวแบบสอบเดียวกัน วิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ จะมีอำนาจการตรวจสอบและมีอัตราความคลาดเคลื่อนประเภทที่ 1 ไม่ต่างกัน ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ถูกปฏิเสธ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 15 อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน)

เงื่อนไขที่	ความยาก (b)	อำนาจจำแนก (a)	RFA	MH	IRT (Z-SA)	IRT (Z-UN)	IRT (χ^2)
28	สูง	สูง	**	0	*	*	*
29		ปานกลาง	**	0	*	*	*
30		ต่ำ	**	0	*	*	*
31	ปานกลาง	สูง	0	0	0	0	0
32		ปานกลาง	0	0	0.003	0	0
33		ต่ำ	0	0.020	0.020	0.020	0.016
34	ต่ำ	สูง	0	0	0	0	0
35		ปานกลาง	0	0	0	0	0
36		ต่ำ	0	0	0	0	0

- * เงื่อนไขนั้นไม่เป็นไปตามข้อตกลงเบื้องต้น
- ** ข้อมูลไม่เหมาะในการตรวจสอบเนื่องจากไม่มี variation ระหว่างข้อมูล

จากตารางที่ 15 เป็นการนำเสนออัตราความคลาดเคลื่อนประเภทที่ 1 ซึ่งหมายถึง การระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA วิธี MH และวิธี IRT แบบ 2 พารามิเตอร์ เมื่อเป็นแบบสอบสั้น (25 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) รายละเอียดมีดังต่อไปนี้

เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี IRT(Z-SA) วิธี IRT (Z-UN) และวิธี IRT (χ^2) พบว่า วิธี IRT(Z-SA) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงที่สุด รองลงมาคือวิธี IRT (Z-UN) และวิธี IRT (χ^2) ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) วิธี IRT(Z-SA) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงที่สุด รองลงมาคือวิธี IRT (Z-UN) และวิธี IRT (χ^2) ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากสูง พบว่าข้อสอบไม่เหมาะสำหรับกาวิเคราะห์ด้วย วิธี RFA วิธี MH และวิธี IRT ทั้ง 3 วิธี

สำหรับกลุ่มข้อสอบที่มีความยากปานกลาง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-SA) พบว่า วิธี IRT(Z-SA) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงที่สุด รองลงมาคือวิธี MH และวิธี RFA ตามลำดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีความยากปานกลาง วิธี IRT(Z-SA) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ตามลำดับ

สำหรับกลุ่มข้อสอบที่มีความยากต่ำ เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-SA) พบว่า วิธี RFA วิธี MH และวิธี IRT(Z-SA) มีอัตราความคลาดเคลื่อนประเภทที่ 1 อยู่ในระดับต่ำที่สุดเท่ากัน กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีความยากต่ำ วิธี RFA วิธี MH และวิธี IRT(Z-SA) สามารถควบคุมไม่ให้มีอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก

สรุป เมื่อพิจารณาจากค่าความยากทั้ง 3 ระดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) วิธี RFA ยังคงควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก ในขณะที่วิธี MH และวิธี IRT (Z-SA) จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง เฉพาะกรณีที่มีข้อสอบนั้นมีความยากปานกลาง แต่โดยภาพรวมแล้ว ทั้ง 5 วิธี มีอัตราความคลาดเคลื่อนประเภทที่ 1 อยู่ในระดับต่ำมาก

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกสูง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-SA) พบว่า วิธี RFA วิธี MH และวิธี IRT(Z-SA) มีอัตราความคลาดเคลื่อนประเภทที่ 1 อยู่ในระดับต่ำที่สุดเท่ากัน กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกสูง วิธี RFA วิธี MH และวิธี IRT(Z-SA) สามารถควบคุมไม่ให้มีอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกปานกลาง เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-SA) พบว่า วิธี IRT(Z-SA) มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่าวิธี MH และวิธี RFA กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีที่มีข้อสอบมีอำนาจจำแนกปานกลาง วิธี IRT(Z-SA) มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกันสูงกว่า วิธี MH และวิธี RFA ในขณะที่วิธี MH และวิธี RFA สามารถควบคุมไม่ให้มีอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก

สำหรับกลุ่มข้อสอบที่มีอำนาจจำแนกต่ำ เมื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ระหว่างวิธี RFA วิธี MH และวิธี IRT(Z-SA) พบว่า วิธี MH และวิธี IRT(Z-SA) มีอัตราความคลาดเคลื่อนประเภทที่ 1 อยู่ในระดับต่ำพอ ๆ กัน ในขณะที่ RFA มีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำที่สุดกล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) เฉพาะกรณีข้อสอบที่มีอำนาจจำแนกต่ำ วิธี MH และวิธี IRT(Z-SA) สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ระดับต่ำพอ ๆ กัน ในขณะที่วิธี RFA สามารถควบคุมไม่ให้เกิดอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก

สรุป เมื่อพิจารณาจากค่าอำนาจจำแนกทั้ง 3 ระดับ กล่าวได้ว่า เมื่อเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างขนาดใหญ่ (1000 คน) วิธี RFA ยังคงควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก ในขณะที่วิธี MH และ IRT (Z-SA) จะมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูง เฉพาะกรณีที่ข้อสอบมีอำนาจจำแนกต่ำ แต่โดยภาพรวมทั้ง 5 วิธี มีอัตราความคลาดเคลื่อนประเภทที่ 1 อยู่ในระดับต่ำมาก

กล่าวโดยสรุป ทั้ง 5 วิธี สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ดีมากเมื่อแบบสอบนั้นเป็นแบบสอบยาว (75 ข้อ) และกลุ่มตัวอย่างมีขนาดใหญ่ (1000 คน)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย