

บทที่ 5

สรุป อภิปรายผลและข้อเสนอแนะ

5.1 วัตถุประสงค์ในการวิจัย

เพื่อเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล(MH) กับวิธีซิบเทสต์ (SIBTEST) เมื่อ

1. ขนาดกลุ่มตัวอย่างย่อยระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็นอัตราส่วนต่อกัน ภายใต้ขนาดกลุ่มตัวอย่างต่างกันที่ระดับความยาวแบบสอบเดียวกัน
2. ขนาดกลุ่มตัวอย่างย่อยระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็นอัตราส่วนต่อกัน ภายใต้ขนาดกลุ่มตัวอย่างเดียวกันแต่มีระดับความยาวแบบสอบต่างกัน

5.2 วิธีดำเนินการวิจัย โดยสังเขปมีขั้นตอนดังนี้

5.2.1.การจำลองข้อมูล

ข้อมูลที่ใช้ในการศึกษาได้จากการจำลองขึ้นด้วยโปรแกรม IRTDATA ซึ่งการศึกษาครั้งนี้กำหนดผู้สอบจำนวน 2,000 คน และความยาวแบบสอบ 90 ข้อ โดยมีค่าอำนาจจำแนกเฉลี่ย 1.0 ค่าความยากเฉลี่ย 0.0 และค่าการเดาเป็น 0.2

กลุ่มตัวอย่างที่จะนำไปวิเคราะห์ในแต่ละเงื่อนไข ได้จากการสุ่มอย่างง่าย แบบไม่ใส่กลับคืนจากข้อมูลที่ได้จำลองขึ้น

5.2.2.การสุ่มข้อสอบ เป็นการสุ่มข้อสอบจากแบบสอบเดิมที่มี 90 ข้อ เพื่อให้ได้แบบสอบที่มีความยาวตามที่ต้องการศึกษา โดยสุ่มข้อสอบให้มีจำนวน 30 ข้อ และ 60 ข้อ ซึ่งแบบสอบทั้ง 2 ขนาดที่สุ่มได้นี้มีค่าอำนาจจำแนกเฉลี่ยและค่าความยากเฉลี่ยใกล้เคียงกับแบบสอบเดิม กล่าวคือ แบบสอบขนาด 30 ข้อ มีค่าอำนาจจำแนกเฉลี่ย 0.99 ค่าความยากเฉลี่ย 0.00 และแบบสอบขนาด 60 ข้อ มีค่าอำนาจจำแนกเฉลี่ย 1.00 ค่าความยากเฉลี่ย 0.00

5.2.3.การสุ่มผู้สอบ เป็นการสุ่มจำนวนผู้สอบ หรือขนาดกลุ่มตัวอย่างตามที่ต้องการศึกษา โดยเป็นการนำข้อมูลที่ได้จากการสุ่มรอบแรกมาสุ่มต่อ กล่าวคือ แบบสอบแต่ละขนาดจะมีผู้สอบ 2,000 คน การสุ่มรอบนี้จะเป็นการสุ่มจำนวนผู้สอบกลุ่มอ้างอิงและกลุ่มเปรียบเทียบทั้งหมด 3 ขนาด คือ 200 600 และ 1,000 คน และในแต่ละขนาดจะมีจำนวนผู้สอบกลุ่มอ้างอิงและกลุ่มเปรียบเทียบมีอัตราส่วนต่อกันเป็น 1 : 1 1 : 0.9 1 : 0.75 และ 1 : 0.5

ดังนั้น แบบสอบแต่ละขนาดจะประกอบด้วยขนาดกลุ่มตัวอย่างที่ศึกษาอยู่ 12 เงื่อนไข เมื่อศึกษาความยาวแบบสอบ 3 ขนาด จึงรวมเงื่อนไขที่ต้องศึกษาทั้งหมด 36 เงื่อนไข

5.2.4. การวิเคราะห์ข้อมูลด้วยวิธีทฤษฎีการตอบสนองข้อสอบ

เริ่มที่การวิเคราะห์ด้วยวิธีทฤษฎีการตอบสนองข้อสอบ (IRT) เพื่อตรวจสอบว่าข้อสอบข้อใดทำหน้าที่ย่างกัน เพื่อใช้เป็นเกณฑ์ในการตัดสินความถูกต้องในการตรวจสอบ และเปรียบเทียบประสิทธิภาพในการตรวจสอบของวิธีแมนเทิล-แฮนส์เซล กับวิธีซิปเทสท์ โดยการตรวจสอบด้วยวิธีทฤษฎีการตอบสนองข้อสอบ จะใช้กับเงื่อนไขที่มีขนาดกลุ่มตัวอย่าง 1,000 คน โดยมีอัตราส่วนระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็น 1:1 ของแบบสอบทุกขนาดรวมเป็น 3 เงื่อนไข ส่วนวิธีแมนเทิล-แฮนส์เซลและวิธีซิปเทสท์ใช้วิเคราะห์ข้อมูลในทุกเงื่อนไข

การวิเคราะห์ด้วยวิธีทฤษฎีการตอบสนองข้อสอบ มีขั้นตอนดังนี้

5.2.4.1. วิเคราะห์องค์ประกอบของแบบสอบด้วยโปรแกรม spssx

5.2.4.2. ประมาณค่าพารามิเตอร์ และค่าความแปรปรวนร่วมของข้อมูลด้วยโปรแกรม BILOG 3.04 โดยต้องแยกกันระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ

5.2.4.3. เที่ยบมาตรฐานของค่าพารามิเตอร์ของกลุ่มอ้างอิงและเปรียบเทียบด้วยโปรแกรม EQUATE 2.0

5.2.4.4. คำนวณพื้นที่ระหว่างโค้งคุณลักษณะข้อสอบและทดสอบค่าสถิติด้วยโปรแกรม AREA

5.2.5. การวิเคราะห์ด้วยวิธีแมนเทิล-แฮนส์เซล และวิธีซิปเทสท์

เป็นการคำนวณค่า α_{MH} และทดสอบนัยสำคัญ χ^2 และการคำนวณค่า β_{SIB} และทดสอบนัยสำคัญด้วยสถิติ Z โดยใช้การวิเคราะห์ด้วยโปรแกรมย่อย SIBIN และ SIBTEST

5.2.6. คำนวณอัตราความถูกต้องในการตรวจสอบ อัตราความคลาดเคลื่อนชนิดที่ 1 อัตราความคลาดเคลื่อนชนิดที่ 2

5.2.7. การทดสอบความแตกต่างของอัตราการตรวจสอบ อัตราความคลาดเคลื่อนชนิดที่ 1 และอัตราความคลาดเคลื่อนชนิดที่ 2 ระหว่างวิธี MH กับวิธี SIBTEST ด้วยสถิติ Z

5.2.8. สรุปผลจากการวิเคราะห์เปรียบเทียบระหว่างวิธี MH กับวิธี SIBTEST

5.3.สรุปผลการวิจัย

5.3.1. การทดสอบข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกมิติ พบว่าแบบสอบขนาด 30 ข้อ 60 ข้อ และ 90 ข้อ มีข้อมูลสนับสนุนว่ามีความเป็นเอกมิติ จึงสามารถที่จะใช้วิธีทฤษฎีการตอบสนองข้อสอบในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้

5.3.2. ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีทฤษฎีการตอบสนองข้อสอบพบว่าแบบสอบที่มีระดับความยาว 30 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกัน 2 ข้อ คือ ข้อ 10 และ ข้อ 30 ส่วนแบบสอบที่มีระดับความยาว 60 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกัน 2 ข้อ คือ ข้อ 24 และ ข้อ 59 และแบบสอบที่มีระดับความยาว 90 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกัน 4 ข้อ คือ ข้อ 23 ข้อ 34 ข้อ 35 และข้อ 87

5.3.3. ผลการเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เซล กับวิธีชิบเทสท์ เมื่อใช้ขนาดกลุ่มตัวอย่างระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็นอัตราส่วนต่อกัน ภายใต้อัตราขนาดกลุ่มตัวอย่างต่างกันที่ระดับความยาวแบบสอบเดียวกัน ปรากฏผลดังแสดงในตารางที่ 30

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 30 สรุปผลการเปรียบเทียบอัตราการตรวจสอบ อัตราความคลาดเคลื่อนชนิดที่ 1 และอัตราความคลาดเคลื่อนชนิดที่ 2 ระหว่างวิธี MH กับวิธี SIBTEST

ความยาวแบบสอบ	ขนาดกลุ่มตัว-อย่าง	อัตราส่วน	วิธี MH			วิธี SIBTEST		
			Power	Type 1 error	Type 2 error	Power	Type 1 error	Type 2 error
30	200	1:1	0	0	0	0	0	0
		1:0.9	0	0	0	0	0	0
		1:0.75	0	0	0	0	0	0
		1:0.5	0	+	0	0	-	0
	600	1:1	0	-	0	0	+	0
		1:0.9	0	0	0	0	0	0
		1:0.75	0	-	0	0	+	0
		1:0.5	0	-	0	0	+	0
	1000	1:1	0	0	0	0	0	0
		1:0.9	0	-	0	0	+	0
		1:0.75	0	-	0	0	+	0
		1:0.5	0	-	0	0	+	0
60	200	1:1	0	-	0	0	+	0
		1:0.9	0	0	0	0	0	0
		1:0.75	0	0	0	0	0	0
		1:0.5	-	-	+	+	+	-
	600	1:1	0	0	0	0	0	0
		1:0.9	0	+	0	0	-	0
		1:0.75	0	0	0	0	0	0
		1:0.5	0	-	0	0	+	0

ความยาวแบบทดสอบ	ขนาดกลุ่มตัวอย่าง	อัตราส่วน	วิธี MH			วิธี SIBTEST		
			Power	Type 1 error	Type 2 error	Power	Type 1 error	Type 2 error
60	1000	1:1	0	0	0	0	0	0
		1:0.9	0	+	0	0	-	0
		1:0.75	0	0	0	0	0	0
		1:0.5	0	0	0	0	0	0
90	200	1:1	0	-	0	0	+	0
		1:0.9	0	-	0	0	+	0
		1:0.75	0	-	0	0	+ *	0
		1:0.5	0	-	0	0	+	0
600	600	1:1	0	0	0	0	0	0
		1:0.9	0	-	0	0	+	0
		1:0.75	0	+	0	0	-	0
		1:0.5	+	+	-	-	-	+
1000	1000	1:1	0	0	0	0	0	0
		1:0.9	0	0	0	0	0	0
		1:0.75	0	0	0	0	0	0
		1:0.5	0	+	0	0	-	0

*p<.05

- 0 หมายถึงมีค่าเท่ากับอีกวิธีหนึ่ง
- + หมายถึงมีค่ามากกว่าอีกวิธีหนึ่ง
- หมายถึงมีค่าน้อยกว่าอีกวิธีหนึ่ง

5.3.3.1 ที่ระดับความยาวแบบสอบ 30 ข้อ อัตราการตรวจสอบระหว่างวิธีแมนเทิล-แฮนส์เชล กับวิธีชิบเทสท์ ที่ขนาดกลุ่มตัวอย่างทุกขนาดและทุกอัตราส่วน มีค่าเท่ากับคือ 0.50 ส่วนอัตราความคลาดเคลื่อนชนิดที่ 1 ของแต่ละวิธีที่อัตราส่วนต่างๆ ภายในในกลุ่มตัวอย่างแต่ละขนาดจะมีค่าเท่ากัน หรือใกล้เคียงกัน เมื่อใช้กลุ่มตัวอย่าง 200 คน วิธีแมนเทิล-แฮนส์เชล กับวิธีชิบเทสท์ มีอัตราความคลาดเคลื่อนชนิดที่ 1 เท่ากัน แต่เมื่อขนาดกลุ่มตัวอย่างเป็น 600 คน และ 1000 คน วิธีชิบเทสท์ มีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่า แต่หลังจากทดสอบด้วยสถิติ Z พบว่าไม่มีหลักฐานเพียงพอที่จะสรุปว่าวิธี SIBTEST มีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธี MH และอัตราความคลาดเคลื่อนชนิดที่ 2 ของวิธีแมนเทิล-แฮนส์เชล กับวิธีชิบเทสท์ ที่ทุกขนาดกลุ่มตัวอย่างและทุกอัตราส่วนมีค่าเท่ากับ 0.50 ทั้งสองวิธี

5.3.3.2 ที่ระดับความยาวแบบสอบ 60 ข้อ อัตราการตรวจสอบระหว่างวิธีแมนเทิล-แฮนส์เชล กับวิธีชิบเทสท์ยังคงมีค่าเท่ากันในทุกขนาดกลุ่มตัวอย่าง และทุกอัตราส่วน โดยเฉพาะเมื่อใช้ขนาดกลุ่มตัวอย่าง 1,000 คน ที่อัตราส่วน 1 : 1 และ 1 : 0.9 มีอัตราการตรวจสอบสูงที่สุด คือ สามารถตรวจสอบพบข้อสอบที่ทำหน้าที่ต่างกันได้ถูกต้องร้อยละ 100 ส่วนที่ขนาดกลุ่มตัวอย่างและอัตราส่วนอื่นๆ จะมีอัตราการตรวจสอบเท่ากัน คือ 0.50 นั่นคือ สามารถตรวจสอบได้ถูกต้องร้อยละ 50 ส่วนอัตราความคลาดเคลื่อนชนิดที่ 1 ของแต่ละวิธีที่อัตราส่วนต่างๆ ภายในกลุ่มตัวอย่างแต่ละขนาดจะมีค่าใกล้เคียงกัน ระหว่างวิธีแมนเทิล-แฮนส์เชล กับวิธีชิบเทสท์ จะมีอัตราความคลาดเคลื่อนชนิดที่ 1 ใกล้เคียงกันในทุกเงื่อนไข โดยมักจะพบว่าวิธีชิบเทสท์มีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธีแมนเทิล-แฮนส์เชลเล็กน้อย อัตราความคลาดเคลื่อนชนิดที่ 1 จะสูงที่สุดเมื่อใช้กลุ่มตัวอย่าง 600 คน และจะมีค่าต่ำที่สุดเมื่อใช้กลุ่มตัวอย่าง 1000 คน ที่อัตราส่วน 1 : 1 อัตราความคลาดเคลื่อนชนิดที่ 2 ระหว่างวิธีแมนเทิล-แฮนส์เชลกับวิธีชิบเทสท์มีค่าเท่ากันเกือบทุกเงื่อนไข

5.3.3.3 ที่ระดับความยาวแบบสอบ 90 ข้อ อัตราการตรวจสอบระหว่างวิธีแมนเทิล-แฮนส์เชลกับวิธีชิบเทสท์ จะมีค่าเท่ากันเกือบทุกเงื่อนไข เมื่อใช้ขนาดกลุ่มตัวอย่าง 200 คน ที่อัตราส่วน 1 : 1 และ 1 : 0.9 อัตราการตรวจสอบต่ำที่สุดคือ 0.0 เมื่อใช้ขนาดกลุ่มตัวอย่าง 1000 คน ที่อัตราส่วน 1 : 1 และ 1 : 0.9 อัตราการตรวจสอบสูงกว่าเงื่อนไขอื่นๆ นอกจากที่กล่าวนี้แล้ว ทุกเงื่อนไขมีอัตราการตรวจสอบเท่ากับ 0.25 อัตราความคลาดเคลื่อนชนิดที่ 1 ของ

แต่ละวิธีที่อัตราส่วนต่าง ๆ ภายในแต่ละขนาดกลุ่มตัวอย่างจะมีค่าใกล้เคียงกัน วิธีแมนเทิล-แฮนส์เซลจะมีอัตราความคลาดเคลื่อนชนิดที่ 1 สูงเมื่อใช้ขนาดกลุ่มตัวอย่าง 600 คน ส่วนวิธีชิบเทสต์จะมีอัตราความคลาดเคลื่อนชนิดที่ 1 สูงเมื่อใช้ขนาดกลุ่มตัวอย่าง 200 คน และเมื่อใช้ขนาดกลุ่มตัวอย่าง 1000 คน ที่อัตราส่วน 1 : 1 และ 1 : 0.9 ทั้งสองวิธีมีอัตราความคลาดเคลื่อนชนิดที่ 1 เท่ากับ 0.00 ระหว่างอัตราส่วนจะมีอัตราความคลาดเคลื่อนชนิดที่ 1 ใกล้เคียงกัน เมื่อขนาดกลุ่มตัวอย่างเท่ากับ 600 คน และ 1000 คน แต่ที่ขนาดกลุ่มตัวอย่าง 200 คน วิธีชิบเทสต์จะมีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธีแมนเทิล-แฮนส์เซล ซึ่งเมื่อทดสอบด้วยสถิติ Z พบว่าเกือบทุกเงื่อนไขไม่มีหลักฐานเพียงพอที่จะสรุปว่าวิธี SIBTEST มีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธี MH ยกเว้นที่อัตราส่วน 1 : 0.5 พบว่าวิธี SIBTEST มีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธี MH อย่างมีนัยสำคัญที่ระดับ 0.05

อัตราความคลาดเคลื่อนชนิดที่ 2 ระหว่างวิธีแมนเทิล-แฮนส์เซลกับวิธีชิบเทสต์จะมีค่าเท่ากันเกือบทุกเงื่อนไข โดยที่ขนาดกลุ่มตัวอย่าง 200 คน ที่อัตราส่วน 1 : 1 และ 1 : 0.9 อัตราความคลาดเคลื่อนชนิดที่ 2 สูงที่สุดมีค่าเท่ากับ 1.00 ทั้งสองวิธี ส่วนที่ขนาดกลุ่มตัวอย่าง 1000 คน ที่อัตราส่วน 1 : 1 และ 1 : 0.9 อัตราความคลาดเคลื่อนชนิดที่ 2 ต่ำกว่าเงื่อนไขอื่นๆ คือมีค่าเท่ากับ 0.50 นอกจากนี้ที่กล่าวนี้แล้วเงื่อนไขอื่นๆ มีอัตราความคลาดเคลื่อนชนิดที่ 2 เท่ากับ 0.75

ผลจากการวิเคราะห์ของวิธีแมนเทิล-แฮนส์เซลกับวิธีชิบเทสต์จะพบว่าเมื่อเพิ่มขนาดกลุ่มตัวอย่างจาก 200 คน เป็น 600 คน อัตราการตรวจสอบยังคงเท่าเดิม แต่เมื่อเพิ่มขนาดกลุ่มตัวอย่างเป็น 1000 คน ที่อัตราส่วน 1 : 1 และ 1 : 0.9 อัตราตรวจสอบจะดีที่สุดเมื่อใช้แบบสอบที่ระดับความยาว 60 และ 90 ข้อ สำหรับอัตราความคลาดเคลื่อนชนิดที่ 1 จะมีค่าต่ำที่สุดเมื่อใช้ขนาดกลุ่มตัวอย่าง 1000 คน และจะมีค่าสูงที่สุดเมื่อใช้ขนาดกลุ่มตัวอย่าง 600 คน ส่วนอัตราความคลาดเคลื่อนชนิดที่ 2 แม้จะใช้กลุ่มตัวอย่างขนาดต่างๆ กัน ก็ยังคงมีค่าเท่าเดิมหรือใกล้เคียงกัน โดยเมื่อกลุ่มตัวอย่างขนาด 1000 คน มีอัตราความคลาดเคลื่อนชนิดที่ 2 ต่ำกว่ากลุ่มตัวอย่างขนาดอื่น

เมื่อเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เซลกับวิธีชิบเทสท์ โดยพิจารณาจากอัตราการตรวจสอบ ความคลาดเคลื่อนชนิดที่ 1 และความคลาดเคลื่อนชนิดที่ 2 เมื่อขนาดกลุ่มตัวอย่างต่างกันที่ระดับความยาวแบบสอปเดียวกันพบว่า อัตราการตรวจสอบ อัตราความคลาดเคลื่อนชนิดที่ 1 และอัตราความคลาดเคลื่อนชนิดที่ 2 ของวิธีแมนเทิล-แฮนส์เซลกับวิธีชิบเทสท์ ส่วนมากจะมีค่าเท่ากันหรือใกล้เคียงกันมาก เมื่อทดสอบความแตกต่างของอัตราการตรวจสอบ อัตราความคลาดเคลื่อนชนิดที่ 1 และอัตราความคลาดเคลื่อนชนิดที่ 2 ระหว่างวิธี MH กับวิธี SIBTEST ด้วยสถิติ Z พบว่าเกือบทุกเงื่อนไขไม่สามารถปฏิเสธสมมติฐานศูนย์ นั่นคือวิธี SIBTEST มีอัตราการตรวจสอบ อัตราความคลาดเคลื่อนชนิดที่ 1 และอัตราความคลาดเคลื่อนชนิดที่ 2 เท่ากับวิธี MH มีเพียงเงื่อนไขเดียวคือ เมื่อความยาวแบบสอป 90 ข้อ ขนาดกลุ่มตัวอย่าง 200 คน ที่อัตราส่วนระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็น 1 : 0.5 ที่วิธี SIBTEST มีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธี MH อย่างมีนัยสำคัญที่ระดับ .05 ดังนั้นจึงอาจกล่าวได้ว่าทั้งวิธีแมนเทิล-แฮนส์เซลและวิธีชิบเทสท์มีประสิทธิภาพเท่าเทียมกันในการตรวจสอบการทำหน้าที่ต่างกัน

5.3.4. ผลการเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เซลกับวิธีชิบเทสท์ เมื่อใช้ขนาดกลุ่มตัวอย่างระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็นอัตราส่วนต่อกัน ภายใต้ขนาดกลุ่มตัวอย่างเดียวกัน แต่มีระดับความยาวแบบสอปต่างกัน

5.3.4.1 เมื่อเพิ่มความยาวแบบสอปจาก 30 ข้อ เป็น 60 ข้อ อัตราการตรวจสอบของทั้งวิธีแมนเทิล-แฮนส์เซลและวิธีชิบเทสท์เกือบทุกเงื่อนไขจะมีค่าเท่าเดิม ยกเว้นเงื่อนไขที่ขนาดกลุ่มตัวอย่าง 1000 คน ที่อัตราส่วน 1 : 1 และ 1 : 0.9 อัตราการตรวจสอบจะเพิ่มขึ้น สำหรับอัตราความคลาดเคลื่อนชนิดที่ 1 ของทั้งสองวิธีจะมีค่าลดลง ที่ขนาดกลุ่มตัวอย่าง 200 คน แต่จะมีค่าเพิ่มขึ้นเมื่อขนาดกลุ่มตัวอย่างเป็น 600 คน ส่วนที่ขนาดกลุ่มตัวอย่าง 1000 คน อัตราความคลาดเคลื่อนชนิดที่ 1 ของวิธีแมนเทิล-แฮนส์เซลจะมีค่าเพิ่มขึ้น ส่วนวิธีชิบเทสท์จะมีค่าลดลง ยกเว้นที่อัตราส่วน 1 : 1 ไม่พบความคลาดเคลื่อนชนิดที่ 1 ทั้งสองวิธี ส่วนอัตราความคลาดเคลื่อนชนิดที่ 2 ของทั้งวิธีแมนเทิล-แฮนส์เซลและวิธีชิบเทสท์ส่วนมากยังคงมีค่าเท่าเดิม ยกเว้นที่ขนาดกลุ่มตัวอย่าง 1000 คน ที่อัตราส่วน 1 : 1 และ 1 : 0.9 ทั้งสองวิธีมีอัตราความคลาดเคลื่อนชนิดที่ 2 ลดลงเหลือ 0.00

5.3.4.2 เมื่อเพิ่มความยาวแบบสอบจาก 60 ข้อ เป็น 90 ข้อ อัตราการตรวจสอบของทั้งวิธีแมนเทิล-แฮนส์เชลและวิธีซิปเทสท์จะมีค่าลดลง ส่วนอัตราความคลาดเคลื่อนชนิดที่ 1 จะมีค่าลดลงทั้งสองวิธี ยกเว้นที่ขนาดกลุ่มตัวอย่าง 200 คน วิธีซิปเทสท์จะมีอัตราความคลาดเคลื่อนชนิดที่ 1 เพิ่มขึ้น และมีค่ามากกว่าวิธีแมนเทิล-แฮนส์เชล ส่วนที่ขนาดกลุ่มตัวอย่าง 1000 คน ที่อัตราส่วน 1 : 1 ยังคงไม่พบความคลาดเคลื่อนชนิดที่ 1 ทั้งสองวิธี สำหรับอัตราความคลาดเคลื่อนชนิดที่ 2 จะมีค่าเพิ่มขึ้นทั้งสองวิธี

เมื่อเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เชลกับวิธีซิปเทสท์ โดยพิจารณาจากอัตราการตรวจสอบ ความคลาดเคลื่อนชนิดที่ 1 และความคลาดเคลื่อนชนิดที่ 2 ที่ขนาดกลุ่มตัวอย่างเดียวกันแต่ที่ระดับความยาวแบบสอบต่างกัน พบว่าอัตราการตรวจสอบระหว่างวิธีแมนเทิล-แฮนส์เชลกับวิธีซิปเทสท์มีค่าเท่ากันในทุกระดับความยาวแบบสอบ สำหรับอัตราความคลาดเคลื่อนชนิดที่ 1 ที่ระดับความยาวแบบสอบขนาดต่างๆ ส่วนมากจะมีค่าใกล้เคียงกันระหว่างสองวิธี ยกเว้นที่ความยาวแบบสอบ 90 ข้อ ที่ขนาดกลุ่มตัวอย่าง 200 คน วิธีซิปเทสท์จะมีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธีแมนเทิล-แฮนส์เชล ส่วนอัตราความคลาดเคลื่อนชนิดที่ 2 ระหว่างสองวิธีจะมีค่าเท่ากันที่ทุกระดับความยาวของแบบสอบ เมื่อทดสอบความแตกต่างของอัตราการตรวจสอบ อัตราความคลาดเคลื่อนชนิดที่ 1 และอัตราความคลาดเคลื่อนชนิดที่ 2 ระหว่างวิธี MH กับวิธี SIBTEST ด้วยสถิติ Z พบว่าเกือบทุกเงื่อนไขไม่สามารถปฏิเสธสมมติฐานศูนย์ หมายถึงวิธี SIBTEST มีอัตราการตรวจสอบ อัตราความคลาดเคลื่อนชนิดที่ 1 และอัตราความคลาดเคลื่อนชนิดที่ 2 เท่ากับวิธี MH มีเพียงเงื่อนไขเดียวที่วิธี SIBTEST มีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธี MH อย่างมีนัยสำคัญที่ระดับ .05 ดังนั้นจึงอาจกล่าวได้ว่าวิธีซิปเทสท์กับวิธีแมนเทิล-แฮนส์เชลกับมีประสิทธิภาพเท่าเทียมกันในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

จุฬาลงกรณ์มหาวิทยาลัย

5.4.อภิปรายผล

5.4.1. ผลจากการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีแมนเทิล-แฮนส์เซลกับวิธีชิบเทสท์ พบว่าอัตราการตรวจสอบของทั้งสองวิธีเท่าเทียมกัน ซึ่งสอดคล้องกับผลการศึกษาของ Shealy และ Stout (1993), Narayanan และ Swaminathan (1994) ทั้งนี้เพราะทั้งสองวิธีนี้ต่างก็ใช้ระดับคะแนนรวมจากการตอบข้อสอบเป็นเกณฑ์ในการจับคู่เปรียบเทียบระหว่างผู้สอบต่างกลุ่ม อีกทั้งในการคำนวณทั้งสองวิธีต่างก็จับคู่เปรียบเทียบที่ระดับความสามารถเหมือนกัน และอาจเนื่องมาจากการสุ่มผู้สอบ ที่ได้ผู้สอบกลุ่มอ้างอิงและกลุ่มเปรียบเทียบมีการกระจายของความสามารถใกล้เคียงกัน จึงทำให้ทั้งสองวิธีมีประสิทธิภาพเท่าเทียมกัน ส่วนอัตราความคลาดเคลื่อนชนิดที่ 1 หรือการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกัน ทั้งที่ความจริงข้อสอบทำหน้าที่ไม่ต่างกันของวิธีแมนเทิล-แฮนส์เซลกับวิธีชิบเทสท์ พบว่าโดยส่วนมากวิธีชิบเทสท์จะมีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธีแมนเทิล-แฮนส์เซลเล็กน้อย ซึ่งสอดคล้องกับผลการศึกษาของ Narayanan และ Swaminathan (1994) ซึ่งอาจมีสาเหตุจากคะแนนเฉลี่ยจากแบบสอบที่ศึกษาระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบแตกต่างกัน ($\bar{Y}_{RK} \neq \bar{Y}_{FK}$) จึงทำให้เกิดการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกัน (DIF) ทั้งที่ความจริงข้อสอบทำหน้าที่ไม่ต่างกัน (no DIF) ดังที่ Shealy และ Stout (1993) ได้กล่าวไว้ และการที่วิธี MH มีอัตราความคลาดเคลื่อนชนิดที่ 1 ไม่มาก ก็อาจเนื่องมาจากการกระจายของความสามารถไม่แตกต่างกันมากหรือใกล้เคียงกัน ดังข้อค้นพบของ Roussos และ Stout (1996) กล่าวคือเมื่อการกระจายความสามารถเท่ากัน จะไม่มีปัญหาการเพิ่มสูงของอัตราความคลาดเคลื่อนชนิดที่ 1 ของวิธี MH ความคลาดเคลื่อนชนิดที่ 1

5.4.2. ขนาดกลุ่มตัวอย่างมีผลกระทบต่ออัตราการตรวจสอบและอัตราความคลาดเคลื่อนชนิดที่ 1 กล่าวคือ เมื่อขนาดกลุ่มตัวอย่าง 200 คน และ 600 คน สามารถตรวจสอบพบข้อสอบที่ทำหน้าที่ต่างกันได้ถูกต้องร้อยละ 50 ซึ่งสอดคล้องกับผลการศึกษาของ Mazor (1992) ที่พบว่าเมื่อใช้ขนาดกลุ่มตัวอย่างน้อยกว่า 500 คน จะสามารถระบุข้อสอบที่ทำหน้าที่ต่างกันได้ถูกต้องร้อยละ 50 แต่เมื่อเพิ่มขนาดกลุ่มตัวอย่างเป็น 1000 คน ทำให้สามารถตรวจสอบพบข้อสอบที่ทำหน้าที่ต่างกันได้ถูกต้องร้อยละ 100 ส่วนอัตราความคลาดเคลื่อนชนิดที่ 1 เมื่อใช้ขนาดกลุ่มตัวอย่าง 200 คน จะมีอัตราความคลาดเคลื่อนชนิดที่ 1 อยู่ในระดับหนึ่ง และเมื่อขนาดกลุ่มตัวอย่างเพิ่มเป็น 600 คน ทำให้อัตราความคลาดเคลื่อนชนิดที่ 1 เพิ่มขึ้นทั้งสองวิธี แต่เมื่อเพิ่ม

ขนาดกลุ่มตัวอย่างเป็น 1,000 คน อัตราความคลาดเคลื่อนชนิดที่ 1 ลดลง กล่าวคือ มีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันเป็นร้อยละ 0 ทั้งนี้เป็นเพราะทั้งวิธีแมนเทิล-แฮนส์เซลและวิธีชิบเทสต์ต่างก็มีพื้นฐานการกระจายแบบ asymptotic ซึ่งต้องใช้จำนวนผู้สอบขนาดใหญ่ (Roussos and Stout, 1996) ดังนั้นเมื่อใช้กลุ่มตัวอย่างน้อย จึงทำให้เกิดอัตราความคลาดเคลื่อนชนิดที่ 1 สูง อัตราการตรวจสอบต่ำกว่า เมื่อใช้กลุ่มตัวอย่างขนาดใหญ่ ซึ่งมีอัตราความคลาดเคลื่อนชนิดที่ 1 ต่ำและอัตราการตรวจสอบสูง

อย่างไรก็ตามในทางปฏิบัติทั้งวิธีแมนเทิล-แฮนส์เซลและวิธีชิบเทสต์ต่างก็ไม่เคร่งครัดกับขนาดกลุ่มตัวอย่างมากนัก แม้จะใช้กลุ่มตัวอย่างน้อยก็สามารถที่จะดำเนินการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ดังผลจากการวิเคราะห์พบว่าเมื่อกลุ่มอ้างอิงมีขนาด 200 คน และกลุ่มเปรียบเทียบมีขนาด 100 คน ก็สามารถตรวจสอบพบข้อสอบที่ทำหน้าที่ต่างกันได้ถูกต้องร้อยละ 50 ซึ่งสอดคล้องกับคำกล่าวของ Hill (1990 Cited in Mazor et al. 1992) ว่าสามารถใช้ขนาดกลุ่มตัวอย่างขนาด 100 คน และ 300 คน สำหรับผู้สอบกลุ่มใดกลุ่มหนึ่งหรือทั้งสองกลุ่ม เป็นการประหยัดทั้งเวลาและค่าใช้จ่าย กล่าวคือ ใช้ขนาดกลุ่มตัวอย่างน้อย จำนวนรอบในการวิเคราะห์น้อย จึงทำให้นักศึกษานิยมใช้โดยเฉพาะวิธีแมนเทิล-แฮนส์เซล เป็นวิธีที่ใช้อย่างแพร่หลาย

5.4.3. ระดับความยาวของแบบสอบ มีผลกระทบต่ออัตราการตรวจสอบ อัตราความคลาดเคลื่อนชนิดที่ 1 และอัตราความคลาดเคลื่อนชนิดที่ 2 ของวิธีแมนเทิล-แฮนส์เซลและวิธีชิบเทสต์ กล่าวคือ เมื่อเพิ่มความยาวแบบสอบจาก 30 ข้อ เป็น 60 ข้อ ที่ขนาดกลุ่มตัวอย่าง 200 คน และ 600 คน อัตราการตรวจสอบยังคงเท่าเดิม ซึ่งสอดคล้องกับผลการศึกษาของ Roger และ Swaminathan (1993) ที่ศึกษาจากแบบสอบขนาด 40 ข้อ และ 80 ข้อ ใช้ขนาดกลุ่มตัวอย่าง 500 คน แต่ที่ขนาดกลุ่มตัวอย่าง 1,000 คน อัตราการตรวจสอบเพิ่มขึ้น หรือสามารถระบุข้อสอบที่ทำหน้าที่ต่างกันได้ถูกต้องร้อยละ 100 ซึ่งสอดคล้องกับคำกล่าวของ Swaminathan และ Roger (1990) ว่าแบบสอบที่ยาวและกลุ่มตัวอย่างขนาดใหญ่ ทำให้มีโอกาสที่ดีในการตรวจสอบพบข้อสอบที่ทำหน้าที่ต่างกัน และเมื่อความยาวแบบสอบเป็น 90 ข้อ พบว่าอัตราการตรวจสอบลดลง ซึ่งขัดแย้งกับคำกล่าวข้างต้น ทั้งนี้ อาจเนื่องมาจากแบบสอบขนาด 90 ข้อ เมื่อเปรียบเทียบค่าความยากเฉลี่ยของข้อสอบระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบพบว่ามีค่าเท่ากัน ซึ่งแสดงว่าทั้งสองกลุ่มมีความสามารถโดยเฉลี่ยใกล้เคียงกัน จึงทำให้ตรวจสอบพบข้อสอบ

ที่ทำหน้าที่ต่างกันได้ถูกต้องลดลง อีกทั้งข้อสอบที่ทำหน้าที่ต่างกันนั้น มีค่าอำนาจจำแนกค่อนข้างต่ำ ซึ่ง Roger และ Swaminathan (1993) Narayanan และ Swaminathan (1994) พบว่าเมื่อข้อสอบมีค่าอำนาจจำแนกสูง จะทำให้อัตราการตรวจสอบสูงที่สุด คือ สามารถตรวจสอบพบข้อสอบที่ทำหน้าที่กันได้อย่างถูกต้องมากที่สุด

เมื่อพิจารณาระดับความยาวของแบบสอบที่มีผลกระทบต่ออัตราความคลาดเคลื่อนชนิดที่ 1 หรือการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกัน ทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ไม่ต่างกัน ในบางเงื่อนไขเมื่อแบบสอบมีความยาวมากขึ้น อัตราความคลาดเคลื่อนชนิดที่ 1 จะลดลงอันเป็นผลตามที่คาดหวังไว้ ซึ่งสอดคล้องกับผลการศึกษาของ Uttaro และ Millsap (1994) แต่ในบางเงื่อนไขเมื่อแบบสอบมีความยาวมากขึ้น คือที่ระดับความยาวแบบสอบ 60 ข้อ อัตราความคลาดเคลื่อนชนิดที่ 1 เพิ่มขึ้น ซึ่งสอดคล้องกับผลการศึกษาของ Swaminathan และ Roger (1990) ที่เป็นเช่นนี้อาจเนื่องมาจากการกระจายของความสามารถระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบแตกต่างกัน หรือข้อสอบเหล่านั้นเป็นข้อที่มีค่าอำนาจจำแนกสูงแต่มีค่าความยากต่ำ จึงทำให้ทั้งวิธี SIBTEST และวิธี MH ระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันเพิ่มขึ้น

เมื่อพิจารณาระดับความยาวของแบบสอบที่มีผลกระทบต่ออัตราความคลาดเคลื่อนชนิดที่ 2 หรือการระบุผิดพลาดว่าข้อสอบทำหน้าที่ไม่ต่างกัน ทั้งที่ความเป็นจริงข้อสอบทำหน้าที่ต่างกัน ที่ความยาวแบบสอบ 30 ข้อ และ 60 ข้อ อัตราความคลาดเคลื่อนชนิดที่ 2 เท่ากัน แต่เมื่อแบบสอบมีความยาว 90 ข้อ อัตราความคลาดเคลื่อนชนิดที่ 2 เพิ่มขึ้น ซึ่งสัมพันธ์กับอัตราการตรวจสอบที่ลดลงเมื่อใช้แบบสอบขนาด 90 ข้อ ที่เป็นเช่นนี้อาจเนื่องมาจากข้อสอบที่ทำหน้าที่ต่างกันนั้นมีค่าดัชนีของการทำหน้าที่ต่างกันของข้อสอบต่ำ (มีพื้นที่ที่เกิดจากความแตกต่างระหว่างโค้งคุณลักษณะของผู้สอบสองกลุ่มย่อยมีค่าน้อย) และส่วนมากเป็นข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูป (Nonuniform DIF) จึงทำให้เกิดการระบุผิดพลาดว่าข้อสอบทำหน้าที่ไม่ต่างกันทั้งที่ความจริงข้อสอบทำหน้าที่ต่างกัน

5.4.4 จากผลการวิเคราะห์ในทฤษฎีพบว่าเป็นไปได้โดยส่วนมากทั้งวิธีแมนเทิล-แฮนส์เชล และวิธีซิปเทสท์ที่มีอัตราการตรวจสอบได้ถูกต้องร้อยละ 25 ถึงร้อยละ 50 ซึ่งถือว่าค่อนข้างต่ำ ทั้งนี้ อาจเนื่องมาจากจำนวนข้อสอบที่ทำหน้าที่ต่างกัน (DIF) ในแบบสอบแต่ละขนาดมีจำนวนน้อย ซึ่ง Narayanan และ Swaminathan (1994) ได้ข้อค้นพบว่า สัดส่วนของข้อสอบที่ DIF ต่อข้อสอบที่ no DIF ทั้งหมดภายในแบบสอบมีผลกระทบต่ออัตราการตรวจสอบ กล่าวคือเมื่อสัดส่วนของข้อสอบที่ DIF ภายในแบบสอบเพิ่มขึ้น ทำให้อัตราการตรวจสอบได้ถูกต้องมากขึ้น

5.5. ข้อเสนอแนะ

5.5.1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

5.5.1.1. จากผลการวิจัยทำให้ได้ข้อค้นพบว่าวิธีแมนเทิล-แฮนส์เชลกับวิธีซิปเทสท์ มีอัตราการตรวจสอบเท่าเทียมกัน มีอัตราความคลาดเคลื่อนชนิดที่ 1 ใกล้เคียงกันและโดยส่วนมากวิธีซิปเทสท์ จะมีอัตราความคลาดเคลื่อนชนิดที่ 1 มากกว่าวิธีแมนเทิล-แฮนส์เชลเล็กน้อย และมีอัตราความคลาดเคลื่อนชนิดที่ 2 เท่ากัน อาจกล่าวได้ว่าทั้งสองวิธีนี้มีประสิทธิภาพเท่าเทียมกันในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ แต่วิธีที่ควรจะไปใช้มากกว่าคือวิธีแมนเทิล-แฮนส์เชลเพราะมีอัตราความคลาดเคลื่อนชนิดที่ 1 น้อยกว่าวิธีซิปเทสท์

5.5.1.2. ขนาดกลุ่มตัวอย่างที่เป็นเงื่อนไขที่ดีที่สุดคือ กลุ่มอ้างอิงและกลุ่มเปรียบเทียบมีขนาด 1,000 คน ทั้งนี้ เพราะสามารถตรวจสอบพบว่าข้อสอบทำหน้าที่ต่างกันได้ถูกต้องร้อยละ 50 ถึง 100 และไม่พบว่ามีผลกระทบคลาดว่าข้อสอบทำหน้าที่ต่างกัน และเงื่อนไขที่ตรงลงมาคือ การใช้กลุ่มตัวอย่างขนาด 200 คน โดยอัตราส่วนระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบทุกอัตราส่วน ทำให้อัตราการตรวจสอบและอัตราความคลาดเคลื่อนเท่ากันหรือใกล้เคียงกัน โดยสามารถตรวจสอบได้ถูกต้องร้อยละ 50 และมีผลกระทบคลาดว่าข้อสอบทำหน้าที่ต่างกันไม่เกินร้อยละ 3.6 ในทางปฏิบัติเงื่อนไขที่เหมาะสมและประหยัดคือ ขนาดกลุ่มตัวอย่าง 200 คน มีอัตราส่วนระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเท่ากับ 1 : 0.75

5.5.1.3. ความยาวของแบบสอบ เมื่อพิจารณาถึงอัตราการตรวจสอบและอัตราความคลาดเคลื่อน แบบสอบขนาด 30 ข้อ และ 60 ข้อ ให้ค่าต่างๆ ใกล้เคียงกัน จึงสามารถที่จะใช้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ดีทั้งแบบสอบขนาด 30 ข้อ และแบบสอบขนาด 60 ข้อ ทั้งนี้ ขึ้นกับเนื้อหาและวัตถุประสงค์ในการสร้างแบบสอบ เงื่อนไขที่เหมาะสมและประหยัดคือ การใช้แบบสอบขนาด 60 ข้อ และกลุ่มตัวอย่างขนาด 200 คน โดยมีอัตราส่วนระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเท่ากับ 1 : 0.75 ทั้งนี้เพราะสามารถตรวจสอบได้ถูกต้องร้อยละ 50 และมีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ไม่ต่างกันร้อยละ 1.7

5.5.2. ข้อเสนอแนะในการศึกษาเพิ่มเติม

5.5.2.1. ในการศึกษาครั้งนี้กลุ่มอ้างอิงและกลุ่มเปรียบเทียบได้มาโดยการสุ่ม จึงควรจะได้ทำการศึกษาว่าหากจำแนกกลุ่มอ้างอิงและกลุ่มเปรียบเทียบตามเพศ สถานที่ศึกษาหรืออื่นๆ โดยยังคงใช้ขนาดกลุ่มอ้างอิงและกลุ่มเปรียบเทียบเป็นอัตราส่วนต่อกัน เพื่อเป็นการตรวจสอบและยืนยันข้อความรู้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

5.5.2.2. เนื่องจากค่าดัชนีของการทำหน้าที่ต่างกันของข้อสอบมีผลต่อการตรวจสอบทางข้อสอบที่ทำหน้าที่ต่างกัน กล่าวคือ เมื่อดัชนีมีค่าต่ำ มักจะตรวจสอบพบได้ยากหรือมีการระบุผิดพลาด จึงควรจะศึกษาเกี่ยวกับขนาดกลุ่มตัวอย่างและความยาวแบบสอบที่มีผลต่อการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันที่มีค่าดัชนีของการทำหน้าที่ต่างกันในระดับต่างๆ เพื่อประโยชน์ในการเลือกใช้เงื่อนไขที่ให้ผลดีที่สุด เพราะในทางปฏิบัติเราย่อมไม่ทราบมาก่อนว่าข้อสอบข้อใดมีค่าดัชนีของการทำหน้าที่ต่างกันระดับใด

5.5.2.3. ความแตกต่างของค่าความยากเฉลี่ยระหว่างกลุ่มมีผลต่อการตรวจสอบการทำหน้าที่ต่างกัน กล่าวคือ หากค่าความยากเฉลี่ยระหว่างกลุ่มไม่มีความแตกต่างกัน หรือแตกต่างกันน้อย ทำให้ความถูกต้องในการตรวจสอบน้อยและมีการระบุผิดพลาด ดังนั้น จึงควรศึกษาว่าควรใช้กลุ่มตัวอย่างขนาดเท่าใด ที่แต่ละระดับของความแตกต่างของค่าความยากเฉลี่ยระหว่างกลุ่ม เพื่อประโยชน์ในการเลือกนำไปใช้

5.5.2.4 จากผลการวิจัยพบว่าอัตราการตรวจสอบได้ถูกต้องโดยส่วนมากคิดเป็นร้อยละ 25 ถึงร้อยละ 50 ซึ่งถือว่าค่อนข้างต่ำ ทั้งนี้อาจเป็นเพราะจำนวนข้อสอบที่ DIF ในแบบสอบแต่ละขนาดมีจำนวนน้อย ดังนั้นจึงน่าจะศึกษาเกี่ยวกับสัดส่วนของข้อสอบที่ DIF ภายในแบบสอบที่มีผลกระทบต่อประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

5.5.2.5 น่าจะศึกษาในกรณีที่อัตราส่วนระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบต่างกัน มาก ๆ ภายใต้กลุ่มตัวอย่างขนาดใหญ่ เช่น ขนาดกลุ่มตัวอย่าง 1000 คน ที่อัตราส่วน 1 : 0.25 และ 1 : 0.10 เพื่อดูประสิทธิภาพในการตรวจสอบ เพราะในทางปฏิบัติจะมีกรณีเช่นนี้ เช่น การสอบคัดเลือกเข้าศึกษาต่อในคณะครุศาสตร์ คณะวิศวกรรมศาสตร์ เป็นต้น



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย