

## บทที่ 5

### สรุป อภิปรายผลและข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอเนกประหว่างวิธีแมนเทิล-แฮนส์เซล (MH) กับวิธีถดถอยโลจิสติก (LR) โดยมีวัตถุประสงค์เฉพาะดังนี้

1. เพื่อศึกษาข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกประหว่าง เมื่อมีการจัดกลุ่มความสามารถ ผู้สอบออกเป็นกลุ่มผู้สอบที่มีความสามารถสูง กลุ่มผู้สอบที่มีความสามารถปานกลาง และกลุ่มผู้สอบที่มีความสามารถต่ำ โดยใช้วิธีทฤษฎีการตอบสนองข้อสอบ (IRT) ตรวจสอบ

2. เพื่อเปรียบเทียบอัตราความถูกต้องและอัตราความคลาดเคลื่อนของผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอเนกประหว่างวิธีแมนเทิล-แฮนส์เซล กับวิธีถดถอยโลจิสติก

ข้อมูลที่ใช้ในการศึกษาได้จากการจำลองขึ้นด้วยโปรแกรม IRTDATA โดยกำหนดจำนวนผู้สอบ 2,000 คน ความยาวข้อสอบ 40 ข้อ และค่าความสามารถผู้สอบแบ่งออกเป็น 3 ระดับ คือ กลุ่มผู้สอบที่มีความสามารถสูง กลุ่มผู้สอบที่มีความสามารถปานกลาง และกลุ่มผู้สอบที่มีความสามารถต่ำ ค่าความยากของข้อสอบแบ่งออกเป็น 3 ระดับ คือ กลุ่มข้อสอบที่มีค่าความยากสูง กลุ่มข้อสอบที่มีค่าความยากปานกลาง และกลุ่มข้อสอบที่มีค่าความยากต่ำ ค่าอำนาจจำแนกของข้อสอบแบ่งออกเป็น 3 ระดับ คือ กลุ่มข้อสอบที่ค่าอำนาจจำแนกสูง กลุ่มข้อสอบที่ค่าอำนาจจำแนกปานกลาง และกลุ่มข้อสอบที่ค่าอำนาจจำแนกต่ำ ส่วนค่าการเดากำหนดให้เท่ากัน คือ 0.2 รวมเงื่อนไขที่ศึกษาทั้งหมด 27 เงื่อนไข ซึ่งแต่ละเงื่อนไขเป็นการจำลองข้อสอบต่างชุดกันตามระดับความสามารถผู้สอบ และภายในเงื่อนไขเดียวกันที่จำลองขึ้นนั้นใช้วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ 3 วิธีคือ วิธี IRT วิธี MH และวิธี LR

กลุ่มตัวอย่างที่ใช้วิเคราะห์ในแต่ละเงื่อนไขได้จากการสุ่มจำนวนผู้สอบด้วยวิธีสุ่มอย่างง่ายจากข้อมูลที่ได้จำลองขึ้น โดยสุ่มจำนวนผู้สอบกลุ่มข้างอิงและกลุ่มเปรียบเทียบมีสัดส่วนเท่ากัน คือ กลุ่มละ 1,000 คน ในทุกเงื่อนไขทั้งหมด 27 เงื่อนไข

ในการวิจัยครั้งนี้ทำการวิเคราะห์ข้อมูล ดังนี้

1. ทำการวิเคราะห์ข้อมูลเบื้องต้น เพื่อหาค่าสถิติพื้นฐานของข้อสอบทุกเงื่อนไขข้อสอบ และวิเคราะห์องค์ประกอบของข้อสอบเพื่อตรวจสอบความเป็นเอกมิติของข้อสอบโดยใช้โปรแกรมสำเร็จรูป SPSS/PC+

2. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูป โดยใช้วิธีทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นวิธีมาตรฐานที่ระบุว่าข้อสอบข้อใดทำหน้าที่ต่างกัน เพื่อให้เป็นเกณฑ์ในการประเมินอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีแมนเทิล-แฮนส์เชล กับวิธีถดถอยโลจิสติก โดยทำการวิเคราะห์ทุกเงื่อนไขที่ศึกษารวมทั้งหมด 27 เงื่อนไข ขั้นตอนในการวิเคราะห์มีดังนี้

2.1) ประมาณค่าพารามิเตอร์ และค่าความแปรปรวนร่วมของข้อมูลโดยใช้โปรแกรม BILOG 3.04 ซึ่งข้อมูลแยกกันในแต่ละแพ้มระหว่างกลุ่มอ้างอิง กับกลุ่มเปรียบเทียบ

2.2) เทียบมาตรฐานค่าพารามิเตอร์ของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ โดยใช้โปรแกรม EQUATE 2.0

2.3) คำนวณพื้นที่ระหว่างโค้งคุณลักษณะข้อสอบ และทดสอบค่าสถิติ โดยใช้โปรแกรม AREA

3. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูปด้วยวิธีแมนเทิล-แฮนส์เชล โดยใช้โปรแกรมสำเร็จรูป MH-DIF เพื่อคำนวณค่าดัชนี  $\alpha_{MH}$  ดัชนี  $MH_{DIF}$  และการทดสอบนัยสำคัญด้วยไคสแควร์

4. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูปด้วยวิธีถดถอยโลจิสติก โดยใช้โปรแกรม SPSS/PC+ เพื่อคำนวณค่าปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบกับความสามารถของผู้สอบในการทำนายโอกาสการตอบข้อสอบถูกหรือผิด และการทดสอบนัยสำคัญด้วยค่าไคสแควร์

5. คำนวณอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีแมนเทิล-แฮนส์เชล กับวิธีถดถอยโลจิสติก และทดสอบความแตกต่างของอัตราความถูกต้องและอัตราความคลาดเคลื่อนของทั้งสองวิธีด้วยสถิติทดสอบ Z

### สรุปผลการวิจัย

1. ผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูปด้วยวิธี IRT

1.1 สำหรับกลุ่มผู้สอบที่ความสามารถสูง ข้อสอบที่ตรวจพบการทำหน้าที่ต่างกันแบบอเนกรูปมากที่สุด คือข้อสอบที่มีค่าความยากสูง ค่าอำนาจจำแนกสูง รองลงมาคือ ข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกสูง และข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง ตามลำดับ

1.2 สำหรับกลุ่มผู้สอบที่มีความสามารถปานกลาง ข้อสอบที่ตรวจพบการทำหน้าที่ต่างกันแบบอนกรุปมากที่สุด คือข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกสูง รองลงมาคือข้อสอบที่มีค่าความยากต่ำ ค่าอำนาจจำแนกปานกลาง และข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง ตามลำดับ

1.3 สำหรับกลุ่มผู้สอบที่มีความสามารถต่ำ ข้อสอบที่ตรวจพบการทำหน้าที่ต่างกันแบบอนกรุปมากที่สุด คือข้อสอบที่มีค่าความยากต่ำ ค่าอำนาจจำแนกสูง รองลงมาคือ ข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง ตามลำดับ

2. ผลการเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เชล กับวิธีดอดอยโลจิสติก เมื่อจัดกลุ่มความสามารถของผู้สอบ ค่าความยากของข้อสอบ และค่าอำนาจจำแนกของข้อสอบต่างกัน

2.1 สำหรับกลุ่มผู้สอบที่มีความสามารถสูง

2.1.1 โดยภาพรวมวิธี MH กับวิธี LR มีประสิทธิภาพการตรวจสอบเท่าเทียมกัน

2.1.2 วิธี LR มีอัตราความถูกต้องประเภทที่ 1 มากกว่าวิธี MH อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง

2.1.3 วิธี LR มีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าวิธี MH อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง

2.2 สำหรับกลุ่มผู้สอบที่มีความสามารถปานกลาง

2.2.1 โดยภาพรวมวิธี MH กับวิธี LR มีประสิทธิภาพการตรวจสอบเท่าเทียมกัน

2.2.2 วิธี LR มีอัตราความถูกต้องประเภทที่ 2 มากกว่าวิธี MH อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง

2.2.3 วิธี LR มีอัตราความคลาดเคลื่อนประเภทที่ 2 ต่ำกว่าวิธี MH อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง

2.3 สำหรับกลุ่มผู้สอบที่มีความสามารถต่ำ

2.3.1 โดยภาพรวมวิธี MH กับวิธี LR มีประสิทธิภาพการตรวจสอบเท่าเทียมกัน

2.3.2 วิธี LR มีอัตราความถูกต้องประเภทที่ 2 มากกว่าวิธี MH อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง

2.3.3 วิธี LR มีอัตราความคลาดเคลื่อนประเภทที่ 2 ต่ำกว่าวิธี MH อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง



## อภิปรายผล

1. ผลจากการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูป ในกลุ่มผู้สอบที่มีความสามารถสูง พบว่าข้อสอบที่ตรวจพบการทำหน้าที่ต่างกันแบบอเนกรูปมากที่สุดเป็นข้อสอบที่มีค่าความยากสูงและค่าความยากปานกลาง และมีค่าอำนาจจำแนกสูง ซึ่งข้อค้นพบดังกล่าวไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ แต่สอดคล้องกับผลการศึกษาของ Rogers และ Swaminathan (1993) ที่ค้นพบว่าข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกสูงจะทำให้มีอัตราการทำข้อสอบสูงสุด คือตรวจพบข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูปได้มากที่สุด และสอดคล้องกับผลการศึกษาของ Sudweeks และ Tolman (1990) ที่พบว่าข้อสอบที่ทำหน้าที่ต่างกันส่วนใหญ่เป็นข้อสอบยาก

นอกจากนี้กลุ่มผู้สอบที่มีความสามารถสูง เมื่อทำข้อสอบที่มีค่าความยากสูง และข้อสอบที่มีค่าความยากปานกลาง ซึ่งเป็นค่าความยากของข้อสอบที่เหมาะสมกับระดับความสามารถผู้สอบ และข้อสอบที่ทำมีค่าอำนาจจำแนกสูง ทำให้ผู้สอบมีความไวในการตอบข้อสอบ ซึ่งอาจมีผลทำให้เกิดการได้เปรียบและเสียเปรียบในการตอบข้อสอบได้ถูกต้องระหว่างกลุ่มผู้สอบ ดังนั้นในการสร้างข้อสอบจึงควรให้ความระมัดระวังปัจจัยที่อาจทำให้เกิดการทำหน้าที่ต่างกัน เช่น ลักษณะผู้สอบกลุ่มย่อย ได้แก่ เพศ ภูมิภาค วัฒนธรรม เป็นต้น และในการสร้างแบบสอบควรพิจารณาเกี่ยวกับ ลักษณะของภาษา คำที่ใช้ เนื้อหาของหลักสูตร เพื่อจะได้นำมาปรับปรุงหรือหลีกเลี่ยงการเกิด DIF ให้น้อยที่สุด ทั้งนี้เพื่อให้ได้แบบสอบที่มีความยุติธรรมต่อกลุ่มผู้สอบมากที่สุด

2. ผลจากการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูป ในกลุ่มผู้สอบที่มีความสามารถปานกลาง พบว่าข้อสอบที่ตรวจพบการทำหน้าที่ต่างกันแบบอเนกรูปมากที่สุด เป็นข้อสอบที่มีค่าความยากปานกลางและมีค่าอำนาจจำแนกสูง ซึ่งข้อค้นพบดังกล่าวไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ แต่สอดคล้องกับผลการศึกษาของ Rogers และ Swaminathan (1993) ที่ค้นพบว่าข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกสูงจะทำให้มีอัตราการทำข้อสอบสูงสุด คือตรวจพบข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูปได้มากที่สุด และสอดคล้องกับผลการศึกษาของ Mazor (1992) ที่พบว่าข้อสอบที่ทำหน้าที่ต่างกัน ส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจจำแนกสูง

นอกจากนี้กลุ่มผู้สอบที่มีความสามารถปานกลาง เมื่อทำข้อสอบที่มีค่าความยากปานกลาง ซึ่งเป็นค่าความยากที่เหมาะสมกับระดับความสามารถผู้สอบ และข้อสอบที่ทำมีค่าอำนาจจำแนกสูง ทำให้ผู้สอบมีความไวในการตอบข้อสอบ ซึ่งอาจมีผลทำให้เกิดการได้เปรียบและเสียเปรียบในการตอบข้อสอบได้ถูกต้องระหว่างกลุ่มผู้สอบ ดังนั้นในการสร้างข้อสอบจึงควรให้

ความระมัดระวังปัจจัยที่อาจทำให้เกิดการทำหน้าที่ต่างกัน เช่น ลักษณะผู้สอบกลุ่มย่อย ได้แก่ เพศ ภูมิภาค วัฒนธรรม เป็นต้น และในการสร้างแบบสอบควรให้ความระมัดระวังและพิจารณาเกี่ยวกับ ลักษณะของภาษา คำที่ใช้ เนื้อหาของหลักสูตร เพื่อจะได้นำมาปรับปรุงหรือหลีกเลี่ยงการเกิด DIF ให้น้อยที่สุด ทั้งนี้เพื่อให้ได้แบบสอบที่มีความยุติธรรมต่อกลุ่มผู้สอบมากที่สุด

3. ผลจากการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอนุกรม ในกลุ่มผู้สอบที่มีความสามารถต่ำ พบว่าข้อสอบที่ทำหน้าที่ต่างกันแบบอนุกรมส่วนใหญ่เป็นข้อสอบที่มีค่าความยากต่ำและค่าความยากปานกลางและส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจจำแนกสูง ซึ่งข้อค้นพบดังกล่าวสอดคล้องกับสมมติฐานที่ตั้งไว้ และสอดคล้องกับผลการศึกษาของ Narayanan และ Swaminathan (1996) ที่ค้นพบว่าข้อสอบที่มีค่าความยากต่ำ ค่าอำนาจจำแนกสูง ตรวจพบการทำหน้าที่ต่างกันของข้อสอบแบบอนุกรมมากที่สุด และสอดคล้องกับผลการศึกษาของ Mazor (1994) ที่ค้นพบว่าข้อสอบที่ทำหน้าที่ต่างกันแบบอนุกรม ส่วนใหญ่เป็นข้อสอบที่ง่ายหรือมีค่าความยากต่ำ

นอกจากนี้กลุ่มผู้สอบที่มีความสามารถต่ำ เมื่อทำข้อสอบที่มีค่าความยากต่ำและข้อสอบที่มีค่าความยากปานกลาง ซึ่งเป็นค่าความยากที่เหมาะสมกับระดับความสามารถผู้สอบ และข้อสอบที่ทำมีค่าอำนาจจำแนกสูง ทำให้ผู้สอบมีความไวในการตอบข้อสอบ ซึ่งอาจมีผลทำให้เกิดการได้เปรียบและเสียเปรียบในการตอบข้อสอบได้ถูกต้องระหว่างกลุ่มผู้สอบ ดังนั้นในการสร้างข้อสอบจึงควรให้ความระมัดระวังปัจจัยที่อาจทำให้เกิดการทำหน้าที่ต่างกัน เช่นลักษณะผู้สอบกลุ่มย่อย ได้แก่ เพศ ภูมิภาค วัฒนธรรม เป็นต้น และในการสร้างแบบสอบควรให้ความระมัดระวังเกี่ยวกับ ลักษณะของภาษา คำที่ใช้ เนื้อหาของหลักสูตร เพื่อจะได้นำมาปรับปรุงหรือหลีกเลี่ยงการเกิด DIF ให้น้อยที่สุด ทั้งนี้เพื่อให้ได้แบบสอบที่มีความยุติธรรมต่อกลุ่มผู้สอบมากที่สุด

4. ผลการเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เชลกับวิธีถดถอยโลจิสติก พบว่าโดยภาพรวมทั้งสองวิธีมีประสิทธิภาพการตรวจสอบเท่ากัน ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอนุกรมด้วยวิธีแมนเทิล-แฮนส์เชลนั้น ใช้เกณฑ์การคัดเลือกผู้สอบเข้ากลุ่มโดยพิจารณาจากค่าเฉลี่ยของคะแนนรวมทั้งหมดเป็นเกณฑ์ในการคัดเลือกผู้สอบ ผู้สอบที่ได้คะแนนรวมสูงกว่าค่าเฉลี่ยจัดให้เป็นกลุ่มสูง ส่วนผู้สอบที่ได้คะแนนรวมน้อยกว่าหรือเท่ากับค่าเฉลี่ยจัดให้เป็นกลุ่มต่ำ จากนั้นจึงวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอนุกรมในกลุ่มสูงและกลุ่มต่ำ ซึ่งการวิเคราะห์ด้วยวิธีดังกล่าวนี้สอดคล้องกับผลการศึกษาของ Mazor (1994) ที่พบว่าเมื่อวิเคราะห์การทำหน้าที่ต่างกันแบบอนุกรมโดยแบ่งผู้สอบออกเป็นกลุ่มสูงและกลุ่มต่ำ จะทำให้วิธีแมนเทิล-แฮนส์เชลตรวจสอบ

ข้อสอบทำหน้าที่ต่างกันได้เพิ่มขึ้น และไม่ทำให้อัตราความคลาดเคลื่อนประเภทที่ 1 เพิ่มขึ้น จึงทำให้ผลการตรวจสอบข้อสอบทำหน้าที่ต่างกันแบบอนเนกรูปด้วยวิธีแมนเทิล-แฮนส์เชลมีประสิทธิภาพการตรวจสอบเท่าเทียมกับวิธีดอดอยโลจิสติก

5. ผลการเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เชลกับวิธีดอดอยโลจิสติก ในกลุ่มผู้สอบที่มีความสามารถสูง พบว่าวิธีดอดอยโลจิสติกมีอัตราความถูกต้องประเภทที่ 1 มากกว่าวิธี วิธีแมนเทิล-แฮนส์เชล ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง ทั้งนี้อาจเนื่องมาจากวิธีแมนเทิล-แฮนส์เชล ตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันแบบอนเนกรูปได้ดีเฉพาะในกรณีข้อสอบยากและข้อสอบง่าย ซึ่งสอดคล้องกับข้อค้นพบของ Rogers และ Swaminathan (1993) กับ Narayanan และ Swaminathan (1996) ที่พบว่าในกรณีข้อสอบง่ายและข้อสอบยาก วิธีแมนเทิล-แฮนส์เชลจะตรวจพบข้อสอบทำหน้าที่ต่างกันแบบอนเนกรูปได้สูง แต่ถ้าข้อสอบมีค่าความยากปานกลางวิธีแมนเทิล-แฮนส์เชลจะตรวจสอบได้น้อยมาก

สำหรับอัตราความคลาดเคลื่อนในการตรวจสอบ พบว่าวิธีแมนเทิล-แฮนส์เชลมีอัตราความคลาดเคลื่อนประเภทที่ 1 มากกว่าวิธีดอดอยโลจิสติก ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง ทั้งนี้อาจเป็นเพราะอัตราความคลาดเคลื่อนที่เพิ่มขึ้นสัมพันธ์กับอัตราความถูกต้องในการตรวจสอบที่ลดลง และจากข้อค้นพบของ Rogers และ Swaminathan (1993) กับ Narayanan และ Swaminathan (1996) ที่พบว่าวิธีแมนเทิล-แฮนส์เชลตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันแบบอนเนกรูปได้สูงเฉพาะในกรณีข้อสอบง่ายและข้อสอบยาก ส่วนข้อสอบที่มีค่าความยากปานกลางวิธีแมนเทิล-แฮนส์เชลตรวจสอบได้น้อยมาก ซึ่งอาจมีผลทำให้วิธีแมนเทิล-แฮนส์เชลมีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันสูงกว่าวิธีดอดอยโลจิสติก ในข้อสอบที่มีค่าความยากปานกลาง

6. ผลการเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เชลกับวิธีดอดอยโลจิสติก ในกลุ่มผู้สอบที่มีความสามารถปานกลาง พบว่าวิธีดอดอยโลจิสติกมีอัตราความถูกต้องประเภทที่ 2 มากกว่าวิธีแมนเทิล-แฮนส์เชล ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง ทั้งนี้เป็นเพราะวิธีแมนเทิล-แฮนส์เชล ตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันแบบอนเนกรูปได้สูงเฉพาะในกรณีข้อสอบง่ายและข้อสอบยาก แต่ถ้าข้อสอบมีค่าความยากปานกลางวิธีแมนเทิล-แฮนส์เชลจะตรวจสอบได้น้อยมาก (Rogers and Swaminathan, 1993 ; Narayanan and Swaminathan, 1996) และอาจเนื่องมาจากการกระจายความสามารถของผู้สอบระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบมีค่าเท่ากันหรือใกล้เคียงกัน มีผลต่ออัตราความถูกต้องประเภทที่ 2 หรือการตรวจสอบข้อสอบทำหน้าที่ต่างกันได้ถูกต้อง ดังข้อค้นพบของ Narayanan และ



- Swaminathan (1996) กล่าวคือเมื่อการกระจายความสามารถระหว่างกลุ่มผู้สอบเท่ากัน วิธีแมนเทิล-แฮนส์เชลตรวจสอบข้อสอบทำหน้าที่ต่างกันแบบอนุกรมได้บ่อยมากในข้อสอบที่มีค่าความยากปานกลาง ซึ่งข้อค้นพบดังกล่าวอาจทำให้ผลการตรวจสอบข้อสอบทำหน้าที่ต่างกันได้ถูกต้องของวิธีแมนเทิล-แฮนส์เชลต่ำกว่าวิธีถดถอยโลจิสติก ในข้อสอบที่มีค่าความยากปานกลาง สำหรับอัตราความคลาดเคลื่อนในการตรวจสอบ พบว่าวิธีแมนเทิล-แฮนส์เชลมีอัตราความคลาดเคลื่อนประเภทที่ 2 มากกว่าวิธีถดถอยโลจิสติก ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง ทั้งนี้เนื่องมาจากอัตราความคลาดเคลื่อนที่เพิ่มขึ้นนั้นสัมพันธ์กับอัตราความถูกต้องในการตรวจสอบที่ลดลง ซึ่งจากข้อค้นพบในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง พบว่าวิธีแมนเทิล-แฮนส์เชลมีอัตราความถูกต้องประเภทที่ 2 ต่ำกว่าวิธีถดถอยโลจิสติก ดังนั้นเมื่อตรวจสอบอัตราความคลาดเคลื่อนประเภทที่ 2 ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง จึงพบว่าวิธีแมนเทิล-แฮนส์เชลมีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ไม่ต่างกัน ทั้งที่ความจริงข้อสอบทำหน้าที่ต่างกันสูงกว่าวิธีถดถอยโลจิสติก

7. ผลการเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เชลกับวิธีถดถอยโลจิสติก ในกลุ่มผู้สอบที่มีความสามารถต่ำ พบว่าวิธีถดถอยโลจิสติกมีอัตราความถูกต้องประเภทที่ 2 มากกว่า วิธีแมนเทิล-แฮนส์เชล ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง ทั้งนี้เป็นเพราะวิธีแมนเทิล-แฮนส์เชล ตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันแบบอนุกรมได้สูงเฉพาะในกรณีข้อสอบง่ายและข้อสอบยาก แต่ถ้าข้อสอบมีค่าความยากปานกลางวิธีแมนเทิล-แฮนส์เชล จะตรวจสอบได้น้อยมาก (Rogers and Swaminathan, 1993 ; Narayanan and Swaminathan, 1996) ซึ่งอาจมีผลทำให้วิธีแมนเทิล-แฮนส์เชลตรวจสอบข้อสอบทำหน้าที่ต่างกันได้ถูกต้อง ต่ำกว่า วิธีถดถอยโลจิสติกในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง อีกทั้งอาจเนื่องมาจากข้อสอบที่ทำหน้าที่ต่างกันนั้นมีค่าดัชนีของการทำหน้าที่ต่างกันต่ำ หรือมีพื้นที่ที่เกิดจากความแตกต่างระหว่างโค้งคุณลักษณะข้อสอบของผู้สอบสองกลุ่มย่อยมีค่าน้อย ซึ่งมีผลต่ออัตราความถูกต้องประเภทที่ 2 หรือการตรวจสอบข้อสอบทำหน้าที่ต่างกันได้ถูกต้อง ดังข้อค้นพบของ Narayanan และ Swaminathan (1996) กล่าวคือ เมื่อค่าดัชนีของการทำหน้าที่ต่างกันมีค่าสูง จะตรวจพบข้อสอบทำหน้าที่ต่างกันแบบอนุกรมได้มากกว่าดัชนีที่มีค่าต่ำ และพบว่าวิธีถดถอยโลจิสติกมีอัตราความถูกต้องสูงกว่าวิธีแมนเทิล-แฮนส์เชลในทุกระดับของค่าดัชนีการทำงานที่ต่างกัน ซึ่งจากข้อค้นพบดังกล่าวอาจมีผลทำให้วิธีถดถอยโลจิสติกมีอัตราความถูกต้องประเภทที่ 2 มากกว่าวิธีแมนเทิล-แฮนส์เชลในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง

สำหรับอัตราความคลาดเคลื่อนในการตรวจสอบ พบว่าวิธีแมนเทล-แฮนส์เซลมีอัตราความคลาดเคลื่อนประเภทที่ 2 มากกว่าวิธีดอดยโลจิสติก ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง ทั้งนี้เนื่องมาจากอัตราความคลาดเคลื่อนที่เพิ่มขึ้นนั้นสัมพันธ์กับอัตราความถูกต้องในการตรวจสอบที่ลดลง ซึ่งจากข้อค้นพบในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง พบว่าวิธีแมนเทลแฮนส์เซลมีอัตราความถูกต้องประเภทที่ 2 ต่ำกว่าวิธีดอดยโลจิสติก ดังนั้นเมื่อตรวจสอบอัตราความคลาดเคลื่อนประเภทที่ 2 ในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกปานกลาง จึงพบว่าวิธีแมนเทล-แฮนส์เซลมีการระบุผิดพลาดว่าข้อสอบทำหน้าที่ไม่ต่างกัน ทั้งที่ความจริงข้อสอบทำหน้าที่ต่างกันสูงกว่าวิธีดอดยโลจิสติก

8. จากการตรวจสอบข้อสอบในกลุ่มผู้สอบที่มีความสามารถต่ำ พบว่าวิธีแมนเทล-แฮนส์เซลตรวจพบข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปได้ถูกต้องส่วนใหญ่อ้อยละ 0 ในขณะที่วิธีดอดยโลจิสติกตรวจพบข้อสอบทำหน้าที่ต่างกันได้ถูกต้องส่วนใหญ่อ้อยละ 100 แต่เมื่อทดสอบด้วยสถิติ Z แล้วพบว่าไม่สามารถปฏิเสธสมมติฐานศูนย์ได้ ทั้งนี้อาจเนื่องมาจากจำนวนข้อสอบที่ทำหน้าที่ต่างกันในแต่ละเงื่อนไขมีจำนวนน้อย ซึ่งข้อค้นพบของ Narayanan และ Swaminathan (1994) พบว่า สัดส่วนของข้อสอบที่ทำหน้าที่ต่างกันภายในข้อสอบมีผลกระทบต่ออัตราความถูกต้องในการตรวจสอบ กล่าวคือเมื่อสัดส่วนของข้อสอบที่ทำหน้าที่ต่างกันภายในข้อสอบเพิ่มขึ้น จะทำให้อัตราอัตราความถูกต้องในการตรวจสอบสูงขึ้น ดังนั้นน่าจะได้มีการศึกษาเพิ่มเติมเกี่ยวกับตัวแปรสัดส่วนของจำนวนข้อสอบที่ทำหน้าที่ต่างกันภายในแบบสอบ

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



## ข้อเสนอแนะ

### 1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 จากผลการวิจัยพบว่า โดยภาพรวมวิธีแมนเทิล-แฮนส์เชลกับวิธีถดถอยโลจิสติกมีประสิทธิภาพการตรวจสอบข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปเท่าเทียมกัน ในทุกระดับความสามารถผู้สอบ แต่บางเงื่อนไขของข้อสอบพบว่าวิธีถดถอยโลจิสติกมีประสิทธิภาพสูงกว่าวิธีแมนเทิล-แฮนส์เชล กล่าวคือวิธีถดถอยโลจิสติกมีอัตราความถูกต้องสูงกว่าวิธีแมนเทิล-แฮนส์เชล และมีอัตราความคลาดเคลื่อนต่ำกว่าวิธีแมนเทิล-แฮนส์เชล ดังนั้นวิธีที่มีความเหมาะสมและควรนำไปใช้ตรวจสอบข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปคือ วิธีถดถอยโลจิสติก

1.2 สำหรับการสอบคัดเลือกนั้น ต้องการให้ได้ผู้สอบที่มีความสามารถมากที่สุด หรือผู้สอบนั้นต้องมีความสามารถสูง ซึ่งข้อสอบที่เหมาะสมต่อการนำไปใช้ในการสอบคัดเลือก คือข้อสอบที่มีค่าความยากสูง ค่าอำนาจจำแนกสูง แต่ข้อสอบลักษณะดังกล่าวจากงานวิจัยพบว่าตรวจพบจำนวนข้อสอบทำหน้าที่ต่างกันมากที่สุด ดังนั้นในการสร้างข้อสอบคัดเลือกควรระมัดระวังปัจจัยที่เกี่ยวกับลักษณะของผู้สอบกลุ่มย่อย เช่น เพศ สถานศึกษา ภูมิภาค เป็นต้น เพราะอาจส่งผลต่อการทำข้อสอบที่มีค่าความยากสูง ค่าอำนาจจำแนกสูง และควรให้ความระมัดระวังเกี่ยวกับ ลักษณะของภาษา คำที่ใช้ เนื้อหาของหลักสูตร เพื่อจะได้นำมาปรับปรุงหรือหลีกเลี่ยงการเกิด DIF ให้น้อยที่สุด ทั้งนี้เพื่อให้เกิดความยุติธรรมต่อผู้สอบมากที่สุด

1.3 จากผลการวิจัยพบว่าในกลุ่มผู้สอบที่มีความสามารถปานกลาง จะพบข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปเป็นจำนวนมากในข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกสูง ดังนั้นในการสร้างข้อสอบเพื่อใช้สอบกับผู้สอบที่มีความสามารถปานกลาง ควรระมัดระวังปัจจัยเกี่ยวกับลักษณะของผู้สอบกลุ่มย่อย เช่น เพศ สถานศึกษา ภูมิภาค เป็นต้น เพราะอาจส่งผลต่อการทำข้อสอบที่มีค่าความยากปานกลาง ค่าอำนาจจำแนกสูง ทั้งนี้เพื่อให้เกิดความยุติธรรมต่อผู้สอบมากที่สุด

1.4 จากผลการวิจัยพบว่าในกลุ่มผู้สอบที่มีความสามารถต่ำ จะพบข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปเป็นจำนวนมากในข้อสอบที่มีค่าความยากต่ำ ค่าอำนาจจำแนกสูง ดังนั้นในการสร้างข้อสอบเพื่อใช้สอบกับผู้สอบที่มีความสามารถต่ำ ควรระมัดระวังปัจจัยที่เกี่ยวกับลักษณะของผู้สอบกลุ่มย่อย เช่น เพศ สถานศึกษา ภูมิภาค เป็นต้น เพราะอาจส่งผลต่อการทำข้อสอบที่มีค่าความยากต่ำ ค่าอำนาจจำแนกสูง ทั้งนี้เพื่อให้เกิดความยุติธรรมต่อผู้สอบมากที่สุด

## 2. ข้อเสนอแนะในการศึกษาเพิ่มเติม

2.1 ควรมีการศึกษาเพิ่มเติมเกี่ยวกับตัวแปรผู้สอบกลุ่มย่อยเช่น เพศ ภูมิภาค สถานศึกษา กับตัวแปรกลุ่มความสามารถผู้สอบ เพื่อศึกษาว่ามีผลต่อการระบุจำนวนข้อสอบที่ทำหน้าที่ต่างกันมากน้อยเพียงใด

2.2 จากผลการวิจัยพบว่าจำนวนข้อสอบที่ทำหน้าที่ต่างกันในข้อสอบมีผลต่อการตรวจสอบการทำหน้าที่ต่างกันแบบอนุกรม กล่าวคือเมื่อสัดส่วนของจำนวนข้อสอบที่ทำหน้าที่ต่างกันภายในแบบสอบเพิ่มขึ้น จะทำให้อัตราความถูกต้องในการตรวจสอบสูงขึ้น ดังนั้นจึงควรศึกษาเพิ่มเติมเกี่ยวกับสัดส่วนของจำนวนข้อสอบที่ทำหน้าที่ต่างกันในข้อสอบ เพื่อเปรียบเทียบผลการตรวจสอบเกี่ยวกับการใช้จำนวนข้อสอบที่ทำหน้าที่ต่างกัน

2.3 การกระจายความสามารถของกลุ่มผู้สอบระหว่างกลุ่มผู้สอบอ้างอิงกับกลุ่มเปรียบเทียบมีผลต่อการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ดังนั้นควรมีการศึกษาเพิ่มเติมเกี่ยวกับการกระจายความสามารถระหว่างกลุ่มผู้สอบสองกลุ่มย่อย กับข้อสอบลักษณะต่างๆ ว่ามีผลต่ออัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบมากน้อยเพียงใด

2.4 วิธีถดถอยโลจิสติกเป็นวิธีที่อยู่บนพื้นฐานของโมเดลถดถอยโลจิสติก ซึ่งเป็นโมเดลที่มีความยืดหยุ่น มีพื้นฐานเป็นแบบจำลองสามารถเพิ่มตัวแปรความสามารถเข้าไปในแบบจำลองได้ ดังนั้นควรศึกษาเพิ่มเติมเกี่ยวกับการนำวิธีถดถอยโลจิสติกมาใช้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ