บทที่ 5

สรุป อภิปราย และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์หลักเพื่อเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบชนิดพหุมิติ ระหว่างวิธีแมนเทล-แฮนส์เซล(MH)กับวิธีถดถอยโลจิสติก(LR)เมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกัน ซึ่งได้แก่ คะแนนรวม คะแนนแบบสอบย่อยและคะแนนหลายแบบสอบย่อย ส่วนวัตถุประสงค์เฉพาะของการวิจัย คือ เพื่อเปรียบเทียบประสิทธิภาพด้านอัตราความถูกต้องและอัตราความคลาดเคลื่อนของผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบชนิดพหุมิติโดยใช้วิธีในการตรวจสอบทั้งสองวิธีภายใต้เกณฑ์จับคู่เปรียบเทียบ 3 เกณฑ์

กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2540 ในโรงเรียนสังกัดสำนักงานการประถมศึกษากรุงเทพมหานคร จำนวน 1,076 คน

เครื่องมือที่ใช้ในการวิจัย คือ แบบทดสอบความสามารถทางคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 เป็นข้อสอบแบบเลือกตอบ 4 ตัวเลือกและมีความยาวแบบสอบ 80 ข้อ โดยมีเนื้อหาในการวัด 4 องค์ประกอบ ได้แก่ ความรู้ความเข้าใจทางคณิตศาสตร์ ความสามารถด้านการคำนวณ ความสามารถด้านการพิจารณาผลลัพธ์ที่สมเหตุสมผลและความสามารถด้านการแก้โจทย์ปัญหา

เกณฑ์ที่ใช้ในการวิเคราะห์ว่าข้อสอบทำหน้าที่ต่างกัน มีดังนี้

1. วิธีแมนเทล-แฮนส์เซล ได้แก่ ค่า $\alpha_{MH} \neq 1.0$ และการทดสอบนัยสำคัญด้วย χ^2_{MH} ที่ระดับ .05

2. วิธีถดถอยโลจิสติก ได้แก่ ค่าอิทธิพลจากปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบและความสามารถของผู้สอบหรือค่าอิทธิพลจากกลุ่มผู้สอบ และการทดสอบนัยสำคัญด้วย χ^2 ที่ระดับ .05

3. วิธี SIBTEST ได้แก่ ค่า $\beta_U > 0$ และการทดสอบนัยสำคัญด้วย Z ที่ระดับ .05

ในการวิจัยครั้งนี้ดำเนินการวิเคราะห์ข้อมูล ดังนี้

1. วิเคราะห์ค่าสถิติเบื้องต้น เพื่อให้ทราบลักษณะของผลการตอบข้อสอบและการแจกแจงของตัวแปรโดยใช้โปรแกรมคอมพิวเตอร์ SPSS/PC$^+$

2. วิเคราะห์คุณภาพข้อสอบตามทฤษฎีแบบดั้งเดิม (Classical Test Theory) เพื่อวิเคราะห์ค่าความยาก (p) ค่าอำนาจจำแนก(r) ของข้อสอบและค่าความเที่ยงของแบบสอบ โดยใช้โปรแกรมวิเคราะห์ข้อสอบ IRT(BAY) 1.0(ศิริชัย กาญจนวาสีและคณิต ไข่มุกด์, 2535)และโปรแกรมวิเคราะห์ข้อสอบ CTIA/ Item(สุพัฒน์ สุกมลสันต์, 2535)

3. ตรวจสอบความตรงเชิงโครงสร้าง (Construct Validity) โดยวิธีวิเคราะห์องค์ประกอบเชิงยืนยัน(Confirmatory Factor Analysis) โดยใช้โปรแกรมคอมพิวเตอร์ลิสเรล(LISREL 8.10)

4. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทล - แฮนส์เซล (MH) และวิธีซิบเทสท์ (SIBTEST) แยกตามการใช้เกณฑ์จับคู่เปรียบเทียบ ได้แก่ คะแนนรวมและคะแนนแบบสอบย่อย โดยนำผลการตอบข้อสอบที่ได้จากการเก็บข้อมูลมาคำนวณหาดัชนี α_{MH} และดัชนี β_{SIB} และทดสอบนัยสำคัญ ซึ่งค่าเหล่านี้สามารถคำนวณได้พร้อมกันด้วยโปรแกรม SIBTEST ของ Stout และ Roussos(1992)

5. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธีถดถอยโลจิสติก (LR) แยกตามการใช้เกณฑ์จับคู่เปรียบเทียบ ซึ่งได้แก่ คะแนนรวม คะแนนแบบสอบย่อย และคะแนนหลายแบบสอบย่อย โดยนำแฟ้มข้อมูลผลการตอบข้อสอบของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบไปวิเคราะห์ด้วยโปรแกรม SPSS/PC$^+$

6. คำนวณอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบ โดยนำผลการวิเคราะห์ด้วยวิธีแมนเทล - แฮนส์เซล (MH) และวิธีถดถอยโลจิสติก (LR) มาเปรียบเทียบกับผลการตรวจสอบด้วยวิธีซิบเทสท์ (SIBTEST) โดยจะใช้เกณฑ์ว่า ข้อสอบที่ถูกตรวจพบว่าทำหน้าที่ต่างกันด้วยวิธีซิบเทสท์ถือว่าเป็นข้อสอบที่ทำหน้าที่ต่างกัน(DIF) หากวิธีแมนเทล-แฮนส์เซลหรือวิธีถดถอยโลจิสติกตรวจพบข้อสอบที่ทำหน้าที่ต่างกันตรงกับที่ตรวจพบด้วยวิธีซิบเทสท์ ถือว่าวิธีนั้นมีความถูกต้องในการตรวจสอบ แต่ถ้าตรวจพบว่าข้อสอบข้อใดทำหน้าที่ต่างกันนอกเหนือจากที่ตรวจพบด้วยวิธีซิบเทสท์ ถือว่ามีความคลาดเคลื่อนในการตรวจสอบ

7. ทดสอบอัตราความถูกต้องและอัตราความคลาดเคลื่อนในการตรวจสอบ ระหว่างวิธีแมนเทล - แฮนส์เซล (MH) กับวิธีถดถอยโลจิสติก(LR) ซึ่งแยกตามเกณฑ์จับคู่เปรียบเทียบด้วยค่าสถิติ Z ที่ระดับนัยสำคัญ .05

สรุปผลการวิจัย

1. ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบขนิดพหุมิติ พบว่า

1.1 วิธีชิบเทสท์ ซึ่งเป็นวิธีที่นำมาใช้ประเมินประสิทธิภาพของการตรวจสอบมีข้อสอบที่ตรวจพบว่าทำหน้าที่ต่างกัน จำนวน 16 ข้อ (21.33%)

1.2 วิธีแมนเทล - แฮนส์เซล ใช้เกณฑ์จับคู่เปรียบเทียบ 2 เกณฑ์ ได้แก่ คะแนนรวมและคะแนนแบบสอบย่อยพบว่า วิธีแมนเทล-แฮนส์เซลเมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่ (MH_{TOT}) ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันจำนวน 15 ข้อ (20%) และเมื่อใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่ (MH_{SUB}) ตรวจพบจำนวน 14 ข้อ (18.67%)

1.3 วิธีถดถอยโลจิสติก ใช้เกณฑ์จับคู่เปรียบเทียบ 3 เกณฑ์ ได้แก่ คะแนนรวม คะแนนแบบสอบย่อยและคะแนนหลายแบบสอบย่อยพบว่า วิธีถดถอยโลจิสติกเมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่ (LR_{TOT}) ตรวจพบข้อสอบที่หน้าที่ต่างกันจำนวนมากที่สุด จำนวน 20 ข้อ (26.67%) รองลงมาคือ ใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่ (LR_{SUB}) ตรวจพบจำนวน 17 ข้อ (22.67%) และตรวจพบน้อยที่สุดคือ เมื่อใช้คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่ ($LR_{MULTIPLE}$) ตรวจพบจำนวน 13 ข้อ (17.33%)

2. ผลการเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบขนิดพหุมิติระหว่างวิธีแมนเทล-แฮนส์เซล (MH) กับวิธีถดถอยโลจิสติก (LR) สรุปได้ดังนี้

2.1 เมื่อใช้คะแนนรวม (total test score) เป็นเกณฑ์จับคู่เปรียบเทียบ วิธี MH มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี LR เนื่องจากมีอัตราความถูกต้องประเภทที่ 1 มากกว่าและมีอัตราความคลาดเคลื่อนประเภทที่ 1 น้อยกว่าอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2.2 เมื่อใช้คะแนนแบบสอบย่อย (subtest score) เป็นเกณฑ์จับคู่เปรียบเทียบ ทั้งสองวิธีมีประสิทธิภาพในการตรวจสอบไม่แตกต่างกัน

3. ผลการเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบขนิดพหุมิติด้วยวิธีถดถอยโลจิสติก ระหว่างการใช้คะแนนหลายแบบสอบย่อย (multiple subtest scores) เป็นเกณฑ์จับคู่กับการใช้คะแนนรวม (total test score) และคะแนนแบบสอบย่อย (subtest score) เป็นเกณฑ์จับคู่เปรียบเทียบ สรุปได้ดังนี้

3.1 วิธี $LR_{MULTIPLE}$ มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี LR_{TOT} เนื่องจากมีอัตราความถูกต้องประเภทที่ 1 มากกว่าและมีอัตราความคลาดเคลื่อนประเภทที่ 1 น้อยกว่าอย่างมีนัยสำคัญที่ระดับ 0.05

3.2 วิธี $LR_{MULTIPLE}$ มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี LR_{SUB} เนื่องจากมีอัตราความถูกต้องประเภทที่ 1 มากกว่าและมีอัตราความคลาดเคลื่อนประเภทที่ 1 น้อยกว่าอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

อภิปรายผลการวิจัย

การวิจัยครั้งนี้เป็นการเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบชนิดพหุมิติ ระหว่างวิธีแมนเทล - แฮนส์เซล (MH) กับวิธีถดถอยโลจิสติก (LR) เมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกัน ได้แก่ คะแนนรวม คะแนนแบบสอบย่อยและคะแนนหลายแบบสอบย่อย ซึ่งพิจารณาจากอัตราความถูกต้องและอัตราความคลาดเคลื่อนของผลการตรวจสอบ จากผลการวิจัยจะเห็นว่า มีทั้งส่วนที่สอดคล้องและไม่สอดคล้องกับสมมติฐานการวิจัย ผู้วิจัยจึงได้พิจารณาและนำเสนอรายละเอียดต่างๆ ดังต่อไปนี้

1. ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบชนิดพหุมิติ ระหว่างวิธีแมนเทล-แฮนส์เซลกับวิธีถดถอยโลจิสติกเมื่อใช้เกณฑ์จับคู่เปรียบเทียบ ได้แก่ คะแนนรวม คะแนนแบบสอบย่อย และคะแนนหลายแบบสอบย่อย ซึ่งมีระดับการศึกษา 6 ลักษณะ คือ MH_{TOT}, MH_{SUB}, LR_{TOT}, LR_{SUB} และ $LR_{MULTIPLE}$ พบว่า วิธีถดถอยโลจิสติกเมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่เปรียบเทียบเป็นวิธีที่ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันมากที่สุด และวิธีที่ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันน้อยที่สุด คือ วิธีถดถอยโลจิสติกเมื่อใช้คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่ ซึ่งข้อค้นพบดังกล่าวสอดคล้องกับผลการศึกษาของ Clauser และคณะ(1991), Ackerman(1992), Clauser และคณะ(1996)ที่พบว่า การตรวจสอบการทำหน้าที่ต่างกันในแบบสอบที่มีมิติในการวัดซับซ้อนโดยใช้คะแนนรวมเป็นเกณฑ์จับคู่นั้นไม่เหมาะสม เนื่องจากมีแนวโน้มสูงที่ระบุว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความจริงทำหน้าที่ไม่ต่างกัน นั่นคือ เกิดอัตราความคลาดเคลื่อนประเภทที่ 1 และผลการศึกษาของ Clauser และคณะ(1996) พบว่า เมื่อใช้คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่จะมีประสิทธิภาพในการตรวจสอบมากกว่าการใช้คะแนนรวมและคะแนนแบบสอบย่อย โดยอัตราความคลาดเคลื่อนประเภทที่ 1 ลดลงและจำนวนข้อสอบที่ถูกระบุว่าทำหน้าที่ต่างกันมีจำนวนลดลง

นอกจากนี้ผลการวิจัยยังสอดคล้องกับผลการศึกษาของ Mazor และ Clauser (1995) ที่พบว่าเมื่อนำตัวแทนความสามารถทางภาษาเข้ามาเป็นเกณฑ์จับคู่เปรียบเทียบร่วมกับคะแนนรวมในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่มีการวัดหลายมิติ (multidimensional item) โดย

การเพิ่มเข้าไปในสมการถดถอยโลจิสติก พบว่า ข้อสอบที่เคยถูกระบุว่าทำหน้าที่ต่างกันเมื่อตรวจสอบโดยใช้คะแนนรวมเป็นเกณฑ์จับคู่ถูกตัดทิ้งไป ทั้งนี้เนื่องจากข้อสอบเหล่านี้มีลักษณะเป็นข้อสอบพหุมิติซึ่งมีคุณลักษณะในการวัดหลายมิติ นั่นคือ ตัวแทนความสามารถทางภาษาอาจเป็นคุณ-ลักษณะที่สองที่เกี่ยวข้องกับข้อสอบบางข้อในลักษณะที่เป็นมิติที่ก่อกวนซึ่งไม่ใช่จุดมุ่งหมายในการวัด แต่ในขณะเดียวกันความสัมพันธ์เหล่านั้นกับข้อสอบอื่นอาจจะเป็นปัจจัยหลักของคุณ-ลักษณะแฝงที่ต้องการวัด เป็นผลทำให้กลุ่มผู้สอบที่มีความสามารถเท่ากันมีโอกาสตอบข้อสอบข้อนั้นได้ถูกไม่เท่ากัน ข้อสอบลักษณะนี้จึงถูกระบุอย่างผิด ๆ ว่าทำหน้าที่ต่างกัน ในสภาพการณ์เช่นนี้ Mazor และ Clauser เสนอแนะให้นำตัวแทนความสามารถหลากหลาย (multiple ability) เข้ามาวิเคราะห์ด้วยเพราะจะทำให้การตรวจสอบมีความถูกต้องมากขึ้น ซึ่งวิธีที่เหมาะสมในการตรวจสอบ คือ วิธีถดถอยโลจิสติก เนื่องจากมีความยืดหยุ่นสูงในการนำตัวแทนความสามารถที่หลากหลายเข้าสู่การวิเคราะห์

2. ผลการเปรียบเทียบประสิทธิภาพของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบชนิดพหุมิติระหว่างวิธีแมนเทล-แฮนส์เซลกับวิธีถดถอยโลจิสติกเมื่อใช้เกณฑ์จับคู่คะแนนรวมและคะแนนแบบสอบย่อย พบว่า เมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่เปรียบเทียบ วิธีแมนเทล-แฮนส์เซลมีประสิทธิภาพในการตรวจสอบมากกว่าวิธีถดถอยโลจิสติกซึ่งไม่สอดคล้องกับสมมติฐานการวิจัย และเมื่อใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบพบว่า ทั้งสองวิธีมีประสิทธิภาพในการตรวจสอบไม่แตกต่างกันซึ่งไม่สอดคล้องกับสมมติฐานการวิจัยอย่างไรก็ตาม ข้อค้นพบดังกล่าวสอดคล้องกับผลการศึกษาของ Clauser และคณะ(1996) ที่พบว่า การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบในแบบสอบที่มีมิติการวัดซับซ้อน ด้วยวิธีแมนเทล-แฮนส์เซลและวิธีถดถอยโลจิสติกจะให้ผลการตรวจสอบคล้ายคลึงกัน นั่นคือ เมื่อเปรียบเทียบทั้งสองวิธีโดยใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบสามารถตรวจสอบได้ดีเช่นเดียวกับการใช้คะแนนรวมเป็นเกณฑ์จับคู่ ซึ่งชี้ให้เห็นว่าประสิทธิภาพในการตรวจสอบของทั้งสองวิธีไม่แตกต่างกัน

3. เปรียบเทียบประสิทธิภาพการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบชนิดพหุมิติด้วยวิธีถดถอยโลจิสติกเมื่อใช้เกณฑ์จับคู่เปรียบเทียบซึ่งได้แก่ คะแนนรวม คะแนนแบบสอบย่อยและคะแนนหลายแบบสอบย่อยพบว่า การใช้คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่จะมีประสิทธิภาพในการตรวจสอบมากกว่าการใช้คะแนนรวมและคะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ ข้อค้นพบนี้สอดคล้องกับผลการศึกษาของ Mazor และ Clauser (1995) ที่พบว่า วิธีถดถอยโลจิสติกสามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ดีกว่าวิธีแมนเทล -

แฮนส์เซล ในแบบสอบที่มีตัวแปรความสามารถสองตัวหรือมากกว่า ทั้งนี้เป็นเพราะความยืดหยุ่นของวิธีถดถอยโลจิสติกที่มีพื้นฐานอยู่บนสมการถดถอยโลจิสติก ซึ่งสามารถเพิ่มตัวแปรความสามารถหลาย ๆ ตัวเข้าไปในการวิเคราะห์ได้ จึงทำให้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ดียิ่งขึ้น

นอกจากนี้ผลการศึกษาของ Clauser และคณะ(1996) ยังได้สนับสนุนว่า คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบที่เหมาะสมที่สุดในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบที่มีมิติการวัดซับซ้อน ทั้งนี้เนื่องจากมิติในการวัดแต่ละมิติที่มีอยู่หลากหลายในแบบสอบชนิดนี้มีการเปลี่ยนแปลงข้ามระหว่างข้อสอบทำให้ผลการตอบสนองข้อสอบบางข้อได้รับผลกระทบจากคุณลักษณะอื่นที่ไม่ใช่คุณลักษณะแฝงหลักที่ต้องการวัด ดังนั้นหากใช้คะแนนรวมเป็นเกณฑ์จับคู่จะระบุอย่างผิด ๆ ว่าข้อสอบเหล่านี้ทำหน้าที่ต่างกัน ทำให้เกิดอัตราความคลาดเคลื่อนประเภทที่ 1 สำหรับเกณฑ์จับคู่คะแนนแบบสอบย่อยเป็นในลักษณะเช่นเดียวกันแต่การตรวจสอบจะให้ผลดีกว่าเพราะการจัดชุดข้อสอบให้เป็นแบบสอบย่อยมีมิติในการวัดเพียงมิติเดียว ซึ่งคุณลักษณะที่เข้ามาก่อกวนจะถูกจัดให้เป็นมิติในการวัดหลักของแบบสอบย่อยที่สร้างขึ้น จะทำให้ผลการตรวจสอบเกิดความผิดพลาดน้อยลง

4. ผลการเปรียบเทียบประสิทธิภาพของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบชนิดพหุมิติในแต่ละวิธีเมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกัน พบว่า ทั้งสองวิธีให้ผลคล้ายคลึงกัน โดยเมื่อใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่จะตรวจพบข้อสอบที่ทำหน้าที่ต่างกันน้อยกว่าการใช้คะแนนรวมเป็นเกณฑ์ และเมื่อพิจารณาอัตราความคลาดเคลื่อนประเภทที่ 1 ซึ่งหมายถึง การระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกันทั้งที่ความจริงข้อสอบทำหน้าที่ไม่ต่างกัน พบว่า อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบด้วยวิธีแมนเทล-แฮนส์เซลเมื่อใช้คะแนนรวมและคะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่มีค่าเท่ากัน ส่วนวิธีถดถอยโลจิสติกเมื่อใช้คะแนนรวมเป็นเกณฑ์จับคู่มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่าการใช้คะแนนแบบสอบย่อย ข้อค้นพบเหล่านี้สอดคล้องกับผลการศึกษาของ Clauser และคณะ (1996) ที่พบว่า ในแบบสอบที่มีจุดมุ่งหมายในการวัดเพียงมิติเดียว การใช้คะแนนรวมเป็นเกณฑ์ที่ในการจับคู่กลุ่มผู้สอบในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบมีความเหมาะสมที่สุด ทั้งนี้เพราะเป็นตัวแทนความสามารถที่มีความเที่ยงในการวัดคุณลักษณะแฝงมากที่สุด ส่วนในแบบสอบที่มีมิติในการวัดหลายมิติ พบว่าการใช้คะแนนรวมเป็นเกณฑ์จับคู่นั้นทำให้เกิดความผิดพลาดของผลการตรวจสอบเนื่องจากรวมเอาผลการตอบข้อสอบที่ไม่สัมพันธ์กับคุณลักษณะแฝงที่ต้องการวัดเป็นข้อสอบที่ถูกตรวจพบว่าทำหน้าที่ต่างกัน แต่เมื่อใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จะให้ผลการตรวจ

สอบดีกว่า เนื่องจากแต่ละแบบสอบย่อยมีนิยามและจุดมุ่งหมายในการวัดปรากฏชัดเจนในตารางโครงสร้างซึ่งยืนยันว่าเป็นเอกมิติทำให้มีความเที่ยงสูงในการวัดมิตินั้น ๆ ดังนั้นผลการตรวจสอบการทำหน้าที่ต่างกันเมื่อใช้คะแนนแบบสอบย่อยเป็นเกณฑ์จึงเกิดความผิดพลาดน้อยกว่า

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

จากผลการวิจัย พบว่าการตรวจสอบการทำหน้าที่ต่างกันในแบบสอบชนิดพหุมิติด้วยวิธีถดถอยโลจิสติก เมื่อใช้คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ ตรวจพบจำนวนข้อสอบที่ทำหน้าที่ต่างกันน้อยที่สุด และเมื่อเปรียบเทียบกับการใช้เกณฑ์จับคู่คะแนนรวมและคะแนนแบบสอบย่อยซึ่งตรวจสอบด้วยวิธีถดถอยโลจิสติกเช่นเดียวกัน พบว่า การใช้คะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่มีประสิทธิภาพในการตรวจสอบมากกว่าการใช้คะแนนรวมและคะแนนแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ ดังนั้นวิธีที่มีความเหมาะสมและควรนำไปใช้ตรวจสอบการทำหน้าที่ต่างกัน ทั้งในแบบสอบเอกมิติ (unidimensional test) และแบบสอบพหุมิติ (multidimensional test) คือ วิธีถดถอยโลจิสติก เมื่อใช้คะะแนนหลายแบบสอบย่อยเป็นเกณฑ์จับคู่เปรียบเทียบ

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. เนื่องจากการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบพหุมิติ มีปัญหาในการหาวิธีเหมาะสมที่สุดที่จะใช้เป็นวิธีมาตรฐานที่จะระบุว่าข้อสอบข้อใดทำหน้าที่ต่างกัน เพื่อใช้เป็นเกณฑ์ในการประเมินประสิทธิภาพการตรวจสอบระหว่างวิธีแมนเทล-แฮนส์เซลและวิธีถดถอยโลจิสติกซึ่งในการวิจัยครั้งนี้ใช้วิธีซิบเทสท์ พบว่าผลการวิจัยโดยส่วนใหญ่สอดคล้องกับงานวิจัยต่าง ๆ ที่ศึกษาเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ดังนั้นจึงน่าจะมีการศึกษาเพิ่มเติมในการนำวิธีอื่นมาใช้เป็นวิธีมาตรฐานเพื่อให้ได้ผลการตรวจสอบที่มีประสิทธิภาพที่สุด ซึ่งวิธีที่น่าศึกษา ได้แก่ วิธีทฤษฎีการตอบสนองข้อสอบชนิดพหุมิติ

2. การใช้เกณฑ์จับคู่เปรียบเทียบของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในแบบสอบพหุมิติยังเป็นประเด็นที่น่าศึกษาเพิ่มเติมเนื่องจากมีผลต่อประสิทธิภาพในการตรวจสอบจากงานวิจัยที่ผ่านมาซึ่งศึกษาเกี่ยวกับการหาทางเลือกในการใช้เกณฑ์จับคู่กลุ่มผู้สอบ พบว่าส่วนใหญ่ใช้เกณฑ์ภายในเป็นเกณฑ์จับคู่ผู้สอบ ดังนั้นน่าจะมีการนำเอาเกณฑ์ภายนอกเข้ามาใช้ในการจับคู่ผู้สอบหรืออาจจะนำไปเป็นเกณฑ์จับคู่ร่วมกับเกณฑ์ภายในด้วย ทั้งนี้เพื่อหาทางเลือก

ในการใช้เกณฑ์จับคู่กลุ่มผู้สอบที่เหมาะสมที่สุดในการการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

3. น่าที่จะมีการศึกษาเพิ่มเติมเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในแบบสอบพหุมิติเมื่อจำแนกตามประเภทของการทำหน้าที่ต่างกัน ซึ่งได้แก่ การทำหน้าที่ต่างกันแบบเอกรูป (uniform DIF) และอเนกรูป (nonuniform DIF) เพื่อศึกษาว่าวิธีใด ภายใต้การใช้เกณฑ์จับคู่เปรียบเทียบใดที่สามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบชนิดพหุมิติได้ดีที่สุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย