

บทที่ 5

สรุปและอภิปราย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ ระหว่างวิธีการวิเคราะห์องค์ประกอบกำลังกับวิธีถดถอยโลจิสติก ภายใต้เงื่อนไข 18 เงื่อนไข $(3 \times 3 \times 2)$ คือ ขนาดของกลุ่มตัวอย่าง 3 ขนาด (2,000 คน 1,000 คน และ 300 คน) ความยาวของแบบทดสอบ 3 ขนาด (40 ข้อ 30 ข้อ และ 20 ข้อ) และเกณฑ์การจับคู่ 2 เกณฑ์ (คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย) เมื่อใช้ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยวิธีจับคู่สหสัมพันธ์เป็นเกณฑ์สำหรับการเปรียบเทียบประสิทธิภาพของวิธีตรวจสอบ

ข้อมูลที่ใช้เป็นข้อมูลทุติยภูมิ ซึ่งเป็นผลการตอบข้อสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติ ปีการศึกษา 2546 วิชาภาษาไทย ชั้นประถมศึกษาปีที่ 6 โดยผู้วิจัยขอความร่วมมือจากเจ้าหน้าที่สำนักทดสอบทางการศึกษา กระทรวงศึกษาธิการ ในการสุ่มผลการตอบข้อสอบกลุ่มตัวอย่าง เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ ปีการศึกษา 2546 ที่เข้าสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติ ได้มาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) แบบไม่กำหนดสัดส่วน ซึ่งแบ่งเป็น ดี พอใช้ และปรับปรุง มีนักเรียนเป็นหน่วยการสุ่ม ใช้กลุ่มตัวอย่างทั้งหมดจำนวน 2,000 คน แบ่งเป็นเพศชาย 1,000 คน และเพศหญิง 1,000 คน

วิเคราะห์ข้อมูลค่าสถิติพื้นฐานโดยใช้โปรแกรม SPSS ตรวจสอบความตรงเชิงโครงสร้าง โดยวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โดยใช้โปรแกรม LISREL 8.50 และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้โปรแกรม SIBTEST, LISREL 8.50 และ SPSS

สรุปผลการวิจัย

1. ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ ปรากฏว่า
 - 1.1 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับ
 - 1.1.1 วิธีการวิเคราะห์องค์ประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกันจำนวน 24 ข้อ
 - 1.1.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 18 ข้อ

1.2 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนน
แบบทดสอบย่อย

1.2.1 วิธีการวิเคราะห์องค์ประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 26 ข้อ

1.2.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 18 ข้อ

1.3 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวม
ทั้งฉบับ

1.3.1 วิธีการวิเคราะห์องค์ประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 17 ข้อ

1.3.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 15 ข้อ

1.4 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนน
แบบทดสอบย่อย

1.4.1 วิธีการวิเคราะห์องค์ประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 22 ข้อ

1.4.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 15 ข้อ

1.5 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวม
ทั้งฉบับ

1.5.1 วิธีการวิเคราะห์องค์ประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 10 ข้อ

1.5.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 10 ข้อ

1.6 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนน
แบบทดสอบย่อย

1.6.1 วิธีการวิเคราะห์องค์ประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 6 ข้อ

1.6.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 6 ข้อ

1.7 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวม
ทั้งฉบับ

1.7.1 วิธีการวิเคราะห์องค์ประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 20 ข้อ

1.7.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 10 ข้อ

1.8 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนน
แบบทดสอบย่อย

1.8.1 วิธีการวิเคราะห์หองคัประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 23 ข้อ

1.8.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 10 ข้อ

1.9 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวม
ทั้งฉบับ

1.9.1 วิธีการวิเคราะห์หองคัประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 17 ข้อ

1.9.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 9 ข้อ

1.10 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนน
แบบทดสอบย่อย

1.10.1 วิธีการวิเคราะห์หองคัประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 21 ข้อ

1.10.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 9 ข้อ

1.11 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวม
ทั้งฉบับ

1.11.1 วิธีการวิเคราะห์หองคัประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 9 ข้อ

1.11.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 6 ข้อ

1.12 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนน
แบบทดสอบย่อย

1.12.1 วิธีการวิเคราะห์หองคัประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 6 ข้อ

1.12.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 8 ข้อ

1.13 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวม
ทั้งฉบับ

1.13.1 วิธีการวิเคราะห์หองคัประกอบจำกัด ตรวจพบข้อสอบทำหน้าที่ต่างกัน
จำนวน 19 ข้อ

1.13.2 วิธีถดถอยโลจิสติก ตรวจพบข้อสอบทำหน้าที่ต่างกัน จำนวน 5 ข้อ

1.14 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนแบบทดสอบย่อย

1.14.1 วิธีการวิเคราะห์ห้อยค์ประกอบจำกัด ตรวจสอบข้อสอบทำห้ำที่ต่างกัน
จำนวน 18 ข้อ

1.14.2 วิธีดคดอยโลจิสติก ตรวจสอบข้อสอบทำห้ำที่ต่างกัน จำนวน 5 ข้อ

1.15 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวม
ทั้งฉบับ

1.15.1 วิธีการวิเคราะห์ห้อยค์ประกอบจำกัด ตรวจสอบข้อสอบทำห้ำที่ต่างกัน
จำนวน 14 ข้อ

1.15.2 วิธีดคดอยโลจิสติก ตรวจสอบข้อสอบทำห้ำที่ต่างกัน จำนวน 5 ข้อ

1.16 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนน
แบบทดสอบย่อย

1.16.1 วิธีการวิเคราะห์ห้อยค์ประกอบจำกัด ตรวจสอบข้อสอบทำห้ำที่ต่างกัน
จำนวน 14 ข้อ

1.16.2 วิธีดคดอยโลจิสติก ตรวจสอบข้อสอบทำห้ำที่ต่างกัน จำนวน 5 ข้อ

1.17 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวม
ทั้งฉบับ

1.17.1 วิธีการวิเคราะห์ห้อยค์ประกอบจำกัด ตรวจสอบข้อสอบทำห้ำที่ต่างกัน
จำนวน 8 ข้อ

1.17.2 วิธีดคดอยโลจิสติก ตรวจสอบข้อสอบทำห้ำที่ต่างกัน จำนวน 2 ข้อ

1.18 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนน
แบบทดสอบย่อย

1.18.1 วิธีการวิเคราะห์ห้อยค์ประกอบจำกัด ตรวจสอบข้อสอบทำห้ำที่ต่างกัน
จำนวน 5 ข้อ

1.18.2 วิธีดคดอยโลจิสติก ตรวจสอบข้อสอบทำห้ำที่ต่างกัน จำนวน 4 ข้อ

2. ผลการเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำห้ำที่ต่างกันของข้อสอบ
ในแบบทดสอบพหุมิติระหว่างวิธีการวิเคราะห์ห้อยค์ประกอบจำกัด (RFA) กับวิธีดคดอยโลจิสติก
(LR) ภายใต้เงื่อนไขที่ศึกษา 18 เงื่อนไข ปรากฏว่า

2.1 เมื่อพิจารณาจากอัตราความถูกต้องประเภทที่ 1 และอัตราความคลาดเคลื่อน
ประเภทที่ 1 ปรากฏว่า ทั้งสองวิธีมีประสิทธิภาพในการตรวจสอบไม่แตกต่างกัน 2 เงื่อนไข คือ

2.1.1 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนแบบทดสอบย่อย

2.1.2 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนแบบทดสอบย่อย

2.2 วิธี LR มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี RFA เนื่องจากวิธี LR มีอัตราความถูกต้องประเภทที่ 1 มากกว่า และอัตราความคลาดเคลื่อนประเภทที่ 1 น้อยกว่าวิธี RFA อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 16 เงื่อนไข คือ

2.2.1 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.2.2 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.2.3 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.2.4 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.2.5 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.2.6 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับ

2.2.7 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.2.8 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.2.9 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับ

2.3 เมื่อพิจารณาจากอัตราความถูกต้องประเภทที่ 2 และอัตราความคลาดเคลื่อนประเภทที่ 2 ปรากฏว่า ทั้งสองวิธีมีประสิทธิภาพในการตรวจสอบไม่แตกต่างกัน 5 เงื่อนไข คือ

2.3.1 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมและคะแนนแบบทดสอบย่อย

2.3.2 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนแบบทดสอบย่อย

2.3.3 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมและคะแนนแบบทดสอบย่อย

2.4 วิธี RFA มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี LR เนื่องจากวิธี LR มีอัตราความถูกต้องประเภทที่ 2 มากกว่า และอัตราความคลาดเคลื่อนประเภทที่ 2 น้อยกว่าวิธี LR อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 2 เงื่อนไข คือ

2.4.1 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนแบบทดสอบย่อย

2.4.2 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับ

2.5 วิธี LR มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี RFA เนื่องจากวิธี LR มีอัตราความถูกต้องประเภทที่ 2 มากกว่า และอัตราความคลาดเคลื่อนประเภทที่ 2 น้อยกว่าวิธี RFA อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 11 เงื่อนไข คือ

2.5.1 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับ

2.5.2 เมื่อกลุ่มตัวอย่าง 2,000 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.5.3 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับ

2.5.4 เมื่อกลุ่มตัวอย่าง 1,000 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.5.5 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 40 ข้อ และใช้เกณฑ์การจับคู่คะแนนแบบทดสอบย่อย

2.5.6 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 30 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.5.7 เมื่อกลุ่มตัวอย่าง 300 คน แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย

2.6 โดยภาพรวมวิธี LR มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี RFA อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

อภิปรายผลการวิจัย

1. ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ ระหว่างวิธีการวิเคราะห์องค์ประกอบจำกัด (RFA) กับวิธีถดถอยโลจิสติก (LR) ภายใต้เงื่อนไขที่ศึกษา 18 เงื่อนไข ปรากฏว่า ทั้งสองวิธีมีประสิทธิภาพในการตรวจสอบไม่แตกต่างกัน 2 เงื่อนไข คือ เมื่อกลุ่มตัวอย่างขนาดกลาง (1,000 คน) แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนแบบทดสอบย่อย และเมื่อกลุ่มตัวอย่างขนาดเล็ก (300 คน) แบบทดสอบ 20 ข้อ และใช้เกณฑ์การจับคู่คะแนนแบบทดสอบย่อย ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 12 และ 18 ทั้งนี้อาจเป็นเพราะในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธี RFA ไม่ขึ้นอยู่กับขนาดของกลุ่มตัวอย่าง สอดคล้องกับผลการศึกษาของ ออร์ต (Oort, 1998) ปรากฏว่า อำนาจการตรวจสอบและอัตราความคลาดเคลื่อนของวิธี RFA ไม่ขึ้นอยู่กับขนาดกลุ่มตัวอย่างและความยาวของแบบทดสอบ ซึ่งจากกรณีนี้ขนาดของแบบทดสอบเท่ากัน คือ 20 ข้อ สอดคล้องกับผลการศึกษาของ เคลาสเซอร์ และคณะ (Clauser et al., 1996) ปรากฏว่า การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบที่มีมิติการวัดซับซ้อน ด้วยวิธีแมนเทิล-แฮนส์เซลและวิธีถดถอยโลจิสติก จะให้ผลการตรวจสอบคล้ายกัน นั่นคือ เมื่อเปรียบเทียบทั้งสองวิธีโดยใช้คะแนนแบบทดสอบย่อยเป็นเกณฑ์การจับคู่เปรียบเทียบสามารถตรวจสอบได้ดีเช่นเดียวกับการใช้คะแนนรวมเป็นเกณฑ์การจับคู่ ซึ่งชี้ให้เห็นว่าประสิทธิภาพในการตรวจสอบของทั้งสองวิธีไม่แตกต่างกัน

2. ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ ระหว่างวิธีการวิเคราะห์องค์ประกอบจำกัด (RFA) กับวิธีถดถอยโลจิสติก (LR) ภายใต้เงื่อนไขที่ศึกษา 18 เงื่อนไข ปรากฏว่า วิธี LR มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี RFA 16 เงื่อนไข คือ เมื่อกลุ่มตัวอย่างขนาดใหญ่ (2,000 คน) มี 6 เงื่อนไข คือ แบบทดสอบทั้ง 3 ขนาด และเกณฑ์การจับคู่ทั้ง 2 เกณฑ์ สำหรับกลุ่มตัวอย่างขนาดกลาง (1,000 คน) มี 5 เงื่อนไข คือ แบบทดสอบทั้ง 3 ขนาด และเกณฑ์การจับคู่ทั้ง 2 เกณฑ์ ยกเว้นแบบทดสอบ 20 ข้อ และเกณฑ์การจับคู่คะแนนแบบทดสอบย่อย ส่วนกลุ่มตัวอย่างขนาดเล็ก (300 คน) มี 5 เงื่อนไข คือ แบบทดสอบทั้ง 3 ขนาด และเกณฑ์การจับคู่ทั้ง 2 เกณฑ์ ยกเว้นแบบทดสอบ 20 ข้อ และเกณฑ์การจับคู่คะแนนแบบทดสอบย่อย ซึ่งข้อค้นพบดังกล่าวไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะว่าในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ วิธี LR มีประสิทธิภาพในการตรวจสอบมากกว่า เกือบทุกเงื่อนไข เมื่อใช้กลุ่มตัวอย่างขนาดใหญ่ ขนาดกลาง และขนาดเล็ก โดยพิจารณาจากอัตราความถูกต้อง และอัตราความคลาดเคลื่อน สอดคล้องกับผลการศึกษาของ เมเซอร์ (Mazor, 1992) สวามินาธาน และโรเจอร์ (Swaminathan & Roger, 1990) ปรากฏว่า เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้นจะทำให้อำนาจการตรวจสอบมากขึ้น

สอดคล้องกับผลการศึกษาของ อัททาโร และมิลแซพ (Uttaro & Millsap, 1994) ปรากฏว่า แบบทดสอบที่มีความยาวและกลุ่มตัวอย่างขนาดใหญ่ ทำให้มีโอกาสตรวจพบข้อสอบทำหน้าที่ต่างกันได้มากขึ้น สอดคล้องกับผลการศึกษาของ เคลาเซอร์ และคณะ (Clauser et al., 1996) ปรากฏว่า การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบที่มีมิติการวัดซับซ้อน ระหว่างวิธีแมนเทล-แฮนส์เชลกับวิธีถดถอยโลจิสติก ให้ผลการตรวจสอบคล้ายคลึงกัน นั่นคือ เมื่อเปรียบเทียบทั้งสองวิธีโดยใช้คะแนนแบบทดสอบย่อยเป็นเกณฑ์การจับคู่สามารถตรวจสอบได้ดีเช่นเดียวกับการใช้คะแนนรวมทั้งฉบับเป็นเกณฑ์การจับคู่ ซึ่งชี้ให้เห็นว่าทั้งสองวิธีมีประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบไม่แตกต่างกัน และผลการศึกษาของ เคลาเซอร์ และคณะ (Clauser et al., 1996) ปรากฏว่า แบบทดสอบที่มีจุดมุ่งหมายในการวัดเพียงมิติเดียวนั้น การใช้คะแนนรวมทั้งฉบับในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบมีความเหมาะสมที่สุด เพราะเป็นตัวแทนความสามารถที่มีความตรงในการวัดคุณลักษณะแฝงมากที่สุด ส่วนแบบทดสอบที่มีการวัดในหลายมิติ การใช้คะแนนรวมทั้งฉบับในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบทำให้เกิดความผิดพลาดของผลการตรวจสอบ เนื่องจากคะแนนผลการสอบที่ไม่สัมพันธ์กับคุณลักษณะแฝงที่ต้องการวัดเป็นข้อสอบที่ตรวจพบว่าทำหน้าที่ต่างกัน แต่เมื่อใช้คะแนนแบบทดสอบย่อยจะให้ผลการตรวจสอบดีกว่า นอกจากนี้ เมเซอร์ และเคลาเซอร์ (Mazor & Clauser, 1995) เสนอแนะให้นำตัวแทนความสามารถหลากหลาย (Multiple Ability) เข้ามาวิเคราะห์ด้วยเพราะจะทำให้การตรวจสอบมีความถูกต้องมากขึ้น ซึ่งวิธีที่เหมาะสมในการตรวจสอบคือ วิธีถดถอยโลจิสติก เนื่องจากมีความยืดหยุ่นสูงในการนำตัวแทนความสามารถที่หลากหลายเข้าสู่การวิเคราะห์ ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ ระหว่างวิธีการวิเคราะห์หองค์ประกอบจำกัด (RFA) กับวิธีถดถอยโลจิสติก(LR) ภายใต้งี้นไขที่ศึกษา 18 งี้นไข ปรากฏว่า โดยภาพรวมวิธี LR มีประสิทธิภาพในการตรวจสอบมากกว่าวิธี RFA ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติด้วยวิธี LR มีอัตราความถูกต้องประเภทที่ 1, 2 มากกว่าและอัตราความคลาดเคลื่อนประเภทที่ 1, 2 น้อยกว่าวิธี RFA เกือบทุกงี้นไขที่ศึกษา สอดคล้องกับผลการศึกษาของ เมเซอร์ และเคลาเซอร์ (Mazor & Clauser, 1995) ปรากฏว่า วิธีถดถอยโลจิสติกสามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ดีกว่าวิธีแมนเทล-แฮนส์เชล ในแบบทดสอบที่มีตัวแปรความสามารถสองตัวหรือมากกว่า ทั้งนี้เป็นเพราะความยืดหยุ่นของวิธีถดถอยโลจิสติกที่มีพื้นฐานอยู่บนสมการถดถอยโลจิสติก ซึ่งสามารถเพิ่มตัวแปรความสามารถหลาย ๆ ตัวเข้าไปในการวิเคราะห์ จึงทำให้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ดียิ่งขึ้น

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. จากผลการวิจัยปรากฏว่า การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีการวิเคราะห์ห้วงค์ประกอบจำกัดกับวิธีดัดดอยโลจิสติก ภายใต้เงื่อนไข 18 เงื่อนไข (3x3x2) คือ ขนาดของกลุ่มตัวอย่าง 3 ขนาด (2,000 คน 1,000 คน และ 300 คน) ความยาวของแบบทดสอบ 3 ขนาด (40 ข้อ 30 ข้อ และ 20 ข้อ) และใช้เกณฑ์การจับคู่ 2 เกณฑ์ (คะแนนรวมทั้งฉบับและคะแนนแบบทดสอบย่อย) เมื่อใช้ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยวิธีชิปเทสท์เป็นเกณฑ์สำหรับการเปรียบเทียบประสิทธิภาพของวิธีตรวจสอบ โดยภาพรวมวิธีดัดดอยโลจิสติกมีประสิทธิภาพในการตรวจสอบมากกว่าวิธีการวิเคราะห์ห้วงค์ประกอบจำกัด เนื่องจากวิธีดัดดอยโลจิสติกมีอัตราความถูกต้องประเภทที่ 1, 2 มากกว่า และมีอัตราความคลาดเคลื่อนประเภทที่ 1, 2 น้อยกว่าวิธีการวิเคราะห์ห้วงค์ประกอบจำกัด ดังนั้น วิธีที่เหมาะสมและควรนำไปใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ คือ วิธีดัดดอยโลจิสติก

2. ขนาดของกลุ่มตัวอย่าง มีผลต่อการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ คือ เมื่อกลุ่มตัวอย่างขนาดใหญ่ตรวจพบข้อสอบทำหน้าที่ต่างกัน ได้ดีกว่ากลุ่มตัวอย่างขนาดเล็ก

3. ความยาวของแบบทดสอบ มีผลต่อการตรวจพบข้อสอบทำหน้าที่ต่างกันของข้อสอบ คือ แบบทดสอบที่มีขนาดความยาวมากจะตรวจพบข้อสอบทำหน้าที่ต่างกัน ได้ดีกว่าแบบทดสอบที่มีขนาดความยาวน้อย

4. เกณฑ์การจับคู่ สำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ ที่เหมาะสมในการตรวจสอบคือ เกณฑ์การจับคู่คะแนนแบบทดสอบย่อย

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. จากการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ มีปัญหาในการหาวิธีที่เหมาะสมที่สุดที่ใช้เป็นวิธีมาตรฐานที่จะระบุว่าข้อสอบข้อใดทำหน้าที่ต่างกัน เพื่อใช้เป็นเกณฑ์ในการประเมินประสิทธิภาพการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบระหว่างวิธีการวิเคราะห์ห้วงค์ประกอบจำกัดกับวิธีดัดดอยโลจิสติก ซึ่งในการศึกษาครั้งนี้ผู้วิจัยใช้วิธีชิปเทสท์ ดังนั้นจึงน่าจะศึกษาเพิ่มเติมในการนำวิธีอื่นมาใช้เป็นวิธีมาตรฐานเพื่อให้ได้ผลการตรวจสอบที่มีประสิทธิภาพที่สุด เช่น วิธีทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นต้น

2. ควรมีการศึกษาวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ วิชาอื่น ๆ เช่น คณิตศาสตร์ สังคมศาสตร์ และภาษาอังกฤษ เป็นต้น เมื่อจำแนกกลุ่มผู้เข้าสอบ

ตามตัวแปรต่าง ๆ เช่น ภูมิปัญญา ศาสนา เชื้อชาติ ภาษา อายุ เป็นต้น ว่าแบบทดสอบแต่ละวิชาจะส่งผลให้พบข้อสอบทำหน้าที่ย่างกัน(DIF) แตกต่างกันหรือไม่

3. ควรมีการศึกษาเพิ่มเติมเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติ เมื่อจำแนกตามประเภทของการทำหน้าที่ต่างกัน ซึ่งได้แก่ การทำหน้าที่ต่างกันแบบเอกรูป (Unifrom DIF) และอนเอกรูป (Nonunifrom DIF) เพื่อศึกษาว่าวิธีใด ภายใต้การใช้เกณฑ์จับคู่แบบใดที่สามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบพหุมิติได้ดีที่สุด