

บทที่ 5

สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันแบบไม่สมมาตรของข้อสอบ ระหว่างวิธีแมนเทิล-แฮนส์เชล แบบปกติ กับ วิธีแมนเทิล-แฮนส์เชล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ โดยใช้ข้อมูลผลการตอบแบบสอบวัดความสามารถในการอ่านภาษาไทย ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2539 จำนวน 1,200 คน จำแนกเป็นเพศชาย 600 คน และเพศหญิง 600 คน โดยมีวัตถุประสงค์เฉพาะ ดังนี้

เพื่อเปรียบเทียบสัดส่วนการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สมมาตร ระหว่างวิธีแมนเทิล-แฮนส์เชล แบบปกติ กับ วิธีแมนเทิล-แฮนส์เชล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ เมื่อใช้จำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สมมาตรที่ตรวจพบ ตามวิธี IRT area เป็นเกณฑ์

กลุ่มตัวอย่างในการวิจัยครั้งนี้ เป็นนักเรียนชายและนักเรียนหญิง ชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2539 ของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา อำเภอเมือง จังหวัดชลบุรี จำนวน 1,200 คน จำแนกเป็นเพศชาย 600 คน และเพศหญิง 600 คน ซึ่งได้มาโดยวิธีสุ่มตัวอย่างแบบหลายขั้นตอน และมีนักเรียนเป็นหน่วยการสุ่ม

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล เป็นแบบสอบวัดความสามารถในการอ่านภาษาไทย ชั้นมัธยมศึกษาปีที่ 1 ชนิดเลือกตอบ 4 ตัวเลือก จำนวน 75 ข้อ ซึ่งผู้วิจัยสร้างขึ้น การวิเคราะห์ข้อมูล ทำการวิเคราะห์ข้อมูลเบื้องต้นเพื่อประมาณค่าความสามารถของผู้สอบรายคน ($\hat{\theta}$) และค่าความยากของข้อสอบรายข้อ (\hat{b}) ตามทฤษฎีการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ โดยใช้โปรแกรมคอมพิวเตอร์ BILOG Ver 3.04 (1990) ต่อจากนั้นคำนวณค่า Exact Unsigned Area Measure (EUA) ในกรณี $a_1 \neq a_2$ เพื่อหาจำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สมมาตร ตามวิธี IRT area แบบ 2 พารามิเตอร์ โดยใช้โปรแกรมคอมพิวเตอร์ 3 โปรแกรม ได้แก่ BILOG Ver 3.04 (1990), EQUATE Ver 2.0 (1991) และ IRTDIF Ver 1.0 (1992) การคำนวณค่าสถิติ $MH-\chi^2$ เพื่อทดสอบการทำหน้าที่ต่างกันของข้อสอบ ตามวิธีแมนเทิล-แฮนส์เชล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ โดยใช้โปรแกรมคอมพิวเตอร์ MHDIF (1994) และขั้นตอน

สุดท้ายเป็นการคำนวณค่าสถิติ $MH-\chi^2$ เพื่อทดสอบการทำหน้าที่ต่างกันของข้อสอบ ตามวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ โดยการวิเคราะห์ข้อมูลจากการจำลองกลุ่มตัวอย่างจากข้อมูลจริง จำนวน 10 ครั้ง ขนาดกลุ่มตัวอย่าง 1,000 คน จำแนกเป็นเพศชาย 500 คน และเพศหญิง 500 คน การวิเคราะห์ข้อมูลใช้โปรแกรมคอมพิวเตอร์ MHDIF (1994)

สรุปผลการวิจัย

1. วิธี IRT area แบบ 2 พารามิเตอร์ ตรวจสอบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ในแบบสอบวัดความสามารถในการอ่านภาษาไทย จำนวน 18 ข้อ ได้แก่ ข้อที่ 1, 3, 5, 18, 20, 23, 28, 38, 40, 41, 42, 43, 45, 52, 57, 61, 64, 71

ลักษณะของข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่พบ เมื่อพิจารณาลักษณะเนื้อหาของสาระของข้อสอบ พบว่า เป็นข้อสอบที่มีเนื้อหาสาระทั่วไป จำนวน 17 ข้อ ได้แก่ ข้อที่ 1, 3, 5, 18, 20, 23, 28, 38, 40, 41, 42, 43, 45, 52, 57, 61, 64 เป็นข้อสอบที่มีเนื้อหาสาระเกี่ยวข้องกับเพศหญิงมากกว่าเพศชาย จำนวน 1 ข้อ ได้แก่ ข้อที่ 71 เมื่อพิจารณาค่าความยากของข้อสอบ พบว่า เป็นข้อสอบยาก จำนวน 2 ข้อ ได้แก่ ข้อที่ 5, 20 เป็นข้อสอบยากง่ายปานกลาง จำนวน 13 ข้อ ได้แก่ ข้อที่ 3, 18, 23, 28, 38, 42, 43, 45, 52, 57, 61, 64, 71 และเป็นข้อสอบง่าย จำนวน 3 ข้อ ได้แก่ 1, 40, 41

ข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจสอบ ตามวิธี IRT area แบบ 2 พารามิเตอร์ จำนวน 18 ข้อนี้ ใช้เป็นเกณฑ์สำหรับเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันแบบไม่สม่ำเสมอของข้อสอบ ระหว่างวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับวิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ โดยมีหลักเกณฑ์การตัดสินใจ ข้อสอบทำหน้าที่ต่างกันที่ตรวจสอบ ตามวิธีแมนเทิล-แฮนส์เซล ว่าเป็นข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ คือ สอดคล้องกับข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ตามวิธี IRT area แบบ 2 พารามิเตอร์

2. วิธีแมนเทิล-แฮนส์เซล แบบปกติ ตรวจสอบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ในแบบสอบฉบับเดียวกัน สอดคล้องกับวิธี IRT area แบบ 2 พารามิเตอร์ จำนวน 6 ข้อ ได้แก่ ข้อที่ 1, 41, 45, 52, 61, 71 เมื่อพิจารณาค่า MH D-DIF พบว่าเป็นข้อสอบทำหน้าที่ต่างกันโดยเข้าข้างกลุ่มผู้สอบเพศหญิง จำนวน 3 ข้อ ได้แก่ ข้อที่ 45, 52, 71 และเป็น

ข้อสอบทำหน้าที่ต่างกันโดยเข้าข้างกลุ่มผู้สอบเพศชาย จำนวน 3 ข้อ ได้แก่ ข้อที่ 1, 41, 61 เมื่อพิจารณาลักษณะเนื้อหาสาระของข้อสอบ พบว่าเป็นข้อสอบที่มีเนื้อหาสาระทั่วไป จำนวน 5 ข้อ ได้แก่ ข้อที่ 1, 41, 43, 52, 61 เป็นข้อสอบที่มีเนื้อหาสาระเกี่ยวข้องกับเพศหญิงมากกว่าเพศชาย จำนวน 1 ข้อ ได้แก่ ข้อที่ 71 เมื่อพิจารณาค่าความยากของข้อสอบ พบว่า เป็นข้อสอบยากง่ายปานกลาง จำนวน 4 ข้อ ได้แก่ ข้อที่ 45, 52, 61, 71 และเป็นข้อสอบง่าย จำนวน 2 ข้อ ได้แก่ ข้อที่ 1, 41 ส่วนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอที่ตรวจไม่พบ (False Negative) จำนวน 12 ข้อ ได้แก่ ข้อที่ 3, 5, 18, 20, 23, 28, 38, 40, 42, 43, 57, 64

3. วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ ในแบบสอบฉบับเดียวกัน สอดคล้องกับวิธี IRT area แบบ 2 พารามิเตอร์ จำนวน 11 ข้อ ได้แก่ ข้อที่ 1, 3, 18, 28, 40, 41, 43, 45, 52, 61, 71 เมื่อพิจารณาลักษณะเนื้อหาสาระของข้อสอบ พบว่าเป็นข้อสอบที่มีเนื้อหาสาระทั่วไป จำนวน 10 ข้อ ได้แก่ ข้อที่ 1, 3, 18, 28, 40, 41, 43, 45, 52, 61 เป็นข้อสอบที่มีเนื้อหาสาระเกี่ยวข้องกับเพศหญิงมากกว่าเพศชาย จำนวน 1 ข้อ ได้แก่ ข้อที่ 71 เมื่อพิจารณาค่าความยากของข้อสอบ พบว่า เป็นข้อสอบยากง่ายปานกลาง จำนวน 8 ข้อ ได้แก่ ข้อที่ 3, 18, 28, 43, 45, 52, 61, 71 และเป็นข้อสอบง่าย จำนวน 3 ข้อ ได้แก่ ข้อที่ 1, 40, 41 ส่วนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอที่ตรวจไม่พบ (False Negative) จำนวน 7 ข้อ ได้แก่ 5, 20, 23, 38, 42, 57, 64

ในจำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอที่ตรวจพบ 11 ข้อ เป็นข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอที่สอดคล้องกับวิธีแมนเทิล-แฮนส์เซล แบบปกติ ทั้ง 6 ข้อ ได้แก่ ข้อที่ 1, 41, 45, 52, 61, 71 และเป็นข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอที่ตรวจพบเพิ่มขึ้น 5 ข้อ ได้แก่ ข้อที่ 3, 18, 28, 40, 43

4. การพิจารณาเปรียบเทียบผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอระหว่างวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ พบว่า วิธีแมนเทิล-แฮนส์เซล แบบปกติ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ ได้สอดคล้องกับวิธี IRT area 33.33 % (6 ข้อ ใน 18 ข้อ) ส่วนวิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ ได้สอดคล้องกับวิธี IRT area 61.11 % (11 ข้อ ใน 18 ข้อ) โดยเป็นข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ

ที่สอดคล้องกับวิธีแมนเทล-แฮนส์เซล แบบปกติ ทั้ง 6 ข้อ และเป็นข้อสอบทำหน้าที่ต่างกัน แบบไม่สม่ำเสมอที่ตรวจพบ เพิ่มขึ้นอีก จำนวน 5 ข้อ (27.77 %) ได้แก่ ข้อที่ 3, 18, 28, 40, 43 เมื่อพิจารณาลักษณะเนื้อหาสาระของข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่เพิ่มขึ้น พบว่า เป็นข้อสอบที่มีเนื้อหาสาระทั่วไปทั้ง 5 ข้อ เมื่อพิจารณาค่าความยากของข้อสอบ พบว่า เป็นข้อสอบยากง่ายปานกลาง 4 ข้อ ได้แก่ ข้อที่ 3, 18, 28, 43 และเป็นข้อสอบง่าย 1 ข้อ ได้แก่ ข้อที่ 40 และเมื่อพิจารณาจากกราฟแสดงโค้งลักษณะข้อสอบของกลุ่มผู้สอบ สองกลุ่ม พบว่า ข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่เพิ่มขึ้นทั้ง 5 ข้อ มีโค้ง ลักษณะข้อสอบของกลุ่มผู้สอบสองกลุ่มตัดกัน ในบริเวณช่วงความสามารถ -3 ถึง +3 ทั้งหมด โดยตัดกัน ในบริเวณใกล้ ๆ จุดกลางของช่วงความสามารถ จำนวน 3 ข้อ กล่าวคือ ตัดกัน ในบริเวณช่วงความสามารถ -1 ถึง +1 จำนวน 2 ข้อ ได้แก่ ข้อที่ 18, 28 และตัดกัน ในบริเวณช่วงความสามารถ -2 ถึง +2 จำนวน 1 ข้อ ได้แก่ ข้อที่ 43

ผู้วิจัยได้ทดสอบสัดส่วนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจพบ ระหว่างวิธี แมนเทล-แฮนส์เซล แบบปกติ กับ วิธีแมนเทล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ เมื่อใช้จำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ ตรวจพบ ตามวิธี IRT area เป็นเกณฑ์ โดยใช้ค่าสถิติ Z - Statistic พบว่า มีนัยสำคัญทาง สถิติที่ระดับ .05 แสดงว่า วิธีแมนเทล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและ ความยากของข้อสอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ได้จำนวนมากกว่า วิธีแมนเทล-แฮนส์เซล แบบปกติ ซึ่งเป็นไปตามสมมุติฐานการวิจัย

5. การเปรียบเทียบผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ระหว่างวิธี แมนเทล-แฮนส์เซล แบบปกติ กับ วิธีแมนเทล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ โดยการวิเคราะห์ข้อมูลจากการจำลองกลุ่มตัวอย่างจากข้อมูล จริง จำนวน 10 ครั้ง เพื่อตรวจสอบความคงเส้นคงวาของสัดส่วนการตรวจพบข้อสอบ ทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ระหว่างวิธีแมนเทล-แฮนส์เซล แบบปกติ กับ วิธีแมนเทล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ เมื่อใช้จำนวน ข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอที่ตรวจพบ ตามวิธี IRT area เป็นเกณฑ์ พบว่า วิธีแมนเทล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ตรวจพบจำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ได้จำนวนมากกว่า วิธีแมนเทล-แฮนส์เซล แบบปกติ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 6 ครั้ง และทั้งสอง

วิธีตรวจพบจำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ไม่แตกต่างกันที่ระดับนัยสำคัญทางสถิติ .05 จำนวน 4 ครั้ง

อภิปรายผลการวิจัย

จากผลการวิจัยครั้งนี้ มีประเด็นสำคัญที่ควรนำมาอภิปรายผล ดังต่อไปนี้

1. วิธีแมนเทิล-แฮนส์เซล แบบปกติ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ สอดคล้องกับวิธี IRT area แบบ 2 พารามิเตอร์ ในระดับปานกลาง (33.33 %) ซึ่งผลการศึกษาจากสถานการณ์จำลองของ Mazor, et al. (1994) พบว่า วิธีแมนเทิล-แฮนส์เซล แบบปกติ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ในระดับปานกลางค่อนข้างสูง (68 %) เหตุที่ผลการศึกษาแตกต่างกันไปบ้าง คงเนื่องมาจากการศึกษาของ Mazor, et al. (1994) เป็นการศึกษาจากสถานการณ์จำลอง ซึ่งจำลองข้อมูลผลการตอบข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบ ข้อมูลจำลองจึงสอดคล้องกับทฤษฎีการตอบสนองข้อสอบ ทำให้ผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอในสถานการณ์จำลอง สูงกว่าผลการศึกษาด้วยข้อมูลจริง

เมื่อพิจารณาค่าความยากของข้อสอบ พบว่า ข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ที่ตรวจพบทั้งหมด เป็นข้อสอบยากง่ายปานกลางและข้อสอบง่าย ส่วนข้อสอบที่ยาก ตรวจไม่พบการทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ซึ่งสอดคล้องกับงานวิจัยของ Clauser, et al, (1991b) เหตุที่เป็นเช่นนี้ก็เนื่องมาจากว่า ข้อสอบยากมีคนตอบถูกน้อย เมื่อแจกแจงจำนวนผู้สอบตามระดับความสามารถแล้ว สัดส่วนจำนวนผู้ตอบข้อสอบถูกในระดับความสามารถสูง ๆ มีจำนวนน้อย ทำให้ค่าสถิติ MH χ^2 ที่คำนวณได้มีค่าต่ำ โอกาสมีนัยสำคัญทางสถิติจึงน้อย เป็นผลให้ไม่สามารถตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ชนิดที่เป็นข้อสอบยาก

2. วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ สอดคล้องกับวิธี IRT area แบบ 2 พารามิเตอร์ ในระดับปานกลางค่อนข้างสูง (61.11 %) ซึ่งสอดคล้องกับผลการศึกษาจากสถานการณ์จำลองของ Mazor, et al. (1994) ที่พบว่า เมื่อวิเคราะห์ข้อมูลโดยแบ่งกลุ่มผู้สอบตามผลการสอบของผู้สอบ ออกเป็น กลุ่มผู้สอบที่มีผลการสอบสูง กับกลุ่มผู้สอบที่มีผลการสอบต่ำ ทำให้วิธีแมนเทิล-แฮนส์เซล ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ได้ในระดับปานกลางค่อนข้างสูง (82 %) สำหรับการวิจัยครั้งนี้ได้เพิ่มสารสนเทศเรื่องค่าความยากของข้อสอบเข้าไปในแบบแผนการวิเคราะห์ข้อมูลด้วย ทำให้วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่ม

ความสามารถของผู้สอบและความยากของข้อสอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่
สม่าเสมอได้เพิ่มขึ้น

อย่างไรก็ตาม เมื่อพิจารณาค่าความยากของข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอที่
ตรวจพบแล้ว พบว่า เป็นข้อสอบยากง่ายปานกลางและข้อสอบง่าย เช่นเดียวกับที่ตรวจพบตาม
วิธีแมนเทิล-แฮนส์เชล แบบปกติ ที่เป็นเช่นนี้ก็ด้วยเหตุผลเดียวกับข้อที่ 1

3. วิธีแมนเทิล-แฮนส์เชล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของ
ข้อสอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ เพิ่มขึ้นจาก วิธีแมนเทิล-แฮนส์เชล
แบบปกติ อีก 27.77 % ซึ่งสอดคล้องกับผลการศึกษาศึกษาจากสถานการณ์จำลองของ Mazor, et al.
(1994) ที่พบว่า เมื่อวิเคราะห์ข้อมูลโดยแบ่งกลุ่มผู้สอบตามผลการสอบของผู้สอบ ออกเป็น
กลุ่มผู้สอบที่มีผลการสอบสูงกับกลุ่มผู้สอบที่มีผลการสอบต่ำ ทำให้วิธีแมนเทิล-แฮนส์เชล
ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ ได้เพิ่มขึ้น 14 % ที่เป็นเช่นนี้ก็เพราะว่า
โอกาสในการตรวจพบข้อสอบทำหน้าที่ต่างกัน ตามวิธีแมนเทิล-แฮนส์เชล ขึ้นอยู่กับฟังก์ชัน
ค่าความสามารถของผู้สอบและความยากของข้อสอบ ดังนั้น การเพิ่มตัวแปรความสามารถของ
ผู้สอบ โดยการแบ่งกลุ่มผู้สอบตามค่าประมาณความสามารถของผู้สอบ ออกเป็นกลุ่มผู้สอบที่มี
ความสามารถสูง กับกลุ่มผู้สอบที่มีความสามารถต่ำ และการเพิ่มตัวแปรความยากของข้อสอบ
โดยการแบ่งข้อสอบตามค่าประมาณความยากของข้อสอบ เป็นกลุ่มข้อสอบยาก กลุ่มข้อสอบยาก
ง่ายปานกลาง และกลุ่มข้อสอบง่าย เข้าไปในแบบแผนการวิเคราะห์ข้อมูล ทำให้กลุ่มผู้สอบที่ถูก
แบ่งมีความเป็นเอกพันธ์มากขึ้น ความแปรปรวนของกลุ่มอ้างอิง ($Var R_m$) ซึ่งเป็นตัวหารใน
สูตร $MH-\chi^2$ ลดลง ส่งผลให้ค่าสถิติ $MH-\chi^2$ ที่คำนวณได้เพิ่มขึ้น จึงมีโอกาที่จะมีนัยสำคัญ
ทางสถิติมากขึ้น ทำให้สามารถตรวจพบจำนวนข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ ได้
เพิ่มขึ้น

เมื่อพิจารณาค่าความยากของข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ พบว่าข้อสอบที่
ตรวจพบเพิ่มขึ้นทั้งหมด เป็นข้อสอบยากง่ายปานกลางและข้อสอบง่าย ซึ่งสอดคล้องกับผลการ
ศึกษาศึกษาจากสถานการณ์จำลองของ Mazor, et al. (1994) และเป็นการยืนยันว่า วิธีแมนเทิล-
แฮนส์เชล ไม่ไวต่อการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ ชนิดที่เป็นข้อสอบยาก
ในสถานการณ์จริง แม้ว่าได้พัฒนาแบบแผนการวิเคราะห์ข้อมูล ตามวิธีแมนเทิล-แฮนส์เชล
แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ แล้วก็ตาม แสดงให้เห็นว่า
ค่าความยากของข้อสอบมีผลต่อการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ ดังนั้น
ทางแก้ในกรณีนี้ คงต้องแก้โดยการสุ่มจำนวนผู้สอบที่มีความสามารถสูงให้มากขึ้น เพื่อให้

สัดส่วนผู้ตอบข้อสอบถูกในระดับที่มีความสามารถสูง ๆ มีมากขึ้น อันอาจทำให้วิธีแมนเทิล-แฮนส์เซล สามารถตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ชนิดที่เป็นข้อสอบยาก ได้ดียิ่งขึ้น

เมื่อพิจารณาจากกราฟแสดงโค้งลักษณะข้อสอบของกลุ่มผู้สอบสองกลุ่ม พบว่าข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ที่ตรวจพบเพิ่มขึ้นทั้ง 5 ข้อ เป็นข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ชนิดที่โค้งลักษณะข้อสอบของกลุ่มสนใจกับกลุ่มอ้างอิง ตัดกันบริเวณใกล้ ๆ จุดกลางของช่วงความสามารถ จำนวน 3 ข้อ ซึ่งข้อสอบลักษณะนี้วิธีแมนเทิล-แฮนส์เซล แบบปกติ ตรวจไม่ค่อยพบ (Swaminathan and Rogers, 1990) เหตุที่วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ตรวจพบข้อสอบในลักษณะนี้เพิ่มขึ้น ก็เนื่องจากว่าสถิติ $MH-\chi^2$ เป็นสถิติที่มีเครื่องหมาย ดังนั้น เมื่อแบ่งกลุ่มผู้สอบออกเป็น 2 กลุ่ม ตามค่าความสามารถของผู้สอบและแบ่งข้อสอบตามค่าความยากของข้อสอบ ทำให้จุดที่แบ่งความสามารถของผู้สอบอยู่ใกล้เดียวกับจุดกลางของช่วงความสามารถของผู้สอบ ซึ่งเป็นบริเวณปฏิสัมพันธ์ของโค้งลักษณะข้อสอบระหว่างกลุ่มสนใจกับกลุ่มอ้างอิง ทำให้ความแตกต่างทางลบ (negative difference) กับ ความแตกต่างทางบวก (positive difference) ของอัตราส่วนเต็มต่อร่วม (common odds ratio) ในแต่ละระดับความสามารถ มีโอกาสหักล้างกันน้อยลง ทำให้ค่าสถิติ $MH-\chi^2$ ที่คำนวณได้สูงขึ้น โอกาสที่จะมีนัยสำคัญทางสถิติจึงมากขึ้น ส่งผลให้วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ สามารถตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ในลักษณะนี้ได้เพิ่มขึ้น

4. การตรวจสอบความคงเส้นคงวาของสัดส่วนการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ระหว่างวิธีแมนเทิล-แฮนส์เซล แบบปกติ กับ แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ โดยการวิเคราะห์ข้อมูลจากการจำลองกลุ่มตัวอย่างจากข้อมูลจริง จำนวน 10 ครั้ง พบว่าวิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ได้จำนวนมากกว่าวิธีแมนเทิล-แฮนส์เซล แบบปกติ อย่างมีนัยสำคัญ จำนวน 6 ครั้ง จึงเป็นการยืนยันข้อค้นพบที่ได้จากการวิเคราะห์ข้อมูลจากกลุ่มตัวอย่างว่า วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ได้ไวกว่า วิธีแมนเทิล-แฮนส์เซล แบบปกติ ส่วนเหตุที่ผลการศึกษาจากการจำลองกลุ่มตัวอย่างจากข้อมูลจริง วิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของ

ข้อสอบ ไม่สามารถตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอได้มากกว่าวิธีแมนเทล-แฮนส์เซิล แบบปกติ อย่างมีนัยสำคัญทางสถิติ ทั้ง 10 ครั้ง คงเนื่องจากว่าวิธีแมนเทล-แฮนส์เซิล ใช้คะแนนที่สังเกตได้ ตามทฤษฎีการสอบแบบดั้งเดิม เป็นเกณฑ์การจับคู่กลุ่มผู้สอบย่อยสองกลุ่ม จึงทำให้ค่าสถิติไค-สแควร์ที่คำนวณได้เปลี่ยนไปตามกลุ่มตัวอย่าง เมื่อกลุ่มตัวอย่างเปลี่ยนแปลงไป ผลการตรวจพบข้อสอบทำหน้าที่ต่างกัน ก็อาจเปลี่ยนแปลงไป ซึ่งสอดคล้องกับผลการศึกษาของ Hambleton และ Jones (1993) ที่ว่า ผลการตรวจพบข้อสอบทำหน้าที่ต่างกัน ของวิธีแมนเทล-แฮนส์เซิล ไม่ค่อยคงเส้นคงวานัก อย่างไรก็ตาม แม้ว่าการจำลองกลุ่มตัวอย่างที่ใช้ศึกษามีขนาดเล็กกว่าการศึกษาเกี่ยวกับกลุ่มตัวอย่าง วิธีแมนเทล-แฮนส์เซิล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ก็ยังตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ได้ไวกว่า วิธีแมนเทล-แฮนส์เซิล แบบปกติ ดังนั้น ในกรณีที่ศึกษากับกลุ่มตัวอย่างขนาดใหญ่กว่างานวิจัยครั้งนี้ ผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ของวิธีแมนเทล-แฮนส์เซิล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ย่อมต้องไวกว่า วิธีแมนเทล-แฮนส์เซิล แบบปกติ เนื่องจากอำนาจของค่าสถิติ $MH-\chi^2$ เพิ่มขึ้นตามขนาดของกลุ่มตัวอย่าง (Mazor, et al., 1992 ; Hambleton, et al., 1993)

5. การที่ตรวจพบข้อสอบข้อหนึ่งข้อใด มีดัชนี (index) แสดงการทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ แสดงว่า เรายังสรุปแน่นอนไม่ได้ว่า ข้อสอบข้อนั้นทำหน้าที่ต่างกันโดยเข้าข้างกลุ่มผู้สอบเพศหญิง หรือเข้าข้างกลุ่มผู้สอบเพศชาย แต่ต้องพิจารณาข้อสอบข้อนั้นตามระดับความสามารถของผู้สอบว่า ในระดับความสามารถใดกลุ่มผู้สอบเพศหญิงเป็นฝ่ายได้เปรียบ หรือในระดับความสามารถใดกลุ่มผู้สอบเพศชายเป็นฝ่ายได้เปรียบ ดังนั้น แม้มีผลการศึกษา (สุพรรณ สุกมลสันต์, 2534 ; กาญจนา วจินสุนทร, 2538) เสนอแนะว่า กลุ่มผู้สอบเพศหญิงมีแนวโน้มได้เปรียบในข้อสอบทางด้านภาษา แต่เมื่อพิจารณาจากผลการวิจัยครั้งนี้แล้ว มีข้อสังเกตว่า ข้อสอบทางด้านภาษานั้น กลุ่มผู้สอบเพศหญิงไม่ได้เป็นฝ่ายได้เปรียบเสมอไป เนื่องจากในบางระดับความสามารถกลุ่มผู้สอบเพศชายอาจเป็นฝ่ายได้เปรียบ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. เมื่อพิจารณาภายใต้เงื่อนไขตัวแปรเพศ วิธีแมนเทล-แฮนส์เซิล แบบปกติ สามารถตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ได้สอดคล้องกับวิธี IRT area 33.33 % (6 ข้อ ใน 18 ข้อ) และวิธีแมนเทล-แฮนส์เซิล แบบแบ่งกลุ่มความสามารถของผู้สอบและความ

ยากของข้อสอบ สามารถตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่ำเสมอ ได้สอดคล้องกับวิธี IRT area 61.11 % (11 ข้อ ใน 18 ข้อ) ถือว่าเป็นระดับประสิทธิผลที่ใช้ได้ เนื่องจากในสภาพความเป็นจริงแล้ว ผู้ปฏิบัติส่วนใหญ่ไม่ใช้วิธี IRT area ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ คงเนื่องมาจากว่ากระบวนการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบสลับซับซ้อน เสียค่าใช้จ่ายในการคำนวณสูงและต้องใช้กับกลุ่มตัวอย่างขนาดใหญ่ แต่นิยมใช้วิธีแมนเทิล-แฮนส์เชล ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในแบบสอบวิชาต่าง ๆ เพราะมีความสะดวกในทางปฏิบัติ เสียค่าใช้จ่ายในการคำนวณไม่แพงและสามารถใช้ได้กับกลุ่มตัวอย่างขนาดไม่ใหญ่นัก

ดังนั้น ผู้ปฏิบัติที่น่าวิธีแมนเทิล-แฮนส์เชลไปใช้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในโครงการทดสอบสำคัญ ๆ ซึ่งต้องการผลการตรวจสอบที่มีความถูกต้องสูงขึ้น ก็ควรเลือกใช้วิธีแมนเทิล-แฮนส์เชล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ แต่ทั้งนี้ต้องระวังเรื่อง การระบุผิดพลาดว่าข้อสอบทำหน้าที่ต่างกัน ทั้ง ๆ ที่ ความจริงแล้วเป็นข้อสอบทำหน้าที่ไม่ต่างกัน ด้วย

ส่วนประเด็นเรื่องการแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ซึ่งต้องใช้ค่าประมาณความสามารถของผู้สอบและความยากของข้อสอบ ปัจจุบันมีโปรแกรมคอมพิวเตอร์ BICAL จำหน่ายโดยทั่วไป โปรแกรมคอมพิวเตอร์นี้สามารถหาค่าประมาณความสามารถของผู้สอบ และค่าความยากของข้อสอบ ตามทฤษฎีการตอบสนองข้อสอบ แบบ หนึ่ง พารามิเตอร์ หรือที่เรียกว่า Rasch model ดังนั้น ผู้ปฏิบัติอาจนำค่าความสามารถของผู้สอบและค่าความยากของข้อสอบที่คำนวณได้ไปใช้ กับวิธีแมนเทิล-แฮนส์เชล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ประกอบกับการวิเคราะห์ข้อสอบตาม Rasch model ใช้กลุ่มตัวอย่างประมาณ 200 คนต่อกลุ่ม เท่านั้น

2. ผู้ปฏิบัติควรเลือกใช้วิธีแมนเทิล-แฮนส์เชล ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในกรณีที่มีกลุ่มตัวอย่างขนาดไม่ใหญ่นัก เป็นต้นว่า ประมาณ 200 คนต่อกลุ่ม หรือในกรณีที่ต้องการนำผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบไปใช้เพื่อพัฒนาข้อสอบเท่านั้น กลุ่มตัวอย่างขนาด 100 คนต่อกลุ่ม ก็ใช้ได้ (Hills, 1989) ซึ่งในกรณีอย่างนี้สามารถใช้วิธีแมนเทิล-แฮนส์เชล เป็นทางเลือกแทนวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบได้ เนื่องจากวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบมีข้อจำกัด ในเรื่องขนาดกลุ่มตัวอย่าง เช่น วิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ อาจต้องใช้กลุ่มตัวอย่างขนาด 400 คนต่อกลุ่มขึ้นไป หรือแบบ 3 พารามิเตอร์ อาจต้องใช้กลุ่มตัวอย่างขนาด 1,000 คนต่อกลุ่มขึ้นไป เป็นต้น

อีกทั้งวิธีแมนเทิล-เฮนส์เซล มีกระบวนการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบไม่ยุ่งยาก ซับซ้อน และเสียค่าใช้จ่ายในการดำเนินการต่ำกว่าวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ

3. ผู้ปฏิบัติควรคำนึงว่า การทำหน้าที่ต่างกันของข้อสอบ เป็นเพียงดัชนี (index) ตัวหนึ่ง ในการประเมินความสำคัญของข้อสอบ เท่านั้น ดังนั้น การตีความผลการวิเคราะห์ทางสถิติ จึงควรกระทำด้วยความระมัดระวัง การที่ข้อสอบข้อใดข้อหนึ่งทำหน้าที่ต่างกัน ไม่ได้หมายความว่า ข้อสอบข้อนั้นไม่ดีเสมอไป การนำดัชนีตัวนี้ไปใช้คัดเลือกข้อสอบออกจากแบบสอบหรือคลังข้อสอบ คงต้องพิจารณาวัตถุประสงค์ในการวัดของแบบสอบและเนื้อหาสาระของข้อสอบ เป็นต้นว่า แบบสอบฉบับนั้นมีจุดมุ่งหมายในการวัดอะไร ถ้าแบบสอบมีจุดมุ่งหมายในการวัดความรู้ทางด้านอาหาร หรือคัดเลือกช่างทำอาหารหรือช่างทำขนม ข้อสอบที่มีเนื้อหาสาระเกี่ยวกับการทำอาหารหรือการทำขนม ก็อาจไม่ใช่ข้อสอบที่สำคัญเข้าข้างเพศหญิง แต่เป็นข้อสอบที่มีความเหมาะสมกับผู้สอบกลุ่มย่อยทุกกลุ่ม

ข้อเสนอแนะในการวิจัย

จากประเด็นต่าง ๆ ที่พบในการวิจัยครั้งนี้ ผู้วิจัยเสนอหัวข้อที่น่าจะได้มีการวิจัย ดังต่อไปนี้

1. ศึกษาการทำหน้าที่ต่างกันของข้อสอบ โดยใช้ตัวแปรจำแนกกลุ่มของประชากรอื่น ๆ เช่น เชื้อชาติ ศาสนา สถานภาพทางเศรษฐกิจและสังคม เป็นต้น เพื่อเปรียบเทียบกับผลการวิจัยครั้งนี้

2. เนื่องจากผลการวิจัยครั้งนี้ วิธีแมนเทิล-เฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ ยังไม่สามารถตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ ชนิดที่เป็นข้อสอบยาก ดังนั้น จึงควรมีการศึกษาวิจัยในประเด็นผลของค่าความยากของข้อสอบ ต่อการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ โดยให้มีข้อสอบยาก ข้อสอบยากง่าย ปานกลาง และข้อสอบง่าย ในสัดส่วนใกล้เคียงกัน

3. ควรมีการศึกษาผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอ ด้วยวิธีแมนเทิล-เฮนส์เซล กับกลุ่มตัวอย่างขนาดต่าง ๆ และจำนวนข้อสอบในแต่ละกลุ่มขนาดต่าง ๆ เพื่อเปรียบเทียบผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบไม่สม่าเสมอเพิ่มเติมต่อไป