

บทที่ 5

สรุปและอภิปรายผล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาประสิทธิภาพของโมเดลในการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายใน ที่พัฒนาตามทฤษฎีเซาว์ปัญญาของสเปียร์แมน และปัจจัยที่มีอิทธิพลต่อลักษณะการแจกแจงค่าดัชนีวัดความสอดคล้องซึ่งได้แก่ จำนวนกลุ่มตัวอย่าง จำนวนตัวดัชนีชี้วัดต่อองค์ประกอบ ให้ได้รูปแบบที่สอดคล้องกับทฤษฎีและหาเกณฑ์ที่เหมาะสมในการตัดสินใจสำหรับการตรวจสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ในการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายใน โดยการเปรียบเทียบกับโมเดลที่นิยมใช้อยู่ในปัจจุบัน

กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 2 ปีการศึกษา 2548 ที่เข้าสอบวัดผลสัมฤทธิ์ทางการเรียนเพื่อประเมินคุณภาพการศึกษาระดับชาติ จำนวน 114 โรงเรียน มีนักเรียนที่เข้าสอบวิชาภาษาไทย วิชาคณิตศาสตร์ วิชาสังคม ศาสนาและวัฒนธรรม วิทยาศาสตร์ และภาษาอังกฤษ จำนวน 5,011 คน

การวิเคราะห์ข้อมูลและแปลความหมายผลการวิเคราะห์ข้อมูล ในการศึกษาวิเคราะห์ประสิทธิภาพของโมเดลและการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบพหุมิติโดยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์การวิจัยไว้ 3 ตอนดังนี้

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลสถิติพื้นฐาน

ตอนที่ 2 การตรวจสอบประสิทธิภาพของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยัน

ตอนที่ 3 การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายใน

สรุปผลการวิจัย

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐานของตัวแปรสังเกตได้

1. ค่าสถิติพื้นฐานของตัวแปรที่นำมาใช้เป็นกลุ่มตัวอย่าง เป็นการวิเคราะห์ข้อมูลจากกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงานคณะกรรมการเขตพื้นที่การศึกษามัธยมศึกษา เขต 2 จำนวน 2,000 คน โดยแบ่งเป็นกลุ่มตัวอย่างขนาด 100, 200, 400, 800, 1,200, 1,600, และ 2,000 คน โดยมีจุดมุ่งหมายเพื่อหาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด ค่าสูงสุด ค่าความโค้ง ค่าความเบ้ ของตัวแปรสังเกตได้ที่ใช้ในการตรวจสอบประสิทธิภาพ

ของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยัน ทั้งนี้เพื่อศึกษาลักษณะการแจกแจงของตัวแปร โดยผู้วิจัยนำเสนอในภาพรวม จำแนกตามขนาดของกลุ่มตัวอย่าง พบว่าตัวแปรทุกตัวมีการแจกแจงแบบปกติ ทั้งนี้เพราะค่าความเบ้ ความโด่งมีค่าเข้าใกล้ศูนย์ตามข้อเสนอแนะของ ไคลน์ (Kline, 2005) ได้กล่าวไว้ในกรณีการแจกแจงตัวแปรเป็นแบบปกติเมื่อค่าความเบ้ต่ำกว่า 3 และค่าความโด่งมีค่าต่ำกว่า 10 ซึ่งยอมรับได้ว่าข้อมูลมีการแจกแจงปกติตามรายตัวแปร ดังนั้นผู้วิจัยจึงใช้ข้อมูลดังกล่าวนี้ในการวิเคราะห์เพื่อตรวจสอบสมมติฐาน โดยไม่มีการเปลี่ยนแปลงข้อมูลแต่อย่างใด

2. การวิเคราะห์คุณภาพของข้อสอบแสดงค่าพารามิเตอร์อำนาจจำแนกและค่าความยากของข้อสอบการวิเคราะห์คุณภาพของแบบสอบที่นำมาใช้เป็นกลุ่มตัวอย่างในการหาประสิทธิภาพของโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันและการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบสำหรับแบบสอบวิชาภาษาไทย วิชาคณิตศาสตร์ วิชาวิทยาศาสตร์ วิชาสังคม ศาสนาและวัฒนธรรม และวิชาภาษาอังกฤษ โดยทำการวิเคราะห์ตามทฤษฎีแนวใหม่ (Item Response Theory) โดยใช้โปรแกรม BILOG 3.0 และเลือกโมเดลแบบ 2 พารามิเตอร์ซึ่งถือว่ามีเหมาะสมที่สุด เลือกข้อสอบตัวอย่างสำหรับนำมาใช้ในการวิเคราะห์ที่มีค่าอำนาจจำแนกตั้งแต่ 0.5 ขึ้นไป ซึ่งถือว่าเป็นข้อสอบที่สามารถจำแนกกลุ่มผู้สอบ ได้ดีสำหรับนำมาใช้เป็นกลุ่มตัวอย่างประกอบด้วยวิชาภาษาไทย 15 ข้อ วิชาคณิตศาสตร์ 15 ข้อ วิชาวิทยาศาสตร์ 15 ข้อ วิชาสังคม ศาสนาและวัฒนธรรม 15 ข้อ และวิชาภาษาอังกฤษ 15 ข้อ รวมทั้งหมด 75 ข้อ โดยมีรายละเอียดดังนี้

2.1 การวิเคราะห์คุณภาพของตัวแปรข้อสอบแสดงค่าพารามิเตอร์อำนาจจำแนกและค่าความยากของข้อสอบ สำหรับวิชาภาษาไทย พบว่า ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 0.770 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.148 โดยมีลักษณะเบ้ขวา ค่าความยากเฉลี่ยมีค่าเท่ากับ 0.689 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 16.7 ค่าไค-สแควร์เฉลี่ยมีค่าเท่ากับ 22.3 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 16.7 มีลักษณะเบ้ขวา

2.2 การวิเคราะห์คุณภาพของตัวแปรข้อสอบแสดงค่าพารามิเตอร์อำนาจจำแนกและค่าความยากของข้อสอบ สำหรับวิชาคณิตศาสตร์พบว่า ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.634 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.461 โดยมีลักษณะเบ้ขวา ค่าความยากเฉลี่ยมีค่าเท่ากับ 2.336 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.531 มีลักษณะเบ้ซ้าย ค่าไค-สแควร์เฉลี่ยมีค่าเท่ากับ 68.253 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 88.21 มีลักษณะเบ้ขวา

2.3 การวิเคราะห์คุณภาพของตัวแปรข้อสอบแสดงค่าพารามิเตอร์อำนาจจำแนกและค่าความยากของข้อสอบ สำหรับวิชาวิทยาศาสตร์ พบว่า ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 0.969 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.412 โดยมีลักษณะเบ้ขวา ค่าความยากเฉลี่ยมีค่าเท่ากับ 2.054 ค่า

เบี่ยงเบนมาตรฐานเท่ากับ 1.090 มีลักษณะเบ้ซ้าย ค่าโค-สแควร์เฉลี่ยมีค่าเท่ากับ 21.2 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 26.3 มีลักษณะเบ้ขวา

2.4 การวิเคราะห์คุณภาพของตัวแปรข้อสอบแสดงค่าพารามิเตอร์อำนาจจำแนกและค่าความยากของข้อสอบสำหรับวิชาสังคม ศาสนาและวัฒนธรรมพบว่า ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 0.934 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.278 โดยมีลักษณะเบ้ขวา ค่าความยากเฉลี่ยมีค่าเท่ากับ 1.478 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.637 มีลักษณะเบ้ขวา ค่าโค-สแควร์เฉลี่ยมีค่าเท่ากับ 14.1 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 12.3 มีลักษณะเบ้ขวา

2.5 การวิเคราะห์คุณภาพของตัวแปรข้อสอบแสดงค่าพารามิเตอร์อำนาจจำแนกและค่าความยากของข้อสอบสำหรับวิชาภาษาอังกฤษ พบว่า ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.999 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.996 โดยมีลักษณะเบ้ขวา ค่าความยากเฉลี่ยมีค่าเท่ากับ 2.161 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.554 มีลักษณะเบ้ซ้าย ค่าโค-สแควร์เฉลี่ยมีค่าเท่ากับ 231.6 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 230.2 มีลักษณะเบ้ขวา

3. การตรวจสอบคุณภาพของแบบสอบในด้านความเที่ยงตรงเชิงโครงสร้างโดยการวิเคราะห์องค์ประกอบเชิงยืนยันชั้นความมารถทางการเรียน

การตรวจสอบความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) ทำการวิเคราะห์และตรวจสอบความตรงเชิงโครงสร้างขององค์ประกอบโดยแบ่งข้อสอบเป็นกลุ่มย่อย 5 องค์ประกอบ องค์ประกอบละ 10 ข้อ จำนวน 50 ข้อสำหรับนำมาใช้ในการหาประสิทธิภาพของโมเดลประกอบด้วยข้อสอบวัดความสามารถด้านภาษาไทย ความสามารถด้านคณิตศาสตร์ ความสามารถด้านวิทยาศาสตร์ ความสามารถด้านสังคม ศาสนาและวัฒนธรรม และความสามารถด้านภาษาอังกฤษ

3.1 โมเดลองค์ประกอบวัดความสามารถด้านภาษาไทยซึ่งประกอบด้วยข้อสอบจำนวน 10 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันชั้นของ โมเดลวัดความสามารถด้านภาษาไทยประกอบด้วยตัวแปรความสามารถด้านภาษาไทยข้อ 8 ความสามารถด้านภาษาไทยข้อ 9 ความสามารถด้านภาษาไทยข้อ 10 ความสามารถด้านภาษาไทยข้อ 16 ความสามารถด้านภาษาไทยข้อ 27 ความสามารถด้านภาษาไทยข้อ 29 ความสามารถด้านภาษาไทยข้อ 33 ความสามารถด้านภาษาไทยข้อ 35 ความสามารถด้านภาษาไทยข้อ 36 และ ความสามารถด้านภาษาไทยข้อ 39 โดยมีค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .027 - .424 โดยข้อสอบวัดความสามารถด้านภาษาไทยที่มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ข้อสอบวัดความสามารถด้านภาษาไทยข้อที่ 27 มีค่าเท่ากับ .424 สำหรับข้อสอบวัดความสามารถด้านภาษาไทยที่มีน้ำหนักองค์ประกอบน้อยที่สุดได้แก่ข้อสอบวัดความสามารถด้านภาษาไทยข้อที่ 8 มีค่าเท่ากับ .027 และมีค่าสัมประสิทธิ์การ

พยากรณ์ (R^2) อยู่ระหว่าง .001- .180 โดยสัมประสิทธิ์การพยากรณ์ที่มีค่ามากที่สุดได้แก่ ข้อสอบวัดความสามารถด้านภาษาไทยข้อที่ 39 เท่ากับ .180 และค่าสัมประสิทธิ์การพยากรณ์ที่มีค่าน้อยที่สุดคือข้อสอบวัดความสามารถด้านภาษาไทยข้อที่ 8 ค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .001 โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี ซึ่งพิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 26.264 ค่าความน่าจะเป็น (p) เท่ากับ .611 ท้องศาอิสระ (df) เท่ากับ 29 และมีค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .027 - .424 และมีค่า R^2 อยู่ระหว่าง .001 - 1.80

3.2 โมเดลองค์ประกอบวัดความสามารถด้านคณิตศาสตร์ซึ่งประกอบด้วยข้อสอบจำนวน 10 ข้อผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลวัดความสามารถด้านคณิตศาสตร์ประกอบด้วยตัวแปร ความสามารถด้านคณิตศาสตร์ข้อ 4 ความสามารถด้านคณิตศาสตร์ข้อ 7 ความสามารถด้านคณิตศาสตร์ข้อ 13 ความสามารถด้านคณิตศาสตร์ข้อ 15 ความสามารถด้านคณิตศาสตร์ข้อ 19 ความสามารถด้านคณิตศาสตร์ข้อ 23 ความสามารถด้านคณิตศาสตร์ข้อ 28 ความสามารถด้านคณิตศาสตร์ข้อ 30 ความสามารถด้านคณิตศาสตร์ข้อ 31 และ ความสามารถด้านคณิตศาสตร์ข้อ 38 โดยมีค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .007 - .269 โดยข้อสอบวัดความสามารถด้านคณิตศาสตร์ที่มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ข้อสอบวัดความสามารถด้านคณิตศาสตร์ข้อที่ 15 มีค่าเท่ากับ -.269 สำหรับข้อสอบวัดความสามารถด้านคณิตศาสตร์ที่มีน้ำหนักองค์ประกอบน้อยที่สุดได้แก่ข้อสอบวัดความสามารถด้านคณิตศาสตร์ข้อที่ 4 มีค่าเท่ากับ .007 และมีค่าสัมประสิทธิ์การพยากรณ์ (R^2) อยู่ระหว่าง .000- .072 โดยสัมประสิทธิ์การพยากรณ์ที่มีค่ามากที่สุดได้แก่ ข้อสอบวัดความสามารถด้านคณิตศาสตร์ข้อที่ 15 เท่ากับ .017 และค่าสัมประสิทธิ์การพยากรณ์ที่มีค่าน้อยที่สุดคือข้อสอบวัดความสามารถด้านคณิตศาสตร์ข้อที่ 4 ค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .000 โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีซึ่งพิจารณาจากค่าไค-สแควร์มีค่าเท่ากับ 28.003 ค่าความน่าจะเป็นเท่ากับ .621 ท้องศาอิสระเท่ากับ 31 ซึ่งมีค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .007 - .269 และมีค่า R^2 อยู่ระหว่าง .000 - .072

3.3 โมเดลองค์ประกอบวัดความสามารถด้านวิทยาศาสตร์ซึ่งประกอบด้วยข้อสอบจำนวน 10 ข้อผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลวัดความสามารถด้านวิทยาศาสตร์ประกอบด้วยตัวแปร ความสามารถด้านวิทยาศาสตร์ข้อ 2 ความสามารถด้านวิทยาศาสตร์ข้อ 6 ความสามารถด้านวิทยาศาสตร์ข้อ 12 ความสามารถด้านวิทยาศาสตร์ข้อ 15 ความสามารถด้านวิทยาศาสตร์ข้อ 21 ความสามารถด้านวิทยาศาสตร์ข้อ 24 ความสามารถด้านวิทยาศาสตร์ข้อ 26 ความสามารถด้านวิทยาศาสตร์ข้อ 29 ความสามารถด้านวิทยาศาสตร์ข้อ 31 และ ความสามารถด้านวิทยาศาสตร์ข้อ 40 โดยมีค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .070 - .446 โดยข้อสอบวัดความสามารถด้านวิทยาศาสตร์ที่มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ข้อสอบวัด

ความสามารถด้านวิทยาศาสตร์ข้อที่ 40 มีค่าเท่ากับ .446 สำหรับข้อสอบวัดความสามารถด้านวิทยาศาสตร์ที่มีน้ำหนักองค์ประกอบน้อยที่สุด ได้แก่ข้อสอบวัดความสามารถด้านวิทยาศาสตร์ข้อที่ 24 มีค่าเท่ากับ -.070 และมีค่าสัมประสิทธิ์การพยากรณ์ (R^2) อยู่ระหว่าง .000 - .199 โดยสัมประสิทธิ์การพยากรณ์ที่มีค่ามากที่สุด ได้แก่ ข้อสอบวัดความสามารถด้านวิทยาศาสตร์ข้อที่ 24 เท่ากับ .199 และค่าสัมประสิทธิ์การพยากรณ์ที่มีค่าน้อยที่สุดคือข้อสอบวัดความสามารถด้านวิทยาศาสตร์ข้อที่ 29 ค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .000 โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดีซึ่งมีค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .070 - .446 และมีค่า R^2 อยู่ระหว่าง .000 - .199 และพบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี ซึ่งพิจารณาจากค่าไค-สแควร์มีค่าเท่ากับ 34.645 ค่าความน่าจะเป็นเท่ากับ .216 ที่องศาอิสระเท่ากับ 29

3.4 โมเดลองค์ประกอบวัดความสามารถด้านสังคมศึกษา ศาสนา และวัฒนธรรมซึ่งประกอบด้วยข้อสอบจำนวน 10 ข้อ ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลวัดความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรม องค์ประกอบความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรม มีจำนวน 10 ข้อประกอบด้วยตัวแปร ความสามารถด้านสังคมศึกษาข้อ 2 ความสามารถด้านสังคมศึกษาข้อ 4 ความสามารถด้านสังคมศึกษาข้อ 5 ความสามารถด้านสังคมศึกษาข้อ 9 ความสามารถด้านสังคมศึกษาข้อ 11 ความสามารถด้านสังคมศึกษาข้อ 13 ความสามารถด้านสังคมศึกษาข้อ 22 ความสามารถด้านข้อ 25 ความสามารถด้านสังคมศึกษาข้อ 27 และ ความสามารถด้านสังคมศึกษาข้อ 28 ซึ่งมีค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .087 - .407 โดยข้อสอบวัดความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรม ที่มีน้ำหนักองค์ประกอบมากที่สุด ได้แก่ข้อสอบวัดความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรม ข้อที่ 9 มีค่าเท่ากับ .407 สำหรับข้อสอบวัดความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรม ที่มีน้ำหนักองค์ประกอบน้อยที่สุด ได้แก่ข้อสอบวัดความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรมข้อที่ 11 มีค่าเท่ากับ .087 และมีค่าสัมประสิทธิ์การพยากรณ์ (R^2) อยู่ระหว่าง .000 - .166 โดยสัมประสิทธิ์การพยากรณ์ที่มีค่ามากที่สุด ได้แก่ ข้อสอบวัดความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรมข้อที่ 9 เท่ากับ .166 และค่าสัมประสิทธิ์การพยากรณ์ที่มีค่าน้อยที่สุดคือข้อสอบวัดความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรมข้อที่ 11 ค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .008 โมเดลการวัดความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรม พบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี ซึ่งพิจารณาจากค่าไค-สแควร์มีค่าเท่ากับ 20.329 ค่าความน่าจะเป็นเท่ากับ .959 ที่องศาอิสระเท่ากับ 33

3.5 โมเดลองค์ประกอบวัดความสามารถด้านภาษาอังกฤษซึ่งประกอบด้วยข้อสอบจำนวน 10 ข้อ ผลการวิเคราะห์องค์ประกอบเชิงยืนยันชั้นของ โมเดลวัดความสามารถด้านภาษาอังกฤษประกอบด้วยตัวแปร ความสามารถด้านภาษาอังกฤษข้อ 8 ความสามารถด้านภาษาอังกฤษข้อ 10 ความสามารถด้านภาษาอังกฤษข้อ 12 ความสามารถด้านภาษาอังกฤษข้อ 15 ความสามารถด้านภาษาอังกฤษข้อ 21 ความสามารถด้านภาษาอังกฤษข้อ 25 ความสามารถด้านภาษาอังกฤษข้อ 28 ความสามารถด้านภาษาอังกฤษข้อ 29 ความสามารถด้านภาษาอังกฤษข้อ 30 และ ความสามารถด้านภาษาอังกฤษข้อ 31 โดยมีค่าน้ำหนักองค์ประกอบอยู่ระหว่าง .039 - .405 โดยข้อสอบวัดความสามารถด้านภาษาอังกฤษที่มีน้ำหนักองค์ประกอบมากที่สุด ได้แก่ข้อสอบวัดความสามารถด้านภาษาอังกฤษข้อที่ 21 มีค่าเท่ากับ .405 สำหรับข้อสอบวัดความสามารถด้านภาษาอังกฤษที่มีน้ำหนักองค์ประกอบน้อยที่สุด ได้แก่ข้อสอบวัดความสามารถด้านภาษาอังกฤษข้อที่ 12 มีค่าเท่ากับ .039 และมีค่าสัมประสิทธิ์การพยากรณ์ (R^2) อยู่ระหว่าง .058 - .164 โดยสัมประสิทธิ์การพยากรณ์ที่มีค่ามากที่สุด ได้แก่ ข้อสอบวัดความสามารถด้านภาษาอังกฤษข้อที่ 21 เท่ากับ .164 และค่าสัมประสิทธิ์การพยากรณ์ที่มีค่าน้อยที่สุดคือข้อสอบวัดความสามารถด้านภาษาอังกฤษข้อที่ 8 ค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .058 โมเดลการวัดความสามารถด้านภาษาอังกฤษพบว่า โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี ซึ่งพิจารณาจากค่าไค-สแควร์มีค่าเท่ากับ 33.370 ค่าความน่าจะเป็นเท่ากับ .401 ที่องศาอิสระเท่ากับ 32

ตอนที่ 2 การตรวจสอบประสิทธิภาพของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยัน

1. การประสิทธิภาพของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง อันดับสอง และโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง อันดับสองแฝงภายใน เพื่อหาโมเดลวัดที่เหมาะสมกับข้อมูลเชิงประจักษ์ การวิเคราะห์ข้อมูลเพื่อหาประสิทธิภาพของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันเป็นการศึกษาประสิทธิภาพของโมเดลการวิเคราะห์องค์ประกอบ ความทฤษฎีสององค์ประกอบของชาร์ล สเปียร์แมน โดยการเปรียบเทียบค่าของดัชนีวัดความสอดคล้อง (Fit Indexes) ใน 4 โมเดล คือ

โมเดลที่ 1 การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง

โมเดลที่ 2 การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง

โมเดลที่ 3 การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน (Nested Factors)

โมเดลที่ 4 การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแบบแฝงภายใน

(Seconds Order Nested Factors)

1.1 ผลการวิเคราะห์องค์ประกอบความสามารถทางการเรียน ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 เขตพื้นที่การศึกษาเขต 2 จังหวัดบุรีรัมย์ โดยใช้โมเดลการวิเคราะห์องค์ประกอบ

เชิงยืนยันอันดับหนึ่ง มีรายละเอียดดังนี้แบบสอบวัดความสามารถด้านภาษาไทยมีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางภาษาไทยข้อที่ 24 มีค่าเท่ากับ .138 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .016 แบบสอบวัดความสามารถด้านคณิตศาสตร์มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางคณิตศาสตร์ ข้อที่ 15 มีค่าเท่ากับ .113 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .013 แบบสอบวัดความสามารถด้านวิทยาศาสตร์มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางวิทยาศาสตร์ข้อที่ 24 มีค่าเท่ากับ .129 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .017 แบบสอบวัดความสามารถด้านสังคมศึกษา ศาสนา และวัฒนธรรม มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางสังคมศึกษา ศาสนาและวัฒนธรรมข้อที่ 5 มีค่าเท่ากับ .286 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .082 แบบสอบวัดความสามารถด้านภาษาอังกฤษ มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางภาษาอังกฤษข้อที่ 12 มีค่าเท่ากับ .120 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .014

การวิเคราะห์องค์ประกอบเชิงยืนยัน โมเดลการวัดความสามารถด้านทางการเรียนพบว่า โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์น้อย ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 2428.632 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .850 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .792 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .716 ดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเท่ากับ .012 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .037 ที่องศาอิสระเท่ากับ 1170

1.2 ประสิทธิภาพของโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง องค์ประกอบผลสัมฤทธิ์ทางการเรียนของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 เขตพื้นที่การศึกษาเขต 2 จังหวัดบุรีรัมย์ พบว่าน้ำหนักองค์ประกอบแบบทดสอบความสามารถทางเรียนทั้ง 5 องค์ประกอบ มีค่าอยู่ระหว่าง .044 - .395 โดยน้ำหนักองค์ประกอบเรียงลำดับจากมากไปหาน้อยได้แก่แบบทดสอบความสามารถด้านคณิตศาสตร์ มีค่าเท่ากับ .395 แบบทดสอบความสามารถด้านภาษาอังกฤษ มีค่าเท่ากับ .261 แบบทดสอบความสามารถด้านวิทยาศาสตร์ มีค่าเท่ากับ .238 แบบทดสอบวัดความสามารถด้านภาษาไทย มีค่าเท่ากับ .198 และแบบทดสอบความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรม มีค่าเท่ากับ .044 โดยมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .156, .068, .057, .039 และ .002 ตามลำดับ

โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์น้อย ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 2433.115 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI)

มีค่าเท่ากับ .863 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .851 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .697 ดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเท่ากับ .012 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่เหมาะสม (RMSEA) มีค่าเท่ากับ .037 ที่องศาอิสระเท่ากับ 1175

1.3 ประสิทธิภาพของโมเดลการวิเคราะห์หองค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในองค์ประกอบความสามารถทางการเรียนของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 เขตพื้นที่การศึกษาเขต 2 จังหวัดบุรีรัมย์ มีรายละเอียดดังนี้ แบบสอบวัดความสามารถด้านภาษาไทยมีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางภาษาไทยข้อที่ 24 มีค่าเท่ากับ .138 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .016 แบบสอบวัดความสามารถด้านคณิตศาสตร์มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางคณิตศาสตร์ ข้อที่ 15 มีค่าเท่ากับ .113 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .013 แบบสอบวัดความสามารถด้านวิทยาศาสตร์มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางวิทยาศาสตร์ข้อที่ 24 มีค่าเท่ากับ .129 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .017 แบบสอบวัดความสามารถด้านสังคมศึกษา ศาสนา และวัฒนธรรม มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางสังคมศึกษา ศาสนา และวัฒนธรรมข้อที่ 5 มีค่าเท่ากับ .286 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .082 แบบสอบวัดความสามารถด้านภาษาอังกฤษ มีน้ำหนักองค์ประกอบมากที่สุดได้แก่ ข้อสอบวัดความสามารถทางภาษาอังกฤษข้อที่ 12 มีค่าเท่ากับ .120 และมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .014

โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1663.119 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .908 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .636 ดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเท่ากับ .008 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่เหมาะสม (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1125

1.4 ประสิทธิภาพของโมเดลการวิเคราะห์หองค์ประกอบเชิงยืนยันอันดับสองแฝงภายในองค์ประกอบผลสัมฤทธิ์ทางการเรียนของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 เขตพื้นที่การศึกษาเขต 2 จังหวัดบุรีรัมย์ พบว่าน้ำหนักองค์ประกอบแบบทดสอบความสามารถทางเรียนทั้ง 5 องค์ประกอบ มีค่าอยู่ระหว่าง .044 - .395 โดยน้ำหนักองค์ประกอบเรียงลำดับจากมากไปหาน้อยได้แก่แบบทดสอบความสามารถด้านคณิตศาสตร์ มีค่าเท่ากับ .395 แบบทดสอบความสามารถด้านภาษาอังกฤษ มีค่าเท่ากับ .261 แบบทดสอบความสามารถด้านวิทยาศาสตร์ มีค่าเท่ากับ .238

แบบทดสอบวัดความสามารถด้านภาษาไทย มีค่าเท่ากับ .198 และแบบทดสอบความสามารถด้านสังคมศึกษา ศาสนาและวัฒนธรรม มีค่าเท่ากับ .044 โดยมีค่าสัมประสิทธิ์การพยากรณ์เท่ากับ .156, .068, .057, .039 และ .002 ตามลำดับ

โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1658.181 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .908 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .634 ดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเท่ากับ .008 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1130

การเปรียบเทียบค่าดัชนีวัดความสอดคล้องของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยัน อันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน และ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน พบว่า โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด โดยมีค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1658.181 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .643 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1130 เมื่อทำการทดสอบความแตกต่างของ โมเดล โดยค่าไค - สแควร์พบว่า โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน และ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในแตกต่างจาก โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองอย่างมีนัยสำคัญทางสถิติ $p < .01$

2. การหาประสิทธิภาพของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันเมื่อเปลี่ยนขนาดกลุ่มตัวอย่าง

2.1 การหาประสิทธิภาพของ โมเดล โดยการเปรียบเทียบดัชนีวัดความกลมกลืนของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 100 ตัวอย่าง พบว่า โมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1674.043 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .634 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่า

เท่ากับ .462 คำนี้อัตราที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .036 ที่องศาอิสระเท่ากับ 1125 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1745.060 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .791 คำนี้อัตราที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (CFI) มีค่าเท่ากับ .088 คำนี้อัตราที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .046 ที่องศาอิสระเท่ากับ 1170 และมีโมเดลที่ไม่สามารถประมาณค่าพารามิเตอร์ได้คือ

2.4 การหาประสิทธิภาพของโมเดลโดยการเปรียบเทียบดัชนีวัดความกลมกลืนของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 800 ตัวอย่าง พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1658.181 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 คำนี้อัตราที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1130 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1663.119 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 คำนี้อัตราที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 770.451 และ 763.513 ที่องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ

2.5 การหาประสิทธิภาพของโมเดลโดยการเปรียบเทียบดัชนีวัดความกลมกลืนของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 1,200 ตัวอย่าง พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง

แฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1642.649 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .945 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .744 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .019 ที่องศาอิสระเท่ากับ 1130 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1644.203 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .945 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .741 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .020 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 1209.541 และ 1207.962 ที่องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ

2.6 การหาประสิทธิภาพของโมเดลโดยการเปรียบเทียบดัชนีวัดความกลมกลืนของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 1,600 ตัวอย่าง พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1729.321 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .956 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .769 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .018 ที่องศาอิสระเท่ากับ 1130 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1745.267 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .956 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .761 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .019 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 1594.032 และ 1578.077 ที่องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ

2.7 การหาประสิทธิภาพของโมเดลโดยการเปรียบเทียบดัชนีวัดความกลมกลืนของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 2,000 ตัวอย่าง พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1778.386 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .964 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .793 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .017 ที่องศาอิสระเท่ากับ 1130 รองลงมาได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1804.753 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .963 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .783 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .017 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 2008.655 และ 1982.248 ที่องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ

ผลของกราฟเปรียบเทียบแสดงความสัมพันธ์ของค่าไค-สแควร์ ของการวิเคราะห์ของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในจำแนกตามขนาดของกลุ่มตัวอย่าง พบว่าเมื่อเพิ่มขนาดของจำนวนตัวแปรต่อองค์ประกอบเส้นกราฟมีแนวโน้มเพิ่มขึ้นใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ในขณะที่ค่าไค-สแควร์มีแนวโน้มคงที่ใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน

3. การหาประสิทธิภาพของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันเมื่อเปลี่ยนขนาดข้อสอบต่อองค์ประกอบ

3.1 การหาประสิทธิภาพของโมเดลโดยการเปรียบเทียบดัชนีวัดความสอดคล้องของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบ

เชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 3 ข้อต่อองค์ประกอบ พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 127.753 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .979 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .818 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .027 ที่องศาอิสระเท่ากับ 80 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 138.145 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .977 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .760 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .032 ที่องศาอิสระเท่ากับ 75 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 130.275 และ 119.833 ที่องศาอิสระต่างกันเท่ากับ 5 และ 10 ตามลำดับ

3.2 การหาประสิทธิภาพของโมเดลโดยการเปรียบเทียบดัชนีวัดความสอดคล้อง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 5 ข้อต่อองค์ประกอบ พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 377.107 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .963 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .774 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 255 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 396.025 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .961 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .694 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .027 ที่องศาอิสระเท่ากับ 250 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 257.364 และ 238.446 ที่องศาอิสระต่างกันเท่ากับ 15 และ 20 ตามลำดับ

3.3 การหาประสิทธิภาพของ โมเดล โดยการเปรียบเทียบดัชนีวัดความสอดคล้อง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 8 ข้อต่อองค์ประกอบ พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1052.783 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .936 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .678 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .025 ที่องศาอิสระเท่ากับ 705 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1029.901 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .937 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .690 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .025 ที่องศาอิสระเท่ากับ 695 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 569.106 และ 591.988 ที่องศาอิสระต่างกันเท่ากับ 30 และ 40 ตามลำดับ

3.4 การหาประสิทธิภาพของ โมเดล โดยการเปรียบเทียบดัชนีวัดความสอดคล้องของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 10 ข้อต่อองค์ประกอบ พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1658.181 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .643 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1130 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1663.119 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .636 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดล

วิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง แฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 770.451 และ 756.513 ที่องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ

3.5 การหาประสิทธิภาพของ โมเดล โดยการเปรียบเทียบดัชนีวัดความสอดคล้องของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 15 ตัวแปรต่อองค์ประกอบ พบว่า โมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 3763.503 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .883 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .582 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .023 ที่องศาอิสระเท่ากับ 2630 รองลงมาได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 3766.498 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .883 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .579 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .023 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง แฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 1359.058 และ 1356.513 ที่องศาอิสระต่างกันเท่ากับ 65 และ 70 ตามลำดับ

กราฟเปรียบเทียบแสดงความสัมพันธ์ของค่าไค-สแควร์ ของการวิเคราะห์ของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในจำแนกตามจำนวนตัวแปรต่อองค์ประกอบ พบว่าเมื่อเพิ่มขนาดของจำนวนตัวแปรต่อองค์ประกอบเส้นกราฟมีแนวโน้มเพิ่มขึ้น โดยที่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองมีค่าไค-สแควร์สูงกว่า โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน

ตอนที่ 3 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบพหุมิติโดยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายใน

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในข้อสอบที่ให้คะแนนแบบสองค่า มีข้อจำกัดอยู่บนเงื่อนไขของข้อตกลงความเป็นมิติเดียวของแบบทดสอบ แนวคิดดังกล่าวได้ถูกนักวิจัยหลายคนวิพากษ์วิจารณ์ว่ามีจุดด้อยขาดความเหมาะสม (Ackerman, 1992, p. 67; Ackerman, & Evan, 1994, p. 329; Cammilli, 1992, p. 130; Mazor, Hambleton, & Clauser, 1998, p. 358; Oshima, & Miller, 1992, p. 237; Oshima, Raju, & Flowers, 1997, p. 253; Stout et al., 1997, p. 196) เนื่องจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนทั่วไปมักวัดความสามารถที่มีลักษณะเด่นหลายความสามารถประกอบกัน ซึ่งความสามารถในแต่ละมิตินิธิพลในการตอบข้อสอบ ดังนั้นการใช้คะแนนรวมของแบบทดสอบตามทฤษฎีการทดสอบแบบมาตรฐานเดิม หรือค่าประมาณความสามารถตามทฤษฎีการตอบข้อสอบ ภายใต้เงื่อนไขของข้อตกลงความเป็นมิติเดียวของแบบทดสอบ จึงไม่น่าเพียงพอที่จะนำมาใช้เป็นเกณฑ์การจับคู่ความสามารถของผู้สอบ ถ้าผู้สอบมาจากกลุ่มที่แตกต่างกันภายใต้การแจกแจงความสามารถหลายมิติ และข้อสอบสามารถจำแนกระดับความสามารถของผู้สอบในหลายมิติดังกล่าว การใช้แบบแผนการให้คะแนนความสามารถมิติเดียวจะมีผลทำให้ข้อสอบทำหน้าที่เบี่ยงเบน (Ackerman, 1992, p. 67) นั่นคือ เกิดความผิดพลาดของการเลือกใช้โมเดลสำหรับกาวิเคราะห์ ซึ่งจะส่งผลให้การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงเกินปกติ ดังนั้นในแบบทดสอบที่วัดคุณลักษณะซับซ้อน (Complex Traits) น่าจะใช้ความสามารถในทุกมิติเป็นเกณฑ์การจับคู่ เพื่อให้อัตราความคลาดเคลื่อนประเภทที่ 1 ลดลง ซึ่งในการวิจัยผู้วิจัยจึงได้เลือกใช้เทคนิคการวิเคราะห์องค์ประกอบในการทดสอบการทำหน้าที่ต่างกันของข้อสอบที่มีหลายองค์ประกอบซึ่ง Jensen (1980, pp. 446 – 450) กล่าวว่าข้อสอบที่มีน้ำหนักองค์ประกอบ (Factor Loading) ซึ่งมีขนาดและจำนวนไม่เท่ากันจากผู้สอบต่างกลุ่มกันเป็นข้อสอบที่มีความลำเอียง ซึ่งหมายความว่าข้อสอบนั้นไม่ได้วัดความสามารถเดียวกันในผู้สอบเดียวกันในผู้สอบแต่ละกลุ่ม โดยผู้วิจัยได้เลือกใช้โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในในการตรวจสอบเพราะมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ซึ่งได้แสดงผลการวิเคราะห์โดยมีรายละเอียดตามตาราง

1. ค่าสถิติพื้นฐานของคะแนนแบบสอบที่นำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน

1.1 ค่าสถิติพื้นฐานคะแนนที่ได้จากการทำแบบทดสอบวัดความสามารถทางการเรียน ด้านภาษาไทยของกลุ่มตัวอย่างที่ใช้ในทดสอบการทำหน้าที่ต่างกันของข้อสอบ พบว่าคะแนนที่ได้จากการทำแบบทดสอบวัดความสามารถทางการเรียนด้านภาษาไทยกลุ่มตัวอย่างที่เป็นเพศ

ความสามารถด้านภาษาอังกฤษข้อ 29 เท่ากับ 0.300 มีค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.459 และเมื่อพิจารณารายข้อพบว่าแบบทดสอบวัดความสามารถด้านภาษาอังกฤษ กลุ่มตัวอย่างที่เป็นเพศหญิงที่มีคะแนนเฉลี่ยต่ำที่สุดได้แก่ข้อสอบความสามารถด้านภาษาอังกฤษข้อ 21 มีค่าเฉลี่ย เท่ากับ 0.114 มีค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.318 ในขณะที่กลุ่มตัวอย่างที่เป็นเพศชายมีคะแนนเฉลี่ยต่ำที่สุดได้แก่ข้อสอบความสามารถด้านภาษาอังกฤษข้อ 12 มีค่าเฉลี่ย เท่ากับ 0.150 มีค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.357

2. การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน

2.1 การทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในที่ผู้วิจัยนำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนนมาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกัน โดยใช้กลุ่มตัวอย่างขนาด 2,000 โดยใช้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,048 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 952 คน พบว่า ไม่มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 แต่พบข้อสอบที่ทำหน้าที่ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 3 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และสังคมศาสนาและวัฒนธรรมข้อที่ 5 และข้อสอบวิชาภาษาอังกฤษข้อที่ 28

2.2 การทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในจำแนกตามขนาดกลุ่มตัวอย่าง

2.2.1 การทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีการวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในที่ผู้วิจัยนำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนนมาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกันจากการใช้กลุ่มตัวอย่างขนาด 1,600 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 833 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal

Group) จำนวน 767 คน พบว่า ไม่มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 แต่พบข้อสอบที่ทำหน้าที่ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 2 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และสังคมศาสนาและวัฒนธรรมข้อที่ 5

2.2.2 การทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีการวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในที่ผู้วิจัยนำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนนมาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกัน โดยใช้กลุ่มตัวอย่างขนาด 2,000 โดยใช้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,048 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 952 คน พบว่า ไม่มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 แต่พบข้อสอบที่ทำหน้าที่ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 3 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และสังคมศาสนาและวัฒนธรรมข้อที่ 5 และข้อสอบวิชาภาษาอังกฤษ ข้อที่ 28

2.2.3 การทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีการวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในที่ผู้วิจัยนำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนนมาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกันจากการใช้กลุ่มตัวอย่างขนาด 2,400 โดยใช้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,263 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,137 คน พบว่า ไม่มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 แต่พบข้อสอบที่ทำหน้าที่ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 2 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และวิชาคณิตศาสตร์ข้อที่ 31

2.2.4 การทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีการวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในที่ผู้วิจัยนำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนน

มาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกันจากการใช้กลุ่มตัวอย่างขนาด 2,800 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,449 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,351 คน พบว่า มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 จำนวน 2 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และวิชาคณิตศาสตร์ข้อที่ 31 และพบข้อสอบที่ทำหน้าที่ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 1 ข้อ ได้แก่วิชาภาษาอังกฤษข้อ 12

2.2.5 การทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยใช้วิธีการวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในที่ผู้วิจัยนำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนนมาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกันจากการใช้กลุ่มตัวอย่างขนาด 3,200 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,616 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,584 คน พบว่า มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 จำนวน 2 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และวิชาภาษาอังกฤษข้อ 12 และพบข้อสอบที่ทำหน้าที่ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 2 ข้อ ได้แก่วิชาคณิตศาสตร์ข้อที่ 31 และวิชาภาษาอังกฤษข้อ 31

2.2.6 การทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยใช้วิธีการวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในที่ผู้วิจัยนำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนนมาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกันจากการใช้กลุ่มตัวอย่างขนาด 3,600 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,911 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,689 คน พบว่า มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 จำนวน 1 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และพบข้อสอบที่ทำหน้าที่

ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 3 ข้อ ได้แก่ วิชาคณิตศาสตร์ข้อที่ 31 วิชาภาษาอังกฤษข้อ 31 วิชาภาษาอังกฤษข้อ 12

2.2.7 การทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีการวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในที่ผู้วิจัยนำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนนมาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณ ได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกันจากการใช้กลุ่มตัวอย่างขนาด 4,000 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 2,114 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,886 คน พบว่า มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 จำนวน 1 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และพบข้อสอบที่ทำหน้าที่ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 2 ข้อ ได้แก่ วิชาภาษาอังกฤษข้อ 31 และวิชาภาษาอังกฤษข้อ 12

2.3 การเปรียบเทียบระดับค่าความเชื่อมั่นของแบบทดสอบระหว่างข้อสอบที่มีการทำหน้าที่ต่างกันและข้อสอบที่ไม่มีข้อที่ทำหน้าที่ต่างกัน โดยการวิเคราะห์ G – Coefficient

2.3.1 ค่าความเชื่อมั่น โดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัด

ความสามารถทางการเรียนในภาพรวมเมื่อยังไม่ตัดข้อสอบที่พบว่าทำหน้าที่ต่างกัน พบว่าทำหน้าที่ต่างกันอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .738 และค่าความเชื่อมั่นโดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัดความสามารถทางการเรียนในภาพรวมเมื่อมีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .759

2.3.2 ค่าความเชื่อมั่น โดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัด

ความสามารถทางการเรียนวิชาภาษาไทยเมื่อยังไม่มีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกัน พบว่าทำหน้าที่ต่างกันพบว่าอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .720 และค่าความเชื่อมั่นโดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัดความสามารถทางการเรียนวิชาภาษาไทยเมื่อมีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .754

2.3.3 ค่าความเชื่อมั่น โดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัด

ความสามารถทางการเรียนวิชาคณิตศาสตร์เมื่อยังไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันพบว่าอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .737 และค่าความเชื่อมั่น โดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัด ความสามารถทางการเรียนวิชาคณิตศาสตร์เมื่อมีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันพบว่าอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .786

2.3.4 ค่าความเชื่อมั่นโดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัดความสามารถทางการเรียนวิชาวิทยาศาสตร์ไม่ตัดข้อสอบที่พบว่าทำหน้าที่ต่างกัน พบว่าทำหน้าที่ต่างกันอยู่ในระดับดีโดยมีค่าความเชื่อมั่นเท่ากับ .707

2.3.5 ค่าความเชื่อมั่น โดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัดความสามารถทางการเรียนวิชาสังคม ศาสนาและวัฒนธรรมเมื่อยังไม่ตัดข้อสอบที่พบว่าทำหน้าที่ต่างกัน พบว่าอยู่ในระดับดีโดยมีค่าความเชื่อมั่นเท่ากับ .825 และค่าความเชื่อมั่น โดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัดความสามารถทางการเรียนวิชาสังคม ศาสนาและวัฒนธรรมเมื่อตัดข้อสอบที่พบว่าทำหน้าที่ต่างกัน พบว่าทำหน้าที่ต่างกันอยู่ในระดับดีโดยมีค่าความเชื่อมั่นเท่ากับ .859

2.3.6 ค่าความเชื่อมั่น โดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัดความสามารถทางการเรียนวิชาภาษาอังกฤษเมื่อยังไม่ตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันพบว่าอยู่ในระดับไม่ดีโดยมีค่าความเชื่อมั่นเท่ากับ .413 และค่าความเชื่อมั่น โดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัดความสามารถทางการเรียนวิชาภาษาอังกฤษเมื่อตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันพบว่าอยู่ในระดับไม่ค่อยดีโดยมีค่าความเชื่อมั่นเท่ากับ .604

เมื่อการเปรียบเทียบระดับค่าความเชื่อมั่น โดยการวิเคราะห์ G – Coefficient ของแบบทดสอบวัดความสามารถทางการเรียนเมื่อไม่มีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันและหลังมีการตัดข้อสอบที่ทำหน้าที่ต่างกัน ในภาพรวมทุกวิชาพบว่าหลังจากตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันออกแล้วระดับค่าความเชื่อมั่นของแบบทดสอบเพิ่มขึ้นเท่ากับ .021 โดยที่ข้อสอบก่อนตัดข้อสอบที่มีการทำหน้าที่ต่างกันมีค่าความเชื่อมั่นเท่ากับ .738 และหลังตัดข้อสอบที่ทำหน้าที่ต่างกันออกมีค่าความเชื่อมั่นเท่ากับ .759 ตามลำดับเมื่อพิจารณาวิชาพบว่าข้อสอบที่มีค่าความเชื่อมั่นเพิ่มขึ้นมากที่สุดได้แก่วิชาภาษาอังกฤษหลังจากตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันออกแล้วระดับค่าความเชื่อมั่นของแบบทดสอบเพิ่มขึ้นเท่ากับ .191 โดยที่ข้อสอบก่อนตัดข้อสอบที่มีการทำหน้าที่ต่างกันมีค่าความเชื่อมั่นเท่ากับ .413 และหลังตัดข้อสอบที่ทำหน้าที่ต่างกันออก มีค่าความเชื่อมั่นเท่ากับ .604 ตามลำดับและยังพบว่าวิชาภาษาไทยมีระดับค่าความเชื่อมั่นของแบบทดสอบเพิ่มขึ้นน้อยเท่ากับ .017 โดยที่ข้อสอบก่อนตัดข้อสอบที่มีการทำหน้าที่ต่างกันมีค่าความเชื่อมั่นเท่ากับ .737 และหลังตัดข้อสอบที่ทำหน้าที่ต่างกันออก มีค่าความเชื่อมั่นเท่ากับ .754 ตามลำดับ

อภิปรายผลการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาประสิทธิภาพของ โมเดลในการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายใน ที่พัฒนาตามทฤษฎีเซาว์ปัญญาของชาร์ล สเปียร์แมน และปัจจัย

ที่มีอิทธิพลต่อลักษณะการแจกแจงค่าดัชนีวัดความสอดคล้องซึ่ง ได้แก่ จำนวนกลุ่มตัวอย่าง จำนวน ตัวดัชนีชี้วัดต่อองค์ประกอบ ให้ได้รูปแบบที่สอดคล้องกับทฤษฎีและหาเกณฑ์ที่เหมาะสม ในการตัดสินใจสำหรับการตรวจสอบความสอดคล้องของ โมเดลกับข้อมูลเชิงประจักษ์ในการ วิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายใน โดยการเปรียบเทียบกับ โมเดลที่นิยมใช้อยู่ในปัจจุบัน

กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงาน เขตพื้นที่การศึกษามัธยมศึกษา เขต 2 ปีการศึกษา 2548 ที่เข้าสอบวัดผลสัมฤทธิ์ทางการเรียนเพื่อ ประเมินคุณภาพการศึกษาระดับชาติ จำนวน 114 โรงเรียน โดยเลือกกลุ่มตัวอย่างสำหรับนักเรียนที่ เข้าสอบวิชาภาษาไทย วิชาคณิตศาสตร์ วิชาสังคม ศาสนาและวัฒนธรรม วิทยาศาสตร์ และ ภาษาอังกฤษ จำนวน 5,011 คน

ค่าสถิติพื้นฐานของตัวแปรที่นำมาใช้เป็นกลุ่มตัวอย่าง เป็นการวิเคราะห์ข้อมูลจาก กลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงานคณะกรรมการเขตพื้นที่การศึกษามัธยมศึกษา เขต 2 จำนวน 2,000 คน โดยแบ่งเป็นกลุ่มตัวอย่างขนาด 100, 200, 400, 800, 1,200, 1,600, และ 2,000 คน โดยมีจุดมุ่งหมายเพื่อหาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด ค่าสูงสุด ค่าความโค้ง ค่าความเบ้ ของตัวแปรสังเกตได้ที่ใช้ในการตรวจสอบประสิทธิภาพของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยัน ทั้งนี้เพื่อศึกษาลักษณะการแจกแจงของตัวแปร โดยผู้วิจัย นำเสนอในภาพรวม จำแนกตามขนาดของกลุ่มตัวอย่าง พบว่าตัวแปรทุกตัวมีการแจกแจงแบบปกติ ทั้งนี้เพราะค่าความเบ้ ความโค้งมีค่าเข้าใกล้ศูนย์ตามข้อเสนอแนะของ ไคลน์ (Kline, 2005) ได้ กล่าวว่าในการแจกแจงตัวแปรเป็นแบบปกติเมื่อค่าความเบ้ต่ำกว่า 3 และค่าความโค้งมีค่าต่ำกว่า 10 ซึ่งยอมรับได้ว่าข้อมูลมีการแจกแจงปกติตามรายตัวแปร ดังนี้ผู้วิจัยจึงใช้ข้อมูลดังกล่าวนี้ในการ วิเคราะห์เพื่อตรวจสอบสมมุติฐาน โดยไม่มีการเปลี่ยนแปลงข้อมูลแต่อย่างใด ส่วนเวสต์และคณะ (West et al., 1995) ก็ได้ให้คำแนะนำในการตรวจสอบความเบ้และความโค้งของตัวแปรสังเกตได้ แต่ละตัว ถ้ามีความเบ้มากกว่า 2.00 และมีค่าความโค้งมากกว่า 7.00 แสดงว่า ลักษณะการแจกแจง ข้อมูลไม่เป็นแบบปกติ ควรใช้กลุ่มตัวอย่างขนาดใหญ่ขึ้น แต่ในกรณีที่ใช้กลุ่มตัวอย่างขนาดใหญ่ มาก (มากกว่า 1,000 คน) ไม่ต้องห่วงเรื่องลักษณะการแจกแจงข้อมูลไม่เป็นแบบปกติ (Amemiya & Anderson, 1990 cited in Hu & Bentler, 1999) นอกจากนี้ผู้วิจัยยังต้องพิจารณาว่าถ้า โมเดล องค์ประกอบที่ศึกษามีความซับซ้อน (ประมาณพารามิเตอร์หลายตัว) ก็ต้องใช้กลุ่มตัวอย่างขนาดใหญ่ขึ้นซึ่งในการวิจัยครั้งนี้ผู้วิจัยจึงใช้กลุ่มตัวอย่างในการทดสอบตั้งแต่ 100, 200, 400, 800, 1,200, 1,600 และ 2,000 เพื่อให้สอดคล้องและครอบคลุม รวมทั้งเกิดความแม่นยำในการทดสอบโมเดล โดยของเสนอการอภิปรายผลดังต่อไปนี้

สมมุติฐานข้อที่ 1 การหาประสิทธิภาพของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง อันดับสองและโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง อันดับสองแฝงภายใน เพื่อหาโมเดลวัดที่เหมาะสมกับข้อมูลเชิงประจักษ์ การวิเคราะห์ข้อมูลเพื่อหาประสิทธิภาพของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันเป็นการศึกษาประสิทธิภาพของโมเดลการวิเคราะห์องค์ประกอบตามทฤษฎีสององค์ประกอบของชาร์ล สเปียร์แมนโดยการเปรียบเทียบค่าของดัชนีวัดความสอดคล้อง (Fit Indexes) ใน 4 โมเดล คือ โมเดลที่ 1 การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลที่ 2 การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลที่ 3 การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน (Nested Factors) โมเดลที่ 4 การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแบบแฝงภายใน (Second Order Nested Factors)

ผลการวิเคราะห์องค์ประกอบความสามารถทางการเรียนการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์น้อย ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 2428.632 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .850 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .792 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .716 ดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเท่ากับ .012 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .037 ที่องศาอิสระเท่ากับ 1170

และพบว่าประสิทธิภาพของโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง องค์ประกอบวัดความสามารถทางการเรียน โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์น้อย ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 2433.115 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .863 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .851 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .697 ดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเท่ากับ .012 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .037 ที่องศาอิสระเท่ากับ 1175

สำหรับการหาประสิทธิภาพของโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในองค์ประกอบความสามารถทางการเรียน พบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1663.119 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .908 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .636 ดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่าเท่ากับ .008 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1125 และยัง

พบว่าข้อสอบบางข้อไม่ได้แบ่งน้ำหนักไปยังองค์ประกอบที่ต้องการวัดซึ่งหมายความว่าข้อสอบข้อนี้หรือขีบายตัวแปรแฝงได้น้อยมากหรือไม่ได้อธิบายตัวแปรแฝงนั้นเลยเมื่อมีการใช้โมเดลวิเคราะห์องค์ประกอบแบบแฝงภายใน ซึ่งสอดคล้องกับงานวิจัยของจิกแนค (Gignac, 2005 c) ได้ทำการทดสอบโครงสร้างขององค์ประกอบของแบบทดสอบ WAIS-R โดยใช้วิธีการวิเคราะห์องค์ประกอบแฝงภายในซึ่งที่ผ่านมานี้ในรูปแบบการวิเคราะห์องค์ประกอบเชิงยืนยันในการทดสอบแบบทดสอบวัดเชาว์ปัญญา WAIS-R โดยการวิเคราะห์ความแปรปรวนภายในแบบทดสอบย่อยโดยใช้วิธีออบลิคและการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ความพยายามทำให้เกิดความผิดพลาดในการสรุปที่เหมาะสมกับความสอดคล้องของข้อมูลบนพื้นฐานของแนะนำในการวิเคราะห์องค์ประกอบเชิงยืนยันโดยใช้มาตรฐานของข้อมูล WAIS-R ในการอธิบาย ซึ่งสิ่งที่สำคัญแบบทดสอบ WAIS-R มีแนวคิดที่ดีกว่าในการวัดองค์ประกอบเชิงยืนยันอันดับหนึ่งในองค์ประกอบทั่วไป และกลุ่มของระดับองค์ประกอบ ผลสรุปสามารถอธิบายความสัมพันธ์ทางภาษาและคุณสมบัติของคะแนนเชาว์ปัญญาและเกณฑ์ภายนอกและพบความแตกต่างของแบบสอบย่อย VIQ/PIQ ของคะแนนและเกณฑ์ในการทดสอบว่า แบบสอบย่อยทางคณิตศาสตร์และหน่วยของตัวเศษไม่ได้แบ่งความแปรปรวนให้กับแบบสอบย่อย VIQ โดยขึ้นอยู่กับเชาว์ปัญญาทั่วไป (General Intelligent) ซึ่งมีความพยายามกระตุ้นและส่งเสริมแนวคิดของ โมเดลองค์ประกอบทางเชาว์ปัญญาที่เรียกว่าองค์ประกอบแฝงภายในในการพิจารณาพบที่มีความสอดคล้องของโมเดลที่เหนือกว่า นอกจากนั้น จิกแนค (Gignac, 2006 a) ได้ศึกษาโครงสร้างของการวิเคราะห์องค์ประกอบเชิงยืนยันในการทดสอบความเป็นเอกมิติโดยการใช้โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในซึ่งประกอบด้วยการวิเคราะห์องค์ประกอบทั่วไปในลำดับที่หนึ่ง เป็นแบบทดสอบเชาว์ปัญญาทางภาษา แสดงให้เห็นหลักฐานที่แสดงว่าองค์ประกอบย่อยทางคณิตศาสตร์ไม่ได้แบ่งความแปรปรวนให้กับองค์ประกอบทางภาษา ซึ่งไปขึ้นอยู่กับองค์ประกอบทั่วไป ซึ่งผลสรุปสามารถอธิบายได้ว่ามีประโยชน์น้อยมากในการใช้วิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งและรูปแบบด้วยวิธีออบลิคในงานวิจัยที่เกี่ยวกับเรื่องเชาว์ปัญญา สำหรับการตรวจสอบความสอดคล้องของ โมเดลสามารถอธิบายความกำกวมที่เกิดขึ้นได้อย่างเหมาะสมระหว่างองค์ประกอบทางเชาว์ปัญญาและเกณฑ์ภายนอก

ในขณะที่ประสิทธิภาพของโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในองค์ประกอบความสามารถทางการเรียน โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1658.181 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .908 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .634 ดัชนีรากกำลังสองเฉลี่ย

ของเศษ (RMR) มีค่าเท่ากับ .008 คำนี้อัตราการที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุคูณ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1130

จากการตรวจสอบดัชนีวัดซึ่งประกอบด้วยค่าไค-สแควร์ (χ^2) ค่าดัชนีวัดความกลมกลืน (GFI) ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) คำนี้อัตราการที่สองเฉลี่ยของเศษ (RMR) คำนี้อัตราการที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุคูณ (RMSEA) และค่าองศาอิสระพบว่าโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1658.181 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .908 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .634 คำนี้อัตราการที่สองเฉลี่ยของเศษ (RMR) มีค่าเท่ากับ .008 คำนี้อัตราการที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุคูณ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1130 รองลงมาได้แก่ โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในองค์ประกอบความสามารถทางการเรียน โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี ซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1663.119 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ค่าดัชนีวัดความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ .908 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .636 คำนี้อัตราการที่สองเฉลี่ยของเศษ (RMR) มีค่าเท่ากับ .008 คำนี้อัตราการที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พหุคูณ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1125 ซึ่งสอดคล้องกับผลการวิจัยของโรสเซน (Rossen, 2008) ได้ศึกษาการวิเคราะห์องค์ประกอบในการทดสอบทางเชาว์ปัญญา (The Mayer-Salovey-Caruso Emotion) ซึ่งในการศึกษานี้ได้ค้นหาในการทดสอบการวัดโครงสร้าง โดยการทดสอบซ้ำจากงานวิจัยที่ใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน โดยแสดงโมเดลดังนี้ คือ 1. จำนวน 1 องค์ประกอบและมืองค์ประกอบทั่วไป 2. จำนวน 2 องค์ประกอบเพื่อสะท้อนให้เห็นผลจากประสบการณ์ 3. โมเดลซึ่งประกอบด้วย 4 องค์ประกอบหมุนแบบออบลิค 4. โมเดลแบบ 3 องค์ประกอบหมุนแบบออบลิคซึ่งประกอบด้วยแบบวัดการรับรู้ความรู้สึก, ความเข้าใจทางอารมณ์ และการจัดการอารมณ์ 5. โมเดลองค์ประกอบทั่วไปแบบแฝงภายใน ในด้านการรับรู้ด้านอารมณ์ และหมุนแบบออบลิค ในองค์ประกอบ ด้านความเข้าใจในอารมณ์และการจัดการด้านอารมณ์ ผลสรุปของการวิเคราะห์พบว่า จากการทดสอบเพื่อตรวจสอบผลการศึกษาของ Gignac (2005 b) พบว่าแบบทดสอบ MSCEIT ไม่ได้ถูกวัดในทุกโครงสร้างตามความต้องการของผู้วิจัย ยิ่งกว่านั้น

ผู้วิจัย ยิ่งกว่านั้นการออกแบบโมเดลที่มีการวิเคราะห์องค์ประกอบทั่วไปสามารถทำให้การทดสอบ
 บริสุทธิ์ขึ้นภายใต้ทฤษฎี

การเปรียบเทียบค่าดัชนีวัดความสอดคล้องของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยัน
 อันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิง
 ยืนยันอันดับหนึ่งแฝงภายใน และ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน
 พบว่าโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในมีความสอดคล้องกับข้อมูลเชิง
 ประจักษ์มากที่สุด โดยมีค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1658.181 ค่าความน่าจะเป็นเท่ากับ .000
 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่า
 เท่ากับ .643 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ
 (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1130 เมื่อทำการทดสอบความแตกต่างของโมเดล
 โดยค่าไค - สแควร์พบว่า โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน และ
 โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในแตกต่างจาก โมเดลวิเคราะห์
 องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองอย่างมี
 นัยสำคัญทางสถิติ $p < .01$ ซึ่งสอดคล้องกับผลการศึกษาของจิกแนก (Gigac, 2008) ได้ศึกษา
 รูปแบบของ โมเดลระดับสูงเปรียบเทียบอิทธิพลทางตรงกับ โมเดลลำดับขั้น พบว่า มีบทบาทอย่าง
 มากในการทดสอบและงานวิจัยทางเซวี่ปัญญาในเรื่องแนวคิดค่า g ซึ่งอธิบายโดย โมเดลที่มี
 อิทธิพลตรงความลำดับขั้น ซึ่งจากการวิจัยยังพบว่ามีความเหมือนและความแตกต่างอย่างชัดเจน
 ระหว่างอิทธิพลตรงและอิทธิพลอ้อมในรูปแบบลำดับขั้น จากการทดสอบซ้ำโดยใช้เมตริกซ์
 สหสัมพันธ์จำนวน 5 ครั้ง เมื่อพิสูจน์การวิเคราะห์องค์ประกอบเชิงยืนยันตามแนวคิดของค่า g ใน
 โมเดลระดับสูงที่มีมากกว่าการวิเคราะห์องค์ประกอบเชิงยืนยันในระดับแรก

จากข้อสนับสนุนดังกล่าวมาข้างต้นจึงสรุปได้ว่า โมเดลวิเคราะห์องค์ประกอบเชิงยืนยัน
 อันดับสองที่มีการแสดงอิทธิพลขององค์ประกอบทั่วไป ไปยังคิวแปรสังเกตได้โดยตรงมีความ
 สอดคล้องกับข้อมูลเชิงประจักษ์มากกว่าโมเดลอื่น ๆ ในการทดสอบ

สมมุติฐานข้อที่ 2 การตรวจสอบการแจกแจงดัชนีวัดความสอดคล้องของ โมเดลการ
 วิเคราะห์องค์ประกอบเชิงยืนยันเมื่อเปลี่ยนขนาดกลุ่มตัวอย่าง

การตรวจสอบการแจกแจงประสิทธิภาพดัชนีวัดความสอดคล้องของ โมเดลโดยการ
 เปรียบเทียบดัชนีวัดความกลมกลืนของ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดล
 วิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝง
 ภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 100
 ตัวอย่าง พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์

องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1674.043 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .634 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .148 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .070 ที่องศาอิสระเท่ากับ 1125 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1840.975 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .561 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .011 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .076 ที่องศาอิสระเท่ากับ 1170 และมีโมเดลที่ไม่สามารถประมาณค่าพารามิเตอร์ได้คือ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน

สำหรับกลุ่มตัวอย่างขนาด 200 ตัวอย่าง พบว่า โมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1533.812 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .764 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .348 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .043 ที่องศาอิสระเท่ากับ 1125 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1745.060 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .730 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .118 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .050 ที่องศาอิสระเท่ากับ 1170 และมีโมเดลที่ไม่สามารถประมาณค่าพารามิเตอร์ได้คือ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน

เนื่องจากผู้วิจัยส่วนมากใช้วิธีการประมาณแบบความควรจะเป็นสูงสุด (Maximum Likelihood: ML) โดยปกติวิธี ML มีข้อเสนอแนะว่า ควรใช้กลุ่มตัวอย่างอย่างน้อย 100-200 หน่วยตัวอย่าง หรือกรณีที่ต้องการเปรียบเทียบคุณสมบัติของเครื่องมือวัดระหว่างกลุ่มตัวอย่างต่างกลุ่ม กลุ่มตัวอย่างแต่ละกลุ่มควรมีกว้างละ 100 – 200 หน่วยตัวอย่าง แฟน และ แวง (Fan & Wang, 1998) ได้ศึกษาขนาดกลุ่มตัวอย่างในโมเดล 3 องค์ประกอบ โดยใช้สถานการณ์พบว่า การใช้กลุ่มตัวอย่างขนาด 100-200 หน่วย ตัวอย่าง อาจได้คำตอบที่ไม่เหมาะสมหรือได้ค่าสถิติที่เป็นไปไม่ได้ เช่น ค่าความแปรปรวนเป็นลบ เป็นต้น แต่ถ้าใช้กลุ่มตัวอย่างตั้งแต่ 500 ตัวอย่างขึ้นไป กลับไม่พบค่าที่ไม่เหมาะสม

กลุ่มตัวอย่างขนาด 400 ตัวอย่าง พบว่า โมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 2173.850 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .852 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .462 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .036 ที่องศาอิสระเท่ากับ 1125 รองลงมาได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1745.060 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .791 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .088 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .046 ที่องศาอิสระเท่ากับ 1170 และมีโมเดลที่ไม่สามารถประมาณค่าพารามิเตอร์ได้คือ

กลุ่มตัวอย่างขนาด 800 ตัวอย่าง พบว่า โมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1658.181 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .640 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1130 รองลงมาได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1663.119 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .636 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน พบว่าแตกต่างกันมีค่าเท่ากับ 770.451 และ 763.513 ที่องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ ส่วนข้อตกลงเบื้องต้นข้อสองเป็นเรื่องที่ปฏิบัติได้ยาก แต่วิธี ML มีความแข็งแกร่งต่อการฝ่าฝืนข้อตกลงเบื้องต้นเรื่องนี้ (Chou & Bentler, 1995) เว้นแต่ในกรณีที่ใช้กลุ่มตัวอย่างขนาดเล็กและ โมเดลมีความซับซ้อน ดังนั้นผู้วิจัยควรใช้กลุ่มตัวอย่างอย่างน้อย 100-200 หน่วยตัวอย่างขึ้นไป หรือในกรณีตรวจสอบเครื่องมือวัดที่มีตั้งแต่ 3 องค์ประกอบขึ้นไป ควรใช้กลุ่มตัวอย่างตั้งแต่ 500 หน่วยตัวอย่างขึ้นไป (Aroian & Norris, 2001)

กลุ่มตัวอย่างขนาด 1,200 ตัวอย่าง พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1642.649 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .945 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .744 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .019 ท้องศาอิสระเท่ากับ 1130 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1644.203 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .945 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .741 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .020 ท้องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน พบว่าแตกต่างกันมีค่าเท่ากับ 1209.541 และ 1207.962 ท้องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ

กลุ่มตัวอย่างขนาด 1,600 ตัวอย่าง พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1729.321 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .956 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .769 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .018 ท้องศาอิสระเท่ากับ 1130 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1745.267 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .956 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .761 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .019 ท้องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน พบว่าแตกต่างกันมีค่าเท่ากับ 1594.032 และ 1578.077 ท้องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ

และกลุ่มตัวอย่างขนาด 2,000 ตัวอย่าง พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1778.386 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .964 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .793 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .017 ที่องศาอิสระเท่ากับ 1130 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1804.753 ค่าความน่าจะเป็นเท่ากับ .000 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .963 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .783 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .017 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 2008.655 และ 1982.248 ที่องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ

ผลของกราฟเปรียบเทียบแสดงความสัมพันธ์ของค่าไค-สแควร์ ของการวิเคราะห์ของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในจำแนกตามขนาดของกลุ่มตัวอย่าง พบว่าเมื่อเพิ่มขนาดของจำนวนข้อสอบต่อองค์ประกอบเส้นกราฟมีแนวโน้มเพิ่มขึ้นใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ในขณะที่ค่าไค-สแควร์มีแนวโน้มที่ในโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในและโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและพบว่าในการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในและการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในมาสามารถประมาณค่ากลุ่มตัวอย่างที่มีขนาดต่ำกว่า 400 ได้โปรแกรมได้เสนอว่าค่าเมทริกซ์ของความแปรปรวนมีค่าเป็นลบซึ่ง ชู และเบนท์เลอร์ (Chou & Bentler, 1995) ได้ให้ข้อคิดเห็นและแนะนำของส่วนข้อตกลงเบื้องต้นว่าเป็นเรื่องปฏิบัติได้ยาก ถึงแม้ว่าวิธี ML จะมีความแกร่งต่อการฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับขนาดและจำนวนกลุ่มตัวอย่าง เว้นแต่ในกรณีที่ใช้กลุ่มตัวอย่างขนาดเล็กและโมเดลมีความซับซ้อนน้อย ดังนั้นผู้วิจัยควรใช้กลุ่มตัวอย่างอย่างน้อย 100-200 หน่วยตัวอย่างขึ้นไป หรือในกรณีตรวจสอบเครื่องมือวัดที่มีตั้งแต่ 3 องค์ประกอบขึ้นไป ควรใช้กลุ่มตัวอย่างตั้งแต่ 500 หน่วยตัวอย่างขึ้นไป (Aroian & Norris, 2001)

ซึ่งผลการวิจัยก็สอดคล้องกับโบลลิน (Bollen, 1989) ได้เสนอแนะไว้กว้าง ๆ ว่า การประมาณค่าพารามิเตอร์อิสระ 1 ตัว ต้องใช้หลายหน่วยตัวอย่าง ลินเดแมน มีเรินดา และ โกลด์ (Lindeman, Merenda, & Gold, 1980) เสนอแนะหลักทั่ว ๆ ไปว่า อัตราส่วนระหว่างจำนวนหน่วยตัวอย่างกับจำนวนพารามิเตอร์หรือตัวแปรควรเป็น 20 : 1 ฮู และเบนท์เลอร์ (Hu & Bentler, 1999) เสนอหลักปฏิบัติในเรื่องนี้ว่า ควรใช้จำนวนหน่วยตัวอย่างมากกว่า 15 เท่า ของจำนวนพารามิเตอร์อิสระถ้าลักษณะการแจกแจงข้อมูลเป็นแบบปกติพหุนามและ ความตรง/ความเที่ยงของเครื่องมืออยู่ในเกณฑ์ดี ส่วนเวสต์ และคณะ (West et al., 1995) แนะนำให้ผู้วิจัยตรวจสอบความเบ้และความโด่งของตัวแปรสังเกตได้แต่ละตัว ถ้ามีความเบ้มากกว่า 2.00 และมีค่าความโด่งมากกว่า 7.00 แสดงว่า ลักษณะการแจกแจงข้อมูลไม่เป็นแบบปกติ ควรใช้กลุ่มตัวอย่างขนาดใหญ่ขึ้น แต่ในกรณีที่ใช้กลุ่มตัวอย่างขนาดใหญ่มาก (มากกว่า 1,000 คน) ไม่ต้องห่วงเรื่องลักษณะการแจกแจงข้อมูลไม่เป็นแบบปกติ (Amemiya & Anderson, 1990; Hu & Bentler, 1999) นอกจากนี้ผู้วิจัยยังต้องพิจารณาว่าถ้าโมเดลองค์ประกอบที่ศึกษามีความซับซ้อน(ประมาณพารามิเตอร์หลายตัว) ก็ต้องใช้กลุ่มตัวอย่างขนาดใหญ่ขึ้น

กลุ่มตัวอย่างยังมีขนาดใหญ่มากเท่าใดค่าที่ได้ยิ่งเข้าใกล้ค่าอันนั้นมากเท่านั้น (Bollen, 1989) กล่าวคือ ค่าสถิติไค-สแควร์มีแนวโน้มที่จะมีค่าสูง ทำให้ค่าสถิติไค-สแควร์มีโอกาสให้ค่านัยสำคัญ ($p \leq .05$) (นงลักษณ์ วิรัชชัย, 2542) ซึ่งชี้ว่าโมเดลองค์ประกอบกับข้อมูลเชิงประจักษ์ไม่สอดคล้องกัน ส่วนกลุ่มตัวอย่างขนาดเล็ก (น้อยกว่า 100 หน่วยตัวอย่าง) มีความน่าจะเป็นที่จะปฏิเสธโมเดลที่ถูกต้อง (True Model) เพิ่มขึ้นหรืออาจกล่าวได้ว่า การใช้กลุ่มตัวอย่างขนาดเล็กมีความเสี่ยงในการเกิดความคลาดเคลื่อนประเภท II (Type II Error) เพิ่มขึ้น

คอมเรย์และลี (Comrey & Lee, 1992) โดยได้สรุปจำนวนของกลุ่มตัวอย่างที่เหมาะสมในการวิเคราะห์ปัจจัยไว้ดังนี้ 100 = ไม่ดี, 200 = ดีปานกลาง, 300 = ดี, 500 = ดีมาก, 1,000 หรือมากกว่า = ดีที่สุด ซึ่งทำให้นักวิจัยได้รับแรงผลักดันให้ใช้กลุ่มตัวอย่างมากกว่า 500 (MacCallum, Widaman, Zhang & Hong, 1999, p. 84)

ซึ่งจากผลการวิจัยและข้อสนับสนุนดังกล่าวผู้วิจัยจึงสรุปได้ว่าการคัดเลือกกลุ่มตัวอย่างที่นำมาใช้ในการทดสอบองค์ประกอบ โดยเฉพาะในกรณีที่ใช้โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในควรมีขนาดกลุ่มตัวอย่างขนาด 800 คนขึ้นไปจึงจะเหมาะสม

สมมุติฐานข้อ 3 การตรวจสอบการแจกแจงดัชนีความสอดคล้องของโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันเมื่อเปลี่ยนขนาดข้อสอบต่อองค์ประกอบ

การตรวจสอบการแจกแจงประสิทธิภาพดัชนีความสอดคล้องของโมเดล โดยการเปรียบเทียบดัชนีวัดความกลมกลืนของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดล

วิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับสอง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับสองแฝงภายในของกลุ่มตัวอย่างขนาด 3 ตัว ข้อต่อองค์ประกอบ พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 127.753 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .979 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .818 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .027 ที่องศาอิสระเท่ากับ 80 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 138.145 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .977 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .760 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .032 ที่องศาอิสระเท่ากับ 75 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับสองแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 130.275 และ 119.833 ที่องศาอิสระต่างกันเท่ากับ 5 และ 10 ตามลำดับ

เมื่อพิจารณาจากกลุ่มตัวอย่างขนาด 5 ข้อต่อองค์ประกอบ พบว่าโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุดได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 377.107 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .963 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .774 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 255 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 396.025 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .961 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .694 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .027 ที่องศาอิสระเท่ากับ 250 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับสองแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันชั้นอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 257.364 และ 238.446 ที่องศาอิสระต่างกันเท่ากับ 15 และ 20 ตามลำดับ

กลุ่มตัวอย่างขนาด 8 ข้อต่อองค์ประกอบ พบว่า โมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1052.783 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .936 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .678 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .025 ที่องศาอิสระเท่ากับ 705 รองลงมาได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1029.901 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .937 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .690 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .025 ที่องศาอิสระเท่ากับ 695 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 569.106 และ 591.988 ที่องศาอิสระต่างกันเท่ากับ 30 และ 40 ตามลำดับ

กลุ่มตัวอย่างขนาด 10 ข้อต่อองค์ประกอบ พบว่า โมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1658.181 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .643 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1130 รองลงมาได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 1663.119 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .919 ดัชนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .636 ดัชนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่พอเหมาะ (RMSEA) มีค่าเท่ากับ .024 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าไค-สแควร์ (χ^2) เปรียบเทียบกับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 770.451 และ 756.513 ที่องศาอิสระต่างกันเท่ากับ 40 และ 45 ตามลำดับ

กลุ่มตัวอย่างขนาด 15 ตัวแปรต่อองค์ประกอบ พบว่า โมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด ได้แก่ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 3763.503 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ

.883 คำนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .582 คำนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่เหมาะสม (RMSEA) มีค่าเท่ากับ .023 ที่องศาอิสระเท่ากับ 2630 รองลงมาได้แก่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในซึ่งพิจารณาจากค่าไค-สแควร์ (χ^2) มีค่าเท่ากับ 3766.498 ค่าดัชนีวัดความกลมกลืน (GFI) มีค่าเท่ากับ .883 คำนีวัดความสอดคล้องเปรียบเทียบ (CFI) มีค่าเท่ากับ .579 คำนีวัดรากที่สองของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่เหมาะสม (RMSEA) มีค่าเท่ากับ .023 ที่องศาอิสระเท่ากับ 1125 และเมื่อพิจารณาความแตกต่างของค่าค่าไค-สแควร์ (χ^2) เปรียบเทียบกับโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง กับ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง แฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในพบว่าแตกต่างกันมีค่าเท่ากับ 1359.058 และ 1356.513 ที่องศาอิสระต่างกันเท่ากับ 65 และ 70 ตามลำดับ

กราฟเปรียบเทียบแสดงความสัมพันธ์ของค่าไค-สแควร์ ของการวิเคราะห์ของโมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่ง โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง โมเดลวิเคราะห์ องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายใน โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในจำแนกตามจำนวนตัวแปรต่อองค์ประกอบ พบว่าเมื่อเพิ่มขนาดของจำนวนตัวแปรต่อองค์ประกอบเส้นกราฟมีแนวโน้มเพิ่มขึ้น โดยที่โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองมีค่าไค-สแควร์สูงกว่า โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งแฝงภายในและ โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในซึ่งสอดคล้องกับผลการศึกษาของ จิกแนค (Gigac, 2008) ได้ศึกษารูปแบบของโมเดลระดับสูงเปรียบเทียบอิทธิพลทางตรงกับ โมเดลระดับชั้น เพื่อศึกษาการจำลองข้อมูลเพื่อตรวจสอบค่าดัชนีวัดความสอดคล้องในการวิเคราะห์องค์ประกอบเชิงยืนยันบนพื้นฐานของข้อมูลที่บิดเบือนข้อตกลงในการวิเคราะห์โมเดลสมการ โครงสร้าง โดยการเปรียบเทียบดัชนีวัดความสอดคล้อง ซึ่งได้แก่ CFI GFI TFI NNFT RMSEA SRMR และ Chi/df โดยการจำลองข้อมูลและกลุ่มตัวอย่างเป็น 3 ขนาด 250,500 และ 1,000 ตัวอย่าง โดยแต่ละองค์ประกอบมีขนาดหน้าหน้าองค์ประกอบเท่ากับ .40, .50, .60 และ .80 ซึ่งประกอบด้วย 4 องค์ประกอบและ 8 องค์ประกอบ พร้อมทั้งเปรียบเทียบความสัมพันธ์ของเมตริกซ์ความแปรปรวนและเมตริกซ์สหสัมพันธ์ โดยใช้การประมาณค่า วิธีโลกิสติก แบบตั้งฉากและแบบออบลิค พบว่า ค่าสหสัมพันธ์ของดัชนีวัดความสอดคล้องบางตัวมีค่าต่ำ ยิ่งไปกว่านั้นการบิดเบือนเพียงเล็กน้อยไม่ส่งผลกระทบต่อความไม่สอดคล้องของค่าดัชนี RMSEA และ SRMR แต่จะนำมาซึ่งความไม่สอดคล้องของดัชนี IFI

ซึ่งจากการวิจัยพบว่าค่าดัชนีวัดความสอดคล้องของ โมเดลมีขนาดเพิ่มขึ้นเพื่อเพิ่มจำนวนตัวแปรสังเกตได้ต่อองค์ประกอบจึงต้องนำมาเป็นส่วน หนึ่งในการพิจารณาในการสร้างข้อคำถาม

เพื่ออธิบายโครงสร้างที่ต้องการวัดเพราะอาจเป็นสาเหตุให้ค่าไอ-สแควร์ มีค่าสูงผิดปกติเสี่ยงต่อค่าความผิดพลาดที่เกิดจากโมเดลไม่สอดคล้องกับข้อมูลเชิงประจักษ์

สมมุติฐานข้อที่ 4 การตรวจสอบการทำหน้าที่ต่างกัน โดยการเปรียบเทียบความแตกต่างของน้ำหนักองค์ประกอบ (Factor Loading) ซึ่งเป็นการนำการวิเคราะห์องค์ประกอบมาใช้ในการวิเคราะห์โครงสร้างของแบบสอบแยกตามกลุ่มผู้สอบ ความไม่สอดคล้องกันระหว่างน้ำหนักตัวประกอบบนคุณลักษณะที่มุ่งวัด หรือความแตกต่างของค่าเฉลี่ยคะแนน ตัวประกอบ (Factor Score) ระหว่างกลุ่มผู้สอบย่อมสะท้อนให้เห็นการทำหน้าที่ต่างกันของข้อสอบและแบบสอบ ในการใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) ที่นำมาตรวจสอบการทำหน้าที่ต่างกัน มีจุดอ่อนในเรื่องความไม่สอดคล้องกันระหว่างน้ำหนักตัวประกอบซึ่งอาจเกิดจากความแตกต่างระหว่างความสามารถของกลุ่มได้ จึงควรใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) และนอกจากนี้ยังสามารถใช้ CFA สำหรับตรวจสอบความแตกต่างระหว่างกลุ่มในด้านคุณลักษณะหรือความสามารถหลักและความสามารถรองได้อีกด้วย (ศิริชัย กาญจนวาสี, 2550, หน้า 123; Camilli & Shepard, 1994)

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในข้อสอบที่ให้คะแนนแบบสองค่า มีข้อจำกัดอยู่บนเงื่อนไขของข้อตกลงความเป็นมิติเดียวของแบบทดสอบ แนวคิดดังกล่าวได้ถูกนักวิจัยหลายคนวิพากษ์วิจารณ์ว่ามีจุดด้อยขาดความเหมาะสม (Ackerman, 1992, p. 67; Ackerman & Evan, 1994, p. 329; Camilli, 1992, p. 130; Mazor, Hambleton, & Clauser, 1998, p. 358, Oshima & Miller, 1992, p. 237; Oshima, Raju, & Flowers, 1997, p. 253; Stout et al., 1997, p. 196) เนื่องจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนทั่วไปมักวัดความสามารถที่มีลักษณะเด่นหลาย ความสามารถประกอบกัน ซึ่งความสามารถในแต่ละมิตินี้อธิบายในการตอบข้อสอบ ดังนั้นการใช้คะแนนรวมของแบบทดสอบตามทฤษฎีการทดสอบแบบมาตรฐานเดิม หรือค่าประมาณความสามารถตามทฤษฎีการตอบข้อสอบ ภายใต้งื่อนไขของข้อตกลงความเป็นมิติเดียวของแบบทดสอบ จึงไม่น่าเพียงพอที่จะนำมาใช้เป็นเกณฑ์การจับคู่ความสามารถของผู้สอบ ถ้าผู้สอบมาจากกลุ่มที่แตกต่างกันภายใต้การแจกแจงความสามารถหลายมิติ และข้อสอบสามารถจำแนกระดับความสามารถของผู้สอบในหลายมิติดังกล่าว การใช้แบบแผนการให้คะแนนความสามารถมิติเดียวจะมีผลทำให้ข้อสอบทำหน้าที่เบี่ยงเบน (Ackerman, 1992, p. 67) นั่นคือ เกิดความผิดพลาดของการเลือกใช้โมเดลสำหรับการวิเคราะห์ ซึ่งจะส่งผลให้การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ มีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงเกินปกติ ดังนั้นในแบบทดสอบที่วัดคุณลักษณะซับซ้อน (Complex Traits) น่าจะใช้ความสามารถในทุกมิติเป็นเกณฑ์การจับคู่ เพื่อทำให้อัตราความคลาดเคลื่อนประเภทที่ 1 ลดลง ซึ่งในการวิจัยผู้วิจัยจึงได้เลือกใช้เทคนิคการวิเคราะห์องค์ประกอบ

ในการทดสอบการทำหน้าที่ต่างกันของข้อสอบที่มีหลายองค์ประกอบซึ่ง Jensen (1980, pp. 446 – 450) กล่าวว่าข้อสอบที่มีน้ำหนักองค์ประกอบ (Factor Loading) ซึ่งมีขนาดและจำนวนไม่เท่ากัน จากผู้สอบต่างกลุ่มกันเป็นข้อสอบที่มีความลำเอียง ซึ่งหมายความว่าข้อสอบนั้นไม่ได้วัดความสามารถเดียวกันในผู้สอบเดียวกันในผู้สอบแต่ละกลุ่ม โดยผู้วิจัยได้เลือกใช้โมเดลวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในสำหรับการตรวจสอบเพราะมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด โดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนนมาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกันโดยพบว่าไม่มีข้อสอบที่ทำหน้าที่ต่างกัน โดยใช้กลุ่มตัวอย่างขนาด 2,000 โดยใช้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,048 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 952 คน พบว่าไม่มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในข้อสอบที่วัดความสามารถหลายมิติได้มีนักวิจัยหลายคนให้ความสนใจ ดังเช่น โอชิม่า และมิลเลอร์ (Oshima & Miller, 1992) ได้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในข้อสอบที่วัดความสามารถสองมิติ ภายใต้ทฤษฎีการตอบข้อสอบแบบหลายมิติ โดยใช้ดัชนีการวัดพื้นที่ 4 วิธี คือ ดัชนีการวัดพื้นที่แบบคิดเครื่องหมาย (SA) ดัชนีการวัดพื้นที่แบบไม่คิดเครื่องหมาย (UA) ดัชนีผลรวมกำลังสองแบบคิดเครื่องหมาย (SSOS) และดัชนีผลรวมกำลังสองแบบไม่คิดเครื่องหมาย (USOS) ศึกษาภายใต้เงื่อนไขสัดส่วนของข้อสอบทำหน้าที่เบี่ยงเบน ความแตกต่างของความสามารถหลักและความสามารถแทรกซ้อน ผลการศึกษาพบว่า ข้อสอบแบบหลายมิติไม่จำเป็นต้องทำหน้าที่ต่างกัน ยกเว้น ค่าเฉลี่ยของความสามารถแทรกซ้อนแตกต่างกัน และเมื่อ ไม่มีความสัมพันธ์ระหว่างความสามารถหลักและความสามารถแทรกซ้อน ความแตกต่างของค่าเฉลี่ยของความสามารถหลักไม่ได้เป็นสาเหตุทำให้ข้อสอบทำหน้าที่ต่างกัน โอชิม่า ราจู และเฟลาเวอร์ส (Oshima, Raju & Flowers, 1997) ในขณะเดียวกัน คมศักดิ์ ชื่นชู (2539) ได้ศึกษาผลการวิเคราะห์ความลำเอียงที่ใช้วิธีต่างกันของแบบทดสอบวัดจริยธรรมด้านความซื่อสัตย์ โดยการเปรียบเทียบวิธีหาค่าอำนาจจำแนกของข้อสอบ วิธีวิเคราะห์องค์ประกอบและวิธีการวิเคราะห์ถดถอยจำแนกตามกลุ่มตัวอย่างที่ใช้ภาษาไทยกับไม่ใช้ภาษาไทยในชีวิตประจำวันและระหว่างกลุ่มตัวอย่างที่นับถือศาสนาพุทธและไม่นับถือศาสนาพุทธพบว่า ไม่พบการทำหน้าที่ต่างกัน ในวิธีวิเคราะห์องค์ประกอบแต่พบการทำหน้าที่ต่างกัน ในวิธีวิธีหาค่าอำนาจจำแนกของข้อสอบ

วิธีการวิเคราะห์ถดถอย และรัคเนอร์ และคณะ (Rudner et al., 1984) ได้ศึกษาการทำหน้าที่ต่างกันของข้อสอบมาตรฐานวัดผลสัมฤทธิ์ของสแตนฟอร์ด (Stanford Achievement) โดยใช้วิธีแปลงค่าความยาก วิธีไค-สแควร์ วิธีโค้งคุณลักษณะข้อสอบและวิธีวิเคราะห์องค์ประกอบผลการวิจัยพบว่าวิธีโค้งคุณลักษณะข้อสอบและวิธีแปลงค่าความยากเป็นวิธีที่มีความเหมาะสมสำหรับวิธีวิเคราะห์องค์ประกอบเป็นวิธีที่รองลงมา ซึ่งจากการศึกษาการทำหน้าที่ต่างกัน โดยใช้วิธีวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในที่ผู้วิจัยเสนอแล้วไม่พบการทำหน้าที่ต่างกันของข้อสอบในทุกรายวิชาที่ระดับนัยสำคัญ .01 โดยใช้การเปรียบเทียบน้ำหนักองค์ประกอบอาจมาจากสาเหตุที่ต้องมีการแบ่งน้ำหนักองค์ประกอบไปให้กับองค์ประกอบทั่วไปที่มีอิทธิพลตรงไปยังข้อสอบซึ่งจะทำให้น้ำหนักองค์ประกอบมีค่าน้อยลงกว่าปกติแต่สามารถจำแนกน้ำหนักองค์ประกอบได้สอดคล้องกับข้อมูลเชิงประจักษ์มากขึ้นจึงทำให้พบข้อสอบที่ทำหน้าที่ต่างกันยากขึ้น

สมมุติฐานข้อที่ 5 การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายใน ที่ผู้วิจัยนำมาใช้ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยจะถือว่าข้อสอบมีความลำเอียงก็ต่อเมื่อมีความแตกต่างของน้ำหนักองค์ประกอบในรูปแบบคะแนนมาตรฐานของความสามารถรอง (G) โดยใช้เกณฑ์พิจารณาว่าค่าไค-สแควร์ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าไค-สแควร์ที่เปิดตารางที่ระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งมีค่าเท่ากับ 6.64 แสดงว่าข้อสอบทำหน้าที่ต่างกัน โดยใช้กลุ่มตัวอย่างขนาด 2,000 โดยใช้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,048 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 952 คน พบว่า ไม่มีข้อสอบข้อใดทำหน้าที่ต่างกันที่ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 ซึ่งสอดคล้องกับผลการศึกษาของคมสัน ชื่นชม (2539) ที่ทำการศึกษาวินิจฉัยความลำเอียงที่ใช้วิธีต่างกันของแบบทดสอบวัดจริยธรรมด้านความซื่อสัตย์โดยการใช้วิธีการวิเคราะห์องค์ประกอบเปรียบเทียบกับวิธีค่าอำนาจจำแนกของข้อสอบและวิธีวิเคราะห์การถดถอย พบว่าไม่พบข้อสอบที่มีความลำเอียงในการวิเคราะห์โดยใช้วิธีการวิเคราะห์องค์ประกอบ แต่พบในวิธีอื่น ๆ ที่กล่าวมา สำหรับข้อสอบที่ทำหน้าที่ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 3 ข้อได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และสังคมศาสนาและวัฒนธรรมข้อที่ 5 และข้อสอบวิชาภาษาอังกฤษ ข้อที่ 28 และนอกจากนั้นยังสอดคล้องกับงานวิจัยของรัคเนอร์ และคณะ (Rudner et al., 1984) ที่ศึกษาเปรียบเทียบวิธีวิเคราะห์องค์ประกอบ วิธีการแปลงค่าความยาก วิธีไค – สแควร์ วิธีโค้งคุณลักษณะข้อสอบ พบว่าวิธีวิเคราะห์องค์ประกอบมีความเหมาะสมน้อยที่สุด ซึ่งอาจต่างจากงานวิจัยนี้ที่พบการทำหน้าที่ต่างกันของข้อสอบในเกือบทุกรายวิชาที่นำมาวิเคราะห์

เมื่อทำการทดสอบการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองแฝงภายในจำแนกตามขนาดกลุ่มตัวอย่าง พบว่าที่ขนาดกลุ่มตัวอย่างขนาด 1,600 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 833 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 767 คน พบว่า ไม่มีข้อสอบข้อใดทำหน้าที่ต่างกันว่า ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 แต่พบข้อสอบที่ทำหน้าที่ต่างกันว่าระดับนัยสำคัญ .05 จำนวน 2 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และสังคมศาสตร์และวัฒนธรรมข้อที่ 5 ในขณะที่ใช้กลุ่มตัวอย่างขนาด 2,400 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,263 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,137 คน พบว่า ไม่มีข้อสอบข้อใดทำหน้าที่ต่างกันว่า ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 แต่พบข้อสอบที่ทำหน้าที่ต่างกันว่าระดับนัยสำคัญ .05 จำนวน 2 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และวิชาคณิตศาสตร์ข้อที่ 31 เมื่อเปลี่ยนขนาดกลุ่มตัวอย่างเป็น 2,800 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,449 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,351 คน พบว่า มีข้อสอบข้อใดทำหน้าที่ต่างกันว่า ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 จำนวน 2 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และวิชาคณิตศาสตร์ข้อที่ 31 และพบข้อสอบที่ทำหน้าที่ต่างกันว่าระดับนัยสำคัญ .05 จำนวน 1 ข้อ ได้แก่ วิชาภาษาอังกฤษข้อ 12

สำหรับการทดสอบ โดยใช้กลุ่มตัวอย่างขนาด 3,200 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,616 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,584 คน พบว่า มีข้อสอบข้อใดทำหน้าที่ต่างกันว่า ค่าไค-สแควร์ มีระดับนัยสำคัญทางสถิติเท่ากับ .01 จำนวน 2 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และวิชาภาษาอังกฤษข้อ 12 และพบข้อสอบที่ทำหน้าที่ต่างกันว่าระดับนัยสำคัญ .05 จำนวน 2 ข้อ ได้แก่วิชาคณิตศาสตร์ข้อที่ 31 และวิชาภาษาอังกฤษข้อ 31 ในขณะที่ใช้กลุ่มตัวอย่างขนาด 3,600 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 1,911 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,689 คน พบว่า มีข้อสอบข้อใดทำหน้าที่ต่างกันว่า ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01 จำนวน 1 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และพบข้อสอบที่ทำหน้าที่ต่างกันว่าระดับนัยสำคัญ .05 จำนวน 3 ข้อ ได้แก่วิชาคณิตศาสตร์ข้อที่ 31 วิชาภาษาอังกฤษข้อ 31 วิชาภาษาอังกฤษข้อ 12

และเมื่อใช้กลุ่มตัวอย่างขนาด 4,000 โดยให้กลุ่มตัวอย่างที่เป็นเพศหญิงเป็นกลุ่มอ้างอิง (Reference Group) จำนวน 2,114 คน ใช้เพศชายเป็นกลุ่มเปรียบเทียบ (Focal Group) จำนวน 1,886 คน พบว่า มีข้อสอบข้อใดทำหน้าที่ต่างกันว่า ค่าไค-สแควร์มีระดับนัยสำคัญทางสถิติเท่ากับ .01

จำนวน 1 ข้อ ได้ข้อสอบวิชาภาษาไทยข้อที่ 27 และพบข้อสอบที่ทำหน้าที่ต่างกันที่ระดับนัยสำคัญ .05 จำนวน 2 ข้อ ได้แก่ วิชาภาษาอังกฤษข้อ 31 และวิชาภาษาอังกฤษข้อ 12

จากการพิจารณาการพบข้อสอบที่ทำหน้าที่ต่างกันโดยใช้วิธีวิเคราะห์ห้วงค์ประกอบ อันดับสองแฝงภายในจะเห็นได้ว่าพบข้อสอบในข้อเดิมหรือมีแนวโน้มว่าเป็นข้อเดิมเมื่อเราเปลี่ยนขนาดกลุ่มตัวอย่างซึ่งอาจแตกต่างจากวิธีการตรวจสอบการทำหน้าที่ต่างกันด้วยวิธีอื่น ๆ เมื่อเปลี่ยนวิธีหรือเปลี่ยนขนาดของกลุ่มตัวอย่างจะทำให้ข้อสอบที่พบว่าทำหน้าที่ต่างกันเปลี่ยนแปลงข้อไป

สำหรับการเปรียบเทียบระดับค่าความเชื่อมั่นของแบบทดสอบระหว่างข้อสอบที่มีการทำหน้าที่ต่างกันและข้อสอบที่ไม่มีข้อที่ทำหน้าที่ต่างกันโดยการวิเคราะห์ G - Coefficient พบว่า ค่าความเชื่อมั่นโดยการวิเคราะห์ G - Coefficient ของแบบทดสอบวัด ความสามารถทางการเรียนในภาพรวมเมื่อยังไม่ตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .738 และค่าความเชื่อมั่นเมื่อมีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .759

เมื่อพิจารณาค่าความเชื่อมั่น โดยการวิเคราะห์ G - Coefficient ของแบบทดสอบวัด ความสามารถทางการเรียนเป็นรายวิชาพบว่าวิชาภาษาไทยเมื่อยังไม่มีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันมีค่าความเชื่อมั่นอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .754 และเมื่อมีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .720 ซึ่งพบว่าข้อสอบมีความเชื่อมั่นลดลง (Nandakumar, 1993, p. 294 citing Roznowski, 1987, Reith & Roznowski, 1992) กล่าวว่า การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบครั้งละข้อและนำไปพิจารณาตัดข้อสอบออกจากแบบทดสอบ อาจทำให้ความเที่ยงตรงของแบบทดสอบลดลง เนื่องจากกำจัดข้อสอบที่วัดคุณลักษณะที่เกี่ยวข้อง (Trait-Relevant) กับข้ออื่นออกไป สำหรับวิชาคณิตศาสตร์เมื่อยังไม่มีการตัดข้อสอบที่ทำหน้าที่ต่างกันพบว่าอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .737 และเมื่อมีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันพบว่าอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .786

สำหรับค่าความเชื่อมั่น โดยการวิเคราะห์ G - Coefficient ของแบบทดสอบวัด ความสามารถทางการเรียนวิชาวิทยาศาสตร์ไม่ตัดข้อสอบที่พบว่าทำหน้าที่ต่างกัน พบว่าทำหน้าที่ต่างกันอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .707 สำหรับวิชาสังคม ศาสนาและวัฒนธรรมเมื่อยังไม่ตัดข้อสอบที่พบว่าทำหน้าที่ต่างกัน พบว่าอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .825 และเมื่อตัดข้อสอบที่พบว่าทำหน้าที่ต่างกัน พบว่าทำหน้าที่ต่างกันอยู่ในระดับดี โดยมีค่าความเชื่อมั่นเท่ากับ .859 และวิชาภาษาอังกฤษเมื่อยังไม่ตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันพบว่าอยู่ในระดับไม่ดี โดยมีค่าความเชื่อมั่นเท่ากับ .413 และเมื่อตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันพบว่าอยู่ในระดับไม่พอควร โดยมีค่าความเชื่อมั่นเท่ากับ .604

สำหรับการพิจารณาเปรียบเทียบระดับค่าความเชื่อมั่น โดยการวิเคราะห์ $G - Coefficient$ ของแบบทดสอบวัดความสามารถทางการเรียนเมื่อไม่มีการตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันและ หลังมีการตัดข้อสอบที่ทำหน้าที่ต่างกัน ในภาพรวมทุกวิชาพบว่าหลังจากตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันออกแล้วระดับค่าความเชื่อมั่นของแบบทดสอบเพิ่มขึ้นเท่ากับ .021 โดยที่ข้อสอบก่อนตัดข้อสอบที่มีการทำหน้าที่ต่างกันมีค่าความเชื่อมั่นเท่ากับ .738 และหลังตัดข้อสอบที่ทำหน้าที่ต่างกันออก มีค่าความเชื่อมั่นเท่ากับ .759 ตามลำดับเมื่อพิจารณารายวิชาพบว่าข้อสอบที่มีค่าความเชื่อมั่นเพิ่มขึ้นมากที่สุดได้แก่วิชาภาษาอังกฤษหลังจากตัดข้อสอบที่พบว่าทำหน้าที่ต่างกันออกแล้วระดับค่าความเชื่อมั่นของแบบทดสอบเพิ่มขึ้นเท่ากับ .191 โดยที่ข้อสอบก่อนตัดข้อสอบที่มีการทำหน้าที่ต่างกันมีค่าความเชื่อมั่นเท่ากับ .413 และหลังตัดข้อสอบที่ทำหน้าที่ต่างกันออก มีค่าความเชื่อมั่นเท่ากับ .604 ตามลำดับและยังพบว่าวิชาภาษาไทยมีระดับค่าความเชื่อมั่นของแบบทดสอบเพิ่มขึ้นน้อยเท่ากับ .017 โดยที่ข้อสอบก่อนตัดข้อสอบที่มีการทำหน้าที่ต่างกันมีค่าความเชื่อมั่นเท่ากับ .737 และหลังตัดข้อสอบที่ทำหน้าที่ต่างกันออก มีค่าความเชื่อมั่นเท่ากับ .754 ตามลำดับ

นอกจากนั้น แนนคาคุมาร์ (Nandakumar, 1993, p. 294) เสนอแนะว่าการตรวจสอบ การทำหน้าที่ต่างกันของข้อสอบหลายข้อพร้อมกัน ทำให้สามารถศึกษาการทำหน้าที่ต่างกันของ ข้อสอบแบบขยายผล (DIF Amplification) และการทำหน้าที่ต่างกันของข้อสอบแบบหักล้างกัน (DIF Cancellation) ดังนั้น บางครั้งตรวจสอบไม่พบการทำหน้าที่ต่างกันของข้อสอบรายข้อ แต่เมื่อ พิจารณาเป็นหมวดข้อสอบ (Bundle of Items) หรืออาจพบการทำหน้าที่ต่างกันของหมวดข้อสอบ ได้ และในทำนองกลับกัน เมื่อตรวจสอบพบการทำหน้าที่ต่างกันของข้อสอบรายข้อต่อกลุ่มอ้างอิง และตรวจสอบพบข้อสอบข้ออื่น ๆ ทำหน้าที่ต่างกันต่อกลุ่มสนใจ เมื่อพิจารณาพร้อม ๆ กันทั้ง หมวดข้อสอบ อาจไม่พบการทำหน้าที่ต่างกันของหมวดข้อสอบก็ได้ หลักการเดียวกันนี้สามารถ ขยายผลไปศึกษาการทำหน้าที่ต่างกันของแบบทดสอบได้ด้วย การตรวจสอบการทำหน้าที่ต่างกัน ของข้อสอบ หมวดข้อสอบ และแบบทดสอบ ช่วยให้การพัฒนาแบบทดสอบทำได้อย่างมี ประสิทธิภาพและทำให้แบบทดสอบที่สร้างขึ้นมีความเที่ยงตรง เชิง โครงสร้างสูงขึ้น

ซึ่งเมื่อพิจารณาค่าความเชื่อมั่น โดยการวิเคราะห์ห้วิเคราะห์ $G - Coefficient$ ของ แบบทดสอบหลังจากตัดข้อสอบที่มีการทำหน้าที่แตกต่างกันออกแล้วเปรียบเทียบกับที่ยังไม่ได้ตัด ออก พบว่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งนี้เนื่องจากแต่ละรายวิชาพบข้อสอบที่ทำ หน้าที่ต่างกันน้อยในขณะที่ข้อสอบในแต่ละข้อให้น้ำหนักองค์ประกอบน้อยด้วยเช่นกันรวมทั้ง ขั้นตอนในการคัดเลือกข้อสอบผู้วิจัยมุ่งคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกสูงและจากการศึกษา พบว่าข้อสอบที่พบการทำหน้าที่ต่างกันจะเป็นข้อสอบที่มีลักษณะของค่าอำนาจจำแนกต่ำ

ข้อเสนอแนะสำหรับนำผลการวิจัยไปใช้

1. สำหรับการวิเคราะห์ความเที่ยงตรงเชิงโครงสร้างที่ลักษณะของการวัดความสามารถเฉพาะและความสามารถทั่วไปควรเลือกใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายใน ซึ่งจะให้โมเดลที่สอดคล้องกับข้อมูลเชิงประจักษ์มากกว่า

2. ในการเลือกใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในควรมีกุ่มตัวอย่างประมาณ 800 คนขึ้นไปจะสามารถประมาณค่าได้อย่างสมบูรณ์

3. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในจะสามารถทำได้ต่อจากการวิเคราะห์องค์ประกอบเชิงยืนยันให้ประโยชน์ในแง่ความสะดวก ประหยัดและสามารถตรวจสอบคุณภาพในด้านความเที่ยงตรงของเครื่องมือได้ดียิ่งขึ้น

4. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในจะสามารถทำได้โดยไม่ต้องใช้กลุ่มตัวอย่างขนาดใหญ่เนื่องจากโมเดลมีแนวโน้มไม่แปรเปลี่ยนตามขนาดกลุ่มตัวอย่าง

ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

1. ควรมีการศึกษาการวิเคราะห์องค์ประกอบ โดยใช้วิธีวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในเมื่อมีการฝ่าฝืนกฎในการวิเคราะห์องค์ประกอบเพื่อคุณลักษณะการแจกแจงค่าดัชนีวัดความสอดคล้องเช่นการแจกแจงปกติ การแปรเปลี่ยนน้ำหนักองค์ประกอบ เป็นต้น

2. ควรมีการศึกษาการวิเคราะห์องค์ประกอบ โดยใช้วิธีวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในเมื่อใช้แบบทดสอบที่มีอิทธิพลต่อค่าองค์ประกอบทั่วไปเพื่อเปรียบเทียบน้ำหนักองค์ประกอบของแบบสอบต่างชนิดกัน เช่นแบบสอบวัดผลสัมฤทธิ์ทางการเรียน แบบทดสอบวัดความถนัดทางการเรียน หรือแบบทดสอบทางจิตวิทยา เป็นต้น

3. ควรมีการศึกษาเปรียบเทียบเพื่อดูประสิทธิภาพของเทคนิคการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในกับเทคนิคการวิเคราะห์การทำหน้าที่ต่างกันวิธีอื่น ๆ

4. ควรมีการศึกษาเพื่อหาเกณฑ์ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเมื่อเทคนิคการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยันแฝงภายในเนื่องจากเกณฑ์เดิมที่นำมาใช้ไม่ได้มีการแยกน้ำหนักของอิทธิพลตรงไปยังความสามารถทั่วไปซึ่งอาจจะทำให้เกณฑ์ที่ใช้มีค่าสูงกว่าความเป็นจริง