

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์ 3 ประการ คือ ประการแรกเพื่อวิเคราะห์ค่าความยากของข้อสอบและค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบแบบพหุระดับ ในเงื่อนไขของจำนวนผู้สอบและความยาวของแบบสอบแบบต่างๆ ประการที่สองเพื่อวิเคราะห์ความสัมพันธ์ระหว่างค่าความยากของข้อสอบและค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ ประการที่สามเพื่อวิเคราะห์ความแปรปรวนของค่าความยากของข้อสอบ ระหว่างโรงเรียนภายในสังกัด และระหว่างสังกัดรวมทั้งศึกษาความแปรปรวนของค่าความสามารถของผู้สอบ ระหว่างผู้เรียนภายในโรงเรียนระหว่างโรงเรียนภายในสังกัด และระหว่างสังกัด

กรอบแนวคิดสมมติฐานการวิจัย ผู้วิจัยกำหนดสมมติฐานการวิจัยในครั้งนี้ไว้ 2 ข้อ ดังนี้ ข้อแรกเป็นสมมติฐานเกี่ยวกับความสัมพันธ์ของค่าความยากของข้อสอบและค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบน่าจะมีความสัมพันธ์กันในระดับสูงกับค่าความยากและค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบแบบพหุระดับ สมมติฐานการวิจัยข้อสองเกี่ยวกับความแปรปรวนของค่าความสามารถของผู้สอบ ระหว่างผู้เรียน โรงเรียน และสังกัด ซึ่งน่าจะมีความแปรปรวนในระดับผู้เรียนสูงสุด ระดับโรงเรียนรองลงมา และระดับสังกัดน้อยที่สุด

ประชากรที่ใช้ในการวิจัยเป็นผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 3 ช่วงชั้นที่ 3 (ม.1-ม.3) โรงเรียนในเขตกรุงเทพมหานคร ปีการศึกษา 2549 จำนวน 833,414 คน กลุ่มตัวอย่างที่ใช้ในการวิจัยในครั้งนี้ได้มาจากการสุ่มแบบหลายขั้นตอน เป็นผู้เรียนจากโรงเรียนในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) สังกัดสำนักงานการศึกษากรุงเทพมหานคร (กทม.) และสังกัดคณะกรรมการการศึกษาเอกชน (สช.) สังกัดละ 8 โรงเรียน จากโรงเรียนขนาดใหญ่พิเศษ ขนาดใหญ่ ขนาดกลาง และขนาดเล็ก จำนวนทั้งสิ้น 24 โรงเรียน โดยเก็บข้อมูลโรงเรียนละ 1 ห้องเรียน รวมกลุ่มตัวอย่างที่ใช้ในการวิจัยจำนวนทั้งสิ้น 1,002 คน

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบวัดผลสัมฤทธิ์ทางคณิตศาสตร์ ในช่วงชั้นที่ 3 (ม.3) ซึ่งผู้วิจัยดัดแปลงมาจาก อวยพร เรืองตระกูล (2544) เป็นแบบสอบเลือกตอบ 4 ตัวเลือก จำนวนข้อสอบ 40 ข้อ (20 ข้อหลัก 40 ข้อย่อย) ผ่านการตรวจสอบคุณภาพด้านความตรง คุณภาพรายข้อด้านความยากและค่าอำนาจจำแนกและคุณภาพแบบสอบทั้งฉบับด้านความเที่ยงซึ่งพบว่าข้อสอบในทุกข้อมีความยากง่ายปานกลางและสามารถจำแนกคนเก่งและอ่อนออกจากกันได้ดี รวมถึงความเที่ยงสูง โดยแบบสอบความยาว 20 ข้อ มีค่าความยากง่ายอยู่ระหว่าง 0.34-0.76 และมีค่าอำนาจจำแนกอยู่ระหว่าง 0.30-0.80 โดยข้อสอบทั้งฉบับมีค่าความยากเฉลี่ย 0.653 และค่าอำนาจจำแนกเฉลี่ย 0.652 ส่วนแบบสอบความยาว 40 ข้อ มีค่าความยากง่ายอยู่ระหว่าง 0.43-0.77 และมีค่าอำนาจจำแนกอยู่ระหว่าง 0.32-0.76 โดยข้อสอบทั้งฉบับมีค่าความยากเฉลี่ย 0.590 และค่าอำนาจจำแนกเฉลี่ย 0.585 ในส่วนของคุณภาพของแบบสอบทั้งฉบับด้านความเที่ยงโดยใช้สูตรสัมประสิทธิ์แอลฟาของครอนบาค พบว่ามีค่าความเที่ยงเท่ากับ 0.905 และ 0.921 ตามลำดับ ส่วนการตรวจสอบความเป็นเอกมิติของแบบสอบ พบว่าแบบสอบทั้ง 2 มีความเป็นเอกมิติของแบบสอบตามข้อตกลงเบื้องต้นของการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ

การศึกษาในการวิจัยครั้งนี้แบ่งเป็น 2 ขั้นตอน ขั้นแรกเป็นขั้นการศึกษาความสัมพันธ์ของค่าความยากของข้อสอบและค่าความสามารถของผู้สอบ ข้อมูลเป็นผลการสอบจากกลุ่มตัวอย่างจากเงื่อนไขความยาวของแบบสอบที่ศึกษาในงานวิจัยมี 2 ลักษณะคือ 40 ข้อ และ 20 ข้อ (1ข้อหลัก แยกเป็น 2 ข้อย่อย) ผู้วิจัยจะแบ่งเป็นผลการสอบของแบบสอบความยาว 40 ข้อ และแบบสอบความยาว 20 ข้อ โดยใช้วิธีการสุ่มอย่างง่ายข้อย่อยจำนวน 1 ข้อ จากข้อย่อย 2ข้อในแต่ละข้อหลัก จากนั้นผู้วิจัยใช้วิธีสุ่มอย่างง่ายตามเงื่อนไขจำนวนผู้สอบ คือ 100, 200, 500 และ 1,000 คน โดยเริ่มต้นสุ่มจำนวน 100 คน จากนั้นใช้การสุ่มอย่างง่ายเพิ่มจำนวนให้ครบ 200 คน และ 500 คน และ 1,000 คน และ จึงรวมเป็นเงื่อนไขที่ศึกษา คือ ทฤษฎีการวิเคราะห์ข้อสอบ (CTT, IRT และ HLM) ความยาวแบบสอบ (20 และ 40 ข้อ) และจำนวนผู้สอบ (100, 200, 500 และ 1,000 คน) ทั้งสิ้น 24 กรณี ขั้นที่สองเป็นการศึกษาความแปรปรวนของค่าความยากของข้อสอบระหว่างโรงเรียนภายในสังกัด และระหว่างสังกัด และศึกษาความแปรปรวนของค่าความสามารถของผู้สอบระหว่างผู้เรียน โรงเรียน และสังกัด ในส่วนนี้จะศึกษาโดยใช้ข้อมูลผลการสอบทั้งหมดของกลุ่มตัวอย่างซึ่งเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 จาก 3 สังกัด 24 โรงเรียน จำนวนทั้งสิ้น 1,002 คน จากแบบสอบความยาว 40 ข้อ

การวิเคราะห์ข้อมูลสำหรับขั้นแรกการประมาณค่าความยากของข้อสอบและค่าความสามารถของผู้สอบตามทฤษฎีการทดสอบแบบดั้งเดิม โดยใช้โปรแกรม TAP 6.65 สำหรับการ

ประมาณค่าความยากของข้อสอบและค่าความสามารถของผู้สอบตามทฤษฎีการตอบสนองข้อสอบ โดยใช้โปรแกรม BILOG MG3 และสำหรับการประมาณค่าความยากของข้อสอบและค่าความสามารถของผู้สอบตามการวิเคราะห์ข้อสอบแบบพหุระดับโดยใช้โปรแกรม HLM 6.03 วิเคราะห์พหุระดับ 2 ระดับ คือ ระดับข้อสอบ และระดับผู้สอบ ส่วนการวิเคราะห์ความสัมพันธ์ของค่าความยากของข้อสอบและค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ทั้ง 3 แบบ ใช้วิธีการหาค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน ในแต่ละเงื่อนไขจำนวนผู้สอบ และความยาวแบบสอบ โดยใช้โปรแกรม SPSS for Windows สำหรับขั้นที่สองการวิเคราะห์ความแปรปรวนของค่าความยากของข้อสอบรายข้อ จำนวน 40 ข้อ ระหว่างโรงเรียน และสังกัด ใช้ค่าความยากข้อสอบที่ได้จากการวิเคราะห์ข้อสอบแบบพหุระดับ มาวิเคราะห์ด้วยโมเดลศูนย์โดยการวิเคราะห์แบบพหุระดับด้วยโปรแกรม HLM 6.03 ส่วนการวิเคราะห์ความแปรปรวนของค่าความสามารถของผู้สอบระหว่างผู้เรียนภายในโรงเรียนเดียวกัน ระหว่างโรงเรียน และสังกัด โดยนำค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบแบบพหุระดับ มาวิเคราะห์โมเดลศูนย์โดยการวิเคราะห์แบบพหุระดับในทุกจำนวนผู้สอบเช่นกัน

การนำเสนอผลสรุปการวิจัยแบ่งเป็น 4 ตอน ตอนที่ 1 ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่างและคะแนนจากแบบสอบ ตอนที่ 2 ผลการวิเคราะห์ค่าสถิติพื้นฐานของค่าความยากและความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิม ทฤษฎีการตอบสนองข้อสอบ และการวิเคราะห์ข้อสอบแบบพหุระดับ ตอนที่ 3 ผลการวิเคราะห์ค่าสหสัมพันธ์ของค่าความยากของข้อสอบและค่าความสามารถของผู้สอบ ตอนที่ 4 ผลการศึกษาความแปรปรวนของค่าความยากของข้อสอบและความแปรปรวนของค่าความสามารถของผู้สอบ

สรุปผลการวิจัย

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่างและคะแนนจากแบบสอบ

จากข้อมูลทั้งหมดของกลุ่มตัวอย่างเมื่อนำมาวิเคราะห์จำนวนและร้อยละของกลุ่มตัวอย่าง พบว่าจากโรงเรียนทั้งหมด 24 โรงเรียนมีผู้เรียนที่เป็นกลุ่มตัวอย่างจากโรงเรียนสังกัดคณะกรรมการการศึกษาขั้นพื้นฐาน ร้อยละ 36.93 จากโรงเรียนสังกัดสำนักงานการศึกษา กรุงเทพมหานคร ร้อยละ 34.03 และผู้เรียนจากโรงเรียนสังกัดคณะกรรมการการศึกษาเอกชน ร้อยละ 29.04 จะสังเกตว่าสัดส่วนของผู้เรียนจากทั้งสามสังกัดมีความใกล้เคียงกัน ส่วนลักษณะการกระจายของคะแนนจากแบบสอบความยาว 40 ข้อ และ 20 ข้อ พบว่าการกระจายของคะแนนจากแบบสอบ

ความยาว 40 ข้อ ทุกจำนวนผู้สอบมีลักษณะเบ้ซ้าย เมื่อจำนวนผู้สอบยิ่งมากขึ้นลักษณะการแจกแจงของคะแนนเข้าใกล้โค้งปกติและแบนกว่าโค้งปกติทุกจำนวนผู้สอบโดยผู้สอบจำนวน 200 คน มีลักษณะเบ้ซ้ายและแบนมากที่สุดและสำหรับการกระจายของคะแนนจากแบบสอบความยาว 20 ข้อ ทุกจำนวนผู้สอบ พบว่ามีลักษณะเบ้ซ้ายโดยขนาดตัวอย่างยิ่งมากขึ้นลักษณะการแจกแจงข้อมูลเข้าใกล้โค้งปกติ เช่นเดียวกับกระจายของคะแนนจากแบบสอบความยาว 40 ข้อ และทุกขนาดแบนกว่าโค้งปกติ ยกเว้นจำนวนผู้สอบ 100 คน ซึ่งเบ้ซ้ายมากที่สุดและมีการกระจายสูงกว่าโค้งปกติ เมื่อพิจารณาค่าความคลาดเคลื่อนของค่าความเบ้และค่าความโด่งของทุกจำนวนผู้สอบ และความยาวแบบสอบ พบว่ามีค่าไม่เกิน 2.00 จึงถือว่าคะแนนจากการวัดในแต่ละจำนวนผู้สอบและความยาวแบบสอบมีลักษณะการแจกแจงเป็นโค้งปกติ

ตอนที่ 2 ผลการวิเคราะห์ค่าสถิติพื้นฐานของค่าความยากและความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิม ทฤษฎีการตอบสนองข้อสอบและการวิเคราะห์ข้อสอบแบบพหุระดับ

ในการวิจัยครั้งนี้ผู้วิจัยใช้การวิเคราะห์ข้อสอบแบบพหุระดับ (ระดับข้อสอบและระดับผู้สอบ) ผลการวิเคราะห์คุณภาพข้อสอบซึ่งทราบเพียงค่าความยากของข้อสอบ ผลการวิเคราะห์ พบว่าสำหรับแบบสอบความยาว 40 ข้อ ค่าความยากของข้อสอบเฉลี่ยมีค่าเพิ่มสูงขึ้นและค่าความคลาดเคลื่อนลดลงเข้าใกล้ศูนย์ เมื่อจำนวนผู้สอบมีจำนวนมากขึ้น สำหรับแบบสอบความยาว 20 ข้อ พบว่าค่าความยากของข้อสอบรายข้อของจำนวนผู้สอบ 200, 500 และ 1,000 คน มีค่าค่อนข้างใกล้เคียงกัน ส่วนจำนวนผู้สอบ 100 คน มีค่าความยากเฉลี่ยที่ต่างจากกลุ่มอื่นอย่างชัดเจน โดยค่าเฉลี่ยความยากของแบบสอบที่ได้จากการวิเคราะห์ข้อสอบแบบพหุระดับจะมีค่าเพิ่มขึ้นและใกล้เคียงกับค่าความยากของข้อสอบที่ได้จากการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบมากขึ้น เมื่อจำนวนผู้สอบมากขึ้น

ส่วนการประมาณค่าความสามารถของผู้สอบ สำหรับแบบสอบความยาว 40 ข้อ พบว่าการประมาณค่าความสามารถของผู้สอบด้วยวิธี empirical Bayes estimate มีค่าเฉลี่ยความสามารถของผู้สอบทุกจำนวนผู้สอบเท่ากัน คือ 0.000 ซึ่งการประมาณค่าความสามารถของผู้สอบด้วยวิธี ordinary least square estimate มีค่าเฉลี่ยความสามารถของผู้สอบทุกจำนวนผู้สอบอยู่ระหว่าง 0.040 ถึง 0.210 ผู้สอบจำนวน 100 คน มีค่าเฉลี่ยความสามารถสูงสุด และผู้สอบจำนวน 200 คน มีความแปรปรวนของค่าความสามารถของผู้สอบมากที่สุด โดยค่าเฉลี่ยความสามารถของผู้สอบจะมีค่าใกล้เคียงค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบมากขึ้น

เมื่อจำนวนผู้สอบมีจำนวนมากขึ้น สำหรับแบบสอบความยาว 20 ข้อ พบว่าการประมาณค่าความสามารถของผู้สอบที่ประมาณด้วยวิธี empirical Bayes estimate มีค่าเฉลี่ยความสามารถของผู้สอบทุกจำนวนผู้สอบเป็น 0.000 ซึ่งการประมาณค่าความสามารถของผู้สอบด้วยวิธี ordinary least square estimate มีค่าเฉลี่ยความสามารถของผู้สอบทุกจำนวนผู้สอบอยู่ระหว่าง 0.069 ถึง 0.187 ผู้สอบจำนวนผู้สอบ 100 คน มีค่าเฉลี่ยความสามารถมากที่สุด รองลงมาเป็นผู้สอบจำนวน 500, 1,000 และ 200 คน ตามลำดับ ส่วนผู้สอบจำนวน 1,000 คน มีความแปรปรวนมากที่สุด โดยค่าความสามารถของผู้สอบเฉลี่ยของผู้สอบจำนวน 100 คน มีค่าค่อนข้างต่างจากผู้สอบจำนวน 200, 500 และ 1,000 คน ซึ่งมีค่าที่ใกล้เคียงกัน

เมื่อพิจารณาเงื่อนไขของความยาวแบบสอบ พบว่า เมื่อแบบสอบมีความยาวเพิ่มขึ้นจะส่งผลให้ค่าความยากของข้อสอบและค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบแบบพหุระดับมีค่าใกล้เคียงกับค่าความยากที่ได้จากการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบมากขึ้นในทุกจำนวนกลุ่มผู้สอบ

ตอนที่ 3 ผลการวิเคราะห์ค่าสหสัมพันธ์ของค่าความยากของข้อสอบและค่าความสามารถของผู้สอบ

ผลการวิเคราะห์ค่าสหสัมพันธ์ของการวิเคราะห์ค่าความยากของข้อสอบจากทั้ง 3 แบบ พบว่าค่าความยากของข้อสอบระหว่างทฤษฎีการทดสอบแบบดั้งเดิมกับทฤษฎีการตอบสนองข้อสอบ สำหรับแบบสอบความยาว 20 ข้อ มีความสัมพันธ์กันทางลบอย่างมีนัยสำคัญที่ระดับ .01 ในระดับสูงทุกจำนวนผู้สอบ ยกเว้นผู้สอบจำนวน 100 คน ที่มีความสัมพันธ์อย่างไม่มีนัยสำคัญทางสถิติ แต่สำหรับแบบสอบความยาว 40 ข้อ ค่าความยากของข้อสอบความสัมพันธ์กันทางลบอย่างมีนัยสำคัญที่ระดับ .01 ทุกจำนวนผู้สอบ ส่วนค่าความยากของข้อสอบระหว่างทฤษฎีการทดสอบแบบดั้งเดิมกับการวิเคราะห์ข้อสอบแบบพหุระดับ สำหรับแบบสอบความยาว 20 ข้อ พบว่ามีความสัมพันธ์กันทางลบอย่างมีนัยสำคัญที่ระดับ 0.01 และมีความสัมพันธ์กันในระดับสูงทุกจำนวนผู้สอบ ยกเว้นผู้สอบจำนวน 100 คน ที่มีความสัมพันธ์อย่างไม่มีนัยสำคัญทางสถิติ สำหรับแบบสอบความยาว 40 ข้อ พบว่ามีความสัมพันธ์กันทางลบอย่างมีนัยสำคัญที่ระดับ 0.01 และมีความสัมพันธ์กันในระดับสูง ทุกจำนวนผู้สอบ เช่นเดียวกับความสัมพันธ์ของค่าความยากของข้อสอบระหว่างทฤษฎีการทดสอบแบบดั้งเดิมกับทฤษฎีการตอบสนองข้อสอบ ส่วนความสัมพันธ์ของค่าความยากของข้อสอบระหว่างทฤษฎีการตอบสนองข้อสอบกับการวิเคราะห์ข้อสอบแบบพหุระดับ สำหรับแบบสอบความยาว 20 ข้อ พบว่า มี

ความสัมพันธ์กันทางบวกอย่างมีนัยสำคัญที่ระดับ 0.01 และมีความสัมพันธ์กันในระดับสูง ทุกจำนวนผู้สอบ ยกเว้นผู้สอบจำนวน 100 คน ที่มีความสัมพันธ์อย่างไม่มีนัยสำคัญทางสถิติ สำหรับแบบสอบความยาว 40 ข้อ พบว่า มีความสัมพันธ์กันทางบวกอย่างมีนัยสำคัญที่ระดับ 0.01 และมีความสัมพันธ์กันในระดับสูงทุกจำนวนผู้สอบ

ผลการวิเคราะห์ความสัมพันธ์ระหว่างค่าความสามารถของผู้สอบที่ได้จากทั้ง 3 แบบ พบว่าทั้ง 3 แบบมีความสัมพันธ์กันทางบวกอย่างมีนัยสำคัญทางสถิติในระดับสูงทุกจำนวนผู้สอบ ยกเว้นกรณีของแบบสอบความยาว 20 ข้อ จากผู้สอบจำนวน 200 และ 100 คน ที่ความสัมพันธ์ของค่าความสามารถของผู้สอบระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการทดสอบแบบดั้งเดิมมีความสัมพันธ์ในระดับปานกลาง และค่าสัมประสิทธิ์สหสัมพันธ์จะลดลงเมื่อกลุ่มตัวอย่างมีขนาดเล็กลง สำหรับความสัมพันธ์ของค่าความสามารถของผู้สอบระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการตอบสนองข้อสอบ และระหว่างทฤษฎีการทดสอบแบบดั้งเดิมกับทฤษฎีการตอบสนองข้อสอบ มีค่าสัมประสิทธิ์สหสัมพันธ์ที่ใกล้เคียงกันในทุกจำนวนผู้สอบ และมีค่าลดลงเล็กน้อยเมื่อกลุ่มตัวอย่างกลุ่มอย่างมีขนาด 100 คน ทั้งแบบสอบความยาว 20 และ 40 ข้อ เมื่อในเงื่อนไขของความยาวแบบสอบ พบว่าสำหรับความสัมพันธ์ของค่าความสามารถของผู้สอบระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการทดสอบแบบดั้งเดิม ในกรณีผู้สอบจำนวน 100 และ 200 คน จากแบบสอบความยาว 20 ข้อ นั้นมีความสัมพันธ์ทางบวกในระดับปานกลาง แต่เมื่อเพิ่มความยาวแบบสอบเป็น 40 ข้อ ทำให้ความสัมพันธ์ของค่าความสามารถของผู้สอบระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการทดสอบแบบดั้งเดิมมีความสัมพันธ์ในระดับสูง ซึ่งให้เห็นถึงการวิเคราะห์ข้อสอบแบบพหุระดับไม่เหมาะสำหรับการวิเคราะห์ข้อสอบในสถานการณ์ที่ความยาวแบบสอบสั้น (20 ข้อ) และผู้สอบจำนวนเล็ก (100 คน) แต่ความสัมพันธ์ของค่าความสามารถของผู้สอบระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการตอบสนองข้อสอบ และระหว่างทฤษฎีการตอบสนองข้อสอบกับทฤษฎีการทดสอบแบบดั้งเดิมจะมีค่าสัมประสิทธิ์สหสัมพันธ์ใกล้เคียงกันเมื่อเพิ่มความยาวแบบสอบในทุกจำนวนผู้สอบ

โดยภาพรวม ค่าความยากของข้อสอบที่ได้จากทั้ง 3 แบบ นั้นมีความสัมพันธ์กันสูงอย่างมีนัยสำคัญทางสถิติทั้งแบบสอบความยาว 40 และ 20 ข้อ ทุกจำนวนผู้สอบ ยกเว้นผู้สอบจำนวน 100 คน และแบบสอบความยาว 20 ข้อ ที่ความสัมพันธ์ของค่าความยากของข้อสอบทั้ง 3 แบบไม่มีนัยสำคัญทางสถิติ โดยความสัมพันธ์ของค่าความยากที่ได้จากการวิเคราะห์ข้อสอบระหว่างทฤษฎีการทดสอบแบบดั้งเดิมกับทฤษฎีการตอบสนองข้อสอบมีความสัมพันธ์ทางลบ ส่วนค่าความยากของ

ข้อสอบระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการทดสอบแบบดั้งเดิม และค่าความยากของข้อสอบระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการตอบสนองข้อสอบมีความสัมพันธ์ทางบวก สำหรับความสัมพันธ์ของค่าความสามารถของผู้สอบที่ได้จากทั้ง 3 แบบ นั้นมีความสัมพันธ์กันทางบวกในระดับสูงอย่างมีนัยสำคัญทางสถิติทุกจำนวนผู้สอบทั้งแบบสอบความยาว 40 และ 20 ข้อ ยกเว้นผู้สอบจำนวน 200 และ 100 คน และแบบสอบความยาว 20 ข้อ ที่มีความสัมพันธ์ระดับปานกลาง แต่มีความสัมพันธ์ในระดับสูงเมื่อเพิ่มความยาวแบบสอบเป็น 40 ข้อ

เมื่อพิจารณาเงื่อนไขจำนวนผู้สอบ พบว่าผู้สอบจำนวนเล็ก(100 คน)ส่งผลให้ระดับความสัมพันธ์ของค่าความยากของข้อสอบที่ได้จากทั้ง 3 แบบลดลง โดยเฉพาะกรณีจำนวนผู้สอบ 100 คน และแบบสอบความยาว 20 ข้อ ที่ความสัมพันธ์ของค่าความยากของข้อสอบไม่มีนัยสำคัญทางสถิติ และความสัมพันธ์ของค่าความสามารถของผู้สอบระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการทดสอบแบบดั้งเดิมมีความสัมพันธ์ในระดับปานกลาง ส่วนจำนวนผู้สอบที่มีจำนวนมากกว่า 500 คน มีความสัมพันธ์ระดับสูง แต่สำหรับกรณีความยาวแบบสอบ 40 ข้อ พบความสัมพันธ์ระดับสูงทุกจำนวนผู้สอบ แต่ค่าสัมประสิทธิ์สหสัมพันธ์จะลดลงเล็กน้อยเมื่อกลุ่มตัวอย่างน้อยกว่า 200 คน

เมื่อพิจารณาเงื่อนไขความยาวแบบสอบ พบว่าความยาวแบบสอบส่งผลต่อความสัมพันธ์ของค่าความยากและค่าความสามารถของผู้สอบโดยแบบสอบที่มีจำนวนข้อน้อย (20 ข้อ) และผู้สอบจำนวนน้อย (100 คน) จะส่งผลให้ความสัมพันธ์ของค่าความยากของข้อสอบที่ได้จากทั้ง 3 แบบไม่มีนัยสำคัญทางสถิติ และความสัมพันธ์ของค่าความสามารถของผู้สอบระหว่างการวิเคราะห์ข้อสอบแบบพหุระดับกับทฤษฎีการทดสอบแบบดั้งเดิมอยู่ในระดับปานกลาง หากเป็นแบบสอบที่มีจำนวนข้อมากขึ้น (40 ข้อ) ความสัมพันธ์ของค่าความยากของข้อสอบและค่าความสามารถของผู้สอบที่ได้จากทั้ง 3 แบบจะอยู่ในระดับสูง

ตอนที่ 4 ผลการศึกษาความแปรปรวนของค่าความยากของข้อสอบรายข้อและค่าความสามารถของผู้สอบ

ผลการศึกษาความแปรปรวนของค่าความยากของข้อสอบรายข้อ พบว่า ค่าเฉลี่ยค่าความยากของข้อสอบ 40 ข้อ เป็นค่าที่ไม่ต่างจากศูนย์ (เป็นศูนย์) โดยค่าเฉลี่ยค่าความยากของข้อสอบจำนวน 32 ข้อ จากข้อสอบทั้งหมด 40 ข้อ ไม่มีความผันแปรหรือความแตกต่างระหว่างสังกัด (ร้อยละ 80) และมีข้อสอบจำนวน 8 ข้อ มีความผันแปรหรือความแตกต่างกันระหว่างสังกัด (ร้อยละ 20) จาก

การเปรียบเทียบขนาดของส่วนประกอบความแปรปรวน พบว่ามีความแตกต่างระหว่างสังกัด (ร้อยละ 56 ถึง 100) มากกว่าความแตกต่างระหว่างโรงเรียนภายในสังกัด (ร้อยละ 0 ถึง 49)

ผลการศึกษาความแปรปรวนของค่าความสามารถของผู้สอบ พบว่าค่าเฉลี่ยความสามารถของผู้สอบทุกโรงเรียนเป็นค่าที่ไม่ต่างจากศูนย์ทางสถิติ เมื่อพิจารณาอิทธิพลสุ่ม พบว่าค่าเฉลี่ยความสามารถของผู้สอบมีความผันแปรระหว่างโรงเรียนทั้ง 4 จำนวนผู้สอบ และค่าเฉลี่ยความสามารถของผู้สอบไม่มีความผันแปรระหว่างสังกัด ในผู้สอบจำนวน 100, 200 และ 500 ยกเว้นผู้สอบจำนวน 1,000 คน ที่ค่าเฉลี่ยความสามารถของผู้สอบมีผันแปรระหว่างสังกัด เมื่อพิจารณาเปรียบเทียบขนาดของส่วนประกอบความแปรปรวนพบว่า สำหรับผู้สอบจำนวน 100, 200 และ 500 คน มีความแตกต่างระหว่างผู้สอบภายในโรงเรียนมากที่สุด รองลงมาคือความแตกต่างระหว่างโรงเรียนภายในสังกัด และน้อยที่สุดเป็นความแตกต่างระหว่างสังกัด แต่สำหรับผู้สอบจำนวน 1,000 คน มีความแตกต่างระหว่างผู้สอบภายในโรงเรียนมากที่สุด รองลงมาคือความแตกต่างระหว่างสังกัด และน้อยที่สุดเป็นความแตกต่างระหว่างโรงเรียนภายในสังกัด

อภิปรายผลการวิจัย

การอภิปรายผลการวิจัยนี้ นำเสนอ 3 ประเด็น ประเด็นแรกเป็นการอภิปรายเกี่ยวกับผลการวิเคราะห์ข้อสอบแบบพหุระดับ ประเด็นที่สองอภิปรายเกี่ยวกับผลการวิเคราะห์ความสัมพันธ์ของค่าความยากของข้อสอบและค่าความสามารถของผู้สอบระหว่างทั้ง 3 แบบ ประเด็นที่สามเป็นการอภิปรายเกี่ยวกับความแปรปรวนของความยากของข้อสอบและความสามารถของผู้สอบ ซึ่งมีรายละเอียดดังนี้

1. ผลการวิเคราะห์ข้อสอบแบบพหุระดับ

1.1 ในการศึกษาเงื่อนไขจำนวนผู้สอบและความยาวแบบสอบ พบว่า แนวโน้มของค่าความยากที่ได้จากการวิเคราะห์ข้อสอบแบบพหุระดับจะมีค่าใกล้เคียงกับค่าที่ได้จากการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ เมื่อผู้สอบมีจำนวนมากและความยาวแบบสอบปกติ (40 ข้อ) ซึ่งผลการวิจัยนี้สอดคล้องกับผลการวิจัยของ Kamata (2001), Stone และ Yumoto (2004) โดยสังเกตจากค่าความคลาดเคลื่อนมาตรฐานของค่าความยากจะลดลงเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น สำหรับการวิเคราะห์ข้อสอบแบบพหุระดับซึ่งสามารถประมาณค่าพารามิเตอร์ได้ด้วยวิธี empirical Bayes estimate และวิธี ordinary least square estimate ซึ่งวิธีดังกล่าวจะคล้ายกับวิธี Maximum Likelihood estimate เช่นเดียวกับการประมาณค่าในโปรแกรม Bilog ตามทฤษฎีการตอบสนองข้อสอบโดยมีกระบวนการประมาณค่าพารามิเตอร์โดยคำนวณแบบทวนซ้ำๆ จนกว่าจะได้ค่าคงที่

(Iterative) ซึ่งจะหาค่าที่ค่อนข้างถูกต้อง เมื่อจำนวนกลุ่มผู้สอบและความยาวแบบสอบเพิ่มขึ้น (Hambleton and Swaminathan, 1985: 88-89) และจากคำกล่าวของ Stone และ Yumoto (2004) ที่ว่าการประมาณค่าพารามิเตอร์ด้วยโมเดล RASCH และ IRT ค่าที่ได้จะใกล้เคียงกันเมื่อกลุ่มตัวอย่างมีตั้งแต่ 100 ขึ้นไป

1.2 ผลการวิจัยพบว่าค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี empirical Bayes estimate ทุกจำนวนผู้สอบทั้งแบบสอบความยาว 40 และ 20 ข้อ มีค่าเฉลี่ยเป็น 0 เนื่องจากค่าความสามารถของผู้สอบแต่ละคนที่ประมาณค่าได้ในระดับที่ 2 (ระดับบุคคล) มีการแจกแจงแบบปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ τ สอดคล้องกับแนวคิดของ kamata (2001) ส่วนค่าความสามารถที่ประมาณด้วยวิธี ordinary least square estimate มีค่าเฉลี่ยใกล้เคียง 0 แสดงว่าการแจกแจงค่าความสามารถใกล้เคียงโค้งปกติ

2. ผลการวิเคราะห์ความสัมพันธ์ของค่าความยากของข้อสอบและค่าความสามารถของผู้สอบระหว่าง 3 แบบ

2.1 จากผลการวิจัยพบว่าความสัมพันธ์ระหว่างค่าความยากของข้อสอบที่ได้จากการวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมกับทฤษฎีการตอบสนองข้อสอบ พบว่า มีความสัมพันธ์กันทางลบในระดับสูง ซึ่งผลการวิจัยในส่วนนี้สอดคล้องตามสมมติฐานการวิจัยข้อที่ 1 และผลการวิจัยของ Stage (1998) แต่ไม่สอดคล้องกับผลการวิจัยของเบญจพร ยนต์จักรวิณี, Fan (1998), MacDonald และ Paunonen (2002), Ndlichako และ Roger (1997) และ ชยุดม ภิรมสมบัติ (2547) ซึ่งผลการวิจัยรายงานค่าสหสัมพันธ์เป็นบวก อาจมาเนื่องจากเป้าหมายของการศึกษาเป็นเพียงการศึกษาถึงขนาดของความสัมพันธ์มิได้สนใจทิศทางของความสัมพันธ์ ซึ่งโดยทฤษฎีแล้วค่าความยากของข้อสอบตามทฤษฎีการทดสอบดั้งเดิมมีค่าในช่วง 0 ถึง 1 หากค่าเข้าใกล้ 0 แสดงว่าเป็นข้อสอบที่ยาก ส่วนค่าความยากตามทฤษฎีการตอบสนองข้อสอบมีค่าในช่วง -2.5 ถึง +2.5 (ในทางปฏิบัติ) หากค่าเข้าใกล้ -2.5 แสดงว่าเป็นข้อสอบที่ง่าย และหากมีค่าเข้าใกล้ 2.5 แสดงว่าเป็นข้อที่ยาก จึงมีทำให้ค่าความยากของข้อสอบที่ได้จากทั้ง 2 ทฤษฎีมีความสัมพันธ์ในทางลบซึ่งสอดคล้องกับผลการวิจัยของ Stage (1998)

2.2 จากผลการวิจัยที่พบว่าค่าความยากของข้อสอบที่ได้จากการวิเคราะห์ข้อสอบแบบพหุระดับกับค่าความยากของข้อสอบตามทฤษฎีการตอบสนองข้อสอบมีความสัมพันธ์กันในทางบวก ระดับสูง ผลการวิจัยในส่วนนี้เป็นไปตามสมมติฐานการวิจัยที่ตั้งไว้และสอดคล้องกับแนวคิดของ kamata (2001) ประกอบกับความสัมพันธ์ของค่าความยากของข้อสอบระหว่างการวิเคราะห์ข้อสอบ

แบบพระดัดกับทฤษฎีการทดสอบแบบดั้งเดิมมีความสัมพันธ์ทางลบในระดับสูง ดังนั้นการแปลความหมายของค่าความยากของข้อสอบที่ได้จากการวิเคราะห์ข้อสอบแบบพระดัดจะมีลักษณะคล้ายกับค่าความยากตามทฤษฎีการตอบสนองข้อสอบ ซึ่งค่าความยากของข้อสอบตามการวิเคราะห์ข้อสอบแบบพระดัดจะมีค่าในช่วง $-\infty$ ถึง ∞ ดังนั้นหากข้อสอบมีค่าความยากมากแสดงว่าเป็นข้อสอบที่ยากและถ้าหากค่าน้อยแสดงว่าเป็นข้อสอบที่ง่าย

2.3 จากผลการวิจัยที่พบว่าผู้สอบจำนวน 100 คน และความยาวแบบสอบ 20 ข้อ มีค่าความยากของข้อสอบระหว่างทั้ง 3 แบบสัมพันธ์กันอย่างไม่มีนัยสำคัญทางสถิติ แต่ในกรณีผู้สอบจำนวน 100 คน และความยาวแบบสอบ 40 ข้อ ค่าความยากของข้อสอบระหว่างทั้ง 3 แบบมีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติ สาเหตุอาจเนื่องมาจากค่าความยากของข้อสอบมีความสัมพันธ์ในระดับต่ำประกอบกับจำนวนข้อที่น้อย (20 ข้อ) จึงทำให้ค่าสัมประสิทธิ์สหสัมพันธ์ไม่มีนัยสำคัญทางสถิติ แต่เมื่อจำนวนข้อสอบเพิ่มขึ้น (เป็น 40 ข้อ) ทำให้สามารถวิเคราะห์รูปแบบความสัมพันธ์ของค่าความยากของข้อสอบระหว่าง 2 ทฤษฎี ได้ชัดเจนขึ้น จึงทำให้ค่าสัมประสิทธิ์สหสัมพันธ์มีนัยสำคัญทางสถิติ ดังนั้นจึงสรุปได้ว่าเมื่อแบบสอบมีจำนวนข้อน้อยและจำนวนผู้สอบที่น้อยเกินไปจะส่งผลต่อความถูกต้องและความน่าเชื่อถือของค่าสถิติที่ได้จากการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบและการวิเคราะห์ข้อสอบแบบพระดัด

3. ความแปรปรวนของค่าความยากของข้อสอบรายข้อ

จากผลการวิเคราะห์ความแปรปรวนของค่าความยากของข้อสอบรายข้อระหว่างโรงเรียน และสังกัด พบว่า จำนวนข้อสอบ 32 ข้อ จากทั้งหมด 40 ข้อ คิดเป็นร้อยละ 80 ของข้อสอบที่ไม่มีความผันแปรระหว่างสังกัด ส่วนข้อสอบอีก 8 ข้อ (ร้อยละ 20) ค่าความยากของข้อสอบมีความผันแปรระหว่างสังกัด อาจกล่าวได้ว่าส่วนใหญ่แล้วค่าความยากของข้อสอบนั้นมีความคงที่ไม่แตกต่างระหว่างกลุ่มหรือเป็นค่าคงที่สำหรับผู้สอบทุกคน ซึ่งสอดคล้องกับแนวคิดของ kamata (2001) ส่วนข้อสอบจำนวน 8 ข้อ ซึ่งพบว่ามีค่าความผันแปรระหว่างกลุ่มนั้นอาจเป็น เพราะข้อสอบมีลักษณะของการทำหน้าที่ต่างกันของข้อสอบหรือความลำเอียงของข้อสอบ กล่าวคือ ข้อสอบบางข้ออาจเป็นข้อที่ได้เปรียบหรือเสียเปรียบของผู้สอบบางกลุ่ม ทำให้พบความผันแปรของค่าความยากระหว่างกลุ่ม ซึ่งแบบสอบที่ใช้เป็นเครื่องมือในการวิจัยครั้งนี้ได้ผ่านตรวจสอบความตรงโดยผู้เชี่ยวชาญหลายฝ่าย ทั้งทางวัดผลและการสอนคณิตศาสตร์ ก่อนการทดลองใช้ แต่มิได้ทำการตรวจสอบความไวของการทำหน้าที่ต่างกันของข้อสอบหลังการทดลองใช้ จึงไม่อาจสรุปได้อย่างชัดเจนว่าสาเหตุเป็นเช่นนี้เพราะการทำหน้าที่ต่างกันของข้อสอบบางข้อ หรือไม่

4. ความแปรปรวนของค่าความสามารถของผู้สอบ

จากผลการวิเคราะห์ความแปรปรวนของค่าความสามารถระหว่างผู้สอบ โรงเรียน และสังกัด ซึ่งพบว่าความผันแปรในระหว่างผู้เรียน และระหว่างโรงเรียนทุกจำนวนผู้สอบ แต่ในระหว่าง สังกัดพบความผันแปรเฉพาะผู้สอบจำนวน 1,000 คน อาจเนื่องจากความแตกต่างระหว่างผู้เรียน ระหว่างโรงเรียน หรือเป็นความแตกต่างที่ขึ้นกับค่าความยากและค่าอำนาจของข้อสอบที่ผู้วิจัยได้ ปรับให้มีค่าเท่ากันก่อนการประมาณค่าความสามารถของผู้สอบทั้ง 3 แบบ ซึ่งเป็นข้อจำกัดของการ วิจัยครั้งนี้ แต่ค่าความสามารถของผู้สอบไม่พบความผันแปรระหว่างสังกัดในผู้สอบจำนวน 100, 200 และ 500 คน ทั้งนี้เนื่องจากมาจากกลุ่มสังกัดหรือหน่วยการวิเคราะห์ระดับสังกัด (Under Level Unit) ขนาดเล็กเกินไป ($n=3$), (Hox, 2004) และจำนวนหน่วยย่อยในระดับผู้เรียน ในแต่ละโรงเรียนมีจำนวน น้อยเช่นกัน แต่สำหรับในผู้สอบจำนวน 1,000 คน เป็นผู้สอบที่มาจากทุกขนาดโรงเรียน และมีสัดส่วน จำนวนในแต่ละสังกัดที่ใกล้เคียงกัน ทำให้ในผู้สอบจำนวน 1,000 คน มีจำนวนหน่วยย่อยในระดับ ผู้เรียน ซึ่งอยู่ในแต่ละโรงเรียน แต่ละสังกัด โดยประมาณสังกัด 300 คน จึงทำให้ผลการวิเคราะห์พบ ความผันแปรในระดับสังกัดอย่างมีนัยสำคัญทางสถิติ โดยค่าความสามารถของผู้สอบมีความแตกต่าง ระหว่างผู้เรียนมากที่สุด (ประมาณร้อยละ 55) รองลงมาเป็นการแปรปรวนที่มาจากความแตกต่าง ระหว่างสังกัด (ประมาณร้อยละ 24) และความแปรปรวนที่มาจากความแตกต่างระหว่างโรงเรียนมี ขนาดน้อยที่สุด (ประมาณร้อยละ 20) สอดคล้องตามสมมติฐานการวิจัยข้อที่ 2 สามารถสะท้อนได้ถึง ผลสัมฤทธิ์ของการจัดการเรียนรู้กลุ่มสาระคณิตศาสตร์ของโรงเรียนในเขตกรุงเทพมหานคร ในปัจจุบัน ที่มีความไม่เสมอภาคเกิดขึ้นเป็นผลมาจากความแตกต่างระหว่างผู้เรียนมากที่สุด ความแตกต่าง ระหว่างโรงเรียน และความแตกต่างระหว่างสังกัด ซึ่งผลการวิจัยนี้มีทั้งสอดคล้องผลการวิจัยของอวย พร เรื่องตระกูล และสุวิมล ว่องวานิช (2547) ประกอบกับข้อมูลจากการสังเกตสภาพการทำข้อสอบของ ผู้เรียน จากการสัมภาษณ์ผู้เรียนและครูผู้สอนวิชาคณิตศาสตร์โดยผู้วิจัยเป็นผู้สัมภาษณ์พบว่าส่วน ใหญ่ผู้เรียนในสังกัด สพฐ. และ สช.สามารถทำข้อสอบได้มากกว่าผู้เรียนสังกัด กทม.โดยผู้เรียนใน โรงเรียนในสังกัด กทม. บอกว่า ค่อนข้างยาก ทำไม่ค่อยได้หลายข้อ และครูผู้สอนรายงานว่า เนื่องจาก พื้นฐานวิชาคณิตศาสตร์ของผู้เรียนในสังกัดกทม.ส่วนใหญ่ไม่ค่อยดี และมีผู้เรียนที่เก่งวิชาคณิตศาสตร์ จำนวนน้อย และผู้เรียนส่วนใหญ่มีเป้าหมายศึกษาต่อในสายวิชาชีพมากกว่าสายสามัญ ดังนั้นจึงเป็น เหตุให้ผู้เรียนไม่ค่อยให้ความสำคัญกับเรียนวิชาคณิตศาสตร์มากเท่ากับวิชาอื่นๆ

ข้อเสนอแนะ

ข้อเสนอแนะในเชิงนโยบาย

จากลักษณะของการวิเคราะห์ข้อสอบแบบพหุระดับที่สามารถวิเคราะห์ข้อมูลที่มีโครงสร้างเป็นหน่วยซ้อนเป็นระดับชั้นลดหลั่นกัน และผลการวิเคราะห์มีความถูกต้องและเชื่อถือได้มากกว่าการนำตัวแปรในระดับที่ต่างกันมาวิเคราะห์ในระดับเดียวกัน (ศิริชัย กาญจนวาสี, 2541) ดังนั้นในการสอบระดับชาติ ที่มีผู้สอบจำนวนมากมาจากหลากหลายแหล่ง อาทิมาจากหลายโรงเรียน หลายจังหวัด หรือเขตพื้นที่การศึกษา เช่น การทดสอบ O-NET หรือ NT ควรใช้เทคนิคการวิเคราะห์ข้อสอบแบบพหุระดับในการวิเคราะห์ข้อสอบมากกว่าการวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิม หรือทฤษฎีการตอบสนองข้อสอบ เพราะการวิเคราะห์ข้อสอบแบบพหุระดับนอกจากจะวิเคราะห์ค่าความยากและค่าความสามารถของผู้สอบแล้วยังสามารถวิเคราะห์ถึงคุณลักษณะของผู้สอบที่ต่างกันจะส่งผลให้คะแนนสอบมีความแตกต่างกันหรือไม่ กล่าวคือสามารถนำตัวแปรคุณลักษณะผู้สอบเข้ามาร่วมในการวิเคราะห์เพื่ออธิบายผลการสอบของผู้สอบว่าส่งผลแตกต่างกันหรือไม่อย่างไร อีกทั้งสามารถนำตัวแปรในระดับโรงเรียน จังหวัด หรือสังกัด เข้ามาร่วมพิจารณาไปพร้อมกับการวิเคราะห์ข้อสอบในครั้งเดียว

ข้อเสนอแนะในการนำไปใช้

1. ในระดับโรงเรียน ครูผู้สอนควรใช้การวิเคราะห์ข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิม มากกว่าทฤษฎีการตอบสนองข้อสอบและการวิเคราะห์ข้อสอบแบบพหุระดับ เนื่องจากจำนวนผู้สอบมีน้อย และวิธีการวิเคราะห์ง่าย สะดวกแก่การใช้งานในระดับโรงเรียน ซึ่งประสิทธิภาพของค่าสถิติมีความถูกต้องและเชื่อถือได้
2. ในระดับเขตพื้นที่การศึกษา นักวิชาการหรือนักวัดผลควรใช้การวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบหรือการวิเคราะห์ข้อสอบแบบพหุระดับมากกว่าการวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิม เนื่องจากจำนวนผู้สอบมีจำนวนมาก มาจากหลายโรงเรียน หลายสังกัด และผลการวิเคราะห์จะให้ค่าสถิติมีความถูกต้องและเชื่อถือได้มากกว่าเนื่องมีกระบวนการประมาณค่าด้วยวิธีทวนซ้ำ (Iterative) จนได้ค่าที่ใกล้เคียงค่าที่แท้จริงและมีความคงที่ไม่ผันแปรตามกลุ่มผู้สอบ

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. การวิจัยครั้งนี้เป็นการศึกษาโดยใช้ข้อมูลเชิงประจักษ์ในการวิเคราะห์ตามเงื่อนไขของความยาวแบบสอบและจำนวนผู้สอบ ดังนั้นในการวิจัยครั้งต่อไปควรออกแบบการวิจัยเพื่อศึกษา