



บทที่ 5

สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบประสิทธิผลของวิธีประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซิมัมไลค์ลิฮูด วิธีนิวริสติก และวิธีของเบย์นแบบสอบวัดผลสัมฤทธิ์ และ แบบสอบความถนัด โดยทำการเปรียบเทียบค่าฟังก์ชันสารสนเทศ และความตรงเชิงเกณฑ์สัมพัทธ์ของแบบสอบ ที่เป็นผลจากการประมาณค่าพารามิเตอร์ทั้ง 3 วิธี กลุ่มตัวอย่างประชากรที่ใช้ในการวิจัย ได้แก่ นักเรียนมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2531 สังกัดกรมสามัญศึกษา จังหวัดจันทบุรี จำนวน 1,028 คน ที่ลุ่มมาโดยการสุ่มแบบแบ่งชั้น และในแต่ละชั้น ทำการสุ่มอย่างง่ายโดยใช้ห้องเรียนเป็นหน่วยการสุ่ม จากประชากรนักเรียน 50 ห้อง จำนวนทั้งสิ้น 2,125 คน

เครื่องมือที่ใช้ในการวิจัยเป็นแบบสอบ 2 ฉบับคือ แบบสอบวัดผลสัมฤทธิ์วิชาภาษาไทย (ด้านการใช้ภาษา) สำหรับชั้นมัธยมศึกษาปีที่ 3 แบบเลือกตอบ 5 ตัวเลือก ของสำนักทดสอบทางการศึกษาและจิตวิทยา มศว ประสานมิตร มีความเที่ยง (KR-21) เท่ากับ 0.7567 มีความตรงตามเนื้อเรื่อง และมีเกณฑ์ปกติปลายปีระดับประเทศ และแบบสอบความถนัดจำแนกด้านการใช้ภาษาไทย สำหรับชั้นมัธยมศึกษาปีที่ 3 แบบเลือกตอบ 5 ตัวเลือก ของ อัมพิกา สุริยินทร์ ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย มีค่าความเที่ยง (KR-20) เท่ากับ 0.863 มีค่าความตรงร่วมสมัย 0.503 และมีเกณฑ์ปกติเปอร์เซ็นต์ไทล์

ในการเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูลนั้น ผู้วิจัยได้นำแบบสอบทั้ง 2 ฉบับ ไปทดสอบกับกลุ่มตัวอย่าง ที่ใช้ทดลองสอบ เพื่อศึกษาค่าความเที่ยงของแบบสอบ ตามกรอบทฤษฎีดั้งเดิม เมื่อนำมาใช้ในกลุ่มตัวอย่างที่เปลี่ยนไป รวมทั้งศึกษาปัญหาต่าง ๆ

ที่จะนำมาแก้ไขในการดำเนินการสอบจริง จากนั้นจึงนำแบบสอบทั้ง 2 ฉบับ ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการวิจัย โดยผู้เข้าสอบทุกคนจะได้รับการทดสอบด้วยแบบสอบทั้ง 2 ฉบับ แล้วนำข้อมูลที่ได้นำมาวิเคราะห์ตัวประกอบด้วยโปรแกรม SPSS* เพื่อตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบ นอกจากนี้ยังทำการวิเคราะห์เพื่อหาค่าพารามิเตอร์ของข้อสอบและค่าความสามารถของผู้เข้าสอบ ค่าพารามิเตอร์ที่ถือว่าเป็นผลจากการประมาด้วยวิธีแมกซิมัมไลค์ลิตูด คำนวณได้จากโปรแกรม Logist 5 Version 2.5 ส่วนค่าพารามิเตอร์ที่เป็นผลจากการประมาด้วยวิธีชิวริสติก และวิธีของเบย์นั้น คำนวณได้จากโปรแกรมภาษาฟอร์แทรน หลังจากนั้นนำค่าพารามิเตอร์ของข้อสอบคือ ค่าความยาก ค่าอำนาจจำแนก และค่าการเดา ที่ได้จากวิธีประมาค่าแต่ละวิธีของแบบสอบแต่ละฉบับ มาคำนวณค่าฟังก์ชันสารสนเทศของแบบสอบ (test information function) และเขียนโค้งแสดงสารสนเทศของแบบสอบ (test information curve) แล้วนำค่าฟังก์ชันสารสนเทศแบบสอบไปคำนวณหาค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ ซึ่งในแบบสอบแต่ละฉบับ จะนำค่าฟังก์ชันสารสนเทศแบบสอบที่เป็นผลจากวิธีประมาค่าพารามิเตอร์ทั้ง 3 วิธี มาเปรียบเทียบกันครั้งละ 2 วิธี คือ วิธีของเบย์ เปรียบเทียบกับวิธีแมกซิมัมไลค์ลิตูด วิธีของเบย์ เปรียบเทียบกับวิธีชิวริสติก และวิธีแมกซิมัมไลค์ลิตูด เปรียบเทียบกับวิธีชิวริสติก ส่วนค่าพารามิเตอร์ความสามารถของผู้เข้าสอบแต่ละคน ที่ได้จากการประมาค่า 3 วิธี ในผลการตอบแบบสอบแต่ละฉบับนั้น ได้นำมาหาค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน กับคะแนนเกณฑ์ที่ได้จากผลการสอบปลายภาควิชาภาษาไทย ด้านการใช้ภาษาของผู้เข้าสอบแต่ละคน ซึ่งถือว่าเป็นความตรงเชิงเกณฑ์สัมพัทธ์ของแบบสอบ แล้วทำการเปรียบเทียบความตรงเชิงเกณฑ์สัมพัทธ์ของแบบสอบที่ได้จากวิธีประมาค่าแต่ละวิธี ว่าในแบบสอบแต่ละฉบับ วิธีประมาค่าพารามิเตอร์ความสามารถของผู้เข้าสอบวิธีใด ที่ทำให้แบบสอบมีความตรงเชิงเกณฑ์สัมพัทธ์สูงสุด การเปรียบเทียบทำได้โดยทดสอบนัยสำคัญของความแตกต่างระหว่างค่าสัมประสิทธิ์สหสัมพันธ์ที่เป็นผลจากการประมาด้วยวิธีการทั้ง 3 วิธี และเพื่อให้สามารถสรุปอ้างอิงสู่ค่าพารามิเตอร์ของประชากรได้ เช่นเดียวกับผลการวิจัยตามกรอบทฤษฎีการตอบสนองข้อสอบจึงแปลงค่าสัมประสิทธิ์สหสัมพันธ์เป็นค่าฟังก์ชันฟิชเชอร์ซี (Fisher Z : Zr) แล้วทดสอบรวมด้วยค่าไคสแควร์ (Chi square : χ^2) เมื่อพบความแตกต่างจึงนำมาทดสอบเป็นรายคู่ด้วยค่าสัมประสิทธิ์ฟิชเชอร์ซี (Z - test) ต่อไป

ผลการวิจัยสรุปได้ดังนี้

1. พิจารณาค่าประสิทธิภาพสัมพัทธ์ของแบบสอบวัดผลสัมฤทธิ์ที่เป็นผลจากการประมาณค่าพารามิเตอร์ของข้อสอบ 3 วิธี

1.1 แบบสอบวัดผลสัมฤทธิ์ ที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงกว่า ที่ประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด เฉพาะในกลุ่มตัวอย่างระดับความสามารถปานกลาง และกลุ่มตัวอย่างระดับความสามารถต่ำ และแบบสอบวัดผลสัมฤทธิ์ ที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงกว่า เมื่อได้รับการประมาณด้วยวิธีของเบย์ เฉพาะในกลุ่มตัวอย่างระดับความสามารถสูง

1.2 แบบสอบวัดผลสัมฤทธิ์ที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงกว่า ที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีซีวีริสติก ในกลุ่มตัวอย่างทุกระดับความสามารถ

1.3 แบบสอบวัดผลสัมฤทธิ์ที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงกว่าที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีซีวีริสติก ในกลุ่มตัวอย่างทุกระดับความสามารถ

1.4 เมื่อ เปรียบเทียบคุณภาพของแบบสอบวัดผลสัมฤทธิ์ที่เป็นผลจากการประมาณค่าพารามิเตอร์ทั้ง 3 วิธี พร้อมกับ สรุปได้ว่า แบบสอบที่ทำการประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงสุดในกลุ่มตัวอย่างระดับความสามารถสูง รองลงไปคือ แบบสอบที่ประมาณค่าด้วยวิธีของเบย์และวิธีของซีวีริสติก ตามลำดับ ส่วนแบบสอบที่ทำการประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์ มีคุณภาพสูงสุดในกลุ่มตัวอย่างระดับความสามารถปานกลาง และกลุ่มตัวอย่างระดับความสามารถต่ำ รองลงไปคือ แบบสอบที่ประมาณค่าด้วยวิธีแมกซิมัมไลค์ลิฮูด และวิธีซีวีริสติก ตามลำดับ

2. พิจารณาค่าประสิทธิภาพสัมพัทธ์ของแบบสอบความถนัดที่เป็นผลจากการประมาณค่าพารามิเตอร์ของข้อสอบ 3 วิธี

2.1 แบบสอบความถนัด ที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงกว่า ที่ประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด เฉพาะในกลุ่มตัวอย่างระดับความ

สามารถปานกลาง และกลุ่มตัวอย่างระดับความสามารถต่ำ และแบบสอบความถนัดที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงกว่า ที่ประมาณด้วยวิธีของเบย์ เฉพาะในกลุ่มตัวอย่างระดับความสามารถสูง

2.2 แบบสอบความถนัดที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงกว่า ที่ประมาณด้วยค่าพารามิเตอร์ของข้อสอบด้วยวิธีฮิวริสติก ในกลุ่มตัวอย่างทุกระดับความสามารถ

2.3 แบบสอบความถนัดที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงกว่า ที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีฮิวริสติก เฉพาะในกลุ่มตัวอย่างระดับความสามารถสูง และระดับความสามารถปานกลาง และแบบสอบความถนัดที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีฮิวริสติก มีคุณภาพสูงกว่าที่ประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด เฉพาะในกลุ่มตัวอย่างระดับความสามารถต่ำ

2.4 เมื่อเปรียบเทียบคุณภาพของแบบสอบความถนัดที่เป็นผลจากการประมาณค่าพารามิเตอร์ทั้ง 3 วิธีพร้อมกัน สรุปได้ว่า แบบสอบที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงสุดในกลุ่มตัวอย่างระดับความสามารถสูง รองลงไปคือ แบบสอบที่ประมาณค่าด้วยวิธีของเบย์และวิธีของฮิวริสติกตามลำดับ ส่วนแบบสอบที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงสุดในกลุ่มตัวอย่างระดับความสามารถปานกลาง รองลงไปคือ แบบสอบที่ประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด และวิธีฮิวริสติก ตามลำดับ สำหรับในกลุ่มตัวอย่างระดับความสามารถต่ำ แบบสอบที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงสุด รองลงไปคือ แบบสอบที่ประมาณค่าด้วยวิธีฮิวริสติก และวิธีแมกซิมัมไลค์ลิฮูด ตามลำดับ

3. พิจารณาค่าความตรงเชิง เกณฑ์สัมพันธ์ของแบบสอบ

3.1 แบบสอบวัดผลสัมฤทธิ์มีค่าความตรงร่วมสมัยที่เป็นผลจากความสัมพันธ์ระหว่างค่าพารามิเตอร์ความสามารถของผู้เข้าสอบที่ประมาณได้ทั้ง 3 วิธี กับคะแนนเกณฑ์ผลการสอบปลายภาควิชาภาษาไทยของผู้เข้าสอบ อยู่ระหว่าง 0.7281 ถึง 0.8081 ค่าความตรงดังกล่าวแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยที่ค่าความตรงร่วมสมัยของแบบสอบ ที่เป็นผลจากการประมาณด้วยวิธีของเบย์ มีค่าสูงกว่า ที่เป็นผลจาก

การประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด และวิธีอีวีริสติก อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ส่วนค่าความตรงร่วมสมัยของแบบสอบที่เป็นผลจากการประมาณด้วยวิธีอีวีริสติก และวิธีแมกซิมัมไลค์ลิฮูดไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.2 แบบสอบความถนัด มีค่าความตรงเชิงทำนายที่เป็นผลจากความสัมพันธ์ระหว่างค่าพารามิเตอร์ความสามารถของผู้เข้าสอบที่ประมาณได้ทั้ง 3 วิธี กับคะแนนเกณฑ์ที่เป็นผลการสอบปลายภาควิชาภาษาไทยของผู้เข้าสอบ อยู่ระหว่าง 0.5133 ถึง 0.9346 ค่าความตรงดังกล่าวแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยที่ค่าความตรงเชิงทำนายของแบบสอบที่เป็นผลจากการประมาณด้วยวิธีของเบย์ และวิธีอีวีริสติก มีค่าสูงกว่า ที่เป็นผลจากการประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ส่วนค่าความตรงเชิงทำนายของแบบสอบที่เป็นผลจากการประมาณด้วยวิธีของเบย์ มีค่าสูงกว่า ที่เป็นผลจากการประมาณด้วยวิธีอีวีริสติก อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

อภิปรายผลการวิจัย

1. ผลการวิจัยที่เกี่ยวกับประสิทธิผลของวิธีประมาณค่าพารามิเตอร์ 3 วิธี ในแบบสอบวัดผลสัมฤทธิ์ สามารถพิจารณาได้ 2 ด้าน คือ

1.1 ผลการวิจัยเกี่ยวกับประสิทธิภาพสัมพัทธ์ของแบบสอบวัดผลสัมฤทธิ์ที่เป็นผลจากการประมาณค่าพารามิเตอร์ข้อสอบ 3 วิธี พบว่า มีทั้งที่สอดคล้อง และขัดแย้งกับสมมติฐานที่ตั้งไว้ ซึ่งพิจารณาได้ดังนี้

ผลการวิจัยที่สอดคล้องกับสมมติฐานได้แก่

1.1.1 แบบสอบที่ทำการประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงกว่า แบบสอบที่ทำการประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด เฉพาะในกลุ่มตัวอย่างระดับความสามารถปานกลางและต่ำ

1.1.2 แบบสอบที่ทำการประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีของเบย์มีคุณภาพสูงกว่า แบบสอบที่ทำการประมาณค่าด้วยวิธีฮิวริสติก ในกลุ่มตัวอย่างทุกระดับความสามารถ

1.1.3 แบบสอบที่ทำการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงกว่า แบบสอบที่ทำการประมาณค่าด้วยวิธีฮิวริสติกในกลุ่มตัวอย่างทุกระดับความสามารถ

ส่วนผลการวิจัยที่ไม่สอดคล้องกับสมมติฐานคือ

1.1.4 แบบสอบที่ทำการประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงกว่า แบบสอบที่ทำการประมาณค่าด้วยวิธีของเบย์ ในกลุ่มตัวอย่างระดับความสามารถสูง

จากผลการวิจัยในข้อ 1.1.1 และ 1.1.4 ที่พบว่า ทั้งสอดคล้องและขัดแย้งกับสมมติฐานดังกล่าวข้างต้นนั้น น่าจะมีสาเหตุเนื่องมาจาก การหาประสิทธิภาพสัมพัทธ์ของแบบสอบ ขึ้นอยู่กับค่าฟังก์ชันสารสนเทศของแบบสอบที่นำมาเปรียบเทียบกัน ซึ่งค่าฟังก์ชันสารสนเทศของแบบสอบต่าง เป็นผลรวมของค่าฟังก์ชันสารสนเทศของข้อสอบที่เป็นดัชนีกำหนดคุณภาพของข้อสอบ ที่ได้จากค่าพารามิเตอร์ทุกตัวในแบบจำลอง ดังนั้นค่าฟังก์ชันสารสนเทศของแบบสอบที่เป็นผลจากการประมาณค่าพารามิเตอร์ข้อสอบแต่ละวิธีนั้น ย่อมมีความแตกต่างกันขึ้นอยู่กับแนวคิด กระบวนการและผลการประมาณค่าพารามิเตอร์แต่ละวิธี ซึ่งในเรื่องความแตกต่างของผลการประมาณค่าพารามิเตอร์ 2 วิธีนี้ Swaminathan และ Gifford (1986 : 589-601) กล่าวว่า เมื่อทำการเปรียบเทียบผลการประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์และวิธีแมกซิมัมไลค์ลิฮูดนั้น วิธีของเบย์มีประสิทธิภาพเหนือกว่าเกือบทุกสถานการณ์ ไม่ว่าจะ เป็นการเปรียบเทียบในด้านค่าสัมประสิทธิ์สหสัมพันธ์ หรือค่าความแตกต่างระหว่างค่าเฉลี่ยกำลังสองของค่าประมาณ (the estimate values) กับค่าจริง (the true values) ยิ่งกว่านั้นในบางสถานการณ์ วิธีแมกซิมัมไลค์ลิฮูดไม่สามารถประมาณค่าพารามิเตอร์ให้คงที่ได้ หรือประมาณได้ค่าคงที่ แต่ไม่อยู่ในขอบเขตของค่าพารามิเตอร์ที่จะยอมรับได้ ซึ่งในสถานการณ์เดียวกันนี้ วิธีของเบย์ยังคงให้ผลการประมาณที่ถูกต้องและเหมาะสม เมื่อพิจารณาผลการประมาณค่าพารามิเตอร์ที่ได้ในการวิจัยครั้งนี้

ก็พบว่าสอดคล้องกับคำกล่าวข้างต้นเช่นกัน กล่าวคือ ค่าพารามิเตอร์ทุกตัวในข้อสอบทุกข้อที่ประมาณด้วยวิธีของเบย์ เป็นค่าคงที่ ที่อยู่ในขอบเขตที่เหมาะสม ส่วนค่าพารามิเตอร์ที่ได้จากการประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูดนั้น มีค่าสูงเกินปกติอยู่หลายค่า เช่น มีค่าความยากสูงถึง 7.50 และยังมีข้อที่ประมาณค่าอำนาจจำแนกได้สูงมาก จนโปรแกรมต้องกำหนดให้เท่ากับขอบเขตบนที่ควรจะเป็นคือ 2.00 อีก 4 ข้อด้วยกัน ค่าพารามิเตอร์เหล่านี้ร่วมกันส่งผลให้แบบสอบที่ประมาณด้วยแมกซิมัมไลค์ลิฮูด มีค่าความยากสูงกว่าของเบย์ รวมทั้งยังมีค่าฟังก์ชันสารสนเทศสูงสุด ที่กลุ่มตัวอย่างระดับความสามารถสูงกว่าอีกด้วย กล่าวคือ แบบสอบที่ประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด มีค่าฟังก์ชันสารสนเทศสูงสุดที่กลุ่มตัวอย่างระดับความสามารถสูง ส่วนแบบสอบที่ประมาณด้วยวิธีของเบย์ มีค่าฟังก์ชันสารสนเทศสูงสุดที่กลุ่มตัวอย่างระดับความสามารถปานกลาง ดังนั้นเมื่อนำค่าฟังก์ชันสารสนเทศของแบบสอบที่ได้จากวิธีประมาณค่าพารามิเตอร์ 2 วิธี มาเปรียบเทียบกันในทุกระดับความสามารถ จึงพบว่า แบบสอบที่ทำการประมาณด้วยวิธีของเบย์ มีคุณภาพสูงกว่าที่ประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูดเกือบทุกระดับความสามารถ ยกเว้นในระดับความสามารถสูง แบบสอบที่ประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงกว่า วิธีของเบย์ ซึ่งปรากฏการณ์เช่นนี้ สอดคล้องกับคำกล่าวของ Hambleton และ Swaminathan (1985 : 106-111) ที่ว่า เมื่อข้อสอบมีความยากอยู่ระดับใด ข้อสอบจะมีความแม่นยำในการประมาณค่าความสามารถบุคคลในระดับนั้น ระดับความสามารถนี้อาจคลาดเคลื่อนไปได้บ้าง ขึ้นอยู่กับค่าอำนาจจำแนก และค่าการเดาคำ ดังที่ Warm กล่าวว่า ถ้าข้อสอบมีค่าอำนาจจำแนกสูง และค่าการเดาคำแล้ว ข้อสอบจะมีความแม่นยำในการประมาณค่าความสามารถบุคคลสูง (Warm 19798 : 69-70)

สำหรับผลการวิจัยในข้อ 1.1.2 และ 1.1.3 ที่พบว่า แบบสอบวัดผลสัมฤทธิ์ ที่ทำการประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีของเบย์และวิธีแมกซิมัมไลค์ลิฮูดต่างมีคุณภาพสูงกว่า วิธีอีวริสติก ในทุกระดับความสามารถ และสอดคล้องกับสมมติฐานที่น่าจะเป็นเพราะทั้งวิธีของเบย์และวิธีแมกซิมัมไลค์ลิฮูดนี้ มีกระบวนการประมาณค่าพารามิเตอร์โดยใช้การทำซ้ำ ๆ จนกว่าจะได้ค่าคงที่ (Iterative) จึงทำให้ค่าพารามิเตอร์ที่ประมาณได้ค่อนข้างถูกต้อง ส่วนกระบวนการประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีอีวริสติกเป็นการประมาณอย่างง่าย จากค่าสถิติที่ได้จากการวิเคราะห์ในแนวทฤษฎีดั้งเดิม โดย

ไม่ต้องมีการประมาณค่าอย่างซ้ำๆ ค่าพารามิเตอร์ข้อสอบที่ประมาณได้ จึงอาจคลาดเคลื่อนไปบ้าง ดังนั้นเมื่อตรวจสอบค่าฟังก์ชันสารสนเทศแบบสอบ ซึ่งเป็นดัชนีกำหนดคุณภาพที่ได้จากค่าพารามิเตอร์ข้อสอบทุกค่า จึงทำให้ได้ค่าฟังก์ชันสารสนเทศต่ำกว่าวิธีอื่น ๆ แต่ผลการวิจัยนี้ยังไม่สามารถจะลงสรุปได้ว่า วิธีอิวิริสติก ด้อยประสิทธิภาพกว่าวิธีอื่น ๆ ทั้งหมด ทั้งนี้เพราะ Swaminathan และ Gifford (1983 cited by Hambleton and Swaminathan 1985 : 147) เคยนำวิธีอิวิริสติกไปใช้ พบว่า แม้จะมีความคลาดเคลื่อนบ้างในการประมาณค่าพารามิเตอร์ของข้อสอบ อันเนื่องจากการประมาณอย่างง่าย แต่ค่าพารามิเตอร์ข้อสอบที่ได้นี้ เมื่อนำไปเป็นค่าคงตัวในการประมาณความสามารถของผู้เข้าสอบด้วยวิธีแมกซิมัมไลค์ลิฮูดแบบมีเงื่อนไข (คือประมาณเฉพาะค่าความสามารถอย่างเดียว ซึ่งเป็นกระบวนการที่ลดความยุ่งยาก และปัญหาอันเกิดจากการประมาณพร้อม ๆ กันของค่าพารามิเตอร์ข้อสอบ และค่าความสามารถของผู้สอบ) แล้วค่าความสามารถที่ประมาณได้ค่อนข้างถูกต้อง โดยเฉพาะอย่างยิ่ง เมื่อนำไปประยุกต์ใช้ในสถานการณ์ที่จำนวนข้อสอบและกลุ่มผู้เข้าสอบมีขนาดใหญ่แล้ว ค่าความสามารถที่ประมาณด้วยวิธีอิวิริสติก สามารถนำไปเทียบเคียงกับค่าความสามารถที่ประมาณจากวิธีแมกซิมัมไลค์ลิฮูดได้ ซึ่งจะได้อภิปรายถึงในตอนต่อ ๆ ไป

1.2 ผลการวิจัยเกี่ยวกับความตรงร่วมสมัยของแบบวัดผลสัมฤทธิ์ ที่เป็นผลจากการประมาณค่าพารามิเตอร์ความสามารถของผู้เข้าสอบทั้ง 3 วิธี พบว่า แบบสอบที่ประมาณค่าความสามารถผู้เข้าสอบด้วยวิธีของเบย์ มีความตรงร่วมสมัยสูงกว่า วิธีแมกซิมัมไลค์ลิฮูด และวิธีอิวิริสติก อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ส่วนแบบสอบที่ประมาณด้วยวิธีอิวิริสติกมีความตรงร่วมสมัยสูงกว่าวิธีแมกซิมัมไลค์ลิฮูด อย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งล้วนเป็นผลการวิจัยที่สอดคล้องกับสมมติฐานทั้งสิ้น ทั้งนี้ น่าจะเป็นผลเนื่องมาจาก ผลการประมาณค่าในข้อ 1.1 ที่พบว่า แบบสอบที่ประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีของเบย์ และวิธีอิวิริสติก มีความแม่นยำสูงที่สุดในการประมาณค่าความสามารถบุคคลในระดับความสามารถปานกลาง กล่าวคือ มีค่าฟังก์ชันสารสนเทศสูงสุดในระดับความสามารถปานกลางนั่นเอง แต่ค่าฟังก์ชันสารสนเทศสูงสุดของวิธีของเบย์ มีค่าสูงกว่า ของวิธีอิวิริสติก ส่วนแบบสอบที่ประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด มีค่าฟังก์ชันสารสนเทศสูงสุดใน

ระดับความสามารถสูง ประกอบกับกลุ่มตัวอย่างที่สุ่มมาจากประชากรที่ใช้ในการวิจัยครั้งนี้ มีระดับความสามารถเฉลี่ยอยู่ในระดับปานกลาง ไม่ว่าจะทำการประมาณค่าพารามิเตอร์ ความสามารถด้วยวิธีใด ๆ ในทั้ง 3 วิธี ($\bar{\theta}_{\max} = -0.133$, $\bar{\theta}_{\text{BAY}} = -0.063$, $\bar{\theta}_{\text{HVE}} = -0.106$) ดังนั้นเมื่อนำค่าพารามิเตอร์ความสามารถที่ประมาณได้ในแต่ละวิธีของผู้เข้าสอบมาหาความสัมพันธ์กับคะแนนเกณฑ์ จึงพบว่า แบบสอบที่ประมาณค่าพารามิเตอร์ ความสามารถของผู้เข้าสอบ ด้วยวิธีของเบย์ มีความตรงร่วมสมัยสูงที่สุด รองลงมาคือ วิธี ฮิวริสติก หรือวิธีแมกซิมัมไลค์ลิฮูด

2. ผลการวิจัยที่เกี่ยวกับประสิทธิผลของวิธีประมาณค่าพารามิเตอร์ 3 วิธี ในแบบสอบความถนัดสามารถพิจารณาได้ 2 ด้านคือ

2.1 ผลการวิจัยเกี่ยวกับประสิทธิภาพสัมพันธ์ของแบบสอบความถนัดที่เป็นผลจากการประมาณค่าพารามิเตอร์ข้อสอบ 3 วิธี พบว่า มีทั้งสอดคล้องและขัดแย้งกับสมมติฐานที่ตั้งไว้เช่นกัน ซึ่งพิจารณาได้ดังนี้

ผลการวิจัยที่สอดคล้องกับสมมติฐานได้แก่

2.1.1 แบบสอบที่ทำการประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงกว่า แบบสอบที่ทำการประมาณค่าด้วยวิธีแมกซิมัมไลค์ลิฮูด เฉพาะในกลุ่มตัวอย่างระดับความสามารถปานกลางและต่ำ

2.1.2 แบบสอบที่ทำการประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงกว่า แบบสอบที่ทำการประมาณค่าด้วยวิธีฮิวริสติก ในกลุ่มตัวอย่างทุกระดับความสามารถ

2.1.3 แบบสอบที่ทำการประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีแมกซิมัมไลค์ลิฮูด มีคุณภาพสูงกว่า แบบสอบที่ทำการประมาณค่าด้วยวิธีฮิวริสติก เฉพาะในกลุ่มตัวอย่างระดับความสามารถสูงและปานกลาง

ส่วนผลการวิจัยที่ไม่สอดคล้องกับสมมติฐานคือ

2.1.4 แบบสอบที่ทำการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม

โลคัลลิสต์ มีคุณภาพสูงกว่า แบบสอบที่ประมาณด้วยวิธีของเบย์ เฉพาะในกลุ่มตัวอย่างระดับความสามารถสูง

2.1.5 แบบสอบที่ทำการประมาณค่าพารามิเตอร์ด้วยวิธีฮิวริสติก มีคุณภาพสูงกว่า แบบสอบที่ประมาณด้วยวิธีแมกซิมัมไลคัลลิสต์ เฉพาะในกลุ่มตัวอย่างระดับความสามารถต่ำ

จากผลการวิจัยในข้อ 2.1.1 - 2.1.4 นั้นมีเหตุผลสนับสนุนเช่นเดียวกับที่อภิปรายไว้ในข้อ 1.1 ดังนั้นในข้อ 2.1 นี้ ผู้วิจัยจึงจะอภิปรายเฉพาะข้อ 2.1.5 ซึ่งเป็นผลการวิจัยที่ขัดแย้งกับสมมติฐานในประเด็นที่แตกต่างออกไป กล่าวคือ แบบสอบความถนัดที่ประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีฮิวริสติก มีคุณภาพสูงกว่า ที่ประมาณด้วยวิธีแมกซิมัมไลคัลลิสต์ ในกลุ่มตัวอย่างระดับความสามารถต่ำ ทั้งนี้อาจเนื่องมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัมไลคัลลิสต์ ในแบบสอบความถนัดนี้ ข้อสอบมีค่าพิสัยของค่าความยาก (b) กว้างมาก คือ b สูงสุดเท่ากับ 32.925 ส่วน b ต่ำสุดเท่ากับ -4.748 ซึ่งจากกระบวนการประมาณค่าของวิธีแมกซิมัมไลคัลลิสต์นั้น Lord (1975) กล่าวว่า ถ้าข้อสอบค่อนข้างยากแล้ว ค่าประมาณอำนาจจำแนกจะสูงเกินจริง แต่ถ้าข้อสอบค่อนข้างง่ายแล้ว ค่าประมาณอำนาจจำแนกก็จะต่ำกว่าค่าจริงด้วย ส่วน Warm (1978 : 70) กล่าวว่า ถ้าข้อสอบมีค่าอำนาจจำแนกสูงหรือต่ำ ก็จะมีผลให้ค่าฟังก์ชันสารสนเทศสูงหรือต่ำตามลำดับเช่นกัน ประกอบกับข้อสรุปของ Hambleton และ Swaminathan (1985 : 106 - 111) ที่ว่า ถ้าข้อสอบมีความยากอยู่ในระดับใด แล้วจะมีค่าฟังก์ชันสารสนเทศสูงสุดในระดับความสามารถนั้นๆ ด้วย ดังนั้น ข้อสอบที่ค่อนข้างยากถึงยากมาก จึงมีค่าฟังก์ชันสารสนเทศสูงมาก ในกลุ่มตัวอย่างระดับความสามารถสูง ในขณะที่ข้อสอบที่ค่อนข้างง่ายถึงง่ายมากน่าจะมีค่าฟังก์ชันสารสนเทศสูงในกลุ่มตัวอย่างระดับความสามารถต่ำด้วย แต่ไม่เป็นเช่นนั้น เพราะอิทธิพลค่าอำนาจจำแนกที่ต่ำเกินจริงนั่นเอง และเมื่อนำค่าฟังก์ชันของแบบสอบทั้งฉบับที่เกิดจากผลรวมของค่าฟังก์ชันสารสนเทศ ข้อสอบทุกข้อของแบบสอบความถนัดที่เป็นผลจากการประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีแมกซิมัมไลคัลลิสต์ มาเปรียบเทียบกับวิธีฮิวริสติก ที่มีค่าฟังก์ชันสารสนเทศสูงในกลุ่มตัวอย่างระดับความสามารถปานกลางค่อนข้างต่ำถึงต่ำนั้น จึงพบว่า ในกลุ่มตัวอย่างระดับความสามารถต่ำ แบบสอบ

ความถนัดที่ประมาณด้วยวิธีชีววิสถิติ มีคุณภาพสูงกว่าที่ประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูด

2.2 ผลการวิจัยที่เกี่ยวกับความตรงเชิงทำนายของแบบสอบความถนัด ที่เป็นผลจากการประมาณค่าพารามิเตอร์ความสามารถของผู้เข้าสอบทั้ง 3 วิธี พบว่า แบบสอบที่ประมาณค่าความสามารถของผู้เข้าสอบด้วยวิธีของเบย์ และวิธีชีววิสถิติมีความตรงเชิงทำนายสูงกว่าวิธีแมกซิมัมไลค์ลิฮูด อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ซึ่งสอดคล้องกับสมมติฐาน ส่วนแบบสอบที่ประมาณด้วยวิธีของเบย์มีความตรงเชิงทำนายสูงกว่าวิธีชีววิสถิติ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 นั้น เป็นผลการวิจัยที่ไม่สอดคล้องกับสมมติฐาน ทั้งนี้ น่าจะเป็นผลกระทบมาจากผลการประมาณค่าพารามิเตอร์ข้อสอบในข้อ 2.1 ซึ่งเมื่อพิจารณาค่าฟังก์ชันสารสนเทศของแบบสอบที่เป็นผลจากการประมาณแต่ละวิธีพบว่า การประมาณด้วยวิธีแมกซิมัมไลค์ลิฮูดส่งผลให้มีค่าฟังก์ชันสารสนเทศสูงสุดในกลุ่มตัวอย่างระดับความสามารถสูง ส่วนการประมาณด้วยวิธีของเบย์ และวิธีชีววิสถิติส่งผลให้ค่าฟังก์ชันสารสนเทศแบบสอบสูงสุดในกลุ่มตัวอย่างระดับความสามารถปานกลาง และปานกลางถึงค่อนข้างต่ำ ตามลำดับ อย่างไรก็ตามแม้วิธีของเบย์และวิธีชีววิสถิติ จะมีค่าฟังก์ชันสารสนเทศสูงสุดในระดับความสามารถใกล้เคียงกัน แต่ค่าฟังก์ชันสารสนเทศของแบบสอบที่ได้จากวิธีของเบย์ก็มีค่าสูงกว่าที่ได้จากวิธีชีววิสถิติ นอกจากนี้ ในระดับความสามารถต่ำ ค่าฟังก์ชันสารสนเทศแบบสอบที่ได้จากวิธีชีววิสถิติ มีค่าสูงกว่าที่ได้จากวิธีแมกซิมัมไลค์ลิฮูด ประกอบกับความสามารถของกลุ่มตัวอย่างประชากรที่ใช้ในการวิจัยครั้งนี้ ไม่ว่าจะประมาณด้วยวิธีใด ๆ ก็ตาม ล้วนมีค่าเฉลี่ยอยู่ในระดับปานกลางถึงค่อนข้างต่ำ ($\bar{\theta}_{max} = -0.018$, $\bar{\theta}_{BAY} = -0.038$, $\bar{\theta}_{HVE} = -0.079$) ดังนั้นเมื่อนำค่าพารามิเตอร์ความสามารถของผู้เข้าสอบที่ประมาณได้ในแต่ละวิธี มาหาความสัมพันธ์กับคะแนนเกณฑ์ด้วยวิธีของเบย์ จึงมีความตรงเชิงทำนายสูงที่สุด รองลงมาคือ วิธีชีววิสถิติ และวิธีแมกซิมัมไลค์ลิฮูด ตามลำดับ

จากการอภิปรายผลทั้งหมดที่กล่าวมา จะเห็นว่าผลวิจัยในด้านค่าประสิทธิภาพสัมพัทธ์แบบสอบนั้นค่อนข้างคงที่ ทั้ง ๆ ที่แบบสอบวัดผลสัมฤทธิ์มีความเหมาะสมในการสอบวัดในกลุ่มตัวอย่างระดับความสามารถสูง ส่วนแบบสอบความถนัดมีความเหมาะสมในการสอบวัดกลุ่มตัวอย่างระดับความสามารถปานกลาง กล่าวคือแบบสอบที่ประมาณค่าพารามิเตอร์

ข้อสอบด้วยวิธีแมกซีมัมไลค์ลิตูด มีคุณภาพสูงสุดในกลุ่มตัวอย่างระดับความสามารถสูง ส่วนแบบสอบที่ประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีของเบย์ มีคุณภาพสูงสุดในกลุ่มตัวอย่างระดับความสามารถปานกลางและต่ำ ซึ่งผลการวิจัยข้างต้นนี้ ส่งผลให้แบบสอบที่ประมาณค่าพารามิเตอร์ความสามารถของผู้เข้าสอบด้วยวิธีของเบย์ มีความตรงเชิงเกณฑ์สัมพัทธ์สูงสุด และเป็นผลการวิจัยที่สอดคล้องกับสมมติฐานที่ตั้งไว้ อนึ่ง การที่ประสิทธิภาพของวิธีประมาณค่าค่อนข้างคงที่นั้น อาจเป็นเพราะ ความเหลื่อมซ้อนกัน (Over lap) ของแบบสอบทั้ง 2 ประเภท ซึ่งตามทฤษฎีแล้วไม่สามารถที่จะแยกความแตกต่างของแบบสอบทั้งสองประเภทได้อย่างชัดเจน ยกเว้นในการนำไปใช้ เพราะแบบสอบทั้งสองประเภทจะมีจุดมุ่งหมายในการวัดต่างกัน คือแบบสอบวัดผลสัมฤทธิ์วัดความสามารถที่แสดงผลการเรียนรู้ ส่วนแบบสอบความถนัดวัดความสามารถที่แสดงความพร้อมที่จะเรียนรู้ (อุทุมพร จามรมา 2532 : 171) ความแตกต่างดังกล่าว อาจส่งผลให้แบบสอบแต่ละประเภทมีข้อค้นพบบางประการที่แตกต่างกันออกไป

นอกจากนี้ ผลการวิจัยในส่วนของความตรงเชิงเกณฑ์สัมพัทธ์ เป็นผลเฉพาะในกรณีของกลุ่มตัวอย่างประชากรมีความสามารถปานกลางและค่อนข้างต่ำเท่านั้น จึงไม่อาจสรุปได้แน่ชัดว่า ถ้ากลุ่มตัวอย่างมาจากประชากรที่มีความสามารถระดับสูงแล้ว ผลการวิจัยจะเป็นเช่นใด

อย่างไรก็ตาม ผลการวิจัยที่ได้รับในครั้งนี้ ก็เป็นข้อค้นพบที่มีประโยชน์ค่อนข้างจะครอบคลุมในสภาพการวัดผลการศึกษาในปัจจุบัน ที่ต้องการใช้แบบสอบวัดผลสัมฤทธิ์ และแบบสอบความถนัด ให้มีความแม่นยำในการตรวจสอบ หรือทำนายความสามารถของบุคคลเพื่อการตัดสินใจอย่างหนึ่งอย่างใด โดยเลือกใช้วิธีประมาณค่าที่เหมาะสมกับประเภทของแบบสอบและระดับความสามารถของกลุ่มผู้เข้าสอบ ซึ่งหากผู้ดำเนินการสอบวัดได้ประยุกต์ใช้อย่างถูกต้อง เหมาะสมแล้ว การวัดผลการศึกษา ก็จะเป็นกระบวนการที่มีความหมายและให้สารสนเทศที่เชื่อถือได้

ข้อเสนอแนะ

จากผลการวิจัยครั้งนี้ ทำให้ได้แนวคิดที่อาจจะมีประโยชน์ต่อไป ผู้วิจัยจึงขอเสนอ เป็น 2 ประเด็นคือ

ข้อเสนอแนะในการนำไปใช้

1. การหาคุณภาพตามกรอบทฤษฎีการตอบสนองต่อข้อสอบด้านประสิทธิภาพสัมพัทธ์ของแบบสอบวัดผลสัมฤทธิ์และแบบสอบความถนัด ในกลุ่มตัวอย่างระดับความสามารถสูง ควรเลือกใช้การประมาณค่าพารามิเตอร์ข้อสอบด้วยวิธีแมกซิมัมไลค์ลิฮูด ส่วนในกลุ่มตัวอย่างระดับความสามารถปานกลางและต่ำนั้น ควรเลือกใช้วิธีของเบย์
2. การหาคุณภาพของแบบสอบวัดผลสัมฤทธิ์และแบบสอบความถนัด ตามกรอบทฤษฎีการสอบแบบดั้งเดิม ด้านความตรงเชิงเกณฑ์สัมพัทธ์นั้น เมื่อกลุ่มผู้เข้าสอบมีระดับความสามารถเฉลี่ยปานกลางหรือค่อนข้างต่ำ ควรเลือกใช้การประมาณค่าพารามิเตอร์ความสามารถผู้เข้าสอบด้วยวิธีของเบย์
3. โรงเรียนที่ไม่มีเครื่องคอมพิวเตอร์ และบุคลากรที่มีความรู้ความสามารถในการวิเคราะห์แบบสอบในแนวทฤษฎีการตอบสนองต่อข้อสอบอาจนำข้อมูลผลการสอบบันทึกลงในกระดาษลงรหัส และนำมาขอใช้บริการประมวลผลที่สถาบันบริการคอมพิวเตอร์จุฬาลงกรณ์มหาวิทยาลัยก็ได้
4. ในระดับโรงเรียนขนาดเล็กที่รวมตัวกันเป็นกลุ่มโรงเรียน ไม่มีความสะดวกในเรื่องเครื่องคอมพิวเตอร์ และกลุ่มผู้สอบมีความสามารถค่อนข้างต่ำถึงต่ำนั้น หากต้องการวิเคราะห์ข้อสอบและประมาณความสามารถของผู้เข้าสอบตามกรอบทฤษฎีการตอบสนองต่อข้อสอบ อาจเลือกใช้วิธีฮิวริสติกในการประมาณก็ได้ แต่แบบสอบที่จะนำมาวิเคราะห์ต้องมีความเที่ยงสูงด้วย
5. คุณภาพของแบบสอบขึ้นอยู่กับความสามารถของผู้สร้างแบบสอบ เป็นสำคัญวิธีประมาณค่าทั้งหลายนั้น เป็นเพียงดัชนีที่สนับสนุนให้ผู้สร้างแบบสอบ ได้นำมาพิจารณาในการสร้างแบบสอบเท่านั้น แต่ไม่สามารถช่วยให้แบบสอบมีคุณภาพดีขึ้นได้ ดังนั้น การที่จะนำ

วิธีประมาณค่าต่าง ๆ ในทฤษฎีการตอบสนองต่อข้อสอบไปประยุกต์ใช้ให้ได้ผลที่ถูกต้องและ
คุ้มค่านั้น ต้องเลือกข้อสอบที่มีการสร้างถูกต้องมีหลักการและ เชื่อถือได้

ข้อเสนอแนะในการวิจัย

1. ควรทำการศึกษาวิจัยท่านองนี้ โดยเลือกเฉพาะกลุ่มตัวอย่างจากประชากร
ที่มีความสามารถระดับสูง หรือ เฉพาะกลุ่มความสามารถระดับต่ำมาทำการวิจัย เพื่อให้ได้
ข้อมูลเพิ่มเติม อันจะเป็นแนวคิดในการพัฒนาวิธีประมาณค่าพารามิเตอร์ในทฤษฎีการตอบ
สนองต่อข้อสอบให้มีประสิทธิภาพยิ่งขึ้นต่อไป
2. ควรทำการศึกษาวิจัยท่านองนี้ ในวิชาอื่น ๆ เช่น คณิตศาสตร์ วิทยาศาสตร์
เป็นต้น เพื่อให้ได้ข้อความรู้ที่กว้างขวาง และก่อให้เกิดการอ้างอิงเกี่ยวกับประสิทธิผลของ
วิธีประมาณค่าได้อย่างเป็นทั่วไปมากขึ้น
3. จากงานวิจัยเกี่ยวกับทฤษฎีการตอบสนองต่อข้อสอบ ไม่ว่าจะทำการเปรียบเทียบ
เทียบคุณภาพแบบสอบ วิธีการทดสอบ หรือวิธีการให้คะแนนผลการสอบ เท่าที่ผ่านมาส่วน
ใหญ่ จะใช้วิธีแมกซิมัมไลค์ลิชูด ในการประมาณค่าความสามารถผู้เข้าสอบ และมักจะได้ข้อ
ค้นพบว่าตัวแปรต่าง ๆ ที่นำมาทดสอบสมมติฐานนั้นไม่แตกต่างกัน ทั้งนี้อาจเนื่องมาจาก
ความจำกัดของวิธีประมาณที่ต้องมีการกำจัดผู้ที่ได้คะแนนเต็มหรือคะแนนศูนย์จากการ
ประมาณ ทำให้เหลือผู้เข้าสอบที่มีระดับความสามารถใกล้เคียงกัน จึงเป็นเหตุให้คุณภาพ
ของตัวแปรไม่แตกต่างกันดังกล่าวข้างต้น ดังนั้น ควรมีการวิจัยในท่านองเดิมอีก แต่ใช้วิธี
ของเบย์ในการประมาณค่าความสามารถบุคคลแทนวิธีแมกซิมัมไลค์ลิชูด เพื่อศึกษาว่าผลที่ได้
จะเหมือนหรือต่างกันในด้านใดบ้าง
4. ควรจะมีการวิจัยในท่านองนี้ โดยเปลี่ยนค่าดัชนีที่ใช้ตรวจสอบสมมติฐาน
จากค่าประสิทธิภาพสัมพัทธ์ เป็นค่า ความคลาดเคลื่อนมาตรฐานของแบบสอบแทน
5. ควรทำการวิจัยในประเด็นปัญหาเดียวกันนี้ แต่ใช้เทคนิคมอนติคาร์โล โดย
กำหนดรูปแบบผลการตอบให้มีคนที่ได้คะแนนเต็มหรือคะแนนศูนย์ และมีข้อที่ทุกคนตอบถูกหรือ
ข้อที่ทุกคนตอบผิดในสถานการณ์ที่จำนวนข้อสอบ และผู้เข้าสอบมีขนาดต่าง ๆ กัน เพื่อศึกษา
ประสิทธิผลของวิธีประมาณค่าพารามิเตอร์ เมื่อมีข้อจำกัดในการประมาณต่าง ๆ กัน