

## Effects of Local Item Dependence on Reliability, Item Parameters, Person Ability Estimates and Test Information Under Different Testing Conditions

Sunthorn Thiengam<sup>1</sup>  
Sirichai Karnjanawasee<sup>2</sup>  
Derek Srisukho<sup>3</sup>

### ABSTRACT

The objectives of this research were (1) to study the variation of local item dependence (LID) on the estimation of reliability, item parameters, ability and test information under different testing conditions, and (2) to study the presence of LID in Mathematics and English Test in actual data. Monte Carlo technique was used to simulate the data item response. The conditions were consisted of 9 levels of local item dependence divided into 3 levels (mild, medium, strong) 3 item test length (30, 50, 80) and 3 sample sizes (400, 800, 1200). The total of testing conditions was 81. The 1000 replications used to estimate the item parameters and test statistics in each condition. In addition, analyzing the real data in Mathematics and English Test. Fortran Power Station Version 4.0 was used to generate the items datasets and LDID (A Computer Program for local Dependence Indices for Dichotomous Item 1. Version 1.0.) was used to detect LID. The research results were as follows:

1. The reliability increased when LID increase for all conditions. The Minimum- Maximum values of reliability are 0.671-0.740 for 30 items, 0.773-0.823 for 50 items and 0.845-0.886 for 80 items respectively.

2. According to IRT analysis results for the conditions of 400, 800 and 1200 examinees by 30, 50 and 80 items; It was found that (2.1) for 1PL Model, when LID increase the difficulty and ability values were not different while test information tended to decrease with statistical significant different level at .01. (2.2) for 2PL Model, when LID increase most of all conditions were found that the difficulty decreased when the discrimination, ability and test information tended to increase with statistical significant different level at .05. (2.2) for 3PL Model, when LID increase the difficulty, ability and test information tended to decrease while the discrimination and guessing mostly tended to increase and statistical significant different level at .05.

3. For the real data analysis by 1PL, 2PL and 3PL in Mathematics Test were found that the mean  $|Q_3|$  were 0.035, 0.026, 0.022 and English Test was 0.031, 0.028 and 0.022 respectively. Most of  $|Q_3|$  were closed to absolute expected  $|Q_3|$  as .025. It was shown that those tests were slightly local item dependent.

<sup>1</sup> Ph.D. in Educational Research and Psychology

<sup>2</sup> Advisor

<sup>3</sup> Co-Advisor

# ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ค่าความสามารถของผู้สอบ และค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

สุนทร เทียนงาม<sup>1</sup>  
ศิริชัย กาญจนวาสี<sup>2</sup>  
ดิเรก ศรีสุโข<sup>3</sup>

## บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์ผลของระดับความไม่เป็นอิสระของข้อสอบต่อค่าความเที่ยงของแบบสอบ ค่าพารามิเตอร์ของข้อสอบ ค่าความสามารถของผู้สอบและสารสนเทศของแบบสอบ ด้วยทฤษฎีการตอบสนองข้อสอบ ที่มีจำนวนข้อสอบและขนาดกลุ่มตัวอย่างแตกต่างกัน และเพื่อวิเคราะห์ลักษณะความไม่เป็นอิสระของข้อสอบจากแบบสอบจริง การวิจัยนี้ใช้ระเบียบวิธีวิจัยเชิงทดลอง โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล จำนวนตัวแปรประกอบด้วย ระดับความไม่เป็นอิสระของข้อสอบ 9 ระดับโดยแบ่งเป็น ระดับความไม่เป็นอิสระ 3 ระดับ (ระดับต่ำ, ปานกลาง และสูง) จำนวนข้อสอบ 3 กลุ่ม (30, 50 และ 80 ข้อ) และ ขนาดกลุ่มตัวอย่าง 3 กลุ่ม (400, 800 และ 1,200 คน) รวมจำนวนเงื่อนไขทั้งหมด 81 เงื่อนไข และกำหนดจำนวนรอบที่ใช้การประมาณค่าพารามิเตอร์ในแต่ละเงื่อนไข 1,000 รอบ และใช้วิธีการศึกษาจากข้อมูลจริงในวิชาคณิตศาสตร์และภาษาอังกฤษ โปรแกรมที่ใช้ในการจำลองข้อมูล คือ Fortran Power Station 4.0 และโปรแกรมในการตรวจสอบความไม่เป็นอิสระของข้อสอบ คือ LDID (A Computer Program for Local Dependence Indices for Dichotomous Item Version 1) ผลการวิจัยสรุปได้ดังนี้

1. เมื่อระดับความไม่เป็นอิสระของข้อสอบเพิ่มสูงขึ้น ค่าความเที่ยง (KR20) จะเพิ่มสูงขึ้นในทุกเงื่อนไขของการทดสอบ โดยเมื่อใช้จำนวนข้อสอบ 30 ข้อ ค่าความเที่ยงต่ำสุด-สูงสุดมีค่าเท่ากับ 0.671-0.740 จำนวนข้อสอบ 50 ข้อ ค่าความเที่ยงต่ำสุด-สูงสุดมีค่าเท่ากับ 0.773-0.828 และจำนวนข้อสอบ 80 ข้อ ค่าความเที่ยงต่ำสุด-สูงสุดมีค่าเท่ากับ 0.845-0.886

2. เมื่อวิเคราะห์ข้อสอบตามแนวทฤษฎีการตอบสนองข้อสอบ มีจำนวนผู้สอบ 400 800 และ 1,200 คน ทำข้อสอบ 30 50 และ 80 ข้อ ตามลำดับ (2.1) เมื่อวิเคราะห์โดยใช้โมเดล 1 พารามิเตอร์ พบว่า เมื่อระดับความไม่เป็นอิสระของข้อสอบเพิ่มสูงขึ้น ค่าความยากของข้อสอบแต่ละข้อและค่าความสามารถของผู้สอบแต่ละคนไม่แตกต่างกันเมื่อจำนวนผู้สอบและจำนวนข้อสอบแตกต่างกัน และค่าสารสนเทศของแบบสอบส่วนใหญ่มีแนวโน้มคงที่ (2.2) เมื่อวิเคราะห์โดยใช้โมเดล 2 พารามิเตอร์ พบว่า เมื่อระดับความไม่เป็นอิสระของข้อสอบเพิ่มสูงขึ้น ส่วนใหญ่ค่าความยากของข้อสอบลดลง แต่ค่าอำนาจจำแนก ค่าความสามารถของผู้สอบ และสารสนเทศของแบบสอบมีแนวโน้มสูงขึ้น โดยค่าความยากและค่าอำนาจจำแนกของข้อสอบ ค่าความสามารถของผู้สอบ และสารสนเทศของแบบสอบส่วนใหญ่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ (2.3) เมื่อวิเคราะห์โดยใช้โมเดล 3 พารามิเตอร์ พบว่า เมื่อระดับความไม่เป็นอิสระของข้อสอบเพิ่มสูงขึ้น ค่าความยากของข้อสอบ ค่าความสามารถของผู้สอบ และสารสนเทศของแบบสอบมีแนวโน้มลดลง แต่สำหรับค่าอำนาจจำแนก ค่าการเดาส่วนใหญ่มีแนวโน้มสูงขึ้น โดยค่าความยาก ค่าอำนาจจำแนก และค่าการเดาของข้อสอบ ค่าความสามารถของผู้สอบและสารสนเทศของแบบสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. เมื่อวิเคราะห์ความไม่เป็นอิสระของข้อสอบโดยใช้ข้อมูลจริงด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ สำหรับวิชาคณิตศาสตร์ พบว่า ค่าเฉลี่ย  $|Q_j|$  มีค่าเท่ากับ 0.035, 0.026, 0.022 ตามลำดับ และในวิชาภาษาอังกฤษ พบว่า ค่าเฉลี่ย  $|Q_j|$  มีค่าเท่ากับ 0.033, 0.028, 0.022 ตามลำดับ มีค่าใกล้เคียงกับค่าคาดหวังของ  $|Q_j|$  ซึ่งมีค่าเท่ากับ .025 แสดงว่าข้อสอบส่วนใหญ่มีความไม่เป็นอิสระของข้อสอบเล็กน้อย

<sup>1</sup> ดุษฎีบัณฑิต ภาควิชาวิจัยและจิตวิทยาการศึกษา

<sup>2</sup> อาจารย์ที่ปรึกษา

<sup>3</sup> อาจารย์ที่ปรึกษาร่วม

## ความเป็นมาและความสำคัญของปัญหา

การนำผลคะแนนจากแบบสอบหลายตัวเลือกมาใช้ในการประเมินผลการศึกษา ตลอดจน การตัดสินใจและการคัดเลือกบุคคลเข้าศึกษาต่อ ยังคงเป็นประเด็นที่นักวัดและประเมินผล ยังต้องศึกษา อยู่ในปัจจุบันโดยเฉพาะปัญหาเกี่ยวกับการนำผลคะแนนที่ได้จากแบบสอบมาประมาณค่าความสามารถ ว่ามีความถูกต้องแม่นยำและยุติธรรมหรือไม่

การวิเคราะห์ตามแนวทางทฤษฎีการตอบสนองข้อสอบ (Item Response Theory = IRT) นำมาใช้ในการประมาณค่าความสามารถ ค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ โดยได้กำหนด ข้อตกลงเบื้องต้นข้อหนึ่งไว้ว่า ข้อสอบแต่ละข้อต้องมีความเป็นอิสระจากกัน (Local Item Independence) หากข้อสอบมีการฝ่าฝืนข้อตกลงเบื้องต้นดังกล่าว จะส่งผลต่อความคลาดเคลื่อน ในการประมาณค่าพารามิเตอร์ที่ได้ ได้แก่ การประมาณค่าพารามิเตอร์ของข้อสอบ การประมาณค่า สารสนเทศของข้อสอบและแบบสอบ การประมาณค่าความสามารถของผู้สอบและความคลาดเคลื่อน มาตรฐานของการวัด (Zenisky et. al., 2003) นอกจากนี้ ความไม่เป็นอิสระของข้อสอบยังลด ความตรงของข้อสอบด้วย

ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระของข้อสอบ มักจะใช้ในโมเดลที่มีตัวแปรแฝง สำหรับ ทฤษฎีการตอบสนองข้อสอบได้กำหนดไว้ว่า ความเป็นอิสระระหว่างข้อสอบและผู้สอบ หมายถึง เมื่อมีการควบคุมความสามารถของผู้สอบ ( $\theta$ ) ที่ส่งผลต่อการตอบข้อสอบ หรือให้ความสามารถของ ผู้สอบคงที่ และผลของการตอบข้อสอบแต่ละข้อต้องเป็นอิสระจากกัน (ศิริชัย กาญจนวาสี, 2545; Jiao & Kamata, 2003; Smith, 2005) ความเป็นอิสระของข้อสอบจะพิจารณาจากค่าพารามิเตอร์ ของโมเดลเฉพาะหรือกล่าวอีกนัยหนึ่ง ความเป็นอิสระของข้อสอบคือ ความน่าจะเป็นของการตอบ ข้อสอบใดๆ เป็นอิสระจากข้ออื่น ๆ เมื่อควบคุมความสามารถของผู้สอบ ในทางคณิตศาสตร์ ความเป็นอิสระของข้อสอบเป็นความน่าจะเป็นร่วมของการตอบข้อสอบรายคู่ได้ถูกต้อง (joint probability of correct responses to an item pair) สามารถคำนวณจากสูตร

ในสภาพความเป็นจริงข้อตกลงเบื้องต้นดังกล่าว มักจะถูกฝ่าฝืนเสมอ เช่น ผลจากการตอบ ข้อสอบข้อหนึ่งถูกต้องจะมีผลต่อการตอบข้อสอบข้ออื่นถูกต้องด้วย เช่น จำนวนหรือค่าที่ใช้ร่วมกันในข้อสอบ หรือข้อสอบหลายข้อที่อ่านจากกราฟ แผนภูมิ บทความ จากแหล่งเดียวกัน จะทำให้ค่าพารามิเตอร์ ของข้อสอบ แบบสอบ และคุณลักษณะของผู้สอบมีความผิดพลาดได้ โดยเฉพาะค่าประมาณ ความเที่ยงและค่าอำนาจจำแนกของข้อสอบจะเพิ่มสูงขึ้น และจะส่งผลต่อความถูกต้องในการประมาณ ค่าความสามารถของผู้สอบด้วย และผลของการประมาณค่าพารามิเตอร์และค่าความเที่ยงที่สูงขึ้นนี้ จะส่งผลต่อการเพิ่มความคลาดเคลื่อนประเภทที่ 1 สูงตามไปด้วย และถ้าเป็นการทดสอบแบบ

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

ปรับเหมาะ (computerize adaptive tests) การคำนวณค่าความสามารถของผู้สอบ จะยุติการสอบก่อนกำหนด เนื่องจากการคำนวณค่าความสามารถของผู้สอบคลาดเคลื่อน (Zenisky et. al., 2003; Smith, 2005) นอกจากนี้การฝ่าฝืนข้อตกลงเบื้องต้นดังกล่าว ยังส่งผลต่อการประมาณค่าความสัมพันธ์ระหว่างการตอบข้อสอบของผู้สอบที่มีความสามารถใกล้เคียงกัน ที่มีค่าสูงเกินความเป็นจริง

การศึกษาวิธีการแก้ปัญหาความไม่เป็นอิสระของข้อสอบมีหลายแนวทาง แนวทางหนึ่งคือการศึกษาวิธีการตรวจสอบความไม่เป็นอิสระของข้อสอบ โดย Yen (1984) ได้เสนอการคำนวณ  $Q_3$  เพื่อใช้ในการตรวจสอบความไม่เป็นอิสระของข้อสอบ โดยการคำนวณค่าเศษเหลือด้วยวิธีค่าสถิติ Yen's  $Q_3$  ซึ่งคำนวณจากสูตร  $Q_{3ij} = r_{d_{dj}} ; d_{ik} = u_{ik} - P_i(\hat{\theta}_k)$  เมื่อ  $Q_{3ij}$  คือ ค่าความไม่เป็นอิสระของข้อสอบข้อที่ i และข้อที่ j  $u_{ik}$  คือ ผลการตอบข้อสอบข้อที่ i คนที่ k (ตอบถูก = 1, ตอบผิด = 0) และ  $P_i(\hat{\theta}_k)$  คือ ความน่าจะเป็นของการตอบข้อสอบข้อที่ i ของผู้สอบคนที่ k  $d_{ik}$  คือ ค่าคะแนนเศษเหลือหรือค่าผลต่างระหว่างผลการตอบข้อสอบข้อที่ i ของคนที่ k และความน่าจะเป็นของการตอบข้อสอบของผู้สอบคนที่ k และ  $r_{d_{dj}}$  เป็นค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนเศษเหลือของข้อสอบข้อที่ i และข้อที่ j

จากสภาพปัญหาของการใช้คะแนนจากแบบสอบเพื่อใช้ในการประมาณค่าพารามิเตอร์พอสรุปได้ว่า การประมาณค่าความเที่ยง ค่าความสามารถของผู้สอบ รวมทั้งสารสนเทศของข้อสอบและแบบสอบ ด้วยการประมาณค่าโดยโมเดลทฤษฎีการตอบสนองข้อสอบทั้ง 3 แบบ ได้แก่ โมเดล 1, 2 และ 3 พารามิเตอร์ ยังคงมีความไม่สมบูรณ์ โดยมีการฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับความไม่เป็นอิสระของข้อสอบ ผลการวิเคราะห์ภายใต้เงื่อนไขของความไม่เป็นอิสระของข้อสอบส่งผลกระทบต่อค่าสถิติต่างๆ ย่อมทำให้การประเมินผลการเรียนของนักเรียนเกิดความคลาดเคลื่อนตามไปด้วย จากเหตุผลดังกล่าวจึงเป็นประเด็นสำคัญที่ควรศึกษาถึงผลของระดับความไม่เป็นอิสระของข้อสอบ โมเดลทฤษฎีการตอบสนองข้อสอบ จำนวนข้อสอบและขนาดกลุ่มตัวอย่าง ที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ค่าความสามารถของผู้สอบ รวมทั้งค่าสารสนเทศของข้อสอบและแบบสอบ เพื่อจะได้เป็นแนวทางการบริหารการสอบให้มีความถูกต้องเหมาะสมกับสถานการณ์ กล่าวคือ ถ้าแบบสอบใดมีความเสี่ยงที่จะเกิดความไม่เป็นอิสระของข้อสอบมาก ผู้บริหารการสอบสามารถออกแบบการออกข้อสอบและเลือกใช้เงื่อนไขของการวิเคราะห์ข้อสอบที่สามารถลดความคลาดเคลื่อนของการประมาณค่าได้มากที่สุด นอกจากนี้การวิเคราะห์และการประมาณค่าพารามิเตอร์ต่าง ๆ ตามเงื่อนไขของความไม่เป็นอิสระของข้อสอบที่แตกต่างกัน จะทำให้ครูหรือบุคลากรด้านการศึกษาสามารถเลือกแนวทางในการวิเคราะห์ข้อสอบได้อย่างเหมาะสมโดยพิจารณาจากเงื่อนไขของสถานการณ์ การประมาณค่าพารามิเตอร์ที่มีความคลาดเคลื่อนน้อยที่สุด นอกจากนี้ผลจากการศึกษายังสามารถนำไปกำหนดเกณฑ์ในการประเมินเพื่อตัดสินผลการเรียนให้มีความถูกต้องยุติธรรมมากขึ้น โดยใน

การศึกษาครั้งนี้ผู้วิจัยได้เลือกใช้เทคนิควิธีการจำลองข้อมูลรวมทั้งศึกษาจากข้อมูลจริงประกอบกัน เพื่อให้องค์ความรู้ที่ได้จากทฤษฎีสามารถนำไปใช้ปฏิบัติได้จริง

## วัตถุประสงค์การวิจัย

1. เพื่อวิเคราะห์ผลของระดับความไม่แน่นอนอิสระของข้อสอบด้วยโมเดลทฤษฎีการตอบสนองข้อสอบ ที่มีจำนวนข้อสอบและขนาดกลุ่มตัวอย่างแตกต่างกัน ที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ค่าความสามารถของผู้สอบ และสารสนเทศของแบบสอบ
2. เพื่อวิเคราะห์ลักษณะความไม่แน่นอนอิสระของข้อสอบในแบบสอบวิชาคณิตศาสตร์และภาษาอังกฤษ

## สมมติฐานของการวิจัย

Zenisky et. al. (2003) ศึกษาพบว่า การฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระของข้อสอบจะส่งผลกระทบต่อความผิดพลาดในการประมาณค่าสถิติ ได้แก่ การประมาณค่าพารามิเตอร์ของแบบสอบ การประมาณค่าสารสนเทศของข้อสอบและแบบสอบรวมทั้งการประมาณค่าความสามารถของผู้สอบด้วย Ackerman (1987) ศึกษาพบว่า การประมาณค่าพารามิเตอร์มีแนวโน้มที่เกินความเป็นจริง เมื่อระดับความไม่แน่นอนอิสระของข้อสอบเพิ่มขึ้น Linacre (2000 อ้างถึงใน Smith, 2005) ศึกษาพบว่า เมื่อเพิ่มข้อสอบซ้ำในแบบสอบจะมีผลต่อค่าความเที่ยงของแบบสอบและการประมาณค่าความสามารถของผู้สอบที่สูงขึ้นด้วย ศิริชัย กาญจนวาสี (2545) ได้กล่าวถึงความสัมพันธ์ระหว่างความเที่ยงกับความตรงว่า ความเที่ยงเป็นคุณสมบัติด้านความคงที่หรือคงเส้นคงวาของผลการวัดสามารถประมาณค่าได้จากสัดส่วนความแปรปรวนระหว่างคะแนนจริงกับคะแนนที่สังเกตได้ ส่วนความตรงเป็นคุณสมบัติด้านการวัดได้ตรงตามคุณลักษณะที่มุ่งวัด สามารถประมาณค่าได้จากสัดส่วนความแปรปรวนของคะแนนจริงที่ตรงประเด็นกับคุณลักษณะที่มุ่งวัด

Van (1986 อ้างถึงใน กิ่งกาญจน์ เมฆา, 2540) และ Henly et. al. (1989 อ้างถึงใน สุกพล นิลกลาง, 2541) เปรียบเทียบผลของการประมาณค่าพารามิเตอร์ระหว่างโมเดลทฤษฎีการตอบสนองข้อสอบแบบ 1, 2 และ 3 พารามิเตอร์ พบว่า โมเดลทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ต้องการกลุ่มตัวอย่างจำนวนมากในการวิเคราะห์ แต่ในโมเดลแบบ 1 พารามิเตอร์สามารถใช้วิเคราะห์กับกลุ่มตัวอย่างขนาดเล็กได้ และสามารถประมาณค่าความสามารถของผู้สอบมีคุณภาพดีเท่าเทียมกับการประมาณค่าในโมเดลทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ สุวิมล ดิเรกานันท์ (2537) ศึกษาพบว่า การฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกมิติของแบบสอบซึ่งมีความสัมพันธ์กับความเป็นอิสระของข้อสอบด้วย จะส่งผลต่อค่าอำนาจจำแนกมากขึ้น ค่าความยากน้อยลง และค่าการเดามากขึ้น ตลอดจนค่าสารสนเทศของข้อสอบและแบบสอบน้อยลง

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

Jiao et. al. (2005), Ferrara et. al. (1999) ศึกษาพบว่าในวิชาภาษาอังกฤษและคณิตศาสตร์เป็นวิชาที่มีโอกาสเกิดความไม่เป็นอิสระของข้อสอบมากเนื่องจากมักจะมีข้อสอบหลายข้อถามในเนื้อหาเดียวกัน

จากการศึกษารายงานการวิจัยที่ผ่านมาผู้วิจัยจึงกำหนดสมมติฐานดังนี้

1. เมื่อกำหนดระดับความไม่เป็นอิสระของข้อสอบเพิ่มมากขึ้นภายใต้เงื่อนไข จำนวนข้อสอบ ขนาดกลุ่มตัวอย่าง ที่แตกต่างกันส่งผลทำให้ค่าความเที่ยงของแบบสอบมากขึ้น
2. เมื่อกำหนดระดับความไม่เป็นอิสระของข้อสอบเพิ่มมากขึ้นภายใต้เงื่อนไขของโมเดลการตอบสนองข้อสอบที่มีจำนวนข้อสอบและขนาดกลุ่มตัวอย่าง ที่แตกต่างกันส่งผลทำให้ค่าพารามิเตอร์ของข้อสอบ อันได้แก่ ค่าอำนาจจำแนกมากขึ้น ค่าความยากน้อยลง ค่าการเดามากขึ้น ค่าความสามารถมากขึ้น และค่าสารสนเทศของแบบสอบน้อยลง

## วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ 2 ประการ ประการแรก เพื่อเปรียบเทียบค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ ภายใต้เงื่อนไขระดับความไม่เป็นอิสระของข้อสอบที่แตกต่างกัน โมเดลการตอบสนองข้อสอบ จำนวนข้อสอบ และขนาดกลุ่มตัวอย่างที่แตกต่างกัน ใช้วิธีดำเนินการวิจัยเชิงทดลองโดยการจำลองข้อมูล ประการที่สอง เพื่อศึกษาลักษณะความไม่เป็นอิสระของข้อสอบจากข้อมูลจริง โดยใช้วิชาคณิตศาสตร์และภาษาอังกฤษในระดับช่วงชั้นที่ 2 ระดับประถมศึกษาปีที่ 5 เขตพื้นที่การศึกษาจังหวัดสุพรรณบุรี เขต 1

### 1. วิธีการจำลองข้อมูล (simulation method)

การจำลองข้อมูลใช้โปรแกรมภาษา FORTRAN รุ่น Fortran Power Station Version 4.0 เพื่อใช้ในการจำลองข้อมูล โปรแกรม LDID (A Computer Program for Local Dependence Indices for Dichotomous Items<sup>1</sup>) และ BILOG-MG ใช้เพื่อตรวจสอบระดับความไม่เป็นอิสระของข้อสอบตามเงื่อนไขที่กำหนด อันประกอบด้วย ระดับความไม่เป็นอิสระของข้อสอบ จำนวนข้อสอบ และขนาดกลุ่มตัวอย่าง ด้วยโมเดลทฤษฎีการตอบสนองข้อสอบ 3 โมเดล ได้แก่ โมเดล 1 พารามิเตอร์ (1PL) โมเดล 2 พารามิเตอร์ (2PL) และโมเดล 3 พารามิเตอร์ (3PL) ซึ่งในแต่ละเงื่อนไขมีระดับที่แตกต่างกันดังนี้

1.1 ระดับความไม่เป็นอิสระของข้อสอบ (LID) ประกอบด้วย 9 ระดับโดยแบ่งเป็นกลุ่มระดับต่ำ (mild = L) 3 ระดับ กลุ่มระดับปานกลาง (medium = M) 3 ระดับ และกลุ่มระดับสูง (strong = H) 3 ระดับ (Spray and Ackerman (1987); Ackerman (1987 อ้างถึงใน Smith, 2005)

1.2 จำนวนข้อสอบ (items) แบ่งเป็น 3 ระดับ ได้แก่ 30 ข้อ 50 ข้อ และ 80 ข้อ

1.3 ขนาดกลุ่มตัวอย่าง (sample size) แบ่งเป็น 3 ระดับ ได้แก่ 400, 800 และ 1,200 คน

ดังนั้นจำนวนเงื่อนไขของการวิเคราะห์ข้อมูลที่ได้จากการจำลองข้อมูลจะมีจำนวนทั้งสิ้นเท่ากับ 81 เงื่อนไข และจำลองข้อมูลในแต่ละเงื่อนไข จำนวน 1,000 รอบ (replication) สำหรับขั้นตอนในการจำลองข้อมูลและการตรวจสอบความไม่เป็นอิสระของข้อสอบมีดังนี้

**ขั้นที่ 1** จำลองข้อมูล จากการทำข้อสอบแต่ละชุดด้วยวิธีการแจกแจงปกติหลายตัวแปร (multivariate normal distribution) ที่มีความสัมพันธ์กัน เมื่อจำลองคะแนนที่เป็นมาตรฐานของแต่ละตัวแปรแล้ว จึงแปลงคะแนนมาตรฐานที่ได้เป็น คะแนน 0, 1 ในแต่ละเงื่อนไขจำลองข้อมูล จำนวน 1,000 รอบ (Replications) ในแต่ละรอบได้กำหนดให้โปรแกรมคำนวณค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ค่าความสามารถ และค่าสารสนเทศของแบบสอบ และกำหนดให้โปรแกรมจำลองข้อมูลให้มีระดับความไม่เป็นอิสระของข้อสอบ 9 ระดับโดยแบ่งเป็นกลุ่มย่อย 3 ระดับ ได้แก่ กลุ่มระดับต่ำ (mild = L) กลุ่มระดับปานกลาง (medium = M) และกลุ่มระดับสูง (strong = H) โดยกำหนดจำนวนข้อสอบที่มีความไม่เป็นอิสระของข้อสอบในแต่ละเงื่อนไขเป็น 20% ของการจำลองแบบสอบแต่ละชุด

**ขั้นที่ 2** ตรวจสอบความไม่เป็นอิสระของข้อสอบด้วยสถิติ Yen's  $Q_3$  โดยใช้โปรแกรม LDID ในแต่ละเงื่อนไข ดังปรากฏผลดังนี้

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
 ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

ตาราง 1 ค่าต่ำสุด-สูงสุดของระดับความไม่เป็นอิสระของข้อสอบ ( $|Q_3|$ ) ในกลุ่มระดับต่ำ (L) กลุ่มระดับปานกลาง (M) และกลุ่มระดับสูง (H) ภายใต้เงื่อนไขที่แตกต่างกัน

โมเดล	จำนวนผู้สอบ	ค่าระดับความไม่เป็นอิสระของข้อสอบ ( $ Q_3 $ )*								
		30 ข้อ			50 ข้อ			80 ข้อ		
		L	M	H	L	M	H	L	M	H
<b>1PL</b>										
	400	.0476-	.0530-	.0664-	.0414-	.0489-	.0657-	.0422-	.0492-	.0672-
		.0508	.0611	.0859	.0458	.0582	.0900	.0450	.0598	.0875
	800	.0339-	.0439-	.0600-	.0316-	.0407-	.0575-	.0304-	.0397-	.0597-
		.0403	.0537	.0782	.0363	.0509	.0806	.0350	.0518	.0811
	1,200	.0314-	.0419-	.0583-	.0269-	.0367-	.0547-	.0249-	.0346-	.0553-
		.0374	.0525	.0761	.0321	.0476	.0773	.0301	.0472	.0774
<b>2PL</b>										
	400	.0469-	.0533-	.0628-	.0414-	.0478-	.0509-	.0422-	.0483-	.0515-
		.0506	.0603	.0569	.0454	.0516	.0539	.0447	.0517	.0530
	800	.0336-	.0450-	.0547-	.0317-	.0410-	.0429-	.0303-	.0393-	.0420-
		.0403	.0538	.0397	.0363	.0470	.0479	.0349	.0435	.0430
	1,200	.0312-	.0412-	.0553-	.0270-	.0371-	.0385-	.0249-	.0345-	.0398-
		.0369	.0545	.0386	.0323	.0450	.0453	.0300	.0400	.0419
<b>3PL</b>										
	400	.0464-	.0533-	.0628-	.0413-	.0472-	.0476-	.0419-	.0477-	.0496-
		.0506	.0603	.0569	.0451	.0480	.0488	.0444	.0493	.0515
	800	.0331-	.0441-	.0514-	.0315-	.0406-	.0440-	.0302-	.0387-	.0415-
		.0395	.0509	.0472	.0361	.0444	.0449	.0347	.0413	.0432
	1,200	.0308-	.0409-	.0460-	.0267-	.0472-	.0476-	.0247-	.0341-	.0381-
		.0364	.0511	.0513	.0451	.0480	.0488	.0298	.0379	.0404

\*พิสัยของค่าระดับความไม่เป็นอิสระของข้อสอบ ( $|Q_3|$ ) อยู่ระหว่าง -1.0 ถึง +1.0

จากตาราง 1 ค่าระดับความไม่เป็นอิสระของข้อสอบ ( $|Q_3|$ ) ซึ่งแบ่งเป็นกลุ่มระดับต่ำ (L) กลุ่มระดับปานกลาง (M) และกลุ่มระดับสูง (H) ที่ได้จากการจำลองข้อมูลภายใต้เงื่อนไข โมเดล ทฤษฎีการตอบสนองของข้อสอบ จำนวนข้อสอบและจำนวนผู้สอบที่แตกต่างกันด้วยโปรแกรม LDID พบว่า ค่าระดับความไม่เป็นอิสระของข้อสอบมีแนวโน้มเพิ่มมากขึ้นเมื่อจำนวนข้อสอบเพิ่มขึ้น แต่เมื่อจำนวนผู้สอบเพิ่มมากขึ้น ค่าระดับความไม่เป็นอิสระของข้อสอบมีแนวโน้มน้อยลง และเมื่อ



วิเคราะห์ด้วยโมเดลทฤษฎีการตอบสนองข้อสอบแบบ 1, 2 และ 3 พารามิเตอร์ พบว่า ค่าระดับความไม่เป็นอิสระของข้อสอบมีแนวโน้มน้อยลง ตามลำดับ

**ขั้นที่ 3** นำข้อมูลที่จำลองได้ในแต่ละเงื่อนไข คำนวณค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ค่าความสามารถ ค่าสารสนเทศของแบบสอบ

## 2. วิธีการศึกษาข้อมูลจริง (actual data method)

จากการศึกษาของ Yen (1993 อ้างถึงใน Smith, 2005) และ Ferrara et. al. (1999) ศึกษา พบว่า ความไม่เป็นอิสระของข้อสอบวัดความเข้าใจในการอ่านที่วัดหลายข้อในเนื้อหาเดียวกัน และพบความไม่เป็นอิสระของข้อสอบมากขึ้นในวิชาคณิตศาสตร์ที่ใช้ตาราง กราฟ แผนภูมิร่วมกัน Keller et. al. (2003) ศึกษาการประเมินวิธีการให้คะแนนต่างกันแบบสอบที่มีความไม่เป็นอิสระของข้อสอบโดยศึกษาในรายวิชาการอ่านเพื่อความเข้าใจและวิชาคณิตศาสตร์ ดังนั้นผู้วิจัยจึงสนใจศึกษาข้อมูลจริงเพื่อสำรวจลักษณะความไม่เป็นอิสระของข้อสอบ ซึ่งผู้วิจัยใช้แบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาภาษาอังกฤษและคณิตศาสตร์ในช่วงชั้นที่ 2 ระดับชั้นประถมศึกษาปีที่ 5 ปีการศึกษา 2550 เขตพื้นที่การศึกษาจังหวัดสุพรรณบุรี เขต 1 เนื่องจากเป็นช่วงชั้นที่จะจำแนกความสามารถของนักเรียนได้รวมทั้ง นักเรียนได้ผ่านการเรียนเนื้อหาในหลักสูตรมาแล้ว

## ผลการวิเคราะห์ข้อมูลโดยใช้ข้อมูลจำลอง (simulation)

### 1. การวิเคราะห์ค่าความเที่ยง

เมื่อพิจารณาผลการวิเคราะห์ค่าความเที่ยง เมื่อระดับความไม่เป็นอิสระของข้อสอบเพิ่มมากขึ้น ภายใต้เงื่อนไขของจำนวนข้อสอบและจำนวนผู้สอบที่เท่ากัน พบว่า ค่าความเที่ยงจะเพิ่มขึ้นอย่างต่อเนื่อง โดยเมื่อใช้จำนวนข้อสอบ 30 ข้อ จำนวนผู้สอบ 400, 800 และ 1,200 คน ค่าความเที่ยงต่ำสุด-สูงสุด มีค่าเท่ากับ 0.671-0.739, 0.672-0.740 และ 0.672-0.740 ตามลำดับ เมื่อใช้จำนวนข้อสอบ 50 ข้อ จำนวนผู้สอบ 400, 800 และ 1,200 คน ค่าความเที่ยงต่ำสุด-สูงสุด มีค่าเท่ากับ 0.773-0.828, 0.774-0.828 และ 0.773-0.828 ตามลำดับ และ เมื่อใช้จำนวนข้อสอบ 80 ข้อ จำนวนผู้สอบ 400, 800 และ 1,200 คน ค่าความเที่ยงต่ำสุด-สูงสุด มีค่าเท่ากับ 0.845-0.886 ในทุกจำนวนผู้สอบที่แตกต่างกัน ดังแสดงในตาราง 2

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
 ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

ตาราง 2 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าความเที่ยง เมื่อระดับความไม่เป็นอิสระของข้อสอบแตกต่างกันโดยใช้จำนวนผู้สอบ 400 800 และ 1,200 คน ทำข้อสอบ 30, 50 และ 80 ข้อ

N*	30 ข้อ			N	50 ข้อ			N	80 ข้อ		
	$ Q_3 $	$KR_{20}$	$SD_{KR_{20}}$		$ Q_3 $	$KR_{20}$	$SD_{KR_{20}}$		$ Q_3 $	$KR_{20}$	$SD_{KR_{20}}$
400	0.0476	0.671	0.023	400	0.0476	0.773	0.016	400	0.0476	0.845	0.011
	0.0482	0.680	0.023		0.0482	0.780	0.016		0.0482	0.850	0.011
	0.0508	0.688	0.022		0.0508	0.786	0.015		0.0508	0.855	0.010
	0.0530	0.695	0.021		0.0530	0.793	0.014		0.0530	0.860	0.010
	0.0552	0.703	0.021		0.0552	0.799	0.014		0.0552	0.865	0.010
	0.0611	0.711	0.020		0.0611	0.806	0.014		0.0611	0.870	0.009
	0.0664	0.719	0.019		0.0664	0.812	0.013		0.0664	0.874	0.009
	0.0748	0.728	0.019		0.0748	0.819	0.012		0.0748	0.880	0.008
	0.0859	0.739	0.018		0.0859	0.828	0.012		0.0859	0.886	0.008
800	0.0339	0.672	0.017	800	0.0339	0.774	0.011	800	0.0339	0.845	0.008
	0.0363	0.680	0.016		0.0363	0.780	0.011		0.0363	0.850	0.008
	0.0403	0.688	0.016		0.0403	0.787	0.010		0.0403	0.855	0.007
	0.0439	0.696	0.015		0.0439	0.793	0.010		0.0439	0.860	0.007
	0.0477	0.704	0.015		0.0477	0.800	0.010		0.0477	0.865	0.007
	0.0537	0.712	0.014		0.0537	0.806	0.009		0.0537	0.869	0.007
	0.0600	0.720	0.014		0.0600	0.812	0.009		0.0600	0.874	0.006
	0.0680	0.729	0.013		0.0680	0.820	0.009		0.0680	0.879	0.006
	0.0782	0.740	0.013		0.0782	0.828	0.008		0.0782	0.886	0.006
1,200	0.0314	0.672	0.014	1,200	0.0314	0.773	0.009	1,200	0.0314	0.845	0.006
	0.0332	0.680	0.013		0.0332	0.780	0.009		0.0332	0.850	0.006
	0.0374	0.688	0.013		0.0374	0.787	0.009		0.0374	0.855	0.006
	0.0419	0.696	0.012		0.0419	0.793	0.008		0.0419	0.860	0.006
	0.0460	0.704	0.012		0.0460	0.799	0.008		0.0460	0.865	0.005
	0.0525	0.712	0.012		0.0525	0.806	0.008		0.0525	0.870	0.005
	0.0583	0.720	0.011		0.0583	0.812	0.007		0.0583	0.874	0.005
	0.0659	0.729	0.011		0.0659	0.819	0.007		0.0659	0.880	0.005
	0.0761	0.740	0.010		0.0761	0.828	0.006		0.0761	0.886	0.004

\*N หมายถึง จำนวนผู้สอบ,  $|Q_3|$  หมายถึง ค่าดัชนีวัดระดับความไม่เป็นอิสระของข้อสอบ

## 2. ผลการวิเคราะห์สถิติพื้นฐานโมเดล 1, 2 และ 3 พารามิเตอร์ (1PL, 2PL, 3PL)

จากการวิเคราะห์ข้อมูลด้วยวิธีการจำลองสถานการณ์ตามเงื่อนไขการทดสอบที่กำหนด สามารถกล่าวเป็นภาพรวม โดยสรุปดังนี้

การวิเคราะห์โมเดลทฤษฎีการตอบสนองข้อสอบ พบว่า (1) ในการวิเคราะห์ด้วยโมเดล 1 พารามิเตอร์ เมื่อระดับความไม่เป็นที่พอใจของข้อสอบเพิ่มสูงขึ้น ค่าความยากและค่าความสามารถไม่แตกต่างกัน แต่ค่าสารสนเทศของแบบสอบส่วนใหญ่มีแนวโน้มคงที่ (2) เมื่อวิเคราะห์ด้วยโมเดล 2 พารามิเตอร์ พบว่า ส่วนใหญ่ค่าความยากลดลง แต่ค่าอำนาจจำแนก ค่าความสามารถและสารสนเทศของแบบสอบมีแนวโน้มสูงขึ้น โดยส่วนใหญ่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ (3) เมื่อวิเคราะห์โดยใช้โมเดล 3 พารามิเตอร์ พบว่า ค่าความยาก ค่าความสามารถและสารสนเทศของแบบสอบมีแนวโน้มลดลง แต่สำหรับค่าอำนาจจำแนก ค่าการเดาส่วนใหญ่มีแนวโน้มสูงขึ้นและแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และผลการวิเคราะห์สถิติพื้นฐานจากการจำลองข้อมูลโมเดล 1 พารามิเตอร์ โมเดล 2 พารามิเตอร์ และโมเดล 3 พารามิเตอร์ ดังรายละเอียดในตาราง 3

ตาราง 3 สรุปแนวโน้มการเปลี่ยนแปลงค่า ค่าพารามิเตอร์ของข้อสอบ ค่าความสามารถของผู้สอบ และค่าสารสนเทศของแบบสอบ เมื่อระดับความไม่เป็นที่พอใจของข้อสอบเพิ่มขึ้น

โมเดล	จำนวน ข้อสอบ	จำนวน ผู้สอบ	a'	b	c	$\theta$	TIF <sub>max</sub>
1PL	30 ข้อ	400 คน	-	คงที่	-	คงที่	คงที่
		800 คน	-	คงที่	-	คงที่	คงที่ส่วนใหญ่ (เพิ่มขึ้นหรือ น้อยลงบางส่วน เมื่อ $ Q_3 $ เพิ่มขึ้น)
		1,200 คน	-	คงที่	-	คงที่	คงที่ส่วนใหญ่
	50 ข้อ	400 คน	-	คงที่	-	คงที่	คงที่
		800 คน	-	คงที่	-	คงที่	คงที่
		1,200 คน	-	คงที่	-	คงที่	คงที่ส่วนใหญ่

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
 ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

ตาราง 3 (ต่อ)

โมเดล	จำนวน ข้อสอบ	จำนวน ผู้สอบ	a*	b	c	$\theta$	TIF <sub>max</sub>
	80 ข้อ	400 คน	-	คงที่	-	คงที่	คงที่ส่วนใหญ่
		800 คน	-	คงที่	-	คงที่	คงที่ส่วนใหญ่
		1,200 คน	-	คงที่	-	คงที่	คงที่ส่วนใหญ่
<b>2PL</b>							
	30 ข้อ	400 คน	สูงขึ้น (แนวโน้มเพิ่มขึ้น เมื่อ $ Q_3 $ มากขึ้น)	ไม่คงที่ (เพิ่มขึ้นหรือ ลดลงไม่คงที่เมื่อ $ Q_3 $ มากขึ้น)	-	ไม่คงที่	สูงขึ้น
		800 คน	สูงขึ้น	ไม่คงที่	-	ไม่คงที่	สูงขึ้น
		1200 คน	สูงขึ้น	ไม่คงที่	-	ไม่คงที่	สูงขึ้น
	50 ข้อ	400 คน	สูงขึ้น	คงที่และต่ำลง	-	คงที่และ สูงขึ้น	สูงขึ้น
		800 คน	สูงขึ้น	คงที่และต่ำลง	-	คงที่และ สูงขึ้น	สูงขึ้น
		1,200 คน	สูงขึ้น	คงที่และต่ำลง	-	คงที่และ สูงขึ้น	สูงขึ้น
	80 ข้อ	400 คน	สูงขึ้น	คงที่และต่ำลง	-	คงที่และ สูงขึ้น	สูงขึ้น
		800 คน	สูงขึ้น	คงที่และต่ำลง	-	คงที่และ สูงขึ้น	สูงขึ้น
		1,200 คน	สูงขึ้น	คงที่และต่ำลง	-	คงที่และ สูงขึ้น	สูงขึ้น
<b>3PL</b>							
	30 ข้อ	400 คน	คงที่และต่ำลง	คงที่และต่ำลง	คงที่และ ต่ำลง	ต่ำลงและ คงที่	ต่ำลง
		800 คน	คงที่และต่ำลง	ต่ำลงและคงที่	ไม่คงที่	ต่ำลงและ คงที่	ต่ำลง
		1,200 คน	สูงขึ้นและคงที่	ต่ำลงและคงที่	สูงขึ้นและ ต่ำลง	ต่ำลงและ คงที่	ต่ำลง
	50 ข้อ	400 คน	คงที่และต่ำลง	คงที่และต่ำลง	คงที่และ ต่ำลง	ต่ำลง	ต่ำลง

ตาราง 3 (ต่อ)

โมเดล	จำนวนข้อสอบ	จำนวนผู้สอบ	a <sup>*</sup>	b	c	$\theta$	TIF <sub>max</sub>
		800 คน	สูงขึ้นและต่ำลง	ต่ำลงและคงที่	สูงขึ้นและคงที่	ต่ำลง	ต่ำลง
		1,200 คน	สูงขึ้น	ต่ำลงและคงที่	สูงขึ้นและคงที่	คงที่และสูงขึ้น	คงที่และต่ำลง
80 ข้อ		400 คน	คงที่และต่ำลง	ต่ำลง	สูงขึ้นและต่ำลง	ต่ำลง	ต่ำลง
		800 คน	ไม่คงที่	ต่ำลงและคงที่	สูงขึ้นและคงที่	ต่ำลง	ต่ำลง
		1,200 คน	สูงขึ้น	ต่ำลงและคงที่	สูงขึ้นและต่ำลง	ต่ำลง	สูงขึ้นและต่ำลง

\* a หมายถึง ค่าอำนาจจำแนก, b หมายถึง ค่าความยาก, c หมายถึง ค่าการเดา,  $\theta$  หมายถึง ค่าความสามารถ, TIF<sub>max</sub> หมายถึง ค่าสารสนเทศของแบบสอบสูงสุด

### ผลการวิเคราะห์ข้อมูลโดยการใช้ข้อมูลจริง (actual data)

การวิเคราะห์ข้อมูลจริงจะใช้ข้อมูลคะแนนของนักเรียนระดับชั้นประถมศึกษาปีที่ 5 ช่วงชั้นที่ 2 สำนักงานเขตพื้นที่การศึกษาสุพรรณบุรี เขต 1 ในรายวิชาภาษาอังกฤษและคณิตศาสตร์ว่ามีระดับค่าความไม่เป็นอิสระของข้อสอบมากน้อยเพียงใด เนื่องจากเป็นวิชาที่มีโอกาสเกิดความไม่เป็นอิสระของข้อสอบสูง กล่าวคือ เนื้อหาที่ใช้ในการออกข้อสอบมักเป็นการถามที่ใช้เนื้อหาหรือตารางร่วมกันในการออกข้อสอบ (Jiao et. al., 2005; Ferara et. al., 1999)

สำหรับโครงสร้างเนื้อหาของแบบสอบวิชาคณิตศาสตร์ประกอบด้วย สารการเรียนรู้จำนวนและการดำเนินงาน 17 ข้อ การวัด 7 ข้อ เรขาคณิต 8 ข้อ พีชคณิต 5 ข้อ และการวิเคราะห์ข้อมูลและความน่าจะเป็น 3 ข้อ รวมทั้งฉบับ 40 ข้อ และแบบสอบวิชาภาษาอังกฤษ ประกอบด้วย สารการเรียนรู้ภาษาเพื่อการสื่อสาร 2 ข้อ ภาษาและวัฒนธรรม 3 ข้อ ภาษากับความสัมพันธ์กับการเรียนรู้อื่น 4 ข้อ และภาษากับความสัมพันธ์ชุมชนและโลก 12 ข้อ รวมทั้งฉบับ 40 ข้อ

การวิเคราะห์ค่าสถิติพื้นฐานของดัชนีวัดความไม่เป็นอิสระของข้อสอบได้แก่ ค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ค่าความเบ้ ค่าความโด่ง ของดัชนีวัดความไม่เป็นอิสระของข้อสอบ (ค่า  $Q_3$  และ  $|Q_3|$ )

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
 ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

ตาราง 4 ผลการวิเคราะห์ความไม่เป็นอิสระของข้อสอบวิชาภาษาอังกฤษ จำนวนทั้งหมด 40 ข้อ

สถิติ ทดสอบ	จำนวน ข้อสอบ	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย	ค่าเบี่ยงเบน มาตรฐาน	ค่าความ เบ้	ค่าความ โด่ง
โมเดล 1 พารามิเตอร์ (1PL)							
$Q_3$	780	-0.174	0.185	-0.019	0.039	0.785	3.480
$ Q_3 $	780	.000	.185	.0331	.027	1.576	3.770
โมเดล 2 พารามิเตอร์ (2PL)							
$Q_3$	780	-0.164	0.147	-0.013	0.034	0.225	1.372
$ Q_3 $	780	0.000	0.164	0.028	0.022	1.348	3.154
โมเดล 3 พารามิเตอร์ (3PL)							
$Q_3$	780	-0.170	0.143	-0.007	0.029	-0.116	3.924
$ Q_3 $	780	0.000	0.170	0.022	0.021	2.302	8.105

$$E(Q_3) = -1/n-1 = -.025$$

การวิเคราะห์ความไม่เป็นอิสระของข้อสอบของข้อมูลจริงในวิชาภาษาอังกฤษ พบว่า ดัชนี  $Q_3$  ทั้งโมเดลการวิเคราะห์แบบ 1, 2 และ 3 พารามิเตอร์ พบว่า มีค่า  $Q_3$  มีค่าเฉลี่ยเท่ากับ -0.019, -0.013 และ -0.007 ซึ่งเข้าใกล้ค่าคาดหวัง ( $E(Q_3) = -1/n-1 = -.025$ ) แสดงว่าข้อสอบค่อนข้างมีความเป็นอิสระจากกัน และเมื่อพิจารณาระหว่างโมเดลการวิเคราะห์ พบว่า โมเดล 1 พารามิเตอร์ จะให้ค่าดัชนีความไม่เป็นอิสระของข้อสอบมากที่สุด รองลงมา คือ โมเดล 2 และ 3 พารามิเตอร์ ตามลำดับ โดยค่าดัชนี  $Q_3$  มีค่าเฉลี่ยเท่ากับ -0.019, -0.013 และ -0.007 และ  $|Q_3|$  มีค่าเฉลี่ยเท่ากับ 0.033, 0.028 และ 0.022 ตามลำดับ

เมื่อพิจารณาค่าการกระจายของค่าดัชนีความไม่เป็นอิสระของข้อสอบ พบว่า ค่าดัชนีความไม่เป็นอิสระของข้อสอบทั้งการวิเคราะห์โมเดล 1, 2 และ 3 พารามิเตอร์ มีลักษณะเบ้ทางขวา กล่าวคือ ค่าดัชนีส่วนใหญ่มีค่าต่ำกว่าค่าเฉลี่ย สำหรับค่าความโด่ง พบว่า ค่าดัชนี  $Q_3$  มีลักษณะโด่งมากกว่าปกติ กล่าวคือค่าดัชนีส่วนใหญ่จะเกาะกลุ่มอยู่รอบ ๆ ค่าเฉลี่ย

ตาราง 5 ผลการวิเคราะห์ความไม่เป็นอิสระของข้อสอบวิชาคณิตศาสตร์ จำนวนทั้งหมด 40 ข้อ

สถิติ ทดสอบ	จำนวน ข้อสอบ	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย	ค่าเบี่ยงเบน มาตรฐาน	ค่าความ เบ้	ค่าความ โด่ง
โมเดล 1 พารามิเตอร์ (1PL)							
$Q_3$	780	-0.145	0.138	-0.019	0.040	0.375	0.733
$ Q_3 $	780	0.000	0.145	0.035	0.026	0.991	0.915
โมเดล 2 พารามิเตอร์ (2PL)							
$Q_3$	780	-0.105	0.129	-0.015	0.031	0.215	0.851
$ Q_3 $	780	0.000	0.129	0.026	0.021	1.061	1.012
โมเดล 3 พารามิเตอร์ (3PL)							
$Q_3$	780	-0.116	0.137	-0.010	0.027	0.101	2.102
$ Q_3 $	780	0.000	0.137	0.022	0.019	1.655	4.363

$$E(Q_3) = -1/n-1 = -.025$$

สำหรับวิชาคณิตศาสตร์พบว่า เมื่อพิจารณาข้อนี้  $Q_3$  ทั้งโมเดลการวิเคราะห์แบบ 1, 2 และ 3 พารามิเตอร์ พบว่า มีค่าเฉลี่ยเท่ากับ -0.019, -0.015 และ -0.010 ซึ่งต่ำกว่าศูนย์เล็กน้อย และเข้าใกล้ค่าคาดหวัง ( $E(Q_3) = -1/n-1 = -.025$ ) แสดงว่าข้อสอบมีความไม่เป็นอิสระจากกันเล็กน้อย และเมื่อพิจารณาระหว่างโมเดลการวิเคราะห์ พบว่า โมเดล 1 พารามิเตอร์จะให้ค่าดัชนีความไม่เป็นอิสระของข้อสอบมากที่สุด รองลงมา คือ โมเดล 2 และ 3 พารามิเตอร์ ตามลำดับ โดยค่าดัชนี  $Q_3$  มีค่าเฉลี่ยเท่ากับ -0.019, -0.015 และ -0.010 และ  $|Q_3|$  มีค่าเฉลี่ยเท่ากับ 0.035, 0.026 และ 0.022 ตามลำดับ

เมื่อพิจารณาค่าการกระจายของค่าดัชนีความไม่เป็นอิสระของข้อสอบ พบว่า ค่าดัชนีความไม่เป็นอิสระของข้อสอบทั้งการวิเคราะห์โมเดล 1, 2 และ 3 พารามิเตอร์ มีลักษณะเบ้ทางขวา กล่าวคือ ค่าดัชนีส่วนใหญ่มีค่าต่ำกว่าค่าเฉลี่ย สำหรับค่าความโด่ง พบว่า ค่า  $Q_3$  มีลักษณะใกล้เคียงโค้งปกติ แสดงว่า ค่าดัชนีมีลักษณะกระจาย

ผลการวิเคราะห์ค่าความเที่ยงในแบบสอบวิชาภาษาอังกฤษ จำนวน 40 ข้อ และคณิตศาสตร์ จำนวน 40 ข้อ ซึ่งในแบบสอบประกอบด้วยข้อสอบที่มีทั้งความเป็นอิสระและไม่เป็นอิสระของข้อสอบ ได้ค่าความเที่ยง เท่ากับ .7195 และ .7236 ตามลำดับ จากนั้นวิเคราะห์ค่าดัชนีวัดความไม่เป็นอิสระ

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

ของข้อสอบ และดำเนินการคัดข้อสอบที่มีความไม่เป็นอิสระที่สูงออกจำนวน 5 ข้อ ในแบบสอบวิชาภาษาอังกฤษและคณิตศาสตร์ แล้ววิเคราะห์ค่าความเที่ยง พบว่ามีค่าเท่ากับ .6528 และ .6564 ตามลำดับ และคัดข้อสอบที่มีความเป็นอิสระของข้อสอบออกจำนวน 5 ข้อ ในแบบสอบวิชาภาษาอังกฤษและคณิตศาสตร์ วิเคราะห์ค่าความเที่ยง มีค่าเท่ากับ .7195 และ .7236 ตามลำดับ

## อภิปรายผล ข้อเสนอแนะและการนำผลการวิจัยไปใช้

### 1. อภิปรายผลการวิจัย

ในส่วนของการอภิปรายผลผู้วิจัยจะพิจารณาความสอดคล้องของผลการวิจัยกับสมมติฐานที่กำหนดขึ้นจากทฤษฎีและเอกสารรายงานวิจัยที่เกี่ยวข้องซึ่งกำหนดไว้ดังนี้

1) เมื่อกำหนดระดับความไม่เป็นอิสระของข้อสอบเพิ่มสูงขึ้นภายใต้เงื่อนไขของ จำนวนข้อสอบ ขนาดกลุ่มตัวอย่าง ที่แตกต่างกันจะส่งผลทำให้ค่าความเที่ยงของแบบสอบสูงขึ้น

2) เมื่อกำหนดระดับความไม่เป็นอิสระของข้อสอบเพิ่มสูงขึ้นภายใต้เงื่อนไขของโมเดลการตอบสนองข้อสอบ (IRT Model) จำนวนข้อสอบ ขนาดกลุ่มตัวอย่าง ที่แตกต่างกันจะส่งผลทำให้ค่าพารามิเตอร์ของข้อสอบ อันได้แก่ ค่าอำนาจจำแนกสูงขึ้น ค่าความยากต่ำลง ค่าการเดาเพิ่มขึ้น ค่าความสามารถเพิ่มขึ้น และค่าสารสนเทศของแบบสอบลดลง

#### 1.1 ผลการวิเคราะห์ค่าความเที่ยง

จากการศึกษาของ Smith (2005) ศึกษาพบว่า เมื่อเพิ่มข้อสอบซ้ำในแบบสอบจะมีผลต่อค่าความเที่ยงของแบบสอบที่สูงขึ้น ซึ่งจากการศึกษาในครั้งนี้พบว่าความเที่ยงมีแนวโน้มสูงขึ้นอย่างต่อเนื่องตามระดับของความไม่เป็นอิสระของข้อสอบ นอกจากนี้ Ferrara, et. al. (1999 อ้างถึงใน Smith, 2005); Yen (1993 อ้างถึงใน Smith, 2005) ศึกษาข้อสอบที่มีความไม่เป็นอิสระจากกันเมื่อถามเนื้อหาร่วมกัน มีผลทำให้ค่าความเที่ยงของแบบสอบเพิ่มสูงขึ้น และจากการศึกษาของ Reese, (1999) ศึกษาทฤษฎีการวัดแบบดั้งเดิมกับความไม่เป็นอิสระของข้อสอบในแบบสอบ LSAT (Law School Admission Test questions) พบว่า เมื่อเกิดความไม่เป็นอิสระของข้อสอบสูงสุด ภายใต้เงื่อนไขจำนวนข้อสอบและจำนวนผู้สอบที่เท่ากัน ค่าความเที่ยง ( $\alpha$  reliability) ของแบบสอบสูงกว่าความเป็นจริง และยังพบว่า ค่าสัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียล เพิ่มสูงขึ้นเมื่อความไม่เป็นอิสระของข้อสอบอยู่ในระดับสูง ซึ่งลักษณะของค่าความเที่ยงที่เพิ่มขึ้นนั้นพบว่าเพิ่มขึ้นอย่างต่อเนื่องเช่นกัน กล่าวคือ เมื่อความไม่เป็นอิสระของข้อสอบเป็นระดับศูนย์ ต่ำ ปานกลาง และสูง ค่าความเที่ยง มีค่าเท่ากับ .91, .92, .93 และ .97 จะเห็นว่า มีลักษณะการเพิ่มขึ้นที่คล้ายคลึงกับการศึกษาในครั้งนี้ด้วย โดยเมื่อระดับความไม่เป็นอิสระของข้อสอบเพิ่มมากขึ้นภายใต้เงื่อนไขจำนวนข้อสอบเท่ากันและจำนวนผู้สอบเท่ากัน จะพบว่าค่าความเที่ยงเพิ่มสูงขึ้นเช่นกัน และจากการ



วิเคราะห์ข้อมูลจริงผู้วิจัยได้นำข้อสอบที่มีระดับความไม่เป็นอิสระสูงออกจากแบบสอบทั้งภาษาอังกฤษและคณิตศาสตร์ จำนวน 5 ข้อ พบว่า ได้ค่าความเที่ยงเท่ากับ .6528 และ .6564 ตามลำดับ จากนั้นนำข้อสอบที่มีความเป็นอิสระของข้อสอบออก จำนวน 5 ข้อ พบว่า ได้ค่าความเที่ยงเท่ากับ .7039 และ .6752 ตามลำดับ แสดงให้เห็นว่าเมื่อมีข้อสอบที่มีความไม่เป็นอิสระของข้อสอบอยู่ในแบบสอบจะมีผลต่อค่าเที่ยงที่สูงขึ้น สาเหตุที่ค่าความเที่ยงสูงขึ้นเนื่องมาจากข้อสอบที่เพิ่มขึ้นนั้นมีความสัมพันธ์ซึ่งกันและกัน ข้อคำถามถามในเนื้อหาเดียวกันความคงที่ของคำตอบก็จะมีมากขึ้นซึ่งหมายถึงค่าความเที่ยงนั่นเอง

## 1.2 ผลการวิเคราะห์ตามแนวทฤษฎีการตอบสนองข้อสอบ

### 1.2.1 ผลการประมาณค่าพารามิเตอร์ข้อสอบ ค่าความสามารถและสารสนเทศของแบบสอบ

การศึกษาผลการวิจัยที่ผ่านมาในอดีตกับผลการวิจัยมีความสอดคล้องกับการวิจัยผลการศึกษาในครั้งนี้ โดยพบว่า ลักษณะการเพิ่มขึ้นของค่าความเที่ยงสูงขึ้นตามระดับความไม่เป็นอิสระของข้อสอบที่เพิ่มขึ้น กล่าวคือ เมื่อมีการควบคุมจำนวนข้อสอบที่มีความไม่เป็นอิสระ จำนวนข้อสอบทั้งฉบับ และจำนวนผู้สอบให้เท่ากันแต่ระดับของความไม่เป็นอิสระของข้อสอบเพิ่มขึ้นจะส่งผลต่อค่าความเที่ยงให้สูงขึ้น สอดคล้องกับการศึกษาของ Smith (2005) ศึกษาพบว่าเมื่อเพิ่มข้อสอบซ้ำในแบบสอบจะมีผลต่อค่าความเที่ยงของแบบสอบและการประมาณค่าความสามารถที่สูงขึ้นกว่าความเป็นจริง

สำหรับค่าพารามิเตอร์ของข้อสอบซึ่งได้แก่ ค่าความยาก ค่าอำนาจจำแนก ค่าการเดา ค่าความสามารถ และค่าสารสนเทศของแบบสอบ ในบางเงื่อนไขของการทดสอบมีความแตกต่างกันกับการศึกษาที่ผ่านมา เช่น ค่าความสามารถ ในกรณีการวิเคราะห์ด้วยโมเดล 3 พารามิเตอร์ พบว่า ค่าความสามารถมีแนวโน้มลดลง เมื่อระดับความไม่เป็นอิสระของข้อสอบเพิ่มสูงขึ้น ในขณะที่การวิเคราะห์ด้วยโมเดล 2 พารามิเตอร์ค่าความสามารถมีแนวโน้มสูงขึ้นเรื่อย ๆ ซึ่งสอดคล้องกับผลการศึกษาของ Linacre (2000 อ้างถึงใน Smith, 2005) ที่กล่าวว่า เมื่อเพิ่มข้อสอบซ้ำในแบบสอบจะมีผลต่อค่าความสามารถของผู้สอบที่สูงขึ้น จากการศึกษาดังกล่าวจะเห็นว่ายังไม่ครอบคลุมสถานการณ์ทั้งหมด เพราะยังมีบางสถานการณ์ซึ่งจะเห็นจากผลการวิจัยครั้งนี้ว่าเมื่อวิเคราะห์ด้วยโมเดล 3 พารามิเตอร์ ผลการวิจัยจะแตกต่างกันไป

สำหรับค่าพารามิเตอร์ของข้อสอบที่ได้จากการประมาณค่า โดยเฉพาะค่าความยากในโมเดล 1 พารามิเตอร์ พบว่า มีค่าใกล้เคียงกับ 0 เนื่องจากการประมาณค่าพารามิเตอร์ในโมเดล 1, 2 และ 3 พารามิเตอร์ในเชิงทฤษฎีจะกำหนดค่าพารามิเตอร์ของข้อสอบ อันได้แก่

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

อำนาจจำแนก ค่าความยาก และ ค่าการเดา เป็นคะแนนมาตรฐานเพื่อให้เป็นหน่วยเดียวกันกับค่าความสามารถ (Hambleton & Swaminathan, 1985) สำหรับในโมเดล 1 พารามิเตอร์กำหนดค่าอำนาจจำแนก เป็นค่าคงที่เท่ากับ 1 ค่าการเดา เป็น 0 ซึ่งจากผลการวิจัยนี้มีผลทำให้ค่าเฉลี่ยของค่าความยากมีค่าเข้าใกล้ 0 เมื่อค่าความยากของข้อสอบอยู่ในระดับปานกลาง ดังนั้น อาจมีผลต่อค่าความสามารถของผู้สอบและสารสนเทศของแบบสอบ กล่าวคือ ถ้ามีการกำหนดให้ค่าพารามิเตอร์ความยากอยู่ในระดับยากหรือง่าย ผลของการประมาณค่าความสามารถและค่าสารสนเทศของแบบสอบอาจให้ผลที่แตกต่างกันออกไป ซึ่งจากการศึกษาของ Hambleton และ Swaminathan (1985) ได้ใช้วิธีการประมาณค่าความสามารถของผู้สอบและสารสนเทศของแบบสอบด้วยวิธีของเบย์ ซึ่งจะใช้ค่าพารามิเตอร์ความยากเป็นส่วนหนึ่งในการคำนวณ

### 1.2.2 ความสัมพันธ์ระหว่างค่าสารสนเทศของแบบสอบและค่าความเที่ยง

สำหรับค่าสารสนเทศของข้อสอบแต่ละข้อในทฤษฎีการตอบสนองข้อสอบจะส่งผลต่อค่าสารสนเทศของแบบสอบทั้งฉบับ กล่าวคือ ค่าสารสนเทศของข้อสอบในทฤษฎีการตอบสนองข้อสอบเป็นดัชนีผสมที่สร้างจากคุณลักษณะข้อสอบหลายลักษณะ ได้แก่ ค่าพารามิเตอร์ของข้อสอบและความแปรปรวนคะแนนรายข้อ และค่าสารสนเทศของข้อสอบก็จะส่งผลต่อค่าสารสนเทศของแบบสอบ ซึ่งในทฤษฎีการทดสอบแบบดั้งเดิมค่าความยากและค่าอำนาจจำแนกของข้อสอบแต่ละข้อต่างส่งผลต่อค่าความเที่ยงทั้งฉบับ (ศิริชัย กาญจนวาสี, 2545) ดังนั้น ค่าสารสนเทศของแบบสอบและค่าความเที่ยงจึงมีส่วนที่สัมพันธ์กัน สำหรับการวิเคราะห์ด้วยโมเดลทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ พบว่า ค่าสารสนเทศของแบบสอบมีแนวโน้มสูงขึ้นตามลักษณะการเพิ่มของค่าความเที่ยง แต่เมื่อวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ พบว่า ค่าสารสนเทศของแบบสอบมีแนวโน้มลดลง เนื่องจากมีค่าการเดาเข้ามาเกี่ยวข้อง โดยจะพบว่าค่าการเดาของการวิเคราะห์แบบ 3 พารามิเตอร์มีแนวโน้มสูงขึ้น

### 1.2.3 การใช้วิธีการจำลองข้อมูล

การศึกษาครั้งนี้มีข้อแตกต่างจากงานวิจัยของ Ackerman ในเรื่องวิธีการจำลองข้อมูลซึ่ง Ackerman ใช้วิธีการจำลองข้อมูลที่สัมพันธ์กันแบบลูกโซ่มาร์คอฟ (Markov Chain) ซึ่งลักษณะความสัมพันธ์แบบลูกโซ่มาร์คอฟ จะมีลักษณะความสัมพันธ์ที่เป็นระบบ กล่าวคือ ผลของคำตอบข้อหนึ่งจะมีความสัมพันธ์กับผลการตอบข้อถัดไป เช่น ผลการตอบข้อสอบในข้อ 3 จะมีความสัมพันธ์กับผลการตอบเฉพาะในข้อที่ใกล้เคียงหรือข้อถัดมา คือ ข้อ 2 หรือ ข้อ 4 เท่านั้น โดยจะไม่สนใจความสัมพันธ์ระหว่างข้อ 2 กับข้ออื่น ๆ ซึ่งในสภาพความเป็นจริง ผลการตอบข้อสอบอาจมีความสัมพันธ์กับข้อใดก็ได้ โดยในการศึกษาครั้งนี้ใช้วิธีการจำลองข้อมูลแบบการแจกแจงปกติ

หลายตัวแปรที่มีความสัมพันธ์กัน และจากการศึกษาของ Kim et. al. (2005) ด้วยการตรวจสอบความไม่เป็นอิสระของข้อสอบโดยใช้โปรแกรมการวิเคราะห์ความไม่เป็นอิสระของข้อสอบ (LDID Program) ที่ให้ค่าคะแนนแบบ 2 ค่า ด้วยการตรวจสอบข้อสอบทุกคู่ ซึ่งจะครอบคลุมสถานการณ์และสอดคล้องกับสถานการณ์ที่เป็นจริงมากกว่า

#### 1.2.4 โมเดลที่ใช้ในการวิเคราะห์กับขนาดกลุ่มตัวอย่าง

จากการศึกษาครั้งนี้จะเห็นว่าการประมาณค่าพารามิเตอร์ความยากและความสามารถของผู้สอบโดยใช้การวิเคราะห์แบบโมเดล 1 พารามิเตอร์ พบว่าไม่แตกต่างกันเมื่อระดับความไม่เป็นอิสระของข้อสอบสูงขึ้น อาจกล่าวได้ว่า การประมาณค่าพารามิเตอร์ของข้อสอบ ค่าความสามารถ และสารสนเทศของแบบสอบด้วยโมเดล 1 พารามิเตอร์ค่อนข้างมีความคงที่แม้จะได้รับผลกระทบจากความไม่เป็นอิสระของข้อสอบ ซึ่งจากการศึกษาในอดีตมีการวิจัยจำนวนมากสนับสนุนประเด็นนี้ กล่าวคือ จากการศึกษาของ Duglass (1980 อ้างถึงใน สุพล นิลกลาง, 2541) ศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าพารามิเตอร์ระหว่างโมเดลทฤษฎีการตอบสนองข้อสอบและขนาดกลุ่มตัวอย่าง พบว่า โมเดล 1 พารามิเตอร์ให้การประมาณค่าพารามิเตอร์มีประสิทธิภาพเท่ากับการประมาณค่าแบบ 2, 3 พารามิเตอร์ นอกจากนี้ จากการศึกษาของ สุพล นิลกลาง (2541) ได้ศึกษาเปรียบเทียบคุณลักษณะของแบบสอบเลือกตอบระหว่างกลุ่มตัวอย่างที่มีจำนวน 6 กลุ่ม คือ 200, 300, 450, 675, 1,015, และ 1,610 ด้วยการวิเคราะห์โมเดล 1 พารามิเตอร์ พบว่า ค่าความยาก ค่าความสามารถ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ไม่แตกต่างกัน อย่างไรก็ตามเมื่อพิจารณาเปรียบเทียบสารสนเทศของแบบสอบ ระหว่างโมเดลการวิเคราะห์ข้อสอบ จะเห็นว่าเมื่อระดับความไม่เป็นอิสระของข้อสอบเพิ่มขึ้นค่าสารสนเทศของแบบสอบจะแตกต่างกันในทุกโมเดลการวิเคราะห์และทุกขนาดกลุ่มตัวอย่าง

#### 1.2.5 การแบ่งช่วงค่า $Q_3$

การแบ่งช่วงค่า  $Q_3$  ในการวิจัยครั้งนี้ ใช้แนวคิดจากความแปรปรวน-แปรปรวนร่วม ในการกำหนดซึ่งกำหนดค่าตั้งแต่ 0.10 - 0.90 โดยแบ่งเป็น 3 ระดับ ได้แก่ ระดับต่ำ (mild: Cov. = .10, .20, .30) ระดับปานกลาง (medium: Cov. = .40, .50, .60) และระดับสูง (strong: Cov. = .70, .80, .90) ซึ่งจะเหมาะกับการแปลความหมายที่บ่งบอกระดับความเข้มของความไม่เป็นอิสระของข้อสอบได้ ซึ่งจากการศึกษาในอดีต Ackerman ได้กำหนดระดับความไม่เป็นอิสระของข้อสอบในแต่ละเงื่อนไขในระดับที่แตกต่างกันออกไป เช่น กำหนดระดับความไม่เป็นอิสระของข้อสอบในระดับต่ำ ในการจำลองข้อมูลที่มีขนาดกลุ่มตัวอย่างเท่ากับ 400 คน ใช้  $Q_3 = 0.040$ , 800 คน ใช้  $Q_3 = 0.053$  และ 1,200 คน ใช้  $Q_3 = 0.040$  ในขณะที่ Reese (1995) กำหนด  $Q_3$  ในระดับต่ำ

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

ที่  $Q_3 = 0.00-0.01$  ซึ่งจะเห็นว่ายังไม่มีความแตกต่างที่แน่นอน และจากการศึกษาครั้งนี้ยังพบว่า ค่า  $Q_3$  จะมีค่าที่แตกต่างกัน เมื่อใช้โมเดลการวิเคราะห์ที่แตกต่างกันด้วย กล่าวคือ ค่า  $Q_3$  จะลดลงเมื่อใช้การวิเคราะห์แบบ 2 และ 3 พารามิเตอร์ ในขณะที่การวิเคราะห์แบบ 1 พารามิเตอร์จะให้ค่า  $Q_3$  ที่สูงกว่า

#### 1.2.6 การใช้จำนวนรอบที่พอเพียงในการประมาณค่า

การศึกษาครั้งนี้ใช้จำนวนรอบ ในการประมาณค่ากว่า 1000 รอบ ซึ่งให้ค่าอำนาจการทดสอบสูงและตัวเลขที่ได้มีความถูกต้อง (Stone 1993 อ้างถึงใน Harwell et. al., 1996) จากการศึกษาของ Harwell (1996) ได้เสนอแนะให้มีการใช้จำนวนรอบในการประมาณค่า มาใช้ศึกษา ด้วยวิธีการจำลองข้อมูลในทฤษฎี IRT เพื่อให้การประมาณค่าสถิติต่าง ๆ มีความคงที่และน่าเชื่อถือ

#### 1.2.7 ลักษณะของแบบสอบที่ใช้ในการตรวจสอบ

จากการศึกษาครั้งนี้ผู้วิจัยมุ่งตรวจสอบแบบสอบทั้งฉบับโดยมิได้คำนึงถึงลักษณะข้อสอบ จุดมุ่งหมายของแบบสอบแต่ละประเภท แต่เป็นการศึกษาเพื่อตรวจสอบความไม่เป็นอิสระของข้อสอบทั่วไป เช่น แบบวัดผลสัมฤทธิ์ทางการเรียน แบบสอบปลายภาคเรียน เป็นต้น ซึ่งจะต้องออกข้อสอบให้ครอบคลุมเนื้อหาที่กำหนดไว้ในหลักสูตร เหมาะสำหรับการตรวจสอบความไม่เป็นอิสระทั้งฉบับ แต่ไม่เหมาะสมที่จะใช้ตรวจสอบแบบสอบที่มีลักษณะเป็นชุดข้อสอบ (testlet) ที่มีคำถามตามในเนื้อหาเดียวกันหลาย ๆ ข้อ โดยผู้วิจัยอาจต้องมีการปรับคะแนน หรือ สุ่มข้อสอบที่ใช้คำถามในเนื้อหาพร้อมกันเพื่อจะใช้ในการตรวจสอบความสอดคล้องของการจำลองข้อมูลกับสถานการณ์จริง

จากการศึกษาในครั้งนี้ผู้วิจัยพยายามศึกษาด้วยวิธีการจำลองข้อมูลและศึกษาจากข้อมูลจริงไปพร้อมกัน จากผลการวิเคราะห์ค่าความเที่ยงของแบบสอบในข้อมูลจริง พบว่าเมื่อนำข้อสอบที่เป็นอิสระและไม่เป็นอิสระของข้อสอบออกจากแบบสอบ จะพบว่า ค่าความเที่ยงเปลี่ยนแปลงไป กล่าวคือ เมื่อวิเคราะห์ค่าความเที่ยงแบบสอบที่นำข้อสอบที่มีความไม่เป็นอิสระออก 5 ข้อ พบว่า ค่าความเที่ยงต่ำกว่าแบบสอบที่นำข้อสอบที่มีความเป็นอิสระกันออก 5 ข้อ ซึ่งแสดงให้เห็นว่า ค่าความเที่ยงที่สูงขึ้นนั้นมาจากความไม่เป็นอิสระของข้อสอบนั่นเอง

สำหรับการตรวจสอบในสถานการณ์จากการจำลองข้อมูลและสถานการณ์ที่เป็นจริงมีข้อที่น่าพิจารณาคือ แบบสอบที่สร้างอย่างมีมาตรฐาน โดยมีการวางแผนในการออกข้อสอบ มีโครงสร้างของข้อสอบมักจะไม่เกิดปัญหาเกี่ยวกับความไม่เป็นอิสระของข้อสอบ ดังจะเห็นจากผลการวิเคราะห์ข้อสอบจริงครั้งนี้ พบว่า ค่าดัชนีสัมบูรณ์  $Q_3$  ในวิชาคณิตศาสตร์ ที่วิเคราะห์ด้วย

โมเดล 1PL 2PL และ 3PL มีค่าสูงสุด เท่ากับ 0.145, 0.129 และ 0.137 เมื่อเปรียบเทียบกับค่าที่คาดหวังของ  $Q_3$  จะมีค่าเท่ากับ 0.025 แต่อย่างไรก็ตาม นักวัดผลควรระมัดระวังการสร้างข้อสอบที่มีการใช้ตารางเนื้อหาพร้อมกัน โดยเฉพาะในการถามคำถามหลายข้อเกี่ยวกับข้อมูลในตาราง กราฟ หรือในวิชาภาษาอังกฤษ ในเนื้อหาเดียวกัน ซึ่งเงื่อนไขของการทดสอบนี้อาจไม่สอดคล้องกับการออกข้อสอบลักษณะดังกล่าว แต่จะเหมาะสำหรับการตรวจสอบ ระดับความไม่เป็นอิสระของข้อสอบกับแบบสอบทั่วไป

## 2. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1) การเลือกใช้โมเดลทฤษฎีการตอบสนองข้อสอบในการวิเคราะห์ข้อสอบ การสร้างแบบสอบให้เป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบเป็นสิ่งสำคัญ หากแบบสอบมีคุณภาพจะส่งผลทำให้การประมาณค่าพารามิเตอร์ข้อสอบและค่าความสามารถเป็นค่าที่ถูกต้องแท้จริง นักวัดผลอาจเลือกใช้โมเดลการวิเคราะห์ที่เหมาะสมกับจุดมุ่งหมายของการทดสอบ เช่น แบบสอบทั่วไป ที่มีลักษณะข้อสอบที่โอกาสเกิดความไม่เป็นอิสระของข้อสอบน้อย นักวัดผลอาจเลือกใช้โมเดลการวิเคราะห์ 2 หรือ 3 พารามิเตอร์ได้ แต่ในสถานการณ์ของการวิเคราะห์แบบสอบที่มีโอกาสเสี่ยงต่อการเกิดความไม่เป็นอิสระของข้อสอบสูง เช่น แบบสอบที่ถามคำถามในเนื้อหาเดียวกัน การใช้แผนภูมิ กราฟ เพื่อถามคำถามหลายข้ออาจเลือกใช้โมเดลการวิเคราะห์แบบ 1 พารามิเตอร์ได้ ซึ่งจะเหมาะสมกว่า เกิดความคลาดเคลื่อนในการประมาณค่าน้อยกว่า แต่อย่างไรก็ตามการวิเคราะห์ด้วยโมเดลการวิเคราะห์ 1 พารามิเตอร์ จะไม่นำค่าอำนาจจำแนก และค่าการเดาเข้ามาใช้ในการคำนวณค่าความสามารถจึงทำให้ค่าความสามารถไม่ตรงกับความเป็นจริง

2) จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่าลักษณะแบบสอบที่จะเกิดความไม่เป็นอิสระมักจะเป็นแบบสอบที่มีลักษณะเป็นชุดข้อสอบย่อย (testlet) ซึ่งเป็นแบบสอบที่ใช้คำถามหลายข้อถามในเนื้อหาเดียวกันทั้งฉบับ (Ferrara et. al., 1997; Boyd et. al., 2002) ดังนั้นอาจมีการศึกษาความไม่เป็นอิสระของข้อสอบที่เป็นข้อสอบแบบชุดย่อยโดยวิเคราะห์ในลักษณะความไม่เป็นอิสระระหว่างชุดข้อสอบย่อย โดยเป็นการศึกษาความเป็นอิสระของชุดข้อสอบแต่ละชุด

3) ควรมีการศึกษาความไม่เป็นอิสระของข้อสอบในการทดสอบแบบปรับเหมาะเพราะซึ่งจากการศึกษาของ Zenisky et. al. (2003) และ Smith (2005) พบว่า ผลของระดับความไม่เป็นอิสระของข้อสอบจะส่งผลต่อการประมาณค่าความสามารถของผู้สอบที่คลาดเคลื่อน กล่าวคือเป็นการเพิ่มความคลาดเคลื่อนประเภทที่ 1 สูง เมื่อนำมาใช้ในการทดสอบแบบปรับเหมาะโดยการคำนวณค่าความสามารถของผู้สอบจะยุติการสอบก่อนกำหนด

4) ควรมีการศึกษาเปรียบเทียบจำนวนข้อสอบที่มีความไม่เป็นอิสระในแต่ละแบบสอบว่า จะมีผลต่อค่าดัชนีวัดระดับความไม่เป็นอิสระของข้อสอบมากน้อยต่างกันหรือไม่ โดยกำหนดสัดส่วน

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

ของข้อสอบที่ไม่เป็นอิสระกับจำนวนข้อสอบทั้งฉบับให้แตกต่างกัน เช่น 10% 20% 30% และ 40% เป็นต้น ซึ่งจากการทดลองวิเคราะห์ของผู้วิจัย ได้กำหนดข้อสอบให้เกิดความไม่เป็นอิสระทุกข้อ พบว่า ค่าอำนาจจำแนก และค่าความเที่ยงจะสูงขึ้นอย่างมาก

## รายการอ้างอิง

### ภาษาไทย

- กิ่งกาญจน์ เมฆา. (2540). การศึกษาความสัมพันธ์ระหว่างค่าพารามิเตอร์ของแบบสอบมิติสัมพันธ์แบบซ่อนภาพ ที่วิเคราะห์จากกลุ่มตัวอย่างที่มีขนาดต่างกับขนาดกลุ่มตัวอย่างที่เป็นมาตรฐาน ด้วยวิธีราล์ชโมเดล. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- เกศมณี พัยค์. (2543). การเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ของแบบสอบความเข้าใจในการอ่านที่มีจำนวนข้อและจำนวนกลุ่มตัวอย่างต่างกันที่วิเคราะห์ด้วยวิธีราล์ชโมเดล. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- วิชุดา บัวคง. (2533). การเปรียบเทียบประสิทธิผลของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซิมัมไลค์ลิสต์ วิธีอีวีรอสติก และวิธีของเบย์ในแบบสอบวัดผลสัมฤทธิ์ และแบบสอบความถนัด. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาวิจัย การศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- วินัย วงศ์ฤทัยวัฒนา. (2533). การเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของโมเดลโลจิสติกแบบสองพารามิเตอร์ระหว่างวิธีของเบส์กับวิธีแมกซิมัมไลค์ลิสต์. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2545). ทฤษฎีการทดสอบแนวใหม่ (MODERN TEST THEORYS). พิมพ์ครั้งที่ 2. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2548). ทฤษฎีการทดสอบแบบดั้งเดิม (CLASSICAL TEST THEORY). พิมพ์ครั้งที่ 5. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สุพล นิลกลาง. (2541). การเปรียบเทียบคุณลักษณะของแบบสอบเลือกตอบระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกันที่วิเคราะห์ด้วยโมเดลราล์ช. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- สุวิมล ตีรกันันท์. (2537). การวิเคราะห์ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบตามโมเดลโครงสร้าง ความแปรปรวนร่วม. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

## ภาษาอังกฤษ

- Ackerman, T. A. (1987). *The robustness of logist and bilog IRT estimation programs to violations of local independence*. Retrieved October 8, 2010, from [http://www.act.org/research/reports/pdf/ACT\\_RR87-14.pdf](http://www.act.org/research/reports/pdf/ACT_RR87-14.pdf)>
- Adams, R. J., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement, 21*(1), 1-12.
- Boyd, A. M., et. al. (2003). A comparison of exposure control procedures in CAT systems based on different measurement models for testlets.
- Chen, W., & Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics, 22*, 265-289.
- Du Toit, M. (2003). *IRT from SSI: BILOG-MG MULTILOG PARSCALE TESTFACT*. Lincolnwood: Scientific Software International, Inc.
- Ferrara, S., Huynh, H., & Baghi, H. (1997). Contextual characteristics of locally dependent open - ended item cluster in a large - scale performance assessment. *Applied Measurement in Education, 10*(2), 123-144.
- Ferrara, S., Huynh, H., & Micheals, H. (1999). Contextual explanations of local dependence in item clusters in a large scale hand-on science performance assessment. *Journal of Educational Measurement, 36*(2), 199-140.
- Hambleton, R. K., & Swaminathan, H., (1985). *Item response theory: principles and applications*. (2nd Ed.). Boston: Kluwer Nijhoff Publishing.
- Harwell, M., et. al. (1996). Monte Carlo studies in item response theory. *Applied Psychological Measurement, 20*(2), 101-125.
- Hulin, C. L., et. al. (1983). *Item response theory: Application to psychological measurement*. Homewood, Illinois: Dow Jones-Irwin.
- Ip, E. H. (2001). Testing for local dependency in dichotomous and polytomous item response models. *Psychometrika, 66*(1), 109-132.
- Jiao, H., & Kamata, A. (2003). *Model comparisons in the presence of local item dependence*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicaco, IL. April 21-25.

◆ ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบ ◆  
ค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบ เมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน

- Jiao, H., & Kamata, A. (2005). Modeling local item dependence with the hierarchical generalized linear model. *Journal of Applied Measurement*, 6(3), 311-321.
- Keller, L. A., Swaminayhan, H., & Sireci, S. (2003). Evaluation scoring produces for content dependent item sets. *Applied Measurement in Education*, 16(3), 207-222.
- Kim, D., et. al. (2007). Assessing relative performance of local item dependence (LID) Indexes. Retrieve October 8, 2010, from <http://measuredprogress.com/resources/measurement/framework/materials/07/AERA.NCME/AssessingRelativePerformance.pdf>
- Kim, S. H., Cohen, A. S., & Lin, Y. H. (2005). LDID: A Computer program for local dependence indices for dichotomous items1. Version 1.0
- Lee, Y. W. (2004). *Examining passage-related local item dependence (LID) and measurement construct using  $Q_3$  statistics in an EFL reading comprehension test*. Center for Validity Research/Division of Research and Development, Educational Testing service, Rosedale Road, Princeton. 74-100.
- Reese, L. M. (1995). *The impact of local dependencies on some LSAT outcomes*. LSAC Research Report Series. Law School Admission Council, Newtown, PA: Law School Admission Council Inc.
- Reese, L. M. (1999). *A classical test theory perspective on LSAT local item dependence*. LSAC Research Report Series. Law School Admission Council, Newtown, PA: Law School Admission Council Inc.
- Smith, Jr. V. E. (2005). Effect of item redundancy on rash item and person estimates. *Journal of Applied Measurement*, 6(2), 147-163.
- Spray, J. A., & Ackerman, T. (1987). *The effect of item response dependency on trait or ability dimensionality*. ACT Research Report Series 87-10. Iowa City: American College Testing Program. Retrieved October 8, 2010 from [http://www.act.org/research/reports/pdf/ACT\\_RR87-10.pdf](http://www.act.org/research/reports/pdf/ACT_RR87-10.pdf)
- Yen, W. M. (1984). Effect of local item dependence on the fit and equation performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125-145.



Zenisky, A. L., Hambleton, R. K., & Sireci, S. G. (2003). *Effect of local item dependence on the validity of IRT item, test, and ability statistics*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.

