

การเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำ  
ข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์  
เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน



นางสาวพิมพ์สิริ เกียรติยศราษฎร์

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาครุศาสตรมหาบัณฑิต


สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2549

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A COMPARISON OF EXAMINEES' ABILITIES, ANSWER-CHANGING CHARACTERISTICS  
AND TIME SPENT IN COMPUTERIZED ADAPTIVE TESTING  
UNDER VARYING TESTING CONDITIONS AND EXAMINEES' ABILITY LEVELS



Miss Pimsiri Tienoraset

สถาบันวิทยบริการ

A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Education Program in Educational Measurement and Evaluation  
Department of Educational Research and Psychology

Faculty of Education

Chulalongkorn University

Academic Year 2006

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน

โดย

นางสาว พิมพ์สิริ เขียรนรเศรษฐ์

สาขาวิชา

การวัดและประเมินผลการศึกษา

อาจารย์ที่ปรึกษา

ผู้ช่วยศาสตราจารย์ ดร.เอมอร จังศิริพรปกรณ์

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นักศึกษานิพนธ์ฉบับนี้เป็นส่วน  
หนึ่งของการศึกษาตามหลักสูตรปริญญาโทบัณฑิต

..... คณบดีคณะครุศาสตร์  
(รองศาสตราจารย์ ดร. พงษ์ ศิริบรรณพิทักษ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ  
(ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี)

..... อาจารย์ที่ปรึกษา  
(ผู้ช่วยศาสตราจารย์ ดร. เอมอร จังศิริพรปกรณ์)

..... กรรมการ  
(อาจารย์ ดร. กมลวรรณ ตังธนกันนธ์)

พิมพ์สิริ เรียบนรเศรษฐ์ : การเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน. (A COMPARISON OF EXAMINEES' ABILITIES, ANSWER-CHANGING CHARACTERISTICS AND TIME SPENT IN COMPUTERIZED ADAPTIVE TESTING UNDER VARYING TESTING CONDITIONS AND EXAMINEES' ABILITY LEVELS) อ. ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.เอมอร จังศิริพรปกรณ์, 203 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาโปรแกรมคอมพิวเตอร์วิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ที่สามารถกำหนดเงื่อนไขการทดสอบที่ให้ทวนคำตอบได้โดยวิธีการทวนที่มีการจำกัดช่วงกับการทดสอบที่ไม่ให้ทวนคำตอบ รวมทั้งเปรียบเทียบความแตกต่างของค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน กลุ่มตัวอย่างคือนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2549 โรงเรียนวัดพลับพลายชัย จำนวน 200 คน การเก็บรวบรวมข้อมูลเป็นการทดสอบแบบซีเอทีด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ การวิเคราะห์ข้อมูลใช้การวิเคราะห์ความแปรปรวนทางเดียว และทดสอบภายหลังด้วยวิธีของเชฟเฟนหรือดัชนีแทน ที่ 3 ผลการวิจัยสรุปว่า

1. ผู้สอบที่มีระดับความสามารถแตกต่างกันมีความแตกต่างของค่าความสามารถของผู้สอบ ภายใต้เงื่อนไขการทดสอบที่ไม่ให้ทวนคำตอบกับการทดสอบที่ให้ทวนคำตอบ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่ผู้สอบที่มีระดับความสามารถต่ำจะมีค่าความสามารถในเงื่อนไขการให้ทวนคำตอบมากกว่าเงื่อนไขการไม่ให้ทวนคำตอบมากที่สุด รองมาได้แก่ผู้สอบที่มีระดับความสามารถปานกลาง และสูงตามลำดับ

2. ลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วยังคงคำตอบเดิมไว้ ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ ระหว่างผู้สอบที่มีระดับความสามารถแตกต่างกัน ไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

3. ลักษณะการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบ จากผิดเป็นถูก จากผิดเป็นผิด และจากถูกเป็นผิด ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ ระหว่างผู้สอบที่มีระดับความสามารถแตกต่างกัน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 โดยผู้สอบที่มีระดับความสามารถสูงและปานกลางมีจำนวนข้อที่เปลี่ยนจากผิดเป็นถูกมากกว่าผู้สอบที่มีระดับความสามารถต่ำ ส่วนผู้สอบที่มีระดับความสามารถต่ำมีจำนวนข้อที่เปลี่ยนจากผิดเป็นผิด และจากถูกเป็นผิดมากกว่าผู้สอบที่มีระดับความสามารถสูงและปานกลาง

4. ผู้สอบที่มีระดับความสามารถแตกต่างกันมีความแตกต่างของเวลาที่ใช้ในการทำข้อสอบ ภายใต้เงื่อนไขการทดสอบที่ไม่ให้ทวนคำตอบกับการทดสอบที่ให้ทวนคำตอบ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ภาควิชา ศึกษาศาสตร์และจิตวิทยาการศึกษา

ลายมือชื่อผู้วิจัย..... พิมพ์สิริ เรียบนรเศรษฐ์

สาขาวิชา การวัดและประเมินผลการศึกษา

ลายมือชื่ออาจารย์ที่ปรึกษา..... (อ.เอมอร)

ปีการศึกษา 2549

## 4883729427 : MAJOR Educational Measurement and Evaluation

KEY WORD: COMPUTERIZED ADAPTIVE TESTING / EXAMINEES' ABILITY / ANSWER-CHANGING  
 PIMSIRI TIENORASET : A COMPARISON OF EXAMINEES' ABILITIES, ANSWER-  
 CHANGING CHARACTERISTICS AND TIME SPENT IN COMPUTERIZED ADAPTIVE  
 TESTING UNDER VARYING TESTING CONDITIONS AND EXAMINEES' ABILITY LEVELS.  
 THESIS ADVISOR : ASST. PROF. AIMORN JANGSIRIPORNPAKORN, 203 pp.

The purpose of this research were to develop the computer program of adaptive test of the comprehension of fraction of Prathomsuksa 6 students. The testing contained two conditions; i.e., review allowed within successive 5-item blocks and no item review allowed and to compare the difference of examinees' ability levels, answer changing characteristics and the time spent in computerized adaptive testing under the various levels of examinees' abilities. The samples consisted of 200 Prathomsuksa 6 students in 2006 academic year from Wat Plubplachai School. The computerized adaptive testing program was used to collect the data. The one-way analysis of variance and the Scheffe' or Dunnett' s T3 method for post hoc analysis were employed to analyzed the collected data.

The findings were as follows:

1. There was a significance at .05 level on examinees' ability levels among various ability levels of examinees. After changing the test condition from non-review test to review test, the low-ability examinees received the most different ability more than the medium-ability examinees and the high-ability examinees by order.
2. There was no significance at .05 level on retaining the same answer in answer-changing characteristic among various ability levels of examinees under the review test condition.
3. There was a significance at .05 level on changing the answer: wrong to right, wrong to wrong, right to wrong, in answer-changing characteristic among various ability levels of examinees under the review test condition. The medium-ability examinees and the high-ability examinees had the number of changing answer from wrong to right more than the low-ability examinees. The low-ability examinees had the number of changing answer from wrong to wrong and from right to wrong more than the high-ability examinees and the medium-ability examinees.
4. There was no significance at .05 level on the different time spent among various ability levels of examinees under the review test condition.

Department Educational Research and Psychology

Student's signature.....*Pimsiri Tienoraset*

Field of study Educational Measurement and Evaluation.

Advisor's signature.....*Aimj*

Academic year 2006

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความสามารถ ความช่วยเหลือ และความร่วมมือจากหลายๆ บุคคลสำคัญ โดยเฉพาะอย่างยิ่ง ผศ.ดร.เอมอร จังศิริพรปกรณ์ ผู้เป็นอาจารย์ที่ปรึกษา คอยให้คำแนะนำ สละเวลา ตรวจสอบวิทยานิพนธ์ในทุกขั้นตอนการทำงานจนเสร็จสมบูรณ์ อีกทั้งยังเป็นกำลังใจ และให้ความเมตตาเสมอมา รวมถึงคณะกรรมการสอบวิทยานิพนธ์ทุกท่านที่ให้ความกรุณา แนะนำแนวทางข้อแก้ไขต่างๆ จนทุกอย่างผ่านพ้นอย่างราบรื่นและสวยงาม ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

กราบขอบพระคุณ ดร. รังสรรค์ มณีเล็ก เป็นอย่างยิ่งที่ท่านให้ความเป็นกันเอง สละเวลาอันมีค่าให้เข้าพบทุกครั้ง ให้โปรแกรมการทดสอบแบบ CAT ที่ท่านทำไว้เพื่อให้ผู้วิจัยได้พัฒนาต่อยอด รวมทั้งโทรศัพท์ถามความคืบหน้าและให้ความเอ็นดูแก่ผู้วิจัยตลอดมา ผู้วิจัยซาบซึ้งใจในความสามารถของท่านที่สุด และจะนำความมีอัธยาศัยไมตรีของท่านมาใช้เป็นแบบอย่างต่อไป รวมถึงขอขอบพระคุณ ดร.สุพัฒน์ สุขมลสันต์ ที่เป็นผู้จุดประกาย ให้ความรู้ความเข้าใจในการทำงานของ CAT เป็นอย่างดี

ขอบพระคุณเป็นอย่างมากสำหรับบุคลากรทุกท่านที่เกี่ยวข้องในการเก็บรวบรวมข้อมูลในทุกๆ กระบวนการ ทั้งที่โรงเรียนราชวินิตและโรงเรียนวัดพลับพลาชัย ที่ให้ความเอื้อเฟื้อและความอนุเคราะห์แก่ผู้วิจัยในการเก็บรวบรวมข้อมูล

ขอบคุณความดีและความมีน้ำใจของ พี่เจ ชูพงศ์ อุทยานวุฒิมูล ที่คอยให้ความช่วยเหลือในทุกๆ เรื่อง รวมทั้งการรับหน้าที่เป็นโปรแกรมเมอร์คนเก่ง จนวิทยานิพนธ์ฉบับนี้เสร็จสมบูรณ์ ความสำเร็จครั้งนี้ขอแบ่งให้พี่เจครึ่งหนึ่ง และอีกกำลังคนสำคัญ ขอขอบคุณ พี่นุ้ม สุกัญญา ถาวร ที่คอยให้กำลังใจ และมีน้ำใจให้แก่ผู้วิจัยเสมอมา

ขอบคุณพี่จ๊อบที่ทำหน้าที่บุคคลผู้รู้ สงสัยอะไรโทรหาพี่จ๊อบได้เสมอ พี่ภูมิ (ป.เอก) ที่คอยเป็นที่ปรึกษาการวิเคราะห์ค่าสถิติต่างๆ พี่กีฬที่สุดสวยที่มีเวลาดีๆ ฆ่าๆ ด้วยกันตลอด พี่นกใหญ่(แต่ตัวเล็ก) ของน้องนกเล็ก(แต่ตัวใหญ่) ที่ให้ความกระจ่างในทุกๆ คำถาม รวมทั้ง ฮิม เป็ด พี่นิง พี่ดาว ออ ป้านุ้ย เพื่อนร่วมเอกและร่วมทุกข์ร่วมสุขด้วยกันมาตั้ง 2 ปี ขอขอบคุณจริงๆ

ท้ายที่สุดนี้ที่ลืมไม่ได้ คือขอบคุณมะม้า ปะป๊า และน้องเน ความสำเร็จครั้งนี้ก็อีกหนึ่งขอมอบให้ครอบครัวที่แสนอบอุ่น ครอบครัว “เหียนรณเศรษฐ์” ที่คอยเติมเต็มกำลังใจเมื่อยามรู้สึกท้อแท้ และทำให้ผู้วิจัยมีทุกๆ วันนี้ได้

# สารบัญ

๗

หน้า

บทคัดย่อภาษาไทย .....	ง
บทคัดย่อภาษาอังกฤษ .....	จ
กิตติกรรมประกาศ .....	ฉ
สารบัญ.....	ช
สารบัญตาราง .....	ญ
สารบัญแผนภาพ .....	ฎ
บทที่	
1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
คำถามวิจัย .....	6
วัตถุประสงค์ของการวิจัย .....	7
สมมติฐานของงานวิจัย .....	7
ขอบเขตของการวิจัย .....	9
ข้อจำกัดของการวิจัย.....	9
คำจำกัดความที่ใช้ในการวิจัย .....	9
ประโยชน์ที่คาดว่าจะได้รับ .....	13
2 เอกสารและงานวิจัยที่เกี่ยวข้อง .....	14
ตอนที่ 1 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ.....	14
ประวัติการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ .....	15
ความหมายที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ.....	15
องค์ประกอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ .....	16
การดำเนินการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ.....	18
ประเภทของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ.....	19
การให้คะแนนการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ.....	29
ประโยชน์การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ .....	33
ประเด็นปัญหาของการใช้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ .....	33
ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบ .....	36
ความเป็นมาของทฤษฎีการตอบสนองข้อสอบ .....	36

บทที่	หน้า
หลักการของทฤษฎีการตอบสนองต่อข้อสอบ .....	37
ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองต่อข้อสอบ .....	38
แบบของการวิเคราะห์ข้อสอบ .....	40
ปัจจัยที่มีผลต่อคุณภาพของแบบทดสอบ .....	43
การนำทฤษฎีการตอบสนองข้อสอบไปใช้ .....	46
ตอนที่ 3 ระเบียบวิธีวิจัยและการพัฒนาโปรแกรมคอมพิวเตอร์ทางการศึกษา .....	48
กำหนดปัญหา .....	49
วิเคราะห์ปัญหา .....	49
การออกแบบโปรแกรมคอมพิวเตอร์ .....	50
เลือกภาษาคอมพิวเตอร์และเขียนโปรแกรม .....	51
ทดสอบโปรแกรม .....	52
จัดทำเอกสารประกอบโปรแกรม .....	52
ประเมินคุณภาพของโปรแกรม .....	53
ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง .....	54
งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ .....	55
งานวิจัยที่เกี่ยวข้องกับการเปิดโอกาสให้มีการทวนคำตอบในการทดสอบแบบ ปรับเหมาะกับความสามารถของผู้สอบ .....	65
งานวิจัยที่เกี่ยวข้องกับการให้เปลี่ยนคำตอบในการทำการทดสอบแบบดั้งเดิม .....	70
3 วิธีดำเนินการวิจัย .....	72
ประชากรและกลุ่มตัวอย่าง .....	72
เครื่องมือที่ใช้ในการวิจัย .....	74
การเก็บรวบรวมข้อมูล .....	84
การวิเคราะห์ข้อมูล .....	86
4 ผลการพัฒนาโปรแกรมและผลการวิเคราะห์ข้อมูล .....	98
ตอนที่ 1 ผลการพัฒนาโปรแกรมคอมพิวเตอร์ .....	98
การเตรียมความพร้อมของคอมพิวเตอร์ก่อนการติดตั้งโปรแกรม .....	98
ขั้นตอนการติดตั้งโปรแกรม .....	99
ผลการพัฒนาโปรแกรม .....	99
ตอนที่ 2 ผลการเปรียบเทียบค่าความแตกต่างของค่าความสามารถของผู้สอบ .....	113



บทที่	หน้า
ตอนที่ 3 ผลการเปรียบเทียบลักษณะของการเปลี่ยนคำตอบ .....	117
ตอนที่ 4 ผลการเปรียบเทียบค่าความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ .....	124
5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ .....	129
สรุปผลการวิจัย.....	132
อภิปรายผลการวิจัย .....	137
ข้อเสนอแนะ .....	142
รายการอ้างอิง.....	146
ภาคผนวก.....	152
ภาคผนวก ก ตัวอย่าง: แบบสอบปรนัยวิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ชั้นประถมศึกษาปีที่ 6 จำนวน 40 ข้อ .....	153
ภาคผนวก ข แบบฟอร์มการตรวจสอบความตรงตามเนื้อเรื่อง .....	155
ภาคผนวก ค หนังสือ: ขอเชิญเป็นผู้ทรงคุณวุฒิตรวจสอบเครื่องมือวิจัย .....	158
ภาคผนวก ง หนังสือ: ขอความร่วมมือในการเก็บข้อมูล โรงเรียนราชวินิต .....	160
ภาคผนวก จ หนังสือ: ขอความร่วมมือในการเก็บข้อมูล โรงเรียนวัดพลับพลาชัย .....	162
ภาคผนวก ฉ คู่มือสำหรับผู้สอบ .....	164
ภาคผนวก ช การเขียนโปรแกรม.....	168
ภาคผนวก ซ Output การวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ.....	180
ภาคผนวก ฌ ตัวอย่าง: ข้อมูลไฟล์ผลลัพธ์ของผู้สอบ .....	194
ประวัติผู้เขียนวิทยานิพนธ์ .....	204

## สารบัญตาราง

ญ

ตารางที่		หน้า
1	จำนวนโรงเรียนที่ได้จากการสุ่มชั้นตอนแรก จำแนกเป็นรายโรงเรียน .....	72
2	พาเซทสำหรับสร้างข้อสอบ เรื่องความเข้าใจเศษส่วน ระดับชั้นประถมศึกษาปีที่ 6 .....	74
3	การปรับเทียบค่าพารามิเตอร์ความยาก อำนาจจำแนก และโอกาสในการเดา ข้อสอบ ได้ถูกของข้อสอบทั้ง 40 ข้อเพื่อให้อยู่บนสเกลเดียวกับข้อสอบในคลังเดิม .....	88
4	ระยะเวลาที่นักเรียนกลุ่ม 2 ใช้ในการทำข้อสอบในการทดสอบแบบปรับเหมาะกับความ ความสามารถของผู้สอบด้วยคอมพิวเตอร์ เพื่อกำหนดระยะเวลาสูงสุดในการทำ ข้อสอบให้กับโปรแกรมคอมพิวเตอร์ .....	91
5	การวิเคราะห์ความแปรปรวนทางเดียว (1-WAY ANOVA) .....	94
6	การวิเคราะห์ความแปรปรวนทางเดียว (1-WAY ANOVA) .....	95
7	การวิเคราะห์ความแปรปรวนทางเดียว (1-WAY ANOVA) .....	97
8	จำนวนผู้สอบ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ของค่าความสามารถของผู้สอบ ในเรื่องนี้ในการไม่ให้ทวนคำตอบและการให้ทวนคำตอบ ของผู้สอบในแต่ละระดับ ความสามารถ .....	113
9	จำนวนผู้สอบ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด และค่าต่ำสุด ของค่าความ แตกต่างของค่าความสามารถของผู้สอบ จำแนกตามระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ) .....	114
10	ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนด้วย Levene's test และผลการ วิเคราะห์ความแปรปรวนทางเดียว ของค่าความแตกต่างของค่าความสามารถของ สอบ ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน .....	116
11	ผลของการเปรียบเทียบความแตกต่างเป็นรายคู่ของค่าความแตกต่างของค่าความ สามารถของผู้สอบระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน ด้วยเทคนิค Scheffe .....	116
12	จำนวนผู้สอบที่ให้ผลในกรณีต่างๆ ภายใต้การทดสอบที่ให้ทวนคำตอบของผู้สอบ ในแต่ละระดับความสามารถ .....	117
13	จำนวนผู้สอบ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด และผลการ ทดสอบความเป็นเอกพันธ์ความแปรปรวนด้วย Levene's test และผลการวิเคราะห์ ความแปรปรวนทางเดียว ของลักษณะของการเปลี่ยนคำตอบของผู้สอบ ระหว่าง ผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน .....	119

ตารางที่	หน้า
14 ผลของการเปรียบเทียบความแตกต่างเป็นรายคู่ของลักษณะของการเปลี่ยนคำตอบของผู้สอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกันด้วยเทคนิค Dunnett's T3 .....	120
15 จำนวนผู้สอบ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบของผู้สอบ ในแต่ละระดับความสามารถ .....	125
16 จำนวนผู้สอบ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด และค่าต่ำสุด ของค่าความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ จำแนกตามระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ) .....	126
17 ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนด้วย Levene's test และผลการวิเคราะห์ความแปรปรวนทางเดียว ของค่าความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน .....	128

## สารบัญแผนภาพ

ฎ

แผนภาพที่		หน้า
1	โครงสร้างของการทดสอบแบบปรับเหมาะโดยใช้กลวิธีสองขั้นตอน.....	21
2	โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดชนิดขนาดขั้นคงที่.....	22
3	โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดชนิดขนาดขั้นแปรผัน.....	23
4	โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดชนิดรอบบิน—มอนโร.....	23
5	โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดข้างตัดวิธีสกัดการสะท้อนกลับ ...	24
6	โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดข้างตัด วิธีรักษาการสะท้อนกลับ..	24
7	โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดชนิดที่หลายข้อในแต่ละชั้น.....	25
8	โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดชนิดให้นำน้ำหนักตัวเลือกเพื่อ แยกทาง .....	26
9	โครงสร้างของการทดสอบแบบปรับเหมาะแบบระดับยืดหยุ่น .....	27
10	โครงสร้างของการทดสอบแบบปรับเหมาะแบบปรับระดับขั้น.....	27
11	โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อทดสอบแบบ 1 พารามิเตอร์.....	41
12	โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อทดสอบแบบ 2 พารามิเตอร์.....	42
13	โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อทดสอบแบบ 3 พารามิเตอร์.....	42
14	โค้งสารสนเทศของแบบทดสอบสำหรับวัดความสามารถที่ระดับ 1.0.....	45
15	กรอบแนวคิดในการวิจัย .....	71
16	ผังงานแสดงขั้นตอนการทำงานของคอมพิวเตอร์ตามเงื่อนไขการทดสอบที่ 1.....	81
17	ผังงานแสดงขั้นตอนการทำงานของคอมพิวเตอร์ตามเงื่อนไขการทดสอบที่ 2.....	82
18	แสดงข้อมูลวิทยานิพนธ์.....	99
19	แสดงหน้าจอสําหรับกรอกข้อมูลผู้สอบ และเลือกเงื่อนไขการทดสอบ.....	100
20	ผังงานแสดงขั้นตอนการทำงาน ในส่วนหน้าจอบันทึกข้อมูลผู้ทำการสอบ.....	101
21	แสดงหน้าจอสําหรับการทดสอบภายใต้เงื่อนไขการไม่ให้อวนคำตอบ.....	102
22	แสดงผลลัพธ์ของการทดสอบภายใต้เงื่อนไขการไม่ให้อวนคำตอบ.....	104
23	ผังงานแสดงขั้นตอนการทำงานของการทดสอบภายใต้เงื่อนไขการไม่ให้อวนคำตอบ.....	105
24	หน้าจอสําหรับการทดสอบภายใต้เงื่อนไขการให้อวนคำตอบ.....	106
25	ผลลัพธ์ของการทดสอบภายใต้เงื่อนไขการให้อวนคำตอบ.....	107
26	ผังงานแสดงขั้นตอนการทำงานของการทดสอบภายใต้เงื่อนไขการให้อวนคำตอบ.....	108
27	แสดงหน้าจอสําหรับการกลับไปทวนคำตอบ.....	109

แผนภาพที่	หน้า
28	ผังงานแสดงขั้นตอนการทำงานของกรทวนคำตอบ ..... 110
29	แสดงหน้าจอสำหรับการเปลี่ยนคำตอบ ..... 111
30	ผังงานแสดงขั้นตอนการทำงานของกรเปลี่ยนคำตอบ..... 111
31	รูปตัวอย่างโจทย์..... 112
32	ไฟล์ข้อมูลคลังข้อสอบ ..... 112
33	กราฟแสดงแนวโน้มของค่าความแตกต่างของค่าความสามารถของผู้สอบ (θdif) ..... 115
34	กราฟแสดงแนวโน้มของลักษณะของการเปลี่ยนคำตอบประเภทกรทวนคำตอบ แล้วยังคงคำตอบเดิมไว้ของผู้สอบที่มีระดับความสามารถแตกต่างกัน..... 122
35	กราฟแสดงแนวโน้มของลักษณะของการเปลี่ยนคำตอบประเภทกรทวนคำตอบแล้ว มีการเปลี่ยนคำตอบจากผิดเป็นถูก ของผู้สอบที่มีระดับความสามารถแตกต่างกัน ..... 123
36	กราฟแสดงแนวโน้มของลักษณะของการเปลี่ยนคำตอบประเภทกรทวนคำตอบแล้ว มีการเปลี่ยนคำตอบจากผิดเป็นผิด ของผู้สอบที่มีระดับความสามารถแตกต่างกัน ..... 123
37	กราฟแสดงแนวโน้มของลักษณะของการเปลี่ยนคำตอบประเภทกรทวนคำตอบแล้ว มีการเปลี่ยนคำตอบจากถูกเป็นผิด ของผู้สอบที่มีระดับความสามารถแตกต่างกัน ..... 124
38	กราฟแสดงแนวโน้มของค่าความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของ ผู้สอบ (tdif) ของผู้สอบที่มีระดับความสามารถแตกต่างกัน..... 127
39	การแบ่งส่วนประกอบโปรแกรมการทดสอบแบบปรับเหมาะตาม ..... 132

# บทที่ 1

## บทนำ

### ความเป็นมาและความสำคัญของปัญหา

ในสมัยก่อนการจัดเก็บข้อสอบเพื่อนำกลับมาใช้ นิยมทำโดยการพิมพ์ข้อสอบลงบนกระดาษแข็ง แล้วนำข้อสอบจัดเก็บใส่ในกล่องแยกตามวิชา ต่อมาเมื่อมีการนำคอมพิวเตอร์มาใช้ในการวัดและประเมินผลทางการศึกษา จึงมีวิธีการจัดเก็บข้อสอบด้วยเครื่องคอมพิวเตอร์ และพัฒนาเป็นคลังข้อสอบหรือธนาคารข้อสอบ (item pool หรือ item bank) ซึ่งเป็นการเก็บรวบรวมข้อสอบที่มีคุณสมบัติและค่าสถิติของข้อสอบแตกต่างกันเข้าไว้ด้วยกัน โดยอาจวัดเนื้อหาหรือทักษะที่แตกต่างกัน หรือมีค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบที่แตกต่างกัน เช่น ค่าความยาก, ค่าอำนาจจำแนก หรือ ค่าโอกาสในการเดาข้อสอบได้ถูก รวมทั้งมีประโยชน์ในการเลือกหรือสุ่มข้อสอบไปใช้ตามวัตถุประสงค์ของผู้ใช้ (เสวี ชัดเข้ม, 2541)

ต่อมาได้มีการนำคอมพิวเตอร์มาช่วยในการทดสอบ (computerized testing) แทนการทดสอบโดยใช้กระดาษ หลักการสำคัญคือ การสร้างคลังข้อสอบเก็บไว้ในเครื่องคอมพิวเตอร์ แล้วนำเสนอแบบสอบบนหน้าจอคอมพิวเตอร์ ผู้สอบตอบข้อสอบโดยการกดตัวอักษรหรือตัวเลขบนแป้นพิมพ์ ผู้สอบทุกคนสอบโดยใช้แบบสอบฉบับเดียวกัน ใช้เวลาสอบเท่ากัน และสอบในสถานการณ์เดียวกัน รวมทั้งมีการใช้คอมพิวเตอร์ตรวจให้คะแนนและรายงานผลการสอบ ข้อแตกต่างระหว่างการสอบโดยใช้คอมพิวเตอร์กับการสอบโดยใช้กระดาษที่สำคัญคือ การสอบโดยใช้คอมพิวเตอร์เป็นการใช้หน้าจอคอมพิวเตอร์เป็นสื่อแสดงข้อสอบและตอบข้อสอบโดยการกดแป้นพิมพ์ ส่วนการสอบโดยใช้กระดาษเป็นการใช้หน้ากระดาษเป็นสื่อแสดงข้อสอบและตอบข้อสอบโดยการเขียนด้วยปากกาหรือดินสอ เรียกได้ว่าเป็นยุคแรกเริ่ม (the first generation) ของการนำคอมพิวเตอร์มาใช้ในการบริหารการสอบ (Bunderson and Others, 1993 อ้างถึงใน Linn, 1993)

ปัจจุบันได้มีการพัฒนามาเป็นการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (computerized adaptive testing - CAT) ซึ่งเป็นการใช้คอมพิวเตอร์ช่วยบริหารการสอบทุกขั้นตอน ผู้สอบจะทำแบบสอบที่ออกแบบให้เหมาะสมกับระดับความสามารถของแต่ละคน (Hambleton & Sireci, 1997) โดยหลักการแล้วในการเลือกข้อสอบข้อแรก โปรแกรมคอมพิวเตอร์จะเลือกข้อสอบที่ยากง่ายปานกลางจากคลังข้อสอบ ในกรณีที่ผู้สอบทำได้ โปรแกรมจะเลือกข้อสอบข้อถัดไปที่ยากขึ้น หรือในกรณีที่ผู้สอบทำไม่ได้ โปรแกรมจะเลือกข้อสอบข้อถัดไปที่ง่ายลง โดยที่ผู้สอบไม่สามารถที่จะกลับไปเปลี่ยนคำตอบข้อก่อนหน้าได้ กระบวนการนี้จะดำเนิน

ต่อไปเรื่อยๆ จนกระทั่งสามารถประมาณความสามารถของผู้สอบแต่ละคนได้แล้ว โปรแกรมคอมพิวเตอร์จึงหยุดทำงาน (สุพัฒน์ สุกมลสันต์, 2535)

จะเห็นได้ว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีประโยชน์ต่อการวัดและประเมินผลทางการศึกษาอย่างมาก ศิริชัย กาญจนวาสี (2545) ได้ยกตัวอย่างไว้ดังนี้

- 1) ทำให้การทดสอบมีความคล่องตัว สามารถทดสอบเป็นรายบุคคลตามความพร้อมของผู้สอบ มีการตรวจให้คะแนน และรายงานผลเป็นไปอย่างอัตโนมัติ
- 2) ระบบมีความเป็นมาตรฐาน ส่งเสริมการรักษาความลับของข้อสอบและยุติธรรม
- 3) ช่วยเพิ่มคุณภาพและประสิทธิภาพของการวัดความสามารถ/การเรียนรู้ของผู้สอบได้อย่างรวดเร็วและมีความถูกต้องยิ่งขึ้น
- 4) ช่วยให้การประเมินผลพัฒนาการหรือการติดตามการเปลี่ยนแปลงความสามารถ/การเรียนรู้ เป็นไปอย่างต่อเนื่อง และปฏิบัติได้อย่างสะดวกรวดเร็ว

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เป็นที่นิยมแพร่หลายเป็นอันมากในสหรัฐอเมริกา สำหรับในประเทศไทยมีแนวโน้มว่ามีการนำคอมพิวเตอร์มาช่วยในกระบวนการวัดผลมากขึ้น จึงมีความเป็นไปได้สูงที่จะนำการทดสอบแบบปรับเหมาะมาใช้แทนการสอบโดยใช้กระดาษ (Paper and Pencil Test) โดยเฉพาะอย่างยิ่งการวัดผลสัมฤทธิ์ระหว่างเรียนเพื่อปรับปรุงการเรียนการสอนและการสอนซ่อมเสริม แต่การสร้างแบบทดสอบแบบปรับเหมาะ มีองค์ประกอบที่มีผลต่อประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบได้แก่

- 1) มีคลังข้อสอบขนาดใหญ่ คลังข้อสอบจะต้องประกอบด้วยชุดของข้อสอบที่มุ่งวัดลักษณะหรือความสามารถที่ต้องการได้ ข้อสอบจะต้องครอบคลุมระดับความสามารถต่างๆ กัน มีจำนวนข้อสอบที่มากพออย่างน้อย 100 – 200 ข้อ (Weiss, 1998) ถ้าจะให้มีประสิทธิภาพดีจะต้องมีข้อสอบจำนวน  $2^n$  เมื่อ  $n$  คือจำนวนข้อสอบโดยประมาณต่อคนที่ใช้ในการทดสอบ เช่น ถ้าต้องการสอบทุกคนด้วยข้อสอบประมาณ 10 ข้อ จำนวนข้อสอบในคลังควรมีอย่างน้อย 1,024 ข้อ เป็นต้น แต่ถ้าการทดสอบมีการกำหนดจำนวนข้อที่จะใช้คงที่เหมือนกันทุกคน จำนวนข้อสอบควรมีอย่างน้อย  $[n(n+1)] \div 2$  ข้อ เมื่อ  $n$  คือจำนวนข้อสอบ เช่น ถ้าต้องการสอบทุกคนด้วยข้อสอบ 20 ข้อเท่านั้น จำนวนข้อสอบในคลังควรมีอย่างน้อย 210 ข้อ เป็นต้น (Allen & Yen, 1979) โดยข้อสอบแต่ละข้อควรมีอำนาจจำแนกสูง มีค่าความยากที่ครอบคลุมระดับช่วงต่างๆ อย่างเหมาะสม และควรมีจำนวนข้อสอบเท่าๆ กัน ในแต่ละระดับค่าความยาก
- 2) มีวิธีการคัดเลือกข้อสอบและตรวจให้คะแนนอย่างมีประสิทธิภาพ เพื่อให้การทดสอบแบบปรับเหมาะเป็นไปอย่างมีประสิทธิภาพสูง ควรใช้กลยุทธ์แบบแยกทางแปรผัน บนหมวดของทฤษฎี

การตอบสนองข้อสอบ (IRT) ที่ใช้คอมพิวเตอร์ในการทดสอบ ควรใช้วิธีการคัดเลือกข้อสอบแบบที่ให้สารสนเทศสูงที่สุด (Maximum Information Strategy) และประมาณค่าความสามารถที่แท้จริง ( $\theta$ ) โดยใช้วิธี Maximum Likelihood Estimation หรืออาจใช้วิธี Bayesian Estimation ก็ได้ แต่จากการศึกษาของรังสรรค์ มณีเล็ก (2539) พบว่า การประมาณค่าความสามารถด้วยวิธี Bayesian Estimation มีความตรงเชิงสภาพสูงกว่าการประมาณค่าแบบวิธี Maximum Likelihood Estimation

3) มีเกณฑ์การยุติที่เหมาะสม การยุติการทดสอบควรใช้เกณฑ์ที่ทำให้ผลการวัดที่ได้มีความน่าเชื่อถือทัดเทียมกัน สำหรับผู้สอบทุกคน (Equiprecise) ในทุกระดับความสามารถ ซึ่งการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ โดยใช้รูปแบบแยกทางแปรผัน นิยมใช้เกณฑ์การยุติการสอบ 2 วิธีคือ

3.1 การกำหนดจำนวนข้อสอบ ซึ่งมีข้อดีในการเปรียบเทียบสารสนเทศของแบบสอบได้โดยตรง

3.2 การกำหนดค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ (Hambleton, 1991)

ดังนั้นในการนำเอาการทดสอบแบบปรับเหมาะไปใช้จึงจำเป็นที่จะต้องคำนึงถึงสิ่งเหล่านี้เป็นสำคัญ เพื่อให้ได้แบบทดสอบที่มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบ ในด้านรูปแบบต้องเป็นรูปแบบที่ง่ายในการนำไปปฏิบัติ คือ ง่ายต่อการใช้ ง่ายต่อความเข้าใจของทั้งผู้ใช้และผู้สอบ ในด้านจำนวนข้อสอบต้องคำนึงถึงความยาวของแบบทดสอบ ถ้ามากเกินไปอาจทำให้เสียเวลา เสียค่าใช้จ่ายและทำให้เกิดความเบื่อหน่ายแก่ผู้สอบได้ แต่ถ้าข้อสอบน้อยเกินไป อาจทำให้ได้ข้อมูลในการประมาณค่าความสามารถที่คลาดเคลื่อนสูง ดังนั้นจึงควรใช้จำนวนข้อสอบที่เหมาะสม คือจำนวนน้อยแต่ให้ความตรงและความเที่ยงสูง ทั้งยังต้องประมาณค่าความสามารถได้อย่างถูกต้องแม่นยำ ในด้านวิธีให้คะแนนต้องไม่มีความลำเอียงและสามารถประมาณค่าความสามารถของผู้สอบได้ตรงกับความสามารถที่แท้จริงของผู้สอบมากที่สุด

แต่ในปัจจุบันยังมีประเด็นของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (computerized adaptive testing - CAT) ที่ยังเป็นข้อถกเถียงกันอยู่ นั่นคือ หลังจากที่ผู้สอบทำการทดสอบเสร็จ ผู้สอบควรได้รับโอกาสให้กลับไปทวนคำตอบหรือไม่ เป็นที่ชัดเจนว่าผู้สอบยอมต้องการที่จะกลับไปทวนคำตอบ ซึ่งน่าจะได้มาถึงการประมาณค่าความสามารถที่แม่นยำหลังทวน และด้วยข้อเท็จจริงที่ว่า การเลือกคำตอบครั้งแรกอาจไม่ใช่คำตอบที่ถูกต้องที่สุดเสมอ แต่นักวิจัยที่มีชื่อเสียงบางท่านต่างมีความเห็นว่าไม่ควรเปิดโอกาสให้มีการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เนื่องจากข้อจำกัดของกระบวนการทดสอบ ตัวอย่างเช่น



Wainer (1993) กล่าวว่า สาเหตุที่ไม่ควรให้มีการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ คือจะสูญเสียความมีประสิทธิภาพ และเกิดความลำเอียงในการประมาณค่าความสามารถของผู้สอบ รวมทั้งเกิดความคลาดเคลื่อนในการได้มาถึงความสามารถที่แท้จริงของผู้สอบในกระบวนการการทดสอบ ทำให้ขัดแย้งกับกระบวนการสร้างแบบทดสอบแบบปรับเหมาะที่มีประสิทธิภาพ โดย Wainer ได้ยกตัวอย่างสถานการณ์ที่ทำให้ผู้สอบได้คะแนนสูงขึ้น กล่าวคือ เมื่อผู้สอบตั้งใจที่จะตอบคำถามผิด ดังนั้นจึงทำให้ได้รับข้อสอบข้อต่อไปที่มีระดับความยากต่ำลง และถ้าเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้โดยไม่มีเงื่อนไข ผู้สอบก็จะสามารถกลับไปแก้คำตอบข้อก่อนหน้า ทำให้ข้อสอบแต่ละข้อนั้นไม่สัมพันธ์กับระดับความสามารถที่แท้จริงของผู้สอบ จากงานวิจัยพบว่าผู้สอบจะได้รับการประมาณค่าความสามารถโน้มไปทางด้านบวกหรือได้รับคะแนนที่สูงเกินความสามารถที่แท้จริง และมีความคลาดเคลื่อนสูง

Green, Bock, Humphreys, Linn, and Reckase (1984) กล่าวว่า อาจเกิดความสับสนและความล่าช้าในการทำการทดสอบ จากการที่ผู้สอบต้องคอยนี้่กย้อนถึงการตอบข้อสอบข้อที่ผ่านมา

การต่อต้านการทวนคำตอบใน CAT ข้างต้น จะเห็นได้ว่าเกิดจากปัญหาที่มีลักษณะเฉพาะจากการทวนคำตอบ ดังนั้นถึงแม้ว่าการทวนคำตอบจะก่อให้เกิดผลดีมากมายแต่ก็เฉพาะกรณีที่จะสามารถขจัดปัญหาต่างๆที่เกิดขึ้นให้น้อยลงหรือหมดไปเท่านั้น (Bowles, 2001; Weise, 1996) 3 วิธีการที่นำมาใช้ในการทวนคำตอบ (Stocking, 1997) ที่สามารถควบคุมการประมาณค่าความสามารถที่เกินจริงได้ คือ การทวนที่มีการจำกัดจำนวน (Limited Review), การทวนที่มีการจำกัดช่วง (Block Review) และการทวนในหมวดสิ่งเร้าเดียวกัน (Stimulus-based Review) ซึ่งแต่ละวิธีมีหลักการ ดังนี้

1) การทวนที่มีการจำกัดจำนวน (Limited Review) คือ ผู้สอบอาจย้อนกลับไปทวนคำตอบได้ทั้งหมด แต่จะเปลี่ยนคำตอบได้ในจำนวนที่จำกัดเท่านั้น

2) การทวนที่มีการจำกัดช่วง (Block Review) คือ ผู้สอบจะสามารถทวนคำตอบพร้อมทั้งเปลี่ยนคำตอบภายในช่วงที่จำกัดไว้ได้ทั้งหมด แต่ไม่สามารถที่จะทวนข้ามช่วงได้ ยกตัวอย่างเช่น การจำกัดช่วง 10 ข้อ กล่าวคือผู้สอบจะสามารถกลับมาทวนคำตอบภายในช่วง (ข้อ 1 – 10) ได้หลังจากที่ทำเสร็จข้อ 10 หลังจากนั้นจะไปทำข้อสอบในช่วงถัดไป (ข้อ 11 - 20) และจะสามารถกลับมาทวนคำตอบในช่วงนี้ได้ก็ต่อเมื่อทำเสร็จถึงข้อ 20 แต่ไม่สามารถกลับไปทวนคำตอบในช่วงก่อนหน้านี้ (ข้อ 1 -10) ได้ เป็นต้น

3) การทวนในหมวดสิ่งเร้าเดียวกัน (Stimulus-based Review) มีความคล้ายคลึงกับการทวนที่มีการจำกัดช่วง แต่มีความแตกต่างที่ความยาวของช่วง กล่าวคือ การทวน

ที่มีการจำกัดช่วงนั้นจะแบ่งความยาวของข้อสอบออกเป็นช่วง ช่วงละเท่าๆ กัน แต่การทวนในหมวดสิ่งเร้าเดียวกันจะแบ่งความยาวช่วงตามหมวดสิ่งเร้า เช่น แบบทดสอบเรื่องทศนิยมประกอบด้วยเนื้อหาเรื่อง การแปลงเศษส่วนให้เป็นทศนิยม, การอ่านทศนิยม, การบวกทศนิยม และการลบทศนิยม ดังนั้นจึงแบ่ง หมวดสิ่งเร้าตามเนื้อหาของแบบทดสอบเป็น 4 หมวด กล่าวคือ การแปลงเศษส่วนให้เป็นทศนิยม (ข้อ 1 - 4), การอ่านทศนิยม (ข้อ 5 - 10), การบวกทศนิยม (ข้อ 11 - 15) และการลบทศนิยม (ข้อ 16 - 20) ซึ่งจะเห็นได้ว่าช่วงของการทวนจะมีความยาวของข้อสอบไม่เท่ากัน

Vispole และคณะ (2000) ได้สนใจวิธีการทวนที่มีการจำกัดช่วง (Block Review) จึงทำการศึกษาเปรียบเทียบแต่ละเงื่อนไขการทวน คือ การไม่ให้ทวน, การทวนที่มีการจำกัดช่วง 5 ข้อ, การทวนที่มีการจำกัดช่วง 10 ข้อ และการทวนที่มีการจำกัดช่วง 40 ข้อ พบว่า ในขณะที่ลดขนาดช่วงลง ผู้สอบจะทำเครื่องหมายเพื่อกลับมาทวนเพิ่มขึ้น เปลี่ยนคำตอบมากขึ้น ใช้เวลาในการทวนคำตอบมากขึ้น มีแนวโน้มใช้เวลาในการทำข้อสอบนานขึ้น และค่าความสามารถเฉลี่ยมากกว่าเงื่อนไขการไม่ให้ทวน แต่ทว่าไม่แตกต่างกันมากระหว่างเงื่อนไขการให้ทวนในช่วงต่างๆ ดังนั้น Vispole จึงได้ให้ข้อเสนอว่า การทวนแบบจำกัดช่วงสั้นๆ นั้นทำหน้าที่ได้ดีพอๆ กับการทวนได้ทั้งหมด แต่สามารถลดปัญหาการโกงข้อสอบจากการเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้ เนื่องจากผู้สอบต้องตอบข้อสอบให้เสร็จก่อนภายในช่วง จึงจะสามารถกลับไปทวนคำตอบได้ จึงเป็นการประมาณค่าความสามารถของผู้สอบไว้แล้วครั้งหนึ่ง รวมทั้งยังมีการประมาณค่าความสามารถของผู้สอบระหว่างช่วงด้วยเมื่อผู้สอบทำการทวนเสร็จในช่วง

จากประเด็นการศึกษาเกี่ยวกับการเปิดโอกาสให้ผู้สอบสามารถทำการทวนคำตอบข้างต้น การเปิดโอกาสให้ทวนคำตอบโดยวิธีการทวนที่มีการจำกัดช่วง (Block Review) ของ Vispole พบว่าประเด็นที่ชัดเจน คือสามารถที่จะลดปัญหาจากการโกงข้อสอบในการเปิดโอกาสให้ทวนคำตอบ ส่วนประเด็นที่ยังไม่ชัดเจนคือ ค่าความสามารถของผู้สอบจากการเปิดโอกาสให้มีการทวนคำตอบมีค่ามากกว่าเงื่อนไขการไม่ให้ทวน ในทุกกรณีของผู้สอบที่มีระดับความสามารถแตกต่างกันหรือไม่

ประณาท เทียนศรี (2545) ได้สรุปผลจากงานวิจัยเรื่องปัจจัยและผลการเปลี่ยนคำตอบ พบว่า ข้อสอบที่นักเรียนมักมีการเปลี่ยนคำตอบเมื่อกลับไปทวนส่วนใหญ่เป็นข้อสอบที่มีระดับความยากอยู่ที่ค่อนข้างยากจนถึงยากมาก และนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนดี (ระดับความสามารถสูง) เมื่อมีการเปลี่ยนคำตอบใหม่ก็จะเป็นคำตอบที่ถูกต้องมากกว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่ำ (ระดับความสามารถต่ำ) ซึ่งจะสามารถตั้งข้อสังเกตได้ว่า การทวนคำตอบนั้นมีประโยชน์ต่อผู้สอบทุกคนอย่างเท่าเทียมกันหรือไม่

โดยหลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีเป้าหมายที่จะดำเนินการทดสอบด้วยการคัดเลือกข้อสอบให้เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล ดังนั้นผู้สอบที่มีระดับความสามารถสูงก็จะได้รับข้อสอบที่ยาก ส่วนผู้ที่มีระดับความสามารถต่ำก็จะได้รับคำตอบที่ง่ายดังนั้นการกลับไปทวนคำตอบของผู้สอบที่มีระดับความสามารถอื่นๆ ซึ่งไม่ใช่ระดับความสามารถสูงแล้วทำให้มีการเปลี่ยนคำตอบจากถูกเป็นผิด หรือจากผิดเป็นผิด รวมทั้งมีการใช้เวลาในการทำข้อสอบนานขึ้นเนื่องจากเสียเวลากลับไปทวนคำตอบนั้น จะเกิดผลดีอย่างแท้จริงหรือไม่กับการเปิดโอกาสให้มีการทวนคำตอบในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จากข้างต้นผู้วิจัยจึงสนใจที่จะศึกษาเปรียบเทียบนักเรียนที่มีระดับความสามารถแตกต่างกันว่าจะให้ค่าความสามารถและลักษณะของการเปลี่ยนคำตอบของผู้สอบที่แตกต่างกันหรือไม่อย่างไรในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบแตกต่างกัน กล่าวคือ ภายใต้การทดสอบที่ให้ทวนคำตอบด้วยวิธีการทวนที่มีการจำกัดช่วง (Block Review) กับการทดสอบที่ไม่ให้ทวนคำตอบ เพื่อนำไปเป็นแนวทางในการยอมรับหรือไม่ยอมรับการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ในอนาคตต่อไป

### คำถามของการวิจัย

จากการศึกษาเปรียบเทียบนักเรียนที่มีระดับความสามารถแตกต่างกันว่าจะให้ความแตกต่างของค่าความสามารถของผู้สอบ และความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบแตกต่างกัน กล่าวคือ ภายใต้การทดสอบที่ให้ทวนคำตอบด้วยวิธีการทวนที่มีการจำกัดช่วง (Block Review) กับการทดสอบที่ไม่ให้ทวนคำตอบ รวมทั้งศึกษาเปรียบเทียบลักษณะของการเปลี่ยนคำตอบภายใต้การทดสอบที่ให้ทวนคำตอบ ว่าแตกต่างกันหรือไม่อย่างไร ผู้วิจัยจึงได้ตั้งคำถามของการวิจัยไว้ดังนี้

1) ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบกับการทดสอบที่ไม่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน จะให้ความแตกต่างของค่าความสามารถของผู้สอบแตกต่างกันหรือไม่ เพียงใด

2) ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน จะให้ลักษณะของการเปลี่ยนคำตอบแตกต่างกันหรือไม่ เพียงใด โดยแบ่งลักษณะของการเปลี่ยนคำตอบออกเป็น 2 ประเภท คือ

2.1 การทวนคำตอบแล้วยังคงคำตอบเดิมไว้

2.2 การทวนคำตอบแล้วมีการเปลี่ยนคำตอบ

3) ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบกับการทดสอบที่ไม่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน จะให้ความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบแตกต่างกันหรือไม่ เพียงใด

### วัตถุประสงค์ของการวิจัย

1) เพื่อพัฒนาโปรแกรมคอมพิวเตอร์วิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถม ศึกษปีที่ 6 สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ ที่สามารถกำหนดเงื่อนไขการทดสอบได้ 2 เงื่อนไข คือ การทดสอบที่ให้ทวนคำตอบได้โดยวิธีการทวนที่มีการจำกัดช่วง (Block Review) กับการทดสอบที่ไม่ให้ทวนคำตอบ

2) เพื่อเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบกับการทดสอบที่ไม่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน

3) เพื่อเปรียบเทียบลักษณะของการเปลี่ยนคำตอบ ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน โดยแบ่งลักษณะของการเปลี่ยนคำตอบออกเป็น 2 ประเภท คือ

3.1 การทวนคำตอบแล้วยังคงคำตอบเดิมไว้

3.2 การทวนคำตอบแล้วมีการเปลี่ยนคำตอบ

4) เพื่อเปรียบเทียบความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบกับการทดสอบที่ไม่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน

### สมมติฐานของการวิจัย

จากการศึกษาของประณาท เทียนศรี (2546) ที่ได้ศึกษาปัจจัยและผลจากการเปลี่ยนคำตอบข้อสอบปรนัย โดยพบว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ระดับดีมาก ส่วนใหญ่เปลี่ยนคำตอบใหม่แล้วถูกต้องมากขึ้น ไม่ว่าข้อสอบนั้นจะมีความยากระดับง่าย

หรือระดับปานกลาง ส่วนนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ระดับต่ำกว่าปานกลางและต้องปรับปรุง จะเปลี่ยนคำตอบใหม่แล้วผิดมากกว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ระดับดีกว่า

รวมทั้ง วอลเตอร์ พี. วิสโพเอล, เอมี บี เฮนดริกสัน, ทิมโมธี เบลเลอร์ (2000) ได้ทำการศึกษาเพิ่มเติมเกี่ยวกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่อนุญาตให้ทวนคำตอบแบบชงนั้น พบว่าเวลาที่ใช้ในการสอบ ภายใต้เงื่อนไขการทวนคำตอบแบบต่างๆ รวมทั้งค่าความสามารถของผู้สอบเพิ่มขึ้นเล็กน้อยหลังการทวนคำตอบ คำตอบส่วนใหญ่ก็จากผิดไปถูกมากกว่าก็จากถูกไปผิด ผู้สอบที่เปลี่ยนคำตอบส่วนมากจะได้รับค่าความสามารถเพิ่มขึ้น ผู้วิจัยจึงได้ตั้งสมมติฐานของการวิจัย ดังนี้

1) ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบกับการทดสอบที่ไม่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน น่าจะให้ความแตกต่างของค่าความสามารถของผู้สอบแตกต่างกัน โดยที่ผู้สอบที่มีระดับความสามารถสูงและปานกลาง ค่าความสามารถของผู้สอบภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ น่าจะมากกว่าการทดสอบที่ไม่ให้ทวนคำตอบ ส่วนผู้สอบที่มีระดับความสามารถต่ำ ค่าความสามารถของผู้สอบภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ น่าจะน้อยกว่าการทดสอบที่ไม่ให้ทวนคำตอบ

2) ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน น่าจะให้ลักษณะของการเปลี่ยนคำตอบ ประเภท

2.1 การทวนคำตอบแล้วยังคงคำตอบเดิมไว้ ไม่แตกต่างกันในแต่ละระดับความสามารถของผู้สอบ

2.2 การทวนคำตอบแล้วมีการเปลี่ยนคำตอบ น่าจะแตกต่างกันในแต่ละระดับความสามารถของผู้สอบ โดยผู้สอบที่มีระดับความสามารถสูงและปานกลางน่าจะมีการเปลี่ยนคำตอบจากผิดเป็นถูก แต่ผู้สอบที่มีระดับความสามารถต่ำน่าจะมีการเปลี่ยนคำตอบจากผิดเป็นผิด หรือจากถูกเป็นผิด

3) ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน น่าจะให้ความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบไม่แตกต่าง โดยที่ผู้สอบทุกระดับความสามารถน่าจะใช้เวลาในการทำข้อสอบภายใต้การทดสอบที่ให้ทวนคำตอบมากกว่าการทดสอบที่ไม่ให้ทวนคำตอบ

## ขอบเขตของการวิจัย

1) ประชากรที่ใช้ในการวิจัยครั้งนี้ ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2549 โรงเรียนประถมศึกษา สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) กรุงเทพมหานครจำนวน 36 โรงเรียน รวมนักเรียน 5,592 คน (สำนักรายงานและแผนการศึกษาขั้นพื้นฐาน สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ , 2549)

2) สถานการณ์ที่สนใจศึกษา คือ

### 2.1 สถานการณ์ที่หนึ่ง

ตัวแปรต้น (independent variable) ► ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ตัวแปรตาม (dependent variable) ► ความแตกต่างของค่าความสามารถของผู้สอบ

### 2.2 สถานการณ์ที่สอง

ตัวแปรต้น (independent variable) ► ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ตัวแปรตาม (dependent variable) ► ลักษณะของการเปลี่ยนคำตอบของผู้สอบ

### 2.3 สถานการณ์ที่สาม

ตัวแปรต้น (independent variable) ► ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ตัวแปรตาม (dependent variable) ► ความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ

## ข้อจำกัดของการวิจัย

1) ค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบมีการกระจายเป็นช่วงๆ ด้วยจำนวนข้อสอบที่แตกต่างกัน จึงทำให้การคัดเลือกข้อสอบจากคลังข้อสอบในบางครั้งไม่ได้ข้อสอบตามที่ต้องการและจำเป็นต้องใช้ข้อสอบที่มีคุณลักษณะใกล้เคียงที่สุดขึ้นมาแทน

2) โปรแกรมคอมพิวเตอร์ที่ผู้วิจัยสร้างขึ้น ผู้นำไปใช้และผู้สอบไม่สามารถเปลี่ยนแปลงข้อสอบและเกณฑ์ยุติการทดสอบ 20 ข้อ ที่ผู้วิจัยกำหนดไว้ได้

3) ผู้วิจัยต้องการเจาะจงโรงเรียนประถมศึกษาขนาดใหญ่ ในกรุงเทพมหานคร ที่มีนักเรียนประถมศึกษาปีที่ 6 ทั้งระดับชั้นมากที่สุด เพื่อนำมาใช้ในการหาค่าพารามิเตอร์ของข้อสอบ ตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ เนื่องจากนักเรียนจากโรงเรียนเดียวกัน จะมีสภาพแวดล้อมต่างๆ ในการทำข้อสอบเหมือนกัน ไม่เกิดการได้เปรียบ เสียเปรียบ ในการทำข้อสอบ จึงได้จำนวนนักเรียนคิดเป็น 409 คนเท่านั้น

## คำจำกัดความที่ใช้ในการวิจัย

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ หมายถึง การนำคอมพิวเตอร์มาใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

เรื่องความเข้าใจเศษส่วน ด้วยข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบ ทั้งนี้ข้อสอบที่ถูกเลือกให้แก่ผู้สอบแต่ละข้อนั้นขึ้นอยู่กับผลของการตอบข้อสอบข้อที่ผ่านมา ซึ่งดำเนินการโดยใช้โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

**โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์** หมายถึง การทดสอบโดยใช้คอมพิวเตอร์ ซึ่งผู้สอบจะต้องทำการทดสอบ 2 เงื่อนไขรวม 2 ครั้ง และลำดับของเงื่อนไขการทดสอบจะถูกกำหนดโดยเลขที่ของนักเรียน กล่าวคือ ถ้าเลขที่เป็นเลขคี่ จะทดสอบครั้งที่ 1 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การทดสอบที่ให้ทวนคำตอบ และถ้าเลขที่เป็นเลขคู่ จะทดสอบครั้งที่ 1 ภายใต้การทดสอบที่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ โดยบรรจุคลังข้อสอบวิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 และปรับเหมาะกับความสามารถของผู้สอบโดยใช้รูปแบบแยกทางแปรผัน การประมาณค่าความสามารถของผู้สอบด้วยวิธีการของเบส์ คัดเลือกข้อสอบข้อต่อไปโดยดูค่าสารสนเทศของข้อสอบข้อที่สูงที่สุด (Maximum item information) ยุติการทดสอบแบบคงที่ที่ 20 ข้อ ปฏิบัติการบนระบบวินโดวส์เอ็กซ์พี (Windows XP) และใช้โปรแกรมไมโครซอฟท์ วิวอล เบสิก (Microsoft Visual Basic) ในการเขียน

**คลังข้อสอบ** หมายถึง กลุ่มของข้อสอบที่มีคุณภาพ ซึ่งสร้างขึ้นตามหลักเกณฑ์และถูกเก็บไว้เป็นระบบ ในการเลือกข้อสอบขึ้นมาใช้นั้นต้องเลือกตามเกณฑ์ที่กำหนดขึ้น คลังข้อสอบที่ใช้ในการวิจัยนี้ เป็นกลุ่มข้อสอบที่ใช้วัดความเข้าใจเรื่องเศษส่วน มีข้อสอบทั้งหมดอย่างน้อย 210 ข้อ โดย 200 ข้อ เป็นของ รังสรรค์ มณีเล็ก (2539) อีกอย่างน้อย 10 ข้อ เป็นของผู้วิจัย ลักษณะของข้อสอบเป็นแบบเลือกตอบ 4 ตัวเลือก แต่ละข้อได้รับการวิเคราะห์โดยใช้ทฤษฎีการตอบสนองข้อสอบ มีค่าอำนาจจำแนก ค่าความยากง่าย ค่าโอกาสในการเดาข้อสอบได้ถูก และค่าความสามารถที่ทำให้ข้อสอบมีสารสนเทศสูงสุดกำกับอยู่ โดยจะเก็บข้อสอบและค่าพารามิเตอร์ของข้อสอบแยกกัน เมื่อต้องการจะเรียกข้อสอบขึ้นมาที่หน้าจอคอมพิวเตอร์ ก็จะใช้ค่าพารามิเตอร์เป็นตัวชี้

**เงื่อนไขการทดสอบ** หมายถึง การทดสอบที่ไม่ให้ทวนคำตอบและการทดสอบที่ให้ทวนคำตอบ

**การทดสอบที่ไม่ให้ทวนคำตอบ** หมายถึง การไม่เปิดโอกาสให้ผู้สอบทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ คือเมื่อผู้สอบตอบคำตอบแต่ละข้อเสร็จโปรแกรมคอมพิวเตอร์ก็จะคัดเลือกข้อสอบข้อต่อไป ซึ่งผู้สอบจะไม่

สามารถกลับไปทวนคำตอบในข้อที่ผ่านมา กระบวนการนี้จะดำเนินต่อไปจนกระทั่งผู้สอบทำข้อสอบครบ 20 ข้อ

**การทดสอบที่ให้ทวนคำตอบ** หมายถึง การเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ได้ ด้วยวิธีการทวนที่มีการจำกัดช่วง (Block Review) 5 ข้อ ซึ่งโปรแกรมคอมพิวเตอร์จะเปิดโอกาสให้ผู้สอบสามารถกลับมาทวนคำตอบภายในช่วง 5 ข้อได้หลังจากที่ทำเสร็จข้อ 5 ในแต่ละช่วง หลังจากนั้นจะไปทำข้อสอบในช่วง 5 ข้อถัดไป แต่จะไม่สามารถกลับไปทวนคำตอบในช่วง 5 ข้อนี้ได้อีก กระบวนการนี้จะดำเนินต่อไปจนกระทั่งผู้สอบทำข้อสอบครบ 20 ข้อ

**คำตอบจากการทดสอบที่ให้ทวนคำตอบรอบที่ 1** หมายถึง ผลลัพธ์จากการตอบคำถามของผู้สอบในเงื่อนไขการทดสอบที่ให้ทวนคำตอบ ก่อนที่ผู้สอบจะสามารถกลับไปทวนคำตอบภายในช่วง 5 ข้อได้ในแต่ละช่วง จนครบทั้ง 20 ข้อ ซึ่งโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะบันทึกคำตอบของผู้สอบไว้เป็นคำตอบรอบที่ 1

**คำตอบจากการทดสอบที่ให้ทวนคำตอบรอบที่ 2** หมายถึง ผลลัพธ์จากการตอบคำถามของผู้สอบในเงื่อนไขการทดสอบที่ให้ทวนคำตอบ หลังจากผู้สอบกลับไปทวนคำตอบภายในช่วง 5 ข้อ ในแต่ละช่วง จนครบทั้ง 20 ข้อ ซึ่งโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะบันทึกคำตอบของผู้สอบไว้เป็นคำตอบรอบที่ 2

**ลักษณะของการเปลี่ยนคำตอบของผู้สอบ** หมายถึง ความแตกต่างของคำตอบจากการทดสอบที่ให้ทวนคำตอบรอบที่ 1 และรอบที่ 2 ซึ่งแบ่งลักษณะของการเปลี่ยนคำตอบออกเป็น 2 ประเภท คือ การทวนคำตอบแล้วยังคงคำตอบเดิมไว้ และการทวนคำตอบแล้วมีการเปลี่ยนคำตอบ

**การทวนคำตอบแล้วยังคงคำตอบเดิมไว้** หมายถึง จำนวนข้อที่ไม่มี ความแตกต่างของคำตอบจากการทดสอบที่ให้ทวนคำตอบรอบที่ 1 และรอบที่ 2

**การทวนคำตอบแล้วมีการเปลี่ยนคำตอบ** หมายถึง จำนวนข้อที่มีความแตกต่างของคำตอบจากการทดสอบที่ให้ทวนคำตอบรอบที่ 1 และรอบที่ 2 ซึ่งแบ่งเป็น 3 ลักษณะย่อยคือ จำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นถูก, จำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นผิด และจำนวนข้อที่เปลี่ยนคำตอบจากถูกเป็นผิด



**เวลาที่ใช้ในการทำข้อสอบ** หมายถึง ระยะเวลาที่ใช้ในการทำข้อสอบทั้ง 20 ข้อ ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ และการทดสอบที่ไม่ให้ทวนคำตอบ โดยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะเริ่มต้นที่เวลาตั้งแต่ข้อสอบข้อแรกปรากฏที่หน้าจอคอมพิวเตอร์ และหยุดการบันทึกเมื่อผู้สอบตอบข้อที่ 20 เสร็จ

**ความแตกต่างของเวลาที่ใช้ในการทำข้อสอบ** หมายถึง ผลต่างของเวลาที่ใช้ในการทำข้อสอบทั้ง 20 ข้อ ภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบและการทดสอบที่ไม่ให้ทวนคำตอบ ( $\text{เวลา}_{\text{ให้ทวน}} - \text{เวลา}_{\text{ไม่ให้ทวน}}$ ) ซึ่งถ้าความแตกต่างออกมาเป็นค่าบวก จะแสดงถึง เวลาที่ใช้ในการทำข้อสอบของผู้สอบภายใต้เงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบ แต่ถ้าความแตกต่างออกมาเป็นค่าลบ จะแสดงถึง เวลาที่ใช้ในการทำข้อสอบของผู้สอบภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบน้อยกว่าการไม่ให้ทวนคำตอบ

**ค่าความสามารถของผู้สอบ** หมายถึง ตัวเลขที่แสดงถึงความสามารถของผู้สอบแต่ละบุคคล โดยใช้วิธีของเบส์แบบปรับใหม่ ซึ่งวัดได้จากการทำการทดสอบในโปรแกรมคอมพิวเตอร์ทั้ง 2 ครั้ง ภายใต้เงื่อนไขการให้ทวนและไม่ให้ทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ หน่วยที่ใช้คือ  $\theta$  ซึ่งมีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 สำหรับค่าความสามารถมีค่าตั้งแต่  $-\infty$  ถึง  $+\infty$  แต่ผลการวิเคราะห์ส่วนใหญ่จะให้ค่าอยู่ในช่วง  $-3$  ถึง  $+3$

**ความแตกต่างของค่าความสามารถของผู้สอบ** หมายถึง ผลต่างของค่าความสามารถของผู้สอบแต่ละบุคคลภายใต้เงื่อนไขการให้ทวนและไม่ให้ทวนคำตอบ ( $\theta_{\text{ให้ทวน}} - \theta_{\text{ไม่ให้ทวน}}$ ) ซึ่งถ้าความแตกต่างออกมาเป็นค่าบวก จะแสดงถึง ค่าความสามารถของผู้สอบภายใต้เงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบ แต่ถ้าความแตกต่างออกมาเป็นค่าลบ จะแสดงถึง ค่าความสามารถของผู้สอบภายใต้เงื่อนไขการให้ทวนคำตอบน้อยกว่าการไม่ให้ทวนคำตอบ

**ระดับความสามารถของผู้สอบ** หมายถึง ค่าความสามารถของผู้สอบแต่ละบุคคล โดยใช้วิธีของเบส์แบบปรับใหม่ ซึ่งวัดได้จากการทำการทดสอบในโปรแกรมคอมพิวเตอร์ ภายใต้เงื่อนไขการไม่ให้ทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ในการศึกษาครั้งนี้ผู้วิจัยแบ่งระดับความสามารถออกเป็น 3 ระดับ คือ ความสามารถต่ำ ( $\theta \leq -0.7$ ) ซึ่งคิดเป็นพื้นที่ใต้โค้งปกติประมาณ 0.25 หรือประมาณ 25% ของกลุ่มตัวอย่าง, ความสามารถปานกลาง ( $-0.7 < \theta < 0.7$ ) ซึ่งคิดเป็นพื้นที่ใต้โค้งปกติประมาณ

0.5 หรือประมาณ 50% ของกลุ่มตัวอย่าง และความสามารถสูง ( $\theta \geq 0.7$ ) ซึ่งคิดเป็นพื้นที่ใต้โค้งปกติประมาณ 0.25 หรือประมาณ 25% ของกลุ่มตัวอย่าง

### ประโยชน์ที่คาดว่าจะได้รับ

1) ได้โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ วิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 สำหรับ ที่สามารถกำหนดเงื่อนไขการทดสอบได้ 2 เงื่อนไข คือ การทดสอบที่ให้ทวนคำตอบได้โดยวิธีการทวนที่มีการจำกัดช่วง (Block Review) กับการทดสอบที่ไม่ให้ทวนคำตอบ ซึ่งสามารถนำไปใช้ทดสอบนักเรียนได้ในทุกโรงเรียน

2) เพื่อให้ นักวัดผลทางการศึกษา ครู และอาจารย์ ได้ข้อมูลสำหรับเป็นแนวทางเพื่อเลือกใช้เงื่อนไขการทดสอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เพื่อให้การทดสอบนั้นมีประสิทธิภาพสูงสุด นั่นคือ ควรหรือไม่ควรเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้โดยวิธีการทวนที่มีการจำกัดช่วง(Block Review) เนื่องจากผลของการวิจัยในครั้งนี้จะได้ข้อค้นพบว่าการทวนคำตอบจะส่งผลให้ค่าความสามารถของผู้สอบมากกว่าการไม่ให้ทวนในผู้สอบทุกระดับความสามารถหรือไม่ รวมทั้งลักษณะของการเปลี่ยนคำตอบของผู้สอบ มีการส่งผลในด้านบวกหรือด้านลบสำหรับเงื่อนไขการทดสอบที่ให้ทวนคำตอบ

3) ได้แนวทางในการพัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ที่สามารถกำหนดเงื่อนไขให้ผู้สอบสามารถทวนคำตอบได้โดยวิธีการทวนที่มีการจำกัดช่วง (Block Review) เพื่อเผยแพร่องค์ความรู้เกี่ยวกับการเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ให้แพร่หลายออกไปมากกว่าเดิม เนื่องจากในปัจจุบันการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ได้ถูกนำมาใช้ในหลายๆการทดสอบ ดังนั้นผู้ที่มีหน้าที่เกี่ยวข้องกับการจัดการเรียนการสอน โดยเฉพาะผู้ที่มีหน้าที่โดยตรงเกี่ยวกับการวัดผลทางการศึกษาจึงควรที่จะมีความรู้ความเข้าใจในการทดสอบลักษณะนี้อย่างชัดเจน

จุฬาลงกรณ์มหาวิทยาลัย

## บทที่ 2

### เอกสารและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาเอกสาร และงานวิจัยต่างๆ เกี่ยวกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์(Computerized Adaptive Testing – CAT) ซึ่งผู้วิจัยได้นำเสนอเนื้อหาออกเป็น 4 ตอน ดังนี้

ตอนที่ 1 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบ

ตอนที่ 3 ระเบียบวิธีวิจัยและการพัฒนาโปรแกรมคอมพิวเตอร์ทางการศึกษา

ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

โดยในแต่ละตอนมีรายละเอียด ดังต่อไปนี้

#### ตอนที่ 1 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบเป็นการใช้ข้อสอบจากคลังข้อสอบเชอร์บีญญาด้วยแบบสอบ แบบสอบแต่ละชุดมีการออกแบบให้เหมาะสมกับความสามารถของผู้สอบแต่ละคน ซึ่งเรียกว่า “Tailored Test” คำนี้ใช้เป็นครั้งแรก โดย W.W. Turnbull (1951) ต่อมา ไวส์ได้เสนอให้ใช้คำว่า “Adaptive Test” (Weiss, 1974 อ้างถึงใน ศิริชัย กาญจนวาสี, 2545) ซึ่งมีหลักการทั่วไปของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบคือ การให้ผู้สอบเริ่มทำข้อสอบที่มีระดับความยากปานกลางก่อน ถ้าผลการตอบถูกข้อถัดไปจะมีความยากเพิ่มขึ้น แต่หากตอบข้อสอบข้อนั้นผิดข้อต่อไปที่ได้รับก็จะง่ายลง การทดสอบจะดำเนินการเช่นนี้ไปจนถึงสิ้นสุดการทดสอบ (Green and Others, 1984) ซึ่งผู้วิจัยได้นำเสนอเนื้อหาในตอนที่ 1 ออกเป็น 8 ตอนย่อย ดังนี้

1.1 ประวัติของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

1.2 ความหมายที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

1.3 องค์ประกอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

1.4 การดำเนินการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

1.5 ประเภทของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

1.6 การให้คะแนนการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

1.7 ประโยชน์การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

1.8 ประเด็นปัญหาของการใช้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

### 1.1 ประวัติของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ในเริ่มแรกนั้น การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบได้ดำเนินการสอบโดยให้ผู้สอบเขียนตอบลงในกระดาษคำตอบ ซึ่งมีข้อจำกัดในเรื่องการคัดเลือกข้อสอบข้อถัดไปหลังจากที่ผู้สอบตอบข้อสอบข้อแรกไปแล้ว ประกอบกับวิทยาการทางคอมพิวเตอร์ได้มีการพัฒนาก้าวหน้าไปมาก ดังนั้นการดำเนินการสอบด้วยคอมพิวเตอร์จึงได้รับความนิยมมากขึ้นตามลำดับ เนื่องจากรักษาความลับของข้อสอบได้เป็นอย่างดี การบันทึกคำตอบและการแสดงผลการทดสอบสามารถทำได้ง่ายและรวดเร็วกว่าการใช้กระดาษคำตอบและการตรวจข้อสอบแบบเดิม การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบต้องการคลังข้อสอบขนาดใหญ่ ข้อสอบในคลังดังกล่าวต้องมีค่าพารามิเตอร์ประจำข้อสอบแต่ละข้อ ได้แก่ ค่าอำนาจจำแนก (a), ค่าความยาก (b), ค่าโอกาสการเดาคำตอบถูก (c) และค่าสารสนเทศของข้อสอบ (item information function) แสดงคุณภาพไว้เพื่อสะดวกต่อการเลือกใช้ข้อสอบดังกล่าว (Hambleton and Swaminathan, 1985)

### 1.2 ความหมายที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

แบบสอบปรับเปลี่ยน (Adaptive Test) หมายถึง แบบสอบที่สามารถปรับเปลี่ยนความยากของข้อสอบแต่ละข้อให้เหมาะสมกับระดับความสามารถของผู้สอบแต่ละบุคคล (Lord, 1971)

แบบสอบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Test) หมายถึง แบบสอบที่ประกอบด้วยชุดของข้อสอบแตกต่างกันที่จัดให้สำหรับผู้สอบต่างกัน โดยใช้การคัดเลือกข้อสอบสำหรับผู้สอบแต่ละคนตามระดับความสามารถที่มุ่งวัดของบุคคลนั้น (Weiss, 1988)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing) หมายถึง การทดสอบที่ใช้แบบสอบต่างชุดกันสำหรับผู้สอบต่างกัน โดยมีการคัดเลือกข้อสอบที่มีความยากเหมาะสมกับระดับความสามารถของผู้สอบ (ศิริชัย กาญจนวาสี, 2545)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerize Adaptive Testing) หมายถึง การนำคอมพิวเตอร์มาใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

### 1.3 องค์ประกอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบเป็นวิธีการทดสอบที่มีแบบแตกต่างกันไปจากวิธีการทดสอบแบบประเพณีนิยม เนื่องจากการทดสอบวิธีนี้จะต้องกำหนดองค์ประกอบต่าง ๆ ที่ใช้ในการดำเนินการทดสอบไว้ล่วงหน้าทั้งสิ้น ซึ่งองค์ประกอบดังกล่าวนี้มีทั้งสิ้น 6 องค์ประกอบ (Hambleton, 1991) ได้แก่ โมเดลการตอบข้อสอบ, กลุ่มข้อสอบ, ระดับในการเริ่มต้น, การคัดเลือกข้อสอบ, การให้คะแนนหรือการประมาณค่าความสามารถของผู้สอบ และเกณฑ์การยุติการทดสอบ

#### 1.3.1 โมเดลการตอบข้อสอบ

ข้อสอบที่ใช้กับการทดสอบวิธีอาจเป็นข้อสอบที่ผ่านการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบหรือตามประเพณีนิยมมาแล้ว แต่หากเป็นไปได้ควรวิเคราะห์ข้อสอบตามแบบทฤษฎีการตอบสนองข้อสอบเพราะจะได้ค่าสถิติที่คงที่มากกว่า ทฤษฎีการตอบสนองตามแบบทฤษฎีต่อข้อสอบมีข้อตกลงเบื้องต้นว่าคุณลักษณะที่ทดสอบต้องเป็นมิติเดียวกัน ค่าพารามิเตอร์ของข้อสอบมีข้อตกลงเบื้องต้นที่ข้ามกลุ่ม และมีความเป็นอิสระในการตอบข้อสอบ (Lord and Novick, 1978) ดังนั้นการทดสอบวิธีนี้จึงสามารถประยุกต์ได้กับโมเดลการตอบข้อสอบที่มี 1, 2 และ 3 พารามิเตอร์ การเลือกใช้แบบใดควรพิจารณาธรรมชาติของข้อสอบและต้องเหมาะสมกับข้อมูลการตอบข้อสอบด้วย แต่การเลือกโมเดลการตอบข้อสอบที่จะใช้สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ไวส์และยอส (Weiss and Yoes) กล่าวว่า โมเดลโลจิสติกแบบ 3 พารามิเตอร์มีความเหมาะสมที่สุด ซึ่งสอดคล้องกับ Green, Bock, Linn และ Lord ซึ่งกล่าวว่า เพื่อให้ได้การประมาณค่าความสามารถของผู้สอบที่ถูกต้องมากที่สุดควรใช้โมเดลโลจิสติกแบบ 3 พารามิเตอร์

นอกจากนี้ จากการศึกษาของทอร์นไดค์ (Thorndike, 1983) เกี่ยวกับแบบสอบปรับเปลี่ยนชี้ให้เห็นว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบขึ้นอยู่กับจำนวนข้อและค่าอำนาจจำแนก กล่าวคือ ข้อสอบที่มีค่าอำนาจจำแนกสูงชี้้นค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าก็จะลดลง โดยปกติแล้วถ้าข้อสอบที่มีค่าอำนาจจำแนกโดยเฉลี่ย 0.7 ควรใช้ข้อสอบประมาณ 10 ถึง 15 ข้อ (ข้อ)

#### 1.3.2 กลุ่มข้อสอบ

เนื่องจากการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) ข้อสอบควรมีจำนวนมากและกลุ่มข้อสอบ (item pool) จะต้องประกอบด้วยข้อสอบที่ประมาณค่า

ลักษณะเฉพาะของข้อสอบตามทฤษฎีการตอบสนองของข้อสอบ พารามิเตอร์ของข้อสอบทุกข้อต้องได้รับการจัดให้อยู่ในมาตรวัดเดียวกันโดยใช้กระบวนการที่เหมาะสม ข้อสอบในกลุ่มข้อสอบจะต้องครอบคลุมช่วงของระดับความสามารถของประชากรและวัดความสามารถมิติเดียวกัน เออร์รี่ (Urry, 1977 อ้างถึงใน สายชล ออบทม, 2531) กล่าวว่ากลุ่มข้อสอบควรจะประกอบด้วยข้อสอบอย่างน้อย 100 ข้อ ค่าอำนาจจำแนกของข้อสอบเกิน 0.8 ระดับความยากกระจายอยู่ในช่วง  $-3.00$  ถึง  $+3.00$  และค่าอำนาจจำแนกอาจลดต่ำลงให้ช่วงระหว่าง 0 ถึง  $+2.5$

### 1.3.3 ระดับในการเริ่มต้น

ระดับในการเริ่มต้น (entry level) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบผู้สอบแต่ละคนจะตอบข้อสอบต่างกัน ดังนั้นจึงสามารถที่จะให้ผู้สอบเริ่มต้นทำข้อสอบที่มีระดับความยากที่ต่างกันได้ การเริ่มทำข้อสอบที่คลาดเคลื่อนไปจากความสามารถที่เหมาะสมไม่มีผลกระทบต่อการประมาณความสามารถเท่าใดนัก แต่ถ้าให้ผู้สอบเริ่มทำข้อสอบที่ตรงกับความสามารถของตนก็จะทำให้ลดจำนวนข้อที่จะใช้ในการทดสอบลงได้ ไวเนอร์ (Wainer, 1990) และแฮมเบิลตัน (Hambleton, 1991) กล่าวว่า ข้อสอบข้อแรกควรมีค่าความยากปานกลาง

### 1.3.4 การคัดเลือกข้อสอบ

การคัดเลือกข้อสอบ (item selection rule) จะเลือกโดยวิธีใช้ค่าสารสนเทศของข้อสอบที่สูงที่สุด ณ ระดับความสามารถของผู้สอบ (Maximum item information) โดยคำนวณค่าสารสนเทศข้อสอบข้อที่สูงที่สุดจากสูตร

$$I(\theta) = \frac{1 - c_i}{\left(c_i + e^{DL_i}\right)\left(1 - e^{-DL_i}\right)^2} \quad \text{เมื่อ } L_i = a_i(\theta - b_i) \\ D = 1.7$$

จากการศึกษาพบว่า วิธีใช้ค่าสารสนเทศของข้อสอบที่สูงที่สุดจะให้ค่าความตรงเชิงสภาพสูงที่สุด (รังสรรค์ มณีเล็ก, 2540) การเลือกข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบนอกจากจะพิจารณาข้อที่ให้สารสนเทศสูงสุดแล้ว ผลการศึกษาเชิงประจักษ์ของลอร์ด (Lord, 1971) ซึ่งชี้ให้เห็นว่าคะแนนจากแบบสอบจะสะท้อนให้เห็นความสามารถของแต่ละบุคคลได้แม่นยำเมื่อความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของแต่ละบุคคลเป็น 0.50 ดังนั้นแบบสอบจะมีความแม่นยำของการวัดต่ำเมื่อความน่าจะเป็นของการตอบข้อสอบถูกของแต่ละบุคคลมีค่ามากหรือน้อยกว่า 0.50 และจะมีความคลาดเคลื่อนมาตรฐานในการวัดสูง (Weiss, 1974)

### 1.3.5 การให้คะแนนหรือการประมาณค่าความสามารถของผู้สอบ

เพื่อให้ได้คะแนนหรือการประมาณค่าความสามารถของผู้สอบที่ถูกต้องมากที่สุด โดยยึดตามทฤษฎีการตอบสนองข้อสอบ (IRT) กลวิธีที่เหมาะสมมี 2 วิธีคือ

- 1) การประมาณค่าโดยวิธีเบย์ส (Bayesian updating) (Lord, 1980)
- 2) การประมาณค่าความเป็นไปได้สูงสุด (maximum likelihood) (Weiss, 1982)

จากการศึกษาพบว่าการประมาณค่าโดยวิธีประมาณค่าความเป็นไปได้สูงสุดและการประมาณค่าโดยวิธีเบย์สไม่แตกต่างกันเมื่อผู้สอบมีมากกว่า 2,000 คน แต่การประมาณค่าโดยวิธีเบย์สจะให้ผลดีกว่าเมื่อผู้สอบน้อยกว่า 500 คน

### 1.3.6 เกณฑ์การยุติการทดสอบ

เกณฑ์การยุติการทดสอบ (terminal criterion) ขึ้นอยู่กับการทดสอบของแต่ละบุคคลและแบบของการทดสอบปรับเปลี่ยนที่เลือกใช้ แต่เกณฑ์การยุติการทดสอบที่นิยมใช้มี 2 วิธีคือ

- 1) การกำหนดจำนวนข้อสอบ เป็นการกำหนดจำนวนข้อสอบให้แก่ผู้สอบทุกคน เช่น ถ้ากำหนดเกณฑ์ที่ 25 ข้อ ก็คือถ้าผู้สอบทำข้อสอบครบ 25 ข้อ ก็จะยุติการทดสอบนั้น ซึ่งการกำหนดเกณฑ์นี้มีข้อดีในการเปรียบเทียบสารสนเทศของแบบสอบได้โดยตรง แต่มีข้อจำกัดคือ อาจทำให้การวัดผลครั้งนั้นมีความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบสูง

- 2) การกำหนดค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ (Hambleton, 1991) จากการศึกษโดยใช้ Partial Credit Model พบว่า การยุติการทดสอบโดยดูจากค่าความคลาดเคลื่อนในการวัด จะให้ผลดีกว่าการยุติการสอบโดยการกำหนดจำนวนข้อสอบ (Dodd, 1993) และจากการศึกษาพบว่าการกำหนดค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.3 จะให้ค่าความตรงร่วมสมัยสูงสุด (รังสรรค์ มณีเล็ก, 2540)

## 1.4 การดำเนินการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ผู้สอบจะสอบทีละข้อ ผลการตอบข้อสอบแต่ละข้อจะถูกประมาณความสามารถเพื่อคัดเลือกข้อสอบให้เหมาะสมกับระดับความสามารถและมีเนื้อหาในบริเขตเดียวกันกับข้อเดิมเพื่อใช้เป็นกาทดสอบที่จะตอบต่อไป กระบวนการจะดำเนินเช่นนี้จนสิ้นสุดการทดสอบ วิธีการนี้เหมาะที่จะใช้กับเครื่องคอมพิวเตอร์ (Urry, 1977; Weiss and Kingsbury, 1984)

การดำเนินการทดสอบเริ่มต้นโดยการให้ผู้สอบนั่งหน้าจอคอมพิวเตอร์ที่มีหน่วยขับเคลื่อนบันทึกที่สามารถนำแผ่นบันทึกใส่หรือถอดออกมาได้ ในแผ่นบันทึกข้อมูลจะมีโปรแกรมการทดสอบที่สร้างขึ้นมาตามองค์ประกอบต่างๆของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบดังกล่าวข้างต้น ผู้ดำเนินการทดสอบจัดเตรียมเครื่องให้อยู่ในสภาพที่พร้อมจะทดสอบเครื่องคอมพิวเตอร์จะให้คำแนะนำในการตอบและรับข้อมูลส่วนตัวของนักเรียน เช่น ชื่อ—ชื่อสกุล เลขประจำตัว ชั้น/ห้อง เกรดเฉลี่ย เป็นต้น หลังจากนั้นเครื่องคอมพิวเตอร์จะแสดงข้อสอบข้อแรกบนจอภาพเพื่อให้ผู้สอบได้ตอบ ข้อสอบข้อแรกนี้เป็นข้อสอบที่มีระดับความยากปานกลาง ถ้าผู้สอบตอบข้อสอบข้อแรกถูกข้อสอบข้อถัดไปก็จะเป็นข้อสอบที่มีระดับความยากเพิ่มขึ้น แต่ถ้าตอบผิดข้อถัดไปจะเป็นข้อที่มีระดับความยากลดลง จากนั้นก็จะประมาณค่าความสามารถที่แท้จริงของผู้สอบ การทดสอบจะดำเนินต่อไปตามแนวทางต่อไปนี้

- 1) หลังจากประมาณค่าความสามารถของผู้สอบแล้วเครื่องคอมพิวเตอร์ก็จะคัดเลือกข้อสอบข้อใหม่ที่มีประโยชน์ในการกำหนดความสามารถจริงของผู้สอบ
- 2) แสดงข้อสอบข้อใหม่บนจอภาพ และผู้สอบตอบข้อสอบโดยการเลือกตัวเลือกที่ต้องการแล้วพิมพ์ตัวเลือกที่ต้องการ คอมพิวเตอร์จะรับคำตอบแล้วตรวจข้อสอบ
- 3) ประมาณค่าความสามารถของผู้สอบอีกครั้งหนึ่งตามผลของการตอบข้อสอบข้อผ่านมา
- 4) พิจารณาว่าผู้สอบควรจะทำข้อสอบข้อต่อไปหรือไม่ตามเกณฑ์การยุติการทดสอบที่กำหนดไว้ ถ้าผู้สอบจะต้องทำข้อสอบข้อต่อไปนี้ก็ย้อนกลับไปขั้นที่ 1 อีกครั้ง
- 5) เมื่อสิ้นสุดการทดสอบ สารสนเทศเกี่ยวกับการตอบสนองข้อสอบของผู้สอบก็จะถูกบันทึกไว้ในเครื่องคอมพิวเตอร์และแสดงผลการทดสอบให้ผู้สอบทราบ

### 1.5 ประเภทของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแบ่งออกเป็น 2 ประเภทใหญ่ ๆ (Weiss ,1984 ; Hambleton and Swaminathan,1985 อ้างถึงใน สุพัฒน์ สุขมลสันต์, 2535) และในแต่ละประเภทยังแบ่งย่อยได้อีกดังนี้

- 1) กลวิธีสองขั้นตอน (two-stage strategies)
- 2) กลวิธีหลายขั้นตอน (multi – stage strategies) ซึ่งแบ่งเป็น
  - 2.1) แบบแยกทางคงที่ (fixed branching mode) ได้แก่
    - 2.1.1) แบบพีระมิต (pyramidal model) ได้แก่
      - (1) แบบพีระมิตชนิดขั้นคงที่ (constant step size pyramidal model)
      - (2) แบบพีระมิตชนิดขนาดขั้นแปรผัน (variable step size pyramidal model)



- (3) แบบพีระมิดชนิดข้างตัด (truncated pyramidal model)
- (4) แบบพีระมิดชนิดหลายข้อในแต่ละชั้น (multiple – item pyramidal model)
- (5) แบบพีระมิดชนิดให้น้ำหนักตัวเลือกเพื่อแยกทาง (differential response option branching pyramidal mode)
  - 2.1.2) แบบระดับยืดหยุ่น (flexilevel model)
  - 2.1.3) แบบปรับระดับขั้น (stepadaptive model)
- 2.2) แบบแยกทางแปรผัน (variable branching mode) ได้แก่
  - 2.2.1) กลวิธีของเบส์ (Bayesian strategies)
  - 2.2.2) กลวิธีประมาณค่าความเป็นไปได้สูงสุด (maximum likelihood estimation strategies)

### 1) กลวิธีสองขั้นตอน (two-stage strategies)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้กลวิธีสองขั้นตอนจะใช้ข้อสอบ 2 ชุด ข้อสอบชุดแรกเป็นแบบสอบแยกทาง (routing test) ซึ่งผู้สอบทุกคนต้องสอบเหมือนกัน เพื่อแยกระดับความสามารถเพื่อที่จะไปตอบข้อสอบในชุดที่สอง โดยปกติข้อสอบชุดแรกจะมีประมาณ 10 ข้อ ส่วนข้อสอบชุดที่สองเป็นแบบสอบหลัก (main test) ซึ่งจะแบ่งเป็นข้อสอบชุดย่อย ๆ หลายชุดตามระดับความยาก โดยปกติจะมีชุดละประมาณ 20–30 ข้อ ผู้สอบแต่ละคนจะได้รับข้อสอบชุดที่สองไม่เหมือนกัน ผู้สอบที่ได้รับการประเมินจากข้อสอบชุดแรกว่ามีความสามารถสูงจะได้รับข้อสอบชุดที่สองที่ยาก ส่วนผู้สอบที่มีความสามารถต่ำจะได้รับข้อสอบชุดที่ง่าย (Lord, 1971; Weiss and Betz, 1973)

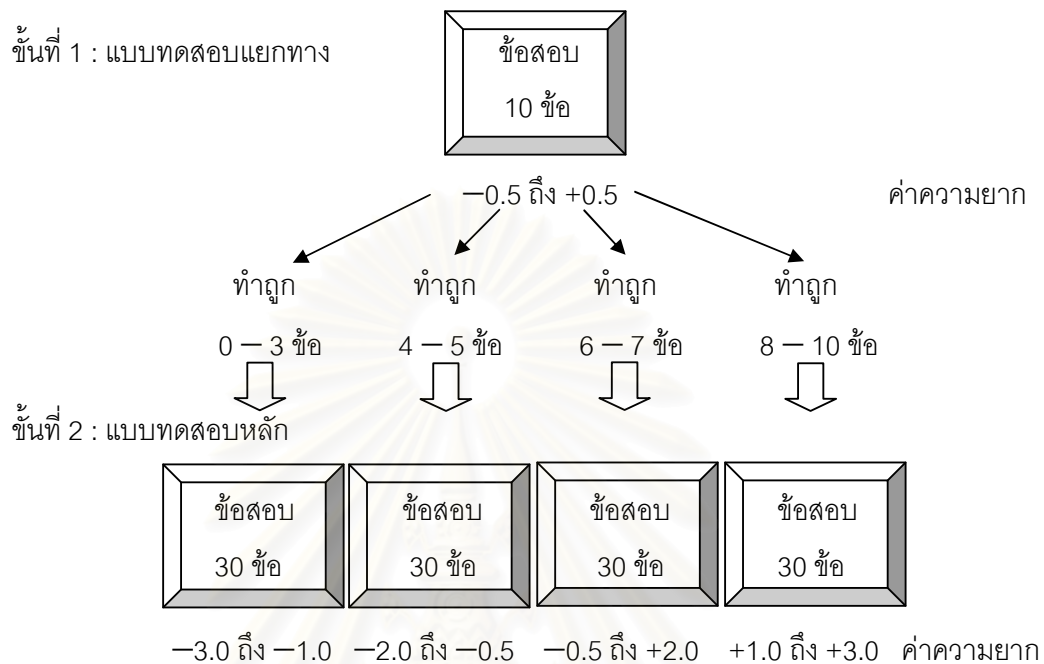
ลอร์ด (Lord, 1980) ได้แนะนำว่าข้อสอบชุดย่อย (subtest) แต่ละฉบับของแบบสอบหลักควรมีข้อสอบบางส่วนที่คาบเกี่ยวกัน (overlap) ในเรื่องของระดับความยากของข้อสอบรายชื่อ

- ถ้าให้  $m$  คือ จำนวนข้อสอบที่คาบเกี่ยวกัน
- $L$  คือ จำนวนฉบับชุดย่อยในขั้นที่สอง
- $n$  คือ จำนวนข้อสอบทั้งหมด
- $n_1$  คือ จำนวนข้อสอบในชุดย่อยแต่ละฉบับในขั้นที่สอง

ดังนั้น

$$m = \frac{n_1 L - n}{L - 1}$$

โครงสร้างของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้  
กลวิธีสองขั้นตอนดังกล่าวแล้ว อาจแสดงให้เห็นได้ด้วยแผนภาพที่ 1 ดังต่อไปนี้



แผนภาพที่ 1 โครงสร้างของการทดสอบแบบปรับเหมาะโดยใช้กลวิธีสองขั้นตอน

จากแผนภาพที่ 1 ผู้สอบทุกคนต้องทำข้อสอบในขั้นที่ 1 ซึ่งมีอยู่ 10 ข้อ ผู้สอบที่  
ทำข้อสอบถูก 0–3 ข้อ จะไปทำข้อสอบขั้นที่ 2 ที่ง่ายที่สุด ส่วนผู้สอบที่ทำข้อสอบขั้นที่ 1 ถูก 4–5  
ข้อ 6–7 หรือ 8–10 ข้อ จะแยกไปทำข้อสอบในขั้นที่ 2 ที่ยากขึ้นตามลำดับ

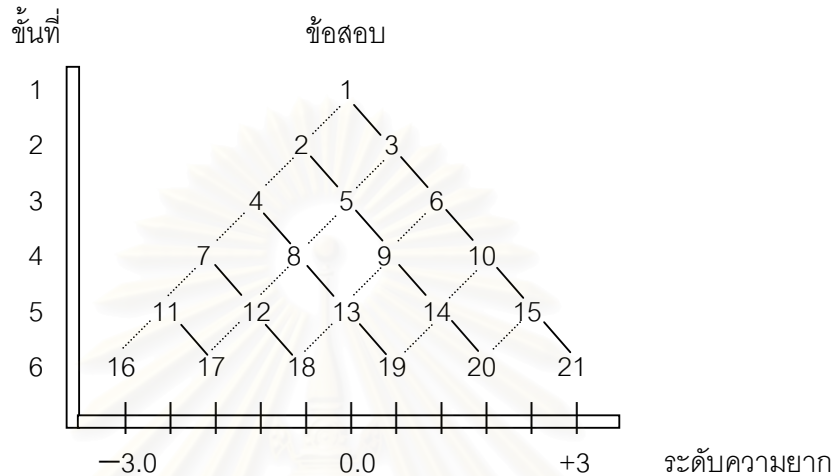
## 2) กลวิธีหลายขั้นตอน (multi – stage strategies)

### 2.1) แบบแยกทางคงที่

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้กลวิธีหลาย  
ขั้นตอนแบบแยกทางคงที่ หมายถึงการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบที่มี  
หลายขั้นตอน ในแต่ละขั้นตอนอาจจะมีข้อสอบเพียงข้อเดียวหรือหลายข้อก็ได้ โดยมีแบบหรือ  
เส้นทางในการตอบข้อสอบที่แน่นอน ซึ่งแบ่งเป็นแบบย่อย ๆ คือ

2.1.1) **แบบพีระมิด** การทดสอบปรับเปลี่ยนแบบพีระมิดจำแนกเป็นหลาย  
แบบ เช่น

(1) แบบพหุมิติชนิดขั้นคองที่ ได้แก่ การทดสอบปรับเหมาะที่จัดเรียงข้อสอบ เป็นรูปสามเหลี่ยมโดยมีจำนวนข้อสอบในแต่ละข้อเท่ากับลำดับที่ของชั้น กล่าวคือ ชั้นที่ 1 จะมีข้อสอบหนึ่งข้อ ชั้นที่ 2 จะมีข้อสอบสองข้อ ในชั้นที่ 5 จะมีข้อสอบห้าข้อ ดังนั้นการทดสอบปรับเปลี่ยนพหุมิติที่มี 6 ชั้น จะมีข้อสอบทั้งหมด 21 ข้อ ดังแผนภาพที่ 2

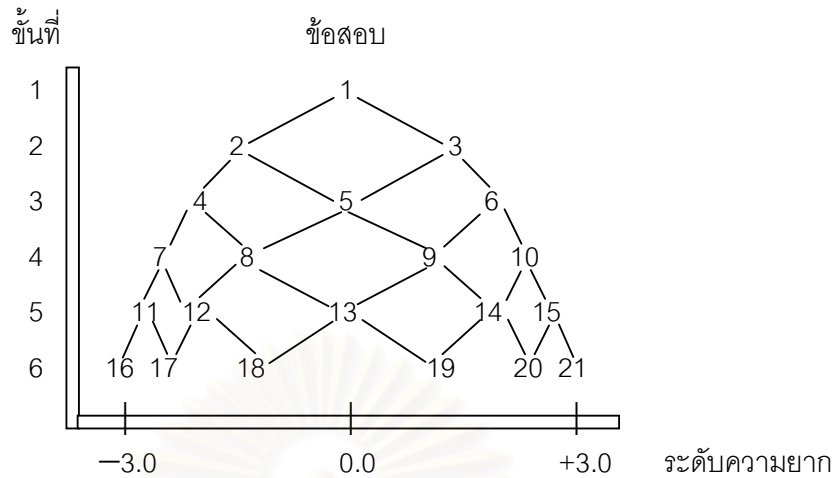


แผนภาพที่ 2 โครงสร้างของการทดสอบแบบปรับเหมาะแบบพหุมิติชนิดขนาดขั้นคองที่

จากแผนภาพที่ 2 แขนงอนเป็นระดับความยากของข้อสอบ (b) ที่สัมพันธ์กับข้อสอบในการทดสอบปรับ เปลี่ยนแบบพหุมิติ ซึ่งจะเห็นว่าระดับความยากมีค่าอยู่ระหว่าง  $-3.0$  ถึง  $+3.0$  โดยแบ่งออกเป็น 10 ช่วงเท่าๆ กัน ฉะนั้นระดับความยากจะห่างกันช่วงละ  $0.6$  ข้อสอบที่อยู่ติดกันภายในชั้นมีค่าเท่ากันตลอด คือ  $1.2$

ในการตอบแบบสอบ ผู้สอบจะต้องตอบข้อสอบชั้นละหนึ่งข้อ โดยเริ่มตอบข้อที่อยู่บนยอดสามเหลี่ยมซึ่งมีความยากปานกลางก่อน ถ้าตอบถูกขั้นต่อไปจะแยกไปตอบข้อที่ยากขึ้น แต่ถ้าตอบผิดขั้นต่อไปจะไปตอบข้อที่ง่ายกว่า กระบวนการทดสอบจะเป็นเช่นนี้จนถึงขั้นสุดท้าย

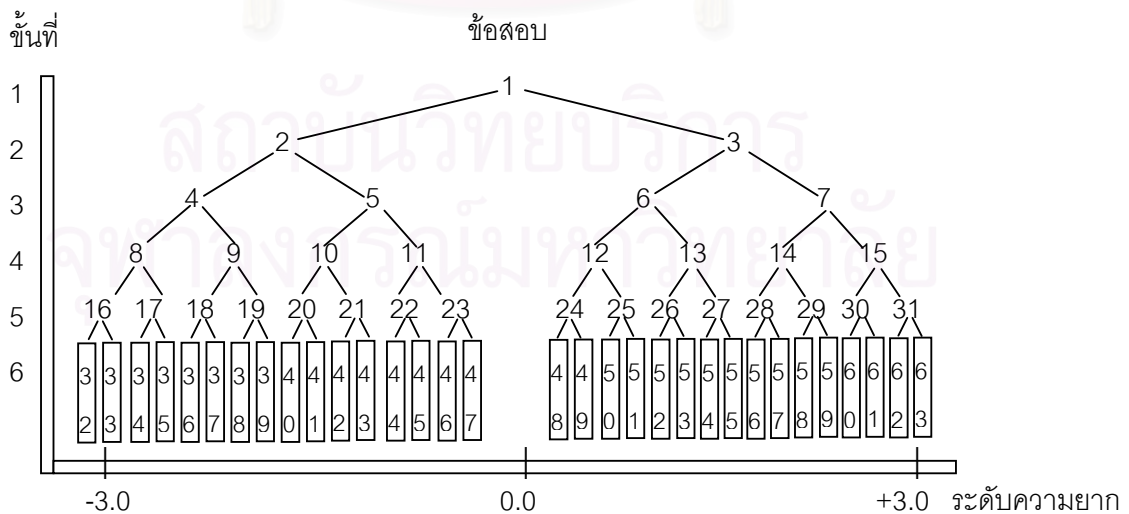
(2) แบบพหุมิติชนิดขั้นแปรผัน เนื่องจากลอร์ด (Lord, 1971) มีความเห็นว่าแบบพหุมิติชนิดขนาดขั้นคองที่ขาดประสิทธิภาพเกี่ยวกับความไวในการกำหนดเส้นทางการตอบข้อทดสอบ จึงได้กำหนดช่วงห่างระดับความยากของข้อที่อยู่ติดกันภายในชั้นให้มีค่าไม่เท่ากัน โดยให้ข้อที่อยู่กลาง ๆ มีช่วงห่างของระดับความยากมาก แล้วค่อย ๆ ลดลงในข้อที่ค่อนข้างง่ายหรือยาก การเพิ่มหรือลดนั้นจะเพิ่มหรือลดไปยังตำแหน่งกึ่งกลางระหว่าง 2 ข้อในชั้นที่มาก่อน หรือระหว่างข้อในชั้นที่มาก่อนกับระดับความยากสูงสุดหรือต่ำสุด ดังตัวอย่างแสดงไว้ในแผนภาพที่ 3



แผนภาพที่ 3 โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดชนิดขนาดขั้นแปรผัน

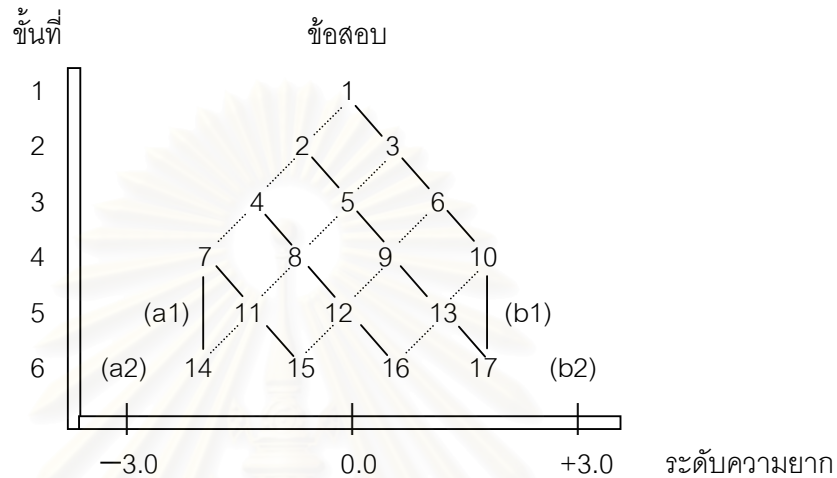
จากแผนภาพที่ 3 ในขั้นที่ 1 มีข้อสอบเพียงข้อเดียวคือ ข้อที่ 1 ซึ่งมีระดับความยากเท่ากับ 0.00 ในขั้นที่ 2 มีข้อสอบ 2 ข้อ คือข้อที่ 2 มีระดับความยากอยู่กึ่งกลางระหว่าง -3.00 (ระดับความยากต่ำสุด) กับ 0.00 (ระดับความยากของข้อที่ 1) ซึ่งก็คือ -1.50 ข้อที่ 3 มีระดับความยากอยู่กึ่งกลางระหว่าง +3.00 (ระดับความยากสูงสุด) กับ 0.00 (ระดับความยากของข้อที่ 1) ซึ่งก็คือ 1.50 ในขั้นที่ 3 มีข้อสอบสามข้อคือข้อที่ 4 มีระดับความยากอยู่กึ่งกลางระหว่าง -3.00 (ระดับความยากเท่ากับต่ำสุด) กับ -1.50 (ระดับความยากของข้อที่ 2) ซึ่งก็คือ -2.25 ข้อที่ 5 มีระดับความยากเท่ากับ 0.00 ข้อที่ 6 มีระดับความยากอยู่กึ่งกลางระหว่าง 1.50 (ระดับความยากของข้อที่ 3) กับ 3.00 (ระดับความยากสูงสุด) ซึ่งก็คือ 2.25 สำหรับในข้ออื่นๆ ก็ในทำนองเดียวกัน

ต่อมาลอร์ด (Lord, 1971) ได้ใช้วิธีการเพิ่มจำนวนข้อในแต่ละชั้นเป็นสองเท่าของจำนวนข้อสอบในขั้นก่อน และให้ชื่อว่าแบบรอบิน-มอนโร (Robin-Monro model) ดังแผนภาพที่ 4



แผนภาพที่ 4 โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดชนิดรอกบิน-มอนโร

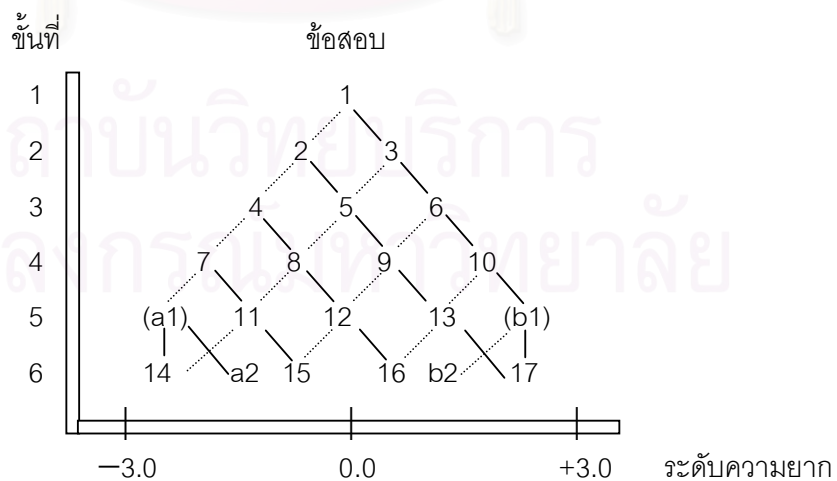
(3) แบบพีระมิดชนิดข้างตัด เป็นแบบที่มัสซิโอ (Mussio, 1973 อ้างถึงใน Weiss, 1974) เป็นผู้เสนอแบบนี้ขึ้นมา เพื่อให้ลดจำนวนข้อสอบลงจากแบบพีระมิดชนิดขนาดชั้นคงที่ด้วยวิธีสกัดการสะท้อนกลับและวิธีการรักษาการสะท้อนกลับ (reflecting and retaining barriers) ดังตัวอย่างแสดงไว้ในแผนภาพที่ 5 และแผนภาพที่ 6



แผนภาพที่ 5 โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดข้างตัด

วิธีสกัดการสะท้อนกลับ

จากแผนภาพที่ 5 การตอบข้อสอบในชั้นที่ 1 ถึงชั้นที่ 3 มีวิธีตอบเช่นเดียวกับแบบพีระมิดชนิดขนาดชั้นคงที่ จนถึงจุดสกัดในชั้นที่ 4 ถ้าตอบข้อสอบที่ 7 ถูก ก็จะแยกไปตอบข้อที่ 11 แต่ถ้าตอบข้อที่ 7 ผิด ก็จะไปตอบข้อที่ 14 และในทำนองเดียวกัน ถ้าตอบข้อที่ 10 ผิดก็จะไปตอบข้อที่ 17

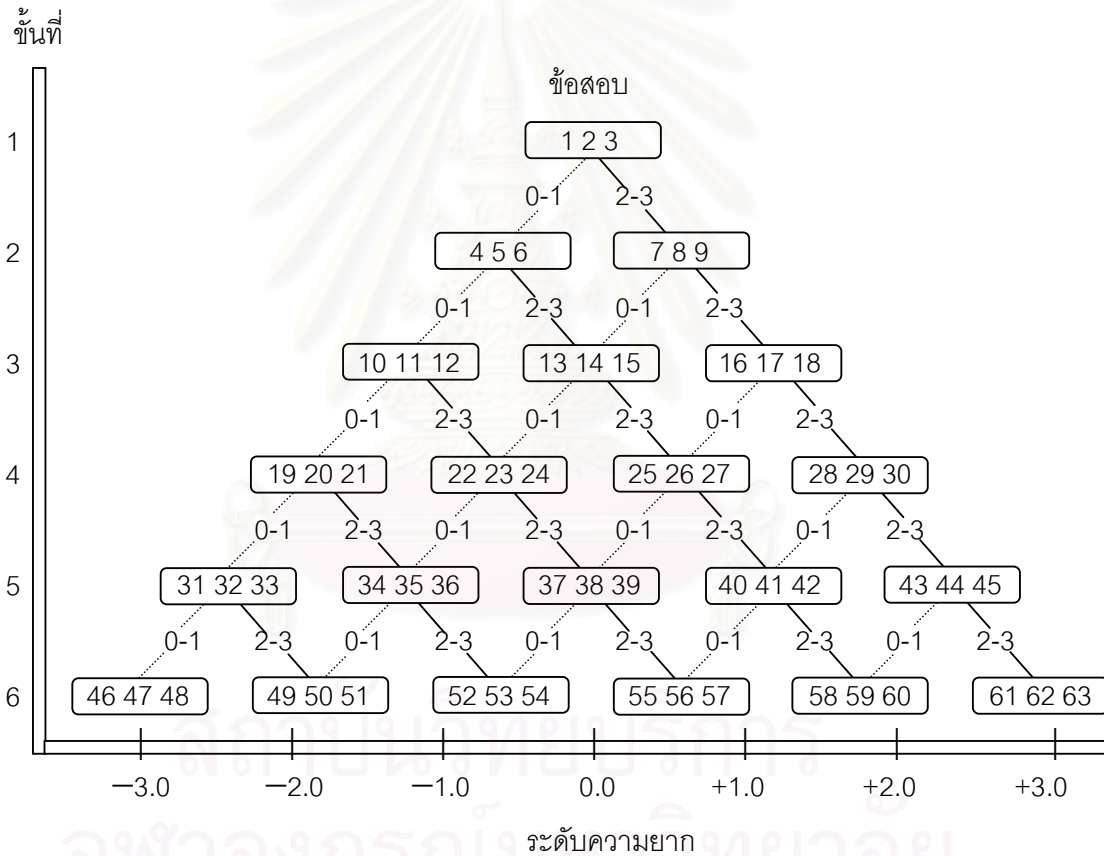


แผนภาพที่ 6 โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิดข้างตัด

วิธีรักษาการสะท้อนกลับ

จากแผนภาพที่ 6 การตอบข้อสอบในชั้นที่ 1 ถึงชั้นที่ 3 มีวิธีตอบเช่นเดียวกับแบบพีระมิตชนิดขนาดชั้นคงที่ จนถึงจุดสกัดในชั้นที่ 4 ถ้าตอบข้อสอบข้อที่ 7 ถูก ก็จะแยกไปตอบข้อที่ 11 แต่ถ้าตอบข้อที่ 7 ผิด ก็จะไปตอบข้อที่ a1 ซึ่งมีระดับความยากเท่ากับข้อที่ 7 และถ้าตอบข้อที่ a1 ผิดอีกก็จะไปตอบข้อที่ 14 แต่ถ้าตอบข้อที่ a1 ถูก ก็จะแยกไปตอบข้อที่ a2 ซึ่งมีระดับความยากเท่ากับข้อที่ 11

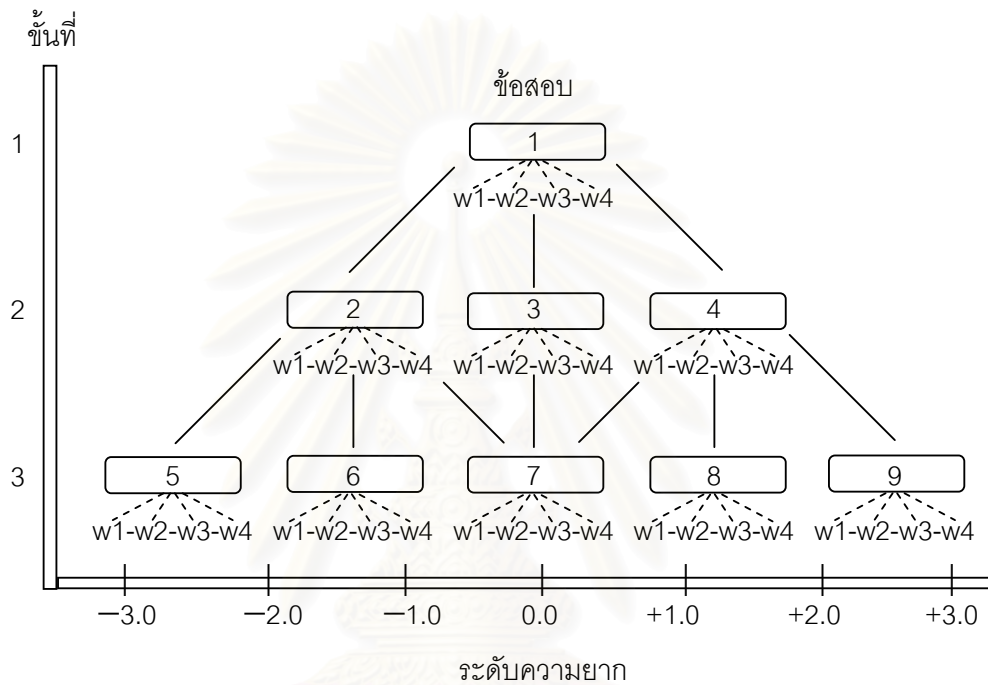
(4) แบบพีระมิตชนิดหลายข้อในแต่ละชั้น เป็นแบบที่แครทเวิร์ท, ฮอยเซอร์ และลินน์ (Krathworth and Huyser 1956 ; Linn 1969 , cited in Weiss 1974 : 105) เป็นผู้เสนอ โดยให้มีจำนวนข้อในแต่ละชั้นมากกว่าหนึ่งข้อ ทั้งนี้ เพื่อให้การตอบเพื่อแยกทางได้ถูกต้องมากยิ่งขึ้น ดังตัวอย่างแผนภาพที่ 7



แผนภาพที่ 7 โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิตชนิดที่หลายข้อในแต่ละชั้น

จากแผนภาพที่ 7 ในชั้นที่ 1 ถ้าผู้สอบตอบข้อสอบถูก 0 ถึง 1 ข้อ จะแยกไปตอบข้อสอบข้อที่ 4, 5 และ 6 ถ้าตอบข้อสอบถูก 2 ถึง 3 ข้อ ก็จะแยกไปตอบข้อสอบข้อที่ 7, 8 และ 9 สำหรับในชั้นอื่น ๆ ก็ในทำนองเดียวกัน

(5) แบบพีระมิตชนิดให้น้ำหนักตัวเลือกเพื่อแยกทาง เบย์รอฟฟ์และซีเลย์ (Bayroff and Seeley, 1968; อ้างอิงโดย Weiss, 1974) เป็นผู้เสนอ มีแบบคล้ายกับแบบที่ใช้ข้อสอบหลาย ๆ ข้อในแต่ละชั้น กล่าวคือแทนที่จะใช้ข้อสอบหลาย ๆ ข้อในแต่ละชั้นก็ใช้ตัวเลือกของข้อสอบแทน รูปแบบนี้เหมาะสำหรับข้อสอบเลือกตอบที่สามารถกำหนดน้ำหนักคะแนนของตัวเลือกได้ ดังแผนภาพที่ 8



แผนภาพที่ 8 โครงสร้างของการทดสอบแบบปรับเหมาะแบบพีระมิตชนิดให้น้ำหนักตัวเลือกเพื่อแยกทาง

### 2.1.2) แบบระดับยี่ดหุ่่น

การทดสอบแบบนี้เป็นแบบที่ลอร์ดคิดขึ้นมาในครั้งแรกเพื่อใช้กับการทดสอบที่ใช้กระดาษกับดินสอ (paper and pencil test) โดยแบ่งข้อสอบออกเป็น 2 กลุ่ม คือข้อสอบกลุ่มง่ายกับข้อสอบกลุ่มยาก และคัดเลือกข้อที่มีระดับความยากอยู่กึ่งกลางระหว่างข้อสอบกลุ่มง่ายกับข้อสอบกลุ่มยากไว้ยากต่างหากอีกหนึ่งข้อเพื่อเป็นข้อเริ่มต้นสำหรับผู้สอบทุกคน ภายในข้อสอบกลุ่มง่ายจะเรียงลำดับข้อสอบจากข้อที่ยากที่สุดภายในกลุ่มนี้ไปหาข้อที่ง่าย ข้อสอบกลุ่มง่ายนี้พิมพ์ด้วยหมึกสีแดง ส่วนข้อสอบในกลุ่มยากจะเรียงลำดับจากข้อที่ง่ายที่สุดภายในกลุ่มนี้ไปหาข้อที่ยากและพิมพ์ด้วยหมึกสีน้ำเงิน ดังตัวอย่างในแผนภาพที่ 9

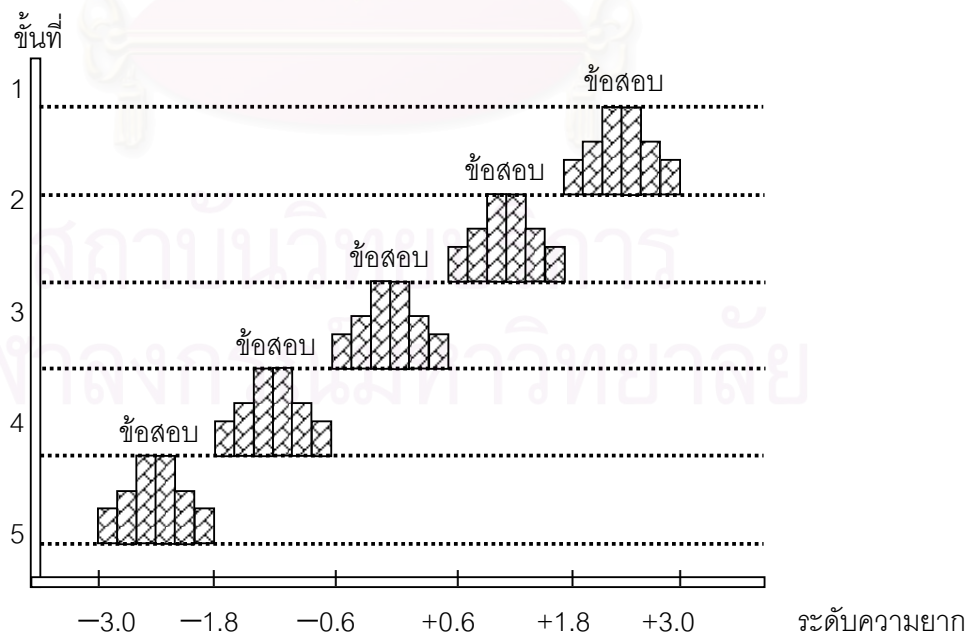
ข้อ 0 (มีระดับความยากอยู่กึ่งกลางระหว่างข้อสอบกลุ่มง่ายกับกลุ่มยาก)	
(ข้อสอบกลุ่มง่าย) (พิมพ์สีแดง)	(ข้อสอบกลุ่มยาก) (พิมพ์สีน้ำเงิน)
ข้อ 1. (ข้อที่ยากที่สุดในกลุ่มนี้)	ข้อ 1. (ข้อที่ง่ายที่สุดในกลุ่มนี้)
ข้อ 2.	ข้อ 2.
ข้อ 3.	ข้อ 3.
.	.
ข้อ 37. (ข้อที่ง่ายที่สุดในกลุ่มนี้)	ข้อ 37. (ข้อที่ยากที่สุดในกลุ่มนี้)

แผนภาพที่ 9 โครงสร้างของการทดสอบแบบปรับเหมาะแบบระดับยี่ดหุน

สำหรับกระดาษคำตอบมีการออกแบบเป็นพิเศษ กล่าวคือ เมื่อนักเรียนตอบถูก จุดสีน้ำเงินก็จะปรากฏขึ้นแล้วผู้สอบจะต้องไปเลือกตอบข้อสอบในกลุ่มสีน้ำเงิน (กลุ่มยาก) ในข้อที่มีหมายเลขต่ำสุดที่ยังไม่ได้ทำ แต่ถ้าตอบผิดจุดสีแดงก็จะปรากฏขึ้นแล้วผู้สอบจะต้องไปเลือกตอบข้อสอบในกลุ่มสีแดง (กลุ่มง่าย) ในข้อที่มีหมายเลขต่ำสุดที่ยังไม่ได้ทำ ถ้าให้  $k$  เป็นจำนวนข้อสอบทั้งหมด ผู้สอบจะต้องตอบข้อสอบเท่ากับ  $\frac{k+1}{2}$  ข้อ (Lord 1971 : 120)

### 2.1.3) แบบปรับระดับขั้น

มีการปรับปรุงมาจากการทดสอบแบบระดับยี่ดหุน โดยการแบ่งระดับความยากของข้อสอบออกเป็นช่วงๆ ในแต่ละช่วงระดับความยากจะมีข้อสอบหลายข้อ (Water and Bayroff, 1971)



แผนภาพที่ 10 โครงสร้างของการทดสอบแบบปรับเหมาะแบบปรับระดับขั้น



## 2.2) แบบแยกทางแปรผัน

การทดสอบปรับเปลี่ยนโดยใช้กลวิธีหลายขั้นตอนตามแบบแยกทางแปรผัน หมายถึง รูปแบบการตอบข้อสอบหลายขั้นตอนที่ไม่ได้กำหนดข้อสอบและเส้นทางในการตอบข้อสอบไว้ล่วงหน้าว่าถ้าผู้สอบตอบถูกต้องจะต้องไปทำข้อสอบข้อใดหรือ ถ้าตอบผิดจะต้องไปทำข้อสอบข้อใดต่อไป ในแบบแยกทางแปรผันนี้จะไม่มีขนาดของชั้น แต่จะดำเนินการทดสอบจากกลุ่มข้อสอบที่คำนวณระดับความยากและค่าอำนาจจำแนกของข้อสอบไว้แล้ว กฎการคัดเลือกข้อสอบข้อถัดไปเพื่อให้เหมาะสมกับความสามารถของผู้สอบนั้นจะใช้แบบทางคณิตศาสตร์เข้ามาช่วย ได้แก่ การใช้กลวิธีของเบส์ และกลวิธีความเป็นไปได้สูงสุด ซึ่งมีรายละเอียดดังนี้

**2.2.1) กลวิธีของเบส์** กลวิธีของเบส์นี้จะประยุกต์ทฤษฎีของเบส์ในกระบวนการตอบข้อสอบและอาศัยเครื่องคอมพิวเตอร์ในการทดสอบ จึงมักเรียกว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งมีลำดับขั้นตอนดังนี้

1) ในแต่ละชั้นของการทดสอบจะมีการประมาณค่าความสามารถของผู้สอบและความคลาดเคลื่อนมาตรฐานการประมาณค่าไว้จากข้อมูลต่างๆ ที่มีอยู่เกี่ยวกับผู้สอบและข้อสอบ

2) คัดเลือกข้อสอบจากกลุ่มข้อสอบที่ได้คำนวณค่าพารามิเตอร์ของข้อสอบไว้แล้ว ข้อสอบในกลุ่มข้อสอบทุกข้อที่ยังไม่ได้นำมาให้ผู้สอบคนนั้นๆ สอบจะถือว่าเป็นข้อสอบที่มีโอกาสนำมาใช้สอบได้ กระบวนการนี้แสดงให้เห็นว่าข้อสอบข้อใดๆ ในกลุ่มข้อสอบที่นำมาใช้ในการทดสอบกับผู้สอบคนใดๆ ก็ตามจะเป็นข้อสอบที่มีระดับความยากใกล้เคียงกับระดับความสามารถของผู้สอบมากที่สุด หลังจากดำเนินการสอบโดยใช้ข้อสอบที่คัดเลือกไว้ก็จะประมาณความสามารถของผู้สอบไว้ก่อน แล้วการตอบถูกหรือการตอบผิดจะนำมารวมกันเพื่อคำนวณโดยใช้ทฤษฎีบทของเบส์หาค่าความสามารถในภายหลัง การประมาณค่าครั้งหลังนี้เป็น การประมาณค่าที่ปรับจากค่าที่ได้คำนวณไว้แต่เดิม กระบวนการนี้จะสิ้นสุดลงเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถมีค่าน้อยกว่าที่กำหนดไว้ล่วงหน้า

**2.2.2) กลวิธีประมาณค่าความเป็นไปได้สูงสุด** วิธีนี้เป็นวิธีที่เออร์รี่ (Urry, 1970) ได้พัฒนาขึ้นโดยใช้วิธีการประมาณค่าความเป็นไปได้สูงสุดตามทฤษฎีการตอบสนองข้อสอบ ซึ่งมีวิธีการประมาณค่าความสามารถของผู้สอบที่คล้ายคลึงกับกระบวนการตามทฤษฎีบทของเบส์ แม้ว่าเหตุผลทางคณิตศาสตร์ต่างกัน หลังจากที่ยุ้สอบตอบข้อสอบเพียง 1 ข้อ ก็จะนำผลการสอบไปแก้สมการความเป็นไปได้สูงสุด และจะได้ค่าประมาณความสามารถและความคลาดเคลื่อนมาตรฐาน ข้อสอบข้อถัดไปที่เลือกมาใช้ทดสอบจะเป็นข้อสอบจากกลุ่มข้อสอบและเป็นข้อสอบที่มีระดับความยากเหมาะสมกับความสามารถของผู้สอบมากที่สุด เมื่อผู้สอบตอบข้อสอบข้อนั้นแล้วก็จะมีการประเมินค่าความสามารถของผู้สอบทันทีจากข้อมูลการตอบทั้งหมดของผู้สอบ

ซึ่งรวมถึงข้อสอบข้อสุดท้ายที่ผู้สอบตอบ จากนั้นก็ประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐานใหม่โดยใช้สมการความเป็นไปได้สูงสุด จนกว่าการทดสอบจะสิ้นสุดตามเกณฑ์ที่กำหนดไว้

## 1.6 การให้คะแนนการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

การให้คะแนนการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบอาจทำได้หลายวิธี (Weiss, 1974) ดังนี้

1. ให้คะแนนตามระดับความยากของข้อที่ยากที่สุดที่ตอบถูก (most difficulty of all items answered correctly) เป็นการให้คะแนนโดยดูจากข้อสอบทั้งหมดที่ผู้สอบสามารถตอบได้ถูกต้อง หากข้อใดที่มีระดับความยากสูงสุดที่ผู้สอบตอบถูกถือว่าเป็นค่าความสามารถของผู้สอบคนนั้น ระดับความยากดังกล่าวนี้ ควรจะเป็นระดับความยากที่ได้จากการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ
2. ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก (average difficulty of all items answered correctly) เป็นการให้คะแนนโดยการนำระดับความยากของข้อสอบทุกข้อที่ผู้สอบตอบถูกมาหาค่าเฉลี่ย ค่าเฉลี่ยความยากนี้ถือเป็นค่าความสามารถของผู้สอบคนนั้น
3. ให้คะแนนตามค่าเฉลี่ยความยากของทุกข้อที่ทำ (average difficulty) วิธีการให้คะแนนแบบนี้ลอร์ดเป็นผู้เสนอ (Lord, 1970) วิธีการนี้คือนำระดับความยากของข้อสอบทุกข้อที่ผู้สอบได้ตอบโดยไม่คำนึงว่าจะตอบถูกหรือผิดมาหาค่าเฉลี่ย ค่าเฉลี่ยนี้ถือเป็นค่าความสามารถของผู้สอบคนนั้น
4. ให้คะแนนตามระดับความยากของข้อสุดท้ายที่ตอบ (difficulty of the final item) นั่นคือ ไม่ว่าจะขั้นสุดท้ายจะตอบถูกหรือไม่ก็ตาม จะถือว่าระดับความยากของข้อสอบข้อสุดท้ายที่ผู้สอบตอบคือค่าความสามารถของผู้สอบคนนั้น
5. ให้คะแนนตามระดับความยากของข้อสอบที่ต่อจากขั้นสุดท้ายที่ตอบ (difficulty of  $(N + 1)^{th}$  item) ซึ่งมีลักษณะคล้ายกับวิธีที่ 4 แต่เป็นการตรวจสอบความสามารถต่อไปอีกขั้นหนึ่ง กล่าวคือ ถ้าขั้นสุดท้ายผู้สอบสามารถตอบถูก คะแนนในขั้นต่อไปก็จะเพิ่มขึ้น แต่ถ้าขั้นสุดท้ายตอบผิดขั้นต่อไปคะแนนก็จะลดลง
6. ให้คะแนนตามกลวิธีของเบย์ ซึ่งมีหลายวิธี แต่วิธีที่นิยมใช้กันมากคือวิธีที่เรียกว่า “วิธีของเบย์ที่ปรับปรุงให้ทันสมัย” (Bayesian updating) ซึ่งสามารถประมาณค่าความสามารถของผู้สอบได้ค่อนข้างคงที่ โดยมีสูตรดังนี้

## 6.1 ในกรณีที่ตอบข้อสอบถูก

$$\theta_{m+1} = \theta_m + (1-c) \left( \frac{\sigma_m^2}{\sqrt{\frac{1}{a^2} + \sigma_m^2}} \right) \left( \frac{O(D)}{c + (1-c)A(-D)} \right)$$

$$\sigma_{m+1}^2 = \sigma_m^2 \left( 1 - \left( \frac{1-c}{1 + \frac{1}{a^2 \sigma_m^2}} \right) \left( \frac{O(D)}{B} \right) \left( \frac{(1-c)O(D)}{B} - D \right) \right)$$

เมื่อ  $D = \frac{b - \theta_m}{\sqrt{\frac{1}{a^2} + \sigma_m^2}}$  และ  $B = c + (1-c) \times A(-D)$

## 6.2 ในกรณีที่ตอบข้อสอบผิด

$$\theta_{m+1} = \theta_m - \left( \frac{\sigma_m^2}{\sqrt{\frac{1}{a^2} + \sigma_m^2}} \right) \left( \frac{O(D)}{A(D)} \right)$$

$$\sigma_{m+1}^2 = \sigma_m^2 \left( 1 - \left( \frac{O(D)}{1 + \frac{1}{a^2 \sigma_m^2}} \right) \left( \frac{O(D)}{A(D)} + D \right) \div A(D) \right) \quad (\text{Owen, 1975})$$

เมื่อ		
$\theta_m$	แทน	ความสามารถของผู้สอบที่ประมาณได้ก่อนตอบข้อสอบ ข้อที่ $m+1$ ตามปกติแล้ว ถ้าไม่ทราบค่าความสามารถเบื้องต้นของผู้สอบ จะกำหนดให้ $\theta_m$ เท่ากับ 0.000
$\sigma_m^2$	แทน	ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อสอบข้อที่ $m+1$ ตามปกติแล้วถ้าไม่ทราบค่าความแปรปรวนดังกล่าวมาก่อน จะกำหนดให้ $\delta_m^2$ เท่ากับ 1.000
$\theta_{m+1}$	แทน	ความสามารถของผู้สอบที่ประมาณได้หลังจากตอบข้อสอบข้อที่ $m+1$ แล้ว
$\sigma_{m+1}^2$	แทน	ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบหลังจากตอบข้อสอบข้อที่ $m+1$ แล้ว
a	แทน	ค่าอำนาจจำแนกของข้อสอบข้อที่ $m+1$
b	แทน	ค่าความยากของข้อสอบข้อที่ $m+1$
c	แทน	ค่าโอกาสในการเดาข้อสอบถูกของข้อสอบข้อที่ $m+1$
D	แทน	จุดบนแกน X
O(D)	แทน	ค่าออร์ดิเนต (Ordinate) ของโค้งปกติที่จุด D
A(D)	แทน	พื้นที่ใต้โค้งปกติจากค่า D ต่ำสุดจนถึงจุด D

7. ให้คะแนนโดยวิธีการประมาณค่าความเป็นไปได้สูงสุด (Maximum Likelihood) ซึ่งมีหลายวิธี แต่วิธีที่นิยมมากที่สุดคือวิธีความเป็นไปได้แบบมีเงื่อนไข (Conditional Maximum Likelihood) การประมาณค่าโดยวิธีนี้มีข้อจำกัด กล่าวคือ ถ้าผู้สอบตอบข้อสอบถูกต้องหรือผิดหมด จะไม่สามารถประมาณค่าได้ ขั้นตอนในการให้คะแนนมีดังนี้ (Hambleton and Swaminathan, 1985)

ขั้นที่ 1 ประมาณค่าความสามารถเริ่มต้น ( $\theta_m = 0; m = 0$ ) โดยใช้สูตร ดังนี้

$$\theta_0 = \ln \left( \frac{r_a}{k - r_a} \right)$$

เมื่อ  $r_a = \sum a_i U_i$

และ

$U_1 = 1$  เมื่อตอบข้อสอบถูก

$U_1 = 0$  เมื่อตอบข้อสอบผิด

$a_i$  คือค่าอำนาจจำแนกของข้อสอบข้อที่  $i$

$k$  คือจำนวนข้อสอบทั้งหมดที่มีผู้สอบตอบ

ขั้นที่ 2 หาค่า  $P_i(\theta_m)$  และ  $Q_i(\theta_m)$  โดยใช้สูตร ดังนี้

$$P_i(\theta_m) = c_i + (1 - c_i) \frac{e^{Da_i(\theta_m - b_i)}}{Da_i(\theta_m - b_i)}$$

$$Q_i(\theta_m) = 1 - P_i(\theta_m)$$

$$= \frac{1 - c_i}{1 + e^{Da_i(\theta_m - b_i)}}$$

- เมื่อ
- $\theta_m$  คือความสามารถของผู้สอบที่ประมาณค่าได้ในครั้งที่  $m$
  - $a_i$  คือค่าอำนาจจำแนกของข้อสอบข้อที่  $i$
  - $b_i$  คือค่าความยากของข้อสอบข้อที่  $i$
  - $c_i$  คือค่าโอกาสในการเดาข้อสอบได้ถูกของข้อสอบข้อที่  $i$
  - $D$  คือค่าคงที่ซึ่งมีค่าประมาณ 1.7
  - $e$  คือค่าคงที่ซึ่งมีค่าประมาณ 2.7182818

ขั้นที่ 3 หาค่าปรับแก้  $h_m$  โดยใช้สูตร ดังนี้

$$h_m = \frac{D[r_a - \sum P_i(\theta_m)]}{-D^2 \sum P_i(\theta_m) Q_i(\theta_m)}$$

ขั้นที่ 4 ประมาณค่าความสามารถของผู้สอบใหม่  $\theta_{m+1}$  โดยใช้สูตร ดังนี้

$$\theta_{m+1} = \theta_m - h_m$$

ขั้นที่ 5 ทำการคำนวณซ้ำในขั้นที่ 2, 3 และ 4 จนกระทั่ง  $h_m$  เข้าใกล้ศูนย์

$$(h_m < 0.001)$$

## 1.7 ประโยชน์การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

1.7.1 ใช้เวลาในการทดสอบน้อยลงกว่าการทดสอบด้วยกระดาษแบบเดิม เนื่องจากผู้สอบได้รับข้อสอบจำนวนน้อยกว่า ซึ่งข้อสอบที่ผู้สอบได้รับนี้เป็นข้อสอบที่มีค่าความยากง่ายเหมาะสมกับความสามารถของผู้สอบ (Urry, 1997; Moreno and others, 1984; Weiss, 1982)

1.7.2 สามารถรายงานผลได้ทันทีหลังจากเสร็จสิ้นการทดสอบ เนื่องจากการทดสอบแบบปรับเหมาะมักนิยมใช้เครื่องคอมพิวเตอร์เข้ามาเป็นตัวช่วยในการดำเนินการทดสอบ ดังนั้นคอมพิวเตอร์จะรับหน้าที่ในการตรวจข้อสอบ, คำนวณค่าความสามารถของผู้สอบ และคำนวณค่าความคลาดเคลื่อน เมื่อการทดสอบยุติลงก็สามารถแสดงผลได้ทันที

1.7.3 ผู้สอบมีเจตคติที่ดีต่อการทดสอบเพราะผู้สอบแต่ละคนจะได้ข้อสอบที่เหมาะสมกับความสามารถของตนเอง ผู้สอบที่มีความสามารถสูงก็ไม่รู้สึกเบื่อหน่ายในการทำข้อสอบที่ง่ายเกินไป ส่วนผู้ที่มีความสามารถต่ำก็ไม่เกิดความตึงเครียดในการทำข้อสอบที่ยากเกินไป (Lee, 1987; Hulin and others, 1983)

1.7.4 สามารถทดสอบนักเรียนเป็นรายบุคคลได้โดยไม่ต้องรอสอบพร้อมๆ กัน

1.7.5 ช่วยรักษาความลับและป้องกันการทุจริตในการสอบได้เป็นอย่างดี เพราะข้อสอบทั้งหมดถูกเก็บเอาไว้ในลักษณะของคลังข้อสอบ และในการทดสอบนักเรียนที่นั่งติดกันก็ไม่สามารถลอกคำตอบกันได้เนื่องจากจะได้รับข้อสอบไม่เหมือนกัน

1.7.6 เพิ่มความเที่ยงตรงและความเชื่อมั่นของการทดสอบได้ เพราะการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนั้นจะตัดข้อสอบที่ไม่เหมาะสมกับความสามารถของผู้สอบออกไป ซึ่งการทดสอบด้วยข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบจะช่วยให้ความเที่ยงตรงและความเชื่อมั่นของการวัดมีค่าสูงขึ้น (Larkin and Weiss, 1974; Weiss, 1984)

## 1.8 ประเด็นปัญหาของการใช้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ปัจจุบันยังมีประเด็นของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (computerized adaptive testing - CAT) ที่ยังเป็นข้อถกเถียงกันอยู่ นั่นคือหลังจากที่ผู้สอบทำการทดสอบเสร็จ ผู้สอบควรได้รับโอกาสให้กลับไปทวนคำตอบหรือไม่ ซึ่งเป็นที่ชัดเจนว่าผู้สอบยอมต้องการที่จะกลับไปทวนคำตอบ ซึ่งน่าจะได้มาถึงการประมาณค่าความสามารถที่แม่นยำหลังทวน และด้วยข้อเท็จจริงที่ว่า การเลือกคำตอบครั้งแรกอาจไม่ใช่คำตอบที่ถูกต้องที่สุดเสมอ แต่นักวิจัยที่มีชื่อเสียงบางท่านต่างมีความเห็นว่าไม่ควรเปิดโอกาสให้มีการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เนื่องจากข้อจำกัดของกระบวนการการทดสอบ ตัวอย่างเช่น

Wainer (1993) กล่าวว่า ที่ไม่ควรให้มีการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เพราะ จะสูญเสียความมีประสิทธิภาพ และเกิดความลำเอียงในการประมาณค่าความสามารถของผู้สอบ รวมทั้งเกิดความคาดเคลื่อนในการได้มาถึงความสามารถที่แท้จริงของผู้สอบในกระบวนการการทดสอบ ทำให้ขัดแย้งกับกระบวนการสร้างแบบทดสอบแบบปรับเหมาะที่มีประสิทธิภาพ โดย Wainer ได้ยกตัวอย่างสถานการณ์ที่ทำให้ผู้สอบได้คะแนนสูงขึ้น กล่าวคือ เมื่อผู้สอบตั้งใจที่จะตอบคำถามผิด ดังนั้นจึงทำให้ได้รับข้อสอบข้อต่อไปที่มีระดับความยากต่ำลง และถ้าเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้โดยไม่มีเงื่อนไข ผู้สอบก็จะสามารถกลับไปแก้คำตอบข้อก่อนหน้า ทำให้ข้อสอบแต่ละข้อนั้นไม่สัมพันธ์กับระดับความสามารถที่แท้จริงของผู้สอบ จากงานวิจัยพบว่าผู้สอบจะได้รับการประมาณค่าความสามารถโน้มไปทางด้านบวกหรือได้รับคะแนนที่สูงเกินความสามารถที่แท้จริง และมีความคาดเคลื่อนสูง

Green, Bock, Humphreys, Linn, and Reckase (1984) กล่าวว่า อาจเกิดความสับสนและความล่าช้าในการทำการทดสอบ จากการที่ผู้สอบต้องคอยนี้่กย้อนถึงการตอบข้อสอบข้อที่ผ่านมา

การต่อต้านการทวนคำตอบใน CAT ข้างต้น จะเห็นได้ว่าเกิดจากปัญหาที่มีลักษณะเฉพาะจากการทวนคำตอบ ดังนั้นถึงแม้ว่าการทวนคำตอบจะก่อให้เกิดผลดีมากมายแต่ก็เฉพาะกรณีที่จะสามารถขจัดปัญหาต่างๆที่เกิดขึ้นให้น้อยลงหรือหมดไปเท่านั้น (Bowles, 2001; Weise, 1996) 3 วิธีการที่นำมาใช้ในการทวนคำตอบ (Stocking, 1997) ที่สามารถควบคุมการประมาณค่าความสามารถที่เกินจริงได้ คือ การทวนที่มีการจำกัดจำนวน (Limited Review), การทวนที่มีการจำกัดช่วง (Block Review) และการทวนในหมวดสิ่งเร้าเดียวกัน (Stimulus-based Review) ซึ่งแต่ละวิธีมีหลักการ ดังนี้

1) การทวนที่มีการจำกัดจำนวน (Limited Review) คือ ผู้สอบอาจย้อนกลับไปทวนคำตอบได้ทั้งหมด แต่จะเปลี่ยนคำตอบได้ในจำนวนที่จำกัดเท่านั้น

2) การทวนที่มีการจำกัดช่วง (Block Review) คือ ผู้สอบจะสามารถทวนคำตอบพร้อมทั้งเปลี่ยนคำตอบภายในช่วงที่จำกัดไว้ได้ทั้งหมด แต่ไม่สามารถที่จะทวนข้ามช่วงได้ ยกตัวอย่างเช่น การจำกัดช่วง 10 ข้อ กล่าวคือผู้สอบจะสามารถกลับมาทวนคำตอบภายในช่วง (ข้อ 1 – 10) ได้หลังจากที่ทำเสร็จข้อ 10 หลังจากนั้นจะไปทำข้อสอบในช่วงถัดไป (ข้อ 11 - 20) และจะสามารถกลับมาทวนคำตอบในช่วงนี้ได้ก็ต่อเมื่อทำเสร็จถึงข้อ 20 แต่ไม่สามารถกลับไปทวนคำตอบในช่วงก่อนหน้านี้ (ข้อ 1 - 10) ได้ เป็นต้น

3) การทวนในหมวดสิ่งเร้าเดียวกัน (Stimulus-based Review) มีความคล้ายคลึงกับการทวนที่มีการจำกัดช่วง แต่แตกต่างตรงความยาวของช่วง กล่าวคือ การทวนที่มี

การจำกัดช่วงนั้นจะแบ่งความยาวของข้อสอบออกเป็นช่วง ช่วงละเท่าๆ กัน แต่การทวนในหมวดสิ่งเร้าเดียวกันจะแบ่งความยาวช่วงตามหมวดสิ่งเร้า เช่น แบบทดสอบเรื่องทศนิยม ประกอบด้วยเนื้อหาเรื่อง การแปลงเศษส่วนให้เป็นทศนิยม, การอ่านทศนิยม, การบวกทศนิยม และการลบทศนิยม ดังนั้นจึงแบ่ง หมวดสิ่งเร้าตามเนื้อหาของแบบทดสอบเป็น 4 หมวด กล่าวคือ การแปลงเศษส่วนให้เป็นทศนิยมบวก (ข้อ 1 - 4), การอ่านทศนิยม (ข้อ 5 - 10), การบวกทศนิยม (ข้อ 11 - 15) และการลบทศนิยม (ข้อ 16 - 20) ซึ่งจะเห็นได้ว่าช่วงของการทวนจะมีความยาวของข้อสอบไม่เท่ากัน

Vispole และคณะ (2000) ได้สนใจวิธีการทวนที่มีการจำกัดช่วง (Block Review) จึงทำการศึกษาเปรียบเทียบแต่ละเงื่อนไขการทวน คือ การไม่ให้ทวน, การทวนที่มีการจำกัดช่วง 5 ข้อ, การทวนที่มีการจำกัดช่วง 10 ข้อ และการทวนที่มีการจำกัดช่วง 40 ข้อ พบว่า ในขณะที่ลดขนาดช่วงลง ผู้สอบจะทำเครื่องหมายเพื่อกลับมาทวนเพิ่มขึ้น, เปลี่ยนคำตอบมากขึ้น, ใช้เวลาในการทวนคำตอบมากขึ้น, มีแนวโน้มใช้เวลาในการทำข้อสอบนานขึ้น และค่าความสามารถเฉลี่ยมากกว่าเงื่อนไขการไม่ให้ทวน แต่ทว่าไม่แตกต่างกันมากระหว่างเงื่อนไขการให้ทวนในช่วงต่างๆ ดังนั้น Vispole จึงได้ให้ข้อเสนอว่า การทวนแบบจำกัดช่วงอาจทำหน้าที่ได้ดีพอๆ กับการทวนได้ทั้งหมด แต่สามารถลดปัญหาการโกงข้อสอบจากการเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้ เนื่องจากผู้สอบต้องตอบข้อสอบให้เสร็จก่อนภายในช่วง จึงจะสามารถกลับไปทวนคำตอบได้ จึงเป็นการประมาณค่าความสามารถของผู้สอบไว้แล้วครั้งหนึ่ง รวมทั้งยังมีการประมาณค่าความสามารถของผู้สอบระหว่างช่วงด้วยเมื่อผู้สอบทำการทวนเสร็จในช่วง

ดังนั้นโดยหลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีเป้าหมายที่จะดำเนินการทดสอบด้วยการคัดเลือกข้อสอบให้เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล ดังนั้นผู้สอบที่มีระดับความสามารถสูงก็จะได้รับข้อสอบที่ยาก ส่วนผู้มีระดับความสามารถต่ำก็จะได้รับคำตอบที่ง่ายดังนั้นการกลับไปทวนคำตอบของผู้สอบที่มีระดับความสามารถอื่นๆ ซึ่งไม่ใช่ระดับความสามารถสูงแล้วทำให้มีการเปลี่ยนคำตอบจากถูกเป็นผิด หรือจากผิดเป็นผิด รวมทั้งมีการใช้เวลาในการทำข้อสอบนานขึ้นเนื่องจากเสียเวลากลับไปทวนคำตอบนั้น จะเกิดผลดีอย่างแท้จริงหรือไม่กับการเปิดโอกาสให้มีการทวนคำตอบในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ดังนั้นจึงควรมีการวิจัยเกี่ยวกับประเด็นนี้เพื่อนำไปเป็นแนวทางในการยอมรับหรือไม่ยอมรับการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ในอนาคตต่อไป



## ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ ทฤษฎีนี้เสนอแนวคิดที่ว่า ความน่าจะเป็นของการตอบข้อสอบได้ถูกต้องขึ้นอยู่กับความสามารถจริงของผู้สอบ และคุณลักษณะของผู้สอบ อันประกอบด้วยพารามิเตอร์ความยาก อำนาจจำแนก และโอกาสการเดาข้อสอบได้ถูก ระบบความสัมพันธ์ดังกล่าว สามารถแสดงได้ด้วยโมเดลการตอบสนองข้อสอบ ซึ่งอาจเป็นแบบโมเดล 1 พารามิเตอร์ โมเดล 2 พารามิเตอร์ หรือโมเดล 3 พารามิเตอร์ (ศิริชัย กาญจนวาสี, 2545) ซึ่งผู้วิจัยได้นำเสนอเนื้อหาในตอนที่ 2 ออกเป็น 6 ตอนย่อย ดังนี้

- 2.1 ความเป็นมาของทฤษฎีการตอบสนองข้อสอบ
- 2.2 หลักการของทฤษฎีการตอบสนองข้อสอบ
- 2.3 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ
- 2.4 แบบของการวิเคราะห์ข้อสอบ
- 2.5 ปัจจัยที่มีผลต่อคุณภาพของแบบทดสอบ
- 2.6 การนำทฤษฎีการตอบสนองข้อสอบไปใช้

### 2.1 ความเป็นมาของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการทดสอบแบบประเพณีนิยม (Classical Test Theory) เป็นทฤษฎีที่ไม่เหมาะสมกับการทดสอบแบบปรับเหมาะเนื่องจากค่าความเชื่อมั่น ความแม่นยำ และคุณภาพของแบบทดสอบแบบประเพณีนิมนั้นเหมาะสำหรับเมื่อผู้สอบทุกคนได้ตอบข้อสอบชุดเดียวกัน แต่การทดสอบแบบปรับเหมาะผู้สอบแต่ละคนจะตอบข้อสอบต่างกันจึงจำเป็นต้องใช้ทฤษฎีที่เหมาะสมเช่น ทฤษฎีความสามารถแฝง (Latent Trait Theory) (Birnbaum, 1958 อ้างถึงใน Green and Others, 1984) ซึ่งต่อมาเรียกว่าทฤษฎีการตอบสนองข้อสอบ (Item Response Theory; IRT) ทฤษฎีนี้อธิบายถึงความสัมพันธ์ระหว่างการตอบข้อสอบกับระดับความสามารถที่มีอยู่ภายในตัวผู้สอบโดยใช้ตัวแบบทางคณิตศาสตร์ที่กำหนดด้วยฟังก์ชันที่เรียกว่าฟังก์ชันลักษณะของข้อสอบ (item characteristic function) หรือโค้งลักษณะเฉพาะของข้อสอบ (Item Characteristic Curve; ICC) เพื่อที่จะกำหนดความสัมพันธ์ระหว่างความสามารถ (traits) ของข้อสอบและการตอบข้อสอบจากฟังก์ชันทางคณิตศาสตร์ ความสัมพันธ์นี้สามารถดัดแปลงได้หลายแบบตามลักษณะข้อตกลงเบื้องต้นเกี่ยวกับทฤษฎีการตอบสนองข้อสอบ

แต่อย่างไรก็ตาม หากนักทดสอบไม่สามารถทำการวิเคราะห์ข้อสอบตามแบบทฤษฎีการตอบสนองข้อสอบตามแบบดังกล่าว หรือเพราะขาดความชำนาญในการแปล

ความหมายของผลลัพธ์จากการวิเคราะห์ ในกรณีนี้นักทดสอบก็อาจนำการวิเคราะห์ข้อสอบตามแบบประเพณีนิยมมาใช้ได้ หากได้ทำการแก้ไขจุดอ่อนของการวิเคราะห์แล้ว

จุดเริ่มต้นของทฤษฎีการตอบสนองข้อสอบเริ่มจากผลงานของเฟอร์กูสัน (Ferguson) และลอร์ลีย์ (Lawley) ในปี ค.ศ.1942 และ ค.ศ.1943 (Warm, 1978) ซึ่งถือว่าเป็นผลงานที่แสดงแนวคิดและหลักการเบื้องต้นที่สำคัญของทฤษฎีการตอบสนองข้อสอบซึ่งเดิมเรียกว่าทฤษฎีความสามารถแฝง ในระยะต่อมาได้มีนักทดสอบอีกหลายคนที่มีแนวคิดเดียวกันนี้ เช่น บรอกเดนต (Brogdent,1946) ทักเกอร์ (Tucker,1946) แครร์รอล (Carrol,1950) ครอนบาค และวอร์ริงตัน (Cronbach and Warrington, 1952) ต่อมาในปี ค.ศ. 1952 ลอร์ดได้เสนอทฤษฎีโค้งลักษณะเฉพาะของข้อสอบ และกล่าวว่าพฤติกรรมคำตอบข้อสอบแต่ละข้อของผู้สอบสามารถอธิบายได้ในเชิงคณิตศาสตร์ที่เราคุ้นเคยกันเป็นอย่างดีแล้ว คือ โค้งลักษณะเฉพาะของข้อสอบแต่ละข้อซึ่งมีลักษณะเป็นโค้งความถี่สะสม หรือที่เรียกว่าแบบโค้งรูปตัวเอสปกติ (normal curve model) ซึ่งแบบนี้กล่าวถึงค่าพารามิเตอร์ 2 ตัวคือ ค่าความยากและค่าอำนาจจำแนก แต่เนื่องจากแบบโค้งรูปตัวเอสปกติต้องใช้การคำนวณที่ยุ่งยากซับซ้อน และขณะนั้นยังขาดโปรแกรมคอมพิวเตอร์ที่จำเป็นต้องใช้ในการวิเคราะห์ข้อมูลตามความคิดดังกล่าว ลอร์ดจึงไม่ได้พัฒนาแนวคิดของตนต่อไป

ต่อมาในปี ค.ศ. 1960 ราช (Rasch) ได้เสนอแนวคิดเช่นเดียวกับลอร์ดแต่อาศัยค่าพารามิเตอร์เดียว คือใช้ค่าความยากของข้อสอบเพียงค่าเดียว เรียกกันว่า การวิเคราะห์ข้อสอบตามแบบราช (Rasch model) จากการเสนอแนวคิดของราชทำให้ลอร์ดกลับมาสนใจทฤษฎีนี้อีกครั้งในปี ค.ศ. 1965

ในปี ค.ศ. 1968 เบิร์นเบอร์ (Birnbaum) ได้ศึกษาแนวคิดเดิมที่ลอร์ดเคยเสนอไว้แล้วแต่ใช้แบบลอจิสติกแทน และใช้พารามิเตอร์ 2 ตัว คือ ค่าความยากและค่าอำนาจจำแนก ซึ่งเป็นแบบที่ง่ายกว่าของลอร์ดจึงเป็นทฤษฎีที่นิยมใช้กันแพร่หลายและได้รับการพัฒนาขึ้นเรื่อยๆจนเรียก ทฤษฎีการทดสอบนี้ว่า “ทฤษฎีการตอบสนองข้อสอบ” และเรียกวิธีการวิเคราะห์ข้อสอบรายข้อนี้ว่า การวิเคราะห์ข้อสอบตามแบบลอจิสติก 1 ซึ่งสามารถใช้ได้กับพารามิเตอร์ 1 ตัวและมีการวิเคราะห์ข้อสอบตามแนวคิดเดียวกันแต่ซับซ้อนมากขึ้นโดยใช้พารามิเตอร์ 2 ตัวและ 3 ตัว

## 2.2 หลักการของทฤษฎีการตอบสนองข้อสอบ

หลักการที่สำคัญของทฤษฎีก็คือ ผลการทดสอบของผู้สอบจากแบบทดสอบใดๆ ขึ้นอยู่กับความสามารถของผู้สอบหรือคะแนนของผู้สอบจากแบบทดสอบใดๆ นั้นสามารถพยากรณ์หรืออธิบายได้จากความสามารถของคนๆ นั้น ดังนั้นทฤษฎีนี้จึงเป็นทฤษฎีที่อธิบาย

ความสัมพันธ์ระหว่างปริมาณของการกระทำหรือคะแนนของผู้สอบกับปริมาณความสามารถของผู้สอบ ซึ่งแสดงได้ดังสมการต่อไปนี้

$$P = f(\theta)$$

เมื่อ	P	แทน ผลการทดสอบ (performance)
	$\theta$	แทน ความสามารถ (ability หรือ trait)
	f	แทน ความสัมพันธ์ (function)

นอกจากนี้ ทฤษฎีนี้ยังมุ่งหาความสัมพันธ์ระหว่างความสามารถที่แท้จริงกับการตอบสนองข้อสอบ ซึ่งในที่นี้ความสามารถที่แท้จริงนั้นคือคุณลักษณะที่วัดได้ด้วยแบบทดสอบ เช่น ความรู้ ความเข้าใจหรือทักษะในด้านต่างๆ ส่วนการตอบสนองข้อสอบ (response) หมายถึงการตอบข้อสอบแบบปรนัยที่เมื่อตอบถูกต้อง 1 คะแนน และตอบผิดได้ 0 คะแนน

จากความสัมพันธ์ในสมการดังกล่าวเป็นการแสดงความสัมพันธ์ทั่วๆ ไปยังไม่เฉพาะเจาะจง จึงมีผู้สนใจศึกษาและพัฒนาเพื่อให้ได้สมการที่สามารถอธิบายลักษณะเฉพาะของข้อสอบแต่ละข้อที่สามารถอธิบายความสัมพันธ์ของสองตัวแปรนี้ได้เด่นชัดขึ้น โดยอาศัยข้อตกลงบางประการและวิธีการทางคณิตศาสตร์

นอกจากนี้แล้ว ถ้าข้อตกลงเบื้องต้นของการทดสอบตรงตามที่ทฤษฎีกำหนดไว้ ผลที่ได้จากการคำนวณจะมีลักษณะที่ดีดังนี้

1. ค่าพารามิเตอร์ของข้อสอบ คือค่าความยาก ค่าอำนาจจำแนก และค่าโอกาสในการเดาข้อสอบได้ถูกจะเป็นค่าที่ไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบ ไม่ว่าจะนำไปสอบกับผู้ใดก็ตาม ลักษณะเช่นนี้เรียกว่า ความเป็นอิสระจากกลุ่มตัวอย่าง (Person-free test calibration)

2. เมื่อทราบลักษณะการตอบข้อสอบแต่ละข้อของผู้สอบคนใด เราสามารถจะคำนวณหาค่าความสามารถที่แท้จริงของบุคคลนั้นได้ ค่าความสามารถที่แท้จริงนี้มีความสัมพันธ์โดยตรงกับคะแนนจริง การคำนวณความสามารถที่แท้จริงนั้นอาจใช้ข้อสอบข้อใดก็ได้ที่วัดสิ่งเดียวกันลักษณะเช่นนี้เรียกว่า ความเป็นอิสระของข้อสอบ (item-free person measurement)

### 2.3 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีข้อตกลงเบื้องต้นที่สำคัญอยู่ 4 ประการ คือ (Warm, 1978)

2.3.1 การรู้คำตอบที่ถูกต้อง (the know-correct assumption) หมายความว่าในการตอบข้อสอบข้อใดข้อหนึ่งนั้น ถ้าผู้สอบรู้คำตอบที่ถูกต้องแล้วจะตอบข้อสอบข้อนั้นถูก หรือถ้าผู้สอบตอบข้อสอบข้อใดข้อหนึ่งผิดแสดงว่าผู้สอบไม่ทราบคำตอบที่ถูกต้องของข้อสอบข้อนั้น

2.3.2 แบบทดสอบมิติเดียว (unidimensional test) หมายความว่าข้อสอบแต่ละข้อในแบบทดสอบฉบับนั้นจะต้องวัดความสามารถหรือคุณลักษณะเดียวกัน หรือมีความเป็นเอกพันธ์ (Homogeneous items) กัน ทั้งนี้เพื่อให้ง่ายต่อการรวมคะแนนและการแปลความหมายของคะแนนผลการสอบ การตรวจสอบว่าแบบทดสอบฉบับหนึ่งใช้วัดความสามารถด้านเดียวหรือไม่ นั้นมีวิธีการทดสอบได้หลายอย่าง แต่ที่นิยมใช้กันมากมี 2 วิธีคือ

1) วิธีทดสอบความสัมพันธ์ (baseline test) โดยการหาค่าความสัมพันธ์ระหว่างค่าน้ำหนักองค์ประกอบรายข้อ (factor loading) ขององค์ประกอบ (factor) ที่หนึ่งกับค่าสหสัมพันธ์แบบไบซีเรียล (biserial correlation) ของข้อสอบรายข้อกับคะแนนรวม หากค่าสัมประสิทธิ์สหสัมพันธ์มีค่าสูงมากกว่า 0.80 แสดงว่าแบบทดสอบนั้นวัดเพียงมิติเดียว

2) วิธีทดสอบค่ากรรมที่เจาะจง (eigenvalue test) โดยการวิเคราะห์ตัวประกอบ (factor analysis) ของข้อสอบทั้งฉบับ หากปรากฏว่าค่ากรรมที่เจาะจงที่ 1 มีค่าสูง แล้วแตกต่างจากค่ากรรมที่เจาะจงที่ 2 มาก และค่ากรรมที่เจาะจงที่ 2 แตกต่างจากค่ากรรมที่เจาะจงที่ 3 และค่ากรรมที่เจาะจงอื่นๆ ไม่มากแสดงว่าแบบทดสอบนั้นวัดเพียงมิติเดียว

2.3.3 ความเป็นอิสระของข้อสอบ (local independence) หมายความว่าข้อสอบแต่ละข้อจะต้องมีอิสระจากกันในเชิงเนื้อหา และจะปรากฏอยู่ในส่วนใดของแบบทดสอบก็ได้ ดังนั้นโอกาสที่ผู้สอบจะทำข้อสอบข้อใดได้ถูกต้องไม่มีผลต่อการตอบข้อสอบข้ออื่นๆ ในแบบทดสอบนั้น

2.3.4 โค้งลักษณะของข้อสอบ (item characteristic curve) หมายความว่าโอกาสที่ผู้สอบจะตอบข้อสอบข้อใดข้อหนึ่งได้ถูกต้องหรือไม่ขึ้นอยู่กับโค้งลักษณะเฉพาะของข้อสอบของแต่ละแบบที่ใช้หรือไม่ขึ้นอยู่กับลักษณะการแจกแจงความสามารถของประชากร นั่นคือรูปร่างของโค้งลักษณะเฉพาะของข้อสอบไม่ได้ขึ้นอยู่กับลักษณะการกระจายของความสามารถของกลุ่มผู้สอบ

ถ้าเป็นไปได้ตามข้อตกลงดังกล่าวข้างต้น การวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบจะให้ข้อดีหรือข้อได้เปรียบดังนี้

1) การประมาณค่าความสามารถของผู้สอบแต่ละคนจะถูกต้องหรือไม่ขึ้นอยู่กับคำตอบข้อสอบที่เลือกมาจากกลุ่มข้อสอบที่วัดความสามารถด้านเดียวกัน

2) ค่าพารามิเตอร์ของข้อสอบจะไม่ผันแปรไปตามกลุ่มตัวอย่างผู้สอบ

3) ผลการทดสอบจะให้การประมาณค่าความสามารถที่ถูกต้องแน่นอนในแต่ละระดับความสามารถ และสามารถประมาณความคลาดเคลื่อนมาตรฐานของการวัดของแต่ละบุคคลได้ด้วย

## 2.4 แบบของการวิเคราะห์ข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีแบบในการวิเคราะห์ข้อสอบที่นิยมใช้กันทั่วไป 3 แบบ แต่ละแบบอาศัยแนวคิดของโค้งลอจิสติก (logistic curve) คือ

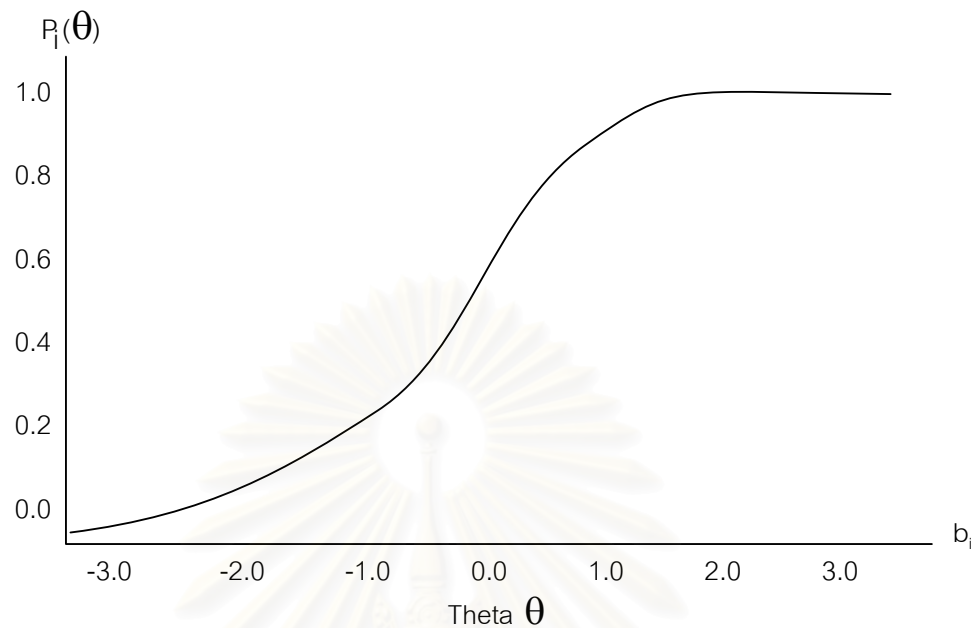
### 2.4.1 แบบ 1 พารามิเตอร์

ราชได้พัฒนาทฤษฎีการตอบสนองข้อสอบและเสนอการวิเคราะห์ข้อสอบแบบ 1 พารามิเตอร์ในปี ค.ศ. 1960 โดยฟังก์ชันของแบบ 1 พารามิเตอร์นี้สามารถอธิบายได้ด้วยพารามิเตอร์ของข้อสอบเพียงตัวเดียวคือค่าความยาก ( $b_i$ ) จึงนิยมเรียกการวิเคราะห์ข้อสอบแบบนี้ว่า แบบราช (Rasch model) ซึ่งตรงกับแบบที่ใช้พารามิเตอร์เดียวที่เบรินแฮมพัฒนาขึ้นและสามารถเขียนเป็นสมการได้ดังนี้

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}} ; i = 1, 2, 3, \dots, n$$

ในเมื่อ  $P_i(\theta)$  คือความน่าจะเป็นที่ผู้สอบมีความสามารถ  $\theta$  จะตอบข้อสอบข้อที่  $i$  ได้ถูกต้อง  
 $D$  คือค่าคงที่ของ scale factor เพื่อปรับค่าของฟังก์ชันลอจิสติกให้ใกล้เคียงกับค่าที่ได้จากฟังก์ชันรูปตัวเอสปกติ มีค่าประมาณ 1.7  
 $b_i$  คือค่าความยากของข้อสอบข้อที่  $i$   
 $\theta$  คือระดับความสามารถที่แท้จริงที่คำนวณจากคะแนนรวมแต่ปรับหน่วยเป็นมาตรฐาน  
 $e$  คือค่าคงที่ซึ่งมีค่าประมาณ 2.7182818

การวิเคราะห์ข้อสอบแบบ 1 พารามิเตอร์ (1-parameter model) นี้มีข้อตกลงว่าข้อสอบทุกข้อไม่มีโอกาสของการเดาถูก ( $c_i = 0$ ) และค่าอำนาจจำแนกของข้อสอบ ( $a_i$ ) ทุกข้อเท่ากันหมด โค้งลักษณะเฉพาะของข้อสอบแบบนี้แสดงไว้ในแผนภาพที่ 11



แผนภาพที่ 11 โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อทดสอบแบบ 1 พารามิเตอร์

#### 2.4.2 แบบ 2 พารามิเตอร์

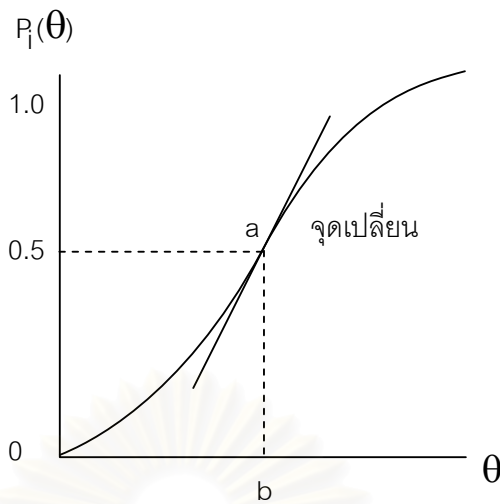
ในปี ค.ศ. 1968 เบิร์นแฮมได้เสนอการวิเคราะห์ข้อสอบแบบลอจิสติกที่มีค่าพารามิเตอร์ 2 ตัวคือ ค่าความยาก ( $b_i$ ) และค่าอำนาจจำแนก ( $a_i$ ) และสามารถเขียนเป็นสมการได้ดังนี้

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} ; i = 1, 2, 3, \dots, n$$

เมื่อ  $D$  = ค่าคงที่ของ scale factor มีค่าประมาณ 1.7

$a_i$  = ค่าอำนาจจำแนกของข้อสอบข้อที่  $i$

การวิเคราะห์ข้อสอบแบบ 2 พารามิเตอร์ (2-parameter model) นี้มีข้อตกลงว่าข้อสอบทุกข้อไม่มีโอกาสของการเดาถูก คือค่า  $c_i = 0$  โค้งลักษณะเฉพาะของข้อสอบแบบนี้มีลักษณะดังแผนภาพที่ 12



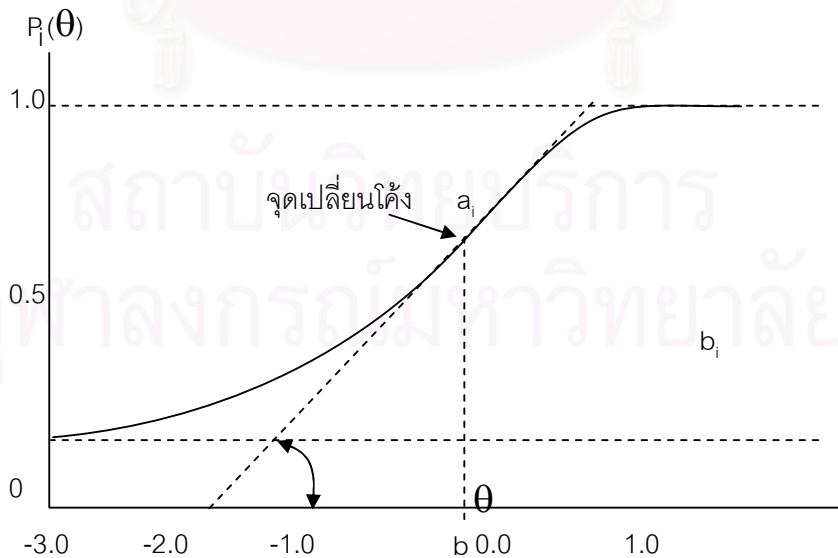
แผนภาพที่ 12 ไค้ลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อทดสอบแบบ 2 พารามิเตอร์

### 2.4.3 แบบ 3 พารามิเตอร์

การวิเคราะห์ข้อสอบแบบ 3 พารามิเตอร์ (3-parameter model) นี้ดัดแปลงมาจากแบบที่ใช้ 2 พารามิเตอร์ พารามิเตอร์ที่เพิ่มขึ้นคือค่าโอกาสในการเดาข้อสอบได้ถูก ( $c_i$ ) และมีสมการดังนี้

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} ; i = 1, 2, 3, \dots, n$$

เมื่อ  $c_i$  = ค่าโอกาสในการเดาข้อสอบได้ถูก หมายถึงค่าโอกาสของการตอบข้อสอบถูกโดยการเดาของผู้สอบที่มีความสามารถน้อย (อ่อน)



แผนภาพที่ 13 ไค้ลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อทดสอบแบบ 3 พารามิเตอร์

## 2.5 ปัจจัยที่มีผลต่อคุณภาพของแบบทดสอบ

ตามแนวคิดของทฤษฎีการตอบสนองข้อสอบเชื่อว่าแบบทดสอบจะมีคุณภาพดีเพียงใดขึ้นอยู่กับคุณภาพของข้อสอบแต่ละข้อว่าสามารถทดสอบสิ่งที่ต้องการทดสอบได้มากน้อยเพียงใด หรือกล่าวอีกนัยหนึ่งก็คือข้อสอบสามารถทดสอบได้ถูกต้องแม่นยำมากเพียงใดซึ่งขึ้นอยู่กับปัจจัยหลายประการ เช่น

### 2.5.1 ความถูกต้องแม่นยำ

ในการประมาณค่าความสามารถของผู้สอบ โดยปกติไม่ว่าจะเป็นการวัดหรือการประมาณค่าใดๆ ก็ตามจะต้องมีความคลาดเคลื่อนอยู่เสมอ โดยเฉพาะอย่างยิ่งในการวัดหรือการประมาณค่าความสามารถในทางการศึกษาและจิตวิทยาจะมีความคลาดเคลื่อนค่อนข้างสูง ถ้าความคลาดเคลื่อนในการประมาณค่ามีค่าสูง ความถูกต้องแม่นยำ (accuracy) ในการประมาณค่าก็จะต่ำ ซึ่งสามารถเขียนเป็นสมการแสดงความสัมพันธ์ระหว่างความสามารถจริง ( $\theta$ ) ความสามารถที่ได้จากการประมาณค่า ( $\hat{\theta}$ ) และค่าความคลาดเคลื่อน (E) ได้ดังนี้ (Lord and Novick 1968: 175)

$$E = \hat{\theta} - \theta$$

จากสมการนี้จะเห็นว่าถ้าค่าความคลาดเคลื่อนมีค่าน้อยหรือเข้าใกล้ศูนย์ ความสามารถที่ได้จากการประมาณค่าก็จะเข้าใกล้ความสามารถจริง นั่นคือ มีความถูกต้องแม่นยำในการประมาณความสามารถของผู้สอบสูง

### 2.5.2 ความคลาดเคลื่อนมาตรฐานในการประมาณค่า

ความคลาดเคลื่อนมาตรฐานในการประมาณค่า (standard error of estimate; SEE) คือส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในการประมาณค่า ซึ่งแตกต่างจากความคลาดเคลื่อนมาตรฐานในการวัด (standard error of measurement; SEM) กล่าวคือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าเป็นความคลาดเคลื่อนของค่าความสามารถของผู้สอบที่คลาดเคลื่อนไปจากความสามารถจริง ( $E = \hat{\theta} - \theta$ ) ส่วนความคลาดเคลื่อนมาตรฐานในการวัดคือความคลาดเคลื่อน (E) ที่เกิดจากคะแนนดิบหรือคะแนนที่ได้จากการวัด (X) ที่คลาดเคลื่อนไปจากคะแนนจริง (T) นั่นคือ

$$E = X - T$$



ความคลาดเคลื่อนมาตรฐานในการประมาณค่ามีความสัมพันธ์กับค่าฟังก์ชันสารสนเทศของแบบทดสอบ (test information function) คือ  $SEE = 1/\sqrt{I(\theta)}$  ค่าฟังก์ชันสารสนเทศของแบบทดสอบจะเป็นตัวบ่งชี้ถึงคุณภาพของการประมาณค่าและยังสามารถใช้แทนค่าความเชื่อมั่นหรือค่าความคลาดเคลื่อนมาตรฐานในการวัดในทฤษฎีการทดสอบแบบประเพณีนิยมได้ด้วย (Hambleton and Cook, 1977) กรีนและคณะ (Green and Others, 1984) ได้ชี้ให้เห็นว่าความคลาดเคลื่อนมาตรฐานในการวัดมีข้อตกลงที่ไม่ถูกต้องที่ว่าความแปรปรวนของคะแนนความคลาดเคลื่อนของทุกๆ คะแนนมีเท่ากัน ดังนั้นการทดสอบแบบปรับเหมาะจึงควรใช้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าแทน แต่การที่จะคำนวณหาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าได้จำเป็นต้องรู้ค่าฟังก์ชันสารสนเทศของแบบทดสอบซึ่งต้องอาศัยค่าฟังก์ชันสารสนเทศของข้อสอบ (item information function)

### 2.5.3 ฟังก์ชันสารสนเทศของข้อสอบ

ฟังก์ชันสารสนเทศของข้อสอบคือค่าที่แสดงถึงความถูกต้องแม่นยำในการประมาณค่าความสามารถจริง ของผู้สอบในการตอบข้อสอบแต่ละข้อ ซึ่งสามารถคำนวณได้จากสูตรต่อไปนี้

$$I(\theta, u_i) = P_i'(\theta)^2 / P_i(\theta)Q_i(\theta)$$

เมื่อ	$I(\theta, u_i)$	คือฟังก์ชันสารสนเทศของข้อสอบ
	$P_i'(\theta)$	คือความชันของโค้งลักษณะเฉพาะของข้อสอบที่ระดับความสามารถ $\theta$
	$P_i(\theta)$	คือความน่าจะเป็นที่ผู้สอบที่มีความสามารถ $\theta$ จะตอบข้อสอบข้อที่ $i$ ได้ถูก
และ	$Q_i(\theta) = 1 - P_i(\theta)$	

จากสมการข้างต้นจะเห็นว่าฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อขึ้นอยู่กับความชันของโค้งลักษณะเฉพาะของข้อสอบ ถ้าโค้งลักษณะเฉพาะของข้อสอบชันมากขึ้นในขณะที่ความแปรปรวนของการตอบข้อสอบถูกน้อยลง โค้งสารสนเทศของข้อสอบที่ระดับความสามารถนั้นๆ จะยิ่งสูงขึ้น ความสูงของโค้งสารสนเทศของข้อสอบอยู่ที่ระดับความสามารถใดแสดงว่าสามารถจำแนกระดับความสามารถของผู้สอบได้ดี ณ ระดับความสามารถนั้น

### 2.5.4 ฟังก์ชันสารสนเทศของแบบทดสอบ

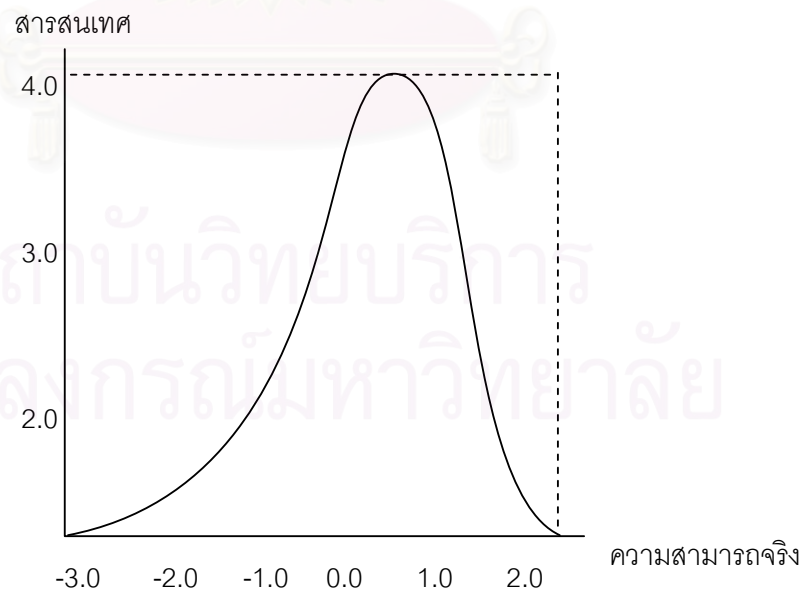
ฟังก์ชันสารสนเทศของแบบทดสอบเป็นส่วนกลับกันกับกำลังสองของความยาวของช่วงความเชื่อมั่น ซึ่งเป็นผลมาจากการประมาณค่าความสามารถของผู้สอบจากการตอบข้อสอบทั้งฉบับ คำนี้นำแสดงถึงความถูกต้องแม่นยำในการประมาณค่าความสามารถจริงของผู้สอบแบบทดสอบทั้งฉบับว่ามีมากน้อยเพียงใด ดังนั้นโค้งสารสนเทศของแบบทดสอบจึงเป็นดัชนีนำแสดงถึงความถูกต้องแน่นอนของค่าความสามารถของผู้สอบ ซึ่งแสดงในรูปสมการได้ดังนี้

$$I(\theta) = \sum_{i=1}^n i(\theta, u_i) = \sum_{i=1}^n \frac{P_i'(\theta)^2}{P_i(\theta)Q_i(\theta)}$$

เมื่อ  $I(\theta)$  คือ ฟังก์ชันสารสนเทศของแบบทดสอบ  
 $I(\theta, u_i)$  คือ ฟังก์ชันสารสนเทศของข้อสอบ

ถ้าเรามีกลุ่มข้อสอบที่ทราบค่าสารสนเทศของข้อสอบเราสามารถสร้างแบบทดสอบให้มีโค้งสารสนเทศของแบบทดสอบ ณ ความสามารถระดับหนึ่งที่เราต้องการได้ เช่น การสร้างแบบทดสอบเพื่อคัดเลือกนักเรียนเข้าศึกษาต่อ ก็ต้องสร้างแบบทดสอบที่มีประสิทธิภาพสูงสุดที่ระดับความสามารถสูงๆ นั่นคือ ให้มีโค้งสารสนเทศของแบบทดสอบสูง ณ ระดับความสามารถสูงๆ นั่นเอง

ตัวอย่างโค้งสารสนเทศของแบบทดสอบที่เหมาะสมสำหรับวัดความสามารถที่ระดับความสามารถ 1.0 ซึ่งจะมีลักษณะดังแผนภาพที่ 14



แผนภาพที่ 14 โค้งสารสนเทศของแบบทดสอบสำหรับวัดความสามารถที่ระดับ 1.0

## 2.6 การนำทฤษฎีการตอบสนองข้อสอบไปใช้

ทฤษฎีการทดสอบแนวใหม่จึงน่าจะให้ผลการวัดที่ชัดเจนและตรงประเด็นมากขึ้น อันจะเป็นประโยชน์ต่อการพัฒนาข้อสอบและแบบสอบให้สามารถนำไปใช้ในสถานการณ์ต่างๆ ของการทดสอบได้อย่างมีประสิทธิภาพ สามารถระบุแนวโน้มของคะแนนจริง คุณภาพของแบบสอบตามเงื่อนไขของการทดสอบ รวมทั้งการประเมินความสามารถที่แท้จริงของผู้สอบ และการบรรยายพฤติกรรมกรรมการตอบสนองของผู้สอบได้เป็นอย่างดี ผู้วิจัยจึงได้สรุปประเด็นในการนำทฤษฎีการตอบสนองข้อสอบไปใช้ไว้ดังนี้

2.6.1 ใช้ในการวิเคราะห์ข้อสอบ (Item analysis) นับแต่อดีตจนถึงปัจจุบันเรานิยมใช้ทฤษฎีการทดสอบแบบเดิมวิเคราะห์คุณภาพของข้อสอบเป็นรายข้อและแบบทดสอบทั้งฉบับ แต่เนื่องจากทฤษฎีดังกล่าวมีข้อจำกัดบางประการ เช่น ค่าสถิติของข้อสอบและแบบทดสอบที่ได้จากการวิเคราะห์จะแปรเปลี่ยนไปตามลักษณะของกลุ่มผู้สอบ ในการเปรียบเทียบความสามารถของผู้สอบนั้นจำเป็นต้องใช้แบบทดสอบฉบับเดียวกันหรือแบบทดสอบคู่ขนาน นอกจากนั้นผลการวิเคราะห์ก็ไม่ได้บ่งบอกว่าผู้สอบสามารถทำอะไรได้บ้างหรือมีข้อบกพร่องในเรื่องใด เพราะได้มีการนำเอาผลการทดสอบของผู้สอบแต่ละคนไปเปรียบเทียบกับกลุ่มผู้สอบ ทำให้มองไม่เห็นในรายละเอียดย่อย

ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบซึ่ง ได้แก่ ค่าอำนาจจำแนก ค่าความยากง่ายและค่าการเดานั้นจะเป็นอิสระจากกลุ่มผู้สอบ ไม่แปรผันไปตามความสามารถของกลุ่มผู้สอบ ขณะเดียวกันความสามารถของผู้สอบก็จะเป็นอิสระจากข้อสอบ เช่นเดียวกัน การที่ผู้สอบจะตอบข้อสอบได้ถูกต้องหรือไม่นั้น ไม่ได้ขึ้นอยู่กับลักษณะของกลุ่มผู้สอบ แต่ขึ้นอยู่กับโค้งลักษณะข้อสอบ (item characteristic curve) และใช้ค่าสารสนเทศของแบบทดสอบ ซึ่งเป็นสิ่งที่แสดงความถูกต้องของการวัดแทนค่าความเชื่อมั่นของแบบทดสอบตามทฤษฎีการทดสอบแบบเดิม ทั้งนี้เนื่องจากส่วนกลับของรากที่สองของค่าสารสนเทศของแบบทดสอบก็คือ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

2.6.2 ใช้สร้างธนาคารข้อสอบ (Item bank) เนื่องจากค่าพารามิเตอร์ของข้อสอบที่ได้จากการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบนั้นจะไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ดังนั้น จึงสามารถที่จะเก็บข้อสอบพร้อมทั้งค่าพารามิเตอร์ประจำข้อไว้ใช้ในโอกาสต่อไปได้ การเก็บข้อสอบพร้อมทั้งค่าพารามิเตอร์ที่เป็นตัวบ่งชี้คุณภาพไว้อย่างเป็นระบบระเบียบสะดวกต่อการนำมาใช้ตามความต้องการ ตลอดจนสามารถเพิ่มหรือลดข้อสอบออกจากระบบไว้โดยข้อสอบที่เพิ่มเข้าไปในระบบนั้น หากถูกวิเคราะห์โดยใช้ผลการทดสอบจากกลุ่มตัวอย่างคนละกลุ่มกับข้อสอบชุดเดิมแล้ว ข้อสอบชุดใหม่ที่จะนำไปเก็บในระบบจำเป็นต้องได้รับการแปลงคะแนน (equating) ให้อยู่บนสเกลเดียวกันกับข้อสอบชุดเดิม นอกจากนั้นยังต้องมีการปรับปรุงข้อสอบให้

มีความเป็นปัจจุบันอยู่เสมอ การดำเนินการตามลักษณะที่กล่าวมาก็คือ การสร้างธนาคารข้อสอบนั่นเอง ซึ่งในปัจจุบันได้มีการนำเอาคอมพิวเตอร์เข้ามาใช้ประโยชน์ในการสร้างธนาคารข้อสอบซึ่งช่วยให้ได้รับความสะดวกเป็นอย่างยิ่งในการดำเนินการ

2.6.3 ใช้ศึกษาลักษณะการตอบแบบทดสอบ (response pattern) ในการศึกษาความถดถอยระหว่างค่าเศษเหลือของการตอบ(response residual) กับค่าความยากง่ายของข้อสอบ ถ้าเส้นถดถอยโค้งขึ้น แสดงว่าผู้สอบตอบข้อสอบด้วยการเดา ถ้าเส้นโค้งลงแสดงว่าผู้สอบตอบข้อสอบด้วยความสะเพร่า ถ้าค่าจากสมการถดถอยระหว่างค่าเศษเหลือของการตอบกับตำแหน่งของผู้สอบเป็นบวก แสดงว่า ผู้สอบตอบข้อสอบอย่างไม่เป็นระบบ ถ้าค่าเป็นลบแสดงว่าผู้สอบข้อสอบไม่ทันเวลา (Wright, 1977 อ้างถึงใน สุพัฒน์ สุขมลสันต์, 2535)

2.6.4 ใช้ตรวจสอบความลำเอียงของข้อสอบ (item bias) ทฤษฎีการตอบสนองข้อสอบมีความเชื่อว่า ในการทดสอบครั้งหนึ่งครั้งใดก็ตาม ผู้สอบที่มีความสามารถเท่ากันย่อมมีความน่าจะเป็นในการตอบข้อสอบข้อเดียวกันได้ถูกต้องเท่ากัน แต่ก็มีบางกรณีที่ผู้สอบซึ่งมีความสามารถเท่าเทียมกัน แต่มีความน่าจะเป็นในการตอบข้อเดียวกันได้ถูกต้องไม่เท่ากัน ลักษณะเช่นนี้ถือว่าข้อสอบดังกล่าวมีความลำเอียง (bias) เป็นข้อสอบที่เอื้อต่อการสอบของผู้สอบคนหนึ่งมากกว่าอีกคนหนึ่ง ในการที่ผู้สอบซึ่งมีความสามารถเท่าเทียมกันในมิติที่ถูกวัด แต่ความน่าจะเป็นในการตอบถูกต่างกัน แสดงให้เห็นว่า ข้อสอบดังกล่าวไม่ได้วัดเพียงความสามารถเดียวที่ผู้สอบมีอยู่เท่านั้น แต่ยังวัดความสามารถด้านอื่นอีกที่ผู้สอบซึ่งมีโอกาสตอบถูกสูงกว่ามีอยู่ แต่ผู้สอบที่มีความน่าจะเป็นในการตอบถูกต่ำกว่าจะไม่มี ทฤษฎีการตอบสนองข้อสอบสามารถบอกได้อย่างดีว่าข้อสอบข้อใดมีความลำเอียง แต่อย่างไรก็ตามในเรื่องของสาเหตุของความลำเอียงของข้อสอบ ตลอดจนเกณฑ์ที่เหมาะสมที่จะชี้ความลำเอียงของข้อสอบก็เป็น เรื่องที่ยังไม่มีความชัดเจนเท่าที่ควร

2.6.5 ใช้ในการปรับเทียบคะแนน (equating) ในกรณีที่มีผู้สอบกลุ่มเดียว แต่ตอบแบบทดสอบหลายฉบับที่วัดสิ่งเดียวกัน ในเวลาที่ใกล้เคียงกัน หรือผู้สอบหลายกลุ่มที่มีความสามารถต่างกันแต่ตอบแบบทดสอบที่วัดสิ่งเดียวกัน หรือผู้สอบหลายๆกลุ่มที่มีความสามารถต่างกัน และแต่ละกลุ่มของผู้สอบก็ตอบแบบทดสอบแตกต่างกันออกไป แต่แบบทดสอบดังกล่าวแต่ละฉบับมีข้อสอบที่ร่วมกันอยู่บางส่วน ทั้ง 3 สถานการณ์ที่กล่าวมานี้ หากต้องการทราบว่าผู้สอบใดมีความสามารถมากกว่า ก็สามารถทำได้โดยการปรับเทียบเทียบคะแนน

ในการปรับเทียบคะแนนนั้นจำแนกออกเป็นประเภทใหญ่ๆได้ 2 ประเภท คือ การปรับเทียบคะแนนตามแนวนอน (horizontal equating) ซึ่งเป็นการเป็นแบบทดสอบหลายฉบับที่มีค่าความยากง่ายใกล้เคียงกันกับกลุ่มผู้สอบกลุ่มเดียวกัน แล้วเทียบคะแนนของแบบทดสอบอีกชุดหนึ่ง ส่วนอีกประเภทหนึ่งนั้นเป็นการปรับเทียบคะแนนตามแนวตั้ง (vertical equating) ซึ่งเป็นการ

ใช้แบบทดสอบหลายฉบับที่มีค่าความยากง่ายต่างกัน นำไปทดสอบกับผู้สอบหลายๆกลุ่ม การปรับเทียบคะแนนก็จะเทียบตามระดับความสามารถของผู้สอบ (สงบ ลักษณะ, 2525)

2.6.6 ใช้ในการออกแบบแบบทดสอบที่ดีที่สุด (best test design) แต่เดิมนั้น ผู้สร้างแบบทดสอบไม่สามารถระบุได้ก่อนสอบว่า ผลของการทดสอบด้วยแบบทดสอบที่สร้างขึ้น จะมีความถูกต้องหรือความคลาดเคลื่อนในการวัดเพียงไร จะทราบได้ก็ต่อเมื่อการทดสอบได้ผ่านพ้นไปแล้วเท่านั้น จากประเด็นนี้ จึงนำไปสู่แนวคิดในการออกแบบแบบทดสอบที่ดีที่สุดโดยจะมีการออกแบบแบบทดสอบไว้ล่วงหน้าก่อนการทดสอบ ทั้งนี้ก็เพื่อให้แบบทดสอบดังกล่าววัดได้ถูกต้องแม่นยำ มีความคลาดเคลื่อนในการวัดน้อยที่สุด

แนวคิดที่สำคัญในการออกแบบแบบทดสอบก็คือ การจัดชุดของแบบทดสอบให้เหมาะสมกับความสามารถของผู้สอบ ดังนั้นหากผู้สร้างแบบทดสอบมีข้อสังเกตเกี่ยวกับผู้สอบละเอียดมากเท่าไร การออกแบบแบบทดสอบก็จะมีประสิทธิภาพมากเท่านั้น แบบแผนของแบบทดสอบที่ดีที่สุดจะขึ้นอยู่กับความสัมพันธ์ระหว่างลักษณะของแบบทดสอบกับลักษณะของกลุ่มผู้สอบที่ทำให้เกิดความคลาดเคลื่อนในการวัดมีค่าต่ำสุด ซึ่งไรท์และสโตน (Wright and Stone, 1979) ได้ใช้โมเดลของราส์ช (Rasch model) ที่ใช้เฉพาะค่าความยากง่ายของข้อสอบเพียงอย่างเดียว ในการออกแบบแบบทดสอบ

2.6.7 ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (adaptive testing) หลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนั้น ก็คือ การจัดข้อสอบให้มีความเหมาะสมกับความสามารถของผู้สอบ ผู้สอบที่มีความสามารถต่ำก็จะได้ข้อสอบที่ง่าย ส่วนผู้สอบที่มีความสามารถสูงก็จะได้ข้อสอบที่ยาก สมบัติในด้านความไม่แปรเปลี่ยนของค่าพารามิเตอร์ของข้อสอบไปตามกลุ่มข้อสอบตามทฤษฎีการตอบสนองของข้อสอบได้ถูกนำมาใช้ประโยชน์อย่างมากในการทดสอบลักษณะนี้ นอกจากนั้นยังสามารถกำหนดค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบในแต่ละครั้งและผู้สอบแต่ละคนได้อีกด้วย การทดสอบลักษณะนี้ใช้เวลาในการทดสอบน้อยลงกว่าเดิม ผู้สอบไม่เกิดความเบื่อหน่ายและยังมีเจตคติที่ดีต่อการสอบมากกว่าการสอบแบบเดิม

### ตอนที่ 3 ระเบียบวิธีวิจัยและการพัฒนาโปรแกรมคอมพิวเตอร์ทางการศึกษา

การวิจัยและการพัฒนา (RESEARCH & DEVELOPMENT) เป็นระเบียบวิธีการวิจัยหนึ่งที่แตกต่างกันจากระเบียบวิธีการวิจัยอื่นๆ เป็นการวิจัยที่มุ่งเน้นการคิดค้นเพื่อประดิษฐ์สิ่งศึกษามากขึ้น และนโยบายทางการศึกษาก็นำการปฏิรูปการศึกษาโดยมีนโยบายว่าโรงเรียนทุกโรงเรียนควรมีการเรียนการสอนคอมพิวเตอร์ให้กับนักเรียน มีการสร้างบทเรียนคอมพิวเตอร์ช่วยสอนออกมาจำหน่ายมาก และยังมีสื่อทางการศึกษาอีกมากที่ใช้คอมพิวเตอร์ ทำให้มีผู้สนใจศึกษา

และทำการวิจัยเกี่ยวกับการใช้คอมพิวเตอร์มาช่วยในการเรียนการสอนมากขึ้น ซึ่งจากการศึกษา งานวิจัยที่เกี่ยวข้องกับการสร้างโปรแกรมคอมพิวเตอร์เพื่อใช้ในการศึกษา และจากหนังสือการสร้างโปรแกรมที่เกี่ยวข้อง ผู้วิจัยจึงขอเสนอแนวทางในการใช้ระเบียบวิธีการวิจัยและการพัฒนา โปรแกรมคอมพิวเตอร์ทางการศึกษา ตามขั้นตอนดังต่อไปนี้ (วันพร บัณฑิต และธนาวรรณ จันทร์รัตน์ไพบูลย์, 2537 อ้างถึงใน สายชล ออบทม, 2539)

- 3.1 กำหนดปัญหา
- 3.2 วิเคราะห์ปัญหา
- 3.3 ออกแบบโปรแกรมคอมพิวเตอร์
- 3.4 เลือกภาษาคอมพิวเตอร์และเขียนโปรแกรม
- 3.5 ทดสอบโปรแกรม
- 3.6 จัดทำเอกสารประกอบการใช้โปรแกรม
- 3.7 ประเมินคุณภาพและบำรุงรักษาโปรแกรม

โดยรายละเอียดการดำเนินงานของแต่ละขั้นตอนมีดังนี้

### 3.1 กำหนดปัญหา

เป็นกระบวนการหรือขั้นตอนที่ใช้ในการระบุให้ชัดเจนว่าผู้จะทำกรวิจัย หรือ ผู้วิจัยมีความประสงค์ที่จะศึกษาอะไร เมื่อได้ปัญหาที่จะทำวิจัยแล้วผู้วิจัยแล้วต้องการกำหนดประเด็นที่จะศึกษาที่เกี่ยวข้องกับปัญหาให้ชัดเจน (พลากร กรพิทักษ์, 2532)

### 3.2 วิเคราะห์ปัญหา

เป็นขั้นตอนที่จะทำการศึกษาถึงรายละเอียดต่างๆ ของงานที่สั่งให้คอมพิวเตอร์ ทำตั้งแต่การทำความเข้าใจกับงานว่าอะไรคือสิ่งที่ต้องการให้คอมพิวเตอร์ทำจริงๆ ผลที่ได้จากคอมพิวเตอร์เป็นอะไร และถ้าต้องการผลที่ได้นั้นจะต้องมีขั้นตอนการคำนวณอย่างไร มีสูตรอะไรบ้าง ซึ่งส่วนที่สำคัญที่จะต้องวิเคราะห์และแจกแจงมีดังนี้

3.2.1 สิ่งที่ต้องการ เป็นการพิจารณาอย่างกว้างๆ ถึงงานที่ต้องการคอมพิวเตอร์ ทำเช่น ต้องการให้คำนวณคะแนนเฉลี่ย ต้องการให้พิมพ์ค่าการคำนวณที่ได้ ต้องการให้แสดงข้อคำถาม ต้องการให้สุ่มตัวเลข ต้องการให้บันทึกค่าที่คำนวณได้ ซึ่งงานแต่ละชนิดที่ต้องการให้คอมพิวเตอร์ทำควรเขียนเป็นข้อๆ ให้ชัดเจน การพิจารณาสิ่งที่ต้องการอาจดูได้จากคำสั่งหรือ ปัญหาที่จะทำว่าต้องการให้คอมพิวเตอร์ทำอะไรบ้าง

3.2.2 ผลลัพธ์ที่ต้องการแสดง เป็นการวิเคราะห์ถึงลักษณะของการรายงานหรือแบบของผลลัพธ์ที่ต้องการให้คอมพิวเตอร์แสดงผลออกมาว่าควรจะมีลักษณะอย่างไร มีรายละเอียดที่ต้องการในรายงานมากน้อยเพียงใด ปัญหาหรืองานบางอย่างอาจไม่สามารถกำหนดลักษณะของรายงานออกมาได้ชัดเจนว่าต้องการรายงานอย่างไร มีรายละเอียดอย่างไร ผู้วิจัยจะต้องทำการศึกษาระบบเดิมหรือระบบที่สร้างมีรูปแบบใดบ้างที่ต้องการออกรายงาน เพื่อความสะดวกของผู้นำผลลัพธ์ไปใช้ การวิเคราะห์ผลลัพธ์ หรือรายงานนั้นเป็นส่วนสำคัญและจะต้องพิจารณาอย่างละเอียดรอบคอบ เพราะการวิเคราะห์รายงานได้ดีนั้น จะทำให้เราทราบถึงจุดมุ่งหมายที่ต้องการให้คอมพิวเตอร์ทำ และจะได้หาวิธีที่นำไปสู่จุดมุ่งหมายนั้นได้ ซึ่งเป็นการกำหนดขอบเขตของงานที่จะทำนั่นเอง ในการวิเคราะห์ผลลัพธ์อาจวางรูปแบบออกมาอย่างคร่าวๆ เหมือนกับที่จะให้คอมพิวเตอร์แสดงผลออกมา

3.2.3 ข้อมูลที่ต้องนำเข้า เป็นขั้นตอนที่หลังจากได้ลักษณะของงานที่แน่นอนแล้วก็มาวิเคราะห์หาค่าผลลัพธ์ออกมานั้นแล้วข้อมูลที่จะนำเข้าควรมีอะไรบ้าง มีลักษณะหรือรูปแบบอย่างไร เพื่อให้ได้ผลลัพธ์ตามที่ต้องการ ซึ่งการวิเคราะห์ลักษณะข้อมูลนั้นนอกจากจะดูจากรายงานแล้ว บางครั้งต้องคำนึงถึงขั้นตอนการประมวลผลเพื่อให้ได้รายงานตามที่ต้องการด้วย เพราะค่าบางค่าในรายงานได้มากจากการคำนวณไม่ใช่เป็นค่าที่ได้มาจากข้อมูล ดังนั้นการวิเคราะห์ข้อมูลจะวิเคราะห์ว่าข้อมูลที่จะทำให้เกิดผลลัพธ์นั้นๆ ได้มาจากไหน เป็นค่าคงที่หรือเป็นค่าที่ได้มาจากการคำนวณ

3.2.4 ตัวแปรที่ใช้ในการกำหนดชื่อแทนความหมายของข้อมูลต่างๆ เพื่อความสะดวกในการอ้างอิงข้อมูลนั้น และรวมไปถึงการเขียนโปรแกรมด้วย การตั้งชื่อตัวแปรที่ใช้ในงานหรือปัญหาใดๆ ควรตั้งชื่อให้มีความหมาย และเกี่ยวข้องกับข้อมูลถ้าเป็นไปได้ควรตั้งอยู่ภายใต้กฎเกณฑ์ของภาษาคอมพิวเตอร์ที่ใช้ในการเขียนโปรแกรม โดยทั่วไป การตั้งชื่อตัวแปรจะพิจารณาความหมายของข้อมูล ว่าตรงกับคำใดในภาษาอังกฤษ แล้วนำมาดัดแปลง หรือย่อให้เข้ากับหลักเกณฑ์ของภาษาคอมพิวเตอร์ที่ใช้

3.2.5 วิธีการประมวลผล เป็นการบอกขั้นตอนของวิธีการหรือการคำนวณเพื่อให้ได้ผลลัพธ์ตามที่ต้องการ โดยเริ่มตั้งแต่การสั่งให้เครื่องรับข้อมูลแล้วนำไปประมวลผลแสดงผลลัพธ์ออกมาขั้นตอนนี้ต้องแสดงการทำงานที่ต่อเนื่องตามลำดับจึงต้องจัดลำดับก่อน-หลังให้ถูกต้อง ในขั้นตอนนี้ถ้ายังกระทำให้ละเอียดก็จะช่วยให้เขียนโปรแกรมได้ง่ายขึ้น

### 3.3 การออกแบบโปรแกรมคอมพิวเตอร์

ขั้นตอนนี้เป็น การออกแบบรายละเอียดขั้นตอนสำหรับโปรแกรม การกำหนดลำดับและความครบถ้วนของขั้นตอนของโปรแกรมทั้งโปรแกรมให้ถูกต้อง โดยไม่มีการออกแบบไว้

ล่วงหน้านั้นย่อมเป็นการเสี่ยงต่อการผิดพลาดในการเขียนโปรแกรมได้ง่ายขึ้น ดังนั้น งานในส่วนนี้เป็นการออกแบบโปรแกรมว่าควรจะต้องมีขั้นตอนตั้งแต่ต้นจนจบอะไรบ้าง มีลำดับขั้นตอนก่อน-หลังอย่างไร วิธีการในการออกแบบขั้นตอนนี้ได้มีผู้คิดไว้หลายวิธีด้วยกัน เช่น การเขียนเป็นอัลกอริทึม ซึ่งจะอธิบายการทำงานแต่ละขั้นตอนด้วยคำหรือประโยคที่สื่อความหมายให้คนเข้าใจได้ง่ายหรืออาจจะเป็นรูปประโยคคล้ายกับคำสั่งของภาษาคอมพิวเตอร์ ซึ่งเรียกว่า Pseudo code หรือการเขียนผังงาน (Flow chart) ซึ่งใช้ลักษณะที่เป็นรูปแทนความหมายต่างๆ และเพื่อให้การอ้างถึงข้อมูลและผลลัพธ์ในขั้นตอนต่างๆ ที่ออกแบบสะดวกและถูกต้องในการเขียนขั้นตอนของโปรแกรม จึงควรที่จะกำหนดชื่อหรือรายการข้อมูลและผลลัพธ์ต่างๆ ขึ้นด้วย ดังนั้นการออกแบบขั้นตอนสำหรับโปรแกรมออกจากการเขียนคำสั่ง ซึ่งแต่ละคำสั่งเหล่านี้จะต้องเขียนให้ถูกต้องตามกฎเกณฑ์ของภาษาคอมพิวเตอร์และนอกจากนี้ยังได้ประโยชน์อีกต่อหนึ่งคือการเก็บผลการออกแบบนี้ไว้จะช่วยในการทำความเข้าใจขั้นตอนของโปรแกรมในภายหลังได้สะดวกกว่าการทำความเข้าใจจากตัวโปรแกรมโดยตรงมาก

### 3.4 เลือกภาษาคอมพิวเตอร์และเขียนโปรแกรม

เป็นการเปลี่ยนขั้นตอนการทำงานต่างๆ ที่แสดงไว้ในผังงานมาเขียนให้อยู่ในรูปของภาษาคอมพิวเตอร์ภาษาใดภาษาหนึ่งตามต้องการ การจะเลือกภาษาคอมพิวเตอร์ภาษาใดนั้นขึ้นอยู่กับลักษณะและประเภทของงานนั้นๆว่าเป็นงานที่ควรใช้ภาษาใด เช่น ถ้าเป็นงานด้านวิทยาศาสตร์และคณิตศาสตร์ควรใช้ภาษาฟอร์แทน ถ้าเป็นงานด้านธุรกิจ ควรใช้ภาษาโคบอลถ้าเป็นงานประเภทที่ต้องการพิมพ์รายงานมากๆ หรือคล้ายกับการออกแบบฟอร์มควรใช้อาร์พีจี แต่ไม่ว่าจะเลือกภาษาใดก็ตาม โดยทั่วไปแล้วลักษณะของโปรแกรมที่ดีควรเป็นโปรแกรมที่แบ่งเป็นส่วนๆ ได้ชัดเจน และในแต่ละส่วนจะเขียนด้วยคำสั่งที่ไม่ยุ่งยากมากนัก นอกจากนี้ยังควรคำนึงถึงขีดจำกัดของเครื่องและตัวแปรภาษาของเครื่องคอมพิวเตอร์ (Compiler) ที่ใช้ และข้อสำคัญในการเลือกใช้โปรแกรมคือ ความถนัดและความชำนาญของผู้เขียนโปรแกรมว่าสามารถใช้ภาษาที่เลือกนั้นได้หรือไม่ และในการเขียนคำสั่งด้วยภาษาคอมพิวเตอร์นั้นยังต้องคำนึงถึงกฎเกณฑ์และหลักของภาษาที่ใช้ให้ถูกต้อง เพราะถ้ามีข้อผิดพลาดเกิดขึ้น ซึ่งเรียกว่า Syntax error โปรแกรมแปลภาษาจะไม่สามารถแปลความหมายของคำสั่งนั้นได้ ขั้นตอนปฏิบัติการโปรแกรมก็เกิดขึ้นไม่ได้ โปรแกรมแปลภาษาส่วนมากจะรายงาน หรือแสดงข้อความออกมาให้ทราบว่าการเกิดผิดพลาดตรงคำสั่งไหนเพราะอะไร เพื่อจะช่วยให้การแก้ไขข้อผิดพลาดนั้นสะดวกยิ่งขึ้น



### 3.5 ทดสอบโปรแกรม

เมื่อโปรแกรมที่เขียนผ่านการแปลภาษาและได้ผลลัพธ์เป็น Object code แล้วควรมีการทดสอบโปรแกรมก่อนที่จะทำการ Link ให้เป็น Execute file เพราะการทำงานของเครื่องคอมพิวเตอร์ตามคำสั่งของโปรแกรมนั้นไม่ได้หมายความว่าจะได้ผลลัพธ์ถูกต้องเสมอไป อาจจะมีขั้นตอนที่ไม่ถูกต้องทำให้ได้ผลลัพธ์ที่ผิดได้ ฉะนั้นเพื่อให้ได้โปรแกรมที่ใช้งานได้ถูกต้องตามจุดประสงค์ที่ตั้งไว้และมีความเชื่อถือได้ก็ควรที่จะต้องทดสอบโปรแกรมเสียก่อน วิธีการทดสอบโปรแกรมนี้กระทำได้โดยการสั่งให้เครื่องทำงานตามคำสั่งในโปรแกรมถ้าโปรแกรมนั้นมีการกำหนดให้เครื่องรับข้อมูลเข้าไปประมวลผลก็จะต้องนำข้อมูลตัวอย่างหรือข้อมูลจริงส่งไป แล้วนำผลลัพธ์ที่ได้ไปตรวจสอบกับผลลัพธ์ที่มีความถูกต้องที่ได้มาจากวิธีการอื่น เช่น ส่งข้อมูลเข้าไปประมวลผลหลายๆ ชุด ถ้าเปรียบเทียบแล้วได้ผลลัพธ์ตรงกัน ก็จะยอมรับว่าโปรแกรมนั้นใช้งานได้ แต่ถ้าผลลัพธ์ไม่ตรงกันจะต้องพิจารณาว่า ความผิดพลาดนั้นเกิดขึ้นที่ตรงไหนที่ข้อมูลหรือที่ตัวโปรแกรมถ้าผิดที่ข้อมูลก็แก้ไขข้อมูลให้ถูกต้องแล้วนำเข้าไปให้เครื่องคำนวณใหม่ถ้าผิดที่ตัวโปรแกรมก็ต้องดูว่าผิดที่ขั้นตอนไหน และทำการแก้ไขโปรแกรมแล้วทำการทดสอบโปรแกรมใหม่อีกครั้ง ข้อผิดพลาดที่เกิดจากโปรแกรม สามารถแบ่งได้ 3 แบบคือ

1) การให้รหัสผิดรูปแบบภาษา (Syntax error) ภาษาคอมพิวเตอร์แต่ละภาษามีกฎเกณฑ์การให้รหัสเฉพาะตัวที่แน่นอนในการเขียนรหัสโปรแกรม ถ้าให้รหัสผิดกฎเกณฑ์ของภาษาเครื่องก็จะให้ข่าวสารเกี่ยวกับข้อมูลผิดพลาดออกมาเพื่อให้ผู้เขียนโปรแกรมรู้ตำแหน่งที่ผิดได้ง่ายขึ้น

2) ข้อผิดพลาดที่เกิดขึ้นกับการใช้งานแต่ละครั้ง (Run-time-error) โปรแกรมที่ใส่รหัสถูกต้องไม่จำเป็นว่าจะทำงานได้สำเร็จเสมอไป ทั้งนี้อาจจะมีข้อผิดพลาดที่เกิดจากการวิ่งโปรแกรม(Running program) ได้ เช่น ในการคำนวณมีการหารค่าใดค่าหนึ่งที่ด้วยตัวแปรที่มีค่าเป็นศูนย์ หรือกรณีที่ตัวแปรที่ให้หารากที่สองมีค่าเป็นลบ ข้อผิดพลาดเช่นนี้โปรแกรมก็จะให้ข่าวสารเกี่ยวกับการผิดพลาดออกมาเพื่อให้ผู้เขียนโปรแกรมรู้และแก้ไขได้ถูกต้อง

3) ข้อผิดพลาดทางตรรกะ (Logical error) ข้อผิดพลาดทางตรรกะเกิดขึ้นเนื่องจากการกำหนดลักษณะของงานหรือการกำหนดวิธีการและขั้นตอนการทำงานผิดหรือไม่ตรงกับจุดประสงค์ของงาน ข้อผิดพลาดนี้ถ้าเกิดขึ้นเครื่องจะไม่บอกข่าวสารมาให้ทำให้ผู้เขียนโปรแกรมแก้ไขข้อผิดพลาดได้ยาก

### 3.6 จัดทำเอกสารประกอบโปรแกรม

เนื่องจากการใช้คอมพิวเตอร์ในการประมวลผลต่างๆ ส่วนมากจะเป็นการทำงานต่อเนื่องและใช้เวลายาวนาน ฉะนั้นโปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นก็จะถูกใช้ไปตามกำหนด

ระยะเวลาหนึ่ง และเมื่อใช้ไประยะหนึ่งอาจมีการเปลี่ยนแปลง ทำให้โปรแกรมนั้นอาจจะไม่เหมาะสม กับการใช้งาน ซึ่งจะต้องมีการพัฒนาโปรแกรมขึ้นมาใหม่ หรืออาจจะแก้ไขโปรแกรมให้มีความเหมาะสมกับการใช้งานมากขึ้น ถ้าได้มีการทำเอกสารประกอบการพัฒนาโปรแกรมไว้ ก็จะเป็นการสะดวกในการใช้เป็นแนวทางในการแก้ไขโปรแกรม ในการทำเอกสารนั้นควรเริ่มทำและรวบรวมตั้งแต่ขั้นตอนแรกของการพัฒนาโปรแกรมมาตามลำดับ มิฉะนั้นอาจทำให้ลืมนำได้ โดยเฉพาะการพัฒนาโปรแกรมที่มีขนาดใหญ่ซึ่งใช้เวลานาน เอกสารสำหรับโปรแกรมเขียนได้ 2 แบบดังนี้

3.6.1 เอกสารสำหรับผู้ใช้โปรแกรม ผู้ใช้โปรแกรมอาจไม่มีความรู้เกี่ยวกับโปรแกรมเลย ดังนั้นเนื้อหาของเอกสารควรประกอบด้วย

- 1) แสดงรายละเอียดขอบเขตความสามารถของโปรแกรม เช่น โปรแกรมนี้จะสามารถทำอะไรได้บ้างและมีข้อกำหนดอะไร
- 2) แสดงรายละเอียดของข้อมูลที่ใช้กับโปรแกรม
- 3) แสดงรายละเอียดของคำสั่งผลลัพธ์ที่ได้จากโปรแกรม
- 4) แสดงรายละเอียดของคำสั่งที่จะใช้ เพื่อให้โปรแกรมเริ่มต้นทำงาน
- 5) แสดงรายละเอียดของการโต้ตอบระหว่างผู้ใช้กับโปรแกรม
- 6) แสดงความหมายต่างๆ ของข่าวสารที่โปรแกรมแสดงออกมา

3.6.2 เอกสารแสดงเทคนิคที่ใช้ในโปรแกรม เอกสารแสดงเทคนิคในการใช้โปรแกรมจะมีประโยชน์มากสำหรับผู้เขียนโปรแกรมผู้อื่นที่จะดัดแปลง ปรับปรุงแก้ไขโปรแกรมเพื่อใช้งานในลักษณะที่ใกล้เคียงกัน ฉะนั้นเอกสารฉบับนี้ควรประกอบด้วยรายละเอียดของส่วนอธิบายต่าง ๆ ที่ใช้ในโปรแกรม ส่วนนี้จะแสดงให้เห็นทราบว่าโปรแกรมควรประกอบด้วย เนื้อหาของปัญหา การวิเคราะห์ปัญหา สูตรหรือทฤษฎีที่ใช้ในขั้นตอนๆ ของโปรแกรม รูปแบบของข้อมูลและผลลัพธ์ เนื้อหาโปรแกรม รายละเอียดการใช้และข้อจำกัดของโปรแกรม ตลอดจนตัวอย่างของการทำงาน of โปรแกรม

### 3.7 ประเมินคุณภาพของโปรแกรม

การประเมินคุณภาพของโปรแกรมผู้วิจัยทำการประเมิน 2 แนวทางคือ

3.7.1 ประเมินคุณภาพของโปรแกรมโดยผู้พัฒนาโปรแกรม เป็นการประเมินระบบการทำงานภายในโปรแกรม โดยประเมินด้านต่างๆ ดังนี้

- 1) ความทนทานต่อความผิดพลาดของผู้ใช้โปรแกรม โดยการทดลองทำให้เกิดข้อผิดพลาด แล้วดูว่า โปรแกรมมีวิธีป้องกันข้อผิดพลาดหรือไม่ อย่างไร

- 2) ความเชื่อถือได้ของโปรแกรม โดยการตรวจสอบว่าโปรแกรมที่สร้างขึ้นเมื่อใช้ครั้งแรกและครั้งต่อไปนั้นมีความเชื่อถือได้หรือไม่
- 3) ความเร็วของโปรแกรม เป็นการประเมินโปรแกรมในเรื่องความเร็วของโปรแกรม
- 4) ความสามารถของโปรแกรมที่สามารถบันทึกข้อมูลของผู้ใช้ในการใช้โปรแกรมแต่ละด้าน เป็นการวิเคราะห์ความสามารถของโปรแกรมที่สามารถจัดเก็บข้อมูลพื้นฐานของผู้ใช้โปรแกรมโดยอัตโนมัติ เช่น ข้อมูล เลขประจำตัว ระดับชั้น วิชาที่เรียน
- 5) ความถูกต้องในการสั่งงานตามต้องการ เป็นการประเมินโปรแกรมว่าสามารถทำงานได้ถูกต้องตามต้องการหรือไม่

3.7.2 ประเมินคุณภาพของโปรแกรมโดยผู้ใช้โปรแกรม เป็นการประเมินโปรแกรมในเรื่องผลย้อนกลับของผู้ใช้โปรแกรม ในด้านต่างๆ ดังนี้

- 1) ความชัดเจนและความสอดคล้องของคู่มือการใช้โปรแกรม
- 2) รูปแบบการใช้โปรแกรม เป็นการประเมินโปรแกรมในด้านการรับข้อมูลการดำเนินงานของโปรแกรม ความรู้พื้นฐานของผู้ใช้โปรแกรม การแสดงผล และขั้นตอนการใช้งาน
- 3) ประสิทธิภาพและประโยชน์ของโปรแกรมโดยส่วนรวม เป็นการประเมินโปรแกรมในด้านประสิทธิภาพ ผลที่ได้รับ และประโยชน์ของโปรแกรม

จะเห็นได้ว่าการพัฒนาโปรแกรมคอมพิวเตอร์ทางการทดสอบนั้น จำเป็นที่จะต้องมีระเบียบวิธีวิจัยและการพัฒนาที่ชัดเจน เพื่อความถูกต้องของผลลัพธ์ที่ได้จากการทดสอบ ซึ่งได้มาถึงประสิทธิภาพและความถูกต้องแม่นยำของการทดสอบ ในปัจจุบันประเภทและเทคนิคการใช้ รวมถึงขนาดและจำนวนการใช้คอมพิวเตอร์นั้นมีการพัฒนาอยู่ตลอดเวลา และโปรแกรมคอมพิวเตอร์ทางการทดสอบก็มีส่วนให้การทดสอบมีความคล่องตัว สะดวกและมีประสิทธิภาพสูงขึ้น ซึ่งในอนาคตอันใกล้เทคโนโลยีการทดสอบจะมีการใช้กันอย่างแพร่หลายมากขึ้น ทั้งการใช้สำหรับการผลิตข้อสอบ การพัฒนาค้างข้อสอบ การตรวจข้อสอบ การวิเคราะห์ข้อสอบ การเปรียบเทียบคะแนนระหว่างแบบสอบ การแปลความหมายของการทดสอบ การรายงานผลของการทดสอบ รวมถึงการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบอีกด้วย

#### ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ นั้น พอสรุปเป็นประเด็นต่างๆได้ ดังนี้

#### 4.1 งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

วอเตอร์ (Water, 1964) ได้เปรียบเทียบแบบทดสอบพีระมิดชนิด 5 ชั้น ใช้รูปแบบการแยกทางแบบคงที่ ระยะห่างของความยากง่ายเท่ากับ .01 แยกทางแบบถูกเพิ่ม 1 ผิดลด 1 กับแบบทดสอบแบบเดิม 4 ชุด ชุดละ 5 ข้อ ผลจากการศึกษาพบว่า คะแนนจากการทดสอบโดยใช้พีระมิดจะมีความสัมพันธ์กับความสามารถของผู้สอบสูงกว่าการทดสอบแบบเดิม และการกระจายของคะแนนมีลักษณะเป็นรูปสี่เหลี่ยมผืนผ้าและกระจายมากกว่าการทดสอบแบบเดิม

เบย์รอฟ และ ซีเลย์ (Bayroff and Seeley, 1967) ได้หาความสัมพันธ์ของผลการทดสอบโดยใช้แบบทดสอบพีระมิดและการทดสอบแบบเดิม โดยใช้แบบทดสอบพีระมิดชนิด 8 ชั้น 2 ฉบับ ใช้ผู้สอบ 102 คน ดำเนินการสอบด้วยคอมพิวเตอร์ และผู้สอบต้องตอบแบบทดสอบแบบเดิมอีก 2 ฉบับ วัดด้านภาษา 50 ข้อ ด้านจำนวน 40 ข้อ ผลจากการศึกษา พบว่า สหสัมพันธ์ของผลการทดสอบจากแบบทดสอบทั้งสองลักษณะ มีค่าเท่ากับ 0.87 และ 0.79 และเมื่อใช้ข้อสอบจำนวนเท่ากัน การทดสอบด้วยแบบทดสอบพีระมิดจะมีค่าสหสัมพันธ์สูงกว่าการทดสอบแบบเดิม

วูด (Wood, 1969) ได้ศึกษาความสัมพันธ์ระหว่างผลจากการทดสอบจากแบบทดสอบพีระมิดกับการทดสอบแบบเดิม โดยใช้แบบทดสอบพีระมิดชนิด 4 5 และ 6 ชั้น รูปแบบการแยกทางแบบถูกเพิ่ม 1 ผิดลด 1 ขนาด ชั้นคงที่ ระยะห่างของความยากง่ายเท่ากับ .05 ให้คะแนนตามจำนวนข้อที่ถูก โดยทดสอบกับกลุ่มตัวอย่าง 91 คน ผลจากการศึกษา พบว่า สหสัมพันธ์ระหว่างคะแนนรวมทั้ง 3 ฉบับ จากแบบทดสอบพีระมิดกับระดับผลการเรียนเฉลี่ยเท่ากับ .68

ลอร์ด (Lord, 1971) พบว่า เมื่อขนาดของช่วงห่าง ของความยากง่ายของข้อสอบมากขึ้น จะทำให้ค่าสารสนเทศลดลงทุกระดับความสามารถ แบบสอบจะมีประสิทธิภาพสูงสุดเมื่อมีช่วงห่างของความยากง่ายอยู่ระหว่าง 0.033 ถึง 0.067 และจะไม่มีประสิทธิภาพเมื่อใช้ช่วงห่างของความยากง่ายใกล้เคียงกับ 0.1

เออร์รี่ (Urry, 1977) ได้ศึกษาโดยใช้ข้อมูลจำลอง พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนั้น จะไม่ค่อยมีประสิทธิภาพหากเลือกใช้โมเดลการวิเคราะห์ข้อมูลที่ไม่เหมาะสม เช่น ค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบจะลดลงเมื่อเทียบกับการใช้แบบทดสอบที่มีความยาวเท่ากันแต่เลือกใช้โมเดลการวิเคราะห์ข้อสอบที่ดีกว่า โมเดลของราสช์ (Rasch) นับว่าไม่ค่อยเหมาะสมในการนำมาใช้กับการวิเคราะห์ข้อสอบแบบเลือกตอบทั้งนี้

เนื่องจากโมเดลของราล์ฟมีข้อตกลงเบื้องต้นว่าค่าอำนาจจำแนกของข้อสอบแต่ละข้อต้องคงที่ และค่าการเดาต้องมีค่าน้อยมากหรือเป็นศูนย์ ซึ่งตามสภาพความเป็นจริงแล้ว นับว่าเป็นเรื่องที่ยากมากที่จะทำให้ข้อสอบแต่ละข้อมีลักษณะตามข้อตกลงดังกล่าวที่กล่าวมา โมเดลโลจิสติก 3 พารามิเตอร์นับว่าความเหมาะสมมากกว่าโมเดลอื่นๆ รวมทั้งพบว่า คลังข้อสอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบควรมีข้อสอบอย่างน้อย 100 ข้อ และข้อสอบที่จะช่วยยกระดับคุณภาพของการวัดนั้น ควรมีค่าอำนาจจำแนกมากกว่า 0.8 ค่าความยากอยู่ในช่วง  $-2.0$  ถึง  $2.0$  และค่าการเดาน้อยกว่า 0.3

ไวส์ (Weiss, 1982) ได้ศึกษาพบว่าการประมาณค่าโดยวิธีความเป็นไปได้สูงสุด และการประมาณค่าโดยวิธีเบย์ไม่แตกต่างกันเมื่อผู้สอบมากกว่า 2,000 คน แต่การประมาณค่าโดยวิธีเบย์จะให้ผลดีกว่าเมื่อจำนวนผู้สอบน้อยกว่า 500 คน

ธอร์นไคค์ (Thronidike, 1982) พบว่าในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้แบบทดสอบโดยใช้แบบทดสอบรูปพีระมิดนั้น เมื่อจำนวนข้อเพิ่มขึ้น (ทั้งแบบขนาดข้อคงที่และขนาดข้อแปรผัน) จะทำให้ประสิทธิภาพของการประมาณค่าความสามารถของผู้สอบสูงขึ้น และหากแบบทดสอบมีจำนวนข้อมาก และข้อสอบมีค่าอำนาจจำแนกสูงขึ้นค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าจะลดลง

เวทเซล, ดักลาส และ แม็คไบรด์ (Wetzel, Douglas, and McBride, 1983) ได้ศึกษาผลของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ของข้อสอบ โดยใช้กลยุทธ์ในการเลือกข้อสอบต่างวิธีกัน ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์และการทดสอบแบบเดิมโดยศึกษาจากข้อมูลที่จำลองขึ้น โดยจำลองข้อมูลให้มีความคลาดเคลื่อนมากและน้อยกว่าความเป็นจริง ผลการศึกษา พบว่า ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ของข้อสอบมีผลน้อยมากต่อประสิทธิผล (efficiency) และความถูกต้องในการวัด (precision)

ฮัลลิน และคณะ (Hulin and others, 1983) พบว่าในการเลือกข้อสอบข้อแรกให้กับผู้สอบนั้น พบว่าลักษณะของข้อสอบข้อแรกนั้นไม่มีผลต่อความถูกต้องในการประมาณค่าความสามารถของผู้สอบเมื่อสิ้นสุดการสอบ รวมทั้งการใช้เกณฑ์การคัดเลือกข้อสอบโดยการจับคู่ระหว่างค่าความสามารถของผู้สอบกับค่า  $\theta_{max}$  เพื่อแก้ปัญหาการใช้ข้อสอบมากเกินไปนั้น พบว่ามีการสูญเสียสารสนเทศไปมากกว่าคัดเลือกข้อสอบโดยวิธีสารสนเทศสูงสุด ( $I_{max}$ ) อย่างไรก็ตามให้สารสนเทศมากกว่าการทดสอบแบบเดิม กล่าวคือถ้าความสามารถของผู้สอบต่ำกว่า –

1.25 แล้ว การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะให้สารสนเทศมากเป็น 2 เท่าของการทดสอบแบบเดิม

แฮนกินส์ (Hankins, 1986) ได้ทำการศึกษาผลของการกำหนดข้อสอบเริ่มต้นแน่นอน และกำหนดข้อสอบเริ่มต้นแปรผัน ต่อค่าความลำเอียงและสารสนเทศของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ผลการศึกษาพบว่าทั้ง 2 วิธี ไม่มีความลำเอียงในการประมาณค่าความสามารถของผู้สอบที่มีความสามารถปานกลาง แต่จะมีความลำเอียงถ้าผู้สอบมีความสามารถสูง หรือต่ำมาก ๆ ทั้ง 2 วิธี ให้ได้สารสนเทศเป็นรูปสมมาตรที่โด่ง ซึ่งแสดงให้เห็นว่าการทดสอบด้วยวิธีการกำหนดข้อสอบเริ่มต้นที่ต่างกันมีประสิทธิภาพเท่าเทียมกัน แต่ทั้ง 2 วิธี เลือกข้อสอบได้ต่างกัน การกำหนดข้อสอบเริ่มต้นแบบแปรผันใช้จำนวนข้อสอบมากกว่า

แบล็กมอร์ (Blackmore, 1986) ได้ศึกษาเปรียบเทียบการทดสอบที่ดำเนินการสอบด้วยคอมพิวเตอร์ (Computerize) การทดสอบแบบปรับเหมาะ กับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computer adaptive) และการทดสอบแบบดั้งเดิม คือ การเขียนตอบ (Paper and Pencil) โดยทดลองกับนักเรียนเกรด 12 โรงเรียนมัธยมการ์เดนโกรฟ ในรัฐแคลิฟอร์เนีย จำนวน 24 คน ในปี 1986 โดยแบ่งเป็น 3 กลุ่มๆ ละ 8 คน เพื่อดำเนินการทดสอบด้วยวิธีการทดสอบ 3 วิธีดังกล่าว การทดลองใช้แบบสอบวัดความถนัด มี 6 ฉบับย่อย คือ การให้เหตุผลทางภาษา ความสามารถทางตัวเลข การให้เหตุผลเชิงนามธรรม การให้เหตุผลเชิงกล มิติสัมพันธ์และการใช้ภาษา แต่ละกลุ่มวิเคราะห์ข้อมูลโดยใช้การทดสอบความแปรปรวนและการทดสอบแม็คเนียร์ (McNemar test) ผลการศึกษาพบว่าคะแนนจากข้อสอบย่อยทุกฉบับจากการดำเนินการสอบทั้ง 3 วิธี ไม่แตกต่างกัน แต่จะแตกต่างกันตามคุณลักษณะของผู้เข้าสอบ เช่น เชื้อชาติและเพศ จากการให้แบบสอบถามและสัมภาษณ์อย่างมีนัยสำคัญทางสถิติ

ลี (Lee, 1986) ได้ศึกษาความเป็นไปได้ของการนำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มาใช้ในการวัดผลสัมฤทธิ์ทางการเรียนโดยสร้างแบบทดสอบ เอ็ม เอ เอ ที (microcomputerized adaptive achievement test, MAAT) จากคลังข้อสอบที่สร้างโดยแผนกคลังข้อสอบของ พี อี ที เอ (pitt educational testing aids, PETA) กลุ่มตัวอย่างได้แก่นิสิตระดับบัณฑิตศึกษาจำนวน 3 ห้องเรียนที่กำลังเรียนวิจัยการศึกษาเบื้องต้น วิธีการทดสอบแบบปรับเหมาะ 2 วิธี ได้แก่ วิธีเอส พี อาร์ ที ของวอลท์ (Wald's sequential probability ratio test) และแบบเฟล็กซ์ิเลเวลของลอร์ด (Lord's flexilevel) นอกจากนั้นยังมีการวัดเจตคติของผู้สอบต่อการทดสอบ เอ็ม เอ เอ ที มาใช้วัดสัมฤทธิ์ทางการเรียน คะแนนจากการ

ทดสอบด้วยคอมพิวเตอร์มีความสัมพันธ์ทางบวกกับผลการทดสอบกลางเทอม ผู้สอบมีเจตคติที่ดีต่อการทดสอบแบบ เอ็ม เอ เอ ที่

แกรริสัน และ บวมการ์เตน (Garrison and Baumarten, 1986) ได้ทดลองใช้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เพื่อวัดทักษะคณิตศาสตร์ของนักศึกษา จำนวน 60 คน โดยมีการเปรียบเทียบความคงเส้นคงวาของค่าความสามารถของผู้สอบจากการทดสอบแบบเดิมและการทดสอบแบบปรับเหมาะกับความสามารถผู้สอบด้วยคอมพิวเตอร์ ผลจากการศึกษาพบว่า การทดสอบแบบปรับเหมาะดังกล่าวมีประสิทธิภาพในการวัดสูงกว่าการทดสอบแบบเดิม ความยาวของแบบทดสอบลดลงจากเดิม ผู้สอบมีเจตคติที่ดีต่อการสอบสูงกว่าการทดสอบแบบเดิม

อลายา และ โคช (Ayala and Koch, 1987) ได้เปรียบเทียบการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างโมเดลการตอบสนองแบบนามบัญญัติ (Nominal response model) กับโมเดลโลจิสติก 3 พารามิเตอร์ โดยศึกษาจากข้อมูลที่จำลองขึ้น ผลจากการศึกษา พบว่า ความสามารถที่ประมาณได้จากการทดสอบแบบปรับเหมาะดังกล่าวทั้ง 2 โมเดลมีความสัมพันธ์กันสูง

ไวส์ และ แม็คไบรด์ (Weiss and McBrede, 1987) ได้ศึกษาเกี่ยวกับวิธีการให้คะแนน หรือวิธีการประมาณค่าความสามารถของผู้สอบตามวิธีการของเบย์ที่ปรับใหม่ (bayesian updating) พบว่า เป็นวิธีการประมาณค่าที่ไม่ลำเอียง รวมทั้งความสามารถจริงกับความสามารถที่ได้จากการประมาณค่าตามวิธีของเบย์มีความสัมพันธ์กันสูงมาก

โฮ (Ho, 1988) ได้เปรียบเทียบวิธีการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ 3 วิธี คือ ประมาณค่าความสามารถของผู้สอบโดยวิธีของเบย์ (bayesian) วิธีโมดัลเบย์ (modal-bayesian) และวิธีความเป็นไปได้สูงสุด (maximum likelihood) โดยใช้โปรแกรมในการทดสอบนี้กับคลังข้อสอบ 3 ขนาด คือ 86, 71 และ 56 ข้อ ลักษณะของคลังข้อสอบมี 3 แบบคือ ข้อสอบได้มาโดยการสุ่ม ข้อสอบที่มีค่าอำนาจสูงๆ และข้อสอบที่ง่ายๆ ทดลองกับผู้สอบที่มีความสามารถ 3 ระดับ คือ สูง กลาง และต่ำ จุดเริ่มต้นในการทดสอบนั้นใช้จุดเริ่มต้นเดียวกัน วิเคราะห์ข้อมูลโดยการวิเคราะห์ความแปรปรวน 3 ทาง แบบวัดซ้ำ ซึ่งตัวแปรตาม ได้แก่ จำนวนข้อสอบที่ใช้ความแปรปรวนของความคลาดเคลื่อนเฉลี่ย และค่าสัมบูรณ์ของความแตกต่างระหว่างค่าความสามารถโดยประมาณ ซึ่งตัวแปรตาม 2 ตัวแรกนั้นใช้สำหรับประเมินประสิทธิภาพของการทดสอบ ตัวที่ 3 ใช้ประเมินความถูกต้องของการทดสอบ ผลการศึกษาพบว่า

- 1) วิธีการประมาณค่าแบบโมดัลเบย์มีประสิทธิภาพสูงสุด วิธีการประมาณค่าของเบย์มีความเชื่อมั่นสูง แต่วิธีการประมาณค่าความเป็นไปได้สูงสุดไม่ค่อยคงเส้นคงวา
- 2) ขนาดของคลังข้อสอบมีผลน้อยมาก ยกเว้นกลุ่มผู้สอบที่มีความสามารถต่ำ และใช้ข้อสอบเริ่มต้นที่มีค่าความยากง่ายสูงกว่าความสามารถจริง
- 3) การประมาณค่าความสามารถจะมีความถูกต้องมากยิ่งขึ้น ถ้าข้อสอบที่ใช้เริ่มต้นสอบนั้นมีค่าความยากง่ายเท่ากับหรือต่ำกว่าความสามารถจริงของผู้สอบ
- 4) มีปฏิสัมพันธ์ระหว่างชนิดของคลังข้อสอบกับวิธีการประมาณค่าความสามารถและชนิดของคลังข้อสอบกับระดับความสามารถจริงของผู้สอบ
- 5) มีความแตกต่างกันน้อยมากระหว่างคุณลักษณะทางสถิติของข้อสอบที่ไม่เคยถูกเลือกใช้กับข้อสอบที่ถูกเลือกใช้บ่อย

สเกจจ์ และ สตีเวนสัน (Skaggs and Stevenson, 1989) ได้ศึกษาเปรียบเทียบผลของการประมาณค่าพารามิเตอร์ต่างๆ ด้วยวิธีของเบส์ (pseudo-batesian) กับวิธีความเป็นไปได้สูงสุด (joint maximum likelihood) โดยใช้โมเดล การวิเคราะห์แบบ 3 พารามิเตอร์วิเคราะห์ข้อมูล จากผลการศึกษา พบว่า ถ้าใช้กลุ่มตัวอย่าง ขนาดใหญ่ ตั้งแต่ 2000 คนขึ้นไป และข้อสอบตั้งแต่ 35 ข้อขึ้นไปแล้ว การประมาณค่าด้วย วิธีของเบส์ จะให้ความคงเส้นคงวา (consistency) และความถูกต้องแม่นยำ (accuracy) มากกว่า แต่ถ้าใช้ข้อสอบน้อยกว่า 15 ข้อแล้วการประมาณค่าด้วยวิธีความเป็นไปได้สูงสุดจะให้ความถูกต้องแม่นยำมากกว่า

ชาง (Chang, 1990) ได้ศึกษาวิธีการแยกทางในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ กรณีที่การทดสอบดังกล่าวมีหลายเนื้อหา ซึ่งอาจเป็นไปได้ที่แบบทดสอบดังกล่าวจะมีหลายมิติ ดังนั้นจึงได้จัดทำแบบทดสอบเป็นฉบับย่อย ๆ และได้เปรียบเทียบวิธีการแยกทางไปสู่แบบทดสอบย่อย 4 วิธี คือ วิธีใช้ค่าสหสัมพันธ์พหุคูณของคะแนนดิบชุดที่ให้ค่าสหสัมพันธ์ที่ต่ำกว่าจะได้รับ การคัดเลือก วิธีคัดเลือกแบบทดสอบย่อยด้วยการสุ่ม วิธีสุดท้ายคือ ไม่มีการแยกทางซึ่งใช้เป็นฐานในการเปรียบเทียบกับวิธีอื่นๆ กลุ่มตัวอย่างได้มาโดยการจำลองข้อมูล 1000 คน สำหรับการวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบ และ 200 คน สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ใช้แบบทดสอบ 4 ชุดย่อย แต่ละชุดย่อยมีข้อสอบ 100 ข้อ แบบทดสอบทั้ง 4 ชุดย่อยมีสหสัมพันธ์อยู่ในช่วง .14 ถึง .69 ผลการศึกษาพบว่า วิธีที่ใช้ค่าสหสัมพันธ์สูงจะให้ผลดีกว่าวิธีอื่นๆ โดยเฉพาะการทดสอบขั้นสุดท้ายที่แบบทดสอบชุดย่อยที่ชุด 3 และชุดที่ 4 ถูกดำเนินการสอบ อย่างไรก็ตามความแตกต่างระหว่างการแยกทางทั้ง 4 วิธีก็ไม่เด่นชัดพอที่จะอ้างถึงในการนำไปใช้ จำนวนชุดย่อยของข้อสอบและค่าสหสัมพันธ์ระหว่างชุดย่อยของข้อสอบควรเป็นตัวแปรที่จะต้องศึกษาต่อไป



อโยลา และคณะ (Alaya and others, 1990) ได้เปรียบเทียบการทดสอบโดยใช้แบบทดสอบเพิลิกซีเลเวลกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ประมาณค่าความสามารถด้วยวิธีของเบส์ ใช้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถเป็นเกณฑ์ยุติการสอบ โดยเปรียบเทียบที่จุดต่างๆ กันของค่าของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ผลจากการศึกษา พบว่า ค่าความสามารถของผู้สอบที่ได้จากการทดสอบสองวิธีมีความถูกต้องแม่นยำไม่ต่างกัน

เบอร์กสตรอม และ สตาห์ล (Bergstrom and Stahl, 1992) ได้ศึกษาขนาดที่พอเหมาะของคลังข้อสอบที่จะทำให้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีประสิทธิภาพที่ระดับต่างๆ ของความสามารถของผู้สอบ ตัวแปรที่ศึกษา คือ ลักษณะเฉพาะของเนื้อหา ความยาวของแบบทดสอบและเกณฑ์ยุติการสอบ ผลจากการศึกษาพบว่าขนาดของคลังข้อสอบที่พอเพียงนั้นขึ้นอยู่กับเกณฑ์ยุติการสอบ

คาร์ลสัน (Carlson, 1993) ได้ศึกษาประสิทธิภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้วิธีการเลือกข้อสอบต่างกัน 4 วิธี เกณฑ์ยุติการสอบ 2 เกณฑ์คือ ตอบข้อสอบครบ 20 ข้อและค่าความสามารถของผู้สอบโดยใช้กลยุทธ์ GSSS ของเซียวโอ (Xiao's golden section search strategy) ดำเนินการสอบโดยใช้โปรแกรมที่เขียนโดยภาษาฟอร์แทรน วิเคราะห์ข้อมูลโดยใช้สถิติสหสัมพันธ์ของเพียร์สันและเคนดอลล์เทาบี (pearson and kendall tau-b correlation) ผลจากการศึกษา พบว่า ผลจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับคะแนนรวมทั้งฉบับมีความสัมพันธ์กันสูงมาก วิธีการเลือกข้อสอบจากคลังข้อสอบทั้ง 4 วิธีให้ผลไม่ต่างกัน อย่างไรก็ตามการเลือกข้อสอบด้วยวิธีใช้ค่าสารสนเทศของข้อสอบที่สูงสุด (maximum information) ดูเหมือนจะดีกว่าวิธีอื่นๆ

ลินซ์ และ เบอร์กสตรอม (Lunz and Bergstrom, 1994) ได้ศึกษารูปแบบของการดำเนินการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ต่อผลการสอบรูปแบบดังกล่าว ได้แก่ ค่าความยากง่ายของข้อสอบเริ่มต้นต่างกัน 3 ระดับ (ง่าย ปานกลาง ยาก) เป้าหมายของความยากง่ายของแบบทดสอบต่างกัน (โอกาสในการตอบถูกร้อยละ 50 60 70) ความยาวของแบบทดสอบ (50, 100, 240 ข้อ) และโอกาสในการควบคุมแบบทดสอบ (ข้ามข้อ ทบทวนข้อเดิม, หน่วงเวลาในการตอบ และไม่เงื่อนไข) กลุ่มตัวอย่าง ได้แก่ นิสิตเทคนิคการแพทย์ วิเคราะห์ข้อมูลด้วยสถิติการวิเคราะห์ความแปรปรวนร่วม โดยใช้ผลการทดสอบแบบเดิมเป็นตัวแปรร่วม ผลจากการศึกษา พบว่าผลการทดลองจะต่างกันเมื่อโอกาสในการควบคุมแบบทดสอบต่างกัน ปฏิสัมพันธ์ของตัวแปรอิสระทั้งหลายไม่ส่งผลต่อตัวแปรตาม

สต็อกกิ้ง และ เลวิส (M. Stocking and C. Lewis 1995) พบว่าวิธีการแบบมีเงื่อนไขเป็นวิธีการที่ดีที่สุด มีความถูกต้องแม่นยำของการวัดจากผลงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบทั้งในประเทศและต่างประเทศ พบว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีคุณภาพไม่แตกต่างกับการทดสอบแบบประเพณีนิยม แต่ใช้จำนวนข้อสอบและเวลาในการทดสอบน้อยกว่า จะเห็นได้ว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีประโยชน์ในด้านอื่นๆอีก เช่น ทำให้ผู้สอบมีเจตคติที่ดี รวมทั้งให้ผู้สอบมีแรงจูงใจ ความตั้งใจ และความสนใจในการสอบมากยิ่งขึ้น และมีความเป็นไปได้ที่จะนำเอาการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ หรือการทดสอบแบบซีไอที (CAT) มาใช้แทนการทดสอบแบบประเพณีนิยม

คาลอห์น และสเปร์ (C. Kalohn and A. Spray, 1998) ได้ทำการศึกษาผลกระทบของการคัดเลือกข้อสอบต่ออัตราการใช้ข้อสอบซ้ำในการทดสอบด้วยคอมพิวเตอร์ โดยมีสมมติฐานการวิจัย คือ อัตราการใช้ข้อสอบซ้ำมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แต่ผลกระทบของความแตกต่างมีเล็กน้อย การศึกษานี้เป็นการจำลองข้อสอบ คลังข้อสอบมีจำนวนข้อสอบ 1,235 ข้อ อัตราการใช้ข้อสอบซ้ำจะถูกคำนวณภายใต้เงื่อนไข 2 เงื่อนไข คือ อัตราการใช้ข้อสอบซ้ำ 20% และ 10% ผลการศึกษาพบว่า วิธีการคัดเลือกข้อสอบทั้งสองเงื่อนไขจะมีความถูกต้องเหมือนกัน แต่ถ้าข้อสอบในคลังข้อสอบมีจำนวนมากซึ่งมีสารสนเทศน้อย ควรใช้อัตราการใช้ข้อสอบซ้ำ 10%

จาไวเออร์ และไวเซนท์ (Javier and Vicent, 1998) ได้ทำการเปรียบเทียบวิธีการควบคุมอัตราการใช้ข้อสอบซ้ำในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ โดยใช้คอมพิวเตอร์ ผลการศึกษาพบว่า Restricted Method ใช้ในการลดอัตราการใช้ข้อสอบซ้ำได้สูงสุด และ Progressive Method ลดจำนวนข้อสอบภายในคลังข้อสอบที่ยังไม่ได้ใช้ ซึ่งการควบคุมอัตราการใช้ข้อสอบซ้ำมีจุดมุ่งหมายที่สำคัญ 2 ประการ คือ เพื่อป้องกันอัตราการใช้ข้อสอบซ้ำที่มากเกินไป และ เพิ่มอัตราการใช้ข้อสอบที่ยังไม่ได้ใช้ภายในคลังข้อสอบ

งามนิศย์ ธาตุทอง และ ปรีชา เครือวัลย์ (2527) พบว่าผู้สอบส่วนใหญ่มีเจตคติที่ดีต่อการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ โดยผู้สอบมีเจตคติต่อการทดสอบแบบเดิมแต่ใช้คอมพิวเตอร์ (computerized adaptive testing - CAT) และการทดสอบแบบเดิมต่างกัน

จิราพร ไกรสรศิวเวท (2529) ได้ศึกษาความสัมพันธ์ระหว่างรูปแบบบางประการของการทดสอบเทเลอรรูปปริมาตรกับความสามารถทางการเรียน โดยใช้การทดสอบเทเลอรรูปปริมาตรที่มี 8 ชั้น 3 แบบ คือ แบบที่ใช้ขนาดชั้นคงที่ ขนาดชั้นแปรผัน และแบบข้างตัด และใช้วิธีการให้คะแนน 3 วิธี คือ ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ได้ทำ ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก และให้คะแนนทุกข้อ โดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 348 คน ผลการศึกษาพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการทดสอบเทเลอรรูปปริมาตรทั้ง 3 แบบ โดยให้คะแนนทั้ง 3 วิธี กับคะแนนจากการทดสอบแบบดั้งเดิม มีความสัมพันธ์กันในทางบวกอย่างมีนัยสำคัญทางสถิติทุกค่า และเมื่อพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ภายใน ระหว่างคะแนนที่ได้จากการทดสอบเทเลอรรูปปริมาตรทั้ง 3 แบบ ให้คะแนนทั้ง 3 วิธี ปรากฏว่ามีค่าสหสัมพันธ์เป็นบวกอย่างมีนัยสำคัญทางสถิติทุกค่า

นันทิยา พึ่งคำ (2531) ได้ทำการเปรียบเทียบคุณภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับการทดสอบแบบเดิม กลุ่มตัวอย่างได้แก่นักเรียนชั้นม.3 จำนวน 120 คน แบบทดสอบแบบเดิมมีข้อสอบ 40 ข้อ ส่วนข้อสอบในคลังข้อสอบที่ใช้สำหรับการทดสอบแบบปรับเหมาะนั้นมี 361 ข้อ ข้อสอบทั้งสองลักษณะวัดความสามารถด้านคำศัพท์ภาษาอังกฤษ ผลการศึกษาพบว่าในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่าและใช้ข้อสอบน้อยกว่าการทดสอบแบบเดิม แต่ยังให้ความตรงในการประมาณค่าความสามารถเหมือนเดิม

เสรี เปรมชวลิตโรจน์ (2531) ได้ศึกษาความตรงในการจำแนกความรู้ และการประมาณค่าความสามารถในวิชาคณิตศาสตร์ โดยใช้แบบสอบเฟล็กซีเลเวลจำนวน 2 ฉบับ และแบบสอบรูปปริมาตรขนาดชั้นแปรผันจำนวน 2 ฉบับ โดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 5,987 คน ผลการศึกษาพบว่าความตรงในการจำแนกความรู้เมื่อใช้ข้อสอบที่สร้างขึ้นเป็นเกณฑ์ ปรากฏว่า ข้อสอบแบบเฟล็กซีเลเวล และรูปปริมาตรแบบขนาดชั้นแปรผันทุกฉบับ มีความตรงในการจำแนกความรู้ได้ และเมื่อเปรียบเทียบค่าความตรงแต่ละแบบ พบว่า แบบสอบแบบเฟล็กซีเลเวลทั้งสองฉบับ มีความตรงในการจำแนกความรู้ได้สูงกว่าการทดสอบรูปปริมาตรอยู่หนึ่งฉบับ

เกรียงศักดิ์ สุวรรณภาค (2532) ได้ศึกษาความสัมพันธ์ระหว่างรูปแบบของแบบทดสอบแยกกลุ่ม กับความสามารถทางการเรียนในวิชาคณิตศาสตร์โดยทำการศึกษาในตัวแปร 2 ตัว คือ รูปแบบของแบบทดสอบ และวิธีการให้คะแนน โดยทำการศึกษากับนักเรียนชั้น

มัธยมศึกษาปีที่ 3 จำนวน 2,455 คน ผลการศึกษาพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการสอบที่ใช้แบบทดสอบแยกกลุ่มแต่ละฉบับ ซึ่งตรวจให้คะแนนในแต่ละวิธีกับความสามารถทางการเรียน มีค่าเป็นบวกอย่างมีนัยสำคัญทางสถิติทุกค่าและผลการทดสอบความแตกต่างของค่าสหสัมพันธ์ดังกล่าวเป็นรายคู่ พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติทุกคู่ ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจให้คะแนนทั้ง 2 วิธีดังกล่าว โดยแยกวิเคราะห์ตามแบบทดสอบแยกกลุ่มแต่ละแบบ มีค่าเป็นบวกอย่างมีนัยสำคัญทางสถิติทุกค่า และผลการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ดังกล่าว พบว่า เมื่อพิจารณาในแบบทดสอบแยกกลุ่มแต่ละแบบ ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจให้คะแนนทั้ง 2 วิธีมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติทุกคู่

ต่าย เชียงฉี (2534) ได้ทำการเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบเทเลอรูปปริมาตรที่มีรูปแบบ จำนวนชั้น และวิธีการให้คะแนนที่แตกต่างกัน โดยใช้วิธีมอนติคาร์โล ผลการศึกษาพบว่าการประมาณค่าความสามารถตามวิธีการของเบส์(bayesian) และวิธีความเป็นไปได้สูงสุด แบบมีเงื่อนไข(condition maximum likelihood) นั้นให้ผลการประมาณที่ไม่แตกต่างกัน การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบรูปพีระมิดที่มีรูปแบบ จำนวนชั้นและวิธีการให้คะแนนที่ต่างกันจะส่งผลต่อความคลาดเคลื่อนมาตรฐานในการประมาณค่า นอกจากนี้ยังพบว่าค่าความตรงเชิงโครงสร้างของแบบสอบพีระมิดแบบขนาดชั้นคงที่ที่มี 16 ชั้น ประมาณค่าความสามารถด้วยวิธีของเบส์มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงสุด

ตริงใจ พูลผลอำนวย (2534) ได้พัฒนาแบบสอบเฉพาะบุคคลในวิชาวิทยาศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้คอมพิวเตอร์ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 4,492 คน ผลการศึกษาพบว่าแบบสอบเฉพาะบุคคลที่ทดสอบโดยใช้คอมพิวเตอร์ และผู้สอบแต่ละคนทำ 10 ข้อ มีความสัมพันธ์กันแบบสอบผลสัมฤทธิ์วิชาวิทยาศาสตร์ทั้งฉบับ 55 ข้อ ซึ่งใช้การเขียนตอบ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.7818 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบสอบเฉพาะบุคคลเท่ากับ 0.443

สายชล อบทม (2538) ได้พัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ที่ใช้รูปแบบทางแยกแปรผัน การประมาณค่าความสามารถของผู้สอบโดยวิธีของเบส์ ยุติการทดสอบเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า

0.3 ประกอบด้วยโปรแกรมย่อย 2 โปรแกรม คือ โปรแกรมย่อยการพิมพ์ข้อสอบและโปรแกรมย่อยการทดสอบที่ปฏิบัติการบนระบบวินโดว์ 95

รังสรรค์ มณีเล็ก (2540) ทำการศึกษาผลของความยากง่ายของข้อสอบข้อแรกที่ใช้ในการทดสอบ เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ วิธีการประมาณค่าความสามารถของผู้สอบเกณฑ์ยุติการทดสอบและความสามารถของผู้สอบต่อความตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ผลการศึกษาพบว่าวิธีการประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการทดสอบ และความสามารถของผู้สอบมีผลต่อค่าความตรงเชิงสภาพของการทดสอบ และความยากง่ายของข้อสอบข้อแรก เกณฑ์ยุติการทดสอบและความสามารถของผู้สอบต่างก็มีผลต่อจำนวนข้อสอบที่ใช้ในการทดสอบ

กนกวรรณ รัตนธน (2544) ทำการเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน พบว่าผู้สอบที่มีระดับความสามารถและอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น จะส่งผลให้ความตรงตามสภาพของการทดสอบและความคงที่ในการประมาณค่าลดลง อีกทั้งผู้สอบที่มีความสามารถสูงขึ้นจะทำให้แนวโน้มของการประมาณค่าความสามารถสูงขึ้นด้วย รวมถึงเมื่อจำนวนครั้งที่ทำการทดสอบเพิ่มขึ้นและมีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น ก็จะทำให้แนวโน้มของการประมาณค่าความสามารถสูงขึ้นด้วย

ปัจจุบันได้มีการพัฒนามาเป็นการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (computerized adaptive testing - CAT) ซึ่งเป็นการใช้คอมพิวเตอร์ช่วยบริหารการสอบทุกขั้นตอน ผู้สอบจะทำแบบสอบที่ออกแบบให้เหมาะสมกับระดับความสามารถของแต่ละคน (Hambleton & Sireci, 1997) โดยหลักการแล้วในการเลือกข้อสอบข้อแรก โปรแกรมคอมพิวเตอร์จะเลือกข้อสอบที่ยากง่ายปานกลางจากคลังข้อสอบ ในกรณีที่ผู้สอบทำได้โปรแกรมจะเลือกข้อสอบข้อถัดไปที่ยากขึ้น หรือในกรณีที่ผู้สอบทำไม่ได้ โปรแกรมจะเลือกข้อสอบข้อถัดที่ง่ายลง โดยที่ผู้สอบไม่สามารถที่จะกลับไปเปลี่ยนคำตอบข้อก่อนหน้าได้ กระบวนการนี้จะดำเนินต่อไปเรื่อย ๆ จนกระทั่งสามารถประมาณความสามารถของผู้สอบแต่ละคนได้แล้ว โปรแกรมคอมพิวเตอร์จึงหยุดทำงาน รวมถึงการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีประโยชน์ต่อการวัดและประเมินผลทางการศึกษาในหลายด้าน เช่น ทำให้การทดสอบมีความคล่องตัว สามารถทดสอบเป็นรายบุคคลตามความพร้อมของผู้สอบ มีการตรวจให้คะแนน และ

รายงานผลเป็นไปอย่างอัตโนมัติ ช่วยเพิ่มคุณภาพและประสิทธิภาพของการวัดความสามารถ/ การเรียนรู้ของผู้สอบได้อย่างรวดเร็วและมีความถูกต้องยิ่งขึ้น การสร้างแบบทดสอบแบบปรับเหมาะมีองค์ประกอบที่มีผลต่อประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบได้แก่ มีคลังข้อสอบขนาดใหญ่ มีวิธีการคัดเลือกข้อสอบและตรวจให้คะแนนอย่างมีประสิทธิภาพ เพื่อให้การทดสอบแบบปรับเหมาะเป็นไปอย่างมีประสิทธิภาพสูง ควรใช้กลยุทธ์แบบแยกทางแปรผัน บนหมวดของทฤษฎีการตอบสนองข้อสอบ (IRT) ที่ใช้คอมพิวเตอร์ในการทดสอบ ควรใช้วิธีการคัดเลือกข้อสอบแบบที่ให้สารสนเทศสูงที่สุด (Maximum Information Strategy) และประมาณค่าความสามารถที่แท้จริง ( $\theta$ ) โดยใช้วิธี Bayesian Estimation เนื่องจากมีความตรงเชิงสภาพสูงกว่าการประมาณค่าแบบวิธี Maximum Likelihood Estimation รวมทั้งมีเกณฑ์การยุติที่เหมาะสม ดังนั้นในการนำเอาการทดสอบแบบปรับเหมาะไปใช้จึงจำเป็นที่จะต้องคำนึงถึงสิ่งเหล่านี้เป็นสำคัญ เพื่อให้ได้แบบทดสอบที่มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบ ในด้านรูปแบบต้องเป็นรูปแบบที่ง่ายในการนำไปปฏิบัติ คือ ง่ายต่อการใช้ ง่ายต่อความเข้าใจของทั้งผู้ใช้และผู้สอบ ในด้านจำนวนข้อสอบต้องคำนึงถึงความยาวของแบบทดสอบ ถ้ามากเกินไปอาจทำให้เสียเวลา เสียค่าใช้จ่ายและทำให้เกิดความเบื่อหน่ายแก่ผู้สอบได้ แต่ถ้าข้อสอบน้อยเกินไป อาจทำให้ได้ข้อมูลในการประมาณค่าความสามารถที่คลาดเคลื่อนสูง ดังนั้นจึงควรใช้จำนวนข้อสอบที่เหมาะสม คือจำนวนน้อยแต่ให้ความตรงและความเที่ยงสูง ทั้งยังต้องประมาณค่าความสามารถได้อย่างถูกต้องแม่นยำ ในด้านวิธีให้คะแนนต้องไม่มีความลำเอียงและสามารถประมาณค่าความสามารถของผู้สอบได้ตรงกับความสามารถที่แท้จริงของผู้สอบมากที่สุด

#### 4.2 งานวิจัยที่เกี่ยวข้องกับการเปิดโอกาสให้มีการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ลูนซ์, เบิร์กสตรอม และไรท์ (Lunz, Bergstrom, and Wright, 1992) ได้ทำการศึกษาผลกระทบจากการอนุญาตให้ผู้สอบสามารถทวนคำตอบและเปลี่ยนคำตอบได้โดยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยทำการสุ่มกลุ่มตัวอย่างนักเรียนให้ได้รับอนุญาตให้ทวนคำตอบได้ โดยมีคำชี้แจงในแบบสอบถามว่าจะต้องตอบคำถามทุกข้อให้เสร็จหลังจากนั้นจึงสามารถทวนคำตอบและเปลี่ยนคำตอบได้ โดยกลุ่มตัวอย่างที่ไม่ถูกสุ่มจะไม่ได้รับอนุญาตให้ทวนคำตอบ ภายใต้เงื่อนไขดังกล่าวพบว่า ค่าประมาณความสามารถของผู้สอบก่อนและหลังทวนมีค่าสหสัมพันธ์ .98 ค่าเฉลี่ยประสิทธิภาพของแบบสอบลดลง 1% หลังการทวนคำตอบ ผู้สอบประมาณ 32% ได้รับค่าประมาณความสามารถสูงขึ้นแต่ไม่มีผลต่อการเปลี่ยนสถานะการสอบผ่านหรือไม่ผ่าน โดยกลุ่มที่อนุญาตให้มีการทวนได้คะแนนสูงกว่ากลุ่มที่ไม่ได้รับอนุญาตให้มีการทวนเล็กน้อย

ลูนซ์, สทาห์ล และเบอร์กส์ตรอม (Lunz, Stahl, and Bergstrom, 1993) ได้ทำการประเมินผลกระทบจากการอนุญาตให้ผู้สอบสามารถทวนคำตอบและเปลี่ยนคำตอบได้โดยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยการทวนจะสามารถทำได้หลังจากการดำเนินการสอบเสร็จ และคำนวณหาค่าความสามารถของผู้สอบ, ความแน่นอนของการวัดผลและความเที่ยงของแบบสอบในแต่ละรายทั้งก่อนและหลังการทวนคำตอบ พบว่าค่าประมาณความสามารถของผู้สอบเพิ่มขึ้นเล็กน้อยหลังการทวนคำตอบ และค่าสหสัมพันธ์ระหว่างก่อนทวนและหลังทวนเท่ากับ .99 ความแน่นอนของการวัดผลขึ้นอยู่กับจำนวนข้อที่เปลี่ยนคำตอบ, แนวทางของการเปลี่ยนคำตอบและความสามารถของผู้สอบ ส่วนความเที่ยงของแบบสอบไม่แตกต่างกัน

เวเนอร์ (1993) กล่าวว่า ที่ไม่ควรให้มีการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เพราะ จะสูญเสียความมีประสิทธิภาพ และเกิดความลำเอียงในการประมาณค่าความสามารถของผู้สอบ รวมทั้งเกิดความคลาดเคลื่อนในการได้มาถึงความสามารถที่แท้จริงของผู้สอบในกระบวนการทดสอบ ทำให้ขัดแย้งกับกระบวนการสร้างแบบทดสอบแบบปรับเหมาะที่มีประสิทธิภาพ โดย Wainer ได้ยกตัวอย่างสถานการณ์ที่ทำให้ผู้สอบได้คะแนนสูงขึ้น กล่าวคือ เมื่อผู้สอบตั้งใจที่จะตอบคำถามผิด ดังนั้นจึงทำให้ได้รับข้อสอบข้อต่อไปที่มีระดับความยากต่ำลง และถ้าเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้โดยไม่มีเงื่อนไข ผู้สอบก็จะสามารถกลับไปแก้คำตอบข้อก่อนหน้า ทำให้ข้อสอบแต่ละข้อนั้นไม่สัมพันธ์กับระดับความสามารถที่แท้จริงของผู้สอบ ผู้สอบจะได้รับการประมาณค่าความสามารถโน้มไปทางด้านบวกหรือได้รับคะแนนที่สูงเกินความสามารถที่แท้จริง และมีความคลาดเคลื่อนสูง

วอลเตอร์ พี. วิสโปเอล, โทมัส อาร์. ร็อคคลิน, เทียนเหยาหวัง และ ทิมโมธี เบลเลอร์ (Walter P. Vispoel, Thomas R. Rocklin, Tianyou Wang, and Timothy Bleiler, 1996) ได้ทำการศึกษาผลกระทบจากการอนุญาตให้ผู้สอบสามารถกลับไปทวนคำตอบและเปลี่ยนคำตอบได้ พบว่าผู้สอบจะได้รับการประเมินค่าความสามารถที่สูงกว่าความสามารถที่แท้จริง โดยอ้างหลักเกณฑ์ของไวเนอร์ในกรณีที่ผู้สอบตั้งใจตอบคำถามผิดก่อนทวนและทำตามความสามารถจริงระหว่างการทวน ผลลัพธ์จากการจำลองสถานการณ์ปรากฏว่า มีบางกรณีที่ประเมินค่าความสามารถได้สูงขึ้นและบางกรณีที่ประมาณค่าความสามารถได้ต่ำลงพอๆ กัน ดังนั้นประเด็นค่าความสามารถที่เพิ่มขึ้นได้นั้นอาจขึ้นอยู่กับกระบวนการที่ใช้ในการประเมินค่าความสามารถ (maximum likelihood หรือ Bayesian), ค่าความสามารถที่แท้จริงของผู้สอบ, ค่าเบี่ยงเบนมาตรฐานของการประมาณค่าความสามารถ, ความสามารถในการใช้แผนการนี้ของผู้เข้าสอบและประเภทการตัดสินใจในการประเมินค่าความสามารถ

ไวส์, เอสแอล (Wise, S. L., 1996) ได้ทำการศึกษาค้นคว้าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่อนุญาตให้มีการทวนคำตอบ พบว่าเกิดผลกระทบ 5 ด้านดังนี้ โอกาสของความเป็นอิสระของข้อสอบที่กระทบต่ออีกคำตอบหนึ่ง, ประสิทธิภาพของแบบสอบที่ลดลง, ผลลัพธ์ของแบบสอบที่กระทบต่อแบบแผนการดำเนินการสอบ, เพิ่มเวลาในการสอบ และต้องทำการพัฒนาแบบสอบที่ซับซ้อนมากยิ่งขึ้น

แฟร์รารา สตีเว่น, อัลเบิร์ต ฟรานเชส, กิลมาร์ติน เดโบราห์, นอตต์ เทเรซ่า และ ไมเคิล (Ferrara Steven, Albert Frances, Gilmartin Deborah, Knott Teresa, and Michaels, 1996) ได้ศึกษาถึงประโยชน์ของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบอนุญาตให้ทวนคำตอบโดยพิจารณาอย่างลึกซึ้งถึงข้อมูลของผู้สอบในการทวนที่จะตัดสินใจเลือกคำตอบเดิมหรือเปลี่ยนคำตอบ โดยวิธีการนี้ได้นำมาใช้กับนักเรียน 29 คนในชั้นมัธยมศึกษาปีที่ 1 และ 2 ผลการวิจัยพบว่าผู้เข้าสอบมีแนวโน้มที่จะได้รับการประเมินค่าความสามารถที่สูงขึ้นเกินความเป็นจริงและทำให้สถานการณ์สอบเปลี่ยนแปลง (สอบได้หรือสอบตก) การทวนทำให้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใช้เวลาเพิ่มขึ้น 20-30 นาที ข้อมูลผลลัพธ์ที่ได้จากการทวนนั้นมีความสำคัญต่อข้อสอบ ผลลัพธ์ที่ได้ยังเสนอแนะว่าผู้สอบไม่มีความรู้ในการโกงการตอบโดยวิธีการตรวจสอบความยากของคำตอบหรือรูปแบบแผนเฉพาะในการเลือกข้อสอบ

วอลเตอร์ พี. วิสโปเอล (Walter P. Vispoel, 1998) ได้ทำการศึกษาผลความแตกต่างในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ระหว่างการอนุญาตให้ทวนคำตอบและการไม่อนุญาตให้ทวนคำตอบ พบว่าการอนุญาตให้ทวนคำตอบนั้นมีข้อสอบจำนวนเปอร์เซ็นต์เล็กน้อยที่มีการเปลี่ยนคำตอบ (5.23%) ข้อที่ถูกเปลี่ยนคำตอบจากผิดไปถูกมีจำนวนมากกว่าข้อที่เปลี่ยนจากถูกไปผิด (คิดเป็นสัดส่วน 2.92 : 1) ผู้เข้าสอบส่วนมาก (66.5%) เปลี่ยนคำตอบอย่างน้อย 1 ข้อ ผู้สอบที่เปลี่ยนคำตอบเกือบทั้งหมดได้รับค่าประมาณความสามารถที่สูงขึ้น (คิดเป็นสัดส่วน 2.55 : 1) ผู้สอบส่วนมากมีความต้องการทวนคำตอบแต่การอนุญาตจะทำให้ผู้สอบใช้เวลาในการสอบเพิ่มขึ้น 41%

วอลเตอร์ พี. วิสโปเอล, ซารา เจ. คลอจ และ ทิมโมธี เบลเลอร์ (Walter P. Vispoel, Sara J. Clough, and Timothy Bleiler, 1999) ที่นักวิจัยสาขาวัดและประเมินผลทางการศึกษาได้ทำการศึกษาเพิ่มเติมในการทวนคำตอบแบบแบ่งช่วง เนื่องจากทวนคำตอบแบบแบ่งช่วงนั้นไม่สามารถป้องกันค่าความสามารถที่เกินจริงได้ทั้งหมด เนื่องจากการใช้หลักการตัดสินใจความยากของข้อสอบว่าเมื่อใดควรเปลี่ยนคำตอบข้อก่อนหน้า จึงได้ทดลองหาผลลัพธ์จาก 6 สถานการณ์ที่ผู้สอบใช้ความยากของข้อสอบในการตัดสินใจเปลี่ยนคำตอบในการทดสอบแบบ



ปรับเหมาะด้วยคอมพิวเตอร์และเปรียบเทียบกับผลลัพธ์ที่ได้จากผู้สอบที่ทวนและเปลี่ยนคำตอบตามปกติ เงื่อนไขสถานการณ์แตกต่างกันไปเมื่อผู้สอบถูกกระตุ้นให้เปลี่ยนคำตอบและทราบวิธีในกระบวนการเลือกข้อสอบที่แน่นอน ซึ่งพบว่าผู้สอบจะได้ค่าเฉลี่ยที่ดีเมื่อทำการทวนและเปลี่ยนคำตอบตามปกติ แผนการที่ดีที่สุดคือหนึ่งในผู้สอบทราบวิธีในกระบวนการเลือกข้อสอบและถูกกระตุ้นให้เปลี่ยนคำตอบต่อเมื่อไม่แน่ใจความถูกต้องของคำตอบและแน่ใจในการตัดสินใจระดับความยากของข้อสอบ

เจ. โอเลีย, เจ. เลวเอลต้า, ซิเมเนซ และเอบาด (J.Olea, J.Revuelta, Ximenez, and Abad, 2000) ได้ทำการทดสอบความสามารถทางคำศัพท์ภาษาอังกฤษของนักศึกษาชาวสเปนด้วยการทดสอบด้วยคอมพิวเตอร์ (แบบปกติและแบบปรับเหมาะ) ภายใต้เงื่อนไขการทวน (อนุญาตให้ทวนคำตอบและไม่อนุญาตให้ทวนคำตอบ) โดยมุ่งการตรวจสอบผลก่อนการทวนและหลังการทวน ผลลัพธ์ 2 สิ่งหลักที่ได้รับหลังการทวนคือ จำนวนคำตอบที่ถูกและค่าประมาณความสามารถมีค่าเพิ่มขึ้น และลดระดับความกังวลใจของผู้สอบ การสอบที่อนุญาตให้ทวนคำตอบและไม่อนุญาตให้ทวนคำตอบนั้นไม่มีความแตกต่างกันในค่าความคลาดเคลื่อนในการวัด

วอลเตอร์ พี. วิสโพเอล, เอมี บี เฮนดริกสัน และทิมโมธี เบลเลอร์ (Walter P. Vispoel, Amy B. Hendrickson, and Timothy Bleiler, 2000) ได้ทำการศึกษาเพิ่มเติมเกี่ยวกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่อนุญาตให้ทวนคำตอบแบบช่วงนั้น ได้ทดลองศึกษาใน 4 ลักษณะคือ ไม่อนุญาตให้ทวนคำตอบ, อนุญาตให้ทวนคำตอบแบบแบ่งช่วง ช่วงละ 5 ข้อ, อนุญาตให้ทวนคำตอบแบบแบ่งช่วง ช่วงละ 10 ข้อ, อนุญาตให้ทวนคำตอบหลังจากทำแบบสอบเสร็จสิ้นทั้ง 40 ข้อ ผลปรากฏว่าไม่มีความแตกต่างในค่าประมาณความสามารถทางภาษา, ความคลาดเคลื่อนในการวัดหรือเวลาที่ใช้ในการสอบ ภายใต้เงื่อนไขการทวนคำตอบแบบต่างๆ ค่าประมาณความสามารถเพิ่มขึ้นเล็กน้อยหลังการทวนคำตอบ คำตอบส่วนใหญ่แก้จากผิดไปถูกมากกว่าแก้จากถูกไปผิด ผู้สอบที่เปลี่ยนคำตอบส่วนมากจะได้รับค่าความสามารถเพิ่มขึ้น ผู้สอบเกือบทุกคนระบุว่าได้รับโอกาสพอเพียงในการทวนคำตอบ และการทวนแบบจำกัดช่วงสั้น (5 ข้อ) อาจทำหน้าที่ได้ดีพอๆ กับการทวนได้ทั้งหมด แต่สามารถลดปัญหาการโกงข้อสอบจากการเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้ เนื่องจากผู้สอบต้องตอบข้อสอบให้เสร็จก่อนภายในช่วง จึงจะสามารถกลับไปทวนคำตอบได้ จึงเป็นการประมาณค่าความสามารถของผู้สอบไว้แล้วครั้งหนึ่ง รวมทั้งยังมีการประมาณค่าความสามารถของผู้สอบระหว่างช่วงด้วยเมื่อผู้สอบทำการทวนเสร็จในช่วง

เซกุลละ-วาคุระ, อาร์ และ บริโท, ซี (Sekula-Wacura, R, and Brito, C , 2000) ได้ทำการศึกษาผลจากการดำเนินการจัดการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ของ คณะกรรมการเอเอสซีที ซึ่งศึกษาถึงผลกระทบในการอนุญาตให้ทวนคำตอบทั้ง 3 วิธีการที่นำมาใช้ในการทวนคำตอบและสามารถควบคุมการประมาณค่าความสามารถที่เกินจริงได้ คือ การทวนที่มีการจำกัดจำนวน (Limited Review), การทวนที่มีการจำกัดช่วง (Block Review) และการทวนในหมวดสิ่งเร้าเดียวกัน (Stimulus-based Review) โดยสรุปแล้วพบว่า การอนุญาตให้ทวนคำตอบทั้ง 3 วิธีนั้นส่งผลดีต่อผู้เข้าสอบทั้งในด้านสภาพทางจิตวิทยาของผู้สอบและด้านการสามารถวัดความสามารถของผู้สอบได้แม่นยำมากขึ้น

ซาเวียร์ เรวเอลต้า, เอ็ม. คาร์เมน ไชเมเนซ และจูลิโอ โอลเลีย (Javier Revuelta, M. Carmen Ximenez, and Julio Olea, 2003) ได้ทำการศึกษาคูณสมบัติทางจิตวิทยาของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ระหว่างความสัมพันธ์ของกระบวนการเลือกข้อสอบและโอกาสในการทวนคำตอบ ในแง่ของความกระวนกระวายและความสบายใจของผู้สอบ สมมติฐานที่ต่างกันทั้ง 2 ข้อถูกนำมาใช้ออกแบบคุณสมบัติแบบสอบ สังเกตการควบคุมและสังเกตประสิทธิภาพ การดำเนินการทดสอบประกอบด้วย 3 ประเภท (แบบปรับเหมาะ, แบบปรับเหมาะอย่างง่าย, และแบบตายตัว) และมี 4 เงื่อนไขการทวน (ไม่อนุญาตให้ทวนคำตอบ, อนุญาตให้ทวนคำตอบหลังจากทำแบบสอบเสร็จสิ้น, อนุญาตให้ทวนคำตอบแบบแบ่งช่วง ช่วงละ 5 ข้อ, อนุญาตให้ทวนคำตอบแบบข้อต่อข้อ) ซึ่งใช้กับกลุ่มตัวอย่างที่เป็นนักศึกษามหาวิทยาลัย 557 คน ผลกระทบนัยสำคัญทางสถิติถูกพบในเรื่องความแม่นยำของแบบสอบประเภทต่างๆ คำตอบหลังการทวนเพิ่มค่าประมาณความสามารถและเพิ่มเวลาในการสอบ ไม่พบผลกระทบจิตวิทยาทางความกระวนกระวายใจ ผู้สอบที่ได้รับอนุญาตให้ทวนคำตอบทุกแบบคำนึงถึงความสำคัญในความเป็นไปได้ของการทวนมากกว่าผู้เข้าสอบที่ไม่ได้รับอนุญาตให้ทวนคำตอบ ผลลัพธ์นี้สอดคล้องกับข้อมูลคุณสมบัติของผู้สอบที่หาไว้ก่อนหน้านี้ และจุดประเด็นที่ควรพิจารณาในสาขาของการสอบที่ควรอนุญาตให้มีการทวน

ประเด็นของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (computerized adaptive testing - CAT) ที่เป็นข้อถกเถียงกันอย่างมาก นั่นคือหลังจากที่ผู้สอบทำการทดสอบเสร็จ ผู้สอบควรได้รับโอกาสให้กลับไปทวนคำตอบหรือไม่ ซึ่งเป็นที่ชัดเจนว่าผู้สอบยอมต้องการที่จะกลับไปทวนคำตอบ เนื่องจากน่าจะได้มาถึงการประมาณค่าความสามารถที่แม่นยำหลังทวน และด้วยข้อเท็จจริงที่ว่า การเลือกคำตอบครั้งแรกอาจไม่ใช่คำตอบที่ถูกต้องที่สุดเสมอ แต่นักวิจัยที่มีชื่อเสียงบางท่านต่างมีความเห็นที่ไม่ควรเปิดโอกาสให้มีการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

เนื่องจากข้อจำกัดของกระบวนการทดสอบ ตัวอย่างเช่น จะสูญเสียความมีประสิทธิภาพ และเกิดความลำเอียงในการประมาณค่าความสามารถของผู้สอบ รวมทั้งเกิดความคลาดเคลื่อนในการได้มาถึงความสามารถที่แท้จริงของผู้สอบในกระบวนการทดสอบ ทำให้ขัดแย้งกับกระบวนการสร้างแบบทดสอบแบบปรับเหมาะที่มีประสิทธิภาพ กล่าวคือเมื่อผู้สอบตั้งใจจะตอบคำถามผิด ดังนั้นจึงทำให้ได้รับข้อสอบข้อต่อไปที่มีระดับความยากต่ำลง และถ้าเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้โดยไม่มีเงื่อนไข ผู้สอบก็จะสามารถกลับไปแก้คำตอบข้อก่อนหน้า ทำให้ข้อสอบแต่ละข้อนั้นไม่สัมพันธ์กับระดับความสามารถที่แท้จริงของผู้สอบ จากงานวิจัยพบว่าผู้สอบจะได้รับการประมาณค่าความสามารถโน้มไปทางด้านบวกหรือได้รับคะแนนที่สูงเกินความสามารถที่แท้จริง และมีความคลาดเคลื่อนสูง วิธีการที่นำมาใช้ในการทวนคำตอบที่สามารถควบคุมการประมาณค่าความสามารถที่เกินจริงได้ คือ การทวนที่มีการจำกัดช่วง (Block Review) โดยผู้สอบจะสามารถทวนคำตอบพร้อมทั้งเปลี่ยนคำตอบภายในช่วงที่จำกัดไว้ได้ทั้งหมด แต่ไม่สามารถที่จะทวนข้ามช่วงได้ ยกตัวอย่างเช่น การจำกัดช่วง 10 ข้อ กล่าวคือผู้สอบจะสามารถกลับมาทวนคำตอบภายในช่วง (ข้อ 1 – 10) ได้หลังจากที่ทำเสร็จข้อ 10 หลังจากนั้นจะไปทำข้อสอบในช่วงถัดไป (ข้อ 11 - 20) และจะสามารถกลับมาทวนคำตอบในช่วงนี้ได้ก็ต่อเมื่อทำเสร็จถึงข้อ 20 แต่ไม่สามารถกลับไปทวนคำตอบในช่วงก่อนหน้านี้อีก (ข้อ 1 - 10) ได้ เป็นต้น โดยการทวนแบบจำกัดช่วงสั้นๆ (5 ข้อ) อาจทำหน้าที่ได้ดีพอๆ กับการทวนได้ทั้งหมด แต่สามารถลดปัญหาการโกงข้อสอบจากการเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้ เนื่องจากผู้สอบต้องตอบข้อสอบให้เสร็จก่อนภายในช่วง จึงจะสามารถกลับไปทวนคำตอบได้ จึงเป็นการประมาณค่าความสามารถของผู้สอบไว้แล้วครั้งหนึ่ง รวมทั้งยังมีการประมาณค่าความสามารถของผู้สอบระหว่างช่วงด้วย เมื่อผู้สอบทำการทวนเสร็จในช่วง

#### 4.3 งานวิจัยที่เกี่ยวข้องกับการให้เปลี่ยนคำตอบในการทำการทดสอบแบบดั้งเดิม

กานดา ทองวัฒนะ (2522) ได้ทำการศึกษาการเปลี่ยนคำตอบในการทำข้อสอบชนิดเลือกตอบ พบว่า คะแนนรวมหลังการเปลี่ยนคำตอบเพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ .001, จำนวนการเปลี่ยนคำตอบเมื่อข้อสอบง่าย และจำนวนการเปลี่ยนคำตอบเมื่อข้อสอบยาก ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

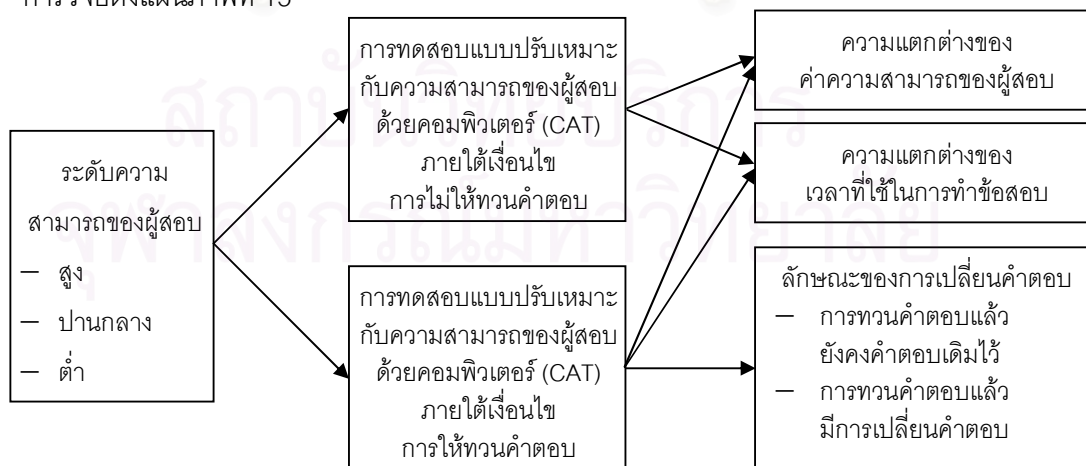
ขวัญใจ อุตราภรณ์ (2527) ได้ทำการศึกษาการเปลี่ยนคำตอบในการทำแบบสอบของนักเรียนเมื่อสอบซ้ำ จำนวนข้อสอบ 45 ข้อ (ครั้งแรกยังไม่ให้เปลี่ยนคำตอบ ส่วนครั้งหลังให้มีการเปลี่ยนคำตอบได้) พบว่า นักเรียนทุกคนมีการเปลี่ยนคำตอบมากกว่า 1 ข้อ พิสัยของการ

เปลี่ยนคำตอบมีจำนวนอยู่ระหว่าง 4 ถึง 39 ข้อ และนักเรียนที่มีความสามารถต่ำจะเปลี่ยนคำตอบมากกว่านักเรียนที่มีความสามารถสูงอย่างมีนัยสำคัญที่ระดับ .001

ประณาท เทียนศรี (2546) ได้ทำการศึกษาปัจจัยและผลจากการเปลี่ยนคำตอบข้อสอบปรนัยของนักเรียนชั้นประถมศึกษาปีที่ 5 โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายประถม พบว่า ในข้อสอบวิชาคณิตศาสตร์ข้อสอบที่นักเรียนจะเปลี่ยนคำตอบมากที่สุดจะเป็นข้อสอบวัดพฤติกรรมด้านความเข้าใจ โดยส่วนใหญ่จะมีค่าอำนาจจำแนกดีมากที่สุด และมีค่าความยากระดับง่าย และปานกลาง นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ระดับดีมาก ส่วนใหญ่เปลี่ยนคำตอบใหม่แล้วถูกต้องมากขึ้น ไม่ว่าจะข้อสอบนั้นจะมีความยากระดับง่าย หรือระดับปานกลาง ส่วนนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ระดับต่ำกว่าปานกลางและต้องปรับปรุง จะเปลี่ยนคำตอบใหม่แล้วผิดมากกว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ระดับดีกว่า

จากงานวิจัยทั้ง 3 งานนี้แสดงให้เห็นถึงปัจจัยในการเปลี่ยนคำตอบของผู้สอบ ซึ่งอาจเกิดจากความยากง่ายของข้อสอบ วิชาที่สอบ ประเภทของพฤติกรรมที่วัด หรือแม้กระทั่งระดับความสามารถของผู้สอบ ซึ่งเห็นได้ชัดเจนว่าในการทดสอบแบบประเพณีนิยมที่เน้นการใช้แบบสอบพร้อมฉบับเดียวและตอบคำถามโดยใช้ปากกาหรือดินสอ นั้น ให้ผลที่ว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนอยู่ในระดับปานกลางถึงดีมากจะมีการเปลี่ยนคำตอบแล้วส่งผลดีต่อตัวผู้สอบ (ได้คะแนนสูงขึ้น) มากกว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนในระดับต่ำกว่าปานกลางและต้องปรับปรุง เนื่องจากนักเรียนในกลุ่มหลังส่วนใหญ่จะมีเปลี่ยนคำตอบใหม่แล้วผิดมากกว่ากลุ่มแรก

จากการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง ทำให้ผู้วิจัยได้กรอบแนวคิดในการวิจัยดังแผนภาพที่ 15



แผนภาพที่ 15 กรอบแนวคิดในการวิจัย

### บทที่ 3

#### วิธีดำเนินการวิจัย

##### ประชากรและกลุ่มตัวอย่าง

ประชากร ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2549 โรงเรียนประถมศึกษา จังหวัดกรุงเทพมหานคร สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) จำนวน 36 โรงเรียน รวมนักเรียน 5,592 คน (สำนักนโยบายและแผนการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ, 2549)

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ ได้มาจากระชากรที่มีการสุ่มแบบหลายขั้นตอน (multistage random) โดยมีขั้นตอนการสุ่มดังนี้

ขั้นตอนแรก สุ่มโดยใช้สำนักงานเขตพื้นที่การศึกษาซึ่งมีทั้งหมด 3 เขต คือเขต 1, เขต 2 และเขต 3 เป็นหน่วยในการสุ่ม โดยการสุ่มอย่างง่าย (simple random sampling) ได้สำนักงานเขตพื้นที่การศึกษารุงเทพมหานคร เขต 1 ซึ่งมีโรงเรียนที่มีชั้นประถมศึกษาปีที่ 6 จำนวน 16 โรงเรียน รวม 2,873 คน จำแนกเป็นรายชื่อโรงเรียนตามตารางที่ 1

ตารางที่ 1 จำนวนโรงเรียนที่ได้จากการสุ่มขั้นตอนแรก จำแนกเป็นรายชื่อโรงเรียน

ชื่อโรงเรียน	แขวง / เขต	จำนวน ห้องเรียนชั้น ประถมศึกษา ปีที่ 6 (ห้อง)	จำนวน นักเรียนชั้น ประถมศึกษา ปีที่ 6 (คน)
1) ดาราคาม	พระโขนง / คลองเตย	5	187
2) ทุ่งมหาเมฆ	ทุ่งวัดดอน / เขตสาทร	5	184
3) ประถมนนทรี	ช่องนนทรี / ยานนาวา	4	153
4) พญาไท	ทุ่งพญาไท / ราชเทวี	9	375
5) พระตำหนักสวนกุหลาบ	พระบรมมหาราชวัง / พระนคร	3	90
6) มหาวิทยาลัยวัด	ตลาดน้อย / สัมพันธวงศ์	—	57
7) ราชวินิต	ดุสิต / ดุสิต	10	409
8) วัดชนะสงคราม	ชนะสงคราม / พระนคร	2	59
9) วัดชัยชนะสงคราม	สัมพันธวงศ์ / สัมพันธวงศ์	1	41
10) วัดदान	บางโพธิ์ / ยานนาวา	1	36
11) วัดพลับพลาชัย	ป้อมปราบ / ป้อมปราบศัตรูพ่าย	6	224
12) วัดโสมนัส	วัดโสมนัส / ป้อมปราบศัตรูพ่าย	3	69

ชื่อโรงเรียน	แขวง / เขต	จำนวน ห้องเรียนชั้น ประถมศึกษา ปีที่ 6 (ห้อง)	จำนวน นักเรียนชั้น ประถมศึกษา ปีที่ 6 (คน)
13) สายน้ำทิพย์	คลองตัน / คลองเตย	—	342
14) อนุบาลพิบูลเวศม์	พระโขนงเหนือ / วัฒนา	7	267
15) อนุบาลวัดปรีณายก	บ้านพานถม / พระนคร	4	175
16) อนุบาลสามเสนวิทยาลัย	สามเสนใน / พญาไท	5	205
	<b>รวม</b>		<b>2,873</b>

ขั้นตอนที่สอง สุ่มโรงเรียนในเขตพื้นที่การศึกษากรุงเทพมหานคร เขต 1 โดยการสุ่มอย่างง่าย (simple random sampling) ได้เป็นโรงเรียนวัดพลับพลาชัย ซึ่งมีนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2549 จำนวน 6 ห้องเรียน คิดเป็นกลุ่มตัวอย่าง 224 คน กลุ่มนักเรียนที่ใกล้เคียงกับกลุ่มตัวอย่าง ซึ่งผู้วิจัยใช้เป็นส่วนหนึ่งในกระบวนการดำเนินการวิจัย คือ

#### 1) นักเรียนกลุ่ม 1

ใช้ในการหาค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ของแบบสอบปรนัยวิชา คณิตศาสตร์ เรื่องความเข้าใจเศษส่วน เพื่อพัฒนาคลังข้อสอบวิชา คณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ระดับชั้นประถมศึกษาปีที่ 6 ให้ครบอย่างน้อย 210 ข้อ ซึ่งได้มาจากการเลือกแบบเจาะจงโรงเรียนประถมศึกษาขนาดใหญ่ ในเขตพื้นที่การศึกษา กรุงเทพมหานคร เขต 1 ได้เป็นโรงเรียนราชวินิต ซึ่งมีนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 10 ห้องเรียน คิดเป็นนักเรียน 409 คน

#### 2) นักเรียนกลุ่ม 2

ใช้ในการกำหนดระยะเวลาสูงสุดที่เหมาะสมในการทำข้อสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing - CAT) ให้กับกลุ่มตัวอย่าง รวมทั้งตรวจสอบและปรับปรุงความถูกต้องของโปรแกรมคอมพิวเตอร์ ซึ่งได้มาจากการเลือกแบบเจาะจง โดยเลือกนักเรียนที่กำลังศึกษาอยู่ในระดับชั้นประถมศึกษาปีที่ 6 คอร์สคณิตเข้มเปิดเทอมปลาย ปีการศึกษา 2549 ในโรงเรียนกวดวิชา พี.พี.เซ็นเตอร์ ของทางผู้วิจัย คิดเป็นนักเรียน 9 คน ซึ่งจะเลือกนักเรียนที่มีความสามารถในวิชาคณิตศาสตร์แตกต่างกัน (ต่ำ, ปานกลาง และสูง) โดยอาศัยคะแนนทดสอบฟรีเทสต์ วิชาคณิตศาสตร์ของทางโรงเรียนเป็นตัวกำหนด

## เครื่องมือที่ใช้ในการวิจัย


1) คลังข้อสอบวิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ของ รั้งสรรพค์ มณีเล็ก (2540) ซึ่งเป็นแบบสอบปรนัย 4 ตัวเลือก จำนวน 200 ข้อ ซึ่งสร้างจากฟาเซตดังนี้


ตารางที่ 2 ฟาเซตสำหรับสร้างข้อสอบ เรื่องความเข้าใจเศษส่วน ระดับชั้นประถมศึกษาปีที่ 6

จุดประสงค์ / เนื้อหา / ข้อ	ฟาเซต	
	คำถาม	คำตอบ
1. มีความเข้าใจพื้นฐานเกี่ยวกับเรื่องเศษส่วน		
1.1 ความหมายของเศษส่วน	<u>สัญลักษณ์</u> 1. รูปที่แบ่งเป็นส่วนที่เท่า ๆ กัน 2. เศษส่วน	<u>สัญลักษณ์</u> 1. รูปที่แบ่งเป็นส่วนเท่า ๆ กัน 2. เศษส่วน
1.2 ชนิดของเศษส่วน	<u>ชนิด</u> 1. เศษส่วนแท้ 2. เศษเกิน 3. จำนวนคละ <u>จัดกลุ่ม</u> 1. ประเภทเดียวกัน 2. คนละประเภท <u>สมบัติ</u> 1. เศษส่วนแท้ 2. เศษเกิน 3. จำนวนคละ <u>การจัดกลุ่ม</u> 1. เข้าพวก 2. ไม่เข้าพวก	<u>ชนิด</u> 1. เศษส่วนแท้ 2. เศษเกิน 3. จำนวนคละ <u>จำนวนเศษส่วนที่เสนอ</u> 1. เดี่ยว $(\frac{1}{2})$ 2. กลุ่ม $(\frac{1}{2}, \frac{1}{4}, 4/5)$
1.3 เศษส่วนของจำนวนนับ	<u>การนำเสนอ</u> 1. เศษส่วนของจำนวนนับ 2. รูปที่แบ่งเป็นส่วนเท่า ๆ กัน	<u>การนำเสนอ</u> 1. เศษส่วนของจำนวนนับ 2. รูปที่แบ่งเป็นส่วนเท่า ๆ กัน

จุดประสงค์ / เนื้อหา / ข้อ	พาเซท	
	คำถาม	คำตอบ
	<u>ลักษณะ</u>	<u>ลักษณะ</u>
	1. สัญลักษณ์	1. สัญลักษณ์
	2. ความหมาย	2. ความหมาย
	<u>การเปรียบเทียบ</u>	
	1. เท่ากัน	
	2. ไม่เท่ากัน	
1.4 เศษส่วนอย่างต่ำ	<u>ชนิด</u>	<u>ชนิด</u>
	1. เศษส่วนแท้	1. เศษส่วนแท้
	2. เศษเกิน	2. เศษเกิน
	3. จำนวนคละ	3. จำนวนคละ
	<u>สัญลักษณ์เศษส่วน</u>	<u>วิธีการ</u>
	1. ตัวเลข	1. นำจำนวนเต็มหารหรือคูณ
	2. ตัวอักษร(จัดไม่จัดกลุ่ม)	2. นำเศษส่วนที่มีเศษและ
	<u>วิธีการ</u>	ส่วนเท่ากันหารหรือคูณ
	1. นำจำนวนเต็มหาร หรือ	3. นำเศษส่วนที่มีเศษส่วน
	คูณ	และส่วนไม่เท่ากันหารหรือ
	2. นำเศษส่วนที่มีเศษและ	คูณ
	ส่วนเท่ากันหรือคูณ	
	3. นำเศษส่วนที่มีเศษและ	
	ส่วนไม่เท่ากันหรือคูณ	
1.5 การแปลงเศษเกินเป็น	<u>ชนิด</u>	<u>ชนิด</u>
จำนวนคละ	1. เศษเกิน	1. เศษเกิน
	2. จำนวนคละ	2. จำนวนคละ
	<u>การเปรียบเทียบ</u>	<u>การระบุค่าที่เท่ากัน</u>
	1. เท่ากัน	1. ระบุเฉพาะเศษหรือส่วน
	2. ไม่เท่ากัน	2. ระบุทั้งเศษและส่วน
	<u>สัญลักษณ์แทนจำนวน</u>	<u>การแจกแจงองค์ประกอบ</u>
	ตัวเลข/ตัวอักษร	แจกแจง/ไม่แจกแจง



จุดประสงค์ / เนื้อหา / ข้อ	ฟาเซท	
	คำถาม	คำตอบ
2. เมื่อกำหนดเศษส่วนให้ สามารถเขียนเป็นเศษส่วน ที่มีค่าเท่าเดิมโดยที่ตัวเศษ หรือตัวส่วนมีค่าตามที่ กำหนดให้ได้	ชนิด 1. เศษส่วนแท้ 2. เศษเกิน 3. จำนวนคละ การเปรียบเทียบ 1. เท่ากัน 2. ไม่เท่ากัน ลักษณะแทนจำนวน 1. รูป 2. สัญลักษณ์	
3. เมื่อกำหนดเศษส่วนสอง จำนวนสามารถ เปรียบเทียบและใช้ สัญลักษณ์ $>$ , $<$ หรือ $=$ ได้	ลักษณะของเศษส่วน 1. เศษเท่ากัน 2. ส่วนเท่ากัน 3. เศษน้อยกว่า/มากกว่า ครึ่งหนึ่งของส่วน สัญลักษณ์ 1. ตัวเลข 2. ตัวอักษร 3. รูป การเรียงลำดับเศษส่วน 1. มากไปหาน้อย 2. น้อยไปหามาก	เครื่องหมายที่ใช้ 1. มากกว่า 2. น้อยกว่า 3. เท่ากับ ลักษณะของเศษส่วน 1. เศษเท่ากัน 2. ส่วนเท่ากัน 3. เศษน้อยกว่า/มากกว่า ครึ่งหนึ่งของส่วน
4. เมื่อกำหนดเศษส่วนใดๆ สามารถบวก, ลบ, คูณ และหารระคนได้ - การบวกและการลบ	รูปที่นำเสนอ 1. แรกตามจำนวน  = $\frac{4}{8}$	สมบัติ 1. สลับที่ 2. เปลี่ยนกลุ่ม

จุดประสงค์ / เนื้อหา / ข้อ	ฟาเซท	
	คำถาม	คำตอบ
	<p>2. แรเงาเป็นสัดส่วน</p>  $= \frac{1}{2}$	<p><u>การนำเสนอ</u></p> <ol style="list-style-type: none"> <li>จำนวนโดด</li> <li>จำนวนหลายจำนวนที่กระทำการกันอยู่</li> </ol>
	<p><u>ตำแหน่งที่ให้เติมส่วนที่หายไป</u></p> <ol style="list-style-type: none"> <li>เศษ/ตัวคูณเศษ</li> <li>เศษส่วนที่มีส่วนไม่เท่ากัน</li> </ol> <p><u>สัญลักษณ์</u></p> <ol style="list-style-type: none"> <li>ตัวเลข</li> <li>อักษร</li> </ol>	
- การคูณ	<p><u>ความสัมพันธ์การบวกและคูณ</u></p> <ol style="list-style-type: none"> <li>จำนวนครั้งของการบวก</li> <li>เศษส่วนที่เท่ากัน</li> <li>เศษส่วนที่เท่ากันที่นำมาบวก</li> </ol> <p><u>สัญลักษณ์</u></p> <ol style="list-style-type: none"> <li>ตัวเลข</li> <li>อักษร</li> </ol>	<p><u>ชนิดของจำนวน</u></p> <ol style="list-style-type: none"> <li>จำนวนเต็ม</li> <li>เศษส่วน</li> </ol>
	<p><u>ส่วนที่ให้เติม</u></p> <ol style="list-style-type: none"> <li>ตัวตั้ง</li> <li>ตัวคูณ</li> <li>เศษหรือส่วนของผลลัพธ์</li> </ol>	
- การหาร	<p><u>ประโยคสัญลักษณ์</u></p> <ol style="list-style-type: none"> <li>จำนวนเต็ม</li> <li>เศษส่วน</li> <li>ตัวอักษร</li> </ol>	<p><u>สัญลักษณ์</u></p> <ol style="list-style-type: none"> <li>ตัวเลข</li> <li>อักษร</li> <li>รูป</li> </ol>

จุดประสงค์ / เนื้อหา / ข้อ	ฟาเซท	
	คำถาม	คำตอบ
	<p><u>ส่วนที่ให้เดิม</u></p> <p>4. เศษเท่ากัน</p> <p>5. ส่วนเท่ากัน</p> <p>6. เศษน้อยกว่า/มากกว่า ครึ่งหนึ่งของส่วน</p>	
	<p><u>สัญลักษณ์</u></p> <p>1. ตัวตั้ง</p> <p>2. ตัวหาร</p> <p>3. ผลลัพธ์</p>	
	<p><u>การเปรียบเทียบ</u></p> <p>1. เท่ากัน</p> <p>2. ไม่เท่ากัน</p>	
- การบวก, ลบ, คูณ, หาร ระคน	<p><u>เครื่องหมาย</u></p> <p>1. บวก</p> <p>2. ลบ</p> <p>3. คูณ</p> <p>4. หาร</p>	<p><u>ชนิด</u></p> <p>1. จำนวนเต็ม</p> <p>2. เศษส่วน</p> <p><u>สัญลักษณ์</u></p> <p>1. ตัวเลข</p> <p>2. อักขร</p>
	<p><u>ชนิด</u></p> <p>1. จำนวนเต็ม</p> <p>2. เศษส่วน</p>	<p><u>วิธี</u></p> <p>3. รูป</p>
	<p><u>สัญลักษณ์</u></p> <p>1. ตัวเลข</p> <p>2. อักขร</p> <p>3. รูป</p>	<p>1. หา ค.ร.น.</p> <p>2. แปลงจำนวนคละ</p> <p>3. กลับเศษเป็นส่วน</p> <p>4. เปลี่ยนเครื่องหมายจาก</p> <p><math>\div \leftrightarrow \times</math></p>
	<p><u>การเปรียบเทียบ</u></p> <p>1. เท่ากัน</p>	<p><u>สมบัติ</u></p> <p>1. สลับที่</p>

จุดประสงค์ / เนื้อหา / ข้อ	ฟาเซท	
	คำถาม	คำตอบ
	2. ไม่เท่ากัน (ต่างจาก)	2. เปลี่ยนกลุ่ม 3. แจกแจง
5. เมื่อกำหนดเศษซ้อนอย่างง่ายที่เฉพาะตัวเศษหรือตัวส่วนที่มีการบวกลบคูณหารอย่างใดอย่างหนึ่งให้สามารถหาผลลัพธ์ได้	<p><u>เครื่องหมาย</u></p> <ol style="list-style-type: none"> <li>1. บวก</li> <li>2. ลบ</li> <li>3. คูณ</li> <li>4. หาร</li> </ol> <p><u>ชนิดของจำนวน</u></p> <ol style="list-style-type: none"> <li>1. เศษส่วน</li> <li>2. จำนวนเต็ม</li> </ol> <p><u>สัญลักษณ์</u></p> <ol style="list-style-type: none"> <li>1. ตัวเลข</li> <li>2. ตัวอักษร</li> </ol> <p><u>ส่วนที่ให้เติม</u></p> <ol style="list-style-type: none"> <li>1. ตัวตั้ง</li> <li>2. ตัวหาร</li> </ol> <p><u>การแปลงรูป</u></p> <ol style="list-style-type: none"> <li>1. เศษส่วน ⇔ เศษซ้อน</li> <li>2. เศษซ้อน ⇔ เศษส่วน</li> </ol>	<p><u>ชนิดของจำนวน</u></p> <ol style="list-style-type: none"> <li>1. เศษซ้อน</li> <li>2. เศษส่วนธรรมดา</li> <li>3. จำนวนเต็ม</li> </ol>
6. เมื่อกำหนดโจทย์ปัญหาเกี่ยวกับการบวก ลบ คูณ หาร เศษส่วนให้สามารถแสดงวิธีทำได้	<p><u>เครื่องหมายการกระทำ</u></p> <ol style="list-style-type: none"> <li>1. บวก</li> <li>2. ลบ</li> <li>3. คูณ</li> <li>4. หาร</li> </ol> <p><u>ชนิดของจำนวน</u></p> <ol style="list-style-type: none"> <li>1. จำนวนเต็ม</li> <li>2. เศษส่วน</li> </ol>	<p><u>ลักษณะการตอบ</u></p> <ol style="list-style-type: none"> <li>1. ประโยคสัญลักษณ์</li> <li>2. คำตอบในรูปของการกระทำ</li> </ol>

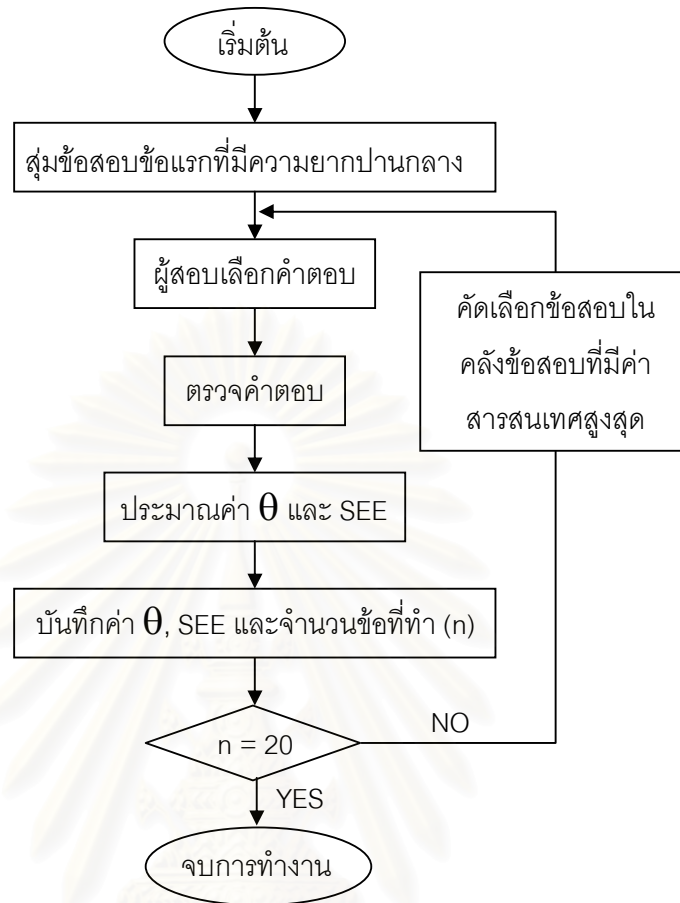
จุดประสงค์ / เนื้อหา / ข้อ	พาเซท	
	คำถาม	คำตอบ
	สถานการณ์	
	1. เงิน	
	2. งาน	
	3. สิ่งของ	
	4. พื้นที่	
	5. ความยาว	
	6. น้ำหนัก	
	7. เวลา	
	8. ปริมาตร	

2) แบบสอบปรนัยวิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วนซึ่งทางผู้วิจัยสร้างขึ้น ลักษณะของแบบสอบเป็นข้อสอบเลือกตอบ 4 ตัวเลือก จำนวน 40 ข้อ สำหรับคัดเลือกเพื่อเพิ่มข้อสอบอย่างน้อย 10 ข้อเข้าคลังข้อสอบ ด้วยวิธีการปรับเทียบค่าพารามิเตอร์ด้วยวิธีค่าเฉลี่ยและซิกมา (Mean and Sigma Method) โดยใช้โปรแกรม BILOG-MG สำหรับการวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์

3) โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งผู้สอบจะต้องทำการทดสอบ 2 เงื่อนไข รวม 2 ครั้ง และลำดับของเงื่อนไขการทดสอบจะถูกกำหนดโดยเลขที่ของนักเรียน กล่าวคือ ถ้าเลขที่เป็นเลขคี่ จะทดสอบครั้งที่ 1 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การทดสอบที่ให้ทวนคำตอบ และถ้าเลขที่เป็นเลขคู่ จะทดสอบครั้งที่ 1 ภายใต้การทดสอบที่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ

3.1) เงื่อนไขการทดสอบที่ 1 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ

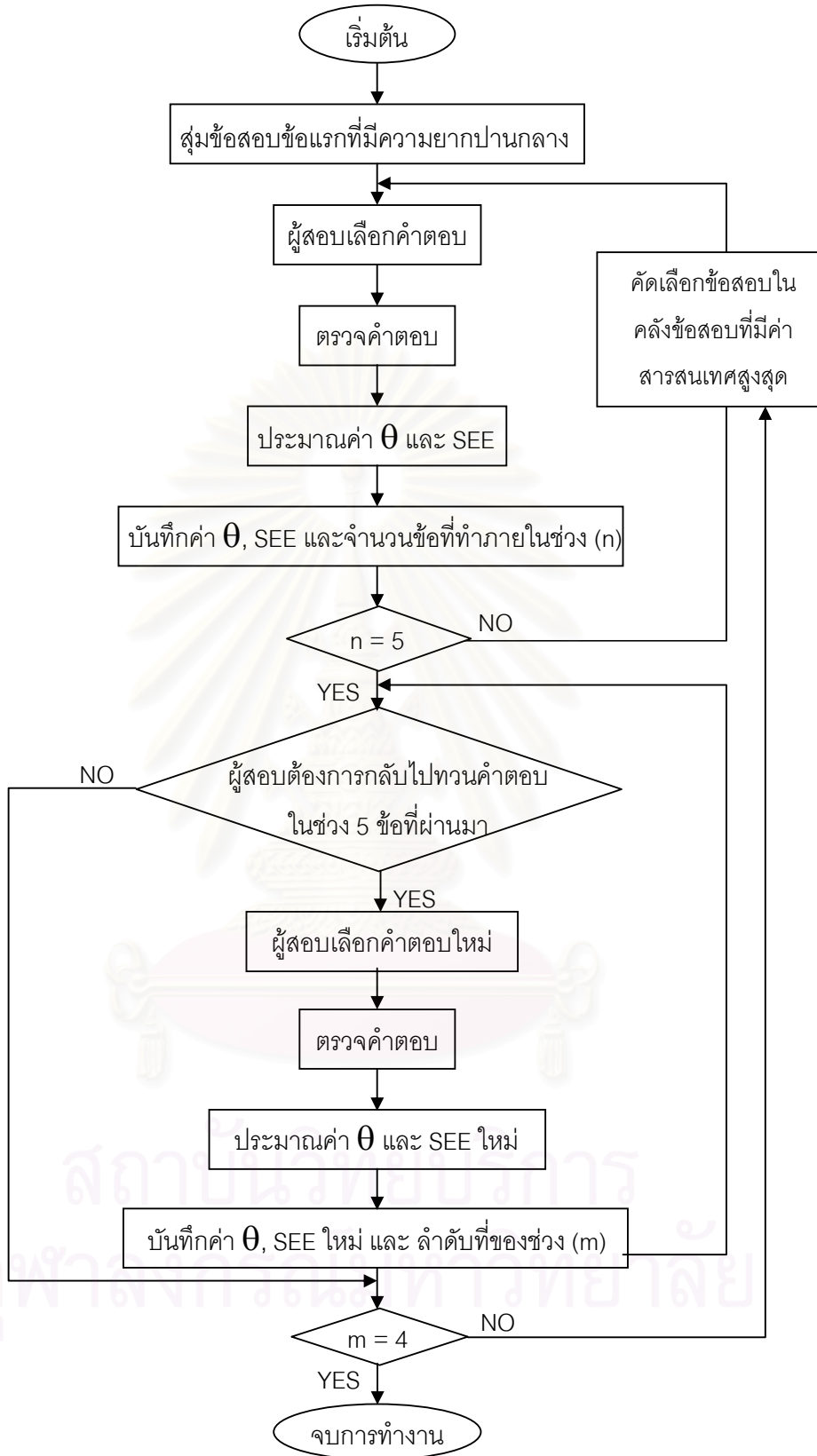
ผู้สอบเริ่มทำข้อสอบข้อแรกโดยการสุ่มข้อสอบที่มีระดับความยากปานกลาง เมื่อผู้สอบตอบโปรแกรมก็จะคำนวณเพื่อประมาณค่าความสามารถ และความคลาดเคลื่อนมาตรฐานของการประมาณค่า เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีความเหมาะสมกับความสามารถของผู้สอบ กระบวนการนี้จะดำเนินต่อไปจนกระทั่งผู้สอบทำข้อสอบครบ 20 ข้อ



แผนภาพที่ 16 ผังงานแสดงขั้นตอนการทำงานของคอมพิวเตอร์ตามเงื่อนไขการทดสอบที่ 1

3.2) เงื่อนไขการทดสอบที่ 2 ภายใต้การทดสอบที่ให้ทวนคำตอบโดยวิธีการทวนคำตอบแบบจำกัดช่วง (Block Review) ซึ่งผู้วิจัยกำหนดใช้ช่วง 5 ข้อ

ผู้สอบเริ่มทำข้อสอบข้อแรกโดยการสุ่มข้อสอบที่มีระดับความยากปานกลาง เมื่อผู้สอบตอบโปรแกรมก็จะคำนวณเพื่อประมาณค่าความสามารถ และความคลาดเคลื่อนมาตรฐานของการประมาณค่า เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีความเหมาะสมกับความสามารถของผู้สอบ ซึ่งผู้สอบสามารถทำเครื่องหมายหน้าข้อที่ต้องการกลับไปทวนคำตอบในแต่ละข้อได้ จากนั้นเมื่อผู้สอบทำข้อสอบครบ 5 ข้อ ผู้สอบจะสามารถกลับไปทวนคำตอบภายในช่วง 5 ข้อที่ทำเสร็จได้ โดยผู้วิจัยได้ทำการเขียนโปรแกรมในการประมาณค่าความสามารถ และความคลาดเคลื่อนมาตรฐานของการประมาณค่าใหม่หลังจากที่ผู้สอบทำการทวนและเปลี่ยนคำตอบเสร็จ จากนั้นผู้สอบจะทำการยืนยัน เพื่อที่จะทำข้อสอบในช่วง 5 ข้อถัดไป ซึ่งจะไม่สามารถกลับมาทวนในช่วง 5 ข้อที่ผ่านมาได้อีก กระบวนการนี้จะดำเนินต่อไปจนกระทั่งผู้สอบทำข้อสอบครบ 20 ข้อ



แผนภาพที่ 17 ผังงานแสดงขั้นตอนการทำงานของคอมพิวเตอร์ตามเงื่อนไขการทดสอบที่ 2

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองเพื่อเปรียบเทียบความแตกต่างของค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ภายใต้การทดสอบที่ให้ทวนคำตอบกับไม่ให้ทวนคำตอบ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน ซึ่งมีภาพรวมของวิธีดำเนินการวิจัย ดังนี้





## การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลเพื่อเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์(Computerized Adaptive Testing - CAT) เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน ได้แบ่งออกเป็น 3 ขั้นตอน คือ

**ขั้นตอนแรก** เป็นการทดสอบด้วยแบบสอบปรนัยวิชา คณิตศาสตร์ เรื่องความเข้าใจเศษส่วน จำนวน 40 ข้อ สำหรับคัดเลือกเพื่อเพิ่มข้อสอบอย่างน้อย 10 ข้อเข้าคลังข้อสอบ ซึ่งผู้วิจัยได้ดำเนินการดังนี้

- 1) สร้างแบบสอบปรนัย 4 ตัวเลือก วิชา คณิตศาสตร์ เรื่องความเข้าใจเศษส่วน จำนวน 40 ข้อ (ข้อสอบที่นำมาทดลองใช้ 30 ข้อ และข้อสอบร่วมที่สุ่มจากคลังข้อสอบ 10 ข้อ)
- 2) นำแบบสอบที่นำมาทดลองใช้ 30 ข้อ ไปให้อาจารย์ที่ปรึกษาตรวจสอบก่อนนำไปให้ผู้เชี่ยวชาญตรวจสอบความตรงตามเนื้อเรื่อง (Content Validity)
- 3) นำแบบสอบมาปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญ
- 4) นำหนังสือจากบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัยไปให้ผู้อำนวยการโรงเรียนราชวินิต ที่ถูกเลือกเป็นนักเรียนกลุ่ม 1 เพื่อขอทดสอบนักเรียนทั้งระดับชั้นประถมศึกษาปีที่ 6 จำนวน 409 คน และชี้แจงวัตถุประสงค์ของการวิจัย
- 5) ประสานงานกับครูผู้สอนประจำวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 เพื่อชี้แจงวัตถุประสงค์ของการวิจัยและนัดหมายวันและเวลาที่จะทดสอบนักเรียน
- 6) นำแบบสอบปรนัย 4 ตัวเลือก วิชา คณิตศาสตร์ เรื่องความเข้าใจเศษส่วน จำนวน 40 ข้อ มาทำการทดสอบนักเรียนที่เป็นนักเรียนกลุ่ม 1
- 7) นำผลการตอบแบบสอบของนักเรียนกลุ่ม 1 มาตรวจให้คะแนน และทำการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ คือ ค่าอำนาจจำแนก, ค่าความยากและค่าโอกาสเดาข้อสอบได้ถูก
- 8) คัดเลือกข้อสอบที่ดีมี ค่าอำนาจจำแนก (a) อยู่ระหว่าง +0.50 ถึง +2.50 และค่าความยาก (b) อยู่ระหว่าง -2.50 ถึง +2.50 และค่าโอกาสเดาข้อสอบได้ถูก (c) ไม่เกิน 0.30 และคำนวณค่าคงที่สำหรับปรับเทียบค่าพารามิเตอร์ด้วยวิธีค่าเฉลี่ยและซิกมาเพื่อบันทึกข้อสอบเข้าคลังเพิ่มอย่างน้อย 10 ข้อ รวมเป็นข้อสอบในคลังอย่างน้อยจำนวน 210 ข้อ

**ขั้นตอนที่สอง** เป็นการทดสอบแบบซีเอที (Computerized Adaptive Testing - CAT) ด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เพื่อ

หาระยะเวลาสูงสุดที่เหมาะสมในการทำข้อสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing - CAT) เพื่อกำหนดระยะเวลาสูงสุดในการทำข้อสอบให้กับกลุ่มตัวอย่าง รวมทั้งตรวจสอบและปรับปรุงความถูกต้องของโปรแกรมคอมพิวเตอร์ ซึ่งผู้สอบจะต้องทำการทดสอบ 2 เงื่อนไข รวม 2 ครั้ง และลำดับของเงื่อนไขการทดสอบจะถูกกำหนดโดยเลขที่ของนักเรียน กล่าวคือ ถ้าเลขที่เป็นเลขคี่ จะทดสอบครั้งที่ 1 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การทดสอบที่ให้ทวนคำตอบ และถ้าเลขที่เป็นเลขคู่ จะทดสอบครั้งที่ 1 ภายใต้การทดสอบที่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ ซึ่งเป็นแบบปรนัย 4 ตัวเลือก ที่ใช้รูปแบบแยกทางแปรผัน การประมาณค่าความสามารถของผู้สอบด้วยวิธีการของเบย์ส์ คัดเลือกข้อสอบข้อต่อไปโดยดูค่าสารสนเทศของข้อสอบข้อที่สูงที่สุด (Maximum item information) ยุติการทดสอบแบบคงที่ที่ 20 ข้อ ปฏิบัติการบนระบบวินโดวส์เอ็กซ์พี (Windows XP) ซึ่งผู้วิจัยได้ดำเนินการดังนี้

- 1) ทำจดหมายถึงผู้ปกครองของนักเรียนที่ถูกเลือกเป็นนักเรียนกลุ่ม 2 เพื่อขอทดสอบนักเรียนนอกเวลาเรียน และชี้แจงวัตถุประสงค์ของการวิจัย โดยนักเรียนที่ถูกเลือกเป็นนักเรียนกลุ่ม 2 จะได้ของรางวัลตอบแทนเป็นกระเป๋าถือของทางโรงเรียนกวดวิชา
- 2) ผู้วิจัยชี้แจงเกี่ยวกับคำสั่งและวิธีการตอบข้อสอบแก่ผู้สอบจนเข้าใจ ผลการตอบของผู้สอบจะถูกเก็บไว้ที่หน้าจอเครื่องคอมพิวเตอร์
- 3) ผู้วิจัยทำการบันทึกข้อมูลการตอบและค่าระดับความสามารถของผู้สอบที่ถูกเก็บไว้ที่หน้าจอคอมพิวเตอร์ทุกเครื่อง
- 4) ผู้วิจัยสอบถามความคิดเห็นของนักเรียนกลุ่ม 2 เกี่ยวกับโปรแกรมคอมพิวเตอร์โดยใช้คำถามปลายเปิด เพื่อรับทราบข้อดี ข้อเสีย รวมทั้งจุดที่ควรปรับปรุง
- 5) นำข้อมูลทีของนักเรียนกลุ่ม 2 แต่ละคนที่ได้มาวิเคราะห์ทางสถิติ เพื่อหาระยะเวลาสูงสุดที่เหมาะสมในการทำข้อสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing - CAT) โดยนำมาเป็นตัวกำหนดระยะเวลาสูงสุดในการทำข้อสอบให้กับกลุ่มตัวอย่าง รวมทั้งตรวจสอบและปรับปรุงความถูกต้องของโปรแกรมคอมพิวเตอร์

**ขั้นตอนที่สาม** เป็นการทดสอบแบบซีเอที (Computerized Adaptive Testing - CAT) ด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งผู้สอบจะต้องทำการทดสอบ 2 เงื่อนไข รวม 2 ครั้ง และลำดับของเงื่อนไขการทดสอบจะถูกเลือกโดยการสุ่ม กล่าวคือ ครั้งที่ 1 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การทดสอบที่ให้ทวนคำตอบ หรือครั้งที่ 1 ภายใต้การทดสอบที่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การ

ทดสอบที่ไม่ให้หวนคำตอบ ซึ่งเป็นแบบปรนัย 4 ตัวเลือก ที่ใช้รูปแบบแยกทางแปรผัน การประมาณค่าความสามารถของผู้สอบด้วยวิธีการของเบส์ คัดเลือกข้อสอบข้อต่อไปโดยดูค่าสารสนเทศของข้อสอบข้อที่สูงที่สุด (Maximum item information) ยุติการทดสอบแบบคงที่ที่ 20 ข้อ ปฏิบัติการบนระบบวินโดวส์เอ็กซ์พี (Windows XP)

1) นำหนังสือจากบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย ไปให้ผู้อำนวยการโรงเรียนวัดพลับพลาชัย ที่ถูกสุ่มเป็นกลุ่มตัวอย่าง เพื่อขอทดสอบนักเรียนทั้งระดับชั้นประถมศึกษาปีที่ 6 จำนวน 224 คน และชี้แจงวัตถุประสงค์ของการวิจัย

2) ประสานงานกับครูผู้สอนประจำวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 และเจ้าหน้าที่ที่ดูแลห้องคอมพิวเตอร์ของทางโรงเรียน เพื่อชี้แจงวัตถุประสงค์ของการวิจัยและนัดหมายวันและเวลาที่จะทดสอบนักเรียน (รวม 12 วัน 2 วัน ต่อ 1ห้อง)

3) ผู้วิจัยชี้แจงเกี่ยวกับคำสั่งและวิธีการตอบข้อสอบแก่ผู้สอบจนเข้าใจ ผลการตอบของผู้สอบรวมทั้งข้อมูลในการทำการทดสอบของผู้สอบจะถูกเก็บไว้ที่หน้าจอเครื่องคอมพิวเตอร์

4) ผู้วิจัยทำการบันทึกข้อมูลในการทำการทดสอบของผู้สอบที่ถูกเก็บไว้ที่หน้าจอคอมพิวเตอร์ทุกเครื่อง

5) นำข้อมูลที่ของกลุ่มตัวอย่าง แต่ละคนที่ได้มาวิเคราะห์ทางสถิติ เพื่อเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing – CAT) เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน

### **การวิเคราะห์ข้อมูล**

ในการวิเคราะห์ข้อมูลเพื่อเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing – CAT) เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน ผู้วิจัยได้แบ่งออกเป็น 3 ขั้นตอน คือ

**ขั้นตอนแรก** สำหรับการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และปรับเทียบค่าพารามิเตอร์ของข้อสอบเพื่อบันทึกค่าคลัง โดยดำเนินการ ดังนี้

1. คำนวณค่าพารามิเตอร์ความยาก(a), อำนาจจำแนก(b) และโอกาสในการเดาข้อสอบได้ถูก(c) ของข้อสอบทั้ง 40 ข้อ โดยใช้โปรแกรม BILOG-MG สำหรับการวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์

2. ทำการปรับเทียบค่าพารามิเตอร์ด้วยวิธีค่าเฉลี่ยและซิกมา (Mean and Sigma Method) เพื่อที่จะบันทึกข้อสอบเข้าคลังเพิ่ม โดยใช้โปรแกรม Microsoft Office Excel 2003

2.1 ทำการปรับเทียบค่าพารามิเตอร์อำนาจจำแนก (a) ความยาก (b) และโอกาสในการเดาข้อสอบได้ถูก (c) ซึ่งมีขั้นตอนดังนี้

(1) กรอกราค่าพารามิเตอร์อำนาจจำแนก ( $a_{YC}$ ) ความยาก ( $b_{YC}$ ) และโอกาสในการเดาข้อสอบได้ถูก ( $c_{YC}$ ) พร้อมทั้งคำนวณค่าเฉลี่ยพารามิเตอร์ความยาก ( $\bar{b}_{YC}$ ) และส่วนเบี่ยงเบนมาตรฐาน ( $S_{YC}$ ) ของข้อสอบร่วมจำนวน 10 ข้อจากคลังข้อสอบ สมมติเรียกว่า แบบสอบฉบับ Y

(2) คำนวณค่าพารามิเตอร์อำนาจจำแนก ( $a_{XC}$ ) ค่าพารามิเตอร์ความยาก ( $b_{XC}$ ) และโอกาสในการเดาข้อสอบได้ถูกค่า ( $c_{XC}$ ) พร้อมทั้งค่าเฉลี่ยพารามิเตอร์ความยาก ( $\bar{b}_{XC}$ ) และส่วนเบี่ยงเบนมาตรฐาน ( $S_{XC}$ ) ของข้อสอบทั้งฉบับจำนวน 40 ข้อ สมมติเรียกว่า แบบสอบฉบับ X

(3) จากข้อ (1) และ (2) คำนวณค่าคงที่  $\beta$

สูตร 
$$\beta = \bar{b}_{YC} - \bar{b}_{XC}$$

(4) คำนวณค่าคงที่  $\alpha$

สูตร 
$$\alpha = \frac{S_{YC}}{S_{XC}}$$

โดยที่  $S_{YC}$  คือ ส่วนเบี่ยงเบนมาตรฐานของ b จากแบบสอบฉบับ Y

$S_{XC}$  คือ ส่วนเบี่ยงเบนมาตรฐานของ b จากแบบสอบฉบับ X

(5) คำนวณค่าประมาณพารามิเตอร์อำนาจจำแนกของข้อสอบฉบับ X ทุกข้อที่ปรับเทียบให้อยู่บนสเกลของข้อสอบร่วม ( $a^*_{YC}$ )

สูตร 
$$a^*_{YC} = \frac{a_{XC}}{\alpha}$$

โดยที่  $a_{XC}$  คือ ค่าประมาณพารามิเตอร์อำนาจจำแนกของข้อสอบทั้งฉบับจำนวน 40 ข้อ

(6) คำนวณค่าประมาณพารามิเตอร์ความยากของข้อสอบฉบับ X ทุกข้อที่ปรับเทียบให้อยู่บนสเกลของข้อสอบร่วม ( $b^*_{YC}$ )

สูตร 
$$b^*_{YC} = \alpha b_{XC} - \beta$$

โดยที่  $b_{XC}$  คือ ค่าประมาณพารามิเตอร์ความยากของข้อสอบทั้งฉบับจำนวน 40 ข้อ



ข้อ เดิม	ข้อ คลัง	แบบสอบฉบับ Y			แบบสอบฉบับ X			เปรียบเทียบ			ข้อ ใหม่
		a	b	c	$a_{XC}$	$b_{XC}$	$c_{XC}$	$a^*_{YC}$	$b^*_{YC}$	$c^*_{YC}$	
7	181	0.723	-0.873	0.135	1.299	-1.124	0.142	1.177	-1.354	0.142	เดิม
8					1.280	-1.214	0.189	1.159	-1.453	0.189	202
9					1.263	1.473	0.120	1.144	1.513	0.120	203
10					0.530	-1.773	0.189	0.480	-2.070	0.189	ไม่ผ่าน
11	122	1.218	1.052	0.211	1.006	-0.361	0.237	0.911	-0.512	0.237	เดิม
12					1.297	-0.579	0.179	1.175	-0.752	0.179	204
13					0.908	-0.659	0.174	0.822	-0.841	0.174	205
14					1.094	0.630	0.174	0.991	0.583	0.174	206
15	4	0.979	2.211	0.132	0.824	0.527	0.247	0.746	0.469	0.247	เดิม
16					1.132	-0.515	0.238	1.025	-0.682	0.238	207
17					1.881	0.845	0.088	1.704	0.820	0.088	208
18	8	0.810	0.191	0.239	0.975	-0.950	0.186	0.883	-1.162	0.186	เดิม
19					0.852	-1.156	0.160	0.772	-1.389	0.160	209
20					0.787	1.142	0.280	0.713	1.148	0.280	210
21					1.235	-1.011	0.151	1.119	-1.229	0.151	211
22					1.014	-0.929	0.213	0.918	-1.139	0.213	212
23					0.824	-1.026	0.168	0.746	-1.246	0.168	213
24					1.126	-0.713	0.167	1.020	-0.900	0.167	214
25					1.244	0.878	0.258	1.127	0.856	0.258	215
26					0.832	0.467	0.160	0.754	0.403	0.160	216
27					1.384	-0.294	0.162	1.254	-0.438	0.162	217
28					1.832	0.139	0.163	1.659	0.040	0.163	218
29	95	0.834	0.648	0.224	1.189	-0.147	0.268	1.077	-0.275	0.268	เดิม
30					1.289	0.320	0.319	1.168	0.240	0.319	ไม่ผ่าน
31					0.781	-1.959	0.202	0.707	-2.276	0.202	219
32					0.636	-1.013	0.236	0.576	-1.231	0.236	220
33					0.829	-0.372	0.236	0.751	-0.524	0.236	221
34					0.884	-0.764	0.178	0.801	-0.956	0.178	222
35					0.914	0.107	0.100	0.828	0.005	0.100	223
36					1.565	1.455	0.123	1.418	1.493	0.123	224
37					1.992	0.978	0.083	1.804	0.967	0.083	30
38					1.186	0.349	0.257	1.074	0.272	0.257	117
39	119	1.076	2.645	0.192	0.818	1.810	0.307	0.741	1.885	0.307	เดิม
40	138	1.266	1.539	0.199	1.202	-0.416	0.212	1.089	-0.572	0.212	เดิม
		$\bar{b}_{YC} = 1.079$			$\bar{b}_{XC} = -0.217$			$\alpha$		$\beta$	
		$S_{YC} = 1.197$			$S_{XC} = 1.084$			$= 1.104$		$= 0.113$	

**หมายเหตุ** 1)  $a^*_{YC} = a_{XC} / 1.104$      $b^*_{YC} = 1.104b_{XC} - 0.113$      $c^*_{YC} = (c_{XC} + c) / 2$

2) ข้อ "ไม่ผ่าน" หรือ "ออก" จะขีดเส้นใต้ค่าพารามิเตอร์ที่ไม่ผ่านเกณฑ์ โดยเกณฑ์

ค่าพารามิเตอร์ที่นำข้อเข้าคลังคือ  $0.50 < a^*_{YC} < 2.50$ ,  $-2.50 < b^*_{YC} < 2.50$  และ  $c^*_{YC} \leq 0.30$

จากตารางที่ 3 พบว่ามีข้อสอบรวมในคลังที่ค่าพารามิเตอร์ไม่ผ่านเกณฑ์ 2 ข้อ คือข้อสอบข้อที่ 3 และ 6 (ตรงกับข้อในคลังที่ 177 และ 30) โดยข้อที่ 3 มีค่าอำนาจจำแนกเท่ากับ 0.176 (น้อยกว่า 0.50) จึงไม่ผ่านเกณฑ์ และข้อที่ 6 มีค่าอำนาจจำแนกเท่ากับ 0.438 (น้อยกว่า 0.50) และค่าความยากเท่ากับ  $-2.725$  (น้อยกว่า  $-2.50$ ) จึงไม่ผ่านเกณฑ์

มีข้อสอบที่ผู้วิจัยออกใหม่ที่ค่าพารามิเตอร์ไม่ผ่านเกณฑ์ 4 ข้อ คือข้อสอบข้อที่ 4, 5, 10 และ 30 โดยข้อที่ 4 มีค่าอำนาจจำแนกเท่ากับ 0.371 (น้อยกว่า 0.50) ข้อที่ 5 มีค่าอำนาจจำแนกเท่ากับ 0.287 (น้อยกว่า 0.50) ข้อที่ 10 มีค่าอำนาจจำแนกเท่ากับ 0.480 (น้อยกว่า 0.50) ข้อที่ 30 มีค่าโอกาสในการเดาข้อสอบได้ถูกเท่ากับ 0.319 (มากกว่า 0.30) จึงไม่ผ่านเกณฑ์

ดังนั้นจึงเหลือข้อสอบที่มีค่าพารามิเตอร์ผ่านเกณฑ์ทั้งหมด 26 ข้อ ที่จะบันทึกเข้าคลังในโปรแกรมคอมพิวเตอร์ คือข้อสอบข้อที่ 1, 8, 9, 12 – 14, 16, 17, 19 – 28 และ 31 – 38 (นำเข้าคลังใหม่เป็นข้อสอบข้อที่ 201-224 เนื่องจากข้อสอบรวมที่ตรงกับข้อในคลังข้อที่ 177 และ 30 ไม่ผ่านเกณฑ์ จึงนำข้อ 37 และ 38 ไปแทนที่ข้อในคลัง โดยข้อสอบที่มีค่าพารามิเตอร์ผ่านเกณฑ์มีค่าอำนาจจำแนกอยู่ในช่วง 0.576 ถึง 1.804 ค่าความยากอยู่ในช่วง  $-2.276$  ถึง 2.212 และมีค่าโอกาสในการเดาข้อสอบได้ถูกอยู่ในช่วง 0.066 ถึง 0.280

**ขั้นตอนที่สอง** สำหรับการวิเคราะห์หาระยะเวลาสูงสุดที่เหมาะสมในการทำข้อสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing – CAT) โดยนำมาเป็นตัวกำหนดระยะเวลาสูงสุดในการทำข้อสอบให้กับโปรแกรมคอมพิวเตอร์ โดยดำเนินการ ดังนี้

1. พิจารณาว่าเงื่อนไขการทดสอบใดที่ใช้เวลาสูงสุดในการทำข้อสอบ (การทดสอบที่ไม่ให้ทวนคำตอบ หรือการทดสอบที่ให้ทวนคำตอบ)
2. นำเวลาสูงสุด (มีหน่วยเป็นวินาที) จากข้อ 1 ที่โปรแกรมคอมพิวเตอร์บันทึกไว้ของกลุ่มตัวอย่างมาหาค่าเฉลี่ย ( $\bar{X}$ ) และส่วนเบี่ยงเบนมาตรฐาน (SD)
3. นำค่าที่ได้จากข้อ 2 นำมาเป็นตัวกำหนดระยะเวลาสูงสุดในการทำข้อสอบให้กับโปรแกรมคอมพิวเตอร์ (Spending Time - มีหน่วยเป็นนาที)

สูตร

$$\text{Spending Time} = \bar{X} + \text{SD} + 5 \text{ นาที}$$

4 ได้ระยะเวลาสูงสุดที่เหมาะสมในการทำข้อสอบในการทดสอบแบบปรับ  
 เหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เท่ากับ 30 นาที โดยมีผลการวิเคราะห์ดัง  
 ตารางที่ 4

**ตารางที่ 4** ระยะเวลาที่นักเรียนกลุ่ม 2 ใช้ในการทำข้อสอบในการทดสอบแบบปรับเหมาะกับความ  
 ความสามารถของผู้สอบด้วยคอมพิวเตอร์ เพื่อกำหนดระยะเวลาสูงสุดในการทำ  
 ข้อสอบให้กับโปรแกรมคอมพิวเตอร์

คะแนนทดสอบ ปริเทสต์ เต็ม 100 คะแนน	ระดับ ความ สามารถ	ระยะเวลาที่ใช้ในการทำข้อสอบ 20 ข้อ (วินาที)		
		เงื่อนไข การไม่ให้ ทวนคำตอบ	เงื่อนไข การให้ ทวนคำตอบ	เงื่อนไข ที่ใช้เวลามากกว่า
91	สูง	894	1068	1068 (ทวน)
85		612	855	855 (ทวน)
80		1066	1245	1245 (ทวน)
76	ปานกลาง	1006	1111	1111 (ทวน)
75		1173	1455	1455 (ทวน)
72		954	1158	1158 (ทวน)
63	ต่ำ	1457	1304	1457 (ไม่ทวน)
61		1275	1315	1315 (ทวน)
50		1164	1374	1374 (ทวน)
	$\bar{X}$	1067	1209	1226
		(17 นาที 47 วินาที)	(20 นาที 9 วินาที)	(20 นาที 26 วินาที)
	SD	242	183	199
		(4 นาที 2 วินาที)	(3 นาที 3 วินาที)	(3 นาที 19 วินาที)
			$\bar{X} + SD + 5$ นาที	1725
				(28 นาที 45 วินาที)
			<b>ระยะเวลาสูงสุด</b>	<b>30 นาที</b>

**หมายเหตุ** 1) ระยะเวลาสูงสุด (Spending Time) =  $\bar{X} + SD + 5$  นาที  
 (มีหน่วยเป็นนาที และทำการประมาณค่าเป็นจำนวนเต็มสิบ)

จากตารางที่ 4 พบว่าโดยภาพรวมระยะเวลาที่ใช้ในการทำข้อสอบวิชา  
 คณิตศาสตร์ เรื่องความเข้าใจเศษส่วน จำนวน 20 ข้อ ในการทดสอบแบบปรับเหมาะกับความ  
 ความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ทางผู้วิจัยสร้างขึ้นนั้น นักเรียนส่วนใหญ่ในทุกๆระดับ



ความสามารถ (สูง ปานกลาง และต่ำ) จะใช้เวลาในการทำข้อสอบในเงื่อนไขการทวนคำตอบ ( $\bar{X} = 1209$  วินาที หรือ 20 นาที 9 วินาที) มากกว่าเงื่อนไขการไม่ให้ทวนคำตอบ ( $\bar{X} = 1067$  วินาที หรือ 17 นาที 47 วินาที) มีเพียง 1 คน ในระดับความสามารถต่ำเท่านั้น ที่ใช้เวลาในการทำข้อสอบในเงื่อนไขการทวนคำตอบ (1304 วินาที หรือ 21 นาที 44 วินาที) น้อยกว่าเงื่อนไขการไม่ให้ทวนคำตอบ (1457 วินาที หรือ 24 นาที 17 วินาที)

จากการวิเคราะห์เงื่อนไขที่ใช้ระยะเวลาในการทำข้อสอบมากกว่าพบว่าอยู่ในช่วง 855 ถึง 1457 วินาที หรือ 14 นาที 15 วินาที ถึง 24 นาที 17 วินาที มีค่าเฉลี่ยเท่ากับ 20 นาที 26 วินาที และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 3 นาที 19 วินาที

จากการคำนวณหาระยะเวลาสูงสุดที่เหมาะสมเพื่อนำไปกำหนดระยะเวลาสูงสุดในการทำข้อสอบให้กับโปรแกรมคอมพิวเตอร์พบว่า ระยะเวลาสูงสุดที่เหมาะสมเท่ากับ 30 นาที

**ขั้นตอนที่สาม** สำหรับการวิเคราะห์เพื่อเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะ กับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing - CAT) เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน ซึ่งแบ่งการวิเคราะห์ตามสถานการณ์ที่สนใจศึกษา คือ

#### สถานการณ์ที่หนึ่ง มีการดำเนินการดังนี้

1.1 ใช้การประมาณผลต่างค่าเฉลี่ยของประชากร 2 กลุ่ม ที่ไม่เป็นอิสระต่อกัน (t-test) โดยใช้โปรแกรม SPSS for Windows เพื่อเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ภายใต้ข้อมูลที่จับคู่ไว้แล้ว คือ

▶▶ ค่าความสามารถของผู้สอบภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ

$(X_1, X_2, X_3, \dots, X_n)$  จากตัวอย่างขนาด  $n$

และ ▶▶ ค่าความสามารถของผู้สอบภายใต้เงื่อนไขการให้ทวนคำตอบ

$(Y_1, Y_2, Y_3, \dots, Y_n)$  จากตัวอย่างขนาด  $n$

ของผู้สอบในแต่ละระดับความสามารถ (สูง, ปานกลาง และต่ำ)

ซึ่งค่าผลต่างรายคู่ ( $d$ ) หาได้จาก

$$d_1 = X_1 - Y_1$$

$$d_2 = X_2 - Y_2$$

.

.

$$d_n = X_n - Y_n$$

ขนาดตัวอย่างในที่นี้จะเท่ากับจำนวนคู่ของข้อมูล เท่ากับ  $n$  นั่นเอง โดยที่  $\bar{d}$  คือ ค่าเฉลี่ยของผลต่างรายคู่ จะมีการแจกแจงที่ (t-distribution) ที่ระดับชั้นความเสรีหรือ df เป็น  $n - 1$  ดังนั้น

$$\Pr \left[ \bar{d} - t_{\left(n-1, \frac{\alpha}{2}\right)} S_{\bar{d}} < \mu_1 - \mu_2 < \bar{d} + t_{\left(n-1, \frac{\alpha}{2}\right)} S_{\bar{d}} \right] = 1 - \alpha$$

และช่วงความเชื่อมั่น  $(1 - \alpha)100\%$  ของ  $\mu_1 - \mu_2$  คือ

$$\bar{d} - t_{\left(n-1, \frac{\alpha}{2}\right)} S_{\bar{d}} < \mu_1 - \mu_2 < \bar{d} + t_{\left(n-1, \frac{\alpha}{2}\right)} S_{\bar{d}}$$

1.2 ใช้การวิเคราะห์ความแปรปรวนทางเดียว (One-Way ANOVA) โดยใช้โปรแกรม SPSS for Windows เพื่อเปรียบเทียบความแตกต่างของความแตกต่างของค่าความสามารถของผู้สอบ ภายใต้ตัวแปรต้น

▶ ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ความผันแปรทั้งหมด (SST) = ความผันแปรระหว่างกลุ่ม (SSTrt) + ความผันแปรภายในกลุ่ม (SSE)

โดยที่ SST = ความผันแปรทั้งหมด (Total Sum Square)

SSB = SSTrt = ความผันแปรระหว่างกลุ่มต่างๆ (Between-Groups Sum Square)

SSW = SSE = ความผันแปรภายในกลุ่มต่างๆ (Within-Groups Sum Square)

การทดสอบสมมติฐาน

$H_0 : \mu_1 = \mu_2 = \dots = \mu_k \quad (k = 3)$

$H_1 : \text{มี } \mu_i \neq \mu_j \text{ อย่างน้อย 1 คู่; } i \neq j$

$$\text{สถิติทดสอบ } F = \frac{\text{MSTrt}}{\text{MSE}} = \frac{\text{MSB}}{\text{MSW}}$$

เขตปฏิเสธ จะปฏิเสธสมมติฐาน  $H_0$  ถ้า  $F > F_{1-\alpha; a-1, ab(m-1)}$  หรือ Sig < ระดับนัยสำคัญ ( $\alpha$ ) ซึ่งการวิเคราะห์ สามารถสรุปได้ดังตารางที่ 5 ดังนี้

ตารางที่ 5 ตารางการวิเคราะห์ความแปรปรวนทางเดียว (1-WAY ANOVA)

แหล่งความแปรผัน	Df	Sum Square	Sum Square	F
ระหว่างกลุ่มต่างๆ	$k - 1$	SSB	MSB	$\frac{MSB}{MSW}$
ภายในกลุ่มต่างๆ	$n - k$	SSW	MSW	MSW
รวม	$n - 1$	SST		

ถ้าหากวิเคราะห์ความแปรปรวนแล้วผลการทดสอบ พบว่าปฏิเสธสมมติฐาน  $H_0$  ให้ทำการทดสอบรายคู่หรือการทดสอบภายหลัง (Post hoc Comparison) ต่อ เพื่อค้นหาว่ามีค่าเฉลี่ยคู่ใดบ้างที่ต่างกัน

กรณี 1 ถ้าทราบว่าประชากรทั้ง  $k$  กลุ่มมีค่าความแปรปรวนไม่แตกต่างกัน (equal variances assumed) ให้ใช้วิธีการของเชฟเฟ (Scheffe')

กรณี 2 ถ้าทราบว่าประชากรทั้ง  $k$  กลุ่มมีค่าความแปรปรวนต่างกัน (equal variances not assumed) ให้ใช้วิธีการของดันเนท ที3 (Dunnett' s T3)

### สถานการณ์ที่สอง มีการดำเนินการดังนี้

ใช้การวิเคราะห์ความแปรปรวนทางเดียว (One-Way ANOVA) โดยใช้โปรแกรม SPSS for Windows เพื่อเปรียบเทียบลักษณะของการเปลี่ยนคำตอบของผู้สอบ (จำนวนข้อที่ไม่มี ความแตกต่างของคำตอบ จำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นถูก จำนวนข้อที่เปลี่ยนคำตอบจากผิด เป็นผิด และจำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นผิด) ภายใต้ตัวแปรต้น

▶▶ ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ซึ่งแบ่งการวิเคราะห์ออกเป็น 4 ครั้ง ตามตัวแปรตาม (ลักษณะของการเปลี่ยน คำตอบของผู้สอบ ซึ่งมี 4 กรณีย่อย) ได้ดังนี้

1) เปรียบเทียบจำนวนข้อที่ไม่มี ความแตกต่างของคำตอบ ภายใต้ตัวแปรต้น

▶▶ ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

2) เปรียบเทียบจำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นถูก ภายใต้ตัวแปรต้น

▶▶ ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

3) เปรียบเทียบจำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นผิด ภายใต้ตัวแปรต้น

▶▶ ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

4) เปรียบเทียบจำนวนข้อที่เปลี่ยนคำตอบจากถูกเป็นผิด ภายใต้ตัวแปรต้น

▶▶ ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ความผันแปรทั้งหมด (SST) = ความผันแปรระหว่างกลุ่ม (SSTrt) + ความผันแปรภายในกลุ่ม (SSE)

โดยที่ SST = ความผันแปรทั้งหมด (Total Sum Square)

SSB = SSTrt = ความผันแปรระหว่างกลุ่มต่างๆ (Between-Groups Sum Square)

SSW = SSE = ความผันแปรภายในกลุ่มต่างๆ (Within-Groups Sum Square)

การทดสอบสมมติฐาน

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k \quad (k = 3)$$

$$H_1 : \text{มี } \mu_i \neq \mu_j \text{ อย่างน้อย 1 คู่; } i \neq j$$

$$\text{สถิติทดสอบ } F = \frac{MSTrt}{MSE} = \frac{MSB}{MSW}$$

เขตปฏิเสธ จะปฏิเสธสมมติฐาน  $H_0$  ถ้า  $F > F_{1-\alpha; a-1, ab(m-1)}$  หรือ  $\text{Sig} < \text{ระดับ}$   
 นัยสำคัญ ( $\alpha$ ) ซึ่งการวิเคราะห์ สามารถสรุปได้ดังตารางที่ 6 ดังนี้

**ตารางที่ 6** ตารางการวิเคราะห์ความแปรปรวนทางเดียว (1-WAY ANOVA)

แหล่งความแปรผัน	Df	Sum Square	Sum Square	F
ระหว่างกลุ่มต่างๆ	$k - 1$	SSB	MSB	$\frac{MSB}{MSW}$
ภายในกลุ่มต่างๆ	$n - k$	SSW	MSW	MSW
<b>รวม</b>	$n - 1$	SST		

ถ้าหากวิเคราะห์ความแปรปรวนแล้วผลการทดสอบ พบว่าปฏิเสธสมมติฐาน  $H_0$  ให้ทำการทดสอบรายคู่หรือการทดสอบภายหลัง (Post hoc Comparison) ต่อ เพื่อค้นหาว่ามีค่าเฉลี่ยคู่ใดบ้างที่ต่างกัน

กรณี 1 ถ้าทราบว่าประชากรทั้ง  $k$  กลุ่มมีค่าความแปรปรวนไม่แตกต่างกัน (equal variances assumed) ให้ใช้วิธีการของเชฟเฟ (Scheffe')

กรณี 2 ถ้าทราบว่าประชากรทั้ง  $k$  กลุ่มมีค่าความแปรปรวนต่างกัน (equal variances not assumed) ให้ใช้วิธีการของดันเนต ที3 (Dunnett' s T3)

**สถานการณ์ที่สาม** มีการดำเนินการดังนี้

1.1 ใช้การประมาณผลต่างค่าเฉลี่ยของประชากร 2 กลุ่ม ที่ไม่เป็นอิสระต่อกัน (t-test) โดยใช้โปรแกรม SPSS for Windows เพื่อเปรียบเทียบความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ภายใต้ข้อมูลที่จับคู่ไว้แล้ว คือ

▶▶ เวลาที่ใช้ในการทำข้อสอบของผู้สอบภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ

$(X_1, X_2, X_3, \dots, X_n)$  จากตัวอย่างขนาด  $n$

และ ▶▶ เวลาที่ใช้ในการทำข้อสอบของผู้สอบภายใต้เงื่อนไขการให้ทวนคำตอบ

$(Y_1, Y_2, Y_3, \dots, Y_n)$  จากตัวอย่างขนาด  $n$

ของผู้สอบในแต่ละระดับความสามารถ (สูง, ปานกลาง และต่ำ)

$$\begin{aligned} \text{ซึ่งค่าผลต่างรายคู่ (d) หาได้จาก} \quad d_1 &= X_1 - Y_1 \\ d_2 &= X_2 - Y_2 \\ &\vdots \\ &\vdots \\ d_n &= X_n - Y_n \end{aligned}$$

ขนาดตัวอย่างในที่นี้จะเท่ากับจำนวนคู่ของข้อมูล เท่ากับ  $n$  นั่นเอง โดยที่  $\bar{d}$  คือค่าเฉลี่ยของผลต่างรายคู่ จะมีการแจกแจงที่ (t-distribution) ที่ระดับชั้นความเสรีหรือ  $df$  เป็น  $n - 1$  ดังนั้น

$$\Pr \left[ \bar{d} - t_{\left(n-1, \frac{\alpha}{2}\right)} S_{\bar{d}} < \mu_1 - \mu_2 < \bar{d} + t_{\left(n-1, \frac{\alpha}{2}\right)} S_{\bar{d}} \right] = 1 - \alpha$$

และช่วงความเชื่อมั่น  $(1 - \alpha)100\%$  ของ  $\mu_1 - \mu_2$  คือ

$$\bar{d} - t_{\left(n-1, \frac{\alpha}{2}\right)} S_{\bar{d}} < \mu_1 - \mu_2 < \bar{d} + t_{\left(n-1, \frac{\alpha}{2}\right)} S_{\bar{d}}$$

1.2 ใช้การวิเคราะห์ความแปรปรวนทางเดียว (One-Way ANOVA) โดยใช้โปรแกรม SPSS for Windows เพื่อเปรียบเทียบความแตกต่างของความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ภายใต้ตัวแปรต้น

▶▶ ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ความผันแปรทั้งหมด (SST) = ความผันแปรระหว่างกลุ่ม (SSTr) + ความผันแปรภายในกลุ่ม (SSE)

โดยที่ SST = ความผันแปรทั้งหมด (Total Sum Square)

SSB = SSTr = ความผันแปรระหว่างกลุ่มต่างๆ (Between-Groups Sum Square)

SSW = SSE = ความผันแปรภายในกลุ่มต่างๆ (Within-Groups Sum Square)

การทดสอบสมมติฐาน

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k \quad (k = 3)$$

$$H_1 : \text{มี } \mu_i \neq \mu_j \text{ อย่างน้อย 1 คู่; } i \neq j$$

$$\text{สถิติทดสอบ } F = \frac{\text{MSTrt}}{\text{MSE}} = \frac{\text{MSB}}{\text{MSW}}$$

เขตปฏิเสธ จะปฏิเสธสมมติฐาน  $H_0$  ถ้า  $F > F_{1-\alpha; a-1, ab(m-1)}$  หรือ  $\text{Sig} < \text{ระดับ}$   
นัยสำคัญ ( $\alpha$ ) ซึ่งการวิเคราะห์ สามารถสรุปได้ดังตารางที่ 7 ดังนี้

ตารางที่ 7 ตารางการวิเคราะห์ความแปรปรวนทางเดียว (1-WAY ANOVA)

แหล่งความแปรผัน	Df	Sum Square	Sum Square	F
ระหว่างกลุ่มต่างๆ	$k - 1$	SSB	MSB	$\frac{\text{MSB}}{\text{MSW}}$
ภายในกลุ่มต่างๆ	$n - k$	SSW	MSW	MSW
รวม	$n - 1$	SST		

ถ้าหากวิเคราะห์ความแปรปรวนแล้วผลการทดสอบ พบว่าปฏิเสธสมมติฐาน  $H_0$  ให้ทำการทดสอบรายคู่หรือการทดสอบภายหลัง (Post hoc Comparison) ต่อ เพื่อค้นหาว่ามีค่าเฉลี่ยคู่ใดบ้างที่ต่างกัน

กรณี 1 ถ้าทราบว่าประชากรทั้ง  $k$  กลุ่มมีค่าความแปรปรวนไม่แตกต่างกัน (equal variances assumed) ให้ใช้วิธีการของเชฟเฟ (Scheffe')

กรณี 2 ถ้าทราบว่าประชากรทั้ง  $k$  กลุ่มมีค่าความแปรปรวนต่างกัน (equal variances not assumed) ให้ใช้วิธีการของดันเนท ที3 (Dunnett' s T3)

## บทที่ 4

### ผลการพัฒนาโปรแกรมและผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบค่าความสามารถของผู้สอบ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน โดยผู้วิจัยเสนอผลการวิเคราะห์ข้อมูลเป็น 4 ตอน ตามลำดับ ดังนี้

ตอนที่ 1 ผลการพัฒนาโปรแกรมคอมพิวเตอร์

ตอนที่ 2 ผลการเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ

ตอนที่ 3 ผลการเปรียบเทียบลักษณะของการเปลี่ยนคำตอบของผู้สอบ

ตอนที่ 4 ผลการเปรียบเทียบความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ

#### ตอนที่ 1 ผลการพัฒนาโปรแกรมคอมพิวเตอร์

ผลการพัฒนาโปรแกรมคอมพิวเตอร์วิชาคณิตศาสตร์ เรื่อง ความเข้าใจเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ ที่สามารถกำหนดเงื่อนไขการทดสอบได้ 2 เงื่อนไข คือ การทดสอบที่ให้ทวนคำตอบได้โดยวิธีการทวนที่มีการจำกัดช่วง (Block Review) กับการทดสอบที่ไม่ให้ทวนคำตอบ ผู้วิจัยแบ่งขั้นตอนการนำโปรแกรมไปใช้งานออกเป็น 3 ตอนย่อยดังนี้

1.1 การเตรียมความพร้อมของคอมพิวเตอร์ก่อนการติดตั้งโปรแกรม

1.2 ขั้นตอนการติดตั้งโปรแกรม

1.3 ผลการพัฒนาโปรแกรม

#### 1.1 การเตรียมความพร้อมของคอมพิวเตอร์ก่อนการติดตั้งโปรแกรม

คุณสมบัติของเครื่องคอมพิวเตอร์ที่จะติดตั้งโปรแกรม

1. ใช้ระบบปฏิบัติการ Microsoft Windows98 หรือสูงกว่า
2. ติดตั้งโปรแกรม .NET Framework Version 1.1 หรือสูงกว่า
3. มีพื้นที่ Harddisk ว่างสูงกว่า 10 เมกะไบต์
4. มีหน่วยความจำอย่างน้อย 64 เมกะไบต์

## 1.2 ขั้นตอนการติดตั้งโปรแกรม

1. เปิด Folder ชื่อว่า CAT จากแผ่น CD โปรแกรมทดสอบแบบปรับเหมาะของผู้วิจัย
2. สร้าง Folder ใหม่ที่ Drive C: ชื่อ CAT
3. คัดลอก CATProject.EXE จากแผ่น CD ไปวางใน Folder CAT ในเครื่องคอมพิวเตอร์
4. Double Click CATProject.EXE เพื่อเริ่มการใช้งาน

## 1.3 ผลการพัฒนาโปรแกรม

### 1.3.1 หน้าจอแสดงข้อมูลวิทยานิพนธ์

แสดงข้อมูลของลักษณะโปรแกรมและข้อมูลผู้พัฒนา ให้กดปุ่มเริ่มเพื่อทำการเริ่มต้นการทดสอบ



แผนภาพที่ 18 แสดงข้อมูลวิทยานิพนธ์

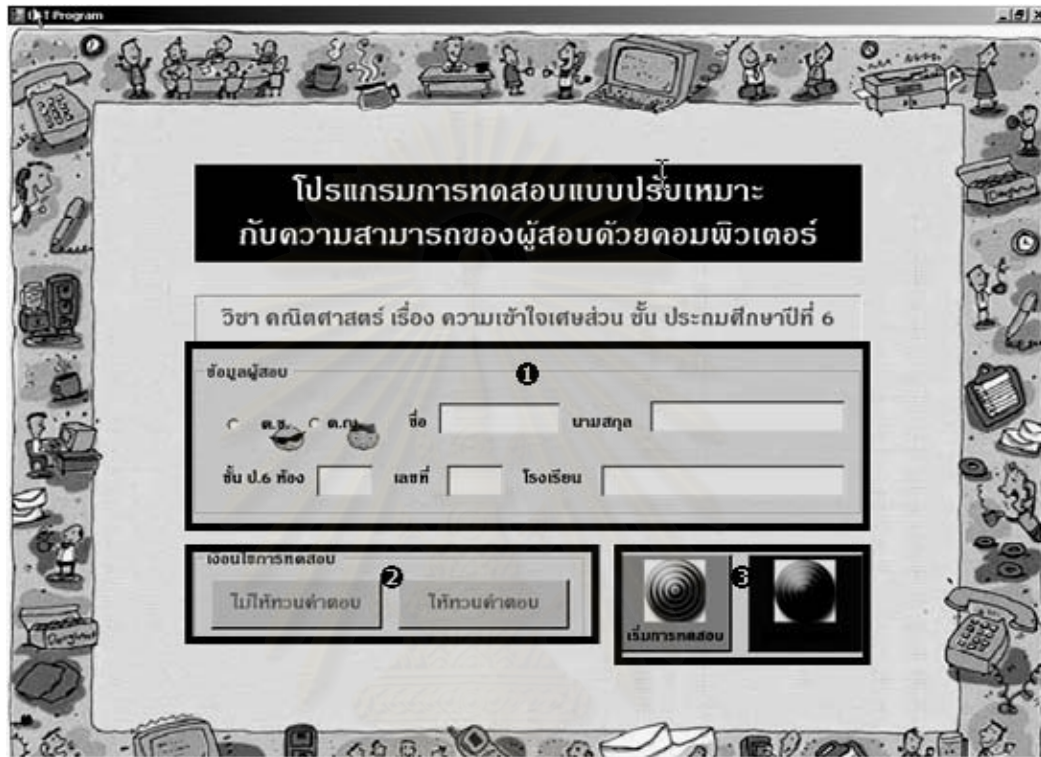
### 1.3.2 หน้าจอบันทึกข้อมูลผู้ทำการทดสอบ

ลักษณะการแสดงผล แบ่งเป็น 3 ส่วน ดังนี้

1. ข้อมูลผู้สอบ ประกอบด้วย เพศ, ชื่อ, นามสกุล, ห้องและเลขที่ของนักเรียน



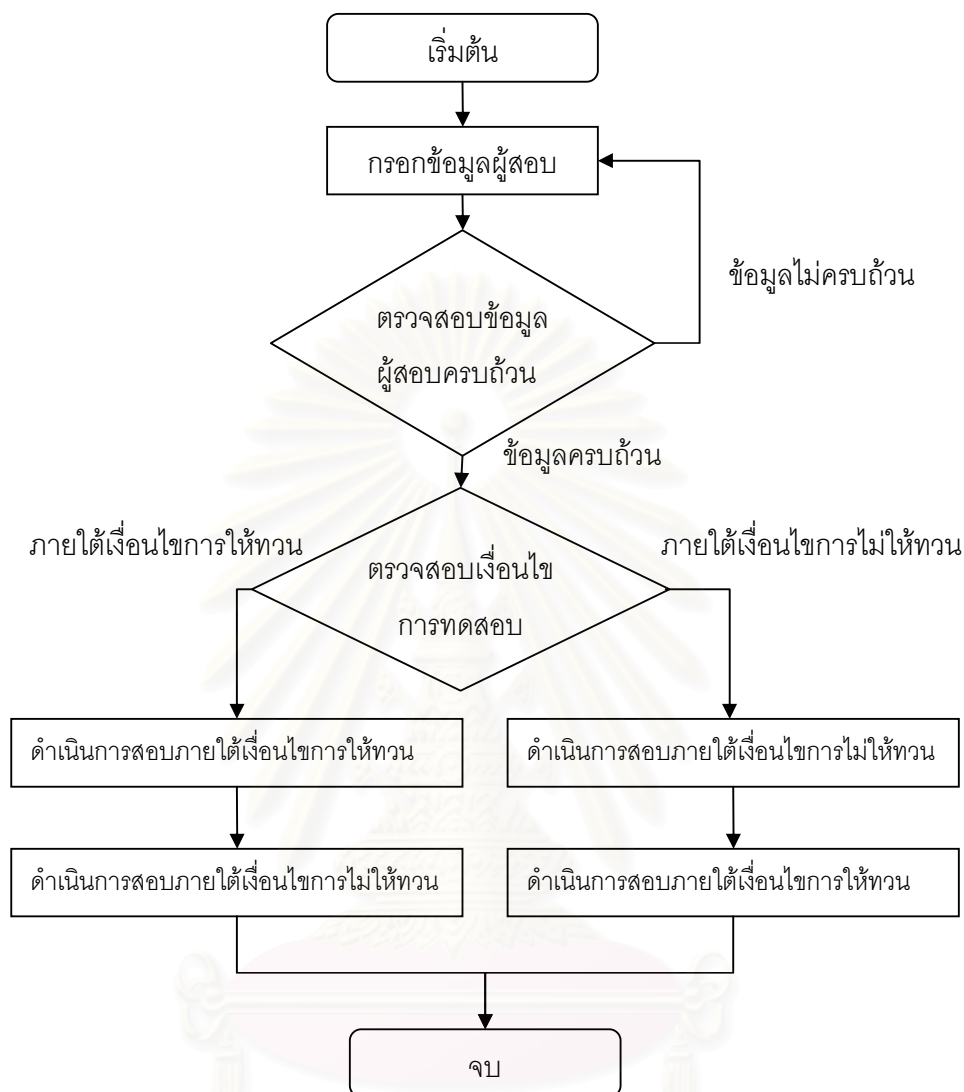
2. เงื่อนไขการทดสอบว่าจะดำเนินการทดสอบภายใต้เงื่อนไขการให้ทวนคำตอบก่อน หรือหลังการทดสอบภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ
3. กดปุ่มเริ่มการทำสอบ (สีเขียว) และปุ่มออกจากระบบ (สีแดง)



แผนภาพที่ 19 แสดงหน้าจอสำหรับกรอกข้อมูลผู้สอบ และเลือกเงื่อนไขการทดสอบ

### ผังงานขั้นตอนการทำงาน

1. เริ่มต้น
2. กรอกข้อมูลผู้สอบ
3. ตรวจสอบข้อมูลผู้สอบ ถ้าไม่ครบถ้วนให้ย้อนกลับไปทำข้อ 2
4. ตรวจสอบเงื่อนไขการทดสอบ
5. ผู้สอบสามารถดำเนินการสอบได้ 2 วิธีคือ
  - a. ดำเนินการสอบภายใต้เงื่อนไขการไม่ให้ทวนคำตอบจากนั้น ดำเนินการสอบภายใต้เงื่อนไขการให้ทวนคำตอบ
  - b. ดำเนินการสอบภายใต้เงื่อนไขการให้ทวนคำตอบจากนั้น ดำเนินการสอบภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ
6. จบการทดสอบ



แผนภาพที่ 20 ผังงานแสดงขั้นตอนการทำงาน ในส่วนหน้าจอบันทึกข้อมูลผู้ทำการสอบ

### 1.3.3 หน้าจอโจทย์ทดสอบภายใต้เงื่อนไขการไม่ให้ทุนคำตอบ

ลักษณะการแสดงผล แบ่งเป็น 6 ส่วน ดังนี้

1. หมายเลขลำดับข้อ
2. คำชี้แจง
3. โจทย์
4. เวลาที่เหลือในการทดสอบ
5. คำตอบที่เลือก
6. ยืนยันคำตอบเพื่อทำข้อถัดไป



แผนภาพที่ 21 แสดงหน้าจอสำหรับการทดสอบภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ

### ผังงานขั้นตอนการทำงาน

1. เริ่มต้น
2. กำหนดค่าความสามารถผู้สอบเริ่มต้น 0.5
3. สุ่มเลือกข้อสอบข้อแรกจากในคลังข้อสอบ
4. คำนวณค่าความสามารถของผู้สอบ และค่าความแปรปรวนในการประมาณค่าความสามารถ โดย
  - a. ในกรณีที่ตอบถูก

$$\theta_{m+1} = \theta_m + (1-c) \left( \frac{\sigma_m^2}{\sqrt{\frac{1}{a^2} + \sigma_m^2}} \right) \left( \frac{O(D)}{c + (1-c)A(-D)} \right)$$

$$\sigma_{m+1}^2 = \sigma_m^2 \left( 1 - \left( \frac{1-c}{1 + \frac{1}{a^2 \sigma_m^2}} \right) \left( \frac{O(D)}{B} \right) \left( \frac{(1-c)O(D)}{B} - D \right) \right)$$

เมื่อ  $D = \frac{b - \theta_m}{\sqrt{\frac{1}{a^2} + \sigma_m^2}}$  และ  $B = c + (1 - c) \times A(-D)$

b. ในกรณีที่ตอบผิด

$$\theta_{m+1} = \theta_m - \left( \frac{\sigma_m^2}{\sqrt{\frac{1}{a^2} + \sigma_m^2}} \right) \left( \frac{O(D)}{A(D)} \right)$$

$$\sigma_{m+1}^2 = \sigma_m^2 \left( 1 - \left( \frac{O(D)}{1 + \frac{1}{a^2 \sigma_m^2}} \right) \left( \frac{O(D)}{A(D)} + D \right) \div A(D) \right) \quad (\text{Owen, 1975})$$

โดยที่

$\theta_m$  แทน ความสามารถของผู้สอบที่ประมาณได้ก่อนตอบข้อสอบ ข้อที่  $m+1$  ตามปกติแล้ว  
ถ้าไม่ทราบค่าความสามารถเบื้องต้นของผู้สอบ ก็จะกำหนดให้  $\theta_m$  เท่ากับ 0.000

$\sigma_m^2$  แทน ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อสอบข้อที่  
 $m+1$  ตามปกติแล้วถ้าไม่ทราบค่าความแปรปรวนดังกล่าวมาก่อน ก็จะกำหนดให้  $\delta_m^2$  เท่ากับ  
1.000

$\theta_{m+1}$  แทน ความสามารถของผู้สอบที่ประมาณได้หลังจากตอบข้อสอบข้อที่  $m+1$  แล้ว

$\sigma_{m+1}^2$  แทน ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบหลังจากตอบข้อสอบ  
ข้อที่  $m+1$  แล้ว

a แทน ค่าอำนาจจำแนกของข้อสอบข้อที่  $m+1$

b แทน ค่าความยากของข้อสอบข้อที่  $m+1$

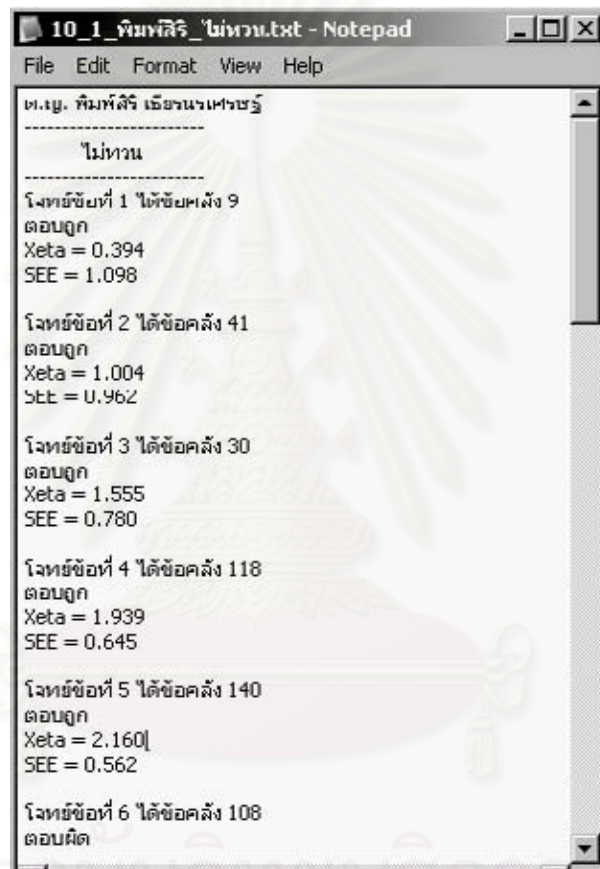
c แทน ค่าโอกาสในการเดาข้อสอบถูกของข้อสอบข้อที่  $m+1$

D แทน จุดบนแกน X

O(D) แทน ค่าออร์ดิเนต (Ordinate) ของโค้งปกติที่จุด D

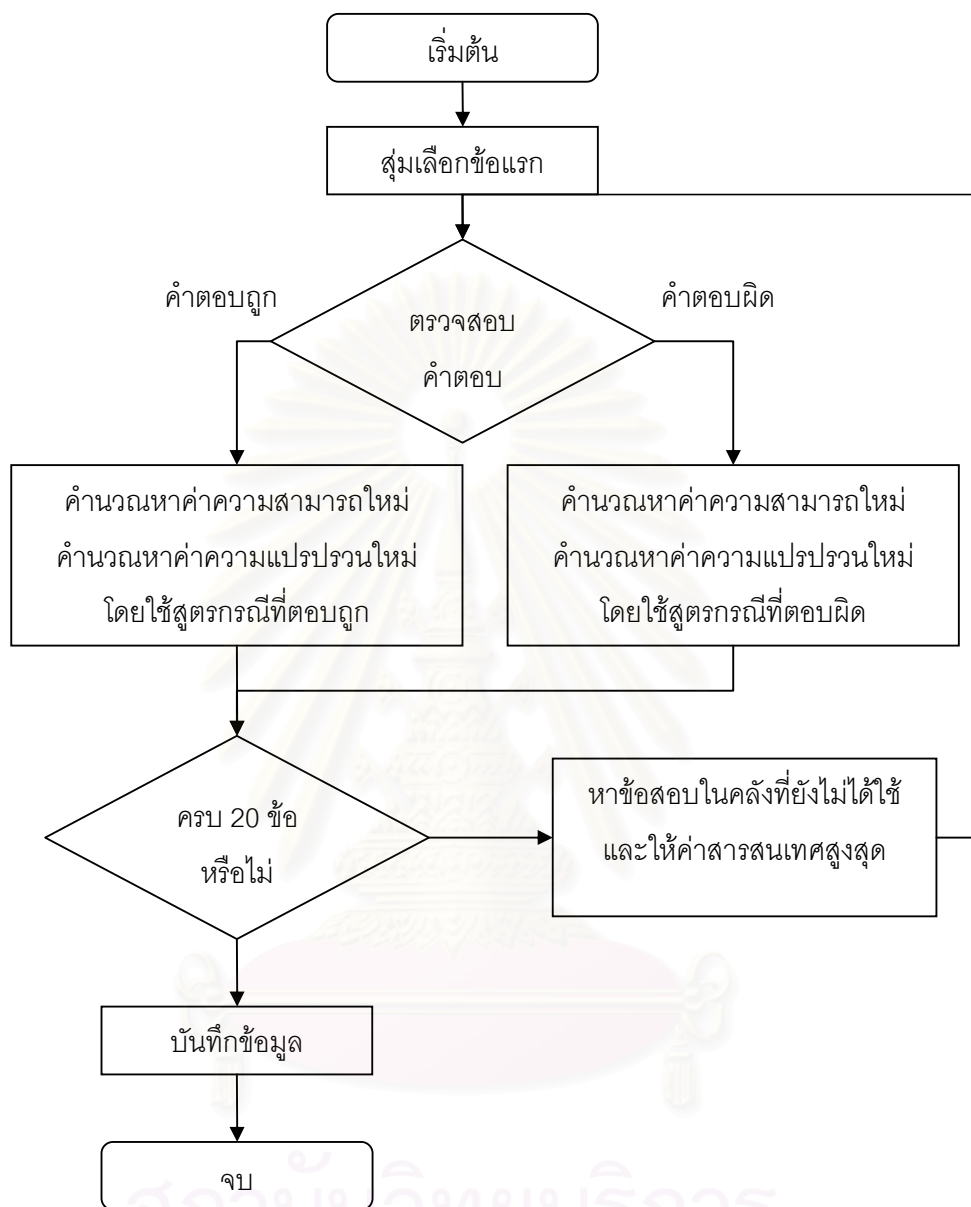
A(D) แทน พื้นที่ใต้โค้งปกติจากค่า D ต่ำสุดจนถึงจุด D

5. ตรวจสอบข้อสอบว่าทำครบ 20 ข้อหรือไม่ถ้ายังไม่ครบให้ทำข้อ 6 ถ้าครบแล้วให้ทำข้อ 7
6. เลือกข้อสอบข้อถัดไปจากคลังข้อสอบที่ให้ค่าสารสนเทศสูงสุดจากสูตรแล้วย้อนกลับไปทำข้อ 4
7. ทำการบันทึกข้อมูลผลการสอบลงในเครื่องคอมพิวเตอร์ ผลลัพธ์จะอยู่ใน Folder C:\CAT\Result\ โครงสร้างชื่อไฟล์มีลักษณะ ดังนี้ AA\_BB\_CC\_ไม่ทวน.txt  
AA แทนเลขที่, BB แทนห้อง, CC แทนชื่อ



แผนภาพที่ 22 แสดงผลลัพธ์ของการทดสอบภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ

8. จบการทำงาน



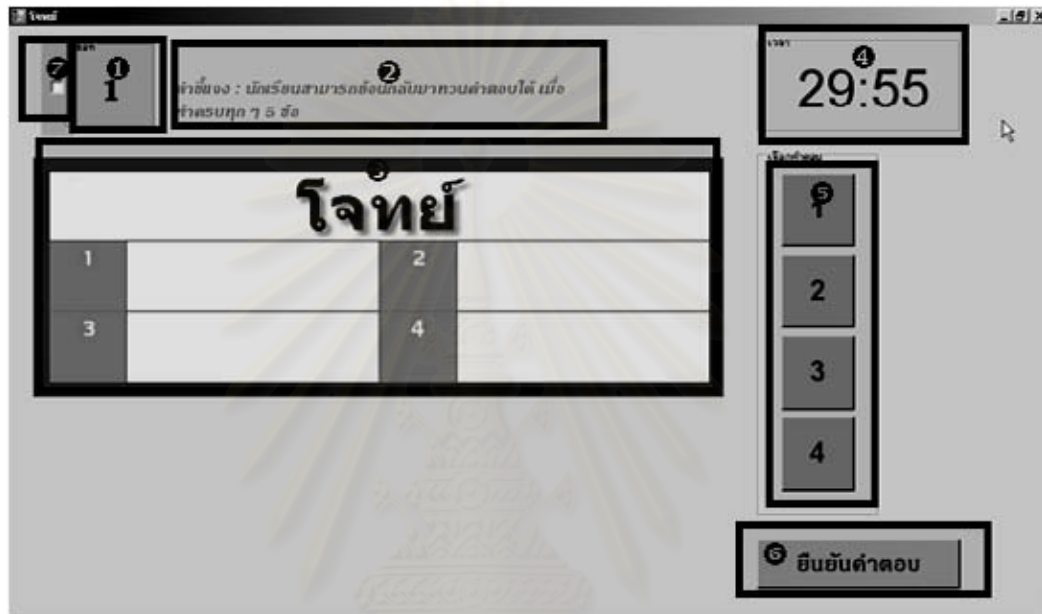
แผนภาพที่ 23 ผังงานแสดงขั้นตอนการทำงานของ การทดสอบ  
ภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ

### 1.3.4 หน้าจอโจทย์ทดสอบภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ

ลักษณะการแสดงผล แบ่งเป็น 7 ส่วน ดังนี้

1. หมายเลขลำดับข้อ
2. คำชี้แจง

3. โจทย์
4. เวลาที่เหลือในการทดสอบ
5. คำตอบที่เลือก
6. ยืนยันคำตอบเพื่อทำข้อถัดไป
7. ทำสัญลักษณ์เพื่อช่วยเตือนในการทวน

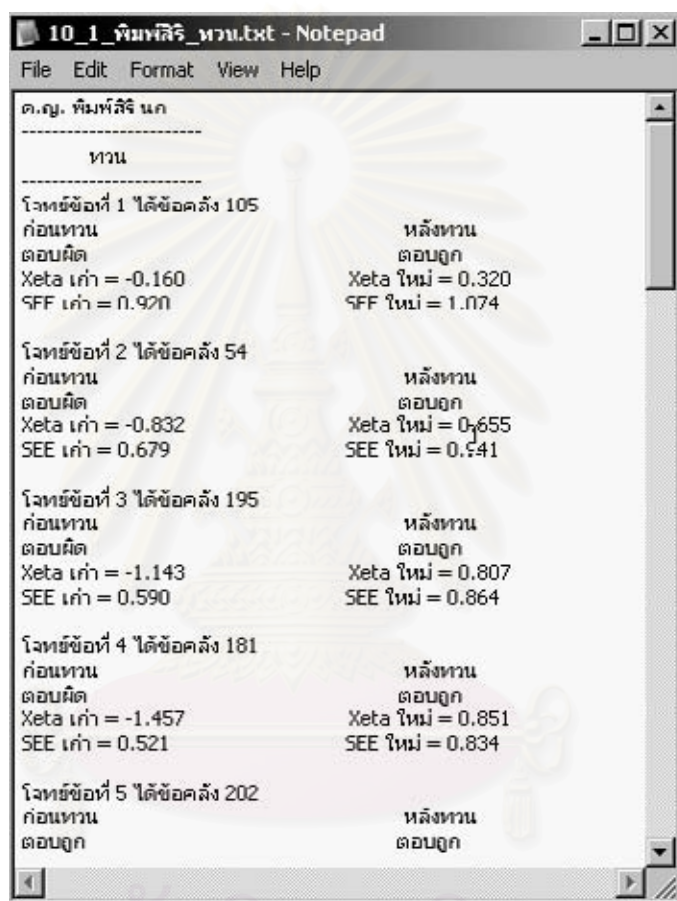


แผนภาพที่ 24 แสดงหน้าจอสำหรับการทดสอบภายใต้เงื่อนไขการให้ทวนคำตอบ

### ผังงานขั้นตอนการทำงาน

1. เริ่มต้น
2. กำหนดค่าความสามารถผู้สอบเริ่มต้น 0.5
3. สุ่มเลือกข้อสอบข้อแรกจากในคลังข้อสอบ
4. ตรวจเช็คคำตอบ (เหมือนในกรณีภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ)
5. ตรวจสอบข้อสอบว่าทำครบ 5 ข้อหรือไม่ ถ้ายังให้ข้ามไปทำข้อ 8
6. ทำการทวนคำตอบ จากนั้นคำนวณค่าความสามารถและค่าความแปรปรวนของ Block ใหม่
7. ตรวจสอบข้อสอบว่าทำครบ 20 ข้อไม่ถ้ายังไม่ครบให้ทำข้อ 8 ถ้าครบแล้วให้ทำข้อ 9
8. เลือกข้อสอบข้อถัดไปจากคลังข้อสอบที่ให้ค่าสารสนเทศสูงสุดจากสูตรแล้วย้อนกลับไปทำข้อ 4

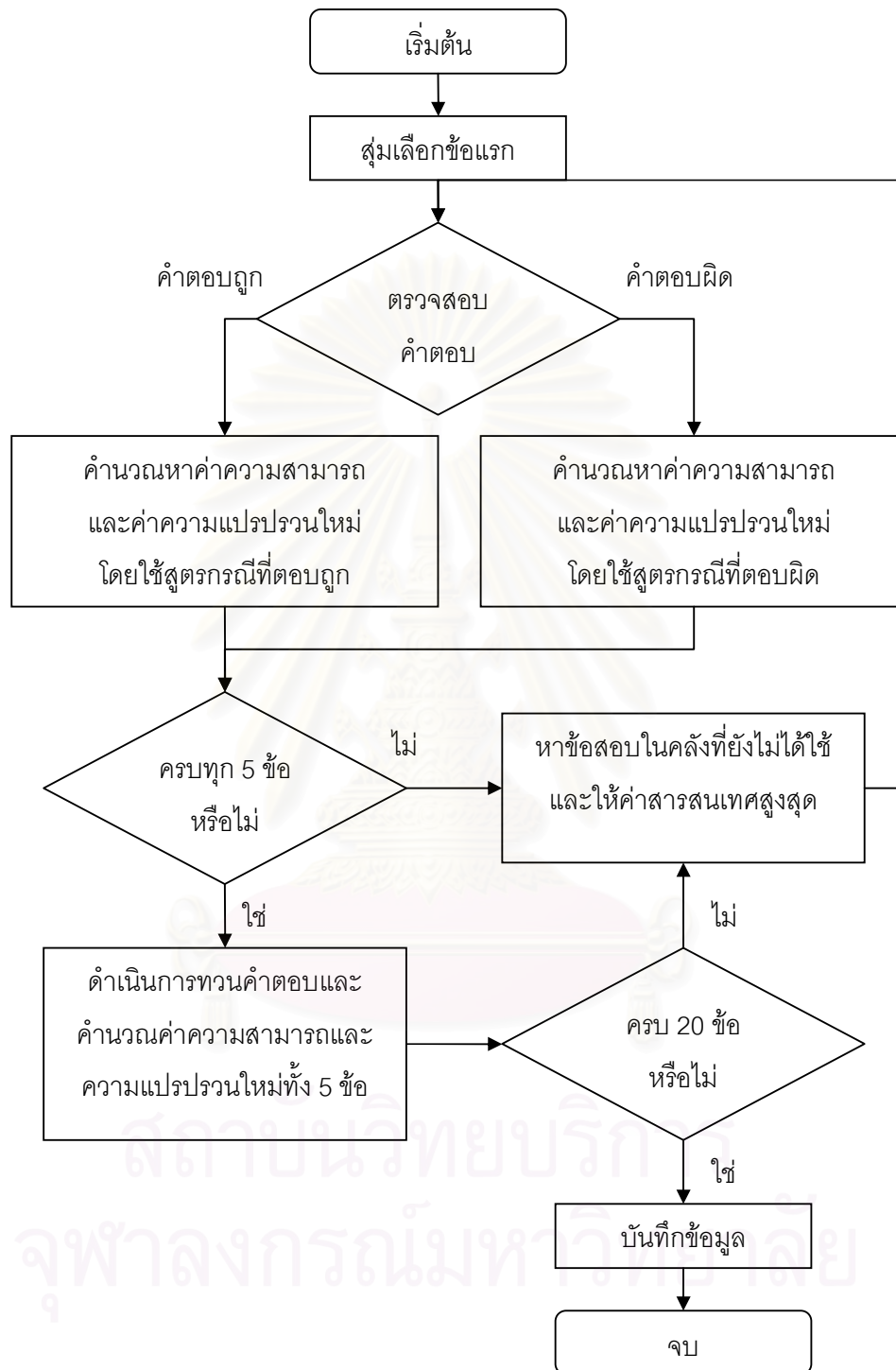
9. ทำการบันทึกข้อมูลผลการทดลองในเครื่องคอมพิวเตอร์ ผลลัพธ์จะอยู่ใน Folder C:\CAT\Result\ โครงสร้างชื่อไฟล์มีลักษณะ ดังนี้ AA\_BB\_CC\_ทวน.txt
- AA แทนเลขที่
- BB แทนห้อง
- CC แทนชื่อ



แผนภาพที่ 25 แสดงผลลัพธ์ของการทดสอบภายใต้เงื่อนไขการให้ทวนคำตอบ

10. จบการทำงาน





แผนภาพที่ 26 ผังงานแสดงขั้นตอนการทำงานของ การทดสอบภายใต้เงื่อนไขการให้ทวนคำตอบ

### 1.3.5 หน้าจอการเลือกข้อที่ต้องการกลับไปทวน

ลักษณะการแสดงผล แบ่งเป็น 3 ส่วน ดังนี้

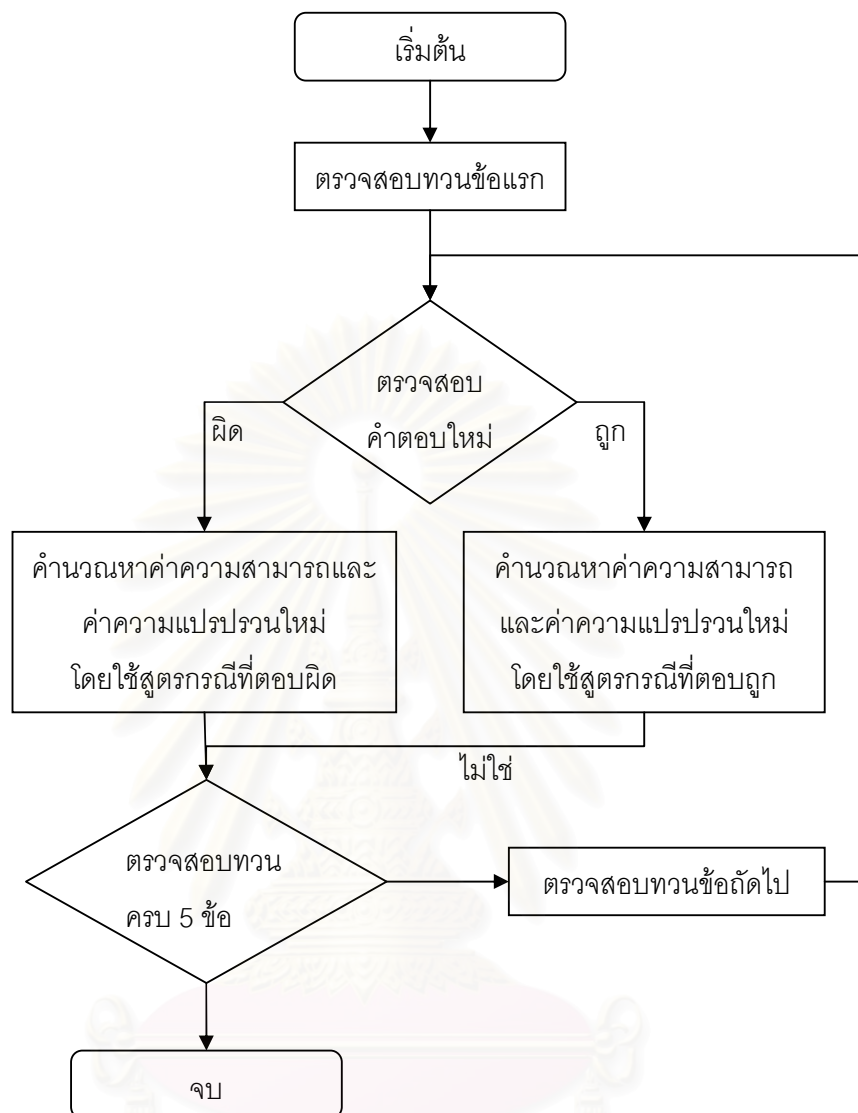
1. เครื่องหมายถูกที่กำกับหน้าข้อ จะมีเครื่องหมายถูกตามที่ถูกทดสอบระบุไว้
2. ปุ่มกดเพื่อกลับไปทวนข้อที่ต้องการ
3. ยืนยันคำตอบทั้งหมดเพื่อทำข้อสอบช่วงถัดไป



แผนภาพที่ 27 แสดงหน้าจอสำหรับการกลับไปทวนคำตอบ

#### ผังงานขั้นตอนการทำงาน

1. เริ่มต้น (เมื่อกดปุ่มยืนยันคำตอบทั้งหมด)
2. ตรวจสอบทวนข้อแรก
3. ตรวจเช็คคำตอบ (เหมือนในกรณีภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ)
4. บันทึกค่าความสามารถและค่าความแปรปรวนใหม่
5. ตรวจสอบข้อสอบว่าทำครบ 5 ข้อหรือไม่ ถ้ายังให้ตรวจสอบข้อถัดไป แล้วย้อนไปทำข้อ 3
6. จบ

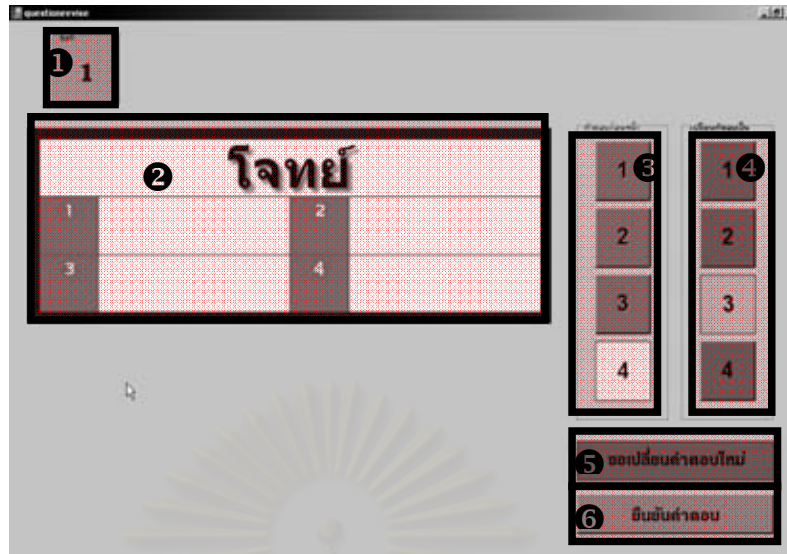


แผนภาพที่ 28 ผังงานแสดงขั้นตอนการทำงานของกรทวนคำตอบ

### 1.3.6 หน้าจอการเปลี่ยนคำตอบ

ลักษณะการแสดงผล แบ่งเป็น 6 ส่วน ดังนี้

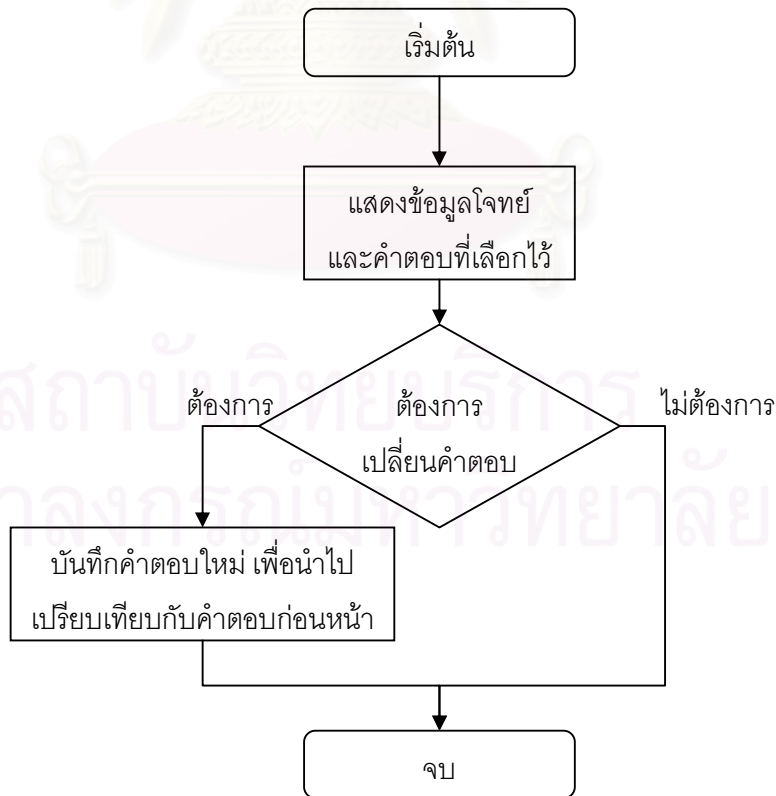
1. หมายเลขลำดับข้อ
2. โจทย์
3. คำตอบก่อนหน้า
4. คำตอบใหม่
5. ปุ่มขอเปลี่ยนคำตอบ
6. ยืนยันคำตอบ



แผนภาพที่ 29 แสดงหน้าจอสำหรับการเปลี่ยนคำตอบ

### ผังงานขั้นตอนการทำงาน

1. เริ่มต้น
2. ถ้าต้องการเปลี่ยนคำตอบให้บันทึกคำตอบก่อนหน้าและคำตอบที่เปลี่ยนเพื่อนำไปวิเคราะห์
3. จบ



แผนภาพที่ 30 ผังงานแสดงขั้นตอนการทำงานของ การเปลี่ยนคำตอบ

### 1.3.7 คลังข้อสอบ

#### ส่วนแสดงผล

จากการวิเคราะห์ผู้วิจัยได้เลือกใช้รูปภาพชนิด JPEG มาทำเป็นรูปแสดงข้อมูลข้อสอบ โดยมีส่วนหัวเป็นโจทย์ และด้านล่างเป็นตารางสี่เหลี่ยม 2 คอลัมน์ 2 บรรทัด บรรจุกุตัวเลือกทั้ง 4 ข้อ มีลักษณะดังแผนภาพที่ 31

โจทย์			
1	ข้อมูลตัวเลือก 1	2	ข้อมูลตัวเลือก 2
3	ข้อมูลตัวเลือก 3	4	ข้อมูลตัวเลือก 4

แผนภาพที่ 31 รูปตัวอย่างโจทย์

#### ส่วนเก็บคุณสมบัติข้อสอบ

จากการวิเคราะห์ผู้วิจัยได้เลือกภาษา XML มาช่วยใช้ในการเก็บข้อมูล ซึ่งปัจจุบันเป็นที่นิยมแพร่หลายและงานต่อการใช้งาน อีกทั้งยังสามารถทำความเข้าใจได้ง่าย โดยในฐานข้อมูลมีการเก็บลำดับข้อในคลัง ค่า 3-Parameter (a,b,c) และตัวเลือกที่ถูกต้องเก็บไว้ มีลักษณะดังภาพ

```

- <Question>
  <No>1</No>
  <Avalue>1.089</Avalue>
  <Bvalue>0.535</Bvalue>
  <Cvalue>0.155</Cvalue>
  <Answer>4</Answer>
</Question>
- <Question>
  <No>2</No>
  <Avalue>1.064</Avalue>
  <Bvalue>0.667</Bvalue>
  <Cvalue>0.152</Cvalue>
  <Answer>3</Answer>
</Question>
- <Question>
  <No>3</No>
  <Avalue>0.856</Avalue>
  <Bvalue>2.214</Bvalue>
  <Cvalue>0.268</Cvalue>
  <Answer>2</Answer>
</Question>

```

แผนภาพที่ 32 ไฟล์ข้อมูลคลังข้อสอบ

## ตอนที่ 2 ผลการเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ

การวิเคราะห์ผลการเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ตามเงื่อนไขการทดสอบ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน ได้ผลการวิเคราะห์ดังตารางที่ 8

**ตารางที่ 8** จำนวนผู้สอบ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ของค่าความสามารถของผู้สอบ และผลของการเปรียบเทียบค่าความสามารถของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบ ของผู้สอบในแต่ละระดับความสามารถ

ระดับ ความสามารถ ของผู้สอบ	จำนวน ผู้สอบ (คน)	ค่าความสามารถของผู้สอบ ( $\theta$ )								t	Sig.
		เงื่อนไขการไม่ให้ทวน คำตอบ				เงื่อนไขการให้ทวน คำตอบ					
		Min	Max	$\bar{X}$	SEE	Min	Max	$\bar{X}$	SEE		
สูง $\theta \geq 0.7$	72	0.701	2.605	1.383	0.232	-0.688	3.127	1.414	0.227	0.416	.679
ปานกลาง $-0.7 < \theta < 0.7$	67	-0.688	0.652	-0.036	0.264	-1.407	2.154	0.189	0.261	2.661*	.010
ต่ำ $\theta \leq -0.7$	61	-2.308	-0.715	-1.274	0.298	-2.555	1.117	-0.856	0.281	5.333*	.000
<b>รวม</b>	200	-2.308	2.605	0.097	0.261	-2.555	3.127	0.312	0.256		

\*p < .05

จากตารางที่ 8 ค่าความสามารถของผู้สอบในเงื่อนไขการไม่ให้ทวนคำตอบ พบว่าผู้สอบที่อยู่ในระดับความสามารถสูงจะมีค่าความสามารถอยู่ในช่วง 0.701 ถึง 2.605 ผู้สอบที่อยู่ในระดับความสามารถปานกลางจะมีค่าความสามารถอยู่ในช่วง -0.688 ถึง 0.652 และผู้สอบที่อยู่ในระดับความสามารถต่ำจะมีค่าความสามารถอยู่ในช่วง -2.308 ถึง -0.715

ค่าความสามารถของผู้สอบในเงื่อนไขการให้ทวนคำตอบ พบว่าผู้สอบที่อยู่ในระดับความสามารถสูงจะมีค่าความสามารถอยู่ในช่วง -0.688 ถึง 3.127 ผู้สอบที่อยู่ในระดับความสามารถปานกลางจะมีค่าความสามารถอยู่ในช่วง -1.407 ถึง 2.154 และผู้สอบที่อยู่ในระดับความสามารถต่ำจะมีค่าความสามารถอยู่ในช่วง -2.555 ถึง 1.117

ค่าเฉลี่ยของค่าความสามารถของผู้สอบ พบว่าค่าเฉลี่ยของค่าความสามารถของผู้สอบรวมทุกระดับความสามารถในเงื่อนไขการให้ทวนคำตอบ ( $\bar{X} = 0.312$ ) มีค่ามากกว่าเงื่อนไขการไม่ให้ทวนคำตอบ ( $\bar{X} = 0.097$ )

ผลการเปรียบเทียบค่าเฉลี่ยของค่าความสามารถของผู้สอบ ระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบ ของผู้สอบที่มีระดับความสามารถสูง พบว่าผู้สอบที่มีระดับความสามารถสูง มีค่าเฉลี่ยของค่าความสามารถของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และผู้สอบที่มีระดับความสามารถปานกลาง และต่ำ พบว่าผู้สอบที่มีระดับความสามารถปานกลาง และต่ำ มีค่าเฉลี่ยของค่าความสามารถของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ พบว่าในเงื่อนไขการให้ทวนคำตอบมีค่าน้อยกว่าเงื่อนไขการไม่ให้ทวนคำตอบ ในทุกระดับความสามารถ (สูง, ปานกลาง และต่ำ) จึงส่งผลให้ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบรวมทุกระดับความสามารถในเงื่อนไขการให้ทวนคำตอบ ( $SEE = 0.256$ ) มีค่าน้อยกว่าเงื่อนไขการไม่ให้ทวนคำตอบ ( $SEE = 0.261$ )

**ตารางที่ 9** จำนวนผู้สอบ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด และค่าต่ำสุด ของความแตกต่างของค่าความสามารถของผู้สอบ จำแนกตามระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ระดับ ความสามารถ ของผู้สอบ	จำนวน นักเรียน	ความแตกต่างของค่าความสามารถของผู้สอบ ( $\theta_{dif}$ )			
		Min	Max	$\bar{X}$	SD
สูง	72	-1.901	1.294	0.031	0.637
ปานกลาง	67	-1.375	1.659	0.225	0.692
ต่ำ	61	-1.048	2.621	0.418	0.679
<b>รวม</b>	<b>200</b>	<b>-1.901</b>	<b>2.621</b>	<b>0.214</b>	<b>0.684</b>

**หมายเหตุ**

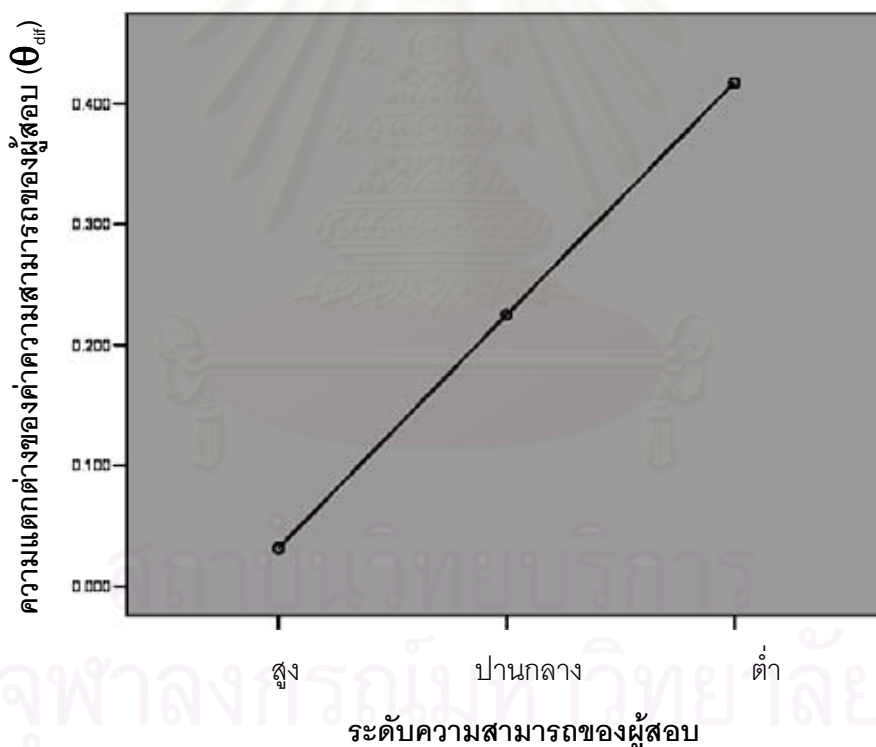
1)  $\theta_{dif} = \theta_{ทวน} - \theta_{ไม่ทวน}$

2) กรณีที่  $\theta_{dif}$  มีค่าเป็นบวก แสดงว่า  $\theta_{ทวน}$  มีค่ามากกว่า  $\theta_{ไม่ทวน}$

$\theta_{dif}$  มีค่าเป็นลบ แสดงว่า  $\theta_{ทวน}$  มีค่าน้อยกว่า  $\theta_{ไม่ทวน}$

จากตารางที่ 9 พบว่า เงื่อนไขการให้ทวนคำตอบส่งผลให้ผู้สอบในระดับความสามารถสูงมีความแตกต่างของค่าความสามารถของผู้สอบ อยู่ในช่วง -1.901 ถึง 1.294 ผู้สอบระดับความสามารถปานกลางอยู่ในช่วง -1.375 ถึง 1.659 และผู้สอบระดับความสามารถต่ำอยู่ในช่วง -1.048 ถึง 2.621 โดยที่ความแตกต่างของค่าความสามารถของผู้สอบที่ต่ำที่สุดอยู่ในกลุ่มผู้สอบที่มีความสามารถระดับสูง ( $\theta_{dif} = -1.901$ ) และความแตกต่างของค่าความสามารถของผู้สอบที่สูงที่สุดอยู่ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ ( $\theta_{dif} = 2.621$ )

ผู้สอบในทุกระดับความสามารถมีค่าเฉลี่ยความแตกต่างของค่าความสามารถเป็นค่าบวก ( $\theta_{dif} = 0.214$ ) ซึ่งแสดงว่าผู้สอบส่วนใหญ่จะได้ค่าความสามารถในเงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบ โดยที่ผู้สอบที่มีระดับความสามารถต่ำจะได้ค่าความสามารถในเงื่อนไขการให้ทวนคำตอบเพิ่มขึ้นจากเงื่อนไขการไม่ให้ทวนคำตอบมากที่สุด ( $\theta_{dif} = 0.418$ ) รองลงมาคือผู้สอบที่มีระดับความสามารถปานกลาง ( $\theta_{dif} = 0.225$ ) และน้อยที่สุดคือผู้สอบที่มีระดับความสามารถสูง ( $\theta_{dif} = 0.031$ )



แผนภาพที่ 33 กราฟแสดงแนวโน้มของความแตกต่างของค่าความสามารถของผู้สอบ ( $\theta_{dif}$ ) ของผู้สอบที่มีระดับความสามารถแตกต่างกัน

จากแผนภาพที่ 33 พบว่าความแตกต่างของค่าความสามารถของผู้สอบ ( $\theta_{dif}$ ) มีแนวโน้มสูงขึ้น (มีค่าเป็นบวกมากขึ้น) เมื่อผู้สอบมีระดับความสามารถต่ำลง



**ตารางที่ 10** ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนด้วย Levene's test และผลการวิเคราะห์ความแปรปรวนทางเดียว ของความแตกต่างของค่าความสามารถของผู้สอบ ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน

แหล่งความแปรปรวน	df	SS	MS	F	Sig.
ระหว่างกลุ่ม	2	4.943	2.472	5.527*	0.005
ภายในกลุ่ม	197	88.102	0.447		
<b>รวม</b>	199	93.045			

Levene Statistic	df1	df2	Sig.
0.624	2	197	0.537

\*p < .05

จากตารางที่ 10 ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนของความแตกต่างของค่าความสามารถของผู้สอบ ระหว่างผู้สอบในแต่ละระดับความสามารถ ไม่พบว่ามี ความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 แสดงว่า ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ ไม่ถูกฝ่าฝืน

ผลการวิเคราะห์ความแปรปรวนทางเดียวของความแตกต่างของค่าความสามารถของผู้สอบ ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน พบว่าผู้สอบที่มีระดับความสามารถแตกต่างกันมีความแตกต่างของค่าความสามารถของผู้สอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

**ตารางที่ 11** ผลของการเปรียบเทียบความแตกต่างเป็นรายคู่ของความแตกต่างของค่าความสามารถของผู้สอบระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน ด้วยเทคนิค Scheffe

ระดับความสามารถ	ระดับความสามารถ		
	สูง	ปานกลาง	ต่ำ
สูง	—	-0.194	-0.386*
ปานกลาง	—	—	-0.193
ต่ำ	—	—	—

\*p < .05

จากตารางที่ 11 เมื่อทดสอบความแตกต่างเป็นรายคู่ พบว่าค่าความสามารถของผู้สอบแตกต่างกันในระดับความสามารถของผู้สอบที่แตกต่างกัน 1 คู่ คือ ผู้สอบที่มีระดับความสามารถสูง กับ ผู้สอบที่มีระดับความสามารถต่ำ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

### ตอนที่ 3 ผลการเปรียบเทียบลักษณะของการเปลี่ยนคำตอบ

การวิเคราะห์ผลการเปรียบเทียบลักษณะของการเปลี่ยนคำตอบ ภายใต้การทดสอบที่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถแตกต่างกัน โดยแบ่งลักษณะของการเปลี่ยนคำตอบออกเป็น 2 ประเภท ได้แก่

1) การทวนคำตอบแล้วยังคงคำตอบเดิมไว้ ซึ่งวัดจากจำนวนข้อที่ไม่มีความแตกต่างของคำตอบจากการทดสอบที่ให้ทวนคำตอบรอบที่ 1 และรอบที่ 2

2) การทวนคำตอบแล้วมีการเปลี่ยนคำตอบ ซึ่งแบ่งเป็น 3 ลักษณะย่อยคือ จำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นถูก, จำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นผิด และจำนวนข้อที่เปลี่ยนคำตอบจากถูกเป็นผิด

ได้ผลการวิเคราะห์ดังตารางต่อไปนี้

ตารางที่ 12 จำนวนและร้อยละของผู้สอบที่ให้ผลในกรณีต่างๆ ภายใต้การทดสอบที่ให้ทวนคำตอบของผู้สอบในแต่ละระดับความสามารถ

ผล	กรณี	จำนวนผู้สอบ (คน)			
		สูง	ปานกลาง	ต่ำ	รวม
การทำ เครื่องหมาย	ไม่ทำเครื่องหมายเพื่อกลับไปทวนเลย	23 31.94%	22 32.84%	29 47.54%	74 37.00%
	ทำเครื่องหมายเพื่อกลับไปทวนอย่างน้อย 1 ข้อ	49 68.06%	45 67.16%	32 52.46%	126 63.00%
การเลือก ทวน	ไม่กลับไปทวนเลย	7 9.72%	6 8.96%	11 18.03%	24 12.00%
	กลับไปทวนอย่างน้อย 1 ข้อ	65 90.28%	61 91.04%	50 81.97%	176 88.00%

ผล	กรณี	จำนวนผู้สอบ (คน)			
		สูง	ปานกลาง	ต่ำ	รวม
ลักษณะของ	การทวนคำตอบแล้วยังคงคำตอบเดิมไว้	35	26	29	90
		48.61%	38.81%	47.54%	45.00%
การเปลี่ยนคำตอบ	การทวนคำตอบแล้วมีการเปลี่ยนคำตอบ	37	41	32	110
	อย่างน้อย 1 ข้อ แบ่งเป็น 3 ลักษณะย่อยคือ	51.39%	61.19%	52.46%	55.00%
	1) มีการเปลี่ยนคำตอบจากผิดเป็นถูก	35	29	28	92
	2) มีการเปลี่ยนคำตอบจากผิดเป็นผิด	18	24	26	68
	3) มีการเปลี่ยนคำตอบจากถูกเป็นผิด	9	10	13	32
	<b>รวม</b>	<b>72</b>	<b>67</b>	<b>61</b>	<b>200</b>

จากตารางที่ 12 ผลจากการทำเครื่องหมาย พบว่า ผู้สอบส่วนใหญ่มีการทำเครื่องหมายเพื่อกลับไปทวนอย่างน้อย 1 ข้อ (63.00%) โดยผู้สอบที่ไม่ทำเครื่องหมายเพื่อกลับไปทวนเลยที่มีจำนวนมากที่สุด คือผู้สอบมีระดับความสามารถต่ำ (47.54%) และน้อยที่สุด คือผู้สอบมีระดับความสามารถสูง (31.94%) ส่วนผู้สอบที่ทำเครื่องหมายเพื่อกลับไปทวนอย่างน้อย 1 ข้อที่มีจำนวนมากที่สุด คือผู้สอบที่มีระดับความสามารถสูง (68.06%) และน้อยที่สุด คือผู้สอบมีระดับความสามารถต่ำ (52.46%)

ผลจากการเลือกการกลับไปทวน พบว่า ผู้สอบส่วนใหญ่เลือกกลับไปทวนอย่างน้อย 1 ข้อ (88.00%) โดยผู้สอบที่ไม่กลับไปทวนเลยมีจำนวนมากที่สุด คือผู้สอบมีระดับความสามารถต่ำ (18.03%) และน้อยที่สุด คือผู้สอบมีระดับความสามารถสูง (9.72%) ส่วนผู้สอบที่กลับไปทวนอย่างน้อย 1 ข้อ ที่มีจำนวนมากที่สุด คือผู้สอบที่มีระดับความสามารถสูง (90.28%) และน้อยที่สุด คือผู้สอบมีระดับความสามารถต่ำ (81.97%)

ผลจากลักษณะการเปลี่ยนคำตอบ พบว่า ผู้สอบส่วนใหญ่เมื่อทวนคำตอบแล้วมีการเปลี่ยนคำตอบอย่างน้อย 1 ข้อ (55.00%) โดยผู้สอบที่ทวนคำตอบแล้วยังคงคำตอบเดิมไว้มากที่สุด คือผู้สอบมีระดับความสามารถสูง (48.61%) และน้อยที่สุด คือผู้สอบมีระดับความสามารถปานกลาง (38.81%) ส่วนผู้สอบที่ทวนคำตอบแล้วมีการเปลี่ยนคำตอบอย่างน้อย 1 ข้อที่มีจำนวนมากที่สุด คือผู้สอบที่มีระดับความสามารถปานกลาง (61.19%) และน้อยที่สุด คือผู้สอบมีระดับความสามารถสูง (51.39%) โดยผู้สอบที่ทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากผิดเป็นถูกมี

จำนวนมากที่สุด (92 คน) รองมาคือเปลี่ยนคำตอบจากผิดเป็นผิด (68 คน) และน้อยที่สุดคือเปลี่ยนคำตอบจากถูกเป็นผิด (32 คน)

**ตารางที่ 13** จำนวนผู้สอบ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด และผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนด้วย Levene's test และผลการวิเคราะห์ความแปรปรวนทางเดียว ของลักษณะของการเปลี่ยนคำตอบของผู้สอบระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน

การทวน คำตอบแล้ว	ระดับ ความสามารถ ของผู้สอบ	จำนวน ผู้สอบ (คน)	จำนวนข้อสอบ (ข้อ)				Levene's test		F-test	
			Min	Max	$\bar{X}$	SD	F	Sig.	F	Sig.
ยังคงคำตอบ เดิมไว้	สูง	72	14	20	18.54	17.72	2.732	.068	.532	.588
	ปานกลาง	67	15	20	18.39	1.60				
	ต่ำ	61	14	20	18.23	1.91				
	รวม	200	14	20	18.40	1.74				
มีการเปลี่ยน คำตอบจาก ผิดเป็นถูก	สูง	35	1	4	2.17	0.95	4.340	.016	10.810*	.000
	ปานกลาง	29	1	4	2.24	0.69				
	ต่ำ	28	1	3	1.39	0.57				
	รวม	92	1	4	1.96	0.85				
มีการเปลี่ยน คำตอบจาก ผิดเป็นผิด	สูง	18	1	2	1.11	0.32	6.369	.003	8.660*	.000
	ปานกลาง	24	1	3	1.38	0.58				
	ต่ำ	26	1	4	1.88	0.82				
	รวม	68	1	4	1.50	0.70				
มีการเปลี่ยน คำตอบจาก ถูกเป็นผิด	สูง	9	1	1	1.00	0.00	43.795	.000	6.204*	.006
	ปานกลาง	10	1	1	1.00	0.00				
	ต่ำ	13	1	3	1.54	0.66				
	รวม	32	1	3	1.22	0.49				

\*p < .05

**ตารางที่ 14** ผลของการเปรียบเทียบความแตกต่างเป็นรายคู่ของลักษณะของการเปลี่ยนคำตอบของผู้สอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบระหว่างผู้สอบที่มีระดับความสามารถ ของผู้สอบแตกต่างกันด้วยเทคนิค Dunnett's T3

การทวน คำตอบแล้ว	ระดับความสามารถ	ระดับความสามารถ		
		สูง	ปานกลาง	ต่ำ
มีการเปลี่ยน	สูง	—	-0.070	0.779*
คำตอบจาก	ปานกลาง	—	—	0.849*
ผิดเป็นถูก	ต่ำ	—	—	—
มีการเปลี่ยน	สูง	—	-0.264	-0.774*
คำตอบจาก	ปานกลาง	—	—	-0.510*
ผิดเป็นผิด	ต่ำ	—	—	—
มีการเปลี่ยน	สูง	—	0.000	-0.538*
คำตอบจาก	ปานกลาง	—	—	-0.538*
ถูกเป็นผิด	ต่ำ	—	—	—

\* $p < .05$

จากตารางที่ 13 พบว่า ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนของลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วยังคงคำตอบเดิมไว้ ระหว่างผู้สอบในแต่ละระดับความสามารถ ไม่พบว่ามี ความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 แสดงว่า ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ไม่ถูกฝ่าฝืน และผลการวิเคราะห์ความแปรปรวนทางเดียวของการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วยังคงคำตอบเดิมไว้ ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน ไม่พบว่ามี ความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 ระหว่างผู้สอบที่มีระดับความสามารถแตกต่างกัน โดยที่จำนวนข้อที่ผู้สอบยังคงคำตอบเดิมไว้มีค่าอยู่ในช่วง 14 ถึง 20 ข้อ และมีค่าเฉลี่ยเท่ากับ 18.40 ข้อ

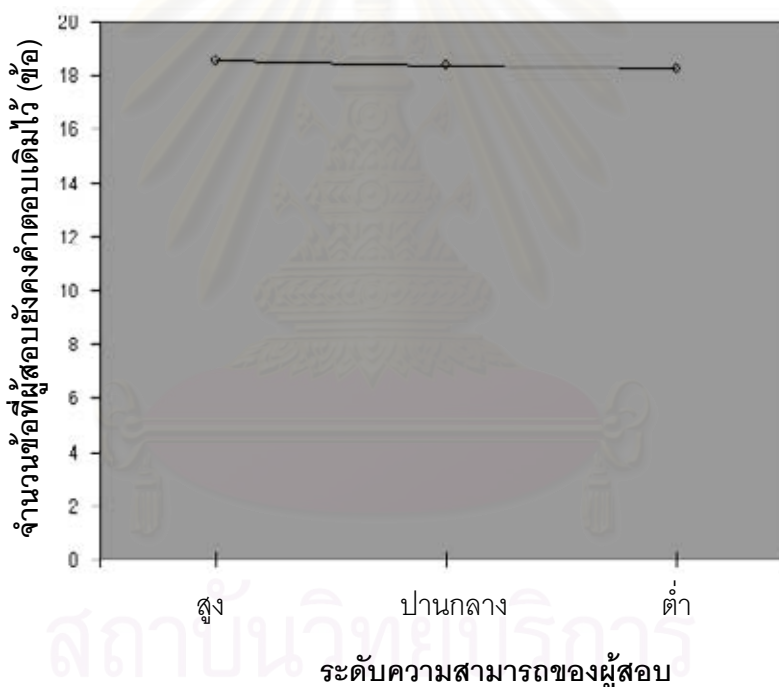
จากตารางที่ 13 และ 14 พบว่า ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนของลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากผิดเป็นถูก ระหว่างผู้สอบในแต่ละระดับความสามารถ พบว่ามี ความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 แสดงว่า ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ถูกฝ่าฝืน และผลการวิเคราะห์ความแปรปรวนทางเดียวของลักษณะการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากผิดเป็นถูก ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน พบว่ามี ความ

แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ระหว่างผู้สอบที่มีระดับความสามารถแตกต่างกัน เมื่อทดสอบความแตกต่างเป็นรายคู่ของลักษณะของการเปลี่ยนคำตอบของผู้สอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากผิดเป็นถูก พบว่าจำนวนข้อที่มีการเปลี่ยนคำตอบจากผิดเป็นถูก จากผิดเป็นผิด และจากถูกเป็นผิด ของผู้สอบ แตกต่างกันในระดับความสามารถของผู้สอบที่แตกต่างกัน 2 คู่ คือ ผู้สอบที่มีระดับความสามารถสูง กับ ผู้สอบที่มีระดับความสามารถต่ำ และ ผู้สอบที่มีระดับความสามารถปานกลาง กับ ผู้สอบที่มีระดับความสามารถต่ำ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่จำนวนข้อที่ผู้สอบมีการเปลี่ยนคำตอบจากผิดเป็นถูก มีค่าอยู่ในช่วง 1 ถึง 4 ข้อ และมีค่าเฉลี่ยเท่ากับ 1.96 ข้อ โดยผู้สอบที่มีระดับความสามารถปานกลางมีการเปลี่ยนคำตอบจากผิดเป็นถูกมากที่สุด (2.24 ข้อ) รองลงมาคือผู้สอบที่มีระดับความสามารถสูง (2.17 ข้อ) และน้อยสุดคือผู้สอบที่มีระดับความสามารถต่ำ (1.39 ข้อ)

ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนของลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากผิดเป็นผิด ระหว่างผู้สอบในแต่ละระดับความสามารถ พบว่ามีความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 แสดงว่า ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ถูกฝ่าฝืน และผลการวิเคราะห์ความแปรปรวนทางเดียวของลักษณะการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากผิดเป็นผิด ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน พบว่ามีความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 ระหว่างผู้สอบที่มีระดับความสามารถแตกต่างกัน เมื่อทดสอบความแตกต่างเป็นรายคู่ของลักษณะของการเปลี่ยนคำตอบของผู้สอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากผิดเป็นผิด พบว่าจำนวนข้อที่มีการเปลี่ยนคำตอบจากผิดเป็นถูก จากผิดเป็นผิด และจากถูกเป็นผิด ของผู้สอบ แตกต่างกันในระดับความสามารถของผู้สอบที่แตกต่างกัน 2 คู่ คือ ผู้สอบที่มีระดับความสามารถสูง กับ ผู้สอบที่มีระดับความสามารถต่ำ และ ผู้สอบที่มีระดับความสามารถปานกลาง กับ ผู้สอบที่มีระดับความสามารถต่ำ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่จำนวนข้อที่ผู้สอบมีการเปลี่ยนคำตอบจากผิดเป็นผิด มีค่าอยู่ในช่วง 1 ถึง 4 ข้อ และมีค่าเฉลี่ยเท่ากับ 1.50 ข้อ โดยผู้สอบที่มีระดับความสามารถต่ำมีการเปลี่ยนคำตอบจากผิดเป็นผิดมากที่สุด (1.88 ข้อ) รองลงมาคือผู้สอบที่มีระดับความสามารถปานกลาง (1.38 ข้อ) และน้อยสุดคือผู้สอบที่มีระดับความสามารถสูง (1.11 ข้อ)

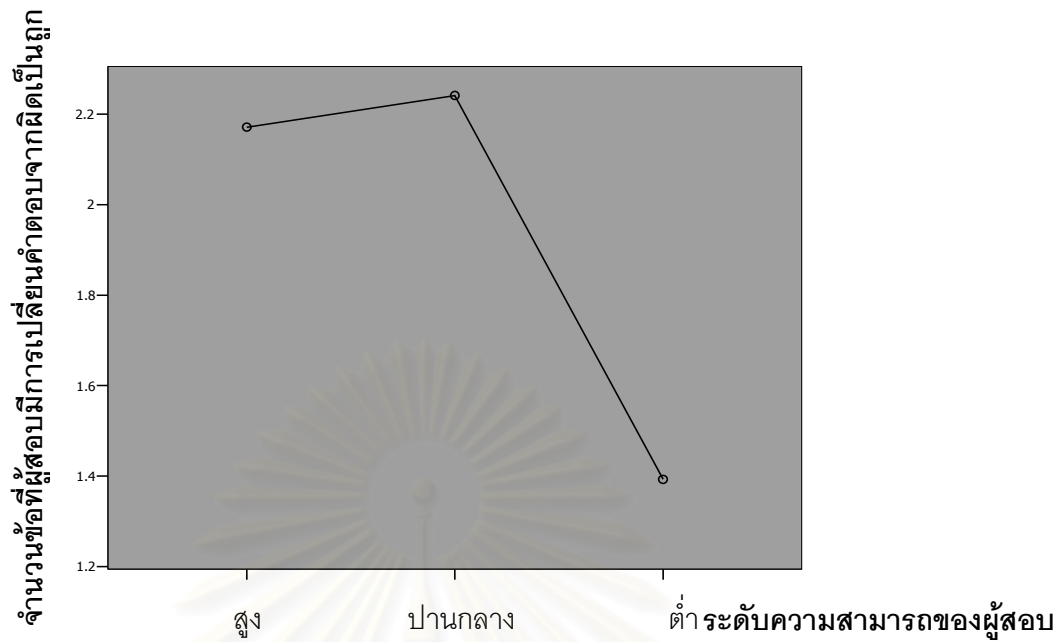
ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนของลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากถูกเป็นผิด ระหว่างผู้สอบในแต่ละระดับความสามารถ พบว่ามีความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 แสดงว่า ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ถูกฝ่าฝืน และผลการวิเคราะห์ความแปรปรวนทางเดียวของลักษณะ

การเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากถูกเป็นผิด ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน พบว่ามีความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 ระหว่างผู้สอบที่มีระดับความสามารถแตกต่างกัน เมื่อทดสอบความแตกต่างเป็นรายคู่ของลักษณะของการเปลี่ยนคำตอบของผู้สอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากถูกเป็นผิด พบว่าจำนวนข้อที่มีการเปลี่ยนคำตอบจากผิดเป็นถูก จากผิดเป็นผิด และจากถูกเป็นผิดของผู้สอบ แตกต่างกันในระดับความสามารถของผู้สอบที่แตกต่างกัน 2 คู่ คือ ผู้สอบที่มีระดับความสามารถสูง กับ ผู้สอบที่มีระดับความสามารถต่ำ และ ผู้สอบที่มีระดับความสามารถปานกลาง กับ ผู้สอบที่มีระดับความสามารถต่ำ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่จำนวนข้อที่ผู้สอบมีการเปลี่ยนคำตอบจากถูกเป็นผิด มีค่าอยู่ในช่วง 1 ถึง 3 ข้อ และมีค่าเฉลี่ยเท่ากับ 1.22 ข้อ โดยผู้สอบที่มีระดับความสามารถต่ำมีการเปลี่ยนคำตอบจากถูกเป็นผิดมากที่สุด (1.54 ข้อ) และน้อยที่สุดคือผู้สอบที่มีระดับความสามารถปานกลาง และสูงซึ่งมีค่าเฉลี่ยเท่ากัน (1.00 ข้อ)



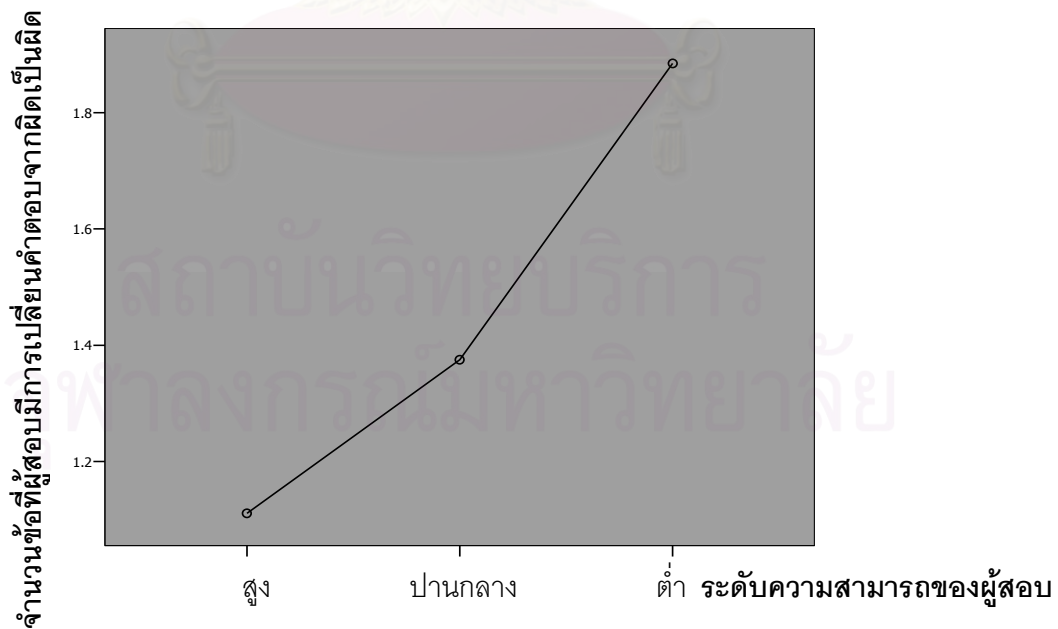
แผนภาพที่ 34 กราฟแสดงแนวโน้มของลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วยังคงคำตอบเดิมไว้ของผู้สอบที่มีระดับความสามารถแตกต่างกัน

จากแผนภาพที่ 34 พบว่าจำนวนข้อที่ผู้สอบยังคงคำตอบเดิมไว้ในกลุ่มผู้สอบที่มีระดับความสามารถสูง ปานกลาง และต่ำมีค่าไม่แตกต่างกัน แต่มีแนวโน้มต่ำลง (จำนวนข้อน้อยลง) เพียงเล็กน้อย เมื่อผู้สอบมีระดับความสามารถต่ำลง



แผนภาพที่ 35 กราฟแสดงแนวโน้มของลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบ แล้วมีการเปลี่ยนคำตอบจากผิดเป็นถูก ของผู้สอบที่มีระดับความสามารถแตกต่างกัน

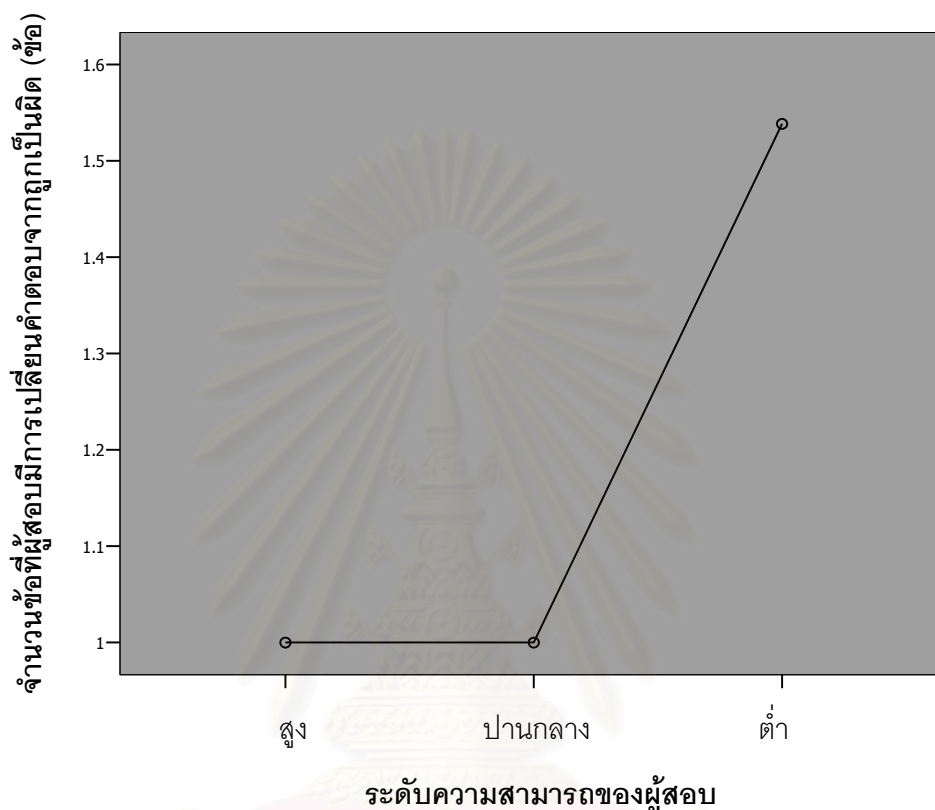
จากแผนภาพที่ 35 พบว่าจำนวนข้อที่ผู้สอบมีการเปลี่ยนคำตอบจากผิดเป็นถูก ในกลุ่มผู้สอบที่มีระดับความสามารถสูงและปานกลางมีค่าไม่แตกต่างกันและมีแนวโน้มต่ำลง (จำนวนข้อน้อยลง) เมื่อผู้สอบมีระดับความสามารถต่ำ



แผนภาพที่ 36 กราฟแสดงแนวโน้มของลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบ แล้วมีการเปลี่ยนคำตอบจากผิดเป็นผิด ของผู้สอบที่มีระดับความสามารถแตกต่างกัน



จากแผนภาพที่ 36 พบว่าจำนวนข้อที่ผู้สอบมีการเปลี่ยนคำตอบจากผิดเป็นผิด ในกลุ่มผู้สอบที่มีระดับความสามารถสูงและปานกลางมีค่าไม่แตกต่างกัน และมีแนวโน้มสูงขึ้น (จำนวนข้อมากขึ้น) เมื่อผู้สอบมีระดับความสามารถต่ำ



แผนภาพที่ 37 กราฟแสดงแนวโน้มของลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบ แล้วมีการเปลี่ยนคำตอบจากถูกเป็นผิด ของผู้สอบที่มีระดับความสามารถแตกต่างกัน

จากแผนภาพที่ 37 พบว่าจำนวนข้อที่ผู้สอบมีการเปลี่ยนคำตอบจากถูกเป็นผิด ในกลุ่มผู้สอบที่มีระดับความสามารถสูงและปานกลางมีค่าไม่แตกต่างกัน และมีแนวโน้มสูงขึ้น (จำนวนข้อมากขึ้น) เมื่อผู้สอบมีระดับความสามารถต่ำ

#### ตอนที่ 4 ผลการเปรียบเทียบความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ

การวิเคราะห์ผลการเปรียบเทียบความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ตามเงื่อนไขการทดสอบ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน ได้ผลการวิเคราะห์ดังตารางที่ 15 ต่อไปนี้

**ตารางที่ 15** จำนวนผู้สอบ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ของเวลาที่ใช้ในการทำข้อสอบ และผลของการเปรียบเทียบเวลาที่ใช้ในการทำข้อสอบของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบ ของผู้สอบในแต่ละระดับความสามารถ

ระดับ ความสามารถ ของผู้สอบ	จำนวน ผู้สอบ (คน)	เวลาที่ใช้ในการทำข้อสอบของผู้สอบ (วินาที)								t	Sig.
		เงื่อนไขการไม่ให้ทวน คำตอบ ( $t_{\text{ไม่ทวน}}$ )				เงื่อนไขการให้ทวน คำตอบ ( $t_{\text{ทวน}}$ )					
		Min	Max	$\bar{X}$	SD	Min	Max	$\bar{X}$	SD		
สูง $\theta \geq 0.7$	72	585	1398	1017	209	765	1680	1230	187	8.194*	.000
		9/45	23/18	16/57	3/29	12/45	28/0	20/30	3/7		
ปานกลาง $-0.7 < \theta < 0.7$	67	444	1739	1008	206	733	1731	1192	200	7.057*	.000
		7/24	28/59	16/48	3/26	12/13	28/51	19/52	3/20		
ต่ำ $\theta \leq -0.7$	61	515	1537	922	223	619	1710	1111	252	6.537*	.000
		8/35	25/37	15/22	3/43	10/19	28/30	18/31	4/12		
<b>รวม</b>	200	444	1537	985	216	619	1731	1181	217		
		7/24	25/37	16/25	3/36	10/19	28/30	19/41	3/37		

\* $p < .05$

**หมายเหตุ** ตัวเลขที่พิมพ์ตัวเอียง หมายถึง นาที/วินาที

จากตารางที่ 15 เวลาที่ใช้ในการทำข้อสอบของผู้สอบในเงื่อนไขการไม่ให้ทวนคำตอบ พบว่าผู้สอบที่อยู่ในระดับความสามารถสูงจะใช้เวลาในการทำข้อสอบอยู่ในช่วง 585 ถึง 1398 วินาที หรือ 9 นาที 45 วินาที ถึง 23 นาที 18 วินาที ผู้สอบที่อยู่ในระดับความสามารถปานกลางจะใช้เวลาในการทำข้อสอบอยู่ในช่วง 444 ถึง 1739 วินาที หรือ 7 นาที 24 วินาที ถึง 28 นาที 59 วินาที และผู้สอบที่อยู่ในระดับความสามารถต่ำจะใช้เวลาในการทำข้อสอบอยู่ในช่วง 515 ถึง 1537 วินาที หรือ 8 นาที 35 วินาที ถึง 25 นาที 37 วินาที

เวลาที่ใช้ในการทำข้อสอบของผู้สอบในเงื่อนไขการให้ทวนคำตอบ พบว่าผู้สอบที่อยู่ในระดับความสามารถสูงจะใช้เวลาในการทำข้อสอบอยู่ในช่วง 765 ถึง 1680 วินาที หรือ 12 นาที 45 วินาที ถึง 28 นาที ผู้สอบที่อยู่ในระดับความสามารถปานกลางจะใช้เวลาในการทำข้อสอบอยู่ในช่วง 733 ถึง 1731 วินาที หรือ 12 นาที 13 วินาที ถึง 28 นาที 51 วินาที และผู้สอบที่อยู่ในระดับความสามารถต่ำจะใช้เวลาในการทำข้อสอบอยู่ในช่วง 619 ถึง 1710 วินาที หรือ 10 นาที 19 วินาที ถึง 28 นาที 30 วินาที

ค่าเฉลี่ยของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ พบว่าในเงื่อนไขการให้ทวนคำตอบมีค่ามากกว่าเงื่อนไขการไม่ให้ทวนคำตอบ ในทุกระดับความสามารถ (สูง, ปานกลาง และต่ำ) จึงส่งผลให้ค่าเฉลี่ยของเวลาที่ใช้ในการทำข้อสอบของผู้สอบรวมทุกระดับความสามารถในเงื่อนไขการให้ทวนคำตอบ ( $\bar{X} = 1181$  วินาที หรือ 19 นาที 41 วินาที) มีค่ามากกว่าเงื่อนไขการไม่ให้ทวนคำตอบ ( $\bar{X} = 985$  วินาที หรือ 16 นาที 25 วินาที)

ผลการเปรียบเทียบค่าเฉลี่ยของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบ ของผู้สอบที่มีระดับความสามารถสูง ปานกลาง และต่ำ พบว่าผู้สอบที่มีระดับความสามารถสูง ปานกลาง และต่ำ มีค่าเฉลี่ยของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

**ตารางที่ 16** จำนวนผู้สอบ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด และค่าต่ำสุด ของความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ จำแนกตามระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

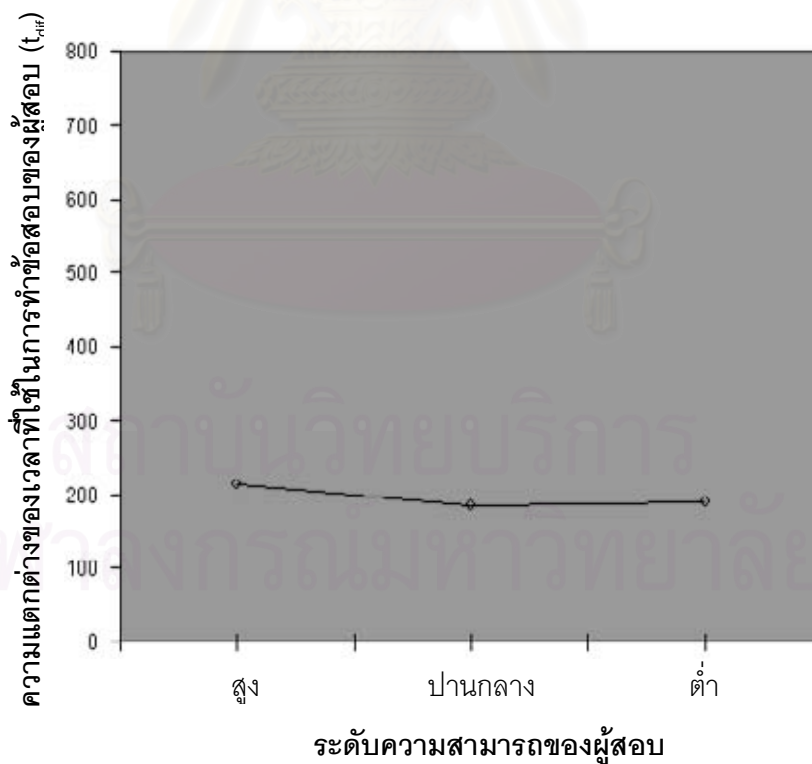
ระดับ ความสามารถ ของผู้สอบ	จำนวน นักเรียน	ความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ( $t_{dif}$ ) (วินาที)			
		Min	Max	$\bar{X}$	SD
สูง	72	-410	676	214	221
		-6/50	11/16	3/34 (21.25%)	3/41
ปานกลาง	67	-453	697	184	213
		-7/33	11/37	3/4 (18.25%)	3/33
ต่ำ	61	-218	840	189	226
		-3/38	14/0	3/9 (20.50%)	3/46
รวม	200	-453	840	196	219
		-7/33	14/0	3/16 (19.90%)	3/39

**หมายเหตุ**

- $t_{dif} = t_{ทวน} - t_{ไม่ทวน}$
- กรณีที่  $t_{dif}$  มีค่าเป็นบวก แสดงว่า  $t_{ทวน}$  มีค่ามากกว่า  $t_{ไม่ทวน}$   
 $t_{dif}$  มีค่าเป็นลบ แสดงว่า  $t_{ทวน}$  มีค่าน้อยกว่า  $t_{ไม่ทวน}$
- ตัวเลขที่พิมพ์ตัวเอียง หมายถึง นาที/วินาที

จากตารางที่ 16 พบว่า เงื่อนไขการให้ทวนคำตอบส่งผลให้ผู้สอบในระดับความสามารถสูงมีความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ อยู่ในช่วง -410 ถึง 676 วินาที หรือ -6 นาที 50 วินาที ถึง 11 นาที 16 วินาที ผู้สอบระดับความสามารถปานกลางอยู่ในช่วง -453 ถึง 697 วินาที หรือ -7 นาที 33 วินาที ถึง 11 นาที 37 วินาที และผู้สอบระดับความสามารถต่ำอยู่ในช่วง -218 ถึง 840 วินาที หรือ -3 นาที 38 วินาที ถึง 14 นาที โดยที่ความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบที่ต่ำที่สุด และสูงที่สุดอยู่ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ ( $t_{diff} = -453$  วินาที หรือ -7 นาที 33 วินาที และ 840 วินาที หรือ 14 นาที ตามลำดับ)

ผู้สอบในทุกระดับความสามารถมีค่าเฉลี่ยความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบเป็นค่าบวก ( $t_{diff} = 196$  วินาที หรือ 3 นาที 16 วินาที) คิดเป็น 19.90% ซึ่งแสดงว่าผู้สอบส่วนใหญ่จะใช้เวลาในการทำข้อสอบในเงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบ โดยที่ผู้สอบที่มีระดับความสามารถสูงจะใช้เวลาในการทำข้อสอบในเงื่อนไขการให้ทวนคำตอบเพิ่มขึ้นจากเงื่อนไขการไม่ให้ทวนคำตอบมากที่สุด ( $t_{diff} = 214$  วินาที หรือ 3 นาที 34 วินาที) คิดเป็น 21.25% รองลงมาคือผู้สอบที่มีระดับความสามารถต่ำ ( $t_{diff} = 189$  วินาที หรือ 3 นาที 9 วินาที) คิดเป็น 20.50% และน้อยที่สุดคือผู้สอบที่มีระดับความสามารถปานกลาง ( $t_{diff} = 184$  วินาที หรือ 3 นาที 4 วินาที) คิดเป็น 18.25%



แผนภาพที่ 38 กราฟแสดงแนวโน้มของความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ( $t_{diff}$ ) ของผู้สอบที่มีระดับความสามารถแตกต่างกัน

จากแผนภาพที่ 38 พบว่าความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบในกลุ่มผู้สอบที่มีระดับความสามารถสูง ปานกลาง และต่ำมีค่าไม่แตกต่างกัน แต่มีแนวโน้มสูงขึ้น เมื่อผู้สอบมีระดับความสามารถในช่วงปานกลางถึงสูง และมีแนวโน้มต่ำลง เมื่อผู้สอบมีระดับความสามารถในช่วงต่ำถึงปานกลาง

**ตารางที่ 17** ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนด้วย Levene's test และผลการวิเคราะห์ความแปรปรวนทางเดียว ของความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน

แหล่งความแปรปรวน	df	SS	MS	F	Sig.
ระหว่างกลุ่ม	2	35562	17781	0.368	0.693
ภายในกลุ่ม	197	9527507	48363		
<b>รวม</b>	199	9563069			

Levene Statistic	df1	df2	Sig.
0.360	2	197	0.698

\*p < .05

จากตารางที่ 17 ผลการทดสอบความเป็นเอกพันธ์ความแปรปรวนของความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ระหว่างผู้สอบในแต่ละระดับความสามารถ ไม่พบว่ามี ความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 แสดงว่า ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ไม่ถูกฝ่าฝืน

ผลการวิเคราะห์ความแปรปรวนทางเดียวของความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ระหว่างผู้สอบที่มีระดับความสามารถแตกต่างกัน พบว่าผู้สอบที่มีระดับความสามารถแตกต่างกันมีความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

## บทที่ 5

### สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาโปรแกรมคอมพิวเตอร์วิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถม ศึกษาปีที่ 6 สำหรับการทดสอบแบบปรับเหมาะกับความ สามารถของผู้สอบด้วยคอมพิวเตอร์ ที่สามารถกำหนดเงื่อนไขการทดสอบได้ 2 เงื่อนไข คือ การทดสอบที่ให้หาคำตอบได้โดยวิธีการทวนที่มีการจำกัดช่วง (Block Review) กับการทดสอบที่ไม่ให้หาคำตอบ รวมทั้งเปรียบเทียบความแตกต่างของค่าความสามารถ ลักษณะของการเปลี่ยน คำตอบ และความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความ สามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อระดับความสามารถของผู้สอบแตกต่างกัน

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2549 จำนวน 6 ห้องเรียน โรงเรียนวัดพลับพลาชัย ซึ่งคิดเป็นกลุ่มตัวอย่าง 224 คน ได้มาจากประชากรที่มีการสุ่มแบบหลายขั้นตอน (multistage random)

กลุ่มนักเรียนที่ใกล้เคียงกับกลุ่มตัวอย่าง ซึ่งผู้วิจัยใช้เป็นส่วนหนึ่งในกระบวนการ ดำเนินการวิจัย คือ

- 1) นักเรียนกลุ่ม 1 ใช้ในการหาค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ของแบบสอบปรนัยวิชา คณิตศาสตร์ เรื่องความเข้าใจ เศษส่วน เพื่อพัฒนาค้างข้อสอบวิชา คณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ระดับชั้น ประถมศึกษาปีที่ 6 ให้ครบอย่างน้อย 210 ข้อ ซึ่งได้มาจากการเลือกแบบเจาะจงโรงเรียน ประถมศึกษาขนาดใหญ่ ในเขตพื้นที่การศึกษากทม.เขต 1 ได้เป็นโรงเรียนราชวินิต ซึ่งมีนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 10 ห้องเรียน คิดเป็นนักเรียน 409 คน

- 2) นักเรียนกลุ่ม 2 ใช้ในการกำหนดระยะเวลาสูงสุดที่เหมาะสมในการทำ ข้อสอบในการทดสอบแบบปรับเหมาะกับความ สามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing - CAT) ให้กับกลุ่มตัวอย่าง รวมทั้งตรวจสอบและปรับปรุง ความถูกต้องของโปรแกรมคอมพิวเตอร์ ซึ่งได้มาจากการเลือกแบบเจาะจง โดยเลือกนักเรียนที่ กำลังศึกษาอยู่ในระดับชั้นประถมศึกษาปีที่ 6 คอร์สคณิตเข้มเปิดเทอมปลาย ปีการศึกษา 2549 ในโรงเรียนกวดวิชา พี.พี.เซ็นเตอร์ ของทางผู้วิจัย คิดเป็นนักเรียน 9 คน ซึ่งจะเลือกนักเรียนที่มี ความสามารถในวิชาคณิตศาสตร์แตกต่างกัน (ต่ำ, ปานกลาง และสูง) โดยอาศัยคะแนนทดสอบพ รีเทสต์ วิชาคณิตศาสตร์ของทางโรงเรียนเป็นตัวกำหนด

สมมติฐานการวิจัยครั้งนี้ คือ

1) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบ และระดับความสามารถของผู้สอบแตกต่างกัน น่าจะส่งผลให้ค่าความสามารถของผู้สอบมีความแตกต่างกัน โดยที่ผู้สอบที่มีระดับความสามารถสูงและปานกลาง ค่าความสามารถของผู้สอบภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ น่าจะมากกว่าการทดสอบที่ไม่ให้ทวนคำตอบ ส่วนผู้สอบที่มีระดับความสามารถต่ำ ค่าความสามารถของผู้สอบภายใต้เงื่อนไขการทดสอบที่ให้ทวนคำตอบ น่าจะน้อยกว่าการทดสอบที่ไม่ให้ทวนคำตอบ

2) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ภายใต้การทดสอบที่ให้ทวนคำตอบ น่าจะส่งผลให้การทวนคำตอบแล้วยังคงคำตอบเดิมไว้ไม่แตกต่างกันในแต่ละระดับความสามารถของผู้สอบ ส่วนการทวนคำตอบแล้วมีการเปลี่ยนคำตอบ น่าจะแตกต่างกันในแต่ละระดับความสามารถของผู้สอบ โดยผู้สอบที่มีระดับความสามารถสูงและปานกลางน่าจะมีการเปลี่ยนคำตอบจากผิดเป็นถูก แต่ผู้สอบที่มีระดับความสามารถต่ำน่าจะมีการเปลี่ยนคำตอบจากผิดเป็นผิด หรือจากถูกเป็นผิด

3) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบ และระดับความสามารถของผู้สอบแตกต่างกัน น่าจะส่งผลให้เวลาที่ใช้ในการทำข้อสอบของผู้สอบไม่แตกต่างกัน โดยที่ผู้สอบทุกระดับความสามารถน่าจะใช้เวลาในการทำข้อสอบภายใต้การทดสอบที่ให้ทวนคำตอบมากกว่าการทดสอบที่ไม่ให้ทวนคำตอบ

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ คลังข้อสอบวิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ของ รังสรรค์ มณีเล็ก (2540) ซึ่งเป็นแบบสอบปรนัย 4 ตัวเลือก จำนวน 200 ข้อ, แบบสอบปรนัยวิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน 4 ตัวเลือก จำนวน 40 ข้อ ซึ่งทางผู้วิจัยสร้างขึ้น สำหรับคัดเลือกเพื่อเพิ่มข้อสอบอย่างน้อย 10 ข้อเข้าคลังข้อสอบ ด้วยวิธีการปรับเทียบค่าพารามิเตอร์ด้วยวิธีค่าเฉลี่ยและซิกมา (Mean and Sigma Method) โดยใช้โปรแกรม BILOG-MG สำหรับการวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งผู้สอบจะต้องทำการทดสอบ 2 เงื่อนไข รวม 2 ครั้ง และลำดับของเงื่อนไขการทดสอบจะถูกกำหนดโดยเลขที่ของนักเรียน กล่าวคือ ถ้าเลขที่เป็นเลขคี่ จะทดสอบครั้งที่ 1 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การทดสอบที่ให้ทวนคำตอบ และถ้าเลขที่เป็นเลขคู่ จะทดสอบครั้งที่ 1 ภายใต้การทดสอบที่ให้ทวนคำตอบ และครั้งที่ 2 ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบ

การเก็บรวบรวมข้อมูล ได้แบ่งออกเป็น 3 ขั้นตอน คือ

**ขั้นตอนแรก** เป็นการทดสอบด้วยแบบสอบปรนัยวิชา คณิตศาสตร์ เรื่อง ความเข้าใจเศษส่วน จำนวน 40 ข้อ สำหรับคัดเลือกเพื่อเพิ่มข้อสอบอย่างน้อย 10 ข้อเข้าคลังข้อสอบ โดยนำไปใช้กับนักเรียนกลุ่ม 1

**ขั้นตอนที่สอง** เป็นการทดสอบแบบซีเอที (Computerized Adaptive Testing - CAT) ด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เพื่อหาระยะเวลาสูงสุดที่เหมาะสมให้กับกลุ่มตัวอย่างในการทดสอบแบบซีเอที โดยนำไปใช้กับนักเรียนกลุ่ม 2

**ขั้นตอนที่สาม** เป็นการทดสอบแบบซีเอที (Computerized Adaptive Testing - CAT) ด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เพื่อนำมาวิเคราะห์ตามวัตถุประสงค์ของการวิจัย โดยนำไปใช้กับกลุ่มตัวอย่าง

การวิเคราะห์ข้อมูล ได้แบ่งออกเป็น 3 ขั้นตอน คือ

**ขั้นตอนแรก** สำหรับการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โดยใช้โปรแกรม BILOG-MG และปรับเทียบค่าพารามิเตอร์ของข้อสอบเพื่อบันทึกเข้าคลัง ด้วยวิธีค่าเฉลี่ยและซิกมา (Mean and Sigma Method) เพื่อที่จะบันทึกข้อสอบเข้าคลังเพิ่ม โดยใช้โปรแกรม Microsoft Office Excel 2003

**ขั้นตอนที่สอง** สำหรับการวิเคราะห์หาระยะเวลาสูงสุดที่เหมาะสมในการทำข้อสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing - CAT) โดยนำมาเป็นตัวกำหนดระยะเวลาสูงสุดในการทำข้อสอบให้กับโปรแกรมคอมพิวเตอร์ โดยใช้สูตร  $Spending\ Time = \bar{X} + SD + 5\ นาที$

**ขั้นตอนที่สาม** สำหรับการวิเคราะห์เพื่อเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing - CAT) เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน ใช้การวิเคราะห์ค่าสถิติพื้นฐานเพื่อทำการวิเคราะห์ข้อมูลเบื้องต้น และใช้การวิเคราะห์ความแปรปรวนทางเดียว (One-Way ANOVA) โดยใช้โปรแกรม SPSS for Windows เพื่อเปรียบเทียบตัวแปรตามภายใต้ตัวแปรต้น ต่อมาใช้วิธีการทดสอบของเชฟเฟ (Schffe) ทำการเปรียบเทียบความแตกต่างเป็นรายคู่ กรณีที่ทราบว่าการที่ประชากรทั้ง k กลุ่มมีค่าความแปรปรวนไม่แตกต่างกัน (equal variances assumed) และใช้วิธีการทดสอบของดันเนท (Dunnett's T3) ทำการเปรียบเทียบความแตกต่างเป็นรายคู่



กรณีที่ทราบว่าเป็นประชากรทั้ง  $k$  กลุ่มมีค่าความแปรปรวนต่างกัน (equal variances not assumed) ตามสถานการณ์ที่สนใจศึกษา ดังนี้

สถานการณ์ที่หนึ่ง

ตัวแปรต้น (independent variable) ► ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ตัวแปรตาม (dependent variable) ► ความแตกต่างของค่าความสามารถของผู้สอบ

สถานการณ์ที่สอง

ตัวแปรต้น (independent variable) ► ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ตัวแปรตาม (dependent variable) ► ลักษณะของการเปลี่ยนคำตอบของผู้สอบ

สถานการณ์ที่สาม

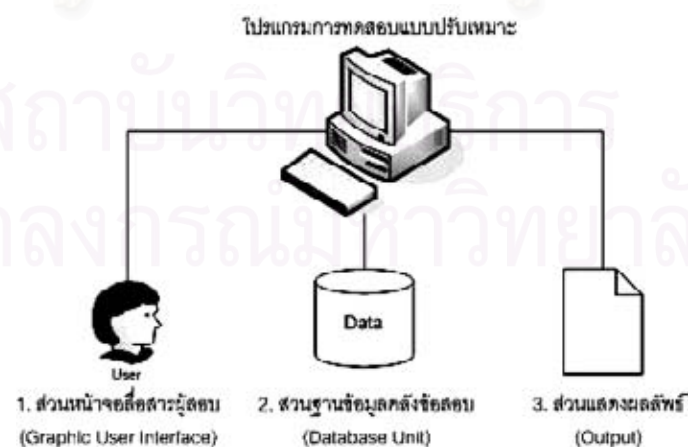
ตัวแปรต้น (independent variable) ► ระดับความสามารถของผู้สอบ (สูง, ปานกลาง และต่ำ)

ตัวแปรตาม (dependent variable) ► ความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ

### สรุปผลการวิจัย

จากผลการพัฒนาโปรแกรมคอมพิวเตอร์และผลการวิเคราะห์ข้อมูลสามารถสรุปผลการวิจัยตามวัตถุประสงค์ของการวิจัยได้ดังนี้

1. โปรแกรมคอมพิวเตอร์วิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ที่สามารถกำหนดเงื่อนไขการทดสอบได้ 2 เงื่อนไข คือ การทดสอบภายใต้เงื่อนไขการให้ทวนคำตอบ โดยวิธีการทวนที่มีการจำกัดช่วง (Block Review) กับเงื่อนไขการไม่ให้ทวนคำตอบ ซึ่งโปรแกรมแบ่งลักษณะความสามารถการทำงานออกเป็น 3 ส่วนดังแผนภาพที่ 39



แผนภาพที่ 39 ภาพแสดงการแบ่งส่วนประกอบโปรแกรมการทดสอบแบบปรับเหมาะตามลักษณะการทำงาน

**1.1 ส่วนหน้าจอสื่อสารผู้สอบ** เป็นส่วนที่ใช้ติดต่อสื่อสารกับผู้สอบ 2 ทาง คือ รับและแสดงผลข้อมูล มีลักษณะเป็นหน้าต่างที่ทำหน้าที่รองรับค่า (input) ของข้อมูลผู้สอบ และแสดงผล (output) ให้ผู้สอบในเวลาเดียวกัน

### 1.1.1 ความสามารถด้านรับค่า (input)

- โปรแกรมสามารถรับค่าอักขระ (Character) ได้ ซึ่งเป็นข้อมูลส่วนตัวของผู้สอบ ได้แก่ ชื่อนามสกุล ชั้น เลขที่ โรงเรียน
- ผู้สอบสามารถเลือกคำตอบ (Radio button) และเปลี่ยนคำตอบก่อนทำข้อสอบข้อถัดไปได้
- ในการทดสอบภายใต้เงื่อนไขการให้ทวนคำตอบ เมื่อผู้สอบทำข้อสอบครบทุก ๆ 5 ข้อแล้วผู้สอบสามารถกดปุ่มข้อสอบเก่าที่ได้ทำการทดสอบผ่านมาแล้ว เพื่อกลับไปทวนได้
- ในการทดสอบภายใต้เงื่อนไขการให้ทวนคำตอบ เมื่อผู้สอบกลับมาทวนข้อสอบเก่าสามารถกดปุ่มขอเปลี่ยนคำตอบ และเลือกคำตอบใหม่ได้

### 1.1.2 ความสามารถด้านแสดงผล (output)

- โปรแกรมสามารถแสดงเวลาที่เหลือในการทดสอบได้
- โปรแกรมสามารถเลือกข้อสอบได้ตรงตามความสามารถของผู้สอบ
- โปรแกรมสามารถแสดงค่าความสามารถของผู้สอบเมื่อทำการทดสอบเสร็จ

**1.2 ส่วนฐานข้อมูลคลังข้อสอบ** เป็นส่วนที่ผู้วิจัยทำการจัดเก็บคลังข้อสอบที่จะนำมาใช้ในการทดสอบ ซึ่งสามารถพัฒนาโปรแกรมการทดสอบโดยเพิ่มข้อสอบหรือลบข้อสอบออกจากคลังได้ และสามารถแก้ไขค่า 3-parameter ของคลังข้อสอบได้

**1.3 ส่วนแสดงผลลัพท์** เป็นส่วนที่บรรยายข้อมูลการทำข้อสอบของผู้สอบตั้งแต่เริ่มจนจบ และบรรยายค่าความสามารถและค่าความแปรปรวนหลังการตอบคำถามแต่ละข้อ

### 1.3.1 ส่วนแสดงผลลัพท์การทดสอบภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ

- โปรแกรมสามารถสร้างแฟ้มข้อมูล (text file) เพื่อบรรยายผลการสอบ ของผู้สอบ ซึ่งมีชื่อนามสกุล ของผู้สอบได้
- โปรแกรมสามารถแสดงการตอบคำถามของผู้สอบในแต่ละข้อได้ ว่าตอบถูกหรือผิด
- โปรแกรมสามารถแสดงค่าความสามารถและค่าความแปรปรวนของผู้สอบหลังทำการทดสอบแต่ละข้อได้
- โปรแกรมสามารถแสดงเวลาที่ใช้ในการสอบของผู้สอบได้
- โปรแกรมสามารถสรุปจำนวนข้อสอบที่ตอบผิดและจำนวนข้อสอบที่ตอบถูกได้

### 1.3.2 ส่วนแสดงผลลัพธ์การทดสอบภายใต้เงื่อนไขการให้ทวน คำตอบ

- โปรแกรมสามารถสร้างแฟ้มข้อมูล (text file) เพื่อบรรยายผลการสอบ ของผู้สอบ ซึ่งมีชื่อนามสกุล ของผู้สอบได้
- โปรแกรมสามารถแสดงการตอบคำถามของผู้สอบในแต่ละข้อได้ ว่าตอบถูกหรือผิดทั้งก่อนและหลังการทวน
- โปรแกรมสามารถแสดงค่าความสามารถและค่าความแปรปรวนของผู้สอบหลังทำการทดสอบแต่ละข้อได้
- โปรแกรมสามารถแสดงค่าความสามารถและค่าความแปรปรวนของผู้สอบก่อนและหลังการทวน เพื่อนำมาเปรียบเทียบได้
- โปรแกรมสามารถแสดงเวลาที่ใช้ในการสอบของผู้สอบได้
- โปรแกรมสามารถสรุปจำนวนข้อสอบที่ตอบผิดและจำนวนข้อสอบที่ตอบถูกได้
- โปรแกรมสามารถสรุปจำนวนคำตอบจากถูกเป็นผิดได้
- โปรแกรมสามารถสรุปจำนวนคำตอบจากผิดเป็นถูกได้
- โปรแกรมสามารถสรุปจำนวนคำตอบจากผิดเป็นผิดได้
- โปรแกรมสามารถสรุปจำนวนข้อที่มีการเปลี่ยนคำตอบได้
- โปรแกรมสามารถสรุปจำนวนข้อที่มีการทำสัญลักษณ์เพื่อกลับไปทวนได้ โดยแยกเป็นการกลับไปทวนข้อที่ทำสัญลักษณ์และการกลับไปทวนข้อที่ไม่ได้ทำสัญลักษณ์

2. จากการสอบถามปากเปล่าและสังเกตพฤติกรรมของผู้สอบ พบว่า ผู้สอบส่วนใหญ่มีความเห็นในด้านบวกต่อโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ทั้งในด้านความสะดวกในการใช้งาน รวมทั้งรูปแบบของโปรแกรมและคู่มือในการใช้งาน (ดูในภาคผนวก ข) มีความเหมาะสม อีกทั้งผู้สอบส่วนใหญ่มีความตั้งใจ สนใจ และกระตือรือร้นในการทำการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เป็นอย่างมาก โดยเมื่อผู้สอบทำการทดสอบเสร็จเรียบร้อยแล้ว ผู้สอบจะริบดูผลการทดสอบที่โปรแกรมแสดงผลให้ดูทางหน้าจอ รวมทั้งมีการจดผลการทดสอบแล้วนำผลการทดสอบมาเปรียบเทียบระหว่างผู้สอบว่าใครได้ค่าความสามารถเท่าไร และมีผู้สอบบางคนขอทำการทดสอบใหม่อีกครั้ง อีกทั้งผู้สอบส่วนมากให้ข้อเสนอแนะว่าอยากให้มีการจัดการทดสอบแบบนี้ในการสอบกลางภาคและปลายภาคด้วย

3. จากการเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ตามเงื่อนไขการทดสอบ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน พบว่า

3.1 ผลการเปรียบเทียบค่าเฉลี่ยของค่าความสามารถของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบ ของผู้สอบที่มีระดับความสามารถสูง พบว่าผู้สอบที่มีระดับความสามารถสูง มีค่าเฉลี่ยของค่าความสามารถของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แต่ผู้สอบที่มีระดับความสามารถปานกลาง และต่ำ พบว่าผู้สอบที่มีระดับความสามารถปานกลาง และต่ำ มีค่าเฉลี่ยของค่าความสามารถของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.2 ผู้สอบที่มีระดับความสามารถแตกต่างกันมีความแตกต่างของค่าความสามารถของผู้สอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยเมื่อทดสอบความแตกต่างเป็นรายคู่ พบว่าค่าความสามารถของผู้สอบแตกต่างกันในระดับความสามารถของผู้สอบที่แตกต่างกัน 1 คู่ คือ ผู้สอบที่มีระดับความสามารถสูง กับ ผู้สอบที่มีระดับความสามารถต่ำ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งความแตกต่างของค่าความสามารถของผู้สอบ ( $\theta_{diff}$ ) มีแนวโน้มสูงขึ้น (มีค่าเป็นบวกมากขึ้น) เมื่อผู้สอบมีระดับความสามารถต่ำลง

3.3 ผู้สอบส่วนใหญ่จะได้ค่าความสามารถในเงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบเฉลี่ย 0.214 โดยที่ผู้สอบที่มีระดับความสามารถต่ำจะได้ค่าความสามารถในเงื่อนไขการให้ทวนคำตอบเพิ่มขึ้นจากเงื่อนไขการไม่ให้ทวนคำตอบมากที่สุด รองลงมาได้แก่ ผู้สอบที่มีระดับความสามารถปานกลาง และผู้สอบที่มีระดับความสามารถสูงตามลำดับ

4. จากการเปรียบเทียบลักษณะของการเปลี่ยนคำตอบ ภายใต้การทดสอบที่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถแตกต่างกัน โดยแบ่งลักษณะของการเปลี่ยนคำตอบออกเป็น 2 ประเภท ได้แก่

4.1 การทวนคำตอบแล้วยังคงคำตอบเดิมไว้ ซึ่งวัดจากจำนวนข้อที่ไม่มี ความแตกต่างของคำตอบจากการทดสอบที่ให้ทวนคำตอบรอบที่ 1 และรอบที่ 2 พบว่า

4.1.1 ลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วยังคงคำตอบเดิมไว้ ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน ไม่พบว่ามี ความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 ระหว่างผู้สอบที่มีระดับความสามารถแตกต่างกัน แต่มีแนวโน้มต่ำลง (จำนวนข้อน้อยลง) เพียงเล็กน้อย เมื่อผู้สอบมีระดับความสามารถต่ำลง

4.1.2 จำนวนข้อที่ผู้สอบยังคงคำตอบเดิมไว้มีค่าอยู่ในช่วง 14 ถึง 20 ข้อ และมีค่าเฉลี่ยเท่ากับ 18.40 ข้อ

4.2 การทวนคำตอบแล้วมีการเปลี่ยนคำตอบ ซึ่งแบ่งเป็น 3 ลักษณะย่อยคือ จำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นถูก, จำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นผิด และจำนวนข้อที่เปลี่ยนคำตอบจากถูกเป็นผิด พบว่า

4.2.1 ลักษณะการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วมีการเปลี่ยนคำตอบจากผิดเป็นถูก จากผิดเป็นผิด และจากถูกเป็นผิด ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน พบว่ามีความแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 โดยเมื่อทดสอบความแตกต่างเป็นรายคู่ พบว่าจำนวนข้อที่มีการเปลี่ยนคำตอบจากผิดเป็นถูก จากผิดเป็นผิด และจากถูกเป็นผิด ของผู้สอบ แตกต่างกันในระดับความสามารถของผู้สอบที่แตกต่างกัน 2 คู่ คือ ผู้สอบที่มีระดับความสามารถสูง กับ ผู้สอบที่มีระดับความสามารถต่ำ และ ผู้สอบที่มีระดับความสามารถปานกลาง กับ ผู้สอบที่มีระดับความสามารถต่ำ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยจำนวนข้อที่ผู้สอบมีการเปลี่ยนคำตอบจากผิดเป็นถูกมีแนวโน้มต่ำลง (จำนวนข้อน้อยลง) เมื่อผู้สอบมีระดับความสามารถต่ำ ส่วนจำนวนข้อที่ผู้สอบมีการเปลี่ยนคำตอบจากผิดเป็นผิด และจากถูกเป็นผิดมีแนวโน้มสูงขึ้น (จำนวนข้อมากขึ้น) เมื่อผู้สอบมีระดับความสามารถต่ำ

4.2.2 ผู้สอบมีการเปลี่ยนคำตอบจากผิดเป็นถูกเฉลี่ย 1.96 ข้อ เปลี่ยนคำตอบจากผิดเป็นผิดและจากถูกเป็นผิด เฉลี่ย 1.50 และ 1.22 ข้อ ตามลำดับ โดยที่ผู้สอบที่มีระดับความสามารถต่ำมีการเปลี่ยนคำตอบจากผิดเป็นถูกน้อยที่สุด และผู้สอบที่มีระดับความสามารถต่ำมีการเปลี่ยนคำตอบจากผิดเป็นผิดและจากถูกเป็นผิดมากที่สุด

5. จากการเปรียบเทียบความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ตามเงื่อนไขการทดสอบ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน พบว่า

5.1 ผลการเปรียบเทียบค่าเฉลี่ยของค่าความสามารถของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบ ของผู้สอบที่มีระดับความสามารถสูง พบว่าผู้สอบที่มีระดับความสามารถสูง มีค่าเฉลี่ยของค่าความสามารถของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และผู้สอบที่มีระดับความสามารถปานกลาง และต่ำ พบว่าผู้สอบที่มีระดับความสามารถปานกลาง และต่ำ มีค่าเฉลี่ยของค่าความสามารถของผู้สอบระหว่างเงื่อนไขการไม่ให้ทวนคำตอบและการให้ทวนคำตอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

5.2 ผู้สอบที่มีระดับความสามารถแตกต่างกันมีความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แต่มีแนวโน้ม

สูงขึ้น เมื่อผู้สอบมีระดับความสามารถในช่วงปานกลางถึงสูง และมีแนวโน้มต่ำลง เมื่อผู้สอบมีระดับความสามารถในช่วงต่ำถึงปานกลาง

5.3 ผู้สอบส่วนใหญ่จะใช้เวลาในการทำข้อสอบในเงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบเฉลี่ย 19.90% โดยที่ผู้สอบที่มีระดับความสามารถสูงจะใช้เวลาในการทำข้อสอบในเงื่อนไขการให้ทวนคำตอบเพิ่มขึ้นจากเงื่อนไขการไม่ให้ทวนคำตอบมากที่สุด รองลงมาได้แก่ ผู้สอบที่มีระดับความสามารถต่ำ และผู้สอบที่มีระดับความสามารถปานกลางตามลำดับ

### อภิปรายผลการวิจัย

ผลการวิจัยเกี่ยวกับการเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน มีประเด็นสำคัญที่สามารถอภิปรายได้ดังนี้

1. โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ทางผู้วิจัยได้พัฒนาขึ้นนั้น ผู้วิจัยได้มีการพัฒนาต่อยอดจากโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ของรังสรรค์ มณีเล็ก (2540) ดังนี้

1.1 เปลี่ยนระบบปฏิบัติการจาก DOS เป็น Windows ทำให้มีความสะดวกและง่ายต่อการใช้งาน การพิมพ์ตัวหนังสือหรือการแทรกรูปภาพทำได้ง่ายขึ้นและมีความสวยงามมากขึ้น

1.2 เพิ่มเติมในเรื่องเงื่อนไขของการทดสอบ กล่าวคือโปรแกรมของผู้วิจัยสามารถกำหนดเงื่อนไขการทดสอบได้ 2 เงื่อนไข คือ ภายใต้การทดสอบที่ไม่ให้ทวนคำตอบและภายใต้การทดสอบที่ให้ทวนคำตอบโดยการทวนที่มีการจำกัดช่วง (Block Review) 5 ข้อ กล่าวคือ ผู้สอบจะสามารถทวนคำตอบพร้อมทั้งเปลี่ยนคำตอบภายในช่วงที่จำกัดไว้ได้ทั้งหมด แต่ไม่สามารถที่จะทวนข้ามช่วงได้ ซึ่งเป็นวิธีการที่ Vispole (2000) ได้คิดขึ้น และจากงานวิจัยของเขาพบว่า การทวนแบบจำกัดช่วงสั้นๆ นั้นทำหน้าที่ได้ดีพอๆ กับการทวนได้ทั้งหมด แต่สามารถลดปัญหาการโกงข้อสอบจากการเปิดโอกาสให้ผู้สอบสามารถทวนคำตอบได้ เนื่องจากผู้สอบต้องตอบข้อสอบให้เสร็จก่อนภายในช่วง จึงจะสามารถกลับไปทวนคำตอบได้ จึงเป็นการประมาณค่าความสามารถของผู้สอบไว้แล้วครั้งหนึ่ง รวมทั้งยังมีการประมาณค่าความสามารถของผู้สอบระหว่างช่วงด้วยเมื่อผู้สอบทำการทวนเสร็จในช่วง

1.3 ผู้วิจัยได้มีการพัฒนาคลังข้อสอบที่บรรจุในโปรแกรม เพิ่มเติมจากคลังเดิม เนื่องจากโปรแกรมของรังสรรค์ มณีเล็ก (2540) มีการยุติการทดสอบแบบแปรผันแต่ด้วยค่าสารสนเทศจากผลการวิจัยที่ผู้สอบต้องการนั้นจำเป็นจะต้องมีการเปรียบเทียบค่าระหว่างผู้สอบด้วยกัน ดังนั้นจึงต้องให้ผู้สอบแต่ละคนมีจำนวนข้อในการทำข้อสอบที่เท่ากัน มุ่งวัดลักษณะเดียวกัน ผู้วิจัยจึงให้โปรแกรมยุติการทดสอบแบบคงที่ที่ 20 ข้อ และจำเป็นที่จะต้องพัฒนาคลังข้อสอบคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 ของรังสรรค์ มณีเล็ก (2540) จำนวน 200 ข้อ เพิ่มเติม เนื่องจากการทดสอบที่มีการกำหนดจำนวนข้อที่จะใช้คงที่เหมือนกันทุกคน จำนวนข้อสอบควรมีอย่างน้อย  $[n(n+1)] \div 2$  ข้อ เมื่อ  $n$  คือจำนวนข้อสอบ (Allen & Yen, 1979) กล่าวคือ ผู้วิจัยต้องการสอบทุกคนด้วยข้อสอบ 20 ข้อเท่ากันจำนวนข้อสอบในคลังจึงควรมีอย่างน้อย 210 ข้อ ผู้วิจัยจึงได้พัฒนาคลังข้อสอบโดยออกแบบสอบปรนัยวิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ซึ่งเป็นข้อสอบเลือกตอบ 4 ตัวเลือก จำนวน 40 ข้อ (ออกใหม่ 30 ข้อ สุ่มร่วมจากคลังเดิม 10 ข้อ) สำหรับคัดเลือกเพื่อเพิ่มข้อสอบใหม่อย่างน้อย 10 ข้อเข้าคลังข้อสอบ ด้วยวิธีการปรับเทียบค่าพารามิเตอร์ด้วยวิธีค่าเฉลี่ยและซิกมา (Mean and Sigma Method) ผลปรากฏว่ามีข้อสอบรวมในคลังที่ค่าพารามิเตอร์ไม่ผ่านเกณฑ์ 2 ข้อ คือข้อสอบข้อที่ 3 และ 6 (ตรงกับข้อในคลังที่ 177 และ 30) โดยข้อที่ 3 มีค่าอำนาจจำแนกเท่ากับ 0.176 (น้อยกว่า 0.50) จึงไม่ผ่านเกณฑ์ และข้อที่ 6 มีค่าอำนาจจำแนกเท่ากับ 0.438 (น้อยกว่า 0.50) และค่าความยากเท่ากับ  $-2.725$  (น้อยกว่า  $-2.50$ ) จึงไม่ผ่านเกณฑ์ รวมทั้งมีข้อสอบที่ผู้วิจัยออกใหม่ที่มีค่าพารามิเตอร์ไม่ผ่านเกณฑ์ 4 ข้อ คือ ข้อสอบข้อที่ 4, 5, 10 และ 30 โดยข้อที่ 4 มีค่าอำนาจจำแนกเท่ากับ 0.371 (น้อยกว่า 0.50) ข้อที่ 5 มีค่าอำนาจจำแนกเท่ากับ 0.287 (น้อยกว่า 0.50) ข้อที่ 10 มีค่าอำนาจจำแนกเท่ากับ 0.480 (น้อยกว่า 0.50) ข้อที่ 30 มีค่าโอกาสในการเดาข้อสอบได้ถูกเท่ากับ 0.319 (มากกว่า 0.30) จึงไม่ผ่านเกณฑ์ ดังนั้นจึงเหลือข้อสอบที่มีค่าพารามิเตอร์ผ่านเกณฑ์ทั้งหมด 26 ข้อ ที่จะบันทึกเข้าคลังในโปรแกรมคอมพิวเตอร์ คือข้อสอบข้อที่ 1, 8, 9, 12, 14, 16, 17, 19, 28 และ 31-38 นำเข้าคลังใหม่เป็นข้อสอบข้อที่ 201-224 และเนื่องจากข้อสอบรวมที่ตรงกับข้อในคลังข้อที่ 177 และ 30 ไม่ผ่านเกณฑ์ จึงนำข้อ 37 และ 38 ไปแทนที่ข้อในคลังเดิม ทำให้สุดท้ายคลังข้อสอบที่ผู้วิจัยได้พัฒนาต่อยอดคั้งนั้นนี้มีข้อสอบรวม 224 ข้อ และเนื่องจากผู้วิจัยมีการยุติข้อสอบคงที่ที่ 20 ข้อ ดังนั้นในการเก็บรวบรวมข้อมูลผู้วิจัยจึงกำหนดเกณฑ์ความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 กล่าวคือถ้าผู้สอบที่ทำการทดสอบแล้วมีความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบสูงกว่า 0.3 จะไม่ถูกนำมาใช้ในการเก็บรวบรวมข้อมูลในครั้งนี้ ทั้งนี้จากการศึกษาของ รังสรรค์ มณีเล็ก (2540) พบว่าถ้ากำหนดความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 จะให้ความตรงเชิงสภาพสูง

2. ผู้สอบส่วนใหญ่มีความเห็นในด้านบวกต่อโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ทั้งในด้านความสะดวกในการใช้งาน รวมทั้งรูปแบบของโปรแกรมและคู่มือในการใช้งาน (ดูในภาคผนวก ค) มีความเหมาะสม อีกทั้งผู้สอบส่วนใหญ่มีความตั้งใจ สนใจ และกระตือรือร้นในการทำการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เป็นอย่างมาก โดยเมื่อผู้สอบทำการทดสอบเสร็จเรียบร้อยแล้ว ผู้สอบจะรีบดูผลการทดสอบที่โปรแกรมแสดงผลให้ดูทางหน้าจอ รวมทั้งมีการจดผลการทดสอบแล้วนำผลการทดสอบมาเปรียบเทียบระหว่างผู้สอบว่าใครได้ค่าความสามารถเท่าไร และมีผู้สอบบางคนขอทำการทดสอบใหม่อีกครั้ง อีกทั้งผู้สอบส่วนมากให้ข้อเสนอแนะว่าอยากให้มีการจัดการทดสอบแบบนี้ในการสอบกลางภาคและปลายภาคด้วย ทั้งนี้เนื่องจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เป็นการสอบรูปแบบใหม่ที่แตกต่างจากการสอบแบบประเมินนิยมอย่างเห็นได้ชัด กล่าวคือ ผู้สอบจะรู้สึกเสมือนว่ามีการโต้ตอบกับโปรแกรมคอมพิวเตอร์ ทำให้การสอบไม่น่าเบื่อ ซึ่งสอดคล้องกับการศึกษาของ ตรังใจ พูลผลอำนวย (2534) ที่ว่าการจัดการทดสอบโดยใช้คอมพิวเตอร์จะทำให้นักเรียนเกิดความสนใจ และมีความกระตือรือร้นในการทดสอบ

3. การเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ตามเงื่อนไขการทดสอบ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน พบว่า

3.1 ผู้สอบทุกระดับความสามารถจะได้ค่าความสามารถในเงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบ ทั้งนี้เนื่องจากการตัดสินใจเลือกคำตอบในครั้งแรก ผู้สอบส่วนใหญ่อาจเลือกคำตอบโดยความไม่แน่ใจ เพราะตัวเลือกทั้งตัวที่เป็นคำตอบและตัวลวงมีความน่าเชื่อถือ น่าสนใจ โดดเด่นกันมาก ซึ่งหลักการในการตั้งคำถาม และตัวเลือกแบบปรนัย เลือกตอบที่ดีข้อหนึ่งคือ การสร้างตัวลวงให้น่าเชื่อถือและน่าสนใจ แต่เมื่อมีการเปิดโอกาสให้ผู้สอบสามารถกลับไปทวนคำตอบได้ทำให้ผู้สอบมีโอกาสที่จะตรวจทานและตัดสินใจเลือกคำตอบที่ถูกต้องที่สุดได้อีกครั้งหนึ่ง ทั้งนี้เนื่องจากการทำงานใดๆ ก็ตามการตรวจทานถือว่าเป็นปัจจัยสำคัญที่ทำให้มีโอกาสในการหาข้อผิดพลาดในการทำงานเจอ และหลังจากที่พบข้อผิดพลาดในงานนั้นแล้ว ก็ทำให้สามารถที่จะนำข้อผิดพลาดมาปรับปรุง แก้ไข หรือพัฒนาการทำงานนั้นให้ดีขึ้น เช่นเดียวกับการเลือกข้อสอบให้ได้คำตอบที่ถูกต้องมากที่สุด ส่วนหนึ่งมาจากการทวนคำตอบเป็นการตรวจพบทวนและเพิ่มความรอบคอบในการหาคำตอบที่น่าจะถูกต้องที่สุดในข้อนั้นๆ ซึ่งเมื่อผู้สอบสามารถเลือกคำตอบที่ถูกต้องได้เพิ่มมากขึ้นจึงส่งผลให้ค่าความสามารถของผู้สอบในเงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบในทุกระดับความสามารถ ดังนั้นการ



ทดสอบที่ให้ทวนคำตอบด้วยวิธีการทวนที่มีการจำกัดช่วง (Block Review) จึงเป็นการส่งผลดีต่อผู้สอบในทุกระดับความสามารถ

3.2 ผู้สอบที่มีระดับความสามารถต่ำจะได้ค่าความสามารถในเงื่อนไขการให้ทวนคำตอบเพิ่มขึ้นจากเงื่อนไขการไม่ให้ทวนคำตอบมากที่สุด รองลงมาได้แก่ ผู้สอบที่มีระดับความสามารถปานกลาง และผู้สอบที่มีระดับความสามารถสูงตามลำดับ ทั้งนี้เนื่องจากผู้สอบที่มีระดับความสามารถปานกลางและสูงส่วนมากมีการทวนคำตอบแล้วคงคำตอบเดิมไว้มากกว่าผู้สอบที่มีระดับความสามารถต่ำ จึงส่งผลให้มีการเปลี่ยนแปลงของค่าความสามารถของผู้สอบน้อยกว่าผู้สอบที่มีระดับความสามารถต่ำ

ซึ่งสอดคล้องกับสมมติฐานที่ผู้วิจัยตั้งไว้ว่าผู้สอบที่มีระดับความสามารถสูงและปานกลางจะมีค่าความสามารถสูงขึ้นภายใต้เงื่อนไขการให้ทวนคำตอบ แต่มีส่วนที่ขัดแย้งกับสมมติฐานที่ผู้วิจัยตั้งไว้ในส่วนของผู้สอบที่มีระดับความสามารถต่ำน่าจะมีค่าความสามารถภายใต้เงื่อนไขการให้ทวนคำตอบน้อยกว่าการไม่ให้ทวนคำตอบ จึงทำให้เห็นได้อย่างชัดเจนว่าการทดสอบที่ให้ทวนคำตอบด้วยวิธีการทวนที่มีการจำกัดช่วง (Block Review) เป็นการส่งผลดีต่อผู้สอบในทุกระดับความสามารถ และจะเอื้อประโยชน์ต่อผู้สอบที่มีระดับความสามารถต่ำมากที่สุด

รวมทั้งผลที่ได้ยังสอดคล้องกับการศึกษาของ วอลเตอร์ พี. วิสโปเอล (Walter P. Vispoel, 1998) ที่ว่าการอนุญาตให้ทวนคำตอบนั้น ผู้สอบเกือบทั้งหมดได้รับค่าประมาณความสามารถเพิ่มขึ้นเล็กน้อยหลังการทวนคำตอบ

4. การเปรียบเทียบลักษณะของการเปลี่ยนคำตอบ ภายใต้การทดสอบที่ให้ทวนคำตอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถแตกต่างกัน โดยแบ่งลักษณะของการเปลี่ยนคำตอบออกเป็น 2 ประเภทได้แก่

4.1 การทวนคำตอบแล้วยังคงคำตอบเดิมไว้ ซึ่งวัดจากจำนวนข้อที่ไม่มี ความแตกต่างของคำตอบจากการทดสอบที่ให้ทวนคำตอบรอบที่ 1 และรอบที่ 2 พบว่า ลักษณะของการเปลี่ยนคำตอบประเภทการทวนคำตอบแล้วยังคงคำตอบเดิมไว้ ระหว่างผู้สอบที่มีระดับความสามารถของผู้สอบแตกต่างกัน ไม่พบว่ามี ความแตกต่างกัน ซึ่งสอดคล้องสมมติฐานที่ผู้วิจัยตั้งไว้ ทั้งนี้เนื่องจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing) เป็นการทดสอบที่ใช้แบบสอบต่างชุดกันสำหรับผู้สอบต่างกัน โดยมีการคัดเลือกข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบ (ศิริชัย กาญจนวาสี, 2545) จึงทำให้ผู้สอบได้รับข้อสอบที่มีค่าความยากในแต่ละกลุ่มระดับความสามารถเสมือนเทียบอยู่บนมาตรฐานเดียวกัน

กล่าวคือมีความยากเหมาะสมกับระดับความสามารถของผู้สอบในระดับความสามารถของตนเอง เหมือนกันในแต่ละกลุ่ม จึงส่งผลให้การทวนคำตอบแล้วยังคงคำตอบเดิมไว้ไม่แตกต่างกันในแต่ละระดับความสามารถของผู้สอบ

4.2 การทวนคำตอบแล้วมีการเปลี่ยนคำตอบ ซึ่งแบ่งเป็น 3 ลักษณะย่อยคือ จำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นถูก, จำนวนข้อที่เปลี่ยนคำตอบจากผิดเป็นผิด และจำนวนข้อที่เปลี่ยนคำตอบจากถูกเป็นผิด พบว่าจำนวนข้อที่มีการเปลี่ยนคำตอบจากผิดเป็นถูกจากผิดเป็นผิด และจากถูกเป็นผิด ของผู้สอบ แตกต่างกันในระดับความสามารถของผู้สอบที่แตกต่างกัน 2 คู่ คือ ผู้สอบที่มีระดับความสามารถสูง กับ ผู้สอบที่มีระดับความสามารถต่ำ และผู้สอบที่มีระดับความสามารถปานกลาง กับ ผู้สอบที่มีระดับความสามารถต่ำ โดยที่ผู้สอบที่มีระดับความสามารถต่ำมีการเปลี่ยนคำตอบจากผิดเป็นถูกน้อยที่สุด และมีการเปลี่ยนคำตอบจากผิดเป็นผิดและจากถูกเป็นผิดมากที่สุด ซึ่งสอดคล้องกับสมมติฐานที่ผู้วิจัยตั้งไว้ ทั้งนี้เนื่องจากผู้สอบที่มีระดับความสามารถต่ำปานกลางถึงสูง เป็นผู้ที่มีความพร้อมทางการเรียน โดยอาจมีความรู้และหรือมีประสบการณ์ รวมทั้งการทำความเข้าใจบทเรียนได้แม่นยำ น้อยกว่าผู้สอบที่มีระดับความสามารถปานกลางถึงสูง ซึ่งการมีข้อมูลมากและหลากหลายน้อยกว่าจึงส่งผลทำให้ผู้สอบที่มีระดับความสามารถต่ำมีการเปลี่ยนคำตอบจากผิดเป็นถูกน้อยที่สุด และมีการเปลี่ยนคำตอบจากผิดเป็นผิดและจากถูกเป็นผิดมากที่สุด ซึ่งสอดคล้องกับการศึกษาของ ประณาท เทียนศรี (2546) ที่ว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ระดับดีมาก ส่วนใหญ่เปลี่ยนคำตอบใหม่แล้วถูกต้องมากขึ้น แต่นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ระดับต่ำกว่าปานกลางและต้องปรับปรุง จะเปลี่ยนคำตอบใหม่แล้วผิดมากกว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ระดับดีกว่า

5. การเปรียบเทียบความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ตามเงื่อนไขการทดสอบเมื่อระดับความสามารถของผู้สอบแตกต่างกัน พบว่า ผู้สอบที่มีระดับความสามารถแตกต่างกันมีความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบไม่แตกต่างกัน โดยที่ผู้สอบส่วนใหญ่จะใช้เวลาในการทำข้อสอบในเงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบ ซึ่งสอดคล้องกับสมมติฐานที่ผู้วิจัยตั้งไว้ ทั้งนี้เนื่องจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing) เป็นการทดสอบที่ใช้แบบสอบต่างชุดกันสำหรับผู้สอบต่างกัน โดยมีการคัดเลือกข้อสอบที่มีความยากเหมาะสมกับระดับความสามารถของผู้สอบ (ศิริชัย กาญจนวาสี, 2545) ทำให้ผู้สอบได้รับข้อสอบที่มีค่าความยากในแต่ละกลุ่มระดับความสามารถเสมือนเทียบอยู่บนมาตรฐานเดียวกัน กล่าวคือมีความยากเหมาะสมกับระดับความสามารถของผู้สอบในระดับ

ความสามารถของตนเองเหมือนกันในแต่ละกลุ่ม รวมทั้งผู้วิจัยได้กำหนดจำนวนข้อสอบคงที่ที่ 20 ข้อ จึงส่งผลให้ความแตกต่างของเวลาที่ใช้ในการทำข้อสอบของผู้สอบไม่แตกต่างกันในแต่ละระดับความสามารถของผู้สอบ และเมื่อผู้สอบได้รับโอกาสในการกลับไปทวนคำตอบได้ ทำให้ผู้สอบมีโอกาที่จะตรวจทานและตัดสินใจเลือกคำตอบที่ถูกต้องที่สุดได้อีกครั้งหนึ่ง จึงต้องใช้เวลาในการตรวจทบทวนและเพิ่มความรอบคอบในการหาคำตอบที่น่าจะถูกที่สุดที่สุดในข้อนั้นๆ ในเงื่อนไขการให้ทวนคำตอบมากกว่าการไม่ให้ทวนคำตอบในทุกระดับความสามารถของผู้สอบ ซึ่งสอดคล้องกับการศึกษาของ วอลเตอร์ พี. วิสโพเอล (Walter P. Vispoel, 2000) ที่ว่าการอนุญาตให้ทวนคำตอบโดยวิธีการจำกัดช่วงนั้น จะส่งผลให้เวลาที่ผู้สอบใช้ในการทำข้อสอบมีระยะเวลาเพิ่มมากขึ้น ไม่ว่าจะจำกัดที่ช่วงกี่ข้อก็ตาม

### ข้อเสนอแนะ

ในการวิจัยนี้ผู้วิจัยได้แบ่งข้อเสนอแนะออกเป็น 2 ตอน คือ

**ตอนที่ 1** ข้อเสนอแนะในการวิจัยครั้งนี้ ซึ่งแบ่งออกเป็น 3 ตอนย่อย คือ

- 1.1 ข้อเสนอแนะเชิงวิชาการ
- 1.2 ข้อเสนอแนะเชิงนโยบาย
- 1.3 ข้อเสนอแนะในการปฏิบัติ

**ตอนที่ 2** ข้อเสนอแนะในการวิจัยครั้งถัดไป

**ตอนที่ 1** ข้อเสนอแนะในการวิจัยครั้งนี้ ซึ่งแบ่งออกเป็น 3 ตอนย่อย คือ

#### 1.1 ข้อเสนอแนะเชิงวิชาการ

1.1.1 จากการวิจัยในครั้งนี้รวมทั้งการวิจัยเกี่ยวกับการทวนคำตอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ผ่านมา ทำให้เห็นถึงข้อสรุปที่ชัดเจนว่า ในการจัดการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ภายใต้เงื่อนไขการให้ผู้สอบสามารถกลับไปทวนคำตอบได้ด้วยวิธีการทวนที่มีการจำกัดช่วงสั้นๆ และมีการยุติข้อสอบแบบคงที่นั้น จะส่งผลดีต่อตัวผู้สอบมากกว่าการจัดการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบในปัจจุบันที่ยังคงไม่เปิดโอกาสให้ผู้สอบกลับไปทวนคำตอบได้ กล่าวคือถ้ามีการให้ผู้สอบสามารถกลับไปทวนคำตอบได้อีกครั้งนั้น ก็จะทำให้ผู้สอบได้รับโอกาสในการกลับไปคิดทบทวนเพื่อเลือกคำตอบที่ดีที่สุด ถูกต้องที่สุด และตรงกับความสามารถที่แท้จริงมากที่สุด รวมทั้งลดความกังวลในการทำข้อสอบ การอ่านโจทย์ไม่รอบคอบ การลังเลใจ หรือแม้กระทั่งความสะเพร่าในการเลือกคำตอบในครั้งแรกอีกด้วย ดังนั้นจึงควรมีการ

เปลี่ยนรูปแบบการทำงานของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์โดยเปิดโอกาสให้ผู้สอบกลับไปทบทวนและสามารถแก้ไขคำตอบได้อีกครั้ง ภายใต้วงจำกัดสั้นๆ

1.1.2 การศึกษาและวิจัยทางการทดสอบแนวใหม่ ควรให้ความสำคัญ ในการใช้เทคโนโลยีคอมพิวเตอร์สำหรับการทดสอบให้มากขึ้น เนื่องจากเทคโนโลยีคอมพิวเตอร์ สามารถเข้ามาช่วยให้การทดสอบบนพื้นฐานของทฤษฎีการตอบสนองข้อสอบมีความเป็นไปได้ มี ประสิทธิภาพสูงและถูกต้องแม่นยำมากยิ่งขึ้น นอกจากนี้ยังสามารถประมวลผลได้อย่างรวดเร็ว รายงานผลได้ทันทีและมีความถูกต้องแม่นยำว่าการทดสอบแบบประเพณีนิยม ดังนั้นจึงมีความ เหมาะสมกับสังคมในยุคโลกาภิวัตน์

## 1.2 ข้อเสนอแนะเชิงนโยบาย

1.2.1 เนื่องจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ มีข้อดีในหลายๆ ด้าน เช่น ใช้เวลาในการทดสอบน้อยลงกว่าการทดสอบด้วย กระดาษแบบเดิม สามารถรายงานผลได้ทันทีหลังจากเสร็จสิ้นการทดสอบ ทำให้ผู้สอบมีเจตคติที่ดี ต่อการทดสอบเพราะผู้สอบแต่ละคนจะได้ข้อสอบที่เหมาะสมกับความสามารถของตนเอง ผู้สอบที่ มีความสามารถสูงก็ไม่รู้สึกเบื่อหน่ายในการทำข้อสอบที่ง่ายเกินไป ส่วนผู้ที่มีความสามารถต่ำก็ไม่ เกิดความตึงเครียดในการทำข้อสอบที่ยากเกินไป อีกทั้งยังสามารถทดสอบนักเรียนเป็นรายบุคคล ได้โดยที่ไม่ต้องรอสอบพร้อมๆ กัน ช่วยรักษาความลับและป้องกันการทุจริตในการสอบได้เป็น อย่างดี เพราะข้อสอบทั้งหมดถูกเก็บเอาไว้ในลักษณะของคลังข้อสอบ และในการทดสอบนักเรียน ที่นั่งติดกันก็ไม่สามารถลอกคำตอบกันได้เนื่องจากจะได้รับข้อสอบไม่เหมือนกัน ดังนั้นภาครัฐบาล ควรเห็นความสำคัญของการจัดการทดสอบในลักษณะนี้ และส่งเสริมให้โรงเรียนที่มีความพร้อม และมีศักยภาพทางการเรียนการสอนคอมพิวเตอร์ได้เริ่มพัฒนาการทดสอบแบบปรับเหมาะ กับ ความสามารถของผู้สอบด้วยคอมพิวเตอร์ให้แพร่หลายมากยิ่งขึ้น ทั้งในด้านการวัดผลระหว่างภาค กลางภาค ปลายภาค หรือแม้แต่การสอบแข่งขัน หรือสอบเข้าต่างๆ โดยเริ่มจากการให้โรงเรียนที่มี ศักยภาพเป็นโรงเรียนแกนนำ และค่อยกระจายสู่โรงเรียนที่มีศักยภาพรองลงมา

1.2.2 ควรมีการเผยแพร่องค์ความรู้เกี่ยวกับการทดสอบแบบปรับเหมาะ กับความสามารถของผู้สอบให้แพร่หลายในประเทศไทยมากขึ้นกว่าเดิม เนื่องจากในอนาคตการ นำคอมพิวเตอร์มาช่วยในการทดสอบ (computerized testing) แทนการทดสอบโดยใช้กระดาษ การใช้คอมพิวเตอร์ตรวจให้คะแนนและรายงานผลการทดสอบ รวมทั้งการทดสอบแบบปรับเหมาะ กับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะเป็นส่วนหนึ่งในการทดสอบต่างๆ อย่างแน่นอน

เนื่องจากสามารถประมาณค่าความสามารถของผู้สอบได้อย่างแม่นยำ มีความคลาดเคลื่อนต่ำ ดังนั้นผู้ที่มีหน้าที่เกี่ยวข้องกับการจัดการเรียนการสอน โดยเฉพาะผู้ที่มีหน้าที่โดยตรงเกี่ยวกับการวัดผลทางการศึกษาจึงควรที่จะมีความรู้ความเข้าใจในการทดสอบลักษณะนี้ รวมทั้งองค์กรต่างๆ ที่เกี่ยวข้องกับการทดสอบทางการศึกษาคควรมีการพัฒนาบุคลากรเพื่อให้มีองค์ความรู้ทางศาสตร์นี้เพิ่มมากขึ้น

### 1.3 ข้อเสนอแนะในการปฏิบัติ

1.3.1 โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ วิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 ภายใต้การทดสอบที่ให้ทวนคำตอบด้วยวิธีการทวนที่มีการจำกัดช่วง (Block Review) ที่ผู้วิจัยพัฒนาขึ้นเพื่อวัดความเข้าใจในเรื่องเศษส่วน ครูผู้สอนสามารถนำไปใช้กับผู้สอบที่กำลังศึกษาอยู่ในระดับชั้นประถมศึกษาปีที่ 6 ในทุกๆ โรงเรียนที่มีระบบการเรียนการสอนด้วยคอมพิวเตอร์ได้

1.3.2 เนื่องจากโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ วิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยพัฒนาขึ้นนี้เป็นโปรแกรมสำเร็จรูป ดังนั้นผู้นำไปใช้และผู้สอบจึงไม่สามารถเปลี่ยนแปลงคลังข้อสอบและเกณฑ์ยุติการทดสอบ 20 ข้อ ที่ผู้วิจัยกำหนดไว้ได้ แต่ครูผู้สอนที่สนใจและมีทักษะความรู้ในด้านการเขียนโปรแกรมสามารถนำเอา CODE การเขียนโปรแกรม (ภาคผนวก ข) ของทางผู้วิจัยไปประยุกต์ใช้เพื่อเปลี่ยนแปลงระบบการทำงานของ การทดสอบ หรือแก้ไขข้อสอบในคลังเพื่อให้ตรงกับความต้องการของผู้สอบที่จะจัดการทดสอบในแต่ละวิชาได้

### ตอนที่ 2 ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 การวิจัยครั้งนี้เป็นการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ผู้สอบสามารถกลับไปทวนคำตอบได้ด้วยวิธีการทวนที่มีการจำกัดช่วง (Block Review) ที่ใช้กลุ่มข้อสอบที่ตรวจให้คะแนนแบบถูก-ผิด จากการศึกษาของนักวิจัยเกี่ยวกับแบบทดสอบแบบประเพณีนิยม พบว่าการตรวจให้คะแนนแบบหลายค่าจะให้สารสนเทศในการทดสอบมากกว่าการตรวจให้คะแนนแบบถูก-ผิด ดังนั้นจึงควรมีการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ผู้สอบสามารถกลับไปทวนคำตอบได้ด้วยวิธีการทวนที่มีการจำกัดช่วง (Block Review) ที่ใช้กับกลุ่มข้อสอบที่ให้คะแนนแบบหลายค่าเพื่อให้ได้สารสนเทศในการทดสอบที่มากขึ้น

2.2 ควรมีศึกษาเปรียบเทียบผลจากการอนุญาตให้มีการทวนคำตอบด้วยวิธีการทวนที่มีการจำกัดช่วง (Block Review) ทั้งในด้านค่าความสามารถของผู้สอบ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ระหว่างที่มีการยุติข้อสอบคงที่กับแบบแปรผัน

2.3 ควรมีศึกษาตรวจสอบความตรงตามสภาพ หรือ ร่วมสมัย (Concurrent validity) โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน หรือค่าความสามารถของผู้สอบที่วัดได้จากโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบที่สร้างขึ้น กับคะแนน หรือค่าความสามารถของผู้สอบที่วัดได้จากเครื่องมือมาตรฐานอื่น ซึ่งสามารถวัดสิ่งนั้นได้ในสภาพปัจจุบัน เพื่อเป็นการตรวจสอบความถูกต้องและความแม่นยำของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบที่สร้างขึ้นในการวัดสิ่งที่ต้องการจะวัด



สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

## รายการอ้างอิง

ภาษาไทย

- กนกวรรณ รัตนธน. 2544. การเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบแตกต่างกัน. วิทยานิพนธ์ปริญญา มหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย.
- กานดา ทองวัฒนะ. 2522. การเปลี่ยนคำตอบในการทำข้อสอบชนิดเลือกตอบ. วิทยานิพนธ์ ปริญญา มหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย.
- เกรียงศักดิ์ สุวรรณภาค. 2532. ความสัมพันธ์ระหว่างรูปแบบบางประการของแบบทดสอบแยกกลุ่มกับความสามารถทางการเรียน ในวิชาคณิตศาสตร์. วิทยานิพนธ์มหาบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ขวัญใจ อุดราภรณ์. 2527. การเปลี่ยนคำตอบในการสอบซ้ำ ของแบบสอบเลือกตอบ วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาตอนต้น. วิทยานิพนธ์ปริญญา มหาบัณฑิต จุฬาลงกรณ์ มหาวิทยาลัย.
- จิราพร ไกรสรศิริเวท. 2529. ความสัมพันธ์ระหว่างรูปแบบบางประการของแบบทดสอบรูปปริมาตร กับความสามารถทางการเรียนในวิชาคณิตศาสตร์. วิทยานิพนธ์มหาบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ตริงใจ พูนผลอำนวย. (2534). พัฒนาแบบสอบเฉพาะบุคคลในวิชาวิทยาศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้คอมพิวเตอร์. วิทยานิพนธ์ปริญญา มหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย.
- ต่าย เชียงฉี. 2534. การศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบเทเลอร์รูปปริมาตรที่มีรูปแบบ จำนวนชั้นและวิธีการให้คะแนนที่แตกต่างกันโดยใช้วิธีมอนติคาร์โล. วิทยานิพนธ์ดุชะฎิบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ
- นันทิยา พึ่งคำ. 2531. การเปรียบเทียบคุณภาพการทดสอบแบบซีไอทีและแบบประเพณีนิยมในการวัดความสามารถด้านคำศัพท์ภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3. วิทยานิพนธ์ปริญญา มหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย.
- ประณาท เทียนศรี. 2546. ปัจจัยและผลการเปลี่ยนคำตอบข้อสอบปรนัยของนักเรียนชั้นประถมศึกษาปีที่ 5 โรงเรียนสาธิต จุฬาลงกรณ์มหาวิทยาลัย. วารสารค殊ศาสตร์ 31 : 10-23.

- พลากร กรพิทักษ์. 2532. การพัฒนาโปรแกรมคอมพิวเตอร์สำหรับการจัดเก็บข้อสอบโดยใช้ไมโครคอมพิวเตอร์. วิทยานิพนธ์มหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย.
- ระพีพรรณ ศรีวิเชียร. 2534. การพัฒนาแบบสอบปรับเหมาะกับความความสามารถของผู้สอบในวิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3. วิทยานิพนธ์ปริญญามหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย.
- รังสรรค์ มณีเล็ก. 2540. ผลของตัวแปรบางตัวต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบ ที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์. วิทยานิพนธ์การศึกษาศาสตรดุษฎีบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ศิริชัย กาญจนวาสี และคณิต ไช้มุกต์. 2535. การพัฒนาโปรแกรมคอมพิวเตอร์สำหรับวิเคราะห์ข้อสอบและประมาณค่าความสามารถของผู้สอบตามทฤษฎีการตอบสนองข้อสอบด้วยวิธีของเบส์. กรุงเทพมหานคร : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. 2538. การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive testing). กรุงเทพมหานคร : ศูนย์ตำราและเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. 2541. การปรับเทียบคะแนนระหว่างแบบสอบ (Test Equating). กรุงเทพมหานคร : ศูนย์ตำราและเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. 2545. ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory). พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- สงบ ลักษณะ. 2525. การเทียบคะแนนระหว่างแบบทดสอบ. วารสารการวัดผลศึกษา 4 (กันยายน-ธันวาคม) : 21-32.
- สายชล อบทม. 2539. การพัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์. วิทยานิพนธ์ปริญญามหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย.
- สุพัฒน์ สุกมลสันต์. 2535. ธนาคารข้อสอบและการทดสอบปรับเปลี่ยนด้วยคอมพิวเตอร์. กรุงเทพมหานคร : บริษัทพิมพ์ดี จำกัด.
- เสรี ชัดเข้ม. 2541. การใช้คอมพิวเตอร์ในการวัดและประเมินผลทางการศึกษา. เอกสารหลักสูตรวิทยาศาสตรมหาบัณฑิต. ชลบุรี: สาขาวิชาเทคโนโลยีการวัดทางการศึกษา (หลักสูตรใหม่ พ.ศ. 2542) ภาควิชาวิจัยและวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา. (อัดสำเนา)



เสรี ปรมชวลิตโรจน์. 2531. การศึกษาความเที่ยงตรงในการจำแนกความรอบรู้และการประมาณค่าความสามารถในวิชาคณิตศาสตร์ของแบบทดสอบเฟล็กซีเลเวลและแบบทดสอบรูปปริมาตรที่ใช้ขนาดชั้นแปรผัน. ปรินซ์นิพนธ์มหาบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ.

ภาษาอังกฤษ

Bergstrom, B. A., and Lunz, M. E. 1992. Confidence in Pass/Fail Decisions for Computer Adaptive and Paper and Pencil Examinations. Evaluation and the Health Professions. 15 (April) : 453-464.

Bergstrom, B. A., and Stahl, J. A. 1992. Assessing Existing Item Bank Depth for Computer Adaptive Testing. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.

Blackmore, L. M. 1986. Computerized Adaptive and Pencil and Paper Test Administration: A Comparative Study in High School Setting. Dissertation Abstracts International 47 (April) : 2554 A.

Chang, S. 1990. A Comparative Study of Item Exposure Control Methods in Computerized Adaptive Testing (Act, Test Security). Dissertation Abstracts International 59 (September) : 3418 A.

De Ayala, R. J. 1992. The Nominal Response Model in Computerized Adaptive Testing. Applied Psychological Measurement 16 : 327-343.

De Ayala, R. J., Dodd, G. B., and Koch, R. W. 1990. A Simulation and Comparison of Flexi level and Bayesian Computerized Adaptive Testing. Journal of Educational Measurement 27 : 227-239.

Dodd, B. G., Koch, W. R., and De Ayala, R. J. 1993. Computerized Adaptive Testing using the Partial Credit Model: Effects of Item Pool Characteristics and Different Stopping Rules. Educational and Psychological Measurement 53 : 61-77.

Dodd, B. G. 1995. Computerized Adaptive Testing with Polytomous Item. Applied Psychological Measurement 19 : 5-19.

Ferrara, S., Frances, A., Gilmartin, D., Knott, T., Michaels, H., Pollack, J., Schuder, T., Vaeth, R., and Wise, S. 1996. A qualitative study of the information examinees consider during item review on a computer-adaptive test. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.

- Green, B. F., Bock, R. D., Humphreys, L. G., Linn, R. L., and Reckase, M. D. 1984. Technical Guidelines for Assessing Computerized Adaptive Tests. Journal of Educational Measurement 21 : 347-360.
- Hambleton, R. K., Swaminathan H., and Rogers, H. J. 1991. Fundamentals of Item Response Theory. London.
- Hulin, C. L., Darsgow, F., and Parsons. 1983. Item Response theory: Application to Psychological Measurement. Homewood IL : Dow-Jones.
- Hankins, J. A. 1986. The Effort of Variable Entry on Bias and Information of the Bayesian Adaptive Testing Procedure. Dissertation Abstracts International 47(August) : 3013 A.
- Ho, R. G. 1988. Using MicroCAT in Computerized Adaptive Testing: A Comparison of Three Adaptive Testing Strategies. Dissertation Abstracts International 50(February) : 421 A.
- Javier revulta, M., Carmen Ximenez, and Julio Olea. 2003. Psychometric and psychological effects of item selection and review on Computerized Adaptive Testing. Educational and Psychological Measurement 63, 5 (October): 791-808.
- Kalohn, J. C., and Spray, J. A. 1998. Test Security and Item Exposure Control for Computer – Based Examination. Performance of a Computerized Classification Test for Professional Certification.
- Kingsbury, G. G. 1996. Item review and adaptive testing. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.
- Lord, F. M. 1971. A Theoretical Study of the Measurement Effectiveness of Flexi level Test. Educational and Psychological Measurement 31 : 805-813.
- Lord, F. M. 1980. Applications of Item Response Theory to Practical Testing Problem. New Jersey : Lawrence Erlbaum Associates.
- Lunz, M. E., Bergstrom, B. A., and Wright, B. D. 1992. The Effect of Review on Student Ability and Test Efficiency for Computerized Adaptive Tests. Applied Psychological Measurement 16(January) : 33-40.

- Lunz, M. E., Stahl, J. A., and Bergstrom, B. A. 1993. Test Targeting and Precision Before and After Review on Computerized Adaptive Tests. Paper presented at the annual meeting of the National Council of Measurement in Education in Atlanta, Georgia.
- Lunz, M. E., and Stone, G. E. 1994. The effect of review on the psychometric characteristics of computerized adaptive tests. Applied Measurement in Education 7(March) : 211-222.
- Olea, J., Revuelta, J., Ximenez, M. C., and Abad, F. J. 2000. Psychometric and psychological effects of review on computerized fixed and adaptive tests. Educational and Psychological Measurement 21 : 157-173.
- Owen, and Roger J. 1975. A Bayesian Sequential Procedure or Quintal Response in the Context of Adaptive Mental Testing. Journal of the American Statistical Association 70(June) : 351-356.
- Revuelta, J., and Ponsoda, V. 1998. Comparison of Item Exposure Control Methods in Computerized Adaptive Testing. Journal of Education Measurement 34(April) : 311-327.
- Sekula-Wacura, R., and Brito, C. 2000. A Review of CAT Review. Laboratory Medicine 31 (August) : 442-444.
- Stone, G. E., and Lunz, M. E. 1993. The Effect of Review on the Psychometric Characteristics of Computerized Adaptive Tests. Applied Measurement in Education 7 : 211-222.
- Styles, I., and Andrich, D. 1993. Linking the Standard and Advanced Forms of the Raven' Progressive Matrices in Both the pencil-and-paper and Computer Adaptive Testing Formats. Educational and Psychological Measurement 53 : 906-924.
- Stocking, M. L., and Lewis, C. 1997. Corraling Item Exposure Conditional on Ability in Computerized Adaptive Testing. Journal of Education and Behavioral Statistics 23(January) : 57-75.
- Thorndike, R. K. 1982. Applied Psychometrics. Boston: Houghton Miffling Company.
- Urry, V. W. 1977. Tailored Testing: A Successful Application of Latent Trait Theory. Journal of Education Measurement 14(February) : 181-196.

- Vispoel, W. P. 1998. Reviewing and changing answers on computer-adaptive and self-adaptive vocabulary tests. Journal of Educational Measurement 35 : 328-345.
- Vispoel, W. P., Rocklin, T. R., Wang, T., and Bleiler, T. 1999. Can examinees use a review option to obtain positively biased ability estimates on a computerized adaptive test. Journal of Educational Measurement 36: 141-157.
- Vispoel, W. P., Hendrickson, A. B., Bleiler, T., Widiatmo, H., Shari, S., and Ihrig, D. 1999. Limiting answer review and change on computerized adaptive vocabulary tests: Psychometric and attitudinal results. Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, Canada.
- Vispoel, W. P., Hendrickson, A. B., and Bleiler, T. 2000. Limiting answer review and change on computerized adaptive vocabulary tests: Psychometric and attitudinal results. Journal of Educational Measurement 37 : 21-38.
- Wainer, H., Dorans, J., Flaugher, R., Green, F., Mislevy, J., Steinberg, L., and Thissen, D. 1993. Test Validity Measurement. Research Report. Psychometric Methods Program, Department of Psychology, University of Minnesota.
- Weiss, D. J., and Kingsbury, G. G. 1984. Application of Computerized Adaptive testing to Education Problem. Journal of Education Measurement 21(April): 361-375.
- Wise, S. L. 1995. Item review and answer changing in computerized adaptive tests. Paper presented at the Third European Conference on Psychological Assessment, Trier, Germany.
- Zwick, R., Thayer, T. D., and Wingersky, M. 1995. Effect of Rasch Calibration on Ability and DIF Estimation in Computer Adaptive Tests. Journal of Educational Measurement 32 : 341-363.



ภาคผนวก

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก ก

ตัวอย่าง: แบบสอบปรนัยวิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ชั้นประถมศึกษาปีที่ 6  
จำนวน 40 ข้อ

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก

# แบบสอบปรนัยวิชาคณิตศาสตร์

เรื่อง ความเข้าใจเศษส่วน      ชั้นประถมศึกษาปีที่ 6      จำนวน 40 ข้อ


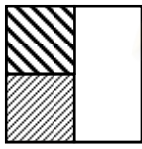

ข้อสอบที่นำมาทดลองใช้ 30 ข้อ และข้อสอบรวมที่สุ่มจากคลังข้อสอบของ ดร.รังสรรค์ มณีเล็ก  
จำนวน 10 ข้อ

\*\*หมายเหตุ ข้อสอบรวมที่สุ่มจากคลังจะมีสัญลักษณ์  ที่ข้อนั้นๆ

วัตถุประสงค์เพื่อ:      เพิ่มข้อสอบเข้าคลังอย่างน้อย 10 ข้อ จากเดิมที่มีอยู่ในคลัง 200 ข้อ ด้วย  
วิธีการปรับเทียบค่าพารามิเตอร์ด้วยวิธีค่าเฉลี่ยและซิกมา (Mean and  
Sigma Method)

จุดประสงค์ที่ 1      มีความเข้าใจพื้นฐานเกี่ยวกับเศษส่วน

## 1.1 ความหมายของเศษส่วน

ข้อ 2  133	 พื้นที่  มีความหมายตรงกับข้อใด				
	ตอบ	①	$\frac{1}{2} \times \frac{1}{2}$	②	$\frac{1}{3} \times \frac{1}{3}$
		③	$\frac{1}{2} \times \frac{1}{4}$	④	$\frac{1}{4} \times \frac{1}{4}$

## 1.2 ชนิดของเศษส่วน

ข้อ 4	ข้อใดจัดเป็นเศษส่วนชนิดเดียวกัน				
	ตอบ	①	$\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$	②	$\frac{2}{5}, \frac{6}{3}, \frac{4}{7}$
		③	$\frac{3}{7}, \frac{4}{8}, \frac{5}{9}$	④	$\frac{1}{2}, \frac{1}{5}, \frac{1}{5}$

จุดประสงค์ที่ 4 เมื่อกำหนดเศษส่วนใดๆ สามารถบวก, ลบ, คูณ และหารระคนได้

## 4.1 การบวกและลบ

ข้อ 16	$2\frac{1}{2} + 4 = \square + \frac{1}{2}$ ต้องเติมจำนวนใดลงใน $\square$				
	ตอบ	①	5	②	6
		③	7	④	8

## 4.2 การคูณและการหาร

ข้อ 22	$\frac{3}{4} \times 4$ มีความหมายตรงกับข้อใด			
ตอบ ๓	๑	$\frac{12}{4} + \frac{12}{4} + \frac{12}{4} + \frac{12}{4}$	๒	$\frac{12}{4} \times \frac{12}{4} \times \frac{12}{4} \times \frac{12}{4}$
	๓	$\frac{3}{4} + \frac{3}{4} + \frac{3}{4} + \frac{3}{4}$	๔	$\frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4}$

ข้อ 23	$1\frac{1}{3} \times \frac{\square}{5} = \frac{8}{15}$ ต้องเติมจำนวนใดลงใน $\square$			
ตอบ ๑	๑	2	๒	3
	๓	4	๔	5

จุดประสงค์ที่ 6 เมื่อกำหนดโจทย์ปัญหาเกี่ยวกับการบวก ลบ คูณ หาร เศษส่วนให้ สามารถแสดงวิธีทำได้

ข้อ 34	ที่ดินแปลงหนึ่งมีพื้นที่ 5 ไร่ แบ่งปลูกไม้ผล $3\frac{3}{4}$ ไร่ ปลูกผักสวนครัว $\frac{3}{5}$ ไร่ ที่เหลือเป็นสนามหญ้าอีกเท่าไร			
ตอบ ๔	๑	$5 - 3\frac{3}{4} + \frac{3}{5} = \square$	๒	$3\frac{3}{4} + \frac{3}{5} - 5 = \square$
	๓	$5 - \left(3\frac{3}{4} - \frac{3}{5}\right) = \square$	๔	$5 - \left(3\frac{3}{4} + \frac{3}{5}\right) = \square$

ข้อ 40 ๑ 138	เลี้ยงปลาไว้ 150 ตัว ขายครั้งแรกไปครึ่งหนึ่ง ครั้งที่สองขายไป $\frac{1}{5}$ ของปลาที่เหลือ ครั้งที่สองขายปลาไปเท่าไร โจทย์ข้อนี้เขียนเป็นประโยคสัญลักษณ์ได้อย่างไร			
ตอบ ๔	๑	$\frac{1}{2} \times 150 = \square$	๒	$\frac{1}{5} \times 150 = \square$
	๓	$\frac{1}{2} + \left(\frac{1}{5} \times 150\right) = \square$	๔	$\frac{1}{5} \times \left(\frac{1}{2} \times 150\right) = \square$





**ภาคผนวก ข**

แบบฟอร์มการตรวจสอบความตรงตามเนื้อเรื่อง

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข

## ตารางตรวจสอบความตรงตามเนื้อเรื่อง (Content Validity)

### ของแบบสอบปรนัยวิชาคณิตศาสตร์

เรื่อง ความเข้าใจเศษส่วน      ชั้นประถมศึกษาปีที่ 6      จำนวน 40 ข้อ

ข้อสอบที่นำมาทดลองใช้ 30 ข้อ และข้อสอบร่วมที่สุ่มจากคลังข้อสอบของ ดร.รังสรรค์ มณีเล็ก  
จำนวน 10 ข้อ

\*\* หมายเหตุ ข้อสอบร่วมที่สุ่มจากคลังจะมีสัญลักษณ์ **8** ที่ข้อนั้นๆ

วัตถุประสงค์เพื่อ:      เพิ่มข้อสอบเข้าคลังอย่างน้อย 10 ข้อ จากเดิมที่มีอยู่ในคลัง 200 ข้อ ด้วย  
วิธีการปรับเทียบค่าพารามิเตอร์ด้วยวิธีค่าเฉลี่ยและซิกมา (Mean and  
Sigma Method)

ชื่อผู้เชี่ยวชาญ

⇒ **คำชี้แจง** : โปรดพิจารณาตรวจสอบความตรงตามเนื้อเรื่องของข้อคำถาม 30 ข้อ ตามแต่ละ  
จุดประสงค์ว่ามีความสอดคล้องหรือไม่เพียงใด โดยเขียนเครื่องหมาย  $\surd$  ลงในช่องว่างตามระดับ  
ที่ท่านพิจารณาเห็นสมควร

\*\* หมายเหตุ\*\*

-1 หมายถึง ไม่สอดคล้อง

0 หมายถึง ไม่แน่ใจ

1 หมายถึง สอดคล้อง

จุดประสงค์ที่	ข้อที่	ระดับความเห็น			ข้อเสนอแนะอื่นๆ
		-1	0	1	
1. มีความเข้าใจพื้นฐานเกี่ยวกับเศษส่วน					
1.1 ความหมายของเศษส่วน	1				
1.2 ชนิดของเศษส่วน	4				
1.3 เศษส่วนของจำนวนนับ	5				
1.4 เศษส่วนอย่างต่ำ	8				
	9				
1.5 การแปลงเศษเกินเป็นจำนวนคละ	10				

จุดประสงค์ที่	ข้อที่	ระดับความเห็น			ข้อเสนอแนะอื่นๆ
		-1	0	1	
2. เมื่อกำหนดเศษส่วนให้สามารถเขียนเป็นเศษส่วนที่มีค่าเท่าเดิมโดยที่ตัวเศษหรือตัวส่วนมีค่าตามที่กำหนดให้ได้	12				
3. เมื่อกำหนดเศษส่วนมาให้ สามารถเปรียบเทียบและใช้สัญลักษณ์ $>$ , $<$ หรือ $=$ ได้	13				
	14				
4. เมื่อกำหนดเศษส่วนใดๆ สามารถบวก, ลบ, คูณ และหารระคนได้ 4.1 การบวกและลบ	16				
	17				
	19				
	20				
	21				
4.2 การคูณและหาร	22				
	23				
	24				
	25				
	26				
	27				
5. เมื่อกำหนดเศษซ้อนอย่างง่ายที่เฉพาะตัวเศษหรือตัวส่วนที่มีการบวกลบคูณหารอย่างใดอย่างหนึ่งให้สามารถหาผลลัพธ์ได้	28				
	30				
6. เมื่อกำหนดโจทย์ปัญหาเกี่ยวกับการบวก ลบ คูณ หาร เศษส่วนให้สามารถแสดงวิธีทำได้	31				
	32				
	33				
	34				
	35				
	36				
	37				
	38				



ภาคผนวก ค

หนังสือ: ขอเชิญเป็นผู้ทรงคุณวุฒิตรวจสอบเครื่องมือวิจัย

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ค

ที่ ศร ๐๕๑๒.๖(๒๗๕๕)/



ภาควิชาวิจัยและจิตวิทยาการศึกษา  
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย  
ถนนพญาไท เขตปทุมวัน กรุงเทพฯ ๑๐๓๓๐

๒๖ มกราคม ๒๕๕๙

เรื่อง ขอเชิญเป็นผู้ทรงคุณวุฒิตรวจสอบเครื่องมือวิจัย  
เรียน ผู้ช่วยศาสตราจารย์ ดร. ณัฐภรณ์ หลาวทอง  
สิ่งที่ส่งมาด้วย เครื่องมือที่ใช้ในการวิจัย

ด้วย นางสาวพิมพ์สิริ เขียวนรเศรษฐ์ นิสิตระดับปริญญาโท สาขาการวัดและประเมินผลการศึกษาภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อยู่ในระหว่างการทำวิทยานิพนธ์ เรื่อง “การเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน” โดยมีผู้ช่วยศาสตราจารย์ ดร.เอมอร จังศิริพรปกรณ์ เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์ ในการนี้จึงใคร่ขอเชิญท่านเป็นผู้ทรงคุณวุฒิตรวจสอบเครื่องมือเพื่อใช้ในการทำวิทยานิพนธ์ดังกล่าว

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์เป็นผู้ทรงคุณวุฒิตรวจสอบเครื่องมือวิจัยด้วย จักเป็นพระคุณยิ่ง

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

หัวหน้าภาควิชาวิจัยและจิตวิทยาการศึกษา

ภาควิชาวิจัยและจิตวิทยาการศึกษา

โทรศัพท์ ๐๘-๕๐๐๙-๑๗๐๗

โทรสาร ๐-๒๒๑๘-๒๕๗๘



ภาคผนวก ง

หนังสือ: ขอความร่วมมือในการเก็บข้อมูลทำวิทยานิพนธ์  
โรงเรียนราชวินิต

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ง

ที่ ศธ ๐๕๑๒.๖(๒๗๕๕)/



คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย  
ถนนพญาไท เขตปทุมวัน กรุงเทพฯ ๑๐๓๓๐

๒ กุมภาพันธ์ ๒๕๕๐

เรื่อง ขอความร่วมมือในการเก็บข้อมูลทำวิทยานิพนธ์

เรียน ผู้อำนวยการโรงเรียนราชวินิต

สิ่งที่ส่งมาด้วย แบบสอบถามวิจัยวิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน

ด้วย นางสาวพิมพ์สิริ เขียรนรเศรษฐ์ นิสิตชั้นปริญญาโท สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อยู่ในระหว่างดำเนินการทำวิทยานิพนธ์ เรื่อง “การเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยนคำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน” โดยมีผู้ช่วยศาสตราจารย์ ดร.เอมอร จังศิริพรปกรณ์ เป็นอาจารย์ที่ปรึกษา ซึ่งจะต้องมีการจัดทำคลังข้อสอบคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 ในกรณีนี้จึงมีความประสงค์จะขอความร่วมมือในการเก็บข้อมูล (จัดการสอบ) นักเรียนชั้นประถมศึกษาปีที่ 6 ทั้งระดับชั้นในโรงเรียนของท่านด้วยแบบสอบคนละ 1 ชุด จำนวน 40 ข้อ โดยขอความอนุเคราะห์ในการประสานงานกับอาจารย์ผู้สอนวิชาคณิตศาสตร์ในระดับชั้นประถมศึกษาปีที่ 6 เพื่อให้ นางสาวพิมพ์สิริ เขียรนรเศรษฐ์ จัดเก็บข้อมูลด้วยตนเอง

จึงเรียนมาเพื่อโปรดให้ความร่วมมือกับนิสิตในการเก็บข้อมูลทำวิทยานิพนธ์ด้วย ทั้งนี้เพื่อเป็นประโยชน์สำหรับนิสิตและประโยชน์ในเชิงวิชาการต่อไป จักขอขอบพระคุณยิ่ง

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

หัวหน้าภาควิชาวิจัยและจิตวิทยาการศึกษา

ภาควิชาวิจัยและจิตวิทยาการศึกษา

โทรศัพท์ ๐๘-๙๗๗๘๘๒๘๑

โทรสาร ๐-๒๒๑๘-๒๕๗๘



ภาคผนวก จ

หนังสือ: ขอความร่วมมือในการเก็บข้อมูลทำวิทยานิพนธ์  
โรงเรียนวัดปลับปลาชัย

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก จ

ที่ ศธ ๐๕๑๒.๖(๒๗๕๕)/



คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย  
ถนนพญาไท เขตปทุมวัน กรุงเทพฯ ๑๐๓๓๐

๒ กุมภาพันธ์ ๒๕๕๐

เรื่อง ขอความร่วมมือในการเก็บข้อมูลทำวิทยานิพนธ์

เรียน ผู้อำนวยการโรงเรียนวัดพลับพลาชัย

ด้วย นางสาวพิมพ์สิริ เขียรนรเศรษฐ์ นิสิตชั้นปริญญาโท สาขาวิชาการวัดและ  
ประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อยู่ใน  
ระหว่างดำเนินการทำวิทยานิพนธ์ เรื่อง “การเปรียบเทียบค่าความสามารถ ลักษณะของการเปลี่ยน  
คำตอบ และเวลาที่ใช้ในการทำข้อสอบของผู้สอบ ในการทดสอบแบบปรับเหมาะกับความสามารถของ  
ผู้สอบด้วยคอมพิวเตอร์ เมื่อเงื่อนไขการทดสอบและระดับความสามารถของผู้สอบแตกต่างกัน” โดยมี  
ผู้ช่วยศาสตราจารย์ ดร.เอมอร จังศิริพรภรณ์ เป็นอาจารย์ที่ปรึกษา ซึ่งจะดำเนินการจัดสอบนักเรียนใน  
วิชาคณิตศาสตร์ เรื่องความเข้าใจเศษส่วน ระดับชั้นประถมศึกษาปีที่ 6 ด้วยวิธีการทดสอบแบบปรับ  
เหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized adaptive testing - CAT) ซึ่งเป็น  
การสอบผ่านหน้าจอคอมพิวเตอร์

ในการนี้จึงมีความประสงค์จะขอความร่วมมือในการเก็บข้อมูล (จัดการสอบ) นักเรียนชั้น  
ประถมศึกษาปีที่ 6 ทั้งระดับชั้นในโรงเรียนของท่าน ด้วยการทดสอบติดต่อกัน 2 รอบ รอบละ 20 ข้อ รวม 1  
ชั่วโมงต่อห้อง ซึ่งทางผู้วิจัยเล็งเห็นว่าโรงเรียนของท่านมีศักยภาพในด้านระบบการเรียนการสอน  
คอมพิวเตอร์ ทางคณะจึงขอความอนุเคราะห์ในการประสานงานกับอาจารย์ผู้สอนวิชาคณิตศาสตร์ใน  
ระดับชั้นประถมศึกษาปีที่ 6 และศูนย์คอมพิวเตอร์ของทางโรงเรียน เพื่อให้ นางสาวพิมพ์สิริ เขียรนรเศรษฐ์  
จัดเก็บข้อมูลด้วยตนเอง

จึงเรียนมาเพื่อโปรดให้ความร่วมมือกับนิสิตในการเก็บข้อมูลทำวิทยานิพนธ์ด้วย ทั้งนี้เพื่อเป็น  
ประโยชน์สำหรับนิสิตและประโยชน์ในเชิงวิชาการต่อไป จักขอขอบพระคุณยิ่ง

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

หัวหน้าภาควิชาวิจัยและจิตวิทยาการศึกษา

ภาควิชาวิจัยและจิตวิทยาการศึกษา

โทรศัพท์ ๐๘-๙๗๗๘๘๒๘๑, ๐-๒๒๑๘-๒๕๗๘



ภาคผนวก จ

คู่มือสำหรับผู้สอบ

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ฉ

# คู่มือการใช้โปรแกรมสำหรับผู้สอบ

หน้าจอแรก : เบื้องหน้ากรอกข้อมูลผู้สอบ

โปรแกรมการทดสอบแบบปรับเหมาะ  
กับความสามารถของผู้สอบด้วยคอมพิวเตอร์

วิชา คณิตศาสตร์ เรื่อง ความเข้าใจเศษส่วน ชั้น ประถมศึกษาปีที่ 6

ข้อมูลผู้สอบ

ค.ช.  ค.ญ. ชื่อ  นามสกุล

ชั้น ป.6 ห้อง  เลขที่  โรงเรียน

เงื่อนไขการทดสอบ

- ❶ ระบุเพศ หรือพิมพ์ชื่อ นามสกุล ห้อง เลขที่ ของนักเรียน ลงในบริเวณสีขาว
- ❷ นักเรียนที่มีเลขที่ เป็นจำนวนคู่ (1,3,5,7,9,...) ให้กดปุ่ม "ไม่ให้ทวนคำตอบ"
- นักเรียนที่มีเลขที่ เป็นจำนวนคี่ (2,4,6,8,10,...) ให้กดปุ่ม "ให้ทวนคำตอบ"
- ❸ กดปุ่มเริ่มการทดสอบ (สีเขียว)

❖ หน้าจอไม่ให้ทวนคำตอบ : ข้อสอบ 20 ข้อ

1 29:53 โจทย์

1 24:48 โจทย์

กับคำตอบ

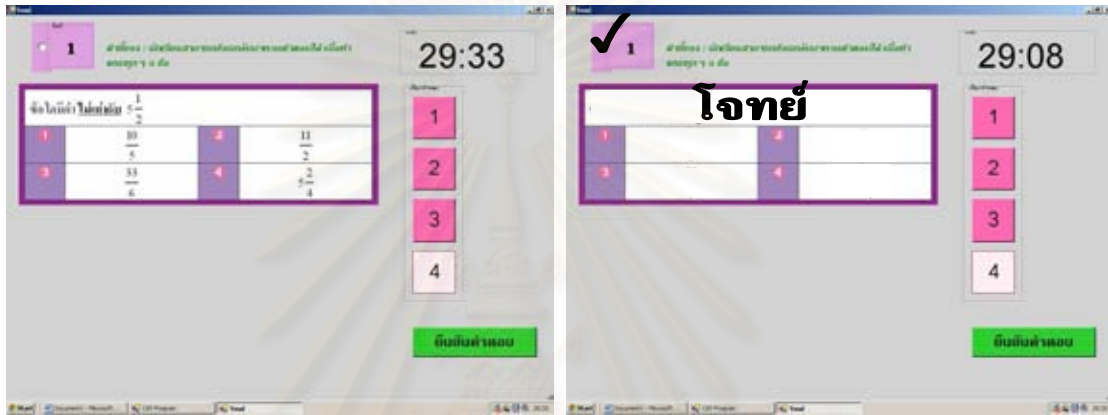
❶ กดปุ่มเลือกคำตอบ (สี่ขมพู)

❷ กดปุ่มยืนยันคำตอบ (สี่ขียว)

★ **หมายเหตุ:** ถ้าห้กเรียหกดปุ่มยืนยันคำตอบแล้ว จะไม่สามารถกลับมาแก้ไขคำตอบในโจทย์ข้อนี้ได้  
อีก

❸ ทำเซ่หห้จหครบ 20 ข้อ

◊ **หน้าจอให้ทวหคำตอบ :** ข้อสอบ 20 ข้อ (สามารถทวหคำตอบทุกจ 5 ข้อ)



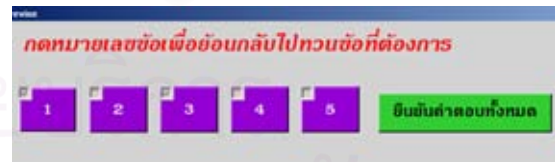
❶ กดปุ่มเลือกคำตอบ (สี่ขมพู)

❷ ถ้าห้กเรียหยังไม่แก้หใจในคำตอบ ห้กเรียหสามารถทำเครื่องขมายเพื่อเป็หสัญลักษณ์เตือนได้ โดยการกดช่องว่างสีขวหห้ข้อห้จ (เมื่อกดแล้วจะข้หเครื่องขมายถูกห้ข้อ)

❸ จากห้หกดปุ่มยืนยันคำตอบ (สี่ขียว)

❹ ทำเซ่หห้จหครบ 5 ข้อ

❺ เมื่อครบทุกจ 5 ข้อ จะข้หหน้าจอดังห้ →  
(ข้อที่มีเครื่องขมายถูกห้ข้อ แปลว่า  
เป็หข้อที่ห้กเรียหยังไม่แก้หใจในคำตอบ)



❻ ห้กเรียหสามารถกดขมายเลขข้อห้จเพื่อย้อนกลับไปทวหข้อที่ต้องการ (สามารถทวหได้ห้จ5ข้อ)  
หรือขมายข้อที่ห้กเรียหต้องการ (มีเครื่องขมายถูก) ก็ได้

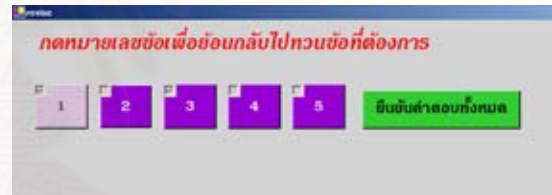
◆ หน้าจอเมื่อกดหมายเลขข้อเพื่อกลับไปทบทวน



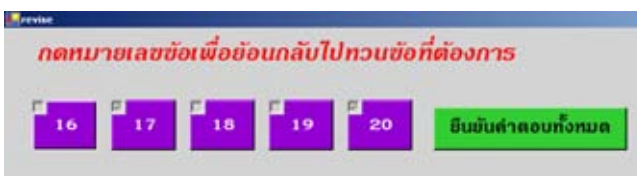
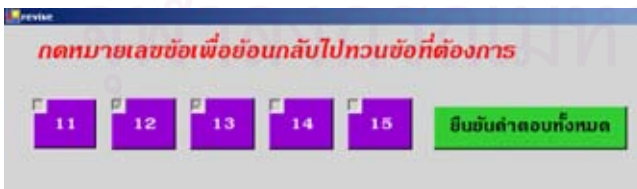
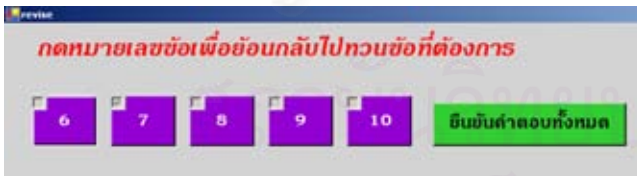
8 ✎ **ต้องการคำตอบเดิม** กดปุ่มยืนยันคำตอบ เปลี่ยนคำตอบใหม่ (สีเขียว)

✎ **ต้องการเปลี่ยนคำตอบ** กดปุ่มขอ (สีม่วง) จะขึ้นปุ่มกดคำตอบใหม่ให้ เมื่อคลิกเลือกคำตอบใหม่เสร็จแล้วให้กดปุ่มยืนยันคำตอบ (สีเขียว)

9 จะกลับไปยังหน้าจออื่นครั้ง (ข้อที่กลับไปทบทวนแล้วจะไม่สามารถกลับไปทบทวนได้อีก) จะคลิกสามารถเลือกทบทวนข้ออื่นๆ ได้อีก แต่ถ้าไม่ต้องการกลับไปทบทวนในช่วง 5 ข้อที่เหลือ ให้กดปุ่มยืนยันคำตอบทั้งหมด (สีเขียว) ซึ่งเมื่อกดปุ่มยืนยันคำตอบแล้ว จะไม่สามารถกลับมาทบทวนในช่วง 5 ข้อนี้ได้



10 จะทำซ้ำๆ เช่นนี้จนครบ 20 ข้อ (สามารถกลับไปทบทวนทุกๆ 5 ข้อ)





ภาคผนวก ซ

การเขียนโปรแกรมแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ซ

โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พัฒนาโดยใช้ภาษา VB.NET ซึ่งตัวโปรแกรมประกอบด้วย 4 หน้าต่างหลักดังนี้

1. **หน้าต่าการทำงานหลัก (Mainmenu.vb)** เมื่อโปรแกรมเริ่มต้นทำงาน หน้าต่านี้จะเป็นส่วนแรกท่แสดงประกอบด้วย ส่วนแสดงข้อมูลวิทยานิพนธ์ ส่วนรับข้อมูลผู้สอบ ส่วนเลือกประเภทการสอบ ประกอบด้วย 7 โปรแกรมย่อย ดังนี้

### 1.1 Sub Quit\_Click สำหรับออกจากระบบ

### 1.2 Sub Start\_Click สำหรับเริ่มทำการทดสอบ

```
Private Sub Start_Click(ByVal sender As System.Object, ByVal e As System.EventArgs) Handles Start.Click
```

```
Dim niDec As New System.Globalization.NumberFormatInfo
```

```
niDec.NumberDecimalDigits = 3
```

```
If validatedata() Then
```

```
Dim question1 As Form = New question(Me)
```

```
Dim question2 As Form = New question(Me)
```

```
Call moderevise_CheckedChanged(Nothing, Nothing)
```

```
If v_moderevise Then
```

```
Call cleararray()
```

```
question2.ShowDialog()
```

```
Call exporttofile(True)
```

```
MessageBox.Show("ค่าความสามารถของผู้สอบแบบทวน =" & ability.ToString("f", niDec))
```

```
MessageBox.Show("กดปุ่ม OK เพื่อเริ่มการทดสอบแบบไม่ทวน")
```

```
Call cleararray()
```

```
v_moderevise = False
```

```
question1.ShowDialog()
```

```
Call exporttofile(False)
```

```
MessageBox.Show("ค่าความสามารถของผู้สอบแบบไม่ทวน =" & ability.ToString("f", niDec))
```

```
Else
```

```
Call cleararray()
```

```
question1.ShowDialog()
```

```
Call exporttofile(False)
```

```
MessageBox.Show("ค่าความสามารถของผู้สอบแบบไม่ทวน =" & ability.ToString("f", niDec))
```

```
MessageBox.Show("กดปุ่ม OK เพื่อเริ่มการทดสอบแบบทวน")
```

```

v_moderevise = True
Call cleararray()
question2.ShowDialog()
Call exporttofile(True)
MessageBox.Show("ค่าความสามารถของผู้สอบแบบทวน =" & ability.ToString("f", niDec))
End If
question1.Dispose()
question2.Dispose()
Array.Clear(d_chosen, 0, 225)
End If
End Sub

```

1.3 Function **validatedata** สำหรับตรวจสอบข้อมูลผู้สอบว่าครบตามเงื่อนไขที่ต้องการหรือไม่

1.4 Sub **moderevise\_CheckedChanged** สำหรับกำหนดเงื่อนไขในการสอบ

1.5 Sub **cleararray** สำหรับล้างข้อมูลผลการสอบในระบบ

1.6 Sub **exporttofile** สำหรับบันทึกข้อมูลผลการสอบลงไฟล์

```

Private Sub exporttofile(ByVal mode As Boolean)
Dim path As String
Dim a As Integer = 1
Dim correctno As Integer
Dim wrongno As Integer
Dim wantrevise As Integer
Dim t_f As Integer
Dim f_t As Integer
Dim f_f As Integer
Dim revisecheck As Integer
Dim revisedontcheck As Integer
Dim changeanswer As Integer
Dim niDec As New System.Globalization.NumberFormatInfo
niDec.NumberDecimalDigits = 3
If mode Then
path = "C:\CAT\Result" & TextBox2.Text & "_" & TextBox1.Text & "_" & TxtName.Text & "_ทวน.txt"

```



```

Else
    path = "C:\CAT\Result" & TextBox2.Text & "_" & TextBox1.Text & "_" & TxtName.Text & "_ไม่ทวน.txt"
End If

Dim sw As StreamWriter
Dim dir As Directory

Dim I As Integer
Dim pass As Boolean = False
If Not dir.Exists("C:\CAT\Result") Then
    dir.CreateDirectory("C:\CAT\Result")
End If
If File.Exists(path) = False Then
    sw = File.CreateText(path)
Else
    sw = File.AppendText(path)
End If
If RadioButton1.Checked Then
    sw.Write("ต.ช. ")
Else
    sw.Write("ต.ญ. ")
End If
sw.WriteLine(TxtName.Text & " " & TxtSname.Text)
If mode Then
    sw.WriteLine("-----")
    sw.WriteLine("      ทวน      ")
    sw.WriteLine("-----")
    For I = 1 To 20
        sw.WriteLine("โจทย์ข้อที่ " & I & " ได้ข้อคั้ง " & u_getquestion(I))
        sw.WriteLine("ก่อนทวน      หลังทวน")
        If d_answerkey(u_getquestion(I)) = u_firstanswer(I) Then
            sw.Write("ตอบถูก      ")
        Else
            sw.Write("ตอบผิด      ")
        End If
    Next I
End If

```

```

End If
If u_secondanswer(l) > 0 Then
    If d_answerkey(u_getquestion(l)) = u_secondanswer(l) Then
        sw.WriteLine("ตอบถูก")
    Else
        sw.WriteLine("ตอบผิด")
    End If
Else
    If d_answerkey(u_getquestion(l)) = u_firstanswer(l) Then
        sw.WriteLine("ตอบถูก")
    Else
        sw.WriteLine("ตอบผิด")
    End If
End If
If u_secondanswer(l) > 0 And u_secondanswer(l) <> u_firstanswer(l) Then
    changeanswer = changeanswer + 1
    If u_secondanswer(l) <> d_answerkey(u_getquestion(l)) And u_firstanswer(l) <>
d_answerkey(u_getquestion(l)) Then
        f_f = f_f + 1
    End If
    If u_secondanswer(l) = d_answerkey(u_getquestion(l)) And u_firstanswer(l) <>
d_answerkey(u_getquestion(l)) Then
        f_t = f_t + 1
    End If
    If u_secondanswer(l) <> d_answerkey(u_getquestion(l)) And u_firstanswer(l) =
d_answerkey(u_getquestion(l)) Then
        t_f = t_f + 1
    End If
End If
If u_wantrevise(l) Then
    wantrevise = wantrevise + 1
End If
If u_viewrevise(l) Then
    If u_wantrevise(l) Then
        revisecheck = revisecheck + 1

```

```

Else
    revisedontcheck = revisedontcheck + 1
End If

End If

sw.WriteLine("Xeta เก่า = " & u_ability_old(l).ToString("f", niDec))
sw.WriteLine("          Xeta ใหม่ = " & u_ability(l).ToString("f", niDec))
sw.WriteLine("SEE เก่า = " & u_variant_old(l).ToString("f", niDec))
sw.WriteLine("          SEE ใหม่ = " & u_variant(l).ToString("f", niDec))
sw.WriteLine("")

Next

sw.WriteLine("มีการเปลี่ยนคำตอบจากผิดเป็นถูก " & f_t & " ข้อ")
sw.WriteLine("มีการเปลี่ยนคำตอบจากผิดเป็นผิด " & f_f & " ข้อ")
sw.WriteLine("มีการเปลี่ยนคำตอบจากถูกเป็นผิด " & t_f & " ข้อ")
sw.WriteLine("มีการเปลี่ยนคำตอบรวม " & changeanswer & " ข้อ")
sw.WriteLine("ทำเครื่องหมายเพื่อกลับไปทวน " & wantrevise & " ข้อ")
sw.WriteLine("กลับไปทวนข้อที่ทำเครื่องหมาย " & revisecheck & " ข้อ")
sw.WriteLine("กลับไปทวนข้อที่ไม่ได้ทำเครื่องหมาย " & revisedontcheck & " ข้อ")
sw.WriteLine("เวลาในการทดสอบ " & Int(u_testtime / 60) & " นาที " & u_testtime Mod 60 & " วินาที")

Else
sw.WriteLine("-----")
sw.WriteLine("          ไม่ทวน          ")
sw.WriteLine("-----")

For I = 1 To 20
    sw.WriteLine("โจทย์ข้อที่ " & I & " ได้ข้อคลัง " & u_getquestion(I))
    If d_answerkey(u_getquestion(I)) = u_firstanswer(I) Then
        sw.WriteLine("ตอบถูก")
        correctno = correctno + 1
    Else
        sw.WriteLine("ตอบผิด")
        wrongno = wrongno + 1
    End If

sw.WriteLine("Xeta = " & u_ability(I).ToString("f", niDec))
sw.WriteLine("SEE = " & u_variant(I).ToString("f", niDec))

```

```

        sw.WriteLine("")
    Next
    sw.WriteLine("ตอบถูก " & correctno & " ข้อ ตอบผิด " & wrongno & " ข้อ")
    sw.WriteLine("เวลาในการทดสอบ " & Int(u_testtime / 60) & " นาที " & u_testtime Mod 60 & " วินาที")
End If
sw.Flush()
sw.Close()

End Sub

```

### 1.7 Sub Button1\_Click สำหรับเปิดหน้าต่างแสดงข้อมูลวิทยานิพนธ์

2. หน้าต่างโจทย์ (Question.vb) เป็นหน้าต่างในการคัดเลือกโจทย์ที่ปรับเหมาะกับความสามารถของผู้สอบจากคลังข้อสอบ และทำการตรวจเช็คคำตอบค่าความค่าความสามารถและค่าความแปรหลังผู้สอบเลือกคำตอบ ประกอบด้วย 12 โปรแกรมย่อย ดังนี้

2.1 Sub question\_Load สำหรับคัดลอกโจทย์จากคลังเข้าสู่โปรแกรม, คัดเลือกโจทย์ข้อแรกและตั้งค่าความสามารถและค่าความแปรปรวนเริ่มต้น

2.2 Sub Timer1\_Tick สำหรับคำนวณเวลาที่เหลือในการทดสอบ และออกจากการทดสอบเมื่อครบตามเวลาที่กำหนด

```

Private Sub Timer1_Tick(ByVal sender As Object, ByVal e As System.EventArgs) Handles
Timer1.Tick
    q_testtime = q_testtime - 1
    If q_testtime Mod 60 < 10 Then
        Label11.Text = Int(q_testtime / 60) & ":0" & q_testtime Mod 60
    Else
        Label11.Text = Int(q_testtime / 60) & ":" & q_testtime Mod 60
    End If
    If (q_testtime Mod 60) = 0 Then
        If checkterminate(0, SEE, q_testtime) Then
            MyBase.Close()
        End If
    End If
End If

```

End Sub

2.3 Sub Button1\_Click สำหรับกดยืนยันคำตอบเพื่อทำข้อถัดไป

2.4 Sub question\_KeyPress สำหรับรับค่าตัวเลขทางคีย์บอร์ด ซึ่งรับได้ 4 ค่า คือ 1,2,3 และ 4

2.5 Sub RadioButton\_CheckedChanged สำหรับรับค่าตัวเลือกทางเมาส์

2.6 Sub question\_Closed สำหรับล้างค่าลำดับข้อของข้อสอบ และเงื่อนไขการสอบ

2.7 Function checkanswer สำหรับเปรียบเทียบคำตอบของผู้สอบว่าตรงกับเฉลยหรือไม่

2.8 Sub BAY สำหรับคำนวณค่าความสามารถและค่าความแปรปรวนข้อผู้สอบหลังตอบคำถาม

Private Sub BAY(ByVal InputQ As Integer)

Dim D, A, t, tD, ord, area, areaD As Double

Call para(InputQ)

$D = (d\_bvalue(InputQ) - ability) / ((1 / (d\_avalue(InputQ) ^ 2) + VAR0) ^ (0.5))$

$t = 1 / (1 + (P1 * D))$

$tD = 1 / (1 - (P1 * D))$

$ord = 0.3989 / (exp ^ (0.5 * D * D))$

$area = 1 - (ord * ((C1 * t) + (C2 * t * t) + (C3 * t * t * t) + (C4 * t * t * t * t) + (C5 * t * t * t * t * t)))$

$areaD = 1 - (ord * ((C1 * tD) + (C2 * tD * tD) + (C3 * tD * tD * tD) + (C4 * tD * tD * tD * tD) + (C5 * tD * tD * tD * tD * tD)))$

$A = d\_cvalue(InputQ) + (1 - d\_cvalue(InputQ)) * areaD$

If checkanswer(InputQ) Then

$ABIT0 = ABIT0 + (((1 - d\_cvalue(InputQ)) * VAR0 * ord) / (((1 / (d\_avalue(InputQ) * d\_avalue(InputQ))) + VAR0) ^ 0.5) * (d\_cvalue(InputQ) + (1 - d\_cvalue(InputQ)) * areaD))$

$VAR0 = VAR0 * (1 - (1 - d\_cvalue(InputQ)) / (1 + (1 / (d\_avalue(InputQ) * d\_avalue(InputQ) * VAR0)))) * (ord / A) * (((1 - d\_cvalue(InputQ)) * ord) / A) - D)$

ability = ABIT0

u\_ability(c\_running) = ABIT0

u\_variant(c\_running) = VAR0 ^ 0.5

```

u_ability_old(c_running) = ABIT0
u_variant_old(c_running) = VAR0 ^ 0.5
Else
ABIT0 = ABIT0 - (VAR0 / (((1 / (d_avalu(InputQ) * d_avalu(InputQ))) + VAR0) ^ 0.5)) * (ord / area)
VAR0 = VAR0 * (1 - (ord / (1 + (1 / (d_avalu(InputQ) * d_avalu(InputQ) * VAR0)))) * ((ord / area) +
D) / (area))
ability = ABIT0
u_ability(c_running) = ABIT0
u_variant(c_running) = VAR0 ^ 0.5
u_ability_old(c_running) = ABIT0
u_variant_old(c_running) = VAR0 ^ 0.5
End If
End Sub

```

### 2.9 Sub para สำหรับเตรียมค่าพารามิเตอร์เพื่อใช้ในการคำนวณ Bay

```

Private Sub para(ByVal InputQ As Integer)
Dim P, Q, IIF, TIF As Double
P = d_cvalue(InputQ) + ((1 - d_cvalue(InputQ)) * (exp ^ (1.7 * d_avalu(InputQ) * (ability -
d_bvalue(InputQ)))) / _
(1 + (exp ^ (1.7 * d_avalu(InputQ) * (ability - d_bvalue(InputQ))))))
Q = 1 - P
IIF = ((1.7 * d_avalu(InputQ) * Q * (P - d_cvalue(InputQ)) / (1 - d_cvalue(InputQ))) ^ 2) / (P * Q)
SEE = VAR0 ^ 0.5
End Sub

```

### 2.10 Function calselectnext สำหรับคัดเลือกโจทย์ที่เหมาะสมกับความสามารถของผู้สอบ

จากคลังข้อสอบ

```

Private Function calselectnext(ByVal cal As Integer) As Integer
Dim i As Integer
Dim Ps, Qs, Iif As Double
Dim choosing As Integer
Dim lmax As Double = -50
For i = 1 To 224
Ps = d_cvalue(i) + ((1 - d_cvalue(i)) * (exp ^ (1.7 * d_avalu(i) * (ability - d_bvalue(i)))) / _
(1 + (exp ^ (1.7 * d_avalu(i) * (ability - d_bvalue(i))))))
Qs = 1 - Ps

```

```

lif = (((1.7 * d_avalve(i) * Qs * (Ps - d_cvalue(i))) / (1 - d_cvalue(i))) ^ 2) / (Ps * Qs)
If (lmax < lif And Not d_chosen(i)) Then
    lmax = lif
    choosing = i
End If
Next
d_chosen(choosing) = True
'Go to next question
calselectnext = choosing

End Function

```

## 2.11 Function checkterminate สำหรับตรวจสอบเงื่อนไขการจบการทดสอบ

## 2.12 Sub displayquestion สำหรับแสดงโจทย์

```

Private Sub displayquestion(ByVal question2 As Integer, ByVal currentQ As Integer)
    Dim executing_assembly As System.Reflection.Assembly = _
        Me.GetType.Assembly.GetEntryAssembly()
    ' Get our namespace.
    Dim my_namespace As String = _
        executing_assembly.GetName().Name.ToString()
    Dim picture_stream As Stream
    Dim PicQuestion As Bitmap
    Dim SD As Double
    Dim picname As String
    Dim niDec As New System.Globalization.NumberFormatInfo
    niDec.NumberDecimalDigits = 4
    CheckBox1.Checked = False
    If question2 < 10 Then
        picname = "00" & question2
    Else
        If question2 < 100 Then
            picname = "0" & question2
        Else
            picname = question2
        End If
    End If

```

```

End If
If v_moderevise Then
    Label1.Visible = False
    Label3.Visible = True
    CheckBox1.Visible = True
Else
    Label1.Visible = True
    Label3.Visible = False
    CheckBox1.Visible = False
End If
picture_stream = executing_assembly.GetManifestResourceStream(my_namespace _
    + "." + picname + ".JPG")
If Not (picture_stream Is Nothing) Then
    PicQuestion = New Bitmap(picture_stream)
    PictureBox1.Image = PicQuestion
    picture_stream.Close()
End If
SD = VAR0 ^ 0.5
Label15.Text = ABIT0.ToString("f", niDec)
Label8.Text = d_cvalue(u_getquestion(c_running))
Label5.Text = d_avalue(u_getquestion(c_running))
Label6.Text = d_bvalue(u_getquestion(c_running))
Label10.Text = u_getquestion(c_running)
Label14.Text = SD.ToString("f", niDec)
Label2.Text = c_running

```

End Sub

**3. หน้าต่างเลือกข้อคำถามในกรณีการทวน (Revise.vb)** เป็นหน้าต่างที่จะแสดงเมื่อผู้สอบทำโจทย์ครบทุกๆ 5 ข้อ เพื่อเปิดโอกาสให้ผู้สอบได้กลับไปทวนคำตอบของ 5 ข้อที่ผ่านมา ประกอบด้วย 7 โปรแกรมย่อย ดังนี้

3.1 Sub Button6\_Click เพื่อยืนยันคำตอบทั้งหมดและทำข้อสอบช่วงถัดไป

3.2 Sub Gotorevise\_Click ตรวจสอบเช็คว่าคุณสอบมีการเปลี่ยนคำตอบหรือไม่



3.3 Sub revise\_Closed คำนวณค่าความสามารถและค่าความแปรปรวนใหม่หลังผู้สอบทำการทวน

3.4 Function checkanswer สำหรับเปรียบเทียบคำตอบของผู้สอบหลังทวนว่าตรงกับเฉลยหรือไม่

3.5 Sub BAY สำหรับคำนวณค่าความสามารถและค่าความแปรปรวนข้อผู้สอบหลังตอบคำถาม

3.6 Sub para สำหรับเตรียมค่าพารามิเตอร์เพื่อใช้ในการคำนวณ Bay

4. หน้าต่างโจทย์สำหรับการทวน (Questionrevise.vb) เป็นหน้าต่างแสดงโจทย์และคำตอบที่ผู้สอบได้ตอบไปแล้วเพื่อทำการทวน ประกอบด้วย 6 โปรแกรมย่อยคือ

4.1 Sub questionrevise\_Load สำหรับตั้งค่าเริ่มต้นเพื่อการทวนคำตอบ

4.2 Sub displayquestion สำหรับแสดงโจทย์ข้อที่ต้องการทวน

4.3 Sub CheckBox1\_CheckedChanged สำหรับระบุว่าต้องการเปลี่ยนคำตอบ

4.4 Sub RadioButton2\_CheckedChanged สำหรับรับค่าคำตอบใหม่หลังการทวน

4.5 Sub RadioButton\_CheckedChanged สำหรับแสดงคำตอบที่ผู้สอบเลือกไว้ก่อนหน้าการทวน

4.6 Sub Button55\_Click สำหรับยืนยันการทวนคำตอบเพื่อทำการทวนคำตอบข้อถัดไป



ภาคผนวก ซ

Output การวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข

BILOG-MG V3.0

REV 19990104.1300

THIS SEPTEMBER 2004 EDITION OF BILOG-MG VERSION 3.0 IS FOR ACADEMIC  
RESEARCH

AND EDUCATIONAL USE ONLY. THE PROGRAM EXPIRES ON 12/31/2007.

BILOG-MG ITEM MAINTENANCE PROGRAM: LOGISTIC ITEM RESPONSE MODEL

DISTRIBUTED BY

SCIENTIFIC SOFTWARE INTERNATIONAL, INC.  
7383 N. LINCOLN AVENUE, SUITE 100  
CHICAGO, IL 60646  
(800) 247-6113  
(847) 675-0720  
WWW: <http://www.ssicentral.com>

PROGRAM COPYRIGHT HELD BY SCIENTIFIC SOFTWARE INTERNATIONAL, INC. 2002

DISTRIBUTION OR USE UNAUTHORIZED BY SSI, INC. IS PROHIBITED

1 \*\*\* BILOG-MG ITEM MAINTENANCE PROGRAM \*\*\*

\*\*\* PHASE 1 \*\*\*

TEST  
>COMMENT  
By nok

>GLOBAL DFName = 'C:\Documents and  
Settings\Administrator\Desktop\cat.dat',

NParm = 3;

FILE ASSIGNMENT AND DISPOSITION  
=====

SUBJECT DATA INPUT FILE	C:\DOCUMENTS AND SETTINGS\ADMINISTRATOR\DESKTOP\ CAT.DAT	
BILOG-MG MASTER DATA FILE	MF.DAT	WILL BE CREATED FROM DATA FILE
CALIBRATION DATA FILE	CF.DAT	WILL BE CREATED FROM DATA FILE
ITEM PARAMETERS FILE	IF.DAT	WILL BE CREATED THIS RUN
CASE SCALE-SCORE FILE	SF.DAT	
CASE WEIGHTING		NONE EMPLOYED
ITEM RESPONSE MODEL		3 PARAMETER LOGISTIC

NORMAL METRIC (I.E., D = 1.7)

&gt;LENGTH NITems = (40);

TEST LENGTH SPECIFICATIONS  
=====

MAIN TEST LENGTHS: 40

&gt;INPUT NTOtal = 40,

NALt = 1000,

NIDchar = 3;

DATA INPUT SPECIFICATIONS  
=====

NUMBER OF FORMAT LINES	1
NUMBER OF ITEMS IN INPUT STREAM	40
NUMBER OF RESPONSE ALTERNATIVES	1000
NUMBER OF SUBJECT ID CHARACTERS	3
NUMBER OF GROUPS	1
NUMBER OF TEST FORMS	1
TYPE OF DATA	SINGLE-SUBJECT DATA, NO CASE

WEIGHTS

MAXIMUM SAMPLE SIZE FOR ITEM CALIBRATION 10000000

ALL SUBJECTS INCLUDED IN RUN

&gt;ITEMS ;

TEST SPECIFICATIONS  
=====

&gt;TEST1 TName = 'TEST0001',

INumber = (1(1)40);

TEST NUMBER: 1 TEST NAME: TEST0001

NUMBER OF ITEMS: 40

ITEM NUMBER	ITEM NAME	ITEM NUMBER	ITEM NAME	ITEM NUMBER	ITEM NAME	ITEM NUMBER	ITEM NAME
1	ITEM0001	14	ITEM0014	27	ITEM0027	40	ITEM0040
2	ITEM0002	15	ITEM0015	28	ITEM0028		
3	ITEM0003	16	ITEM0016	29	ITEM0029		
4	ITEM0004	17	ITEM0017	30	ITEM0030		
5	ITEM0005	18	ITEM0018	31	ITEM0031		
6	ITEM0006	19	ITEM0019	32	ITEM0032		
7	ITEM0007	20	ITEM0020	33	ITEM0033		
8	ITEM0008	21	ITEM0021	34	ITEM0034		
9	ITEM0009	22	ITEM0022	35	ITEM0035		
10	ITEM0010	23	ITEM0023	36	ITEM0036		
11	ITEM0011	24	ITEM0024	37	ITEM0037		
12	ITEM0012	25	ITEM0025	38	ITEM0038		
13	ITEM0013	26	ITEM0026	39	ITEM0039		



ITEM	11	12	13	14	15	16	17	18	19	20
TRIED	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RIGHT	1.0	1.0	0.0	0.0	0.0	1.0	0.0	1.0	1.0	0.0

ITEM	21	22	23	24	25	26	27	28	29	30
TRIED	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RIGHT	1.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0

ITEM	31	32	33	34	35	36	37	38	39	40
TRIED	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RIGHT	1.0	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	1.0

400 OBSERVATIONS READ FROM FILE: C:\DOCUMENTS AND SETTINGS\  
ADMINISTRATOR\DESKTOP\CAT.DAT  
400 OBSERVATIONS WRITTEN TO FILE: MF.DAT

## ITEM STATISTICS FOR SUBTEST TEST0001

CORRELATION		ITEM*TEST					
ITEM	NAME	#TRIED	#RIGHT	PCT	LOGIT/1.7	PEARSON	BISERIAL
1	ITEM0001	400.0	46.0	11.5	1.20	0.237	0.389
2	ITEM0002	400.0	179.0	44.8	0.12	0.328	0.412
3	ITEM0003	400.0	274.0	68.5	-0.46	0.001	0.002
4	ITEM0004	400.0	268.0	67.0	-0.42	0.229	0.297
5	ITEM0005	400.0	325.0	81.2	-0.86	0.114	0.165
6	ITEM0006	400.0	352.0	88.0	-1.17	0.205	0.333
7	ITEM0007	400.0	334.0	83.5	-0.95	0.466	0.697
8	ITEM0008	400.0	344.0	86.0	-1.07	0.417	0.650
9	ITEM0009	400.0	92.0	23.0	0.71	0.330	0.457
10	ITEM0010	400.0	335.0	83.8	-0.96	0.243	0.364
11	ITEM0011	400.0	279.0	69.8	-0.49	0.441	0.580
12	ITEM0012	400.0	293.0	73.2	-0.59	0.503	0.676
13	ITEM0013	400.0	291.0	72.8	-0.58	0.441	0.591
14	ITEM0014	400.0	176.0	44.0	0.14	0.433	0.545
15	ITEM0015	400.0	211.0	52.7	-0.06	0.343	0.430
16	ITEM0016	400.0	294.0	73.5	-0.60	0.426	0.574
17	ITEM0017	400.0	118.0	29.5	0.51	0.501	0.663
18	ITEM0018	400.0	317.0	79.2	-0.79	0.417	0.590
19	ITEM0019	400.0	323.0	80.8	-0.84	0.381	0.550
20	ITEM0020	400.0	183.0	45.8	0.10	0.236	0.296
21	ITEM0021	400.0	325.0	81.2	-0.86	0.459	0.666
22	ITEM0022	400.0	320.0	80.0	-0.82	0.415	0.594
23	ITEM0023	400.0	314.0	78.5	-0.76	0.391	0.550
24	ITEM0024	400.0	300.0	75.0	-0.65	0.470	0.641
25	ITEM0025	400.0	179.0	44.8	0.12	0.339	0.426
26	ITEM0026	400.0	192.0	48.0	0.05	0.392	0.492
27	ITEM0027	400.0	263.0	65.8	-0.38	0.523	0.676
28	ITEM0028	400.0	216.0	54.0	-0.09	0.566	0.710
29	ITEM0029	400.0	296.0	74.0	-0.62	0.435	0.589
30	ITEM0030	400.0	193.0	48.2	0.04	0.507	0.636
31	ITEM0031	400.0	363.0	90.8	-1.34	0.278	0.487
32	ITEM0032	400.0	312.0	78.0	-0.74	0.315	0.441
33	ITEM0033	400.0	277.0	69.2	-0.48	0.396	0.519

34	ITEM0034	400.0	299.0	74.8	-0.64	0.421	0.572
35	ITEM0035	400.0	208.0	52.0	-0.05	0.493	0.618
36	ITEM0036	400.0	88.0	22.0	0.74	0.282	0.394
37	ITEM0037	400.0	103.0	25.8	0.62	0.467	0.633
38	ITEM0038	400.0	222.0	55.5	-0.13	0.414	0.520
39	ITEM0039	400.0	161.0	40.2	0.23	0.139	0.176
40	ITEM0040	400.0	282.0	70.5	-0.51	0.478	0.632

---

536 BYTES OF NUMERICAL WORKSPACE USED OF 8192000 AVAILABLE IN PHASE-1  
8176 BYTES OF CHARACTER WORKSPACE USED OF 2048000 AVAILABLE IN PHASE-1  
1

BILOG-MG V3.0  
REV 19990329.1300

BILOG-MG ITEM MAINTENANCE PROGRAM: LOGISTIC ITEM RESPONSE MODEL

THIS SEPTEMBER 2004 EDITION OF BILOG-MG VERSION 3.0 IS FOR ACADEMIC  
RESEARCH AND EDUCATIONAL USE ONLY. THE PROGRAM EXPIRES ON 12/31/2004.

\*\*\* BILOG-MG ITEM MAINTENANCE PROGRAM \*\*\*

\*\*\* PHASE 2 \*\*\*

TEST

>CALIB ACCel = 1.0000,

TPRior,

GPRior,

FLOat;

CALIBRATION PARAMETERS  
=====

MAXIMUM NUMBER OF EM CYCLES: 20

MAXIMUM NUMBER OF NEWTON CYCLES: 2

CONVERGENCE CRITERION: 0.0100

ACCELERATION CONSTANT: 1.0000

LATENT DISTRIBUTION: NORMAL PRIOR FOR EACH GROUP

PLOT EMPIRICAL VS. FITTED ICC'S: NO

DATA HANDLING: DATA ON SCRATCH FILE

CONSTRAINT DISTRIBUTION ON ASYMPTOTES: YES

CONSTRAINT DISTRIBUTION ON SLOPES: YES

CONSTRAINT DISTRIBUTION ON THRESHOLDS: YES

SOURCE OF ITEM CONSTRAINT DISTRIBUTION

MEANS AND STANDARD DEVIATIONS: PROGRAM DEFAULTS

ITEM CONSTRAINTS IF PRESENT  
WILL BE UPDATED EACH CYCLE

1

\*\*\*\*\*

## CALIBRATION OF MAINTEST

TEST0001

\*\*\*\*\*

METHOD OF SOLUTION:

EM CYCLES (MAXIMUM OF 20)  
FOLLOWED BY NEWTON-RAPHSON STEPS (MAXIMUM OF 2)

QUADRATURE POINTS AND PRIOR WEIGHTS:

	1	2	3	4	5
POINT	-0.4000E+01	-0.3429E+01	-0.2857E+01	-0.2286E+01	-0.1714E+01
WEIGHT	0.7648E-04	0.6387E-03	0.3848E-02	0.1673E-01	0.5245E-01
	6	7	8	9	10
POINT	-0.1143E+01	-0.5714E+00	-0.8882E-15	0.5714E+00	0.1143E+01
WEIGHT	0.1186E+00	0.1936E+00	0.2280E+00	0.1936E+00	0.1186E+00
	11	12	13	14	15
POINT	0.1714E+01	0.2286E+01	0.2857E+01	0.3429E+01	0.4000E+01
WEIGHT	0.5245E-01	0.1673E-01	0.3848E-02	0.6387E-03	0.7648E-04

CONSTRAINT DISTRIBUTIONS ON ITEM PARAMETERS  
(THRESHOLDS, NORMAL; SLOPES, LOG-NORMAL; GUESSING, BETA)

ITEM	THRESHOLDS		SLOPES		ASYMPTOTES	
	MU	SIGMA	MU	SIGMA	ALPHA	BETA
ITEM0001	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0002	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0003	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0004	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0005	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0006	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0007	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0008	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0009	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0010	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0011	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0012	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0013	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0014	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0015	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0016	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0017	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0018	0.000	2.000	1.000	1.649	1.02	20.98



ITEM0019	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0020	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0021	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0022	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0023	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0024	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0025	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0026	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0027	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0028	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0029	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0030	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0031	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0032	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0033	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0034	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0035	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0036	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0037	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0038	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0039	0.000	2.000	1.000	1.649	1.02	20.98
ITEM0040	0.000	2.000	1.000	1.649	1.02	20.98

-----

[E-M CYCLES]

-2 LOG LIKELIHOOD = 16041.111

CYCLE 1; LARGEST CHANGE= 0.70302

-2 LOG LIKELIHOOD = 15965.228

UPDATED PRIOR ON ASYMPOTOTES; ALPHA & BETA = 1.29324 20.70676  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.40288 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.56277 2.00000

CYCLE 2; LARGEST CHANGE= 0.29770

-2 LOG LIKELIHOOD = 15948.305

UPDATED PRIOR ON ASYMPOTOTES; ALPHA & BETA = 1.28213 20.71787  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.37838 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.61505 2.00000

CYCLE 3; LARGEST CHANGE= 0.62597

-2 LOG LIKELIHOOD = 15926.544

UPDATED PRIOR ON ASYMPOTOTES; ALPHA & BETA = 1.50271 20.49729  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.33683 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.59129 2.00000

CYCLE 4; LARGEST CHANGE= 0.25496

-2 LOG LIKELIHOOD = 15911.646

UPDATED PRIOR ON ASYMPOTOTES; ALPHA & BETA =	1.70225	20.29775
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.28716	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.57280	2.00000

CYCLE 5; LARGEST CHANGE= 0.15942

-2 LOG LIKELIHOOD = 15901.447

UPDATED PRIOR ON ASYMPOTOTES; ALPHA & BETA =	2.01358	19.98642
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.25051	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.52813	2.00000

CYCLE 6; LARGEST CHANGE= 0.10471

-2 LOG LIKELIHOOD = 15893.800

UPDATED PRIOR ON ASYMPOTOTES; ALPHA & BETA =	2.41882	19.58118
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.21565	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.48241	2.00000

CYCLE 7; LARGEST CHANGE= 0.23504

-2 LOG LIKELIHOOD = 15895.531

UPDATED PRIOR ON ASYMPOTOTES; ALPHA & BETA =	2.83015	19.16985
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.15818	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.40262	2.00000

CYCLE 8; LARGEST CHANGE= 0.27915

-2 LOG LIKELIHOOD = 15880.979

UPDATED PRIOR ON ASYMPOTOTES; ALPHA & BETA =	3.25456	18.74544
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.13331	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.36590	2.00000

CYCLE 9; LARGEST CHANGE= 0.29172

-2 LOG LIKELIHOOD = 15878.475

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA = 3.62450 18.37550  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.09749 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.30015 2.00000

CYCLE 10; LARGEST CHANGE= 0.11317

-2 LOG LIKELIHOOD = 15872.466

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA = 3.93924 18.06076  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.07080 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.25871 2.00000

CYCLE 11; LARGEST CHANGE= 0.15645

-2 LOG LIKELIHOOD = 15868.029

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA = 4.17922 17.82078  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.05167 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.23766 2.00000

CYCLE 12; LARGEST CHANGE= 0.26263

-2 LOG LIKELIHOOD = 15873.317

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA = 4.35503 17.64497  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.04170 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.21708 2.00000

CYCLE 13; LARGEST CHANGE= 0.30284

-2 LOG LIKELIHOOD = 15865.651

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA = 4.48865 17.51135  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.02687 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.20365 2.00000

CYCLE 14; LARGEST CHANGE= 0.21704

-2 LOG LIKELIHOOD = 15864.413

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA = 4.59309 17.40691  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.02432 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.20486 2.00000

CYCLE 15; LARGEST CHANGE= 0.27738

-2 LOG LIKELIHOOD = 15866.173

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA =	4.66278	17.33722
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.02457	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.19477	2.00000

CYCLE 16; LARGEST CHANGE= 0.15184

-2 LOG LIKELIHOOD = 15863.684

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA =	4.71484	17.28516
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.02214	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.19868	2.00000

CYCLE 17; LARGEST CHANGE= 0.16226

-2 LOG LIKELIHOOD = 15863.556

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA =	4.75138	17.24862
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.01994	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.19905	2.00000

CYCLE 18; LARGEST CHANGE= 0.22737

-2 LOG LIKELIHOOD = 15863.865

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA =	4.77665	17.22335
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.02125	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.20005	2.00000

CYCLE 19; LARGEST CHANGE= 0.11593

-2 LOG LIKELIHOOD = 15862.775

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA =	4.79561	17.20439
UPDATED PRIOR ON LOG SLOPES; MEAN & SD =	-0.01853	0.50000
UPDATED PRIOR ON THRESHOLDS; MEAN & SD =	-0.20266	2.00000

CYCLE 20; LARGEST CHANGE= 0.00536

[NEWTON CYCLES]

UPDATED PRIOR ON ASYMPTOTES; ALPHA & BETA = 4.80903 17.19097  
 UPDATED PRIOR ON LOG SLOPES; MEAN & SD = -0.01823 0.50000  
 UPDATED PRIOR ON THRESHOLDS; MEAN & SD = -0.20400 2.00000

-2 LOG LIKELIHOOD: 15862.5413

CYCLE 21; LARGEST CHANGE= 0.00869

INTERVAL COUNTS FOR COMPUTATION OF ITEM CHI-SQUARES

-----  
 11. 19. 51. 52. 65. 66. 60. 46. 30.  
 -----

INTERVAL AVERAGE THETAS

-----  
 -2.312 -1.542 -1.111 -0.686 -0.202 0.165 0.669 1.136 1.897  
 -----

1

SUBTEST TEST0001; ITEM PARAMETERS AFTER CYCLE 21

ITEM	INTERCEPT S.E.	SLOPE S.E.	THRESHOLD S.E.	LOADING S.E.	ASYMPTOTE S.E.	CHISQ (PROB)	DF
ITEM0001	-2.470 0.589*	1.173 0.348*	2.106 0.236*	0.761 0.226*	0.066 0.017*	4.2 (0.7557)	7.0
ITEM0002	-1.449 0.483*	1.644 0.458*	0.882 0.112*	0.854 0.238*	0.277 0.036*	13.5 (0.0951)	8.0
ITEM0003	0.150 0.159*	0.194 0.049*	-0.772 0.869*	0.191 0.048*	0.270 0.100*	16.3 (0.0616)	9.0
ITEM0004	0.172 0.159*	0.410 0.085*	-0.419 0.424*	0.379 0.078*	0.235 0.095*	3.1 (0.9604)	9.0
ITEM0005	0.713 0.124*	0.317 0.068*	-2.249 0.635*	0.302 0.065*	0.214 0.094*	8.8 (0.4578)	9.0
ITEM0006	1.146 0.130*	0.484 0.100*	-2.366 0.503*	0.436 0.090*	0.195 0.089*	7.4 (0.4895)	8.0
ITEM0007	1.460 0.168*	1.299 0.188*	-1.124 0.138*	0.792 0.115*	0.142 0.067*	10.4 (0.0344)	4.0
ITEM0008	1.553 0.205*	1.280 0.235*	-1.214 0.168*	0.788 0.145*	0.189 0.082*	3.7 (0.7118)	6.0
ITEM0009	-1.861 0.467*	1.263 0.351*	1.473 0.153*	0.784 0.218*	0.120 0.027*	2.7 (0.9495)	8.0
ITEM0010	0.939 0.124*	0.530 0.088*	-1.773 0.358*	0.468 0.078*	0.189 0.086*	3.3 (0.9119)	8.0
ITEM0011	0.363 0.159*	1.006 0.210*	-0.361 0.194*	0.709 0.148*	0.237 0.083*	2.5 (0.8712)	6.0

ITEM0012	0.752 0.140*	1.297 0.234*	-0.579 0.140*	0.792 0.143*	0.179 0.070*	1.2 (0.9782)	6.0
ITEM0013	0.599 0.129*	0.908 0.134*	-0.659 0.180*	0.672 0.099*	0.174 0.073*	4.9 (0.7698)	8.0
ITEM0014	-0.689 0.245*	1.094 0.247*	0.630 0.133*	0.738 0.166*	0.174 0.049*	5.4 (0.6148)	7.0
ITEM0015	-0.434 0.232*	0.824 0.184*	0.527 0.206*	0.636 0.142*	0.247 0.065*	4.9 (0.8452)	9.0
ITEM0016	0.582 0.151*	1.132 0.201*	-0.515 0.174*	0.749 0.133*	0.238 0.077*	4.6 (0.7109)	7.0
ITEM0017	-1.589 0.395*	1.881 0.448*	0.845 0.076*	0.883 0.210*	0.088 0.023*	2.7 (0.9101)	7.0
ITEM0018	0.926 0.139*	0.975 0.165*	-0.950 0.189*	0.698 0.118*	0.186 0.081*	3.6 (0.8204)	7.0
ITEM0019	0.985 0.126*	0.852 0.122*	-1.156 0.200*	0.649 0.093*	0.160 0.074*	7.0 (0.3230)	6.0
ITEM0020	-0.899 0.324*	0.787 0.229*	1.142 0.228*	0.619 0.180*	0.280 0.055*	15.3 (0.0823)	9.0
ITEM0021	1.249 0.152*	1.235 0.188*	-1.011 0.145*	0.777 0.118*	0.151 0.068*	8.4 (0.1333)	5.0
ITEM0022	0.942 0.146*	1.014 0.166*	-0.929 0.196*	0.712 0.117*	0.213 0.085*	8.5 (0.2874)	7.0
ITEM0023	0.845 0.126*	0.824 0.132*	-1.026 0.206*	0.636 0.102*	0.168 0.077*	3.3 (0.8517)	7.0
ITEM0024	0.802 0.134*	1.126 0.185*	-0.713 0.154*	0.748 0.123*	0.167 0.072*	3.9 (0.6864)	6.0
ITEM0025	-1.093 0.366*	1.244 0.333*	0.878 0.136*	0.779 0.209*	0.258 0.043*	3.3 (0.9160)	8.0
ITEM0026	-0.388 0.176*	0.832 0.151*	0.467 0.162*	0.640 0.116*	0.160 0.053*	14.3 (0.0733)	8.0
ITEM0027	0.407 0.135*	1.384 0.219*	-0.294 0.115*	0.810 0.129*	0.162 0.055*	7.4 (0.2853)	6.0
ITEM0028	-0.255 0.192*	1.832 0.387*	0.139 0.087*	0.878 0.185*	0.163 0.042*	2.1 (0.9104)	6.0
ITEM0029	0.560 0.164*	1.189 0.228*	-0.471 0.179*	0.765 0.147*	0.268 0.081*	6.8 (0.4467)	7.0
ITEM0030	-0.412 0.183*	1.289 0.237*	0.320 0.106*	0.790 0.145*	0.139 0.043*	4.7 (0.6979)	7.0
ITEM0031	1.529 0.176*	0.781 0.156*	-1.959 0.322*	0.615 0.123*	0.202 0.091*	9.9 (0.1964)	7.0
ITEM0032	0.644 0.145*	0.636 0.122*	-1.013 0.304*	0.537 0.103*	0.236 0.097*	15.4 (0.0521)	8.0
ITEM0033	0.308 0.161*	0.829 0.160*	-0.372 0.231*	0.638 0.124*	0.236 0.087*	8.8 (0.2682)	7.0
ITEM0034	0.676 0.131*	0.884 0.138*	-0.764 0.190*	0.663 0.103*	0.178 0.078*	2.0 (0.9581)	7.0





**ภาคผนวก ฅ**

ตัวอย่าง: ข้อมูลไฟล์ผลลัพธ์ของผู้สอบ  
ของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์  
ภายใต้เงื่อนไขการไม่ให้ทวนคำตอบ และการให้ทวนคำตอบ

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก ฉ

ด.ช. พรเทพ เหมชัยเจริญกิจ (2\_1\_พรเทพ\_ไม่ทวน.txt)

-----  
ไม่ทวน  
-----

โจทย์ข้อที่ 1 ได้ข้อคั้ง 201

ตอบผิด

Xeta=-0.0838419967281679

SD=0.945823983012809

โจทย์ข้อที่ 2 ได้ข้อคั้ง 91

ตอบถูก

Xeta=0.272104374550452

SD=0.837784854180824

โจทย์ข้อที่ 3 ได้ข้อคั้ง 41

ตอบถูก

Xeta=0.732303975466623

SD=0.766468651967635

โจทย์ข้อที่ 4 ได้ข้อคั้ง 208

ตอบถูก

Xeta=1.16401931912421

SD=0.652101123560067

โจทย์ข้อที่ 5 ได้ข้อคั้ง 70

ตอบถูก

Xeta=1.48855803324827

SD=0.585390916577118

โจทย์ข้อที่ 6 ได้ข้อคั้ง 39

ตอบถูก

Xeta=1.66770663881358

SD=0.537915977087573

โจทย์ข้อที่ 7 ได้ข้อคคลัง 151

ตอบถูก

Xeta=1.85573077387427

SD=0.493407110868936

โจทย์ข้อที่ 8 ได้ข้อคคลัง 34

ตอบถูก

Xeta=2.0059788067504

SD=0.460272615820894

โจทย์ข้อที่ 9 ได้ข้อคคลัง 103

ตอบถูก

Xeta=2.11448113430473

SD=0.435565972363308

โจทย์ข้อที่ 10 ได้ข้อคคลัง 150

ตอบถูก

Xeta=2.19144386345527

SD=0.415204130004348

โจทย์ข้อที่ 11 ได้ข้อคคลัง 121

ตอบผิด

Xeta=2.01452514997552

SD=0.384319906960933

โจทย์ข้อที่ 12 ได้ข้อคคลัง 224

ตอบผิด



สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

Xeta=1.78422833575616

SD=0.349524323103981

โจทย์ข้อที่ 13 ได้ข้อคลัง 48

ตอบผิด

Xeta=1.59334856249234

SD=0.316530355939863

โจทย์ข้อที่ 14 ได้ข้อคลัง 110

ตอบถูก

Xeta=1.66353984295422

SD=0.305196365646312

โจทย์ข้อที่ 15 ได้ข้อคลัง 134

ตอบถูก

Xeta=1.72780883722002

SD=0.294917605583392

โจทย์ข้อที่ 16 ได้ข้อคลัง 129

ตอบถูก

Xeta=1.77971461362137

SD=0.285891990806928

โจทย์ข้อที่ 17 ได้ข้อคลัง 18

ตอบถูก

Xeta=1.82725991688522

SD=0.279240617719985

โจทย์ข้อที่ 18 ได้ข้อคลัง 113

ตอบถูก

Xeta=1.86281766964185

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

SD=0.272048583182017

โจทย์ข้อที่ 19 ได้ข้อคั้ง 106

ตอบถูก

Xeta=1.90790037187468

SD=0.266718351788456

โจทย์ข้อที่ 20 ได้ข้อคั้ง 37

ตอบถูก

Xeta=1.94567442940347

SD=0.263501162914148

ตอบถูก 16 ข้อ ตอบผิด4 ข้อ

เวลาในการทดสอบ 13 นาที 18 วินาที

ด.ช. พรเทพ เจริญเจริญกิจ (2\_1\_พรเทพ\_ทวน.txt)

-----  
ทวน  
-----

โจทย์ข้อที่ 1 ได้ข้อคั้ง 200

ก่อนทวน

หลังทวน

ตอบถูก

ตอบถูก

Xeta เก่า =0.327525121428286

Xeta ใหม่ =0.327525103767614

SD เก่า =0.93427622229185

SD ใหม่ =0.934276229056045

โจทย์ข้อที่ 2 ได้ข้อคั้ง 218

ก่อนทวน

หลังทวน

ตอบผิด

ตอบถูก

Xeta เก่า =-0.433384752215491

Xeta ใหม่ =0.707618864061576

SD เก่า =0.66960200633174

SD ใหม่ =0.806991216706844

โจทย์ข้อที่ 3 ได้ข้อคคลัง 54

ก่อนทวน

หลังทวน

ตอบถูก

ตอบถูก

Xeta เก่า = -0.166632537460042

Xeta ใหม่ = 0.846579749297371

SD เก่า = 0.629323708462545

SD ใหม่ = 0.737441461895417

โจทย์ข้อที่ 4 ได้ข้อคคลัง 123

ก่อนทวน

หลังทวน

ตอบถูก

ตอบถูก

Xeta เก่า = 0.0981423553377224

Xeta ใหม่ = 1.00082079165312

SD เก่า = 0.59327599801953

SD ใหม่ = 0.671564839992282

โจทย์ข้อที่ 5 ได้ข้อคคลัง 57

ก่อนทวน

หลังทวน

ตอบถูก

ตอบถูก

Xeta เก่า = 0.285696983885758

Xeta ใหม่ = 1.0971523216542

SD เก่า = 0.558368534219566

SD ใหม่ = 0.627174489274643

โจทย์ข้อที่ 6 ได้ข้อคคลัง 30

ก่อนทวน

หลังทวน

ตอบผิด

ตอบถูก

Xeta เก่า = 0.674577916919471

Xeta ใหม่ = 1.38105738573234

SD เก่า = 0.49563757881919

SD ใหม่ = 0.540370754831472

โจทย์ข้อที่ 7 ได้ข้อคคลัง 124

ก่อนทวน

หลังทวน

ตอบถูก

ตอบถูก

Xeta เก่า = 0.903971736579644

Xeta ใหม่ = 1.49800004539861

SD เก่า = 0.448906042682828

SD ใหม่ = 0.489915813107521

โจทย์ข้อที่ 8 ได้ข้อคคลัง 131

ก่อนทวน	หลังทวน
ตอบถูก	ตอบถูก
Xeta เก่า =1.07310211723507	Xeta ใหม่ =1.59527036161973
SD เก่า =0.416421946926635	SD ใหม่ =0.451984213753333

โจทย์ข้อที่ 9 ได้ข้อคล้ง 96

ก่อนทวน	หลังทวน
ตอบถูก	ตอบถูก
Xeta เก่า =1.20352025424113	Xeta ใหม่ =1.67690021357485
SD เก่า =0.393455020685611	SD ใหม่ =0.423265508702609

โจทย์ข้อที่ 10 ได้ข้อคล้ง 118

ก่อนทวน	หลังทวน
ตอบถูก	ตอบถูก
Xeta เก่า =1.40448190692111	Xeta ใหม่ =1.80979468068163
SD เก่า =0.364790145062979	SD ใหม่ =0.379052049681455

โจทย์ข้อที่ 11 ได้ข้อคล้ง 140

ก่อนทวน	หลังทวน
ตอบถูก	ตอบถูก
Xeta เก่า =1.91593072171058	Xeta ใหม่ =1.91593064077386
SD เก่า =0.350656591181768	SD ใหม่ =0.350656625961048

โจทย์ข้อที่ 12 ได้ข้อคล้ง 108

ก่อนทวน	หลังทวน
ตอบถูก	ตอบถูก
Xeta เก่า =2.03967908967605	Xeta ใหม่ =2.03967904020177
SD เก่า =0.328285227115916	SD ใหม่ =0.328285256262801

โจทย์ข้อที่ 13 ได้ข้อคล้ง 149

ก่อนทวน	หลังทวน
---------	---------

ตอบผิด

Xeta เก่า =1.93691858348084  
SD เก่า =0.308436037779591

ตอบผิด

Xeta ใหม่ =1.93691853644045  
SD ใหม่ =0.308436067612484

โจทย์ข้อที่ 14 ได้ข้อคล้ง 32

ก่อนทวน

ตอบถูก

Xeta เก่า =1.99269451425157  
SD เก่า =0.295726494372612

หลังทวน

ตอบถูก

Xeta ใหม่ =1.99269439601724  
SD ใหม่ =0.295726548609707

โจทย์ข้อที่ 15 ได้ข้อคล้ง 130

ก่อนทวน

ตอบถูก

Xeta เก่า =2.04109201168444  
SD เก่า =0.285187078105593

หลังทวน

ตอบถูก

Xeta ใหม่ =2.04109183884875  
SD ใหม่ =0.2851871511832

โจทย์ข้อที่ 16 ได้ข้อคล้ง 86

ก่อนทวน

ตอบถูก

Xeta เก่า =2.13396669455316  
SD เก่า =0.275497225618953

หลังทวน

ตอบถูก

Xeta ใหม่ =2.13396662388756  
SD ใหม่ =0.27549724503266

โจทย์ข้อที่ 17 ได้ข้อคล้ง 78

ก่อนทวน

ตอบถูก

Xeta เก่า =2.19164242478795  
SD เก่า =0.269066656707129

หลังทวน

ตอบถูก

Xeta ใหม่ =2.19164236460511  
SD ใหม่ =0.269066675454137

โจทย์ข้อที่ 18 ได้ข้อคล้ง 93

ก่อนทวน

ตอบถูก

หลังทวน

ตอบถูก

Xetaเก่า = 2.26826478534895

Xetaใหม่ = 2.26826466324239

SDเก่า = 0.263659799831867

SDใหม่ = 0.263659834543588

โจทย์ข้อที่ 19 ได้ข้อคล้ง 10

ก่อนทวน

หลังทวน

ตอบผิด

ตอบผิด

Xetaเก่า = 2.17079646207984

Xetaใหม่ = 2.17079641237251

SDเก่า = 0.250739218037186

SDใหม่ = 0.25073927648834

โจทย์ข้อที่ 20 ได้ข้อคล้ง 84

ก่อนทวน

หลังทวน

ตอบถูก

ตอบถูก

Xetaเก่า = 2.21053562553365

Xetaใหม่ = 2.21053559157453

SDเก่า = 0.246670949262576

SDใหม่ = 0.246671005961249

มีการเปลี่ยนคำตอบจากผิดเป็นถูก 2 ข้อ

มีการเปลี่ยนคำตอบจากผิดเป็นผิด 0 ข้อ

มีการเปลี่ยนคำตอบจากถูกเป็นผิด 0 ข้อ

มีการเปลี่ยนคำตอบรวม 2 ข้อ

ทำเครื่องหมายเพื่อกลับไปทวน 12 ข้อ

กลับไปทวนข้อที่ทำเครื่องหมาย 12 ข้อ

กลับไปทวนข้อที่ไม่ได้ทำเครื่องหมาย 0 ข้อ

เวลาในการทดสอบ 22 นาที 6 วินาที

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



## ประวัติผู้เขียนวิทยานิพนธ์

นางสาวพิมพ์สิริ เขียรนรเศรษฐ์ เกิดเมื่อวันที่ 17 มีนาคม พ.ศ.2526 สำเร็จการศึกษา ครุศาสตรบัณฑิต สาขาวิชามัธยมศึกษา วิชาเอกวิทยาศาสตร์ทั่วไปและคณิตศาสตร์ จากจุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2548 และเข้าศึกษาต่อในระดับบัณฑิตศึกษา สาขาการวัดและประเมินผล การศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปัจจุบันเป็นครู และผู้บริหารโรงเรียนกวดวิชาพี.พี.เซ็นเตอร์



สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย