

การเปรียบเทียบความตรงตามสภาพและดวงคงที่ในการประมาณค่าความสามารถของผู้สอบในการ
ทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบ
และอัตราการใช้ข้อสอบซ้ำแตกต่างกัน



นางสาวกนกวรรณ รัตนธน

สถาบันวิทยบริการ
วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา


คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2544

ISBN 974-030-564-4

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

**A COMPARISON OF CONCURRENT VALIDITY AND STABILITY IN THE ESTIMATION
OF EXAMINEES' ABILITY IN COMPUTERIZED ADAPTIVE TESTING WITH DIFFERENCE
OF EXAMINEES' ABILITY AND ITEM EXPOSURE RATES**



Miss Kanokwan Ratanathon

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

**Faculty of Education
Chulalongkorn University**

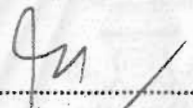
Academic year 2001

ISBN 974-030-564-4

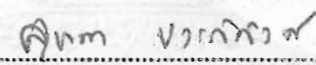
I 20603745

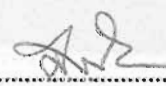
หัวข้อวิทยานิพนธ์ การเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถ
 ของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วย
 คอมพิวเตอร์เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบแตกต่างกัน
 โดย นางสาวกนกวรรณ รัตนชน
 สาขาวิชา การวัดและประเมินผลการศึกษา
 อาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี
 อาจารย์ที่ปรึกษาร่วม ผู้ช่วยศาสตราจารย์ ดร.ภาวิณี ศรีสุขวัฒนานันท์


คณะกรรมการ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยรับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาโท

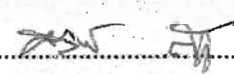

คณบดีคณะครุศาสตร์
 (รองศาสตราจารย์ ดร.ไพฑูรย์ สินลารัตน์)


คณะกรรมการสอบวิทยานิพนธ์


ประธานกรรมการ
 (ผู้ช่วยศาสตราจารย์ ดร.สุชาดา บวรกิตติวงศ์)


อาจารย์ที่ปรึกษา
 (รองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)


อาจารย์ที่ปรึกษาร่วม
 (ผู้ช่วยศาสตราจารย์ ดร.ภาวิณี ศรีสุขวัฒนานันท์)


กรรมการ
 (รองศาสตราจารย์ ดร.ศิริเดช สุขีวะ)


กรรมการ
 (อาจารย์ ดร.เอมอร จังศิริพรกรณ์)

กนกวรรณ รัตนธน : การเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของ
 ผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของ
 ผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน (A COMPARISON OF CONCURRENT VALIDITY AND
 STABILITY IN THE ESTIMATION OF EXAMINEES' ABILITY IN COMPUTERIZED ADAPTIVE
 TESTING WITH DIFFERENCE OF EXAMINEES' ABILITY AND ITEM EXPOSURE RATES) อ.ที่ปรึกษา :
 รศ.ดร.ศิริชัย กาญจนวาสี. อ.ที่ปรึกษาร่วม : ผศ.ดร.ภาวิณี ศรีสุขวัฒนานนท์. 157 หน้า. ISBN 974-030-564-4

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถ
 ของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบ และ
 อัตราการใช้ข้อสอบซ้ำแตกต่างกัน กลุ่มตัวอย่างคือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2543 จังหวัด
 สุพรรณบุรี จำนวน 319 คน เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบวัดความสามารถทางภาษาอังกฤษ คลังข้อสอบวิชาภาษา
 อังกฤษ และโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (MicroCAT) วิเคราะห์
 ข้อมูลโดยใช้ค่าสถิติเบื้องต้น ค่าสหสัมพันธ์ ค่าโคสแควร์ และการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ

ผลการวิจัยพบว่า

1. ระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำที่แตกต่างกัน ทำให้ความตรงตามสภาพในการ
 ประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน
 โดยผู้สอบที่มีระดับความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น ส่งผลให้ความตรงตามสภาพของการทดสอบลดลง
2. ระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำที่แตกต่างกัน ทำให้ความคงที่ในการประมาณค่า
 ความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน โดยผู้สอบที่มี
 ระดับความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น ส่งผลให้ความคงที่ในการประมาณค่าความสามารถของผู้สอบลด
 ลง
3. ปฏิสัมพันธ์ระหว่างระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำร่วมกันส่งผลต่อความคงที่ในการ
 ประมาณค่าความสามารถของผู้สอบ โดยผู้สอบที่มีระดับความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น จะทำให้แนวโน้ม
 ของการประมาณค่าความสามารถของผู้สอบสูงขึ้นด้วย
4. ปฏิสัมพันธ์ระหว่างจำนวนครั้งที่ทำการทดสอบและอัตราการใช้ข้อสอบซ้ำร่วมกันส่งผลต่อความคงที่ในการ
 ประมาณค่าความสามารถของผู้สอบ เมื่อจำนวนครั้งที่ทำการทดสอบเพิ่มขึ้นและมีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น จะทำให้แนว
 โหมดของการประมาณค่าความสามารถของผู้สอบสูงขึ้นด้วย

ภาควิชา ศึกษาศาสตร์
 สาขาวิชา การวัดและประเมินผลการศึกษา
 ปีการศึกษา 2544

ลายมือชื่อนิติ
 ลายมือชื่ออาจารย์ที่ปรึกษา
 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม
(Handwritten signatures and initials)

4283651527 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION

KEY WORD : CONCURRENTE VALIDITY / STABILITY /COMPUTERIZED ADAPTIVE TESTING

KANOKWAN RATANATHON : A COMPARISON OF CONCURRENT VALIDITY AND STABILITY IN THE ESTIMATION OF EXAMINEES' ABILITY IN COMPUTERIZED ADAPTIVE TESTING WITH DIFFERENCE OF EXAMINEES' ABILITY AND ITEM EXPOSURE RATES. THESIS ADVISOR : ASSD.PROF.SIRICHAJ KANJANAWASEE, PH.D. THESIS CO - ADVISOR : ASSIST.PROF.PAWINEE SRISUKWATANANAN. 157 pp. ISBN 974-030-564-4

The purpose of this research was to compare concurrent validity and stability in the estimation of examinees' ability in computerized adaptive testing with difference of examinees' ability and item exposure rates. The sample was 319 students of Matthayom 3 in the second semester of 2000 academic year in Suphanburi Province. The research instruments were English ability test, English item pools and Computer program for Computerized adaptive testing (MicroCAT). Data were analyzed by using basic statistical methods, correlation coefficient , chi - square and analysis of variance on repeated measurement.

The findings of this research were :

1. The differences in examinees' ability and item exposure rate was affected on the concurrent validity in the estimation of examinees' ability, examinees with high ability and higher item exposure rate were affected on the decreasing in concurrent validity.
2. The difference in examinees' ability and item exposure rate was affected on the stability in the estimation of examinees' ability, examinees with high ability and higher item exposure rate were affected on the decreasing in concurrent validity.
3. Interaction between examinees' ability and item exposure rate were affected on stability in the estimation of examinees' ability, examinees with high ability and higher item exposure rate were caused higher tendency in the estimation of examinees' ability.
4. Interaction between test repeatability and item exposure rate were affected on stability in the estimation of examinees' ability, higher test repeatability and higher item exposure rate were caused higher tendency in the estimation of examinees' ability.

Department Education Research

Field of Study Education Measurement and Evaluation

Academic Year 2001

Student's signature Kanokwan R.

Advisor's signature Sirichai R.

Co-Advisor's signature Pawinee Srisk

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดีเป็นเพราะได้รับความกรุณาช่วยเหลืออย่างดีจากบุคคลหลายท่าน โดยเฉพาะบุคคลแรกที่ผู้วิจัยขอกราบขอบพระคุณรองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้เสียสละเวลาอันมีค่ากรุณาให้คำปรึกษา แนะนำ ชี้แนะแนวทางในการทำวิจัยครั้งนี้ ตรวจสอบแก้ไขข้อบกพร่องต่างๆ ด้วยความเอาใจใส่ และเป็นกำลังใจแก่ผู้วิจัยอย่างยิ่งตลอดมาจนงานวิจัยสำเร็จลุล่วงด้วยดี รวมทั้งผู้ช่วยศาสตราจารย์ ดร.ภาวิณี ศรีสุขวัฒน์นันท์ อาจารย์ที่ปรึกษา ร่วมวิทยานิพนธ์ ที่ท่านได้กรุณาอนุเคราะห์ให้ผู้วิจัยได้ใช้โปรแกรม MicroCAT และชี้แนะแนวทางที่เป็นประโยชน์ในการทำวิจัยครั้งนี้

ขอกราบขอบพระคุณศาสตราจารย์ ดร.นงลักษณ์ วิรัชชัย รองศาสตราจารย์ ดร.สุวิมล ว่องวานิช ผู้ช่วยศาสตราจารย์ ดร.สุวัฒนา สุวรรณเขตนิกม และรองศาสตราจารย์พวงแก้ว ปุณยกนก ที่ให้ข้อเสนอแนะและคำปรึกษาที่เป็นประโยชน์จากการเรียนวิชาสัมมนาการวิจัย นอกจากนี้ผู้วิจัยขอกราบขอบพระคุณคณาจารย์ภาควิชาวิจัยการศึกษาทุกท่าน ที่ได้ประสิทธิ์ประสาทความรู้แก่ผู้วิจัย

ขอกราบขอบพระคุณ ดร.รังสรรค์ มณีเล็ก ที่ให้คำข้อเสนอแนะที่เป็นประโยชน์อย่างยิ่งต่อการวิจัยครั้งนี้ ผู้ช่วยศาสตราจารย์ ดร.แมนมาศ ลีลสัตยกุล ที่ให้ใช้คลังข้อสอบวิชาภาษาอังกฤษ ผู้อำนวยการสำนักทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ และอาจารย์ชัยวัฒน์ วัชรมัย ที่ให้ความช่วยเหลือและให้คำแนะนำในการนำแบบสอบวัดความสามารถทางภาษาอังกฤษมาใช้ในงานวิจัยครั้งนี้

ขอขอบพระคุณผู้บริหารโรงเรียน ผู้ช่วยฝ่ายวิชาการ หัวหน้าศูนย์คอมพิวเตอร์ และนักเรียนโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา จังหวัดสุพรรณบุรี ที่ให้ความร่วมมือเป็นอย่างดี ในการเก็บรวบรวมข้อมูลสำหรับการวิจัยในครั้งนี้ และขอขอบคุณบัณฑิตวิทยาลัยที่ได้มอบทุนอุดหนุนการทำวิจัยในครั้งนี้

ขอขอบคุณไพรัช จุลบุญญาสิทธิ อนุชา ลีมบุพศิริพร รักชนก ยี่สุนศรี ราชัย จำทรัพย์ ทิพวรรณ ประเสริฐอำไพสกุล จุฬาลักษณ์ ชันธบุตร สุธิสา นามเหล่า สดใส ขณะรัตน์และเพื่อน ๆ ภาควิชาวิจัยการศึกษาทุกท่านที่ให้ความช่วยเหลือ และเป็นกำลังใจให้ตลอดมา

ท้ายนี้ ผู้วิจัยใคร่ขอกราบขอบพระคุณ อาจารย์สุรศักดิ์ เข็มเงิน นางปณิตดา เข็มเงิน ผู้เป็นมารดา และน้องชาย ผู้ให้กำลังใจและสนับสนุนส่งเสริมการศึกษาของผู้วิจัยอย่างยิ่งตลอดมา

กนกวรรณ รัตนธน

ผู้วิจัย

สารบัญ

| | หน้า |
|--|------|
| บทคัดย่อภาษาไทย..... | ง |
| บทคัดย่อภาษาอังกฤษ..... | จ |
| กิตติกรรมประกาศ..... | ฉ |
| สารบัญ..... | ช |
| สารบัญตาราง..... | ฌ |
| สารบัญแผนภาพ..... | ฎ |
| บทที่ | |
| 1 บทนำ..... | 1 |
| ความเป็นมาและความสำคัญของปัญหา..... | 1 |
| คำถามวิจัย..... | 5 |
| วัตถุประสงค์การวิจัย..... | 6 |
| สมมติฐานการวิจัย..... | 6 |
| ขอบเขตการวิจัย..... | 7 |
| ข้อตกลงเบื้องต้น..... | 8 |
| คำจำกัดความที่ใช้ในการวิจัย..... | 8 |
| ข้อจำกัดของการวิจัย..... | 9 |
| ประโยชน์ที่คาดว่าจะได้รับ..... | 10 |
| 2 เอกสารและงานวิจัยที่เกี่ยวข้อง..... | 11 |
| ทฤษฎีการตอบสนองข้อสอบ..... | 11 |
| ลักษณะทั่วไป..... | 11 |
| หลักการของทฤษฎี..... | 13 |
| ข้อตกลงเบื้องต้นของทฤษฎี..... | 13 |
| โมเดลตามทฤษฎี..... | 15 |
| ปัจจัยที่มีผลต่อคุณภาพของการทดสอบ..... | 18 |

สารบัญ (ต่อ)

| บทที่ | หน้า |
|---|------|
| การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ..... | 21 |
| ความหมาย..... | 21 |
| หลักการของการทดสอบ..... | 21 |
| ประเภทของการทดสอบ..... | 22 |
| การให้คะแนนการทดสอบ..... | 32 |
| องค์ประกอบของการทดสอบ..... | 38 |
| การดำเนินการทดสอบ..... | 40 |
| ลักษณะของระบบการทดสอบไมโครแคท (MicroCAT)..... | 41 |
| ลักษณะทั่วไป..... | 41 |
| การใช้ระบบการทดสอบ..... | 43 |
| งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ..... | 46 |
| 3 วิธีดำเนินการวิจัย..... | 52 |
| เครื่องมือที่ใช้ในการวิจัย..... | 53 |
| การพัฒนาโปรแกรมสำหรับการทดสอบ..... | 58 |
| การเก็บรวบรวมข้อมูล..... | 61 |
| การวิเคราะห์ข้อมูล..... | 63 |
| 4 ผลการวิเคราะห์ข้อมูล..... | 70 |
| 5 สรุป อภิปรายและข้อเสนอแนะ..... | 100 |
| สรุปผลการวิจัย..... | 100 |
| อภิปรายผลการวิจัย..... | 104 |
| ข้อเสนอแนะ..... | 107 |
| รายการอ้างอิง..... | 110 |
| ภาคผนวก..... | 114 |
| ประวัติผู้เขียนวิทยานิพนธ์..... | 157 |

สารบัญตาราง

| ตาราง | หน้า |
|---|------|
| 1 จำนวนกลุ่มตัวอย่าง จำแนกเป็นรายโรงเรียน..... | 52 |
| 2 จำนวนนักเรียนกลุ่มตัวอย่าง จำแนกเป็นรายโรงเรียน และตามระดับความสามารถ | 53 |
| 3 โครงสร้างของข้อสอบวัดความสามารถทางภาษาอังกฤษ..... | 54 |
| 4 ผลการวิเคราะห์แบบสอบวัดความสามารถทางภาษาอังกฤษ..... | 55 |
| 5 โครงสร้างของคลังข้อสอบวิชาภาษาอังกฤษ จำแนกตามจุดประสงค์การวัด..... | 56 |
| 6 เนื้อหาในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ | 58 |
| 7 กลุ่มตัวอย่าง และเนื้อหาในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ | 62 |
| 8 จำนวนกลุ่มตัวอย่างที่นำมาวิเคราะห์..... | 63 |
| 9 ตารางการวิเคราะห์ความแปรปรวนสองทางแบบวัดซ้ำ..... | 69 |
| 10 จำนวนผู้สอบ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบ และจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ จำนวน 5 ครั้ง..... | 71 |
| 11 ค่าความตรงตามสภาพ และค่าไคสแควร์สำหรับการทดสอบความแตกต่างของค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ | 75 |
| 12 การทดสอบความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถต่ำ | 79 |
| 13 การทดสอบความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถปานกลาง | 79 |
| 14 การทดสอบความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถสูง | 80 |
| 15 ความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถของผู้สอบ..... | 81 |

สารบัญตาราง (ต่อ)

| ตาราง | หน้า |
|--|------|
| 16 ความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ จำแนกตามอัตราการใช้ข้อสอบซ้ำ..... | 81 |
| 17 การทดสอบความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับ ความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถของผู้สอบ..... | 82 |
| 18 การทดสอบความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับ ความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถของผู้สอบ..... | 83 |
| 19 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการ ทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตาม แหล่งความแปรปรวน..... | 84 |
| 20 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการ การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมี ความสามารถต่ำ และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์..... | 87 |
| 21 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการ การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมี ความสามารถปานกลาง และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์..... | 88 |
| 22 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการ การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมี ความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์..... | 89 |
| 23 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการ การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมี ความสามารถต่ำ และมีอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์..... | 90 |
| 24 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการ การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมี | |

สารบัญตาราง (ต่อ)

| ตาราง | หน้า |
|-------|---|
| | ความสามารถปานกลาง และมีอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์..... 91 |
| 25 | ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์..... 92 |
| 26 | ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถต่ำ และมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์..... 93 |
| 27 | ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถปานกลาง และมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์..... 94 |
| 28 | ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์..... 95 |
| 29 | ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่อมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์..... 96 |
| 30 | ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่อมีอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์..... 97 |
| 31 | ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่อมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์..... 98 |

สารบัญแผนภาพ

แผนภาพที่

หน้า

| | | |
|----|---|----|
| 1 | แนวคิดของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ..... | 2 |
| 2 | โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 1 พารามิเตอร์..... | 16 |
| 3 | โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 2 พารามิเตอร์..... | 17 |
| 4 | โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 3 พารามิเตอร์..... | 18 |
| 5 | โครงสร้างของการทดสอบโดยใช้ยุทธวิธีสองขั้นตอน..... | 24 |
| 6 | โครงสร้างของการทดสอบรูปแบบปิรามิดขนาดชั้นคงที่..... | 24 |
| 7 | โครงสร้างของการทดสอบรูปแบบปิรามิดขนาดชั้นแปรผัน..... | 25 |
| 8 | โครงสร้างของการทดสอบรูปแบบปิรามิดชนิดรอกบิน - มอนโร..... | 26 |
| 9 | โครงสร้างของการทดสอบรูปแบบปิรามิดข้างตัดตามวิธีสกัดการสะท้อนกลับ..... | 26 |
| 10 | โครงสร้างของการทดสอบรูปแบบปิรามิดข้างตัดตามวิธีรักษาการสะท้อนกลับ..... | 27 |
| 11 | โครงสร้างของการทดสอบรูปแบบปิรามิดมีหลายข้อในแต่ละชั้น..... | 28 |
| 12 | โครงสร้างของการทดสอบรูปแบบปิรามิดชนิดให้น้ำหนักตัวเลือกข้อข้อสอบเพื่อแยกทาง..... | 29 |
| 13 | โครงสร้างของการทดสอบรูปแบบเพลิกซีเลเวล..... | 29 |
| 14 | โครงสร้างของการทดสอบรูปแบบปรับระดับขั้น..... | 30 |
| 15 | ผังงานแสดงขั้นตอนการทำงานของคอมพิวเตอร์ในการดำเนินการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ในรอบที่ 1..... | 59 |
| 16 | ผังงานแสดงขั้นตอนการทำงานของคอมพิวเตอร์ในการดำเนินการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ในรอบที่ 2-5..... | 60 |
| 17 | ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถ และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์..... | 76 |

สารบัญแผนภาพ (ต่อ)

| แผนภาพที่ | | หน้า |
|-----------|---|------|
| 18 | ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถ และมีอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซนต์..... | 77 |
| 19 | ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถ และมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซนต์..... | 77 |
| 20 | ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถ และมีอัตราการใช้ข้อสอบซ้ำ..... | 78 |
| 21 | ค่าความสามารถเฉลี่ย จำแนกตามระดับความสามารถและอัตราการใช้ข้อสอบซ้ำ..... | 85 |
| 22 | ค่าความสามารถเฉลี่ย จำแนกตามครั้งที่ทำการทดสอบและอัตราการใช้ข้อสอบซ้ำ..... | 86 |
| 23 | ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถต่ำ และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซนต์..... | 87 |
| 24 | ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถปานกลาง และมีอัตราการใช้ ข้อสอบซ้ำ 10 เปอร์เซนต์..... | 88 |
| 25 | ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซนต์..... | 89 |
| 26 | ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถต่ำ และอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซนต์..... | 90 |
| 27 | ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถปานกลาง และมีอัตราการใช้ ข้อสอบซ้ำ 20 เปอร์เซนต์..... | 91 |
| 28 | ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซนต์..... | 92 |

บทที่ 1

บทนำ

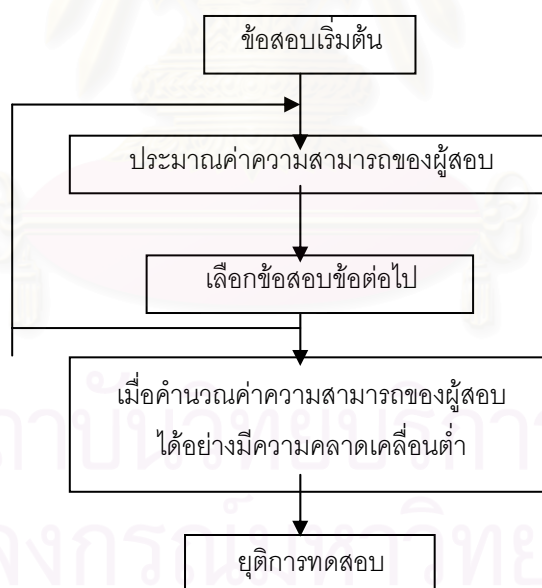
ความเป็นมาและความสำคัญของปัญหา

การวัดผลและประเมินผลทางการศึกษาเป็นองค์ประกอบที่สำคัญของกระบวนการจัดการศึกษา เนื่องจากสิ่งที่ได้จากการวัดผลและประเมินผลนั้นทำให้เราทราบว่าผลของการจัดการศึกษาบรรลุเป้าหมายที่กำหนดไว้หรือไม่มากน้อยเพียงไร โดยธรรมชาติการวัดพฤติกรรม หรือลักษณะทางจิตวิทยาไม่สามารถวัดได้โดยตรงเหมือนกับการวัดทางกายภาพ ต้องอาศัยการวัดทางอ้อม จึงมักมีความ คลาดเคลื่อนเกิดขึ้น ดังนั้น นักวัดผลจึงได้พยายามสร้างเครื่องมือและพัฒนาทฤษฎีการวัดต่างๆ ขึ้นมา เพื่อที่จะอธิบายหรือทำนายคุณลักษณะที่แท้จริงของบุคคล ให้มีความถูกต้องและเชื่อถือได้มากที่สุด วิธีวัดผลที่ใช้กันอยู่นั้นมีหลายชนิดด้วยกัน แต่ที่ได้รับความนิยมมากก็คือแบบสอบ ทั้งนี้อาจเนื่องมาจากกระบวนการสร้างและการดำเนินการทดสอบ ตลอดจนการแปลผลที่ได้จากการวัดนั้นค่อนข้างสะดวกและชัดเจน (รังสรรค์ มณีเล็ก, 2540)

การทดสอบแบบประเพณีนิยม (traditional testing) เป็นการใช้แบบสอบที่สร้างขึ้นตามทฤษฎีว่าด้วยการแปลความหมายผลการวัดแบบดั้งเดิม (classical testing theory) เพื่อวัด คุณลักษณะที่สนใจ เช่น ผลสัมฤทธิ์ ความสามารถ ความถนัด เจตคติ เป็นต้น การดำเนินการทดสอบจัดกระทำโดยกำหนดให้ผู้สอบทุกคนต้องทำข้อสอบทุกข้อเหมือนกันในเวลาเดียวกัน โดย แบบสอบจะประกอบด้วยข้อสอบที่ยากและง่าย ผลที่ได้จากการทดสอบ นักวัดผลจะกำหนดในรูปของคะแนน และใช้คะแนนตัดสินความสามารถของผู้สอบ ปัญหาอย่างหนึ่งที่เกิดขึ้นคือ คะแนนที่ได้อาจไม่สามารถอธิบายความสามารถที่แท้จริง หรือคุณลักษณะภายในของผู้สอบได้กับบุคคลที่มีพื้นฐานที่ แตกต่างกัน ทั้งนี้เพราะข้อสอบที่สร้างอาจไม่เหมาะสมกับความสามารถของผู้สอบทุกคน เช่น ข้อสอบจำนวนหนึ่งจะยากเกินไปสำหรับผู้สอบที่มีความสามารถต่ำ ผู้สอบกลุ่มนี้จึงต้องเดาคำตอบเป็นส่วนใหญ่ ในขณะที่ผู้สอบกลุ่มที่มีความสามารถสูงจำเป็นต้องทำข้อสอบบางส่วนที่ง่ายเกินไปสำหรับความสามารถของตน ผลการตอบจึงมิได้ให้สารสนเทศที่เป็นประโยชน์เกี่ยวกับผู้สอบเท่าที่ควร อันนำไปสู่การเกิดความคลาดเคลื่อนของการวัดที่สูงขึ้นได้ (ศิริชัย กาญจนวาสี, 2538) ทำให้นักวัดผลได้พยายามหาทางแก้ไขข้อจำกัดของการทดสอบที่อาศัยทฤษฎีการทดสอบแบบประเพณีนิยม ด้วยการพยายามจัด ข้อสอบให้เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล โดยมีแนวคิดว่าผู้ที่มีความสามารถต่ำควรจะทำได้ทำข้อสอบข้อ

ที่ง่าย และผู้ที่มีความสามารถสูงควรจะได้ทำข้อสอบข้อที่ยาก จากแนวคิดนี้ทำให้เกิดการพัฒนาการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive testing) โดยมี เป้าหมายที่จะดำเนินการทดสอบด้วยการคัดเลือกข้อสอบให้เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล โดยอาศัยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory, IRT) มาใช้เป็นพื้นฐานในการสร้างแบบสอบและวิธีการทดสอบ ซึ่งจะมีการคัดเลือกข้อสอบที่สามารถให้สารสนเทศอันเป็นประโยชน์เกี่ยวกับความสามารถของผู้สอบแต่ละคน ทำให้การทดสอบเป็นไปอย่างมีประสิทธิภาพ สามารถรายงานผลการทดสอบที่มีความคลาดเคลื่อนต่ำด้วยการใช้จำนวนข้อสอบที่สั้นลง ประหยัดเวลาของการทดสอบ (สายชล อบทม, 2539)

หลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ เป็นการคัดเลือก ข้อสอบสำหรับแต่ละบุคคลโดยอยู่บนพื้นฐานของผลการตอบข้อสอบข้อที่ผ่านมาของผู้สอบนั้น เมื่อผู้สอบทำข้อสอบข้อแรกซึ่งได้จากคลังข้อสอบแล้ว จะมีการวิเคราะห์ระดับความสามารถหรือประเมินความสามารถของผู้สอบเบื้องต้น เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีค่าความยาก และอำนาจจำแนกเหมาะสมที่จะใช้วัดระดับความสามารถของผู้สอบ และประมาณค่าความสามารถของผู้สอบใหม่ จากนั้นก็จะคัดเลือกข้อสอบที่เหมาะสมข้อต่อไป โดยอาศัยหลักการที่ว่าถ้าการทำข้อที่ผ่านมาถูกข้อถัดไปจะยากขึ้น แต่ถ้าการทำข้อที่ผ่านมาผิดข้อถัดไปจะง่ายลง กระบวนการนี้จะดำเนินต่อไปเรื่อยๆจนสามารถประมาณระดับความสามารถของผู้สอบได้อย่างเชื่อถือได้ตามเกณฑ์ที่กำหนดไว้ การทดสอบก็จะยุติลง (ศิริชัย กาญจนวาสี, 2538) ดังแผนภาพที่ 1



แผนภาพที่ 1 แนวคิดของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนับว่ามีประโยชน์ต่อการวัดผลการศึกษาเป็นอย่างมาก เนื่องจากประหยัดค่าใช้จ่าย ทั้งนี้เพราะใช้จำนวนข้อสอบน้อยกว่าการทดสอบแบบประเพณีนิยมและไม่จำเป็นต้องสร้างข้อสอบใหม่ทุกครั้งที่มีการสอบ เนื่องจากมีข้อสอบที่สร้างเก็บไว้ในคลังข้อสอบ ผู้สอบไม่เกิดความเหน็ดเหนื่อยและความเบื่อหน่ายในการสอบ เพราะใช้เวลาในการทดสอบน้อยลงกว่าเดิม และผู้สอบจะถูกกำหนดให้ทำข้อสอบที่มีความยากง่ายเหมาะสมกับความสามารถของผู้สอบ

ในปัจจุบันได้มีการนำเอาคอมพิวเตอร์มาประยุกต์ใช้ซึ่งเรียกว่า การทดสอบแบบ ปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์หรือการทดสอบแบบซี เอ ที (Computerized Adaptive Testing : CAT) โดยให้คอมพิวเตอร์ทำการคัดเลือกข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบ และทำการประมาณค่าความสามารถของผู้สอบ ซึ่งโปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นเพื่อการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบที่ได้รับความนิยมสูงในสหรัฐอเมริกา และแคนาดา คือ โปรแกรมการทดสอบไมโครแคท (MicroCAT) พัฒนาโดย Assessment System Corporation (1995) เป็นระบบการทดสอบที่มีความสามารถของการพัฒนากระบวนการทดสอบ ตั้งแต่การสร้างข้อสอบและจัดเก็บข้อสอบเป็นคลังข้อสอบ สามารถจัดชุดแบบสอบจากข้อสอบที่มีอยู่ใน คลังข้อสอบ ตลอดจนการทดสอบและการประมวลผล ซึ่งระบบการทดสอบของไมโครแคท (MicroCAT) มีการพัฒนาเป็นลำดับจนถึงรุ่นที่ 3 (version 3) (ภาวิณี ศรีสุขวัฒนานันท์, 2543)

ตั้งแต่มีการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ มีผู้ให้ความสนใจวิธีดำเนินการทดสอบ ทำการศึกษา ค้นคว้า และพัฒนารูปแบบของการทดสอบเพื่อให้เป็นการทดสอบที่มี ประสิทธิภาพ ซึ่งผลการวิจัยมีด้วยกันหลายประเด็นด้วยกัน เช่น เรื่องโมเดลการตอบสนองข้อสอบ เออร์รี่ (Urry, 1977) พบว่าโมเดลการตอบสนองข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์มีความเหมาะสมมากกว่าโมเดลอื่น และพบว่าขนาดของคลังข้อสอบที่วัดคุณลักษณะเดียวกันควรมีอย่างน้อย 100 ข้อ ข้อสอบภายในคลังข้อสอบควรมีค่าอำนาจจำแนก (a) อย่างน้อย 0.8 ค่าความยากง่าย (b) อยู่ระหว่าง -2.0 ถึง 2.0 และค่าการเดา (c) ควรน้อยกว่า 0.3 เรื่องความยากง่ายของการเสนอข้อสอบข้อแรก โฮ (Ho, 1989) เสนอให้เลือกข้อสอบข้อแรกที่มีความยากง่ายเท่ากับหรือต่ำกว่าความสามารถจริงของผู้สอบ เพราะการประมาณค่าความสามารถของผู้สอบจะมีความถูกต้องมากยิ่งขึ้น และแฮนกินส์ (Hankins, 1987) พบว่าการเสนอข้อสอบข้อแรกแบบกำหนดแน่นอน (fixed entry procedure) และการกำหนดแบบแปรผัน (variable entry procedure) ไม่มีผลให้เกิดความลำเอียงในการประมาณค่าความสามารถของผู้สอบ เรื่องเกณฑ์การคัดเลือกข้อสอบ ฮูลินและคณะ (Hulin and other, 1983 อ้างถึงใน รังสรรค์ มณีเล็ก, 2540) พบว่าการใช้เกณฑ์การ

คัดเลือก ข้อสอบโดยวิธีสารสนเทศสูงสุด (Maximum Information) ดีกว่าการใช้เกณฑ์การคัดเลือกข้อสอบโดยวิธีอื่น เรื่องวิธีการประมาณค่าความสามารถของผู้สอบ ต่าย เชื้อยงฉี (2534) พบว่าการประมาณค่าความสามารถตามวิธีของเบส์ (Bayesian) และวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (Conditional maximum likelihood) นั้นให้ผลการประมาณค่าไม่แตกต่างกัน และพบว่าการทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบ รูปแบบที่มีรูปแบบ จำนวนชั้น และวิธีการให้คะแนนต่างกัน จะส่งผลต่อความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ เรื่องประสิทธิภาพการทดสอบ แบล็กมอร์ (Blackmore, 1986) พบว่าคะแนนที่ได้จากการทดสอบด้วยคอมพิวเตอร์ การทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบด้วยคอมพิวเตอร์ และการทดสอบแบบประเพณีนิยมไม่แตกต่างกัน และนันทิยา พึ่งคำ (2531) พบว่า การทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่าการทดสอบแบบประเพณีนิยม และใช้จำนวนข้อสอบน้อยกว่าการทดสอบแบบประเพณีนิยม นอกจากนี้ รังสรรค์ มณีเล็ก (2540) พบว่า วิธีการประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการทดสอบ และระดับความสามารถของผู้สอบ มีผลต่อความตรงตามสภาพ และความยากง่ายของ ข้อสอบข้อแรก เกณฑ์การคัดเลือกข้อสอบข้อแรก วิธีการประมาณค่าความสามารถของผู้สอบ เกณฑ์ ยุติการทดสอบ และระดับความสามารถของผู้สอบ ส่งผลต่อจำนวนข้อสอบที่ใช้ในการทดสอบ เรื่องอัตราการใช้ข้อสอบซ้ำ ชาง และทวู (Chang and Twu, 1998) พบว่าวิธีการควบคุมอัตราการใช้ข้อสอบซ้ำด้วยวิธีของซิมสัน และเฮ็ทเทอร์ (Simpson and Hetter) มีประสิทธิภาพ ส่วนการควบคุมอัตราการใช้ข้อสอบซ้ำด้วยวิธีของสต็อกกิงและเลวิส (Stocking and Lewis, 1998) สามารถป้องกันไม่ให้ผู้สอบจำข้อสอบได้ คาลอห์น และสเปร (Kalohn and Spray, 1998) พบว่า วิธีการคัดเลือก ข้อสอบโดยมีอัตราการใช้ข้อสอบซ้ำ 10 และ 20 เปอร์เซ็นต์ มีความถูกต้องเหมือนกัน

จากผลการวิจัยที่กล่าวมาข้างต้น พบว่าประเด็นที่เกี่ยวกับการทำให้การทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบมีประสิทธิภาพสูงสุด คือ รูปแบบของการทดสอบ โมเดลการตอบสนองข้อสอบ ขนาดของคลังข้อสอบ ระดับความยากง่ายของการเสนอข้อสอบข้อแรก เกณฑ์การ คัดเลือกข้อสอบ และวิธีการประมาณค่าความสามารถของผู้สอบ ส่วนประเด็นที่ยังไม่ค่อยชัดเจนคือ อัตราการใช้ข้อสอบซ้ำ ซึ่งหลักการของการทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบด้วยคอมพิวเตอร์เป็นการคัดเลือกข้อสอบจากคลังข้อสอบให้เหมาะสมกับความความสามารถของผู้สอบ โดยที่ ผู้สอบแต่ละบุคคลจะได้รับข้อสอบที่แตกต่างกันออกไป แต่ในทางปฏิบัติแล้วพบว่าข้อสอบภายในคลังข้อสอบบางข้อถูกใช้บ่อย ขณะที่บางข้อไม่ได้ใช้ ข้อสอบที่ถูกใช้บ่อยหรือมีอัตราการใช้ข้อสอบซ้ำสูงจึงเป็นสิ่งที่ไม่พึงปรารถนา ถ้าผู้สอบทำการทดสอบหลายๆ ครั้งด้วยการ

ทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์โดยใช้คลังข้อสอบเดียวกัน อาจทำให้ผู้สอบถูกถามด้วยคำถามเหมือนๆ กัน จะทำให้เกิดการเรียนรู้จากการตอบหรือเกิดการเรียนรู้จากการที่ผู้สอบที่ทำการทดสอบแล้วนำไปพูดคุยระหว่างเพื่อนที่มีความสามารถใกล้เคียงกันซึ่งยังไม่ได้ทำการทดสอบ ทำให้ไม่ได้วัดคุณลักษณะภายในของผู้สอบอย่างแท้จริง เป็นสาเหตุให้ความตรงของการทดสอบลดลง (Revuelta & Ponsoda, 1998) ดังนั้นในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ในการทดสอบครั้งต่อไปนั้นจะต้องมีการควบคุมอัตราการใช้ข้อสอบซ้ำ เพื่อป้องกันผู้สอบที่สามารถจดจำข้อสอบได้นอกจากนี้ สต็อกกิ้งและเลวิส (Stocking & Lewis, 1998) กล่าวว่า ระดับความสามารถของผู้สอบจะมีความสัมพันธ์กับอัตราการใช้ข้อสอบซ้ำด้วย ถ้าระดับความสามารถของผู้สอบสูงหรือต่ำกว่าค่าเฉลี่ยความยากของข้อสอบ ซึ่งจะส่งผลต่อความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ นั่นคือ ระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำที่แตกต่างกัน จะทำให้การประมาณค่าความสามารถของผู้สอบมีความตรงตามสภาพ และความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกันหรือไม่ เพียงใด และจากการศึกษาวิจัยที่เกี่ยวข้องยังไม่พบว่ามีผู้ใดศึกษาเรื่องนี้มาก่อน ดังนั้นผู้วิจัยจึงมีความสนใจที่จะเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน ด้วยโปรแกรมไมโครแคท (MicroCAT)

คำถามวิจัย

การวิจัยครั้งนี้ มีคำถามวิจัยดังนี้

1. ระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำที่แตกต่างกัน จะทำให้การประมาณค่าความสามารถของผู้สอบมีความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกันหรือไม่ เพียงใด
2. ระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำที่แตกต่างกัน จะทำให้การประมาณค่าความสามารถของผู้สอบมีความคงที่ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกันหรือไม่ เพียงใด

วัตถุประสงค์การวิจัย

1. เพื่อเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบ และ อัตราการใช้ข้อสอบซ้ำแตกต่างกัน
2. เพื่อเปรียบเทียบความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบ และ อัตราการใช้ข้อสอบซ้ำแตกต่างกัน

สมมุติฐานการวิจัย

จากหลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ มีเป้าหมายที่จะดำเนินการทดสอบด้วยการคัดเลือกข้อสอบให้เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล ดังนั้นผู้สอบที่มีระดับความสามารถสูงก็จะได้รับข้อสอบที่ยาก ส่วนผู้สอบที่มีระดับความสามารถต่ำก็จะได้รับข้อสอบที่ง่าย สต็อกกิงและเลวิส (Stocking and Lewis, 1998) กล่าวว่า ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้นจะต้องมีการควบคุมอัตราการใช้ข้อสอบซ้ำ เพื่อป้องกันผู้สอบที่สามารถจำข้อสอบได้ ทำให้การประมาณค่าความสามารถของผู้สอบมีความถูกต้องแม่นยำ โดยที่การทดสอบแต่ละครั้งควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 20 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา นอกจากนี้ คาลอห์น และสเปรย์ (Kalohn and Spray, 1998) พบว่าผู้สอบที่มีระดับความสามารถสูงควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 10 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา ดังนั้นผู้วิจัยจึงตั้งสมมุติฐานการวิจัยดังต่อไปนี้

1. ระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน น่าจะส่งผลให้ความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน โดยผู้สอบที่มีระดับความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำสูงขึ้นไป น่าจะส่งผลให้ความตรงตามสภาพของการทดสอบลดลง
2. ระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน น่าจะส่งผลให้ความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถ

ของผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน โดยผู้สอบที่มีระดับความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำสูงขึ้น น่าจะส่งผลให้ความคงที่ในการประมาณค่าความสามารถของผู้สอบลดลง

ขอบเขตของการวิจัย

1. แบบสอบวัดความสามารถทางภาษาอังกฤษ ระดับชั้นมัธยมศึกษาตอนต้น สำหรับการวิจัยนี้เป็นแบบวัดผลสัมฤทธิ์ทางการเรียน ซึ่งพัฒนาโดยกรมวิชาการ กระทรวงศึกษาธิการ (2540) เพื่อใช้ในการประเมินคุณภาพทางการศึกษา ลักษณะของแบบสอบเป็นข้อสอบเลือกตอบ 4 ตัวเลือก จำนวน 50 ข้อ

2. คลังข้อสอบ เป็นกลุ่มข้อสอบที่ใช้วัดความรู้ความเข้าใจในวิชาภาษาอังกฤษ ระดับชั้นมัธยมศึกษาตอนต้น ในโครงการวิจัยเรื่องการพัฒนาคลังข้อสอบรายวิชาตามหลักสูตรมัธยมศึกษา ซึ่งพัฒนาโดยแมนมาศ ลีลสัตยกุล (2540) ลักษณะของแบบสอบเป็นข้อสอบ เลือกตอบ 4 ตัวเลือก จำนวน 231 ข้อ มีค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.400 ถึง 2.500 ค่าความยาก (b) อยู่ระหว่าง -3.000 ถึง 3.000 และค่าการเดา (c) อยู่ระหว่าง 0.020 ถึง 0.370

3. โปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ คือ โปรแกรมการทดสอบไมโครแคท (MicroCAT) รุ่นที่ 3 (Version 3) พัฒนาโดย Assessment System Corporation (1995) กำหนดรูปแบบของการทดสอบเป็นหลายขั้นตอน (multi - stage strategies) แบบทางแยกแปรผัน (variable branching model) โดยเริ่มต้นการทดสอบจากการ คัดเลือกข้อสอบที่มีค่าความยากระดับปานกลาง ประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบย์ส์ (Bayesian) การคัดเลือกข้อสอบข้อต่อไป โดยดูค่าสารสนเทศของข้อสอบข้อที่สูงที่สุด (maximum information) และยุติการทดสอบโดยพิจารณาความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่า 0.3

4. ตัวแปรที่ศึกษาในครั้งนี้ประกอบด้วย

4.1 ตัวแปรอิสระ ได้แก่

4.1.1 ระดับความสามารถทางภาษาอังกฤษของผู้สอบ จำแนกเป็น 3 ระดับดังนี้

- 1) ความสามารถต่ำ ($\theta \leq -0.75$)
- 2) ความสามารถปานกลาง ($-0.75 < \theta < 0.75$)
- 3) ความสามารถสูง ($\theta \geq 0.75$)

4.1.2 อัตราการใช้ข้อสอบซ้ำ กำหนดไว้เป็น 3 ระดับดังนี้

- 1) อัตราการใช้ข้อสอบซ้ำ เท่ากับ 10 เปอร์เซ็นต์
- 2) อัตราการใช้ข้อสอบซ้ำ เท่ากับ 20 เปอร์เซ็นต์
- 3) อัตราการใช้ข้อสอบซ้ำ เท่ากับ 30 เปอร์เซ็นต์

4.2 ตัวแปรตาม ได้แก่

4.2.1 ความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบ

4.2.2 ความคงที่ในการประมาณค่าความสามารถของผู้สอบ

ข้อตกลงเบื้องต้น

ความเข้าใจในวิธีการทดสอบด้วยเครื่องคอมพิวเตอร์ของผู้สอบทุกคน สามารถควบคุมได้ด้วยคำแนะนำในการฝึกทำข้อสอบที่ปรากฏบนจอภาพก่อนเริ่มทำการทดสอบ

คำจำกัดความที่ใช้ในการวิจัย

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ หมายถึง การทดสอบความรู้ทางภาษาอังกฤษ ด้วยข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบและดำเนินการสอบโดยใช้โปรแกรมการทดสอบไมโครแคท (MicroCAT) รุ่นที่ 3 (Version 3) ซึ่งพัฒนาโดย Assessment System Corporation (1995)

ความสามารถทางภาษาอังกฤษ หมายถึง คะแนนที่ได้จากการทดสอบด้วยแบบสอบวัดความสามารถทางภาษาอังกฤษ ระดับชั้นมัธยมศึกษาตอนต้น ซึ่งพัฒนาโดยกรมวิชาการ กระทรวง ศึกษาธิการ (2540) ลักษณะของแบบสอบเป็นข้อสอบเลือกตอบ 4 ตัวเลือก จำนวน 50 ข้อ

ความตรงตามสภาพ หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถที่ได้จากการทดสอบด้วยแบบสอบวัดความสามารถทางภาษาอังกฤษกับค่าความสามารถเฉลี่ยที่ได้จากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

ความคงที่ในการประมาณค่าความสามารถของผู้สอบ หมายถึง ความสอดคล้องของ ค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 5

ครั้ง ซึ่งคำนวณได้จากกราฟวิเคราะห์ความแปรปรวนแบบวัดซ้ำของการประมาณค่าความสามารถของผู้สอบ

อัตราการใช้ข้อสอบซ้ำ หมายถึง สัดส่วนระหว่างจำนวนข้อสอบที่ใช้ซ้ำในการทดสอบแบบปรับ
 เหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ กับจำนวนข้อสอบทั้งหมดที่ผู้สอบได้รับในการทดสอบครั้งที่
 ผ่านมา ซึ่งการวิจัยครั้งนี้กำหนดอัตราการใช้ข้อสอบซ้ำ 3 ระดับ คืออัตราการใช้ข้อสอบซ้ำ เท่ากับ 10 20 และ
 30 เปอร์เซ็นต์

ค่าความสามารถของผู้สอบ (θ) หมายถึง ระดับความสามารถของผู้สอบแต่ละบุคคล ซึ่งวัดใน
 หน่วย θ ที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 สำหรับค่าความสามารถมีค่าตั้งแต่ $-\infty$ ถึง
 $+\infty$ ในการศึกษาครั้งนี้ผู้วิจัยแบ่งระดับความสามารถออกเป็น 3 ระดับ ดังนี้ ความสามารถต่ำ ($\theta \leq -$
 0.75) ความสามารถปานกลาง ($-0.75 < \theta < 0.75$) และ ความสามารถสูง ($\theta \geq 0.75$)

คลังข้อสอบ หมายถึง กลุ่มของข้อสอบซึ่งสร้างขึ้นตามหลักเกณฑ์ที่ถูกต้องและถูกเก็บไว้
 เป็นระบบ คลังข้อสอบที่ใช้ในการวิจัยครั้งนี้เป็นกลุ่มข้อสอบที่ใช้วัดความรู้ความเข้าใจในวิชาภาษาอังกฤษ
 ระดับชั้นมัธยมศึกษาตอนต้น ลักษณะของแบบสอบเป็นข้อสอบเลือกตอบ 4 ตัวเลือก จำนวน 231 ข้อ มีค่า
 อำนาจจำแนก (a) อยู่ระหว่าง 0.400 ถึง 2.500 ค่าความยาก (b) อยู่ระหว่าง -3.000 ถึง 3.000 และค่าการ
 เดลา (c) อยู่ระหว่าง 0.020 ถึง 0.370

โปรแกรมการทดสอบไมโครแคท (MicroCAT) หมายถึง ระบบการทดสอบด้วยคอมพิวเตอร์ที่มี
 ความสามารถของการพัฒนากระบวนการทดสอบ ตั้งแต่การสร้างข้อสอบและจัดเก็บ ข้อสอบเป็นคลังข้อสอบ
 สามารถจัดชุดแบบสอบจากข้อสอบที่มีอยู่ในคลังข้อสอบ ตลอดจนการทดสอบและประมวลผล โดยอาศัย
 คอมพิวเตอร์ที่จะทำให้เกิดกระบวนการของระบบ พัฒนาโดย Assessment Systems Corporation (1995)

ข้อจำกัดของการวิจัย

ค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบมีการกระจายเป็นช่วงๆ ด้วยจำนวนข้อสอบที่ แตกต่างกัน
 จึงทำให้การคัดเลือกข้อสอบจากคลังข้อสอบในบางครั้งไม่ได้ข้อสอบตามที่ต้องการและ จำเป็นต้องใช้ข้อสอบที่
 มีคุณลักษณะใกล้เคียงขึ้นมาแทน

ประโยชน์ที่คาดว่าจะได้รับ

ผลการวิจัยครั้งนี้ ก่อให้เกิดประโยชน์ดังนี้

1. เป็นการขยายองค์ความรู้ในเชิงทฤษฎีเกี่ยวกับความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์
2. เพื่อให้ให้นักวัดผลการศึกษา ครู และอาจารย์ได้ข้อมูลสำหรับเป็นแนวทางในการเลือกใช้รูปแบบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่มีประสิทธิภาพ นั่นคือ การทดสอบต้องมีความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบ เนื่องจากผลของการวิจัยในครั้งนี้จะได้ข้อค้นพบว่าตัวแปรใดบ้างที่เกี่ยวข้องกับความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์
3. ได้เผยแพร่ความรู้เกี่ยวกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ให้แพร่หลายออกไปมากกว่าเดิม เนื่องจากในปัจจุบันคอมพิวเตอร์เข้ามามีบทบาทต่อการจัดการเรียนการสอนในสถานศึกษาระดับต่างๆ มากขึ้น ดังนั้นผู้ที่ทำหน้าที่เกี่ยวข้องกับการจัดการศึกษา โดยเฉพาะผู้ที่มีหน้าที่โดยตรงเกี่ยวกับการวัดผลการศึกษาควรที่จะมีความรู้ความเข้าใจในการทดสอบลักษณะนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาเอกสาร งานวิจัย และตำราต่างๆ ในเรื่องที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะ กับความสามารถของผู้สอบด้วยคอมพิวเตอร์หรือการทดสอบแบบซี เอ ที (CAT) ซึ่งผู้วิจัยได้ นำเสนอเนื้อหา ออกเป็น 4 ตอน ดังนี้

- ตอนที่ 1 ทฤษฎีการตอบสนองข้อสอบ
- ตอนที่ 2 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์
- ตอนที่ 3 โปรแกรมการทดสอบไมโครแคท (MicroCAT)
- ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ในแต่ละตอนมีรายละเอียดดังต่อไปนี้

ตอนที่ 1 ทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ เป็นทฤษฎีที่อธิบายความสัมพันธ์ระหว่างลักษณะหรือความสามารถ ที่มีอยู่ภายในตัวบุคคลกับพฤติกรรมการตอบสนองข้อสอบของบุคคลนั้น ทฤษฎีนี้มีความเชื่อว่าพฤติกรรม การตอบสนองข้อสอบของผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงจะถูกกำหนดโดยคุณลักษณะ (trait) หรือ ความสามารถ (ability) ที่มีอยู่ในตัวบุคคลซึ่งไม่สามารถสังเกตได้ และยังเชื่อว่าค่าพารามิเตอร์ต่างๆ ของ ข้อสอบไม่ว่าจะเป็นค่าความยาก (b) ค่าอำนาจจำแนก (a) หรือค่าการเดา (c) ของข้อสอบแต่ละข้อเป็น คุณลักษณะที่มีอยู่ประจำ และคงที่พอสมควรในตัวข้อสอบนั้นจริง ฉะนั้นค่าพารามิเตอร์ของข้อสอบจึงไม่แปร เปลี่ยนไปตามกลุ่มตัวอย่าง และค่าความสามารถของผู้สอบซึ่งเชื่อว่าเป็นคุณลักษณะที่มีอยู่ในตัวผู้สอบนั้นจริง จึงไม่ควรแปรเปลี่ยนไปตามค่าความยากของข้อสอบ เนื่องจากความสามารถของผู้สอบเป็นคุณลักษณะซึ่งไม่ สามารถวัดหรือสังเกตได้โดยตรง นักวัดผลทางการศึกษาจึงได้พยายามหาความสัมพันธ์ระหว่างปริมาณการ กระทำข้อสอบ หรือคะแนน (test performance or score) กับปริมาณความสามารถ (ability) โดยแสดง ความสัมพันธ์ดังกล่าวในลักษณะโมเดลทางคณิตศาสตร์ที่กำหนดด้วยฟังก์ชันทางคณิตศาสตร์ เรียกว่า ฟังก์ชัน

ลักษณะข้อสอบ (Item Characteristic Function) หรือเรียกว่าโค้งคุณลักษณะของข้อสอบ (Item Characteristic Curve) หรือฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) (ศิริชัย กาญจนวาสี, 2538) ซึ่งมีการพัฒนาขึ้นหลายโมเดลด้วยกัน เช่น นอร์มัล ออจีฟ โมเดล (Normal Ogive Model) โลจิสติก โมเดล (Logistic Model) และราสช์ โมเดล (Rasch Model) เป็นต้น แต่ละโมเดลนี้จะมีจำนวนพารามิเตอร์ของข้อสอบแตกต่างกันไป กล่าวคือ แบบพารามิเตอร์หนึ่งตัวจะแสดงเฉพาะ พารามิเตอร์ค่าความยาก (difficulty) เท่านั้น แบบพารามิเตอร์สองตัวจะเพิ่มพารามิเตอร์ค่าอำนาจจำแนก (discrimination) เข้าไปอีกตัวหนึ่ง และแบบพารามิเตอร์สามตัวจะมีทั้งพารามิเตอร์ค่าความยาก ค่าอำนาจจำแนก และค่าการเดา สำหรับการเลือกโมเดลของทฤษฎีการตอบข้อสอบมาใช้นั้น จะต้องเลือกให้เหมาะสมกับสถานการณ์ต่างๆ ตามข้อตกลงเบื้องต้นของแต่ละโมเดล โดยทฤษฎีมีจุดเด่นหลายประการ ดังนี้ (Hambleton and Swaminathan, 1985)

ประการแรก ค่าพารามิเตอร์ของข้อสอบเป็นอิสระจากกลุ่มผู้สอบที่ใช้ในการประมาณค่า กล่าวคือ ค่าพารามิเตอร์ของข้อสอบแต่ละข้อ ได้แก่ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c) เป็นค่าคงที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ถึงแม้ว่าผู้สอบต่างกลุ่มกัน ค่าพารามิเตอร์ของข้อสอบจะมีค่าคงเดิมเสมอ และในทำนองเดียวกันค่าพารามิเตอร์ความสามารถของผู้สอบเป็นอิสระจากข้อสอบที่ใช้ในการประมาณค่า จะไม่แปรเปลี่ยนตามค่าพารามิเตอร์ของข้อสอบ ไม่ว่าผู้สอบจะทำข้อสอบข้อใด หรือได้ทำข้อสอบเมื่อใดก็ตาม

ประการที่สอง การเปรียบเทียบความสามารถของผู้สอบ จะไม่ขึ้นอยู่กับคำถามในแบบสอบ ถึงแม้ว่าผู้สอบจะใช้ข้อคำถามต่างกัน ก็สามารถนำค่าความสามารถมาเปรียบเทียบกันได้ ทั้งนี้เนื่องจากค่าความสามารถที่ประมาณค่ามาได้ นั้น เป็นคะแนนโลจิท (Logit) ซึ่งอยู่ในมาตรวัดเดียวกัน

ประการที่สาม การรายงานคุณภาพของข้อสอบในรูปค่าอินฟอร์เมชัน (information) สามารถรายงานได้ทั้งเป็นรายข้อและทั้งฉบับ ค่าอินฟอร์เมชันเป็นตัวบ่งชี้ถึงความถูกต้องแม่นยำ (accuracy) ในการประมาณค่าความสามารถ สามารถนำมาใช้แทนค่าความเที่ยง และค่าความคลาดเคลื่อนมาตรฐานในการวัดได้

จากคุณลักษณะดังกล่าว จึงทำให้นักวัดผลนำทฤษฎีการตอบสนองข้อสอบไปประยุกต์ใช้ในการวัดผลทางการศึกษาและจิตวิทยาในหลายๆ เรื่องด้วยกัน เช่น การสร้างคลังข้อสอบ (Item banking) การพัฒนาแบบทดสอบ (Test Development) การหาความลำเอียงของข้อสอบ (Item Bias) การทดสอบอิงเกณฑ์ (Criterion – Reference Testing) การเทียบคะแนนการสอบ (Test score Equating) และการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive testing)

หลักการของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ เป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างคุณลักษณะภายในหรือความสามารถที่มีอยู่ในตัวบุคคลกับพฤติกรรมกรรมการตอบสนองข้อสอบของบุคคลนั้นว่ามีโอกาสตอบข้อสอบถูกมากน้อยเพียงไร ทฤษฎีนี้มีพื้นฐานความเชื่อว่า พฤติกรรมการตอบสนองต่อข้อสอบของผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่อยู่ภายในตัวบุคคลซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง (ศิริชัย กาญจนวาสี, 2542) ดังนั้นทฤษฎีจึงอธิบายความสัมพันธ์ระหว่างปริมาณของการกระทำหรือคะแนนของผู้สอบกับระดับความสามารถของผู้สอบ ซึ่งแสดงได้ดังสมการต่อไปนี้ (สุพัฒน์ สุขมลสันต์, 2539)

$$P = f(\theta)$$

| | | | |
|-------|----------|-----|---------------------------------|
| เมื่อ | P | แทน | ผลการทดสอบ (performance) |
| | θ | แทน | ความสามารถ (ability หรือ trait) |
| | f | แทน | ความสัมพันธ์ (function) |

จากความสัมพันธ์ในสมการดังกล่าวเป็นการแสดงความสัมพันธ์ที่ทุกๆ ไปยังไม่เฉพาะเจาะจง จึงมีผู้สนใจศึกษาและพัฒนาเพื่อให้ได้สมการที่สามารถอธิบายลักษณะเฉพาะของข้อสอบแต่ละข้อที่สามารถอธิบายความสัมพันธ์ของสองตัวแปรนี้ได้เด่นชัดขึ้น โดยอาศัยข้อตกลงเบื้องต้นบางประการและวิธีการทางคณิตศาสตร์

ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (IRT) มีข้อตกลงเบื้องต้น ดังต่อไปนี้ (Hambleton and Swaminathan, 1985)

1. ความเป็นมิติเดียว (Unidimension) คุณลักษณะภายในหรือความสามารถของผู้สอบมีอยู่หลายประการ ซึ่งคุณลักษณะภายในหรือความสามารถแต่ละอย่างสามารถกำหนดพฤติกรรมการตอบสนองข้อสอบ ถ้าผลการตอบสนองข้อสอบหรือคะแนนของผู้สอบสามารถอธิบายได้คุณลักษณะเดียวกันก็ถือว่ามีความเป็นมิติเดียว โมเดลของการตอบสนองข้อสอบมีข้อตกลงว่า ข้อสอบแต่ละข้อในแบบสอบนั้นวัดความสามารถ

หรือคุณลักษณะเดียวกัน แต่ข้อตกลงข้อนี้ไม่เข้มงวดนัก ถ้ามีลักษณะเด่นที่จะวัดองค์ประกอบใดองค์ประกอบหนึ่งก็ถือว่าเป็นมิติเดียวเช่นกัน

วิธีการตรวจสอบว่าแบบสอบนั้นวัดในมิติเดียวหรือไม่นั้น มีวิธีการทดสอบได้หลายอย่างแต่ที่นิยมมากมี 2 วิธี คือ (Warm, 1978 อ้างถึงใน สุพัชญ์ สุขมลสันต์, 2539)

1.1 โดยการหาค่าความสัมพันธ์ระหว่างค่าน้ำหนักองค์ประกอบรายข้อ (factor loading) ขององค์ประกอบที่หนึ่งกับค่าสหสัมพันธ์แบบไบซีเรียล (biserial correlation) ของข้อสอบรายข้อกับคะแนนรวม หากค่าสหสัมพันธ์มีค่ามากกว่า .80 ก็แสดงว่าแบบสอบนั้นวัดมิติเดียว

1.2 โดยการวิเคราะห์องค์ประกอบ (factor analysis) ของข้อสอบทั้งฉบับ ถ้าค่าไอเกน (eigen value) ค่าสูงสุดแตกต่างจากค่าอื่นอย่างชัดเจนหรือไม่ ถ้าแตกต่างก็แสดงว่าแบบสอบนั้นวัดมิติเดียว

2. ความเป็นอิสระในการตอบข้อสอบ (Local Independence) หมายถึง การตอบ ข้อสอบข้อต่างๆ ในแบบสอบของผู้สอบนั้นมีความเป็นอิสระในเชิงสถิติ กล่าวคือ การตอบข้อสอบข้อหนึ่งๆ จะไม่มีผลกระทบต่อคำตอบข้อสอบข้ออื่นๆ ในแบบสอบ เนื้อหาของข้อสอบข้อหนึ่งๆ จะต้องไม่มีเงื่อนไขในการตอบข้อสอบข้ออื่นๆ แมคโดแนล (McDonald, 1980a, 1980b, 1982 อ้างถึงใน Hambleton and Swaminathan, 1985) ได้นิยามของความเป็นมิติเดียวของข้อสอบว่าควรอยู่บนพื้นฐานของหลักการความเป็นอิสระในการตอบข้อสอบ

3. โค้งลักษณะของข้อสอบ (Item Characteristic Curves : ICC) เป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงถึงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบนั้นได้ถูกต้องกับระดับความสามารถที่วัดได้โดยชุดของข้อสอบหรือแบบสอบ จะเห็นได้ว่าโอกาสที่ผู้สอบจะตอบข้อสอบถูก (probability) จะขึ้นอยู่กับโค้งลักษณะของข้อสอบ (ICC) ซึ่งเป็นอิสระจากการกระจายของความสามารถของผู้สอบ นั่นคือ โอกาสที่ผู้สอบจะตอบข้อสอบถูกไม่ขึ้นอยู่กับจำนวนของผู้สอบที่มีความสามารถเหมือนกัน รูปร่างของโค้งลักษณะของข้อสอบ (ICC) ในแต่ละข้อมีคุณสมบัติไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ดังนั้นจึงทำให้โอกาสในการตอบข้อสอบถูกในแต่ละข้อไม่แปรเปลี่ยน

นอกจากนี้แล้ว ถ้าข้อตกลงเบื้องต้นของการทดสอบตรงตามที่ทฤษฎีกำหนดไว้ ผลที่ได้จากการคำนวณจะมีลักษณะที่ดีดังนี้

1. ค่าพารามิเตอร์ของข้อสอบ คือ ค่าความยาก (a) ค่าอำนาจจำแนก (b) และค่าการเดา (c) จะเป็นค่าที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ไม่ว่าจะนำไปสอบกับผู้ใดก็ตาม ลักษณะเช่นนี้เรียกว่าความเป็นอิสระจากกลุ่มตัวอย่าง (person free test calibration)

2. เมื่อทราบลักษณะการตอบข้อสอบแต่ละข้อของผู้สอบ เราสามารถจะคำนวณหาค่าความสามารถที่แท้จริงของผู้สอบคนนั้นได้ ค่าความสามารถที่แท้จริงนี้มีความสัมพันธ์กับโดยตรงกับคะแนนจริง การคำนวณความสามารถที่แท้จริงนั้นอาจใช้ข้อสอบข้อใดก็ได้ที่วัดสิ่งเดียวกัน ลักษณะเช่นนี้เรียกว่า ความเป็นอิสระของข้อสอบ (item free person measurement)

โค้งลักษณะของข้อสอบมีหลายรูปแบบตามความเชื่อเกี่ยวกับความสัมพันธ์ระหว่างระดับความสามารถกับโอกาสในการตอบข้อสอบได้ถูกต้อง จึงทำให้มีโมเดลของโค้งลักษณะของข้อสอบขึ้นหลายโมเดล ดังจะกล่าวต่อไปนี้

โมเดลของทฤษฎีการตอบสนองข้อสอบ

โค้งลักษณะของข้อสอบแสดงความสัมพันธ์ระหว่างระดับความสามารถกับโอกาสในการตอบข้อสอบได้ถูกต้อง ทำให้เกิดโมเดลของโค้งลักษณะของข้อสอบขึ้นหลายโมเดล แต่รูปแบบที่นิยมใช้กันแพร่หลายคือ รูปแบบโลจิสติก (Logistic Model) ซึ่งนิยมใช้กันทั่วไป 3 รูปแบบ ดังนี้ (Hambleton and Swaminathan, 1985)

รูปแบบโลจิสติกที่มี 1 พารามิเตอร์ (One – Parameter Logistic Model)

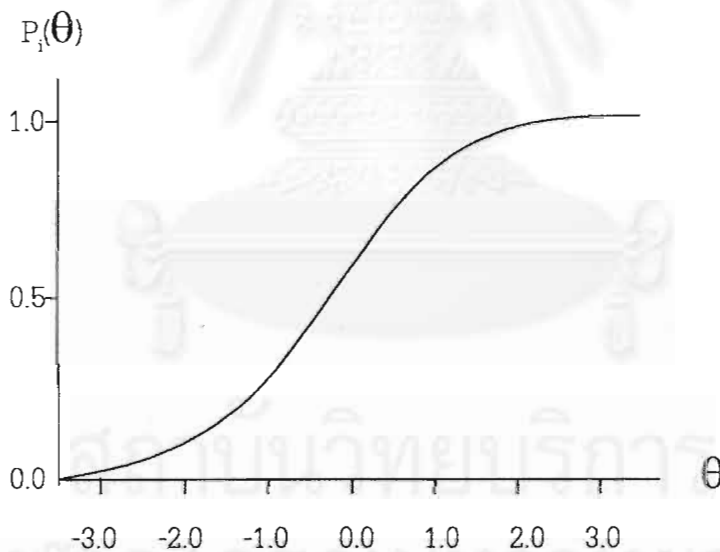
ราล์ฟได้พัฒนาทฤษฎีการตอบสนองข้อสอบ และเสนอการวิเคราะห์ข้อสอบแบบ 1 พารามิเตอร์ในปี ค.ศ.1960 โดยฟังก์ชันของแบบ 1 พารามิเตอร์นี้สามารถอธิบายได้ด้วยพารามิเตอร์ของข้อสอบเพียงตัวเดียวคือ ค่าความยาก (b) เขียนเป็นสมการได้ดังนี้

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

เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นของผู้สอบที่มีความสามารถ (θ) จะตอบข้อสอบข้อที่ i ได้ถูกต้อง
 D แทน ค่าคงที่ของ scale factor เพื่อปรับค่าของฟังก์ชันโลจิสติกให้ใกล้เคียงกับค่าที่ได้จากฟังก์ชันรูปตัวเอสปกติ มีค่าประมาณ 1.7

- b_i แทน ระดับความยากของข้อสอบข้อที่ i ที่แสดงถึงระดับความสามารถที่จุดโค้งลักษณะข้อสอบมีความชันมากที่สุด ในทางปฏิบัติค่าความยากจะมีค่าอยู่ระหว่าง -2 ถึง $+2$
- θ แทน ระดับความสามารถที่แท้จริงที่คำนวณจากคะแนนรวม โดยการปรับให้เป็นคะแนนที่มีค่าเฉลี่ยเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 โดยมีพิสัยการกระจายเท่ากับ $-\alpha$ ถึง $+\alpha$ แต่ในทางปฏิบัติระดับความสามารถจะมีค่าอยู่ระหว่าง -3 ถึง $+3$
- e แทน ค่าคงที่ซึ่งมีค่าประมาณ 2.7182818

การวิเคราะห์ข้อสอบแบบ 1 พารามิเตอร์ (1 - parameter model) นี้มีข้อตกลงว่าข้อสอบทุกข้อไม่มีโอกาสของการเดาถูก ($c_i = 0$) และค่าอำนาจจำแนกของข้อสอบ (a_i) ทุกข้อเท่ากันหมด โค้งลักษณะเฉพาะของข้อสอบแสดงในแผนภาพที่ 2



แผนภาพที่ 2 โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 1 พารามิเตอร์

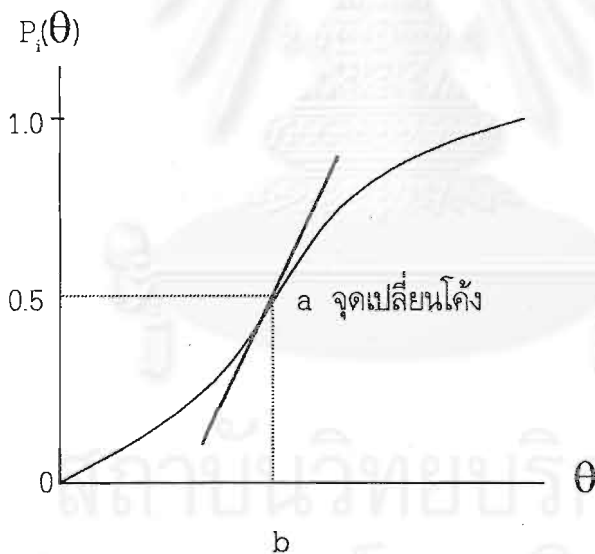
รูปแบบโลจิสติกที่มี 2 พารามิเตอร์ (Two - Parameter Logistic Model)

เบิร์นบาม (Birnbau, 1957, 1958a, 1958b, 1968) ได้เสนอการวิเคราะห์ข้อสอบแบบโลจิสติกที่มีพารามิเตอร์ 2 ตัว คือ ค่าความยาก (b) และค่าอำนาจจำแนก (a) และสามารถเขียนเป็นสมการได้ดังนี้

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

เมื่อ a_i แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ i

การวิเคราะห์ข้อสอบแบบ 2 พารามิเตอร์ (2 - parameter model) นี้มีข้อตกลงว่า ข้อสอบทุกข้อไม่มีโอกาสของการเดาถูก คือ ค่า $c_i = 0$ โคงลักษณะเฉพาะของข้อสอบแบบนี้มีลักษณะดังแผนภาพที่ 3



แผนภาพที่ 3 โคงลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 2 พารามิเตอร์

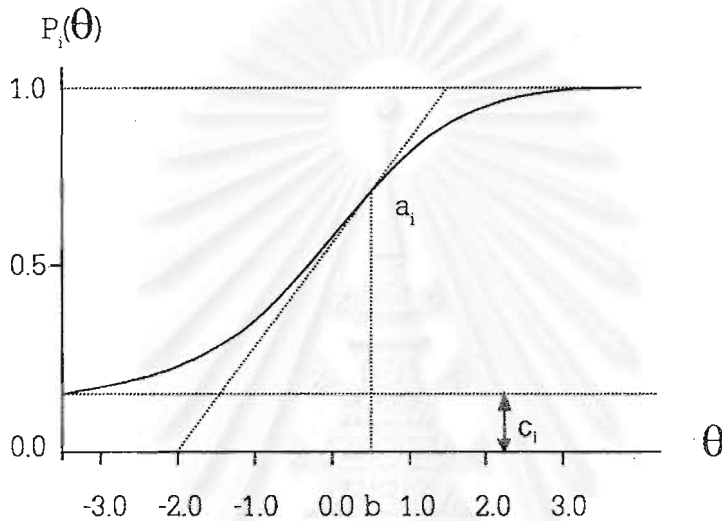
รูปแบบโลจิสติกที่มี 3 พารามิเตอร์ (Three - Parameter Logistic Model)

การวิเคราะห์ข้อสอบแบบ 3 พารามิเตอร์ (3 - parameter model) นี้ได้ดัดแปลงมาจากรูปแบบโลจิสติกที่มี 2 พารามิเตอร์ พารามิเตอร์ที่เพิ่มขึ้น คือ ค่าการเดา (c_i) และมีสมการดังนี้

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

เมื่อ c_i แทน ค่าการเดา

โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 3 พารามิเตอร์ มีลักษณะดัง
แผนภาพที่ 4



แผนภาพที่ 4 โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 3 พารามิเตอร์

ปัจจัยที่มีผลต่อคุณภาพของแบบสอบ

ตามแนวคิดของทฤษฎีการตอบสนองข้อสอบเชื่อว่าแบบสอบจะมีคุณภาพดีเพียงใดขึ้นอยู่กับคุณภาพของข้อสอบแต่ละข้อว่าสามารถทดสอบสิ่งที่ต้องการทดสอบได้มากน้อยเพียงใด หรือกล่าวอีกนัยหนึ่งก็คือ ข้อสอบสามารถทดสอบได้ถูกต้องแม่นยำมากน้อยเพียงใด ซึ่งขึ้นอยู่กับปัจจัยหลายประการ (สุพัฒน์ สุกมลสันต์, 2539)

ความถูกต้องแม่นยำ (accuracy)

ในการประมาณค่าความสามารถของผู้สอบ โดยปกติไม่ว่าจะเป็นการวัดหรือการประมาณค่าใดๆ ก็ตามจะต้องมีความคลาดเคลื่อนอยู่เสมอ โดยเฉพาะอย่างยิ่งในการวัดหรือการประมาณค่าความสามารถในทางการศึกษาและจิตวิทยาจะมีความคลาดเคลื่อนค่อนข้างสูง ถ้าความคลาดเคลื่อนในการประมาณค่ามีค่าสูง ความถูกต้องแม่นยำ (accuracy) ในการประมาณค่าก็จะต่ำ ซึ่งสามารถเขียน

เป็นสมการแสดงความสัมพันธ์ระหว่างความสามารถจริง (θ) ความสามารถที่ได้จากการประมาณค่า ($\hat{\theta}$) และค่าความคลาดเคลื่อน (E) ได้ดังนี้ (Lord and Novick, 1968)

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

จากสมการนี้จะเห็นว่าถ้าความคลาดเคลื่อนมีค่าน้อยหรือเข้าใกล้ศูนย์ความสามารถที่ได้จากการประมาณค่าก็จะเข้าใกล้ความสามารถจริง นั่นคือ มีความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบสูง

ความคลาดเคลื่อนมาตรฐานในการประมาณค่า (standard error of estimate : SEE)

ความคลาดเคลื่อนมาตรฐานในการประมาณค่า คือ ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในการประมาณค่า ซึ่งแตกต่างจากความคลาดเคลื่อนมาตรฐานในการวัด (standard error of measurement : SEM) กล่าวคือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าเป็นความคลาดเคลื่อนของค่าความสามารถของผู้สอบที่คลาดเคลื่อนไปจากความสามารถจริง ($E = \hat{\theta} - \theta$) ส่วนความคลาดเคลื่อนมาตรฐานในการวัด คือ ความคลาดเคลื่อน (E) ที่เกิดจากคะแนนดิบหรือคะแนนที่ได้จากการวัด (X) ที่คลาดเคลื่อนไปจากคะแนนจริง (T) นั่นคือ $E = X - T$

ความคลาดเคลื่อนมาตรฐานในการประมาณค่ามีความสัมพันธ์กับค่าฟังก์ชันสารสนเทศของแบบสอบ (test information function) คือ $SEE = 1 / \sqrt{I(\theta)}$ ค่าฟังก์ชันสารสนเทศของแบบสอบจะเป็นตัวบ่งชี้ถึงคุณภาพของการประมาณค่าและยังสามารถใช้แทนค่าความเที่ยงหรือ ค่าความคลาดเคลื่อนมาตรฐานในการวัด ในทฤษฎีการทดสอบแบบประเพณีนิยมได้ด้วย (Hambleton and Cook, 1977 อ้างถึงใน สุพัฒน์ สุขมลสันต์, 2539)

ฟังก์ชันสารสนเทศของข้อสอบ (item information function)

ฟังก์ชันสารสนเทศของข้อสอบ คือ ค่าที่แสดงถึงความถูกต้องแม่นยำในการประมาณค่าความสามารถจริง (θ_i) ของผู้สอบในการตอบข้อสอบแต่ละข้อ ซึ่งสามารถคำนวณได้จากสูตรต่อไปนี้

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

เมื่อ $I(\theta, u_i)$ แทน ฟังก์ชันสารสนเทศของข้อสอบ

$P_i'(\theta)$ แทน ความชันของโค้งลักษณะเฉพาะของข้อสอบที่ระดับความสามารถ θ

$$P_i(\theta) \text{ แทน ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ } \theta \text{ จะตอบข้อสอบ}$$

$$\text{ข้อที่ } i \text{ ได้ถูก}$$

$$Q_i(\theta) \text{ แทน } 1 - P_i(\theta)$$

จากสมการข้างต้น จะเห็นได้ว่า ฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อขึ้นอยู่กับความชันของโค้งลักษณะเฉพาะของข้อสอบ ถ้าโค้งลักษณะเฉพาะของข้อสอบชันมากขึ้นในขณะที่ความแปรปรวนของการตอบข้อสอบถูกน้อยลง โค้งสารสนเทศของข้อสอบที่ระดับความสามารถนั้นๆ จะยิ่งสูงขึ้น ความสูงของโค้งสารสนเทศของข้อสอบอยู่ที่ระดับความสามารถใดแสดงว่าสามารถจำแนกระดับความสามารถของผู้สอบได้ดี ณ ระดับความสามารถนั้น

ฟังก์ชันสารสนเทศของแบบสอบ (test information function)

ฟังก์ชันสารสนเทศของแบบสอบเป็นสัดส่วนกลับกันกับกำลังสองของความยาวของช่วงความเที่ยง ซึ่งเป็นผลมาจากการประมาณค่าความสามารถของผู้สอบจากการตอบข้อสอบทั้งหมด ค่านี้แสดงถึงความถูกต้องแม่นยำในการประมาณค่าความสามารถจริง (θ) ของผู้สอบแบบสอบทั้งหมดว่ามีมากน้อยเพียงใด ดังนั้นโค้งสารสนเทศของแบบสอบจึงเป็นตัวบ่งชี้ถึงความถูกต้องแม่นยำของค่าความสามารถที่ประมาณได้ ซึ่งแสดงในรูปสมการได้ดังนี้

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

เมื่อ $I(\theta)$ แทน ฟังก์ชันสารสนเทศของแบบสอบ

$I(\theta, u_i)$ แทน ฟังก์ชันสารสนเทศของข้อสอบ

ถ้าเรามีกลุ่มข้อสอบที่ทราบค่าสารสนเทศของข้อสอบ เราสามารถสร้างแบบสอบให้มี โค้งสารสนเทศของแบบสอบ ณ ความสามารถระดับใดระดับหนึ่งที่เราต้องการ เช่นการสร้างแบบสอบเพื่อคัดเลือคนักเรียนเข้าศึกษาต่อ ก็ต้องสร้างแบบสอบที่มีประสิทธิภาพสูงสุดที่ระดับความสามารถสูงๆ นั่นคือ ให้มีโค้งสารสนเทศของแบบสอบสูง ณ ระดับความสามารถสูงๆ

ตอนที่ 2 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

ความหมายของการทดสอบแบบปรับเหมาะ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing) หมายถึง การทดสอบที่ใช้แบบสอบต่างชุดกันสำหรับผู้สอบต่างกัน โดยมีการคัดเลือกข้อสอบที่มีความยาก เหมาะสมกับระดับความสามารถของผู้สอบ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ มีความแตกต่างที่สำคัญจากการทดสอบแบบประเพณีนิยม ตรงที่การทดสอบแบบประเพณีนิยมเป็นการสร้างแบบสอบขึ้นมาชุดเดียวให้ครอบคลุมองค์ประกอบของคุณลักษณะหรือเนื้อหาที่ต้องการวัด มีค่าความยากของข้อสอบที่ หลากหลายโดยไม่มีการกำหนดสัดส่วนของค่าความยากที่แน่นอน โดยทั่วไปนิยมให้มีค่าความยาก (b) อยู่ระหว่าง 0.2 – 0.8 สำหรับนำไปใช้สอบกับผู้สอบกลุ่มใดก็ได้ ไม่ว่าจะผู้สอบจะมีความสามารถสูง ปานกลางหรือต่ำ แต่การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบเป็นการใช้ข้อสอบจากคลัง ข้อสอบมาสร้างเป็นแบบสอบ แบบสอบแต่ละชุดมีการออกแบบให้เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล (ศิริชัย กาญจนวาสี, 2538) ซึ่งการทดสอบดังกล่าวนี้มีชื่อและวิธีการแตกต่างกันออกไป เช่น adaptive testing , branched testing , individualized testing , response contingent testing , tailor testing computerized testing , programmed testing ,computerized automate testing , sequential item testing ชื่อแรกที่ใช้เรียกการทดสอบลักษณะนี้ คือ การทดสอบเทเลอร์ (tailor testing) ซึ่งวิลเลียม ดับบลิว เทอร์นบูล (William W. tumbull) เป็นผู้คิดขึ้นในปี 1951 (Lord, 1980)

หลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

หลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ เป็นการคัดเลือก ข้อสอบ สำหรับแต่ละบุคคลโดยอยู่บนพื้นฐานของผลการตอบข้อสอบข้อที่ผ่านมาของผู้สอบนั้น เมื่อ ผู้สอบทำข้อสอบข้อแรกซึ่งได้จากคลังข้อสอบแล้ว จะมีการวิเคราะห์ระดับความสามารถหรือประเมินความสามารถของผู้สอบเบื้องต้น เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีค่าความยากและอำนาจจำแนก เหมาะสมที่จะใช้วัดระดับความสามารถของผู้สอบและทำการประมาณค่าระดับความสามารถของผู้สอบใหม่ จากนั้นก็จะเลือกข้อสอบที่เหมาะสมข้อต่อไป โดยอาศัยหลักการที่ว่า ถ้าการทำข้อที่ผ่านมาถูก ข้อถัดไปจะยากขึ้น แต่ถ้าการทำข้อที่ผ่านมา

มาผิดข้อถัดไปจะง่ายลง กระบวนการนี้จะดำเนินต่อไปเรื่อยๆจนสามารถประมาณค่าระดับความสามารถของผู้สอบได้อย่างเชื่อถือได้ตามเกณฑ์ที่กำหนดไว้ การทดสอบก็จะยุติลง (ศิริชัย กาญจนวาสี, 2538)

ประเภทของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ไวส์ (Weiss, 1974) ได้แบ่งวิธีการคัดเลือกข้อสอบที่มีระดับความยากเหมาะสมกับความสามารถของผู้สอบเป็น 2 วิธี คือ

1. ยุทธวิธีสองขั้นตอน (Two – Stage Strategies)
2. ยุทธวิธีหลายขั้นตอน (Multi – Stage Strategies)
 - 2.1 รูปแบบแยกทางคงที่ (Fixed Branching Model)
 - 2.1.1 รูปแบบปิรามิด (Pyramidal Modal) ได้แก่
 - 1) รูปแบบปิรามิดขนาดขั้นคงที่ (Constant Step Size Pyramidal Model)
 - 2) รูปแบบปิรามิดขนาดขั้นแปรผัน (Variable Step Size Pyramidal Model)
 - 3) รูปแบบปิรามิดข้างตัด (Truncated Pyramidal Model)
 - 4) รูปแบบปิรามิดแบบมีหลายข้อในแต่ละขั้น (Multi – Item Pyramidal Model)
 - 5) รูปแบบปิรามิดแบบให้น้ำหนักตัวเลือกเพื่อแยกทาง (Differential Response Option Branching Pyramidal Model)
 - 2.1.2 รูปแบบเฟล็กซิเลเวล (Flexilevel Model)
 - 2.1.3 รูปแบบปรับระดับขั้น (Stradaptive Model)
 - 2.2 รูปแบบแยกทางแปรผัน (Variable Branching Model)
 - 2.2.1 ยุทธวิธีแบบเบย์ส์ (Bayesian Strategies)
 - 2.2.2 ยุทธวิธีความเป็นไปได้สูงสุด (Maximum likelihood estimation Strategies)

รายละเอียดของการทดสอบแบบต่างๆ พอสรุปได้ดังนี้

ยุทธวิธีสองขั้นตอน

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้ยุทธวิธีสองขั้นตอน โดยมีการดำเนินการสอบเป็น 2 ขั้นตอน ขั้นตอนแรกทำการทดสอบด้วยแบบสอบกำหนดทิศทาง (routing system) ซึ่งเป็นแบบสอบที่มีจำนวนข้อสอบประมาณ 10 ข้อ ประกอบด้วยข้อสอบที่มีค่าเฉลี่ย ความยากอยู่ระดับปานกลาง ผู้สอบทุกคนต้องตอบข้อสอบเหมือนกัน เพื่อกำหนดทิศทางของผู้สอบแต่ละคนว่าจะต้องทำการทดสอบในขั้นตอนที่สองด้วยแบบสอบที่มีความยากระดับใด หลังจากทราบทิศทางหรือระดับความสามารถเบื้องต้นของผู้สอบแล้วก็ทำการทดสอบในขั้นตอนที่สอง ด้วยแบบสอบวัดผล (Measurement Test) ซึ่งจะแบ่งเป็นแบบสอบชุดย่อย ๆ หลายชุดตามระดับความยาก โดยปกติจะมีชุดละประมาณ 20 – 30 ข้อ ผู้สอบแต่ละคนจะได้รับแบบสอบชุดที่สองไม่เหมือนกัน ผู้สอบที่ได้รับการประเมินจากการทดสอบในขั้นตอนแรกว่ามีความสามารถสูงก็จะได้แบบสอบชุดที่สองที่ยาก ผู้สอบที่มีความสามารถปานกลางจะได้รับแบบสอบชุดที่สองที่มีระดับความยากปานกลาง และผู้สอบที่มีความสามารถต่ำก็จะได้แบบสอบชุดที่สองที่ง่าย (Lord, 1971)

โครงสร้างของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้ยุทธวิธีสอง ขั้นตอน แสดงดังแผนภาพที่ 5

ยุทธวิธีหลายขั้นตอน

การทดสอบแบบหลายขั้นตอน เป็นการทดสอบมากกว่า 2 ขั้นตอน โดยมีการจัดโครงสร้างของขั้นตอนการทดสอบ มีการจัดเรียงข้อสอบตามความยาก ในแต่ละขั้นตอนของการทดสอบมีวิธีดำเนินการแตกต่างกัน หลักการโดยทั่วไปเริ่มต้นจากข้อสอบที่มีความยากปานกลาง ถ้าตอบถูกข้อต่อไปจะยากขึ้น แต่ถ้าตอบผิดข้อต่อไปจะง่ายลง การทดสอบจะดำเนินเช่นนี้ไปเรื่อยและจะยุติตามเกณฑ์ที่กำหนดไว้ การทดสอบแบบหลายขั้นตอนนี้พอจะจำแนกเป็นแบบต่างๆ กันได้ดังนี้ (ศิริชัย กาญจนวาสี, 2538)

1. แบบทางแยกคงที่

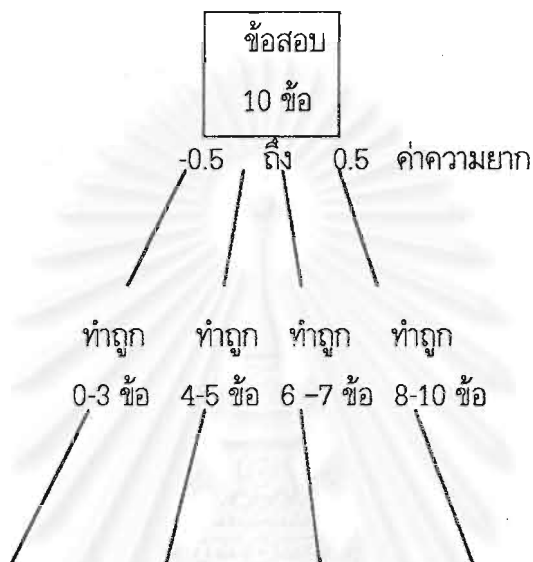
เป็นการทดสอบที่กำหนดทางแยกของการเลือกข้อสอบไว้คงที่ตายตัว มีรูปแบบของการดำเนินการทดสอบแบบทางแยกคงที่นี้หลายรูปแบบดังนี้

1.1 รูปแบบปิรามิด จำแนกเป็นหลายแบบ ได้แก่

1.1.1 รูปแบบปิรามิดขนาดขั้นคงที่ เป็นการจัดโครงสร้างข้อสอบตามลำดับ

ความยากเป็นรูปสามเหลี่ยมหรือเรียกว่าทางแยกขึ้นลง (up-and-down branching) โดยมีจำนวนข้อทดสอบในแต่ละชั้นเท่ากับลำดับที่ของชั้น กล่าวคือ ชั้นที่ 1 จะมีข้อสอบ 1 ข้อ ชั้นที่ 2 จะมีข้อสอบ 2 ข้อ ในชั้นที่ 6 จะมีข้อสอบ 6 ข้อ ดังแผนภาพที่ 6

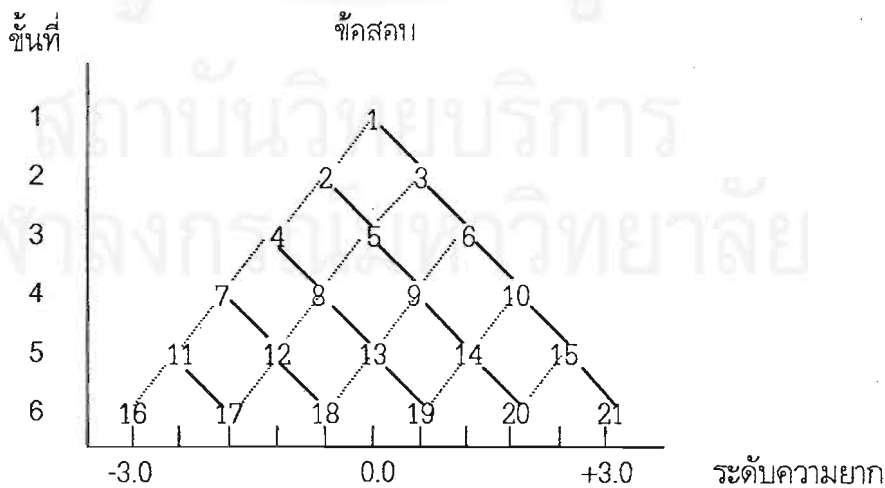
ขั้นตอนที่ 1



ขั้นตอนที่ 2



แผนภาพที่ 5 โครงสร้างของการทดสอบโดยใช้ยุทธวิธีสองขั้นตอน

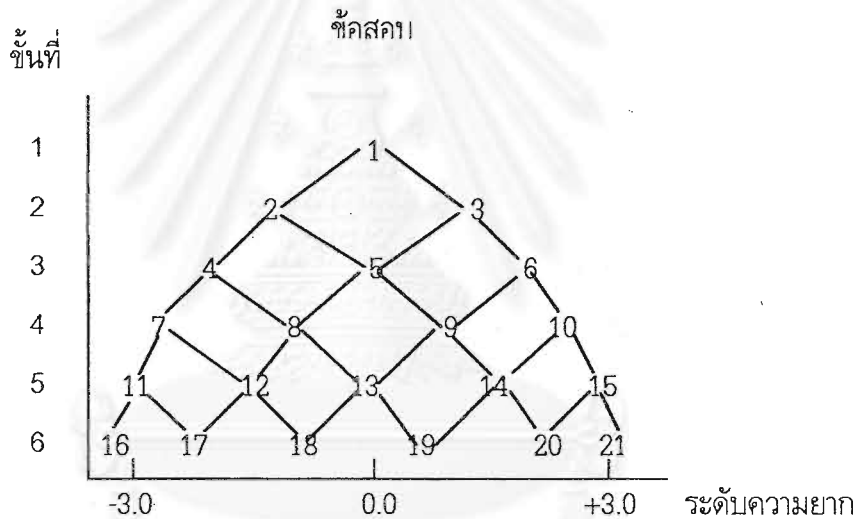


แผนภาพที่ 6 โครงสร้างของการทดสอบรูปแบบปิรามิดขนาดชั้นคงที่

ในการตอบแบบสอบ ผู้สอบจะต้องตอบข้อสอบชั้นละหนึ่งข้อ โดยเริ่มตอบข้อที่อยู่บนยอดสามเหลี่ยมซึ่งมีความยากปานกลางก่อน ถ้าตอบถูกชั้นต่อไปจะแยกไปตอบข้อที่ยากขึ้น แต่ถ้าตอบผิดชั้นต่อไปจะไปตอบข้อที่ง่ายกว่า กระบวนการทดสอบจะเป็นเช่นนี้จนถึงชั้นสุดท้าย

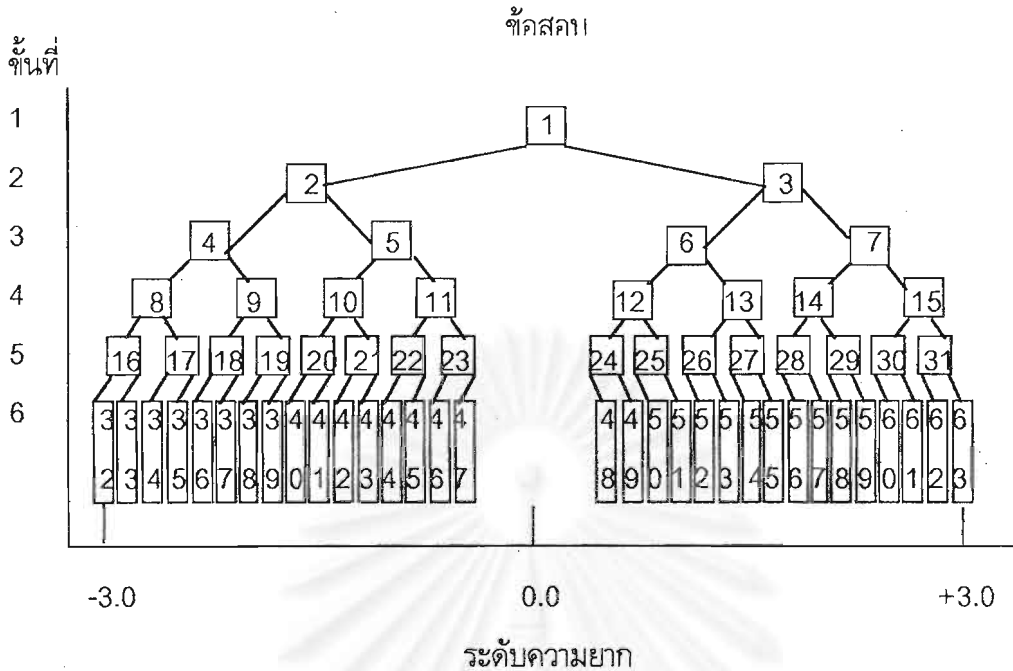
1.1.2 รูปแบบปิรามิดขนาดชั้นแปรผัน เป็นการจัดโครงสร้างข้อสอบตามลำดับความยาก โดยข้อสอบแต่ละข้อมีค่าความยากเพิ่มขึ้นหรือลดลงไม่คงที่ หรือเรียกว่า วิธี H - L โดยมีรูปแบบการปรับค่าความยากของข้อสอบเป็น 2 แบบดังนี้

1) H - L Method กำหนดช่วงห่างระดับความยากของข้อที่อยู่ติดกันภายในชั้นให้มีค่าไม่เท่ากัน โดยให้ข้อที่อยู่กลาง ๆ มีช่วงห่างของระดับความยากมาก แล้วค่อย ๆ ลดลงในข้อที่ค่อนข้างง่ายหรือยาก ดังแผนภาพที่ 7



แผนภาพที่ 7 โครงสร้างของการทดสอบรูปแบบปิรามิดขนาดชั้นแปรผัน

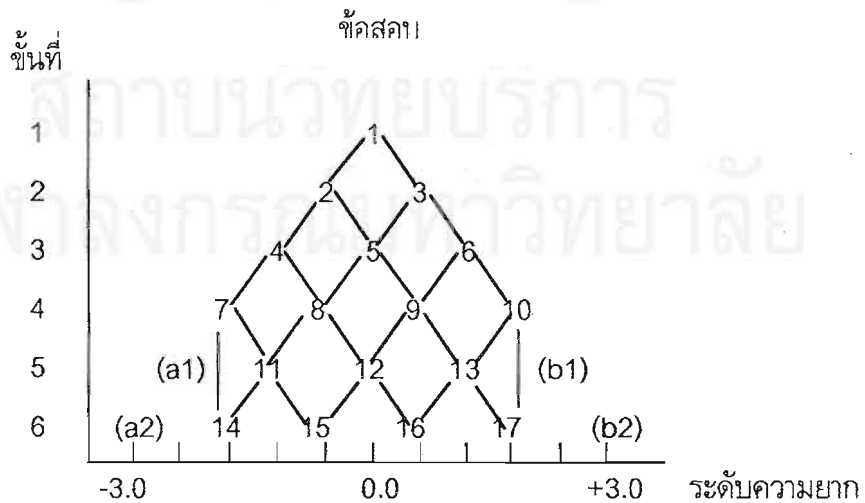
2) Robbins - Monro Procedure วิธีของ Robbins และ Monro นี้มีการปรับช่วงการเพิ่มขึ้นหรือลดลงของค่าความยากให้แคบลงอย่างต่อเนื่องตลอดการทดสอบ ซึ่งเป็นการทำให้การเปลี่ยนแปลงค่าความยากของข้อสอบหดตัวลงระหว่างการทดสอบ ดังแผนภาพที่ 8



แผนภาพที่ 8 โครงสร้างของการทดสอบรูปปิรามิดชนิดรอกบิน - มอนโร

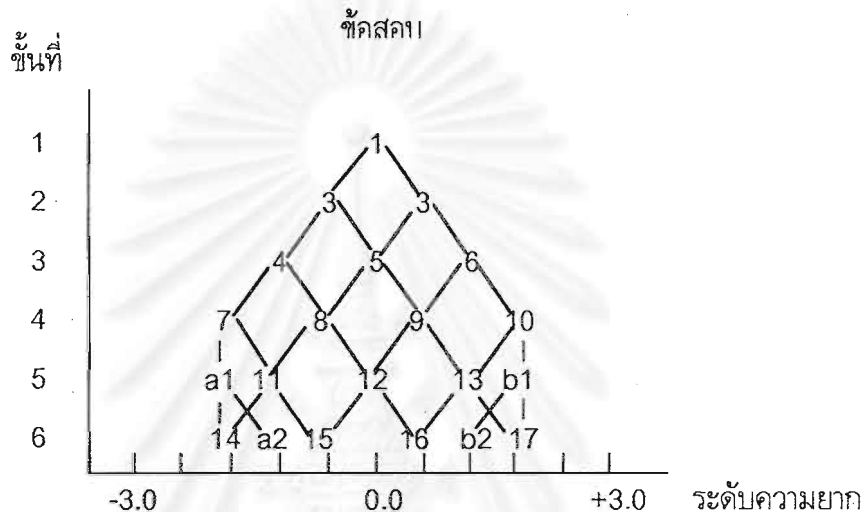
การจัดชุดข้อสอบสำหรับการทดสอบลักษณะนี้ จึงต้องเตรียมข้อสอบที่มีช่วงห่างของค่าความยากไม่เท่ากันในแต่ละชั้นของการทดสอบ โดยในชั้นตอนแรกๆ ต้องใช้ข้อสอบที่มีความยากปานกลาง ข้อสอบในชั้นตอนนั้นจะมีช่วงห่างของค่าความยากค่อนข้างกว้าง ชั้นตอนต่อไป ข้อสอบที่ง่ายลงหรือยากขึ้น จะต้องมีส่วนห่างของค่าความยากที่แคบลง ตามลำดับ

1.1.3 รูปแบบปิรามิดข้างตัด เป็นแบบที่ลดจำนวนข้อสอบลงจากแบบปิรามิดขนาดชั้นคงที่ด้วยวิธีการสกัดการสะท้อนกลับและวิธีรักษาการสะท้อนกลับ ดังแผนภาพที่ 9 และ 10



แผนภาพที่ 9 โครงสร้างของการทดสอบรูปแบบปิรามิดข้างตัดตามวิธีสกัดการสะท้อนกลับ

จากแผนภาพที่ 9 การตอบข้อสอบในชั้นที่ 1 ถึงชั้นที่ 3 มีวิธีตอบเช่นเดียวกับแบบปิรามิดขนาดชั้นคงที่ จนถึงจุดสกัดในชั้นที่ 4 ถ้าตอบข้อสอบข้อที่ 7 ถูก ก็จะแยกไปตอบข้อที่ 11 แต่ถ้าตอบข้อที่ 7 ผิด ก็จะไปตอบข้อที่ 14 และในทำนองเดียวกัน ถ้าตอบข้อที่ 10 ผิดก็จะไป ก็จะแยกไปตอบข้อที่ 13 แต่ถ้าตอบข้อที่ 10 ถูกก็จะไปตอบข้อที่ 17

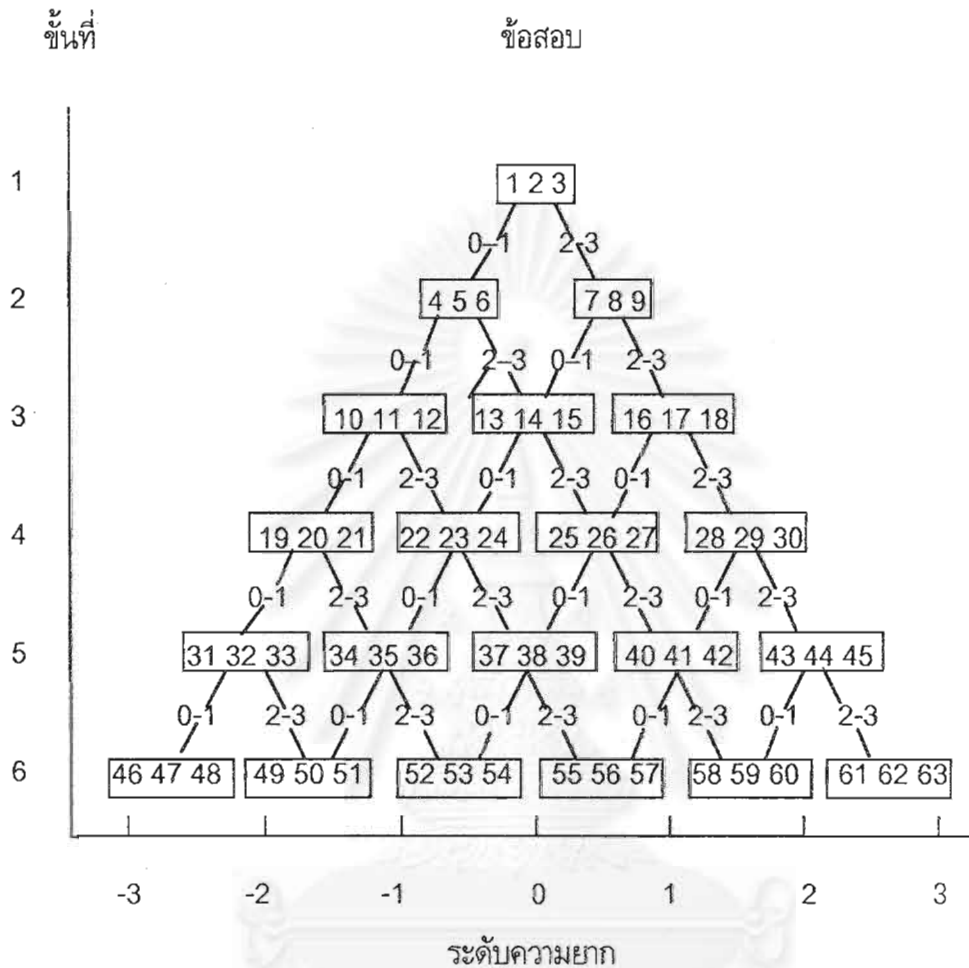


แผนภาพที่ 10 โครงสร้างของการทดสอบรูปแบบปิรามิดข้างตัดตามวิธีการสลับ

จากแผนภาพที่ 10 การตอบข้อสอบในชั้นที่ 1 ถึงชั้นที่ 3 มีวิธีตอบเช่นเดียวกับแบบพีระมิดชนิดขนาดชั้นคงที่ 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

1.1.4 รูปแบบปิรามิดแบบมีหลายข้อในแต่ละชั้น เป็นแบบที่มีจำนวนข้อในแต่ละชั้นมากกว่าหนึ่งข้อ ทั้งนี้เพื่อให้การตอบเพื่อแยกทางได้ถูกต้องมากยิ่งขึ้น ดังแผนภาพที่ 11

จากแผนภาพที่ 11 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



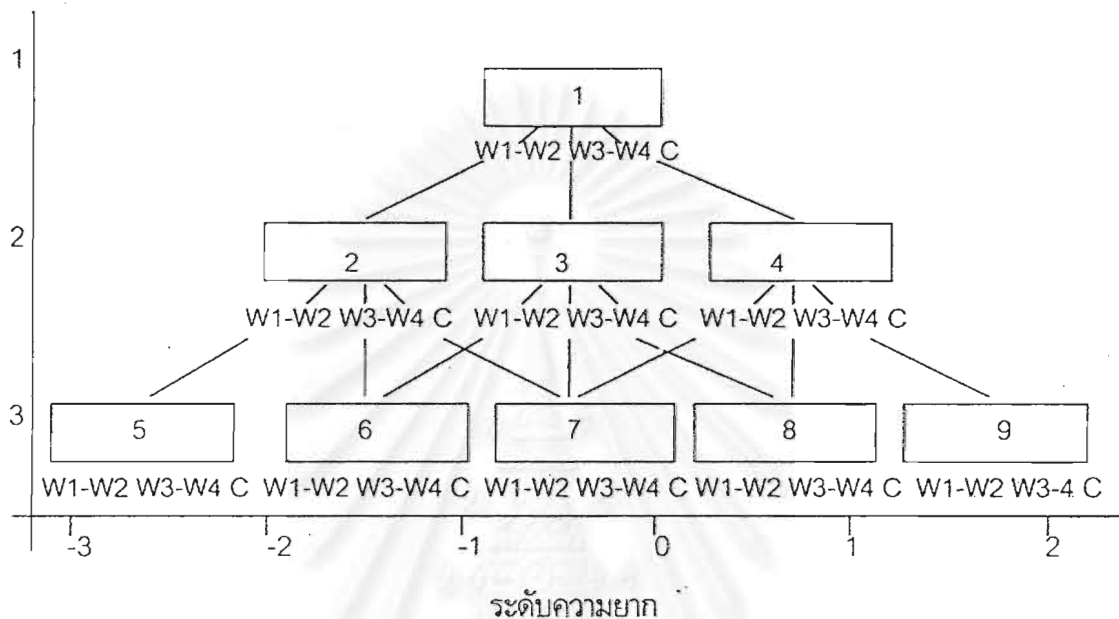
แผนภาพที่ 11 โครงสร้างการทดสอบรูปแบบปิรามิดมีหลายข้อในแต่ละชั้น

1.1.5 รูปแบบปิรามิดชนิดนี้ให้ให้นักตัวเลือกของข้อสอบเพื่อแยกทาง เป็นแบบที่คล้ายกับแบบที่ใช้ข้อสอบหลายๆ ข้อในแต่ละชั้น คือ แทนที่จะใช้ข้อสอบหลายๆ ข้อในแต่ละชั้นก็ใช้ตัวเลือกของข้อสอบแทนรูปแบบนี้เหมาะสำหรับข้อสอบเลือกตอบที่สามารถกำหนดน้ำหนักคะแนนของตัวเลือกได้ ดังแผนภาพที่ 12

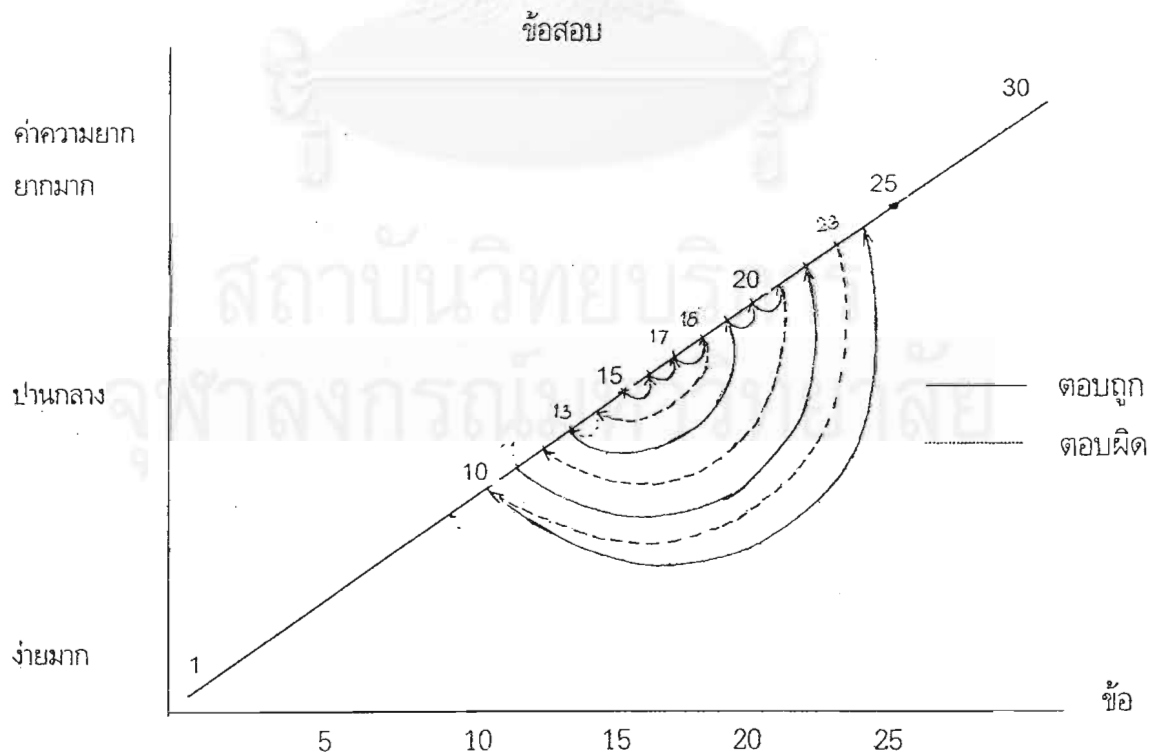
1.2 รูปแบบเฟลิกซ์เลเวล

การทดสอบแบบนี้เป็นแบบที่ลอร์ดคิดขึ้นมาในครั้งแรกเพื่อใช้กับการทดสอบที่ใช้กระดาษกับดินสอ (paper and pencil test) ซึ่งประกอบด้วยชุดของข้อสอบจำนวนหนึ่ง ข้อสอบแต่ละข้อมีช่วงห่างของค่าความยากเท่ากัน โดยทำการเรียงจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด

การทดสอบเริ่มต้นด้วยการให้ผู้ตอบทำข้อสอบข้อที่มีความยากปานกลาง ถ้าตอบถูกก็จะไปตอบข้อที่ยากขึ้น แต่ถ้าตอบผิดข้อต่อไปก็จะง่ายลง ดังแผนภาพที่ 13



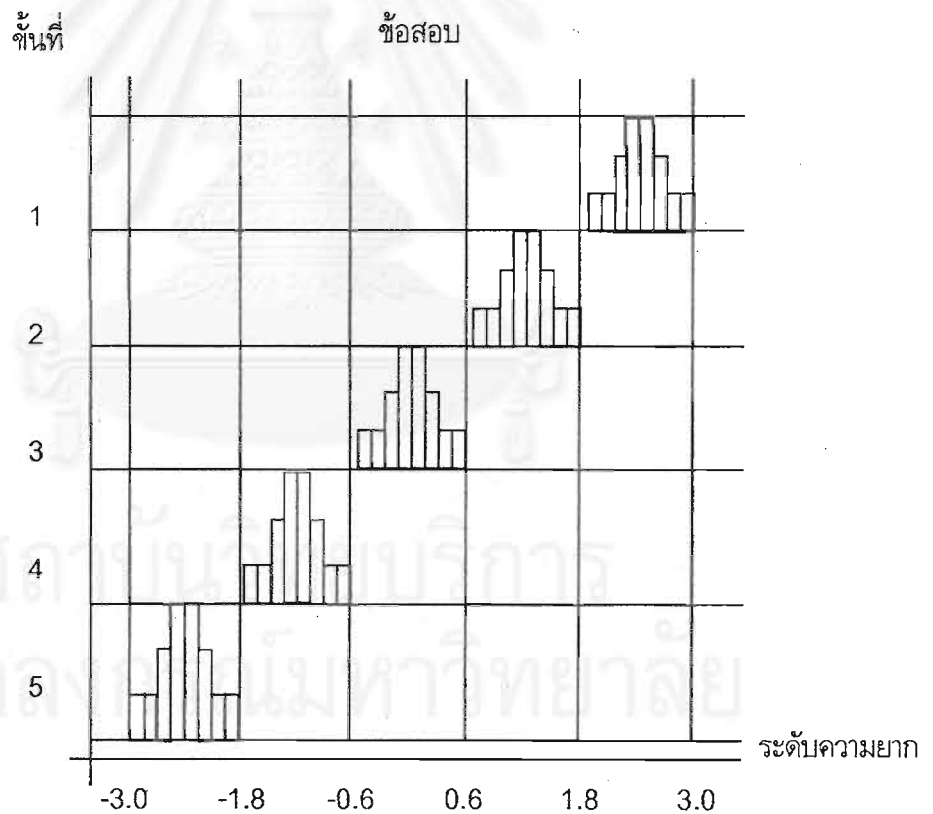
แผนภาพที่ 12 โครงสร้างของการทดสอบรูปแบบปิรามิดชนิดให้น้ำหนักตัวเลือกข้อสอบเพื่อแยกทาง



แผนภาพที่ 13 โครงสร้างของการทดสอบรูปแบบเฟลิกซีเลเวล

1.3 รูปแบบปรับระดับขั้น แบบสอบปรับระดับแบบแบ่งขั้นพัฒนามาจากแบบสอบปรับระดับแบบเพิลิกซิเลเวล และมีลักษณะคล้ายกับลักษณะการใช้แบบสอบวัดเชาว์ปัญญาของ Binet คือมีการดำเนินการทดสอบโดยใช้ข้อสอบจากคลังข้อสอบที่มีการจัดแบ่งความยากของข้อสอบเป็นระดับขั้น (stratified) แต่ละระดับขั้นประกอบด้วยชุดของข้อสอบที่มีค่าความยากอยู่ในช่วงที่กำหนดเดียวกัน เช่น ค่าความยากอยู่ระหว่าง .20 ถึง .30, .31 ถึง .40 เป็นต้น

การใช้แบบสอบปรับระดับขั้นนี้เริ่มต้นที่ข้อสอบระดับความยากใดก็ได้ ถ้าทำข้อนั้นถูก ข้อต่อไปจะเป็นข้อที่มีอำนาจจำแนกสูงสุดที่มีความยากถัดขึ้น แต่ถ้าทำข้อนั้นผิด ข้อต่อไปจะเป็นข้อที่มีอำนาจจำแนกสูงสุดที่มีความง่ายถัดลงไป การทดสอบจะดำเนินเช่นนี้ไปเรื่อยจนถึงความยากระดับเพดาน (ceiling stratum) ซึ่งผู้สอบไม่สามารถทำข้อสอบชุดนั้นได้เลยสักข้อ หรือทำคะแนนได้ไม่เกินคะแนนที่ได้จากการเดา (chance score) ดังแผนภาพที่ 10



แผนภาพที่ 14 โครงสร้างของการทดสอบรูปแบบปรับระดับขั้น

2. รูปแบบแยกทางแปรผัน

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้กลวิธีหลายขั้นตอนตามแบบทางแยกแปรผัน หมายถึง รูปแบบการตอบข้อสอบหลายขั้นตอนที่ไม่ได้กำหนดข้อสอบและเส้นทางในการตอบข้อสอบไว้ล่วงหน้าว่าถ้าผู้สอบตอบถูกจะต้องไปทำข้อสอบข้อใดหรือถ้าตอบผิดจะต้องไปทำข้อใดต่อไป ในแบบทางแยกแปรผันนี้จะไม่มีความซับซ้อน แต่จะดำเนินการทดสอบจากกลุ่มข้อสอบที่คำนวณระดับความยากและค่าอำนาจจำแนกของข้อสอบไว้แล้ว กฎการคัดเลือกข้อทดสอบถัดไปเพื่อให้เหมาะสมกับความสามารถของผู้สอบนั้นจะใช้แบบทางคณิตศาสตร์เข้ามาช่วย ได้แก่ การใช้กลวิธีของเบส์และกลวิธีความเป็นไปได้สูงสุด ซึ่งมีรายละเอียดดังนี้

2.1 ยุทธวิธีของเบส์

ยุทธวิธีของเบส์นี้ประยุกต์ทฤษฎีของเบส์ในกระบวนการตอบข้อสอบและอาศัยเครื่องคอมพิวเตอร์ในการทดสอบ จึงมักเรียกว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ โดยใช้เครื่องคอมพิวเตอร์ช่วยมีขั้นตอนดังนี้

- 1) ในแต่ละขั้นของการทดสอบจะมีการประมาณค่าความสามารถของผู้สอบและความคลาดเคลื่อนมาตรฐานในการประมาณค่าไว้จากข้อมูลต่างๆ ที่มีอยู่เกี่ยวกับผู้สอบและข้อสอบ
- 2) คัดเลือกข้อสอบจากกลุ่มข้อสอบที่ได้คำนวณค่าพารามิเตอร์ของข้อสอบไว้แล้ว ข้อสอบในกลุ่มข้อสอบทุกข้อที่ยังไม่ได้นำมาให้ผู้สอบคนนั้นๆ สอบจะถือว่าเป็นข้อสอบที่มีโอกาสนำมาใช้สอบได้ กระบวนการนี้แสดงให้เห็นว่าข้อสอบข้อใดๆ ในกลุ่มข้อสอบที่นำมาใช้ในการทดสอบกับผู้สอบคนใดก็ตาม จะเป็นข้อสอบที่มีระดับความยากใกล้เคียงกับระดับความสามารถของผู้สอบมากที่สุด หลังจากดำเนินการสอบโดยใช้ข้อสอบที่คัดเลือกไว้ก็จะประมาณความสามารถของผู้สอบไว้ก่อน แล้วการตอบถูกหรือการตอบผิดจะนำมารวมกันเพื่อคำนวณโดยใช้ทฤษฎีบทของเบส์หาค่าความสามารถในภายหลัง การประมาณค่าครั้งหลังนี้เป็น การประมาณค่าที่ปรับจากค่าที่ได้คำนวณไว้แต่เดิม กระบวนการนี้จะสิ้นสุดลงเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถมีค่าน้อยกว่าที่กำหนดไว้ล่วงหน้า

2.2 ยุทธวิธีการประมาณค่าความเป็นไปได้สูงสุด

วิธีนี้ได้พัฒนาขึ้นโดยใช้วิธีการประมาณค่าความเป็นไปได้สูงสุดตามทฤษฎีการตอบสนองข้อสอบ ซึ่งมีวิธีการประมาณค่าความสามารถของผู้สอบที่คล้ายคลึงกับกระบวนการตามทฤษฎีบทของเบส์ แม้ว่าเหตุผลทางคณิตศาสตร์ต่างกัน หลังจากที่ผู้สอบตอบข้อสอบเพียง 1 ข้อก็จะนำผลการทดสอบไปแก้สมการความเป็นไปได้สูงสุด และจะได้ค่าประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐาน ข้อสอบข้อถัดไปที่เลือกมาใช้ทดสอบจะเป็นข้อสอบจากกลุ่มข้อสอบและเป็นข้อสอบที่มีระดับความยาก

เหมาะสมกับความสามารถของผู้สอบมากที่สุด เมื่อผู้สอบตอบข้อสอบข้อนั้นแล้วก็จะมีค่าประเมินค่าความสามารถของผู้สอบทันทีจากข้อมูลการตอบทั้งหมดของผู้สอบ ซึ่งรวมถึงข้อสอบข้อสุดท้ายที่ผู้สอบตอบ จากนั้นก็ประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐานใหม่โดยใช้สมการความเป็นไปได้สูงสุด จนกว่าการทดสอบจะสิ้นสุดตามเกณฑ์ที่กำหนดไว้

การให้คะแนนการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

การให้คะแนนการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีหลายวิธี คือ (Weiss, 1974)

1. การให้คะแนนตามระดับความยากของข้อที่ยากที่สุดที่ตอบถูก (the most difficult of items answered correctly) เป็นการให้คะแนนโดยดูจากข้อสอบที่มีระดับความยากสูงสุดที่ผู้สอบสามารถตอบถูกก็จะถือว่าเป็นค่าความสามารถของผู้สอบ
2. การให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก (average difficulty of all items answered correctly) เป็นการให้คะแนนโดยการนำระดับความยากของข้อสอบทุกข้อที่ผู้สอบตอบถูกมาหาค่าเฉลี่ย ค่าเฉลี่ยความยากนั้นถือเป็นค่าความสามารถของผู้สอบ
3. การให้คะแนนตามค่าเฉลี่ยความยากของทุกข้อที่ทำ (average difficulty) คือ นำระดับความยากของข้อสอบทุกข้อที่ผู้สอบได้ตอบโดยไม่คำนึงว่าจะตอบถูกหรือผิดมาหาค่าเฉลี่ย ค่าเฉลี่ยนี้ถือเป็นค่าความสามารถของผู้สอบ
4. การให้คะแนนตามระดับความยากของข้อสุดท้ายที่ตอบ (difficulty of the final item) คือ ไม่ว่าจะข้้นสุดท้ายจะตอบถูกหรือผิด จะถือว่าระดับความยากของข้อสอบข้อสุดท้ายที่ผู้สอบตอบคือค่าความสามารถของผู้สอบ
5. การให้คะแนนตามระดับความยากของข้อสอบที่ต่อจากข้อสุดท้ายที่ตอบ (difficulty of $(N+1)^{th}$ item) ถ้าข้้นสุดท้ายที่ผู้สอบสามารถตอบถูก คะแนนในข้้นต่อไปก็จะเพิ่มขึ้น แต่ถ้าข้้น สุดท้ายตอบผิด ข้้นต่อไปคะแนนก็จะลดลง
6. การให้คะแนนตามวิธีของเบย์ส์ (Bayesian Estimation) ถ้าผู้สอบทำข้อสอบถูกหมดหรือผิดหมด การประมาณค่าความสามารถด้วยวิธีความเป็นไปได้สูงสุด (maximum likelihood) จะทำได้ไม่ดี เมื่อนำสารสนเทศกายแรกของการแจกแจงความสามารถของกลุ่มผู้สอบมาใช้ วิธีของเบย์ส์ จะมีความเหมาะสมในการประมาณค่าความสามารถ การประมาณค่าความสามารถกายแรกของผู้สอบในกลุ่มจะต้องตั้งสมมติฐานเกี่ยวกับสารสนเทศความสามารถของผู้สอบคนหนึ่งที่ไม่แตกต่างไปจากกลุ่ม ผู้สอบคนอื่นๆ (Novick, Lewis &

Jackson, 1973 อ้างถึงใน Hambleton and Swaminathan, 1990) การกำหนดค่าสารสนเทศภายใต้แรก ต้องมีการกำหนดค่าการแจกแจงของความสามารถ (θ_a) คือ

$$\theta_a \sim N(\mu, \phi)$$

เมื่อ $N(\mu, \phi)$ เป็นการกระจายแบบปกติ มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ ϕ Owen (1975) เสนอให้กำหนด $\mu = 0$ และ $\phi = 1$ สำหรับการทดสอบแบบ ปรับเหมาะกับความสามารถของผู้สอบ

การกำหนดรูปร่างของการแจกแจงเป็นรูปโค้งปกติทำให้มีความสะดวก เบอร์บัวม (Birnbaum) เชื่อว่า การแจกแจงภายใต้แรกของความสามารถเป็นฟังก์ชันโลจิสติก ซึ่งมีสมการดังนี้

$$f(\theta) = \frac{\exp(\theta)}{[1 + \exp(\theta)]^2}$$

หัวใจของวิธีการประมาณค่าความสามารถตามวิธีของเบย์ คือ ทฤษฎีของเบย์ ซึ่งมีความสัมพันธ์กับ Condition probability และ marginal probability สูตรคือ

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

ในบริบทของการประมาณค่าความสามารถ A อาจกำหนดให้เป็น θ_a และ B เป็น การตอบข้อสอบ n ข้อ ได้คะแนน μ จากสมการ $P(B|A)$ สามารถแสดงได้เป็น

$$P(\theta_a|\mu) = \frac{P(\mu|\theta_a)P(\theta_a)}{P(\mu)}$$

เมื่อ θ_a เป็นตัวแปรต่อเนื่อง ดังนั้นจึงเป็นฟังก์ชันความหนาแน่น (density function) ซึ่งอาจจะทำให้สับสนระหว่าง $P(\theta_a)$ กับ $P(\theta)$ ซึ่งเป็นฟังก์ชันการตอบสนองของข้อสอบ จึงใช้ $f(\theta_a)$ แทน $P(\theta)$ จึงเขียนสมการใหม่ได้คือ

$$f(\theta_a | \mu) = \frac{f(\mu | \theta_a) f(\theta_a)}{f(\mu)}$$

เมื่อ $f(\mu)$ เป็นค่าคงที่

$f(\theta_a | \mu)$ เป็นความหนาแน่นภายหลัง (posterior density) ของ θ_a

$f(\theta_a)$ เป็นความหนาแน่นภายหลังแรก (prior density) ของ θ_a

จาก $f(\theta_a | \mu)$ สามารถกำหนดให้เป็นฟังก์ชันความเป็นไปได้สูงสุดได้จากสมการ

$$f(\theta_a | \mu) \propto L(\mu | \theta_a) f(\theta_a)$$

วิธีการประมาณค่าความสามารถของผู้สอบตามวิธีการของเบส์นี้ได้ถูกนำไปประยุกต์เป็นวิธีต่างๆ อีกหลายวิธี แต่วิธีที่ได้รับความนิยมมาก คือ วิธีของเบส์ที่ปรับเปลี่ยน (bayesian updating) ซึ่งเสนอโดยโอเวน (Owen, 1975) มีวิธีการประมาณค่า ดังนี้

ในกรณีที่ตอบข้อสอบถูก

$$\theta_{m+1} = \theta_m + (1 + c_i) \left[\frac{\sigma_m^2}{\sqrt{\frac{1}{a_i^2} + \sigma_m^2}} \left[\frac{\phi(D)}{c_i + (1 - c_i)\Phi(-D)} \right] \right]$$

และ

$$\sigma_m^2 + 1 = \sigma_m^2 \left[1 - \frac{1 - c_i}{1 + \frac{1}{a_i^2 \sigma_m^2}} \left| \frac{\phi(D)}{A} \right| \left| \frac{(1 - c_i \phi(D))}{A} - D \right| \right]$$

ในกรณีที่ตอบข้อสอบผิด

$$\theta_{m+1} = \theta_m - \left[\frac{\sigma_m^2}{\sqrt{\frac{1}{a_i^2} + \sigma_m^2}} \left[\frac{\phi(D)}{\Phi(D)} \right] \right]$$

และ

$$\sigma_m^2 + 1 = \sigma_m^2 \left[1 - \frac{\phi(D)}{1 + \frac{1}{a_i^2 \sigma_m^2}} \frac{\frac{\phi(D)}{\Phi(D)} + D}{\Phi(D)} \right]$$

เมื่อ $\phi(D)$ คือ ค่าตั้งฉากของโค้งปกติ ณ จุด D

$\Phi(D)$ คือ พื้นที่ใต้โค้งปกติจากค่า D ต่ำสุดจนถึงจุด D

$\Phi(D)$

$$D = \frac{b_i - \theta_m}{\sqrt{\frac{1}{a_i^2} + \sigma_m^2}}$$

$$A = c_i + (1 - c_i)\Phi(-D)$$

เมื่อ θ_m คือ ความสามารถของผู้สอบที่ประมาณค่าได้ก่อนตอบข้อสอบข้อที่ m+1 โดยปกติถ้าเป็นการประมาณค่าครั้งแรกที่ยังไม่ทราบความสามารถของผู้สอบจะให้ $\theta_m = 0.0000$

σ_m^2 คือ ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อสอบข้อที่ m+1 โดยปกติถ้าเป็นการประมาณค่าครั้งแรกที่ยังไม่ทราบค่าความแปรปรวนจะให้ $\theta_m = 1.0000$

θ_{m+1} คือ ความสามารถของผู้สอบที่ประมาณค่าได้เมื่อตอบข้อสอบข้อที่ m+1 แล้ว

$\sigma_m^2 + 1$ คือ ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบเมื่อตอบข้อสอบข้อที่ m+1 แล้ว

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ m+1

b_i คือ ระดับความยากของข้อสอบข้อที่ m+1

c_i คือ ค่าการเดาของข้อสอบข้อที่ m+1

7. การให้คะแนนโดยวิธีการประมาณค่าความเป็นไปได้สูงสุด (maximum likelihood)

ซึ่งมีหลายวิธี แต่วิธีที่นิยมมาก คือ วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (conditional maximum likelihood) การประมาณค่าโดยวิธีนี้มีข้อจำกัด คือ ถ้าผู้สอบตอบข้อสอบถูกหมดหรือผิดหมด จะไม่สามารถประมาณค่าได้ โดยมีสูตรดังนี้ (Hambleton and Swaminathan, 1990)

ถ้าหากกลุ่มของผู้สอบ N คน ทำการทดสอบด้วยข้อสอบ n ข้อ ที่ทราบค่าพารามิเตอร์ของข้อสอบ แล้วก็จะทำให้สามารถประมาณค่าความสามารถของผู้สอบได้ (θ_a) เมื่อ $a = 1, \dots, N$

$$U_a = [U_{1a}, U_{2a}, \dots, U_{na}]$$

เมื่อ $a = 1, \dots, N$

U เป็นเวกเตอร์การตอบข้อสอบของผู้สอบคนที่ a และ $[\theta_1, \dots, \theta_n]$ แสดงเวกเตอร์ของความสามารถของผู้สอบ N คน ฟังก์ชัน likelihood สำหรับเวกเตอร์ U ของผู้สอบ N คน เมื่อ $U = [U_1, U_2, \dots, U_n]$ คือ

$$L(U|\theta) \equiv L(U_1, U_2, \dots, U_N | \theta_1, \theta_2, \dots, \theta_N) = \prod_{a=1}^N L(U_a | \theta_a)$$

$$\prod_{a=1}^N \prod_{i=1}^n L(U_{ia} | \theta) = \prod_{a=1}^N \prod_{i=1}^n P_{ia}^{u_{ia}} Q_{ia}^{1-u_{ia}}$$

เมื่อ $P_{ia} \equiv P_i(\theta_a)$

logarithm ของฟังก์ชัน likelihood คือ

$$\ln L(U_1, U_2, \dots, U_N | \theta) = \sum_{a=1}^N \sum_{i=1}^n [U_{ia} \ln P_{ia} + (1 - U_{ia}) \ln(1 - P_{ia})]$$

การประมาณค่าความเป็นไปได้สูงสุดของ $\theta_1, \dots, \theta_N$ คือ การแก้สมการ likelihood

$$\frac{\partial}{\partial \theta_a} \ln L(U_1, U_2, \dots, U_N | \theta) = 0$$

เมื่อ $a = 1, \dots, N$

สมการเหล่านี้สามารถแสดงได้คือ

$$\frac{\partial \ln L}{\partial \theta_a} = \sum_{i=1}^n \frac{\partial \ln L}{\partial P_{ia}} \frac{\partial P_{ia}}{\partial \theta_a}$$

$$\begin{aligned}
&= \sum_{i=1}^n \left(\frac{U_{ia}}{P_{ia}} - \frac{1-U_{ia}}{-1P_{ia}} \right) \frac{\partial P_{ia}}{\partial \theta_a} \\
&= \sum_{i=1}^n \left(\frac{U_{ia} - P_{ia}}{P_{ia} Q_{ia}} \right) \frac{\partial P_{ia}}{\partial \theta_a} = 0
\end{aligned}$$

สำหรับโมเดล 1 พารามิเตอร์ สมการ likelihood คือ

$$\frac{\partial \ln L}{\partial \theta_a} \equiv D \sum_{i=1}^n (U_{ia} - P_{ia}) = 0$$

หรือ

$$D \left(r_a - \sum_{i=1}^n P_{ia} \right) = 0$$

เมื่อ $r_a = \sum U_{ia}$ คือ จำนวนข้อสอบที่ผู้สอบคนที่ a ทำถูก

$$\frac{\partial^2 \ln L}{\partial \theta_a^2} = -D^2 \sum_{i=1}^n P_i Q_i$$

ค่าเริ่มต้นสำหรับ θ_a , θ_{0a} สำหรับผู้สอบคนที่ a คือ

$$\begin{aligned}
\theta_{0a} &= \ln \left[\frac{r_a}{(n - r_a)} \right] \\
&\equiv \theta_0
\end{aligned}$$

ประมาณค่าความสามารถของผู้สอบใหม่ โดยใช้สูตร

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

หาค่าปรับแก้ (h_m) โดยใช้สูตร

$$h_m = D \left[\frac{r - \sum_{i=1}^n P_i(\theta_m)}{-D^2 \sum_{i=1}^n P_i(\theta_m) Q_i(\theta_m)} \right]$$

เมื่อค่า $|h_m|$ มีค่าน้อยกว่า ε (นิยมใช้ 0.001) กระบวนการทำซ้ำก็จะยุติลง

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

องค์ประกอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Weiss and Kingsbury, 1984)

1. โมเดลการตอบสนองข้อสอบ (item response model) ภายใต้บริบทของทฤษฎีการตอบสนองข้อสอบแล้ว การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์หรือการทดสอบแบบซี เอ ที (CAT) นั้นจะใช้โมเดลโลจิสติก (logistic model) 1 2 หรือ 3 พารามิเตอร์ หรือใช้โมเดลโอจิว (Ogive Model) ก็ได้ การเลือกโมเดลมาใช้ในการวิเคราะห์ข้อมูลนั้นขึ้นอยู่กับลักษณะข้อสอบว่ามีลักษณะอย่างไร มีความเหมาะสมกับโมเดลใด

2. คลังข้อสอบ (item pool) ประกอบด้วยชุดของข้อสอบที่มุ่งวัดลักษณะหรือความสามารถที่ต้องการ ข้อสอบจะต้องครอบคลุมช่วงของระดับความสามารถต่างๆ กัน มีจำนวนข้อสอบที่มากพอ โดยข้อสอบแต่ละข้อควรมีอำนาจจำแนกสูง มีค่าความยากที่ครอบคลุมช่วงระดับต่างๆ อย่างเหมาะสม และควรมีจำนวนข้อสอบเท่าๆ กันในแต่ละระดับค่าความยาก

3. ความยากง่ายของข้อสอบข้อแรก (entry level) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ข้อสอบข้อแรกจะมีระดับความยากง่ายแตกต่างกันสำหรับผู้สอบแต่ละคน การเริ่มทำข้อสอบที่คลาดเคลื่อนไปจากความสามารถจริงไม่มีผลกระทบต่อการประมาณค่าความสามารถ แต่ถ้าหากมีการเลือกข้อสอบข้อแรกให้กับผู้สอบอย่างเหมาะสมแล้วก็จะช่วยประหยัดเวลาในการทดสอบได้มาก เพราะลดจำนวนข้อสอบลง

4. เกณฑ์การคัดเลือกข้อสอบ (item selection rule) มีวิธีการเลือกข้อสอบจาก คลังข้อสอบที่มีประสิทธิภาพ 2 วิธีการ คือ วิธีการเลือกข้อสอบที่มีค่าสารสนเทศสูงสุด (maximum information item selection) และวิธีการเลือกข้อสอบแบบเบย์ส์ (bayesian item selection) ซึ่งวิธีแรกจะให้ค่าสารสนเทศสูงสุด

ส่วนวิธีหลังจะให้ค่าความแปรปรวนที่คาดหวังของการประมาณค่าความสามารถต่ำสุด ทั้ง 2 วิธีจะคล้ายคลึงกัน เนื่องจากว่าค่าสารสนเทศสูงสุดและความแปรปรวนที่คาดหวังของการประมาณค่ามีความสัมพันธ์กัน

5. วิธีการให้คะแนน (scoring method) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ มีวิธีการให้คะแนนหลายวิธีด้วยกัน คือ (Weiss, 1974)

5.1 การให้คะแนนตามระดับความยากของข้อที่ยากที่สุดที่ตอบถูก เป็นการให้คะแนนโดยดูจากข้อสอบที่มีระดับความยากสูงที่สุดที่ผู้สอบสามารถตอบได้ถูกต้อง

5.2 การให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก เป็นการให้คะแนนโดยการนำระดับความยากของข้อสอบทุกข้อที่ผู้สอบตอบถูกมาหาค่าเฉลี่ย

5.3 การให้คะแนนตามค่าเฉลี่ยความยากของทุกข้อที่ทำ เป็นการนำระดับความยากของข้อสอบทุกข้อที่ผู้สอบได้ตอบโดยไม่คำนึงว่าจะตอบถูกหรือไม่มาหาค่าเฉลี่ย

5.4 การให้คะแนนตามระดับความยากของข้อสุดท้ายที่ตอบ ไม่ว่าจะขั้นสุดท้ายจะตอบถูกหรือไม่

5.5 การให้คะแนนตามระดับความยากของข้อสอบที่ต่อจากข้อสุดท้ายที่ตอบ กล่าวคือ ถ้าขั้นสุดท้ายที่ผู้สอบสามารถตอบถูก คะแนนในขั้นต่อไปก็จะเพิ่มขึ้น แต่ถ้าขั้นสุดท้ายตอบผิดขั้นต่อไป คะแนนก็จะลดลง

5.6 การให้คะแนนตามวิธีของเบย์ มีหลายวิธี แต่วิธีที่นิยมใช้กันมาก คือ วิธีของเบย์ที่ปรับปรุงให้ทันสมัย (Bayesian updating) ซึ่งสามารถประมาณค่าความสามารถของผู้สอบได้ค่อนข้างคงที่

5.7 การให้คะแนนโดยวิธีการประมาณค่าความเป็นไปได้สูงสุด (maximum likelihood) มีหลายวิธี แต่วิธีที่นิยมใช้กันมากคือวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (Conditional maximum likelihood) การประมาณค่าโดยวิธีนี้มีข้อจำกัด คือ ถ้าผู้สอบตอบข้อสอบถูกหมดหรือผิดหมด จะไม่สามารถประมาณค่าได้

6. เกณฑ์ยุติการทดสอบ (termination criterion) ลักษณะสำคัญของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ คือ การทดสอบจะดำเนินไปเรื่อยๆจนกระทั่งถึงเกณฑ์ที่ระบุไว้ให้ยุติการทดสอบ การดำเนินการสอบก็จะยุติลง เกณฑ์ที่ใช้กันอยู่ทั่วไปมีอยู่ 2 เกณฑ์ คือ

6.1 การกำหนดจำนวนข้อสอบคงที่ เป็นการกำหนดจำนวนข้อสอบที่ใช้ในการทดสอบให้คงที่สำหรับทุกคน เช่น กำหนดให้การทดสอบแบบปรับเหมาะใช้ข้อสอบจำนวน 25 ข้อ ($n = 25$) เมื่อผู้สอบทำข้อสอบได้ครบ 25 ข้อ การทดสอบก็จะยุติลง เกณฑ์นี้ค่อนข้างเป็นประโยชน์ในการศึกษาภายใต้สถานการณ์

จำลองแบบมอนติ คาร์โล (Monte Carlo Simulation) เนื่องจากจำนวนข้อสอบเท่ากันทำให้สามารถเปรียบเทียบสารสนเทศของแบบสอบได้โดยตรง ในทางปฏิบัติการกำหนดให้ทุกคนทำข้อสอบจำนวนเท่ากัน อาจมีคุณภาพของการวัดผลได้แตกต่างกัน (ศิริชัย กาญจนวาสี, 2538)

6.2 การกำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ในทางปฏิบัติการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ ควรให้ดำเนินการไปเรื่อยๆจนกว่าการประมาณค่า θ มีความคลาดเคลื่อนมาตรฐาน (Standard Error; SE(θ)) ลดต่ำลงจนถึงระดับที่ยอมรับได้ การทดสอบจึงควรรุยติลง

การดำเนินการทดสอบ

การดำเนินการทดสอบเริ่มต้นโดยการให้ผู้สอบนั่งหน้าจอคอมพิวเตอร์ที่มีเครื่องขับแผ่นบันทึก (disk drive) ในแผ่นบันทึกข้อมูลจะมีโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (MicroCAT) ที่มีองค์ประกอบต่างๆ ตามที่กำหนด ผู้ดำเนินการทดสอบจัดเตรียมเครื่องให้อยู่ในสภาพที่พร้อมจะทดสอบ และชี้แจงเกี่ยวกับวิธีการตอบข้อสอบ จากนั้นเครื่องคอมพิวเตอร์จะแสดงข้อสอบข้อแรกบนจอภาพ เพื่อให้ผู้สอบได้ตอบ ข้อสอบข้อแรกนี้จะเป็นข้อสอบที่มีระดับความยากง่ายปานกลาง ถ้าผู้สอบตอบข้อสอบข้อแรกถูกข้อสอบข้อต่อไปจะเป็นข้อสอบที่มีระดับความยากเพิ่มขึ้น แต่ถ้าผู้สอบข้อสอบข้อแรกผิด ข้อสอบข้อต่อไปจะเป็นข้อสอบที่มีระดับความยากลดลง จากนั้นก็จะประมาณค่าความสามารถของผู้สอบ การทดสอบจะดำเนินต่อไปในตามแนวทางต่อไปนี้ (สุพัฒน์ สุกมลสันต์, 2539)

1. หลังจากประมาณค่าความสามารถของผู้สอบแล้ว เครื่องคอมพิวเตอร์ก็จะคัดเลือก ข้อสอบข้อใหม่ที่มีความยากง่ายใกล้เคียงกับความสามารถของผู้สอบ
2. ผู้สอบตอบข้อสอบข้อใหม่ที่ปรากฏในหน้าจอคอมพิวเตอร์ คอมพิวเตอร์รับคำตอบและตรวจคำตอบ
3. ประมาณค่าความสามารถของผู้สอบอีกครั้งหนึ่งตามผลการตอบข้อสอบข้อที่ผ่านมา
4. พิจารณาว่าผู้สอบควรทำข้อสอบต่อไปหรือไม่ตามเกณฑ์ยุติการทดสอบที่กำหนดไว้
5. ถ้าผู้สอบต้องทำสอบข้อต่อไปก็ย้อนกลับไปขั้นที่ 1 อีกครั้ง

เมื่อสิ้นสุดการทดสอบ สารสนเทศเกี่ยวกับการตอบข้อสอบของผู้สอบก็จะถูกบันทึกไว้ในเครื่องคอมพิวเตอร์ และแสดงผลการทดสอบให้ผู้สอบทราบ

ตอนที่ 3 ลักษณะของระบบการทดสอบไมโครแคท (MicroCAT)

ระบบการทดสอบไมโครแคท (MicroCAT) เป็นระบบที่พัฒนาขึ้นเพื่อใช้กับไมโครคอมพิวเตอร์ (microcomputer) ในการพัฒนา การดำเนินการ การให้คะแนน และการวิเคราะห์ ซึ่งเป็นการทดสอบที่เอื้อต่อการทดสอบแบบประเพณีนิยมด้วย ระบบการทดสอบไมโครแคท (MicroCAT) ประกอบด้วย 3 ระบบ อันได้แก่ ระบบการพัฒนาค้างข้อสอบ ระบบการทดสอบด้วยคอมพิวเตอร์ และระบบการวิเคราะห์ข้อสอบและแบบสอบ (Assessment System Corporation, 1995)

ระบบที่ 1 ระบบการพัฒนาค้างข้อสอบ เป็นการพัฒนาค้างข้อสอบและแบบสอบ (ภาวิณี ศรีสุข วัฒนานันท์, 2543)

1. ระบบการพัฒนาค้างข้อสอบ (Item development)

MicroCAT Graphics item Banker (GIB) เป็นโปรแกรมที่ใช้ในการจัดเก็บ ข้อสอบเข้า ค้างข้อสอบ สามารถตรวจสอบ แก้ไขข้อสอบที่จัดเก็บไว้แล้วได้ ในแต่ละคลังข้อสอบสามารถเก็บข้อสอบได้ 999 ข้อ

2. ระบบการพัฒนาแบบสอบ (Test development)

การพัฒนาแบบสอบสามารถสร้างแบบสอบได้หลายรูปแบบ แต่ที่นิยมสร้างมี 3 รูปแบบ คือ

2.1 แบบสอบประเพณีนิยม สร้างโดยกำหนดกลุ่มข้อสอบและมีจำนวนข้อใน แบบสอบที่แน่นอน ผลการสอบใช้การนับจำนวนข้อที่ถูกต้อง แบบสอบชนิดนี้ง่ายต่อการจัดทำเหมาะสำหรับใช้ผลการสอบในการจัดกลุ่ม หรือเปรียบเทียบความสามารถระหว่างกลุ่มผู้สอบ แต่ไม่มีประสิทธิภาพพอที่จะบอกระดับความสามารถที่แท้จริงของบุคคล

2.2 แบบสอบแบบปรับเหมาะกับความสามารถของผู้สอบ แบบสอบชนิดนี้มีประสิทธิภาพในการจัดกลุ่มคน และบอกระดับความสามารถที่แตกต่างกันระหว่างบุคคล หลักการสำคัญของการสร้างแบบสอบชนิดนี้จะออกแบบให้แบบสอบมีระดับความยากง่ายที่เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล ซึ่งทำให้ได้สารสนเทศจากการสอบที่สามารถบ่งชี้ความสามารถของบุคคลนั้นได้ชัดเจนขึ้น

2.3 แบบสอบที่กำหนดจำนวนข้อ และลักษณะค่าตัวแปร หรือสร้างแบบสอบที่ใช้ แนวคิดรูปแบบปรับระดับขั้น (stradaptive)

ระบบการทดสอบไมโครแคท (MicroCAT) สามารถใช้ได้ทั้งจอสีและจอโมโนโครมในการดำเนินการทดสอบ การสร้างข้อสอบสามารถสร้างเป็นเท็กซ์ไฟล์ (textfile) หรือกราฟิก (graphic) ก็ได้ นอกจากนี้ ระบบการทดสอบสามารถนำเสนอข้อสอบได้หลายรูปแบบรวมทั้งลักษณะของการตอบด้วย ซึ่งรูปแบบปกติระบบการทดสอบจะแสดงข้อสอบบนหน้าจอคอมพิวเตอร์โดยบรรทัดล่างสุดจะแสดงข้อความเพื่อรับคำตอบของผู้สอบ โดยผู้สอบสามารถเลือกตัวเลือก โดยการกดเลข 1-6 หรือ กด A-F จากนั้นกด Enter ระบบจะทำการตรวจและบันทึกคะแนนไว้ หลังจากที่ถูกกด Enter ข้อสอบข้อต่อไปจะปรากฏขึ้นมา

โปรแกรมการทดสอบไมโครแคท (MicroCAT) ยังมีความสามารถในการสร้างตัวอักษรพิเศษขึ้นมาใหม่ได้ถึง 124 ตัวและโปรแกรมยังอำนวยความสะดวกในการพัฒนาแบบสอบแบบประเพณีนิยมโดยใช้เครื่องมือ (template) ซึ่งเครื่องมือสามารถสร้างแบบสอบได้ทั้งการถามตอบอย่างง่ายและการกำหนดความยาวของแบบสอบ ในการใช้เครื่องมือนี้จะมีโปรแกรมทำการคอมไพล์ (compile) แบบสอบนั้นด้วย นอกจากนี้โปรแกรมการทดสอบไมโครแคท (MicroCAT) ยังมีโปรแกรมย่อย Importer ซึ่งอำนวยความสะดวกในการนำข้อสอบที่สร้างด้วยโปรแกรมอื่น เช่น word processing เข้ามาไว้ในคลังข้อสอบของระบบได้ และโปรแกรมย่อย Exporter ที่อำนวยความสะดวกในการนำ ข้อสอบจากคลังข้อสอบมาทำการแก้ไขหรือเพิ่มเติมด้วยโปรแกรมอื่น เช่น word processing นอกจากนี้ยังมีโปรแกรมที่ใช้ในการพิมพ์แบบสอบ และคลังข้อสอบลงกระดาษอีกด้วย

ระบบที่ 2 ระบบการสอบผ่านเครื่องคอมพิวเตอร์ ระบบการทดสอบเป็นแบบอัตโนมัติขณะที่ทำการทดสอบจะมีการบันทึกคะแนนของผู้สอบลงในแฟ้มข้อมูล และสามารถรายงานผลได้ทันที ระบบสามารถสร้างรูปแบบการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้เครื่องมือ (template) หรือใช้ภาษา MCATL (Minnesota Computerized Adaptive Testing Language) และสามารถเพิ่ม effect เสียงดนตรีประกอบการเสนอข้อสอบได้ด้วย

ระบบที่ 3 ระบบการวิเคราะห์ข้อสอบและแบบสอบ ประกอบด้วยโปรแกรมต่างๆ ดังนี้

1. The Data Collection Program ใช้เก็บข้อมูลที่เป็นผลการสอบจากคอมพิวเตอร์ แฟ้มข้อมูลผลลัพธ์ มีนามสกุลเป็น .DAT

2. The MicroCAT Item and Test Analysis Program (ITEMAN) ให้ผลการวิเคราะห์แบบประเพณีนิยมซึ่งจะช่วยพิจารณาในด้านความเที่ยงของแบบสอบ จากผลการวิเคราะห์ รายข้อและให้ค่าสถิติบรรยาย ได้แก่ ค่าเฉลี่ย ค่าความแปรปรวน ค่าความเบ้ และค่าความโด่ง สัมประสิทธิ์ความเที่ยง ค่าความคลาดเคลื่อนมาตรฐานของการวัด และสหสัมพันธ์ภายในระหว่างแบบสอบย่อยในชุดของแบบสอบ

3. The Rasch Model Item Calibration Program (RASCAL) วิเคราะห์ข้อสอบตามแนวทฤษฎีการสอบข้อสอบ (IRT) แบบราสซ์โมเดล มีค่าพารามิเตอร์ประจำข้อสอบเพียงค่าเดียว คือ ค่าความยากของข้อสอบ (b) โปรแกรมการทดสอบไมโครแคท (MicroCAT) ได้พัฒนาการวิเคราะห์ข้อสอบแบบราสซ์โมเดลโดยอาศัยพื้นฐานการพัฒนามาจากแนวทฤษฎีของ Wright และ Stone (1997) จากแนวทฤษฎีนี้พบว่าค่าพารามิเตอร์ประจำข้อสอบที่ได้มีความคงที่มากกว่าค่าพารามิเตอร์ที่ได้จากการวิเคราะห์แบบโมเดลโลจิสติก

4. The MicroCAT Item Parameter Estimation Program (ASCAL) ประมาณค่าความสามารถของผู้สอบและค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก (b) อำนาจจำแนก (a) และค่าการเดา (c)

5. Test Pre-Evaluation Program (PREVAL) ประเมินศักยภาพของแบบสอบก่อนที่จะดำเนินการทดสอบ (การใช้ฟังก์ชันสารสนเทศของแบบสอบ) การประมาณค่าพารามิเตอร์ตามทฤษฎีการตอบสนองของข้อสอบ (IRT)

6. Test Validation Program (VALID) โปรแกรมจะทำการคำนวณ bivariate และการถดถอยพหุคูณสำหรับการวิเคราะห์ความตรงของแบบสอบ

การใช้ระบบการทดสอบไมโครแคท (MicroCAT)

โปรแกรมที่ใช้ในระบบการทดสอบไมโครแคท (MicroCAT) แบ่งออกเป็น 3 รูปแบบ รูปแบบและโปรแกรมสามารถทำงานได้โดยผ่านรายการคำสั่ง (menu) รายการคำสั่งถูกจัดเตรียมไว้สำหรับรูปแบบทั้งหมด ดังนั้นจึงสามารถใช้โปรแกรมการทดสอบไมโครแคท (MicroCAT) ได้อย่างง่าย ถึงแม้ว่าจะไม่มีประสบการณ์ในการใช้คอมพิวเตอร์มาก่อน การเลือกรายการคำสั่งทำโดยเลื่อนแถบสว่างมายังคำสั่งที่ต้องการ และกด Enter การเลื่อนแถบสว่างสามารถทำได้ 3 วิธี คือ กดปุ่ม Space Bar เลื่อนปุ่มลูกศรขึ้น-ลง และโดยการพิมพ์คำสั่ง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ย่อ ดังนี้

คำสั่งหลักในโปรแกรมไมโครแคท (MicroCAT) รายการคำสั่งหลักจะมีคำอธิบายคำสั่งโดย

| | |
|---------------|---|
| ITEM BANKER | เป็นคำสั่งที่ใช้ในการพัฒนาคลังข้อสอบ เพื่อใช้ในการสร้างและพัฒนาแบบสอบ |
| COMPUTER TEST | เป็นคำสั่งที่ใช้ในรูปแบบการทดสอบด้วยคอมพิวเตอร์ สำหรับการดำเนินการและการให้คะแนนการทดสอบด้วยคอมพิวเตอร์ |
| ITEM ANALYSIS | เป็นคำสั่งที่ใช้ในการวิเคราะห์ข้อสอบและแบบสอบ |
| PRINT | เป็นคำสั่งที่ใช้ในการพิมพ์แบบสอบประเพณีนิยมและคลังข้อสอบ |
| QUIT | เป็นคำสั่งที่ออกจากโปรแกรม |

รายการคำสั่งหลักประกอบด้วยรายการคำสั่งย่อย ดังนี้

รายการคำสั่ง "การพัฒนาคลังข้อสอบ" ประกอบด้วย 10 คำสั่งย่อย ดังนี้

| | |
|-------------|--|
| ITEM BANKER | เป็นคำสั่งที่ใช้ในการพัฒนาคลังข้อสอบแบบกราฟฟิก (graphic) โดยใช้ในการเพิ่มและแก้ไขข้อสอบ |
| MAKE FONT | เป็นคำสั่งที่ใช้ในการสร้างรูปแบบตัวอักษร ซึ่งสามารถสร้างรูปแบบตัวอักษรได้ 124 แบบ |
| CREAT | เป็นคำสั่งที่ใช้ในการสร้างแบบสอบโดยใช้ template |
| EDIT fname | เป็นคำสั่งที่ใช้ในการแก้ไขไฟล์ |
| IMPORT | เป็นคำสั่งที่เปลี่ยนข้อสอบที่เป็นเท็กซ์ไฟล์ (textfile) ให้มาอยู่ในระบบคลังข้อสอบของ MicroCAT |
| EXPORT | เป็นการสร้างข้อสอบในรูปแบบเท็กซ์ไฟล์ (textfile) |
| UTILITY | เป็นคำสั่งอรรถประโยชน์ |
| CONVERT | เป็นคำสั่งที่เปลี่ยนไฟล์ (file) รูปแบบกราฟฟิก (graphic) เป็นที่ไอเอฟ (TIF) |
| TESTBUILD | เป็นคำสั่งที่ใช้ในการสร้างแบบสอบแบบประเพณีนิยม |
| QUIT | เป็นคำสั่งที่ใช้ในการกลับไปยังรายการคำสั่งหลัก |

รายการคำสั่ง “รูปแบบการทดสอบด้วยคอมพิวเตอร์” เข้ามาโดยการเลือกรายการคำสั่ง COMPUTER TEST ที่รายการคำสั่งหลัก ประกอบด้วย 4 คำสั่งย่อย ดังนี้

| | |
|----------|---|
| TESTONE | เป็นคำสั่งที่ใช้ในการสั่งให้มีการทดสอบครั้งเดียวและจะทำการบันทึกคะแนนไว้ ส่วนรายการคำสั่ง TESTMANY สามารถพิมพ์คำสั่งได้ที่ DOS prompt ซึ่งเป็นคำสั่งที่ใช้ในการทดสอบกับผู้สอบหลายคน และในแต่ละครั้งจะมีการบันทึกคะแนน |
| PASSWORD | เป็นคำสั่งที่ใช้ในการเปลี่ยนรหัสผ่าน |
| QUIT | เป็นคำสั่งที่ใช้ในการกลับไปยัง รายการคำสั่งหลัก |

รายการคำสั่ง “การวิเคราะห์ข้อสอบและแบบสอบ” ประกอบด้วย 7 คำสั่งย่อย ดังนี้

| | |
|------------|--|
| ANALYSIS | เป็นคำสั่งที่ใช้ในการวิเคราะห์ข้อสอบและแบบสอบ |
| REESTIMATE | เป็นคำสั่งที่ใช้ในการประมาณค่าพารามิเตอร์ ในรูปแบบการวิเคราะห์แบบราสซีโมเดล |
| ESTIMATE | เป็นคำสั่งที่ใช้ในการประมาณค่า 2 – 3 พารามิเตอร์ในรูปแบบการวิเคราะห์โลจิสติก 2 – 3 พารามิเตอร์ |
| EVALUATE | เป็นคำสั่งที่ใช้ในการประเมินคุณภาพของแบบสอบก่อนที่จะมีการดำเนินการทดสอบ |
| VALIDATE | เป็นคำสั่งที่ใช้ในการคำนวณค่าสหสัมพันธ์และสมการการถดถอยพหุคูณ |
| SCORE | เป็นคำสั่งที่ใช้ในการคำนวณค่าความสามารถสำหรับการทดสอบ |
| QUIT | เป็นคำสั่งที่ใช้ในการกลับไปยัง รายการคำสั่งหลัก |

รายการคำสั่ง “พิมพ์เอกสาร” ประกอบด้วยคำสั่งย่อยดังต่อไปนี้

| | |
|-----------------|--|
| PRINT TEST/BANK | เป็นคำสั่งที่ใช้ในการสั่งพิมพ์แบบสอบประเพณีนิยมซึ่งประกอบด้วยข้อสอบที่บรรจุอยู่ในคลังข้อสอบ นอกจากนี้สามารถสั่งพิมพ์คลังข้อสอบทั้งหมดได้ด้วย |
| QUIT | เป็นคำสั่งที่ใช้ในการกลับไปยัง รายการคำสั่งหลัก |

ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ มีผู้ให้ความสนใจวิธีดำเนินการทดสอบ ทำการศึกษา ค้นคว้า และพัฒนารูปแบบของการทดสอบเพื่อให้เกิดการทดสอบที่มีประสิทธิภาพ ซึ่งผลการวิจัย มีดังนี้

จิราพร ไกรสรศิวเวท (2529) ได้ศึกษาความสัมพันธ์ระหว่างรูปแบบบางประการของการทดสอบเทเลอรรูปปรีรามิดกับความสามารถทางการเรียน โดยใช้การทดสอบเทเลอรรูปปรีรามิดที่มี 8 ชั้น 3 แบบ คือ แบบที่ใช้ขนาดชั้นคงที่ ขนาดชั้นแปรผัน และแบบข้างตัด และใช้วิธีการให้คะแนน 3 วิธี คือ ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ได้ทำ ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก และให้คะแนนทุกข้อ โดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 348 คน ผลการศึกษาพบว่า

ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการทดสอบเทเลอรรูปปรีรามิดทั้ง 3 แบบ โดยให้คะแนนทั้ง 3 วิธี กับคะแนนจากการทดสอบแบบดั้งเดิม มีความสัมพันธ์กันในทางบวกอย่างมีนัยสำคัญทางสถิติทุกค่า และเมื่อพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ภายใน ระหว่างคะแนนที่ได้จากการทดสอบเทเลอรรูปปรีรามิดทั้ง 3 แบบ ให้คะแนนทั้ง 3 วิธีปรากฏว่ามีค่าสหสัมพันธ์เป็นบวกอย่างมีนัยสำคัญทางสถิติทุกค่า

นันทิยา พึ่งคำ (2531) ได้ทำการเปรียบเทียบคุณภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับการทดสอบแบบเดิม กลุ่มตัวอย่างได้แก่ นักเรียนชั้น ม.3 จำนวน 120 คน แบบทดสอบแบบเดิมมีข้อสอบ 40 ข้อ ส่วนข้อสอบในคลังข้อสอบที่ใช้สำหรับการทดสอบแบบปรับเหมาะนั้นมี 361 ข้อ ข้อสอบทั้งสองลักษณะวัดความสามารถด้านคำศัพท์ภาษาอังกฤษ ผลการศึกษาพบว่า

ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่าและใช้ข้อสอบน้อยกว่าการทดสอบแบบเดิม แต่ยังคงให้ความตรงในการประมาณค่าความสามารถเหมือนเดิม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เสรี ปรมชวลิตโรจน์ (2531) ได้ศึกษาความตรงในการจำแนกความรอบรู้ และการประมาณค่าความสามารถในวิชาคณิตศาสตร์ โดยใช้แบบสอบแฟล็กซีเลเวลจำนวน 2 ฉบับ และแบบสอบรูปปริมาตรขนาดชั้นแปรผันจำนวน 2 ฉบับ โดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษา ปีที่ 3 จำนวน 5,987 คน ผลการศึกษาพบว่า

ความตรงในการจำแนกความรอบรู้ เมื่อใช้ข้อสอบที่สร้างขึ้นเป็นเกณฑ์ ปรากฏว่า ข้อสอบแบบแฟล็กซีเลเวล และรูปปริมาตรแบบขนาดชั้นแปรผันทุกฉบับ มีความตรงในการจำแนกความรอบรู้ได้ และเมื่อเปรียบเทียบค่าความตรงแต่ละแบบ พบว่า แบบสอบแบบแฟล็กซีเลเวลทั้งสองฉบับ มีความตรงในการจำแนกความรอบรู้ได้สูงกว่าการทดสอบรูปปริมาตรอยู่หนึ่งฉบับ

เกรียงศักดิ์ สุวรรณภาค (2532) ได้ศึกษาความสัมพันธ์ระหว่างรูปแบบของแบบทดสอบแยกกลุ่ม กับความสามารถทางการเรียนในวิชาคณิตศาสตร์โดยทำการศึกษาในตัวแปร 2 ตัว คือ รูปแบบของแบบทดสอบ และวิธีการให้คะแนน โดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 2,455 คน ผลการศึกษาพบว่า

ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการสอบที่ใช้แบบทดสอบแยกกลุ่มแต่ละฉบับ ซึ่งตรวจให้คะแนนในแต่ละวิธี กับความสามารถทางการเรียน มีค่าเป็นบวกอย่างมีนัยสำคัญทางสถิติทุกค่า และผลการทดสอบความแตกต่างของค่าสหสัมพันธ์ดังกล่าวเป็นรายคู่ พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติทุกคู่

ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจให้คะแนนทั้ง 2 วิธีดังกล่าว โดยแยกวิเคราะห์ตามแบบทดสอบแยกกลุ่มแต่ละแบบ มีค่าเป็นบวกอย่างมีนัยสำคัญทางสถิติทุกค่า และผลการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ดังกล่าว พบว่า เมื่อพิจารณาในแบบทดสอบแยกกลุ่มแต่ละแบบ ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจให้คะแนนทั้ง 2 วิธีมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติทุกคู่

ต่าย เชียงฉวี (2534) ได้ทำการเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบเทเลอร์รูปปริมาตรที่มีรูปแบบ จำนวนชั้น และวิธีการให้คะแนนที่แตกต่างกันโดยใช้วิธีมอนติคาร์โล ผลการศึกษาพบว่า

การประมาณค่าความสามารถตามวิธีการของเบย์ส์ (bayesian) และวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (condition maximum likelihood) นั้นให้ผลการประมาณที่ไม่แตกต่างกัน การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบรูปปริมาตรที่มีรูปแบบ จำนวนชั้นและวิธีการให้คะแนนที่ต่างกันจะส่งผลต่อ

ความคลาดเคลื่อนมาตรฐานในการประมาณค่า นอกจากนี้ยังพบว่าค่าความตรงเชิงโครงสร้างของแบบสอบพีระมิดแบบขนาดขั้นคงที่มี 16 ขั้น ประมาณค่าความสามารถด้วยวิธีของเบสส์มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงสุด

ตริงใจ พูลผลอำนาจ (2534) ได้พัฒนาแบบสอบเฉพาะบุคคลในวิชาวิทยาศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้คอมพิวเตอร์ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 4,492 คน ผลการศึกษาพบว่า

แบบสอบเฉพาะบุคคลที่ทดสอบโดยใช้คอมพิวเตอร์ และผู้สอบแต่ละคนทำ 10 ข้อ มีความสัมพันธ์กับแบบสอบผลสัมฤทธิ์วิชาวิทยาศาสตร์ทั้งฉบับ 55 ข้อ ซึ่งใช้การเขียนตอบ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.7818 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบสอบเฉพาะบุคคลเท่ากับ 0.443

สายชล อบทม (2538) ได้พัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่ใช้รูปแบบทางแยกแปรผัน การประมาณค่าความสามารถของผู้สอบโดยวิธีของเบสส์ ยุติการทดสอบเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 ประกอบด้วยโปรแกรมย่อย 2 โปรแกรม คือ โปรแกรมย่อยการพิมพ์ข้อสอบและโปรแกรมย่อยการทดสอบที่ปฏิบัติการบนระบบวินโดวส์ 95

รังสรรค์ มณีเล็ก (2540) ทำการศึกษาผลของความยากง่ายของข้อสอบข้อแรกที่ใช้ในการทดสอบ เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ วิธีการประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการทดสอบและความสามารถของผู้สอบต่อความตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ผลการศึกษาพบว่า

วิธีการประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการทดสอบ และความสามารถของผู้สอบมีผลต่อค่าความตรงเชิงสภาพของการทดสอบ และความยากง่ายของข้อสอบข้อแรก เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ วิธีการประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการทดสอบและความสามารถของผู้สอบต่างก็มีผลต่อจำนวนข้อสอบที่ใช้ในการทดสอบ

เออร์รี่ (Urry, 1977) พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนั้นจะไม่ค่อยมีประสิทธิภาพหากเลือกใช้โมเดลการวิเคราะห์ข้อมูลไม่ค่อยเหมาะสม ซึ่งโมเดลโลจิสติก 3 พารามิเตอร์มีความเหมาะสมมากกว่าโมเดลอื่นๆ และ ขนาดของคลังข้อสอบนั้นควรมีอย่างน้อย 100 ข้อ ส่วนข้อสอบในคลังข้อสอบนั้น ควรมีค่าอำนาจจำแนกมากกว่า .8 ค่าความยากง่ายอยู่ระหว่าง -2.0 ถึง 2.0 และค่าการเดาควรมีน้อยกว่า .3

แบล็กมอร์ (Blackmore, 1987) ได้ศึกษาเปรียบเทียบการทดสอบที่ดำเนินการสอบด้วยคอมพิวเตอร์ (Computerize) การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (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 การทดลองใช้แบบสอบวัดความถนัดมี 6 ฉบับย่อย คือ การให้เหตุผลทางภาษา ความสามารถทางตัวเลข การให้เหตุผลเชิงนามธรรม การให้เหตุผลเชิงกล มิติสัมพันธ์ และการใช้ภาษา แต่ละกลุ่มวิเคราะห์ข้อมูลโดยใช้การทดสอบความแปรปรวนและการทดสอบแม็คเนมาร์ (McNemar test) ผลการศึกษาพบว่า

คะแนนจากข้อสอบย่อยทุกฉบับจากการดำเนินการสอบทั้ง 3 วิธี ไม่แตกต่างกัน แต่จะ แตกต่างกันตามคุณลักษณะของผู้เข้าสอบ เช่น เชื้อชาติ และเพศ จากการใช้แบบสอบถามและสัมภาษณ์อย่างมีนัยสำคัญทางสถิติ

โฮ (Ho, 1988) ได้เปรียบเทียบวิธีการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ 3 วิธี คือ ประเมินค่าความสามารถของผู้สอบโดยวิธีของเบย์ (Bayesian) วิธีโมดัลเบย์ (modal bayesian) และวิธีความเป็นไปได้สูงสุด (maximum likelihood) โดยใช้โปรแกรมในการทดสอบนี้กับคลังข้อสอบ 3 ขนาด คือ 86, 71 และ 56 ข้อ ลักษณะของคลังข้อสอบมี 3 แบบ คือ ข้อสอบได้มาโดยการสุ่มข้อสอบที่มีค่าอำนาจจำแนกสูงๆ และข้อสอบที่ง่าย ทดลองกับผู้สอบที่มีความสามารถ 3 ระดับ คือ สูง กลาง และต่ำ จุดเริ่มต้นในการทดสอบนั้นใช้จุดเริ่มต้นเดียวกัน วิเคราะห์ข้อมูลโดยการวิเคราะห์ความแปรปรวน 3 ทาง แบบวัดซ้ำ ซึ่งตัวแปรตาม ได้แก่ จำนวนข้อสอบที่ใช้ ความแปรปรวนของความคลาดเคลื่อนเฉลี่ย และค่าสัมบูรณ์ของความแตกต่างระหว่างค่าความสามารถที่คาดหวังกับค่าความสามารถโดยประมาณ ซึ่งตัวแปรตาม 2 ตัวแรกนั้นใช้สำหรับประเมินประสิทธิภาพของการทดสอบ ตัวที่ 3 ใช้ประเมินความถูกต้องของการทดสอบ ผลการศึกษาพบว่า

- 1) วิธีการประมาณค่าแบบโมเดลเบย์ส (model bayesian) มีประสิทธิภาพสูงสุดวิธีการประมาณค่าของเบย์ส (Bayesian) มีความเที่ยงสูง แต่วิธีการประมาณค่าความเป็นไปสูงสุด (maximum likelihood) ไม่คงเส้นคงวา
- 2) ขนาดของคลังข้อสอบมีผลน้อยมาก
- 3) การประมาณค่าความสามารถจะมีความถูกต้องมากขึ้น ถ้าข้อสอบที่ใช้เริ่มต้นสอบนั้นมีค่าความยากง่ายเท่ากับหรือต่ำกว่าความสามารถจริงของผู้สอบ
- 4) มีปฏิสัมพันธ์ระหว่างชนิดของคลังข้อสอบกับวิธีการประมาณค่าความสามารถและชนิดของคลังข้อสอบกับระดับความสามารถของผู้สอบ
- 5) มีความแตกต่างกันน้อยมากระหว่างคุณลักษณะทางสถิติของข้อสอบที่ไม่เคยถูกเลือกใช้กับข้อสอบที่ถูกเลือกใช้บ่อย

แฮนกินส์ (Hankins, 1986) ได้ทำการศึกษาผลของการกำหนดข้อสอบเริ่มต้นแน่นอน และกำหนดข้อสอบเริ่มต้นแปรผัน ต่อค่าความลำเอียงและสารสนเทศของการทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบ ผลการศึกษาพบว่า

ทั้ง 2 วิธี ไม่มีความลำเอียงในการประมาณค่าความสามารถของผู้สอบที่มีความสามารถปานกลาง แต่จะมีความลำเอียงถ้าผู้สอบมีความสามารถสูงหรือต่ำมาก ๆ ทั้ง 2 วิธี ให้ได้สารสนเทศเป็นรูปสมมาตรที่โด่ง ซึ่งแสดงให้เห็นว่าการทดสอบด้วยวิธีการกำหนดข้อสอบเริ่มต้นที่ต่างกันมีประสิทธิภาพเท่าเทียมกัน แต่ทั้ง 2 วิธี เลือก ข้อสอบได้ต่างกัน การกำหนดข้อสอบเริ่มต้นแบบแปรผันใช้จำนวนข้อสอบมากกว่า

คาลอห์น และสเปร (C. Kalohn and A. Spray, 1998) ได้ทำการศึกษาผลกระทบของการคัดเลือกข้อสอบต่ออัตราการใช้ข้อสอบซ้ำในการทดสอบด้วยคอมพิวเตอร์ โดยมีสมมุติฐานการวิจัย คือ อัตราการใช้ข้อสอบซ้ำมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แต่ผลกระทบของความแตกต่างมีเล็กน้อย การศึกษานี้เป็นการจำลองข้อมูล คลังข้อสอบมีจำนวนข้อสอบ 1,235 ข้อ อัตราการใช้ข้อสอบซ้ำจะถูกคำนวณภายใต้เงื่อนไข 2 เงื่อนไข คือ อัตราการใช้ข้อสอบซ้ำ 20% และ 10% ผลการศึกษาพบว่า

วิธีการคัดเลือกข้อสอบทั้งสองเงื่อนไขจะมีความถูกต้องเหมือนกัน แต่ถ้าข้อสอบในคลังข้อสอบมีจำนวนมากซึ่งมีสารสนเทศน้อย ควรใช้อัตราการใช้ข้อสอบซ้ำ 10%

จาไวเออร์ และ ไวเซนท์ (Javier and Vicent , 1998) ได้ทำการเปรียบเทียบวิธีการควบคุมอัตราการใช้ข้อสอบซ้ำในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ โดยใช้คอมพิวเตอร์ ผลการศึกษาพบว่า

Restricted Method ใช้ในการลดอัตราการใช้ข้อสอบซ้ำได้สูงสุด และ Progressive method ลดจำนวนข้อสอบภายในคลังข้อสอบที่ยังไม่ได้ใช้ ซึ่งการควบคุมอัตราการใช้ข้อสอบซ้ำมีจุดมุ่งหมายที่สำคัญ 2 ประการ คือ เพื่อป้องกันอัตราการใช้ข้อสอบซ้ำที่มากเกินไป และ เพิ่มอัตราการใช้ข้อสอบที่ยังไม่ได้ใช้ภายในคลังข้อสอบ

ซาง และทวู (W. Chang and Y. Twu, 1998) ได้ทำการเปรียบเทียบวิธีการควบคุมอัตราการใช้ข้อสอบซ้ำในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยทำการเปรียบเทียบวิธีการควบคุมอัตราการใช้ข้อสอบซ้ำ 5 วิธี ในการประมาณค่าความสามารถของผู้สอบ การควบคุมอัตราการใช้ข้อสอบซ้ำแต่ละวิธีจะดำเนินการในขบวนการของการคัดเลือกข้อสอบและการทดสอบแบบซี เอ ที (CAT) การศึกษานี้เป็นการจำลองข้อมูลโดยใช้วิธีมอนติคาโร (Monte Carlo simulation) โดยการพิจารณาภายใต้ขนาดของคลังข้อสอบที่แตกต่างกัน อัตราการใช้ข้อสอบซ้ำสูงสุดที่แตกต่างกัน และความคลาดเคลื่อนมาตรฐานของการวัด ผลการศึกษาพบว่า

วิธีการแบบมีเงื่อนไขที่พัฒนาโดย สต็อกกิ้ง และเลวิส (M. Stocking and C. Lewis, 1995) เป็นวิธีการที่ดีที่สุด มีความถูกต้องแม่นยำของการวัด

จากผลงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบทั้งในประเทศและต่างประเทศ พบว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีคุณภาพไม่แตกต่างกับการทดสอบแบบประเพณีนิยม แต่ใช้จำนวนข้อสอบและเวลาในการทดสอบน้อยกว่า จะเห็นได้ว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีประโยชน์ในด้านอื่นๆ อีก เช่น ทำให้ผู้สอบมีเจตคติที่ดี รวมทั้งให้ผู้สอบมีแรงจูงใจ ความตั้งใจ และความสนใจในการสอบมากยิ่งขึ้น และมีความเป็นไปได้ที่จะนำเอาการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ หรือการทดสอบแบบซีเอที (CAT) มาใช้แทนการทดสอบแบบประเพณีนิยม

จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง เพื่อศึกษาผลที่เกิดขึ้นจากตัวแปรอิสระ 2 ตัวแปร คือ ระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำ ที่มีต่อความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์หรือการทดสอบแบบซี เอ ที (CAT) โดยมีวิธีดำเนินการวิจัยดังนี้

ประชากรและกลุ่มตัวอย่าง

ประชากร ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2543 ใน โรงเรียนมัธยมศึกษาขนาดใหญ่ สังกัดกรมสามัญศึกษาจังหวัดสุพรรณบุรี จำนวน 8 โรงเรียน จำนวนนักเรียน 3,476 คน

กลุ่มตัวอย่าง ได้มาจากประชากรที่มีการสุ่มแบบหลายขั้นตอน (multi stage random sampling) โดยมีขั้นตอนในการสุ่มดังนี้

ขั้นตอนแรก สุ่มโรงเรียน 6 โรงเรียนจากโรงเรียนมัธยมศึกษาขนาดใหญ่ สังกัด กรมสามัญศึกษาจังหวัดสุพรรณบุรี ซึ่งมีทั้งหมด 8 โรงเรียน โดยการสุ่มอย่างง่าย (simple random sampling) ผู้วิจัยใช้แบบสอบถามวัดความสามารถทางภาษาอังกฤษทดสอบนักเรียนทุกคน จำแนกเป็นรายโรงเรียนดังตารางที่ 1

ตารางที่ 1 จำนวนกลุ่มตัวอย่าง จำแนกเป็นรายโรงเรียน

| ลำดับที่ | โรงเรียน | จำนวน (คน) |
|----------|-------------------------------|------------|
| 1 | กรรณสูตศึกษาลัย | 477 |
| 2 | บรรหารแจ่มใสวิทยา 1 | 351 |
| 3 | ธรรมโชติศึกษาลัย | 331 |
| 4 | สามชุกรัตนโกศาราม | 423 |
| 5 | บรรหารแจ่มใสวิทยา 3 | 416 |
| 6 | บางปลาม้า “สูงสูมารมดุงวิทย์” | 459 |
| รวม | | 2,457 |

ขั้นตอนที่สอง สุ่มนักเรียนที่ได้จากการสุ่มในขั้นตอนที่หนึ่ง โดยการสุ่มแบบแบ่งชั้น (stratified random sampling) ซึ่งใช้ระดับความสามารถทางภาษาอังกฤษเป็นชั้นของการสุ่ม ได้แก่ ความสามารถต่ำ ปานกลาง และสูง ผู้วิจัยต้องการกลุ่มตัวอย่างกลุ่มละ 120 คน รวมกลุ่มตัวอย่าง ทั้งหมดที่สุ่มได้ 360 คน จำแนกเป็น โรงเรียนได้ดังตารางที่ 2

ตารางที่ 2 จำนวนนักเรียนกลุ่มตัวอย่าง จำแนกเป็นรายโรงเรียน และตามระดับความสามารถ

| ลำดับที่ | โรงเรียน | ระดับความสามารถ | | | จำนวนนักเรียน |
|----------|-------------------------------|-----------------|---------|-----|---------------|
| | | ต่ำ | ปานกลาง | สูง | |
| 1 | กรรณสูตศึกษาลัย | 20 | 20 | 20 | 60 |
| 2 | บรรหารแจ่มใสวิทยา 1 | 20 | 20 | 20 | 60 |
| 3 | ธรรมโชติศึกษาลัย | 20 | 20 | 20 | 60 |
| 4 | สามชุกรัตนโกคาราม | 20 | 20 | 20 | 60 |
| 5 | บรรหารแจ่มใสวิทยา 3 | 20 | 20 | 20 | 60 |
| 6 | บางปลาหมอ “สูงสูมารมดุจวิทย์” | 20 | 20 | 20 | 60 |
| | รวม | 120 | 120 | 120 | 360 |

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย มีดังนี้

1. แบบสอบวัดความสามารถทางภาษาอังกฤษ ระดับชั้นมัธยมศึกษาตอนต้น เป็นแบบสอบวัดผลสัมฤทธิ์ทางการเรียน ซึ่งพัฒนาโดยกรมวิชาการ กระทรวงศึกษาธิการ (2540) เพื่อใช้ในการประเมินคุณภาพทางการศึกษา ลักษณะของแบบสอบเป็นข้อสอบเลือกตอบ 4 ตัวเลือก จำนวน 50 ข้อ เครื่องมือชุดนี้ไว้สำหรับทดสอบนักเรียนที่เป็นกลุ่มตัวอย่าง เพื่อแบ่งออกเป็น 3 กลุ่ม คือ กลุ่มที่มีความสามารถทางภาษาอังกฤษต่ำ ปานกลาง สูง และใช้คะแนนจากแบบสอบฉบับนี้ในการศึกษาความตรงตามสภาพด้วย โดยแบบสอบมีโครงสร้างของสมรรถภาพที่วัดและจำนวนข้อสอบ ดังตารางที่ 3

ตารางที่ 3 โครงสร้างของข้อสอบวัดความสามารถทางภาษาอังกฤษ

| สมรรถภาพ / สมรรถภาพย่อย | จำนวนข้อ | ข้อที่ |
|--|----------|----------------|
| 1. ความสามารถในการส่งสาร | 30 | |
| 1.1 ใช้ภาษาพูดได้ตามสถานการณ์ที่กำหนด | 7 | 1-7 |
| 1.2 ใช้ภาษาพูดได้ตามกาลเทศะและวัฒนธรรมการใช้ภาษา | 8 | 8-15 |
| 1.3 ใช้คำที่มีความหมายทางไวยากรณ์ได้ในข้อความ | 5 | 16-20 |
| 1.4 ใช้ภาษาเขียนได้ตามสถานการณ์ที่กำหนด | 10 | 21-30 |
| 2. ความสามารถในการรับสาร | 20 | |
| 2.1 ระบุรายละเอียดของสารที่กำหนด | 8 | 31-37,47 |
| 2.2 ระบุความหมายของคำในสารที่กำหนด | 2 | 41-42 |
| 2.3 ตีความจากสารที่กำหนด | 4 | 44,46,49-50 |
| 2.4 สรุปความของสารที่กำหนด | 6 | 38-40,43,45,48 |
| รวม | 50 | |

ผู้วิจัยนำแบบสอบวัดความสามารถทางภาษาอังกฤษมาทดลองใช้ โดยนำไปทดสอบกับ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 โรงเรียนสงวนหญิง และโรงเรียนบางลิ่ววิทยา จังหวัดสุพรรณบุรี จำนวน 254 คน นำผลการทดสอบมาวิเคราะห์หาคุณภาพของแบบสอบโดยใช้โปรแกรมการทดสอบไมโครแคท (MicroCAT) รุ่นที่ 3 (Version 3) ได้ผลการวิเคราะห์ ดังตารางที่ 4

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4 ผลการวิเคราะห์แบบสอบวัดความสามารถทางภาษาอังกฤษ

| รายการ | ค่าสถิติ |
|--------------------------------|----------|
| คะแนนเต็ม | 50 |
| ค่ามัธยิมเลขคณิต (\bar{X}) | 27.409 |
| ส่วนเบี่ยงเบนมาตรฐาน (SD) | 8.038 |
| ค่าต่ำสุด (Min) | 12 |
| ค่าสูงสุด (Max) | 48 |
| พิสัย (Rang) | 36 |
| มัธยฐาน (Med) | 27 |
| ความเบ้ (Sk) | 0.354 |
| ความโด่ง (Ku) | -0.448 |
| ความคลาดเคลื่อนมาตรฐาน (SEM) | 3.099 |
| ความเที่ยง (alpha) | 0.851 |
| ค่าอำนาจจำแนก (a) | 1.3487 |
| ค่าความยาก (b) | .4312 |
| ค่าการเดา (c) | .25 |

จากตารางที่ 4 พบว่าแบบสอบวัดความสามารถทางภาษาอังกฤษ มีคะแนนเต็ม 50 คะแนน โดยมีคะแนนต่ำสุดเท่ากับ 12 คะแนน และมีคะแนนสูงสุดเท่ากับ 48 คะแนน มีค่าเฉลี่ยของคะแนนเท่ากับ 27.409 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 8.038 ลักษณะการกระจายของคะแนนมีลักษณะเบ้ขวาเล็กน้อย และมีความโด่งเท่ากับ -0.448 แบบสอบมีความเที่ยงเท่ากับ .85 มีค่าความคลาดเคลื่อนมาตรฐานในการวัดเท่ากับ 3.099 แบบสอบมีค่าอำนาจจำแนกเท่ากับ 1.3487 มีค่าความยากเท่ากับ .4312 และค่าการเดาเท่ากับ .25

2. คลังข้อสอบ เป็นกลุ่มข้อสอบที่ใช้วัดความรู้ความเข้าใจในวิชาภาษาอังกฤษ ระดับชั้นมัธยมศึกษาตอนต้น ในโครงการวิจัยเรื่องการพัฒนาคลังข้อสอบรายวิชาตามหลักสูตรมัธยมศึกษา ซึ่งพัฒนาโดยแมนมาศ ลีลสัตยกุล (2540) ลักษณะของแบบสอบเป็นข้อสอบแบบเลือกตอบ 4 ตัวเลือก จำนวน 231 ข้อ มีค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.400 ถึง 2.500 ค่าความยาก (b) อยู่ระหว่าง -

มีค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.400 ถึง 2.500 ค่าความยาก (b) อยู่ระหว่าง -3.000 ถึง 3.000 และค่าการเดา (c) อยู่ระหว่าง 0.020 ถึง 0.370 โดยมีจำนวนข้อสอบจำแนกตามจุดประสงค์การวัด ดังตารางที่ 5

ตารางที่ 5 โครงสร้างของคลังข้อสอบวิชาภาษาอังกฤษ จำแนกตามจุดประสงค์การวัด

| ที่ | จุดประสงค์ | จุดประสงค์เชิงพฤติกรรม | ข้อที่ | จำนวนข้อ |
|-----|--------------|--|--|----------|
| 1 | ทักษะการพูด | | | 59 |
| | | 1. เพื่อทดสอบความสามารถในการจับใจความสำคัญ | 8-9 | 2 |
| | | 2. เพื่อทดสอบความสามารถในการระบุรายละเอียดที่สำคัญ | 1-7 140-145 | 13 |
| | | 3. เพื่อทดสอบความรู้ความสามารถของนักเรียนด้านการพูดให้ถูกต้องเหมาะสม | 10-23 110-129 , 146-155 | 44 |
| 2 | ทักษะการอ่าน | | | 50 |
| | | 1. เพื่อทดสอบความเข้าใจความหมายของคำหรือข้อความในบริบท | 210,211,216, 221,226 | 5 |
| | | 2. เพื่อทดสอบความสามารถในการถ่ายโอนความรู้จากสิ่งพิมพ์หรือข้อมูลที่น่าเสนอด้วยรูปแบบต่างๆ | 177,183,193,203, 212,222,227 | 7 |
| | | 3. เพื่อทดสอบความเข้าใจเนื้อเรื่องที่เป็นใจความสำคัญของเรื่องที่อ่าน | 174-175,178-182, 194-195,198-202, 204-207,209,214 217-218,225 | 24 |
| | | 4. เพื่อทดสอบความสามารถในการอ่านเพื่อหารายละเอียดของเรื่องที่อ่าน | 176,196,208,213, 223,228-231 | 9 |
| | | 5. เพื่อทดสอบความเข้าใจ ความหมายของสิ่งที่อ่านจากข้อความที่มีความหมายแฝงอยู่ในลักษณะคำพูด สำนวน ความเห็น | 215,224 | 2 |
| | | 6. เพื่อทดสอบความสามารถในการสรุปความเรื่องที่อ่าน | 197,219,220 | 3 |

ตารางที่ 5 (ต่อ)

| ที่ | จุดประสงค์ | จุดประสงค์เชิงพฤติกรรม | ข้อที่ | จำนวนข้อ |
|-----|---------------|---|-----------------------------------|----------|
| 3 | ทักษะการเขียน | เพื่อทดสอบความรู้ ความสามารถของ ผู้เรียนเกี่ยวกับความหมายของคำศัพท์ที่ ใช้ในบริบทต่าง ๆ | 24-41,86-109, 156-161,164-173 | 58 |
| 4 | ไวยากรณ์ภาษา | เพื่อทดสอบความรู้ ความสามารถเกี่ยวกับ กฎเกณฑ์และการใช้ไวยากรณ์ภาษาของ นักเรียน | 42-85,130-139, 162-163,184-192 | 65 |

3. โปรแกรมสำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ คือ โปรแกรมการทดสอบไมโครแคท (MicroCAT) รุ่นที่ 3 (Version 3) พัฒนาโดย Assessment System Corporation (1995) กำหนดรูปแบบของการทดสอบเป็นหลายขั้นตอน (multi – stage strategies) แบบทางแยกแปรผัน (variable branching model) โดยเริ่มต้นการทดสอบจากการคัดเลือกข้อสอบที่มีค่าความยากระดับปานกลาง ประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบย์ส์ (Bayesian) การคัดเลือกข้อสอบข้อต่อไป โดยดูค่าสารสนเทศของข้อสอบข้อที่สูงที่สุด (maximum information) และยุติการทดสอบโดยพิจารณาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่า 0.3 และได้เพิ่มเงื่อนไขเพื่อให้สอดคล้องกับงานวิจัย มี ทั้งหมด 9 เงื่อนไข ซึ่งได้จากตัวแปรอิสระ 2 ตัวแปร คือ

3.1 ระดับความสามารถทางภาษาอังกฤษของผู้สอบ จำแนกเป็น 3 ระดับ ได้แก่

- 1) ความสามารถต่ำ ($\theta \leq -0.75$)
- 2) ความสามารถปานกลาง ($-0.75 < \theta < 0.75$)
- 3) ความสามารถสูง ($\theta \geq 0.75$)

3.2 อัตราการใช้ข้อสอบซ้ำ 3 ระดับ กำหนดไว้เป็น 3 ระดับ ได้แก่

- 1) อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์
- 2) อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์
- 3) อัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์

เงื่อนไขของการทดสอบจึงประกอบด้วย 9 เงื่อนไข ดังตารางที่ 6

ตารางที่ 6 เงื่อนไขในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

| เงื่อนไข | ระดับความสามารถทาง ภาษาอังกฤษของผู้สอบ | อัตราการให้ข้อสอบซ้ำ (เปอร์เซ็นต์) |
|----------|---|---------------------------------------|
| 1 | ต่ำ | 10 |
| 2 | ต่ำ | 20 |
| 3 | ต่ำ | 30 |
| 4 | ปานกลาง | 10 |
| 5 | ปานกลาง | 20 |
| 6 | ปานกลาง | 30 |
| 7 | สูง | 10 |
| 8 | สูง | 20 |
| 9 | สูง | 30 |

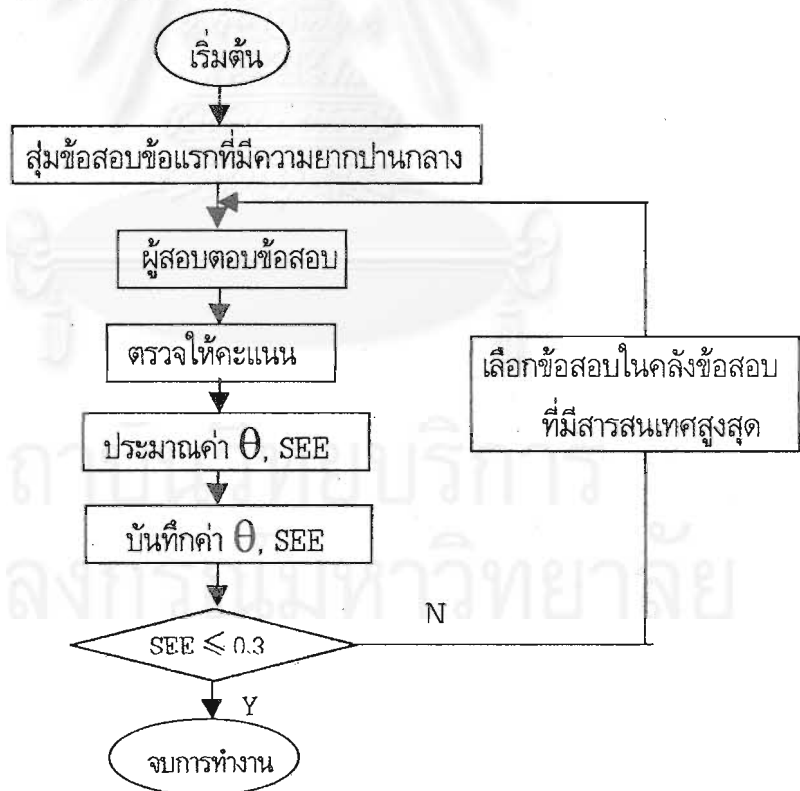
การพัฒนาโปรแกรมสำหรับการทดสอบ

เนื่องจากเครื่องมือที่ใช้ในการวิจัยประกอบด้วย แบบสอบวัดความสามารถทางภาษาอังกฤษ ระดับชั้นมัธยมศึกษาปีที่ 3 ซึ่งพัฒนาโดยกรมวิชาการ กระทรวงศึกษาธิการ (2540) คลังข้อสอบวิชาภาษาอังกฤษ ระดับมัธยมศึกษาตอนต้น ซึ่งพัฒนาโดยแมนมาศ ลีลสัตยกุล (2540) และโปรแกรมคอมพิวเตอร์ สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ได้แก่ โปรแกรมการทดสอบไมโครแคท (MicroCAT)

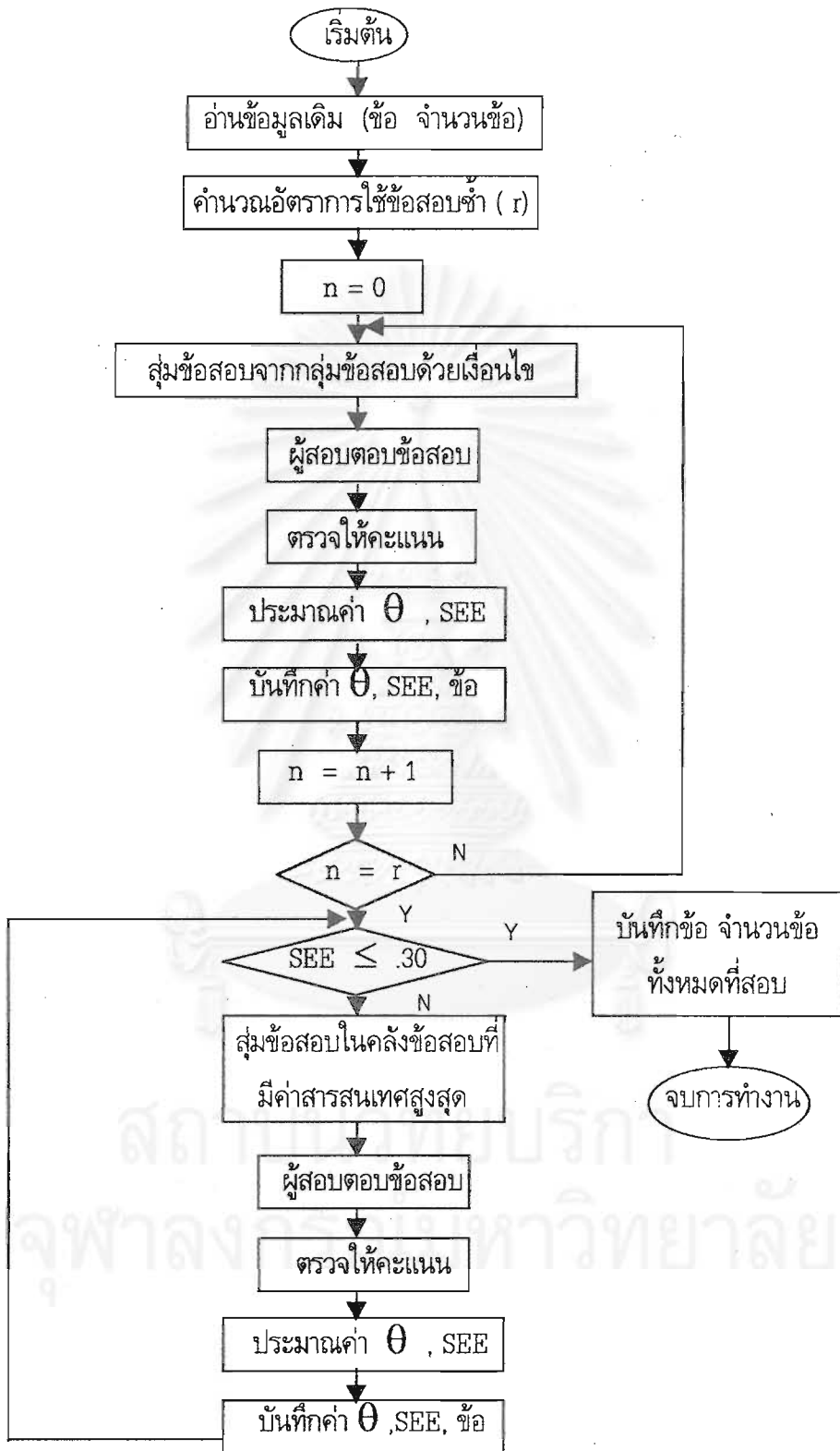
การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์หรือการทดสอบแบบซี เอ ที (CAT) นั้นดำเนินการทดสอบโดยให้ผู้สอบนั่งสอบหน้าจอคอมพิวเตอร์และตอบข้อสอบครั้งละ 1 ข้อ โดยคอมพิวเตอร์เป็นผู้เลือกข้อสอบจากคลังข้อสอบขึ้นมาให้ผู้สอบได้ตอบข้อสอบที่ถูกเลือกขึ้นมานั้นจะมีความเหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล ในการดำเนินการทดสอบลักษณะนี้จำเป็นต้องมีโปรแกรมสั่งให้คอมพิวเตอร์จัดกระทำกับข้อมูลในลักษณะต่างๆ ตามตัวแปรอิสระที่ต้องการศึกษา ซึ่งผู้วิจัยได้พัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบด้วยโปรแกรมการทดสอบไมโครแคท (MicroCAT) โดยใช้ภาษา MCATL (Minnesota Computerized Adaptive testing Language) ในการพัฒนาโปรแกรมมีขั้นตอนดำเนินการดังนี้

การทดสอบในรอบที่ 1 เมื่อผู้สอบเริ่มทำข้อสอบข้อแรกโดยการคัดเลือกข้อสอบที่มีระดับความยากปานกลาง เมื่อผู้สอบตอบก็ทำการประมาณค่าความสามารถ และความคลาดเคลื่อนมาตรฐานของการประมาณค่า เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีความเหมาะสมกับความสามารถของผู้สอบ กระบวนการนี้จะดำเนินต่อไปเรื่อยๆจนกระทั่งค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าน้อยกว่าหรือเท่ากับ 0.3 การทดสอบก็จะยุติลง ดังแสดงในแผนภาพที่ 15

การทดสอบในรอบที่ 2 - 5 ซึ่งเป็นการทดสอบที่มีอัตราการใช้ข้อสอบซ้ำ เริ่มต้นทำการอ่านข้อมูลเดิม ได้แก่ ข้อ และจำนวนข้อของการทดสอบในครั้งที่ผ่านมา เพื่อทำการคำนวณจำนวนข้อที่จะใช้ซ้ำ โดยผู้วิจัยทำการเขียนโปรแกรมในการอ่านข้อมูลนั้นด้วยโปรแกรม Visual Basic เริ่มต้นการทดสอบโดยการคัดเลือกข้อสอบจากการกลุ่มข้อสอบที่ได้รับในครั้งที่ผ่านมาตามจำนวนข้อที่ได้คำนวณไว้ ทำการประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐานของการประมาณค่า จากนั้นทำการคัดเลือกข้อสอบภายในคลังข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบ กระบวนการนี้จะดำเนินต่อไปเรื่อยๆจนกระทั่งค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าน้อยกว่าหรือเท่ากับ 0.3 การทดสอบก็จะยุติลง ดังแสดงในแผนภาพที่ 16



แผนภาพที่ 15 ผังงานแสดงขั้นตอนการทำงานของคอมพิวเตอร์ในการทดสอบรอบที่ 1



แผนภาพที่ 16 ผังงานแสดงขั้นตอนการทำงานของคอมพิวเตอร์ในการดำเนินการทดสอบแบบ
ปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ในรอบที่ 2 - 5

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลเพื่อศึกษาผลที่เกิดขึ้นต่อค่าความตรงตามสภาพ และความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้น ได้แบ่งออกเป็น 2 ขั้นตอน ขั้นตอนแรก เป็นการทดสอบความสามารถทางภาษาอังกฤษด้วยแบบสอบวัดความสามารถทางภาษาอังกฤษเพื่อจัดกลุ่มตัวอย่างออกเป็น 3 กลุ่ม คือ กลุ่มที่มีระดับความสามารถสูง ปานกลาง และต่ำ ขั้นตอนที่สอง เป็นการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 5 ครั้ง ซึ่งผู้วิจัยได้ดำเนินการดังนี้

1. นำหนังสือจากบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัยไปให้สำนักงานสามัญศึกษาจังหวัดสุพรรณบุรี เพื่อขอทดสอบนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในโรงเรียนที่ได้รับการสุ่มเป็นกลุ่ม ตัวอย่าง จำนวน 6 โรงเรียน

2. ประสานงานกับโรงเรียนที่สุ่มเป็นกลุ่มตัวอย่าง เพื่อชี้แจงวัตถุประสงค์ของการวิจัยและนัดหมายวัน เวลาเพื่อทดสอบนักเรียน

3. นำแบบสอบวัดความสามารถทางภาษาอังกฤษ ระดับชั้นมัธยมศึกษาตอนต้น มาทำการทดสอบนักเรียนที่เป็นกลุ่มตัวอย่าง

4. นำผลการตอบแบบสอบถามมาตรวจให้คะแนน และทำการประมาณค่าความสามารถของผู้สอบด้วยโปรแกรม BILOG 3 ทำการจัดแบ่งกลุ่มตัวอย่างตามระดับความสามารถออกเป็น 3 ระดับ ดังนี้

4.1 ความสามารถทางภาษาอังกฤษต่ำ ($\theta \leq -0.75$)

4.2 ความสามารถทางภาษาอังกฤษปานกลาง ($-0.75 < \theta < 0.75$)

4.3 ความสามารถทางภาษาอังกฤษสูง ($\theta \geq 0.75$)

5. เก็บคะแนนค่าความสามารถจากการทดสอบในข้อ 4 ไว้เพื่อนำไปใช้เป็นคะแนนเกณฑ์ในการศึกษาความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

6. สุ่มเงื่อนไขของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งมีทั้งหมด 9 เงื่อนไข ให้กับกลุ่มตัวอย่าง ดังตารางที่ 7

ตารางที่ 7 กลุ่มตัวอย่าง และเงื่อนไขในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

| เงื่อนไข | นักเรียนกลุ่มตัวอย่าง | จำนวน (คน) |
|----------|---|------------|
| 1 | โรงเรียนวรรณสุตศึกษาลัย และ โรงเรียนบรรหารแจ่มใสวิทยา 1 | 40 |
| 2 | โรงเรียนธรรมโชติศึกษาลัยและโรงเรียนสามชุกรัตนโกศาราม | 40 |
| 3 | โรงเรียนบรรหารแจ่มใสวิทยา 3 และโรงเรียนบางปลาม้า “สูงสูดผดุงวิทย์” | 40 |
| 4 | โรงเรียนวรรณสุตศึกษาลัย และ โรงเรียนบรรหารแจ่มใสวิทยา 1 | 40 |
| 5 | โรงเรียนธรรมโชติศึกษาลัยและโรงเรียนสามชุกรัตนโกศาราม | 40 |
| 6 | โรงเรียนบรรหารแจ่มใสวิทยา 3 และโรงเรียนบางปลาม้า “สูงสูดผดุงวิทย์” | 40 |
| 7 | โรงเรียนวรรณสุตศึกษาลัย และ โรงเรียนบรรหารแจ่มใสวิทยา 1 | 40 |
| 8 | โรงเรียนธรรมโชติศึกษาลัยและโรงเรียนสามชุกรัตนโกศาราม | 40 |
| 9 | โรงเรียนบรรหารแจ่มใสวิทยา 3 และโรงเรียนบางปลาม้า “สูงสูดผดุงวิทย์” | 40 |

7. ผู้วิจัยดำเนินการทดสอบกลุ่มตัวอย่างโดยใช้คอมพิวเตอร์ในโรงเรียนที่เป็นกลุ่มตัวอย่างโรงเรียนละ 2 ห้องเรียน เพื่อความสะดวกในการเก็บข้อมูล

8. ผู้วิจัยชี้แจงเกี่ยวกับวิธีการตอบข้อสอบแก่ผู้สอบจนเข้าใจ ผลการตอบของผู้สอบแต่ละคนจะถูกเก็บบันทึกไว้ในจานแม่เหล็ก

9. ดำเนินการทดสอบจำนวน 5 ครั้ง โดยเว้นระยะห่างในการทดสอบแต่ละครั้งประมาณ 1 สัปดาห์

10. ตรวจสอบความสมบูรณ์ครบถ้วนของข้อมูล จากการดำเนินการสอบโดยใช้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 5 ครั้ง กับนักเรียนกลุ่มตัวอย่างจำนวน 360 คน หลังจากที่ได้ตรวจสอบความครบถ้วนของข้อมูล ปรากฏว่า มีจำนวนนักเรียนที่เป็นกลุ่มตัวอย่างมีข้อมูลของการทดสอบไม่ครบทั้ง 5 ครั้ง จึงต้องตัดออกไม่นำมาวิเคราะห์ จำนวน 41 คน คิดเป็นร้อยละ 11.39 ของจำนวนกลุ่มตัวอย่าง จึงเหลือกลุ่มตัวอย่างจำนวน 319 คน คิดเป็นร้อยละ 88.61 ของจำนวนกลุ่มตัวอย่างทั้งหมด มีรายละเอียดดังตารางที่ 8

ตารางที่ 8 จำนวนกลุ่มตัวอย่างที่นำมาวิเคราะห์

| ระดับความสามารถทางภาษาอังกฤษ | อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | จำนวนกลุ่มตัวอย่าง | จำนวนกลุ่มตัวอย่างที่นำมาวิเคราะห์ |
|------------------------------|------------------------------------|--------------------|------------------------------------|
| ต่ำ | 10 | 40 | 24 |
| ต่ำ | 20 | 40 | 27 |
| ต่ำ | 30 | 40 | 33 |
| ปานกลาง | 10 | 40 | 40 |
| ปานกลาง | 20 | 40 | 40 |
| ปานกลาง | 30 | 40 | 40 |
| สูง | 10 | 40 | 40 |
| สูง | 20 | 40 | 40 |
| สูง | 30 | 40 | 35 |
| รวม | | 360 | 319 |

10. นำข้อมูลของผู้สอบแต่ละคนที่บันทึกไว้ในงานแม่เหล็ก คือ การประมาณค่าความสามารถของผู้สอบทั้ง 5 ครั้ง มาวิเคราะห์ทางสถิติ เพื่อศึกษาผลของตัวแปรอิสระต่อตัวแปรตาม

การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลนั้นได้วิเคราะห์ค่าของตัวแปรตาม ได้แก่ ความตรงตามสภาพ และความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ เพื่อศึกษาผลที่เกิดขึ้นอันเนื่องมาจากตัวแปรอิสระ ได้แก่ ระดับความสามารถทางภาษาอังกฤษของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ โดยดำเนินการดังนี้

1. คำนวณค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ จากสูตร (Glass and Stanley, 1970)

$$r_{xy} = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

เมื่อ r_{xy} แทน ความสัมพันธ์ระหว่างค่าความสามารถจากแบบสอบวัดความทาง
ภาษาอังกฤษกับค่าความสามารถเฉลี่ยจากการทดสอบแบบปรับ
เหมาะกับความสามารถด้วยคอมพิวเตอร์

n แทน จำนวนผู้สอบ

x แทน ค่าความสามารถโดยประมาณจากแบบสอบวัดความสามารถทาง
ภาษาอังกฤษ

y แทน ค่าความสามารถเฉลี่ยจากการทดสอบแบบปรับเหมาะกับความ
สามารถของผู้สอบด้วยคอมพิวเตอร์

2. แปลงค่าความตรงตามสภาพให้เป็นคะแนนฟิชเชอร์ซี (Fisher's Z) จากสูตร
(Glass and Stanley, 1970)

$$Z_r = 0.5 [\ln (1+r) + \ln (1-r)]$$

$$r = \frac{(e^{2z} - 1)}{(e^{2z} + 1)}$$

เมื่อ Z_r แทน คะแนนฟิชเชอร์ซี

r แทน ค่าความตรงเชิงโครงสร้าง

3. ทดสอบความแตกต่างของค่าความตรงเชิงสภาพ โดยใช้ค่าไคสแควร์ (Chi square) จากสูตร
(Glass and Stanley, 1970)

$$\sum_{i=1}^j (Z_{r_j} - Z_r)^2 (n_j - 3) \longrightarrow \chi_{j-1}^2; \alpha$$

$$Z_r = \frac{\sum_{i=1}^j (n_j - 3) Z_{r_j}}{\sum_{i=1}^j (n_j - 3)}$$

เมื่อ Z_{r_j} แทน ค่าฟิชเชอร์ซี (Fisher's Z) ของกลุ่มผู้สอบ j

n_j แทน จำนวนผู้สอบกลุ่มที่ j

4. เปรียบเทียบความแตกต่างของค่าความตรงตามสภาพเป็นรายคู่ (Glass and Stanley, 1970)

$$Z = \frac{Z_{r1} - Z_{r2}}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

เมื่อ Z แทน สถิติทดสอบซี (Z test)

Z_{r1}, Z_{r2} แทน ค่าพีชเชอร์ซี ซึ่งแปลงมาจากค่าสหสัมพันธ์ r_1 และ r_2
ตามลำดับ

n แทน ขนาดของกลุ่มตัวอย่าง

5. การวิเคราะห์ความแปรปรวนทางเดียวแบบวัดซ้ำ เพื่อเปรียบเทียบค่าเฉลี่ยของค่าความสามารถแบบวัดซ้ำ มีสูตรในการคำนวณตามลำดับดังนี้ (Laurence G Grimm, 1993)

5.1 คำนวณ Total Sum of Squares (SS_T) จากสูตร

$$SS_T = \sum X^2 - \frac{(\sum X)^2}{N}$$

เมื่อ $\sum X^2$ แทน ผลรวมกำลังสองของข้อมูล

$(\sum X)^2$ แทน ผลรวมของข้อมูล

N แทน จำนวนข้อมูล

5.2 คำนวณ Sum of Squares Between – Groups (SS_{BG}) จากสูตร

$$SS_{BG} = \frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \dots + \frac{(\sum X_k)^2}{n_k}$$

เมื่อ $(\sum X_k)^2$ แทน ผลรวมของข้อมูลในการวัดซ้ำครั้งที่ k

n_k แทน จำนวนของข้อมูลในการวัดซ้ำครั้งที่ k

5.3 คำนวณ Sum of Squares Within – Groups (SS_w) จากสูตร

$$SS_w = \sum X^2 - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \dots + \frac{(\sum X_k)^2}{n_k} \right]$$

5.4 คำนวณ Sum of Squares Between – Subject (SS_{BS}) จากสูตร

$$SS_{BS} = \frac{(\sum P_1)^2}{k} + \frac{(\sum P_2)^2}{k} + \dots + \frac{(\sum P_k)^2}{k} - \left[\frac{(\sum X)^2}{N} \right]$$

เมื่อ $(\sum P_k)^2$ แทน ผลรวมคะแนนของผู้สอบแต่ละบุคคลในการวัดซ้ำ
 k แทน จำนวนครั้งที่ทำการวัดซ้ำ

5.5 คำนวณ Sum of Squares Error (SS_{error}) จากสูตร

$$SS_{error} = SS_W - SS_{BS}$$

5.6 คำนวณ F จากสูตร

$$F = \frac{MS_{BG}}{MS_{error}}$$

เมื่อ $MS_{BG} = SS_{BG}/df_{BG}$
 $MS_{error} = SS_{error}/df_{error}$

6. วิเคราะห์ความแปรปรวนแบบสองทางแบบวัดซ้ำ เพื่อทดสอบผลของตัวแปรอิสระ (ศิริชัย พงษ์วิชัย, 2543)

6.1 โมเดลเชิงคณิตศาสตร์

โมเดลเชิงคณิตศาสตร์ในรูปของสมการเชิงเส้นตรงแบบบวก สำหรับการวิเคราะห์ความแปรปรวนจำแนกแบบสองทาง เมื่อมีการวัดซ้ำ กำหนดไว้ดังนี้ คือ

$$X_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad \begin{array}{l} i = 1, 2, \dots, r \\ j = 1, 2, \dots, c \\ k = 1, 2, \dots, n \end{array}$$

เมื่อ μ แทน ค่าเฉลี่ยโดยทั่วไป (general grand mean)

α_i แทน ค่าคงที่ซึ่งแสดงผลกระทบ (main effect) ของตัวแปรอิสระที่ 1 ระดับที่ i

β_j แทน ค่าคงที่ซึ่งแสดงผลกระทบ (main effect) ของตัวแปรอิสระที่ 2 ระดับที่ j

$(\alpha\beta)_{ij}$ แทน ค่าคงที่ซึ่งแสดงผลกระทบ (interaction) จากตัวแปรอิสระที่ 1 ระดับที่ i และตัวแปรอิสระที่ 2 ระดับที่ j

ε_{ijk} แทน ค่าความคลาดเคลื่อนของค่าสังเกต X_{ijk} จากค่าเฉลี่ย และถ้า μ_{ij} แทนค่าเฉลี่ยของตัวแปรอิสระที่ 1 ระดับที่ i ทั้งหมด และในตัวแปรอิสระที่ 2 ระดับที่ j ทั้งหมด

$$\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

ดังนั้นโมเดลเชิงคณิตศาสตร์ อาจกำหนดได้อีกลักษณะหนึ่ง ดังนี้คือ

$$X_{ijk} = \mu_{ij} + \varepsilon_{ijk}$$

6.2 การวิเคราะห์ความแปรปรวนแบบสองทางแบบวัดซ้ำ แบ่งความแปรปรวนทั้งหมดออกเป็น 5 ส่วนดังนี้คือ

- 1) ความแปรปรวนที่เกิดจากความแตกต่างระหว่างระดับต่าง ๆ ของปัจจัยที่ 1
- 2) ความแปรปรวนที่เกิดจากความแตกต่างระหว่างระดับต่าง ๆ ของปัจจัยที่ 2
- 3) ความแปรปรวนที่เกิดจากความแตกต่างระหว่างส่วนประกอบ (Subtotal) ของตัวแปรอิสระทั้ง 2
- 4) ความแปรปรวนที่เกิดจากความแตกต่างระหว่างระดับต่าง ๆ ของตัวแปรอิสระที่ 1 และระดับต่าง ๆ ของตัวแปรอิสระที่ 2
- 5) ความแปรปรวนที่เกิดจากความแตกต่างภายในตัวแปรอิสระทั้งสอง ซึ่งเป็นความแปรปรวนที่ไม่ทราบสาเหตุ หรือเป็นความคลาดเคลื่อน

ความแปรปรวนที่จำแนกมาทั้ง 5 กลุ่ม สามารถคำนวณได้ตามประเภทของแหล่งผันแปรโดยคำนวณในรูปของผลรวมกำลังสอง (Sum of Square :SS) โดยมีการคำนวณตามลำดับต่อไปนี้

- 1) แหล่งความแปรปรวนทั้งหมด (Total Variation : SST)

$$SST = \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^{n_{ij}} X_{ijk}^2 - CT$$

$$CT = \left(\sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^{n_{ij}} X_{ijk} \right)^2 / n$$

เมื่อ r แทน จำนวนกลุ่มที่แบ่งด้วยตัวแปรอิสระที่ 1
 c แทน จำนวนกลุ่มที่แบ่งด้วยตัวแปรอิสระที่ 2
 X_{ijk} แทน ค่าของข้อมูลลำดับที่ k จากกลุ่มที่ i ของตัวแปรอิสระที่ 1 กลุ่มที่ j
 ของตัวแปรอิสระที่ 2
 n แทน จำนวนข้อมูลทั้งหมด

2) แหล่งความแปรปรวนที่เกิดจากความแตกต่างระหว่างระดับต่างๆ ของตัวแปรอิสระที่ 1 (SSR) คือ การพิจารณาถึงความแปรปรวนของตัวแปรแถว (Row) ซึ่งสามารถคำนวณได้ดังนี้

$$SSR = \sum_{i=1}^r (X_{i..}^2 / n_{i.}) - CT$$

เมื่อ $n_{i.}$ แทน จำนวนข้อมูลทั้งหมดของตัวแปรอิสระที่ 1 สำหรับกลุ่มที่ i

3) แหล่งความแปรปรวนที่เกิดจากความแตกต่างระหว่างระดับต่างๆ ของตัวแปรอิสระที่ 2 (SSC) คือ การพิจารณาถึงความแปรปรวนของตัวแปรหลัก (Column) ซึ่งสามารถคำนวณได้ดังนี้

$$SSC = \sum_{j=1}^c (X_{.j.}^2 / n_{.j}) - CT$$

เมื่อ $n_{.j}$ แทน จำนวนข้อมูลทั้งหมดของตัวแปรอิสระที่ 2 สำหรับกลุ่มที่ j

4) แหล่งความแปรปรวนที่เกิดจากความแตกต่างระหว่างส่วนประกอบของตัวแปรอิสระ (SSS) เป็นการพิจารณาถึงความแปรปรวนระหว่างส่วนประกอบ (Subtotal) ของตัวแปรอิสระซึ่งส่วนประกอบของตัวแปรทั้ง 2 ที่ประกอบเป็นตาราง $r \times c$ ซึ่งสามารถคำนวณได้ดังนี้คือ

$$SSS = \sum_{i=1}^r \sum_{j=1}^c X_{ij.}^2 / n_{ij} - CT$$

เมื่อ n_{ij} แทน จำนวนข้อมูลทั้งหมดของตัวแปรอิสระที่ 1 กลุ่มที่ i และตัวแปรอิสระที่ 2 สำหรับกลุ่มที่ j

5) แหล่งความแปรปรวนที่เกิดจากความแตกต่างร่วมระหว่างตัวแปรอิสระที่ 1 และ 2 (SSI) เป็นการพิจารณาถึงความแปรปรวนของผลกระทบร่วม (interaction) ระหว่างตัวแปรด้านแถวและหลัก ซึ่งสามารถคำนวณได้ดังนี้

$$SSI = SSS - SSR - SSC$$

6) แหล่งความแปรปรวนที่ไม่ทราบสาเหตุ (SSE) เป็นการพิจารณาถึงความแปรปรวนที่เป็นความคลาดเคลื่อน (Error) ซึ่งสามารถคำนวณได้ดังนี้

$$SSE = SST - SSS$$

การวิเคราะห์ความแปรปรวนจากแหล่งความแปรปรวนต่างๆ สามารถสรุปได้ดังตารางที่ 9

ตารางที่ 9 ตารางการวิเคราะห์ความแปรปรวนแบบสองทางแบบวัดซ้ำ

| สาเหตุของความแปรปรวน | df | SS | MS | F |
|--|------------|------------|--------------------------|-------------|
| ระหว่างกลุ่มต่างๆ ในตัวแปรอิสระที่ 1 (ตัวแปรด้านแถว) | r-1 | SSR | MSR = SSR/r-1 | F = MSR/MSE |
| ระหว่างกลุ่มต่างๆ ในตัวแปรอิสระที่ 2 (ตัวแปรด้านหลัก) | c-1 | SSC | MSC = SSC/c-1 | F = MSC/MSE |
| ระหว่างส่วนประกอบของตัวแปร ผลกระทบร่วม (Interaction) | rc-1 | SSS | MSS = SSS/rc-1 | F = MSS/MSE |
| | (r-1)(c-1) | SSI | MSI = SSI/(r-1) (c-1) | F = MSI/MSE |
| ความคลาดเคลื่อน (Error) | n-rc | SSE | MSE = SSE/n-rc | |
| รวม | n-1 | SST | | |

6. เปรียบเทียบความแตกต่างของค่าเฉลี่ยจำนวนข้อสอบเป็นรายคู่โดยใช้วิธีเอสของเชฟเฟ (Scheffe's S Method)

$$S = \sqrt{(k-1)F_{\alpha, v_1, v_2} MS_e \left[\sum_{j=1}^k \frac{(c_j)^2}{n_j} \right]}$$

เมื่อ F แทน ค่า F ที่ขึ้นความเป็นอิสระ v_1 และ v_2 ที่ระดับนัยสำคัญ α

k แทน จำนวนระดับการทดลอง

c_j แทน สัมประสิทธิ์การเปรียบเทียบ

n_j แทน จำนวนคะแนนในระดับการทดลอง j

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน โดยผู้วิจัยเสนอผลการวิเคราะห์ข้อมูลเป็น 3 ตอนตามลำดับดังนี้

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลเบื้องต้น

ตอนที่ 2 ผลการเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

1. ผลการคำนวณค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำ แตกต่างกัน และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพของการทดสอบ

2. ผลการคำนวณค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพของการทดสอบ

ตอนที่ 3 ผลการเปรียบเทียบความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

1. ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามแหล่งความแปรปรวน

2. ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน

3. ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำและจำนวนครั้งที่ทำการทดสอบแตกต่างกัน

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลเบื้องต้น

การวิเคราะห์ข้อมูลเบื้องต้นในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 5 ครั้ง ได้ผลการวิเคราะห์ดังตารางที่ 10

ตารางที่ 10 จำนวนผู้สอบ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบและจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 5 ครั้ง

| อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ ของผู้สอบ | จำนวน ครั้งที่ ผู้สอบ | ค่าความสามารถ | | | | จำนวนข้อสอบ | | | | | |
|---------------------------------|----------------------------------|-----------------------------|---------------|---------|-----------|---------|-------------|-----|-----------|---------|--------|--|
| | | | min | max | \bar{X} | SD | min | max | \bar{X} | SD | | |
| 10 | ต่ำ | 1 | 24 | -2.2120 | -.7410 | -1.7193 | .4686 | 7 | 13 | 9.7083 | 1.7315 | |
| | | 2 | 24 | -2.6050 | -.7920 | -1.8800 | .5494 | 10 | 32 | 20.6667 | 6.0265 | |
| | | 3 | 24 | -2.7250 | -.7850 | -1.9344 | .4932 | 9 | 40 | 21.1667 | 9.3095 | |
| | | 4 | 24 | -2.6180 | -.9400 | -1.7993 | .4371 | 7 | 29 | 14.4091 | 6.5948 | |
| | | 5 | 24 | -2.4690 | -.6060 | -1.6805 | .5355 | 8 | 20 | 14.4737 | 4.2995 | |
| | ปานกลาง | 1 | 40 | -2.2650 | -.0960 | -1.2999 | .5617 | 7 | 16 | 10.7436 | 2.2562 | |
| | | 2 | 40 | -2.5200 | .5290 | -1.2975 | .7363 | 8 | 33 | 15.2308 | 5.7604 | |
| | | 3 | 40 | -2.7590 | .1330 | -1.5907 | .7166 | 7 | 39 | 18.6250 | 8.1451 | |
| | | 4 | 40 | -2.4810 | .4460 | -1.3256 | .7437 | 7 | 37 | 14.4737 | 6.9816 | |
| | | 5 | 40 | -2.5630 | .1780 | -1.3864 | .7258 | 8 | 21 | 13.3636 | 3.3243 | |
| สูง | 1 | 40 | -2.1220 | .4770 | -.3758 | .5588 | 7 | 17 | 11.0789 | 2.4756 | | |
| | 2 | 40 | -.9670 | .8950 | -.1318 | .4451 | 8 | 18 | 12.1538 | 2.3003 | | |
| | 3 | 40 | -1.2790 | .8140 | -.0709 | .4078 | 8 | 15 | 11.2703 | 1.9386 | | |
| | 4 | 40 | -1.0280 | .8430 | -.0319 | .4377 | 9 | 18 | 11.6667 | 1.7970 | | |
| | 5 | 40 | -.8360 | 1.0010 | .1646 | .4211 | 8 | 15 | 10.8421 | 2.2243 | | |
| รวม | | | 104 | | | | | | | | | |

ตารางที่ 10 (ต่อ)

| อัตราการใช้ข้อสอบซ้ำ (%) | ระดับความสามารถของผู้สอบ | ครั้งที่ | จำนวนผู้สอบ | ค่าความสามารถ | | | | จำนวนข้อสอบ | | | | |
|--------------------------|--------------------------|----------|-------------|---------------|--------|-----------|-------|-------------|-----|-----------|--------|--|
| | | | | min | max | \bar{X} | SD | min | max | \bar{X} | SD | |
| 20 | ต่ำ | 1 | 27 | -2.2650 | -.8240 | -1.6585 | .4072 | 6 | 15 | 9.8519 | 2.0885 | |
| | | 2 | 27 | -2.4340 | -.4530 | -1.7233 | .5729 | 8 | 28 | 16.1481 | 5.3111 | |
| | | 3 | 27 | -2.3350 | -.3020 | -1.6676 | .4516 | 7 | 24 | 13.5200 | 4.6558 | |
| | | 4 | 27 | -2.5690 | -.4900 | -1.7215 | .5437 | 8 | 27 | 13.8462 | 4.5580 | |
| | | 5 | 27 | -2.2340 | -.7220 | -1.6702 | .4500 | 7 | 18 | 11.0526 | 2.7177 | |
| | ปานกลาง | 1 | 40 | -2.1920 | .0580 | -1.1551 | .6752 | 7 | 25 | 10.7500 | 3.4844 | |
| | | 2 | 40 | -2.6360 | -.0950 | -1.2956 | .5912 | 7 | 21 | 13.1500 | 2.9400 | |
| | | 3 | 40 | -2.6670 | .1270 | -1.2706 | .5913 | 9 | 32 | 13.8649 | 5.5835 | |
| | | 4 | 40 | -2.4220 | .2650 | -1.2668 | .6985 | 9 | 21 | 14.6667 | 3.8359 | |
| | | 5 | 40 | -2.3600 | .3650 | -1.2665 | .6595 | 9 | 18 | 11.5758 | 2.1799 | |
| | สูง | 1 | 40 | -1.611 | 1.2790 | .0324 | .6362 | 8 | 16 | 11.6500 | 1.8886 | |
| | | 2 | 40 | -1.476 | .9650 | .0898 | .5744 | 8 | 21 | 12.2750 | 2.5918 | |
| | | 3 | 40 | -1.669 | 1.0700 | .0456 | .5753 | 9 | 19 | 11.2250 | 2.2129 | |
| | | 4 | 40 | -1.139 | 1.4370 | .1815 | .6161 | 8 | 16 | 11.4615 | 1.9848 | |
| | | 5 | 40 | -.8960 | 1.2390 | .2454 | .4445 | 8 | 16 | 10.6842 | 1.9879 | |
| รวม | | | 107 | | | | | | | | | |

ตารางที่ 10 (ต่อ)

| อัตราการใช้ข้อสอบซ้ำ (%) | ระดับความสามารถของผู้สอบ | ครั้งที่ | จำนวนผู้สอบ | ค่าความสามารถ | | | | จำนวนข้อสอบ | | | |
|--------------------------|--------------------------|----------|-------------|---------------|--------|-----------|-------|-------------|-----|-----------|--------|
| | | | | min | max | \bar{X} | SD | min | max | \bar{X} | SD |
| 30 | ต่ำ | 1 | 33 | -2.4410 | -.8400 | -1.8963 | .3891 | 7 | 24 | 11.6061 | 4.3513 |
| | | 2 | 33 | -2.5070 | -.3230 | -1.8466 | .5129 | 9 | 22 | 15.0303 | 3.3956 |
| | | 3 | 33 | -2.4610 | -.3030 | -1.8208 | .4532 | 7 | 24 | 14.3939 | 4.3007 |
| | | 4 | 33 | -2.3670 | -.1080 | -1.7567 | .5563 | 7 | 24 | 14.9697 | 4.5927 |
| | | 5 | 33 | -2.1830 | -.0870 | -1.5431 | .5241 | 7 | 25 | 11.6250 | 3.1186 |
| | ปานกลาง | 1 | 40 | -2.6310 | .2600 | -.8929 | .8615 | 7 | 27 | 11.6500 | 4.3059 |
| | | 2 | 40 | -2.3560 | .4320 | -.7466 | .7383 | 7 | 22 | 11.8421 | 3.5982 |
| | | 3 | 40 | -2.1200 | .7380 | -.7654 | .8178 | 8 | 26 | 12.4872 | 3.7267 |
| | | 4 | 40 | -1.974 | .4230 | -.6777 | .6445 | 7 | 21 | 11.7895 | 2.9515 |
| | | 5 | 40 | -1.6960 | .9100 | -.3607 | .6467 | 7 | 19 | 10.4615 | 2.5634 |
| | สูง | 1 | 35 | -1.8130 | .3880 | -.2596 | .5063 | 8 | 17 | 10.5937 | 2.4344 |
| | | 2 | 35 | -1.482 | .7300 | -.3240 | .5764 | 7 | 18 | 11.7059 | 2.7582 |
| | | 3 | 35 | -1.5200 | .5830 | -.2920 | .5413 | 7 | 20 | 11.1143 | 2.6764 |
| | | 4 | 35 | -1.4200 | .7790 | -.2813 | .5342 | 8 | 17 | 10.9118 | 2.2207 |
| | | 5 | 35 | -.9340 | 1.645 | .0407 | .6175 | 8 | 15 | 11.1765 | 1.8169 |
| รวม | | | 108 | | | | | | | | |

จากตารางที่ 10 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 5 ครั้ง เมื่ออัตราการใช้ข้อสอบซ้ำเป็น 10 เปอร์เซนต์ พบว่า ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำจะมีค่าความสามารถต่ำสุดอยู่ระหว่าง -2.7250 ถึง -2.2120 มีค่าความสามารถสูงสุดอยู่ระหว่าง -.9400 ถึง -.6060 และมีค่าความสามารถเฉลี่ยอยู่ระหว่าง -1.9344 ถึง -1.6805 จำนวนข้อสอบที่ใช้ต่ำสุดอยู่ระหว่าง 7 ถึง 10 ข้อ จำนวนข้อสอบสูงสุดอยู่ระหว่าง 13 ถึง 40 ข้อ และจำนวนข้อสอบเฉลี่ยอยู่ระหว่าง 9 ถึง 21 ข้อ ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษปานกลางจะมีค่าความสามารถต่ำสุดอยู่ระหว่าง -2.7590 ถึง -

และมีค่าความสามารถเฉลี่ยอยู่ระหว่าง -.3240 ถึง .0407 จำนวนข้อสอบที่ใช้ต่ำสุดอยู่ระหว่าง 7 ถึง 8 ข้อ จำนวนข้อสอบสูงสุดอยู่ระหว่าง 15 ถึง 20 ข้อ และจำนวนข้อสอบเฉลี่ยอยู่ระหว่าง 11 ถึง 12 ข้อ

ตอนที่ 2 ผลการเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

1. ผลการคำนวณค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน

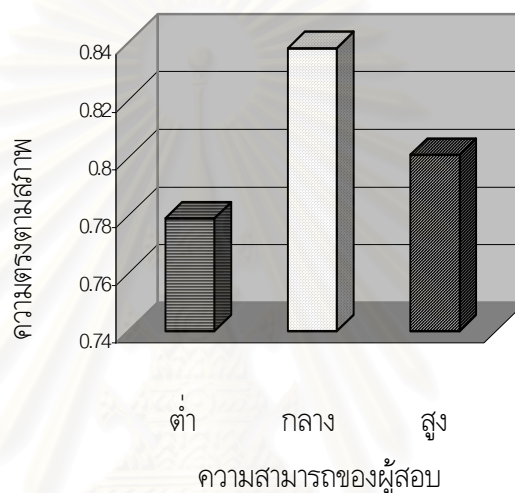
การคำนวณค่าความตรงตามสภาพของการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถเฉลี่ยที่ได้จากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับค่าความสามารถที่ได้จากแบบสอบวัดความสามารถทางภาษาอังกฤษ และศึกษาความแตกต่างของค่าความตรงตามสภาพโดยใช้ค่าไคสแควร์ ได้ผลการวิเคราะห์ดังตารางที่ 11

ตารางที่ 11 ค่าความตรงตามสภาพ และค่าไคสแควร์สำหรับการทดสอบความแตกต่างของค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

| อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | ระดับ ความสามารถ | จำนวน ผู้สอบ | ความตรงตามสภาพ | | χ^2 |
|---------------------------------------|---------------------|-----------------|----------------|--------|----------|
| | | | r | z | |
| 10 | ต่ำ | 24 | .779 | 1.0428 | .4529 |
| | ปานกลาง | 40 | .838 | 1.2144 | |
| | สูง | 40 | .801 | 1.1014 | |
| 20 | ต่ำ | 27 | .602 | .6962 | 2.0451 |
| | ปานกลาง | 40 | .770 | 1.0203 | |
| | สูง | 40 | .631 | .7431 | |
| 30 | ต่ำ | 33 | .508 | .5600 | .2621 |
| | ปานกลาง | 40 | .592 | .6807 | |
| | สูง | 35 | .534 | .5957 | |
| | | | | | 15.7292* |

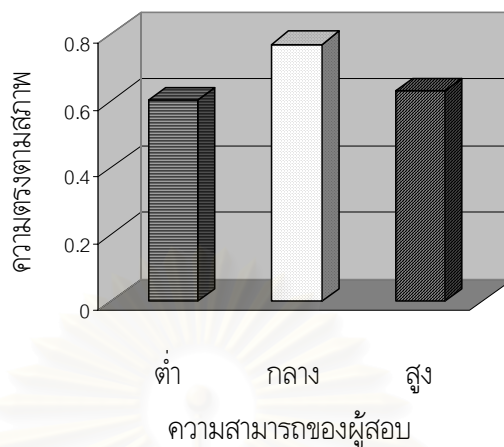
* $p < .05$

จากตารางที่ 11 พบว่า เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ ปานกลาง และสูง มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ .779 .838 และ .801 ตามลำดับ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพดังกล่าว พบว่า ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ ปานกลาง และสูง มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 17



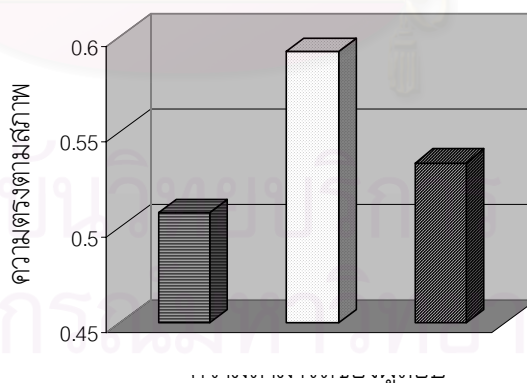
แผนภาพที่ 17 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถ และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์

เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ ปานกลาง และสูง มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ .602 .770 และ .631 ตามลำดับ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพดังกล่าว พบว่า ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ ปานกลาง และสูง มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 18



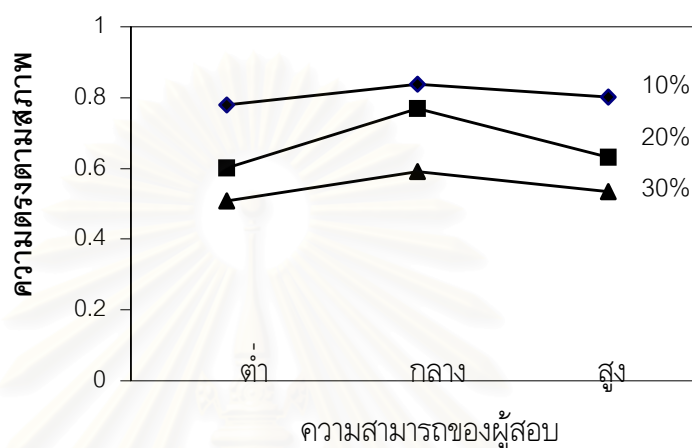
แผนภาพที่ 18 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถ และมีอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์

เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ ปานกลาง และสูง มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ .508 .592 และ .534 ตามลำดับ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพดังกล่าว พบว่า ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ ปานกลาง และสูงมีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงใน แผนภาพที่ 19



แผนภาพที่ 19 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถ และมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์

เมื่อทำการทดสอบภาพรวมของค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความ
 ความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า มีอย่างน้อย 1 คู่ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ
 .05 ดังแสดงในแผนภาพที่ 20



แผนภาพที่ 20 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความ
 ความสามารถของผู้สอบ
 ด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถ และอัตราการใช้ข้อสอบซ้ำ

2. ผลการทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับ
 ความสามารถของผู้สอบด้วยคอมพิวเตอร์

จากผลการทดสอบภาพรวมของค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับ
 ความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่ามีอย่างน้อย 1 คู่ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ
 .05 ผู้วิจัยจึงนำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความ
 ความสามารถของผู้สอบด้วย
 คอมพิวเตอร์ มาทดสอบความแตกต่างเป็นรายคู่ จำแนกตามระดับความสามารถของ ผู้สอบ ซึ่งผลการ
 ทดสอบปรากฏดังตารางที่ 12 - 14

สถาบันวิทยบริการ
 จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 12 การทดสอบความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความ
ความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถต่ำ

| อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | ความตรงตามสภาพ | อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | | |
|---------------------------------------|----------------|------------------------------------|------|------|
| | | 10 | 20 | 30 |
| 10 | .779 | - | .177 | .271 |
| 20 | .602 | | - | .094 |
| 30 | .508 | | | - |

จากตารางที่ 12 เมื่อนำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความ
ความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถทางภาษาอังกฤษต่ำมาทำการทดสอบความ
แตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 20 และ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการทดสอบไม่
แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 13 การทดสอบความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความ
ความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถปานกลาง

| อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | ความตรงตามสภาพ | อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | | |
|---------------------------------------|----------------|------------------------------------|------|-------|
| | | 10 | 20 | 30 |
| 10 | .838 | - | .068 | .246* |
| 20 | .770 | | - | .178 |
| 30 | .592 | | | - |

* $p < .05$

จากตารางที่ 13 เมื่อนำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความ
ความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถทางภาษาอังกฤษปานกลางมาทำการทดสอบความ
แตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ จะมีค่าความตรงตามสภาพของการทดสอบไม่แตกต่าง

จากอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์อย่างมีนัยสำคัญทางสถิติ และอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ จะมีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างจากอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์อย่างมีนัยสำคัญทางสถิติ ส่วนอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ จะมีค่าความตรงตามสภาพของการทดสอบแตกต่างจากอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตารางที่ 14 ความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถสูง

| อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | ความตรงตามสภาพ | อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | | |
|---------------------------------------|----------------|------------------------------------|------|-------|
| | | 10 | 20 | 30 |
| 10 | .801 | - | .173 | .267* |
| 20 | .631 | | - | .094 |
| 30 | .534 | | | - |

* $p < .05$

จากตารางที่ 14 เมื่อนำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีระดับความสามารถทางภาษาอังกฤษสูงมาทำการทดสอบความแตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ จะมีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างจากอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์อย่างมีนัยสำคัญทางสถิติ และอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ จะมีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างจากอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์อย่างมีนัยสำคัญทางสถิติ ส่วนอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ จะมีค่าความตรงตามสภาพของการทดสอบแตกต่างจากอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. ผลการคำนวณค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ การคำนวณค่าความตรงตามสภาพของการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถเฉลี่ยที่ได้จากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

กับค่าความสามารถที่ได้จากแบบสอบวัดความสามารถทางภาษาอังกฤษ โดยจำแนกตามตัวแปรอิสระ 2 ตัวแปร คือ ระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ ผลการวิเคราะห์ปรากฏดังตารางที่ 15 - 16

ตารางที่ 15 ความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วย
คอมพิวเตอร์ จำแนกตามระดับความสามารถของผู้สอบ

| ระดับความสามารถของผู้สอบ | จำนวนผู้สอบ | ความตรงตามสภาพ | |
|--------------------------|-------------|----------------|--------|
| | | r | z |
| ต่ำ | 84 | .614 | .7153 |
| ปานกลาง | 120 | .770 | 1.0203 |
| สูง | 115 | .666 | .8035 |
| รวม | 319 | | |

จากตารางที่ 15 พบว่า ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ ปานกลาง และสูง มีความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ .614 .770 และ .666 ตามลำดับ

เมื่อใช้อัตราการใช้ข้อสอบซ้ำเป็นเกณฑ์ในการศึกษาความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ได้ผลการวิเคราะห์ ดังตารางที่ 16

ตารางที่ 16 ความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วย
คอมพิวเตอร์ จำแนกตามอัตราการใช้ข้อสอบซ้ำ

| อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | จำนวนผู้สอบ | ความตรงตามสภาพ | |
|---------------------------------------|-------------|----------------|--------|
| | | r | z |
| 10 | 104 | .943 | 1.7645 |
| 20 | 107 | .924 | 1.6157 |
| 30 | 108 | .848 | 1.2490 |
| รวม | 319 | | |

จากตารางที่ 16 พบว่า เมื่ออัตราการใช้ข้อสอบซ้ำเป็น 10 20 และ 30 เปอร์เซนต์ มีความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เท่ากับ .943 .924 และ .848 ตามลำดับ

4. ผลการทดสอบความแตกต่างของค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ

เมื่อนำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มาทดสอบความแตกต่างเป็นรายคู่ในแต่ละระดับของตัวแปรอิสระ 2 ตัวแปร คือ ระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ ซึ่งผลการทดสอบปรากฏดังตารางที่ 17-18

ตารางที่ 17 การทดสอบความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามระดับความสามารถของผู้สอบ

| ระดับความสามารถของผู้สอบ | ความตรงตามสภาพ | ระดับความสามารถของผู้สอบ | | |
|--------------------------|----------------|--------------------------|---------|------|
| | | ต่ำ | ปานกลาง | สูง |
| ต่ำ | .614 | - | .156* | .052 |
| ปานกลาง | .770 | | - | .104 |
| สูง | .666 | | | - |

* p < .05

จากตารางที่ 17 พบว่า ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำจะมีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างจาก ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษปานกลาง อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วน ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำจะมีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างจากผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษสูงอย่างมีนัยสำคัญทางสถิติ และผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษปานกลางจะมีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างจาก ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษสูงอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 18 การทดสอบความแตกต่างของความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความ
ความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามอัตราการใช้ข้อสอบซ้ำ

| อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | ความตรงตามสภาพ | อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์) | | |
|---------------------------------------|----------------|------------------------------------|------|-------|
| | | 10 | 20 | 30 |
| 10 | .943 | - | .019 | .095* |
| 20 | .924 | | - | .076* |
| 30 | .848 | | | - |

* $p < .05$

จากตารางที่ 18 พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ จะมีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ไม่แตกต่างจากอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์อย่างมีนัยสำคัญทางสถิติ ส่วนอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ จะมีค่าความตรงตามสภาพของการทดสอบแตกต่างจากอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ จะมีค่าความตรงตามสภาพของการทดสอบแตกต่างจากอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตอนที่ 3 ผลการเปรียบเทียบความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

1. ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ทั้ง 5 ครั้ง จำแนกตามแหล่งความแปรปรวน

การวิเคราะห์ความแปรปรวน เพื่อพิจารณาความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามแหล่งความแปรปรวน ซึ่งได้แก่ ระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ โดยใช้สถิติวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (Repeated Measurement ANOVA) และใช้วิธีการทดสอบของ Scheffe' ทำการเปรียบเทียบความแตกต่างเป็นรายคู่ เมื่อพบว่าผลการวิเคราะห์ความแปรปรวนแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ผลการวิเคราะห์ปรากฏดังตารางที่ 19

ตารางที่ 19 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามแหล่งความแปรปรวน

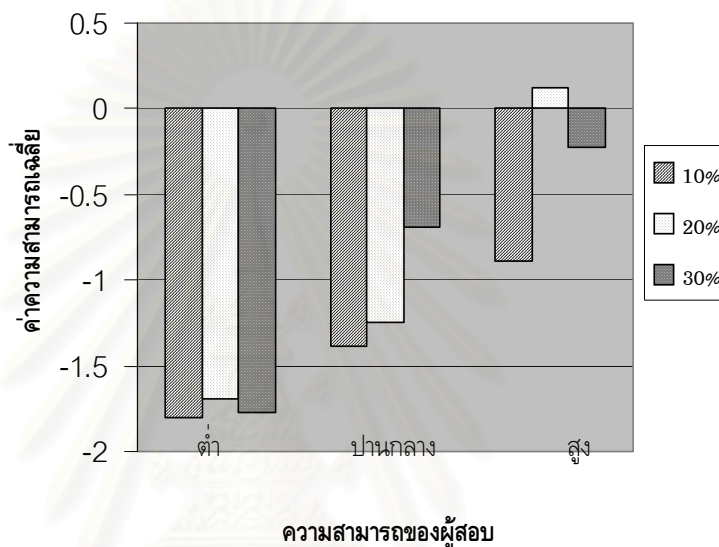
| แหล่งความแปรปรวน | df | SS | MS | F |
|--------------------------|------|----------|---------|------------|
| ระหว่างกลุ่ม | | | | |
| อัตราการใช้ข้อสอบซ้ำ (A) | 2 | 11.069 | 5.535 | 15.584** |
| ภายในกลุ่ม | | | | |
| ระดับความสามารถ (B) | 2 | 729.861 | 364.931 | 1027.545** |
| ครั้งที่ทำการทดสอบ (C) | 4 | 5.939 | 1.485 | 4.181** |
| A X B | 4 | 49.527 | 12.382 | 34.864** |
| A X C | 8 | 6.843 | .855 | 2.409** |
| B X C | 8 | 1.639 | .205 | .577 |
| A X B X C | 16 | 4.583 | .286 | .807 |
| ความคลาดเคลื่อน | 1550 | 550.480 | .355 | |
| รวม | 1594 | 1376.456 | | |

** $p < .01$

จากตารางที่ 19 เป็นการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามแหล่งความแปรปรวนพบว่า เมื่อพิจารณาผลของตัวแปรอัตราการใช้ข้อสอบซ้ำ และระดับความสามารถทางภาษาอังกฤษของผู้สอบที่ส่งผลต่อการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ถ้าไม่พิจารณาปฏิสัมพันธ์ระหว่างตัวแปรแล้ว พบว่า ตัวแปรอิสระทั้ง 2 ตัวแปร ต่างก็ส่งผลต่อการประมาณค่าความสามารถของผู้สอบอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

หากพิจารณาปฏิสัมพันธ์ระหว่างตัวแปรอิสระที่ส่งผลต่อการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า อัตราการใช้ ข้อสอบซ้ำ และระดับความสามารถของผู้สอบ ร่วมกันส่งผลต่อการประมาณค่าความสามารถของผู้สอบอย่างมีนัยสำคัญ

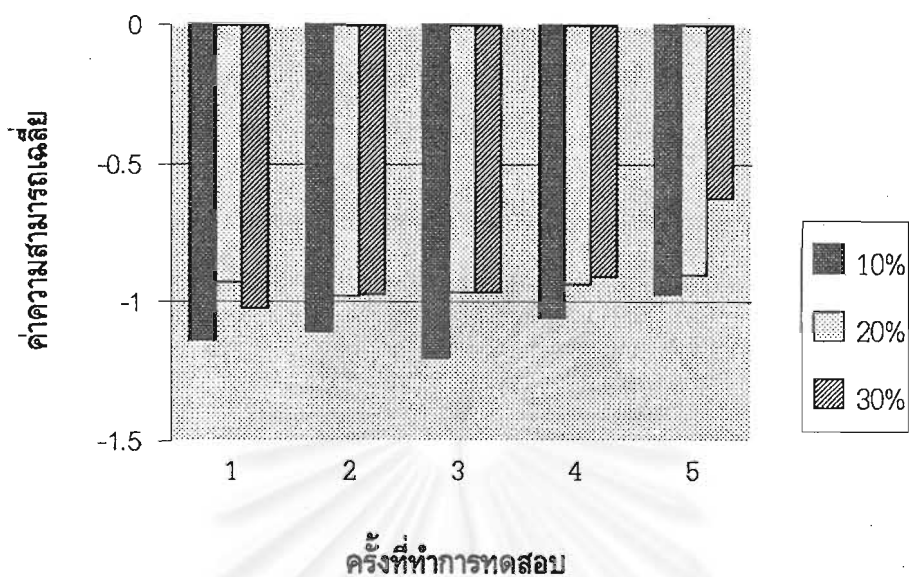
ทางสถิติที่ระดับ .01 และจำนวนครั้งที่ทำการทดสอบและอัตราการใช้ข้อสอบซ้ำร่วมกันส่งผลต่อการประมาณค่าความสามารถของผู้สอบอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ผู้วิจัยได้นำค่าความสามารถเฉลี่ยของผู้สอบจำแนกตามระดับความสามารถและอัตราการใช้ข้อสอบซ้ำเพื่อแสดงแนวโน้มของค่าความสามารถเฉลี่ยของผู้สอบ แสดงดังแผนภาพที่ 21



แผนภาพที่ 21 ความสามารถเฉลี่ยของผู้สอบ จำแนกตามความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำ

จากแผนภาพที่ 21 พบว่า เมื่อผู้สอบมีระดับความสามารถทางภาษาอังกฤษสูงขึ้น และมีอัตราการใช้ข้อสอบซ้ำสูงขึ้น จะทำให้ค่าความสามารถเฉลี่ยของผู้สอบมีแนวโน้มสูงขึ้น

ในการพิจารณาปฏิสัมพันธ์ของตัวแปรจำนวนครั้งที่ทำการทดสอบ และอัตราการใช้ข้อสอบซ้ำ ซึ่งร่วมกันส่งผลต่อความคงที่ในการประมาณค่าความสามารถของผู้สอบนั้น ผู้วิจัยได้นำค่าความสามารถเฉลี่ยของผู้สอบจำแนกตามอัตราการใช้ข้อสอบซ้ำและจำนวนครั้งที่ทำการทดสอบเพื่อแสดงแนวโน้มของค่าความสามารถเฉลี่ยของผู้สอบ แสดงดังแผนภาพที่ 22



แผนภาพที่ 22 ค่าสามารถเฉลี่ยของผู้สอบ จำแนกตามอัตราการให้ข้อสอบซ้ำและจำนวนครั้งที่ทำการทดสอบ

จากแผนภาพที่ 22 พบว่า เมื่อจำนวนครั้งที่ทำการทดสอบมากขึ้น และมีอัตราการให้ข้อสอบซ้ำที่สูงขึ้น จะทำให้แนวโน้มของค่าความสามารถเฉลี่ยของผู้สอบสูงขึ้นด้วย

2. ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ทั้ง 5 ครั้ง เมื่อระดับความสามารถของผู้สอบ และอัตราการให้ข้อสอบซ้ำแตกต่างกัน

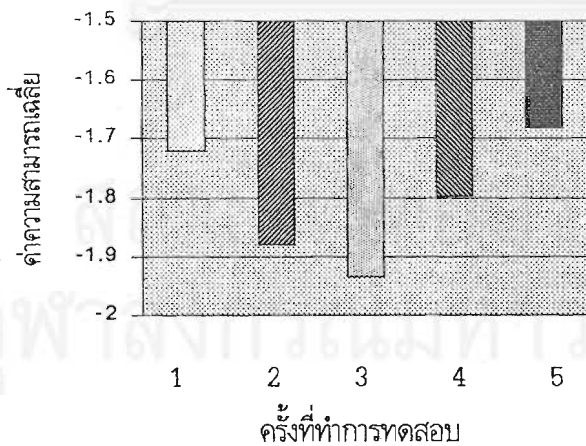
การวิเคราะห์ความแปรปรวน เพื่อพิจารณาความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถและอัตราการให้ข้อสอบซ้ำแตกต่างกัน โดยใช้สถิติวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (Repeated Measurement ANOVA)

เมื่อทำการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างระดับความสามารถของผู้สอบเมื่ออัตราการให้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ ปรากฏผลการวิเคราะห์ดังตารางที่ 20 - 22

ตารางที่ 20 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถต่ำ และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์

| แหล่งความแปรปรวน | df | SS | MS | F |
|-------------------|-----|--------|-------|-------|
| ระหว่างกลุ่ม | 4 | 1.311 | 0.325 | 2.321 |
| ภายในกลุ่ม | 115 | 28.576 | | |
| - ระหว่างบุคคล | 23 | 9.266 | | |
| - ความคลาดเคลื่อน | 138 | 19.310 | 0.140 | |
| รวม | 119 | 29.887 | | |

จากตารางที่ 20 เป็นผลการวิเคราะห์ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อผู้สอบมีความสามารถต่ำ และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้งไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 23

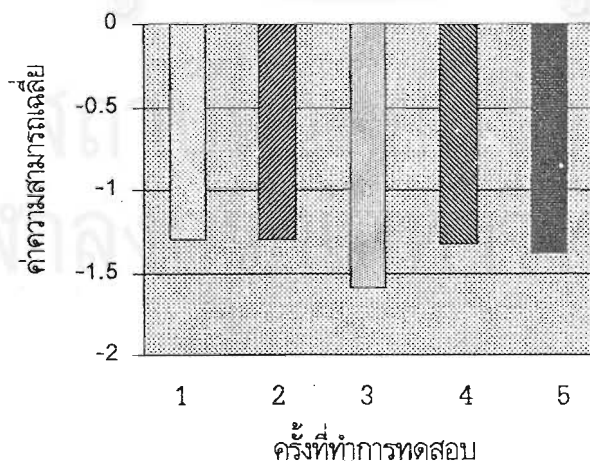


แผนภาพที่ 23 ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถต่ำและมีอัตราการใช้ข้อสอบซ้ำ 10 %

ตารางที่ 21 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถปานกลางและมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์

| แหล่งความแปรปรวน | df | SS | MS | F |
|-------------------|-----|--------|-------|-------|
| ระหว่างกลุ่ม | 4 | 2.424 | 0.606 | 1.883 |
| ภายในกลุ่ม | 195 | 95.595 | | |
| - ระหว่างบุคคล | 39 | 45.389 | | |
| - ความคลาดเคลื่อน | 156 | 50.206 | 0.322 | |
| รวม | 199 | 98.018 | | |

จากตารางที่ 21 เป็นผลการวิเคราะห์ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อผู้สอบมีความสามารถปานกลาง และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้งไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 24

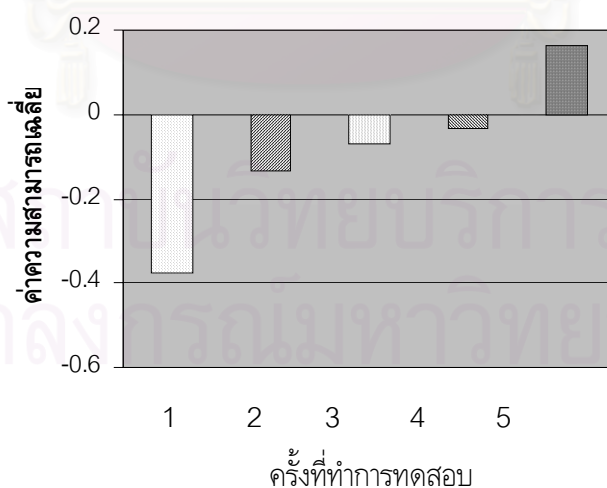


แผนภาพที่ 24 ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถปานกลาง และมีอัตราการใช้ข้อสอบซ้ำ 10 %

ตารางที่ 22 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|-----|--------|-------|-------|
| ระหว่างกลุ่ม | 4 | 0.516 | 0.129 | 1.047 |
| ภายในกลุ่ม | 195 | 45.524 | | |
| - ระหว่างบุคคล | 39 | 26.317 | | |
| - ความคลาดเคลื่อน | 156 | 19.207 | 0.123 | |
| รวม | 199 | 46.040 | | |

จากตารางที่ 22 เป็นผลการวิเคราะห์ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อผู้สอบมีความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้ง ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 25



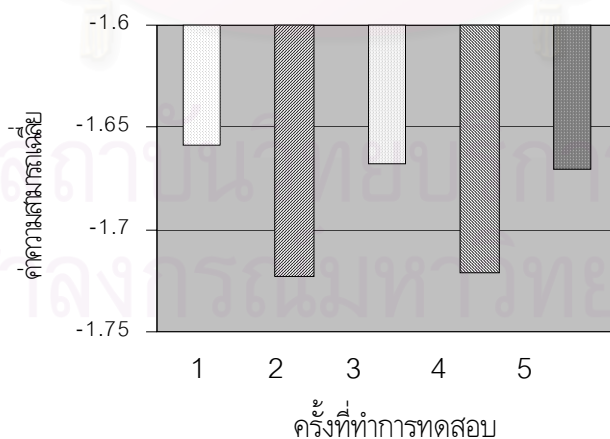
แผนภาพที่ 25 ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำ 10 %

เมื่อทำการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ระหว่างระดับความสามารถของผู้สอบ เมื่ออัตราการการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ ปรากฏผลดังตารางที่ 23 -25

ตารางที่ 23 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถต่ำ และมีอัตราการการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|-----|--------|-------|-------|
| ระหว่างกลุ่ม | 4 | 0.108 | 0.027 | 0.133 |
| ภายในกลุ่ม | 130 | 31.096 | | |
| - ระหว่างบุคคล | 26 | 9.968 | | |
| - ความคลาดเคลื่อน | 104 | 21.128 | 0.383 | |
| รวม | 134 | 31.204 | | |

จากตารางที่ 23 เป็นผลการวิเคราะห์ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อผู้สอบมีความสามารถต่ำและมีอัตราการการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้ง ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 26

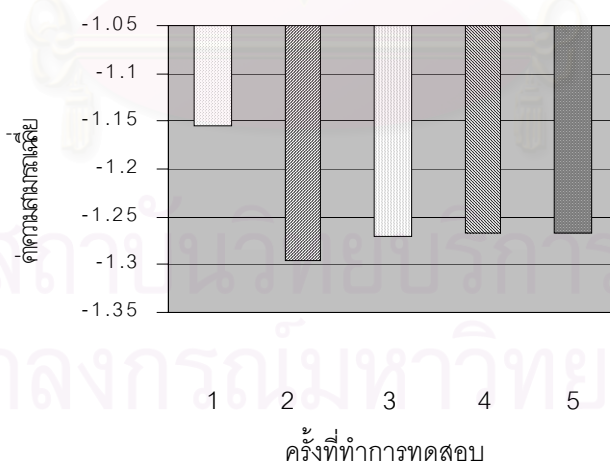


แผนภาพที่ 26 ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถต่ำและมีอัตราการการใช้ข้อสอบซ้ำ 20 %

ตารางที่ 24 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถปานกลาง และมีอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|-----|--------|-------|-------|
| ระหว่างกลุ่ม | 4 | 0.482 | 0.121 | 0.633 |
| ภายในกลุ่ม | 195 | 81.034 | | |
| - ระหว่างบุคคล | 39 | 51.237 | | |
| - ความคลาดเคลื่อน | 156 | 29.797 | 0.191 | |
| รวม | 199 | 81.516 | | |

จากตารางที่ 24 เป็นผลการวิเคราะห์ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อผู้สอบมีความสามารถปานกลาง และมีอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้งไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แสดงในแผนภาพที่ 27



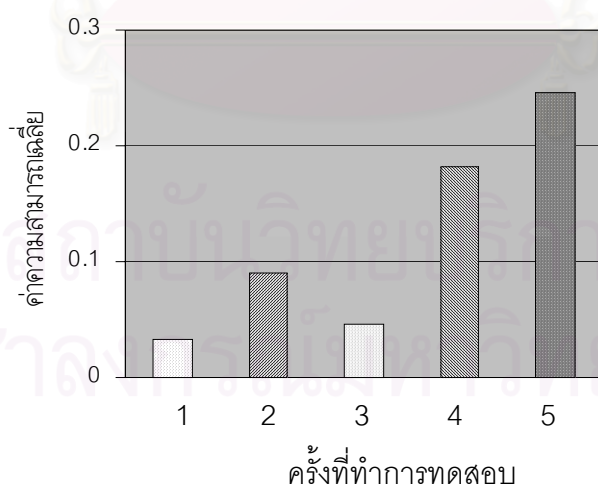
แผนภาพที่ 27 ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถปานกลางและมีอัตราการใช้ ข้อสอบซ้ำ 20 %

ตารางที่ 25 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบ
ปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถสูงและมี
อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|-----|--------|-------|-------|
| ระหว่างกลุ่ม | 4 | 1.345 | 0.673 | 3.44* |
| ภายในกลุ่ม | 195 | 64.075 | | |
| - ระหว่างบุคคล | 39 | 33.472 | | |
| - ความคลาดเคลื่อน | 156 | 30.603 | 0.196 | |
| รวม | 199 | 65.420 | | |

* $p < .05$

จากตารางที่ 25 เป็นผลการวิเคราะห์ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบ
ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อผู้สอบมีความสามารถสูงและมี
อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบ
ในการทดสอบทั้ง 5 ครั้งมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในแผนภาพที่ 28



แผนภาพที่ 28 ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำ 20 %

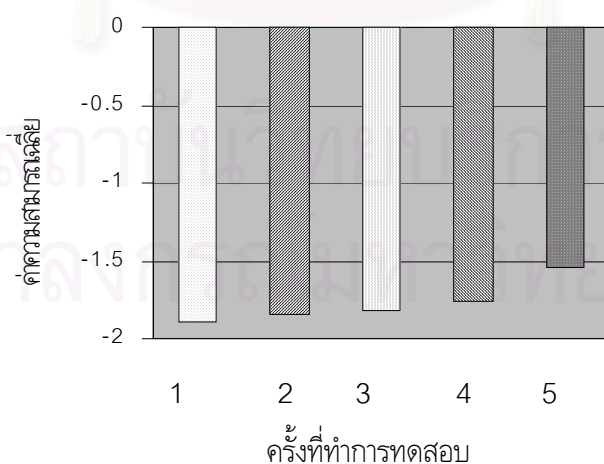
เมื่อทำการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างระดับความสามารถของผู้สอบ เมื่ออัตราการการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ ปรากฏผลการวิเคราะห์ดังตารางที่ 26 - 28

ตารางที่ 26 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถต่ำ และมีอัตราการการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|-----|--------|-------|--------|
| ระหว่างกลุ่ม | 4 | 2.509 | 0.627 | 4.053* |
| ภายในกลุ่ม | 160 | 38.532 | | |
| - ระหว่างบุคคล | 32 | 18.723 | | |
| - ความคลาดเคลื่อน | 128 | 19.809 | 0.155 | |
| รวม | 164 | 41.040 | | |

* $p < .05$

จากตารางที่ 26 เป็นผลการวิเคราะห์ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อผู้สอบมีความสามารถต่ำและมีอัตราการการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้งมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 29



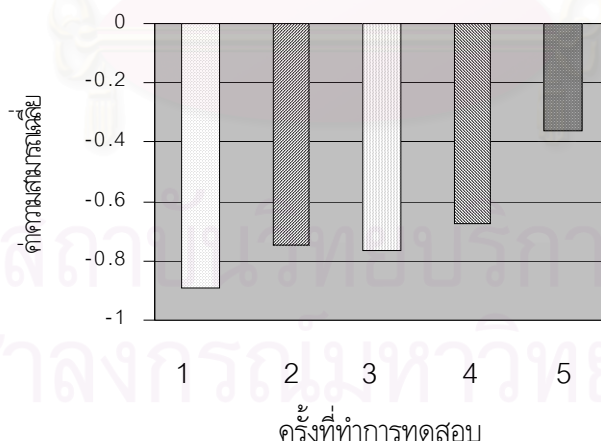
แผนภาพที่ 29 ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถต่ำและมีอัตราการการใช้ข้อสอบซ้ำ 30 %

ตารางที่ 27 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบ
ปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถปานกลาง
และมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|-----|---------|-------|--------|
| ระหว่างกลุ่ม | 4 | 6.346 | 1.587 | 9.723* |
| ภายในกลุ่ม | 195 | 108.792 | | |
| - ระหว่างบุคคล | 39 | 83.338 | | |
| - ความคลาดเคลื่อน | 156 | 25.454 | | |
| รวม | 199 | 115.138 | 0.163 | |

* $p < .05$

จากตารางที่ 27 เป็นผลการวิเคราะห์ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบ
ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อผู้สอบมีความสามารถปานกลาง
และมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5
ครั้งมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 30



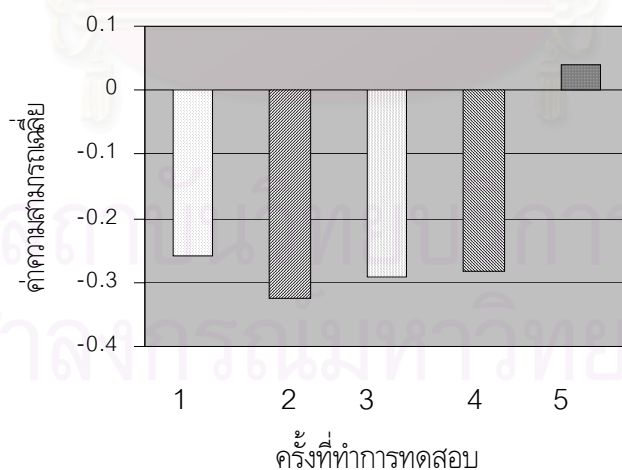
แผนภาพที่ 30 ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถปานกลาง และมีอัตราการใช้ ข้อสอบซ้ำ
30 %

ตารางที่ 28 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อผู้สอบมีความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|-----|--------|-------|--------|
| ระหว่างกลุ่ม | 4 | 3.037 | 0.759 | 3.466* |
| ภายในกลุ่ม | 175 | 52.968 | | |
| - ระหว่างบุคคล | 35 | 30.042 | | |
| - ความคลาดเคลื่อน | 140 | 28.926 | 0.214 | |
| รวม | 179 | 56.005 | | |

* $p < .05$

จากตารางที่ 28 เป็นผลการวิเคราะห์ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อผู้สอบมีความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้งมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 31



แผนภาพที่ 31 ค่าเฉลี่ยความสามารถของผู้สอบที่มีความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำ 30 %

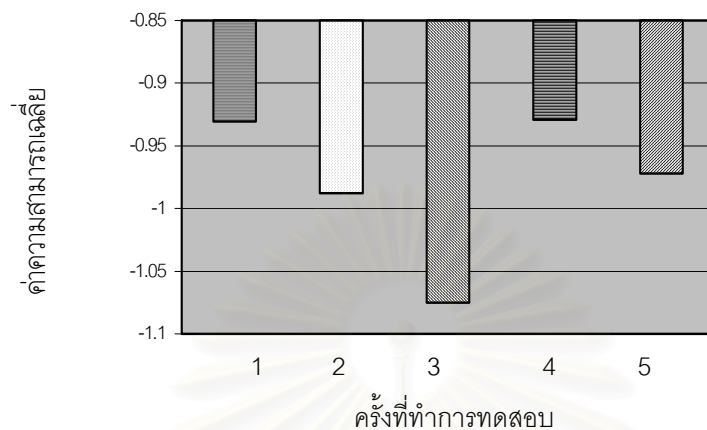
3. ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ และจำนวนครั้งที่ทำการทดสอบแตกต่างกัน

การวิเคราะห์ความแปรปรวน เพื่อพิจารณาความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ ข้อสอบและจำนวนครั้งที่ทำการทดสอบแตกต่างกัน โดยใช้สถิติวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (Repeated Measurement ANOVA) ปรากฏผลการวิเคราะห์ดังตารางที่ 29 - 31

ตารางที่ 29 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|-----|---------|-------|-------|
| ระหว่างกลุ่ม | 4 | 1.474 | 0.369 | 0.082 |
| ภายในกลุ่ม | 515 | 446.246 | | |
| - ระหว่างบุคคล | 103 | 354.741 | | |
| - ความคลาดเคลื่อน | 412 | 91.505 | 4.502 | |
| รวม | 519 | 447.720 | | |

จากตารางที่ 29 เป็นการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบครั้งที่ 1 ถึง 5 ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 32



แผนภาพที่ 32 ค่าเฉลี่ยความสามารถของผู้สอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์

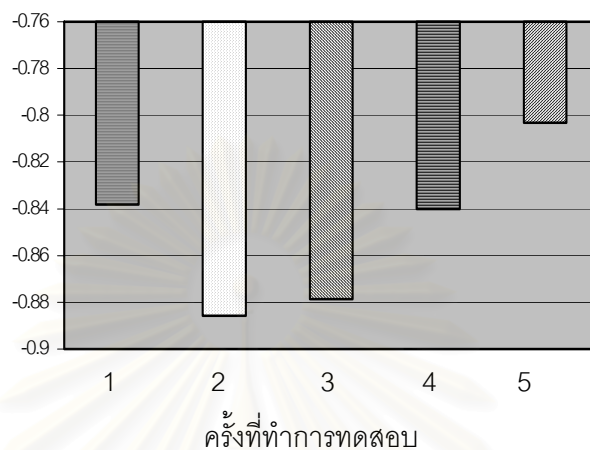
เมื่อทำการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ ปรากฏผลการวิเคราะห์ดังตารางที่ 30

ตารางที่ 30 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|-----|---------|-------|-------|
| ระหว่างกลุ่ม | 4 | 0.480 | 0.12 | 0.642 |
| ภายในกลุ่ม | 535 | 496.009 | | |
| - ระหว่างบุคคล | 107 | 416.015 | | |
| - ความคลาดเคลื่อน | 428 | 79.994 | 0.187 | |
| รวม | 539 | 496.489 | | |

จากตารางที่ 30 เป็นการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่อ

อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบครั้งที่ 1 ถึง 5 ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 33



แผนภาพที่ 33 ค่าเฉลี่ยความสามารถของผู้สอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์

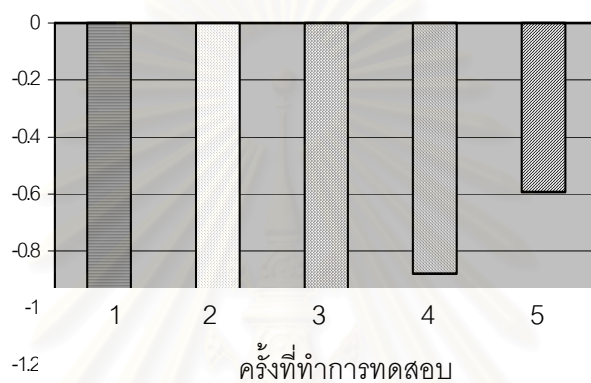
เมื่อทำการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ ปรากฏผลการวิเคราะห์ดังตารางที่ 31

ตารางที่ 31 ผลการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์

| แหล่งความแปรปรวน | Df | SS | MS | F |
|-------------------|---------|-----|--------|---------|
| ระหว่างกลุ่ม | 11.098 | 4 | 2.7744 | 17.212* |
| ภายในกลุ่ม | 415.100 | 535 | | |
| - ระหว่างบุคคล | 346.112 | 107 | | |
| - ความคลาดเคลื่อน | 68.988 | 428 | 0.161 | |
| รวม | 426.198 | 539 | | |

* $p < .05$

จากตารางที่ 31 เป็นการวิเคราะห์ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ระหว่างจำนวนครั้งที่ทำการทดสอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ พบว่า การประมาณค่าความสามารถของผู้สอบในการทดสอบครั้งที่ 1 ถึง 5 มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงใน แผนภาพที่ 34



แผนภาพที่ 34 ค่าเฉลี่ยความสามารถของผู้สอบ เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์

บทที่ 5

สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน โดยใช้โปรแกรมการทดสอบไมโครแคท (MicroCAT)

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2543 ในโรงเรียนมัธยมศึกษาขนาดใหญ่ สังกัดกรมสามัญศึกษา จำนวน 6 โรงเรียน จำนวน นักเรียน 319 คน

สมมุติฐานการวิจัย คือ ระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำที่ต่างกัน น่าจะส่งผลให้ความตรงตามสภาพ และความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน โดยที่ผู้สอบที่มีระดับความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำสูงขึ้น น่าจะส่งผลให้ความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถลดลง

เครื่องมือที่ใช้ในการวิจัย ประกอบด้วยแบบสอบวัดความสามารถทางภาษาอังกฤษ ระดับชั้นมัธยมศึกษาตอนต้น ซึ่งพัฒนาโดยกรมวิชาการ กระทรวงศึกษาธิการ (2540) คลังข้อสอบวิชาภาษาอังกฤษ ระดับชั้นมัธยมศึกษาตอนต้น ซึ่งพัฒนาโดยแมนมาศ สีสัตยกุล (2540) และโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ได้แก่ โปรแกรมการทดสอบไมโครแคท (MicroCAT)

การเก็บรวบรวมข้อมูล นำแบบสอบวัดความสามารถทางภาษาอังกฤษทดสอบนักเรียนที่เป็นกลุ่มตัวอย่าง ทำการประมาณค่าความสามารถของผู้สอบด้วยโปรแกรม BILOG 3 และจัดกลุ่มตัวอย่างตามระดับความสามารถต่ำ ปานกลาง และสูง จากนั้นทำการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 5 ครั้ง

การวิเคราะห์ข้อมูล ทำการวิเคราะห์ข้อมูลเบื้องต้น โดยการวิเคราะห์ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบและจำนวนข้อสอบที่ใช้ในการทดสอบ วิเคราะห์ค่าความตรงตามสภาพของการทดสอบ

โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถเฉลี่ยที่ได้จากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับค่าความสามารถที่ได้จากการทดสอบด้วยแบบสอบวัดความสามารถทางภาษาอังกฤษ จำแนกตามระดับความสามารถและอัตราการใช้ข้อสอบซ้ำ และเปรียบเทียบค่าความตรงตามสภาพของการทดสอบเป็นรายคู่ วิเคราะห์ความแปรปรวนของค่าความสามารถ เพื่อพิจารณาความคงที่ในการประมาณค่าความสามารถของผู้สอบ จำแนกตามแหล่งความแปรปรวน โดยใช้สถิติวิเคราะห์ความแปรปรวนแบบการวัดซ้ำ (Repeated Measurement ANOVA) และใช้วิธีการทดสอบของเชฟเฟ (Schffe') ทำการเปรียบเทียบความแตกต่างเป็นรายคู่

สรุปผลการวิจัย

จากผลการวิเคราะห์ข้อมูล สามารถสรุปผลการวิจัยตามวัตถุประสงค์ของการวิจัยได้ดังนี้

1. จากการเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำ แตกต่างกัน พบว่า

1.1 เมื่ออัตราการใช้ข้อสอบซ้ำแตกต่างกันในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ทำให้ค่าความตรงตามสภาพของการทดสอบแตกต่างกัน ระหว่างอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ กับอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ และระหว่างอัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ กับอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ และค่าความตรงตามสภาพของการทดสอบจะมีค่าสูงสุด เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ รองลงมาคือ อัตราการใช้ข้อสอบซ้ำ 20 และ 30 เปอร์เซ็นต์ ตามลำดับ

1.2 ผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษปานกลางจะมีค่าความตรงตามสภาพของการทดสอบสูงสุดในทุกอัตราการใช้ข้อสอบซ้ำ (10 20 และ 30 %) รองลงมาได้แก่ ผู้สอบที่มีความสามารถทางภาษาอังกฤษสูงและความสามารถทางภาษาอังกฤษต่ำ ตามลำดับ

1.3 ระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำที่ต่างกัน ทำให้ความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของ

ผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน โดยผู้สอบที่มีความสามารถสูงและมีอัตราการใช้ ข้อสอบซ้ำที่สูงขึ้น ส่งผลให้ความตรงตามสภาพของการทดสอบลดลง

2. จากการเปรียบเทียบความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ แตกต่างกัน พบว่า

2.1 ระดับความสามารถทางภาษาอังกฤษของผู้สอบ ส่งผลต่อการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ผู้สอบที่มีความสามารถสูงจะมีค่าความสามารถจากการทดสอบทั้ง 5 ครั้งแตกต่างกัน และมีแนวโน้มของการประมาณค่าความสามารถสูงขึ้น

2.2 อัตราการใช้ข้อสอบซ้ำ ส่งผลต่อการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 นั่นคืออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ ทำให้การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้งมีแนวโน้มสูงขึ้น

2.3 เมื่อพิจารณาปฏิสัมพันธ์ระหว่างตัวแปรระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำ ที่ส่งผลต่อความคงที่ในการประมาณค่าความสามารถของผู้สอบ พบว่า ระดับความสามารถทางภาษาอังกฤษของผู้สอบและอัตราการใช้ข้อสอบซ้ำ ร่วมกันส่งผลต่อความคงที่ในการประมาณค่าความสามารถของผู้สอบ กล่าวคือ เมื่อผู้สอบมีระดับความสามารถทางภาษาอังกฤษสูง และมีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น จะทำให้การประมาณค่าความสามารถของผู้สอบสูงมีแนวโน้มสูงขึ้นด้วย

2.4 เมื่อพิจารณาปฏิสัมพันธ์ระหว่างตัวแปรจำนวนครั้งที่ทำการทดสอบและอัตราการใช้ข้อสอบซ้ำ ที่ส่งผลต่อความคงที่ในการประมาณค่าความสามารถของผู้สอบ พบว่า ครั้งที่ทำการทดสอบและอัตราการใช้ข้อสอบซ้ำ ร่วมกันส่งผลต่อความคงที่ในการประมาณค่าความสามารถของผู้สอบ กล่าวคือ เมื่อจำนวนครั้งที่ทำการทดสอบมากขึ้น และมีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น จะทำให้แนวโน้มของการประมาณค่าความสามารถของผู้สอบสูงขึ้นด้วย

2.5 เมื่อพิจารณาความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ แตกต่างกัน พบว่า

2.5.1 เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ การประมาณค่าความสามารถของผู้สอบจากการทดสอบทั้ง 5 ครั้ง ของผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ ปานกลาง และสูง มีความคงที่

2.5.2 เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ การประมาณค่าความสามารถของผู้สอบจากการทดสอบทั้ง 5 ครั้ง ของผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ และปานกลางมีความคงที่ ส่วนผู้สอบที่มีความสามารถทางภาษาอังกฤษสูงจะมีการประมาณค่าความสามารถของผู้สอบไม่คงที่ โดยมีแนวโน้มของการประมาณค่าความสามารถของผู้สอบสูงขึ้น

2.5.3 เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ การประมาณค่าความสามารถของผู้สอบจากการทดสอบทั้ง 5 ครั้งของผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษต่ำ ปานกลาง และสูง ไม่คงที่ โดยมีแนวโน้มของการประมาณค่าความสามารถของผู้สอบสูงขึ้น

2.6 เมื่อพิจารณาความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ และครั้งที่ทำการทดสอบ แตกต่างกัน พบว่า

2.6.1 เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์และ 20 เปอร์เซ็นต์ การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้งไม่มีความแตกต่างกัน

2.6.2 เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ การประมาณค่าความสามารถของผู้สอบในการทดสอบทั้ง 5 ครั้งมีความแตกต่างกัน โดยที่การทดสอบในครั้งที่ 4 และ 5 การประมาณค่าความสามารถมีแนวโน้มสูงขึ้นจากการทดสอบในครั้งที่ 1

อภิปรายผลการวิจัย

ผลการวิจัยเกี่ยวกับการเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำแตกต่างกัน มีประเด็นสำคัญที่สามารถอภิปรายได้ดังนี้

1. การเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า ผู้สอบที่มีความสามารถปานกลางมีความตรงตามสภาพของการทดสอบสูงกว่าผู้สอบที่มีความสามารถต่ำและสูง ทั้งนี้เนื่องจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะทำการคัดเลือกข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบโดยดูจากค่าสารสนเทศของข้อสอบข้อที่สูงที่สุด และพบว่าภายในคลังข้อสอบประกอบด้วย ข้อสอบที่มีค่าสารสนเทศของข้อสอบข้อที่สูงที่สุดกระจายอยู่ในช่วงของระดับความสามารถปานกลาง ($-0.75 < \theta < 0.75$) มากกว่าการกระจายของข้อสอบที่มีค่าสารสนเทศของข้อสอบข้อที่สูงที่สุดในช่วงของระดับความสามารถต่ำ ($\theta \leq -0.75$) และความสามารถสูง ($\theta \geq 0.75$) ดังนั้นจึงทำให้ผู้สอบที่มีความสามารถปานกลางได้รับข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบมากกว่าผู้สอบที่มีความสามารถต่ำ และสูง จึงทำให้การประมาณค่าความสามารถของผู้สอบที่มีความสามารถปานกลางมีความถูกต้องแม่นยำมากกว่าการประมาณค่าความสามารถของผู้สอบที่มีความสามารถต่ำ และสูง ซึ่งความถูกต้องแม่นยำเกี่ยวข้องโดยตรงกับความตรงตามสภาพของการทดสอบ

2. การเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบซ้ำแตกต่างกัน โดยผู้วิจัยนำค่าความสามารถที่ได้จากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มาเปรียบเทียบกับเกณฑ์ คือ ค่าความสามารถที่ได้จากการทดสอบด้วยแบบสอบวัดความสามารถ พบว่าค่าความตรงตามสภาพของการทดสอบมีความแตกต่างกัน ซึ่งผลการวิจัยนี้สอดคล้องกับข้อค้นพบของ รีเวิลต้าและ พอนโซดา (Revuelta and Ponsoda, 1998) พบว่า การที่ผู้สอบทำการทดสอบหลาย ๆ ครั้งด้วยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยใช้คลังข้อสอบเดิม ส่งผลให้ความตรงตามสภาพของการทดสอบลดลง ทั้งนี้เนื่องจากการทดสอบที่มีอัตราการใช้ข้อสอบซ้ำ เป็นการเปิดโอกาสให้ผู้สอบได้รับข้อสอบซ้ำจากการทดสอบในครั้งที่ผ่านมา และถ้าอัตราการใช้ข้อสอบซ้ำยิ่งสูงขึ้น ก็จะทำให้ผู้สอบได้รับข้อสอบซ้ำจำนวนมากขึ้น ซึ่งจะช่วยให้ผู้สอบสามารถจดจำข้อสอบเหล่านั้นได้ และเกิดการเรียนรู้จากการตอบข้อสอบ ทำให้ไม่ให้เกิดจาก

คุณลักษณะภายในของผู้สอบอย่างแท้จริงเมื่อใช้ข้อสอบเดิม ส่งผลให้ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบสูงขึ้น นั่นคือมีความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบต่ำลง และค่าความถูกต้อง แม่นยำในการประมาณค่าความสามารถของผู้สอบนั้นเกี่ยวข้องกับโดยตรงกับความตรงตามสภาพของการทดสอบ ดังนั้นการทดสอบที่มีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น ก็จะทำให้ค่าความตรงตามสภาพของการทดสอบลดลง

3. การเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ แตกต่างกัน โดยผู้วิจัยนำค่าความสามารถของผู้สอบในแต่ละระดับความสามารถ และอัตราการใช้ ข้อสอบซ้ำที่ได้จากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ มาทำการเปรียบเทียบกับเกณฑ์ คือค่าความสามารถของผู้สอบที่ได้จากการทดสอบด้วยแบบสอบวัดความสามารถ พบว่า ค่าความตรงตามสภาพของการทดสอบมีความแตกต่างกัน โดยผู้สอบที่มีระดับความสามารถทางภาษาอังกฤษสูง และมีอัตราการใช้ข้อสอบซ้ำสูงขึ้นไป จะส่งผลให้ความตรงตามสภาพของการทดสอบลดลง ซึ่งผลการวิจัยนี้ตรงกับสมมุติฐานที่กำหนดไว้ และสอดคล้องกับงานวิจัยของ สต็อกกิงและเลวิส (Stocking and Lewis, 1998) พบว่า ระดับความสามารถของผู้สอบมีความสัมพันธ์กับอัตราการใช้ข้อสอบซ้ำ ถ้าระดับความสามารถของผู้สอบสูงหรือต่ำกว่าค่าเฉลี่ยความยากของข้อสอบ ซึ่งจะส่งผลต่อความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ นั่นคือความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบนั้นเกี่ยวข้องกับโดยตรงกับความตรงตามสภาพ และเนื่องจากค่าความสามารถของผู้สอบนั้นจะส่งผลต่อค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกต้อง และเมื่อ ผู้สอบได้รับอัตราการใช้ข้อสอบซ้ำที่ต่างกัน นั่นคือจำนวนข้อสอบซ้ำที่ผู้สอบได้รับการทดสอบในครั้งที่ผ่านมามีความแตกต่างกัน ซึ่งจะส่งผลต่อความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องแตกต่างกัน โดย ผู้สอบที่มีระดับความสามารถสูง และได้รับข้อสอบที่มีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น โอกาสที่จะทำ ให้ ผู้สอบสามารถจดจำข้อสอบได้สูง และมีความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องมากขึ้น ซึ่งไม่ได้เกิดจากคุณลักษณะภายในของผู้สอบอย่างแท้จริง ทำให้ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบมีค่าสูงขึ้น ทำให้ค่าความตรงตามสภาพของการทดสอบต่ำลง ดังนั้น ผู้สอบที่มีความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น จะส่งผลให้ความตรงตามสภาพของการทดสอบลดลง

สถาบันวิจัยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4. การเปรียบเทียบความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 5 ครั้ง เมื่อระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำแตกต่างกัน พบว่า ระดับความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำแตกต่างกันในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ส่งผลต่อความคงที่ในการประมาณค่าความสามารถของผู้สอบ โดยผู้สอบที่มีระดับความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำสูงขึ้น ส่งผลให้ความคงที่ในการประมาณค่าความสามารถของผู้สอบลดลง ผลการวิจัยนี้ตรงตามสมมุติฐานที่กำหนดไว้ และสอดคล้องกับงานวิจัยของสต็อกกิง และเลวิส (Stocking and Lewis, 1998) พบว่า ระดับความสามารถของผู้สอบจะมีความสัมพันธ์กับอัตราการใช้ข้อสอบซ้ำ ถ้าระดับความสามารถของผู้สอบสูงหรือต่ำกว่าค่าเฉลี่ยความยากของข้อสอบ ซึ่งส่งผลต่อความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ และพบว่าในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้นจะต้องมีการควบคุมอัตราการใช้ข้อสอบซ้ำเพื่อป้องกันผู้สอบที่สามารถจดจำข้อสอบได้ ซึ่งจะส่งผลต่อความน่าจะเป็นในการตอบข้อสอบได้ถูกต้อง ทำให้ค่าความสามารถของผู้สอบเปลี่ยนแปลงไปจากค่าความสามารถที่แท้จริง โดยเฉพาะอย่างยิ่งถ้าผู้สอบที่มีระดับความสามารถสูงและได้รับข้อสอบที่มีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น ทำให้ผู้สอบสามารถจดจำข้อสอบที่ซ้ำได้มากขึ้น และเกิดการเรียนรู้จากข้อสอบเหล่านั้น ส่งผลให้ความน่าจะเป็นในการตอบ ข้อสอบได้ถูกต้องมีมากขึ้น ทำให้ค่าความสามารถของผู้สอบในการทดสอบครั้งต่อ ๆ ไปมีแนวโน้มสูงขึ้นด้วย จึงทำให้ความคงที่ในการประมาณค่าความสามารถของผู้สอบลดลง

5. การเปรียบเทียบความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 5 ครั้ง พบว่าจำนวนครั้งที่ทำการทดสอบและอัตราการใช้ข้อสอบซ้ำที่แตกต่างกัน ส่งผลต่อความคงที่ในการประมาณค่าความสามารถของผู้สอบ โดยการทดสอบที่มีอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ การประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ทั้ง 5 ครั้งมีความแตกต่างกัน ทั้งนี้เนื่องจากในการทดสอบแต่ละครั้งผู้สอบจะได้รับข้อสอบซ้ำ 30 เปอร์เซ็นต์ของจำนวนข้อสอบที่ได้ทำในการทดสอบครั้งที่ผ่านมา โอกาสที่ผู้สอบจะสามารถจดจำข้อสอบได้มีมากขึ้น ทำให้ความน่าจะเป็นที่ผู้สอบจะสามารถตอบข้อสอบได้ถูกต้องสูงขึ้นด้วย ส่งผลให้ค่าความสามารถของผู้สอบในการทดสอบแต่ละครั้งเปลี่ยนแปลงไป และมีแนวโน้มสูงขึ้น ดังนั้นในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ควรจำกัดให้มีอัตราการใช้ข้อสอบซ้ำไม่เกิน 20 เปอร์เซ็นต์ เพื่อให้การประมาณค่าความสามารถของผู้สอบมีความคงที่ และถูกต้องแม่นยำ ซึ่งสอดคล้องกับงานวิจัยของสต็อกกิง และเลวิส (Stocking and Lewis, 1998) พบว่า ในการทดสอบแบบปรับเหมาะ

ความสามารถของผู้สอบมีการประมาณค่าความสามารถของผู้สอบที่ความถูกต้องแม่นยำนั้น การทดสอบแต่ละครั้งควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 20 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา

6. ในเรื่องความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า ผู้สอบที่มีระดับความสามารถปานกลาง และระดับความสามารถต่ำ ควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 20 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา และผู้สอบที่มีระดับความสามารถสูง ควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 10 เปอร์เซ็นต์ของการจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา ซึ่งสอดคล้องกับงานวิจัยของสต็อกกิง และเลวิส (Stocking and Lewis, 1998) พบว่า ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีการประมาณค่าความสามารถของผู้สอบที่ความถูกต้องแม่นยำนั้น การทดสอบแต่ละครั้งควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 20 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา นอกจากนี้ คาลอห์น และสเปรย์ (Kalohn and Spray, 1998) พบว่าผู้สอบที่มีระดับความสามารถสูงควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 10 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา ทั้งนี้เนื่องจากผู้สอบที่มีระดับความสามารถสูง และได้รับข้อสอบที่มีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้น ทำให้ผู้สอบได้รับข้อสอบซ้ำจากการทดสอบครั้งที่ผ่านมาสูงขึ้น สามารถจดจำ ข้อสอบที่ซ้ำได้มากขึ้น และเกิดการเรียนรู้จากข้อสอบเหล่านั้น ส่งผลให้การประมาณค่าความสามารถของผู้สอบเปลี่ยนแปลงไปและมีแนวโน้มสูงขึ้น จึงควรจำกัดให้ผู้สอบที่มีระดับความสามารถสูงมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 10 เปอร์เซ็นต์ เพื่อความคงที่และถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ

ข้อเสนอแนะ

จากผลการวิจัยที่ได้ ผู้วิจัยใคร่จะเสนอแนะสำหรับผู้ที่จะนำผลการวิจัยนี้ไปใช้ และผู้ที่คิดจะค้นคว้าวิจัยเรื่องในทำนองนี้ต่อไป ดังนี้

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. จากผลการวิจัยได้ชี้ให้เห็นว่าอัตราการใช้ข้อสอบซ้ำส่งผลกระทบต่อความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ดังนั้นในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ไม่ควรใช้อัตราการใช้ข้อสอบซ้ำ หากมีอัตราการใช้ข้อสอบซ้ำก็ไม่ควรเกิน 20 เปอร์เซ็นต์สำหรับผู้สอบที่มีระดับความสามารถ

ต่ำ และปานกลาง และไม่ควรเกิน 10 เปอร์เซ็นต์สำหรับผู้สอบที่มีระดับความสามารถสูง จึงจะทำให้การทดสอบมีความถูกต้องมีความตรงตามสภาพ และการประมาณค่าความสามารถของผู้สอบมีความคงที่

2. ควรนำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มาใช้ในการทดสอบให้กว้างขวางขึ้น ทั้งนี้เพราะการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ มีข้อดีว่าการทดสอบแบบเขียนตอบที่ใช้กันอยู่ เนื่องจากประหยัดค่าใช้จ่าย ทั้งนี้เพราะใช้จำนวนข้อสอบน้อย และไม่จำเป็นต้องสร้างข้อสอบใหม่ทุกครั้งที่มีการทดสอบ เนื่องจากมีข้อสอบเก็บไว้ในคลังข้อสอบ ผู้สอบไม่เกิดความเหน็ดเหนื่อยและเบื่อหน่ายในการสอบ เพราะใช้เวลาในการทดสอบน้อยลงกว่าเดิม ข้อสอบที่ใช้มีความยากง่ายเหมาะสมกับความสามารถของผู้สอบ

3. การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์หรือการทดสอบแบบซี เอ ที เป็นการทดสอบรายบุคคลที่ค่อนข้างมีความถูกต้องแม่นยำ การทดสอบแต่ละครั้งไม่จำเป็นต้องทำพร้อมกันทุกคน ดังนั้น จำนวนของเครื่องคอมพิวเตอร์จึงไม่ใช่ปัญหาอุปสรรคสำหรับการทดสอบลักษณะนี้ เมื่อผู้สอบต้องการที่จะทดสอบก็สามารถมาทำการทดสอบได้ด้วยตนเอง ผลการทดสอบแต่ละครั้งจะถูกบันทึกไว้ซึ่งจะเป็นแหล่งข้อมูลแหล่งหนึ่งที่จะช่วยให้ครูผู้สอนสามารถติดตามความก้าวหน้าทางการเรียนการสอนได้

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. อัตราการใช้ข้อสอบซ้ำส่งผลต่อความตรงตามสภาพ และความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถด้วยคอมพิวเตอร์ ดังนั้นควรมีการศึกษาวิธีการคัดเลือกข้อสอบที่สามารถควบคุมอัตราการใช้ข้อสอบซ้ำ เช่น วิธีการคัดเลือกข้อสอบเพื่อควบคุมอัตราการใช้ข้อสอบซ้ำด้วยวิธีของซิมสันและเฮตเตอร์ (Symson and Hetter Method) กับวิธีการคัดเลือกข้อสอบโดยวิธีสต็อกกิงและเลวิส (Stocking and Lewis Method) ว่าวิธีการคัดเลือกข้อสอบวิธีใดสามารถทำให้การทดสอบมีประสิทธิภาพสูงสุด

2. ควรมีการศึกษาการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับคลังข้อสอบเนื้อหาวิชาอื่นที่ไม่ใช่วิชาภาษาอังกฤษ

3. ควรมีการศึกษาวិธีการที่จะนำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ไปใช้ในลักษณะเครือข่าย (Wide area network : WAN)

4. ควรมีการศึกษาเกี่ยวกับการกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อว่าจะส่งผลในลักษณะใดต่อการประมาณค่าความสามารถของผู้สอบ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

ภาษาไทย

- วิชาการ, กรม. (2540). **แบบสอบวัดความสามารถทางภาษาอังกฤษ**. กรุงเทพมหานคร : กระทรวงศึกษาธิการ.
- เกรียงศักดิ์ สุวรรณภาค. (2532). **ความสัมพันธ์ระหว่างรูปแบบบางรูปแบบของแบบทดสอบแยกกลุ่มกับความสามารถทางการเรียน ในวิชาคณิตศาสตร์**. วิทยานิพนธ์การศึกษาศาสตรมหาบัณฑิต. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- จิราพร ไกรสรคิวงเท. (2529). **ความสัมพันธ์ระหว่างรูปแบบบางประการของการทดสอบรูปปริมาตรกับความสามารถทางการเรียนวิชาคณิตศาสตร์**. วิทยานิพนธ์การศึกษาศาสตรมหาบัณฑิต. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ตริงใจ พูนผลอำนวย. (2534). **พัฒนาแบบสอบเฉพาะบุคคลในวิชาวิทยาศาสตร์สำหรับ นักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้คอมพิวเตอร์**. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. จุฬาลงกรณ์มหาวิทยาลัย.
- ต่าย เชียงฉี. (2534). **การศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบเทเลอร์รูปปริมาตรที่มีรูปแบบ จำนวนชั้นและวิธีการให้คะแนนที่แตกต่างกันโดยใช้วิธีมอนติคาร์โล**. วิทยานิพนธ์การศึกษาศาสตรดุษฎีบัณฑิต. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- นันทิยา พึ่งคำ. (2531). **การเปรียบเทียบคุณภาพการทดสอบแบบซีเอทีและแบบประเพณีนิยมในการวัดความสามารถด้านคำศัพท์ภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3**. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. จุฬาลงกรณ์มหาวิทยาลัย.
- แมนมาศ ลีลสัตยกุล. (2540). **เอกสารการพัฒนาคั้งข้อสอบรายวิชาตามหลักสูตรมัธยมศึกษา. คณะศึกษาศาสตร์. มหาวิทยาลัยเกษตรศาสตร์**
- ภาวิณี ศรีสุขวัฒนานันท์. (2543). **เอกสารการประชุมเชิงปฏิบัติการ “เทคนิคการออกข้อสอบ” รุ่นที่ 5**. คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์.
- รังสรรค์ มณีเล็ก. (2540). **ผลของตัวแปรบางตัวต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์**. วิทยานิพนธ์การศึกษาศาสตรดุษฎีบัณฑิต. มหาวิทยาลัยศรีนครินทรวิโรฒ.

- ศิริชัย กาญจนวาสี. (2538). **การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive testing)**. กรุงเทพมหานคร : ศูนย์ทดสอบทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
- ศิริชัย กาญจนวาสี. (2538). **รวมบทความวิชาทฤษฎีการวัดและประเมิน**. คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย พงษ์วิชัย. (2543). **การวิเคราะห์ข้อมูลทางสถิติด้วยคอมพิวเตอร์**. พิมพ์ครั้งที่ 9. กรุงเทพมหานคร : จุฬาลงกรณ์มหาวิทยาลัย.
- สายชล ออบทม. (2539). **การพัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์**. วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ จุฬาลงกรณ์มหาวิทยาลัย
- สุพัฒน์ สุกมลสันต์. (2539). **ธนาคารข้อทดสอบและการทดสอบปรับเปลี่ยนด้วยคอมพิวเตอร์**. กรุงเทพมหานคร : บริษัทพิมพ์ดี จำกัด.
- เสรี ปรมชวลิตโรจน์. (2531). **การศึกษาความเที่ยงตรงในการจำแนกความรู้และการประมาณค่าความสามารถในวิชาคณิตศาสตร์ของแบบทดสอบเพลิกซีเลเวลและแบบทดสอบรูปปิรามิดที่ใช้ขนาดชั้นแปรผัน**. วิทยานิพนธ์การศึกษาศาสตรมหาบัณฑิต. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ไอลดา คล้ายสำริด. (2538). **การเปรียบเทียบความคลาดเคลื่อนในการประมาณค่าประสิทธิภาพสัมพัทธ์ในการประมาณค่าและความสามารถของผู้สอบ ในแบบสอบวิชาคณิตศาสตร์ที่ใช้วิธีต่างกัน**. วิทยานิพนธ์การศึกษามหาบัณฑิต. มหาวิทยาลัยศรีนครินทรวิโรฒ.

ภาษาอังกฤษ

- Assessment Systems Corporation. (1995). **User's Manual for the MicroCAT Testing System (third Edition)**. Minesota : Assessment Systems Corporation.
- Blackmore, L.M. (1987). "Computerized, Computerized Adaptive and Pencil -and- Paper Test Administration : A Comparative Study in High School Setting" **Dissertation Abstracts International** 47(07) : 2554 - A.
- Chang, S. (1998). " A Comparative Study of Item Exposure Control Methods in Computerized Adaptive Testing (Act , Test Security)" **Dissertation Abstracts International** 59(09) : 3418.

- Glass, G.V. and J.C.Stanley. (1970). **Statistical Methods in Education and Psychology**. Englewood Cliffs,NJ. : Prentice-Hall.
- Grimm, G.L. (1993). **Statistical Applications for the Behavioral Sciences**. New York : John Wiley & sons, Inc.
- Hambleton, R.K. and H. Swaminathan. (1985). **Item Response Theory : Principles and Application**. Boston : Kluner – Nijhoff Publishing.
- Hankins, J.A. (1987). "The Effort of Variable Entry on Bias and Information of the Bayesian Adaptive Testing Procedure" **Dissertation Abstracts International**. 47(08) : 3013 – A.
- Ho, R.G. (1989). "Using MicroCAT in Computerized Adaptive Testing : A Comparison of Three Adaptive Testing Strategies" **Dissertation Abstracts International**. 50(2) : 421 – A
- 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**.
- Lord, F.M. and M.R. Novick. (1968). **Statistical Theories of Mental Test Score**. Massachusetts : Addison Wesley Publishing Company.
- Lord, F.M. and M.R. Novick. (1980). **Applications of Item Response Theory to Practical Testing Problem**. New Jersey : Lawrence Erlbaum Associates.
- Owen, Roger J. (1975). "A Bayesian Sequential Procedure for Quantal Response in the Context of Adaptive Mental Testing" **Journal of the American Statistical Association** 70(350) : 351-356; June.
- Revuelta, J. & Ponsoda,V. (1998). "A Comparison of Item Exposure Control Methods in Computerized Adaptive Testing" **Journal of Education Measurement** 34(4) : 311 – 327.
- Stocking, M.L. & Lewis, C. (1998). "Controlling Item Exposure Conditional on Ability in Computerized Adaptive Testing" **Journal of Education and Behavioral Statistics** 23(1) : 57 – 75.

- Urry, V.W. (1977). "Tailored Testing : A Successful Application of Latent Trait Theory" **Journal of Education Measurement** 14(2) : 181 - 196.
- Weiss, D.J. (1974). "Strategies of Adaptive Measurement" **Research Report 74 - 5**.
Psychometric Method Program. Department of Psychology, University of Minnesota, Minneapolis.
- Weiss, D.J. and G.G. Kingsbury. (1984). "Application of Computerized Adaptive testing to Education Problem" **Journal of Education Measurement** 21(4) : 361 - 375.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ที่ ทม 0302(2702)/



ภาควิชาวิจัยการศึกษา
คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย

23 สิงหาคม 2543

เรื่อง ขอความร่วมมือในการวิจัย

เรียน ผู้ช่วยศาสตราจารย์ ดร.แมนมาศ ลีลลัตยกุล

เนื่องด้วย นางสาวกนกวรรณ รัตนธณ นิสิตชั้นปริญญาโท สาขาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กำลังดำเนินการวิจัย เพื่อเสนอเป็นวิทยานิพนธ์เรื่อง "การเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์" โดยมีรองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี เป็นอาจารย์ที่ปรึกษา ในการนี้ นิสิตจำเป็นต้องใช้คลังข้อสอบวิชาภาษาอังกฤษระดับชั้นมัธยมศึกษาปีที่ 1 เพื่อประกอบการทำวิจัย

จึงเรียนมาเพื่อขอความอนุเคราะห์ข้อมูล และขอขอบคุณเป็นอย่างสูงมา ณ โอกาสนี้ด้วย

ขอแสดงความนับถือ

(รองศาสตราจารย์พวงแก้ว ปุณยกนก)

หัวหน้าภาควิชาวิจัยการศึกษา

สถาบันวิจัยการศึกษา
จุฬาลงกรณ์มหาวิทยาลัย



ที่ ทม 0302(2702)/

ภาควิชาวิจัยการศึกษา
คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย

24 สิงหาคม 2543

เรื่อง ขอความร่วมมือในการวิจัย

เรียน ผู้อำนวยการสำนักงานทดสอบทางการศึกษา

เนื่องด้วย นางสาวกนกวรรณ รัตนธน นิสิตชั้นปริญญาโทมหาบัณฑิต สาขา การวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์ มหาวิทยาลัย กำลังดำเนินการวิจัยเพื่อเสนอเป็นวิทยานิพนธ์เรื่อง “การเปรียบเทียบความตรงตาม สภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความ สามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อความสามารถของผู้สอบ และอัตราการให้ข้อสอบซ้ำ แตกต่างกัน” โดยมีรองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี เป็นอาจารย์ที่ปรึกษา ในกรณีนี้ นิสิต จำเป็นต้องใช้แบบสอบวัดความสามารถทางภาษาอังกฤษระดับชั้นมัธยมศึกษาตอนต้น เพื่อนำมา ทดสอบนักเรียนที่เป็นกลุ่มตัวอย่าง และแบ่งกลุ่มตัวอย่างออกเป็น 3 กลุ่ม คือกลุ่มที่มีระดับความ สามารถสูง ปานกลาง และต่ำ

จึงเรียนมาเพื่อขอความอนุเคราะห์และอนุญาตให้ใช้แบบสอบดังกล่าว ขอขอบคุณ เป็นอย่างสูงมา ณ โอกาสนี้ด้วย

ขอแสดงความนับถือ

(รองศาสตราจารย์พวงแก้ว ปุณยกันก)

หัวหน้าภาควิชาวิจัยการศึกษา

ที่ ทม 0302(2702)/

ภาควิชาวิจัยการศึกษา
คณะครุศาสตร์

ที่ ทม 0302(2702)/



ภาควิชาวิจัยการศึกษา
คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย

24 สิงหาคม 2543

เรื่อง ขอความร่วมมือในการวิจัย

เรียน หัวหน้าภาควิชาการศึกษา

เนื่องด้วย นางสาวกนกวรรณ รัตนธน นิสิตชั้นปริญญาโทบัณฑิต สาขาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กำลังดำเนินการวิจัยเพื่อเสนอเป็นวิทยานิพนธ์เรื่อง "การเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบด้วยคอมพิวเตอร์" โดยมีรองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี เป็นอาจารย์ที่ปรึกษา ในการนี้ นิสิตจำเป็นต้องขอรับคำปรึกษาจากผู้เชี่ยวชาญทางด้านระบบการทดสอบด้วยคอมพิวเตอร์ (โปรแกรม MicroCAT) จึงขอเรียนเชิญผู้เชี่ยวชาญศาสตราจารย์ ดร.ภาวิณี ศรีสุขวัฒนานันท์ เป็นอาจารย์ที่ปรึกษาร่วมวิทยานิพนธ์

จึงเรียนมาเพื่อทราบ และโปรดพิจารณาอนุญาต

ขอแสดงความนับถือ

(รองศาสตราจารย์พวงแก้ว ปุณยกนก)

หัวหน้าภาควิชาวิจัยการศึกษา

สถาบันวิจัยการศึกษาคณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย



ที่ ทม 0302(2702)/590

ภาควิชาวิจัยการศึกษา
คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย

7 พฤศจิกายน 2543


เรื่อง ขอความร่วมมือในการวิจัย

เรียน ผู้อำนวยการโรงเรียนอุ้มทอง

เนื่องด้วย นางสาวกนกวรรณ รัตนธน์ นิสิตชั้นปริญญาโท สาขา การวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์ มหาวิทยาลัย กำลังดำเนินการวิจัยเพื่อเสนอเป็นวิทยานิพนธ์เรื่อง "การเปรียบเทียบความตรงตาม สภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความ สามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำ แตกต่างกัน" โดยมีรองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี เป็นอาจารย์ที่ปรึกษา ในการนี้ นิสิต จำเป็นต้องทำการทดลองใช้เครื่องมือ โดยนำแบบสอบวัดความสามารถทางภาษาอังกฤษทดสอบ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 2 ห้องเรียน

จึงเรียนมาเพื่อขอความอนุเคราะห์จากท่านได้โปรดพิจารณาอนุญาตให้ นางสาวกนกวรรณ รัตนธน์ ได้ทำการทดลองใช้เครื่องมือดังกล่าว และขอขอบคุณเป็นอย่างสูงมา ณ โอกาสนี้ด้วย

ขอแสดงความนับถือ


(ผู้ช่วยศาสตราจารย์ ดร. สุชาดา บวรกิตติวงศ์)

หัวหน้าภาควิชาวิจัยการศึกษา



ที่ ทม0302(2770.0603)2298

สำนักงานฝ่ายจัดการศึกษา คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย
ถนนพญาไท กรุงเทพฯ 10330

10 พฤศจิกายน 2543

เรื่อง ขอความร่วมมือในการเก็บข้อมูลวิจัย และขอใช้บริการคอมพิวเตอร์

เรียน ผู้อำนวยการโรงเรียนธรรมโชติศึกษาลัย

สิ่งที่ส่งมาด้วย เครื่องมือที่ใช้ในการวิจัย

ด้วย นางสาวกนกวรรณ รัตนธณ นิสิตชั้นปริญญาโทบัณฑิต ภาคศึกษาวิจัยการศึกษา สาขาวิชา การวัดและประเมินผลการศึกษา อยู่ในระหว่างการทำนิพนธ์วิทยานิพนธ์เรื่อง "การเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความ สามารถของผู้สอบด้วยคอมพิวเตอร์เพื่อความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน" โดยมี รองศาสตราจารย์ ดร.ศิริชัย ภาณุจนวนาสี และ ผู้ช่วยศาสตราจารย์ ดร.ภาวิณี ศรีสุขวัฒนานันท์ เป็นอาจารย์ ที่ปรึกษา ในการนี้ นิสิตมีความจำเป็นต้องขอความร่วมมือในการเก็บข้อมูลวิจัยกับกลุ่มตัวอย่างซึ่งเป็นนักเรียนชั้น มัธยมศึกษาปีที่ 3 พร้อมนี้ขอความอนุเคราะห์ในการขอใช้เครื่องคอมพิวเตอร์ของโรงเรียนด้วย ทั้งนี้ นิสิตจะ ประสานงานในรายละเอียดต่อไป

จึงเรียนมาเพื่อขอความอนุเคราะห์จากท่านโปรดอนุญาตให้ นางสาวกนกวรรณ รัตนธณ ดำเนินการ ดังกล่าว เพื่อประโยชน์ทางวิชาการต่อไป และขอขอบคุณมาในโอกาสนี้

ขอแสดงความนับถือ

1363/2543 11/11/2000

(รองศาสตราจารย์ ดร.เจียรณี นิมมวอล)

รองคณบดีฝ่ายจัดการศึกษา

ปฏิบัติราชการแทนคณบดีคณะครุศาสตร์

สำนักงานฝ่ายจัดการศึกษา (ระดับบัณฑิตศึกษา)

โทร. 218-2682

โปรแกรมที่ใช้ในการประมาณค่าความสามารถของผู้สอบ

```
>TITLE TESTING;  
>COMMENTS ITEMS FOR ABILITY GROUP;  
>GLOBAL DFNAME='C:\EXAM1.DAT',KFNAME='C:\EXAM1.DAT',  
    NPARAM=2,OMITS,SAVE;  
>SAVE COV='C:\EXAM1.COV';  
>LENGTH NITEMS=50;  
>INPUT NTOT=50,SAMPLE=1498,NALT=4,NIDC=4,KFNAME='C:\EXAM1.DAT',  
    OFNAME='C:\EXAM1.DAT';  
    (4A1,T5,50A1)  
>TEST TNAME=EXAM1;  
>CALIB FLOAT;  
>SCORE RSC=3;
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

โปรแกรมที่ใช้ในการวิเคราะห์คุณภาพแบบสอบภาษาอังกฤษ

```
>TITLE TESTING;  
>COMMENTS ITEMS FOR ABILITY GROUP;  
>GLOBAL DFNAME='C:\TRYOUT2.DAT',KFNAME='C:\TRYOUT2.DAT',  
    NPARAM=2,OMITS,SAVE;  
>SAVE COV='C:\TRYOUT2.COV';  
>LENGTH NITEMS=50;  
>INPUT NTOT=50,SAMPLE=254,NALT=4,NIDC=4,KFNAME='C:\TRYOUT2.DAT',  
    OFNAME='C:\TRYOUT2.DAT';  
    (4A1,T5,50A1)  
>TEST TNAME=TRYOUT2;  
>CALIB FLOAT;  
>SCORE RSC=3;
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารของผู้สอบด้วยคอมพิวเตอร์

โดยใช้ MicroCAT รอบที่ 1

TEST PRETEST

Declare String: IsRight

Declare String: FileExam

DECLARE INTEGER : M

#DEF001

@userID = @FreeResponse

Set keepFileUser = @UserID + "1.OUT"

@KeepFileName = KeepFileUser

#DEF002

@Keepline("Examinee ID : ",@UserID)

@Keepline("Examinee Name : ",@FreeResponse)

!Declare real : x

SETSCORE @Bayesian(Mean, Variance, Popmean, Popvar)

Set Variance = 1

AUTOEXEC

! Convert Error Variance to Standard Error

SET StdError = @SQRT(Variance)

IF @ItemResponse = @ItemKey

IsRight = " X "

ELSE

IsRight = " "

ENDIF

ENDAUTOEXEC

@KeepLine

@KeepLine(" Std")

@KeepLine(" Item # Item ID Resp Key Correct Ability Error")

@KeepLine(" -----")

@KeepLine

```
AutoKeep(@ItemsAdministered,@ItemID,@ItemResponse, &
  @ItemKey," ", IsRight, " ", Mean," ", StdError,@NL)
SET Popmean = 0
SET Popvar = 1
TERMINATE $DONE (StdError <= 0.3)
SEARCH Mean
  #ENG001
  #ENG002

  #ENG229
ENDSEARCH
$DONE @Keep(@NL, @ItemsAdministered, "items were administered.", @NL)
@KeepLine ("The final Bayesian posterior mean was: ", Mean)
@KeepLine ("The final Bayesian posterior variance was: ", Variance)
@KeepLine
ENDTEST
```



```

AutoKeep(@ItemsAdministered,@ItemID,@ItemResponse, &
  @ItemKey," ", IsRight, " ", Mean," ", StdError,@NL)
SET Popmean = 0
SET Popvar = 1
TERMINATE $DONE (StdError <= 0.3) OR (@ITEMSADMINISTERED = 1)
  SEARCH MEAN
    {I A.EXM}
  ENDSEARCH
$DONE
TERMINATE $DONE1 (StdError <= 0.3)
  SEARCH Mean.
    {I B.EXM}
  ENDSEARCH
$DONE1 @Keep(@NL, @ItemsAdministered, "items were administered.", @NL)
@KeepLine ("The final Bayesian posterior mean was: ", Mean)
@KeepLine ("The final Bayesian posterior variance was: ", Variance)
@KeepLine
ENDTEST

```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบ

| Item | a | b | c | Item | a | b | c |
|------|-------|--------|-------|------|-------|--------|-------|
| 1 | 0.927 | -0.860 | 0.230 | 28 | 1.868 | 0.507 | 0.310 |
| 2 | 1.255 | -0.962 | 0.310 | 29 | 1.575 | 3.000 | 0.250 |
| 3 | 1.032 | -0.556 | 0.280 | 30 | 1.444 | 3.000 | 0.280 |
| 4 | 1.742 | 1.138 | 0.180 | 31 | 1.882 | 0.789 | 0.220 |
| 5 | 1.514 | 0.505 | 0.180 | 32 | 0.761 | -0.135 | 0.240 |
| 6 | 1.694 | -0.301 | 0.230 | 33 | 2.500 | 3.000 | 0.260 |
| 7 | 1.172 | -1.063 | 0.190 | 34 | 1.793 | 1.301 | 0.220 |
| 8 | 0.451 | -0.069 | 0.340 | 35 | 0.802 | -3.000 | 0.260 |
| 9 | 2.500 | 3.000 | 0.230 | 36 | 1.118 | -1.103 | 0.210 |
| 10 | 1.664 | -2.275 | 0.250 | 37 | 1.366 | 1.961 | 0.340 |
| 11 | 1.104 | -1.033 | 0.220 | 38 | 0.959 | -0.601 | 0.190 |
| 12 | 1.056 | -1.635 | 0.230 | 39 | 0.403 | -2.111 | 0.220 |
| 13 | 0.949 | 0.199 | 0.210 | 40 | 1.207 | -0.993 | 0.220 |
| 14 | 1.836 | 1.251 | 0.330 | 41 | 0.969 | -1.743 | 0.250 |
| 15 | 1.308 | 1.356 | 0.190 | 42 | 1.321 | -0.024 | 0.200 |
| 16 | 0.786 | -0.992 | 0.290 | 43 | 1.079 | -0.495 | 0.240 |
| 17 | 1.483 | -1.120 | 0.200 | 44 | 1.172 | 1.079 | 0.110 |
| 18 | 1.656 | 3.000 | 0.180 | 45 | 1.191 | 0.079 | 0.260 |
| 19 | 0.892 | -0.578 | 0.230 | 46 | 1.225 | -0.090 | 0.150 |
| 20 | 1.744 | 3.000 | 0.080 | 47 | 0.431 | 0.119 | 0.370 |
| 21 | 2.119 | 1.019 | 0.200 | 48 | 0.505 | 1.130 | 0.290 |
| 22 | 1.723 | 3.000 | 0.150 | 49 | 0.769 | -1.052 | 0.260 |
| 23 | 0.400 | 0.121 | 0.370 | 50 | 1.100 | 0.980 | 0.100 |
| 24 | 1.737 | 3.000 | 0.050 | 51 | 1.494 | 1.636 | 0.340 |
| 25 | 1.342 | 1.640 | 0.320 | 52 | 0.841 | -0.393 | 0.190 |
| 26 | 1.616 | 3.000 | 0.200 | 53 | 0.813 | -0.619 | 0.280 |
| 27 | 1.998 | 0.449 | 0.260 | 54 | 0.545 | -1.103 | 0.270 |

| Item | a | b | c | Item | a | b | c |
|------|-------|--------|-------|------|-------|--------|-------|
| 55 | 1.639 | 0.921 | 0.180 | 84 | 0.781 | -1.445 | 0.240 |
| 56 | 0.532 | 1.428 | 0.370 | 85 | 1.080 | 1.289 | 0.180 |
| 57 | 1.395 | 0.263 | 0.220 | 86 | 1.900 | -0.410 | 0.170 |
| 58 | 2.225 | -0.311 | 0.150 | 87 | 0.784 | -1.339 | 0.260 |
| 59 | 1.671 | 2.565 | 0.100 | 88 | 1.852 | 1.557 | 0.240 |
| 60 | 1.375 | 1.536 | 0.080 | 89 | 1.727 | 3.000 | 0.200 |
| 61 | 1.094 | 0.612 | 0.350 | 90 | 1.575 | 0.906 | 0.280 |
| 62 | 1.054 | -0.582 | 0.200 | 91 | 1.624 | -1.072 | 0.200 |
| 63 | 1.758 | 0.768 | 0.310 | 92 | 0.764 | 0.194 | 0.260 |
| 64 | 1.239 | 1.327 | 0.300 | 93 | 1.929 | 0.437 | 0.250 |
| 65 | 1.421 | -0.083 | 0.240 | 94 | 1.523 | -0.092 | 0.200 |
| 66 | 0.876 | -1.377 | 0.250 | 95 | 1.921 | -0.749 | 0.150 |
| 67 | 1.172 | -1.386 | 0.230 | 96 | 1.535 | -1.740 | 0.180 |
| 68 | 1.634 | 0.228 | 0.360 | 97 | 1.503 | -0.529 | 0.150 |
| 69 | 1.207 | 0.392 | 0.250 | 98 | 1.186 | -1.482 | 0.230 |
| 70 | 0.725 | -1.142 | 0.270 | 99 | 0.649 | 0.530 | 0.190 |
| 71 | 0.561 | -0.286 | 0.300 | 100 | 1.781 | -0.050 | 0.220 |
| 72 | 1.334 | 1.600 | 0.150 | 101 | 1.492 | 3.000 | 0.360 |
| 73 | 1.030 | 1.461 | 0.160 | 102 | 1.220 | -0.031 | 0.160 |
| 74 | 0.923 | 1.605 | 0.220 | 103 | 0.452 | 1.025 | 0.370 |
| 75 | 1.420 | -0.920 | 0.210 | 104 | 1.693 | 3.000 | 0.100 |
| 76 | 1.688 | 3.000 | 0.240 | 105 | 1.227 | 1.020 | 0.340 |
| 77 | 1.695 | 3.000 | 0.230 | 106 | 1.079 | -0.883 | 0.190 |
| 78 | 1.203 | -0.472 | 0.290 | 107 | 0.719 | -3.000 | 0.270 |
| 79 | 0.845 | -0.270 | 0.260 | 108 | 1.804 | 3.000 | 0.100 |
| 80 | 2.014 | 0.582 | 0.190 | 109 | 0.636 | -3.000 | 0.20 |
| 81 | 0.881 | 0.246 | 0.240 | 110 | 0.548 | -2.059 | 0.260 |
| 82 | 1.307 | 2.072 | 0.320 | 111 | 0.465 | 0.146 | 0.320 |
| 83 | 1.562 | 1.599 | 0.340 | 112 | 1.841 | 1.438 | 0.320 |

| Item | a | b | c | Item | a | b | c |
|------|-------|--------|-------|------|-------|--------|-------|
| 113 | 1.508 | 2.375 | 0.250 | 142 | 1.691 | 0.152 | 0.150 |
| 114 | 0.497 | -1.865 | 0.260 | 143 | 2.025 | -1.596 | 0.210 |
| 115 | 1.631 | 0.210 | 0.350 | 144 | 0.701 | -1.350 | 0.240 |
| 116 | 1.287 | -1.049 | 0.29 | 145 | 1.850 | 0.302 | 0.130 |
| 117 | 1.391 | -0.823 | 0.150 | 146 | 1.171 | -2.406 | 0.240 |
| 118 | 0.727 | -1.083 | 0.260 | 147 | 2.021 | -0.237 | 0.150 |
| 119 | 1.008 | -1.628 | 0.260 | 148 | 0.814 | -1.247 | 0.250 |
| 120 | 0.828 | -0.017 | 0.220 | 149 | 1.691 | 1.705 | 0.130 |
| 121 | 1.314 | -1.467 | 0.250 | 150 | 1.591 | 0.607 | 0.230 |
| 122 | 2.086 | 0.556 | 0.280 | 151 | 1.623 | 1.809 | 0.320 |
| 123 | 1.255 | -1.883 | 0.240 | 152 | 1.128 | 2.129 | 0.280 |
| 124 | 1.107 | -1.428 | 0.240 | 153 | 0.706 | -1.813 | 0.250 |
| 125 | 0.787 | -1.133 | 0.260 | 154 | 0.985 | 0.185 | 0.310 |
| 126 | 1.113 | -2.051 | 0.270 | 155 | 0.723 | -1.269 | 0.240 |
| 127 | 0.949 | -1.077 | 0.270 | 156 | 0.497 | -3.000 | 0.270 |
| 128 | 1.323 | -0.238 | 0.320 | 157 | 1.043 | -0.088 | 0.290 |
| 129 | 1.955 | 1.208 | 0.370 | 158 | 1.858 | 0.261 | 0.320 |
| 130 | 0.754 | 1.599 | 0.170 | 159 | 1.519 | 3.000 | 0.200 |
| 131 | 0.561 | 0.988 | 0.240 | 160 | 0.887 | -1.259 | 0.200 |
| 132 | 0.642 | -0.986 | 0.260 | 161 | 1.441 | -0.939 | 0.270 |
| 133 | 0.610 | -2.076 | 0.260 | 162 | 1.050 | -0.451 | 0.110 |
| 134 | 0.569 | -1.451 | 0.280 | 163 | 1.063 | 1.230 | 0.210 |
| 135 | 0.476 | -0.514 | 0.330 | 164 | 0.519 | -1.453 | 0.280 |
| 136 | 0.815 | 0.240 | 0.290 | 165 | 1.283 | 1.209 | 0.190 |
| 137 | 1.542 | 3.000 | 0.260 | 166 | 1.625 | 1.786 | 0.240 |
| 138 | 1.842 | -0.239 | 0.100 | 167 | 0.853 | -1.161 | 0.250 |
| 139 | 0.792 | -0.369 | 0.250 | 168 | 0.947 | 0.086 | 0.350 |
| 140 | 1.528 | 0.223 | 0.130 | 169 | 0.942 | -0.425 | 0.140 |
| 141 | 0.991 | 2.135 | 0.160 | 170 | 0.701 | 0.175 | 0.330 |

| Item | a | b | c | Item | a | b | c |
|------|-------|--------|-------|------|-------|--------|-------|
| 171 | 0.636 | 0.246 | 0.130 | 201 | 1.064 | -0.031 | 0.130 |
| 172 | 0.630 | 1.263 | 0.240 | 202 | 0.939 | -0.136 | 0.180 |
| 173 | 1.270 | 1.420 | 0.230 | 203 | 1.080 | -0.028 | 0.170 |
| 174 | 0.602 | -0.145 | 0.310 | 204 | 1.155 | 0.056 | 0.330 |
| 175 | 1.680 | 1.233 | 0.340 | 205 | 1.890 | 1.688 | 0.370 |
| 176 | 1.608 | 1.268 | 0.220 | 206 | 1.096 | -0.983 | 0.220 |
| 177 | 0.947 | 0.688 | 0.310 | 207 | 1.934 | 0.602 | 0.260 |
| 178 | 1.598 | 2.436 | 0.230 | 208 | 1.364 | 1.804 | 0.310 |
| 179 | 1.702 | 1.029 | 0.340 | 209 | 0.640 | 0.462 | 0.360 |
| 180 | 2.007 | 0.605 | 0.350 | 210 | 1.019 | -0.007 | 0.210 |
| 181 | 0.670 | -1.180 | 0.270 | 211 | 0.593 | -2.175 | 0.250 |
| 182 | 0.447 | 0.906 | 0.350 | 212 | 1.761 | 1.101 | 0.370 |
| 183 | 1.725 | 3.000 | 0.110 | 213 | 1.779 | 3.000 | 0.190 |
| 184 | 0.925 | -0.721 | 0.230 | 214 | 1.776 | 3.000 | 0.190 |
| 185 | 1.488 | 3.000 | 0.320 | 215 | 1.500 | -0.031 | 0.340 |
| 186 | 1.755 | 3.000 | 0.090 | 216 | 1.469 | 3.000 | 0.280 |
| 187 | 1.790 | 3.000 | 0.240 | 217 | 0.649 | 1.318 | 0.240 |
| 188 | 1.689 | 3.000 | 0.150 | 218 | 1.583 | 1.779 | 0.20 |
| 190 | 1.814 | 3.000 | 0.020 | 219 | 1.783 | 3.000 | 0.200 |
| 191 | 0.720 | 1.287 | 0.350 | 220 | 1.872 | 3.000 | 0.130 |
| 192 | 2.500 | 3.000 | 0.240 | 221 | 2.198 | -0.857 | 0.070 |
| 193 | 0.782 | -0.777 | 0.200 | 222 | 1.769 | 3.000 | 0.210 |
| 194 | 1.018 | 1.553 | 0.360 | 223 | 1.565 | 1.790 | 0.330 |
| 195 | 1.966 | 1.085 | 0.250 | 224 | 1.812 | 3.000 | 0.260 |
| 196 | 1.189 | 0.858 | 0.290 | 225 | 1.708 | 1.890 | 0.320 |
| 197 | 1.192 | 0.070 | 0.210 | 226 | 1.478 | -0.395 | 0.170 |
| 198 | 1.764 | -0.088 | 0.180 | 227 | 2.275 | -0.053 | 0.180 |
| 199 | 1.228 | 0.095 | 0.250 | 228 | 1.028 | -0.188 | 0.150 |
| 200 | 1.291 | 0.466 | 0.170 | 229 | 1.463 | 0.279 | 0.230 |

การประมาณค่าความสามารถของผู้สอบในการทดสอบด้วยแบบสอบวัดความสามารถทางภาษาอังกฤษ
และการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ จำนวน 5 ครั้ง

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 1 | 10 | ต่ำ | -1.117 | -1.075 | -1.583 | -1.466 | -1.475 | -1.589 | -1.4376 |
| 2 | 10 | ต่ำ | -0.79 | -1.266 | -1.283 | -1.588 | -1.068 | -1.301 | -1.3012 |
| 3 | 10 | ต่ำ | -1.547 | -2.081 | -1.578 | -2.363 | -2.024 | -1.598 | -1.9288 |
| 4 | 10 | ต่ำ | -1.125 | -2.061 | -2.297 | -2.026 | -1.33 | -1.109 | -1.7646 |
| 5 | 10 | ต่ำ | -1.202 | -1.914 | -2.275 | -2.001 | -1.772 | -0.606 | -1.7136 |
| 6 | 10 | ต่ำ | -1.427 | -1.914 | -1.579 | -1.332 | -2.309 | -1.622 | -1.7512 |
| 7 | 10 | ต่ำ | -1.427 | -2.088 | -2.46 | -1.912 | -1.912 | -2.224 | -2.1192 |
| 8 | 10 | ต่ำ | -0.79 | -2.032 | -1.454 | -0.785 | -1.215 | -0.654 | -1.228 |
| 9 | 10 | ต่ำ | -1.626 | -2.192 | -2.416 | -2.391 | -1.586 | -2.469 | -2.2108 |
| 10 | 10 | ต่ำ | -1.465 | -1.076 | -0.894 | -2.625 | -1.418 | -1.638 | -1.5302 |
| 11 | 10 | ต่ำ | -1.427 | -1.861 | -1.609 | -2.141 | -1.639 | -0.653 | -1.5806 |
| 12 | 10 | ต่ำ | -1.427 | -0.741 | -0.792 | -1.412 | -1.504 | -2.283 | -1.3464 |
| 13 | 10 | ต่ำ | -1.246 | -1.797 | -1.72 | -2.197 | -1.904 | -1.905 | -1.9046 |
| 14 | 10 | ต่ำ | -1.547 | -1.508 | -2.597 | -1.675 | -1.927 | -1.925 | -1.9264 |
| 15 | 10 | ต่ำ | -1.125 | -1.975 | -0.995 | -2.451 | -2.247 | -1.181 | -1.7698 |
| 16 | 10 | ต่ำ | -1.746 | -1.979 | -1.692 | -2.326 | -1.91 | -2.37 | -2.0554 |
| 17 | 10 | ต่ำ | -1.246 | -0.759 | -2.561 | -1.085 | -2.34 | -1.686 | -1.6862 |
| 18 | 10 | ต่ำ | -1.465 | -2.212 | -2.1 | -2.165 | -0.94 | -1.854 | -1.8542 |
| 19 | 10 | ต่ำ | -1.426 | -1.035 | -1.949 | -2.725 | -1.904 | -1.649 | -1.8524 |
| 20 | 10 | ต่ำ | -1.836 | -1.861 | -2.272 | -2.055 | -2.028 | -2.097 | -2.0626 |
| 21 | 10 | ต่ำ | -1.744 | -2.192 | -2.005 | -2.235 | -1.592 | -2.143 | -2.0334 |
| 22 | 10 | ต่ำ | -1.697 | -1.861 | -1.91 | -1.888 | -2.045 | -2.071 | -1.955 |
| 23 | 10 | ต่ำ | -1.626 | -2.135 | -2.605 | -1.425 | -2.618 | -1.633 | -2.0832 |
| 24 | 10 | ต่ำ | -1.836 | -1.648 | -2.496 | -2.157 | -2.476 | -2.071 | -2.1696 |
| 25 | 10 | ปานกลาง | -0.732 | -1.607 | -0.418 | -1.712 | -2.06 | -1.803 | -1.52 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 26 | 10 | ปานกลาง | -0.72 | -1.456 | -1.939 | -2.157 | -2.026 | -1.632 | -1.842 |
| 27 | 10 | ปานกลาง | -0.706 | -1.936 | -1.962 | -2.414 | -1.505 | -1.715 | -1.9064 |
| 28 | 10 | ปานกลาง | -0.68 | -1.29 | -1.114 | -2.759 | -1.07 | -1.907 | -1.628 |
| 29 | 10 | ปานกลาง | 0.24 | -0.358 | 0.529 | -1.349 | -2.164 | -0.836 | -0.8356 |
| 30 | 10 | ปานกลาง | -0.642 | -1.26 | -1.352 | -0.931 | -1.326 | -2.091 | -1.392 |
| 31 | 10 | ปานกลาง | -0.613 | -2.081 | -1.305 | -2.183 | -1.902 | -1.848 | -1.8638 |
| 32 | 10 | ปานกลาง | -0.61 | -1.914 | -2.454 | -2.443 | -0.406 | -0.821 | -1.6076 |
| 33 | 10 | ปานกลาง | -0.585 | -1.914 | -1.328 | -1.631 | -1.43 | -2.11 | -1.6826 |
| 34 | 10 | ปานกลาง | -0.582 | -0.835 | -2.256 | -1.542 | -1.417 | -1.678 | -1.5456 |
| 35 | 10 | ปานกลาง | -0.555 | -1.709 | -2.07 | -1.7 | -1.839 | -1.83 | -1.8296 |
| 36 | 10 | ปานกลาง | -0.539 | -1.158 | -1.456 | -0.883 | -1.358 | -1.213 | -1.2136 |
| 37 | 10 | ปานกลาง | -0.486 | -1.618 | -0.6 | -1.734 | -1.033 | -1.827 | -1.3624 |
| 38 | 10 | ปานกลาง | -0.478 | -1.861 | -0.174 | -2.073 | -2.152 | -2.261 | -1.7042 |
| 39 | 10 | ปานกลาง | -0.475 | -1.914 | -2.52 | -0.421 | -1.861 | -1.145 | -1.5722 |
| 40 | 10 | ปานกลาง | -0.467 | -1.425 | -1.084 | -2.73 | -1.891 | -0.753 | -1.5766 |
| 41 | 10 | ปานกลาง | -0.437 | -2.081 | -0.911 | -2.043 | -1.678 | -1.678 | -1.6782 |
| 42 | 10 | ปานกลาง | -0.03 | -0.901 | -1.61 | -0.829 | -1.296 | -0.253 | -0.9778 |
| 43 | 10 | ปานกลาง | -0.386 | -0.687 | -1.649 | -0.882 | -1.664 | -1.221 | -1.2206 |
| 44 | 10 | ปานกลาง | -0.323 | -1.554 | -1.159 | -1.435 | 0.02 | -0.807 | -0.987 |
| 45 | 10 | ปานกลาง | -0.299 | -1.546 | -1.964 | -1.534 | -2.179 | -2.563 | -1.9572 |
| 46 | 10 | ปานกลาง | -0.267 | -1.554 | -1.613 | -2.047 | -1.124 | -1.878 | -1.6432 |
| 47 | 10 | ปานกลาง | -0.262 | -1.626 | -0.911 | -1.155 | -2.481 | -1.491 | -1.5328 |
| 48 | 10 | ปานกลาง | -0.732 | -1.577 | -1.915 | -2.533 | -1.941 | -2.17 | -2.0272 |
| 49 | 10 | ปานกลาง | -0.2 | -0.889 | -0.72 | -1.335 | -0.754 | -0.924 | -0.9244 |
| 50 | 10 | ปานกลาง | -0.162 | -0.948 | -1.275 | -2.199 | 0.159 | -1.066 | -1.0658 |
| 51 | 10 | ปานกลาง | -0.706 | -1.914 | -2.143 | -2.302 | -1.641 | -1.958 | -1.9916 |
| 52 | 10 | ปานกลาง | -0.127 | -1.618 | -1.607 | -2.239 | -1.289 | -1.153 | -1.5812 |
| 53 | 10 | ปานกลาง | -0.732 | -2.265 | -2.1 | -1.657 | -2.353 | -1.848 | -2.0446 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 54 | 10 | ปานกลาง | -0.106 | -0.873 | -0.9 | -0.09 | -0.876 | -1.482 | -0.8442 |
| 55 | 10 | ปานกลาง | -0.041 | -0.759 | -1.336 | -1.302 | -1.063 | 0.178 | -0.8564 |
| 56 | 10 | ปานกลาง | -0.03 | -0.751 | -1.648 | -1.683 | -1.099 | -0.022 | -1.0406 |
| 57 | 10 | ปานกลาง | 0.011 | -0.722 | -1.201 | -0.71 | -0.598 | -2.103 | -1.0668 |
| 58 | 10 | ปานกลาง | -0.72 | -1.366 | -1.284 | -2.23 | -2.332 | -2.228 | -1.888 |
| 59 | 10 | ปานกลาง | 0.084 | -1.056 | -1.811 | -1.565 | 0.107 | -0.367 | -0.9384 |
| 60 | 10 | ปานกลาง | 0.084 | -1.556 | 0.124 | -2.208 | -1.637 | -2.232 | -1.5018 |
| 61 | 10 | ปานกลาง | 0.24 | -0.799 | -1.386 | -0.493 | -0.31 | -0.213 | -0.6402 |
| 62 | 10 | ปานกลาง | 0.338 | -0.282 | -0.254 | -0.846 | -0.62 | -0.501 | -0.5006 |
| 63 | 10 | ปานกลาง | 0.43 | -0.096 | 0.43 | 0.133 | 0.446 | -0.08 | 0.1666 |
| 64 | 10 | ปานกลาง | -0.386 | -0.241 | -1.554 | -1.78 | -1.382 | -1.925 | -1.3764 |
| 65 | 10 | สูง | 1.515 | -0.323 | -0.527 | 0.128 | 0.142 | 0.097 | -0.0966 |
| 66 | 10 | สูง | 0.866 | -0.136 | -0.895 | -0.608 | -0.229 | -1.773 | -0.7282 |
| 67 | 10 | สูง | 0.873 | -0.092 | -0.857 | 0.261 | -0.456 | -0.323 | -0.2934 |
| 68 | 10 | สูง | 0.908 | 0.127 | -0.222 | 0.127 | 0.011 | 0.0107 | 0.0107 |
| 69 | 10 | สูง | 0.962 | -0.527 | -0.958 | -0.293 | -0.482 | -0.106 | -0.4732 |
| 70 | 10 | สูง | 0.985 | -1.33 | -0.399 | -0.294 | -0.231 | -0.07 | -0.4648 |
| 71 | 10 | สูง | 1.066 | -0.156 | -0.536 | -0.38 | -0.852 | -0.504 | -0.4856 |
| 72 | 10 | สูง | 1.068 | -0.164 | 0.084 | 0.119 | -0.025 | 0.045 | 0.0118 |
| 73 | 10 | สูง | 1.095 | 0.023 | -0.222 | 0.1 | -0.034 | -0.68 | -0.1626 |
| 74 | 10 | สูง | 0.985 | 0.044 | -0.868 | -0.731 | -1.279 | -0.855 | -0.7378 |
| 75 | 10 | สูง | 1.137 | 0.044 | -0.307 | -0.257 | -0.326 | -0.211 | -0.2114 |
| 76 | 10 | สูง | 1.161 | -0.111 | -0.364 | 0.341 | -0.067 | -0.457 | -0.1316 |
| 77 | 10 | สูง | 0.962 | -2.122 | -0.806 | -0.616 | -0.537 | -0.012 | -0.8186 |
| 78 | 10 | สูง | 1.243 | 0.05 | -0.265 | -0.417 | 0.041 | -0.168 | -0.1518 |
| 79 | 10 | สูง | 1.381 | 0.163 | -0.477 | -0.686 | 0.094 | -0.289 | -0.239 |
| 80 | 10 | สูง | 1.45 | -0.087 | -0.906 | -0.546 | -0.485 | -0.483 | -0.5014 |
| 81 | 10 | สูง | 0.908 | -0.131 | -0.956 | -0.486 | -1.028 | -0.836 | -0.6874 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 82 | 10 | สูง | 1.517 | 0.184 | -0.599 | 0.264 | -0.26 | -0.103 | -0.1028 |
| 83 | 10 | สูง | 1.518 | 0.127 | 0.016 | -0.003 | -0.003 | 0.388 | 0.105 |
| 84 | 10 | สูง | 1.539 | 0.024 | -0.425 | -0.242 | 0.01 | -0.442 | -0.215 |
| 85 | 10 | สูง | 1.553 | -0.545 | -0.013 | -0.013 | -0.45 | 0.214 | -0.1614 |
| 86 | 10 | สูง | 1.584 | -0.03 | 0.091 | -0.037 | -0.41 | -0.642 | -0.2056 |
| 87 | 10 | สูง | 1.717 | 0.034 | 0.163 | -0.162 | 0.191 | -0.091 | 0.027 |
| 88 | 10 | สูง | 1.737 | -0.155 | 0.461 | 0.222 | 0.461 | 0.393 | 0.2764 |
| 89 | 10 | สูง | 1.819 | 0.309 | -0.362 | -0.474 | -0.167 | -0.174 | -0.1736 |
| 90 | 10 | สูง | 1.848 | 0.196 | 0.961 | 0.102 | 0.349 | 0.102 | 0.342 |
| 91 | 10 | สูง | 1.904 | -0.116 | -0.236 | 0.218 | 0.659 | 0.131 | 0.1312 |
| 92 | 10 | สูง | 1.911 | 0.122 | -0.092 | 0.337 | -0.469 | 0.309 | 0.0414 |
| 93 | 10 | สูง | 1.919 | 0.051 | 0.552 | -0.438 | 0.441 | -0.042 | 0.1128 |
| 94 | 10 | สูง | 2.029 | -0.694 | 0.353 | -0.093 | 0.614 | 0.152 | 0.0664 |
| 95 | 10 | สูง | 2.038 | -0.792 | -0.132 | 0.355 | 0.144 | 0.291 | -0.0268 |
| 96 | 10 | สูง | 2.081 | 0.236 | 1.001 | 0.672 | 0.843 | 0.073 | 0.565 |
| 97 | 10 | สูง | 2.161 | 0.477 | 0.902 | 0.528 | 0.814 | 0.589 | 0.662 |
| 98 | 10 | สูง | 2.288 | 0.178 | 0.448 | 0.252 | 0.124 | -0.056 | 0.1892 |
| 99 | 10 | สูง | 1.186 | 0.006 | -0.163 | -0.335 | -0.221 | -0.485 | -0.2396 |
| 100 | 10 | สูง | 2.374 | 0.477 | 0.895 | 0.389 | 0.907 | 0.57 | 0.6476 |
| 101 | 10 | สูง | 2.386 | 0.073 | -0.055 | -0.153 | 0.452 | 0.187 | 0.1008 |
| 102 | 10 | สูง | 2.468 | 0.271 | 0.247 | 0.471 | 0.411 | -0.282 | 0.2394 |
| 103 | 10 | สูง | 1.481 | 0.353 | -0.967 | -0.155 | 0.134 | -0.155 | -0.158 |
| 104 | 10 | สูง | 3.135 | 0.433 | 0.727 | 0.795 | 0.752 | 0.378 | 0.617 |
| 105 | 20 | ต่ำ | -1.336 | -1.975 | -2.404 | -1.712 | -2.569 | -2.167 | -2.1654 |
| 106 | 20 | ต่ำ | -0.915 | -1.638 | -2.051 | -1.594 | -1.425 | -2.172 | -1.776 |
| 107 | 20 | ต่ำ | -1.697 | -1.407 | -2.365 | -1.83 | -2.467 | -2.165 | -2.0468 |
| 108 | 20 | ต่ำ | -0.925 | -1.112 | -1.164 | -1.245 | -1.399 | -1.586 | -1.3012 |
| 109 | 20 | ต่ำ | -1.11 | -1.914 | -1.586 | -1.628 | -1.574 | -1.673 | -1.675 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 110 | 20 | ต่ำ | -1.836 | -1.881 | -2.353 | -1.555 | -2.251 | -2.028 | -2.0136 |
| 111 | 20 | ต่ำ | -1.697 | -2.265 | -0.66 | -2.072 | -2.36 | -2.22 | -1.9154 |
| 112 | 20 | ต่ำ | -1.619 | -2.179 | -1.973 | -1.456 | -2.149 | -1.588 | -1.869 |
| 113 | 20 | ต่ำ | -1.456 | -1.668 | -2.335 | -1.2 | -0.669 | -0.9 | -1.3544 |
| 114 | 20 | ต่ำ | -1.409 | -2.088 | -2.27 | -1.076 | -1.343 | -0.84 | -1.5234 |
| 115 | 20 | ต่ำ | -1.336 | -1.755 | -1.627 | -1.94 | -1.921 | -1.811 | -1.8108 |
| 116 | 20 | ต่ำ | -1.184 | -1.621 | -1.163 | -1.913 | -1.566 | -1.565 | -1.5656 |
| 117 | 20 | ต่ำ | -1.744 | -1.979 | -1.866 | -2.2 | -2.102 | -1.854 | -2.0002 |
| 118 | 20 | ต่ำ | -1.088 | -1.709 | -1.227 | -2.218 | -1.717 | -1.751 | -1.7244 |
| 119 | 20 | ต่ำ | -1.056 | -0.848 | -1.913 | -1.752 | -1.794 | -1.577 | -1.5768 |
| 120 | 20 | ต่ำ | -1.051 | -1.499 | -1.769 | -1.643 | -1.992 | -2.234 | -1.8274 |
| 121 | 20 | ต่ำ | -1.079 | -2.081 | -0.917 | -2.081 | -1.51 | -1.842 | -1.6862 |
| 122 | 20 | ต่ำ | -1.11 | -1.914 | -1.835 | -0.991 | -2.277 | -2.016 | -1.8066 |
| 123 | 20 | ต่ำ | -1.291 | -1.591 | -1.723 | -2.265 | -1.966 | -1.909 | -1.8908 |
| 124 | 20 | ต่ำ | -1.336 | -2.013 | -2.143 | -1.986 | -1.784 | -1.982 | -1.9815 |
| 125 | 20 | ต่ำ | -0.915 | -1.525 | -0.838 | -2.335 | -0.49 | -1.49 | -1.3356 |
| 126 | 20 | ต่ำ | -1.091 | -1.26 | -2.246 | -1.349 | -1.461 | -1.012 | -1.4656 |
| 127 | 20 | ต่ำ | -0.865 | -0.824 | -1.557 | -1.694 | -1.065 | -0.722 | -1.1724 |
| 128 | 20 | ต่ำ | -0.82 | -1.943 | -1.466 | -1.876 | -2.047 | -1.833 | -1.833 |
| 129 | 20 | ต่ำ | -0.793 | -0.865 | -2.192 | -1.586 | -1.965 | -1.663 | -1.6542 |
| 130 | 20 | ต่ำ | -0.789 | -1.338 | -0.453 | -0.302 | -0.657 | -0.832 | -0.7164 |
| 131 | 20 | ต่ำ | -1.547 | -1.888 | -2.434 | -1.527 | -1.961 | -1.663 | -1.8946 |
| 132 | 20 | ปานกลาง | -0.747 | -1.814 | -2.068 | -1.642 | -1.914 | -1.859 | -1.8594 |
| 133 | 20 | ปานกลาง | -0.716 | -1.844 | -1.68 | -1.336 | -1.461 | -1.949 | -1.654 |
| 134 | 20 | ปานกลาง | -0.689 | -1.499 | -1.528 | -1.935 | -2.061 | -2.342 | -1.873 |
| 135 | 20 | ปานกลาง | -0.684 | -1.662 | -1.092 | -1.692 | -1.486 | -1.483 | -1.483 |
| 136 | 20 | ปานกลาง | -0.673 | -1.425 | -1.948 | -1.453 | -1.645 | -1.618 | -1.6177 |
| 137 | 20 | ปานกลาง | -0.67 | -1.783 | -1.323 | -1.653 | -1.593 | -1.823 | -1.635 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 138 | 20 | ปานกลาง | -0.623 | -2.068 | -1.762 | -1.386 | -2.422 | -1.91 | -1.9096 |
| 139 | 20 | ปานกลาง | -0.61 | -1.885 | -2.107 | -0.972 | -1.996 | -1.289 | -1.6498 |
| 140 | 20 | ปานกลาง | -0.575 | -0.979 | -0.095 | -0.734 | -0.346 | -1.209 | -0.6726 |
| 141 | 20 | ปานกลาง | -0.575 | -2.018 | -1.708 | -1.732 | -1.563 | -1.957 | -1.7956 |
| 142 | 20 | ปานกลาง | -0.557 | -1.803 | -1.703 | -1.598 | -1.679 | -1.563 | -1.6692 |
| 143 | 20 | ปานกลาง | -0.546 | -1.913 | -0.848 | -1.687 | -1.62 | -1.682 | -1.55 |
| 144 | 20 | ปานกลาง | -0.517 | -0.605 | -0.399 | -2.166 | -1.644 | -2.148 | -1.3924 |
| 145 | 20 | ปานกลาง | -0.747 | -1.456 | -1.798 | -2.667 | -1.726 | -2.36 | -2.0014 |
| 146 | 20 | ปานกลาง | -0.483 | -1.004 | -2.306 | -1.447 | -2.01 | -1.915 | -1.7364 |
| 147 | 20 | ปานกลาง | -0.452 | -1.75 | -1.187 | -1.663 | -1.874 | -1.512 | -1.5972 |
| 148 | 20 | ปานกลาง | -0.451 | -1.253 | -1.835 | -1.035 | -2.122 | -1.845 | -1.618 |
| 149 | 20 | ปานกลาง | -0.426 | -1.607 | -0.905 | -1.425 | -1.608 | -1.339 | -1.3768 |
| 150 | 20 | ปานกลาง | -0.382 | -2.081 | -1.881 | -1.997 | -0.937 | -0.921 | -1.3314 |
| 151 | 20 | ปานกลาง | -0.365 | -1.16 | -2.636 | -1.303 | -0.546 | -1.411 | -1.4112 |
| 152 | 20 | ปานกลาง | -0.325 | -2.192 | -1.456 | -1.632 | -1.425 | -1.676 | -1.4922 |
| 153 | 20 | ปานกลาง | -0.313 | -0.167 | -0.522 | -0.379 | -0.026 | -0.114 | -0.2416 |
| 154 | 20 | ปานกลาง | -0.313 | -0.444 | -1.344 | -0.155 | -0.75 | -0.673 | -0.6732 |
| 155 | 20 | ปานกลาง | -0.313 | -1.093 | -1.102 | -1.1 | -0.631 | -0.095 | -0.8042 |
| 156 | 20 | ปานกลาง | -0.575 | -1.518 | -1.786 | -1.665 | -1.681 | -1.662 | -1.6624 |
| 157 | 20 | ปานกลาง | -0.296 | -0.759 | -1.296 | -1.505 | -0.631 | -0.878 | -1.0138 |
| 158 | 20 | ปานกลาง | -0.29 | -0.961 | -0.882 | -0.826 | -1.211 | -1.112 | -0.9984 |
| 159 | 20 | ปานกลาง | -0.262 | -0.605 | -0.873 | -0.521 | -1.894 | -0.961 | -0.9708 |
| 160 | 20 | ปานกลาง | -0.239 | -0.796 | -1.619 | -0.968 | -1.648 | -0.535 | -1.1132 |
| 161 | 20 | ปานกลาง | -0.126 | -0.073 | -1.09 | -1.275 | -1.525 | -1.379 | -1.0684 |
| 162 | 20 | ปานกลาง | -0.051 | -0.543 | -1.445 | -1.167 | -0.333 | -0.97 | -0.8916 |
| 163 | 20 | ปานกลาง | -0.119 | -1.004 | -0.457 | -0.471 | -0.882 | -0.367 | -0.5268 |
| 164 | 20 | ปานกลาง | 0.136 | -1.194 | -0.729 | -0.186 | -0.842 | -0.604 | -0.5902 |
| 165 | 20 | ปานกลาง | -0.313 | -0.167 | -1.82 | -1.627 | -2.097 | -1.114 | -1.365 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 166 | 20 | ปานกลาง | 0.173 | 0.058 | -0.82 | -1.704 | 0.034 | -0.339 | -0.5542 |
| 167 | 20 | ปานกลาง | 0.344 | -0.146 | -0.67 | -1.062 | -0.443 | -1.017 | -0.6676 |
| 168 | 20 | ปานกลาง | 0.401 | -0.582 | -0.509 | 0.127 | 0.049 | -0.228 | -0.2286 |
| 169 | 20 | ปานกลาง | 0.475 | -0.301 | -0.449 | -0.377 | 0.265 | 0.365 | -0.0994 |
| 170 | 20 | ปานกลาง | 0.637 | -0.155 | -0.896 | -1.276 | -1.283 | -1.628 | -1.0476 |
| 171 | 20 | ปานกลาง | -0.64 | -1.954 | -1.251 | -1.532 | -1.464 | -1.536 | -1.5474 |
| 172 | 20 | สูง | 0.758 | -1.08 | -0.193 | 0.204 | -0.469 | 0.362 | -0.2352 |
| 173 | 20 | สูง | 0.779 | -0.204 | -0.811 | 0.089 | 0.326 | 0.166 | -0.0868 |
| 174 | 20 | สูง | 0.952 | 0.098 | 0.589 | -0.176 | 0.317 | -0.028 | 0.16 |
| 175 | 20 | สูง | 1.015 | -0.155 | 0.026 | -0.155 | -0.115 | -0.048 | -0.0894 |
| 176 | 20 | สูง | 1.031 | -1.19 | -0.741 | 0.166 | 0.098 | 0.166 | -0.3002 |
| 177 | 20 | สูง | 1.07 | 0.206 | -0.023 | 0.864 | 0.63 | 0.323 | 0.4 |
| 178 | 20 | สูง | 1.079 | 0.016 | 0.8 | 0.56 | 0.725 | 0.089 | 0.438 |
| 179 | 20 | สูง | 1.111 | -0.1 | 0.66 | 0.59 | -0.051 | 0.274 | 0.2746 |
| 180 | 20 | สูง | 1.234 | -0.342 | 0.285 | 0.3 | 0.499 | 0.252 | 0.1988 |
| 181 | 20 | สูง | 1.245 | -1.081 | -0.204 | -0.359 | -0.206 | -0.004 | -0.3708 |
| 182 | 20 | สูง | 1.307 | -0.053 | 0.508 | 1.07 | 1.027 | 0.864 | 0.6832 |
| 183 | 20 | สูง | 1.324 | -0.326 | -0.005 | -0.307 | 0.786 | -0.128 | 0.004 |
| 184 | 20 | สูง | 1.347 | -0.069 | -0.392 | -0.386 | -0.925 | -0.069 | -0.499 |
| 185 | 20 | สูง | 1.362 | -0.713 | -0.06 | 0.093 | -0.812 | 0.183 | -0.2618 |
| 186 | 20 | สูง | 1.466 | 0.019 | -0.549 | -0.204 | 0.173 | 0.571 | 0.002 |
| 187 | 20 | สูง | 1.641 | 0.43 | 0.064 | 0.064 | 0.167 | -0.093 | 0.1264 |
| 188 | 20 | สูง | 1.716 | 0.039 | -0.18 | 0.06 | -0.46 | 0.064 | -0.0954 |
| 189 | 20 | สูง | 1.74 | 0.178 | -0.172 | -0.194 | 0 | 0.045 | -0.0286 |
| 190 | 20 | สูง | 1.75 | -0.525 | 0.381 | 0.127 | 0.201 | 0.32 | 0.1008 |
| 191 | 20 | สูง | 1.804 | -0.236 | 0.364 | -0.402 | -0.305 | -0.801 | -0.276 |
| 192 | 20 | สูง | 1.81 | 0.154 | 0.047 | 0.062 | 0.278 | 0.1352 | 0.1352 |
| 193 | 20 | สูง | 1.848 | -0.006 | 0.228 | 0.153 | 0.071 | 0.1115 | 0.1115 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 194 | 20 | สูง | 1.851 | 0.527 | 0.868 | 0.218 | 0.676 | 0.709 | 0.5996 |
| 195 | 20 | สูง | 1.866 | 0.265 | 0.659 | -0.313 | -0.081 | 0.313 | 0.1686 |
| 196 | 20 | สูง | 1.893 | 0.22 | 0.965 | -0.24 | 0.308 | 0.893 | 0.4292 |
| 197 | 20 | สูง | 1.916 | 0.484 | 0.748 | 0.093 | 1.068 | 0.03 | 0.4846 |
| 198 | 20 | สูง | 1.976 | 1.135 | 0.505 | 0.565 | 0.977 | 0.634 | 0.7632 |
| 199 | 20 | สูง | 2.048 | 0.407 | -0.001 | 0.509 | 1.437 | 0.48 | 0.5664 |
| 200 | 20 | สูง | 2.075 | 0.585 | 0.744 | 0.634 | 0.677 | 0.686 | 0.6652 |
| 201 | 20 | สูง | 0.758 | -1.611 | -1.476 | -0.541 | -0.437 | 0.073 | -0.7984 |
| 202 | 20 | สูง | 2.124 | 0.089 | 0.886 | 0.493 | 0.926 | 1.03 | 0.6842 |
| 203 | 20 | สูง | 2.312 | 0.802 | 0.358 | 1.037 | 0.38 | 0.792 | 0.6738 |
| 204 | 20 | สูง | 2.717 | 1.018 | 0.199 | 0.725 | 0.59 | 0.725 | 0.6514 |
| 205 | 20 | สูง | 3.233 | 1.279 | 0.249 | 0.826 | 0.659 | 0.753 | 0.7532 |
| 206 | 20 | สูง | 2.501 | 0.648 | 0.156 | -0.088 | -0.209 | 0.049 | 0.1112 |
| 207 | 20 | สูง | 2.501 | 0.648 | -0.156 | -0.088 | -0.209 | 0.049 | 0.0488 |
| 208 | 20 | สูง | 0.758 | -1.073 | 0.642 | -0.457 | -0.315 | -0.199 | -0.2804 |
| 209 | 20 | สูง | 0.779 | 0.429 | -0.648 | -1.669 | -1.139 | 1.239 | -0.3576 |
| 210 | 20 | สูง | 0.779 | 0.339 | -1.203 | -1.449 | -1.038 | -0.896 | -0.8494 |
| 211 | 20 | สูง | 1.44 | 0.048 | -0.527 | -0.65 | 1.035 | -0.297 | -0.0782 |
| 212 | 30 | ต่ำ | -0.805 | -1.031 | -0.398 | -0.303 | -0.108 | -0.403 | -0.4486 |
| 213 | 30 | ต่ำ | -0.96 | -2.25 | -2.07 | -2.461 | -1.6 | -0.844 | -1.8456 |
| 214 | 30 | ต่ำ | -1.304 | -1.974 | -1.982 | -1.433 | -0.806 | -1.481 | -1.5352 |
| 215 | 30 | ต่ำ | -1.447 | -1.925 | -1.886 | -1.554 | -1.9 | -1.474 | -1.748 |
| 216 | 30 | ต่ำ | -1.367 | -2.352 | -2.239 | -1.801 | -2.18 | -1.449 | -2.0042 |
| 217 | 30 | ต่ำ | -1.746 | -1.88 | -2.252 | -2.254 | -2.1 | -2.16 | -2.13 |
| 218 | 30 | ต่ำ | -1.312 | -1.685 | -2.37 | -1.835 | -2.05 | -1.775 | -1.9434 |
| 219 | 30 | ต่ำ | -1.744 | -2.321 | -1.765 | -1.911 | -2.1 | -1.863 | -1.991 |
| 220 | 30 | ต่ำ | -1.645 | -1.677 | -1.661 | -1.744 | -1.78 | -1.749 | -1.7226 |
| 221 | 30 | ต่ำ | -1.547 | -1.945 | -1.823 | -1.975 | -2.13 | -1.794 | -1.9332 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 222 | 30 | ต่ำ | -1.414 | -2.183 | -2.134 | -2.379 | -1.73 | -0.717 | -1.8278 |
| 223 | 30 | ต่ำ | -1.39 | -2.026 | -2.021 | -1.959 | -0.807 | -1.109 | -1.5844 |
| 224 | 30 | ต่ำ | -1.319 | -2.198 | -2.507 | -1.202 | -0.117 | -0.087 | -1.2222 |
| 225 | 30 | ต่ำ | -1.312 | -2.441 | -1.494 | -2.169 | -2.25 | -1.991 | -2.0698 |
| 226 | 30 | ต่ำ | -1.31 | -2.062 | -0.323 | -1.681 | -1.64 | -1.802 | -1.5022 |
| 227 | 30 | ต่ำ | -1.29 | -2.13 | -1.393 | -0.953 | -1.92 | -1.803 | -1.639 |
| 228 | 30 | ต่ำ | -1.278 | -2.073 | -2.252 | -2.005 | -2.18 | -1.974 | -2.0976 |
| 229 | 30 | ต่ำ | -1.255 | -2.347 | -1.579 | -2.275 | -1.86 | -2.101 | -2.0316 |
| 230 | 30 | ต่ำ | -1.252 | -2.036 | -2.193 | -2.054 | -2.24 | -1.836 | -2.0712 |
| 231 | 30 | ต่ำ | -1.239 | -0.84 | -1.964 | -1.549 | -1.79 | -1.527 | -1.5346 |
| 232 | 30 | ต่ำ | -1.235 | -2.421 | -2.215 | -2.109 | -1.75 | -1.126 | -1.924 |
| 233 | 30 | ต่ำ | -1.232 | -2.032 | -1.979 | -2.251 | -2.16 | -1.838 | -2.0512 |
| 234 | 30 | ต่ำ | -1.244 | -1.547 | -1.337 | -1.587 | -1.33 | -1.438 | -1.4468 |
| 235 | 30 | ต่ำ | -1.2 | -1.715 | -2.05 | -1.951 | -1.88 | -1.56 | -1.8304 |
| 236 | 30 | ต่ำ | -1.16 | -1.429 | -1.868 | -1.308 | -1.82 | -0.521 | -1.3892 |
| 237 | 30 | ต่ำ | -1.126 | -1.877 | -1.554 | -2.166 | -1.92 | -1.515 | -1.8072 |
| 238 | 30 | ต่ำ | -1.116 | -1.451 | -1.585 | -1.489 | -1.68 | -1.906 | -1.6226 |
| 239 | 30 | ต่ำ | -1.746 | -2.006 | -2.318 | -2.102 | -2.04 | -2.117 | -2.117 |
| 240 | 30 | ต่ำ | -1.067 | -2.067 | -2.298 | -2.181 | -1.66 | -2.183 | -2.0774 |
| 241 | 30 | ต่ำ | -1.065 | -1.85 | -1.731 | -1.557 | -1.68 | -1.492 | -1.6626 |
| 242 | 30 | ต่ำ | -1.645 | -1.977 | -2.169 | -1.955 | -2.37 | -1.598 | -2.0132 |
| 243 | 30 | ต่ำ | -1.744 | -1.699 | -2.427 | -2.331 | -2.34 | -1.679 | -2.0946 |
| 244 | 30 | ต่ำ | -0.825 | -1.13 | -1.101 | -1.602 | -2.06 | -2.011 | -1.5806 |
| 245 | 30 | ปานกลาง | -0.74 | -2.306 | -2.089 | -1.843 | -1.589 | -1.146 | -1.7946 |
| 246 | 30 | ปานกลาง | -0.694 | -2.078 | -1.985 | -1.946 | -1.58 | -0.897 | -1.6974 |
| 247 | 30 | ปานกลาง | -0.582 | -0.911 | -0.698 | -1.405 | -1.297 | -1.107 | -1.0836 |
| 248 | 30 | ปานกลาง | -0.58 | -2.328 | -2.145 | -1.653 | -1.921 | -1.696 | -1.9486 |
| 249 | 30 | ปานกลาง | -0.578 | -0.906 | -1.432 | -1.71 | -1.356 | -1.495 | -1.3798 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 250 | 30 | ปานกลาง | -0.538 | -2.306 | -1.409 | -2.065 | -1.583 | -1.664 | -1.8054 |
| 251 | 30 | ปานกลาง | -0.505 | -1.794 | -1.642 | -1.507 | -1.181 | 0.058 | -1.2132 |
| 252 | 30 | ปานกลาง | 0.505 | -2.571 | -2.356 | -1.697 | -1.974 | -1.352 | -1.99 |
| 253 | 30 | ปานกลาง | -0.463 | -0.911 | -0.9 | -1.474 | -0.913 | -0.88 | -1.0156 |
| 254 | 30 | ปานกลาง | -0.463 | -1.686 | -1.254 | -1.593 | -1.151 | -0.781 | -1.293 |
| 255 | 30 | ปานกลาง | -0.392 | -0.642 | -0.73 | -1.611 | -1.104 | -0.801 | -0.9776 |
| 256 | 30 | ปานกลาง | -0.212 | -0.694 | -0.929 | -2.12 | -0.971 | -1.066 | -1.156 |
| 257 | 30 | ปานกลาง | -0.07 | -1.079 | -0.99 | -1.227 | -0.532 | 0.014 | -0.7628 |
| 258 | 30 | ปานกลาง | -0.025 | -0.865 | -0.403 | -0.836 | -0.775 | -0.878 | -0.7514 |
| 259 | 30 | ปานกลาง | -0.025 | -0.23 | -0.196 | 0.324 | -0.15 | -0.144 | -0.0792 |
| 260 | 30 | ปานกลาง | 0.028 | 0.141 | 0.006 | -0.416 | -0.327 | -0.574 | -0.234 |
| 261 | 30 | ปานกลาง | 0.077 | 0.26 | 0.432 | 0.208 | 0.305 | 0.547 | 0.3504 |
| 262 | 30 | ปานกลาง | 0.124 | -1.481 | -0.482 | -0.616 | 0.142 | 0.11 | -0.4654 |
| 263 | 30 | ปานกลาง | 0.216 | -0.289 | -0.812 | -0.519 | -0.004 | 0.071 | -0.3106 |
| 264 | 30 | ปานกลาง | 0.259 | -0.203 | -0.628 | -0.823 | -0.119 | 0.184 | -0.3178 |
| 265 | 30 | ปานกลาง | 0.277 | 0.144 | 0.088 | 0.055 | -0.067 | -0.804 | -0.1168 |
| 266 | 30 | ปานกลาง | 0.351 | -0.905 | -0.734 | -0.188 | -0.673 | 0.1 | -0.48 |
| 267 | 30 | ปานกลาง | 0.352 | -0.514 | -0.979 | -0.712 | -0.561 | -0.579 | -0.669 |
| 268 | 30 | ปานกลาง | 0.394 | -0.531 | -0.668 | -0.739 | -0.639 | -0.759 | -0.6672 |
| 269 | 30 | ปานกลาง | 0.714 | -2.631 | -1.744 | -2.03 | -1.941 | -0.019 | -1.673 |
| 270 | 30 | ปานกลาง | 0.43 | -0.13 | -0.855 | -0.294 | -0.522 | -0.493 | -0.4588 |
| 271 | 30 | ปานกลาง | 0.44 | -0.216 | 0.313 | 0.738 | -0.64 | 0.642 | 0.1674 |
| 272 | 30 | ปานกลาง | 0.442 | -0.024 | 0.378 | 0.011 | 0.176 | 0.341 | 0.1764 |
| 273 | 30 | ปานกลาง | 0.456 | -1.202 | -0.384 | 0.224 | -0.111 | -0.13 | -0.3206 |
| 274 | 30 | ปานกลาง | 0.463 | -1.482 | 0.054 | 0.176 | -0.818 | 0.127 | -0.3886 |
| 275 | 30 | ปานกลาง | 0.477 | -0.217 | 0.204 | 0.125 | 0.423 | 0.236 | 0.1542 |
| 276 | 30 | ปานกลาง | 0.48 | -2.251 | -1.559 | -1.488 | -0.163 | 0.052 | -1.0818 |
| 277 | 30 | ปานกลาง | 0.484 | -0.009 | 0.106 | 0.33 | 0.067 | 0.91 | 0.2808 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 278 | 30 | ปานกลาง | 0.487 | -1.295 | -0.759 | -0.731 | -0.654 | -0.2 | -0.7278 |
| 279 | 30 | ปานกลาง | 0.51 | -0.994 | -1.065 | -0.683 | -0.872 | -0.164 | -0.7556 |
| 280 | 30 | ปานกลาง | 0.52 | -0.442 | -0.067 | -0.276 | -0.706 | 0.122 | -0.2738 |
| 281 | 30 | ปานกลาง | 0.531 | 0.023 | -0.826 | -0.284 | -0.78 | -0.509 | -0.4752 |
| 282 | 30 | ปานกลาง | 0.596 | -0.153 | 0.122 | -0.112 | -0.074 | 0.257 | 0.008 |
| 283 | 30 | ปานกลาง | 0.714 | 0.122 | -0.621 | 0.197 | 0.004 | 0.249 | -0.0098 |
| 284 | 30 | ปานกลาง | 0.746 | -0.128 | -0.225 | -0.407 | -0.475 | -0.308 | -0.3086 |
| 285 | 30 | สูง | 0.765 | 0.013 | -0.491 | -0.65 | -0.453 | -0.934 | -0.503 |
| 286 | 30 | สูง | 0.788 | -0.002 | 0.120 | 0.200 | -0.007 | 0.244 | 0.111 |
| 287 | 30 | สูง | 0.812 | -0.03 | 0.359 | 0.282 | 0.034 | 0.301 | 0.1892 |
| 288 | 30 | สูง | 1.332 | -1.181 | -0.888 | -1.231 | -0.763 | 0.177 | -0.7772 |
| 289 | 30 | สูง | 2.356 | 0.267 | 0.623 | -0.468 | -0.065 | 0.118 | 0.095 |
| 290 | 30 | สูง | 0.941 | -1.002 | -0.38 | 0.057 | 0.199 | -0.255 | -0.2762 |
| 291 | 30 | สูง | 0.987 | -0.21 | -0.496 | -0.911 | -0.328 | 0.155 | -0.358 |
| 292 | 30 | สูง | 1.000 | -0.184 | -0.302 | -0.183 | -0.572 | -0.183 | -0.2848 |
| 293 | 30 | สูง | 1.025 | -0.115 | -0.247 | -0.293 | 0.581 | 0.739 | 0.133 |
| 294 | 30 | สูง | 1.029 | -1.212 | -0.692 | -0.537 | -0.108 | -0.106 | -0.531 |
| 295 | 30 | สูง | 1.064 | -0.069 | -0.42 | -0.08 | -0.864 | -0.595 | -0.4056 |
| 296 | 30 | สูง | 1.069 | -0.26 | -1.118 | -1.52 | -0.421 | 0.372 | -0.5894 |
| 297 | 30 | สูง | 1.069 | -1.075 | -0.52 | -0.502 | -0.333 | 0.549 | -0.3762 |
| 298 | 30 | สูง | 1.075 | 0.005 | -0.984 | -0.976 | -1.378 | -0.608 | -0.7882 |
| 299 | 30 | สูง | 1.091 | -0.241 | 0.526 | 0.068 | 0.018 | -0.735 | -0.0728 |
| 300 | 30 | สูง | 1.136 | 0.073 | -0.075 | -0.411 | -1.42 | -0.476 | -0.4618 |
| 301 | 30 | สูง | 1.209 | 0.038 | -0.096 | -0.37 | -0.106 | 0.177 | -0.0714 |
| 302 | 30 | สูง | 1.303 | -0.003 | -1.16 | -1.508 | -1.195 | -0.671 | -0.9074 |
| 303 | 30 | สูง | 1.332 | -0.759 | 0.111 | 0.006 | -0.139 | 0.009 | -0.1544 |
| 304 | 30 | สูง | 1.344 | -0.288 | -0.507 | -0.082 | -0.502 | -0.291 | -0.334 |
| 305 | 30 | สูง | 1.533 | 0.281 | -0.469 | 0.075 | -0.126 | 0.236 | -0.0006 |

| ลำดับที่ | อัตราการใช้ ข้อสอบซ้ำ (%) | ระดับความ สามารถ | θ_a | θ_1 | θ_2 | θ_3 | θ_4 | θ_5 | $\bar{\theta}$ |
|----------|------------------------------|---------------------|------------|------------|------------|------------|------------|------------|----------------|
| 306 | 30 | สูง | 1.575 | 0.039 | 0.295 | 0.178 | -0.088 | 0.303 | 0.1454 |
| 307 | 30 | สูง | 1.464 | -0.006 | 0.185 | -0.007 | 0.036 | 0.462 | 0.134 |
| 308 | 30 | สูง | 1.758 | 0.388 | 0.73 | 0.015 | 0.405 | 0.986 | 0.5048 |
| 309 | 30 | สูง | 1.992 | 0.287 | 0.022 | 0.341 | 0.779 | 1.496 | 0.585 |
| 310 | 30 | สูง | 2.546 | 0.015 | 0.346 | 0.455 | 0.54 | 1.645 | 0.6002 |
| 311 | 30 | สูง | 1.059 | -0.842 | -1.26 | -0.3 | -0.657 | -0.408 | -0.6934 |
| 312 | 30 | สูง | 1.174 | -0.278 | 0.182 | 0.583 | 0.541 | 0.152 | 0.236 |
| 313 | 30 | สูง | 1.136 | -0.164 | -0.994 | -1.065 | -0.683 | -0.872 | -0.7556 |
| 314 | 30 | สูง | 0.941 | 0.054 | -1.482 | 0.176 | -0.818 | 0.127 | -0.3886 |
| 315 | 30 | สูง | 1.303 | -0.82 | -0.559 | -0.57 | -0.635 | -0.545 | -0.6258 |
| 316 | 30 | สูง | 1.174 | -0.009 | 0.106 | 0.33 | 0.067 | 0.91 | 0.2808 |
| 317 | 30 | สูง | 1.697 | -0.006 | 0.027 | -0.126 | -0.147 | -0.303 | -0.111 |
| 318 | 30 | สูง | 0.896 | -1.813 | -1.005 | -0.912 | -0.456 | -0.241 | -0.8854 |
| 319 | 30 | สูง | 0.987 | 0.023 | -0.826 | -0.284 | -0.78 | -0.509 | -0.4752 |

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
10 เปอร์เซนต์ ในรอบที่ 1

Test data for 5081 on PRETEST 12/19/00

Examinee ID : 5081

Examinee Name : warunee

| Item | Item ID | Resp | Key | Correct | Std | |
|------|---------|------|-----|---------|---------|-------|
| | | | | | Ability | Error |
| 1 | ENG061 | 1 | 2 | | -0.900 | 0.650 |
| 2 | ENG223 | 3 | 2 | | -1.307 | 0.495 |
| 3 | ENG146 | 3 | 3 | X | -1.168 | 0.454 |
| 4 | ENG094 | 2 | 4 | | -1.362 | 0.402 |
| 5 | ENG099 | 1 | 1 | X | -1.280 | 0.383 |
| 6 | ENG017 | 1 | 1 | X | -1.169 | 0.372 |
| 7 | ENG098 | 4 | 4 | X | -1.008 | 0.368 |
| 8 | ENG120 | 2 | 2 | X | -0.897 | 0.357 |
| 9 | ENG078 | 2 | 2 | X | -0.814 | 0.347 |
| 10 | ENG100 | 2 | 2 | X | -0.701 | 0.338 |
| 11 | ENG089 | 4 | 1 | | -0.798 | 0.311 |
| 12 | ENG164 | 1 | 1 | X | -0.745 | 0.305 |
| 13 | ENG140 | 1 | 3 | | -0.799 | 0.289 |

13items were administered.

The final Bayesian posterior mean was: -0.799

The final Bayesian posterior variance was: 0.083

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
10 เปอร์เซนต์ ในรอบที่ 2

Test data for 5081 on PRETEST 01/09/01

Examinee ID : 5081

Examinee Name : warunee

| Item | Item ID | Resp | Key | Correct | Std | |
|------|---------|------|-----|---------|---------|-------|
| | | | | | Ability | Error |
| 1 | ENG061 | 1 | 2 | | -0.900 | 0.650 |
| 2 | ENG007 | 2 | 3 | | -1.254 | 0.565 |
| 3 | ENG101 | 1 | 1 | X | -1.116 | 0.539 |
| 4 | ENG119 | 4 | 3 | | -1.347 | 0.481 |
| 5 | ENG070 | 1 | 1 | X | -1.231 | 0.466 |
| 6 | ENG043 | 3 | 3 | X | -1.102 | 0.456 |
| 7 | ENG039 | 4 | 4 | X | -0.994 | 0.443 |
| 8 | ENG002 | 2 | 2 | X | -0.902 | 0.434 |
| 9 | ENG011 | 2 | 2 | X | -0.813 | 0.422 |
| 10 | ENG228 | 3 | 2 | | -0.924 | 0.391 |
| 11 | ENG109 | 1 | 1 | X | -0.840 | 0.383 |
| 12 | ENG208 | 4 | 4 | X | -0.770 | 0.374 |
| 13 | ENG165 | 3 | 1 | | -0.854 | 0.360 |
| 14 | ENG127 | 2 | 4 | | -1.015 | 0.341 |
| 15 | ENG012 | 1 | 3 | | -1.157 | 0.326 |
| 16 | ENG126 | 1 | 3 | | -1.330 | 0.307 |
| 17 | ENG122 | 4 | 4 | X | -1.294 | 0.304 |
| 18 | ENG044 | 1 | 3 | | -1.386 | 0.295 |

18items were administered.

The final Bayesian posterior mean was: -1.386

The final Bayesian posterior variance was: 0.087

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
10 เปอร์เซนต์ ในรอบที่ 3

Test data for 5081 on PRETEST 01/19/01

Examinee ID : 5081

Examinee Name : warunee

| Item | Item ID | Resp | Key | Correct | Std Ability | Std Error |
|------|---------|------|-----|---------|-------------|-----------|
| 1 | ENG061 | 1 | 2 | | -0.900 | 0.650 |
| 2 | ENG007 | 3 | 3 | X | -0.705 | 0.611 |
| 3 | ENG089 | 1 | 1 | X | -0.398 | 0.575 |
| 4 | ENG140 | 3 | 3 | X | -0.090 | 0.512 |
| 5 | ENG200 | 4 | 1 | | -0.363 | 0.432 |
| 6 | ENG100 | 2 | 2 | X | -0.242 | 0.408 |
| 7 | ENG006 | 2 | 3 | | -0.439 | 0.362 |
| 8 | ENG120 | 2 | 2 | X | -0.372 | 0.348 |
| 9 | ENG103 | 1 | 2 | | -0.467 | 0.322 |
| 10 | ENG078 | 2 | 2 | X | -0.421 | 0.313 |
| 11 | ENG097 | 2 | 3 | | -0.493 | 0.298 |

11items were administered.

The final Bayesian posterior mean was: -0.493

The final Bayesian posterior variance was: 0.089

ผลการทดสอบแบบปรับเหมาะกับความสามารรถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
10 เปอร์เซนต์ ในรอบที่ 4

Test data for 5081 on PRETEST 02/06/01

Examinee ID : 5081

Examinee Name : warunee

| Item | Item ID | Resp | Key | Correct | Std Ability | Error |
|------|---------|------|-----|---------|-------------|-------|
| 1 | ENG061 | 1 | 2 | | -0.900 | 0.650 |
| 2 | ENG098 | 4 | 4 | X | -0.571 | 0.587 |
| 3 | ENG094 | 4 | 4 | X | -0.431 | 0.540 |
| 4 | ENG228 | 2 | 2 | X | -0.236 | 0.506 |
| 5 | ENG145 | 4 | 4 | X | 0.003 | 0.488 |
| 6 | ENG148 | 1 | 2 | | -0.179 | 0.425 |
| 7 | ENG068 | 1 | 3 | | -0.338 | 0.389 |
| 8 | ENG049 | 1 | 2 | | -0.444 | 0.368 |
| 9 | ENG164 | 1 | 1 | X | -0.393 | 0.357 |
| 10 | ENG165 | 3 | 1 | | -0.497 | 0.342 |
| 11 | ENG017 | 1 | 1 | X | -0.455 | 0.331 |
| 12 | ENG043 | 3 | 3 | X | -0.415 | 0.324 |
| 13 | ENG081 | 4 | 4 | X | -0.365 | 0.320 |
| 14 | ENG105 | 4 | 4 | X | -0.285 | 0.315 |
| 15 | ENG045 | 3 | 4 | | -0.359 | 0.302 |
| 16 | ENG131 | 3 | 3 | X | -0.310 | 0.299 |

16items were administered.

The final Bayesian posterior mean was: -0.310

The final Bayesian posterior variance was: 0.090

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
10 เปอร์เซนต์ ในรอบที่ 5

Test data for 5081 on PRETEST 02/13/01

Examinee ID : 5081

Examinee Name : warunee

| Item | Item ID | Resp | Key | Correct | Std | |
|------|---------|------|-----|---------|---------|-------|
| | | | | | Ability | Error |
| 1 | ENG061 | 1 | 2 | | -0.900 | 0.650 |
| 2 | ENG098 | 4 | 4 | X | -0.571 | 0.587 |
| 3 | ENG089 | 4 | 1 | | -0.865 | 0.482 |
| 4 | ENG120 | 2 | 2 | X | -0.702 | 0.456 |
| 5 | ENG100 | 2 | 2 | X | -0.533 | 0.434 |
| 6 | ENG140 | 3 | 3 | X | -0.312 | 0.406 |
| 7 | ENG006 | 3 | 3 | X | -0.196 | 0.389 |
| 8 | ENG200 | 4 | 1 | | -0.351 | 0.350 |
| 9 | ENG103 | 2 | 2 | X | -0.241 | 0.346 |
| 10 | ENG097 | 4 | 3 | | -0.350 | 0.323 |
| 11 | ENG078 | 2 | 2 | X | -0.310 | 0.314 |
| 12 | ENG143 | 1 | 1 | X | -0.201 | 0.313 |
| 13 | ENG217 | 4 | 4 | X | -0.145 | 0.311 |
| 14 | ENG161 | 3 | 2 | | -0.213 | 0.292 |

14 items were administered.

The final Bayesian posterior mean was: -0.213

The final Bayesian posterior variance was: 0.085

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
20 เปอร์เซนต์ ในรอบที่ 1

Test data for 2088 on PRETEST 01/24/01

Examinee ID : 2088

Examinee Name : sutarat

| Item | Item ID | Resp | Key | Correct | Std Ability | Error |
|------|---------|------|-----|---------|----------------|-------|
| 1 | ENG061 | 1 | 2 | | -0.900 | 0.650 |
| 2 | ENG223 | 3 | 2 | | -1.307 | 0.495 |
| 3 | ENG146 | 3 | 3 | X | -1.168 | 0.454 |
| 4 | ENG094 | 3 | 4 | | -1.362 | 0.402 |
| 5 | ENG099 | 4 | 1 | | -1.602 | 0.360 |
| 6 | ENG126 | 1 | 3 | | -1.752 | 0.338 |
| 7 | ENG010 | 2 | 1 | | -1.973 | 0.305 |
| 8 | ENG149 | 2 | 1 | | -2.088 | 0.292 |

8 items were administered.

The final Bayesian posterior mean was: -2.088

The final Bayesian posterior variance was: 0.085

สถาบันวิจัยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
20 เปอร์เซนต์ ในรอบที่ 2

Test data for 2088 on PRETEST 01/29/01

Examinee ID : 2088

Examinee Name : sutarat

| Item | Item ID | Resp | Key | Correct | Std Ability | Error |
|------|---------|------|-----|---------|----------------|-------|
| 1 | ENG061 | 1 | 2 | | -0.900 | 0.650 |
| 2 | ENG223 | 3 | 2 | | -1.307 | 0.495 |
| 3 | ENG017 | 2 | 1 | | -1.501 | 0.442 |
| 4 | ENG124 | 3 | 1 | | -1.672 | 0.405 |
| 5 | ENG101 | 3 | 1 | | -1.791 | 0.382 |
| 6 | ENG129 | 1 | 4 | | -1.936 | 0.361 |
| 7 | ENG012 | 1 | 3 | | -2.016 | 0.348 |
| 8 | ENG044 | 1 | 3 | | -2.088 | 0.338 |
| 9 | ENG122 | 3 | 4 | | -2.147 | 0.328 |
| 10 | ENG127 | 3 | 4 | | -2.191 | 0.320 |
| 11 | ENG038 | 2 | 4 | | -2.289 | 0.312 |
| 12 | ENG110 | 1 | 1 | X | -2.266 | 0.310 |
| 13 | ENG156 | 3 | 3 | X | -2.231 | 0.309 |
| 14 | ENG070 | 1 | 1 | X | -2.173 | 0.313 |
| 15 | ENG087 | 1 | 4 | | -2.209 | 0.309 |
| 16 | ENG069 | 2 | 1 | | -2.242 | 0.304 |
| 17 | ENG163 | 2 | 1 | | -2.270 | 0.299 |

17items were administered.

The final Bayesian posterior mean was: -2.270

The final Bayesian posterior variance was: 0.090

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
20 เปอร์เซนต์ ในรอบที่ 3

Test data for 2088 on PRETEST 02/05/01

Examinee ID : 2088

Examinee Name : sutarut

| Item | Item ID | Resp | Key | Correct | Std | |
|------|---------|------|-----|---------|---------|-------|
| | | | | | Ability | Error |
| 1 | ENG061 | 2 | 2 | X | 0.443 | 0.830 |
| 2 | ENG163 | 1 | 1 | X | 0.526 | 0.795 |
| 3 | ENG069 | 1 | 1 | X | 0.583 | 0.771 |
| 4 | ENG083 | 2 | 4 | | 0.066 | 0.571 |
| 5 | ENG140 | 1 | 3 | | -0.373 | 0.452 |
| 6 | ENG089 | 4 | 1 | | -0.618 | 0.385 |
| 7 | ENG098 | 3 | 4 | | -0.831 | 0.336 |
| 8 | ENG094 | 3 | 4 | | -0.996 | 0.307 |
| 9 | ENG120 | 1 | 2 | | -1.076 | 0.293 |

9items were administered.

The final Bayesian posterior mean was: -1.076

The final Bayesian posterior variance was: 0.086

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
20 เปอร์เซนต์ ในรอบที่ 4

Test data for 2088 on PRETEST 02/06/01

Examinee ID : 2088

Examinee Name : sutarat



| Item | Item ID | Resp | Key | Correct | Std Ability | Std Error |
|------|---------|------|-----|---------|-------------|-----------|
| 1 | ENG061 | 2 | 2 | X | 0.443 | 0.830 |
| 2 | ENG083 | 3 | 4 | | -0.061 | 0.618 |
| 3 | ENG200 | 2 | 1 | | -0.434 | 0.499 |
| 4 | ENG100 | 1 | 2 | | -0.695 | 0.435 |
| 5 | ENG078 | 3 | 2 | | -0.919 | 0.392 |
| 6 | ENG017 | 3 | 1 | | -1.109 | 0.356 |
| 7 | ENG146 | 2 | 3 | | -1.394 | 0.305 |
| 8 | ENG099 | 1 | 1 | X | -1.343 | 0.295 |

8 items were administered.

The final Bayesian posterior mean was: -1.343

The final Bayesian posterior variance was: 0.087

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
20 เปอร์เซ็นต์ ในรอบที่ 5

Test data for 2088 on PRETEST 02/12/01

Examinee ID : 2088

Examinee Name : SUTARAT



| Item | Item ID | Resp | Key | Correct | Ability | Std Error |
|------|---------|------|-----|---------|---------|--------------|
| 1 | ENG061 | 2 | 2 | X | 0.443 | 0.830 |
| 2 | ENG083 | 2 | 4 | | -0.061 | 0.618 |
| 3 | ENG140 | 3 | 3 | X | 0.199 | 0.537 |
| 4 | ENG148 | 1 | 2 | | -0.072 | 0.449 |
| 5 | ENG089 | 4 | 1 | | -0.402 | 0.374 |
| 6 | ENG098 | 3 | 4 | | -0.657 | 0.323 |
| 7 | ENG094 | 1 | 4 | | -0.840 | 0.296 |

7 items were administered.

The final Bayesian posterior mean was: -0.840

The final Bayesian posterior variance was: 0.088

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
30 เปอร์เซนต์ ในรอบที่ 1

Test data for 3057 on PRETEST 12/21/00

Examinee ID : 3057

Examinee Name : pittaya



| Item | Item ID | Resp | Key | Correct | Std Ability | Error |
|------|---------|------|-----|---------|----------------|-------|
| 1 | ENG200 | 2 | 1 | | -0.737 | 0.711 |
| 2 | ENG061 | 3 | 2 | | -1.041 | 0.566 |
| 3 | ENG223 | 3 | 2 | | -1.325 | 0.456 |
| 4 | ENG146 | 3 | 3 | X | -1.202 | 0.421 |
| 5 | ENG094 | 2 | 4 | | -1.366 | 0.379 |
| 6 | ENG099 | 1 | 1 | X | -1.292 | 0.362 |
| 7 | ENG017 | 4 | 1 | | -1.405 | 0.338 |
| 8 | ENG124 | 3 | 1 | | -1.521 | 0.319 |
| 9 | ENG126 | 1 | 3 | | -1.650 | 0.303 |
| 10 | ENG010 | 4 | 1 | | -1.850 | 0.277 |

10items were administered.

The final Bayesian posterior mean was: -1.850

The final Bayesian posterior variance was: 0.077

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
30 เปอร์เซ็นต์ ในรอบที่ 2

Test data for 3057 on PRETEST 01/04/01

Examinee ID : 3057

Examinee Name : pittaya

| Item | Item ID | Resp | Key | Correct | Std Ability | Error |
|------|---------|------|-----|---------|-------------|-------|
| 1 | ENG061 | 1 | 2 | | -0.900 | 0.650 |
| 2 | ENG223 | 3 | 2 | | -1.307 | 0.495 |
| 3 | ENG146 | 3 | 3 | X | -1.168 | 0.454 |
| 4 | ENG078 | 2 | 2 | X | -1.023 | 0.442 |
| 5 | ENG120 | 4 | 2 | | -1.175 | 0.406 |
| 6 | ENG164 | 4 | 1 | | -1.302 | 0.375 |
| 7 | ENG101 | 1 | 1 | X | -1.234 | 0.366 |
| 8 | ENG070 | 2 | 1 | | -1.364 | 0.346 |
| 9 | ENG127 | 1 | 4 | | -1.469 | 0.332 |
| 10 | ENG012 | 1 | 3 | | -1.568 | 0.319 |
| 11 | ENG129 | 3 | 4 | | -1.691 | 0.306 |
| 12 | ENG149 | 1 | 1 | X | -1.663 | 0.301 |
| 13 | ENG122 | 1 | 4 | | -1.731 | 0.293 |

13 items were administered.

The final Bayesian posterior mean was: -1.731

The final Bayesian posterior variance was: 0.086

ผลการทดสอบแบบปรับเหมาะกับความสามารรถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
30 เปอร์เซนต์ ในรอบที่ 3

Test data for 3057 on PRETEST 01/09/01

Examinee ID : 3057

Examinee Name : pittaya

| Item | Item ID | Resp | Key | Correct | Std | |
|------|---------|------|-----|---------|---------|-------|
| | | | | | Ability | Error |
| 1 | ENG061 | 3 | 2 | | -0.900 | 0.650 |
| 2 | ENG223 | 3 | 2 | | -1.307 | 0.495 |
| 3 | ENG146 | 3 | 3 | X | -1.168 | 0.454 |
| 4 | ENG078 | 2 | 2 | X | -1.023 | 0.442 |
| 5 | ENG007 | 4 | 3 | | -1.191 | 0.411 |
| 6 | ENG119 | 2 | 3 | | -1.325 | 0.383 |
| 7 | ENG126 | 1 | 3 | | -1.531 | 0.355 |
| 8 | ENG010 | 1 | 1 | X | -1.497 | 0.344 |
| 9 | ENG039 | 1 | 4 | | -1.565 | 0.332 |
| 10 | ENG044 | 3 | 3 | X | -1.521 | 0.328 |
| 11 | ENG043 | 3 | 3 | X | -1.449 | 0.328 |
| 12 | ENG011 | 1 | 2 | | -1.509 | 0.317 |
| 13 | ENG208 | 4 | 4 | X | -1.447 | 0.317 |
| 14 | ENG109 | 3 | 1 | | -1.495 | 0.309 |
| 15 | ENG163 | 4 | 1 | | -1.550 | 0.302 |
| 16 | ENG069 | 1 | 1 | X | -1.511 | 0.300 |
| 17 | ENG002 | 3 | 2 | | -1.557 | 0.291 |

17items were administered.

The final Bayesian posterior mean was: -1.557

The final Bayesian posterior variance was: 0.085

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
30 เปอร์เซนต์ ในรอบที่ 4

Test data for 3057 on PRETEST 01/15/01

Examinee ID : 3057

Examinee Name : pittaya

| Item | Item ID | Resp | Key | Correct | Std | |
|------|---------|------|-----|---------|---------|-------|
| | | | | | Ability | Error |
| 1 | ENG061 | 3 | 2 | | -0.900 | 0.650 |
| 2 | ENG223 | 3 | 2 | | -1.307 | 0.495 |
| 3 | ENG146 | 3 | 3 | X | -1.168 | 0.454 |
| 4 | ENG078 | 2 | 2 | X | -1.023 | 0.442 |
| 5 | ENG007 | 4 | 3 | | -1.191 | 0.411 |
| 6 | ENG120 | 2 | 2 | X | -1.045 | 0.401 |
| 7 | ENG164 | 3 | 1 | | -1.189 | 0.369 |
| 8 | ENG101 | 4 | 1 | | -1.339 | 0.348 |
| 9 | ENG070 | 1 | 1 | X | -1.275 | 0.341 |
| 10 | ENG127 | 1 | 4 | | -1.384 | 0.327 |
| 11 | ENG012 | 2 | 3 | | -1.486 | 0.315 |
| 12 | ENG129 | 3 | 4 | | -1.612 | 0.301 |
| 13 | ENG122 | 2 | 4 | | -1.683 | 0.293 |

13items were administered.

The final Bayesian posterior mean was: -1.683

The final Bayesian posterior variance was: 0.086

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีอัตราการใช้ข้อสอบ
30 เปอร์เซนต์ ในรอบที่ 5

Test data for 3057 on PRETEST 01/29/01

Examinee ID : 3057

Examinee Name : pittaya

| Item | Item ID | Resp | Key | Correct | Std | |
|------|---------|------|-----|---------|---------|-------|
| | | | | | Ability | Error |
| 1 | ENG061 | 3 | 2 | | -0.900 | 0.650 |
| 2 | ENG223 | 3 | 2 | | -1.307 | 0.495 |
| 3 | ENG146 | 3 | 3 | X | -1.168 | 0.454 |
| 4 | ENG078 | 2 | 2 | X | -1.023 | 0.442 |
| 5 | ENG098 | 3 | 4 | | -1.184 | 0.389 |
| 6 | ENG094 | 1 | 4 | | -1.330 | 0.354 |
| 7 | ENG099 | 1 | 1 | X | -1.268 | 0.340 |
| 8 | ENG017 | 2 | 1 | | -1.372 | 0.319 |
| 9 | ENG126 | 1 | 3 | | -1.515 | 0.302 |
| 10 | ENG010 | 1 | 1 | X | -1.492 | 0.295 |

10items were administered.

The final Bayesian posterior mean was: -1.492

The final Bayesian posterior variance was: 0.087

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวกนกวรรณ รัตนธน สำเร็จการศึกษาปริญญาตรีครุศาสตรบัณฑิต ตามโครงการ
 ครูทายาท (หลักสูตร 4 ปี) วิชาเอกคอมพิวเตอร์ศึกษา คณะวิทยาศาสตร์และเทคโนโลยี สถาบันราชภัฏ
 เพชรบุรี ในปีการศึกษา 2539 และเข้าศึกษาต่อในหลักสูตรครุศาสตรมหาบัณฑิต ที่จุฬาลงกรณ์
 มหาวิทยาลัย เมื่อ พ.ศ.2542 ปัจจุบันรับราชการตำแหน่ง อาจารย์ 1 ระดับ 4 ที่โรงเรียนหนองหญ้าไซ
 วิทยา อำเภอหนองหญ้าไซ จังหวัดสุพรรณบุรี



สถาบันวิทยบริการ
 จุฬาลงกรณ์มหาวิทยาลัย