# การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ประพล เปรมทองสุข

ดุษฎีนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
กันยายน 2560
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา

คณะกรรมการควบคุมดุษฎีนิพนธ์และคณะกรรมการสอบดุษฎีนิพนธ์ ได้พิจารณาดุษฎีนิพนธ์ของ ประพล เปรมทองสุข ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา ของมหาวิทยาลัยบูรพาได้

คณะกรรมการควบคุมดุษฎีนิพนธ์

.....อาจารย์ที่ปรึกษาหลัก
(รองศาสตราจารย์ ดร. เสรี ชัดแช้ม)

.....อาจารย์ที่ปรึกษาร่วม
(ดร. ปิยะทิพย์ ประดุจพรม)

คณะกรรมการสอบดุษฎีนิพนธ์

.....ประธาน
(ผู้ช่วยศาสตราจารย์ ดร. สุรชัย มีชาญ)

.....กรรมการ
(รองศาสตราจารย์ ดร. เสรี ชัดแช้ม)

.....กรรมการ
(ดร. ปิยะทิพย์ ประดุจพรม)

.....กรรมการ
(ดร. โสพส สุขานนท์สวัสดิ์)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. พูลพงศ์ สุขสว่าง)

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา อนุมัติให้รับดุษฎีนิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา ของมหาวิทยาลัยบูรพา

.....คณบดีวิทยาลัยวิทยาการวิจัย
(ผู้ช่วยศาสตราจารย์ ดร. สุชาดา กรเพชรปาณี) และวิทยาการปัญญา
วันที่...8.....เดือน...มิถุนายน.....พ.ศ. 2560

# กิตติกรรมประกาศ

ดุษฎีนิพนธ์ฉบับนี้สำเร็จลุล่วงเป็นอย่างดี จากความกรุณาของรองศาสตราจารย์ ดร. เสรี ชัดแช้ม อาจารย์ที่ปรึกษาหลัก และ อาจารย์ ดร. ปิยะทิพย์ ประดุจพรม อาจารย์ที่ปรึกษาร่วม ที่ได้ถ่ายทอดความรู้ ให้คำแนะนำในการแก้ไข ตลอดจนคอยกระตุ้นและติดตามการดำเนินการวิจัย จนดุษฎีนิพนธ์ฉบับนี้สำเร็จได้ จึงขอกราบขอบพระคุณอาจารย์ที่ปรึกษาทั้งสองเป็นอย่างสูง

ขอขอบพระคุณอาจารย์ ดร. ทิพย์ ขำอยู่ และ อาจารย์ ดร. โสพส สุขานนท์สวัสดิ์ ที่อนุเคราะห์คลังข้อสอบที่ใช้ในการศึกษา ตลอดจนให้คำแนะนำเบื้องต้นจนได้แนวทางการศึกษาในหัวข้อเรื่องนี้

ขอขอบพระคุณผู้ช่วยศาสตราจารย์ ดร. พีรพัฒน์ ยางกลาง ผู้ช่วยศาสตราจารย์ ดร. สิริชัย ดีเลิศ และ อาจารย์ ดร. สุชาดา สกลกิจรุ่งโรจน์ ซึ่งเป็นผู้เชี่ยวชาญในการประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และให้ข้อเสนอแนะในการปรับปรุงจนทำให้โปรแกรมมีความสมบูรณ์ยิ่งขึ้น

ขอขอบคุณอาจารย์ศศิกานต์ เฮ้าจัตุรัส และอาจารย์พัชรินทร์ พระราช ที่ให้ความอนุเคราะห์คัดเลือกข้อสอบที่ใช้ในการศึกษาวิจัย ตลอดจนตรวจทานความถูกต้องของไวยากรณ์ภาษาอังกฤษในดุษฎีนิพนธ์ฉบับนี้

สุดท้ายขอขอบพระคุณบิดา มารดา ภรรยา และบุตรสาว ซึ่งแม้ไม่มีส่วนเกี่ยวข้องในการดำเนินการวิจัย แต่กำลังใจและคำกระตุ้นที่ได้รับ กลับมีส่วนสำคัญอย่างยิ่งจนทำให้ดุษฎีนิพนธ์ฉบับนี้สำเร็จลุล่วงได้

ประโยชน์ที่ได้รับจากดุษฎีนิพนธ์ฉบับนี้ ขอมอบให้กับคณาจารย์และผู้มีพระคุณทุกท่าน ทั้งที่ได้เอ่ยนามและไม่ได้เอ่ยนาม ที่มีส่วนบ่มเพาะความรู้ของผู้วิจัยให้เติบโตขึ้นจนถึงปัจจุบัน

ประพล เปรมทองสุข

53810282: สาขาวิชา: การวิจัยและสถิติทางวิทยาการปัญญา
ปร.ด. (การวิจัยและสถิติทางวิทยาการปัญญา)
คำสำคัญ: การคัดเลือกข้อสอบข้อถัดไป/ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์/ เกณฑ์ของเฮอร์วิคซ์/ การควบคุมการใช้ข้อสอบ

ประพล เปรมทองสุข: การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (DEVELOPMENT OF THE NEXT ITEM SELECTION PROCEDURE USING HURWICZ CRITERION WITH THE ITEM EXPOSURE CONTROL FOR COMPUTERIZED ADAPTIVE TESTING) คณะกรรมการควบคุมดุษฎีนิพนธ์: เสรี ชัดแช้ม, ค.ด., ปิยะทิพย์ ประดุจพรม, Ph.D. 216 หน้า. ปี พ.ศ. 2560.

การวิจัยนี้มีวัตถุประสงค์ 1) เพื่อพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ 2) เพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 4 วิธี คือ 2.1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (วิธีการ MIC) 2.2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (วิธีการ RDM) 2.3) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ (วิธีการ HC) และ 2.4) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ (วิธีการ HC-Ex) โดยเปรียบเทียบในด้านการประมาณค่าความสามารถของผู้สอบ ด้านความยาวของแบบทดสอบ และด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 3) เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ และ 4) เพื่อศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ กลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาตรี ปีการศึกษา 2559 จำนวน 30 คน และวิเคราะห์ข้อมูลด้วยการวิเคราะห์สหสัมพันธ์แบบเพียร์สัน

ผลการวิจัยปรากฏว่า วิธีการคัดเลือกข้อสอบข้อถัดไปและมีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้น มีประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบและลดความยาวของแบบทดสอบได้ โดยพิจารณาจากค่าความลำเอียงเฉลี่ย ปรากฏว่า วิธีการ RDM วิธีการ HC และวิธีการ HC-Ex มีประสิทธิภาพสูงสุดเทียบเท่ากัน เมื่อเปรียบเทียบด้านความยาวของแบบทดสอบ ปรากฏว่า วิธีการ HC มีประสิทธิภาพสูงสุดเทียบเท่ากับวิธีการ RDM และเมื่อเปรียบเทียบด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ปรากฏว่า วิธีการ MIC มีประสิทธิภาพสูงสุด และโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น (วิธีการ HC-Ex) มีความเหมาะสมในการใช้งานระดับมากถึงมากที่สุด ในด้านความสะดวกในการใช้โปรแกรม ความถูกต้อง ลักษณะทั่วไปของโปรแกรม และความชัดเจนของคู่มือ และค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น (วิธีการ HC-Ex) กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ปรากฏว่า มีความสัมพันธ์กันทางบวก ($p < .01$)

53810282: MAJOR: RESEARCH AND STATISTICS IN COGNITIVE SCIENCE
Ph.D. (RESEARCH AND STATISTICS IN COGNITIVE SCIENCE)
KEYWORDS: NEXT ITEM SELECTION/ COMPUTERIZED ADAPTIVE TESTING/
HURWICZ CRITERION/ ITEM EXPOSURE CONTROL
PRAPON PREMTHONGSUK: DEVELOPMENT OF THE NEXT ITEM SELECTION PROCEDURE USING HURWICZ CRITERION WITH THE ITEM EXPOSURE CONTROL FOR COMPUTERIZED ADAPTIVE TESTING. ADVISORY COMMITTEE: SEREE CHADCHAM, Ph.D., PIYATHIP PRADUJPROM, Ph.D. 216 P. 2017.

This research aimed at 1) developing the next item selection procedure using Hurwicz Criterion with the item exposure control; 2) comparing the efficiency of four the next item selection procedures: 2.1) Maximum Information Criterion (MIC), 2.2) Risk Decision Making (RDM), 2.3) Hurwicz Criterion (HC), and 2.4) Hurwicz Criterion with the item exposure control (HC-Ex) by comparing the efficiency of the ability estimation of examinees, the test length, and the number of item with the item exposure rate more than 0.2; 3) developing the computerized adaptive testing program using the next item selection procedure using Hurwicz Criterion with the item exposure control, and 4) studying the correlation of the ability estimation of the examinees via the computerized adaptive testing and the observed score via the computerized testing. The samples were 30 undergraduate students in academic year 2016, and Pearson's Product Moment Correlation was used to analyze the data.

The results showed that the developed next item selection procedure using Hurwicz Criterion with the item exposure control had the efficiency of ability estimation and reduction of the test length, considering the average bias, RDM, HC, and HC-Ex were equally efficient. By comparing the efficiency of the test length, HC was found to be as efficient as that of RDM. For the comparison of the efficiency of the number of item with the item exposure rate of more than 0.2, MIC was the most efficient. The developed computerized adaptive testing program using the developed HC-Ex was suitable for use at the high and very high level in terms of convenience, reliability, program features, and instruction manual. The ability estimation of the examinees via the computerized adaptive testing using the developed HC-Ex and the observed score via the computerized testing were positively correlated ($p < .01$).

# สารบัญ


หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
สารบัญ ..... ฉ
สารบัญตาราง ..... ซ
สารบัญภาพ ..... ฎ
บทที่
1 บทนำ ..... 1
ความเป็นมาและความสำคัญของปัญหา ..... 1
วัตถุประสงค์ของการวิจัย ..... 6
กรอบแนวคิดการวิจัย ..... 6
สมมติฐานของการวิจัย ..... 9
ขอบเขตของการวิจัย ..... 9
ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย ..... 11
นิยามศัพท์เฉพาะ ..... 12
2 เอกสารและงานวิจัยที่เกี่ยวข้อง ..... 17
ตอนที่ 1 การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 17
ตอนที่ 2 ทฤษฎีการตัดสินใจและงานวิจัยที่เกี่ยวข้อง ..... 44
ตอนที่ 3 การศึกษาในสถานการณ์จำลองแบบมอนติคาร์โล ..... 51
ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 53
3 วิธีดำเนินการวิจัย ..... 62
ระยะที่ 1 การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์
และมีการควบคุมการใช้ข้อสอบ ..... 63
ระยะที่ 2 การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ..... 71
ระยะที่ 3 การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 77
ระยะที่ 4 การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบ
ที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับ
คะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ..... 91
4 ผลการวิจัย ..... 95
ตอนที่ 1 ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์
และมีการควบคุมการใช้ข้อสอบ ..... 96
ตอนที่ 2 ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ..... 98
ตอนที่ 3 ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 107

# สารบัญ (ต่อ)


บทที่ หน้า

ตอนที่ 4 ผลการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ..... 123

5 สรุปผลและอภิปรายผล ..... 126

สรุปผลการวิจัย ..... 126

อภิปรายผล ..... 129

ข้อเสนอแนะในการนำผลการวิจัยไปใช้ ..... 135

ข้อเสนอแนะสำหรับการวิจัยต่อไป ..... 135

บรรณานุกรม ..... 136

ภาคผนวก ..... 143

ภาคผนวก ก รายชื่อผู้เชี่ยวชาญ ..... 144

ภาคผนวก ข ผลการศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 146

ภาคผนวก ค คลังข้อสอบจำลองที่ใช้ในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 173

ภาคผนวก ง คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี ..... 187

ภาคผนวก จ แบบประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยผู้เชี่ยวชาญ ..... 208

ภาคผนวก ฉ แบบประเมินโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยผู้สอบ ..... 211

ภาคผนวก ช ตัวอย่างการแสดงผลการทดสอบจากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 213

ประวัติย่อของผู้วิจัย ..... 216

# สารบัญตาราง


| ตารางที่ | | หน้า |
|---|---|---|
| 2-1 | ผลตอบแทนจากการลงทุนในบริษัท A, B และ C | 46 |
| 2-2 | ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ กรณีที่ผู้สอบตอบถูก และกรณีที่ผู้สอบตอบผิด เมื่อได้รับข้อสอบข้อที่ 1 ถึงข้อที่ 5 | 47 |
| 3-1 | ตัวอย่างการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ | 69 |
| 3-2 | ตัวอย่างการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด | 73 |
| 3-3 | ตัวอย่างการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง | 75 |
| 3-4 | สัญลักษณ์และความหมายที่ใช้ในการวิเคราะห์แผนผังบริบท | 78 |
| 3-5 | สัญลักษณ์และความหมายที่ใช้ในผังการไหลของข้อมูล | 79 |
| 3-6 | แฟ้มข้อมูลข้อสอบ | 84 |
| 3-7 | แฟ้มข้อมูลรายการคำตอบ | 84 |
| 3-8 | แฟ้มข้อมูลผู้สอบ | 85 |
| 3-9 | แฟ้มข้อมูลการทดสอบ | 85 |
| 3-10 | แฟ้มข้อมูลสาขาวิชา | 85 |
| 3-11 | แฟ้มข้อมูลคณะวิชา | 86 |
| 3-12 | แฟ้มข้อมูลข้อสอบการทดสอบด้วยคอมพิวเตอร์ | 86 |
| 3-13 | ลำดับการทดสอบ | 93 |
| 4-1 | ค่าสถิติพื้นฐานของค่าความสามารถที่แท้จริงของผู้สอบ | 98 |
| 4-2 | ค่าสถิติพื้นฐานของค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบจำลอง | 99 |
| 4-3 | ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยและค่าความลำเอียงเฉลี่ยของวิธีการคัดเลือกข้อสอบข้อถัดไป | 100 |
| 4-4 | ผลการวิเคราะห์ความแตกต่างของค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยและค่าความลำเอียงเฉลี่ย จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป | 101 |
| 4-5 | ผลการเปรียบเทียบพหุคูณของค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยจำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป | 101 |
| 4-6 | ผลการเปรียบเทียบพหุคูณของค่าความลำเอียงเฉลี่ย จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป | 102 |
| 4-7 | ค่าสถิติพื้นฐานความยาวของแบบทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป | 103 |
| 4-8 | ผลการวิเคราะห์ความแตกต่างของความยาวของแบบทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป | 103 |

# สารบัญตาราง (ต่อ)


ตารางที่ หน้า

4-9 ผลการเปรียบเทียบค่าเฉลี่ยของอันดับเป็นรายคู่ของความยาวของแบบทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป ..... 104

4-10 จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป ..... 105

4-11 ผลการวิเคราะห์ความแตกต่างของจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป ..... 105

4-12 ผลการเปรียบเทียบค่าเฉลี่ยของอันดับเป็นรายคู่ของจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป ..... 106

4-13 สรุปผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 106

4-14 เกณฑ์ประเมินระดับความสามารถของผู้สอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 118

4-15 เกณฑ์ประเมินระดับความสามารถของผู้สอบ สำหรับการทดสอบด้วยคอมพิวเตอร์..... 119

4-16 ผลการประเมินโปรแกรมด้านความสะดวกในการใช้โปรแกรมโดยผู้เชี่ยวชาญ ..... 119

4-17 ผลการประเมินโปรแกรมด้านความถูกต้องในการใช้งานโดยผู้เชี่ยวชาญ..... 120

4-18 ผลการประเมินโปรแกรมด้านลักษณะทั่วไปของโปรแกรมโดยผู้เชี่ยวชาญ ..... 120

4-19 ผลการประเมินคู่มือการใช้โปรแกรมโดยผู้เชี่ยวชาญ ..... 121

4-20 ผลการประเมินโปรแกรมด้านความสะดวกในการใช้โปรแกรมโดยผู้สอบ..... 122

4-21 ผลการประเมินโปรแกรมด้านลักษณะทั่วไปของโปรแกรมโดยผู้สอบ..... 122

4-22 ความยาวของแบบทดสอบ เวลาที่ใช้ในการทดสอบ และผลการทดสอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์ ..... 123

4-23 ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างค่าประมาณความสามารถของผู้สอบกับคะแนนรวมของผู้สอบ ..... 125

ข-1 ค่าความสามารถที่แท้จริงของผู้สอบชุดที่ 1 ค่าประมาณความสามารถของผู้สอบ และความยาวของแบบทดสอบของการศึกษาในสถานการณ์จำลอง การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 147

ค-1 ค่าพารามิเตอร์ จำนวนการใช้ข้อสอบ และอัตราการใช้ข้อสอบ ของการศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ..... 174

# สารบัญภาพ


| ภาพที่ | | หน้า |
|---|---|---|
| 1-1 | กรอบแนวคิดการวิจัย | 8 |
| 2-1 | กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ | 20 |
| 2-2 | โค้งคุณลักษณะของข้อสอบแบบ 1 พารามิเตอร์ | 38 |
| 2-3 | โค้งคุณลักษณะของข้อสอบแบบ 2 พารามิเตอร์ | 39 |
| 2-4 | โค้งคุณลักษณะของข้อสอบแบบ 3 พารามิเตอร์ | 40 |
| 2-5 | กรอบแนวคิดการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีการสะท้อนข้อมูลย้อนกลับ | 57 |
| 3-1 | ขั้นตอนการดำเนินการวิจัย | 62 |
| 3-2 | ขั้นตอนการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป | 63 |
| 3-3 | ตัวอย่างขั้นตอนการควบคุมการใช้ข้อสอบโดยใช้วิธีการสุ่มตัวอย่างแบบมีระบบ | 67 |
| 3-4 | ขั้นตอนการศึกษาในสถานการณ์จำลองเพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป | 71 |
| 3-5 | ขั้นตอนพัฒนาโปรแกรมการทดสอบ | 78 |
| 3-6 | แผนผังบริบทของโปรแกรมการทดสอบ | 79 |
| 3-7 | แผนผังการไหลของข้อมูลระดับ 0 | 80 |
| 3-8 | แผนผังการไหลของข้อมูลระดับ 1 ของกระบวนการจัดการคลังข้อสอบ | 81 |
| 3-9 | แผนผังการไหลของข้อมูลระดับ 1 ของกระบวนการจัดการทดสอบด้วยโปรแกรม | 82 |
| 3-10 | ความสัมพันธ์ระหว่างข้อมูล | 86 |
| 3-11 | โครงสร้างหน้าจอทั่วไปของโปรแกรมการทดสอบ | 87 |
| 3-12 | โครงสร้างหน้าจอการกรอกข้อมูลเบื้องต้นของผู้สอบ | 87 |
| 3-13 | โครงสร้างหน้าจอการทดสอบ | 87 |
| 3-14 | โครงสร้างหน้าจอการรายงานผลการทดสอบ | 87 |
| 3-15 | โครงสร้างหน้าจอ Login | 88 |
| 3-16 | โครงสร้างหน้าจอการจัดการข้อสอบ | 88 |
| 3-17 | โครงสร้างหน้าจอรายงานผลการจัดการข้อสอบ | 88 |
| 3-18 | ขั้นตอนการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ | 91 |
| 4-1 | หน้าหลักของโปรแกรมการทดสอบ | 108 |
| 4-2 | หน้าจอ Login เข้าโปรแกรมในส่วนการจัดการข้อสอบ | 109 |
| 4-3 | หน้าจอการตั้งรหัสกลุ่มผู้สอบ | 110 |
| 4-4 | หน้าจอการจัดการคลังข้อสอบ | 110 |

# สารบัญภาพ (ต่อ)


| ภาพที่ | | หน้า |
|---|---|---|
| 4-5 | หน้าจอการเพิ่มข้อสอบ | 111 |
| 4-6 | หน้าจอแก้ไขข้อสอบ | 112 |
| 4-7 | หน้าจอการจัดการข้อสอบของการทดสอบด้วยคอมพิวเตอร์ | 113 |
| 4-8 | หน้าจอแรกในส่วนรายงานผลสอบ | 114 |
| 4-9 | หน้าจอรายงานผลสอบ | 114 |
| 4-10 | หน้าจอเริ่มต้นการทดสอบ | 115 |
| 4-11 | หน้าจอกรอกข้อมูลผู้สอบและเลือกแบบทดสอบ | 116 |
| 4-12 | หน้าจอข้อสอบ | 117 |
| 4-13 | หน้าจอรายงานผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ | 117 |
| 4-14 | หน้าจอรายงานผลการทดสอบด้วยคอมพิวเตอร์ | 118 |
| ช-1 | ตัวอย่างรายงานผลสอบของผู้สอบทั้งหมด | 214 |
| ช-2 | ตัวอย่างผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ | 214 |
| ช-3 | ตัวอย่างผลการทดสอบด้วยคอมพิวเตอร์ | 215 |

# บทที่ 1
# บทนำ

## ความเป็นมาและความสำคัญของปัญหา

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นการทดสอบที่มีการจัดข้อสอบให้เหมาะสมกับความสามารถของผู้สอบ ซึ่งพิจารณาจากผลการตอบข้อสอบในทุกข้อที่ผ่านมา กล่าวคือ เมื่อผู้สอบทำข้อสอบข้อแรกหรือชุดเริ่มต้นเสร็จลง จะนำผลการตอบมาวิเคราะห์ระดับความสามารถของผู้สอบ เพื่อใช้คัดเลือกข้อสอบข้อถัดไปที่มีความเหมาะสมกับความสามารถของผู้สอบ โดยการวิเคราะห์นี้จะใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เป็นพื้นฐาน การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในช่วงเริ่มต้นจะเป็นการทดสอบแบบสองขั้นตอน (Two-Stage Strategies) โดยขั้นตอนแรกเป็นการทดสอบแบบแยกทาง (Routing Test) ผลการทดสอบในขั้นนี้จะนำไปใช้เลือกระดับความยากของชุดข้อสอบในขั้นถัดไป ขั้นตอนที่ 2 การทดสอบแบบวัดผล (Measurement Test) เพื่อวัดความสามารถของผู้สอบ ซึ่งผู้สอบแต่ละคนจะได้ชุดข้อสอบแตกต่างกันขึ้นอยู่กับระดับความยากที่ประเมินได้จากขั้นตอนแรก ต่อมาเมื่อคอมพิวเตอร์มีประสิทธิภาพมากยิ่งขึ้น จึงมีการพัฒนาการทดสอบแบบหลายขั้นตอน (Multi-Stage Strategies) ซึ่งแบ่งได้ 2 แบบ คือ 1) การทดสอบแบบทางแยกคงที่ (Fixed-Branching Model) เป็นการทดสอบที่ผู้สอบจะได้ข้อสอบแตกต่างกัน ขึ้นอยู่กับผลการสอบในทุกข้อที่ผ่านมา และมีการจัดรูปแบบหรือเส้นทางในการทดสอบอย่างแน่นอนไว้แล้ว เช่น การทดสอบแบบพีระมิด (Pyramid Model) การทดสอบแบบยืดหยุ่น (Flexilevel Model) และ 2) การทดสอบแบบทางแยกแปรผัน (Variable Branching Model) มีลักษณะคล้ายกับแบบทางแยกคงที่ แต่ไม่มีการจัดรูปแบบหรือเส้นทางในการทดสอบอย่างแน่นอน การคัดเลือกข้อสอบข้อถัดไปจะใช้ผลการประเมินความสามารถของผู้สอบที่ได้จากผลการตอบข้อสอบในทุกข้อที่ผ่านมา การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในปัจจุบันส่วนใหญ่ จะเป็นการทดสอบแบบทางแยกแปรผัน (de Ayala, 2009, pp. 373-381; Wainer, Dorans, Flaugher, Green, & Mislevy, 2000, pp. 105-112)

ขั้นตอนการดำเนินการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีทั้งหมด 5 ขั้นตอน คือ ขั้นตอนที่ 1 การสร้างคลังข้อสอบ (Create Item Bank) เป็นขั้นตอนในการจัดทำข้อสอบ เพื่อใช้ในการทดสอบ โดยข้อสอบที่จัดทำขึ้นนี้ต้องเป็นไปตามทฤษฎีการตอบสนองข้อสอบ ขั้นตอนที่ 2 การคัดเลือกข้อสอบข้อแรก (First Item Selection) เป็นการเลือกข้อสอบที่มีความยากระดับปานกลางให้ผู้สอบ แล้วจะนำผลสอบที่ได้ไปประมาณค่าความสามารถของผู้สอบ เพื่อใช้ในการดำเนินการสอบขั้นต่อไป ขั้นตอนที่ 3 การคัดเลือกข้อสอบข้อถัดไป (Next Item Selection) เป็นการคัดเลือกข้อสอบที่มีความสอดคล้องหรือเหมาะสมกับระดับความสามารถของผู้สอบ โดยอาศัยผลการตอบข้อสอบในทุกข้อที่ผ่านมา ซึ่งมีหลายวิธีการ เช่น วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด วิธีการคัดเลือกข้อสอบโดยใช้ค่าสารสนเทศฟิชเชอร์สูงสุด ขั้นตอนที่ 4 การประมาณค่าความสามารถของผู้สอบ (Calculate Possible Ability Level) เป็นการคำนวณค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น โดยอาศัยผลการตอบข้อสอบในทุกข้อที่ผ่านมา ซึ่งมี

หลายวิธี เช่น วิธีความน่าจะเป็นแบบมีเงื่อนไข วิธีเบย์แบบปรับใหม่ และขั้นตอนที่ 5 เกณฑ์ยุติการทดสอบ (Termination Criterion) เป็นการกำหนดเกณฑ์สิ้นสุดของการทดสอบ โดยการทดสอบจะดำเนินการซ้ำตั้งแต่ขั้นตอนที่ 3 ถึงขั้นตอนที่ 5 จนกระทั่งการทดสอบเป็นไปตามเกณฑ์ยุติการทดสอบที่กำหนดไว้ ซึ่งเกณฑ์มีหลายเกณฑ์ เช่น ผู้สอบทำข้อสอบได้ครบตามจำนวนข้อที่กำหนด หรือผู้สอบทำการทดสอบจนค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถมีค่าต่ำกว่าเกณฑ์ที่กำหนด (Thompson & Weiss, 2011)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่กล่าวไว้ข้างต้น ขั้นตอนการคัดเลือกข้อสอบข้อถัดไป นับเป็นขั้นตอนที่มีความสำคัญ เนื่องจากการคัดเลือกข้อสอบข้อถัดไปที่มีความเหมาะสมกับความสามารถของผู้สอบ จะช่วยลดความยาวของแบบทดสอบ และเพิ่มประสิทธิภาพการประมาณค่าความสามารถของผู้สอบ (Luecht & Sireci, 2011, pp. 9-10; Lunz, Bergstrom, & Wright, 1992) หลักการทั่วไปของวิธีการคัดเลือกข้อสอบข้อถัดไป จะใช้ผลการตอบข้อสอบในข้อที่ผ่านมา มาใช้ในการคัดเลือกข้อสอบข้อถัดไป กล่าวคือ ถ้าตอบถูก ผู้สอบจะได้ข้อสอบข้อถัดไปที่มีค่าความยากของข้อสอบเพิ่มขึ้น ในทางกลับกัน ถ้าตอบผิด ผู้สอบจะได้ข้อสอบที่มีค่าความยากของข้อสอบลดลง แต่ทั้งนี้วิธีการคัดเลือกข้อสอบข้อถัดไปแต่ละวิธีจะมีรายละเอียดแตกต่างกันขึ้นอยู่กับผู้พัฒนา เช่น วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) ซึ่งวิธีการนี้จะนำค่าพารามิเตอร์ของข้อสอบทั้ง 3 ค่า ได้แก่ ค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบ และค่าการเดาของข้อสอบ มาพิจารณาร่วมกับค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น โดยจะเลือกข้อสอบที่มีค่าสารสนเทศสูงสุด ณ ระดับค่าความสามารถของผู้สอบที่ตรงหรือใกล้เคียงที่สุด เป็นข้อสอบข้อถัดไป (Birnbaum, 1968 cited in van der Linden & Glas, 2002, pp. 9-10) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้สารสนเทศฟิชเชอร์สูงสุด (Maximum Fisher Information: MFI) เสนอโดย Weiss (1982) วิธีการนี้จะนำค่าพารามิเตอร์ของข้อสอบมาร่วมพิจารณาด้วย คล้ายกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด แต่จะคัดเลือกข้อสอบที่มีค่าสารสนเทศฟิชเชอร์สูงสุดมาเป็นข้อสอบข้อถัดไป วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยงของ โสพส สุขานนท์ สวัสดิ์, เสรี ชัดแช้ม และกฤษณะ ชินสาร (2556) เป็นวิธีการที่นำทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (Decision Making Under Risk) ตามเกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุด (Maximum Expected Monetary Value: EMV) โดยจะเลือกข้อสอบข้อที่มีค่าสัมประสิทธิ์การกระจาย (Coefficient of Variation) สูงสุดมาเป็นข้อสอบข้อถัดไป

ทฤษฎีการตัดสินใจ (Decision Theory) เป็นทฤษฎีหนึ่งที่ใช้เป็นเครื่องมือในการตัดสินใจเพื่อให้ผู้ตัดสินใจมีโอกาสตัดสินใจผิดพลาดน้อยลง หรือตัดสินใจได้ถูกต้องมากขึ้นตามสภาวการณ์ต่างกัน ซึ่งทฤษฎีการตัดสินใจแบ่งสภาวการณ์ที่ต้องตัดสินใจได้ทั้งหมด 4 ประเภท คือ 1) การตัดสินใจในสภาวการณ์ที่แน่นอน (Certainly Decision Making) เป็นการตัดสินใจที่ผู้ตัดสินใจมีข้อมูลเกี่ยวกับเรื่องที่ต้องการตัดสินใจอยู่ครบถ้วน และทราบถึงผลของเหตุการณ์ที่จะเกิดขึ้นได้เป็นอย่างดี 2) การตัดสินใจในสภาวการณ์ที่ไม่แน่นอน (Uncertainly Decision Making) เป็นการตัดสินใจที่ผู้ตัดสินใจมีข้อมูลเกี่ยวกับเรื่องที่จะตัดสินใจไม่เพียงพอ ทำให้ไม่ทราบผลของเหตุการณ์ที่จะเกิดขึ้น และไม่สามารถคาดคะเนหรือกำหนดค่าความน่าจะเป็นของแต่ละเหตุการณ์ได้

3) การตัดสินใจในสภาวการณ์ความเสี่ยง (Risk Decision Making) เป็นการตัดสินใจที่ผู้ตัดสินใจมีข้อมูลเกี่ยวกับเรื่องที่จะตัดสินใจไม่เพียงพอ ทำให้ไม่ทราบผลลัพธ์ของเหตุการณ์ที่จะเกิดขึ้น แต่สามารถคาดคะเนหรือกำหนดค่าความน่าจะเป็นในแต่ละเหตุการณ์ได้ และ 4) การตัดสินใจในสภาวการณ์การแข่งขัน (Competition Decision Making) เป็นการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ประกอบด้วยผู้แข่งขันตั้งแต่ 2 ฝ่ายขึ้นไป แต่ละฝ่ายจะนำกลยุทธ์การแข่งขันมาใช้เพื่อแสวงหาผลประโยชน์ให้กับตนเอง และคอยขัดขวางผลประโยชน์ของฝ่ายตรงข้าม (กิตติ ภักดีวัฒนะกุล และพนิดา พานิชกุล, 2554, หน้า 541-554; Taylor, 2009, pp. 539-572)

การตัดสินใจในสภาวการณ์ที่ไม่แน่นอนเป็นการตัดสินใจที่ผู้ตัดสินใจมีข้อมูลไม่เพียงพอ ทำให้ไม่ทราบหรือไม่สามารถคาดคะเนผลลัพธ์ของเหตุการณ์นั้น ๆ ได้ หลักเกณฑ์ที่นำมาใช้ตัดสินใจในสภาวการณ์นี้ ส่วนใหญ่จึงเป็นดุลยพินิจและประสบการณ์ที่ผ่านมาในอดีตของผู้ตัดสินใจ ซึ่งมีหลายหลักเกณฑ์ เช่น เกณฑ์แมกซิแมกซ์ (Maximax Criterion) เป็นเกณฑ์ที่ผู้ตัดสินใจมองโลกในแง่ดี ผู้ตัดสินใจจะเลือกทางเลือกที่ให้ผลตอบแทนสูงสุด โดยคาดว่าทุกทางเลือกจะให้ผลตอบแทนในระดับที่สูงที่สุดแก่ตนเอง ซึ่งเป็นการพิจารณาในกรณีที่ดีที่สุดที่อาจจะเกิดขึ้นได้ เกณฑ์แมกซิมิน (Maximin Criterion) เป็นเกณฑ์ที่ผู้ตัดสินใจมองโลกในแง่ร้าย ผู้ตัดสินใจจะเลือกทางเลือกที่ให้ผลตอบแทนสูงสุด โดยคาดว่าทุกทางเลือกจะให้ผลตอบแทนในระดับที่ต่ำที่สุดแก่ตนเอง ซึ่งเป็นการพิจารณาในกรณีที่เลวร้ายที่สุดที่อาจจะเกิดขึ้นได้ เกณฑ์ของลาปลาซ (Laplace Criterion) เป็นเกณฑ์ที่ผู้ตัดสินใจจะเลือกทางเลือกที่สูงที่สุด โดยคิดว่าทุกทางเลือกจะให้ผลตอบแทนในระดับปานกลาง ซึ่งจะพิจารณาจากค่าเฉลี่ยของผลตอบแทนของทางเลือกนั้น ๆ แล้วเลือกทางเลือกที่มีเฉลี่ยสูงที่สุด และเกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion) เป็นเกณฑ์ที่ผสมผสานระหว่างเกณฑ์แมกซิแมกซ์และเกณฑ์แมกซิมิน ซึ่งเปิดโอกาสให้ผู้ตัดสินใจระบุระดับของผลตอบแทนที่ผู้ตัดสินใจคิดว่าเหมาะสม โดยระบุจากค่าสัมประสิทธิ์ของการมองโลกในแง่ดี (Coefficient of Optimism: $\alpha$) ซึ่งมีค่าตั้งแต่ 0 ถึง 1 ถ้ากำหนด $\alpha$ มีค่าใกล้เคียงหรือเท่ากับ 1 แสดงว่า ผู้ตัดสินใจมองโลกในแง่ดี ผู้ตัดสินใจคาดว่าทุกทางเลือกจะให้ผลตอบแทนในระดับสูงสุด ผลการตัดสินใจจะเหมือนกับการใช้เกณฑ์แมกซิแมกซ์ ในทางกลับกัน ถ้ากำหนด $\alpha$ มีค่าใกล้เคียงหรือเท่ากับ 0 แสดงว่า ผู้ตัดสินใจมองโลกในแง่ร้าย ผู้ตัดสินใจคาดว่าทุกทางเลือกจะให้ผลตอบแทนในระดับต่ำสุด ผลการตัดสินใจจะเหมือนกับการใช้เกณฑ์แมกซิมิน เกณฑ์ของเฮอร์วิคซ์นี้ จะพิจารณาจากค่าเฉลี่ยถ่วงน้ำหนักระหว่างผลตอบแทนที่มีระดับสูงสุดกับระดับต่ำสุด โดยใช้ค่า $\alpha$ เป็นค่าถ่วงน้ำหนัก แล้วจะตัดสินใจเลือกทางเลือกที่มีค่าเฉลี่ยถ่วงน้ำหนักสูงสุด (กิตติ ภักดีวัฒนะกุล และพนิดา พานิชกุล, 2554, หน้า 541-545; Taylor, 2009, pp. 542-543)

ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ตามเกณฑ์ของเฮอร์วิคซ์ข้างต้น สามารถนำมาประยุกต์กับวิธีการคัดเลือกข้อสอบข้อถัดไปให้มีความเหมาะสมกับระดับความสามารถของผู้สอบได้ โดยตั้งเงื่อนไขการคัดเลือกข้อสอบข้อถัดไปให้มีความสอดคล้องกับเกณฑ์ของเฮอร์วิคซ์ ซึ่งกำหนดให้ทางเลือกที่ใช้ในการตัดสินใจ คือ ข้อสอบในคลังข้อสอบทุกข้อที่ยังไม่ถูกเลือกมาใช้ในการทดสอบ ผลตอบแทนในแต่ละทางเลือก คือ ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ (Standard Error of Estimate: SEE) และกำหนดค่าสัมประสิทธิ์ของการมองโลกในแง่ดี ($\alpha$) โดยอาศัยหลักการของข้อสอบที่มีความเหมาะสมกับผู้สอบ คือ ข้อสอบที่มี

ค่าความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้สอบ นั่นคือ หากข้อสอบข้อใดมีค่าความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้สอบ จะกำหนดค่า $\alpha$ ให้มีค่ามาก (มองโลกในแง่ดี เพราะค่า SEE จะมีค่าน้อยลง เนื่องจากได้รับข้อสอบที่มีความเหมาะสมกับผู้สอบ) ในทางกลับกัน หากข้อสอบข้อใดมีค่าความยากของข้อสอบแตกต่างจากค่าความสามารถของผู้สอบมาก จะกำหนดค่า $\alpha$ ให้มีค่าน้อย (มองโลกในแง่ร้าย เพราะค่า SEE จะมีค่ามากขึ้น เนื่องจากได้รับข้อสอบที่ไม่เหมาะสมกับผู้สอบ) จากนั้นจะคำนวณค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE โดยใช้ค่า SEE ในกรณีที่คาดว่าผู้สอบจะตอบข้อสอบถูกและกรณีที่คาดว่าผู้สอบจะตอบข้อสอบผิด และใช้ค่า $\alpha$ เป็นค่าถ่วงน้ำหนัก จากเกณฑ์ของเฮอร์วิคซ์จะเลือกทางเลือกที่มีค่าเฉลี่ยถ่วงน้ำหนักสูงสุด ซึ่งในที่นี้คือ ข้อสอบที่มีค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE ที่ต่ำสุด มาเป็นข้อสอบข้อถัดไป ดังนั้น เกณฑ์ของเฮอร์วิคซ์จึงเป็นเกณฑ์ที่สามารถนำมาประยุกต์กับการคัดเลือกข้อสอบข้อถัดไปได้ โดยข้อสอบที่ได้รับเลือกจากการใช้เกณฑ์นี้จะมีความเหมาะสมกับระดับความสามารถของผู้สอบ เพราะเป็นข้อสอบที่มีค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบโดยเฉลี่ยต่ำสุด

วิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จะคัดเลือกข้อสอบที่มีค่าสารสนเทศของข้อสอบสูงสุด ณ ระดับความสามารถของผู้สอบขณะนั้น ซึ่งข้อสอบที่มีค่าสารสนเทศของข้อสอบสูง คือ ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง และมีค่าความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้สอบ ณ ขณะนั้น เป็นผลให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงถูกเลือกใช้ในการทดสอบบ่อยครั้งเกินไป ส่งผลให้ผู้สอบในรุ่นถัดมาสามารถคาดการณ์ล่วงหน้าได้ว่า จะได้ข้อสอบข้อใดในการทดสอบ (Chang & van der Linden, 2003; Wainer et al., 2000, pp. 119-120) ด้วยเหตุนี้ จึงมีผู้เสนอวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีกระบวนการในการควบคุมความถี่ในการใช้ข้อสอบแต่ละข้อให้มีความเหมาะสม ซึ่งเรียกว่า การควบคุมการใช้ข้อสอบ (Item Exposure Control) โดย Georgiadou, Triantafillou, and Economides (2007) ได้จำแนกวิธีการควบคุมการใช้ข้อสอบออกเป็น 5 ประเภท ได้แก่ 1) วิธีการคัดเลือกอย่างสุ่ม (Randomization Strategies) วิธีการในกลุ่มนี้จะอาศัยการสุ่มข้อสอบในกลุ่มข้อสอบที่สนใจ เช่น กลุ่มข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง หรือกลุ่มข้อสอบที่มีค่าความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้สอบมาเป็นพื้นฐานในการพัฒนา วิธีการในกลุ่มนี้ เช่น วิธีการ 5-4-3-2-1 (McBride & Martin, 1983) วิธีการ Within 0.1 Logits (Lunz & Stahl, 1998) 2) วิธีการคัดเลือกอย่างมีเงื่อนไข (Conditional Selection Strategies) วิธีการในกลุ่มนี้จะควบคุมการใช้ข้อสอบด้วยค่าพารามิเตอร์ควบคุมการใช้ข้อสอบ (Exposure Control Parameter) ซึ่งเป็นค่าความน่าจะเป็นของข้อสอบที่ได้รับเลือกจะถูกจัดให้แก่ผู้สอบ มีค่าตั้งแต่ 0 ถึง 1 ถ้าค่านี้เข้าใกล้ 1 แสดงว่า ข้อสอบข้อที่ได้รับเลือกนั้นมีโอกาสสูงที่จะถูกจัดให้แก่ผู้สอบ วิธีการในกลุ่มนี้ได้แก่ วิธีการซิมสันเฮกเตอร์ (Sympson-Hetter: SH) เสนอโดย Sympson and Hetter (1985) 3) วิธีการคัดเลือกตามระดับชั้น (Stratified Strategies) วิธีการในกลุ่มนี้จะควบคุมการใช้ข้อสอบด้วยการแบ่งคลังข้อสอบออกเป็นชั้น ๆ แล้วบังคับให้เลือกข้อสอบจากทุกชั้นมาใช้ในการทดสอบ วิธีการในกลุ่มนี้ได้แก่ วิธีการ a-Stratified Strategy (ASTR) เสนอโดย Chang and Ying (1999) 4) วิธีการคัดเลือกแบบผสม (Combined Strategies) วิธีการในกลุ่มนี้จะนำวิธีการควบคุมการใช้ข้อสอบตั้งแต่ 2 วิธีการขึ้นไปมาใช้ร่วมกัน วิธีการในกลุ่มนี้ได้แก่ วิธี Progressive Restricted เสนอโดย Revuelta

and Ponsoda (1998) และ 5) วิธีการคัดเลือกแบบปรับเหมาะหลายขั้นตอน (Multiple Stage Adaptive Testing Designs) วิธีการในกลุ่มนี้จะดำเนินการร่วมกันระหว่างการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบแบบเขียนตอบบนกระดาษ โดยใช้ผลจากการทดสอบแบบเขียนตอบบนกระดาษมาเป็นข้อมูลในการคัดเลือกข้อสอบข้อถัดไป แล้วหาค่าความสามารถของผู้สอบด้วยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ วิธีการในกลุ่มนี้ได้แก่ วิธี Computerize Adaptive Sequential Testing (CAST) ซึ่งเสนอโดย Luecht, Nungester, and Hadadi (1996)

วิธีการควบคุมการใช้ข้อสอบ ในกลุ่มวิธีการคัดเลือกอย่างสุ่ม (Randomization Strategies) ที่กล่าวข้างต้น จะอาศัย “การสุ่ม” เพื่อควบคุมการใช้ข้อสอบในกลุ่มที่สนใจ (กลุ่มข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง) ไม่ให้ข้อสอบถูกเลือกใช้ในการทดสอบบ่อยครั้งเกินไป ตามแนวคิดนี้จึงนำวิธีการสุ่มตัวอย่างแบบมีระบบ (Systematic Random Sampling) ซึ่งเป็นวิธีการสุ่มตัวอย่างวิธีหนึ่ง มาประยุกต์กับวิธีการควบคุมการใช้ข้อสอบ โดยจะคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงบางข้อ ด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ ไปใช้ในการคัดเลือกข้อสอบข้อถัดไป ซึ่งการสุ่มข้อสอบนี้จะดำเนินการทุกครั้งในขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ทำให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงถูกหมุนเวียนนำไปใช้ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไป วิธีการนี้จึงช่วยสกัดไม่ให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงข้อใดข้อหนึ่ง เข้าสู่ขั้นตอนการคัดเลือกข้อสอบข้อถัดไปได้ทุกครั้ง เป็นผลให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงจะถูกเลือกใช้ในการทดสอบลดลง

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวกับวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่ายังไม่พบการนำทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ตามเกณฑ์ของเฮอร์วิคซ์ ที่พิจารณาค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถที่ต่ำสุด และนำค่าความแตกต่างระหว่างค่าความยากของข้อสอบกับค่าประมาณความสามารถของผู้สอบมาเป็นค่าสัมประสิทธิ์ของการมองโลกในแง่ดี มาใช้กับวิธีการคัดเลือกข้อสอบข้อถัดไป ซึ่งการดำเนินการตามนี้ ข้อสอบข้อถัดไปที่ได้รับ จะเป็นข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบ เนื่องจากค่าประมาณความสามารถที่คำนวณได้จากผลการตอบของข้อสอบที่ได้รับนี้ จะมีความคลาดเคลื่อนจากค่าความสามารถที่แท้จริงของผู้สอบโดยเฉลี่ยต่ำสุด แต่ทั้งนี้ พบว่างานวิจัยที่ใกล้เคียงสุดเป็นงานวิจัยของ โสพส สุขานนท์สวัสดิ์ และคณะ (2556) ได้นำทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยงมาประยุกต์กับวิธีการคัดเลือกข้อสอบข้อถัดไป ซึ่งมีการคำนวณที่ซับซ้อนมากกว่าเมื่อเทียบกับการนำเกณฑ์ของเฮอร์วิคซ์มาประยุกต์กับวิธีการคัดเลือกข้อสอบข้อถัดไป เนื่องจาก ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยงต้องอาศัยค่าความน่าจะเป็นในการคำนวณ จึงมีความซับซ้อนยุ่งยากมากกว่าการใช้เกณฑ์ของเฮอร์วิคซ์ที่พิจารณาจากค่าเฉลี่ยถ่วงน้ำหนักเท่านั้น

วิธีการคัดเลือกอย่างสุ่มเป็นกลุ่มวิธีที่มีการดำเนินการซับซ้อนน้อยกว่า เมื่อเทียบกับวิธีการในกลุ่มอื่นๆ เนื่องจากใช้เพียงการสุ่มข้อสอบจากกลุ่มข้อสอบที่สนใจเท่านั้น วิธีการในกลุ่มนี้จึงง่ายต่อการนำไปใช้ควบคุมการใช้ข้อสอบ แต่มีข้อจำกัดตรงที่ไม่รับรองว่า จะควบคุมการใช้ข้อสอบได้ตามที่ต้องการหรือไม่ เนื่องจากวิธีการในกลุ่มนี้จะสุ่มข้อสอบจากกลุ่มข้อสอบที่สนใจ ซึ่งมีข้อสอบจำนวนน้อย เช่น ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง 5 อันดับแรก หรือข้อสอบที่มีค่าความยากของ

ข้อสอบต่างจากค่าความสามารถของผู้สอบไม่เกิน 0.1 จึงทำให้อาจมีข้อสอบบางข้อยังถูกเลือกไปใช้ทดสอบมากเกินไป ด้วยเหตุนี้จึงได้นำวิธีการสุ่มตัวอย่างแบบมีระบบมาใช้กับวิธีการควบคุมการใช้ข้อสอบ โดยการสุ่มข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ ทำให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงมีโอกาสถูกเลือกไปใช้เท่า ๆ กัน นอกจากข้อสอบที่สุ่มได้นี้ ยังนำข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบต่ำถึงปานกลางมาใช้ในการคัดเลือกข้อสอบข้อถัดไปด้วย ทำให้ข้อสอบที่ใช้ในการคัดเลือกข้อสอบข้อถัดไปมีจำนวนเพิ่มมากขึ้น โอกาสที่จะเลือกข้อสอบซ้ำกันจึงลดลง

จากเหตุผลดังกล่าวข้างต้น ทำให้ผู้วิจัยสนใจที่จะพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอนตามเกณฑ์ของเฮอร์วิค และมีการควบคุมการใช้ข้อสอบด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ เพื่อให้ได้วิธีการคัดเลือกข้อสอบข้อถัดไปที่มีประสิทธิภาพสูงขึ้น ทั้งในด้านการประมาณค่าความสามารถของผู้สอบ ด้านความยาวของแบบทดสอบ และด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 แล้วนำวิธีการคัดเลือกข้อสอบข้อถัดไปและมีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้น ไปพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวัดความสามารถทางภาษาอังกฤษ ของนักศึกษาชั้นปี 2 คณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร ต่อไป

## วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
2. เพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 4 วิธี คือ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง 3) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ และ 4) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ โดยเปรียบเทียบประสิทธิภาพด้าน 1) การประมาณค่าความสามารถของผู้สอบ 2) ความยาวของแบบทดสอบ และ 3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2
3. เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ
4. เพื่อศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้นกับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์

## กรอบแนวคิดการวิจัย

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีแนวคิดในการพัฒนาดังนี้

ผู้วิจัยได้ศึกษาทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน (Uncertainly Decision Making) ตามเกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion) แล้วนำมาประยุกต์ในวิธีการคัดเลือกข้อสอบ

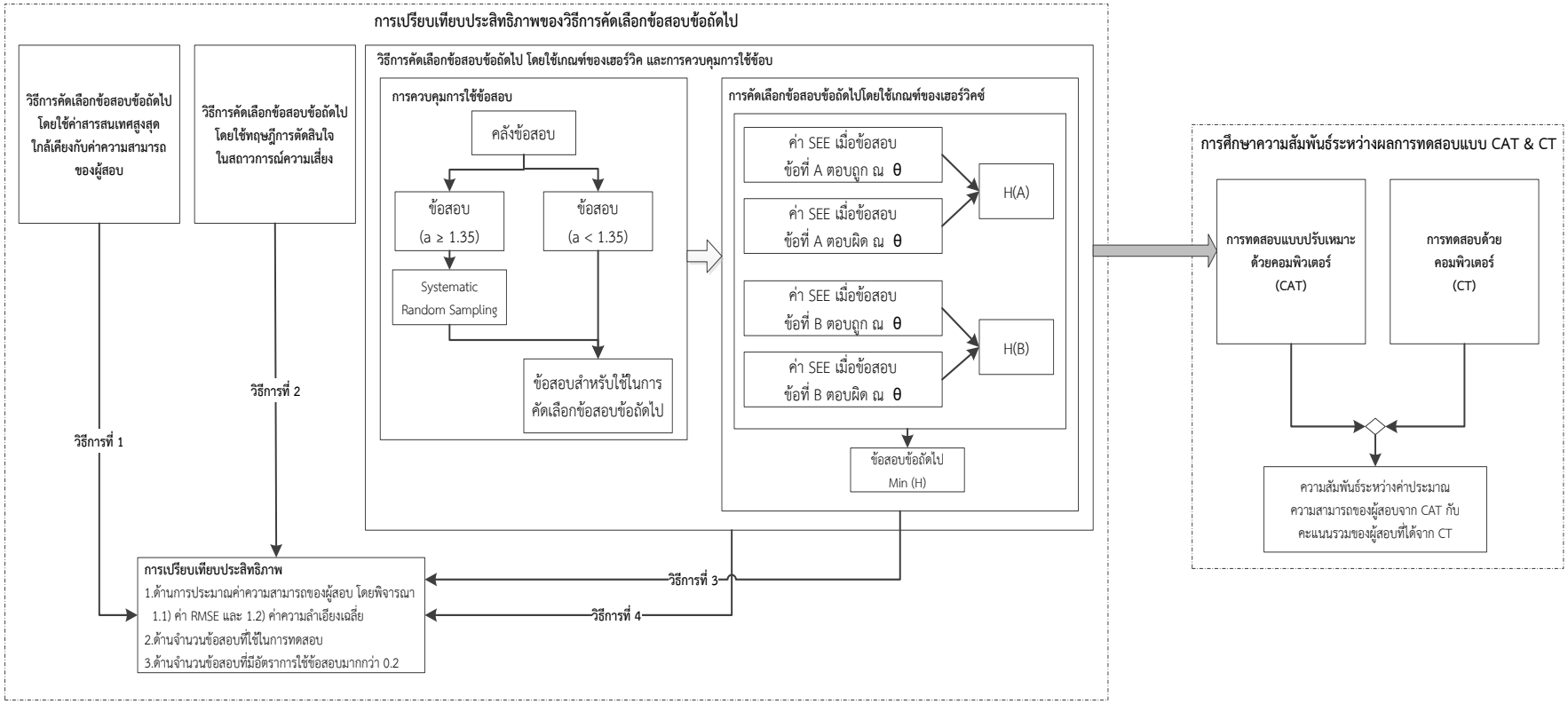
ข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยคัดเลือกข้อสอบในคลังข้อสอบที่ยังไม่เคยถูกใช้ในการทดสอบ และเป็นข้อสอบที่มีค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ (Standard Error of Estimate: SEE) ต่ำสุดมาเป็นข้อสอบข้อถัดไป ซึ่งการคำนวณค่าเฉลี่ยถ่วงน้ำหนักนี้จะอาศัยค่า SEE ทั้งกรณีที่คาดว่าผู้สอบตอบข้อสอบถูก และกรณีที่คาดว่าผู้สอบตอบข้อสอบผิด และใช้ผลต่างระหว่างค่าความยากของข้อสอบกับค่าประมาณความสามารถของผู้สอบขณะนั้นมาเป็นค่าถ่วงน้ำหนัก ดังนั้น ข้อสอบที่ได้รับเลือกจากวิธีการนี้จึงเป็นข้อสอบที่ให้ค่า SEE โดยเฉลี่ยต่ำสุด

วิธีการควบคุมการใช้ข้อสอบในการวิจัยนี้ ได้พัฒนาโดยนำแนวคิดของวิธีการควบคุมการใช้ข้อสอบในกลุ่มวิธีการคัดเลือกอย่างสุ่ม (Randomization Strategies) มาพัฒนาร่วมกับวิธีการสุ่มตัวอย่างแบบมีระบบ (Systematic Random Sampling) โดยในแต่ละครั้งก่อนถึงขั้นตอนการคัดเลือกข้อสอบข้อถัดไป จะสุ่มข้อสอบด้วยวิธีการสุ่มตัวอย่างแบบมีระบบกับข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง คือ มีค่าอำนาจจำแนกของข้อสอบตั้งแต่ 1.35 ขึ้นไป (Baker, 2001, p. 34) แล้วนำข้อสอบที่สุ่มได้นี้รวมกับข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบต่ำถึงปานกลาง (มีค่าอำนาจนำแนกของข้อสอบไม่เกิน 1.35) ไปใช้ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ซึ่งการสุ่มข้อสอบนี้จะต้องสุ่มใหม่ทุกครั้งก่อนการคัดเลือกข้อสอบข้อถัดไป ทำให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ถูกหมุนเวียนสลับกันเข้าสู่ขั้นตอนการคัดเลือกข้อสอบข้อถัดไป เป็นผลให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง มีโอกาสถูกเลือกใช้ซ้ำน้อยลง

การตรวจสอบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปและการควบคุมการใช้ข้อสอบที่ได้พัฒนาขึ้น ดำเนินการโดยเปรียบเทียบกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) ซึ่งเป็นวิธีการที่นิยมใช้ทั่วไป และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (โสพส สุขานนท์สวัสดิ์ และคณะ, 2556) ซึ่งเป็นวิธีการที่ใช้แนวคิดของทฤษฎีการตัดสินใจคล้ายคลึงกับวิธีการที่ผู้วิจัยพัฒนา มาเป็นฐานในการเปรียบเทียบ โดยนำมาเปรียบเทียบกับวิธีการที่ได้พัฒนาขึ้นทั้งในกรณีที่ไม่มีการควบคุมการใช้ข้อสอบ และกรณีที่มีการควบคุมการใช้ข้อสอบ ในด้าน 1) การประมาณค่าความสามารถของผู้สอบ 2) ความยาวของแบบทดสอบ และ 3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 (Doong, 2009)

จากนั้น นำวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้น ไปพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี และนำไปใช้ในการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้นกับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT)

จากที่กล่าวมาข้างต้น สามารถสรุปแนวคิดการวิจัยได้ ดังภาพที่ 1-1




ภาพที่ 1-1 กรอบแนวคิดการวิจัยเรื่อง การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

จากภาพที่ 1-1 แสดงกรอบแนวคิดการวิจัย ซึ่งมีสัญลักษณ์ที่ใช้ประกอบดังนี้

| | | |
|---|---|---|
| CAT | คือ | การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ |
| CT | คือ | การทดสอบด้วยคอมพิวเตอร์ |
| θ | คือ | ค่าความสามารถของผู้สอบ |
| a | คือ | ค่าอำนาจจำแนกของข้อสอบ |
| SEE | คือ | ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ |
| H(A) | คือ | ค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของข้อสอบข้อที่ A |
| Min(H) | คือ | ค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณความสามารถที่มีค่าต่ำสุด ซึ่งใช้คัดเลือกข้อสอบข้อถัดไป |
| RMSE | คือ | ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย |

## สมมติฐานของการวิจัย

1. ด้านการประมาณค่าความสามารถของผู้สอบ แบ่งได้ 2 ข้อ ดังนี้

1.1 ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์มีค่าน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

1.2 ค่าความลำเอียงเฉลี่ยของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์มีค่าน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

2. ด้านความยาวของแบบทดสอบของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์มีจำนวนน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

3. ด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ มีจำนวนน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์

4. ค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีความสัมพันธ์ทางบวกกับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์

## ขอบเขตของการวิจัย

การวิจัยนี้มุ่งพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ แล้วนำวิธีการคัดเลือก

ข้อสอบข้อถัดไปที่พัฒนาขึ้น ทั้งในกรณีที่มีและไม่มีการควบคุมการใช้ข้อสอบมาเปรียบเทียบประสิทธิภาพกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง จากนั้นพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปและมีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้น มาใช้วัดความรู้ในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนของนักศึกษาในระดับปริญญาตรี และสุดท้ายจะนำโปรแกรมการทดสอบที่พัฒนาขึ้นนี้มาศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ซึ่งมีขอบเขตในการวิจัยดังนี้

1. ด้านโปรแกรม ที่ใช้ในการวิจัยนี้มี 3 โปรแกรม ได้แก่

1.1 โปรแกรม MATLAB เป็นภาษาคอมพิวเตอร์ชั้นสูง ที่มีโครงสร้างของภาษาที่ไม่ซับซ้อน เข้าใจง่าย และมีฟังก์ชันทางคณิตศาสตร์และสถิติให้เลือกใช้ค่อนข้างครบถ้วน โปรแกรม MATLAB ได้นำมาใช้ในการเขียนโปรแกรมสำหรับศึกษาในสถานการณ์จำลองแบบมอนติคาร์โล (Monte Carlo Simulation Study) ซึ่งเป็นการศึกษาโดยการจำลองสถานการณ์เสมือนจริง ผลการศึกษาสามารถนำไปอ้างอิงกับการศึกษาในสถานการณ์จริงได้ (Harwell, Stone, Hsu, & Kirisci, 1996) ในที่นี้จะเขียนโปรแกรมเพื่อจำลองสถานการณ์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปที่แตกต่างกัน 4 วิธี แล้วนำผลที่ได้จากการศึกษามาเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ตามวัตถุประสงค์ของการวิจัยข้อที่ 2

1.2 โปรแกรม WinGen3 เป็นโปรแกรมสำเร็จรูปที่ใช้จำลองค่าความสามารถที่แท้จริงของผู้สอบ ($\theta$) และค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (Han, 2007) ข้อมูลที่จำลองด้วยโปรแกรม WinGen3 นี้จะใช้เป็นข้อมูลเริ่มต้นในการเขียนโปรแกรมด้วย MATLAB สำหรับจำลองสถานการณ์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยข้อมูลที่จำลองขึ้นมี 2 ส่วน ได้แก่

1.2.1 ค่าความสามารถที่แท้จริงของผู้สอบ ($\theta$) สร้างขึ้นโดยการสุ่มจากการแจงแจงแบบปกติมาตรฐาน ซึ่งมีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 โดยจะสุ่มค่าความสามารถที่แท้จริงของผู้สอบทั้งหมด 10,000 ค่า แบ่งเป็นชุด ๆ ละ 1,000 ค่า

1.2.2 ค่าพารามิเตอร์ของข้อสอบ ตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โดยจำลองข้อสอบทั้งหมด 500 ข้อ ซึ่งกำหนดการสุ่มพารามิเตอร์แต่ละตัว ดังต่อไปนี้

ก. ค่าอำนาจจำแนกของข้อสอบ (Discrimination Power Parameter: a) สุ่มจากการแจกแจงแบบยูนิฟอร์ม มีค่าอยู่ในช่วง 0.50 ถึง 2.50

ข. ค่าความยากของข้อสอบ (Difficulty Parameter: b) สุ่มจากการแจกแจงแบบยูนิฟอร์ม มีค่าอยู่ในช่วง -2.50 ถึง 2.50

ค. ค่าการเดาของข้อสอบ (Guess Parameter: c) สุ่มจากการแจกแจงแบบยูนิฟอร์ม มีค่าไม่เกิน 0.30

1.3 โปรแกรม Notepad++ เป็นโปรแกรมสำหรับพัฒนา Web Application รองรับการเขียนโปรแกรมด้วยภาษา PHP ร่วมกับการใช้ Tools ของ Cloud9 เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ตามวัตถุประสงค์ของการวิจัยข้อที่ 3 โดยโปรแกรม

การทดสอบที่พัฒนานี้ ออกแบบให้รองรับการทดสอบด้วยคอมพิวเตอร์ได้ เพื่อใช้ในการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ตามวัตถุประสงค์ของการวิจัยข้อที่ 4

2. ด้านประชากร ที่ใช้ในการวิจัยตามวัตถุประสงค์ของการวิจัยข้อที่ 4 คือ นักศึกษาชั้นปีที่ 2 คณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร ประจำปีการศึกษา 2559 จำนวนทั้งสิ้น 775 คน

3. ด้านตัวแปร ที่ใช้ในการวิจัยแบ่งตามวัตถุประสงค์ของการวิจัยได้ 2 ส่วน ดังนี้

ส่วนที่ 1 ตัวแปรตามวัตถุประสงค์ของการวิจัยข้อที่ 2 มีดังนี้

ตัวแปรต้น คือ วิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 4 วิธี ได้แก่

1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด

2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง

3) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์

4) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

ตัวแปรตาม คือ ประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มี 3 ด้าน ได้แก่

1) ด้านการประมาณค่าความสามารถของผู้สอบพิจารณาจากค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) (หน่วยวัดเป็นคะแนนมาตรฐาน)และค่าความลำเอียงเฉลี่ย (Average Bias) (หน่วยวัดเป็นคะแนนมาตรฐาน)

2) ด้านความยาวของแบบทดสอบ (หน่วยวัดเป็นจำนวนข้อ)

3) ด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 (หน่วยวัดเป็นจำนวนข้อ)

ส่วนที่ 2 ตัวแปรตามวัตถุประสงค์ของการวิจัยข้อที่ 4 มี 2 ตัวแปร ได้แก่

1) ค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (หน่วยวัดเป็นคะแนนมาตรฐาน)

2) คะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (หน่วยวัดเป็นคะแนน)

## ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. ได้วิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ ที่สามารถลดความยาวของแบบทดสอบ เมื่อเทียบกับการทดสอบด้วยคอมพิวเตอร์ จึงทำให้ประหยัดเวลาและแรงงาน รวมทั้งงบประมาณที่ใช้ในการจัดการทดสอบ

2. ได้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีการควบคุมการใช้ข้อสอบสำหรับใช้วัดความสามารถทางภาษาอังกฤษของนักศึกษา ในระดับปริญญาตรี

3. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สามารถนำไปใช้ทดสอบวัดความรู้ทางการศึกษาในรายวิชาใดก็ได้ เพียงแต่ข้อสอบที่ใช้ต้องมีลักษณะการตรวจให้คะแนน 2 ค่า คือ ถูก-ผิด หรือ 0-1 และมีรายการคำตอบอย่างน้อย 4 รายการ

4. แนวทางในการศึกษาเพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เช่น การคัดเลือกข้อสอบข้อแรก การคัดเลือกข้อสอบข้อถัดไป การประมาณค่าความสามารถของผู้สอบ และเกณฑ์ยุติการทดสอบ

## นิยามศัพท์เฉพาะ

การทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) หมายถึง กระบวนการทดสอบที่นำคอมพิวเตอร์มาใช้ในการทดสอบแทนการทดสอบบนกระดาษดินสอ ผู้สอบทุกคนจะได้แบบทดสอบชุดเดียวกัน จำนวน 40 ข้อ ซึ่งประกอบด้วยข้อสอบจำนวนหนึ่งที่ผ่านกระบวนการสร้างและคัดเลือกตามเกณฑ์คุณภาพที่กำหนดไว้แล้ว คะแนนที่ได้ คือ จำนวนข้อสอบที่ผู้สอบตอบถูก

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) หมายถึง กระบวนการทดสอบที่นำคอมพิวเตอร์มาใช้ในการทดสอบแทนการทดสอบบนกระดาษดินสอ ซึ่งประกอบด้วยข้อสอบจำนวนหนึ่งที่ผ่านกระบวนการสร้าง และคัดเลือกตามเกณฑ์คุณภาพที่กำหนดไว้แล้ว โดยผู้สอบแต่ละคนจะได้แบบทดสอบต่างชุดกัน เริ่มสอบจากข้อสอบที่มีค่าความยากปานกลางก่อน หากผู้สอบตอบถูก ให้จัดข้อสอบข้อถัดไปที่มีค่าความยากสูงขึ้น แต่หากผู้สอบตอบผิดให้จัดข้อสอบข้อถัดไปที่มีค่าความยากน้อยลง ทำเช่นนี้จนกว่าจะสิ้นสุดการทดสอบ คะแนนที่ได้ คือ ค่าประมาณความสามารถของผู้สอบ

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) หมายถึง ทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างความสามารถที่มีอยู่ภายในบุคคลกับผลการตอบข้อสอบหรือข้อคำถามโดยใช้โค้งคุณลักษณะของข้อสอบ (Item Characteristic Curve: ICC) ซึ่งโค้งคุณลักษณะของข้อสอบมีหลายลักษณะขึ้นอยู่กับโมเดล (Model) ที่ใช้อธิบายความสัมพันธ์ดังกล่าว โมเดลที่นิยมใช้ ได้แก่ โมเดลแบบ 1 พารามิเตอร์ (One – Parameter Model) โมเดลแบบ 2 พารามิเตอร์ (Two – Parameter Model) และโมเดลแบบ 3 พารามิเตอร์ (Three – Parameter Model) ซึ่งในการวิจัยนี้เลือกใช้โมเดลแบบ 3 พารามิเตอร์

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing Programs) หมายถึง โปรแกรมคอมพิวเตอร์ที่ได้พัฒนาขึ้นตามหลักการของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) และหลักการทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน (Uncertainly Decision Making) ตามเกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion) และมีการควบคุมการใช้ข้อสอบ (Item Exposure Control) ด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ (Systematic Random Sampling)

ค่าประมาณความสามารถของผู้สอบ (Ability Estimate) หมายถึง ค่าที่ใช้สำหรับประมาณค่าความสามารถที่แท้จริงของผู้สอบ (True Ability) แสดงถึงระดับความสามารถของผู้สอบ โดยคำนวณจากผลการตอบข้อสอบในทุกข้อที่ผ่านมา ตามทฤษฎีการตอบสนองข้อสอบ ค่าความสามารถที่แท้จริงของผู้สอบมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติมีค่าอยู่ระหว่าง -3.00 ถึง 3.00 ถ้า

ค่าประมาณความสามารถของผู้สอบมีค่าเท่ากับ -3.00 แสดงว่า ผู้สอบมีความสามารถต่ำมาก แต่ถ้าค่าประมาณความสามารถของผู้สอบมีค่าเท่ากับ 3.00 แสดงว่า ผู้สอบมีความสามารถสูงมาก

คะแนนรวมของผู้สอบ หมายถึง คะแนนที่ผู้สอบได้รับจากการทดสอบ หรือเป็นคะแนนสังเกตได้ของผู้สอบ (Observed Score) คำนวณจากจำนวนข้อสอบทั้งหมดที่ผู้สอบตอบถูกในการทดสอบด้วยคอมพิวเตอร์

ค่าความยากของข้อสอบ (Difficulty Parameter: b) หมายถึง ค่าที่แสดงถึงระดับความสามารถของผู้สอบที่จุดเปลี่ยนโค้งของโค้งคุณลักษณะของข้อสอบ ซึ่งมีค่าอยู่ระหว่าง -∞ ถึง +∞ แต่ในทางปฏิบัติค่าความยากของข้อสอบมีค่าตั้งแต่ -2.50 ถึง 2.50 โดยข้อสอบที่มีค่าความยากของข้อสอบเท่ากับ -2.50 แสดงว่า ข้อสอบง่ายมาก และข้อสอบที่มีค่าความยากของข้อสอบเท่ากับ 2.50 แสดงว่า ข้อสอบยากมาก

ค่าอำนาจจำแนกของข้อสอบ (Discrimination Power Parameter: a) หมายถึง ค่าที่เป็นสัดส่วนโดยตรงกับความชันของโค้งคุณลักษณะของข้อสอบ ณ จุดเปลี่ยนโค้ง มีค่าระหว่าง -∞ ถึง +∞ สำหรับค่าอำนาจจำแนกของข้อสอบที่เป็นลบ (-) แสดงว่า ข้อสอบไม่ดี ไม่สามารถจำแนกผู้สอบได้ ต้องตัดข้อสอบข้อนั้นทิ้ง ค่าอำนาจจำแนกของข้อสอบที่เป็นศูนย์ (0) แสดงว่า ข้อสอบไม่มีอำนาจจำแนก และอำนาจจำแนกของข้อสอบที่เป็นบวก (+) แสดงว่า ข้อสอบดี สามารถจำแนกผู้สอบได้ ในทางปฏิบัติข้อสอบที่ถูกคัดเลือกไว้ จะมีค่าอำนาจจำแนกของข้อสอบตั้งแต่ 0.50 ถึง 2.50

ค่าการเดาของข้อสอบ (Guess Parameter: c) หมายถึง ค่าความน่าจะเป็นของผู้สอบที่มีความสามารถต่ำ แต่ตอบข้อสอบข้อนั้นได้ถูกต้อง หรือค่าที่แสดงโอกาสการตอบข้อสอบถูก โดยผู้ที่ไม่มีความรู้ในเรื่องนั้น ๆ สำหรับค่าการเดาของข้อสอบมีค่าตั้งแต่ 0 ถึง 1 ในทางปฏิบัติข้อสอบที่ถูกเลือกไว้ จะมีค่าการเดาของข้อสอบไม่เกิน 0.30

วิธีการคัดเลือกข้อสอบข้อถัดไป (Next Item Selection) หมายถึง วิธีการคัดเลือกข้อสอบในคลังข้อสอบที่ยังไม่ถูกเลือกใช้ในการทดสอบ ซึ่งเป็นข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบ ซึ่งเป็นขั้นตอนหนึ่งในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) หมายถึง วิธีการคัดเลือกข้อสอบข้อถัดไปวิธีหนึ่ง โดยข้อสอบที่ถูกคัดเลือกเป็นข้อสอบข้อถัดไป คือ ข้อสอบที่มีค่าสารสนเทศสูงสุด ณ ระดับความสามารถที่ตรงหรือใกล้เคียงกับความสามารถของผู้สอบขณะนั้น

วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (Risk Decision Making: RDM) หมายถึง วิธีการคัดเลือกข้อสอบข้อถัดไปวิธีการหนึ่ง ที่ใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง ตามเกณฑ์ของค่าความคาดหวังสูงสุด เป็นพื้นฐาน โดยข้อสอบที่ถูกคัดเลือกเป็นข้อสอบข้อถัดไป คือ ข้อสอบที่มีค่าสัมประสิทธิ์การกระจายสูงที่สุด

วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion: HC) หมายถึง วิธีการคัดเลือกข้อสอบข้อถัดไปวิธีการหนึ่ง ที่ใช้ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ตามเกณฑ์ของเฮอร์วิคซ์ เป็นพื้นฐาน โดยข้อสอบที่ถูกคัดเลือกเป็นข้อสอบข้อถัดไป คือ ข้อสอบที่มีค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบต่ำสุด

วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ (Hurwicz Criterion with Item Exposure Control: HC-Ex) หมายถึง วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบที่ใช้วิธีการสุ่มตัวอย่างแบบมีระบบ

ทฤษฎีการตัดสินใจ (Decision Theory) หมายถึง หลักเกณฑ์ หรือเครื่องมือต่าง ๆ ที่เข้ามาช่วยในการตัดสินใจ เพื่อทำให้ผู้ตัดสินใจมีโอกาสผิดพลาดน้อยลงหรือการตัดสินใจได้ถูกต้องมากขึ้นในสภาวการณ์ต่าง ๆ ที่สามารถทำการประเมินได้ และใช้เกณฑ์หรือเครื่องมือดังกล่าวพิจารณาทางเลือกที่ดีที่สุด

การตัดสินใจในสภาวการณ์ที่ไม่แน่นอน (Uncertainly Decision Making) หมายถึง การตัดสินใจที่ผู้ตัดสินใจมีข้อมูลเกี่ยวกับเรื่องที่ตัดสินใจไม่เพียงพอ ทำให้ไม่สามารถคาดการณ์ผลลัพธ์ที่จะเกิดขึ้นได้อย่างแน่นอน การตัดสินใจจึงขึ้นอยู่กับดุลยพินิจ หรือประสบการณ์ของผู้ตัดสินใจ

เกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion) หมายถึง เกณฑ์การตัดสินใจเกณฑ์หนึ่งในการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน โดยตัดสินใจเลือกทางเลือกที่ให้ค่าเฉลี่ยถ่วงน้ำหนักของผลตอบแทนที่สูงสุด ซึ่งค่าเฉลี่ยถ่วงน้ำหนักของแต่ละทางเลือกนี้ จะคำนวณมาจากผลตอบแทนสูงสุดกับผลตอบแทนต่ำสุดของทางเลือกนั้น ๆ และใช้ค่าสัมประสิทธิ์ของการมองโลกในแง่ดี ($\alpha$) มาเป็นค่าถ่วงน้ำหนัก ซึ่งค่านี้มีค่าตั้งแต่ 0 ถึง 1 เป็นค่าที่ถูกกำหนดจากผู้ตัดสินใจ โดยคาดการณ์ว่าผลตอบแทนที่จะได้อยู่ในระดับใด กล่าวคือ ถ้าค่า $\alpha$ เข้าใกล้ 1 แสดงว่า ผู้ตัดสินใจคาดว่าผลตอบแทนที่จะได้อยู่ในระดับดี (มองโลกในแง่ดี) แต่ถ้าค่า $\alpha$ เข้าใกล้ 0 แสดงว่า ผู้ตัดสินใจคาดว่าผลตอบแทนที่จะได้อยู่ในระดับต่ำ (มองโลกในแง่ร้าย)

การเลือกตัวอย่าง (Sampling) หมายถึง กระบวนการเลือกตัวอย่างจากประชากร เพื่อให้กลุ่มตัวอย่างที่ได้เป็นตัวแทนที่ดีของประชากร และสามารถใช้ข้อมูลจากกลุ่มตัวอย่างเป็นข้อมูลอ้างอิงสู่ประชากรได้อย่างสมเหตุสมผล

การสุ่มตัวอย่างแบบมีระบบ (Systematic Random Sampling) หมายถึง วิธีการสุ่มตัวอย่างวิธีหนึ่งที่ใช้กับประชากรที่มีจำนวนมาก ที่มีรายชื่อของทุกหน่วยประชากร และมีการจัดเรียงลำดับตามรายชื่อไว้แล้ว การสุ่มนี้ทุกหน่วยประชากรมีโอกาสถูกเลือกเท่า ๆ กัน ดำเนินการโดยแบ่งประชากรออกเป็นช่วง ๆ ที่เท่ากัน ซึ่งอาจแบ่งตามสัดส่วนของขนาดกลุ่มตัวอย่างกับประชากร แล้วสุ่มจากประชากรช่วงแรก ส่วนประชากรช่วงถัดไปให้นับจากค่าสัดส่วนที่คำนวณไว้

ประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป (Efficiency of the Next Item Selection Procedures) หมายถึง ความสามารถของวิธีการคัดเลือกข้อสอบข้อถัดไป สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งจะพิจารณาในด้านการประมาณค่าความสามารถของผู้สอบ ด้านความยาวของแบบทดสอบ และด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2

ประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบ (Estimative Efficiency) หมายถึง ความสามารถของวิธีการประมาณค่าความสามารถของผู้สอบให้เท่ากับหรือใกล้เคียงกับค่าความสามารถที่แท้จริงของผู้สอบมากที่สุด ซึ่งจะพิจารณาจากค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error) และค่าความลำเอียงเฉลี่ย (Average Bias)

ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) หมายถึง ค่าบอกความถูกต้องของค่าประมาณความสามารถของผู้สอบจำนวนหนึ่ง ซึ่งมีค่าเท่ากับผลต่างกำลังสองเฉลี่ย

ระหว่างค่าประมาณความสามารถกับค่าความสามารถที่แท้จริงของผู้สอบ ถ้าค่าความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเข้าใกล้ศูนย์ แสดงว่า ค่าประมาณความสามารถของผู้สอบมีค่าใกล้เคียงกับค่าความสามารถที่แท้จริงของผู้สอบ คือ ค่าประมาณความสามารถมีความถูกต้องมาก

ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) หมายถึง ค่าบอกความถูกต้องของค่าประมาณความสามารถของผู้สอบจำนวนหนึ่ง ซึ่งมีค่าเท่ากับค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ถ้าค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเข้าใกล้ศูนย์ แสดงว่า ค่าประมาณความสามารถของผู้สอบมีค่าใกล้เคียงกับค่าความสามารถที่แท้จริงของผู้สอบ

ค่าความลำเอียงเฉลี่ย (Average Bias) หมายถึง ค่าบอกความเที่ยงตรงของค่าประมาณความสามารถของผู้สอบจำนวนหนึ่ง มีค่าเท่ากับค่าเฉลี่ยของผลต่างระหว่างค่าประมาณความสามารถและค่าความสามารถที่แท้จริงของผู้สอบแต่ละคน ใช้บอกทิศทางการประมาณค่าความสามารถของผู้สอบว่า ให้ผลสูงหรือต่ำกว่าค่าความสามารถที่แท้จริงของผู้สอบมากน้อยเพียงใด

ขนาดคลังข้อสอบ (Item Pool Size) หมายถึง จำนวนข้อสอบทั้งหมดที่บรรจุในคลังข้อสอบ ซึ่งเป็นแหล่งรวบรวมข้อสอบ โดยข้อสอบทั้งหมดที่บรรจุในคลังข้อสอบต้องเป็นข้อสอบมีค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ ที่เหมาะสม

ความยาวของแบบทดสอบ (Test Length) หมายถึง ปริมาณของข้อสอบในแบบทดสอบที่ต้องการจัดให้แก่ผู้สอบ

การใช้ข้อสอบ (Item Exposure) หมายถึง การจัดข้อสอบที่ผ่านการคัดเลือกแล้วให้แก่ผู้สอบ

อัตราการใช้ข้อสอบ (Item Exposure Rate) หมายถึง อัตราส่วนระหว่างจำนวนการใช้ข้อสอบข้อนั้น ๆ ต่อจำนวนผู้สอบทั้งหมดที่ทำการทดสอบ ซึ่งมีค่าตั้งแต่ 0 ถึง 1 ถ้าอัตราการใช้ข้อสอบมีค่าเข้าใกล้ 1 แสดงว่า ข้อสอบข้อนั้นถูกจัดให้แก่ผู้สอบบ่อยครั้งเกินไป ในทางปฏิบัติคลังข้อสอบไม่ควรมีข้อสอบที่มีค่าอัตราการใช้ข้อสอบมากกว่า 0.20 หรือหากมีต้องมีจำนวนน้อยที่สุด

การควบคุมการใช้ข้อสอบ (Item Exposure Control) หมายถึง การควบคุมการจัดข้อสอบที่ผ่านการคัดเลือกแล้วให้แก่ผู้สอบด้วยวิธีการใดวิธีการหนึ่ง เพื่อไม่ให้ข้อสอบข้อใดข้อหนึ่งถูกจัดให้แก่ผู้สอบบ่อยครั้งเกินไป คือ ควบคุมไม่ให้มีข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.20

ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ (Standard Error of the Ability Estimate: SEE) หมายถึง ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถที่แท้จริงของผู้สอบ ซึ่งมีค่าเท่ากับรากที่สองของส่วนกลับของฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function)

เกณฑ์ยุติการทดสอบ (Termination Criterion) หมายถึง ข้อกำหนดเพื่อใช้สิ้นสุดการทดสอบ ซึ่งมีหลายเกณฑ์ แต่เกณฑ์ที่เป็นที่นิยม ได้แก่ กำหนดจำนวนข้อสอบคงที่ เช่น กำหนดจำนวนข้อสอบ 25 ข้อ ถ้าผู้สอบทำข้อสอบครบ 25 ข้อ การทดสอบจะยุติลง หรือกำหนดค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบที่เหมาะสม เช่น กำหนดค่า SEE ต่ำกว่า 0.30 แสดงว่า ผู้สอบจะดำเนินการทดสอบไปเรื่อย ๆ จนกว่าค่า SEE มีค่าลดลงต่ำกว่า 0.30 การทดสอบจึงยุติลง

นักศึกษา หมายถึง ผู้เรียนระดับชั้นปีที่ 2 คณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร ประจำปีการศึกษา 2559

# บทที่ 2
# เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยนี้เพื่อพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แล้วเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้นทั้งในกรณีที่มีและไม่มีการควบคุมการใช้ข้อสอบ กับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง โดยเปรียบเทียบประสิทธิภาพด้าน 1) การประมาณค่าความสามารถของผู้สอบ 2) ความยาวของแบบทดสอบ และ 3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จากนั้นพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปและมีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้น แล้วศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์

ในบทนี้ นำเสนอแนวคิดหรือทฤษฎีบทพื้นฐานที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รวมทั้งทฤษฎีการตัดสินใจ เพื่อใช้เป็นฐานความคิดในการศึกษาและดำเนินการวิจัย แบ่งการนำเสนอออกเป็น 4 ตอน ดังนี้

ตอนที่ 1 การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
ตอนที่ 2 ทฤษฎีการตัดสินใจและงานวิจัยที่เกี่ยวข้อง
ตอนที่ 3 การศึกษาในสถานการณ์จำลองแบบมอนติคาร์โล
ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

## ตอนที่ 1 การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การทดสอบโดยใช้คอมพิวเตอร์ (Computer Based Testing: CBT) เป็นการนำคอมพิวเตอร์มาใช้ในการจัดการการทดสอบ โดยแบ่งออกเป็น 4 ยุค คือ 1) การทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) เป็นยุคที่นำคอมพิวเตอร์มาใช้ในการทดสอบแทนการทดสอบบนกระดาษ ซึ่งผู้สอบแต่ละคนจะได้แบบทดสอบชุดเดียวกัน มีเวลาในการทดสอบเท่ากัน และสอบในสถานที่เดียวกัน 2) ยุคการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นยุคที่นำคอมพิวเตอร์มาใช้ในการทดสอบ โดยผู้สอบแต่ละคนจะได้แบบทดสอบต่างชุดกัน ซึ่งมีขั้นตอนในการคัดเลือกข้อสอบที่มีระดับความยากของข้อสอบเหมาะสมกับระดับความสามารถของผู้สอบ และความยาวของแบบทดสอบอาจไม่เท่ากัน 3) การวัดอย่างต่อเนื่อง (Continuous Measurement: CM) เป็นยุคที่นำการเรียนการสอน รวมทั้งการทดสอบมาใช้ในระบบคอมพิวเตอร์เดียวกัน โดยสลับกันไปมาระหว่างบทเรียน กล่าวคือ เมื่อเรียนจบเรื่องใดเรื่องหนึ่ง จะทำการทดสอบเรื่องนั้นทันที และ 4) การวัดอย่างเชี่ยวชาญ (Intelligent Measurement: IM) เป็นยุคที่นำคอมพิวเตอร์มาใช้ในการจัดการเรียนการสอน รวมทั้งการทดสอบครบทุกขั้นตอน กล่าวคือ นำคอมพิวเตอร์มาใช้ในการให้คะแนนแบบทดสอบประเภทเขียนตอบที่มีคำตอบแน่นอน

แล้วทำการวิเคราะห์จุดแข็งและจุดอ่อน เพื่อนำไปเสนอแนะกับผู้เรียนหรือผู้สอบ (Bunderson, Inouye, & Olsen, 1989 cited in Linn, 1990, pp. 367-407)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ เป็นการใช้ข้อสอบจากคลังข้อสอบมาสร้างเป็นแบบทดสอบที่มีความเหมาะสมกับความสามารถของผู้สอบแต่ละคน กล่าวคือ ผู้สอบจะได้ข้อสอบที่มีระดับความยากไม่ยากหรือไม่ง่ายเกินไป ทำให้ผู้สอบไม่เกิดความเครียดเมื่อได้รับข้อสอบที่มีความยากเกินไป หรือเกิดความเบื่อหน่ายเมื่อได้รับข้อสอบที่ง่ายเกินไป และทำให้ความยาวของแบบทดสอบลดลง และตามทฤษฎีการตอบสนองข้อสอบ ข้อสอบที่มีระดับความยากเหมาะสมกับผู้สอบจะให้ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบลดลง เมื่อเทียบกับการทดสอบแบบประเพณีนิยม (Antal, Eros, & Imre, 2010)

**การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นการทดสอบโดยใช้คอมพิวเตอร์เป็นเครื่องมือในการทดสอบแทนการทดสอบบนกระดาษ และมีกระบวนการในการจัดข้อสอบให้มีความเหมาะสมกับความสามารถของผู้สอบ กล่าวคือ เมื่อเริ่มการทดสอบ ผู้สอบจะได้ข้อสอบที่มีค่าความยากของข้อสอบปานกลาง จากนั้นจะนำผลการตอบของผู้สอบมาวิเคราะห์หรือประเมินระดับความสามารถของผู้สอบ เพื่อนำมาใช้ในการจัดข้อสอบข้อถัดไปให้แก่ผู้สอบ โดยอาศัยหลักการว่า หากผู้สอบตอบข้อสอบถูก ข้อสอบข้อถัดไปที่จะจัดให้กับผู้สอบจะมีค่าความยากของข้อสอบเพิ่มขึ้น ในทางกลับกัน หากผู้สอบตอบข้อสอบผิด ข้อสอบข้อถัดไปจะมีค่าความยากของข้อสอบลดลง กระบวนการนี้จะดำเนินต่อไปจนกระทั่งสิ้นสุดการทดสอบ และจากความสามารถของคอมพิวเตอร์ในปัจจุบัน การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะช่วยอำนวยความสะดวกให้แก่ผู้ควบคุมการทดสอบ และผู้สอบเป็นอย่างมาก โดยคอมพิวเตอร์จะเข้ามาอำนวยการทุกขั้นตอนในการทดสอบ ตั้งแต่การสร้างคลังข้อสอบ การคัดเลือกข้อสอบข้อแรก การคัดเลือกข้อสอบข้อถัดไป การประมาณค่าความสามารถของผู้สอบ และเกณฑ์ยุติการทดสอบ ซึ่งจะอาศัยทฤษฎีการตอบสนองข้อสอบเป็นพื้นฐาน

ในที่นี้จะอธิบายกระบวนการและองค์ประกอบที่สำคัญในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ตลอดจนปัญหาและข้อจำกัดต่าง ๆ ที่พบในการทดสอบ เพื่อใช้เป็นฐานความรู้ในการดำเนินการวิจัยต่อไป ซึ่งมีรายละเอียดต่าง ๆ ดังต่อไปนี้

**1. ประเภทของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบ่งได้ 2 ประเภท ได้แก่ การทดสอบแบบสองขั้นตอน (Two-Stage Strategies) และการทดสอบแบบหลายขั้นตอน (Multi-Stage Strategies) ซึ่งมีรายละเอียดดังนี้ (de Ayala, 2009, pp. 373-381; Wainer, Dorans, Flaugher, Green, & Mislevy, 2000, pp. 105-112)

**1.1 การทดสอบแบบสองขั้นตอน** (Two-Stage Strategies)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบบสองขั้นตอนเป็นการทดสอบที่มีการดำเนินการ 2 ขั้นตอน คือ ขั้นตอนที่ 1 การทดสอบแบบแยกทาง (Routing Test) เป็นการทดสอบเพื่อหาระดับความสามารถของผู้สอบเบื้องต้น ผลจากการทดสอบในขั้นตอนนี้ จะถูกนำไปใช้เลือกแบบทดสอบในขั้นตอนที่สอง หากผู้สอบได้รับคะแนนสูง แสดงว่า ผู้สอบมีความสามารถ

มาก ดังนั้น ในขั้นตอนที่สองผู้สอบจะได้แบบทดสอบที่มีระดับความยากมาก แต่หากผู้สอบได้รับคะแนนต่ำ ในขั้นตอนที่สอง ผู้สอบจะได้แบบทดสอบที่มีระดับความยากลดหลั่นลงมา การทดสอบแบบทางแยกนี้ ประกอบด้วยข้อสอบประมาณ 10 ข้อ ที่มีค่าความยากของข้อสอบระดับปานกลาง ผู้สอบทุกคนจะได้ข้อสอบชุดเดียวกัน เหมือนกันทุกข้อ ขั้นตอนที่ 2 การทดสอบแบบวัดผล (Measurement Test) เป็นการทดสอบเพื่อวัดความสามารถของผู้สอบ ผู้สอบแต่ละคนจะได้แบบทดสอบต่างชุดกัน ขึ้นอยู่กับระดับความสามารถของผู้สอบ ที่ประเมินได้จากผลการทดสอบในขั้นตอนที่ 1 การทดสอบแบบวัดผลนี้ ข้อสอบในแต่ละแบบทดสอบประกอบด้วยข้อสอบประมาณ 20-30 ข้อ ซึ่งข้อสอบในแบบทดสอบเดียวกัน มีค่าความยากของข้อสอบใกล้เคียงกัน

**1.2 การทดสอบแบบหลายขั้นตอน** (Multi-Stage Strategies)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบบหลายขั้นตอนเป็นการทดสอบที่มีการดำเนินการมากกว่า 2 ขั้นตอน ในแต่ละขั้นตอนผู้สอบจะได้ข้อสอบแตกต่างกัน ขึ้นอยู่กับค่าความสามารถของผู้สอบ ที่ถูกประมาณขึ้น ณ ขณะนั้น หลักการทั่วไปของการทดสอบแบบหลายขั้นตอน คือ ผู้สอบจะได้ข้อสอบข้อแรกที่มีค่าความยากของข้อสอบปานกลาง ถ้าผู้สอบตอบข้อสอบถูก ผู้สอบจะได้ข้อสอบข้อถัดไปที่มีค่าความยากของข้อสอบเพิ่มขึ้น แต่ถ้าผู้สอบตอบข้อสอบผิด ผู้สอบจะได้ข้อสอบข้อถัดไปที่มีค่าความยากของข้อสอบลดลง การทดสอบจะดำเนินการเช่นนี้เรื่อย ๆ จนกว่าจะยุติการทดสอบตามเกณฑ์ที่ตั้งไว้ การทดสอบแบบหลายขั้นตอน แบ่งได้ 2 แบบ ดังนี้

1.2.1 การทดสอบแบบทางแยกคงที่ (Fixed-Branching Model)

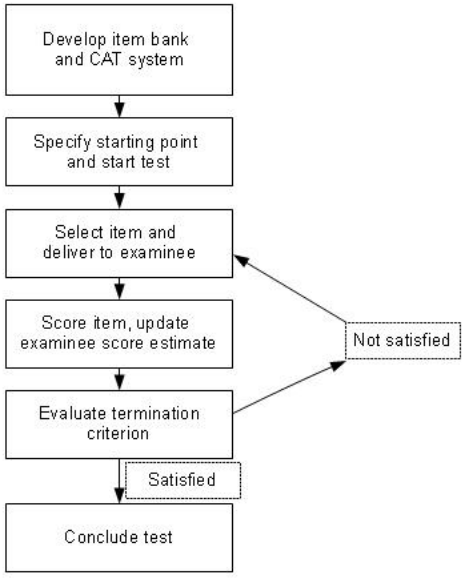
การทดสอบแบบทางแยกคงที่ เป็นการทดสอบที่ผู้สอบแต่ละคนจะได้ข้อสอบแตกต่างกัน ขึ้นอยู่กับผลการสอบในข้อที่ผ่านมาก หากผู้สอบตอบข้อสอบในข้อที่ผ่านมาถูก ผู้สอบจะได้ข้อสอบข้อถัดไปที่มีความยากของข้อสอบเพิ่มขึ้น ในทางตรงกันข้าม หากผู้สอบตอบข้อสอบในข้อที่ผ่านมาผิด ผู้สอบจะได้ข้อสอบข้อถัดไปที่มีความยากของข้อสอบลดลง ซึ่งการทดสอบนี้ จะจัดรูปแบบหรือเส้นทางในการทดสอบอย่างแน่นอนไว้ล่วงหน้า เช่น การทดสอบแบบพีระมิด (Pyramid Model) การทดสอบแบบยืดหยุ่น (Flexilevel Model)

1.2.2 การทดสอบแบบทางแยกแปรผัน (Variable Branching Model)

การทดสอบแบบทางแยกแปรผัน เป็นการทดสอบที่มีลักษณะคล้ายกับการทดสอบแบบทางแยกคงที่ แต่ไม่มีการจัดรูปแบบหรือเส้นทางในการทดสอบอย่างแน่นอนไว้ล่วงหน้า คือ ไม่ว่าผู้สอบจะตอบข้อสอบในข้อที่ผ่านมาถูกหรือผิด จะไม่ทราบอย่างแน่นอนว่าผู้สอบจะได้ข้อสอบข้อถัดไปเป็นข้อใด การคัดเลือกข้อสอบข้อถัดไปพิจารณาจากค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น และค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบ และค่าการเดาของข้อสอบ (ขึ้นอยู่กับโมเดลการตอบสนองข้อสอบที่เลือกใช้) และอาศัยโมเดลทางคณิตศาสตร์มาช่วยวิเคราะห์ เช่น วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) วิธีการคัดเลือกข้อสอบข้อถัดไป ด้วยใช้ค่าสารสนเทศฟิชเชอร์สูงสุด (Maximum Fisher Information: MFI) หรือวิธีการคัดเลือกข้อสอบข้อถัดไปที่ระดับความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้สอบ

**2. กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีทั้งหมด 5 ขั้นตอน (Thompson & Weiss, 2011) คือ 1) การสร้างคลังข้อสอบ 2) การคัดเลือกข้อสอบข้อแรก 3) การคัดเลือกข้อสอบข้อถัดไป 4) การประมาณค่าความสามารถของผู้สอบ และ 5) เกณฑ์ยุติการทดสอบ โดยกระบวนการทดสอบในแต่ละขั้นตอนมีแนวคิดและวิธีการหลายแบบขึ้นอยู่กับผู้ออกแบบจะเป็นผู้กำหนด สามารถแสดงดังภาพที่ 2-1 (Thompson & Weiss, 2011)




ภาพที่ 2-1 กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Thompson & Weiss, 2011, p. 2)

กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แต่ละกระบวนการมีรายละเอียดดังนี้

**2.1 การสร้างคลังข้อสอบ** (Create Item Bank)

คลังข้อสอบ (Item Pool) เป็นการรวบรวมข้อสอบที่มีคุณภาพจำนวนมาก และครอบคลุมเนื้อหาอย่างสมบูรณ์เก็บไว้อย่างเป็นระบบ และสามารถนำมาใช้ได้อย่างสะดวก ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผู้สอบแต่ละคนจะได้ข้อสอบที่แตกต่างกันขึ้นอยู่กับระดับความสามารถของผู้สอบ ดังนั้น คุณภาพของคลังข้อสอบจึงมีผลอย่างมากต่อประสิทธิภาพของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ คุณภาพของคลังข้อสอบควรพิจารณาจากขนาดของคลังข้อสอบ ความครอบคลุมเนื้อหาของข้อสอบ ตลอดจนคุณลักษณะทางจิตมิติ (Wainer et al., 2000, pp. 37-45; Parshall, Spray, Kalohn, & Davey, 2002, pp. 169-171)

คลังข้อสอบจะต้องประกอบด้วยชุดของข้อสอบที่มุ่งวัดลักษณะ หรือความสามารถที่ต้องการได้ ซึ่งต้องมีข้อสอบจำนวนมากเพียงพอที่จะทำให้ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ กระจายอยู่ทุกระดับความสามารถของผู้สอบ และทุกระดับเนื้อหาที่จะทดสอบ นักวิจัยหลายท่านจึงเสนอจำนวนข้อสอบที่

เพียงพอในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เช่น Embetson and Reise (2000, p. 264) เสนอให้มีข้อสอบประมาณ 100 ข้อ Weiss (2011) เสนอให้มีข้อสอบไม่น้อยกว่า 200 ข้อ Thompson and Weiss (2011) เสนอให้มีข้อสอบประมาณ 400 ข้อ และ de Ayala (2009, p. 376) เสนอว่าคลังข้อสอบควรมีข้อสอบจำนวน 8 - 12 เท่า ของความยาวของแบบทดสอบ กล่าวคือ ถ้าในการทดสอบใช้ข้อสอบประมาณ 25 ข้อ คลังข้อสอบควรมีข้อสอบ 200 - 300 ข้อ

Molina, Pareja, and Sanmartin (2008) ได้เสนอขั้นตอนในการสร้างคลังข้อสอบที่นำมาใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เพื่อให้ได้ประสิทธิภาพมากขึ้น ซึ่งประกอบด้วย 5 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 กำหนดกฎเกณฑ์ คำอธิบายต่าง ๆ รวมถึงคำสั่งที่เกี่ยวข้อง ที่คาดว่าจะส่งผลทางจิตวิทยาระหว่างดำเนินการทดสอบ

ขั้นตอนที่ 2 เลือกข้อสอบที่เหมาะสมและตรงตามกฎเกณฑ์ที่กำหนดในคลังข้อสอบที่ได้จากการบริหารจัดการทดสอบโดยใช้คอมพิวเตอร์หรือการจัดการทดสอบโดยใช้กระดาษคำตอบ

ขั้นตอนที่ 3 พัฒนาฐานข้อมูลของคลังข้อสอบ จากชุดข้อสอบที่เลือกจากการจัดการทดสอบ โดยใช้คอมพิวเตอร์หรือการจัดการทดสอบด้วยกระดาษคำตอบ เพื่อนำมาใช้ในการบริหารการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ขั้นตอนที่ 4 แยกข้อสอบที่ได้จากการบริหารจัดการทดสอบโดยใช้คอมพิวเตอร์หรือการทดสอบด้วยกระดาษคำตอบ โดยนำข้อสอบที่เลือกแล้วไปจัดเก็บคนละส่วนกัน และนำข้อสอบที่เลือกมาแล้วมาเลือกอีกครั้ง จึงนำไปพัฒนาเป็นคลังข้อสอบที่จะนำไปใช้

ขั้นตอนที่ 5 วิเคราะห์ความแตกต่างของผลกระทบทางจิตวิทยาของกลุ่มข้อสอบที่ได้จากการบริหารจัดการทดสอบด้วยคอมพิวเตอร์ หรือการทดสอบด้วยกระดาษคำตอบกับกลุ่มข้อสอบในคลังข้อสอบที่พัฒนาขึ้น ซึ่งนำไปเป็นข้อมูลในการพัฒนาปรับปรุงคลังข้อสอบใหม่ให้มีคุณภาพยิ่งขึ้น

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) มีข้อตกลงเบื้องต้นที่สำคัญข้อหนึ่ง คือ ความเป็นเอกมิติ (Unidimensionality) ข้อสอบแต่ละข้อในแบบทดสอบควรวัดความสามารถหรือคุณลักษณะของผู้สอบเพียงลักษณะเดียว แต่ในทางปฏิบัติแบบทดสอบส่วนใหญ่ประกอบไปด้วยข้อสอบที่วัดเนื้อหาหลาย ๆ ด้าน ดังนั้น เพื่อความสอดคล้องกับข้อตกลงเบื้องต้นนี้ จึงแก้ปัญหาด้วยการแบ่งคลังข้อสอบใหญ่ ให้เป็นคลังข้อสอบย่อย โดยแต่ละคลังข้อสอบย่อยบรรจุข้อสอบที่วัดคุณลักษณะด้านเดียวกัน ซึ่งการสร้างคลังข้อสอบมากกว่าหนึ่งคลัง หรือแบ่งคลังข้อสอบใหญ่ออกเป็นคลังข้อสอบย่อย ควรพิจารณาเพิ่มเติมในประเด็นต่อไปนี้ ขนาดของคลังข้อสอบ การทับซ้อนของคลังข้อสอบ วิธีเข้าสู่แต่ละคลังข้อสอบ การหมุนเวียนใช้แต่ละคลังข้อสอบและเกณฑ์การออกจากคลังข้อสอบ การควบคุมการใช้ข้อสอบ (Item Exposure Control) และการควบคุมสัดส่วนเนื้อหาของข้อสอบ (Content Balancing)

**2.2 การคัดเลือกข้อสอบข้อแรก** (First Item Selection)

การทดสอบเริ่มต้นด้วยการคัดเลือกข้อสอบข้อแรก (Initial item) ในขั้นนี้โดยทั่วไปนิยมใช้ข้อสอบที่มีความยากระดับปานกลาง แบ่งวิธีการคัดเลือกเป็น 2 กรณีดังนี้

กรณีที่ 1 เมื่อประชากรผู้สอบมีความสามารถใกล้เคียงกัน (homogeneous) หรือไม่ทราบข้อมูลเกี่ยวกับผลสัมฤทธิ์ที่ผ่านมาของผู้สอบ ข้อสอบข้อแรกควรเป็นข้อสอบที่มีค่าความยากของข้อสอบปานกลาง

กรณีที่ 2 เมื่อประชากรผู้สอบมีความสามารถแตกต่างกัน (heterogeneous) หรือมีข้อมูลเกี่ยวกับระดับการศึกษาของผู้สอบ เช่น ทราบว่าผู้สอบอยู่ในระดับมัธยมปลาย ข้อสอบข้อแรกควรเป็นข้อสอบที่มีค่าความยากของข้อสอบปานกลางของแต่ละระดับการศึกษาของผู้สอบ

ทั้งนี้ยังมีผู้เสนอแนวคิดเกี่ยวกับการคัดเลือกข้อสอบข้อแรก ดังนี้

Lord (1977) พบว่า การคัดเลือกข้อสอบข้อแรกไม่ค่อยมีความสำคัญในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ โดยทำการศึกษาในสถานการณ์จำลอง ซึ่งจำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีความยาวของแบบทดสอบ 25 ข้อขึ้นไป เมื่อเริ่มต้นการทดสอบด้วยข้อสอบข้อแรกที่มีค่าความยากของข้อสอบต่างกันหลาย ๆ ค่า ปรากฏว่า ระดับค่าความยากของข้อสอบที่ต่างกัน ไม่มีผลต่อความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ จึงสรุปได้ว่า การคัดเลือกข้อสอบข้อแรกไม่มีผลต่อแบบทดสอบที่มีความยาวของแบบทดสอบจำนวนข้อมาก

Chalhoub-Deville, Alcaya, and Lozier (1996, p. 36) ได้เสนอวิธีการ โดยเริ่มจากให้ผู้สอบประเมินความสามารถของตนเองก่อนการคัดเลือกข้อสอบข้อแรก โดยเตรียมข้อสอบที่มีความยากระดับต่าง ๆ กันกระจายครอบคลุมค่าความสามารถของผู้สอบ และผู้สอบต้องตอบข้อสอบอย่างน้อยที่สุด คือ ตอบถูก 1 ข้อ และตอบผิด 1 ข้อ เพื่อนำไปใช้ประมาณค่าความสามารถของผู้สอบ

Fan and Zhu (1999 cited in Fan & Zhu, 2002) เสนอการใช้ข้อมูลสารสนเทศที่เกี่ยวกับความสามารถของผู้สอบ เช่น ผลการเรียนในรายวิชาที่เกี่ยวข้อง หรือผลการเรียนเฉลี่ย แล้วนำมาใช้คัดเลือกข้อสอบข้อแรก สามารถลดอัตราการใช้ข้อสอบที่มีค่าความยากปานกลางได้ และเสนอรายงานว่า การทดสอบเบื้องต้นก่อนการทดสอบจริง เพื่อนำผลมากำหนดจุดเริ่มต้นในการทดสอบมีผลทำให้ความยาวของแบบทดสอบลดลง เมื่อเทียบกับการทดสอบที่เริ่มต้นด้วยข้อสอบที่มีความยากปานกลาง และการทดสอบที่ใช้สารสนเทศจากผลการเรียนในรายวิชาที่เกี่ยวข้อง หรือผลการเรียนเฉลี่ย

Chang, Qian, and Ying (2001) เสนอแนะว่า ควรใช้ข้อสอบเทียม 3 ข้อเพื่อหาค่าสารสนเทศของผู้สอบเบื้องต้น โดยข้อสอบข้อแรกเป็นข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบ ($a_1$) เท่ากับ 1 ส่วนค่าความยากของข้อสอบ ($b_1$) จะสุ่มเลือกจากการแจกแจงโค้งปกติ N(0,1) และค่าการเดาของข้อสอบ (c) เท่ากับ 0.20 ถ้าผู้สอบตอบข้อสอบเทียมข้อแรกถูก ข้อสอบข้อที่ 2 จะมีความยากมากขึ้นเป็น $b_2=b_1+2$ แต่ถ้าผู้สอบตอบข้อสอบข้อแรกผิด ข้อสอบข้อที่ 2 จะมีความยากลดลงเป็น $b_2=b_1-2$ ต่อมาข้อสอบข้อที่ 3 จะมีการดำเนินการคล้ายกัน ซึ่งข้อสอบเทียมทั้ง 3 ข้อ จะมีค่าอำนาจจำแนกของข้อสอบ และค่าการเดาของข้อสอบ เท่ากัน เมื่อดำเนินการครบ 3 ข้อ จะนำผลการตอบข้อสอบเทียมทั้ง 3 ข้อ มาประมาณค่าความสามารถของผู้สอบ ด้วยวิธีการประมาณค่าความเป็นไปได้สูงสุด (Maximum Likelihood Estimation: MLE) ซึ่งจะใช้วิธีการประมาณค่าความเป็นไปได้สูงสุด ตั้งแต่เริ่มต้น จนสิ้นสุดการทดสอบ และถ้าผลการตอบข้อสอบเทียมทั้ง 3 ข้อได้ผลถูกหมด หรือผิดหมด ต้องเริ่มต้นทำข้อสอบเทียมใหม่ตามขั้นตอนข้างต้น

การคัดเลือกข้อสอบข้อแรก ส่วนใหญ่นิยมคัดเลือกข้อสอบที่มีค่าความยากของข้อสอบในระดับปานกลางมาให้ผู้สอบใช้เป็นข้อสอบข้อแรก โดยไม่ได้คำนึงถึงระดับความสามารถของผู้สอบ เป็นผลทำให้ข้อสอบที่มีค่าความยากของข้อสอบปานกลางถูกเลือกใช้ในการทดสอบบ่อยครั้งจนเกินไป ด้วยเหตุนี้การคัดเลือกข้อสอบข้อแรก จึงใช้วิธีสุ่มเลือกจากข้อสอบที่มีค่าความยากของข้อสอบหลายระดับ หรือใช้ข้อมูลสารสนเทศที่เกี่ยวข้องของผู้สอบ เช่น คะแนนสอบย่อยก่อนสอบจริง หรือผลการเรียนเฉลี่ย มาช่วยในการคัดเลือกข้อสอบข้อแรกเพื่อให้ได้ข้อสอบที่มีค่าความยากของข้อสอบสอดคล้องกับระดับความสามารถของผู้สอบ

**2.3 การคัดเลือกข้อสอบข้อถัดไป** (Next Item Selection)

การคัดเลือกข้อสอบข้อถัดไปเป็นองค์ประกอบหนึ่งที่มีความสำคัญในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เนื่องจากการได้ข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบจะทำให้ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ และความยาวของแบบทดสอบลดลง (Luecht & Sireci, 2011, pp. 9-10; Lunz et al., 1992) โดยทั่วไป การคัดเลือกข้อสอบที่มีความเหมาะสมกับผู้สอบ จะเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบ (a) สูงสุด ค่าความยากของข้อสอบ (b) ใกล้เคียงกับค่าความสามารถของผู้สอบ และค่าการเดาของข้อสอบ (c) เข้าใกล้ศูนย์ ในขณะเดียวกัน การคัดเลือกข้อสอบต้องคำนึงถึงการควบคุมการใช้ข้อสอบ (Item Exposure Control) เพื่อควบคุมไม่ให้มีการใช้ข้อสอบข้อใดข้อหนึ่งมากหรือน้อยเกินไป ซึ่งส่งผลต่อความปลอดภัยในการทดสอบ และคำนึงถึงการควบคุมสัดส่วนเนื้อหาของข้อสอบ (Content Balancing) เพื่อควบคุมกระบวนการคัดเลือกข้อสอบ ทำให้ผู้สอบทุกคนได้รับข้อสอบครบทุกหัวข้อเรื่อง รายละเอียดของการคัดเลือกข้อสอบข้อถัดไป การควบคุมการใช้ข้อสอบ และการควบคุมสัดส่วนเนื้อหาของข้อสอบ แสดงได้ดังนี้

**การคัดเลือกข้อสอบข้อถัดไป** (Next Item Selection)

วิธีการคัดเลือกข้อสอบแต่ละวิธีจะมีลักษณะ จุดเด่นจุดด้อยแตกต่างกัน ได้แก่

**วิธีการที่ 1** วิธีการคัดเลือกข้อสอบข้อถัดไปที่ระดับความยาก (b) ใกล้เคียงกับค่าความสามารถของผู้สอบ กล่าวคือ เมื่อทำการประมาณค่าความสามารถของผู้สอบแล้ว จะนำค่าความสามารถของผู้สอบมาเปรียบเทียบกับค่าความยากของข้อสอบแต่ละข้อในคลังข้อสอบ แล้วจะเลือกข้อสอบที่มีค่าความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้สอบในขณะนั้น เป็นข้อสอบข้อถัดไป โดยไม่คำนึงถึงค่าอำนาจจำแนกของข้อสอบ และค่าการเดาของข้อสอบ วิธีนี้เป็นวิธีที่ใช้กันมาตั้งแต่เริ่มต้นพัฒนาการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้แนวคิดพื้นฐานของการปรับเหมาะ นั่นคือ ข้อสอบในข้อถัดไปต้องเป็นข้อสอบที่มีค่าความยากของข้อสอบเหมาะสมกับความสามารถของผู้สอบ (Lilley, Barker, & Britton, 2004)

**วิธีการที่ 2** วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) เสนอโดย Birnbaum (1968 cited in van der Linden & Glas, 2002, pp. 9-10) วิธีการนี้เป็นวิธีการคัดเลือกข้อสอบข้อถัดไปที่ใช้อย่างกว้างขวาง โดยคำนึงถึงค่าพารามิเตอร์ของข้อสอบทั้ง 3 ค่า ได้แก่ ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ วิธีการนี้ดำเนินการโดยนำค่าพารามิเตอร์ของข้อสอบทุกข้อในคลังข้อสอบ

ที่ยังไม่ถูกเลือกใช้ในการทดสอบ มาคำนวณค่าสารสนเทศสูงสุดของข้อสอบ ($m_i$) สามารถคำนวณได้จาก

$$m_i = b_i + \frac{1}{Da_i} \ln\left[\frac{1+\sqrt{1+8c_i}}{2}\right] \quad (1)$$

เมื่อ $m_i$ คือ สารสนเทศสูงสุดของข้อสอบข้อที่ i
D คือ ค่าคงที่มีค่าเท่ากับ 1.7
$a_i$ คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i
$b_i$ คือ ค่าความยากของข้อสอบข้อที่ i
$c_i$ คือ ค่าการเดาของข้อสอบข้อที่ i

จากนั้น จะเปรียบเทียบค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น กับค่าสารสนเทศสูงสุดของข้อสอบ แต่ละข้อที่คำนวณได้จากสมการที่ 1 ข้อสอบข้อใดที่มีค่าสารสนเทศสูงสุดของข้อสอบใกล้เคียงกับค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น มากที่สุด จะถูกคัดเลือกเป็นข้อสอบข้อถัดไป

**วิธีการที่ 3** วิธีการคัดเลือกข้อสอบข้อถัดไป ด้วยใช้ค่าสารสนเทศฟิชเชอร์สูงสุด (Maximum Fisher Information: MFI) เสนอโดย Weiss (1982) วิธีการนี้จะประมาณค่าความสามารถของผู้สอบแล้ว จะนำค่าประมาณความสามารถที่ได้ไปคำนวณหาค่าสารสนเทศฟิชเชอร์และเลือกข้อสอบที่มีค่าสารสนเทศฟิชเชอร์สูงสุดมาเป็นข้อสอบข้อถัดไป ซึ่งค่าสารสนเทศฟิชเชอร์ ณ ตำแหน่งความสามารถขณะนั้น สามารถคำนวณได้จาก

$$I_i(\theta) = \frac{2.89a_i^2(1-c_i)}{[c_i+e^{1.7a_i(\theta-b_i)}][1+e^{-1.7a_i(\theta-b_i)}]^2} \quad (2)$$

เมื่อ $I_i(\theta)$ คือ สารสนเทศฟิชเชอร์ของข้อสอบข้อที่ i
$a_i$ คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i
$b_i$ คือ ค่าความยากของข้อสอบข้อที่ i
$c_i$ คือ ค่าการเดาของข้อสอบข้อที่ i

**วิธีการที่ 4** การคัดเลือกข้อสอบข้อถัดไปโดยใช้สารสนเทศคูลแบค-ไลเบลอร์ (Kullback-Leibler Information: KL) ซึ่งนำเสนอโดย Chang and Ying (1996) ด้วยการใช้ฟังก์ชันที่แปลงค่าความสามารถของผู้สอบ ($\hat{\theta}$) เป็นค่าความสามารถที่แท้จริงของผู้สอบ ($\theta$) ที่ข้อสอบ n ข้อ และใช้พื้นที่ใต้เส้นโค้งของฟังก์ชันเป็นเกณฑ์ในการคัดเลือกข้อสอบ กล่าวคือ ข้อสอบที่มีพื้นที่มากที่สุด ซึ่งมีค่าเท่ากับ $KL_i(\theta||\hat{\theta})$ จะถูกเลือกเป็นข้อสอบข้อถัดไป โดยมีฟังก์ชันดังนี้

$$KL_i(\hat{\theta}) = \int_{-\infty}^{\infty} KL_i(\theta||\hat{\theta})\, d\theta \quad (3)$$

เมื่อ $(\hat{\theta}_i||\hat{\theta}_u) = (\hat{\theta}_i - \frac{z}{\sqrt{n}},\ \hat{\theta}_i + \frac{z}{\sqrt{n}})$

$\frac{z}{\sqrt{n}}$ คือ ลิมิตความเชื่อมั่น (Confidence Limit)

**วิธีการที่ 5** วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้สารสนเทศฟิชเชอร์ที่มีการถ่วงน้ำหนัก (Fisher Interval Information: FII) นำเสนอโดย Veerkamp and Berger (1997) หรือเรียกอีกอย่างหนึ่งว่า การคัดเลือกข้อสอบข้อถัดไปโดยใช้สารสนเทศถ่วงน้ำหนักทั่วไป (General Weighted Information criterion: GWIC) เป็นการนำค่าสารสนเทศของฟิชเชอร์ที่มากเกินช่วงระดับคุณลักษณะมาคิดเป็นค่าเฉลี่ยถ่วงน้ำหนัก (Weighted Average) โดยเลือกข้อสอบที่มีค่าเฉลี่ยถ่วงน้ำหนักสูงสุดมาเป็นข้อสอบข้อถัดไป คำนวณได้ดังนี้

$$GWIC_i(\theta) = \int_{-\infty}^{\infty} W(\theta) I_i(\theta)\, d\theta \qquad (4)$$

เมื่อ $W(\theta)$ คือ ฟังก์ชันถ่วงน้ำหนัก

$I_i(\theta)$ คือ สารสนเทศของฟิชเชอร์

**วิธีการที่ 6** วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้สารสนเทศคูลแบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (Kullback-Leibler Information with a Posterior Distribution: KLP) นำเสนอโดย Veldkamp and van der Linden (2002) เป็นฟังก์ชันที่แก้ไขฟังก์ชันสารสนเทศฟิชเชอร์ในค่าสารสนเทศคูลเบค-ไลเบลอร์ $KL_i(\theta||\hat{\theta})$ โดยเลือกข้อสอบที่มีค่า KLP สูงสุด ตามสมการต่อไปนี้

$$(\hat{\theta}_n, U_n) = \int_{-\infty}^{\infty} P(\theta|U_n) KL_i(\theta||\hat{\theta}_n)\, d\theta \qquad (5)$$

เมื่อ $P(\theta|U_n)$ คือ การแจกแจงคุณลักษณะภายหลัง หลังจากที่ข้อสอบ n ข้อถูกใช้สอบ

$(\hat{\theta}_i||\hat{\theta}_u) = (\hat{\theta}_i - \frac{z}{\sqrt{n}},\ \hat{\theta}_i + \frac{z}{\sqrt{n}})$

$\frac{z}{\sqrt{n}}$ คือ ลิมิตความเชื่อมั่น (Confidence Limit)

**วิธีการที่ 7** วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้สารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (Fisher Information with a Posterior Distribution: FIP) นำเสนอโดย Chen and Ankenmann (2004) เป็นการกำหนดการแจกแจงหลังจากหาค่าถ่วงน้ำหนัก ซึ่งเป็นฟังก์ชันที่เหมาะสมในการพิจารณาคุณลักษณะทั้งหมด โดยที่เลือกข้อสอบที่มีค่า FIP สูงสุด มาเป็นข้อสอบข้อถัดไป ตามสมการต่อไปนี้

$$FIP_i(X_n) = GWIC_i(X_n) = \int_{-\infty}^{\infty} P(\theta|x_n) I_i(\theta)\, d\theta \qquad (6)$$

เมื่อ $P(\theta|x_n)$ คือ การแจกแจงคุณลักษณะภายหลัง หลังจากที่ข้อสอบ n ข้อถูกใช้สอบ

$I_i(\theta)$ คือ สารสนเทศของฟิชเชอร์

**วิธีการที่ 8** วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ประสิทธิภาพของข้อมูล (Efficiency Balanced Information: EBI) นำเสนอโดย Han (2012) เป็นวิธีที่เลือกข้อสอบข้อถัดไปจากการประเมินช่วงของค่าความสามารถของผู้สอบ โดยประเมินจากประสิทธิภาพของค่าสารสนเทศ ซึ่งหาได้จากค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ และการกำหนดค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถที่คาดว่าจะเกิดขึ้นจากค่าความสามารถของผู้สอบหลังจากผู้สอบตอบข้อสอบข้อที่ j โดยเลือกข้อสอบที่มีค่า EBI สูงสุดเป็นข้อสอบข้อถัดไป คำนวณได้ตามสมการต่อไปนี้

$$EBI_i(\hat{\theta}) = \left(1+\frac{1}{I_i(\hat{\theta})}\right)\int_{\hat{\theta}-2\varepsilon_j}^{\hat{\theta}+2\varepsilon_j} I_i(\theta)\,d\theta \qquad (7)$$

**วิธีการที่ 9** วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง นำเสนอโดย โสพส สุขานนท์สวัสดิ์ และคณะ (2556) เป็นวิธีการคัดเลือกข้อสอบข้อถัดไปที่ใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (Decision Making Under Risk) ตามเกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุด (Maximum expected Monetary Value: EMV) โดยเลือกข้อสอบข้อที่มีค่าสัมประสิทธิ์การกระจาย (Coefficient of Variation) สูงสุดมาเป็นข้อสอบข้อถัดไป ซึ่งการคำนวณค่าความคาดหวังสูงสุด จะคำนวณโดยนำค่าประมาณความสามารถของผู้สอบก่อนที่จะตอบข้อสอบที่มีอยู่ในคลังข้อสอบ ทั้งในกรณีที่คาดว่าผู้สอบจะตอบข้อสอบถูก และกรณีที่คาดว่าผู้สอบจะตอบข้อสอบผิด มาพิจารณาคัดเลือกข้อสอบข้อถัดไปร่วมกัน โดยไม่นำข้อสอบที่ใช้ไปแล้วมาพิจารณา คำนวณได้ตามสมการต่อไปนี้

$$\mathrm{EMV_i}(\theta)=(\mathrm{P_i}(\theta)\times\theta_{\mathrm{true}})+(\mathrm{Q_i}(\theta)\times\theta_{\mathrm{false}}) \qquad (8)$$

เมื่อ $\mathrm{EMV_i}(\theta)$ คือ ค่าคาดหวังสูงสุดของข้อสอบข้อที่ i

$\mathrm{P_i}(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อที่ i ถูก

$\mathrm{Q_i}(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อที่ i ผิด

$\theta_{\mathrm{true}}$ คือ ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ถูก

$\theta_{\mathrm{false}}$ คือ ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ผิด

หลังจากหาค่า EMV จากสมการที่ 8 แล้ว ให้หาค่าการกระจาย คำนวณจาก

$$\sigma_{\mathrm{i(EMV)}}(\theta)=\sqrt{((\theta_{\mathrm{true}}-\mathrm{EMV_i}(\theta))^2\times\mathrm{P_i}(\theta))+((\theta_{\mathrm{false}}-\mathrm{EMV_i}(\theta))^2\times\mathrm{Q_i}(\theta))} \qquad (9)$$

เมื่อ $\sigma_{\mathrm{i(EMV)}}(\theta)$ คือ ค่าการกระจายของข้อสอบข้อที่ i

$EMV_i(\theta)$ คือ ค่าคาดหวังสูงสุดของข้อสอบข้อที่ i
$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อที่ i ถูก
$Q_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อที่ i ผิด
$\theta_{true}$ คือ ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ถูก
$\theta_{false}$ คือ ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ผิด

หลังจากหาค่าการกระจายจากสมการที่ 9 แล้ว ให้หาค่าสัมประสิทธิ์การกระจาย คำนวณได้ดังนี้

$$CV_i(\theta) = \left|\frac{\sigma_{i(EMV)}(\theta)}{EMV_i(\theta)}\right| \quad (10)$$

เมื่อ $CV_i(\theta)$ คือ ค่าสัมประสิทธิ์การกระจายของข้อสอบข้อที่ i
$\sigma_{i(EMV)}(\theta)$ คือ ค่าการกระจายข้อมูลของข้อสอบข้อที่ i
$EMV_i(\theta)$ คือ ค่าความคาดหวังสูงสุดของข้อสอบข้อที่ i

วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง จะคัดเลือกข้อสอบที่มีค่าสัมประสิทธิ์การกระจาย (จากสมการที่ 10) มากที่สุด มาเป็นข้อสอบข้อถัดไป

**การควบคุมการใช้ข้อสอบ** (Item Exposure Control)

การใช้ข้อสอบ (Item Exposure) คือ การจัดข้อสอบที่ผ่านการคัดเลือกแล้วให้ผู้สอบ โดยเป็นส่วนหนึ่งของขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ขั้นตอนการคัดเลือกข้อสอบข้อถัดไป จะคัดเลือกข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบ ซึ่งเป็นข้อสอบที่ให้ค่าสารสนเทศสูงสุด ณ ระดับความสามารถของผู้สอบขณะนั้น

ค่าสารสนเทศของข้อสอบจะมีค่าเพิ่มขึ้น ใน 2 กรณี คือ 1) ค่าอำนาจจำแนกของข้อสอบมีค่าเพิ่มขึ้น หรือ 2) ค่าความยากของข้อสอบมีค่าใกล้เคียงกับค่าความสามารถของผู้สอบ ด้วยเหตุนี้ ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ส่วนใหญ่จะคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ไปใช้ในการทดสอบ เป็นผลให้ข้อสอบบางข้อถูกนำไปใช้ในการทดสอบบ่อยครั้งเกินไป (Over-Exposure Item) ขณะเดียวกัน ข้อสอบบางข้อถูกนำไปใช้ในการทดสอบน้อย หรือไม่เคยถูกนำไปใช้ในการทดสอบ (Under-Utilized Item) การมีข้อสอบที่ถูกนำไปใช้ในการทดสอบบ่อยครั้งเกินไป จะส่งผลกระทบต่อความปลอดภัยของการทดสอบ (Test Security) กล่าวคือ เมื่อมีข้อสอบที่ถูกนำไปใช้ในการทดสอบบ่อยครั้งเกินไป จนรู้กันในกลุ่มผู้สอบ เป็นผลทำให้ผู้สอบในรุ่นถัดมาสามารถคาดการณ์ล่วงหน้าได้ว่าตนเองจะได้ข้อสอบข้อใด ทำให้สามารถตอบข้อสอบข้อดังกล่าวได้โดยไม่ได้ใช้ความสามารถของตนเอง คะแนนสอบจึงขาดความถูกต้อง ผู้สอบสามารถทำคะแนนได้มากเกินความสามารถของตนเอง ซึ่งส่งผลกระทบต่อความปลอดภัยของการทดสอบและประสิทธิภาพของการใช้ข้อสอบในคลังข้อสอบ (Pool Utilization) ขณะเดียวกัน กรณีที่ข้อสอบถูก

จัดให้ผู้สอบใช้น้อยหรือไม่ถูกจัดให้ผู้สอบใช้เลย ทำให้เสียค่าใช้จ่ายในการจัดเตรียมคลังข้อสอบโดยเปล่าประโยชน์

ดังนั้นจึงมีการศึกษาเกี่ยวกับวิธีการควบคุมการใช้ข้อสอบหลายวิธี โดย Georgiadou et al. (2007) ได้จำแนกวิธีการควบคุมการใช้ข้อสอบ ได้ 5 ประเภท ดังนี้

**1. วิธีการคัดเลือกอย่างสุ่ม** (Randomization Strategies)

วิธีการควบคุมการใช้ข้อสอบในกลุ่มนี้จะทำการคัดเลือกข้อสอบที่เหมาะสมมา 1 กลุ่ม หลังจากนั้นจะทำการสุ่มเลือกข้อสอบข้อใดข้อหนึ่งจากกลุ่มข้อสอบนี้มาให้ผู้สอบใช้ จึงมีการดำเนินการค่อนข้างง่าย ไม่ซับซ้อน แต่ไม่รับรองว่าจะสามารถควบคุมการใช้ข้อสอบได้ตามต้องการวิธีการในกลุ่มนี้เป็นวิธีการคัดเลือกข้อสอบข้อถัดไป พร้อมทั้งเป็นการควบคุมการใช้ข้อสอบอีกด้วย วิธีการในกลุ่มนี้ได้แก่

**วิธีการที่ 1** วิธี 5-4-3-2-1 เป็นวิธีที่เสนอโดย McBride and Martin (1983) วิธีนี้ในขั้นแรกจะเลือกข้อสอบที่มีค่าสารสนเทศสูงสุดมา 5 ข้อ ข้อสอบข้อแรกที่จะจัดให้ผู้สอบจะเป็นข้อสอบที่สุ่มจากข้อสอบที่มีค่าสารสนเทศสูงสุด 5 ข้อนี้ ถัดมาข้อสอบข้อที่ 2 จะสุ่มเลือกมาจากข้อสอบที่มีค่าสารสนเทศสูงสุดที่เหลืออีก 4 ข้อ ซึ่งจะดำเนินการลักษณะนี้ จนถึงข้อที่ 5 หลังจากข้อที่ 5 ก็จะเลือกข้อสอบที่ให้ค่าสารสนเทศสูงสุด ณ ค่าประมาณความสามารถขณะนั้นมาใหม่อีก 5 ข้อ แล้วการจัดข้อสอบข้อที่ 6 ก็จะดำเนินการตั้งแต่ขั้นตอนแรกอีกครั้ง ซึ่งกระบวนการนี้จะกระทำจนกระทั่งสิ้นสุดการทดสอบ ข้อดีของวิธีการนี้เป็นกระบวนการที่ไม่ยุ่งยาก แต่มีข้อเสีย คือ อาจจะไม่สามารถควบคุมการใช้ข้อสอบได้ดีพอ

**วิธีการที่ 2** วิธี Randomesque Strategy เสนอโดย Kingsbury and Zara (1989) เป็นวิธีการที่คล้ายกับวิธี 5-4-3-2-1 โดยเริ่มแรกจะเลือกข้อสอบที่มีค่าสารสนเทศสูงสุดมาเป็นจำนวนที่สุ่มจากเลขไม่เกิน 10 ข้อ หลังจากนั้นจะสุ่มเลือกข้อสอบเพียงข้อเดียวจากกลุ่มข้อสอบที่มีค่าสารสนเทศสูงสุดนี้ มาเป็นข้อสอบข้อถัดไปให้กับผู้สอบ โดยวิธีการนี้จะช่วยลดอัตราการใช้ข้อสอบของกลุ่มผู้สอบที่มีความสามารถใกล้เคียงกัน

**วิธีการที่ 3** วิธี Within 0.1 Logits Strategy เสนอโดย Lunz and Stahl (1998) เป็นวิธีควบคุมการใช้ข้อสอบ โดยสนใจกลุ่มข้อสอบที่มีค่าความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้สอบ กล่าวคือ ในขั้นแรกวิธีการนี้จะเลือกข้อสอบที่มีค่าความยากของข้อสอบเท่ากับค่าความสามารถของผู้สอบขณะนั้น แต่ถ้าเลือกไม่ได้ จะสุ่มเลือกข้อสอบข้อใดข้อหนึ่งจากกลุ่มข้อสอบที่มีค่าความยากของข้อสอบแตกต่างจากค่าความสามารถของผู้สอบขณะนั้น ไม่เกิน 0.1

**2. วิธีการคัดเลือกอย่างมีเงื่อนไข** (Conditional Selection Strategies)

วิธีการควบคุมการใช้ข้อสอบกลุ่มนี้ เป็นวิธีการควบคุมความน่าจะเป็นของข้อสอบที่ได้รับเลือกแล้วให้มีการใช้ข้อสอบเป็นไปตามที่กำหนด วิธีนี้รับรองว่าอัตราการใช้ข้อสอบจะไม่สูงเกินกว่าค่าที่กำหนด โดยควบคุมอัตราการใช้ข้อสอบผ่านค่าพารามิเตอร์ควบคุมการใช้ข้อสอบ (Exposure Control Parameter) แต่มีข้อเสียคือ วิธีการมีความยุ่งยากซับซ้อนมาก เนื่องจากการคำนวณค่าพารามิเตอร์ควบคุมการใช้ข้อสอบ ต้องทำในสถานการณ์จำลองที่มีการทำซ้ำจำนวนมาก และเมื่อองค์ประกอบต่าง ๆ ในการทดสอบเปลี่ยนแปลงไป เช่น มีการเพิ่ม-ลดข้อสอบในคลังข้อสอบ จำเป็นต้องคำนวณค่าพารามิเตอร์ควบคุมการใช้ข้อสอบใหม่ทุกครั้ง วิธีการในกลุ่มนี้ ได้แก่

**วิธีการที่ 1** วิธีการซิมสันเฮกเตอร์ (Sympson-Hetter: SH) เสนอโดย Sympson and Hetter (1985) ได้เสนอวิธีการควบคุมการใช้ข้อสอบโดยใช้ความถี่ของข้อสอบที่จัดให้กับกลุ่มผู้สอบจำนวนมาก โดยสร้างขึ้นในสถานการณ์จำลองมาเปรียบเทียบกับอัตราการใช้ข้อสอบเป้าหมาย เพื่อหาค่าพารามิเตอร์ควบคุมการใช้ข้อสอบ ขบวนการนี้ทำซ้ำ ๆ กันจนกระทั่งค่าพารามิเตอร์ควบคุมการใช้ข้อสอบทุกข้อจะมีค่าระหว่าง 0 กับ 1 ซึ่งพารามิเตอร์นี้จะใช้สำหรับการคัดเลือกข้อสอบในการทดสอบสถานการณ์จริง โดยสร้างตัวเลขสุ่มจากการแจกแจงเหมือนกันเปรียบเทียบกับค่าพารามิเตอร์ควบคุมการใช้ข้อสอบ ถ้าค่าพารามิเตอร์ควบคุมการใช้ข้อสอบมีค่ามากกว่าตัวเลขสุ่มจะจัดข้อสอบให้แก่ผู้สอบ แต่ถ้าค่าพารามิเตอร์ควบคุมการใช้ข้อสอบมีค่าน้อยกว่าตัวเลขสุ่มจะทำการคัดเลือกข้อสอบใหม่แทน ซึ่งผลจากการศึกษา พบว่า วิธีนี้สามารถควบคุมการใช้ข้อสอบได้อย่างมีประสิทธิภาพ

**วิธีการที่ 2** วิธี Extended Sympson-Hetter Strategy (ESH) เสนอโดย Stocking (1993) ได้ปรับปรุงวิธีการซิมสันเฮกเตอร์ โดยกำหนดค่าพารามิเตอร์ควบคุมการใช้ข้อสอบจากการประมาณค่าระดับความสามารถของผู้สอบ แทนการแจกแจงค่าความสามารถของผู้สอบที่เป็นของวิธีการเดิม ซึ่งวิธีการนี้จะทำให้ข้อสอบแต่ละข้อมีค่าพารามิเตอร์ควบคุมการใช้ข้อสอบที่แต่ละค่าความสามารถของผู้สอบ และผลจากการทดสอบ ปรากฏว่า วิธีการที่ปรับปรุงใหม่นี้สามารถควบคุมการใช้ข้อสอบได้อย่างมีประสิทธิภาพ

**วิธีการที่ 3** วิธี Davey-Parshall Strategy (DP) เสนอโดย Davey and Parshall (1995) ได้นำเสนอวิธีการใหม่ที่ใช้ควบคุมการใช้ข้อสอบ โดยมีแนวคิดคล้ายกับวิธีการซิมสันเฮกเตอร์ วิธีการนี้เน้นลดการทับซ้อนของข้อสอบระหว่างกลุ่มผู้สอบ โดยสร้างพารามิเตอร์ควบคุมการใช้ข้อสอบ และสร้างเงื่อนไขค่าพารามิเตอร์ในเมทริกซ์เฉียง โดยควบคุมความถี่ในการใช้ข้อสอบจากการเปรียบเทียบเป็นรายคู่ในสถานการณ์จำลอง เพื่อควบคุมการใช้ข้อสอบ

**วิธีการที่ 4** วิธี Restricted Maximum Information Strategy เสนอโดย Revuelta and Ponsoda (1998) วิธีการนี้เป็นวิธีการควบคุมการใช้ข้อสอบ โดยกำหนดอัตราการใช้ข้อสอบสูงสุดไว้ล่วงหน้า ข้อสอบข้อใดที่มีอัตราการใช้ข้อสอบสูงกว่าอัตราการใช้ข้อสอบสูงสุดที่กำหนดไว้จะไม่นำมาพิจารณาในการคัดเลือกข้อสอบข้อถัดไป วิธีการนี้จะใช้วิธีการคัดเลือกข้อสอบที่มีค่าสารสนเทศสูงสุดใกล้เคียงกับค่าประมาณความสามารถของผู้สอบ (MIC) ควบคู่ด้วย

**3. วิธีการคัดเลือกตามระดับชั้น** (Stratified Strategies)

วิธีการควบคุมการใช้ข้อสอบในกลุ่มนี้ มีแนวคิดโดยแบ่งข้อสอบในคลังข้อสอบเป็นชั้น ๆ ตามค่าอำนาจจำแนกของข้อสอบ และควบคุมการใช้ข้อสอบโดยให้เลือกข้อสอบจากทุกชั้นมาใช้ในการทดสอบ วิธีการนี้จะเพิ่มโอกาสให้ข้อสอบที่มีค่าสารสนเทศต่ำมีโอกาสถูกเลือกมาใช้สอบเพิ่มมากขึ้น วิธีการนี้ได้แก่

**วิธีการที่ 1** a-Stratified Strategy (ASTR) เสนอโดย Chang and Ying (1999) เป็นวิธีการที่แบ่งคลังข้อสอบออกเป็นชั้น ๆ ตามค่าอำนาจจำแนกของข้อสอบ กล่าวคือ ข้อสอบที่อยู่ในชั้นที่ 1 จะเป็นข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบต่ำสุด ส่วนข้อสอบที่อยู่ในชั้นถัดไปจะเป็นข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงกว่าข้อสอบในชั้นที่ 1 ถัดไปเรื่อย ๆ จนถึงข้อสอบในชั้นสุดท้ายจะเป็นข้อสอบที่ค่าอำนาจจำแนกของข้อสอบสูงสุด โดยผู้สอบจะต้องทำข้อสอบในทุกชั้น ซึ่งข้อสอบที่

ได้รับในแต่ละชั้นจะเป็นข้อสอบที่มีค่าค่าความยากของข้อสอบใกล้เคียงกับค่าประมาณความสามารถของผู้สอบ

**วิธีการที่ 2** a-Stratified with b-Blocking (BASTR) เสนอโดย Chang, Qian, and Ying (2001) เป็นวิธีการที่ปรับปรุงจากวิธีการ a-Stratified strategy (ASTR) กล่าวคือ เริ่มแรกจะแบ่งคลังข้อสอบออกเป็นชั้น ๆ ตามค่าความยากของข้อสอบ ชั้นแรกเป็นข้อสอบที่มีค่าความยากของข้อสอบต่ำสุด ถัดไปค่าความยากของข้อสอบในแต่ละชั้นจะสูงขึ้น จนชั้นสุดท้ายจะเป็นข้อสอบที่มีค่าความยากของข้อสอบสูงสุด แล้วในแต่ละชั้นจะแบ่งข้อสอบออกเป็นกลุ่ม ๆ ตามค่าอำนาจจำแนกของข้อสอบ นั่นคือ กลุ่มแรกของแต่ละชั้นจะเป็นข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบต่ำสุด กลุ่มถัดไปจะมีค่าอำนาจจำแนกของข้อสอบสูงขึ้นเรื่อย ๆ จนกลุ่มสุดท้ายในแต่ละชั้นจะมีค่าอำนาจจำแนกของข้อสอบสูงสุด โดยผู้สอบจะได้ข้อสอบในชั้นที่มีค่าความยากของข้อสอบที่เหมาะสมกับตัวผู้สอบ การเลื่อนไปทำข้อสอบในชั้นถัดไปจะต้องเป็นไปตามเงื่อนไขที่กำหนด

**4. วิธีการคัดเลือกแบบผสม** (Combined Strategies)

วิธีการควบคุมการใช้ข้อสอบในกลุ่มนี้ เป็นการนำวิธีการควบคุมการใช้ข้อสอบตั้งแต่ 2 วิธีการขึ้นไป มาใช้ร่วมกัน ซึ่งจะนำจุดเด่นของแต่ละวิธีมาใช้ อย่างเช่น วิธี Progressive Restricted เสนอโดย Revuelta and Ponsoda (1998) ซึ่งวิธีการนี้จะควบคุมการคัดเลือกใช้ข้อสอบในคลังข้อสอบให้มีความสมดุล โดยนำวิธีการ Restricted Maximum Information ซึ่งเป็นวิธีที่สามารถลดอัตราการใช้ข้อสอบของข้อสอบที่ถูกเลือกใช้บ่อยให้มีค่าต่ำลงได้ มาใช้ร่วมกับวิธีการ Progressive ซึ่งเป็นวิธีที่เพิ่มอัตราการใช้ข้อสอบของข้อสอบที่ถูกเลือกใช้น้อยให้มีค่าสูงขึ้นได้

**5. วิธีการคัดเลือกแบบปรับเหมาะหลายขั้นตอน** (Multiple Stage Adaptive Test Designs)

วิธีการในกลุ่มนี้ถูกพัฒนาขึ้น โดยรวมลักษณะของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์และการทดสอบแบบดั้งเดิม (เขียนตอบบนกระดาษ) เข้าด้วยกัน คือ จะใช้กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในการหาค่าความสามารถของผู้สอบ และใช้ผลที่ได้จากการทดสอบเบื้องต้นบนกระดาษหรือการทดสอบคู่ขนาน (Parallel Test) มาเป็นข้อมูลในการคัดเลือกข้อสอบข้อถัดไปที่เหมาะสมกับผู้สอบ วิธีการที่ใช้แนวคิดนี้ ได้แก่ วิธี Computerize Adaptive Sequential Testing (CAST) ซึ่งเสนอโดย Luecht et al. (1996)

**การควบคุมสัดส่วนเนื้อหาของข้อสอบ** (Content Balancing)

การกำหนดคุณลักษณะของข้อสอบ (Content Specification) เป็นสิ่งสำคัญอย่างยิ่ง โดยเฉพาะการวัดผลสัมฤทธิ์ทางการศึกษา ในการทดสอบแบบดั้งเดิม แบบทดสอบที่เป็นมาตรฐานส่วนใหญ่ถูกสร้างตามรายละเอียดในตารางการกำหนดคุณลักษณะของข้อสอบ (Table of Content Specification) ซึ่งผู้สอบทุกคนจะได้แบบทดสอบชุดเดียวกัน ประกอบด้วยข้อสอบครบทุกคุณลักษณะที่ต้องการวัดตามที่ออกแบบไว้ ด้วยเหตุนี้ จึงเป็นข้อจำกัดในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เนื่องจากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จะคัดเลือกข้อสอบที่ให้ค่าสารสนเทศสูงสุดที่ค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น ซึ่งไม่ได้พิจารณาขอบเขตเนื้อหา (Content Area) ของข้อสอบ เป็นผลให้ผู้สอบบางคนไม่ได้รับข้อสอบครบทุกขอบเขตเนื้อหาของข้อสอบ ยกตัวอย่างเช่น ในการวัดความสามารถด้านการคำนวณ การบวก การลบ การคูณ และ

การหาร ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผู้สอบอาจได้รับข้อสอบเฉพาะ การบวกและการลบ แต่ไม่ได้รับข้อสอบการคูณและการหาร ส่งผลให้ผู้ที่ไม่มีความรู้เรื่องการคูณ และการหาร จะได้ผลการประเมินความสามารถสูงเกินจริง ในขณะเดียวกัน ผู้สอบที่มีความรู้ในเรื่องการคูณและการหาร จะถูกประเมินความสามารถต่ำเกินจริง ดังนั้น เพื่อรับประกันว่าผู้สอบแต่ละคนจะได้ข้อสอบที่เป็นตัวแทนตามสัดส่วนของแต่ละขอบเขตเนื้อหาครบถ้วน และเพื่อให้แน่ใจว่าผู้สอบทุกคนได้รับข้อสอบในแต่ละเนื้อหาเท่ากัน ในการคัดเลือกข้อสอบข้อถัดไปจึงต้องคำนึงถึงการควบคุมสัดส่วนเนื้อหาของข้อสอบ (Wainer et al., 2000, pp. 120-121; Thompson & Weiss, 2011)

Kingsbury and Zara (1989) ได้เสนอวิธีการควบคุมสัดส่วนเนื้อหาของข้อสอบ ซึ่งวิธีการนี้จะเข้าแทรกแซงกระบวนการการคัดเลือกข้อสอบข้อถัดไป เพื่อให้ผู้สอบได้รับข้อสอบครบทุกหัวข้อเรื่อง และมีจำนวนข้อสอบในแต่ละหัวข้อเรื่องเป็นไปตามสัดส่วนที่ระบุ วิธีการ Kingsbury-Zara มีรายละเอียดดังนี้

ขั้นที่ 1 กำหนดสัดส่วนของข้อสอบที่ต้องการในแต่ละหัวข้อเรื่อง

ขั้นที่ 2 คำนวนสัดส่วนของข้อสอบในแต่ละหัวข้อเรื่องที่ผู้สอบได้รับ โดยคำนวณจากจำนวนข้อสอบในแต่ละหัวข้อเรื่องที่ผู้สอบได้รับหารด้วยจำนวนข้อสอบทั้งหมดในหัวข้อเรื่องนั้น คูณ 100

ขั้นที่ 3 คำนวนผลต่างของสัดส่วนที่ได้จากขั้นที่ 1 กับขั้นที่ 2

ขั้นที่ 4 คัดเลือกจากข้อสอบในหัวข้อเรื่องที่มีผลต่างมากสุด เป็นข้อสอบข้อถัดไป

(แสดงขั้นตอนการควบคุมสัดส่วนเนื้อหาของข้อสอบ ตามวิธีของ Kingsbury-Zara ในหน้า 84)

**2.4 การประมาณค่าความสามารถของผู้สอบ** (Calculate Possible Ability Levels)

วิธีประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบที่นิยมใช้มี 3 วิธี คือ 1) วิธีของเบส์ 2) วิธีของเบส์แบบปรับใหม่ และ 3) วิธีความน่าจะเป็นแบบมีเงื่อนไข ซึ่งทั้งสามวิธีนี้ให้ประสิทธิภาพการประมาณค่าความสามารถของผู้สอบสูง และมีรายละเอียดของแต่ละวิธีดังนี้

**2.4.1 วิธีของเบส์** (Bayesian)

การประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ จะอาศัยทฤษฎีของเบส์ (Bayes's Theorem) คำนวณได้จาก

$$f(\theta/U) = K.L(U/\theta).f(\theta) \qquad (11)$$

เมื่อ $f(\theta/U)$ คือ Posterior Distribution ของ $\theta$

$L(U/\theta)$ คือ Likelihood Function ของ vector U (item response)

$f(\theta)$ คือ Prior Distribution ของ $\theta$

K คือ Constant

**2.4.2 วิธีเบส์แบบปรับใหม่** (Bayesian Updating) เป็นวิธีการประมาณค่าความสามารถของผู้สอบ ที่นำเสนอโดย Owen (1975) และเป็นวิธีการที่ได้รับความนิยมนำไปใช้อย่างแพร่หลายเนื่องจากประมาณค่าความสามารถของผู้สอบได้ค่อนข้างคงที่

กรณีที่ตอบข้อสอบถูก คำนวณได้ดังสมการต่อไปนี้

$$\theta_{m+1} = \theta_m + (1-c)\left(\frac{\sigma_m^2}{\sqrt{\frac{1}{a^2}+\sigma_m^2}}\right)\left(\frac{O(D)}{c+(1-c)A(-D)}\right) \tag{12}$$

$$\sigma_{m+1}^2 = \sigma_m^2\left(1-\left(\frac{1-c}{1+\frac{1}{a^2\sigma_m^2}}\right)\left(\frac{O(D)}{B}\right)\left(\frac{(1-c)O(D)}{B}-D\right)\right) \tag{13}$$

กรณีที่ตอบข้อสอบผิด สามารถคำนวณได้ดังสมการต่อไปนี้

$$\theta_{m+1} = \theta_m - \left(\frac{\sigma_m^2}{\sqrt{\frac{1}{a^2}+\sigma_m^2}}\right)\left(\frac{O(D)}{A(D)}\right) \tag{14}$$

$$\sigma_{m+1}^2 = \sigma_m^2\left(1-\left(\frac{O(D)}{1+\frac{1}{a^2\sigma_m^2}}\right)\left(\frac{O(D)}{A(D)}+D\right)\div A(D)\right) \tag{15}$$

เมื่อ
$$D = \frac{b-\theta_m}{\sqrt{\frac{1}{a^2}+\sigma_m^2}} \tag{16}$$

$$B = c+(1-c)(A(-D)) \tag{17}$$

เมื่อ $\theta_m$ คือ ค่าความสามารถของผู้สอบที่ประมาณได้ก่อนตอบข้อสอบ ข้อที่ m+1 ตามปกติแล้ว เมื่อเริ่มทำข้อสอบจะไม่ทราบค่าความสามารถเบื้องต้นของผู้สอบ จึงกำหนดให้ $\theta_m$ เท่ากับ 0.000

$\sigma_m^2$ คือ ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อสอบข้อที่ m+1 ตามปกติแล้ว เมื่อเริ่มทำข้อสอบจะไม่ทราบค่าความแปรปรวนดังกล่าวมาก่อน จึงกำหนดให้มีค่าเท่ากับ 1.000

$\theta_{m+1}$ คือ ค่าความสามารถของผู้สอบที่ประมาณค่าความสามารถของผู้สอบหลังจากตอบข้อสอบข้อที่ m+1

$\sigma_{m+1}^2$ คือ ค่าความแปรปรวนในการประมาณค่าความสามารถของผู้สอบที่ประมาณค่าความสามารถของผู้สอบหลังจากตอบข้อสอบข้อที่ m+1

$a_i$ คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ m+1

$b_i$ คือ ค่าความยากของข้อสอบข้อที่ m+1

$c_i$ คือ ค่าการเดาของข้อสอบข้อที่ m+1

D คือ จุดบนแกน x

O(D) คือ ค่าออร์ดิเนท (Ordinate) ของโค้งปกติที่จุด D

A(D) คือ พื้นที่ใต้โค้งปกติจากค่า D ต่ำสุดจนถึงจุด D

**2.4.3 วิธีความน่าจะเป็นแบบมีเงื่อนไข** (Conditional Maximum Likelihood)

วิธีการประมาณค่าความสามารถของผู้สอบ ด้วยวิธีความน่าจะเป็นสูงสุด (Maximum Likelihood: ML) มีอยู่หลายวิธี แต่วิธีที่ได้รับนิยมมากที่สุด คือ วิธีความน่าจะเป็นแบบมีเงื่อนไข (Conditional ML) การประมาณค่าความสามารถโดยวิธีนี้ มีข้อจำกัด คือ ถ้าผู้สอบตอบข้อสอบถูกหมด หรือผิดหมด จะไม่สามารถประมาณค่าความสามารถได้ (Hambleton & Swaminathan, 1985, pp. 81-88) ขั้นตอนในการประมาณค่าความสามารถ สำหรับการทดสอบแบบปรับเหมาะมีดังนี้

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

สูตรประมาณค่า $\theta$ แสดงดังสมการต่อไปนี้

$$\theta_0 = \ln\left(\frac{r_a}{k-r_a}\right) \qquad (18)$$

เมื่อ $r_a = \sum a_i u_i$

$u_i$ = 1 เมื่อตอบข้อสอบถูก และ $u_i$ = 0 เมื่อตอบข้อสอบผิด

$a_i$ = ค่าอำนาจจำแนกของข้อสอบข้อที่ i

k = จำนวนข้อสอบทั้งหมดที่มีผู้สอบตอบ

ขั้นที่ 2 หาค่า $P_i(\theta_m)$ และ $Q_i(\theta_m)$ จาก

$$P_i(\theta_m) = c_i + (1-c_i)\frac{e^{Da_i(\theta_m-b_i)}}{Da_i(\theta_m-b_i)} = \frac{1-c_i}{1+e^{Da_i(\theta_m-b_i)}} \qquad (19)$$

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

เมื่อ $\theta_m$ คือ ค่าประมาณความสามารถของผู้สอบในครั้งที่ m

D คือ ค่าคงที่มีค่าเท่ากับ 1.7

$a_i$ คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

$b_i$ คือ ค่าความยากของข้อสอบข้อที่ i

$c_i$ คือ ค่าการเดาของข้อสอบข้อที่ i

ขั้นที่ 3 หาค่าปรับแก้ $h_m$ ดังสมการต่อไปนี้

$$h_m = \frac{D\left[r_a - \sum P_i(\theta_m)\right]}{-D^2 \sum P_i(\theta_m) Q_i(\theta_m)} \quad (20)$$

ขั้นที่ 4 ประมาณค่าความสามารถของผู้สอบใหม่ $\theta_{m+1}$ ได้ดังสมการต่อไปนี้

$$\theta_{m+1} = \theta_m - h_m \quad (21)$$

ขั้นที่ 5 ประมาณค่า $\theta$ ซ้ำจนบรรลุตามเงื่อนไข

ทำการคำนวณซ้ำในขั้นที่ 2, 3 และ 4 จนกระทั่ง $h_m$ เข้าใกล้ศูนย์ ($h_m < 0.001$) หรือบรรลุผลตามเงื่อนไขของเกณฑ์ยุติการทดสอบ

**2.5 เกณฑ์ยุติการทดสอบ** (Termination Criterion)

เกณฑ์ยุติการทดสอบในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีให้เลือกใช้หลากหลายดังนี้ (de Ayala, 2009, pp. 378-379; Thompson & Weiss, 2011; Wainer et al., 2000, p. 112)

2.5.1 เกณฑ์ยุติการทดสอบเมื่อผู้สอบใช้ข้อสอบในคลังข้อสอบครบทุกข้อ

2.5.2 เกณฑ์ยุติการทดสอบแบบความยาวคงที่ (Fixed-Length) ผู้สอบทุกคนจะได้ข้อสอบจำนวนเท่ากัน ซึ่งจะกำหนดจำนวนข้อสอบไว้ล่วงหน้า เช่น กำหนดให้การทดสอบยุติลง เมื่อผู้สอบทำข้อสอบครบ 25 ข้อ

2.5.3 เกณฑ์ยุติการทดสอบแบบความยาวยืดหยุ่น (Variable-Length) ผู้สอบแต่ละคนจะได้ข้อสอบจำนวนไม่เท่ากัน ซึ่งเลือกพิจารณาจากค่าดังต่อไปนี้

2.5.3.1 ค่าประมาณความสามารถของผู้สอบ การทดสอบจะยุติเมื่อค่าประมาณความสามารถของผู้สอบในแต่ละครั้งมีการเปลี่ยนแปลงน้อยมาก หรือไม่เปลี่ยนแปลงเลย

2.5.3.2 ค่าสารสนเทศของข้อสอบ (Item Information) การทดสอบจะยุติเมื่อคลังข้อสอบเหลือแต่ข้อสอบที่มีค่าสารสนเทศของข้อสอบน้อยกว่าค่าที่กำหนด เช่น การทดสอบจะดำเนินต่อไป จนกระทั่งคลังข้อสอบเหลือเฉพาะข้อสอบที่มีค่าสารสนเทศของข้อสอบน้อยกว่า 0.90 ซึ่งเกณฑ์ยุติการทดสอบนี้จะเรียกว่า เกณฑ์สารสนเทศน้อยสุด (Minimum Information Criterion)

2.5.3.3 ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ (Standard Error of Ability Estimate: SEE) โดยผู้สอบจะได้รับข้อสอบ จนกว่าค่า SEE มีค่าน้อยกว่าค่าที่กำหนด เช่น การทดสอบจะดำเนินต่อไป จนกระทั่งค่า SEE มีค่าน้อยกว่า 0.30 ซึ่งเกณฑ์ยุติการทดสอบนี้จะเรียกว่า เกณฑ์ค่าความคลาดเคลื่อนมาตรฐานน้อยสุด (Minimum Standard Error Criterion)

**3. ประโยชน์ของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีประโยชน์ดังต่อไปนี้ (Wainer et al., 2000, pp. 11-16)

3.1 ทำให้การทดสอบมีความคล่องตัว สามารถทดสอบเป็นรายบุคคลตามความพร้อมของผู้สอบ มีการตรวจให้คะแนน และรายงานผลเป็นไปอย่างอัตโนมัติ

3.2 ระบบการทดสอบมีความเป็นมาตรฐาน การตรวจข้อสอบมีความชัดเจนและถูกต้อง ไม่มีปัญหาที่เกิดจากกระดาษคำตอบ ความยาวของแบบทดสอบมีความเหมาะสมกับความสามารถของผู้สอบ

3.3 ช่วยเพิ่มคุณภาพและประสิทธิภาพของการวัดความสามารถ/การเรียนรู้ของผู้สอบได้อย่างรวดเร็วและมีความถูกต้องยิ่งขึ้น

3.4 สามารถตรวจให้คะแนน และส่งผลย้อนกลับ (Feedback) แก่ผู้สอบได้ทันที

3.5 ช่วยให้การประเมินผลพัฒนาการหรือการติดตามการเปลี่ยนแปลงความสามารถ/การเรียนรู้เป็นไปอย่างต่อเนื่อง และปฏิบัติได้อย่างสะดวกรวดเร็ว

**ทฤษฎีทางการวัด**

ทฤษฎีทางการวัดเป็นการศึกษาคุณลักษณะภายในมนุษย์ซึ่งไม่สามารถสังเกตได้โดยตรง อันนำไปสู่ความเข้าใจพฤติกรรมภายนอกของมนุษย์ที่แสดงออกมา ทฤษฎีทางการวัดและการประเมินผลมีอยู่ด้วยกันหลายทฤษฎี โดยแต่ละทฤษฎีจะมีความเหมาะสมในการนำไปใช้แตกต่างกัน ซึ่งในที่นี้จะกล่าวเฉพาะทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) โดยมีรายละเอียดดังนี้

**1. ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)**

ทฤษฎีการตอบสนองข้อสอบ (เป็นทฤษฎีการทดสอบแนวใหม่ ซึ่งจำแนกได้ 2 ประเภท คือ 1) ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Dichotomous IRT) และ 2) ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous IRT) โดยในที่นี้จะกล่าวถึงเฉพาะทฤษฎีการตอบสนองข้อสอบแบบให้คะแนน 2 ค่าเท่านั้น

ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า เป็นโมเดลการตอบสนองข้อสอบที่ใช้กับการตรวจคะแนนรายข้อแบบ 2 ค่า เช่น ข้อสอบหรือข้อคำถามที่ตรวจให้คะแนนแบบ 0/1 (ตอบผิดได้ 0 ตอบถูกได้ 1) หรือ แบบใช่/ไม่ใช่ เป็นต้น โดยทั่วไปการนำโมเดลการตอบสนองข้อสอบไปใช้ในการศึกษา ต้องมีการตรวจสอบข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบ มี 4 ข้อ ได้แก่ (de Ayala, 2009, pp. 20-21; Hambleton, Swaminathan, & Rogers, 1991, pp. 9-12; Wainer et al., 2000, pp. 63-68) ดังนี้

1.1 ความเป็นเอกมิติ (Unidimensionality) หมายถึง ข้อคำถามหรือข้อสอบทุกข้อในเครื่องมือ (แบบทดสอบ) มุ่งวัดคุณลักษณะภายในของผู้สอบเพียงด้านใดด้านหนึ่ง การตรวจสอบความเป็นเอกมิติของแบบทดสอบ สามารถดำเนินการโดยใช้เทคนิคการวิเคราะห์ทางสถิติ ได้แก่ การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis) เพื่อคำนวณค่าไอแกน (Eigen Value) แล้วคำนวณอัตราส่วนระหว่างค่าไอเกนขององค์ประกอบแรกกับองค์ประกอบถัดไป ถ้าอัตราส่วนที่คำนวณได้มีค่าสูง แสดงว่า เครื่องมือหรือแบบทดสอบวัดคุณลักษณะเด่นเดียว (Single Dominant Factor) หรือใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) เพื่อตรวจสอบยืนยันว่า เครื่องมือหรือแบบทดสอบมุ่งวัดเพียงคุณลักษณะเดียว

1.2 ความเป็นอิสระ (Local Independence) มี 2 ลักษณะ ดังนี้

1.2.1 ความเป็นอิสระระหว่างข้อสอบ หมายถึง เมื่อมีการควบคุมความสามารถของผู้สอบให้คงที่แล้ว ผลการตอบข้อสอบแต่ละข้อของผู้สอบ จะต้องเป็นอิสระต่อกัน นั่นคือ ถ้าสุ่มผู้สอบ

ที่มีความสามารถ θ ขึ้นมา 1 คน ซึ่งได้ตอบข้อสอบ k ข้อ และให้ $U_j$ เป็นผลการตอบข้อสอบข้อที่ j เมื่อ j = 1, 2, ..., k เมื่อควบคุม θ ของผู้สอบ แล้วผลการตอบข้อสอบในแต่ละคู่ต้องไม่มีความสัมพันธ์กัน (อิสระต่อกัน) ซึ่งสอดคล้องกับสมการทางสถิติ ดังนี้

$$P(U_1, U_2, ..., U_k|\ \theta) = P(U_1|\theta)\ P(U_2|\theta)\ ...\ P(U_k|\theta) = \sum_{j=1}^{k} P(U_j|\theta) \quad (22)$$

ถ้าผลการตอบข้อสอบของผู้สอบคนหนึ่งเป็นอิสระต่อกัน แล้วค่าความน่าจะเป็นของแบบแผนในการตอบข้อสอบทุกข้อ จะเท่ากับผลคูณระหว่างค่าความน่าจะเป็นของผลการตอบข้อสอบแต่ละข้อ

1.2.2 ความเป็นอิสระระหว่างผู้สอบ หมายถึง เมื่อมีการควบคุมความสามารถของผู้สอบแต่ละคนให้คงที่แล้ว ผลการตอบข้อสอบระหว่างผู้สอบแต่ละคน จะต้องเป็นอิสระต่อกัน นั่นคือ ถ้าสุ่มข้อสอบขึ้นมา 1 ข้อ ในการตอบข้อสอบของผู้สอบ n คน และให้ $U_i$ เป็นผลการตอบข้อสอบของคนที่ i เมื่อ i = 1, 2, ..., n เมื่อควบคุม θ ของผู้สอบ แล้วผลการตอบข้อสอบของผู้สอบแต่ละคู่ต้องไม่มีความสัมพันธ์กัน (อิสระต่อกัน) ซึ่งสอดคล้องกับสมการทางสถิติ ดังนี้

$$P(U_1, U_2, ..., U_n|\ \theta) = P(U_1|\theta)\ P(U_2|\theta)\ ...\ P(U_n|\theta) = \sum_{i=1}^{n} P(U_i|\theta) \quad (23)$$

ถ้าผลการตอบข้อสอบข้อเดียวกันของผู้สอบแต่ละคนมีความเป็นอิสระต่อกัน แล้วความน่าจะเป็นของแบบแผนในการตอบข้อสอบข้อนั้นของผู้สอบทุกคน จะเท่ากับผลคูณระหว่างค่าความน่าจะเป็นของผลการตอบข้อสอบข้อนั้นของผู้สอบแต่ละคน

การตรวจสอบความเป็นอิสระระหว่างข้อสอบและผู้สอบ ดำเนินการโดยพิจารณาเมทริกซ์ความแปรปรวนและความแปรปรวนร่วม (Variance-Covariance Matrix) หรือพิจารณาเมทริกซ์สหสัมพันธ์ (Correlation Matrix) ของผลการตอบข้อสอบรายข้อ สำหรับผู้สอบที่มีช่วงความสามารถเท่ากัน โดยค่านอกแนวทแยงมุมควรมีค่าต่ำหรือเข้าใกล้ 0

1.3 โมเดลการตอบสนองข้อสอบ (Item Response Models) เป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็นของการตอบข้อสอบถูกกับความสามารถของผู้สอบ ซึ่งเป็นไปตามเส้นโค้งคุณลักษณะของข้อสอบ โดยจะกล่าวไว้ในหัวข้อถัดไป

1.4 การทดสอบที่ไม่แข่งขันด้านเวลา (Nonspeeded Test Administration) กล่าวคือ ทฤษฎี IRT ถือว่าความสามารถของผู้สอบเป็นปัจจัยที่สำคัญต่อผลการทดสอบ ความเร็วในการตอบจะต้องไม่มีอิทธิพลต่อผลการตอบ การจัดการทดสอบจึงต้องไม่อยู่ในสถานการณ์ที่มีการแข่งขันด้วยเวลา ผู้สอบที่มีความสามารถจะต้องมีเวลาเพียงพอในการทำข้อสอบ (Power Test Administration) การตรวจสอบความเหมาะสมในด้านเวลา สำหรับการดำเนินการสอบ สามารถทำได้โดยพิจารณาจากร้อยละของจำนวนผู้สอบที่ทำข้อสอบได้ครบทุกข้อต่อผู้สอบทั้งหมด ถ้าเวลาในการดำเนินการสอบมีความเหมาะสม ผู้สอบส่วนใหญ่ (ร้อยละ 80) สามารถตอบข้อสอบได้ครบหรือเกือบครบทุกข้อ หรือพิจารณาได้จากการเปรียบเทียบระหว่างความแปรปรวนของจำนวนข้อสอบที่

เว้น กับความแปรปรวนของจำนวนข้อสอบที่ตอบผิด ถ้าอัตราส่วนของความแปรปรวนมีค่าเข้าใกล้ 0 แสดงว่า เวลาในการดำเนินการสอบมีความเหมาะสม

**2. โมเดลการตอบสนองข้อสอบ (Item Response Model)**

โมเดลการตอบสนองข้อสอบ เป็นโมเดลทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างโอกาสตอบข้อสอบถูก ($P_i$) กับความสามารถที่มีอยู่ภายในของผู้สอบ ($\theta$) ซึ่งอยู่ในรูปของฟังก์ชันทางคณิตศาสตร์ 2 แบบ คือ ฟังก์ชันโลจิสติก (Logistic Function) และฟังก์ชันปกติสะสม (Normal Ogive Function) โดยโมเดลทั้งสองฟังก์ชันนี้ให้ผลลัพธ์ของการประมาณค่าใกล้เคียงกันมาก แต่ฟังก์ชันโลจิสติกจะได้รับความนิยมมากกว่า เนื่องจากมีลักษณะสูตรทางคณิตศาสตร์ที่คำนวณง่ายและสะดวกกว่าฟังก์ชันปกติสะสม โมเดลโลจิสติกมีหลายลักษณะขึ้นอยู่กับพารามิเตอร์ที่ปรากฏอยู่ในโมเดล ซึ่งมีตั้งแต่ 1 พารามิเตอร์ จนถึง 3 พารามิเตอร์ ซึ่งแต่ละโมเดลมีรายละเอียดดังนี้

**2.1 โมเดลโลจิสติกหนึ่งพารามิเตอร์ (One-Parameter Logistic Model: 1PL)**

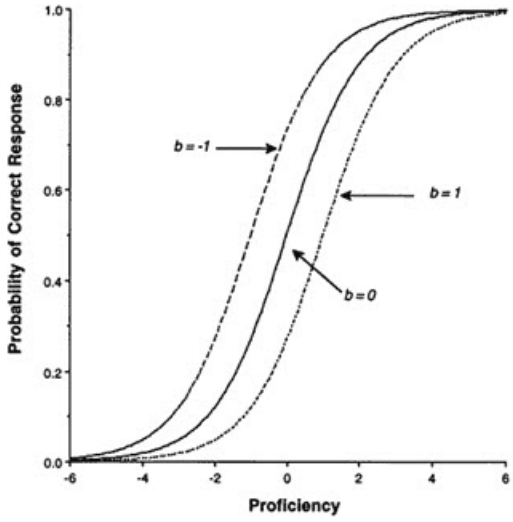
โมเดลโลจิสติกหนึ่งพารามิเตอร์ หรือราส์ชโมเดล (Rasch Model) เป็นโมเดลที่แสดงความน่าจะเป็นของการตอบข้อสอบถูก สำหรับผู้สอบที่มีค่าความสามารถ ($\theta$) โดยมีข้อตกลงเบื้องต้นว่า ข้อสอบแต่ละข้อมีค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (Discrimination Power Parameter: a) เท่ากับ 1 มีค่าพารามิเตอร์การเดาของข้อสอบ (Guess Parameter: c) เท่ากับ 0 แต่มีความแตกต่างเฉพาะค่าพารามิเตอร์ความยากของข้อสอบ (Difficulty Parameter: b) เท่านั้น โมเดลนี้จะเหมาะสำหรับข้อสอบแบบอิงเกณฑ์ที่มีสลับซับซ้อน สามารถเขียนฟังก์ชันโลจิสติกได้ดังนี้

$$P_i(\theta) = \frac{1}{1 + e^{-(\theta - b_i)}} \tag{24}$$

เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบซึ่งมีความสามารถ $\theta$ ตอบข้อสอบข้อที่ i ได้ถูกต้อง

$\theta$ แทน ระดับความสามารถของผู้สอบ

$b_i$ แทน ค่าความยากของข้อสอบข้อที่ i




ภาพที่ 2-2 โค้งคุณลักษณะของข้อสอบแบบ 1 พารามิเตอร์ (Wainer et al., 2000, p. 65)

จากภาพที่ 2-2 แสดงโค้งคุณลักษณะของข้อสอบ (Item Characteristic Curves: ICCs) จากโมเดลโลจิสติกหนึ่งพารามิเตอร์ ซึ่งเปรียบเทียบข้อสอบจำนวน 3 ข้อ ที่มีค่าความยากของข้อสอบ (b) แตกต่างกัน คือ -1, 0 และ 1 ตามลำดับ จากโค้งคุณลักษณะแสดงให้เห็นว่า 1) ค่าความน่าจะเป็นในการตอบข้อสอบถูกจะค่อย ๆ เพิ่มขึ้นตามค่าความสามารถของผู้สอบ 2) โค้งคุณลักษณะของข้อสอบทั้ง 3 เส้น มีความชันไม่แตกต่างกัน ถึงแม้มีค่าความยากของข้อสอบต่างกัน และ 3) จุดเปลี่ยนโค้งของเส้นโค้งคุณลักษณะของข้อสอบทั้ง 3 เส้น จะอยู่ตรงค่าความน่าจะเป็นในการตอบถูกเท่ากับ 0.5 เสมอ

**2.2 โมเดลโลจิสติกสองพารามิเตอร์ (Two-Parameter Logistic Model: 2PL)**

โมเดลโลจิสติกสองพารามิเตอร์ เป็นโมเดลที่แสดงความน่าจะเป็นของการตอบข้อสอบถูก สำหรับผู้สอบที่มีค่าความสามารถ $\theta$ โดยมีข้อตกลงเบื้องต้นว่า ข้อสอบแต่ละข้อมีพารามิเตอร์การเดาของข้อสอบเท่ากับ 0 และมีความแตกต่างกันของพารามิเตอร์ความยากของข้อสอบกับพารามิเตอร์อำนาจจำแนกของข้อสอบ สามารถเขียนฟังก์ชันโลจิสติกได้ดังนี้

$$P_i(\theta) = \frac{1}{1 + e^{-Da_i(\theta - b_i)}} \qquad (25)$$

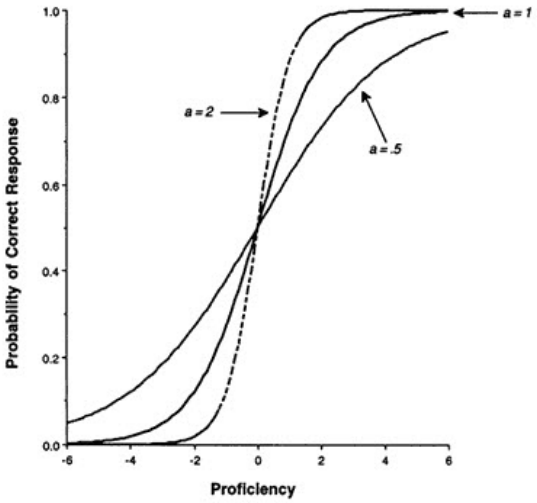
เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบซึ่งมีความสามารถ $\theta$ ตอบข้อสอบข้อที่ i ได้ถูกต้อง

$\theta$ แทน ระดับความสามารถของผู้สอบ

$b_i$ แทน ค่าความยากของข้อสอบข้อที่ i

$a_i$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ i

D แทน ค่าคงที่ เท่ากับ 1.7




ภาพที่ 2-3 โค้งคุณลักษณะของข้อสอบแบบ 2 พารามิเตอร์ (Wainer et al., 2000, p. 66)

จากภาพที่ 2-3 แสดงโค้งคุณลักษณะของข้อสอบ จากโมเดลโลจิสติกสองพารามิเตอร์ ซึ่งเปรียบเทียบข้อสอบจำนวน 3 ข้อ ที่มีค่าความยากของข้อสอบ (b) ไม่แตกต่างกัน แต่มีค่าอำนาจจำแนกของข้อสอบ (a) แตกต่างกัน คือ 0.5, 1 และ 2 ตามลำดับ จากโค้งคุณลักษณะแสดงให้เห็นว่าความชันของเส้นโค้งจะเปลี่ยนไปตามค่าอำนาจจำแนกของข้อสอบ กล่าวคือ ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง แล้วเส้นโค้งคุณลักษณะของข้อสอบจะมีความชันมาก และจุดเปลี่ยนโค้งของเส้นโค้งคุณลักษณะทั้ง 3 เส้น จะอยู่ตรงกับค่าความน่าจะเป็นในการตอบถูก เท่ากับ 0.5 เช่นเดียวกับโค้งคุณลักษณะข้อสอบ แบบ 1 พารามิเตอร์

**2.3 โมเดลโลจิสติกสามพารามิเตอร์ (Three-Parameter Logistic Model: 3PL)**

โมเดลโลจิสติกสามพารามิเตอร์ เป็นโมเดลที่แสดงความน่าจะเป็นของการตอบข้อสอบถูก สำหรับผู้สอบที่มีค่าความสามารถ $\theta$ โดยมีข้อตกลงเบื้องต้นว่า ข้อสอบแต่ละข้อ มีพารามิเตอร์ของข้อสอบที่แตกต่างกันทั้ง 3 พารามิเตอร์ ได้แก่ ค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบ และค่าการเดาของข้อสอบ โมเดลนี้จึงมีความเหมาะสำหรับข้อสอบที่มีรายการคำตอบแบบหลายตัวเลือก เนื่องจากผู้สอบสามารถเดาคำตอบได้ สามารถเขียนฟังก์ชันโลจิสติกได้ดังนี้

$$P_i(\theta) = c_i + \frac{(1-c_i)}{1+e^{-Da_i(\theta-b_i)}} \qquad (26)$$

เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบซึ่งมีความสามารถ $\theta$ ตอบข้อสอบข้อที่ i ได้ถูกต้อง

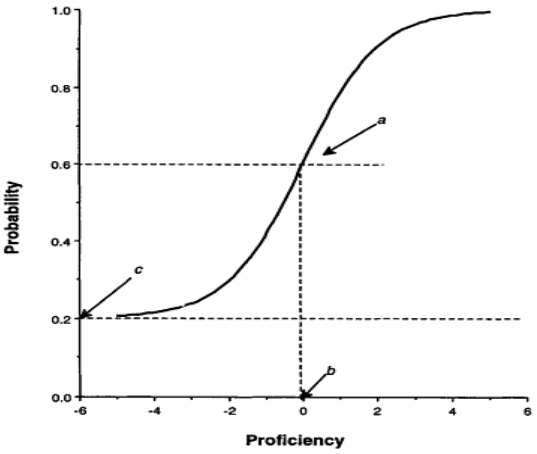
$\theta$ แทน ระดับความสามารถของผู้สอบ

$b_i$ แทน ค่าความยากของข้อสอบข้อที่ i

$a_i$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ i

$c_i$ แทน ค่าการเดาของข้อสอบข้อที่ i

D แทน ค่าคงที่ เท่ากับ 1.7




ภาพที่ 2-4 โค้งคุณลักษณะของข้อสอบแบบ 3 พารามิเตอร์ (Wainer et al., 2000, p. 67)

จากภาพที่ 2-4 แสดงโค้งคุณลักษณะของข้อสอบ จากโมเดลโลจิสติกสสามพารามิเตอร์ของข้อสอบ 1 ข้อ จากโค้งคุณลักษณะจะเห็นว่าเส้นโค้งไม่ได้เริ่มจากศูนย์ แต่เริ่มจากค่าการเดาของข้อสอบ นั่นแสดงให้เห็นว่าผู้สอบมีโอกาสตอบข้อสอบถูก ทั้งที่ไม่มีความสามารถในเรื่องที่ต้องการวัด การตอบข้อสอบถูกนั้นเกิดจากการเดาของผู้สอบ และจะสังเกตเห็นว่าจุดเปลี่ยนโค้งของเส้นโค้งคุณลักษณะจะไม่อยู่ที่ค่าความน่าจะเป็นในการตอบถูก เท่ากับ 0.5 เหมือนกับสองโมเดลแรก เนื่องจากโค้งคุณลักษณะของข้อสอบไม่ได้เริ่มจากศูนย์ ส่งผลให้โอกาสในการตอบข้อสอบถูก มีค่าสูงขึ้น

**3. คุณสมบัติของความไม่แปรเปลี่ยนของค่าพารามิเตอร์ (Invariance)**

เมื่อโมเดลการตอบสนองข้อสอบมีความสอดคล้องกับข้อมูล (Model-Data Fit) จะทำให้เกิดคุณสมบัติที่สำคัญของทฤษฎีการตอบสนองข้อสอบ คือ คุณสมบัติความไม่แปรเปลี่ยน ซึ่งมี 2 ลักษณะ ดังนี้

3.1 ความไม่แปรเปลี่ยนของพารามิเตอร์ของข้อสอบ (Item Parameter) นั้นคือ ค่าพารามิเตอร์ของข้อสอบจะไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบ (โค้งคุณลักษณะข้อสอบ ICC จะมีลักษณะเดียวกัน ทุกกลุ่มความสามารถของผู้สอบ)

3.2 ความไม่แปรเปลี่ยนของพารามิเตอร์ความสามารถของผู้สอบ (Ability Parameter) นั้นคือ ค่าพารามิเตอร์ความสามารถของผู้สอบจะไม่เปลี่ยนแปลงไปตามชุดของข้อสอบ

**4. ฟังก์ชันสารสนเทศของข้อสอบและแบบทดสอบ (Item and Test Information)**

สารสนเทศข้อสอบเป็นเครื่องบ่งชี้คุณภาพข้อสอบ ซึ่งจะให้ค่าสูงสุดเมื่อข้อสอบมีค่าความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้สอบ มีค่าอำนาจจำแนกสูง และมีค่าการเดาของข้อสอบน้อยมากจนเข้าใกล้ศูนย์ ค่าสารสนเทศของข้อสอบแสดงเป็นฟังก์ชันทางคณิตศาสตร์ ได้ดังนี้ (Wainer et al., 2000, p. 73)

$$I_i(\theta) = \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)} \quad (27)$$

เมื่อ $I_i(\theta)$ คือ สารสนเทศที่ได้รับจากข้อสอบข้อที่ i สำหรับผู้สอบที่มีความสามารถ $\theta$

$P_i'(\theta)$ คือ อนุพันธ์ของ $P_i(\theta)$หรือความชันของฟังก์ชันการตอบสนองข้อสอบข้อที่ i ณ ตำแหน่งความสามารถ $\theta$

$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบซึ่งมีความสามารถ $\theta$ ตอบข้อสอบข้อที่ i ได้ถูกต้อง

$Q(\theta)$ คือ $1-P_i(\theta)$

เพื่อความสะดวกในการคำนวณค่าสารสนเทศของข้อสอบ ในกรณีที่เป็นโมเดลโลจิสติกแบบสามพารามิเตอร์ สมการที่ 27 สามารถเขียนใหม่ได้ดังนี้

$$I_i(\theta) = \frac{D^2 a_i^2 [P_i(\theta) - c_i]^2 Q_i(\theta)}{(1-c_i)^2 P_i(\theta)} \quad (28)$$

เมื่อ $I_i(\theta)$ คือ สารสนเทศที่ได้รับจากข้อสอบข้อที่ i สำหรับผู้สอบที่มีความสามารถ $\theta$

$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบซึ่งมีความสามารถ $\theta$ ตอบข้อสอบข้อที่ i ได้ถูกต้อง

$Q(\theta)$ คือ $1-P_i(\theta)$

$a_i$ คือ ค่าอำนาจจำแนกของข้อสอบ ข้อที่ i

$c_i$ คือ ค่าการเดาของข้อสอบ ข้อที่ i

D คือ ค่าคงที่ เท่ากับ 1.7

ค่าสารสนเทศข้อสอบแต่ละข้อ เมื่อนำมารวมเข้าด้วยกัน จะได้ค่าสารสนเทศของแบบทดสอบ ณ ตำแหน่งความสามารถเดียวกัน ให้ความแม่นยำในการประมาณค่าความสามารถของผู้สอบว่า เมื่อทำแบบทดสอบทั้งฉบับ ผลการทดสอบบอกถึงผู้สอบมีความสามารถอยู่ ณ ตำแหน่งใด แสดงเป็นฟังก์ชันทางคณิตศาสตร์ได้ดังนี้

$$I(\theta) = \sum_{i=1}^{n} I_i(\theta) \quad (29)$$

เมื่อ $I(\theta)$ คือ สารสนเทศแบบทดสอบสำหรับผู้สอบที่มีความสามารถ $\theta$

$I_i(\theta)$ คือ สารสนเทศที่ได้รับจากข้อสอบข้อที่ i สำหรับผู้สอบที่มีความสามารถ $\theta$

n คือ จำนวนข้อสอบในแบบทดสอบ

**5. ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ (Standard Error of the Ability Estimate)**

ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ เป็นค่าส่วนเบี่ยงเบนมาตรฐานของการแจกแจงความน่าจะเป็นของค่าประมาณความสามารถ $\theta$ ซึ่งเป็นสัดส่วนผกผันกับความถูกต้องแม่นยำของการประมาณค่าความสามารถ หรือค่าสารสนเทศแบบทดสอบ (de Ayala, 2009, pp. 27-31) แสดงดังฟังก์ชันต่อไปนี้

$$\mathrm{SEE}(\hat{\theta}) = \frac{1}{\sqrt{\mathrm{I}(\theta)}} \tag{30}$$

เมื่อ $\mathrm{SEE}(\hat{\theta})$ คือ ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ

$\mathrm{I}(\theta)$ คือ สารสนเทศแบบทดสอบสำหรับผู้สอบที่มีความสามารถ $\theta$

จากฟังก์ชันข้างต้น จะเห็นได้ว่า ถ้าค่าสารสนเทศแบบทดสอบมีค่าสูง จะทำให้ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถนั้นจะมีค่าต่ำ (ที่ระดับความสามารถ $\theta$ เดียวกัน)

**6. การประเมินประสิทธิภาพการประมาณค่าความสามารถของผู้สอบ**

ประสิทธิภาพการประมาณค่าความสามารถของผู้สอบ หมายถึง ความสามารถของวิธีกำหนดในการประมาณค่าความสามารถของผู้สอบให้ใกล้เคียงหรือเท่ากับค่าความสามารถที่แท้จริงของผู้สอบมากที่สุด ซึ่งการประมาณค่าความสามารถว่าใกล้เคียงหรือเท่ากับค่าความสามารถที่แท้จริงของผู้สอบหรือไม่ สามารถพิจารณาได้จากค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย หรือค่าความลำเอียงเฉลี่ย ซึ่งมีนักวิชาการหลายท่านที่นำค่าทั้ง 3 นี้ มาใช้เปรียบเทียบประสิทธิภาพการประมาณค่าความสามารถของผู้สอบ ของวิธีการต่าง ๆ ในการดำเนินการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ไม่ว่าจะเป็นวิธีการคัดเลือกข้อสอบข้อถัดไป วิธีการควบคุมการใช้ข้อสอบ หรือวิธีการควบคุมสัดส่วนเนื้อหาของข้อสอบ นักวิชาการเหล่านั้น ได้แก่ Belov, Armstrong, and Weissman (2008); Cheng and Chang (2009); Cheng, Chang, Douglas, and Guo (2009); Hau and Chang (2001) และ Leung et al. (2002, 2003)

**6.1 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย** (Mean Square Error: MSE) หมายถึง ค่าบอกความถูกต้องของค่าประมาณความสามารถของผู้สอบ แสดงเป็นผลต่างกำลังสองเฉลี่ยระหว่างค่าประมาณความสามารถและค่าความสามารถที่แท้จริงของผู้สอบ ซึ่งค่า MSE ขนาดเล็กบอกถึงการประมาณค่าความสามารถมีความแตกต่างจากค่าความสามารถที่แท้จริงของผู้สอบน้อยมาก สะท้อนให้เห็นถึงประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป (Yan, von Davier, & Lewis, 2014, p. 115)

$$\mathrm{MSE} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left( \hat{\theta}_{\mathrm{i}} - \theta_{\mathrm{i}} \right)^2 \tag{31}$$

เมื่อ n คือ จำนวนผู้สอบทั้งหมด
$\hat{\theta}_i$ คือ ค่าประมาณความสามารถของผู้สอบคนที่ i
$\theta_i$ คือ ค่าความสามารถที่แท้จริงของผู้สอบคนที่ i

**6.2 ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)** หมายถึง ค่าบอกความถูกต้องของค่าประมาณความสามารถของผู้สอบ แสดงเป็นค่ารากที่สองของผลต่างกำลังสองเฉลี่ยระหว่างค่าประมาณความสามารถและค่าความสามารถที่แท้จริงของผู้สอบ นั่นคือ ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ดังนั้นค่า RMSE ขนาดเล็กบอกถึงการประมาณค่าความสามารถมีความแตกต่างจากค่าความสามารถที่แท้จริงของผู้สอบน้อยมาก (Yan et al., 2014, p. 115)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_i - \theta_i\right)^2} \qquad (32)$$

เมื่อ n คือ จำนวนผู้สอบทั้งหมด
$\hat{\theta}_i$ คือ ค่าประมาณความสามารถของผู้สอบคนที่ i
$\theta_i$ คือ ค่าความสามารถที่แท้จริงของผู้สอบคนที่ i

**6.3 ค่าความลำเอียงเฉลี่ย** (Average Bias) หมายถึง ค่าบอกความเที่ยงตรงของค่าประมาณความสามารถของผู้สอบ แสดงเป็นผลต่างเฉลี่ยระหว่างค่าประมาณความสามารถและค่าความสามารถที่แท้จริงของผู้สอบ สามารถบอกทิศทางการประมาณค่าว่าให้ผลสูงหรือต่ำกว่าค่าความสามารถที่แท้จริงของผู้สอบ ถ้าค่าความลำเอียงเฉลี่ยเข้าใกล้ศูนย์ จะสะท้อนให้เห็นถึงประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป (Yan et al., 2014, p. 115)

$$\text{Average Bias} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_i - \theta_i\right) \qquad (33)$$

เมื่อ n คือ จำนวนผู้สอบทั้งหมด
$\hat{\theta}_i$ คือ ค่าประมาณความสามารถของผู้สอบคนที่ i
$\theta_i$ คือ ค่าความสามารถที่แท้จริงของผู้สอบคนที่ i

**7. การประเมินประสิทธิภาพของการใช้ข้อสอบในคลังข้อสอบ**

ประสิทธิภาพของการใช้ข้อสอบในคลังข้อสอบ หมายถึง ความสามารถของวิธีกำหนดที่ทำให้ข้อสอบทุกข้อในคลังข้อสอบมีโอกาสนำไปใช้ในการทดสอบ โดยพิจารณาจากอัตราการใช้ข้อสอบซึ่งเป็นอัตราส่วนระหว่างจำนวนการใช้ข้อสอบต่อจำนวนผู้สอบทั้งหมด หากข้อสอบข้อใดมีอัตราการใช้ข้อสอบมากกว่า 0.2 ข้อสอบข้อนั้น จะเป็นข้อสอบที่มีอัตราการใช้ข้อสอบสูง (Over-Exposure Item) อัตราการใช้ข้อสอบคำนวณได้ดังนี้ (Ozturk & Dogan, 2015)

$$\text{อัตราการใช้ข้อสอบ} = \frac{\text{จำนวนการใช้ข้อสอบ}}{\text{จำนวนผู้สอบทั้งหมด}} \quad (34)$$

คลังข้อสอบใดประกอบด้วยข้อสอบที่มีอัตราการใช้ข้อสอบสูงจำนวนมาก แสดงว่าในการทดสอบจะมีข้อสอบบางข้อถูกจัดให้ผู้สอบใช้ในการทดสอบบ่อยครั้งจนเกินไป ซึ่งจะส่งผลต่อความปลอดภัยในการทดสอบ เพราะผู้สอบในรุ่นถัดไปสามารถคาดเดาล่วงหน้าได้ว่าตนเองจะได้ข้อสอบข้อใด ดังนั้น จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จึงสะท้อนให้เห็นถึงประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปในด้านการควบคุมการใช้ข้อสอบ (van der Linden & Glas, 2002, pp. 45-47) ด้วยเหตุนี้ การคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ นอกจากคำนึงถึงการคัดเลือกข้อสอบที่มีค่าความยากของข้อสอบเหมาะสมกับค่าความสามารถของผู้สอบแล้ว ยังต้องคำนึงถึงการควบคุมการใช้ข้อสอบ (Item Exposure Control) ไม่ให้มีหรือให้มีข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำนวนน้อยที่สุดเท่าที่ทำได้ วิธีการควบคุมการใช้ข้อสอบมีหลายวิธีการ เช่น วิธี 5-4-3-2-1, วิธีการซิมสันเฮกเตอร์ หรือวิธีการ a-Stratified Strategy (Chang & Ying, 1999; McBride & Martin, 1983; Sympson & Hetter, 1985)

## ตอนที่ 2 ทฤษฎีการตัดสินใจและงานวิจัยที่เกี่ยวข้อง

การตัดสินใจเป็นสิ่งสำคัญอย่างหนึ่งในการดำเนินงานต่าง ๆ โดยเฉพาะทางธุรกิจ ให้เป็นไปตามวัตถุประสงค์ที่ตั้งไว้ ทฤษฎีการตัดสินใจมีแนวคิดในการคัดเลือกผลลัพธ์ที่ให้ผลตอบแทนที่ดีที่สุดในสภาวการณ์ใดสภาวการณ์หนึ่ง จะต้องมีองค์ประกอบพื้นฐานในการตัดสินใจดังนี้ (Taylor, 2009, pp. 538-539)

**องค์ประกอบของการตัดสินใจ**

1. ทางเลือก (Alternative Acts) การตัดสินใจแก้ไขปัญหาส่วนใหญ่ จะพบว่ามีทางเลือกในการแก้ไขปัญหาอย่างน้อยที่สุด 2 ทางเลือก ซึ่งทางเลือกจะหมายถึง การกำหนดทางเลือกของกระทำทั้งหมดเท่าที่ผู้ตัดสินใจจะสามารถพิจารณาได้ เช่น บริษัทหนึ่ง ต้องการขนส่งสินค้าจากจังหวัดเชียงใหม่มายังกรุงเทพ โดยมีวัตถุประสงค์เพื่อให้ต้นทุนในการขนส่งต่ำสุด โดยกำหนดทางเลือกในการขนส่งที่เป็นไปได้จำนวน 3 ทางเลือก ได้แก่ การขนส่งทางบก โดยใช้รถไฟ การขนส่งทางอากาศโดยใช้เครื่องบิน หรือการขนส่งทางน้ำโดยใช้เรือ เป็นต้น

2. สภาวการณ์ (State of Nature) หรือเหตุการณ์ (Event) เป็นสถานการณ์ที่เกิดขึ้นอยู่ในปัจจุบันหรืออนาคตแบ่งได้ 2 ประเภท ได้แก่ สภาวการณ์ที่แน่นอน คือ สภาวการณ์ที่ทราบว่าจะเกิดขึ้นแน่นอนในอนาคต และสภาวการณ์ที่ไม่แน่นอน คือ สภาวการณ์ที่ไม่ทราบว่าจะมีอะไรเกิดขึ้นบางในอนาคต แต่บางครั้งอาจจะทราบโอกาสในการเกิดของสภาวการณ์นั้น ๆ

3. ผลตอบแทน (Payoff) เป็นผลลัพธ์ที่เกิดจากการตัดสินใจปฏิบัติทางเลือกใดทางเลือกหนึ่ง ซึ่งเมื่อเกิดผลลัพธ์จากการตัดสินใจเลือกทางเลือกใดทางเลือกหนึ่งแล้ว ผู้ตัดสินใจจะนำผลลัพธ์ดังกล่าวนั้นไปประเมินค่าเป็นตัวเลข ค่าที่เป็นตัวเลขนั้น เรียกกว่า ผลตอบแทน

ผลตอบแทนในรูปของค่าเสียโอกาส คือ ผลตอบแทนที่ผู้ตัดสินใจควรจะได้รับแต่ไม่ได้รับ เนื่องจากผู้ตัดสินใจไม่ได้เลือกทางเลือกที่ดีที่สุด ค่าเสียโอกาสเป็นผลต่างระหว่างผลตอบแทนจากทางเลือกที่ดีที่สุด กับผลตอบแทนจากทางเลือกอื่นๆ ในแต่ละเหตุการณ์

**ประเภทของการตัดสินใจ**

การตัดสินใจเลือกทางเลือกที่ดีที่สุด มีหลายวิธีจำแนกตามสภาวการณ์ที่แวดล้อมการตัดสินใจอยู่ในขณะนั้น โดยแบ่งได้เป็น 4 ประเภท (กิตติ ภักดีวัฒนะกุล และพนิดา พานิชกุล, 2554, หน้า 541-554; Taylor, 2009, pp. 539-572) มีรายละเอียดดังนี้

**1. การตัดสินใจในสภาวการณ์ที่แน่นอน** (Certainly Decision Making) หมายถึง การตัดสินใจที่ผู้ตัดสินใจมีข้อมูลเกี่ยวกับเรื่องที่ต้องการตัดสินใจครบสมบูรณ์ และทราบถึงผลลัพธ์จากเหตุการณ์ที่คาดว่าจะเกิดขึ้นได้เป็นอย่างดี

**2. การตัดสินใจในสภาวการณ์ที่ไม่แน่นอน** (Uncertainly Decision Making) หมายถึง การตัดสินใจที่ผู้ตัดสินใจมีข้อมูลที่เกี่ยวข้องกับเรื่องที่ต้องการตัดสินใจไม่เพียงพอ ทำให้ไม่สามารถทราบถึงผลลัพธ์ของเหตุการณ์ที่คาดว่าจะเกิดขึ้นได้อย่างแน่นอน นอกจากนี้ผู้ตัดสินใจยังไม่สามารถคาดคะเน หรือกำหนดค่าความน่าจะเป็นในแต่ละเหตุการณ์ได้ ดังนั้น หลักเกณฑ์ที่นำมาใช้ในการตัดสินใจโดยส่วนใหญ่จึงเป็นดุลยพินิจ และประสบการณ์ที่ผ่านมาในอดีตเพื่อช่วยในการตัดสินใจ ดังนี้

2.1 เกณฑ์แมกซิแมกซ์ (Maximax Criteria) หมายถึง การตัดสินใจเลือกผลตอบแทนที่สูงที่สุดจากผลตอบแทนที่สูงที่สุดในแต่ละทางเลือก กล่าวคือ ผู้ตัดสินใจจะเลือกผลตอบแทนที่ดีที่สุดของแต่ละทางเลือกออกมาก่อน จากนั้นจึงเลือกทางเลือกที่ให้ผลตอบแทนมากที่สุด

2.2 เกณฑ์แมกซิมิน (Maximin Criteria) หมายถึง การตัดสินใจเลือกผลตอบแทนที่สูงที่สุดจากผลตอบแทนที่ต่ำสุดในแต่ละทางเลือก กล่าวคือ ผู้ตัดสินใจจะเลือกผลตอบแทนที่มีค่าต่ำที่สุดในแต่ละทางเลือกมาก่อน จากนั้นจึงเลือกทางเลือกที่ให้ผลตอบแทนมากที่สุด วิธีนี้จะนำมาใช้กับปัญหาการตัดสินใจที่เกี่ยวข้องกับรายได้

2.3 เกณฑ์ของลาปลาซ (Laplace Criteria) หมายถึง การตัดสินใจเลือกทางเลือกที่ดีที่สุด โดยพิจารณาจากค่าคาดหวังของผลตอบแทนสูงสุดหากเป็นการตัดสินใจเกี่ยวกับรายได้ (หรือพิจารณาจากค่าคาดหวังของผลตอบแทนต่ำสุดหากเป็นการตัดสินใจเกี่ยวกับต้นทุน) วิธีนี้จะต้องมีการกำหนดความน่าจะเป็นของแต่ละเหตุการณ์ โดยสมมติให้ทุกเหตุการณ์ “มีค่าความน่าจะเป็นเท่ากัน”

2.4 เกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criteria) หมายถึง การตัดสินใจเลือกทางเลือกที่ดีที่สุด โดยพิจารณาจากผลตอบแทนเฉลี่ยถ่วงน้ำหนักสูงสุด โดยผลตอบแทนเฉลี่ยถ่วงน้ำหนัก ในแต่ละทางเลือกของการตัดสินใจ สามารถคำนวณได้ดังนี้

ผลตอบแทนเฉลี่ยถ่วงน้ำหนัก = $\alpha$ × (ผลตอบแทนสูงสุด) + (1- $\alpha$)×(ผลตอบแทนต่ำสุด)

โดยที่ $\alpha$ คือ สัมประสิทธิ์ของการมองโลกในแง่ดี (Coefficient of Optimism)

ค่าสัมประสิทธิ์ของการมองโลกในแง่ดี หรือค่าถ่วงน้ำหนักของผลตอบแทนสูงสุด เป็นค่าที่ผู้ตัดสินใจใช้ดุลยพินิจกำหนดขึ้นเอง โดยคาดคะเนจากข้อมูลหรือประสบการณ์ในอดีต ซึ่งมีค่าตั้งแต่ 0 ถึง 1 ถ้าค่า $\alpha$ มีค่าเข้าใกล้ 1 หมายความว่า ผู้ตัดสินใจค่อนข้างจะให้น้ำหนักกับผลตอบแทนสูงมากกว่าผลตอบแทนต่ำ ในทางตรงข้าม ถ้าค่า $\alpha$ มีค่าเข้าใกล้ 0 หมายความว่า ผู้ตัดสินใจค่อนข้างจะให้น้ำหนักกับผลตอบแทนต่ำมากกว่าผลตอบแทนสูง

ขั้นตอนในการตัดสินใจ โดยใช้เกณฑ์ของเฮอร์วิคซ์ มี 3 ขั้น ดังนี้

ขั้นที่ 1 กำหนดค่าสัมประสิทธิ์ $\alpha$

ขั้นที่ 2 คำนวณผลตอบแทนเฉลี่ยถ่วงน้ำหนักของแต่ละทางเลือก

ขั้นที่ 3 ตัดสินใจเลือกทางเลือกที่ให้ผลตอบแทนเฉลี่ยถ่วงน้ำหนักสูงสุด

ตัวอย่างการตัดสินใจในการจัดการดำเนินงานทางธุรกิจ โดยใช้เกณฑ์ของเฮอร์วิคซ์

สมมติ นักลงทุนรายหนึ่ง มีเงินลงทุนอยู่ 1 ล้านบาท ได้คัดเลือกหุ้นที่น่าสนใจเพื่อการลงทุนไว้ 3 รายการ คือ หุ้นของบริษัท A, B และ C ตามลำดับ โดยผลตอบแทนจากการลงทุนขึ้นอยู่กับสภาวการณ์ทางเศรษฐกิจ ซึ่งมี 3 ระดับ ได้แก่ สภาวการณ์ที่เศรษฐกิจดี สภาวการณ์ที่เศรษฐกิจปานกลาง และสภาวการณ์ที่เศรษฐกิจแย่ ซึ่งสามารถคำนวณผลตอบแทนจากการลงทุนของแต่ละบริษัท ภายใต้สภาวการณ์ทางเศรษฐกิจต่างกัน ได้ดังนี้

ตารางที่ 2-1 ผลตอบแทนจากการลงทุนในบริษัท A, B และ C

| สภาวการณ์ทางเศรษฐกิจ | ผลตอบแทน (ล้านบาท) | | |
|---|---|---|---|
| | บริษัท A | บริษัท B | บริษัท C |
| เศรษฐกิจดี | 2.2 | 3.0 | 2.6 |
| เศรษฐกิจปานกลาง | 1.7 | 2.0 | 2.1 |
| เศรษฐกิจแย่ | 1.3 | 0.8 | 1.1 |

จากขั้นตอนในการตัดสินใจ โดยใช้เกณฑ์ของเฮอร์วิคซ์ จะได้ว่า

ขั้นที่ 1 กำหนดค่าสัมประสิทธิ์ $\alpha$

สมมติกำหนดค่า $\alpha$ เท่ากับ 0.8 แสดงว่า นักลงทุนรายนี้มองโลกค่อนข้างดีคาดคะเนว่า ไม่ว่าจะเลือกลงทุนในบริษัทใดก็ตาม จะให้ผลตอบแทนที่ดีมากกว่าไม่ดี

ขั้นที่ 2 คำนวณผลตอบแทนเฉลี่ยถ่วงน้ำหนักของแต่ละทางเลือก

ผลตอบแทนเฉลี่ยถ่วงน้ำหนักของบริษัท A

คือ (0.8×2.2) + (0.2×1.3) = 2.02

ผลตอบแทนเฉลี่ยถ่วงน้ำหนักของบริษัท B

คือ (0.8×3.0) + (0.2×0.8) = 2.56

ผลตอบแทนเฉลี่ยถ่วงน้ำหนักของบริษัท C

คือ (0.8×2.6) + (0.2×1.1) = 2.30

ขั้นที่ 3 ตัดสินใจเลือกทางเลือกที่ให้ผลตอบแทนเฉลี่ยถ่วงน้ำหนักสูงสุด

จากการคำนวณผลตอบแทนเฉลี่ยถ่วงน้ำหนักของแต่ละบริษัท ปรากฏว่า บริษัท B ให้ผลตอบแทนเฉลี่ยถ่วงน้ำหนักสูงสุด ดังนั้น นักลงทุนรายนี้ ควรตัดสินใจเลือกลงทุนกับบริษัท B

จากตัวอย่างข้างต้น เมื่อนำเกณฑ์ของเฮอร์วิคซ์มาประยุกต์ในการคัดเลือกข้อสอบข้อถัดไป สามารถดำเนินการได้เช่นเดียวกัน โดยทางเลือกแต่ละทางเลือก คือ ข้อสอบทุกข้อในคลังข้อสอบที่ยังไม่ถูกเลือกใช้ในการทดสอบ สภาวการณ์ที่จะเกิดขึ้นมีเพียง 2 สภาวการณ์ คือ ผู้สอบตอบข้อสอบถูก และผู้สอบตอบข้อสอบผิด และเกณฑ์การตัดสินใจที่เลือกทางเลือกที่ให้ผลตอบแทนเฉลี่ยถ่วงน้ำหนักสูงสุด เปลี่ยนเป็นตัดสินใจเลือกข้อสอบที่ให้ค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบต่ำสุด สามารถแสดงตัวอย่างการนำเกณฑ์ของเฮอร์วิคซ์มาใช้ในการคัดเลือกข้อสอบข้อถัดไป ได้ดังนี้

สมมติ คลังข้อสอบมีข้อสอบเหลือ 5 ข้อ ได้แก่ ข้อสอบข้อที่ 1 ถึงข้อที่ 5 ซึ่งจะนำมาใช้ในการคัดเลือกข้อสอบข้อถัดไป และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ (SEE) ทั้งกรณีที่ผู้สอบตอบถูก และกรณีที่ผู้สอบตอบผิด ของข้อสอบแต่ละข้อ (จากสมการที่ 28 – 30 และจากสมการที่ 12 - 17) แสดงดังตารางที่ 2-2

ตารางที่ 2-2 ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ กรณีที่ผู้สอบตอบถูก และกรณีที่ผู้สอบตอบผิด เมื่อได้รับข้อสอบข้อที่ 1 ถึงข้อที่ 5

| สภาวการณ์ | ค่า SEE | | | | |
|---|---|---|---|---|---|
| | ข้อที่ 1 | ข้อที่ 2 | ข้อที่ 3 | ข้อที่ 4 | ข้อที่ 5 |
| ผู้สอบตอบถูก | 2.75 | 2.90 | 2.30 | 2.32 | 9.63 |
| ผู้สอบตอบผิด | 3.24 | 2.65 | 2.46 | 1.06 | 1.05 |

จากขั้นตอนในการตัดสินใจ โดยใช้เกณฑ์ของเฮอร์วิคซ์ จะได้ว่า

ขั้นที่ 1 กำหนดค่าสัมประสิทธิ์ $\alpha$

สมมติกำหนดค่า $\alpha$ เท่ากับ 0.75

ขั้นที่ 2 คำนวณผลตอบแทนเฉลี่ยถ่วงน้ำหนักของข้อสอบแต่ละข้อ

ค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE สำหรับข้อสอบข้อที่ 1

คือ (0.75×3.24) + (0.25×2.75) = 3.12

ค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE สำหรับข้อสอบข้อที่ 2

คือ (0.75×2.90) + (0.25×2.65) = 2.84

ค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE สำหรับข้อสอบข้อที่ 3

คือ (0.75×2.46) + (0.25×2.30) = 2.42

ค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE สำหรับข้อสอบข้อที่ 4

คือ (0.75×2.32) + (0.25×1.06) = 2.01

ค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE สำหรับข้อสอบข้อที่ 5

คือ (0.75×9.63) + (0.25×1.05) = 7.49

ขั้นที่ 3 ตัดสินใจเลือกข้อสอบที่ให้ค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE ต่ำสุด

จากการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของข้อสอบแต่ละข้อ ปรากฏว่า ข้อสอบข้อที่ 4 ให้ค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE ต่ำสุด ดังนั้น ข้อสอบข้อถัดไป คือ ข้อสอบข้อที่ 4

2.5 เกณฑ์ Minimax Regret Criteria หมายถึง การตัดสินใจเลือกค่าเสียโอกาสที่ต่ำที่สุดจากค่าเสียโอกาสที่สูงที่สุดในแต่ละทางเลือก กล่าวคือ ผู้ตัดสินใจเลือกค่าเสียโอกาสที่สูงที่สุดในแต่ละทางเลือกออกมา และเลือกค่าเสียโอกาสที่มีค่าต่ำที่สุด

**3. การตัดสินใจในสภาวการณ์ความเสี่ยง** (Risk Decision Making) หมายถึง การตัดสินใจที่ผู้ตัดสินใจมีข้อมูลที่เกี่ยวข้องกับเรื่องที่ต้องการตัดสินใจไม่เพียงพอ ทำให้ไม่สามารถทราบถึงผลลัพธ์จากเหตุการณ์ที่คาดว่าจะเกิดขึ้นได้อย่างแน่นอน แต่สามารถคาดคะเนหรือกำหนดค่าความน่าจะเป็นในแต่ละเหตุการณ์ได้ โดยค่าความน่าจะเป็นที่กำหนดขึ้นนี้ อาจได้มาจากการศึกษาข้อมูลในอดีตจากเหตุการณ์ที่มีความคล้ายคลึงกัน หรือจากดุลยพินิจของผู้ตัดสินใจมาเป็นพื้นฐานในการคาดคะเน การตัดสินใจในสภาวการณ์ความเสี่ยงนี้ แบ่งออกได้ 2 กรณี ดังนี้

3.1 ค่าคาดหวังของผลตอบแทนสูงสุด (Maximum Expected Payoff Criterion) การตัดสินใจที่ผู้ตัดสินใจทราบความน่าจะเป็นของเหตุการณ์ที่จะเกิดขึ้น แล้วจะตัดสินใจเลือกทางเลือกที่มีค่าคาดหวังของผลตอบแทนสูงสุด โดยค่าคาดหวังของผลตอบแทน สามารถคำนวณได้จากการหาผลรวมของผลคูณระหว่างค่าตอบแทนของแต่ละเหตุการณ์ในทางเลือกนั้นกับความน่าจะเป็นของเหตุการณ์นั้น

3.2 ค่าคาดหวังของค่าเสียโอกาสต่ำที่สุด (Minimum Expected Opportunity Loss Criterion) การตัดสินใจที่ผู้ตัดสินใจทราบความน่าจะเป็นของเหตุการณ์ที่จะเกิดขึ้น และผู้ตัดสินใจจะตัดสินใจเลือกทางเลือกที่มีค่าคาดหวังของค่าเสียโอกาสต่ำสุด โดยค่าคาดหวังของค่าเสียโอกาสสามารถคำนวณได้จากการหาผลรวมของผลคูณระหว่างค่าเสียโอกาสของแต่ละเหตุการณ์ในทางเลือกนั้นกับความน่าจะเป็นของเหตุการณ์นั้น

**4. การตัดสินใจในสภาวการณ์การแข่งขัน** (Competition Decision Making) หรือเรียกอีกชื่อหนึ่งว่า ทฤษฎีเกม (Game Theory) หมายถึง การตัดสินใจในสภาวการณ์ที่ไม่แน่นอน โดยประกอบด้วยคู่แข่งขันตั้งแต่ 2 ฝ่ายขึ้นไป แต่ละฝ่ายจะนำกลยุทธ์การแข่งขันมาใช้เพื่อให้ได้มาซึ่งชัยชนะเหนือคู่แข่งขัน เช่น การแข่งขันในการแย่งส่วนแบ่งทางการตลาด โดยทฤษฎีเกมสามารถแบ่งได้ 2 รูปแบบ ดังนี้

4.1 เกมกลยุทธ์เดียว (Pure Strategy Game) หมายถึง เกมการแข่งขันที่แต่ละฝ่ายต่างเลือกใช้กลยุทธ์เดียวตลอดการแข่งขัน โดยไม่สนใจว่าอีกฝ่ายจะใช้กลยุทธ์อะไร ทฤษฎีเกมรูปแบบนี้จะใช้เกณฑ์แมกซิแมกซ์ และมินนิแมกซ์

4.2 เกมกลยุทธ์ผสม (Mixed Strategy Game) หมายถึง เกมการแข่งขันที่แต่ละฝ่ายเลือกใช้กลยุทธ์ในการแข่งขันหลายกลยุทธ์ผสมกันในเกมการแข่งขันเดียวกัน เพื่อให้ฝ่ายของตนอยู่เหนือคู่แข่งขัน ทฤษฎีเกมรูปแบบนี้จะใช้เกณฑ์ความน่าจะเป็น วิธีทางพีชคณิต วิธีกราฟ และการโปรแกรมเชิงเส้น

**ขั้นตอนในการตัดสินใจ**

Anderson, Sweeney, Williams, Camm, and Martin (2012, pp. 3-6) ได้เสนอขั้นตอนของกระบวนการตัดสินใจเป็น 7 ขั้นตอน ดังนี้

1. ระบุปัญหา (Define the Problem) เป็นขั้นตอนแรกในกระบวนการตัดสินใจ โดยผู้ตัดสินใจต้องกำหนดปัญหาและวิเคราะห์สาเหตุของปัญหาให้เกิดความชัดเจน การระบุปัญหาได้ถูกต้องจะส่งผลต่อการดำเนินการในขั้นต่อ ๆ ไปของกระบวนการตัดสินใจ

2. กำหนดทางเลือก (Identify the Alternatives) เป็นขั้นตอนถัดมาหลังจากผู้ตัดสินใจสามารถกำหนดปัญหาได้ชัดเจนแล้ว ซึ่งการกำหนดทางเลือกจะต้องกลั่นกรองข้อมูลต่าง ๆ ที่เกี่ยวข้องกับปัญหาทั้งหมด โดยอาจใช้ความคิดเห็นจากผู้มีประสบการณ์ หรือประสบความสำเร็จในด้านนั้น ๆ หรือใช้ข้อมูลที่ได้จากการจัดการประชุมย่อย ร่วมกับความรู้ ความสามารถ ความคิดสร้างสรรค์ และประสบการณ์ของผู้ตัดสินใจมาพัฒนาเป็นทางเลือก

3. กำหนดเกณฑ์ (Determine the Criteria) เป็นขั้นตอนในการกำหนดเกณฑ์ที่ใช้ในการพิจารณาทางเลือกต่าง ๆ ที่จะใช้แก้ปัญหา

4. ประเมินทางเลือก (Evaluate the Alternatives) เป็นขั้นตอนของการประเมินผลทางเลือกต่าง ๆ ซึ่งเป็นแนวทางการนำปัญหาไปสู่การแก้ไข ในขั้นตอนนี้ผู้ตัดสินใจจะวิเคราะห์และประเมินว่าทางเลือกใดสามารถแก้ไขปัญหาได้ดีที่สุด ทางเลือกใดควรจะดำเนินการก่อนและหลัง และมีกระบวนการพิจารณาถึงผลดีและผลเสียในแต่ละทางเลือกด้วย

5. เลือกทางเลือก (Choose an Alternative) เป็นขั้นตอนในการนำทางเลือกต่าง ๆ มาเปรียบเทียบว่าทางเลือกใดจะเหมาะสมและเป็นไปได้ที่สุด แล้วเลือกทางเลือกนั้นมาใช้แก้ปัญหา

6. ดำเนินการตามทางเลือกที่ตัดสินใจ (Implement the Selected Alternative) หลังจากเลือกทางเลือกที่ดีที่สุด และมีความเหมาะสมที่สุดแล้ว จะนำผลการตัดสินใจสู่การปฏิบัติและประเมินผลต่อไป

7. ประเมินผลลัพธ์ (Evaluate the Results) เป็นขั้นตอนสุดท้ายของกระบวนการตัดสินใจ การประเมินผลเป็นการพิจารณาคุณค่าของผลงานและความแตกต่างระหว่างผลการปฏิบัติงานกับเกณฑ์ และมาตรฐานที่ได้เลือกจากทางเลือกที่ตัดสินใจ โดยผู้ตัดสินใจควรเปรียบเทียบผลงานกับเกณฑ์ และมาตรฐานว่ามีความแตกต่างกันหรือไม่ ความแตกต่างนั้นมีความสำคัญมากน้อยเพียงใด จะก่อนให้เกิดความเสียหายหรือไม่อย่างไร

**งานวิจัยที่เกี่ยวข้องกับทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน โดยใช้เกณฑ์ของเฮอร์วิคซ์**

Pazek and Rozman (2009) ได้ใช้ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน เพื่อใช้ตัดสินใจผลิตสินค้าทางการเกษตร กรณีศึกษา การผลิตน้ำมันฟักทอง (Oil Pumpkin) โดยใช้เกณฑ์การตัดสินใจที่แตกต่างกัน 5 เกณฑ์ ได้แก่ 1) เกณฑ์แมกซิมิน (Maximin Criterion) 2) เกณฑ์แมกซิแมกซ์ (Maximax Criterion) 3) เกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion) โดยใช้สัมประสิทธิ์การมองโลกในแง่ดีเท่ากับ 0.7 4) เกณฑ์มินนิแมกซ์ (Minimax Criterion) และ 5) เกณฑ์ของลาปลาซ (Laplace Criterion) สำหรับตัดสินใจเลือกขนาดพื้นที่เพื่อปลูกฟักทองที่ต่างกัน 3 ขนาด ได้แก่ 5 เฮกเตอร์, 3 เฮกเตอร์ และ 1 เฮกเตอร์ ตามลำดับ และจะพิจารณาผลตอบแทนที่ได้รับจาก

การขายน้ำมันฟักทองที่ผลิตได้ ภายใต้โอกาสทางการตลาด (Market Opportunities) ที่ 100%, 85% และ 50% ตามลำดับ ผลการศึกษา ปรากฏว่า เกณฑ์แมกซิแมกซ์ เฮอร์วิคซ์ และลาปลาซ ให้ผลการตัดสินใจเหมือนกัน คือ เลือกพื้นที่ปลูกฟักทอง 5 เฮกเตอร์ เกณฑ์แมกซิมิน เลือกพื้นที่ปลูกฟักทอง 1 เฮกเตอร์ และเกณฑ์มินนิแมกซ์ เลือกพื้นที่ปลูกฟักทอง 3 เฮกเตอร์

Groenewald and Pretorius (2011) ได้ศึกษากลยุทธ์ในการตัดสินใจเลือกพอร์ตการลงทุน (Portfolio) โดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ที่แตกต่างกัน 8 เกณฑ์ คือ 1) เกณฑ์น้อยสุด (Min Criterion) 2) เกณฑ์มากสุด (Max Criterion) 3) เกณฑ์น้อยสุด/สูงสุด (Min/Max Criterion) 4) เกณฑ์แมกซิแมกซ์ (Maximax Criterion) 5) เกณฑ์แมกซิมิน (Maximin Criterion) 6) เกณฑ์ของลาปลาซ (Laplace Criterion) 7) เกณฑ์มินนิแมกซ์ (Minimax Criterion) และ 8) เกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion) ซึ่งจะพิจารณาผลตอบแทนรายเดือนของตลาดเงิน (the Monthly Returns of the Money Market) ภายใต้ตัวแบบการจัดพอร์ตของมาร์โควิทส์ (Markowitz Portfolio Model) ผลการศึกษา ปรากฏว่า การตัดสินใจเลือกพอร์ตการลงทุน โดยใช้เกณฑ์ของเฮอร์วิคซ์มีความเหมาะสมที่สุด

Gaspars-Wieloch (2014) ได้ปรับปรุงเกณฑ์ของเฮอร์วิคซ์ให้มีประสิทธิภาพยิ่งขึ้น สำหรับการตัดสินใจในกรณีที่มีทางเลือกที่ผลตอบแทนสูงสุดมีค่าต่างจากผลตอบแทนต่ำสุดมากเกินไป ถ้าใช้เกณฑ์ของเฮอร์วิคซ์ส่วนใหญ่จะตัดสินใจเลือกทางเลือกนี้ ซึ่งเป็นทางเลือกที่ไม่เหมาะสม เพราะมีความเสี่ยงมากเกินไป ดังนั้นจึงปรับปรุงเกณฑ์ของเฮอร์วิคซ์ โดยอาศัยแนวคิดตามเกณฑ์ของลาปลาซ ซึ่งดำเนินการโดยใช้ค่าเฉลี่ยของผลตอบแทนในกลุ่มสูงกับผลตอบแทนในกลุ่มต่ำ แทนการใช้ค่าผลตอบแทนสูงสุดและต่ำสุดในแบบเดิม จากนั้นนำเกณฑ์ของเฮอร์วิคที่ได้ปรับปรุงนี้ ใช้ในการตัดสินใจเลือกโครงการลงทุน (Investment Project) ปรากฏว่า โครงการลงทุนที่ตัดสินใจเลือกมีความเหมาะสม เมื่อเทียบกับการตัดสินใจโดยใช้เกณฑ์เฮอร์วิคซ์ดั่งเดิม

Gaspars-Wieloch (2015) ได้เสนอเกณฑ์ใหม่ภายใต้ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน เพื่อค้นหากลยุทธ์ที่มีความเหมาะสมที่สุด โดยใช้ค่าสัมประสิทธิ์ของการมองโลกในแง่ดี (หรือสัมประสิทธิ์ของการมองโลกในแง่ร้าย) และมีการดำเนินการ 2 ขั้นตอน ได้แก่ 1) กำหนดสภาวการณ์ที่คาดว่าจะเกิดขึ้นจริง (The True Scenario Forecasting) ซึ่งขึ้นอยู่กับความต้องการของผู้ตัดสินใจ และ 2) การเลือกทางเลือกที่เหมาะสม (The Appropriate Alternative Selection) ซึ่งพิจารณาจากเมทริกซ์ของผลตอบแทนภายใต้สภาวการณ์จริง หรือภายใต้สภาวการณ์ที่คาดว่าจะเกิดขึ้นจริง เกณฑ์ที่พัฒนาขึ้นนี้ตั้งอยู่บนแนวคิดว่า สภาวการณ์ต่าง ๆ ที่คาดว่าจะเกิดขึ้น ซึ่งแสดงอยู่ในเมทริกซ์ของผลตอบแทน จะเกิดขึ้นจริงเพียงหนึ่งสภาวการณ์เท่านั้น

จากการศึกษางานวิจัยที่เกี่ยวข้องกับทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน โดยใช้เกณฑ์ของเฮอร์วิคซ์ ปรากฏว่า งานวิจัยในอดีตเป็นการนำเกณฑ์ต่าง ๆ ในทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอนมาใช้ตัดสินใจในการดำเนินงานทางธุรกิจ เช่น การตัดสินใจในการลงทุน การตัดสินใจในการผลิต ซึ่งผลจากการวิจัยชี้ให้เห็นว่าผลลัพธ์ในการตัดสินใจตามเกณฑ์ของเฮอร์วิคซ์มีความเหมาะสมมากกว่าการตัดสินใจโดยใช้เกณฑ์อื่น ๆ และในระยะต่องานวิจัยจะมีทิศทางมุ่งพัฒนาเกณฑ์ของเฮอร์วิคซ์ให้มีประสิทธิภาพมากยิ่งขึ้น

## ตอนที่ 3 การศึกษาในสถานการณ์จำลองแบบมอนติคาร์โล

การศึกษาในสถานการณ์จำลองแบบมอนติคาร์โล (Monte Carlo Simulation Study) โดยทั่วไปจะเรียกว่าการศึกษาในสถานการณ์จำลอง (Simulation Study) เป็นการศึกษาโดยการจำลองสถานการณ์เสมือนจริงและดำเนินการทดลองเป็นระบบภายใต้เงื่อนไขที่กำหนด ผลการศึกษาสามารถนำไปอ้างอิงได้กับการทดลองเชิงประจักษ์ ซึ่งจะช่วยแก้ไขปัญหาที่ไม่สามารถปฏิบัติได้ในการทดลองเชิงประจักษ์

การศึกษาในสถานการณ์จำลองเป็นวิธีหนึ่งที่นิยมใช้ในการหาคำตอบทางสถิติ โดยอาศัยการสุ่มจำนวนเพื่อนำมาสร้างค่าของตัวแปรให้เหมือนในสถานการณ์จริง และใช้การทดลองซ้ำหลาย ๆ ครั้ง ด้วยคอมพิวเตอร์ เพื่อให้ได้ค่าที่แน่นอนที่จะเป็นข้อสรุป หรือใช้อธิบายปรากฏการณ์ต่าง ๆ ในสถานการณ์จริง ซึ่งการศึกษาในสถานการณ์จำลองนี้เป็นที่นิยมมาก ในการแก้ปัญหาทางคณิตศาสตร์ วิศวกรรมศาสตร์ หรือทางสถิติ (Frenkel, 2004; Raychaudhuri, 2008)

การศึกษาในสถานการณ์จำลองมีความจำเป็นต่อการพัฒนาการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เป็นอย่างมาก เพราะเป็นขั้นตอนแรกเพื่อตรวจสอบความคุ้มค่าก่อนการพัฒนาแบบทดสอบ โดยเฉพาะอย่างยิ่งผู้พัฒนาที่ยังมีประสบการณ์น้อย (Thompson & Weiss, 2011) นอกจากนี้ยังจำเป็นต่อการปรับปรุงวิธีการเดิมให้มีประสิทธิภาพยิ่งขึ้น หรือพัฒนาวิธีการใหม่ ๆ เพราะช่วยลดความซับซ้อน ลดเวลา และลดค่าใช้จ่ายลงได้มาก เมื่อเทียบกับการทดสอบโดยใช้ข้อมูลจริง เช่น งานของ van der Linden (2003) ได้ปรับปรุงวิธีการซิมสันเฮกเตอร์ เพื่อลดจำนวนรอบในการควบคุมการใช้ข้อสอบ หรืองานของ Gao and Chen (2005) ได้เปรียบเทียบการประมาณค่าพารามิเตอร์ของข้อสอบระหว่างวิธีมาร์จินัลไลค์ลิฮูดสูงสุด (Marginal Maximum Likelihood Estimation) กับวิธีมาร์จินัลเบย์โมเดล (Marginal Bayes Model Estimation) ในโมเดลโลจิสติก 3 พารามิเตอร์

การศึกษาในสถานการณ์จำลองมีขั้นตอนในการดำเนินการคล้ายคลึงกับการศึกษาในสถานการณ์จริง โดย Harwell et al. (1996) ได้เสนอขั้นตอนของการศึกษาในสถานการณ์จำลองตามพื้นฐานทฤษฎีการตอบสนองข้อสอบ ดังนี้

1. การกำหนดปัญหา (Formulating the Problem) โดยผู้วิจัยจะเป็นผู้กำหนดปัญหาที่สนใจ การตั้งคำถามของการวิจัย สมมติฐานของการวิจัย หรือสิ่งที่ต้องการวัด โดยอาศัยการทบทวนวรรณกรรมที่เกี่ยวข้อง ไม่ว่าจะเป็นทฤษฎีที่เกี่ยวข้อง หรือผลการวิจัยในอดีต

2. การออกแบบการศึกษาในสถานการณ์จำลอง (Designing a Monte Carlo Study) มีหลักการเช่นเดียวกันกับการออกแบบการศึกษาในสถานการณ์จริง ซึ่งมีขั้นตอนย่อยที่ต้องพิจารณาดังนี้

2.1 การกำหนดและการตั้งค่าของตัวแปรต้น (Selecting the Independent Variables and Their Values) จะถูกกำหนดตามคำถามของการวิจัย หรือสมมติฐานของการวิจัยที่ตั้งไว้ ซึ่งผู้ศึกษาต้องกำหนดและตั้งค่าของตัวแปรต้นนั้น ๆ เพื่อใช้ในการจำลองข้อมูลในขั้นตอนถัดไป เช่น การศึกษาผลกระทบของความยาวของแบบทดสอบ (Test Lengths) ที่มีผลต่อความแปรปรวนในการแจกแจงก่อนของค่าประมาณอำนาจจำแนกของข้อสอบ (Prior Variances of the Distribution of Discriminantion Parameter Estimation) ซึ่งตัวแปรต้นในการศึกษานี้ คือ

ความยาวของแบบทดสอบ โดยกำหนดค่าเป็นจำนวนเต็ม ซึ่งเป็นตัวเลขสุ่มที่จะจำลองขึ้นในขั้นตอนถัดไป

2.2 การกำหนดรูปแบบการทดลอง (Selecting an Experimental Design) จะถูกกำหนดตามลักษณะของตัวแปรต้น และวัตถุประสงค์ของการศึกษา โดยทั่วไปจะใช้การทดลองแบบแฟคเทอเรียล (Factorial Design) ซึ่งเป็นการศึกษาอิทธิพลของปัจจัย (ตัวแปรต้น) ตั้งแต่ 2 ปัจจัยขึ้นไปพร้อม ๆ กัน แบ่งแยกเป็นการศึกษาอิทธิพลของปัจจัยแต่ละตัว (Main Effect) และการศึกษาอิทธิพลร่วม (Interaction Effect) ระหว่างปัจจัย ซึ่งไม่ว่าผู้ศึกษาจะเลือกใช้การทดลองรูปแบบใด จะมีผลต่อโครงสร้างในการวิเคราะห์ผลลัพธ์ เช่น ถ้าพิจารณาปัจจัย 2 ตัว แต่ละปัจจัยมีค่า 3 ระดับ ต้องใช้การทดลองแบบแฟคเทอเรียล $3^k$ ($3^k$ Factorial Design) แล้วในการวิเคราะห์ผลลัพธ์จะต้องวิเคราะห์ผลทั้งหมด 3×3 = 9 รูปแบบ

2.3 การกำหนดตัวแปรตาม (Selecting Dependent Variables) จะกำหนดตามวัตถุประสงค์ของการศึกษา ซึ่งตัวแปรตามเป็นตัวแปรที่มีการเปลี่ยนแปลงค่าตามการเปลี่ยนแปลงของตัวแปรต้น

2.4 การกำหนดจำนวนของการทดลองซ้ำ (Selecting the Number of Replications) จะเทียบได้กับการกำหนดขนาดกลุ่มตัวอย่างของการทดลองในสถานการณ์จริง ดังนั้นจึงสามารถนำเกณฑ์ต่าง ๆ ที่ใช้ในการกำหนดกลุ่มตัวอย่างในการดำเนินการวิจัยทั่วไป มาใช้กำหนดจำนวนของการทดลองซ้ำได้ ทั้งนี้การกำหนดจำนวนในการทดลองซ้ำมาก ๆ (กลุ่มตัวอย่างมีขนาดใหญ่) จะทำให้ค่าความแปรปรวนของตัวอย่างในการประมาณค่าพารามิเตอร์มีค่าน้อย ซึ่งเป็นสิ่งที่นักสถิติต้องการ เพราะทำให้ผลการวิเคราะห์ทางสถิติมีความเชื่อมั่นสูง (Reliable Result) ดังนั้นโดยทั่วไปการทดลองซ้ำในสถานการณ์จำลองจะกำหนดในระดับพันรอบ

2.5 การกำหนดโมเดลทางคณิตศาสตร์ (Formulating the Mathematical Model) จะมีผลโดยตรงต่อการจำลองข้อมูลของการศึกษาในสถานการณ์จำลอง โดยทฤษฎีการตอบสนองข้อสอบมีโมเดลให้เลือกใช้หลายแบบ เช่น โมเดลปกติสะสม (Normal Ogive Model) หรือโมเดลโลจิสติก (Logistic Model) ซึ่งผู้ศึกษาควรเลือกให้สอดคล้องกับปัญหาหรือวัตถุประสงค์ของการศึกษา

3. การเขียนหรือเลือกใช้โปรแกรมคอมพิวเตอร์ (Writing or Selecting Computer Programs) โดยทั่วไปการศึกษาในสถานการณ์จำลองจะอาศัยโปรแกรมคอมพิวเตอร์ในการจำลองข้อมูล ประมาณค่าพารามิเตอร์ และวิเคราะห์ผลลัพธ์ที่ต้องการ ซึ่งผู้ศึกษาสามารถเขียนโปรแกรมใหม่ขึ้นมาใช้เอง โดยใช้โปรแกรมภาษาคอมพิวเตอร์ทั่วไป เช่น ภาษาปาสคาล (Pascal) ภาษาฟอร์แทรน (Fortran) หรือภาษา MATLAB หรือในกรณีที่ผู้ศึกษาไม่มีความถนัดในการเขียนโปรแกรมอาจเลือกใช้โปรแกรมสำเร็จรูป เช่น โปรแกรม WinGen3 (Han, 2007) โปรแกรม PARDSIM (Yoes, 1997) นอกจากนี้ มีโปรแกรมเฉพาะสำหรับจำลองสถานการณ์ของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เช่น โปรแกรม FireStar (Choi, 2009) โปรแกรม CATSim (Weiss & Guyer, 2012) เป็นต้น การจำลองข้อมูลตามทฤษฎีการตอบสนองข้อสอบ สามารถแบ่งได้ 2 ส่วน ดังนี้

3.1 การจำลองผลการตอบข้อสอบ (Generating Item Responses) ของผู้สอบในแต่ละข้อ ซึ่งมี 2 ค่า คือ 0 (ตอบผิด) หรือ 1 (ตอบถูก) โดยผู้ศึกษาสามารถเขียนโปรแกรมเพื่อจำลองผลการตอบข้อสอบได้ด้วยแนวคิดดังนี้ (Thompson & Weiss, 2011)

สมมติว่า ผู้สอบที่มีความสามารถปานกลาง ($\theta$ = 0.0) มีโอกาสที่จะตอบคำถามถูกเท่ากับ 0.75 (คำนวณจากฟังก์ชันการตอบสนองข้อสอบ ซึ่งแทนด้วยค่า $\theta$ และค่าพารามิเตอร์ของข้อสอบข้อนั้น) แล้วสุ่มเลขที่มีการแจกแจงแบบยูนิฟอร์ม (Uniform distribution) ในช่วง 0 ถึง 1 มา 1 ค่า ถ้าเลขสุ่มมีค่าน้อยกว่าหรือเท่ากับ 0.75 ให้กำหนดผลการตอบข้อสอบเป็นตอบถูก แต่ถ้าเลขสุ่มมีค่ามากกว่า 0.75 ให้กำหนดผลการตอบข้อสอบเป็นตอบผิด

3.2 การประมาณพารามิเตอร์ของข้อสอบ (Estimating Model Parameters) ในสถานการณ์จริงจะนำผลการตอบของผู้สอบจำนวนมากไปประมาณค่าพารามิเตอร์ของข้อสอบ โดยใช้โปรแกรมสำเร็จรูป เช่น BILOG MULTILOG หรือ XCALIBRE ด้วยวิธีการเดียวกันนี้ สำหรับการศึกษาในสถานการณ์จำลองจะใช้ชุดผลการตอบข้อสอบที่จำลองขึ้น (จากข้อ 3.1) มาประมาณค่าพารามิเตอร์ของข้อสอบด้วยโปรแกรมสำเร็จรูปเช่นเดียวกับการประมาณค่าในสถานการณ์จริง แต่เพื่อความสะดวกสามารถใช้โปรแกรม WinGen3 ซึ่งเป็นโปรแกรมที่สามารถจำลองค่าความสามารถของผู้สอบ ผลการตอบข้อสอบ และค่าประมาณพารามิเตอร์ของข้อสอบได้ในคราวเดียวกัน

4. การวิเคราะห์ผลการศึกษาในสถานการณ์จำลอง (Analyzing the Results of a Monte Carlo Study) ส่วนใหญ่ใช้สถิติเชิงพรรณนา (Descriptive Statistic) เช่น ค่าเฉลี่ย ตาราง แจกแจงความถี่ หรือกราฟ ซึ่งเป็นการแสดงผลลัพธ์โดยรวม ดังนั้น ผู้ศึกษาสามารถใช้สถิติเชิงอนุมาน (Inferential Statistic) ที่เหมาะสม เช่น การวิเคราะห์ความแปรปรวน การวิเคราะห์การถดถอย มาวิเคราะห์ร่วมกับสถิติเชิงพรรณนา ซึ่งจะทำให้ผลการวิเคราะห์มีความน่าเชื่อถือยิ่งขึ้น

## ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

งานวิจัยที่เกี่ยวข้องกับการวิจัยนี้ แบ่งได้ 2 ส่วน ดังนี้

**1. งานวิจัยที่เกี่ยวข้องกับวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

โสพส สุขานนท์สวัสดิ์ และคณะ (2556) ได้เสนอวิธีการคัดเลือกข้อสอบข้อถัดไปที่ใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (Decision Making Under Risk) ตามเกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุด (Maximum Expected Monetary Value: EMV) โดยจะเลือกข้อสอบข้อที่มีค่าสัมประสิทธิ์การกระจาย (Coefficient of Variation) สูงสุด มาเป็นข้อสอบข้อถัดไป ซึ่งจากการศึกษาเปรียบเทียบวิธีการนี้กับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด กับกลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาตรี มหาวิทยาลัยสวนดุสิต กรุงเทพมหานคร ที่ลงทะเบียนวิชาภาษาอังกฤษเพื่อทักษะการเรียนของนักศึกษาในระดับปริญญาตรี จำนวน 280 คน ผลการวิจัยสรุปได้ว่า จำนวนข้อสอบที่ใช้และเวลาในการทดสอบของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจมีค่าน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

Chen and Ankenman (2004) ได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 4 วิธี ได้แก่ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศของฟิชเชอร์ (Fisher information: F) 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (Fisher Information with a Posterior Distribution: FP) 3) วิธีการคัดเลือก

ข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศของคูลแบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (Kullback-Leibler Information with a Posterior Distribution: KP) และ 4) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้การสุ่มข้อสอบแบบสมบูรณ์ (Completely Randomized Item Selection: RN) ซึ่งจะเปรียบเทียบประสิทธิภาพการประมาณค่าความสามารถของผู้สอบ และการใช้ข้อสอบ โดยจะศึกษาภายใต้เงื่อนไข 3 เงื่อนไข ได้แก่ 1) ใช้เฉพาะค่าสารสนเทศของแต่ละวิธีการคัดเลือกข้อสอบข้อถัดไป 2) ใช้ค่าสารสนเทศ และมีการควบคุมสัดส่วนเนื้อหาของข้อสอบ (Content Balancing) และ 3) ใช้ค่าสารสนเทศ การควบคุมสัดส่วนเนื้อหาของข้อสอบ และการควบคุมการใช้ข้อสอบ (Item Exposure Control) ในการทดสอบที่ใช้ข้อสอบน้อยกว่า 10 ข้อ ปรากฏว่า วิธี FP และ KP มีแนวโน้มในการประมาณค่าความสามารถของผู้สอบดีกว่าวิธี F ภายใต้เงื่อนไขข้อที่ 1 แต่ไม่สามารถสรุปได้ว่าวิธี FP และ KP มีประสิทธิภาพดีกว่าวิธี F โดยเฉพาะเมื่อมีการควบคุมการใช้ข้อสอบ ส่วนในการทดสอบที่ใช้ข้อสอบมากกว่า 10 ข้อ วิธี FP KP และวิธี F มีประสิทธิภาพใกล้เคียงกันไม่ว่าจะอยู่ภายใต้เงื่อนไขใด แต่วิธี F มีการใช้ข้อสอบสูงกว่าวิธีอื่น ๆ เล็กน้อย

Lilley et al. (2004) ได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เพื่อวัดความรู้ทางภาษาอังกฤษของนักศึกษานานาชาติ ใน University of Hertfordshire ประเทศอังกฤษ โปรแกรมการทดสอบนี้ ใช้คลังข้อสอบขนาด 250 ข้อ การคัดเลือกข้อสอบข้อแรกใช้วิธีการสุ่มข้อสอบที่มีค่าความยากของข้อสอบ ตั้งแต่ -0.50 ถึง 0.50 ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปที่ระดับความยากของข้อสอบ (b) ใกล้เคียงกับค่าความสามารถของผู้สอบ การประมาณค่าความสามารถของผู้สอบใช้วิธีความน่าจะเป็นสูงสุด (Maximum Likelihood) และเกณฑ์ยุติการทดสอบใช้เกณฑ์ความยาวคงที่ (Fixed-Length) ที่ 10 ข้อ การประเมินการใช้งานโปรแกรม จากกลุ่มตัวอย่างซึ่งเป็นอาจารย์ 11 คน และนักศึกษานานาชาติ 27 คน ปรากฏว่า การใช้งานโปรแกรมเป็นที่น่าพอใจอยู่ในระดับดี (Mean = 3.9-4.5) และทำการเปรียบเทียบผลการทดสอบระหว่างการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) กับการทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) ปราฏว่า จำนวนข้อสอบที่ผู้สอบตอบถูกระหว่างการทดสอบแบบ CAT กับการทดสอบแบบ CT ไม่มีความแตกต่างกันอย่างมีนัยสำคัญที่ .01 และมีความสัมพันธ์กันที่ระดับนัยสำคัญ .001 แสดงให้เห็นว่า ผลการทดสอบแบบ CAT มีความเท่าเทียมและสอดคล้องในทางเดียวกันกับการทดสอบแบบ CT

Costa, Karino, Moura, and Andrade (2009) ได้เปรียบเทียบวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 3 วิธีการ ได้แก่ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์คูลแบค-ไลเบลอร์ (Kullback-Leibler Criterion: KL) และ 3) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าคาดหวังของค่าสารสนเทศสูงสุด (Maximum Expected Information Criterion: MEI) โดยใช้คลังข้อสอบวิชาภาษาอังกฤษ จำนวน 246 ข้อ และศึกษาเปรียบเทียบในสถานการณ์จำลอง 5 กรณี คือ กรณีที่ 1 เปรียบเทียบในด้านความยาวของแบบทดสอบ โดยใช้เกณฑ์ยุติการทดสอบที่ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถน้อยกว่า 0.4 ปรากฏว่า ทั้ง 3 วิธีการ ใช้ข้อสอบโดยเฉลี่ยใกล้เคียงกัน คือ 23, 22 และ 21 ข้อ ตามลำดับ กรณีที่ 2 ศึกษาความสัมพันธ์ระหว่างค่าความสามารถที่แท้จริงกับค่าประมาณความสามารถของ

ผู้สอบ เมื่อกำหนดค่าประมาณความสามารถเริ่มต้นเท่ากับ 0.00 และใช้เกณฑ์ยุติการทดสอบที่ 25 ข้อ ปรากฏว่า ทั้ง 3 วิธีการ มีค่าความสัมพันธ์สูงมากใกล้เคียงกัน กรณีที่ 3) เปรียบเทียบผลการทดสอบ เมื่อกำหนดค่าประมาณความสามารถเริ่มต้นแตกต่างกัน 3 ค่า คือ -1.50, 0.00 และ 1.50 และใช้เกณฑ์ยุติการทดสอบที่ 25 ข้อ ปรากฏว่า เมื่อเสร็จสิ้นการทดสอบ ทั้ง 3 วิธีการ ให้ค่าประมาณความสามารถใกล้เคียงกัน กรณีที่ 4 เปรียบเทียบผลการทดสอบ เมื่อกำหนดข้อสอบข้อเริ่มต้นมีค่าความยากของข้อสอบแตกต่างกัน 3 แบบ คือ มีค่าน้อยกว่า, เท่ากับ และมากกว่าค่าความสามรถจริงของผู้สอบ ปรกฏว่า ทั้ง 3 วิธีการ ให้ค่าประมาณความสามารถใกล้เคียงกัน และกรณีที่ 5 ศึกษาประสิทธิภาพการประมาณค่าความสามารถของผู้สอบ ซึ่งพิจารณาจากค่าความลำเอียงเฉลี่ย และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เมื่อกำหนดค่าความสามารถที่แท้จริงของผู้สอบแตกต่างกัน 10 ค่า ปรากฏว่า ทั้ง 3 วิธีการให้ค่าความลำเอียงเฉลี่ย และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยใกล้เคียงกัน

Fan, Wang, Chang, and Douglas (2012) ได้เสนอวิธีการคัดเลือกข้อสอบข้อถัดไป วิธีการใหม่ 2 วิธีการ ได้แก่ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดต่อหน่วยเวลา (Maximum Information per Time Unit) ซึ่งพัฒนาจากวิธีการคัดเลือกข้สอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) และ 2) วิธีการ Time Weighted a-Stratification ซึ่งพัฒนามาจากวิธีการควบคุมการใช้ข้อสอบ a-Stratified Strategy (ASTR) โดยวิธีการใหม่ที่ได้พัฒนาขึ้น จะคัดเลือกข้อสอบข้อถัดไป โดยพิจารณาจากเวลาที่ใช้ในการทำข้อสอบควบคู่กับค่าสารสนเทศของข้อสอบ เพื่อแก้ไขปัญหาที่ผู้สอบบางคนได้รับข้อสอบที่ใช้เวลาในการทำข้อสอบนานกว่าคนอื่นๆ หลังจากนั้น นำวิธีการที่พัฒนาขึ้นมาเปรียบเทียบกับวิธีการดั้งเดิมด้วยการศึกษาในสถานการณ์จำลอง จำนวน 1,000 ครั้ง และใช้คลังข้อสอบขนาด 500 ข้อ ปรากฏว่า การดำเนินการทดสอบด้วยวิธีการใหม่มีสมดุลด้านเวลาเพิ่มขึ้น แต่ยังคงมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบ และการควบคุมการใช้ข้อสอบใกล้เคียงกับวิธีการดั้งเดิม

Murphy, Dodd, and Vaughn (2010) ได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 3 วิธี ได้แก่ 1) วิธี Maximum Fisher's Information (MFI) 2) วิธี Maximum Posterior Weighted Information (MPWI) และ 3) วิธี Minimum Expected Posterior Variance (MEPV) ในกรณีที่คลังข้อสอบถูกแบ่งเป็นคลังข้อสอบย่อย (Testlet) ภายใต้โมเดลการตอบสนองข้อสอบ (IRT) และโมเดลการตอบสนองข้อสอบที่ใช้แบบทดสอบย่อย (Testlet Response Theory: TRT) ดำเนินการวิจัยโดยใช้การศึกษาในสถานการจำลอง โดยจำลองแบบแผนการตอบของผู้สอบ จำนวน 10 กลุ่ม แต่ละกลุ่มมีผู้สอบ 1,000 คน ผลการศึกษาปรากฏว่า เมื่อใช้โมเดล IRT กับข้อมูลที่ฝ่าฝืนข้อตกลงเบื้องต้น ความเป็นอิสระระหว่างข้อสอบ ผลการประมาณค่าความสามารถของผู้สอบมีแนวโน้มที่จะมีค่ามากกว่าค่าจริง และผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ปรากฏว่า วิธีการทั้ง 3 วิธี มีประสิทธิภาพไม่แตกต่างกัน เนื่องจากทั้ง 3 วิธี ต่างทำให้ผู้สอบแต่ละคนได้รับข้อสอบจำนวนมาก (ประมาณ 50 ข้อ) เป็นเหตุให้ค่าประมาณความสามารถของผู้สอบลู่เข้าหาค่าความสามารถที่แท้จริงของผู้สอบ

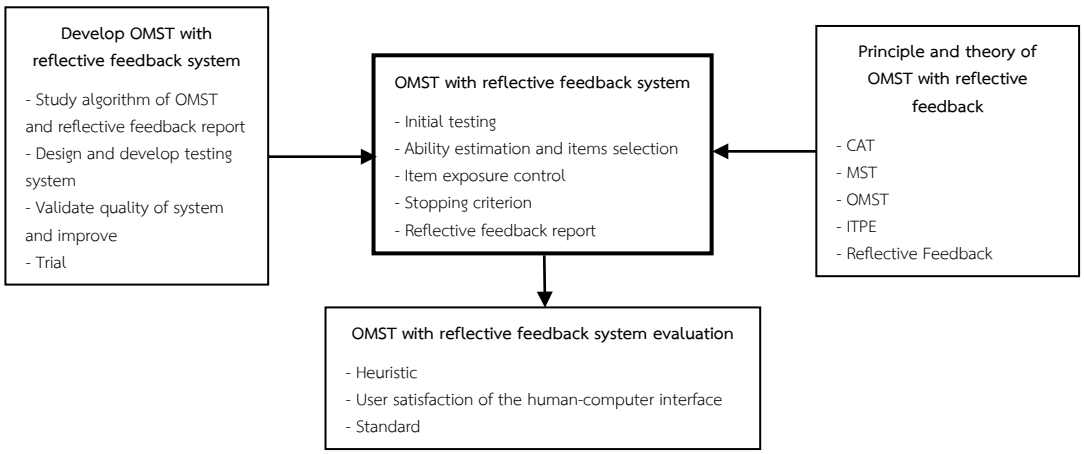
Wang, Kuo, Chao, and Tsai (2012) ได้พัฒนาและประเมินผลระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวัดความสามารถทางภาษาจีน ภายใต้กรอบมาตรฐาน

การประเมินความสามารถทางภาษาในกลุ่มสหภาพยุโรป (The Common European Framework of Reference: CEFR) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นี้ ถูกพัฒนาในรูปแบบระบบการทดสอบบนเว็บ (Web – Base Test System) สำหรับวัดความสามารถทางภาษาจีนในระดับ A1 และ A2 (จาก 6 ระดับ ได้แก่ A1, A2, B1, B2, C1 และ C2 ตามลำดับ) โดยเป็นการวัดทักษะการฟังของผู้สอบ ประกอบด้วยคำถาม 2 ลักษณะ ได้แก่ 1) คำถามวัดความเข้าใจในการฟัง (Listening Comprehension Item) มีลักษณะให้ผู้สอบฟังวลี หรือบทสนทนา แล้วให้เลือกคำตอบที่ถูกที่สุดเพียงข้อเดียว จากตัวเลือกทั้งหมด 4 ตัวเลือก ซึ่งคำถามแต่ละข้อถูกจำกัดเวลาในการตอบ หากหมดเวลา ผู้สอบจะได้รับคำถามข้อถัดไปทันที 2) คำถามวัดความเข้าใจในการฟัง-มอง (Visual-Listening Comprehension Item) มีลักษณะให้ผู้สอบฟังวลี หรือบทสนทนา แล้วให้เลือกรูปภาพที่สอดคล้องกับวลี หรือบทสนทนานั้น ซึ่งมีตัวเลือก (รูปภาพ) ทั้งหมด 4 ตัวเลือก กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนระดับมัธยม เกรด 5 ถึงเกรด 10 ของโรงเรียน Grace Christian Chinese School ในประเทศฟิลิปปินส์ ซึ่งนักเรียนเกรด 5 ถึงเกรด 7 ได้ทดสอบในระดับ A1 ส่วนนักเรียนเกรด 8 ถึงเกรด 10 ได้ทดสอบในระดับ A2 ผลการวิเคราะห์ความเที่ยงของคำถาม ปรากฏว่า คำถามใน A1 และ A2 มีค่าสัมประสิทธิ์แอลฟามากกว่า 0.8 แสดงว่า คำถามใน A1 และ A2 มีความเที่ยงสูง

Koedsri, Lawthong, and Ngudgratoke (2014) ได้ศึกษาประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ที่มีความยาวแบบยืดหยุ่น (Variable-Length) โดยใช้โมเดลการตอบสนองข้อสอบที่ใช้แบบทดสอบย่อย (Testlet Response Model: TRT) และใช้วิธีการแบ่งกลุ่มค่าอำนาจจำแนกของข้อสอบแบบถ่วงน้ำหนักที่มีการบังคับ (The Constraint-Weighted a-Stratification Method: CWA) เป็นวิธีการคัดเลือกข้อสอบข้อถัดไป เนื่องจากความเป็นอิสระระหว่างข้อสอบเป็นข้อตกลงเบื้องต้นที่สำคัญของทฤษฎีการตอบสนองข้อสอบแบบดั้งเดิม (Traditional Item Response Theory) การละเมิดข้อตกลงนี้มีผลต่อการประมาณค่าความสามารถของผู้สอบ ดังนั้น โมเดล TRT จึงถูกพัฒนาขึ้นเพื่อแก้ข้อบกพร่องนี้ ด้วยเหตุนี้ การวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษาความถูกต้องของวิธีการคัดเลือกข้อสอบข้อถัดไป ด้วยวิธี CWA ในโมเดล TRT ที่ใช้กับแบบทดสอบที่มีความยาวยืดหยุ่น ซึ่งวิธี CWA เป็นวิธีการคัดเลือกข้อสอบข้อถัดไป ที่มีประสิทธิภาพสูง สามารถควบคุมการใช้ข้อสอบ และควบคุมสัดส่วนเนื้อหาของข้อสอบ ได้พร้อมกันด้วยวิธีเดียว ดำเนินการวิจัยด้วยการศึกษาในสถานการณ์จำลองแบบมอนติคาร์โล ประกอบด้วยคลังข้อสอบจำลอง 1,000 ข้อ และค่าความสามารถที่แท้จริงของผู้สอบ 1,000 คน โดยศึกษาเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 3 วิธี ได้แก่ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้การคัดเลือกอย่างสุ่ม (Randomization Criterion: RAN) 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) และวิธีการแบ่งกลุ่มค่าอำนาจจำแนกของข้อสอบแบบถ่วงน้ำหนักที่มีการบังคับ (The Constraint-Weighted a-Stratification Method: CWA) ผลปรากฏว่า สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีความยาวแบบยืดหยุ่น และใช้โมเดลการตอบสนองข้อสอบที่ใช้แบบทดสอบย่อย วิธี CWA เป็นวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีประสิทธิภาพสูงสุด เมื่อเทียบกับวิธีการ RAN และวิธีการ MIC

Chaimongkol, Pasiphol, and Kanjanawasee (2016) ได้เสนอกรอบแนวคิดในการพัฒนาการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีการสะท้อนข้อมูลย้อนกลับ (Reflective

Feedback) ให้ผู้สอบภายหลังการทดสอบเสร็จสิ้น พร้อมทั้งนำวิธี On-the-Fly Assemble Multistage Adaptive testing (OMST) มาใช้ในการคัดเลือกข้อสอบข้อถัดไป ซึ่งวิธี OMST นอกจากเป็นวิธีการคัดเลือกข้อสอบข้อถัดไปแล้ว ยังเป็นวิธีที่ใช้ปรับปรุงประสิทธิภาพการประมาณค่าความสามารถของผู้สอบ คลังข้อสอบที่นำมาใช้เป็นข้อสอบสำหรับการทดสอบมาตรฐานวิชาชีพไอที ระดับไอพี (Information Technology Porfessional Examination: Level IP) โดยกรอบแนวคิดในการพัฒนา ประกอบด้วย 1) การศึกษาหลักการและทฤษฎีของวิธี OMST ที่มีการสะท้อนข้อมูลย้อนกลับ (Principle and Theory of OMST with Reflective Feedback) 2) การพัฒนาระบบวิธี OMST ที่มีการสะท้อนข้อมูลย้อนกลับ (Develop OMST with Reflective Feedback) 3) ระบบวิธี OMST ที่มีการสะท้อนข้อมูลย้อนกลับ (OMST with Reflective Feedback System) และ 4) การประเมินระบบวิธี OMST ที่มีการสะท้อนข้อมูลย้อนกลับ (OMST with Reflective Feedback System Evaluation) แสดงกรอบแนวคิดดังภาพที่ 2-5




ภาพที่ 2-5 กรอบแนวคิดการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีการสะท้อนข้อมูลย้อนกลับ เสนอโดย Chaimongkol et al. (2016)

Tseng (2016) ได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวัดคำศัพท์ภาษาอังกฤษ เพื่อทดแทนการทดสอบแบบกระดาษและดินสอ (Paper and Pencil: P&P) ซึ่งมีความยุ่งยากและเสียค่าใช้จ่ายมาก การดำเนินการพัฒนาแบ่งเป็น 3 ส่วน ได้แก่ ส่วนที่ 1 การพัฒนาคลังข้อสอบ ประกอบด้วยข้อสอบจำนวนทั้งสิ้น 240 ข้อ มีลักษณะเป็นรายการคำตอบแบบ 4 ตัวเลือก แบ่งเป็นชุดข้อสอบแบบคู่ขนานได้ 5 ชุด โดย 4 ชุดแรกมีข้อสอบชุดละ 45 ข้อ และชุดสุดท้ายเป็นข้อสอบพื้นฐาน 15 ข้อ วิเคราะห์ข้อสอบโดยใช้โมเดลโลจิสติกแบบหนึ่งพารามิเตอร์ หรือโมเดลราส์ช (One-Parameter Logistic Model or Rasch Model) กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนระดับมัธยมปลายในไต้หวัน จำนวน 1,536 คน ผลการวิเคราะห์ความเที่ยงปรากฏว่า ข้อสอบในแต่ละชุดมีความเที่ยงสูง และโมเดลราส์ชมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ส่วนที่ 2 การพัฒนาขั้นตอนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ประกอบด้วยวิธีการคัดเลือกข้อสอบข้อถัดไปโดยการสุ่มข้อสอบหนึ่งข้อ จากกลุ่มข้อสอบที่มีค่าความยากของข้อสอบ

แตกต่างจากค่าประมาณความสามารถของผู้สอบในช่วง -0.5 ถึง 0.5 ใช้วิธีของเบส์ (Bayesian) ในการประมาณค่าความสามารถของผู้สอบ และใช้เกณฑ์ยุติการทดสอบ 2 แบบ คือ เกณฑ์ยุติการทดสอบแบบความยาวคงที่ (Fixed-Length) ที่ 30 ข้อ ร่วมกับเกณฑ์ยุติการทดสอบแบบความยาวยืดหยุ่น (Variable-Length) ที่ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ น้อยกว่า 0.3 และส่วนที่ 3 การจำแนกกลุ่มผู้สอบผ่านกับไม่ผ่าน จากการประเมินผลการใช้งานโปรแกรม ปรากฏว่า โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวัดคำศัพท์ภาษาอังกฤษ ใช้ข้อสอบเพียง 1 ใน 3 ของการทดสอบแบบกระดาษและดินสอ รวมทั้งการประมาณค่าความสามารถของผู้สอบมีความแม่นยำสูง

จากการศึกษางานวิจัยที่เกี่ยวข้องกับวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่า ในช่วงเริ่มแรก ส่วนใหญ่เป็นงานวิจัยที่เกี่ยวข้องกับการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปขึ้นใหม่ ส่วนงานวิจัยในช่วงต่อมา ส่วนใหญ่เป็นการศึกษาเพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีผู้พัฒนาเอาไว้แล้ว โดยมีจุดมุ่งหมายเพื่อให้ทราบข้อดี หรือข้อจำกัดที่เกิดขึ้นในการนำวิธีการนั้น ๆ ไปใช้งานจริง ซึ่งการศึกษาโดยใช้ข้อมูลจริง หรือศึกษาในสถานการณ์จำลอง ควบคู่กับการนำวิธีการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้นมาแล้ว ไปใช้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เพื่อวัดความรู้ความสามารถของผู้สอบในด้านใดด้านหนึ่ง

**2. งานวิจัยที่เกี่ยวข้องกับวิธีการควบคุมการใช้ข้อสอบ ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

Leung, Chang, and Hau (2000) ได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบระหว่างวิธีการ a-Stratified Strategy (ASTR) กับวิธีการ a-Stratified with b-Blocking (BASTR) ในกรณีที่ค่าอำนาจจำแนกของข้อสอบกับค่าความยากของข้อสอบมีความสัมพันธ์กัน โดยดำเนินการศึกษาในสถานการณ์จำลอง ด้วยการจำลองค่าความสามารถที่แท้จริงของผู้สอบ 5,000 ค่า แต่ใช้คลังข้อสอบจริง ซึ่งเป็นข้อสอบคณิตศาสตร์ระดับเบื้องต้น จำนวน 700 ข้อ ผลการศึกษา ปรากฏว่า วิธีการ BASTR มีประสิทธิภาพในการควบคุมการใช้ข้อสอบได้ดีกว่าวิธีการ ASTR

Chang, Qian, and Ying (2001) ได้เสนอวิธี a-Stratified Multistage Computerized Adaptive Testing with b Blocking (BASTR) เพื่อควบคุมการใช้ข้อสอบ ซึ่งวิธีการนี้ปรับปรุงมาจากวิธี a-Stratified Adaptive Testing (ASTR) โดยมีแนวคิด คือ ทำการแบ่งข้อสอบในคลังข้อสอบออกเป็นชั้น ๆ ตามค่าพารามิเตอร์อำนาจการจำแนก และค่าพารามิเตอร์ความยากของข้อสอบ เมื่อได้ชั้นของข้อสอบตามที่ต้องการแล้ว จะทำการคัดเลือกข้อสอบในแต่ละชั้นให้ผู้สอบ ซึ่งผู้สอบต้องทำข้อสอบในทุกชั้น ผลการศึกษา ปรากฏว่า วิธีการ BSTR ให้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถที่แท้จริงกับค่าประมาณความสามารถของผู้สอบสูงกว่า มีค่าความลำเอียงเฉลี่ยและค่าความคลาดเคลื่อนกำลังสองเฉลี่ยน้อยกว่า เมื่อเทียบกับวิธีการ ASTR แสดงว่า วิธีการ BASTR มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าวิธีการ ASTR นอกจากนี้ วิธีการ BASTR สามารถลดอัตราการใช้ข้อสอบได้ร้อยละ 74 เมื่อเทียบกับวิธีการ ASTR

Hau and Chang (2001) ได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบ 3 วิธี ได้แก่ วิธี 1) Descending a-Stratified (ADSTR) 2) วิธี Non Systematic Stratified (NSTR) และ 3) วิธี a-Stratified Adaptive Testing (ASTR) ผลการศึกษาปรากฏว่า วิธี ASTR และวิธี ADSTR ให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยและค่าความลำเอียงเฉลี่ยมีแนวโน้มลดลงเหมือนกัน เมื่อข้อสอบที่ใช้ในการทดสอบมีจำนวนเพิ่มขึ้น แต่ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธี ASTR มีค่าน้อยกว่าวิธี ADSTR ส่วนวิธี NSTR ให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยอยู่กึ่งกลางระหว่างวิธี ASTR กับวิธี ADSTR

Leung, Chan, and Hau (2002) ได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบ 3 วิธี ได้แก่ 1) วิธี a-Stratified Adaptive Testing (ASTR) 2) วิธีการร่วมระหว่างวิธี a-Stratified Adaptive Testing (ASTR) กับวิธี Sympson-Hetter (ASTR-SH) และ 3) วิธีการร่วมระหว่างวิธี Maximum Information กับวิธี Sympson-Hetter (MI-SH) ผลการศึกษาปรากฏว่า ค่าประมาณความสามารถของผู้สอบที่ได้จาก 3 วิธี มีความสัมพันธ์กับค่าความสามารถที่แท้จริงของผู้สอบสูง ค่าความลำเอียงเฉลี่ยมีค่าเข้าใกล้ 0 ในกรณีที่แบบทดสอบมีจำนวน 40 ข้อ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธี MI-SH มีค่าน้อยกว่าวิธี ASTR และวิธี ASTR-SH แต่ในกรณีที่แบบทดสอบมีจำนวน 60 ข้อ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธี ASTR มีค่าน้อยกว่าวิธี ASTR-SH กับวิธี MI-SH ส่วนจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบสูง ปรากฏว่า วิธี ASTR มีจำนวนน้อยสุด เมื่อเทียบกับวิธี ASTR-SH กับวิธี MI-SH ส่วนระยะเวลาที่ใช้ในการคำนวณค่าพารามิเตอร์ควบคุมการใช้ข้อสอบ ปรากฏว่า วิธี ASTR-SH ใช้เวลาน้อยกว่าวิธี MI-SH ส่วนการควบคุมสัดส่วนเนื้อหาของข้อสอบ ปรากฏว่า วิธี ASTR-SH สามารถควบคุมสัดส่วนเนื้อหาของข้อสอบได้ดีกว่าวิธี MI-SH

Chang and van der Linden (2003) ได้ปรับปรุงวิธีการ a-Stratified Adaptive Testing (ASTR) ซึ่งเป็นวิธีการควบคุมการใช้ข้อสอบ โดยแบ่งข้อสอบในคลังข้อสอบออกเป็นชั้น ๆ เพื่อกระจายการคัดเลือกข้อสอบข้อถัดไปให้ทั่วทุกชั้น โดยวิธีการนี้มีจุดด้อยตรงที่ไม่มีการระบุจำนวนชั้นที่เหมาะสมในการแบ่งคลังข้อสอบ และไม่มีความชัดเจนในการจำแนกข้อสอบลงในแต่ละชั้น ดังนั้นจึงทำการปรับปรุงโดยอาศัยการโปรแกรมเชิงเส้น เพื่อหาจำนวนชั้นที่เหมาะสมในการแบ่งข้อสอบในคลังข้อสอบ ทั้งนี้หลังจากการปรับปรุง พบว่า ข้อสอบในคลังข้อสอบมีอัตราการใช้ข้อสอบน้อยลง เมื่อเทียบกับวิธีการดั้งเดิม และวิธีการใหม่นี้ยังคงประมาณค่าความสามารถของผู้สอบได้อย่างมีประสิทธิภาพ

Chang and Ansley (2003) ได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบ 5 วิธี ได้แก่ 1) วิธีการ 5-4-3-2-1 2) วิธีการ Sympson Hetter (SH) 3) วิธีการ Davey-Parshall (DP) 4) วิธีการ Stocking and Lewis Unconditional Multinomial (SL) และ 5) วิธีการ Stocking and Lewis Conditional Multinomial (SLC) โดยศึกษาในสถานการจำลอง กำหนดคลังข้อสอบจำลองขนาด 360 ข้อ และ 720 ข้อ และกำหนดอัตราการใช้ข้อสอบที่พึงปรารถนาไว้ที่ 0.1 และ 0.2 ยกเว้นวิธีการ 5-4-3-2-1 ที่ไม่ต้องกำหนดอัตราการใช้ข้อสอบที่ปรารถนาไว้ล่วงหน้า ใช้วิธี Maximum Likelihood Estimation (MLE) เป็นวิธีการประมาณค่าความสามารถของผู้สอบ กำหนดความสามารถของผู้สอบเริ่มต้นที่ 0 และกำหนดเกณฑ์ยุติการทดสอบแบบความยาวคงที่ จำนวน 30

ข้อ ผลการศึกษาแสดงให้เห็นว่า 1) วิธีการ 5-4-3-2-1 ไม่รับรองว่าจะสามารถควบคุมการใช้ข้อสอบได้ตามต้องการ ส่วนวิธี SH และ SL มีประสิทธิภาพการควบคุมการใช้ข้อสอบเท่าเทียมกัน ในกรณีที่ใช้คลังข้อสอบขนาดเล็ก ภายใต้อัตราการใช้ข้อสอบที่พึงปรารถนาเท่ากับ 0.1 และเมื่อเพิ่มอัตราการใช้ข้อสอบที่พึงปรารถนาเท่ากับ 0.2 วิธีการ SH มีประสิทธิภาพการควบคุมการใช้ข้อสอบสูงสุด ส่วนวิธี SLC มีประสิทธิภาพสูงเป็นที่น่าพอใจ ในการควบคุมการใช้ข้อสอบ 2) ในกรณีที่คลังข้อสอบมีขนาดใหญ่ไม่เพียงพอ (360 ข้อ) วิธีการ 5-4-3-2-1 ไม่รับรองการควบคุมการใช้ข้อสอบ ส่วนวิธีอื่น ๆ ปรากฏว่าอัตราการใช้ข้อสอบเฉลี่ยอยู่ในระดับที่ยอมรับได้ แต่เมื่อเพิ่มขนาดคลังข้อสอบเป็น 2 เท่า (720 ข้อ) ปรากฏว่า วิธีการ SH และ SL มีประสิทธิภาพลดลง แต่วิธีการ DP และ SLC กลับมีประสิทธิภาพการควบคุมการใช้ข้อสอบเพิ่มขึ้น

Lee, IP, and Fuht (2008) ได้ปรับปรุงวิธีการ a-Stratified Adaptive Testing (ASTR) เพื่อควบคุมการใช้ข้อสอบสำหรับแบบทดสอบหลายมิติ (Multi-Dimensional Test) โดยใช้เวคเตอร์ของค่าพารามิเตอร์มาทดแทนค่าพารามิเตอร์ค่าเดียวในหนึ่งมิติ ซึ่งหลังจากการปรับปรุง ปรากฏว่าวิธีการใหม่สามารถควบคุมการใช้ข้อสอบได้เป็นอย่างดี และยังสามารถประมาณค่าความสามารถของผู้สอบได้อย่างมีประสิทธิภาพ

Deng, Ansley, and Chang (2010) ได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการ 3 วิธี ที่ใช้ในการควบคุมการใช้ข้อสอบ คือ 1) a-Stratified Adaptive Testing (ASTR) 2) การคัดเลือกอย่างสุ่ม (RAN) และ 3) The Low a-Stratified (USTR) โดยพิจารณาจากค่าความแปรปรวนของความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ และค่าความน่าเชื่อถือของค่าประมาณความสามารถของผู้สอบ ผลจากการศึกษา ปรากฏว่า วิธี USTR สามารถลดความแปรปรวนของค่าความคลาดเคลื่อนในการประมาณค่าความสามารถได้เมื่อคลังข้อสอบมีขนาดเล็ก และการประมาณค่าความสามารถของผู้สอบทั้ง 3 วิธี มีประสิทธิภาพไม่แตกต่างกัน

Leroux, Lopez, Hembry, and Dodd (2013) ได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบระหว่างวิธีการ Progressive-Restricted Standard Error (PR-SE) กับวิธีการควบคุมการใช้ข้อสอบทั่วไป คือ วิธีการ Randomesque วิธีการ Sympson-Hetter (SH) และวิธีการที่ไม่มีการควบคุมการใช้ข้อสอบ ภายใต้โมเดลโลจิสติก 3 พารามิเตอร์ และเปรียบเทียบในกรณีคลังข้อสอบขนาดเล็กเทียบกับขนาดใหญ่ ผลจากการศึกษา ปรากฏว่า วิธีการ PR-SE สามารถเข้าควบคุมการใช้ข้อสอบของข้อสอบในคลังได้เกือบทุกข้อ แต่วิธีการ Randamesque กับวิธี SH สามารถเข้าควบคุมการใช้ข้อสอบของข้อสอบในคลังได้เพียง 52% และ 80% ในกรณีที่คลังข้อสอบมีขนาดใหญ่และเล็ก ตามลำดับ

Ozturk and Dogan (2015) ได้ศึกษาเปรียบเทียบประสิทธิภาพการประมาณค่าความสามารถของผู้สอบ และการควบคุมการใช้ข้อสอบ ระหว่างวิธีการควบคุมการใช้ข้อสอบ 3 วิธี ได้แก่ 1) วิธี Randomesque 2) วิธี Sympson-Hetter และ 3) วิธี Fade-Away โดยใช้คลังข้อสอบจำลอง ด้วยการสุ่มค่าอำนาจจำแนกของข้อสอบ (a) จากการแจกแจงแบบยูนิฟอร์มในช่วง 0.50 ถึง 2.00 สุ่มค่าการเดาของข้อสอบ (c) จากการแจกแจงแบบยูนิฟอร์มในช่วง 0.05 ถึง 0.20 ส่วนการจำลองค่าความยากของข้อสอบ (b) จะแบ่งได้ 2 ส่วน คือ 1) คลังข้อสอบที่มีค่าความยากของข้อสอบปานกลาง ด้วยการสุ่มค่า b จากการแจกแจงแบบยูนิฟอร์มในช่วง -3.00 ถึง 3.00 และ 2)

คลังข้อสอบที่มีค่าความยากของข้อสอบสูง ด้วยการสุ่มค่า b จากการแจกแจงแบบปกติ N(2, 1.5) ผลการศึกษาปรากฏว่า ประสิทธิภาพการประมาณค่าความสามารถของผู้สอบทั้ง 3 วิธี ไม่มีความแตกต่างกัน แต่ วิธี Fade-Away สามารถควบคุมการใช้ข้อสอบได้ดีกว่าวิธี Randomesque และวิธี Sympson-Hetter

จากการศึกษางานวิจัยที่เกี่ยวข้องกับวิธีการควบคุมการใช้ข้อสอบ ปรากฏว่า ในช่วงเริ่มแรกจะเป็นงานวิจัยที่เกี่ยวข้องกับการพัฒนาวิธีการขึ้นใหม่ รวมทั้งปรับปรุงประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบที่มีผู้พัฒนาขึ้นแล้ว ให้มีประสิทธิภาพเพิ่มมากขึ้น และในระยะต่อมางานวิจัยส่วนใหญ่จะเป็นการศึกษาในสถานการณ์จำลอง เพื่อเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบที่นำไปใช้ในสถานการณ์ต่าง ๆ เพื่อให้ทราบข้อดี หรือข้อจำกัดของวิธีการนั้นๆ

การทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องกับวิธีการคัดเลือกข้อสอบข้อถัดไป และวิธีการควบคุมการใช้ข้อสอบ ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่า ยังไม่พบงานวิจัยใดที่นำเกณฑ์ของเฮอร์วิคซ์มาประยุกต์กับวิธีการคัดเลือกข้อสอบข้อถัดไป และนำวิธีการสุ่มตัวอย่างแบบมีระบบมาประยุกต์กับวิธีการควบคุมการใช้ข้อสอบ ดังนั้น ผู้วิจัยจึงมีความสนใจที่จะพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ โดยจะพัฒนาให้วิธีการนี้มีประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบ ด้านความยาวของแบบทดสอบ และจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2

# บทที่ 3
# วิธีดำเนินการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จากนั้นนำวิธีการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้น ทั้งในกรณีที่มีและไม่มีการควบคุมการใช้ข้อสอบ มาเปรียบเทียบประสิทธิภาพกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง ซึ่งพิจารณาด้าน 1) การประมาณค่าความสามารถของผู้สอบ 2) ความยาวของแบบทดสอบ และ 3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จากนั้นพัฒนาโปรแกรมการทดสอบ ซึ่งสามารถรองรับทั้งการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และการทดสอบด้วยคอมพิวเตอร์ แล้วนำโปรแกรมการทดสอบที่พัฒนาได้มาศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ จึงแบ่งวิธีดำเนินการวิจัยได้ 4 ระยะ ได้แก่

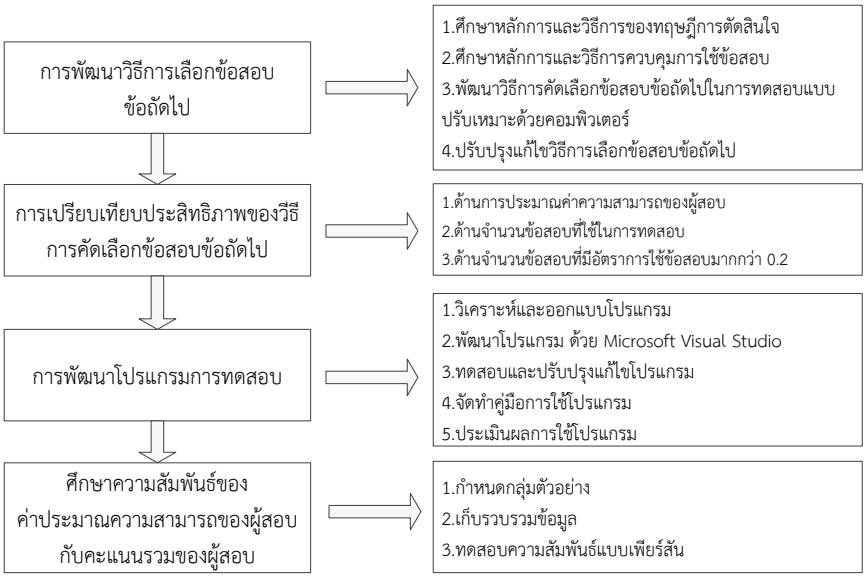
ระยะที่ 1 การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

ระยะที่ 2 การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป

ระยะที่ 3 การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ระยะที่ 4 การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์

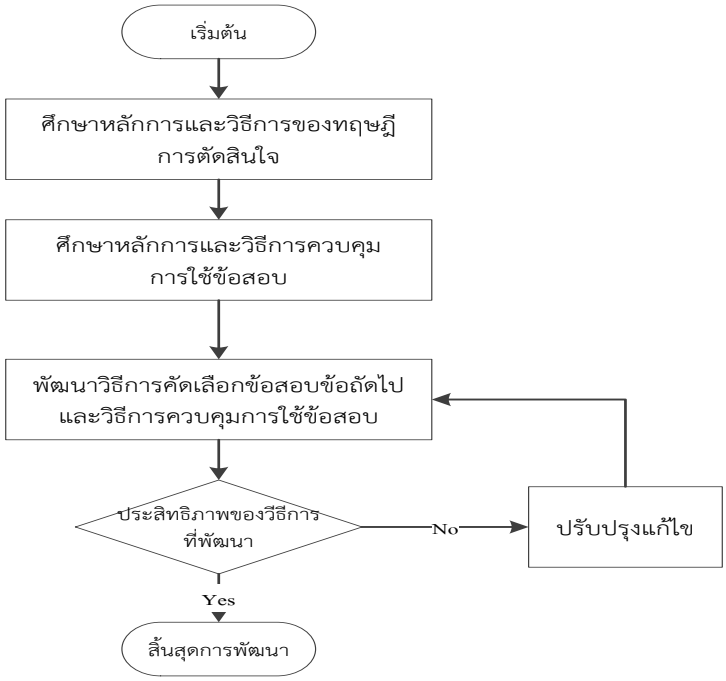
การวิจัยนี้สามารถแสดงขั้นตอนการดำเนินการวิจัยได้ดังภาพที่ 3-1




ภาพที่ 3-1 ขั้นตอนการดำเนินการวิจัย

**ระยะที่ 1 การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ**

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ตามเกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบโดยใช้วิธีการสุ่มตัวอย่างแบบมีระบบ ซึ่งเขียนแผนผังแสดงขั้นตอนในการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้ดังนี้




ภาพที่ 3-2 ขั้นตอนการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป

จากภาพที่ 3-2 สามารถอธิบายรายละเอียดในแต่ละขั้นตอนได้ดังนี้

**ขั้นตอนที่ 1 ศึกษาหลักการและวิธีการของทฤษฎีการตัดสินใจ** เป็นการศึกษาเอกสารและวรรณกรรมที่เกี่ยวข้องกับทฤษฎีการตัดสินใจ เพื่อเป็นฐานความรู้มาใช้กับวิธีการคัดเลือกข้อสอบข้อถัดไป จากการศึกษาทฤษฎีการตัดสินใจในสภาวการณ์ทั้ง 4 ประเภท ได้แก่ 1) การตัดสินใจในสภาวการณ์ที่แน่นอน 2) การตัดสินใจในสภาวการณ์ที่ไม่แน่นอน 3) การตัดสินใจในสภาวการณ์ความเสี่ยง และ 4) การตัดสินใจในสภาวการณ์การแข่งขัน ปรากฏว่า ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน (Uncertainly Decision Making) และในสภาวการณ์ความเสี่ยง (Risk Decision Making) สามารถนำมาประยุกต์กับวิธีการคัดเลือกข้อสอบข้อถัดไปได้ แต่ทั้งนี้ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน มีแนวคิดและการนำไปใช้ที่ง่ายกว่าเมื่อเทียบกับทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง เนื่องจากทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยงต้องอาศัยค่าความน่าจะเป็นในการคำนวณ จึงมีความยุ่งยากในการคำนวณมากกว่าทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ซึ่งใช้เพียงค่าเฉลี่ยหรือพิจารณาจากค่าผลตอบแทนสูงสุด-ต่ำสุด (Taylor,

2009, pp. 539-548) นอกจากนี้เกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion) เป็นเกณฑ์ที่มีความยืดหยุ่นในการนำไปใช้งานสูงกว่าเมื่อเทียบกับเกณฑ์อื่น ๆ เนื่องจากในการคำนวณตามเกณฑ์ของเฮอร์วิคซ์สามารถปรับเปลี่ยนน้ำหนักของผลตอบแทนในแต่ละทางเลือก (สัมประสิทธิ์ของการมองโลกในแง่ดี) ให้สอดคล้องกับดุลยพินิจของผู้ตัดสินใจได้ ซึ่งเกณฑ์อื่น ๆ ในทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอนไม่สามารถทำได้ (Taylor, 2009, p. 542) ด้วยเหตุนี้ ผู้วิจัยจึงสนใจนำทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ตามเกณฑ์ของเฮอร์วิคซ์ มาปรับใช้กับวิธีการคัดเลือกข้อสอบข้อถัดไป

**ขั้นตอนที่ 2 ศึกษาหลักการและวิธีการควบคุมการใช้ข้อสอบ** เป็นการศึกษาเอกสารและวรรณกรรมที่เกี่ยวข้องกับวิธีการควบคุมการใช้ข้อสอบ เพื่อเป็นฐานความรู้มาใช้กับวิธีการควบคุมการใช้ข้อสอบ ซึ่งจากการเปรียบเทียบข้อดีของวิธีการควบคุมการใช้ข้อสอบทั้ง 5 ประเภท ได้แก่ 1) วิธีการคัดเลือกอย่างสุ่ม 2) วิธีการคัดเลือกอย่างมีเงื่อนไข 3) วิธีการคัดเลือกตามระดับชั้น 4) วิธีการคัดเลือกแบบผสม และ 5) วิธีการคัดเลือกแบบปรับเหมาะหลายขั้นตอน ปรากฏว่า แนวคิดและการนำไปใช้ของวิธีการควบคุมการใช้ข้อสอบในประเภทวิธีการคัดเลือกอย่างสุ่ม (Randomization Strategies) มีความยุ่งยากซับซ้อนน้อยกว่าเมื่อเทียบกับวิธีการควบคุมการใช้ข้อสอบประเภทอื่น ๆ เพราะใช้เพียงการสุ่มข้อสอบ ทำให้ข้อสอบข้อใดข้อหนึ่งไม่ถูกเลือกใช้ในการทดสอบบ่อยครั้งจนเกินไป (Georgiadou et al., 2007) โดยอาศัยแนวคิดของการสุ่มข้อสอบนี้จึงนำวิธีการสุ่มตัวอย่างแบบมีระบบมาใช้กับวิธีการควบคุมการใช้ข้อสอบ

**ขั้นตอนที่ 3 พัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป และวิธีการควบคุมการใช้ข้อสอบ** โดยการนำผลการวิเคราะห์ที่ได้จากการศึกษาในขั้นตอนที่ 1 - 2 มาพัฒนาเป็นวิธีการคัดเลือกข้อสอบข้อถัดไป และวิธีการควบคุมการใช้ข้อสอบในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งในทางปฏิบัติวิธีการควบคุมการใช้ข้อสอบจะดำเนินการก่อนการคัดเลือกข้อสอบข้อถัดไป จึงแสดงรายละเอียดในการพัฒนาวิธีการควบคุมการใช้ข้อสอบ และวิธีการคัดเลือกข้อสอบข้อถัดไป ตามลำดับดังนี้

**3.1 การพัฒนาวิธีการควบคุมการใช้ข้อสอบ**

จากการศึกษาวิธีการคัดเลือกข้อสอบข้อถัดไป ปรากฏว่า วิธีการคัดเลือกข้อสอบข้อถัดไป ส่วนใหญ่จะเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงมาเป็นข้อสอบข้อถัดไป เนื่องจากข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง เป็นข้อสอบที่ให้ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบต่ำ ทำให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงถูกเลือกใช้บ่อยเกินไป ซึ่งจะส่งผลต่อความปลอดภัยของการทดสอบ ด้วยเหตุนี้ผู้วิจัยจึงพัฒนาวิธีการควบคุมการใช้ข้อสอบ โดยใช้แนวคิดของวิธีการควบคุมการใช้ข้อสอบในกลุ่มวิธีการคัดเลือกอย่างสุ่ม (Randomization Strategies) มาเป็นฐานในการพัฒนา ซึ่งจากการศึกษาเห็นว่าวิธีการสุ่มตัวอย่างแบบมีระบบ (Systematic Random Sampling) สามารถใช้เพื่อควบคุมการใช้ข้อสอบแต่ละข้อให้มีความเหมาะสมได้ โดยจะสุ่มข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงบางข้อเท่านั้น ที่นำไปใช้ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ซึ่งจะสุ่มด้วยวิธีการสุ่มแบบมีระบบและจะสุ่มข้อสอบใหม่ทุกครั้งก่อนถึงขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ทำให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ถูกหมุนเวียนนำไปใช้ในการคัดเลือกข้อสอบข้อถัดไป เป็นผลให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบ

สูงข้อใดข้อหนึ่ง ถูกเลือกใช้ลดลง ขั้นตอนในการควบคุมการใช้ข้อสอบด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ มีดังนี้

ขั้นที่ 1 แบ่งข้อสอบในคลังข้อสอบ (หรือข้อสอบในหัวข้อเรื่องถัดจากข้อสอบที่ทำขณะนั้น) ออกเป็น 2 กลุ่ม คือ ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบต่ำถึงปานกลาง ซึ่งมีค่าน้อยกว่า 1.35 (Baker, 2001, p. 34) กำหนดให้เป็นกลุ่ม A และข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ซึ่งมีค่าตั้งแต่ 1.35 ขึ้นไป กำหนดให้เป็นกลุ่ม B

ขั้นที่ 2 สุ่มเลือกข้อสอบในกลุ่ม B ด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ ซึ่งมีขั้นตอนการสุ่มดังนี้

ขั้นที่ 2.1 กำหนดจำนวนข้อสอบที่ต้องการสุ่มในกลุ่ม B สมมติว่าต้องการสุ่มข้อสอบจำนวน n ข้อ โดยที่ n เป็นจำนวนเต็มที่สุ่มจากช่วงจำนวนตั้งแต่ 5 ถึง 10 ข้อ

การกำหนดช่วงจำนวนข้อสอบสุ่มตั้งแต่ 5 ถึง 10 ข้อนี้ อาศัยแนวคิดจากงานของ McBride & Martin (1983) ที่เสนอวิธี 5-4-3-2-1 ซึ่งเป็นวิธีการควบคุมการใช้ข้อสอบ โดยข้อสอบข้อถัดไปเป็นข้อสอบที่สุ่มจากข้อสอบที่มีค่าสารสนเทศของข้อสอบสูงสุดมา 5 ข้อ และจากงานของ Kingsbury & Zara (1989) ที่เสนอวิธี Randomesque strategy ซึ่งเป็นวิธีการควบคุมการใช้ข้อสอบ โดยข้อสอบข้อถัดไปจะเป็นข้อสอบที่สุ่มจากข้อสอบที่มีค่าสารสนเทศของข้อสอบสูงสุดไม่เกิน 10 ข้อ แสดงให้เห็นว่า วิธีการควบคุมการใช้ข้อสอบจะพิจารณาจากข้อสอบประมาณ 5-10 ข้อเท่านั้น ดังนั้น ผู้วิจัยจึงกำหนดจำนวนข้อสอบสุ่ม 5 - 10 ข้อ

ขั้นที่ 2.2 หาช่วงของการสุ่ม (k) โดยคำนวณจาก $k=\frac{N}{n}$ เมื่อ N คือ จำนวนข้อสอบทั้งหมดในกลุ่ม B และ n คือ ข้อสอบที่ต้องสุ่ม

ขั้นที่ 2.3 หาข้อสอบที่ใช้เป็นตัวสุ่มเริ่มต้น (r) โดยสุ่มค่า r จากค่าตั้งแต่ 1 ถึง k จากนั้นจึงเลือกข้อสอบโดยเริ่มตั้งแต่ข้อที่ r, r+k, r+2k ตามลำดับจนครบ n ข้อ

ขั้นที่ 3 รวมข้อสอบในกลุ่ม A กับข้อสอบที่สุ่มได้ในขั้นที่ 2 เข้าด้วยกัน ซึ่งข้อสอบทั้งหมดในขั้นนี้ จะนำไปใช้ในกระบวนการคัดเลือกข้อสอบข้อถัดไป

ตัวอย่างแสดงขั้นตอนการควบคุมการใช้ข้อสอบโดยใช้วิธีการสุ่มตัวอย่างแบบมีระบบ

กำหนดให้ข้อสอบทั้งหมดในหัวข้อเรื่องถัดไปมีจำนวน 100 ข้อ แสดงได้ดังภาพที่ 3-3 (ก) ขั้นตอนในการควบคุมการใช้ข้อสอบมีรายละเอียดดังนี้

ขั้นที่ 1 แบ่งข้อสอบในหัวข้อเรื่องถัดไปออกเป็น 2 กลุ่ม คือ กลุ่ม A เป็นข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบ (a) ต่ำถึงปานกลาง มีจำนวน 50 ข้อ และกลุ่ม B เป็นข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง มีจำนวน 50 ข้อ แสดงดังภาพที่ 3-3 (ข)

ขั้นที่ 2 สุ่มเลือกข้อสอบในกลุ่ม B ด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ มีดังนี้

ขั้นที่ 2.1 กำหนดจำนวนข้อสอบที่ต้องการสุ่มในกลุ่ม B เป็นจำนวน 8 ข้อ (กำหนดจากการสุ่มตัวเลขในช่วง 5 ถึง 10)

ขั้นที่ 2.2 หาช่วงของการสุ่ม $k = \frac{50}{8} = 6.25$ ตัดเศษได้ k = 6

ขั้นที่ 2.3 กำหนดตัวสุ่มเริ่มต้น r = 2 (กำหนดจากการสุ่มตัวเลขในช่วง 1 ถึง k = 6) ดังนั้น ข้อสอบที่ถูกเลือก คือ ข้อสอบลำดับที่ 2, ลำดับที่ 8, ลำดับที่ 14 จนถึงข้อสอบลำดับที่ 44 ซึ่งมีข้อสอบทั้งหมด 8 ข้อ แสดงได้จากข้อสอบที่ขีดเส้นใต้ ในภาพที่ 3-3 (ค)

ขั้นที่ 3 รวมข้อสอบในกลุ่ม A กับข้อสอบที่สุ่มได้ในขั้นที่ 2 เข้าด้วยกัน ซึ่งจะเป็นข้อสอบทั้งหมดที่ใช้ในกระบวนการคัดเลือกข้อสอบข้อถัดไป แสดงได้ในภาพที่ 3-3 (ง)

| ข้อที่ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | ... | 100 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | 1.07 | 0.51 | 0.90 | 1.70 | 1.65 | 2.15 | 1.45 | 2.40 | 0.70 | 1.22 | 0.91 | 1.59 | ... | 2.05 |

(ก)

**ข้อสอบกลุ่ม A**

| ข้อที่ | 2 | 30 | 54 | 32 | 60 | 14 | 55 | ... | 77 |
|---|---|---|---|---|---|---|---|---|---|
| a | 0.51 | 0.52 | 0.54 | 0.55 | 0.57 | 0.58 | 0.62 | ... | 1.34 |

ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบต่ำถึงปานกลาง (a < 1.35)

**ข้อสอบกลุ่ม B**

| ข้อที่ | 73 | 7 | 69 | 61 | 87 | 19 | 39 | ... | 91 |
|---|---|---|---|---|---|---|---|---|---|
| a | 1.44 | 1.45 | 1.45 | 1.48 | 1.52 | 1.53 | 1.53 | ... | 2.49 |

ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง (a ≥ 1.35)

(ข)

**ข้อสอบกลุ่ม B**

| ลำดับ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | ... | 49 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ข้อที่ | 73 | 7 | 69 | 61 | 87 | 19 | 39 | 52 | 12 | 98 | 5 | 80 | 33 | 4 | 57 | 56 | 17 | 63 | 72 | 64 | 51 | ... | 91 |
| a | 1.44 | 1.45 | 1.45 | 1.48 | 1.52 | 1.53 | 1.53 | 1.57 | 1.59 | 1.61 | 1.65 | 1.67 | 1.70 | 1.70 | 1.72 | 1.72 | 1.73 | 1.78 | 1.79 | 1.82 | 1.83 | ... | 2.49 |

(ค)

**ข้อสอบกลุ่ม A**

| ข้อที่ | 2 | 30 | 54 | 32 | 60 | 14 | 55 | ... | 77 |
|---|---|---|---|---|---|---|---|---|---|
| a | 0.51 | 0.52 | 0.54 | 0.55 | 0.57 | 0.58 | 0.62 | ... | 1.34 |

**ข้อสอบที่สุ่มได้จากข้อสอบกลุ่ม B**

| 7 | 52 | 4 | 64 | 97 | 6 | 49 | 8 |
|---|---|---|---|---|---|---|---|
| 1.45 | 1.57 | 1.70 | 1.82 | 1.94 | 2.15 | 2.26 | 2.40 |

(ง)

ภาพที่ 3-3 ตัวอย่างขั้นตอนการควบคุมการใช้ข้อสอบโดยใช้วิธีการสุ่มตัวอย่างแบบมีระบบ

**3.2 การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป**

จากการศึกษาทฤษฎีการตัดสินใจ ปรากฏว่า ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน (Uncertainty Decision Making) ตามเกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion) มีความเหมาะสมที่จะนำมาประยุกต์เป็นวิธีการคัดเลือกข้อสอบข้อถัดไปโดยเกณฑ์ของเฮอร์วิคซ์จะตัดสินใจเลือกทางเลือก (Alternative Acts) ที่ให้ค่าเฉลี่ยถ่วงน้ำหนักของผลตอบแทนสูงสุด แต่เมื่อนำมาประยุกต์กับวิธีการคัดเลือกข้อสอบข้อถัดไป เป็นการเลือกข้อสอบที่ให้ค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถต่ำสุด มาเป็นข้อสอบข้อถัดไป โดยข้อสอบที่จะนำมาหาค่าเฉลี่ยถ่วงน้ำหนักนี้ ต้องเป็นข้อสอบที่ไม่เคยใช้สอบมาก่อน และเป็นข้อสอบที่ถูกคัดเลือกจากวิธีการควบคุมการใช้ข้อสอบในขั้นตอน 3.1 (ข้างต้น) สูตรคำนวณค่าเฉลี่ยถ่วงน้ำหนักค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ H(i) ตามเกณฑ์ของเฮอร์วิคซ์ มีดังนี้

$$H(i) = \alpha \times \min(SEE_{true}(\hat{\theta}), SEE_{false}(\hat{\theta})) + (1-\alpha) \times \max(SEE_{true}(\hat{\theta}), SEE_{false}(\hat{\theta})) \quad (35)$$

โดยที่ $H(i)$ คือ ค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ณ ข้อสอบข้อที่ i

$SEE_{true}(\hat{\theta})$ คือ ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถในกรณีที่ผู้สอบตอบถูก

$SEE_{false}(\hat{\theta})$ คือ ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ในกรณีที่ผู้สอบตอบผิด

min คือ ค่าต่ำสุดที่เลือกจากค่า $SEE_{true}(\hat{\theta})$ และค่า $SEE_{false}(\hat{\theta})$ ขึ้นอยู่กับค่าใดต่ำกว่ากัน

max คือ ค่าสูงสุดที่เลือกจากค่า $SEE_{true}(\hat{\theta})$ และค่า $SEE_{false}(\hat{\theta})$ ขึ้นอยู่กับค่าใดสูงกว่ากัน

$\hat{\theta}$ คือ ค่าประมาณความสามารถของผู้สอบ

$\alpha$ คือ สัมประสิทธิ์ของการมองโลกในแง่ดี มีเท่ากับ $1-\frac{1}{5.5}|\hat{\theta} - b|$

b คือ ค่าความยากของข้อสอบ

หมายเหตุ สัมประสิทธิ์ของการมองโลกในแง่ดี ($\alpha$) หรือค่าถ่วงน้ำหนักของผลตอบแทนสูงสุด เป็นค่าที่ผู้ตัดสินใจใช้ดุลยพินิจกำหนดขึ้นเองโดยอาศัยข้อมูลหรือประสบการณ์ที่ผ่านมา ซึ่งมีค่าตั้งแต่ 0 ถึง 1 (กิตติ ภักดีวัฒนะกุล และพนิดา พานิชกุล, 2554, หน้า 541)

ในที่นี้กำหนดค่า $\alpha$ จากที่ทราบว่า ข้อสอบที่มีคุณภาพสูงเป็นข้อสอบที่มีค่าความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้สอบ โดยค่าความยากของข้อสอบในทางปฏิบัติจะมีค่าอยู่ระหว่าง -2.5 ถึง 2.5 ส่วนค่าความสามารถของผู้สอบในทางปฏิบัติจะมีค่าอยู่ระหว่าง -3 ถึง 3 ซึ่งค่าความยากของข้อสอบและค่าความสามารถของผู้สอบถูกวัดในสเกลเดียวกัน ดังนั้น ค่าความยากของข้อสอบและค่าความสามารถของผู้สอบจึงมีค่าแตกต่างกันต่ำสุด 0 หน่วย (คือ ค่าความยากของข้อสอบมีค่าเท่ากับค่าความสามารถของผู้สอบ) และมีค่าแตกต่างกันสูงสุด 5.5 หน่วย

(คือ ค่าความยากของข้อสอบมีค่าต่ำสุด -2.5 และค่าความสามารถของผู้สอบมีค่าสูงสุด 3 หรือค่าความยากของข้อสอบมีค่าสูงสุด 2.5 และค่าความสามารถของผู้สอบมีค่าต่ำสุด -3) และโดยการเทียบสัดส่วนตามหลักคณิตศาสตร์ จึงกำหนดให้ $\alpha$ มีค่าเท่ากับ 1 - $\frac{1}{5.5}|\hat{\theta} - b|$

ตัวอย่างการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ ในวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ ซึ่งกำหนดค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น $\hat{\theta}$ = 1.0 และจากข้อสอบทั้งหมดที่ผ่านวิธีการควบคุมการใช้ข้อสอบ (แสดงตัวอย่างใน 3.1 การพัฒนาวิธีการควบคุมการใช้ข้อสอบ ตามภาพที่ 3-3 (ง)) การคำนวณค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ ตามสมการที่ 35 แสดงได้ดังตารางที่ 3-1

ตารางที่ 3-1 ตัวอย่างการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | SEE | | H |
|---|---|---|---|---|---|---|
| | a | b | c | กรณีผู้สอบตอบถูก | กรณีผู้สอบตอบผิด | |
| 2 | 0.51 | 2.08 | 0.10 | 2.75 | 3.24 | 2.85 |
| 30 | 0.52 | -0.04 | 0.15 | 2.90 | 2.65 | 2.69 |
| 54 | 0.54 | 1.33 | 0.06 | 2.30 | 2.46 | 2.31 |
| : | : | : | : | : | : | : |
| 77 | 1.34 | -0.12 | 0.18 | 2.32 | 1.06 | 1.32 |
| 7 | 1.45 | -1.49 | 0.11 | 9.63 | 1.05 | 4.93 |
| 52 | 1.57 | -1.25 | 0.09 | 8.51 | 0.87 | 4.00 |
| 4* | 1.70 | 1.35 | 0.06 | 0.83 | 1.85 | 0.90 |
| 64 | 1.82 | -0.79 | 0.08 | 6.28 | 0.70 | 2.51 |
| 97 | 1.94 | 1.39 | 0.19 | 0.74 | 3.50 | 0.93 |
| 6 | 2.15 | 1.57 | 0.25 | 0.73 | 6.95 | 1.36 |
| 49 | 2.26 | -1.62 | 0.07 | 43.45 | 0.57 | 21.00 |
| 8 | 2.40 | 0.32 | 0.06 | 2.23 | 0.87 | 1.03 |

หมายเหตุ a หมายถึง ค่าอำนาจจำแนกของข้อสอบ
b หมายถึง ค่าความยากของข้อสอบ
c หมายถึง ค่าการเดาของข้อสอบ
SEE หมายถึง ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ
H หมายถึง ค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE
* หมายถึง ข้อสอบที่มีค่า H ต่ำสุด

จากตารางที่ 3-1 แสดงตัวอย่างการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ ซึ่งเริ่มจากการคำนวณค่าประมาณความสามารถของผู้สอบ ทั้งในกรณีที่ผู้สอบตอบถูก และกรณีที่ผู้สอบตอบผิด จากสมการที่ 12 ถึงสมการที่ 17 หลังจากนั้นนำค่าประมาณความสามารถของผู้สอบ มาคำนวณค่าสารสนเทศของแบบทดสอบ และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ จากสมการที่ 29 และสมการที่ 30 ตามลำดับ ซึ่งจะคำนวณค่า SEE ทั้งกรณีที่ผู้สอบตอบถูก และกรณีที่ผู้สอบตอบผิด แล้วนำไปใช้คำนวณค่าเฉลี่ยถ่วงน้ำหนัก (H) ตามสมการที่ 35 ซึ่งจากการพิจารณาค่า H ปรากฏว่า ข้อสอบที่มีค่า H ต่ำสุด คือ ข้อที่ 4 ดังนั้น ข้อสอบข้อถัดไปที่ผู้สอบจะได้รับ คือ ข้อสอบข้อที่ 4

**ขั้นตอนที่ 4 ปรับปรุงและแก้ไขวิธีการคัดเลือกข้อสอบข้อถัดไปและวิธีการควบคุมการใช้ข้อสอบ** เป็นการนำวิธีการคัดเลือกข้อสอบข้อถัดไปและวิธีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้นไปทดสอบความถูกต้อง หรือนำไปพัฒนาวิธีการให้มีประสิทธิภาพมากยิ่งขึ้น โดยการนำไปทดลองใช้ในสถานการณ์จำลอง แล้วปรับปรุงจนกว่าจะได้วิธีการที่มีความเหมาะสม ซึ่งมีรายละเอียดดังนี้

**4.1 การปรับปรุงและแก้ไขข้อผิดพลาดที่เกิดจากการเขียนโปรแกรม**

ข้อผิดพลาดที่พบจากการเขียนโปรแกรมมี 2 ลักษณะ ได้แก่ 1) ข้อผิดพลาดจากการเขียนคำสั่งของโปรแกรมผิดรูปแบบไวยกรณ์ของภาษา สามารถตรวจสอบได้ง่ายขณะประมวลผล โดยโปรแกรมจะรายงานข้อผิดพลาดที่เกิดขึ้น ซึ่งส่วนใหญ่เกิดขึ้นจากการพิมพ์ผิด แก้ไขโดยการพิมพ์ให้ถูกต้อง และ 2) ข้อผิดพลาดจากการใช้คำสั่งผิดตรรกะหรือผิดเงื่อนไข ซึ่งทำให้ผลลัพธ์จากการกระทำตามเงื่อนไขนั้นเกิดความผิดพลาด เช่น การตรวจสอบเงื่อนไข ใช้คำว่า “or” กับ “and” สลับกัน ซึ่งข้อผิดพลาดลักษณะนี้ตรวจสอบโดยพิจารณาค่าของตัวแปรที่ได้รับ หรือพิจารณาผลลัพธ์จากการประมวลผลโปรแกรมที่ไม่เป็นไปตามที่คาดไว้ หากพบจะตรวจสอบผลลัพธ์จากการประมวลผลทีละคำสั่ง แล้วแก้ไขให้ถูกต้อง

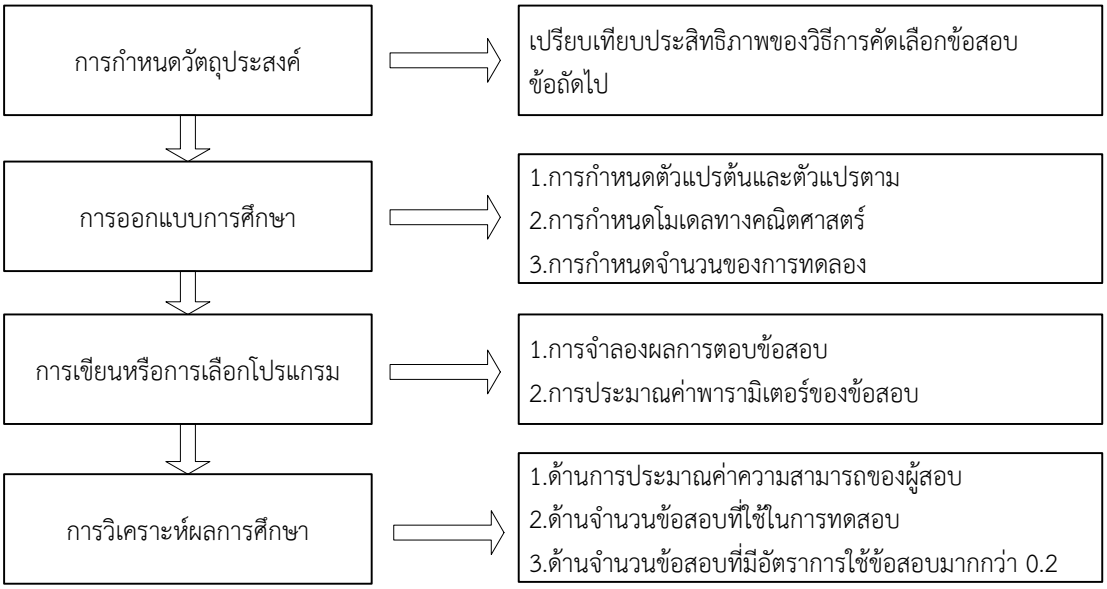
**4.2 การปรับปรุงและแก้ไขวิธีการให้มีประสิทธิภาพมากขึ้น**

การปรับปรุงและแก้ไขวิธีการให้มีประสิทธิภาพมากขึ้น เป็นการนำแนวคิดจากหลายส่วน เช่น ทฤษฎีการตอบสนองข้อสอบ ทฤษฎีการตัดสินใจ วิธีการคัดเลือกข้อสอบข้อถัดไป วิธีการควบคุมการใช้ข้อสอบ หรือวิธีการสุ่มตัวอย่างในระเบียบวิธีการวิจัย มาปรับปรุงวิธีการคัดเลือกข้อสอบข้อถัดไปและวิธีการควบคุมการใช้ข้อสอบ จนกระทั่งได้วิธีการที่มีประสิทธิภาพเป็นที่น่าพอใจตามรายละเอียดที่ให้ไว้ในขั้นตอนที่ 3 เช่น การกำหนดค่าถ่วงน้ำหนักหรือค่าสัมประสิทธิ์ของการมองโลกในแง่ดี ในช่วงเริ่มต้นการพัฒนา ได้กำหนดให้ใช้ค่าอำนาจจำแนกของข้อสอบ แต่เมื่อทดลองใช้ในสถานการณ์จำลอง ปรากฏว่า ผลการประมาณค่าความสามารถของผู้สอบมีความคลาดเคลื่อนจากค่าความสามารถที่แท้จริงของผู้สอบมาก จึงปรับปรุงการกำหนดค่าสัมประสิทธิ์ของการมองโลกในแง่ดีใหม่ โดยกำหนดจากค่าผลต่างระหว่างค่าความยากของข้อสอบกับค่าความสามารถของผู้สอบ แล้วทดลองใช้ในสถานการณ์จำลองอีกครั้ง ปรากฏว่า การประมาณค่าความสามารถของผู้สอบมีความคลาดเคลื่อนลดลงจากค่าความสามารถที่แท้จริงของผู้สอบเป็นอย่างมาก

## ระยะที่ 2 การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป

การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปในการวิจัยนี้ ใช้การศึกษาในสถานการณ์จำลอง เพื่อให้แน่ใจว่าความแตกต่างของประสิทธิภาพที่เกิดขึ้น (ตัวแปรตาม) เป็นผลมาจากความแตกต่างของวิธีการคัดเลือกข้อสอบข้อถัดไป (ตัวแปรต้น) ไม่ได้เกิดจากตัวแปรแทรกซ้อนอื่น ๆ เช่น อารมณ์ของผู้สอบ หรือสภาพแวดล้อมขณะทดสอบ ซึ่งการควบคุมตัวแปรแทรกซ้อนให้คงที่ตลอดการศึกษาในสถานการณ์จริงเป็นไปได้ยากมาก นอกจากนี้ การวัดประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบนั้น จำเป็นต้องทราบค่าความสามารถที่แท้จริงของผู้สอบ ซึ่งในสถานการณ์จริงไม่สามารถหาค่าได้

การศึกษาในสถานการณ์จำลอง ดำเนินการตามขั้นตอนที่เสนอโดย Harwell et al. (1996) ซึ่งสามารถเขียนแผนผังแสดงขั้นตอนการศึกษาได้ดังนี้




ภาพที่ 3-4 ขั้นตอนการศึกษาในสถานการณ์จำลองเพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป

แต่ละขั้นตอนการศึกษาในสถานการณ์จำลองมีรายละเอียดดังนี้

**ขั้นตอนที่ 1 การกำหนดวัตถุประสงค์** ของการศึกษาในสถานการณ์จำลอง เป็นการศึกษาเพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ระหว่าง 4 วิธีการ ได้แก่ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง 3) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ และ 4) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ ซึ่งจะเปรียบเทียบประสิทธิภาพด้าน 1) การประมาณค่าความสามารถของผู้สอบ 2) ความยาวของแบบทดสอบ และ 3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2

**ขั้นตอนที่ 2 การออกแบบการศึกษา** ดำเนินการดังนี้

**2.1 กำหนดตัวแปรต้นและตัวแปรตาม** จากวัตถุประสงค์ของการศึกษาในสถานการณ์จำลอง

ตัวแปรต้น คือ วิธีการคัดเลือกข้อสอบข้อถัดไป มี 4 วิธีการ ได้แก่

1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด

2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง

3) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์

4) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

ตัวแปรตาม คือ ประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ซึ่งพิจารณาด้าน

1) การประมาณค่าความสามารถของผู้สอบ

1.1) ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

1.2) ค่าความลำเอียงเฉลี่ย (Average Bias)

2) ความยาวของแบบทดสอบ (Test Length)

3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2

**2.2 การกำหนดโมเดลทางคณิตศาสตร์** ของการศึกษาในสถานการณ์จำลองจะใช้ทฤษฎีการตอบสนองข้อสอบ โดยเลือกโมเดลโลจิสติก 3 พารามิเตอร์ มาใช้เป็นพื้นฐาน สำหรับจำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งมีขั้นตอนในการทดสอบดังนี้

**ขั้นที่ 1 การคัดเลือกข้อสอบข้อแรก** โดยการสุ่มเลือกข้อสอบในคลังข้อสอบที่มีค่าความยากของข้อสอบตั้งแต่ -1.00 ถึง 1.00

**ขั้นที่ 2 การคัดเลือกข้อสอบข้อถัดไป** จะเลือกใช้วิธีการใดวิธีการหนึ่งใน 4 วิธีการข้างต้น ขึ้นอยู่กับการจำลองสถานการณ์ของการทดสอบขณะนั้นว่า กำลังวัดประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปวิธีการใด ซึ่งแต่ละวิธีการผู้วิจัยได้เขียนโปรแกรมขึ้นเองตามรายละเอียดของแต่ละวิธีการดังนี้

วิธีการที่ 1 วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) วิธีนี้จะเลือกข้อสอบในคลังข้อสอบที่ยังไม่ถูกเลือกใช้สอบ และเป็นข้อสอบที่ให้ค่าสารสนเทศสูงสุด ณ ระดับความสามารถที่ตรงหรือใกล้เคียงกับค่าความสามารถของผู้สอบขณะนั้น มาเป็นข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดคำนวณได้จาก

$$m_i = b_i + \frac{1}{Da_i} \ln\left[\frac{1+\sqrt{1+8c_i}}{2}\right] \qquad (36)$$

เมื่อ $m_i$ คือ สารสนเทศสูงสุดของข้อสอบข้อที่ i

D คือ ค่าคงที่มีค่าเท่ากับ 1.7

$a_i$ คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i
$b_i$ คือ ค่าความยากของข้อสอบข้อที่ i
$c_i$ คือ ค่าการเดาของข้อสอบข้อที่ i

ตัวอย่างแสดงวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด โดยกำหนดค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น $\hat{\theta}$ = 1.0 พิจารณาข้อสอบที่ใช้ในการคัดเลือกข้อสอบข้อถัดไป จำนวน 10 ข้อ ซึ่งประกอบด้วยค่าพารามิเตอร์ของข้อสอบ และสามารถคำนวณค่าสารสนเทศสูงสุด ตามสมการที่ 36 ได้ดังตารางที่ 3-2

ตารางที่ 3-2 ตัวอย่างการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | $m_i$ |
|---|---|---|---|---|
| | a | b | c | |
| 1 | 1.07 | 1.34 | 0.09 | 1.41 |
| 2 | 0.51 | 2.08 | 0.10 | 2.26 |
| 3 | 0.90 | 1.75 | 0.09 | 1.84 |
| 4 | 1.70 | 1.35 | 0.06 | 1.39 |
| 5 | 1.65 | -2.20 | 0.14 | -2.13 |
| 6 | 2.15 | 1.57 | 0.25 | 1.65 |
| 7 | 1.45 | -1.49 | 0.11 | -1.42 |
| 8 | 2.40 | 0.32 | 0.06 | 0.34 |
| 9 | 0.70 | -1.86 | 0.26 | -1.59 |
| 10* | 1.22 | 1.03 | 0.14 | 1.13 |

หมายเหตุ a หมายถึง ค่าอำนาจจำแนกของข้อสอบ
b หมายถึง ค่าความยากของข้อสอบ
c หมายถึง ค่าการเดาของข้อสอบ
$m_i$ หมายถึง ค่าสารสนเทศสูงสุดของข้อสอบข้อที่ i
* หมายถึง ข้อสอบที่มีค่า $m_i$ ใกล้เคียงกับค่า $\hat{\theta}$ มากที่สุด

จากตารางที่ 3-2 แสดงตัวอย่างการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด เมื่อพิจารณาค่า $m_i$ ของข้อสอบแต่ละข้อ ปรากฏว่า ข้อที่ 10 เป็นข้อสอบที่มีค่า $m_i$ ใกล้เคียงกับค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น ($\hat{\theta}$ = 1.0) มากที่สุด ดังนั้น ข้อสอบข้อที่ 10 จึงถูกคัดเลือกเป็นข้อสอบข้อถัดไป

วิธีการที่ 2 วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (Decision Making Under Risk) ของโสพส สุขานนท์สวัสดิ์ และคณะ (2556) ซึ่งการดำเนินการจะเริ่มจากหาค่าคาดหวังสูงสุด (EMV) ของข้อสอบในคลังข้อสอบเฉพาะข้อที่ยังไม่ถูกเลือกใช้สอบ ซึ่งการคำนวณค่าความคาดหวังสูงสุดได้จากสมการ ดังนี้

$$EMV_i(\theta) = (P_i(\theta) \times \theta_{true}) + (Q_i(\theta) \times \theta_{false}) \quad (37)$$

เมื่อ $EMV_i(\theta)$ คือ ค่าคาดหวังสูงสุดของข้อสอบข้อที่ i
$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อที่ i ถูก
$Q_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อที่ i ผิด
$\theta_{true}$ คือ ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ถูก
$\theta_{false}$ คือ ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ผิด

หลังจากหาค่า EMV จากสมการที่ 37 แล้ว ให้หาค่าการกระจาย คำนวณจาก

$$\sigma_{i(EMV)}(\theta) = \sqrt{((\theta_{true} - EMV_i(\theta))^2 \times P_i(\theta)) + ((\theta_{false} - EMV_i(\theta))^2 \times Q_i(\theta))} \quad (38)$$

เมื่อ $\sigma_{i(EMV)}(\theta)$ คือ ค่าการกระจายของข้อสอบข้อที่ i
$EMV_i(\theta)$ คือ ค่าคาดหวังสูงสุดของข้อสอบข้อที่ i
$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อที่ i ถูก
$Q_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบข้อที่ i ผิด
$\theta_{true}$ คือ ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ถูก
$\theta_{false}$ คือ ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ผิด

หลังจากหาค่าการกระจายจากสมการที่ 38 แล้ว ให้หาค่าสัมประสิทธิ์การกระจาย คำนวณได้ดังนี้

$$CV_i(\theta) = \left| \frac{\sigma_{i(EMV)}(\theta)}{EMV_i(\theta)} \right| \quad (39)$$

เมื่อ $CV_i(\theta)$ คือ ค่าสัมประสิทธิ์การกระจายของข้อสอบข้อที่ i
$\sigma_{i(EMV)}(\theta)$ คือ ค่าการกระจายข้อมูลของข้อสอบข้อที่ i
$EMV_i(\theta)$ คือ ค่าความคาดหวังสูงสุดของข้อสอบข้อที่ i

โดยจะเลือกข้อสอบที่ให้ค่าสัมประสิทธิ์การกระจายมากที่สุด มาเป็นข้อสอบข้อถัดไป

ตัวอย่างแสดงวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง โดยกำหนดค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น $\hat{\theta}$ = 1.0 พิจารณาข้อสอบที่ใช้ในการคัดเลือกข้อสอบข้อถัดไป จำนวน 10 ข้อ ซึ่งประกอบด้วยค่าพารามิเตอร์ของข้อสอบ และสามารถคำนวณค่าคาดหวังสูงสุด (สมการที่ 37) ค่าการกระจาย (สมการที่ 38) และค่าสัมประสิทธิ์การกระจาย (สมการที่ 39) ได้ดังตารางที่ 3-3

ตารางที่ 3-3 ตัวอย่างการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | EMV | $\sigma_{EMV}$ | CV |
|---|---|---|---|---|---|---|
| | a | b | c | | | |
| 1 | 1.07 | 1.34 | 0.09 | 0.95 | 0.51 | 0.54 |
| 2 | 0.51 | 2.08 | 0.10 | 0.98 | 0.29 | 0.30 |
| 3 | 0.90 | 1.75 | 0.09 | 0.94 | 0.43 | 0.45 |
| 4* | 1.70 | 1.35 | 0.06 | 0.86 | 0.59 | 0.68 |
| 5 | 1.65 | -2.20 | 0.14 | 1.01 | 0.03 | 0.03 |
| 6 | 2.15 | 1.57 | 0.25 | 0.86 | 0.45 | 0.52 |
| 7 | 1.45 | -1.49 | 0.11 | 1.03 | 0.09 | 0.09 |
| 8 | 2.40 | 0.32 | 0.06 | 1.30 | 0.35 | 0.27 |
| 9 | 0.70 | -1.86 | 0.26 | 1.02 | 0.19 | 0.18 |
| 10 | 1.22 | 1.03 | 0.14 | 0.99 | 0.53 | 0.54 |

หมายเหตุ a หมายถึง ค่าอำนาจจำแนกของข้อสอบ
b หมายถึง ค่าความยากของข้อสอบ
c หมายถึง ค่าการเดาของข้อสอบ
EMV หมายถึง ค่าคาดหวังสูงสุดของข้อสอบข้อที่ i
$\sigma_{EMV}$ หมายถึง ค่าการกระจาย
CV หมายถึง ค่าสมบูรณ์ของสัมประสิทธิ์การกระจาย
* หมายถึง ข้อสอบที่มีค่า |CV| มากที่สุด

จากตารางที่ 3-3 แสดงตัวอย่างการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง เมื่อพิจารณาค่า CV (โดยไม่คำนึงถึงเครื่องหมาย + และ -) ของข้อสอบแต่ละข้อ ปรากฏว่า ข้อที่ 4 เป็นข้อสอบที่มีค่า CV มากที่สุด ดังนั้น ข้อสอบข้อที่ 4 จึงถูกคัดเลือกเป็นข้อสอบข้อถัดไป

วิธีการที่ 3 วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion) ตามรายละเอียดที่ให้ไว้ในระยะที่ 1 ซึ่งวิธีการที่ 3 นี้ จะไม่มีการควบคุมการใช้ข้อสอบ

วิธีการที่ 4 วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ (Hurwicz Criterion with Item Exposure Control) ซึ่งวิธีการที่ 4 นี้ จะนำวิธีการสุ่มตัวอย่างแบบมีระบบมาควบคุมการใช้ข้อสอบ ตามรายละเอียดที่ให้ไว้ในระยะที่ 1

**ขั้นที่ 3 การประมาณค่าความสามารถของผู้สอบ** เลือกใช้วิธีการประมาณค่าของเบส์แบบปรับใหม่ (Bayesian Updating) โดยการนำผลการตอบข้อสอบ และค่าพารามิเตอร์ของข้อสอบในทุกข้อที่ผ่านมา ได้แก่ ค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบ และค่าการเดาของข้อสอบ มาคำนวณค่าประมาณความสามารถของผู้สอบจากสมการที่ 12 -17

**ขั้นที่ 4 เกณฑ์ยุติการทดสอบ** ใช้เกณฑ์เมื่อค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบมีค่าต่ำกว่า 0.3 หรือทำข้อสอบครบ 30 ข้อ

**2.3 การกำหนดจำนวนของการทดลองซ้ำ** เทียบได้กับการกำหนดขนาดกลุ่มตัวอย่างของการศึกษาในสถานการณ์จริง แต่การศึกษาในสถานการณ์จำลองยังไม่มีกฎเกณฑ์แน่นอนในการกำหนดจำนวนของการทดลองซ้ำ แต่โดยทั่วไปการศึกษาในสถานการณ์จำลองที่ใช้ทฤษฎีการตอบสนองข้อสอบเป็นพื้นฐานจะกำหนดจำนวนมาก ๆ เช่น งานของ Leung et al. (2002) กำหนดจำนวนทดลองซ้ำ 10,000 ค่า งานของ Chang and Harris (2002) กำหนดจำนวนทดสอบซ้ำ 3,000 ค่า งานของ Chen and Doong (2003) และ Fan et al. (2012) กำหนดจำนวนทดลองซ้ำ 1,000 ค่า และงานของ Barrada, Mazuela, and Olea (2006) กำหนดจำนวนทดลองซ้ำ 5,000 ค่า ดังนั้น ในการวิจัยนี้ได้กำหนดจำนวนทดลองซ้ำ 10,000 ค่า ซึ่งการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปในแต่ละวิธี ได้ดำเนินการทำซ้ำ 10 รอบ ทุกรอบจะใช้คลังข้อสอบขนาด 500 ข้อ ชุดเดียวกัน แต่ใช้ค่าความสามารถที่แท้จริงของผู้สอบทีละชุด ๆ ละ 1,000 ค่า รวมทั้งหมด 10,000 ค่า ซึ่งค่าความสามารถที่แท้จริงของผู้สอบนี้ ถูกจำลองขึ้นโดยใช้โปรแกรม WinGen3 โดยกำหนดให้โปรแกรมสุ่มข้อมูลจากเลขสุ่มที่มีการแจกแจงแบบปกติ N (0, 1) จำนวน 10,000 ค่า แล้วแบ่ง 10 ชุด ๆ ละ 1,000 ค่า

**ขั้นตอนที่ 3 การเขียนหรือเลือกใช้โปรแกรม** การศึกษาในสถานการณ์จำลองนี้ เป็นการจำลองสถานการณ์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้โปรแกรม MATLAB ซึ่งผู้วิจัยได้เขียนคำสั่งให้โปรแกรมดำเนินการตามขั้นตอนในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เหมือนในสถานการณ์จริงทุกประการ (รายละเอียดแสดงในขั้นตอนที่ 2) แต่ข้อมูลที่ใช้ในการศึกษา ได้แก่ ค่าความสามารถที่แท้จริงของผู้สอบ ผลการตอบข้อสอบ และคลังข้อสอบ เป็นข้อมูลที่จำลองขึ้น โดยค่าความสามารถที่แท้จริงของผู้สอบจำลองขึ้นโดยใช้เลขสุ่ม (รายละเอียดแสดงใน 2.3) ส่วนผลการตอบข้อสอบและคลังข้อสอบ จำลองขึ้นโดยมีรายละเอียดดังนี้

**3.1 การจำลองผลการตอบข้อสอบ** เนื่องจากในสถานการณ์จำลองไม่มีผู้สอบที่แท้จริง ดังนั้น ในการเขียนโปรแกรมจำลองสถานการณ์การทดสอบ จึงต้องจำลองผลการตอบข้อสอบแต่ละข้อของผู้สอบ ซึ่งดำเนินการโดยให้โปรแกรมสุ่มเลขที่มีการแจกแจงแบบยูนิฟอร์ม U (0, 1) มา 1 ค่า แล้วนำเลขสุ่มนี้มาเปรียบเทียบกับค่าความน่าจะเป็นที่ผู้สอบ ที่มีความสามารถ $\theta$ ตอบข้อสอบข้อที่ i ได้ถูกต้อง ($P_i(\theta)$) ตามโมเดลโลจิสติกสามพารามิเตอร์ ซึ่งคำนวณได้จากสมการที่ 26 แล้วผลการตอบข้อสอบจะกำหนดได้จากเงื่อนไขดังนี้ (Thompson & Weiss, 2011)

กรณีที่ ค่า $P_i(\theta) \geq$ เลขสุ่ม ให้กำหนดผลการตอบข้อสอบเป็น 1 (ตอบถูก)

กรณีที่ ค่า $P_i(\theta) <$ เลขสุ่ม ให้กำหนดผลการตอบข้อสอบเป็น 0 (ตอบผิด)

**3.2 การประมาณค่าพารามิเตอร์ของข้อสอบ** เนื่องจากการศึกษาในสถานการณ์จำลอง ไม่มีคลังข้อสอบที่แท้จริง ดังนั้น จึงสร้างคลังข้อสอบจำลอง โดยข้อสอบแต่ละข้อจะมีเพียงค่าพารามิเตอร์ของข้อสอบ 3 ค่า ได้แก่ ค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบ และค่าการเดาของข้อสอบ ซึ่งจำลองขึ้นจากการสุ่มตัวเลขด้วยโปรแกรม WinGen3 โดยกำหนดให้โปรแกรมสุ่มเลขที่มีรายละเอียดดังนี้

ก. ค่าความยากของข้อสอบ (b-parameter) สุ่มจากการแจกแจงแบบยูนิฟอร์มมีค่าอยู่ในช่วง -2.50 ถึง 2.50

ข. ค่าอำนาจจำแนกของข้อสอบ (a-parameter) สุ่มจากการแจกแจงแบบยูนิฟอร์ม มีค่าอยู่ในช่วง 0.50 ถึง 2.50

ค. ค่าการเดาของข้อสอบ (c-parameter) สุ่มจากการแจกแจงแบบยูนิฟอร์ม มีค่าไม่เกิน 0.30

ผู้วิจัยกำหนดขนาดของคลังข้อสอบจำลอง 500 ข้อ ดังนั้น ต้องสุ่มเลขทั้งหมด 500 ข้อ แต่ละข้อประกอบด้วยเลขสุ่ม 3 ค่า ซึ่งใช้แทนค่าพารามิเตอร์ของข้อสอบ 3 ค่า โดยใช้โปรแกรม WinGen3 (คลังข้อสอบจำลอง แสดงในภาคภาคผนวก ค) ทั้งนี้ การกำหนดขนาดคลังข้อสอบ 500 ข้อนี้ กำหนดตามวิธีดำเนินการวิจัยของ Fan et al. (2012)

การศึกษาในสถานการณ์จำลอง การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นี้ กำหนดให้แต่ละวิธีการคัดเลือกข้อสอบข้อถัดไป ทำการทดลองซ้ำ 10 รอบ โดยทุกรอบใช้คลังข้อสอบจำลอง 500 ข้อชุดเดียวกัน แต่ใช้ค่าความสามารถที่แท้จริงของผู้สอบทีละชุด (ชุดละ 1,000 ค่า) เริ่มจากรอบที่ 1 ใช้ค่าความสามารถที่แท้จริงของผู้สอบชุดที่ 1 รอบที่ 2 ใช้ค่าความสามารถที่แท้จริงของผู้สอบชุดที่ 2 จนกระทั่งครบ 10 รอบ แล้วในแต่ละรอบได้เก็บข้อมูล ได้แก่ ค่าประมาณความสามารถของผู้สอบ ($\hat{\theta}$) ความยาวของแบบทดสอบ และอัตราการใช้ข้อสอบในคลังข้อสอบ เพื่อใช้เปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปทั้ง 4 วิธี

**ขั้นตอนที่ 4 การวิเคราะห์ผลการศึกษา** ดำเนินการโดยรวบรวมผลการศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปที่แตกต่างกัน 4 วิธี โดยทุกวิธีได้ดำเนินการซ้ำ 10 รอบ ซึ่งในแต่ละรอบจะได้ค่าประมาณความสามารถของผู้สอบ 1,000 ค่า ความยาวของแบบทดสอบ 1,000 ค่า และอัตราการใช้ข้อสอบในคลังข้อสอบ 500 ค่า แล้วนำข้อมูลทั้งหมดไปใช้เพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปในด้านต่อไปนี้

4.1 การประมาณค่าความสามารถของผู้สอบ ได้แก่

4.1.1 ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (ข้อมูลและการคำนวณค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย แสดงในภาคผนวก ข)

4.1.2 ค่าความลำเอียงเฉลี่ย (ข้อมูลและการคำนวณค่าความลำเอียงเฉลี่ย แสดงในภาคผนวก ข)

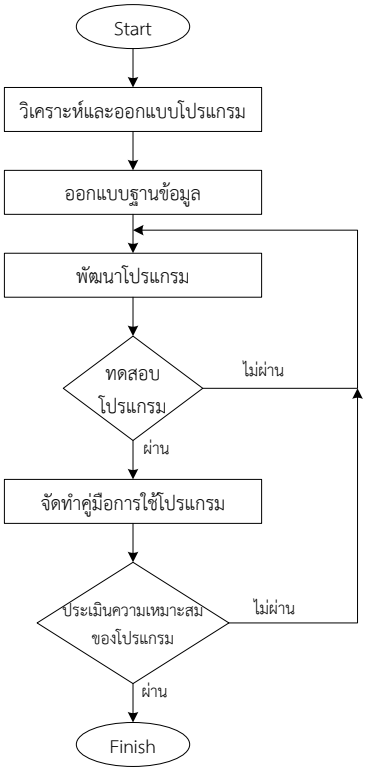
4.2 ความยาวของแบบทดสอบ (ข้อมูลแสดงในภาคผนวก ข)

4.3 จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 (ข้อมูลและการคำนวณค่าอัตราการใช้ข้อสอบ แสดงในภาคผนวก ค)

## ระยะที่ 3 การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น จะเป็นการทดสอบแบบทางแยกแปรผัน (Variable Branching Model) ซึ่งนำวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้นในการวิจัยระยะที่ 1 มาใช้ในสถานการณ์จริง โดยใช้กระบวนการพัฒนาซอฟต์แวร์แบบน้ำตก (Waterfall Model) นอกจากนี้ เพื่อให้สอดคล้องกับวัตถุประสงค์ของการวิจัยข้อที่ 4 คือ การศึกษาความสัมพันธ์ระหว่างค่าประมาณ

ความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ จึงออกแบบให้โปรแกรมที่จะพัฒนาขึ้นสามารถรองรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์และการทดสอบด้วยคอมพิวเตอร์ได้ โดยโปรแกรมการทดสอบนี้มีขั้นตอนการพัฒนา ดังภาพที่ 3-5



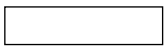
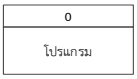
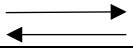

ภาพที่ 3-5 ขั้นตอนพัฒนาโปรแกรมการทดสอบ

ขั้นตอนในการพัฒนาโปรแกรมการทดสอบแต่ละขั้นตอน มีรายละเอียดดังต่อไปนี้

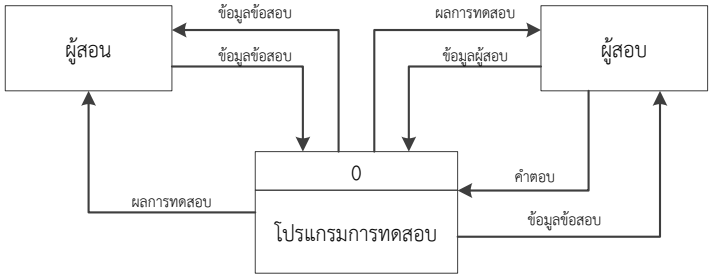
**1. การวิเคราะห์และออกแบบโปรแกรม** มีขั้นตอนย่อยดังนี้

**1.1 การวิเคราะห์แผนผังบริบท** (Context Diagram) เป็นการวิเคราะห์เพื่อหา Source Destination ที่เกี่ยวข้องกับโปรแกรมการทดสอบ โดยกำหนดสัญลักษณ์ที่ใช้ในการวิเคราะห์แผนผังบริบทดังนี้

ตารางที่ 3-4 สัญลักษณ์และความหมายที่ใช้ในการวิเคราะห์แผนผังบริบท

| สัญลักษณ์ | ความหมาย |
|---|---|
|  | ผู้ใช้ที่มีความเกี่ยวข้องกับโปรแกรม ไม่ว่าจะเป็นผู้ป้อนข้อมูลเข้าสู่โปรแกรม หรือผู้รับข้อมูลออกจากโปรแกรม |
|  | โปรแกรมที่ทำการพัฒนา |
|  | ทิศทางการไหล หรือการติดต่อของข้อมูลในโปรแกรม |

Source Destination ที่เกี่ยวข้องกับโปรแกรมการทดสอบ ได้แก่ ผู้สอน นักศึกษา ซึ่งสามารถเขียนแผนผังบริบทแสดงทิศทางการรับส่งข้อมูลในโปรแกรมได้ดังนี้




ภาพที่ 3-6 แผนผังบริบทของโปรแกรมการทดสอบ

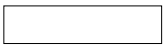
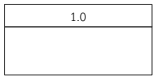
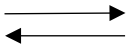
จากภาพที่ 3-6 แสดงการรับส่งข้อมูลระหว่างโปรแกรมกับ Source Destination ซึ่ง Source Destination มี 2 แหล่ง ได้แก่

ผู้สอน หมายถึง ครู อาจารย์ หรือบุคลากร ที่มีหน้าที่ในการจัดการเรียนการสอนให้กับผู้สอบ โดยเป็นผู้ที่สามารถเพิ่ม ลบ หรือแก้ไขข้อสอบในคลังข้อสอบได้ รวมทั้งสามารถตรวจสอบผลการทดสอบได้

ผู้สอบ หมายถึง นักเรียน นักศึกษา หรือบุคคลอื่น ๆ ในสถานศึกษาที่มีหน้าที่เรียนรู้และเข้าทดสอบเพื่อวัดความรู้ในเรื่องนั้น ตามที่โปรแกรมการทดสอบมุ่งหมายที่จะวัด

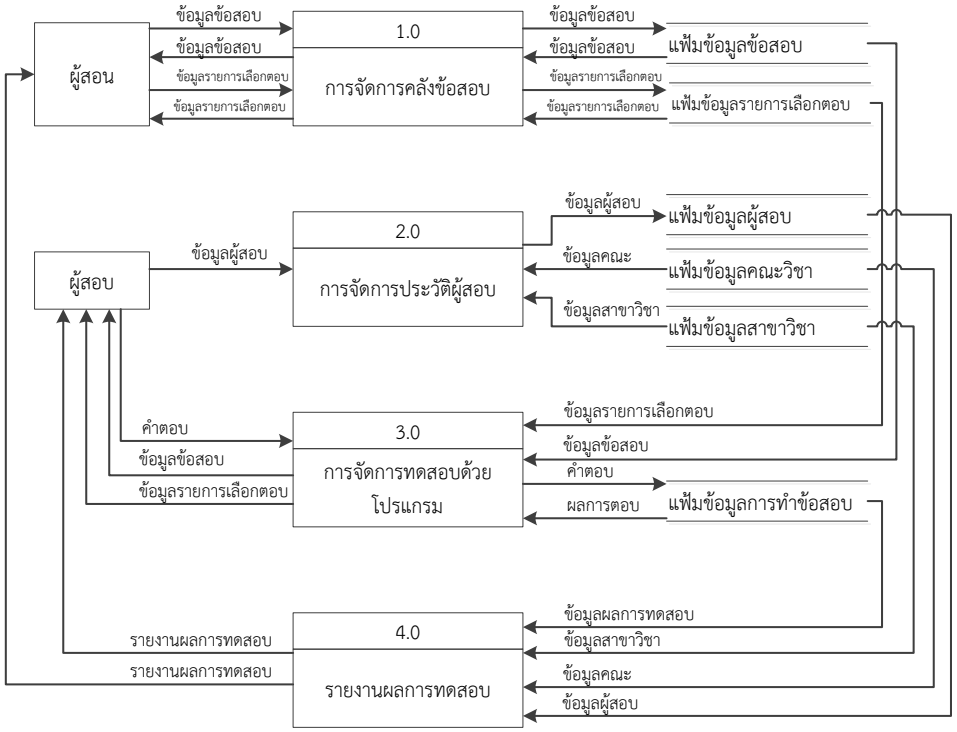
**1.2 การจัดทำผังการไหลของข้อมูล** (Data Flow Diagram: DFD) เป็นการวิเคราะห์ให้เห็นภาพรวมในการทำงานของโปรแกรม โดยจะแสดงรายละเอียดการรับหรือส่งข้อมูลภายในโปรแกรม ซึ่งมีสัญลักษณ์ที่ใช้ในการแสดงผังการไหลของข้อมูลดังนี้

ตารางที่ 3-5 สัญลักษณ์และความหมายที่ใช้ในผังการไหลของข้อมูล

| สัญลักษณ์ | ความหมาย |
|---|---|
|  | ผู้ใช้ที่มีความเกี่ยวข้องกับโปรแกรม ไม่ว่าจะเป็นผู้ป้อนข้อมูลเข้าสู่โปรแกรม หรือผู้รับข้อมูลออกจากโปรแกรม |
|  | การประมวลผลข้อมูลที่เกิดขึ้นในโปรแกรม |
| | แหล่งเก็บข้อมูล เช่น ไฟล์ ตารางจัดเก็บข้อมูล |
|  | ทิศทางการไหลของข้อมูล โดยเขียนข้อความที่บ่งบอกการไหลของข้อมูล |

การจัดทำผังการไหลของข้อมูลสามารถแบ่งขั้นตอนย่อยตามระดับชั้นได้ดังนี้

**1.2.1 แผนผังการไหลของข้อมูลระดับ 0** (Data Flow Diagram Level 0) เป็นแผนผังแสดงการรับส่งข้อมูลที่อยู่ในกระบวนการหลักของแผนผังบริบท ได้แก่ การจัดการคลังข้อสอบ การจัดประวัติผู้สอบ การจัดการทดสอบด้วยโปรแกรม และรายงานผลการทดสอบ ซึ่งแต่ละกระบวนการมีรายละเอียดได้ภาพที่ 3-7




ภาพที่ 3-7 แผนผังการไหลของข้อมูลระดับ 0

จากภาพที่ 3-7 แสดงการรับส่งข้อมูลของกระบวนการหลักในโปรแกรมการทดสอบ ซึ่งแต่ละกระบวนการหลักมีรายละเอียดดังนี้

กระบวนการหลักที่ 1 การจัดการคลังข้อสอบ เป็นกระบวนการที่ใช้จัดเก็บข้อมูล ซึ่งเป็นข้อสอบที่ใช้ในการวัดความรู้ของผู้สอบ โดยข้อสอบที่เก็บเข้าคลังข้อสอบนี้ ต้องเป็นข้อสอบที่มีรายการคำตอบอย่างน้อย 4 ตัวเลือก และผ่านการวิเคราะห์คุณภาพของข้อสอบตามหลักทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Dichotomous IRT) แบบ 3 พารามิเตอร์ ข้อสอบแต่ละข้อต้องมีค่าความยากของข้อสอบ (b) ตั้งแต่ -2.50 ถึง 2.50 มีค่าอำนาจจำแนกของข้อสอบ (a) ตั้งแต่ 0.5 ถึง 2.50 และมีค่าการเดาของข้อสอบไม่เกิน 0.30

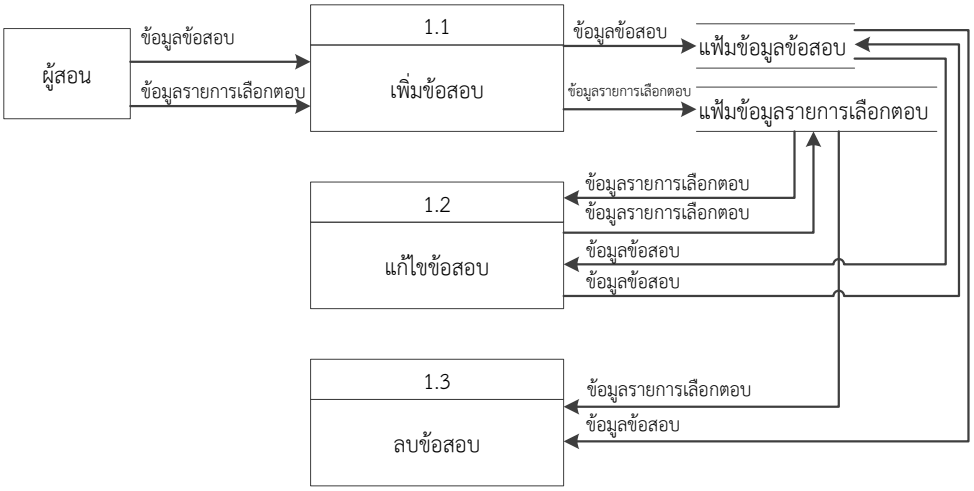
กระบวนการหลักที่ 2 การจัดการประวัติผู้สอบ เป็นกระบวนการที่ใช้จัดเก็บข้อมูลซึ่งเป็นประวัติทั่วไปของผู้สอบ

กระบวนการหลักที่ 3 การจัดการทดสอบด้วยโปรแกรม เป็นกระบวนการดำเนินการทดสอบ ซึ่งมี 2 แบบ ได้แก่ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์

กระบวนการหลักที่ 4 การรายงานผลการสอบ เป็นกระบวนการที่นำข้อมูล ซึ่งเป็นผลการสอบ มาจัดทำเป็นรายงานแล้วนำเสนอให้ผู้สอบ หรือผู้เกี่ยวข้องทราบ

**1.2.2 แผนผังการไหลของข้อมูลระดับ 1** (Data Flow Diagram Level 1) เป็นแผนผังที่แสดงรายละเอียดการรับส่งของข้อมูลในกระบวนการหลัก ซึ่งในที่นี้มีกระบวนการหลัก 2 กระบวน ได้แก่

กระบวนการหลักที่ 1 การจัดการคลังข้อสอบ มี 3 กระบวนการย่อย คือ 1.1) การเพิ่มข้อสอบ 1.2) การแก้ไขข้อสอบ และ 1.3) การลบข้อสอบ โดยแสดงได้ดังภาพที่ 3-8




ภาพที่ 3-8 แผนผังการไหลของข้อมูลระดับ 1 ของกระบวนการจัดการคลังข้อสอบ

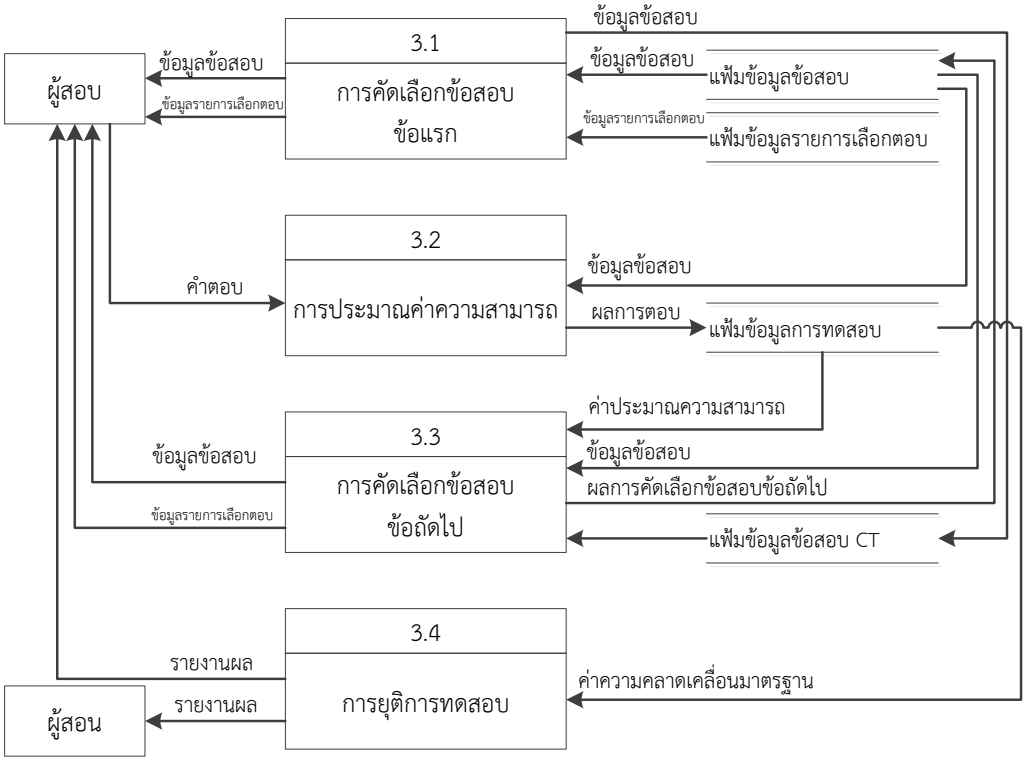
จากภาพที่ 3-8 แสดงแผนผังการรับส่งข้อมูลในกระบวนการจัดการคลังข้อสอบ ซึ่งมีกระบวนการย่อย 3 กระบวนการ คือ

1.1) การเพิ่มข้อสอบ เป็นการเพิ่มข้อสอบใหม่ลงในคลังข้อสอบ โดยข้อสอบที่เพิ่มลงไปต้องเป็นข้อสอบที่มีรายการคำตอบอย่างน้อย 4 ตัวเลือก มีค่าความยากของข้อสอบ ตั้งแต่ -2.50 ถึง 2.50 มีค่าอำนาจจำแนกของข้อสอบ ตั้งแต่ 0.50 ถึง 2.50 และมีค่าการเดาของข้อสอบ ไม่เกิน 0.30

1.2) การแก้ไขข้อสอบ เป็นการแก้ไขข้อสอบเดิมที่มีอยู่ในคลังข้อสอบ โดยข้อสอบที่ได้รับการแก้ไขแล้ว ต้องเป็นข้อสอบที่มีรายการคำตอบอย่างน้อย 4 ตัวเลือก มีค่าความยากของข้อสอบ ตั้งแต่ -2.50 ถึง 2.50 มีค่าอำนาจจำแนกของข้อสอบ ตั้งแต่ 0.50 ถึง 2.50 และมีค่าการเดาของข้อสอบ ไม่เกิน 0.30

1.3) การลบข้อสอบ เป็นการลบข้อสอบที่ไม่ต้องการออกจากคลังข้อสอบ

กระบวนการหลักที่ 3 การจัดการทดสอบด้วยโปรแกรม เป็นกระบวนการจัดการทดสอบ ซึ่งมี 4 กระบวนการย่อย ได้แก่ 3.1) การคัดเลือกข้อสอบข้อแรก 3.2) การประมาณค่าความสามารถของผู้สอบ 3.3) การคัดเลือกข้อสอบข้อถัดไป และ 3.4) ยุติการทดสอบ โดยแสดงได้ดังภาพที่ 3-9




ภาพที่ 3-9 แผนผังการไหลของข้อมูลระดับ 1 ของกระบวนการจัดการทดสอบด้วยโปรแกรม

ภาพที่ 3-9 แสดงการรับส่งข้อมูลในกระบวนการจัดการทดสอบด้วยโปรแกรม ซึ่งมีกระบวนการย่อย 4 กระบวนการ โดยแต่ละกระบวนการมี 2 แบบ ขึ้นอยู่กับการทดสอบนั้นว่าเป็นการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) หรือเป็นการทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) ซึ่งมีรายละเอียดดังนี้ คือ

3.1) การคัดเลือกข้อสอบข้อแรก เป็นการจัดข้อสอบข้อแรกให้แก่ผู้สอบ มี 2 วิธีการขึ้นอยู่กับการทดสอบ ได้แก่

ก. การทดสอบแบบ CAT ข้อสอบข้อแรก จะคัดเลือกด้วยการสุ่มข้อสอบในคลังข้อสอบที่มีค่าความยากของข้อสอบระหว่าง -1.0 ถึง 1.0

ข. การทดสอบแบบ CT ข้อสอบข้อแรก จะจัดข้อสอบที่กำหนดไว้แล้วให้แก่ผู้สอบ ซึ่งผู้สอบทุกคนจะได้ข้อสอบข้อแรกข้อเดียวกัน

3.2) การประมาณค่าความสามารถ เป็นการประมาณค่าความสามารถของผู้สอบของการทดสอบแบบ CAT โดยใช้วิธีของเบส์แบบปรับใหม่ ซึ่งการประมาณค่าจะอาศัยผลการตอบ

ข้อสอบในทุกข้อที่ผ่านมา (คำนวณได้ตามสมการที่ 12-17) ส่วนการทดสอบแบบ CT ไม่มีกระบวนการนี้

3.3) การคัดเลือกข้อสอบข้อถัดไป เป็นการจัดข้อสอบข้อถัดไปให้แก่ผู้สอบ มี 2 วิธีการ ขึ้นอยู่กับการทดสอบ ได้แก่

ก. การทดสอบแบบ CAT ข้อสอบข้อถัดไป จะเป็นข้อสอบที่ถูกคัดเลือกด้วยวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

ข. การทดสอบแบบ CT ข้อสอบข้อถัดไป จะเป็นข้อสอบที่กำหนดไว้แล้ว โดยข้อสอบที่จัดให้ผู้สอบจะมีค่าความยากของข้อสอบคละกัน และผู้สอบทุกคนจะได้ข้อสอบข้อถัดไปข้อเดียวกัน

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ นอกจากต้องคำนึงถึงวิธีการคัดเลือกข้อสอบข้อถัดไปและการควบคุมการใช้ข้อสอบแล้ว ยังต้องคำนึงถึงการควบคุมสัดส่วนเนื้อหาของข้อสอบ (Content Balancing) ซึ่งเป็นการควบคุมเนื้อหาของข้อสอบที่จัดให้แก่ผู้สอบ เพื่อป้องกันไม่ให้ผู้สอบได้รับข้อสอบที่มีเนื้อหาด้านใดด้านหนึ่งเพียงด้านเดียว (Thompson & Weiss, 2011)

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นี้ ผู้วิจัยได้เลือกวิธีการของ Kingsbury-Zara (1989) มาควบคุมสัดส่วนเนื้อหาของข้อสอบ ซึ่งวิธีการนี้จะเข้าแทรกแซงกระบวนการการคัดเลือกข้อสอบข้อถัดไป เพื่อให้ผู้สอบได้รับข้อสอบครบทุกหัวข้อเรื่องและมีจำนวนข้อสอบในแต่ละหัวข้อเรื่องเป็นไปตามสัดส่วนที่ระบุ วิธีการ Kingsbury-Zara มีรายละเอียดดังนี้

ขั้นที่ 1 กำหนดสัดส่วนของข้อสอบที่ต้องการในแต่ละหัวข้อเรื่อง เนื่องจากผู้สอบแต่ละคนต้องได้รับข้อสอบครบทุกหัวข้อเรื่อง และมีจำนวนข้อสอบในแต่ละหัวข้อเรื่องเท่ากัน ซึ่งคลังข้อสอบที่ใช้ในการพัฒนาโปรแกรมการทดสอบ แบ่งเนื้อหาได้ 4 หัวข้อเรื่อง ดังนั้น ผู้วิจัยจึงกำหนดสัดส่วนของข้อสอบที่ต้องการในแต่ละหัวข้อเรื่องเป็น 25%, 25%, 25% และ 25% ตามลำดับ

ขั้นที่ 2 คำนวนสัดส่วนของข้อสอบในแต่ละหัวข้อเรื่องที่ผู้สอบได้รับ โดยคำนวณจากจำนวนข้อสอบในแต่ละหัวข้อเรื่องที่ผู้สอบได้รับหารด้วยจำนวนข้อสอบทั้งหมดในหัวข้อเรื่องนั้น คูณ 100 เช่น ในระหว่างการทดสอบ ผู้สอบทำข้อสอบไปแล้ว 7 ข้อ เป็นข้อสอบในหัวข้อเรื่องที่ 1 ถึง 4 จำนวน 2 ข้อ, 3 ข้อ, 2 ข้อ และ 2 ข้อ ตามลำดับ ซึ่งในคลังข้อสอบมีข้อสอบในหัวข้อเรื่องที่ 1 ถึง 4 จำนวน 118 ข้อ, 104 ข้อ, 107 ข้อ และ 92 ข้อ ตามลำดับ ดังนั้น สัดส่วนของข้อสอบที่ผู้สอบได้รับในหัวข้อเรื่องที่ 1 ถึง 4 คือ $\frac{2}{118}\times 100$, $\frac{3}{104}\times 100$, $\frac{2}{107}\times 100$, และ $\frac{2}{92}\times 100$ คิดเป็น 1.7%, 2.9%, 1.9% และ 2.2% ตามลำดับ

ขั้นที่ 3 คำนวนผลต่างของสัดส่วนที่ได้จากขั้นที่ 1 กับขั้นที่ 2 ดังนั้น ข้อสอบในหัวข้อเรื่องที่ 1 มีผลต่าง 25% - 1.7% = 23.3% ข้อสอบในหัวข้อเรื่องที่ 2 มีผลต่าง 25% - 2.9% = 22.1% ข้อสอบในหัวข้อเรื่องที่ 3 มีผลต่าง 25% - 1.9% = 23.1% และข้อสอบในหัวข้อเรื่องที่ 4 มีผลต่าง 25% - 2.2% = 22.8%

ขั้นที่ 4 คัดเลือกจากข้อสอบในหัวข้อเรื่องที่มีผลต่างมากสุด เป็นข้อสอบข้อถัดไป ซึ่งในขั้นที่ 3 ข้อสอบในหัวข้อเรื่องที่ 1 มีผลต่างมากที่สุด ดังนั้น ข้อสอบข้อถัดไปจะคัดเลือกจากข้อสอบในหัวข้อเรื่อง 1

3.4) ยุติการทดสอบ เป็นกระบวนการที่ทำให้การทดสอบยุติลง มี 2 วิธีการ ขึ้นอยู่กับการทดสอบ ได้แก่

ก. การทดสอบแบบ CAT เกณฑ์ยุติการทดสอบ ใช้เกณฑ์เมื่อค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบมีค่าน้อยกว่า 0.3

ข. การทดสอบแบบ CT เกณฑ์ยุติการทดสอบ ใช้เกณฑ์เมื่อผู้สอบทำข้อสอบครบ 40 ข้อ

2. **ออกแบบฐานข้อมูล**

การออกแบบฐานข้อมูลของโปรแกรมการทดสอบในการวิจัยนี้ มี 7 แฟ้มข้อมูล ได้แก่ 1) แฟ้มข้อมูลข้อสอบ 2) แฟ้มข้อมูลรายการคำตอบ 3) แฟ้มข้อมูลผู้สอบ 4) แฟ้มข้อมูลการทดสอบ 5) แฟ้มข้อมูลสาขาวิชา 6) แฟ้มข้อมูลคณะวิชา และ 7) แฟ้มข้อมูลข้อสอบการทดสอบด้วยคอมพิวเตอร์ โดยแต่ละแฟ้มข้อมูลมีรายละเอียดดังนี้

2.1 แฟ้มข้อมูลข้อสอบ มีทั้งหมด 7 Fields ดังตารางที่ 3-6

ตารางที่ 3-6 แฟ้มข้อมูลข้อสอบ (Examination)

| Key | Field Name | Field Type | Size | Note | Sample |
|---|---|---|---|---|---|
| PK | ExamID | Int | - | ข้อสอบข้อที่ | 1 |
| | ExamSection | Varchar | 100 | เนื้อหาของข้อสอบ | Reading, Listening, Grammar |
| | Question | Varchar | Max | โจทย์ | 2+3=? |
| | A_Parameter | Varchar | 7 | ค่าอำนาจจำแนกของข้อสอบ | 0.57 |
| | B_Parameter | Varchar | 7 | ค่าความยากของข้อสอบ | -1.62 |
| | C_Parameter | Varchar | 6 | ค่าการเดาของข้อสอบ | 0.12 |
| | CorrectAnswer | Varchar | 1 | คำตอบที่ถูกต้อง | 2 |

2.2 แฟ้มข้อมูลรายการคำตอบ มีทั้งหมด 3 Fields ดังตารางที่ 3-7

ตารางที่ 3-7 แฟ้มข้อมูลรายการคำตอบ (Examination Choice)

| Key | Field Name | Field Type | Size | Note | Sample |
|---|---|---|---|---|---|
| PK | ExamID | Int | - | ข้อสอบข้อที่ | 1 |
| Pk | ChoiceID | Int | - | รายการคำตอบที่ | 1 |
| | Choice | Varchar | Max | รายการคำตอบ | red |

2.3 แฟ้มข้อมูลผู้สอบ มีทั้งหมด 10 Fields ดังตารางที่ 3-8

ตารางที่ 3-8 แฟ้มข้อมูลผู้สอบ (StudentInfo)

| Key | Field Name | Field Type | Size | Note | Sample |
|---|---|---|---|---|---|
| PK | M_Runno | Int | - | Running Number | 1 |
| | StudentID | Varchar | 50 | รหัสนักศึกษา | 58000001 |
| FK | FacultyID | Int | - | รหัสคณะวิชา | 12 |
| FK | MajorID | Int | - | รหัสสาขาวิชา | 761 |
| | StudentName | Varchar | 500 | ชื่อ-นามสกุลนักศึกษา | อารี ปกป้อง |
| | StartDateTime | Varchar | 50 | วัน-เวลา เริ่มการสอบ | 14-06-2017 13:07:35 |
| | EndDateTime | Varchar | 50 | วัน-เวลา สิ้นสุดการสอบ | 14-06-2017 14:12:35 |
| | Correct | Int | - | จำนวนข้อถูก | 5 |
| | InCorrect | Int | - | จำนวนข้อผิด | 2 |
| | AbilityValue | Numeric | (18,5) | ค่าประมาณความสามารถ | 1.57 |

2.4 แฟ้มข้อมูลการทำข้อสอบ มีทั้งหมด 6 Fields ดังตารางที่ 3-9

ตารางที่ 3-9 แฟ้มข้อมูลการทดสอบ (Test)

| Key | Field Name | Field Type | Size | Note | Sample |
|---|---|---|---|---|---|
| PK | T_Runno | Int | - | Running Number | 1 |
| FK | M_Runno | Int | - | Running Number ของแฟ้มข้อมูลผู้สอบ | 1 |
| | ExamSection | Varchar | 100 | เนื้อหาของข้อสอบ | Reading, Listening, Grammar |
| | ExamID | Int | - | ข้อสอบข้อที่ | 1 |
| | UserAnswer | Varchar | 1 | คำตอบที่ผู้สอบเลือก | 2 |
| | IsCorrect | Bit | - | ผลการตอบข้อสอบ | 1=True, 0=False |

2.5 แฟ้มข้อมูลสาขาวิชา มีทั้งหมด 3 Fields ดังตารางที่ 3-10

ตารางที่ 3-10 แฟ้มข้อมูลสาขาวิชา (Major)

| Key | Field Name | Field Type | Size | Note | Sample |
|---|---|---|---|---|---|
| PK | MajorID | Int | - | รหัสสาขาวิชา | 1 |
| FK | FacultyID | Int | - | รหัสคณะวิชา | 1 |
| | MajorName | Varchar | 500 | ชื่อสาขาวิชา | การตลาด |

2.6 แฟ้มข้อมูลคณะวิชา มีทั้งหมด 2 Fields ดังตารางที่ 3-11

ตารางที่ 3-11 แฟ้มข้อมูลคณะวิชา (Faculty)

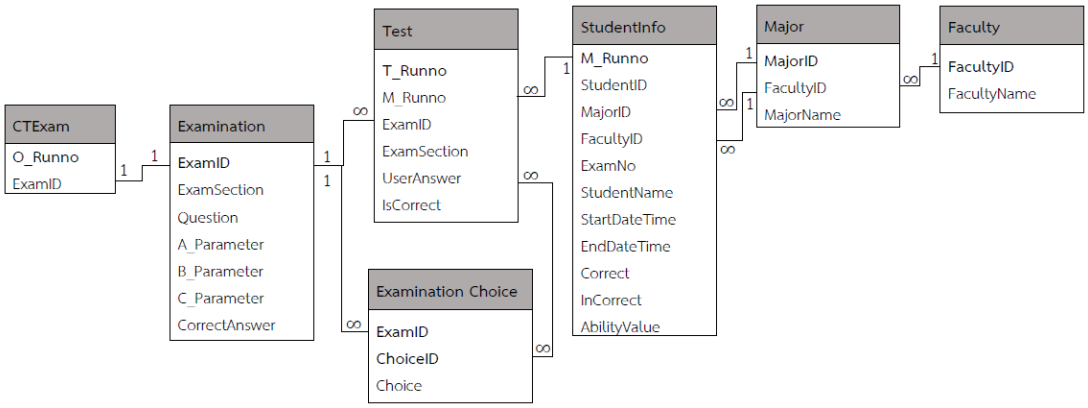
| Key | Field Name | Field Type | Size | Note | Sample |
|---|---|---|---|---|---|
| PK | FacultyID | Int | - | รหัสคณะวิชา | 1 |
| | FacultyName | Varchar | 500 | ชื่อคณะวิชา | วิทยาการจัดการ |

2.7 แฟ้มข้อมูลข้อสอบการทดสอบด้วยคอมพิวเตอร์ มีทั้งหมด 2 Fields ดังตารางที่ 3-12

ตารางที่ 3-12 แฟ้มข้อมูลข้อสอบการทดสอบด้วยคอมพิวเตอร์ (CTExam)

| Key | Field Name | Field Type | Size | Note | Sample |
|---|---|---|---|---|---|
| PK | O_Runno | int | - | Running Number | 1 |
| | ExamID | Int | - | ข้อสอบข้อที่ | 1 |

จากฐานข้อมูลของโปรแกรมการทดสอบข้างต้น สามารถเขียนแผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (Database Diagram) ได้ตามภาพที่ 3-10

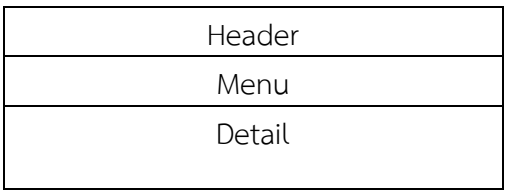



ภาพที่ 3-10 ความสัมพันธ์ระหว่างข้อมูล (Database Diagram)

**3. การพัฒนาโปรแกรม**

การพัฒนาโปรแกรมการทดสอบในการวิจัยนี้ เป็นการพัฒนาในรูปแบบของ Web Application โดยใช้โปรแกรม Notepad++ ที่รองรับการเขียนโปรแกรมด้วยภาษา PHP และใช้ MySQL เป็นฐานข้อมูล และใช้ภาษา SQL เขียนคำสั่งในการเชื่อมโยงข้อมูลในฐานข้อมูล การออกแบบโครงสร้างหน้าจอของโปรแกรมการทดสอบมีดังนี้

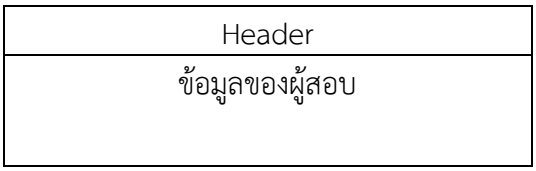
3.1 โครงสร้างหน้าจอทั่วไปของโปรแกรมการทดสอบ เป็นการแจ้งรายละเอียดต่าง ๆ ที่เกี่ยวข้องกับโปรแกรมการทดสอบ แสดงได้ตามภาพที่ 3-11




ภาพที่ 3-11 โครงสร้างหน้าจอทั่วไปของโปรแกรมการทดสอบ

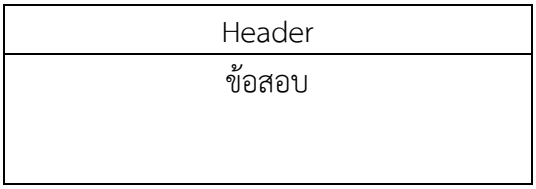
3.2 โครงสร้างหน้าจอการทดสอบของผู้สอบ เป็นการแสดงหน้าจอการดำเนินการทดสอบสำหรับผู้สอบ มีรายละเอียดดังนี้

3.2.1 โครงสร้างหน้าจอการกรอกข้อมูลเบื้องต้นของผู้สอบ เป็นหน้าจอที่ให้ผู้สอบกรอกข้อมูลเบื้องต้นของตนเอง เช่น ชื่อ นามสกุล เลขที่ประจำตัวสอบ เป็นต้น แสดงได้ตามภาพที่ 3-12




ภาพที่ 3-12 โครงสร้างหน้าจอการกรอกข้อมูลเบื้องต้นของผู้สอบ

3.2.2 โครงสร้างหน้าจอการทดสอบ เป็นหน้าจอที่แสดงข้อสอบให้แก่ผู้สอบ ซึ่งจะแสดงข้อสอบทีละข้อ แสดงได้ตามภาพที่ 3-13



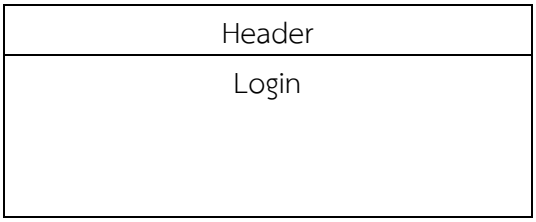

ภาพที่ 3-13 โครงสร้างหน้าจอการทดสอบ

3.2.3 โครงสร้างหน้าจอการรายงานผลการทดสอบ เป็นหน้าจอที่แสดงรายงานผลการทดสอบให้แก่ผู้สอบ เช่น คะแนนการทดสอบ จำนวนข้อที่ใช้ในการทดสอบ หรือเวลาที่ใช้ในการทดสอบ แสดงได้ตามภาพที่ 3-14

ภาพที่ 3-14 โครงสร้างหน้าจอการรายงานผลการทดสอบ

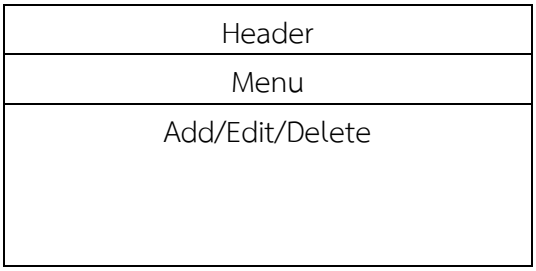
3.3 โครงสร้างหน้าจอการจัดการข้อสอบ เป็นหน้าจอแสดงการปรับปรุงข้อสอบในคลังข้อสอบ สำหรับอาจารย์ผู้สอนหรือบุคคลที่ได้รับมอบหมายให้สามารถปรับปรุงข้อสอบได้ ซึ่งมีรายละเอียดดังนี้

3.3.1 โครงสร้างหน้าจอการ Login เป็นหน้าจอที่ให้อาจารย์ผู้สอน หรือบุคคลที่ได้รับมอบหมาย กรอกชื่อผู้ใช้ (Username) และรหัสผ่าน (Password) เพื่อเข้าสู่ส่วนการปรับปรุงข้อสอบ แสดงได้ตามภาพที่ 3-15



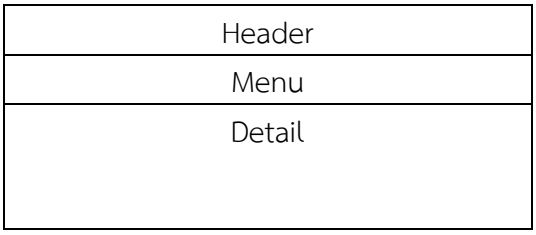

ภาพที่ 3-15 โครงสร้างหน้าจอ Login

3.3.2 โครงสร้างหน้าจอการจัดการข้อสอบ เป็นหน้าจอที่แสดงรายละเอียดในการเพิ่มข้อสอบ ลบข้อสอบ หรือแก้ไขข้อสอบ ในคลังข้อสอบ แสดงได้ตามภาพที่ 3-16




ภาพที่ 3-16 โครงสร้างหน้าจอการจัดการข้อสอบ

3.3.3 โครงสร้างหน้าจอการรายงานผลการจัดการข้อสอบ เป็นหน้าจอที่แสดงรายงานผลที่เกี่ยวข้องกับการจัดการข้อสอบครั้งล่าสุด เช่น จำนวนข้อสอบทั้งหมด ค่าต่ำสุด-สูงสุดของค่าพารามิเตอร์ของข้อสอบทั้งหมด แสดงได้ตามภาพที่ 3-17




ภาพที่ 3-17 โครงสร้างหน้าจอรายงานผลการจัดการข้อสอบ

**4. การทดสอบและแก้ไขโปรแกรม**

เมื่อพัฒนาโปรแกรมการทดสอบเป็นที่เรียบร้อยแล้ว ผู้วิจัยร่วมกับโปรแกรมเมอร์ได้ตรวจสอบโปรแกรมการทดสอบ เพื่อหาจุดบกพร่อง ได้แก่ พิมพ์เครื่องหมายการดำเนินการทางคณิตศาสตร์ผิด (+ − × ÷) พิมพ์ตัวแปรผิด รวมทั้งการใช้คำสั่งในการเขียนโปรแกรมผิดพลาด ซึ่งได้ตรวจสอบทีละคำสั่ง แล้วแก้ไขโปรแกรมให้ถูกต้อง นอกจากนี้ เพื่อให้มั่นใจว่าโปรแกรมการทดสอบที่พัฒนาขึ้นมีความถูกต้อง สามารถใช้งานได้จริง จึงรันโปรแกรมการทดสอบ แล้วนำผลการทดสอบเทียบกับผลการทดสอบที่ได้จากการศึกษาในสถานการณ์จำลอง ปรากฏว่า โปรแกรมการทดสอบที่พัฒนาขึ้นให้ผลการทดสอบเท่ากับผลการทดสอบจากการศึกษาในสถานการณ์จำลอง จึงมั่นใจได้ว่าโปรแกรมการทดสอบที่พัฒนาขึ้นไม่มีความผิดพลาด จากนั้นจึงนำโปรแกรมเสนอต่ออาจารย์ที่ปรึกษาเพื่อขอความคิดเห็น แล้วนำข้อคิดเห็นที่ได้มาปรับปรุงโปรแกรมให้มีความสมบูรณ์ยิ่งขึ้นต่อไป

**5. การจัดทำคู่มือการใช้โปรแกรม**

หลังจากทดสอบและแก้ไขโปรแกรมเป็นที่น่าพอใจแล้ว ได้จัดทำคู่มือในการใช้โปรแกรมการทดสอบนี้ เพื่ออธิบายวิธีการใช้โปรแกรม โดยแบ่งออกเป็น 2 ส่วน ได้แก่ ส่วนที่ 1 การจัดการข้อสอบ ประกอบด้วย การตั้งรหัสกลุ่มผู้สอบ การจัดการข้อสอบ การรายงานผลสอบ (ต่ออาจารย์) และการออกจากระบบ และส่วนที่ 2 การจัดการสอบ ประกอบด้วย เริ่มต้นการทดสอบ กรอกข้อมูลผู้สอบและเลือกแบบทดสอบ การดำเนินการสอบ และการรายงานผลสอบ (ต่อผู้สอบ) แสดงรายละเอียดในภาคผนวก ง

**6. การประเมินโปรแกรมและคู่มือการใช้โปรแกรม**

การประเมินโปรแกรมและคู่มือการใช้โปรแกรม ดำเนินการหลังจากทดสอบ ปรับปรุงแก้ไขและจัดทำคู่มือการใช้โปรแกรมเป็นที่เรียบร้อยแล้ว โดยผู้วิจัยนำโปรแกรมและคู่มือการใช้โปรแกรมไปประเมินโปรแกรมและคู่มือการใช้โปรแกรม ซึ่งแบ่งได้ 2 ส่วน คือ การประเมินโดยผู้เชี่ยวชาญ และการประเมินโดยผู้สอบ โดยมีรายละเอียดดังนี้

6.1 การประเมินโดยผู้เชี่ยวชาญ

การประเมินโดยผู้เชี่ยวชาญ ได้กำหนดเกณฑ์ว่าผู้เชี่ยวชาญต้องมีความรู้ทางด้านการวัดผล หรือด้านการพัฒนาโปรแกรมคอมพิวเตอร์ และมีประสบการณ์ทางด้านการวัดผล หรือด้านการพัฒนาโปรแกรมคอมพิวเตอร์ ไม่น้อยกว่า 10 ปี และมีวุฒิการศึกษาไม่ต่ำกว่าระดับปริญญาโท ผู้เชี่ยวชาญประกอบด้วย

1) ผู้ช่วยศาสตราจารย์ ดร. พีรพัฒน์ ยางกลาง อาจารย์ประจำคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร

2) ผู้ช่วยศาสตราจารย์ ดร. สิริชัย ดีเลิศ อาจารย์ประจำคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร

3) อาจารย์ ดร. สุชาดา สกลกิจรุ่งโรจน์ อาจารย์ประจำสำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช

ผู้เชี่ยวชาญได้ประเมินโปรแกรมและคู่มือการใช้โปรแกรมใน 4 ด้าน ได้แก่ 1) ด้านความสะดวกในการนำไปใช้ 2) ด้านความถูกต้องในการใช้งาน 3) ด้านลักษณะทั่วไป

ของโปรแกรม และ 4) ด้านคู่มือการใช้โปรแกรม โดยใช้แบบประเมินผลที่ผู้วิจัยพัฒนาขึ้น ซึ่งมีลักษณะเป็นมาตรประมาณค่า 5 ระดับ ซึ่งมีเกณฑ์การให้คะแนน ดังนี้

5 หมายถึง เหมาะสมมากที่สุด
4 หมายถึง เหมาะสมมาก
3 หมายถึง เหมาะสมปานกลาง
2 หมายถึง เหมาะสมน้อย
1 หมายถึง เหมาะสมน้อยที่สุด

เกณฑ์การแปลความหมายของคะแนนเฉลี่ย ดังนี้

คะแนนเฉลี่ยตั้งแต่ 4.50 - 5.00 หมายถึง โปรแกรมมีความเหมาะสมมากที่สุด
คะแนนเฉลี่ยตั้งแต่ 3.50 - 4.49 หมายถึง โปรแกรมมีความเหมาะสมมาก
คะแนนเฉลี่ยตั้งแต่ 2.50 - 3.49 หมายถึง โปรแกรมมีความเหมาะสมปานกลาง
คะแนนเฉลี่ยตั้งแต่ 1.50 - 2.49 หมายถึง โปรแกรมมีความเหมาะสมน้อย
คะแนนเฉลี่ยตั้งแต่ 1.00 - 1.49 หมายถึง โปรแกรมมีความเหมาะสมน้อยที่สุด

(แบบประเมินโปรแกรมโดยผู้เชี่ยวชาญ แสดงในภาคผนวก จ)

6.2 การประเมินโดยผู้สอบ

การประเมินโดยผู้สอบ เป็นการประเมินความคิดเห็นในด้านความสะดวกในการนำไปใช้และด้านลักษณะทั่วไปของโปรแกรม โดยกลุ่มตัวอย่างที่ทดลองใช้ เป็นนักศึกษาคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร จำนวน 30 คน คัดเลือกด้วยวิธีการเลือกตัวอย่างตามสะดวก โดยใช้แบบประเมินผลที่ผู้วิจัยพัฒนาขึ้น ในลักษณะมาตรประมาณค่า 5 ระดับ โดยมีเกณฑ์การให้คะแนน ดังนี้

5 หมายถึง เหมาะสมมากที่สุด
4 หมายถึง เหมาะสมมาก
3 หมายถึง เหมาะสมปานกลาง
2 หมายถึง เหมาะสมน้อย
1 หมายถึง เหมาะสมน้อยที่สุด

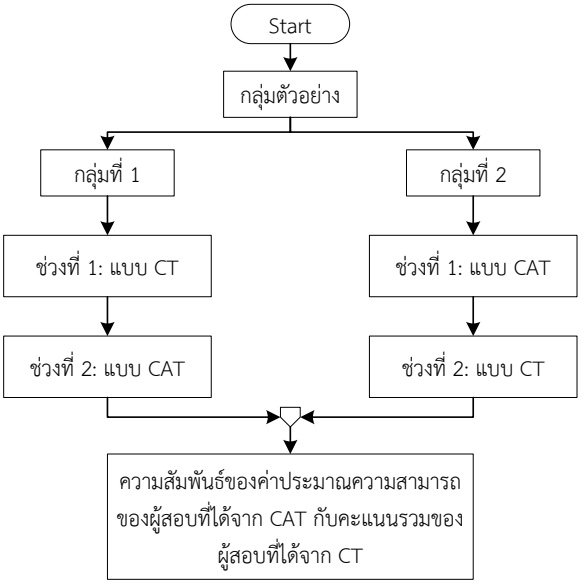
เกณฑ์การแปลความหมายของคะแนนเฉลี่ย ดังนี้

คะแนนเฉลี่ยตั้งแต่ 4.50 - 5.00 หมายถึง โปรแกรมมีความเหมาะสมมากที่สุด
คะแนนเฉลี่ยตั้งแต่ 3.50 - 4.49 หมายถึง โปรแกรมมีความเหมาะสมมาก
คะแนนเฉลี่ยตั้งแต่ 2.50 - 3.49 หมายถึง โปรแกรมมีความเหมาะสมปานกลาง
คะแนนเฉลี่ยตั้งแต่ 1.50 - 2.49 หมายถึง โปรแกรมมีความเหมาะสมน้อย
คะแนนเฉลี่ยตั้งแต่ 1.00 - 1.49 หมายถึง โปรแกรมมีความเหมาะสมน้อยที่สุด

(แบบประเมินโปรแกรมโดยผู้สอบ แสดงในภาคผนวก ฉ)

**ระยะที่ 4 การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์**

การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) ของการวิจัยนี้ จะดำเนินการตามงานวิจัยของ Liley, Barker, and Britton (2004) ที่ได้ศึกษาจากความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ โดยสามารถแสดงแผนการดำเนินการได้ตามภาพที่ 3-18




ภาพที่ 3-18 ขั้นตอนการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์

จากภาพที่ 3-18 แสดงขั้นตอนการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบ CAT กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบแบบ CT โดยผลการศึกษานี้จะแสดงให้เห็นว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีการวัดผลที่สอดคล้องเป็นไปในทิศทางเดียวกับการทดสอบด้วยคอมพิวเตอร์หรือไม่ ด้วยเหตุนี้นักศึกษาแต่ละคนในกลุ่มตัวอย่าง จึงต้องทำการทดสอบทั้ง 2 แบบ โดยใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นในระยะที่ 3 เป็นเครื่องมือที่ใช้ในการวิจัย การดำเนินการวิจัยในระยะนี้ จึงแบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม โดยกลุ่มที่ 1 ดำเนินการทดสอบแบบ CT ก่อน แล้วจึงทำการทดสอบแบบ CAT ส่วนในกลุ่มที่ 2 ดำเนินการทดสอบแบบ CAT ก่อน แล้วจึงทำการทดสอบแบบ

CT ทั้งนี้เพื่อป้องกันปัญหาเรื่องลำดับของการทดสอบ (Order Effect) ซึ่งอาจส่งผลต่อผลการทดสอบได้ หลังจากนั้นจึงนำผลการทดสอบทั้ง 2 แบบ มาทดสอบความสัมพันธ์ วิธีดำเนินการวิจัยในระยะนี้มีรายละเอียดดังนี้

**1. กลุ่มตัวอย่าง**

กลุ่มตัวอย่างที่ใช้ในการวิจัยนี้เป็นนักศึกษาระดับปริญญาตรี ชั้นปีที่ 2 ประจำปีการศึกษา 2559 ในคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร คัดเลือกมาด้วยวิธีการเลือกตัวอย่างแบบตามสะดวกจำนวน 30 คน

**2. เครื่องมือที่ใช้ในการวิจัย**

เครื่องมือที่ใช้ในการวิจัย คือ โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี ซึ่งเป็นโปรแกรมในรูปแบบ Web Application โดยผู้ใช้สามารถเลือกการทดสอบได้ 2 แบบ คือ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และการทดสอบด้วยคอมพิวเตอร์

**3. การเก็บรวมรวมข้อมูล**

การเก็บรวบรวมข้อมูลในการวิจัยนี้ แบ่งการดำเนินการได้ 2 ส่วน คือ ส่วนที่ 1 การเตรียมเอกสารที่เกี่ยวข้องในการเก็บรวมรวมข้อมูล และส่วนที่ 2 การดำเนินการทดสอบ ซึ่งมีรายละเอียดดังนี้

ส่วนที่ 1 การเตรียมเอกสารที่เกี่ยวข้องในการเก็บรวบรวมข้อมูล มีขั้นตอนดังนี้

ขั้นที่ 1 จัดเตรียมเอกสารต่าง ๆ เพื่อออกหนังสือขอความอนุเคราะห์เก็บรวบรวมข้อมูลเพื่อการวิจัย เช่น เค้าโครงการวิจัยฉบับย่อ เครื่องมือ หรือแบบรายงานการผ่านจริยธรรมการวิจัย ให้เสร็จสิ้นภายในวันที่ 1 สิงหาคม พ.ศ. 2559

ขั้นที่ 2 ขอออกหนังสือขอความอนุเคราะห์เก็บรวบรวมข้อมูลเพื่อการวิจัย จากวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา ถึงผู้บริหารคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร ในวันที่ 6 สิงหาคม พ.ศ. 2559

ขั้นที่ 3 จัดส่งหนังสือขอความอนุเคราะห์เก็บรวมรวมข้อมูลเพื่อการวิจัย ถึงผู้บริหารคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร ในวันที่ 22 สิงหาคม พ.ศ. 2559

ส่วนที่ 2 การดำเนินการทดสอบ โดยใช้โปรแกรมการทดสอบที่ผู้วิจัยได้พัฒนาขึ้นตามรายละเอียดที่ให้ไว้ในระยะที่ 3 เพื่อใช้ทดสอบความรู้ทางภาษาอังกฤษกับนักศึกษาชั้นปีที่ 2 คณะวิทยาการจัดการ มหาวิทยาศิลปากร มีขั้นตอนในการดำเนินการดังนี้

ขั้นที่ 1 การจัดเตรียมคลังข้อสอบ ในการทดสอบนี้ใช้คลังข้อสอบที่เป็นข้อสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน โดยคลังข้อสอบทั้งหมดได้รับความอนุเคราะห์จากอาจารย์ประจำหลักสูตรภาษาอังกฤษ มหาวิทยาลัยสวนดุสิต ซึ่งข้อสอบแต่ละข้อ แบ่งได้ 2 ส่วน ได้แก่ ส่วนที่ 1 เนื้อหาของข้อสอบ ประกอบด้วย โจทย์ รายการคำตอบแบบ 4 ตัวเลือก พร้อมเฉลยตัวเลือกที่ถูกต้อง มีทั้งหมด 421 ข้อ แบ่งเนื้อหาของการทดสอบได้ 4 ส่วน ได้แก่ 1) ชนิดของคำ (Part of Speech) จำนวน 118 ข้อ 2) การใช้บริบทบ่งชี้เพื่อค้นหาความหมายของคำศัพท์ยาก (Using Context Clue) จำนวน 104 ข้อ 3) การเรียนรู้ความหมายคำศัพท์จากโครงสร้างของคำศัพท์ (Word Formations) จำนวน 107 ข้อ และ 4) การสรุปความ (Making Inferences) จำนวน 92 ข้อ และ

ส่วนที่ 2 ค่าพารามิเตอร์ของข้อสอบ ข้อสอบทุกข้อในคลังข้อสอบเป็นข้อสอบที่ผ่านการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ ประกอบด้วยค่าพารามิเตอร์ 3 ค่า ได้แก่ ค่าอำนาจจำแนกของข้อสอบ มีค่าตั้งแต่ 0.50 ถึง 2.50 ค่าความยากของข้อสอบ มีค่าตั้งแต่ -2.50 ถึง 2.50 และค่าการเดาของข้อสอบ มีค่าไม่เกิน 0.30 ข้อสอบทั้งหมดในคลังข้อสอบจะถูกนำไปใช้ในโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นตามระยะที่ 3

ขั้นที่ 2 การดำเนินการทดสอบ โดยใช้โปรแกรมการทดสอบที่ผู้วิจัยพัฒนาขึ้น (จากระยะที่ 3) เพื่อใช้ทดสอบความรู้ทางภาษาอังกฤษของนักศึกษา ในวันที่ 1 กันยายน พ.ศ. 2559 ณ ห้องปฏิบัติการคอมพิวเตอร์1 ชั้น 2 อาคารเรียนรวม1 คณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร ซึ่งมีขั้นตอนย่อยในการดำเนินการดังนี้

ขั้นที่ 2.1 แบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม กลุ่มละ 15 คน

ขั้นที่ 2.2 ชี้แจงวัตถุประสงค์ของการทดสอบและวิธีการใช้โปรแกรมอย่างละเอียด เพื่อให้กลุ่มตัวอย่างเห็นความสำคัญของการทดสอบ และทำการทดสอบด้วยความตั้งใจ

ขั้นที่ 2.3 ดำเนินการทดสอบ โดยใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี โดยกลุ่มตัวอย่างทั้ง 2 กลุ่ม ต้องดำเนินการทดสอบทั้ง 2 แบบ คือ การทดสอบแบบ CT และการทดสอบแบบ CAT ซึ่งกลุ่มที่ 1 ดำเนินการทดสอบแบบ CT ก่อน แล้วจึงดำเนินการทดสอบแบบ CAT ส่วนกลุ่มที่ 2 ดำเนินการทดสอบแบบ CAT ก่อน แล้วจึงดำเนินการทดสอบแบบ CT แสดงได้ดังตารางที่ 3-13

ตารางที่ 3-13 ลำดับการทดสอบ

| | สอบช่วงที่ 1 | พัก | สอบช่วงที่ 2 |
|---|---|---|---|
| | 9.00 น. - 10.00 น. | | 10.30 น. - 11.30 น. |
| กลุ่มที่ 1 | แบบ CT | 30 นาที | แบบ CAT |
| กลุ่มที่ 2 | แบบ CAT | | แบบ CT |

**4. การวิเคราะห์ข้อมูล**

การวิเคราะห์ข้อมูลของการวิจัย แบ่งได้ 3 ส่วน ดังนี้

4.1 การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 4 วิธี คือ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง 3) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ และ 4) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ โดยเปรียบเทียบประสิทธิภาพด้าน 1) การประมาณค่าความสามารถของผู้สอบ 2) ความยาวของแบบทดสอบ และ 3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ซึ่งวิเคราะห์ด้วยสถิติดังนี้

4.1.1 การเปรียบเทียบประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบ โดยพิจารณาจากค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error:

RMSE) และค่าความลำเอียงเฉลี่ย (Average Bias) ซึ่งเป็นข้อมูลแบบต่อเนื่อง (Continuous Data) และเมื่อทดสอบความแตกต่างของค่าความแปรปรวน ปรากฏว่า ค่า RMSE และค่าความลำเอียงเฉลี่ยของแต่ละกลุ่ม (วิธีการ) มีค่าความแปรปรวนแตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .05 ซึ่งไม่เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way Analysis of Variance: One-way ANOVA) การวิจัยนี้ จึงใช้การวิเคราะห์ Brown-Forsythe และหากผลการวิเคราะห์มีนัยสำคัญทางสถิติ จะทำการเปรียบเทียบพหุคูณ (Multiple Comparison) ด้วยการทดสอบของแทมเฮน (Tamhane's Test)

4.1.2 การเปรียบเทียบประสิทธิภาพด้านความยาวของแบบทดสอบ ซึ่งเป็นข้อมูลแบบไม่ต่อเนื่อง (Discrete Data) จึงใช้สถิตินอนพาราเมตริก ด้วยการทดสอบ Kruskal-Wallis และหากผลการวิเคราะห์มีนัยสำคัญทางสถิติ จะทำการเปรียบเทียบค่าเฉลี่ยของอันดับเป็นรายคู่ ด้วยการทดสอบ Dunn-Bonferroni

4.1.3 การเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ซึ่งเป็นข้อมูลแบบไม่ต่อเนื่อง (Discrete Data) จึงใช้สถิตินอนพาราเมตริก ด้วยการทดสอบ Kruskal-Wallis และหากผลการวิเคราะห์มีนัยสำคัญทางสถิติ จะทำการเปรียบเทียบค่าเฉลี่ยของอันดับเป็นรายคู่ ด้วยการทดสอบ Dunn-Bonferroni

4.2 การประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยผู้เชี่ยวชาญและผู้สอบ ใช้สถิติเชิงบรรยาย ได้แก่ ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

4.3 การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ใช้สถิติเชิงบรรยาย ได้แก่ ความถี่ (Frequency) ค่าสูงสุด (Max) ค่าต่ำสุด (Min) ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ทั้งนี้ค่าประมาณความสามารถของผู้สอบและคะแนนรวมของผู้สอบเป็นข้อมูลแบบต่อเนื่อง จึงใช้การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน (Pearson's Product Moment Correlation)

# บทที่ 4
# ผลการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จากนั้นนำวิธีการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้น ทั้งในกรณีที่มีและไม่มีการควบคุมการใช้ข้อสอบ มาเปรียบเทียบประสิทธิภาพกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง ซึ่งพิจารณาด้าน 1) การประมาณค่าความสามารถของผู้สอบ 2) ความยาวของแบบทดสอบ และ 3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จากนั้นพัฒนาโปรแกรมการทดสอบ ซึ่งสามารถรองรับทั้งการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และการทดสอบด้วยคอมพิวเตอร์ แล้วนำโปรแกรมการทดสอบที่พัฒนาได้มาศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ผู้วิจัยแบ่งการนำเสนอผลการวิจัยเป็น 4 ตอน ดังนี้

ตอนที่ 1 ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

ตอนที่ 2 ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป

ตอนที่ 3 ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ตอนที่ 4 ผลการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์

เพื่อความเข้าใจตรงกัน ผู้วิจัยได้กำหนดสัญลักษณ์ที่ใช้ในการวิจัย ดังนี้

| | | |
|---|---|---|
| *n* | หมายถึง | ขนาดกลุ่มตัวอย่าง |
| *Mean* | หมายถึง | ค่าเฉลี่ยเลขคณิต |
| *SD* | หมายถึง | ส่วนเบี่ยงเบนมาตรฐาน |
| *Median* | หมายถึง | ค่ามัธยฐาน |
| *Max* | หมายถึง | ค่าสูงสุด |
| *Min* | หมายถึง | ค่าต่ำสุด |
| $\theta$ | หมายถึง | ค่าความสามารถที่แท้จริงของผู้สอบ |
| $\hat{\theta}$ | หมายถึง | ค่าประมาณความสามารถของผู้สอบ |
| a | หมายถึง | ค่าอำนาจจำแนกของข้อสอบ |
| b | หมายถึง | ค่าความยากของข้อสอบ |
| c | หมายถึง | ค่าการเดาของข้อสอบ |
| *Asymp F* | หมายถึง | ค่าสถิติ *F* จากการวิเคราะห์ Brown-Forsythe |

| | | |
|---|---|---|
| $\chi^2$ | หมายถึง | ค่าไคสแควร์ |
| *df1* | หมายถึง | ค่าองศาอิสระ ค่าที่ 1 |
| *df2* | หมายถึง | ค่าองศาอิสระ ค่าที่ 2 |
| *p-value* | หมายถึง | ระดับนัยสำคัญ $\alpha$ ที่จะปฏิเสธ $H_0$ |
| *RMSE* | หมายถึง | ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย |
| MIC | หมายถึง | วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด |
| RDM | หมายถึง | วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง |
| HC | หมายถึง | วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ |
| HC-Ex | หมายถึง | วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ |
| CAT | หมายถึง | การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ |
| CT | หมายถึง | การทดสอบด้วยคอมพิวเตอร์ |

## ตอนที่ 1 ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในการวิจัยนี้เป็นการนำทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ตามเกณฑ์ของเฮอร์วิคซ์มาใช้ในการคัดเลือกข้อสอบข้อถัดไปให้มีความเหมาะสมกับระดับความสามารถของผู้สอบ และใช้วิธีการสุ่มตัวอย่างแบบมีระบบ ซึ่งเป็นหนึ่งในวิธีการสุ่มตัวอย่างที่อาศัยความน่าจะเป็นมาควบคุมการใช้ข้อสอบ โดยวิธีการที่พัฒนาขึ้นมีการดำเนินการที่เรียบง่าย ไม่ยุ่งยากซับซ้อน ในทางปฏิบัติวิธีการควบคุมการใช้ข้อสอบที่ได้พัฒนาขึ้น ถูกดำเนินการก่อนวิธีการคัดเลือกข้อสอบข้อถัดไป ดังนั้น ผู้วิจัยจึงแสดงรายละเอียดของวิธีการควบคุมการใช้ข้อสอบ ต่อด้วยวิธีการคัดเลือกข้อสอบข้อถัดไป ดังนี้

1. ผลการพัฒนาวิธีการควบคุมการใช้ข้อสอบด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ สามารถสรุปเป็นขั้นตอนในการดำเนินการได้ดังนี้

ขั้นที่ 1 แบ่งข้อสอบในคลังข้อสอบ (หรือข้อสอบในหัวข้อเรื่องถัดจากข้อสอบที่ทำขณะนั้น) ออกเป็น 2 กลุ่ม คือ ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบต่ำถึงปานกลาง ซึ่งมีค่าน้อยกว่า 1.35 (Baker, 2001, p. 34) กำหนดให้เป็นกลุ่ม A และข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ซึ่งมีค่าตั้งแต่ 1.35 ขึ้นไป กำหนดให้เป็นกลุ่ม B

ขั้นที่ 2 สุ่มเลือกข้อสอบในกลุ่ม B ด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ ซึ่งมีขั้นตอนการสุ่มดังนี้

ขั้นที่ 2.1 กำหนดจำนวนข้อสอบที่ต้องการสุ่มในกลุ่ม B สมมติว่าต้องการสุ่มข้อสอบจำนวน n ข้อ โดยที่ n เป็นจำนวนเต็มที่สุ่มจากช่วงจำนวนตั้งแต่ 5 ถึง 10 ข้อ

ขั้นที่ 2.2 หาช่วงของการสุ่ม (k) โดยคำนวณจาก $k=\frac{N}{n}$ เมื่อ N คือ จำนวนข้อสอบทั้งหมดในกลุ่ม B และ n คือ ข้อสอบที่ต้องสุ่ม

ขั้นที่ 2.3 หาข้อสอบที่ใช้เป็นตัวสุ่มเริ่มต้น (r) โดยสุ่มค่า r จากค่าตั้งแต่ 1 ถึง k จากนั้นจึงเลือกข้อสอบโดยเริ่มตั้งแต่ข้อที่ r, r+k, r+2k ตามลำดับจนครบ n ข้อ

ขั้นที่ 3 รวมข้อสอบในกลุ่ม A กับข้อสอบที่สุ่มได้ในขั้นที่ 2 เข้าด้วยกัน ซึ่งข้อสอบทั้งหมดในขั้นนี้ จะนำไปใช้ในกระบวนการคัดเลือกข้อสอบข้อถัดไป

(ตัวอย่างแสดงขั้นตอนการควบคุมการใช้ข้อสอบ แสดงในหน้า 65 )

2. ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ สามารถสรุปเป็นขั้นตอนในการดำเนินการได้ดังนี้

ขั้นที่ 1 นำข้อสอบที่ผ่านการควบคุมการใช้ข้อสอบ (จากข้อ 1.1) มาคำนวณหาค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ $H(i)$ ตามเกณฑ์ของเฮอร์วิคซ์ ซึ่งมีสูตรคำนวณดังนี้

$$H(i) = \alpha \times \min(SEE_{true}(\hat{\theta}), SEE_{false}(\hat{\theta})) + (1-\alpha) \times \max(SEE_{true}(\hat{\theta}), SEE_{false}(\hat{\theta})) \qquad (40)$$

โดยที่ $H(i)$ คือ ค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ณ ข้อสอบข้อที่ i

$SEE_{true}(\hat{\theta})$ คือ ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถในกรณีที่ผู้สอบตอบถูก

$SEE_{false}(\hat{\theta})$ คือ ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ในกรณีที่ผู้สอบตอบผิด

min คือ ค่าต่ำสุดที่เลือกจากค่า $SEE_{true}(\hat{\theta})$ และค่า $SEE_{false}(\hat{\theta})$ ขึ้นอยู่กับค่าใดต่ำกว่ากัน

max คือ ค่าสูงสุดที่เลือกจากค่า $SEE_{true}(\hat{\theta})$ และค่า $SEE_{false}(\hat{\theta})$ ขึ้นอยู่กับค่าใดสูงกว่ากัน

$\hat{\theta}$ คือ ค่าประมาณความสามารถของผู้สอบ

$\alpha$ คือ สัมประสิทธิ์ของการมองโลกในแง่ดี มีเท่ากับ $1-\frac{1}{5.5}|\hat{\theta} - b|$

b คือ ค่าความยากของข้อสอบ

ขั้นที่ 2 เลือกข้อสอบที่มีค่า $H(i)$ ต่ำสุด มาเป็นข้อสอบข้อถัดไป

(ตัวอย่างการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE แสดงในหน้า 69)

## ตอนที่ 2 ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป

การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการนำวิธีการคัดเลือกข้อสอบข้อถัดไป และการควบคุมการใช้ข้อสอบที่ผู้วิจัยได้พัฒนาขึ้น มาเปรียบเทียบประสิทธิภาพกับวิธีการคัดเลือกข้อสอบข้อถัดไปที่นิยมใช้กันทั่วไป และวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีแนวคิดคล้ายคลึงกัน ซึ่งเปรียบเทียบวิธีการทั้งหมด 4 วิธี ได้แก่

1. วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC)

2. วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (Risk Decision Making: RDM)

3. วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ (Hurwicz Criterion: HC)

4. วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ (Hurwicz Criterion with Item Exposure Control: HC-Ex)

โดยเปรียบเทียบประสิทธิภาพด้าน 1) การประมาณค่าความสามารถของผู้สอบ 2) ความยาวของแบบทดสอบ และ 3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ซึ่งจะดำเนินการด้วยการศึกษาในสถานการณ์จำลอง โดยจำลองสถานการณ์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปที่แตกต่างกันใน 4 วิธีการข้างต้น การทดสอบในแต่ละวิธีการ ดำเนินการทำซ้ำทั้งหมด 10 รอบ โดยทุกรอบใช้คลังข้อสอบจำลอง ขนาด 500 ข้อ ชุดเดียวกัน แต่ใช้ค่าความสามารถที่แท้จริงของผู้สอบที่ละชุด ชุดละ 1,000 ค่า โดยการทดสอบรอบที่ 1 ใช้ค่าความสามารถที่แท้จริงของผู้สอบชุดที่ 1 การทดสอบรอบที่ 2 ใช้ค่าความสามารถที่แท้จริงของผู้สอบชุดที่ 2 จนกระทั่งถึงการทดสอบรอบที่ 10 ใช้ค่าความสามารถที่แท้จริงของผู้สอบชุดที่ 10 ซึ่งการคัดเลือกข้อสอบข้อถัดไปทุกวิธีการ ใช้ค่าความสามารถที่แท้จริงของผู้สอบ 10 ชุดนี้ เหมือนกัน

ค่าความสามารถที่แท้จริงของผู้สอบ ถูกจำลองขึ้นด้วยโปรแกรม WinGen3 โดยกำหนดให้โปรแกรมสุ่มข้อมูลจากเลขสุ่มที่มีการแจกแจงแบบปกติ N (0, 1) จำนวน 10,000 ค่า แบ่งเป็นชุด ๆ ละ 1,000 ค่า ได้ทั้งหมด 10 ชุด แสดงดังตารางที่ 4-1

ตารางที่ 4-1 ค่าสถิติพื้นฐานของค่าความสามารถที่แท้จริงของผู้สอบ

| ค่าความสามารถที่แท้จริง ($\theta$) | *n* | *Min* | *Max* | *Mean* | *SD* | *Skewness* | *Kurtosis* |
|---|---|---|---|---|---|---|---|
| ชุดที่ 1 | 1,000 | -3.224 | 4.255 | .007 | 1.004 | .047 | .294 |
| ชุดที่ 2 | 1,000 | -3.200 | 3.308 | .035 | 1.044 | -.013 | -.069 |
| ชุดที่ 3 | 1,000 | -3.301 | 3.463 | .031 | 1.023 | -.045 | -.005 |
| ชุดที่ 4 | 1,000 | -3.386 | 3.065 | .010 | 1.003 | -.059 | .255 |
| ชุดที่ 5 | 1,000 | -2.784 | 3.080 | .004 | 0.938 | -.047 | -.133 |
| ชุดที่ 6 | 1,000 | -2.841 | 3.247 | .034 | 1.000 | -.075 | -.115 |

ตารางที่ 4-1 (ต่อ)

| ค่าความสามารถที่แท้จริง ($\theta$) | *n* | *Min* | *Max* | *Mean* | *SD* | *Skewness* | *Kurtosis* |
|---|---|---|---|---|---|---|---|
| ชุดที่ 7 | 1,000 | -3.598 | 3.418 | -.002 | 1.004 | .019 | .237 |
| ชุดที่ 8 | 1,000 | -3.249 | 3.098 | .013 | 1.016 | .080 | -.004 |
| ชุดที่ 9 | 1,000 | -3.222 | 3.509 | .004 | 1.019 | .008 | -.061 |
| ชุดที่ 10 | 1,000 | -3.530 | 3.617 | .037 | 1.035 | .106 | .024 |

จากตารางที่ 4-1 แสดงค่าสถิติพื้นฐานของค่าความสามารถที่แท้จริงของผู้สอบ ($\theta$) ซึ่งถูกจำลองขึ้นด้วยโปรแกรม WinGen3 จำนวน 10,000 ค่า แบ่งเป็นชุด ชุดละ 1,000 ค่า ปรากฏว่า ค่าความสามารถที่แท้จริงของผู้สอบที่จำลองขึ้นทั้ง 10 ชุด มีลักษณะเป็นโค้งระฆังคว่ำ (การแจกแจงแบบปกติ) มีค่าเฉลี่ยเข้าใกล้ 0 ค่าส่วนเบี่ยงเบนมาตรฐานเข้าใกล้ 1 ค่าความเบ้ (*Skewness*) และค่าความโด่ง (*Kurtosis*) เข้าใกล้ 0 (รายละเอียดค่าความสามารถที่แท้จริงของผู้สอบชุดที่ 1 แสดงในภาคผนวก ข)

คลังข้อสอบจำลอง ถูกจำลองขึ้นจากการสุ่มตัวเลขด้วยโปรแกรม WinGen3 เพื่อจำลองค่าพารามิเตอร์ของข้อสอบ จำนวน 500 ข้อ แต่ละข้อประกอบด้วยค่าพารามิเตอร์ 3 ค่า ได้แก่ 1) อำนาจจำแนกของข้อสอบ 2) ค่าความยากของข้อสอบ และ 3) ค่าการเดาของข้อสอบ แสดงรายละเอียดดังตารางที่ 4-2

ตารางที่ 4-2 ค่าสถิติพื้นฐานของค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบจำลอง

| ค่าพารามิเตอร์ของข้อสอบ | *n* | *Min* | *Max* | *Mean* | *SD* |
|---|---|---|---|---|---|
| ค่าอำนาจจำแนกของข้อสอบ (a) | 500 | 0.502 | 2.497 | 1.479 | 0.575 |
| ค่าความยากของข้อสอบ (b) | 500 | -2.477 | 2.499 | -0.101 | 1.455 |
| ค่าการเดาของข้อสอบ (c) | 500 | 0.001 | 0.300 | 0.149 | 0.086 |

จากตารางที่ 4-2 แสดงค่าสถิติพื้นฐานของค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบจำลองที่จำลองขึ้นจากโปรแกรม WinGen3 จำนวน 500 ข้อ แต่ละข้อประกอบด้วยค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) และค่าการเดาของข้อสอบ (c) มีความสอดคล้องกับข้อกำหนดในการจำลองค่าพารามิเตอร์ของข้อสอบ คือ ค่าอำนาจจำแนกของข้อสอบมีค่าอยู่ในช่วง 0.50 ถึง 2.50 ค่าความยากของข้อสอบมีค่าอยู่ในช่วง -2.50 ถึง 2.50 และค่าการเดาของข้อสอบมีค่าไม่เกิน 0.30 (รายละเอียดคลังข้อสอบจำลอง แสดงในภาคผนวก ค)

ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป แบ่งได้ 3 ประเด็น ดังนี้

**1. การเปรียบเทียบประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบ**

การเปรียบเทียบประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบ เป็นการเปรียบเทียบความแม่นยำในการประมาณความสามารถของผู้สอบ ($\hat{\theta}$) ที่ได้รับจากการศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีวิธีการคัดเลือกข้อสอบข้อถัดไปแตกต่างกัน 4 วิธี ได้แก่ วิธีการ MIC, วิธีการ RDM, วิธีการ HC และวิธีการ HC-Ex ซึ่งทุกวิธีการถูกทำซ้ำ 10 รอบ โดยใช้คลังข้อสอบจำลองขนาด 500 ข้อ ชุดเดียวกัน แต่ใช้ค่าความสามารถที่แท้จริงของผู้สอบทีละชุด ชุดละ 1,000 ค่า ผลการศึกษาในแต่ละรอบ ทุกวิธีการคัดเลือกข้อสอบข้อถัดไปจะได้ค่าประมาณความสามารถของผู้สอบ 1,000 ค่า แล้วนำมาคำนวณค่ารากที่สองของคลาดเคลื่อนกำลังสองเฉลี่ย (*RMSE*) และค่าความลำเอียงเฉลี่ย (*Average Bias*) แสดงดังตารางที่ 4-3 (รายละเอียดค่าประมาณความสามารถของผู้สอบ การคำนวณค่า RMSE และค่าความลำเอียงเฉลี่ยแสดงในภาคผนวก ข)

ตารางที่ 4-3 ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความลำเอียงเฉลี่ยของวิธีการคัดเลือกข้อสอบข้อถัดไป

| รอบที่ | *n* | MIC | | RDM | | HC | | HC-Ex | |
|---|---|---|---|---|---|---|---|---|---|
| | | *RMSE* | *Average Bias* | *RMSE* | *Average Bias* | *RMSE* | *Average Bias* | *RMSE* | *Average Bias* |
| 1 | 1,000 | 0.583 | 0.055 | 0.312 | -0.007 | 0.330 | -0.002 | 0.348 | 0.001 |
| 2 | 1,000 | 0.573 | 0.081 | 0.301 | -0.005 | 0.340 | 0.006 | 0.362 | 0.005 |
| 3 | 1,000 | 0.564 | 0.053 | 0.307 | 0.018 | 0.346 | 0.011 | 0.344 | 0.001 |
| 4 | 1,000 | 0.576 | 0.044 | 0.315 | 0.008 | 0.329 | 0.002 | 0.348 | -0.005 |
| 5 | 1,000 | 0.504 | 0.033 | 0.284 | 0.002 | 0.317 | 0.012 | 0.327 | 0.006 |
| 6 | 1,000 | 0.526 | 0.031 | 0.302 | -0.001 | 0.331 | 0.007 | 0.339 | 0.014 |
| 7 | 1,000 | 0.560 | 0.059 | 0.329 | -0.002 | 0.322 | -0.012 | 0.335 | 0.001 |
| 8 | 1,000 | 0.620 | 0.072 | 0.298 | -0.004 | 0.323 | 0.008 | 0.354 | -0.001 |
| 9 | 1,000 | 0.555 | 0.088 | 0.307 | -0.006 | 0.341 | 0.009 | 0.322 | -0.006 |
| 10 | 1,000 | 0.580 | 0.061 | 0.305 | 0.002 | 0.327 | -0.005 | 0.333 | -0.026 |
| *Mean* | | 0.564 | 0.058 | 0.306 | 0.000 | 0.331 | 0.004 | 0.341 | -0.001 |
| *SD* | | 0.030 | 0.018 | 0.011 | 0.007 | 0.009 | 0.007 | 0.012 | 0.010 |

จากตารางที่ 4-3 แสดงค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความลำเอียงเฉลี่ยของวิธีการคัดเลือกข้อสอบข้อถัดไป ปรากฏว่า วิธีการ RDM มีค่า *RMSE* น้อยที่สุด (*Mean* = 0.306) รองลงมา ได้แก่ วิธีการ HC, HC-Ex และ MIC ตามลำดับ (*Mean* = 0.331, 0.341 และ 0.564 ตามลำดับ) แสดงให้เห็นว่า วิธีการ RDM มีประสิทธิภาพการประมาณค่าความสามารถของผู้สอบสูงสุด และเมื่อพิจารณาค่าความลำเอียงเฉลี่ย ปรากฏว่า วิธีการที่มีค่าความลำเอียงเฉลี่ยเข้าใกล้ 0 มากที่สุด คือ วิธีการ RMD (*Mean* = 0.000) รองลงมา ได้แก่ วิธีการ HC-Ex, HC และ MIC ตามลำดับ (*Mean* = -0.001, 0.004 และ 0.058 ตามลำดับ) แสดงให้เห็นว่า วิธีการ RDM มีความแม่นยำในการประมาณค่าความสามารถของผู้สอบสูงสุด ส่วนวิธีการ HC และวิธีการ MIC มี

ค่าประมาณความสามารถของผู้สอบสูงกว่าค่าความสามารถที่แท้จริงของผู้สอบเล็กน้อย (มีค่าบวก) แต่วิธีการ HC-Ex มีค่าประมาณความสามารถของผู้สอบต่ำกว่าค่าความสามารถที่แท้จริงของผู้สอบเล็กน้อย (มีค่าลบ) และแสดงผลการทดสอบสมมติฐานทางสถิติได้ดังตารางที่ 4-4

ตารางที่ 4-4 ผลการวิเคราะห์ความแตกต่างของค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยและค่าความลำเอียงเฉลี่ย จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป

| | *Asymp F* | *df1* | *df2* | *p-value* |
|---|---|---|---|---|
| *RMSE* | 414.261 | 3 | 16.097 | .000** |
| *Average Bias* | 54.655 | 3 | 21.212 | .000** |

** $p < .01$

จากตารางที่ 4-4 แสดงผลการวิเคราะห์ความแตกต่างของค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (*RMSE*) และค่าความลำเอียงเฉลี่ย (*Average Bias*) จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป 4 วิธี ด้วยการวิเคราะห์ Brown-Forsythe เมื่อพิจารณาจากค่า *RMSE* ปรากฏว่า มีวิธีการคัดเลือกข้อสอบข้อถัดไปอย่างน้อย 1 คู่ ที่มีค่า *RMSE* แตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 และเมื่อพิจารณาค่าความลำเอียงเฉลี่ย ปรากฏว่า มีวิธีการคัดเลือกข้อสอบข้อถัดไปอย่างน้อย 1 คู่ ที่มีค่าความลำเอียงเฉลี่ยแตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 ดังนั้น ผู้วิจัยจึงเปรียบเทียบพหุคูณของค่า *RMSE* และค่าความลำเอียงเฉลี่ย ด้วยการทดสอบของแทมเฮน แสดงดังตารางที่ 4-5 และ 4-6 ตามลำดับ

ตารางที่ 4-5 ผลการเปรียบเทียบพหุคูณของค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยจำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป

| วิธีการ | *Mean* | ผลต่างเฉลี่ย (*Mean Difference*) | | | |
|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex |
| MIC | .564 | - | .258** | .233** | .223** |
| RDM | .306 | | - | -.025** | -.035** |
| HC | .331 | | | - | -.010 |
| HC-Ex | .341 | | | | - |

** $p < .01$

จากตารางที่ 4-5 แสดงผลการเปรียบเทียบพหุคูณของค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (*RMSE*) จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป 4 วิธี ปรากฏว่า วิธีการที่มีค่า *RMSE* แตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 มีทั้งหมด 5 คู่ ได้แก่ วิธีการ MIC กับ RDM, วิธีการ MIC กับ HC, วิธีการ MIC กับ HC-Ex, วิธีการ RDM กับ HC และวิธีการ RDM กับ HC-Ex ส่วนวิธีการ HC กับ HC-Ex มีค่า RMSE ไม่แตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 เมื่อพิจารณาจากค่าเฉลี่ยของ *RMSE* จึงสรุปได้ว่า ค่า *RMSE* ของวิธีการ RDM มีค่าต่ำสุด รองลงมา ได้แก่ วิธีการ HC ซึ่งมีค่า

ไม่แตกต่างกับวิธีการ HC-Ex ส่วนวิธี MIC มีค่า RMSE สูงสุด แสดงให้เห็นว่าวิธีการ HC มีค่า *RMSE* น้อยกว่าวิธีการ MIC ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อ 1.1 แต่วิธีการ HC มีค่า *RMSE* ไม่น้อยกว่าวิธีการ MIC และ HC-Ex ซึ่งไม่เป็นไปตามสมมติฐานของการวิจัยข้อ 1.1 ดังนั้น ผลการศึกษาจึงเป็นไปตามสมมติฐานของการวิจัยข้อ 1.1 บางส่วน

ตารางที่ 4-6 ผลการเปรียบเทียบพหุคูณของค่าความลำเอียงเฉลี่ย จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป

| วิธีการ | *Mean* | ผลต่างเฉลี่ย (*Mean Difference*) | | | |
|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex |
| MIC | .058 | - | .058** | .054** | .059** |
| RDM | .000 | | - | -.004 | .001 |
| HC | .004 | | | - | .005 |
| HC-Ex | -.001 | | | | - |

** $p < .01$

จากตารางที่ 4-6 แสดงผลการเปรียบเทียบพหุคูณของค่าความลำเอียงเฉลี่ย จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป ปรากฏว่า วิธีการที่มีค่าความลำเอียงเฉลี่ยแตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 มีทั้งหมด 3 คู่ ได้แก่ วิธีการ MIC กับ RDM, วิธีการ MIC กับ HC และวิธีการ MIC กับ HC-Ex วิธีการที่มีค่าความลำเอียงเฉลี่ยไม่แตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 มี 3 คู่ ได้แก่ วิธีการ RDM กับ HC, วิธีการ RDM กับ HC-Ex และวิธีการ HC กับ HC-Ex เมื่อพิจารณาจากค่าเฉลี่ยของค่าความลำเอียงเฉลี่ย สรุปได้ว่า วิธีการ RDM, HC และ HC-Ex มีค่าความลำเอียงเฉลี่ยไม่แตกต่างกัน และมีค่าต่ำกว่าวิธีการ MIC แสดงให้เห็นว่า วิธีการ HC มีค่าความลำเอียงเฉลี่ยน้อยกว่าวิธีการ MIC ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อ 1.2 แต่วิธีการ HC มีค่าความลำเอียงเฉลี่ยเทียบเท่ากับวิธีการ RDM และ HC-Ex ซึ่งไม่เป็นไปตามสมมติฐานของการวิจัยข้อ 1.2 ดังนั้น ผลการศึกษาจึงเป็นไปตามสมมติฐานของการวิจัยข้อ 1.2 บางส่วน

**2. การเปรียบเทียบประสิทธิภาพด้านความยาวของแบบทดสอบ**

การเปรียบเทียบประสิทธิภาพด้านความยาวของแบบทดสอบ โดยพิจารณาจากความยาวของแบบทดสอบ จากวิธีการคัดเลือกข้อสอบข้อถัดไปที่แตกต่างกัน 4 วิธี ได้แก่ วิธีการ MIC, วิธีการ RDM, วิธีการ HC และวิธีการ HC-Ex โดยศึกษาในสถานการณ์จำลอง ซึ่งทุกวิธีการถูกทำซ้ำ 10 รอบ โดยใช้คลังข้อสอบจำลองขนาด 500 ข้อ ชุดเดียวกัน แต่ใช้ค่าความสามารถที่แท้จริงของผู้สอบทีละชุด ชุดละ 1,000 ค่า ผลการศึกษาในแต่ละรอบ ทุกวิธีการคัดเลือกข้อสอบข้อถัดไป จะได้ค่าความยาวของแบบทดสอบของแต่ละค่า $\theta$ รวม 1,000 ค่า แสดงได้ดังตารางที่ 4-7 (รายละเอียดค่าความยาวของแบบทดสอบ แสดงในภาคผนวก ข)

ตารางที่ 4-7 ค่าสถิติพื้นฐานความยาวของแบบทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป

| รอบที่ | *n* | *Median* | | | |
|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex |
| 1 | 1,000 | 11 | 7 | 7 | 9 |
| 2 | 1,000 | 11 | 7 | 7 | 9 |
| 3 | 1,000 | 11 | 7 | 7 | 9 |
| 4 | 1,000 | 11 | 7 | 7 | 9 |
| 5 | 1,000 | 11 | 7 | 7 | 9 |
| 6 | 1,000 | 11 | 7 | 7 | 9 |
| 7 | 1,000 | 11 | 7 | 7 | 9 |
| 8 | 1,000 | 11 | 7 | 7 | 9 |
| 9 | 1,000 | 11 | 7 | 7 | 10 |
| 10 | 1,000 | 11 | 7 | 7 | 9 |
| โดยรวม | 10,000 | 11 | 7 | 7 | 9 |

จากตารางที่ 4-7 แสดงค่าสถิติพื้นฐานของความยาวของแบบทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป ปรากฏว่า วิธีการ RDM และ HC มีความยาวของแบบทดสอบน้อยสุดเทียบเท่ากัน (*Median* = 7) รองลงมา ได้แก่ วิธีการHC-Ex (*Median* = 9) และวิธีการ MIC (*Median* = 11) ตามลำดับ และแสดงผลการทดสอบสมมติฐานทางสถิติได้ดังตารางที่ 4-8

ตารางที่ 4-8 ผลการวิเคราะห์ความแตกต่างของความยาวของแบบทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป

| วิธีการ | *Median* | $\chi^2$ | *df* | *p-value* |
|---|---|---|---|---|
| MIC | 11 | 13,398.43 | 3 | .000** |
| RDM | 7 | | | |
| HC | 7 | | | |
| HC-Ex | 9 | | | |

***p < .01*

จากตารางที่ 4-8 แสดงผลการวิเคราะห์ความแตกต่างของความยาวของแบบทดสอบจำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป 4 วิธี ด้วยการทดสอบ Kruskal-Wallis ปรากฏว่า มีวิธีการคัดเลือกข้อสอบข้อถัดไปอย่างน้อย 1 คู่ ที่มีความยาวของแบบทดสอบแตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 ดังนั้น จึงเปรียบเทียบค่าเฉลี่ยของอันดับเป็นรายคู่ ด้วยการทดสอบ Dunn-Bonferroni แสดงดังตารางที่ 4-9

ตารางที่ 4-9 ผลการเปรียบเทียบค่าเฉลี่ยของอันดับของความยาวของแบบทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไปเป็นรายคู่

| วิธีการ | ค่าเฉลี่ยของอันดับ | ผลต่างค่าเฉลี่ยของอันดับ | | | |
|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex |
| MIC | 28,734.58 | - | 15,209.19** | 15,137.87** | 4,589.26** |
| RDM | 13,525.39 | | - | -71.32 | -10,619.93** |
| HC | 13,596.71 | | | - | -10,548.61** |
| HC-Ex | 24,145.32 | | | | - |

***p< .01*

จากตารางที่ 4-9 แสดงผลการเปรียบเทียบค่าเฉลี่ยของอันดับของความยาวของแบบทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไปเป็นรายคู่ โดยนำความยาวของแบบทดสอบทั้งหมด 40,000 ค่า (ได้จาก 4 วิธีการ × 10 รอบ × 1,000 ค่า) มาจัดเรียงอันดับจากน้อยไปมาก แล้วนำอันดับมาคำนวณค่าเฉลี่ย และทดสอบเปรียบเทียบค่าเฉลี่ยของอันดับเป็นรายคู่ ด้วยการทดสอบ Dunn-Bonferroni ปรากฏว่า วิธีการที่มีความยาวของแบบทดสอบแตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 มีทั้งหมด 5 คู่ ได้แก่ วิธีการ MIC กับ RDM, วิธีการ MIC กับ HC, วิธีการ MIC กับ HC-Ex, วิธีการ RDM กับ HC-Ex และวิธีการ HC กับ HC-Ex ส่วนวิธีการ RDM กับ HC มีความยาวของแบบทดสอบไม่แตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 เมื่อพิจารณาจากค่าเฉลี่ยของอันดับ สรุปได้ว่า ค่าเฉลี่ยของอันดับจากวิธีการ RDM มีค่าต่ำสุด ซึ่งไม่แตกต่างกับวิธีการ HC รองลงมา ได้แก่ วิธีการ HC-Ex และ MIC ตามลำดับ แสดงให้เห็นว่า วิธีการ HC มีความยาวของแบบทดสอบน้อยกว่าวิธีการ HC-Ex และ MIC ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อ 2 แต่วิธีการ HC มีความยาวของแบบทดสอบเทียบเท่ากับวิธีการ RDM ซึ่งไม่เป็นไปตามสมมติฐานของการวิจัยข้อที่ 2 ดังนั้น ผลการศึกษาจึงเป็นไปตามสมมติฐานของการวิจัยข้อที่ 2 บางส่วน

**3. การเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2**

การเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 โดยพิจารณาจากจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จากวิธีการคัดเลือกข้อสอบข้อถัดไปที่แตกต่างกัน 4 วิธี ได้แก่ วิธีการ MIC, วิธีการ RDM, วิธีการ HC และวิธีการ HC-Ex โดยศึกษาในสถานการณ์จำลอง ซึ่งทุกวิธีการถูกทำซ้ำ 10 รอบ โดยใช้คลังข้อสอบจำลองขนาด 500 ข้อ ชุดเดียวกัน แต่ใช้ค่าความสามารถที่แท้จริงของผู้สอบที่ละชุด ชุดละ 1,000 ค่า ผลจากการศึกษาในแต่ละรอบ ทุกวิธีการคัดเลือกข้อสอบข้อถัดไป จะได้จำนวนการใช้ข้อสอบ (ครั้ง) ของข้อสอบแต่ละข้อในคลังข้อสอบจำลอง แล้วนำมาคำนวณอัตราการใช้ข้อสอบ ซึ่งจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 แสดงได้ดังตารางที่ 4-10 (รายละเอียดจำนวนการใช้ข้อสอบ การคำนวณอัตราการใช้ข้อสอบ แสดงในภาคผนวก ค)

ตารางที่ 4-10 จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป

| รอบที่ | MIC | RDM | HC | HC-Ex |
|---|---|---|---|---|
| 1 | 0 | 10 | 11 | 3 |
| 2 | 0 | 11 | 10 | 1 |
| 3 | 0 | 10 | 10 | 1 |
| 4 | 0 | 10 | 10 | 3 |
| 5 | 1 | 10 | 11 | 2 |
| 6 | 0 | 11 | 9 | 3 |
| 7 | 0 | 9 | 11 | 1 |
| 8 | 0 | 9 | 10 | 2 |
| 9 | 0 | 10 | 10 | 2 |
| 10 | 0 | 0 | 10 | 2 |
| *Median* | 0 | 10 | 10 | 2 |

จากตารางที่ 4-10 แสดงจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป ปรากฏว่า วิธีการ MIC มีข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 น้อยสุด (*Median* = 0) รองลงมา ได้แก่ วิธีการ HC-Ex (*Median* = 2) ส่วนวิธีการ RDM และ HC มีข้อสอบที่มีอัตราการใช้ข้อสอบสูงสุดเท่ากัน (*Median* = 10) และแสดงผลการทดสอบสมมติฐานทางสถิติได้ดังตารางที่ 4-11

ตารางที่ 4-11 ผลการวิเคราะห์ความแตกต่างของจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป

| วิธีการ | *Median* | $\chi^2$ | *df* | *p-value* |
|---|---|---|---|---|
| MIC | 0 | 30.486 | 3 | .000** |
| RDM | 10 | | | |
| HC | 10 | | | |
| HC-Ex | 2 | | | |

*** p < .01*

จากตารางที่ 4-11 แสดงผลการวิเคราะห์ความแตกต่างของจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป 4 วิธี ด้วยการทดสอบ Kruskal-Wallis ปรากฏว่า มีวิธีการคัดเลือกข้อสอบข้อถัดไปอย่างน้อย 1 คู่ ที่มีจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 แตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 ดังนั้น จึงเปรียบเทียบค่าเฉลี่ยของอันดับเป็นรายคู่ของจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ด้วยการทดสอบ Dunn-Bonferroni แสดงดังตารางที่ 4-12

ตารางที่ 4-12 ผลการเปรียบเทียบค่าเฉลี่ยของอันดับของจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไปเป็นรายคู่

| วิธีการ | ค่าเฉลี่ยของอันดับ | ผลต่างค่าเฉลี่ยของอันดับ | | | |
|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex |
| MIC | 6.20 | - | -21.55** | -25.50** | -10.15** |
| RDM | 27.75 | | - | -3.95 | 11.40** |
| HC | 31.70 | | | - | 15.35** |
| HC-Ex | 16.35 | | | | - |

***p* < .01*

จากตารางที่ 4-12 แสดงผลการเปรียบเทียบค่าเฉลี่ยของอันดับของจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไปเป็นรายคู่ โดยมีข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำนวน 40 ค่า (ได้จาก 4 วิธีการ × 10 รอบ) มาจัดเรียงอันดับจากน้อยไปมาก แล้วนำอันดับมาคำนวณค่าเฉลี่ย และทดสอบเปรียบเทียบค่าเฉลี่ยของอันดับเป็นรายคู่ด้วยการทดสอบ Dunn-Bonferroni ปรากฏว่า วิธีการที่มีจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 แตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 มีทั้งหมด 5 คู่ ได้แก่ วิธีการ MIC กับ RDM, วิธีการ MIC กับ HC, วิธีการ MIC กับ HC-Ex, วิธีการ RDM กับ HC-Ex และวิธีการ HC กับ HC-Ex ส่วนวิธีการ RDM กับ HC มีจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบไม่แตกต่างกัน ที่ระดับนัยสำคัญทางสถิติ .01 เมื่อพิจารณาจากค่าเฉลี่ยของอันดับ สรุปได้ว่าค่าเฉลี่ยของอันดับจากวิธีการ MIC มีค่าต่ำสุด รองลงมา ได้แก่ วิธีการ HC-Ex ส่วนวิธีการ RDM มีค่าสูงสุด ซึ่งไม่แตกต่างกับวิธีการ HC แสดงให้เห็นว่า วิธีการ HC-Ex มีข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำนวนน้อยกว่าวิธีการ HC และ RDM ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 3 แต่วิธีการ HC-Ex มีข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 จำนวนไม่น้อยกว่าวิธีการ MIC ซึ่งไม่เป็นไปตามสมมติฐานของการวิจัยข้อ 3 ดังนั้น ผลการศึกษาจึงเป็นไปตามสมมติฐานของการวิจัยข้อที่ 3 บางส่วน

ตารางที่ 4-13 สรุปผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

| ประสิทธิภาพ | MIC | RDM | HC | HC-Ex |
|---|---|---|---|---|
| ด้านการประมาณค่าความสามารถของผู้สอบ | | | | |
| - พิจารณาจากค่า RMSE | อันดับ4 | อันดับ1 | อันดับ2 | อันดับ2 |
| - พิจารณาจากค่าความลำเอียงเฉลี่ย | อันดับ4 | อันดับ1 | อันดับ1 | อันดับ1 |
| ด้านความยาวของแบบทดสอบ | อันดับ4 | อันดับ1 | อันดับ1 | อันดับ3 |
| ด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 | อันดับ1 | อันดับ3 | อันดับ3 | อันดับ2 |

จากผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สรุปได้ว่า ประสิทธิภาพด้านการประมาณความสามารถของผู้สอบ เมื่อพิจารณาจากค่า RMSE ปรากฏว่า วิธีการ RDM มีประสิทธิภาพสูงสุด รองลงมา ได้แก่ วิธีการ HC ซึ่งมีประสิทธิภาพเทียบเท่ากับวิธีการ HC-Ex ส่วนวิธีการ MIC มีประสิทธิภาพต่ำสุด แต่เมื่อพิจารณาจากค่าความลำเอียงเฉลี่ย ปรากฏว่า วิธีการ RDM, HC และ HC-Ex มีประสิทธิภาพสูงสุดเทียบเท่ากัน ส่วนวิธีการ MIC มีประสิทธิภาพต่ำสุด ส่วนประสิทธิภาพด้านความยาวของแบบทดสอบ ปรากฏว่า วิธีการ RDM และ HC มีประสิทธิภาพสูงสุดเทียบเท่ากัน รองลงมา ได้แก่ วิธีการ HC-Ex และวิธีการ MIC ตามลำดับ ส่วนประสิทธิภาพด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ปรากฏว่า วิธีการ MIC มีประสิทธิภาพสูงสุด รองลงมา ได้แก่ วิธีการ HC-Ex ส่วนวิธีการ RDM และ HC มีประสิทธิภาพต่ำสุดเทียบเท่ากัน

ผลการศึกษาแสดงให้เห็นว่า วิธีการ HC ซึ่งเป็นวิธีการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้น มีประสิทธิภาพสูงในด้านการประมาณค่าความสามารถของผู้สอบ และด้านความยาวของแบบทดสอบ ถึงแม้จะมีประสิทธิภาพต่ำกว่าวิธีการ RDM แต่มีประสิทธิภาพสูงกว่าวิธีการ MIC ซึ่งเป็นวิธีการที่นิยมใช้ทั่วไป ส่วนวิธีการ HC-Ex ซึ่งเป็นวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีการควบคุมการใช้ข้อสอบ ถึงแม้จะมีประสิทธิภาพต่ำกว่าวิธีการ MIC แต่สามารถควบคุมการใช้ข้อสอบได้ค่อนข้างมีประสิทธิภาพ เนื่องจากข้อสอบในคลังข้อสอบขนาด 500 ข้อ มีเพียง 2 ข้อเท่านั้น ที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ดังนั้น วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้น จึงเป็นอีกทางเลือกหนึ่งสำหรับนำไปใช้คัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

## ตอนที่ 3 ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในการวิจัยนี้ เป็นการนำวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ มาพัฒนาโปรแกรมการทดสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี ในรูปแบบของ Web Application โดยโปรแกรมการทดสอบนี้ได้ออกแบบให้สามารถเลือกการทดสอบได้ 2 แบบ คือ 1) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และ 2) การทดสอบด้วยคอมพิวเตอร์ ซึ่งสามารถแบ่งการนำเสนอผลการพัฒนาโปรแกรมได้ 2 ประเด็น ดังนี้

1. ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
   1.1 การจัดการข้อสอบ
   1.2 การจัดการสอบ
2. ผลการประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
   2.1 ผลการประเมินโดยผู้เชี่ยวชาญ
   2.2 ผลการประเมินโดยผู้สอบ

ในแต่ละประเด็นมีรายละเอียดดังนี้

**1. ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นี้ เป็นการทดสอบเนื้อหาส่วนหนึ่งของรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ในระดับปริญญาตรี ซึ่งครอบคลุมเนื้อหา 4 ด้าน คือ 1) ชนิดของคำ (Part of Speech) 2) การใช้บริบทบ่งชี้เพื่อค้นหาความหมายของคำศัพท์ยาก (Using Context Clue) 3) การเรียนรู้ความหมายคำศัพท์จากโครงสร้างของคำศัพท์ (Word Formations) และ 4) การสรุปความ (Making Inferences) ซึ่งผู้ใช้งานเข้าใช้โปรแกรมได้จาก catctexam.com สามารถแสดงหน้าหลักของโปรแกรมได้ดังภาพที่ 4-1

ภาพที่ 4-1 หน้าหลักของโปรแกรมการทดสอบ

จากภาพที่ 4-1 แสดงหน้าหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งในหน้าหลักของโปรแกรม ประกอบด้วยเมนูหลัก 4 เมนู ดังนี้

1) หน้าหลัก เป็นเมนูแสดงหน้าแรกของโปรแกรมการทดสอบ

2) การทดสอบ เป็นเมนูสำหรับผู้สอบ เพื่อทดสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ในระดับปริญญาตรี

3) ติดต่อเรา เป็นเมนูแสดงข้อมูลของวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา และข้อมูลของผู้พัฒนาโปรแกรม

4) เข้าสู่ระบบ เป็นเมนูสำหรับครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการข้อสอบ แสดงการจัดการข้อสอบ รวมทั้งรายงานผลการทดสอบของผู้สอบทั้งหมด

รายละเอียดของโปรแกรมการทดสอบนี้ แบ่งได้ 2 ส่วน ดังนี้

**1.1 การจัดการข้อสอบ**

การจัดการข้อสอบเป็นส่วนที่ให้ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการข้อสอบ ใส่ข้อสอบที่ผ่านการคัดเลือกเป็นไปตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ลงในโปรแกรมเพื่อเป็นคลังข้อสอบ หรือทำการลบ แก้ไขข้อสอบเดิมที่มีอยู่ในคลังข้อสอบ หลังจากผู้ใช้เลือก “เข้าสู่ระบบ” ในหน้าหลัก โปรแกรมจะเข้าสู่หน้าจอแรกของส่วนการจัดการข้อสอบ แสดงได้ดังภาพที่ 4-2

คลังข้อสอบที่ใช้ในโปรแกรมนี้ เป็นข้อสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ซึ่งได้รับความอนุเคราะห์ข้อสอบทั้งหมด จากอาจารย์ประจำหลักสูตรภาษาอังกฤษ มหาวิทยาลัยสวนดุสิต โดยข้อสอบทั้งหมดมีจำนวน 421 ข้อ ประกอบด้วย โจทย์ รายการคำตอบแบบ 4 ตัวเลือก และค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ ซึ่งแบ่งข้อสอบได้ 4 ชุด ได้แก่ ชุดที่ 1 ชนิดของคำ (Part of Speech) จำนวน 118 ข้อ ชุดที่ 2 การใช้บริบทบ่งชี้เพื่อค้นหาความหมายของคำศัพท์ยาก (Using Context Clue) จำนวน 104 ข้อ ชุดที่ 3 การเรียนรู้ความหมายคำศัพท์จากโครงสร้างของคำศัพท์ (Word Formations) จำนวน 107 ข้อ และ ชุดที่ 4 การสรุปความ (Making Inferences) จำนวน 92 ข้อ

ภาพที่ 4-2 หน้าจอ Login เข้าโปรแกรมในส่วนการจัดการข้อสอบ

จากภาพที่ 4-2 แสดงหน้าจอ Login เข้าส่วนการจัดการข้อสอบ โดยผู้วิจัยกำหนดรหัสผ่านดังนี้

User Name: admin
Password: @dmin

เมื่อผู้ใช้กรอกรหัสผ่านเรียบร้อยแล้ว หน้าจอจะเข้าสู่ส่วนการจัดการข้อสอบ ประกอบด้วยเมนูย่อย ได้แก่ 1) ตั้งรหัสกลุ่มผู้สอบ 2) การจัดการข้อสอบ 3) รายงานผลสอบ และ 4) ออกจากระบบ โดยแต่ละเมนูมีรายละเอียดดังนี้

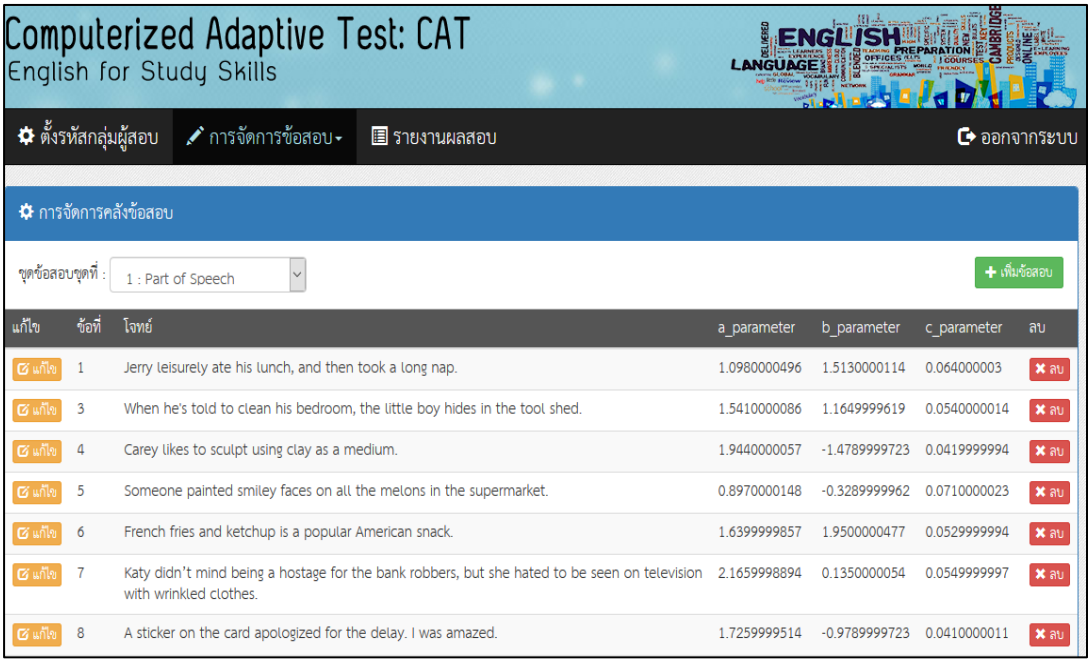
**1) ตั้งรหัสกลุ่มผู้สอบ** สำหรับให้ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการข้อสอบ ทำการตั้งรหัสกลุ่มผู้สอบใหม่ เพื่อป้องกันไม่ให้บุคคลภายนอกเข้ามาใช้โปรแกรมได้ หรือในกรณีที่มี

ผู้สอบหลายกลุ่ม ในแต่ละกลุ่มใช้ข้อสอบต่างชุดกัน จึงจำเป็นต้องตั้งรหัสกลุ่มผู้สอบใหม่ เพื่อป้องกันผู้สอบเข้าสอบผิดกลุ่ม เมื่อผู้ใช้ตั้งรหัสใหม่เรียบร้อยแล้วให้กดปุ่ม “บันทึก” ซึ่งแสดงหน้าจอได้ดังภาพที่ 4-3

ภาพที่ 4-3 หน้าจอการตั้งรหัสกลุ่มผู้สอบ

**2) การจัดการข้อสอบ** สำหรับให้ครู อาจารย์ หรือผู้เกี่ยวข้อง ทำการจัดการข้อสอบ ซึ่งมีเมนูย่อย 2 ส่วน ได้แก่

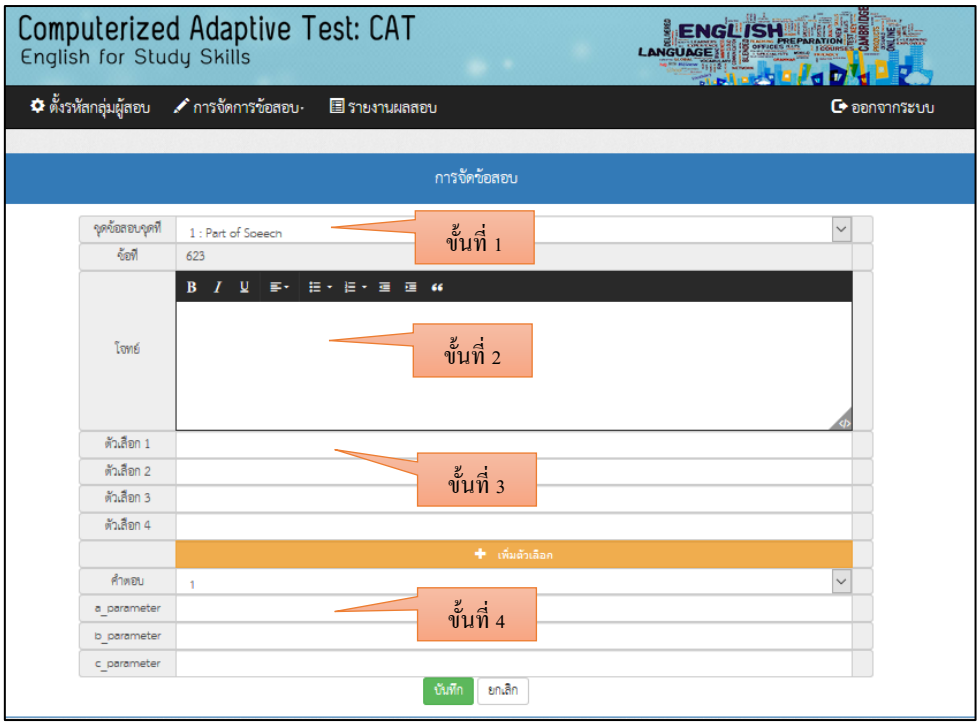
**2.1) การจัดการคลังข้อสอบ** เป็นการเพิ่ม ลบ และแก้ไขข้อสอบที่อยู่ในคลังข้อสอบ ซึ่งหน้าจอการจัดการคลังข้อสอบ แสดงได้ดังภาพที่ 4-4




ภาพที่ 4-4 หน้าจอการจัดการคลังข้อสอบ

**2.1.1 การเพิ่มข้อสอบ**

การเพิ่มข้อสอบลงในโปรแกรม โดยให้ผู้ใช้กดปุ่ม “เพิ่มข้อสอบ” แล้วโปรแกรมจะเข้าสู่หน้าจอดังภาพที่ 4-5




ภาพที่ 4-5 หน้าจอการเพิ่มข้อสอบ

จากภาพที่ 4-5 แสดงหน้าจอการเพิ่มข้อสอบลงในโปรแกรม โดยข้อสอบที่ใช้ในโปรแกรมนี้ต้องเป็นข้อสอบที่มีรายการคำตอบอย่างน้อย 4 ตัวเลือก และผ่านการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ จึงมีค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ (ผู้สนใจสามารถคำนวณค่าพารามิเตอร์ของข้อสอบได้จากโปรแกรม Xcalibre) มีขั้นตอนการเพิ่มข้อสอบดังนี้

**ขั้นที่ 1 เลือกชุดข้อสอบ**ที่ต้องการเพิ่มข้อสอบเข้าไป

**ขั้นที่ 2 กรอกรายละเอียดโจทย์**

**ขั้นที่ 3 กรอกรายละเอียดรายการคำตอบ**ที่ละรายการ ในกรณีมีรายการคำตอบมากกว่า 4 รายการให้กดปุ่ม “เพิ่มตัวเลือก” พร้อมระบุรายการคำตอบที่เป็นคำตอบ

**ขั้นที่ 4 กรอกรายละเอียดค่าพารามิเตอร์ของข้อสอบ** ได้แก่ ค่าอำนาจจำแนกของข้อสอบ (Discrimination Power Parameter: a_parameter) ต้องมีค่าระหว่าง -2.50 ถึง 2.50 ค่าความยากของข้อสอบ (Difficulty Parameter: b_parameter) ต้องมีค่าตั้งแต่ 0.50 ถึง 2.50 และค่าการเดาของข้อสอบ (Guess Parameter: c_parameter) ต้องมีค่าไม่เกิน 0.30

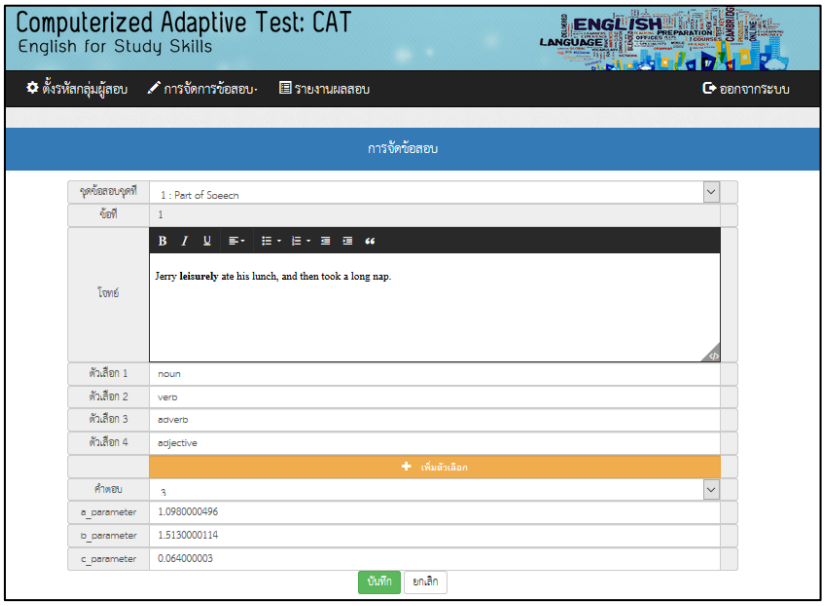
เมื่อกรอกรายละเอียดข้อสอบที่ต้องการเพิ่มเรียบร้อยแล้ว ให้กดปุ่ม “บันทึกข้อมูล” แต่หากไม่ต้องการบันทึก ให้กดปุ่ม “ยกเลิก”

**2.1.2 การลบข้อสอบ**

การลบข้อสอบออกจากคลังข้อสอบที่อยู่ในโปรแกรม ดำเนินการโดยไปหน้าจอการจัดการคลังข้อสอบ ดังภาพที่ 4-4 จากนั้นให้เลือกชุดของข้อสอบที่ต้องการลบข้อสอบออก แล้วทำการกดปุ่ม “ลบ” ซึ่งปรากฏอยู่ด้านท้ายของข้อสอบที่ต้องการลบ

**2.1.3 การแก้ไขข้อสอบ**

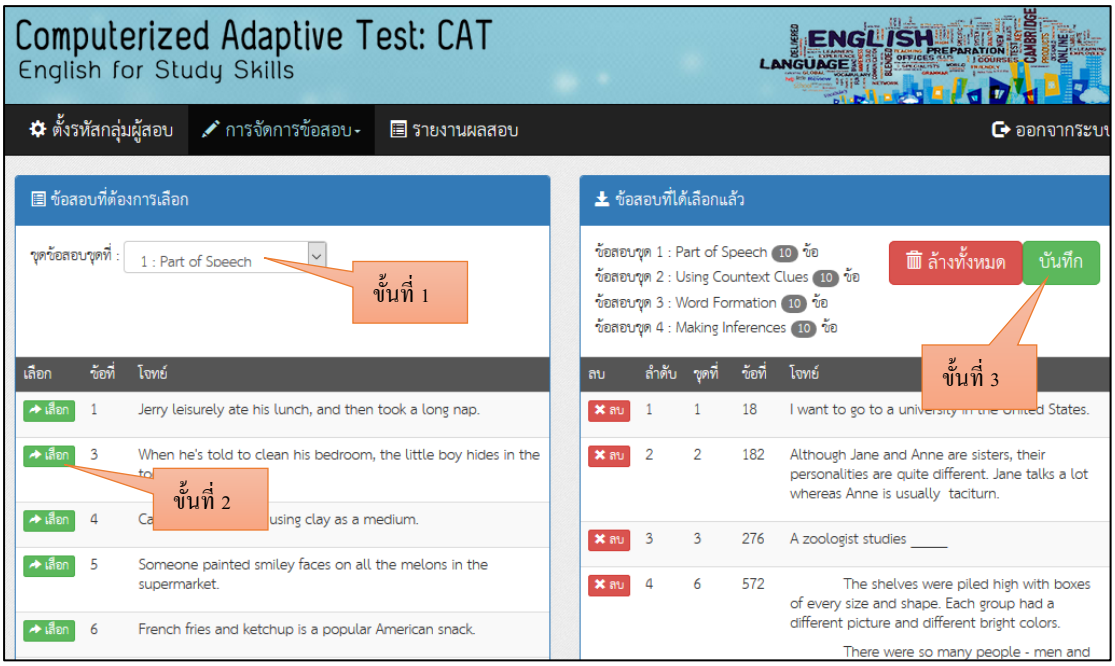
การแก้ไขข้อสอบเดิมในคลังข้อสอบ ดำเนินการโดยไปที่หน้าการจัดการคลังข้อสอบดังภาพที่ 4-4 จากนั้นให้เลือกชุดข้อสอบที่ต้องการแก้ไขข้อสอบ แล้วทำการกดปุ่ม “แก้ไข” ซึ่งปรากฏอยู่ด้านหน้าของข้อสอบที่ต้องการแก้ไข แล้วจะได้หน้าจอ ดังภาพที่ 4-6




ภาพที่ 4-6 หน้าจอแก้ไขข้อสอบ

จากภาพที่ 4-6 แสดงหน้าจอแก้ไขข้อสอบ เมื่อผู้ใช้แก้ไขข้อสอบตามที่ต้องการเรียบร้อยแล้ว ให้กดปุ่ม “บันทึก” หากไม่ต้องการบันทึกให้กดปุ่ม “ยกเลิก”

**2.2) การจัดการข้อสอบแบบ CT** เป็นการจัดชุดข้อสอบในการทดสอบด้วยคอมพิวเตอร์ (Computerized Testing) ซึ่งผู้สอบทุกคนจะได้ข้อสอบชุดเดียวกัน ดำเนินการโดยให้ผู้ใช้ (ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดข้อสอบ) คัดเลือกข้อสอบจำนวน 40 ข้อ จากคลังข้อสอบโดยคัดเลือกข้อสอบชุดละ 10 ข้อ จากข้อสอบทั้งหมด 4 ชุด ประกอบด้วยข้อสอบที่มีค่าความยากของข้อสอบจากน้อยไปมาก ในการจัดการชุดข้อสอบแบบ CT แสดงดังภาพที่ 4-7




ภาพที่ 4-7 หน้าจอการจัดการข้อสอบของการทดสอบด้วยคอมพิวเตอร์

จากภาพที่ 4-7 แสดงหน้าจอการจัดการข้อสอบแบบ CT โดยให้เลือกข้อสอบ จำนวน 40 ข้อ จากคลังข้อสอบ ซึ่งมีขั้นตอนดังนี้

**ขั้นที่ 1 เลือกชุดข้อสอบ**

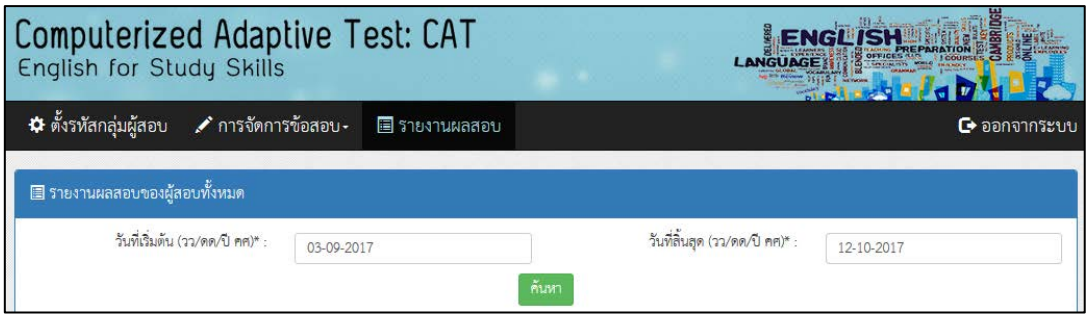
**ขั้นที่ 2 เลือกข้อสอบ** โดยกดปุ่ม “เลือก” ที่อยู่ด้านหน้าของข้อสอบข้อที่ต้องการ โดยหลังจากกดเลือกแล้ว ข้อสอบข้อนั้นจะไปปรากฏด้านล่างสุดในตารางด้านขวา แสดงว่า ข้อสอบข้อนั้นได้ถูกเลือกไปใช้ในการทดสอบแบบ CT แล้ว จากนั้นดำเนินการเลือกข้อสอบในชุดอื่นให้ครบตามที่ต้องการ

ในกรณีที่เลือกข้อสอบผิดข้อ ผู้ใช้สามารถลบโดยการกดปุ่ม “ลบ” หน้าข้อสอบที่ต้องการลบได้ หรือกดปุ่ม “ล้างทั้งหมด” ซึ่งเป็นการลบข้อสอบทุกข้อที่ได้เลือกมาแล้ว

**ขั้นที่ 3 บันทึกข้อมูล** เมื่อเลือกข้อสอบได้ครบตามต้องการแล้ว กดปุ่ม “บันทึก” เพื่อบันทึกข้อสอบทั้งหมดในการทดสอบแบบ CT

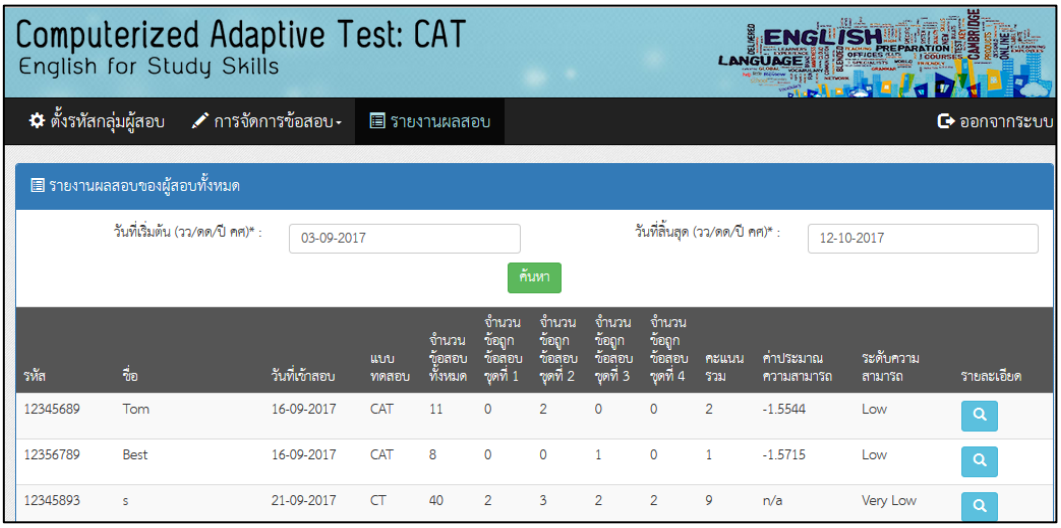
**3) รายงานผลสอบ**

การรายงานผลสอบ เป็นการรายงานผลสอบของผู้สอบทุกคนให้ครู อาจารย์ ทราบเพื่อใช้ในการประเมินผลการเรียนรู้ของผู้สอบ โดยกดปุ่ม “รายงานผลสอบ” จะปรากฏหน้าจอดังภาพที่ 4-8




ภาพที่ 4-8 หน้าจอแรกในส่วนรายงานผลสอบ

จากภาพที่ 4-8 แสดงหน้าจอแรกในส่วนรายงานผลสอบ โดยให้ระบุช่วงวัน แล้วกดปุ่ม “ค้นหา” โปรแกรมจะแสดงผลสอบของผู้สอบทั้งหมดที่สอบภายในช่วงวันที่ระบุ ดังภาพที่ 4-9




ภาพที่ 4-9 หน้าจอรายงานผลสอบ

จากภาพที่ 4-9 แสดงหน้าจอรายงานผลสอบของผู้สอบทั้งหมดที่สอบภายในวันที่ผู้ใช้ระบุ โดยโปรแกรมจะแสดงข้อมูลเป็นรายคน ได้แก่ รหัสนักศึกษา ชื่อ-สกุล วันที่เข้าสอบ คะแนนของผู้สอบในแต่ละหัวข้อเรื่อง คะแนนรวม ค่าประมาณความสามารถ และระดับความสามารถของผู้สอบ

นอกจากนี้ หากต้องการทราบรายละเอียดในการสอบของผู้สอบแต่ละคน ให้กดปุ่ม ในช่องรายละเอียด ซึ่งอยู่ด้านท้ายของผู้สอบที่ต้องการทราบ โปรแกรมจะเชื่อมข้อมูลไปยังรายงานผลการสอบของผู้สอบคนนั้น

**4) ออกจากระบบ**

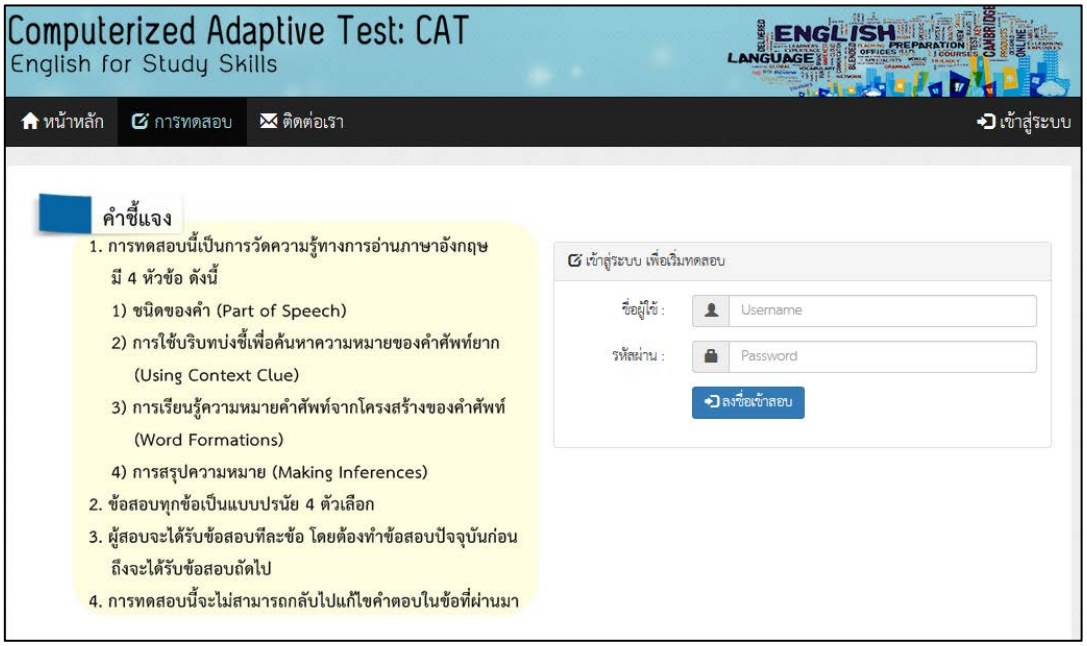
ออกจากระบบ เป็นเมนูสุดท้ายในส่วนการเข้าระบบ เมื่อผู้ใช้กดปุ่ม “ออกจากระบบ” โปรแกรมจะกลับไปสู่หน้าจอหลักของโปรแกรม

**1.2 การจัดการสอบ**

การจัดการสอบนี้ เป็นส่วนที่ให้นิสิต นักศึกษา เข้าใช้โปรแกรมเพื่อทดสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ในระดับปริญญาตรี โดยผู้สอบจะได้ข้อสอบครบทุกหัวข้อเรื่อง ซึ่งวิธีการคัดเลือกข้อสอบมาให้ผู้สอบนั้น ขึ้นอยู่กับแบบทดสอบที่ผู้สอบเลือก มีทั้งสิ้น 2 แบบ คือ การทดสอบด้วยคอมพิวเตอร์ และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ขั้นตอนการทดสอบมีดังต่อไปนี้

**1.2.1 เริ่มต้นการทดสอบ**

เมื่อผู้สอบต้องการเริ่มทดสอบ ให้เลือก “การทดสอบ” ซึ่งเป็นเมนูหนึ่งในหน้าหลักของโปรแกรม หลังจากนั้นจะปรากฏหน้าจอดังภาพที่ 4-10




ภาพที่ 4-10 หน้าจอเริ่มต้นการทดสอบ

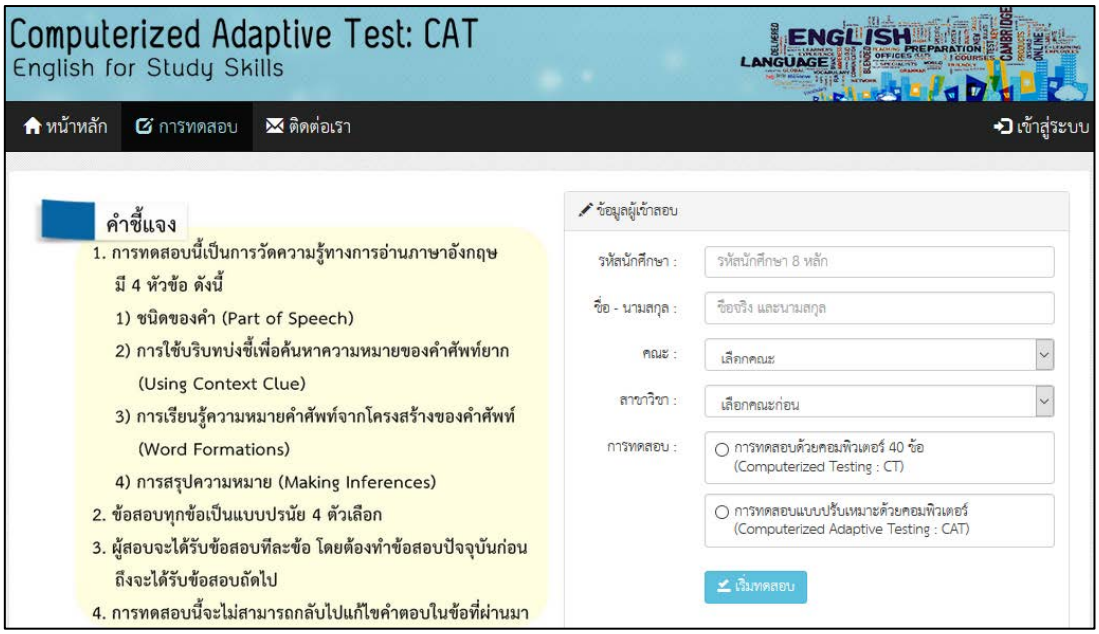
จากภาพที่ 4-10 แสดงหน้าจอเริ่มต้นการทดสอบ ซึ่งจะแสดงคำชี้แจงในการสอบให้ผู้สอบทราบ และให้ผู้สอบกรอกรหัสผ่านเพื่อเข้าสอบ ซึ่งผู้วิจัยกำหนดรหัสผ่านเบื้องต้นไว้ ดังนี้

| |
|---|
| ผู้ใช้: test<br>รหัสผ่าน: test |

โดยรหัสผ่านนี้กำหนดเพื่อป้องกันปัญหาเกี่ยวกับการทุจริตการสอบ ป้องกันไม่ให้บุคคลภายนอก หรือนักศึกษากลุ่มอื่น เข้าดูข้อสอบได้ ซึ่งผู้ใช้งาน ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการทดสอบ สามารถเปลี่ยนรหัสผ่านได้จาก ตั้งรหัสกลุ่มผู้สอบ แสดงตามภาพที่ 4-3

**1.2.2 กรอกข้อมูลผู้สอบ และเลือกแบบทดสอบ**

หลังจากผู้สอบกรอกรหัสผ่านเพื่อเข้าสอบเรียบร้อยแล้ว จะปรากฏหน้าจอดังภาพที่ 4-11

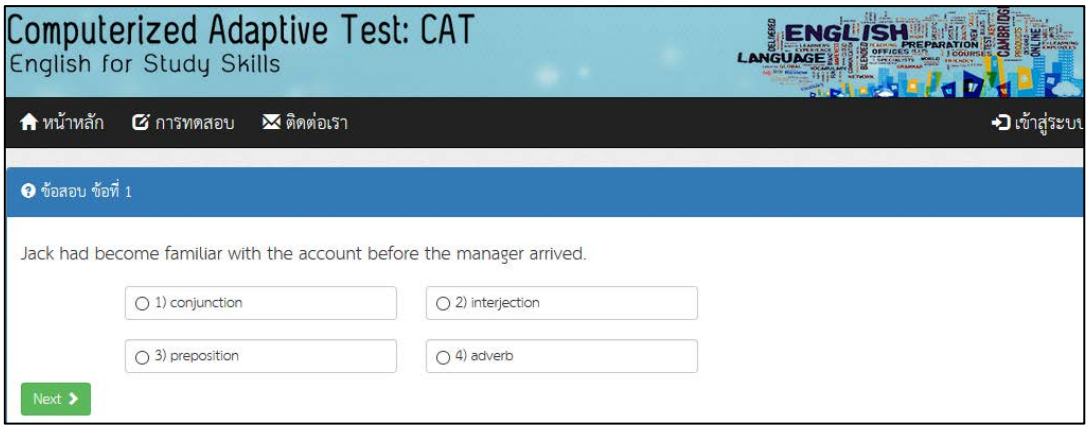



ภาพที่ 4-11 หน้าจอกรอกข้อมูลผู้สอบ และเลือกแบบทดสอบ

จากภาพที่ 4-11 แสดงหน้ากรอกข้อมูลผู้สอบ โดยให้ผู้สอบกรอกรหัสนักศึกษา ชื่อ-สกุล คณะ และสาขาวิชา ของตนเองลงไป พร้อมทั้งเลือกแบบทดสอบ ซึ่งในโปรแกรมนี้กำหนดแบบทดสอบให้ผู้สอบเลือก 2 แบบ ได้แก่ 1) การทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) และ 2) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) โดยผู้สอบต้องเลือกแบบทดสอบแบบใดแบบหนึ่งเท่านั้น หลักจากนั้นให้กดปุ่ม “เริ่มทดสอบ” เพื่อดำเนินการสอบต่อไป

**1.2.3 ดำเนินการสอบ**

หลังจากผู้สอบเลือกแบบทดสอบเรียบร้อยแล้ว จะเข้าสู่การดำเนินการสอบ โดยจะปรากฏหน้าจอดังภาพที่ 4-12




ภาพที่ 4-12 หน้าจอข้อสอบ

จากภาพที่ 4-12 แสดงหน้าจอข้อสอบ โดยโปรแกรมจะคัดเลือกข้อสอบมาให้ผู้สอบขึ้นอยู่กับแบบทดสอบที่ผู้สอบเลือก หลังจากนั้นให้ผู้สอบเลือกคำตอบ แล้วกดปุ่ม “Next” เพื่อรับข้อสอบข้อถัดไป ซึ่งการดำเนินการทดสอบจะดำเนินต่อไป จนสิ้นสุดการทดสอบ

**1.2.4 รายงานผลสอบรายบุคคล**

หลังจากผู้สอบดำเนินการทดสอบเสร็จสิ้นแล้ว โปรแกรมจะรายงานผลการทดสอบให้ผู้สอบทราบทันที ซึ่งการรายงานผลการทดสอบมี 2 แบบ ขึ้นอยู่กับแบบทดสอบที่ผู้สอบเลือก (CAT หรือ CT) สามารถแสดงได้ดังภาพที่ 4-13 และ 4-14

ภาพที่ 4-13 หน้าจอรายงานผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

จากภาพที่ 4-13 แสดงหน้าจอรายงานผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยแสดงรายละเอียดของผู้สอบ ได้แก่ ชื่อ-สกุล (Student Name) รหัสนักศึกษา (Student No)

คณะ (Faculty) และสาขาวิชา (Major) และแสดงรายละเอียดเกี่ยวกับการสอบ ได้แก่ วันที่สอบ (Date) เวลาที่ใช้ในการสอบ (Time) ความยาวของแบบทดสอบ (Test Length) และจำนวนข้อสอบที่ตอบถูก (Correct) ค่าประมาณความสามารถของผู้สอบ (Ability Estimate) และระดับความสามารถของผู้สอบ (Ability Level) พร้อมทั้งแสดงตารางการแจกแจงจำนวนข้อสอบที่ตอบถูกและค่าร้อยละ จำแนกตามชุดข้อสอบ ทั้งนี้เกณฑ์การประเมินระดับความสามารถของผู้สอบมีรายละเอียดดังตารางที่ 4-14

ตารางที่ 4-14 เกณฑ์ประเมินระดับความสามารถของผู้สอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

| ค่าประมาณความสามารถของผู้สอบ | ระดับความสามารถ |
|---|---|
| มากกว่า 2.0000 | Very High |
| 1.0001 ถึง 2.0000 | High |
| 0.5001 ถึง 1.0000 | Rather High |
| -0.4999 ถึง 0.5000 | Normal |
| -1.4999 ถึง -0.5000 | Rather Low |
| -2.0000 ถึง -1.5000 | Low |
| ต่ำกว่า -2.0000 | Very Low |

ภาพที่ 4-14 หน้าจอรายงานผลการทดสอบด้วยคอมพิวเตอร์

จากภาพที่ 4-14 แสดงหน้าจอรายงานผลการทดสอบด้วยคอมพิวเตอร์ โดยแสดงรายละเอียดของผู้สอบ และรายละเอียดเกี่ยวกับการสอบ เหมือนกับรายงานผลการทดสอบแบบ

ปรับเหมาะด้วยคอมพิวเตอร์ แต่แตกต่างตรงผลการทดสอบด้วยคอมพิวเตอร์ จะพิจารณาจากคะแนนรวมของผู้สอบ ซึ่งเป็นจำนวนข้อสอบที่ตอบถูก และปรับเทียบระดับความสามารถ โดยใช้เกณฑ์ดังนี้

ตารางที่ 4-15 เกณฑ์เทียบระดับความสามารถของผู้สอบ สำหรับการทดสอบด้วยคอมพิวเตอร์

| คะแนนรวม (ร้อยละ) | จำนวนข้อสอบที่ตอบถูก | ระดับความสามารถ |
|---|---|---|
| ตั้งแต่ 80% ขึ้นไป | ตั้งแต่ 32 ข้อ ขึ้นไป | Very High |
| 70-79% | 28-31 ข้อ | High |
| 60-69% | 24-27 ข้อ | Rather High |
| 50-59% | 20-23 ข้อ | Normal |
| 40-49% | 16-19 ข้อ | Rather Low |
| 30-39% | 12-15 ข้อ | Low |
| ต่ำกว่า 30% | ต่ำกว่า 12 ข้อ | Very Low |

**2. ผลการประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

ผลการประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี นำเสนอได้ 2 ส่วน ดังนี้

**2.1 ผลการประเมินโดยผู้เชี่ยวชาญ**

การประเมินโดยผู้เชี่ยวชาญ ผู้วิจัยได้นำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ พร้อมคู่มือการใช้โปรแกรมส่งให้ผู้เชี่ยวชาญตรวจสอบ โดยผู้เชี่ยวชาญทั้ง 3 คน ได้ประเมินโปรแกรมและคู่มือการใช้โปรแกรมใน 4 ด้าน คือ 1) ความสะดวกในการใช้โปรแกรม 2) ความถูกต้องในการใช้งาน 3) ลักษณะทั่วไปของโปรแกรม และ 4) คู่มือการใช้โปรแกรม ซึ่งประเมินในระหว่างวันที่ 1 - 11 สิงหาคม พ.ศ. 2559 สามารถแสดงผลลัพธ์การประเมินโปรแกรมและคู่มือการใช้โปรแกรมได้ดังนี้

ตารางที่ 4-16 ผลการประเมินโปรแกรมด้านความสะดวกในการใช้โปรแกรมโดยผู้เชี่ยวชาญ

| รายประเมิน | *Mean* | *SD* | ระดับความเหมาะสม |
|---|---|---|---|
| 1. การเข้าถึงโปรแกรมทำได้ง่ายและสะดวก | 4.67 | 0.58 | มากที่สุด |
| 2. โปรแกรมมีเมนูที่เข้าถึงส่วนต่าง ๆ ของโปรแกรมได้ง่ายและสะดวก | 4.33 | 0.58 | มาก |
| 3. การสอบโดยใช้โปรแกรมมีความสะดวกกว่าเมื่อเทียบกับการสอบบนกระดาษ | 5.00 | 0.00 | มากที่สุด |
| 4. โปรแกรมมีการแสดงผลการสอบทันที เมื่อทดสอบเสร็จ | 5.00 | 0.00 | มากที่สุด |
| 5. การแสดงผลสอบมีรายละเอียดครบถ้วน | 4.33 | 0.58 | มาก |
| สรุปผล | 4.67 | 0.94 | มากที่สุด |

จากตารางที่ 4-16 แสดงผลการประเมินโปรแกรมด้านความสะดวกในการใช้โปรแกรมโดยผู้เชี่ยวชาญ แสดงให้เห็นว่า โปรแกรมมีความเหมาะสมในด้านความสะดวกในการใช้งาน อยู่ในระดับมากที่สุด (*Mean* = 4.67)

ตารางที่ 4-17 ผลการประเมินโปรแกรมด้านความถูกต้องในการใช้งานโดยผู้เชี่ยวชาญ

| รายประเมิน | *Mean* | *SD* | ระดับความเหมาะสม |
|---|---|---|---|
| 1. โปรแกรมสามารถเพิ่ม ลบ หรือแก้ไข ข้อสอบได้ | 5.00 | 0.00 | มากที่สุด |
| 2. โปรแกรมสามารถจัดการทดสอบได้ตรงตามวัตถุประสงค์หรือเงื่อนไขของการทดสอบได้ | 5.00 | 0.00 | มากที่สุด |
| 3. โปรแกรมสามารถประมาณค่าความสามารถของผู้สอบถูกต้อง | 4.67 | 0.58 | มากที่สุด |
| 4. โปรแกรมสามารถบันทึกผลการทดสอบได้ถูกต้อง | 5.00 | 0.00 | มากที่สุด |
| สรุปผล | 4.91 | 0.30 | มากที่สุด |

จากตารางที่ 4-17 แสดงผลการประเมินโปรแกรมด้านความถูกต้องในการใช้งานโดยผู้เชี่ยวชาญ แสดงให้เห็นว่า โปรแกรมมีความเหมาะสมในด้านความถูกต้องในการใช้งาน อยู่ในระดับมากที่สุด (*Mean* = 4.91)

ตารางที่ 4-18 ผลการประเมินโปรแกรมด้านลักษณะทั่วไปของโปรแกรมโดยผู้เชี่ยวชาญ

| รายประเมิน | *Mean* | *SD* | ระดับความเหมาะสม |
|---|---|---|---|
| 1. การออกแบบโปรแกรมมีความน่าสนใจ | 4.00 | 0.00 | มาก |
| 2. การจัดรูปแบบหน้าจอง่ายต่อการใช้งาน | 5.00 | 0.00 | มากที่สุด |
| 3. การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว | 4.33 | 0.58 | มาก |
| 4. โปรแกรมมีระบบป้องกันการทำงานผิดพลาดของผู้ใช้ทุกขั้นตอน เช่น มีข้อเสนอแนะปรากฏขึ้นเมื่อใช้เมาส์ชี้ หรือเมื่อผู้ใช้ ใช้งานผิด | 4.33 | 0.58 | มาก |
| สรุปผล | 4.42 | 0.51 | มาก |

จากตารางที่ 4-18 แสดงผลการประเมินโปรแกรมด้านลักษณะทั่วไปของโปรแกรมโดยผู้เชี่ยวชาญ แสดงให้เห็นว่า โปรแกรมมีความเหมาะสมในด้านลักษณะทั่วไปของโปรแกรม อยู่ในระดับมาก (*Mean* = 4.42)

ตารางที่ 4-19 ผลการประเมินคู่มือการใช้โปรแกรมโดยผู้เชี่ยวชาญ

| รายประเมิน | *Mean* | *SD* | ระดับความเหมาะสม |
|---|---|---|---|
| 1. คู่มือมีคำอธิบายวัตถุประสงค์ของโปรแกรมได้อย่างชัดเจน | 5.00 | 0.00 | มากที่สุด |
| 2. คู่มือการใช้โปรแกรมแสดงวิธีการใช้งานอย่างมีลำดับขั้นตอน | 5.00 | 0.00 | มากที่สุด |
| 3. ภาษาที่ใช้มีความเข้าใจง่าย | 4.67 | 0.58 | มากที่สุด |
| 4. ภาพประกอบสามารถอธิบายขั้นตอนการใช้โปรแกรมได้อย่างชัดเจน | 4.33 | 0.58 | มาก |
| 5. หลังจากอ่านคู่มือแล้ว ผู้ใช้มีความมั่นใจว่าสามารถใช้โปรแกรมได้ | 4.67 | 0.58 | มากที่สุด |
| สรุปผล | 4.73 | 0.46 | มากที่สุด |

จากตารางที่ 4-19 แสดงผลการประเมินคู่มือการใช้โปรแกรมโดยผู้เชี่ยวชาญ แสดงให้เห็นว่า คู่มือการใช้โปรแกรมมีความเหมาะสม อยู่ในระดับมากที่สุด (*Mean* = 4.73)

นอกจากนี้ ผู้เชี่ยวชาญยังมีข้อเสนอแนะเพิ่มเติม ดังนี้

1. ควรปรับคู่มือให้มีความชัดเจนเกี่ยวกับการได้มาของค่าพารามิเตอร์ของข้อสอบ
2. ควรปรับหน้าหลักของโปรแกรมและหน้าการทดสอบให้อยู่ในหน้าเดียวกัน เพื่อความสะดวกในการทดสอบ
3. ที่อยู่ URL ของเว็ปไซต์ควรสั้นกว่านี้ เพราะผู้ใช้งานมีโอกาสพิมพ์ผิดมาก หากไม่มี link มาให้อาจใช้ bit.ly เพื่อย่อชื่อ URL
4. หน้าจอเข้าสู่ระบบ เมื่อเข้าใช้งานแล้ว ไม่ปรากฏแถบเมนูด้านบน ทำให้ไปยังเมนูอื่น ๆ ไม่ได้
5. หน้าจอหลักควรมีรายละเอียดที่มาของโปรแกรม ก่อนบอกรายละเอียดว่ามีการทดสอบให้เลือก 2 การทดสอบ
6. ถ้าผู้สอบคนเดิมเข้าใช้งาน มีการเก็บผลการทดสอบไว้หรือไม่ ผู้สอบสามารถค้นหาข้อมูลเพื่อเปรียบเทียบการทดสอบของตนเองได้หรือไม่
7. ควรมีเมนู Help หรือมี Link เพื่อดาวน์โหลดคู่มือการใช้โปรแกรมบนเว็ปไซต์ ควรแยกคู่มือสำหรับผู้สอบ และสำหรับครู-อาจารย์ คนละฉบับ
8. ควรมีหน้าจอสำหรับลงทะเบียนผู้เข้าสอบ เพราะถ้าไม่มีรหัสมาให้ก่อน จะไม่สามารถเข้าใช้งานได้
9. เป็นโปรแกรมที่มีประโยชน์ในการพัฒนาทักษะการเรียนรู้ ควรปรับโปรแกรมให้สมบูรณ์ และนำไปใช้จริง

**2.2 ผลการประเมินโดยผู้สอบ**

การประเมินโดยผู้สอบ เป็นการประเมินความคิดเห็นในด้านความสะดวกในการนำไปใช้และด้านลักษณะทั่วไปของโปรแกรม โดยทดลองใช้กับนักศึกษาคณะวิทยาการจัดการ มหาวิทยาลัย

ศิลปากร จำนวน 30 คน ในระหว่างวันที่ 15 - 16 สิงหาคม พ.ศ. 2559 สามารถแสดงผลลัพธ์การประเมินโปรแกรมได้ดังนี้

ตารางที่ 4-20 ผลการประเมินโปรแกรมด้านความสะดวกในการใช้โปรแกรมโดยผู้สอบ

| รายประเมิน | *Mean* | *SD* | ระดับความเหมาะสม |
|---|---|---|---|
| 1. การเข้าถึงโปรแกรมทำได้ง่ายและสะดวก | 4.41 | 0.63 | มาก |
| 2. โปรแกรมมีเมนูที่เข้าถึงส่วนต่าง ๆ ของโปรแกรมได้ง่ายและสะดวก | 4.31 | 0.66 | มาก |
| 3. การสอบโดยใช้โปรแกรมมีความสะดวกกว่าเมื่อเทียบกับการสอบบนกระดาษ | 4.53 | 0.68 | มากที่สุด |
| 4. โปรแกรมมีการแสดงผลการสอบทันที เมื่อทดสอบเสร็จ | 4.57 | 0.63 | มากที่สุด |
| 5. การแสดงผลสอบมีรายละเอียดครบถ้วน | 4.47 | 0.63 | มาก |
| สรุปผล | 4.46 | 0.64 | มาก |

จากตารางที่ 4-20 แสดงผลการประเมินโปรแกรมด้านความสะดวกในการใช้โปรแกรมโดยผู้สอบ แสดงให้เห็นว่า โปรแกรมมีความเหมาะสมในด้านความสะดวกในการใช้โปรแกรม อยู่ในระดับมาก (*Mean* = 4.46)

ตารางที่ 4-21 ผลการประเมินโปรแกรมด้านลักษณะทั่วไปของโปรแกรมโดยผู้สอบ

| รายประเมิน | *Mean* | *SD* | ระดับความเหมาะสม |
|---|---|---|---|
| 1. การออกแบบโปรแกรมมีความน่าสนใจ | 3.90 | 0.66 | มาก |
| 2. การจัดรูปแบบหน้าจอง่ายต่อการใช้งาน | 4.40 | 0.62 | มาก |
| 3. การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว | 4.03 | 0.81 | มาก |
| 4. โปรแกรมมีระบบป้องกันการทำงานผิดพลาดของผู้ใช้ทุกขั้นตอน เช่น มีข้อเสนอแนะปรากฏขึ้นเมื่อใช้เมาส์ชี้ หรือเมื่อผู้ใช้ ใช้งานผิด | 3.80 | 0.81 | มาก |
| สรุปผล | 4.03 | 0.76 | มาก |

จากตารางที่ 4-21 แสดงผลการประเมินโปรแกรมด้านลักษณะทั่วไปของโปรแกรมโดยผู้สอบ แสดงให้เห็นว่า โปรแกรมมีความเหมาะสมในด้านลักษณะทั่วไปของโปรแกรม อยู่ในระดับมาก (*Mean* = 4.03)

นอกจากนี้ ผู้สอบมีข้อเสนอแนะเพิ่มเติม ดังนี้

1. ขนาดตัวอักษรในเว็ปไซต์มีขนาดเล็ก ไม่สะดวกในการอ่าน
2. ปรับปรุงรูปภาพให้มีสีสันน่าสนใจมากกว่านี้
3. อยากให้โปรแกรมสามารถอธิบายองค์ประกอบของประโยค อย่างเช่น หลักแกรมม่า และควรเพิ่มการเฉลยคำตอบในข้อที่ตอบผิด

4. ปรับปรุงโปรแกรมให้สามารถกลับไปแก้ไขคำตอบได้

5. ปรับปรุงโปรแกรมให้สามารถบอกเวลาที่เหลือในการทดสอบได้

จากผลการประเมินโปรแกรมและคู่มือการใช้โปรแกรมทั้งจากผู้เชี่ยวชาญและผู้สอบ ผู้วิจัยได้นำข้อเสนอแนะมาปรับปรุงแก้ไขโปรแกรมให้มีความเหมาะสมยิ่งขึ้น แต่จากผลการประเมินแสดงให้เห็นว่า โปรแกรมและคู่มือการใช้โปรแกรมมีความเหมาะสมเพียงพอที่จะนำไปใช้ทดสอบจริงได้

## ตอนที่ 4 ผลการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์

การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคและมีการควบคุมการใช้ข้อสอบ กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ โดยกลุ่มตัวอย่างเป็นนักศึกษาชั้นปีที่ 2 ในคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร ซึ่งคัดเลือกด้วยวิธีการเลือกตัวอย่างแบบตามสะดวก จำนวน 30 คน ดำเนินการเก็บข้อมูลในวันที่ 1 กันยายน พ.ศ. 2559 โดยใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี เป็นเครื่องมือที่ใช้ในการวิจัย โดยได้ผลการวิจัยดังนี้

ตารางที่ 4-22 ความยาวของแบบทดสอบ เวลาที่ใช้ในการทดสอบ และผลการทดสอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์

| คนที่ | จำนวนข้อสอบที่ใช้ในการทดสอบ (ข้อ) | | เวลาที่ใช้ในการทดสอบ (นาที: วินาที) | | ผลการทดสอบ | |
|---|---|---|---|---|---|---|
| | CAT | CT | CAT | CT | CAT ($\hat{\theta}$) | CT (คะแนนรวม) |
| 1 | 12 | 40 | 04:01 | 13:48 | -0.77 | 15 |
| 2 | 8 | 40 | 14:41 | 26:23 | 0.26 | 19 |
| 3 | 9 | 40 | 07:19 | 23:54 | 0.25 | 21 |
| 4 | 10 | 40 | 10:52 | 13:20 | 1.88 | 25 |
| 5 | 7 | 40 | 10:27 | 16:23 | -0.95 | 8 |
| 6 | 14 | 40 | 07:06 | 16:21 | -1.46 | 15 |
| 7 | 15 | 40 | 37:22 | 10:47 | 1.68 | 30 |
| 8 | 7 | 40 | 13:45 | 43:14 | -0.82 | 26 |
| 9 | 11 | 40 | 12:27 | 19:01 | 0.62 | 20 |
| 10 | 12 | 40 | 13:26 | 17:54 | -1.38 | 11 |
| 11 | 10 | 40 | 25:20 | 15:52 | -0.41 | 20 |
| 12 | 11 | 40 | 09:29 | 18:58 | -0.02 | 21 |
| 13 | 16 | 40 | 10:21 | 20:45 | 0.00 | 19 |

ตารางที่ 4-22 (ต่อ)

| คนที่ | จำนวนข้อสอบที่ใช้ในการทดสอบ (ข้อ) | | เวลาที่ใช้ในการทดสอบ (นาที: วินาที) | | ผลการทดสอบ | |
|---|---|---|---|---|---|---|
| | CAT | CT | CAT | CT | CAT ($\hat{\theta}$) | CT (คะแนนรวม) |
| 14 | 9 | 40 | 07:45 | 31:37 | -0.54 | 24 |
| 15 | 11 | 40 | 14:53 | 13:51 | 1.22 | 25 |
| 16 | 6 | 40 | 17:33 | 18:54 | 1.60 | 31 |
| 17 | 6 | 40 | 03:42 | 21:25 | -1.13 | 13 |
| 18 | 8 | 40 | 02:50 | 11:02 | -0.85 | 26 |
| 19 | 10 | 40 | 03:14 | 24:57 | -1.42 | 18 |
| 20 | 9 | 40 | 02:23 | 19:29 | -1.10 | 18 |
| 21 | 7 | 40 | 16:01 | 17:41 | 0.54 | 16 |
| 22 | 8 | 40 | 15:41 | 18:55 | -0.72 | 21 |
| 23 | 13 | 40 | 07:05 | 13:42 | -1.56 | 21 |
| 24 | 7 | 40 | 15:41 | 16:43 | -0.82 | 29 |
| 25 | 12 | 40 | 04:44 | 17:47 | 0.11 | 20 |
| 26 | 10 | 40 | 09:14 | 18:31 | 0.03 | 13 |
| 27 | 7 | 40 | 07:39 | 28:33 | -0.99 | 20 |
| 28 | 14 | 40 | 15:25 | 18:04 | -0.10 | 18 |
| 29 | 8 | 40 | 13:24 | 16:13 | 0.60 | 24 |
| 30 | 8 | 40 | 06:31 | 22:24 | 1.17 | 27 |
| *Min* | 6 | 40 | 02:23 | 10:47 | -1.56 | 8 |
| *Max* | 16 | 40 | 37:22 | 43:14 | 1.88 | 31 |
| *Mean* | 9.91 | 40 | 11:01 | 19:33 | -0.16 | 20.47 |
| *SD* | 2.73 | 0.00 | 07:20 | 06:34 | 1.00 | 5.59 |

หมายเหตุ CAT หมายถึง การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ (HC-Ex)

CT หมายถึง การทดสอบด้วยคอมพิวเตอร์

จากตารางที่ 4-22 แสดงความยาวของแบบทดสอบ เวลาที่ใช้ในการทดสอบ และผลการทดสอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์ ปรากฏว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีความยาวของของแบบทดสอบเฉลี่ย 9.91 ข้อ ในขณะที่การทดสอบด้วยคอมพิวเตอร์มีความยาวของแบบทดสอบ 40 ข้อ แสดงให้เห็นว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ มีความยาวของแบบทดสอบคิดเป็นร้อยละ 24.78 ของการทดสอบด้วยคอมพิวเตอร์

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใช้เวลาในการทดสอบเฉลี่ย 11:01 นาที ขณะที่การทดสอบด้วยคอมพิวเตอร์ใช้เวลาในการทดสอบเฉลี่ย 19:33 นาที แสดงให้เห็นว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ ใช้เวลาในการทดสอบคิดเป็นร้อยละ 56.35 ของการทดสอบด้วยคอมพิวเตอร์

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้ค่าประมาณความสามารถของผู้สอบ ($\hat{\theta}$) เฉลี่ยเท่ากับ -0.16 หมายความว่า ผู้สอบมีความสามารถโดยเฉลี่ยอยู่ในระดับปานกลาง (เกณฑ์การประเมินระดับความสามารถของผู้สอบ สำหรับการทดสอบแบบ CAT แสดงในตารางที่ 4-14) เช่นเดียวกับการทดสอบด้วยคอมพิวเตอร์ได้คะแนนรวมเฉลี่ยเท่ากับ 20.47 หมายความว่า ผู้สอบมีความสามารถโดยเฉลี่ยอยู่ในระดับปานกลาง (เกณฑ์การประเมินความสามารถของผู้สอบ สำหรับการทดสอบแบบ CT แสดงในตารางที่ 4-15) แสดงให้เห็นว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ ให้ผลการทดสอบใกล้เคียงกับการทดสอบด้วยคอมพิวเตอร์

ตารางที่ 4-23 ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างค่าประมาณความสามารถของผู้สอบกับคะแนนรวมของผู้สอบ

| | CAT | CT |
|---|---|---|
| CAT | 1 | |
| CT | .535** | 1 |

*** p < .01*

จากตารางที่ 4-23 แสดงค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ปรากฏว่า ค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ มีความสัมพันธ์เชิงบวก ขนาดปานกลาง ที่ระดับนัยสำคัญทางสถิติ .01 แสดงว่า ผลการทดสอบแบบ CAT มีความสอดคล้องกับผลการทดสอบแบบ CT นั่นคือ ผู้สอบที่ได้รับค่าประมาณความสามารถจากการทดสอบแบบ CAT สูง (หรือต่ำ) จะได้คะแนนรวมจากการทดสอบแบบ CT สูง (หรือต่ำ) ด้วย ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อ 4

# บทที่ 5
# สรุปผลและอภิปรายผล

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบโดยใช้วิธีการสุ่มตัวอย่างแบบมีระบบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แล้วนำวิธีการที่ได้พัฒนาขึ้นทั้งในกรณีที่มีและไม่มีการควบคุมการใช้ข้อสอบ มาเปรียบเทียบประสิทธิภาพกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง พร้อมทั้งนำมาพัฒนาโปรแกรมการทดสอบความสามารถทางภาษาอังกฤษ การดำเนินการวิจัยแบ่งเป็น 4 ระยะ ได้แก่ 1) การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจตามเกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบโดยใช้วิธีการสุ่มตัวอย่างแบบมีระบบ 2) การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยศึกษาในสถานการณ์จำลองแบบมอนติคาร์โลด้วยการจำลองค่าความสามารถที่แท้จริงของผู้สอบ 10,000 ค่า และจำลองค่าพารามิเตอร์ของข้อสอบ 500 ค่า วิเคราะห์ข้อมูลด้วยค่าเฉลี่ยเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย ค่าความลำเอียงเฉลี่ย การวิเคราะห์ Brown-Forsythe และ Kruskal-Wallis 3) การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี และ 4) การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ โดยใช้กลุ่มตัวอย่างเป็นนักศึกษาปริญญาตรี ชั้นปีที่ 2 ในคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร ซึ่งคัดเลือกตัวอย่างแบบตามสะดวก จำนวน 30 คน โดยวิเคราะห์ข้อมูลด้วยการวิเคราะห์สหสัมพันธ์แบบเพียร์สัน

## สรุปผลการวิจัย

1. ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ เริ่มจากวิธีการควบคุมการใช้ข้อสอบ จนถึงวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ ดังนี้

1.1 ผลการพัฒนาวิธีการควบคุมการใช้ข้อสอบด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ สามารถสรุปเป็นขั้นตอนดำเนินการดังนี้

ขั้นที่ 1 แบ่งข้อสอบในคลังข้อสอบ (หรือข้อสอบในหัวข้อเรื่องถัดจากข้อสอบที่ทำขณะนั้น) ออกเป็น 2 กลุ่ม คือ ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบต่ำถึงปานกลาง ซึ่งมีค่าน้อยกว่า 1.35 กำหนดให้เป็นกลุ่ม A และข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ซึ่งมีค่าตั้งแต่ 1.35 ขึ้นไป กำหนดให้เป็นกลุ่ม B

ขั้นที่ 2 สุ่มเลือกข้อสอบในกลุ่ม B ด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ ซึ่งมีขั้นตอนการสุ่มดังนี้

ขั้นที่ 2.1 กำหนดจำนวนข้อสอบที่ต้องการสุ่มในกลุ่ม B สมมติว่าต้องการสุ่มข้อสอบจำนวน n ข้อ โดยที่ n เป็นจำนวนเต็มที่สุ่มจากช่วงจำนวนตั้งแต่ 5 ถึง 10 ข้อ

ขั้นที่ 2.2 หาช่วงของการสุ่ม (k) โดยคำนวณจาก $k=\frac{N}{n}$ เมื่อ N คือ จำนวนข้อสอบทั้งหมดในกลุ่ม B และ n คือ ข้อสอบที่ต้องสุ่ม

ขั้นที่ 2.3 หาข้อสอบที่ใช้เป็นตัวสุ่มเริ่มต้น (r) โดยสุ่มค่า r จากค่าตั้งแต่ 1 ถึง k จากนั้นจึงเลือกข้อสอบโดยเริ่มตั้งแต่ข้อที่ r, r+k, r+2k ตามลำดับจนครบ n ข้อ

ขั้นที่ 3 รวมข้อสอบในกลุ่ม A กับข้อสอบที่สุ่มได้ในขั้นที่ 2 เข้าด้วยกัน ซึ่งข้อสอบทั้งหมดในขั้นนี้ จะนำไปใช้ในกระบวนการคัดเลือกข้อสอบข้อถัดไป

1.2 ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ สามารถสรุปขั้นตอนดำเนินการดังนี้

ขั้นที่ 1 นำข้อสอบที่ผ่านการควบคุมการใช้ข้อสอบ (จากข้อ 1.1) มาคำนวณหาค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ $H(i)$ ตามเกณฑ์ของเฮอร์วิคซ์ ซึ่งมีสูตรคำนวณดังนี้

$$H(i) = \alpha \times \min(SEE_{true}(\hat{\theta}), SEE_{false}(\hat{\theta})) + (1-\alpha) \times \max(SEE_{true}(\hat{\theta}), SEE_{false}(\hat{\theta})) \quad (41)$$

โดยที่ $H(i)$ คือ ค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ณ ข้อสอบข้อที่ i

$SEE_{true}(\hat{\theta})$ คือ ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถในกรณีที่ผู้สอบตอบถูก

$SEE_{false}(\hat{\theta})$ คือ ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ในกรณีที่ผู้สอบตอบผิด

min คือ ค่าต่ำสุดที่เลือกจากค่า $SEE_{true}(\hat{\theta})$ และค่า $SEE_{false}(\hat{\theta})$ ขึ้นอยู่กับค่าใดต่ำกว่ากัน

max คือ ค่าสูงสุดที่เลือกจากค่า $SEE_{true}(\hat{\theta})$ และค่า $SEE_{false}(\hat{\theta})$ ขึ้นอยู่กับค่าใดสูงกว่ากัน

$\hat{\theta}$ คือ ค่าประมาณความสามารถของผู้สอบ

$\alpha$ คือ สัมประสิทธิ์ของการมองโลกในแง่ดี มีเท่ากับ $1-\frac{1}{5.5}|\hat{\theta} - b|$

b คือ ค่าความยากของข้อสอบ

ขั้นที่ 2 เลือกข้อสอบที่มีค่า $H(i)$ ต่ำสุด เป็นข้อสอบข้อถัดไป

2. ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป เป็นการนำวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (วิธีการ MIC) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (วิธีการ RDM) มาเปรียบเทียบประสิทธิภาพกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบที่ได้

พัฒนาขึ้น ซึ่งจะเปรียบเทียบทั้งในกรณีที่ไม่มีการควบคุมการใช้ข้อสอบ (วิธีการ HC) และกรณีที่มีการควบคุมการใช้ข้อสอบ (วิธีการ HC-Ex) สรุปผลการวิจัยได้ดังนี้

2.1 ผลการเปรียบเทียบประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบ ซึ่งพิจารณาจากค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) ปรากฏว่า วิธีการ RDM มีประสิทธิภาพสูงสุด รองลงมา ได้แก่ วิธีการ HC ซึ่งมีประสิทธิภาพเทียบเท่ากับวิธีการ HC-Ex ส่วนวิธีการ MIC มีประสิทธิภาพต่ำสุด ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อ 1.1 บางส่วน แต่เมื่อพิจารณาจากค่าความลำเอียงเฉลี่ย ปรากฏว่า วิธีการ RDM, HC และ HC-Ex มีประสิทธิภาพสูงสุดเทียบเท่ากัน ส่วนวิธีการ MIC มีประสิทธิภาพต่ำสุด ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 1.2 บางส่วน

2.2 ผลการเปรียบเทียบประสิทธิภาพด้านความยาวของแบบทดสอบ ปรากฏว่า วิธีการ RDM และวิธีการ HC มีประสิทธิภาพสูงสุดเทียบเท่ากัน รองลงมา ได้แก่ วิธีการ HC-Ex และวิธีการ MIC ตามลำดับ ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 2 บางส่วน

2.3 ผลการเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ปรากฏว่า วิธีการ MIC มีประสิทธิภาพสูงสุด รองลงมา ได้แก่ วิธีการ HC-Ex ส่วนวิธีการ RDM มีประสิทธิภาพต่ำสุดเทียบเท่ากับวิธีการ HC ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 3 บางส่วน

3. ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ

3.1 ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวัดความรู้ในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนของนักศึกษาในระดับปริญญาตรี โดยพัฒนาอยู่ในรูปของ Web Application จึงสามารถใช้งานผ่านอุปกรณ์ทั่วไปที่เชื่อมต่ออินเทอร์เน็ตได้ ซึ่งโปรแกรมออกแบบให้ผู้ใช้สามารถเลือกการทดสอบได้ 2 แบบ คือ 1) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และ 2) การทดสอบด้วยคอมพิวเตอร์ ทั้งนี้โปรแกรมแบ่งการใช้งานได้ 2 ส่วน คือ ส่วนที่ 1 การจัดการทดสอบ เป็นส่วนที่ให้ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการสอบนำข้อสอบที่ผ่านการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการตอบสนองข้อสอบมาใส่ในโปรแกรมเพื่อใช้เป็นคลังข้อสอบ รวมทั้งการลบ และแก้ไขข้อสอบเดิมที่อยู่ในคลังข้อสอบ นอกจากนี้ยังมีการรายงานผลการทดสอบของผู้สอบทั้งหมด เพื่อให้ครู อาจารย์ นำผลการทดสอบไปใช้ประเมินผลการเรียนรู้ต่อไป ส่วนที่ 2 การจัดการสอบ เป็นส่วนที่ให้ผู้สอบ คือ นิสิต นักศึกษา ใช้โปรแกรมนี้เพื่อทดสอบวัดความสามารถทางภาษาอังกฤษ มีการรายงานผลการทดสอบทันทีที่ผู้สอบทำการทดสอบเสร็จสิ้น

3.2 ผลการประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบ่งได้ 2 ส่วน คือ 1) ผลการประเมินโดยผู้เชี่ยวชาญ ปรากฏว่า โปรแกรมมีความเหมาะสมอยู่ในระดับมากถึงมากที่สุด ทั้งในด้านความสะดวกในการใช้โปรแกรม ความถูกต้องลักษณะทั่วไปของโปรแกรม และความชัดเจนของคู่มือ และ 2) ผลการประเมินโดยผู้สอบ ปรากฏว่า โปรแกรมมีความเหมาะสมอยู่ในระดับมาก ทั้งในด้านความสะดวกในการใช้โปรแกรม และด้านลักษณะทั่วไปของโปรแกรม

4. ผลการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ได้พัฒนาขึ้นกับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ จากกลุ่มตัวอย่างจำนวน 30 คน

4.1 การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) กลุ่มตัวอย่างใช้เวลาในการทดสอบเฉลี่ย 11:01 นาที ใช้ข้อสอบเฉลี่ย 9.91 ข้อ และมีค่าประมาณความสามารถของผู้สอบ ($\hat{\theta}$) เฉลี่ย -0.14 ส่วนการทดสอบด้วยคอมพิวเตอร์ (CT) กลุ่มตัวอย่างใช้เวลาในการทดสอบเฉลี่ย 19:33 นาที ใช้ข้อสอบ 40 ข้อ และมีคะแนนรวมของผู้สอบเฉลี่ย 20.41 ข้อ แสดงให้เห็นว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใช้เวลาและจำนวนข้อสอบลดลง เมื่อเทียบกับการทดสอบด้วยคอมพิวเตอร์

4.2 ผลการทดสอบความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ปรากฏว่า ผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับผลการทดสอบด้วยคอมพิวเตอร์มีความสัมพันธ์เชิงบวก ขนาดปานกลาง ที่ระดับนัยสำคัญทางสถิติ .01 ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 4

## อภิปรายผล

การอภิปรายผลการวิจัย แบ่งได้ 4 ประเด็น ได้แก่ 1) การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ 2) การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 3) การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และ 4) การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ มีรายละเอียดดังนี้

**1. การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ**

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นการทดสอบที่มีการจัดข้อสอบให้เหมาะสมกับความสามารถของผู้สอบ ซึ่งพิจารณาจากผลการตอบข้อสอบในทุกข้อที่ผ่านมา กล่าวคือ เมื่อผู้สอบทำข้อสอบข้อแรกหรือชุดเริ่มต้นเสร็จลง จะนำผลการตอบมาวิเคราะห์ระดับความสามารถของผู้สอบ เพื่อใช้คัดเลือกข้อสอบข้อถัดไปที่มีความเหมาะสมกับความสามารถของผู้สอบ โดยใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เป็นพื้นฐาน ขั้นตอนการทดสอบแบบ CAT มี 5 ขั้นตอน คือ ขั้นตอนที่ 1 การสร้างคลังข้อสอบ (Create Item Bank) เป็นการจัดทำข้อสอบที่ใช้ในการทดสอบ ซึ่งต้องสอดคล้องกับทฤษฎี IRT ขั้นตอนที่ 2 การคัดเลือกข้อสอบข้อแรก (First Item Selection) เป็นการคัดเลือกข้อสอบข้อแรก โดยส่วนใหญ่ใช้ข้อสอบที่มีค่าความยากของข้อสอบระดับปานกลาง ขั้นตอนที่ 3 การคัดเลือกข้อสอบข้อถัดไป (Next Item Selection) โดยคัดเลือกข้อสอบให้มีความเหมาะสมกับระดับความสามารถของผู้สอบ ซึ่งพิจารณาจากผลการตอบข้อสอบในทุกข้อที่ผ่านมา ขั้นตอนที่ 4 การประมาณค่าความสามารถของผู้สอบ (Calculate Possible Ability Level) โดยอาศัย

ผลการตอบข้อสอบในทุกข้อที่ผ่านมา และขั้นตอนที่ 5 เกณฑ์ยุติการทดสอบ (Termination Criterion) เป็นการกำหนดเกณฑ์สิ้นสุดการทดสอบ โดยการทดสอบจะดำเนินการซ้ำตั้งแต่ขั้นตอนที่ 3 ถึงขั้นตอนที่ 5 จนกระทั่งการทดสอบเป็นไปตามเกณฑ์ยุติการทดสอบที่กำหนดไว้

การคัดเลือกข้อสอบข้อถัดไปเป็นหนึ่งในขั้นตอนที่สำคัญในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เนื่องจาก ผลการตอบข้อสอบจะถูกนำไปใช้ในการประมาณค่าความสามารถของผู้สอบ หากผู้สอบได้รับข้อสอบที่มีความยากเหมาะสมกับระดับความสามารถของตนเอง ผลการประมาณค่าความสามารถของผู้สอบจะมีความถูกต้องแม่นยำยิ่งขึ้น ตามทฤษฎี IRT ข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบต้องเป็นข้อสอบที่ให้ค่าสารสนเทศของข้อสอบสูง โดยทั่วไปข้อสอบจะมีค่าสารสนเทศของข้อสอบสูงขึ้น เมื่อข้อสอบข้อนั้นมีค่าอำนาจจำแนกของข้อสอบสูง มีค่าความยากของข้อสอบใกล้เคียงกับระดับความสามารถของผู้สอบ และมีค่าการเดาของข้อสอบเข้าใกล้ศูนย์ หากแบบทดสอบที่ผู้สอบได้รับประกอบด้วยข้อสอบที่มีค่าสารสนเทศของข้อสอบสูง จะทำให้ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบต่ำลง (ศิริชัย กาญจนวาสี, 2555, หน้า 63-68; Wainer et al., 2000, pp. 63-64)

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในอดีต จึงคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง มีค่าความยากของข้อสอบใกล้เคียงกับระดับความสามารถของผู้สอบ คือ คัดเลือกข้อสอบที่ทำให้ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบต่ำสุด มาเป็นข้อสอบข้อถัดไป อาศัยหลักการนี้มาพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป โดยเลือกทฤษฎีการตัดสินใจ (Decision Theory) มาใช้พัฒนาควบคู่กับทฤษฎี IRT ซึ่งทฤษฎีการตัดสินใจจะใช้ข้อมูลในอดีต หรือใช้ดุลยพินิจหรือประสบการณ์ของผู้ตัดสินใจมาช่วยตัดสินใจเลือกทางเลือกที่มีความเหมาะสมที่สุด โดยในแต่ละทางเลือก จะวิเคราะห์หาผลลัพธ์ที่คาดว่าจะเกิดขึ้น เมื่อผู้ตัดสินใจเลือกทางเลือกนั้นไว้ล่วงหน้า แล้วตัดสินใจเลือกทางเลือกที่ให้ผลลัพธ์ (ที่คาดว่าจะเกิดขึ้น) ที่ดีที่สุดแก่ผู้ตัดสินใจ ผลการตัดสินใจจึงมีความผิดพลาดน้อยลง เพราะพิจารณาจากผลลัพธ์ที่คาดว่าจะเกิดขึ้น (Taylor, 2009, pp. 538-539)

การนำเกณฑ์ของเฮอร์วิคซ์ ซึ่งเป็นเกณฑ์หนึ่งภายใต้ทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน มาใช้คัดเลือกข้อสอบข้อถัดไป โดยคำนวณค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ (SEE) ของข้อสอบทุกข้อในคลังข้อสอบ ที่ยังไม่ถูกเลือกใช้ในการทดสอบ โดยคำนวณค่า SEE ทั้งในกรณีที่คาดว่าผู้สอบจะตอบข้อสอบถูก และกรณีที่คาดว่าผู้สอบจะตอบข้อสอบผิด แล้วนำค่า SEE ทั้งสองกรณี มาคำนวณค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE จากนั้นตัดสินใจเลือกข้อสอบข้อที่มีค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE ต่ำสุด เป็นข้อสอบข้อถัดไป ตามกระบวนการนี้เกณฑ์ของเฮอร์วิคซ์จึงมีความเหมาะสมสำหรับใช้คัดเลือกข้อสอบข้อถัดไป เนื่องจาก ผู้สอบจะได้ข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบ (เป็นข้อสอบที่มีค่าเฉลี่ยของค่า SEE ต่ำ) นอกจากนี้ เกณฑ์ของเฮอร์วิคซ์เป็นเกณฑ์ที่มีประสิทธิภาพค่อนข้างสูง เมื่อเทียบกับเกณฑ์อื่น ๆ ในทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน สอดคล้องกับงานวิจัยของ Groenewald and Pretorius (2011) ที่ได้ศึกษากลยุทธ์ในการตัดสินใจเลือกพอร์ตการลงทุน โดยใช้เกณฑ์ที่แตกต่างกันของการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ผลปรากฏว่า การตัดสินใจโดยใช้เกณฑ์ของเฮอร์วิคซ์ให้

ผลตอบแทนรายเดือนของตลาดเงิน (the Monthly Returns of the Money Market) ที่เหมาะสมที่สุด

การคัดเลือกข้อสอบข้อถัดไป ส่วนใหญ่จะคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง มาเป็นข้อสอบข้อถัดไป ดังนั้น ข้อสอบเหล่านี้จึงถูกเลือกใช้ในการทดสอบบ่อยครั้งจนเกินไป เป็นผลให้ผู้สอบในรุ่นถัดไป สามารถคาดการณ์ข้อสอบที่ตนเองจะได้ล่วงหน้าได้ ในการวิจัยนี้ ได้พัฒนาวิธีการควบคุมการใช้ข้อสอบโดยใช้วิธีการสุ่มตัวอย่างแบบมีระบบ (Systematic Random Sampling) มาควบคุมการใช้ข้อสอบ โดยข้อสอบที่นำไปใช้ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไป คือ ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบต่ำถึงปานกลาง และข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงจำนวน 5-10 ข้อ ซึ่งถูกสุ่มจากกลุ่มข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ ด้วยกระบวนการนี้ ผู้สอบที่มีความสามารถเท่ากันหรือใกล้เคียงกัน จึงมีโอกาสได้รับข้อสอบข้อเดียวกันน้อยลง เนื่องจาก ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงถูกหมุนเวียน (จากการสุ่ม) นำไปใช้ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไป วิธีการนี้จึงลดจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ลงได้ (Georgiadou et al., 2007)

ประเด็นหนึ่งที่สำคัญในการพัฒนาวิธีการควบคุมการใช้ข้อสอบโดยใช้วิธีการสุ่มตัวอย่างแบบมีระบบ คือ จำนวนข้อสอบที่สุ่มจากข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ ควรสุ่มข้อสอบจำนวนเท่าใดจึงจะเหมาะสม โดยในช่วงเริ่มต้นของการพัฒนา ได้ทดลองกำหนดจำนวนข้อสอบสุ่มไว้ที่ 10% หรือ 20% ของจำนวนข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงทั้งหมดที่มีในคลังข้อสอบ ปรากฏว่า เมื่อกำหนดจำนวนข้อสอบสุ่มน้อยลง มีผลทำให้ข้อสอบที่มีอัตราการใช้ข้อสอบสูงกว่า 0.2 มีแนวโน้มลดจำนวนลง (บ่งชี้ว่า จำนวนข้อสอบที่ถูกใช้ซ้ำบ่อยลดลง ซึ่งเป็นข้อดี) แต่ในขณะเดียวกันค่า RMSE และความยาวของแบบทดสอบกับมีแนวโน้มเพิ่มขึ้น (บ่งชี้ว่า ประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบและด้านความยาวของแบบทดสอบลดลง ซึ่งเป็นข้อเสีย) ซึ่งเป็นผลมาจากเมื่อกำหนดจำนวนข้อสอบที่สุ่มไปใช้ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไปให้มีจำนวนมากขึ้น ข้อสอบที่ไม่เหมาะสมกับผู้สอบ เช่น ข้อสอบที่มีค่าความยากของข้อสอบต่างจากค่าความสามารถของผู้สอบมาก จึงมีโอกาสถูกคัดเลือกไปให้ผู้สอบเพิ่มมากขึ้น ดังนั้น ในการพัฒนาวิธีการควบคุมการใช้ข้อสอบนี้ จึงกำหนดจำนวนข้อสอบที่จะสุ่มให้มีจำนวนน้อยตั้งแต่ 5-10 ข้อ โดยอาศัยแนวคิดจากวิธีการ 5-4-3-2-1 และวิธีการ Randomesque (McBride & Martin, 1983; Kingsbury & Zara, 1989)

**2. การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป**

การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ดำเนินการโดยศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปที่แตกต่างกัน 4 วิธี ได้แก่ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (วิธีการ MIC) 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (วิธีการ RDM) 3) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ (วิธีการ HC) และ 4) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ (วิธีการ HC-Ex) ซึ่งเปรียบเทียบประสิทธิภาพด้าน 1) การประมาณค่าความสามารถของผู้สอบ 2) ความยาวของ

แบบทดสอบ และ 3) จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 โดยแบ่งการอภิปรายผลการวิจัยเป็น 3 ประเด็นย่อย ดังนี้

2.1 การเปรียบเทียบประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบ ปรากฏว่า เมื่อพิจารณาจากค่า RMSE วิธีการ RDM มีประสิทธิภาพสูงสุด รองลงมา ได้แก่ วิธีการ HC ซึ่งมีประสิทธิภาพเทียบเท่ากับวิธีการ HC-Ex ส่วนวิธีการ MIC มีประสิทธิภาพต่ำสุด ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 1.1 บางส่วน แต่เมื่อพิจารณาจากค่าความลำเอียงเฉลี่ย วิธีการ RDM, HC และ HC-Ex มีประสิทธิภาพสูงสุดเทียบเท่ากัน ส่วนวิธีการ MIC มีประสิทธิภาพต่ำสุด แสดงให้เห็นว่า วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ ได้แก่ วิธีการ RDM, HC และ HC-Ex มีประสิทธิภาพสูงกว่าวิธีการ MIC ซึ่งเป็นวิธีการที่นิยมใช้ทั่วไป เนื่องจาก วิธีการ RDM, HC และ HC-Ex ต่างถูกพัฒนาขึ้นโดยใช้ทฤษฎีการตัดสินใจ ควบคู่กับทฤษฎี IRT ซึ่งทฤษฎีการตัดสินใจจะคำนวณผลตอบแทนที่คาดว่าจะได้รับของทุกทางเลือก แล้วตัดสินใจเลือกทางเลือกที่ให้ผลตอบแทนที่ดีที่สุด เมื่อนำหลักการนี้มาใช้กับการคัดเลือกข้อสอบข้อถัดไป อย่างเช่น วิธีการ HC กับวิธีการ HC-Ex ที่พัฒนาขึ้น ได้คำนวณค่า SEE ที่คาดว่าจะเกิดขึ้น เมื่อผู้สอบทำข้อสอบแต่ละข้อในคลังข้อสอบ โดยคำนวณค่า SEE ทั้งกรณีที่ผู้สอบตอบถูก และกรณีที่ผู้สอบตอบผิด แล้วนำมาคำนวณค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE และตัดสินใจเลือกข้อสอบที่มีค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE ต่ำสุด เป็นข้อสอบข้อถัดไป ด้วยกระบวนการนี้ ข้อสอบข้อถัดไปที่ผู้สอบได้รับจึงเป็นข้อสอบที่มีค่า SEE ต่ำสุด ประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบ จึงสูงกว่าวิธีการ MIC ที่พิจารณาเพียงค่าสารสนเทศสูงสุด ณ ความสามารถที่ตรงหรือใกล้เคียงกับค่าความสามารถของผู้สอบ ณ ขณะนั้น (Taylor, 2009, pp. 538-539; Birnbaum, 1968 cited in van der Linden & Glas, 2002, pp. 9-10)

เมื่อพิจารณาวิธีการ RDM, HC และ HC-Ex ซึ่งใช้ทฤษฎีการตัดสินใจในการพัฒนา ผลการวิจัยแสดงให้เห็นว่า วิธีการ HC และ HC-Ex ที่ผู้วิจัยพัฒนาขึ้นมีประสิทธิต่ำกว่าวิธีการ RDM เป็นผลมาจาก วิธีการ RDM เป็นวิธีการที่พัฒนาจากทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (Risk Decision Making) ซึ่งใช้ค่าความน่าจะเป็นมาคำนวณผลตอบแทนที่คาดว่าจะเกิดขึ้น ผลลัพธ์ที่ได้จึงมีความแม่นยำมากกว่า วิธีการ HC และ HC-Ex ที่พัฒนาจากทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน (Uncertainly Decision Making) ที่ใช้เพียงค่าเฉลี่ยถ่วงน้ำหนักมาคำนวณผลตอบแทนที่คาดว่าจะเกิดขึ้น (Taylor, 2009, pp. 539-547) แต่ทั้งนี้ การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป ผู้วิจัยมุ่งหมายพัฒนาโดยใช้ทฤษฎีการตัดสินใจ ที่มีการคำนวณไม่ยุ่งยากซับซ้อนเท่ากับการใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง ซึ่งเกณฑ์ของเฮอร์วิคซ์ในทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอน ใช้ค่าเฉลี่ยถ่วงน้ำหนัก ถึงแม้จะมีความแม่นยำน้อยกว่า แต่ให้ประสิทธิภาพใกล้เคียงกัน เห็นได้จาก เมื่อพิจารณาจากค่าความลำเอียงเฉลี่ย ประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบของวิธีการ RDM, HC และ HC-Ex มีประสิทธิภาพสูงเทียบเท่ากัน

2.2 การเปรียบเทียบประสิทธิภาพด้านความยาวของแบบทดสอบ ปรากฏว่า วิธีการ RDM มีประสิทธิภาพสูงสุดเทียบเท่ากับวิธีการ HC รองลงมา ได้แก่ วิธีการ HC-Ex และ MIC ตามลำดับ ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 2 บางส่วน แสดงให้เห็นว่า วิธีการ HC ที่พัฒนาขึ้นใช้ข้อสอบในการทดสอบจำนวนน้อย เนื่องจาก วิธีการนี้ใช้ทฤษฎีการตัดสินใจในสภาวการณ์

ที่ไม่แน่นอน ตามเกณฑ์ของเฮอร์วิคซ์ โดยเลือกข้อสอบที่มีค่าเฉลี่ยถ่วงน้ำหนักของค่า SEE ต่ำสุด เป็นข้อสอบข้อถัดไป ทำให้ข้อสอบที่ผู้สอบได้รับมีความเหมาะสมกับระดับความสามารถของผู้สอบ สอดคล้องกับงานวิจัยของ โสพส สุขานนท์สวัสดิ์ และคณะ (2556) ได้พัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง ผลการวิจัยปรากฏว่า วิธีการที่พัฒนาขึ้นสามารถลดจำนวนข้อสอบและเวลาที่ใช้ในการทดสอบลงได้

2.3 การเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ปรากฏว่า วิธีการ MIC มีประสิทธิภาพสูงสุด รองลงมา ได้แก่ วิธีการ HC-Ex และวิธีการ RDM มีประสิทธิภาพต่ำสุดเทียบเท่าวิธีการ HC ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 3 บางส่วน แสดงให้เห็นว่า วิธีการ HC-Ex เป็นวิธีการที่มีประสิทธิภาพค่อนข้างสูง สามารถควบคุมข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ให้มีจำนวนลดลงได้ โดยข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ซึ่งวิธีการนี้ ทำให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง มีโอกาสถูกเลือกใช้ในการทดสอบน้อยลง โดยการสุ่มข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงจำนวน 5-10 ข้อ ด้วยวิธีการสุ่มตัวอย่างแบบมีระบบ แล้วนำข้อสอบที่สุ่มได้นี้ไปใช้ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ซึ่งการสุ่มข้อสอบนี้ จะทำการสุ่มใหม่ทุกครั้งก่อนเข้าสู่ขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ทำให้ข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง ถูกหมุนเวียนนำไปใช้ในการคัดเลือกข้อสอบข้อถัดไป โอกาสถูกเลือกใช้จึงลดลง ซึ่งสอดคล้องกับงานวิจัยของ Ozturk and Dogan (2015) ที่ได้เปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบ ปรากฏว่า วิธีการ Randomesque เป็นวิธีการควบคุมการใช้ข้อสอบด้วยวิธีการสุ่ม สามารถลดจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ให้น้อยลงได้

**3. การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในงานวิจัยนี้ เป็นโปรแกรมสำหรับวัดความรู้ในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี มีลักษณะเป็น Web Application จึงสามารถใช้ผ่านอุปกรณ์ทั่วไปที่เชื่อมต่อกับอินเทอร์เน็ตได้ ซึ่งโปรแกรมสามารถเลือกการทดสอบได้ 2 แบบ คือ 1) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการทดสอบที่ผู้สอบแต่ละคนจะได้ข้อสอบแตกต่างกัน ขึ้นอยู่กับความสามารถของผู้สอบ และ 2) การทดสอบด้วยคอมพิวเตอร์ เป็นการทดสอบที่ผู้สอบทุกคนจะได้ข้อสอบข้อเดียวกัน และมีจำนวนข้อสอบเท่ากัน

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น ในแต่ละขั้นตอนมีจุดเด่นดังนี้ ขั้นตอนที่ 1 การสร้างคลังข้อสอบ โปรแกรมนี้รองรับข้อสอบที่มีลักษณะการตรวจให้คะแนน 2 ค่า คือ ถูก-ผิด หรือ 0-1 และมีรายการคำตอบอย่างน้อย 4 รายการ จึงสามารถนำโปรแกรมไปปรับใช้วัดความรู้ทางการศึกษาในรายวิชาอื่นได้ ที่ข้อสอบมีลักษณะเดียวกัน ขั้นตอนที่ 2 การคัดเลือกข้อสอบข้อแรก คัดเลือกโดยสุ่มข้อสอบในคลังข้อสอบที่มีค่าความยากของข้อสอบระหว่าง -1.0 ถึง 1.0 ซึ่งวิธีการนี้มีประสิทธิภาพสูงกว่าวิธีการเลือกข้อสอบใดข้อหนึ่ง ที่มีค่าความยากของข้อสอบปานกลาง มาเป็นข้อสอบข้อแรก เพราะสามารถลดปัญหาข้อสอบข้อแรกถูกเลือกใช้บ่อยครั้งเกินไปได้ เมื่อเทียบกับการเลือกข้อสอบที่มีค่าความยากของข้อสอบระดับปานกลางข้อใดข้อหนึ่ง ขั้นตอนที่ 3 การคัดเลือกข้อสอบข้อถัดไป เลือกใช้วิธีการ HC-Ex ซึ่งเป็นวิธีการที่มีประสิทธิภาพ

สามารถประมาณค่าความสามารถของผู้สอบได้แม่นยำ ลดความยาวของแบบทดสอบ และลดจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ได้ ขั้นตอนที่ 4 การประมาณค่าความสามารถของผู้สอบ เลือกใช้วิธีของเบส์แบบปรับใหม่ (Bayesian Updating) ซึ่งเป็นวิธีที่ได้รับความนิยมนำไปใช้อย่างแพร่หลาย และเป็นวิธีที่สามารถคำนวณประมาณค่าความสามารถของผู้สอบได้ ในกรณีที่ผู้สอบตอบถูกทุกข้อ หรือตอบผิดทุกข้อ ซึ่งวิธีการอื่นไม่สามารถคำนวณค่าได้ และขั้นตอนที่ 5 เกณฑ์ยุติการทดสอบ เลือกใช้เกณฑ์ที่ค่า SEE น้อยกว่า 0.3 ซึ่งเป็นเกณฑ์ที่มีความน่าเชื่อถือ และลดความยาวของแบบทดสอบได้มาก (Lilley et al., 2004; Thompson & Weiss, 2011; Wainer et al., 2000, pp. 105-113)

การคัดเลือกข้อสอบข้อถัดไปที่ใช้ในโปรแกรมนี้ นอกจากต้องคำนึงถึงการควบคุมการใช้ข้อสอบ ยังต้องคำนึงถึงการควบคุมสัดส่วนเนื้อหาของข้อสอบ (Content Balancing ) เพื่อป้องกันไม่ให้ผู้สอบได้รับข้อสอบที่มีเนื้อหา (หัวข้อเรื่อง) ด้านใดด้านหนึ่งเพียงด้านเดียว โดยเฉพาะอย่างยิ่งการวัดผลการเรียนรู้ของผู้เรียน ซึ่งมีเนื้อหาในการเรียนที่หลากหลาย ซึ่งในโปรแกรมนี้ควบคุมสัดส่วนเนื้อหาของข้อสอบ ด้วยวิธีการของ Kingsbury-Zara (1989) ดังนั้น โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นนี้ จึงมีประสิทธิภาพสูงเมื่อนำไปใช้งานจริง เนื่องจาก ใช้วิธีการ HC-Ex ที่มีประสิทธิภาพสูงในการประมาณค่าความสามารถของผู้สอบ ลดความยาวของแบบทดสอบ มีการควบคุมการใช้ข้อสอบ และมีการควบคุมสัดส่วนเนื้อหาของข้อสอบ

**4. การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปที่ได้พัฒนาขึ้นกับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์**

ผลการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปที่ได้พัฒนาขึ้นกับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ที่ใช้ข้อสอบคงที่ จำนวน 40 ข้อ ปรากฏว่า เป็นไปตามสมมติฐานของการวิจัยข้อ 4 นั่นคือ ผลการทดสอบทั้งสองแบบมีความสัมพันธ์เชิงบวก ขนาดปานกลาง ซึ่งแสดงให้เห็นว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์ให้ผลลัพธ์ในการทดสอบเป็นไปในทิศทางเดียวกัน แต่ทั้งนี้ ความสัมพันธ์ที่ศึกษาได้ มีขนาดปานกลาง ซึ่งเป็นผลมาจากวิธีการคัดเลือกข้อสอบข้อถัดไปที่ใช้ในโปรแกรมมีการควบคุมการใช้ข้อสอบ และควบคุมสัดส่วนเนื้อหาของข้อสอบ ทำให้ประสิทธิภาพในการคัดเลือกข้อสอบลดลง เนื่องจาก การควบคุมการใช้ข้อสอบและการควบคุมสัดส่วนเนื้อหาของข้อสอบ จะเข้าแทรกแซงกระบวนการคัดเลือกข้อสอบ (Thompson & Weiss, 2011) ซึ่งสอดคล้องกับผลการวิจัยของ Lilley et al. (2004) ที่ได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวัดความรู้ทางภาษาอังกฤษของนักศึกษานานาชาติ ใน University of Hertfordshire ประเทศอังกฤษ ผลปรากฏว่า ผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ธ์กับผลการทดสอบด้วยคอมพิวเตอร์ มีความสัมพันธ์เชิงบวก ขนาดปานกลาง

## ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. ผู้ที่สนใจพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สามารถนำผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์ในการวิจัยนี้ ไปใช้เป็นแนวทางในการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปวิธีการใหม่ ที่ใช้เกณฑ์อื่นในทฤษฎีการตัดสินใจในสภาวการณ์ที่ไม่แน่นอนได้ เช่น ใช้เกณฑ์แมกซิแมก เกณฑ์แมกซิมิน หรือเกณฑ์ลาปลาซ

2. ผู้ที่สนใจพัฒนาวิธีการควบคุมการใช้ข้อสอบ ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สามารถนำผลการพัฒนาวิธีการควบคุมการใช้ข้อสอบโดยใช้วิธีการสุ่มตัวอย่างแบบมีระบบของการวิจัยนี้ ไปใช้เป็นแนวทางในการพัฒนาวิธีการควบคุมการใช้ข้อสอบวิธีการใหม่ ที่ใช้วิธีการสุ่มตัวอย่างวิธีการอื่นๆ ได้ เช่น วิธีการสุ่มตัวอย่างแบบแบ่งชั้นภูมิ (Stratified Random Sampling)

3. ครู อาจารย์ หรือผู้ที่เกี่ยวข้องในการจัดการสอบ สามารถนำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี ที่ได้จากการวิจัยนี้ ไปใช้วัดความสามารถทางภาษาอังกฤษ ในระดับปริญญาตรี หรือนำไปประยุกต์กับการวัดผลการศึกษาในรายวิชาอื่นได้

## ข้อเสนอแนะสำหรับการวิจัยต่อไป

1. การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ เป็นการพัฒนากระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไปเท่านั้น ดังนั้นควรพัฒนากระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในขั้นตอนอื่นด้วย เช่น การสร้างคลังข้อสอบ การคัดเลือกข้อสอบข้อแรก การประมาณค่าความสามารถของผู้สอบ หรือเกณฑ์ยุติการทดสอบ

2. วิธีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้นในการวิจัยนี้ มุ่งเน้นเข้าควบคุมการใช้ข้อสอบที่มีอัตราการใช้ข้อสอบสูง (Over-Exposure Item) ให้ลดลงได้ แต่ไม่ได้ควบคุมข้อสอบที่ถูกเลือกใช้น้อยหรือไม่ถูกเลือกใช้เลย (Under-Utilized Item) ให้มีอัตราการใช้ข้อสอบเพิ่มขึ้น จึงควรพัฒนาวิธีการควบคุมการใช้ข้อสอบ ให้ครอบคลุมกรณีข้อสอบที่ถูกเลือกใช้น้อยหรือไม่ถูกเลือกใช้เลย

3. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นในการวิจัยนี้ ใช้กับข้อสอบที่มีลักษณะการตรวจให้คะแนน 2 ค่า คือ ถูก-ผิด หรือ 0-1 และมีรายการคำตอบอย่างน้อย 4 ตัวเลือก และไม่รองรับข้อสอบวัดการฟังในรูปแบบเสียง หรือข้อสอบวัดการอ่านในรูปแบบที่หนึ่งบทความใช้กับข้อสอบหลายข้อ จึงควรพัฒนาโปรแกรมให้รองรับรูปแบบการทดสอบที่หลากหลายมากกว่านี้

4. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นในการวิจัยนี้ เหมาะสำหรับข้อสอบที่ผ่านการวิเคราะห์ข้อสอบตามหลักทฤษฎีการตอบสนองข้อสอบมาแล้ว นั่นคือ เป็นข้อสอบที่ทราบค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ ดังนั้นเพื่อความสะดวกในการใช้งาน จึงควรพัฒนาโปรแกรมให้สามารถวิเคราะห์ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบได้

# บรรณานุกรม


กิตติ ภักดีวัฒนะกุล และพนิดา พานิชกุล. (2554). *การวิเคราะห์เชิงปริมาณเพื่อการตัดสินใจ*. กรุงเทพฯ: เคทีพี คอมพ์ แอนด์ คอนซัลท์.

ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการตอบสนองแนวใหม่* (พิมพ์ครั้งที่ 4). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

โสพส สุขานนท์สวัสดิ์, เสรี ชัดแช้ม และกฤษณะ ชินสาร. (2556). การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์. *วิทยาการวิจัยและวิทยาการปัญญา, 10*(2), 71-85.

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Martin, K. (2012). *An Introduction to Management Science: Quantitative Approaches to Decision Making* (13th ed.). Mason, OH: South-Western Cengage Learning.

Antal, M., Eros, L., & Imre, A. (2010). Computerized adaptive testing: implementation issues. *Informatica, 2*(2), 168-183.

Baker, F. B. (2001). *The Basics of Item Response Theory* (2nd ed.). USA: ERIC Clearinghouse on Assessment and Evaluation.

Barrada, J. R., Mazuela, P., & Olea, J. (2006). Maximum information stratification method for controlling item exposure in computerized adaptive testing. *Psicothema, 18*(1), 156-159.

Belov, D. I., Armstrong, R. D., & Weissman, A. (2008). A Monte Carlo Approach for Adaptive Testing With Content Constraints. *Applied Psychological Measurement, 32*(6), 431-446.

Chaimongkol, N., Pasiphol, S., & Kanjanawasee, S. (2016). Computerized Adaptive Testing with Reflective Feedback: A Conceptual Framework. *Procedia - Social and Behavioral Sciences, 217*, 806-812.

Chalhoub-Deville, M., Alcaya, C., & Lozier, V. M. (1996). *An Operational Framework for Constructing a Computer-Adaptive Test of L2 Reading Ability: Theoretical and Practical Issues*. Minneapolis, MN: University of Minnesota.

Chang, H. H., Qian, J., & Ying, Z. (2001). a-Stratified Multistage Computerized Adaptive Testing with b Blocking. *Applied Psychological Measurement, 25*(4), 333-341.

Chang, H. H., & van der Linden, W. J. (2003). Optimal Stratification of Item Pools in a-Stratified Computerized Adaptive Testion. *Applied Psychological Measurement*, *27*(4), 262 – 274.

Chang, H. H., & Ying, Z. (1996). A global information approach to computerized adaptive testing. *Applied Psychological Measurement*, *20*, 213–229.

Chang, H. H., & Ying, Z. (1999). A-stratified multistage computerized adaptive testing. *Applied Psychological Measurement, 23*(3), 211-222.

Chang, S. W., & Harris, D. J., (2002). *Redeveloping the exposure control parameters of CAT items when a pool is modified*. Paper presented at the annual meeting of the American Educational Research Association. New Orleans, LA: IACAT.

Chang, S. W., & Ansley, T. N. (2003). A Comparative Study of Item Exposure Control Methods in Computerized Adaptive Testing. *Journal of Educational Measurement, 40*(1), 71-103.

Chen, S. Y., & Ankenmann, R. D. (2004). Effects of practical constraints on item selection rules at the early stages of computerized adaptive testing. *Journal of Educational Measurement, 41*(2), 149-174.

Chen, S. Y., & Doong, S. H. (2003). *Predicting Item Exposure Parameters in Computerized Adaptive Testing*. Paper presented at the 2003 annual meeting of the American Educational Research Association. Chicago, IL: IACAT.

Cheng, Y., & Chang, H. H. (2009). The maximum priority index method for severely constrained item selection in computerized adaptive testing. *Br J Math Stat Psychol, 62*(2), 369-383.

Cheng. Y., Chang, H. H., Douglas, J., & Guo, F. (2009). Constraint-Weighted a-Stratification for Computerized Adaptive Testing With Nonstatistical Constraints. *Educational and Psychological Measurement, 69*(1). 35-49.

Choi, S. W. (2009). Firestar: Computerized adaptive testing (CAT) simulation program for polytomous IRT Models (Computer software). *Applied Psychological Measurement, 33*, 644–645.

Costa, D. R., Karino, C. A., Moura, F. A. S., & Andrade, D. F. (2009). A comparison of three methods of item selection for computerized adaptive testing. In D. J. Weiss (Ed.), *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing*. Retrieved from www.psych.umn.edu/psylabs/CATCentral/

Davey, T., & Parshall, C. G. (1995). New algorithms for item selection and exposure control with computerized adaptive testing. In *the annual meeting of the American Educational Research Association*. San Francisco, CA: IACAT.

de Ayala, R. J. (2009). *The theory and practice of item response theory*. London: The Guilford Press.

Deng, H., Ansley, T., & Chang, H. H. (2010). Stratified and Maximum Information Item Selection Procedures in Computer Adaptive Testing. *Journal of Education Measurement, 47*(2), 202-226.

Doong, S. H. (2009). A Knowledge-Based Approach for Item Exposure Control in Computerized Adaptive Testing. *Journal of Educational and Behavioral Statistics, 34*(4), 530-558.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.

Fan, M., & Zhu, D. (2002). *A Further Study on Adjusting CAT Item Selection Starting Point for Individual Examinees*. Paper presented at the annual meeting of the American Educational Research Association. New Orleans, LA: IACAT.

Fan, Z., Wang, C., Chang, H. H., & Douglas, J., (2012). Utilizing Response Time Distributions for Item Selection in CAT. *Journal of Educational and Behavioral Statistics, 37*(5), 655-670.

Frenkel, D. (2004). Introduction to Monte Carlo Methods. *NIC Series, 23*, 29-60.

Gao, F., & Chen, L. (2005). Bayesian or Non-Bayesian: A Comparison Study of Item Parameter Estimation in the Three-Parameter Logistic Model. *Applied Measurement in Education, 18*(4), 351-380.

Gaspars-Wieloch, H. (2014). Modifications of the Hurwicz's decision rule. *Central European Journal of Operations Research, 22*(4), 779-794.

Gaspars-Wieloch, H. (2015). On a decision rule supported by a forecasting stage based on the decision maker's coefficient of optimism. *Central European Journal of Operations Research, 23*(3), 579-594.

Georgiadou, E., Triantafillou, E., & Economides, A. (2007). A review of item exposure control strategies for computerized adaptive testing developed from 1983 to 2005. *Journal of Technology, Learning, and Assessment, 5*(8), 1-38.

Groenewald, M. E., & Pretorius, P. D. (2011). Comparison of Decision-making under Uncertainty Investment Strategies with the Money Market. *Journal of Financial Studies and Research*, 1-16.

Harwell, M., Stone, C. A., Hsu, T. C., & Kirisci, L. (1996). Monte Carlo Studies in Item Response Theory. *Applied Psychological Measurement, 20*(2), 101-125.

Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Boston: Kluwer Nijhoff.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, California: SAGE.

Han, K. T. (2007). WinGen: Windows software that generates IRT parameters and item responses. *Applied Psychological Measurement, 31*(5), 457-459.

Han, K. T. (2012). An Efficiency Balanced Information Criterion for Item Selection in Computerized Adaptive Testing. *Journal of Educational Measurement, 49*(3), 22-246.

Hau, K. T., & Chang, H. H. (2001). Item Selection in Computerized Adaptive Testing: Should More Discriminating Item be Used First?. *Journal of Educational Measurement, 38*(3), 249-266.

Kingsbury, G. G., & Zara, A. R. (1989). Procedures for selecting items for computerized adaptive tests. *Applied Measurement in Education, 2*(4), 359-375.

Koedsri, A., Lawthong, N., & Ngudgratoke, S. (2014). Efficiency of Item Selection Method in Variable-Length Computerized Adaptive Testing for the Testlet Response Model: Constraint-Weighted A-Statification Method. *Procedia - Social and Behavioral Sciences, 116*, 1890-1895.

Lee, Y. H., Ip, E. H., & Fuht, C. D. (2008). A strategy for Controlling Item Exposure in Multidimensional Computerized Adaptive Testing. *Educational and Psychological Measurement, 68*(2), 215-232.

Leroux, A. J., Lopez, M., Hembry, I., & Dodd, B. G. (2013). A Comparison of Exposure Control Procedures in CATs Using the 3PL Model. *Educational and Psychological Measurement, 73*(5), 857-874.

Leung, C. K., Chang, H. H., & Hau, K. T. (2000). *Solving Complex Constraints in a-Sratified Computerized Adaptive Testing Designs*. Paper presented at the National Council on Measurement in Education Annual Meeting 2000. New Orleans, LA: NCME.

Leung, C. K., Chang, H. H., & Hau, K. T. (2002). Item Selection in Computerized Adaptive Testing: Improving the a-Stratified Design with the Sympson-Heter Algorithm. *Applied Psychological Measurement, 26*(4), 376-392.

Leung, C. K., Chang, H. H., & Hau, K. T. (2003). Incorporation of Content Balancing Requirements in Stratification Designs for Computerized Adaptive Testing. *Educational and Psychological Measurement, 63*(2), 257-270.

Lilley, M., Barker, T., & Britton, C. (2004). The development and evaluation of a software prototype for computer-adaptive testing. *Computers & Education, 43*, 109-123.

Linn, R. L. (1993). *Educational Measurement* (3rd ed.). Phoenix, AZ: The Oryx Press.

Lord, F. M. (1971). A theoretical study of the measurement effectiveness of flexilevel test. *Educational and Psychological Measurement*, *31*, 805-813.

Lord, F. M. (1977). A broad-range tailored test of verbal ability. *Applied Psychological Measurement, 1*, 95-100.

Luecht, R. M., Nungester, R. J., & Hadadi, A. (1996). *Heuristics based CAT: Balancing item information, content, and exposure*. Paper presented at the annual meeting of the National Council on Measurement in Education. New York: MCNE.

Luecht, R. M., & Sireci, S. G. (2011). *A Preview of Models for Computer-Based Testing.* New York: The College Board. Retrieved from https://research. collegeboard.org/publications/content/2012/05/review-models-computer-based-testing

Lunz, M. E., Bergstrom, B. A., & Wright, B. D. (1992). The Effect of Review on Student Ability and Test Efficiency for Computerized Adaptive Tests. *Applied Psychological Measurement, 16*(1), 33-40.

Lunz, M. E., & Stahl, J.A. (1998). Patterns of item exposure using a randomized CAT algorithm. Paper presented at the annual meeting of the National Council on Measurement in Education. San Diego, CA: NCME.

McBride, J. R., & Martin, J. T. (1983). Reliability and validity of adaptive ability tests in a military setting. In D. J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing* (pp. 223-226). New York: Academic Press.

Molina, J. G., Pareja, I., & Sanmartín, J. (2008). Modeling item banking: Analysis and design of a computerized system. *Revista Electrónica de Metodología Aplicada, 13*(2), 1-14.

Murphy, D. L., Dodd, B. G., & Vaughn, B. K. (2010). A Comparison of Item Selection Techniques for Testlet. *Applied Psychological Measurement, 34*(6), 424-437.

Owen, R. 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.

Ozturk, N. B., & Dogan, N. (2015). Investigating Item Exposure Control Methods in Computerized Adaptive Testing. *Educational Sciences: Theory & Practice, 15*(1), 85-98.

Parshall, C. G., Spray, J. A., Kalohn, J. C., & Davey, T. (2002). *Practical Considerations in Computer-Based Testing*. New York: Springer-Verlag.

Pazek, K., & Rozman, C. (2009). Decision making under conditions of uncertainty in agriculture: a case study of oil crops. *Poljoprivreda, 15*(1), 45-50.

Polit, D. F., & Beck, C. T. (2004). *Nursing Research: Principles and Methods*. Philadelphia, PA: Lippincott Williams & Wilkins.

Raychaudhuri, S. (2008). Introduction to Monte Carlo Simulation. In S. J. Mason, R. R. Hill, L. Monch, O. Rose, T. Jefferson, & J. W. Fowler (eds.), *Proceedings of the 2008 Winter Simulation Conference* (pp. 91-100). Boomfield, CO: IEEE Conferece.

Revuelta, J., & Ponsoda, V. (1998). A comparison of item exposure control methods in computerized adaptive testing. *Journal of Educational Measurement, 35*(4). 311-327.

Stocking, M. L. (1993). *Controlling item exposure rates in a realistic adaptive testing paradigm* (ETS Research Report RR-93-2). Princeton, NJ: Educational Testing Service.

Sympson, J. B., & Hetter, R. D. (1985). Controlling item-exposure rates in computerized adaptive testing. In Proceedings of the 27th annual meeting of the Military Testing Association (pp. 973-977). San Diego, CA: Navy Personnel Research and Development Center.

Taylor, B. W. (2009). *Introduction to Management Science* (11th ed.). Upper Saddle River, NJ: Prentice Hall.

Thompson, N. A., & Weiss, D. J. (2011). A Framework for the Development of Computerized Adaptive Tests. *Practical Assessment, Research & Evaluation, 16*(1), 1-9.

Tseng, W. T. (2016). Measuring English Vocabulary Size via Computerized Adaptive Testing. *Computers & Education, 97*, 69-85.

van der Linden, W. J. (2003). Some Alternatives to Sympson-Hetter Item-Exposure Control in Computerized Adaptive Testing. *Journal of Educational and Behavioral Statistics, 28*(3), 249-265.

van der Linden, W. J., & Glas, G. A. W. (2002). *Computerized Adaptive Testing: Theory and Practice*. New York: Kluwer Academic.

Veerkamp, W. J. J., & Berger, M. P. F. (1997). Some new item selection criteria for adaptive testing. *Journal of Educational and Behavioral Statistics, 22*, 203-226.

Veldkamp, B. P., & van der Linden, W. J. (2002). Multidimensional adaptive testing with constraints on test content. *Psychometrika, 67*(4), 575-588.

Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., & Mislevy, R. J. (2000). *Computerized Adaptive Testing: A Primer* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Wang, H. P., Kuo, B. C., Chao, R. C., & Tsai, Y. H. (2012). The Development and Evaluation of a Computerized Adaptive Testing System for Chinese Proficiency – Base on CEFR. *Procedia-Social and Behavioral Sciences, 64*, 34-42.

Weiss, D. J. (1982). Improving measurement quality and efficiency with adaptive testing. *Applied Psychological Measurement, 6*, 473-492.

Weiss, D. J. (2011). Better Data From Better Measurements Using Computerized Adaptive Testing. *Journal of Methods and Measurement in the Social Sciences, 2*(1), 1-27.

Weiss, D. J., & Guyer, R. (2012). Manual for CATSim: Comprehensive simulation of computerized adaptive testing. Saint Paul, MN: Assessment Systems Corporation.

Yan, D., von Davier, A. A., & Lewis, C. (2014). Computerized Multistage Testing: Theory and Applications. Princeton, NJ: CRC Press.

Yoes, M. (1997). PARDSIM (Computer software). Saint Paul, MN: Assessment Systems Corporation.

# ภาคผนวก

# ภาคผนวก ก
# รายชื่อผู้เชี่ยวชาญ

# รายชื่อผู้เชี่ยวชาญ

รายชื่อผู้เชี่ยวชาญสำหรับการประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีดังนี้

1) ผู้ช่วยศาสตราจารย์ ดร. พีรพัฒน์ ยางกลาง อาจารย์ประจำคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร

2) ผู้ช่วยศาสตราจารย์ ดร. สิริชัย ดีเลิศ อาจารย์ประจำคณะวิทยาการจัดการ มหาวิทยาลัยศิลปากร

3) อาจารย์ ดร. สุชาดา สกลกิจรุ่งโรจน์ อาจารย์ประจำสำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช

# ภาคผนวก ข
# ผลการศึกษาในสถานการณ์จำลอง
# การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ข1. ค่าความสามารถที่แท้จริงของผู้สอบชุดที่ 1 ค่าประมาณความสามารถของผู้สอบและความยาวของแบบทดสอบของการศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ข2. การคำนวณค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย

ข3. การคำนวณค่าความลำเอียงเฉลี่ย

**ผลการศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

ตารางที่ ข-1 ค่าความสามารถที่แท้จริงของผู้สอบชุดที่ 1 ค่าประมาณความสามารถของผู้สอบ และความยาวของแบบทดสอบ ที่ได้จากการศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 1 | 0.552 | 0.609 | 0.606 | 0.554 | 0.235 | 6 | 7 | 6 | 6 |
| 2 | 0.151 | 0.558 | 0.387 | 0.181 | 0.361 | 11 | 8 | 9 | 11 |
| 3 | -0.280 | -0.314 | -0.449 | -0.543 | -0.407 | 9 | 8 | 6 | 12 |
| 4 | 0.003 | 0.074 | 0.446 | 0.141 | 0.179 | 13 | 8 | 8 | 11 |
| 5 | -0.785 | -0.666 | -0.786 | -0.661 | -1.115 | 9 | 6 | 6 | 10 |
| 6 | -1.377 | -1.118 | -1.489 | -1.325 | -1.070 | 15 | 6 | 8 | 8 |
| 7 | -0.115 | -0.508 | 0.040 | -0.285 | -0.412 | 9 | 10 | 5 | 10 |
| 8 | -0.255 | 0.055 | -0.288 | -0.285 | -0.678 | 11 | 8 | 8 | 8 |
| 9 | -0.429 | -0.637 | -0.681 | -0.371 | 0.136 | 13 | 7 | 7 | 8 |
| 10 | -1.564 | -0.204 | -1.185 | -1.485 | -0.793 | 17 | 5 | 11 | 6 |
| 11 | -1.932 | -0.714 | -1.699 | -1.539 | -1.733 | 13 | 8 | 8 | 11 |
| 12 | 2.336 | 1.667 | 1.416 | 2.104 | 1.946 | 12 | 6 | 10 | 15 |
| 13 | 0.204 | 0.180 | 0.503 | 0.416 | 0.191 | 11 | 8 | 6 | 9 |
| 14 | 0.114 | 0.332 | 0.213 | 0.325 | 0.212 | 11 | 10 | 5 | 11 |
| 15 | 0.854 | 1.372 | 0.734 | 0.625 | 0.645 | 11 | 7 | 6 | 11 |
| 16 | 0.560 | 0.577 | 0.636 | 0.572 | 0.549 | 6 | 5 | 4 | 14 |
| 17 | -0.016 | -0.393 | -0.188 | 0.114 | -0.214 | 12 | 9 | 13 | 7 |
| 18 | 0.808 | 0.430 | 0.548 | 0.776 | 0.911 | 10 | 7 | 6 | 8 |
| 19 | -0.605 | -0.396 | -0.713 | -1.068 | -0.735 | 11 | 6 | 7 | 8 |
| 20 | 0.783 | -0.354 | 0.468 | 0.681 | 0.870 | 9 | 6 | 6 | 5 |
| 21 | -2.145 | -1.510 | -2.041 | -1.878 | -1.621 | 9 | 8 | 8 | 8 |
| 22 | 1.210 | 1.738 | 0.746 | 0.639 | 1.210 | 7 | 5 | 5 | 12 |
| 23 | 0.157 | 0.485 | 0.149 | 0.445 | 0.427 | 9 | 9 | 6 | 9 |
| 24 | 0.243 | -0.160 | 0.048 | 0.616 | -0.105 | 11 | 9 | 6 | 15 |
| 25 | -0.734 | -0.936 | -0.671 | -0.557 | -0.778 | 8 | 8 | 8 | 8 |
| 26 | -1.079 | -0.744 | -1.465 | -0.979 | -0.977 | 16 | 6 | 10 | 7 |
| 27 | -0.742 | -0.616 | -0.556 | -1.217 | -0.895 | 9 | 8 | 6 | 16 |
| 28 | 4.255 | 3.709 | 2.640 | 1.595 | 1.798 | 30 | 11 | 12 | 23 |
| 29 | -0.653 | -0.431 | -0.449 | -0.319 | -1.096 | 9 | 8 | 7 | 7 |
| 30 | -0.903 | -0.578 | -1.174 | -0.568 | -0.855 | 7 | 5 | 5 | 12 |
| 31 | -0.126 | -0.230 | 0.036 | -0.529 | -0.284 | 12 | 11 | 8 | 10 |
| 32 | 1.821 | 2.161 | 1.278 | 1.897 | 1.712 | 14 | 7 | 11 | 8 |
| 33 | 2.285 | 2.676 | 1.917 | 2.238 | 1.953 | 18 | 7 | 10 | 14 |
| 34 | -0.060 | 1.183 | 0.012 | -0.267 | 0.403 | 7 | 10 | 6 | 9 |
| 35 | -0.533 | -0.483 | -1.232 | -0.580 | 0.120 | 11 | 6 | 6 | 13 |
| 36 | -0.979 | -1.091 | -0.967 | -0.004 | -1.238 | 16 | 5 | 6 | 7 |
| 37 | 0.172 | -0.349 | 0.034 | 0.329 | -0.090 | 9 | 9 | 5 | 11 |
| 38 | -1.715 | -1.859 | -1.555 | -1.711 | -1.908 | 12 | 6 | 9 | 11 |
| 39 | 0.056 | 0.854 | -0.241 | -0.199 | -0.539 | 13 | 8 | 5 | 11 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 40 | 0.352 | 0.964 | 0.755 | 0.472 | 0.329 | 7 | 5 | 8 | 7 |
| 41 | 0.265 | -0.063 | 0.295 | 0.307 | 0.136 | 8 | 8 | 7 | 8 |
| 42 | 1.412 | 0.410 | 1.244 | 1.557 | 1.663 | 12 | 6 | 12 | 14 |
| 43 | -1.190 | 0.216 | -0.962 | -1.340 | 0.201 | 13 | 5 | 6 | 11 |
| 44 | -1.174 | -1.321 | -1.013 | -1.013 | -1.231 | 8 | 7 | 7 | 8 |
| 45 | -1.093 | -1.258 | -1.348 | -0.713 | -0.970 | 11 | 5 | 5 | 8 |
| 46 | -0.073 | -0.216 | -0.316 | -0.454 | -0.617 | 12 | 6 | 5 | 7 |
| 47 | 0.240 | 0.761 | 0.280 | 0.180 | 0.170 | 10 | 7 | 6 | 10 |
| 48 | -1.422 | -1.759 | -1.544 | -1.211 | -1.456 | 15 | 6 | 7 | 8 |
| 49 | 0.110 | -0.218 | 0.050 | 0.201 | 0.689 | 13 | 10 | 7 | 12 |
| 50 | 0.560 | 1.766 | 0.571 | 0.382 | 0.433 | 12 | 6 | 6 | 10 |
| 51 | -0.039 | -0.363 | -0.254 | -0.245 | -0.042 | 13 | 9 | 5 | 10 |
| 52 | 0.207 | -0.398 | -0.014 | 0.009 | 0.189 | 12 | 8 | 5 | 7 |
| 53 | -0.107 | -0.215 | -0.379 | -0.378 | 0.041 | 11 | 6 | 6 | 5 |
| 54 | -1.670 | -0.968 | -1.839 | -1.058 | -1.444 | 14 | 9 | 9 | 10 |
| 55 | -1.559 | -1.137 | -1.335 | -1.391 | -1.748 | 8 | 5 | 9 | 9 |
| 56 | -1.487 | -0.371 | -1.121 | -1.455 | -1.354 | 12 | 5 | 13 | 7 |
| 57 | -1.250 | 0.269 | -0.667 | -0.720 | -1.319 | 10 | 7 | 6 | 13 |
| 58 | 1.238 | 1.005 | 1.454 | 1.033 | 0.631 | 9 | 7 | 6 | 10 |
| 59 | -0.543 | -0.334 | -0.269 | -0.865 | -0.619 | 9 | 8 | 9 | 10 |
| 60 | 1.388 | 1.713 | 1.481 | 1.015 | 1.388 | 15 | 7 | 10 | 10 |
| 61 | 0.620 | 0.845 | 0.331 | 0.613 | 0.331 | 13 | 7 | 8 | 8 |
| 62 | -0.225 | 0.263 | -0.314 | -0.123 | -0.026 | 12 | 6 | 6 | 11 |
| 63 | 0.947 | 0.506 | 0.690 | 1.150 | 1.036 | 12 | 6 | 7 | 18 |
| 64 | -1.248 | -0.900 | -1.191 | -1.604 | -1.177 | 13 | 5 | 10 | 7 |
| 65 | 1.081 | 1.284 | 1.177 | 0.644 | 0.928 | 13 | 5 | 5 | 7 |
| 66 | -0.270 | -0.212 | -0.300 | -0.720 | 0.944 | 11 | 7 | 6 | 7 |
| 67 | -1.271 | -0.555 | -1.454 | -1.220 | -1.578 | 12 | 6 | 10 | 8 |
| 68 | -1.989 | -0.812 | -1.820 | -2.111 | -1.621 | 8 | 9 | 8 | 13 |
| 69 | -2.006 | -2.312 | -2.121 | -1.949 | -1.594 | 9 | 7 | 10 | 11 |
| 70 | -0.919 | -0.985 | -0.846 | -0.810 | -0.956 | 9 | 5 | 7 | 10 |
| 71 | 0.479 | 0.944 | 0.653 | 0.620 | 1.265 | 10 | 5 | 5 | 8 |
| 72 | 0.039 | -0.441 | 0.202 | -0.337 | -0.101 | 14 | 10 | 6 | 9 |
| 73 | -0.170 | -0.113 | 0.033 | -0.285 | -0.257 | 11 | 9 | 5 | 10 |
| 74 | -0.034 | -0.206 | -0.595 | 0.225 | -0.001 | 13 | 6 | 7 | 8 |
| 75 | 0.137 | -0.055 | 0.857 | 0.005 | -0.276 | 8 | 5 | 7 | 13 |
| 76 | -0.511 | -0.457 | -0.613 | -0.625 | -0.967 | 11 | 7 | 8 | 8 |
| 77 | -0.311 | -0.716 | 0.443 | -0.513 | -0.386 | 10 | 9 | 9 | 13 |
| 78 | 0.461 | 0.447 | 0.609 | -0.027 | 0.557 | 14 | 7 | 4 | 7 |
| 79 | 1.084 | 1.554 | 1.152 | 1.097 | 1.015 | 15 | 8 | 15 | 8 |
| 80 | -1.766 | -1.917 | -1.275 | -2.012 | -1.783 | 7 | 6 | 11 | 10 |
| 81 | 1.738 | 0.916 | 1.972 | 1.545 | 1.683 | 12 | 8 | 10 | 11 |
| 82 | 0.671 | 0.521 | 0.660 | 0.757 | 0.419 | 9 | 6 | 9 | 10 |
| 83 | 1.751 | 1.978 | 1.300 | 1.610 | 1.362 | 12 | 7 | 14 | 9 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า θ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 84 | 0.337 | 0.462 | 0.162 | -0.137 | 0.224 | 7 | 9 | 6 | 7 |
| 85 | 0.695 | 0.763 | 0.734 | 0.700 | 0.252 | 11 | 7 | 5 | 9 |
| 86 | 0.333 | 0.156 | 0.008 | 0.416 | 0.609 | 10 | 9 | 6 | 11 |
| 87 | -1.351 | -1.469 | -0.991 | -1.404 | -1.310 | 9 | 5 | 11 | 6 |
| 88 | -0.001 | -0.324 | 0.036 | -0.324 | -0.121 | 12 | 12 | 11 | 12 |
| 89 | 0.963 | -0.196 | 1.144 | 1.022 | 0.971 | 14 | 7 | 6 | 7 |
| 90 | -0.856 | -1.135 | -0.697 | -0.857 | -0.855 | 7 | 6 | 8 | 8 |
| 91 | 1.915 | 2.106 | 2.193 | 1.169 | 1.619 | 14 | 8 | 16 | 9 |
| 92 | -0.220 | -0.311 | -0.577 | -0.672 | -0.071 | 8 | 5 | 6 | 10 |
| 93 | 0.421 | -0.013 | -0.412 | 0.027 | 0.296 | 11 | 7 | 6 | 7 |
| 94 | -1.792 | -1.073 | -1.534 | -1.212 | -1.760 | 14 | 6 | 11 | 11 |
| 95 | -0.456 | 0.290 | -0.437 | -0.672 | -0.385 | 14 | 5 | 6 | 12 |
| 96 | 0.041 | -0.346 | 0.027 | -0.040 | -0.096 | 10 | 9 | 8 | 8 |
| 97 | 0.374 | 0.179 | -0.371 | 0.284 | -0.042 | 10 | 6 | 5 | 7 |
| 98 | -0.491 | 0.221 | -0.338 | -0.268 | -0.615 | 10 | 7 | 6 | 8 |
| 99 | -0.099 | 0.246 | 0.035 | 0.198 | 0.270 | 11 | 9 | 5 | 8 |
| 100 | -0.301 | -0.945 | 0.040 | -0.340 | -0.082 | 13 | 10 | 8 | 9 |
| 101 | 1.399 | 1.338 | 1.095 | 1.186 | 1.265 | 13 | 5 | 19 | 13 |
| 102 | 1.056 | 0.901 | 0.746 | 1.206 | 0.803 | 12 | 5 | 7 | 14 |
| 103 | -1.001 | -0.336 | -1.055 | -1.009 | -0.663 | 11 | 8 | 11 | 10 |
| 104 | -0.212 | 0.375 | -0.242 | -0.620 | -0.647 | 15 | 9 | 5 | 9 |
| 105 | 0.144 | -0.079 | -0.034 | 0.147 | 0.601 | 11 | 8 | 7 | 8 |
| 106 | 1.035 | 1.229 | 1.328 | 1.266 | 0.796 | 13 | 7 | 7 | 15 |
| 107 | -0.748 | -0.126 | -0.739 | -0.995 | -0.823 | 10 | 7 | 7 | 8 |
| 108 | 1.822 | 1.983 | 1.537 | 1.929 | 1.629 | 15 | 7 | 10 | 17 |
| 109 | 1.043 | 0.680 | 1.012 | 0.692 | 0.807 | 11 | 7 | 7 | 7 |
| 110 | -0.511 | -0.158 | -0.321 | -0.543 | -0.024 | 10 | 6 | 8 | 13 |
| 111 | 1.380 | 1.063 | 1.485 | 1.079 | 1.581 | 10 | 7 | 7 | 17 |
| 112 | -1.649 | -1.193 | -1.447 | -1.395 | -1.689 | 11 | 6 | 9 | 11 |
| 113 | 0.460 | 0.535 | 0.662 | 0.242 | 0.221 | 9 | 6 | 5 | 7 |
| 114 | -1.311 | -1.584 | -1.587 | -0.910 | -1.283 | 8 | 7 | 6 | 12 |
| 115 | -0.515 | -0.148 | -0.470 | -0.327 | -0.633 | 15 | 6 | 6 | 7 |
| 116 | -0.432 | -1.314 | -0.360 | -0.680 | -0.823 | 12 | 8 | 6 | 8 |
| 117 | -1.277 | -0.777 | -0.655 | -1.120 | -1.120 | 11 | 6 | 8 | 9 |
| 118 | -0.457 | 0.191 | 0.031 | -0.517 | -0.119 | 13 | 10 | 6 | 10 |
| 119 | 1.878 | 2.020 | 1.556 | 2.363 | 1.740 | 8 | 7 | 15 | 16 |
| 120 | -0.758 | -0.984 | -0.980 | -0.416 | -0.930 | 7 | 6 | 4 | 10 |
| 121 | -0.597 | -0.695 | -0.674 | -0.418 | -0.701 | 6 | 7 | 6 | 7 |
| 122 | -0.301 | -0.762 | -0.659 | -0.428 | -0.684 | 11 | 8 | 5 | 8 |
| 123 | 0.399 | 1.549 | 0.496 | -0.006 | 0.311 | 17 | 8 | 7 | 10 |
| 124 | -0.063 | -0.268 | 0.037 | 0.111 | 0.350 | 12 | 9 | 12 | 15 |
| 125 | -0.720 | -1.282 | -0.459 | -0.841 | -0.828 | 17 | 6 | 7 | 8 |
| 126 | -0.262 | -0.100 | -0.798 | -0.367 | -0.055 | 7 | 5 | 8 | 13 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า θ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 127 | 1.272 | 1.126 | 1.281 | 1.355 | 0.691 | 13 | 8 | 13 | 8 |
| 128 | 2.055 | 1.112 | 2.299 | 1.573 | 2.010 | 12 | 8 | 12 | 15 |
| 129 | 0.950 | 1.345 | 1.117 | 0.546 | 0.612 | 9 | 6 | 6 | 9 |
| 130 | -1.308 | -1.481 | -1.271 | -1.319 | -1.343 | 9 | 10 | 7 | 7 |
| 131 | -0.331 | -0.824 | 0.254 | -0.273 | -0.958 | 11 | 8 | 6 | 8 |
| 132 | -0.763 | -0.611 | -0.979 | -1.036 | -0.513 | 7 | 4 | 9 | 8 |
| 133 | -0.050 | -0.988 | 0.466 | 0.031 | -0.678 | 12 | 8 | 6 | 11 |
| 134 | -0.448 | -0.339 | 0.383 | -0.456 | -0.178 | 13 | 10 | 13 | 10 |
| 135 | 0.733 | 0.611 | 0.849 | 0.803 | 0.845 | 7 | 5 | 7 | 8 |
| 136 | 0.426 | 0.552 | 0.455 | 0.402 | 0.649 | 11 | 6 | 10 | 6 |
| 137 | 0.331 | 0.510 | 0.165 | 0.486 | 0.279 | 11 | 9 | 7 | 8 |
| 138 | 0.164 | 0.219 | 0.057 | 0.015 | 0.629 | 11 | 8 | 5 | 8 |
| 139 | 0.014 | 0.051 | 0.027 | -0.274 | -0.217 | 11 | 9 | 14 | 9 |
| 140 | -1.479 | -1.037 | -1.706 | -1.405 | -1.359 | 11 | 6 | 7 | 8 |
| 141 | -0.326 | 0.224 | 0.239 | -0.038 | -0.669 | 12 | 9 | 8 | 7 |
| 142 | 0.224 | -0.041 | 0.476 | 0.016 | 0.002 | 15 | 6 | 7 | 8 |
| 143 | -1.128 | -1.365 | -1.106 | -1.275 | -0.978 | 9 | 5 | 7 | 12 |
| 144 | -2.085 | -1.931 | -1.704 | -1.714 | -1.684 | 11 | 6 | 9 | 10 |
| 145 | -0.052 | -0.422 | -0.264 | 0.215 | 0.020 | 11 | 9 | 6 | 7 |
| 146 | -0.644 | -0.342 | -0.778 | -0.239 | -0.114 | 10 | 5 | 5 | 7 |
| 147 | 1.108 | 1.316 | 0.425 | 1.462 | 1.295 | 11 | 9 | 11 | 11 |
| 148 | 0.606 | 0.620 | 0.235 | 0.472 | 0.173 | 10 | 6 | 6 | 9 |
| 149 | -0.828 | -0.212 | -1.188 | -0.745 | -0.478 | 9 | 5 | 6 | 9 |
| 150 | 1.188 | 1.446 | 1.242 | 1.204 | 1.519 | 13 | 7 | 7 | 16 |
| 151 | -0.675 | -0.307 | -0.691 | -0.773 | -0.902 | 10 | 8 | 9 | 7 |
| 152 | -0.081 | -0.356 | 0.523 | 0.188 | -0.397 | 11 | 8 | 6 | 8 |
| 153 | 0.154 | 0.540 | -0.160 | 0.368 | 0.385 | 11 | 11 | 7 | 8 |
| 154 | 1.146 | 0.484 | 0.726 | 1.137 | 0.935 | 11 | 6 | 9 | 11 |
| 155 | -0.230 | -0.195 | -0.451 | 0.114 | -0.570 | 12 | 6 | 13 | 11 |
| 156 | -0.805 | -0.999 | -0.765 | -0.684 | -0.097 | 9 | 7 | 6 | 12 |
| 157 | 0.381 | 0.420 | 0.217 | 0.157 | 0.401 | 8 | 10 | 6 | 10 |
| 158 | -1.034 | -1.114 | -0.979 | -1.136 | -1.221 | 9 | 5 | 7 | 7 |
| 159 | 0.748 | 0.701 | 0.855 | 0.541 | 0.736 | 9 | 5 | 6 | 12 |
| 160 | 0.356 | 0.566 | -0.294 | 0.827 | 0.883 | 12 | 7 | 5 | 8 |
| 161 | -0.719 | -0.863 | -0.991 | -0.674 | -1.005 | 12 | 5 | 8 | 10 |
| 162 | 0.248 | 1.212 | 0.686 | 0.150 | 0.171 | 19 | 7 | 6 | 6 |
| 163 | -0.795 | -0.487 | -0.697 | -0.351 | -1.046 | 16 | 6 | 7 | 12 |
| 164 | 1.340 | 0.878 | 1.222 | 1.522 | 1.214 | 14 | 7 | 12 | 11 |
| 165 | 0.171 | 0.770 | -0.322 | 0.236 | 0.261 | 10 | 7 | 8 | 11 |
| 166 | 1.888 | 2.381 | 1.961 | 1.626 | 1.224 | 16 | 8 | 13 | 4 |
| 167 | 0.520 | 0.234 | 0.033 | 0.542 | 0.458 | 7 | 9 | 6 | 7 |
| 168 | -0.148 | 0.606 | 0.036 | -0.056 | 0.152 | 9 | 9 | 7 | 8 |
| 169 | 1.412 | 0.993 | 1.185 | 1.582 | 1.569 | 16 | 6 | 11 | 14 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 170 | -0.952 | -1.272 | -0.915 | -0.804 | -0.850 | 8 | 5 | 8 | 6 |
| 171 | -1.172 | -0.961 | -1.353 | -0.819 | -0.995 | 14 | 5 | 6 | 17 |
| 172 | 1.202 | 0.253 | 1.346 | 0.971 | 1.467 | 7 | 7 | 6 | 13 |
| 173 | -0.781 | -1.107 | -0.786 | -0.361 | -1.554 | 10 | 6 | 9 | 10 |
| 174 | -0.786 | -0.865 | -0.814 | -0.793 | -1.102 | 9 | 7 | 4 | 7 |
| 175 | -0.081 | -0.529 | -0.022 | -0.026 | 0.280 | 9 | 10 | 8 | 10 |
| 176 | -1.912 | -1.998 | -0.916 | -1.873 | -1.646 | 13 | 6 | 10 | 11 |
| 177 | -1.704 | -0.018 | -1.664 | -0.681 | -0.698 | 10 | 6 | 6 | 9 |
| 178 | 0.520 | -0.078 | 0.642 | 0.254 | 0.598 | 6 | 7 | 6 | 8 |
| 179 | 0.745 | 0.116 | 0.592 | 0.390 | 0.095 | 10 | 7 | 7 | 11 |
| 180 | 0.761 | 0.307 | 0.469 | 0.629 | 0.533 | 10 | 6 | 7 | 9 |
| 181 | 1.494 | 1.601 | 2.012 | 1.025 | 0.958 | 9 | 8 | 7 | 7 |
| 182 | 1.261 | 0.711 | 1.338 | 1.776 | 0.738 | 10 | 7 | 14 | 8 |
| 183 | -0.797 | -1.226 | -1.090 | -1.330 | -0.879 | 10 | 5 | 7 | 7 |
| 184 | -1.422 | -0.974 | -1.746 | -1.325 | -1.309 | 10 | 6 | 10 | 9 |
| 185 | 0.264 | 0.868 | 0.284 | 0.542 | -0.006 | 14 | 7 | 6 | 11 |
| 186 | 0.251 | -0.235 | 0.063 | 0.144 | 0.534 | 15 | 8 | 5 | 8 |
| 187 | -1.250 | -1.544 | -0.872 | -1.149 | -1.450 | 8 | 8 | 6 | 7 |
| 188 | 0.637 | 0.537 | 0.424 | 0.859 | 0.354 | 11 | 8 | 5 | 14 |
| 189 | 2.705 | 3.716 | 2.381 | 1.566 | 1.632 | 30 | 8 | 12 | 14 |
| 190 | 0.478 | 0.519 | 0.679 | 0.933 | 0.372 | 7 | 5 | 6 | 11 |
| 191 | -1.207 | -1.170 | -1.151 | -1.275 | -1.488 | 10 | 9 | 9 | 12 |
| 192 | 0.163 | 0.507 | -0.012 | 0.272 | -0.046 | 14 | 9 | 5 | 11 |
| 193 | -0.302 | -0.173 | -0.281 | -0.181 | -0.479 | 10 | 8 | 8 | 8 |
| 194 | 0.959 | 0.984 | 0.875 | 1.027 | 0.773 | 8 | 7 | 6 | 9 |
| 195 | 0.272 | 0.643 | 0.015 | 0.002 | 0.221 | 9 | 10 | 5 | 9 |
| 196 | -1.656 | -1.147 | -1.727 | -1.699 | -1.542 | 8 | 7 | 9 | 12 |
| 197 | 0.072 | -0.113 | 0.014 | 0.024 | 0.076 | 17 | 10 | 5 | 8 |
| 198 | 0.734 | 0.807 | 0.757 | 1.169 | 0.259 | 9 | 5 | 9 | 14 |
| 199 | 0.657 | 1.257 | 1.150 | 0.323 | 0.680 | 11 | 7 | 10 | 8 |
| 200 | -0.891 | -0.528 | -0.912 | -1.287 | -0.680 | 13 | 6 | 6 | 13 |
| 201 | 0.628 | 0.672 | 0.569 | 0.697 | 0.631 | 11 | 7 | 5 | 11 |
| 202 | 1.416 | 0.803 | 0.843 | 1.571 | 0.895 | 13 | 5 | 8 | 15 |
| 203 | -0.326 | 0.169 | -0.730 | -0.392 | -0.117 | 11 | 7 | 7 | 15 |
| 204 | -1.454 | -1.267 | -1.321 | -1.043 | -1.876 | 17 | 6 | 10 | 9 |
| 205 | -1.140 | -0.351 | -0.757 | -0.743 | -1.551 | 12 | 6 | 9 | 13 |
| 206 | 0.915 | 1.570 | 0.566 | 0.781 | 1.233 | 12 | 7 | 7 | 8 |
| 207 | -1.403 | -0.999 | -0.703 | -1.270 | -1.620 | 16 | 11 | 5 | 8 |
| 208 | -0.452 | -0.908 | -0.247 | -0.349 | -0.333 | 9 | 8 | 8 | 9 |
| 209 | 1.323 | 0.092 | 0.916 | 1.210 | 1.486 | 14 | 7 | 7 | 10 |
| 210 | -0.547 | -0.193 | -0.768 | -0.800 | -0.927 | 10 | 6 | 6 | 7 |
| 211 | -1.223 | -1.659 | -1.703 | -0.999 | -1.394 | 10 | 6 | 7 | 6 |
| 212 | 0.062 | 1.212 | -0.051 | -0.316 | 0.471 | 13 | 9 | 5 | 9 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 213 | -0.019 | -0.436 | -0.385 | -0.012 | -0.006 | 9 | 9 | 7 | 15 |
| 214 | 0.470 | 1.583 | 0.275 | 0.480 | 0.664 | 9 | 8 | 6 | 9 |
| 215 | 0.203 | -0.156 | 0.035 | -0.349 | -0.217 | 12 | 9 | 6 | 17 |
| 216 | -1.177 | -1.314 | -0.898 | -1.358 | -1.450 | 14 | 6 | 11 | 7 |
| 217 | 1.455 | 1.430 | 1.152 | 1.244 | 1.643 | 11 | 8 | 9 | 17 |
| 218 | -0.239 | 0.317 | -0.040 | 0.218 | -0.196 | 12 | 8 | 6 | 10 |
| 219 | 0.081 | 0.500 | -0.249 | 0.148 | 0.633 | 12 | 9 | 10 | 17 |
| 220 | 2.091 | 1.801 | 1.660 | 1.139 | 2.366 | 11 | 7 | 15 | 12 |
| 221 | 1.692 | 2.657 | 1.866 | 1.919 | 1.761 | 16 | 7 | 8 | 11 |
| 222 | 0.757 | 0.915 | 0.631 | 0.915 | 0.398 | 10 | 5 | 6 | 14 |
| 223 | 0.606 | 0.092 | 0.742 | -0.030 | 0.579 | 11 | 6 | 8 | 7 |
| 224 | 2.333 | 2.319 | 1.140 | 1.944 | 1.746 | 15 | 7 | 8 | 14 |
| 225 | -0.240 | -0.451 | -0.553 | 0.110 | -0.289 | 12 | 7 | 5 | 13 |
| 226 | 0.720 | 0.777 | 0.703 | 1.019 | 0.918 | 13 | 6 | 8 | 8 |
| 227 | -0.167 | -0.190 | -0.027 | -0.313 | 0.002 | 13 | 8 | 9 | 12 |
| 228 | 0.107 | -0.092 | 0.031 | 0.110 | -0.401 | 10 | 9 | 13 | 12 |
| 229 | -1.443 | -1.195 | -0.947 | -1.365 | -1.997 | 10 | 5 | 8 | 12 |
| 230 | -1.347 | -1.195 | -1.281 | -1.160 | -1.253 | 9 | 9 | 10 | 11 |
| 231 | -0.896 | -0.682 | -0.994 | -0.138 | -0.498 | 9 | 5 | 5 | 10 |
| 232 | -2.485 | -1.561 | -2.618 | -1.031 | -1.813 | 11 | 8 | 11 | 15 |
| 233 | -0.039 | 0.170 | -0.460 | -0.066 | 0.125 | 10 | 6 | 5 | 10 |
| 234 | -0.761 | 0.232 | -0.775 | -0.999 | -0.994 | 10 | 6 | 7 | 6 |
| 235 | 0.988 | 0.092 | 1.137 | 1.139 | 1.035 | 13 | 6 | 15 | 7 |
| 236 | 0.725 | 0.689 | 0.501 | 0.281 | 0.521 | 11 | 7 | 8 | 8 |
| 237 | -1.372 | -1.731 | -1.781 | -1.100 | -0.784 | 9 | 6 | 7 | 9 |
| 238 | 0.912 | 0.425 | 0.066 | 0.402 | 0.879 | 14 | 9 | 7 | 10 |
| 239 | -0.130 | 0.055 | -0.402 | -0.352 | -0.538 | 9 | 5 | 5 | 11 |
| 240 | 0.318 | 0.052 | 0.007 | 0.235 | 0.321 | 12 | 9 | 6 | 9 |
| 241 | -0.577 | -0.844 | -0.487 | -0.396 | -0.472 | 10 | 8 | 7 | 10 |
| 242 | 0.266 | -0.327 | 0.119 | 0.008 | 0.804 | 10 | 10 | 6 | 12 |
| 243 | -0.255 | -0.778 | 0.038 | -0.760 | -0.425 | 10 | 8 | 5 | 11 |
| 244 | -0.019 | 0.116 | -0.544 | -0.060 | 0.213 | 11 | 5 | 10 | 11 |
| 245 | -0.346 | -0.485 | -0.299 | -0.500 | 0.075 | 14 | 8 | 5 | 11 |
| 246 | -1.284 | -1.609 | -0.958 | -1.067 | -1.217 | 8 | 6 | 5 | 8 |
| 247 | 0.881 | 1.035 | 0.807 | 0.916 | 0.984 | 10 | 7 | 5 | 12 |
| 248 | 1.893 | 1.112 | 2.196 | 1.851 | 1.498 | 12 | 7 | 9 | 9 |
| 249 | 0.236 | -0.298 | 0.636 | 0.750 | 0.134 | 9 | 5 | 10 | 9 |
| 250 | 1.183 | 0.043 | 1.390 | 0.853 | 1.483 | 13 | 8 | 6 | 14 |
| 251 | 0.095 | -0.422 | 0.413 | 0.104 | 0.482 | 11 | 8 | 5 | 8 |
| 252 | 0.611 | 0.238 | 0.914 | 0.540 | 0.395 | 11 | 8 | 6 | 10 |
| 253 | -0.173 | -0.161 | -0.224 | -0.286 | -0.242 | 14 | 12 | 5 | 15 |
| 254 | 0.920 | 1.945 | 0.849 | 0.686 | 0.796 | 18 | 5 | 5 | 13 |
| 255 | -1.199 | -1.245 | -0.780 | -0.893 | -1.276 | 12 | 5 | 11 | 15 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 256 | 1.547 | 1.015 | 1.335 | 1.070 | 0.990 | 11 | 7 | 8 | 8 |
| 257 | -0.422 | -0.350 | -0.259 | -0.325 | -0.783 | 9 | 8 | 7 | 9 |
| 258 | -1.691 | -0.589 | -1.719 | -1.276 | -1.503 | 7 | 8 | 7 | 12 |
| 259 | 0.460 | 0.482 | 0.275 | 0.473 | 0.401 | 12 | 7 | 6 | 11 |
| 260 | -0.970 | 0.439 | -0.317 | -1.200 | -0.795 | 11 | 7 | 6 | 7 |
| 261 | -0.094 | -0.460 | -0.243 | -0.278 | 0.152 | 12 | 10 | 5 | 8 |
| 262 | -0.178 | -0.819 | -0.053 | -0.064 | 0.149 | 11 | 9 | 5 | 13 |
| 263 | 0.158 | 0.583 | 0.291 | 0.071 | -0.004 | 12 | 7 | 6 | 7 |
| 264 | 0.843 | 0.738 | 0.693 | 1.454 | 0.750 | 10 | 7 | 13 | 11 |
| 265 | -1.292 | -0.788 | -1.707 | -1.538 | -1.117 | 12 | 6 | 8 | 10 |
| 266 | 0.611 | 0.042 | 0.304 | 0.966 | 0.419 | 11 | 7 | 5 | 6 |
| 267 | 0.266 | 0.937 | 0.293 | 0.142 | 0.112 | 10 | 9 | 7 | 8 |
| 268 | -1.825 | -1.853 | -2.112 | -1.445 | -1.726 | 18 | 6 | 7 | 12 |
| 269 | -1.319 | -1.460 | -1.155 | -0.608 | -0.957 | 9 | 5 | 9 | 9 |
| 270 | -0.126 | -0.202 | -0.263 | -0.076 | -0.004 | 10 | 9 | 6 | 9 |
| 271 | 0.481 | 0.997 | 0.502 | 0.604 | 0.298 | 12 | 7 | 5 | 9 |
| 272 | 0.780 | 0.520 | 0.679 | 0.568 | 0.291 | 11 | 6 | 8 | 13 |
| 273 | 1.776 | 1.122 | 1.424 | 1.770 | 1.422 | 11 | 6 | 10 | 10 |
| 274 | 1.162 | 1.522 | 1.378 | 1.099 | 1.260 | 12 | 9 | 6 | 12 |
| 275 | 0.561 | 0.819 | 0.436 | 0.540 | 0.714 | 7 | 8 | 6 | 8 |
| 276 | -0.608 | 0.510 | -0.450 | -0.822 | -0.689 | 12 | 7 | 6 | 8 |
| 277 | -0.722 | -0.383 | -0.981 | -0.898 | -0.419 | 13 | 8 | 10 | 7 |
| 278 | 0.749 | 0.770 | 0.470 | 0.934 | 0.498 | 11 | 8 | 6 | 10 |
| 279 | 0.497 | 1.550 | 0.043 | 0.971 | 0.535 | 16 | 10 | 6 | 9 |
| 280 | 0.046 | -0.235 | 0.042 | 0.326 | -0.407 | 11 | 8 | 7 | 7 |
| 281 | -1.514 | -0.057 | -1.419 | -1.710 | -1.016 | 12 | 6 | 9 | 9 |
| 282 | 2.104 | 2.146 | 1.701 | 1.063 | 1.915 | 12 | 8 | 15 | 12 |
| 283 | 0.158 | 0.821 | -0.052 | -0.275 | 0.175 | 10 | 10 | 6 | 8 |
| 284 | 0.360 | 0.978 | 0.669 | 0.167 | 0.348 | 12 | 5 | 6 | 8 |
| 285 | -1.581 | -1.780 | -1.518 | -1.178 | -1.644 | 9 | 5 | 5 | 6 |
| 286 | -0.850 | 0.253 | -0.578 | -0.805 | -0.244 | 18 | 8 | 6 | 11 |
| 287 | -0.421 | -0.473 | -0.942 | 0.298 | -0.131 | 7 | 6 | 5 | 17 |
| 288 | -0.480 | -0.503 | -0.017 | -0.672 | -0.167 | 11 | 11 | 6 | 7 |
| 289 | 0.027 | 0.187 | 0.032 | -0.145 | -0.110 | 15 | 9 | 5 | 10 |
| 290 | -0.120 | -0.043 | -0.428 | -0.032 | -0.162 | 10 | 10 | 4 | 9 |
| 291 | -0.544 | -0.768 | -0.461 | -0.289 | -0.191 | 10 | 8 | 5 | 11 |
| 292 | -0.685 | -0.371 | -0.986 | -0.606 | -0.678 | 9 | 5 | 5 | 7 |
| 293 | 1.406 | 0.963 | 1.022 | 1.776 | 1.026 | 11 | 6 | 13 | 12 |
| 294 | 1.483 | 0.931 | 0.799 | 1.683 | 1.693 | 9 | 6 | 16 | 8 |
| 295 | 0.900 | 0.224 | 0.814 | 1.162 | 1.079 | 13 | 5 | 11 | 14 |
| 296 | -0.252 | -0.892 | -0.179 | -0.732 | 0.022 | 9 | 9 | 6 | 14 |
| 297 | -0.027 | 0.189 | 0.050 | -0.011 | 0.257 | 11 | 9 | 11 | 7 |
| 298 | 0.238 | 0.518 | -0.003 | 0.262 | 0.478 | 14 | 11 | 7 | 9 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 299 | -0.015 | -0.094 | -0.266 | 0.353 | -0.016 | 14 | 7 | 6 | 8 |
| 300 | -0.034 | -0.382 | -0.659 | 0.123 | 0.401 | 12 | 8 | 9 | 9 |
| 301 | -0.046 | -1.435 | -0.660 | 0.168 | 0.017 | 7 | 8 | 5 | 7 |
| 302 | -0.283 | -0.163 | -0.351 | -0.051 | -0.452 | 12 | 8 | 8 | 10 |
| 303 | -1.023 | -0.799 | -1.659 | -0.441 | -0.646 | 9 | 6 | 7 | 17 |
| 304 | -0.659 | 0.187 | -0.506 | -0.195 | -0.459 | 15 | 6 | 7 | 6 |
| 305 | -0.116 | -0.382 | -0.289 | -0.129 | -0.044 | 9 | 9 | 5 | 9 |
| 306 | 0.167 | -0.217 | -0.034 | 0.245 | 0.110 | 11 | 9 | 5 | 14 |
| 307 | -1.155 | -0.665 | -1.094 | -0.955 | -1.091 | 10 | 5 | 5 | 9 |
| 308 | -1.154 | -1.631 | -0.976 | -1.209 | -0.548 | 12 | 6 | 10 | 10 |
| 309 | -0.529 | -0.461 | -0.230 | 0.002 | -0.721 | 7 | 12 | 5 | 10 |
| 310 | 1.049 | 1.158 | 1.378 | 1.092 | 0.609 | 16 | 8 | 5 | 9 |
| 311 | 0.681 | 0.643 | 0.826 | 0.616 | 0.490 | 9 | 5 | 6 | 7 |
| 312 | -2.929 | -1.994 | -2.122 | -1.737 | -2.002 | 11 | 9 | 8 | 12 |
| 313 | -0.517 | -0.667 | -0.278 | -0.577 | -0.332 | 10 | 7 | 6 | 7 |
| 314 | -1.540 | -0.900 | -1.294 | -0.916 | -1.386 | 9 | 6 | 6 | 10 |
| 315 | -0.862 | -0.169 | -1.116 | -0.801 | -0.907 | 15 | 6 | 6 | 6 |
| 316 | 0.237 | -0.071 | -0.213 | 0.750 | 0.176 | 9 | 8 | 6 | 13 |
| 317 | -0.558 | -1.125 | -0.248 | -0.334 | -0.572 | 15 | 10 | 5 | 7 |
| 318 | -0.799 | -1.140 | -0.795 | -0.277 | -0.771 | 11 | 6 | 6 | 9 |
| 319 | -0.195 | 0.271 | -0.318 | -0.195 | -0.491 | 12 | 7 | 5 | 8 |
| 320 | -0.508 | 0.661 | -0.577 | -0.540 | -0.945 | 10 | 8 | 9 | 7 |
| 321 | -0.208 | 0.457 | -0.575 | -0.217 | -0.129 | 10 | 9 | 5 | 8 |
| 322 | -0.710 | -0.555 | -0.674 | -0.540 | -0.672 | 12 | 7 | 6 | 7 |
| 323 | 0.818 | 0.709 | 0.363 | 0.606 | 0.819 | 9 | 6 | 7 | 10 |
| 324 | -0.316 | -0.430 | -0.033 | -0.506 | -0.179 | 8 | 8 | 5 | 10 |
| 325 | -1.080 | -0.672 | -0.897 | -0.526 | -0.901 | 7 | 5 | 6 | 5 |
| 326 | 1.246 | 1.148 | 1.123 | 1.094 | 0.862 | 13 | 6 | 5 | 6 |
| 327 | -0.016 | -0.376 | -0.369 | 0.271 | 0.243 | 12 | 6 | 6 | 11 |
| 328 | -0.519 | -0.149 | -0.700 | -0.512 | -0.941 | 14 | 6 | 5 | 16 |
| 329 | 0.796 | 0.411 | 0.642 | 0.428 | 0.443 | 10 | 6 | 7 | 9 |
| 330 | 1.798 | 2.798 | 1.991 | 1.744 | 0.998 | 21 | 8 | 7 | 8 |
| 331 | -1.357 | -1.293 | -1.717 | -1.560 | -0.714 | 11 | 6 | 8 | 8 |
| 332 | 1.010 | 0.881 | 0.782 | 1.080 | 1.020 | 14 | 6 | 7 | 11 |
| 333 | -0.887 | -0.536 | -0.436 | -1.199 | -0.882 | 7 | 6 | 12 | 8 |
| 334 | -0.992 | -1.563 | -0.978 | -0.940 | -0.858 | 7 | 6 | 9 | 9 |
| 335 | 0.457 | 0.993 | 0.329 | 0.572 | 1.054 | 17 | 8 | 4 | 10 |
| 336 | -0.811 | -0.405 | -0.713 | -1.124 | -0.062 | 10 | 7 | 5 | 13 |
| 337 | -0.013 | -0.219 | -0.409 | 0.172 | 0.017 | 13 | 7 | 6 | 12 |
| 338 | 0.321 | -0.462 | 0.322 | 0.434 | 0.418 | 13 | 7 | 5 | 8 |
| 339 | 0.574 | 0.976 | 0.524 | 0.236 | 0.180 | 10 | 8 | 6 | 9 |
| 340 | -0.167 | 0.325 | -0.269 | -0.371 | 0.055 | 13 | 8 | 8 | 10 |
| 341 | -1.308 | -1.199 | -0.765 | -1.626 | -1.417 | 10 | 6 | 10 | 10 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 342 | 0.633 | 0.268 | 0.487 | 0.813 | 0.700 | 12 | 8 | 5 | 10 |
| 343 | -0.368 | -0.051 | -0.783 | -0.521 | -0.824 | 11 | 5 | 5 | 11 |
| 344 | 0.013 | 0.434 | -0.447 | 0.234 | -0.120 | 10 | 7 | 6 | 9 |
| 345 | -0.356 | -0.395 | 0.289 | -0.824 | -0.118 | 15 | 9 | 6 | 7 |
| 346 | 1.053 | 0.996 | 0.619 | 0.936 | 1.584 | 8 | 6 | 6 | 16 |
| 347 | -0.672 | -0.931 | -0.413 | -0.586 | -0.509 | 8 | 6 | 8 | 7 |
| 348 | 0.502 | 0.713 | 0.429 | 0.464 | 0.785 | 9 | 8 | 6 | 8 |
| 349 | 1.027 | 1.221 | 0.637 | 0.819 | 0.291 | 12 | 6 | 5 | 13 |
| 350 | -0.226 | -0.733 | -0.209 | -0.600 | 0.152 | 10 | 10 | 5 | 11 |
| 351 | 0.484 | 1.484 | 0.443 | 0.810 | 0.578 | 14 | 7 | 5 | 8 |
| 352 | -0.075 | 0.005 | -0.027 | 0.161 | -0.330 | 11 | 11 | 10 | 8 |
| 353 | 0.298 | 0.928 | -0.241 | 0.326 | 0.129 | 12 | 8 | 5 | 11 |
| 354 | -0.980 | -1.370 | -0.778 | -1.211 | -0.822 | 6 | 5 | 7 | 9 |
| 355 | 0.536 | 0.810 | 0.963 | 0.746 | 0.343 | 7 | 7 | 6 | 7 |
| 356 | -1.240 | -0.740 | -1.050 | -0.867 | -1.402 | 6 | 6 | 5 | 8 |
| 357 | 0.223 | -0.136 | 0.225 | 0.164 | -0.170 | 11 | 8 | 7 | 9 |
| 358 | 2.174 | 1.734 | 2.039 | 1.943 | 1.880 | 11 | 7 | 13 | 17 |
| 359 | 0.056 | -0.238 | 0.043 | -0.273 | 0.247 | 9 | 9 | 6 | 8 |
| 360 | -1.006 | -0.696 | -0.683 | -0.756 | -0.428 | 9 | 6 | 8 | 12 |
| 361 | -0.592 | -0.177 | -0.590 | -0.147 | -0.733 | 9 | 8 | 7 | 8 |
| 362 | -0.812 | -0.927 | -0.720 | -1.279 | -0.207 | 12 | 7 | 9 | 12 |
| 363 | -0.254 | -0.726 | -0.617 | -0.143 | -0.427 | 6 | 7 | 5 | 14 |
| 364 | -0.585 | -0.211 | -0.679 | -0.430 | -0.476 | 9 | 6 | 5 | 8 |
| 365 | -1.048 | -0.914 | -0.992 | -1.118 | -0.503 | 8 | 5 | 7 | 12 |
| 366 | -0.877 | -1.068 | 0.685 | -0.839 | -0.843 | 9 | 6 | 7 | 10 |
| 367 | -1.570 | -1.009 | -1.697 | -1.338 | -1.728 | 7 | 5 | 6 | 11 |
| 368 | -0.292 | -0.398 | -0.624 | 0.111 | -0.048 | 10 | 7 | 5 | 14 |
| 369 | 2.437 | 4.821 | 2.509 | 2.568 | 1.628 | 30 | 9 | 11 | 18 |
| 370 | -0.570 | -0.489 | 0.036 | -0.520 | -0.556 | 8 | 11 | 7 | 12 |
| 371 | 1.183 | 0.588 | 0.860 | 1.036 | 1.173 | 8 | 7 | 7 | 13 |
| 372 | 1.219 | 1.210 | 1.147 | 0.905 | 1.354 | 10 | 6 | 6 | 11 |
| 373 | -0.076 | -0.745 | -0.305 | -0.356 | -0.276 | 11 | 6 | 6 | 7 |
| 374 | -0.229 | -0.341 | -0.202 | -0.333 | 0.000 | 8 | 10 | 6 | 12 |
| 375 | 1.030 | 1.135 | 0.802 | 1.027 | 0.733 | 11 | 5 | 8 | 7 |
| 376 | 0.235 | -0.240 | 0.496 | -0.061 | -0.216 | 11 | 8 | 7 | 11 |
| 377 | -0.442 | -0.180 | -0.268 | -0.720 | -0.211 | 9 | 9 | 6 | 10 |
| 378 | 0.029 | -0.185 | 0.015 | 0.144 | 0.159 | 13 | 10 | 7 | 16 |
| 379 | -1.154 | -1.262 | -1.262 | -1.207 | -1.426 | 14 | 6 | 7 | 12 |
| 380 | -0.356 | -0.884 | -0.230 | -0.052 | -0.154 | 7 | 12 | 13 | 15 |
| 381 | 0.612 | 0.951 | 0.505 | 0.636 | 0.006 | 9 | 6 | 6 | 7 |
| 382 | -0.566 | -0.066 | -0.431 | -0.774 | -0.233 | 8 | 9 | 9 | 7 |
| 383 | -1.072 | -1.138 | -0.969 | -1.106 | -0.860 | 8 | 6 | 6 | 7 |
| 384 | 1.665 | 1.330 | 1.715 | 0.960 | 1.659 | 16 | 9 | 8 | 10 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า θ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 385 | 0.371 | 0.164 | 0.500 | 0.151 | 0.329 | 10 | 8 | 5 | 9 |
| 386 | -2.042 | -1.501 | -1.792 | -1.438 | -1.641 | 15 | 6 | 13 | 15 |
| 387 | -2.577 | -1.808 | -2.062 | -1.598 | -2.401 | 10 | 7 | 8 | 10 |
| 388 | 0.160 | 0.047 | -0.083 | -0.047 | 0.068 | 16 | 9 | 4 | 9 |
| 389 | -0.308 | 0.408 | -0.643 | -0.161 | -0.541 | 8 | 8 | 10 | 13 |
| 390 | 1.489 | 1.112 | 1.127 | 1.534 | 1.777 | 9 | 8 | 14 | 15 |
| 391 | -0.530 | -0.476 | -0.440 | -0.638 | -0.972 | 11 | 5 | 5 | 7 |
| 392 | 1.469 | 1.576 | 1.053 | 1.552 | 1.607 | 11 | 6 | 6 | 11 |
| 393 | 0.981 | 0.584 | 0.891 | 0.532 | 1.178 | 13 | 5 | 4 | 8 |
| 394 | 0.342 | -0.109 | 0.738 | 0.046 | 0.060 | 16 | 5 | 6 | 9 |
| 395 | -0.919 | -1.631 | -0.706 | -0.849 | -1.001 | 10 | 6 | 9 | 6 |
| 396 | -1.301 | -1.362 | -0.726 | -1.254 | -1.225 | 11 | 7 | 9 | 9 |
| 397 | -1.370 | -1.120 | -0.919 | -1.556 | -1.039 | 10 | 6 | 8 | 11 |
| 398 | 1.124 | 0.430 | 0.808 | 1.531 | 0.703 | 12 | 5 | 7 | 9 |
| 399 | 2.421 | 2.506 | 1.877 | 2.232 | 2.433 | 20 | 7 | 14 | 18 |
| 400 | -1.937 | -1.595 | -1.990 | -1.745 | -0.960 | 13 | 7 | 8 | 12 |
| 401 | 1.835 | 2.309 | 1.710 | 0.999 | 1.998 | 17 | 8 | 6 | 11 |
| 402 | 0.476 | -0.005 | 0.041 | 0.412 | 0.419 | 9 | 8 | 9 | 7 |
| 403 | 3.057 | 2.819 | 2.567 | 1.040 | 1.951 | 23 | 10 | 30 | 12 |
| 404 | -2.329 | -1.613 | -2.248 | -1.637 | -2.196 | 10 | 7 | 10 | 8 |
| 405 | 0.342 | 0.447 | 0.393 | 0.183 | 0.188 | 10 | 8 | 7 | 11 |
| 406 | -0.227 | -0.701 | 0.282 | -0.201 | 0.220 | 7 | 8 | 5 | 7 |
| 407 | 0.686 | 0.213 | 1.138 | 0.895 | 0.310 | 12 | 6 | 6 | 8 |
| 408 | 0.769 | 0.536 | 0.797 | 0.401 | 0.838 | 9 | 6 | 6 | 9 |
| 409 | -0.774 | -0.572 | -0.696 | -0.413 | -0.202 | 8 | 6 | 4 | 14 |
| 410 | 1.047 | 0.600 | 1.139 | 0.346 | 1.164 | 13 | 6 | 7 | 9 |
| 411 | 0.773 | 1.219 | 0.377 | 1.040 | 0.472 | 10 | 6 | 30 | 13 |
| 412 | -1.023 | -0.977 | -1.326 | -0.825 | -1.090 | 19 | 4 | 9 | 8 |
| 413 | 2.494 | 0.939 | 2.208 | 1.648 | 2.006 | 13 | 7 | 12 | 11 |
| 414 | -0.119 | -0.879 | -0.256 | -0.416 | -0.528 | 9 | 9 | 5 | 9 |
| 415 | -0.015 | -0.485 | -0.258 | 0.207 | 0.273 | 10 | 9 | 6 | 9 |
| 416 | -0.236 | -0.577 | 0.275 | -0.331 | 0.020 | 19 | 7 | 7 | 9 |
| 417 | -0.073 | 0.523 | 0.237 | -0.091 | -0.402 | 11 | 9 | 4 | 9 |
| 418 | 0.111 | 0.150 | -0.215 | 0.212 | -0.581 | 11 | 10 | 6 | 7 |
| 419 | 1.463 | 1.025 | 1.559 | 1.052 | 1.502 | 14 | 7 | 7 | 11 |
| 420 | -0.395 | -1.117 | -0.245 | -0.046 | -0.571 | 9 | 9 | 6 | 7 |
| 421 | 0.471 | -0.118 | 0.857 | 0.413 | 0.595 | 13 | 5 | 5 | 15 |
| 422 | -0.035 | 0.403 | 0.039 | -0.150 | 1.187 | 14 | 9 | 7 | 10 |
| 423 | -0.643 | -0.561 | -0.657 | -0.760 | -0.213 | 7 | 8 | 5 | 12 |
| 424 | -0.853 | -0.467 | -0.387 | -0.277 | 0.860 | 12 | 8 | 6 | 11 |
| 425 | 1.405 | 1.020 | 0.799 | 1.606 | 1.621 | 7 | 6 | 14 | 10 |
| 426 | 0.250 | 0.539 | 0.653 | 0.298 | 0.696 | 12 | 6 | 8 | 8 |
| 427 | -0.931 | -0.423 | -0.744 | -1.087 | -0.991 | 12 | 6 | 8 | 8 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 428 | -0.638 | -0.105 | -0.303 | -0.212 | -1.146 | 14 | 8 | 8 | 12 |
| 429 | 0.272 | 0.167 | 0.476 | 0.141 | -0.051 | 14 | 6 | 6 | 9 |
| 430 | -1.568 | -1.274 | -1.624 | -1.565 | -1.496 | 17 | 6 | 8 | 17 |
| 431 | 0.605 | 0.464 | 0.668 | 0.270 | 0.877 | 11 | 6 | 6 | 9 |
| 432 | 1.842 | 2.061 | 1.857 | 1.608 | 1.121 | 15 | 7 | 14 | 12 |
| 433 | -0.546 | -1.149 | -0.283 | -0.484 | -0.646 | 16 | 10 | 10 | 7 |
| 434 | 1.086 | 0.948 | 1.456 | 0.999 | 1.420 | 8 | 7 | 6 | 15 |
| 435 | 0.276 | 0.481 | -0.030 | 0.609 | 0.440 | 15 | 9 | 12 | 8 |
| 436 | -1.175 | -1.351 | -1.176 | -1.690 | -0.485 | 8 | 5 | 7 | 7 |
| 437 | 0.729 | 0.664 | 0.407 | 0.504 | 0.159 | 15 | 9 | 8 | 10 |
| 438 | 0.319 | -0.197 | 0.220 | 0.006 | 0.638 | 10 | 9 | 5 | 10 |
| 439 | 0.298 | 0.648 | 0.017 | 0.244 | 0.201 | 11 | 8 | 6 | 10 |
| 440 | -2.169 | -1.855 | -1.689 | -2.157 | -0.974 | 14 | 8 | 11 | 11 |
| 441 | -0.738 | -1.200 | -0.428 | -0.914 | -0.515 | 8 | 5 | 6 | 9 |
| 442 | 0.016 | 0.216 | 0.015 | -0.330 | 0.310 | 10 | 9 | 9 | 10 |
| 443 | -0.037 | -0.570 | 0.059 | -0.310 | -0.227 | 9 | 8 | 9 | 9 |
| 444 | 0.055 | -0.576 | 0.033 | 0.152 | 0.410 | 10 | 9 | 7 | 9 |
| 445 | 0.161 | -0.150 | 0.024 | -0.041 | 0.260 | 15 | 9 | 14 | 15 |
| 446 | -1.052 | -0.928 | -1.123 | -1.214 | -0.720 | 10 | 6 | 7 | 12 |
| 447 | 1.085 | 0.051 | 1.388 | 1.156 | 1.496 | 11 | 7 | 9 | 11 |
| 448 | 1.026 | 0.448 | 1.250 | 1.512 | 0.824 | 8 | 6 | 6 | 7 |
| 449 | 0.244 | 0.856 | 0.274 | 0.354 | 0.245 | 12 | 8 | 6 | 10 |
| 450 | 0.521 | -0.055 | 0.737 | 0.325 | 0.242 | 21 | 7 | 6 | 11 |
| 451 | 1.261 | 1.714 | 1.647 | 1.171 | 1.279 | 13 | 7 | 8 | 13 |
| 452 | -0.386 | 0.434 | -0.430 | -0.199 | -0.637 | 12 | 6 | 5 | 8 |
| 453 | 0.037 | -0.184 | -0.503 | 0.145 | 0.550 | 10 | 8 | 10 | 8 |
| 454 | 1.014 | 1.719 | 0.437 | 1.177 | 0.800 | 11 | 8 | 9 | 9 |
| 455 | -0.134 | 0.159 | 0.020 | -0.027 | -0.168 | 10 | 9 | 6 | 8 |
| 456 | 0.879 | 0.690 | 0.514 | 1.113 | 0.500 | 10 | 7 | 6 | 11 |
| 457 | 0.968 | 0.892 | 0.866 | 1.052 | 1.562 | 12 | 5 | 6 | 13 |
| 458 | -1.414 | -1.419 | -1.542 | -1.240 | -0.994 | 10 | 7 | 9 | 10 |
| 459 | 0.161 | 0.698 | 0.455 | -0.274 | 0.189 | 8 | 6 | 6 | 10 |
| 460 | 0.259 | 0.378 | 0.544 | 0.613 | 0.182 | 9 | 6 | 5 | 9 |
| 461 | -0.493 | 0.169 | -0.544 | -0.154 | -0.114 | 8 | 5 | 7 | 12 |
| 462 | 1.062 | -0.090 | 0.759 | 1.142 | 0.576 | 8 | 5 | 6 | 7 |
| 463 | 0.978 | 0.072 | 0.820 | 1.197 | 1.107 | 11 | 5 | 7 | 9 |
| 464 | -0.552 | -0.167 | -0.783 | -0.664 | -0.473 | 17 | 6 | 6 | 12 |
| 465 | 1.580 | 1.555 | 1.913 | 1.112 | 1.071 | 16 | 8 | 6 | 10 |
| 466 | -1.071 | -1.387 | -1.091 | -0.973 | -1.140 | 9 | 8 | 7 | 10 |
| 467 | -1.213 | -1.207 | -0.974 | -1.128 | -0.755 | 14 | 5 | 7 | 7 |
| 468 | 0.064 | 0.115 | 0.740 | -0.021 | 0.638 | 14 | 5 | 4 | 10 |
| 469 | -0.091 | -0.745 | -0.251 | 0.110 | -0.352 | 7 | 8 | 5 | 7 |
| 470 | 0.815 | 1.129 | 0.566 | 1.101 | 1.105 | 12 | 7 | 5 | 8 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 471 | 0.689 | 0.594 | 0.430 | 0.105 | 0.346 | 10 | 9 | 12 | 9 |
| 472 | -0.582 | -0.301 | 0.379 | -0.169 | -0.475 | 10 | 6 | 6 | 7 |
| 473 | -0.005 | 1.163 | 0.042 | -0.608 | -0.397 | 11 | 10 | 4 | 6 |
| 474 | 0.227 | 0.395 | 0.214 | 0.628 | -0.069 | 13 | 10 | 6 | 9 |
| 475 | 0.778 | 0.682 | 0.757 | 0.807 | 0.742 | 9 | 5 | 5 | 10 |
| 476 | -1.280 | -1.207 | -0.975 | -1.305 | -1.327 | 12 | 6 | 7 | 10 |
| 477 | -1.046 | -0.356 | -0.987 | -0.598 | -1.128 | 16 | 5 | 8 | 12 |
| 478 | -0.156 | -0.301 | -0.737 | -0.004 | 0.026 | 12 | 5 | 7 | 9 |
| 479 | 0.266 | 0.640 | 0.216 | 0.270 | 0.022 | 9 | 10 | 6 | 15 |
| 480 | 2.091 | 1.307 | 2.157 | 1.670 | 1.016 | 16 | 7 | 15 | 9 |
| 481 | 0.600 | 0.328 | 0.802 | 0.016 | 0.404 | 15 | 5 | 5 | 8 |
| 482 | -0.925 | -0.975 | -0.687 | -1.197 | -0.795 | 9 | 7 | 7 | 10 |
| 483 | -1.382 | -0.603 | -1.753 | -1.529 | -1.430 | 18 | 6 | 11 | 7 |
| 484 | 0.780 | 1.327 | 0.742 | 1.203 | 0.262 | 10 | 6 | 7 | 10 |
| 485 | -1.064 | -1.100 | -0.747 | -0.818 | -0.714 | 7 | 7 | 5 | 8 |
| 486 | -0.416 | -0.100 | 0.252 | -0.591 | -0.703 | 16 | 9 | 5 | 12 |
| 487 | -1.676 | -1.203 | -0.732 | -1.562 | -1.390 | 12 | 9 | 8 | 7 |
| 488 | -0.379 | -0.278 | 0.244 | -0.502 | -0.272 | 12 | 8 | 9 | 6 |
| 489 | 0.366 | 0.657 | 0.043 | 0.248 | 0.230 | 14 | 10 | 6 | 7 |
| 490 | -1.335 | -1.065 | -1.661 | -1.197 | -1.167 | 10 | 6 | 7 | 9 |
| 491 | -0.391 | -1.350 | -0.391 | -0.676 | -0.567 | 7 | 7 | 6 | 6 |
| 492 | 0.718 | 0.164 | 0.672 | 0.955 | 0.146 | 11 | 7 | 6 | 8 |
| 493 | 0.290 | 1.699 | 0.761 | 0.247 | -0.205 | 18 | 5 | 6 | 14 |
| 494 | 0.758 | 1.477 | 0.921 | 0.836 | 0.701 | 13 | 5 | 6 | 8 |
| 495 | 0.920 | 1.192 | 0.474 | 1.057 | 0.494 | 12 | 6 | 6 | 12 |
| 496 | -0.479 | 0.956 | -0.246 | -0.638 | -0.237 | 10 | 9 | 8 | 14 |
| 497 | 0.129 | 0.419 | -0.267 | 0.248 | 0.032 | 12 | 10 | 6 | 11 |
| 498 | 0.514 | 0.554 | 0.650 | 0.335 | 1.150 | 12 | 6 | 6 | 14 |
| 499 | 0.422 | 0.091 | 0.317 | 0.123 | 0.337 | 11 | 6 | 9 | 9 |
| 500 | 0.025 | -0.188 | 0.318 | -0.432 | -0.190 | 9 | 11 | 9 | 10 |
| 501 | -0.394 | 0.119 | -0.606 | -0.720 | -0.306 | 11 | 8 | 6 | 12 |
| 502 | -0.760 | -0.240 | -0.466 | -1.121 | -0.629 | 11 | 7 | 8 | 5 |
| 503 | 0.284 | 0.443 | 0.238 | -0.036 | 0.446 | 9 | 10 | 8 | 16 |
| 504 | 1.021 | 1.040 | 0.940 | 1.351 | 0.592 | 12 | 5 | 7 | 10 |
| 505 | -0.004 | -0.105 | -0.383 | -0.161 | 0.033 | 12 | 8 | 7 | 8 |
| 506 | -0.406 | -0.681 | -0.686 | 0.042 | -0.493 | 7 | 6 | 13 | 7 |
| 507 | -0.729 | -0.598 | -0.317 | -0.769 | -0.673 | 11 | 10 | 4 | 10 |
| 508 | 0.280 | 0.130 | 0.438 | 0.681 | 0.016 | 10 | 7 | 6 | 6 |
| 509 | 1.976 | 1.601 | 2.273 | 2.059 | 1.996 | 9 | 9 | 10 | 17 |
| 510 | 0.849 | 0.582 | 1.401 | 0.624 | 1.082 | 10 | 10 | 6 | 7 |
| 511 | -0.411 | -0.527 | -0.694 | 0.204 | 0.266 | 8 | 6 | 6 | 9 |
| 512 | 2.585 | 1.570 | 2.010 | 2.231 | 2.375 | 17 | 7 | 14 | 17 |
| 513 | -1.095 | -1.277 | -1.181 | -1.326 | -1.221 | 7 | 6 | 7 | 10 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า θ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 514 | 0.820 | -0.488 | 0.844 | 0.811 | 0.601 | 11 | 5 | 6 | 6 |
| 515 | -0.244 | -0.164 | -0.387 | 0.058 | 0.059 | 15 | 7 | 6 | 8 |
| 516 | 1.385 | 1.408 | 1.264 | 1.147 | 0.718 | 13 | 9 | 9 | 8 |
| 517 | 0.772 | -0.798 | 0.574 | 0.619 | 0.641 | 11 | 7 | 6 | 7 |
| 518 | -0.725 | -0.848 | -0.696 | -0.586 | -0.954 | 12 | 6 | 11 | 7 |
| 519 | -1.713 | -1.328 | -1.278 | -1.187 | -1.462 | 13 | 7 | 10 | 10 |
| 520 | 0.302 | -0.195 | 0.219 | 0.447 | 0.163 | 13 | 11 | 5 | 6 |
| 521 | 0.990 | 0.196 | 0.989 | 1.050 | 0.511 | 12 | 6 | 7 | 9 |
| 522 | -2.188 | -1.919 | -1.723 | -2.039 | -1.756 | 14 | 7 | 12 | 12 |
| 523 | -0.955 | -0.864 | -1.172 | -0.978 | -0.596 | 9 | 5 | 6 | 7 |
| 524 | -0.038 | -0.099 | -0.150 | 0.077 | -0.270 | 9 | 8 | 5 | 11 |
| 525 | 0.062 | 0.497 | -0.217 | 0.030 | 0.168 | 13 | 10 | 6 | 7 |
| 526 | 0.076 | 0.162 | 0.045 | -0.347 | -0.065 | 8 | 7 | 6 | 10 |
| 527 | 2.370 | 2.597 | 1.864 | 2.243 | 1.952 | 18 | 9 | 8 | 12 |
| 528 | 1.144 | 1.252 | 0.922 | 1.141 | 0.864 | 11 | 5 | 9 | 8 |
| 529 | 1.038 | 1.629 | 0.806 | 0.696 | 1.289 | 14 | 5 | 7 | 15 |
| 530 | 0.003 | 0.640 | 0.043 | 0.001 | -0.006 | 11 | 9 | 7 | 16 |
| 531 | 1.070 | 0.908 | 0.975 | 0.933 | 0.709 | 15 | 6 | 6 | 7 |
| 532 | -1.051 | -1.283 | -1.176 | -0.703 | -1.023 | 13 | 5 | 12 | 8 |
| 533 | -1.285 | -0.586 | -0.981 | -1.124 | -1.285 | 9 | 5 | 7 | 8 |
| 534 | -1.416 | -1.700 | -1.431 | -1.397 | -1.186 | 11 | 5 | 9 | 9 |
| 535 | -0.378 | -0.100 | -0.696 | -0.396 | -0.091 | 11 | 6 | 7 | 11 |
| 536 | -2.057 | -1.981 | -1.690 | -1.602 | -1.916 | 7 | 6 | 11 | 13 |
| 537 | 0.948 | 1.328 | 1.114 | 1.011 | 0.857 | 14 | 6 | 9 | 11 |
| 538 | -0.049 | -0.136 | 0.453 | -0.142 | 0.525 | 10 | 7 | 8 | 9 |
| 539 | -0.562 | -0.963 | -0.403 | -0.349 | -0.462 | 8 | 9 | 8 | 16 |
| 540 | 0.215 | 0.934 | 0.149 | 0.489 | 0.451 | 13 | 9 | 7 | 13 |
| 541 | 1.075 | 0.655 | 1.182 | 1.104 | 1.379 | 11 | 5 | 6 | 12 |
| 542 | -0.520 | -0.745 | -0.294 | -0.747 | -0.496 | 12 | 7 | 6 | 13 |
| 543 | -1.159 | -1.054 | -0.975 | -0.595 | -1.015 | 11 | 6 | 7 | 12 |
| 544 | 0.389 | 0.482 | 0.444 | 0.245 | -0.104 | 14 | 9 | 5 | 11 |
| 545 | -0.379 | 0.550 | 0.033 | -0.765 | -0.519 | 13 | 9 | 8 | 7 |
| 546 | -0.473 | -1.141 | -0.402 | -0.663 | -0.764 | 11 | 5 | 6 | 8 |
| 547 | 1.889 | 1.795 | 1.138 | 1.063 | 1.156 | 17 | 6 | 15 | 21 |
| 548 | -2.053 | -1.537 | -1.724 | -2.078 | -1.733 | 12 | 8 | 11 | 13 |
| 549 | -0.869 | -0.618 | -0.697 | -1.182 | -1.233 | 11 | 6 | 7 | 7 |
| 550 | -0.505 | -1.117 | -0.921 | -0.746 | -0.584 | 10 | 6 | 7 | 7 |
| 551 | -0.815 | -0.886 | -1.352 | -0.925 | -0.710 | 18 | 5 | 9 | 9 |
| 552 | 0.397 | 0.752 | 0.614 | 0.391 | -0.010 | 11 | 7 | 6 | 11 |
| 553 | 0.238 | 0.287 | 0.036 | 0.490 | 0.354 | 11 | 8 | 7 | 7 |
| 554 | -1.334 | -1.132 | -1.375 | -1.148 | -1.433 | 8 | 5 | 10 | 5 |
| 555 | 0.863 | 0.716 | 0.731 | 0.489 | 0.992 | 12 | 6 | 6 | 9 |
| 556 | -0.897 | -1.358 | -0.970 | -1.105 | -0.576 | 13 | 6 | 7 | 10 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 557 | 0.557 | 0.481 | 0.769 | 0.712 | 0.721 | 10 | 5 | 5 | 8 |
| 558 | 1.699 | 0.975 | 1.333 | 1.775 | 1.470 | 13 | 7 | 15 | 10 |
| 559 | 0.727 | 0.073 | 0.656 | 0.666 | 0.487 | 12 | 8 | 7 | 9 |
| 560 | -0.529 | -0.109 | -0.460 | 0.253 | -0.099 | 13 | 9 | 8 | 8 |
| 561 | 0.218 | 0.773 | 0.056 | -0.078 | 0.700 | 10 | 10 | 5 | 14 |
| 562 | 0.186 | 0.077 | -0.233 | -0.143 | 0.396 | 9 | 10 | 7 | 15 |
| 563 | 0.372 | 0.411 | 0.428 | 0.012 | 0.330 | 12 | 6 | 8 | 9 |
| 564 | -0.510 | -0.712 | -0.597 | -0.670 | -0.920 | 10 | 6 | 5 | 8 |
| 565 | 0.857 | 0.825 | 0.840 | 0.891 | 1.035 | 9 | 7 | 6 | 12 |
| 566 | 0.207 | 0.743 | 0.004 | -0.358 | 0.218 | 10 | 10 | 7 | 11 |
| 567 | 0.705 | 1.189 | 0.793 | 1.086 | 1.277 | 13 | 5 | 9 | 14 |
| 568 | 0.729 | 1.166 | 1.331 | 1.051 | 0.713 | 17 | 7 | 6 | 8 |
| 569 | 0.498 | -0.530 | 0.663 | 0.888 | 0.293 | 9 | 7 | 8 | 10 |
| 570 | -1.543 | -0.884 | -1.720 | -1.394 | -1.594 | 9 | 7 | 7 | 8 |
| 571 | 1.936 | 1.529 | 1.897 | 1.612 | 1.993 | 12 | 7 | 8 | 11 |
| 572 | -2.648 | -1.221 | -2.076 | -2.524 | -1.822 | 13 | 7 | 9 | 12 |
| 573 | 0.137 | -0.022 | -0.022 | -0.057 | 0.295 | 10 | 9 | 9 | 8 |
| 574 | -1.658 | -2.269 | -1.109 | -1.737 | -1.555 | 8 | 8 | 8 | 12 |
| 575 | -0.622 | -0.783 | -0.461 | -0.517 | -0.829 | 6 | 8 | 7 | 10 |
| 576 | -0.580 | -0.722 | -0.567 | -0.732 | -0.630 | 9 | 8 | 6 | 9 |
| 577 | -0.932 | -0.555 | -0.993 | -0.937 | -0.696 | 7 | 5 | 7 | 13 |
| 578 | -0.802 | -0.162 | -1.248 | -0.983 | -0.866 | 12 | 6 | 7 | 9 |
| 579 | -1.272 | -0.122 | -1.417 | -1.129 | -1.255 | 10 | 6 | 7 | 10 |
| 580 | -0.359 | -0.083 | -0.306 | -0.291 | -0.472 | 11 | 7 | 6 | 10 |
| 581 | 1.744 | 2.180 | 1.538 | 1.730 | 1.622 | 14 | 8 | 15 | 11 |
| 582 | -0.386 | -0.776 | -0.691 | -0.366 | -0.297 | 10 | 8 | 5 | 11 |
| 583 | -0.724 | -0.728 | -0.828 | -0.230 | -0.862 | 10 | 7 | 5 | 7 |
| 584 | 0.253 | -0.886 | 0.724 | 0.329 | 0.642 | 9 | 7 | 5 | 8 |
| 585 | 0.522 | 0.886 | 0.601 | 0.714 | 0.200 | 11 | 7 | 7 | 9 |
| 586 | 0.371 | 1.098 | 0.320 | 0.466 | 0.178 | 13 | 8 | 6 | 8 |
| 587 | 1.390 | 1.060 | 0.947 | 0.837 | 1.033 | 9 | 5 | 6 | 10 |
| 588 | -0.415 | -0.706 | -0.033 | -0.291 | -0.857 | 8 | 10 | 6 | 11 |
| 589 | -2.053 | -1.199 | -1.541 | -1.999 | -2.096 | 17 | 7 | 8 | 12 |
| 590 | 0.286 | 0.375 | 0.449 | 0.330 | 0.269 | 7 | 9 | 19 | 8 |
| 591 | 0.192 | 0.245 | 0.052 | 0.354 | 0.421 | 9 | 12 | 6 | 14 |
| 592 | -0.558 | -0.575 | -0.776 | -0.309 | -0.422 | 11 | 6 | 7 | 11 |
| 593 | 0.418 | 0.590 | 0.271 | 0.169 | -0.037 | 7 | 8 | 8 | 10 |
| 594 | 0.012 | 0.252 | 0.031 | 0.203 | -0.073 | 13 | 11 | 6 | 7 |
| 595 | -1.419 | -0.568 | -1.416 | -1.364 | -1.230 | 7 | 6 | 8 | 11 |
| 596 | 1.435 | 1.268 | 1.260 | 0.944 | 0.566 | 10 | 6 | 6 | 12 |
| 597 | 0.673 | 0.160 | 0.517 | 0.739 | 0.749 | 11 | 7 | 11 | 7 |
| 598 | -0.421 | -0.500 | -0.787 | -0.331 | -0.663 | 8 | 5 | 8 | 8 |
| 599 | 1.395 | 1.918 | 1.415 | 0.087 | 1.103 | 14 | 6 | 5 | 14 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 600 | -1.541 | -0.590 | -1.067 | -1.268 | -1.481 | 6 | 7 | 10 | 8 |
| 601 | 0.293 | -0.165 | 0.512 | 0.114 | 0.531 | 10 | 6 | 13 | 9 |
| 602 | 0.073 | -0.193 | 0.136 | -0.044 | -0.041 | 7 | 9 | 5 | 8 |
| 603 | 0.480 | 0.196 | 0.326 | 1.021 | -0.037 | 19 | 7 | 6 | 9 |
| 604 | 0.151 | -0.406 | 0.199 | 0.612 | 0.205 | 10 | 8 | 6 | 12 |
| 605 | 0.051 | -0.355 | 0.508 | -0.005 | 0.199 | 10 | 8 | 7 | 10 |
| 606 | -0.015 | -0.710 | 0.047 | -0.309 | 0.377 | 10 | 8 | 7 | 9 |
| 607 | 1.547 | 2.245 | 1.452 | 1.373 | 1.444 | 14 | 7 | 10 | 10 |
| 608 | 0.012 | -0.193 | 0.296 | -0.044 | 0.411 | 14 | 8 | 9 | 9 |
| 609 | -1.074 | 0.190 | -1.230 | -1.211 | -1.580 | 15 | 6 | 5 | 9 |
| 610 | 0.127 | -0.104 | 0.324 | 0.399 | -0.395 | 11 | 8 | 5 | 15 |
| 611 | -0.577 | -1.073 | -0.042 | -0.523 | -0.384 | 11 | 8 | 6 | 19 |
| 612 | 0.111 | -0.430 | -0.200 | -0.293 | 0.129 | 20 | 8 | 5 | 9 |
| 613 | -0.694 | -0.230 | -0.823 | -0.309 | -0.412 | 12 | 8 | 8 | 6 |
| 614 | -1.347 | -1.273 | -1.160 | -1.225 | -1.256 | 6 | 7 | 12 | 6 |
| 615 | -0.636 | -0.563 | -0.785 | -0.776 | -0.960 | 9 | 5 | 9 | 10 |
| 616 | -0.081 | 0.230 | 0.485 | -0.096 | -0.300 | 11 | 6 | 5 | 8 |
| 617 | 0.374 | 0.105 | 0.440 | 0.508 | -0.022 | 17 | 7 | 6 | 8 |
| 618 | 0.150 | 0.393 | 0.135 | 0.204 | -0.122 | 11 | 11 | 6 | 13 |
| 619 | 1.107 | 1.035 | 1.148 | 1.747 | 0.762 | 11 | 6 | 12 | 8 |
| 620 | 0.217 | 0.067 | 0.057 | 0.044 | 0.456 | 13 | 10 | 6 | 8 |
| 621 | -0.638 | -0.893 | -0.352 | -0.908 | -0.785 | 6 | 7 | 10 | 7 |
| 622 | 0.193 | 0.242 | -0.030 | 0.007 | 0.147 | 8 | 8 | 5 | 7 |
| 623 | -1.848 | -1.942 | -1.935 | -1.129 | -2.117 | 10 | 7 | 12 | 9 |
| 624 | -0.669 | -0.429 | -0.617 | -0.292 | -0.506 | 10 | 6 | 7 | 10 |
| 625 | 1.684 | 1.297 | 1.693 | 1.449 | 1.445 | 11 | 7 | 14 | 15 |
| 626 | -2.316 | -1.072 | -2.136 | -1.986 | -1.761 | 13 | 7 | 11 | 15 |
| 627 | 0.159 | 0.116 | -0.211 | 0.090 | 0.379 | 17 | 10 | 5 | 11 |
| 628 | 0.555 | 0.653 | 0.731 | 0.614 | 0.683 | 9 | 5 | 5 | 9 |
| 629 | 0.759 | 0.808 | 0.840 | 0.924 | 0.943 | 11 | 7 | 6 | 9 |
| 630 | -0.546 | -0.031 | -0.460 | -0.319 | -0.656 | 11 | 6 | 8 | 11 |
| 631 | -0.006 | 0.655 | 0.264 | 0.081 | 0.277 | 6 | 7 | 5 | 6 |
| 632 | 0.355 | -0.127 | 0.338 | -0.340 | 0.117 | 15 | 11 | 7 | 14 |
| 633 | 2.298 | 2.149 | 1.718 | 1.655 | 1.262 | 13 | 8 | 8 | 8 |
| 634 | 1.194 | 1.121 | 1.639 | 1.027 | 1.603 | 12 | 7 | 6 | 10 |
| 635 | 0.174 | -0.364 | -0.033 | 0.154 | -0.228 | 9 | 9 | 4 | 9 |
| 636 | -1.263 | -1.366 | -1.138 | -1.165 | -0.498 | 5 | 8 | 7 | 6 |
| 637 | 1.113 | 1.813 | 0.714 | 1.166 | 0.818 | 13 | 6 | 15 | 7 |
| 638 | 1.310 | 1.564 | 1.173 | 1.554 | 1.142 | 13 | 8 | 13 | 7 |
| 639 | -0.543 | -0.380 | -0.697 | -0.461 | -0.480 | 13 | 6 | 5 | 8 |
| 640 | -0.310 | -0.574 | -0.692 | -0.516 | -0.299 | 12 | 7 | 6 | 9 |
| 641 | -0.235 | 0.794 | -0.290 | -0.794 | -0.194 | 12 | 8 | 5 | 7 |
| 642 | -0.689 | -0.402 | -0.691 | -0.942 | -1.018 | 9 | 6 | 7 | 11 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 643 | 1.607 | 1.943 | 1.367 | 1.921 | 1.432 | 19 | 7 | 8 | 10 |
| 644 | -2.609 | -2.129 | -2.038 | -2.105 | -2.064 | 11 | 7 | 10 | 9 |
| 645 | 1.101 | 1.594 | 1.172 | 0.915 | 0.908 | 9 | 5 | 9 | 8 |
| 646 | -0.451 | -0.081 | -0.745 | -0.045 | -0.215 | 10 | 7 | 8 | 13 |
| 647 | 1.027 | 0.672 | 1.477 | 0.510 | 0.694 | 11 | 7 | 8 | 6 |
| 648 | 0.704 | 1.186 | 0.667 | 0.697 | 0.401 | 16 | 6 | 20 | 8 |
| 649 | -0.193 | -0.963 | 0.008 | 0.345 | -0.337 | 13 | 9 | 9 | 14 |
| 650 | 0.700 | 0.832 | 0.743 | 0.615 | 0.553 | 10 | 8 | 6 | 13 |
| 651 | 1.311 | 0.684 | 0.822 | 1.197 | 1.314 | 10 | 5 | 7 | 14 |
| 652 | -2.070 | -2.683 | -1.543 | -1.357 | -2.341 | 9 | 7 | 8 | 9 |
| 653 | -0.217 | 0.269 | -0.221 | -0.327 | -0.729 | 10 | 9 | 6 | 11 |
| 654 | -0.267 | -0.103 | -0.046 | -0.136 | 0.342 | 14 | 8 | 5 | 11 |
| 655 | -0.300 | -0.730 | 0.732 | -0.514 | -0.283 | 19 | 5 | 7 | 7 |
| 656 | 0.728 | 1.219 | 0.507 | 0.572 | 0.119 | 8 | 6 | 7 | 9 |
| 657 | -0.471 | -0.124 | 0.410 | -0.491 | 0.143 | 11 | 9 | 5 | 7 |
| 658 | 0.438 | 0.362 | 0.376 | 0.327 | 0.157 | 12 | 8 | 5 | 7 |
| 659 | -0.177 | 0.071 | -0.504 | 0.153 | 0.016 | 10 | 8 | 7 | 15 |
| 660 | -0.750 | -0.058 | -0.604 | -0.615 | -0.921 | 11 | 7 | 6 | 11 |
| 661 | -1.030 | -0.038 | -0.713 | -0.870 | -1.175 | 14 | 6 | 8 | 10 |
| 662 | 0.844 | 1.063 | 0.284 | 0.955 | 1.108 | 9 | 7 | 6 | 9 |
| 663 | 1.359 | 1.074 | 1.562 | 0.999 | 0.809 | 9 | 8 | 6 | 8 |
| 664 | 1.263 | 0.598 | 1.154 | 1.160 | 1.826 | 10 | 7 | 8 | 13 |
| 665 | -1.295 | -0.913 | -1.391 | -1.109 | -1.001 | 11 | 6 | 6 | 7 |
| 666 | -0.359 | -0.489 | -0.698 | -0.167 | -0.510 | 14 | 6 | 5 | 11 |
| 667 | 0.513 | 0.447 | 0.274 | 0.406 | 0.187 | 13 | 8 | 6 | 13 |
| 668 | 0.479 | 0.443 | 0.030 | 0.683 | 0.274 | 11 | 10 | 5 | 9 |
| 669 | 1.591 | 0.995 | 1.682 | 1.832 | 1.347 | 9 | 7 | 14 | 13 |
| 670 | -0.529 | -0.681 | -0.169 | -0.336 | -0.597 | 9 | 10 | 8 | 10 |
| 671 | -1.430 | -2.295 | -1.433 | -1.419 | -1.202 | 9 | 6 | 7 | 12 |
| 672 | 0.705 | 0.635 | 0.889 | 0.362 | 0.674 | 10 | 6 | 6 | 7 |
| 673 | 0.081 | 0.640 | 0.217 | 0.378 | -0.115 | 10 | 7 | 6 | 6 |
| 674 | 0.650 | 1.232 | 0.901 | 0.738 | 0.077 | 16 | 5 | 6 | 12 |
| 675 | 1.756 | 2.145 | 1.473 | 1.575 | 1.878 | 17 | 6 | 12 | 12 |
| 676 | 1.008 | 1.146 | 1.177 | 0.987 | 1.236 | 16 | 5 | 7 | 10 |
| 677 | -0.743 | -0.954 | -0.524 | -1.399 | -0.703 | 9 | 8 | 8 | 8 |
| 678 | 0.084 | 0.241 | 0.034 | -0.023 | -0.055 | 17 | 9 | 6 | 5 |
| 679 | -0.555 | -0.373 | -0.280 | -0.349 | -0.044 | 9 | 11 | 6 | 13 |
| 680 | -0.251 | -0.234 | -0.207 | 0.270 | 0.065 | 13 | 11 | 6 | 8 |
| 681 | -2.098 | -2.304 | -1.726 | -1.638 | -2.111 | 15 | 7 | 12 | 11 |
| 682 | 0.778 | 0.726 | 0.859 | 0.817 | 1.006 | 12 | 5 | 5 | 9 |
| 683 | -0.435 | -0.059 | 0.008 | -0.328 | -0.685 | 11 | 9 | 8 | 10 |
| 684 | 0.107 | 0.504 | 0.056 | 0.255 | 0.135 | 16 | 10 | 6 | 6 |
| 685 | 1.860 | 1.995 | 1.856 | 1.628 | 1.695 | 13 | 7 | 30 | 15 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 686 | 0.179 | 0.208 | 0.001 | -0.286 | 0.532 | 10 | 10 | 5 | 7 |
| 687 | 0.134 | 0.228 | 0.089 | -0.340 | -0.529 | 11 | 10 | 6 | 6 |
| 688 | 2.530 | 3.749 | 2.486 | 1.573 | 2.870 | 30 | 9 | 12 | 17 |
| 689 | -1.150 | -0.291 | -1.171 | -1.219 | -1.242 | 13 | 5 | 6 | 8 |
| 690 | 1.509 | 0.654 | 1.257 | 0.709 | 1.169 | 12 | 6 | 5 | 15 |
| 691 | 0.210 | 0.813 | 0.035 | 0.368 | 0.389 | 12 | 9 | 7 | 12 |
| 692 | -0.678 | -0.962 | -0.475 | -1.012 | -1.109 | 13 | 10 | 10 | 7 |
| 693 | 1.321 | 1.964 | 1.732 | 1.454 | 1.283 | 16 | 7 | 14 | 9 |
| 694 | -2.183 | -2.188 | -2.144 | -0.723 | -2.297 | 12 | 8 | 5 | 11 |
| 695 | 0.370 | 0.702 | 0.652 | -0.422 | 0.016 | 10 | 5 | 6 | 6 |
| 696 | -0.656 | -0.489 | -0.652 | -0.309 | -0.564 | 14 | 9 | 11 | 6 |
| 697 | 1.090 | 1.928 | 1.231 | 1.557 | 0.701 | 14 | 7 | 6 | 12 |
| 698 | -2.408 | -0.955 | -1.407 | -1.711 | -2.187 | 9 | 8 | 11 | 13 |
| 699 | 0.397 | 0.066 | 0.032 | 0.261 | 0.268 | 14 | 9 | 9 | 8 |
| 700 | 0.713 | 1.126 | 0.887 | 0.751 | 0.611 | 8 | 5 | 5 | 9 |
| 701 | -1.481 | -1.092 | -1.302 | -1.144 | -1.252 | 11 | 10 | 11 | 6 |
| 702 | 0.269 | 0.115 | 0.415 | 0.456 | 0.311 | 12 | 9 | 6 | 9 |
| 703 | -1.517 | -1.144 | -1.554 | -1.998 | -1.224 | 9 | 7 | 8 | 11 |
| 704 | -0.870 | -1.132 | -0.564 | -0.552 | -0.151 | 8 | 8 | 5 | 13 |
| 705 | -0.749 | -1.005 | -0.931 | -0.633 | -0.121 | 10 | 6 | 7 | 7 |
| 706 | -0.196 | 0.195 | 0.397 | 0.022 | -0.275 | 10 | 9 | 5 | 11 |
| 707 | -2.181 | -2.640 | -2.050 | -1.863 | -2.133 | 8 | 7 | 9 | 11 |
| 708 | -1.190 | -1.775 | -1.351 | -1.166 | -1.099 | 11 | 5 | 8 | 8 |
| 709 | 0.389 | 0.659 | 0.230 | 0.645 | 0.399 | 15 | 8 | 5 | 10 |
| 710 | -0.201 | -0.119 | -0.418 | 0.000 | 0.114 | 8 | 7 | 7 | 12 |
| 711 | 0.807 | 1.045 | 0.536 | 0.817 | 0.762 | 9 | 6 | 5 | 7 |
| 712 | 0.132 | -0.879 | 0.135 | 0.435 | 0.291 | 7 | 11 | 5 | 8 |
| 713 | -0.361 | 0.054 | -0.702 | -0.671 | -0.413 | 10 | 7 | 5 | 6 |
| 714 | -0.677 | -1.438 | -0.575 | -0.632 | -0.721 | 14 | 7 | 7 | 8 |
| 715 | -0.035 | -0.368 | -0.035 | 0.085 | 0.049 | 12 | 10 | 4 | 8 |
| 716 | -2.832 | -0.821 | -1.756 | -1.748 | -1.479 | 10 | 9 | 12 | 15 |
| 717 | -0.881 | -0.869 | -1.234 | -0.482 | -1.062 | 10 | 6 | 5 | 11 |
| 718 | -0.082 | 0.174 | -0.496 | -0.317 | -0.035 | 10 | 10 | 5 | 13 |
| 719 | -0.724 | -0.878 | -0.715 | -0.755 | -0.021 | 8 | 7 | 9 | 11 |
| 720 | 1.643 | 1.887 | 0.919 | 1.659 | 2.075 | 10 | 5 | 7 | 13 |
| 721 | 0.911 | 0.816 | 0.916 | 0.326 | 0.881 | 11 | 8 | 5 | 8 |
| 722 | 0.134 | -0.668 | -0.216 | 1.139 | 0.541 | 8 | 11 | 15 | 10 |
| 723 | -0.594 | -1.149 | -0.651 | -0.608 | 0.037 | 13 | 8 | 4 | 11 |
| 724 | 1.417 | 1.427 | 1.213 | 1.074 | 1.497 | 14 | 9 | 7 | 10 |
| 725 | -1.051 | -1.349 | -1.177 | -1.118 | -0.968 | 7 | 5 | 7 | 6 |
| 726 | 0.674 | 0.584 | 0.660 | 0.029 | 0.472 | 9 | 6 | 6 | 9 |
| 727 | 0.516 | 0.831 | 0.239 | 0.693 | 0.585 | 9 | 9 | 5 | 6 |
| 728 | 0.992 | 0.162 | 0.867 | 0.581 | 0.685 | 8 | 5 | 7 | 8 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 729 | 1.135 | 1.021 | 0.944 | 1.418 | 1.197 | 13 | 5 | 16 | 15 |
| 730 | -3.224 | -4.074 | -2.664 | -2.289 | -2.584 | 30 | 9 | 10 | 15 |
| 731 | 0.888 | 0.998 | 0.879 | 1.044 | 0.706 | 12 | 5 | 6 | 9 |
| 732 | -0.387 | -0.975 | -0.026 | -0.527 | -0.390 | 10 | 10 | 6 | 8 |
| 733 | -0.432 | -0.643 | -0.906 | -0.394 | -0.149 | 12 | 6 | 7 | 7 |
| 734 | 0.023 | -0.447 | -0.265 | 0.387 | -0.341 | 8 | 9 | 9 | 7 |
| 735 | 0.447 | 0.409 | 0.242 | 0.619 | 0.245 | 14 | 8 | 5 | 7 |
| 736 | -0.262 | -0.878 | -0.205 | -0.634 | -0.226 | 13 | 10 | 8 | 9 |
| 737 | 0.480 | 0.079 | 0.657 | 0.248 | 0.536 | 12 | 6 | 6 | 7 |
| 738 | -1.822 | -2.298 | -1.060 | -1.845 | -1.782 | 11 | 8 | 9 | 9 |
| 739 | 1.683 | 2.225 | 1.518 | 1.776 | 1.936 | 19 | 8 | 11 | 14 |
| 740 | 0.291 | 0.321 | 0.443 | 0.232 | 0.293 | 14 | 8 | 5 | 10 |
| 741 | 1.731 | 1.626 | 1.538 | 1.825 | 1.115 | 11 | 7 | 13 | 12 |
| 742 | 0.977 | 0.757 | 1.166 | 0.713 | 0.448 | 10 | 7 | 5 | 6 |
| 743 | 1.067 | 0.081 | 1.395 | 0.599 | 0.343 | 9 | 7 | 7 | 11 |
| 744 | 0.090 | 0.701 | -0.032 | 0.444 | 0.282 | 9 | 10 | 6 | 14 |
| 745 | -0.370 | 0.106 | 0.178 | -0.779 | -0.473 | 11 | 7 | 5 | 8 |
| 746 | 0.166 | -0.188 | 0.192 | -0.506 | 0.554 | 11 | 9 | 5 | 7 |
| 747 | 1.631 | 1.583 | 2.002 | 1.186 | 1.036 | 11 | 7 | 10 | 9 |
| 748 | -0.463 | -0.248 | -0.023 | 0.105 | -0.044 | 12 | 9 | 5 | 9 |
| 749 | 0.919 | 0.017 | 0.744 | 1.148 | 0.862 | 13 | 5 | 6 | 6 |
| 750 | -0.241 | -0.374 | 0.042 | -0.408 | -0.180 | 13 | 9 | 5 | 13 |
| 751 | -0.744 | -1.157 | -0.750 | -0.840 | -0.535 | 13 | 6 | 7 | 10 |
| 752 | -0.914 | -0.487 | -1.175 | -0.856 | -1.230 | 12 | 6 | 8 | 11 |
| 753 | -0.655 | -0.622 | -0.792 | -0.573 | -0.802 | 14 | 6 | 5 | 10 |
| 754 | 0.898 | 1.265 | 0.702 | 0.934 | 0.936 | 13 | 6 | 6 | 9 |
| 755 | -0.020 | -0.482 | -0.227 | -0.193 | 0.009 | 10 | 8 | 7 | 7 |
| 756 | -0.038 | 1.332 | -0.451 | 0.444 | 0.168 | 10 | 7 | 6 | 18 |
| 757 | 0.414 | 0.671 | 0.247 | 0.955 | 0.555 | 10 | 10 | 6 | 11 |
| 758 | 0.124 | 0.041 | 0.150 | 0.003 | 0.098 | 8 | 8 | 5 | 9 |
| 759 | 0.546 | 0.854 | 0.863 | -0.052 | 0.527 | 11 | 5 | 7 | 6 |
| 760 | 0.546 | 0.599 | 0.836 | 0.397 | 0.517 | 15 | 5 | 5 | 9 |
| 761 | -1.259 | 0.156 | -1.192 | -1.413 | -0.756 | 13 | 5 | 8 | 10 |
| 762 | -1.059 | -1.647 | -0.761 | -1.260 | -1.306 | 11 | 7 | 8 | 8 |
| 763 | -0.004 | -0.273 | 0.058 | -0.415 | 0.197 | 10 | 11 | 6 | 12 |
| 764 | 0.659 | 1.788 | 0.433 | 0.889 | 0.797 | 19 | 9 | 4 | 9 |
| 765 | -0.125 | 0.396 | -0.257 | 0.016 | -0.291 | 10 | 9 | 6 | 7 |
| 766 | -0.698 | -0.153 | -0.373 | -0.785 | -0.595 | 12 | 8 | 8 | 10 |
| 767 | 0.524 | 1.202 | 0.396 | 0.273 | 0.433 | 14 | 5 | 8 | 6 |
| 768 | 0.931 | 1.291 | 0.508 | 0.399 | 0.988 | 12 | 7 | 6 | 7 |
| 769 | -0.571 | 0.134 | -0.407 | -0.389 | -0.460 | 10 | 6 | 5 | 9 |
| 770 | 0.106 | -0.237 | -0.306 | 0.007 | 0.016 | 12 | 8 | 5 | 11 |
| 771 | 1.371 | 0.680 | 0.736 | 1.803 | 1.564 | 11 | 5 | 10 | 11 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า θ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 772 | -0.035 | 0.093 | -0.032 | -0.306 | 0.059 | 9 | 8 | 7 | 14 |
| 773 | 0.791 | 0.852 | 1.159 | 0.696 | 1.093 | 9 | 8 | 5 | 10 |
| 774 | -0.431 | -1.669 | -0.418 | -0.626 | -0.454 | 12 | 9 | 13 | 11 |
| 775 | 0.280 | 0.695 | 0.041 | -0.349 | 0.174 | 6 | 9 | 6 | 7 |
| 776 | -0.279 | -0.051 | -0.385 | -0.334 | 0.002 | 16 | 8 | 6 | 7 |
| 777 | 0.220 | 0.693 | -0.030 | 0.319 | -0.027 | 13 | 9 | 5 | 11 |
| 778 | -0.522 | -1.110 | -0.828 | -0.774 | -0.486 | 8 | 7 | 9 | 7 |
| 779 | 0.803 | 0.664 | 0.625 | 0.666 | 0.729 | 11 | 6 | 6 | 12 |
| 780 | 0.442 | 0.536 | 0.755 | 0.209 | 0.027 | 16 | 7 | 6 | 12 |
| 781 | -1.532 | -1.020 | -1.393 | -1.412 | -1.078 | 8 | 6 | 5 | 10 |
| 782 | 0.641 | 0.285 | 1.151 | 0.149 | 0.765 | 11 | 6 | 7 | 10 |
| 783 | -1.898 | -1.571 | -2.049 | -1.206 | -1.664 | 10 | 7 | 13 | 8 |
| 784 | -0.570 | -0.210 | -0.688 | -1.174 | -0.447 | 13 | 7 | 8 | 7 |
| 785 | 0.162 | 0.257 | -0.195 | 0.252 | 0.664 | 13 | 9 | 5 | 9 |
| 786 | -0.858 | -0.364 | -0.671 | -0.761 | -0.960 | 9 | 7 | 10 | 10 |
| 787 | -0.923 | -0.467 | -0.778 | -0.841 | -0.879 | 8 | 6 | 7 | 6 |
| 788 | -0.519 | -0.726 | -0.618 | -0.749 | -0.750 | 11 | 8 | 6 | 11 |
| 789 | -0.009 | 0.647 | -0.306 | 0.110 | -0.145 | 12 | 8 | 9 | 13 |
| 790 | 0.639 | 0.227 | 1.202 | 0.224 | 0.526 | 11 | 9 | 6 | 8 |
| 791 | -0.613 | -0.338 | -1.173 | -0.515 | -0.572 | 13 | 5 | 7 | 10 |
| 792 | -0.082 | -0.692 | 0.033 | -0.028 | 0.117 | 10 | 9 | 9 | 13 |
| 793 | 0.667 | 0.197 | 0.840 | 0.822 | 0.355 | 11 | 9 | 6 | 7 |
| 794 | -1.393 | -1.087 | -1.614 | -1.511 | -1.158 | 9 | 6 | 7 | 10 |
| 795 | 1.630 | 1.336 | 1.415 | 1.581 | 2.021 | 7 | 6 | 12 | 15 |
| 796 | -1.184 | -1.257 | -0.828 | -1.408 | -1.209 | 8 | 7 | 7 | 12 |
| 797 | 2.227 | 2.255 | 1.875 | 1.928 | 1.583 | 17 | 7 | 8 | 8 |
| 798 | -0.325 | -0.072 | -0.261 | -0.491 | 0.037 | 9 | 9 | 9 | 10 |
| 799 | -0.676 | -0.384 | -1.349 | -0.793 | -0.784 | 8 | 5 | 4 | 6 |
| 800 | 0.474 | 0.627 | 0.733 | 0.737 | 1.055 | 13 | 6 | 9 | 8 |
| 801 | 0.745 | 0.277 | -0.001 | 0.954 | 0.948 | 14 | 7 | 6 | 8 |
| 802 | -0.262 | 0.117 | 0.228 | 0.181 | 0.229 | 10 | 9 | 7 | 5 |
| 803 | 1.106 | 0.679 | 0.966 | 0.897 | 1.740 | 11 | 7 | 7 | 12 |
| 804 | -0.641 | -0.966 | -0.676 | -0.799 | -0.469 | 11 | 6 | 5 | 13 |
| 805 | 2.677 | 1.914 | 2.131 | 2.457 | 2.907 | 19 | 8 | 9 | 13 |
| 806 | -0.003 | -0.066 | 0.668 | 0.261 | -0.142 | 12 | 6 | 6 | 9 |
| 807 | -0.491 | -0.939 | -0.788 | -0.328 | 0.179 | 10 | 5 | 8 | 7 |
| 808 | 0.400 | 1.153 | 0.675 | 0.034 | 0.591 | 13 | 5 | 6 | 9 |
| 809 | -0.517 | -0.879 | -0.697 | -0.750 | 0.295 | 13 | 7 | 9 | 13 |
| 810 | 0.923 | 2.184 | 0.968 | 0.917 | 0.771 | 15 | 6 | 6 | 8 |
| 811 | -0.099 | 0.361 | -0.452 | -0.331 | 0.089 | 14 | 6 | 7 | 7 |
| 812 | -0.411 | -0.423 | -0.785 | -0.346 | -0.574 | 17 | 5 | 5 | 13 |
| 813 | -0.368 | -0.257 | -0.312 | -0.397 | -0.386 | 8 | 11 | 6 | 18 |
| 814 | -0.266 | -0.078 | -0.206 | -0.292 | -0.075 | 14 | 9 | 5 | 10 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 815 | 0.401 | -0.063 | 0.134 | 0.351 | 0.332 | 10 | 9 | 7 | 10 |
| 816 | 0.191 | 0.397 | 0.216 | -0.040 | -0.343 | 10 | 10 | 8 | 17 |
| 817 | 1.420 | 0.905 | 0.616 | 1.792 | 1.665 | 12 | 6 | 15 | 11 |
| 818 | 0.589 | 0.711 | 0.606 | 0.112 | 0.743 | 12 | 8 | 9 | 6 |
| 819 | 0.928 | 0.678 | 1.002 | 1.411 | 0.489 | 9 | 8 | 16 | 7 |
| 820 | 0.677 | -0.435 | 0.639 | 0.342 | 0.478 | 13 | 7 | 8 | 6 |
| 821 | 0.232 | -0.055 | 0.386 | -0.128 | -0.220 | 17 | 8 | 5 | 11 |
| 822 | -0.563 | 0.527 | -0.459 | -0.438 | -0.671 | 16 | 9 | 5 | 11 |
| 823 | 0.983 | 1.745 | 0.815 | 0.966 | 1.103 | 11 | 7 | 5 | 8 |
| 824 | 0.780 | -0.056 | 0.615 | 0.766 | 0.708 | 7 | 6 | 8 | 14 |
| 825 | 1.016 | 0.233 | 0.996 | 1.211 | 0.094 | 8 | 7 | 7 | 9 |
| 826 | 1.538 | 2.220 | 0.890 | 1.647 | 1.747 | 15 | 6 | 12 | 17 |
| 827 | 0.576 | -0.768 | 0.289 | 0.571 | 0.742 | 10 | 9 | 7 | 8 |
| 828 | -0.150 | 0.308 | -0.413 | 0.032 | -0.155 | 12 | 7 | 6 | 6 |
| 829 | 0.136 | -0.092 | -0.320 | 0.326 | 0.152 | 15 | 7 | 5 | 9 |
| 830 | 1.386 | 0.485 | 1.197 | 1.855 | 1.034 | 9 | 7 | 16 | 7 |
| 831 | -0.273 | -0.690 | -0.791 | -0.104 | -0.083 | 9 | 7 | 6 | 16 |
| 832 | -0.344 | -0.315 | 0.046 | -0.349 | 0.021 | 10 | 9 | 8 | 14 |
| 833 | 0.228 | 1.307 | 0.032 | 0.120 | -0.070 | 10 | 9 | 7 | 10 |
| 834 | -2.348 | -2.165 | -2.268 | -2.044 | -1.705 | 12 | 9 | 11 | 14 |
| 835 | 0.237 | 0.080 | 0.056 | -0.049 | 0.375 | 13 | 8 | 8 | 9 |
| 836 | 0.780 | 0.209 | 0.398 | 0.330 | 0.562 | 12 | 10 | 19 | 16 |
| 837 | -0.193 | 0.056 | -0.601 | -0.077 | 0.241 | 12 | 6 | 9 | 10 |
| 838 | -0.830 | -0.179 | -0.158 | -0.945 | -0.643 | 6 | 8 | 10 | 7 |
| 839 | -0.914 | -0.739 | -0.433 | -0.538 | -0.542 | 15 | 8 | 5 | 17 |
| 840 | 0.437 | 0.122 | 0.470 | 0.354 | 0.067 | 12 | 9 | 6 | 8 |
| 841 | 0.117 | 0.992 | 0.722 | -0.163 | 0.558 | 12 | 6 | 7 | 10 |
| 842 | 0.583 | 0.966 | 0.236 | 0.330 | 0.542 | 8 | 6 | 9 | 6 |
| 843 | 0.825 | 0.719 | 0.796 | 0.756 | 0.328 | 13 | 5 | 7 | 9 |
| 844 | 0.832 | 1.408 | 1.121 | 0.976 | 0.654 | 16 | 6 | 6 | 8 |
| 845 | 0.198 | -0.226 | 0.245 | 0.020 | 0.246 | 12 | 7 | 5 | 10 |
| 846 | 0.259 | 1.088 | 0.325 | -0.152 | 0.194 | 13 | 7 | 5 | 9 |
| 847 | -0.114 | -0.182 | -0.376 | 0.010 | -0.294 | 10 | 10 | 5 | 13 |
| 848 | 0.113 | 0.311 | -0.033 | -0.344 | 0.231 | 9 | 10 | 9 | 8 |
| 849 | 1.610 | 1.941 | 1.464 | 1.697 | 1.590 | 12 | 8 | 12 | 15 |
| 850 | -1.476 | -0.869 | -1.250 | -1.738 | -1.012 | 6 | 8 | 8 | 17 |
| 851 | 0.269 | -0.113 | 0.211 | 0.206 | -0.252 | 11 | 10 | 6 | 8 |
| 852 | 2.380 | 2.350 | 2.350 | 1.648 | 1.107 | 11 | 8 | 12 | 12 |
| 853 | -0.815 | -1.191 | -0.436 | -0.755 | -0.542 | 9 | 8 | 8 | 8 |
| 854 | 0.349 | 0.480 | 0.199 | 0.681 | 0.275 | 12 | 9 | 7 | 5 |
| 855 | -0.630 | 0.763 | -0.266 | -0.336 | -0.669 | 14 | 7 | 6 | 9 |
| 856 | 1.727 | 2.176 | 2.117 | 1.747 | 1.843 | 20 | 7 | 12 | 16 |
| 857 | 1.075 | 0.757 | 0.972 | 1.207 | 0.956 | 10 | 7 | 10 | 10 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 858 | -1.394 | -0.133 | -1.618 | -1.395 | -1.377 | 13 | 6 | 9 | 11 |
| 859 | -1.167 | -0.659 | -0.890 | -1.271 | -1.279 | 13 | 5 | 10 | 8 |
| 860 | 1.646 | 1.539 | 1.283 | 1.449 | 1.342 | 10 | 7 | 16 | 13 |
| 861 | 0.330 | 0.251 | 0.035 | 0.255 | 0.636 | 9 | 9 | 6 | 9 |
| 862 | -0.533 | -0.589 | -0.522 | -0.469 | -0.300 | 10 | 7 | 9 | 8 |
| 863 | -0.268 | 0.472 | -0.263 | -0.745 | -0.160 | 16 | 8 | 5 | 7 |
| 864 | 0.253 | 0.014 | 0.429 | 0.101 | 0.000 | 16 | 8 | 5 | 8 |
| 865 | 0.800 | -0.148 | 0.806 | 1.065 | 0.762 | 19 | 7 | 15 | 10 |
| 866 | -0.071 | -0.305 | -0.360 | -0.020 | 0.167 | 7 | 8 | 7 | 11 |
| 867 | -0.909 | -0.799 | -0.787 | -1.209 | -0.810 | 8 | 9 | 7 | 7 |
| 868 | -0.238 | -0.740 | -0.703 | -0.003 | 0.446 | 6 | 6 | 5 | 8 |
| 869 | 1.434 | 0.925 | 1.146 | 1.600 | 1.321 | 9 | 6 | 12 | 9 |
| 870 | -2.842 | -2.618 | -2.672 | -2.262 | -2.312 | 12 | 9 | 11 | 17 |
| 871 | -0.426 | -1.332 | -0.257 | -0.843 | -0.844 | 10 | 9 | 7 | 10 |
| 872 | -0.092 | -0.224 | -0.272 | -0.053 | -0.270 | 10 | 8 | 11 | 12 |
| 873 | -1.646 | -0.337 | -1.634 | -1.890 | -1.628 | 11 | 6 | 9 | 15 |
| 874 | 0.307 | 0.750 | 0.324 | 0.105 | 0.642 | 10 | 8 | 5 | 8 |
| 875 | -0.090 | 0.558 | -0.402 | -0.004 | -0.241 | 13 | 5 | 6 | 11 |
| 876 | -0.442 | -0.753 | -0.408 | -0.733 | -0.312 | 13 | 11 | 9 | 8 |
| 877 | -0.810 | -1.686 | -0.258 | -0.708 | -0.827 | 12 | 10 | 6 | 5 |
| 878 | 2.472 | 2.335 | 2.301 | 2.638 | 2.154 | 16 | 10 | 13 | 12 |
| 879 | 0.922 | 0.114 | 1.473 | 0.713 | 1.124 | 9 | 7 | 6 | 7 |
| 880 | 0.063 | -0.466 | 0.032 | -0.342 | 0.210 | 12 | 9 | 10 | 8 |
| 881 | 0.800 | 1.103 | 0.263 | 0.919 | 0.762 | 11 | 8 | 7 | 10 |
| 882 | -0.196 | -0.803 | 0.033 | -0.355 | -0.113 | 15 | 9 | 5 | 11 |
| 883 | -0.160 | -0.137 | 0.300 | -0.027 | 0.252 | 9 | 7 | 6 | 7 |
| 884 | -0.686 | -1.042 | -0.760 | -0.753 | -0.506 | 7 | 6 | 6 | 7 |
| 885 | -2.539 | -1.707 | -1.702 | -2.271 | -1.951 | 12 | 6 | 11 | 12 |
| 886 | 0.683 | 0.438 | 0.706 | 1.017 | 0.489 | 9 | 5 | 7 | 10 |
| 887 | 0.756 | 1.237 | 0.731 | 0.700 | 1.060 | 13 | 6 | 5 | 14 |
| 888 | -1.107 | 0.193 | -0.788 | -1.087 | -1.020 | 12 | 5 | 8 | 8 |
| 889 | 0.374 | 0.264 | 0.221 | 0.699 | 0.707 | 12 | 9 | 5 | 10 |
| 890 | -0.942 | -0.989 | -0.976 | -0.928 | -1.551 | 14 | 6 | 7 | 8 |
| 891 | -1.035 | -0.695 | -1.454 | -1.225 | -1.020 | 8 | 6 | 7 | 8 |
| 892 | -0.402 | -0.639 | -0.231 | -0.295 | -0.430 | 11 | 9 | 5 | 6 |
| 893 | 1.293 | 2.195 | 1.910 | 0.985 | 1.690 | 22 | 8 | 9 | 19 |
| 894 | 2.951 | 11.352 | 2.492 | 2.738 | 1.863 | 30 | 10 | 14 | 8 |
| 895 | -0.001 | -0.257 | -0.493 | 0.330 | 0.439 | 10 | 10 | 19 | 8 |
| 896 | -0.001 | 0.146 | -0.368 | 0.011 | 0.065 | 11 | 5 | 7 | 8 |
| 897 | 0.849 | 0.964 | 1.240 | 1.169 | 0.718 | 9 | 6 | 16 | 9 |
| 898 | 0.124 | 0.128 | 0.235 | 0.480 | -0.123 | 11 | 8 | 6 | 9 |
| 899 | 2.204 | 1.416 | 2.141 | 1.974 | 1.284 | 11 | 7 | 10 | 19 |
| 900 | -0.883 | -0.997 | -1.166 | -0.795 | -1.415 | 10 | 7 | 12 | 7 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 901 | 0.851 | 1.341 | 0.665 | 0.777 | 1.213 | 15 | 6 | 8 | 9 |
| 902 | -0.076 | -0.472 | -0.626 | 0.229 | -0.426 | 12 | 7 | 7 | 8 |
| 903 | 1.721 | 2.416 | 1.129 | 1.588 | 1.634 | 17 | 9 | 12 | 11 |
| 904 | -0.195 | -0.102 | -0.323 | -0.027 | 0.079 | 13 | 7 | 7 | 11 |
| 905 | -0.728 | -0.951 | -0.868 | -0.792 | -1.041 | 13 | 6 | 5 | 7 |
| 906 | -1.450 | -2.054 | -1.114 | -1.111 | -1.185 | 12 | 7 | 7 | 9 |
| 907 | 1.358 | 1.385 | 1.656 | 1.598 | 0.874 | 10 | 7 | 12 | 9 |
| 908 | -1.809 | -1.577 | -1.798 | -1.714 | -1.433 | 11 | 6 | 9 | 6 |
| 909 | -1.801 | -1.423 | -1.795 | -0.998 | -1.339 | 17 | 9 | 9 | 8 |
| 910 | 0.216 | -0.525 | -0.032 | 0.347 | -0.023 | 11 | 9 | 5 | 9 |
| 911 | -0.049 | -0.622 | -0.058 | 0.234 | 0.183 | 10 | 8 | 6 | 10 |
| 912 | 1.327 | 1.308 | 1.534 | 1.490 | 1.838 | 10 | 7 | 13 | 13 |
| 913 | 0.914 | 1.446 | 0.751 | 1.525 | 0.309 | 10 | 5 | 12 | 8 |
| 914 | -0.199 | 0.239 | -0.303 | -0.092 | -0.108 | 11 | 8 | 8 | 10 |
| 915 | 1.383 | 1.398 | 1.458 | 1.776 | 1.026 | 8 | 7 | 13 | 10 |
| 916 | -0.765 | -0.688 | -0.985 | -0.628 | -0.266 | 17 | 6 | 8 | 7 |
| 917 | 1.523 | 0.980 | 1.529 | 1.546 | 1.029 | 12 | 8 | 13 | 12 |
| 918 | -0.202 | -0.365 | -0.219 | -0.360 | -0.035 | 13 | 12 | 5 | 14 |
| 919 | 0.713 | 1.154 | 0.696 | 0.412 | 0.926 | 10 | 6 | 8 | 6 |
| 920 | 0.418 | 0.506 | 0.490 | 0.570 | 0.797 | 11 | 6 | 8 | 12 |
| 921 | 0.183 | 0.208 | 0.222 | -0.349 | 0.313 | 12 | 10 | 8 | 7 |
| 922 | 1.816 | 2.484 | 1.966 | 0.896 | 1.496 | 27 | 9 | 6 | 10 |
| 923 | 1.934 | 2.392 | 1.956 | 1.671 | 1.765 | 9 | 8 | 8 | 16 |
| 924 | -0.494 | -0.211 | -0.608 | -0.619 | -0.540 | 13 | 7 | 9 | 8 |
| 925 | 0.512 | 0.105 | 0.471 | 0.938 | 0.910 | 16 | 7 | 6 | 11 |
| 926 | 0.239 | 0.485 | 0.621 | 0.648 | -0.011 | 9 | 6 | 5 | 10 |
| 927 | -1.496 | -1.172 | -0.927 | -0.346 | -1.687 | 11 | 5 | 6 | 8 |
| 928 | -0.373 | -0.389 | -0.369 | -0.716 | 0.781 | 9 | 6 | 5 | 8 |
| 929 | -0.964 | -1.267 | -0.821 | -0.840 | -0.588 | 7 | 7 | 7 | 9 |
| 930 | 0.638 | 0.791 | 0.734 | 0.372 | 0.340 | 7 | 6 | 9 | 7 |
| 931 | -0.008 | 0.045 | -0.262 | -0.150 | 0.186 | 14 | 9 | 13 | 16 |
| 932 | 0.161 | 0.020 | -0.031 | -0.023 | 0.248 | 11 | 10 | 6 | 5 |
| 933 | 0.248 | 0.058 | 0.512 | 0.366 | 0.949 | 15 | 6 | 7 | 8 |
| 934 | 0.860 | 0.670 | 0.732 | 0.759 | 1.219 | 9 | 5 | 5 | 7 |
| 935 | -0.727 | -0.644 | -0.257 | -0.879 | -0.971 | 5 | 9 | 11 | 7 |
| 936 | -1.132 | -1.311 | -1.195 | -1.098 | -0.951 | 15 | 6 | 9 | 9 |
| 937 | -0.137 | 0.100 | 0.266 | -0.012 | -0.030 | 10 | 8 | 7 | 11 |
| 938 | -1.157 | -1.101 | -0.778 | -1.040 | -1.062 | 8 | 6 | 7 | 7 |
| 939 | -0.329 | -0.347 | -0.688 | -0.248 | -0.097 | 9 | 6 | 5 | 8 |
| 940 | 0.471 | -0.062 | 0.190 | 0.563 | 0.345 | 7 | 8 | 7 | 11 |
| 941 | 1.916 | 1.864 | 1.597 | 1.637 | 2.086 | 20 | 7 | 12 | 13 |
| 942 | 0.931 | 1.119 | 1.342 | 0.888 | 0.358 | 10 | 7 | 7 | 14 |
| 943 | -1.436 | 0.814 | -1.449 | -1.199 | -1.486 | 6 | 7 | 7 | 10 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า $\theta$ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 944 | 0.218 | -0.207 | 0.292 | -0.160 | 0.741 | 14 | 9 | 7 | 7 |
| 945 | 0.192 | 0.299 | 0.460 | -0.173 | -0.213 | 11 | 6 | 5 | 9 |
| 946 | 0.306 | 0.716 | 0.654 | 0.116 | 0.134 | 7 | 6 | 9 | 8 |
| 947 | 0.619 | 0.828 | 0.506 | 0.430 | 0.453 | 10 | 5 | 6 | 8 |
| 948 | 0.575 | 0.583 | 0.593 | 0.638 | 0.459 | 10 | 7 | 6 | 8 |
| 949 | -1.082 | -0.594 | -1.195 | -0.919 | -1.295 | 10 | 5 | 6 | 11 |
| 950 | -0.485 | -0.369 | -0.631 | -0.722 | -0.707 | 13 | 5 | 6 | 16 |
| 951 | 0.101 | 0.128 | 0.267 | 0.283 | 0.053 | 15 | 10 | 7 | 7 |
| 952 | 0.823 | 0.772 | 0.318 | 0.641 | 1.060 | 7 | 11 | 6 | 12 |
| 953 | -0.518 | -0.783 | -0.652 | -0.434 | -0.461 | 6 | 6 | 7 | 10 |
| 954 | -1.573 | -0.938 | -1.484 | -1.453 | -1.605 | 9 | 6 | 13 | 13 |
| 955 | 1.364 | 2.404 | 1.325 | 0.999 | 1.307 | 21 | 7 | 6 | 13 |
| 956 | -1.358 | -1.293 | -1.157 | -1.318 | -1.409 | 16 | 7 | 7 | 8 |
| 957 | 0.769 | 1.454 | 0.794 | 1.645 | 1.204 | 12 | 5 | 7 | 18 |
| 958 | 0.569 | 0.339 | 0.652 | 0.143 | 0.463 | 10 | 7 | 8 | 12 |
| 959 | -0.549 | -0.126 | -0.364 | -0.141 | -0.124 | 16 | 8 | 5 | 11 |
| 960 | 0.566 | 0.672 | 0.279 | 0.147 | 0.174 | 15 | 8 | 7 | 16 |
| 961 | 0.604 | -0.154 | 0.691 | 0.108 | 0.693 | 12 | 5 | 5 | 9 |
| 962 | 0.983 | 0.967 | 0.633 | 0.927 | 1.012 | 12 | 6 | 6 | 8 |
| 963 | -1.410 | -1.287 | -1.011 | -1.416 | -1.165 | 8 | 4 | 7 | 9 |
| 964 | 0.681 | 0.248 | 0.849 | 0.601 | 0.740 | 13 | 5 | 6 | 6 |
| 965 | 0.266 | -0.042 | 0.045 | 0.105 | 0.322 | 13 | 10 | 9 | 9 |
| 966 | 0.489 | 0.053 | 0.243 | 0.432 | 0.172 | 17 | 12 | 5 | 9 |
| 967 | -0.947 | -1.109 | -1.096 | -0.752 | -0.818 | 9 | 5 | 8 | 15 |
| 968 | -0.037 | 0.054 | -0.236 | -0.145 | -0.378 | 12 | 10 | 5 | 13 |
| 969 | -0.113 | 0.098 | 0.046 | -0.178 | -0.005 | 10 | 8 | 9 | 11 |
| 970 | -0.389 | -0.801 | -0.257 | -0.339 | -0.180 | 10 | 9 | 6 | 8 |
| 971 | 1.300 | 0.740 | 1.094 | 1.685 | 0.546 | 10 | 8 | 7 | 14 |
| 972 | 0.735 | 0.748 | 0.754 | 1.070 | 0.176 | 7 | 5 | 13 | 9 |
| 973 | -0.017 | 0.846 | -0.197 | -0.128 | 0.001 | 12 | 9 | 6 | 7 |
| 974 | -0.608 | -0.196 | -0.770 | -0.892 | -0.773 | 13 | 6 | 7 | 8 |
| 975 | 0.553 | 1.032 | 0.652 | 0.253 | 0.347 | 12 | 7 | 6 | 21 |
| 976 | -0.962 | -1.272 | -0.524 | -0.588 | -1.051 | 7 | 11 | 8 | 11 |
| 977 | 0.252 | 0.037 | 0.273 | 0.197 | 0.375 | 14 | 7 | 7 | 10 |
| 978 | 1.263 | 1.224 | 1.347 | 1.739 | 1.238 | 10 | 7 | 10 | 16 |
| 979 | -1.938 | -1.332 | -1.230 | -2.021 | -1.562 | 10 | 7 | 11 | 10 |
| 980 | 0.501 | 0.109 | 0.445 | 0.513 | 0.016 | 15 | 7 | 6 | 6 |
| 981 | 0.302 | -0.352 | 0.032 | 0.639 | 0.553 | 10 | 9 | 5 | 8 |
| 982 | 0.967 | 1.148 | 1.094 | 0.973 | 1.326 | 11 | 8 | 6 | 13 |
| 983 | -1.337 | -1.776 | -0.829 | -1.403 | -0.951 | 11 | 6 | 10 | 9 |
| 984 | 0.672 | 0.731 | 0.235 | 0.729 | 0.433 | 9 | 8 | 6 | 9 |
| 985 | -0.712 | -1.146 | -0.429 | -0.831 | -0.191 | 8 | 5 | 6 | 10 |
| 986 | -1.252 | -1.210 | -1.147 | -1.200 | -1.330 | 11 | 9 | 7 | 13 |

ตารางที่ ข-1 (ต่อ)

| ที่ | ค่า θ | ค่าประมาณ $\hat{\theta}$ | | | | ความยาวของแบบทดสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 987 | -1.459 | -1.009 | -1.544 | -1.493 | -1.439 | 9 | 6 | 10 | 8 |
| 988 | 0.647 | 0.995 | 0.460 | 0.583 | -0.057 | 11 | 8 | 6 | 12 |
| 989 | -0.551 | -0.354 | -0.572 | -0.769 | -0.712 | 16 | 8 | 4 | 19 |
| 990 | -1.455 | -1.490 | -1.547 | -1.416 | -1.116 | 16 | 6 | 10 | 10 |
| 991 | -1.108 | -1.299 | -0.547 | -0.693 | -0.909 | 14 | 8 | 12 | 6 |
| 992 | 0.077 | -0.447 | -0.029 | 0.213 | 0.228 | 8 | 10 | 4 | 8 |
| 993 | 1.030 | 1.534 | 0.881 | 0.768 | 0.892 | 13 | 5 | 6 | 16 |
| 994 | -0.725 | -1.160 | -0.369 | -0.623 | -1.023 | 10 | 8 | 9 | 13 |
| 995 | -0.138 | -0.253 | -0.227 | 0.745 | -0.187 | 11 | 8 | 6 | 12 |
| 996 | 0.263 | 1.234 | 0.249 | 0.228 | 0.468 | 17 | 7 | 6 | 16 |
| 997 | 0.158 | 0.219 | -0.376 | 0.369 | -0.381 | 14 | 7 | 7 | 7 |
| 998 | -0.136 | 0.639 | -0.369 | -0.346 | -0.425 | 11 | 6 | 6 | 9 |
| 999 | -0.744 | -1.140 | -0.952 | -0.672 | -0.593 | 15 | 6 | 6 | 8 |
| 1000 | 0.461 | -0.265 | 0.332 | 0.486 | 0.313 | 10 | 8 | 6 | 8 |
| *Mean* | 0.007 | 0.062 | 0.000 | 0.006 | 0.009 | 11.43 | 7.27 | 7.56 | 9.95 |
| *SD* | 1.004 | 1.073 | 0.944 | 0.932 | 0.925 | 3.20 | 1.61 | 2.87 | 2.86 |
| *RMSE* | | 0.583 | 0.312 | 0.330 | 0.348 | | | | |
| *Average Bias* | | 0.055 | -0.007 | -0.002 | 0.001 | | | | |

จากตารางที่ ข-1 แสดงค่าความสามารถที่แท้จริงของผู้สอบ (θ) ชุดที่ 1 ซึ่งใช้เป็นกลุ่มตัวอย่างของการศึกษาในสถานการณ์จำลอง การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ดำเนินการศึกษาซ้ำทั้งหมด 10 รอบ ทุกรอบใช้คลังข้อสอบจำลองขนาด 500 ข้อ ชุดเดียวกัน แต่ใช้ค่าความสามารถที่แท้จริงของผู้สอบทีละชุด ชุดละ 1,000 ค่า รวมทั้งหมด 10,000 ค่า ผลที่ได้รับจากการศึกษาแต่ละครั้ง จะได้ค่าประมาณความสามารถของผู้สอบ ($\hat{\theta}$) และความยาวของแบบทดสอบของค่า θ แต่ละค่า ดังนั้น ในแต่ละวิธีการคัดเลือกข้อสอบข้อถัดไป จะได้ค่าประมาณความสามารถของผู้สอบ และความยาวของแบบทดสอบ อย่างละ 10,000 ค่า เนื่องด้วยข้อจำกัดในการแสดงผล จึงขอแสดงผลเฉพาะชุดที่ 1 เท่านั้น (ดังตารางที่ ข-1) ค่าประมาณความสามารถของผู้สอบนี้ นำไปคำนวณหาค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความลำเอียงเฉลี่ย เพื่อเปรียบเทียบประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบ (แสดงตัวอย่างการคำนวณในหัวข้อถัดไป) ส่วนความยาวของแบบทดสอบ จะนำไปใช้เปรียบเทียบประสิทธิภาพในด้านความยาวของแบบทดสอบ

**การคำนวณค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย**

การคำนวณค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) ดำเนินการโดยนำค่าความสามารถที่แท้จริงของผู้สอบ (θ) และค่าประมาณความสามารถของผู้สอบ ($\hat{\theta}$) ที่ได้รับจากการศึกษาในสถานการณ์จำลอง (แสดงในตารางที่ ข-1) มาแทนค่าในสมการที่ 32

แสดงตัวอย่างการคำนวณค่า RMSE ของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (MIC) ได้ดังนี้

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_i - \theta_i\right)^2}$$

เมื่อ n คือ จำนวนผู้สอบทั้งหมด
$\hat{\theta}_i$ คือ ค่าประมาณความสามารถของผู้สอบคนที่ i
$\theta_i$ คือ ค่าความสามารถที่แท้จริงของผู้สอบคนที่ i

แทนค่า

$$\text{RMSE} = \sqrt{\frac{1}{1{,}000}\sum_{i=1}^{1{,}000}\left(\hat{\theta}_i - \theta_i\right)^2}$$

$$= \sqrt{\frac{1}{1{,}000}\left[\left(\hat{\theta}_1\text{-}\theta_1\right)^2+\left(\hat{\theta}_2\text{-}\theta_2\right)^2+(\hat{\theta}_3\text{-}\theta_3)^2+\cdots+\left(\hat{\theta}_{1{,}000}\text{-}\theta_{1{,}000}\right)^2\right]}$$

$$= \sqrt{\frac{1}{1{,}000}[(0.609\text{-}0.552)^2+(0.558\text{-}0.151)^2+((\text{-}0.314)\text{-}(\text{-}0.280))^2+\cdots+((\text{-}0.265)\text{-}0.461)^2]}$$

$$= 0.583$$

**การคำนวณค่าความลำเอียงเฉลี่ย**

การคำนวณค่าความลำเอียงเฉลี่ย (Average Bias) ดำเนินการโดยนำค่าความสามารถที่แท้จริงของผู้สอบ ($\theta$) และค่าประมาณความสามารถของผู้สอบ ($\hat{\theta}$) ที่ได้รับจากการศึกษาในสถานการณ์จำลอง (แสดงในตารางที่ ข-1) มาแทนค่าในสมการที่ 33

แสดงตัวอย่างการคำนวณค่าความลำเอียงเฉลี่ยของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (MIC) ได้ดังนี้

$$\text{Average Bias} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_i - \theta_i\right)$$

เมื่อ n คือ จำนวนผู้สอบทั้งหมด
$\hat{\theta}_i$ คือ ค่าประมาณความสามารถของผู้สอบคนที่ i
$\theta_i$ คือ ค่าความสามารถที่แท้จริงของผู้สอบคนที่ i

แทนค่า

$$\text{Average Bias} = \frac{1}{1{,}000}\sum_{i=1}^{1{,}000}\left(\hat{\theta}_i - \theta_i\right)$$

$$= \frac{1}{1{,}000}\left[\left(\hat{\theta}_1\text{-}\theta_1\right)+\left(\hat{\theta}_2\text{-}\theta_2\right)+\left(\hat{\theta}_3\text{-}\theta_3\right) + \cdots + \left(\hat{\theta}_{1{,}000}\text{-}\theta_{1{,}000}\right)\right]$$

$$= \frac{1}{1{,}000}\left[(0.609\text{-}0.552) + (0.558\text{-}0.151) + ((\text{-}0.314)\text{-}(\text{-}0.280)) + \cdots + ((\text{-}0.265)\text{-}0.461)\right]$$

$$= 0.055$$

ส่วนการคำนวณค่า RMSE และค่าความลำเอียงเฉลี่ยของวิธีการคัดเลือกข้อสอบข้อถัดไป วิธีอื่น ๆ สามารถคำนวณได้ในทำนองเดียวกัน

# ภาคผนวก ค
# คลังข้อสอบจำลองของการศึกษาในสถานการณ์จำลอง การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ค1. ค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบจำลอง จำนวนการใช้ข้อสอบ และอัตราการใช้ข้อสอบ จากการศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ค2. การคำนวณอัตราการใช้ข้อสอบ

**คลังข้อสอบจำลองที่ใช้ในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

ตารางที่ ค-1 ค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบจำลอง จำนวนการใช้ข้อสอบ และอัตราการใช้ข้อสอบ จากการศึกษาในสถานการณ์จำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 1 | 1.07 | 1.336 | 0.088 | 21 | 0 | 0 | 3 | 0.021 | 0 | 0 | 0.003 |
| 2 | 0.514 | 2.075 | 0.102 | 12 | 0 | 0 | 0 | 0.012 | 0 | 0 | 0 |
| 3 | 0.901 | 1.754 | 0.085 | 13 | 0 | 1 | 0 | 0.013 | 0 | 0.001 | 0 |
| 4 | 1.701 | 1.354 | 0.062 | 15 | 0 | 0 | 7 | 0.015 | 0 | 0 | 0.007 |
| 5 | 1.654 | -2.201 | 0.142 | 2 | 0 | 0 | 1 | 0.002 | 0 | 0 | 0.001 |
| 6 | 2.146 | 1.565 | 0.245 | 20 | 0 | 0 | 3 | 0.02 | 0 | 0 | 0.003 |
| 7 | 1.448 | -1.487 | 0.109 | 21 | 0 | 0 | 4 | 0.021 | 0 | 0 | 0.004 |
| 8 | 2.397 | 0.318 | 0.055 | 67 | 340 | 405 | 152 | 0.067 | 0.34 | 0.405 | 0.152 |
| 9 | 0.704 | -1.858 | 0.257 | 4 | 1 | 0 | 0 | 0.004 | 0.001 | 0 | 0 |
| 10 | 1.221 | 1.031 | 0.141 | 21 | 3 | 0 | 0 | 0.021 | 0.003 | 0 | 0 |
| 11 | 0.911 | 2.073 | 0.288 | 3 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 |
| 12 | 1.592 | -0.438 | 0.075 | 43 | 17 | 4 | 46 | 0.043 | 0.017 | 0.004 | 0.046 |
| 13 | 0.839 | 1.08 | 0.083 | 8 | 0 | 0 | 0 | 0.008 | 0 | 0 | 0 |
| 14 | 0.583 | 1.566 | 0.171 | 4 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 |
| 15 | 2.259 | 1.514 | 0.242 | 1 | 0 | 4 | 14 | 0.001 | 0 | 0.004 | 0.014 |
| 16 | 0.99 | -2.398 | 0.052 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 17 | 1.73 | 0.371 | 0.256 | 27 | 12 | 7 | 21 | 0.027 | 0.012 | 0.007 | 0.021 |
| 18 | 0.993 | -0.716 | 0.164 | 41 | 5 | 6 | 4 | 0.041 | 0.005 | 0.006 | 0.004 |
| 19 | 1.526 | 0.371 | 0.017 | 21 | 15 | 10 | 65 | 0.021 | 0.015 | 0.01 | 0.065 |
| 20 | 0.666 | 0.987 | 0.219 | 24 | 3 | 8 | 6 | 0.024 | 0.003 | 0.008 | 0.006 |
| 21 | 1.251 | 1.46 | 0.248 | 7 | 0 | 0 | 0 | 0.007 | 0 | 0 | 0 |
| 22 | 0.659 | -2.094 | 0.3 | 8 | 1 | 0 | 0 | 0.008 | 0.001 | 0 | 0 |
| 23 | 2.216 | 1.405 | 0.222 | 12 | 0 | 0 | 14 | 0.012 | 0 | 0 | 0.014 |
| 24 | 0.845 | 2.499 | 0.257 | 15 | 0 | 0 | 0 | 0.015 | 0 | 0 | 0 |
| 25 | 1.138 | -0.635 | 0.19 | 43 | 7 | 6 | 3 | 0.043 | 0.007 | 0.006 | 0.003 |
| 26 | 1.268 | -2.419 | 0.294 | 4 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 |
| 27 | 2.212 | -2.385 | 0.141 | 1 | 0 | 0 | 3 | 0.001 | 0 | 0 | 0.003 |
| 28 | 2.184 | 0.612 | 0.009 | 25 | 396 | 357 | 145 | 0.025 | 0.396 | 0.357 | 0.145 |
| 29 | 1.144 | -2.181 | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 30 | 0.517 | -0.035 | 0.148 | 28 | 3 | 5 | 2 | 0.028 | 0.003 | 0.005 | 0.002 |
| 31 | 1.178 | -0.493 | 0.116 | 41 | 9 | 6 | 1 | 0.041 | 0.009 | 0.006 | 0.001 |
| 32 | 0.55 | 0.187 | 0.275 | 37 | 4 | 6 | 4 | 0.037 | 0.004 | 0.006 | 0.004 |
| 33 | 1.698 | -0.335 | 0.205 | 53 | 12 | 6 | 29 | 0.053 | 0.012 | 0.006 | 0.029 |
| 34 | 1.143 | -0.446 | 0.08 | 83 | 2 | 3 | 2 | 0.083 | 0.002 | 0.003 | 0.002 |
| 35 | 2.467 | -1.396 | 0.109 | 11 | 6 | 65 | 61 | 0.011 | 0.006 | 0.065 | 0.061 |
| 36 | 2.308 | -0.565 | 0.074 | 29 | 314 | 154 | 146 | 0.029 | 0.314 | 0.154 | 0.146 |
| 37 | 1.056 | -1.51 | 0.295 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 38 | 2.455 | 1.259 | 0.23 | 37 | 3 | 94 | 39 | 0.037 | 0.003 | 0.094 | 0.039 |
| 39 | 1.531 | 2.219 | 0.142 | 11 | 0 | 0 | 0 | 0.011 | 0 | 0 | 0 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 40 | 2.097 | -1.89 | 0.293 | 7 | 0 | 0 | 5 | 0.007 | 0 | 0 | 0.005 |
| 41 | 1.178 | 1.52 | 0.04 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 42 | 0.963 | -0.768 | 0.142 | 26 | 17 | 29 | 118 | 0.026 | 0.017 | 0.029 | 0.118 |
| 43 | 2.366 | -0.537 | 0.2 | 122 | 4 | 5 | 2 | 0.122 | 0.004 | 0.005 | 0.002 |
| 44 | 0.82 | 1.971 | 0.2 | 9 | 0 | 0 | 0 | 0.009 | 0 | 0 | 0 |
| 45 | 1.112 | -1.241 | 0.141 | 26 | 0 | 0 | 0 | 0.026 | 0 | 0 | 0 |
| 46 | 0.803 | 1.947 | 0.112 | 1 | 0 | 0 | 1 | 0.001 | 0 | 0 | 0.001 |
| 47 | 1.926 | -1.482 | 0.225 | 4 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 |
| 48 | 1.194 | -1.523 | 0.036 | 1 | 0 | 0 | 18 | 0.001 | 0 | 0 | 0.018 |
| 49 | 1.886 | -1.621 | 0.074 | 1 | 6 | 18 | 48 | 0.001 | 0.006 | 0.018 | 0.048 |
| 50 | 2.264 | -1.471 | 0.225 | 6 | 0 | 12 | 37 | 0.006 | 0 | 0.012 | 0.037 |
| 51 | 2.347 | 0.146 | 0.171 | 44 | 6 | 4 | 38 | 0.044 | 0.006 | 0.004 | 0.038 |
| 52 | 1.828 | -1.251 | 0.086 | 24 | 0 | 0 | 11 | 0.024 | 0 | 0 | 0.011 |
| 53 | 1.572 | -1.637 | 0.138 | 9 | 2 | 11 | 24 | 0.009 | 0.002 | 0.011 | 0.024 |
| 54 | 2.485 | 1.329 | 0.056 | 18 | 1 | 6 | 10 | 0.018 | 0.001 | 0.006 | 0.01 |
| 55 | 0.541 | 2.46 | 0.144 | 15 | 0 | 0 | 0 | 0.015 | 0 | 0 | 0 |
| 56 | 0.618 | 0.764 | 0.116 | 12 | 18 | 6 | 28 | 0.012 | 0.018 | 0.006 | 0.028 |
| 57 | 1.723 | -0.606 | 0.28 | 10 | 8 | 3 | 10 | 0.01 | 0.008 | 0.003 | 0.01 |
| 58 | 1.717 | -0.246 | 0.248 | 117 | 5 | 7 | 3 | 0.117 | 0.005 | 0.007 | 0.003 |
| 59 | 1.148 | 1.935 | 0.011 | 10 | 0 | 0 | 0 | 0.01 | 0 | 0 | 0 |
| 60 | 1.084 | -1.186 | 0.094 | 12 | 0 | 0 | 0 | 0.012 | 0 | 0 | 0 |
| 61 | 0.573 | -1.005 | 0.189 | 55 | 0 | 0 | 1 | 0.055 | 0 | 0 | 0.001 |
| 62 | 1.48 | 0.345 | 0.272 | 40 | 7 | 2 | 9 | 0.04 | 0.007 | 0.002 | 0.009 |
| 63 | 1.067 | 0.677 | 0.284 | 42 | 12 | 3 | 17 | 0.042 | 0.012 | 0.003 | 0.017 |
| 64 | 1.783 | -0.79 | 0.079 | 16 | 12 | 3 | 55 | 0.016 | 0.012 | 0.003 | 0.055 |
| 65 | 1.819 | 1.922 | 0.073 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 66 | 1.217 | -1.896 | 0.152 | 3 | 0 | 5 | 23 | 0.003 | 0 | 0.005 | 0.023 |
| 67 | 2.488 | -0.062 | 0.086 | 58 | 7 | 185 | 134 | 0.058 | 0.007 | 0.185 | 0.134 |
| 68 | 2.21 | 2.229 | 0.192 | 9 | 0 | 0 | 2 | 0.009 | 0 | 0 | 0.002 |
| 69 | 1.903 | 1.801 | 0.008 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0.002 |
| 70 | 1.451 | -0.926 | 0.171 | 15 | 13 | 34 | 95 | 0.015 | 0.013 | 0.034 | 0.095 |
| 71 | 2.426 | -0.119 | 0.283 | 133 | 8 | 6 | 6 | 0.133 | 0.008 | 0.006 | 0.006 |
| 72 | 1.204 | 0.323 | 0.057 | 46 | 17 | 8 | 90 | 0.046 | 0.017 | 0.008 | 0.09 |
| 73 | 1.792 | -0.979 | 0.26 | 8 | 7 | 5 | 6 | 0.008 | 0.007 | 0.005 | 0.006 |
| 74 | 1.441 | -0.833 | 0.052 | 18 | 8 | 4 | 11 | 0.018 | 0.008 | 0.004 | 0.011 |
| 75 | 0.745 | 0.595 | 0.095 | 130 | 11 | 6 | 69 | 0.13 | 0.011 | 0.006 | 0.069 |
| 76 | 1.858 | -2.431 | 0.114 | 1 | 0 | 0 | 1 | 0.001 | 0 | 0 | 0.001 |
| 77 | 1.26 | -0.123 | 0.182 | 21 | 5 | 9 | 19 | 0.021 | 0.005 | 0.009 | 0.019 |
| 78 | 1.339 | -0.761 | 0.01 | 22 | 8 | 5 | 3 | 0.022 | 0.008 | 0.005 | 0.003 |
| 79 | 0.697 | -0.638 | 0.084 | 37 | 6 | 5 | 7 | 0.037 | 0.006 | 0.005 | 0.007 |
| 80 | 0.923 | 0.374 | 0.019 | 46 | 15 | 7 | 67 | 0.046 | 0.015 | 0.007 | 0.067 |
| 81 | 1.668 | 1.954 | 0.19 | 8 | 0 | 0 | 0 | 0.008 | 0 | 0 | 0 |
| 82 | 0.846 | 1.423 | 0.086 | 6 | 2 | 0 | 14 | 0.006 | 0.002 | 0 | 0.014 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 83 | 0.714 | -0.829 | 0.114 | 25 | 3 | 6 | 7 | 0.025 | 0.003 | 0.006 | 0.007 |
| 84 | 1.246 | -0.573 | 0.103 | 28 | 5 | 5 | 8 | 0.028 | 0.005 | 0.005 | 0.008 |
| 85 | 2.368 | 0.775 | 0.211 | 17 | 13 | 71 | 66 | 0.017 | 0.013 | 0.071 | 0.066 |
| 86 | 2.117 | 1.701 | 0.147 | 1 | 0 | 0 | 7 | 0.001 | 0 | 0 | 0.007 |
| 87 | 1.516 | -1.513 | 0.086 | 4 | 0 | 0 | 5 | 0.004 | 0 | 0 | 0.005 |
| 88 | 0.876 | -2.099 | 0.219 | 4 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 |
| 89 | 0.672 | -2.313 | 0.005 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 90 | 0.791 | -1.26 | 0.267 | 34 | 0 | 0 | 0 | 0.034 | 0 | 0 | 0 |
| 91 | 2.492 | 2.261 | 0.149 | 7 | 15 | 10 | 0 | 0.007 | 0.015 | 0.01 | 0 |
| 92 | 2.37 | 0.91 | 0.16 | 17 | 8 | 67 | 52 | 0.017 | 0.008 | 0.067 | 0.052 |
| 93 | 1.043 | -0.999 | 0.165 | 9 | 3 | 2 | 3 | 0.009 | 0.003 | 0.002 | 0.003 |
| 94 | 2.075 | -1.037 | 0.284 | 14 | 1 | 0 | 19 | 0.014 | 0.001 | 0 | 0.019 |
| 95 | 1.089 | -0.798 | 0.09 | 30 | 2 | 8 | 7 | 0.03 | 0.002 | 0.008 | 0.007 |
| 96 | 1.254 | 1.207 | 0.148 | 44 | 2 | 0 | 0 | 0.044 | 0.002 | 0 | 0 |
| 97 | 1.94 | 1.387 | 0.188 | 11 | 0 | 0 | 5 | 0.011 | 0 | 0 | 0.005 |
| 98 | 1.614 | -2.12 | 0.029 | 1 | 0 | 0 | 2 | 0.001 | 0 | 0 | 0.002 |
| 99 | 0.941 | 0.287 | 0.288 | 43 | 5 | 4 | 5 | 0.043 | 0.005 | 0.004 | 0.005 |
| 100 | 2.047 | 1.142 | 0.202 | 20 | 0 | 0 | 13 | 0.02 | 0 | 0 | 0.013 |
| 101 | 2.151 | -1.49 | 0.228 | 3 | 0 | 1 | 23 | 0.003 | 0 | 0.001 | 0.023 |
| 102 | 2.122 | -2.333 | 0.227 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0.005 |
| 103 | 2.28 | -1.469 | 0.055 | 19 | 0 | 18 | 51 | 0.019 | 0 | 0.018 | 0.051 |
| 104 | 2.045 | 0.862 | 0.182 | 51 | 9 | 6 | 41 | 0.051 | 0.009 | 0.006 | 0.041 |
| 105 | 2.451 | -1.074 | 0.214 | 19 | 3 | 29 | 62 | 0.019 | 0.003 | 0.029 | 0.062 |
| 106 | 1.419 | 0.993 | 0.181 | 41 | 2 | 4 | 6 | 0.041 | 0.002 | 0.004 | 0.006 |
| 107 | 0.838 | -0.001 | 0.245 | 15 | 5 | 4 | 4 | 0.015 | 0.005 | 0.004 | 0.004 |
| 108 | 1.598 | -2.237 | 0.204 | 2 | 0 | 0 | 2 | 0.002 | 0 | 0 | 0.002 |
| 109 | 1.758 | 0.215 | 0.254 | 174 | 8 | 5 | 24 | 0.174 | 0.008 | 0.005 | 0.024 |
| 110 | 0.959 | 1.098 | 0.026 | 4 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 |
| 111 | 1.258 | 1.076 | 0.242 | 23 | 1 | 0 | 1 | 0.023 | 0.001 | 0 | 0.001 |
| 112 | 0.897 | 2.199 | 0.234 | 10 | 0 | 0 | 0 | 0.01 | 0 | 0 | 0 |
| 113 | 1.788 | -0.621 | 0.203 | 13 | 12 | 7 | 23 | 0.013 | 0.012 | 0.007 | 0.023 |
| 114 | 0.863 | -1.879 | 0.056 | 3 | 1 | 0 | 0 | 0.003 | 0.001 | 0 | 0 |
| 115 | 0.994 | 1.977 | 0.117 | 4 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 |
| 116 | 1.251 | -0.573 | 0.053 | 27 | 7 | 10 | 21 | 0.027 | 0.007 | 0.01 | 0.021 |
| 117 | 2.213 | -0.522 | 0.214 | 65 | 19 | 4 | 81 | 0.065 | 0.019 | 0.004 | 0.081 |
| 118 | 1.449 | 1.679 | 0.132 | 7 | 0 | 0 | 0 | 0.007 | 0 | 0 | 0 |
| 119 | 0.723 | 1.517 | 0.168 | 6 | 2 | 0 | 0 | 0.006 | 0.002 | 0 | 0 |
| 120 | 0.868 | -2.182 | 0.201 | 7 | 0 | 0 | 0 | 0.007 | 0 | 0 | 0 |
| 121 | 2.315 | -2.477 | 0.163 | 3 | 0 | 0 | 2 | 0.003 | 0 | 0 | 0.002 |
| 122 | 1.261 | -0.304 | 0.059 | 61 | 6 | 6 | 49 | 0.061 | 0.006 | 0.006 | 0.049 |
| 123 | 1.248 | -1.874 | 0.229 | 5 | 1 | 0 | 0 | 0.005 | 0.001 | 0 | 0 |
| 124 | 1.139 | -1.696 | 0.294 | 11 | 0 | 0 | 0 | 0.011 | 0 | 0 | 0 |
| 125 | 1.146 | -1.383 | 0.044 | 10 | 1 | 0 | 4 | 0.01 | 0.001 | 0 | 0.004 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 126 | 1.606 | -1.524 | 0.043 | 1 | 2 | 0 | 5 | 0.001 | 0.002 | 0 | 0.005 |
| 127 | 0.906 | 2.113 | 0.255 | 9 | 0 | 0 | 0 | 0.009 | 0 | 0 | 0 |
| 128 | 1.418 | 0.086 | 0.117 | 104 | 10 | 5 | 15 | 0.104 | 0.01 | 0.005 | 0.015 |
| 129 | 2.179 | 1.341 | 0.147 | 16 | 43 | 2 | 25 | 0.016 | 0.043 | 0.002 | 0.025 |
| 130 | 1.581 | -0.96 | 0.019 | 14 | 16 | 3 | 25 | 0.014 | 0.016 | 0.003 | 0.025 |
| 131 | 1.473 | -0.818 | 0.086 | 17 | 27 | 3 | 15 | 0.017 | 0.027 | 0.003 | 0.015 |
| 132 | 2.303 | -2.419 | 0.188 | 5 | 0 | 0 | 4 | 0.005 | 0 | 0 | 0.004 |
| 133 | 2.468 | -0.599 | 0.1 | 21 | 470 | 347 | 124 | 0.021 | 0.47 | 0.347 | 0.124 |
| 134 | 1.605 | -0.306 | 0.256 | 68 | 7 | 3 | 10 | 0.068 | 0.007 | 0.003 | 0.01 |
| 135 | 0.571 | 0.2 | 0.179 | 54 | 6 | 3 | 6 | 0.054 | 0.006 | 0.003 | 0.006 |
| 136 | 1.417 | -1.367 | 0.004 | 8 | 0 | 0 | 8 | 0.008 | 0 | 0 | 0.008 |
| 137 | 2.305 | -0.485 | 0.1 | 63 | 257 | 202 | 117 | 0.063 | 0.257 | 0.202 | 0.117 |
| 138 | 1.451 | -0.133 | 0.238 | 12 | 3 | 6 | 6 | 0.012 | 0.003 | 0.006 | 0.006 |
| 139 | 0.573 | 2.208 | 0.022 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 140 | 1.689 | -1.04 | 0.254 | 30 | 5 | 0 | 6 | 0.03 | 0.005 | 0 | 0.006 |
| 141 | 1.372 | 0.554 | 0.104 | 44 | 19 | 5 | 16 | 0.044 | 0.019 | 0.005 | 0.016 |
| 142 | 0.775 | 0.18 | 0.017 | 23 | 5 | 4 | 1 | 0.023 | 0.005 | 0.004 | 0.001 |
| 143 | 2.011 | 0.692 | 0.22 | 37 | 4 | 6 | 51 | 0.037 | 0.004 | 0.006 | 0.051 |
| 144 | 2.137 | -2.143 | 0.016 | 1 | 10 | 6 | 6 | 0.001 | 0.01 | 0.006 | 0.006 |
| 145 | 1.402 | -1.067 | 0.008 | 28 | 10 | 0 | 15 | 0.028 | 0.01 | 0 | 0.015 |
| 146 | 2.151 | -1.219 | 0.144 | 47 | 0 | 2 | 36 | 0.047 | 0 | 0.002 | 0.036 |
| 147 | 1.481 | -1.906 | 0.153 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 148 | 1.149 | 0.156 | 0.02 | 51 | 3 | 6 | 59 | 0.051 | 0.003 | 0.006 | 0.059 |
| 149 | 1.331 | 0.252 | 0.193 | 8 | 1 | 3 | 9 | 0.008 | 0.001 | 0.003 | 0.009 |
| 150 | 0.6 | -1.081 | 0.17 | 3 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 |
| 151 | 2.051 | -1.418 | 0.214 | 12 | 0 | 0 | 11 | 0.012 | 0 | 0 | 0.011 |
| 152 | 2.261 | -0.579 | 0.295 | 12 | 17 | 6 | 57 | 0.012 | 0.017 | 0.006 | 0.057 |
| 153 | 0.959 | 2.185 | 0.05 | 13 | 0 | 0 | 0 | 0.013 | 0 | 0 | 0 |
| 154 | 1.62 | -2.224 | 0.234 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0.001 |
| 155 | 2.165 | -0.939 | 0.221 | 29 | 10 | 3 | 45 | 0.029 | 0.01 | 0.003 | 0.045 |
| 156 | 2.442 | 2.404 | 0.28 | 12 | 3 | 3 | 0 | 0.012 | 0.003 | 0.003 | 0 |
| 157 | 1.1 | -1.376 | 0.175 | 31 | 0 | 0 | 0 | 0.031 | 0 | 0 | 0 |
| 158 | 1.722 | 1.54 | 0.007 | 19 | 3 | 23 | 15 | 0.019 | 0.003 | 0.023 | 0.015 |
| 159 | 1.049 | 1.441 | 0.001 | 14 | 0 | 2 | 14 | 0.014 | 0 | 0.002 | 0.014 |
| 160 | 0.874 | -0.248 | 0.22 | 64 | 5 | 4 | 5 | 0.064 | 0.005 | 0.004 | 0.005 |
| 161 | 2.246 | 0.066 | 0.147 | 82 | 5 | 42 | 91 | 0.082 | 0.005 | 0.042 | 0.091 |
| 162 | 1.454 | 1.738 | 0.132 | 4 | 0 | 0 | 2 | 0.004 | 0 | 0 | 0.002 |
| 163 | 0.62 | 1.155 | 0.052 | 10 | 0 | 7 | 8 | 0.01 | 0 | 0.007 | 0.008 |
| 164 | 2.233 | -2.048 | 0.234 | 9 | 0 | 0 | 8 | 0.009 | 0 | 0 | 0.008 |
| 165 | 1.716 | 0.254 | 0.273 | 13 | 5 | 9 | 18 | 0.013 | 0.005 | 0.009 | 0.018 |
| 166 | 0.939 | -2.189 | 0.119 | 8 | 0 | 0 | 0 | 0.008 | 0 | 0 | 0 |
| 167 | 1.89 | 1.918 | 0.177 | 3 | 0 | 0 | 3 | 0.003 | 0 | 0 | 0.003 |
| 168 | 2.471 | -2.456 | 0.055 | 4 | 8 | 2 | 0 | 0.004 | 0.008 | 0.002 | 0 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 169 | 1.316 | -0.93 | 0.113 | 12 | 2 | 3 | 14 | 0.012 | 0.002 | 0.003 | 0.014 |
| 170 | 1.903 | 0.943 | 0.233 | 31 | 1 | 4 | 14 | 0.031 | 0.001 | 0.004 | 0.014 |
| 171 | 1.875 | 1.751 | 0.298 | 11 | 0 | 0 | 2 | 0.011 | 0 | 0 | 0.002 |
| 172 | 1.094 | -2.435 | 0.047 | 3 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 |
| 173 | 1.495 | 1.676 | 0.077 | 11 | 0 | 0 | 0 | 0.011 | 0 | 0 | 0 |
| 174 | 2.206 | -0.859 | 0.145 | 63 | 17 | 8 | 70 | 0.063 | 0.017 | 0.008 | 0.07 |
| 175 | 0.995 | 0.24 | 0.14 | 28 | 8 | 3 | 5 | 0.028 | 0.008 | 0.003 | 0.005 |
| 176 | 1.95 | 1.257 | 0.198 | 2 | 0 | 0 | 8 | 0.002 | 0 | 0 | 0.008 |
| 177 | 2.001 | -1.101 | 0.102 | 21 | 7 | 0 | 44 | 0.021 | 0.007 | 0 | 0.044 |
| 178 | 0.874 | -1.442 | 0.059 | 9 | 5 | 0 | 0 | 0.009 | 0.005 | 0 | 0 |
| 179 | 0.989 | 2.055 | 0.192 | 12 | 0 | 0 | 0 | 0.012 | 0 | 0 | 0 |
| 180 | 0.502 | -1.979 | 0.091 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 181 | 1.036 | -1.121 | 0.196 | 14 | 0 | 0 | 0 | 0.014 | 0 | 0 | 0 |
| 182 | 1.632 | 1.077 | 0.122 | 24 | 3 | 0 | 6 | 0.024 | 0.003 | 0 | 0.006 |
| 183 | 1.314 | 1.712 | 0.045 | 4 | 0 | 0 | 6 | 0.004 | 0 | 0 | 0.006 |
| 184 | 1.617 | 1.616 | 0.137 | 23 | 0 | 0 | 0 | 0.023 | 0 | 0 | 0 |
| 185 | 0.915 | -0.046 | 0.163 | 139 | 5 | 7 | 4 | 0.139 | 0.005 | 0.007 | 0.004 |
| 186 | 0.931 | -1.862 | 0.201 | 8 | 0 | 0 | 0 | 0.008 | 0 | 0 | 0 |
| 187 | 2.054 | -0.404 | 0.098 | 67 | 9 | 13 | 112 | 0.067 | 0.009 | 0.013 | 0.112 |
| 188 | 1.841 | -0.384 | 0.201 | 148 | 10 | 6 | 33 | 0.148 | 0.01 | 0.006 | 0.033 |
| 189 | 2.22 | -0.195 | 0.004 | 66 | 240 | 525 | 210 | 0.066 | 0.24 | 0.525 | 0.21 |
| 190 | 2.484 | 0.812 | 0.037 | 24 | 613 | 336 | 121 | 0.024 | 0.613 | 0.336 | 0.121 |
| 191 | 1.328 | -0.782 | 0.05 | 13 | 13 | 2 | 119 | 0.013 | 0.013 | 0.002 | 0.119 |
| 192 | 2.184 | -0.889 | 0.092 | 8 | 18 | 8 | 87 | 0.008 | 0.018 | 0.008 | 0.087 |
| 193 | 2.455 | 0.843 | 0.085 | 32 | 165 | 311 | 109 | 0.032 | 0.165 | 0.311 | 0.109 |
| 194 | 2.294 | 1.133 | 0.077 | 12 | 164 | 68 | 61 | 0.012 | 0.164 | 0.068 | 0.061 |
| 195 | 1.192 | 0.736 | 0.158 | 18 | 5 | 4 | 11 | 0.018 | 0.005 | 0.004 | 0.011 |
| 196 | 1.4 | -1.339 | 0.125 | 40 | 5 | 0 | 0 | 0.04 | 0.005 | 0 | 0 |
| 197 | 1.206 | 1.85 | 0.193 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 198 | 2.186 | -0.325 | 0.275 | 46 | 8 | 5 | 63 | 0.046 | 0.008 | 0.005 | 0.063 |
| 199 | 2.05 | 0.718 | 0.171 | 11 | 7 | 5 | 67 | 0.011 | 0.007 | 0.005 | 0.067 |
| 200 | 1.769 | 1.5 | 0.271 | 11 | 0 | 0 | 0 | 0.011 | 0 | 0 | 0 |
| 201 | 1.796 | 0.873 | 0.151 | 31 | 8 | 4 | 31 | 0.031 | 0.008 | 0.004 | 0.031 |
| 202 | 1.838 | -0.62 | 0.176 | 52 | 9 | 4 | 31 | 0.052 | 0.009 | 0.004 | 0.031 |
| 203 | 1.556 | -1.839 | 0.144 | 3 | 0 | 0 | 3 | 0.003 | 0 | 0 | 0.003 |
| 204 | 1.377 | -0.016 | 0.268 | 90 | 11 | 9 | 3 | 0.09 | 0.011 | 0.009 | 0.003 |
| 205 | 1.438 | -0.639 | 0.222 | 18 | 8 | 1 | 11 | 0.018 | 0.008 | 0.001 | 0.011 |
| 206 | 1.315 | 0.637 | 0.267 | 24 | 10 | 4 | 21 | 0.024 | 0.01 | 0.004 | 0.021 |
| 207 | 2.023 | 1.753 | 0.124 | 7 | 0 | 0 | 3 | 0.007 | 0 | 0 | 0.003 |
| 208 | 0.897 | 1.303 | 0.117 | 14 | 0 | 0 | 0 | 0.014 | 0 | 0 | 0 |
| 209 | 1.897 | -2.443 | 0.083 | 1 | 0 | 0 | 2 | 0.001 | 0 | 0 | 0.002 |
| 210 | 2.096 | -2.392 | 0.111 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 211 | 2.33 | -1.611 | 0.208 | 10 | 0 | 9 | 18 | 0.01 | 0 | 0.009 | 0.018 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 212 | 0.873 | -0.574 | 0.048 | 20 | 6 | 6 | 7 | 0.02 | 0.006 | 0.006 | 0.007 |
| 213 | 0.64 | 1.14 | 0.269 | 4 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 |
| 214 | 0.608 | 0.81 | 0.146 | 11 | 4 | 9 | 3 | 0.011 | 0.004 | 0.009 | 0.003 |
| 215 | 2.323 | -2.06 | 0.082 | 0 | 4 | 9 | 11 | 0 | 0.004 | 0.009 | 0.011 |
| 216 | 1.499 | -1.691 | 0.26 | 14 | 0 | 0 | 0 | 0.014 | 0 | 0 | 0 |
| 217 | 2.497 | -0.889 | 0.1 | 7 | 176 | 272 | 125 | 0.007 | 0.176 | 0.272 | 0.125 |
| 218 | 2.256 | 0.938 | 0.27 | 28 | 10 | 11 | 36 | 0.028 | 0.01 | 0.011 | 0.036 |
| 219 | 1.73 | -0.737 | 0.132 | 25 | 20 | 5 | 19 | 0.025 | 0.02 | 0.005 | 0.019 |
| 220 | 0.898 | 1.345 | 0.143 | 4 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 |
| 221 | 1.816 | -2.06 | 0.177 | 1 | 0 | 0 | 2 | 0.001 | 0 | 0 | 0.002 |
| 222 | 1.956 | -1.053 | 0.292 | 27 | 0 | 0 | 11 | 0.027 | 0 | 0 | 0.011 |
| 223 | 1.557 | 2.474 | 0.225 | 19 | 0 | 0 | 0 | 0.019 | 0 | 0 | 0 |
| 224 | 1.91 | 2.07 | 0.266 | 14 | 0 | 0 | 0 | 0.014 | 0 | 0 | 0 |
| 225 | 1.117 | 1.953 | 0.158 | 8 | 0 | 0 | 0 | 0.008 | 0 | 0 | 0 |
| 226 | 1.163 | 2.379 | 0.163 | 13 | 0 | 0 | 0 | 0.013 | 0 | 0 | 0 |
| 227 | 2.003 | 2.457 | 0.076 | 11 | 9 | 0 | 1 | 0.011 | 0.009 | 0 | 0.001 |
| 228 | 1.367 | 1.115 | 0.244 | 15 | 0 | 0 | 1 | 0.015 | 0 | 0 | 0.001 |
| 229 | 2.252 | -0.029 | 0.206 | 66 | 7 | 43 | 89 | 0.066 | 0.007 | 0.043 | 0.089 |
| 230 | 1.407 | -2.059 | 0.282 | 3 | 0 | 0 | 1 | 0.003 | 0 | 0 | 0.001 |
| 231 | 0.862 | -1.369 | 0.283 | 34 | 0 | 0 | 0 | 0.034 | 0 | 0 | 0 |
| 232 | 1.875 | -0.593 | 0.285 | 37 | 15 | 3 | 21 | 0.037 | 0.015 | 0.003 | 0.021 |
| 233 | 2.058 | -1.796 | 0.11 | 11 | 0 | 0 | 12 | 0.011 | 0 | 0 | 0.012 |
| 234 | 2.319 | 0.332 | 0.074 | 22 | 59 | 165 | 123 | 0.022 | 0.059 | 0.165 | 0.123 |
| 235 | 2.428 | 0.969 | 0.133 | 11 | 18 | 113 | 60 | 0.011 | 0.018 | 0.113 | 0.06 |
| 236 | 2.1 | -2.375 | 0.101 | 1 | 0 | 0 | 3 | 0.001 | 0 | 0 | 0.003 |
| 237 | 0.515 | -0.609 | 0.285 | 13 | 5 | 3 | 4 | 0.013 | 0.005 | 0.003 | 0.004 |
| 238 | 2.165 | 2.325 | 0.244 | 11 | 0 | 0 | 1 | 0.011 | 0 | 0 | 0.001 |
| 239 | 0.907 | 2.261 | 0.293 | 8 | 0 | 0 | 0 | 0.008 | 0 | 0 | 0 |
| 240 | 1.543 | 1.902 | 0.005 | 6 | 0 | 0 | 1 | 0.006 | 0 | 0 | 0.001 |
| 241 | 1.403 | -1.81 | 0.148 | 2 | 0 | 0 | 2 | 0.002 | 0 | 0 | 0.002 |
| 242 | 0.879 | 1.089 | 0.01 | 8 | 0 | 1 | 1 | 0.008 | 0 | 0.001 | 0.001 |
| 243 | 2.315 | 1.629 | 0.172 | 11 | 0 | 24 | 13 | 0.011 | 0 | 0.024 | 0.013 |
| 244 | 1.825 | 0.779 | 0.276 | 12 | 4 | 5 | 22 | 0.012 | 0.004 | 0.005 | 0.022 |
| 245 | 1.455 | -2.184 | 0.16 | 2 | 0 | 0 | 1 | 0.002 | 0 | 0 | 0.001 |
| 246 | 1.59 | 0.773 | 0.103 | 24 | 9 | 4 | 20 | 0.024 | 0.009 | 0.004 | 0.02 |
| 247 | 0.892 | -0.705 | 0.149 | 42 | 5 | 7 | 5 | 0.042 | 0.005 | 0.007 | 0.005 |
| 248 | 2.469 | -1.594 | 0.299 | 3 | 0 | 14 | 28 | 0.003 | 0 | 0.014 | 0.028 |
| 249 | 2.017 | -2.063 | 0.056 | 5 | 0 | 0 | 8 | 0.005 | 0 | 0 | 0.008 |
| 250 | 0.658 | 1.332 | 0.151 | 18 | 3 | 2 | 0 | 0.018 | 0.003 | 0.002 | 0 |
| 251 | 2.003 | -2.406 | 0.293 | 2 | 0 | 0 | 6 | 0.002 | 0 | 0 | 0.006 |
| 252 | 1.177 | -1.002 | 0.19 | 38 | 0 | 0 | 0 | 0.038 | 0 | 0 | 0 |
| 253 | 1.201 | -1.189 | 0.141 | 26 | 1 | 0 | 0 | 0.026 | 0.001 | 0 | 0 |
| 254 | 2.2 | -1.398 | 0.031 | 12 | 46 | 6 | 60 | 0.012 | 0.046 | 0.006 | 0.06 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 255 | 1.548 | -1.758 | 0.045 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0.004 |
| 256 | 0.615 | -2.464 | 0.077 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 257 | 0.545 | 1.678 | 0.263 | 1 | 0 | 0 | 3 | 0.001 | 0 | 0 | 0.003 |
| 258 | 1.125 | 0.633 | 0.101 | 78 | 3 | 4 | 5 | 0.078 | 0.003 | 0.004 | 0.005 |
| 259 | 1.866 | 0.849 | 0.078 | 57 | 10 | 2 | 49 | 0.057 | 0.01 | 0.002 | 0.049 |
| 260 | 1.695 | 1.719 | 0.077 | 3 | 0 | 0 | 1 | 0.003 | 0 | 0 | 0.001 |
| 261 | 2.165 | 0.397 | 0.06 | 43 | 36 | 9 | 107 | 0.043 | 0.036 | 0.009 | 0.107 |
| 262 | 0.666 | 0.265 | 0.002 | 98 | 7 | 8 | 3 | 0.098 | 0.007 | 0.008 | 0.003 |
| 263 | 2.14 | -2.313 | 0.034 | 1 | 0 | 0 | 6 | 0.001 | 0 | 0 | 0.006 |
| 264 | 1.35 | 0.171 | 0.033 | 61 | 4 | 6 | 24 | 0.061 | 0.004 | 0.006 | 0.024 |
| 265 | 2.414 | -0.655 | 0.188 | 31 | 22 | 39 | 105 | 0.031 | 0.022 | 0.039 | 0.105 |
| 266 | 2.407 | -1.044 | 0.238 | 2 | 2 | 24 | 53 | 0.002 | 0.002 | 0.024 | 0.053 |
| 267 | 1.306 | -0.294 | 0.236 | 118 | 5 | 5 | 5 | 0.118 | 0.005 | 0.005 | 0.005 |
| 268 | 0.825 | 1.892 | 0.232 | 7 | 2 | 0 | 0 | 0.007 | 0.002 | 0 | 0 |
| 269 | 2.038 | -2.452 | 0.013 | 6 | 2 | 0 | 3 | 0.006 | 0.002 | 0 | 0.003 |
| 270 | 0.818 | 1.86 | 0.235 | 8 | 1 | 0 | 0 | 0.008 | 0.001 | 0 | 0 |
| 271 | 1.459 | -1.204 | 0.171 | 36 | 0 | 0 | 1 | 0.036 | 0 | 0 | 0.001 |
| 272 | 2.448 | 1.96 | 0.098 | 4 | 40 | 24 | 8 | 0.004 | 0.04 | 0.024 | 0.008 |
| 273 | 0.832 | -0.923 | 0.192 | 40 | 3 | 4 | 5 | 0.04 | 0.003 | 0.004 | 0.005 |
| 274 | 2.285 | -0.844 | 0.064 | 59 | 317 | 119 | 126 | 0.059 | 0.317 | 0.119 | 0.126 |
| 275 | 1.758 | 1.306 | 0.055 | 16 | 2 | 2 | 11 | 0.016 | 0.002 | 0.002 | 0.011 |
| 276 | 0.92 | -1.514 | 0.259 | 17 | 0 | 0 | 0 | 0.017 | 0 | 0 | 0 |
| 277 | 0.795 | -0.312 | 0.214 | 93 | 7 | 1 | 7 | 0.093 | 0.007 | 0.001 | 0.007 |
| 278 | 0.642 | -1.376 | 0.15 | 19 | 1 | 0 | 0 | 0.019 | 0.001 | 0 | 0 |
| 279 | 0.832 | -2.19 | 0.066 | 0 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 |
| 280 | 1.09 | 0.815 | 0.189 | 25 | 8 | 3 | 3 | 0.025 | 0.008 | 0.003 | 0.003 |
| 281 | 1.788 | -0.287 | 0.23 | 16 | 7 | 6 | 37 | 0.016 | 0.007 | 0.006 | 0.037 |
| 282 | 1.199 | -0.524 | 0.151 | 14 | 9 | 5 | 5 | 0.014 | 0.009 | 0.005 | 0.005 |
| 283 | 1.687 | -0.881 | 0.255 | 20 | 11 | 7 | 8 | 0.02 | 0.011 | 0.007 | 0.008 |
| 284 | 2.239 | -1.018 | 0.125 | 16 | 2 | 9 | 78 | 0.016 | 0.002 | 0.009 | 0.078 |
| 285 | 1.689 | -1.801 | 0.244 | 2 | 0 | 0 | 3 | 0.002 | 0 | 0 | 0.003 |
| 286 | 1.278 | 2.433 | 0.013 | 14 | 0 | 0 | 0 | 0.014 | 0 | 0 | 0 |
| 287 | 1.115 | -1.449 | 0.235 | 39 | 1 | 0 | 0 | 0.039 | 0.001 | 0 | 0 |
| 288 | 0.862 | -0.586 | 0.06 | 32 | 5 | 6 | 3 | 0.032 | 0.005 | 0.006 | 0.003 |
| 289 | 1.052 | -1.702 | 0.296 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 290 | 1.242 | 1.052 | 0.199 | 13 | 1 | 0 | 1 | 0.013 | 0.001 | 0 | 0.001 |
| 291 | 1.592 | 0.443 | 0.136 | 75 | 6 | 12 | 30 | 0.075 | 0.006 | 0.012 | 0.03 |
| 292 | 1.947 | 1.88 | 0.007 | 11 | 29 | 4 | 10 | 0.011 | 0.029 | 0.004 | 0.01 |
| 293 | 2.258 | -1.089 | 0.102 | 33 | 78 | 81 | 82 | 0.033 | 0.078 | 0.081 | 0.082 |
| 294 | 2.025 | -0.695 | 0.192 | 69 | 10 | 6 | 36 | 0.069 | 0.01 | 0.006 | 0.036 |
| 295 | 1.695 | -0.082 | 0.147 | 10 | 5 | 4 | 41 | 0.01 | 0.005 | 0.004 | 0.041 |
| 296 | 0.758 | -2.3 | 0.048 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 297 | 1.177 | -2.423 | 0.281 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 298 | 1.166 | 0.409 | 0.078 | 50 | 4 | 4 | 9 | 0.05 | 0.004 | 0.004 | 0.009 |
| 299 | 1 | -1.098 | 0.125 | 6 | 0 | 0 | 0 | 0.006 | 0 | 0 | 0 |
| 300 | 1.115 | -0.806 | 0.143 | 49 | 5 | 2 | 4 | 0.049 | 0.005 | 0.002 | 0.004 |
| 301 | 1.319 | -0.388 | 0.053 | 46 | 5 | 6 | 222 | 0.046 | 0.005 | 0.006 | 0.222 |
| 302 | 1.337 | 1.9 | 0.099 | 12 | 0 | 0 | 0 | 0.012 | 0 | 0 | 0 |
| 303 | 1.492 | -1.612 | 0.178 | 0 | 5 | 0 | 3 | 0 | 0.005 | 0 | 0.003 |
| 304 | 2.177 | 0.384 | 0.029 | 46 | 150 | 156 | 146 | 0.046 | 0.15 | 0.156 | 0.146 |
| 305 | 2.495 | -1.499 | 0.113 | 8 | 11 | 42 | 49 | 0.008 | 0.011 | 0.042 | 0.049 |
| 306 | 1.378 | 1.965 | 0.283 | 3 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 |
| 307 | 1.471 | 1.951 | 0.13 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 308 | 1.721 | -1.047 | 0.082 | 12 | 6 | 0 | 28 | 0.012 | 0.006 | 0 | 0.028 |
| 309 | 1.97 | -2.461 | 0.18 | 1 | 0 | 0 | 1 | 0.001 | 0 | 0 | 0.001 |
| 310 | 1.824 | 2.017 | 0.148 | 7 | 0 | 0 | 1 | 0.007 | 0 | 0 | 0.001 |
| 311 | 1.956 | 0.516 | 0.256 | 89 | 10 | 23 | 28 | 0.089 | 0.01 | 0.023 | 0.028 |
| 312 | 1.925 | -0.56 | 0.163 | 8 | 14 | 5 | 56 | 0.008 | 0.014 | 0.005 | 0.056 |
| 313 | 2.121 | -2.278 | 0.217 | 2 | 0 | 0 | 2 | 0.002 | 0 | 0 | 0.002 |
| 314 | 0.941 | 0.526 | 0.282 | 58 | 9 | 4 | 5 | 0.058 | 0.009 | 0.004 | 0.005 |
| 315 | 1.595 | -1.629 | 0.29 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 316 | 1.648 | -2.298 | 0.034 | 1 | 0 | 0 | 2 | 0.001 | 0 | 0 | 0.002 |
| 317 | 0.592 | 1.885 | 0.233 | 7 | 2 | 0 | 0 | 0.007 | 0.002 | 0 | 0 |
| 318 | 1.253 | -1.904 | 0.099 | 7 | 0 | 0 | 1 | 0.007 | 0 | 0 | 0.001 |
| 319 | 2.068 | 0.434 | 0.015 | 49 | 65 | 63 | 125 | 0.049 | 0.065 | 0.063 | 0.125 |
| 320 | 1.213 | 1.61 | 0.148 | 6 | 0 | 0 | 0 | 0.006 | 0 | 0 | 0 |
| 321 | 0.95 | 0.095 | 0.155 | 95 | 6 | 3 | 5 | 0.095 | 0.006 | 0.003 | 0.005 |
| 322 | 0.576 | -1.209 | 0.167 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 323 | 1.995 | 0.771 | 0.221 | 26 | 4 | 7 | 36 | 0.026 | 0.004 | 0.007 | 0.036 |
| 324 | 1.686 | -1.996 | 0.12 | 6 | 0 | 0 | 3 | 0.006 | 0 | 0 | 0.003 |
| 325 | 1.844 | 1.294 | 0.078 | 8 | 0 | 0 | 9 | 0.008 | 0 | 0 | 0.009 |
| 326 | 1.547 | -2.284 | 0.112 | 5 | 0 | 0 | 6 | 0.005 | 0 | 0 | 0.006 |
| 327 | 1.341 | 2.356 | 0.144 | 10 | 0 | 0 | 0 | 0.01 | 0 | 0 | 0 |
| 328 | 0.66 | 0.899 | 0.265 | 44 | 8 | 4 | 5 | 0.044 | 0.008 | 0.004 | 0.005 |
| 329 | 0.659 | 1.182 | 0.286 | 16 | 0 | 0 | 1 | 0.016 | 0 | 0 | 0.001 |
| 330 | 1.658 | -0.154 | 0.202 | 40 | 11 | 8 | 28 | 0.04 | 0.011 | 0.008 | 0.028 |
| 331 | 2.323 | -1.909 | 0.072 | 2 | 41 | 23 | 20 | 0.002 | 0.041 | 0.023 | 0.02 |
| 332 | 0.533 | 0.934 | 0.22 | 46 | 4 | 7 | 6 | 0.046 | 0.004 | 0.007 | 0.006 |
| 333 | 1.671 | 1.365 | 0.151 | 1 | 0 | 0 | 3 | 0.001 | 0 | 0 | 0.003 |
| 334 | 1.784 | 0.101 | 0.296 | 22 | 5 | 8 | 17 | 0.022 | 0.005 | 0.008 | 0.017 |
| 335 | 0.819 | -1.93 | 0.136 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 336 | 1.291 | 1.08 | 0.033 | 23 | 11 | 0 | 78 | 0.023 | 0.011 | 0 | 0.078 |
| 337 | 0.706 | -0.258 | 0.109 | 112 | 5 | 3 | 7 | 0.112 | 0.005 | 0.003 | 0.007 |
| 338 | 1.219 | -0.327 | 0.076 | 66 | 3 | 8 | 5 | 0.066 | 0.003 | 0.008 | 0.005 |
| 339 | 2.266 | -1.905 | 0.069 | 7 | 9 | 0 | 17 | 0.007 | 0.009 | 0 | 0.017 |
| 340 | 1.789 | -1.414 | 0.13 | 7 | 0 | 0 | 12 | 0.007 | 0 | 0 | 0.012 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 341 | 2.295 | 2.148 | 0.103 | 5 | 23 | 9 | 2 | 0.005 | 0.023 | 0.009 | 0.002 |
| 342 | 1.201 | -1.39 | 0.255 | 21 | 0 | 0 | 0 | 0.021 | 0 | 0 | 0 |
| 343 | 2.076 | -0.557 | 0.24 | 86 | 11 | 5 | 49 | 0.086 | 0.011 | 0.005 | 0.049 |
| 344 | 1.535 | -1.811 | 0.157 | 1 | 0 | 0 | 4 | 0.001 | 0 | 0 | 0.004 |
| 345 | 1.33 | 0.243 | 0.214 | 11 | 7 | 2 | 6 | 0.011 | 0.007 | 0.002 | 0.006 |
| 346 | 0.948 | -0.361 | 0.252 | 88 | 6 | 2 | 4 | 0.088 | 0.006 | 0.002 | 0.004 |
| 347 | 0.538 | -0.121 | 0.068 | 86 | 5 | 4 | 5 | 0.086 | 0.005 | 0.004 | 0.005 |
| 348 | 0.965 | 1.565 | 0.001 | 18 | 2 | 0 | 2 | 0.018 | 0.002 | 0 | 0.002 |
| 349 | 1.221 | 0.65 | 0.153 | 45 | 6 | 3 | 7 | 0.045 | 0.006 | 0.003 | 0.007 |
| 350 | 2.188 | -1.753 | 0.146 | 5 | 1 | 1 | 17 | 0.005 | 0.001 | 0.001 | 0.017 |
| 351 | 2.221 | 0.922 | 0.107 | 7 | 3 | 9 | 68 | 0.007 | 0.003 | 0.009 | 0.068 |
| 352 | 2.23 | -1.969 | 0.277 | 6 | 0 | 0 | 6 | 0.006 | 0 | 0 | 0.006 |
| 353 | 0.552 | 0.351 | 0.036 | 61 | 5 | 2 | 8 | 0.061 | 0.005 | 0.002 | 0.008 |
| 354 | 2.19 | -0.994 | 0.013 | 26 | 314 | 215 | 92 | 0.026 | 0.314 | 0.215 | 0.092 |
| 355 | 1.886 | 1.364 | 0.006 | 17 | 193 | 49 | 39 | 0.017 | 0.193 | 0.049 | 0.039 |
| 356 | 2.475 | 1.498 | 0.198 | 12 | 20 | 84 | 17 | 0.012 | 0.02 | 0.084 | 0.017 |
| 357 | 2.236 | -1.674 | 0.036 | 9 | 65 | 40 | 37 | 0.009 | 0.065 | 0.04 | 0.037 |
| 358 | 1.294 | -1.845 | 0.222 | 9 | 0 | 0 | 0 | 0.009 | 0 | 0 | 0 |
| 359 | 0.889 | 2.031 | 0.146 | 5 | 0 | 0 | 0 | 0.005 | 0 | 0 | 0 |
| 360 | 1.619 | -1.326 | 0.019 | 2 | 3 | 0 | 20 | 0.002 | 0.003 | 0 | 0.02 |
| 361 | 1.173 | -1.56 | 0.1 | 10 | 0 | 0 | 0 | 0.01 | 0 | 0 | 0 |
| 362 | 0.945 | 2.425 | 0.02 | 11 | 0 | 0 | 0 | 0.011 | 0 | 0 | 0 |
| 363 | 2.089 | 1.802 | 0.049 | 14 | 66 | 12 | 14 | 0.014 | 0.066 | 0.012 | 0.014 |
| 364 | 0.64 | -1.479 | 0.248 | 12 | 0 | 0 | 0 | 0.012 | 0 | 0 | 0 |
| 365 | 2.406 | -1.413 | 0.055 | 7 | 133 | 147 | 58 | 0.007 | 0.133 | 0.147 | 0.058 |
| 366 | 2.222 | -0.882 | 0.084 | 21 | 20 | 10 | 96 | 0.021 | 0.02 | 0.01 | 0.096 |
| 367 | 1.251 | -0.128 | 0.168 | 34 | 6 | 4 | 11 | 0.034 | 0.006 | 0.004 | 0.011 |
| 368 | 2.352 | -2.047 | 0.045 | 0 | 31 | 14 | 19 | 0 | 0.031 | 0.014 | 0.019 |
| 369 | 1.776 | 1.082 | 0.075 | 20 | 5 | 0 | 25 | 0.02 | 0.005 | 0 | 0.025 |
| 370 | 0.871 | 2.436 | 0.054 | 20 | 0 | 0 | 0 | 0.02 | 0 | 0 | 0 |
| 371 | 0.956 | -0.73 | 0.143 | 16 | 6 | 4 | 4 | 0.016 | 0.006 | 0.004 | 0.004 |
| 372 | 0.689 | -1.209 | 0.186 | 14 | 1 | 0 | 0 | 0.014 | 0.001 | 0 | 0 |
| 373 | 1.081 | -0.702 | 0.032 | 35 | 4 | 6 | 3 | 0.035 | 0.004 | 0.006 | 0.003 |
| 374 | 0.834 | 1.059 | 0.032 | 5 | 0 | 0 | 0 | 0.005 | 0 | 0 | 0 |
| 375 | 1.006 | 1.624 | 0.14 | 9 | 4 | 0 | 0 | 0.009 | 0.004 | 0 | 0 |
| 376 | 2.484 | 0.968 | 0.104 | 15 | 115 | 299 | 81 | 0.015 | 0.115 | 0.299 | 0.081 |
| 377 | 1.727 | 0.558 | 0.119 | 9 | 19 | 9 | 40 | 0.009 | 0.019 | 0.009 | 0.04 |
| 378 | 2.281 | 1.972 | 0.07 | 6 | 16 | 11 | 7 | 0.006 | 0.016 | 0.011 | 0.007 |
| 379 | 1.329 | -1.825 | 0.162 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0.001 |
| 380 | 1.754 | 1.978 | 0.11 | 3 | 0 | 0 | 1 | 0.003 | 0 | 0 | 0.001 |
| 381 | 0.89 | 0.275 | 0.052 | 61 | 10 | 7 | 7 | 0.061 | 0.01 | 0.007 | 0.007 |
| 382 | 2.145 | 0.544 | 0.016 | 40 | 206 | 94 | 137 | 0.04 | 0.206 | 0.094 | 0.137 |
| 383 | 1.386 | 2.046 | 0.034 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 384 | 2.476 | -2.417 | 0.08 | 2 | 8 | 2 | 8 | 0.002 | 0.008 | 0.002 | 0.008 |
| 385 | 2.169 | 2.346 | 0.137 | 10 | 8 | 1 | 0 | 0.01 | 0.008 | 0.001 | 0 |
| 386 | 1.146 | 1.855 | 0.048 | 11 | 0 | 0 | 0 | 0.011 | 0 | 0 | 0 |
| 387 | 2.079 | 1.753 | 0.062 | 19 | 26 | 4 | 7 | 0.019 | 0.026 | 0.004 | 0.007 |
| 388 | 1.875 | -1.233 | 0.028 | 19 | 1 | 0 | 33 | 0.019 | 0.001 | 0 | 0.033 |
| 389 | 0.744 | 0.789 | 0.022 | 24 | 8 | 5 | 9 | 0.024 | 0.008 | 0.005 | 0.009 |
| 390 | 1.328 | -1.722 | 0.081 | 5 | 1 | 0 | 20 | 0.005 | 0.001 | 0 | 0.02 |
| 391 | 2.153 | 0.975 | 0.275 | 57 | 14 | 7 | 22 | 0.057 | 0.014 | 0.007 | 0.022 |
| 392 | 2.198 | -0.068 | 0.143 | 17 | 6 | 38 | 107 | 0.017 | 0.006 | 0.038 | 0.107 |
| 393 | 1.031 | -2.368 | 0.221 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 394 | 2.09 | -1.7 | 0.141 | 7 | 0 | 1 | 18 | 0.007 | 0 | 0.001 | 0.018 |
| 395 | 1.109 | -0.433 | 0.028 | 19 | 1 | 5 | 13 | 0.019 | 0.001 | 0.005 | 0.013 |
| 396 | 1 | -1.64 | 0.009 | 4 | 1 | 0 | 0 | 0.004 | 0.001 | 0 | 0 |
| 397 | 0.837 | 2.061 | 0.112 | 3 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 |
| 398 | 0.713 | 1.607 | 0.224 | 5 | 0 | 0 | 0 | 0.005 | 0 | 0 | 0 |
| 399 | 0.602 | 0.825 | 0.095 | 20 | 5 | 2 | 6 | 0.02 | 0.005 | 0.002 | 0.006 |
| 400 | 1.924 | -0.089 | 0.298 | 41 | 9 | 4 | 44 | 0.041 | 0.009 | 0.004 | 0.044 |
| 401 | 2.019 | 0.803 | 0.209 | 12 | 0 | 7 | 45 | 0.012 | 0 | 0.007 | 0.045 |
| 402 | 1.397 | 0.41 | 0.224 | 30 | 7 | 2 | 3 | 0.03 | 0.007 | 0.002 | 0.003 |
| 403 | 2.354 | 0.006 | 0.024 | 42 | 30 | 625 | 168 | 0.042 | 0.03 | 0.625 | 0.168 |
| 404 | 1.981 | 0.975 | 0.035 | 34 | 18 | 18 | 77 | 0.034 | 0.018 | 0.018 | 0.077 |
| 405 | 1.641 | -0.23 | 0.011 | 45 | 8 | 6 | 81 | 0.045 | 0.008 | 0.006 | 0.081 |
| 406 | 0.785 | 2.019 | 0.238 | 6 | 0 | 0 | 0 | 0.006 | 0 | 0 | 0 |
| 407 | 0.557 | -0.911 | 0.079 | 38 | 12 | 7 | 5 | 0.038 | 0.012 | 0.007 | 0.005 |
| 408 | 1.333 | 1.682 | 0.239 | 3 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 |
| 409 | 1.624 | 1.948 | 0.129 | 10 | 0 | 0 | 0 | 0.01 | 0 | 0 | 0 |
| 410 | 2.004 | -2.321 | 0.046 | 1 | 0 | 0 | 5 | 0.001 | 0 | 0 | 0.005 |
| 411 | 0.781 | -2.208 | 0.244 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 412 | 1.576 | 0.345 | 0.145 | 25 | 20 | 8 | 29 | 0.025 | 0.02 | 0.008 | 0.029 |
| 413 | 0.911 | 1.543 | 0.277 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 414 | 1.186 | -2.359 | 0.166 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 415 | 1.376 | 0.837 | 0.217 | 34 | 5 | 4 | 9 | 0.034 | 0.005 | 0.004 | 0.009 |
| 416 | 0.514 | -0.827 | 0.092 | 75 | 10 | 9 | 6 | 0.075 | 0.01 | 0.009 | 0.006 |
| 417 | 2.116 | -0.638 | 0.053 | 31 | 55 | 21 | 113 | 0.031 | 0.055 | 0.021 | 0.113 |
| 418 | 0.507 | 2.494 | 0.281 | 14 | 0 | 0 | 4 | 0.014 | 0 | 0 | 0.004 |
| 419 | 1.385 | 1.47 | 0.062 | 18 | 0 | 0 | 1 | 0.018 | 0 | 0 | 0.001 |
| 420 | 1.45 | 1.218 | 0.292 | 25 | 0 | 0 | 1 | 0.025 | 0 | 0 | 0.001 |
| 421 | 0.683 | -0.536 | 0.082 | 52 | 4 | 1 | 3 | 0.052 | 0.004 | 0.001 | 0.003 |
| 422 | 1.713 | 0.74 | 0.148 | 65 | 12 | 6 | 9 | 0.065 | 0.012 | 0.006 | 0.009 |
| 423 | 0.945 | 1.219 | 0.053 | 47 | 1 | 2 | 4 | 0.047 | 0.001 | 0.002 | 0.004 |
| 424 | 0.959 | -1.421 | 0.272 | 24 | 0 | 0 | 0 | 0.024 | 0 | 0 | 0 |
| 425 | 0.546 | 2.041 | 0.244 | 11 | 0 | 0 | 0 | 0.011 | 0 | 0 | 0 |
| 426 | 2.156 | -2.371 | 0.275 | 2 | 0 | 0 | 4 | 0.002 | 0 | 0 | 0.004 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 427 | 2.196 | 0.917 | 0.281 | 10 | 4 | 8 | 30 | 0.01 | 0.004 | 0.008 | 0.03 |
| 428 | 1.175 | 1.858 | 0.19 | 5 | 0 | 0 | 0 | 0.005 | 0 | 0 | 0 |
| 429 | 2.276 | -2.432 | 0.067 | 2 | 3 | 0 | 3 | 0.002 | 0.003 | 0 | 0.003 |
| 430 | 1.32 | 0.226 | 0.275 | 36 | 4 | 4 | 4 | 0.036 | 0.004 | 0.004 | 0.004 |
| 431 | 0.684 | -2.038 | 0.039 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 432 | 0.568 | 0.859 | 0.128 | 28 | 12 | 6 | 4 | 0.028 | 0.012 | 0.006 | 0.004 |
| 433 | 1.747 | 2.185 | 0.107 | 3 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 |
| 434 | 1.858 | -1.249 | 0.297 | 22 | 0 | 0 | 8 | 0.022 | 0 | 0 | 0.008 |
| 435 | 1.272 | -0.164 | 0.285 | 34 | 3 | 5 | 11 | 0.034 | 0.003 | 0.005 | 0.011 |
| 436 | 1.229 | 1.545 | 0.061 | 8 | 2 | 0 | 5 | 0.008 | 0.002 | 0 | 0.005 |
| 437 | 1.287 | -0.656 | 0.09 | 31 | 2 | 6 | 8 | 0.031 | 0.002 | 0.006 | 0.008 |
| 438 | 1.118 | -0.852 | 0.132 | 19 | 2 | 3 | 7 | 0.019 | 0.002 | 0.003 | 0.007 |
| 439 | 0.543 | -1.13 | 0.177 | 5 | 0 | 0 | 0 | 0.005 | 0 | 0 | 0 |
| 440 | 0.646 | -0.762 | 0.196 | 36 | 5 | 6 | 6 | 0.036 | 0.005 | 0.006 | 0.006 |
| 441 | 1.335 | -2.304 | 0.14 | 1 | 0 | 0 | 3 | 0.001 | 0 | 0 | 0.003 |
| 442 | 1.178 | 1.918 | 0.19 | 10 | 0 | 0 | 0 | 0.01 | 0 | 0 | 0 |
| 443 | 1.705 | -1.742 | 0.286 | 1 | 0 | 0 | 1 | 0.001 | 0 | 0 | 0.001 |
| 444 | 1.994 | -1.926 | 0.191 | 7 | 0 | 0 | 1 | 0.007 | 0 | 0 | 0.001 |
| 445 | 0.596 | 0.233 | 0.275 | 104 | 5 | 12 | 5 | 0.104 | 0.005 | 0.012 | 0.005 |
| 446 | 1.374 | 2.196 | 0.157 | 5 | 0 | 0 | 0 | 0.005 | 0 | 0 | 0 |
| 447 | 2.003 | -0.673 | 0.177 | 28 | 12 | 5 | 41 | 0.028 | 0.012 | 0.005 | 0.041 |
| 448 | 0.906 | 1.706 | 0.248 | 2 | 9 | 0 | 0 | 0.002 | 0.009 | 0 | 0 |
| 449 | 0.703 | -0.32 | 0.282 | 76 | 5 | 5 | 6 | 0.076 | 0.005 | 0.005 | 0.006 |
| 450 | 1.335 | 1.077 | 0.122 | 21 | 3 | 0 | 10 | 0.021 | 0.003 | 0 | 0.01 |
| 451 | 0.618 | 0.045 | 0.277 | 5 | 4 | 3 | 5 | 0.005 | 0.004 | 0.003 | 0.005 |
| 452 | 0.762 | 0.257 | 0.176 | 43 | 7 | 3 | 9 | 0.043 | 0.007 | 0.003 | 0.009 |
| 453 | 1.365 | 0.132 | 0.265 | 176 | 4 | 5 | 5 | 0.176 | 0.004 | 0.005 | 0.005 |
| 454 | 0.51 | -1.8 | 0.284 | 6 | 0 | 0 | 0 | 0.006 | 0 | 0 | 0 |
| 455 | 1.077 | -0.206 | 0.203 | 20 | 8 | 3 | 6 | 0.02 | 0.008 | 0.003 | 0.006 |
| 456 | 2.465 | -1.479 | 0.04 | 10 | 136 | 98 | 71 | 0.01 | 0.136 | 0.098 | 0.071 |
| 457 | 2.03 | -0.053 | 0.067 | 29 | 9 | 62 | 120 | 0.029 | 0.009 | 0.062 | 0.12 |
| 458 | 2.204 | 0.567 | 0.131 | 18 | 33 | 39 | 78 | 0.018 | 0.033 | 0.039 | 0.078 |
| 459 | 2.162 | 0.383 | 0.179 | 26 | 28 | 9 | 70 | 0.026 | 0.028 | 0.009 | 0.07 |
| 460 | 0.728 | 0.42 | 0.273 | 114 | 4 | 4 | 4 | 0.114 | 0.004 | 0.004 | 0.004 |
| 461 | 2.036 | 0.888 | 0.269 | 34 | 9 | 9 | 28 | 0.034 | 0.009 | 0.009 | 0.028 |
| 462 | 2.105 | -2.369 | 0.095 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 463 | 1.309 | -1.688 | 0.122 | 10 | 0 | 0 | 3 | 0.01 | 0 | 0 | 0.003 |
| 464 | 0.848 | -0.296 | 0.068 | 64 | 5 | 5 | 4 | 0.064 | 0.005 | 0.005 | 0.004 |
| 465 | 1.272 | -1.491 | 0.267 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 466 | 1.318 | -2.384 | 0.282 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0.001 |
| 467 | 0.523 | -1.711 | 0.296 | 11 | 0 | 0 | 0 | 0.011 | 0 | 0 | 0 |
| 468 | 1.994 | 1.905 | 0.049 | 9 | 0 | 0 | 4 | 0.009 | 0 | 0 | 0.004 |
| 469 | 1.332 | -1.198 | 0.014 | 13 | 5 | 0 | 101 | 0.013 | 0.005 | 0 | 0.101 |

ตารางที่ ค-1 (ต่อ)

| ข้อที่ | ค่าพารามิเตอร์ของข้อสอบ | | | จำนวนการใช้ข้อสอบ (ครั้ง) | | | | อัตราการใช้ข้อสอบ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | MIC | RDM | HC | HC-Ex | MIC | RDM | HC | HC-Ex |
| 470 | 1.895 | 0.102 | 0.173 | 36 | 10 | 6 | 45 | 0.036 | 0.01 | 0.006 | 0.045 |
| 471 | 0.775 | -2.003 | 0.258 | 5 | 0 | 0 | 0 | 0.005 | 0 | 0 | 0 |
| 472 | 0.985 | -2.458 | 0.21 | 2 | 0 | 0 | 0 | 0.002 | 0 | 0 | 0 |
| 473 | 1.17 | 1.081 | 0.049 | 4 | 0 | 0 | 2 | 0.004 | 0 | 0 | 0.002 |
| 474 | 0.866 | -0.155 | 0.211 | 60 | 9 | 6 | 4 | 0.06 | 0.009 | 0.006 | 0.004 |
| 475 | 0.916 | -0.892 | 0.259 | 28 | 7 | 11 | 2 | 0.028 | 0.007 | 0.011 | 0.002 |
| 476 | 0.571 | 1.571 | 0.033 | 13 | 0 | 1 | 2 | 0.013 | 0 | 0.001 | 0.002 |
| 477 | 1.259 | 2.128 | 0.093 | 8 | 1 | 0 | 0 | 0.008 | 0.001 | 0 | 0 |
| 478 | 0.772 | 1.864 | 0.147 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 479 | 1.08 | -0.454 | 0.038 | 32 | 14 | 3 | 2 | 0.032 | 0.014 | 0.003 | 0.002 |
| 480 | 2.214 | -0.91 | 0.181 | 44 | 8 | 7 | 43 | 0.044 | 0.008 | 0.007 | 0.043 |
| 481 | 2.244 | 1.582 | 0.122 | 7 | 69 | 27 | 11 | 0.007 | 0.069 | 0.027 | 0.011 |
| 482 | 0.693 | -2.237 | 0.178 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 483 | 0.589 | 2.01 | 0.078 | 11 | 0 | 0 | 0 | 0.011 | 0 | 0 | 0 |
| 484 | 1.372 | 2.291 | 0.064 | 13 | 0 | 0 | 0 | 0.013 | 0 | 0 | 0 |
| 485 | 2.102 | -1.425 | 0.017 | 3 | 0 | 4 | 36 | 0.003 | 0 | 0.004 | 0.036 |
| 486 | 1.782 | -1.689 | 0.046 | 2 | 0 | 0 | 6 | 0.002 | 0 | 0 | 0.006 |
| 487 | 2.362 | -1.268 | 0.245 | 10 | 0 | 18 | 20 | 0.01 | 0 | 0.018 | 0.02 |
| 488 | 1.278 | -1.335 | 0.283 | 8 | 0 | 0 | 0 | 0.008 | 0 | 0 | 0 |
| 489 | 1.796 | -0.368 | 0.186 | 90 | 3 | 3 | 24 | 0.09 | 0.003 | 0.003 | 0.024 |
| 490 | 1.298 | 2.271 | 0.018 | 13 | 0 | 0 | 0 | 0.013 | 0 | 0 | 0 |
| 491 | 1.82 | 2.498 | 0.295 | 16 | 0 | 0 | 0 | 0.016 | 0 | 0 | 0 |
| 492 | 1.22 | 0.453 | 0.044 | 43 | 9 | 8 | 68 | 0.043 | 0.009 | 0.008 | 0.068 |
| 493 | 1.429 | -2.197 | 0.164 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 494 | 1.131 | 1.477 | 0.145 | 14 | 0 | 0 | 0 | 0.014 | 0 | 0 | 0 |
| 495 | 2.291 | -1.092 | 0.166 | 16 | 6 | 16 | 18 | 0.016 | 0.006 | 0.016 | 0.018 |
| 496 | 1.325 | 0.352 | 0.053 | 38 | 5 | 8 | 272 | 0.038 | 0.005 | 0.008 | 0.272 |
| 497 | 0.907 | -1.943 | 0.285 | 1 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 |
| 498 | 0.84 | 2.161 | 0.188 | 12 | 1 | 0 | 0 | 0.012 | 0.001 | 0 | 0 |
| 499 | 0.918 | -0.101 | 0.031 | 91 | 4 | 4 | 9 | 0.091 | 0.004 | 0.004 | 0.009 |
| 500 | 1.841 | -1.537 | 0.083 | 7 | 0 | 0 | 0 | 0.007 | 0 | 0 | 0 |
| *Mean* | 1.479 | -0.101 | 0.149 | 11.43 | 7.27 | 7.56 | 9.95 | | | | |
| *SD* | 0.575 | 1.455 | 0.086 | 3.20 | 1.61 | 2.87 | 2.86 | | | | |
| จำนวนข้อสอบที่มีอัตราการใช้ข้อสอบ > 0.2 | | | | 0 | 10 | 11 | 3 | | | | |

**การคำนวณอัตราการใช้ข้อสอบ**

การคำนวณอัตราการใช้ข้อสอบ ดำเนินการโดยหาอัตราส่วนระหว่างจำนวนการใช้ข้อสอบหารด้วยจำนวนผู้สอบทั้งหมด โดยสามารถคำนวณจากสมการที่ 34 ได้ดังนี้

$$\text{อัตราการใช้ข้อสอบ} = \frac{\text{จำนวนการใช้ข้อสอบ}}{\text{จำนวนผู้สอบทั้งหมด}}$$

ยกตัวอย่างการคำนวณอัตราการใช้ข้อสอบ ของข้อสอบข้อที่ 1 ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (MIC) ซึ่งข้อสอบข้อที่ 1 ถูกนำไปใช้ในการทดสอบทั้งหมด 21 ครั้ง จากผู้สอบทั้งหมด 1,000 คน ดังนั้น อัตราการใช้ข้อสอบของข้อสอบข้อที่ 1 คือ $\frac{21}{1{,}000} = 0.021$

# ภาคผนวก ง
# คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี

# คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี

# User Guide for Computerized Adaptive Testing English for Study Skills in Bachelor Degree

สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
สิงหาคม 2560
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา

# คำนำ

คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี เล่มนี้แบ่งออกเป็น 2 ส่วน คือ 1) การจัดการข้อสอบ เป็นส่วนที่ให้ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการข้อสอบ นำข้อสอบที่ผ่านการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) มาใส่ในโปรแกรมเพื่อใช้เป็นคลังข้อสอบ รวมทั้งการลบ และแก้ไขข้อสอบเดิมที่อยู่ในคลังข้อสอบ นอกจากนี้ยังมีการรายงานผลการทดสอบของผู้สอบทั้งหมดเพื่อให้ครู อาจารย์ นำผลการทดสอบไปใช้ประเมินผลการเรียนรู้ต่อไป 2) การจัดการสอบ เป็นส่วนที่ให้ผู้สอบ คือ นิสิต นักศึกษา ใช้โปรแกรมนี้เพื่อทดสอบวัดความสามารถทางการอ่านภาษาอังกฤษ มีการรายงานผลการทดสอบทันทีที่ผู้สอบทำการทดสอบเสร็จสิ้น

คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เล่มนี้นอกจากจะใช้ประกอบการใช้โปรแกรมแล้ว ยังเหมาะสมกับผู้ที่สนใจพัฒนาโปรแกรมเกี่ยวกับการวัดผลอีกด้วย และหากคู่มือการใช้โปรแกรมเล่มนี้มีข้อผิดพลาดประการใด ผู้วิจัยต้องขออภัยมา ณ ที่นี้ด้วย

ประพล เปรมทองสุข

# สารบัญ


| เรื่อง | หน้า |
|---|---|
| 1 คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ | 1 |
| วัตถุประสงค์ของโปรแกรม | 1 |
| ประโยชน์ที่คาดว่าจะได้รับ | 1 |
| 2 วิธีการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ | 2 |
| ส่วนที่ 1 การจัดการข้อสอบ | 4 |
| การตั้งรหัสผู้สอบ | 5 |
| การจัดการข้อสอบ | 5 |
| การรายงานผลสอบ | 10 |
| การออกจากระบบ | 11 |
| ส่วนที่ 2 การจัดการสอบ | 11 |
| เริ่มต้นการทดสอบ | 11 |
| กรอกข้อมูลผู้สอบและเลือกแบบทดสอบ | 12 |
| การดำเนินการสอบ | 13 |
| การรายงานผลสอบรายบุคคล | 14 |

## สารบัญภาพ


| ภาพ | | หน้า |
|---|---|---|
| 1 | หน้าหลักของโปรแกรมการทดสอบ | 3 |
| 2 | หน้าจอ Login เข้าโปรแกรมในส่วนการจัดการข้อสอบ | 4 |
| 3 | หน้าจอการตั้งรหัสผู้สอบ | 5 |
| 4 | หน้าจอการจัดการคลังข้อสอบ | 6 |
| 5 | หน้าจอการเพิ่มข้อสอบ | 6 |
| 6 | หน้าจอแก้ไขข้อสอบ | 8 |
| 7 | หน้าจอการจัดการข้อสอบของการทดสอบด้วยคอมพิวเตอร์ | 9 |
| 8 | หน้าจอแรกในส่วนรายงานผลสอบ | 10 |
| 9 | หน้าจอรายงานผลสอบ | 10 |
| 10 | หน้าจอเริ่มต้นการทดสอบ | 12 |
| 11 | หน้าจอกรอกข้อมูลผู้สอบและเลือกแบบทดสอบ | 13 |
| 12 | หน้าจอข้อสอบ | 14 |
| 13 | หน้าจอรายงานผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ | 14 |
| 14 | หน้าจอรายงานผลการทดสอบด้วยคอมพิวเตอร์ | 16 |

# คู่มือการใช้โปรแกรมการการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
# (User Guide for Computerized Adaptive Testing)

คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี เล่มนี้ แบ่งออกเป็น 2 ส่วน คือ 1) การจัดการข้อสอบเป็นส่วนที่ให้ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการข้อสอบ นำข้อสอบที่ผ่านการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) มาใส่ในโปรแกรมเพื่อใช้เป็นคลังข้อสอบ และการรายงานผลการทดสอบของผู้สอบทั้งหมด เพื่อให้ครู อาจารย์ นำผลการทดสอบของผู้สอบแต่ละคน ไปประเมินผลการเรียนรู้ต่อไป 2) การจัดการสอบ เป็นส่วนที่ให้ผู้สอบ คือ นิสิต นักศึกษา มาใช้โปรแกรมนี้เพื่อทดสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ซึ่งเมื่อผู้สอบทำการทดสอบเสร็จสิ้น จะรายงานผลการทดสอบให้ผู้สอบทราบทันที

**วัตถุประสงค์ของโปรแกรม**

1. เพื่อใช้จัดทำคลังข้อสอบสำหรับใช้ทดสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน
2. เพื่อใช้ทดสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน

**ประโยชน์ที่คาดว่าจะได้รับ**

1. สามารถนำโปรแกรมไปใช้งานตามวัตถุประสงค์ได้อย่างมีประสิทธิภาพ
2. สามารถนำไปเป็นแนวทางในการพัฒนาโปรแกรมการทดสอบผู้เรียนในรายวิชาอื่น ๆ ได้

**วิธีการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นี้ เป็น Web Application ที่พัฒนาขึ้นบนโปรแกรม Notepad++ ร่วมกับการใช้ Tools ของ Cloud9 โดยพัฒนาให้สามารถใช้งานบนอุปกรณ์ที่เชื่อมต่อกับอินเทอร์เน็ตได้ เช่น คอมพิวเตอร์ โน๊ตบุ๊ค หรือสมาร์ทโฟน

โปรแกรมการทดสอบนี้ ถูกออกแบบให้ผู้ใช้งานสามารถเลือกรูปแบบการทดสอบได้ 2 แบบ ดังนี้

1) การทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) เป็นการทดสอบที่นำคอมพิวเตอร์มาใช้แทนการทดสอบบนกระดาษ ผู้สอบทุกคนจะได้รับแบบทดสอบชุดเดียวกัน จำนวน 40 ข้อ โดยคัดเลือกข้อสอบไว้ล่วงหน้าแล้ว ซึ่งประกอบด้วยข้อสอบที่มีค่าความยากของข้อสอบหลายระดับ คละกัน

2) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นการทดสอบด้วยคอมพิวเตอร์ ซึ่งผู้สอบแต่ละคนจะได้รับแบบทดสอบต่างชุดกัน ขึ้นอยู่กับความสามารถของผู้สอบ เริ่มต้นจากข้อสอบที่มีค่าความยากปานกลางก่อน หากผู้สอบตอบถูก ให้จัดข้อสอบข้อถัดไปที่มีค่าความยากสูงขึ้น แต่หากผู้สอบตอบผิด ให้จัดข้อสอบข้อถัดไปที่มีค่าความยากน้อยลง ทำเช่นนี้จนกว่าจะสิ้นสุดการทดสอบ

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนนี้ จะครอบคลุมเนื้อหา 4 ด้าน คือ 1) ชนิดของคำ (Part of Speech) 2) การใช้บริบทบ่งชี้เพื่อค้นหาความหมายของคำศัพท์ยาก (Using Context Clue) 3) การเรียนรู้ความหมายคำศัพท์จากโครงสร้างของคำศัพท์ (Word Formations) และ 4) การสรุปความ (Making Inferences) ซึ่งผู้ใช้งานจะเข้าใช้โปรแกรมได้จาก http://catctexam.com สามารถแสดงหน้าหลักของโปรแกรมการได้ดังภาพที่ 1

3

ภาพที่ 1 หน้าหลักของโปรแกรมการทดสอบ

จากภาพที่ 1 แสดงหน้าหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งในหน้าหลักของโปรแกรม ประกอบด้วยเมนูหลัก 4 เมนู ดังนี้

1) หน้าหลัก เป็นเมนูแสดงหน้าแรกของโปรแกรมการทดสอบ

2) การทดสอบ เป็นเมนูสำหรับผู้สอบ เพื่อทดสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ในระดับปริญญาตรี

3) ติดต่อเรา เป็นเมนูแสดงข้อมูลของวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา และข้อมูลของผู้พัฒนาโปรแกรม

4) เข้าสู่ระบบ เป็นเมนูสำหรับครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการข้อสอบ แสดงการจัดการข้อสอบ รวมทั้งรายงานผลการทดสอบของผู้สอบทั้งหมด

**ส่วนที่ 1 การจัดการข้อสอบ**

การจัดการข้อสอบเป็นส่วนที่ให้ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการข้อสอบ ใส่ข้อสอบที่ผ่านการคัดเลือกเป็นไปตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ลงในโปรแกรมเพื่อเป็นคลังข้อสอบ หรือทำการลบ แก้ไขข้อสอบเดิมที่มีอยู่ในคลังข้อสอบ หลังจากผู้ใช้เลือก “เข้าสู่ระบบ” ในหน้าหลัก โปรแกรมจะเข้าสู่หน้าจอแรกของส่วนการจัดการข้อสอบ แสดงได้ดังภาพที่ 2

คลังข้อสอบที่ใช้ในโปรแกรมนี้ เป็นข้อสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ซึ่งได้รับความอนุเคราะห์ข้อสอบทั้งหมด จากอาจารย์ประจำหลักสูตรภาษาอังกฤษ มหาวิทยาลัยสวนดุสิต โดยข้อสอบทั้งหมดมีจำนวน 421 ข้อ ประกอบด้วย โจทย์ รายการคำตอบแบบ 4 ตัวเลือก และค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ ซึ่งแบ่งข้อสอบได้ 4 ชุด ได้แก่ ชุดที่ 1 ชนิดของคำ (Part of Speech) จำนวน 118 ข้อ ชุดที่ 2 การใช้บริบทบ่งชี้เพื่อค้นหาความหมายของคำศัพท์ยาก (Using Context Clue) จำนวน 104 ข้อ ชุดที่ 3 การเรียนรู้ความหมายคำศัพท์จากโครงสร้างของคำศัพท์ (Word Formations) จำนวน 107 ข้อ และ ชุดที่ 4 การสรุปความ (Making Inferences) จำนวน 92 ข้อ

ภาพที่ 2 หน้าจอ Login เข้าโปรแกรมในส่วนการจัดการข้อสอบ

จากภาพที่ 2 แสดงหน้าจอ Login เข้าส่วนการจัดการข้อสอบ โดยผู้วิจัยกำหนดรหัสผ่านดังนี้

User Name: admin
Password: @dmin

เมื่อผู้ใช้กรอกรหัสผ่านเรียบร้อยแล้ว หน้าจอจะเข้าสู่ส่วนการจัดการข้อสอบ ประกอบด้วยเมนูย่อย ได้แก่ 1) ตั้งรหัสกลุ่มผู้สอบ 2) การจัดการข้อสอบ 3) รายงานผลสอบ และ 4) ออกจากระบบ โดยแต่ละเมนูมีรายละเอียดดังนี้

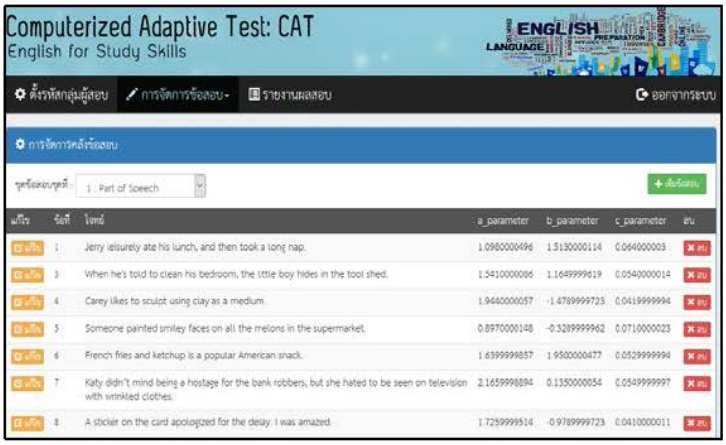
1) ตั้งรหัสกลุ่มผู้สอบ สำหรับให้ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการข้อสอบ ทำการตั้งรหัสกลุ่มผู้สอบใหม่ เพื่อป้องกันไม่ให้บุคคลภายนอกเข้ามาใช้โปรแกรมได้ หรือในกรณีที่มีผู้สอบหลายกลุ่ม ในแต่ละกลุ่มใช้ข้อสอบต่างชุดกัน จึงจำเป็นต้องตั้งรหัสกลุ่มผู้สอบใหม่ เพื่อป้องกันผู้สอบเข้าสอบผิดกลุ่ม เมื่อผู้ใช้ตั้งรหัสใหม่เรียบร้อยแล้วให้กดปุ่ม “บันทึก” ซึ่งแสดงหน้าจอได้ดังภาพที่ 3

ภาพที่ 3 หน้าจอการตั้งรหัสกลุ่มผู้สอบ

2) การจัดการข้อสอบ สำหรับให้ครู อาจารย์ หรือผู้เกี่ยวข้อง ทำการจัดการข้อสอบ ซึ่งมีเมนูย่อย 2 ส่วน ได้แก่

2.1) การจัดการคลังข้อสอบ เป็นการเพิ่ม ลบ และแก้ไขข้อสอบที่อยู่ในคลังข้อสอบ ซึ่งหน้าจอการจัดการคลังข้อสอบ แสดงได้ดังภาพที่ 4

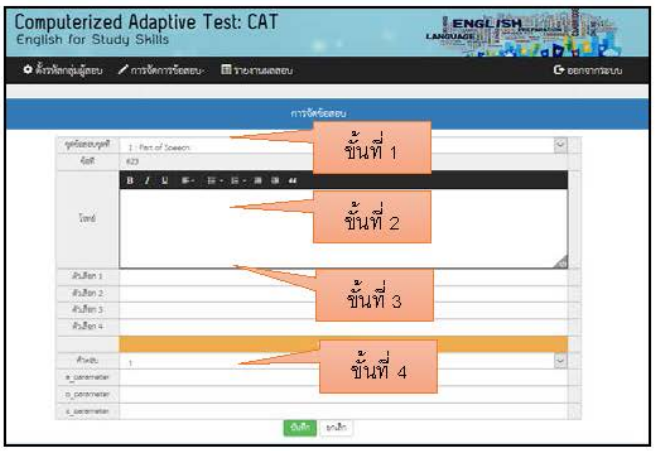
6




ภาพที่ 4 หน้าจอการจัดการคลังข้อสอบ

**2.1.1 การเพิ่มข้อสอบ**

การเพิ่มข้อสอบลงในโปรแกรม โดยให้ผู้ใช้กดปุ่ม “เพิ่มข้อสอบ” แล้วโปรแกรมจะเข้าสู่หน้าจอดังภาพที่ 5




ภาพที่ 5 หน้าจอการเพิ่มข้อสอบ

จากภาพที่ 5 แสดงหน้าจอการเพิ่มข้อสอบลงในโปรแกรม โดยข้อสอบที่ใช้ในโปรแกรมนี้ต้องเป็นข้อสอบที่มีรายการคำตอบอย่างน้อย 4 ตัวเลือก และผ่านการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ จึงมีค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ (ผู้สนใจสามารถคำนวณค่าพารามิเตอร์ของข้อสอบได้จากโปรแกรม Xcalibre) มีขั้นตอนการเพิ่มข้อสอบดังนี้

**ขั้นที่ 1 เลือกชุดข้อสอบ**ที่ต้องการเพิ่มข้อสอบเข้าไป

**ขั้นที่ 2 กรอกรายละเอียดโจทย์**

**ขั้นที่ 3 กรอกรายละเอียดรายการคำตอบ**ที่ละรายการ ในกรณีมีรายการคำตอบมากกว่า 4 รายการให้กดปุ่ม “เพิ่มตัวเลือก” พร้อมระบุรายการคำตอบที่เป็นคำตอบ

**ขั้นที่ 4 กรอกรายละเอียดค่าพารามิเตอร์ของข้อสอบ** ได้แก่ ค่าอำนาจจำแนกของข้อสอบ (Discrimination Power Parameter: a_parameter) ต้องมีค่าระหว่าง -2.50 ถึง 2.50 ค่าความยากของข้อสอบ (Difficulty Parameter: b_parameter) ต้องมีค่าตั้งแต่ 0.50 ถึง 2.50 และค่าการเดาของข้อสอบ (Guess Parameter: c_parameter) ต้องมีค่าไม่เกิน 0.30

เมื่อกรอกรายละเอียดข้อสอบที่ต้องการเพิ่มเรียบร้อยแล้ว ให้กดปุ่ม “บันทึกข้อมูล” แต่หากไม่ต้องการบันทึก ให้กดปุ่ม “ยกเลิก”

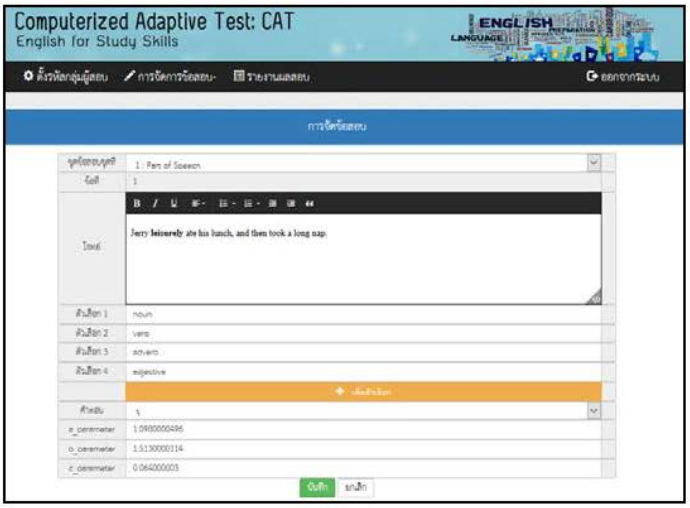
**2.1.2 การลบข้อสอบ**

การลบข้อสอบออกจากคลังข้อสอบที่อยู่ในโปรแกรม ดำเนินการโดยไปหน้าจอการจัดการคลังข้อสอบ ดังภาพที่ 4 จากนั้นให้เลือกชุดของข้อสอบที่ต้องการลบข้อสอบออก แล้วทำการกดปุ่ม “ลบ” ซึ่งปรากฏอยู่ด้านท้ายของข้อสอบที่ต้องการลบ

**2.1.3 การแก้ไขข้อสอบ**

การแก้ไขข้อสอบเดิมในคลังข้อสอบ ดำเนินการโดยไปที่หน้าการจัดการคลังข้อสอบ ดังภาพที่ 4 จากนั้นให้เลือกชุดข้อสอบที่ต้องการแก้ไขข้อสอบ

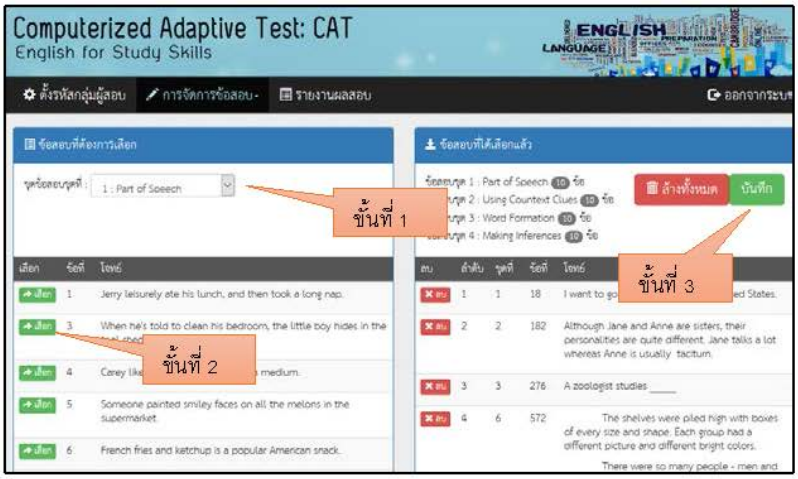
แล้วทำการกดปุ่ม “แก้ไข” ซึ่งปรากฏอยู่ด้านหน้าของข้อสอบที่ต้องการแก้ไข แล้วจะได้หน้าจอ ดังภาพที่ 6




ภาพที่ 6 หน้าจอแก้ไขข้อสอบ

จากภาพที่ 6 แสดงหน้าจอแก้ไขข้อสอบ เมื่อผู้ใช้แก้ไขข้อสอบตามที่ต้องการเรียบร้อยแล้ว ให้กดปุ่ม “บันทึก” หากไม่ต้องการบันทึกให้กดปุ่ม “ยกเลิก”

**2.2 การจัดการข้อสอบแบบ CT** เป็นการจัดชุดข้อสอบในการทดสอบด้วยคอมพิวเตอร์ (Computerized Testing) ซึ่งผู้สอบทุกคนจะได้รับข้อสอบชุดเดียวกัน ดำเนินการโดยให้ผู้ใช้ (ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดข้อสอบ) คัดเลือกข้อสอบจำนวน 40 ข้อ จากคลังข้อสอบ โดยคัดเลือกข้อสอบชุดละ 10 ข้อ จากข้อสอบทั้งหมด 4 ชุด ประกอบด้วยข้อสอบที่มีค่าความยากของข้อสอบจากน้อยไปมาก ในการจัดการชุดข้อสอบแบบ CT แสดงดังภาพที่ 7




ภาพที่ 7 หน้าจอการจัดการข้อสอบของการทดสอบด้วยคอมพิวเตอร์

จากภาพที่ 7 แสดงหน้าจอการจัดการข้อสอบแบบ CT โดยให้เลือกข้อสอบ จำนวน 40 ข้อ จากคลังข้อสอบ ซึ่งมีขั้นตอนดังนี้

**ขั้นที่ 1 เลือกชุดข้อสอบ**

**ขั้นที่ 2 เลือกข้อสอบ** โดยกดปุ่ม “เลือก” ที่อยู่ด้านหน้าของข้อสอบข้อที่ต้องการ โดยหลังจากกดเลือกแล้ว ข้อสอบข้อนั้นจะไปปรากฏด้านล่างสุดในตารางด้านขวา แสดงว่า ข้อสอบข้อนั้นได้ถูกเลือกไปใช้ในการทดสอบแบบ CT แล้ว จากนั้นดำเนินการเลือกข้อสอบในชุดอื่นให้ครบตามที่ต้องการ

ในกรณีที่เลือกข้อสอบผิดข้อ ผู้ใช้สามารถลบโดยการกดปุ่ม “ลบ” หน้าข้อสอบที่ต้องการลบได้ หรือกดปุ่ม “ล้างทั้งหมด” ซึ่งเป็นการลบข้อสอบทุกข้อที่ได้เลือกมาแล้ว

**ขั้นที่ 3 บันทึกข้อมูล** เมื่อเลือกข้อสอบได้ครบตามต้องการแล้ว กดปุ่ม “บันทึก” เพื่อบันทึกข้อสอบทั้งหมดในการทดสอบแบบ CT

**3. การรายงานผลสอบ**

การรายงานผลสอบ เป็นการรายงานผลสอบของผู้สอบทุกคนให้ครู อาจารย์ทราบเพื่อใช้ในการประเมินผลการเรียนรู้ของผู้สอบ โดยกดปุ่ม “รายงานผลสอบ” จะปรากฏหน้าจอดังภาพที่ 8

ภาพที่ 8 หน้าจอแรกในส่วนรายงานผลสอบ

จากภาพที่ 8 แสดงหน้าจอแรกในส่วนรายงานผลสอบ โดยให้ระบุช่วงวันแล้วกดปุ่ม “ค้นหา” โปรแกรมจะแสดงผลสอบของผู้สอบทั้งหมดที่สอบภายในช่วงวันที่ระบุ ดังภาพที่ 9

ภาพที่ 9 หน้าจอรายงานผลสอบ

จากภาพที่ 9 แสดงหน้าจอรายงานผลสอบของผู้สอบทั้งหมดที่สอบภายในวันที่ผู้ใช้ระบุ โดยโปรแกรมจะแสดงข้อมูลเป็นรายคน ได้แก่ รหัสนักศึกษา ชื่อ-สกุล วันที่เข้าสอบ คะแนนของผู้สอบในแต่ละหัวข้อ คะแนนรวม ค่าประมาณความสามารถ และระดับความสามารถของผู้สอบ

นอกจากนี้ หากต้องการทราบรายละเอียดในการสอบของผู้สอบแต่ละคน ให้กดปุ่ม ในช่องรายละเอียด ซึ่งอยู่ด้านท้ายของผู้สอบที่ต้องการทราบ โปรแกรมจะเชื่อมข้อมูลไปยังรายงานผลการสอบของผู้สอบคนนั้น

**4. การออกจากระบบ**

ออกจากระบบ เป็นเมนูสุดท้ายในส่วนการเข้าระบบ เมื่อผู้ใช้กดปุ่ม “ออกจากระบบ” โปรแกรมจะกลับไปสู่หน้าจอหลักของโปรแกรม

**ส่วนที่ 2 การจัดการสอบ**

การจัดการสอบนี้ เป็นส่วนที่ให้นิสิต นักศึกษา เข้าใช้โปรแกรมเพื่อทดสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ในระดับปริญญาตรี โดยผู้สอบจะได้รับข้อสอบครบทุกหัวข้อ ซึ่งวิธีการคัดเลือกข้อสอบมาให้ผู้สอบนั้น ขึ้นอยู่กับแบบทดสอบที่ผู้สอบเลือก มีทั้งสิ้น 2 แบบ คือ การทดสอบด้วยคอมพิวเตอร์ และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ขั้นตอนการทดสอบมีดังต่อไปนี้

**1. เริ่มต้นการทดสอบ**

เมื่อผู้สอบต้องการเริ่มทดสอบ ให้เลือก “การทดสอบ” ซึ่งเป็นเมนูหนึ่งในหน้าหลักของโปรแกรม หลังจากนั้นจะปรากฏหน้าจอดังภาพที่ 10

ภาพที่ 10 หน้าจอเริ่มต้นการทดสอบ

จากภาพที่ 10 แสดงหน้าจอเริ่มต้นการทดสอบ ซึ่งจะแสดงคำชี้แจงในการสอบให้ผู้สอบทราบ และให้ผู้สอบกรอกรหัสผ่านเพื่อเข้าสอบ ซึ่งผู้วิจัยกำหนดรหัสผ่านเบื้องต้นไว้ ดังนี้

ผู้ใช้: test
รหัสผ่าน: test

โดยรหัสผ่านนี้กำหนดเพื่อป้องกันปัญหาเกี่ยวกับการทุจริตการสอบ ป้องกันไม่ให้บุคคลภายนอก หรือนักศึกษากลุ่มอื่น เข้าดูข้อสอบได้ ซึ่งผู้ใช้งาน ครู อาจารย์ หรือผู้เกี่ยวข้องในการจัดการทดสอบ สามารถเปลี่ยนรหัสผ่านได้จาก ตั้งรหัสผู้สอบ แสดงตามภาพที่ 2

**2. กรอกข้อมูลผู้สอบและเลือกแบบทดสอบ**

หลังจากผู้สอบกรอกรหัสผ่านเพื่อเข้าสอบเรียบร้อยแล้ว จะปรากฏหน้าจอดังภาพที่ 11

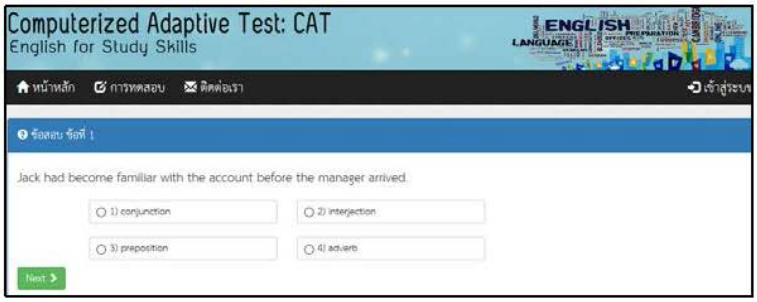
13

ภาพที่ 11 หน้าจอกรอกข้อมูลผู้สอบ และเลือกแบบทดสอบ

จากภาพที่ 11 แสดงหน้ากรอกข้อมูลผู้สอบ โดยให้ผู้สอบกรอกรหัสนักศึกษา ชื่อ-สกุล คณะ และสาขาวิชา ของตนเองลงไป พร้อมทั้งเลือกแบบทดสอบ ซึ่งในโปรแกรมนี้กำหนดแบบทดสอบให้ผู้สอบเลือก 2 แบบ ได้แก่ 1) การทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) และ 2) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) โดยผู้สอบต้องเลือกแบบทดสอบแบบใดแบบหนึ่งเท่านั้น หลักจากนั้นให้กดปุ่ม “เริ่มทดสอบ” เพื่อดำเนินการสอบต่อไป

**3. การดำเนินการสอบ**

หลังจากผู้สอบเลือกแบบทดสอบเรียบร้อยแล้ว จะเข้าสู่การดำเนินการสอบ โดยจะปรากฏหน้าจอดังภาพที่ 12




ภาพที่ 12 หน้าจอข้อสอบ

จากภาพที่ 12 แสดงหน้าจอข้อสอบ โดยโปรแกรมจะคัดเลือกข้อสอบมาให้ผู้สอบ ขึ้นอยู่กับแบบทดสอบที่ผู้สอบเลือก หลังจากนั้นให้ผู้สอบเลือกคำตอบ แล้วกดปุ่ม “Next” เพื่อรับข้อสอบข้อถัดไป ซึ่งการดำเนินการทดสอบจะดำเนินต่อไปจนสิ้นสุดการทดสอบ

**4. รายงานผลสอบรายบุคคล**

หลังจากผู้สอบดำเนินการทดสอบเสร็จสิ้นแล้ว โปรแกรมจะรายงานผลการทดสอบให้ผู้สอบทราบทันที ซึ่งการรายงานผลการทดสอบมี 2 แบบ ขึ้นอยู่กับแบบทดสอบที่ผู้สอบเลือก (CAT หรือ CT) สามารถแสดงได้ดังภาพที่ 13 และ 14

ภาพที่ 13 หน้าจอรายงานผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

จากภาพที่ 13 แสดงหน้าจอรายงานผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยแสดงรายละเอียดของผู้สอบ ได้แก่ ชื่อ-สกุล (Student Name) รหัสนักศึกษา (Student No) คณะ (Faculty) และสาขาวิชา (Major) และแสดงรายละเอียดเกี่ยวกับการสอบ ได้แก่ วันที่สอบ (Date) เวลาที่ใช้ในการสอบ (Time) จำนวนข้อสอบที่ใช้ในการทดสอบ (Test Length) และจำนวนข้อสอบที่ตอบถูก (Correct) ค่าประมาณความสามารถของผู้สอบ (Ability Estimate) และระดับความสามารถของผู้สอบ (Ability Level) พร้อมทั้งแสดงตารางการแจกแจงจำนวนข้อสอบที่ตอบถูก และค่าร้อยละ จำแนกตามชุดข้อสอบ ทั้งนี้เกณฑ์การประเมินระดับความสามารถของผู้สอบมีรายละเอียดดังตารางที่ 1

ตารางที่ 1 เกณฑ์ประเมินระดับความสามารถ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

| ค่าประมาณความสามารถของผู้สอบ | ระดับความสามารถ |
|---|---|
| มากกว่า 2.0000 | Very High |
| 1.0001 ถึง 2.0000 | High |
| 0.5001 ถึง 1.0000 | Rather High |
| -0.4999 ถึง 0.5000 | Normal |
| -1.4999 ถึง -0.5000 | Rather Low |
| -2.0000 ถึง -1.5000 | Low |
| ต่ำกว่า 2.0000 | Very Low |

16

ภาพที่ 14 หน้าจอรายงานผลการทดสอบด้วยคอมพิวเตอร์

จากภาพที่ 14 แสดงหน้าจอรายงานผลการทดสอบด้วยคอมพิวเตอร์ โดยแสดงรายละเอียดของผู้สอบ และรายละเอียดเกี่ยวกับการสอบ เหมือนกับรายงานผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แต่แตกต่างตรงผลการทดสอบด้วยคอมพิวเตอร์ จะพิจารณาจากคะแนนรวมของผู้สอบ ซึ่งเป็นจำนวนข้อสอบที่ตอบถูก และปรับเทียบระดับความสามารถ โดยใช้เกณฑ์ดังนี้

ตารางที่ 2 เกณฑ์เทียบระดับความสามารถ สำหรับการทดสอบด้วยคอมพิวเตอร์

| คะแนนรวม (ร้อยละ) | จำนวนข้อสอบที่ตอบถูก | ระดับความสามารถ |
|---|---|---|
| ตั้งแต่ 80% ขึ้นไป | ตั้งแต่ 32 ข้อ ขึ้นไป | Very High |
| 70-79% | 28-31 ข้อ | High |
| 60-69% | 24-27 ข้อ | Rather High |
| 50-59% | 20-23 ข้อ | Normal |
| 40-49% | 16-19 ข้อ | Rather Low |
| 30-39% | 12-15 ข้อ | Low |
| ต่ำกว่า 30% | ต่ำกว่า 12 ข้อ | Very Low |

# ภาคผนวก จ
# แบบประเมินโปรแกรมและคู่มือการใช้โปรแกรม
# การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับผู้เชี่ยวชาญ

**แบบประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี (สำหรับผู้เชี่ยวชาญ)**

**คำชี้แจง**

แบบประเมินโปรแกรมและคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ฉบับนี้ มีจุดมุ่งหมายเพื่อทราบความคิดเห็นที่มีต่อการใช้งานโปรแกรมการทดสอบ รวมทั้งคู่มือการใช้โปรแกรม เพื่อนำผลที่ได้มาเป็นข้อมูลในการพัฒนาโปรแกรมต่อไป

**ตอนที่ 1** การประเมินโปรแกรมและคู่มือการใช้โปรแกรม

โปรดทำเครื่องหมาย ✓ ในช่องที่ตรงกับความคิดเห็นของท่าน หลังจากท่านได้ศึกษาคู่มือการใช้โปรแกรม และทดลองใช้โปรแกรมแล้ว โดยเกณฑ์ประเมินแบ่งได้ 5 ระดับดังนี้

ระดับ 5 หมายถึง โปรแกรมมีความเหมาะสมมากที่สุด
ระดับ 4 หมายถึง โปรแกรมมีความเหมาะสมมาก
ระดับ 3 หมายถึง โปรแกรมมีความเหมาะสมปานกลาง
ระดับ 2 หมายถึง โปรแกรมมีความเหมาะสมน้อย
ระดับ 1 หมายถึง โปรแกรมมีความเหมาะสมน้อยที่สุด

| รายการประเมิน | ระดับความเหมาะสม | | | | |
|---|---|---|---|---|---|
| | 5 | 4 | 3 | 2 | 1 |
| 1. ความสะดวกในการใช้โปรแกรม | | | | | |
| 1.1 การเข้าถึงโปรแกรมทำได้ง่ายและสะดวก | | | | | |
| 1.2 โปรแกรมมีเมนูที่เข้าถึงส่วนต่าง ๆ ของโปรแกรมได้ง่ายและสะดวก | | | | | |
| 1.3 การสอบโดยใช้โปรแกรมมีความสะดวกกว่าเมื่อเทียบกับการสอบบนกระดาษ | | | | | |
| 1.4 โปรแกรมมีการแสดงผลการสอบทันที เมื่อทดสอบเสร็จ | | | | | |
| 1.5 การแสดงผลสอบมีรายละเอียดครบถ้วน | | | | | |
| 2. ความถูกต้องในการใช้งาน | | | | | |
| 2.1 โปรแกรมสามารถเพิ่ม ลบ หรือแก้ไข ข้อสอบได้ | | | | | |
| 2.2 โปรแกรมสามารถจัดการทดสอบได้ตรงตามวัตถุประสงค์ หรือเงื่อนไขของการทดสอบได้ | | | | | |
| 2.3 โปรแกรมสามารถประมาณค่าความสามารถของผู้สอบถูกต้อง | | | | | |
| 2.4 โปรแกรมสามารถบันทึกผลการทดสอบได้ถูกต้อง | | | | | |
| 3. ลักษณะทั่วไปของโปรแกรม | | | | | |
| 3.1 การออกแบบโปรแกรมมีความน่าสนใจ | | | | | |
| 3.2 การจัดรูปแบบหน้าจอง่ายต่อการใช้งาน | | | | | |
| 3.3 การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว | | | | | |

| รายการประเมิน | ระดับความเหมาะสม | | | | |
|---|---|---|---|---|---|
| | 5 | 4 | 3 | 2 | 1 |
| 3.4 โปรแกรมมีระบบป้องกันการทำงานผิดพลาดของผู้ใช้ทุกขั้นตอน เช่น มีข้อเสนอแนะปรากฏขึ้นเมื่อใช้เมาส์ชี้ หรือเมื่อผู้ใช้ ใช้งานผิด | | | | | |
| 4. ความชัดเจนของคู่มือการใช้โปรแกรม | | | | | |
| 4.1 คู่มือมีคำอธิบายวัตถุประสงค์ของโปรแกรมได้อย่างชัดเจน | | | | | |
| 4.2 คู่มือการใช้โปรแกรมแสดงวิธีการใช้งานอย่างมีลำดับขั้นตอน | | | | | |
| 4.3 ภาษาที่ใช้มีความเข้าใจง่าย | | | | | |
| 4.4 ภาพประกอบสามารถอธิบายขั้นตอนการใช้โปรแกรมได้อย่างชัดเจน | | | | | |
| 4.5 หลังจากอ่านคู่มือแล้ว ผู้ใช้มีความมั่นใจว่าสามารถใช้โปรแกรมได้ | | | | | |

**ตอนที่ 2** ข้อเสนอแนะอื่นๆ

.....

.....

.....

.....

.....

# ภาคผนวก ฉ
# แบบประเมินโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
# สำหรับผู้สอบ

**แบบประเมินโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**
**รายวิชาภาษาอังกฤษเพื่อทักษะการเรียน ระดับปริญญาตรี**
**(สำหรับผู้สอบ)**

**คำชี้แจง**

แบบประเมินโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ฉบับนี้ มีจุดมุ่งหมายเพื่อทราบความคิดเห็นที่มีต่อการใช้งานโปรแกรมการทดสอบ รวมทั้งคู่มือการใช้โปรแกรม เพื่อนำผลที่ได้มาเป็นข้อมูลในการพัฒนาโปรแกรมต่อไป

**ตอนที่ 1** การประเมินโปรแกรม

โปรดทำเครื่องหมาย ✓ ในช่องที่ตรงกับความคิดเห็นของท่าน หลังจากท่านได้ศึกษาคู่มือการใช้โปรแกรม และทดลองใช้โปรแกรมแล้ว โดยเกณฑ์ประเมินแบ่งได้ 5 ระดับดังนี้

ระดับ 5 หมายถึง โปรแกรมมีความเหมาะสมมากที่สุด
ระดับ 4 หมายถึง โปรแกรมมีความเหมาะสมมาก
ระดับ 3 หมายถึง โปรแกรมมีความเหมาะสมปานกลาง
ระดับ 2 หมายถึง โปรแกรมมีความเหมาะสมน้อย
ระดับ 1 หมายถึง โปรแกรมมีความเหมาะสมน้อยที่สุด

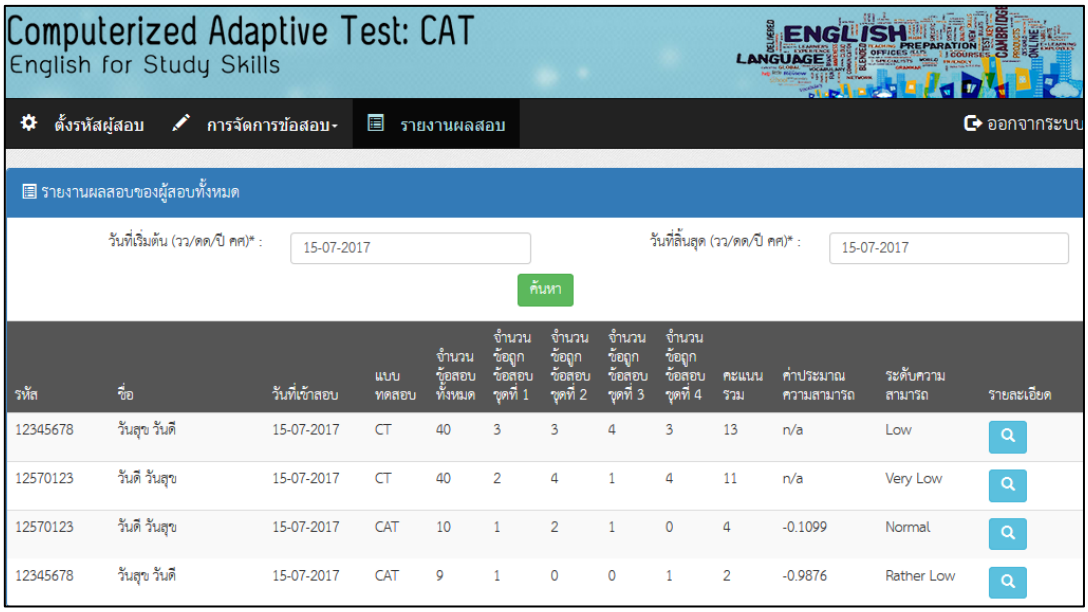
| รายการประเมิน | ระดับความเหมาะสม | | | | |
|---|---|---|---|---|---|
| | 5 | 4 | 3 | 2 | 1 |
| 1. ความสะดวกในการใช้โปรแกรม | | | | | |
| 1.1 การเข้าถึงโปรแกรมทำได้ง่ายและสะดวก | | | | | |
| 1.2 โปรแกรมมีเมนูที่เข้าถึงส่วนต่าง ๆ ของโปรแกรมได้ง่ายและสะดวก | | | | | |
| 1.3 การสอบโดยใช้โปรแกรมมีความสะดวกกว่าเมื่อเทียบกับการสอบบนกระดาษ | | | | | |
| 1.4 โปรแกรมมีการแสดงผลการสอบทันที เมื่อทดสอบเสร็จ | | | | | |
| 1.5 การแสดงผลสอบมีรายละเอียดครบถ้วน | | | | | |
| 2. ลักษณะทั่วไปของโปรแกรม | | | | | |
| 2.1 การออกแบบโปรแกรมมีความน่าสนใจ | | | | | |
| 2.2 การจัดรูปแบบหน้าจอง่ายต่อการใช้งาน | | | | | |
| 2.3 การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว | | | | | |
| 2.4 โปรแกรมมีระบบป้องกันการทำงานผิดพลาดของผู้ใช้ทุกขั้นตอน เช่น มีข้อเสนอแนะปรากฏขึ้นเมื่อใช้เมาส์ชี้ หรือเมื่อผู้ใช้ ใช้งานผิด | | | | | |

**ตอนที่ 2** ข้อเสนอแนะอื่นๆ

.....
.....
.....
.....

# ภาคผนวก ช
# ตัวอย่างการแสดงผลการทดสอบจากโปรแกรม การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

# ตัวอย่างการแสดงผลการทดสอบจากโปรแกรม
# การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์




| รหัส | ชื่อ | วันที่เข้าสอบ | แบบทดสอบ | จำนวนข้อสอบทั้งหมด | จำนวนข้อถูกข้อสอบชุดที่ 1 | จำนวนข้อถูกข้อสอบชุดที่ 2 | จำนวนข้อถูกข้อสอบชุดที่ 3 | จำนวนข้อถูกข้อสอบชุดที่ 4 | คะแนนรวม | ค่าประมาณความสามารถ | ระดับความสามารถ | รายละเอียด |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12345678 | วันสุข วันดี | 15-07-2017 | CT | 40 | 3 | 3 | 4 | 3 | 13 | n/a | Low | |
| 12570123 | วันดี วันสุข | 15-07-2017 | CT | 40 | 2 | 4 | 1 | 4 | 11 | n/a | Very Low | |
| 12570123 | วันดี วันสุข | 15-07-2017 | CAT | 10 | 1 | 2 | 1 | 0 | 4 | -0.1099 | Normal | |
| 12345678 | วันสุข วันดี | 15-07-2017 | CAT | 9 | 1 | 0 | 0 | 1 | 2 | -0.9876 | Rather Low | |

ภาพที่ ช-1 ตัวอย่างรายงานผลสอบของผู้สอบทั้งหมด

| ชุดที่ | ชื่อชุด | จำนวนข้อทั้งหมด | จำนวนข้อถูก | ร้อยละ |
|---|---|---|---|---|
| 1 | Part of Speech | 3 | 1 | 33.33 |
| 2 | Using Context Clues | 2 | 2 | 100 |
| 3 | Word Formation | 3 | 1 | 33.33 |
| 4 | Making Inferences | 2 | 0 | 0 |
| | รวม | 10 | 4 | 40 |

ภาพที่ ช-2 ตัวอย่างผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

English Reading Skills Report

Computerized Testing : CT

| | | | |
|---|---|---|---|
| Student Name | : วันสุข วันดี | Student No | : 12345678 |
| Faculty | : Management Science | Major | : Business Management and English |
| Date | : 15 Jul 2017 | Time | : 0 Hours 01 Minutes 16 Seconds |
| Test Length | : 40 items | Correct | : 13 |
| | | Ability Level | : Low |

| ชุดที่ | ชื่อชุด | จำนวนข้อทั้งหมด | จำนวนข้อถูก | ร้อยละ |
|---|---|---|---|---|
| 1 | Part of Speech | 10 | 3 | 30 |
| 2 | Using Context Clues | 10 | 3 | 30 |
| 3 | Word Formation | 10 | 4 | 40 |
| 4 | Making Inferences | 10 | 3 | 30 |
| | รวม | 40 | 13 | 32.5 |

<< กลับสู่หน้าหลัก

ภาพที่ ช-3 ตัวอย่างผลการทดสอบด้วยคอมพิวเตอร์