

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

โสฬส สุขานนท์สวัสดิ์ และ เสรี ชัดแจ้ง

ศูนย์การวัดผลประยุกต์ วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

กฤษณะ ชินสาร

คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยนำทฤษฎีการตัดสินใจมาประยุกต์กับขั้นตอนการคัดเลือกข้อสอบข้อถัดไป และเปรียบเทียบผลการทดสอบด้วยวิธีที่พัฒนาขึ้นกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับความสามารถของผู้สอบ โดยการพิจารณาจำนวนข้อสอบที่ใช้ในการทดสอบและเวลาที่ใช้ในการทดสอบ

กลุ่มตัวอย่างในการวิจัยเป็นนักศึกษาระดับปริญญาตรี ภาคเรียนที่ 1 ปีการศึกษา 2555 มหาวิทยาลัยราชภัฏสวนดุสิต กรุงเทพมหานคร ที่ลงทะเบียนเรียนวิชาภาษาอังกฤษเพื่อทักษะการเรียน จำนวน 280 คน ผู้วิจัยได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เป็นส่วนหนึ่งของการวิจัย ผลการวิจัยปรากฏว่า วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจมีประสิทธิภาพดีกว่า ทั้งในด้านจำนวนข้อสอบที่ใช้ในการทดสอบมีจำนวนน้อยกว่าและใช้เวลาในการทดสอบน้อยกว่าวิธีการที่ใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับความสามารถของผู้สอบ

คำสำคัญ: การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์, การคัดเลือกข้อสอบข้อถัดไป, ทฤษฎีการตัดสินใจ

Development of the Next Item Selection Procedure Using Decision Theory in Computerized Adaptive Testing

Soros Sukhanonsawat and Seree Chadcham

Centre for Applied Measurement

College of Research Methodology and Cognitive Science

Burapha University, Thailand

Krisana Chinnasarn

Faculty of Informatics, Burapha University, Thailand

Abstract

The objective of this research was to apply decision theory to the process of next-item selection in computerized adaptive testing, and to compare results with those obtained using the maximum information criterion method for next-item selection. Comparisons were made on the basis of the number of test items in the examination, and the time spent on the examination.

For this study, carried out in the first semester of 2555, the sample was 280 undergraduate students from Suan Dusit Rajabhat University, Bangkok who were enrolling in the English for Study Skills courses. A computerized adaptive testing program was developed as part of the research. It was found that using decision theory had two advantages: fewer test items were necessary, and less examination time was required than using the maximum information criterion method for next-item selection process.

Keywords: computerized adaptive testing, the next-item selection, decision making theory

ความนำ

การทดสอบเป็นกระบวนการหนึ่งของการจัดการเรียนการสอน เป็นกระบวนการวัดผลสัมฤทธิ์การเรียนรู้ของผู้เรียนหลังจากเรียนจบหลักสูตร โดยทั่วไปจะใช้แบบทดสอบชนิดใช้กระดาษเขียนตอบ (Paper pencil test) แบบทดสอบที่นำมาใช้ในการวัดผลสัมฤทธิ์การเรียนรู้มีอยู่หลายรูปแบบ เช่น แบบทดสอบแบบหลายตัวเลือก แบบทดสอบแบบจับคู่ แบบทดสอบแบบเขียนตอบ เป็นต้น ผู้สอบจะได้ทำแบบทดสอบชุดเดียวกันและใช้เวลาในการทดสอบเท่ากัน ซึ่งเป็นแบบทดสอบที่สร้างขึ้นมาให้ครอบคลุมองค์ประกอบของคุณลักษณะหรือเนื้อหาที่ต้อง เป็นแบบทดสอบที่มีค่าความยากของข้อสอบที่หลากหลาย ไม่มีการกำหนดสัดส่วนของค่าความยากที่แน่นอน นิยมใช้ข้อสอบที่มีค่าความยากของข้อสอบที่ระดับ 0.20 ถึง 0.80 สามารถเรียกอีกอย่างหนึ่งว่าการทดสอบแบบประเพณีนิยม (Classical testing)

การที่ผู้สอบได้ทำแบบทดสอบชุดเดียวกันและทดสอบในเวลาเดียวกัน ความแม่นยำของคะแนนที่ได้จะไม่เหมือนกัน กล่าวคือ ผู้สอบบางคนอาจจะได้คะแนนที่มีความคลาดเคลื่อนน้อย แต่คะแนนของบางคนอาจมีความคลาดเคลื่อนมาก ดังนั้นการจัดการทดสอบควรเป็น การจัดการทดสอบเป็นรายบุคคล โดยการคัดเลือกข้อสอบที่มีความเหมาะสมกับผู้สอบ เพื่อประมาณค่าความสามารถของผู้สอบได้อย่างแม่นยำและมีความคลาดเคลื่อนต่ำจึงยุติการทดสอบ เรียกว่าการทดสอบแบบปรับเหมาะตามระดับความสามารถของผู้สอบ (Adaptive testing)

การทดสอบแบบปรับเหมาะตามระดับความสามารถของผู้สอบ เป็นการทดสอบที่ใช้แบบทดสอบแตกต่างกับสำหรับผู้สอบแต่ละคน โดยมีการคัดเลือกข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบ (Weiss, 1974) ทำให้จำนวนข้อสอบที่ใช้ในการทดสอบน้อยลง ประหยัดเวลาในการทดสอบ นอกจากนั้นยังช่วยผ่อนคลายความเครียดของผู้สอบได้ เช่น ผู้สอบที่มีความสามารถต่ำไม่เครียดกับการทำข้อสอบที่ยากเกินความสามารถของตนเอง ขณะที่ผู้สอบที่มีความสามารถสูงไม่จำเป็นต้องทำข้อสอบที่ง่ายเกินไป ซึ่งอาจจะส่งผลให้ผู้สอบขาดความระมัดระวังในการทำการทดสอบ การคัดเลือกข้อสอบที่จะนำมาใช้ในการทดสอบแบบปรับเหมาะ เป็นขั้นตอนที่มีความสำคัญขั้นตอนหนึ่ง กล่าวคือ หากพิจารณาคัดเลือกข้อสอบที่ไม่มีความเหมาะสมกับระดับความสามารถของผู้สอบ ไม่ว่าจะเป็นการคัดเลือกข้อสอบข้อเริ่มต้นหรือการคัดเลือกข้อสอบข้อถัดไป จะส่งผลต่อการประมาณค่าความสามารถของผู้สอบและทำให้มีความคลาดเคลื่อนสูงตามไปด้วย ดังนั้นการคัดเลือกข้อสอบที่จะนำมาใช้จะต้องพิจารณาจากปัจจัยต่าง ๆ ของผู้สอบร่วมด้วย เช่น คะแนนดิบหรือเกรดเฉลี่ยของผู้สอบในการคัดเลือกข้อสอบข้อเริ่มต้น หากเป็นการพิจารณาคัดเลือกข้อสอบข้อถัดไปอาจจะพิจารณาจากค่าความสามารถของผู้สอบที่ได้จากการตอบข้อสอบข้อสอบข้อที่ผ่านมา ซึ่งสามารถนำทฤษฎีการตัดสินใจมาประยุกต์ใช้ได้

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นการทดสอบที่จัดข้อสอบให้เหมาะสมกับความสามารถของผู้สอบ ซึ่งอยู่บนพื้นฐานของการตอบข้อสอบข้อที่ผ่านมาของผู้สอบ เมื่อผู้สอบทำข้อสอบข้อเริ่มต้นหรือชุดเริ่มต้น (ขึ้นอยู่กับการออกแบบ) แล้วนำผลการตอบข้อสอบมาวิเคราะห์หรือประเมินระดับความสามารถของผู้สอบเพื่อที่จะคัดเลือกข้อสอบข้อถัดไปที่เหมาะสมโดยอาศัยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เป็นพื้นฐาน (Weiss, 1974) การดำเนินการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มี 5 ขั้นตอน คือ ขั้นตอนที่ 1 การสร้างคลังข้อสอบ (Create item bank) เป็นขั้นตอนที่ต้องพิจารณานาขนาดของคลังข้อสอบและเงื่อนไขต่าง ๆ ที่สอดคล้องกับหลักการของทฤษฎีการตอบสนองข้อสอบ เพื่อคัดเลือกข้อสอบที่ตรงตามเงื่อนไขที่กำหนดมาจัดเก็บในคลังข้อสอบ ขั้นตอนที่ 2 การคัดเลือกข้อสอบข้อเริ่มต้น (First item selection) เป็นขั้นตอนที่ต้องพิจารณาคัดเลือกข้อสอบข้อเริ่มต้นที่มีความเหมาะสมกับผู้สอบ ซึ่งควรเป็นข้อสอบที่มีค่าความยากของข้อสอบระดับปานกลาง โดยการเลือกข้อสอบที่มีค่าความยากของข้อสอบใกล้เคียงกับค่าเฉลี่ยของค่าความยากของข้อสอบทั้งหมด

ขั้นตอนที่ 3 การคัดเลือกข้อสอบข้อถัดไป (Next item selection) เป็นขั้นตอนที่ต้องพิจารณาคัดเลือกข้อสอบตามเงื่อนไขที่กำหนด โดยอาศัยผลการตอบข้อสอบข้อก่อนหน้ามาร่วมพิจารณาด้วย ขั้นตอนที่ 4 การประมาณค่าความสามารถของผู้สอบ (Calculate possible ability levels) เป็นการประมาณค่าความสามารถของผู้สอบหลังจากตอบข้อสอบแล้ว ขั้นตอนที่ 5 เกณฑ์การยุติการทดสอบ (Termination criterion) คือ การสิ้นสุดการทดสอบ เมื่อผู้สอบทดสอบครบตามเงื่อนไขหรือเกณฑ์ที่กำหนด การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะขาดขั้นตอนใดขั้นตอนหนึ่งไม่ได้ เนื่องจากแต่ละขั้นตอนมีการดำเนินการที่มีความสัมพันธ์กัน (Thompson & Weiss, 2011; Wainer, 2000)

ทฤษฎีการตัดสินใจ (Decision theory) เป็นทฤษฎีที่อธิบายถึงการนำหลักเกณฑ์หรือเครื่องมือต่าง ๆ เข้ามาช่วยในการตัดสินใจ เพื่อให้ผู้ตัดสินใจมีโอกาสผิดพลาดน้อยลง ภายใต้สถานการณ์ต่าง ๆ ที่สามารถประเมินได้ โดยใช้กฎเกณฑ์หรือเครื่องมือดังกล่าวมาพิจารณาทางเลือกที่เหมาะสม ถ้ามีทางเลือกเพียงทางเลือกเดียวการตัดสินใจก็ไม่เกิดขึ้น เนื่องจากต้องเลือกตามวิธีทางเลือกที่มีอยู่ ไม่ต้องมีการเปรียบเทียบว่าผลลัพธ์หรือความเหมาะสมของทางเลือกใดดีกว่า ถ้ามีวิธีที่มีความเหมาะสมมากกว่าหนึ่งทางเลือกแล้ว จะต้องมีการตัดสินใจเลือกทางเลือกที่มีความเหมาะสมมากที่สุด (Hansson, 2005) หลักการของทฤษฎีการตัดสินใจ เป็นการนำเงื่อนไขและหลักเกณฑ์ของแต่ละกรณีมาใช้ในการพิจารณา เพื่อหาทางเลือกที่มีความเหมาะสมมากที่สุด ซึ่งสามารถแบ่งวิธีการตัดสินใจออกเป็น 3 ประเภท คือ 1) การตัดสินใจในสภาพแวดล้อมภายใต้ความแน่นอน เป็นวิธีที่ผู้ตัดสินใจทราบถึงความแน่นอนของสิ่งที่จะตามมาของทุกทางเลือก โดยผู้ตัดสินใจจะเลือกทางเลือกที่มีความเหมาะสมมากที่สุด 2) การตัดสินใจในสภาพแวดล้อมภายใต้ความไม่แน่นอน เป็นวิธีที่ผู้ตัดสินใจไม่ทราบถึงความน่าจะเป็นของการเกิดเหตุการณ์ต่าง ๆ ดังนั้น การตัดสินใจจึงอาศัยสารสนเทศที่มีอยู่ เจตคติส่วนตัวของผู้ตัดสินใจและความเหมาะสมตามเงื่อนไขที่มีอยู่ 3) การตัดสินใจภายใต้สภาพแวดล้อมที่มีความเสี่ยง เป็นวิธีที่ผู้ตัดสินใจทราบถึงความน่าจะเป็นของการเกิดเหตุการณ์ต่าง ๆ โดยพิจารณาจากค่าคาดหวังของผลที่ได้จากผลคูณของค่าความน่าจะเป็นของเหตุการณ์กับผลตอบแทนที่ได้รับของเหตุการณ์นั้น

จากหลักการของทฤษฎีการตัดสินใจข้างต้น สามารถประยุกต์กับการจัดการทดสอบได้ โดยในการเลือกแบบทดสอบจะพิจารณาทั้งกลุ่มที่ได้คะแนนสูงและกลุ่มที่ได้คะแนนต่ำ หากเป็นการเลือกข้อสอบจะพิจารณาทั้งกรณีที่ตอบข้อสอบถูกและกรณีที่ตอบข้อสอบผิด ซึ่งตามหลักการคัดเลือกแบบทดสอบหรือข้อสอบที่จะนำมาใช้ในการทดสอบนั้น จะต้องมีการกระจายของข้อมูลสูง จากเหตุผลข้างต้นประกอบกับการศึกษาเอกสารและงานวิจัยทั้งในและต่างประเทศ เกี่ยวกับวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่า ยังไม่มีการนำค่าประมาณความสามารถของผู้สอบก่อนที่จะตอบข้อสอบที่ยังไม่ได้ใช้ในคลังข้อสอบ กรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นถูกและกรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นผิด มาพิจารณาคัดเลือกข้อสอบข้อถัดไปร่วมกัน ผู้วิจัยจึงสนใจที่จะพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) เพื่อนำเสนอวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีประสิทธิภาพมากยิ่งขึ้น ด้านจำนวนข้อสอบที่ใช้ในการทดสอบและเวลาที่ใช้ในการทดสอบ

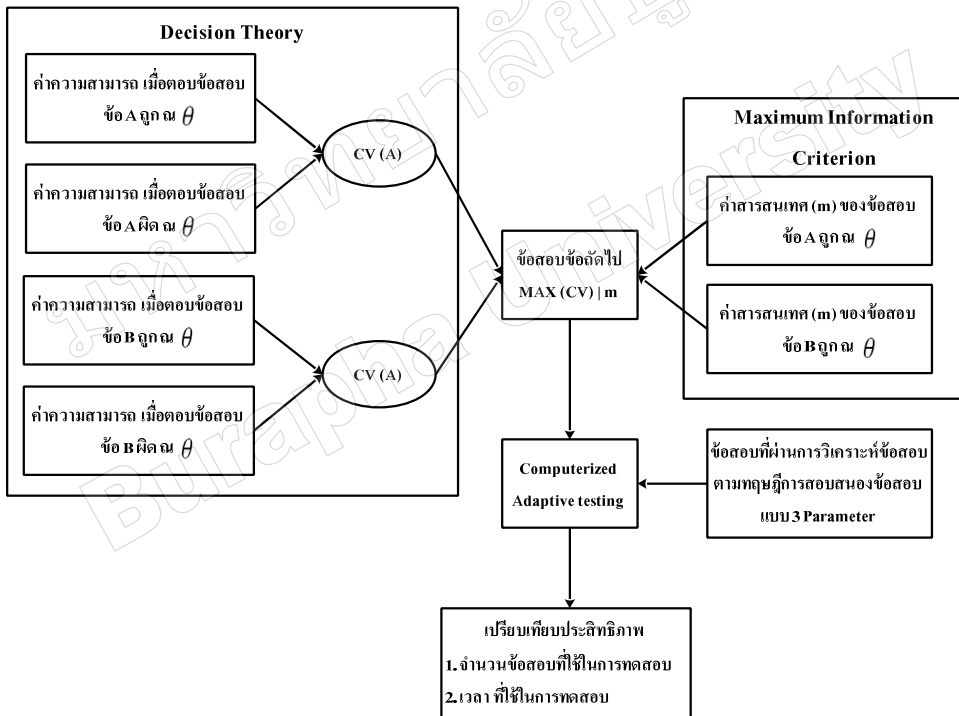
วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
2. เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้ทฤษฎีการตัดสินใจในการคัดเลือกข้อสอบข้อถัดไป

3. เพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ด้านจำนวนข้อสอบที่ใช้ในการทดสอบ และด้านเวลาที่ใช้ในการทดสอบ ระหว่างวิธีการใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ (Maximum Information Criterion: MI) กับวิธีการใช้ทฤษฎีการตัดสินใจ (Decision theory: DT)

กรอบแนวคิดการวิจัย

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้ทฤษฎีการตัดสินใจนี้ เป็นการนำวิธีการตัดสินใจภายใต้ความเสี่ยง (Decision making under risk) มาประยุกต์ใช้ในการคัดเลือกข้อสอบข้อถัดไป โดยการนำค่าประมาณความสามารถของผู้สอบก่อนที่จะตอบข้อสอบที่มีอยู่ในคลังข้อสอบ กรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นถูกและกรณีที่คาดว่าจะตอบข้อสอบข้อนั้นผิด มาพิจารณาคัดเลือกข้อสอบข้อถัดไปร่วมกัน โดยไม่นำข้อสอบที่ใช้ไปแล้วมาพิจารณา และเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่ใช้ในการทดสอบ และด้านเวลาที่ใช้ในการทดสอบระหว่างวิธีการที่พัฒนาขึ้นกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ กรอบแนวคิดการวิจัยแสดงดังภาพที่ 1



- θ คือ ความสามารถปัจจุบันของผู้สอบ
- MAX (CV) คือ ค่าสัมประสิทธิ์การกระจายสูงสุด ซึ่งเป็นค่าที่ใช้ในการเลือกข้อสอบข้อถัดไป
- m คือ ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ

ภาพที่ 1 กรอบแนวคิดการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

สมมติฐานการวิจัย

1. จำนวนข้อสอบที่ใช้ในการทดสอบของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจมีจำนวนน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ
2. เวลาที่ใช้ในการทดสอบของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ

วิธีดำเนินการวิจัย

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจ

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีขั้นตอนการดำเนินการ ดังนี้

1. ศึกษาหลักการและวิธีการที่เกี่ยวกับทฤษฎีการตัดสินใจ ร่วมกับทฤษฎีการตอบสนองข้อสอบ (Item response theory) เพื่อค้นหาวิธีการตัดสินใจที่เหมาะสมสำหรับการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้อย่างมีประสิทธิภาพ

2. วิเคราะห์เนื้อหาและหลักการของทฤษฎีการตัดสินใจ เพื่อคัดเลือกวิธีการที่มีความเหมาะสมที่จะนำไปพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป ซึ่งในที่นี้ ผู้วิจัยเลือกวิธีการตัดสินใจภายใต้สภาพแวดล้อมที่มีความเสี่ยงตามเกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุด (Maximum Expected Monetary Value: EMV) กรณีที่ใช้สัมประสิทธิ์การกระจายมาพัฒนาเป็นวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

3. พัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป โดยนำผลการวิเคราะห์เนื้อหาและหลักการของทฤษฎีการตัดสินใจ มาพัฒนาเป็นวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งมีขั้นตอน ดังนี้

- 3.1 ศึกษาความเป็นมาและคุณลักษณะของวิธีการตัดสินใจภายใต้สภาพแวดล้อมที่มีความเสี่ยงตามเกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุด กรณีที่ใช้ค่าสัมประสิทธิ์การกระจายอย่างละเอียด

- 3.2 ทดลองใช้หลักการและสูตรของวิธีการตัดสินใจภายใต้สภาพแวดล้อมที่มีความเสี่ยงตามเกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุดกรณีที่ใช้ค่าสัมประสิทธิ์การกระจาย โดยการจำลองเหตุการณ์ขึ้นมาทดลอง ซึ่งวิธีการตัดสินใจดังกล่าวจะต้องอาศัยความน่าจะเป็น และผลตอบแทนแต่ละด้านของแต่ละเหตุการณ์ร่วมกัน เพื่อหาความคาดหวังสูงสุด (EMV) ค่าส่วนเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจาย

- 3.3 นำหลักการและสูตรของวิธีการตัดสินใจภายใต้สภาพแวดล้อมที่มีความเสี่ยงตามเกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุดกรณีที่ใช้สัมประสิทธิ์การกระจายมาใช้ร่วมกับหลักการคัดเลือกข้อสอบมาพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ

4. นำวิธีการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้นไปทดสอบความถูกต้อง และหาข้อผิดพลาดที่เกิดขึ้นจากการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป โดยการเปรียบเทียบกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีขั้นตอนดังนี้

1. วิเคราะห์และออกแบบโปรแกรม เป็นการกำหนดโครงสร้างและลักษณะในการทำงานของโปรแกรม มีขั้นตอนการดำเนินการ ดังนี้

1.1 วิเคราะห์แผนผังบริบท (Context diagram) เป็นขั้นตอนการวิเคราะห์เพื่อหา Source destination ที่เกี่ยวข้องกับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

1.2 การจัดทำผังการไหลของข้อมูล (Data flow diagram) เป็นขั้นตอนการวิเคราะห์ให้เห็นภาพรวมของโปรแกรมทั้งข้อมูลและการทำงาน ซึ่งทำให้ง่ายต่อการศึกษาของผู้ที่สนใจครั้งต่อไป

2. ออกแบบฐานข้อมูล เป็นขั้นตอนการสร้างแฟ้มข้อมูลที่ใช้ในการเก็บข้อมูลให้สอดคล้องกับโครงสร้างและลักษณะในการทำงานของโปรแกรม

3. พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นขั้นตอนการพัฒนาโปรแกรมในรูปแบบของ Web application โดยใช้ภาษา PHP ในการพัฒนาโปรแกรม ใช้ MySQL เป็นฐานข้อมูล และใช้ภาษา SQL เป็นคำสั่งในการเชื่อมโยงข้อมูลในฐานข้อมูล

4. ทดสอบและแก้ไขโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นขั้นตอนการตรวจสอบหาจุดบกพร่องของโปรแกรม เพื่อนำจุดบกพร่องไปแก้ไขต่อไป โดยนำไปทดลองใช้กับอาจารย์และนักศึกษามหาวิทยาลัยราชภัฏสวนดุสิต กรุงเทพมหานคร

5. จัดทำคู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นขั้นตอนจัดทำคู่มือการใช้โปรแกรมเพื่อเป็นแนวทางการนำไปใช้จริง

6. ประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นขั้นตอนที่นำโปรแกรมและคู่มือไปทดลองใช้ เพื่อประเมินความเหมาะสมของโปรแกรม โดยผู้เชี่ยวชาญที่มีความรู้ทางด้านการศึกษา การศึกษา และด้านการพัฒนาโปรแกรมคอมพิวเตอร์ มีประสบการณ์ไม่น้อยกว่า 10 ปี

การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป

การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ในด้านจำนวนข้อสอบที่ใช้ในการทดสอบ และด้านเวลาที่ใช้ในการทดสอบ ระหว่างวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศที่ใกล้เคียงกับค่าความสามารถของผู้สอบกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ โดยทดสอบกับกลุ่มตัวอย่างจำนวน 280 คน นำเสนอเป็นแผนผังการดำเนินการดังภาพที่ 2

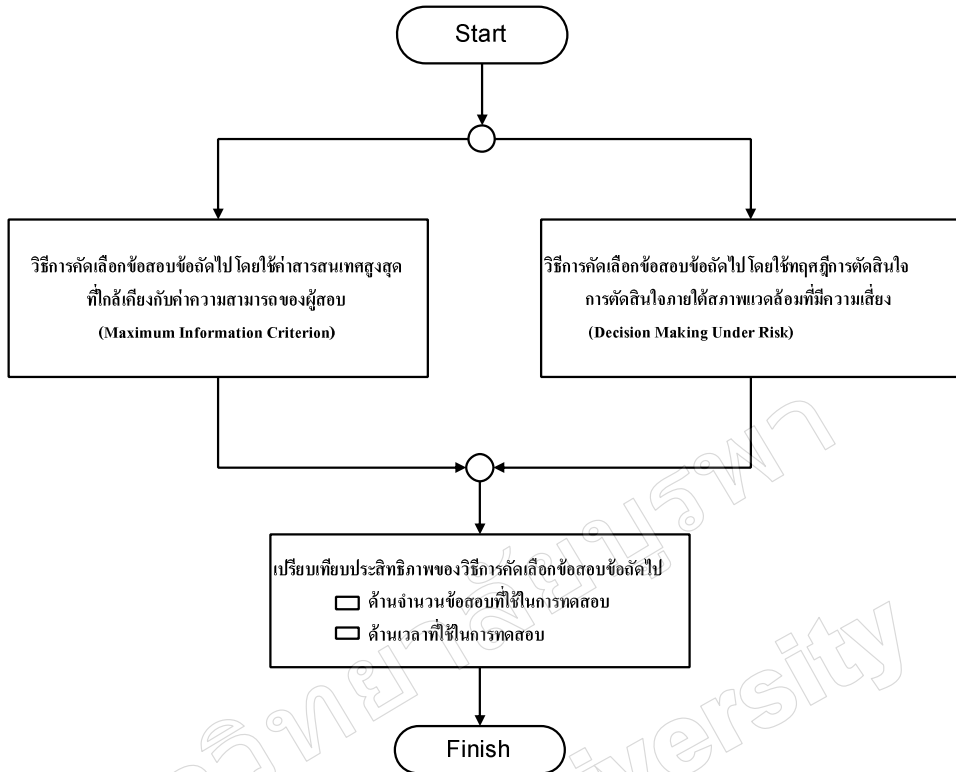
ประชากรและกลุ่มตัวอย่าง

ประชากรในการวิจัยเพื่อพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นนักศึกษาระดับปริญญาตรีที่ลงทะเบียนเรียนวิชาภาษาอังกฤษเพื่อทักษะการเรียน ในภาคเรียนที่ 1 ปีการศึกษา 2555 มหาวิทยาลัยราชภัฏสวนดุสิต กรุงเทพมหานคร จำนวน 1,200 คน

กลุ่มตัวอย่างใช้วิธีการสุ่มอย่างง่ายด้วยคอมพิวเตอร์จากรายชื่อนักศึกษาที่ลงทะเบียนเรียน จำนวน 280 คน สุ่มแบ่งออกเป็น 2 กลุ่ม ดังนี้

กลุ่มที่ 1 ดำเนินการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ จำนวน 140 คน

กลุ่มที่ 2 ดำเนินการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ จำนวน 140 คน



ภาพที่ 2 ขั้นตอนการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป

การเก็บรวบรวมข้อมูล

1. อธิบายเงื่อนไขหลักของการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์อย่างละเอียดให้แก่ นักศึกษากลุ่มตัวอย่างทั้ง 2 กลุ่ม พร้อมเอกสารวิธีการทดสอบ
2. ดำเนินการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับนักศึกษาทั้ง 2 กลุ่ม โดยจัดการทดสอบผ่านทาง Website (www.engcat.net) ซึ่งนักศึกษามีสามารถทำการทดสอบ ณ สถานที่ใดก็ได้ที่สามารถเข้าถึงเว็บไซต์ ระหว่างวันที่ 1 ตุลาคม พ.ศ. 2555 ถึง 12 ตุลาคม พ.ศ. 2555

หลังจากเสร็จสิ้นการทดสอบตามวันที่กำหนด นำผลการทดสอบของนักศึกษาทั้ง 2 กลุ่ม มาวิเคราะห์ประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ในด้านจำนวนข้อสอบที่ใช้ในการทดสอบและด้านเวลาที่ใช้ในการทดสอบ ระหว่างวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ

การวิเคราะห์ข้อมูล

1. การตรวจสอบประสิทธิภาพโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ข้อมูลจากผู้ทดลองใช้ที่เป็นอาจารย์และผู้นักศึกษา วิเคราะห์ด้วยสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ส่วนข้อมูลจากผู้เชี่ยวชาญ วิเคราะห์ด้วยดัชนีความสอดคล้อง
2. การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ในด้านจำนวนข้อสอบที่ใช้ในการทดสอบและด้านเวลาที่ใช้ในการทดสอบ ระหว่างวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ วิเคราะห์ดังนี้

2.1 จำนวนข้อสอบที่ใช้ในการสอบ วิเคราะห์ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และสถิติทดสอบไค-สแควร์ (χ^2) แบบการทดสอบความเป็นอิสระ

2.2 เวลาที่ใช้ในการทดสอบ วิเคราะห์ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และสถิติ t - test

ผลการวิจัย

ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่า ได้สูตรในการคัดเลือกข้อสอบข้อถัดไป สามารถเขียนเป็นสมการได้ ดังนี้

$$CV_i(\theta) = \frac{\sigma_i(EMV)(\theta)}{EMV_i(\theta)} ; i=1, 2, 3, \dots, n$$

โดยที่

$$\sigma_i(EMV)(\theta) = \sqrt{\left((\theta_{true} - EMV_i(\theta))^2 \times P_i(\theta) \right) + \left((\theta_{false} - EMV_i(\theta))^2 \times Q_i(\theta) \right)}$$

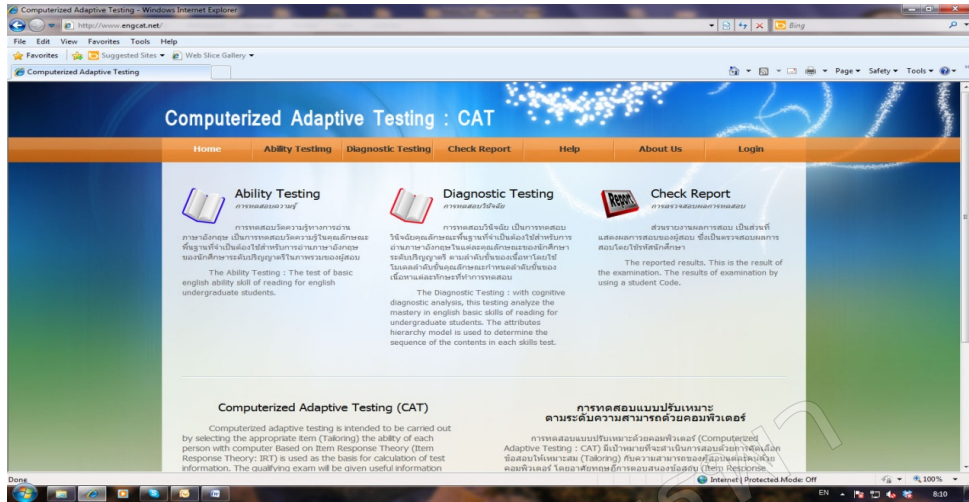
$$EMV_i(\theta) = (P_i(\theta) \times \theta_{true}) + (Q_i(\theta) \times \theta_{false})$$

เมื่อ

$CV_i(\theta)$	หมายถึง	ค่าสัมประสิทธิ์การกระจายของข้อสอบที่ i
$EMV_i(\theta)$	หมายถึง	ค่าความคาดหวังสูงสุดในระดับความสามารถของผู้สอบ ณ ขณะนั้นของข้อสอบที่ i
$\sigma_i(EMV)(\theta)$	หมายถึง	ค่าการกระจายข้อมูล ในระดับความสามารถของผู้สอบ ณ ขณะนั้นของข้อสอบที่ i
$P_i(\theta)$	หมายถึง	ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบที่ i ถูก
$Q_i(\theta)$	หมายถึง	$1 - P(\theta)$
θ_{true}	หมายถึง	ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ถูก
θ_{false}	หมายถึง	ค่าประมาณความสามารถเมื่อผู้สอบตอบข้อสอบข้อที่ i ผิด

ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ ปรากฏว่า โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีคุณภาพอยู่ในเกณฑ์ดีเป็นที่ยอมรับของผู้เชี่ยวชาญ ผู้ทดลองใช้ที่เป็นอาจารย์และผู้ทดลองใช้ที่เป็นนักศึกษา และมีความเหมาะสมที่จะนำไปใช้งาน หน้าจอแรกของโปรแกรม แสดงดังภาพที่ 3

จากภาพที่ 3 แสดงหน้าจอแรกของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ประกอบด้วยโปรแกรมทดสอบ 2 โปรแกรม คือ 1) โปรแกรมทดสอบความรู้ทางการอ่านภาษาอังกฤษ 2) โปรแกรมทดสอบวินิจัยคุณลักษณะทางการอ่านภาษาอังกฤษ



ภาพที่ 3 หน้าจอแรกของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไประหว่างวิธีการใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ ด้านจำนวนข้อสอบที่ใช้ในการทดสอบและด้านเวลาที่ใช้ในการทดสอบสามารถสรุปได้ ดังนี้

1. ด้านจำนวนข้อสอบที่ใช้ในการทดสอบ ผลการวิเคราะห์ข้อมูลแสดงดังตารางที่ 1 และตารางที่ 2

ตารางที่ 1 จำนวนข้อสอบที่ใช้ในการทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป

วิธีการคัดเลือกข้อสอบ ข้อถัดไป	จำนวน ตัวอย่าง (n)	จำนวนข้อสอบที่ใช้ ในการทดสอบ (ข้อ)		ค่าเฉลี่ย (M)	ส่วนเบี่ยงเบน มาตรฐาน (SD)
		มากที่สุด (Max)	น้อยที่สุด (Min)		
ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่า ความสามารถของผู้สอบ (MI)	140	21	8	12.86	2.65
ทฤษฎีการตัดสินใจ (DT)	140	10	6	6.84	0.64

ตารางที่ 1 แสดงผลข้อมูลพื้นฐานของวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้านจำนวนข้อสอบที่ใช้ในการทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป สามารถอธิบายได้ดังนี้

- 1) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบมีผู้สอบจำนวน 140 คน จำนวนข้อสอบที่ใช้ในการทดสอบประมาณ 12 – 13 ข้อ (ค่าเฉลี่ย = 12.86 ข้อ)
- 2) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ มีผู้สอบจำนวน 140 คน จำนวนข้อสอบที่ใช้ในการทดสอบประมาณ 6 – 7 ข้อ (ค่าเฉลี่ย = 6.84 ข้อ)

ตารางที่ 2 ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ด้านจำนวนข้อสอบที่ใช้ในการทดสอบ

วิธีการคัดเลือกข้อสอบ ข้อถัดไป	จำนวน ตัวอย่าง (n)	จำนวนข้อสอบ เฉลี่ย (M)	Chi-square	p-value
ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถ ของผู้สอบ (MI)	140	12.86	263.95**	0.00
ทฤษฎีการตัดสินใจ (DT)	140	6.84		

** $p < .01$

จากตารางที่ 2 ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้านจำนวนข้อสอบที่ใช้ในการทดสอบ ปรากฏว่า จำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้ทฤษฎีการตัดสินใจในการคัดเลือกข้อสอบข้อถัดไป ใช้จำนวนข้อสอบในการทดสอบน้อยกว่าวิธีใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. ด้านเวลาที่ใช้ในการทดสอบ ผลการวิเคราะห์ข้อมูลแสดงดังตารางที่ 3 และตารางที่ 4 ดังนี้

ตารางที่ 3 เวลาที่ใช้ในการทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป

วิธีการคัดเลือกข้อสอบ ข้อถัดไป	จำนวน ตัวอย่าง (n)	เวลาที่ใช้ ในการทดสอบ (นาที)		ค่าเฉลี่ย (M)	ส่วน เบี่ยงเบน มาตรฐาน (SD)
		มากที่สุด (Max)	น้อยที่สุด (Min)		
ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่า ความสามารถของผู้สอบ (MI)	140	38.72	1.10	8.83	6.76
ทฤษฎีการตัดสินใจ (DT)	140	21.95	0.72	4.82	4.29

จากตารางที่ 3 แสดงผลข้อมูลพื้นฐานของวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบ แบบปรับเหมาะด้วยคอมพิวเตอร์ ด้านเวลาที่ใช้ในการทดสอบ จำแนกตามวิธีการคัดเลือกข้อสอบข้อถัดไป สามารถอธิบายได้ ดังนี้

1) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบในการคัดเลือกข้อสอบข้อถัดไป มีผู้สอบจำนวน 140 คน ใช้เวลาในการทดสอบประมาณ 8.83 นาที (8 นาที 49.80 วินาที)

2) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ที่ใช้ทฤษฎีการตัดสินใจในการคัดเลือกข้อสอบข้อถัดไป มีผู้สอบจำนวน 140 คน ใช้เวลาประมาณ 4.82 นาที (4 นาที 49.20 วินาที)

ตารางที่ 4 ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ด้านเวลาที่ใช้ในการทดสอบ

วิธีการคัดเลือกข้อสอบ ข้อถัดไป	จำนวน ตัวอย่าง (n)	เวลาเฉลี่ย (M)	ส่วนเบี่ยงเบน มาตรฐาน (SD)	t-test	p-value
ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่า ความสามารถของผู้สอบ (MI)	140	8.83	6.76	5.92	0.00
ทฤษฎีการตัดสินใจ (DT)	140	4.82	4.29		

**p < .01

จากตารางที่ 4 ผลการเปรียบเทียบประสิทธิภาพของวิธีการเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้านเวลาที่ใช้ในการทดสอบ ปรากฏว่า เวลาที่ใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้ทฤษฎีการตัดสินใจในการคัดเลือกข้อสอบข้อถัดไป ใช้น้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศที่ใกล้เคียงกับค่าความสามารถของผู้สอบสูงสุด อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ระหว่างวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบกับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ ด้านจำนวนข้อสอบที่ใช้ในการทดสอบและด้านเวลาที่ใช้ในการทดสอบข้างต้น ชี้ให้เห็นว่า วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ มีประสิทธิภาพดีกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ

อภิปรายผล

การคัดเลือกข้อสอบข้อถัดไป เป็นขั้นตอนที่มีความสำคัญขั้นตอนหนึ่งในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เนื่องจากข้อสอบที่เลือกมาเป็นข้อสอบข้อถัดไปนั้นส่งผลต่อการประมาณค่าความสามารถของผู้สอบ ดังนั้นควรเลือกข้อสอบที่มีความสอดคล้องกับความสามารถของผู้สอบ ณ ขณะนั้น (Hulin, Drasgow & Parson, 1983) ซึ่งจะต้องพิจารณาตามเงื่อนไขในการคัดเลือกข้อสอบที่กำหนดขึ้น ขณะนี้ยังไม่มีการนำค่าประมาณความสามารถของผู้สอบก่อนที่จะตอบข้อสอบที่ยังไม่ได้ใช้ในคลังข้อสอบ กรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นถูกและกรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นผิดมาพิจารณาคัดเลือกข้อสอบข้อถัดไปร่วมกัน โดยการพิจารณาผลที่ได้จากการนำค่าประมาณความสามารถของผู้สอบก่อนที่จะตอบข้อสอบที่ยังไม่ได้ใช้ในคลังข้อสอบ กรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นถูกและกรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นผิดมาพิจารณาร่วมกันนั้นสอดคล้องกับหลักการของทฤษฎีการตัดสินใจ (Decision theory) ซึ่งเป็นทฤษฎีที่อธิบายถึงการนำหลักเกณฑ์หรือเครื่องมือต่าง ๆ เข้ามาช่วยในการตัดสินใจ ซึ่งจะพิจารณาผลที่ได้จากการประเมินภายใต้สถานการณ์ต่าง ๆ ทำให้ผู้ตัดสินใจเลือกทางเลือกที่มีความเหมาะสมมากที่สุดและมีโอกาสผิดพลาดน้อยลง (Hansson, 2005)

วิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้ทฤษฎีการตัดสินใจที่ผู้วิจัยพัฒนาขึ้นสามารถลดจำนวนข้อสอบที่ใช้ในการทดสอบลง และใช้เวลาในการทดสอบน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ ซึ่งสอดคล้องกับสมมุติฐานการวิจัย

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในการวิจัยนี้ เป็นการพัฒนาโปรแกรมในรูปแบบของ Web application โดยใช้ภาษา PHP ในการพัฒนาโปรแกรม ใช้ MySQL เป็นฐานข้อมูลและใช้ภาษา SQL เป็นคำสั่งในการเชื่อมโยงข้อมูลในฐานข้อมูล ผู้วิจัยพัฒนาโปรแกรมคอมพิวเตอร์ตามหลักการของวงจรการพัฒนา

ซอฟต์แวร์ (Software Development Life Cycle: SDLC) (Elliott, 2004) ซึ่งผู้สอบสามารถเข้ามาทำการทดสอบที่ใด เวลาใดก็ได้ แบ่งการทำงานออกเป็น 4 ส่วน คือ 1) ส่วนบริหารจัดการข้อสอบ เป็นส่วนที่สามารถเพิ่มข้อสอบ แก้ไขข้อสอบหรือลบข้อสอบออกจากคลังข้อสอบได้ตลอดเวลา ซึ่งข้อสอบที่จะนำมาใช้กับโปรแกรมนี้ต้องผ่านการวิเคราะห์ข้อสอบตามหลักการของทฤษฎีการตอบสนองข้อสอบ (Item response theory) แบบ 3 parameter มีค่าความยากของข้อสอบ (b -parameter) ระหว่าง -2.50 ถึง +2.50 ค่าอำนาจจำแนกของข้อสอบ (a -parameter) ระหว่าง 0.50 ถึง +2.50 และค่าการเดาของข้อสอบ (c -parameter) ไม่เกิน 0.30 2) ส่วนบริหารจัดการทดสอบ เป็นส่วนที่ผู้สอบเข้ามาทำการทดสอบระดับความสามารถทางการอ่านภาษาอังกฤษ นอกจากนั้นยังมีโปรแกรมจับวันและเวลาในการทดสอบด้วย ซึ่งโปรแกรมดังกล่าวจะตรวจสอบวันและเวลาที่ผู้สอบเข้ามาทำการทดสอบรวมถึงเวลาที่ใช้ในการทดสอบ หลังจากผู้สอบทำการทดสอบเสร็จแล้วโปรแกรมจะรายงานผลการทดสอบในส่วนนี้ด้วย 3) ส่วนรายงานผลการทดสอบ เป็นส่วนที่ผู้สอบสามารถตรวจสอบผลการทดสอบภายหลังจากออกจากส่วนของการจัดการทดสอบแล้ว โดยให้ผู้สอบใส่รหัสของผู้สอบที่ใช้ในการทดสอบ โปรแกรมจะแสดงผลการทดสอบของผู้สอบออกมา 4) ส่วนความช่วยเหลือ (Help) เป็นส่วนที่อธิบายวิธีใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เบื้องต้น

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในการวิจัยนี้ผ่านการตรวจสอบและทดลองใช้งานจากผู้ทดลองใช้ที่เป็นอาจารย์ ผู้ทดลองใช้ที่เป็นนักศึกษา และผู้เชี่ยวชาญแล้ว ซึ่งได้ประเมินโปรแกรมว่ามีคุณภาพอยู่ในเกณฑ์ดี มีความเหมาะสมที่จะนำไปใช้เป็นเป็นเครื่องมือในการจัดการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จึงทำให้โปรแกรมที่พัฒนาขึ้นมีประสิทธิภาพ

การเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจกับวิธีการใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบซึ่งเป็นวิธีการคัดเลือกข้อสอบที่นิยมใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่ใช้ในการทดสอบและด้านเวลาที่ใช้ในการทดสอบ ผลการวิจัยแสดงให้เห็นว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจมีประสิทธิภาพดีกว่าวิธีการใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ อาจจะเป็นเพราะว่า การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ เป็นวิธีการคัดเลือกข้อสอบข้อถัดไปที่นำผลการประมาณค่าความสามารถของผู้สอบก่อนทั้งกรณีที่คาดว่าจะตอบข้อสอบข้อนั้นถูกและกรณีที่คาดว่าจะตอบข้อสอบข้อนั้นผิด (ไม่นำข้อสอบที่ผู้สอบคนนั้นใช้แล้วมาพิจารณา) และคัดเลือกข้อสอบที่มีค่าสัมประสิทธิ์การกระจายสูงสุดที่ระดับความสามารถ ณ ขณะนั้นมาเป็นข้อสอบข้อถัดไป สาเหตุที่เลือกข้อสอบที่มีค่าสัมประสิทธิ์การกระจายสูงสุด เพราะตามหลักการคัดเลือกข้อสอบหรือแบบทดสอบนั้น ข้อสอบหรือแบบทดสอบที่ดีจะต้องมีค่าการกระจายของข้อมูลหรือสัมประสิทธิ์การกระจายของข้อมูลสูง จึงทำให้สามารถใช้ข้อสอบจำนวนน้อยลงกว่าวิธีดั้งเดิม เมื่อใช้จำนวนข้อสอบน้อยลงย่อมส่งผลให้เวลาที่ใช้ในการทดสอบในการทดสอบของผู้สอบแต่ละคนข้อยกลงไปด้วย

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. สามารถนำวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ ไปประยุกต์ในการคัดเลือกข้อสอบหรือแบบทดสอบชนิดอื่น ๆ ได้ เช่น แบบทดสอบสอบชนิดจับคู่ แบบทดสอบชนิดถูก-ผิด หรือแบบทดสอบชนิดเติมคำ เป็นต้น
2. สามารถนำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ไปประยุกต์กับการทดสอบทักษะการอ่านภาษาอังกฤษในหน่วยงานต่าง ๆ เช่น การทดสอบวัดระดับความรู้ทักษะการอ่านภาษาอังกฤษของผู้บริหารหรือพนักงานในบริษัท
3. สามารถนำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ไปประยุกต์กับการทดสอบวัดความรู้ในวิชาอื่น ๆ ได้

4. สามารถเป็นแนวทางการศึกษาถึงวิธีการพัฒนากระบวนการของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เช่น การสร้างคลังข้อสอบ การคัดเลือกข้อสอบข้อเริ่มต้น การประมาณค่าความสามารถของผู้สอบ และเกณฑ์การยุติการทดสอบ เป็นต้น

ข้อเสนอแนะในการวิจัยต่อไป

1. การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้ทฤษฎีการตัดสินใจ เป็นการพัฒนากระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไปเท่านั้น ดังนั้นควรศึกษาพัฒนากระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในขั้นตอนอื่นด้วย เช่น การสร้างคลังข้อสอบ การคัดเลือกข้อสอบข้อเริ่มต้น การประมาณค่าความสามารถของผู้สอบ หรือเกณฑ์การยุติการทดสอบ เป็นต้น

2. การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้ทฤษฎีการตัดสินใจ เป็นการพัฒนาโปรแกรมที่นำมาใช้กับการทดสอบวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้เท่านั้น และยังไม่สามารถใช้รูปภาพหรือสัญลักษณ์ทางคณิตศาสตร์กับข้อสอบได้ ดังนั้น ควรพัฒนาโปรแกรมให้รองรับการทดสอบวิชาอื่นและสามารถใช้รูปภาพ หรือสัญลักษณ์ทางคณิตศาสตร์ในข้อสอบได้ด้วย

3. ข้อสอบที่นำมาใช้ในการวิจัยนี้ เป็นข้อสอบแบบเลือกตอบชนิด 4 ตัวเลือก มีลักษณะตรวจให้คะแนนแบบ 2 ค่า (ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน) เท่านั้น ดังนั้น ควรนำไปศึกษากับข้อสอบที่มีลักษณะการตรวจให้คะแนนแบบหลายค่า เพื่อให้ได้ค่าสารสนเทศในการทดสอบมากขึ้น

4. การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจนี้ ยังไม่มีการตรวจสอบอัตราส่วนการใช้ข้อสอบซ้ำในการทดสอบแต่ละครั้ง ดังนั้น ควรมีการวิจัยตรวจสอบอัตราส่วนการใช้ข้อสอบซ้ำในการทดสอบแต่ละครั้งด้วย

เอกสารอ้างอิง

- Ayala, R. J. D. (2009). *The Theory and practice of item response theory*. New York: Guilford Press.
- Chen, S. & 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., Ankenmann, R. D., & Spray, J. A. (2003). *The relationship between item exposure and test overlap in computerized adaptive testing*. *Journal of Educational Measurement*, 40(2), 129-145.
- Chen, S. Y. & Lei, P. W. (2005) *Controlling Item Exposure and Test Overlap in Computerized Adaptive Testing*. *Applied Psychological Measurement* 29(3), 204-217.
- Dodd, B. G., Koch, W. R. & Ayala, D. (1993). *Computerized Adaptive Testing using the Partial Credit Model: Effect of Item Pool Characteristic and Different Stopping Rule*. *Education and Psychological Management*, 53(4), 61 – 77.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Boston: Kluwer Nijhoff.

- Hansson.S. O. (2005). *Decision Theory A Brief Introduction*. Stockholm : Department of Philosophy and the History of Technology, Royal Institute of Technology.
- Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Application to psychological measurement*. Illinois: Dow Jones-Irwin.
- McAlpine, M. (2002). *A Summary of Methods of Item Analysis*. London: Robert Clark Centre for Technological Education, University of Glasgow.
- Thompson, N. A. & Weiss, D. J. (2011). *A Framework for the Development of Computerized Adaptive Tests*. *Practical Assessment, Research & Evaluation* 16(1), 1-9.
- Urry, V. W. (1977). *Tailored Testing : A Successful Application of Latent Trait Theory*. *Journal of Education Measurament*, 14, 181 – 106.
- Wainer, H. (2000). *Computerized Adaptive Testing : A Primer*. New Jersey : Lawrence Erlbaum Associates.
- Weiss, D. J. (1974). *Strategies of Adaptive Ability Measurement*. Research Report. Minneapolis : University of Minnesote, Department of Psychology, Psychometric Methods Program.

มหาวิทยาลัยบูรพา
Burapha University