โปรแกรมต้นแบบของระบบแบบทดสอบออนไลน์ที่ปรับเปลี่ยนตามความสามารถของผู้สอบ โดยใช้แบบจำลองการทดสอบอัตราส่วนความน่าจะเป็นเชิงลำดับ กรณีศึกษา: ภาษาอังกฤษขั้นพื้นฐาน

A Prototype of Online Adaptive Testing System Using Sequential Probability Ratio Test Model Case Study: Basic English

ธนาวุธ บุญลิปตานนท์[1] กมลชนก ขำแก้ว[2] อรยา ปรีชาพานิช[3*] และสุดา เธียรมนตรี[3]

Thanawut Bunliptanon[1], Kamolchanok Khamkaew[2], Oraya Preechapanich[3*] and Suda Thernmontri[3]

บทคัดย่อ

ระบบแบบทดสอบที่ปรับเปลี่ยนตามความสามารถของผู้สอบโดยใช้คอมพิวเตอร์เป็นวิธีการทดสอบที่ใช้ความสามารถของคอมพิวเตอร์ในการคัดเลือกและนำเสนอข้อสอบที่สอดคล้องกับผลการประเมินระดับความสามารถของผู้สอบแต่ละคน งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบแบบทดสอบที่ปรับเปลี่ยนตามความสามารถของผู้สอบโดยใช้แบบจำลองการทดสอบอัตราส่วนความน่าจะเป็นเชิงลำดับ ซึ่งเหมาะกับการสร้างแบบทดสอบตามระดับความสามารถของผู้สอบแต่ละคนจากคลังข้อสอบขนาดเล็ก ระบบนี้ได้พัฒนาโดยใช้โปรแกรมภาษา PHP Javascript และ CSS ร่วมกับระบบจัดการฐานข้อมูล MySQL และได้สร้างคลังข้อสอบภาษาอังกฤษขั้นพื้นฐานจากข้อสอบ O-NET เพื่อใช้ในการทดสอบประสิทธิภาพของระบบตามหลักการวิศวกรรมซอฟต์แวร์และข้อกำหนดคุณลักษณะความต้องการใช้งานระบบ จากนั้นจึงได้ทำการประเมินผลการใช้งานระบบโดยใช้แบบสอบถามเกี่ยวกับส่วนต่อประสานกับผู้ใช้ ความสะดวกในการใช้งาน ความถูกต้อง และความสอดคล้องของระบบกับความต้องการใช้งาน สรุปได้ว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพอยู่ในระดับน่าพึงพอใจและสามารถนำไปประยุกต์ใช้งานได้จริง

คำสำคัญ: การทดสอบที่ปรับเปลี่ยนตามความสามารถของผู้สอบ ระบบแบบทดสอบออนไลน์ แบบจำลองการทดสอบอัตราส่วนความน่าจะเป็นเชิงลำดับ

---

[1] ครูพิเศษสอน วิทยาลัยสารพัดช่างภูเก็ต ภูเก็ต 83000

[2] ครูพิเศษสอน วิทยาลัยการอาชีพตะกั่วป่า พังงา 82110

[3] ผศ., สาขาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ พัทลุง 93210

[1] Contract Teacher, Phuket Polytechnic College, Phuket, 83000

[2] Contract Teacher, Takuapa Industrial and Community Education, Phang-nga, 82110

[3] Asst. Prof., Department of Computer and Information Technology, Faculty of Science, Thaksin University, Phatthalung, 93210

[*] Corresponding author: Tel.: 074-609600 ext. 2567. E-mail address: ms_oraya@yahoo.com

Abstract

Computerized adaptive testing (CAT) system is a method of assessment which the computer selects and presents test items to examinees according to the estimated examinee's ability levels. The objective of this research aims to develop the CAT system with Sequential Probability Ratio Test (SPRT) model that has ability to generate the customize test items from a small-scale item bank. This proposed system was developed by PHP, JavaScript, CSS and MySQL. The Basic English item bank in this prototype is constructed from O-Net examination for system performance testing based on software engineering principles and system requirement specifications. This system was assessed using a questionnaire examining user interface, ease of use, validation and functional requirement. The result of the questionnaires showed that the efficiency of this system is satisfied and it is possible to deploy this system for our students practically.

Keywords: Adaptive Testing, Online Testing System, Sequential Probability Ratio Test Model

บทนำ

แบบทดสอบเป็นเครื่องมือประเภทหนึ่งที่นิยมใช้ในการวัดผลสัมฤทธิ์ทางการศึกษา เนื่องจากแบบทดสอบจะใช้เป็นตัวแทนเนื้อหาการเรียนรู้ของผู้เรียนที่เปรียบเสมือนสิ่งเร้าที่นำไปกระตุ้นผู้เรียนให้แสดงพฤติกรรมตอบสนองออกมา แบบทดสอบที่นิยมใช้กันโดยทั่วไปคือ แบบทดสอบประเพณีนิยมหรือเรียกว่าแบบทดสอบใช้กระดาษและดินสอ (Paper-Pencil Test : P&P Test) ที่ผู้สอบจะได้รับแบบทดสอบชุดเดียวกันและมีจำนวนข้อสอบเท่ากัน ทำให้ผู้สอบบางคนจะต้องทำข้อสอบที่ยากหรือง่ายเกินไป ซึ่งอาจส่งผลต่อความถูกต้องและแม่นยำในการวัดผล [1] จากข้อจำกัดของทฤษฎีการทดสอบแบบมาตรฐานเดิม จึงได้มีการพัฒนาเป็นการทดสอบแบบปรับเปลี่ยนตามความสามารถของผู้สอบ (Adaptive Testing) ซึ่งในการทดสอบแต่ละครั้งผู้สอบจะได้แบบทดสอบที่เหมาะสมกับความสามารถของผู้สอบแต่ละคน โดยมีการคัดเลือกข้อสอบข้อถัดไปที่มีระดับความยากง่ายแปรผันไปตามผลที่ได้จากการประเมินความสามารถในการทำข้อสอบข้อที่ผ่านมาจนกว่าเงื่อนไขของการยุติการทดสอบจะเป็นจริง [2]

ในช่วงหลายปีที่ผ่านมานี้ประเทศไทยและประเทศสมาชิกอาเซียนได้มีการเตรียมพร้อมในการจัดตั้งประชาคมเศรษฐกิจอาเซียน (ASEAN Economic Community หรือ AEC) โดยกฎบัตรอาเซียนข้อ 34 ได้บัญญัติไว้ว่า ภาษาที่ใช้ในการทำงานของอาเซียนคือภาษาอังกฤษ (The Working Language of ASEAN Shall be English.) ดังนั้นภาษาอังกฤษจึงเป็นเครื่องมือที่สำคัญสำหรับพลเมืองอาเซียนในการสื่อสารสร้างสัมพันธ์สู่โลกกว้างของภูมิภาคอาเซียน โลกแห่งการแข่งขันไร้ขอบเขตภูมิศาสตร์และวัฒนธรรม [3] จากผลสำรวจทักษะความสามารถของคนไทยในการใช้ภาษาอังกฤษซึ่งเป็นเครื่องมือสำหรับการแข่งขันบนเวทีภูมิภาคอาเซียนและในตลาดโลกหลายครั้งพบว่า ประเทศไทยยังอยู่ในอันดับท้าย ๆ เมื่อเทียบกับประเทศอื่น ๆ เช่น ในปี พ.ศ. 2559 ดัชนีวัดระดับความรู้ทางภาษาอังกฤษของ Education First (EF English Proficiency Index- EPI) ของคนไทยอยู่ในกลุ่มประเทศที่มีความสามารถด้านภาษาอังกฤษระดับต่ำมาก นั่นคืออยู่ในลำดับที่ 56 จากทั้งหมด 72 ประเทศที่ได้รับการสำรวจ [4]

ด้วยเหตุนี้ผู้วิจัยจึงมีแนวความคิดที่จะพัฒนาโปรแกรมต้นแบบของระบบแบบทดสอบออนไลน์ที่ปรับเปลี่ยนตามความสามารถของผู้สอบ กรณีศึกษา: ภาษาอังกฤษขั้นพื้นฐาน เพื่อให้ระบบสร้างแบบทดสอบโดยการคัดเลือกข้อสอบที่เหมาะสมกับระดับความสามารถของผู้สอบแต่ละคน ส่งผลให้สามารถลดจำนวนข้อสอบและลดระยะเวลาในการทดสอบให้น้อยลง เนื่องจากผู้สอบไม่จำเป็นต้องทำข้อสอบจนครบทุกข้อเหมือนในอดีต การวิจัยครั้งนี้ผู้วิจัยได้เลือกใช้แบบจำลองการทดสอบอัตราส่วนความน่าจะเป็นเชิงลำดับ (Sequential Probability Ratio Test : SPRT) ซึ่งเป็นแบบจำลองประเภทหนึ่งที่ใช้ในการทดสอบแบบปรับเปลี่ยนตามความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ (Computer Adaptive Testing : CAT)

และเป็นการทดสอบที่ใช้การวิเคราะห์ตามทฤษฎีการตอบสนองรายข้อที่มีความเหมาะสมสำหรับการสร้างแบบทดสอบด้วยวิธีการสุ่มจากคลังข้อสอบขนาดเล็ก (Small-Scale Item Bank) [5-6] โดยคาดหวังว่าโปรแกรมต้นแบบนี้จะเป็นช่องทางหนึ่งสำหรับนิสิตเพื่อปรับพื้นฐานและฝึกทักษะภาษาอังกฤษของตนเองให้ดียิ่งขึ้น นอกจากนั้นแล้วอาจารย์ผู้สอนในรายวิชาภาษาอังกฤษหรือในรายวิชาอื่น ๆ ที่ต้องการสร้างระบบแบบทดสอบออนไลน์ที่ปรับเปลี่ยนตามความสามารถของผู้สอบสามารถนำระบบดังกล่าวไปประยุกต์ใช้งานในอนาคต

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

แบบทดสอบที่ปรับเปลี่ยนตามความสามารถของผู้สอบ

แบบทดสอบที่ปรับเปลี่ยนตามความสามารถของผู้สอบ (Adaptive Testing) หมายถึง แบบทดสอบที่ประกอบด้วยชุดของข้อสอบที่แตกต่างกันสำหรับผู้สอบแต่ละคนโดยใช้หลักการคัดเลือกข้อสอบตามระดับความสามารถที่มุ่งวัดของบุคคลนั้น [7] นั่นคือการทดสอบแบบปรับเปลี่ยนตามความสามารถของผู้สอบจะมีความแตกต่างที่สำคัญจากการทดสอบแบบประเพณีนิยมซึ่งมีการสร้างแบบสอบขึ้นมาชุดเดียวให้ครอบคลุมองค์ประกอบของคุณลักษณะหรือเนื้อหาที่ต้องการวัด มีค่าความยากของข้อสอบที่หลากหลายโดยไม่มีการกำหนดสัดส่วนของค่าความยากที่แน่นอน โดยทั่วไปนิยมให้มีค่าความยาก (p) อยู่ระหว่าง 0.2 - 0.8 สำหรับนำไปใช้สอบกับผู้สอบกลุ่มใดก็ได้ไม่ว่าผู้สอบจะมีความสามารถสูง ปานกลาง หรือต่ำ นั่นคือผู้สอบทุกคนจะต้องทำข้อสอบชุดเดียวกันที่มีจำนวนข้อเท่ากันและเหมือนกันหมดทุกข้อ แต่การทดสอบแบบปรับเปลี่ยนตามความสามารถของผู้สอบเป็นการใช้ข้อสอบจากคลังข้อสอบมาสร้างแบบสอบแบบสอบแต่ละชุดโดยมีการออกแบบให้เหมาะสมกับความสามารถของผู้สอบแต่ละคน การทดสอบแบบปรับเปลี่ยนตามความสามารถของผู้สอบจะช่วยลดจำนวนข้อสอบและเวลาของการสอบ ทั้งนี้การคัดเลือกข้อสอบที่มีระดับความยากพอเหมาะกับผู้สอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory Models : IRT) ย่อมให้สารสนเทศที่สำคัญเกี่ยวกับความสามารถที่แท้จริงของผู้สอบ ข้อสอบแต่ละข้อจากคลังข้อสอบที่นำมาใช้แต่ละคนจึงถูกเลือกหรือเสมือนเป็นการปรับให้เหมาะสมกับความสามารถของผู้สอบ ดังนั้นในขณะที่มีการทดสอบผู้สอบแต่ละคนจึงทำข้อสอบแตกต่างกันและจำนวนข้อของการสอบไม่จำเป็นต้องเท่ากัน [2, 8]

แบบจำลองการทดสอบอัตราส่วนความน่าจะเป็นเชิงลำดับ

แบบจำลองการทดสอบอัตราส่วนความน่าจะเป็นเชิงลำดับ (Sequential Probability Ratio Test Model : SPRT Model) เป็นแบบจำลองหนึ่งสำหรับระบบ CAT ที่มีการวิเคราะห์ตามทฤษฎีการตอบสนองรายข้อ มีการจัดข้อสอบให้เหมาะสมกับผู้สอบแต่ละคน และการบริหารการสอบทุกขั้นตอนทำโดยใช้คอมพิวเตอร์ โดยแบบจำลอง SPRT จะเหมาะสมกับการสร้างแบบทดสอบด้วยวิธีการสุ่มจากคลังข้อสอบขนาดเล็ก ในขณะที่แบบจำลองอื่น ๆ เช่น แบบจำลองโลจิสติกแบบหนึ่งพารามิเตอร์ สองพารามิเตอร์ หรือสามพารามิเตอร์ เหมาะสำหรับระบบที่มีคลังข้อสอบขนาดใหญ่และจำเป็นต้องใช้กลุ่มตัวอย่างขนาดใหญ่เพื่อประมาณการค่าพารามิเตอร์ที่เกี่ยวข้องกับการจัดการคลังข้อสอบ [5-6] โดยแบบจำลอง SPRT ใช้สูตรทางคณิตศาสตร์ดังต่อไปนี้

$$\text{LBM}\left(\text{Lower Bound Mastery}\right)=\frac{1-\beta}{\alpha}$$

$$\text{UBN}\left(\text{Upper Bound Non}-\text{mastery}\right)=\frac{\beta}{1-\alpha}$$

$$\text{ความน่าจะเป็นที่จะตอบถูก } \mathbf{PR}=\frac{P_m^s(1-P_m)^f}{P_{nm}^s(1-P_{nm})^f}$$

โดยที่

P_m คือ ความน่าจะเป็นที่รู้จริงและตอบคำถามถูก

P_{nm} คือ ความน่าจะเป็นที่ไม่รู้จริงแต่ตอบคำถามถูก

s คือ จำนวนข้อสอบที่ตอบถูกจากข้อสอบทั้งหมด

f คือ จำนวนข้อสอบที่ตอบผิดจากข้อสอบทั้งหมด

α คือ ข้อผิดพลาดแบบที่ 1 คือ ตัดสินว่า “รู้จริง” แต่ในความเป็นจริงแล้ว “ไม่รู้จริง”

β คือ ข้อผิดพลาดแบบที่ 2 คือ ตัดสินว่า “ไม่รู้จริง” แต่ในความเป็นจริงแล้ว “รู้จริง”

เมื่อผู้สอบตอบคำถามระบบก็จะทำการประมวลผลหาค่า PR ในทันที ถ้าค่า PR มีค่ามากกว่าหรือเท่ากับ LBM ผลที่ได้จากการสอบจะตัดสินว่า “รู้จริง” และการสอบก็จะสิ้นสุดลง ในทางกลับกัน ผลการสอบจะถูกตัดสินว่า “ไม่รู้จริง” โดยถ้าค่า PR มีค่ามากกว่า UBN แต่น้อยกว่า LBM ระบบจะทำการสุ่มข้อสอบมาให้ทำอีกครั้ง แต่ถ้าค่า PR มีค่าน้อยกว่าหรือเท่ากับ UBN ผู้สอบจะถูกตัดสินว่า “ไม่รู้จริง”

งานวิจัยที่เกี่ยวข้อง

ปัจจุบันได้มีการประยุกต์ใช้เทคโนโลยีคอมพิวเตอร์และเครือข่ายอินเทอร์เน็ตในการพัฒนาระบบ CAT ที่หลากหลาย เช่น การทดสอบแบบปรับเหมาะที่มีการสะท้อนข้อมูลย้อนกลับในการทดสอบมาตรฐานวิชาชีพไอที [9] ถัดมา คือ การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย [10] และการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการจัดสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 6 [11] งานวิจัยที่กล่าวมาแล้วข้างต้นมักจะใช้ทฤษฎี IRT ที่จำเป็นจะต้องมีการใช้ข้อสอบและจำนวนผู้สอบค่อนข้างมากเพื่อวิเคราะห์ค่าพารามิเตอร์ต่าง ๆ สำหรับการจัดการคลังข้อสอบให้เป็นไปตามมาตรฐาน ดังนั้นผู้วิจัยจึงมีแนวคิดในการพัฒนาโปรแกรมต้นแบบของระบบ CAT โดยใช้แบบจำลอง SPRT ซึ่งสามารถประยุกต์กับการสร้างและการจัดการคลังข้อสอบขนาดเล็กสำหรับใช้ในระบบแบบทดสอบออนไลน์ที่สามารถปรับเปลี่ยนข้อสอบให้สอดคล้องกับความสามารถของผู้สอบแต่ละคนได้เช่นกัน

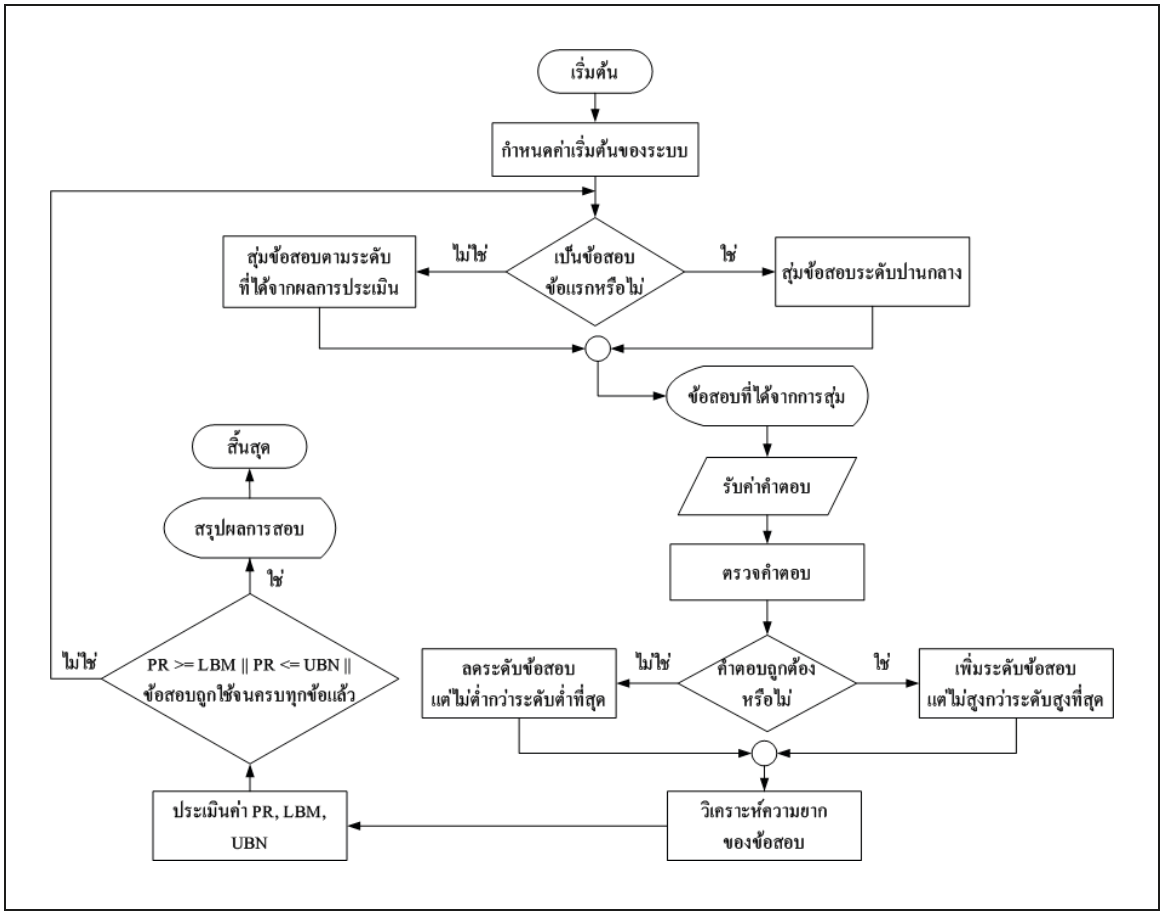
การดำเนินการวิจัย

ผู้วิจัยได้ดำเนินการวิจัย ประกอบด้วย การวิเคราะห์และออกแบบโปรแกรมต้นแบบ รวมไปถึงการพัฒนาและการทดสอบโปรแกรมต้นแบบ โดยมีรายละเอียดดังนี้

การวิเคราะห์และออกแบบโปรแกรมต้นแบบ

จากการทบทวนวรรณกรรมและสารสนเทศที่เกี่ยวข้อง ทั้งในส่วนของทฤษฎีและแนวคิดในการพัฒนาโปรแกรมต้นแบบ สามารถสรุปผลการวิเคราะห์และออกแบบโปรแกรมต้นแบบของระบบ CAT ในส่วนของโมดูลการทำข้อสอบซึ่งเป็นโมดูลหลักของงานวิจัยนี้ในรูปแบบของผังงานโปรแกรม (Program Flowchart) เพื่ออธิบายขั้นตอนการทำงานโมดูลการทำข้อสอบซึ่งใช้อัลกอริทึมที่มีพื้นฐานจากแบบจำลอง SPRT (ภาพที่ 1) เริ่มต้นจากการกำหนดค่าเริ่มต้นของโมดูลการทำข้อสอบ จากนั้นจึงเป็นการสร้างชุดข้อสอบโดยการตรวจสอบสถานะในการทำข้อสอบว่าเป็นการทำข้อสอบข้อแรกหรือไม่ เพื่อทำการสุ่มข้อสอบจากคลังข้อสอบตามหลักการที่ได้อธิบายไว้แล้วก่อนหน้านี้ หลังจากนั้นเมื่อผู้สอบได้ส่งคำตอบของข้อสอบแต่ละข้อแล้วโปรแกรมต้นแบบจะประเมินการตอบสนองข้อสอบด้วย ค่า θ_j และค่า PR ซึ่งในกรณีที่ผู้สอบตอบคำถามได้ถูกต้อง โปรแกรมต้นแบบจะเลือกข้อสอบข้อถัดไปที่มีระดับความยากเพิ่มขึ้นอีก 1 ระดับ ในทางตรงกันข้ามหาก

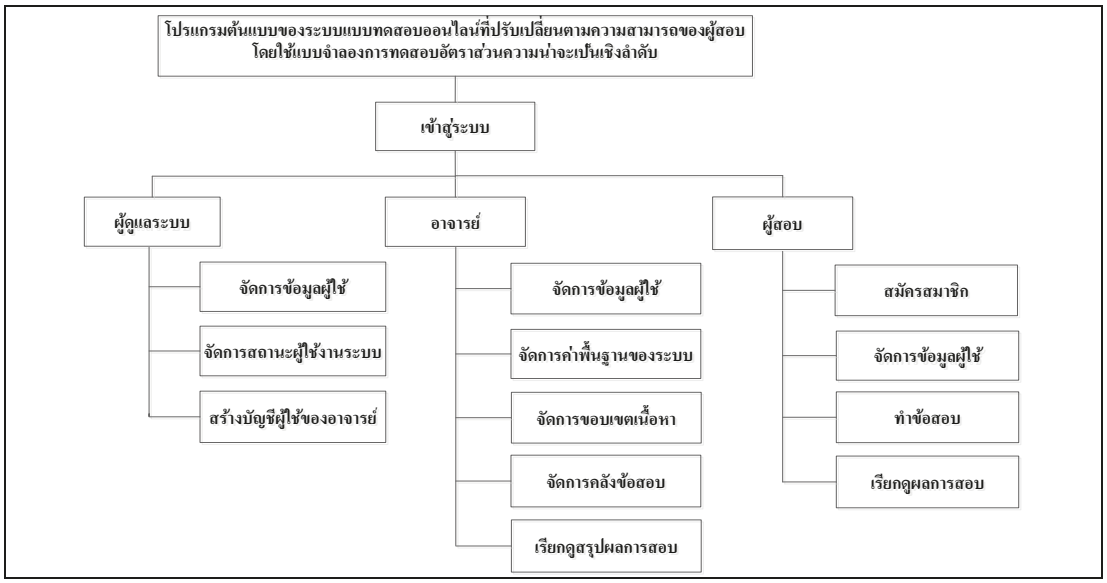
ผู้สอบตอบคำถามผิด ข้อสอบข้อถัดไปก็จะมีระดับความง่ายลดลงไปอีก 1 ระดับ การทำข้อสอบของผู้สอบแต่ละคนจะมีจำนวนข้อสอบมากน้อยแตกต่างกันขึ้นอยู่กับผลที่ได้จากการตรวจสอบเงื่อนไขที่ใช้ในการยุติการทดสอบนั่นคือ PR ≥ LBM หรือ PR ≤ UBN หรือจนกว่าข้อสอบในคลังข้อสอบได้ถูกใช้จนครบทุกข้อแล้ว



ภาพที่ 1 ผังงานโปรแกรมของโมดูลการทำข้อสอบ

การพัฒนาและทดสอบโปรแกรมต้นแบบ

จากขั้นตอนการวิเคราะห์และออกแบบโปรแกรมต้นแบบ ผู้วิจัยได้กำหนดโครงสร้างระบบของโปรแกรมต้นแบบโดยมีโมดูลต่างๆ จำแนกตามกลุ่มของผู้ใช้งานได้เป็น 3 กลุ่มคือ ผู้ดูแลระบบ อาจารย์ และผู้สอบ (ภาพที่ 2) จากนั้นผู้วิจัยได้พัฒนาโปรแกรมต้นแบบโดยใช้โปรแกรมภาษา PHP Javascript และ CSS ร่วมกับระบบจัดการฐานข้อมูล MySQL และได้สร้างคลังข้อสอบจำลองจากข้อสอบ O-NET ของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ในรายวิชาภาษาอังกฤษ จำนวน 300 ข้อ เพื่อใช้ในการทดสอบโปรแกรมต้นแบบตามหลักการวิศวกรรมซอฟต์แวร์และข้อกำหนดเกี่ยวกับความต้องการของระบบ

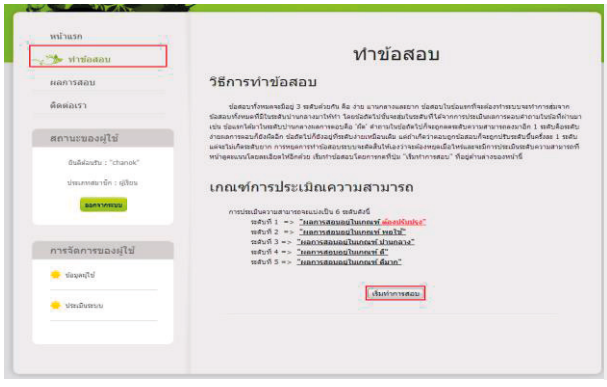


ภาพที่ 2 โครงสร้างของโปรแกรมต้นแบบในภาพรวม

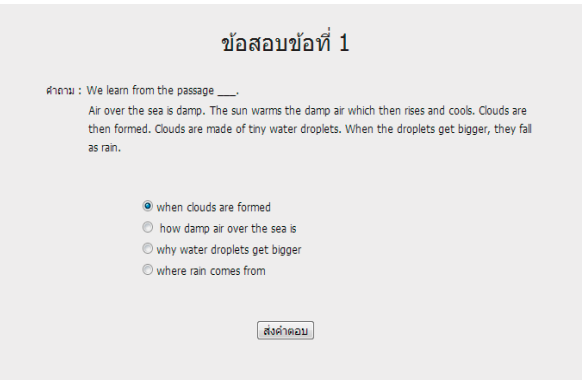
ผลการวิจัยและการอภิปรายผล

ผลที่ได้จากการพัฒนาโปรแกรมต้นแบบ

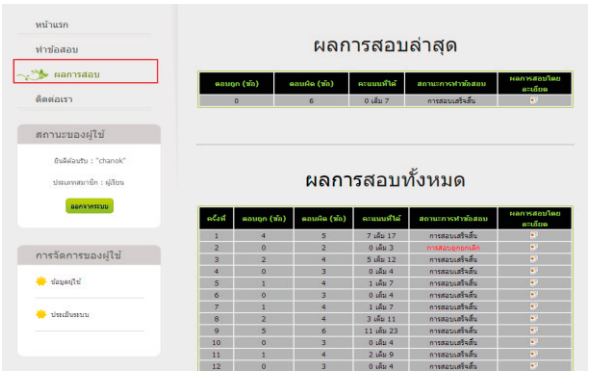
ผลที่ได้จากการพัฒนาโปรแกรมต้นแบบของระบบแบบทดสอบออนไลน์ที่ปรับเปลี่ยนตามความสามารถของผู้สอบโดยใช้แบบจำลองการทดสอบอัตราส่วนความน่าจะเป็นเชิงลำดับ กรณีศึกษา: ภาษาอังกฤษขั้นพื้นฐาน สามารถแสดงตัวอย่างหน้าจอการใช้งานของผู้สอบได้ดังนี้ หน้าหลักสำหรับการใช้งานโมดูลการทำข้อสอบ (ภาพที่ 3) หน้าจอการทำข้อสอบและส่งคำตอบ (ภาพที่ 4) หน้าจอสำหรับเรียกดูผลการสอบที่ผ่านมาของตนเอง (ภาพที่ 5) และหน้าจอแสดงรายละเอียดของผลการสอบแต่ละครั้ง (ภาพที่ 6)



ภาพที่ 3 หน้าหลักของการทำข้อสอบ



ภาพที่ 4 หน้าจอการทำข้อสอบและส่งคำตอบ



ภาพที่ 5 หน้าจอเรียกดูผลการสอบที่ผ่านมา

ภาพที่ 6 หน้าจอแสดงรายละเอียดผลการสอบแต่ละครั้ง

ผลการประเมินการใช้งานโปรแกรมต้นแบบ

กลุ่มตัวอย่างที่ใช้ในการประเมินผลการใช้งานโปรแกรมต้นแบบคือ นิสิตปริญญาตรีคณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ จำนวน 96 คน มีหลักเกณฑ์การให้คะแนนแบ่งเป็น 5 ระดับจากระดับน้อยที่สุดไปหาระดับมากที่สุด มีคะแนนตั้งแต่ 1-5 คะแนนตามลำดับ ผู้วิจัยได้ใช้เกณฑ์ของลิเคิร์ท (Likert) ในการสรุปผลคะแนนจากแบบสอบถาม ดังตารางที่ 1

ตารางที่ 1 สรุปผลการประเมินการใช้งานโปรแกรมต้นแบบ

รายการประเมิน	ผลการประเมิน	
	Mean	**S.D.**
1. ส่วนต่อประสานกับผู้ใช้งาน	4.42	0.58
2. ความสะดวกในการใช้งาน	4.52	0.57
3. ความถูกต้องของระบบ	4.51	0.57
4. ความสอดคล้องของระบบกับความต้องการใช้งาน	4.62	0.51

ผลการประเมินความพึงพอใจในการใช้งานระบบพบว่าผู้ใช้มีความพึงพอใจในส่วนต่อประสานกับผู้ใช้งานอยู่ในระดับมาก ส่วนประเด็นด้านความสะดวกในการใช้งาน ความถูกต้องของระบบ และความสอดคล้องของระบบกับความต้องการใช้งานอยู่ในระดับมากที่สุด ซึ่งจากความคิดเห็นและข้อเสนอแนะในแบบสอบถามพบว่า ระบบที่พัฒนาขึ้นสามารถใช้งานได้ง่าย เนื่องจากมีคำอธิบายประกอบการใช้งานอย่างชัดเจน และมีการเชื่อมโยงหน้าจอต่าง ๆ อย่างเป็นระบบ ส่วนในเรื่องของส่วนต่อประสานกับผู้ใช้งานพบว่า การใช้สี รวมไปถึงขนาดและตำแหน่งของวัตถุที่ใช้ในบางหน้าจอยังไม่สอดคล้องกับขนาดของพื้นที่ที่ใช้ในการนำเสนอข้อมูล

ผลการประเมินคุณภาพระบบโดยผู้เชี่ยวชาญ

การประเมินคุณภาพของระบบจากผู้เชี่ยวชาญ 3 คน ด้วยวิธีการประเมินผลระบบแบบฮิวริสติค (Heuristic Evaluation) โดยแบ่งประเด็นการสอบถามออกเป็น 3 ประเด็นหลักคือ คุณภาพของสารสนเทศ คุณภาพของระบบ และคุณภาพของการบริการ ดังตารางที่ 2

ตารางที่ 2 สรุปผลการประเมินคุณภาพของระบบ

ประเด็นที่ใช้ในการประเมินคุณภาพของระบบ	ผลการประเมิน	
	Mean	**S.D.**
1. **คุณภาพของสารสนเทศ** (ความแม่นยำ (Accuracy) ความน่าเชื่อถือ (Reliability) ความสมบูรณ์ (Completeness) และความตรงประเด็น (Relevance))	4.58	0.49
2. **คุณภาพของระบบ** (การใช้งานได้ (Usability) สภาพพร้อมใช้งาน (Availability) การออกแบบส่วนต่อประสานกับผู้ใช้ (User Interface Design) และความยืดหยุ่น (Flexibility))	4.42	0.49
3. **คุณภาพของการบริการ** (ความปลอดภัยและความเป็นส่วนตัว (Security and Privacy) และความสามารถในการตอบสนอง (Responsiveness))	4.67	0.50

จากตารางข้างต้นจะเห็นได้ว่าผลการประเมินคุณภาพของระบบในประเด็นด้านคุณภาพของสารสนเทศและคุณภาพของการบริการจะอยู่ในระดับมากที่สุด ส่วนประเด็นด้านคุณภาพของระบบจะอยู่ในระดับมาก โดยผู้เชี่ยวชาญได้ให้ข้อเสนอแนะเกี่ยวกับการออกแบบส่วนต่อประสานกับผู้ใช้ว่า ระบบควรมีการใช้สีสันและการแบ่งพื้นที่ในหน้าจอให้มีความสมดุลและมีจุดเด่นมากกว่านี้ ซึ่งผู้วิจัยจะนำผลสะท้อนกลับเหล่านั้นไปวิเคราะห์เพื่อปรับปรุงระบบให้ดียิ่งขึ้นในอนาคต

สรุปผลการวิจัย

งานวิจัยนี้เป็นการพัฒนาโปรแกรมต้นแบบของระบบแบบทดสอบออนไลน์ โดยผู้วิจัยได้เลือกใช้แบบจำลองการทดสอบอัตราส่วนความน่าจะเป็นเชิงลำดับ เพื่อสร้างโมดูลการทดสอบแบบปรับเปลี่ยนตามความสามารถของผู้สอบโดยใช้คอมพิวเตอร์และเป็นการทดสอบที่ใช้การวิเคราะห์ตามทฤษฎีการตอบสนองรายข้อที่มีความเหมาะสมกับคลังข้อสอบขนาดเล็ก โดยผู้สอบจะต้องสมัครสมาชิกเพื่อเข้าใช้งานระบบ ระบบจะสุ่มข้อสอบข้อแรกที่มีความยากในระดับปานกลางเพื่อประเมินระดับความสามารถของผู้สอบ จากนั้นระบบจะทำการคัดเลือกข้อสอบข้อถัดไปที่มีระดับความยากง่ายแปรผันไปตามผลที่ได้จากการประเมินความสามารถในการทำข้อสอบข้อที่ผ่านมาวนรอบไปโดยอัตโนมัติ จนกว่าเงื่อนไขของการยุติการทดสอบจะเป็นจริงหรือจนกว่าข้อสอบในคลังข้อสอบจะถูกใช้จนครบทุกข้อแล้ว ระบบจึงทำการสรุปผลการสอบเพื่อแจ้งให้ผู้สอบรับทราบ จากผลการประเมินคุณภาพของระบบพบว่าโปรแกรมต้นแบบที่พัฒนาขึ้นมีคุณภาพและมีความสอดคล้องกับข้อกำหนดเกี่ยวกับความต้องการของระบบ ดังนั้นในอนาคตจึงอาจประยุกต์โปรแกรมต้นแบบนี้เพื่อสร้างแบบทดสอบที่ปรับเปลี่ยนตามความสามารถของผู้สอบในรายวิชาพื้นฐานที่มีผู้สอบเป็นจำนวนมากและใช้ข้อสอบในรูปแบบปรนัยหลายตัวเลือก (Multiple Choices) เช่น รายวิชาพื้นฐานทางด้านภาษาศาสตร์ คณิตศาสตร์ วิทยาศาสตร์ รวมไปถึงด้านมนุษยศาสตร์และสังคมศาสตร์ เป็นต้น

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนสนับสนุนการทำวิจัยจากงบประมาณแผ่นดิน มหาวิทยาลัยทักษิณ ประจำปีงบประมาณ พ.ศ. 2559 ทำให้การทำวิจัยนี้สำเร็จลุล่วงด้วยดี ผู้วิจัยจึงขอขอบคุณมา ณ โอกาสนี้

เอกสารอ้างอิง

[1] Lord, F. M. (1980). **Application of Item Response Theory to Practical Testing Problems**. New Jersey: Lawrence Erlbaum Associates Publishers.

[2] Kanjanawasee, S. (2012). **Modern Test Theories**. Ed 4. Bangkok: Chulalongkorn University Press.

[3] Onwimon, S. (2011). **ASEAN Diary** (Online). Retrieved 15 August 2015, from http://blog.eduzones.com/wigi/81880.

[4] EF Education First. **The World's Largest Ranking of Countries by English Skills** (Online). Retrieved 12 March 2017, from http://www.ef.co.th/EPI.

[5] Frick, T.W. (1990). A comparison of three decision models for adapting the length of computer-based mastery test. **Journal of Educational Computing Research Systems**. 5(1), 89-114.

[6] Tao, Y.H., Wu, Y.L. and Chang, H.Y. (2008). A Practical Computer Adaptive Testing Model for Small-Scale Scenarios. **Educational Technology & Society**. 11, 259-274.

[7] Weiss, D.J. (1983). **New Horizons in Testing: Latent Trait Test Theory and Computerized Adaptive Testing**. New York, NY: Academic Press.

[8] Wong, K., Leung, K., Kwan, R. and Tsang. P. (2010). E-Learnig: Developing a Simple Web-Based Intelligent Tutoring System Using Cognitive Diagnostic Assessment and Adaptive Testing Technology. **Springer-Verlag Berlin Heidelberg**. 23-34.

[9] Chaimonglol, N., Pasiphol, S. and Kanjanawasee, S. (2016). Computerized Adaptive Testing with Reflective Feedback in Information Technology Professional Examination. **Information Technology Journal**. 12(2), 58-64.

[10] Sakolkijrungroj, S., Chadcham, S. and Sudhasani, S. (2015). The Development of Computerized Adaptive Testing Program for Thai Happiness Scale. **Research Methodology & Cognitive Science**. 13(1), 1-17.

[11] Plumjai, N., Tinnaworn, P. and Sukhanonsawat, S. (2015). Development of Computerized Adaptive Testing Program for O-NET at the Grade Twelve Level. **Research Methodology & Cognitive Science**. 13(2), 109-125.