

ความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบ
แบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบ
ด้วยคอมพิวเตอร์ สำหรับการทดสอบรายวิชาภาษาอังกฤษ
เพื่อทักษะการเรียนรู้ระดับปริญญาตรี
The Correlation between Ability Estimation of Examinees from
Computerized Adaptive Testing and Total Score of Examinees from
Computerized Testing for English Test
in Undergraduate Study Skills

รัชกฤช ธนพัฒน์^{1*} และปิยะทิพย์ ประดุงพรม^{2**}

¹อาจารย์ ดร. สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช ratchakrit.tan@stou.ac.th

²อาจารย์ ดร. วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา piyatip10@hotmail.com

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ สำหรับการทดสอบรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ระดับปริญญาตรี ซึ่งการทดสอบดังกล่าวทดสอบผ่านระบบทดสอบที่ผู้วิจัยพัฒนาขึ้นในรูปแบบ Web Application โดยผู้สอบสามารถเลือกการทดสอบได้ 2 ประเภท คือ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และการทดสอบด้วยคอมพิวเตอร์ การดำเนินการวิจัยแบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม กลุ่มละ 15 คน โดยกลุ่มที่ 1 ดำเนินการทดสอบด้วยคอมพิวเตอร์ (CT) ก่อน แล้วจึงทำการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT-MIC) ในขณะที่ กลุ่มที่ 2 ดำเนินการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ก่อน แล้วจึงทำการทดสอบด้วยคอมพิวเตอร์ ทั้งนี้เพื่อป้องกันปัญหาเรื่องลำดับของการทดสอบ ซึ่งอาจส่งผลต่อผลการทดสอบได้ จากนั้นจึงนำผลการทดสอบทั้ง 2 ประเภทมาทดสอบความสัมพันธ์ การวิเคราะห์ข้อมูลใช้สถิติเชิงบรรยาย ได้แก่ ความถี่ ค่าสูงสุด ค่าต่ำสุด ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน และการวิเคราะห์สหสัมพันธ์แบบเพียร์สัน ผลการวิจัยปรากฏว่าค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ มีความสัมพันธ์เชิงบวกขนาดสูงเท่ากับ .837 ที่ระดับนัยสำคัญทางสถิติ .01 แสดงว่า ผลการทดสอบแบบ CAT-MIC มีความสอดคล้องกับผลการทดสอบแบบ CT นั่นคือ ผู้สอบที่ได้รับค่าประมาณความสามารถจากการทดสอบแบบ CAT สูง(หรือต่ำ) จะได้คะแนนรวมจากการทดสอบแบบ CT สูง(หรือต่ำ) ด้วย

คำสำคัญ : การทดสอบด้วยคอมพิวเตอร์, การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์, ความสัมพันธ์

Abstract

The objective of the study was to investigate the correlation between ability estimation of examinees from computerized adaptive testing using the next item selection procedure using maximum information criterion and total score of examinees from computerized testing for english test in undergraduate study skills. this study was tested and developed in web application form, the examinees can choose 2 types of the test from computerized adaptive testing and computerized testing. the sample of examinees was divided with 2 groups and 15 people in each group, the first group started with computerized testing (CT) and computerized adaptive testing (CAT-MIC) respectively, the second group started with computerized adaptive testing and

computerized testing respectively, this process was used for avoid order effect which affect to testing. the results were used for data analysis using descriptive statistics such as frequency, max, min, mean, standard deviation and pearson's product moment correlation. The finding indicated that ability estimation of examinees from computerized adaptive testing and total score of examinees from computerized testing were at 0.837 high positive correlation and statistically significant at 0.01 level. the results form CAT-MIC conformed to CT therefore ability estimation of examinees from CAT was high (or low) and total score from CT was also high (or low).

Keywords : computerized testing, computerized adaptive testing, correlation

บทนำ

การทดสอบด้วยคอมพิวเตอร์ (Computerized Testing--CT) เป็นการนำคอมพิวเตอร์มาใช้ในการทดสอบแทนการใช้กระดาษคำตอบ โดยผู้เข้าสอบทำแบบทดสอบเดียวกัน ใช้เวลาเท่ากัน และสอบในสถานที่เดียวกัน องค์ประกอบของการทดสอบด้วยคอมพิวเตอร์ต้องมีระบบคลังข้อสอบเพื่อใช้ในการจัดเก็บข้อสอบและบริหารจัดการข้อสอบสำหรับทดสอบ ในขณะที่ความสามารถของระบบทดสอบด้วยคอมพิวเตอร์แต่ละระบบจะแตกต่างกันไปขึ้นอยู่กับวิธีการออกแบบ ความจำเป็น และความต้องการของผู้ใช้ระบบทดสอบนั้นๆ เป็นหลัก

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing--CAT) ประกอบไปด้วยองค์ความรู้ 2 ด้าน คือทฤษฎีการวัดผลการศึกษาและเทคโนโลยีคอมพิวเตอร์ วิธีดำเนินการจัดการการทดสอบจะคัดเลือกข้อสอบจากคลังข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบซึ่งอยู่บนพื้นฐานของการตอบข้อสอบข้อที่ผ่านมาของผู้สอบส่งผลให้ผู้สอบแต่ละคนได้ข้อสอบที่แตกต่างกันและจำนวนข้อสอบที่ไม่เท่ากัน วิธีการดำเนินการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ประกอบไปด้วย 5 ขั้นตอนคือ 1) การสร้างคลังข้อสอบ (create item bank) 2) การคัดเลือกข้อสอบข้อเริ่มต้น (first item selection) เป็นขั้นตอนที่ต้องพิจารณาคัดเลือกข้อสอบข้อเริ่มต้นที่มีความเหมาะสมกับผู้สอบโดยทั่วไป จากงานวิจัยพบว่าการคัดเลือกข้อสอบข้อแรกไม่ค่อยมีความสำคัญหรือมีผลต่อการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Hua & Chang, 2001) โดยทั่วไปจะใช้ข้อสอบที่มีความยากระดับปานกลางเป็นข้อสอบข้อเริ่มต้น 3) การเลือกข้อสอบข้อถัดไป (next item selection) เป็นขั้นตอนที่ต้องพิจารณาคัดเลือกข้อสอบตามเงื่อนไขที่กำหนด โดยอาศัยผลการตอบข้อสอบข้อก่อนหน้ามาพิจารณาด้วย 4) การประมาณค่าความสามารถของข้อสอบ (calculate possible ability levels) เป็นการประมาณค่าความสามารถของผู้สอบหลังจากตอบข้อสอบแล้ว ซึ่งวิธีการประมาณค่าดังกล่าวที่นิยมใช้กันคือ Maximum Likelihood Estimation, Bayesian Estimation และ Bayesian Updating Estimation และ 5) เกณฑ์การยุติการทดสอบ (Termination Criterion) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะสิ้นสุดลงเมื่อผลการตอบข้อสอบข้อที่ผ่านมาเป็นไปตามเกณฑ์ยุติการทดสอบ เกณฑ์ยุติการทดสอบอาจเป็นจำนวนข้อสอบที่กำหนดไว้ล่วงหน้า เวลาที่ใช้ในการทดสอบคะแนนจุดตัดแยกระดับความสามารถของกลุ่มผู้สอบ ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ เป็นต้น (Reckase, 2009, pp. 311-313)

ความแตกต่างระหว่างการทดสอบด้วยคอมพิวเตอร์กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีประสิทธิภาพเหนือกว่าการทดสอบด้วยคอมพิวเตอร์ในประเด็นต่อไปนี้ (1) ลดจำนวนข้อสอบที่ใช้ในการทดสอบของแต่ละคน จึงเป็นการประหยัดเวลาของ การทดสอบ ซึ่งสามารถลดจำนวนข้อสอบลงได้ประมาณ 50% หรือถ้าใช้เวลาในการทดสอบเท่ากัน จะทำให้วัดความสามารถได้ละเอียดกว่าหรือหลายมิติกว่า เช่น การวัดความถนัดพหุในด้านต่างๆ (2) ได้สารสนเทศสูงสุดเกี่ยวกับความสามารถของผู้สอบ กล่าวคือความคลาดเคลื่อนของการวัด (Measurement Error) จะลดลงกว่าวิธีการทดสอบด้วยคอมพิวเตอร์ การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบจึงมีประสิทธิภาพเนื่องจากประหยัดเวลา ลดจำนวนข้อสอบลงได้โดยไม่ทำให้คุณภาพของการทดสอบลดลง แต่กลับมีแนวโน้มที่จะเพิ่มคุณภาพของการวัดให้สูงขึ้น ทั้งในแง่ความเที่ยง (Reliability) และความตรง (Validity) อย่างไรก็ตามการทดสอบที่มีคุณภาพสูงขึ้นนี้จะได้มาก็ด้วยความพยายามของการสร้างระบบการทดสอบแนวใหม่และการใช้เทคโนโลยีที่เหมาะสมมาช่วยในการทดสอบ (Thompson & Weiss, 2011)

ในปัจจุบันสถาบันอุดมศึกษาหลายแห่ง ได้นำการทดสอบด้วยคอมพิวเตอร์ไปใช้ในการทดสอบวัดความรู้วิชาภาษาอังกฤษ ของนิสิตนักศึกษา ไม่ว่าจะเป็นการทดสอบเพื่อวัดผลและประเมินผลประจำภาคเรียน หรือการใช้ในการทดสอบเพื่อวัดความรู้ภาษาอังกฤษก่อนที่นิสิตนักศึกษาจะจบการศึกษา เนื่องจากมีความสะดวกในการบริหารจัดการการทดสอบ

รวมทั้งสามารถตรวจสอบข้อสอบของนักศึกษาจำนวนมากได้อย่างมีประสิทธิภาพในเวลาอย่างรวดเร็ว ในขณะที่เดียวกับการนำระบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้ทดสอบก็มีข้อดีดังที่ได้กล่าวมาแล้วข้างต้น ดังนั้นหากจะใช้ระบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ทดแทนการทดสอบด้วยคอมพิวเตอร์ เพื่อประหยัดเวลาสอบ ลดความเมื่อยล้าเบื่อหน่ายของผู้สอบ โดยการลดจำนวนข้อสอบลง แต่ไม่ทำให้ประสิทธิภาพของการทดสอบลดลงนั้น จึงต้องทดลองทดสอบด้วยระบบทดสอบทั้ง 2 ประเภทดังกล่าวเพื่อหาความสัมพันธ์ของผลสอบว่าสอดคล้องกันหรือไม่

ดังนั้นผู้วิจัยจึงต้องการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ สำหรับการทดสอบรายวิชาภาษาอังกฤษว่ามีความสัมพันธ์เป็นไปในทิศทางเดียวกันหรือไม่ อย่างไร ซึ่งการดำเนินการดังกล่าวอ้างอิงการตามกระบวนการวิจัยของ (Liley, Barker, and Britton ,2004)

วัตถุประสงค์ของการวิจัย

เพื่อศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ สำหรับการทดสอบรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ระดับปริญญาตรี

วิธีดำเนินการวิจัย

วิธีการดำเนินการวิจัยเริ่มต้นจากการพัฒนาระบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และระบบทดสอบด้วยคอมพิวเตอร์ในรูปแบบของ Web Application ซึ่งรายละเอียดที่เกี่ยวข้องกับอัลกอริทึม (algorithm) ในการพัฒนาระบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ของงานวิจัยนี้ มีดังนี้

1) การคัดเลือกข้อสอบข้อเริ่มต้น (first item selection) เลือกข้อสอบที่มีระดับความยากระดับปานกลาง เพื่อให้การประมาณค่าความสามารถของผู้สอบได้ถูกต้องมากขึ้น (Thompson & Weiss, 2011)

2) การคัดเลือกข้อสอบข้อถัดไป (next item selection) ใช้เกณฑ์สารสนเทศสูงสุด (Maximum Information Criterion) เป็นวิธีการเลือกข้อสอบที่เป็นที่นิยม (Linden & Glas, 2010, pp.11) เนื่องจากเป็นการคัดเลือกข้อสอบข้อถัดไปที่น่าปัจจัยของการเดาข้อสอบมาพิจารณาร่วมกับค่าความสามารถของผู้สอบขณะนั้น ซึ่งจะคัดเลือกข้อสอบที่มีค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบมาเป็นข้อสอบข้อถัดไป ค่าสารสนเทศสูงสุดของข้อสอบสามารถคำนวณได้จากสมการที่ 1

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

เมื่อ m_i คือค่าสารสนเทศสูงสุด, $D=1.70$, $\ln$ คือ Natural logarithm, a_i คือค่าอำนาจจำแนกของข้อสอบ, b_i คือค่าความยากของข้อสอบ, c_i คือค่าการเดาของข้อสอบ

3) การประมาณค่าความสามารถของผู้สอบ (Ability Estimation of Examinees) เป็นการประมาณค่าความสามารถของผู้สอบหลังจากตอบข้อสอบแล้ว ซึ่งใช้การประมาณค่าด้วยวิธีของเบส์ปรับใหม่ (Bayesian Updating) สามารถคำนวณได้จากสมการที่ 2-5 (De Ayala, 2013)

(1) ในกรณีที่ตอบข้อสอบถูก

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

$$\sigma_{m+1}^2 = \sigma_m^2 \left(1 - \left(\frac{1-c}{1+\frac{1}{a^2\sigma_m^2}} \right) \left(\frac{O(D)}{B} \right) \left(\frac{(1-c)O(D)}{B} - D \right) \right) \quad (3)$$

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

และ $B = c + (1 - c) \times A(-D)$

(2) ในกรณีที่ตอบข้อสอบผิด

$$\theta_{m+1} = \theta_m - \left(\frac{\sigma_m^2}{\sqrt{\frac{1}{a^2} + \sigma_m^2}} \right) \left(\frac{O(D)}{A(D)} \right) \quad (4)$$

$$\sigma_{m+1}^2 = \sigma_m^2 \left(1 - \left(\frac{O(D)}{1+\frac{1}{a^2\sigma_m^2}} \right) \left(\frac{O(D)}{A(D)} + D \right) \div A(D) \right) \quad (5)$$

เมื่อ

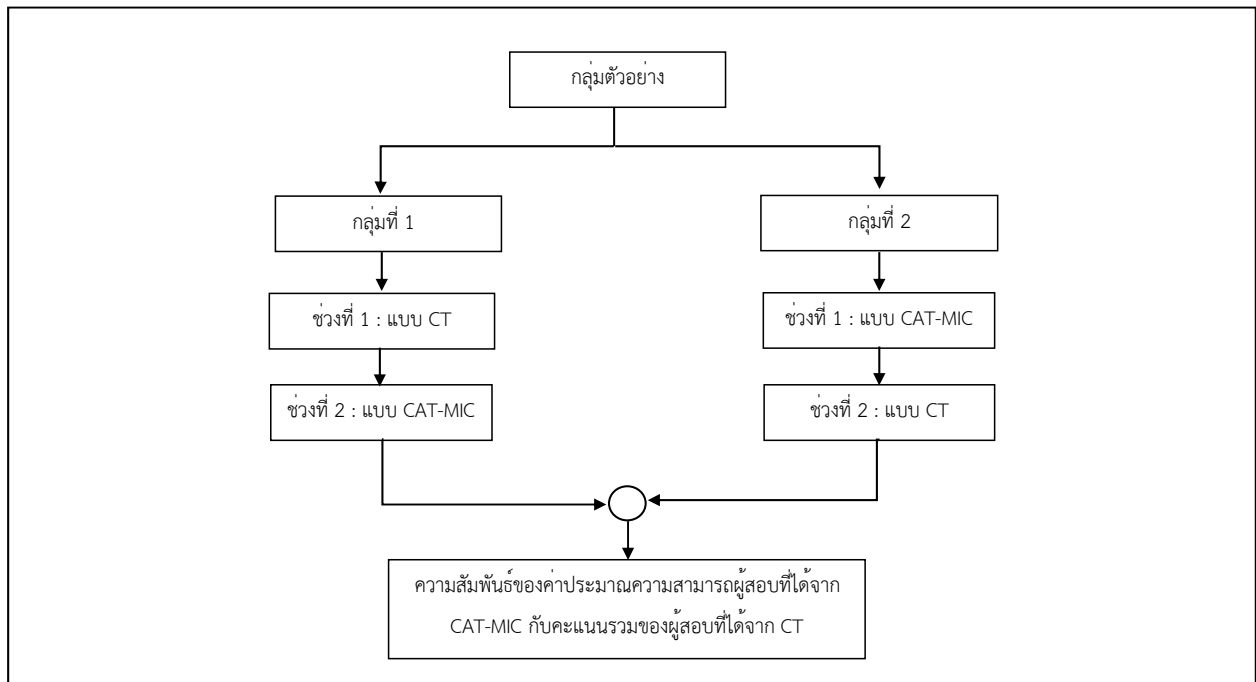
θ_m	=	ความสามารถของผู้สอบที่ประมาณได้ก่อนตอบข้อสอบที่ $m + 1$ ตามปกติ แล้วเมื่อเริ่มทำข้อสอบจะไม่ทราบค่าความสามารถเบื้องต้นของผู้สอบจึงกำหนดให้ θ_m เท่ากับ 0.000
σ_m^2	=	ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อสอบข้อที่ $m + 1$ ตามปกติแล้วเมื่อเริ่มทำข้อสอบจะไม่ทราบค่าความแปรปรวนดังกล่าวมาก่อน จึงกำหนดให้ σ_m^2 เท่ากับ 1.000
θ_{m+1}	=	ความสามารถของผู้สอบที่ประมาณได้หลังจากตอบข้อสอบข้อที่ $m + 1$
σ_{m+1}^2	=	ความแปรปรวนในการประมาณค่าความสามารถของหลังจากตอบข้อสอบข้อที่ $m + 1$
a_i	=	ค่าอำนาจจำแนกของข้อสอบข้อที่ $m + 1$
b_i	=	ค่าความยากของข้อสอบข้อที่ $m + 1$
c_i	=	ค่าโอกาสในการเดาข้อสอบถูกของข้อสอบข้อที่ $m + 1$
D	=	จุดบนแกน X
$O(D)$	=	ค่าออร์ดิเนต (Ordinate) ของโค้งปกติที่จุด D
$A(D)$	=	พื้นที่ใต้โค้งปกติจากค่า D ต่ำสุดจนถึงจุด D

4) เกณฑ์การยุติการทดสอบ (Termination Criterion) ผู้สอบจะต้องทำแบบทดสอบไปเรื่อยๆ จนกว่าค่าความคลาดเคลื่อนมาตรฐานมีค่าลดลงถึงระดับที่กำหนดการทดสอบจึงยุติ โดยกำหนดค่าความคลาดเคลื่อนมาตรฐานน้อยกว่า 0.30 (Thompson & Weiss, 2011) การหาความคลาดเคลื่อนมาตรฐาน สามารถคำนวณได้จากสมการที่ 6

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (6)$$

เมื่อ $SSE(\theta)$ = ความคลาดเคลื่อนมาตรฐานของการประมาณค่า θ
 $I(\theta)$ = สารสนเทศของแบบสอบที่ให้สำหรับผู้มีความสามารถ θ

จากนั้นศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (CAT-MIC) กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT) ซึ่งสามารถแสดงขั้นตอนการดำเนินการได้ตามภาพ 1



ภาพ 1 ขั้นตอนการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์

จากภาพ 1 แสดงขั้นตอนการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบ CAT-MIC กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบแบบ CT ซึ่งผลการศึกษานี้จะแสดงให้เห็นว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีการวัดผลที่สอดคล้องเป็นไปในทิศทางเดียวกับการทดสอบด้วยคอมพิวเตอร์หรือไม่ ด้วยเหตุนี้ นักศึกษาแต่ละคนในกลุ่มตัวอย่าง จึงต้องทำการทดสอบทั้ง 2 แบบ โดยใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นดังกล่าวข้างต้นเป็นเครื่องมือที่ใช้ในการวิจัย

การดำเนินการวิจัยที่เกี่ยวข้องกับการทดสอบ ดำเนินการโดยแบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม โดยกลุ่มที่ 1 ดำเนินการทดสอบแบบ CT ก่อน แล้วจึงทำการทดสอบแบบ CAT-MIC ส่วนในกลุ่มที่ 2 ดำเนินการทดสอบแบบ CAT-MIC ก่อน แล้วจึงทำการทดสอบแบบ CT ทั้งนี้เพื่อป้องกันปัญหาเรื่องลำดับของการทดสอบ (Order Effect) ซึ่งอาจส่งผลต่อผลการทดสอบได้ หลังจากนั้นจึงนำผลที่ได้รับจากการทดสอบทั้ง 2 แบบ มาทดสอบความสัมพันธ์ ซึ่งมีรายละเอียดดังนี้

1. กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในงานวิจัยนี้เป็นนักศึกษาชั้นปีที่ 3 คณะการบัญชี มหาวิทยาลัยอีสเทิร์นเอเซีย คัดเลือกตัวอย่างแบบตามสะดวกจำนวน 30 คน

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย คือ ระบบทดสอบรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ระดับปริญญาตรี ซึ่งเป็นโปรแกรมในรูปแบบ Web Application ซึ่งผู้สอบสามารถเลือกการทดสอบได้ 2 ประเภท คือ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และการทดสอบด้วยคอมพิวเตอร์

3. การเก็บรวบรวมข้อมูล

ดำเนินการทดสอบโดยใช้ระบบทดสอบที่ผู้วิจัยได้พัฒนาขึ้นดังกล่าวทดสอบรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ระดับปริญญาตรี มีขั้นตอนในการดำเนินการดังนี้

ขั้นตอนที่ 1 การจัดเตรียมคลังข้อสอบ ในการทดสอบนี้ใช้คลังข้อสอบที่เป็นข้อสอบในรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ โดยคลังข้อสอบทั้งหมดได้รับความอนุเคราะห์จากอาจารย์ประจำหลักสูตรภาษาอังกฤษ มหาวิทยาลัยสวนดุสิต ซึ่งข้อสอบแต่ละข้อ แบ่งได้ 2 ส่วน ได้แก่

ส่วนที่ 1 เนื้อหาของข้อสอบ ประกอบด้วย โจทย์ รายการคำตอบแบบ 4 ตัวเลือก พร้อมเฉลย ตัวเลือกที่ถูกต้อง 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

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

จากนั้นให้ผู้เชี่ยวชาญด้านภาษาจากหลักสูตรภาษาอังกฤษ มหาวิทยาลัยสวนดุสิต คัดเลือกข้อสอบที่เป็นตัวแทนของเนื้อหาการทดสอบทั้ง 4 ส่วนดังกล่าว ออกมาส่วนละ 10 ข้อ รวมทั้งสิ้น 40 ข้อเพื่อใช้ในการทดสอบประเภทการทดสอบด้วยคอมพิวเตอร์ (CT) ในขณะทำการทดสอบประเภท CAT-MIC ใช้ข้อสอบทั้งหมดที่อยู่ในคลังข้อสอบในการทดสอบ

ขั้นตอนที่ 2 ดำเนินการทดสอบความรู้รายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ระดับปริญญาตรี มีขั้นตอนย่อยในการดำเนินการดังนี้

ขั้นตอนที่ 2.1 แบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม กลุ่มละ 15 คน

ขั้นตอนที่ 2.2 ชี้แจงวัตถุประสงค์ของการทดสอบและวิธีการใช้โปรแกรมอย่างละเอียดเพื่อให้กลุ่มตัวอย่างเห็นความสำคัญของการทดสอบ และทำการทดสอบด้วยความตั้งใจ

ขั้นตอนที่ 2.3 ดำเนินการทดสอบรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ระดับปริญญาตรี โดยกลุ่มตัวอย่างทั้ง 2 กลุ่ม ต้องดำเนินการทดสอบทั้ง 2 แบบ คือ การทดสอบแบบ CT และการทดสอบ CAT-MIC ซึ่งกลุ่มที่ 1 ดำเนินการทดสอบแบบ CT ก่อน แล้วจึงดำเนินการทดสอบแบบ CAT-MIC ส่วนกลุ่มที่ 2 ดำเนินการทดสอบแบบ CAT-MIC ก่อน แล้วจึงดำเนินการทดสอบแบบ CT แสดงได้ดังตาราง 1

ตาราง 1 ลำดับการทดสอบ

	สอบช่วงที่ 1	พัก	สอบช่วงที่ 2
	9.00 น. - 11.00 น.		13.00 น. - 15.00 น.
กลุ่มที่ 1	แบบ CT	2 ชั่วโมง	แบบ CAT-MIC
กลุ่มที่ 2	แบบ CAT-MIC		แบบ CT

4. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลของการวิจัย จะดำเนินการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ใช้สถิติเชิงบรรยาย ได้แก่ ความถี่ (frequency) ค่าสูงสุด (max) ค่าต่ำสุด (min) ค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (standard deviation) ทั้งนี้ค่าประมาณความสามารถของผู้สอบและคะแนนรวมของผู้สอบเป็นข้อมูลแบบต่อเนื่อง จึงใช้การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน (Pearson's Product Moment Correlation)

5. เกณฑ์การประเมินผลการทดสอบแสดงได้ดังตาราง 2-3 (Han, 2012)

ตาราง 2 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

ตาราง 3 เกณฑ์ประเมินระดับความสามารถของผู้สอบ 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

ผลการวิจัย

ผลการพัฒนาระบบทดสอบรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ระดับปริญญาตรี ซึ่งเป็นโปรแกรมในรูปแบบ Web Application สามารถแสดงตัวอย่างโปรแกรมบางส่วนได้ดังภาพ 2-6

The screenshot shows a web application interface for exam preparation. At the top, there is a navigation bar with links: Home, Account, Exam, About Us, Contact Us, and a Sign Out button. Below the navigation bar, the main heading is "การทำแบบทดสอบ" (Exam Preparation). Under this heading, user information is displayed: รหัส: 41112003, ชื่อ-นามสกุล: นาย รัชกฤษ ธนพัฒน์, สาขา: วิศวกรรมคอมพิวเตอร์, คณะ: วิศวกรรมศาสตร์, มหาวิทยาลัย: KMITL. Below the user information, there is a section for selecting the exam type. It includes a dropdown menu for "ปีทดสอบ" (Exam Year) set to 2018, a dropdown menu for "วิชา" (Subject) set to ภาษาอังกฤษ, and radio buttons for "ประเภทการทดสอบ" (Exam Type) with options CT and CAT. A blue button labeled "ทำข้อสอบ" (Take Exam) is located at the bottom of the form.

ภาพ 2 หน้าโปรแกรมหลักการทำแบบทดสอบ

จากภาพ 2 ให้ผู้สอบเลือกวิชาสอบ และประเภทการทดสอบ จากนั้นคลิกเมนู“ทำข้อสอบ” การทำแบบทดสอบผู้สอบสามารถเลือกประเภทการทดสอบได้ 2 แบบดังนี้

- 1) การทำแบบทดสอบของการทดสอบแบบ CAT
- 2) การทำแบบทดสอบของการทดสอบแบบ CT

หากผู้สอบเลือกประเภทการทดสอบแบบ CAT โปรแกรมจะปรากฏหน้าจอแบบทดสอบตามวิชาที่เลือก ดังภาพ 3

ความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ สำหรับการทดสอบรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ในระดับปริญญาตรี

The screenshot shows the CAT test interface. At the top, there is a navigation bar with links: Home, Account, Exam, About Us, Contact Us, and a Sign Out button. The main heading is "การทดสอบแบบ CAT". Below this, there is a question: "โจทย์ ข้อที่ 2. Don't be **cowed** by other's attitudes. Speak your mind." There are four radio button options: 1. resolute, 2. **determined** (selected), 3. bullied, and 4. concluded. At the bottom, there are two buttons: "<< ข้อที่แล้ว" (Previous Question) and "ข้อถัดไป >>" (Next Question).

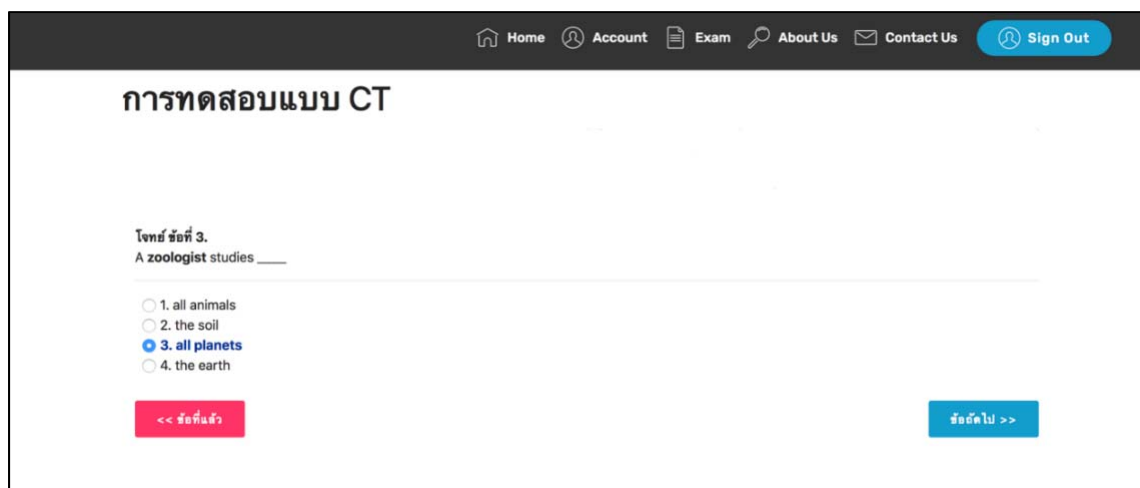
ภาพ 3 หน้าโปรแกรมการทำแบบทดสอบตามวิชาที่เลือกสอบของการทดสอบประเภท CAT

จากภาพ 3 การทำแบบทดสอบของการทดสอบประเภท CAT เมื่อผู้สอบคลิกตัวเลือกของข้อสอบแต่ละข้อแล้วให้คลิกที่เมนู“ข้อถัดไป>>”เพื่อทำข้อสอบข้อถัดไป หากต้องการแก้ไขข้อสอบที่ทำไปแล้วข้อก่อนหน้าสามารถคลิกที่เมนู“<<ข้อที่แล้ว”ให้ผู้สอบทำข้อสอบไปเรื่อยๆจนกระทั่งโปรแกรมขึ้นเมนู“ส่งกระดาษคำตอบ” จากนั้นให้ผู้สอบคลิกส่งกระดาษคำตอบแล้วโปรแกรมจะปรากฏ ดังภาพ 4

ผลการทดสอบแบบ CAT				
ปีที่ทดสอบ: 2018 วิชา: ภาษาอังกฤษ ประเภทการทดสอบ: CAT				
ชื่อแบบทดสอบ: ข้อสอบวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ในระดับปริญญาตรี ปี 2018				
Student Name	: รัชกฤษ ธนพัฒน์	Student No.	: 41112003	
Faculty	: วิศวกรรมศาสตร์	Major	: วิศวกรรมคอมพิวเตอร์	
Date	: 26/08/2018	Time	: 0 Hours 3 Minutes 25 Seconds	
Testing	: 20 Items	Correct	: 6	
Ability Estimate	: -1.4565	Ability Level	: Rather Low	
ชุดที่	ชื่อชุด	จำนวนข้อทั้งหมด	จำนวนข้อถูก	ถูกร้อยละ
1	Part of Speech	5	1	20.00
2	Using Context Clues	5	2	40.00
3	Word Formation	5	0	0.00
4	Making Inferences	5	3	60.00
รวม		20	6	30.00

ภาพ 4 หน้าโปรแกรมสรุปผลการสอบรายบุคคลของการทดสอบประเภท CAT

จากภาพ 4 โปรแกรมจะแสดงรายละเอียดข้อมูลส่วนตัวของผู้สอบ วันที่ทำข้อสอบ เวลาที่ใช้ในการทำข้อสอบ ค่าประมาณความสามารถของผู้สอบ ระดับความสามารถของผู้สอบ จำนวนข้อสอบที่ได้รับ รวมถึงจำนวนข้อที่ผู้สอบทำได้อีกต้อง หากผู้สอบเลือกประเภทการทดสอบแบบ CT โปรแกรมจะปรากฏหน้าจอแบบทดสอบตามวิชาที่เลือก ดังภาพ 5



ภาพ 5 หน้าโปรแกรมการทำแบบทดสอบตามวิชาที่เลือกสอบของการทดสอบประเภท CT

จากภาพ 5 การทำแบบทดสอบของการทดสอบประเภท CT เมื่อผู้สอบคลิกตัวเลือกของข้อสอบแต่ละข้อแล้วให้คลิกที่เมนู“ข้อถัดไป>>”เพื่อทำข้อสอบข้อถัดไป หรือคลิกที่เมนู“<<ข้อที่แล้ว”หากต้องการแก้ไขการทำข้อสอบข้อก่อนหน้า และให้ผู้สอบทำข้อสอบไปจนกระทั่งโปรแกรมขึ้นเมนู“ส่งกระดาษคำตอบ” จากนั้นให้ผู้สอบคลิกส่งกระดาษคำตอบ แล้วโปรแกรมจะปรากฏ ดังภาพ 6

ผลการทดสอบแบบ CT				
ปีที่ทดสอบ: 2018 วิชา: ภาษาอังกฤษ ประเภทการทดสอบ: CT				
ชื่อแบบทดสอบ: ข้อสอบวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ในระดับปริญญาตรี ปี 2018				
Student Name	: รัชกฤษ อนุพัฒน์	Student No.	: 41112003	
Faculty	: วิศวกรรมศาสตร์	Major	: วิศวกรรมคอมพิวเตอร์	
Date	: 26/08/2018	Time	: 0 Hours 9 Minutes 30 Seconds	
Testing	: 40 Items	Correct	: 10	
Ability Estimate	: N/A	Ability Level	: Very Low	
ชุดที่	ชื่อชุด	จำนวนข้อทั้งหมด	จำนวนข้อถูก	ถูกร้อยละ
1	Part of Speech	10	4	40.00
2	Using Context Clues	10	2	20.00
3	Word Formation	10	3	30.00
4	Making Inferences	10	1	10.00
รวม		40	10	25.00

ภาพ 6 หน้าโปรแกรมสรุปผลการสอบของผู้สอบของการทดสอบประเภท CT

จากภาพ 6 โปรแกรมจะแสดงรายละเอียดข้อมูลส่วนตัวของผู้สอบ วันที่ทำข้อสอบ เวลาที่ใช้ในการทำข้อสอบ ระดับความสามารถของผู้สอบ จำนวนข้อสอบที่ได้รับ รวมถึงจำนวนข้อที่ผู้สอบทำได้ถูกต้อง

ผลการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ได้ผลการวิจัยดังนี้

ตาราง 4 ความยาวของแบบทดสอบ เวลาที่ใช้ในการทดสอบ และผลการทดสอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์

คนที่	ความยาวของแบบทดสอบ (ข้อ)		เวลาที่ใช้ในการทดสอบ (นาที: วินาที)		ผลการทดสอบ	
	CAT-MIC	CT	CAT-MIC	CT	CAT-MIC (θ)	CT (คะแนนรวม)
1	15	40	11:24	27:24	-0.46	23
2	20	40	13:22	25:44	0.41	22
3	18	40	07:36	15:20	-0.78	18
4	15	40	11:18	26:59	-0.63	17
5	17	40	18:48	27:19	1.34	28
6	18	40	15:56	24:26	-1.18	17
7	19	40	12:19	21:34	0.20	20
8	15	40	17:22	22:31	-1.20	18
9	16	40	09:16	13:15	-0.89	19
10	15	40	18:12	27:50	0.56	25
11	15	40	13:20	34:58	-0.46	23
12	15	40	16:50	31:41	1.42	30
13	18	40	10:47	13:05	-0.88	19
14	20	40	07:54	16:33	-0.12	20
15	15	40	14:30	21:30	0.29	23
16	20	40	09:41	11:26	-1.15	16
17	15	40	13:50	21:23	-1.29	16
18	18	40	21:36	30:48	0.21	20
19	15	40	06:24	37:28	-1.14	19
20	15	40	15:38	18:37	-0.05	30
21	19	40	07:32	20:16	-0.55	19
22	18	40	13:32	26:06	0.75	31
23	15	40	16:10	26:57	-0.22	20
24	15	40	09:28	16:11	0.58	24
25	20	40	16:28	21:17	1.00	24
26	15	40	18:32	25:07	1.01	28
27	15	40	14:28	29:01	1.11	27
28	15	40	09:33	20:11	-0.64	19
29	18	40	15:08	23:24	0.24	22
30	18	40	17:32	22:17	-0.81	16
Min	15	40	06:24	11:26	-1.29	16
Max	20	40	21:36	37:28	1.42	31
Mean	16.73	40	13:20	23:21	-0.11	21.77
SD	1.95	0.00	03:55	06:19	0.83	4.45

หมายเหตุ CAT-MIC หมายถึง การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้การคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด
CT หมายถึง การทดสอบด้วยคอมพิวเตอร์

จากตาราง 4 แสดงความยาวของแบบทดสอบ เวลาที่ใช้ในการทดสอบ และผลการทดสอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับ การทดสอบด้วยคอมพิวเตอร์ ปรากฏว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีความยาวของของแบบทดสอบเฉลี่ย 16.73 ข้อ ในขณะที่การทดสอบด้วยคอมพิวเตอร์มีความยาวของแบบทดสอบ 40 ข้อ แสดงให้เห็นว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด มีความยาวของแบบทดสอบคิดเป็นร้อยละ 41.83 ของการทดสอบด้วยคอมพิวเตอร์

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใช้เวลาในการทดสอบเฉลี่ย 13:20 นาที ขณะที่การทดสอบด้วยคอมพิวเตอร์ใช้เวลาในการทดสอบเฉลี่ย 23:21 นาที แสดงให้เห็นว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ใช้เวลาในการทดสอบคิดเป็นร้อยละ 56.87 ของการทดสอบด้วยคอมพิวเตอร์

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้ค่าประมาณความสามารถของผู้สอบ ($\hat{\theta}$) เฉลี่ยเท่ากับ -0.11 หมายความว่า ผู้สอบมีความสามารถโดยเฉลี่ยอยู่ในระดับปานกลาง (เกณฑ์การประเมินระดับความสามารถของผู้สอบ สำหรับการทดสอบแบบ CAT แสดงในตาราง 2) เช่นเดียวกับการทดสอบด้วยคอมพิวเตอร์ได้คะแนนรวมเฉลี่ยเท่ากับ 21.77 หมายความว่า ผู้สอบมีความสามารถโดยเฉลี่ยอยู่ในระดับปานกลาง (เกณฑ์การประเมินความสามารถของผู้สอบ สำหรับการทดสอบแบบ CT แสดงในตาราง 3) แสดงให้เห็นว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดให้ผลการทดสอบใกล้เคียงกับการทดสอบด้วยคอมพิวเตอร์

ตาราง 5 ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างค่าประมาณความสามารถของผู้สอบกับคะแนนรวมของผู้สอบ

	CAT-MIC	CT
CAT-MIC	1	
CT	.837**	1

** $p < .01$

จากตาราง 5 แสดงค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ปรากฏว่า ค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ มีความสัมพันธ์เชิงบวกขนาดสูงเท่ากับ .837 ที่ระดับนัยสำคัญทางสถิติ .01

สรุปและอภิปรายผล

ผลการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นกับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ จากกลุ่มตัวอย่างจำนวน 30 คน สรุปได้ดังนี้

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT-MIC) กลุ่มตัวอย่างใช้เวลาในการทดสอบเฉลี่ย 13:20 นาที ใช้ข้อสอบเฉลี่ย 16.73 ข้อ และมีค่าประมาณความสามารถของผู้สอบ ($\hat{\theta}$) เฉลี่ย -0.11 ส่วนการทดสอบด้วยคอมพิวเตอร์ (CT) กลุ่มตัวอย่างใช้เวลาในการทดสอบเฉลี่ย 23:21 นาที ใช้ข้อสอบ 40 ข้อ และมีคะแนนรวมของผู้สอบเฉลี่ย 21.77 ข้อ แสดงให้เห็นว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใช้เวลาและจำนวนข้อสอบลดลง เมื่อเทียบกับการทดสอบด้วยคอมพิวเตอร์

ผลการทดสอบความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ปรากฏว่า ผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับผลการทดสอบด้วยคอมพิวเตอร์มีความสัมพันธ์เชิงบวก ขนาดสูง ที่ระดับนัยสำคัญทางสถิติ .01 แสดงว่า ผลการทดสอบแบบ CAT-MIC มีความสอดคล้องกับผลการทดสอบแบบ CT นั่นคือ ผู้สอบที่ได้รับค่าประมาณความสามารถจากการทดสอบแบบ CAT สูง (หรือต่ำ) จะได้คะแนนรวมจากการทดสอบแบบ CT สูง (หรือต่ำ) ด้วย

จากผลการวิจัยข้างต้นพบว่าระบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น สามารถคัดเลือกข้อสอบที่มีความเหมาะสมต่อระดับความสามารถของผู้สอบ ส่งผลให้ประสิทธิภาพของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถประมาณค่าความสามารถของผู้สอบได้ใกล้เคียงกับความสามารถที่แท้จริงของผู้สอบ โดยใช้จำนวนข้อสอบที่ใช้ทดสอบน้อยลงแต่ไม่ลดประสิทธิภาพของการทดสอบเมื่อเปรียบเทียบกับ การทดสอบด้วยคอมพิวเตอร์ที่ผู้สอบต้องทำข้อสอบทุกข้อ ซึ่งสอดคล้องกับ (Phankokkrud & Woraratpanya, 2012) และ (Reckase, 2009, pp. 311-313)

อย่างไรก็ตามการวิจัยนี้เป็นการศึกษาความสัมพันธ์ของระบบทดสอบทั้ง 2 ระบบโดยใช้วิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ระดับปริญญาตรีในการทดสอบ เพื่อให้เห็นทิศทางการสอดคล้องของระบบสอบเท่านั้น ดังนั้นเพื่อให้เกิดความมั่นใจในการใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) ทดแทนการทดสอบด้วยคอมพิวเตอร์ (CT) จึงจำเป็นต้องทดสอบ

กับรายวิชาอื่น ๆ ที่หลากหลาย และกลุ่มผู้สอบที่หลากหลายเพื่อให้เกิดความแม่นยำและความน่าเชื่อถือในการตัดสินผลการ
ทดสอบของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ข้อเสนอแนะ

วิธีการคัดเลือกข้อสอบข้อถัดไปที่ใช้ในงานวิจัยนี้ยังมีจุดอ่อนคือไม่สามารถเลือกข้อสอบได้กระจายทุกค่าความยากของ
ข้อสอบเนื่องจากสมการที่ใช้เลือกข้อสอบขึ้นอยู่กับค่าประมาณความสามารถของผู้สอบในขณะนั้น ส่งผลให้ข้อสอบที่ไม่อยู่ใน
เงื่อนไขอาจไม่ถูกเลือกมาให้ผู้สอบทำการทดสอบเลย ดังนั้นการพัฒนาการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถ
เลือกใช้วิธีการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีอื่นๆที่นอกเหนือจากที่ใช้ในงานวิจัยนี้ได้ เช่น การใช้วิธีทางปัญญาประดิษฐ์มา
ประยุกต์ใช้ในการเลือกข้อสอบข้อถัดไปซึ่งจะส่งผลให้การเลือกข้อสอบกระจายตัวได้ดี และเปิดโอกาสให้ข้อสอบทุกข้อสามารถถูก
เลือกนำมาใช้ในการทดสอบได้

กิตติกรรมประกาศ

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. ดวงเพ็ญ เจตนพิพัฒน์พงษ์ อาจารย์ ดร. ชีรนวิษ์ สุขวิทย์ศิริบุญ และอาจารย์อารียา
ศรีประเสริฐ ที่เป็นผู้เชี่ยวชาญในการประเมินโปรแกรมระบบทดสอบฯ และให้ข้อเสนอแนะในการปรับปรุง จนทำให้โปรแกรม
มีความสมบูรณ์ยิ่งขึ้น รวมทั้งขอขอบพระคุณ คณาจารย์หลักสูตรภาษาอังกฤษ มหาวิทยาลัยสวนดุสิต ที่ให้ความอนุเคราะห์ข้อสอบ
รายวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ในระดับปริญญาตรีมาใช้ในการวิจัยครั้งนี้

บรรณานุกรม

- De Ayala, R. J. (2013). *The theory and practice of item response theory*. London: The Guilford Press.
- Han, K. T. (2012). An efficiency balanced information criterion for item selection in computerized adaptive testing. *Journal of Educational Measurement*, 49(3), 225-246.
- 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.
- Liley, M., Barker, T., & Britton, C. (2004). The development and evaluation of a software prototype for computerized-adaptive testing. *Computers & Education*, 43(43), 109-123.
- Phankokkruad, M., & Woratpanya, K. (2012). Item classification algorithm for computer adaptive testing based on web services. *International Journal of Engineering Research and Applications*, 2(1), 835-847.
- Reckase, M. (2009). *Multidimensional item response theory* (Vol. 150). New York: Springer.
- Thompson, N. A., & Weiss, D. J. (2011). A framework for the development of computerized adaptive tests. *Practical Assessment, Research & Evaluation*, 16(1), 1-9.
- Van der Linden, W. J., & Glas, G. A. W. (2010). *Computerized adaptive testing: theory and practice*. New York: Kluwer Academic.

