จุฬาลงกรณ์มหาวิทยาลัย

เงินอุดหนุนทั่วไปจากรัฐบาล
ประจำปีงบประมาณ 2561

รายงานผลการวิจัย

เรื่อง ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

Title: Computerized Adaptive Testing System in Information and Communication Technology Literacy Skills for 21st Century of Undergraduate Students

โดย

รองศาสตราจารย์ ดร. โชติกา ภาษีผล
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

2561

กิตติกรรมประกาศ

การดำเนินงานของโครงการวิจัยนี้สำเร็จลุล่วงไปด้วยดี โดยได้รับทุนอุดหนุนทั่วไปจากรัฐบาลประจำปีงบประมาณ 2561 ผู้วิจัยขอขอบพระคุณเป็นอย่างสูงในการสนับสนุนงบประมาณในครั้งนี้ ส่วนการดำเนินการในแต่ละส่วนของโครงการนี้ได้รับความร่วมมือจากผู้มีส่วนช่วยเหลือหลายฝ่ายในแต่ละระยะของการดำเนินการวิจัย เริ่มตั้งแต่ คณะกรรมการดำเนินงานสร้างข้อคำถามแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ทั้งจากสาขาวิชาการวัดและประเมินผล และสาขาวิชาเทคโนโลยีการศึกษา ทำให้ได้ข้อสอบที่มีสถานการณ์ที่หลากหลาย และมีความสร้างสรรค์ตามประสบการณ์ของกรรมการแต่ละท่าน ผู้ทรงคุณวุฒิด้านการวัดและประเมินผล และด้านเทคโนโลยีการศึกษาที่ให้ความกรุณาตรวจสอบความสอดคล้องระหว่างองค์ประกอบ นิยามและข้อคำถามของแบบวัด อีกทั้งตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ผู้วิจัยจึงขอขอบคุณทุกท่านที่ให้ความร่วมมือเป็นอย่างดี และให้ข้อเสนอแนะอันมีประโยชน์ในการแก้ไขปรับปรุงแบบวัดและระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ให้มีคุณภาพมากยิ่งขึ้น

ขอขอบคุณคณาจารย์ในสถาบันอุดมศึกษาที่เป็นตัวอย่างวิจัยทุกท่านที่ให้ความร่วมมือและช่วยติดต่อประสานงานในการเก็บรวบรวมข้อมูล และที่สำคัญที่จะขาดไม่ได้ คือตัวอย่างนิสิตนักศึกษาระดับปริญญาตรีที่เป็นตัวอย่างวิจัยทุกคนที่ให้ความร่วมมือในทดลองใช้แบบวัด และระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รวมทั้งประเมินความพึงพอใจต่อระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ทำให้ได้ข้อมูลที่เป็นประโยชน์ในการปรับปรุงแก้ไขแบบวัดและระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ให้มีคุณภาพมากยิ่งขึ้น

ผู้วิจัยหวังเป็นอย่างยิ่งว่า ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ที่พัฒนาขึ้นจะเป็นนวัตกรรมทางการวัดและประเมินผลการศึกษาที่เป็นประโยชน์ในการพัฒนาคุณลักษณะของนิสิตนักศึกษาของประเทศให้มีความพร้อมในทักษะด้านนี้ และสามารถนำงานวิจัยนี้ไปพัฒนาต่อยอดงานวิจัยอื่นๆที่เป็นประโยชน์ต่อประเทศต่อไป

เลขหมู่
เลขทะเบียน 018073
วัน เดือน ปี 18 มี.ค. 62

บทคัดย่อ

การวิจัยเรื่องนี้มีวัตถุประสงค์หลัก 3 ประการ ได้แก่ 1) เพื่อพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี 2) เพื่อพัฒนาข้อสอบสำหรับจัดทำคลังข้อสอบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี และ 3) เพื่อพัฒนาและตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยแบ่งการดำเนินงานวิจัยเป็น 5 ระยะ ได้แก่ ระยะที่ 1 การพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ฯ ระยะที่ 2 การพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21ฯ ระยะที่ 3 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ฯ ระยะที่ 4 การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ฯ ไปทดลองใช้ และระยะที่ 5 การตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ฯ กลุ่มตัวอย่างที่ทดลองใช้แบบวัด ได้แก่ นิสิตนักศึกษาระดับปริญญาตรีจากสถาบันอุดมศึกษาของรัฐบาลในกรุงเทพมหานคร จำนวนทั้งสิ้น 1,672 คน กลุ่มตัวอย่างที่ทดลองใช้ระบบการทดสอบฯ ได้แก่ นิสิตนักศึกษาระดับปริญญาตรีจากสถาบันอุดมศึกษาของรัฐบาลในกรุงเทพมหานคร จำนวน 217 คน การวิเคราะห์ข้อมูลเชิงปริมาณ ใช้ ค่าร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน t test การวิเคราะห์คุณภาพของข้อสอบตามทฤษฎี CTT และ IRT การวิเคราะห์ข้อมูลเชิงคุณภาพ ใช้ การวิเคราะห์เนื้อหา (content analysis) ผลการวิจัยสรุปได้ ดังนี้

1.องค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ประกอบด้วย 5 องค์ประกอบ ได้แก่ 1) การเข้าถึงสารสนเทศ (information accessibility) 2) การจัดการสารสนเทศ (information management) 3) การบูรณาการสารสนเทศ (information integration) 4) การประเมินสารสนเทศ (information evaluation) 5) การสื่อสารสารสนเทศ (information communication)

2. การสร้างข้อสอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ตามนิยามเชิงปฏิบัติการเป็นข้อสอบเลือกตอบ 5 ตัวเลือก องค์ประกอบละ 52 ข้อ รวมทั้งสิ้น 260 ข้อ ข้อสอบส่วนใหญ่ผ่านการประเมินความตรงเชิงเนื้อหา คิดเป็นร้อยละ 79.61 ผลการวิเคราะห์ค่าความเที่ยง (R) ของแบบสอบทั้ง 10 ชุดมีค่าระหว่าง .550 ถึง .725 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อตามทฤษฎี IRT พบว่า โมเดลเอกมิติ 2PL เป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลคำตอบของแบบสอบทุกฉบับมากที่สุด ผลการคัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบ ได้ข้อสอบจำนวน 212 ข้อ โดยมีค่าเฉลี่ยพารามิเตอร์ความยากข้อสอบเท่ากับ .18 (SD.=2.06) และค่าเฉลี่ยพารามิเตอร์การจำแนกเท่ากับ .70 (SD.=.39) ในภาพรวมสรุปได้ว่าข้อสอบที่นำเข้าคลังข้อสอบแบบปรับเหมาะเป็นข้อสอบยากปานกลางและจำแนกได้ปานกลาง

3. การพัฒนาระบบในการวิจัยครั้งนี้ใช้ภาษา PHP เนื่องจากเป็นระบบการทดสอบที่สามารถใช้งานผ่านระบบออนไลน์ โปรแกรมที่พัฒนาประกอบด้วย 4 ส่วน ได้แก่ 1) การลงทะเบียน 2) การสร้างชุดข้อสอบ 3) การทดสอบ ซึ่งประกอบด้วย การประมาณค่าความสามารถ การคัดเลือกข้อสอบ และการยุติการทดสอบ 4) การรายงานผลการทดสอบ และมีการพัฒนาคู่มือการใช้ระบบการทดสอบฯ แบ่งออกเป็น 2 ส่วน คือ 1) คู่มือการใช้งานสำหรับผู้ใช้งานทั่วไป (user) และ 2) คู่มือการใช้งานสำหรับผู้จัดการระบบ (admin)

4. การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้กับนิสิตนักศึกษาสายสังคมศาสตร์ และสายวิทยาศาสตร์ พบว่าไม่ว่านิสิตนักศึกษาชายหรือหญิงมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด แต่เมื่อพิจารณาทักษะด้านที่สูงสุด พบว่า นิสิตนักศึกษาชายมีคะแนนเฉลี่ยด้านการเข้าถึงสารสนเทศสูงที่สุด ในขณะที่นิสิตนักศึกษาหญิงมีคะแนนเฉลี่ยด้านการสื่อสารสารสนเทศสูงที่สุด

5.ผลการตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยผู้ทรงคุณวุฒิ พบว่าในภาพรวม และในแต่ละด้านย่อย ได้แก่ ความมีประโยชน์ ความเป็นไปได้ในการนำไปใช้ ความเหมาะสม และความถูกต้อง มีคุณภาพอยู่ในระดับมากที่สุด รวมทั้งผลการประเมินความพึงพอใจของนิสิตนักศึกษาต่อระบบการทดสอบด้านภาพรวมของระบบ หน้าจอของระบบ การเรียนรู้ของผู้ใช้งาน และความสามารถของระบบ อยู่ในระดับมากเช่นกัน

[28202843

Abstract

This study has three main purposes, including (1) to develop components for measuring information and communication technology (ICT) literacy of undergraduate students in the 21st century, (2) to develop items for an item bank of the ICT literacy of the undergraduate students in the 21st century, and (3) to develop and examine the quality of a computerized adaptive testing (CAT) system for measuring the ICT literacy of the undergraduate students in the 21st century. Researchers designed research procedures including 5 phrases: First phrase was developing the components of ICT literacy for the 21st century; Second phrase was developing items and an item bank for measuring the ICT literacy for the 21st century; Third phrase was developing a CAT system for measuring the ICT literacy for the 21st century; Forth phrase was piloting the CAT system; and Fifth phrase was examining the quality of the CAT system. Data was collected from 1,672 undergraduate students for the piloting and 217 students for testing the system. All participants attended public universities in Bangkok. The data were analyzed qualitatively and quantitatively including content analysis, descriptive analysis, t-test, and classical test theory and item response theory modeling. The results were described as follows.

(1) There were five components of the ICT literacy including (1) Information accessibility, (2) Information management, (3) Information integration, (4) Information evaluation, and (5) Information communication.

(2) There were 260 Initial ICT literacy items created according to the definitions of the five components (52 items each). Most initial items (79.61%) showed content validity evidence by content expert judgment. After the item revision, the items were assembled to ten test forms and their Cronbach reliabilities ranged between .550 and .725. Two-parameter item response models were best fit for all test forms. The revised 212 items (with a mean of item difficulty parameters of .18 and a mean of item discrimination parameters of .70) qualified according to item quality criterions and stored in the item bank of CAT system. This means that overall the difficulty and discrimination of qualified items were moderate.

(3) The CAT system was designed for online purpose using PHP. The system had four stages: (1) Registration, (2) Test assembly design, (3) Test delivery, including parameter estimation, item selection, and test termination, and (4) Test score report. Accompanying manuals of the CAT system were provided for users and system administers.

(4) The results from piloting the CAT system showed that both female and male students in social sciences and science had the lowest mean scores on the information management component. However, females had the highest mean score on the information communication component, while males had the highest one on the information accessibility component.

(5) The CAT system was evaluated by a group of experts about its utility, practical uses, appropriateness, and accuracy. The evaluation results showed that the quality of the system reached the highest levels for all aspects. Moreover, the students were satisfied the functions of CAT system in aspects of good display, easy-to-use design, and high performance.

สารบัญ

หน้า

กิตติกรรมประกาศ	ก
บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ค
สารบัญ	ง
สารบัญตาราง	ฉ
สารบัญภาพ	ฌ
บทที่ 1 บทนำ	1
ความสำคัญและที่มาของปัญหา	1
วัตถุประสงค์ของการวิจัย	3
ขอบเขตของการวิจัย	3
นิยามศัพท์	4
ประโยชน์ที่ได้รับจากการวิจัย	6
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	7
ตอนที่ 1 ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21	7
ตอนที่ 2 มโนทัศน์เกี่ยวกับการวิเคราะห์คุณภาพข้อสอบเพื่อจัดทำคลังข้อสอบ	25
ตอนที่ 3 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized adaptive testing: CAT)	36
ตอนที่ 4 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์	62
ตอนที่ 5 กรอบแนวคิดการวิจัย	70
บทที่ 3 วิธีดำเนินการวิจัย	75
ระยะที่ 1 การพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	75
ระยะที่ 2 การพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	78
ระยะที่ 3 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	91
ระยะที่ 4 การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้	92
ระยะที่ 5 การตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	93

สารบัญ

	หน้า
บทที่ 4 ผลการวิจัย	98
ตอนที่ 1 ผลการพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	99
ตอนที่ 2 ผลการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	110
ตอนที่ 3 ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	143
ตอนที่ 4 ผลการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้	163
ตอนที่ 5 ผลการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	170
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ	181
สรุปผลการวิจัย	183
อภิปรายผลการวิจัย	187
ข้อเสนอแนะ	196
รายการอ้างอิง	199
ภาคผนวก	208
ภาคผนวก 1 รายนามผู้ทรงคุณวุฒิในการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21	209
ภาคผนวก 2 รายนามผู้ทรงคุณวุฒิในการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศ	212
ภาคผนวก 3 รายนามผู้ทรงคุณวุฒิในการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21	214
ภาคผนวก 4 ตัวอย่างแบบประเมินความตรงตามเนื้อหา	216
ภาคผนวก 5 แบบประเมินระบบการทดสอบอิงมาตรฐาน	235
ภาคผนวก 6 ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ	239
ภาคผนวก 7 ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ (print out)	241
ภาคผนวก 8 ภาพการเก็บข้อมูลระยะการพัฒนาข้อสอบ	245
ภาคผนวก 9 ภาพการเก็บข้อมูลระยะการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศ	248

สารบัญ

	หน้า
บทที่ 4 ผลการวิจัย	98
ตอนที่ 1 ผลการพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	99
ตอนที่ 2 ผลการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	110
ตอนที่ 3 ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	143
ตอนที่ 4 ผลการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้	163
ตอนที่ 5 ผลการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	170
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ	181
สรุปผลการวิจัย	183
อภิปรายผลการวิจัย	187
ข้อเสนอแนะ	196
รายการอ้างอิง	199
ภาคผนวก	208
ภาคผนวก 1 รายนามผู้ทรงคุณวุฒิในการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21	209
ภาคผนวก 2 รายนามผู้ทรงคุณวุฒิในการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศ	212
ภาคผนวก 3 รายนามผู้ทรงคุณวุฒิในการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21	214
ภาคผนวก 4 ตัวอย่างแบบประเมินความตรงตามเนื้อหา	216
ภาคผนวก 5 แบบประเมินระบบการทดสอบอิงมาตรฐาน	235
ภาคผนวก 6 ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ	239
ภาคผนวก 7 ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ (print out)	241
ภาคผนวก 8 ภาพการเก็บข้อมูลระยะการพัฒนาข้อสอบ	245
ภาคผนวก 9 ภาพการเก็บข้อมูลระยะการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศ	248

สารบัญตาราง

หน้า

ตาราง 2.1	การสังเคราะห์องค์ประกอบของทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร	16
ตาราง 2.2	ตัวอย่างการทดสอบจำแนกตามองค์ประกอบที่ต้องการวัด	19
ตาราง 2.3	ตัวอย่างการทดสอบจำแนกตามมิติและลักษณะข้อคำถาม	20
ตาราง 3.1	ตัวอย่างแบบตรวจสอบรายการในการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	76
ตาราง 3.2	ตัวอย่างการวิเคราะห์ร้อยละของความถี่ของแบบตรวจสอบรายการในการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี	77
ตาราง 3.3	ประชากรในการวิจัย	78
ตาราง 3.4	ผลการสุ่มตัวอย่างโดยใช้สาขาวิชาเป็นหน่วยในการสุ่ม	79
ตาราง 3.5	ผลการกำหนดขนาดตัวอย่าง	80
ตาราง 3.6	เปรียบเทียบการกำหนดขนาดตัวอย่างตามแผนและผลการเก็บข้อมูลจริง	80
ตาราง 3.7	ตัวอย่างในการเก็บข้อมูลจำแนกตามสถาบันและสาขาวิชา	81
ตาราง 3.8	ตัวอย่างแบบประเมินความสอดคล้องระหว่างข้อคำถามกับองค์ประกอบที่ 1 การเข้าถึงสารสนเทศ(Information accessibility)	82
ตาราง 3.9	จำนวนข้อสอบและจำนวนผู้ทรงคุณวุฒิในการประเมินข้อสอบจำแนกตามชุดแบบทดสอบ	83
ตาราง 3.10	รายละเอียดองค์ประกอบ นิยามและจำนวนของข้อสอบในแบบทดสอบ	83
ตาราง 3.11	แผนการจัดข้อสอบจำแนกออกเป็น 10 ชุด	84
ตาราง 3.12	จำนวนตัวอย่างวิจัยสำหรับการทดลองใช้ระบบการทดสอบ	92
ตาราง 4.1	การสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร...	100
ตาราง 4.2	การสังเคราะห์นิยามตามองค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร	102
ตาราง 4.3	สรุปจำนวนข้อสอบจำแนกตามองค์ประกอบของการวัด	111
ตาราง 4.4	ผลการตรวจสอบความตรงเชิงเนื้อหาจำแนกตามองค์ประกอบ	116
ตาราง 4.5	ตัวอย่างข้อสอบในองค์ประกอบที่ 1 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง	117
ตาราง 4.6	ตัวอย่างข้อสอบในองค์ประกอบที่ 2 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง	118
ตาราง 4.7	ตัวอย่างข้อสอบในองค์ประกอบที่ 3 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง	119
ตาราง 4.8	ตัวอย่างข้อสอบในองค์ประกอบที่ 4 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง	120

สารบัญตาราง

		หน้า
ตาราง 4.9	ตัวอย่างข้อสอบในองค์ประกอบที่ 5 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง	121
ตาราง 4.10	ผลการวิเคราะห์คุณภาพของแบบสอบ (Test) ทั้งฉบับตามทฤษฎี CTT จำแนกตามองค์ประกอบและชุดข้อสอบ	122
ตาราง 4.11	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 1 ตามทฤษฎี CTT	123
ตาราง 4.12	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 2 ตามทฤษฎี CTT	124
ตาราง 4.13	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 3 ตามทฤษฎี CTT	125
ตาราง 4.14	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 4 ตามทฤษฎี CTT	126
ตาราง 4.15	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 5 ตามทฤษฎี CTT	127
ตาราง 4.16	โครงสร้างแบบสอบก่อนพัฒนาเป็นคลังข้อสอบแบบปรับเหมาะ	128
ตาราง 4.17	ผลการเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดล 1pl, 2pl, และ 3pl	129
ตาราง 4.18	ผลการเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดลเอกมิติและโมเดลพหุมิติ	131
ตาราง 4.19	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 1 ตามทฤษฎี IRT	133
ตาราง 4.20	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 2 ตามทฤษฎี IRT	134
ตาราง 4.21	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 3 ตามทฤษฎี IRT	135
ตาราง 4.22	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 4 ตามทฤษฎี IRT	136
ตาราง 4.23	ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 5 ตามทฤษฎี IRT	137
ตาราง 4.24	ผลการคัดข้อสอบออกจำแนกตามองค์ประกอบและชุดแบบทดสอบ	138
ตาราง 4.25	ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 1 การเข้าถึงสารสนเทศ	139
ตาราง 4.26	ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 2 การจัดการสารสนเทศ	140
ตาราง 4.27	ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 3 การบูรณาการสารสนเทศ	141

สารบัญตาราง

		หน้า
ตาราง 4.28	ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 4 การประเมินสารสนเทศ	142
ตาราง 4.29	ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 5 การสื่อสารสารสนเทศ	143
ตาราง 4.30	ค่าสถิติพื้นฐานของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาสายสังคมศาสตร์	164
ตาราง 4.31	ค่าสถิติพื้นฐานของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบของนิสิตนักศึกษาสายสังคมศาสตร์	165
ตาราง 4.32	ค่าสถิติพื้นฐานของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาสายวิทยาศาสตร์	167
ตาราง 4.33	ค่าสถิติพื้นฐานของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบของนิสิตนักศึกษาสายวิทยาศาสตร์	169
ตาราง 4.34	ผลการตรวจสอบความตรงเชิงเนื้อเรื่องของแบบประเมินระบบแบบอิงมาตรฐาน	171
ตาราง 4.35	ผลการตรวจสอบคุณภาพของระบบการทดสอบ	174
ตาราง 4.36	ผลการตรวจสอบความตรงเชิงเนื้อเรื่องของแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์	175
ตาราง 4.37	ผลการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบ	179

สารบัญภาพ

		หน้า
ภาพ 2.1	กรอบแนวคิดแห่งศตวรรษที่ 21 (Partnership for 21st skill,2007).........................	8
ภาพ 2.2	กรอบแนวคิดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารพัฒนาโดยศูนย์บริการการทดสอบทางการศึกษา (Educational Testing Service: ETS)............................	11
ภาพ 2.3	โค้งคุณลักษณะข้อสอบแบบ 1 พารามิเตอร์ (Wainer et al.,2000)............................	27
ภาพ 2.4	โค้งคุณลักษณะข้อสอบแบบ 2 พารามิเตอร์ (Wainer et al.,2000)............................	28
ภาพ 2.5	โค้งคุณลักษณะข้อสอบแบบ 3 พารามิเตอร์ (Wainer et al.,2000)............................	29
ภาพ 2.6	กระบวนการทำงานของการทดสอบแบบปรับเหมาะ..	38
ภาพ 2.7	โมเดลการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional IRT Model) (Cheng, Patton & Shao, 2014)..	42
ภาพ 2.8	ลักษณะของโมเดลการตอบสนองข้อสอบแบบพหุมิติระหว่างข้อสอบ A (Between-Items MIRT Model) และโมเดลการตอบสนองข้อสอบแบบพหุมิติภายในข้อสอบ B (Within-Items MIRT Model)...	49
ภาพ 2.9	กรอบแนวคิดการวิจัย..	74
ภาพ 3.1	ตัวอย่างกระดาษคำตอบที่ใช้ในการทดสอบ...	85
ภาพ 3.2	โมเดลการตอบสนองข้อสอบเอกมิติ..	89
ภาพ 3.3	โมเดลการตอบสนองข้อสอบพหุมิติ..	90
ภาพ 3.4	กรอบการดำเนินการพัฒนาระบบระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ...	97
ภาพ 4.1	ตัวอย่างของข้อสอบในองค์ประกอบที่ 1..	112
ภาพ 4.2	ตัวอย่างของข้อสอบในองค์ประกอบที่ 2..	112
ภาพ 4.3	ตัวอย่างของข้อสอบในองค์ประกอบที่ 3..	113
ภาพ 4.4	ตัวอย่างของข้อสอบในองค์ประกอบที่ 4..	114
ภาพ 4.5	ตัวอย่างของข้อสอบในองค์ประกอบที่ 5..	115
ภาพ 4.6	หน้าหลักระบบการทดสอบ...	144
ภาพ 4.7	การลงทะเบียน..	144
ภาพ 4.8	ฐานข้อมูลสำหรับการสร้างชุดข้อสอบ..	145
ภาพ 4.9	ชุดข้อสอบสำหรับการทดสอบขั้นแรก..	145
ภาพ 4.10	ข้อมูลของผู้เข้ารับการทดสอบเมื่อเข้าสู่ระบบ...	146
ภาพ 4.11	ข้อมูลของผู้เข้ารับการทดสอบหลังการอัปเดต...	146
ภาพ 4.12	หน้าแสดงรายละเอียดคำชี้แจง...	147
ภาพ 4.13	ข้อสอบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21.............	147
ภาพ 4.14	การประมาณค่าความสามารถ..	148
ภาพ 4.15	หน้าแสดงผลเมื่อการทดสอบเสร็จสิ้น..	150
ภาพ 4.16	ตัวอย่างรายงานผลการทดสอบ...	151
ภาพ 4.17	ตัวอย่างหน้าการประเมินคุณภาพระบบการทดสอบ...	152

สารบัญภาพ

		หน้า
ภาพ 4.18	หน้าแสดงการลงทะเบียนสำเร็จ	153
ภาพ 4.19	หน้าสำหรับเข้าใช้งานระบบ	153
ภาพ 4.20	รายละเอียดหน้าการจัดการของ Admin	157
ภาพ 4.21	กราฟค่าเฉลี่ยของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 (θ) จำแนกตามเพศของนิสิตนักศึกษาสายสังคมศาสตร์	164
ภาพ 4.22	ค่าเฉลี่ยของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบจำแนกตามเพศของนิสิตนักศึกษาสายสังคมศาสตร์	166
ภาพ 4.23	กราฟค่าเฉลี่ยของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 (θ) จำแนกตามเพศของนิสิตนักศึกษาสายวิทยาศาสตร์	168
ภาพ 4.24	ค่าเฉลี่ยของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบจำแนกตามเพศของนิสิตนักศึกษาสายวิทยาศาสตร์	167
ภาพ 4.25	แบบประเมินระบบแบบอิงมาตรฐาน	173
ภาพ 4.26	แบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์...	178

บทที่ 1
บทนำ

ความสำคัญและที่มาของปัญหา

เครือข่ายองค์กรความร่วมมือเพื่อทักษะการเรียนรู้ในศตวรรษที่ 21 (2552) ได้ศึกษาและพัฒนาระบบการศึกษาขึ้นมาภายใต้ชื่อว่า "ทักษะแห่งศตวรรษที่ 21" (21^{st} Century Skills) ซึ่งมีความแตกต่างจากระบบการศึกษาเดิมอย่างสิ้นเชิง ระบบการศึกษาแบบใหม่นี้ ไม่เพียงแต่มุ่งบรรลุเป้าหมายในการเรียนรู้ แต่ต้องมีการปลูกฝังคุณลักษณะที่ยั่งยืนในอนาคต และทัศนคติที่ดีในการเรียนรู้ด้วย พื้นฐานของหลักสูตรจะเน้นที่ความรู้ ทักษะ การเคารพตนเอง และการยอมรับคุณค่าในชีวิต ส่วนการประเมินผล เป็นการประเมินกระบวนการเรียนรู้ของผู้เรียนแต่ละคน และครูอยู่ในฐานะเป็นนักพัฒนารูปแบบการเรียน การฝึกทักษะ รู้จักใช้ ICT และอินเตอร์เน็ตเป็นตัวช่วยสำคัญในการสอน (Dede, 2009; Longworth, 2003) อาจกล่าวได้ว่าทักษะแห่งศตวรรษที่ 21 ช่วยเตรียมความพร้อมให้แก่ผู้เรียน โดยฝึกให้ผู้เรียนรู้จักคิด เรียนรู้ ทำงาน แก้ปัญหา สื่อสาร และร่วมมือทำงานกับผู้อื่นได้อย่างมีประสิทธิภาพไปตลอดชีวิต ทั้งในความแตกต่างทางวัฒนธรรมและวิถีชีวิตที่หลากหลาย และสามารถสื่อสารด้วยภาษาอื่นที่ไม่ใช่ภาษาอังกฤษได้ พร้อมทั้งมีความรู้พื้นฐานทางด้านการเงิน เศรษฐกิจ ธุรกิจ และการประกอบกิจการ (Bellanca & Brandt, 2010) โดยกรอบการเรียนรู้ของทักษะแห่งศตวรรษที่ 21 ได้แบ่งออกเป็น 2 ส่วน คือ ส่วนผลการเรียนรู้ของผู้เรียน ประกอบด้วยสาระวิชาหลัก (ภาษาอังกฤษ ภาษาสำคัญของโลก ศิลปะ คณิตศาสตร์ เศรษฐศาสตร์ วิทยาศาสตร์ ภูมิศาสตร์ ประวัติศาสตร์ และการปกครองและความเป็นพลเมืองที่ดี), คุณลักษณะ (ความรู้เรื่องโลก ความรู้ด้านการเงิน เศรษฐกิจ ธุรกิจ และการเป็นผู้ประกอบการ ความรู้ด้านการเป็นพลเมืองที่ดี ความรู้ด้านสุขภาพ และความรู้ด้านสิ่งแวดล้อม), ทักษะด้านการเรียนรู้และนวัตกรรม (ความคิดสร้างสรรค์และนวัตกรรม การคิดอย่างมีวิจารณญาณและการแก้ปัญหา และการสื่อสารและความร่วมมือ), ทักษะสารสนเทศ สื่อและเทคโนโลยี (ทักษะด้านสารสนเทศ ทักษะด้านสื่อ และทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร) และทักษะชีวิตและอาชีพ (ความยืดหยุ่นและความสามารถในการปรับตัว การริเริ่มและการกำกับดูแลตนเองได้ ทักษะด้านสังคมและทักษะข้ามวัฒนธรรม การมีผลงานและความรับผิดชอบตรวจสอบได้ และภาวะผู้นำและความรับผิดชอบ) และส่วนระบบสนับสนุนการศึกษา ประกอบด้วย มาตรฐานการเรียนรู้ในศตวรรษที่ 21 การประเมินผลทักษะการเรียนรู้ในศตวรรษที่ 21 หลักสูตรและวิธีการสอนในศตวรรษที่ 21 การพัฒนาวิชาชีพในศตวรรษที่ 21 และบรรยากาศการเรียนรู้ในศตวรรษที่ 21

ปัจจุบันโลกมีการเปลี่ยนแปลงอย่างรวดเร็ว โดยปัจจัยหนึ่งที่มีความสำคัญในการใช้ชีวิตในยุคปัจจุบันคือสารสนเทศ สื่อและเทคโนโลยี ซึ่งสอดคล้องกับทักษะการเรียนรู้ในศตวรรษที่ 21 ที่มีการแบ่งทักษะออกเป็นด้านสารสนเทศ สื่อ และเทคโนโลยี สำหรับด้านสารสนเทศมุ่งเน้นการเข้าถึงสารสนเทศอย่างมีประสิทธิภาพและประสิทธิผล มีทักษะในการประเมินสารสนเทศ สามารถใช้สารสนเทศได้อย่างถูกต้องและสร้างสรรค์ สำหรับด้านสื่อมุ่งเน้นความเข้าใจว่าข้อความจากสื่อมีจุดมุ่งหมายอะไร มีความเข้าใจในจริยธรรมและกฎหมายในการเข้าถึงและใช้สื่อ สำหรับด้านเทคโนโลยีสารสนเทศและการสื่อสารมุ่งเน้นการใช้เทคโนโลยีเป็นเครื่องมือในการวิจัย จัดการ ประเมิน และติดต่อสื่อสารสารสนเทศต่างๆ รวมถึงการใช้เทคโนโลยีดิจิทัล เช่น คอมพิวเตอร์ PDAs การเข้าถึงเครือข่ายสังคมออนไลน์อย่างเหมาะสม และการจัดการ บูรณาการ ประเมิน และสร้างสรรค์สารสนเทศเพื่อประโยชน์ทางด้านเศรษฐกิจ นอกจากนี้ยังต้องคำนึงถึงจริยธรรมและกฎหมายต่างๆ ในการเข้าถึงและใช้งานเทคโนโลยีสารสนเทศ (Partnership for 21st Century Skills, 2007)

สำหรับทักษะการเรียนรู้ในศตวรรษที่ 21 ด้านสารสนเทศ สื่อ และเทคโนโลยีสารสนเทศและการสื่อสาร มีประเด็นซึ่งเป็นที่น่าสนใจคือประเด็นด้านเทคโนโลยีสารสนเทศและการสื่อสาร (information and communication technology: ICT) ประกอบกับประเทศไทยกำลังก้าวเข้าสู่ยุคประเทศไทย 4.0 ซึ่งต้องใช้เทคโนโลยีสารสนเทศและการสื่อสารในการขับเคลื่อน และเพื่อให้แนวทางการดำเนินงานด้านเทคโนโลยีสารสนเทศและการสื่อสารบรรลุวัตถุประสงค์ จึงได้มีการจัดตั้งกระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคมขึ้น เพื่อให้สอดคล้องกับแนวทางการขับเคลื่อนประเทศไทยเข้าสู่ยุค 4.0 จากข้อมูลดังกล่าวจะเห็นได้ว่าเทคโนโลยีสารสนเทศและการสื่อสาร (ICT) เป็นสิ่งจำเป็นในการพัฒนาประเทศ ดังนั้นจึงควรมีการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร เพื่อเป็นการเตรียมความพร้อมในการพัฒนาบุคลากรในระดับอุดมศึกษาของประเทศให้มีความรู้ความเข้าในเทคโนโลยีสารสนเทศและการสื่อสาร และพร้อมนำความรู้ไปใช้ในการพัฒนา ศึกษา และวิจัย เพื่อพัฒนาประเทศให้ก้าวทันกระแสโลก ซึ่งทักษะดังกล่าวมีความสอดคล้องกับทักษะการเรียนรู้ในศตวรรษที่ 21 ด้านสารสนเทศ สื่อ และเทคโนโลยีสารสนเทศและการสื่อสาร ดังนั้นการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารจึงเป็นสิ่งจำเป็นสำหรับการศึกษาในระดับมหาวิทยาลัย (Educational Testing Service, 2002) เนื่องจากผู้เรียนต้องใช้เทคโนโลยีในการศึกษาค้นคว้าและพัฒนาองค์ความรู้ใหม่ๆ อยู่ตลอดเวลา หากผู้เรียนขาดทักษะการใช้งานเทคโนโลยีสารสนเทศและการสื่อสาร อาจทำให้ผู้เรียนไม่สามารถคิดค้นและพัฒนานวัตกรรมและความรู้ใหม่ๆ ได้ ด้วยเหตุนี้จึงควรมีการวัดทักษะการใช้งานเทคโนโลยีสารสนเทศและการสื่อสารขึ้น เพื่อประเมินผู้เรียนว่ามีความพร้อมก่อนจบการศึกษา และสามารถนำความรู้ด้านเทคโนโลยีสารสนเทศและการสื่อสารไปประยุกต์ใช้ในการประกอบอาชีพหลังจบการศึกษาต่อไป

การศึกษาการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารได้มีงานวิจัยที่พัฒนาแบบวัด/แบบประเมินทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารของผู้เรียนทั้งในระดับการศึกษาขั้นพื้นฐานและระดับอุดมศึกษา เครื่องมือที่ใช้วัดมีทั้งแบบวัดที่เป็นข้อสอบและแบบประเมินรายการที่ให้ผู้เรียนประเมินตนเองที่มีทั้งแบบทำลงในกระดาษและแบบทำในคอมพิวเตอร์ โดยยึดองค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารตามที่ได้ศึกษาและสังเคราะห์ขึ้น การใช้รูปแบบเครื่องมือที่ผ่านมามีทั้งจุดแข็งและจุดอ่อนที่แตกต่างกัน การใช้แบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารของผู้เรียนดีกว่าการใช้แบบประเมินตนเองเนื่องจากช่วยลดการตอบตามความต้องการของสังคม ทำให้ได้ข้อมูลได้ตรงกับความสามารถที่แท้จริงของผู้สอบ ส่วนการใช้รูปแบบกระดาษและคอมพิวเตอร์ก็ขึ้นอยู่กับปัจจัยความพร้อม แต่แบบวัดที่ทำในรูปคอมพิวเตอร์อาจได้เปรียบในด้านของการบูรณาการ การวิเคราะห์และการนำไปประยุกต์ใช้ได้มากกว่า โดยเฉพาะในเรื่องการใช้เพื่อการทดสอบแบบปรับเหมาะ (computerized adaptive testing: CAT)

การทดสอบแบบปรับเหมาะ คือการทดสอบที่ผู้สอบได้รับแบบทดสอบที่แตกต่างกัน โดยมีดำเนินการคัดเลือกแบบทดสอบที่มีความเหมาะสมกับระดับความสามารถของผู้เข้าสอบแต่ละคน ซึ่งใช้ระบบคอมพิวเตอร์เข้ามาช่วยในการประมาณค่าความสามารถและการคัดเลือกข้อสอบ โดยกระบวนการคำนวณทางคอมพิวเตอร์จะช่วยทำให้การประมาณค่ามีความถูกต้องของคะแนนการทดสอบในเชิงสถิติที่เพิ่มขึ้น (ณภัทร ชัยมงคล, โชติกา ภาษีผล และศิริชัย กาญจนวาสี, 2558; Meijer & Nering, 1999) หลักการในการทดสอบแบบปรับเหมาะคือ ผู้สอบจะได้รับข้อสอบในข้อถัดไปขึ้นอยู่กับผลการตอบข้อสอบในข้อก่อนหน้า โดยข้อสอบข้อแรกจะเป็นข้อสอบที่มีระดับความยากปานกลาง และเมื่อผู้สอบตอบข้อสอบเสร็จเรียบร้อยจะดำเนินการประมาณค่าความสามารถจากผลการตอบข้อสอบ เพื่อทำการเลือกข้อสอบในข้อถัดไป โดยมีหลักคือ หากผู้สอบตอบข้อสอบข้อแรกถูก ข้อสอบข้อถัดไปจะยากขึ้น แต่หากผู้สอบตอบข้อสอบข้อแรกผิด ข้อสอบข้อถัดไปจะง่ายลง โดยผู้สอบจะดำเนินการสอบไปเรื่อยๆ กระทั่งค่าความสามารถมีความคงเส้นคงวาและน่าเชื่อถือหรือ

มีค่าความคลาดเคลื่อนมาตรฐานอยู่ในเกณฑ์ที่กำหนด การทดสอบจะยุติลง (ศิริชัย กาญจนวาสี, 2555; ณภัทร ชัยมงคล, โชติกา ภาษีผล และศิริชัย กาญจนวาสี, 2558) ซึ่งจะช่วยลดความเครียดของผู้สอบไม่ต้องทำข้อสอบที่ยากมากเกินความสามารถ ในขณะเดียวกันผู้สอบที่มีความสามารถสูงก็ไม่จำเป็นที่ต้องทำข้อสอบที่ง่ายมากที่ต่ำกว่าความสามารถ

เนื่องจากการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีในประเทศไทย ยังไม่ได้มีการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ให้พร้อมใช้ การพัฒนาระบบดังกล่าวจะต้องมีการพัฒนาองค์ประกอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารและต้องสร้างข้อสอบเพื่อพัฒนาเป็นคลังข้อสอบตามองค์ประกอบที่วัด ตลอดจนสร้างระบบการทดสอบแบบปรับเหมาะ ผู้วิจัยจึงสนใจพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีในประเทศไทย นอกจากนี้จะต้องดำเนินการประเมินประสิทธิภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีเพื่อตรวจสอบคุณภาพของระบบ โดยระบบที่พัฒนาขึ้นจะเป็นระบบซึ่งอำนวยความสะดวกให้กับผู้เข้ารับการทดสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เนื่องจากเป็นระบบที่ทดสอบผ่านคอมพิวเตอร์ซึ่งมีความแม่นยำสูง ป้องกันการทุจริต เนื่องจากผู้เข้ารับการทดสอบจะได้รับข้อสอบที่แตกต่างกันขึ้นอยู่กับระดับความสามารถ และระบบสามารถรายงานผลการทดสอบให้ผู้สอบทราบผลการทดสอบได้ทันทีเพื่อนำไปใช้ในการพัฒนาและปรับปรุงทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ให้มีประสิทธิภาพยิ่งขึ้นไป

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี
2. เพื่อพัฒนาข้อสอบสำหรับจัดทำคลังข้อสอบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี
3. เพื่อพัฒนาและตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ขอบเขตของการวิจัย

การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ครั้งนี้เป็นการสร้างและตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (computerized adaptive testing) ที่ผู้เข้ารับการทดสอบแต่ละคนได้รับแบบทดสอบที่แตกต่างกัน โดยคัดเลือกข้อสอบจากระดับความสามารถของผู้เข้ารับการทดสอบแต่ละคน ซึ่งใช้ระบบคอมพิวเตอร์ช่วยในการประมาณค่าและการคัดเลือกข้อสอบประกอบด้วย 4 ส่วน องค์ประกอบ ได้แก่ 1) การลงทะเบียน 2) การสร้างชุดข้อสอบ 3) การทดสอบ 4) การรายงานผลการทดสอบ กระบวนการทางคอมพิวเตอร์ช่วยให้มีความถูกต้องของคะแนนการทดสอบในเชิงสถิติที่เพิ่มขึ้น จึงมีความจำเป็นที่จะต้องพัฒนาคลังข้อสอบที่มีข้อสอบจำนวนมากพอที่ผ่านการวิเคราะห์ว่ามีคุณภาพ คลังข้อสอบที่ใช้ในการทดสอบจึงเป็นข้อสอบแบบเลือกตอบ 5 ตัวเลือก ที่มีการให้คะแนนแบบ 0,1

ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาระดับปริญญาตรี เป็นทักษะที่มีความจำเป็นต่อการเรียนรู้ในยุคดิจิทัล (digital) จากการสังเคราะห์เอกสารและงานวิจัยที่เกี่ยวข้องประกอบด้วย 6 องค์ประกอบ ได้แก่ การเข้าถึงสารสนเทศ การจัดการสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ การสื่อสารสารสนเทศ และการสร้างสรรค์สารสนเทศ แต่การศึกษาวิจัยครั้งนี้ ใช้ข้อสอบแบบเลือกตอบจึงไม่สามารถวัดองค์ประกอบการสร้างสรรค์ได้ดี งานวิจัยนี้จึงวัด 5 องค์ประกอบ ได้แก่ 1. การเข้าถึงสารสนเทศ 2. การจัดการสารสนเทศ 3. การบูรณาการสารสนเทศ 4. การประเมินสารสนเทศ และ 5.การสื่อสารสารสนเทศ

การวิเคราะห์คุณภาพข้อสอบเพื่อใช้กับระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ที่พัฒนาขึ้นในครั้งนี้ ใช้ทฤษฎีตอบสนองข้อสอบ (Item Response Theory) โดยเลือกโมเดลที่มีความเหมาะสมที่สุดจากการตรวจสอบความสอดคล้องกับข้อมูลอย่างมีนัยสำคัญทางสถิติ ทำให้ได้โมเดลที่ทำการวิเคราะห์คุณภาพข้อสอบครั้งนี้เป็นแบบเอกมิติ 2 พารามิเตอร์ เป็นหลักในการคัดเลือกข้อสอบเข้าคลังข้อสอบ และมีการวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory) ประกอบการพิจารณาร่วมด้วย

ประชากรในการวิจัยครั้งนี้ คือ นิสิตนักศึกษาระดับปริญญาตรีจากสถาบันอุดมศึกษาของรัฐบาลทั่วประเทศ โดยใช้นิสิตนักศึกษาชั้นปีที่ 3 ทดลองใช้แบบวัดฯ เนื่องจากผ่านการเรียนรู้ในหลักสูตรและใกล้สำเร็จการศึกษา การได้มาซึ่งตัวอย่างทดลองใช้แบบวัดฯ ใช้วิธีการสุ่มสาขาวิชาที่ได้จำแนกไว้โดย สำนักงานคณะกรรมการการอุดมศึกษา (2561) จำนวน 10 สาขาวิชา ได้แก่ 1) ครุศาสตร์ ศึกษาศาสตร์ (Education) 2) ศิลปะศาสตร์ มนุษยศาสตร์ (Arts and humanities) 3) สังคมศาสตร์ วารสารศาสตร์และสารสนเทศ (Social sciences, journalism and information) 4) บริหารธุรกิจ การบริหารการจัดการและนิติศาสตร์ (Business, administration and law) 5) วิทยาศาสตร์ คณิตศาสตร์และสถิติ (Natural sciences, mathematics and statistics) 6) เทคโนโลยีสารสนเทศและการสื่อสาร (Information and Communication Technologies) 7) วิศวกรรมศาสตร์ (Engineering, manufacturing and construction 8) เกษตรศาสตร์ ป่าไม้ การประมงและสัตวแพทย์ (Agriculture, forestry, fisheries and veterinary) 9) สุขภาพและสวัสดิการ (Health and welfare) และ 10) การบริการ (Services) ส่วนการได้มาซึ่งตัวอย่างทดลองใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เป็นการเลือกตัวอย่างที่มีทั้งสายสังคมศาสตร์ และสายวิทยาสาสตร์

นิยามศัพท์

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ หมายถึง การทดสอบด้วยคอมพิวเตอร์ที่ใช้แบบสอบต่างชุดกันสำหรับผู้สอบที่มีความสามารถต่างกัน โดยมีการคัดเลือกข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบ

ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 หมายถึง ความรู้ความเข้าใจและทักษะในการใช้เทคโนโลยีดิจิตอลในศตวรรษที่ 21 ทั้งส่วนที่เป็นอุปกรณ์เทคโนโลยีดิจิตอล เครื่องมือสื่อสารต่าง ๆ รวมทั้งเครือข่ายอินเตอร์เน็ต ประกอบด้วย การเข้าถึง การจัดการ การบูรณาการ การประเมิน และการสื่อสารสารสนเทศ เพื่ออำนวยความสะดวกในการเรียนรู้ โดยอยู่บนพื้นฐานของความเข้าใจด้านกฎหมายและศีลธรรมในการเข้าถึงและใช้ข้อมูล

การเข้าถึงสารสนเทศ (information accessibility) หมายถึง ความรู้และความสามารถในการสืบค้นและรวบรวมสารสนเทศรวมทั้งอธิบายหรือแสดงวิธีการเรียกสารสนเทศที่มีอยู่กลับคืนมาเพื่อการใช้งานในระบบเทคโนโลยีสารสนเทศได้

การจัดการสารสนเทศ (information management) หมายถึง ความรู้และความสามารถในการจัดระบบหมวดหมู่ของสารสนเทศในระบบเทคโนโลยีสารสนเทศให้ง่ายต่อการการเรียกค้นและการประยุกต์ใช้ให้เหมาะสมกับงานแต่ละประเภทได้

การบูรณาการสารสนเทศ (information integration) หมายถึง ความรู้และความสามารถในการแปลความหมาย สรุป เปรียบเทียบ จำแนกความเหมือนและความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆ ในระบบเทคโนโลยีสารสนเทศได้

การประเมินสารสนเทศ (information evaluation) หมายถึง ความรู้และความสามารถในการตัดสินคุณภาพ ความสอดคล้อง ประโยชน์ หรือประสิทธิภาพของสารสนเทศที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร ในระบบเทคโนโลยีสารสนเทศได้

การสื่อสารสารสนเทศ (information communication) หมายถึง ความรู้และความสามารถในการจัดทำและเผยแพร่สารสนเทศในระบบเทคโนโลยีสารสนเทศเพื่อตอบสนองความต้องการของกลุ่มบุคคลและบริบทต่างๆที่เป็นการเคารพสิทธิ์ความเป็นส่วนตัวได้

ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 หมายถึง กลุ่มขององค์ประกอบในการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ที่คัดเลือกข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบและสามารถรายงานผลการทดสอบให้ผู้สอบทราบผลการทดสอบได้ทันที ที่ผู้วิจัยพัฒนาจากคลังข้อสอบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ที่พัฒนาขึ้น โดยใช้หลักการและแนวคิดของทฤษฎีการตอบสนองข้อสอบ ในการสร้างโปรแกรมประกอบด้วย 4 ส่วน ได้แก่ 1) การลงทะเบียน 2) การสร้างชุดข้อสอบ 3) การทดสอบ 4) การรายงานผลการทดสอบ

คลังข้อสอบ หมายถึง การรวบรวมข้อสอบเลือกตอบ 5 ตัวเลือก ที่วัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ที่มีคุณภาพ ตามกระบวนการวิเคราะห์ข้อสอบตามทฤษฎีการตอบข้อสอบ โดยมีการจัดเก็บข้อสอบไว้บนฐานข้อมูลในการทดสอบปรับเหมาะด้วยคอมพิวเตอร์

คุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 หมายถึง ประสิทธิภาพในการใช้งานระบบการทดสอบ ซึ่งมีการประเมินประสิทธิภาพ 2 วิธี ได้แก่ (1) การประเมินระบบแบบอิงมาตรฐาน โดยมีเนื้อหาในการประเมิน 4 องค์ประกอบ ได้แก่ การใช้ประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้อง แบบประเมินมีลักษณะเป็นมาตรประมาณค่า 5 ระดับ จำนวน 14 ข้อ และ (2) การประเมินความพึงพอใจระหว่างอินเตอร์เฟสกับผู้ใช้งาน โดยมีเนื้อหาในการประเมิน 5 ด้าน ได้แก่ ภาพรวมปฏิสัมพันธ์ของระบบหน้าจอ การใช้คำศัพท์และสารสนเทศระบบ การเรียนรู้ และความสามารถของระบบ แบบประเมินมีลักษณะเป็นมาตรจำแนกความหมาย 7 ระดับ จำนวน 20 ข้อ (ข้อ4)

ประโยชน์ที่ได้รับจากการวิจัย

1.ประโยชน์ในทางปฏิบัติ

ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีที่พัฒนาขึ้น สามารถนำไปใช้กับนิสิตนักศึกษาภายใต้การกำกับดูแลของหน่วยงาน/สถาบันที่สนใจ เพื่อการประเมินศักยภาพด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 โดยที่ผู้สอบแต่ละคนได้ทำข้อสอบไม่เหมือนกันขึ้นอยู่กับความสามารถและไม่จำเป็นต้องทำข้อสอบในเวลาเดียวกัน อีกทั้งมีการรายงานผลการสอบทันที ทำให้ผู้สอบสามารถนำไปพัฒนาปรับปรุงเพื่อเพิ่มขีดความสามารถในด้านนี้ต่อไป

2.ประโยชน์ในทางวิชาการ

ได้นวัตกรรมที่เป็นแนวทางในการพัฒนาข้อสอบจากองค์ประกอบที่สังเคราะห์ การเก็บข้อมูลที่ใช้ข้อสอบจำนวนมากเพื่อจัดทำคลังข้อสอบ การวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบที่มีการตรวจสอบโมเดลที่เหมาะสมและการปรับเทียบโดยใช้ข้อสอบร่วมในแบบสอบแต่ละชุด เพื่อให้การประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบมีประสิทธิภาพมากที่สุด

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้องกับการวิจัยเรื่อง ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ผู้วิจัยได้รวบรวมเอกสารและงานวิจัยที่เกี่ยวข้อง โดยแบ่งการนำเสนอเป็น 5 ตอน ดังต่อไปนี้

ตอนที่ 1 ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21
- 1.1 นิยามความหมายของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21
- 1.2 องค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21
- 1.3 งานวิจัยที่เกี่ยวข้องกับทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

ตอนที่ 2 มโนทัศน์เกี่ยวกับการวิเคราะห์คุณภาพข้อสอบเพื่อจัดทำคลังข้อสอบ
- 2.1 การวิเคราะห์ข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ (IRT)
- 2.2 การตรวจสอบความเหมาะสมของโมเดลการตอบสนองข้อสอบกับข้อมูล
- 2.3 การจัดทำคลังข้อสอบ

ตอนที่ 3 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (computerized adaptive testing: CAT)
- 3.1 ความหมายของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
- 3.2 หลักการของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
- 3.3 องค์ประกอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
- 3.4 การประยุกต์ใช้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

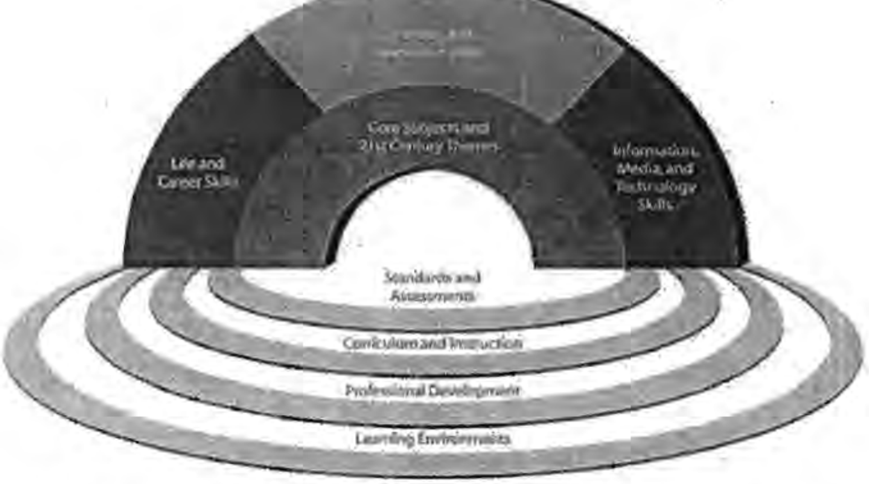
ตอนที่ 4 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
- 4.1 ขั้นตอนการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
- 4.2 การตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
- 4.3 งานวิจัยที่เกี่ยวข้องกับการพัฒนาระบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์

ตอนที่ 5 กรอบแนวคิดการวิจัย

ตอนที่ 1 ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

ในยุคศตวรรษที่ 21 เทคโนโลยีสารสนเทศและการสื่อสารมีความก้าวหน้าไปจากเดิมไปมาก ทำให้มีการเปลี่ยนแปลงวิธีการใช้ชีวิตของมนุษย์จากเดิมทั้งการทำงานและการดำเนินชีวิตประจำวัน เทคโนโลยีสารสนเทศและการสื่อสารที่ก้าวหน้ากำลังเปลี่ยนแปลงลักษณะของทักษะเดิมที่มีอยู่ก่อนและเพิ่มเติมทักษะที่จำเป็นสำหรับการทำงานและการเป็นพลเมืองในศตวรรษนี้ ประเทศต่างๆทั่วโลกได้เห็นความสำคัญของการพัฒนาทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารและเห็นว่าเป็นทักษะที่จำเป็นอย่างมาก ซึ่งเห็นได้จากในกรอบแนวคิดทักษะ

แห่งศตวรรษที่ 21 ที่พัฒนาโดยเครือข่ายองค์กรความร่วมมือเพื่อทักษะแห่งศตวรรษที่ 21 ได้บรรจุทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร (ICT literacy skill) เป็นทักษะหนึ่งที่จำเป็นในศตวรรษที่ 21 ภายใต้กรอบทักษะด้าน สารสนเทศ สื่อ และเทคโนโลยี



ภาพ 2.1 กรอบแนวคิดทักษะแห่งศตวรรษที่ 21 (Partnership for 21st skills, 2007: 1)

จากภาพกรอบแนวคิดทักษะที่จำเป็นแห่งศตวรรษที่ 21 ซึ่งพัฒนาโดยเครือข่ายองค์กรความร่วมมือเพื่อทักษะการเรียนรู้ในศตวรรษที่ 21 แบ่งออกเป็น 3 ทักษะได้แก่ ทักษะด้านการเรียนรู้และสร้างนวัตกรรม (Learning and innovation skill) ทักษะด้านสารสนเทศ สื่อ และ เทคโนโลยี (Information media and technology skills) และ ทักษะด้านชีวิตและอาชีพ (Life and career skill)

เครือข่ายองค์กรความร่วมมือเพื่อทักษะแห่งศตวรรษที่ 21 (Partnership for 21st Century Skills, 2007) ได้กล่าวว่า ผู้เรียนต้องฝึกฝนจนเกิดผลการเรียนรู้ที่เป็นสมรรถนะสำคัญ 3 ประการ เรียกว่า ความสามารถ 3Rs หรือ 3รู้ ได้แก่

รู้ที่หนึ่ง คือ รู้อ่านรู้เขียน (Literacy) ซึ่งมิใช่แค่อ่านออกเขียนได้ (Reading & Writing) แต่ต้องอ่านเข้าใจเขียนรู้เรื่อง คือ เข้าใจความหมายของเรื่องราวต่าง ๆ และสามารถสื่อสารไปยังผู้อื่นได้อย่างถูกต้องเหมาะสม รู้ศัพท์รู้ภาษา

รู้ที่สอง คือ รู้คณิต (Numeracy) ซึ่งมิใช่แค่คิดเลขเป็น (Arithmetic) แต่ต้องสามารถตีความหมายและเข้าใจความคิดต่าง ๆ ที่สื่อสารออกมาในรูปของคณิตศาสตร์ เช่น เลขคณิต พีชคณิต เรขาคณิต ตรีโกณมิติ

รู้ที่สาม คือ รู้ ICT (Information and Communication Technology Literacy) คือ เข้าใจและสามารถใช้เทคโนโลยีสารสนเทศและการสื่อสาร อันเป็นทักษะจำเป็นอย่างยิ่งในโลกปัจจุบัน

นอกจากนี้ รู้ ICT (ICT Literacy) เป็นส่วนหนึ่งของผลการเรียนรู้ทักษะด้านข้อมูลสารสนเทศ สื่อ และเทคโนโลยี (Information, Media and Technology Skills) ได้แก่ การรู้เท่าทันสารสนเทศ (Information Literacy) การรู้เท่าทันสื่อ (Media Literacy) และการรู้ทันไอซีที (ICT: Information, Communication and Technology Literacy) เป็นประสิทธิผลของการประยุกต์ใช้เทคโนโลยี (Apply Technology Efficiency) ได้แก่

(1) ใช้เทคโนโลยีเป็นเครื่องมือเพื่อการวิจัย การจัดการองค์กร การประเมินและการสื่อสารทางสารสนเทศ (2) ใช้เทคโนโลยีดิจิทัล (Computer, PDAs, Media Players etc.) ในการสื่อสารและการสร้างเครือข่าย รวมทั้งการเข้าถึง การบูรณาการ การประเมิน และการสร้างสรรค์ข้อมูล เพื่อให้สามารถทำงานได้อย่างมีประสิทธิภาพในระบบองค์ความรู้ และ (3) มีความรู้พื้นฐานในการประยุกต์ใช้ ICT ได้ตามกรอบแห่งคุณธรรมจริยธรรมที่มีข้อมูลหลากหลายรอบด้าน เพื่อให้เข้าใจในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 จึงได้รวบรวมเนื้อหาที่เกี่ยวข้อง ได้แก่ นิยามความหมาย องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 รวมทั้งงานวิจัยที่เกี่ยวข้องกับทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ดังมีรายละเอียด ดังนี้

1.1 นิยามความหมายของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

งานวิจัยเรื่องนี้ใช้คำว่า "ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร" ที่ตรงกับภาษาอังกฤษ คือ Information and Communication Technology Literacy Skills ดังนั้นการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องอาจพบว่ามีการใช้คำภาษาไทยที่ต่างกัน เช่น การรู้เท่าทันไอซีที แต่มีความหมายมาจากคำเดียวกัน คือ ICT Literacy แต่มาจากภาษาอังกฤษคำเดียวกัน

มีผู้ให้ความหมายของ ICT Literacy ไว้หลายท่านและหลายองค์กร ดังนี้

Educational Testing Service (2002) นิยาม ICT Literacy ว่าเป็นความสามารถในการใช้เทคโนโลยีดิจิตอลเป็นเครื่องมือในการสื่อสารและ/หรือ ใช้เครือข่ายอินเตอร์เน็ตในการแก้ปัญหาจัดการข้อมูลเพื่อการเรียนรู้ ซึ่งประกอบด้วยความสามารถในการใช้เทคโนโลยีเป็นเครื่องมือในการค้นหา จัดการ ประเมิน และ ติดต่อสื่อสาร โดยอยู่บนพื้นฐานของความเข้าใจด้านกฎหมายและศีลธรรมในการเข้าถึงและใช้ข้อมูล

Wijaya and Surendro (2007) ให้ความหมายของ ICT Literacy ไว้ว่า เป็นความสามารถในการใช้เทคโนโลยีดิจิทัล เครื่องมือการสื่อสาร และ/หรือเครือข่ายที่เหมาะสมเพื่อการเรียนรู้หรือการทำงานในสังคมแห่งความรู้ ซึ่งหมายรวมถึงความสามารถในการใช้เทคโนโลยีเป็นเครื่องมือในการทำวิจัย การจัดการ การประเมินผลและการสื่อสารสารสนเทศด้วยความเข้าใจและมีจริยธรรม

Markauskaite (2006) ได้กล่าวถึงนิยามของ ICT Literacy ว่าหมายถึง การมีความรู้ที่จะใช้เทคโนโลยีในการจัดการปัญหาเกี่ยวกับข้อมูล

Claro et al. (2012) ได้ให้ความหมายของ ICT Literacy ว่าหมายถึง ความสามารถในการแก้ปัญหาเกี่ยวกับข้อมูล การสื่อสาร และ ความรู้ในสังคมดิจิตอล

กิดานันท์ มลิทอง (2548) ให้นิยามของการรู้เท่าทันไอซีทีว่าคือ การที่บุคคลมีความรู้และทักษะ เข้าใจอย่างลึกซึ้งในการใช้เทคโนโลยีดิจิทัลประเภทต่างๆโดยเฉพาะอย่างยิ่งคอมพิวเตอร์ และผลิตภัณฑ์ด้านฮาร์ดแวร์และซอฟต์แวร์ เครื่องมือสื่อสารเพื่อการรับส่งสารสนเทศ และช่องทางในการส่งผ่านสารสนเทศ เพื่อสามารถจัดการสารสนเทศด้วยการจัดเก็บ ค้นคืน แปลความหมาย เชื่อมโยง สรุป เปรียบเทียบ ประเมิน และสร้างสรรค์ได้อย่างมีประสิทธิภาพ

น้ำทิพย์ วิภาวิน (2552) ได้กล่าวถึงนิยามของการรู้เท่าทันไอซีทีว่าหมายถึง ความสนใจ ทัศนคติ และความสามารถของแต่ละบุคคลในการใช้เทคโนโลยีดิจิทัลและเครื่องมือการสื่อสารในการเข้าถึงสารสนเทศ การจัดการ การเชื่อมโยง การประเมิน การลำดับเนื้อหา และการสื่อสาร เพื่อใช้ในการทำงานได้อย่างมีประสิทธิภาพ

จากการศึกษา สรุปได้ว่า ICT Literacy หรือ ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร หรือ การรู้เท่าทันไอซีที คือ การที่บุคคลมีความรู้ความเข้าใจและทักษะในการใช้เทคโนโลยีดิจิตอล ทั้งส่วนที่เป็นอุปกรณ์เทคโนโลยีดิจิตอล เครื่องมือสื่อสารต่าง ๆ รวมทั้งเครือข่ายอินเตอร์เน็ตในการเข้าถึง จัดหมวดหมู่สารสนเทศให้สะดวกต่อการใช้งาน ประเมินคุณค่าของสารสนเทศ การสร้างสรรค์ และการสื่อสารได้อย่างชำนาญ เพื่ออำนวยความสะดวกในการเรียนรู้ โดยอยู่บนพื้นฐานของความเข้าใจด้านกฎหมายและศีลธรรมในการเข้าถึงและใช้ข้อมูล

1.2 องค์ประกอบในการวัดทักษะด้านเทคโนโลยี สารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

ผู้วิจัยได้ทำการศึกษาค้นคว้าเกี่ยวกับกรอบแนวคิดและองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศ และการสื่อสาร เพื่อเป็นแนวทางในการนำไปสังเคราะห์กรอบแนวคิดในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 พบว่า มีองค์กรต่างๆ ได้มีการระบุ องค์ประกอบของ ICT literacy ไว้จำนวนหนึ่ง ซึ่งมีรายละเอียดดังต่อไปนี้

จากการศึกษาเอกสารที่เกี่ยวข้องพบว่า ในส่วนของกรอบแนวคิดตามกรอบแนวคิดโดยองค์กรเครือข่ายความร่วมมือเพื่อทักษะแห่งศตวรรษที่ 21 ได้ให้ความสำคัญกับการใช้เทคโนโลยีอย่างมีประสิทธิภาพ ใช้เทคโนโลยีเป็นเครื่องมือในการศึกษาวิจัย จัดการ ประเมินผล สื่อสาร และสร้างเครือข่ายทางสังคมอย่างเหมาะสม ภายใต้ความเข้าใจด้านจริยธรรมและกฎหมายในการเข้าถึงและใช้เทคโนโลยี โดยได้ยึดองค์ประกอบของทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสารตาม ศูนย์บริการการทดสอบทางการศึกษา

Educational Testing Service (2002) ได้แบ่งองค์ประกอบของ ICT literacy ออกเป็น 5 องค์ประกอบที่มีส่วนที่เป็นทักษะและส่วนที่เป็นความรู้ซึ่งได้แก่สมรรถนะด้านพุทธิพิสัยและสมรรถนะด้านเทคนิคอยู่ภายใต้ การรู้เท่าทันเทคโนโลยีสารสนเทศได้แก่

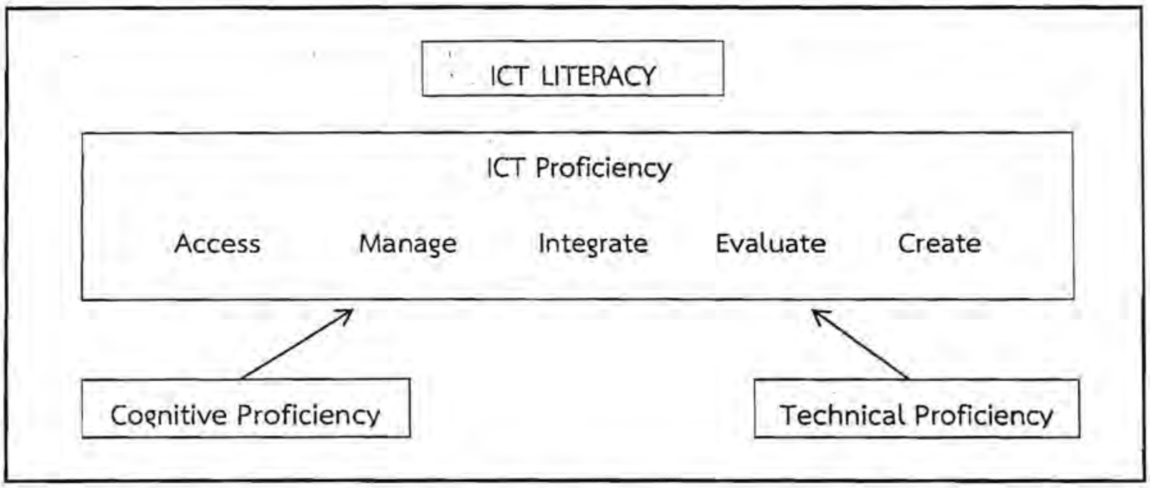
1) การเข้าถึง (Access) คือ การรู้สารสนเทศที่ต้องการและรู้ว่าจะเข้าถึงสารสนเทศดังกล่าวได้อย่างไร

2) การจัดการ (Manage) คือ ความสามารถในการจัดเก็บสารสนเทศที่มีอยู่ไว้เป็นหมวดหมู่ต่างๆง่ายต่อการใช้งาน

3) การบูรณาการ (Integrate) คือ ความสามารถในการแปลความหมาย สรุป เปรียบเทียบ แยกความแตกต่างของข้อมูลที่มีความคล้ายหรือมีรูปแบบการนำเสนอที่หลากหลาย

4) การประเมิน (Evaluate) คือ การตัดสินเกี่ยวกับคุณภาพ ความสัมพันธ์ ประโยชน์ หรือ ประสิทธิภาพของสารสนเทศ

5) การสร้างสรรค์ (Create) คือ ความสามารถในการสร้างสารสนเทศ โดยการปรับเปลี่ยน การประยุกต์ การออกแบบ การประดิษฐ์ หรือการเขียนสารสนเทศ



ภาพ 2.2 กรอบแนวคิดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารพัฒนาโดยศูนย์บริการการทดสอบทางการศึกษา (Educational Testing Service, 2002: 18)

Katz and Macklin (2007) ได้กำหนดองค์ประกอบของ ICT literacy ไว้ 7 องค์ประกอบ ได้แก่

1) การระบุ (Define) คือ การใช้เครื่องมือดิจิทัลในการระบุหรือแสดงข้อมูลที่ต้องการได้อย่างเหมาะสม

2) การเข้าถึง (Access) คือ การรวบรวม หรือ ค้นคืนสารสนเทศ

3) การจัดการ (Manage) คือ การใช้เครื่องมือดิจิทัลในการจัดระบบ หมวดหมู่ของสารสนเทศ

4) การบูรณาการ (Integrate) คือ การอธิบายหรือแปลความหมายของสารสนเทศ เช่น การใช้เครื่องมือดิจิทัลในการสังเคราะห์ การสรุป การเปรียบเทียบ และแยกความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่าง ๆ

5) การประเมิน (Evaluate) คือ การตัดสินระดับสารสนเทศดิจิทัลที่มีความต้องการจำเป็น การตัดสินคุณภาพ ความสัมพันธ์ ประโยชน์ หรือ ความน่าเชื่อถือและมีประสิทธิภาพของข้อมูล

6) การสร้างสรรค์ (Create) คือ การปรับเปลี่ยน ประยุกต์ ออกแบบ หรือ การสร้างสารสนเทศในสิ่งแวดล้อมดิจิทัล

7) การสื่อสาร (Communicate) คือ การเผยแพร่ข้อมูลสารสนเทศและความรู้ไปยังผู้อื่นหรือกลุ่มผู้คน

Markauskaite (2007) ได้กำหนดองค์ประกอบของ ICT literacy ไว้ 9 องค์ประกอบ ได้แก่

1) การวางแผน (Plan) คือ การใช้เครื่องมือในการวางแผนและสนับสนุนในการตัดสินใจ เพื่อวางแผนแก้ปัญหาด้านพุทธิพิสัย เช่น การระบุแนวคิดหลัก การพัฒนาแผนยุทธศาสตร์

2) การเข้าถึง (Access) คือ การทำงานภายในสภาพแวดล้อมที่สืบค้นจากแหล่งข้อมูลทางดิจิทัล การบำรุงรักษาเครื่องคอมพิวเตอร์ เพื่อเลือกเทคนิคและเครื่องมือที่เหมาะสม ในการค้นหาสารสนเทศจากสื่อและแหล่งข้อมูลต่าง ๆ

3) การจัดการ (Manage) คือ การดำเนินการเกี่ยวกับซอฟต์แวร์ต่าง ๆ เช่น การจัดการข้อมูล การออกแบบฐานข้อมูล เป็นต้น เพื่อการจัดทำแผนผังองค์กร การจัดหมวดหมู่และจัดเก็บข้อมูล บันทึกข้อมูลและแหล่งที่มาของข้อมูล เป็นต้น

4) การบูรณาการ (Integrate) คือ การแก้ไขปัญหาโดยการใช้เครื่องมือดิจิทัล โดยการสรุป การเปรียบเทียบ และการจำแนกความแตกต่างของแหล่งข้อมูลและแนวคิดต่าง ๆ

5) การประเมิน (Evaluate) คือ การประเมินความสัมพันธ์ระหว่างทรัพยากรดิจิทัล สารสนเทศ และเครื่องมือ เพื่อใช้ในการกำหนดเกณฑ์การประเมินตัดสินคุณภาพ ประโยชน์ และความสัมพันธ์ของสารสนเทศ และการตัดสินใจ

6) การสร้างสรรค์ (Create) คือ การสร้างกราฟฟิค เอกสาร การนำเสนอ และเว็บเพจ เพื่อนำไปปรับใช้ และสร้างสารสนเทศ นำเสนอความคิดใหม่ ๆ ออกแบบสิ่งประดิษฐ์ และสร้างผลงานอื่น ๆ

7) การสื่อสาร (Communicate) คือ การจัดทำและเผยแพร่ผลงานวิจัย โดยใช้เครื่องมือดิจิทัลและเครือข่ายอินเตอร์เน็ต เพื่อนำเสนอแนวทางแก้ไขในรูปแบบต่าง ๆ แก่หลากหลายบุคคล การเคารพในลิขสิทธิ์ความเป็นส่วนตัว

8) การทำงานร่วมกัน (Collaborate) คือ การสื่อสารผ่านทางอีเมล์และเครื่องมือเครือข่ายต่าง ๆ การทำงานร่วมกันในสภาพแวดล้อมดิจิทัล เพื่อให้การทำงานร่วมกันและการสื่อสารกับบุคคลอื่นในบริบทต่าง ๆ การทำงานเป็นทีม การปรับตัวให้เข้ากับการเรียนรู้ต่าง ๆ

9) การวิเคราะห์และการตัดสิน (Reflect and judge) คือ การจัดการส่วนบุคคล และการใช้เครื่องมือวิเคราะห์ เพื่อตัดสินใจขั้นสุดท้ายและพิจารณากระบวนการที่ใช้ในการแก้ปัญหา

California Emerging Technology Fund (2008) ได้กำหนดองค์ประกอบของ ICT literacy ไว้ 6 องค์ประกอบ ได้แก่

1) การเข้าถึง (Access) คือ การมีความรู้และสามารถรวบรวมสารสนเทศ โดยใช้เครื่องมือดิจิทัลในการสืบค้น ค้นหา และดึงข้อมูลในระบบดิจิทัลได้

2) การจัดการ (Manage) คือ การใช้เครื่องมือดิจิทัลในการจัดระบบ หมวดหมู่ของสารสนเทศ โดยใช้เครื่องมือดิจิทัลดำเนินการจัดเตรียมข้อมูลพื้นฐานสำหรับการเรียกค้นและการประยุกต์ใช้ในอนาคต

3) การบูรณาการ (Integrate) คือ การตีความและแสดงสารสนเทศ ทั้งการสรุปผล การเปรียบเทียบและความแตกต่างระหว่างกัน โดยใช้เครื่องมือดิจิทัลในการสังเคราะห์ สรุป เปรียบเทียบ และ แยกความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆ

4) การประเมิน (Evaluate) คือ การตัดสินเกี่ยวกับคุณภาพ ความสัมพันธ์ หรือประโยชน์ของสารสนเทศดิจิทัล โดยการตัดสินตามความทันสมัย ความเหมาะสมและความเพียงพอของสารสนเทศและแหล่งข้อมูล

5) การสร้างสรรค์ (Create) คือ การสร้างสารสนเทศ โดยการปรับเปลี่ยน ประยุกต์ ออกแบบ หรือการสร้างสารสนเทศในสิ่งแวดล้อมดิจิทัล เพื่ออธิบายเหตุการณ์ แสดงความคิดเห็น สนับสนุนข้อโต้แย้งหรือมุมมอง

6) การสื่อสาร (Communicate) คือ การเผยแพร่ข้อมูลสารสนเทศ เพื่อตอบสนองความต้องการของกลุ่มผู้คน โดยใช้เครื่องมือดิจิทัลในการสื่อสาร ปรับตัว และนำเสนอสารสนเทศอย่างถูกต้องและเหมาะสมในบริบทต่าง ๆ ในสภาพแวดล้อมที่เป็นดิจิทัล

Somerville et al. (2008) ได้กำหนดองค์ประกอบของ ICT literacy 7 องค์ประกอบ ได้แก่

1) การระบุ (Define) คือ การทำความเข้าใจขอบเขตของปัญหา เพื่อที่จะอำนวยความสะดวกในการสืบค้นสารสนเทศ เช่น การแยกแยะคำถามวิจัยให้มีความชัดเจน ครอบคลุม และเฉพาะเจาะจงจากคำถามที่มีปัญหาในวงที่กว้างเกินไปหรือไม่เป็นไปตามความต้องการของสารสนเทศ การสอบถามผู้เชี่ยวชาญเพื่อให้ได้สารสนเทศที่ชัดเจนขึ้น เป็นต้น

2) การเข้าถึง (Access) คือ การรวบรวม และ/หรือดึงสารสนเทศในระบบดิจิทัล โดยที่แหล่งข้อมูลอาจจะเป็นเว็บเพจ ฐานข้อมูล กลุ่มสนทนา อีเมล หรือสื่อสิ่งพิมพ์ออนไลน์ รวมถึงการสร้างและรวมคำที่ใช้ในการค้นหา เพื่อตอบสนองความต้องการของงานวิจัยที่เฉพาะ เจาะจง การเลือกใช้แหล่งข้อมูลอย่างน้อยหนึ่งรายการเพื่อค้นหาข้อมูลที่ต้องการ และการตัดสินใจเลือกแหล่งข้อมูลที่ให้ข้อมูลที่เป็นประโยชน์มากที่สุดสำหรับความต้องการ

3) การประเมิน (Evaluate) คือ การตัดสินว่าสารสนเทศสามารถตอบสนองความต้องการของปัญหาภายใต้อคติ การกำหนดของระยะเวลา รวมถึงความเกี่ยวข้องอื่น ๆ เช่น การพิจารณาว่าเว็บเพจและบทความวารสารออนไลน์นี้มีประโยชน์ การประเมินว่าฐานข้อมูลมีข้อมูลที่เป็นปัจจุบันและตรงประเด็นเหมาะสม และการตัดสินใจว่าสามารถรวบรวมข้อมูลได้ครอบคลุมเรื่องที่ต้องการวิจัย

4) การจัดการ (Manage) คือ การจัดระเบียบสารสนเทศ เพื่อช่วยให้บุคคลอื่นรวมถึงตัวเองสามารถค้นพบได้ในภายหลัง เช่น การจัดหมวดหมู่อีเมลลงในแฟ้มที่มีเนื้อหาอีเมลที่คล้ายกัน การจัดข้อมูลบุคลากรลงในแผนผังองค์กร การจัดเรียงไฟล์อีเมลหรือการส่งข้อมูลคืนฐานข้อมูล เพื่อชี้แจงข้อมูลที่เกี่ยวข้อง เป็นต้น

5) การบูรณาการ (Integrate) คือ การตีความและแสดงข้อมูล เช่น การใช้เครื่องมือดิจิทัลเพื่อสังเคราะห์ สรุป เปรียบเทียบข้อมูลจากแหล่งข้อมูลที่หลากหลาย ทั้งการเปรียบเทียบโฆษณา อีเมล หรือเว็บไซต์ของคู่แข่งด้วยการสรุปข้อมูลในลงตาราง การกำหนดเกณฑ์เฉพาะเพื่อเปรียบเทียบข้อมูลและตัดสินใจเป็นต้น

6) การสร้างสรรค์ (Create) คือ การปรับใช้ ประยุกต์ หรือ สร้างสารสนเทศในระบบดิจิทัล เช่น การแก้ไขและจัดรูปแบบของเอกสารตามข้อกำหนดเฉพาะ การสร้างสไลด์นำเสนอเพื่อสนับสนุนประเด็นในการโต้แย้ง การสร้างและแสดงผลข้อมูลเพื่อชี้แจงความสัมพันธ์ระหว่างตัวแปรทางวิชาการกับตัวแปรทางเศรษฐกิจ เป็นต้น

7) การสื่อสาร (Communicate) คือ การเผยแพร่สารสนเทศให้เหมาะสมกับกลุ่มผู้รับในรูปแบบของดิจิทัลที่มีประสิทธิภาพ เช่น การจัดรูปแบบเอกสารเพื่อให้เป็นประโยชน์กับกลุ่มใดกลุ่มหนึ่ง การเลือกและจัดระเบียบภาพนิ่งสำหรับการนำเสนอที่แตกต่างกันเพื่อผู้ชมที่แตกต่างกัน การออกแบบใบปลิวเพื่อโฆษณากับกลุ่มผู้ใช้ที่แตกต่างกัน เป็นต้น

UNESCO (2008) ได้กำหนดองค์ประกอบของ ICT literacy ไว้ 7 องค์ประกอบ ได้แก่

1) การระบุ (Define) คือ การใช้เครื่องมือดิจิทัลในการระบุหรือแสดงข้อมูลที่ต้องการได้อย่างเหมาะสม

2) การเข้าถึง (Access) คือ การรวบรวม หรือ สืบค้นสารสนเทศ ในระบบดิจิทัล นอกจากนี้ ยังมีความสามารถในการพัฒนาระบบในการค้นหาข้อมูลในฐานข้อมูล

3) การจัดการ (Manage) คือ การจัดระเบียบสารสนเทศตามแบบแผนที่มีอยู่

4) การประเมิน (Evaluate) คือ การตัดสินเกี่ยวกับคุณภาพ ความสัมพันธ์ ประโยชน์ ประสิทธิภาพ อคติ และข้อจำกัดทางด้านระยะเวลาของสารสนเทศ

5) การบูรณาการ (Integrate) คือ การตีความ สรุป เปรียบเทียบ และความแตกต่างกันระหว่างข้อมูลจากแหล่งข้อมูลที่หลากหลาย

6) การสร้างสรรค์ (Create) คือ การสร้างสารสนเทศและความรู้ใหม่ ๆ โดยการปรับใช้ ประยุกต์ ออกแบบ ประดิษฐ์ หรือนำเสนอสารสนเทศในระบบดิจิทัล

7) การสื่อสาร (Communicate) คือ การถ่ายทอดสารสนเทศและความรู้ไปสู่บุคคลหรือกลุ่มต่าง ๆ

Claro et al. (2012) ได้กำหนดมิติของทักษะ ICT literacy ไว้ 3 มิติ และในแต่ละมิติ แบ่งออกเป็น 2 มิติย่อย ได้แก่

1) มิติการจัดการสารสนเทศ (Information fluency)

1.1) ทักษะ ICT ในการจัดหาสารสนเทศ รวมถึงความสามารถในการสืบค้น คัดเลือก ประเมิน จัดระเบียบและจัดการสารสนเทศดิจิทัล

1.2) ทักษะ ICT ในการประมวลสารสนเทศ ซึ่งพิจารณาจากความสามารถในการวิเคราะห์ ปรับเปลี่ยน นำเสนอสารสนเทศ สร้างสรรค์สารสนเทศใหม่ และพัฒนาความคิดของตนเอง

2) มิติการสื่อสารอย่างมีประสิทธิภาพ (Effective communication)

2.1) ทักษะ ICT ในการสื่อสารอย่างมีประสิทธิภาพ ซึ่งพิจารณาจากความสามารถในการส่งสารสนเทศให้แก่ผู้อื่นได้อย่างถูกต้อง ครบถ้วน และใช้เครื่องมือดิจิทัลที่เหมาะสม

2.2) ทักษะ ICT ในการทำงานร่วมกันและปฏิสัมพันธ์ รวมถึงความสามารถในการโต้ตอบผ่านเครือข่ายอินเตอร์เน็ต และการใช้เทคโนโลยีสารสนเทศ เพื่อแลกเปลี่ยนข้อมูลสารสนเทศ เจรจาข้อตกลง และการตัดสินใจร่วมกับเพื่อนร่วมงานด้วยความเคารพซึ่งกันและกัน

3) มิติจริยธรรมและผลกระทบทางสังคม (Ethics and social impact dimension)

3.1) ความสามารถในการประเมินความรับผิดชอบทางด้าน ICT ซึ่งพิจารณาจากความสามารถในการตัดสินใจเกี่ยวกับข้อจำกัดของกฎหมาย จริยธรรม และวัฒนธรรมในการใช้ ICT โดยเฉพาะ และความรับผิดชอบในการใช้ ICT ต่อสังคม โดยทำความเข้าใจความเสี่ยงที่อาจจะเกิดขึ้นได้ในเครือข่ายอินเตอร์เน็ต

3.2) ความสามารถในการประเมินผลกระทบของ ICT ต่อสังคม รวมถึงความสามารถในการเข้าใจ การวิเคราะห์ และการประเมินผลกระทบของ ICT ต่อสังคม เศรษฐกิจ และวัฒนธรรม

Ahmad, Karim, Din, and Albakri (2013) ได้กำหนดองค์ประกอบของ ICT literacy ไว้ 7 องค์ประกอบตามโมเดลสมรรถนะ ICT แห่งศตวรรษที่ 21 ได้แก่ การระบุ (Define) การเข้าถึง (Access) การประเมิน (Evaluate) การจัดการ (Manage) การบูรณาการ (Integrate) การสร้าง (Construct) และ การสื่อสาร (Communicate)

น้ำทิพย์ วิภาวิน (2552) ได้กำหนดองค์ประกอบของทักษะการรู้เท่าทันเทคโนโลยี สารสนเทศและการสื่อสารไว้ 7 องค์ประกอบ ได้แก่

1) การระบุ (Define) คือ การใช้เครื่องมือ ICT เพื่อระบุความต้องการสารสนเทศ

2) การเข้าถึง (Access) คือ การรวบรวม และ/หรือ การค้นหาสารสนเทศในสังคมดิจิทัล

3) การจัดการ (Manage) คือ การใช้เครื่องมือ ICT เพื่อประยุกต์ใช้สารสนเทศในองค์กร

4) การบูรณาการ (Integrate) คือ การสังเคราะห์ สรุปความ เปรียบเทียบ และนำสารสนเทศมาใช้จากหลายแหล่ง

5) การประเมิน (Evaluate) คือ การประเมินระดับของสารสนเทศที่สอดคล้องกับความต้องการ

6) สร้างสรรค์ (Create) คือ การปรับใช้ ประยุกต์ ออกแบบ หรือสร้างสรรค์สารสนเทศ

7) การสื่อสาร (Communicate) คือ การสื่อความสารสนเทศในบริบทสภาพแวดล้อมทางเทคโนโลยีสารสนเทศ

จากองค์ประกอบการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารดังกล่าวข้างต้น พบว่ามีความคล้ายคลึงกัน โดยให้ความสำคัญกับการใช้เทคโนโลยีสารสนเทศและการสื่อสารในการระบุและนำไปสู่การเข้าถึงข้อมูลที่ต้องการ การจัดระเบียบสารสนเทศให้เป็นหมวดหมู่ สามารถค้นคืนข้อมูลที่ต้องการได้ การประเมินสารสนเทศที่มีความจำเป็น ให้ระดับสารสนเทศที่มีคุณภาพ รวมไปถึงการสร้างสรรค์สารสนเทศ การส่งต่อสารสนเทศ การติดต่อสื่อสารในระดับบุคคลและในกลุ่มบุคคล ซึ่งผู้วิจัยได้ทำการสังเคราะห์องค์ประกอบทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสารดังแสดงในตาราง 2.1

ตาราง 2.1 การสังเคราะห์องค์ประกอบของทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร

องค์ประกอบของทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร	Educational Testing Service: ETS (2002)	Katz and Macklin (2007)	Markauskaite (2007)	California Emerging Technology Fund (2008)	Somerville et al. (2008)	UNESSCO (2008)	Claro et al. (2012)	Ahmad et al. (2013)	น้ำทิพย์ วิภาวิน (2552)	รวม (ร้อยละ)
การเข้าถึง (Access)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การจัดการ (Manage)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การบูรณาการ (Integrate)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การประเมิน (Evaluate)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การสร้างสรรค์ (Create)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การสื่อสาร (Communicate)		✓	✓	✓	✓	✓	✓	✓	✓	8 (88.89)
การระบุ (Define)		✓			✓	✓	✓	✓	✓	7 (77.78)
การวางแผน (Plan)			✓							1 (11.11)
การทำงานร่วมกัน (Collaborate)			✓							1 (11.11)
การวิเคราะห์และการตัดสิน (Reflect and judge)			✓							1 (11.11)

จากตารางสังเคราะห์องค์ประกอบทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสารผู้วิจัยสรุปองค์ประกอบได้เป็น 6 องค์ประกอบ โดยใช้เกณฑ์พิจารณาองค์ประกอบที่มีค่าน้ำหนักมากกว่าร้อยละ 80 เท่านั้น เพื่อใช้ในการวัดทักษะการรู้เท่าทันเทคโนโลยี สารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 โดยสรุปคำนิยามที่ได้จากการสังเคราะห์ในแต่ละองค์ประกอบที่ทำการวัดได้แก่

1) การเข้าถึง (Access) คือ การรวบรวม หรือ ค้นคืนสารสนเทศ

2) การจัดการ (Manage) คือ การใช้เครื่องมือดิจิทัลในการจัดระบบ หมวดหมู่ของสารสนเทศ

3) การบูรณาการ (Integrate) คือ การอธิบายหรือแปลความหมายของสารสนเทศ การใช้เครื่องมือดิจิทัลในการสังเคราะห์ สรุป เปรียบเทียบ และ แยกความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆ

4) การประเมิน (Evaluate) คือ การตัดสินระดับสารสนเทศดิจิทัลที่มีความต้องการจำเป็น การตัดสินคุณภาพ ความสัมพันธ์ ประโยชน์ หรือ ความน่าเชื่อถือและมีประสิทธิภาพของข้อมูล

5) การสร้างสรรค์ (Create) คือ การปรับเปลี่ยน ประยุกต์ ออกแบบ หรือ การสร้างสารสนเทศในสิ่งแวดล้อมดิจิทัล

6) การสื่อสาร (Communicate) คือ การเผยแพร่ข้อมูลสารสนเทศและความรู้ไปยังผู้อื่นหรือกลุ่มผู้คน

ผลการสังเคราะห์องค์ประกอบทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร เพื่อให้ได้องค์ประกอบที่เป็นตัวแทนที่ดี ลดความซ้ำซ้อนในการวัด รวมทั้งสามารถสร้างเครื่องมือที่มีการให้คะแนนแบบเลือกตอบ (multiple choices) ที่นำไปใช้ในการทดสอบแบบปรับเหมาะ (computerized adaptive testing: CAT) ที่ใช้ระบบคอมพิวเตอร์เข้ามาช่วยในการประมาณค่าความสามารถและการคัดเลือกข้อสอบ โดยกระบวนการคำนวณทางคอมพิวเตอร์จะช่วยทำให้การประมาณค่ามีความถูกต้องของคะแนนการทดสอบในเชิงสถิติที่เพิ่มขึ้น ทำให้ได้องค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี 5 องค์ประกอบ ไม่มีองค์ประกอบด้านการสร้างสรรค์ เนื่องจากไม่สามารถสร้างองค์ประกอบนี้ได้ดีโดยใช้ข้อสอบแบบเลือกตอบได้ โดยสรุปคำนิยามในแต่ละองค์ประกอบ ดังนี้

1. **การเข้าถึงสารสนเทศ** (information accessibility) หมายถึง การระบุสารสนเทศที่จำเป็นในการสืบค้นและการเรียกสารสนเทศกลับคืนเพื่อการใช้งาน (retrieve information)

2. **การจัดการสารสนเทศ** (information management) หมายถึง การจัดระบบและเก็บรวบรวมสารสนเทศ เพื่อเรียกกลับคืนมาใช้งาน (retrieve) และใช้ซ้ำ (reuse)

3. **การบูรณาการสารสนเทศ** (information integration) หมายถึง การแปลผลและนำเสนอสารสนเทศซึ่งมีความเกี่ยวข้องกับการสรุป เปรียบเทียบ และจำแนกความแตกต่าง

4. **การประเมินสารสนเทศ** (information evaluation) หมายถึง การสะท้อนถึงกระบวนการที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร เพื่อนำมาใช้ในการตัดสินใจซึ่งเกี่ยวข้องกับคุณภาพ ความสอดคล้อง ประโยชน์ หรือประสิทธิภาพของสารสนเทศ

5.การสื่อสารสารสนเทศ (information communication) หมายถึง การแลกเปลี่ยนสารสนเทศโดยการสร้างและเผยแพร่ข้อมูลข่าวสารที่มีความเหมาะสมกับบริบทของผู้รับสารสนเทศ

ผู้วิจัยได้ใช้ทั้ง 5 องค์ประกอบดังกล่าว วัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 โดยพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ต่อไป

1.3 งานวิจัยที่เกี่ยวข้องกับทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

จากการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารพบว่ามีเอกสารงานวิจัยที่เกี่ยวข้องกับการพัฒนาเครื่องมือเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารและเอกสารงานวิจัยที่เกี่ยวข้องกับประสิทธิผลของการใช้ไอซีทีในการจัดการเรียนรู้ เครื่องมือที่ใช้มีหลายรูปแบบ แต่หากเป็นงานวิจัยในประเทศไทยยังไม่พบเอกสารงานวิจัยที่เกี่ยวข้องโดยตรงในด้านของการพัฒนาเครื่องมือสำหรับใช้วัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร ที่พบในประเทศไทยส่วนใหญ่เป็นงานวิจัยการวัดสมรรถนะไอซีที เครื่องมือที่ใช้เป็นแบบสอบถาม และแบบประเมิน ซึ่งต่างจากการวิจัยในครั้งนี้

การวิเคราะห์เอกสารงานวิจัยต่างๆที่เกี่ยวข้องในตอนนี้ จึงแบ่งงานวิจัยเป็น 2 กลุ่ม ได้แก่ กลุ่มที่ 1 เป็นงานวิจัยที่เกี่ยวข้องกับการพัฒนาองค์ประกอบ เครื่องมือสำหรับวัดทักษะด้านเทคโนโลยี สารสนเทศและการสื่อสาร และการตรวจสอบคุณภาพเครื่องมือ กลุ่มที่ 2 เป็นงานวิจัยที่เกี่ยวข้องกับการใช้ไอซีทีในการจัดการเรียนการสอนรวมทั้งการวัดระดับการรู้เท่าทันไอซีที

1) งานวิจัยที่เกี่ยวข้องกับการพัฒนาองค์ประกอบ เครื่องมือสำหรับวัดทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร และการตรวจสอบคุณภาพเครื่องมือ ซึ่งมีรายละเอียดดังต่อไปนี้

Wijaya and Surendro (2007) ทำการศึกษาเกี่ยวกับการพัฒนาเครื่องมือสำหรับวัดการรู้เท่าทันไอซีทีสำหรับวัดระดับการรู้เท่าทันไอซีที ของบุคลากรในสำนักงานของหน่วยงานต่างๆของรัฐบาล ทั้งนี้ในงานวิจัยไม่ได้มีการสุ่มตัวอย่าง เป็นเพียงการศึกษาองค์ประกอบและนิยามตัวบ่งชี้ที่ควรจะต้องวัด และพัฒนาแบบสอบในลักษณะโปรแกรมเสมือน ข้อคำถามเป็นรูปแบบของการให้สถานการณ์และตอบคำถามหรือให้ทำการปฏิบัติงานบางอย่าง แต่ยังไม่ได้มีการพัฒนาตัวโปรแกรมดังกล่าวขึ้นมา ผลการศึกษาพบว่า ในการทำงานของหน่วยงานราชการของรัฐบาล ต้องการบุคลากรที่มีการรู้เท่าทันไอซีที มีความสามารถที่จะคิดวิเคราะห์เกี่ยวกับข้อมูลสารสนเทศและการสื่อสาร เข้าใจและสามารถสร้างแบบฟอร์มข้อมูลแบบต่างๆได้ แล้วก็พบว่ามีความสัมพันธ์สูงในเชิงบวก ระหว่างความสามารถในการใช้ประโยชน์ของไอซีที เพื่อการดำเนินงาน กับ การประสบความสำเร็จในการปฏิบัติงานในสำนักงานในหน่วยงานของรัฐบาล และในอนาคตจะมีการพัฒนาโปรแกรมสำหรับประเมินการรู้เท่าทันไอซีทีสำหรับ สำนักงานในหน่วยงานต่างๆของรัฐบาล ทั้งนี้ในงานวิจัยดังกล่าวได้มีการให้ขอบเขตตัวอย่างพฤติกรรมที่ทำการทดสอบทั้ง 7 องค์ประกอบดังตารางต่อไปนี้

ตาราง 2.2 ตัวอย่างการทดสอบจำแนกตามองค์ประกอบที่ต้องการวัด

ลำดับที่	องค์ประกอบที่ต้องการวัด	ตัวอย่างการทดสอบ
1	Define: การระบุพื้นฐานทักษะด้านคอมพิวเตอร์	การใช้เมาส์ลาก การคลุมดำ การบันทึกข้อมูล การคัดลอกและวางข้อมูลในหน้าเอกสารที่ต้องการ การใช้แท็บเลื่อนลงด้านล่าง การสร้างกล่องรายการ (Combo box)
2	Access: การเข้าถึง	เลือกและ เปิดจดหมายอิเล็กทรอนิกส์จากกล่องจดหมายเข้า อ่านข้อความเชิญประชุมทาง E-mail ส่ง E-mail จดหมายเชิญประชุม เก็บรวบรวมข้อมูลเกี่ยวกับกฎข้อบังคับของรัฐบาล ในการใช้เครื่องมือสืบค้นข้อมูล
3	Manage: การจัดการ	จัดกลุ่ม E-mail ที่มีเนื้อเรื่องเดียวกันให้อยู่ในกลุ่มเดียวกัน แต่ละกลุ่มแยกเป็นกลุ่มละ Folder
4	Integrate: การบูรณาการ	แสดงข้อมูลที่เป็นตัวแทนจากแหล่งสืบค้นข้อมูลต่างๆโดยใช้ Application อื่นที่คล้ายกับ Application ของสำนักงาน แสดงข้อมูลเกี่ยวกับพลเมืองจากแหล่งต่าง ๆ โดยใช้ Application ของสำนักงาน
5	Evaluate: การประเมิน	เปรียบเทียบและตัดสินเลือกข้อมูลจาก แหล่งสืบค้นข้อมูลต่าง ๆ ซึ่งสอดคล้องสัมพันธ์กับหัวข้อที่เกี่ยวข้องกับการวางแผนเชิงกลยุทธ์ระดับภูมิภาค
6	Create: การสร้างสรรค์	สร้างสารสนเทศใหม่ หรือ ความรู้ใหม่เกี่ยวกับ พลเมืองจากหลายแหล่งข้อมูล โดยใช้ Application ของสำนักงาน
7	Communication: การสื่อสาร	ตอบ E-mail เกี่ยวกับนโยบายของรัฐบาลให้ประชาชนได้รับทราบ เลือกข้อมูลที่ถูกต้องเผยแพร่ต่อที่สาธารณะ

Katz and Macklin (2007) ได้ทำการศึกษาการกระตุ้นการเรียนรู้โดยใช้การประเมินเป็นฐานเพื่อพัฒนาทักษะการรู้เท่าทันไอซีทีในระดับอุดมศึกษา มีวัตถุประสงค์ในการวิจัยเพื่อตรวจสอบความตรงของเครื่องมือที่ใช้ประเมินทักษะการรู้เท่าทันไอซีที ตัวอย่างคือนักศึกษาระดับปริญญาบัณฑิตชั้นปีที่1 เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลเป็นแบบวัดที่ใช้อินเทอร์เน็ตเป็นฐาน (Internet-base) เป็นชุดคำสั่งที่มีลำดับขั้นที่แน่นอน ลักษณะมีการให้สถานการณ์และให้ ภาระงาน โดยให้เวลาตอบแต่ละข้อประมาณ 5-15 นาที (Typical applications) ซึ่งพัฒนาโดย Educational Testing Service (ETS) มีเป้าหมายเพื่อพิจารณาทักษะด้านการแก้ปัญหา (Cognitive) วิเคราะห์คุณภาพเครื่องมือโดยการตรวจสอบความตรงเชิงโครงสร้าง (Convergent and discriminant validity) มีการพัฒนาแบบสอบถามและเปรียบเทียบผลการวัดจากแบบสอบถาม ผลวิจัยสรุปได้ว่า การศึกษาครั้งนี้พบว่าแบบสอบที่พัฒนาโดย ETS ได้ให้หลักฐานแสดงความตรงทั้งส่วนที่เป็น Convergent and discriminant validity ซึ่งหากมีการนำไปใช้ทดสอบจะ ทำให้ทราบว่านักศึกษาชั้นปีที่1 มีทักษะในกระบวนการได้รับสารสนเทศอย่างไร มีประโยชน์ต่อการเลือกวิธีการที่เหมาะสมในการฝึกทักษะดังกล่าวลงในหลักสูตรการเรียนการสอนต่อไป

Claro et al. (2012) ทำงานวิจัยเกี่ยวกับศึกษาการประเมินทักษะการรู้เท่าทัน ICT ของนักเรียนวัย 15 ปี โดยในงานวิจัยนี้จะทำการศึกษาและระบุนิยามเชิงปฏิบัติการของทักษะด้านไอซีทีและทำการวัดระดับทักษะการรู้เท่าทันไอซีทีของนักเรียน โดยออกแบบการสร้างเครื่องมือในลักษณะของแบบวัดการปฏิบัติประกอบด้วยข้อคำถามทั้งหมด 57 ข้อคำถาม แบ่งออกเป็น 3 มิติ ได้แก่

ตาราง 2.3 ตัวอย่างการทดสอบจำแนกตามมิติและลักษณะข้อคำถาม

มิติ	มิติย่อย	ลักษณะข้อคำถาม	ตัวอย่างการทดสอบ
มิติการจัดการสารสนเทศ	ทักษะ ICT ในการจัดหาสารสนเทศ	ปรนัย 14 ข้อ	ให้นักเรียนเลือกคำที่ใช้ค้นหาข้อมูลบนอินเตอร์เน็ตได้ดีที่สุด จากตัวเลือก 5 คำ
	ทักษะ ICT ในการประมวลสารสนเทศ	ปรนัยและเขียนตอบแบบเปิด 14 ข้อ	- ให้นักเรียนอ่านและวิเคราะห์ข้อมูลประเภทต่าง ๆ (กราฟ ตาราง ข้อความ) และเลือกการตีความข้อมูลที่ดีที่สุดจาก 5 ตัวเลือก - ให้นักเรียนเลือกภาพที่ดีที่สุดในการแสดงข้อมูล
มิติการสื่อสารอย่างมีประสิทธิภาพ	ทักษะ ICT ในการสื่อสารอย่างมีประสิทธิภาพ	ปรนัยและเขียนตอบแบบเปิด 10 ข้อ	- ให้นักเรียนเลือกวิธีการสื่อสารที่ดีที่สุด โดยคำนึงถึงวัตถุประสงค์และผู้รับ - ให้นักเรียนเขียนและส่งอีเมลในการเชิญเข้าร่วมการรณรงค์ด้านสิ่งแวดล้อม
	ทักษะ ICT ในการทำงานร่วมกันและปฏิสัมพันธ์	ปรนัยและเขียนตอบแบบเปิด 10 ข้อ	- ให้นักเรียนเลือกกระบวนการทำงานที่ดีที่สุดสำหรับผู้ดำเนินการในที่ประชุม เมื่อมีผู้เข้าร่วมรายหนึ่งไม่ปฏิบัติตามกฎ - ให้นักเรียนเขียนคำถามในที่ประชุม เพื่อถามข้อมูลที่เฉพาะ
มิติจริยธรรมและผลกระทบทางสังคม	ความสามารถในการประเมินความรับผิดชอบทางด้าน ICT	ปรนัย 8 ข้อ	- ให้นักเรียนเข้าเว็บเพจและให้สมัครว่าเป็นสมาชิกขององค์กรทางด้านนิเวศวิทยา และให้เลือกเหตุผลว่าทำไมเว็บไซต์นี้ถึงไม่มีความน่าเชื่อถือในการเผยแพร่ข้อมูล - ให้นักเรียนเลือกคำแนะนำที่ดีที่สุดสำหรับเพื่อนที่ได้รับอีเมลจากผู้ที่ไม่รู้จักว่าได้รับรางวัล
	ความสามารถในการประเมินผลกระทบของ ICT ต่อสังคม	เขียนตอบแบบเปิด 1 ข้อ	ให้นักเรียนอธิบายเกี่ยวกับบทความที่ทำให้นึกถึงผลกระทบของอินเตอร์เน็ตในสังคม

ตัวอย่างคือนักเรียนวัย 15 ปี จำนวน 1,185 คน ที่ได้มาจากการสุ่มแบบหลายขั้นตอน วิเคราะห์ข้อมูลตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) แบบวัดมีความเที่ยงอยู่ที่ 0.825 ผลการศึกษาพบว่า ผู้เรียนที่ได้รับการทดสอบส่วนใหญ่มีทักษะการแก้ปัญหาจากภาระงานที่กำหนดให้อยู่ในเกณฑ์ที่ต่ำ ควรจะได้รับการพัฒนาทักษะการรู้เท่าทัน ICT และควรจะมีการบรรจุการพัฒนาทักษะการรู้เท่าทันไอซีทีไว้ในหลักสูตร

Ahmad et al. (2013) ศึกษาเกี่ยวกับการประเมินสมรรถนะด้านไอซีทีของผู้ที่จบการศึกษาระดับปริญญาบัณฑิตโดยใช้โมเดลสมรรถนะไอซีทีแห่งศตวรรษที่ 21 เป็นฐานซึ่งพูดถึงรายละเอียดและวิธีการพัฒนาแบบวัดสมรรถนะไอซีทีที่เรียกว่า "ICTC-Test" (Information and Communication Technology Competency Test) ประกอบด้วย 7 องค์ประกอบที่ใช้ทดสอบ ได้แก่ การระบุ การเข้าถึง การประเมิน การจัดการ การบูรณาการ การสร้างสรรค์ และ การสื่อสาร ซึ่งเป็นแบบวัดที่มีลักษณะเป็นการประเมินภาคปฏิบัติ วัดทั้งในส่วนที่เป็นทักษะด้านพุทธิพิสัยและทักษะด้านเทคนิค โดยลักษณะเครื่องมือประกอบด้วยข้อคำถามหลายรูปแบบทั้งแบบเขียนตอบและแบบให้สถานการณ์พร้อมภาระงาน จะเป็นการให้สถานการณ์สมมติที่มีความสอดคล้องกับสถานการณ์จริงในชีวิตประจำวันเช่น ให้สถานการณ์เรื่องสึนามิที่ประเทศญี่ปุ่น แล้วให้ภาระงานผู้สอบทำการสืบค้นข้อมูล สรุป และเขียนรายงานส่งครูเป็นต้น แต่ไม่พบลักษณะข้อสอบที่เป็นปรนัยหลายตัวเลือก

ตัวอย่างเครื่องมือ

ภาระงานที่ 2

วันที่ 12 มีนาคม 2011 เกิดเหตุการณ์สึนามิขึ้นที่ประเทศญี่ปุ่น ซึ่งเคยเกิดขึ้น ประเทศอินโดนีเซียในปี 2010 ทำให้ชนเผ่า Mentawai เสียชีวิตเป็นจำนวนมาก และ ที่มาเลเซียในปี 2004 นับเป็นภัยธรรมชาติที่มีผลกระทบต่อชีวิตรวมถึงสิ่งก่อสร้างต่างๆ วิทยากรของคุณต้องการให้ คุณสืบค้นข้อมูลเพิ่มเติมที่มีความน่าเชื่อถือซึ่งเกี่ยวข้องกับการเกิดซึนามิที่ประเทศมาเลเซีย และนำข้อมูลที่ได้จากการสืบค้นมานำเสนอหน้าชั้นเรียน

ขั้นตอน

(มีการอ่านขั้นตอนและแนะนำการปฏิบัติให้ผู้เข้าสอบก่อนการดำเนินงาน โดยจะต้องดำเนินงานเป็นลำดับขั้นดังนี้)

A) ให้เขียนชื่อเว็บไซต์ที่เป็นแหล่งข้อมูล (เป็นแบบเขียนตอบ) และให้เขียน คำค้น (เขียนตอบ) โดยให้เขียนได้หลายเว็บ

B) ให้นำข้อมูลที่สืบค้นได้ใส่ใน โฟลเดอร์ที่สร้างขึ้นโดยใช้ชื่อ folder ว่า TSUNAMI

C) อ่านข้อมูลที่สืบค้นได้ เลือกเก็บข้อมูลที่คุณรู้สึกว่ามันมีความสำคัญสำหรับนำเสนอหน้าชั้น

D) พัฒนาสื่อนำเสนอประมาณ 5 นาทีโดยใช้ข้อมูลที่ได้จากการสืบค้น เกี่ยวกับtsunami ในปี 2004 พร้อมทั้งข้อมูลเกี่ยวกับตัวคุณ (ชื่อ และ รหัสนักเรียน) ใน e-mail ที่จะส่งถึงวิทยากร

E) ในสื่อนำเสนอของคุณควรจะต้องมี

1. Link 3 เว็บไซต์ที่เป็นข้อมูลเกี่ยวกับสึนามิที่มาเลเซีย
2. link วิดีโอสึนามิ

ให้คะแนนโดยใช้เกณฑ์การตรวจให้คะแนน(รูบริค) มีการตรวจสอบคุณภาพเครื่องมือทั้งเชิงปริมาณและเชิงคุณภาพ ตรวจสอบความตรงและความเที่ยงของแบบวัด แต่ไม่ได้ระบุรายละเอียดของวิธี

2) งานวิจัยที่เกี่ยวข้องกับการใช้ไอซีทีในการจัดการเรียนการสอนรวมทั้งการวัดระดับการรู้เท่าทันไอซีที

Kiss and Gastelú (2015) ได้ศึกษาเกี่ยวกับการเปรียบเทียบระดับการรู้เท่าทันไอซีทีของนักศึกษาชาวเม็กซิกัน และชาวฮังกาเรียน เพื่อนำมาเป็นข้อมูลพิจารณาการพัฒนาการจัดการเรียนการสอน ตัวอย่างเป็นนักศึกษาจำนวน 720 คน เครื่องมือที่ใช้ในการวัดคือ แบบวัดแบบมาตรประมาณค่า 5ระดับ (Likert scaled) จำนวน 15 ข้อคำถามซึ่งจะให้นักศึกษาทำการประเมินตนเอง วิเคราะห์ข้อมูลโดยใช้สถิติทดสอบ Mann-Whitney test ซึ่งเป็นสถิติที่ใช้ทดสอบความสัมพันธ์สำหรับข้อมูลที่มีมาตรวัดที่ต่ำกว่ามาตรวัดอันตรภาค (Interval scale) ผลการศึกษาพบว่าระดับการรู้เท่าทันไอซีทีมีความแตกต่างกันในแต่ละประเทศโดยนักศึกษาชาวเม็กซิกันผู้หญิงและผู้ชายมีระดับการรู้เท่าทันไอซีทีใกล้เคียงกัน ในขณะที่นักศึกษาชาวฮังกาเรียนมีระดับการรู้เท่าทันไอซีทีแตกต่างกันระหว่างเพศหญิงและเพศชาย และพบว่าชาวเม็กซิกันมีระดับการรู้เท่าทันไอซีทีอยู่ในระดับสูง

Senkbeil and Ihme (2017) ทำงานวิจัยเกี่ยวกับการสร้างและการตรวจสอบทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร (ICT) โดยแบ่งมิติการใช้งาน ICT ออกเป็น 3 มิติ ได้แก่ แรงจูงใจของเครื่องมือ (การค้นหาข้อมูล และ การเรียนรู้และการทำงาน) แรงจูงใจ (ความบันเทิง และ การหนีจากปัญหาชีวิต) และแรงจูงใจในการปฏิสัมพันธ์ทางสังคม (การแลกเปลี่ยนทางสังคม และ การนำเสนอตนเอง) โดยมีการวิเคราะห์ความสามารถในการใช้ ICT และการควบคุมตนเองเกี่ยวกับ ICT ในทางการศึกษา จากกลุ่มตัวอย่างที่เป็นนักเรียนชาวเยอรมัน อายุระหว่าง 16 – 27 ปี จำนวน 323 คน และใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory factor analyses) สนับสนุนโมเดลที่สร้างขึ้นโดยใช้ 5 องค์ประกอบ ได้แก่ แรงจูงใจของเครื่องมือ แรงจูงใจ แรงจูงใจในการปฏิสัมพันธ์ทางสังคม ความสามารถในการใช้ ICT และ การควบคุมตนเองเกี่ยวกับ ICT ในการวัดความรู้ความเข้าใจทาง ICT ผู้วิจัยใช้แบบทดสอบแบบปรนัย จำนวน 60 ข้อ มีลักษณะเป็นปัญหาในชีวิตประจำวัน แบ่งเป็นการเข้าถึงสารสนเทศ จำนวน 23 ข้อ การสร้างสารสนเทศ จำนวน 11 ข้อ การบูรณาการและการจัดการสารสนเทศ จำนวน 16 ข้อ และ การประเมินบูรณาการ จำนวน 10 ข้อ โดยให้นักเรียนทำแบบทดสอบผ่านคอมพิวเตอร์ เพื่อรับมือกับปัญหาในสถานการณ์จริง โดยเลือกคำตอบที่เหมาะสมที่สุด พบว่าแบบทดสอบมีค่าความเที่ยง 0.77 และมีความตรงเชิงโครงสร้าง และ แรงจูงใจของ ICT มีความสัมพันธ์กับทักษะการรู้เท่าทันเทคโนโลยี สารสนเทศ และการสื่อสารและลักษณะเฉพาะของแต่ละบุคคล (พื้นหลังทางสังคมและความต้องการในการรับรู้) แตกต่างกันอย่างมีนัยสำคัญ

ช่อบุญ จิรานุภาพ (2554) ได้ทำการศึกษาเกี่ยวกับการพัฒนาตัวบ่งชี้รวมความสำเร็จการใช้ไอซีทีในการเรียนการสอนเพื่อเสริมสร้างทักษะการรู้ไอซีทีของนักเรียนระดับชั้นมัธยมศึกษาและ ตรวจสอบความสอดคล้องของโมเดลการวัดที่พัฒนาขึ้นกับข้อมูลเชิงประจักษ์ ตัวอย่างประกอบด้วย กลุ่มผู้ทรงคุณวุฒิจำนวน 11 คน กลุ่มผู้บริหารและครูผู้สอนในโรงเรียนมัธยมศึกษาจำนวน 541 คน และนักเรียนในระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 2,199 คน จาก 55 โรงเรียนที่ได้มาจากการสุ่มโรงเรียนทั่วประเทศ ผลการศึกษาได้ตัวบ่งชี้รวมความสำเร็จของการใช้ไอซีทีในการจัดการเรียนการสอนจำนวน 43 ตัวบ่งชี้ ครอบคลุมองค์ประกอบหลัก 4 ด้าน ได้แก่ ด้านบริบทจำนวน 10 ตัวบ่งชี้ ด้านปัจจัยนำเข้า จำนวน 13 ตัวบ่งชี้ ด้านกระบวนการจำนวน 15 ตัวบ่งชี้ และ ด้าน ผลลัพธ์จำนวน 5 ตัวบ่งชี้ และจากการตรวจสอบโมเดลตัวบ่งชี้ พบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ โดย

องค์ประกอบย่อยที่มีค่าน้ำหนักองค์ประกอบสูงที่สุดคือ ความสามารถของนักเรียนในการสร้างสรรค์ชิ้นงานด้วยการใช้ไอซีที

Markauskaite (2006) เปรียบเทียบประสบการณ์เกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสาร และการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารระหว่างเพศชายและเพศหญิง ตัวอย่างวิจัย คือ นักศึกษาระดับปริญญาโทสาขาวิชาการสอนชั้นปีที่ 1 ประเทศออสเตรเลีย จำนวน 120 คน ได้มาจากการเลือกตัวอย่างแบบอาสาสมัคร เครื่องมือวิจัย คือ แบบสอบถามซึ่งมีลักษณะเป็นมาตรประมาณค่า 5 และ 6 ระดับ ประกอบด้วยข้อมูลภูมิหลัง ประสบการณ์เกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสาร และการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสาร วิเคราะห์ข้อมูลโดยใช้สถิติทดสอบทีสำหรับตัวอย่างที่เป็นอิสระต่อกัน ผลการวิจัยพบว่า นักศึกษาเพศชายมีการใช้งานคอมพิวเตอร์ มีความสามารถด้านเทคโนโลยีสารสนเทศและการสื่อสาร มีการศึกษาหาความรู้เกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสาร และเผชิญสถานการณ์ที่ต้องใช้ความรู้เกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสารในการแก้ปัญหามากกว่านักศึกษาเพศหญิงอย่างมีนัยสำคัญทางสถิติ

Madigan, Goodfellow และ Stone (2007) เปรียบเทียบความมั่นใจในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร และการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารระหว่างเพศชายและเพศหญิง ตัวอย่างวิจัย คือ นักศึกษาปริญญาตรีชั้นปีที่ 1 ประเทศสหรัฐอเมริกา จำนวน 289 คน ได้มาจากการเลือกแบบเจาะจง การดำเนินการวิจัยแบ่งเป็น 2 ระยะ คือ ระยะที่ 1 ศึกษาความมั่นใจในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร เก็บรวบรวมข้อมูลโดยใช้แบบสอบถามซึ่งมีลักษณะเป็นมาตรประมาณค่า 5 ระดับ จำนวน 20 ข้อ ระยะที่ 2 ศึกษาการรู้เรื่องการใช้เทคโนโลยีสารสนเทศและการสื่อสาร เก็บรวบรวมข้อมูลโดยมอบหมายงานให้นักศึกษาจัดทำรายงานซึ่งต้องใช้ทักษะเกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสาร จำนวน 10 ทักษะ หากนักศึกษามีทักษะเกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสารทักษะใดให้ 1 คะแนน ถ้านักศึกษาไม่มีทักษะเกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสารทักษะใดให้ 0 คะแนน นั่นคือ มีคะแนนเต็ม 10 คะแนน ทั้งนี้ นักศึกษาที่ส่งรายงานมีจำนวน 38 คน โดยเป็นอาสาสมัครที่ตอบแบบสอบถามในระยะที่ 1 วิเคราะห์ข้อมูลโดยใช้การวิเคราะห์ความแปรปรวนทางเดียว ผลการวิจัยพบว่า นักศึกษาเพศชายมีความมั่นใจในการใช้เทคโนโลยีสารสนเทศและการสื่อสารมากกว่านักศึกษาเพศหญิงอย่างมีนัยสำคัญทางสถิติ แต่อย่างไรก็ตาม นักศึกษาเพศชายและเพศหญิงมีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

Tella และ Mutula (2008) เปรียบเทียบการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารระหว่างเพศชายและเพศหญิง ตัวอย่างวิจัย คือ นักศึกษาปริญญาตรีชั้นปีที่ 1-4 ประเทศบอตสวานา จำนวน 500 คน ได้มาจากการเลือกแบบเจาะจง เครื่องมือวิจัย คือ แบบสอบถามซึ่งมีลักษณะแบบถูกผิด และมาตรประมาณค่า 4 ระดับ ประกอบด้วยข้อมูลภูมิหลัง การรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสาร ประสบการณ์ในการใช้งานคอมพิวเตอร์ ประเภทการใช้งานคอมพิวเตอร์ และการใช้งานห้องสมุด วิเคราะห์ข้อมูลโดยสถิติทีสำหรับตัวอย่างที่เป็นอิสระต่อกัน ผลการวิจัยพบว่า นักศึกษาเพศชายมีประสบการณ์ในการใช้งานคอมพิวเตอร์ มีเวลาในการใช้งานคอมพิวเตอร์ และมีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารสูงว่านักศึกษาเพศหญิงอย่างมีนัยสำคัญทางสถิติ นอกจากนี้พบว่า นักศึกษาที่มีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารสูงมีแนวโน้มในการใช้บริการห้องสมุดสูงกว่านักศึกษาที่มีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารต่ำ

Agbatogun (2010) เปรียบเทียบการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารของครูที่มีเพศ คุณวุฒิทางการศึกษา และวิชาเอกแตกต่างกัน ตัวอย่างวิจัย คือ ครูระดับมัธยมศึกษาในโรงเรียนรัฐบาล ประเทศไนจีเรีย

จำนวน 364 คน ได้มาจากการสุ่มอย่างง่าย เครื่องมือวิจัย คือ แบบวัดการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสาร มีลักษณะเป็นมาตรประมาณค่า 4 ระดับ จำนวน 26 ข้อ วิเคราะห์ข้อมูลโดยใช้สถิติทดสอบทีสำหรับตัวอย่างที่เป็นอิสระต่อกัน และการวิเคราะห์ความแปรปรวนทางเดียว ผลการวิจัยพบว่า ครูที่มีเพศ และคุณวุฒิทางการศึกษาต่างกันมีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ส่วนครูที่มีวิชาเอกต่างกันมีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารแตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยครูกลุ่มวิชาเอกพาณิชยศาสตร์มีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารสูงกว่าครูกลุ่มวิชาเอกวิทยาศาสตร์และสังคมศาสตร์

Ritzhaupt, Liu, Dawson และ Barron (2013) ศึกษาผลของเพศ เชื้อชาติ และเศรษฐสถานะที่มีต่อการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสาร ตัวอย่างวิจัย คือ นักเรียนระดับมัธยมศึกษาตอนต้น จำนวน 5,990 คน ประเทศสหรัฐอเมริกา เครื่องมือวิจัย คือ ระบบการประเมินการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารซึ่งเป็นโปรแกรมคอมพิวเตอร์มีส่วนประกอบ 2 ส่วน คือ ภาระงานมุ่งเน้นการปฏิบัติ (performance-based tasks) จำนวน 67 ชิ้น และแบบสอบประเภทเลือกตอบ จำนวน 40 ข้อ ประกอบด้วยแบบสอบหลายตัวเลือกและแบบถูกผิด วิเคราะห์ข้อมูลโดยใช้การวิเคราะห์ความแปรปรวนหลายตัวแปรสามทาง ผลการวิจัยพบว่า เพศ เชื้อชาติ และเศรษฐสถานะส่งผลต่อการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารอย่างมีนัยสำคัญทางสถิติ โดยที่ปฏิสัมพันธ์ระหว่างเพศ เชื้อชาติ และเศรษฐสถานะไม่มีนัยสำคัญทางสถิติ ทั้งนี้นักเรียนเพศหญิงมีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารสูงกว่านักเรียนเพศชาย นักเรียนผิวขาวมีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารสูงกว่านักเรียนที่มีผิวสีอื่น อีกทั้งนักเรียนที่มีเศรษฐสถานะสูงมีการรู้เรื่องเทคโนโลยีสารสนเทศและการสื่อสารสูงกว่านักเรียนที่มีเศรษฐสถานะต่ำ

จากเอกสารและงานวิจัยที่เกี่ยวข้องในกลุ่มที่ 1 ซึ่งเกี่ยวกับการพัฒนาองค์ประกอบ เครื่องมือสำหรับวัดทักษะด้านเทคโนโลยี สารสนเทศและการสื่อสาร และการตรวจสอบคุณภาพเครื่องมือ และกลุ่มที่ 2 งานวิจัยที่เกี่ยวข้องกับการใช้ไอซีทีในการจัดการเรียนการสอนรวมทั้งการวัดระดับการรู้เท่าทันไอซีที พบว่าเป็นงานวิจัยที่ศึกษาในต่างประเทศเป็นส่วนใหญ่ ยังไม่พบงานวิจัยลักษณะนี้ในประเทศไทย องค์ประกอบที่ใช้ในการวัดจะมีลักษณะเป็นไปแนวทางเดียวกัน ที่วัดทักษะด้านพุทธิพิสัย และด้านเทคนิคเกี่ยวกับไอซีที เครื่องมือที่ใช้ในภาพรวม คือ แบบวัดและแบบประเมินตนเอง ลักษณะข้อคำถามส่วนใหญ่เป็นลักษณะของการให้ภาระงานผ่านทางสถานการณ์สมมติ โดยผ่านโปรแกรมเสมือน ผ่านแอปพลิเคชัน หรือ การสอบออนไลน์ ทั้งนี้ในแต่ละประเทศก็มีลักษณะของการใช้ไอซีที และบริบทในการใช้ไอซีทีแตกต่างกัน ซึ่งหากมีการพัฒนาเครื่องมือสำหรับวัดทักษะด้านไอซีทีในการใช้ไอซีทีเพื่อการเรียนรู้ที่เข้ากับบริบทของประเทศไทยก็น่าที่จะเป็นประโยชน์ในอนาคต สำหรับใช้ในการวัดระดับทักษะด้านไอซีทีของกลุ่มคนที่เข้าสู่วัยทำงาน ว่ามีระดับการรู้เท่าทันไอซีทีอยู่ในระดับใด เพื่อพัฒนากำลังคนให้สามารถดำรงชีวิตในศตวรรษที่ 21 ได้อย่างมีศักยภาพในการพัฒนาประเทศต่อไป

การวิจัยในเรื่องนี้จึงไม่เพียงแต่เป็นการพัฒนาองค์ประกอบเพื่อสร้างเครื่องมือวัดเท่านั้น แต่เป็นการพัฒนาองค์ประกอบและข้อสอบเพื่อนำไปสร้างคลังข้อสอบให้สามารถนำไปใช้ในการทดสอบแบบปรับเหมาะ (computer adaptive testing) ที่ผู้สอบแต่ละคนไม่จำเป็นต้องทำข้อสอบเหมือนกันภายใต้เวลาเดียวกัน แต่สามารถวัดความรู้ความสามารถของผู้สอบได้อย่างมีความแม่นยำ ดังนั้น การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วย Computer ของงานวิจัยนี้ได้ออกแบบให้เหมาะสมกับบริบทการเรียนรู้ในระดับอุดมศึกษาของประเทศไทย และสามารถใช้เป็นแนวทางในการสร้างแบบวัดทักษะอื่นๆที่จำเป็นต่อไป

ตอนที่ 2 มโนทัศน์เกี่ยวกับการวิเคราะห์คุณภาพข้อสอบเพื่อจัดทำคลังข้อสอบ

การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ในการวิจัยนี้ หลังจากที่ได้ข้อสอบตามองค์ประกอบที่สังเคราะห์แล้ว จำเป็นจะต้องตรวจสอบคุณภาพของข้อสอบที่สร้างขึ้นก่อนและหลังจากการนำไปใช้ สำหรับการตรวจสอบคุณภาพข้อสอบหลังนำไปใช้ต้องอาศัยผลการตอบของผู้สอบมาวิเคราะห์ และทฤษฎีที่เป็นที่ยอมรับในปัจจุบัน คือ ทฤษฎีการตอบสนองข้อสอบ ในกรณีที่ข้อสอบเป็นแบบเลือกตอบ มีการตอบสนองเป็นแบบ 0 และ 1 มีประเภทของโมเดลการตอบสนองข้อสอบ (Item Response Model) ที่นิยมเป็น 3 ประเภท และการที่จะวิเคราะห์โดยใช้โมเดลใดต้องมีการตรวจสอบความเหมาะสมของโมเดลการตอบสนองข้อสอบกับข้อมูล หลังจากที่ได้ข้อสอบที่มีคุณภาพผ่านการวิเคราะห์แล้วจึงนำไปจัดทำคลังข้อสอบ ดังนั้น ในตอนนี้ ได้นำเสนอเนื้อหาที่เกี่ยวข้องคือ การวิเคราะห์คุณภาพข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ การตรวจสอบความเหมาะสมของโมเดลการตอบสนองข้อสอบกับข้อมูล และการจัดทำคลังข้อสอบ ดังมีรายละเอียด ดังนี้

2.1 การวิเคราะห์คุณภาพข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ (IRT)

ทฤษฎีการตอบสนองข้อสอบเป็นโมเดลทางคณิตศาสตร์ที่อธิบายความสัมพันธ์ระหว่างความสามารถของผู้สอบกับความน่าจะเป็นในการตอบข้อสอบได้ถูกต้อง เป็นทฤษฎีที่มีความเชื่อว่า พฤติกรรมการตอบสนองข้อสอบเป็นสิ่งที่สังเกตไม่ได้โดยตรง ซึ่งจะถูกกำหนดโดยคุณลักษณะภายใน (Trait) หรือความสามารถ (Ability) ที่อยู่ภายในตัวบุคคล โมเดลที่แสดงความสัมพันธ์ระหว่างระดับความสามารถคุณลักษณะของข้อสอบและโอกาสของการตอบข้อสอบได้ถูก ในรูปของโค้งลักษณะข้อสอบ (Item Characteristic; ICC) ซึ่งมีลักษณะเป็นฟังก์ชันโลจิส (Logistic function) หรือฟังก์ชันปกติสะสม (Normal ogive function) ซึ่งมีการกำหนดลักษณะของข้อสอบด้วยพารามิเตอร์ ความยาก (b) อำนาจจำแนก (a) และโอกาสการเดาข้อสอบถูก (c)

ข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบ

- ความเป็นเอกมิติ (Unidimensionality) ข้อคำถามทุกข้อในแบบสอบมุ่งวัดเพียงคุณลักษณะเดียว หรือ ความสามารถเดียว
- ความเป็นอิสระ (Local Independence) ผลการตอบข้อสอบแต่ละข้อจะต้องเป็นอิสระจากกัน คือ โมเดลการตอบสนองข้อสอบมีเพียงความสามารถที่แท้จริงของผู้สอบเท่านั้นที่มีอิทธิพลต่อการตอบรายข้อ
- ความสัมพันธ์ที่แท้จริงระหว่างคุณลักษณะภายในกับคำตอบที่มาจากข้อคำถามที่วัดคุณลักษณะนั้น (true relationship) คือ ICC สามารถอธิบายคุณลักษณะภายในของผู้สอบได้ (Hambleton, Swaminathan, & Rogers, 1991)

โมเดลการตอบสนองข้อสอบ (Item Response Model) แบ่งออกเป็น 3 ประเภท ดังนี้

1. โมเดลการตอบข้อสนองข้อสอบที่ใช้กับคะแนนแบบ 2 ค่า (Dichotomous) เป็นข้อสอบที่ตรวจให้คะแนน 0,1 (ตอบผิดได้ 0 คะแนน, ตอบถูกได้ 1 คะแนน) เช่น Guttman Perfect Scale, Latent Distance Model, Linear Model, One-Parameter Model, Two-Parameter Model และ Three-Parameter Model เป็นต้น
2. โมเดลการตอบสนองข้อสอบที่ใช้กับคะแนนมากกว่า 2 ค่า (Polytomous) เช่น Norminal Response Model, Grade Response Model, Partical Credit Model เป็นต้น
3. โมเดลการตอบสนองข้อสอบที่ใช้กับคะแนนแบบต่อเนื่อง (Continuous) เช่น Continuous Response Model เป็นต้น

ในที่นี่ขอนำเสนอโมเดลการตอบสนองข้อสอบแบบ 2 ค่า ที่ใช้กันแพร่หลาย ได้แก่

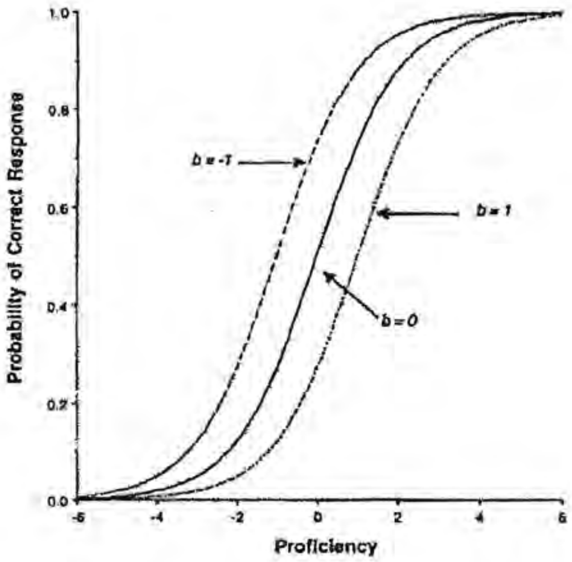
โมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ (One-Parameter Model: 1PL) มีข้อตกลงเบื้องต้นว่า ข้อสอบแต่ละข้อมีพารามิเตอร์ c = 0 (โอกาสการเดาข้อสอบถูก) และพารามิเตอร์ a เท่ากัน (อำนาจจำแนก) แต่มีความแตกต่างกันเฉพาะพารามิเตอร์ b (ความยาก)

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

เมื่อ $P_i(\theta)$ = ความน่าจะเป็นที่ผู้ตอบมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

b_i = ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด θ ที่มีโอกาสตอบข้อสอบถูก 0.50

e = 2.718



ภาพ 2.3 โค้งคุณลักษณะข้อสอบแบบ 1 พารามิเตอร์ (Wainer et al., 2000: 68)

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ มีลักษณะสำคัญ ดังนี้

b_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

a_i = มีค่าคงที่

C_i = 0

โมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (Two-Parameter Model: 2 PL) มีข้อตกลงเบื้องต้นว่า ข้อสอบแต่ละข้อมีพารามิเตอร์ c = 0 (โอกาสการเดาข้อสอบถูก) มีความแตกต่างของ พารามิเตอร์ a และ b (อำนาจจำแนก และความยาก) ดังสมการ

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

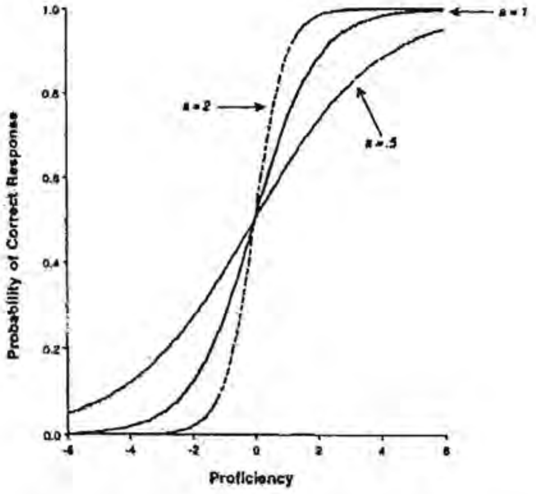
เมื่อ $P_i(\theta)$ = ความน่าจะเป็นที่ผู้ตอบมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

b_i = ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด θ ที่มีโอกาสตอบข้อสอบถูก 0.50

a_i = ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นค่าความชันของ ICC ณ ตำแหน่ง b_i

e = 2.718

D = 1.70



ภาพ 2.4 โค้งคุณลักษณะข้อสอบแบบ 2 พารามิเตอร์ (Wainer et al., 2000: 70)

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ มีลักษณะสำคัญ ดังนี้

b_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

a_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

C_i = 0

โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (Three-Parameter Model: 3 PL) มีข้อตกลงเบื้องต้นว่าข้อสอบแต่ละข้อมีความแตกต่างกันได้ทั้งพารามิเตอร์ a, b, c (ความยาก, อำนาจจำแนกและโอกาสการเดาข้อสอบ) ดังสมการ

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

เมื่อ $P_i(\theta)$ = ความน่าจะเป็นที่ผู้ตอบมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

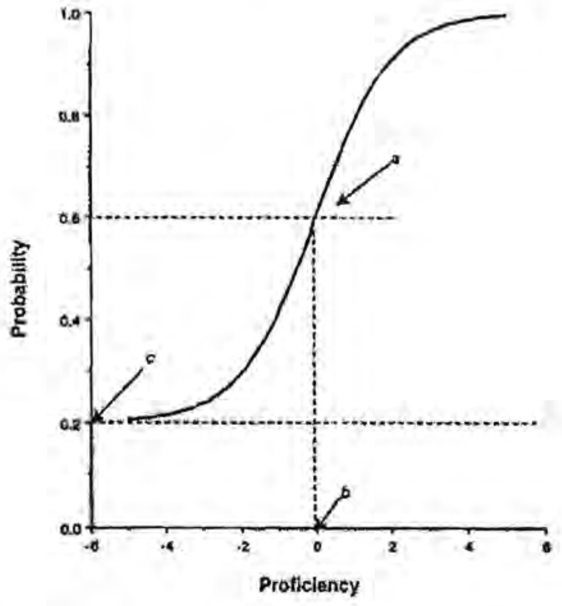
b_i = ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด θ ที่มีโอกาสตอบข้อสอบถูก $\frac{1+C_i}{2}$

a_i = ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นค่าความชันของ ICC ณ ตำแหน่ง b_i

C_i = ค่าพารามิเตอร์โอกาสเดาข้อสอบได้ถูก

e = 2.718

D = 1.70



ภาพ 2.5 โค้งคุณลักษณะข้อสอบแบบ 3 พารามิเตอร์ (Wainer et al., 2000: 71)

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ มีลักษณะสำคัญ ดังนี้

b_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

a_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

C_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

2.2 การตรวจสอบความเหมาะสมของโมเดลการตอบสนองข้อสอบกับข้อมูล

โมเดลการตอบสนองข้อสอบ (item response model) เป็นเครื่องมือสำคัญในการพัฒนาคลังข้อสอบและการทดสอบแบบปรับเหมาะ (computerized adaptive testing) ในปัจจุบัน (Embretson & Reise, 2000; Revicki & Reise, 2015; Wainer et al., 2000; van der Linden & Glas, 2000, 2010; Way & Robin, 2016) เช่น การประยุกต์การทดสอบแบบปรับเหมาะสำหรับประเมินการเรียนรู้จากเกมและสถานการณ์จำลอง (simulation- and game-based assessment: von Davier & Mislevy, 2016) และใช้เป็นรากฐานการพัฒนาการทดสอบแบบปรับเหมาะหลายขั้นตอน (computerized adaptive multistage testing: Yan, von Davier, & Lewis, 2014) รวมทั้งซอฟแวร์คอมพิวเตอร์ต่างๆ เช่น SimulCAT (Han, 2012), MSTGen (Han, 2013), Firestar (Choi, 2018), catR และ mstR (Magis, Yan, & von Davier, 2017)

คุณสมบัติเด่นของโมเดลการตอบสนองข้อสอบที่เป็นประโยชน์ต่อการทดสอบแบบปรับเหมาะมีหลายประการ กล่าวคือ ประการที่ 1 ค่าพารามิเตอร์ความสามารถของผู้สอบไม่แปรเปลี่ยนตามชุดข้อสอบ และค่าพารามิเตอร์ข้อสอบไม่แปรเปลี่ยนตามกลุ่มผู้สอบ ซึ่งความไม่แปรเปลี่ยนนี้จากการผนวกรวมสารสนเทศข้อสอบเข้าไปในกระบวนการประมาณค่าความสามารถ และในขณะเดียวกันก็ผนวกรวมสารสนเทศความสามารถผู้สอบเข้าไปในกระบวนการประมาณค่าพารามิเตอร์ข้อสอบ (Embretson & Reise, 2000; Hambleton et al., 1991)

ประการที่ 2 โมเดลการตอบสนองข้อสอบประมาณค่าความคลาดเคลื่อนของค่าความสามารถผู้สอบแต่ละคนไม่เท่ากัน โดยเฉพาะผู้สอบที่มีความสามารถสูงมากและต่ำมาก ซึ่งโดยปกติมักจะมีค่าความคลาดเคลื่อนสูงกว่าผู้สอบที่มีความสามารถปานกลาง สำหรับการทดสอบแบบปรับเหมาะนั้น ผู้สอบที่มีความสามารถสูงมากและต่ำมาก นอกจากจะมีค่าความคลาดเคลื่อนของความสามารถไม่เท่ากันแล้ว ยังช่วยให้ค่าความคลาดเคลื่อนลดลงด้วยเพราะผู้สอบจะได้ทำข้อสอบที่เหมาะสมกับความสามารถจำนวนเพิ่มขึ้น และทำข้อสอบที่ยากเกินความสามารถจำนวนลดลง นั่นคือ ข้อสอบที่ผู้สอบแต่ละคนตอบนั้นจะทำให้สารสนเทศความสามารถของผู้สอบเพิ่มขึ้น และทำให้ค่าคลาดเคลื่อนของความสามารถลดลง (Hambleton et al., 1991)

ประการที่ 3 เป็นคุณสมบัติสืบเนื่องจากประการที่ 2 กล่าวคือ แนวคิดดั้งเดิมที่ให้ผู้สอบซึ่งโดยธรรมชาติแล้วมีความสามารถแตกต่างกัน ทำแบบสอบที่มีจำนวนข้อสอบคงที่เหมือนกันทุกคน เป็นแนวคิดที่ไม่มีประสิทธิภาพในการประมาณความสามารถของผู้สอบแต่ละคน (Lord, 1980) ข้อสอบที่ให้ผู้สอบแต่ละคนทำนั้นสามารถมีจำนวนน้อยได้ ในขณะที่สามารถอธิบายความสามารถของผู้สอบได้เพิ่มขึ้นและไม่สูญเสียความแม่นยำในการประมาณค่าความสามารถ นั่นคือ แบบสอบฉบับสั้นสามารถให้ผลการวัดคงเส้นคงวามากกว่าแบบสอบฉบับยาว (Embretson & Reise, 2000; Hambleton et al., 1991)

ดังที่กล่าวข้างต้น โมเดลการตอบสนองข้อสอบมีบทบาทสำคัญอย่างมากต่อการทดสอบแบบปรับเหมาะ โมเดลการตอบสนองข้อสอบที่ไม่เหมือนกันย่อมให้ผลการประมาณค่าพารามิเตอร์ต่างๆ แตกต่างกันด้วย โดยทั่วไปแล้ว โมเดลการตอบสนองข้อสอบที่นิยมใช้ในปัจจุบันแบ่งออกเป็น 3 กลุ่มตามจำนวนพารามิเตอร์ข้อสอบที่ประมาณได้จากโมเดล และนิยมเขียนให้อยู่ในรูปฟังก์ชันโลจิสติกส์ (logistic function) มากกว่าฟังก์ชันการแจกแจงปกติมาตรฐาน (standard normal distribution: Raykov & Marcoulides, 2018) ได้แก่ โมเดลตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (three parameter logistic-item response model) หรือ โมเดล 3pl ซึ่งประกอบด้วยพารามิเตอร์ความยาก (item difficulty) พารามิเตอร์การจำแนก (item discrimination) และพารามิเตอร์การเดา (guessing) โมเดลตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (two parameter logistic-item response model) หรือ โมเดล 2pl ซึ่งประกอบด้วยพารามิเตอร์ความยากและ พารามิเตอร์การจำแนก และ

โมเดลตอบสนองข้อสอบแบบ 1 พารามิเตอร์ (one parameter logistic-item response model) หรือ โมเดล 1pl ซึ่งประกอบด้วยพารามิเตอร์ความยากเพียงอย่างเดียว

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}} \tag{2.2.1}$$

สมการ 2.2.1 เป็นโมเดล 3pl ซึ่งแสดงถึงความสัมพันธ์ระหว่างความน่าจะเป็น P ของผู้สอบที่มีความสามารถ θ ที่จะตอบข้อสอบ i ซึ่งมีความยากเท่ากับ b การจำแนกเท่ากับ a และโอกาสเดา c ได้ถูกต้อง ค่า θ ซึ่งแสดงถึงความสามารถของผู้สอบแต่ละคน จะไม่เท่ากัน และส่งผลให้ผู้สอบแต่ละคนมีโอกาสตอบข้อสอบข้อเดียวกันได้ถูกต้องไม่เท่ากันด้วย ผู้สอบที่มีความสามารถมาก (ค่า θ มาก) จะมีโอกาสตอบข้อสอบถูกสูงกว่าผู้สอบที่มีความสามารถน้อย (ค่า θ น้อย)

เมื่อกำหนดให้ค่าโอกาสการเดาเท่ากับ 0 โมเดล 2.2.1 จะกลายเป็นโมเดล 2pl ที่มีเฉพาะพารามิเตอร์ความยาก และการจำแนก ดังสมการ 2.2.2

$$P_i(\theta) = 0 + (1 - 0)\frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}} = \frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}} \tag{2.2.2}$$

สมการ 2.2.2 แสดงให้เห็นว่า พารามิเตอร์ความสามารถของผู้สอบเป็นฟังก์ชันของพารามิเตอร์ความยากและการจำแนกข้อสอบเท่านั้น

เมื่อกำหนดให้พารามิเตอร์การจำแนกเท่ากับ 1 โมเดล 2.2.2 จะกลายเป็นโมเดล 1pl ที่พารามิเตอร์ความสามารถของผู้สอบเป็นฟังก์ชันของความยากข้อสอบ ดังสมการ 2.2.3

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \tag{2.2.3}$$

ความสัมพันธ์ระหว่างสมการ 2.2.1, 2.2.2 และ 2.2.3 แสดงให้เห็นว่า โมเดล 2pl และโมเดล 1pl เป็นโมเดลย่อย (nested model) ที่ซ้อนอยู่ในโมเดล 3pl จึงทำให้ความสอดคล้องกลมกลืนระหว่างโมเดลทั้ง 3 โมเดลกับข้อมูลคำตอบของข้อสอบ ที่จะนำไปใช้ในกระบวนการทดสอบแบบปรับเหมาะสามารถเปรียบเทียบกันได้ เพื่อหาโมเดลที่สอดคล้องกลมกลืนกับข้อมูลคำตอบมากที่สุด หรือหาโมเดลที่สามารถประมาณค่าพารามิเตอร์ความสามารถได้แม่นยำ (precision) มากที่สุด

การเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลทั้ง 3 มีหลักการเช่นเดียวกับการตรวจสอบความไม่แปรเปลี่ยนของค่าพารามิเตอร์ข้อสอบ (Raykov & Marcoulides, 2018) กล่าวคือ ค่าพารามิเตอร์ข้อสอบแต่ละข้อที่ประมาณได้จากโมเดล 3pl จะแตกต่างกัน ดังนั้น

$$c_1 \neq c_2 \neq c_3 \neq \cdots \neq c_I$$
$$a_1 \neq a_2 \neq a_3 \neq \cdots \neq a_I$$
$$b_1 \neq b_2 \neq b_3 \neq \cdots \neq b_I$$

โดยที่ i หมายถึงข้อสอบตั้งแต่ข้อที่ 1 ถึงข้อที่ I และเมื่อต้องการตรวจสอบว่าโมเดล 2pl สอดคล้องกลมกลืนกับข้อมูลคำตอบมากกว่าโมเดล 3pl หรือไม่ จึงหมายถึงการตรวจสอบว่า ค่าโอกาสการเดาของข้อสอบทุกข้อสามารถกำหนดให้เท่ากันเท่ากับ 0 หรือไม่แปรเปลี่ยนได้หรือไม่ ดังนั้น $c_1 = c_2 = c_3 = \cdots = c_i = 0$ และเมื่อโมเดล 2pl สอดคล้องกลมกลืนกับข้อมูลมากกว่าโมเดล 3pl อย่างมีนัยสำคัญ ย่อมแสดงให้เห็นว่า ค่าโอกาสเดาของข้อสอบแต่ละข้อไม่ได้ส่งผลอย่างมีนัยสำคัญต่อการประมาณค่าพารามิเตอร์ความสามารถผู้สอบ

ในทำนองเดียวกัน การเปรียบเทียบว่าโมเดล 1pl สอดคล้องกลมกลืนมากกว่าโมเดล 2pl หรือไม่ จึงเป็นการตรวจสอบว่า เมื่อกำหนดให้ ข้อสอบแต่ละข้อมีพารามิเตอร์การจำแนกเท่ากันเท่ากับ 1 ($a_1 = a_2 = a_3 = \cdots = a_i = 1$) และพารามิเตอร์โอกาสเดาเท่ากันเท่ากับ 0 แล้ว ค่าพารามิเตอร์ความสามารถของผู้สอบที่ประมาณได้จากโมเดล 1pl แม่นยำเพิ่มขึ้นหรือไม่ โมเดล 1pl ที่สอดคล้องกลมกลืนกับข้อมูลมากกว่าโมเดล 2pl จึงแสดงให้เห็นว่า ค่าพารามิเตอร์ความยากข้อสอบแต่ละข้อเพียงพอที่จะให้ประมาณค่าความสามารถผู้สอบ

การเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดล 2 โมเดลสามารถพิจารณาได้จากการทดสอบอัตราส่วนความเป็นไปได้ (likelihood-ratio test) ที่ว่าโมเดลที่สนใจ 2 โมเดลนั้นเหมือนกันหรือไม่ หรือเรียกว่าการทดสอบ LR โดยค่า log-likelihood สูงสุดของแต่ละโมเดลที่มาจากข้อมูลคำตอบเดียวกันและมาจากการประมาณด้วยวิธีความเป็นไปได้สูงสุด (maximum likelihood estimation) จะไม่ได้บอกความซับซ้อนของแต่ละโมเดล แต่เมื่อหาค่าผลต่างของค่า log-likelihood สูงสุดจาก 2 โมเดล ก็สามารถอธิบายในเบื้องต้นว่า โมเดลที่มีจำนวนพารามิเตอร์มากกว่าช่วยทำให้โมเดลนั้นสอดคล้องกลมกลืนกับข้อมูลคำตอบเพิ่มขึ้นหรือไม่ ตัวอย่างเช่น นักวิจัยเปรียบเทียบระหว่างโมเดล 3pl และโมเดล 2pl เพื่อต้องการตรวจสอบว่าพารามิเตอร์การเดาของข้อสอบแต่ละข้อที่เพิ่มเข้าไปในโมเดล 2pl นั้นช่วยทำให้โมเดล 3pl สอดคล้องกลมกลืนกับข้อมูลคำตอบเพิ่มขึ้นหรือไม่ การทดสอบ LR คำนวณได้จากสมการ 2.2.4

$$\Delta l = -2(l_0 - l_1) \tag{2.2.4}$$

โดยที่ l_0 หมายถึงค่า log-likelihood สูงสุดของโมเดลย่อยที่มีจำนวนพารามิเตอร์น้อยกว่า และ l_1 หมายถึงค่า log-likelihood สูงสุดของโมเดลที่มีจำนวนพารามิเตอร์มากกว่า และเนื่องจากค่า Δl มีการแจกแจงโดยประมาณใกล้เคียงกับการแจกแจงแบบ χ^2 ที่เกิดจากการผสมกันของการแจกแจงแบบ χ^2 2 ชุด (mixture of two χ^2 distribution) นัยสำคัญทางสถิติของค่า Δl จึงสามารถตรวจสอบได้จากค่าในตาราง χ^2 ณ องศาอิสระ (degree of freedom) เท่ากับผลต่างขององศาอิสระจาก 2 โมเดล ถ้า Δl มีนัยสำคัญทางสถิติ นั่นหมายความว่า โมเดลที่มีจำนวนพารามิเตอร์มากกว่าเป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลมากกว่า ในทางตรงกันข้าม ถ้า Δl ไม่มีนัยสำคัญทางสถิติ นั่นหมายความว่า ทั้ง 2 โมเดลสอดคล้องกลมกลืนกับข้อมูลเท่ากัน ดังนั้นจึงไม่จำเป็นต้อง

ประมาณค่าความสามารถผู้สอบด้วยค่าพารามิเตอร์ข้อสอบที่เพิ่มเข้ามาในโมเดล และสามารถกำจัดพารามิเตอร์ข้อสอบเหล่านั้นออกไปได้

นอกจากนี้ ค่า –2 x log-likelihood หรือเรียกว่าค่า deviance (G^2) ของแต่ละโมเดล สามารถตีความในเชิงเปรียบเทียบได้ว่า โมเดลใดที่มีค่า G^2 ต่ำที่สุด จะเป็นโมเดลที่มีความแปรปรวนที่ไม่สามารถอธิบายได้หลงเหลือน้อยที่สุด (Wilson & Moore, 2012) นั่นคือ จะเป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลมากที่สุด

สถิติทดลอง LR มีจุดด้อยเช่นเดียวกับการทดสอบ χ^2 เกี่ยวกับความไวต่อกลุ่มตัวอย่างขนาดใหญ่ กล่าวคือ ค่า Δl เพียงเล็กน้อยก็อาจจะมีนัยสำคัญทางสถิติได้ (Werner & Schermelleh-Engel, 2010) ดังนั้นจึงมีดัชนีอีก 2 ตัวที่นิยมใช้คู่กับสถิติทดสอบ LR และ G^2 เพื่อตรวจสอบความสอดคล้องกลมกลืนของโมเดล คือ ดัชนี Akaike information criterion (Akaike, 1974) หรือที่เรียกว่าดัชนี AIC และดัชนี Bayesian information criterion (Schwarz ,1978) หรือที่เรียกว่าดัชนี BIC

ดัชนี AIC แสดงให้เห็นถึงสารสนเทศที่สูญเสียจากการเลือกใช้แต่ละโมเดลในการประมาณค่าพารามิเตอร์ความสามารถผู้สอบ ค่า AIC ของโมเดลย่อยแต่ละโมเดลสามารถเปรียบเทียบกันได้โดยตรง โดยโมเดลที่มีค่า AIC ต่ำที่สุด ซึ่งหมายถึงโมเดลที่สูญเสียสารสนเทศเกี่ยวกับความสามารถผู้สอบน้อยที่สุด จะเป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลมากที่สุด ดัชนี AIC คำนวณได้จากสมการ 2.2.5 โดย p หมายถึงจำนวนพารามิเตอร์ที่ต้องประมาณค่าทั้งหมดในแต่ละโมเดล

$$\text{AIC} = -2 \times (\log - \text{likelihood}) + 2p \tag{2.2.5}$$

ดัชนี BIC ให้สารสนเทศเช่นเดียวกับดัชนี AIC แต่นำขนาดกลุ่มตัวอย่างมาใช้อธิบายสารสนเทศที่สูญเสียในแต่ละโมเดล และปรับลดอิทธิพลของขนาดกลุ่มตัวอย่างที่ส่งผลต่อนัยสำคัญทางสถิติของโมเดล โมเดลที่มีค่า BIC ต่ำที่สุด หมายถึงเป็นโมเดลที่สูญเสียสารสนเทศน้อยที่สุด และเป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลคำตอบมากที่สุดด้วย ดัชนี BIC คำนวณได้จากสมการ 2.2.6 โดย n หมายถึงจำนวนผู้สอบทั้งหมดที่จะนำข้อมูลมาใช้วิเคราะห์

$$\text{BIC} = -2 \times (\log - \text{likelihood}) + p(\log n) \tag{2.2.6}$$

งานวิจัยเรื่องนี้ ได้ตรวจสอบความเหมาะสมของโมเดลการตอบสนองข้อสอบกับข้อมูลเชิงประจักษ์ที่เก็บจากตัวอย่าง และเลือกใช้โมเดลที่เหมาะสมที่สุดจากการวิเคราะห์ตามหลักการที่ได้นำเสนอข้างต้น เพื่อให้ได้ข้อสอบที่มีการประมาณค่าที่ถูกต้องแม่นยำ และนำไปจัดทำคลังข้อสอบเพื่อใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ต่อไป

2.3 การจัดทำคลังข้อสอบ

คลังข้อสอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เป็นการรวบรวมข้อสอบที่มีคุณภาพ ตามกระบวนการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ การทดสอบปรับเหมาะจะมีประสิทธิภาพสูงขึ้นถ้ามีคลังข้อสอบขนาดใหญ่ ข้อสอบดังกล่าวได้รับการประมาณค่าพารามิเตอร์โดยใช้ ทฤษฎีการตอบสนองข้อสอบ

Weiss (1982) เสนอว่าจำนวนข้อสอบในคลังข้อสอบควรมีอย่างน้อย 100-200 ข้อ Allen and Yen (1979) เสนอว่าต้องมีจำนวนข้อสอบจำนวน 2^n ข้อ เมื่อ n = จำนวนข้อสอบที่ใช้ในการทดสอบ แต่ถ้าการทดสอบมีการกำหนดจำนวนข้อที่จะใช้คงที่เหมือนกันทุกคน จำนวนข้อสอบควรมีอย่างน้อย n (n+1)/2

Embretson and Reise (2000) ได้กล่าวเกี่ยวกับคลังข้อสอบ cat ไว้ดังนี้ ความเห็นทั่วไปคือว่า ถ้าการทดสอบด้วย cat นั้นต้องการทดสอบความสามารถของคนที่มีความสามารถตั้งแต่ระดับต่ำไปจนถึงระดับสูง (ประมาณว่าต้องการทดสอบความสามารถของบุคคลที่ครอบคลุมพิสัยความสามารถทั้งหมดอาจจะเป็นตั้งแต่ -3 ถึง +3 หรือตั้งแต่ -4/-5 ไปจนถึง +4/+5 เพราะเขาหมายถึงครอบคลุมพิสัยทั้งหมดตามที่ผู้พัฒนา cat ต้องการ) คลังข้อสอบก็ควรจะมีข้อสอบที่จำแนกได้เป็นอย่างดี และมีความยากครอบคลุมพิสัยความสามารถของบุคคลจำนวนมากพอ (Weiss, 1982) (ตรงนี้หมายความว่า ข้อสอบที่มีคุณสมบัติแบบนี้ ต้องมีมากพอ แต่ไม่ได้หมายความว่าข้อสอบทุกข้อต้องเป็นแบบนี้ เพราะถ้าต้องการให้ข้อสอบทุกข้อมีคุณสมบัติแบบนี้ ก็จะเป็นคลังข้อสอบ cat ในอุดมคติ แล้วก็ Weiss อธิบายด้วยแนวคิดแบบ 2pl irt model ไม่ใช่ 1pl และไม่ใช่ 3pl) และเมื่อคลังข้อสอบมีข้อสอบแบบนี้มากพอ ก็จะทำให้สามารถวัดความสามารถของบุคคลได้อย่างถูกต้อง

ดังนั้นจำนวนข้อสอบในคลังแบบ cat ต้องมีเท่าไรนั้น ยังไม่มีกฎเกณฑ์ตายตัว ซึ่งถ้าให้ประมาณอย่างคร่าวๆ Embretson and Reise (2000) ก็บอกว่า น่าจะสัก 100 ข้อ แต่จำนวน 100 ข้อก็ยังขึ้นอยู่กับประเภทของเนื้อหาที่มุ่งวัดด้วย (เช่นคลังข้อสอบวัดบุคลิกภาพ คลังข้อสอบวัดความสามารถทางคณิตศาสตร์) และบางครั้งก็เป็นเรื่องยากที่จะทำให้ข้อสอบทั้ง 100 ข้อมีคุณสมบัติจำแนกเป็นอย่างดี และมีความยากครอบคลุมทุกช่วงความสามารถ ข้อเสนอของ Embretson and Reise (2000) จึงบอกว่า สำหรับการทดสอบโดยทั่วๆ ไป ที่ไม่ได้มีวัตถุประสงค์การทดสอบเฉพาะ หรือพิเศษ (เช่น คัดเลือกนักเรียนเข้าโครงการความสามารถพิเศษ) คลังข้อสอบ cat ที่มีสัก 100 ข้อก็น่าจะพอใช้ได้

Molina, Pareja, and Sanmartin (2008) ได้เสนอขั้นตอนการสร้างคลังข้อสอบที่จะมาใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เพื่อให้ได้คลังข้อสอบที่มีประสิทธิภาพ ประกอบด้วย 5 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 อธิบายข้อมูลและเกณฑ์ต่าง ๆ ที่เกี่ยวข้องของข้อสอบ (เช่น รูปแบบของข้อสอบ คำอธิบายข้อสอบ) ระหว่างการดำเนินการสอบ

ขั้นตอนที่ 2 คัดเลือกข้อสอบที่เหมาะสม และตรงตามเกณฑ์ที่กำหนดในคลังข้อสอบที่ได้จากการบริหารจัดการสอบโดยใช้คอมพิวเตอร์ หรือการจัดการสอบโดยใช้กระดาษคำตอบ (paper and pencil)

ขั้นตอนที่ 3 พัฒนาฐานข้อมูลของคลังข้อสอบ จากชุดข้อสอบที่คัดเลือกจากการจัดการสอบโดยใช้คอมพิวเตอร์ หรือการจัดการสอบโดยใช้กระดาษคำตอบ เพื่อนำมาใช้ในการบริหารการสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ขั้นตอนที่ 4 การรวบรวมข้อสอบ ทำการแยกข้อสอบที่ได้จากการบริหารการสอบโดยใช้คอมพิวเตอร์ หรือการจัดสอบโดยใช้กระดาษคำตอบ และนำข้อสอบที่ผ่านเกณฑ์การคัดเลือกมาตัดสินใจคัดเลือกอีกครั้ง แล้วนำไปพัฒนาเป็นคลังข้อสอบที่จะนำไปใช้

ขั้นตอนที่ 5 วิเคราะห์ความแตกต่างของกลุ่มข้อสอบโดยใช้คอมพิวเตอร์ หรือการจัดการสอบโดยใช้กระดาษคำตอบ กับกลุ่มข้อสอบในคลังข้อสอบที่พัฒนาขึ้น สามารถนำไปเป็นข้อมูลในการพัฒนาปรับปรุงคลังข้อสอบใหม่ให้มีคุณภาพ

การวิเคราะห์ข้อสอบเพื่อจัดทำคลังข้อสอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ โดยมีความเชื่อว่าข้อสอบที่สร้างขึ้นนั้นวัดตัวแปรคุณลักษณะแฝงหรือระดับความสามารถของผู้สอบ โดยข้อสอบจะประมาณค่าพารามิเตอร์โดยใช้ทฤษฎีการตอบสนองข้อสอบ ซึ่งข้อสอบในคลังข้อสอบจะมีค่าอำนาจจำแนกสูง มีค่าความยากครอบคลุมช่วงความสามารถของผู้สอบ และมีค่าการเดาต่ำ ข้อสอบควรมีอำนาจจำแนกสูง ($a > 0.80$) มีค่าความยากที่ครอบคลุมระดับช่วงต่างๆ อย่างเหมาะสม [$-2.0 < \theta < +2.0$] ค่าสัมประสิทธิ์การเดา ($c < 0.30$) และควรมีข้อสอบเท่าๆกันในแต่ละระดับความยาก (ศิริชัย กาญจนวาสี, 2555)

Urry (1977) กำหนดเกณฑ์ในการคัดเลือกข้อสอบ โดยข้อสอบแต่ละข้อควรมีค่าอำนาจจำแนกของข้อสอบ (a) มีค่าตั้งแต่ 0.50 ถึง 2.50 ค่าความยากของข้อสอบ (b) ที่ครอบคลุมระดับช่วงต่างๆ อย่างเหมาะสม มีค่าตั้งแต่ -2.50 ถึง 2.50 และค่าการเดาของข้อสอบ (c) ไม่ควรเกิน 0.30 และควรมีข้อสอบในจำนวนที่เท่า ๆ กันในแต่ละระดับความยากของข้อสอบ

Reckase (2003) ได้กำหนดเกณฑ์ในการคัดเลือกข้อสอบที่มีคุณภาพตามทฤษฎีตอบสนองข้อสอบ ในการจัดทำคลังข้อสอบได้แก่ ค่าความยากของข้อสอบอยู่ในช่วง -4 ถึง 4 ค่าอำนาจจำแนกของข้อสอบอยู่ระหว่าง 0 ถึง ∞ และค่าการเดาของข้อสอบไม่เกิน 0.50

Baker and Kim (2017) เสนอเกณฑ์อย่างง่ายสำหรับพิจารณาคุณภาพข้อสอบตามทฤษฎีการตอบสนองโดยพิจารณาจากโค้งคุณลักษณะข้อสอบ (item characteristic curve: ICC) เป็นหลัก กล่าวคือ ค่าความสามารถของผู้สอบ (θ) ที่ประมาณได้จากโมเดล 1pl, 2pl และ 3pl ส่วนใหญ่จะมีค่ากระจายอยู่ในช่วง -3 ลอจิต ถึง +3 ลอจิต ดังนั้น ICC จึงสามารถตรวจจับ (capture) หรืออธิบายความแปรปรวนของค่า θ ของผู้สอบทั้งหมดได้ดีในช่วงดังกล่าว ดังนั้น ข้อเสนอแนะเบื้องต้นของ Baker and Kim (2017) เกี่ยวกับค่าความยากข้อสอบที่เหมาะสมจึงอยู่ในช่วง -3 ลอจิต ถึง +3 ลอจิต

สำหรับค่าการจำแนกนั้น Baker and Kim (2017) เสนอว่า ข้อสอบที่มีค่าการจำแนกติดลบสะท้อนว่าผู้สอบที่มีความสามารถสูงมีโอกาสทำข้อสอบถูกน้อยกว่าผู้สอบที่มีความสามารถต่ำกว่า นั่นหมายความว่าสารสนเทศที่ได้จากข้อสอบดังกล่าว ไม่ได้ช่วยอธิบายความแปรปรวนของค่า θ ของผู้สอบ เป็นข้อสอบที่ให้สารสนเทศคลาดเคลื่อน ดังนั้น ข้อสอบที่มีค่าการจำแนกติดลบ จึงเป็นข้อสอบที่ไม่เหมาะสม ข้อสอบที่จำแนกได้เหมาะสม หมายถึงข้อสอบที่สามารถจำแนกผู้สอบที่มีความสามารถใกล้เคียงกันได้หรือไม่ ดังนั้นข้อเสนอแนะของ Baker and Kim (2017) จึงยอมรับในเบื้องต้นกับข้อสอบที่มีค่าการจำแนกมากกว่า 0 ขึ้นไป เพราะเมื่อค่าการ

I 28202843

จำแนกมากกว่า 0 นั้นหมายความว่าข้อสอบเริ่มจำแนกผู้สอบได้บ้างแล้ว แม้ว่าจะเป็นการจำแนกในระดับต่ำมากก็ตาม

Baker and Kim (2017) อธิบายว่า ค่าการเดามีค่าที่เป็นไปได้ระหว่าง 0 ถึง 1.0 แต่ในทางปฏิบัติค่าที่มากกว่า .35 ถือว่าไม่สามารถยอมรับได้ ดังนั้นค่าการเดาที่เหมาะสมของข้อสอบควรอยู่ระหว่าง 0 ถึง .35

Ali and Chang (2014) ตรวจสอบประสิทธิภาพของการทดสอบแบบปรับเหมาะ โดยใช้ดัชนีความเหมาะสม (suitability index) ในกระบวนการปรับเทียบสเกลข้อสอบในระยะเริ่มต้นก่อนเริ่มการทดสอบแบบปรับเหมาะ (pretest items calibration adaptive design) แบบ 3 พารามิเตอร์ ข้อสอบที่ Ali and Chang (2014) เลือกใช้ในการปรับเทียบก่อนการทดสอบแบบปรับเหมาะนั้น เกิดจากการจำลองค่าพารามิเตอร์ a, b และ c ขึ้นจากข้อสอบจริงจำนวน 521 ข้อ ซึ่งมีค่า a ระหว่าง .29 ถึง 2.9 ค่า b ระหว่าง -4.79 ถึง 3.62 และค่า c ระหว่าง 0 ถึง .50

ผลการจำลองค่าพารามิเตอร์ข้อสอบ พบว่า มีลักษณะเดียวกับค่าพารามิเตอร์ของข้อสอบจริง โดยมีค่า b อยู่ระหว่าง -3.0 ถึง +3.0 และมีค่า a อยู่ระหว่าง 0 ถึง 2.0 นั่นหมายความว่า ในกระบวนการปรับสเกลข้อสอบก่อนการทดสอบปรับเหมาะ ยอมรับข้อสอบที่มีค่าการจำแนกเท่ากับ 0 ด้วย

การวิเคราะห์ข้อสอบเพื่อจัดทำคลังข้อสอบในการวิจัยครั้งนี้ ผู้วิจัยพิจารณาคุณภาพข้อสอบจากค่าความคลาดเคลื่อนของการประมาณค่าพารามิเตอร์เป็นหลัก โดยข้อสอบที่มีค่าความเคลื่อนสูงผิดปกติ อันเนื่องมาจากพารามิเตอร์ข้อสอบข้อนั้นๆ ไม่สามารถระบุเป็นค่าที่แน่นอนได้ (not well-specified) และเลือกกำจัดข้อสอบที่มีค่าพารามิเตอร์ความยากสูงผิดปกติและต่ำผิดปกติด้วย โดยผู้วิจัยยอมรับข้อสอบที่มีค่าพารามิเตอร์ความยากต่ำกว่า -3.00 และ สูงกว่า +3.00 ลอจิท เป็นข้อสอบที่มีคุณภาพตรงตามวัตถุประสงค์ของการวิจัย ที่ต้องการใช้ข้อสอบง่ายเพื่อตรวจสอบความรู้พื้นฐานของนักศึกษา และใช้ข้อสอบยากเพื่อท้าทายความสามารถของนักศึกษาเช่นเดียวกับงานวิจัยของ Ali and Chang (2014) และ Mao and Xin (2013) ที่ยอมรับข้อสอบที่มีค่าพารามิเตอร์ความยากต่ำกว่า -3.00 และสูงกว่า +3.00 ลอจิท ในกระบวนการพัฒนาคลังข้อสอบแบบปรับเหมาะขั้นต้น

สำหรับการพิจารณาคุณภาพข้อสอบจากค่าพารามิเตอร์การจำแนก ผู้วิจัยเลือกใช้เกณฑ์ของ Baker and Kim (2017) โดยผู้วิจัยกำจัดข้อสอบที่มีค่าการจำแนกติดลบ และเลือกข้อสอบที่มีค่าการจำแนกตั้งแต่ 0 ขึ้นไป โดยผู้วิจัยยอมรับข้อสอบที่มีค่าการจำแนกใกล้เคียงกับ 0 ด้วย เช่นเดียวกับงานวิจัยของ Ali and Chang (2014) ที่ยอมรับข้อสอบที่มีค่าพารามิเตอร์การจำแนกใกล้เคียง 0 ในขั้นตอนของการพัฒนาคลังข้อสอบขั้นต้น

ตอนที่ 3 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (computerized adaptive testing: CAT)

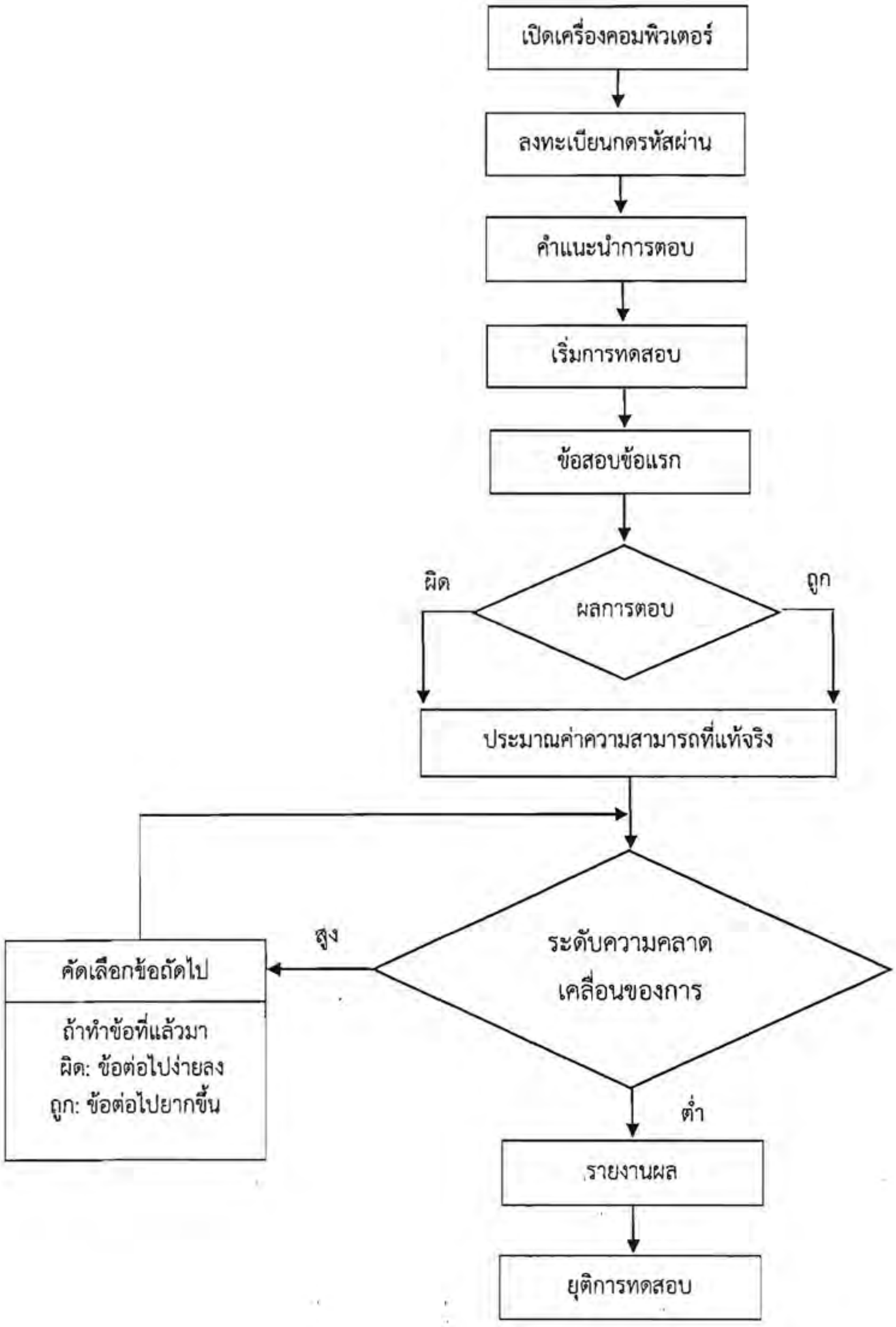
จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องสามารถแบ่งเนื้อหามโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะออกเป็น 4 ประเด็น ได้แก่ 3.1 ความหมายการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 3.2 หลักการของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 3.3 องค์ประกอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และ 3.4 การประยุกต์ใช้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยมีรายละเอียดดังนี้

3.1 ความหมายของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computerized adaptive testing: CAT) คือ การทดสอบที่ผู้สอบได้รับแบบทดสอบที่แตกต่างกัน โดยมีกระบวนการคัดเลือกข้อสอบตามระดับความสามารถของผู้สอบ ซึ่งนำกระบวนการทางคอมพิวเตอร์เข้ามาช่วยในการคัดเลือกข้อสอบและประมาณค่าความสามารถของผู้สอบ ซึ่งกระบวนการคำนวณทางคอมพิวเตอร์ทำให้การคัดเลือกข้อสอบและการประมาณค่าความสามารถมีความถูกต้องและแม่นยำมากยิ่งขึ้น สำหรับแนวคิดของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีการริเริ่มมาตั้งแต่ปี 1905 โดยนำมาประยุกต์ใช้กับแบบวัดเชาว์ปัญญาของ Binet-Simon ซึ่งข้อคำถามมีกระบวนการเลือกให้มีความเหมาะสมกับอายุสมองของผู้เข้ารับการทดสอบ โดยอายุสมองของผู้เข้ารับการทดสอบได้มาจากผลการตอบข้อคำถามจากข้อก่อนหน้า หลังจากนั้นดำเนินการเลือกข้อคำถามให้ผู้เข้ารับการทดสอบตอบกระทั่งสามารถประมาณค่าอายุที่แท้จริงของผู้เข้ารับการทดสอบได้ สำหรับแนวคิดการทดสอบแบบปรับเหมาะมีการนำมาประยุกต์ใช้เป็นเวลายาวนาน โดยการทดสอบที่เห็นได้ชัดเจนคือการทดสอบปากเปล่า (oral test) ซึ่งผู้จัดสอบจะดำเนินการเลือกคำถามที่มีความเหมาะสมกับความสามารถของผู้สอบตามระดับความรู้ความสามารถจากการสอบถาม (ศิริชัย กาญจนวาสี, 2555; van der Linden & Glas, 2000)

3.2 หลักการของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เป็นการคัดเลือกข้อสอบให้มีความเหมาะกับความสามารถของผู้สอบ โดยมีแนวทางในการคัดเลือกข้อสอบของผู้สอบแต่ละคนตามผลการตอบข้อสอบข้อที่ผ่านมา สำหรับการทดสอบผู้สอบจะเริ่มทำข้อสอบข้อแรกจากข้อที่มีค่าความยากปานกลาง หลังจากนั้นจะดำเนินการประมาณค่าความสามารถของผู้สอบจากผลการตอบข้อสอบ เมื่อทราบระดับความสามารถของผู้สอบแล้วจะดำเนินการคัดเลือกข้อสอบข้อถัดไปซึ่งมีระดับความยากและอำนาจจำแนกที่เหมาะสมกับระดับความสามารถของผู้สอบ หลักการคือหากผู้สอบตอบข้อสอบข้อที่ผ่านมาได้ถูกต้อง ข้อสอบข้อถัดไปจะมีความยากขึ้น แต่หากผู้สอบตอบข้อสอบผิด ข้อสอบข้อถัดไปจะมีความยากลดลง การทดสอบจะดำเนินการไปจนกระทั่งการประมาณค่าความสามารถผู้สอบอยู่ในระดับที่มีความน่าเชื่อถือ กล่าวคืออยู่ในเกณฑ์ของความคลาดเคลื่อนที่ยอมรับได้ตามที่มีการกำหนดไว้ จึงดำเนินการยุติการทดสอบ สำหรับขั้นตอนการดำเนินการทดสอบแสดงดังภาพ 2.6 (ศิริชัย กาญจนวาสี, 2555; Čisar, Radosav, Markoski, Pinter, & Čisar, 2010)



ภาพ 2.6 กระบวนการทำงานของการทดสอบแบบปรับเหมาะ

3.3 องค์ประกอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

สำหรับการทดสอบแบบปรับเหมาะมีองค์ประกอบที่สำคัญ 4 องค์ประกอบ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ (initial item) 2) การประมาณค่าความสามารถ (ability estimation) 3) การคัดเลือกข้อสอบ (item selection) และ 4) จุดยุติในการทดสอบ (stopping criteria) โดยแต่ละองค์ประกอบมีรายละเอียดดังนี้

1) จุดเริ่มต้นในการทดสอบ (initial item)

จุดเริ่มต้นในการทดสอบสำหรับการทดสอบแบบปรับเหมาะเป็นสิ่งที่สำคัญอย่างยิ่ง เนื่องจากหากกระบวนการคัดเลือกข้อสอบข้อแรกสำหรับเริ่มต้นการทดสอบไม่เหมาะสมกับระดับความสามารถผู้สอบ อาจส่งผลให้ความสามารถที่ได้จากการประมาณค่ามีความคลาดเคลื่อนสูง ทำให้กระบวนการคัดเลือกข้อสอบข้อถัดไปมีความคลาดเคลื่อนตามมา ดังนั้นจึงได้มีการแบ่งแนวทางสำหรับกระบวนการกำหนดข้อสอบข้อแรกไว้ 2 ประเภท ดังนี้ (ศิริชัย กาญจนวาสี, 2555)

- เมื่อประชากรของผู้สอบมีความสามารถที่ใกล้เคียงกัน (homogenous) หรือไม่ทราบคะแนนผลสัมฤทธิ์ที่ผ่านมาของประชากรผู้สอบ ดังนั้นควรดำเนินการเลือกข้อสอบข้อแรกที่มีระดับความยากปานกลาง เนื่องจากไม่ทราบข้อมูลก่อนหน้า หากดำเนินการนำข้อสอบที่มีความยากหรือง่ายเกินระดับความสามารถของผู้สอบ อาจส่งผลต่อการประมาณค่าความสามารถ

- เมื่อประชากรของผู้สอบมีความสามารถที่แตกต่างกัน (heterogeneous) กล่าวคือ พอมีสารสนเทศซึ่งเกี่ยวข้องกับความสามารถของผู้ทดสอบอยู่บ้าง เช่น คะแนนผลสัมฤทธิ์ ระดับการศึกษา ดังนั้นควรดำเนินการเลือกข้อสอบข้อแรกที่มีความเหมาะสมสารสนเทศที่มีอยู่ เช่น ผู้สอบอยู่ในระดับการศึกษาที่แตกต่างกัน ข้อสอบข้อแรกควรมีระดับความยากที่แตกต่างกัน

สำหรับเกณฑ์ในการคัดเลือกข้อสอบข้อแรกไม่ค่อยมีความสำคัญมากนักในการทดสอบที่มีความยาวของแบบสอบเกิน 25 ข้อ ซึ่งจากการศึกษาด้วยการจำลองข้อมูลโดยใช้ข้อสอบข้อแรกที่มีระดับความยากแตกต่างกัน พบว่า เมื่อดำเนินการใช้ข้อสอบเกิน 25 ข้อ ความแม่นยำในการประมาณค่าความสามารถของผู้สอบไม่เปลี่ยนแปลง ดังนั้นจึงสามารถสรุปได้ว่าระดับความยากของข้อสอบข้อแรกไม่ส่งผลต่อการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในกรณีที่แบบสอบมีความยาวตั้งแต่ 25 ข้อขึ้นไป (ศิริชัย กาญจนวาสี, 2555)

2) การประมาณค่าความสามารถ (ability estimation)

การประมาณค่าความสามารถของผู้สอบตามทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) ซึ่งเป็นที่ยอมรับและมีการนำมาใช้กันอย่างกว้างขวางมี 2 วิธี ได้แก่ 1) วิธีการประมาณค่าด้วยความเป็นไปได้สูงที่สุด (maximum likelihood estimation) และ 2) วิธีการประมาณค่าแบบ เบเซียน (bayesian estimation) (ศิริชัย กาญจนวาสี, 2555)

- วิธีการประมาณค่าด้วยความเป็นไปได้สูงที่สุด (maximum likelihood estimation)

วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด เป็นวิธีการประมาณค่าความสามารถที่ค่อนข้างมีความแม่นยำสูงเมื่อเปรียบเทียบกับวิธีการประมาณค่าด้วยวิธีอื่น แต่การประมาณค่าจะมีความความคลาดเคลื่อนสูงเมื่อผู้เข้ารับการทดสอบมีผลการตอบข้อสอบถูกหรือผิดทั้งหมด โดยมีสูตรในการคำนวณดังนี้ (Hambleton & Swaminathan, 1985)

$$\mathrm{P}(U_{ij} = 1|\theta_i, a_j, b_j, c_j) = c_j + (1 - c_j)\frac{exp\left(a_j(\theta_i - b_j)\right)}{1 + exp\left(a_j(\theta_i - b_j)\right)}$$

$$L_J(\theta) = \prod_{j=1}^{J} \left[P_j(\theta)\right]^{x_j} \left[1 - P_j(\theta)\right]^{1-x_j}$$

$$h_m = \frac{D[r_m - \sum P_i(\theta_m)]}{-D^2 \sum P_i(\theta_m)\, Q_i(\theta_m)}$$

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

- วิธีการประมาณค่าแบบเบเซียน (bayesian estimation)

วิธีการประมาณค่าแบบเบเซียนเป็นอีกวิธีการหนึ่งซึ่งนำมาใช้ร่วมกับวิธีการประมาณค่าด้วยความเป็นไปได้สูงที่สุดในกรณีที่ผู้สอบมีผลการตอบข้อสอบถูกหรือผิดทั้งหมด โดยวิธีซึ่งเป็นที่นิยมคือวิธีการประมาณค่าความสามารถด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) เนื่องจากวิธีการประมาณค่าด้วยความเป็นไปได้สูงที่สุดจะมีความความคลาดเคลื่อนสูงเมื่อผู้เข้ารับการทดสอบมีผลการตอบข้อสอบถูกหรือผิดทั้งหมด ดังนั้นจึงควรใช้วิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง โดยมีสูตรในการคำนวณดังนี้ (Bock & Mislevy, 1982)

$$\bar{\theta_J} = \frac{\sum_{k=1}^{q} X_k L_J(X_k) \cdot W(X_k)}{\sum_{k=1}^{q} L_J(X_k) \cdot W(X_k)}$$

3) การคัดเลือกข้อสอบ (item selection)

การคัดเลือกข้อสอบซึ่งเป็นที่นิยมใช้กันอย่างแพร่หลายมีด้วยกัน 2 วิธี ได้แก่ 1) การคัดเลือกข้อสอบที่มีค่าความยากสอดคล้องกับระดับความสามารถ และ 2) การคัดเลือกข้อสอบที่มีตำแหน่งสารสนเทศสูงสุดที่สอดคล้องกับระดับความสามารถ โดยมีรายละเอียดดังนี้ (ศิริชัย กาญจนวาสี, 2555)

- การคัดเลือกข้อสอบที่มีค่าความยากสอดคล้องกับระดับความสามารถ

สำหรับการคัดเลือกข้อสอบที่มีค่าความยาก (item difficulty) ที่สอดคล้องกับระดับความสามารถของผู้สอบ เริ่มดำเนินการโดยตรวจสอบข้อสอบทุกข้อในคลังที่มีค่าระดับความยากสอดคล้องกับระดับความสามารถของผู้สอบ กล่าวคือพิจารณาจากข้อสอบที่มีค่าความยากใกล้เคียงกับระดับความสามารถของผู้สอบ เมื่อพบข้อสอบแล้วจะดำเนินการนำข้อสอบที่ได้รับการคัดเลือกไปให้ผู้เข้าสอบดำเนินการสอบต่อไป โดยวิธีการดังกล่าวเป็นวิธีการที่ค่อนข้างง่ายไม่ซับซ้อนและเป็นวิธีการที่ประหยัดค่าใช้จ่ายมากที่สุด

- การคัดเลือกข้อสอบที่มีตำแหน่งสารสนเทศสูงสุดที่สอดคล้องกับระดับความสามารถ

สำหรับการคัดเลือกข้อสอบที่มีตำแหน่งสารสนเทศ (item information) สูงสุดที่สอดคล้องกับระดับความสามารถของผู้สอบ เริ่มมีแนวคิดมาจากการทดสอบโดยทั่วไปที่อาจมีการเดา หากดำเนินการพิจารณาเฉพาะค่าความยากอาจทำให้ผู้เข้ารับการสอบได้รับข้อสอบที่ยากเกินไป ดังนั้นจึงได้มีการพัฒนาวิธีการคัดเลือกข้อสอบจากตำแหน่งสารสนเทศสูงสุดที่มีความสอดคล้องกับระดับความสามารถของผู้สอบ โดยข้อสอบแต่ละข้อจะ

ให้สารสนเทศสูงสุด ณ ตำแหน่ง $\theta = m_i$ เมื่อ m_i คือค่าสำหรับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โดยมีสูตรในการคำนวณดังนี้

$$m_i = b_i + \frac{1}{Da_1} ln \left[\frac{1 + \sqrt{1 + 8c_i}}{2}\right]$$

เมื่อ D = ค่าคงที่

Ln = ค่าลอกฐานธรรมชาติ (natural logarithm)

a_i = ค่าอำนาจจำแนกของข้อสอบ

b_i = ค่าความยากของข้อสอบ

c_i = ค่าโอกาสในการเดาข้อสอบถูก

เมื่อ m_i คือ ค่าความสามารถของผู้สอบ (θ) ณ ตำแหน่งที่ข้อสอบข้อนั้นให้สารสนเทศสูงสุด หาก ค่า $m_i > b_i$ และ $c_i > 0$ หมายถึง ข้อสอบมีค่าความยากที่เหมาะสมสำหรับประมาณค่าความสามารถของผู้สอบที่มีความสามารถสูงกว่าความยากของข้อสอบเล็กน้อย ดังนั้น การนำค่า m_i มาใช้ประกอบการพิจารณาคัดเลือกข้อสอบจะช่วยทำให้การคัดเลือกข้อสอบมีประสิทธิภาพมากยิ่งขึ้น

4) เกณฑ์ยุติในการทดสอบ (stopping criteria)

ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ผู้สอบจะทำการทดสอบจนกระทั่งเป็นไปตามเกณฑ์ยุติในการทดสอบจึงหยุดการทดสอบและดำเนินการประมาณค่าความสามารถของผู้สอบในลำดับต่อไป สำหรับเกณฑ์ยุติในการทดสอบที่นิยมใช้กันอย่างแพร่หลายมี 2 วิธี คือ 1) เกณฑ์กำหนดจำนวนข้อสอบคงที่ และ 2) เกณฑ์กำหนดระดับความคลาดเคลื่อนที่ยอมรับได้

• เกณฑ์กำหนดจำนวนข้อสอบคงที่ (fixed length)

ผู้สอบทุกคนจะได้รับข้อสอบที่มีความยาวของแบบสอบเท่ากันทุกคน เช่น ในการทดสอบกำหนดเกณฑ์แบบจำนวนข้อสอบคงที่โดยให้ความยาวแบบสอบเท่ากับ 35 ข้อ ดังนั้นผู้เข้ารับการทดสอบทุกคนจะต้องทำข้อสอบจำนวน 35 ข้อ เมื่อทำครบข้อที่ 35 การทดสอบจะยุติลง เกณฑ์ดังกล่าวมีความเหมาะสมกับการศึกษาความถูกต้องของกระบวนการทดสอบแบบปรับเหมาะโดยใช้การจำลองข้อมูล เนื่องจากเมื่อจำนวนข้อสอบเท่ากันทำให้สามารถเปรียบเทียบสารสนเทศของแบบสอบได้อย่างชัดเจนมากยิ่งขึ้น

• เกณฑ์กำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ (variable length)

แนวคิดของการทดสอบแบบปรับเหมาะ คือ ผู้สอบจะดำเนินการทดสอบจนกระทั่งผู้สอบได้รับการประมาณค่าความสามารถที่มีความคงเส้นคงวา โดยมีค่าความคลาดเคลื่อนมาตรฐาน (standard error) ลดต่ำลงจนถึงระดับที่สามารถยอมรับได้ การทดสอบจึงยุติลง ในการคำนวณค่าความคลาดเคลื่อนมาตรฐานสามารถหาได้จากสูตร

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

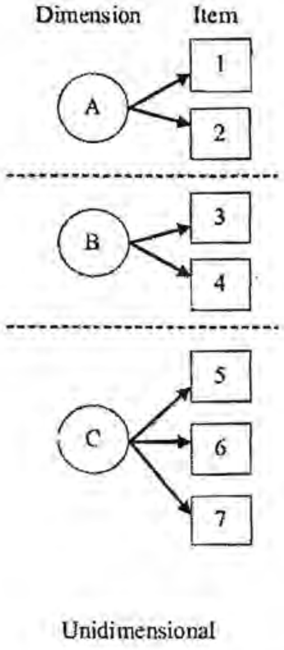
เมื่อ $SE(\theta)$ = ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ θ

$I(\theta)$ = สารสนเทศของแบบสอบของผู้มีความสามารถ θ

เพื่อให้ได้รายละเอียดที่ชัดเจนมากยิ่งขึ้น ผู้วิจัยได้ทบทวนประเภทของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 2 ประเภท ได้แก่ การทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติ และ การทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ พร้อมทั้งงานวิจัยที่เกี่ยวข้อง ดังมีรายละเอียด ดังนี้

3.3.1 การทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติ (Unidimensional computerized adaptive testing)

การทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติเป็นการรวมแนวคิดของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และทฤษฎีการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional Item response theory) ด้วยกัน โดยการทดสอบปรับเหมาะด้วยคอมพิวเตอร์เป็นการทดสอบที่มีประสิทธิภาพในการวัดสูง มีความแม่นยำในการประมาณค่าความสามารถของผู้สอบ และสามารถลดเวลาและจำนวนข้อสอบที่ผู้สอบต้องทำการทดสอบได้ ซึ่งในการประมาณค่าความสามารถของการทดสอบด้วยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) นั้นจะใช้ทฤษฎีการตอบสนองข้อสอบ (Item response theory) เป็นพื้นฐานในการคำนวณคะแนนความสามารถของผู้สอบ ภายใต้ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional) (Segall, 1996 อ้างถึงในทัศน์ศิรินทร์ สว่างบุญ ,2554) ซึ่งเป็นโมเดลการตอบสนองข้อสอบที่ตรวจให้คะแนน 2 ค่าแบบเอกมิติ (Unidimensional IRT Models) โดยตัวแปรคุณลักษณะแฝงหรือระดับความสามารถของผู้สอบมีเพียงองค์ประกอบเดียว หรือมิติเดียว หรือข้อสอบนั้นมุ่งที่จะวัดคุณลักษณะแฝงเพียงคุณลักษณะเดียว ดังภาพ 2.7



ภาพ 2.7 โมเดลการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional IRT Model)
(Cheng, Patton & Shao, 2014)

องค์ประกอบของการทดสอบปรับเหมาะแบบเอกมิติ

1) จุดเริ่มต้นการทดสอบ (starting point)

การทดสอบเริ่มต้นด้วยการคัดเลือกข้อสอบข้อแรก (initial item) การคัดเลือกข้อสอบข้อแรกนิยมใช้ข้อสอบที่มีความยากปานกลางซึ่ง (ศิริชัย กาญจนวาสี, 2555) สามารถแบ่งวิธีการคัดเลือกเป็น 2 กรณี

1.1) เมื่อประชากรผู้สอบมีความสามารถค่อนข้างใกล้เคียงกัน (homogeneous) หรือประชากรผู้สอบที่ไม่มีข้อมูลเกี่ยวกับผลสัมฤทธิ์ที่ผ่านมาควรเริ่มต้นด้วยข้อสอบที่มีค่าความยากปานกลางเท่ากันสำหรับทุกคน

1.2) เมื่อประชากรผู้สอบมีความสามารถค่อนข้างแตกต่างกัน (heterogeneous) หรือเป็นประชากรผู้สอบที่พอมีสารสนเทศเกี่ยวกับระดับการศึกษาของผู้สอบ ควรเริ่มต้นด้วยข้อสอบที่มีค่าความยากปานกลางสำหรับแต่ละระดับการศึกษาของผู้สอบ

2) การประมาณค่าความสามารถในการทดสอบปรับเหมาะแบบเอกมิติ

การประมาณค่าความสามารถในการทดสอบแบบเอกมิติจะใช้การประมาณค่าด้วยวิธี Bayesian และวิธี Maximum likelihood (ML) (Weiss, 1974 อ้างถึงในทัศน์ศิรินทร์ สว่างบุญ, 2554) มีรายละเอียดดังนี้

- การให้คะแนนตามวิธี Bayesian วิธีที่นิยมใช้กันมาก คือ วิธีของเบส์ที่ปรับปรุงให้ทันสมัย (Bayesian updating) ซึ่งสามารถประมาณค่าความสามารถของผู้สอบได้ค่อนข้างคงที่ ซึ่งมีสูตรดังนี้

<u>กรณีที่ตอบถูก</u>

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

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

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

และ B = c + (1+c) x A(-D)

<u>กรณีที่ตอบผิด</u>

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

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

θ_m	แทน	ความสามารถของผู้สอบที่ประมาณค่าได้ก่อนตอบข้อสอบข้อที่ m+1 ซึ่งตามปกติแล้ว ถ้าไม่ทราบค่าความสามารถเบื้องต้นของผู้สูอบมักกำหนดให้เท่ากับ 0
σ_m^2	แทน	ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อที่ m+1 ถ้าไม่ทราบค่าความแปรปรวนดังกล่าวมาก่อนมักกำหนดให้เท่ากับ 1
θ_{m+1}	แทน	ค่าความสามารถของผู้สอบโดยประมาณหลัง จากที่ตอบข้อที่ m+1
σ_{m+1}^2	แทน	ค่าความแปรปรวนในการประมาณค่าความสามารถของผู้สอบเมื่อ ตอบข้อสอบข้อที่ m+1
a	แทน	พารามิเตอร์ค่าอำนาจจำแนกข้อสอบข้อที่ m+1
b	แทน	พารามิเตอร์ค่า ความยากของข้อสอบข้อที่ m+1
c	แทน	พารามิเตอร์ระดับโอกาสการเดาข้อสอบได้ถูก ข้อที่ m+1
D	แทน	จุดบนแกน X
O(D)	แทน	ค่าออร์ดิเนท (ordinate) ของโค้งปกติที่จุด D
A (D)	แทน	พื้นที่ใต้โค้งปกติจากค่าD จนถึงจุด D

- การให้คะแนนโดยวิธีการประมาณค่าความเป็นไปได้สูงสุด (maximum likelihood) วิธีที่นิยมใช้มาก คือ วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (conditional maximum likelihood) การประมาณค่า โดยวิธีนี้มีข้อจำกัด คือ ถ้าผู้สอบตอบข้อสอบถูกหมดหรือผิดหมด จะไม่สามารถประมาณค่าได้ ขั้นตอนในการให้คะแนนมีดังนี้ (Hambleton & Swanminathan, 1985 อ้างถึงใน ทัศน์ศิรินทร์ สว่างบุญ, 2554; ศิริชัย กาญจนวาสี, 2555)

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

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

เมื่อ	r_a	=	$\sum a_i U_i$
	U_i	=	1 เมื่อตอบข้อสอบถูก
	U_i	=	0 เมื่อตอบข้อสอบผิด
	a_i	คือ	อำนาจจำแนกของข้อสอบข้อที่ I
	K	คือ	จำนวนข้อสอบทั้งหมดที่ผู้สอบตอบ

ขั้นที่ 2 หาค่า $P_1(\theta_m)$ และ $Q(\theta_m)$โดยใช้สูตร ดังนี้

$$P_1(\theta_m) = c_i + (1-c_i)\frac{e^{Da_1(\theta_m-b_1)}}{Da_i(\theta_m-b_i)}$$

$$Q(\theta_m) = 1 - P_1(\theta_m)$$

$$= \frac{1-c_i}{1+e^{Da_1(\theta_m-b_1)}}$$

เมื่อ θ_m คือ ความสามารถของผู้สอบที่ประมาณค่าได้ในครั้งที่ m

a_1 คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_1 คือ ค่าความยากของข้อสอบข้อที่ i

c_1 คือ ค่าการเดาของข้อสอบข้อที่ t

D คือ ค่าคงที่ ซึ่งมีค่าเท่ากับ 1.7

e คือ ค่าคงที่ ซึ่งมีค่าเท่ากับ 2.7182

ขั้นที่ 3 หาค่าปรับแก้ (h_m) โดยใช้สูตร

$$h_m = \frac{D[r_m - \sum Pi(\theta_m)]}{-D^2 \sum Pi(\theta_m)\, Qi(\theta_m)}$$

ขั้นที่ 4 ประมาณค่าความสามารถของผู้สอบใหม่ $(\theta_m + 1)$ โดยใช้สูตร

$$\theta_m + 1 = \theta_m - h_m$$

ขั้นที่ 5 ทำการคำนวณซ้ำในขั้นที่ 2,3 และ 4 จนกระทั่ง h_m เข้าใกล้ศูนย์

$$(h_m < 0.001)$$

3) วิธีการคัดเลือกข้อสอบในการทดสอบปรับเหมาะแบบเอกมิติ

แบบทดสอบแบบปรับเหมาะที่ให้สารสนเทศเกี่ยวกับความสามารถของผู้สอบมากที่สุดโดยใช้ข้อสอบจำนวนน้อย เป็นแบบทดสอบปรับเหมาะที่มีประสิทธิภาพ เนื่องจากการเลือกใช้ข้อสอบแต่ละข้อมีความสำคัญต่อการทดสอบ ข้อสอบจะให้สารสนเทศสูงสุดเมื่อมีค่าความยาก (b) ใกล้เคียงค่าความสามารถของผู้สอบมากที่สุด ค่าอำนาจจำแนก (a) สูง ค่าการเดา (c) เข้าใกล้ศูนย์ และเลือกข้อสอบคุณสมบัติดังกล่าวจัดให้แก่ผู้สอบ ซึ่งขณะเดียวกันการเลือกข้อสอบต้องคำนึงถึงการป้องกันการใช้ข้อสอบข้อใดข้อหนึ่งมากเกินไปและเพิ่มการใช้ข้อสอบที่ใช้น้อยหรือไม่ได้ใช้เลย รักษาสมดุลเนื้อหา ร่วมกับการให้ประสิทธิภาพการประมาณค่าความสามารถของผู้สอบไปพร้อมๆ กัน ดังนั้นจุดมุ่งหมายการเลือกข้อสอบว่า เพื่อให้ความแม่นยำในการประมาณค่า

ความสามารถของผู้สอบมากที่สุด การรักษาความปลอดภัยของข้อสอบ และรักษาสมดุลเนื้อหาแบบทดสอบพร้อมกันด้วย (Davis & Dodd, 2003; Parshall et al. 1998 อ้างถึงใน สุดารัตน์ หวลมุกดา, 2550; ศิริชัย กาญจนวาสี, 2555)

วิธีการเลือกข้อสอบแต่ละวิธี มีดังนี้

-การเลือกข้อสอบโดยใช้วิธีการสุ่มเลือกข้อสอบ (Random Item Selection) เนื่องจากวิธีการสุ่มเลือกเปิดโอกาสให้ข้อสอบทุกข้อมีโอกาสถูกเลือกเท่าเทียมกัน วิธีนี้จึงช่วยกระจายการใช้ข้อสอบในคลังข้อสอบ แต่ให้ผลลดประสิทธิภาพการประมาณค่าความสามารถของผู้สอบและไม่ให้สมดุลเนื้อหา (Eggen, 2001)

- การเลือกข้อสอบโดยพิจารณาจากฟังก์ชันสารสนเทศ (Information Function) วิธีนี้สามารถให้ประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูง แต่ข้อสอบที่ให้สารสนเทศสูงมักนำมาใช้บ่อยเกินไป ในขณะที่ข้อสอบบางข้อเลือกใช้น้อยหรือไม่ได้เลือกเลย ข้อสอบใช้มากเกินไปทำให้ผู้สอบเกิดการเรียนรู้หรือข้อสอบเป็นที่รู้จักกันทั่วไปในกลุ่มผู้สอบ เสี่ยงต่อความปลอดภัยของข้อสอบและสูญเสียคุณสมบัติทางจิตวิทยาการทดสอบจากความเที่ยงตรงของข้อสอบลดลง วิธีการเลือกข้อสอบที่นิยม ดังนี้ คือ

การเลือกข้อสอบวิธีสารสนเทศสูงสุด (Maximum Information Item Selection) วิธีสารสนเทศสูงสุดจะเลือกข้อสอบที่ให้สารสนเทศสูงสุด ตามค่าประมาณความสามารถของผู้สอบครั้งหลังสุด เมื่อสารสนเทศแบบทดสอบสูงขึ้น ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถจะลดลง ทำให้ค่าประมาณความสามารถเข้าใกล้ค่าความสามารถจริงของผู้สอบมากที่สุด

การเลือกข้อสอบวิธีเบเซี่ยน (Bayesian Item Selection) วิธีนี้เลือกข้อสอบที่ให้ความแปรปรวนของการประมาณค่าความสามารถของผู้สอบน้อยที่สุด การเลือกข้อสอบทั้งวิธีสารสนเทศสูงสุดและวิธีการเลือกข้อสอบวิธีเบเซี่ยน สามารถลดความคลาดเคลื่อนกำลังสองเฉลี่ยได้ ทั้ง 2 วิธีให้ผลการเลือกข้อสอบใกล้เคียงกันเมื่อความยาวแบบทดสอบมากเพียงพอ

4) การควบคุมการใช้ข้อสอบซ้ำในการทดสอบปรับเหมาะแบบเอกมิติ

การควบคุมการใช้ข้อสอบซ้ำ ของการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ คือการใช้ข้อสอบซ้ำถี่เกินไป เช่น กลุ่มของข้อสอบ ที่มีค่าพารามิเตอร์อำนาจจำแนกสูงๆ เมื่อนำออกมาใช้บ่อยจนรู้จักแพร่หลายในกลุ่มผู้สอบ จะนำไปสู่ปัญหาในเรื่องความปลอดภัยของแบบสอบ ซึ่งเกิดจากการร่วมใช้ข้อสอบ ระหว่างผู้สอบ ผู้สอบจะได้รับข้อมูลของแบบสอบจากผู้ที่เคยสอบก่อนหน้าและผู้สามารถ สอบตอบคำถามข้อนั้นถูกอย่างง่ายดาย โดยการล่วงรู้ข้อสอบก่อน โดยไม่ต้องใช้ความสามารถของตนเอง เพื่อให้เกิดความเหมาะสมในทางปฏิบัติ การทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ควรให้ความสำคัญในเรื่องการควบคุมการใช้ข้อสอบซ้ำ (Item Exposure Control) และ การสร้างความสมดุลของเนื้อหา (Content Balancing) (Chen & Doong, 2008) โดยสามารถแบ่งวิธีการควบคุมการใช้ข้อสอบซ้ำ ดังนี้

4.1). กลุ่มที่ใช้กระบวนการสุ่มเข้ามาช่วยควบคุมการใช้ข้อสอบซ้ำในขั้นตอนวิธีการคัดเลือกข้อสอบ(Randomization) โดยการเลือกกลุ่มข้อสอบที่เหมาะสมที่สุดจากชุดข้อสอบ แทนการเลือกข้อสอบสารสนเทศสูงสุดเพียงข้อเดียว และสุ่มเลือกข้อสอบข้อหนึ่งจากกลุ่มข้อสอบจัดให้กับผู้สอบ วิธีการเลือกอย่างสุ่มไม่รับรองว่าการใช้ข้อสอบจะควบคุมได้ตามเกณฑ์ที่ต้องการ ตัวอย่างวิธีนี้ เช่น วิธี 5-4-3-2-1 (5-4-3-2-1 Procedure) วิธีสุ่มเลือกจากกลุ่มข้อสอบ (Randomesque Procedure) วิธีสุ่มเลือกแบบก้าวหน้า (Progressive Procedure)

4.2) กลุ่มที่ใช้ขั้นตอนวิธีการคัดเลือกข้อสอบอย่างมีเงื่อนไข เป็นวิธีควบคุมความน่าจะเป็นของข้อสอบที่ได้รับการเลือกแล้วให้มีการใช้ข้อสอบเป็นไปตามกำหนด โดยอัตราการใช้ข้อสอบสูงสุดเป็นไปตามกำหนด แต่ขั้นตอนมีความซับซ้อน และต้องศึกษาในสถานการณ์จำลองในด้านจำนวนครั้งสำหรับการปฏิบัติ เมื่อองค์ประกอบต่างๆ ในการทดสอบเปลี่ยนแปลงไป จำเป็นต้องศึกษาในสถานการณ์จำลองก่อนใช้ในสถานการณ์จริง มีวิธีต่างๆ ตัวอย่างวิธีนี้ เช่น วิธี Sympson-Hetter (SH) วิธี Sympson-Hetter แบบมีเงื่อนไข (Conditional Sympson-Hetter)

4.3) กลุ่มที่ใช้วิธีในการแบ่งคลังข้อสอบออกเป็นชั้นๆ โดยใช้วิธีการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ตามระดับชั้นของค่าอำนาจจำแนก (a-Stratified Method: a-STR) เป็นวิธีการเลือกข้อสอบและมีประสิทธิภาพในการควบคุมการใช้ข้อสอบซ้ำโดยอัตโนมัติ เริ่มจากการแบ่งขั้นคลังข้อสอบตามค่าอำนาจจำแนกของข้อสอบ และแบ่งแบบทดสอบเข้าในแต่ละขั้นคลังข้อสอบ ขั้นแรก บรรจุข้อสอบอำนาจจำแนกต่ำสุดขั้นต่อไปบรรจุข้อสอบอำนาจจำแนกสูงกว่าขั้นแรก ทำเช่นนี้เรื่อยๆ จนกระทั่งขั้นสุดท้ายบรรจุข้อสอบค่าอำนาจจำแนกสูงสุด การทดสอบจะดำเนินจากขั้นแรกจนถึงขั้นสุดท้าย วิธีนี้กระจายการใช้ข้อสอบซ้ำในคลังข้อสอบให้เท่าเทียม

5) เกณฑ์การยุติการทดสอบ

ในการทดสอบปรับเหมาะแบบเอกมิติจะสิ้นสุดลงเมื่อผลการตอบข้อสอบที่ผ่านมาเป็นไปตามเกณฑ์ยุติการทดสอบ เกณฑ์การยุติการทดสอบอาจเป็นจำนวนข้อสอบที่กำหนดไว้ล่วงหน้า เวลาที่ใช้ในการทดสอบคะแนนจุดตัดที่แยกระดับความสามารถของกลุ่มผู้สอบ ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ เกณฑ์การยุติการทดสอบมี 2 ลักษณะ คือ กำหนดจำนวนข้อสอบให้คงที่ และ กำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ ดังนี้ (ศิริชัย กาญจนวาสี, 2555)

- กำหนดจำนวนข้อสอบให้คงที่ เป็นการกำหนดจำนวนข้อสอบที่ใช้ในการทดสอบให้คงที่สำหรับทุกคน เช่น กำหนดให้การทดสอบปรับเหมาะใช้ข้อสอบจำนวน 25 ข้อ (n = 25) เมื่อผู้สอบทำข้อสอบได้ครบ 25 ข้อ การทดสอบก็จะยุติลง เกณฑ์นี้ค่อนข้างเป็นประโยชน์ในการศึกษาภายใต้สถานการณ์จำลองแบบ MonteCarlo Simulation เนื่องจากจำนวนข้อสอบ เท่ากันทำให้สามารถเปรียบเทียบสารสนเทศของแบบสอบได้โดยตรง ในทางปฏิบัติการกำหนดให้ทุกคนทำข้อสอบจำนวนเท่ากัน อาจมีคุณภาพของการวัดผลได้แตกต่างกัน ดังนั้นเกณฑ์ในอีกลักษณะหนึ่งจึงพิจารณาถึงความคลาดเคลื่อนมาตรฐานของการวัดผลเป็นสำคัญ

- กำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ ในทางปฏิบัติ การทดสอบปรับเหมาะด้วยคอมพิวเตอร์ ควรให้ดำเนินไปเรื่อยจนกว่าการประมาณค่า (θ) มีความคลาดเคลื่อนมาตรฐาน (Standard Error: SE (θ)) ลดต่ำลงจนถึงระดับที่ยอมรับได้การทดสอบจึงยุติลง ณ เวลานั้น

$$SEE(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

เมื่อ $SE(\theta)$ คือ ความคลาดเคลื่อนมาตรฐานของการประมาณค่า θ

$I(\theta)$ คือ สารสนเทศของแบบสอบที่ให้สำหรับผู้มีความสามารถ θ

3.3.2 การทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional computerized adaptive testing)

แนวคิดทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (MIRT)

แนวคิดทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (*MIRT)* มาจากแนวคิดพื้นฐานสำคัญซึ่งจำแนกออกเป็น *2* กลุ่มแนวคิด โดยกลุ่มแรกเป็นการศึกษาของ Spearman (1927) และ Thurstone (1947) ได้ศึกษาและพบว่าโมเดลการตอบสนองข้อสอบแบบพหุมิติ มีที่มาจากการวิเคราะห์องค์ประกอบ (*Factor Analysis*) เนื่องจากลักษณะของการวิเคราะห์องค์ประกอบจะมีความสัมพันธ์โดยตรงกับการวิเคราะห์ *MIRT* ส่วนอีกกลุ่มเป็นการศึกษาของ Lazarsfeld (1950) ที่กล่าวว่าการตอบสนองข้อสอบแบบพหุมิติ ได้รับอิทธิพลมาจาก *UIRT* เช่นเดียวกับการวิเคราะห์องค์ประกอบ แต่แนวคิดเกี่ยวกับ *UIRT* ในช่วงแรกยังไม่เป็นที่นิยมเหมือนเช่นปัจจุบัน จากแนวคิดทั้งสองอาจกล่าวได้ว่าการตอบสนองข้อสอบแบบพหุมิติ เป็นแนวคิดที่แผ่ขยายมาจาก *UIRT* หรือเป็นกรณีเฉพาะของการวิเคราะห์องค์ประกอบหรือโมเดลสมการเชิงโครงสร้าง (*structural equation modeling*) (ศิริชัย กาญจนวาสี, 2555)

โมเดลการตอบสนองข้อสอบแบบพหุมิติ

โมเดลการตอบสนองข้อสอบแบบพหุมิติ ถือว่าคุณลักษณะแฝงของบุคคลมากกว่า 1 องค์ประกอบ ส่งผลต่อการตอบข้อสอบ ดังนั้นพารามิเตอร์ความสามารถของผู้สอบจึงมีตั้งแต่ 2 พารามิเตอร์ขึ้นไป ดังนั้นการพิจารณาความสามารถหลายมิติของบุคคลน่าจะช่วยทำให้โมเดลมีความสอดคล้องกับข้อมูลได้ดียิ่งขึ้น เมื่อผู้สอบใช้ความสามารถหลายมิติที่แตกต่างกันอย่างเป็นระบบในการตอบข้อสอบย่อมมีความสัมพันธ์กับระดับความยากและอำนาจจำแนกของข้อสอบรายข้อ และมีผลต่อความน่าจะเป็นในการตอบข้อสอบได้ถูกต้อง โดยมีสัญลักษณ์ของค่าต่าง ๆ ในโมเดล (ทัศน์ศิรินทร์ สว่างบุญ, 2554) ดังนี้

X_{ip} = ผลการตอบข้อสอบข้อที่ i ของบุคคล p (0=ผิด ,1=ถูก)

θ_{pm} = ค่าระดับความสามารถหรือคุณลักษณะแฝงของผู้สอบ p (m≥2)

b_i = ค่าความยากของข้อสอบ ข้อที่ i

δ_i = ค่าจุดตัดความง่าย (easiness intercept) ของข้อสอบ ข้อที่ i

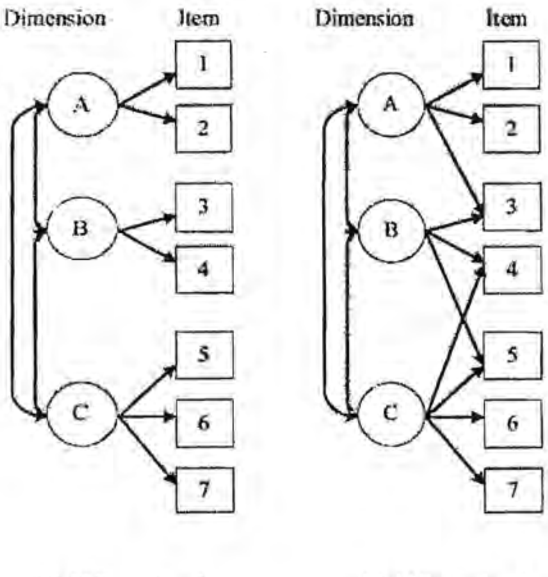
a_{im} = ค่าอำนาจจำแนกของข้อสอบข้อที่ i บนมิติ m

C_i = ค่าโอกาสการเดาข้อสอบได้ถูก (guessing) ของข้อสอบข้อที่ i

ประเภทของโมเดลการตอบสนองข้อสอบแบบพหุมิติ

ในการวิเคราะห์โมเดลการตอบสนองข้อสอบแบบพหุมิติสามารถทำได้ 2 วิธี คือ การวิเคราะห์โดยใช้โมเดลเชิงสำรวจ (exploratory model) และการวิเคราะห์โดยใช้โมเดลเชิงยืนยัน (confirmatory model) ซึ่งมีแนวคิดคล้ายกับการวิเคราะห์องค์ประกอบ (Factor analysis) กล่าวได้ว่า Exploratory IRT Model เป็นการประมาณค่าพารามิเตอร์ของข้อสอบและพารามิเตอร์ความสามารถของผู้สอบที่มีมากกว่า 1 มิติ เพื่อเพิ่มความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ โดยจะไม่มีการกำหนดจำนวนองค์ประกอบของคุณลักษณะแฝงในการวิเคราะห์ แต่จะหาจำนวนองค์ประกอบหรือมิติแฝงของบุคคลที่มีผลต่อการตอบข้อสอบ สำหรับ confirmatory IRT Model เป็นการประมาณค่าพารามิเตอร์ของข้อสอบและพารามิเตอร์ความสามารถของผู้สอบตามจำนวนองค์ประกอบของคุณลักษณะแฝงที่กำหนดไว้ล่วงหน้า

ยกตัวอย่างเช่นแบบสอบที่สร้างขึ้นมุ่งวัดคุณลักษณะแฝงหรือความสามารถของบุคคล 3 ด้าน (3 มิติ หรือ 3 องค์ประกอบ) ความสัมพันธ์ระหว่างคุณลักษณะของข้อสอบและความสามารถของผู้สอบที่ส่งผลต่อความน่าจะเป็นของการตอบข้อสอบได้ถูกต้อง สามารถเขียนโมเดลการวัดแบบพหุมิติ ได้ 2 ลักษณะ ดังภาพ



A B

ที่มา Cheng et al. (2014)

ภาพ 2.8 ลักษณะของโมเดลการตอบสนองข้อสอบแบบพหุมิติระหว่างข้อสอบ A (Between-Items MIRT Model) และโมเดลการตอบสนองข้อสอบแบบพหุมิติภายในข้อสอบ B (Within-Items MIRT Model)

จากภาพ 2.8 ในภาพ A เป็นโมเดลการตอบสนองข้อสอบแบบ 3 มิติ ซึ่งความสามารถทั้ง 3 ด้าน มีความเกี่ยวข้องสัมพันธ์กันโดยความสามารถแต่ละด้านถูกวัดระหว่างข้อสอบต่างชุดกัน (between-items) หรือคุณลักษณะแฝงแต่ละด้านถูกวัดโดยใช้ข้อสอบต่างชุดกัน ส่วนภาพ B เป็นโมเดลการตอบสนองข้อสอบแบบ 3 มิติ ซึ่งความสามารถทั้ง 3 ด้าน มีความเกี่ยวข้องสัมพันธ์กันโดยความสามารถแต่ละด้านถูกวัดภายในข้อสอบชุดเดียวร่วมกัน (within-items)

องค์ประกอบของการทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

องค์ประกอบของการทดสอบปรับเหมาะแบบพหุมิติ มีขั้นตอนและองค์ประกอบคล้ายกับการทดสอบแบบปรับเหมาะแบบเอกมิติ ซึ่งมีรายละเอียดดังนี้

1) คลังข้อสอบสำหรับการทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

ข้อสอบเพื่อจัดทำคลังข้อสอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติจะวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ โดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory Model: MIRT) เนื่องจากการตอบข้อสอบแบบเอกมิติ ในบางครั้งไม่เหมาะสมกับการทดสอบจริง เพราะข้อสอบบางข้อมีการวัดที่มุ่งจะวัดคุณลักษณะแฝงหรือความสามารถของบุคคลมากกว่า 1 องค์ประกอบ หรือ หลายมิติ จากการเปรียบเทียบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) กับการทดสอบปรับเหมาะพหุมิติด้วยคอมพิวเตอร์ (MCAT) นั้น พบว่าโมเดลที่อยู่ภายใต้ทฤษฎีตอบสนองข้อสอบแบบพหุมิติ มีการพัฒนาคลังข้อสอบที่มีข้อสอบจำนวนมากครอบคลุมเนื้อหาที่ต้องการวัด มีการประมาณค่าความสามารถของผู้สอบ มีการคัดเลือกข้อสอบ และมีเกณฑ์ยุติการสอบ ซึ่งสิ่งที่แตกต่างกัน ระหว่าง CAT กับ MCAT คือข้อสอบแต่ละข้อของ MCAT จะสามารถวัดความสามารถของผู้สอบได้หลายอย่าง ข้อสอบแต่ละข้อของ CAT วัดความสามารถของผู้สอบเพียงด้านเดียว การทดสอบแบบ MCAT จึงให้สารสนเทศได้มากกว่าการทดสอบแบบ CAT เมื่อใช้จำนวนข้อสอบที่มีความยาวของแบบทดสอบเท่ากัน ดังนั้น จำนวนข้อสอบในคลังข้อสอบต้องมีจำนวนมากเพียงพอกับทุกระดับความสามารถของผู้สอบ ข้อสอบในคลังข้อสอบควรมีค่าความยากกระจายเต็มระดับความสามารถของประชากรผู้สอบ ค่าอำนาจจำแนกสูง และค่าการเดาเข้าใกล้หรือเท่ากับศูนย์ คลังข้อสอบขนาดใหญ่จะให้สารสนเทศแบบทดสอบได้ดีกว่าคลังข้อสอบขนาดเล็ก

2) การประมาณค่าความสามารถในการทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

การประมาณค่าความสามารถของโมเดลการตอบสนองข้อสอบแบบพหุมิติ (MIRT) มีเกณฑ์การประมาณค่าที่แตกต่างกันหลายแนวทาง เช่น Maximum Likelihood Criterion, Maximum a Posteriori Bayesian Criterion, Newton-Raphson Method และ Least Squares Criterion (Reckase, 2009) แต่โดยทั่วไปการประมาณค่าความสามารถของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบบพหุมิติ (MCAT) ที่นิยมมีอยู่ 2 วิธี คือ วิธีการประมาณค่าแบบความน่าจะเป็นสูงสุด (Maximum Likelihood Estimation) และวิธีการประมาณค่าแบบเบส์ (Bayesian Estimation) (Reckase, 2009) และจากข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระของข้อสอบและผู้สอบในทฤษฎีการตอบสนองข้อสอบ ถ้าตัวแปรความสามารถมีค่าคงที่ การตอบสนองข้อสอบของผู้สอบต่อข้อสอบแต่ละข้อของแบบทดสอบ ถูกคาดหวังว่าเป็นอิสระจากกันในเชิงสถิติ การประมาณค่าความสามารถด้วยเทคนิคการประมาณค่า แบบความน่าจะเป็นสูงสุด (Maximum Likelihood Estimation) ก่อให้เกิดความไม่ชัดเจนในการประมาณค่าของเวกเตอร์ความสามารถ $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_p)$ ณ จุดเริ่มต้นของการ

ทดสอบ แต่วิธีการประมาณค่าแบบเบส์จะหลีกเลี่ยงข้อบกพร่องนี้ แต่ก็ยังไม่มีข้อสรุปเกี่ยวกับปัญหาใน การประมาณค่าความสามารถ ในทางปฏิบัติจึงมีการใช้วิธีการประมาณค่าทั้งแบบความน่าจะเป็น สูงสุดและแบบเบส์ (Frey & Seitz, 2009) และมีวิธีการประมาณค่าความสามารถที่สำคัญ ดังนี้

2.1) วิธีการประมาณค่าแบบความน่าจะเป็นสูงสุด (Maximum likelihood Estimation)

การประมาณค่าความสามารถด้วยวิธี Maximum Likelihood มีวิธีการคำนวณดังสมการ ต่อไปนี้ (Segall, 2010; Reckase, 2009)

$$f(U_{i1} = u_{i1}, U_{i2} = u_{i2}, \dots, U_{in} = u_{in}|\boldsymbol{\theta}) = L(\boldsymbol{u}|\boldsymbol{\theta}) \\ = \prod_{i\in S_n} p_i(\boldsymbol{\theta})^{u_i} q_i(\boldsymbol{\theta})^{1-u_i}$$

เมื่อ $f(U_{i1} = u_{i1} = u_{i2}, \dots, u_{in} = u_{in} / \theta)$ เป็นผลจากฟังชั่นการตอบข้อสอบข้อที่จะนำไปใช้ในการคำนวณในขั้นตอนการคัดเลือกข้อสอบและการให้คะแนน

$L(\mu|\theta)$ คือ ฟังชั่นความน่าจะเป็น

S_n คือ ข้อสอบข้อที่ 1,2,...,n

$P_i(q)$ คือ ความน่าจะเป็นในการตอบข้อสอบถูก

$q_i(q) = 1 - P_i(q)$

วิธีการประมาณค่าแบบนี้มักจะมีปัญหาไม่สามารถประมาณค่าได้ ถ้ามีจำนวนข้อของข้อสอบน้อย (Diao & Reckase, 2009) และจากการศึกษาของ Diao และ Reckase โดยใช้ข้อสอบ 50 ข้อ พบว่า การประมาณค่าด้วยวิธีนี้ใช้เวลานานกว่าจะประสบความสำเร็จในการประมาณค่าความสามารถ (Diao & Reckase, 2009)

2.2) วิธีการประมาณค่าแบบเบส์ (Bayesian Estimation)

การประมาณค่าความสามารถแบบ Bayesian มีวิธีการคำนวณแสดงดังสมการ (Segall, 1996 อ้างถึงในทัศน์ศิรินทร์ สว่างบุญ ,2554; Reckase, 2009)

$$f(\theta|u) \;\; \alpha \;\; L(u|\theta)f(\theta)$$

สำหรับการประมาณค่าแสดงด้วย $\hat{\boldsymbol{\theta}}$ ซึ่งเป็นค่าของ θ ที่ต้องการสำหรับเซตของสมการ $\partial \, \boldsymbol{ln\, f(\theta|u)\partial\theta = 0}$ ที่ดำเนินไปพร้อมกัน p ครั้งโดยที่

$$\frac{\partial}{\partial\theta} lnf(u|\theta) = D\sum_{i=1}^{n} v_i a_i - \theta$$

และ

$$v_i = \frac{[P_i(\theta)-c_i][u_i-P_i(\theta)]}{(1-c_i)P_i(\theta)}$$

และเนื่องจากเซตของสมการนี้ไม่มีการแก้ปัญหารูปแบบปิด ดังนั้นวิธีการวนซ้ำจึงถูกนำมาใช้ ซึ่งในที่นี้จะใช้วิธีของ Newton-Raphson โดยกำหนดให้ **θ** (m) แทนการประมาณลำดับที่ m ของค่า θ ที่สูงที่สุดใน $\boldsymbol{f(\theta|u)}$ ดังนั้นวิธีประมาณค่าที่ดีกว่าคือ ทำให้เป็นรูปทั่วไป โดยให้

$$\theta^{(m+1)} = \theta^{(m)} - \delta^{(m)}$$

โดยที่ $\delta^{(m)}$ เป็นเวกเตอร์ 1xp

$$\delta^{(m)} = \left[J(\theta^{(m)})\right]^{-1} x \frac{\partial}{\partial\theta} lnf(\theta^{(m)}|u)$$

เมทริกซ์ $J(\theta)$ เป็นเมทริกซ์ของอนุพันธ์บางส่วนอันดับสองซึ่งถูกประเมินที่ตำแหน่ง $\theta=\theta^{(m)}$

$$J(\theta) = D^2 \sum a_i a_i' w_i - I$$

โดยที่ $w_i = \frac{Q_i(\theta)[P_i(\theta)-c_i][c_i u_i - P_i^2(\theta)]}{P_i^2(\theta)(1-c_i)^2}$

การประมาณค่าจะดำเนินการด้วยสมการทั้ง 2 จนกระทั่งค่าลู่เข้าหากัน

จากการศึกษาของ Diao และ Reckase โดยใช้ข้อสอบจำนวน 50 ข้อ พบว่า การประมาณค่าด้วยวิธีนี้ประสบความสำเร็จในการประมาณค่าด้วยเวลาที่รวดเร็ว (Diao & Reckase, 2009)

2.3) วิธีการประมาณค่าแบบนิวตันราฟ (Newton-Raphson Method)

วิธีการนี้ใช้ควบคู่ไปกับการประมาณค่าวิธีแบบความน่าจะเป็นสูงสุด ในกรณีที่ไม่สามารถประมาณค่าความสามารถได้ (Segall, 1996 อ้างถึงในทัศน์ศิรินทร์ สว่างบุญ ,2554; Chen, 2009; Frey & Seitz, 2009) โดยมีวิธีการคำนวณดังสมการ (Segall, 1996 อ้างถึงในทัศน์ศิรินทร์ สว่างบุญ ,2554; Chen, 2009)

$$\theta^{(j+1)} = \theta^{(j)} - \delta^{(j)}$$

เมื่อ $\delta^{(j)}$ คือ เวกเตอร์ P × 1 และคำนวณค่าได้ดังสมการ

$$\delta^{(j)} = -\left[H(\theta^{(j)})\right]^{-1}\left(\frac{\partial}{\partial\theta} InL(u/\theta)\right)$$

เมื่อ $H(\theta^{(j)})$ คือ เมทริกซ์ P × P เป็นค่าของอนุพันธ์อันดับสองของฟังก์ชั่นความน่าจะเป็นของ $\theta^{(j)}$ และ $H(\theta)$ แสดงได้ดังสมการ

$$\mathbf{H}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial^2 \ln L}{\partial \theta_1^2} & \frac{\partial^2 \ln L}{\partial \theta_1 \partial \theta_2} & \frac{\partial^2 \ln L}{\partial \theta_1 \partial \theta_p} \\ \frac{\partial^2 \ln L}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 \ln L}{\partial \theta_2^2} & \frac{\partial^2 \ln L}{\partial \theta_2 \partial \theta_p} \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 \ln L}{\partial \theta_p \partial \theta_1} & \frac{\partial^2 \ln L}{\partial \theta_p \partial \theta_2} & \frac{\partial^2 \ln L}{\partial \theta_p^2} \end{bmatrix}$$

และ $\frac{\partial}{\partial \theta} InL(u/\theta)$ เป็นค่าของอนุพันธ์อันดับหนึ่งของฟังก์ชั่นความน่าจะเป็นของ $\theta^{(j)}$ แสดงได้ดังสมการ

$$\frac{\partial}{\partial \boldsymbol{\theta}} \ln L(\mathbf{u}|\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial}{\partial \boldsymbol{\theta}_1} \ln L(\mathbf{u}|\boldsymbol{\theta}) \\ \frac{\partial}{\partial \boldsymbol{\theta}_2} \ln L(\mathbf{u}|\boldsymbol{\theta}) \\ \vdots \\ \frac{\partial}{\partial \boldsymbol{\theta}_\mathrm{p}} \ln L(\mathbf{u}|\boldsymbol{\theta}) \end{bmatrix}$$

2.4) วิธีการประมาณค่าแบบกำลังสองน้อยที่สุด (Least Squares Estimation)

วิธีการนี้พบได้ไม่บ่อยนักในการประมาณค่าพารามิเตอร์ของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ แต่ก็เป็นวิธีที่ใช้ในโปรแกรม NOHARM มีสูตรคำนวณ มีวิธีการคำนวณแสดงได้ดังสมการ (Reckase, 2009)

$$SS_\theta = \sum_{i=1}^{n} (u_i - P(u_i = 1/\theta, a_i, d_i))^2$$

เมื่อ SS_θ คือ ผลรวมของความแตกต่างยกกำลังของแต่ละเวกเตอร์ θ

U_i คือ คะแนนที่ได้จากข้อสอบข้อที่ i (มีค่าเป็น 0 และ 1)

3) วิธีการคัดเลือกข้อสอบในการทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

วิธีการคัดเลือกข้อสอบที่ใช้กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ อาศัยความน่าจะเป็นสูงสุดหรือต่ำสุดจากการประมาณค่าความสามารถ และมีวิธีการคัดเลือกข้อสอบอยู่หลากหลายวิธี แต่ละวิธีสามารถใช้ร่วมกันกับวิธีการประมาณค่าทุกแบบและยังสามารถใช้วิธีการคัดเลือกข้อสอบหลายวิธีร่วมกันได้อีกด้วย (Reckase, 2009) โดยมีวิธีการคัดเลือกที่สำคัญ ดังนี้

3.1) วิธีการคัดเลือกข้อสอบแบบสารสนเทศของฟิชเชอร์ (Maximize the Determinant of the Fisher Information Matrix)

การคัดเลือกข้อสอบด้วยเมทริกซ์สารสนเทศฟิชเชอร์ ดำเนินการภายใต้ เงื่อนไขการประมาณค่าเวกเตอร์ความสามารถที่กำหนดจากเวกเตอร์ค่าเฉลี่ยเท่ากับ 0 และ เวกเตอร์ความสามารถจริง (Tureθ) มีการแจกแจงแบบปกติหลายตัวแปรด้วยเมทริกซ์ความแปรปรวนและความแปรปรวนร่วม (Variance–Covariance

Matrix) ที่เกี่ยวข้องกับเมทริกซ์สารสนเทศของฟิชเชอร์ (Fisher Information Matrix) (Segall, 1996 อ้างถึงในทัศน์ศิรินทร์ สว่างบุญ ,2554; Frey & Seitz, 2009; Reckase, 2009) โครงสร้างทฤษฎีการตอบสนองข้อสอบแบบพหุมิติมองเมทริกซ์สารสนเทศของฟิชเชอร์เป็นเหมือนการวัดปริมาณของสารสนเทศการตอบสนองของข้อสอบและนำเสนอเกี่ยวกับความสามารถภายหลังการตอบข้อสอบข้อนั้นแล้ว (θ) ในทางคณิตศาสตร์ค่าความคาดหวังของอนุพันธ์ที่สองของความน่าจะเป็นลอกลิฮู้ด ที่เกี่ยวข้องกับความสามารถจะมีค่าเป็นลบ (Frey & Seitz, 2009) สามารถแสดงในวิธีการทางานแบบพหุมิติที่นาเสนอโดย Segall ในปี 1996 ได้ดังสมการ (Segall, 1996 อ้างถึงในทัศน์ศิรินทร์ สว่างบุญ ,2554; Frey & Seitz, 2009; Reckase, 2009)

$$I_i(\theta) = -E\left[\frac{\partial^2}{\partial\theta\partial\theta'} log\, f(x_i|\theta))|\theta)\right]$$

จากสมการข้างต้น ถ้าใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ 3 พารามิเตอร์ ตามโมเดลในสมการดังนี้

$$I_i(\theta) = \frac{Q_i(\theta)[P_j(\theta) - c_i]^2}{P_j(\theta)(1 - c_i)} \begin{bmatrix} a_{i1}^2 & a_{i1}a_{i2} & \dots & a_{i1}a_{ip} \\ a_{i1}a_{i2} & a_{i2}^2 & \dots & a_{i2}a_{ip} \\ & \vdots & & \\ a_{i1}a_{ip} & a_{i2}a_{ip} & \dots & a_{ip}^2 \end{bmatrix}$$

ให้ $P_j(\theta)$ แทน ฟังก์ชันการตอบสนองข้อสอบข้อที่ j

$Q_i(\theta)$ แทน $1 - P_j(\theta)$

เมื่อ $Q_i(\theta) = 1 - P_i(\theta)$ เมทริกซ์สารสนเทศของข้อสอบจะเพิ่มขึ้นจากเมทริกซ์สารสนเทศของแบบสอบ โดยยังคงรักษาคุณสมบัติเดิม ในกรณีนี้คุณสมบัติของ MLE มีความสัมพันธ์กับ FI โดยเฉพาะอย่างยิ่งเมื่อกำหนด θ เป็นเวกเตอร์มิติ p และ MLE ของ $\widehat{\theta}^{mle}$ เป็นการแจกแจงที่ $\widehat{\theta}^{mle} \sim N(\theta, I_s^{-1}(\widehat{\theta}^{mle}))$ เมื่อ $I_s^{-1}\widehat{\theta}^{mle}$ เป็นค่าผกผันของเมทริกซ์สารสนเทศจากการประเมิน $\widehat{\theta}^{mle}$ กับแต่ละส่วนที่แสดงความแปรปรวนของมิติความสามารถหนึ่งมิติหรือความแปรปรวนร่วมระหว่างมิติความสามารถสองมิติ

3.2) วิธีการคัดเลือกข้อสอบแบบสารสนเทศของ Kullback-Leibler (Maximize Kullback-Leibler Information) วิธีการ Posterior Expected Kullback–Leibler Information ถูกนำเสนอ ให้ใช้ในการคัดเลือกข้อสอบในการทดสอบแบบปรับเหมาะที่เป็นมิติเดียว Chang and Ying (1996) และต่อมา Veldkamp and van der Linden (2002) ได้เสนอทางเลือกให้ใช้ วิธีการ Posterior Expected Kullback–Leibler Information ในการคัดเลือกข้อสอบแบบปรับเหมาะแบบพหุมิติ (MCAT) โดยมีเกณฑ์การคัดเลือกข้อสอบจากระยะห่างของการกระจายของการตอบสนอง จากข้อสอบข้อที่จะถูกเลือกให้ใช้ทดสอบข้อต่อไป เป็นตำแหน่งที่ใช้ประมาณค่าความสามารถของ ผู้สอบ ($\widehat{\theta}$) กับความสามารถที่แท้จริงของผู้สอบ (θ) ด้วยการคาดหมายเอาจากการตอบข้อสอบ (Mulder & van der Linden, 2010; Reckase, 2009) กรณีข้อสอบ 1 ข้อ คำนวณได้ดังสมการ

$$K_i(\widehat{\theta}\,;\theta) = -E\left[\log\frac{f(U_i|\widehat{\theta})}{f(U_i|\theta)}\right]$$

$$K_i(\hat{\theta};\theta) = P_i(\hat{\theta})\log\frac{P_i(\hat{\theta})}{P_i(\theta)} + Q_i(\hat{\theta})\log\frac{Q_i(\hat{\theta})}{Q_i(\theta)}$$

เมื่อ $P_i(*)$ เป็นความน่าจะเป็นในการตอบสนองข้อสอบข้อที่ i ได้ถูกต้อง

$$Q_i(*) = 1 - P_i(*)$$

$U_i(*)$ เป็นผลการตอบสนองข้อสอบข้อที่ กรณีข้อสอบ n ข้อ

$$K_n(\hat{\theta};\theta) = \sum_{i=1}^{n} K_i$$

ส่วนแนวทางในการคัดเลือกข้อสอบที่แนะนำโดย Veldkamp and van der Linden (2002) คือ เลือกจากข้อสอบที่มีค่าสารสนเทศ KL จากการคาดหวังภายหลังสูงสุด (Maximizes the Posterior Expected Kullback–Leibler Information) ที่ถ่วงน้ำหนักสารสนเทศ ด้วยฟังก์ชั่นความหนาแน่นภายหลังจากการประมาณค่าความสามารถ และ Veldkamp and van der Linden (2002) ใช้ Gauss–Hermite Quadrature ประเมินค่า แต่อย่างไรก็ตาม วิธีการ Quadrature วิธีอื่นๆ ก็สามารถใช้งานได้ดีเช่นกัน (Reckase, 2009)

$$K_i^B(\hat{\theta}^{k-1}) = \int_{\theta} K_i(\theta, \hat{\theta}^{k-1}) f(\theta|u_1, \ldots, u_{k-1}) \partial\theta$$

เมื่อ K_i^B	คือ	สารสนเทศความคาดหวังภายหลังแบบเบส์
$\hat{\theta}^{k-1}$	คือ	ความสามารถที่ประมาณค่าหลังจากข้อสอบ ข้อที่ k-1 สอบไปแล้ว
$f(\theta\|u_1, \ldots, u_{k-1})$	คือ	ความหนาแน่นภายหลังจากการสอบข้อสอบข้อที่ k-1 ผ่านไปแล้ว

แต่อย่างไรก็ตาม วิธีการนี้ค่อนข้างจะมีความซับซ้อนและจากการทดลอง ใช้งานของ Reckase ด้วยโปรแกรม MATLAB ด้วยคุณลักษณะแฝง 3 คุณลักษณะ พบว่า การคำนวณทำได้ช้ามาก (Reckase, 2009)

4) การควบคุมการใช้ข้อสอบซ้ำในการทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

วิธีการควบคุมการใช้ข้อสอบซ้ำสำหรับการทดสอบแบบพหุมิติ มีรายละเอียด ดังนี้

4.1) วิธี Maximum likelihood estimation and D-optimality ในการคัดเลือกข้อสอบสำหรับการทดสอบแบบเอกมิติจะยึดหลักเกณฑ์สารสนเทศสูงสุดของฟิชเชอร์ (maximum Fisher information criterion) เมื่อการทดสอบเป็นแบบพหุมิติและมีลักษณะเป็น 2 มิติ (d = 2) วิธีการคัดเลือกข้อสอบสามารถสร้างขึ้นได้ โดยสมมติให้ $\theta = (\theta_1, \theta_2), a = (a_1, a_2)$ หลังจากที่ผู้สอบทำข้อสอบแล้วจำนวน k ข้อ สารสนเทศฟิชเชอร์ที่ความสามารถ θ หาได้จาก

$$I_k(\theta) = D^2 \begin{bmatrix} \sum_{i=1}^{k} a_{1i}^2 p_1(\theta) q_i(\theta) & \sum_{i=1}^{k} a_{1i} a_{2i} p_1(\theta) q_i(\theta) \\ \sum_{i=1}^{k} a_{2i} a_{1i} p_1(\theta) q_i(\theta) & \sum_{i=1}^{k} a_{2i}^2 p_1(\theta) q_i(\theta) \end{bmatrix}$$

โดยที่ $q_i(\theta) = 1 - p_i(\theta)$

การประมาณค่าความเป็นไปได้สูงสุด (maximum likelihood estimate: MLE) ได้มาเพื่อแทนที่ปริมาณความสามารถ θ ที่ไม่ทราบค่าในการคำนวณค่าสารสนเทศขณะที่ข้อถัดไปหรือข้อที่ k+1 จะต้องถูกเลือก ดังนั้นสารสนเทศที่มากที่สุดสามารถให้ที่ความสามารถ $\hat{\theta}$

ประโยชน์ของวิธี MLE ภายใต้เงื่อนไขกฎเกณฑ์ โดยวิธีนี้จะมีแนวโน้มไปสู่การแลกแลงแบบปกติของตัวแปรพหุ ด้วยการกระจายของเมตริกซ์ $I_{k+1}^{-1}(\theta)$ ซึ่งเป็นส่วนกลับของเมตริกซ์สารสนเทศของฟิชเชอร์ ความเชื่อมั่นสำหรับความสามารถ θ มาจากรูปแบบ ดังนี้

$$\{\theta: (\theta - \hat{\theta})^T I_{k+1}(\hat{\theta})(\theta - \hat{\theta}) \leq cons \tan t\}$$

เพราะปริมาณของพื้นที่ความเชื่อมาน (ellipsoid) เป็นปริมาณที่เป็นสัดส่วนกันกับ $|I_{k+1}(\hat{\theta})|^{1-/2}$ (นั่นคือ) $|\cdot|$ แทน determinant โดยที่

$$|I_{k+1}(\hat{\theta})| = D^4 \left\{ \sum_{i=1}^{k+1} a_{1i}^2 p_i(\hat{\theta}) q_i(\hat{\theta}) \cdot \sum_{i=1}^{k+1} a_{2i}^2 p_i(\hat{\theta}) q_i(\hat{\theta}) - \left(\sum_{i=1}^{k+1} a_{1i} a_{2i} p_i(\hat{\theta}) q_i(\hat{\theta}) \right)^2 \right\}$$

ธรรมชาติของการออกแบบเกณฑ์ คือ การทำให้เกิดพื้นที่ความเชื่อมั่น (ellipsoid) น้อยที่สุดเท่าที่จะเป็นไปได้ คือการเลือกข้อสอบข้อที่ (K+1) ที่ให้ค่า $|I_{k+1}(\hat{\theta})|$ สูงสุดซึ่งเป็นเกณฑ์ที่ได้มาตรฐานของ D-optimality หลังจากทำข้อสอบข้อที่ (K+1) และการตอบข้อสอบนั้นทำให้เกิดการจูงใจจากผู้สอบผลการตอบจะถูกให้คะแนนและความสามรถ θ จะถูกประมาณค่าใหม่ กระบวนการนี้จะถูกทำซ้ำจนกระทั่งการทดสอบยุติและความสามารถ θ ในขั้นสุดท้ายถูกประมาณค่า

4.2) วิธีการคัดเลือกข้อสอบด้วยวิธี a-stratified multistage

การเสนอวิธีการควบคุมการใช้ข้อสอบซ้ำนี้จะใช้การทำหน้าที่ของ a ในการแบ่งลำดับขั้นคลังข้อสอบ นอกเหนือจากกลไกการควบคุม การใช้ข้อสอบซ้ำ วิธีการคัดเลือกข้อสอบจะขึ้นอยู่กับ D-optimality เพียงอย่างเดียวมีแนวโน้มที่จะเลือกข้อสอบที่มี a_1 และ a_2 ใกล้กับขอบเขตบน เมื่อ d=1 ซึ่งหมายความว่าข้อสอบที่มีค่าพารามิเตอร์ a สูง จะมีความเหมาะสมที่จะถูกเลือก เมื่อ d > 1 แนวคิดที่จะแบ่งลำดับขั้นของแบบแผนจะขึ้นอยู่กับ ค่าพารามิเตอร์ a เช่นเดียวกับ d=1 ผลของ b ถูกกำหนดใน $p_i(\hat{\theta})$ และ $q_i(\hat{\theta})$ เมื่อแบบทดสอบมีความก้าวหน้า θ จะมีค่าคงที่ และ |I| คือ หน้าที่ของเวคเตอร์ a ดังนั้นความท้าทายของการคัดเลือกที่สมเหตุสมผล ในการเลือกฟังก์ชันที่เหมาะสมของ a ซึ่ง ทำให้แน่ใจได้ว่าลักษณะของการดำเนินการ a-stratified

สำหรับคลังข้อสอบแบบพหุมิติ (multidimensional item pool) มีความคล้ายคลึงกับคลังข้อสอบแบบเอกมิติ (unidimensional item pool)

Chang and Ying (1996) ได้เสนอกระบวนการของ a-Stratified multistage (ASTR) ไว้ดังนี้

1) การแบ่งแยกข้อสอบในคลังข้อสอบเป็น M ระดับ ตามค่าของพารามิเตอร์ a ของข้อสอบ ชั้นของข้อสอบข้อแรก (first item stratum) จะกำหนดด้วยค่าที่น้อยที่สุด ของ $|a_1, a_2|$

2) แบ่งแยกแต่ละระดับยกเว้นชั้นแรก เป็น 2 ส่วนย่อย ส่วนแรก ประกอบด้วยข้อสอบ $(a_1 > a_2)$ ส่วนที่สองประกอบด้วยข้อสอบ $(a_1 < a_2)$

3) แบ่งแยกข้อสอบเป็น M ขั้น

4) ขั้นแรกเลือกจากชั้นของข้อสอบข้อแรก (first item stratum) จำนวน n_1 ข้อตามเกณฑ์นี้สำหรับการคัดเลือกแต่ละจุดจะเลือกข้อสอบที่มีค่าความยาก พารามิเตอร์ b ที่มีค่าใกล้กับค่าตามสูตร

$$b = \frac{a_1\hat{\theta}_1 + a_2\hat{\theta}_2}{a_1 + a_2}$$

5) ในขั้น ที่ j เมื่อ j > 1 เลือกข้อสอบข้อที่ n_{j1} และ n_{j2} แต่ละส่วนย่อยของชั้น ที่ j เช่น $n_j = n_{j1} + n_{j2}$ ถูกแบ่งชั้น จาดนั้นจับคู่ความยากของข้อคำถามพารามิเตอร์ b ค่าเฉลี่ยถ่วงน้ำหนักของ $\hat{\theta}$

6) ดำเนินการซ้ำในข้อที่ 4 สำหรับ $J = 2, \ldots, M. \sum_{j=2}^{m} n_{j1} = \sum_{j=2}^{m} n_{j2}$

ค่าของ M, n_j , j = 1, ..., M ถูกกำหนดในแบบทดสอบไว้ล่วงหน้า แบบทดสอบสามารถออกแบบไปยังการสิ้นสุดการสอบ เมื่อ $\hat{\theta}$ ไปถึงระดับที่ระบุได้อย่างแม่นยำ นอกจากนั้น เมื่อจำนวนข้อสอบที่ $(a_1 > a_2)$ ถูกปรับเทียบกับจำนวนข้อสอบที่ $(a_1 < a_2)$ แบบสอบจะมีความสมดุลในเรื่องความสามารถในการจำแนกในทั้ง 2 มิติ หรืออาจกล่าวได้ว่า θ_1และ θ_2 ควรได้รับการวัดที่มีความแม่นยำเท่าเทียมกัน

5) กฎการยุติของการทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

กฎการยุติการทดสอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยทั่วไปใช้กฎการยุติ 2 ลักษณะ คือ กฎความยาวคงที่ (Fixed-Length) และ กฎความยาวแปรผัน (Variable-Length)

กฎความยาวคงที่ (Fixed-Length) เป็นการกำหนดจานวนข้อสอบเท่าๆ กัน และจัดให้กับผู้สอบแต่ละคน เมื่อการทดสอบครบตามจำนวนข้อสอบกำหนด การทดสอบจึงยุติ การศึกษาโดยใช้กฎความยาวคงที่ มักสนใจความคลาดเคลื่อนมาตรฐานของค่าประมาณความสามารถของผู้สอบเมื่อสิ้นสุดการทดสอบว่ามากน้อยเท่าไร (Thissen & Mislevy, 1990) กฎนี้ง่ายในการใช้ สามารถทำนายอัตราการใช้ข้อสอบได้ แต่ต้องระวังการใช้กฎความยาวคงที่อาจให้ฟังก์ชันสารสนเทศเป็นโค้งลักษณะแบน ดังนั้น กฎความยาวคงที่อาจไม่ให้ความแม่นยำเท่าเทียมกันที่ระดับความสามารถต่างๆ (Thissen & Wainer, 1990)

กฎความยาวแปรผัน (Variable-Length) เป็นการจัดข้อสอบให้อย่างต่อเนื่องจนระดับความแม่นยำของค่าประมาณความสามารถของแต่ละบุคคล มาถึงระดับความแม่นยำตามกำหนด การทดสอบจึงยุติ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใช้กฎความยาวแปรผันเพื่อกำหนดความยาวแบบทดสอบว่า จำนวนข้อสอบเท่าไรควรจัดให้แก่ผู้สอบระดับความสามารถต่างๆ (Thissen & Mislevy, 1990) ซึ่งการใช้กฎการยุติจากค่าความคลาดเคลื่อนมาตรฐานจะให้ความแม่นยำ ในการวัดผู้สอบทั้งระดับความสามารถสูงและต่ำ (Meijer &

Nering, 1999) โดยทั่วไปจะใช้กฎการยุติการทดสอบจากค่าความคลาดเคลื่อนมาตรฐานกำหนดไว้ที่ 0.3 (Thissen & Wainer, 1990)

การยุติการทดสอบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ มีเกณฑ์ในการยุติการทดสอบที่นำเสนอหลากหลายเกณฑ์ เช่น หยุดเมื่อจำนวนข้อสอบที่ทดสอบครบตามจำนวน หรือเมื่อการวัดมีความแม่นยำตามเกณฑ์ หรือใช้ทั้งสองเกณฑ์ร่วมกัน การศึกษาจากข้อมูลจาลองส่วนใหญ่ การทดสอบแบบปรับเหมาะแบบพหุมิติ จะยุติภายหลังการสอบครบตามจำนวนข้อสอบที่กำหนดไว้ล่วงหน้า แต่อย่างไรก็ตาม การกำหนดจำนวนข้อสอบที่ใช้ในการทดสอบเป็นเกณฑ์การยุติการสอบ ส่งผลต่อความแม่นยำในการประมาณค่าความสามารถที่มีความแปรปรวนระหว่างผู้สอบสูงมาก ผู้สอบบางคนมีความคลาดเคลื่อนมาตรฐานของการวัดจากหนึ่งมิติหรือมากกว่าจะมีขนาดใหญ่ ในขณะที่คนอื่นจะมีขนาดเล็กกว่ามาก และความแม่นยำ ในการประมาณค่าความสามารถยังขึ้นอยู่กับการตอบข้อสอบและคุณลักษณะของคลังข้อสอบ (Frey & Seitz, 2009) การยุติการทดสอบจากความแม่นยำในการวัดเป็นอีกทางเลือกหนึ่งที่ใช้เป็นเกณฑ์การยุติการทดสอบได้ เช่น การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติของ Segall (1996) การทดสอบดำเนินการต่อไปเรื่อยๆ จนกว่าค่าความคลาดเคลื่อนมาตรฐานจะตกอยู่ในช่วงความเชื่อถือได้ (ต่ำกว่าค่าที่กำหนดไว้) การกำหนดเกณฑ์วิธีนี้เป็นที่ยอมรับทั่วไปที่สามารถดำเนินการได้อย่างเหมาะสม การใช้ความแม่นยำในการวัดเป็นเกณฑ์ในการยุติการทดสอบจะมีประสิทธิภาพหรือไม่ขึ้นอยู่กับคุณลักษณะของคลังข้อสอบที่นำมาใช้งานด้วย ถ้าองค์ประกอบของความสามารถของผู้สอบตกอยู่ในขอบเขตภายใต้พื้นที่หลายพื้นที่และวัดด้วยข้อสอบจานวนมากกว่าหนึ่งข้อที่เป็นตัวแทนของเนื้อหาสาระจะทาให้แบบทดสอบนั้นมีความยาวมาก ดังนั้นการใช้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ จึงเป็นการรวมเอาการวัดที่มีความแม่นยำและใช้ข้อสอบจำนวนมาก กรณีที่บรรลุถึงระดับความแม่นยำที่ตั้งเอาไว้ช้า และ ใช้ข้อสอบจำนวนน้อย กรณีที่บรรลุถึงระดับความแม่นยำที่ตั้งเอาไว้รวดเร็ว (Frey & Seitz, 2009)

งานวิจัยที่เกี่ยวข้องกับทดสอบปรับเหมาะด้วยคอมพิวเตอร์

Deng, Ansley and Chang (2010) ได้ศึกษาเรื่อง การแบ่งชั้นและวิธีการคัดเลือกข้อสอบที่มีสารสนเทศสูงสุดในการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ ในการศึกษานี้เป็นการประเมินและเปรียบเทียบกับสามวิธีการคัดเลือกข้อสอบ:maximum Fisher information procedure (F), the a-stratified multistage computer adaptive testing (CAT) (STR) และวิธีการแบ่งชั้นข้อสอบที่จะได้รับการคัดเลือกในชั้นที่สูงขึ้นและข้อสอบในชั้นที่ต่ำ (USTR) พร้อมกับการคัดเลือกข้อสอบแบบสุ่ม (RAN) โดยการเปรียบเทียบส่วนที่เกี่ยวกับการคาดหวังความคลาดเคลื่อนของความแปรปรวน ความน่าเที่ยงของ การประมาณค่าความสามารถและการใช้ข้อสอบผ่านการจำลองการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ ภายใต้เงื่อนไข 9 เงื่อนไข ข้อจำกัดในทางปฏิบัติต่างๆ และขอบเขตการคัดเลือกข้อสอบ ผลการศึกษาพบว่า F มีประโยชน์และมีความแม่นยำและ STR USTR ภายใต้การคัดเลือกข้อสอบไม่มีข้อจำกัดแต่กับการใช้ข้อคำถามที่ไม่ดี USTR ลดความคลาดเคลื่อนของความแปรปรวนสำหรับ STR ภายใต้เงื่อนไขต่างๆ ที่มีการกำหนดในการใช้งานข้อสอบที่มีขนาดเล็ก เมื่อเทียบกับ F, USTR การใช้งานข้อสอบที่เพิ่มขึ้นในขณะที่การมีความแม่นยำเปรียบเทียบในการประมาณค่าความสามารถ; ผลการทดสอบในระดับที่มีความแม่นยำคล้ายกับ F กับการใช้งานที่ดีขึ้นเมื่อข้อสอบได้รับการคัดเลือกภายใต้การ

ควบคุมและการเปิดเผยกับการขอบเขตพื้นที่ในการคัดเลือกที่จำกัด ผลการทดสอบให้ความหมายสำหรับการเลือกวิธีการคัดเลือกข้อสอบที่เหมาะสมที่ประยุกต์ได้

Barrada et al. (2010) ศึกษาเรื่อง วิธีสำหรับการเปรียบเทียบกฎการคัดเลือกข้อสอบในการทดสอบปรับเหมาะสมด้วยคอมพิวเตอร์ ในการศึกษาเป็นการเปรียบเทียบประสิทธิภาพความสัมพันธ์ของ 2 กฎการคัดเลือกข้อสอบในการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ ผลที่เหมือนกันคือความแตกต่างกันในความถูกต้องและความปลอดภัยทำให้มันยากที่จะมีผลต่อข้อสรุปในกฎที่เหมาะสมมากขึ้น การศึกษาครั้งนี้ได้นำเสนอกลยุทธ์ในการดำเนินการเปรียบเทียบระดับโลกหรือมากกว่า 2 กฎของการคัดเลือก ในการศึกษาครั้งนี้แสดงให้เห็นถึงประสิทธิภาพการทำงานของแต่ละกฎการคัดเลือกอัตราการใช้ข้อสอบซ้ำสูงสุดโดยทั้งหมดจะถูกเมื่อเทียบกับกฎอื่น ๆ โดยนำกลยุทธ์ที่ถูกนำมาใช้ในการศึกษาแบบจำลองกับการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ที่มีความยาวคงที่สำหรับการเปรียบเทียบของ 6 กฎการคัดเลือก ได้แก่ the point Fisher information, Fisher information weighted by likelihood, Kullback-Leibler weighted by likelihood, maximum information stratification with blocking, progressive และ วิธี proportional ผลการศึกษาแสดงให้เห็นว่าไม่มีกฎที่เหมาะสมสำหรับค่าทับซ้อนใด ๆ หรือรากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) ในความจริงที่ว่ากฎสำหรับระดับที่กำหนดของการทับซ้อนมี RMSE ต่ำกว่าที่อื่นซึ่งไม่ได้หมายความว่ารูปแบบนี้ถือเป็นอัตราการทับซ้อนกัน การเปรียบเทียบกฎต้องมีการจัดการของอัตราซ้ำสูงสุดวิธีที่ดีที่สุดเป็นกลุ่ม Kullback-Leibler weighted by likelihood, วิธี proportional and the maximum information stratification method with blocking

Ho and Dodd (2012) ศึกษาเรื่อง การคัดเลือกข้อสอบและวิธีการประมาณค่าความสามารถสำหรับการทดสอบปรับเหมาะรูปแบบผสม ในการศึกษานี้เป็นการเปรียบเทียบ 5 วิธีการคัดเลือกข้อสอบที่ใช้ 3 วิธีการประมาณค่าความสามารถในบริบทของการทดสอบปรับเหมาะรูปแบบผสมใน generalized partial credit model วิธีการคัดเลือกข้อที่ใช้คือ maximum posterior weighted information, maximum expected information, maximum posterior weighted Kullback-Leibler information, และวิธีmaximum expected posterior weighted Kullback-Leibler information วิธีการประมาณค่าความสามารถในการตรวจสอบมีการประมาณค่า maximum likelihood estimation (MLE), weighted likelihood estimation (WLE), and expected a posteriori (EAP). ผลการศึกษาชี้ให้เห็นว่าวิธีการคัดเลือกข้อสอบโดยไม่คำนึงถึงฟังก์ชั่นของสารสนเทศที่ใช้ในดำเนินการอย่างเท่าเทียมกันกับวิธีการประเมินความสามารถ ข้อสรุปเกี่ยวกับวิธีการประเมินค่าความสามารถพบว่า MLE เป็นวิธีทางเลือกในทางปฏิบัติและ WLE ควรพิจารณาเมื่อมีข้อมูลที่ไม่ตรงกันระหว่างคลังข้อสอบและการกระจายความสามารถของประชากร EAP สามารถใช้เป็นทางเลือกในการทำงานเมื่อวิธีการกระจายที่เฉพาะเจาะจง

Huang, Chen and Wang (2013) ได้ศึกษาเรื่องการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้การจัดชั้นของโมเดลการตอบสนองข้อสอบแบบขั้นสูง ในวิทยาศาสตร์เกี่ยวกับมนุษย์นั้น ข้อสมมติฐานร่วมกันคือคุณลักษณะแฝงจะมีโครงสร้างแบบลำดับชั้น โมเดลทฤษฎีการตอบสนองข้อสอบแบบขั้นสูงได้รับการพัฒนาเพื่อการเข้าถึงสำหรับลำดับขั้นนี้ ในการศึกษานี้อัลกอริทึมของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) ขึ้นอยู่กับชนิดของโมเดลที่ถูกนำมาใช้ และประสิทธิภาพการทำงานภายใต้ความหลากหลายของสถานการณ์

ที่ถูกตรวจสอบโดยใช้แบบจำลอง ผลการศึกษาพบว่าอัลกอริทึม CAT จะมีประสิทธิภาพมาก วิธีการที่ก้าวหน้าสำหรับการคัดเลือกข้อสอบ และวิธีการของ Sympson และ Hetter กับวิธีการแบบออนไลน์และแบบแช่แข็งในการควบคุมการเปิดเผยข้อสอบ และโมเดล mulitinomial สำหรับการสมดุลเนื้อหาสามารถรักษาความแม่นยำในการวัดที่ดี การควบคุมการเปิดเผยข้อสอบ การสมดุลเนื้อหา ความปลอดภัยของการทดสอบและการใช้คลังข้อสอบได้

Mao and Xin (2013) ได้ศึกษาการประยุกต์วิธี Monte Carlo เพื่อการทดสอบปรับเหมาะด้วยคอมพิวเตอร์วินิจฉัยเชิงปัญญากับเนื้อหาจำกัด ซึ่งในการวิจัยนี้ได้ใช้วิธีการจำลองสถานการณ์แบบ Monte Carlo ซึ่งมีการใช้มาก่อนในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบดั้งเดิม (CAT) ซึ่งในที่นี้ได้ประยุกต์ใช้ CAT กับการวินิจฉัยเชิงปัญญา เพื่อตรวจสอบความสามารถของวิธีนี้เมื่อใช้กับเนื้อหาจำกัดที่หลากหลาย ความสามารถของวิธีการจำลองสถานการณ์แบบ Monte Carlo เป็นการเปรียบเทียบกับความสามารถของดัชนีการจำแนกสูงสุดอย่างกว้างสูงสุดแบบปรับ (MMGDI) ซึ่งทำในสถานการณ์จำลองกับเนื้อหาจำกัดบนจำนวนข้อสอบที่วัดในแต่ละคุณลักษณะ ผลการวิจัยของการทดลองใน 2 สถานการณ์ แสดงให้เห็นว่า 1) วิธี Monte Carlo เติมเต็มสิ่งสำคัญของการทดสอบและให้ผลเกี่ยวกับความแม่นยำของการวัดและผลของการเปิดเผยข้อสอบเป็นที่น่าพอใจ และ 2) วิธี Monte Carlo ให้ผลที่ดีกว่าวิธี MMGDI เมื่อวิธี Monte Carlo ถูกใช้กับลำดับขั้นตอนแบบ Kullback-Leibler ที่ให้น้ำหนักภายหลัง หรือใช้กับการผสมสารสนเทศแบบ Kullback-Leibler ที่ดัชนีการคัดเลือกข้อสอบ โดยภาพรวมอัตราการฟื้นคืนกลับของความรู้ การแจกแจงของการเปิดเผยข้อสอบ และอัตราการใช้ประโยชน์ของคลังข้อสอบถูกพัฒนาปรับปรุงเมื่อวิธี Monte Carlo ถูกนำมาใช้

He, Diao and Hauser (2014) ได้ศึกษาเรื่องการเปรียบเทียบวิธีการคัดเลือกข้อสอบ 4 วิธีสำหรับข้อจำกัดที่รุนแรงของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ การศึกษาครั้งนี้เป็นการเปรียบเทียบวิธีการคัดเลือกข้อสอบ 4 วิธี ที่พัฒนาขึ้นเพื่อใช้กับข้อจำกัดอย่างรุนแรงในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CATs) ข้อจำกัดที่รุนแรงของ CATs อ้างถึงการทดสอบแบบปรับเหมาะที่พยายามจะตอบสนองต่อชุดที่ซับซ้อนของข้อจำกัดที่มักจะไม่ได้ข้อสรุปกัน เช่น)ข้อสอบอาจนำไปสู่ความพึงพอใจของข้อจำกัดหลายอย่างในเวลาเดียวกัน (วิธีการตรวจสอบในการศึกษานี้จะรวมถึงโมเดลการเบี่ยงเบนถ่วงน้ำหนัก (WDM) โมเดลการลงโทษถ่วงน้ำหนัก (WPM) , ดัชนีความสำคัญสูงสุด (WPI) และวิธีการทดสอบเสมือน (STA) นอกจากนี้วิธีการที่ปรับปรุงแก้ไข 2 วิธีของวิธี MPI ถูกนำไปจัดการกับเงื่อนไขกรณีที่ผลในขั้นตอนการคัดเลือกข้อสอบกลายเป็นความผิดปกติในระหว่างการทดสอบ ผลที่ได้ชี้ให้เห็นว่า STA ทำงานได้ดีที่สุดในวิธีการทั้งหมดในแง่ของความถูกต้องในการวัดและการจัดการข้อจำกัดอื่น ๆ สำหรับอีก 3 วิธี พบว่าไม่แตกต่างกันอย่างมีนัยสำคัญในด้านความถูกต้องในการวัดและการจัดการในระดับที่จำกัดที่ขอบเขตล่าง แต่วิธี WPM กลับมีการดำเนินการที่ดีขึ้นอย่างมากในการบริหารจัดการข้อจำกัด โดยรวมกว่าวิธี WDM หรือวิธี MPI

Cheng, Patton and Shao (2014) ได้ศึกษาเรื่อง วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ตามระดับขั้นของค่าอำนาจจำแนกแบบแบ่งชั้น (a-stratified) ที่มีอยู่ในความคลาดเคลื่อนของการทดสอบความถูกต้อง แบบทดสอบปรับเหมาะด้วยคอมพิวเตอร์วิธีตามระดับขั้นของค่าอำนาจจำแนกและค่าความยาก b-blocking (AST) เป็นทางเลือกที่ใช้กันอย่างแพร่หลายของวิธีการคัดเลือกข้อสอบของสารสนเทศสูงสุดของฟิชเชอร์ (MFI) ที่มีประสิทธิภาพสามารถรักษาความสมดุลของการใช้คลังข้อสอบให้

ประมาณค่าคุณลักษณะแฝงที่ถูกต้องในการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ (CAT)อย่างไรก็ตามก่อนหน้านี้มีการเปรียบเทียบวิธีการนี้ในการประมาณค่าพารามิเตอร์ของข้อสอบกับค่าพารามิเตอร์ของประชากรที่แท้จริง ในบทความนี้เสนอประโยชน์ของการตรวจสอบผลการดำเนินงานของวิธี AST ภายใต้เงื่อนไขที่เป็นจริงมากขึ้นในการประมาณค่าพารามิเตอร์ของข้อสอบแทนค่าพารามิเตอร์ที่แท้จริงที่ใช้ในการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ ประสิทธิภาพการทำงานเมื่อเปรียบเทียบกับวิธีการ MFI เมื่อถูกนำมาใช้ร่วมกับ Sympson-Hetter หรือการควบคุมการสุ่มอย่างเปิดเผย ผลการศึกษาพบว่าวิธีการ MFI เมื่อรวมกับการควบคุมการเปิดเผยไวต่อโอกาสที่เกิดขึ้น นความจริงโดยเฉพาะอย่างยิ่งเมื่อการทดสอบความถูกต้องของกลุ่มตัวอย่างที่มีขนาดเล็ก แต่ในทางกลับกัน, AST มีประสิทธิภาพมากขึ้นต่อโอกาสที่เกิดขึ้น สอดคล้องกับการตรวจสอบก่อนหน้านี้ที่ใช้ค่าพารามิเตอร์ข้อสอบจริง AST ให้ผลที่สมดุลในคลังข้อสอบมากขึ้นกับการเสียความแม่นยำขนาดเล็กของ ประมาณค่าคุณลักษณะแฝง การเสียเพียงเล็กน้อยเมื่อมีการทดสอบข้อสอบ 40 ข้อ

สรุปได้ว่าสิ่งที่สำคัญที่สุดในแบบทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติและแบบพหุมิติคือข้อสอบในคลังข้อสอบเพราะ ข้อสอบที่สร้างขึ้นนั้นเป็นสิ่งที่มุ่งวัดตัวแปรคุณลักษณะแฝงหรือระดับความสามารถของผู้สอบ ลักษณะของข้อสอบในคลังข้อสอบเป็นส่วนสำคัญของการทดสอบ ลักษณะการทดสอบปรับเหมาะด้วยคอมพิวเตอร์นั้นนิยมใช้ข้อสอบแบบเลือกตอบ (Multiple choice) ข้อสอบสอบแบบเลือกตอบที่ใช้ในปัจจุบันมีทั้งแบบข้อสอบเดี่ยว (Single item) และกลุ่มข้อสอบที่ใช้สิ่งเร้าร่วมกัน (stimulus) ซึ่งข้อสอบลักษณะนี้จะเรียกว่าแบบสอบฉบับย่อย (Testlet) หรือชุดข้อสอบ (Item Set) และเป็นการนำการให้คะแนนข้อสอบที่มีลักษณะเป็นแบบทดสอบย่อย ข้อสอบในคลังข้อสอบของแบบทดสอบปรับเหมาะด้วยคอมพิวเตอร์จะมีจุดมุ่งหมายแตกต่างกัน ถ้าคลังข้อสอบในแบบทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติจะใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ซึ่งมีข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกมิติ (Unidimension) ข้อสอบแต่ละข้อในแบบสอบนั้นมุ่งวัดความสามารถหรือคุณลักษณะเดียว หรือ มิติเดียว ส่วนแบบทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิตินั้นใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (MIRT) ที่ข้อสอบมุ่งวัดคุณลักษณะแฝงหรือความสามารถของบุคคลมากกว่า 1 องค์ประกอบ หรือ หลายมิติ ดังนั้นการสร้างข้อสอบในคลังข้อสอบเพื่อวัดตัวแปรคุณลักษณะแฝงหรือระดับความสามารถของผู้สอบมีความสำคัญที่ผู้สร้างต้องระบุจุดมุ่งของวัดให้ชัดเจนว่าเป็นแบบมิติเดียวหรือหลายมิติ เพราะข้อสอบในคลังจะส่งผลต่อวิธีการประมาณค่าความสามารถของผู้สอบ วิธีการคัดเลือกข้อสอบ การควบคุมการใช้ข้อสอบซ้ำและกฎการยุติที่แตกต่างกัน

3.4 การประยุกต์ใช้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ปัจจุบันมีการนำแนวคิดการทดสอบแบบปรับเหมาะไปใช้กันอย่างแพร่หลาย โดยสามารถแบ่งกลุ่มการประยุกต์ใช้การทดสอบแบบปรับเหมาะออกเป็น 2 กลุ่ม ได้แก่ การวัดด้านพุทธิพิสัย (cognitive) และการวัดด้านจิตพิสัย (affective) โดยมีรายละเอียดดังนี้

1) การวัดด้านพุทธิพิสัย

ในการวัดด้านพุทธิพิสัยมีการนำการทดสอบแบบปรับเหมาะมาใช้กันอย่างแพร่หลาย ซึ่งตัวอย่างที่เป็นสากลในการนำการทดสอบแบบปรับเหมาะมาใช้ในการทดสอบได้แก่ การทดสอบเพื่อศึกษาต่อในระดับ

บัณฑิตศึกษาและด้านการบริหาร (Graduate Record Examinations: GRE) โดยในการทดสอบจะแบ่งการวัดความรู้ออกเป็น 3 ด้าน ได้แก่ ความสามารถทางภาษาอังกฤษ (Verbal Reasoning) ความสามารถทางคณิตศาสตร์ (Quantitative Reasoning) และความสามารถด้านการเขียนวิเคราะห์ (Analytical Writing)

ในการทดสอบจะใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มาช่วยในการคัดเลือกข้อสอบสำหรับผู้เข้ารับการทดสอบ ซึ่งด้านความสามารถทางภาษาอังกฤษและความสามารถทางคณิตศาสตร์จะมีกระบวนการคัดเลือกข้อสอบที่เหมือนกัน โดยใช้คอมพิวเตอร์ในการเลือกข้อสอบในตอน (section) ถัดไปให้กับผู้เข้ารับการทดสอบ โดยอิงจากความสามารถของผู้สอบในตอนก่อนหน้า สำหรับด้านความสามารถด้านการเขียนวิเคราะห์จะดำเนินการคัดเลือกแบบทดสอบซึ่งเป็นความเรียง (essay) ด้วยระบบคอมพิวเตอร์และทำการตรวจโดยระบบ e-rater (Educational Testing Service, 2016)

2) การวัดด้านจิตพิสัย (affective)

สำหรับการวัดทางจิตพิสัยมีการนำการทดสอบแบบปรับเหมาะไปปรับใช้ในหลายรูปแบบ ดังเช่นทางการแพทย์ ได้มีการนำการทดสอบแบบปรับเหมาะไปปรับใช้ในการวัดการรับรู้ของผู้ป่วยในโรงพยาบาล โดยได้มีการศึกษาเปรียบเทียบกับวิธีการวัดการรับรู้แบบดั้งเดิมที่ไม่ได้มีการปรับข้อคำถามตามระดับความสามารถของผู้ตอบแบบทดสอบ พบว่าการวัดการรับรู้โดยใช้การทดสอบแบบปรับเหมาะใช้จำนวนข้อคำถามน้อยกว่าการวัดแบบดั้งเดิม โดยวิธีการวัดด้วยการทดสอบแบบปรับเหมาะมีประสิทธิภาพมากกว่าวิธีการวัดแบบดั้งเดิมร้อยละ 42 ซึ่งเป็นเป้าหมายหลักของการทดสอบแบบปรับเหมาะคือ ช่วยลดจำนวนข้อคำถามลงโดยที่ไม่ทำให้สูญเสียความแม่นยำ (Chien, Wang, Huang, Lai, & Chow, 2011; Zheng & Chang, 2015)

ตอนที่ 4 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องสามารถแบ่งเนื้อหามโนทัศน์เกี่ยวกับการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ออกเป็น 3 ประเด็น ได้แก่ 4.1 ขั้นตอนการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 4.2 การตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 4.3 งานวิจัยที่เกี่ยวข้องกับการพัฒนาระบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ โดยมีรายละเอียดดังนี้

4.1 ขั้นตอนการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบ่งออกเป็น 6 ขั้นตอน ได้แก่ การวางแผน การพัฒนาคลังข้อสอบ การทดลองใช้และวิเคราะห์ข้อสอบ การกำหนดรูปแบบระบบการทดสอบ การสร้างระบบการทดสอบ และการดูแลระบบการทดสอบ (Birdsall, 2011; Thompson & Weiss, 2011) แต่ละขั้นตอนมีรายละเอียดดังต่อไปนี้

1. การวางแผน

ขั้นตอนแรกของการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ คือ การวางแผนการพัฒนาระบบการทดสอบโดยพิจารณาเกี่ยวกับโปรแกรมที่ใช้สร้างระบบการทดสอบ วางแผนเพื่อกำหนดประเภทของแบบสอบ ความยาวของแบบสอบ ความแม่นยำของคะแนนสอบ การเปิดเผยข้อสอบ ขนาดคลังข้อสอบ คุณสมบัติทางจิตมิติของข้อสอบ อีกทั้งวางแผนการทดลองใช้ข้อสอบและระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ รวมทั้งพิจารณาถึงความคุ้มค่าในการดำเนินงาน

2. การพัฒนาคลังข้อสอบ

การพัฒนาคลังข้อสอบเป็นขั้นตอนในการเขียนข้อสอบ และทบทวนข้อสอบตามผังการออกข้อสอบที่กำหนดไว้ ซึ่งปัจจัยที่ควรนำมาพิจารณาเพื่อกำหนดขนาดคลังข้อสอบ คือ การกระจายของค่าพารามิเตอร์ของข้อสอบ การกระจายของเนื้อหา และการเปิดเผยข้อสอบ โดยข้อสอบในคลังข้อสอบจะต้องมีจำนวนที่เพียงพอสำหรับทุกระดับความยาก และมีการกระจายเนื้อหาที่เหมาะสม

3. การทดลองใช้และวิเคราะห์ข้อสอบ

หลังจากการเขียนข้อสอบ ขั้นตอนต่อไป คือ การนำข้อสอบไปทดลองใช้ และวิเคราะห์คุณภาพของข้อสอบรายข้อโดยเลือกใช้โมเดลการตอบสนองข้อสอบที่เหมาะสมกับการวิจัย ซึ่งค่าพารามิเตอร์ที่นิยมวิเคราะห์ได้แก่ ความยาก และอำนาจจำแนก เพื่อเป็นข้อมูลในพิจารณาตัดข้อสอบที่ไม่มีคุณภาพทิ้ง ปรับปรุงแก้ไขข้อสอบหรือนำข้อสอบไปทดลองใช้อีกครั้ง รวมทั้งพิจารณาเกี่ยวกับค่าความสอดคล้องของโมเดลการตอบสนองข้อสอบที่เลือกใช้

4. การกำหนดรูปแบบระบบการทดสอบ

ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบ่งเป็น 5 องค์ประกอบ ได้แก่ การพัฒนาคลังข้อสอบ การเริ่มต้นการทดสอบ อัลกอริทึมคัดเลือกข้อสอบ อัลกอริทึมประมาณค่าความสามารถ และเกณฑ์การยุติการทดสอบ โดยขั้นตอนการพัฒนาคลังข้อสอบ และการทดลองใช้และวิเคราะห์ข้อสอบเป็นเพียงองค์ประกอบแรกของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ขั้นตอนต่อมาจึงต้องวางแผนการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์อีก 4 องค์ประกอบ ดังนี้

4.1 การเริ่มต้นการทดสอบ เป็นการคัดเลือกข้อสอบข้อแรกให้ผู้สอบตอบคำถามซึ่งแบ่งออกเป็น 3 ลักษณะ ได้แก่ (1) คัดเลือกข้อสอบที่มีค่าความยากเท่ากับ 0.0 เนื่องจากเป็นข้อสอบที่มีความยากปานกลางและเป็นค่ากึ่งกลางของค่าความยากของข้อสอบ นั่นคือผู้สอบทุกคนจะได้รับข้อสอบข้อแรกเหมือนกัน ซึ่งอาจจะส่งผลต่อความปลอดภัยของข้อสอบหรือการเปิดเผยข้อสอบ (2) สุ่มข้อสอบที่อยู่ในช่วงความยากปานกลาง เช่น ข้อสอบที่มีความยากอยู่ในช่วง -0.5 ถึง 0.5 ซึ่งมีข้อสอบเริ่มต้นที่หลากหลาย และ (3) กำหนดข้อสอบเริ่มต้นจากผลการสอบที่ผ่านมา เช่น พิจารณาจากคะแนนเฉลี่ยสะสม และระดับความสามารถจากการทดสอบด้วยระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ผ่านมา

4.2 อัลกอริทึมคัดเลือกข้อสอบ เป็นอัลกอริทึมที่คัดเลือกข้อสอบที่เหมาะสมที่สุดให้ผู้สอบตอบคำถาม แบ่งตามเป้าหมายของแบบสอบ 2 กรณี ได้แก่ (1) มุ่งเน้นความแม่นยำในการประมาณค่าความสามารถของผู้สอบ อัลกอริทึมจะคัดเลือกข้อสอบที่ให้สารสนเทศที่มากที่สุดสำหรับการประมาณค่าความสามารถของผู้สอบแต่ละคน และ (2) มุ่งเน้นจัดกลุ่มผู้สอบบนพื้นฐานของคะแนนจุดตัด ควรใช้วิธีอัตราส่วนความเป็นไปได้ (likelihood ratio approach) ในการออกแบบอัลกอริทึมคัดเลือกข้อสอบ เพื่อประเมินสารสนเทศที่คะแนนจุดตัด อย่างไรก็ตามในการออกแบบอัลกอริทึมคัดเลือกข้อสอบควรพิจารณาเกี่ยวกับการเปิดเผยข้อสอบ ความครอบคลุมของเนื้อหา และพฤติกรรมการเรียนรู้ที่ข้อสอบมุ่งวัด

4.3 อัลกอริทึมประมาณค่าความสามารถ การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะต้องเลือกใช้อัลกอริทึมประมาณค่าความสามารถที่มีประสิทธิภาพ และเหมาะสมกับแบบสอบที่สร้างขึ้น โดยอัลกอริทึมประมาณค่าความสามารถที่มีประสิทธิภาพ ได้แก่ วิธีความเป็นไปได้สูงสุด (maximum likelihood) และวิธีของเบย์ (Bayesian methods)

4.4 เกณฑ์การยุติการทดสอบ เกณฑ์ในการยุติการทดสอบแบ่งเป็น 4 ลักษณะ ได้แก่ (1) การยุติโดยพิจารณาจากการประมาณค่าความสามารถของผู้สอบ ซึ่งจะยุติการทดสอบเมื่อค่าระดับความสามารถเปลี่ยนแปลงน้อยกว่าค่าที่กำหนดไว้ (2) การยุติโดยพิจารณาจากความคลาดเคลื่อนมาตรฐานในการวัด ซึ่งจะยุติการทดสอบเมื่อความคลาดเคลื่อนมาตรฐานในการวัดมีค่าต่ำกว่าที่กำหนดไว้ วิธีการนี้เป็นวิธีที่นิยมใช้มากที่สุด (3) การยุติโดยพิจารณาจากคลังข้อสอบ ซึ่งจะยุติการทดสอบเมื่อข้อสอบที่เหลือในคลังข้อสอบให้สารสนเทศน้อยกว่าที่กำหนดไว้ และ (4) การยุติโดยกำหนดจำนวนข้อสอบและเวลาในการทดสอบ ซึ่งจะยุติการทดสอบเมื่อจำนวนข้อสอบ หรือเวลาในการทดสอบถึงค่าที่กำหนดไว้

5. การสร้างระบบการทดสอบ

การสร้างระบบการทดสอบเป็นขั้นตอนที่รวมทุกองค์ประกอบเพื่อเขียนโปรแกรมระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งจะต้องคำนึงถึงความเป็นไปได้ในการนำไปใช้ การดำเนินการสอบ และความปลอดภัยของระบบการทดสอบ อีกทั้งจะต้องออกแบบรายงานผลการทดสอบที่ให้สารสนเทศที่สำคัญแก่ผู้สอบ เช่น จำนวนคำตอบที่ถูกและจำนวนคำตอบที่ผิด นอกจากนี้ควรนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่สร้างขึ้นไปทดลองใช้เพื่อปรับปรุงแก้ไขข้อผิดพลาดที่อาจจะเกิดขึ้น

6. การดูแลระบบการทดสอบ

หลังจากการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ไปใช้ทดสอบกับผู้สอบระยะหนึ่งแล้ว ควรมีการตรวจสอบระบบการทดสอบเพื่อยืนยันถึงความตรงและความเที่ยงของระบบการทดสอบ รวมทั้งพิจารณาเกี่ยวกับการประมวลผลและการรายงานผล เพื่อปรับปรุงระบบการทดสอบให้มีประสิทธิภาพเช่นเดิม นอกจากนี้ควรมีการปรับปรุงคลังข้อสอบเพื่อสร้างข้อสอบเข้ามาทดแทนข้อสอบเดิมที่มีการเปิดเผยมาก

กล่าวโดยสรุป ขั้นตอนการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบ่งออกเป็น 2 ส่วน ได้แก่ การพัฒนาคลังข้อสอบ ซึ่งมีขั้นตอนการดำเนินงานดังหัวข้อ 2.3 และการออกแบบระบบการทดสอบ ซึ่งการวิจัยนี้ ผู้วิจัยออกแบบและพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์หลังจากได้ข้อสอบที่มีคุณภาพแล้ว โดยเขียนโปรแกรมระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบ่งออกเป็น 4 ส่วน ได้แก่ 1) การลงทะเบียน 2) การสร้างชุดข้อสอบ 3) การทดสอบ และ 4) การรายงานผลการทดสอบ ดังมีรายละเอียดเสนอในบทที่ 4 ตอนที่ 3 ต่อไป

4.2 การตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

สิ่งสำคัญสำหรับการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ คือ การประเมินผลการใช้งานระบบ โดยมีแนวคิดในการประเมินผลการใช้งานระบบการทดสอบหลากหลายวิธี ซึ่งวิธีที่นิยมและมีการนำมาใช้อย่างแพร่หลาย คือ การประเมินระบบแบบอิงมาตรฐาน (standard evaluation) และการประเมินความพึงพอใจระหว่างอินเตอร์เฟสกับผู้ใช้งาน (user satisfaction of the human-computer interface) โดยมีรายละเอียดดังนี้

1. การประเมินระบบแบบอิงมาตรฐาน (standard evaluation)

การประเมินระบบแบบอิงมาตรฐานมีพื้นฐานมาจากแนวคิดของ Stufflebeam ประกอบด้วยเนื้อหาในการประเมิน 4 องค์ประกอบ ได้แก่ การใช้ประโยชน์ (utility) ความเป็นไปได้ (feasibility) ความเหมาะสม

(propriety) และความถูกต้อง (accuracy) ซึ่งแต่ละองค์ประกอบมีรายละเอียดดังนี้ (ศิริชัย กาญจนวาสี, 2555; Economides & Roupas, 2009; Georgiadou, Triantafillou, & Economides, 2007)

1.1 การใช้ประโยชน์ (utility) คือ การประกันความเป็นประโยชน์ของระบบเพื่อตอบสนองต่อความต้องการในการใช้สารสนเทศของผู้ที่เกี่ยวข้องให้มีความครอบคลุม และทันเวลา กล่าวคือ ข้อสอบในคลังข้อสอบจะต้องมีค่าความยากอยู่ในช่วงกว้างและให้สารสนเทศแก่ผู้สอบทุกระดับความสามารถ นอกจากนี้ระบบการทดสอบจะต้องให้ประโยชน์ในการระบุจุดแข็ง และจุดอ่อนของผู้สอบ ระบุระดับความสามารถในแต่ละทักษะ รวมทั้งระบุพฤติกรรมการเรียนรู้ที่ผู้สอบสามารถบรรลุผล และไม่สามารถบรรลุผล

1.2 ความเป็นไปได้ (feasibility) คือ การประกันความสอดคล้องของระบบกับสภาพความเป็นจริงที่มีความเหมาะสมกับสถานการณ์ สามารถยอมรับได้ พร้อมทั้งประหยัดและคุ้มค่า โดยระบบการทดสอบจะต้องมีการแสดงผลที่เข้าใจง่าย ทำให้ผู้สอบแต่ละระดับความสามารถเข้าใจตรงกัน และดึงดูดความสนใจของผู้สอบแต่ละคน

1.3 ความเหมาะสม (propriety) คือ การประกันว่าระบบมีความเหมาะสมโดยเป็นไปตามกฎระเบียบ จรรยาบรรณ และคำนึงถึงสวัสดิภาพของผู้ที่เกี่ยวข้องและได้รับผลกระทบ ทั้งนี้ระบบการทดสอบจะต้องมีระบบรักษาความปลอดภัยที่ป้องกันไม่ให้บุคคลอื่นเข้าถึงข้อมูลเพื่อนำข้อมูลการทดสอบไปใช้ประโยชน์ในทางมิชอบ หรือการแก้ไขข้อมูลที่ทำให้ระบบการทดสอบเกิดความเสียหาย

1.4 ความถูกต้อง (accuracy) คือ การประกันว่าระบบมีการใช้เทคนิคที่มีความเหมาะสม และสามารถให้สารสนเทศที่เพียงพอต่อการใช้งาน อีกทั้งระบบการทดสอบจะต้องมีระบบการป้องกันความผิดพลาดของผู้สอบที่อาจจะเกิดขึ้นโดยไม่ได้ตั้งใจ เช่น มีข้อความแจ้งเตือนทุกครั้งที่มีการใช้งานเครื่องมือต่างๆ ในโปรแกรมการทดสอบ

2. การประเมินความพึงพอใจระหว่างอินเตอร์เฟสกับผู้ใช้งาน (user satisfaction of the human-computer interface)

การประเมินความพึงพอใจระหว่างอินเตอร์เฟสกับผู้ใช้งานเป็นสิ่งสำคัญในการพัฒนาระบบ เนื่องจากหากระบบสามารถตอบสนองการใช้งานจนสามารถทำให้ผู้ใช้งานเกิดความพึงพอใจ จะส่งผลให้ระบบได้รับการสนับสนุนและผลักดันในการนำไปใช้งานจริง ดังนั้นจึงมีผู้พัฒนาเครื่องมือการประเมินความพึงพอใจระหว่างอินเตอร์เฟสกับผู้ใช้งานขึ้นสำหรับประเมินความพึงพอใจของอินเตอร์เฟสกับผู้ใช้งาน โดยการประเมินความพึงพอใจระหว่างอินเตอร์เฟสกับผู้ใช้งานแบ่งเป็น 5 องค์ประกอบ ดังรายละเอียดต่อไปนี้ (ณภัทร ชัยมงคล, โชติกา ภาษีผล และศิริชัย กาญจนวาสี, 2558; Chin, Diehl & Norman 1988; Navas et al., 2007)

2.1 ภาพรวมปฏิสัมพันธ์ของระบบ (overall reaction to the software) คือ ความพึงพอใจในภาพรวมของผู้ใช้งานสำหรับการใช้งานระบบ โดยความพึงพอใจของผู้ใช้งานขึ้นอยู่กับการใช้คำศัพท์ ข้อมูลย้อนกลับจากระบบ การให้ความช่วยเหลือโดยระบบ และการออกแบบหน้าจอ

2.2 หน้าจอ (screen) คือ การที่ผู้ใช้งานสามารถอ่านรายละเอียดบนหน้าจอของระบบได้ง่ายชัดเจน รวมถึงความสวยงามของรูปแบบในการจัดวางและการใช้ประโยชน์ ผู้สอบสามารถเรียนรู้เมนูการใช้งานบน

หน้าจอได้ด้วยตัวเอง นอกจากนี้หน้าจอจะต้องกระตุ้นความสนใจของผู้สอบ และดึงดูดความสนใจของผู้สอบให้สนใจอยู่กับการทดสอบตลอดเวลา

2.3 การใช้คำศัพท์และสารสนเทศระบบ (terminology and system information) คือ การที่ผู้ใช้งานสามารถพบข้อความคำศัพท์ที่มีความสอดคล้องและเหมาะสมเพื่อแจ้งเตือนบนหน้าจอของระบบได้อย่างชัดเจน พร้อมทั้งมีระบบที่สามารถแจ้งเตือนความผิดพลาดซึ่งเกิดประโยชน์ต่อการใช้งานระบบ

2.4 การเรียนรู้ (learning) คือ การที่ผู้ใช้งานเกิดความเข้าใจว่าระบบง่ายต่อการเรียนรู้ ไม่ยุ่งยาก ซับซ้อน รายละเอียดและคำสั่งง่ายต่อการจดจำ

2.5 ความสามารถของระบบ (system capabilities) คือ การที่ผู้ใช้งานสามารถใช้งานระบบได้ง่าย พร้อมทั้งสามารถแก้ไขข้อผิดพลาดของระบบได้อย่างรวดเร็ว รวมถึงมีความสัมพันธ์กับระดับประสบการณ์ของผู้ใช้งาน นอกจากนี้ระบบจะต้องมีเครื่องมือที่สนับสนุนการใช้งานที่หลากหลายและมีประสิทธิภาพ

4.3 งานวิจัยที่เกี่ยวข้องกับการพัฒนาระบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์

สุชาดา กรเพชรปาณี, เสรี ชัดแช้ม, ปิยะทิพย์ ตินวร, และ โสพส สุขานนท์สวัสดิ์ (2557) ได้ศึกษาการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สาหรับการจัดสอบ O-NET สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 ระดับชั้นละ 8 กลุ่มสาระการเรียนรู้ จำนวน 24 กลุ่มสาระการเรียนรู้ วิธีดำเนินการวิจัยแบ่งเป็น 3 ขั้นตอน ดังนี้ ขั้นตอนที่หนึ่ง เป็นการจัดทำคลังข้อสอบโดยใช้ฐานข้อมูล MySQL และคัดเลือกข้อสอบ O-NET ของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ระหว่างปี พ.ศ. 2551-2553 ที่ผ่านการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ โดยใช้โปรแกรมสำเร็จรูป Xcalibre Version 4.1.7 ที่มีคุณภาพตามเกณฑ์บรรจุในคลังข้อสอบ ขั้นตอนที่สอง เป็นการพัฒนาโปรแกรมในรูปแบบของ Web Application โดยใช้ภาษา PHP พัฒนาโปรแกรม ใช้ MySQL เป็นฐานข้อมูลและใช้ภาษา SQL เป็นคำสั่งเชื่อมโยงข้อมูลในฐานข้อมูล พัฒนาโปรแกรมคอมพิวเตอร์ตามหลักการของวงจรการพัฒนาซอฟแวร์ (Software Development Life Circle: SDLC) (Elliott, 2004) และขั้นตอนการดำเนินการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ของ Thompson and Weiss (2011) ที่ใช้กับทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ และขั้นตอนที่สาม เป็นการประเมินประสิทธิภาพของโปรแกรม โดยให้นักเรียนที่ศึกษาในภาคเรียนที่ 1 ปีการศึกษา 2557 จำนวน 1,091 คน จำแนกเป็นระดับชั้นประถมศึกษาปีที่ 6 จำนวน 224 คน ระดับชั้นมัธยมศึกษาปีที่ 3 จานวน 432 คน และระดับชั้นมัธยมศึกษาปีที่ 6 จำนวน 435 คน ทำการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สาหรับการจัดสอบ O-NET ใน website://www.onetcat.net วิเคราะห์ความพึงพอใจของนักเรียนที่ใช้โปรแกรมด้วยสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน ผลการวิจัย มีดังนี้ 1) คลังข้อสอบ O-NET สามารถบรรจุข้อสอบแบบหลายตัวเลือก (ไม่เกิน 5 ตัวเลือก) โดยได้บรรจุข้อสอบ O-NET ที่วิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ และผ่านเกณฑ์การคัดเลือกข้อสอบ ในระดับชั้นประถมศึกษาปีที่ 6 จำนวน 258 ข้อ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 469 ข้อ และระดับชั้นมัธยมศึกษาปีที่ 6 จำนวน 1,197 ข้อ 2) โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET เป็นโปรแกรมการจัดการทดสอบในรูปแบบของ Web Application ใน

website://www.onetcat.net มีคุณภาพอยู่ในเกณฑ์ดี และด้านการวัดและประเมินผลการศึกษา ทั้งในด้านความสะดวกในการใช้งาน ความถูกต้องในการใช้งาน ลักษณะทั่วไปของโปรแกรมและความชัดเจนของคู่มือการใช้โปรแกรม 3) ประสิทธิภาพในการใช้งานของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET อยู่ในเกณฑ์ดี เป็นที่พึงพอใจของกลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 ที่ใช้โปรแกรมในด้านรูปแบบสวยงาม และความสะดวกในการใช้งาน

ญานิศรา มุนินทร์สาคร (2558) ได้ศึกษาการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สาหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 การวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์คุณภาพของข้อสอบ O-NET จัดทำคลังข้อสอบ O-NET และพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 จำนวน 8 กลุ่มสาระการเรียนรู้ วิธีดำเนินการวิจัยแบ่งเป็น4 ขั้นตอน ดังนี้ 1) วิเคราะห์คุณภาพของข้อสอบ O-NET จำนวน 8 กลุ่มสาระการเรียนรู้ 2) จัดทำคลังข้อสอบ O-NET จำนวน 8 กลุ่มสาระการเรียนรู้ 3) พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และ 4) ประเมินความคิดเห็นของผู้ทดลองใช้โปแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งเป็นนักเรียนระดับชั้นประถมศึกษาปีที่ 6 จังหวัดชลบุรี จำนวน 30 คน ผลการวิจัยมีดังนี้ 1) ข้อสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 มีค่าความยากของข้อสอบ (b) ในระดับค่อนข้างยาก 2) คลังข้อสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 บรรจุข้อสอบแบบหลายตัวเลือก (Multiple Choice) ชนิด 4 ตัวเลือก ได้โดยไม่จำกัด ขึ้นอยู่กับขนาดของ Server ซึ่งมีข้อสอบจำนวน 258 ข้อ ที่ผ่านเกณฑ์การวิเคราะห์คุณภาพของข้อสอบตามโมเดลโลจิส แบบ 3 พารามิเตอร์ใน 8 กลุ่มสาระการเรียนรู้ 3) โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 มีความเหมาะสมของโปรแกรมในระดับมากที่สุด ไม่มีปัญหาด้านการนำไปใช้ และเป็นที่ยอมรับของที่ผู้ทดลองใช้โปรแกรม

ดิเรก หอมจันทร์ (2554) ได้ศึกษาเรื่อง การพัฒนาแบบทดสอบปรับเหมาะกับความสามารถของผู้สอบสำหรับการเรียนการสอนแบบอีเลิร์นนิ่ง (E-Learning) รายวิชา 4000107 : เทคโนโลยีสารสนเทศเพื่อชีวิต สำหรับนักศึกษาชั้นปีที่ 1 มหาวิทยาลัยราชภัฏสุรินทร์ เพื่อช่วยในการจัดการเรียนการสอนให้มีความสะดวก และสามารถเรียนรู้ได้ด้วยตัวเองตลอดเวลาจากการพัฒนาแบบทดสอบปรับเหมาะกับความสามารถของผู้สอบสำหรับการเรียนการสอนแบบ อีเลิร์นนิ่ง (E-Learning) รายวิชา 4000107 : เทคโนโลยีสารสนเทศเพื่อชีวิต ได้มีการทดลองใช้งานและทดสอบประสิทธิภาพของโปรแกรมโดยใช้แบบประเมิน จากกลุ่มตัวอย่าง 2 กลุ่ม ได้แก่ กลุ่มผู้เชี่ยวชาญที่มีความรู้ทางด้านการสอนวิชาเทคโนโลยีสารสนเทศเพื่อชีวิต และวิชาที่เกี่ยวข้องกับคอมพิวเตอร์ 5 คน ในการประเมินคุณภาพของระบบ และกลุ่มผู้ใช้งานจำนวน 50 คน ในการประเมินความพึงพอใจของระบบ โดยผลจากการวิเคราะห์ข้อมูลของระบบ พบว่าผู้เชี่ยวชาญประเมินคุณภาพของระบบ อยู่ในระดับดี เนื่องจากระบบสามารถประมวลผลและแสดงผลลัพธ์ได้ตรงตามความต้องการที่กำหนดไว้ โดยได้ค่าเฉลี่ยเท่ากับ 4.26 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.63 ส่วนกลุ่มผู้ใช้มีความพึงพอใจในโปรแกรมอยู่ในระดับดี โดยได้ค่าเฉลี่ยเท่ากับ 4.46 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.63 ดังนั้น สามารถสรุปได้ว่า การพัฒนาแบบทดสอบปรับเหมาะกับความสามารถ

ของผู้สอบสำหรับการเรียนการสอนแบบอีเลิร์นนิ่ง (E-Learning) รายวิชา 4000107 : เทคโนโลยีสารสนเทศเพื่อชีวิต มีคุณภาพและความพึงพอใจอยู่ในระดับที่ดี สามารถนำไปใช้ประกอบการเรียนการสอนได้

ทิพย์ ขำอยู่, เสรี ชัดแช้ม, และ กฤษณะ ชินสาร (2556) ได้ศึกษาเรื่อง การวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยประยุกต์โมเดลลำดับขั้นคุณลักษณะ และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีวัตถุประสงค์เพื่อพัฒนาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษ โดยประยุกต์โมเดลลำดับขั้น คุณลักษณะและใช้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวินิจฉัยพื้นความรู้ทักษะการอ่าน (ขำอยู่, ชัดแช้ม, & ชินสาร, 2556) (ศิริเบญจา, ปราณี, และ สีเขียว, 2556) (สกลกิจรุ่งโรจน์, ชัดแช้ม, และ สุทัศนีย์, 2558)ภาษาอังกฤษของนักศึกษาระดับปริญญาตรี กลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏสวนดุสิต ปีการศึกษา 2555 จำนวน 1,763 คน เครื่องมือที่ใช้ในการวิจัย ได้แก่ ชุดข้อสอบวินิจฉัยทักษะการอ่านภาษาอังกฤษ และโปรแกรมการ ทดสอบวินิจฉัยแบบปรับเหมาะด้วยคอมพิวเตอร์ ผลการวิจัยพบว่า วิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษบนโปรแกรมการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ที่พัฒนาขึ้น สามารถใช้ประโยชน์จากแนวคิดโมเดลลำดับขั้นคุณลักษณะซึ่งเป็นแนวคิดของทฤษฎีทางปัญญา และทฤษฎีการตอบสนองข้อสอบสำหรับการทดสอบเพื่อวินิจฉัยพื้นความรู้ด้านทักษะการอ่านภาษาอังกฤษของนักศึกษา ระดับปริญญาตรีได้ ผลการประเมินโดยผู้เชี่ยวชาญและผู้ใช้งานมีความสอดคล้องกันคือ วิธีการวินิจฉัยที่พัฒนาขึ้นมีความ เหมาะสม สามารถนำไปใช้ประโยชน์สำหรับการทดสอบเพื่อวินิจฉัยพื้นความรู้ด้านทักษะการอ่านภาษาอังกฤษของ นักศึกษาได้

สุนันทา ศิริเบญจา, ไชยรัตน์ ปราณี, และ ดวงใจ สีเขียว (2556) ได้ศึกษา เรื่อง การพัฒนาแบบทดสอบปรับเหมาะโดยใช้คอมพิวเตอร์ สาระเทคโนโลยีสารสนเทศและการสื่อสาร โดยประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 พร้อมทั้งหาประสิทธิภาพของแบบทดสอบปรับเหมาะโดยใช้คอมพิวเตอร์ กลุ่มตัวอย่างที่ใช้ในการศึกษา คือ นักเรียนชั้นมัธยมศึกษา ปีที่ 3 ของโรงเรียนมัธยมศึกษาในอำเภอห้วยคต จังหวัดอุทัยธานี จำนวน 150 คน ได้มาโดยการสุ่มแบบกลุ่ม (Cluster Sampling) เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลเป็นแบบทดสอบ ปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 91 ข้อ และแบบสอบถามความพึงพอใจของนักเรียน จำนวน 20 ข้อผลการวิจัยพบว่า 1) แบบทดสอบปรับเหมาะโดยใช้คอมพิวเตอร์ มีคุณภาพตามทฤษฎีการตอบสนองต่อข้อสอบ (IRT) แบบ 3 พารามิเตอร์ คือ มีค่าอำนาจจำแนกอยู่ระหว่าง 0.583-1.786 ค่าความยากง่าย อยู่ระหว่าง -3.000-2.983 ค่าการเดา อยู่ระหว่าง 0.112-0.294 และมีความเป็นมิติเดียวของข้อสอบจากการตรวจสอบการวิเคราะห์องค์ประกอบ 2) แบบทดสอบปรับเหมาะโดยใช้คอมพิวเตอร์ โดยประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ มีประสิทธิภาพ โดยมีค่าเฉลี่ยฟังก์ชันสารสนเทศของข้อสอบ เท่ากับ 0.398 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ เท่ากับ 13.277 มีค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า อยู่ระหว่าง 0.280-0.299 ค่าความสามารถของผู้สอบรายบุคคล อยู่ระหว่าง -1.322-2.117 และความพึงพอใจของนักเรียน โดยรวมอยู่ในระดับมาก

สุชาดา สกลกิจรุ่งโรจน์, เสรี ชัดแช้ม, และ สมพร สุทัศนีย์ (2558) ได้ศึกษาการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย โดยใช้หลักการของวงจรการพัฒนาระบบเป็นแนวทางในการพัฒนาโปรแกรม คลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทยพัฒนาขึ้นตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งจัดทำขึ้นในงานวิจัยก่อนหน้านี้ ผู้วิจัยวิเคราะห์และออกแบบโปรแกรม ด้วยการสร้าง

แบบจำลองกระบวนการและแบบจำลองข้อมูลรวมทั้งกำหนดรายละเอียดองค์ประกอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ดังนี้ 1) คัดเลือกข้อคำถามเริ่มต้นที่ให้ค่าสารสนเทศสูงที่สุด ณ ตำแหน่งค่าคุณลักษณะแฝงเริ่มต้นเท่ากับ 0.0 2) ประมาณค่าคุณลักษณะแฝงของผู้ทดสอบ ด้วยวิธีการ Expected a Posteriori (EAP) 3) คัดเลือกข้อคำถามข้อถัดไป โดยใช้ผลการประมาณค่าคุณลักษณะแฝงที่ได้จากกระบวนการก่อนหน้ามาเป็นข้อมูลสำหรับการเลือกข้อคำถามข้อถัดไปที่ให้ค่าสารสนเทศสูงที่สุด ณ ตำแหน่งค่าคุณลักษณะแฝงที่ประมาณค่าได้ และ 4) ยุติการทดสอบ โดยกำหนดเกณฑ์การยุติการทดสอบเมื่อความคลาดเคลื่อนมาตรฐานของการประมาณค่าคุณลักษณะแฝงมีค่าต่ำกว่า 0.30 หรือข้อคำถามหมดคลัง โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย สร้างขึ้นในรูปแบบเว็บแอปพลิเคชัน ใช้ซอฟท์แวร์MySQL เป็นระบบในการจัดการฐานข้อมูล ใช้ภาษา PHPในการเขียนโปรแกรม แล้วจึงทดสอบระบบโดยใช้เทคนิคการทดสอบแบบกล่องดำ จากนั้นจึงจัดทำคู่มือการใช้งานแล้วนำโปรแกรมที่พัฒนาขึ้นไปทดลองใช้เพื่อประเมินผลโปรแกรมโดยผู้เชี่ยวชาญและผู้ใช้งานผลการวิจัยปรากฏว่า โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคน ไทย สามารถเข้าใช้งาน ได้ทุก ที่ทุก เวลาเมื่อ มีการเชื่อมต่ออินเทอร์เน็ตผ่านทางเว็บไซต์ http://www.thscat.com/Test การทดลองใช้และผลการประเมินความเหมาะสมของโปรแกรมโดยผู้เชี่ยวชาญและผู้ใช้งานแสดงให้เห็นว่า โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทยเป็นที่ยอมรับของผู้เชี่ยวชาญและผู้ใช้งานเป็นอย่างดี ซึ่งชี้ให้เห็นว่าโปรแกรมมีความเหมาะสมสำหรับการนำไปใช้เพื่อวัดความสุขของคนไทย

Gibbons et al. (2012) ได้ทำการศึกษาเรื่องการพัฒนาการทดสอบปรับเหมาะด้วยคอมพิวเตอร์สำหรับภาวะซึมเศร้า เพื่อเพิ่มความแม่นยำในการตรวจภาวะซึมเศร้าของแพทย์ การพัฒนาการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ใช้แบบวัดความภาวะซึมเศร้าของ Hamilton ซึ่งเป็นแบบวัดมาตราส่วนประมาณค่า จำนวน 24 ข้อ ผลการวิจัยพบว่าการวัดแบบดั้งเดิมเป็นการกำหนดจำนวนข้อคำถามที่ชัดเจนเพื่อให้การผลการวินิจฉัยที่แน่นอน แต่การทดสอบปรับเหมาะด้วยคอมพิวเตอร์ช่วยแก้ไขความไม่แน่นอนของการวินิจฉัยภาวะโรคซึมเศร้าและลดจำนวนข้อคำถามที่ใช้ในการวัด

Anatchkova, Saris-Baglama, Kosinski & Bjorner (2009) ได้ทำการศึกษาเรื่องการพัฒนาและการประเมินเบื้องต้นของการทดสอบด้วยคอมพิวเตอร์ในการประเมินอาการปวดเรื้อรัง โดยการเปรียบเทียบการทดสอบแบบดั้งเดิมและการทดสอบด้วยคอมพิวเตอร์ กลุ่มตัวอย่าง คือ ประชากรของสหรัฐอเมริกา จำนวน 782 คน เพื่อใช้กำหนดความรุนแรงของอาการปวด และนำมาทำการจำลองข้อมูลจริงเพื่อกำหนดเกณฑ์การยุติในการประเมิน ผลการวิจัยพบว่าการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ให้คะแนนในการประเมินที่แม่นยำและแพทย์สามารถนำระบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ไปปรับใช้ในการประเมินอาการปวดและรักษาคนไข้เรื้อรังได้

Susan et al. (2016) ได้ทำการศึกษา เรื่อง การพัฒนาและการตรวจสอบความตรงของการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับภาวะทางจิตจากเหตุการณ์รุนแรง (PTSD) เพื่อเพิ่มการประเมินและการลดการติดตามอาการภาวะทางจิตจากเหตุการณ์รุนแรง โดยการสัมภาษณ์ทหารผ่านศึกษา จำนวน 1,288 คน และการ

ทดสอบทหารผ่านศึกษาจำนวน 1,085 คน พบว่าการทดสอบปรับเหมาะด้วยคอมพิวเตอร์เป็นเครื่องมือที่สามารถประเมินภาวะทางจิตจากเหตุการณ์รุนแรงที่มีประสิทธิภาพและมีความแม่นยำ

จากการวิเคราะห์งานวิจัยที่เกี่ยวข้องกับการพัฒนาระบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ที่ผ่านมาพบว่ามีการพัฒนาระบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์เพื่อนำไปใช้ทางการศึกษาและการสาธารณสุข สำหรับการศึกษาพบว่ามีการพัฒนาระบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์เพื่อใช้ในการประเมินผลสัมฤทธิ์ทางการเรียนทั้งในระดับรายวิชา และระดับชาติ และใช้ในการการวินิจฉัยทักษะ เช่น ทักษะการอ่านภาษาอังกฤษ ส่วนการสาธารณสุขพบว่ามีการพัฒนาระบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์เพื่อใช้วัดคุณลักษณะและอาการของผู้ป่วย เช่น ความสุข ภาวะซึมเศร้า อาการปวดเรื้อรัง และภาวะทางจิตจากเหตุการณ์รุนแรง เป็นต้น

ตอนที่ 5 กรอบแนวคิดการวิจัย

กรอบแนวคิดการวิจัย เรื่อง ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี นี้ ประกอบด้วยตัวแปรหลายส่วนตามระยะการวิจัย เริ่มจากการพัฒนาข้อสอบตามแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องตามที่ได้สังเคราะห์ในตอนที่ 1 แล้วนำมาวิเคราะห์คุณภาพของข้อสอบโดยใช้ทฤษฎีที่น่าเชื่อถือเป็นที่ยอมรับเพื่อให้ได้ข้อสอบที่มีคุณภาพแล้วนำไปจัดไว้ในคลังข้อสอบ จากนั้นจึงดำเนินการพัฒนาและตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารฯ ดังนั้นตัวแปรที่ใช้ในงานวิจัยนี้จึงขอแยกเป็น 5 ส่วน ดังมีรายละเอียด ดังนี้

1. ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21ฯ
2. ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21ฯ
3. คู่มือการใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21ฯ
4. การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21ฯไปทดลองใช้
5. คุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารฯ

1. ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ฯ จากการสังเคราะห์องค์ประกอบทักษะเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เพื่อให้ได้องค์ประกอบที่เป็นตัวแทนที่ดี ลดความซ้ำซ้อนในการวัด รวมทั้งสามารถสร้างเครื่องมือที่มีการให้คะแนนแบบเลือกตอบ ที่นำไปใช้ในการทดสอบแบบปรับเหมาะ ที่ใช้ระบบคอมพิวเตอร์เข้ามาช่วยในการประมาณค่าความสามารถและการคัดเลือกข้อสอบ โดยกระบวนการคำนวณทางคอมพิวเตอร์จะช่วยทำให้การประมาณค่ามีความถูกต้องของคะแนนการ

ทดสอบในเชิงสถิติที่เพิ่มขึ้น ทำให้ได้องค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี 5 องค์ประกอบ ได้แก่ **1. การเข้าถึงสารสนเทศ** (information accessibility) **2. การจัดการสารสนเทศ** (information management) **3. การบูรณาการสารสนเทศ** (information integration) **4. การประเมินสารสนเทศ** (information evaluation) **5.การสื่อสารสารสนเทศ** (information communication) (รายละเอียดจากการสังเคราะห์นำเสนอในตอนที่ 1)

เมื่อสร้างข้อสอบตามโครงสร้างที่กำหนดแล้วมีการนำข้อสอบไปทดลองใช้เพื่อนำผลการตอบมาวิเคราะห์คุณภาพของข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบแล้วจึงคัดข้อสอบที่มีคุณภาพเข้าคลังข้อสอบ โมเดลการตอบสนองข้อสอบมีบทบาทสำคัญอย่างมากต่อการทดสอบแบบปรับเหมาะ โมเดลการตอบสนองข้อสอบที่ไม่เหมือนกันย่อมให้ผลการประมาณค่าพารามิเตอร์ต่างๆ แตกต่างกันด้วย โดยทั่วไปแล้ว โมเดลการตอบสนองข้อสอบที่นิยมใช้ในปัจจุบันแบ่งออกเป็น 3 กลุ่มตามจำนวนพารามิเตอร์ข้อสอบที่ประมาณได้จากโมเดล และนิยมเขียนให้อยู่ในรูปฟังก์ชันโลจิสติกส์ (logistic function) ได้แก่ โมเดลตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (3pl) ซึ่งประกอบด้วยพารามิเตอร์ความยาก (item difficulty) พารามิเตอร์การจำแนก (item discrimination) และพารามิเตอร์การเดา (guessing) โมเดลตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (2pl) ซึ่งประกอบด้วยพารามิเตอร์ความยากและ พารามิเตอร์การจำแนก และโมเดลตอบสนองข้อสอบแบบ 1 พารามิเตอร์ (1pl) ซึ่งประกอบด้วยพารามิเตอร์ความยากเพียงอย่างเดียว การเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดล 2 โมเดลสามารถพิจารณาได้จากการทดสอบอัตราส่วนความเป็นไปได้ (likelihood-ratio test) ที่ว่าโมเดลที่สนใจ 2 โมเดลนั้นเหมือนกันหรือไม่ หรือเรียกว่าการทดสอบ LR โดยค่า log-likelihood สูงสุดของแต่ละโมเดลที่มาจากข้อมูลคำตอบเดียวกันและมาจากการประมาณด้วยวิธีความเป็นไปได้สูงสุด (maximum likelihood estimation) เมื่อได้ข้อสรุปแล้วจึงเลือกโมเดลประมาณค่าที่มีความถูกต้องแม่นยำที่สุด (รายละเอียดแสดงในหัวข้อ 2.2)

การคัดเลือกข้อสอบที่มีคุณภาพเพื่อเข้าสู่คลังข้อสอบแบบปรับเหมาะ ผู้วิจัยพิจารณาคุณภาพข้อสอบจากค่าความคลาดเคลื่อนของการประมาณค่าพารามิเตอร์เป็นหลัก และเลือกกำจัดข้อสอบที่มีค่าพารามิเตอร์ความยากสูงผิดปกติและต่ำผิดปกติด้วย โดยผู้วิจัยยอมรับข้อสอบที่มีค่าพารามิเตอร์ความยากต่ำกว่า -3.00 และ สูงกว่า +3.00 ลอจิท เป็นข้อสอบที่มีคุณภาพตรงตามวัตถุประสงค์ของการวิจัย ที่ต้องการใช้ข้อสอบง่ายเพื่อตรวจสอบความรู้พื้นฐานของนักศึกษา และใช้ข้อสอบยากเพื่อท้าทายความสามารถของนักศึกษา เช่นเดียวกับงานวิจัยของ Ali and Chang (2014) และ Mao and Xin (2013) ที่ยอมรับข้อสอบที่มีค่าพารามิเตอร์ดังกล่าวในกระบวนการพัฒนาคลังข้อสอบแบบปรับเหมาะขั้นต้น ส่วนการพิจารณาคุณภาพข้อสอบจากค่าพารามิเตอร์การจำแนก เลือกใช้เกณฑ์ของ Baker and Kim (2017) โดยผู้วิจัยกำจัดข้อสอบที่มีค่าการจำแนกติดลบ และเลือกข้อสอบที่มีค่าการจำแนกตั้งแต่ 0 ขึ้นไป เช่นเดียวกับงานวิจัยของ Ali and Chang (2014) ที่ยอมรับข้อสอบที่มีค่าพารามิเตอร์การจำแนกใกล้เคียง 0 ในขั้นตอนของการพัฒนาคลังข้อสอบขั้นต้น

2. ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี การออกแบบและพัฒนาระบบการทดสอบฯแบ่งออกเป็น 4 ส่วน ได้แก่

1) การลงทะเบียน กลุ่มตัวอย่างในการวิจัยทุกคนต้องดำเนินการลงทะเบียนก่อนเข้าสู่ระบบเพื่อเป็นการยืนยันตัวบุคคล ระบบจะนำข้อมูลที่ผู้เข้ารับการทดสอบลงทะเบียนไว้มาใช้ในการออกรายงานผลการทดสอบ

2) สร้างชุดข้อสอบ สำหรับชุดข้อสอบขั้นแรก ผู้วิจัยต้องดำเนินการจัดชุดข้อสอบไว้ก่อนล่วงหน้า ดังนั้นจึงต้องมีระบบที่ใช้สำหรับสร้างข้อสอบในขั้นแรก โดยข้อสอบขั้นแรก จะมีจำนวนทั้งสิ้น 5 ข้อ และความครอบคลุมตามเนื้อหาที่กำหนด

3) การทดสอบ สำหรับหน้าจอการทดสอบจะมีองค์ประกอบหลักที่แสดงได้แก่ ชุดข้อสอบที่ผู้สอบกำลังดำเนินการทดสอบ แถบแสดงสถานะเนื้อหาที่กลุ่มตัวอย่างในการวิจัยกำลังดำเนินการทดสอบ ข้อสอบและตัวเลือกสำหรบข้อสอบแต่ละข้อ และระยะเวลาสำหรับการทำแบบทดสอบ โดยมีส่วนที่สำคัญ คือ

- การประมาณค่าความสามารถ สำหรับขั้นตอนการออกแบบระบบการประมาณค่า ใช้การประมาณค่าร่วมกัน 2 แบบได้แก่ วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) และวิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP)

- การคัดเลือกข้อสอบ สำหรับการออกแบบระบบการเลือกข้อสอบสำหรับการวิจัยนี้ ใช้ดัชนีความสำคัญสูงสุด (maximum priority index: MPI) ในการเลือกข้อสอบสำหรับขั้นถัดไป โดยในการเลือกข้อสอบมีการจัดแบ่งคลังข้อสอบตามค่าอำนาจจำแนก (a-stratified) ควบคุมการเปิดเผยข้อสอบ (exposure rate) ด้วยวิธี Sympson & Hetter และควบคุมเนื้อหาของแบบทดสอบ (content balance)

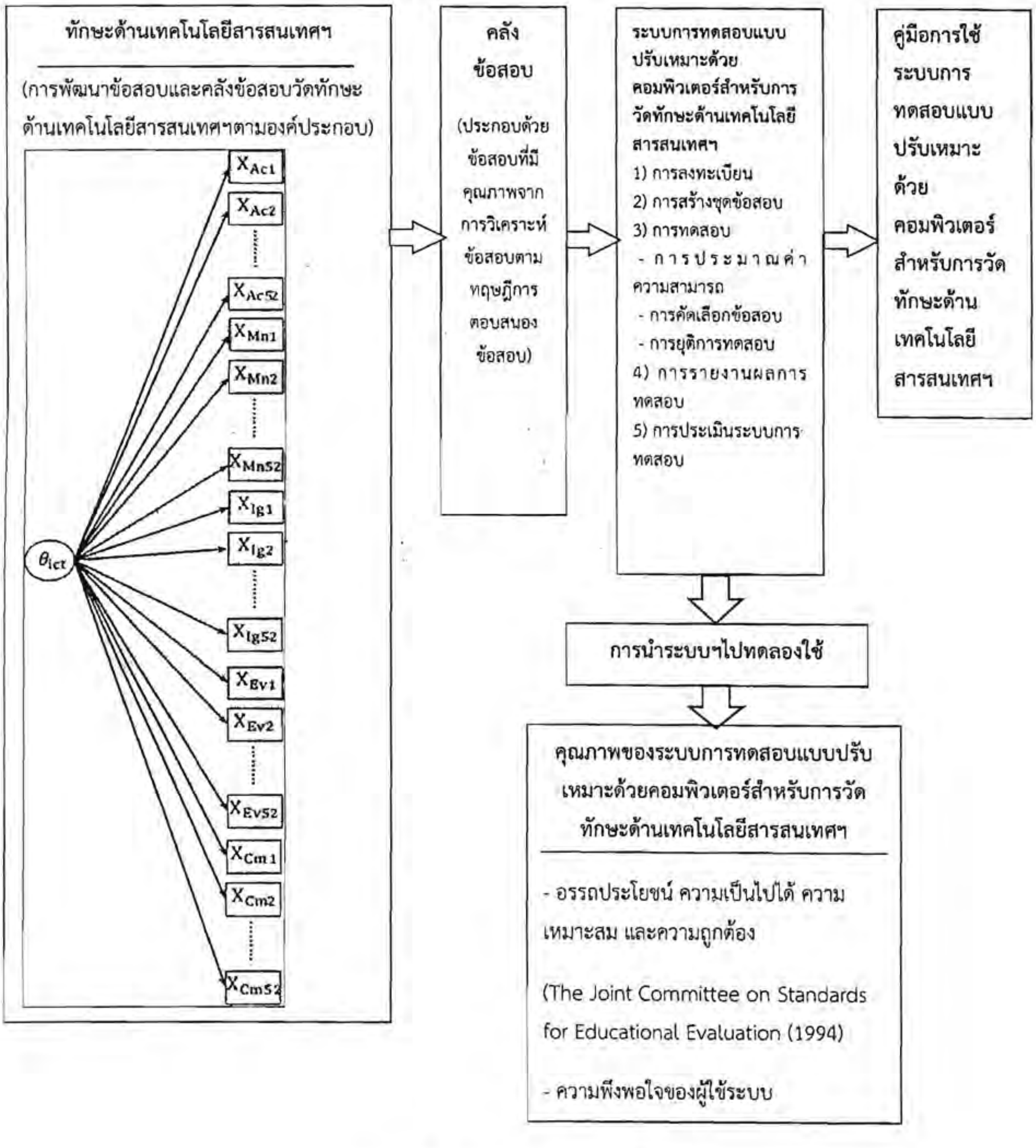
- การยุติการทดสอบ สำหรับเกณฑ์ในการยุติการทดสอบสำหรับการวิจัยครั้งนี้ ใช้ การยุติโดยพิจารณาจากความคลาดเคลื่อนมาตรฐานในการวัด ซึ่งจะยุติการทดสอบเมื่อความคลาดเคลื่อนมาตรฐานในการวัดมีค่าต่ำกว่าที่กำหนดไว้ วิธีการนี้เป็นวิธีที่นิยมใช้มากที่สุด ร่วมกับการยุติโดยกำหนดจำนวนข้อสอบ ซึ่งจะยุติการทดสอบเมื่อจำนวนข้อสอบ ถึงค่าที่กำหนดไว้

4) การรายงานผลการทดสอบ สำหรับรายงานผลการทดสอบจะแสดงผลการตอบข้อสอบทั้งหมด โดยรูปแบบรายงานผลการทดสอบแบ่งออกเป็นสองส่วนได้แก่ การรายงานจำนวนข้อที่ผู้เข้ารับการทดสอบตอบข้อสอบถูกและผิดในแต่ละองค์ประกอบ และส่วนที่สองคือกราฟแสดงสัมพันธ์ของผลการทดสอบในแต่ละองค์ประกอบ

3. คู่มือการใช้ระบบการทดสอบระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 สร้างขึ้นเพื่อให้ผู้เกี่ยวข้องสามารถใช้ระบบฯที่พัฒนาขึ้นได้อย่างถูกต้องแบ่งออกเป็น 2 ส่วน ได้แก่ คู่มือการใช้งานสำหรับผู้ใช้งานทั่วไป (user) และคู่มือการใช้งานสำหรับผู้จัดการระบบ (admin) โดยที่ คู่มือการใช้งานสำหรับผู้ใช้งานทั่วไป มีรายละเอียดเกี่ยวกับการลงทะเบียน การเข้าใช้งานระบบครั้งแรก และการรายงานผลการสอบเมื่อทำข้อสอบครบตามเงื่อนไข ส่วนคู่มือการใช้งานสำหรับผู้จัดการระบบมีรายละเอียดเกี่ยวกับการเข้าสู่ระบบ การจัดการข้อสอบ การแก้ไขข้อสอบ การสร้างชุดข้อสอบ การจัดการสมาชิกและการดูรายงานผลการสอบของสมาชิก

4. การนำระบบฯไปทดลองใช้ เป็นการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้กับนิสิตนักศึกษา ซึ่งเป็นการตรวจสอบระบบฯและทำให้ได้ข้อมูลสารสนเทศที่เป็นประโยชน์เพื่อนำมาปรับปรุงระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยี สารสนเทศและการสื่อสารฯ ให้มีประสิทธิภาพมากยิ่งขึ้น

5. คุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี คุณภาพของระบบประเมินระบบการทดสอบ ใช้การประเมินระบบแบบอิงมาตรฐานตามแนวคิดของ Stufflebeam ประกอบด้วยเนื้อหาในการประเมินจำนวน 4 องค์ประกอบ ได้แก่ การใช้ประโยชน์ (utility) ความเป็นไปได้ (feasibility) ความเหมาะสม (propriety) และความถูกต้อง (accuracy) ร่วมกับการประเมินผลระบบการทดสอบจากผู้ใช้งานที่เป็นผู้สอบ ประเมินความพึงพอใจ 5 ด้าน ได้แก่ 1. ภาพรวมปฏิสัมพันธ์ของระบบ 2. หน้าจอ (screen) 3.การใช้คำศัพท์และสารสนเทศระบบ (terminology and system information) 4. การเรียนรู้ (learning) และ 5.ความสามารถของระบบ (system capabilities) โดยใช้มาตรการวัดตามแนว Osgood 7 ระดับ



ภาพ 2.9 กรอบแนวคิดในการวิจัย

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยเรื่อง "ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี" ได้ดำเนินการวิจัยจำแนกออกเป็น 5 ระยะ คือ

ระยะที่ 1 การพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ระยะที่ 2 การพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ระยะที่ 3 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ระยะที่ 4 การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้

ระยะที่ 5 การตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี
โดยแต่ละระยะมีรายละเอียดของวิธีดำเนินการวิจัยดังนี้

ระยะที่ 1 การพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

การพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยประกอบด้วยกระบวนการย่อย 2 ขั้นตอน คือ 1) การศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องและผลการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี และ 2) การกำหนดนิยามเชิงปฏิบัติในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยมีรายละเอียดของแหล่งข้อมูล ขั้นตอนการดำเนินการวิจัย เครื่องมือที่ใช้ในการวิจัย การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูล ดังนี้

1.1 แหล่งข้อมูล

เอกสารและงานวิจัยที่เกี่ยวข้องกับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีทั้งในประเทศไทยและต่างประเทศโดยรวบรวมมาจากเอกสารจริงและฐานข้อมูลออนไลน์ (Online database)

1.2 ขั้นตอนการดำเนินการวิจัย

1. ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการสร้างแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

2. สังเคราะห์องค์ประกอบการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

3. ระบุนิยามขององค์ประกอบการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

4. ผู้วิจัยดำเนินการประชุมเพื่อตรวจสอบความถูกต้องเหมาะสมขององค์ประกอบและนิยามของการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

5. จัดทำโครงสร้างองค์ประกอบและนิยามของแบบวัดวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เพื่อเป็นแบบแผนในการสร้างข้อสอบในระยะที่ 2

1.3 เครื่องมือที่ใช้ในการวิจัย

แบบตรวจสอบรายการ (Check list) เพื่อสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ดังตัวอย่างในตาราง 3.1

ตาราง 3.1 ตัวอย่างแบบตรวจสอบรายการในการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

องค์ประกอบของทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร	ชื่อผู้แต่ง (พ.ศ.)		
	A (1999)	B (XXXX)	C (YYYY)
การเข้าถึง (Access)	✓	✓	✓
การจัดการ (Manage)	✓	✓	✓
การบูรณาการ (Integrate)	✓	✓	✓
การประเมิน (Evaluate)	✓	✓	✓
การสร้างสรรค์ (Create)	✓	✓	✓
การสื่อสาร (Communicate)		✓	✓
การระบุ (Define)		✓	
การวางแผน (Plan)			✓
การทำงานร่วมกัน (Collaborate)			✓
การวิเคราะห์และการตัดสิน (Reflect and judge)			✓

1.4 การเก็บรวบรวมข้อมูล

1) การศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องและผลการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

2) การกำหนดนิยามเชิงปฏิบัติในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

1.5 การวิเคราะห์ข้อมูล

1. การวิเคราะห์ข้อมูลเชิงคุณภาพ ได้แก่ การวิเคราะห์เนื้อหา (content analysis) เพื่อวิเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

2. การวิเคราะห์ข้อมูลเชิงปริมาณ ได้แก่ ร้อยละของความถี่ จากแบบตรวจสอบรายการในการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ดังตัวอย่างในตาราง 3.2

ตาราง 3.2 ตัวอย่างการวิเคราะห์ร้อยละของความถี่ของแบบตรวจสอบรายการในการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

องค์ประกอบของทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร	ชื่อผู้แต่ง (พ.ศ.)			ความถี่ (ร้อยละ)
	A (1999)	B (XXXX)	C (YYYY)	
การเข้าถึง (Access)	✓	✓	✓	3 (100.00)
การจัดการ (Manage)	✓	✓	✓	3 (100.00)
การบูรณาการ (Integrate)	✓	✓	✓	3 (100.00)
การประเมิน (Evaluate)	✓	✓	✓	3 (100.00)
การสร้างสรรค์ (Create)	✓	✓	✓	3 (100.00)
การสื่อสาร (Communicate)		✓	✓	2 (67.00)
การระบุ (Define)		✓		1 (33.00)
การวางแผน (Plan)			✓	1 (33.00)
การทำงานร่วมกัน (Collaborate)			✓	1 (33.00)
การวิเคราะห์และการตัดสิน (Reflect and judge)			✓	1 (33.00)

ระยะที่ 2 การพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

การพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ประกอบด้วยกระบวนการย่อย 4 ขั้นตอน คือ 1) การสร้างข้อสอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ตามนิยามเชิงปฏิบัติการที่กำหนดไว้ 2) การตรวจสอบความตรงเชิงเนื้อหาก่อนนำไปทดลองใช้ 3) การตรวจสอบคุณภาพของข้อสอบที่ได้จากการทดลองใช้ และ 4) การคัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบ โดยมีรายละเอียดของประชากรและตัวอย่าง ขั้นตอนการดำเนินการวิจัย เครื่องมือที่ใช้ในการวิจัย การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูล ดังนี้

2.1 ประชากรและตัวอย่าง

ประชากร

ประชากรในการวิจัยครั้งนี้ คือ นิสิตและนักศึกษาระดับปริญญาตรีจากสถาบันอุดมศึกษาของรัฐบาลทั่วประเทศ จำนวนทั้งสิ้น 1,422,254 คน ซึ่งมีรายละเอียดดังตาราง 3.3

ตาราง 3.3 ประชากรในการวิจัย

กลุ่มสาขาวิชา	จำนวน	ร้อยละ
1. ครุศาสตร์ ศึกษาศาสตร์ (Education)	147,229	10.35
2. ศิลปะศาสตร์ มนุษยศาสตร์ (Arts and Humanities)	175,633	12.35
3. สังคมศาสตร์ วารสารศาสตร์และสารสนเทศ (Social sciences, Journalism and Information)	170,779	12.01
4. บริหารธุรกิจ การบริหารการจัดการและนิติศาสตร์ (Business, Administration and Law)	412,964	29.04
5. วิทยาศาสตร์ คณิตศาสตร์และสถิติ (Natural Sciences, Mathematics and Statistics)	77,021	5.42
6.เทคโนโลยีสารสนเทศและการสื่อสาร (Information and Communication Technologies)	59,574	4.19
7. วิศวกรรมศาสตร์ (Engineering, Manufacturing and Construction)	173,982	12.23
8. เกษตรศาสตร์ ป่าไม้ การประมงและสัตวแพทย์ (Agriculture, Forestry, Fisheries and Veterinary)	54,168	3.81
9. สุขภาพและสวัสดิการ (Health and Welfare)	88,015	6.19
10. การบริการ (Services)	62,889	4.42
รวม	1,422,254	100.00

(ที่มา : สำนักงานคณะกรรมการการอุดมศึกษา, 2561)

ตัวอย่าง

ตัวอย่างในการวิจัยครั้งนี้ ได้แก่ นิสิตและนักศึกษาระดับปริญญาตรีจากสถาบันอุดมศึกษาของรัฐบาลในกรุงเทพมหานคร จำนวนทั้งสิ้น 2,000 คน โดยมีขั้นตอนการสุ่มและกำหนดขนาดกลุ่มตัวอย่างดังนี้

1) การสุ่มเลือกนิสิตและนักศึกษาโดยใช้สาขาวิชาเป็นหน่วยในการสุ่ม

จากข้อมูลของ สำนักงานคณะกรรมการการอุดมศึกษา (2561) ที่ได้จำแนกสาขาวิชาไว้จำนวน 10 สาขาวิชา ได้แก่ 1) ครุศาสตร์ ศึกษาศาสตร์ (Education) 2) ศิลปะศาสตร์ มนุษยศาสตร์ (Arts and Humanities) 3) สังคมศาสตร์ วารสารศาสตร์และสารสนเทศ (Social Sciences, Journalism and Information) 4) บริหารธุรกิจ การบริหารการจัดการและนิติศาสตร์ (Business, Administration and Law) 5) วิทยาศาสตร์ คณิตศาสตร์และสถิติ (Natural Sciences, Mathematics and Statistics) 6) เทคโนโลยีสารสนเทศและการสื่อสาร (Information and Communication Technologies) 7) วิศวกรรมศาสตร์ (Engineering, Manufacturing and Construction 8) เกษตรศาสตร์ ป่าไม้ การประมงและสัตวแพทย์ (Agriculture, Forestry, Fisheries and Veterinary) 9) สุขภาพและสวัสดิการ (Health and Welfare) และ 10) การบริการ (Services) จากข้อมูลสาขาวิชาที่มีทั้งหมดคือ 10 สาขาวิชา ผู้วิจัยได้สุ่มสาขาวิชามาจำนวน 5 สาขา ได้แก่ 1) ครุศาสตร์ ศึกษาศาสตร์ (Education) 2) สังคมศาสตร์ วารสารศาสตร์และสารสนเทศ(Social Sciences, Journalism and Information) 3) วิทยาศาสตร์ คณิตศาสตร์และสถิติ (Natural Sciences, Mathematics and Statistics) 4) วิศวกรรมศาสตร์ (Engineering, Manufacturing and Construction) และ 5) สุขภาพและสวัสดิการ (Health and Welfare) ดังตาราง 3.4

ตาราง 3.4 ผลการสุ่มตัวอย่างโดยใช้สาขาวิชาเป็นหน่วยในการสุ่ม

ประชากร	ร้อยละ	ผลการสุ่มตัวอย่าง
1. ครุศาสตร์ ศึกษาศาสตร์ (Education)	10.35	✓
2. ศิลปะศาสตร์ มนุษยศาสตร์ (Arts and Humanities)	12.35	
3. สังคมศาสตร์ วารสารศาสตร์และสารสนเทศ (Social Sciences, Journalism and Information)	12.01	✓
4. บริหารธุรกิจ การบริหารการจัดการและนิติศาสตร์ (Business, Administration and Law)	29.04	
5. วิทยาศาสตร์ คณิตศาสตร์และสถิติ (Natural Sciences, Mathematics and Statistics)	5.42	✓
6.เทคโนโลยีสารสนเทศและการสื่อสาร (Information and Communication Technologies)	4.19	
7. วิศวกรรมศาสตร์ (Engineering, Manufacturing and Construction)	12.23	✓
8. เกษตรศาสตร์ ป่าไม้ การประมงและสัตวแพทย์ (Agriculture, Forestry, Fisheries and Veterinary)	3.81	
9. สุขภาพและสวัสดิการ (Health and Welfare)	6.19	✓
10. การบริการ (Services)	4.42	

หมายเหตุ ✓ หมายถึง สาขาที่ได้รับการสุ่ม

2) การกำหนดขนาดของตัวอย่าง

การวิจัยครั้งนี้เป็นการวิจัยเพื่อสร้างและพัฒนาคลังข้อสอบแบบปรับเหมาะซึ่งในการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบนั้นอยู่บนพื้นฐานแนวคิดทฤษฎีการตอบสนองข้อสอบ (Item response theory; IRT) ซึ่งในการออกแบบระบบการทดสอบแบบปรับเหมาะในครั้งนี้จะประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ โดยขนาดตัวอย่างที่เหมาะสมควรมีขนาดมากกว่า 1,200 คน (สุนทร เทียนงาม, 2551) ดังนั้นเพื่อให้ได้ขนาดของตัวอย่างเพียงพอและป้องกันการสูญหายของข้อมูลระหว่างการเก็บข้อมูลจึงได้กำหนดขนาดของตัวอย่างเป็น 2,000 คน ซึ่งจำนวนตัวอย่างที่ได้ของแต่ละสาขาวิชามาจากการคำนวณอิงตามสัดส่วนจริงของจำนวนประชากร โดยมีรายละเอียดดังตาราง 3.5

ตาราง 3.5 ผลการกำหนดขนาดตัวอย่าง

สาขาวิชา	จำนวนประชากร (คน)	จำนวนตัวอย่าง (คน)
1. ครุศาสตร์ ศึกษาศาสตร์ (Education)	147,229	448
2. สังคมศาสตร์ วารสารศาสตร์และสารสนเทศ (Social Sciences, Journalism and Information)	170,779	520
3. วิทยาศาสตร์ คณิตศาสตร์และสถิติ (Natural Sciences, Mathematics and Statistics)	77,021	235
4. วิศวกรรมศาสตร์ (Engineering, Manufacturing and Construction)	173,982	530
5. สุขภาพและสวัสดิการ (Health and Welfare)	88,015	267
รวม	**657,026**	**2,000**

เมื่อกำหนดขนาดตัวอย่างในการเก็บข้อมูล คือ 2,000 คน ผู้วิจัยจึงได้เลือกสถาบันอุดมศึกษาของรัฐบาลในกรุงเทพมหานคร จำนวน 6 สถาบัน ซึ่งครอบคลุมสาขาวิชาที่ได้สุ่มไว้คือ 5 สาขาวิชา โดยจากการเก็บข้อมูลจริงพบว่าสามารถเก็บข้อมูลได้จริง 1,672 คน คิดเป็นร้อยละ 83.60 รายละเอียดดังตาราง 3.6 และ 3.7

ตาราง 3.6 เปรียบเทียบการกำหนดขนาดตัวอย่างตามแผนและผลการเก็บข้อมูลจริง

สาขาวิชา	จำนวนตัวอย่างตามแผน (คน)	จำนวนตัวอย่างที่เก็บข้อมูลได้จริง (คน)
1. ครุศาสตร์ ศึกษาศาสตร์ (Education)	448	486
2. สังคมศาสตร์ วารสารศาสตร์และสารสนเทศ (Social Sciences, Journalism and Information)	520	293
3. วิทยาศาสตร์ คณิตศาสตร์และสถิติ (Natural Sciences, Mathematics and Statistics)	235	325
4. วิศวกรรมศาสตร์ (Engineering, Manufacturing and Construction)	530	300
5. สุขภาพและสวัสดิการ (Health and Welfare)	267	268
รวม	2,000	1,672

ผู้วิจัยได้ดำเนินการเลือกสถาบันอุดมศึกษาของรัฐบาลในกรุงเทพมหานครที่มีสาขาวิชาตามแผนการสุ่มตัวอย่าง จำนวน 6 โดยมีรายละเอียดดังตาราง 3.7

ตาราง 3.7 ตัวอย่างในการเก็บข้อมูลจำแนกตามสถาบันและสาขาวิชา

สถาบัน	สาขาวิชา	คณะ	จำนวน	ร้อยละ
1.จุฬาลงกรณ์มหาวิทยาลัย	ครุศาสตร์ ศึกษาศาสตร์ (Education)	ครุศาสตร์	162	50.60
	วิทยาศาสตร์ คณิตศาสตร์และสถิติ (Natural Sciences, Mathematics and Statistics)	วิทยาศาสตร์	45	14.10
	วิศวกรรมศาสตร์ (Engineering, Manufacturing and Construction)	วิศวกรรมศาสตร์	113	35.30
		รวม	320	100.00
2.มหาวิทยาลัยศรีนครินทรวิโรฒ	สังคมศาสตร์ วารสารศาสตร์และสารสนเทศ (Social Sciences, Journalism and Information)	สังคมศาสตร์	293	76.10
	วิศวกรรมศาสตร์ (Engineering, Manufacturing and Construction)	วิศวกรรมศาสตร์	92	23.90
		รวม	385	100.00
3.มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ	วิทยาศาสตร์ คณิตศาสตร์และสถิติ (Natural Sciences, Mathematics and Statistics)	วิทยาศาสตร์	280	74.70
	วิศวกรรมศาสตร์ (Engineering, Manufacturing and Construction)	วิศวกรรมศาสตร์	89	23.70
		สถาปัตยกรรมศาสตร์	6	1.60
		รวม	375	100.00
4.วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้า	สุขภาพและสวัสดิการ (Health and Welfare)	แพทยศาสตร์	268	100.00
		รวม	268	100.00
5.มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา	ครุศาสตร์ ศึกษาศาสตร์ (Education)	ครุศาสตร์	125	100.00
		รวม	125	100.00
6. มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ฯ	ครุศาสตร์ ศึกษาศาสตร์ (Education)	ครุศาสตร์	199	100.00
		รวม	199	100.00
		รวมทั้งหมด	1,672 คน	

ผู้ให้ข้อมูล

ผู้ให้ข้อมูลในการวิจัยครั้งนี้ ประกอบด้วย 1) กรรมการผู้สร้างข้อสอบและ 2) ผู้ทรงคุณวุฒิในการตรวจสอบความตรง

1) กรรมการผู้สร้างข้อสอบแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี จำนวน 9 ท่าน จากสาขาวิชาด้านการวัดและประเมินผลและสาขาวิชาเทคโนโลยีการศึกษา

2) ผู้ทรงคุณวุฒิผู้ตรวจสอบความสอดคล้องขององค์ประกอบและข้อคำถามของแบบวัดจำนวน 15 ท่าน ได้แก่ อาจารย์ในสถาบันการศึกษาระดับอุดมศึกษาที่มีความเชี่ยวชาญในเนื้อหาทางด้านเทคโนโลยี จำนวน 7 ท่าน และผู้เชี่ยวชาญในด้านการวัดและประเมินผลจำนวน 8 ท่าน

2.2 ขั้นตอนการดำเนินการวิจัย

1) ดำเนินการจัดประชุมเพื่อวางแผนสำหรับการสร้างแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 โดย มีกรรมการผู้ร่วมสร้างแบบวัดทั้ง 9 ท่าน เข้าร่วมการประชุมกับคณะผู้วิจัย ซึ่งกรรมการทุกท่านจะได้รับมอบหมายให้สร้างแบบวัดในทุกองค์ประกอบตามองค์ประกอบที่ผ่านการสังเคราะห์และตรวจสอบจากผู้ทรงคุณวุฒิ

2) ดำเนินการรวบรวมข้อสอบจากกรรมการแล้วทีมผู้วิจัยดำเนินการวิพากษ์ปรับปรุงแก้ไขข้อสอบและจัดชุดเบื้องต้นเพื่อส่งให้ผู้ทรงคุณวุฒิจำนวน 15 ท่านดำเนินการตรวจสอบความสอดคล้องของข้อคำถามและองค์ประกอบในการวัด

3) ปรับปรุง แก้ไขข้อสอบตามข้อเสนอแนะของผู้ทรงคุณวุฒิ

4) ตรวจสอบคุณภาพของข้อสอบที่ได้จากการทดลองใช้

5) คัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบ

2.3 เครื่องมือที่ใช้ในการวิจัย

1) แบบประเมินความสอดคล้องระหว่างข้อคำถามกับองค์ประกอบในการวัด (Item Objective Congruence Index: IOC) โดยมีตัวอย่างดังตาราง 3.8

ตาราง 3.8 ตัวอย่างแบบประเมินความสอดคล้องระหว่างข้อคำถามกับองค์ประกอบที่ 1 การเข้าถึงสารสนเทศ(Information accessibility)

ข้อสอบ	ระดับความคิดเห็น			ข้อเสนอแนะ
	+1	0	-1	
1. เมื่อท่านต้องการทำรายงานวิจัย โดยมีคำถามวิจัย คือ "ในการดูแลสุขภาพจิต การนั่งสมาธิ หรือ การเดินจงกลม จะส่งผลต่อสุขภาพจิตได้ดีกว่ากัน" ท่านจะค้นหาข้อมูลโดยใช้คำต่าง ๆดังนี้ <u>ยกเว้น</u> ข้อใดต่อไปนี้ 1. การดูแลสุขภาพจิต 2. การนั่งสมาธิ 3. การเดินจงกลม 4. ส่งผลต่อสุขภาพจิต 5*. ได้ดีกว่ากัน				

2) แบบทดสอบ จำนวน 260 ข้อ จำแนกเป็น 5 ฉบับ (A, B, C, D และ E) โดยฉบับ A, B และ C มีจำนวนข้อสอบฉบับละ 52 ข้อ โดยทุกข้อเป็นข้อสอบแบบเลือกตอบหลายตัวเลือก (Multiple Choice) จำนวน 5 ตัวเลือก 1 คำตอบถูก โดยมีรายละเอียดของจำนวนข้อสอบและจำนวนผู้ทรงคุณวุฒิในการประเมินข้อสอบและผังข้อสอบ ดังตาราง 3.9 - 3.10

ตาราง 3.9 จำนวนข้อสอบและจำนวนผู้ทรงคุณวุฒิในการประเมินข้อสอบจำแนกตามชุดแบบทดสอบ

แบบทดสอบชุดที่	จำนวนข้อสอบจำแนกตามองค์ประกอบในการวัด (ข้อ)					รวม (ข้อ)	จำนวนผู้ทรงคุณวุฒิ (ท่าน)		รวม (คน)
	1	2	3	4	5		ICT	การวัดฯ	
A	11	11	10	10	10	52	2	1	3
B	11	11	10	10	10	52	1	2	3
C	10	10	11	11	10	52	2	1	3
D	10	10	11	11	10	52	1	2	3
E	10	10	10	10	12	52	1	2	3
รวม	52	52	52	52	52	260	7	8	15

ตาราง 3.10 รายละเอียดองค์ประกอบ นิยามและจำนวนของข้อสอบในแบบทดสอบ

องค์ประกอบ	นิยาม	จำนวน (ข้อ)
องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ (Information accessibility)	ความรู้และความสามารถในการสืบค้นและรวบรวมสารสนเทศ รวมทั้งอธิบายหรือแสดงวิธีการเรียกสารสนเทศที่มีอยู่กลับคืนมาเพื่อการใช้งานในระบบดิจิทัลได้	52
องค์ประกอบที่ 2 การจัดการสารสนเทศ (Information management)	ความรู้และความสามารถในการจัดระบบหมวดหมู่ของสารสนเทศในระบบดิจิทัลให้ง่ายต่อการการเรียกค้นและการประยุกต์ใช้ให้เหมาะสมกับงานแต่ละประเภทได้	52
องค์ประกอบที่ 3 การบูรณาการสารสนเทศ (Information integration)	ความรู้และความสามารถในการแปลความหมาย สรุป เปรียบเทียบ จำแนกความเหมือนและความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆในระบบดิจิทัลได้	52
องค์ประกอบที่ 4 การประเมินสารสนเทศ (Information evaluation)	ความรู้และความสามารถในการตัดสินคุณภาพ ความสอดคล้อง ประโยชน์ หรือประสิทธิภาพของสารสนเทศที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร ในระบบดิจิทัลได้	52
องค์ประกอบที่ 5 การสื่อสารสนเทศ (Information communication)	ความรู้และความสามารถในการจัดทำและเผยแพร่สารสนเทศในระบบดิจิทัลเพื่อตอบสนองความต้องการของกลุ่มบุคคลและบริบทต่างๆที่เป็นการเคารพสิทธิ์ความเป็นส่วนตัวได้	52
	รวม	260

ผังข้อสอบ (Test blue print) จำแนกตามองค์ประกอบ

องค์ประกอบ	ระดับพฤติกรรม						รวม
	ความรู้	ความเข้าใจ	การนำไปใช้	การวิเคราะห์	การประเมินค่า	การสังเคราะห์	
1	2	10	15	15	10	-	52
2	2	10	15	15	10	-	52
3	2	10	15	15	10	-	52
4	2	10	15	15	10	-	52
5	2	10	15	15	10	-	52
รวม	10	50	75	75	50	-	260

2.4 การเก็บรวบรวมข้อมูล

1) ผู้วิจัยและผู้ทรงคุณวุฒิร่วมกันตรวจสอบความถูกต้องเหมาะสมของข้อคำถาม และตัวเลือก ซึ่งได้แก่ ตัวลวงและคำตอบถูก นอกจากนี้ยังร่วมกันพิจารณา รูปแบบการพิมพ์ ภาษาที่ใช้ ตลอดจนความชัดเจนของรูปภาพต่างๆก่อนนำไปจัดพิมพ์เพื่อนำไปจัดชุด 10 ชุด โดยมีรายละเอียดดังตาราง 3.11

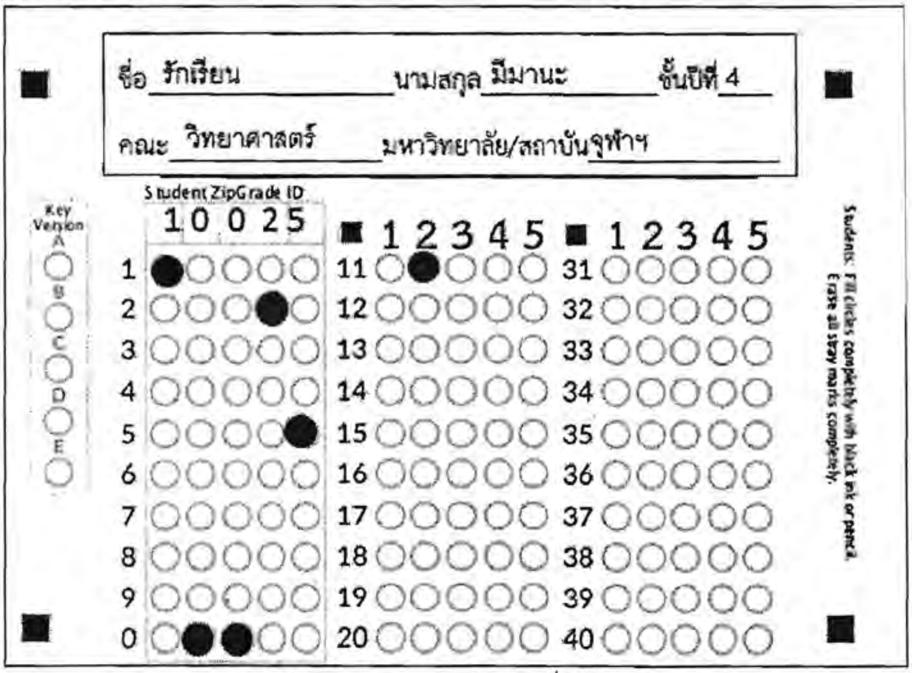
ตาราง 3.11 แผนการจัดข้อสอบจำแนกออกเป็น 10 ชุด

แบบทดสอบฉบับที่	ข้อสอบ	จำนวนข้อสอบในแต่ละองค์ประกอบ					รวม (ข้อ)
		1	2	3	4	5	
1	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35
2	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35
3	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35
4	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35
5	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35
6	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35

แบบทดสอบฉบับที่	ข้อสอบ	จำนวนข้อสอบในแต่ละองค์ประกอบ					รวม (ข้อ)
		1	2	3	4	5	
7	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35
8	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35
9	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35
10	ข้อสอบเฉพาะ	5	5	5	5	5	25
	ข้อสอบร่วม	2	2	2	2	2	10
	รวม	7	7	7	7	7	35
รวมทั้งหมด	ข้อสอบเฉพาะ	50	50	50	50	50	250
	ข้อสอบร่วม	2	2	2	2	2	10
รวมทั้งหมด							260

2) นัดหมายกำหนดการในการขอความร่วมมือเพื่อเก็บข้อมูลการวิจัยจากสถาบันอุดมศึกษาในกรุงเทพมหานครและปริมณฑลตามแผนการกำหนดจำนวนตัวอย่างที่ได้คำนวณไว้

3) ออกแบบกระดาษคำตอบภายใต้การใช้โปรแกรมสำเร็จรูปในการตรวจข้อสอบ คือ Zip Grade ซึ่งเป็นกระดาษคำตอบแบบฝนที่ออกแบบขึ้นภายใต้โปรแกรมสำเร็จรูปจาก www.zipgrade.com ซึ่งช่วยอำนวยความสะดวกในการตรวจข้อสอบผ่าน Zip Grade Application บนมือถือที่รองรับทั้งระบบปฏิบัติการ iOS และ Androids ดังภาพ 3.1



ภาพ 3.1 ตัวอย่างกระดาษคำตอบที่ใช้ในการทดสอบ

4) ดำเนินการจัดการทดสอบตามกำหนดการที่ได้นัดหมาย โดยกำหนดให้การเก็บข้อมูลในแต่ละครั้ง ผู้สอบจะได้รับข้อสอบเพียงคนละ 1 ชุด (ชุดละ 35 ข้อ) โดยใช้ระยะเวลาในการทดสอบ 60 นาที

5) ดำเนินการเก็บรวบรวมกระดาษคำถามและแบบทดสอบกลับคืนเพื่อนำไปวิเคราะห์หาค่าพารามิเตอร์และค่าสถิติพื้นฐานของคะแนนสอบและข้อสอบ

2.5 การวิเคราะห์ข้อมูล

1) การวิเคราะห์ค่าสถิติบรรยาย ได้แก่ ค่าเฉลี่ย ค่าสูงสุด ค่าต่ำสุด ส่วนเบี่ยงเบนมาตรฐาน ของคะแนนสอบ ค่าความยาก อำนาจจำแนก

2) การวิเคราะห์ค่าความตรงตามเนื้อหา (IOC) ซึ่งมีสูตรการคำนวณความสอดคล้องระหว่างข้อคำถามกับองค์ประกอบในการวัด (Item Objective Congruence Index: IOC) ดังนี้

$$\mathrm{IOC} = \frac{\sum \mathrm{R}}{\mathrm{N}}$$

$\sum \mathrm{R}$ = ผลรวมคะแนนความคิดเห็นของผู้ทรงคุณวุฒิ

N = จำนวนผู้ทรงคุณวุฒิ

เกณฑ์การพิจารณา

IOC $\geq$ 0.5 แสดงว่าข้อคำถามวัดวัตถุประสงค์ข้อนั้นจริง หมายความว่า ข้อนั้นมีความตรงตามเนื้อหา

IOC < 0.5 แสดงว่าข้อคำถามไม่ได้วัดวัตถุประสงค์ข้อนั้นจริง หมายความว่าข้อนั้นไม่มีความตรงตามเนื้อหา

3) การวิเคราะห์ค่าความเที่ยงของแบบทดสอบที่มีการให้คะแนนแบบ 0-1 โดยใช้สูตร KR20

$$\mathrm{KR20} = \left[\frac{\mathrm{k}}{\mathrm{k}-1}\right]\left[1 - \frac{\sum \mathrm{p_i q_i}}{\mathrm{S_x^2}}\right]$$

เมื่อ KR20 = สัมประสิทธิ์ความเที่ยงของแบบสอบ

K = จำนวนข้อสอบ

p_i = สัดส่วนของผู้ตอบถูกในข้อที่ i

q_i = สัดส่วนของผู้ตอบผิดในข้อที่ i

S_x^2 = ความแปรปรวนของคะแนนรวม x

4) การวิเคราะห์คุณภาพของข้อสอบตามทฤษฎี CTT

โปรแกรมที่ใช้ในการวิเคราะห์ข้อสอบตามทฤษฎี CTT ได้แก่ Zip Grade, TAP และ MS Excel ซึ่งผลการวิเคราะห์คุณภาพข้อสอบรายข้อจะวิเคราะห์จาก 1) ค่าความยาก (p) และ 2) ค่าอำนาจจำแนก (r) โดยใช้สูตรดังนี้

1) ค่าความยาก (p)

$$P = \frac{R_H + R_L}{N_H + N_L}$$

P คือ ความยากง่าย

R_H คือ จำนวนผู้สอบที่ตอบถูกในกลุ่มคะแนนสูง

R_L คือ จำนวนผู้สอบที่ตอบถูกในกลุ่มคะแนนต่ำ

N_H คือ จำนวนผู้สอบทั้งหมดในกลุ่มคะแนนสูง

N_L คือ จำนวนผู้สอบทั้งหมดในกลุ่มคะแนนต่ำ

2) ค่าอำนาจจำแนก (r)

$$r = \frac{R_H - R_L}{N_H \text{ or } N_L}$$

r คือ ค่าอำนาจจำแนก

R_H คือ จำนวนผู้สอบที่ตอบถูกในกลุ่มคะแนนสูง

R_Lคือ จำนวนผู้สอบที่ตอบถูกในกลุ่มคะแนนต่ำ

N_Hคือ จำนวนผู้สอบทั้งหมดในกลุ่มคะแนนสูง

N_Lคือ จำนวนผู้สอบทั้งหมดในกลุ่มคะแนนต่ำ

5) การวิเคราะห์คุณภาพของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ

การวิเคราะห์คุณภาพข้อสอบในส่วนนี้ เป็นการตรวจสอบความสอดคล้องกลมกลืนของโมเดลการวัด และการประมาณค่าพารามิเตอร์ข้อสอบ ก่อนที่จะคัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์เข้าสู่คลังข้อสอบสำหรับการทดสอบแบบปรับเหมาะในระยะต่อไป ประกอบด้วยขั้นตอนย่อย 3 ขั้นตอนคือ

5.1) การปรับเทียบข้อสอบพร้อมกัน

ตาราง 3.11 แสดงให้เห็นว่า ข้อสอบแบ่งออกเป็น 10 ชุด ชุดละ 35 ข้อ โดยแต่ละ-ชุดมีข้อสอบร่วมจำนวน 10 ข้อ และข้อสอบปกติจำนวน 25 ข้อ รวมทั้งสิ้น 260 ข้อ นอกจากนี้แต่ละชุดยังมีข้อสอบที่วัดแต่ละองค์ประกอบจำนวนเท่ากันด้วย คำตอบของตัวอย่างจากข้อสอบทุกข้อจะถูกให้คะแนนแบบสองค่า (dichotomous scoring) คำตอบถูกในแต่ละข้อจะได้ 1 คะแนน ส่วนคำตอบผิด ตอบมากกว่า 1 ตัวเลือกและข้อมูลขาดหาย (missing response) จะได้ 0 คะแนน จากนั้นจึงวิเคราะห์ความสอดคล้องกลมกลืนระหว่างโมเดลการวัดและข้อมูลคะแนนคำตอบต่อไป

ผู้วิจัยออกแบบการเก็บรวบรวมข้อมูลโดยให้ตัวอย่างทำข้อสอบเพียง 1 ชุด เท่านั้น และให้ข้อสอบทุกชุดมีข้อสอบร่วม ดังนั้นตัวอย่างจะได้ตอบข้อสอบร่วมเหมือนกันทุกคน จากนั้นผู้วิจัยจึงปรับเทียบ (equate) ข้อมูลที่ได้จากข้อสอบแต่ละชุด โดยรวมข้อมูลคำตอบของตัวอย่างทุกคนจากข้อสอบทุกข้อ แล้ววิเคราะห์ข้อมูลพร้อมกัน เรียกวิธีการนี้ว่า การปรับเทียบพร้อมกัน (concurrent calibration method: Wingersky & Lord, 1984) พารามิเตอร์ข้อสอบที่มาจากต่างชุด และค่าพารามิเตอร์ความสามารถตัวอย่างที่ทำข้อสอบบางข้อไม่เหมือนกัน จะถูกประมาณค่าพร้อมกันด้วยการวิเคราะห์เพียงครั้งเดียว ด้วยชุดคำสั่ง ltm (Rizopoulos, 2006) และ mirt (Chalmers, 2012) ในโปรแกรม R โดยที่การแจกแจงพารามิเตอร์

ความสามารถของตัวอย่างที่ทำข้อสอบแต่ละชุด จะสมมติให้มาจากประชากร 10 กลุ่ม และข้อสอบที่ไม่ได้เป็นข้อสอบร่วมจะถูกสมมติให้เป็นข้อสอบที่ไม่ได้ให้ตัวอย่างตอบ (Gonzalez & Wiberg, 2017) การตรวจสอบความถูกต้องแม่นยำของค่าพารามิเตอร์ข้อสอบที่ได้จากวิธีปรับเทียบพร้อมกันของ DeMars (2002) แสดงให้เห็นว่า ค่าพารามิเตอร์ความยากข้อสอบจะสูงกว่าความเป็นจริง (overestimate) เมื่อข้อสอบชุดนั้นเป็นข้อสอบค่อนข้างง่าย และค่าพารามิเตอร์ความยากข้อสอบจะต่ำกว่าความเป็นจริง (underestimate) เมื่อข้อสอบชุดนั้นเป็นข้อสอบค่อนข้างยาก แต่อย่างไรก็ตาม จากการจำลองข้อมูลของ DeMars (2002) ก็แสดงหลักฐานที่สนับสนุนว่า ความลำเอียงของค่าความยากข้อสอบที่เกิดขึ้นมีค่าเล็กน้อย

5.2) การเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลการตอบสนองข้อสอบแบบ 1PL, 2PL และ 3PL

ผู้วิจัยมีมุมมองต่อการสร้างโมเดลการวัด 2 แบบ ได้แก่ มุมมองเอกมิติ และมุมมองพหุมิติ กล่าวคือ มุมมองเอกมิตินี้มีข้อตกลงเบื้องต้นว่า แม้ว่าข้อสอบจะแบ่งเป็น 5 ด้านตามองค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร องค์ประกอบเหล่านั้นสามารถรวมเป็นหนึ่งเดียวกันได้ โดยมีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารเป็นตัวแปรแฝง (latent variable หรือ construct) หรือเป็นมโนทัศน์ฐานราก (underlying concept) ที่ฝังอยู่ในเนื้อหาของแต่ละองค์ประกอบ ดังนั้นโมเดลการวัดที่ผู้วิจัยใช้เปรียบเทียบความสอดคล้องกลมกลืนคือ โมเดลการตอบสนองข้อสอบเอกมิติ (unidimensional item response model) แบบ 1PL, 2PL และ 3PL ดังภาพ 3.2

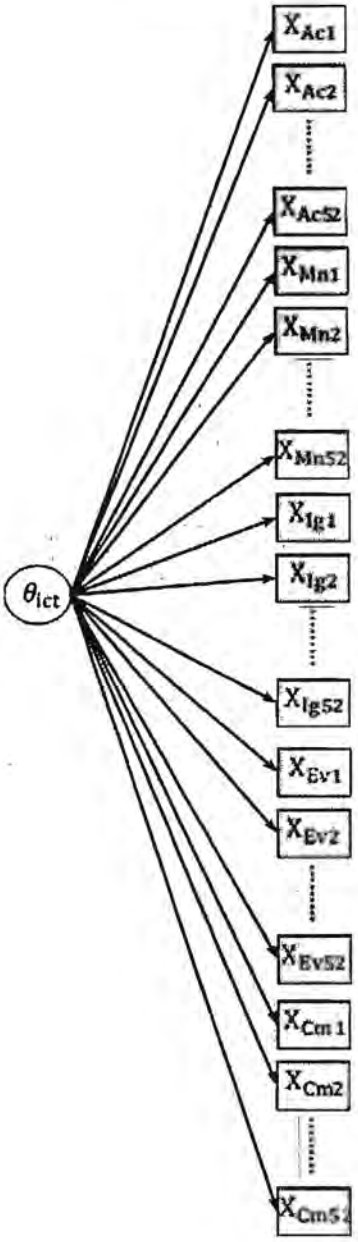
สัญลักษณ์ θ_{ict} ในภาพ 3.2 คือ พารามิเตอร์ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารของตัวอย่างแต่ละคน โดยที่พารามิเตอร์นี้ประมาณค่าได้จากคะแนนคำตอบ (X=0 หรือ 1) ของตัวอย่างแต่ละคนที่มีมาจากข้อสอบทุกข้อ ซึ่งวัดครอบคลุมด้านการเข้าถึงสารสนเทศ (Information Accessibility: Ac) ด้านการจัดการสารสนเทศ (Information Management: Mn) ด้านการบูรณาการสารสนเทศ (Information Integration: Ig) ด้านการประเมินสารสนเทศ (Information Evaluation: Ev) และด้านการสื่อสารสนเทศ (Information Communication: Cm) รูปวงกลมในภาพ 3.2 แสดงตัวแปรแฝงทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารของตัวอย่าง รูปสี่เหลี่ยมแสดงตัวแปรสังเกตได้ หรือคะแนนคำตอบของข้อสอบแต่ละข้อ และลูกศรที่ชี้จากวงกลมไปสี่เหลี่ยมแสดงให้เห็นว่า ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารของตัวอย่าง สามารถวัดได้ด้วยข้อสอบทั้งหมด

ค่าพารามิเตอร์ θ_{ict} และค่าพารามิเตอร์ข้อสอบสามารถประมาณได้จากโมเดล IRT แบบ 1PL, 2PL และ 3PL ดังสมการ 3.1-3.3 จากนั้นจึงเปรียบเทียบว่าโมเดลใดเป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลมากที่สุด

$$P_{\text{i}}(\text{X} = 1|\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \tag{3.1}$$

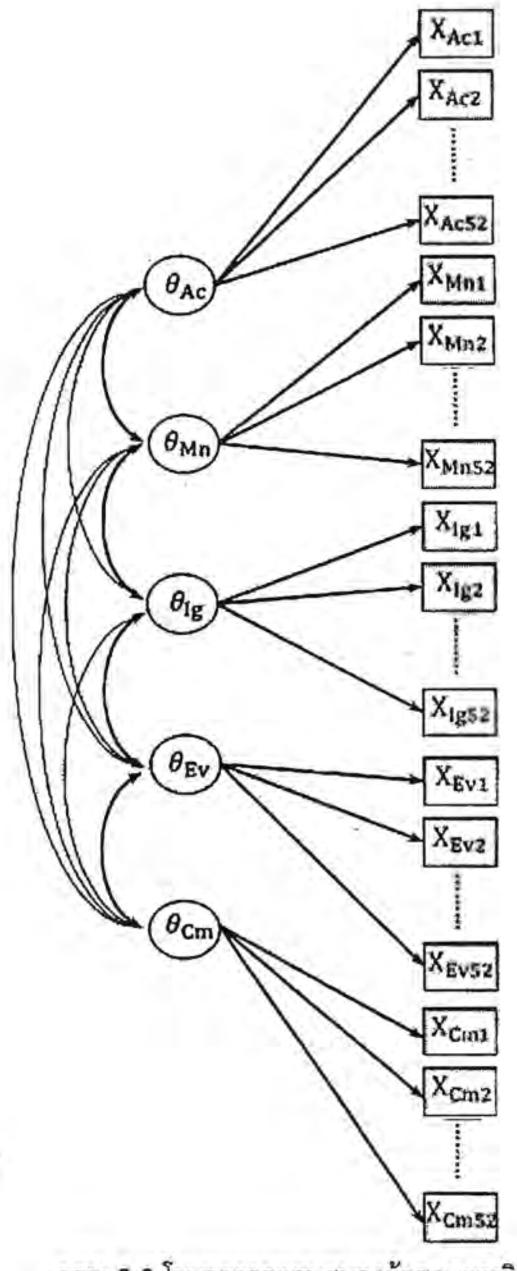
$$P_{\text{i}}(\text{X} = 1|\theta) = \frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}} \tag{3.2}$$

$$P_{\text{i}}(\text{X} = 1|\theta) = c_i + (1 - c_i)\frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}} \tag{3.3}$$



ภาพ 3.2 โมเดลการตอบสนองข้อสอบเอกมิติ

มุมมองที่ 2 คือ มุมมองพหุมิติต่อทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร โดยมุมมองนี้เชื่อว่า ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารประกอบด้วย 5 องค์ประกอบที่สัมพันธ์กันและร่วมกันอธิบายคุณลักษณะของตัวอย่างที่มีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร หรือตัวอย่างที่มีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารจะส่งผลให้มีความรู้และความสามารถใน 5 ด้านด้วย ดังนั้นโมเดลการวัดที่ผู้วิจัยใช้เปรียบเทียบความสอดคล้องกลมกลืนคือ โมเดลการตอบสนองข้อสอบพหุมิติ (multidimensional item response model) แบบ 1PL, 2PL และ 3PL ดังภาพ 3.3



ภาพ 3.3 โมเดลการตอบสนองข้อสอบพหุมิติ

ภาพ 3.3 แสดงให้เห็นว่า ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารประกอบด้วยความรู้และความสามารถเกี่ยวกับการเข้าถึงสารสนเทศ (θ_{Ac}) ด้านการจัดการสารสนเทศ (θ_{Mn}) ด้านการบูรณาการสารสนเทศ (θ_{Ig}) ด้านการประเมินสารสนเทศ (θ_{Ev}) และด้านการสื่อสารสนเทศ (θ_{Cm}) โดยที่ความรู้และความสามารถทั้ง 5 ด้านนี้สัมพันธ์กัน ซึ่งแสดงด้วยสัญลักษณ์เส้นโค้งที่มีหัวลูกศรทั้ง 2 ด้าน

5.3) การคัดเลือกข้อสอบที่มีคุณภาพเข้าคลังข้อสอบแบบปรับเหมาะ

ภายหลังจากผู้วิจัยเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลการวัดเอกมิติแบบ 1PL, 2PL และ 3PL และเปรียบเทียบระหว่างโมเดลเอกมิติและโมเดลพหุมิติ จนได้โมเดลที่สอดคล้องกลมกลืนกับ

ข้อมูลคะแนนคำตอบของตัวอย่างมากที่สุดแล้ว ผู้วิจัยจะคัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์ที่กำหนด โดยจะพิจารณาจากค่าความคลาดเคลื่อนของการประมาณค่าเป็นหลัก

ระยะที่ 3 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยการออกแบบโปรแกรมคอมพิวเตอร์หลังจากที่ได้ข้อสอบที่มีคุณภาพเข้าคลังข้อสอบ ได้แบ่งการดำเนินงานออกเป็น 4 ส่วน คือ 1) การลงทะเบียน 2) การสร้างชุดข้อสอบ 3) การทดสอบ 4) การรายงานผลการทดสอบ การพัฒนาระบบใช้ภาษา PHP เนื่องจากเป็นระบบการทดสอบที่สามารถใช้งานผ่านระบบออนไลน์ โดยมีรายละเอียดแต่ละขั้นตอนมีดังนี้

1) การลงทะเบียน ในการทดสอบกลุ่มตัวอย่างในการวิจัยทุกคนต้องดำเนินการลงทะเบียนก่อนเข้าสู่ระบบ หลังจากนั้นระบบจะนำข้อมูลที่ได้จากการลงทะเบียนมาใช้ในการออกรายงานผลการทดสอบ

2) การสร้างชุดข้อสอบ สำหรับชุดข้อสอบขั้นแรก ผู้วิจัยจะดำเนินการจัดชุดข้อสอบไว้ก่อนล่วงหน้า ดังนั้นจึงต้องมีระบบเพื่อใช้สำหรับสร้างข้อสอบในขั้นแรก ซึ่งข้อสอบขั้นแรก มีจำนวนทั้งสิ้น 5 ข้อ โดยมีความครอบคลุมตามเนื้อหาที่กำหนด สำหรับข้อสอบที่นำมาใช้ในการพัฒนาระบบครั้งนี้ใช้การวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) แบบ 2 พารามิเตอร์ได้แก่ ค่าความยาก (b) และอำนาจจำแนก (a) ซึ่งมีเกณฑ์ในการคัดเลือกข้อสอบในการจัดทำคลัง ได้แก่ ค่าความยากอยู่ในช่วง -4 ถึง 4 และค่าอำนาจจำแนกอยู่ระหว่าง 0 ถึง ∞

3) การทดสอบ สำหรับหน้าจอในการทดสอบ ประกอบด้วยองค์ประกอบหลัก คือ แถบแสดงสถานะเนื้อหาที่กลุ่มตัวอย่างในการวิจัยกำลังดำเนินการทดสอบ ข้อสอบพร้อมตัวเลือกสำหรับทำการทดสอบแต่ละข้อ และระยะเวลาสำหรับการทำแบบทดสอบ ซึ่งสำหรับการวิจัยนี้ใช้เวลาในการทดสอบทั้งสิ้นไม่เกิน 60 นาที โดยมีส่วนที่สำคัญ 3 ส่วน คือ 1) การประมาณค่าความสามารถ สำหรับการประมาณค่าความสามารถของผู้สอบใช้การประมาณค่าร่วมกัน 2 วิธี คือ วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) เมื่อผู้สอบตอบข้อสอบผิดและถูกผสมกัน และวิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) เมื่อผู้สอบตอบข้อสอบถูกหรือผิดทั้งหมด 2) การคัดเลือกข้อสอบ สำหรับการเลือกข้อสอบข้อถัดไปสำหรับการวิจัยนี้ ใช้คัดเลือกข้อสอบที่มีตำแหน่งสารสนเทศ (item information) สูงสุดซึ่งสอดคล้องกับระดับความสามารถของผู้สอบ พร้อมทั้งควบคุมการเปิดเผยข้อสอบ (exposure rate) ไม่เกินร้อยละ 20 และควบคุมเนื้อหาของแบบทดสอบ (content balance) ให้เป็นไปตามตามองค์ประกอบที่กำหนด 3) การยุติการทดสอบ ใช้ค่าความสามารถของผู้เข้ารับการทดสอบที่มีความคงเส้นคงวา โดยที่ค่าความคลาดเคลื่อนมาตรฐาน (standard error) มีค่าน้อยกว่าหรือเท่ากับ 0.3

4) การรายงานผลการทดสอบ สำหรับรายงานผลการทดสอบแบ่งออกเป็น 2 ส่วน คือ 1) รายงานผลการตอบข้อสอบและร้อยละในการตอบข้อสอบถูก และ 2) กราฟเรดาร์จากผลการตอบข้อสอบตามองค์ประกอบ 5 ด้าน ทั้งนี้รูปแบบรายงานประกอบด้วยข้อมูลส่วนตัวของกลุ่มตัวอย่างในการวิจัย ได้แก่ ชื่อนามสกุล ชั้นปี คณะ และมหาวิทยาลัย

ระยะที่ 4 การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้

การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้ มีการดำเนินการดังรายละเอียดต่อไปนี้

ตัวอย่างวิจัย

ตัวอย่างวิจัย คือ นิสิตระดับปริญญาตรีสายสังคมศาสตร์ และสายวิทยาศาสตร์ ชั้นปีที่ 3 จำนวน 217 คน แบ่งเป็นเพศชาย จำนวน 104 คน (ร้อยละ 47.93) และเพศหญิง จำนวน 113 คน (ร้อยละ 52.07) โดยมีนิสิตสายสังคมศาสตร์ จำนวน 119 คน (ร้อยละ 54.84) แบ่งเป็นเพศชาย จำนวน 44 คน (ร้อยละ 20.28) และเพศหญิง จำนวน 75 คน (ร้อยละ 34.56) และมีนิสิตสายวิทยาศาสตร์ จำนวน 98 คน (ร้อยละ 45.16) แบ่งเป็นเพศชาย จำนวน 60 คน (ร้อยละ 27.65) และเพศหญิง จำนวน 38 คน (ร้อยละ 17.51) ดังตาราง 3.12

ตาราง 3.12 จำนวนตัวอย่างวิจัยสำหรับการทดลองใช้ระบบการทดสอบ

สายการเรียน	จำนวนตัวอย่างวิจัย		
	เพศชาย *n* (%)	เพศหญิง *n* (%)	รวม *n* (%)
สังคมศาสตร์	44 (20.28)	75 (34.56)	119 (54.84)
วิทยาศาสตร์	60 (27.65)	38 (17.51)	98 (45.16)
รวม	104 (47.93)	113 (52.07)	217 (100.00)

เครื่องมือวิจัย

เครื่องมือวิจัย คือ ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ที่ผู้วิจัยพัฒนาขึ้น ซึ่งโปรแกรมการทดสอบ แบ่งออกเป็น 4 ส่วน ได้แก่ การลงทะเบียน การสร้างชุดข้อสอบ การทดสอบ และการรายงานผลการทดสอบ โดยวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 จำนวน 5 องค์ประกอบ ได้แก่ การเข้าถึงสารสนเทศ การจัดการสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ และการสื่อสารสารสนเทศ

การเก็บรวบรวมข้อมูล

การทดลองใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี มีขั้นตอนการเก็บรวบรวมข้อมูลดังนี้

1. ประสานงานกับเจ้าหน้าที่ประจำห้องคอมพิวเตอร์เพื่อขอเข้าใช้ห้องคอมพิวเตอร์สำหรับการจัดสอบกับตัวอย่างวิจัย พร้อมทั้งตรวจสอบความเรียบร้อยของอุปกรณ์ภายในห้องคอมพิวเตอร์

2. จัดนิสิตให้นั่งประจำคอมพิวเตอร์ จากนั้นจึงชี้แจงรายละเอียดเกี่ยวกับการทดสอบ และการตอบคำถามในโปรแกรมการทดสอบ

3. ก่อนการทดสอบ นิสิตจะต้องลงทะเบียนการทดสอบผ่านเว็บไซต์ http://www.oneinthai.com/ictc จากนั้นจึงเข้าระบบการทดสอบเพื่อกรอกรายละเอียดเกี่ยวกับผู้เข้ารับการทดสอบ เมื่อกรอกรายละเอียดครบถ้วนแล้ว โปรแกรมการทดสอบจะแสดงคำชี้แจงในการทดสอบ หลังจากอ่านคำชี้แจงจนเข้าใจแล้ว นิสิตจึงเริ่มต้นการทดสอบ โดยมีเวลาในการทดสอบทั้งสิ้น 1 ชั่วโมง หากนิสิตมีข้อสงสัยในการทดสอบหรือมีปัญหาระหว่างการทดสอบสามารถสอบถามผู้คุมสอบได้ตลอดเวลา

4. โปรแกรมการทดสอบจะเสนอแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ให้นิสิตที่ทำแบบสอบเสร็จเพื่อประเมินความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบ หลังจากนิสิตทำแบบประเมินเสร็จแล้ว โปรแกรมการทดสอบจะรายงานผลการทดสอบ โดยรายงานจำนวนข้อถูกและข้อผิดในแต่ละองค์ประกอบ พร้อมทั้งแสดงกราฟความสัมพันธ์ระหว่างคะแนนในแต่ละองค์ประกอบ

การวิเคราะห์ข้อมูล

1. วิเคราะห์ข้อมูลภูมิหลังของตัวอย่างวิจัยโดยใช้สถิติเชิงบรรยาย ได้แก่ ความถี่ และร้อยละ

2. วิเคราะห์ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 และคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ทั้ง 5 องค์ประกอบ ได้แก่ การเข้าถึงสารสนเทศ การจัดการสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ และการสื่อสารสารสนเทศ จำแนกตามสายการเรียนโดยใช้สถิติเชิงบรรยาย ได้แก่ ค่าเฉลี่ยเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด สัมประสิทธิ์การกระจาย ความเบ้ และความโด่ง

3. วิเคราะห์ความแตกต่างของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 และคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบ ระหว่างนิสิตเพศชายและเพศหญิง ในแต่ละสายการเรียนโดยใช้สถิติทดสอบทีสำหรับตัวอย่าง 2 กลุ่ม ที่เป็นอิสระต่อกัน (t-test for independent samples)

ระยะที่ 5 การตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

การตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี มีการดำเนินการดังรายละเอียดต่อไปนี้

ผู้ให้ข้อมูล

ผู้ให้ข้อมูลแบ่งเป็น 2 กลุ่ม คือ ผู้ให้ข้อมูลสำหรับการตรวจสอบคุณภาพของระบบการทดสอบ และผู้ให้ข้อมูลสำหรับการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบ

ผู้ให้ข้อมูลสำหรับการตรวจสอบคุณภาพของระบบการทดสอบ คือ ผู้ทรงคุณวุฒิด้านเทคโนโลยีสารสนเทศและการสื่อสารและผู้ทรงคุณวุฒิด้านการศึกษา จำนวน 3 ท่าน (ภาคผนวก)

ผู้ให้ข้อมูล สำหรับการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบเป็นกลุ่มเดียวกับตัวอย่างวิจัยที่ทดลองใช้ระบบการทดสอบ ซึ่งมีรายละเอียดดังตาราง 3.12

เครื่องมือวิจัย

เครื่องมือวิจัย คือ แบบประเมินระบบแบบอิงมาตรฐาน มีลักษณะเป็นมาตรประมาณค่า 5 ระดับ จำนวน 14 ข้อ และแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ มีลักษณะเป็นมาตรจำแนกความหมาย 7 ระดับ จำนวน 20 ข้อ

1. แบบประเมินระบบแบบอิงมาตรฐาน

แบบประเมินระบบแบบอิงมาตรฐานมีขั้นตอนการสร้าง ดังนี้

1) ศึกษาองค์ประกอบของการประเมินระบบแบบอิงมาตรฐาน และกำหนดนิยามเชิงปฏิบัติการของแต่ละองค์ประกอบ โดยการประเมินระบบแบบอิงมาตรฐานแบ่งเป็น 4 องค์ประกอบ ได้แก่ ความมีประโยชน์ ความเป็นไปได้ในการนำไปใช้ ความเหมาะสม และความถูกต้อง

2) สร้างข้อคำถามตามนิยามเชิงปฏิบัติการ โดยมีข้อคำถามทั้งสิ้น 14 ข้อ

3) ตรวจสอบความตรงเชิงเนื้อเรื่องโดยใช้ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมายของการวัด (Item-Objective Congruence: IOC) โดยให้ผู้ทรงคุณวุฒิด้านการศึกษา จำนวน 3 ท่าน เป็นผู้พิจารณา ซึ่งมีเกณฑ์ในการพิจารณา ดังนี้

+1 หมายถึง แน่ใจว่าข้อคำถามสอดคล้องกับนิยามเชิงปฏิบัติการ

0 หมายถึง ไม่แน่ใจว่าข้อคำถามสอดคล้องกับนิยามเชิงปฏิบัติการ

-1 หมายถึง แน่ใจว่าข้อคำถามไม่สอดคล้องกับนิยามเชิงปฏิบัติการ

2. แบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์

แบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์มีขั้นตอนการสร้าง ดังนี้

1) ศึกษาองค์ประกอบของความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ และกำหนดนิยามเชิงปฏิบัติการของแต่ละองค์ประกอบ โดยความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์แบ่งเป็น 5 องค์ประกอบ ได้แก่ ภาพรวมของระบบ หน้าจอของระบบ คำศัพท์และสารสนเทศระบบ การเรียนรู้ของผู้ใช้งาน และความสามารถของระบบ

2) สร้างข้อคำถามตามนิยามเชิงปฏิบัติการ โดยมีข้อคำถามทั้งสิ้น 20 ข้อ

3) ตรวจสอบความตรงเชิงเนื้อเรื่องโดยใช้ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมายของการวัด (Item-Objective Congruence: IOC) โดยให้ผู้ทรงคุณวุฒิด้านการศึกษา จำนวน 3 ท่าน เป็นผู้พิจารณา ซึ่งมีเกณฑ์ในการพิจารณา ดังนี้

+1 หมายถึง แน่ใจว่าข้อคำถามสอดคล้องกับนิยามเชิงปฏิบัติการ

0 หมายถึง ไม่แน่ใจว่าข้อคำถามสอดคล้องกับนิยามเชิงปฏิบัติการ

-1 หมายถึง แน่ใจว่าข้อคำถามไม่สอดคล้องกับนิยามเชิงปฏิบัติการ

การเก็บรวบรวมข้อมูล

การตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี มีขั้นตอนการเก็บรวบรวมข้อมูล ดังนี้

1. ให้ผู้ทรงคุณวุฒิด้านเทคโนโลยีสารสนเทศและการสื่อสารและผู้ทรงคุณวุฒิด้านการศึกษาประเมินคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยใช้แบบประเมินระบบแบบอิงมาตรฐาน

2. ให้นิสิตที่ทดลองใช้ระบบการทดสอบประเมินความพึงพอใจต่อระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยใช้แบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์

การวิเคราะห์ข้อมูล

1. วิเคราะห์ความตรงเชิงเนื้อเรื่องของแบบประเมินระบบแบบอิงมาตรฐาน และแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์โดยคำนวณค่าดัชนีความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมายของการวัด (IOC)

2. วิเคราะห์ข้อมูลภูมิหลังของตัวอย่างวิจัยสำหรับการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบโดยใช้สถิติเชิงบรรยาย ได้แก่ ความถี่ และร้อยละ

3. วิเคราะห์คุณภาพของระบบการทดสอบและความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบโดยใช้สถิติเชิงบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ยเลขคณิต และส่วนเบี่ยงเบนมาตรฐาน

ขั้นตอนการดำเนินการ (Process)	ผลที่ได้ (Output)
ระยะที่ 1 การพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ	
1. ศึกษาและสังเคราะห์องค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ →	องค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ
↓	
2. ระบุนิยามขององค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ →	นิยามขององค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ
↓	
3. จัดทำโครงสร้างของแบบทดสอบวัดทักษะด้านเทคโนโลยีสารสนเทศฯ →	โครงสร้างของแบบทดสอบตามนิยามขององค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ
ระยะที่ 2 การพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศฯ	
1. สร้างข้อสอบวัดทักษะด้านเทคโนโลยีสารสนเทศฯ →	ข้อสอบวัดทักษะด้านเทคโนโลยีสารสนเทศฯ
↓	
2. ตรวจสอบความตรงเชิงเนื้อหาก่อนนำไปทดลองใช้ →	ค่า IOC ของแบบสอบ
↓	
3. ตรวจสอบคุณภาพของข้อสอบที่ได้จากการทดลองใช้ →	ดัชนีความยากและอำนาจจำแนกตามทฤษฎี CTT และ IRT
↓	
4. คัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบ →	คลังข้อสอบที่มีข้อสอบที่มีคุณภาพตามเกณฑ์
ระยะที่ 3 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ	
1. การลงทะเบียน →	หน้าจอการทำงานในระบบ
↓	
2. การสร้างชุดข้อสอบ →	ชุดข้อสอบในระบบคลังข้อสอบ
↓	
3. การทดสอบ →	ผลการคัดเลือกข้อสอบและยุติการสอบ
↓	
4. การรายงานผลการทดสอบ →	ผลการตอบข้อสอบและกราฟเรดาร์จากผลการตอบข้อสอบตามองค์ประกอบ 5 ด้าน
ระยะที่ 4 การทดลองใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ	
1. การทดลองใช้ระบบฯกับนิสิตนักศึกษา ระดับปริญญาตรี →	คะแนนทักษะด้านเทคโนโลยีสารสนเทศ

ขั้นตอนการดำเนินการ (Process)	ผลที่ได้ (Output)
สายสังคมศาสตร์	และการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบฯ
2. การทดลองใช้ระบบฯกับนิสิตนักศึกษา ระดับปริญญาตรี สายวิทยาศาสตร์	คะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบฯ

ระยะที่ 5 การตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ขั้นตอนการดำเนินการ (Process)	ผลที่ได้ (Output)
1. การตรวจสอบคุณภาพระบบการทดสอบฯ	ผลการประเมินตามแบบประเมินระบบแบบอิงมาตรฐาน
2. การศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบฯ	ผลการประเมินความพึงพอใช้ของผู้ใช้ระบบการทดสอบฯ

ภาพ 3.4 กรอบการดำเนินการพัฒนาระบบระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศฯ

บทที่ 4

ผลการวิจัย

ในการนำเสนอผลการวิจัยครั้งนี้ จำแนกเป็น 5 ตอน ตามระยะดำเนินการวิจัย ดังนี้

ตอนที่ 1 ผลการพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

1.1 ผลการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องและผลการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

1.2 ผลการกำหนดนิยามเชิงปฏิบัติในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ตอนที่ 2 ผลการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

2.1 ผลการสร้างข้อสอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ตามนิยามเชิงปฏิบัติการที่กำหนดไว้

2.2 ผลการตรวจสอบความตรงเชิงเนื้อหาก่อนนำไปทดลองใช้

2.3 ผลการตรวจสอบคุณภาพของข้อสอบที่ได้จากการทดลองใช้

2.4 ผลการคัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบ

ตอนที่ 3 ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

3.1 ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

3.2 คู่มือการใช้ระบบการทดสอบระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

ตอนที่ 4 ผลการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้

4.1 การทดลองใช้ระบบฯกับนิสิตนักศึกษาปริญญาตรี สายสังคมศาสตร์

4.2 การทดลองใช้ระบบฯกับนิสิตนักศึกษาปริญญาตรี สายวิทยาศาสตร์

ตอนที่ 5 ผลการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

5.1 ผลการตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

5.2 ผลการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

โดยผลการวิจัยแต่ละตอนมีรายละเอียดดังนี้

ตอนที่ 1 ผลการพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ผลการพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี จำแนกออกเป็น 2 ประเด็น คือ 1) ผลการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องและผลการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี และ 2) ผลการกำหนดนิยามเชิงปฏิบัติในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยมีรายละเอียดดังนี้

1.1 ผลการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องและผลการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

การศึกษาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษ ที่ 21 จากเอกสารและงานวิจัยที่เกี่ยวข้อง พบว่า มีแนวคิดและงานวิจัยที่นำเสนอองค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ที่นำมาใช้กับนิสิตนักศึกษาปริญญาตรีได้ เพราะอยู่ในช่วงที่เข้าสู่วัยทำงาน และพบว่าองค์ประกอบการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารของเอกสารดังกล่าว มีความคล้ายคลึงกัน โดยให้ความสำคัญกับการใช้เทคโนโลยีสารสนเทศและการสื่อสารในการระบุและนำไปสู่การเข้าถึงข้อมูลที่ต้องการ การจัดระเบียบสารสนเทศให้เป็นหมวดหมู่ สามารถค้นคืนข้อมูลที่ต้องการได้ การประเมินสารสนเทศที่มีความจำเป็น ให้ระดับสารสนเทศที่มีคุณภาพ รวมไปถึงการสร้างสรรค์สารสนเทศ การส่งต่อสารสนเทศ การติดต่อสื่อสารในระดับบุคคลและในกลุ่มบุคคล ตามที่ได้นำเสนอในบทที่ 2 ดังนั้นผลการสังเคราะห์องค์ประกอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร แสดงในตาราง 4.1

ตาราง 4.1 การสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร

องค์ประกอบของทักษะการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร	Educational Testing Service: ETS (2002)	Katz and Macklin (2007)	Markauskaite (2007)	California Emerging Technology Fund (2008)	Somerville et al. (2008)	UNESCO (2008)	Claro et al. (2012)	Ahmad et al. (2013)	น้ำทิพย์ วิภาวิน (2552)	รวม (ร้อยละ)
การเข้าถึง (Access)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การจัดการ (Manage)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การบูรณาการ (Integrate)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การประเมิน (Evaluate)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การสร้างสรรค์ (Create)	✓	✓	✓	✓	✓	✓	✓	✓	✓	9 (100)
การสื่อสาร (Communicate)		✓	✓	✓	✓	✓	✓	✓	✓	8 (88.89)
การระบุ (Define)		✓			✓	✓	✓	✓	✓	7 (77.78)
การวางแผน (Plan)			✓							1 (11.11)
การทำงานร่วมกัน (Collaborate)			✓							1 (11.11)
การวิเคราะห์และการตัดสิน (Reflect and judge)			✓							1 (11.11)

จากตารางสังเคราะห์องค์ประกอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร สรุปองค์ประกอบหลักได้เป็น 6 องค์ประกอบ โดยใช้เกณฑ์พิจารณาองค์ประกอบที่มีค่าน้ำหนักมากกว่าร้อยละ 80 สรุปองค์ประกอบที่ทำการวัดทักษะด้านเทคโนโลยี สารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ได้แก่

1) การเข้าถึง (Access)
2) การจัดการ (Manage)
3) การบูรณาการ (Integrate)
4) การประเมิน (Evaluate)
5) การสร้างสรรค์ (Create)
6) การสื่อสาร (Communicate)

1.2 ผลการกำหนดนิยามเชิงปฏิบัติการในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

การกำหนดนิยามเชิงปฏิบัติการครั้งนี้ ได้จากการสังเคราะห์องค์ประกอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี เพื่อให้ได้องค์ประกอบที่เป็นตัวแทนที่ดี ลดความซ้ำซ้อนในการวัด รวมทั้งสามารถสร้างแบบสอบแบบเลือกตอบที่มีการให้คะแนนแบบ 0 - 1 ที่นำไปใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (computerized adaptive testing: CAT) ที่ใช้ระบบคอมพิวเตอร์เข้ามาช่วยในการประมาณค่าความสามารถและการคัดเลือกข้อสอบ โดยกระบวนการคำนวณทางคอมพิวเตอร์จะช่วยทำให้การประมาณค่ามีความถูกต้องของคะแนนการทดสอบในเชิงสถิติที่เพิ่มขึ้น จึงได้คัดเลือกองค์ประกอบที่มีความเหมาะสมและได้ตัดองค์ประกอบด้านการสร้างสรรค์ออก เนื่องจากไม่เหมาะกับรูปแบบของแบบสอบเลือกตอบที่ต้องมีการประมวลผลการให้คะแนนทันทีและมีความเป็นปรนัยในการสร้างระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ทำให้การวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีในการวิจัยครั้งนี้ ประกอบด้วย 5 องค์ประกอบ โดยสรุปคำนิยามและนิยามเชิงปฏิบัติการในแต่ละองค์ประกอบจากการสังเคราะห์เอกสารและงานวิจัยที่เกี่ยวข้อง ดังแสดงในตาราง 4.2

ตาราง 4.2 การสังเคราะห์นิยามตามองค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร

องค์ประกอบ	แหล่งข้อมูล	นิยามจากแหล่งข้อมูล	นิยามจากการสังเคราะห์	การกำหนดนิยามที่ปรับแก้ไขจากข้อเสนอแนะ
1) การเข้าถึง (Access)	ETS (2002)	การรู้สารสนเทศที่ต้องการและรู้ว่าจะเข้าถึงสารสนเทศดังกล่าวได้อย่างไร	การเข้าถึงสารสนเทศ (information accessibility) หมายถึง การระบุสารสนเทศที่จำเป็นและรู้วิธีการในการเก็บรวบรวม และการเรียกสารสนเทศกลับคืนเพื่อการใช้งาน (retrieve information) วัดจาก 1. การระบุสารสนเทศที่จำเป็น คือ สามารถเข้าใจและอธิบายความหมายของสารสนเทศที่เลือกมาใช้งานได้อย่างเหมาะสม 2. วิธีการในการเก็บรวบรวมสารสนเทศ คือ สามารถเลือกวิธีการในการเก็บรวบรวมสารสนเทศแต่ละประเภทได้	**การเข้าถึงสารสนเทศ** (information accessibility) หมายถึง การรู้สารสนเทศที่จำเป็นในการสืบค้นและการเรียกสารสนเทศกลับคืนเพื่อการใช้งาน (retrieve information)ในระบบดิจิทัล โดยมีนิยามปฏิบัติการคือ **"ความรู้และความสามารถในการสืบค้นและรวบรวมสารสนเทศ รวมทั้งอธิบายหรือแสดงวิธีการเรียกสารสนเทศที่มีอยู่กลับคืนมาเพื่อการใช้งานในระบบดิจิทัลได้"**
	Katz and Macklin (2007)	การรวบรวม หรือ ค้นคืนสารสนเทศ		
	Markauskaite (2007)	การทำงานภายในสภาพแวดล้อมที่สืบค้นจากแหล่งข้อมูลทางดิจิทัล การบำรุงรักษาเครื่องคอมพิวเตอร์ เพื่อเลือกเทคนิคและเครื่องมือที่เหมาะสม ในการค้นหาสารสนเทศจากสื่อและแหล่งข้อมูลต่าง ๆ		
	California Emerging Technology Fund (2008)	การมีความรู้และสามารถรวบรวมสารสนเทศโดยใช้เครื่องมือดิจิทัลในการสืบค้น ค้นหาและดึงข้อมูลในระบบดิจิทัลได้		

องค์ประกอบ	แหล่งข้อมูล	นิยามจากแหล่งข้อมูล	นิยามจากการสังเคราะห์	การกำหนดนิยามที่ปรับแก้ไขจากข้อเสนอแนะ
	Somerville et al. (2008)	การรวบรวม และ/หรือดึงสารสนเทศในระบบดิจิทัล โดยที่แหล่งข้อมูลอาจจะเป็นเว็บเพจ ฐานข้อมูล กลุ่มสนทนา อีเมล หรือสื่อสิ่งพิมพ์ออนไลน์ รวมถึงการสร้างและรวมคำที่ใช้ในการค้นหาเพื่อตอบสนองความต้องการของงานวิจัยที่เฉพาะ เจาะจง การเลือกใช้แหล่งข้อมูลอย่างน้อยหนึ่งรายการเพื่อค้นหาข้อมูลที่ต้องการ และการตัดสินใจเลือกแหล่งข้อมูลที่ให้ข้อมูลที่เป็นประโยชน์มากที่สุดสำหรับความต้องการ	อย่างเหมาะสม 3. การเรียกสารสนเทศกลับคืนเพื่อการใช้งาน คือ สามารถอธิบายหรือแสดงวิธีการเรียกสารสนเทศที่มีอยู่กลับคืนมาเพื่อการใช้งานได้อย่างถูกต้อง	
	UNESCO (2008)	การรวบรวม หรือ สืบค้นสารสนเทศ ในระบบดิจิทัล นอกจากนี้ ยังมีความสามารถในการพัฒนาระบบในการค้นหาข้อมูลในฐานข้อมูล		
	น้ำทิพย์ วิภาวิน (2552)	การรวบรวม และ/หรือ การค้นหาสารสนเทศในสังคมดิจิทัล		
2) การจัดการ (Manage)	ETS (2002)	ความสามารถในการจัดเก็บสารสนเทศที่มีอยู่ไว้เป็นหมวดหมู่ต่างๆง่ายต่อการใช้งาน	การจัดการสารสนเทศ (information management) หมายถึง การจัดระบบและเก็บรวบรวมสารสนเทศเพื่อเรียกกลับคืนมาใช้งาน และสามารถประยุกต์ใช้ระบบสารสนเทศที่มีอยู่พร้อมทั้งจำแนกประเภทได้	**การจัดการสารสนเทศ** (information management) หมายถึง การจัดระบบและเก็บรวบรวมสารสนเทศ เพื่อเรียกกลับคืนมาใช้งาน (retrieve) และใช้ซ้ำ (reuse)ในระบบดิจิทัล
	Katz and Macklin (2007)	การใช้เครื่องมือดิจิทัลในการจัดระบบหมวดหมู่ของสารสนเทศ		
	Markauskaite (2007)	การดำเนินการเกี่ยวกับซอฟต์แวร์ต่าง ๆ เช่น การจัดการข้อมูล การออกแบบฐานข้อมูล		

องค์ประกอบ	แหล่งข้อมูล	นิยามจากแหล่งข้อมูล	นิยามจากการสังเคราะห์	การกำหนดนิยามที่ปรับแก้ไขจากข้อเสนอแนะ
		เป็นต้น เพื่อการจัดทำแผนผังองค์กร การจัดหมวดหมู่และจัดเก็บข้อมูล บันทึกข้อมูลและแหล่งที่มาของข้อมูล เป็นต้น	1. การจัดระบบและเก็บรวบรวมสารสนเทศ เพื่อเรียกกลับคืนมาใช้งาน คือ สามารถบริหารจัดการระบบที่มีอยู่และเลือกใช้วิธีการในการจัดเก็บสารสนเทศที่มีความเหมาะสม เพื่อความสะดวกและรวดเร็วในการเรียกกลับคืนมาใช้งาน 2. สามารถประยุกต์ใช้ระบบสารสนเทศที่มีอยู่พร้อมทั้งจำแนกประเภทได้ คือ สามารถนำระบบสารสนเทศที่มีอยู่มาประยุกต์ใช้ให้มีความเหมาะสมกับงานแต่ละประเภท พร้อมทั้งสามารถจำแนกความแตกต่างของสารสนเทศที่มีอยู่ได้	โดยมีนิยามปฏิบัติการคือ **“ความรู้และความสามารถในการจัดระบบหมวดหมู่ของสารสนเทศในระบบดิจิทัลให้ง่ายต่อการเรียกค้นและการประยุกต์ใช้ให้เหมาะสมกับงานแต่ละประเภทได้"**
	California Emerging Technology Fund (2008)	การใช้เครื่องมือดิจิทัลในการจัดระบบหมวดหมู่ของสารสนเทศ โดยใช้เครื่องมือดิจิทัลดำเนินการจัดเตรียมข้อมูลพื้นฐานสำหรับการเรียกค้นและการประยุกต์ใช้ในอนาคต		
	Somerville et al. (2008)	การจัดระเบียบสารสนเทศ เพื่อช่วยให้บุคคลอื่นรวมถึงตัวเองสามารถค้นพบได้ในภายหลัง เช่น การจัดหมวดหมู่อีเมลลงในแฟ้มที่มีเนื้อหาอีเมลที่คล้ายกัน การจัดข้อมูลบุคลากรลงในแผนผังองค์กร การจัดเรียงไฟล์อีเมล หรือการส่งข้อมูลคืนฐานข้อมูล เพื่อชี้แจงข้อมูลที่เกี่ยวข้อง เป็นต้น		
	UNESCO (2008)	การจัดระเบียบสารสนเทศตามแบบแผนที่มีอยู่		
	น้ำทิพย์ วิภาวิน (2552)	การใช้เครื่องมือ ICT เพื่อประยุกต์ใช้สารสนเทศในองค์กร		

องค์ประกอบ	แหล่งข้อมูล	นิยามจากแหล่งข้อมูล	นิยามจากการสังเคราะห์	การกำหนดนิยามที่ปรับแก้ไขจากข้อเสนอแนะ
3) การบูรณาการ (Integrate)	ETS (2002)	ความสามารถในการแปลความหมาย สรุป เปรียบเทียบ แยกความแตกต่างของข้อมูลที่มีความคล้ายหรือมีรูปแบบการนำเสนอที่หลากหลาย	การบูรณาการสารสนเทศ (information integration) หมายถึง การแปลผลและนำเสนอสารสนเทศซึ่งมีความเกี่ยวข้องกับการสรุป เปรียบเทียบ และจำแนกความแตกต่าง 1. การแปลผลและนำเสนอสารสนเทศ คือ สามารถแปลผลข้อมูลจากสารสนเทศที่มีอยู่ พร้อมทั้งสามารถนำเสนอสารสนเทศในรูปแบบที่มีความเหมาะสมกับประเภทของการใช้งาน 2. การสรุป เปรียบเทียบ และจำแนกความแตกต่าง คือ สามารถสรุปรวมและจัดกลุ่มสารสนเทศแต่ละประเภทที่มีความสอดคล้องกัน พร้อมทั้งสามารถเปรียบเทียบและจำแนกความแตกต่างและความเหมาะสมของสารสนเทศแต่ละประเภทได้	**การบูรณาการสารสนเทศ** (information integration) หมายถึง การแปลผลและนำเสนอสารสนเทศซึ่งมีความเกี่ยวข้องกับการสรุป เปรียบเทียบ และจำแนกความแตกต่างของสารสนเทศในระบบดิจิทัล โดยมีนิยามปฏิบัติการคือ **“ความรู้และความสามารถในการแปลความหมาย สรุป เปรียบเทียบ จำแนกความเหมือนและความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆในระบบดิจิทัลได้”**
	Katz and Macklin (2007)	การอธิบายหรือแปลความหมายของสารสนเทศ เช่น การใช้เครื่องมือดิจิทัลในการสังเคราะห์ การสรุป การเปรียบเทียบ และแยกความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่าง ๆ		
	Markauskaite (2007)	การแก้ไขปัญหาโดยการใช้เครื่องมือดิจิทัลโดยการสรุป การเปรียบเทียบ และการจำแนกความแตกต่างของแหล่งข้อมูลและแนวคิดต่าง ๆ		
	California Emerging Technology Fund (2008)	การตีความและแสดงสารสนเทศ ทั้งการสรุปผล การเปรียบเทียบ และความแตกต่างระหว่างกัน โดยใช้เครื่องมือดิจิทัลในการสังเคราะห์ สรุป เปรียบเทียบ และ แยกความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆ		
	Somerville et al. (2008)	การตีความและแสดงข้อมูล เช่น การใช้เครื่องมือดิจิทัลเพื่อสังเคราะห์ สรุป		

องค์ประกอบ	แหล่งข้อมูล	นิยามจากแหล่งข้อมูล	นิยามจากการสังเคราะห์	การกำหนดนิยามที่ปรับแก้ไขจากข้อเสนอแนะ
		เปรียบเทียบข้อมูลจากแหล่งข้อมูลที่หลากหลาย ทั้งการเปรียบเทียบโฆษณา อีเมล หรือเว็บไซต์ของคู่แข่งด้วยการสรุปข้อมูลในลงตาราง การกำหนดเกณฑ์เฉพาะเพื่อเปรียบเทียบข้อมูลและตัดสินใจเป็นต้น		
	UNESCO (2008)	การตีความ สรุป เปรียบเทียบ และความแตกต่างกันระหว่างข้อมูลจากแหล่งข้อมูลที่หลากหลาย		
	น้ำทิพย์ วิภาวิน (2552)	การสังเคราะห์ สรุปความ เปรียบเทียบ และนำสารสนเทศมาใช้จากหลายแหล่ง		
4) การประเมิน (Evaluate)	ETS (2002)	การตัดสินเกี่ยวกับคุณภาพ ความสัมพันธ์ ประโยชน์ หรือ ประสิทธิภาพของสารสนเทศ	การประเมินสารสนเทศ (information evaluation) หมายถึง การสะท้อนถึงกระบวนการที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร เพื่อนำมาใช้ในการตัดสินใจซึ่งเกี่ยวข้องกับคุณภาพ ความสอดคล้อง ประโยชน์ หรือ ประสิทธิภาพของสารสนเทศ	**การประเมินสารสนเทศ** (information evaluation) หมายถึง การตัดสินคุณค่าของสารสนเทศที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสารในระบบดิจิทัล โดยมีนิยามปฏิบัติการคือ
	Katz and Macklin (2007)	การตัดสินระดับสารสนเทศดิจิทัลที่มีความต้องการจำเป็น การตัดสินคุณภาพ ความสัมพันธ์ ประโยชน์ หรือ ความน่าเชื่อถือ และมีประสิทธิภาพของข้อมูล		
	Markauskaite (2007)	การประเมินความสัมพันธ์ระหว่างทรัพยากรดิจิทัล สารสนเทศ และเครื่องมือ เพื่อใช้ในการกำหนดเกณฑ์การประเมินตัดสินคุณภาพ		

องค์ประกอบ	แหล่งข้อมูล	นิยามจากแหล่งข้อมูล	นิยามจากการสังเคราะห์	การกำหนดนิยามที่ปรับแก้ไขจากข้อเสนอแนะ
		ประโยชน์ และความสัมพันธ์ของสารสนเทศและการตัดสินใจ	1. การสะท้อนถึงกระบวนการที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร คือ สามารถเข้าใจแนวคิดวิธีการในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศที่มีความสอดคล้องกับลักษณะสารสนเทศแต่ละประเภท 2. การตัดสินใจซึ่งเกี่ยวข้องกับคุณภาพ ความสอดคล้อง ประโยชน์ หรือประสิทธิภาพของสารสนเทศ คือ สามารถตัดสินคุณภาพของสารสนเทศที่มีความเหมาะสมกับการใช้งาน โดยคำนึงถึงประโยชน์และประสิทธิภาพในการใช้งานของสารสนเทศแต่ละประเภท	**“ความรู้และความสามารถในการตัดสินคุณภาพ ความสอดคล้อง ประโยชน์ หรือประสิทธิภาพของสารสนเทศที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร ในระบบดิจิทัลได้”**
	California Emerging Technology Fund (2008)	การตัดสินเกี่ยวกับคุณภาพ ความสัมพันธ์ หรือประโยชน์ของสารสนเทศดิจิทัล โดยการตัดสินตามความทันสมัย ความเหมาะสมและความเพียงพอของสารสนเทศและแหล่งข้อมูล		
	Somerville et al. (2008)	การตัดสินว่าสารสนเทศสามารถตอบสนองความต้องการของปัญหา ภายใต้อคติ การกำหนดของระยะเวลา รวมถึงความเกี่ยวข้องอื่น ๆ เช่น การพิจารณาว่าเว็บเพจและบทความวารสารออนไลน์นี้มีประโยชน์ การประเมินว่าฐานข้อมูลมีข้อมูลที่เป็นปัจจุบันและตรงประเด็นเหมาะสม และการตัดสินใจว่าสามารถรวบรวมข้อมูลได้ครอบคลุมเรื่องที่ต้องการวิจัย		
	UNESCO (2008)	การตัดสินเกี่ยวกับคุณภาพ ความสัมพันธ์ ประโยชน์ ประสิทธิภาพ อคติ และข้อจำกัดทางด้านระยะเวลาของสารสนเทศ		
	น้ำทิพย์ วิภาวิน	การประเมินระดับของสารสนเทศที่สอดคล้อง		

องค์ประกอบ	แหล่งข้อมูล	นิยามจากแหล่งข้อมูล	นิยามจากการสังเคราะห์	การกำหนดนิยามที่ปรับแก้ไขจากข้อเสนอแนะ
	(2552)	กับความต้องการ		
5) การสื่อสาร	ETS (2002)	-	การสื่อสารสารสนเทศ (information communication) หมายถึง การสร้างวิธีการและการแลกเปลี่ยนสารสนเทศที่มีความเหมาะสมกับบริบทของผู้รับสารสนเทศ วัดจาก 1. การสร้างสารสนเทศที่มีความเหมาะสมกับบริบทของผู้รับสารสนเทศ คือ สามารถพัฒนาสารสนเทศให้มีความเหมาะสมและสอดคล้องกับบริบทของผู้รับสารสนเทศ 2. การแลกเปลี่ยนสารสนเทศที่มีความเหมาะสมกับบริบทของผู้รับสารสนเทศ คือ สามารถส่งผ่านสารสนเทศ โดยใช้วิธีการหรือช่องทางต่าง ๆ ที่มีความเหมาะสมตามประเภทของสารสนเทศ โดยคำนึงถึงบริบทของผู้รับสารสนเทศเป็นสำคัญ	**การสื่อสารสารสนเทศ** (information communication) หมายถึง การแลกเปลี่ยนสารสนเทศโดยการสร้างและเผยแพร่ข้อมูลข่าวสารที่มีความเหมาะสมกับบริบทของผู้รับสารสนเทศในระบบดิจิทัล โดยมีนิยามปฏิบัติการคือ **“ความรู้และความสามารถในการจัดทำและเผยแพร่สารสนเทศในระบบดิจิทัลเพื่อตอบสนองความต้องการของกลุ่มบุคคลและบริบทต่างๆที่เป็นการเคารพสิทธิ์ความเป็นส่วนตัวได้”**
	Katz and Macklin (2007)	การเผยแพร่ข้อมูลสารสนเทศและความรู้ไปยังผู้อื่นหรือกลุ่มผู้คน		
	Markauskaite (2007)	การจัดทำและเผยแพร่ผลงานวิจัย โดยใช้เครื่องมือดิจิทัลและเครือข่ายอินเตอร์เน็ตเพื่อนำเสนอแนวทางแก้ไขในรูปแบบต่าง ๆ แก่หลากหลายบุคคล การเคารพในลิขสิทธิ์ความเป็นส่วนตัว		
	California Emerging Technology Fund (2008)	การเผยแพร่ข้อมูลสารสนเทศ เพื่อตอบสนองความต้องการของกลุ่มผู้คน โดยใช้เครื่องมือดิจิทัลในการสื่อสาร ปรับตัว และนำเสนอสารสนเทศอย่างถูกต้องและเหมาะสมในบริบทต่าง ๆ ในสภาพแวดล้อมที่เป็นดิจิทัล		
	Somerville (2008)	การเผยแพร่สารสนเทศให้เหมาะสมกับกลุ่มผู้รับในรูปแบบของดิจิทัลที่มีประสิทธิภาพ เช่น การจัดรูปแบบเอกสารเพื่อให้เป็นประโยชน์กับกลุ่มใดกลุ่มหนึ่ง การเลือกและจัดระเบียบภาพนิ่งสำหรับการนำเสนอที่		

องค์ประกอบ	แหล่งข้อมูล	นิยามจากแหล่งข้อมูล	นิยามจากการสังเคราะห์	การกำหนดนิยามที่ปรับแก้ไขจากข้อเสนอแนะ
		แตกต่างกันเพื่อผู้ชมที่แตกต่างกัน การออกแบบใบปลิวเพื่อโฆษณากับกลุ่มผู้ใช้ที่แตกต่างกัน เป็นต้น		
	UNESCO (2008)	การถ่ายทอดสารสนเทศและความรู้ไปสู่บุคคลหรือกลุ่มต่าง ๆ		
	น้ำทิพย์ วิภาวิน (2552)	การสื่อความสารสนเทศในบริบทสภาพแวดล้อมทางเทคโนโลยีสารสนเทศ		

สรุปผลการกำหนดองค์ประกอบและนิยามสามารถสรุปได้เป็น 5 องค์ประกอบดังนี้

1. การเข้าถึงสารสนเทศ (information accessibility) หมายถึง ความรู้และความสามารถในการสืบค้นและรวบรวมสารสนเทศรวมทั้งอธิบายหรือแสดงวิธีการเรียกสารสนเทศที่มีอยู่กลับคืนมาเพื่อการใช้งานในระบบดิจิทัลได้

2. การจัดการสารสนเทศ (information management) หมายถึง ความรู้และความสามารถในการจัดระบบหมวดหมู่ของสารสนเทศในระบบดิจิทัลให้ง่ายต่อการการเรียกค้นและการประยุกต์ใช้ให้เหมาะสมกับงานแต่ละประเภทได้

3. การบูรณาการสารสนเทศ (information integration) หมายถึง ความรู้และความสามารถในการแปลความหมาย สรุป เปรียบเทียบ จำแนกความเหมือนและความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆในระบบดิจิทัลได้

4. การประเมินสารสนเทศ (information evaluation) หมายถึง ความรู้และความสามารถในการตัดสินคุณภาพ ความสอดคล้อง ประโยชน์ หรือประสิทธิภาพของสารสนเทศที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร ในระบบดิจิทัลได้

5.การสื่อสารสารสนเทศ (information communication) หมายถึง ความรู้และความสามารถในการจัดทำและเผยแพร่สารสนเทศในระบบดิจิทัลเพื่อตอบสนองความต้องการของกลุ่มบุคคลและบริบทต่างๆที่เป็นการเคารพสิทธิ์ความเป็นส่วนตัวได้

ผู้วิจัยได้ใช้นิยามเชิงปฏิบัติการทั้ง 5 องค์ประกอบดังกล่าวไปสร้างข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 โดยพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ต่อไป

ตอนที่ 2 ผลการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ผลการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี จำแนกออกเป็น 4 ประเด็นโดยมีรายละเอียดดังต่อไปนี้

2.1. ผลการสร้างข้อสอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ตามนิยามเชิงปฏิบัติการที่กำหนดไว้

ผลการสร้างข้อสอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ตามนิยามเชิงปฏิบัติการที่กำหนดไว้โดยทั้ง 5 องค์ประกอบมีจำนวนข้อสอบเท่ากันคิดเป็นองค์ประกอบละ 52 ข้อ รวมทั้งสิ้น 260 ข้อ รายละเอียดของจำนวนข้อสอบและตัวอย่างข้อสอบดังตาราง 4.3 และภาพ 4.1-4.5

ตาราง 4.3 สรุปจำนวนข้อสอบจำแนกตามองค์ประกอบของการวัด

องค์ประกอบ	นิยาม	จำนวน (ข้อ)
องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ (Information accessibility)	ความรู้และความสามารถในการสืบค้นและรวบรวมสารสนเทศ รวมทั้งอธิบายหรือแสดงวิธีการเรียกสารสนเทศที่มีอยู่กลับคืนมาเพื่อการใช้งานในระบบดิจิทัลได้	52
องค์ประกอบที่ 2 การจัดการสารสนเทศ (Information management)	ความรู้และความสามารถในการจัดระบบหมวดหมู่ของสารสนเทศในระบบดิจิทัลให้ง่ายต่อการการเรียกค้นและการประยุกต์ใช้ให้เหมาะสมกับงานแต่ละประเภทได้	52
องค์ประกอบที่ 3 การบูรณาการสารสนเทศ (Information integration)	ความรู้และความสามารถในการแปลความหมาย สรุป เปรียบเทียบ จำแนกความเหมือนและความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆในระบบดิจิทัลได้	52
องค์ประกอบที่ 4 การประเมินสารสนเทศ (Information evaluation)	ความรู้และความสามารถในการตัดสินคุณภาพ ความสอดคล้อง ประโยชน์ หรือประสิทธิภาพของสารสนเทศที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร ในระบบดิจิทัลได้	52
องค์ประกอบที่ 5 การสื่อสารสารสนเทศ (Information communication)	ความรู้และความสามารถในการจัดทำและเผยแพร่สารสนเทศในระบบดิจิทัลเพื่อตอบสนองความต้องการของกลุ่มบุคคลและบริบทต่างๆที่เป็นการเคารพสิทธิ์ความเป็นส่วนตัวได้	52
	รวม	260

ตัวอย่างข้อสอบองค์ประกอบที่ 1 การเข้าถึงสารสนเทศ (information accessibility)

1.หากท่านต้องการสืบค้นหนังสือเกี่ยวกับ การเพาะเลี้ยงเนื้อเยื่อพืช (Plant tissue culture) จากห้องสมุดผ่านระบบสืบค้นที่เรียกว่า e-library ท่านจะเลือกคำค้นประเภทใดลงในช่องสืบค้น (search) เพื่อให้ได้หนังสือที่ตรงตามความต้องการมากที่สุด

1. เลข ISBN/เลข ISSN ของหนังสือที่สนใจ
2. เทคนิคที่เกี่ยวข้องกับการเพาะพันธุ์พืช
3. รายชื่อผู้แต่งหนังสือและปีพ.ศ.ที่พิมพ์เผยแพร่
4. รายชื่ออาจารย์ผู้สอนรายวิชาการเพาะเลี้ยงเนื้อเยื่อพืช
5. สำนักพิมพ์หนังสือทางด้านการเกษตรและขยายพันธุ์พืช

เฉลย 2

ภาพ 4.1 ตัวอย่างของข้อสอบในองค์ประกอบที่ 1

ตัวอย่างข้อสอบองค์ประกอบที่ 2 การจัดการสารสนเทศ (information management)

12. หากแก้วต้องการบันทึกเอกสารในไฟล์ที่มีขนาดใหญ่มากกว่า 700 MB โดยภายหลังจากการบันทึกไฟล์ลงอุปกรณ์แล้วสามารถลบและบันทึกซ้ำได้ แก้วควรเลือกใช้อุปกรณ์ใดในการจัดเก็บ

1. CD-R
2. CD-RW
3. DVD-R
4. DVD-RW
5. ถูกทั้ง 3 และ 4

เฉลย 4

ภาพ 4.2 ตัวอย่างของข้อสอบในองค์ประกอบที่ 2

ตัวอย่างข้อสอบองค์ประกอบที่ 3 การบูรณาการสารสนเทศ (information integration)

จากข้อมูลสารสนเทศที่กำหนดให้ ในตาราง จงตอบคำถามข้อที่ 20

รุ่น	Spec
X100	Size : 9.7 นิ้ว CPU Speed : 2.15 GHZ Memory : Ram Size 4 GB S Pen ปากกาอัจฉริยะ กล้องหลัง : -
X200	Size : 8.0 นิ้ว CPU Speed : 1.3 GHZ Memory : Ram Size 3 GB กล้องหลัง : 13 ล้านพิกเซล
X300	Size : 9.7 นิ้ว CPU Speed : 1.9 GHZ Memory : Ram Size 3 GB กล้องหลัง : 8 ล้านพิกเซล

20.หากผู้ปกครองของท่านต้องการคำแนะนำในการเลือกซื้อ tablet สำหรับใช้ในการอ่านข่าวสารและรับชมทีวีออนไลน์ ท่านจะแนะนำ tablet รุ่นใดที่มีคุณสมบัติ (spec) เหมาะสมกับการใช้งานมากที่สุด เพราะเหตุใด

1. X200 ดีกว่า X100 เพราะ มีกล้องหลัง
2. X300 ดีกว่า X100 เพราะ มีกล้องหลัง
3. X100 ดีกว่า X200 เพราะ หน้าจอใหญ่กว่า
4. X300 ดีกว่า X200 เพราะ CPU มีความไวกว่า
5. X200 ดีกว่า X300 เพราะ กล้องหลังมีกำลังขยายสูงกว่า

เฉลย 3

ภาพ 4.3 ตัวอย่างของข้อสอบในองค์ประกอบที่ 3

ตัวอย่างข้อสอบองค์ประกอบที่ 4 การประเมินสารสนเทศ (information evaluation)

25.เมื่อท่านได้รับสารสนเทศดังภาพผลิตภัณฑ์ดังกล่าว ท่านจะทราบได้อย่างไรว่าผลิตภัณฑ์นี้มีความน่าเชื่อถือ

ผลิตภัณฑ์อาหารเสริมสกัดเข้มข้น

CR Herbal และ 99Z อาหารเสริมไทย

สูตรต้นตำรับโบราณหลากหลายชนิดผสานด้วยเทคโนโลยีเพื่อให้ได้อาหารเสริมสกัดเข้มข้น

ได้รับมาตรฐาน GMP และ อย.(11 – 1 – 12222 – 2 – 2222)

ช่วยดูแลผู้ป่วยที่ทำเคมีบำบัด หรือการฉายแสงฆ่าโรคร้าย

ซึ่งได้ผลตอบรับจากผู้ป่วยเป็นที่น่าพอใจ

http://www.herbal99Z.net

1. ชื่อเสียงของผลิตภัณฑ์
2. มาตรฐาน GMP ที่ได้รับ
3. ที่อยู่ของเว็บไซต์ที่ชัดเจนสามารถเข้าไปตรวจสอบได้
4. เลข อย.ที่สามารถตรวจสอบได้จากฐานข้อมูลออนไลน์
5. ขึ้นอยู่กับผู้แนะนำว่าเป็นใคร หากเป็นบุคคลที่น่าเชื่อถือแชร์มาให้ก็จะน่าเชื่อถือมากขึ้น

เฉลย 4

ภาพ 4.4 ตัวอย่างของข้อสอบในองค์ประกอบที่ 4

ตัวอย่างข้อสอบองค์ประกอบที่ 5 การสื่อสารสารสนเทศ (information communication)

30. หากท่านวางแผนไว้ว่าหลังจากสำเร็จการศึกษาระดับปริญญาตรี จะศึกษาต่อในประดับปริญญาโทที่ต่างประเทศ ท่านจึงส่งหลักฐานการสมัครเรียนไปยังมหาวิทยาลัยที่ต้องการเข้าศึกษา แต่ก่อนหมดเขตรับสมัคร 2 วัน พบว่าขาดหลักฐานการสมัครเรียนจำนวน 1 รายการ ท่านควรส่งหลักฐานการสมัครเรียนเพิ่มเติมผ่านทางช่องทางใด

1. อีเมล (E-mail)
2. แฟนเพจ (Fanpage)
3. โทรสาร (fax)
4. ไปรษณีย์
5. โทรศัพท์

เฉลย 1

ภาพ 4.5 ตัวอย่างของข้อสอบในองค์ประกอบที่ 5

2.2 ผลการตรวจสอบความตรงเชิงเนื้อหาก่อนนำไปทดลองใช้

จากการตรวจสอบความตรงเชิงเนื้อหาโดยผู้ทรงคุณวุฒิทั้งหมด 15 ท่าน ได้จำแนกแบบทดสอบเป็น 5 ชุด ที่แต่ละชุดประกอบด้วยข้อสอบครบทั้ง 5 องค์ประกอบ และมีผู้ทรงคุณวุฒิตรวจสอบความตรงเชิงเนื้อหาชุดละ 3 ท่าน ผลการตรวจสอบความตรงเชิงเนื้อหา พบว่า ข้อสอบในองค์ประกอบที่ 1 มีค่า IOC อยู่ระหว่าง 0.00 ถึง 1.00 โดยมีข้อสอบที่ผ่านเกณฑ์คิดเป็นร้อยละ 76.92 องค์ประกอบที่ 2 มีค่า IOC อยู่ระหว่าง 0.00 ถึง 1.00 โดยมีข้อสอบที่ผ่านเกณฑ์คิดเป็นร้อยละ 73.08 องค์ประกอบที่ 3 มีค่า IOC อยู่ระหว่าง -0.33 ถึง 1.00 โดยมีข้อสอบที่ผ่านเกณฑ์คิดเป็นร้อยละ 76.92 องค์ประกอบที่ 4 มีค่า IOC อยู่ระหว่าง 0.00 ถึง 1.00 โดยมีข้อสอบที่ผ่านเกณฑ์คิดเป็นร้อยละ 86.54 และองค์ประกอบที่ 5 มีค่า IOC อยู่ระหว่าง 0.67 ถึง 1.00 โดยมีข้อสอบที่ผ่านเกณฑ์คิดเป็นร้อยละ 79.61 ดังตาราง 4.4

ตาราง 4.4 ผลการตรวจสอบความตรงเชิงเนื้อหาจำแนกตามองค์ประกอบ

องค์ประกอบ	การตรวจสอบความตรงเชิงเนื้อหา		รวม
	จำนวนข้อสอบที่ผ่านเกณฑ์ (ร้อยละ)	จำนวนข้อสอบที่ไม่ผ่านเกณฑ์ (ร้อยละ)	
องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ	40 (76.92)	12 (23.08)	52 (100.00)
องค์ประกอบที่ 2 การจัดการสารสนเทศ	38 (73.08)	14 (26.92)	52 (100.00)
องค์ประกอบที่ 3 การบูรณาการสารสนเทศ	40 (76.92)	12 (23.08)	52 (100.00)
องค์ประกอบที่ 4 การประเมินสารสนเทศ	45 (86.54)	7 (13.46)	52 (100.00)
องค์ประกอบที่ 5 การสื่อสารสารสนเทศ	44 (84.62)	8 (15.38)	52 (100.00)
รวม	207 (79.61)	53 (20.39)	260 (100.00)

ภายหลังจากการวิเคราะห์ค่า IOC ผู้วิจัยจึงได้ปรับปรุงข้อสอบตามข้อเสนอแนะของผู้ทรงคุณวุฒิในแต่ละองค์ประกอบ แล้วจึงนำไปทดลองใช้กับกลุ่มตัวอย่างที่เป็น นิสิตนักศึกษา โดยมีตัวอย่างผลการปรับปรุงดังแสดงในตาราง 4.5 - 4.9

ตาราง 4.5 ตัวอย่างข้อสอบในองค์ประกอบที่ 1 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง

องค์ประกอบ	ข้อสอบก่อนการปรับปรุง	ข้อเสนอแนะ
องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ	การที่มีแนวคิดห้องสมุดเสมือน (Virtual Library) จะทำงานได้อย่างมีประสิทธิภาพนั้น ขั้นตอนแรกต้องลดการทำงานของคนและอาศัยเทคโนโลยีการเรียกค้นคืนสารสนเทศที่มีประสิทธิภาพมาทดแทน" จากคำกล่าวข้างต้นท่านเห็นด้วยหรือไม่ เพราะเหตุใด 1. ไม่เห็นด้วยเพราะห้องสมุดยังต้องมีคนในการปฏิบัติหน้าที่อยู่ถึงแม้จะมีเทคโนโลยีเข้ามาช่วยแล้วก็ตาม 2. เห็นด้วย เพราะห้องสมุดเสมือนหลายแห่งที่เปิดให้บริการไม่มีเจ้าหน้าที่คอยให้บริการ เช่น www.vlib.org 3*. เห็นด้วยที่ห้องสมุดเสมือนต้องลดการทำงานของคนในส่วนการให้บริการ แต่ยังจำเป็นต้องมีคนหน้าที่เป็นผู้เชี่ยวชาญคอยรวบรวม วิเคราะห์ ดูแล และพัฒนากระบวนการการสื่อสารและการกระบวนการทำงานของเทคโนโลยีค้นคืนอยู่ 4. เห็นด้วยที่ห้องสมุดเสมือนไม่จำเป็นต้องใช้คน เพราะเทคโนโลยีการค้นคืนสมัยใหม่สามารถพัฒนาการทำงานได้โดยไม่ต้องมีคนมาเกี่ยวข้อง 5. เห็นด้วยเพราะจะได้ประหยัดค่าใช้จ่าย	เหตุผลตัวเลือก 5 สั้นไปผู้สอบอาจจะไม่ตอบ ควรเพิ่มข้อความให้เหตุผล
	ข้อสอบหลังการปรับปรุง	
	การที่มีแนวคิดห้องสมุดเสมือน (Virtual Library) จะทำงานได้อย่างมีประสิทธิภาพนั้น ขั้นตอนแรกต้องลดการทำงานของคน และอาศัยเทคโนโลยีการเรียกค้นคืนสารสนเทศที่มีประสิทธิภาพมาทดแทน" จากคำกล่าวข้างต้นท่านเห็นด้วยหรือไม่ เพราะเหตุใด 1. ไม่เห็นด้วย เพราะ เป็นการสิ้นเปลืองงบประมาณในการบำรุงรักษาระบบเทคโนโลยีสารสนเทศ 2. ไม่เห็นด้วย เพราะห้องสมุดยังต้องมีคนในการปฏิบัติหน้าที่อยู่ ถึงแม้จะมีเทคโนโลยีเข้ามาช่วยแล้วก็ตาม 3. เห็นด้วย เพราะห้องสมุดเสมือนหลายแห่งที่เปิดให้บริการไม่มี เจ้าหน้าที่คอยให้บริการ เช่น www.vlib.org 4. เห็นด้วยที่ห้องสมุดเสมือนไม่จำเป็นต้องใช้คน เพราะเทคโนโลยี การค้นคืนสมัยใหม่สามารถพัฒนาการท างานได้โดยไม่ต้องมีคนมาเกี่ยวข้อง 5*. เห็นด้วย เพราะห้องสมุดเสมือนต้องลดการทำงานของคนในส่วนการให้บริการ แต่ยังจำเป็นต้องมีคนทำหน้าที่รวบรวม วิเคราะห์ ดูแล และพัฒนาระบบเทคโนโลยีการค้นคืน	

ตาราง 4.6 ตัวอย่างข้อสอบในองค์ประกอบที่ 2 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง

<table>
<tr><th>องค์ประกอบ</th><th>ข้อสอบก่อนการปรับปรุง</th><th>ข้อเสนอแนะ</th></tr>
<tr><td rowspan="3">องค์ประกอบที่ 2
การจัดการสารสนเทศ</td><td>บงกชกำลังทำโครงงานเรื่องเทคโนโลยีกับชีวิตประจำวันอยู่ที่ร้านคอมพิวเตอร์ใกล้มหาวิทยาลัย โดยประกอบด้วย รายงาน และหนังสั้น ซึ่งจะต้องนำผลงานที่จัดทำขึ้นไปปรึกษาอาจารย์ที่ปรึกษาและกลับมาแก้ไขจนกว่าผลงานจะเสร็จสมบูรณ์ บงกชควรบันทึกไฟล์งานที่ไว้ในอุปกรณ์ตามข้อใด
1. Hard Disk
2. Floppy Disk
3. CD-RW
4. DVD-R
5*. DVD-RAM</td><td>ปัจจุบันฮาร์ดแวร์ที่นำมาใช้สำหรับบันทึกข้อมูลเปลี่ยนแปลงไป ดังนั้นตัวเลือกควรสอบถามจากความนิยมใช้ในปัจจุบัน เช่น External HDD หรือ Thumb drive เป็นต้น</td></tr>
<tr><td colspan="2">ข้อสอบหลังการปรับปรุง</td></tr>
<tr><td colspan="2">บงกชกำลังทำโครงงานเรื่องเทคโนโลยีกับชีวิตประจำวันอยู่ที่ร้านคอมพิวเตอร์ใกล้มหาวิทยาลัย โดยประกอบด้วย รายงาน และหนังสั้น ซึ่งจะต้องนำผลงานที่จัดทำขึ้นไปปรึกษาอาจารย์ที่ปรึกษาและกลับมาแก้ไขจนกว่าผลงานจะเสร็จสมบูรณ์ บงกชควรบันทึกไฟล์งานไว้ในอุปกรณ์ตามข้อใด
1. CD-RW
2. DVD-R
3. Hard Disk
4*. Flash Drive
5. SD Memory Card</td></tr>
</table>

ตาราง 4.7 ตัวอย่างข้อสอบในองค์ประกอบที่ 3 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง

องค์ประกอบ	ข้อสอบก่อนการปรับปรุง	ข้อเสนอแนะ
องค์ประกอบที่ 3 การบูรณาการสารสนเทศ	จากภาพเป็นร้อยละของผู้ตอบแบบสำรวจ จำแนกตามระดับรายได้เฉลี่ยต่อเดือน ข้อใดสรุปสารสนเทศไม่ถูกต้อง (ที่มา https://www.etda.or.th/publishing-detail/thailand-internet-user-profile-2017.html) 1. ผู้ตอบแบบสำรวจส่วนใหญ่มีรายได้ 20,001-30,000 บาท 2. ผู้มีรายได้ 15,001-20,000 บาท มีจำนวนร้อยละ 15.8 3. ผู้มีรายได้ 70,001-90,000 บาท มีจำนวนน้อยที่สุด 4*. ผู้มีรายได้น้อยกว่า 10,000 บาท มีจำนวนน้อยกว่าผู้มีรายได้ 10,001-15,000 บาท 5. ผู้มีรายได้มากกว่า 70,000 บาท มีจำนวนมากกว่าผู้มีรายได้ 50,001-70,000 บาท	ควรเรียงลำดับความสั้น ยาวของตัวเลือก
	ข้อสอบหลังการปรับปรุง	
	จากภาพเป็นร้อยละของผู้ตอบแบบสำรวจ จำแนกตามระดับรายได้เฉลี่ยต่อเดือน ข้อใดสรุปสารสนเทศไม่ถูกต้อง (ที่มา https://www.etda.or.th/publishing-detail/thailand-internet-user-profile-2017.html) 1. ผู้มีรายได้ 70,001-90,000 บาท มีจำนวนน้อยที่สุด 2. ผู้มีรายได้ 15,001-20,000 บาท มีจำนวนร้อยละ 15.8 3. ผู้ตอบแบบสำรวจส่วนใหญ่มีรายได้ 20,001-30,000 บาท 4*. ผู้มีรายได้น้อยกว่า 10,000 บาท มีจำนวนน้อยกว่าผู้มีรายได้ 10,001-15,000 บาท 5. ผู้มีรายได้มากกว่า 70,000 บาท มีจำนวนมากกว่าผู้มีรายได้ 50,001-70,000 บาท	

ตาราง 4.8 ตัวอย่างข้อสอบในองค์ประกอบที่ 4 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง

องค์ประกอบ	ข้อสอบก่อนการปรับปรุง	ข้อเสนอแนะ
องค์ประกอบที่ 4 **การประเมินสารสนเทศ**	กรณีที่ท่านต้องการทราบอัตราค่าจ้างขั้นต่ำ ข้อมูลจากเว็บไซต์ใดที่มีความน่าเชื่อที่สุด 1*. ตารางอัตราค่าจ้างขั้นต่ำใหม่ ตามประกาศคณะกรรมการค่าจ้าง เรื่อง อัตราค่าจ้างขั้นต่ำ (ฉบับที่ 9) ซึ่งได้ประกาศให้มีผลใช้บังคับ ตั้งแต่วันที่ 1 เมษายน 2561 จาก www.mol.go.th 2. สรุปแล้ว!ปรับขึ้นค่าจ้างปี 61 เพิ่มทุกจังหวัด 5 – 22 บาท จาก http://www.jeban.com/board_all.php 3. แรงงานร้องขึ้นค่าจ้างขั้นต่ำ 360 บาท เท่ากันทั้งประเทศจาก https://mafiasubthai.blogspot.com/ 4. ปรับขึ้นค่าแรงขั้นต่ำทั่วประเทศ 5 – 22 บาท 'ภูเก็ต-ชลบุรี-ระยอง'สูงสุด 330 บาท จาก https://www.lnwshop.com/webboard 5. เพิ่มค่าแรงขั้นต่ำ คำตอบสุดท้ายสำหรับผู้ใช้แรงงานไทย? จาก https://board.postjung.com	โจทย์ต้องการให้พิจารณาชื่อเว็บไซต์ จึงควรพิจารณาเฉพาะชื่อเว็บไซต์ เนื่องจากข้อความจะชี้นำคำตอบ
	ข้อสอบหลังการปรับปรุง	
	กรณีที่ท่านต้องการทราบอัตราค่าจ้างขั้นต่ำ ข้อมูลจากเว็บไซต์ใดที่มีความน่าเชื่อที่สุด 1*. อัตราค่าจ้างขั้นต่ำใหม่เริ่มตั้งแต่วันที่ 1 เมษายน 2561 จาก www.mol.go.th 2. เพิ่มค่าแรงขั้นต่ำ คำตอบสุดท้ายสำหรับผู้ใช้แรงงานไทย? จาก https://board.postjung.com 3. แรงงานร้องขึ้นค่าจ้างขั้นต่ำ 360 บาท เท่ากันทั้งประเทศจาก https://mafiasubthai.blogspot.com/ 4. สรุปแล้ว!ปรับขึ้นค่าจ้างปี 61 เพิ่มทุกจังหวัด 5 – 22 บาท จาก http://www.jeban.com/board_all.php 5. ปรับขึ้นค่าแรงขั้นต่ำทั่วประเทศ 5 – 22 บาท 'ภูเก็ต-ชลบุรี-ระยอง'สูงสุด 330 บาท จาก https://www.lnwshop.com/webboard	

ตาราง 4.9 ตัวอย่างข้อสอบในองค์ประกอบที่ 5 ที่ผู้ทรงคุณวุฒิให้ข้อเสนอแนะและผลการปรับปรุง

องค์ประกอบ	ข้อสอบก่อนการปรับปรุง	ข้อเสนอแนะ
องค์ประกอบที่ 5 การสื่อสารสารสนเทศ	วริณได้รับหน้าที่ให้นำเสนอเรื่อง “วิธีการลดช่องว่างการอยู่ร่วมกันของผู้สูงอายุและเด็กในครอบครัว” โดยมีผู้เข้ารับฟังจำนวน 40 คน เป็นผู้สูงอายุและเด็กซึ่ง มีเงื่อนไขของผู้เข้ารับฟังดังนี้ ผู้สูงอายุ ต้องมีอายุระหว่าง 50-85 ปี และต้องพาหลานของตนมาด้วย 1 คน อายุระหว่าง 5-12 ปี ข้อใดเหมาะสมที่สุดกับการนำเสนอครั้งนี้ 1.การถาม-ตอบระหว่างกิจกรรมเพื่อสร้างความเข้าใจที่ชัดเจน ขจัดความสงสัยของผู้เข้าร่วมกิจกรรมได้ง่าย และมีการสรุปข้อมูลขั้นตอนที่เหมาะสมก่อนเลิกกิจกรรม 2.การนำเสนอด้วยสถานการณ์จริง มีกลุ่มอาสาสมัครแสดงวิธีการหรือขั้นตอนต่าง ๆ ที่ด้านหน้าลานกิจกรรมเพื่อให้ทุกคนเห็นได้อย่างชัดเจน 3.การนำเสนอด้วยข้อความที่อธิบายถึงวิธีการลดช่องว่างการอยู่ร่วมกันของผู้สูงอายุและเด็กในครอบครัวด้วยขั้นตอนที่ชัดเจนผ่านโปรแกรมการนำเสนอข้อมูล 4*การนำเสนอด้วยวีดิทัศน์และข้อความผ่านโปรแกรมการนำเสนอข้อมูล ประกอบการถาม-ตอบระหว่างกิจกรรม 5.การนำเสนอด้วยภาพนิ่งและกราฟิกเพื่อดึงดูดความสนใจของเด็ก ๆ ผ่านโปรแกรมประมวลผลคำ	ตัวเลือกอื่นก็เหมาะสมและปรับคำให้กระชับ
	ข้อสอบหลังการปรับปรุง	
	ในการนำเสนอความรู้เรื่อง “การดูแลสุขภาพของผู้สูงอายุ” โดยมีผู้เข้ารับฟัง จำนวน 40 คน เป็นผู้สูงอายุที่มีอายุ ระหว่าง 50-85 ปี ข้อใดเหมาะสมน้อยที่สุดกับการนำเสนอครั้งนี้ 1.การนำเสนอด้วยภาพนิ่งและกราฟิกเพื่อดึงดูดความสนใจ โปรแกรมประมวลผลคำ 2.การนำเสนอด้วยวีดิทัศน์และข้อความผ่านโปรแกรมการนำเสนอข้อมูล ประกอบการถาม-ตอบระหว่างกิจกรรม 3*.การนำเสนอด้วยข้อความที่อธิบายถึงการดูแลสุขภาพของผู้สูงอายุด้วยขั้นตอนที่ชัดเจนผ่านโปรแกรมการนำเสนอข้อมูล 4.การนำเสนอด้วยสถานการณ์จริง โดยมีกลุ่มอาสาสมัครแสดงวิธีการหรือ ขั้นตอนต่าง ๆ ที่ด้านหน้าลานกิจกรรมเพื่อให้ทุกคนเห็นได้อย่างชัดเจน 5.การถาม-ตอบระหว่างกิจกรรมเพื่อสร้างความเข้าใจที่ชัดเจน ขจัดความ สงสัยของผู้เข้าร่วมกิจกรรมได้ง่าย และมีการสรุปข้อมูลขั้นตอนที่เหมาะสม ก่อนเลิกกิจกรรม	

2.3 ผลการตรวจสอบคุณภาพของข้อสอบที่ได้จากการทดลองใช้

ผลการตรวจสอบคุณภาพของข้อสอบที่ได้จากการทดลองใช้กับนิสิตนักศึกษาตัวอย่างจำนวน 1,672 คน ตามแผนการสุ่มที่ได้จาก 5 สาขาวิชา ได้แก่ 1.ครุศาสตร์ ศึกษาศาสตร์ 2. สังคมศาสตร์ วารสารศาสตร์และสารสนเทศ 3. วิทยาศาสตร์ คณิตศาสตร์และสถิติ 4. วิศวกรรมศาสตร์ และ 5. สุขภาพและสวัสดิการ จาก 6 สถาบัน ได้แก่ 1) จุฬาลงกรณ์มหาวิทยาลัย 2) มหาวิทยาลัยศรีนครินทรวิโรฒ 3) มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ 4) วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้า 5) มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา และ 6) มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ ในพระบรมราชูปถัมภ์ (รายละเอียดนำเสนอในบทที่ 3) จำแนกเป็น 3 ประเด็นคือ 1) ผลการวิเคราะห์คุณภาพของแบบสอบ (Test) ทั้งฉบับตามทฤษฎี CTT 2) ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ตามทฤษฎี CTT และ 3) ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ตามทฤษฎี IRT โดยแต่ละประเด็นมีรายละเอียดดังนี้

2.3.1 ผลการวิเคราะห์คุณภาพของแบบสอบ (Test) ทั้งฉบับตามทฤษฎี CTT

ผลการวิเคราะห์คุณภาพของแบบสอบ (Test) ทั้งฉบับ ตามทฤษฎี CTT ในภาพรวมพบว่า ค่าเฉลี่ยความยาก (p) ของข้อสอบเท่ากับ 0.483 ซึ่งถือว่าอยู่ในระดับปานกลาง ขณะที่ค่าอำนาจจำแนกเฉลี่ย (r) ของข้อสอบเท่ากับ 0.273 ถือว่าแบบสอบจำแนกผู้สอบได้ และค่าความเที่ยง (R^2) ของแบบสอบทั้ง 10 ชุดมีค่าเท่ากับ 0.629 ถือว่ามีความเที่ยงที่อยู่ในเกณฑ์ที่ยอมรับได้ในระดับค่อนข้างสูง รายละเอียดดังตาราง 4.10

ตาราง 4.10 ผลการวิเคราะห์คุณภาพของแบบสอบ (Test) ทั้งฉบับตามทฤษฎี CTT จำแนกตามองค์ประกอบและชุดข้อสอบ

องค์ประกอบ	N	ค่าเฉลี่ยความยาก (p)	ค่าเฉลี่ยอำนาจจำแนก (r)	ความเที่ยง (R^2)
1	52	0.513	0.254	0.629
2	52	0.384	0.261	
3	52	0.491	0.277	
4	52	0.443	0.283	
5	52	0.497	0.293	
รวม	260	0.483	0.273	

แบบทดสอบชุดที่	N	ค่าเฉลี่ยความยาก (p)	ค่าเฉลี่ยอำนาจจำแนก (r)	ความเที่ยง (R)
1	35	0.478	0.299	0.715
2	35	0.474	0.268	0.630
3	35	0.401	0.239	0.550
4	35	0.431	0.262	0.627
5	35	0.416	0.248	0.517
6	35	0.457	0.274	0.695
7	35	0.440	0.253	0.643
8	35	0.417	0.240	0.653
9	35	0.461	0.240	0.537
10	35	0.483	0.306	0.725

2.3.2 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ตามทฤษฎี CTT

องค์ประกอบที่ 1 มีข้อสอบที่ผ่านเกณฑ์จำนวน 28 ข้อ จากทั้งหมด 52 ข้อ คิดเป็นร้อยละ 53.84 องค์ประกอบที่ 2 มีข้อสอบที่ผ่านเกณฑ์จำนวน 32 ข้อ จากทั้งหมด 52 ข้อ คิดเป็นร้อยละ 61.54, องค์ประกอบที่ 3 มีข้อสอบที่ผ่านเกณฑ์จำนวน 40 ข้อ จากทั้งหมด 52 ข้อ คิดเป็นร้อยละ 76.92, องค์ประกอบที่ 4 มีข้อสอบที่ผ่านเกณฑ์จำนวน 34 ข้อ จากทั้งหมด 52 ข้อ คิดเป็นร้อยละ 65.38 และองค์ประกอบที่ 5 มีข้อสอบที่ผ่านเกณฑ์จำนวน 42 ข้อ จากทั้งหมด 52 ข้อ คิดเป็นร้อยละ 80.77 เมื่อรวมทุกองค์ประกอบจำนวน 260 ข้อ มีข้อสอบที่ผ่านเกณฑ์จำนวน 176 ข้อ คิดเป็นร้อยละ 67.70 ดังมีรายละเอียดดังตารางที่ 4.11-4.15

ตาราง 4.11 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 1 ตามทฤษฎี CTT

องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ							
ข้อที่	P	r	ผลประเมิน	ข้อที่	p	r	ผลประเมิน
1	0.835	0.180	×	27	0.615	0.130	×
2	0.241	0.180	×	28	0.152	0.260	×
3	0.522	0.360	✓	29	0.853	0.300	×
4	0.494	0.380	✓	30	0.518	0.500	✓
5	0.631	0.550	✓	31	0.692	-0.040	×
6	0.630	0.200	✓	32	0.343	0.390	✓
7	0.840	0.290	×	33	0.355	0.130	×
8	0.346	0.170	×	34	0.548	0.310	✓
9	0.596	0.290	✓	35	0.734	0.410	✓
10	0.233	0.260	✓	36	0.097	0.280	×
11	0.123	0.080	×	37	0.573	0.380	✓
12	0.900	0.120	×	38	0.335	0.160	×
13	0.417	0.190	×	39	0.493	0.140	×
14	0.823	0.330	×	40	0.579	0.120	×
15	0.352	0.260	✓	41	0.437	0.280	✓
16	0.730	0.240	✓	42	0.656	0.380	✓
17	0.515	0.220	✓	43	0.425	0.160	×
18	0.538	0.260	✓	44	0.350	0.140	×
19	0.389	0.280	✓	45	0.856	0.120	×
20	0.526	0.370	✓	46	0.612	0.420	✓
21	0.597	0.220	✓	47	0.206	0.190	×
22	0.603	0.390	✓	48	0.790	0.310	✓
23	0.798	0.340	✓	49	0.329	0.120	×
24	0.272	0.150	×	50	0.335	0.290	✓
25	0.491	0.250	✓	51*	0.598	0.450	✓
26	0.579	0.300	✓	52*	0.142	0.040	×
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 28 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 53.84)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 24 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 46.16)							

หมายเหตุ * หมายถึง ข้อสอบแกนกลาง

✓ หมายถึง ผ่านเกณฑ์

× หมายถึง ไม่ผ่านเกณฑ์

ตาราง 4.12 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 2 ตามทฤษฎี CTT

องค์ประกอบที่ 2 การจัดการสารสนเทศ							
ข้อที่	P	r	ผลประเมิน	ข้อที่	p	r	ผลประเมิน
53	0.703	0.370	✓	79	0.268	0.480	✓
54	0.560	0.310	✓	80	0.311	0.110	×
55	0.192	0.080	×	81	0.506	0.380	✓
56	0.527	0.320	✓	82	0.658	0.090	×
57	0.752	0.330	✓	83	0.228	0.120	×
58	0.903	0.320	×	84	0.759	0.060	×
59	0.352	0.310	✓	85	0.355	0.470	✓
60	0.187	0.210	×	86	0.301	0.450	✓
61	0.409	0.140	×	87	0.415	0.140	×
62	0.664	0.270	✓	88	0.231	0.340	✓
63	0.041	0.130	×	89	0.134	0.290	×
64	0.535	0.300	✓	90	0.628	0.050	✓
65	0.123	0.220	×	91	0.530	0.200	✓
66	0.229	0.140	×	92	0.243	0.310	✓
67	0.452	0.150	×	93	0.287	0.340	✓
68	0.526	0.440	✓	94	0.768	0.290	✓
69	0.670	0.270	✓	95	0.131	0.050	×
70	0.503	0.420	✓	96	0.512	0.200	✓
71	0.329	0.170	×	97	0.312	0.310	✓
72	0.275	0.220	✓	98	0.483	0.360	✓
73	0.716	0.300	✓	99	0.154	0.200	×
74	0.136	0.200	×	100	0.412	0.200	✓
75	0.461	0.370	✓	101	0.238	0.290	✓
76	0.426	0.340	✓	102	0.464	0.430	✓
77	0.325	0.290	✓	103*	0.648	0.240	✓
78	0.170	0.410	×	104*	0.280	0.120	×
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 32 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 61.54)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 20 ข้อ จากทั้งหมด (ร้อยละ 38.46)							

หมายเหตุ * หมายถึง ข้อสอบแกนกลาง
✓ หมายถึง ผ่านเกณฑ์
× หมายถึง ไม่ผ่านเกณฑ์

ตาราง 4.13 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 3 ตามทฤษฎี CTT

องค์ประกอบที่ 3 การบูรณาการสารสนเทศ							
ข้อที่	P	r	ผลประเมิน	ข้อที่	p	r	ผลประเมิน
105	0.412	0.270	✓	131	0.408	0.370	✓
106	0.780	0.470	✓	132	0.439	0.350	✓
107	0.324	0.270	✓	133	0.542	0.220	✓
108	0.236	0.090	×	134	0.524	0.260	✓
109	0.522	0.500	✓	135	0.722	0.190	×
110	0.903	0.270	×	136	0.379	0.330	✓
111	0.284	0.170	×	137	0.451	0.010	×
112	0.392	0.160	×	138	0.500	0.280	✓
113	0.329	0.160	×	139	0.355	0.480	✓
114	0.767	0.380	✓	140	0.664	0.250	✓
115	0.523	0.270	✓	141	0.567	0.250	✓
116	0.182	0.200	×	142	0.250	0.260	✓
117	0.694	0.280	✓	143	0.274	0.400	✓
118	0.458	0.400	✓	144	0.689	0.250	✓
119	0.494	0.370	✓	145	0.550	0.250	✓
120	0.712	0.370	✓	146	0.587	0.250	✓
121	0.425	0.400	✓	147	0.318	0.260	✓
122	0.796	0.212	✓	148	0.668	0.400	✓
123	0.365	0.220	✓	149	0.793	0.250	✓
124	0.538	0.400	✓	150	0.393	0.130	×
125	0.142	0.100	×	151	0.354	0.140	×
126	0.514	0.320	✓	152	0.587	0.370	✓
127	0.627	0.280	✓	153	0.690	0.300	✓
128	0.295	0.340	✓	154	0.548	0.330	✓
129	0.485	0.360	✓	155*	0.137	-0.080	×
130	0.402	0.280	✓	156*	0.538	0.350	✓
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 40 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 76.92)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 12 ข้อ จากทั้งหมด (ร้อยละ 23.08)							

หมายเหตุ * หมายถึง ข้อสอบแกนกลาง
✓ หมายถึง ผ่านเกณฑ์
× หมายถึง ไม่ผ่านเกณฑ์

ตาราง 4.14 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 4 ตามทฤษฎี CTT

องค์ประกอบที่ 4 การประเมินสารสนเทศ							
ข้อที่	P	r	ผลประเมิน	ข้อที่	p	r	ผลประเมิน
158	0.714	0.380	✓	183	0.457	0.310	✓
158	0.186	0.150	×	184	0.140	0.500	×
159	0.873	0.510	✓	185	0.530	0.120	×
160	0.708	0.580	✓	186	0.426	0.090	×
161	0.269	0.040	×	187	0.313	0.270	✓
162	0.522	0.460	✓	188	0.337	0.140	×
163	0.358	0.180	×	189	0.728	0.240	✓
164	0.534	0.410	✓	190	0.265	0.360	✓
165	0.284	0.300	✓	191	0.253	0.390	✓
166	0.653	0.390	✓	192	0.304	0.370	✓
167	0.170	-0.060	×	193	0.286	0.420	✓
168	0.476	0.400	✓	194	0.378	0.250	✓
169	0.394	0.180	×	195	0.536	0.250	✓
170	0.264	0.390	✓	196	0.439	0.080	×
171	0.235	0.120	×	197	0.512	0.370	✓
172	0.700	0.250	✓	198	0.787	0.420	✓
173	0.215	0.150	×	199	0.506	0.250	✓
174	0.521	0.430	✓	200	0.731	0.250	✓
175	0.556	0.310	✓	201	0.368	0.080	×
176	0.239	0.140	×	202	0.509	0.460	✓
177	0.100	0.060	×	203	0.651	0.430	✓
178	0.159	0.050	×	204	0.761	0.440	✓
179	0.355	0.280	✓	205	0.741	0.420	✓
180	0.289	0.160	×	206	0.741	0.410	✓
181	0.272	0.100	×	207*	0.384	0.460	✓
182	0.664	0.240	✓	208*	0.258	0.350	✓
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 34 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 65.38)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 18 ข้อ จากทั้งหมด (ร้อยละ 34.62)							

หมายเหตุ * หมายถึง ข้อสอบแกนกลาง
✓ หมายถึง ผ่านเกณฑ์
× หมายถึง ไม่ผ่านเกณฑ์

ตาราง 4.15 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 5 ตามทฤษฎี CTT

องค์ประกอบที่ 5 การสื่อสารสารสนเทศ							
ข้อที่	P	r	ผลประเมิน	ข้อที่	p	r	ผลประเมิน
209	0.686	0.400	✓	235	0.622	0.350	✓
210	0.472	0.310	✓	236	0.628	0.260	✓
211	0.703	0.370	✓	237	0.682	0.300	✓
212	0.258	0.290	✓	238	0.463	0.190	×
213	0.351	0.340	✓	239	0.337	0.430	✓
214	0.647	0.440	✓	240	0.614	0.370	✓
215	0.409	0.300	✓	241	0.427	0.230	✓
216	0.386	0.080	×	242	0.771	0.570	✓
217	0.539	0.350	✓	243	0.331	0.210	✓
218	0.755	0.410	✓	244	0.597	0.230	✓
219	0.452	0.410	✓	245	0.567	0.300	✓
220	0.594	0.410	✓	246	0.536	0.350	✓
221	0.817	0.510	×	247	0.664	0.070	×
222	0.582	0.260	✓	248	0.250	0.240	✓
223	0.223	-0.040	×	249	0.593	0.230	✓
224	0.383	0.400	✓	250	0.231	0.300	✓
225	0.155	0.240	×	251	0.337	0.350	✓
226	0.359	0.100	✓	252	0.193	0.070	×
227	0.245	0.160	✓	253	0.431	0.240	✓
228	0.101	0.020	×	254	0.464	0.290	✓
229	0.420	0.280	✓	255	0.638	0.310	✓
230	0.733	0.360	✓	256	0.774	0.330	✓
231	0.686	0.300	✓	257	0.696	0.420	✓
232	0.810	0.430	✓	258	0.651	0.540	✓
233	0.230	0.080	×	259*	0.236	0.140	×
234	0.554	0.300	✓	260*	0.565	0.380	✓
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 42 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 80.77)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 10 ข้อ จากทั้งหมด (ร้อยละ 19.23)							

หมายเหตุ * หมายถึง ข้อสอบแกนกลาง
✓ หมายถึง ผ่านเกณฑ์
× หมายถึง ไม่ผ่านเกณฑ์

การวิเคราะห์คุณภาพของข้อสอบด้วยค่าความยากและค่าอำนาจจำแนกตามทฤษฎีการทดสอบแบบดั้งเดิม มีข้อจำกัดเกี่ยวกับดัชนีคุณภาพข้อสอบแปรเปลี่ยนไปตามกลุ่มผู้สอบและไม่สามารถปรับเทียบข้อสอบในกรณีที่กลุ่มผู้สอบต่างกลุ่มกัน ดังนั้นในการตรวจสอบคุณภาพข้อสอบในงานวิจัยเรื่องนี้ใช้ตัวอย่างมากเพียงพอและมีการใช้ข้อสอบร่วมระหว่างแบบสอบแต่ละชุด จึงได้ใช้วิธีการวิเคราะห์คุณภาพข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ(IRT) เป็นหลักในการคัดเลือกข้อสอบเข้าคลังข้อสอบ ดังมีรายละเอียดของผลการวิเคราะห์ในหัวข้อต่อไป

2.3.3 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ตามทฤษฎี IRT

ก่อนที่จะทำการการวิเคราะห์คุณภาพของข้อสอบรายข้อตามทฤษฎี IRT เพื่อให้ได้สารสนเทศที่มีความถูกต้อง ผู้วิจัยได้ตรวจสอบและเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลการตอบสนองข้อสอบ ทั้งแบบ 1 Pl 2 PL และ 3PL เพื่อเลือกโมเดลที่มีความถูกต้องมากที่สุดแล้วนำไปวิเคราะห์คุณภาพข้อสอบที่มีการปรับเทียบแบบสอบและการประเมินผลคุณภาพข้อสอบ ดังมีรายละเอียด ดังนี้

1) การตรวจสอบและเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลการตอบสนองข้อสอบ

ข้อสอบทั้งหมดถูกจัดเป็นชุด จำนวน 10 ฉบับ ฉบับละ 35 ข้อ โดยแบบสอบทั้ง 10 ฉบับจะมีข้อสอบร่วม (common item) เหมือนกัน 10 ข้อ และมีข้อสอบที่ไม่เหมือนกันอีกชุดละ 25 ข้อ (รายละเอียดของการจัดชุดข้อสอบแสดงดังตาราง 4.16 แบบสอบแต่ละฉบับจัดพิมพ์บนกระดาษและให้ตัวอย่างฝนคำตอบลงบนกระดาษคำตอบ (paper-and-pencil format) ซึ่งเป็นวิธีการดั้งเดิมสำหรับการพัฒนาคลังข้อสอบแบบปรับเหมาะด้วยทฤษฎีการตอบสนองข้อสอบ (Bergstrom & Lunz, 1999) นิสิตนักศึกษาแต่ละคนทำแบบสอบเพียงหนึ่งฉบับเท่านั้น ข้อมูลคำตอบของแบบสอบแต่ละฉบับถูกวิเคราะห์ด้วยโมเดลการวัดแบบ 1pl, 2pl และ 3pl เพื่อตรวจสอบความสอดคล้องกลมกลืนของโมเดล ดังแสดงในตาราง 4.17

ตาราง 4.16 โครงสร้างแบบสอบก่อนพัฒนาเป็นคลังข้อสอบแบบปรับเหมาะ

ฉบับที่	ข้อสอบร่วม (10 ข้อ)	ข้อสอบปกติ (25 ข้อ)	ข้อสอบปกติ (25 ข้อ)	ข้อสอบปกติ (25 ข้อ)	ข้อสอบปกติ (25 ข้อ)	ข้อสอบปกติ (25 ข้อ)	ข้อสอบปกติ (25 ข้อ)	ข้อสอบปกติ (25 ข้อ)	ข้อสอบปกติ (25 ข้อ)	ข้อสอบปกติ (25 ข้อ)	ข้อสอบปกติ (25 ข้อ)
1	✓	✓									
2	✓		✓								
3	✓			✓							
4	✓				✓						
5	✓					✓					
6	✓						✓				
7	✓							✓			
8	✓								✓		
9	✓									✓	
10	✓										✓

ตาราง 4.17 ผลการเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดล 1pl, 2pl, และ 3pl

ฉบับที่ 1	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl	ฉบับที่ 2	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl
n	181	181	181	n	176	176	176
p	35	70	105	p	35	70	105
G^2	7334.24	7172.88	7155.30	G^2	7207.67	7102.00	7074.99
AIC	7406.24	7312.79	7365.31	AIC	7279.67	7242.00	7284.99
BIC	7521.39	7536.69	7701.15	BIC	7393.80	7463.94	7617.89
1pl vs. 2pl	LRT = 161.45**	ΔR^2 = .0220 (2.20%)		1pl vs. 2pl	LRT = 105.66**	ΔR^2 = .0147 (1.47%)	
1pl vs. 3pl	LRT = 178.93**	ΔR^2 = .0244 (2.44%)		1pl vs. 3pl	LRT = 132.68**	ΔR^2 = .0184 (1.84%)	
2pl vs. 3pl	LRT = 17.48	ΔR^2 = .0024 (.24%)		2pl vs. 3pl	LRT = 27.01	ΔR^2 = .0038 (.38%)	
ฉบับที่ 3	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl	ฉบับที่ 4	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl
n	170	170	170	n	167	167	167
p	35	70	105	p	35	70	105
G^2	6678.59	6572.52	6554.39	G^2	6957.20	6888.82	6856.40
AIC	6750.59	6712.52	6764.39	AIC	7029.21	7028.81	7066.40
BIC	6863.48	6932.02	7093.65	BIC	7141.45	7247.07	7393.79
1pl vs. 2pl	LRT = 106.07**	ΔR^2 = .0159 (1.59%)		1pl vs. 2pl	LRT = 68.39**	ΔR^2 = .0098 (.98%)	
1pl vs. 3pl	LRT = 124.20**	ΔR^2 = .0186 (1.86%)		1pl vs. 3pl	LRT = 100.81**	ΔR^2 = .0145 (1.45%)	
2pl vs. 3pl	LRT = 18.13	ΔR^2 = .0027 (.27%)		2pl vs. 3pl	LRT = 32.42	ΔR^2 = .0047 (.47%)	
ฉบับที่ 5	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl	ฉบับที่ 6	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl
n	169	169	169	n	164	164	164
p	35	70	105	p	35	70	105
G^2	6790.38	6705.88	6677.88	G^2	6863.74	6721.30	6700.18
AIC	6862.37	6845.88	6887.88	AIC	6935.73	6861.31	6910.17
BIC	6975.05	7064.97	7216.52	BIC	7047.33	6861.31	7235.66
1pl vs. 2pl	LRT = 84.50**	ΔR^2 = .0124 (1.24%)		1pl vs. 2pl	LRT = 142.43**	ΔR^2 = .0208 (2.08%)	
1pl vs. 3pl	LRT = 112.49**	ΔR^2 = .0166 (1.66%)		1pl vs. 3pl	LRT = 163.56**	ΔR^2 = .0238 (2.38%)	
2pl vs. 3pl	LRT = 28.00	ΔR^2 = .0042 (.42%)		2pl vs. 3pl	LRT = 21.13	ΔR^2 = .0031 (.31%)	
ฉบับที่ 7	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl	ฉบับที่ 8	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl
n	166	166	166	n	164	164	164
p	35	70	105	p	35	70	105
G^2	7073.06	6953.76	6932.14	G^2	6865.30	6723.24	6700.24
AIC	7145.06	7093.75	7142.15	AIC	6937.31	6863.25	6910.23
BIC	7257.09	7311.59	7468.91	BIC	7048.90	7080.24	7235.72
1pl vs. 2pl	LRT = 119.31**	ΔR^2 = .0169 (1.69%)		1pl vs. 2pl	LRT = 142.06**	ΔR^2 = .0207 (2.07%)	
1pl vs. 3pl	LRT = 140.91**	ΔR^2 = .0199 (1.99%)		1pl vs. 3pl	LRT = 165.07**	ΔR^2 = .0240 (2.40%)	
2pl vs. 3pl	LRT = 21.60	ΔR^2 = .0031 (.31%)		2pl vs. 3pl	LRT = 23.02	ΔR^2 = .0034 (.34%)	
ฉบับที่ 9	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl	ฉบับที่ 10	โมเดล 1pl	โมเดล 2pl	โมเดล 3pl
n	160	160	160	n	155	155	155
p	35	70	105	p	35	70	105
G^2	6700.66	6625.22	6605.38	G^2	6404.00	6299.96	6261.52
AIC	6772.66	6765.23	6815.38	AIC	6475.99	6439.97	6471.53
BIC	6883.36	6980.49	7138.27	BIC	6585.56	6653.01	6791.09
1pl vs. 2pl	LRT = 75.43**	ΔR^2 = .0113 (1.13%)		1pl vs. 2pl	LRT = 104.03**	ΔR^2 = .0162 (1.62%)	
1pl vs. 3pl	LRT = 95.28*	ΔR^2 = .0142 (1.42%)		1pl vs. 3pl	LRT = 142.47**	ΔR^2 = .0222 (2.22%)	
2pl vs. 3pl	LRT = 19.85	ΔR^2 = .0030 (.30%)		2pl vs. 3pl	LRT = 38.44	ΔR^2 = .0061 (.61%)	

หมายเหตุ ** $p<.01$ * $p<.05$

ตาราง 4.17 แสดงให้เห็นว่า กลุ่มตัวอย่าง (n) ที่ทำแบบสอบแต่ละฉบับมีจำนวนไม่เท่ากัน ระหว่าง 155 – 181 คน ข้อมูลคำตอบของนิสิตนักศึกษาที่เป็นตัวอย่างแต่ละฉบับถูกตรวจสอบความสอดคล้องกลมกลืนกับโมเดลการวัดทั้ง 3 แบบ โดยที่โมเดลการวัดแต่ละแบบจะมีจำนวนพารามิเตอร์ (p) ที่ต้องประมาณค่าไม่เท่ากัน โมเดล 1pl ประมาณเฉพาะพารามิเตอร์ความยากจำนวน 35 ค่า (ข้อละ 1 ค่า) โมเดล 2pl ประมาณพารามิเตอร์ 70 ค่า (ข้อละ 2 ค่า) ประกอบด้วย พารามิเตอร์ความยากจำนวน 35 ค่า และพารามิเตอร์การจำแนกจำนวน 35 ค่า และโมเดล 3pl ประมาณพารามิเตอร์ 105 ค่า (ข้อละ 3 ค่า) โดยมีพารามิเตอร์การเดาเพิ่มขึ้นมาอีก 35 ค่า จากโมเดล 2pl

ผลการเปรียบเทียบความสอดคล้องกลมกลืนของแบบสอบแต่ละฉบับมีดังนี้ แบบสอบฉบับที่ 1 พบว่า ค่าดัชนี deviance (G^2) โมเดล 3pl มีค่าต่ำที่สุด แสดงว่าโมเดล 3pl สอดคล้องกลมกลืนกับข้อมูลมากที่สุด (best fit) ในขณะที่เมื่อพิจารณาค่า AIC และ BIC พบว่า โมเดล 2pl และโมเดล 1pl กลับเป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลมากที่สุด ตามลำดับ ผลการเปรียบเทียบที่ไม่เป็นไปในทิศทางเดียวกันนี้ ทำให้ผู้วิจัยจำเป็นต้องทดสอบนัยสำคัญของความสอดคล้องกลมกลืนของโมเดลทีละคู่ ด้วยการทดสอบอัตราส่วนความเป็นไปได้ (likelihood-ratio test: LRT) พบว่า โมเดล 2pl สอดคล้องกลมกลืนกับข้อมูลมากกว่าโมเดล 1pl อย่างมีนัยสำคัญทางสถิติ ($p<.01$) แม้ว่าโมเดล 1pl จะมีค่า BIC ต่ำที่สุดก็ตาม และสอดคล้องกลมกลืนกับข้อมูลไม่แตกต่างจากโมเดล 3pl ทั้งๆ ที่โมเดล 3pl มีค่า G^2 ต่ำสุด นั่นหมายความว่า ผลต่างของ G^2 ระหว่างโมเดล 3pl และโมเดล 2pl และผลต่างของ BIC ระหว่างโมเดล 1pl และโมเดล 2pl ไม่มากพอที่จะใช้ยืนยันความเป็นโมเดลที่สอดคล้องกลมกลืนมากที่สุดได้ นอกจากนี้เมื่อพิจารณาสัดส่วนสัมพัทธ์ของความแปรปรวนของค่าพารามิเตอร์ความสามารถตัวอย่างนิสิตนักศึกษา (θ_{ict}) ที่สามารถอธิบายได้ของโมเดลแต่ละคู่ (ΔR^2: de Ayala, 2009; Haberman, 1978) พบว่า พารามิเตอร์การเดาจำนวน 35 ค่าในโมเดล 3pl ไม่ได้เพิ่มความแปรปรวนที่สามารถอธิบายได้อย่างมีนัยสำคัญ เพียงแค่ 0.24% เมื่อเปรียบเทียบกับโมเดล 2pl ดังนั้นจึงสรุปว่า **โมเดล 2pl เป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลคำตอบของแบบสอบฉบับที่ 1 มากที่สุด**

ผลการเปรียบเทียบความสอดคล้องระหว่างโมเดล 1pl, 2pl และ 3pl ของแบบสอบฉบับที่ 2 ถึง 10 ไม่แตกต่างจากแบบสอบฉบับที่ 1 กล่าวคือ ค่า G^2 ของโมเดล 3pl ต่ำที่สุด เมื่อเปรียบเทียบกับโมเดล 2pl และโมเดล 3pl ค่า AIC ของโมเดล 2pl ต่ำที่สุด เมื่อเปรียบเทียบกับโมเดล 3pl และโมเดล 1pl และค่า BIC ของโมเดล 1pl ต่ำที่สุด เมื่อเปรียบเทียบกับโมเดล 2pl และโมเดล 3pl ยกเว้นแบบสอบฉบับที่ 6 ที่โมเดล 2pl มีค่า BIC ต่ำที่สุด และเมื่อตรวจสอบนัยสำคัญของผลการเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลแต่ละคู่ พบว่า เมื่อใช้โมเดล 2pl ซึ่งประกอบด้วยพารามิเตอร์ความยากและพารามิเตอร์การจำแนก สามารถอธิบายความแปรปรวนของความสามารถนิสิตนักศึกษาได้เพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติ ($p<.05$) มากกว่าโมเดล 1pl ที่มีเฉพาะพารามิเตอร์ความยากเพียงอย่างเดียว ในทางตรงกันข้าม เมื่อเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดล 3pl ซึ่งมีพารามิเตอร์การเดาเพิ่มขึ้นอีก 35 ค่าจากโมเดล 2pl พบว่า โดยภาพรวม พารามิเตอร์การเดาทั้ง 35 ค่าไม่ได้ทำให้สัดส่วนความแปรปรวนที่สามารถอธิบายได้เพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติ ($p>.05$) ดังนั้นจึงสรุปว่า **โมเดล 2pl เป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลคำตอบของแบบสอบทุกฉบับมากที่สุด**

ผลการเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลที่แสดงในตาราง 4.17เป็นผลวิเคราะห์จากการวัดแบบเอกมิติ กล่าวคือ ข้อสอบทุกข้อวัดตัวแปรแฝงความสามารถนิสิตนักศึกษาเพียงตัวแปรเดียว (ดูภาพ 3.2 ประกอบ) แต่อย่างไรก็ตาม ความสามารถหรือทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารสามารถพิจารณาเป็น 5 องค์ประกอบย่อยได้ด้วย (ดูภาพ 3.3 ประกอบ) ดังนั้นการเปรียบเทียบความสอดคล้อง

กลมกลืนในขั้นตอนต่อไปจึงเป็นการเปรียบเทียบระหว่างโมเดลเอกมิติ 2pl และโมเดลพหุมิติ 2pl สำหรับแบบสอบทั้ง 10 ฉบับ ผลการตรวจสอบความสอดคล้องกลมกลืนระหว่างโมเดลเอกมิติและโมเดลพหุมิติแสดงดังตาราง 4.18

ตาราง 4.18 ผลการเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดลเอกมิติและโมเดลพหุมิติ

ฉบับที่ 1	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl	ฉบับที่ 2	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl
n	181	181	n	176	176
p	70	80	p	70	80
G^2	7172.89	7242.35	G^2	7102.00	7146.22
AIC	7312.89	7402.35	AIC	7242.00	7307.02
BIC	7536.78	7658.23	BIC	7463.94	7560.66
1D vs. 5D	LRT = -69.46, ΔR^2 = -.0097 (-.97%)		1D vs. 5D	LRT = -45.02	ΔR^2 = -.0090 (-.90%)
ฉบับที่ 3	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl	ฉบับที่ 4	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl
n	170	170	n	167	167
p	70	80	p	70	80
G^2	7102.00	7147.02	G^2	6888.81	6918.96
AIC	7242.00	7307.02	AIC	7028.81	7078.96
BIC	7463.94	7560.66	BIC	7247.07	7328.40
1D vs. 5D	LRT = -45.02	ΔR^2 = -.0063 (-.63%)	1D vs. 5D	LRT = -30.14	ΔR^2 = -.0044 (.44%)
ฉบับที่ 5	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl	ฉบับที่ 6	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl
n	169	169	n	164	164
p	70	80	p	70	80
G^2	6705.88	6736.37	G^2	6721.02	6782.76
AIC	6845.88	6896.37	AIC	6861.30	6942.76
BIC	7064.97	7146.76	BIC	7078.29	7190.75
1D vs. 5D	LRT = -30.49	ΔR^2 = -.0045 (-.45%)	1D vs. 5D	LRT = -61.46	ΔR^2 = -.0092 (-.92%)
ฉบับที่ 7	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl	ฉบับที่ 8	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl
n	166	166	n	164	164
p	70	80	p	70	80
G^2	6953.75	6970.67	G^2	6723.24	6777.39
AIC	7093.75	7130.67	AIC	6863.24	6937.39
BIC	7311.59	7379.63	BIC	7080.24	7185.38
1D vs. 5D	LRT = -16.92	ΔR^2 = -.0024 (-.24%)	1D vs. 5D	LRT = -54.15	ΔR^2 = -.0081 (-.81%)
ฉบับที่ 9	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl	ฉบับที่ 10	โมเดลเอกมิติ 2pl	โมเดลพหุมิติ 2pl
n	160	160	n	155	155
p	70	80	p	70	80
G^2	6625.23	6660.09	G^2	6299.95	6346.65
AIC	6765.23	6820.09	AIC	6439.95	6506.65
BIC	6980.49	7066.10	BIC	6652.99	6750.12
1D vs. 5D	LRT = -34.86	ΔR^2 = -.0053 (-.53%)	1D vs. 5D	LRT = -46.70	ΔR^2 = -.0074 (-.74%)

ผลการเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดลเอกมิติ 2pl และโมเดลพหุมิติ 2pl ในตาราง 4.18 แสดงให้เห็นว่า เมื่อความสามารถ/ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารของนิสิตนักศึกษา (θ_{ict}) ถูกกำหนดให้เป็นความสามารถย่อยเกี่ยวกับการเข้าถึงสารสนเทศ (θ_{Ac}) ด้านการจัดการ

สารสนเทศ (θ_{Mn}) ด้านการบูรณาการสารสนเทศ (θ_{Ig}) ด้านการประเมินสารสนเทศ (θ_{Ev}) และด้านการสื่อสารสารสนเทศ (θ_{Cm}) โมเดลพหุมิติไม่ได้สอดคล้องกลมกลืนกับข้อมูลได้ดีกว่าโมเดลเอกมิติอย่างมีนัยสำคัญสำหรับแบบสอบทุกฉบับ โดยจะเห็นว่า ค่า G^2, AIC และ BIC ของโมเดลเอกมิติมีค่าต่ำกว่าโมเดลพหุมิติ และการตรวจสอบนัยสำคัญของผลการเปรียบเทียบความสอดคล้องกลมกลืนไม่มีนัยสำคัญทางสถิติ ($p>.05$) นั่นคือ ในภาพรวม ค่าความแปรปรวนร่วมของความสามารถย่อยแต่ละองค์ประกอบที่มีต่อกัน (เช่นความแปรปรวนร่วมระหว่าง θ_{Ac} และ θ_{Mn}) ไม่ได้ช่วยเพิ่มสัดส่วนความแปรปรวนที่สามารถอธิบายได้ ดังนั้นจึงสรุปว่า ข้อสอบที่ผู้วิจัยสร้างขึ้นสามารถใช้วัดความสามารถนิสิตนักศึกษาแบบภาพรวมได้ดีกว่าการวัดแบบแยกรายองค์ประกอบ และผู้วิจัยจะใช้โมเดลเอกมิติ 2pl **เพื่อประมาณค่าพารามิเตอร์ข้อสอบก่อนนำเข้าคลังข้อสอบแบบปรับเหมาะต่อไป**

2) การปรับเทียบแบบสอบและการประเมินผลคุณภาพข้อสอบ

ผลวิเคราะห์ในส่วนนี้เกี่ยวข้องกับการปรับเทียบ (equating) แบบสอบทั้ง 10 ฉบับเพื่อให้ค่าพารามิเตอร์ความยากและพารามิเตอร์การจำแนกของข้อสอบทุกข้ออยู่บนสเกลเดียวกัน วิธีการปรับเทียบที่ผู้วิจัยเลือกใช้คือ การปรับเทียบรวม (pooled equating: Lord, 1980) หรือการปรับเทียบพร้อมกัน (concurrent calibration method: Wingersky & Lord, 1984) โดยข้อมูลคำตอบของตัวอย่างนิสิตนักศึกษาจากแบบสอบแต่ละฉบับจะนำมารวมกัน จากนั้นวิเคราะห์ข้อมูลทั้งหมดพร้อมกัน โดยมีข้อสอบร่วมทั้ง 10 ข้อทำหน้าที่เชื่อมโยงข้อสอบปกติอีก 250 ข้อเข้าด้วยกัน และข้อมูลคำตอบขาดหายจะถูกสมมติให้เป็นข้อสอบที่ไม่ให้กำหนดให้นิสิตนักศึกษาทำ

ผลการปรับเทียบข้อสอบทั้ง 260 ข้อ พบว่า ข้อสอบแต่ละข้อมีค่าพารามิเตอร์ความยาก ค่าพารามิเตอร์การจำแนก และค่าความคลาดเคลื่อน แตกต่างกันมาก จึงนำไปสู่การคัดเลือกข้อสอบที่มีคุณภาพเพื่อเข้าสู่คลังข้อสอบแบบปรับเหมาะ ผู้วิจัยพิจารณาคุณภาพข้อสอบจากค่าความคลาดเคลื่อนของการประมาณค่าพารามิเตอร์เป็นหลัก โดยข้อสอบที่มีค่าความเคลื่อนสูงผิดปกติ อันเนื่องมาจากพารามิเตอร์ข้อสอบข้อนั้นๆ ไม่สามารถระบุเป็นค่าที่แน่นอนได้ (not well-specified) และเลือกกำจัดข้อสอบที่มีค่าพารามิเตอร์ความยากสูงผิดปกติและต่ำผิดปกติด้วย โดยผู้วิจัยยอมรับข้อสอบที่มีค่าพารามิเตอร์ความยากต่ำกว่า -3.00 และ สูงกว่า +3.00 ลอจิท เป็นข้อสอบที่มีคุณภาพตรงตามวัตถุประสงค์ของการวิจัย ที่ต้องการใช้ข้อสอบง่ายเพื่อตรวจสอบความรู้พื้นฐานของนิสิตนักศึกษา และใช้ข้อสอบยากเพื่อท้าทายความสามารถของนิสิตนักศึกษา เช่นเดียวกับงานวิจัยของ Ali and Chang (2014) และ Mao and Xin (2013) ที่ยอมรับข้อสอบที่มีค่าพารามิเตอร์ความยากต่ำกว่า -3.00 และสูงกว่า +3.00 ลอจิท ในกระบวนการพัฒนาคลังข้อสอบแบบปรับเหมาะขั้นต้น

สำหรับการพิจารณาคุณภาพข้อสอบจากค่าพารามิเตอร์การจำแนก ผู้วิจัยเลือกใช้เกณฑ์ของ Baker and Kim (2017) โดยผู้วิจัยกำจัดข้อสอบที่มีค่าการจำแนกติดลบ และเลือกข้อสอบที่มีค่าการจำแนกตั้งแต่ 0 ขึ้นไป โดยผู้วิจัยยอมรับข้อสอบที่มีค่าการจำแนกใกล้เคียงกับ 0 ด้วย เช่นเดียวกับงานวิจัยของ Ali and Chang (2014) ที่ยอมรับข้อสอบที่มีค่าพารามิเตอร์การจำแนกใกล้เคียง 0 ในขั้นตอนของการพัฒนาคลังข้อสอบขั้นต้น นั่นหมายความว่า ข้อสอบที่ผ่านเกณฑ์จะหมายถึงข้อสอบที่เมื่อนิสิตนักศึกษาตอบข้อสอบนั้นถูกจะเป็นกลุ่มนิสิตนักศึกษาที่มีความสามารถในภาพรวมสูงกว่านิสิตนักศึกษาที่ตอบข้อสอบข้อนั้นผิด

ผลการประเมินคุณภาพข้อสอบตามเกณฑ์ข้างต้นพบว่า มีข้อสอบที่ถูกตัดทิ้งจำนวน 48 ข้อ คงเหลือข้อสอบที่จะนำเข้าคลังข้อสอบแบบปรับเหมาะจำนวน 212 ข้อ แบ่งเป็นข้อสอบด้านการเข้าถึงสารสนเทศจำนวน 43 ข้อ ด้านการจัดการสารสนเทศ จำนวน 41 ข้อ ด้านการบูรณาการสารสนเทศ จำนวน 43 ข้อ ด้านการประเมินสารสนเทศ จำนวน 40 ข้อ และด้านการสื่อสารสารสนเทศ จำนวน 45 ข้อ ซึ่งจำนวนข้อสอบ 212

ข้อนี้เพียงพอสำหรับการพัฒนาเป็นคลังข้อสอบแบบปรับเหมาะต่อไป (Drasgow, 1989) นอกจากนี้ค่าเฉลี่ยพารามิเตอร์ความยากข้อสอบเท่ากับ .18 (SD.=2.06) และค่าเฉลี่ยพารามิเตอร์การจำแนกเท่ากับ .70 (SD.=.39) นั่นหมายความว่า เมื่อพิจารณาจากเกณฑ์การตีความค่าพารามิเตอร์ข้อสอบของ Baker and Kim (2017) และ Hambleton et al. (1991) พบว่า ในภาพรวม ข้อสอบที่นำเข้าคลังข้อสอบแบบปรับเหมาะเป็นข้อสอบยากปานกลางและจำแนกได้ปานกลาง

ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ตามทฤษฎี IRT แบบ 2PL จำแนกตามองค์ประกอบ องค์ประกอบละ 52 ข้อ รวมเป็นข้อสอบทั้งสิ้น 260 ข้อโดยแต่ละองค์ประกอบมีค่าพารามิเตอร์ความยาก (b) และอำนาจจำแนก (a) ในภาพรวมดังนี้ รายละเอียดดังตาราง 4.19-4.23

ตาราง 4.19 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 1 ตามทฤษฎี IRT

องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ							
ข้อที่	a	b	ผลประเมิน	ข้อที่	a	b	ผลประเมิน
1	0.382	-4.336		27	0.383	-1.210	
2	0.235	4.921		28	0.322	5.502	
3	0.733	-0.092		29	0.470	-3.852	
4	0.715	0.078		30	0.705	-0.045	
5	1.424	-0.513		31	0.501	-1.687	
6	0.191	-2.859		32	0.009	67.838	คัดออก
7	0.564	-3.178		33	0.491	1.308	
8	0.203	3.116		34	0.628	-0.307	
9	0.545	-0.803		35	1.416	-0.945	
10	0.340	3.555		36	-0.163	-13.680	คัดออก
11	0.108	18.109	คัดออก	37	0.899	-0.373	
12	0.236	-9.465		38	0.090	7.580	คัดออก
13	0.202	1.580		39	0.672	0.053	
14	0.740	-2.376		40	1.033	-0.366	
15	0.519	1.158		41	0.452	0.603	
16	0.452	-2.283		42	-0.051	-5.835	คัดออก
17	0.298	-0.181		43	0.690	-1.012	
18	0.429	-0.355		44	-0.217	8.291	คัดออก
19	0.646	0.787		45	0.163	3.848	
20	0.748	-0.135		46	0.874	-0.639	
21	0.146	-2.750		47	0.356	3.861	
22	0.984	-0.532		48	0.733	-2.126	
23	1.029	-1.621		49	0.062	11.383	
24	-0.037	-26.088	คัดออก	50	0.575	1.257	คัดออก
25	0.400	0.068		51*	0.675	-0.791	
26	0.215	-1.443		52*	-0.252	-8.104	คัดออก
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 43 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 82.70)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 9 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 17.30)							

* หมายถึง ข้อสอบแกนกลาง

ตาราง 4.20 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 2 ตามทฤษฎี IRT

องค์ประกอบที่ 2 การจัดการสารสนเทศ							
ข้อที่	a	b	ผลประเมิน	ข้อที่	a	b	ผลประเมิน
53	0.829	-1.196		79	-0.025	-39.073	คัดออก
54	0.527	-0.445		80	-0.119	-6.627	คัดออก
55	0.579	-0.159		81	0.839	0.039	
56	0.763	-1.604		82	0.137	-4.760	
57	0.030	47.454	คัดออก	83	1.130	1.358	
58	0.462	6.940		84	1.517	-1.023	
59	0.427	-0.425		85	-0.014	-40.225	คัดออก
60	0.484	4.136		86	0.926	1.096	
61	0.094	1.935		87	-0.035	-9.581	คัดออก
62	0.022	16.673	คัดออก	88	0.070	17.158	คัดออก
63	0.462	6.940		89	-0.017	-106.26	คัดออก
64	0.427	-0.425		90	1.147	-0.566	
65	0.484	4.136		91	1.012	-0.133	
66	0.119	9.487	คัดออก	92	0.0141	80.5313	คัดออก
67	0.094	1.935		93	0.678	1.487	
68	1.046	-0.098		94	0.547	-2.314	
69	0.396	-1.838		95	-0.240	-7.929	คัดออก
70	0.939	0.013		96	0.105	-0.460	
71	0.129	5.539		97	0.763	1.181	
72	0.310	3.210		98	0.725	0.077	
73	0.853	-1.264		99	0.410	4.260	
74	0.506	3.796		100	0.221	1.586	
75	0.697	0.222		101	0.718	1.771	
76	0.720	0.437		102	0.910	0.162	
77	0.719	1.100		103*	0.594	-1.158	
78	0.931	2.039		104*	0.215	4.372	
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 41 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 78.85)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 11 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 21.15)							

* หมายถึง ข้อสอบแกนกลาง

ตาราง 4.21 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 3 ตามทฤษฎี IRT

องค์ประกอบที่ 3 การบูรณาการสารสนเทศ							
ข้อที่	a	b	ผลประเมิน	ข้อที่	a	b	ผลประเมิน
105	0.416	0.896		131	0.384	1.063	
106	1.356	-1.216		132	0.032	7.681	คัดออก
107	0.511	1.534		133	1.151	-0.111	
108	0.033	34.627	คัดออก	134	1.239	-0.024	
109	1.147	-0.082		135	0.663	-1.556	
110	1.042	-2.574		136	0.745	0.769	
111	0.211	4.392		137	0.702	0.335	
112	0.034	12.810	คัดออก	138	0.298	0.028	
113	0.066	10.642	คัดออก	139	0.589	1.118	
114	0.975	-1.487		140	0.315	-2.214	
115	0.402	-0.323		141	0.653	-0.442	
116	0.371	4.078		142	-0.196	-5.622	คัดออก
117	0.537	-1.699		143	0.541	1.928	
118	0.929	0.133		144	1.199	-0.840	
119	0.648	-0.038		145	0.262	-0.760	
120	1.028	-1.040		146	0.371	-0.965	
121	0.915	0.415		147	0.433	1.849	
122	0.550	-2.610		148	1.277	-0.693	
123	0.395	1.473		149	0.669	-2.175	
124	0.799	-0.196		150	-0.047	-9.142	คัดออก
125	0.046	38.776	คัดออก	151	0.107	5.575	คัดออก
126	0.653	-0.122		152	0.608	-0.654	
127	0.351	-1.550		153	0.587	-1.498	
128	0.931	1.066		154	0.477	-0.454	
129	0.787	0.063		155*	-0.302	-5.713	คัดออก
130	0.223	1.860		156*	0.623	-0.275	
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 43 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 82.70)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 9 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 17.30)							

* หมายถึง ข้อสอบแกนกลาง

ตาราง 4.22 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 4 ตามทฤษฎี IRT

องค์ประกอบที่ 4 การประเมินสารสนเทศ							
ข้อที่	a	b	ผลประเมิน	ข้อที่	a	b	ผลประเมิน
158	-1.123	0.935		183	0.695	0.344	
158	0.169	8.714	คัดออก	184	0.669	3.005	
159	-1.410	2.383		185	0.743	-0.114	
160	-0.688	2.082		186	0.751	0.514	
161	-0.105	-9.414	คัดออก	187	0.355	2.303	
162	1.167	-0.131		188	0.607	1.233	
163	0.165	3.533		189	1.679	-0.832	
164	0.972	-0.202		190	0.059	17.165	คัดออก
165	0.557	1.739		191	-0.065	-16.463	คัดออก
166	1.200	-0.710		192	0.443	1.958	
167	-0.330	-4.963	คัดออก	193	0.124	7.355	คัดออก
168	0.776	0.060		194	0.422	1.244	
169	0.163	2.570		195	0.651	-0.235	
170	1.161	1.017		196	0.751	0.382	
171	0.161	7.270		197	0.791	-0.049	
172	0.427	-2.046		198	1.073	-1.461	
173	0.221	5.934		199	0.352	-0.053	
174	0.876	-0.084		200	0.618	-1.730	
175	0.586	-0.395		201	-0.281	-1.933	ตัดออก
176	0.112	10.303	คัดออก	202	1.018	-0.070	
177	-0.123	-17.876	คัดออก	203	0.999	-0.785	
178	-0.210	-7.959	คัดออก	204	1.281	-1.213	
179	0.437	1.403		205	1.202	-1.152	
180	13.109	0.068	คัดออก	206	1.141	-1.190	
181	3.379	-0.297	คัดออก	207*	0.697	0.838	
182	1.036	-0.734		208*	0.690	1.753	
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 40 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 76.92)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 12 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 23.08)							

* หมายถึง ข้อสอบแกนกลาง

ตาราง 4.23 ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อ (Item) ในองค์ประกอบที่ 5 ตามทฤษฎี IRT

องค์ประกอบที่ 5 การสื่อสารสารสนเทศ							
ข้อที่	a	b	ผลประเมิน	ข้อที่	a	b	ผลประเมิน
209	0.938	-0.956		235	1.228	-0.450	
210	0.495	0.240		236	1.312	-0.455	
211	0.854	-1.132		237	1.847	-0.573	
212	0.596	1.921		238	0.674	0.310	
213	0.783	0.902		239	0.462	1.563	
214	-0.684	1.198		240	0.868	-0.593	
215	0.839	0.439		241	0.228	1.320	
216	-0.154	-3.040	คัดออก	242	0.893	-1.546	
217	-0.368	0.508		243	0.095	7.374	คัดออก
218	-1.193	1.263		244	1.009	-0.463	
219	0.790	0.194		245	0.788	-0.378	
220	1.197	-0.477		246	0.397	-0.371	
221	2.048	-1.243		247	1.734	-0.588	
222	0.565	-0.708		248	0.294	3.819	
223	-0.446	-2.975	คัดออก	249	0.305	-1.252	
224	0.832	0.682		250	0.599	2.171	
225	0.703	2.651		251	0.697	1.089	
226	-0.021	-27.288	คัดออก	252	0.183	7.847	คัดออก
227	0.215	5.287		253	0.243	1.175	
228	-0.142	-15.344	คัดออก	254	0.568	0.246	
229	0.379	0.855		255	0.516	-1.199	
230	0.713	-1.592		256	0.815	-1.747	
231	0.600	-1.430		257	1.015	-1.021	
232	1.616	-1.293		258	1.635	-0.595	
233	-3.5416	-0.350	คัดออก	259*	0.316	3.910	
234	0.593	-0.334		260*	0.832	-0.056	
รวมข้อสอบที่ผ่านเกณฑ์จำนวน 45 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 86.54)							
รวมข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 7 ข้อ จากทั้งหมด 52 ข้อ (ร้อยละ 13.46)							

* หมายถึง ข้อสอบแกนกลาง

2.4 ผลการคัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบ

ผลการคัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบจำแนกออกเป็น 2 ประเด็น คือ 1) ผลการคัดเลือกจำแนกตามองค์ประกอบและชุดแบบทดสอบ และ 2) ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกจำแนกตามองค์ประกอบ โดยมีรายละเอียดดังนี้

2.4.1 ผลการคัดเลือกจำแนกตามองค์ประกอบและชุดแบบทดสอบ

ผลการคัดเลือกทั้งหมด 260 ข้อ มีข้อสอบที่ถูกคัดออกรวม 48 ข้อ รวมเป็นข้อสอบที่คัดเลือกเข้าคลังข้อสอบ จำนวน 212 ข้อ คิดเป็นร้อยละ 81.53 เมื่อจำแนกตามองค์ประกอบและชุดแบบทดสอบเมื่อจำแนกตามองค์ประกอบของแบบทดสอบพบว่า องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ มีข้อสอบที่ถูกคัดออก 9 ข้อ จากทั้งหมด 52 ข้อ องค์ประกอบที่ 2 การจัดการสารสนเทศ มีข้อสอบที่ถูกคัดออก 11 ข้อ จากทั้งหมด 52 ข้อ องค์ประกอบที่ 3 การบูรณาการสารสนเทศ มีข้อสอบที่ถูกคัดออก 9 ข้อ จากทั้งหมด 52 ข้อ องค์ประกอบที่ 4 การประเมินสารสนเทศ มีข้อสอบที่ถูกคัดออก 12 ข้อ จากทั้งหมด 52 ข้อ และองค์ประกอบที่ 5 การประเมินสารสนเทศ มีข้อสอบที่ถูกคัดออก 7 ข้อ จากทั้งหมด 52 ข้อ รายละเอียดดังตาราง 4.24

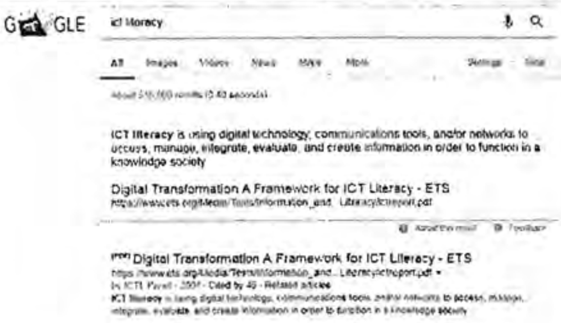
ตาราง 4.24 ผลการคัดข้อสอบออกจำแนกตามองค์ประกอบและชุดแบบทดสอบ

องค์ประกอบที่	แบบสอบฉบับที่ (ข้อ)											รวมคัดออก (ข้อ)
	1	2	3	4	5	6	7	8	9	10	แกนกลาง	
1 การเข้าถึงสารสนเทศ	0	0	1	0	1	0	1	2	2	1	1	9
2 การจัดการสารสนเทศ	1	1	1	0	0	2	2	3	1	0	0	11
3 การบูรณาการสารสนเทศ	1	2	0	0	1	1	0	1	0	2	1	9
4 การประเมินสารสนเทศ	2	0	1	1	4	0	2	1	1	0	0	12
5 การสื่อสารสารสนเทศ	0	1	1	2	1	0	1	0	1	0	0	7
รวม	4	4	4	3	7	3	6	7	5	3	2	48

2.4.2 ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกจำแนกตามองค์ประกอบ

ข้อสอบที่ผ่านการคัดเลือกและคัดออกจำแนกตามองค์ประกอบโดยใช้ค่าพารามิเตอร์ของข้อสอบรายข้อตามทฤษฎี IRT เป็นเกณฑ์ในการคัดข้อสอบในทุกองค์ประกอบโดยมีตัวอย่างข้อสอบในแต่ละองค์ประกอบที่ผ่านการคัดเลือกและคัดออกดังปรากฏในตาราง 4.25-4.29

ตาราง 4.25 ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 1 การเข้าถึงสารสนเทศ

ตัวอย่างข้อสอบที่ผ่านการคัดเลือก ขององค์ประกอบที่ 1	**ตัวอย่างข้อสอบที่ถูกคัดออก ขององค์ประกอบที่ 1**
 5.จากภาพ เมื่อต้องการค้นหาหนังสือที่เกี่ยวข้องกับ ICT Literacy ควรดำเนินการอย่างไรจึงมีความเหมาะสมและรวดเร็วที่สุด 1. คลิก Images เลือก Book 2. คลิก Videos เลือก Book 3. คลิก News เลือก Book 4. คลิก Maps เลือก Book 5. คลิก More เลือก Book	7.กำหนดให้ Array หนึ่งเก็บข้อมูลเมทริกซ์จำนวน 3 เมทริกซ์ โดยแต่ละเมทริกซ์มีขนาด 2 x 2 ดังรูป และต้องเรียกคืนค่า 0.6 และ 0.8 ต้องกำหนดตำแหน่งไฟล์อย่างไร $\begin{bmatrix} 0.2 & 0.3 \\ 0.4 & 0.5 \end{bmatrix}$ เมทริกซ์ที่ 1 $\begin{bmatrix} 0.1 & 0.0 \\ 0.9 & 1.5 \end{bmatrix}$ เมทริกซ์ที่ 2 $\begin{bmatrix} 0.2 & 0.5 \\ 0.6 & 0.8 \end{bmatrix}$ เมทริกซ์ที่ 3 1. [2, ,3] 2. [,2,3] 3. [3,2,] 4. [3, ,2] 5. [3,2,2]
พารามิเตอร์ความยาก (b) = -0.512	พารามิเตอร์ความยาก (b) =-8.104
พารามิเตอร์อำนาจจำแนก (a) = 0.156	พารามิเตอร์อำนาจจำแนก (a) =0.252

ตาราง 4.26 ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 2 การจัดการสารสนเทศ

ตัวอย่างข้อสอบที่ผ่านการคัดเลือก ขององค์ประกอบที่ 2	ตัวอย่างข้อสอบที่ถูกคัดออก ขององค์ประกอบที่ 2
12.หากท่านต้องการ Copy ไฟล์งานของท่านที่ทำไว้ในเครื่องคอมพิวเตอร์เดสก์ท็อปของเพื่อน แต่ท่านไม่ได้นำ Handy Drive ไปด้วย ท่านควรจัดการสถานการณ์นี้อย่างไรด้วยตนเองโดยใช้ประโยชน์จากระบบเครือข่ายที่มีอยู่ได้อย่างรวดเร็ว ครบทุกไฟล์ที่ต้องการอย่างมีประสิทธิภาพ 1.ให้เพื่อนส่งไฟล์งานให้ทางอีเมล์ 2.ให้เพื่อนส่งไฟล์งานให้ทาง Line 3.Login เข้าอีเมล์ตนเอง และส่งไฟล์แนบ (Attached File) 4.ขอยืม Handy Drive เพื่อน เพื่อ Save ไฟล์งานจากคอมพิวเตอร์เพื่อน 5.* ถูกทุกข้อ	10.ขนิษฐาทำงานมา 7 ปี องค์กรไม่มีระบบการจัดการเอกสารและนโยบายการบันทึกเอกสาร ดังนั้น ขนิษฐาควรจัดการเอกสารลักษณะใดที่มีความเหมาะสม เพื่อให้คนอื่นสามารถทำงานที่เกี่ยวข้องต่อได้ หากมีกรณีด่วนที่ไม่สามารถติดต่อตนได้ 1.จัดเก็บเอกสารไว้ในแฟลชไดร์ฟของตนพกไว้ในกระเป๋าเสมอ โดยแยกแฟลชไดร์ฟละ 1 งาน 2.จัดเก็บและบันทึกเอกสารเหมือนปกติที่เคยปฏิบัติ เพราะเป็นรูปแบบที่ใคร ๆ ก็สามารถเข้าใจได้ 3.จัดเก็บเอกสารไว้บนเดสก์ทอปเพื่อความสะดวกในการใช้งาน และบันทึกชื่อเอกสารตามคำแนะนำของโปรแกรม 4.จัดเก็บเอกสารไว้บนคลาวด์ (iCloud) โดยการแยกโฟลเดอร์ตามชื่องานและบันทึกชื่อเอกสารตามวัตถุประสงค์ของเอกสาร 5.จัดเก็บรวมกันไว้ในพาร์ติชันอื่นที่ไม่ใช่พาร์ติชันระบบปฏิบัติการ โดยการแยกโฟลเดอร์ตามชื่องานและบันทึกชื่อเอกสารตามหัวข้อหลัก-หัวข้อย่อยของเอกสาร
พารามิเตอร์ความยาก (b) = -1.195	พารามิเตอร์ความยาก (b) = 47.454
พารามิเตอร์อำนาจจำแนก (a) = 0.828	พารามิเตอร์อำนาจจำแนก (a) =0.0301

ตาราง 4.27 ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 3 การบูรณาการสารสนเทศ

ตัวอย่างข้อสอบที่ผ่านการคัดเลือก ขององค์ประกอบที่ 3	**ตัวอย่างข้อสอบที่ถูกคัดออก ขององค์ประกอบที่ 3**
18.จากหัวข้อข่าวเรื่อง "คนไทยอ่านปีละ 8 บรรทัด" สถิติใหม่ชี้อ่านกันวันละ 66 นาที!!"และแผนภาพแสดงผลการสำรวจที่กำหนดให้ ข้อใดต่อไปนี้กล่าวถึงผลการสำรวจได้ถูกต้อง (ที่มา : https://www.khaosod.co.th/monitor-news/news_16138) 1.วัยรุ่นใช้ระยะเวลาในการอ่านมากกว่าวัยเด็ก 2.หญิงใช้ระยะเวลาในอ่านประมาณ 10 ชม./สัปดาห์ 3.สาเหตุส่วนใหญ่ที่คนไม่ชอบอ่านหนังสือเพราะไม่ชอบหรือไม่สนใจ 4.คนส่วนใหญ่ที่ชอบอ่านหนังสือเพราะครู อาจารย์ส่งเสริมการอ่าน 5.คนนอกเขตเทศบาลมีแนวโน้มชอบอ่านน้อยกว่าคนในเขตเทศบาล	19.หากท่านกำลังท่องอินเทอร์เน็ตแล้วพบว่าเพื่อนของท่านได้แบ่งปัน (Share) ข้อมูลเกี่ยวกับประเภทของพลังงานตามรายละเอียดด้านล่าง ประเภทของพลังงาน จากข้อมูลข้างต้น ข้อใดกล่าวเกี่ยวกับประเภทของพลังงานได้ถูกต้อง 1. พลังงานต้นกำเนิดทุกชนิดจัดเป็นพลังงานหมุนเวียน 2. พลังงานแปรรูปทุกชนิดจัดเป็นพลังงานที่ใช้หมดเปลือง 3. พลังงานที่ใช้หมดเปลืองทุกชนิดจัดเป็นพลังงานทางพาณิชย์ 4. พลังงานนอกพาณิชย์ทุกชนิดจัดเป็นพลังงานหมุนเวียน 5. พลังงานทางพาณิชย์ทุกชนิดจัดเป็นพลังงานต้นกำเนิด
พารามิเตอร์ความยาก (b) = 1.534	พารามิเตอร์ความยาก (b) = 34.627
พารามิเตอร์อำนาจจำแนก (a) = 0.616	พารามิเตอร์อำนาจจำแนก (a) =0.033

ตาราง 4.28 ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 4 การประเมินสารสนเทศ

ตัวอย่างข้อสอบที่ผ่านการคัดเลือกขององค์ประกอบที่ 4	ตัวอย่างข้อสอบที่ถูกคัดออกขององค์ประกอบที่ 4
25.ท่านคิดว่าการกระทำในข้อใดเหมาะสมที่สุดเกี่ยวกับการได้รับข้อมูลและข่าวสารต่าง ๆ บนสังคมออนไลน์ที่นำเสนอหลายด้านหลายแง่มุม 1.อ่านและเพิ่มเติมข้อมูลที่ตนทราบมาก่อนส่งต่อผู้อื่น 2.อ่านและเลือกส่งต่อเฉพาะข้อมูลที่ตนคิดว่าเป็นประโยชน์ต่อผู้อื่น 3.ไม่อ่านเนื่องจากส่วนใหญ่เป็นข้อมูลและข่าวสารที่ส่งต่อกันมา ทำให้สามารถได้รับข้อมูลอันเป็นเท็จได้ 4.อ่านพร้อมตรวจสอบข้อเท็จจริงจากแหล่งข้อมูลอื่นที่เชื่อถือได้ และเลือกส่งต่อเฉพาะข้อมูลที่เป็นประโยชน์ต่อผู้อื่น 5.อ่านรับทราบข้อมูลและข่าวสาร โดยไม่เผยแพร่ต่อทางสื่อสังคมออนไลน์ เพราะมีการส่งต่ออย่างแพร่หลายอยู่แล้วแต่จะเล่าต่อให้เพื่อนหรือครอบครัวรับทราบแทน	23.ข้อใดเป็นการรักษาความปลอดภัยจากเทคโนโลยีที่มีความเสี่ยงต่อการถูกโจรกรรมข้อมูล 1. Encryption 2. Anti Virus Software 3. Digital Signature 4. Contents Filter 5. Biometric Authentication
พารามิเตอร์ความยาก (b) = -0.688	พารามิเตอร์ความยาก (b) = 8.741
พารามิเตอร์อำนาจจำแนก (a) = 2.0820	พารามิเตอร์อำนาจจำแนก (a) =0.169

ตาราง 4.29 ตัวอย่างของข้อสอบที่ผ่านการคัดเลือกและคัดออกขององค์ประกอบที่ 5 การสื่อสารสารสนเทศ

ตัวอย่างข้อสอบที่ผ่านการคัดเลือก ขององค์ประกอบที่ 5	**ตัวอย่างข้อสอบที่ถูกคัดออก ขององค์ประกอบที่ 5**
30.หากท่านได้รับมอบหมายจากอาจารย์ให้จัดทำข้อมูลเพื่อเผยแพร่ให้นิสิตนักศึกษา นักศึกษา และบุคลากรภายในมหาวิทยาลัยรับทราบเกี่ยวกับสิ่งที่ควรกระทำเมื่อทำบัตรประชาชนหาย เพื่อป้องกันการถูกสวมรอยท่านจะเลือกใช้สื่อใดเพื่อนำเสนอข้อมูลดังกล่าวให้น่าสนใจ เข้าใจได้ง่าย ใช้เวลาน้อย และประหยัดค่าใช้จ่าย 1. SMS 2. วีดิทัศน์ 3. วารสาร 4. Infographic 5. Portable Document Format	31.สมพงษ์ได้เข้าร่วมค่ายอาสาเพื่อสร้างโรงเรียนให้แก่ชาวไทยภูเขาที่อาศัยอยู่บริเวณภาคเหนือของประเทศเป็นเวลา 2 สัปดาห์ แต่เขาลืมนำเอกสารเกี่ยวกับการจัดค่ายอาสาจากมหาวิทยาลัยซึ่งอยู่ที่กรุงเทพมหานครไปให้หัวหน้าชุมชนลงนาม เขาจึงขอให้ท่านจัดส่งเอกสารดังกล่าวไปให้ ท่านจะเลือกส่งเอกสารไปให้เขาผ่านทางช่องทางใดจึงจะเหมาะสม 1. โทรสาร (fax) 2. โทรศัพท์ 3. อีเมล (E-mail) 4. บริษัทจัดส่งเอกสาร 5. บริการไปรษณีย์ด่วนพิเศษ (EMS)
พารามิเตอร์ความยาก (b) = 0.8391	พารามิเตอร์ความยาก (b) = -3.040
พารามิเตอร์อำนาจจำแนก (a) = 0.210	พารามิเตอร์อำนาจจำแนก (a) = -0.154

ตอนที่ 3 ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี แบ่งเป็น 2 ประเด็น ได้แก่ 1) ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 และ 2) คู่มือการใช้ระบบการรายงานผลการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ดังมีรายละเอียดดังนี้

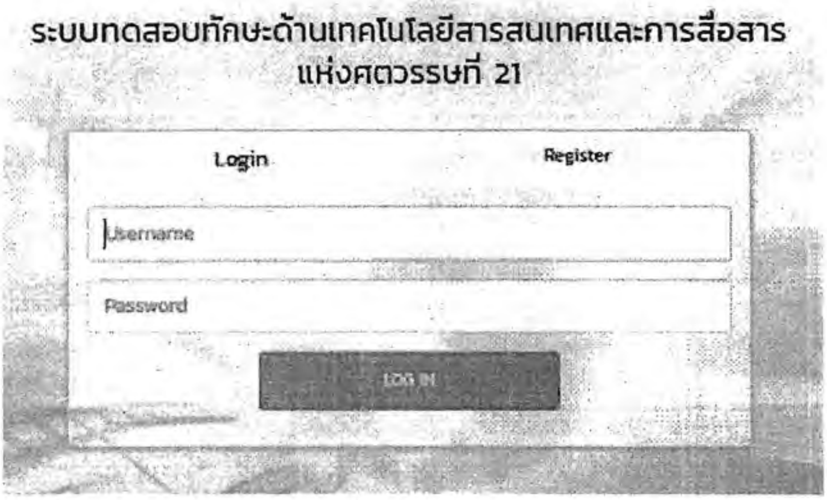
3.1 ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

สำหรับผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ได้พัฒนาโปรแกรมระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบ่งออกเป็น 4 ส่วน ได้แก่ 1) การลงทะเบียน 2) การสร้างชุดข้อสอบ 3) การทดสอบ และ 4) การรายงานผลการทดสอบ โดยแต่ละส่วนมีรายละเอียดดังนี้

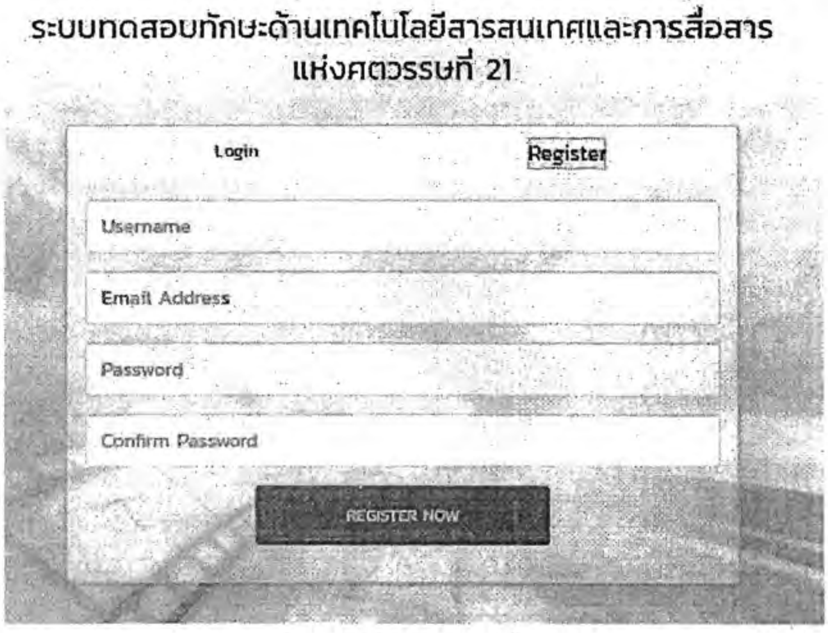
3.1.1 การลงทะเบียน

ผู้เข้ารับการทดสอบต้องดำเนินการลงทะเบียนก่อนเข้ารับการทดสอบ โดยเปิดเว็บไซต์ http://www.oneinthai.com/ictc จะปรากฏข้อมูลดังภาพ 4.6 หลังจากนั้นเลือกเมนู register เพื่อ

ลงทะเบียน โดยผู้เข้ารับการทดสอบต้องกรอกรหัสผู้ใช้ (username) อีเมล (e-mail) และรหัสผ่าน (password) โดยรายละเอียดการลงทะเบียนแสดงดังภาพ 4.7 เมื่อดำเนินการกรอกข้อมูลเรียบร้อย กดปุ่ม register now



ภาพ 4.6 หน้าหลักระบบการทดสอบ

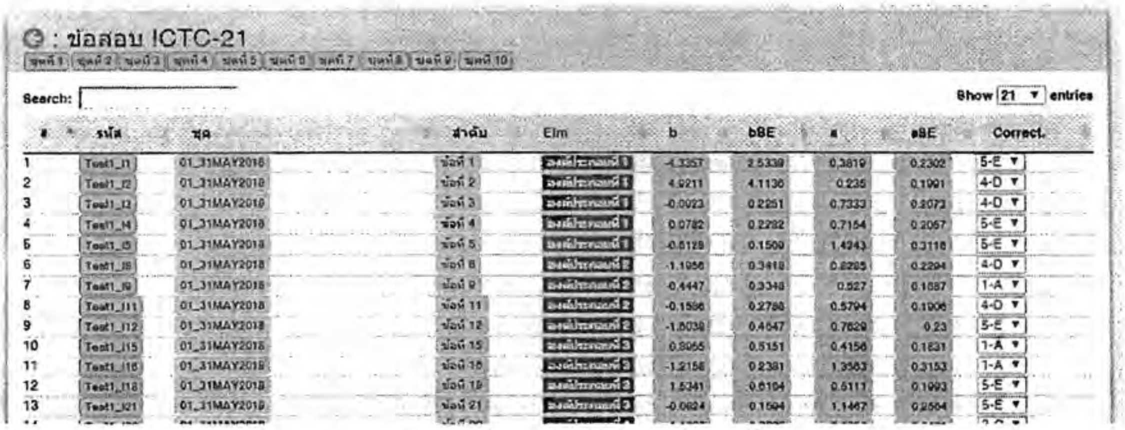


ภาพ 4.7 การลงทะเบียน

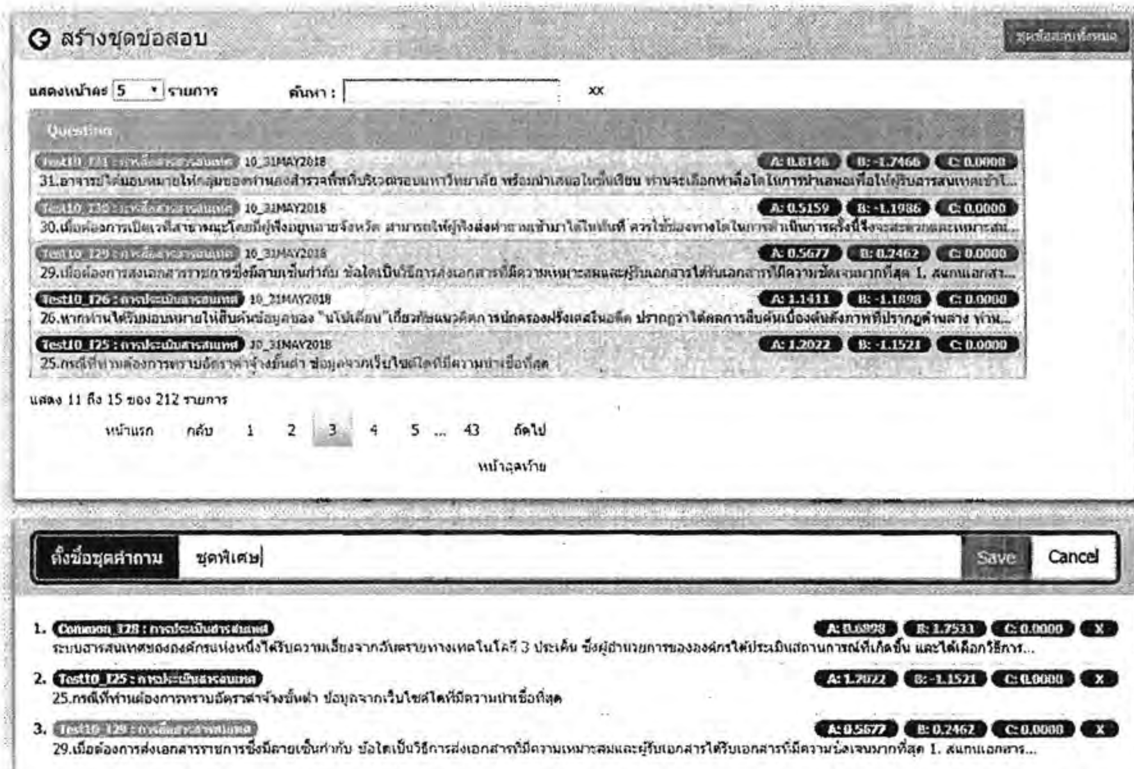
3.1.2 การสร้างชุดข้อสอบ

สำหรับการทดสอบในขั้นแรก (stage) ดำเนินการจัดชุดข้อสอบไว้ล่วงหน้าจำนวน 5 ชุด โดยเป็นข้อสอบที่มีความยาก (b) ปานกลาง หรือมีค่าอยู่ระหว่าง -0.5 ถึง 0.5 ซึ่งข้อสอบแต่ละชุดประกอบด้วยเนื้อหาครบทั้ง 5 องค์ประกอบเพื่อให้ผู้เข้ารับการทดสอบได้รับชุดข้อสอบชุดแรกที่มีเนื้อหาครบถ้วน โดยระบบการสร้างชุดข้อสอบพัฒนาขึ้นเพื่ออำนวยความสะดวกในการจัดชุดข้อสอบ ซึ่งระบบการสร้างชุดข้อสอบจะ

ประกอบด้วยฐานข้อมูลซึ่งมีข้อสอบในคลังจำนวน 212 ข้อ และชุดข้อสอบที่ดำเนินการจัดชุดไว้เรียบร้อยแล้ว จำนวน 5 ชุด เพื่อใช้สำหรับสุ่มให้ผู้เข้ารับการทดสอบดำเนินการทดสอบใน 5 ข้อแรก รายละเอียดแสดงดัง ภาพ 4.8 ถึง 4.9



ภาพ 4.8 ฐานข้อมูลสำหรับการสร้างชุดข้อสอบ

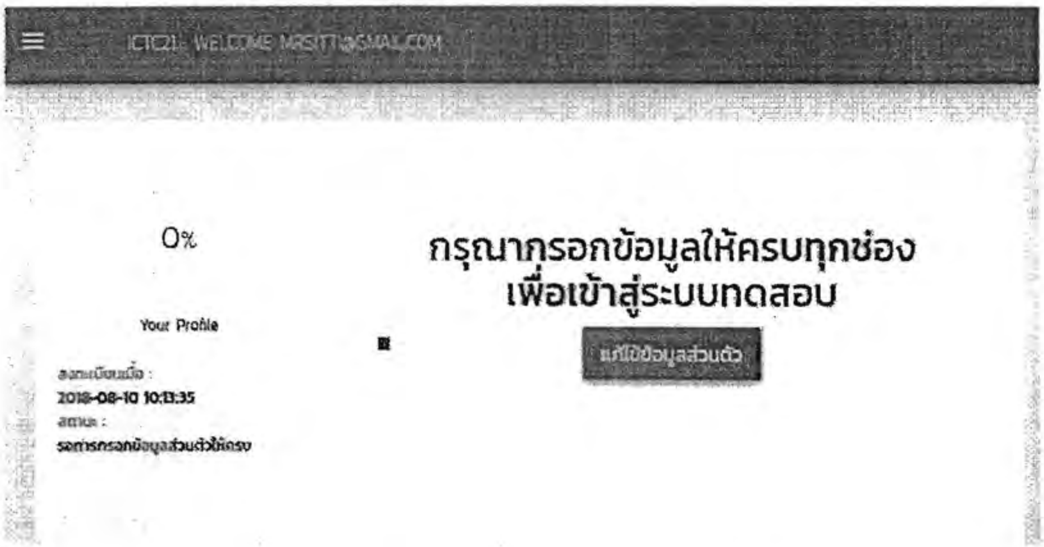


ภาพ 4.9 ชุดข้อสอบสำหรับการทดสอบขั้นแรก

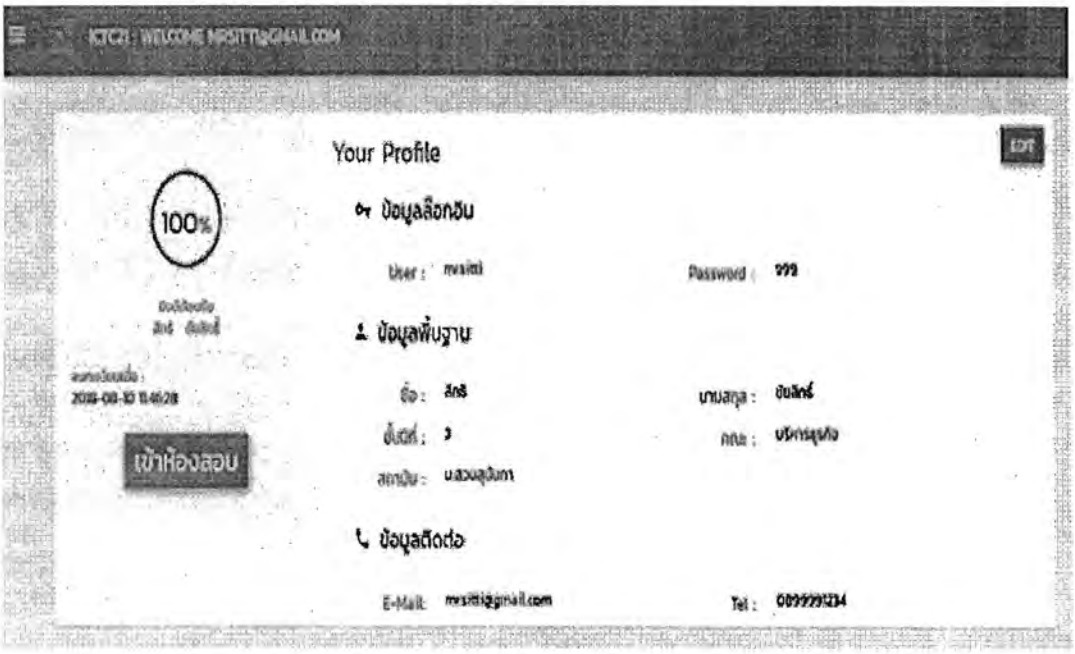
3.1.3 การทดสอบ

สำหรับกระบวนการในการทดสอบ เมื่อผู้เข้ารับการทดสอบเข้าสู่ระบบการทดสอบแล้ว ระบบจะแสดงหน้ารายละเอียดข้อมูลของผู้เข้ารับการทดสอบ ดังภาพ 4.10 โดยผู้เข้ารับการทดสอบต้อง

ดำเนินการกรอกข้อมูลให้ครบถ้วนเพื่อนำไปใช้ในการออกผลรายงานการทดสอบ หลังการทดสอบเสร็จสิ้น โดยผู้เข้ารับการทดสอบดำเนินการคลิกปุ่ม "แก้ไขข้อมูลส่วนตัว" เพื่อแก้ไขข้อมูล หลังจากแก้ไขข้อมูลเรียบร้อยแล้วกดปุ่มอัพเดต "update" ข้อมูลส่วนตัว ข้อมูลจะแสดงรายละเอียดดังภาพ 4.11 และกดปุ่ม "เข้าห้องสอบ" เพื่อเข้าสู่ระบบการทดสอบ

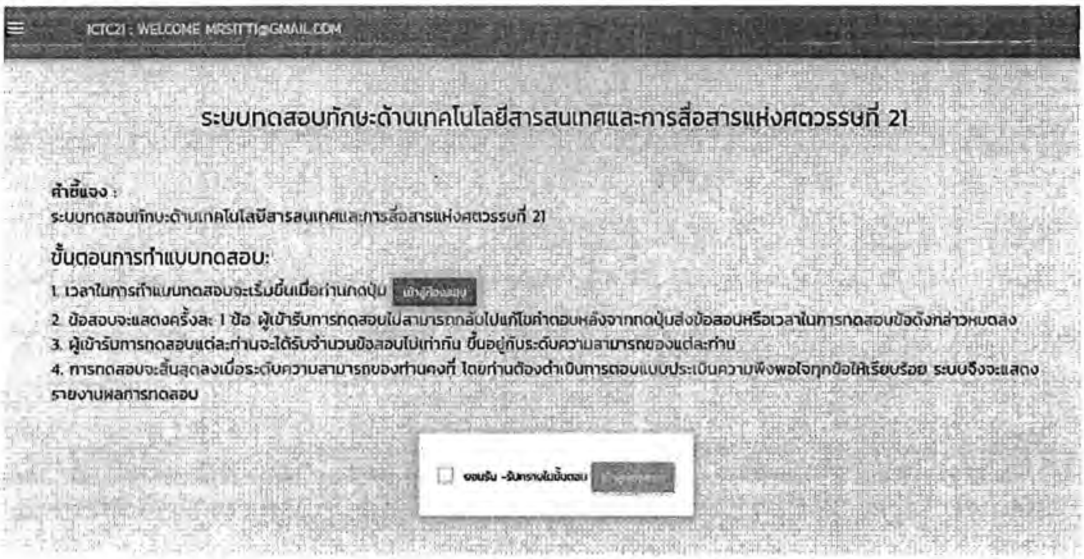


ภาพ 4.10 ข้อมูลของผู้เข้ารับการทดสอบเมื่อเข้าสู่ระบบ



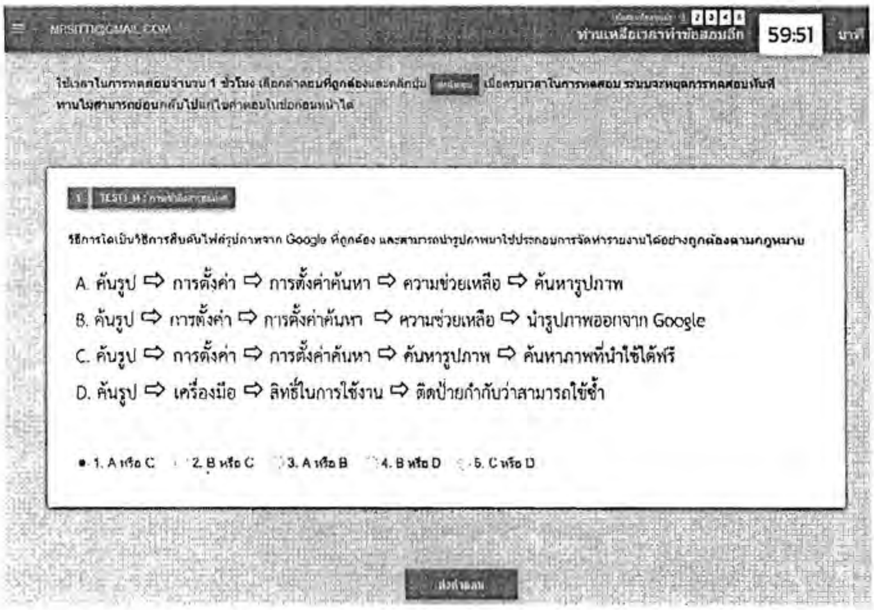
ภาพ 4.11 ข้อมูลของผู้เข้ารับการทดสอบหลังการอัพเดต

เมื่อผู้เข้ารับการทดสอบกดปุ่มเพื่อเข้าห้องสอบเรียบร้อย ระบบจะแสดงหน้าคำชี้แจงรายละเอียดเกี่ยวกับการทดสอบ ซึ่งระบบการทดสอบนี้เป็นระบบที่จัดทำขึ้นเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยรายละเอียดหน้าจอคำชี้แจงแสดงดังภาพ 4.12



ภาพ 4.12 หน้าแสดงรายละเอียดคำชี้แจง

เมื่อผู้เข้ารับการทดสอบอ่านคำชี้แจงเรียบร้อยแล้ว ผู้เข้ารับการทดสอบต้องกดยอมรับและรับทราบขั้นการการทดสอบและเข้าสู่ห้องสอบ และเมื่อกดปุ่มเข้าสู่ห้องสอบแล้วระบบจะแสดงข้อสอบครั้งละ 1 ข้อ ดังภาพ 4.13

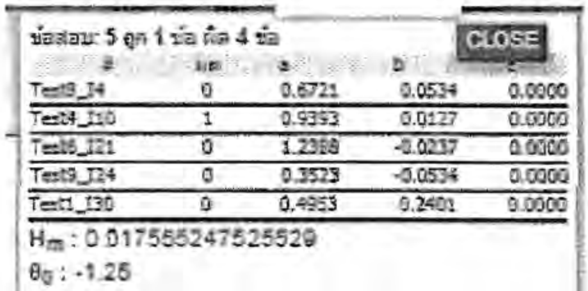


ภาพ 4.13 ข้อสอบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ในการทดสอบผู้เข้ารับการทดสอบจะได้รับข้อสอบครั้งละจำนวน 1 ข้อ โดยมีเนื้อหาครอบคลุมตามทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ทั้ง 5 องค์ประกอบ ได้แก่ 1) การเข้าถึงสารสนเทศ (Information accessibility) 2) การจัดการสารสนเทศ (Information management) 3) การบูรณาการสารสนเทศ (Information integration) 4) การประเมินสารสนเทศ (Information evaluation) และ 5) การสื่อสารสนเทศ (Information communication) โดยองค์ประกอบดังกล่าวจะแสดงอยู่ด้านบนของแบบทดสอบ พร้อมระยะเวลาในการทดสอบ สำหรับข้อสอบในชุดแรก ผู้จัดการทดสอบจะคัดเลือกมาจากคลังข้อสอบทั้งหมด โดยดำเนินการจัดชุดไว้ตามความครอบคลุมของเนื้อหาและเป็นข้อสอบที่มีระดับความยาก (difficulty) ปานกลาง หลังจากผู้เข้ารับการทดสอบตอบข้อสอบจำนวน 5 ข้อ ครบเสร็จเรียบร้อยแล้วครบทุกองค์ประกอบ ระบบจะทำการประมาณค่าความสามารถของผู้สอบ และนำค่าความสามารถที่ได้จากการประมาณค่าไปใช้ในการคัดเลือกข้อสอบในลำดับถัดไป สำหรับการเลือกข้อสอบใช้การพิจารณาจากค่าสารสนเทศสูงที่สุด (Maximum Item Information) และมีการควบคุมการเปิดเผยข้อสอบ (exposure rate) ไม่เกินร้อยละ 20 พร้อมทั้งควบคุมเนื้อหาของแบบทดสอบ (content balance) และระบบจะดำเนินการยุติการทดสอบเมื่อค่าความคลาดเคลื่อนมาตรฐาน (standard error) มีค่าน้อยกว่าหรือเท่ากับ 0.3 โดยรายละเอียดองค์ประกอบที่สำคัญในการทดสอบที่สำคัญมีดังนี้

3.1.3.1 การประมาณค่าความสามารถ

สำหรับการประมาณค่า ใช้การประมาณค่าร่วมกัน 2 วิธี คือ 1) วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) ใช้สำหรับกรณีผู้สอบที่ตอบข้อสอบถูกและผิดผสมกันและ 2) วิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ใช้สำหรับกรณีผู้สอบตอบข้อสอบถูกหรือผิดทุกข้อ โดยกระบวนการประมาณค่าความสามารถผู้สอบในการพัฒนาระบบได้ดำเนินการศึกษากระบวนการ (algorithm) ของโปรแกรม R จากชุดคำสั่งสำเร็จรูป (package) irtoys เพื่อนำมาพัฒนาเป็นระบบออนไลน์ โดยในการพัฒนาใช้ภาษา PHP ในการดำเนินการ ดังภาพ 4.14



ภาพ 4.14 การประมาณค่าความสามารถ

3.1.3.2 การคัดเลือกข้อสอบ

ในการเลือกข้อสอบสำหรับการวิจัยนี้ ใช้การเลือกข้อสอบจากค่าสารสนเทศสูงที่สุด (Maximum Item Information) และมีการควบคุมการเปิดเผยข้อสอบ (exposure rate) ไม่เกินร้อยละ 20

พร้อมทั้งควบคุมเนื้อหาของแบบทดสอบ (content balance) ให้เป็นไปตามองค์ประกอบ 5 องค์ประกอบ โดยขั้นตอนในการคัดเลือกข้อสอบมีดังนี้

1) แบ่งคลังข้อสอบการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ออกเป็น 5 กลุ่ม ตามองค์ประกอบ

2) นำค่าความสามารถ (θ) ที่ได้จากขั้นก่อนหน้าไปคำนวณสารสนเทศของข้อสอบ (item information: $I(\theta)$) ทุกข้อในคลังข้อสอบการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

3) สำหรับการคัดเลือกข้อสอบเริ่มตั้งแต่ข้อที่ 6 เป็นต้นไป โดยพิจารณาจากองค์ประกอบที่ 1 เป็นลำดับแรก ซึ่งจะดำเนินการเลือกข้อสอบข้อที่มีสารสนเทศของข้อสอบสูงที่สุดในองค์ประกอบที่ 1 และหากผู้เข้ารับการทดสอบยังไม่สามารถยุติการทดสอบได้ จะดำเนินการคัดเลือกข้อสอบจากองค์ประกอบที่ 2 ถึงองค์ประกอบที่ 5 เป็นลำดับต่อไป และเมื่อครบทั้ง 5 องค์ประกอบแล้วจะดำเนินการวนกลับมาพิจารณาจากองค์ประกอบที่ 1 อีกครั้งจนกว่าจะเป็นไปตามเกณฑ์การยุติในการทดสอบ ทั้งนี้มีกระบวนการควบคุมการเปิดเผยของข้อสอบไม่ให้มีการเลือกข้อสอบข้อเดิมซ้ำเกินร้อยละ 20 จากผู้เข้ารับการทดสอบทั้งหมด

3.1.3.3 การยุติการทดสอบ

สำหรับการยุติการทดสอบในการทดสอบผู้เข้ารับการทดสอบต้องดำเนินการทดสอบจำนวน 6 ข้อ เป็นอย่างน้อย กล่าวคือผู้เข้ารับการทดสอบต้องทำข้อสอบชุดแรกจำนวน 5 ข้อ ให้ครบตามองค์ประกอบและมีการคัดเลือกข้อสอบมาเพิ่มเติมจำนวน 1 ข้อ หลังจากดำเนินการประมาณค่าความสามารถของผู้สอบเรียบร้อยแล้ว เมื่อพบว่าค่าความสามารถของผู้เข้ารับการทดสอบมีความคงเส้นคงวา โดยที่ค่าความคลาดเคลื่อนมาตรฐาน (standard error) มีค่าน้อยกว่าหรือเท่ากับ 0.3 ระบบจะทำการยุติการทดสอบ และจะแสดงหน้าการประเมินคุณภาพระบบการทดสอบให้ผู้เข้ารับการทดสอบได้ดำเนินการประเมิน โดยผู้เข้ารับการทดสอบต้องทำการประเมินคุณภาพระบบก่อนออกจากระบบ หากไม่ดำเนินการประเมิน ระบบจะไม่แสดงผลรายงาน แต่หากผู้เข้ารับการทดสอบทำข้อสอบแล้วปรากฏว่าค่าความคลาดเคลื่อนมาตรฐานยังมากกว่า 0.3 ระบบจะทำการเลือกข้อสอบมาให้ผู้เข้ารับการทดสอบทำไปเรื่อย ๆ กระทั่งผลการตอบข้อสอบของผู้เข้ารับการทดสอบมีค่าความคลาดเคลื่อนมาตรฐานน้อยกว่าหรือเท่ากับ 0.3 ระบบจะยุติการทดสอบ หากผู้เข้ารับการทดสอบทำข้อสอบไปแล้วทั้งสิ้นจำนวน 35 ข้อแล้ว ความคลาดเคลื่อนมาตรฐานยังไม่น้อยกว่าหรือเท่ากับ 0.3 ระบบจะยุติการทดสอบโดยอัตโนมัติ และจะแสดงหน้าการประเมินคุณภาพระบบการทดสอบทันที ผู้เข้ารับการทดสอบต้องทำการประเมินคุณภาพระบบเพื่อเข้าดูรายงานผลการทดสอบต่อไป โดยรายละเอียดเมื่อผู้เข้ารับการทดสอบทำการทดสอบเสร็จสิ้นแสดงดังภาพ 4.15

ภาพ 4.15 หน้าแสดงผลเมื่อการทดสอบเสร็จสิ้น

3.1.4 การรายงานผลการทดสอบ

สำหรับการออกผลรายงานผลการทดสอบ ผู้เข้ารับการทดสอบจะต้องดำเนินการประเมินคุณภาพระบบด้วยแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ให้เรียบร้อยก่อน โดยผู้เข้ารับการทดสอบต้องทำการประเมินทุกข้อมิเช่นนั้นระบบจะไม่แสดงผลการทดสอบ เมื่อดำเนินการประเมินผลเรียบร้อยแล้ว ระบบจะรายงานผลการทดสอบให้ผู้เข้ารับการทดสอบทราบ โดยรายงานผลการทดสอบแบ่งออกเป็นสองส่วนได้แก่ รายงานผลการทดสอบ คือ การรายงานจำนวนข้อที่ผู้เข้ารับการทดสอบตอบข้อสอบถูกและผิดในแต่ละองค์ประกอบ และส่วนที่สอง คือ กราฟแสดงสัมพันธ์ของผลการทดสอบในแต่ละองค์ประกอบ โดยตัวอย่างรายงานผลการทดสอบแสดงดังภาพ 4.16 และตัวอย่างหน้าการประเมินคุณภาพระบบการทดสอบแสดงดังภาพ 4.17

เลขที่รายงาน : ICT2118081531-1
วันที่ : 2018 Aug Wed

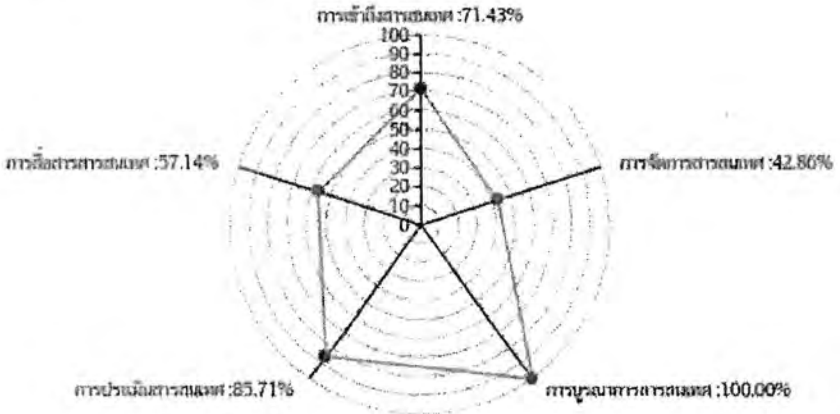
รายงานผลการทดสอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

ชื่อ : สิทธิ นามสกุล : ชัยสิทธิ ชั้นปีที่ : 2

คณะ : บริหารธุรกิจ มหาวิทยาลัย : รามคำแหง

องค์ประกอบ	จำนวนข้อที่ทำ	จำนวนที่ตอบถูก	คิดเป็นร้อยละ
1.การเข้าถึงสารสนเทศ	7	5	71.43%
2.การจัดการสารสนเทศ	7	3	42.86%
3.การบูรณาการสารสนเทศ	7	7	100.00%
4.การประเมินสารสนเทศ	7	6	85.71%
5.การสื่อสารสารสนเทศ	7	4	57.14%
รวม	35	25	71.43%

JS chart by amCharts



ภาพ 4.16 ตัวอย่างรายงานผลการทดสอบ

ภาพ 4.17 ตัวอย่างหน้าการประเมินคุณภาพระบบการทดสอบ

3.2 คู่มือการใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

สำหรับคู่มือการใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แบ่งออกเป็น 2 ส่วน ได้แก่ คู่มือการใช้งานสำหรับผู้ใช้งานทั่วไป (user) และคู่มือการใช้งานสำหรับผู้จัดการระบบ (admin) โดยแต่ละส่วนมีรายละเอียดดังนี้

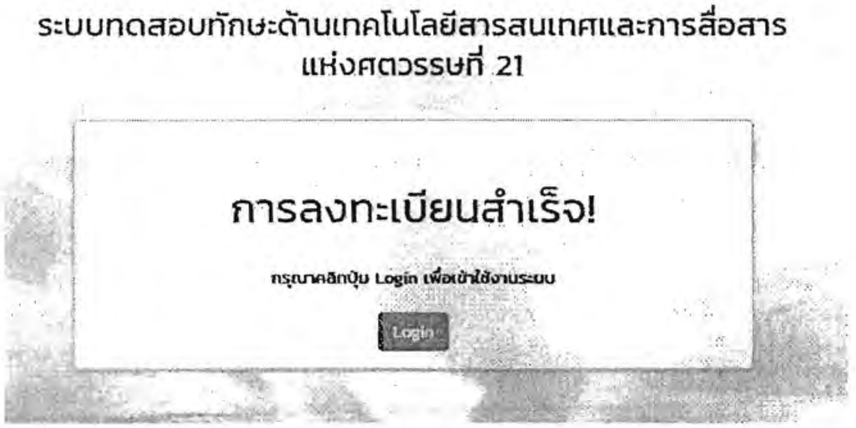
3.2.1 คู่มือการใช้งานสำหรับผู้ใช้งานทั่วไป (user)

<u>การลงทะเบียน</u>

1. เข้าสู่ระบบการทดสอบโดยพิมพ์ Url: http://www.oneinthai.com/ictc

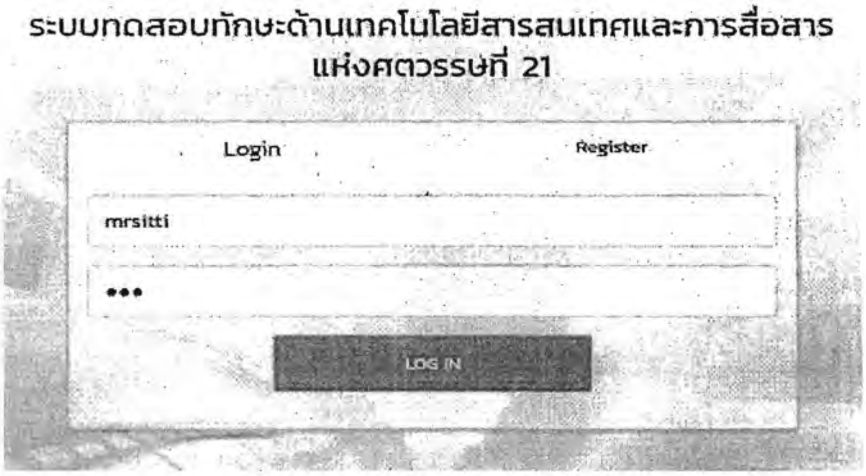
2. คลิกที่ Register เพื่อลงทะเบียนโดย Username และ Password ต้องประกอบด้วยอักษรอย่างน้อย 3 ตัวอักษร หาก Username ที่ต้องการใช้มีผู้ลงทะเบียนแล้ว ระบบจะทำการแจ้งเตือน

3. เมื่อลงทะเบียนเรียบร้อยระบบจะแสดงการลงทะเบียนสำเร็จ ดำเนินการคลิก Login เพื่อไปยังหน้าจอสำหรับเข้าระบบ



ภาพ 4.18 หน้าแสดงการลงทะเบียนสำเร็จ

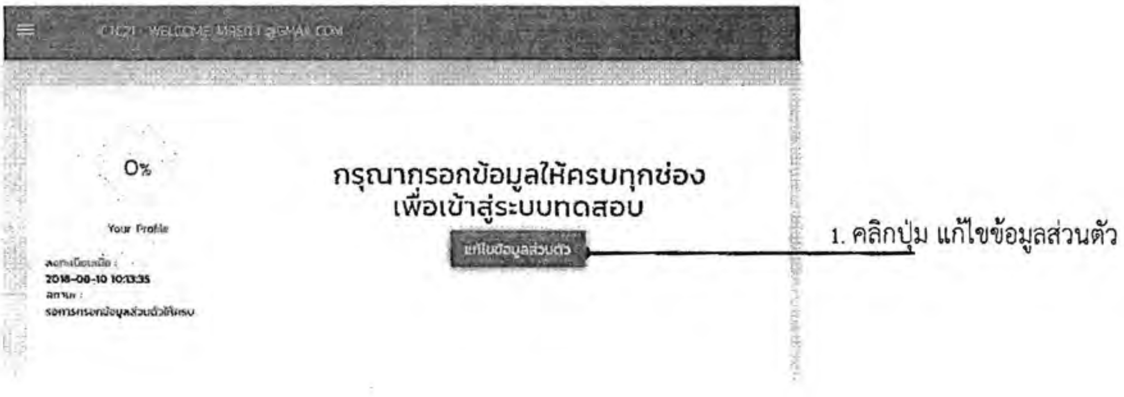
4. หน้าสำหรับเข้าใช้งานระบบ กรอก Username และ Password ให้ถูกต้อง หลังจากนั้นคลิก Log in เพื่อเข้าใช้งานระบบ

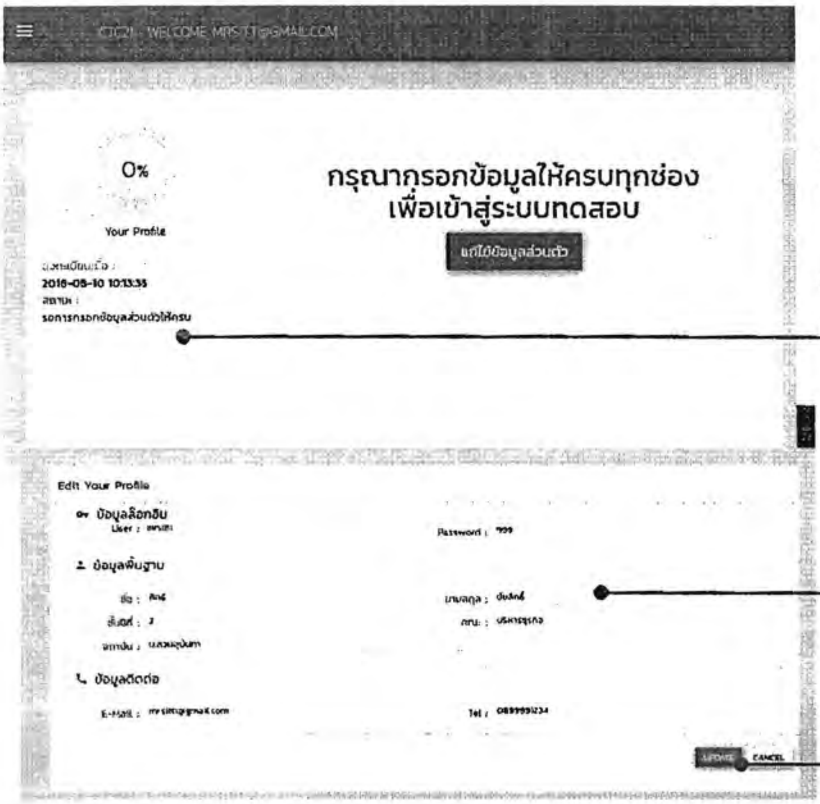


ภาพ 4.19 หน้าสำหรับเข้าใช้งานระบบ

<u>การเข้าใช้งานระบบครั้งแรก</u>

สำหรับการใช้งานระบบครั้งแรกต้องดำเนินการกรอกข้อมูลให้ครบทุกช่อง เพื่อรับสิทธิ์ในการเข้าระบบทดสอบ โดยมีรายละเอียดดังนี้





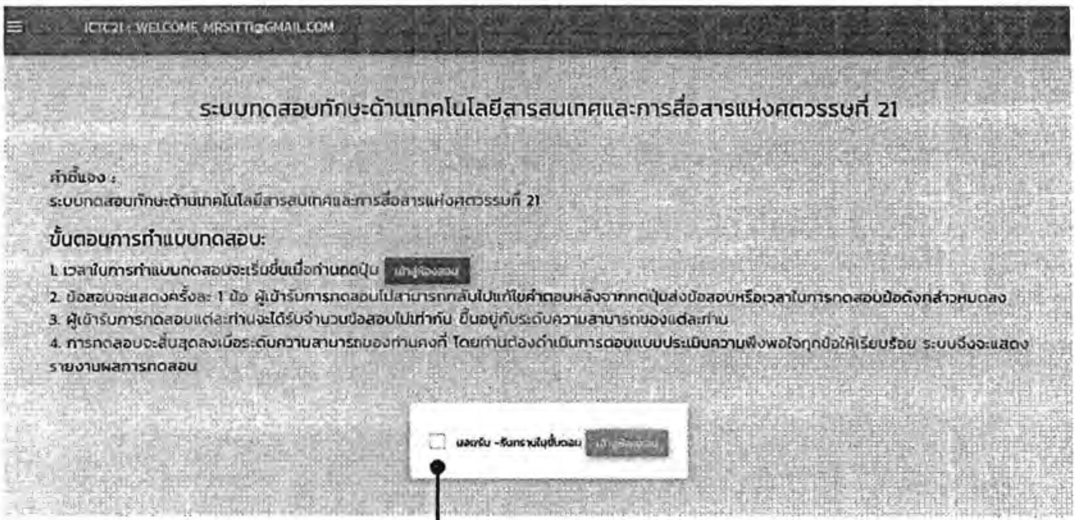
4. เมื่อกรอกข้อมูลครบแล้ว ระบบจะแสดงปุ่ม เข้าห้องสอบ คลิกที่ปุ่ม เข้าห้องสอบ เพื่อเริ่มทำแบบทดสอบ

2. กรอกข้อมูลให้ครบทุกช่อง

3. คลิก Update เพื่อบันทึก

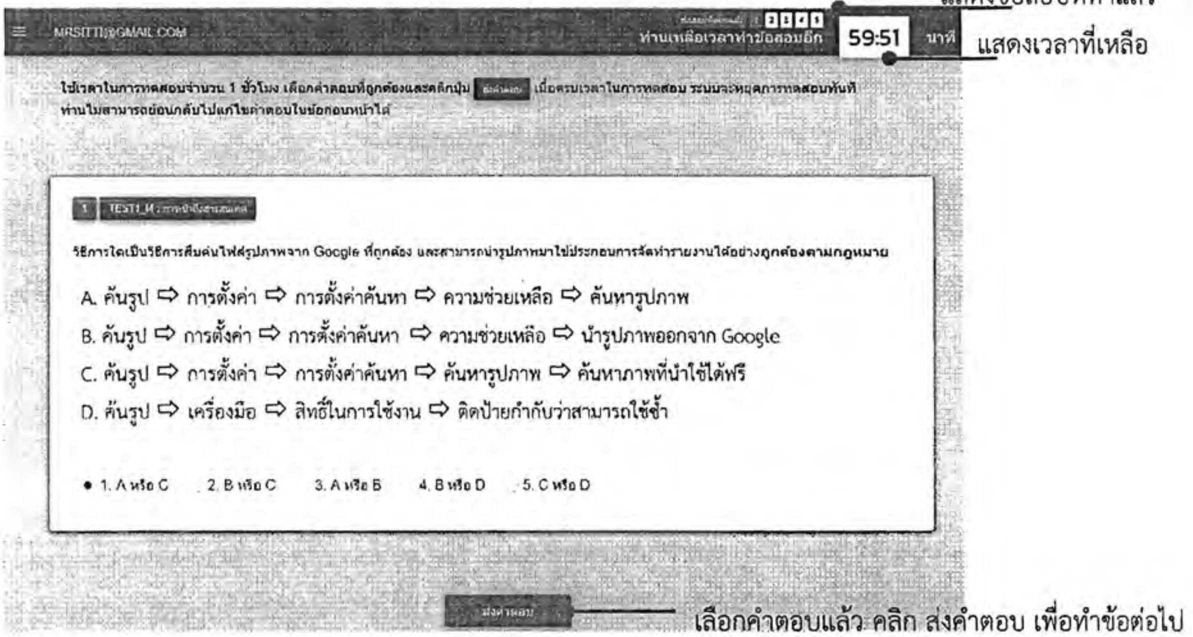
<u>การเข้าสู่ระบบเพื่อเริ่มทำแบบทดสอบ</u>

เมื่อกรอกข้อมูลส่วนตัวครบเรียบร้อยแล้ว สามารถเข้าสู่ระบบทดสอบได้ทันที โดยคลิกที่ปุ่ม เข้าห้องสอบ หลังจากนั้นจะพบหน้าคำชี้แจงก่อนเข้าสู่ระบบทดสอบ โดยรายละเอียดการเข้าสู่ระบบเพื่อเริ่มทำแบบทดสอบมีดังนี้



1. คลิกที่เช็คบ๊อก เพื่อยอมรับ

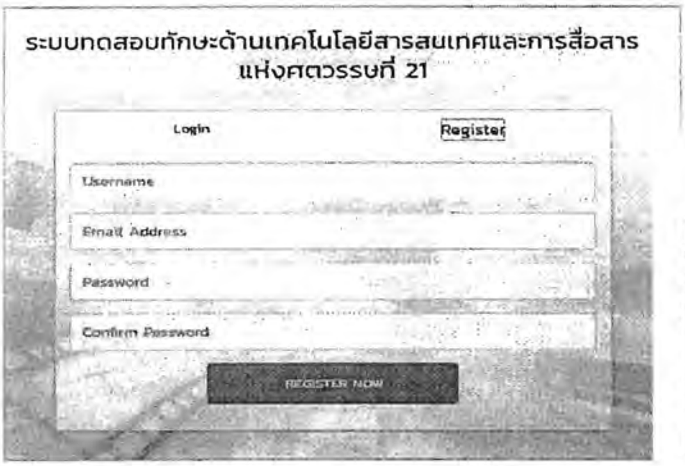
2. คลิกที่ปุ่ม เข้าสู่ห้องสอบ



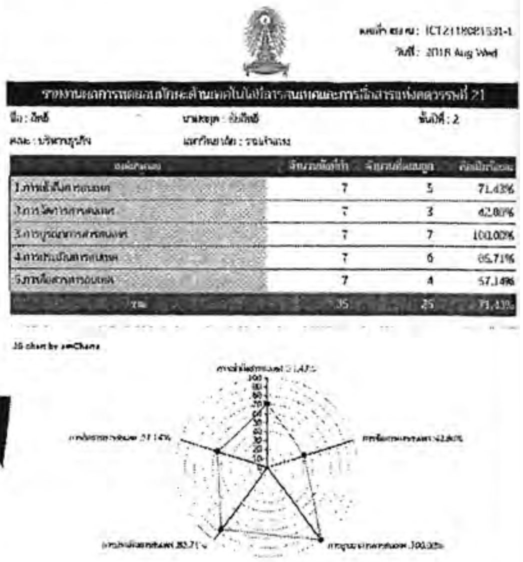
แสดงข้อสอบที่ทำแล้ว
MRSITTI@GMAIL.COM
ท่านเหลือเวลาทำข้อสอบอีก
59:51
นาที
แสดงเวลาที่เหลือ
ใช้เวลาในการทดสอบจำนวน 1 ชั่วโมง เลือกคำตอบที่ถูกต้องและคลิกปุ่ม ส่งคำตอบ เมื่อครบเวลาในการทดสอบ ระบบจะหยุดการทดสอบทันที ท่านไม่สามารถย้อนกลับไปแก้ไขคำตอบในข้อก่อนหน้าได้
1
วิธีการใดเป็นวิธีการสืบค้นไฟล์รูปภาพจาก Google ที่ถูกต้อง และสามารถนำรูปภาพมาใช้ประกอบการจัดทำรายงานได้อย่างถูกต้องตามกฎหมาย
A. ค้นรูป ⇨ การตั้งค่า ⇨ การตั้งค่าค้นหา ⇨ ความช่วยเหลือ ⇨ ค้นหารูปภาพ
B. ค้นรูป ⇨ การตั้งค่า ⇨ การตั้งค่าค้นหา ⇨ ความช่วยเหลือ ⇨ นำรูปภาพออกจาก Google
C. ค้นรูป ⇨ การตั้งค่า ⇨ การตั้งค่าค้นหา ⇨ ค้นหารูปภาพ ⇨ ค้นหาภาพที่นำใช้ได้ฟรี
D. ค้นรูป ⇨ เครื่องมือ ⇨ สิทธิ์ในการใช้งาน ⇨ ติดป้ายกำกับว่าสามารถใช้ซ้ำ
1. A หรือ C
2. B หรือ C
3. A หรือ B
4. B หรือ D
5. C หรือ D
ส่งคำตอบ
เลือกคำตอบแล้ว คลิก ส่งคำตอบ เพื่อทำข้อต่อไป

<u>การรายงานผลการทดสอบ</u>

เมื่อผู้เข้ารับการทดสอบทำแบบทดสอบครบตามเงื่อนไขจะแสดงรายละเอียดดังนี้



หน้าแสดงรายงานผลการสอบ



ตอบแบบสอบถามให้ครบทุกข้อ เพื่อรับผลการสอบ

3.2.2 คู่มือการใช้งานสำหรับผู้จัดการระบบ (admin)

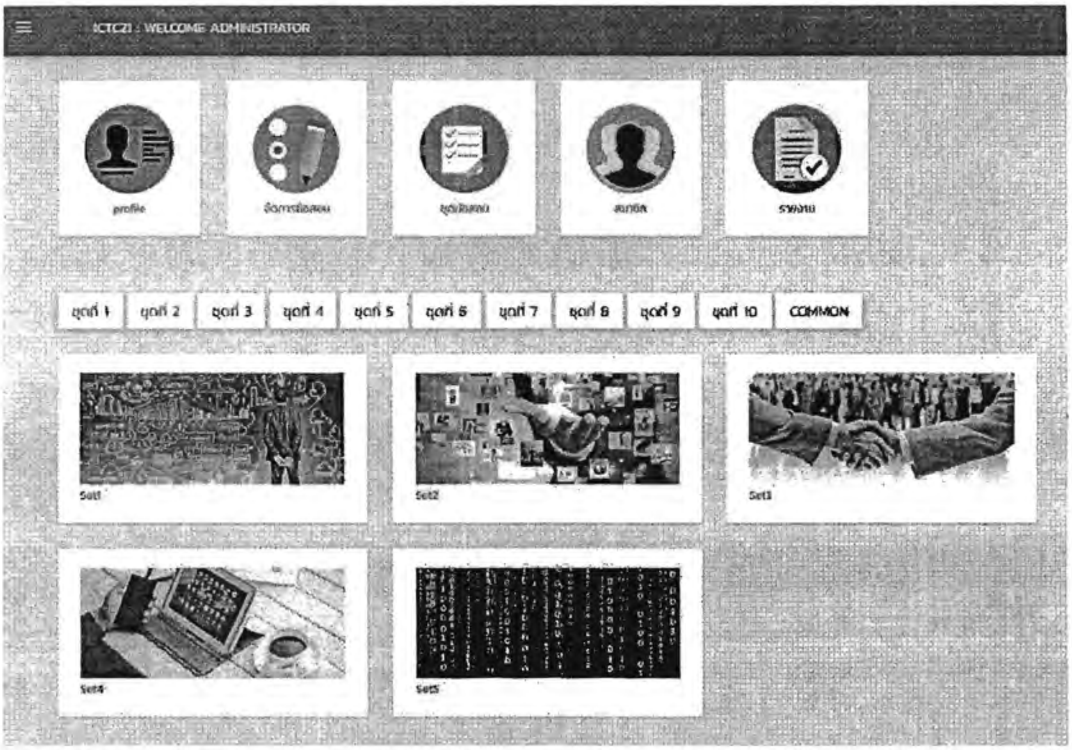
<u>การเข้าสู่ระบบ</u>

1. เข้าสู่ระบบการทดสอบโดยพิมพ์ Url: http://www.oneinthai.com/ictc

2. ใส่ Username และ Password ของ Admin โดยรายละเอียดข้อมูลหน้า Admin ประกอบด้วยข้อมูลดังนี้

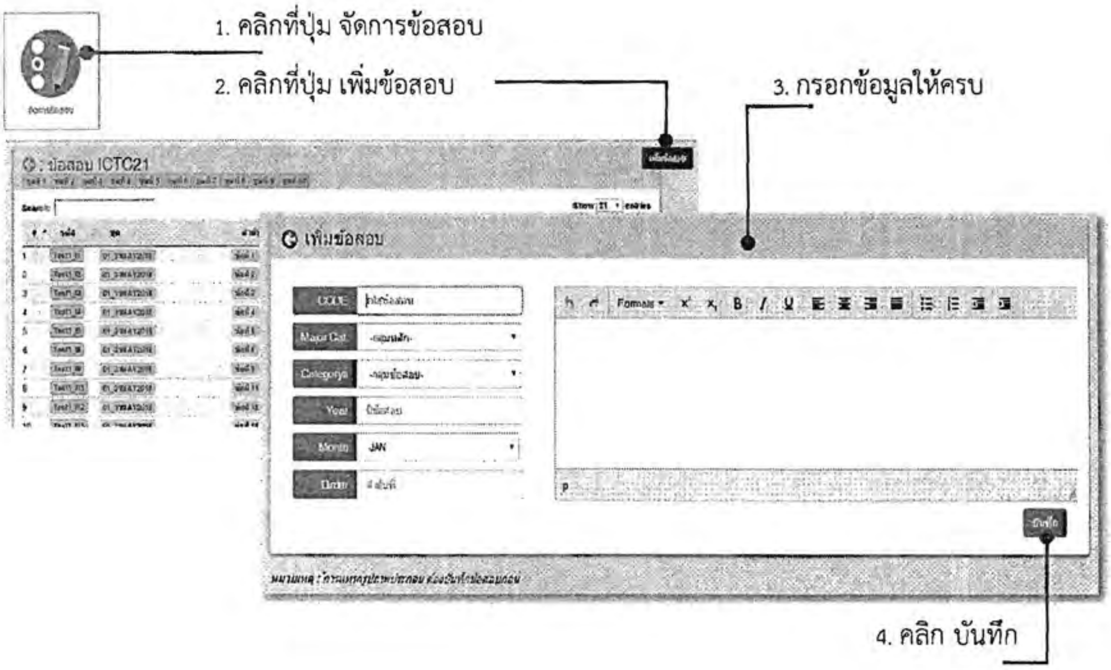
- Profile : แก้ไขข้อมูล & รหัสผ่าน สำหรับ Admin
- จัดการข้อสอบ : เพิ่ม/แก้ไข ข้อสอบในระบบ
- ชุดข้อสอบ : สร้างชุดข้อสอบ

- สมาชิก : จัดการข้อมูลสมาชิกที่ใช้งานระบบ
- รายงาน : ดูผลสอบของผู้สอบแต่ละคน
- ชุดที่ .... : สำหรับตรวจสอบข้อสอบตามชุดต่าง

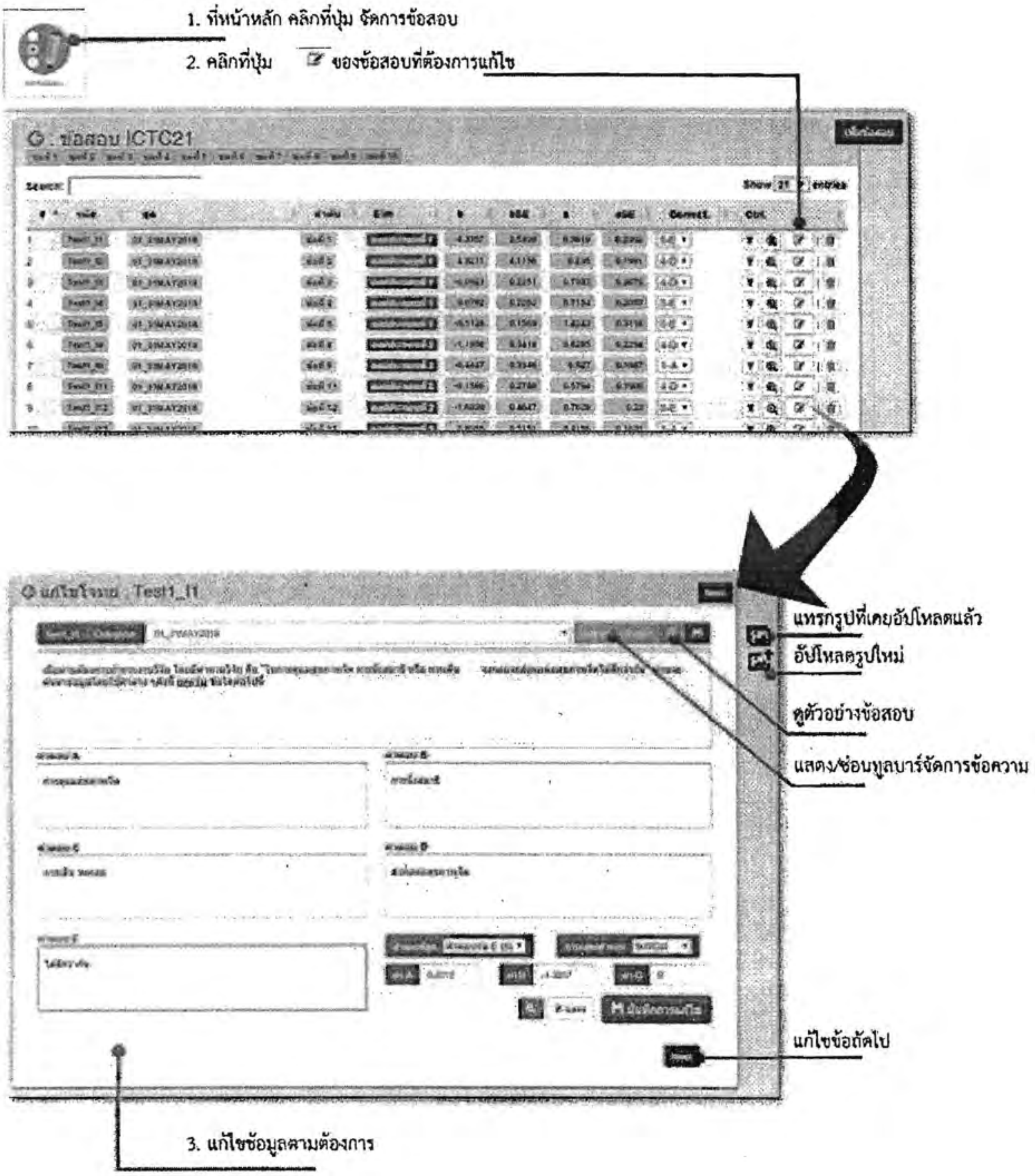


ภาพ 4.20 รายละเอียดหน้าการจัดการของ Admin

<u>การจัดการข้อสอบ</u>



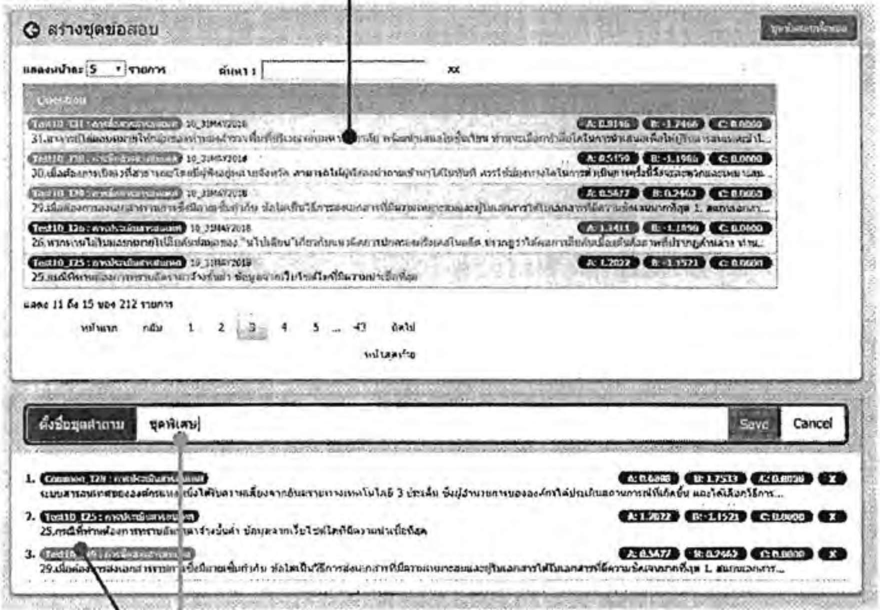
การแก้ไขข้อสอบ



การสร้างชุดข้อสอบ

1. ที่หน้าหลัก คลิกที่ปุ่ม ชุดข้อสอบ

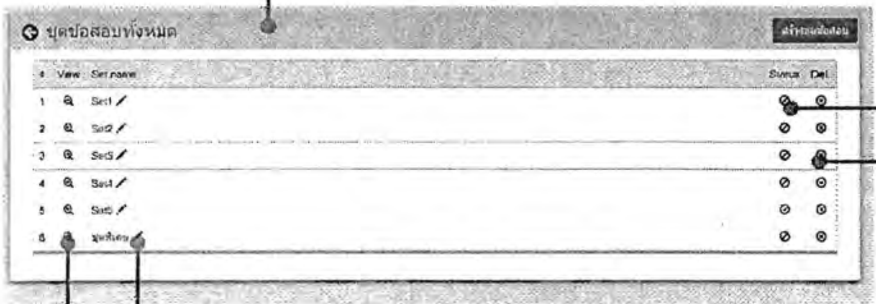
2. คลิกที่ข้อสอบที่ต้องการ



3. รายการข้อสอบที่คลิกเลือกไว้

4. ตั้งชื่อชุดตามต้องการแล้วคลิก Save

5. ระบบแสดงรายชื่อชุดข้อสอบที่มีอยู่



สถานะเปิด/ปิด การใช้ชุดข้อสอบ

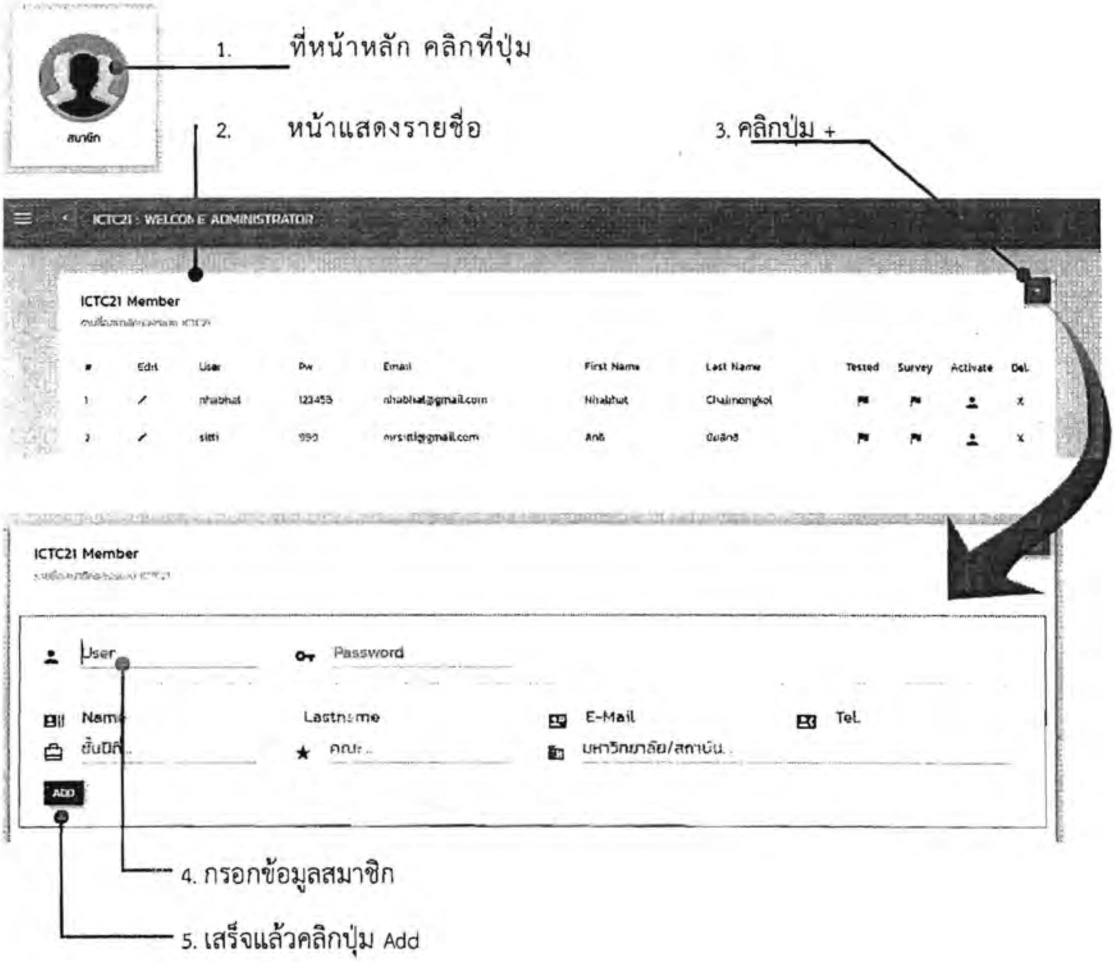
ลบชุดข้อสอบนี้ทิ้ง

คลิก เพื่อแก้ไขชื่อชุด

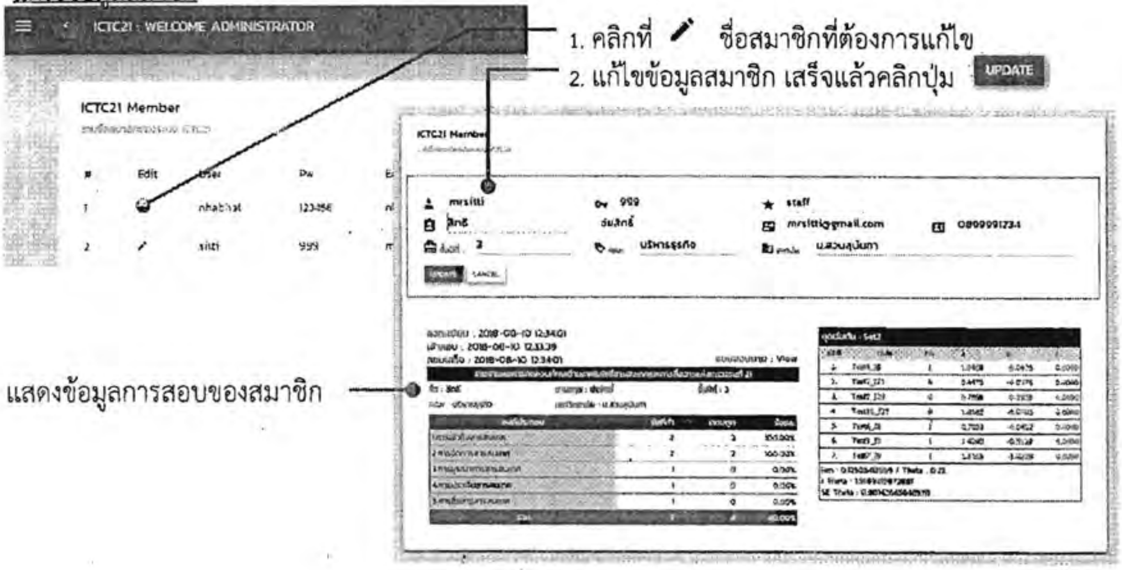
คลิก เพื่อดูตัวอย่างข้อสอบ

การจัดการสมาชิก

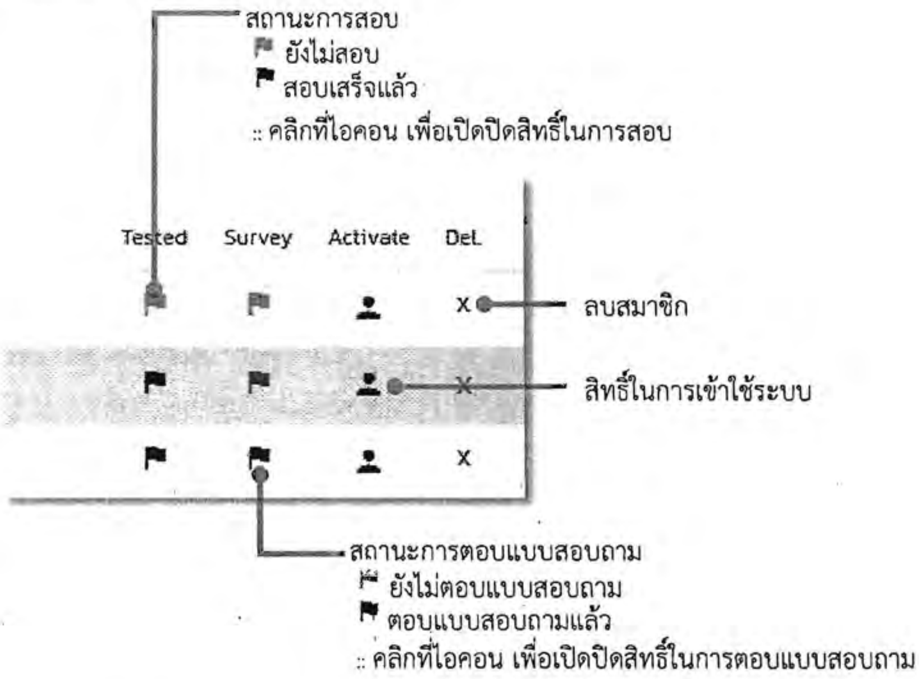
เพิ่มสมาชิกใหม่



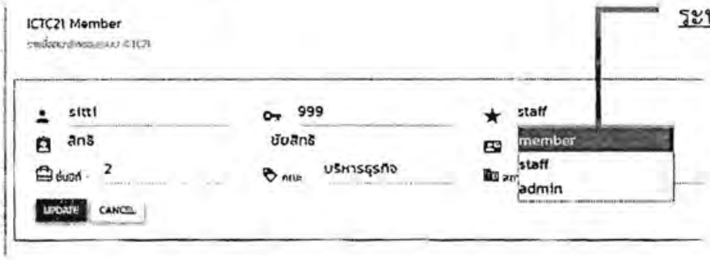
แก้ไขข้อมูลสมาชิก



ปุ่มคำสั่งที่หน้า จัดการสมาชิก



ระดับสิทธิ์ของสมาชิก



ระบบสามารถกำหนดสิทธิ์ได้ 3 ระดับ

1 member คือสมาชิกทั่วไป

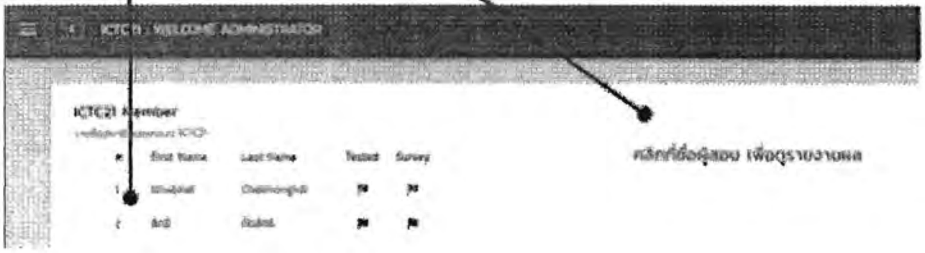
2 staff ระบบจะแสดงคำเฉลยและค่า Theta

3 Admin สามารถจัดการข้อมูลได้ทั้งหมด

<u>ดูรายงานผลการสอบของสมาชิก</u>

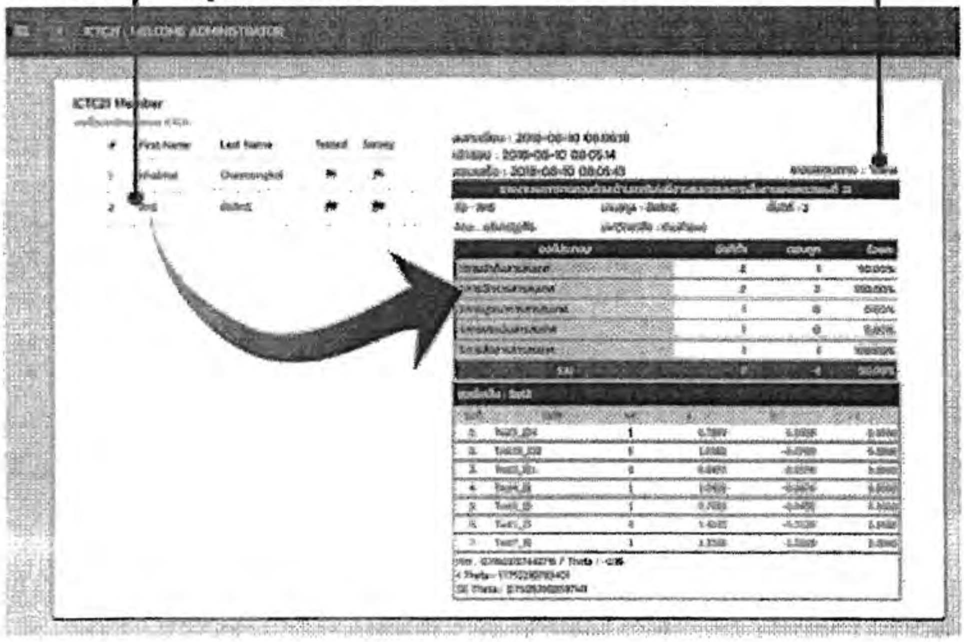
1. ที่หน้าหลัก คลิกที่ปุ่ม รายงาน

2. หน้ารายงานแบ่งเป็นสองส่วน
   ฝั่งซ้ายแสดงรายชื่อสมาชิก ฝั่งขวา แสดงรายงาน



3. คลิกที่ชื่อสมาชิก เพื่อดูรายงานผลการสอบ

ดูข้อมูลการตอบแบบสอบถามของสมาชิก



ตอนที่ 4 ผลการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้

ผู้วิจัยนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้กับนิสิตนักศึกษาสายสังคมศาสตร์ จำนวน 119 คน และสายวิทยาศาสตร์ จำนวน 98 คน ที่สมัครใจเข้าร่วมการทดลองใช้ระบบครั้งนี้ โดยมีผลการทดลองใช้ดังรายละเอียดต่อไปนี้

4.1 การทดลองใช้ระบบฯกับนิสิตนักศึกษาปริญญาตรี สายสังคมศาสตร์

ผู้วิจัยทดลองใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี กับนิสิตนักศึกษาสายสังคมศาสตร์ จำนวน 119 คน แบ่งเป็นชาย จำนวน 44 คน (ร้อยละ 36.97) และหญิง จำนวน 75 คน (ร้อยละ 63.03) โดยผู้วิจัยแบ่งการนำเสนอออกเป็น 2 ประเด็น ได้แก่ (1) ผลการวัดความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 (θ) และ(2) ผลการวัดคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ดังรายละเอียดต่อไปนี้

4.1.1 ผลการวัดความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 (θ)

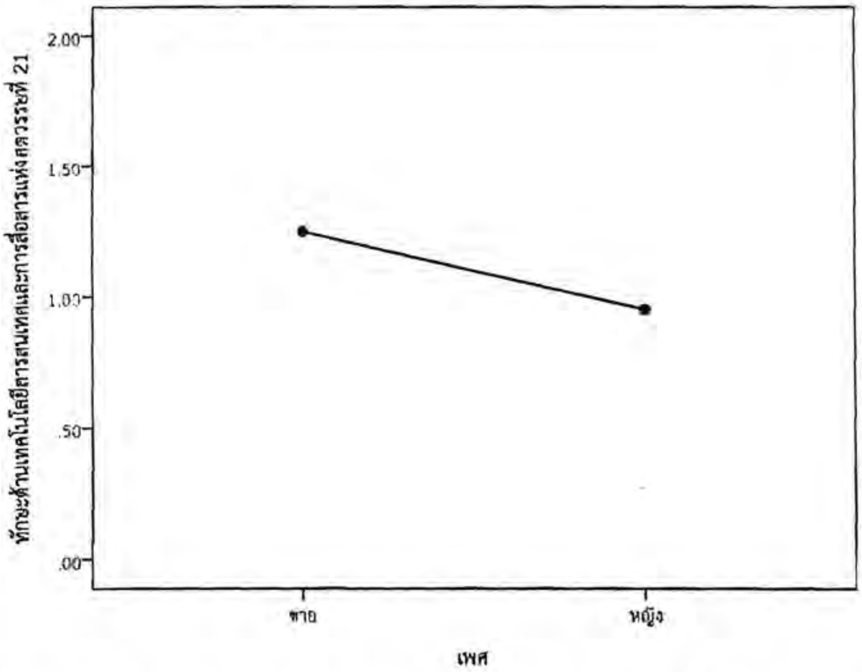
นิสิตนักศึกษามีความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เฉลี่ยเท่ากับ 1.06 (*SD* = 0.78) โดยความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 มีการกระจายใกล้เคียงกับโค้งปกติ (*sk* = -0.16) แสดงว่านิสิตนักศึกษาส่วนใหญ่มีความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ใกล้เคียงกับค่าเฉลี่ย และมีความโด่งใกล้เคียงกับโค้งปกติ (*ku* = -0.08) แสดงว่าความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษามีการแจกแจงใกล้เคียงโค้งปกติ

ผลการวิเคราะห์ค่าสถิติพื้นฐานของความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 จำแนกตามเพศ พบว่า นิสิตนักศึกษาชายมีความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เฉลี่ย เท่ากับ 1.25 (*SD* = 0.88) มีการแจกแจงแบบเบ้ซ้ายเล็กน้อย (*sk* =-0.47) และมีความโด่งมากกว่าโค้งปกติ (*ku* = 0.25) แสดงว่ามีการกระจายน้อยกว่าโค้งปกติ ส่วนนิสิตนักศึกษาหญิงมีความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เฉลี่ย เท่ากับ 0.95 (*SD* = 0.70) มีการแจกแจงใกล้เคียงกับโค้งปกติ (*sk* = -0.10) และมีความโด่งน้อยกว่าโค้งปกติ (*ku* = -0.29) แสดงว่ามีการกระจายมากกว่าโค้งปกติ ดังตาราง 4.30

ผลการทดสอบความแตกต่างระหว่างความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาชายและหญิง พบว่า นิสิตนักศึกษามีความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 สูงกว่านิสิตนักศึกษาหญิงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (117) = 2.03, p = .045, d = 0.38 ดังตาราง 4.30 โดยมีกราฟค่าเฉลี่ยของความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 จำแนกตามเพศ ดังภาพ 4.21

ตาราง 4.30 ค่าสถิติพื้นฐานของความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาสายสังคมศาสตร์

	ความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21						
เพศ	*M*	*SD*	*min*	*max*	*CV%*	*sk (SE)*	*ku (SE)*
ชาย	1.25	0.88	-1.15	3.05	70.40	-0.47 (0.36)	0.25 (0.70)
หญิง	0.95	0.70	-0.82	2.54	73.68	-0.10 (0.28)	-0.29 (0.55)
	t (117) = 2.03, p = .045, d = 0.38						
รวม	1.06	0.78	-1.15	3.05	73.58	-0.16 (0.22)	-0.08 (0.44)



ภาพ 4.21 กราฟค่าเฉลี่ยของความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 (θ) จำแนกตามเพศของนิสิตนักศึกษาสายสังคมศาสตร์

4.1.2 คะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบ

ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แบ่งเป็น 5 องค์ประกอบ ได้แก่ การเข้าถึงสารสนเทศ การจัดการสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ และการสื่อสารสารสนเทศ ซึ่งนิสิตนักศึกษาสายสังคมศาสตร์มีคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบ ดังรายละเอียดต่อไปนี้

นิสิตนักศึกษามีคะแนนด้านการเข้าถึงสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 75.00 (*SD* = 17.37) โดยนิสิตนักศึกษาชาย (*M* = 75.76, *SD* = 17.97) และหญิง (*M* = 74.57, *SD* = 17.12) มีคะแนนด้านการเข้าถึงสารสนเทศไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (117) = 0.36, p = .72, d = 0.07

นิสิตนักศึกษามีคะแนนด้านการจัดการสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 48.92 (*SD* = 14.64) โดยนิสิตนักศึกษาชาย (*M* = 50.43, *SD* = 15.56) และหญิง (*M* = 48.03, *SD* = 14.11) มีคะแนนด้านการจัดการสารสนเทศไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (117) = 0.86, p = .39, d = 0.16

นิสิตนักศึกษามีคะแนนด้านการบูรณาการสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 71.27 (*SD* = 19.76) โดยนิสิตนักศึกษาชาย (*M* = 71.23, *SD* = 20.92) และหญิง (*M* = 71.29, *SD* = 19.19) มีคะแนนด้านการบูรณาการสารสนเทศไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (117) = 0.02, p = .99, d = 0.003

นิสิตนักศึกษามีคะแนนด้านการประเมินสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 72.12 (*SD* = 14.75) โดยนิสิตนักศึกษาชาย (*M* = 73.68, *SD* = 13.86) และหญิง (*M* = 71.20, *SD* = 15.26) มีคะแนนด้านการประเมินสารสนเทศ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (117) = 0.89, p = .38, d = 0.17

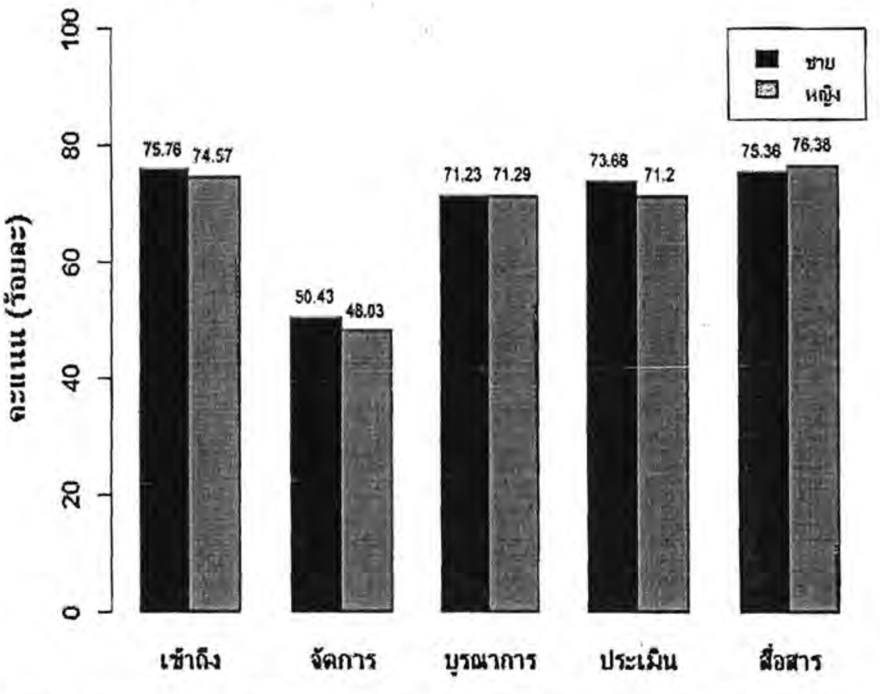
นิสิตนักศึกษามีคะแนนด้านการสื่อสารสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 76.00 (*SD* = 17.54) โดยนิสิตนักศึกษาชาย (*M* = 75.36, *SD* = 19.59) และหญิง (*M* = 76.38, *SD* = 16.35) มีคะแนนด้านการสื่อสารสารสนเทศไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (117) = 0.31, p = .76, d = 0.06

ค่าสถิติพื้นฐานของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบของนิสิตนักศึกษาแสดงดังตาราง 4.31 และภาพ 4.22 จะเห็นได้ว่านิสิตนักศึกษาชายและหญิงมีคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบใกล้เคียงกัน โดยนิสิตนักศึกษาชายมีคะแนนเฉลี่ยด้านการเข้าถึงสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด ในขณะที่นิสิตนักศึกษาหญิงมีคะแนนเฉลี่ยด้านการสื่อสารสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด ทั้งนี้นิสิตนักศึกษาชายและหญิงมีคะแนนเฉลี่ยด้านการเข้าถึงสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ และการสื่อสารสารสนเทศค่อนข้างสูง แต่มีคะแนนเฉลี่ยด้านการจัดการสารสนเทศค่อนข้างต่ำ

ตาราง 4.31 ค่าสถิติพื้นฐานของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบของนิสิตนักศึกษาสายสังคมศาสตร์

	คะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21						
เพศ	*M*	*SD*	*min*	*max*	*CV%*	*sk (SE)*	*ku (SE)*
การเข้าถึงสารสนเทศ							
รวม	75.00	17.37	16.67	100.00	23.16	-0.51 (0.22)	0.20 (0.44)
ชาย	75.76	17.97	16.67	100.00	23.72	-0.90 (0.36)	1.43 (0.70)
หญิง	74.57	17.12	28.57	100.00	22.96	-0.26 (0.28)	-0.44 (0.55)
	t (117) = 0.36, p = .72, d = 0.07						
การจัดการสารสนเทศ							
รวม	48.92	14.64	0.00	85.71	29.93	-0.25 (0.22)	0.66 (0.44)
ชาย	50.43	15.56	14.29	85.71	30.85	-0.46 (0.36)	0.46 (0.70)
หญิง	48.03	14.11	0.00	85.71	28.38	-0.15 (0.28)	1.06 (0.55)
	t (117) = 0.86, p = .39, d = 0.16						

การบูรณาการสารสนเทศ							
รวม	71.27	19.76	14.29	100.00	27.73	-0.68 (0.22)	0.06 (0.44)
ชาย	71.23	20.92	20.00	100.00	29.37	-0.48 (0.36)	-0.64 (0.70)
หญิง	71.29	19.19	14.29	100.00	26.92	-0.83 (0.28)	0.67 (0.55)
$t\,(117) = 0.02, p = .99, d = 0.003$							
การประเมินสารสนเทศ							
รวม	72.12	14.75	28.57	100.00	20.45	-0.16 (0.22)	-0.24 (0.44)
ชาย	73.68	13.86	42.86	100.00	18.81	0.16 (0.36)	-0.46 (0.70)
หญิง	71.20	15.26	28.57	100.00	21.43	-0.27 (0.28)	-0.24 (0.55)
$t\,(117) = 0.89, p = .38, d = 0.17$							
การสื่อสารสารสนเทศ							
รวม	76.00	17.54	0.00	100.00	23.08	-1.00 (0.22)	2.12 (0.44)
ชาย	75.36	19.59	0.00	100.00	26.00	-1.72 (0.36)	4.37 (0.70)
หญิง	76.38	16.35	40.00	100.00	21.41	-0.28 (0.28)	-0.63 (0.55)
$t\,(117) = 0.31, p = .76, d = 0.06$							



ภาพ 4.22 ค่าเฉลี่ยของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบจำแนกตามเพศของนิสิตนักศึกษาสายสังคมศาสตร์

4.2 การทดลองใช้ระบบฯกับนิสิตนักศึกษาปริญญาตรี สายวิทยาศาสตร์

ผู้วิจัยทดลองใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี กับนิสิตนักศึกษาสายวิทยาศาสตร์ จำนวน 98 คน แบ่งเป็นชาย จำนวน 60 คน (ร้อยละ 61.22) และหญิง จำนวน 38 คน (ร้อยละ 38.78) โดยผู้วิจัยแบ่งการนำเสนอออกเป็น 2 ประเด็น ได้แก่ (1) ผลการวัดความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 (θ) และ(2) ผลการวัดคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ดังรายละเอียดต่อไปนี้

4.2.1 ผลการวัดความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 (θ)

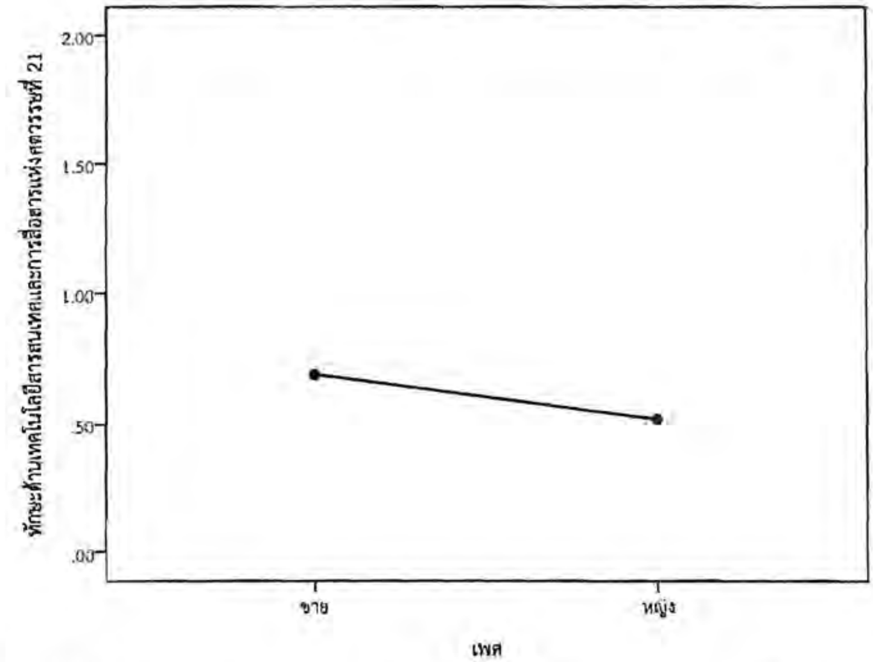
นิสิตนักศึกษามีความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เฉลี่ยเท่ากับ 0.62 (*SD* = 0.75) มีการแจกแจงแบบเบ้ขวา (*sk* = 0.59) และมีความโด่งใกล้เคียงกับโค้งปกติ (*ku* = 0.14) แสดงว่าความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษามีการกระจายใกล้เคียงโค้งปกติ

ผลการวิเคราะห์ค่าสถิติพื้นฐานของความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 จำแนกตามเพศ พบว่า นิสิตนักศึกษาชายมีความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เฉลี่ย เท่ากับ 0.69 (*SD* = 0.73) มีการแจกแจงแบบเบ้ขวาเล็กน้อย (*sk* = 0.62) และมีความโด่งใกล้เคียงกับโค้งปกติ (*ku* = -0.17) แสดงว่าความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาชายมีการกระจายใกล้เคียงโค้งปกติ ส่วนนิสิตนักศึกษาหญิงมีความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เฉลี่ย เท่ากับ 0.52 (*SD* = 0.80) มีการแจกแจงแบบเบ้ขวาเล็กน้อย (*sk* = 0.65) และมีความโด่งมากกว่าโค้งปกติ (*ku* = 0.71) แสดงว่าความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาหญิงมีการกระจายน้อยกว่าโค้งปกติ ดังตาราง 4.32

ผลการทดสอบความแตกต่างระหว่างความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาชายและหญิง พบว่า นิสิตนักศึกษาชายและหญิงมีความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (96) = 1.08, p = .29, d = 0.22 ดังตาราง 4.32 โดยมีกราฟค่าเฉลี่ยของความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 จำแนกตามเพศ ดังภาพ 4.23

ตาราง 4.32 ค่าสถิติพื้นฐานของความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาสายวิทยาศาสตร์

เพศ	ความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21						
	M	*SD*	*min*	*max*	*CV%*	*sk* (*SE*)	*ku* (*SE*)
ชาย	0.69	0.73	-0.47	2.63	105.80	0.62 (0.31)	-0.17 (0.61)
หญิง	0.52	0.80	-1.06	2.45	153.85	0.65 (0.38)	0.71 (0.75)
	t (96) = 1.08, p = .29, d = 0.22						
รวม	0.62	0.75	-1.06	2.63	120.97	0.59 (0.24)	0.14 (0.48)



ภาพ 4.23 กราฟค่าเฉลี่ยของความสามารถทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 (θ) จำแนกตามเพศของนิสิตนักศึกษาสายวิทยาศาสตร์

4.2.2 คะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบ

ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แบ่งเป็น 5 องค์ประกอบ ได้แก่ การเข้าถึงสารสนเทศ การจัดการสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ และการสื่อสารสารสนเทศ ซึ่งนิสิตนักศึกษาสายวิทยาศาสตร์มีคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบ ดังรายละเอียดต่อไปนี้

นิสิตนักศึกษามีคะแนนด้านการเข้าถึงสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 69.64 (SD = 16.94) โดยนิสิตนักศึกษาชาย (M = 72.18, SD = 16.46) และหญิง (M = 65.61, SD = 17.11) มีคะแนนด้านการเข้าถึงสารสนเทศไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (96) = 1.90, p = .06, d = 0.39

นิสิตนักศึกษามีคะแนนด้านการจัดการสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 54.25 (SD = 17.17) โดยนิสิตนักศึกษาชาย (M = 56.51, SD = 16.25) และหญิง (M = 50.68, SD = 18.19) มีคะแนนด้านการจัดการสารสนเทศไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (96) = 1.65, p = .10, d = 0.34

นิสิตนักศึกษามีคะแนนด้านการบูรณาการสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 68.48 (SD = 20.98) โดยนิสิตนักศึกษาชาย (M = 67.72, SD = 21.81) และหญิง (M = 69.67, SD = 19.81) มีคะแนนด้านการบูรณาการสารสนเทศไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (96) = 0.45, p = .66, d = 0.09

นิสิตนักศึกษามีคะแนนด้านการประเมินสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 68.44 (SD = 16.20) โดยนิสิตนักศึกษาชาย (M = 68.78, SD = 17.31) และหญิง (M = 67.89, SD = 14.48) มีคะแนนด้านการประเมินสารสนเทศ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (96) = 0.26, p = .79, d = 0.06

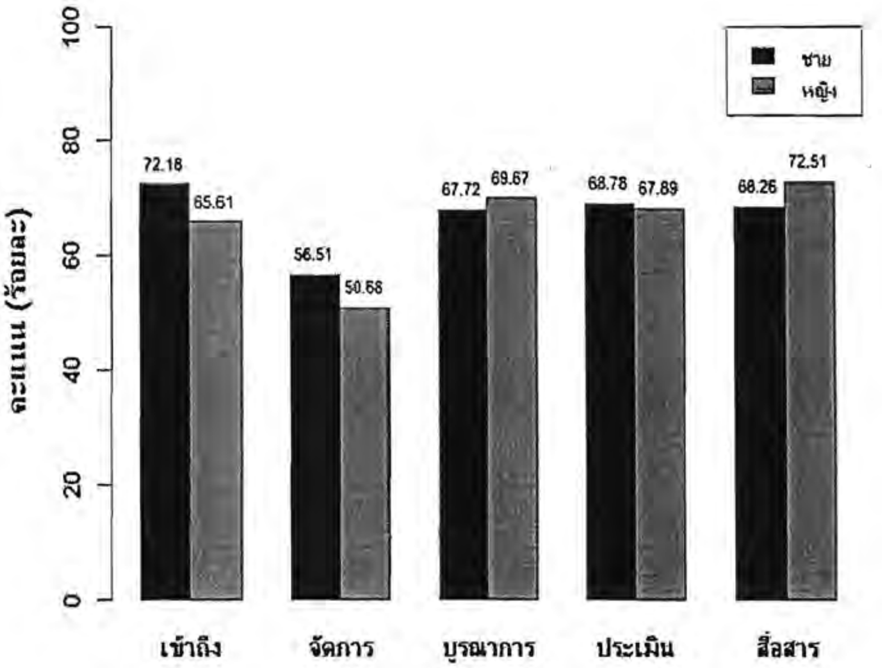
นิสิตนักศึกษามีคะแนนด้านการสื่อสารสารสนเทศเฉลี่ย เท่ากับ ร้อยละ 69.91 (*SD* = 15.75) โดยนิสิตนักศึกษาชาย (*M* = 68.26, *SD* = 15.54) และหญิง (*M* = 72.51, *SD* = 15.94) มีคะแนนด้านการสื่อสารสารสนเทศไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, t (96) = 1.30, p = .20, d = 0.27

ค่าสถิติพื้นฐานของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบของนิสิตนักศึกษาแสดงดังตาราง 4.33 และภาพ 4.24 จะเห็นได้ว่านิสิตนักศึกษาชายและหญิงมีคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบใกล้เคียงกัน โดยนิสิตนักศึกษาชายมีคะแนนเฉลี่ยด้านการเข้าถึงสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด ในขณะที่นิสิตนักศึกษาหญิงมีคะแนนเฉลี่ยด้านการสื่อสารสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด ทั้งนี้นิสิตนักศึกษาชายและหญิงมีคะแนนเฉลี่ยด้านการเข้าถึงสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ และการสื่อสารสารสนเทศค่อนข้างสูง แต่มีคะแนนเฉลี่ยด้านการจัดการสารสนเทศค่อนข้างต่ำ

ตาราง 4.33 ค่าสถิติพื้นฐานของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบของนิสิตนักศึกษาสายวิทยาศาสตร์

	คะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21						
เพศ	*M*	*SD*	*min*	*max*	*CV%*	*sk (SE)*	*ku (SE)*
การเข้าถึงสารสนเทศ							
รวม	69.64	16.94	16.67	100.00	24.33	-0.07 (0.24)	0.15 (0.48)
ชาย	72.18	16.46	16.67	100.00	22.80	-0.26 (0.31)	0.80 (0.61)
หญิง	65.61	17.11	28.57	100.00	26.08	0.24 (0.38)	-0.12 (0.75)
$t(96) = 1.90, p = .06, d = 0.39$							
การจัดการสารสนเทศ							
รวม	54.25	17.17	0.00	100.00	31.65	-0.37 (0.24)	0.49 (0.48)
ชาย	56.51	16.25	16.67	100.00	28.76	-0.04 (0.31)	0.24 (0.61)
หญิง	50.68	18.19	0.00	85.71	35.89	-0.69 (0.38)	0.34 (0.75)
$t(96) = 1.65, p = .10, d = 0.34$							
การบูรณาการสารสนเทศ							
รวม	68.48	20.98	14.29	100.00	30.64	-0.50 (0.24)	-0.33 (0.48)
ชาย	67.72	21.81	14.29	100.00	32.21	-0.48 (0.31)	-0.39 (0.61)
หญิง	69.67	19.81	20.00	100.00	28.43	-0.52 (0.38)	-0.16 (0.75)
$t(96) = 0.45, p = .66, d = 0.09$							
การประเมินสารสนเทศ							
รวม	68.44	16.20	33.33	100.00	23.67	-0.09 (0.24)	-0.65 (0.48)
ชาย	68.78	17.31	33.33	100.00	25.17	-0.10 (0.31)	-0.76 (0.61)
หญิง	67.89	14.48	40.00	100.00	21.33	-0.13 (0.38)	-0.46 (0.75)
$t(96) = 0.26, p = .79, d = 0.06$							

การสื่อสารสารสนเทศ							
รวม	69.91	15.75	33.33	100.00	22.53	-0.37 (0.24)	-0.19 (0.48)
ชาย	68.26	15.54	33.33	100.00	22.77	-0.26 (0.31)	0.002 (0.61)
หญิง	72.51	15.94	33.33	100.00	21.98	-0.59 (0.38)	-0.11 (0.75)
$t(96) = 1.30, p = .20, d = 0.27$							



ภาพ 4.24 ค่าเฉลี่ยของคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แต่ละองค์ประกอบจำแนกตามเพศของนิสิตนักศึกษาสายวิทยาศาสตร์

ตอนที่ 5 ผลการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

5.1 ผลการตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ผู้วิจัยตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยใช้แบบประเมินระบบแบบอิงมาตรฐาน ซึ่งผู้วิจัยพัฒนาแบบประเมินระบบแบบอิงมาตรฐานก่อนนำไปใช้ตรวจสอบคุณภาพของระบบการทดสอบ จึงแบ่งการนำเสนอเป็น 2 ประเด็น ได้แก่ ผลการพัฒนาแบบประเมินระบบแบบอิงมาตรฐาน และผลการตรวจสอบคุณภาพของระบบการทดสอบ ดังรายละเอียดต่อไปนี้

5.1.1 ผลการพัฒนาแบบประเมินระบบแบบอิงมาตรฐาน

แบบประเมินระบบแบบอิงมาตรฐานมีลักษณะเป็นมาตรประมาณค่า 5 ระดับ ประกอบด้วยข้อคำถามจำนวน 14 ข้อ ผลการตรวจสอบความตรงเชิงเนื้อเรื่องของแบบประเมินระบบแบบอิงมาตรฐานพบว่าข้อคำถามทุกข้อผ่านเกณฑ์การประเมิน โดยมีค่าดัชนี IOC อยู่ระหว่าง 0.67 ถึง 1.00 ทั้งนี้ผู้ทรงคุณวุฒิด้านการศึกษา จำนวน 3 ท่าน ให้ข้อเสนอแนะในการปรับปรุงข้อคำถามดังรายละเอียดในตาราง 4.34

ตาราง 4.34 ผลการตรวจสอบความตรงเชิงเนื้อเรื่องของแบบประเมินระบบแบบอิงมาตรฐาน

ข้อคำถาม	IOC	ข้อเสนอแนะ
ด้านความมีประโยชน์ (utility)		
1.ระบบมีการจัดสอบที่มีความครอบคลุมและให้ผลการจัดสอบตรงตามความต้องการของผู้ใช้งาน	0.67	-คำถามมี 2 ประเด็น ควรแยกเป็น 2 ข้อ -น่าจะวัดด้านความเหมาะสม -คำถามควรเปลี่ยนเป็น การจัดสอบของระบบมีความครอบคลุมและให้ผลตรงตามความต้องการของผู้ใช้งาน
2.ระบบมีการสรุปผลการทดสอบที่มีความชัดเจนเข้าใจง่าย	1.00	คำถามควรเปลี่ยนเป็น การรายงานผลการทดสอบของระบบมีความชัดเจน เข้าใจง่าย
3.ผลการทดสอบสามารถนำไปใช้ประโยชน์ในการพัฒนาศักยภาพของผู้สอบได้	1.00	เพิ่มคำว่า “ได้” หลังคำถาม
ด้านความเป็นไปได้ในการนำไปใช้ (feasibility)		
4.ขั้นตอนของระบบมีการเรียกใช้งานง่าย สะดวกและรวดเร็ว	1.00	-
5.การแสดงผลของระบบมีความรวดเร็ว	1.00	-
6.ระบบมีสีสันดึงดูดความสนใจ	1.00	-
7.ระบบมีคู่มือการใช้งานที่ชัดเจน	1.00	-
ด้านความเหมาะสม (propriety)		
8.ระบบมีการแสดงขั้นตอนการทดสอบอย่างชัดเจน	1.00	คำถามควรเปลี่ยนเป็น การแสดงขั้นตอนการทดสอบของระบบมีความชัดเจน
9.ระบบมีการกำหนดเนื้อหาในการทดสอบอย่างเหมาะสม	1.00	คำถามควรเปลี่ยนเป็น การกำหนดเนื้อหาในการทดสอบของระบบมีความเหมาะสม
10.ระยะเวลาในการทดสอบมีความเหมาะสม	1.00	คำถามควรเปลี่ยนเป็น ระยะเวลาในการทดสอบของระบบมีความเหมาะสม
11.ผลการทดสอบครอบคลุมและครบถ้วน	1.00	คำถามควรเปลี่ยนเป็น ผลการทดสอบมีความครอบคลุมและครบถ้วน

ข้อคำถาม	IOC	ข้อเสนอแนะ
ด้านความถูกต้อง (accuracy)		
12.ระบบสามารถจัดการทดสอบได้ตรงตามวัตถุประสงค์	1.00	-
13.ระบบสามารถคำนวณค่าความสามารถของผู้เข้าสอบได้ถูกต้อง	1.00	ควรเปลี่ยนคำว่า "คำนวณ" เป็นคำว่าประมาณค่า
14.ระบบสามารถสรุปผลการทดสอบและบันทึกผลการสอบได้ถูกต้อง	1.00	-

ผู้วิจัยปรับปรุงแก้ไขข้อคำถามตามข้อเสนอแนะของผู้ทรงคุณวุฒิ จากนั้นจึงจัดพิมพ์แบบประเมินระบบแบบอิงมาตรฐานเพื่อให้ผู้ทรงคุณวุฒิด้านเทคโนโลยีสารสนเทศและการสื่อสารและผู้ทรงคุณวุฒิด้านการศึกษาประเมินคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยแบบประเมินระบบแบบอิงมาตรฐานมีลักษณะดังภาพ 4.25

แบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation)

คำชี้แจง

ให้ท่านพิจารณาข้อความต่อไปนี้ แล้วทำเครื่องหมาย ✓ ลงในช่องที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด โดยมีระดับความคิดเห็น 5 ระดับ ดังนี้

1 หมายถึง น้อยที่สุด
2 หมายถึง น้อย
3 หมายถึง ปานกลาง
4 หมายถึง มาก
5 หมายถึง มากที่สุด

คำถาม	ระดับความคิดเห็น				
	1 น้อยที่สุด	2 น้อย	3 ปานกลาง	4 มาก	5 มากที่สุด
ด้านความมีประโยชน์ (utility)					
1.การจัดสอบของระบบมีความครอบคลุมตามความต้องการของผู้ใช้งาน					
2.การรายงานผลการทดสอบของระบบมีความชัดเจน เข้าใจง่าย					
3.ผลการทดสอบสามารถนำไปใช้ประโยชน์ในการพัฒนาศักยภาพของผู้สอบได้					
ด้านความเป็นไปได้ในการนำไปใช้ (feasibility)					
4.ขั้นตอนของระบบมีการเรียกใช้งานง่าย สะดวกและรวดเร็ว					
5.การแสดงผลของระบบมีความรวดเร็ว					
6.ระบบมีสีสันดึงดูดความสนใจ					
7.ระบบมีคู่มือการใช้งานที่ชัดเจน					
ด้านความเหมาะสม (propriety)					
8.การแสดงขั้นตอนการทดสอบของระบบมีความชัดเจน					
9.การกำหนดเนื้อหาในการทดสอบของระบบมีความเหมาะสม					
10.ระยะเวลาในการทดสอบของระบบมีความเหมาะสม					
11.ผลการทดสอบมีความครอบคลุมและครบถ้วน					
ด้านความถูกต้อง (accuracy)					
12.ระบบสามารถจัดการทดสอบได้ตรงตามวัตถุประสงค์					
13.ระบบสามารถประมาณค่าความสามารถของผู้เข้าสอบได้ถูกต้อง					
14.ระบบสามารถสรุปผลการทดสอบและบันทึกผลการสอบได้ถูกต้อง					

ภาพ 4.25 แบบประเมินระบบแบบอิงมาตรฐาน

5.1.2 ผลการตรวจสอบคุณภาพของระบบการทดสอบ

ผลการตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี พบว่า โดยภาพรวมระบบการทดสอบมีคุณภาพอยู่ในระดับมากที่สุด (*M* = 4.95, *SD* = 0.08) หากพิจารณาคุณภาพของระบบการทดสอบในแต่ละด้าน พบว่า คุณภาพด้านความมีประโยชน์ (*M* = 4.89, *SD* = 0.19) ด้านความเป็นไปได้ใน

การนำไปใช้ (M = 5.00, SD = 0.00) ด้านความเหมาะสม (M = 4.92, SD = 0.14) และด้านความถูกต้อง (M = 5.00, SD = 0.00) มีคุณภาพอยู่ในระดับมากที่สุดทั้ง 4 ด้าน โดยมีผลการวิเคราะห์ดังตาราง 4.35

ตาราง 4.35 ผลการตรวจสอบคุณภาพของระบบการทดสอบ

รายการ	น้อยที่สุด n (%)	น้อย n (%)	ปานกลาง n (%)	มาก n (%)	มากที่สุด n (%)
ด้านความมีประโยชน์ (M = 4.89, SD = 0.19)					
1.การจัดสอบของระบบมีความครอบคลุมตามความต้องการของผู้ใช้งาน	0 (0.00)	0 (0.00)	0 (0.00)	1 (33.33)	2 (66.67)
2.การรายงานผลการทดสอบของระบบมีความชัดเจน เข้าใจง่าย	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
3.ผลการทดสอบสามารถนำไปใช้ประโยชน์ในการพัฒนาศักยภาพของผู้สอบได้	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
ด้านความเป็นไปได้ในการนำไปใช้ (M = 5.00, SD = 0.00)					
4.ขั้นตอนของระบบมีการเรียกใช้งานง่ายสะดวกและรวดเร็ว	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
5.การแสดงผลของระบบมีความรวดเร็ว	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
6.ระบบมีสีสันดึงดูดความสนใจ	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
7.ระบบมีคู่มือการใช้งานที่ชัดเจน	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
ด้านความเหมาะสม (M = 4.92, SD = 0.14)					
8.การแสดงขั้นตอนการทดสอบของระบบมีความชัดเจน	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
9.การกำหนดเนื้อหาในการทดสอบของระบบมีความเหมาะสม	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
10.ระยะเวลาในการทดสอบของระบบมีความเหมาะสม	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
11.ผลการทดสอบมีความครอบคลุมและครบถ้วน	0 (0.00)	0 (0.00)	0 (0.00)	1 (33.33)	2 (66.67)
ด้านความถูกต้อง (M = 5.00, SD = 0.00)					
12.ระบบสามารถจัดการทดสอบได้ตรงตามวัตถุประสงค์	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
13.ระบบสามารถประมาณค่าความสามารถของผู้เข้าสอบได้ถูกต้อง	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
14.ระบบสามารถสรุปผลการทดสอบและบันทึกผลการสอบได้ถูกต้อง	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (100.00)
รวม (M = 4.95, SD = 0.08)	0 (0.00)	0 (0.00)	0 (0.00)	2 (4.76)	40 (95.24)

5.2 ผลการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ผู้วิจัยศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีโดยใช้แบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ ซึ่งผู้วิจัยพัฒนาแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ก่อนการนำไปศึกษาความพึงพอใจกับนิสิตนักศึกษาสายสังคมศาสตร์และสายวิทยาศาสตร์ จึงแบ่งการนำเสนอเป็น 2 ประเด็น ได้แก่ ผลการพัฒนาแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ และผลการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบ ดังรายละเอียดต่อไปนี้

5.2.1 ผลการพัฒนาแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์

แบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์มีลักษณะเป็นมาตรจำแนกความหมาย 7 ระดับ ประกอบด้วยข้อคำถามจำนวน 20 ข้อ ผลการตรวจสอบความตรงเชิงเนื้อเรื่องของแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์พบว่าข้อคำถามทุกข้อผ่านเกณฑ์การประเมิน โดยมีค่าดัชนี IOC อยู่ระหว่าง 0.67 ถึง 1.00 ทั้งนี้ผู้ทรงคุณวุฒิด้านการศึกษา จำนวน 3 ท่าน ให้ข้อเสนอแนะในการปรับปรุงข้อคำถามดังรายละเอียดในตาราง 4.36

ตาราง 4.36 ผลการตรวจสอบความตรงเชิงเนื้อเรื่องของแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์

ข้อคำถาม	คำคุณศัพท์		IOC	ข้อเสนอแนะ
ภาพรวมของระบบ				
1.ปฏิสัมพันธ์ในภาพรวมต่อระบบ (overall reaction to the software)	แย่มาก	ดีเยี่ยม	1.00	-
2.ความคาดหวังในภาพรวมต่อระบบ	ผิดหวัง	พอใจ	1.00	-คำถามควรระบุเกี่ยวกับสิ่งที่คาดหวัง -ควรใช้คำว่า พึงพอใจ แทนคำว่า พอใจ
3.ความน่าสนใจในภาพรวมต่อระบบ	น่าเบื่อ	น่าสนใจ	1.00	-
หน้าจอของระบบ (screen)				
4.รูปแบบของข้อความและส่วนประกอบต่าง ๆ บนหน้าจอ	ยากต่อการอ่าน	ง่ายต่อการอ่าน	1.00	-
5.การใช้งานของเครื่องมือ (tools) ต่าง ๆ บนหน้าจอ	ยาก	ง่าย	1.00	-
6.การจัดการสารสนเทศบนหน้าจอ	สับสน	ชัดเจน	1.00	-

ข้อคำถาม	คำคุณศัพท์		IOC	ข้อเสนอแนะ
7.ลำดับการแสดงภาพและส่วนประกอบต่าง ๆ บนหน้าจอ	สับสน	ชัดเจน	1.00	-
คำศัพท์และสารสนเทศระบบ (terminology and system information)				
8.การใช้คำศัพท์เฉพาะในส่วนต่าง ๆ ของระบบ	ผิดพลาด	ถูกต้อง	1.00	
9.คำศัพท์ที่ใช้มีความสัมพันธ์กับเครื่องมือ (tools)	ขัดแย้ง	สัมพันธ์	1.00	-ควรใช้คำว่า ขัดแย้งกัน แทนคำว่า ขัดแย้ง -ควรใช้คำว่า สัมพันธ์กัน แทนคำว่า สัมพันธ์
10.ตำแหน่งของข้อความบนหน้าจอเพื่อป้อนข้อมูล	ปรับปรุง	เหมาะสม	1.00	ควรใช้คำว่า ควรปรับปรุง แทนคำว่า ปรับปรุง
11.การแจ้งเตือนสถานะว่ากำลังดำเนินการอะไรอยู่	ปรับปรุง	เหมาะสม	1.00	-คำถามควรเปลี่ยนเป็น การแจ้งเตือนสถานะขณะกำลังดำเนินการ -ควรใช้คำว่า ควรปรับปรุง แทนคำว่า ปรับปรุง
12.การมีข้อความแจ้งเตือนความผิดพลาด ทำให้สามารถใช้งานระบบได้ง่ายขึ้น	ปรับปรุง	เหมาะสม	1.00	ควรใช้คำว่า ควรปรับปรุง แทนคำว่า ปรับปรุง
การเรียนรู้ของผู้ใช้งาน (learning)				
13.ผู้ใช้งานสามารถเรียนรู้การใช้งานและการทำงานของระบบด้วยตนเอง	ยาก	ง่าย	1.00	-
14.การแสดงเค้าโครงใหม่ (new features) ด้วยตัวอย่างและข้อผิดพลาดทำให้ผู้ใช้งานเข้าใจการทำงานของระบบได้ดีขึ้น	สับสน	ชัดเจน	0.67	-
15.ผู้ใช้งานสามารถจำชื่อและการใช้งานคำสั่งต่าง ๆ ของระบบได้	ยาก	ง่าย	1.00	-
16.ข้อความช่วยเหลือบนหน้าจอทำให้ผู้ใช้งานเข้าใจการทำงานของระบบได้ดียิ่งขึ้น	สับสน	ชัดเจน	1.00	-
17.คำชี้แจงในการใช้งานระบบทำให้	สับสน	ชัดเจน	1.00	-

ข้อคำถาม	คำคุณศัพท์		IOC	ข้อเสนอแนะ
ผู้ใช้งานเข้าใจระบบได้ดีขึ้น				
ความสามารถของระบบ (system capabilities)				
18.ความเร็วระบบ (speed)	ช้า	เร็ว	1.00	-
19.ความคงที่ (reliability) ในการทำงานของระบบ	บกพร่อง	น่าเชื่อถือ	1.00	-คำถามควรเปลี่ยนเป็น "ความน่าเชื่อถือในการทำงานของระบบ" เพื่อให้เข้าใจง่าย และสอดคล้องกับคำคุณศัพท์ -ไม่แน่ใจเกี่ยวกับคำคุณศัพท์ทั้ง 2 คำ
20.ความถูกต้องในการทำงานของระบบ	ผิดพลาด	แม่นยำ	1.00	-

ผู้วิจัยปรับปรุงแก้ไขข้อคำถามตามข้อเสนอแนะของผู้ทรงคุณวุฒิ จากนั้นจึงสร้างแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ไว้ในระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ซึ่งให้นิสิตนักศึกษาประเมินความพึงพอใจต่อระบบการทดสอบหลังจากเสร็จสิ้นการทำแบบสอบ โดยแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ในระบบการทดสอบแสดงดังภาพ 4.26

แบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์
(user satisfaction of the human-computer interface)

คำชี้แจง: โปรดคลิกให้คะแนนที่ ◉ ตามความคิดเห็นของท่าน ดังนี้
***มาตรที่ใช้ในการวัด: 7 ระดับ 1 2 3 4 5 6 7

ข้อ				
ภาพรวมของระบบ				
1. ปฏิสัมพันธ์ในภาพรวมต่อระบบ (overall reaction to the software)	แย่มาก	1 2 3 4 5 6 7	ดีเยี่ยม
2. ความคาดหวังต่อระบบในภาพรวม	ผิดหวัง	1 2 3 4 5 6 7	พึงพอใจ
3. ความน่าสนใจในภาพรวมต่อระบบ	น่าเบื่อ	1 2 3 4 5 6 7	น่าสนใจ
หน้าจอของระบบ (screen)			
4. รูปแบบของข้อความและส่วนประกอบต่าง ๆ บนหน้าจอ	ยากต่อการอ่าน	1 2 3 4 5 6 7	ง่ายต่อการอ่าน
5. การใช้งานของเครื่องมือ (tools) ต่าง ๆ บนหน้าจอ	ยาก	1 2 3 4 5 6 7	ง่าย
6. การจัดการสารสนเทศบนหน้าจอ	สับสน	1 2 3 4 5 6 7	ชัดเจน
7. ลำดับการแสดงภาพและส่วนประกอบต่าง ๆ บนหน้าจอ	สับสน	1 2 3 4 5 6 7	ชัดเจน
คำศัพท์และสารสนเทศระบบ (terminology and system information)			
8. การใช้คำศัพท์เฉพาะในส่วนต่าง ๆ ของระบบ	ผิดพลาด	1 2 3 4 5 6 7	ถูกต้อง
9. คำศัพท์ที่ใช้มีความสัมพันธ์กับเครื่องมือ (tools)	ขัดแย้งกัน	1 2 3 4 5 6 7	สัมพันธ์กัน
10. ตำแหน่งของข้อความบนหน้าจอเพื่อป้อนข้อมูล	ควรปรับปรุง	1 2 3 4 5 6 7	เหมาะสม
11. การแจ้งเตือนสถานะขณะกำลังดำเนินการ	ควรปรับปรุง	1 2 3 4 5 6 7	เหมาะสม
12. การมีข้อความแจ้งเตือนความผิดพลาด ทำให้สามารถใช้งานระบบได้ง่ายขึ้น	ควรปรับปรุง	1 2 3 4 5 6 7	เหมาะสม
การเรียนรู้ของผู้ใช้งาน (learning)			
13. ผู้ใช้งานสามารถเรียนรู้การใช้งานและการทำงานของระบบด้วยตนเอง	ยาก	1 2 3 4 5 6 7	ง่าย
14. การแสดงเค้าโครงใหม่ (new features) ด้วยตัวอย่างและข้อผิดพลาด ทำให้ผู้ใช้งานเข้าใจการทำงานของระบบได้ดีขึ้น	สับสน	1 2 3 4 5 6 7	ชัดเจน
15. ผู้ใช้งานสามารถจำชื่อและการใช้งานคำสั่งต่าง ๆ ของระบบได้	ยาก	1 2 3 4 5 6 7	ง่าย
16. ข้อความช่วยเหลือบนหน้าจอทำให้ผู้ใช้งานเข้าใจการทำงานของระบบได้ดียิ่งขึ้น	สับสน	1 2 3 4 5 6 7	ชัดเจน
17. คำชี้แจงในการใช้งานระบบทำให้ผู้ใช้งานเข้าใจระบบได้ดีขึ้น	สับสน	1 2 3 4 5 6 7	ชัดเจน
ความสามารถระบบ (system capabilities)			
18. ความเร็วระบบ (speed)	ช้า	1 2 3 4 5 6 7	เร็ว
19. ความคงเส้นคงวา (reliability) ในการทำงานของระบบ	ไม่น่าเชื่อถือ	1 2 3 4 5 6 7	น่าเชื่อถือ
20. ความถูกต้องในการทำงานของระบบ	ผิดพลาด	1 2 3 4 5 6 7	แม่นยำ

ภาพ 4.26 แบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์

5.2.2 ผลการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบ

ผู้วิจัยศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีกับนิสิตนักศึกษาสายสังคมศาสตร์ และสายวิทยาศาสตร์ จำนวน 217 คน แบ่งเป็นชาย จำนวน 104 คน (ร้อยละ 47.93) และหญิง จำนวน 113 คน (ร้อยละ 52.07) พบว่า โดยภาพรวม นิสิตนักศึกษามีความพึงพอใจ

ระดับมากต่อระบบการทดสอบ (*M* = 5.63, *SD* = 0.76) ซึ่งมีผลการวิเคราะห์ดังตาราง 4.37 นอกจากนี้ยังพบว่า นิสิตนักศึกษามีความพึงพอใจระดับมากต่อระบบการทดสอบด้านภาพรวมของระบบ (*M* = 5.66, *SD* = 0.85) หน้าจอของระบบ (*M* = 5.72, *SD* = 0.82) การเรียนรู้ของผู้ใช้งาน (*M* = 5.65, *SD* = 0.85) และความสามารถของระบบ (*M* = 5.72, *SD* = 0.89) และค่อนข้างพึงพอใจต่อระบบการทดสอบด้านคำศัพท์และสารสนเทศระบบ (*M* = 5.47, *SD* = 0.94) ผู้วิจัยแปลผลโดยใช้ค่าเฉลี่ยของระดับคะแนนความพึงพอใจเทียบกับเกณฑ์การแปลความหมาย ซึ่งมีเกณฑ์การแปลความหมายความพึงพอใจ ดังนี้ (บุญชม ศรีสะอาด, 2539)

1.00 - 1.50 หมายถึง ไม่พึงพอใจมากที่สุด
1.51 - 2.50 หมายถึง ไม่พึงพอใจมาก
2.51 - 3.50 หมายถึง ค่อนข้างไม่พึงพอใจ
3.51 - 4.50 หมายถึง เฉยๆ
4.51 - 5.50 หมายถึง ค่อนข้างพึงพอใจ
5.51 - 6.50 หมายถึง พึงพอใจมาก
6.51 - 7.00 หมายถึง พึงพอใจมากที่สุด

ตาราง 4.37 ผลการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบ

รายการ	1 *n* (%)	2 *n* (%)	3 *n* (%)	4 *n* (%)	5 *n* (%)	6 *n* (%)	7 *n* (%)
ภาพรวมของระบบ (*M* = 5.66, *SD* = 0.85)							
1.ปฏิสัมพันธ์ในภาพรวมต่อระบบ	0 (0.00)	1 (0.46)	9 (4.15)	20 (9.22)	53 (24.42)	83 (38.25)	51 (23.50)
2.ความคาดหวังต่อระบบในภาพรวม	1 (0.46)	1 (0.46)	4 (1.84)	24 (11.06)	62 (28.57)	68 (31.34)	57 (26.27)
3.ความน่าสนใจในภาพรวมต่อระบบ	0 (0.00)	3 (1.38)	5 (2.30)	21 (9.68)	63 (29.03)	69 (31.80)	56 (25.81)
หน้าจอของระบบ (*M* = 5.72, *SD* = 0.82)							
4.รูปแบบของข้อความและส่วนประกอบต่าง ๆ บนหน้าจอ	0 (0.00)	1 (0.46)	3 (1.38)	18 (8.29)	63 (29.03)	69 (31.80)	63 (29.03)
5.การใช้งานของเครื่องมือต่าง ๆ บนหน้าจอ	0 (0.00)	1 (0.46)	3 (1.38)	30 (13.82)	63 (29.03)	66 (30.41)	54 (24.88)
6.การจัดการสารสนเทศบนหน้าจอ	0 (0.00)	0 (0.00)	3 (1.38)	28 (12.90)	58 (26.73)	68 (31.34)	60 (27.65)
7.ลำดับการแสดงภาพและส่วนประกอบต่าง ๆ บนหน้าจอ	0 (0.00)	0 (0.00)	3 (1.38)	21 (9.68)	54 (24.88)	83 (38.25)	56 (25.81)
คำศัพท์และสารสนเทศระบบ (*M* = 5.47, *SD* = 0.94)							
8.การใช้คำศัพท์เฉพาะในส่วนต่าง ๆ ของระบบ	1 (0.46)	4 (1.84)	15 (6.91)	28 (12.90)	42 (19.35)	75 (34.56)	52 (23.96)
9.คำศัพท์ที่ใช้มีความสัมพันธ์กับเครื่องมือ	0 (0.00)	2 (0.92)	10 (4.61)	23 (10.60)	62 (28.57)	68 (31.34)	52 (23.96)
10.ตำแหน่งของข้อความบนหน้าจอเพื่อป้อนข้อมูล	0 (0.00)	1 (0.46)	2 (0.92)	15 (6.91)	42 (19.35)	88 (40.55)	69 (31.80)
11.การแจ้งเตือนสถานะขณะกำลังดำเนินการ	1 (0.46)	4 (1.84)	10 (4.61)	30 (13.82)	60 (27.65)	75 (34.56)	37 (17.05)
12.การมีข้อความแจ้งเตือนความผิดพลาด ทำให้สามารถใช้งานระบบได้ง่ายขึ้น	7 (3.23)	15 (6.91)	24 (11.06)	31 (14.29)	32 (14.75)	57 (26.27)	51 (23.50)

รายการ	1 *n* (%)	2 *n* (%)	3 *n* (%)	4 *n* (%)	5 *n* (%)	6 *n* (%)	7 *n* (%)
การเรียนรู้ของผู้ใช้งาน (*M* = 5.65, *SD* = 0.85)							
13.ผู้ใช้งานสามารถเรียนรู้การใช้งานและการทำงานของระบบด้วยตนเอง	0 (0.00)	0 (0.00)	4 (1.84)	13 (5.99)	51 (23.50)	84 (38.71)	65 (29.95)
14.การแสดงเค้าโครงใหม่ด้วยตัวอย่างและข้อผิดพลาด ทำให้ผู้ใช้งานเข้าใจการทำงานของระบบได้ดีขึ้น	0 (0.00)	1 (0.46)	2 (0.92)	26 (11.98)	56 (25.81)	81 (37.33)	51 (23.50)
15.ผู้ใช้งานสามารถจำชื่อและการใช้งานคำสั่งต่าง ๆ ของระบบได้	0 (0.00)	2 (0.92)	7 (3.23)	23 (10.60)	66 (30.41)	68 (31.34)	51 (23.50)
16.ข้อความช่วยเหลือบนหน้าจอทำให้ผู้ใช้งานเข้าใจการทำงานของระบบได้ดียิ่งขึ้น	1 (0.46)	1 (0.46)	3 (1.38)	25 (11.52)	64 (29.49)	85 (39.17)	38 (17.51)
17.คำชี้แจงในการใช้งานระบบทำให้ผู้ใช้งานเข้าใจระบบได้ดีขึ้น	1 (0.46)	5 (2.30)	9 (4.15)	24 (11.06)	60 (27.65)	63 (29.03)	55 (25.35)
ความสามารถของระบบ (*M* =5.72 , *SD* = 0.89)							
18.ความเร็วระบบ	0 (0.00)	2 (0.92)	8 (3.69)	12 (5.53)	41 (18.89)	67 (30.88)	87 (40.09)
19.ความคงเส้นคงวาในการทำงานของระบบ	0 (0.00)	2 (0.92)	10 (4.61)	23 (10.60)	62 (28.57)	68 (31.34)	52 (23.96)
20.ความถูกต้องในการทำงานของระบบ	0 (0.00)	2 (0.92)	8 (3.69)	19 (8.76)	66 (30.41)	64 (29.49)	58 (26.73)
รวม (*M* = 5.63, *SD* = 0.76)	12 (0.28)	48 (1.11)	142 (3.27)	454 (10.46)	1120 (25.81)	1449 (33.39)	1115 (25.69)

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยเรื่อง"ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี" มีวัตถุประสงค์การวิจัย 3 ข้อ ได้แก่ (1) เพื่อพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี (2) เพื่อพัฒนาข้อสอบสำหรับจัดทำคลังข้อสอบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี และ (3) เพื่อพัฒนาและตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยดำเนินการวิจัยเป็น 5 ระยะ ดังรายละเอียดต่อไปนี้

ระยะที่ 1 การพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ผู้วิจัยสังเคราะห์และนิยามองค์ประกอบการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี จากเอกสารและงานวิจัยที่เกี่ยวข้องทั้งในประเทศไทยและต่างประเทศ โดยใช้การวิเคราะห์เนื้อหา ความถี่ และร้อยละ

ระยะที่ 2 การพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ประชากรวิจัย คือ นิสิตนักศึกษาระดับปริญญาตรีจากสถาบันอุดมศึกษาของรัฐบาลทั่วประเทศ จำนวนทั้งสิ้น 1,422,254 คน ตัวอย่างวิจัย ได้แก่ นิสิตนักศึกษาระดับปริญญาตรีจากสถาบันอุดมศึกษาของรัฐบาลในกรุงเทพมหานคร จำนวนทั้งสิ้น 1,672 คน ได้มาจากสุ่มเลือกสาขาวิชา จากนั้นจึงเลือกสถาบันอุดมศึกษาและนิสิตนักศึกษาในแต่ละสาขาวิชา โดยกำหนดจำนวนตัวอย่างวิจัยในแต่ละสาขาวิชาตามสัดส่วนจริงของประชากรวิจัย

การดำเนินการวิจัยในระยะที่ 2 มี 5 ขั้นตอน ได้แก่ (1) ดำเนินการจัดประชุมเพื่อวางแผนสำหรับการสร้างแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีกับกรรมการผู้ร่วมสร้างแบบวัด (2) ดำเนินการรวบรวมข้อสอบจากกรรมการเพื่อดำเนินการวิพากษ์ปรับปรุงแก้ไขข้อสอบและจัดชุดเบื้องต้นสำหรับส่งให้ผู้ทรงคุณวุฒิตรวจสอบความสอดคล้องของข้อคำถามและองค์ประกอบในการวัด (3) ปรับปรุง แก้ไขข้อสอบตามข้อเสนอแนะของผู้ทรงคุณวุฒิ (4) ตรวจสอบคุณภาพของข้อสอบที่ได้จากการทดลองใช้ และ (5) คัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบ เครื่องมือวิจัย มี 2 ประเภท ได้แก่ แบบประเมินความสอดคล้องระหว่างข้อคำถามกับองค์ประกอบในการวัดและแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

การวิเคราะห์ข้อมูลในระยะที่ 2 แบ่งเป็น 5 ส่วน ได้แก่ (1) การวิเคราะห์ค่าสถิติบรรยาย ได้แก่ ค่าเฉลี่ย ค่าสูงสุด ค่าต่ำสุด ส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบ (2) การวิเคราะห์ค่าความตรงตามเนื้อหา

โดยใช้ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมายของการวัด (IOC) (3) การวิเคราะห์ค่าความเที่ยงของแบบวัดโดยใช้สูตร KR20 (4) การวิเคราะห์คุณภาพของข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิม ได้แก่ ความยาก และอำนาจจำแนก และ (5) การวิเคราะห์คุณภาพของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ ได้แก่ การปรับเทียบข้อสอบ การเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลการตอบสนองข้อสอบแบบ 1PL, 2PL และ 3PL และการเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลการตอบสนองข้อสอบเอกมิติ และโมเดลการตอบสนองข้อสอบพหุมิติ

ระยะที่ 3 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ผู้วิจัยดำเนินการออกแบบโปรแกรมคอมพิวเตอร์หลังจากที่ได้ข้อสอบที่มีคุณภาพเข้าคลังข้อสอบ โดยแบ่งการดำเนินงานออกเป็น 4 ส่วน ได้แก่ (1) การลงทะเบียน (2) การสร้างชุดข้อสอบ (3) การทดสอบ และ (4) การรายงานผลการทดสอบ ซึ่งพัฒนาระบบโดยใช้ภาษา PHP เนื่องจากเป็นระบบการทดสอบที่สามารถใช้งานผ่านระบบออนไลน์

ระยะที่ 4 การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้ ตัวอย่างวิจัย คือ นิสิตระดับปริญญาตรีสายสังคมศาสตร์ และสายวิทยาศาสตร์ ชั้นปีที่ 3 จำนวน 217 คน เครื่องมือวิจัย คือ ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ในการเก็บรวบรวมข้อมูล ตัวอย่างวิจัยจะต้องลงทะเบียนการทดสอบเพื่อกรอกรายละเอียดเกี่ยวกับผู้เข้ารับการทดสอบก่อนการทดสอบ โดยมีเวลาในการทดสอบทั้งสิ้น 1 ชั่วโมง เมื่อตัวอย่างวิจัยทำแบบสอบเสร็จแล้ว โปรแกรมการทดสอบจะเสนอแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์ให้ตัวอย่างวิจัยประเมินความพึงพอใจต่อระบบการทดสอบ จากนั้นจึงรายงานผลการทดสอบ วิเคราะห์ข้อมูลโดยใช้สถิติเชิงบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ยเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด สัมประสิทธิ์การกระจาย ความเบ้ และความโด่ง รวมทั้งใช้สถิติทดสอบทีสำหรับตัวอย่าง 2 กลุ่ม ที่เป็นอิสระต่อกัน (t-test for independent samples)

ระยะที่ 5 การตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ผู้ให้ข้อมูลแบ่งเป็น 2 กลุ่ม ได้แก่ ผู้ทรงคุณวุฒิให้ข้อมูลสำหรับการตรวจสอบคุณภาพของระบบการทดสอบ และนิสิตระดับปริญญาตรีสำหรับการศึกษาความพึงพอใจของผู้ใช้ที่มีต่อระบบการทดสอบซึ่งเป็นกลุ่มเดียวกับตัวอย่างวิจัยที่ทดลองใช้ระบบการทดสอบ เครื่องมือวิจัย คือ แบบประเมินระบบแบบอิงมาตรฐาน มีลักษณะเป็นมาตรประมาณค่า 5 ระดับ และแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์มีลักษณะเป็นมาตรจำแนกความหมาย 7 ระดับ เก็บรวบรวมข้อมูลโดยให้ผู้ทรงคุณวุฒิประเมินคุณภาพของระบบการทดสอบโดยใช้แบบประเมินระบบแบบอิงมาตรฐาน และให้นิสิตที่ทดลองใช้ระบบการทดสอบ

ประเมินความพึงพอใจต่อระบบการทดสอบโดยใช้แบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟสระหว่างผู้ใช้งานกับคอมพิวเตอร์วิเคราะห์ข้อมูลโดยการคำนวณค่าดัชนีความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมายของการวัด (IOC) และสถิติเชิงบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ยเลขคณิต และส่วนเบี่ยงเบนมาตรฐาน

สรุปผลการวิจัย

ผลการวิจัยสรุปเป็น 5 ประเด็น ตามระยะการดำเนินการวิจัย ได้แก่ (1) การพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี 2) การพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี (3) การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี (4) การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้ และ (5) การตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยการสรุปผลการวิจัย การอภิปรายผลการวิจัย และข้อเสนอแนะ แต่ละหัวข้อมีรายละเอียดดังนี้

1. สรุปผลการพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ผลการวิเคราะห์องค์ประกอบและนิยามขององค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีที่ใช้ในการวิจัยครั้งนี้ สรุปได้เป็น 5 องค์ประกอบดังนี้

1) การเข้าถึงสารสนเทศ (information accessibility) หมายถึง ความรู้และความสามารถในการสืบค้นและรวบรวมสารสนเทศรวมทั้งอธิบายหรือแสดงวิธีการเรียกสารสนเทศที่มีอยู่กลับคืนมาเพื่อการใช้งานในระบบดิจิทัลได้

2) การจัดการสารสนเทศ (information management) หมายถึง ความรู้และความสามารถในการจัดระบบหมวดหมู่ของสารสนเทศในระบบดิจิทัลให้ง่ายต่อการการเรียกค้นและการประยุกต์ใช้ให้เหมาะสมกับงานแต่ละประเภทได้

3) การบูรณาการสารสนเทศ (information integration) หมายถึง ความรู้และความสามารถในการแปลความหมาย สรุป เปรียบเทียบ จำแนกความเหมือนและความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆในระบบดิจิทัลได้

4) การประเมินสารสนเทศ (information evaluation) หมายถึง ความรู้และความสามารถในการตัดสินคุณภาพ ความสอดคล้อง ประโยชน์ หรือประสิทธิภาพของสารสนเทศที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร ในระบบดิจิทัลได้

5) การสื่อสารสารสนเทศ (information communication) หมายถึง ความรู้และความสามารถในการจัดทำและเผยแพร่สารสนเทศในระบบดิจิทัลเพื่อตอบสนองความต้องการของกลุ่มบุคคลและบริบทต่างๆที่เป็นการเคารพสิทธิ์ความเป็นส่วนตัวได้

2. สรุปผลการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

2.1 ผลการสร้างข้อสอบทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ตามนิยามเชิงปฏิบัติการที่ได้วิเคราะห์ไว้ในระยะที่ 1 โดยทั้ง 5 องค์ประกอบประกอบด้วยจำนวนข้อสอบเท่ากันคิดเป็นองค์ประกอบละ 52 ข้อ รวมทั้งสิ้น 260 ข้อ ซึ่งข้อสอบทุกข้อเป็นข้อสอบแบบเลือกตอบหลายตัวเลือก (Multiple-choice) จำนวน 5 ตัวเลือกและมีเพียง 1 คำตอบถูก มีการเกณฑ์การให้คะแนนแบบ 0,1 (Dichotomous scoring) โดยที่ข้อสอบในแต่ละองค์ประกอบมีการออกแบบเพื่อให้วัดพฤติกรรมการเรียนรู้ในระดับความรู้ ความเข้าใจ การนำไปใช้ และการวิเคราะห์ ตามทฤษฎีการเรียนรู้ของบลูม

2.2 ผลการตรวจสอบความตรงเชิงเนื้อหาก่อนนำไปทดลองใช้ ได้กำหนดให้มีการประเมินความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญจำนวนทั้งสิ้น 15 ท่าน โดยแบบทดสอบหนึ่งชุดมีผู้เชี่ยวชาญตรวจสอบความตรงจำนวน 3 ท่าน ประกอบด้วย (1) ผู้เชี่ยวชาญด้านการวัดและประเมิน และ (2) ด้านเทคโนโลยีสารสนเทศ โดยผลการตรวจสอบความตรงในภาพรวม พบว่า ข้อสอบส่วนใหญ่ผ่านการประเมินความตรงเชิงเนื้อหา คิดเป็นร้อยละ 79.61 แต่มีข้อสอบบางข้อมีข้อเสนอแนะให้ปรับปรุง เช่น การปรับปรุงด้านการใช้ภาษาเพื่อให้เข้าใจง่าย การปรับปรุงข้อคำถามเพื่อให้สอดคล้องตามนิยามการวัด การปรับปรุงตัวเลือกถูกและตัวลวงให้มีความสอดคล้องกัน การปรับปรุงรูปภาพแผนภาพให้ชัดเจน เป็นต้น

2.3 ผลการตรวจสอบคุณภาพของข้อสอบที่ได้จากการทดลองใช้ ในการวิจัยครั้งนี้ตรวจสอบคุณภาพของข้อสอบเพื่อนำไปคัดเลือกข้อสอบเพื่อจัดเก็บเข้าคลังข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory, IRT) แต่มีการวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory, CTT) ประกอบการพิจารณาด้วย สรุปผลการวิเคราะห์คุณภาพรายข้อ ดังนี้

1) ผลการวิเคราะห์คุณภาพของข้อสอบและแบบสอบทั้งฉบับ ตามทฤษฎี CTT ได้แบ่งแบบสอบเป็น 10 ชุด แต่ละชุดมีจำนวนข้อเท่ากัน คือ 35 ข้อ ในภาพรวมพบว่า ค่าเฉลี่ยความยาก (p) ของแบบสอบทั้ง 10 ชุดเท่ากับ 0.483 แสดงว่ามีความยากอยู่ในระดับปานกลาง ขณะที่ค่าอำนาจจำแนกเฉลี่ย (r) ของแบบสอบทั้ง 10 ชุดมีค่าเท่ากับ 0.273 แสดงว่าจำแนกผู้สอบพอใช้ได้ และค่าความเที่ยงเฉลี่ย (R) ของแบบสอบทั้ง 10 ชุดมีค่าเท่ากับ 0.725 แสดงว่ามีความเที่ยงในระดับค่อนข้างสูง

2) ผลการวิเคราะห์คุณภาพของข้อสอบรายข้อตามทฤษฎี IRT เพื่อให้ได้สารสนเทศที่มีความถูกต้อง ผู้วิจัยได้ตรวจสอบและเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลการตอบสนองข้อสอบ ทั้งแบบ 1 PL 2 PL และ 3PL เพื่อเลือกโมเดลที่มีความถูกต้องมากที่สุดแล้วนำไปวิเคราะห์คุณภาพข้อสอบที่มีการปรับเทียบแบบสอบและการประเมินผลคุณภาพข้อสอบ สรุปว่า ผลการเปรียบเทียบค่าดัชนี deviance (G^2) ค่า AIC และ BIC ไม่เป็นไปในทิศทางเดียวกัน จึงทดสอบนัยสำคัญของความสอดคล้องกลมกลืนของโมเดลทีละคู่ ด้วยการทดสอบอัตราส่วนความเป็นไปได้ (likelihood-ratio test: LRT) พบว่า **โมเดล 2PL เป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลคำตอบของแบบสอบทุกฉบับมากที่สุด**

ผลการเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดลเอกมิติ 2 PL และโมเดลพหุมิติ 2 PL พบว่า โมเดลพหุมิติไม่ได้สอดคล้องกลมกลืนกับข้อมูลได้ดีกว่าโมเดลเอกมิติอย่างมีนัยสำคัญ สำหรับแบบสอบทุกฉบับ โดยจะเห็นว่า ค่า G^2, AIC และ BIC ของโมเดลเอกมิติมีค่าต่ำกว่าโมเดลพหุมิติ และการตรวจสอบนัยสำคัญของผลการเปรียบเทียบความสอดคล้องกลมกลืนไม่มีนัยสำคัญทางสถิติ ($p>.05$) นั่นคือ ในภาพรวม

ค่าความแปรปรวนร่วมของความสามารถย่อยแต่ละองค์ประกอบที่มีต่อกัน (เช่นความแปรปรวนร่วมระหว่าง θ_{Ac} และ θ_{Mn}) ไม่ได้ช่วยเพิ่มสัดส่วนความแปรปรวนที่สามารถอธิบายได้ ดังนั้นจึงสรุปว่า **ข้อสอบที่สร้างขึ้นสามารถใช้วัดความสามารถในแบบภาพรวมได้ดีกว่าการวัดแบบแยกรายองค์ประกอบ**

ผลการประเมินคุณภาพข้อสอบ ด้วยโมเดลเอกมิติ 2 PL ในขั้นตอนของการพัฒนาคลังข้อสอบขั้นต้น ได้พิจารณาคุณภาพข้อสอบจากค่าความคลาดเคลื่อนของการประมาณค่าพารามิเตอร์เป็นหลัก ข้อสอบที่ตัดออกเป็นข้อสอบที่มีค่าความเคลื่อนสูงผิดปกติ อันเนื่องมาจากพารามิเตอร์ข้อสอบข้อนั้นๆ ไม่สามารถระบุเป็นค่าที่แน่นอนได้ และเลือกตัดข้อสอบที่มีค่าพารามิเตอร์ความยากสูงผิดปกติและต่ำผิดปกติด้วย ส่วนค่าอำนาจจำแนกใช้เกณฑ์ของ Baker and Kim (2017) ที่ตัดข้อสอบที่มีค่าการจำแนกติดลบ และเลือกข้อสอบที่มีค่าการจำแนกตั้งแต่ 0 ขึ้นไป พบว่า มีข้อสอบที่ถูกตัดทิ้งจำนวน 48 ข้อ คงเหลือข้อสอบที่จะนำเข้าคลังข้อสอบแบบปรับเหมาะจำนวน 212 ข้อ ซึ่งจำนวนข้อสอบ 212 ข้อนี้เพียงพอสำหรับการพัฒนาเป็นคลังข้อสอบแบบปรับเหมาะต่อไป (Drasgow, 1989)

2.4 ผลการคัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบ จำแนกตามองค์ประกอบและชุดแบบทดสอบ พบว่า องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ มีข้อสอบที่ถูกคัดเลือกไว้ 43 ข้อ องค์ประกอบที่ 2 การจัดการสารสนเทศ มีข้อสอบที่ถูกคัดเลือกไว้ 41 ข้อ องค์ประกอบที่ 3 การบูรณาการสารสนเทศ มีข้อสอบที่ถูกคัดเลือกไว้ 43 ข้อ องค์ประกอบที่ 4 การประเมินสารสนเทศ มีข้อสอบที่ถูกคัดเลือกไว้ 40 ข้อ และองค์ประกอบที่ 5 การสื่อสารสารสนเทศ มีข้อสอบที่ถูกคัดเลือกไว้ จำนวน 45 ข้อ โดยมีค่าเฉลี่ยพารามิเตอร์ความยากข้อสอบเท่ากับ .18 (SD.=2.06) และค่าเฉลี่ยพารามิเตอร์การจำแนกเท่ากับ .70 (SD.=.39) ในภาพรวมสรุปได้ว่าข้อสอบที่นำเข้าคลังข้อสอบแบบปรับเหมาะเป็นข้อสอบยากปานกลางและจำแนกได้ปานกลาง

3. สรุปผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

3.1 ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 การพัฒนาระบบในการวิจัยครั้งนี้ใช้ภาษา PHP เนื่องจากเป็นระบบการทดสอบที่สามารถใช้งานผ่านระบบออนไลน์ โปรแกรมที่พัฒนาประกอบด้วย 4 ส่วน ได้แก่ 1) การลงทะเบียน 2) การสร้างชุดข้อสอบ 3) การทดสอบ ซึ่งประกอบด้วย การประมาณค่าความสามารถ การคัดเลือกข้อสอบ และการยุติการทดสอบ 4) การรายงานผลการทดสอบ และมีการพัฒนาคู่มือการระบบการทดสอบฯเพื่อให้ผู้เกี่ยวข้องสามารถเข้าใจและนำไปใช้งานได้

3.2 ผลการพัฒนาคู่มือการใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 แบ่งออกเป็น 2 ส่วน คือ 1) คู่มือการใช้งานสำหรับผู้ใช้งานทั่วไป (user) ประกอบด้วย การลงทะเบียน การเข้าใช้งานระบบครั้งแรก การเข้าสู่ระบบเพื่อเริ่มทำแบบทดสอบ และการรายงานผลการทดสอบ และ 2) คู่มือการใช้งานสำหรับผู้จัดการระบบ (admin) ประกอบด้วย การเข้าสู่ระบบ การจัดการข้อสอบ การแก้ไขข้อสอบ การสร้างชุดข้อสอบ และการจัดการสมาชิก และการดูรายงานผลการทดสอบของสมาชิก

4 สรุปผลการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้

4.1 การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้กับนิสิตนักศึกษาสายสังคมศาสตร์ พบว่า นิสิตนักศึกษาชายมีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 สูงกว่านิสิตนักศึกษาหญิงอย่างมีนัยสำคัญทางสถิติ เมื่อพิจารณาจากคะแนนเฉลี่ย พบว่านิสิตนักศึกษาชายมีคะแนนเฉลี่ยด้านการเข้าถึงสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด ในขณะที่นิสิตนักศึกษาหญิงมีคะแนนเฉลี่ยด้านการสื่อสารสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด แต่เมื่อทดสอบความแตกต่างกันตามองค์ประกอบย่อยทั้ง 5 องค์ประกอบ ได้แก่ การเข้าถึงสารสนเทศ การจัดการสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ และการสื่อสารสารสนเทศ พบว่านิสิตนักศึกษาชายและหญิงสายสังคมศาสตร์ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

4.2 การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้กับนิสิตนักศึกษาสายวิทยาศาสตร์ นิสิตนักศึกษาชายสายวิทยาศาสตร์มีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ไม่แตกต่างกับนิสิตนักศึกษาหญิงอย่างมีนัยสำคัญทางสถิติ เมื่อพิจารณาจากคะแนนเฉลี่ย พบว่านิสิตนักศึกษาชายมีคะแนนเฉลี่ยด้านการเข้าถึงสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด ในขณะที่นิสิตนักศึกษาหญิงมีคะแนนเฉลี่ยด้านการสื่อสารสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด แต่เมื่อทดสอบความแตกต่างกันตามองค์ประกอบย่อยทั้ง 5 องค์ประกอบ ได้แก่การเข้าถึงสารสนเทศ การจัดการสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ และการสื่อสารสารสนเทศ พบว่านิสิตนักศึกษาชายและหญิงสายวิทยาศาสตร์ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

5. สรุปผลการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

5.1 ผลการตรวจสอบคุณภาพของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยผู้ทรงคุณวุฒิ พบว่าในภาพรวม ระบบการทดสอบมีคุณภาพอยู่ในระดับมากที่สุด สอดคล้องกับการพิจารณาคุณภาพของระบบการทดสอบในแต่ละด้าน พบว่า คุณภาพด้านความมีประโยชน์ ด้านความเป็นไปได้ในการนำไปใช้ ด้านความเหมาะสม และด้านความถูกต้อง มีคุณภาพอยู่ในระดับมากที่สุดทั้ง 4 ด้าน

5.2 นิสิตนักศึกษามีความพึงพอใจต่อระบบการทดสอบในภาพรวมระดับมาก เมื่อแยกตามรายการย่อย พบว่า นิสิตนักศึกษามีความพึงพอใจต่อระบบการทดสอบด้านภาพรวมของระบบ หน้าจอของระบบ การเรียนรู้ของผู้ใช้งาน และความสามารถของระบบ ในระดับมากเช่นกัน ส่วนความพึงพอใจในระบบการทดสอบด้านคำศัพท์และสารสนเทศระบบ อยู่ในระดับค่อนข้างพึงพอใจ

อภิปรายผลการวิจัย

จากผลการวิจัยข้างต้น นำมาซึ่งการอภิปรายผลการวิจัยได้ 5 ประเด็นตามผลการวิจัย คือ (1) ผลการพัฒนาองค์ประกอบและตัวบ่งชี้ในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี (2) ผลการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี (3) ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี (4) ผลการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีไปทดลองใช้ (5) ผลการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี โดยแต่ละประเด็นมีรายละเอียดดังนี้

1. อภิปรายผลการพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

การพัฒนาองค์ประกอบในการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ในการวิจัยครั้งนี้ได้ทบทวนเอกสารและงานวิจัยทั้งไทยและต่างประเทศที่เกี่ยวข้องกับองค์ประกอบการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร ในช่วงเวลาสิบกว่าปีที่ผ่านมา เพื่อให้ได้องค์ประกอบที่เป็นตัวแทนที่ดีและเป็นสากล สามารถนำไปใช้ในการสร้างแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารได้ 6 องค์ประกอบคือ (1) การเข้าถึงสารสนเทศ (information accessibility) (2) การจัดการสารสนเทศ (information management) (3) การบูรณาการสารสนเทศ (information integration) (4) การประเมินสารสนเทศ (information evaluation) และ (5) การสื่อสารสารสนเทศ (information communication) (6) การสร้างสรรค์ (information creation) แต่การวิจัยเรื่องนี้สร้างแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารจาก 5 องค์ประกอบ ไม่มีองค์ประกอบการสร้างสรรค์ ทั้งนี้เป็นเพราะว่า การวัดการสร้างสรรค์ควรเป็นข้อสอบความเรียงไม่จำกัดคำตอบ แต่ด้วยข้อจำกัดของงานวิจัยที่พัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ต้องมีคำตอบถูกที่แน่นอนและมีการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบไว้ล่วงหน้าในการคัดเลือกข้อสอบให้ผู้สอบทำ จึงไม่สามารถที่จะใช้ข้อสอบความเรียงแบบไม่จำกัดคำตอบเพื่อวัดความคิดสร้างสรรค์ในระบบที่พัฒนาขึ้น ส่วนการสร้างแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารจาก 5 องค์ประกอบ มีนิยามของแต่ละองค์ประกอบที่ได้จากการสังเคราะห์เอกสารและงานวิจัยที่เกี่ยวข้องในแต่ละองค์ประกอบ ดังนี้

องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ (information accessibility) ผลการสังเคราะห์ได้นิยามว่าเป็น ความรู้และความสามารถในการสืบค้นและรวบรวมสารสนเทศรวมทั้งอธิบายหรือแสดงวิธีการเรียกสารสนเทศที่มีอยู่กลับคืนมาเพื่อการใช้งานในระบบดิจิทัลได้ สอดคล้องกับนิยามของ Educational Testing Service (2002) และUNESCO (2008) ที่ระบุว่า การเข้าถึงสารสนเทศ คือ การรู้สารสนเทศที่ต้องการและรู้ว่าจะเข้าถึงสารสนเทศดังกล่าวได้อย่างไร โดยพฤติกรรมบ่งชี้ที่ระบุในแบบวัดที่พัฒนาในงานวิจัยครั้งนี้มุ่งเน้น

การวัดพฤติกรรมการสืบค้นข้อมูลจากเว็บไซต์ ฐานข้อมูลออนไลน์ รวมถึงทักษะการเลือกเข้าถึงข้อมูลผ่านช่องทางต่างๆในระบบอินเตอร์เน็ตและสื่อสังคมออนไลน์ที่น่าเชื่อถือ

องค์ประกอบที่ 2 การจัดการสารสนเทศ (information management) ผลการสังเคราะห์ได้นิยามว่าเป็น ความรู้และความสามารถในการจัดระบบหมวดหมู่ของสารสนเทศในระบบดิจิทัลให้ง่ายต่อการการเรียกค้นและการประยุกต์ใช้ให้เหมาะสมกับงานแต่ละประเภทได้ สอดคล้องนิยามของ Educational Testing Service (2002) และUNESCO (2008) ที่ระบุว่า การจัดการสารสนเทศคือ ความสามารถในการจัดเก็บสารสนเทศที่มีอยู่ไว้เป็นหมวดหมู่ต่างๆง่ายต่อการใช้งาน โดยพฤติกรรมบ่งชี้ที่ระบุในแบบวัดที่พัฒนาในงานวิจัยครั้งนี้มุ่งเน้น การวัดพฤติกรรมการบันทึกข้อมูล การเลือกใช้อุปกรณ์บันทึกข้อมูล ตลอดจนทักษะในการรวบรวมข้อมูลเพื่อให้สะดวกต่อการเรียกใช้ข้อมูล เป็นต้น

องค์ประกอบที่ 3 การบูรณาการสารสนเทศ (information integration) ผลการสังเคราะห์ได้นิยามว่าเป็น ความรู้และความสามารถในการแปลความหมาย สรุป เปรียบเทียบ จำแนกความเหมือนและความแตกต่างของสารสนเทศจากแหล่งข้อมูลต่างๆในระบบดิจิทัลได้ สอดคล้องกับนิยามของ Educational Testing Service (2002) และ UNESCO (2008) ที่ระบุว่า การบูรณาการสารสนเทศคือ ความสามารถในการแปลความหมาย สรุป เปรียบเทียบ แยกความแตกต่างของข้อมูลที่มีความคล้ายหรือมีรูปแบบการนำเสนอที่หลากหลาย โดยพฤติกรรมบ่งชี้ที่ระบุในแบบวัดที่พัฒนาในงานวิจัยครั้งนี้มุ่งเน้น การวัดพฤติกรรมการตีความหมายของข้อมูลในรูปแบบของบทความ แผนภาพ หรือกราฟ จากสื่อต่างๆ การวิเคราะห์เปรียบเทียบสารสารเทศที่ได้รับจากโฆษณา บทความ สื่อสังคมออนไลน์ เป็นต้น

องค์ประกอบที่ 4 การประเมินสารสนเทศ (information evaluation) ผลการสังเคราะห์ได้นิยามว่าเป็น ความรู้และความสามารถในการตัดสินคุณภาพ ความสอดคล้อง ประโยชน์ หรือประสิทธิภาพของสารสนเทศที่นำมาใช้ในการออกแบบและแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศและการสื่อสาร ในระบบดิจิทัลได้ สอดคล้องกับนิยามของ Educational Testing Service (2002) และ UNESCO (2008) ที่ระบุว่า การประเมินสารสนเทศ คือ การตัดสินเกี่ยวกับคุณภาพ ความสัมพันธ์ ประโยชน์ หรือ ประสิทธิภาพของสารสนเทศ โดยพฤติกรรมบ่งชี้ที่ระบุในแบบวัดที่พัฒนาในงานวิจัยครั้งนี้มุ่งเน้น พฤติกรรมการมีวิจารณญาณในการส่งต่อ เผยแพร่ข้อมูลข่าวสาร ตลอดจนตัดสินคุณภาพของข้อมูลในรูปแบบของภาพ ข้อความ บทความโฆษณาและสถานการณ์จำลอง เป็นต้น

องค์ประกอบที่ 5 การสื่อสารสารสนเทศ (information communication) ผลการสังเคราะห์ได้นิยามว่าเป็น ความรู้และความสามารถในการจัดทำและเผยแพร่สารสนเทศในระบบดิจิทัลเพื่อตอบสนองความต้องการของกลุ่มบุคคลและบริบทต่างๆที่เป็นการเคารพสิทธิ์ความเป็นส่วนตัวได้ สอดคล้องกับนิยามของ Somerville et al. (2008) ที่ระบุว่า การสื่อสารสารสนเทศ คือ การเผยแพร่สารสนเทศให้เหมาะสมกับกลุ่มผู้รับในรูปแบบของดิจิทัลที่มีประสิทธิภาพ เช่น การจัดรูปแบบเอกสารเพื่อให้เป็นประโยชน์กับกลุ่มใดกลุ่มหนึ่ง การเลือกและจัดระเบียบภาพนิ่งสำหรับการนำเสนอที่แตกต่างกันเพื่อผู้ชมที่แตกต่างกัน การออกแบบใบปลิวเพื่อโฆษณากับกลุ่มผู้ใช้ที่แตกต่างกัน เป็นต้น

กล่าวโดยสรุปการสร้างแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารจาก 5 องค์ประกอบในการวิจัยครั้งนี้ ได้กำหนดนิยามจากเอกสารงานวิจัยที่เกี่ยวข้องพร้อมทั้งมีการตรวจสอบนิยามโดยใช้ดุลยพินิจของผู้ทรงคุณวุฒิและมีการปรับแก้ให้เหมาะสมก่อนที่จะนำไปสร้างข้อสอบต่อไป

2. อภิปรายผลการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

2.1 การสร้างข้อสอบแบบเลือกตอบ 5 ตัวเลือก วัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ตามนิยามเชิงปฏิบัติการที่สร้างไว้ ทั้ง 5 องค์ประกอบที่มีการออกแบบเพื่อให้วัดพฤติกรรมการเรียนรู้ในระดับความรู้ ความเข้าใจ การนำไปใช้ และการวิเคราะห์ ตามทฤษฎีการเรียนรู้ของบลูม การที่เลือกใช้รูปแบบข้อสอบเป็นแบบเลือกตอบ 5 ตัวเลือก มีข้อดี คือ วัดความสามารถของผู้สอบได้ครอบคลุมเนื้อหาที่ต้องการและวัดความคิดขั้นสูงได้ โดยเฉพาะการตรวจให้คะแนนมีความเป็นปรนัย จึงเหมาะสมกับงานวิจัยครั้งนี้ที่ใช้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ที่ต้องมีการประมาณค่าพารามิเตอร์ของข้อสอบ และผู้สอบทันที อีกทั้งในการวิจัยครั้งนี้ออกแบบการวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบที่มีข้อสอบร่วมในการปรับเทียบ การใช้ข้อสอบแบบเลือกตอบหลายตัวเลือกจึงมีความเหมาะสม

2.2 การตรวจสอบความตรงเชิงเนื้อหาก่อนนำไปทดลองใช้ โดยให้ผู้เชี่ยวชาญพิจารณาความสอดคล้องของเนื้อหาตามนิยามและพฤติกรรมที่วัด พบว่า ข้อสอบส่วนใหญ่ผ่านการประเมินความตรงเชิงเนื้อหา แต่ก็มีข้อสอบบางข้อมีข้อเสนอแนะให้ปรับปรุง เช่น การปรับปรุงด้านการใช้ภาษาเพื่อให้เข้าใจง่าย การปรับปรุงข้อคำถามเพื่อให้สอดคล้องตามนิยามการวัด การปรับปรุงตัวเลือกถูกและตัวลวงให้มีความสอดคล้องกัน การปรับปรุงรูปภาพแผนภาพให้ชัดเจน เป็นต้น สอดคล้องกับ ศิริชัย กาญจนวาสี (2555) ที่ได้กล่าวว่าการใช้โครงสร้างภาษาที่ซับซ้อนเกินไป ข้อคำถามที่ใช้ภาษากำกวม ตลอดจนข้อสอบมีความยากไม่เหมาะสม จะส่งผลกระทบต่อความตรงของข้อสอบ ผู้วิจัยจึงได้ปรับปรุงคุณภาพของข้อสอบให้มีคุณภาพที่ดีขึ้นตามคำแนะนำและยังคงจำนวนข้อเท่าเดิม คือ 260 ข้อ

2.3 การตรวจสอบคุณภาพของข้อสอบที่ได้จากการทดลองใช้ หลังจากที่ได้ปรับแก้ข้อสอบตามคำแนะนำแล้ว ผลการวิเคราะห์คุณภาพของข้อสอบตามทฤษฎี CTT โดยแบบสอบมีทั้งหมดจำนวน 10 ชุด ในภาพรวมพบว่า ค่าเฉลี่ยความยาก (p) ของแบบสอบปานกลาง ขณะที่ค่าอำนาจจำแนกเฉลี่ย (r) ของแบบสอบจำแนกผู้สอบได้พอใช้ สอดคล้องกับ ศิริชัย กาญจนวาสี (2555) ที่กล่าวถึงดัชนีบ่งชี้คุณภาพของข้อสอบอิงกลุ่มว่า ระดับความยากง่ายของข้อสอบ (p) ที่มีค่าระหว่าง 0.2 ถึง 0.8 ถือว่าเป็นข้อสอบที่มีความยากง่ายพอเหมาะ และข้อสอบทั้งฉบับควรมีระดับความยากง่ายเฉลี่ยประมาณ 0.50 แต่เนื่องจากการวิเคราะห์ข้อสอบด้วยทฤษฎีการทดสอบแบบดั้งเดิมมีข้อจำกัดที่การคำนวณค่าดัชนีคุณภาพของข้อสอบแปรเปลี่ยนไปตามกลุ่มผู้สอบประกอบกับข้อจำกัดเกี่ยวกับจำนวนตัวอย่างที่เก็บข้อมูลได้จริงไม่เป็นไปตามแผนในบางสาขาวิชา รวมทั้งการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ต้องประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบโดยใช้ทฤษฎีการตอบสนองข้อสอบ ดังนั้นกระบวนการวิเคราะห์คุณภาพข้อสอบเพื่อนำเข้าคลังข้อสอบแบบปรับเหมาะจึงใช้ทฤษฎีการตอบสนองข้อสอบเป็นหลักในการวิจัยครั้งนี้ ซึ่งทฤษฎีการตอบสนองข้อสอบมีจุดเด่นที่สำคัญ 2 ประการในการวิจัยครั้งนี้ คือ 1) การตรวจสอบความสอดคล้องกลมกลืนของข้อมูลคำตอบกับโมเดลการตอบสนองข้อสอบแบบต่างๆ และ 2) การปรับเทียบแบบสอบ ดังมีรายละเอียดต่อไปนี้

ประการที่ 1 การตรวจสอบความสอดคล้องกลมกลืนของโมเดลการตอบสนองข้อสอบ ผลการวิจัยแสดงให้เห็นว่า โมเดลการตอบสนองข้อสอบแบบเอกมิติ ชนิด 2 พารามิเตอร์ เป็นโมเดลที่สอดคล้องกลมกลืนกับข้อมูลคะแนนคำตอบจากแบบสอบทั้ง 10 ฉบับมากที่สุด ซึ่งผู้วิจัยได้เปรียบเทียบความสอดคล้องกับโมเดลแข่งขัน (completing model) ต่างๆ ไม่ว่าจะเป็นการเปรียบเทียบระหว่างโมเดลการวัดแบบ 1, 2, และ 3 พารามิเตอร์ก่อนในขั้นต้น จนได้ข้อสรุปว่า โมเดลแบบ 2 พารามิเตอร์สอดคล้องกับข้อมูลมากที่สุด จุดเด่นที่น่าสนใจของขั้นตอนนี้ คือ ผู้วิจัยต้องการแสดงให้เห็นว่า การพิจารณาความสอดคล้องของโมเดลจากดัชนีเชิงสัมพันธ์ AIC และ BIC เพียงอย่างเดียวนั้น อาจนำไปสู่ข้อสรุปคลาดเคลื่อนว่า โมเดลแบบ 3 พารามิเตอร์ เป็น

โมเดลที่สอดคล้องกลมกลืนมากที่สุด แต่เมื่อผู้วิจัยใช้การทดสอบอัตราส่วนความเป็นไปได้ (likelihood-ratio (LR) test) กลับพบว่า จำนวนพารามิเตอร์ที่เพิ่มขึ้นของโมเดลแบบ 3 พารามิเตอร์ ไม่ได้เพิ่มสัดส่วนความแปรปรวนที่อธิบายได้ ได้อย่างมีนัยสำคัญทางสถิติ นั่นหมายความว่า ถ้าผู้วิจัยยอมรับว่าโมเดลแบบ 3 พารามิเตอร์ เป็นโมเดลที่สอดคล้องกับข้อมูลมากที่สุด ก็จะเป็นการยอมรับผลวิเคราะห์จากโมเดลไม่ประหยัด (non-parsimonious model) ส่งผลให้อำนาจหรือประสิทธิภาพ (power) ในการประมาณค่าความสามารถของผู้สอบลดลงด้วย (Brown, Templin & Cohen, 2015)

แต่อย่างไรก็ตาม แม้ว่าข้อสรุปจากการทดสอบ LR จะยืนยันว่าโมเดลแบบ 2 พารามิเตอร์เป็นโมเดลที่สอดคล้องกลมกลืนมากที่สุด แต่ก็เป็นการแสดงว่าให้เห็นว่า "ในภาพรวม" ข้อสอบทั้งหมดสอดคล้องกลมกลืนกับโมเดลแบบ 2 พารามิเตอร์ ซึ่งมีความเป็นไปได้ที่ข้อสอบบางข้อ "อาจจะไม่จำเป็น" ต้องประมาณค่าพารามิเตอร์การจำแนกก็ได้ เนื่องจากผลต่างของค่า AIC และ BIC ระหว่างโมเดลแบบ 1 พารามิเตอร์และ 2 พารามิเตอร์ไม่มาก วิธีการที่เป็นไปได้สำหรับตรวจสอบความสอดคล้องกลมกลืนในเชิงลึก ซึ่งอยู่นอกเหนือขอบเขตของการวิจัยครั้งนี้ คือ การตรวจสอบขนาดอิทธิพล (effect size) หรือนัยสำคัญทางการนำไปใช้ (practical significance) ของผลต่างความสอดคล้องกลมกลืนระหว่างโมเดลทั้งสอง ควบคู่ไปกับการทดสอบนัยสำคัญทางสถิติ จากนั้นวิเคราะห์ความสอดคล้องกลมกลืนใหม่อีกครั้งด้วยโมเดลการตอบสนองข้อสอบแบบไฮบริด (hybrid item response model: Raykov & Marcoulides, 2018) แล้วจึงเปรียบเทียบนัยสำคัญของความสอดคล้องกลมกลืนอีกครั้ง

ขั้นต่อมา คือ การเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดลเอกมิติกับโมเดลพหุมิติ โดยภายหลังจากที่ผู้วิจัยได้ข้อสรุปว่า โมเดลแบบ 2 พารามิเตอร์เป็นโมเดลที่สอดคล้องกับข้อมูลมากที่สุดแล้ว จึงตรวจสอบความเป็นพหุมิติของแบบสอบทั้ง 10 ฉบับ ผลการวิจัยแสดงให้เห็นว่า โมเดลเอกมิติสอดคล้องกลมกลืนกับข้อมูลมากกว่าโมเดลพหุมิติ ซึ่งไม่เป็นไปตามผลการสังเคราะห์องค์ประกอบของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารของการวิจัยนี้

สาเหตุที่เป็นไปได้ที่ทำให้ข้อสอบที่ผู้วิจัยพัฒนาขึ้นไม่สามารถแสดงคุณสมบัติความเป็นพหุมิติได้ คือ องค์ประกอบทั้ง 5 ด้าน ได้แก่ การเข้าถึงสารสนเทศ การจัดการสารสนเทศ การบูรณาการสารสนเทศ การประเมินสารสนเทศ และการสื่อสารสนเทศ อาจจะสัมพันธ์กันเองสูง บางองค์ประกอบอาจจะมีเนื้อหาร่วมและต่อเนื่องกันมาก ทำให้ข้อสอบที่ผู้วิจัยพัฒนาขึ้นไม่สามารถกระตุ้นให้ผู้สอบใช้ความสามารถต่างกันตอบข้อสอบที่วัดต่างองค์ประกอบกันได้ ดังนั้น ผู้วิจัยจึงคาดว่า ถ้ายุบรวมบางองค์ประกอบเข้าด้วยกัน เช่น รวมด้านการเข้าถึงสารสนเทศ และการจัดการสารสนเทศ อาจจะทำให้ข้อสอบแสดงความเป็นพหุมิติของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร ได้ชัดเจนมากขึ้น แต่อย่างไรก็ตาม สำหรับการวิจัยในครั้งนี้ แม้ว่าโมเดลเอกมิติสอดคล้องกลมกลืนกับข้อมูลมากกว่าโมเดลพหุมิติ ก็ไม่ได้หมายความว่า ผู้วิจัยปฏิเสธหรือละเลยความเป็นพหุมิติของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสาร และถือเป็นจุดเริ่มต้นที่ดีเพื่อพัฒนาคลังข้อสอบแบบพหุมิติต่อไปในอนาคต

ประการที่ 2 การปรับเทียบแบบสอบ เนื่องจากผู้วิจัยคำนึงถึงความอ่อนล้าของผู้สอบที่น่าจะเกิดขึ้นจากการตอบข้อสอบทั้งหมดจำนวน 260 ข้อ จึงจัดข้อสอบเป็นแบบสอบจำนวน 10 ฉบับ แต่ละฉบับมีข้อสอบร่วม 10 ข้อ และข้อสอบปกติจำนวน 25 ข้อ และให้ผู้สอบแต่ละคนทำข้อสอบเพียงชุดเดียว จากนั้นจึงวิเคราะห์ปรับเทียบข้อสอบทั้งหมดเข้าด้วยกัน ด้วยวิธีการปรับเทียบรวม (pooled equating) หรือการปรับเทียบพร้อมกัน (concurrent calibration method) แล้วประมาณค่าพารามิเตอร์ด้วยวิธี full information maximum likelihood ผลการปรับเทียบแสดงให้เห็นว่า ข้อสอบทุกข้อมีค่าพารามิเตอร์ความยากอยู่บนสเกลลอจิตเดียวกัน นั่นคือ ค่าพารามิเตอร์ความสามารถของผู้สอบที่ทำแบบสอบต่างฉบับ ก็จะอยู่

บนสเกลลอจิตเดียวกันด้วย แนวคิดการปรับเทียบแบบสอบก่อนที่จะนำข้อสอบเข้าคลังข้อสอบแบบปรับเหมาะนี้ เป็นวิธีที่เหมาะสม เพราะทำให้ผู้วิจัยมั่นใจได้ว่า ค่าพารามิเตอร์ต่างๆ นอกจากจะไม่แปรเปลี่ยน (invariance) แล้วยังสามารถเปรียบเทียบกันได้ด้วย วิธีการที่ไม่เหมาะสมคือ การแยกวิเคราะห์รายฉบับแล้วนำข้อสอบเข้าคลัง ซึ่งแม้ว่าข้อสอบเหล่านี้จะมีคุณสมบัติความไม่แปรเปลี่ยนของค่าพารามิเตอร์ความสามารถผู้สอบและค่าพารามิเตอร์ข้อสอบ แต่ค่าพารามิเตอร์เหล่านี้มาจากการวิเคราะห์ (calibration) แยกกัน จึงนำมาเปรียบเทียบกันไม่ได้

หัวใจสำคัญของการปรับเทียบข้อสอบคือ คุณภาพของข้อสอบร่วม (Holland & Rubin, 1982; Kolen & Brennan, 2014) ซึ่งผู้วิจัยพิจารณาจากเนื้อหาของข้อสอบเป็นหลัก แล้วคัดเลือกข้อสอบที่คาดว่าจะเป็นข้อสอบคุณภาพดีให้เป็นข้อสอบร่วม ผลการวิเคราะห์แสดงให้เห็นว่า ข้อสอบร่วมจำนวน 2 ข้อจาก 10 ข้อ ไม่สามารถจำแนกความสามารถของผู้สอบได้ดี และส่งผลเสียต่อประสิทธิภาพต่อการปรับเทียบข้อสอบและความแม่นยำของค่าพารามิเตอร์ต่างๆ ผู้วิจัยจึงตัดข้อสอบร่วมทั้งสองข้อทิ้ง และแม้ว่าข้อสอบร่วมจะเหลือ 8 ข้อ แต่ก็มีจำนวนเพียงพอต่อการวิเคราะห์ (Haberman, 2008; Hanson & Beguin, 2002; Kim, 2006; Kim & Cohen, 1998, 2002; Sykes, 1997; Vale, 1986)

2.4 ผลการคัดเลือกข้อสอบที่มีคุณภาพเพื่อจัดเก็บเข้าสู่ระบบคลังข้อสอบ ผลการวิเคราะห์แสดงให้เห็นว่า ผู้วิจัยคัดเลือกข้อสอบที่มีค่าพารามิเตอร์ความยากระหว่าง -9.46 ถึง +7.27 ลิจิต และมีค่าพารามิเตอร์การจำแนกระหว่าง .10 ถึง 2.3 เข้าคลังข้อสอบแบบปรับเหมาะ ซึ่งข้อสอบเหล่านี้มีคุณสมบัติไม่เป็นไปตามข้อเสนอแนะของ Baker and Kim (2017) และ Hambleton, Swaminathan and Roger (1991) แต่อย่างไรก็ตาม ผู้วิจัยประยุกต์ใช้แนวคิดและหลักการเดียวกันเพื่อพิจารณาคัดเลือกข้อสอบเข้าคลังข้อสอบแบบปรับเหมาะ กล่าวคือ เกณฑ์อย่างง่ายสำหรับพิจารณาคุณภาพข้อสอบตามที่ Baker and Kim (2017) และ Hambleton, Swaminathan and Roger (1991) เสนอแนะ มีแนวคิดพื้นฐานจากโค้งคุณลักษณะข้อสอบ (item characteristic curve: ICC) โดยข้อสอบที่มีคุณภาพและใช้ทดสอบทั่วไป ควรมีค่าพารามิเตอร์ความยากอยู่ในช่วง -3 ถึง +3 ลอจิต เนื่องจาก ICC สามารถตรวจจับความแปรปรวนของค่าพารามิเตอร์ความสามารถของผู้สอบทั้งหมดได้ดี นั่นหมายความว่า ข้อสอบที่มีค่าความยากอยู่ในช่วงดังกล่าว สามารถบรรยายความสัมพันธ์ระหว่างความน่าจะเป็นของการตอบข้อสอบข้อนั้นถูกกับค่าพารามิเตอร์ความสามารถของผู้สอบ โดยผู้สอบที่มีความสามารถสูงจะมีโอกาสตอบข้อสอบข้อนั้นถูก มากกว่าผู้สอบที่มีความสามารถต่ำ เมื่อผู้วิจัยประยุกต์ใช้หลักการนี้กับข้อสอบที่มีค่าความยากต่ำกว่า -3 ลอจิต และสูงกว่า +3 ลอจิต พบว่าข้อสอบเหล่านี้มี ICC ที่สามารถตรวจจับความแปรปรวนของความสามารถผู้สอบได้บางส่วน เช่น ข้อสอบที่มีค่าความยากต่ำกว่า -3 ลอจิต จะมี ICC ที่ตรวจจับความแปรปรวนของผู้สอบที่มีความสามารถต่ำได้ดีกว่าผู้สอบที่มีความสามารถสูง ในขณะที่ข้อสอบที่มีค่าความยากสูงกว่า +3 ลอจิต จะมี ICC ที่ตรวจจับความแปรปรวนของผู้สอบที่มีความสามารถสูงได้ดีกว่าผู้สอบที่มีความสามารถต่ำ ดังนั้นข้อสอบที่มีค่าพารามิเตอร์ความยากน้อยกว่า -3 และมากกว่า +3 ลอจิต จึงเหมาะสมกับผู้สอบไม่เก่งมากและผู้สอบเก่งมาก ตามลำดับ

หลักการที่สนับสนุนแนวคิดของผู้วิจัยข้างต้น พิจารณาได้จาก Embretson and Reise (2000) และ Weiss (1982) กล่าวคือ ข้อสอบสำหรับการทดสอบแบบปรับเหมาะนั้น มีจุดประสงค์เพื่อนำไปใช้ประมาณค่าพารามิเตอร์ความสามารถผู้สอบตั้งแต่ระดับต่ำไปจนถึงระดับสูง ข้อสอบจึงจำเป็นต้องครอบคลุมความสามารถของผู้สอบทั้งหมด นั่นหมายความว่า ข้อสอบที่ผู้วิจัยคัดเลือกเข้าคลังข้อสอบแบบปรับเหมาะจะครอบคลุมผู้สอบที่มีความสามารถแตกต่างกันด้วย

สำหรับค่าพารามิเตอร์การจำแนกนั้น Baker and Kim (2017) อธิบายว่า ข้อสอบที่มีค่าการจำแนกติดลบ เป็นข้อสอบที่ให้สารสนเทศคลาดเคลื่อน ไม่สามารถอธิบายความแปรปรวนของค่าพารามิเตอร์

ความสามารถของผู้สอบได้ Baker and Kim (2017) จึงถือว่า ข้อสอบที่มีค่าพารามิเตอร์การจำแนกมากกว่า 0 ขึ้นไป เป็นคุณสมบัติของข้อสอบในระยะเริ่มแรกของการพัฒนาคลังข้อสอบ ดังนั้น การที่ผู้วิจัยคัดเลือกข้อสอบที่มีค่าพารามิเตอร์การจำแนกตั้งแต่ .10 ขึ้นไป จึงถือว่ายอมรับได้และสอดคล้องกับกระบวนการพัฒนาคลังข้อสอบแบบปรับเหมาะของการวิจัยนี้

ข้ออภิปรายข้างต้นสัมพันธ์กับจำนวนข้อสอบในคลังข้อสอบแบบปรับเหมาะของการวิจัยนี้ เนื่องจากจำนวนข้อสอบที่เหมาะสมคือ จำนวนข้อสอบที่ครอบคลุมค่าพารามิเตอร์ความสามารถของผู้สอบ ที่มักจะมีค่าอยู่ระหว่าง -3 ถึง +3 ลอจิต Embretson and Reise (2000) กล่าวว่า ไม่มีเกณฑ์ตายตัวเพื่อกำหนดจำนวนข้อสอบในคลัง แต่ข้อสอบในคลังควรเป็นข้อสอบจำแนกได้ดี และมีความยากครอบคลุมทุกช่วงความสามารถ และอาจจะมีอย่างน้อย 100 ข้อก็ได้ ขึ้นอยู่กับปัจจัยหลายอย่าง เช่น เนื้อหาของสิ่งที่วัด ดังนั้น เมื่อพิจารณาข้อสอบจำนวน 212 ข้อที่คัดเลือกเข้าคลังแบบปรับเหมาะของการวิจัยนี้ นอกจากจะมากกว่าข้อเสนอแนะข้างต้นแล้ว ผู้วิจัยยังเชื่อว่าจะครอบคลุมทุกช่วงความสามารถของผู้สอบอีกด้วย

3. อภิปรายผลการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการทดสอบที่ผู้เข้ารับการทดสอบจะได้รับข้อสอบที่มีความแตกต่างกัน โดยข้อสอบที่ได้รับจะมีความเหมาะสมตามระดับความสามารถของผู้เข้ารับการทดสอบแต่ละคน โดยมีการนำกระบวนการทางคอมพิวเตอร์มาใช้ในการคัดเลือกข้อสอบและประมาณค่าความสามารถของผู้เข้ารับการทดสอบ โดยหลักการของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ คือ ผู้เข้ารับการทดสอบจะได้รับข้อสอบข้อแรกจากข้อที่มีระดับความยากปานกลาง หากผู้เข้ารับการทดสอบตอบข้อสอบข้อดังกล่าวผิด ข้อสอบข้อถัดไปที่ผู้เข้ารับการทดสอบได้รับจะมีระดับง่ายลง แต่ผู้เข้ารับการทดสอบตอบข้อสอบข้อดังกล่าวถูก ข้อสอบข้อถัดไปจะระดับที่ยากขึ้น การทดสอบจะดำเนินการไปเรื่อย ๆ กระทั่งสามารถประมาณค่าความสามารถของผู้สอบได้คงที่และมีค่าความคลาดเคลื่อนอยู่ในเกณฑ์ที่ยอมรับได้ การทดสอบจะทำการยุติลง (Čisar et al., 2010; ศิริชัย กาญจนวาสี, 2555)

การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ในครั้งนี้ เป็นการพัฒนาระบบการทดสอบแบบออนไลน์ (on-line) ในการพัฒนาระบบต้องมีการถอดขั้นตอนการดำเนินงาน (algorithm) ต่าง ๆ ของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เพื่อนำขั้นตอนดังกล่าวไปใช้ในการเขียนโปรแกรมให้รองรับระบบการใช้งานออนไลน์ ซึ่งเขียนด้วยภาษา PHP ดังนั้นจึงต้องมีการตรวจสอบขั้นตอนการดำเนินงานให้มีความถูกต้อง โดยการตรวจสอบขั้นตอนการดำเนินงาน ดำเนินการตรวจสอบจากโปรแกรม R ด้วยชุดคำสั่งสำเร็จรูป (package) irtoys เนื่องจากเป็นโปรแกรมมาตรฐานที่ได้รับการยอมรับและมีการใช้งานกันอย่างแพร่หลาย และเมื่อตรวจสอบความถูกต้องของขั้นตอนการดำเนินงานเรียบร้อยแล้ว จึงนำมาเขียนเป็นระบบออนไลน์ด้วยภาษา PHP ซึ่งขั้นตอนดังกล่าวค่อนข้างมีความละเอียดและซับซ้อน เนื่องจากขั้นตอนการดำเนินงานของระบบการทดสอบแบบปรับเหมาะ มีกระบวนการคำนวณค่อนข้างมาก ทำให้ระหว่างการพัฒนาระบบต้องใช้เวลาปรับแก้และตรวจสอบนานกว่าการพัฒนาระบบออนไลน์ประเภทอื่นเพื่อให้เกิดความถูกต้องและแม่นยำในการทดสอบมากที่สุด โดยแนวทางการพัฒนาระบบด้วยวิธีดังกล่าวมีความสอดคล้องกับแนวทางการพัฒนาระบบของ ณภัทร ชัยมงคล โชติกา ภาษีผล และศิริชัย กาญจนวาสี (2559)

สำหรับวิธีการประมาณค่าความสามารถ จากการศึกษาพบว่า การประมาณค่าด้วยความเป็นไปได้สูงสุด (MLE) เป็นวิธีการประมาณค่าที่มีความแม่นยำค่อนข้างสูง แต่มีความคลาดเคลื่อนสูงเช่นกันในกรณีที่ผู้

เข้ารับการทดสอบตอบข้อสอบถูกหรือผิดทั้งหมด ในขณะที่วิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (EAP) จะให้ผลแม่นยำกว่าวิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (MLE) ในกรณีที่ผู้เข้ารับการทดสอบตอบข้อสอบถูกหรือผิดทั้งหมด (Bock & Mislevy, 1982; Hambleton & Swaminathan, 1985; Zheng & Chang, 2015) ดังนั้น การประมาณค่าความสามารถของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21ในครั้งนี้ ใช้ทั้ง 2 วิธี คือ เมื่อผู้เข้ารับการทดสอบตอบข้อสอบถูกและผิดผสมกันจะใช้การประมาณค่าด้วยวิธีความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) แต่หากผู้เข้ารับการทดสอบตอบข้อสอบถูกหรือผิดทั้งหมด จะใช้การประมาณค่าความสามารถด้วยวิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ซึ่งการใช้การประมาณค่าทั้งสองวิธีร่วมกัน จะช่วยทำให้การประมาณค่าความสามารถมีความถูกต้องและแม่นยำมากยิ่งขึ้น

4. อภิปรายผลการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ไปทดลองใช้

นิสิตนักศึกษาชายทั้งสายสังคมศาสตร์และสายวิทยาศาสตร์มีแนวโน้มที่มีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ในภาพรวมสูงกว่านิสิตนักศึกษาหญิง เมื่อจำแนกตามองค์ประกอบ พบว่านิสิตนักศึกษาชายและหญิงที่เป็นตัวอย่างสายสังคมศาสตร์และสายวิทยาศาสตร์ มีคะแนนทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติและเป็นไปในทิศทางเดียวกันทั้งสายสังคมศาสตร์และสายวิทยาศาสตร์ โดยนิสิตนักศึกษาชายมีคะแนนเฉลี่ยด้านการเข้าถึงสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด ในขณะที่นิสิตนักศึกษาหญิงมีคะแนนเฉลี่ยด้านการสื่อสารสารสนเทศสูงที่สุด และมีคะแนนเฉลี่ยด้านการจัดการสารสนเทศต่ำที่สุด

การที่นิสิตนักศึกษาชายมีแนวโน้มที่มีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เฉลี่ยสูงกว่านิสิตนักศึกษาหญิงอาจเนื่องมาจากนิสิตนักศึกษาชายมีประสบการณ์การใช้เทคโนโลยีสารสนเทศและการสื่อสาร รวมทั้งมีจำนวนชั่วโมงที่ใช้งานคอมพิวเตอร์มากกว่าเพศหญิง (Tella & Mutula, 2008) อีกทั้งยังมีการศึกษาหาความรู้เกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสาร และเผชิญสถานการณ์ที่ต้องใช้ความรู้เกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสารในการแก้ปัญหามากกว่าหญิง (Markauskaite, 2006) นอกจากนี้ยังพบว่า นิสิตนักศึกษาชายมีความมั่นใจในการใช้เทคโนโลยีสารสนเทศและการสื่อสารมากกว่าหญิง (Madigan, Goodfellow, & Stone, 2007) จึงทำให้นิสิตนักศึกษาชายมีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 สูงกว่าหญิง

อย่างไรก็ตามผลการศึกษาเกี่ยวกับความแตกต่างของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารระหว่างชายและหญิงที่ผ่านมา พบว่ามีผลการศึกษาที่ขัดแย้งกัน กล่าวคือ มีทั้งผลการศึกษาที่พบว่า ผู้ชายมีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารสูงกว่าผู้หญิงอย่างมีนัยสำคัญทางสถิติ (Tella & Mutula, 2008) แต่ Ritzhaupt, Liu, Dawson and Barron (2013) พบว่าผู้หญิงมีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารสูงกว่าผู้ชายอย่างมีนัยสำคัญทางสถิติ ในขณะที่ Madigan, Goodfellow, and Stone (2007) และ Agbatogun (2010) กลับพบว่าผู้ชายและผู้หญิงมีทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ นั่นคือ ความแตกต่างของทักษะด้านเทคโนโลยีสารสนเทศและการ

สื่อสารระหว่างเพศอาจขึ้นอยู่กับระดับของตัวแปรอื่น เช่น เศรษฐานะ สายการเรียน เจตคติต่อเทคโนโลยีสารสนเทศและการสื่อสาร ความสามารถในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร การควบคุมตนเองเกี่ยวกับการใช้เทคโนโลยีสารสนเทศและการสื่อสาร และความต้องการรับรู้ข้อมูล จึงควรมีการศึกษาในการวิจัยครั้งต่อไป เพื่อเป็นแนวทางในการส่งเสริมทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาระดับปริญญาตรี เนื่องจากทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 เป็นสิ่งจำเป็นสำหรับการศึกษาในระดับมหาวิทยาลัย (Educational Testing Service, 2002) จึงต้องมีการเตรียมความพร้อมนิสิตนักศึกษาให้สามารถเรียนรู้ในสังคมที่ต้องใช้เทคโนโลยีสารสนเทศและการสื่อสารในการพัฒนาประเทศ อีกทั้ง นิสิตนักศึกษาจะต้องใช้ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ในการศึกษาค้นคว้า วิจัย และพัฒนาองค์ความรู้ใหม่เพื่อพัฒนาประเทศให้ก้าวทันกระแสโลก และเพื่อให้สอดคล้องกับความต้องการบุคลากรของหน่วยงานต่างๆ ที่ต้องการบุคลากรที่มีการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร รวมทั้งมีความสามารถที่จะคิดวิเคราะห์เกี่ยวกับข้อมูลสารสนเทศและการสื่อสาร (Wijaya & Surendro, 2007)

เมื่อพิจารณาตามองค์ประกอบย่อย พบว่าการที่นิสิตนักศึกษาชายมีประสบการณ์ในการใช้เทคโนโลยีสารสนเทศและการสื่อสารด้านการเข้าถึงสารสนเทศสูงกว่าด้านอื่นๆ แสดงว่านิสิตนักศึกษาชายมีความสามารถในการสืบค้น และเรียกสารสนเทศกลับคืนเพื่อการใช้งานอยู่ในระดับสูง ในขณะที่นิสิตนักศึกษาหญิงมีทักษะในการสื่อสารสารสนเทศ แสดงว่านิสิตนักศึกษาหญิงมีความสามารถในการส่ง และจัดการเกี่ยวกับจดหมายอิเล็กทรอนิกส์ มีความคิดสร้างสรรค์ในการนำเสนองาน รวมทั้งสามารถในการเลือกใช้ซอฟท์แวร์และสร้างสารสนเทศได้อย่างเหมาะสม อยู่ในระดับสูงกว่าด้านอื่นๆ (Ritzhaupt et al., 2013)

ส่วนด้านการจัดการสารสนเทศ พบว่านิสิตนักศึกษาทั้งชายและหญิงมีทักษะด้านนี้ต่ำกว่าทักษะด้านอื่นๆ ทั้งนี้อาจเนื่องจากความรู้เกี่ยวกับไฟล์ข้อมูล และโปรแกรมที่ใช้ในการจัดระบบและเก็บรวบรวมสารสนเทศเป็นข้อสอบที่ส่วนมากวัดเกี่ยวกับการจัดระบบไฟล์ข้อมูล ได้แก่ ไฟล์เอกสาร ไฟล์รูปภาพ ไฟล์เสียง ไฟล์วิดิทัศน์ และการใช้โปรแกรมในการจัดระบบและเก็บรวบรวมสารสนเทศ ซึ่งเป็นความรู้เฉพาะเชิงลึกกว่าองค์ประกอบอื่นๆ

5. อภิปรายผลการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี

ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ที่ผู้วิจัยพัฒนาขึ้นแบ่งเป็น 4 ส่วน ได้แก่ การลงทะเบียน การสร้างชุดข้อสอบ การทดสอบ และการรายงานผลการทดสอบ โดยการทดสอบแบ่งออกเป็น 4 ส่วน ได้แก่ การประมาณค่าความสามารถ การคัดเลือกข้อสอบ การยุติการทดสอบ และการรายงานผลการทดสอบ

ขั้นตอนแรกในการทดสอบ คือ การลงทะเบียน โดยผู้สอบจะต้องดำเนินการลงทะเบียนการทดสอบผ่านระบบอินเทอร์เน็ต เมื่อดำเนินการกรอกข้อมูลเรียบร้อยแล้วจึงจะสามารถดำเนินการทดสอบได้ สำหรับการสร้างชุดข้อสอบ ผู้วิจัยดำเนินการจัดชุดข้อสอบไว้ล่วงหน้าจำนวน 5 ชุด ชุดละ 5 ข้อ ซึ่งเป็นข้อสอบที่มี

ความยากอยู่ระหว่าง -0.5 ถึง 0.5 ในแต่ละชุดประกอบด้วยเนื้อหาครบทั้ง 5 องค์ประกอบ เพื่อใช้สำหรับสุ่มให้ผู้สอบแต่ละคนทำข้อสอบ 5 ข้อแรก หลังจากผู้สอบทำข้อสอบครบ 5 ข้อ ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะประมาณค่าความสามารถของผู้สอบ และนำค่าความสามารถที่ได้จากการประมาณค่าไปใช้ในการคัดเลือกข้อสอบในลำดับถัดไป สำหรับการเลือกข้อสอบใช้การพิจารณาจากค่าสารสนเทศสูงที่สุด (Maximum Item Information) และควบคุมการเปิดเผยข้อสอบ (exposure rate) ให้ไม่เกินร้อยละ 20 พร้อมทั้งควบคุมเนื้อหาของแบบสอบ (content balance)

การประมาณค่าความสามารถของผู้สอบใช้การประมาณค่าร่วมกัน 2 วิธี คือ วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) ใช้สำหรับกรณีผู้สอบที่ตอบข้อสอบถูกและผิดผสมกัน และวิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ใช้สำหรับกรณีผู้สอบที่ตอบข้อสอบถูกหรือผิดทุกข้อ และระบบจะดำเนินการยุติการทดสอบเมื่อค่าความคลาดเคลื่อนมาตรฐาน มีค่าน้อยกว่าหรือเท่ากับ 0.3 แต่ผู้สอบต้องทำข้อสอบอย่างน้อย 6 ข้อ หากผู้สอบทำข้อสอบครบ 35 ข้อ แล้วความคลาดเคลื่อนมาตรฐานยังสูงกว่า 0.3 ระบบจะยุติการทดสอบโดยอัตโนมัติ

Birdsall (2011) และ Thompson and Weiss (2011) ศึกษาและสังเคราะห์งานวิจัยที่เกี่ยวข้องกับการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ พร้อมทั้งเสนอขั้นตอนการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีประสิทธิภาพ ซึ่งแบ่งออกเป็น 6 ขั้นตอน ได้แก่ การวางแผน การพัฒนาคลังข้อสอบ การทดลองใช้และวิเคราะห์ข้อสอบ การกำหนดรูปแบบระบบการทดสอบ การสร้างระบบการทดสอบ และการดูแลระบบการทดสอบ โดยการกำหนดรูปแบบการทดสอบแบ่งออกเป็น 4 ส่วน ได้แก่ การเริ่มต้นการทดสอบ อัลกอริทึมคัดเลือกข้อสอบ อัลกอริทึมประมาณค่าความสามารถ และเกณฑ์การยุติการทดสอบ โดยผู้วิจัยพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี ตามขั้นตอนการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ข้างต้น จึงทำให้ระบบการทดสอบมีคุณภาพในทุกด้านไม่ว่าจะเป็นด้านความมีประโยชน์ ความเป็นไปได้ในการนำไปใช้ ความเหมาะสม และความถูกต้อง อีกทั้งยังทำให้นิสิตพึงพอใจมากต่อระบบการทดสอบในด้านภาพรวมของระบบ หน้าจอของระบบ การเรียนรู้ของผู้ใช้งานและความสามารถของระบบ แต่ค่อนข้างพึงพอใจต่อระบบการทดสอบด้านคำศัพท์และสารสนเทศระบบ ซึ่งน้อยกว่าด้านอื่น เมื่อพิจารณาข้อคำถามที่วัดความพึงพอใจต่อระบบการทดสอบด้านคำศัพท์และสารสนเทศระบบ พบว่า ข้อคำถามเกี่ยวกับข้อความแจ้งเตือนความผิดพลาดมีจำนวนนิสิตนักศึกษาที่ไม่พึงพอใจมากกว่าคำถามข้ออื่น อาจเนื่องมาจากระบบการทดสอบไม่มีการแจ้งเตือนความผิดพลาดเกี่ยวกับการใช้งานระบบการทดสอบ เช่น การขาดหายของสัญญาณอินเทอร์เน็ต หรืออินเทอร์เน็ตมีความเร็วต่ำทำให้การประมวลผลของระบบการทดสอบเกิดความล่าช้า จึงทำให้นิสิตนักศึกษารอผลการทดสอบนานกว่าที่ควรจะเป็น ดังนั้น การวิจัยครั้งต่อไปควรมีการเพิ่มเติมในรายละเอียดดังกล่าวเพื่อทำให้นิสิตนักศึกษามีความพึงพอใจต่อระบบการทดสอบเพิ่มมากขึ้น

ข้อเสนอแนะ

1. ข้อเสนอแนะสำหรับการนำผลวิจัยไปใช้

1.1 การเตรียมความพร้อมนิสิตนักศึกษาให้สามารถเรียนรู้ในสังคมที่ต้องใช้เทคโนโลยีสารสนเทศและการสื่อสารในการพัฒนาประเทศมีความสำคัญอย่างยิ่ง นิสิตนักศึกษาจะต้องใช้ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ในการศึกษาค้นคว้า วิจัย และพัฒนาองค์ความรู้ใหม่เพื่อพัฒนาตนเองและประเทศให้ก้าวทันกระแสการเปลี่ยนแปลงของโลก และเพื่อให้สอดคล้องกับความต้องการบุคลากรของหน่วยงานต่างๆ ที่ต้องการบุคลากรที่มีการรู้เท่าทันเทคโนโลยีสารสนเทศและการสื่อสาร รวมทั้งมีความสามารถที่จะคิดวิเคราะห์เกี่ยวกับข้อมูลสารสนเทศและการสื่อสาร ดังนั้น ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในงานวิจัยนี้ สถาบันการศึกษาในระดับอุดมศึกษาสามารถนำไปเป็นเครื่องมือในการประเมินความรู้ความสามารถดังกล่าวได้เป็นอย่างดี ผลการประเมินที่ได้จะช่วยให้สถาบันการศึกษาในระดับอุดมศึกษานำไปพัฒนาผู้เรียน และจัดการเรียนการสอนให้สอดคล้องกับการเปลี่ยนแปลงของกระแสโลกในศตวรรษที่ 21

1.2 จากผลการศึกษาพบว่า ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรี มีการประเมินคุณภาพระดับมากที่สุดทุกด้าน ประกอบกับนิสิตนักศึกษาได้ทดลองใช้ระบบฯและมีความพึงพอใจต่อระบบการทดสอบทุกด้าน จึงสามารถนำระบบฯไปใช้จริงได้ ดังนั้นอาจารย์ในระดับอุดมศึกษาหรือสถาบันอุดมศึกษาที่ต้องการพัฒนาศักยภาพของนิสิตนักศึกษาในด้านนี้ สามารถนำระบบที่พัฒนาไปใช้หรือใช้เป็นแนวทางในการพัฒนาต่อยอด โดยศึกษาจากการใช้คู่มือของระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ประกอบด้วย 2 ส่วน คือ 1) คู่มือการใช้งานสำหรับผู้ใช้งานทั่วไป (user) ประกอบด้วย การลงทะเบียน การเข้าใช้งานระบบครั้งแรก การเข้าสู่ระบบเพื่อเริ่มทำแบบทดสอบ และการรายงานผลการทดสอบ และ 2) คู่มือการใช้งานสำหรับผู้จัดการระบบ (admin) ประกอบด้วย การเข้าสู่ระบบ การจัดการข้อสอบ การแก้ไขข้อสอบ การสร้างชุดข้อสอบ และการจัดการสมาชิก และการดูรายงานผลการทดสอบของสมาชิก

1.3 การออกแบบระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในงานวิจัยนี้ต้องใช้สัญญาณอินเทอร์เน็ต และไม่จำเป็นที่ผู้สอบต้องทำแบบสอบพร้อมๆกันหรือในช่วงเวลาเดียวกันอันเป็นจุดเด่นของการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แต่ที่ต้องให้ความสำคัญ คือ การเตรียมคอมพิวเตอร์ในสถานที่ใช้สอบจะต้องมีสัญญาณอินเทอร์เน็ตที่มีความเร็วเพียงพอ ถ้าสัญญาณอินเทอร์เน็ตมีความเร็วต่ำหรือสัญญาณอินเตอร์เน็ตหลุด อาจทำให้การประมวลผลของระบบการทดสอบเกิดความล่าช้า ทำให้นิสิตนักศึกษาใช้เวลาในการทดสอบนานกว่าที่ควรจะเป็น ดังนั้นจึงควรให้ความสำคัญกับเรื่องดังกล่าวในการใช้ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ฯที่พัฒนานี้

1.4 ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีที่พัฒนาขึ้นในการวิจัยครั้งนี้เหมาะสำหรับนำไปประยุกต์ใช้กับนิสิตนักศึกษาที่กำลังศึกษาอยู่ในหลักสูตรที่ใช้ภาษาไทย เนื่องจากนิสิตและนักศึกษาในหลักสูตรนานาชาติจะไม่คุ้นเคยกับการใช้คำในรูปแบบของภาษาไทย ดังนั้นในการนำไปใช้จึงควรคำนึงถึงตัวอย่างที่สามารถนำระบบทดสอบที่พัฒนาขึ้นไปเพื่อให้ผลการทดสอบปราศจากความลำเอียง (bias) ที่เกิดขึ้นจากการใช้ภาษา

1.5 ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ของนิสิตนักศึกษาปริญญาตรีที่พัฒนาขึ้นในการวิจัยครั้งนี้ทำให้ได้นวัตกรรมที่เป็นแนวทางสำหรับผู้ที่สนใจพัฒนาระบบการทดสอบแบบปรับเหมาะสามารถนำไปใช้ในการพัฒนาข้อสอบจากองค์ประกอบที่สังเคราะห์ การเก็บข้อมูลที่ใช้ข้อสอบจำนวนมากเพื่อจัดทำคลังข้อสอบ การวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบที่มีการตรวจสอบโมเดลที่เหมาะสมและการปรับเทียบโดยใช้ข้อสอบร่วมในแบบสอบแต่ละชุด เพื่อเพิ่มประสิทธิภาพในการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ

2. ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

2.1 การวิจัยเรื่องนี้สร้างแบบวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารจาก 5 องค์ประกอบ ไม่มีองค์ประกอบการสร้างสรรค์ ทั้งนี้เป็นเพราะว่า ข้อสอบวัดการสร้างสรรค์เทคโนโลยีสารสนเทศ ซึ่งเป็นตัวชี้วัดหนึ่งในทักษะทางเทคโนโลยีสารสนเทศ ควรเป็นข้อสอบความเรียงไม่จำกัดคำตอบ แต่ด้วยข้อจำกัดของงานวิจัยที่พัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ต้องมีคำตอบถูกที่แน่นอนและมีการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบไว้ล่วงหน้าในการคัดเลือกข้อสอบให้ผู้สอบทำ จึงไม่สามารถที่จะใช้ข้อสอบความเรียงแบบไม่จำกัดคำตอบเพื่อวัดความคิดสร้างสรรค์ในระบบที่พัฒนาขึ้น ดังนั้นการวิจัยครั้งต่อไปจึงควรมีการศึกษาเพิ่มเติมในตัวชี้วัดที่เกี่ยวข้องกับการสร้างสรรค์เทคโนโลยีสารสนเทศ โดยใช้ข้อสอบในรูปแบบที่หลากหลายและต้องประยุกต์ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ภายใต้ทฤษฎีการตอบสนองข้อสอบที่มีการให้คะแนนหลายค่า (polytomous scoring) เพื่อให้สามารถวัดทักษะทางเทคโนโลยีและสารสนเทศได้อย่างครบองค์ประกอบและครอบคลุมมากขึ้น

2.2 ผลการวิจัยครั้งนี้แสดงให้เห็นว่า คะแนนคำตอบสอดคล้องกลมกลืนกับโครงสร้างที่มีเพียงองค์ประกอบเดียว ซึ่งไม่สัมพันธ์กับผลการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง ดังนั้น จึงควรมีงานวิจัยที่เน้นการศึกษาองค์ประกอบที่วัดทักษะทางเทคโนโลยีและสารสนเทศ หรืออาจยุบรวมบางองค์ประกอบเข้าด้วยกัน เพื่อปรับปรุงให้ความเป็นพหุมิติของข้อสอบชัดเจนมากขึ้น โดยมีวัตถุประสงค์หลักเพื่อไม่ให้สารสนเทศของแต่ละองค์ประกอบสูญหายไป อันเนื่องมาจากการระบุลักษณะจำเพาะของโมเดลการวัดคลาดเคลื่อน (model misspecification) นอกจากนี้ การยุบรวมองค์ประกอบเข้าด้วยกัน ยังเป็นการตรวจสอบและปรับปรุงผลการวัดให้มีความตรงเชิงโครงสร้างสูงขึ้นอีกด้วย

2.3 การวิจัยเรื่องนี้มีจุดประสงค์สำคัญ คือพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จึงไม่ได้เน้นการศึกษาตัวแปรอิสระที่ทำให้เกิดความแตกต่างของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ดังนั้นการที่นิสิตนักศึกษามีความรู้ความสามารถในด้านนี้ที่แตกต่างกัน นอกจากตัวแปรเพศแล้ว อาจขึ้นอยู่กับระดับของตัวแปรอื่น เช่น ฐานะทางเศรษฐกิจ สายการเรียน เจตคติต่อเทคโนโลยีสารสนเทศและการสื่อสาร ความสามารถในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร การควบคุมตนเองเกี่ยวกับการใช้เทคโนโลยีสารสนเทศและการสื่อสาร และความต้องการรับรู้ข้อมูล จึงควรมีการศึกษา ตัวแปรปัจจัยต่างๆเหล่านี้ และควรมีการพัฒนาเกณฑ์การตัดสินระดับคุณภาพของทักษะด้านนี้ โดยอาจมีการพัฒนาเกณฑ์จุดตัดของทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่21ในการวิจัยครั้งต่อไป

2.4 การวิจัยในอนาคตผู้วิจัยควรให้ความสำคัญกับกลุ่มตัวอย่างที่อยู่ในหลักสูตรนานาชาติที่ไม่ได้มีการใช้ภาษาไทยในการจัดการเรียนการสอน เนื่องจากนิสิตนักศึกษาปริญญาตรีในปัจจุบันมีความหลากหลายไม่ว่าจะเป็นในระดับสาขาวิชา คณะ มหาวิทยาลัย และภูมิภาคตลอดจนความหลากหลายของหลักสูตรที่มีทั้ง

หลักสูตรภาษาไทยและภาษาต่างประเทศ ดังนั้นเพื่อให้การวัดทักษะทางเทคโนโลยีสารสนเทศเป็นไปอย่างทันสมัยและครอบคลุมในอนาคตจึงควรมีการออกแบบแบบวัดที่สามารถใช้ได้ทั้งใน 2 ภาษา คือแบบภาษาไทยและภาษาอังกฤษเพื่อไม่ให้ภาษาเป็นอุปสรรคในการทดสอบจึงควรมีการพัฒนาแบบวัดที่มีมาตรฐานและสามารถนำไปใช้กับกลุ่มตัวอย่างได้หลากหลายมากยิ่งขึ้นในอนาคต

รายการอ้างอิง

ภาษาไทย

กิดานันท์ มลิทอง. (2548). *เทคโนโลยีและการสื่อสารเพื่อการศึกษา*. กรุงเทพมหานคร: ห้างหุ้นส่วนจำกัด อรุณการพิมพ์.

ช่อบุญ จิรานุภาพ. (2554). *การพัฒนาตัวบ่งชี้ความสำเร็จการใช้เทคโนโลยีสารสนเทศและการสื่อสารในการเรียนการสอนเพื่อเสริมสร้าง ทักษะการเรียนรู้เทคโนโลยีสารสนเทศและการสื่อสารของนักเรียนมัธยมศึกษา* (ปริญญานิพนธ์ปริญญาดุษฎีบัณฑิต). กรุงเทพมหานคร: จุฬาลงกรณ์มหาวิทยาลัย.

ญานิศรา มุนินทร์สาคร. (2558). *การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6* (ปริญญานิพนธ์ปริญญามหาบัณฑิต). ชลบุรี: มหาวิทยาลัยบูรพา.

ณภัทร ชัยมงคล, โชติกา ภาษีผล และศิริชัย กาญจนวาสี. (2558). *การทดสอบแบบปรับเหมาะ* (ปริญญานิพนธ์ปริญญาดุษฎีบัณฑิต). กรุงเทพมหานคร: จุฬาลงกรณ์มหาวิทยาลัย.

ณภัทร ชัยมงคล โชติกา ภาษีผล และ ศิริชัย กาญจนวาสี. (2559). การทดสอบแบบปรับเหมาะที่มีการสะท้อนข้อมูลย้อนกลับในการทดสอบมาตรฐานวิชาชีพไอที. *วารสารเทคโนโลยีสารสนเทศ, 12*(2). 58-64.

ดิเรก หอมจันทร์. (2554). *การพัฒนาแบบทดสอบปรับเหมาะกับความสามารถของผู้สอบสำหรับการเรียนการสอนแบบอีเลิร์นนิ่ง (E-learning) รายวิชา 4000107: เทคโนโลยีสารสนเทศเพื่อชีวิต สำหรับนักศึกษาชั้นปีที่ 1 มหาวิทยาลัยราชภัฏสุรินทร์* (ปริญญานิพนธ์ปริญญามหาบัณฑิต). กรุงเทพมหานคร: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

ทัศน์ศิรินทร์ สว่างบุญ. (2554). *การเปรียบเทียบวิธีการคัดเลือกข้อสอบขั้นแรกและลำดับข้อสอบที่มีต่อคุณภาพของการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์* (ปริญญานิพนธ์ปริญญาดุษฎีบัณฑิต). กรุงเทพมหานคร: จุฬาลงกรณ์มหาวิทยาลัย.

ทิพย์ ขำอยู่, เสรี ชัดแช้ม, และ กฤษณะ ชินสาร. (2556). การวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยประยุกต์โมเดลลำดับขั้นคุณลักษณะและการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์. *วิทยาการวิจัยและวิทยาการปัญญา, 10*(2), 54–69.

น้ำทิพย์ วิภาวิน. (2552). การรู้เรื่องการอ่าน การรู้สารสนเทศ และการรู้เท่าทันเทคโนโลยีสารสนเทศ. *วารสารบรรณศาสตร์, 2*(1), 109–123.

บุญชม ศรีสะอาด. (2539). การแปลผลเมื่อใช้เครื่องมือรวบรวมข้อมูลแบบมาตราส่วนประมาณค่า. *วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม, 2*(1), 64–70

ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่* (พิมพ์ครั้งที่ 4). กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สุชาดา กรเพชรปาณี, เสรี ชัดแช้ม, ปิยะทิพย์ ตินวร, และ โสหส สุขานนท์สวัสดิ์. (2557). *การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET*. กรุงเทพมหานคร: สำนักทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน).

สุชาดา สกลกิจรุ่งโรจน์, เสรี ชัดแช้ม, และ สมพร สุทัศนีย์. (2558). การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย. *วิทยาการวิจัยและวิทยาการปัญญา, 13*(1), 1–17.

สุดารัตน์ หวลมุกดา. (2550). *ประสิทธิภาพของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ตามระดับชั้นของค่าอำนาจจำแนกภายใต้เงื่อนไข ต่างกันโดยใช้วิธีการจำลองข้อมูล* (ปริญญานิพนธ์ปริญญาดุษฎีบัณฑิต). กรุงเทพมหานคร: มหาวิทยาลัยศรีนครินทรวิโรฒ.

สำนักงานคณะกรรมการการอุดมศึกษา. (2561). *นักศึกษารวม ปีการศึกษา 2560*. Retrieved from http://www.info.mua.go.th/information/

สุนทร เทียนงาม. (2551). *ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบค่าความสามารถของผู้สอบและสารสนเทศของแบบสอบเมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน* (ปริญญานิพนธ์ปริญญาดุษฎีบัณฑิต). กรุงเทพมหานคร: จุฬาลงกรณ์มหาวิทยาลัย.

สุนันทา ศิริเบญจา, ไชยรัตน์ ปราณี, และ ดวงใจ สีเขียว. (2556). การพัฒนาแบบทดสอบปรับเหมาะโดยใช้คอมพิวเตอร์ สาระเทคโนโลยีสารสนเทศและการสื่อสาร โดยประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ สาหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3. *วารสารวิชาการและวิจัยสังคมศาสตร์, 8*(22), 87–102.

ภาษาอังกฤษ

Agbatogun, A. (2010). Gender, academic qualification and subject discipline differentials of Nigerian teachers' ICT literacy. *Academic Leadership: The Online Journal, 8*(1), Article 18.

Ahmad, M., Karim, A. A., Din, R., & Albakri, I. S. M. A. (2013). Assessing ICT competencies among postgraduate students based on the 21st century ICT competency model. *Asian Social Science, 9*(16), 32–39.

Ali, U. S., & Chang, H. -H. (2014). *An item-driven adaptive design for calibrating pretest items* (Research Report No. RR-14-38). Princeton, NJ: Educational Testing Service. doi:10.1002/ets2.12044

Allen, M. J., & Yen, W. M. (1979). *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transaction on Automatic Control, 19*(6), 716–723.

Anatchkova, M. D., Saris-Baglama, R. N., Kosinski, M., & Bjorner, J. (2009). Development and Preliminary

Testing of a Computerized Adaptive Assessment of Chronic Pain. *The Journal of Pain: Official Journal of the American Pain Society, 10*(9), 932–943.

Baker, F. B., & Kim, S. -H. (2017). *The basics of item response theory using R*. New York: Springer.

Bellanca, J., & Brandt, R. (2010). *21st Century Skills: Rethinking How Students Learn*. Bloomington, IN: Solution Tree.

Barrada et al. (2010). A method for the comparison of item selection rules in computerized adaptive testing. *Applied Psychological Measurement, 34*(6), 438–452.

Bergstrom, B. A., & Lunz, M. E. (1999). CAT for certification and licensure. In F. Drasgow & J. B. Olson-Buchanan (Eds.), *Innovations in computerized assessment* (pp. 67–91). Mahwah, NJ: Lawrence Erlbaum Associates.

Birdsall, M. (2011). *Implementing computer adaptive testing to improve achievement opportunities*. Retrieved from https://assets.publishing.service.gov.uk/government /uploads/system/uploads/attachment_data/file/606023/0411_MichaelBirdsall_impleme nting-computer-testing-_Final_April_2011_With_Copyright.pdf

Bock, R. D., & Mislevy, R. J. (1982). Adaptive EAP estimaion of ability in a microcomputer environment. *Applied Psychological Measurement, 6*(4), 431–444.

Brown, C., Templin, J., & Cohen, A. (2015). Comparing the two- and three-parameter logistic models via likelihood ratio tests: A commonly misunderstood problem. *Applied Psychological measurement, 39*(5), 335–348.

California Emerging Technology Fund. (2008). *California ICT digital literacy assessments and curriculum framework.* Retrieved from http://www.ictliteracy.info/rf.pdf/California%20ICT%20Assessments%20Curriculum%20Framework.pdf.

Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software, 48*(6), 1–29. doi: http://dx.doi.org/10.18637/jss.v048.i06

Chang, H, -H., & Ying, Z. (1996). A global information approach to computerized adaptive testing. *Applied Psychological Measurement, 20*(3), 213–229.

Chen, P. H. (2009). Comparison of adaptive Bayesian estimation and weighted Bayesian estimation in multidimensional computerized adaptive testing. In D. J. Weiss (Ed.), *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing*.

Chen, S. -Y., & Doong, S. H. (2008). Predicting item exposure parameters in computerized adaptive testing. *Br J Math Stat Psychol, 61*(Pt 1), 75–91.doi:10.1348/000711006X 129553.

Cheng, Y., Patton, J. M., & Shao, C. (2014). a-Stratified computerized adaptive testing in the presence of calibration error. *Educational and Psychological Measurement*. *75*(2), 260–283.

Chien, T. W., Wang, W. C., Huang, S. Y., Lai, W. P., & Chow, J. C. (2011). A Web-based computerized adaptive testing (CAT) to assess patient perception in hospitalization. *Journal of medical internet research*, *13*(3), 1–10.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). *Development of a tool measuring user satisfaction of the human-computer interface.* Retrieved from https://www.researchgate.net/publication/248594191_Development_of_an_Instrument_Measuring_User_Satisfaction_of_the_Human-Computer_Interface

Choi, S. W. (2018). Firestar. In W. J. van der Linden (Ed.), *Handbook of item response theory, Volume 3: Applications* (pp. 547–556). Boca Raton, FL: CRC Press.

Čisar, S. M., Radosav, D., Markoski, B., Pinter, R. & Čisar, P. (2010, November). Computer adaptive testing for student's knowledge in C++ exam. Paper presented at 2010 11th International Symposium on Computational Intelligence and Informatics (CINTI).

Claro et al. (2012). Assessment of 21st century ICT skills in Chile: Test design and results from high school level students. *Computers & Education, 59*(3), 1042–1053.

Davis, L. L., & Dodd, B. G. (2003). Item exposure constraints for testlets in the verbal reasoning section of the MCAT. *Applied Psychological Measurement*, *27*, 335–356.

Drasgow, F. (1989). An evaluation of marginal maximum likelihood estimation for the two-parameter logistic model. *Applied Psychological Measurement, 13*(1), 77–90.

de Ayala, R. J. (2009). *Methodology in the social sciences. The theory and practice of item response theory.* New York: Guilford Press.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, *323*(5910), 66–69.

Deng, H., Ansley, T., & Chang, H. -H. (2010). Stratified and maximum information item selection procedures in computer adaptive testing. *Journal of Educational Measurement*, *47*(2), 202–226.

DeMars, C. (2002). Incomplete data and item parameter estimates under JMLE and MML estimation. *Applied Measurement in Education*, 15(1), 15–31.

Diao, Q., and Reckase, M. (2009). Comparison of ability estimation and item selection methods in multidimensional computerized adaptive testing. In D. J. Weiss (Ed.), *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing*.

Economides, A. A. & Roupas, C. (2009). Evaluating computerized adaptive testing systems. In N. Karacapilidis (Ed.), *Solutions and Innovations in Web-Based Technologies for Augmented Learning: Improved Platforms, Tools and Application* (pp. 185–201). Hershey, PA: IGI-Global.

Educational Testing Service. (2002). *Digital transformation a framework for ICT literacy*. Retrieved from www1.ets.org/Media/Tests/Information_and_Communication_ Technology_Literacy/ictreport.pdf.

Educational Testing Service. (2016). *Guide to the use of scores*. Retrieved from https://www.ets.org/s/gre/pdf/gre_guide.pdf.

Eggen, T. J. H. M. (2001). Overexposure and underexposure of item in computerized adaptive testing. *Measurement and Research Department Report*, *1*, 1–21.

Elliott, G. (2004). *Global Business Information Technology: an integrated systems approach*. London: Pearson Education.

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

Frey, A., & Seitz, N., -N. (2009). Multidimensional Adaptive Testing in Educational and Psychological Measurement: Current State and Future Challenges. *Studies in Educational Evaluation*, *35*(2), 89–94.

Georgiadou, E., Triantafillou, E., & Economides, A. (2006). Evaluation parameters for computer-adaptive testing. *BJET, 37*, 261–278.

Gibbons et al. (2012). Development of a computerized adaptive test for depression. *Archives of General Psychiatry*, *69*(11), 1104–1112.

González, J., & Wiberg, M. (2017). *Applying test equating methods using R*. New York: Springer.

Haberman, S. J. (1978). *Analysis of qualitative data*. Cambridge, MA: Academic Press.

Haberman, S. J. (2008). When can subscores have value? *Journal of Educational and Behavioral Statistics, 33*(2), 204–229.

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, CA: Sage.

Han, K. T. (2012). SimulCAT: Windows software for simulating computerized adaptive test administration. *Applied Psychological Measurement*, *36*(1), 64–66.

Han, K. T. (2013). MSTGen: Simulated data generator for multistage testing. *Applied Psychological Measurement, 37*(8), 666–668.

Hanson, B. A., & Beguin, A. A. (2002). Obtaining a common scale for item response theory item parameters using separate versus concurrent estimation in the common item equating design. *Applied Psychological Measurement, 61*(1), 3-24.

He, W., Diao, Q., & Hauser, C. (2014). A comparison of four item-selection methods for severely constraint CATs. *Educational and Psychological Measurement*, *74*(4), 677–696.

Ho, T. -H., & Dodd, B. G. (2012). Item selection and ability estimation procedures for a mixed-format adaptive test. *Applied Measurement in Education, 25*(4), 305–326.

Holland, P. W., & Rubin, D. B. (1982). *Test equating*. Cambridge, MA: Academic Press.

Huang, H. -Y., Wang, W. -C., Chen, P, -H., & Su, C. -M. (2013). Higher-order item response models for hierarchical latent traits. *Applied Psychological Measurement*, *37*(8), 619–637.

Joint Committee on Standards for Educational Evaluations. (1994). *The program evaluation standards: How to assess evaluations of educational programs*. Newbury Park, CA: Sage.

Katz, I. R., & Macklin, A. S. (2007). *Information and communication technology (ICT) literacy: Integration and assessment in higher education*. Retrieved from http://www.iiisci.org/Journal/CV$/sci/pdfs/P890541.pdf.

Kim, J. (2006). Using the distractors categories of multiple-choice items to improve IRT linking. *Journal of Educational Measurement, 43*(3), 193–213.

Kim, S. H., & Cohen, A. S. (1998). A comparison of linking and concurrent calibration under item response theory. *Applied Psychological Measurement, 22*(2), 131–143.

Kim, S. H., & Cohen, A. S. (2002). A comparison of linking and concurrent calibration under the grade response model. *Applied Psychological Measurement*, 26(1), 25–41.

Kiss, G., & Gastelú, C. A. T. (2015). Comparison of the ICT literacy level of the Mexican and Hungarian students in the higher education. *Procedia-Social and Behavioral Sciences, 176*, 824–833.

Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices*. New York: Springer.

Lazarsfeld, P. F. (1950). The logical and mathematical foundation of latent structure analysis. *Measurement and Prediction*, *4*, 362–412.

Longworth, N. (2003). *Lifelong Learning in Action: Transforming Education in the 21st Century*. London: Kogan Page.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Mahwah, NJ: Lawrence Erlbaum Associates.

Madigan, E. M., Goodfellow, M., & Stone, J. A. (2007). Gender, perceptions, and reality: technological literacy among first-year students. *ACM SIGCSE Bulletin, 39*(1), 410-414.

Magis, D., Yan, D., & von Davier, A. A. (2017). *Computerized adaptive and multistage testing with R: Using packages catR and mstR*. New York: Springer.

Mao, X., & Xin, T. (2013). The application of the Monte Carlo approach to cognitive diagnostic computerized adaptive testing with content constrains. *Applied Psychological Measurement, 37*(6), 482–496.

Markauskaite, L. (2006). Gender issues in preservice teachers' training: ICT literacy and online learning. *Australasian Journal of Educational Technology, 22*(1), 1–29.

Markauskaite, L. (2007). Exploring the structure of trainee teachers' ICT literacy: the main components of, and relationships between, general cognitive and technical capabilities. *Educational Technology Research and Development, 55*(6), 547–572.

Meijer, R. R., & Nering, M. L. (1999). Computerized adaptive testing: Overview and Introduction. *Applied Psychological Measurement, 23*(3), 187–194.

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

Mulder, J., & van der Linden,W. J. & Pashley, P.J. (2010). Multidimensional Adaptive Testing with Kullback–Leibler Information Item Selection. In W. J. van der Linden & C. A. W. Glas (Eds.). *Elements of adaptive testing*. New York: Springer.

Navas, H., Osornio, A. L., Baum, A., Gomez, A., Luna, D., & de Quiros, F. G. (2007). Creation and Evaluation of a Terminology Server for the Interactive Coding of Discharge Summaries. *MEDINFO, 12*(1), 650–654.

Partnership for 21st Century Skills. (2007). *P21 framework definitions*. Retrieved from http://www.p21.org/storage/documents/P21_Framework_Definitions.pdf.

Raykov, T., & Marcoulides, G. A. (2018). *A course in item response theory and modeling with Stata*. College Station, TX: Stata Press.

Reckase, M. D. (2003, April). *Item pool design for computerized adaptive tests*. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago.

Reckase, M. D. (2009). *Multidimensional item response theory*. New York: Springer-Verlag.

Revicki, D. A., & Reise, S. P. (2015). *Handbook of item response theory modeling: Applications to typical performance assessment*. New York: Routledge.

Ritzhaupt, A. D., Liu, F., Dawson, K., & Barron, A. E. (2013). Differences in student information and communication technology literacy based on socio-economic status, ethnicity, and gender: Evidence of a digital divide in Florida schools. *Journal of Research on Technology in Education, 45*(4), 291–307.

Rizopoulos, D. (2006). ltm: An R package for Latent Variable Modelling and Item Response Theory Analyses. *Journal of Statistical Software, 17*(5), 1–25.

Segall, D. O. (1996). Multidimensional adaptive testing. *Psychometrika. 61*(2), 331–354.

Segall, D. O. (2010). Principles of Multidimensional Adaptive Testing. In W. J. van der Linden & C. A. W. Glas (Eds.). *Elements of adaptive testing*. New York: Springer.

Senkbeil, M. & Ihme, J. M. (2017). Motivational factors predicting ICT literacy: First evidence on the structure of an ICT motivation inventory. *Computers & Education, 108*, 145–158.

Somerville et al. (2008). The ETS iSkillsTM assessment: A digital age tool. *The Electronic Library, 26*(2), 158–171.

Spearman, C. (1927). *The abilities of man*. London: Macmillan.

Susan et al. (2016). Development and validation of computerized adaptive test for PTSD (P-CAT). *Psychiatric Services, 67*(10), 1116–1123.

Sykes, R. C. (1997). Guidelines for the selection of anchor items for mixed (or single) item format tests. Monterrey, CA: CTB/McGraw-Hill.

Tella, A., & Mutula, S. M. (2008). Gender differences in computer literacy among undergraduate students at the University of Botswana: implications for library use. *Malaysian Journal of Library & Information Science*, *13*(1), 59–76.

Thissen, D., & Mislevy, R. J. (1990). Testing algorithms. In H. Wainer, N. J. Dorans, R. Flaugher, B. F. Green, R. J. Mislevy, L. Steinberg & D. Thissen (Eds.), *Computerized adaptive testing: A primer* (pp. 103–135). Hillsdale, NJ: Erlbaum.

Thissen, D., & Wainer, H. (1990). Confident envelopes for item response theory. *Journal of Educational and Behavioral Statistics*, *15*(2), 113–128.

Thomson, N. A., & Weiss, D. J. (2011). A framework for the development of computerized adaptive tests. Practical Assessment, Research & Evaluation, 16(1), 1–9.

Thurstone, L. L. (1947). *Multiple factor analysis: A development and expansion of vectors of the mind*. University of Chicago Press, Chicago.

UNESCO. (2008). *Strategy framework for promoting ICT literacy in the Asia-Pacific region*. Retrieved from http://www2.unescobkk.org/elib/publications/188/promotingICT_literacy.pdf.

Urry, V. (1977). Tailored testing: A successful application of latent trait theory. *Journal of Education Measuarement*, *14*, 181–196.

Vale, D. C. (1986). Linking item parameters onto a common scale. *Applied Psychological Measurement, 10*(4), 333–344.

van der Linden, W. J., & Glas, C. A. W. (2000). *Computerized adaptive testing: Theory and practice*. Dordrechi, The Netherlands: Kluwer Academic.

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

von Davier, A. A., & Mislevy, R. J. (2016). Design and modeling frameworks for 21st-century simulation- and game-based assessments. In C. S. Wells & M. Faulkner-Bond (Eds.), *Educational measurement: from foundations to future* (pp. 239–256). New York: Guilford Press.

Wainer et al. (2000). *Computer adaptive testing: A primer* (2nd ed.). Mahwah, NJ: Erlbaum.

Way, W. D., & Robin, F. (2016). The history of computer-based testing. In C. S. Wells & M. Faulkner-Bond (Eds.), *Educational measurement: from foundations to future* (pp. 185–207). New York: Guilford Press.

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

Werner, C., & Schermelleh-Engel, K. (2010). *Introduction to structural equation modeling with LISREL*. Retrieved from http://www.psychologie.uzh.ch/dam/jcr:ffffffff-b371-2797-0000-00000fda8f29/chisquare_diff_en.pdf.

Wijaya, S. W., & Surendro, K. (2007). *ICT literacy as an indicator of e-government Readiness*. Retrieved from http://academia.edu/1044238/ICT_literacy_as_an_indicator_of_e-government_Readiness.

Wilson, M., & Moore, S. (2012). An explanative modeling approach to measurement of reading comprehension. In J. P. Sabatini, T. O'Reilly & E. R. Albro (Eds.), *Reaching an understanding: Innovations in how we view reading assessment* (pp. 147–168). Lanham, MD: Rowman & Littlefield Education.

Wingersky, M. S., & Lord, F. M. (1984). *An investigation of methods for reducing sampling error in certain IRT procedures*. Retrieved from https://conservancy.umn.edu/bitstream /handle/11299/101951/v08n3p347.pdf?sequence=1

Yan, D., von Davier, A.A., & Lewis, C. (2014). *Computerized Multistage Testing: Theory and Applications*. London: Chapman and Hall.

Zheng, Y., & Chang, H. H. (2015). On-the-fly assembled multistage adaptive testing. *Applied Psychological Measurement, 39*(2), 1–15.

ภาคผนวก

ภาคผนวก 1

รายนามผู้ทรงคุณวุฒิในการพัฒนาข้อสอบและคลังข้อสอบเพื่อวัดทักษะด้านเทคโนโลยีสารสนเทศและการ
สื่อสารแห่งศตวรรษที่ 21

รายนามผู้ทรงคุณวุฒิด้านเทคโนโลยีสารสนเทศ

1. ผู้ช่วยศาสตราจารย์ ดร.ดรัณภพ เพียรจัด	อาจารย์ประจำคณะครุศาสตร์ มหาวิทยาลัยราชภัฏราชนครินทร์
2. อาจารย์ ดร.วารภรณ์ สินถาวร	อาจารย์ประจำคณะครุศาสตร์ มหาวิทยาลัยราชภัฏราชนครินทร์
3. ผู้ช่วยศาสตราจารย์ ดร.ทิพยา จินตโกวิท	อาจารย์ประจำ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
4. อาจารย์ ดร.ศิพาณี ชิตประสิทธิ์ชัย	อาจารย์ประจำ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
5. อาจารย์ ดร.ธนพล เจนสุทธิเวชกุล	อาจารย์ประจำ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
6. ผู้ช่วยศาสตราจารย์ ดร.เอกรัฐ หล่อพิเพียร	อาจารย์ประจำ คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
7. อาจารย์ ดร.ชุติวัฒน์ สุวัตถิพงศ์	อาจารย์ประจำ สำนักเทคโนโลยีการศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช

รายนามผู้ทรงคุณวุฒิด้านการวัดและประเมินผลการศึกษา

1. รองศาสตราจารย์ ดร.กมลวรรณ ตังธนกานนท์	อาจารย์ประจำคณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. ผู้ช่วยศาสตราจารย์ ดร.สังวรณ์ งัดกระโทก	อาจารย์ประจำสาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช
3. อาจารย์ ดร.อนุสรณ์ เกิดศรี	อาจารย์ประจำสำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช
4. รท.นพ.อนุพงษ์ กันธิวงศ์	อาจารย์ประจำวิทยาลัยแพทยศาสตร์ พระมงกุฎเกล้า
5. ผู้ช่วยศาสตราจารย์ ดร.เรืองเดช ศิริกิจ	อาจารย์ประจำสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ
6. ผู้ช่วยศาสตราจารย์ ดร.เมษา นวลศรี	อาจารย์ประจำคณะศึกษาศาสตร์ มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ฯ
7. ผู้ช่วยศาสตราจารย์ ดร.ฉัตรศิริ ปิยะพิมลสิทธิ์	อาจารย์ประจำคณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
8. อาจารย์ ดร. อัญชลี สุขในสิทธิ์	อาจารย์ประจำคณะสังคมศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

ภาคผนวก 2

รายนามผู้ทรงคุณวุฒิในการพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

รายนามผู้ทรงคุณวุฒิ

1. รองศาสตราจารย์ ดร.กมลวรรณ ตังธนกานนท์	อาจารย์ประจำคณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. ผู้ช่วยศาสตราจารย์ ดร.เมษา นวลศรี	อาจารย์ประจำคณะศึกษาศาสตร์ มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ฯ
3. อาจารย์ ดร.สุกัญญา บุญศรี	อาจารย์ประจำคณะครุศาสตร์ มหาวิทยาลัยราชมงคลธัญญบุรี

ภาคผนวก 3

รายนามผู้ทรงคุณวุฒิในการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

รายนามผู้ทรงคุณวุฒิ

1. อาจารย์ ดร.อนุสรณ์ เกิดศรี	อาจารย์ประจำสำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช
2. อาจารย์ ดร.ชุติวัฒน์ สุวัตถิพงศ์	อาจารย์ประจำ สำนักเทคโนโลยีการศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช
3. อาจารย์ ดร.กฤตยา ทองผาสุก	อาจารย์ประจำ คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ภาคผนวก 4

ตัวอย่างข้อสอบและกระดาษคำตอบแบบฝน

1. เมื่อท่านต้องการทำรายงานวิจัย โดยมีคำถามวิจัย คือ "ในการดูแลสุขภาพจิต การนั่งสมาธิ หรือ การเดินจงกลมจะส่งผลต่อสุขภาพจิตได้ดีกว่ากัน" ท่านจะค้นหาข้อมูลโดยใช้คำต่าง ๆดังนี้ <u>ยกเว้น</u> ข้อใดต่อไปนี้

1. การดูแลสุขภาพจิต
2. การนั่งสมาธิ
3. การเดินจงกลม
4. ส่งผลต่อสุขภาพจิต
5*. ได้ดีกว่ากัน

2.หากท่านได้รับมอบหมายจากอาจารย์ให้นำวีดิทัศน์ (video) ไฟล์เสียง และรูปภาพจากอินเทอร์เน็ตมาตัดต่อเป็นภาพยนตร์สำหรับการรณรงค์เกี่ยวกับการใช้เทคโนโลยีสารสนเทศอย่างมีวิจารณญาณบนเว็บไซต์ของมหาวิทยาลัย ท่านควรเลือกใช้วีดิทัศน์ ไฟล์เสียง และรูปภาพที่มีเครื่องหมายตามข้อใดจึงจะถือว่าไม่เป็นการละเมิดลิขสิทธิ์

1. TM
2. Ⓟ
3. ©
4*. 🅭
5. ®

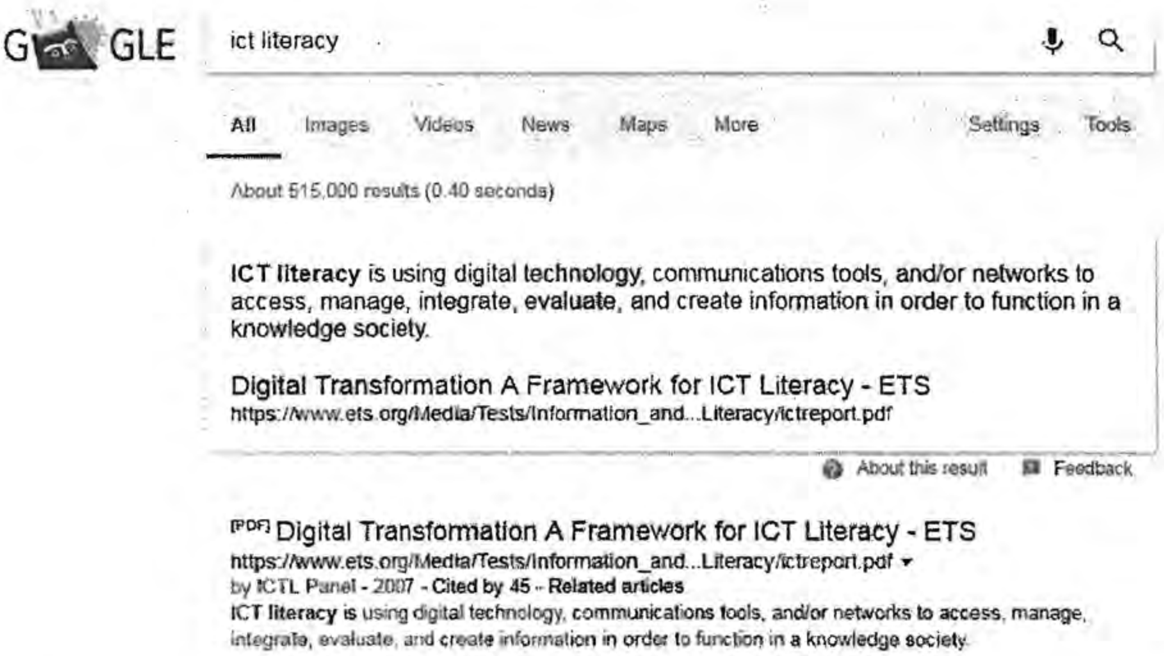
3.เว็บไซต์ซึ่งมีการตั้งค่าการใช้งาน bookmark มีความแตกต่างจากการใช้งาน extensions อย่างไร

1. bookmarks ใช้เพื่อเรียกโปรแกรมเสริมเหมือนกับ extensions
2. bookmarks ใช้เพื่อเรียกใช้โปรแกรมเสริม extensions ใช้เพื่อเรียกเปิดหน้าเว็บไซต์ใหม่
3. bookmarks ใช้เพื่อเรียกใช้โปรแกรมเสริม extensions ใช้เพื่อเรียกใช้งานเว็บไซต์ซึ่งเคยบันทึกไว้
4*. extensions ใช้เพื่อเรียกใช้โปรแกรมเสริม bookmarks ใช้เพื่อเรียกใช้งานเว็บไซต์ซึ่งเคยบันทึกไว้
5. extensions ใช้เพื่อเรียกเปิดหน้าเว็บไซต์ใหม่ bookmarks ใช้เพื่อเรียกใช้งานเว็บไซต์ซึ่งเคยบันทึกไว้

4.วิธีการใดเป็นวิธีการสืบค้นไฟล์รูปภาพจาก Google ที่ถูกต้อง และสามารถนำรูปภาพมาใช้ประกอบการจัดทำรายงานได้อย่างถูกต้องตามกฎหมาย

A. ค้นรูป ⇨ การตั้งค่า ⇨ การตั้งค่าค้นหา ⇨ ความช่วยเหลือ ⇨ ค้นหารูปภาพ
B. ค้นรูป ⇨ การตั้งค่า ⇨ การตั้งค่าค้นหา ⇨ ความช่วยเหลือ ⇨ นำรูปภาพออกจาก Google
C. ค้นรูป ⇨ การตั้งค่า ⇨ การตั้งค่าค้นหา ⇨ ค้นหารูปภาพ ⇨ ค้นหาภาพที่นำใช้ได้ฟรี
D. ค้นรูป ⇨ เครื่องมือ ⇨ สิทธิ์ในการใช้งาน ⇨ ติดป้ายกำกับว่าสามารถใช้ซ้ำ

1. A หรือ C
2. B หรือ C
3. A หรือ B
4. B หรือ D
5*. C หรือ D



5. จากภาพ เมื่อต้องการค้นหาหนังสือที่เกี่ยวข้องกับ ICT Literacy ควรดำเนินการอย่างไรจึงมีความเหมาะสมและรวดเร็วที่สุด

1. คลิก Images เลือก Book
2. คลิก Videos เลือก Book
3. คลิก News เลือก Book
4. คลิก Maps เลือก Book
5*. คลิก More เลือก Book

6.ปัจจุบันหลายเว็บไซต์มีการใช้งานโปรโตคอล (protocol) https แทน http เพิ่มมากขึ้น ซึ่งโปรโตคอลดังกล่าวจะปรากฏอยู่หน้า url เว็บไซต์ เช่น https://google.co.th ข้อใดกล่าวถึงความแตกต่างของการใช้งานโปรโตคอลทั้งสองรูปแบบได้ถูกต้องที่สุด

1. โปรโตคอล https มีความสะดวกในการใช้งานมากกว่าโปรโตคอล http
2. โปรโตคอล https มีการเพิ่มความเร็วในการใช้งานเว็บไซต์ให้รวดเร็วมากยิ่งขึ้น
3*. โปรโตคอล https มีการป้องกันความปลอดภัยสำหรับการเข้าถึงข้อมูลมากยิ่งขึ้น
4. โปรโตคอล http มีความซับซ้อนมากจึงมีการพัฒนาโปรโตคอล https ขึ้นมาทดแทน
5. โปรโตคอล http มีการเข้าถึงเว็บไซต์ผิดพลาดเป็นจำนวนมาก จึงมีการพัฒนาโปรโตคอล https ขึ้นมาทดแทน

7.กำหนดให้ Array หนึ่งเก็บข้อมูลเมทริกซ์จำนวน 3 เมทริกซ์ โดยแต่ละเมทริกซ์มีขนาด 2 x 2 ดังรูป และต้องเรียกคืนค่า 0.6 และ 0.8 ต้องกำหนดตำแหน่งไฟล์อย่างไร

$\begin{bmatrix} 0.2 & 0.3 \\ 0.4 & 0.5 \end{bmatrix}$ เมทริกซ์ที่ 1

$\begin{bmatrix} 0.1 & 0.0 \\ 0.9 & 1.5 \end{bmatrix}$ เมทริกซ์ที่ 2

$\begin{bmatrix} 0.2 & 0.5 \\ 0.6 & 0.8 \end{bmatrix}$ เมทริกซ์ที่ 3

1*. [2, ,3]   2. [,2,3]   3. [3,2,]   4. [3, ,2]   5. [3,2,2]

8.วินัยเป็นเจ้าหน้าที่ฝ่ายการตลาด เขาต้องนำเสนอแผนการขายสินค้าให้กับผู้บริหาร โปรแกรมในข้อใดต่อไปนี้เหมาะสมกับการนำเสนอข้อมูลของวินัยมากที่สุด

1. Microsoft Access   2. Microsoft Excel   3. Microsoft Word
4*. Microsoft PowerPoint   5. Microsoft Visio

9.วิธีการจัดเก็บรวบรวมสารสนเทศใด<u>ไม่เหมาะสม</u>ในอนาคต

1.* จัดเก็บในแผ่น CD-ROM
2. จัดเก็บในเครื่องคอมพิวเตอร์
3.จัดเก็บในไฟล์ข้อมูลอิเล็กทรอนิกส์ เช่น Excel
4. จัดเก็บในฮาร์ดดิกส์แบบพกพา(External Hard disk)
5. จัดเก็บในฐานข้อมูลในระบบคลาวด์ (Cloud computing)

10.ขนิษฐาทำงานมา 7 ปี องค์กรไม่มีระบบการจัดการเอกสารและนโยบายการบันทึกเอกสาร ดังนั้น ขนิษฐาควรจัดการเอกสารลักษณะใดที่มีความเหมาะสม เพื่อให้คนอื่นสามารถทำงานที่เกี่ยวข้องต่อได้ หากมีกรณีด่วนที่ไม่สามารถติดต่อตนได้

1.จัดเก็บเอกสารไว้ในแฟลชไดร์ฟของตนพกไว้ในกระเป๋าเสมอ โดยแยกแฟลชไดร์ฟละ 1 งาน
2.จัดเก็บและบันทึกเอกสารเหมือนปกติที่เคยปฏิบัติ เพราะเป็นรูปแบบที่ใคร ๆ ก็สามารถเข้าใจได้
3.จัดเก็บเอกสารไว้บนเดสก์ทอปเพื่อความสะดวกในการใช้งาน และบันทึกชื่อเอกสารตามคำแนะนำของโปรแกรม
4.จัดเก็บเอกสารไว้บนคลาวด์ (iCloud) โดยการแยกโฟลเดอร์ตามชื่องานและบันทึกชื่อเอกสารตามวัตถุประสงค์ของเอกสาร
5*.จัดเก็บรวมกันไว้ในพาร์ติชันอื่นที่ไม่ใช่พาร์ติชันระบบปฏิบัติการ โดยการแยกโฟลเดอร์ตามชื่องานและบันทึกชื่อเอกสารตามหัวข้อหลัก-หัวข้อย่อยของเอกสาร

11. ข้อใดเรียงลำดับวิธีการในการจัดโครงสร้างการจัดเก็บข้อมูลสารสนเทศให้ปลอดภัยได้อย่างเหมาะสม

A: การจัดแบ่งโฟลเดอร์

B: การตั้งชื่อไฟล์และโฟลเดอร์

C: การตั้งรหัสและกำหนดสิทธิ์การเข้าถึงข้อมูลในโฟลเดอร์

D: การแยกประเภทไฟล์ข้อมูล

E: การจัดเรียงโครงสร้างความสัมพันธ์และระดับชั้นของโฟลเดอร์

1. E, A, B, C, D
2. C, E, D, B, A
3. A, C, E, B, D
4*. B, D, A, E, C
5. C, A, E, D, B

12. หากท่านต้องการจัดเก็บ/บันทึกเอกสารเดิมเป็นเอกสารอื่นที่มีชื่อแตกต่างกันโดยเอกสารนี้เป็นเอกสารที่เกี่ยวกับ ตารางคำนวณ (spreadsheet) ที่จัดอยู่ในชุดของ Microsoft Office ท่านจะเลือก Click เมนูย่อยและแฟ้มเอกสารที่ได้จะมีนามสกุลใด

ตัวเลือก	เมนูย่อย	นามสกุล
1	save	.xls
2	save	.doc
3	save	.ppt
4	save as	.ppt
5*	save as	.xls

13.หากท่านได้รับมอบหมายให้รวบรวมไฟล์รูปภาพและไฟล์เสียงต่าง ๆ ในการนำเสนอผลงาน ซึ่งมีจำนวนมากและมีนามสกุลไฟล์ที่หลากหลาย ท่านจะจัดแฟ้มตามข้อใดจึงจะถูกต้องที่สุด

ตัวเลือก	แฟ้มรูปภาพ				แฟ้มเสียง			
1.	.aac	.gif	.jpg	.raw	.flac	.ogg	.png	.wav
2.	.flac	.png	.psd	.tif	.aac	.cda	.m4a	.wma
3*.	.gif	.jpg	.png	.raw	.cda	ogg	.mp3	.wav
4.	.jpg	.ogg	.psd	.tif	.cda	.png	.raw	.wma
5.	.cda	.gif	.jpg	.raw	.aac	.m4a	.tif	.wav

	A	B
1	ลำดับที่	ชื่อนักเรียน
2	4	ดี
3	2	บี
4	1	เอ
5	3	ซี

14.จากภาพของโปรแกรม MS Excel เมื่อท่านเลือกคอลัมน์ลำดับที่ และใช้คำสั่ง Sort Smallest to Largest พร้อมกับเงื่อนไข Continue with Current Selection ผลที่ได้จะเป็นไปตามข้อใด

1.

	A	B
1	ลำดับที่	ชื่อนักเรียน
2	1	เอ
3	2	บี
4	3	ซี
5	4	ดี

2*.

	A	B
1	ลำดับที่	ชื่อนักเรียน
2	1	ดี
3	2	บี
4	3	เอ
5	4	ซี

3.

	A	B
1	ลำดับที่	ชื่อนักเรียน
2	4	ดี
3	3	ซี
4	2	บี
5	1	เอ

4.

	A	B
1	ลำดับที่	ชื่อนักเรียน
2	4	ดี
3	3	บี
4	2	เอ
5	1	ซี

5.

	A	B
1	ลำดับที่	ชื่อนักเรียน
2	1	เอ
3	3	ซี
4	2	บี
5	4	ดี

15.หากอาจารย์วิชาสถิติในชีวิตประจำวันมอบหมายให้ท่านเลือกรูปแบบการนำเสนอข้อมูลปริมาณการใช้ไฟฟ้าของมหาวิทยาลัยในแต่ละเดือนย้อนหลัง 5 ปี ท่านจะเลือกนำเสนอข้อมูลในรูปแบบใดจึงจะเหมาะสมกับข้อมูลดังกล่าว

1*. กราฟเส้น　2. แผนภูมิแท่ง　3. ตารางแจกแจงความถี่
4. ฮีสโทแกรม　5. แผนภาพการกระจาย

16. ข้อใดเป็นการเปรียบเทียบข้อมูลสารสนเทศในเหตุการณ์พายุถล่มที่ฮ่องกงในเดือนสิงหาคม พ.ศ.2560 ได้เหมาะสมที่สุด

1*. ยอดผู้เสียชีวิต กับผู้บาดเจ็บ
2. ปริมาณน้ำฝน กับจำนวนผู้ว่างงาน
3. อายุของประชากร กับรายได้ต่อหัวของประชากร
4. จำนวนครัวเรือนที่ค้างชำระค่าไฟฟ้า ระยะเวลาที่ไฟฟ้าดับ
5. จำนวนศูนย์พักพิงฉุกเฉิน กับจำนวนทรัพย์สินของทางการที่สูญหาย

17.เจ้าหน้าที่ท่านหนึ่งมีข้อมูลปริมาณน้ำกักเก็บในเขื่อน 3 แห่ง ระยะเวลา 3 ปี และข้อมูลปริมาณน้ำฝนที่ตกเฉลี่ยระยะเวลา 3 ปี และต้องการนำเสนอสารสนเทศให้แก่ผู้มีส่วนเกี่ยวข้องกับการบริหารจัดการน้ำใน 3 ประเด็นได้แก่

(1) การเปรียบเทียบปริมาณน้ำกักเก็บเฉลี่ยในเขื่อน 3 แห่ง ระยะเวลา 3 ปี
(2) สัดส่วนปริมาณน้ำฝนที่ตกเฉลี่ยจำนวน 3 ปี และ
(3) ความสัมพันธ์ระหว่างปริมาณน้ำฝนที่ตกเฉลี่ยและปริมาณน้ำที่กักเก็บในเขื่อน ระยะเวลา 3 ปี

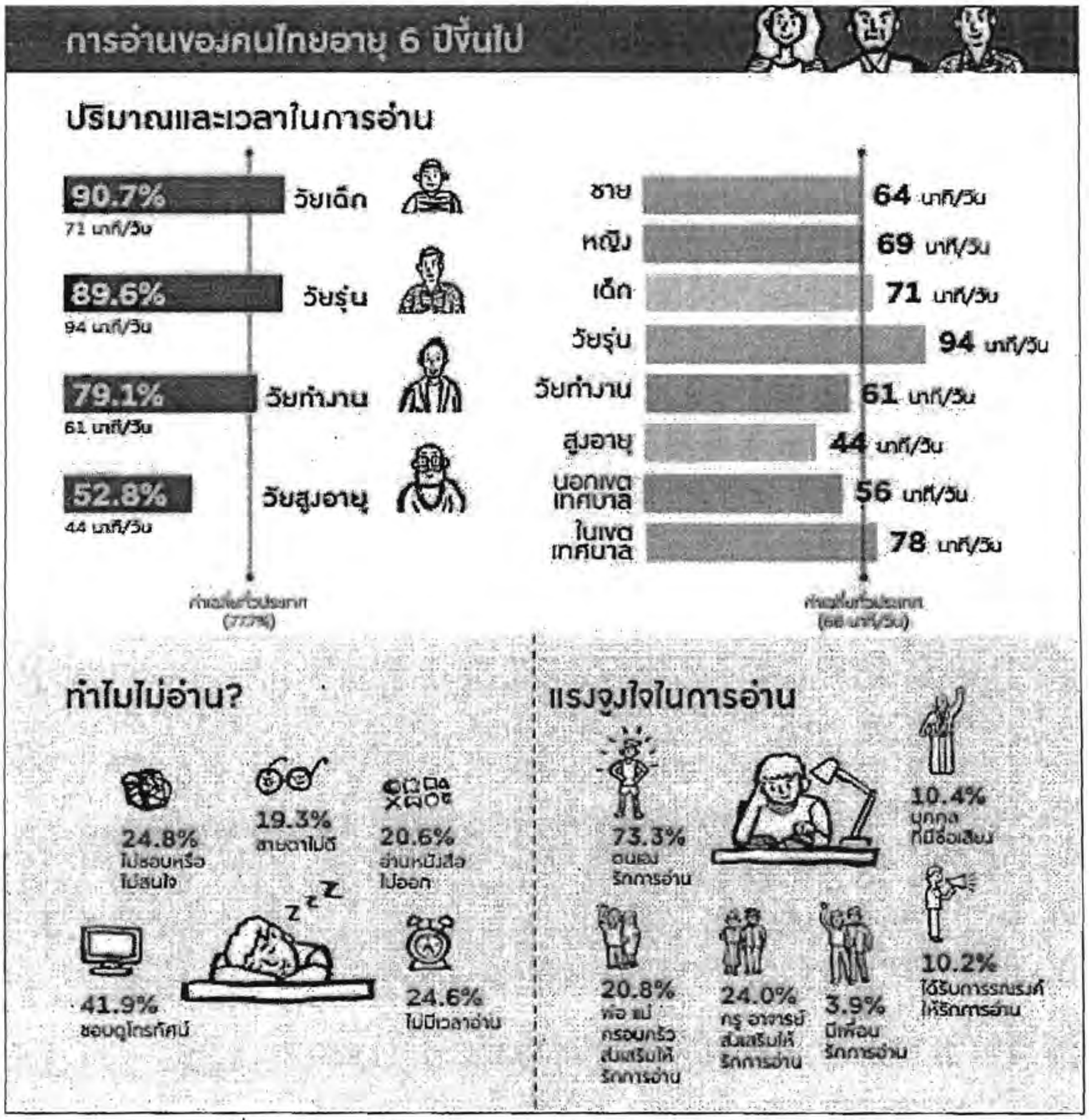
เมื่อกำหนดวิธีการนำเสนอสารสนเทศจำแนกตามประเด็นดังตาราง

ประเด็นนำเสนอ	กราฟ/แผนภูมิ			
	แผนภูมิแท่ง	แผนภูมิวงกลม	กราฟเส้น	แผนผังการกระจาย
A) เปรียบเทียบปริมาณน้ำกักเก็บเฉลี่ยในเขื่อน 3 แห่ง ระยะเวลา 3 ปี	/		/	
B) สัดส่วนปริมาณน้ำฝนที่ตกเฉลี่ยจำนวน 3 ปี	/	/		
C) ความสัมพันธ์ระหว่างปริมาณน้ำฝนที่ตกเฉลี่ยและปริมาณน้ำที่กักเก็บในเขื่อน ระยะเวลา 3 ปี			/	/

ข้อใดถูกต้อง

1. A　2. B　3. A และ C
4*. B และ C　5. A, B และ C

18. จากหัวข้อข่าวเรื่อง "คนไทยอ่านปีละ 8 บรรทัด" สถิติใหม่ชี้อ่านกันวันละ 66 นาที!!"และแผนภาพแสดงผลการสำรวจที่กำหนดให้ ข้อใดต่อไปนี้กล่าวถึงผลการสำรวจได้ถูกต้อง



(ที่มา : https://www.khaosod.co.th/monitor-news/news_16138)

1. วัยรุ่นใช้ระยะเวลาในการอ่านมากกว่าวัยเด็ก
2. เพศหญิงใช้ระยะเวลาในอ่านประมาณ 10 ชม./สัปดาห์
3. สาเหตุส่วนใหญ่ที่คนไม่ชอบอ่านหนังสือเพราะไม่ชอบหรือไม่สนใจ
4. คนส่วนใหญ่ที่ชอบอ่านหนังสือเพราะครู อาจารย์ส่งเสริมการอ่าน
5*.คนนอกเขตเทศบาลมีแนวโน้มชอบอ่านน้อยกว่าคนในเขตเทศบาล

19.หากท่านกำลังท่องอินเทอร์เน็ตแล้วพบว่าเพื่อนของท่านได้แบ่งปัน (Share) ข้อมูลเกี่ยวกับประเภทของพลังงานตามรายละเอียดด้านล่าง

https://www.baanjomyut.com/library_3/extension-1/energy_and_quality_of_life/02.html

ประเภทของพลังงาน

เนื่องจากความหลากหลายของพลังงานที่มีอยู่เป็นส่วนหนึ่งที่ทำให้ผู้ศึกษาอาจเกิดความเข้าใจสับสน จึงได้มีนักวิชาการพยายามที่จะจำแนกพลังงานเพื่อให้ง่ายต่อการศึกษามากยิ่งขึ้น โดยพลังงานสามารถจำแนกได้ดังนี้

1. จำแนกตามแหล่งที่ได้มา แบ่งออกเป็น 2 ประเภท ได้แก่

1.1 พลังงานต้นกำเนิด (Primary energy) หมายถึง แหล่งพลังงานที่เกิดขึ้นหรือมีอยู่แล้วตามธรรมชาติ สามารถนำมาใช้ประโยชน์ได้โดยตรง ได้แก่ น้ำ แสงแดด ลม เชื้อเพลิงตามธรรมชาติ เช่น น้ำมันดิบ ถ่านหิน ก๊าซธรรมชาติ พลังงานความร้อนใต้พิภพ แร่นิวเคลียร์ ไม้ฟืน แกลบ ชานอ้อย เป็นต้น

1.2 พลังงานแปรรูป (Secondary energy) หมายถึง สภาวะของพลังงานซึ่งได้มาโดยการนำพลังงานต้นกำเนิดมาแปรรูป ปรับปรุง ปรุงแต่ง ให้อยู่ในรูปที่สามารถนำไปใช้ประโยชน์ในลักษณะต่าง ๆ ตามความต้องการ เช่น พลังงานไฟฟ้า ผลิตภัณฑ์ปิโตรเลียม ถ่านไม้ ก๊าซปิโตรเลียมเหลว เป็นต้น

2. จำแนกตามแหล่งที่นำมาใช้ประโยชน์ แบ่งออกเป็น 2 ประเภท ได้แก่

2.1 พลังงานหมุนเวียน (Renewable energy resources) หมายถึง แหล่งพลังงานที่ใช้แล้วหมุนเวียนกลับมาให้ใช้เป็นประจำ เช่น น้ำ แสงแดด ลม เป็นต้น

2.2 พลังงานที่ใช้หมดเปลือง (Non - renewable energy resources) หมายถึง พลังงานที่เมื่อใช้หมดแล้วไม่สามารถสร้างขึ้นมาใหม่หรือหามาทดแทนโดยธรรมชาติได้ทันความต้องการในเวลาอันรวดเร็ว ได้แก่ น้ำมัน ก๊าซธรรมชาติ ถ่านหิน เป็นต้น

3. จำแนกตามลักษณะทางการค้า แบ่งออกเป็น 2 ประเภท ได้แก่

3.1 พลังงานทางพาณิชย์ (Commercial energy) หมายถึง พลังงานที่มีการซื้อขายกันในวงกว้างและดำเนินการผลิตในลักษณะอุตสาหกรรม เช่น น้ำมันปิโตรเลียม ก๊าซธรรมชาติ ถ่านหิน แร่นิวเคลียร์ ไฟฟ้า เป็นต้น

3.2 พลังงานนอกพาณิชย์ (Non - commercial energy) หมายถึง พลังงานที่มีการซื้อขายกันในวงแคบและดำเนินการผลิตในลักษณะกิจกรรมในครัวเรือน ใช้กันมากในชนบท เช่น ฟืน แกลบ ชานอ้อย และมูลสัตว์ เป็นต้น

จากข้อมูลข้างต้น ข้อใดกล่าวเกี่ยวกับประเภทของพลังงานได้ถูกต้อง

1. พลังงานต้นกำเนิดทุกชนิดจัดเป็นพลังงานหมุนเวียน
2*. พลังงานแปรรูปทุกชนิดจัดเป็นพลังงานที่ใช้หมดเปลือง
3. พลังงานที่ใช้หมดเปลืองทุกชนิดจัดเป็นพลังงานทางพาณิชย์
4. พลังงานนอกพาณิชย์ทุกชนิดจัดเป็นพลังงานหมุนเวียน
5. พลังงานทางพาณิชย์ทุกชนิดจัดเป็นพลังงานต้นกำเนิด

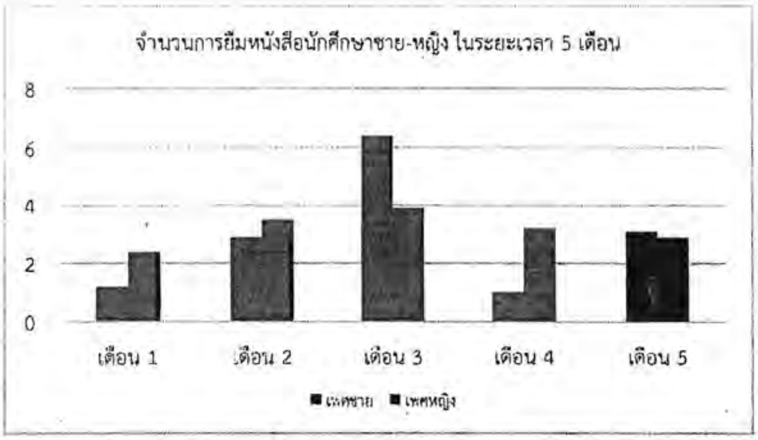
จากข้อมูลสารสนเทศที่กำหนดให้ ในตาราง จงตอบคำถามข้อที่ 20

รุ่น	Spec
X100	Size : 9.7 นิ้ว CPU Speed : 2.15 GHZ Memory : Ram Size 4 GB S Pen ปากกาอัจฉริยะ กล้องหลัง : -
X200	Size : 8.0 นิ้ว CPU Speed : 1.3 GHZ Memory : Ram Size 3 GB กล้องหลัง : 13 ล้านพิกเซล
X300	Size : 9.7 นิ้ว CPU Speed : 1.9 GHZ Memory : Ram Size 3 GB กล้องหลัง : 8 ล้านพิกเซล

20. หากผู้ปกครองของท่านต้องการคำแนะนำในการเลือกซื้อ tablet สำหรับใช้ในการอ่านข่าวสารและรับชมทีวีออนไลน์ ท่านจะแนะนำ tablet รุ่นใดที่มีคุณสมบัติ (spec) เหมาะสมกับการใช้งานมากที่สุด เพราะเหตุใด

1. X200 ดีกว่า X100 เพราะ มีกล้องหลัง
2. X300 ดีกว่า X100 เพราะ มีกล้องหลัง
3*. X100 ดีกว่า X200 เพราะ หน้าจอใหญ่กว่า
4. X300 ดีกว่า X200 เพราะ CPU มีความไวกว่า
5. X200 ดีกว่า X300 เพราะ กล้องหลังมีกำลังขยายสูงกว่า

21.จากแผนภูมิแสดงจำนวนนักศึกษายืมหนังสือจากห้องสมุดแห่งหนึ่ง ระหว่างเดือนที่ 1-5 จำแนกตามเพศ ข้อใดอธิบายแปลผลสารสนเทศได้ถูกต้อง และถ้าต้องการนำเสนอด้วยรูปแบบอื่นที่ให้สารสนเทศเช่นเดียวกับกราฟแท่งนี้ ควรเลือกแบบใด



ตัวเลือก	แปลผล	รูปแบบการนำเสนออื่นที่ใช้ทดแทนแผนภูมิแท่ง		
		กราฟเรดาร์	กราฟเส้น	แผนภูมิวงกลม
1.	จำนวนการยืมหนังสือของเพศหญิงมีความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานใกล้เคียงกัน			/
2.	จำนวนการยืมหนังสือของเพศหญิงมีส่วนเบี่ยงเบนมาตรฐานมากกว่าเพศชาย	/	/	/
3.	จำนวนการยืมหนังสือของเพศชายมีส่วนเบี่ยงเบนมาตรฐานน้อยกว่าเพศหญิง	/		/
4.	จำนวนการยืมหนังสือของเพศชายมีความแปรปรวนน้อยกว่าเพศหญิง		/	/
5*.	จำนวนการยืมหนังสือของเพศชายมีความแปรปรวนมากกว่าเพศหญิง		/	

22.ท่านคิดว่า พฤติกรรมใดในข้อต่อไปนี้ มีความเสี่ยงที่จะถูกแฮก (Hack) ข้อมูลน้อยที่สุด
1. ใช้ระบบจำรหัสผ่านอัตโนมัติ
2. เชื่อม WiFi ฟรีในที่สาธารณะ
3*. เปลี่ยนแปลงรหัสผ่านทุกเดือน
4. กดเข้าเว็บไซต์ที่ได้รับรางวัล 5 ล้าน
5. ดาวน์โหลดไฟล์เข้าอีเมลที่ร้านอินเทอร์เน็ต

23. ข้อใดเป็นการรักษาความปลอดภัยจากเทคโนโลยีที่มีความเสี่ยงต่อการถูกโจรกรรมข้อมูล

1*. Encryption 2. Anti Virus Software 3.Digital Signature
4. Contents Filter 5. Biometric Authentication

24. เมื่อท่านเข้าใช้งานอินเตอร์เน็ตและเกิดปัญหาดังภาพ ข้อใดเป็นวิธีการแก้ไขปัญหาที่ดีที่สุดเป็นลำดับแรก

1. ตรวจสอบระบบการเก็บ Cookies
2. ตรวจสอบระบบการประมวลผล (CPU)
3*. ตรวจสอบระบบการเชื่อมต่ออินเตอร์เน็ต
4. ตรวจสอบระบบการทำงานของเครื่องคอมพิวเตอร์
5. ตรวจสอบระบบการใช้งานของเบราว์เซอร์ (Browser)

25. ท่านคิดว่าการกระทำในข้อใดเหมาะสมที่สุดเกี่ยวกับการได้รับข้อมูลและข่าวสารต่าง ๆ บนสังคมออนไลน์ที่นำเสนอหลายด้านหลายแง่มุม
1.ไม่อ่านเนื่องจากส่วนใหญ่เป็นข้อมูลและข่าวสารที่ส่งต่อกันมา ทำให้สามารถได้รับข้อมูลอันเป็นเท็จได้
2.อ่านและเลือกส่งต่อเฉพาะข้อมูลที่ตนคิดว่าเป็นประโยชน์ต่อผู้อื่น
3.อ่านและเพิ่มเติมข้อมูลที่ตนทราบมาก่อนส่งต่อผู้อื่น
4*.อ่านพร้อมตรวจสอบข้อเท็จจริงจากแหล่งข้อมูลอื่นที่เชื่อถือได้ และเลือกส่งต่อเฉพาะข้อมูลที่เป็นประโยชน์ต่อผู้อื่น
5.อ่านรับทราบข้อมูลและข่าวสาร โดยไม่เผยแพร่ต่อทางสื่อสังคมออนไลน์ เพราะมีการส่งต่ออย่างแพร่หลายอยู่แล้ว แต่จะเล่าต่อให้เพื่อนหรือครอบครัวรับทราบแทน

26.บริษัทประกันภัยแห่งหนึ่งต้องการพัฒนาระบบสารสนเทศที่เป็นการใช้งานเครือข่ายภายในเท่านั้น เพื่อป้องกันบุคคลภายนอกเข้ามาลักลอบข้อมูลลูกค้าและข้อมูลทางการตลาด และอีกทั้งยังต้องการป้องอันตรายทางเทคโนโลยีที่จากโปรแกรมขนาดเล็กที่สร้างมารบกวนการทำงานและบันทึกประวัติ การใช้งานคอมพิวเตอร์ของผู้ใช้งานได้ จากสถานการณ์ดังกล่าวนี้จึงมีการรักษาความปลอดภัยทางเทคโนโลยีด้วยการกำหนดให้มี ID/Password มีการใช้ Anti-Virus Software จากความต้องการของบริษัทประกันและวิธีการป้องกันอันตรายทางเทคโนโลยี ท่านคิดว่าการรักษาความปลอดภัยทางเทคโนโลยีของบริษัทมีความสอดคล้องกับความต้องของบริษัทหรือไม่เพราะเหตุใด

1. มีความเหมาะสม เนื่องจาก ID/Password สามารถป้องกันการใช้งานจากบุคคลภายนอกและพิสูจน์ได้ว่าข้อมูลนั้นมาจากผู้ส่งข้อมูลจริง และ Anti-Virus Software มีความจำเป็นต่อการป้องกันไวรัสเข้ามาก่อกวนระบบคอมพิวเตอร์ของบริษัทได้

2*. มีความเหมาะสม เนื่องจาก ID/Password สามารถป้องกันการใช้ระบบคอมพิวเตอร์จากบุคคลภายนอก และ Anti-Virus Software มีความจำเป็นต่อการป้องกันไวรัสเข้ามาก่อกวนระบบคอมพิวเตอร์ของบริษัทได้

3. ไม่เหมาะสม เนื่องจากต้องใช้ Content filter จึงจะสามารถป้องกันการใช้ระบบคอมพิวเตอร์จากบุคคลภายนอกได้

4. ไม่เหมาะสม เนื่องจากต้องใช้ Digital Signature จึงจะสามารถป้องกันการใช้ระบบคอมพิวเตอร์จากบุคคลภายนอกได้

5. ไม่เหมาะสม เนื่องจากต้องใช้ Content filter และ Digital Signature จึงจะสามารถป้องกันการใช้ระบบคอมพิวเตอร์จากบุคคลภายนอกได้

27.บริษัทรถยนต์แห่งหนึ่งต้องการนำเสนอโมเดลรถยนต์แห่งอนาคตของบริษัท ซึ่งต้องการให้ผู้ชมได้สัมผัสถึงโมเดลรถยนต์แบบสามมิติ ร่วมกับการนำเสนอผลการใช้พลังงานเชื้อเพลิงเมื่อเทียบกับรถยนต์ใน 5 ปีที่ผ่านมา บนจอภาพขนาดใหญ่ เพื่อให้ผู้ชมเห็นได้อย่างชัดเจน นอกจากนี้ยังเพิ่มการมีส่วนร่วมให้ผู้ชมด้วยการใช้สมาร์ทโฟนผ่านหน้าจอได้โดยตรง จากความต้องการดังกล่าวจึงได้กำหนดลักษณะสารสนเทศไว้ดังนี้

A. นำเสนอโมเดลรถยนต์แบบสามมิติด้วยรูปแบบ Computer graphic และ QR code
B. นำเสนอโมเดลรถยนต์แบบสามมิติด้วยรูปแบบ Computer graphic และ Barcode
C. นำเสนอกราฟเส้นเพื่อนำเสนอการเปลี่ยนแปลงการใช้พลังงานที่ลดลงในระยะ 5 ปี ซึ่งบันทึกเป็นรูปแบบไฟล์ภาพนิ่ง .png
D. นำเสนอกราฟแท่งเพื่อนำเสนอการเปลี่ยนแปลงการใช้พลังงานที่ลดลงในระยะ 5 ปี ซึ่งบันทึกเป็นรูปแบบไฟล์ภาพนิ่ง .gif

จงพิจารณาว่า<u>สารสนเทศแบบใดบ้าง</u>ที่มีความเหมาะสมในการนำมาใช้งานสำหรับสถานการณ์ดังกล่าว

1*. A และ C
2. A และ D
3. B และ C
4. C และ D
5. A, C และ D

28.ระบบสารสนเทศขององค์กรแห่งหนึ่งได้รับความเสี่ยงจากอันตรายทางเทคโนโลยี 3 ประเด็น ซึ่งผู้อำนวยการขององค์กรได้ประเมินสถานการณ์ที่เกิดขึ้น และได้เลือกวิธีการรักษาความปลอดภัยด้านทางเทคโนโลยีสารสนเทศ ดังตาราง

อันตรายทางเทคโนโลยี	การรักษาความปลอดภัยทางเทคโนโลยี			
	Firewall	Anti-Virus software	Encryption	ID/ Password
A. การทำงานถูกรบกวนจากการเปิดหน้าต่างโฆษณาบ่อยครั้ง		/		/
B. การลักลอบใช้งานเครือข่ายขององค์กรจากบุคคลภายนอก	/	/		
C. การลักลอบอ่านข้อมูลภายในเครื่องคอมพิวเตอร์ขององค์กร			/	/

จงพิจารณาว่าการแก้ไขด้านเทคโนโลยีสารสนเทศ ข้อใดถูกต้อง

1. A
2*. C
3. A และ B
4. B และ C
5. A, B และ C

29.เมื่อท่านได้รับมอบหมายให้แนะนำวิธีการออกกำลังกายให้กับผู้สูงอายุ ท่านควรเลือกใช้สื่อประเภทใดเพื่อใช้ประกอบการให้คำแนะนำ
1.แผ่นพับ
2*. วีดิทัศน์
3. Infographic
4. Microsoft Word
5. Portable Document Format

30. มหาวิทยาลัยแห่งหนึ่งมีระบบอินเตอร์เน็ตซึ่งมีความรวดเร็วและเสถียร ข้อใดต่อไปนี้เป็นระบบการทดสอบที่เหมาะสมที่สุดสำหรับมหาวิทยาลัยแห่งนี้ซึ่งสอดคล้องกับทักษะการใช้งานสารสนเทศในศตวรรษที่ 21
1. pool testing
2*. on-line testing
3. automatic testing
4. paper-based testing
5. computer-based testing

31.ข้อมูลประเภทใดต่อไปนี้ไม่ควรใช้ในการแลกเปลี่ยนข้อมูลมากที่สุด
1.ข้อมูลรายการสินค้า
2.ข้อมูลส่วนตัวพนักงาน
3*.ข้อมูลรหัสชื่อ และรหัสผ่าน
4.ข้อมูลรายงานผลประกอบการของบริษัท
5.ข้อมูลการเงินของบริษัทที่สามารถเปิดเผยได้

32. หากท่านเป็นตัวแทนของสมาชิกในกลุ่มในการส่งรายงานกลุ่มแก่อาจารย์ประจำวิชาผ่านทางอีเมล (E-mail) ซึ่งท่านต้องการส่งรายงานดังกล่าวให้สมาชิกในกลุ่มโดยที่ไม่ต้องการให้อาจารย์ทราบ ท่านจะกรอก E-mail address ของสมาชิกในกลุ่มในช่องใด
1. To (ถึง)
2. Cc (ส่งสำเนาถึง)
3*. Bcc (ส่งสำเนาลับถึง)
4. To (ถึง) และ Cc (ส่งสำเนาถึง)
5. To (ถึง) และ Bcc (ส่งสำเนาลับถึง)

33.ท่านคิดว่ารูปแบบการนำเสนอใดมีความเหมาะสมต่อการนำเสนอรายงานเกี่ยวกับความหมายของคำว่า "สมรรถนะ" ให้ผู้สนใจเข้าใจง่ายและครอบคลุมความหมาย

ผลการสืบค้นข้อมูลความหมายของ "สมรรถนะ"

1. Katz & Green (1992 อ้างถึงใน ชัชวาลย์ ทัตศิวัช, 2554) สมรรถนะ หมายถึง ความสามารถในการกระทำสิ่งใดสิ่งหนึ่งที่ระดับความชำนาญ รวมทั้งความรู้ ทักษะ เจตคติ และค่านิยม
2. ศิริชัย กาญจนวาสี และคณะ (อ้างถึงใน ชุตินันท์ จันทรเสนานนท์, 2553) สมรรถนะ หมายถึง ความรู้(Knowledge) ทักษะ(Skill) ความสามารถ(Ability) และคุณลักษณะ (Attribute) ซึ่งมีอยู่ในแต่ละบุคคลที่บูรณาการขึ้นมาเป็นกลุ่มพฤติกรรมจนเกิดเป็นความสามารถอันเด่นชัดที่เพียงพอในการทำงาน(Action) ได้อย่างมีมาตรฐาน (Standard) จนงานประสบความสำเร็จ โดยสมรรถนะของบุคคลเป็นสิ่งที่สั่งสมมาจากการศึกษา ประสบการณ์และการฝึกฝน
3. สำนักงานคณะกรรมการข้าราชการพลเรือน (2558) สมรรถนะ คือ คุณลักษณะเชิงพฤติกรรม ที่เป็นผลมาจากความรู้ ทักษะ/ความสามารถ และคุณลักษณะอื่น ๆ ที่ทำให้บุคคลสามารถสร้างผลงานได้โดดเด่นกว่าเพื่อนร่วมงานอื่น ๆ ในองค์กร

แหล่งที่มา: รายงานผลการประเมินสมรรถนะผู้ใหญ่ของประเทศไทย ปี 2557 (สกศ.)

1. Katz & Green (1992 อ้างถึงใน ชัชวาลย์.ทัตศิวัช, 2554) สมรรถนะ หมายถึง ความสามารถในการกระทำสิ่งใดสิ่งหนึ่งที่ระดับความชำนาญ รวมทั้งความรู้ ทักษะ เจตคติ และค่านิยม

ศิริชัย กาญจนวาสี และคณะ (อ้างถึงใน ชุตินันท์ จันทรเสนานนท์, 2553) สมรรถนะ หมายถึง ความรู้(Knowledge) ทักษะ ความสามารถ และคุณลักษณะของแต่ละบุคคลที่บูรณาการขึ้นมาเป็นกลุ่มพฤติกรรมจนเกิดเป็นความสามารถอันเด่นชัด

สำนักงานคณะกรรมการข้าราชการพลเรือน (ก.พ.) (2558) สมรรถนะ คือ คุณลักษณะเชิงพฤติกรรม ที่เป็นผลมาจากความรู้ ทักษะ/ความสามารถ และคุณลักษณะอื่น ๆ

2. จากความหมายของคำว่าสมรรถนะนำเสนอข้างต้นสามารถสรุปได้ว่า สมรรถนะหมายถึง คุณลักษณะของบุคคลที่จะทำให้งานที่ทำประสบผลสำเร็จมีมาตรฐานซึ่งคุณลักษณะดังกล่าวจะซ่อนอยู่ในตัวบุคคล

3. สมรรถนะ(Competency) คือ คุณลักษณะของบุคคลที่จะทำให้งานที่ทำประสบผลสำเร็จมีมาตรฐาน

4.

รายชื่อ	ความหมายของสมรรถนะ					
	ความรู้	ความสามารถ	ทักษะ	คุณลักษณะ	เจตคติ	ค่านิยม
Katz & Green	✓	-	✓	-	✓	✓
ศิริชัย กาญจนวาสี และคณะ	✓	✓	✓	✓	-	-
ก.พ.	✓	✓	✓	✓	-	-

5*.

รายชื่อ	ความหมายของสมรรถนะ					
	ความรู้	ความสามารถ	ทักษะ	คุณลักษณะ	เจตคติ	ค่านิยม
Katz & Green	✓	-	✓	-	✓	✓
ศิริชัย กาญจนวาสี และคณะ	✓	✓	✓	✓	-	-
ก.พ.	✓	✓	✓	✓	-	-

จากตารางข้างต้น กล่าวได้ว่าสมรรถนะ คือ คุณลักษณะของบุคคลที่จะทำให้งานที่ทำประสบผลสำเร็จมีมาตรฐานซึ่งคุณลักษณะดังกล่าวจะซ่อนอยู่ในตัวบุคคล โดยจะประกอบด้วยความรู้ ความสามารถ ทักษะ และคุณลักษณะ

34.เมื่อท่านต้องการส่งเอกสารที่มีความปลอดภัย โดยฝ่ายรับต้องได้รับเฉพาะเอกสารที่ถูกเข้ารหัส ภายใต้ Public key cryptography จงพิจารณาขั้นตอนการดำเนินการส่งข้อมูลระหว่างฝ่ายรับข้อมูลและฝ่ายส่งข้อมูล ดังตาราง

ข้อ	ฝ่ายรับข้อมูล		ฝ่ายส่งข้อมูล
A.	ดำเนินการ Register public key กับ CA	→	เข้ารหัสเอกสารโดยใช้ Public key จากฝ่ายรับข้อมูล
B.	ดำเนินการถอดรหัสโดยใช้ Secret key ที่ทำการ Register กับ CA	→	ส่งข้อมูลที่ถูกเข้ารหัสให้ฝ่ายรับข้อมูล
C.	ดำเนินการถอดรหัสโดยใช้ Secret key และ General key ที่ทำการ Register กับ CA	→	เข้ารหัสเอกสารโดยใช้ Public key จากฝ่ายรับข้อมูล
D.	ดำเนินการถอดรหัสโดยใช้ Public key ที่ทำการ Register กับ CA	→	ส่งข้อมูลที่ถูกเข้ารหัสให้ฝ่ายรับข้อมูล

ข้อใดถูกต้อง

1*. A และ C
2. B และ C
3. C และ D
4. A,B และ C
5. B, C และ D

35.หากพนักงานคนหนึ่งต้องการส่งอีเมลให้แก่บุคคล 2 กลุ่ม เป็นดังนี้ *กลุ่มแรก* คือ เพื่อนร่วมงานจำนวน 5 คน ซึ่งเพื่อนร่วมงานแต่ละคนจะเห็นรายชื่อเพื่อนร่วมงานอีก 4 คนที่ได้รับข้อมูล และ*กลุ่มสอง* คือ ผู้บังคับบัญชา โดยไม่ต้องการให้เพื่อนร่วมทั้ง 5 คนทราบว่ามีการส่งอีเมลดังกล่าวให้ผู้บังคับบัญชาเมื่อกำหนดตารางแสดงการกำหนดค่าปลายทางจำแนกตามกลุ่มผู้รับข้อมูล

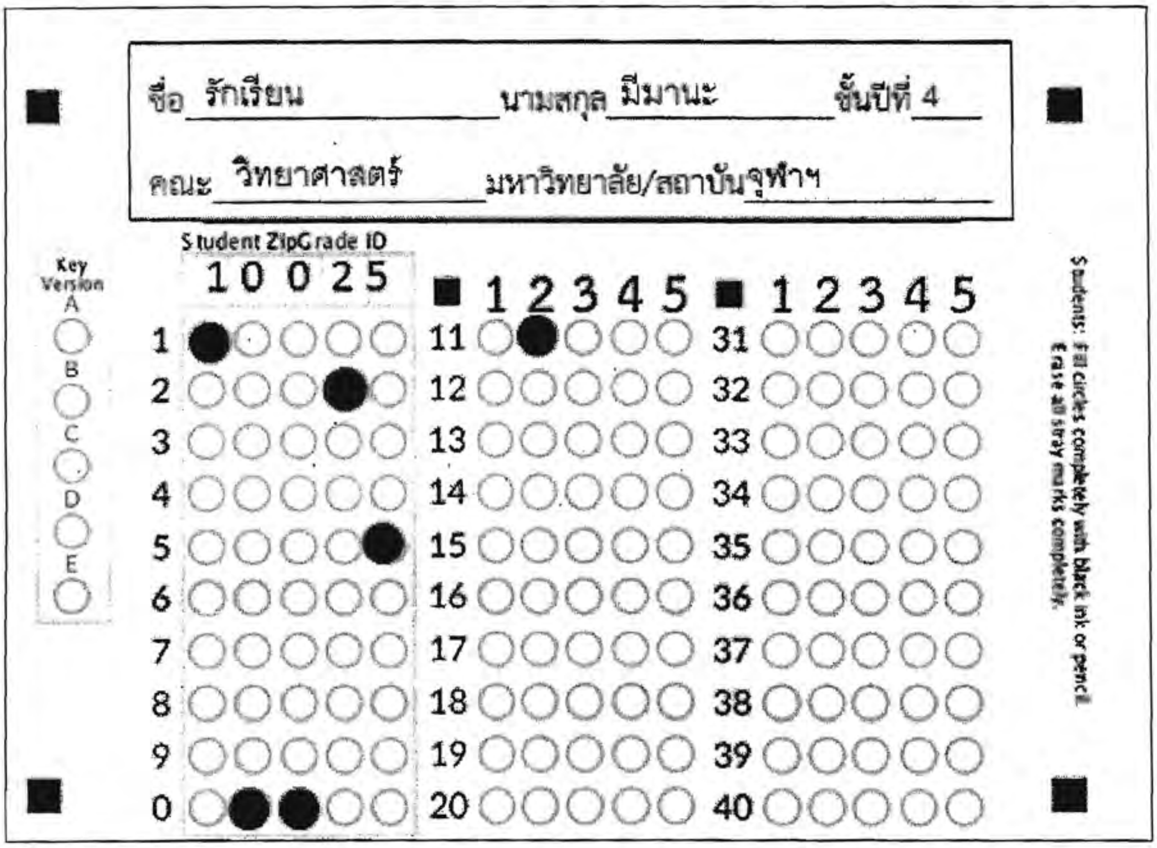
ข้อ	กลุ่ม	ค่าปลายทาง		
		To	cc	Bcc
A.	ผู้บังคับบัญชา		/	
B.	ผู้บังคับบัญชา			/
C.	เพื่อนร่วมงาน	/		
D.	เพื่อนร่วมงาน			/

ข้อใดถูกต้อง

1. A
2. C
3*. B และ C
4. C และ D
5. A, B และ D

****

ตัวอย่างกระดาษคำตอบแบบฝน



ภาคผนวก 5

ตัวอย่างแบบประเมินความตรงตามเนื้อหา (IOC)

แบบประเมินความสอดคล้องระหว่างข้อคำถามและองค์ประกอบของการทดสอบ

รายละเอียดแบบทดสอบ

รูปแบบข้อสอบ : ข้อสอบหลายตัวเลือก 5 ตัวเลือก (คำตอบถูก 1 คำตอบ)

จำนวน : 50 ข้อ

ตัวอย่าง : นิสิต/นักศึกษา ระดับปริญญาตรีชั้นปีที่ 3-4

เนื้อหาแบบทดสอบ : ทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21

องค์ประกอบ :

องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ	จำนวน 10 ข้อ	(ข้อที่1-10)
องค์ประกอบที่ 2 การจัดการสารสนเทศ	จำนวน 10 ข้อ	(ข้อที่11-20)
องค์ประกอบที่ 3 การบูรณาการสารสนเทศ	จำนวน 10 ข้อ	(ข้อที่21-30)
องค์ประกอบที่ 4 การประเมินสารสนเทศ	จำนวน 10 ข้อ	(ข้อที่31-40)
องค์ประกอบที่ 5 การสื่อสารสนเทศ	จำนวน 10 ข้อ	(ข้อที่41-50)

องค์ประกอบที่ 1 การเข้าถึงสารสนเทศ (Information accessibility)

การเข้าถึงสารสนเทศ ประกอบด้วย 1) การระบุสารสนเทศที่จำเป็น คือ สามารถเข้าใจและอธิบายความหมายของสารสนเทศที่เลือกมาใช้งานได้อย่างเหมาะสม 2) วิธีการในการเก็บรวบรวมสารสนเทศ คือ สามารถเลือกวิธีการในการเก็บรวบรวมสารสนเทศแต่ละประเภทได้อย่างเหมาะสม และ 3) การเรียกสารสนเทศกลับคืนเพื่อการใช้งาน คือ สามารถอธิบายหรือแสดงวิธีการเรียกสารสนเทศที่มีอยู่กลับคืนมาเพื่อการใช้งานได้อย่างถูกต้อง

ข้อสอบ	ระดับความคิดเห็น			ข้อเสนอแนะ
	+1	0	-1	
1.ปัจจุบันหลายเว็บไซต์มีการใช้งานโปรโตคอล (protocol) https แทน http เพิ่มมากขึ้น ซึ่งโปรโตคอลดังกล่าวจะปรากฏอยู่หน้า url เว็บไซต์ เช่น https://google.co.th ข้อใดกล่าวถึงความแตกต่างของการใช้งานโปรโตคอลทั้งสองรูปแบบได้ถูกต้องที่สุด 1. โปรโตคอล http มีความซับซ้อนมากจึงมีการพัฒนาโปรโตคอล https ขึ้นมาทดแทน 2. โปรโตคอล http มีการเข้าถึงเว็บไซต์ผิดพลาดเป็นจำนวนมาก จึงมีการพัฒนาโปรโตคอล https ขึ้นมาทดแทน 3*. โปรโตคอล https มีการป้องกันความปลอดภัยสำหรับการเข้าถึงข้อมูลมากยิ่งขึ้น 4. โปรโตคอล https มีการเพิ่มความเร็วในการใช้งานเว็บไซต์ให้รวดเร็วมากยิ่งขึ้น 5. โปรโตคอล https มีความสะดวกในการใช้งานมากกว่าโปรโตคอล http				
2.เมื่อต้องการส่งไฟล์ขนาด 256 MB ผ่านทางจดหมายอิเล็กทรอนิกส์ (e-mail) วิธีใดเป็นวิธีการที่เหมาะสมที่สุด 1. ดำเนินการบีบอัดไฟล์ (zip) ก่อนส่งไปยังผู้รับ 2*. อัพโหลดขึ้นระบบ could และส่งลิงค์ (link) ไปยังผู้รับ 3. แบ่งไฟล์ออกเป็น 2 ส่วน และส่งไปยังผู้รับทีละไฟล์เพื่อไม่ให้เกินปริมาณที่กำหนด 4. ไม่สามารถดำเนินการส่งได้เนื่องจากไฟล์ขนาดใหญ่เกินปริมาณที่กำหนด 5. แนบไฟล์ไปกับจดหมายอิเล็กทรอนิกส์เหมือนการส่งไฟล์ทั่วไป				

ข้อสอบ	ระดับความคิดเห็น			ข้อเสนอแนะ
	+1	0	-1	
3.เครื่องคอมพิวเตอร์มีการเก็บไฟล์ที่มีความสำคัญไว้ที่ drive อื่น เช่น D:/ ซึ่งไม่ใช่ drive C:/ ข้อใดกล่าวถูกต้องที่สุดเกี่ยวกับแนวทางดังกล่าว 1*. เพื่อป้องกันไฟล์สูญหายหากระบบปฏิบัติการ (OS) เกิดปัญหา 2. เพื่อป้องกันไฟล์จากการถูกโจมตี (hack) จากผู้ไม่หวังดี 3. เพื่อป้องกันไฟล์ถูกทำลายจากโปรแกรมอื่นที่ไม่เกี่ยวข้อง 4. เพื่อป้องกันไฟล์จากการถูกเข้ารหัส (encrypt) โดยระบบปฏิบัติการ (OS) 5. เพื่อป้องกันไฟล์จากการถูกรบกวนโดยโปรแกรมที่มีการใช้หน่วยประมวลผล (CPU) จำนวนมาก				
4.เมื่อดำเนินการติดตั้งโปรแกรมในเครื่องคอมพิวเตอร์เรียบร้อยแล้ว วิธีการใดเป็นวิธีการที่สามารถทำให้การเรียกใช้โปรแกรมดังกล่าวดำเนินการได้สะดวกและรวดเร็วที่สุด 1. สร้าง Folder และเก็บโปรแกรมดังกล่าวไว้ที่ Desktop 2. สร้าง Folder และเก็บโปรแกรมดังกล่าวไว้ที่ Drive D 3*. สร้าง Shortcut และเก็บโปรแกรมดังกล่าวไว้ที่ Desktop 4. สร้าง Shortcut และเก็บโปรแกรมดังกล่าวไว้ที่ Drive D 5. สร้าง Shortcut และเก็บโปรแกรมดังกล่าวไว้ที่ Drive C				
5.เว็บไซต์ซึ่งมีการตั้งค่าการใช้งาน bookmark มีความแตกต่างจากการใช้งาน extensions อย่างไร 1. extensions ใช้สำหรับการเรียกเปิดหน้าเว็บไซต์ใหม่ bookmarks ใช้สำหรับการเรียกใช้งานเว็บไซต์ซึ่งเคยบันทึกไว้ 2*. extensions ใช้สำหรับการเรียกใช้โปรแกรมเสริม bookmarks ใช้สำหรับการเรียกใช้งานเว็บไซต์ซึ่งเคยบันทึกไว้ 3. bookmarks ใช้สำหรับการเรียกใช้โปรแกรมเสริม extensions ใช้สำหรับการเรียกใช้งานเว็บไซต์ซึ่งเคยบันทึกไว้ 4. bookmarks ใช้สำหรับการเรียกใช้โปรแกรมเสริม extensions ใช้สำหรับการเรียกเปิดหน้าเว็บไซต์ใหม่ 5. bookmarks ใช้สำหรับเรียกโปรแกรมเสริมเหมือนกับ extensions				

ภาคผนวก 6

แบบประเมินระบบการทดสอบอิงมาตรฐาน

แบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation)

คำชี้แจง

ให้ท่านพิจารณาข้อความต่อไปนี้ แล้วทำเครื่องหมาย ✓ ลงในช่องที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด โดยมีระดับความคิดเห็น 5 ระดับ ดังนี้

1 หมายถึง น้อยที่สุด

2 หมายถึง น้อย

3 หมายถึง ปานกลาง

4 หมายถึง มาก

5 หมายถึง มากที่สุด

คำถาม	ระดับความคิดเห็น				
	1 น้อยที่สุด	2 น้อย	3 ปานกลาง	4 มาก	5 มากที่สุด
ด้านความมีประโยชน์ (utility)					
1.การจัดสอบของระบบมีความครอบคลุมตามความต้องการของผู้ใช้งาน					
2.การรายงานผลการทดสอบของระบบมีความชัดเจน เข้าใจง่าย					
3.ผลการทดสอบสามารถนำไปใช้ประโยชน์ในการพัฒนาศักยภาพของผู้สอบได้					
ด้านความเป็นไปได้ในการนำไปใช้ (feasibility)					
4.ขั้นตอนของระบบมีการเรียกใช้งานง่าย สะดวกและรวดเร็ว					
5.การแสดงผลของระบบมีความรวดเร็ว					
6.ระบบมีสีสันดึงดูดความสนใจ					
7.ระบบมีคู่มือการใช้งานที่ชัดเจน					
ด้านความเหมาะสม (propriety)					
8.การแสดงขั้นตอนการทดสอบของระบบมีความชัดเจน					
9.การกำหนดเนื้อหาในการทดสอบของระบบมีความเหมาะสม					
10.ระยะเวลาในการทดสอบของระบบมีความเหมาะสม					
11.ผลการทดสอบมีความครอบคลุมและครบถ้วน					
ด้านความถูกต้อง (accuracy)					
12.ระบบสามารถจัดการทดสอบได้ตรงตามวัตถุประสงค์					
13.ระบบสามารถประมาณค่าความสามารถของผู้เข้าสอบได้ถูกต้อง					
14.ระบบสามารถสรุปผลการทดสอบและบันทึกผลการสอบได้ถูกต้อง					

ภาคผนวก 7

ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ (print out)

ค่าพารามิเตอร์ข้อสอบ

Form	Item no	Code	b	SE_b	a	SE_a	Form	Item no	Code	b	SE_b	a	SE_a
common	6	Cac1	-0.79	0.11	0.67	0.07	3	3	f3ac3	1.58	1.77	0.20	0.20
common	13	Cmn1	-1.16	0.15	0.59	0.07	3	4	f3ac4	-2.38	0.78	0.74	0.28
common	14	Cmn2	4.37	1.34	0.22	0.07	3	5	f3ac5	1.16	0.59	0.52	0.23
common	20	Cig2	-0.27	0.09	0.62	0.07	3	8	f3mn1	6.94	7.34	0.46	0.51
common	27	Cev1	0.84	0.11	0.70	0.07	3	9	f3mn2	-0.42	0.41	0.43	0.21
common	28	Cev2	1.75	0.19	0.69	0.08	3	10	f3mn3	4.14	2.59	0.48	0.32
common	34	Ccm1	3.91	0.88	0.32	0.07	3	12	f3mn5	1.93	4.44	0.09	0.19
common	35	Ccm2	-0.06	0.07	0.83	0.08	3	15	f3ig1	-0.32	0.41	0.40	0.21
1	1	f1ac1	-4.34	2.53	0.38	0.23	3	16	f3ig2	4.08	2.86	0.37	0.26
1	2	f1ac2	4.92	4.11	0.24	0.20	3	18	f3ig3	-1.70	0.70	0.54	0.23
1	3	f1ac3	-0.09	0.23	0.73	0.21	3	19	f3ig4	0.13	0.20	0.93	0.28
1	4	f1ac4	0.08	0.23	0.72	0.21	3	21	f3ig5	-0.04	0.26	0.65	0.23
1	5	f1ac5	-0.51	0.16	1.42	0.31	3	23	f3ev2	0.06	0.22	0.78	0.25
1	8	f1mn1	-1.20	0.34	0.83	0.23	3	24	f3ev3	2.57	3.31	0.16	0.20
1	9	f1mn2	-0.44	0.33	0.53	0.19	3	25	f3ev4	1.02	0.28	1.16	0.35
1	11	f1mn4	-0.16	0.28	0.58	0.19	3	26	f3ev5	7.27	0.47	0.16	0.23
1	12	f1mn5	-1.60	0.46	0.76	0.23	3	29	f3cm1	0.19	0.23	0.79	0.26
1	15	f1ig1	0.90	0.52	0.42	0.18	3	30	f3cm2	-0.48	0.18	1.20	0.34
1	16	f1ig2	-1.22	0.24	1.36	0.32	3	31	f3cm3	-1.24	0.21	2.05	0.60
1	18	f1ig3	1.53	0.62	0.51	0.20	3	32	f3cm4	-0.71	0.37	0.57	0.23
1	21	f1ig5	-0.08	0.16	1.15	0.26	4	1	f4ac1	-2.28	1.14	0.45	0.23
1	22	f1ev1	-1.12	0.30	0.94	0.24	4	2	f4ac2	-0.18	0.54	0.30	0.20
1	24	f1ev3	-1.41	0.21	2.38	0.70	4	3	f4ac3	-0.36	0.41	0.43	0.21
1	25	f1ev4	-0.69	0.14	2.08	0.47	4	4	f4ac4	0.79	0.35	0.65	0.23
1	29	f1cm1	-0.96	0.27	0.94	0.25	4	5	f4ac5	-0.14	0.23	0.75	0.24
1	30	f1cm2	0.24	0.32	0.50	0.19	4	8	f4mn1	-0.10	0.18	1.05	0.30
1	31	f1cm3	-1.13	0.32	0.85	0.23	4	9	f4mn2	-1.84	1.05	0.40	0.22
1	32	f1cm4	1.92	0.67	0.60	0.22	4	10	f4mn3	0.01	0.19	0.94	0.27
1	33	f1cm5	0.90	0.29	0.78	0.22	4	11	f4mn4	5.54	8.72	0.13	0.20
2	1	f2ac1	-2.86	2.83	0.19	0.19	4	12	f4mn5	3.21	2.30	0.31	0.23
2	2	f2ac2	-3.18	1.34	0.56	0.26	4	15	f4ig1	-1.04	0.29	1.03	0.30
2	3	f2ac3	3.12	3.01	0.20	0.19	4	16	f4ig2	0.41	0.21	0.92	0.27
2	4	f2ac4	-0.80	0.39	0.55	0.21	4	18	f4ig3	-2.61	1.16	0.55	0.26
2	5	f2ac5	3.55	2.31	0.34	0.22	4	19	f4ig4	1.47	0.84	0.40	0.21
2	8	f2mn1	-2.69	0.77	0.98	0.36	4	21	f4ig5	-0.20	0.22	0.80	0.24
2	9	f2mn2	0.98	0.39	0.66	0.23	4	22	f4ev1	-2.05	1.08	0.43	0.22
2	10	f2mn3	3.72	2.19	0.40	0.25	4	23	f4ev2	5.93	6.34	0.22	0.24
2	12	f2mn5	-1.67	0.79	0.44	0.20	4	24	f4ev3	-0.08	0.20	0.88	0.26
2	15	f2ig1	-2.57	0.70	1.04	0.37	4	25	f4ev4	-0.40	0.32	0.59	0.23
2	16	f2ig2	4.39	4.27	0.21	0.20	4	29	f4cm1	0.68	0.27	0.83	0.26
2	21	f2ig5	-1.49	0.37	0.97	0.28	4	30	f4cm2	2.65	1.03	0.70	0.31
2	22	f2ev1	-0.13	0.16	1.17	0.29	4	32	f4cm4	5.29	5.51	0.22	0.23
2	23	f2ev2	3.53	4.16	0.16	0.19	5	1	f5ac1	-2.75	3.83	0.15	0.20
2	24	f2ev3	-0.20	0.18	0.97	0.26	5	2	f5ac2	-0.53	0.21	0.98	0.28
2	25	f2ev4	1.74	0.71	0.56	0.23	5	3	f5ac3	-1.62	0.43	1.03	0.33
2	26	f2ev5	-0.71	0.20	1.20	0.32	5	5	f5ac5	0.07	0.40	0.40	0.21
2	29	f2cm1	-0.68	0.20	1.20	0.33	5	8	f5mn1	-1.26	0.38	0.85	0.27
2	30	f2cm2	0.84	0.53	0.44	0.21	5	9	f5mn2	3.80	2.14	0.51	0.30
2	32	f2cm4	-0.37	0.33	0.51	0.20	5	10	f5mn3	0.22	0.25	0.70	0.24
2	33	f2cm5	-1.19	0.26	1.26	0.34	5	11	f5mn4	0.44	0.27	0.72	0.24
3	2	f3ac2	-9.46	2.36	0.24	0.32	5	12	f5mn5	1.10	0.41	0.72	0.25

Form	Item no	Code	b	SE_b	a	SE_a	Form	Item no	Code	b	SE_b	a	SE_a
5	16	f5ig2	-0.12	0.26	0.65	0.23	8	4	f8ac4	0.05	0.25	0.67	0.21
5	18	f5ig3	-1.55	0.98	0.35	0.21	8	5	f8ac5	-0.37	0.19	1.03	0.26
5	19	f5ig4	1.07	0.33	0.93	0.29	8	10	f8mn3	-0.57	0.20	1.15	0.28
5	21	f5ig5	0.06	0.22	0.79	0.25	8	11	f8mn4	-0.13	0.18	1.01	0.26
5	24	f5ev3	1.40	0.76	0.44	0.22	8	15	f8ig1	-2.21	1.44	0.31	0.20
5	29	f5cm1	0.85	0.62	0.38	0.21	8	16	f8ig2	-0.44	0.29	0.65	0.21
5	30	f5cm2	-1.59	0.55	0.71	0.26	8	19	f8ig4	1.93	0.78	0.54	0.23
5	31	f5cm3	-1.43	0.57	0.60	0.24	8	21	f8ig5	-0.84	0.22	1.20	0.30
5	32	f5cm4	-1.29	0.26	1.62	0.47	8	22	f8ev1	1.96	0.94	0.44	0.21
6	1	f6ac1	-1.44	1.46	0.21	0.18	8	24	f8ev3	1.24	0.67	0.42	0.20
6	2	f6ac2	-1.21	0.74	0.38	0.19	8	25	f8ev4	-0.23	0.27	0.65	0.21
6	3	f6ac3	5.50	4.29	0.32	0.26	8	26	f8ev5	0.38	0.25	0.75	0.23
6	4	f6ac4	-3.85	2.02	0.47	0.26	8	29	f8cm1	-0.46	0.20	1.01	0.26
6	5	f6ac5	-0.05	0.24	0.71	0.22	8	30	f8cm2	-0.38	0.24	0.79	0.23
6	8	f6mn1	2.04	0.56	0.93	0.32	8	31	f8cm3	-0.37	0.44	0.40	0.19
6	11	f6mn4	0.04	0.21	0.84	0.23	8	32	f8cm4	-0.59	0.15	1.73	0.41
6	12	f6mn5	-4.76	6.68	0.14	0.19	8	33	f8cm5	3.82	2.73	0.29	0.21
6	15	f6ig1	1.86	1.63	0.22	0.19	9	1	f9ac1	0.60	0.45	0.45	0.23
6	16	f6ig2	1.06	0.63	0.38	0.19	9	2	f9ac2	-1.01	0.42	0.69	0.26
6	19	f6ig4	-0.11	0.17	1.15	0.27	9	4	f9ac4	3.85	5.20	0.16	0.22
6	21	f6ig5	-0.02	0.16	1.24	0.29	9	8	f9mn1	1.49	0.57	0.68	0.28
6	22	f6ev1	-0.73	0.24	1.04	0.26	9	9	f9mn2	-2.31	1.05	0.55	0.26
6	23	f6ev2	0.34	0.25	0.69	0.22	9	11	f9mn4	-0.46	1.77	0.10	0.20
6	24	f6ev3	3.00	1.19	0.67	0.30	9	12	f9mn5	1.18	0.42	0.76	0.28
6	25	f6ev4	-0.11	0.24	0.74	0.22	9	15	f9ig1	-0.76	0.87	0.26	0.21
6	26	f6ev5	0.51	0.25	0.75	0.23	9	16	f9ig2	-0.96	0.70	0.37	0.22
6	29	f6cm1	-0.33	0.31	0.59	0.21	9	18	f9ig3	1.85	1.02	0.43	0.24
6	30	f6cm2	-0.45	0.18	1.23	0.29	9	19	f9ig4	-0.69	0.21	1.28	0.36
6	31	f6cm3	-0.46	0.18	1.31	0.30	9	21	f9ig5	-2.18	0.86	0.67	0.29
6	32	f6cm4	-0.57	0.15	1.85	0.42	9	22	f9ev1	-0.05	0.22	0.79	0.26
6	33	f6cm5	0.31	0.26	0.67	0.21	9	23	f9ev2	-1.46	0.39	1.07	0.34
7	1	f7ac1	-1.69	0.75	0.50	0.22	9	24	f9ev3	-0.05	0.46	0.35	0.21
7	3	f7ac3	1.31	0.61	0.49	0.21	9	25	f9ev4	-1.73	0.70	0.62	0.26
7	4	f7ac4	-0.31	0.28	0.63	0.22	9	29	f9cm1	-1.25	1.02	0.31	0.22
7	5	f7ac5	-0.94	0.22	1.42	0.36	9	30	f9cm2	2.17	0.93	0.60	0.28
7	8	f7mn1	1.36	0.32	1.13	0.32	9	31	f9cm3	1.09	0.43	0.70	0.27
7	9	f7mn2	-1.02	0.22	1.52	0.38	9	33	f9cm5	1.17	1.19	0.24	0.21
7	11	f7mn4	1.10	0.31	0.93	0.27	10	1	f10ac1	-0.64	0.25	0.87	0.24
7	15	f7ig1	-1.56	0.55	0.66	0.24	10	2	f10ac2	3.86	2.49	0.36	0.23
7	16	f7ig2	0.77	0.30	0.75	0.24	10	3	f10ac3	-2.13	0.69	0.73	0.26
7	18	f7ig3	0.34	0.26	0.70	0.23	10	5	f10ac5	1.26	0.53	0.58	0.22
7	19	f7ig4	0.03	0.53	0.30	0.19	10	8	f10mn1	0.08	0.25	0.72	0.22
7	21	f7ig5	1.12	0.46	0.59	0.22	10	9	f10mn2	4.26	2.64	0.41	0.26
7	22	f7ev1	2.30	1.35	0.36	0.21	10	10	f10mn3	1.59	1.53	0.22	0.19
7	23	f7ev2	1.23	0.48	0.61	0.23	10	11	f10mn4	1.77	0.59	0.72	0.25
7	24	f7ev3	-0.83	0.18	1.68	0.42	10	12	f10mn5	0.16	0.21	0.91	0.24
7	29	f7cm1	1.56	0.74	0.46	0.21	10	18	f10ig3	-0.65	0.34	0.61	0.21
7	30	f7cm2	-0.59	0.25	0.87	0.25	10	19	f10ig4	-1.50	0.57	0.59	0.22
7	31	f7cm3	1.32	1.25	0.23	0.19	10	21	f10ig5	-0.45	0.39	0.48	0.20
7	32	f7cm4	-1.55	0.44	0.89	0.27	10	22	f10ev1	-0.07	0.19	1.02	0.26
8	2	f8ac2	-0.37	0.22	0.90	0.24	10	23	f10ev2	-0.78	0.24	1.00	0.27

Form	Item no	Code	b	SE_b	a	SE_a
10	24	f10ev3	-1.21	0.26	1.28	0.33
10	25	f10ev4	-1.15	0.26	1.20	0.31
10	26	f10ev5	-1.19	0.28	1.14	0.30
10	29	f10cm1	0.25	0.31	0.57	0.21
10	30	f10cm2	-1.20	0.54	0.52	0.21
10	31	f10cm3	-1.75	0.51	0.81	0.27
10	32	f10cm4	-1.02	0.27	1.01	0.27
10	33	f10cm5	-0.59	0.16	1.63	0.39
10	24	f10ev3	-1.21	0.26	1.28	0.33
10	25	f10ev4	-1.15	0.26	1.20	0.31
10	26	f10ev5	-1.19	0.28	1.14	0.30
10	29	f10cm1	0.25	0.31	0.57	0.21
10	30	f10cm2	-1.20	0.54	0.52	0.21
10	31	f10cm3	-1.75	0.51	0.81	0.27
10	32	f10cm4	-1.02	0.27	1.01	0.27
10	33	f10cm5	-0.59	0.16	1.63	0.39

Summary

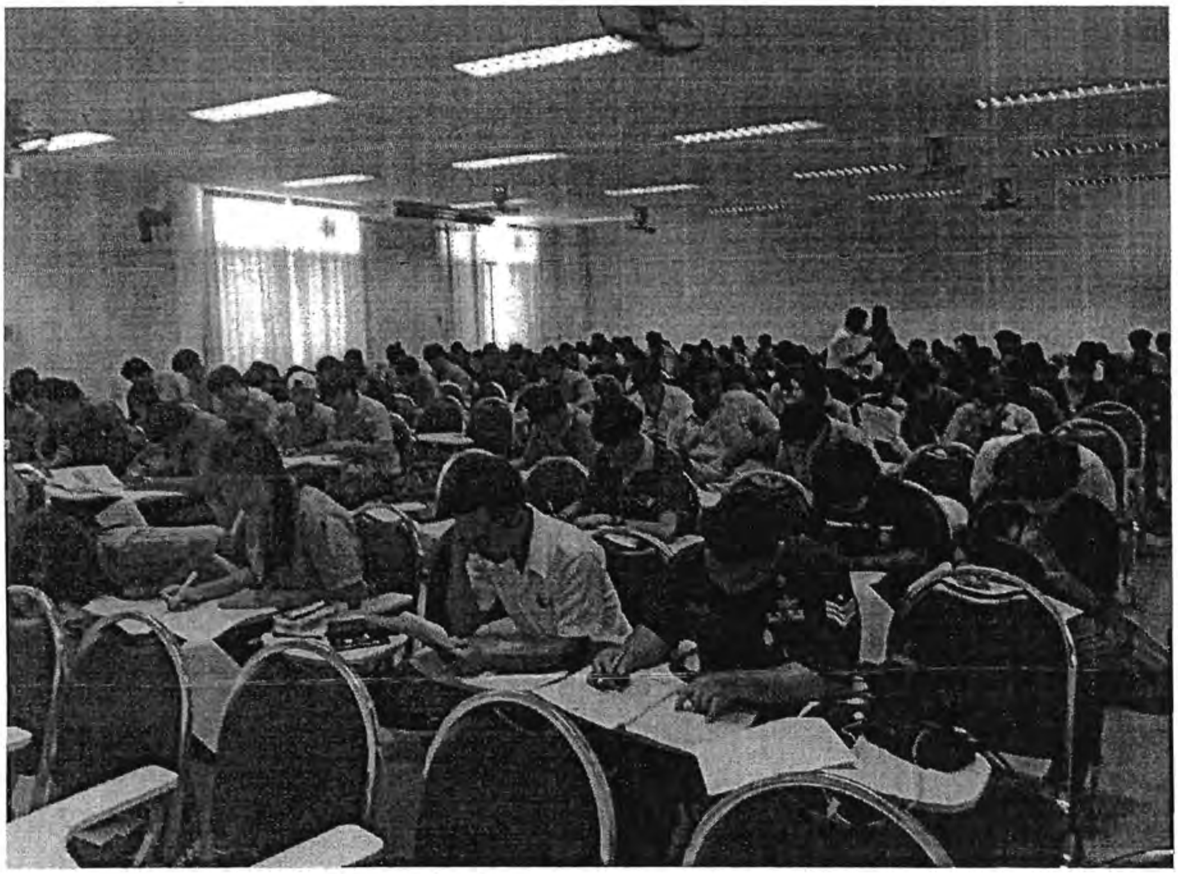
Remaining items	Total
D1: Accessibility	43
D2: Management	41
D3: Integration	43
D4: Evaluation	40
D5: Communication	45
Total	212 items

Descriptive stat of b and a parameters

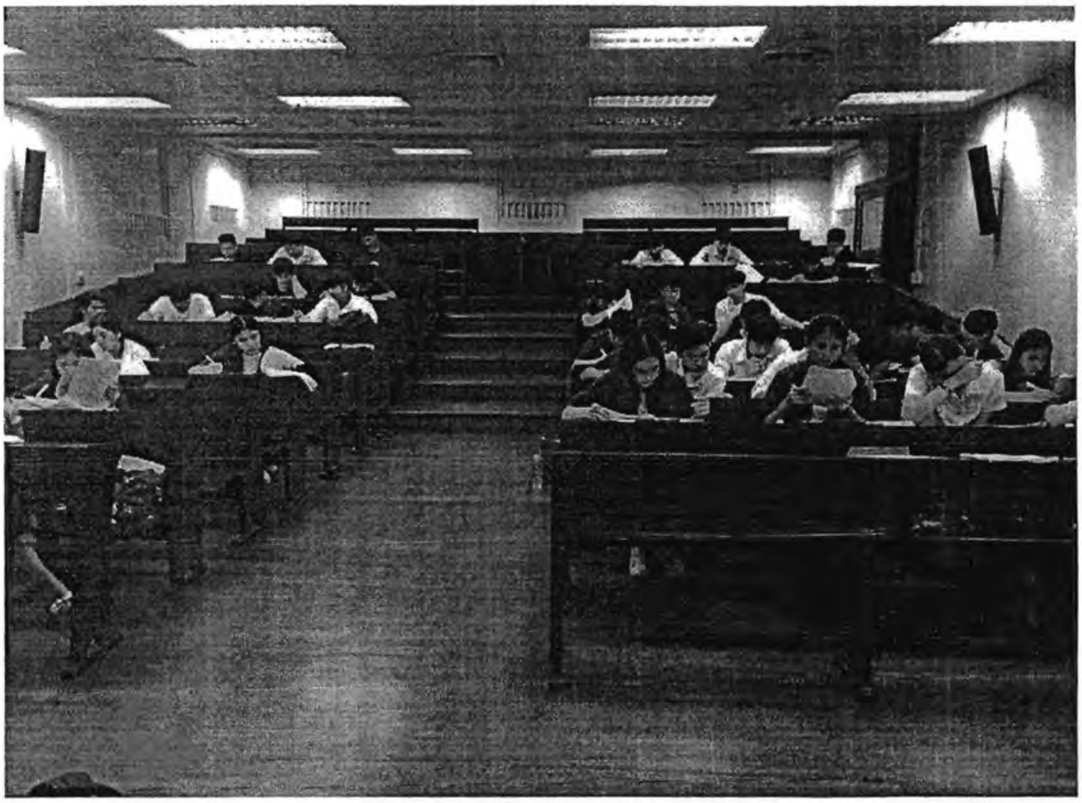
	b (item difficulty)	a (Item discrimination)
mean	0.19	0.70
sd	2.06	0.39
max	7.27	2.38
min	-9.46	0.09

ภาคผนวก 8

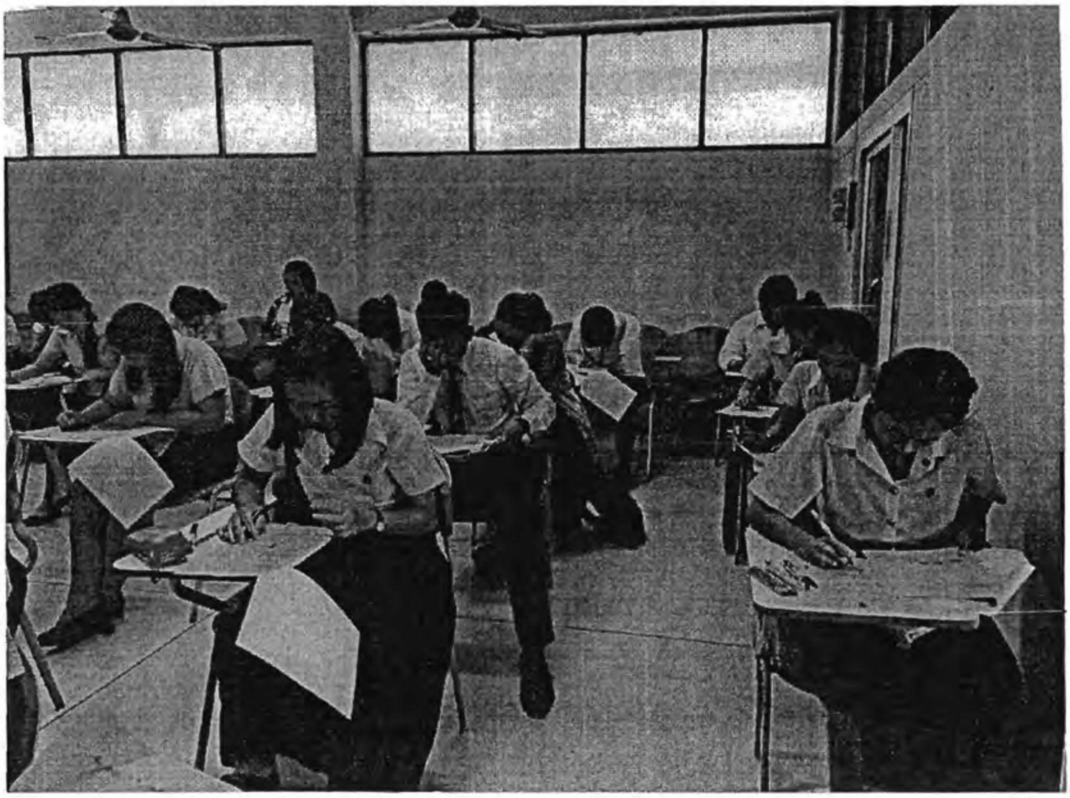
ภาพการเก็บข้อมูลระยะการพัฒนาข้อสอบ



มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ



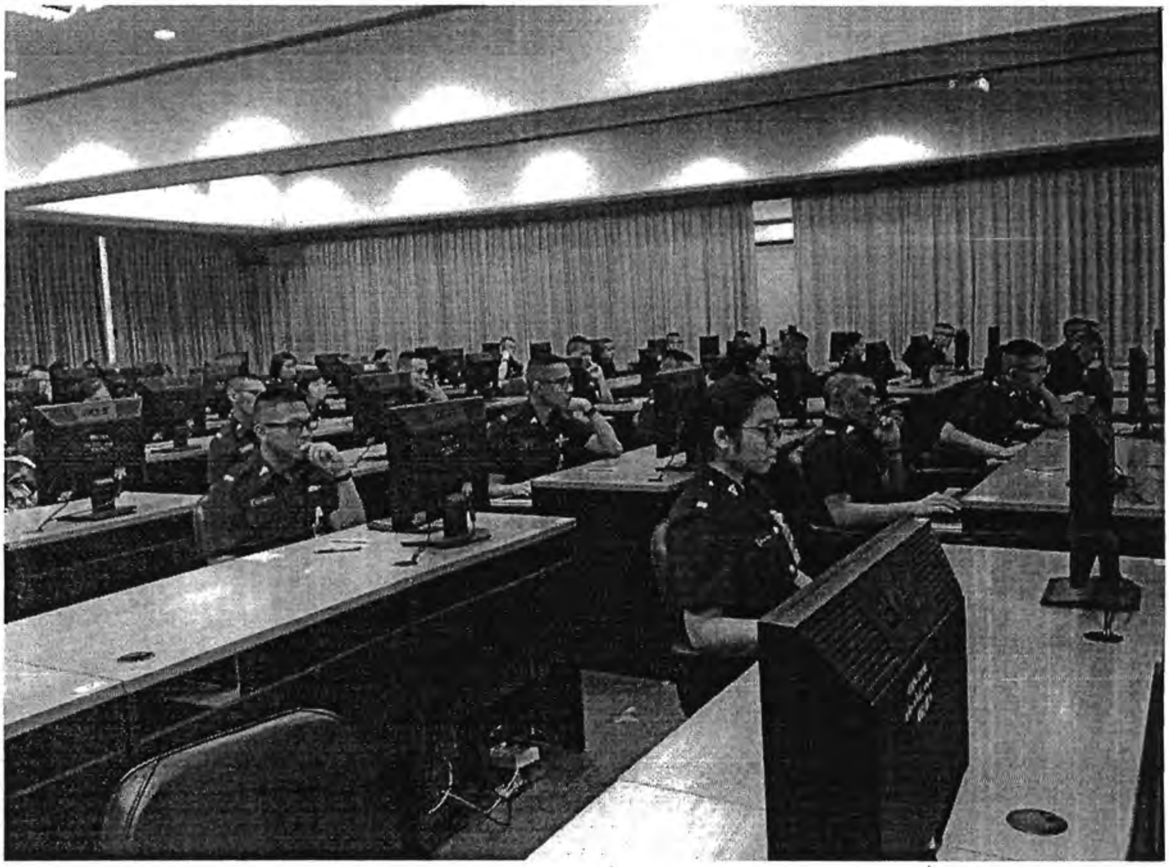
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย



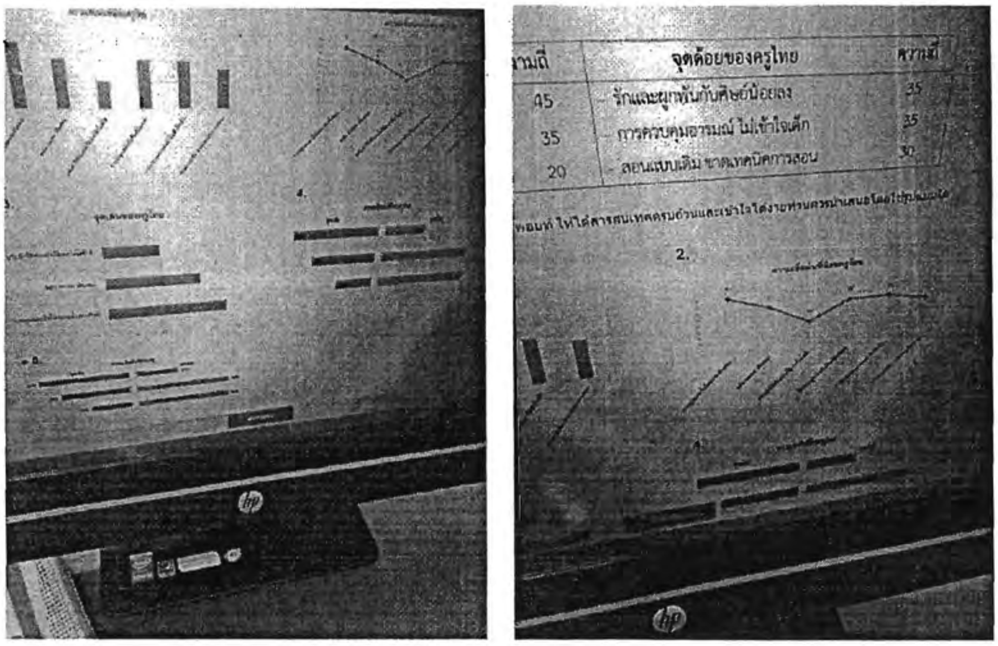
คณะศึกษาศาสตร์ มหาวิทยาลัยราชภัฏวไลยอลงกรณ์

ภาคผนวก 9

ภาพการเก็บข้อมูลระยะการนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการวัดทักษะด้านเทคโนโลยีสารสนเทศและการสื่อสารแห่งศตวรรษที่ 21 ไปทดลองใช้



วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้า



ภาพหน้าจอระบบการทดสอบที่ใช้จริง

รายนามคณะผู้วิจัย

1. รองศาสตราจารย์ ดร.โชติกา	ภาษีผล	หัวหน้าโครงการวิจัย อาจารย์ประจำภาควิชาวิจัยและจิตวิทยาการศึกษา สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. ดร.ณภัทร	ชัยมงคล	ผู้ช่วยนักวิจัย นักเทคโนโลยีสารสนเทศ อาวุโส สถาบันระหว่างประเทศเพื่อการค้าและการพัฒนา (สคพ.)
3. อาจารย์ ดร. จุฑาภรณ์	มาสันเทียะ	ผู้ช่วยนักวิจัย อาจารย์ประจำภาควิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง
4. นายวีรภัทร์	สุขศิริ	ผู้ช่วยนักวิจัย Doctoral student Quantitative Methods and Evaluation Graduate School of Education, University of California, Berkeley
5. นายกิตติทัศน์	หวานฉ่ำ	ผู้ช่วยนักวิจัย นิสิตระดับดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
6. นางสาวรังสิมาภรณ์	หนูน้อย	ผู้ช่วยนักวิจัย นิสิตระดับดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย