

Development of Computer Software for Criterion-referenced Item Analysis

Nhabhat Chaimongkol¹

Received: May 8, 2026 – Revised: June 25, 2026 – Accepted: June 25, 2026

Abstract

This research aimed to develop and assess the quality of computer software for criterion-referenced item analysis, designed to align with Thailand's standards-based basic education curriculum. A research and development (R&D) methodology was employed, comprising two phases: Phase 1, software development, and Phase 2, software quality assessment. The research instruments included a heuristic evaluation checklist for assessing the appropriateness of the software interface and a user experience questionnaire. The participants who tested the software consisted of 50 users. The results revealed the following: 1) The developed software could serve as a replacement for existing software whose analytical outputs may be difficult to interpret for users without a background in educational measurement and evaluation, while also being compatible with current operating systems. The software developed in this study, named i-ACT (Item Analysis for Criterion-Referenced Test), was developed using the Visual Basic programming language and can be used without installation (portable software). The key item analysis results generated by the software include the difficulty index, discrimination index, and reliability coefficient, based on criterion-referenced test analysis concepts. 2) Regarding software quality assessment, users reported a positive experience with the software. Users rated the software highly in terms of performance efficiency ($M = 6.26$, $SD = 0.89$) and ease of learning ($M = 6.24$, $SD = 0.71$). In addition, users perceived the developed software as appealing ($M = 6.20$, $SD = 0.81$).

Keywords: Software Development, Criterion-referenced Test, Item Analysis, Standards-based Curriculum

¹ Corresponding Author,

Department of Educational Research and Psychology, Faculty of Education, Chulalongkorn University, Bangkok 10330, Thailand.
nhabhat.c@chula.ac.th

การพัฒนาซอฟต์แวร์คอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบ แบบอิงเกณฑ์

ณภัทร ชัยมงคล¹

รับต้นฉบับ : 8 พฤษภาคม 2569 – รับแก้ไข : 25 มิถุนายน 2569 – ตอบรับตีพิมพ์ : 25 มิถุนายน 2569

บทคัดย่อ

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาและประเมินคุณภาพซอฟต์แวร์คอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ที่สอดคล้องกับหลักสูตรอิงมาตรฐานของการศึกษาขั้นพื้นฐานในประเทศไทย โดยใช้ระเบียบวิธีวิจัยและพัฒนา การวิจัยแบ่งออกเป็น 2 ระยะ ได้แก่ ระยะที่ 1 การพัฒนาซอฟต์แวร์ และระยะที่ 2 การประเมินคุณภาพซอฟต์แวร์ เครื่องมือที่ใช้ในการวิจัย ประกอบด้วย แบบประเมินอิวิริสติกสำหรับประเมินความเหมาะสมของอินเตอร์เฟซของซอฟต์แวร์ และแบบประเมินประสบการณ์ผู้ใช้ ตัวอย่างวิจัยที่ทดลองใช้งานซอฟต์แวร์มีจำนวน 50 คน ผลการวิจัยพบว่า 1) ซอฟต์แวร์ที่พัฒนาขึ้นสามารถทดแทนซอฟต์แวร์เดิมที่การแปลผลการวิเคราะห์อาจเข้าใจยากสำหรับผู้ที่ไม่มีความรู้ด้านการวัดและการประเมินผลการศึกษา รวมถึงรองรับระบบปฏิบัติการคอมพิวเตอร์ใหม่ในปัจจุบัน ซอฟต์แวร์ที่พัฒนาขึ้นในการวิจัยครั้งนี้ชื่อว่า i-ACT (Item Analysis for Criterion-Referenced Test) เป็นซอฟต์แวร์ที่พัฒนาขึ้นด้วยภาษา Visual Basic สามารถใช้งานได้โดยไม่ต้องติดตั้งลงบนเครื่องคอมพิวเตอร์ (portable software) ผลการวิเคราะห์ข้อสอบที่สำคัญจากซอฟต์แวร์ประกอบด้วย ค่าความยาก ค่าอำนาจจำแนก และค่าความเที่ยงตามแนวคิดการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ และ 2) การประเมินคุณภาพซอฟต์แวร์ ผู้ใช้งานมีประสบการณ์ที่ดีต่อการใช้งานโดยซอฟต์แวร์ที่พัฒนาขึ้นมีประสิทธิภาพในการทำงานเชิงบวก ($M = 6.26, SD = 0.89$) และผู้ใช้งานสามารถเรียนรู้วิธีการใช้งานซอฟต์แวร์ได้ง่าย ($M = 6.24, SD = 0.71$) นอกจากนี้ ผู้ใช้งานเห็นว่าซอฟต์แวร์ที่พัฒนาขึ้นมีความน่าสนใจ ($M = 6.20, SD = 0.81$)

คำสำคัญ : การพัฒนาซอฟต์แวร์, การทดสอบแบบอิงเกณฑ์, การวิเคราะห์ข้อสอบ, หลักสูตรอิงมาตรฐาน

¹ ผู้รับผิดชอบบทความหลัก,

ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพฯ 10330
nhabhat.c@chula.ac.th

บทนำ

ภายใต้บริบททางการศึกษาของประเทศไทย แนวทางการวัดและประเมินผลการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 กำหนดให้ผู้เรียนจะต้องมีความรู้รอบด้านตามมาตรฐานการเรียนรู้และตัวชี้วัดที่กำหนดในแต่ละกลุ่มสาระ จำนวน 8 กลุ่ม อีกทั้ง ต้องมีความสามารถในการอ่าน เขียน คิดวิเคราะห์ มีคุณลักษณะอันพึงประสงค์ และเข้าร่วมกิจกรรมเพื่อพัฒนาผู้เรียนตามที่หลักสูตรกำหนด (สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ, 2557) โดยเครื่องมือที่มีการนำมาใช้ในการประเมินผู้เรียนกันอย่างกว้างขวาง คือ ข้อสอบแบบหลายตัวเลือก (multiple choice) เนื่องจากสามารถสร้างให้ครอบคลุมเนื้อหาได้อย่างมีประสิทธิภาพและนำไปพัฒนาให้เป็นแบบสอบมาตรฐานได้ (ศิริชัย กาญจนวาสี, 2556) ดังนั้น เพื่อให้แบบสอบมีคุณภาพและเป็นที่ยอมรับจึงต้องมีการตรวจสอบคุณภาพข้อสอบรายข้อเพื่อให้มั่นใจได้ว่าข้อสอบมีมาตรฐานตามหลักการ

การตรวจสอบคุณภาพของข้อสอบด้วยการวิเคราะห์ข้อสอบ (item analysis) ตามทฤษฎีการทดสอบแบบดั้งเดิมมีตัวบ่งชี้คุณภาพที่สำคัญ ได้แก่ ค่าความยาก (item difficulty) และอำนาจจำแนก (item discrimination) ซึ่งในการวิเคราะห์ข้อสอบแบบหลายตัวเลือกสามารถจำแนกออกเป็น 2 ประเภท คือ สำหรับแบบสอบอิงกลุ่ม (norm-referenced test) และสำหรับแบบสอบอิงเกณฑ์ (criterion-referenced test) (Latif et al., 2016; ศิริชัย กาญจนวาสี, 2556) อย่างไรก็ตาม ในการตัดสินผลการเรียนตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 ใช้แนวทางการตัดสินผู้เรียนแบบอิงเกณฑ์ กล่าวคือ ผู้เรียนต้องผ่านรายวิชาไม่น้อยกว่าร้อยละ 50 และมีการกำหนดระดับผลการเรียนในระบบต่าง ๆ เช่น ระบบตัวเลข ระบบตัวอักษร ระบบร้อยละ (สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ, 2557)

จากแนวคิดดังกล่าวจึงได้มีผู้พัฒนาซอฟต์แวร์วิเคราะห์ข้อสอบแบบอิงเกณฑ์ขึ้นในประเทศไทยหรือเรียกว่า ซอฟต์แวร์ B-Index เพื่อช่วยอำนวยความสะดวกให้แก่ครูผู้สอนในการวิเคราะห์คุณภาพแบบสอบโดยใช้วิธีวิเคราะห์ข้อสอบแบบอิงเกณฑ์ด้วยดัชนีจำแนก B (B-Index) โดยสามารถจำแนกผู้สอบได้วาระอบรู้หรือไม่รอบรู้เป็นรายจุดประสงค์ ซึ่งมีการกำหนดเกณฑ์การผ่านจุดประสงค์ร้อยละ 50 เป็นค่าเริ่มต้น สอดคล้องกับการตัดสินผลการเรียนตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 นอกจากนี้ ยังสามารถนำไปใช้วิเคราะห์กับแบบสอบแบบอิงกลุ่มได้ในกรณีที่มีการคัดเลือกผู้สอบกลุ่มเก่งและกลุ่มอ่อนจำนวนร้อยละ 27 ทั้งสองกลุ่มตามแนวทางการประเมินแบบอิงกลุ่ม และผู้สอบกลุ่มอ่อนไม่ผ่านเกณฑ์ประเมิน (สาคร แสงผิง, 2547) ซอฟต์แวร์ดังกล่าวได้รับการเผยแพร่และนำไปใช้ในการวิเคราะห์ข้อสอบแบบอิงเกณฑ์อย่างแพร่หลายในประเทศไทย อย่างไรก็ตาม ซอฟต์แวร์ดังกล่าวรองรับการทำงานบนระบบปฏิบัติการ Windows 98, NT, 2000 และ XP ปัจจุบันระบบปฏิบัติการได้มีการพัฒนาเวอร์ชันใหม่ เช่น Windows 7, 8 และ 10 ทำให้บางฟังก์ชันไม่สามารถใช้งานได้โดยมีประสิทธิภาพ อีกทั้งการนำเข้าข้อมูลของซอฟต์แวร์ดังกล่าวต้องกรอกข้อมูลผลการตอบของผู้สอบรายบุคคลซึ่งไม่สะดวกต่อการใช้งาน จากข้อจำกัดดังกล่าวได้นักวิชาการพยายามพัฒนาฟอร์มจาก Microsoft Excel เพื่อใช้ในการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ แต่ขั้นตอนการใช้งานอาจยังมีข้อจำกัดหลายส่วน รวมถึงรายละเอียดการแปลผลและการนำผลการวิเคราะห์ไปใช้งาน

ด้วยเหตุนี้ ผู้วิจัยจึงมีแนวคิดในการพัฒนาซอฟต์แวร์คอมพิวเตอร์สำหรับวิเคราะห์ข้อสอบแบบอิงเกณฑ์ที่สอดคล้องกับหลักสูตรอิงมาตรฐานของการศึกษาขั้นพื้นฐานในประเทศไทย เพื่อลดข้อจำกัดของซอฟต์แวร์วิเคราะห์ข้อสอบแบบอิงเกณฑ์ดั้งเดิมที่อาจไม่ได้รับการอัปเดตให้สอดคล้องกับเทคโนโลยีที่เปลี่ยนไปในปัจจุบัน เนื่องจากการวัดและประเมินผลตามหลักสูตรการศึกษาขั้นพื้นฐานของประเทศไทยพิจารณาผลลัพธ์การเรียนรู้แบบอิงเกณฑ์ การวิเคราะห์คุณภาพข้อสอบรายข้อจึงควรใช้การวิเคราะห์ข้อสอบแบบอิงเกณฑ์เพื่อให้ความสอดคล้องกัน นอกจากนี้ ซอฟต์แวร์คอมพิวเตอร์ที่ผู้วิจัยพัฒนายังสามารถรองรับการทำงานบนระบบปฏิบัติการในปัจจุบัน เช่น Microsoft Windows 10 และสามารถให้สารสนเทศจากผลการวิเคราะห์ข้อสอบ เช่น ความยาก อำนาจจำแนก และส่งออก (export) ข้อมูลจากผลการวิเคราะห์ เพื่อนำไปใช้ประโยชน์ในส่วนที่เกี่ยวข้องต่อไป โดยผู้ใช้งานสามารถนำผลการวิเคราะห์ดังกล่าวไปใช้ในการพัฒนาข้อสอบและแบบสอบให้มีคุณภาพมากยิ่งขึ้น รวมถึงสามารถนำผลการวิเคราะห์ของผู้เรียนไปใช้ในการพัฒนาผู้เรียนรายบุคคลให้มีคุณภาพมากยิ่งขึ้นต่อไป

วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาซอฟต์แวร์คอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ที่สอดคล้องกับหลักสูตรอิงมาตรฐานของการศึกษาขั้นพื้นฐานในประเทศไทย
2. เพื่อประเมินคุณภาพซอฟต์แวร์คอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ที่สอดคล้องกับหลักสูตรอิงมาตรฐานของการศึกษาขั้นพื้นฐานในประเทศไทย

วิธีดำเนินการวิจัย

การวิจัยในครั้งนี้ใช้ระเบียบวิธีวิจัยและพัฒนา (research and development) แบ่งการศึกษาออกเป็น 2 ระยะ ได้แก่ ระยะที่ 1 การพัฒนาซอฟต์แวร์ และระยะที่ 2 การประเมินคุณภาพซอฟต์แวร์ โดยมีรายละเอียดแต่ละระยะดังนี้

ระยะที่ 1 การพัฒนาซอฟต์แวร์

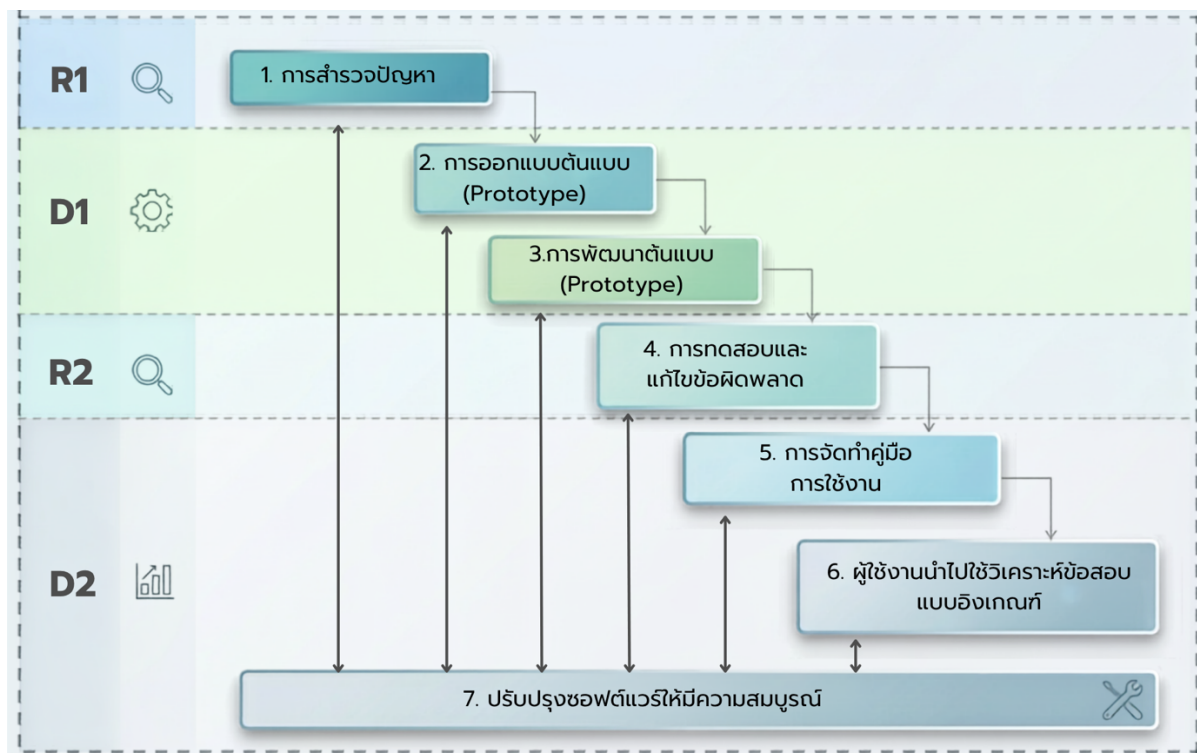
ผู้วิจัยได้ดำเนินการศึกษาเอกสารที่เกี่ยวข้องและสามารถสรุปประเด็นสำคัญในการพัฒนาซอฟต์แวร์คอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ที่สอดคล้องกับหลักสูตรอิงมาตรฐานของการศึกษาขั้นพื้นฐานในประเทศไทยออกเป็น 3 ประเด็น ดังนี้ (1) สามารถใช้งานในระบบออนไลน์และรองรับระบบปฏิบัติการ Windows 10 ขึ้นไป (2) สามารถวิเคราะห์คุณภาพข้อสอบรายข้อและทั้งฉบับตามแนวคิดของการวิเคราะห์แบบสอบอิงเกณฑ์ และ (3) สามารถรายงานผลการวิเคราะห์ในภาพรวมและรายบุคคลพร้อมส่งออก (export) ผลการวิเคราะห์ในรูปแบบไฟล์มาตรฐาน PDF

สำหรับการพัฒนาซอฟต์แวร์คอมพิวเตอร์ในการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ที่สอดคล้องกับหลักสูตรอิงมาตรฐานของการศึกษาขั้นพื้นฐานในประเทศไทยพัฒนาด้วยภาษา Visual Basic เนื่องจากเป็นภาษาที่ไม่ซับซ้อนใช้งานง่ายสะดวกต่อการปรับแก้ไขหากเกิดความผิดพลาด (debug) ของซอฟต์แวร์ และสามารถใช้งานได้ดีกับระบบปฏิบัติการ Windows ซึ่งเป็นระบบปฏิบัติการที่ผู้ใช้งานในประเทศไทยส่วนใหญ่เลือกใช้งาน อีกทั้งการปรับแก้ไขหน้าจอรวมถึง

ผลการวิเคราะห์สามารถดำเนินการได้อย่างรวดเร็ว และยังสามารถเชื่อมโยงการใช้งานกับซอฟต์แวร์อื่น เช่น Microsoft Excel สำหรับกระบวนการในการพัฒนาและทดสอบซอฟต์แวร์มีรายละเอียดดังนี้

ภาพ 1

ขั้นตอนการพัฒนาซอฟต์แวร์



ภาพ 1 แสดงให้เห็นขั้นตอนการพัฒนาซอฟต์แวร์โดยแบ่งออกเป็น 7 ขั้นตอน ตามแนวคิดวงจรการพัฒนาระบบ (System Development Life Cycle: SDLC) ซึ่งช่วยให้การพัฒนาซอฟต์แวร์มีการวิเคราะห์และออกแบบอย่างเป็นระบบ อีกทั้งยังช่วยให้ผู้พัฒนาซอฟต์แวร์ได้ดำเนินการตรวจสอบและปรับปรุงการใช้งานให้มีความเหมาะสมและสอดคล้องกับการใช้งานจริง (Conger, 2011; Rao et al., 2025) การพัฒนาซอฟต์แวร์เริ่มต้นจากขั้นที่ 1 การวิเคราะห์ปัญหาและประเมินข้อจำกัดของซอฟต์แวร์การวิเคราะห์คุณภาพข้อสอบแบบอิงเกณฑ์ดั้งเดิม ซึ่งเป็นการวิจัยขั้นต้นหรือสำรวจปัญหา (R1) จากนั้น ขั้นที่ 2 และ 3 ดำเนินการออกแบบและพัฒนาต้นแบบ (prototype) ซอฟต์แวร์เพื่อแก้ไขข้อจำกัดในขั้นที่ 1 ซึ่งเป็นขั้นการวิจัยเพื่อพัฒนาต้นแบบ (D1)

ขั้นที่ 4 ดำเนินการทดสอบและแก้ไขข้อผิดพลาดของซอฟต์แวร์ก่อนนำไปทดลองใช้ โดยประเมินความถูกต้องและเหมาะสมในการวิเคราะห์ ซึ่งเป็นขั้นการวิจัยและทดลองใช้งานต้นแบบเพื่อปรับปรุงแก้ไขข้อจำกัดต่าง ๆ (R2) ในขั้นตอนที่ 5 6 และ 7 เป็นการจัดทำคู่มือการใช้งานซอฟต์แวร์และนำซอฟต์แวร์ไปให้ผู้ใช้งานใช้ในการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ หลังจากนั้นนำข้อคิดเห็นจากผู้ใช้งานมาปรับปรุงซอฟต์แวร์ตามข้อสังเกต ซึ่งเป็นขั้นการพัฒนาและปรับปรุงซอฟต์แวร์ให้สมบูรณ์มากยิ่งขึ้นและเตรียมพร้อมเผยแพร่และขยายผลการใช้งานต่อไป (D2)

ระยะที่ 2 การประเมินคุณภาพซอฟต์แวร์

การประเมินคุณภาพซอฟต์แวร์แบ่งออกเป็น 2 ส่วน ได้แก่ (1) การประเมินคุณภาพซอฟต์แวร์ก่อนนำไปทดลองใช้งานด้วยการประเมินอินเทอร์เฟซของซอฟต์แวร์ (User Interface Evaluation: UI) จากแบบประเมินฮิวริสติก ประกอบด้วยรายการประเมิน 10 รายการ ในรูปแบบมาตราประมาณค่า 5 ระดับ ตั้งแต่ 1 (ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง) ถึง 5 (มีความจำเป็นสูงมาก / ควรปรับปรุงแก้ไขทันที) โดยให้ผู้เชี่ยวชาญประเมินความถูกต้องและเหมาะสมของซอฟต์แวร์ก่อนนำไปทดลองใช้ ซึ่งประกอบไปด้วยผู้เชี่ยวชาญ 5 ท่าน แบ่งเป็นผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา จำนวน 3 ท่าน และผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศ จำนวน 2 ท่าน หลังจากนั้นนำผลการประเมินและข้อสังเกตมาปรับปรุงซอฟต์แวร์ให้มีคุณภาพมากยิ่งขึ้น หลังจากปรับปรุงคุณภาพซอฟต์แวร์ตามข้อสังเกตเรียบร้อยแล้วจึงดำเนินการจัดทำคู่มือการใช้งานซอฟต์แวร์เพื่อให้ตัวอย่างวิจัยได้ทดลองใช้งานต่อไป และ (2) การประเมินคุณภาพซอฟต์แวร์หลังการนำไปใช้งานด้วยการประเมินประสบการณ์ผู้ใช้ (user experience evaluation: UX) จากแบบประเมินจำนวน 19 รายการในรูปแบบมาตราประมาณค่าความหมายตามแนวคิดของออสกูด จำนวน 7 ระดับ การประเมินประกอบด้วยตัวอย่างวิจัยที่เป็นอาสาสมัครในการทดลองใช้งานซอฟต์แวร์ จำนวน 50 คน ซึ่งเป็นบุคลากรทางการศึกษาในสังกัดคณะกรรมการการศึกษาขั้นพื้นฐาน แบ่งเป็นครูผู้สอนระดับประถมศึกษา จำนวน 25 คน และครูผู้สอนระดับมัธยมศึกษา จำนวน 25 คน

แบบประเมินทั้งสองส่วนผ่านการประเมินคุณภาพด้วยการตรวจสอบความตรงเชิงเนื้อหาโดยใช้ความคิดเห็นของผู้เชี่ยวชาญ จำนวน 5 ท่าน ผลการประเมิน พบว่า ผู้เชี่ยวชาญทุกท่านเห็นไปในทิศทางเดียวกันว่า แบบประเมินฮิวริสติกสามารถวัดได้ตรงตามสิ่งที่มุ่งวัดหรือคิดเป็นร้อยละ 100 สำหรับแบบประเมินประสบการณ์ผู้ใช้ผู้เชี่ยวชาญเห็นไปในทิศทางเช่นเดียวกันว่าสามารถวัดได้ตรงตามสิ่งที่มุ่งวัด โดยมีข้อสังเกตจำนวน 3 ขอรายการ คือ 1 8 และ 11 ที่ควรปรับปรุงคำคุณศัพท์ให้สอดคล้องกับนิยามมากยิ่งขึ้นก่อนนำไปใช้งาน ผู้วิจัยจึงได้ดำเนินการปรับแก้คำคุณศัพท์ที่ใช้ในการประเมินตามข้อสังเกตของผู้เชี่ยวชาญก่อนนำแบบประเมินไปใช้งาน

ผลการวิจัย

ผลการวิจัยสามารถแบ่งได้เป็น 2 ตอน ตามระยะการวิจัย ดังนี้

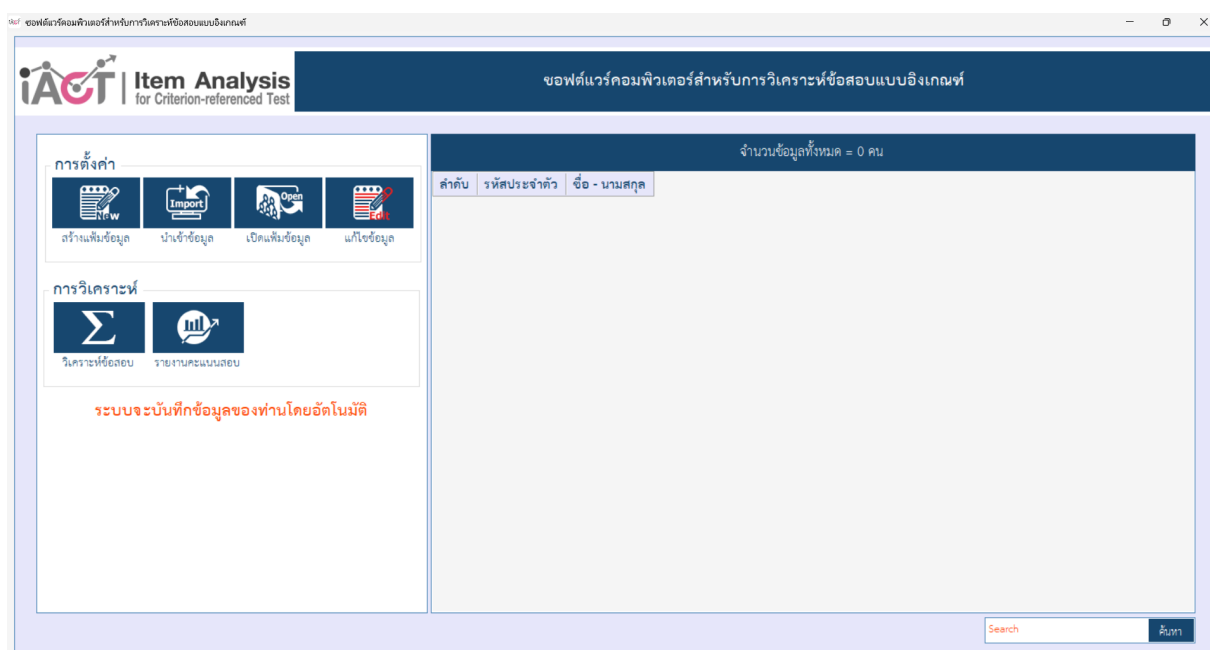
ตอนที่ 1 ผลการพัฒนาซอฟต์แวร์

ผลการพัฒนาซอฟต์แวร์ตามแนวคิดวงจรการพัฒนาระบบมีดังนี้ ขั้นที่ 1 ผลการวิเคราะห์ปัญหาพบว่า ปัจจุบันซอฟต์แวร์ที่ใช้วิเคราะห์ข้อสอบแบบอิงเกณฑ์ในประเทศไทยปัจจุบันไม่รองรับระบบปฏิบัติการเวอร์ชันใหม่ อีกทั้ง การนำเข้าและแปลความผลการวิเคราะห์มีความซับซ้อนทำให้เข้าใจยากสำหรับผู้ที่ไม่มีความรู้ด้านการวัดและประเมินผลการศึกษา นอกจากนี้ ซอฟต์แวร์การวิเคราะห์ในต่างประเทศต้องเสียค่าใช้จ่ายเป็นรายเดือนหรือรายปี และมีข้อจำกัดทางภาษาแก่ผู้ใช้งานในประเทศไทยที่ไม่เชี่ยวชาญภาษาอังกฤษ ซอฟต์แวร์ที่พัฒนาขึ้นจึงลดข้อจำกัดด้านความซับซ้อนในการใช้งาน การแสดงผลอยู่ในรูปแบบภาษาไทยที่เข้าใจง่าย รองรับระบบปฏิบัติการเวอร์ชันใหม่และเป็นซอฟต์แวร์ที่เปิดให้ใช้งานฟรีโดยไม่เสียค่าใช้จ่าย

ขั้นที่ 2 ผลการออกแบบหน้าจอและฟังก์ชันสอดคล้องกับผลการวิเคราะห์ปัญหา โดยซอฟต์แวร์เน้นการแสดงผลในรูปแบบภาษาไทยที่เอื้ออำนวยต่อการใช้งาน สามารถนำเข้าข้อมูล (import) ในการวิเคราะห์ทำให้เกิดความสะดวกรวดเร็ว ไม่จำเป็นต้องเขียนคำสั่งในการวิเคราะห์ อีกทั้ง ผลการวิเคราะห์อยู่ในรูปแบบรายงานแปลผลที่เข้าใจง่ายและสามารถส่งออก (export) ผลการวิเคราะห์ในรูปแบบไฟล์มาตรฐาน PDF ขั้นที่ 3 ผลการพัฒนาซอฟต์แวร์แสดงตัวอย่างหน้าจอต่างภาพ 2-8 ขั้นที่ 4 ผลการทดสอบและแก้ไขข้อผิดพลาดของซอฟต์แวร์ก่อนนำไปทดลองใช้จะนำเสนอรายละเอียดในส่วนที่ 2 การประเมินคุณภาพซอฟต์แวร์ส่วนการประเมินอินเตอร์เฟซของซอฟต์แวร์ต่อไป ขั้นที่ 5 ผลการพัฒนาคู่มือการใช้งานซอฟต์แวร์ซึ่งได้มีการพัฒนาคู่มือการใช้งานให้สอดคล้องกับผลการทดสอบการใช้งานซอฟต์แวร์และแก้ไขข้อผิดพลาดของซอฟต์แวร์ก่อนนำไปให้ตัวอย่างวิจัยทดลองใช้ต่อไป ขั้นที่ 6 การนำซอฟต์แวร์ไปทดลองใช้กับตัวอย่างวิจัยที่เป็นอาสาสมัครในการทดลองใช้งานซอฟต์แวร์ จำนวน 50 คน ผลการประเมินจะนำเสนอในส่วนที่ 2 การประเมินคุณภาพซอฟต์แวร์ส่วนการประเมินประสิทธิภาพผู้ใช้อย่างต่อไป และขั้นที่ 7 ผลการปรับปรุงซอฟต์แวร์ให้มีความสมบูรณ์มากยิ่งขึ้นดำเนินการหลังจากการนำซอฟต์แวร์ไปทดลองใช้กับตัวอย่างวิจัย โดยซอฟต์แวร์มีหน้าจอในการทำงานหลัก ดังนี้

ภาพ 2

หน้าจอหลักของซอฟต์แวร์คอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบแบบอิงเกณฑ์



ภาพ 2 แสดงให้เห็นถึงหน้าจอหลักของซอฟต์แวร์คอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ (Item Analysis for Criterion-referenced Test: i-ACT) ซึ่งเป็นการออกแบบหน้าจอที่เรียบง่ายสะดวกกับผู้ใช้งานตามหลักของการออกแบบส่วนเชื่อมต่อผู้ใช้งาน (user interface: UI) ที่ดี โดยในการวิเคราะห์ข้อมูลแบ่งขั้นตอนในการวิเคราะห์ออกเป็น 2 ส่วน ได้แก่ ส่วนที่ 1 การตั้งค่า และส่วนที่ 2 การวิเคราะห์ข้อมูล ก่อนเริ่มใช้งานการวิเคราะห์ต้องเริ่มต้นจากการตั้งค่าและการจัดเตรียมข้อมูลสำหรับการวิเคราะห์ โดยทำการสร้างแฟ้มข้อมูลเพื่อกำหนดเป็น

ค่าเริ่มต้นในการวิเคราะห์ ผู้ใช้งานสามารถเลือกได้ว่าต้องการวิเคราะห์ข้อสอบประเภทปรนัยหรือ
อัตนัย กรณีที่ผู้ใช้งานเลือกวิเคราะห์ข้อสอบแบบปรนัย ผู้ใช้งานจะต้องกรอกข้อมูลเบื้องต้นที่
จำเป็นในการวิเคราะห์ ได้แก่ ชื่อแฟ้มข้อมูล ชื่อวิชา รหัสวิชา ปีการศึกษา จำนวนข้อ จำนวน
วัตถุประสงค์ เฉลย การตั้งค่าวัตถุประสงค์ ประเภทคำตอบ ซึ่งมีให้เลือก 3 รูปแบบ ได้แก่ 1 2 3...
ก ข ค... และ A B C... และจำนวนตัวเลือก ซึ่งกำหนดให้เลือกตั้งแต่ 2 ถึง 5 ตัวเลือก

สำหรับการกำหนดเฉลย ผู้ใช้งานต้องกำหนดรูปแบบเฉลยคำตอบของข้อสอบแต่ละข้อตาม
รูปแบบที่ซอฟต์แวร์กำหนด จากนั้น จึงจะสามารถกำหนดวัตถุประสงค์ในการทดสอบได้ โดย
ผู้ใช้งานสามารถกำหนดจำนวนข้อและเกณฑ์การผ่านรายวัตถุประสงค์ รวมถึงเกณฑ์การผ่านของ
คะแนนรวมทั้งฉบับ เมื่อผู้ใช้งานกรอกข้อมูลเรียบร้อยแล้ว สามารถกำหนดรายละเอียดของข้อสอบ
ได้ว่าแต่ละข้อต้องการวัดในวัตถุประสงค์ใด เมื่อกำหนดรายละเอียดครบถ้วนแล้วจึงดำเนินการ
บันทึกข้อมูลเพื่อเสร็จสิ้นการตั้งค่าและดำเนินการในขั้นถัดไป

กรณีที่ผู้ใช้งานเลือกวิเคราะห์ข้อสอบแบบอัตนัยจะต้องกรอกข้อมูล ได้แก่ ชื่อแฟ้มข้อมูล
ชื่อวิชา รหัสวิชา ปีการศึกษา จำนวนข้อ จำนวนวัตถุประสงค์ คะแนนเต็มแต่ละข้อ ซึ่งสามารถ
กรอกได้ตั้งแต่ 1 ถึง 99 คะแนน และดำเนินการตั้งค่าวัตถุประสงค์ต่อไป รายละเอียดดังภาพ 3

ภาพ 3

หน้าจอการกำหนดค่าเริ่มต้นในการวิเคราะห์ข้อมูล (แบบอัตนัย)

รายละเอียดสำหรับการวิเคราะห์

ข้อสอบเขียนตอบ (อัตนัย)

ชื่อแฟ้มข้อมูล :

ชื่อวิชา :

รหัสวิชา :

ปีการศึกษา :

จำนวนข้อ :

คะแนนเต็มแต่ละข้อ

จำนวนจุดประสงค์ :

ตั้งค่าจุดประสงค์

ยกเลิก

บันทึก

เมื่อเลือกประเภทข้อสอบและกำหนดข้อมูลพื้นฐานเรียบร้อยแล้ว ลำดับถัดไป ผู้ใช้งาน
สามารถเลือกรูปแบบการนำเข้าข้อมูลจากหน้าต่างหลัก ซึ่งซอฟต์แวร์รองรับการนำเข้าไฟล์
นามสกุล .csv (แบบ Comma delimited) โดยเรียงลำดับของผู้ตอบและคำตอบไว้ไฟล์ดังกล่าวให้
เรียบร้อยแล้ว นอกจากนี้ ผู้ใช้งานยังสามารถกรอกข้อมูลในซอฟต์แวร์ได้โดยตรงเพื่อทำการวิเคราะห์
หลังจากนั้นดำเนินการบันทึกข้อมูล เมื่อทำการบันทึกข้อมูลเรียบร้อยแล้วผลการบันทึกข้อมูลจะ
แสดงผลในหน้าจอหลัก หลังจากนั้นจึงจะสามารถเริ่มต้นการวิเคราะห์ข้อมูลได้

หากผู้ใช้งานต้องการแก้ไขข้อมูลพื้นฐานที่ใช้ในการวิเคราะห์ ผู้ใช้งานสามารถเลือกแก้ไขข้อมูลได้โดยเลือกแก้ไขรายละเอียดสำหรับการวิเคราะห์ หากผู้ใช้งานต้องการแก้ไขข้อมูลของผู้ตอบที่นำเข้าในซอฟต์แวร์ ผู้ใช้งานสามารถเลือกที่ชื่อหรือคำตอบของผู้ตอบ จากนั้น เลือกแก้ไขข้อมูลรายบุคคลเพื่อกลับไปแก้ไขข้อมูลหรือลบข้อมูลของผู้ตอบดังกล่าวออกจากชุดข้อมูลที่ใช้ในการวิเคราะห์

สำหรับการวิเคราะห์ข้อมูลของซอฟต์แวร์ ผู้ใช้สามารถดำเนินการวิเคราะห์ข้อมูลได้โดยเลือกที่วิเคราะห์ข้อสอบ ซึ่งซอฟต์แวร์จะแสดงตัวเลือกย่อยว่าผู้ใช้งานต้องการวิเคราะห์ข้อสอบแบบเลือกตอบ (ปรนัย) หรือแบบเขียนตอบ (อัตนัย) ทั้งนี้ ผู้ใช้งานจะต้องเลือกรูปแบบการวิเคราะห์ให้สอดคล้องกับลักษณะของข้อมูลที่นำเข้า

กรณีที่ผู้ใช้งานต้องการวิเคราะห์ข้อสอบแบบเลือกตอบ (ปรนัย) ซอฟต์แวร์จะทำการวิเคราะห์ข้อมูล และรายงานผลการวิเคราะห์ ดังภาพ 4 โดยมีรายละเอียดในการวิเคราะห์ประกอบด้วย ค่าสถิติพื้นฐานของชุดข้อมูล ได้แก่ จำนวนผู้สอบ จำนวนข้อสอบ ค่าเฉลี่ยคะแนนสอบ ส่วนเบี่ยงเบนมาตรฐาน ค่าความเที่ยงทั้งฉบับตามแนวคิดการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ และคุณภาพรายข้อของข้อสอบจำแนกตามวัตถุประสงค์ของการสอบ ทั้งผลการวิเคราะห์ตัวเลือกถูกและตัวเลือกหลง ได้แก่ ค่าความยาก และค่าอำนาจจำแนกด้วยดัชนี B (B-index) ประกอบการแปลผล นอกจากนี้ ผู้ใช้งานสามารถเลือกส่งออกข้อมูลเพื่อนำไปใช้งานในรูปแบบไฟล์มาตรฐาน PDF

ภาพ 4

ตัวอย่างการรายงานผลการวิเคราะห์ข้อสอบแบบเลือกตอบ (ปรนัย)

รายงานผลการวิเคราะห์ข้อสอบเลือกตอบ (ปรนัย) แบบอิงเกณฑ์				
จำนวนผู้สอบ (N) : 10 คน		ค่าเฉลี่ยคะแนนสอบ (M) : 2.90	ค่าความเที่ยงทั้งฉบับ (Reliability) : 0.74	
จำนวนข้อสอบ (I) : 10 ข้อ		ส่วนเบี่ยงเบนมาตรฐาน (SD) : 1.79		
จุดประสงค์/ข้อ	ตัวเลือก	ค่าความยาก (P)	ค่าอำนาจจำแนก (B-Index)	การแปลผลรายตัวเลือก
จุดประสงค์ 1				
ข้อ 1.	1*	0.30	0.33	ค่อนข้างยาก อำนาจจำแนกดีพอใช้
	2	0.60	-0.17	ความยากและอำนาจจำแนกใช้ได้
	3	0.10	-0.17	ความยากและอำนาจจำแนกใช้ได้
ข้อ 2.	1	0.30	-0.08	ความยากและอำนาจจำแนกใช้ได้
	2*	0.40	0.58	ความยากปานกลาง อำนาจจำแนกดีมาก
	3	0.30	-0.50	ความยากและอำนาจจำแนกใช้ได้
ข้อ 3.	1	0.50	-0.83	ความยากและอำนาจจำแนกใช้ได้
	2*	0.20	0.50	ค่อนข้างยาก อำนาจจำแนกดีมาก
	3	0.30	0.33	ความยากใช้ได้ อำนาจจำแนกใช้ไม่ได้
ข้อ 4.	1	0.40	0.17	ความยากใช้ได้ อำนาจจำแนกใช้ไม่ได้
	2	0.40	-0.67	ความยากและอำนาจจำแนกใช้ได้

หากผู้ใช้งานต้องการวิเคราะห์ข้อสอบแบบเขียนตอบ (อัตนัย) ซอฟต์แวร์จะแสดงผลการวิเคราะห์ ดังภาพ 5 โดยผลการวิเคราะห์ประกอบด้วย จำนวนผู้สอบ จำนวนข้อสอบ ค่าเฉลี่ยคะแนนสอบ ส่วนเบี่ยงเบนมาตรฐาน ค่าความเที่ยงทั้งฉบับตามแนวคิดการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ คุณภาพรายข้อของข้อสอบจำแนกตามวัตถุประสงค์ของการสอบ และการแปลผล ซึ่งผู้ใช้งานสามารถเลือกส่งออกข้อมูลเพื่อนำไปใช้งานในรูปแบบไฟล์มาตรฐาน PDF ได้เช่นเดียวกัน

ภาพ 5

ตัวอย่างการรายงานผลการวิเคราะห์ข้อสอบแบบเขียนตอบ (อัตนัย)

รายงานผลการวิเคราะห์ข้อสอบเขียนตอบ (อัตนัย) แบบอิงเกณฑ์				
จำนวนผู้สอบ (N) : 9 คน		ค่าเฉลี่ยคะแนนสอบ (M) : 24.44		ค่าความเที่ยงทั้งฉบับ (Reliability) : 0.99
จำนวนข้อสอบ (I) : 5 ข้อ		ส่วนเบี่ยงเบนมาตรฐาน (SD) : 6.98		
จุดประสงค์/ข้อ	คะแนนเต็ม	ค่าความยาก (P)	ค่าอำนาจจำแนก (R)	การแปลผลรายข้อ
จุดประสงค์ 1				
ข้อ 1.	5	0.61	0.58	ค่อนข้างง่าย อำนาจจำแนกดีมาก
ข้อ 2.	5	0.63	0.69	ค่อนข้างง่าย อำนาจจำแนกดีมาก
จุดประสงค์ 2				
ข้อ 3.	10	0.33	0.29	ค่อนข้างยาก อำนาจจำแนกพอใช้ได้
ข้อ 4.	10	0.39	0.34	ค่อนข้างยาก อำนาจจำแนกดีพอใช้
ข้อ 5.	10	0.23	0.60	ค่อนข้างยาก อำนาจจำแนกดีมาก

นอกจากผลการวิเคราะห์คุณภาพข้อสอบรายข้อและแบบสอบทั้งฉบับแล้ว ผู้ใช้งานยังสามารถเรียกดูรายงานคะแนนรายบุคคลและคะแนนของผู้สอบทั้งหมด โดยสามารถเลือกที่รายงานคะแนนสอบ ซอฟต์แวร์จะปรากฏตัวเลือกการรายงานผลการสอบรายบุคคลและคะแนนของผู้สอบทั้งหมดให้ผู้ใช้งานเลือก ส่วนรายงานผลคะแนนสอบรายบุคคลประกอบด้วย ข้อมูลผู้สอบ และข้อมูลการสอบ ได้แก่ ชื่อวิชา รหัสวิชา ปีการศึกษา คะแนนเต็ม จำนวนจุดประสงค์ เกณฑ์ผ่าน คะแนนที่ได้ในแต่ละข้อ คะแนนรวม ร้อยละของคะแนนที่ได้ และผลการประเมิน รายละเอียดตัวอย่างรายงานคะแนนสอบของผู้เรียนรายบุคคลแบบปรนัยและแบบอัตนัยแสดง ดังภาพ 6 และ 7 ตามลำดับ

หากผู้ใช้งานต้องการทราบรายงานผลในภาพรวมของผู้สอบทุกคน สามารถเลือกดูรายงานได้เช่นเดียวกัน โดยรายงานประกอบด้วยผลคะแนนที่ได้ ร้อยละ และผลการประเมินตามเกณฑ์ที่กำหนดของผู้เรียนแต่ละคน รายละเอียดตัวอย่างการรายงานผลของคะแนนในภาพรวมสำหรับข้อสอบแบบปรนัยแสดง ดังภาพ 8

ภาพ 6

ตัวอย่างการรายงานคะแนนสอบของผู้เรียนรายบุคคล (แบบปรนัย)

รายงานผลคะแนนสอบรายบุคคล

ชื่อ - นามสกุล : ง รหัสประจำตัว : 107

ชื่อวิชา : English รหัสวิชา : 2757301 ปีการศึกษา : 2569

คะแนนเต็ม : 10 จำนวนจุดประสงค์ : 2

จุดประสงค์/ข้อ	คะแนนที่ได้	คะแนนรวม	คะแนนรวม คิดเป็นร้อยละ	ผลประเมิน
จุดประสงค์ 1	คะแนนเต็ม = 5	เกณฑ์ผ่าน = 40		
ข้อ 1.	1	3	60	ผ่านจุดประสงค์
ข้อ 2.	1			
ข้อ 3.	0			
ข้อ 4.	0			
ข้อ 5.	1			
จุดประสงค์ 2	คะแนนเต็ม = 5	เกณฑ์ผ่าน = 60		
ข้อ 6.	0	2	40	ไม่ผ่านจุดประสงค์
ข้อ 7.	1			
ข้อ 8.	0			
ข้อ 9.	0			
ข้อ 10.	1			
ทั้งฉบับ		5	50	ผ่านรายวิชา

หมายเหตุ : เกณฑ์ผ่านร้อยละทั้งฉบับ = 50

ภาพ 7

ตัวอย่างการรายงานคะแนนสอบของผู้เรียนรายบุคคล (แบบอัตนัย)

รายงานผลคะแนนสอบรายบุคคล

ชื่อ - นามสกุล : f รหัสประจำตัว : 6

ชื่อวิชา : คณิตศาสตร์ รหัสวิชา : 21021 ปีการศึกษา : 2568

คะแนนเต็ม : 40 จำนวนจุดประสงค์ : 2

จุดประสงค์/ข้อ	คะแนนที่ได้	คะแนนรวม	คะแนนรวม คิดเป็นร้อยละ	ผลประเมิน
จุดประสงค์ 1	คะแนนเต็ม = 10	เกณฑ์ผ่าน = 50		
ข้อ 1.	1	3	30	ไม่ผ่านจุดประสงค์
ข้อ 2.	2			
จุดประสงค์ 2	คะแนนเต็ม = 30	เกณฑ์ผ่าน = 50		
ข้อ 3.	5	15	50	ผ่านจุดประสงค์
ข้อ 4.	5			
ข้อ 5.	5			
ทั้งฉบับ		18	45	ไม่ผ่านรายวิชา

หมายเหตุ : เกณฑ์ผ่านร้อยละทั้งฉบับ = 50

ภาพ 8

ตัวอย่างการรายงานผลคะแนนภาพรวม (แบบปรนัย)

รายงานผลคะแนนทั้งหมด

ชื่อวิชา : English รหัสวิชา : 2757301 ปีการศึกษา : 2569

จำนวนข้อสอบ (I) : 10 ข้อ จำนวนจุดประสงค์ : 2 คะแนนเต็ม : 10

จำนวนผู้สอบ (N) : 10 คน ค่าเฉลี่ยคะแนนสอบ (M) : 2.90 ส่วนเบี่ยงเบนมาตรฐาน (SD) : 1.79

ลำดับที่	รหัสประจำตัว	ชื่อ - นามสกุล	จุดประสงค์ 1		จุดประสงค์ 2		ทั้งฉบับ	
			คะแนนที่ได้ (ร้อยละ)	ผลประเมิน 40*	คะแนนที่ได้ (ร้อยละ)	ผลประเมิน 60*	คะแนนที่ได้ (ร้อยละ)	ผลประเมิน 50*
1	101	A	3 (60)	ผ่าน	0 (0)	ตก	3 (30)	ตก
2	102	b	1 (20)	ตก	1 (20)	ตก	2 (20)	ตก
3	103	c	3 (60)	ผ่าน	2 (40)	ตก	5 (50)	ผ่าน
4	104	d	1 (20)	ตก	1 (20)	ตก	2 (20)	ตก
5	105	e	0 (0)	ตก	1 (20)	ตก	1 (10)	ตก
6	106	f	0 (0)	ตก	2 (40)	ตก	2 (20)	ตก

ตอนที่ 2 การประเมินคุณภาพซอฟต์แวร์

สำหรับการประเมินคุณภาพซอฟต์แวร์คอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ที่สอดคล้องกับหลักสูตรอิงมาตรฐานของการศึกษาขั้นพื้นฐานในประเทศไทย หรือซอฟต์แวร์ i-ACT ผู้วิจัยแบ่งการประเมินคุณภาพออกเป็น 2 ส่วน ได้แก่ การประเมินคุณภาพซอฟต์แวร์ก่อนนำไปทดลองใช้งานด้วยการประเมินอินเตอร์เฟซของซอฟต์แวร์ (user interface evaluation: UI) จากแบบประเมินฮิวริสติก และการประเมินคุณภาพซอฟต์แวร์หลังการนำไปใช้งานด้วยการประเมินประสบการณ์ผู้ใช้ (user experience evaluation: UX) จากแบบประเมินประสบการณ์ผู้ใช้ รายละเอียดมีดังนี้

2.1 ผลการประเมินคุณภาพซอฟต์แวร์ก่อนนำไปทดลองใช้งานด้วยการประเมินอินเตอร์เฟซของซอฟต์แวร์

ผลการประเมินคุณภาพซอฟต์แวร์ก่อนนำไปทดลองใช้งานด้วยการประเมินอินเตอร์เฟซของซอฟต์แวร์ จากแบบประเมินฮิวริสติก เป็นการประเมินความจำเป็นและข้อสังเกตของอินเตอร์เฟซซอฟต์แวร์ โดยใช้ข้อความสำหรับการประเมินในรูปแบบมาตรฐานค่า 5 ระดับตามแนวคิดของลิเคิร์ต จำนวน 10 ข้อ เพื่อประเมินว่า อินเตอร์เฟซมีความจำเป็นต่อการปรับปรุงก่อนนำไปทดลองใช้งาน การประเมินดำเนินการโดยผู้เชี่ยวชาญ จำนวน 5 ท่าน จำแนกเป็นผู้เชี่ยวชาญด้านเทคโนโลยี 2 ท่าน และผู้เชี่ยวชาญด้านการวัดและประเมินผล 3 ท่าน โดยผู้วิจัยได้นำซอฟต์แวร์ที่พัฒนาขึ้นไปให้ผู้เชี่ยวชาญพิจารณาและประเมินคุณภาพก่อนนำซอฟต์แวร์ไปให้ตัวอย่างวิจัยทดลองใช้ในการวิเคราะห์ข้อสอบ ซึ่งเกณฑ์สำหรับแปลผลแบบประเมินคุณภาพมีดังนี้ (Galavi et al., 2024; Nielsen, 1994) ได้แก่ 1.00 – 1.49: ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง 1.50 – 2.49: มีความจำเป็นน้อย / ปรับเปลี่ยนหรือไม่ก็ได้ 2.50 – 3.49: มีความจำเป็นระดับปานกลาง / ควรปรับปรุงแก้ไข 3.50 – 4.49: มีความจำเป็นสูง / ควรปรับปรุงแก้ไขโดยเร็ว และ 4.50 – 5.00: มีความจำเป็นสูงมาก / ควรปรับปรุงแก้ไขทันที

ตาราง 1

ผลการประเมินแบบอิวิริสติก

ข้อ	รายการประเมิน	M	SD	แปลผล
1	การแสดงสถานะ	1.60	0.55	มีความจำเป็นน้อย / ปรับเปลี่ยนหรือไม่ก็ได้
2	ความสอดคล้องระหว่างซอฟต์แวร์คอมพิวเตอร์กับการนำไปปฏิบัติจริง	1.60	0.89	มีความจำเป็นน้อย / ปรับเปลี่ยนหรือไม่ก็ได้
3	การควบคุมการใช้งานเพื่อป้องกันความผิดพลาด	1.60	0.55	มีความจำเป็นน้อย / ปรับเปลี่ยนหรือไม่ก็ได้
4	ความเป็นมาตรฐานและความคงเสถียร	1.00	0.00	ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง
5	การป้องกันความผิดพลาด	1.00	0.00	ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง
6	การจัดจำข้อมูลการใช้งาน	1.40	0.55	ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง
7	ความยืดหยุ่นและประสิทธิภาพในการทำงาน	1.40	0.55	ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง
8	การออกแบบที่เรียบง่ายสวยงาม	1.40	0.55	ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง
9	การแจ้งเตือนและให้ความช่วยเหลือผู้ใช้งานจากความผิดพลาด	1.40	0.55	ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง
10	ข้อมูลให้ความช่วยเหลือและเอกสารที่เกี่ยวข้อง	1.20	0.45	ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง

จากตาราง 1 พบว่า ผลการประเมินอินเตอร์เฟซของซอฟต์แวร์ i-ACT จากแบบประเมินอิวิริสติกด้วยผู้เชี่ยวชาญ 5 ท่าน ภาพรวมอยู่ในระดับดีมาก โดยผลการประเมินแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มที่ไม่พบความจำเป็นหรือปัญหาที่ต้องปรับปรุง ได้แก่ รายการประเมินที่ 4 ถึง 10 ซึ่งมีค่าเฉลี่ยอยู่ระหว่าง 1.00 ถึง 1.40 ประกอบไปด้วย ความเป็นมาตรฐานและความคงเสถียร ($M = 1.00$, $SD = 0.00$) การป้องกันความผิดพลาด ($M = 1.00$, $SD = 0.00$) ข้อมูลให้ความช่วยเหลือและเอกสารที่เกี่ยวข้อง ($M = 1.20$, $SD = 0.45$) การจัดจำข้อมูลการใช้งาน ($M = 1.40$, $SD = 0.55$) ความยืดหยุ่นและประสิทธิภาพในการทำงาน ($M = 1.40$, $SD = 0.55$) และการแจ้งเตือนและให้ความช่วยเหลือผู้ใช้งานจากความผิดพลาด ($M = 1.40$, $SD = 0.55$) และกลุ่มที่มีความจำเป็นเล็กน้อย สามารถปรับเปลี่ยนหรือไม่ปรับเปลี่ยนก็ได้ ได้แก่ รายการประเมินที่ 1 ถึง 3 โดยมีค่าเฉลี่ยเท่ากัน ได้แก่ การแสดงสถานะ ($M = 1.60$, $SD = 0.55$) ความสอดคล้องระหว่างซอฟต์แวร์คอมพิวเตอร์กับการนำไปปฏิบัติจริง ($M = 1.60$, $SD = 0.89$) และการควบคุมการใช้งานเพื่อป้องกันความผิดพลาด ($M = 1.60$, $SD = 0.55$)

2.2 ผลการประเมินคุณภาพซอฟต์แวร์หลังการนำไปใช้งานด้วยการประเมินประสบการณ์ผู้ใช้

การประเมินคุณภาพซอฟต์แวร์หลังการนำไปใช้งานด้วยการประเมินประสบการณ์ผู้ใช้ (user experience evaluation: UX) จากแบบประเมินประสบการณ์ผู้ใช้ซอฟต์แวร์ i-ACT โดยผู้ใช้

มาตรฐานความหมายตามแนวคิดของออสกู๊ด 7 ระดับ จำนวน 19 ข้อ ซึ่งมีคำคุณศัพท์ในการประเมินทั้งทางบวกและทางลบ โดยตัวอย่างวิจัย คือ บุคลากรทางการศึกษาสังกัดคณะกรรมการการศึกษาขั้นพื้นฐาน จำนวน 50 คน แบ่งเป็นครูผู้สอนระดับประถมศึกษา 25 คน และระดับมัธยมศึกษา 25 คน เกณฑ์การพิจารณาไขาคากลางของมาตรวัด ซึ่งเท่ากับ 3.50 หากผลการประเมินมีค่าตั้งแต่ 3.50 ขึ้นไปหมายความว่า ตัวอย่างวิจัยมีประสพการณ์ที่ดีในทางบวกต่อการใช้งานซอฟต์แวร์ โดยมีผลการประเมินดังตาราง 2

ตาราง 2

ผลการประเมินประสพการณ์ผู้ใช้

ข้อ	มาตรวัด 7 ระดับ (1-7)		ระดับ ประถมศึกษา		ระดับ มัธยมศึกษา		M	SD
			(n=25)		(n=25)			
			M	SD	M	SD		
1	เข้าใจยาก	- เข้าใจง่าย	6.04	0.84	6.44	0.58	6.24	0.71
2*	สร้างสรรค์	- ลอกเลียนแบบ	5.08	1.61	5.48	1.45	5.28	1.53
3*	เรียนรู้ง่าย	- เรียนรู้ยาก	5.04	1.57	6.04	1.02	5.54	1.30
4*	มีคุณค่า	- ไร้คุณค่า	5.96	1.40	6.40	0.58	6.18	0.99
5	น่าเบื่อ	- น่าสนใจ	6.04	0.98	6.36	0.64	6.20	0.81
6	อึดอัด	- สบายใจ	5.88	1.01	6.08	1.22	5.98	1.12
7*	เร็ว	- ช้า	5.12	1.79	6.20	1.08	5.66	1.44
8*	แนวใหม่	- แบบดั้งเดิม	5.12	1.76	5.80	0.91	5.46	1.34
9	ติดขัด	- ราบรื่น	5.80	0.87	6.24	0.83	6.02	0.85
10*	ดี	- แย่	5.76	1.51	6.40	0.58	6.08	1.05
11	ซับซ้อน	- ไม่ซับซ้อน	5.52	1.12	5.84	1.31	5.68	1.22
12	ไม่น่าใช้งาน	- น่าใช้งาน	5.96	1.02	6.12	0.83	6.04	0.93
13	ล้าสมัย	- ล้ำสมัย	5.80	0.96	6.08	0.81	5.94	0.89
14	น่าตื่นตา	- น่าชื่นชม	6.12	0.88	6.16	0.69	6.14	0.79
15*	ปลอดภัย	- ไม่ปลอดภัย	5.60	1.55	6.16	1.03	5.88	1.29
16*	จูงใจ	- ไม่จูงใจ	5.64	1.22	6.04	0.61	5.84	0.92
17*	ตรงตาม ความคาดหวัง	- ไม่ตรงตาม ความคาดหวัง	5.68	1.31	5.96	0.79	5.82	1.05
18	ขาด ประสิทธิภาพ	- มี ประสิทธิภาพ	6.08	1.12	6.44	0.65	6.26	0.89
19*	ชัดเจน	- คลุมเครือ	5.40	1.71	6.28	0.74	5.84	1.23
ภาพรวม			5.67	0.97	6.13	0.54	5.90	0.76

หมายเหตุ * หมายถึง มาตรที่มีการให้คะแนนแบบกลับ

ตาราง 2 ผลการประเมินประสิทธิภาพผู้ใช้งานซอฟต์แวร์ i-ACT โดยใช้มาตรฐานความหมายตามแนวคิดของออสกูด 7 ระดับ พบว่า ตัวอย่างวิจัยมีประสิทธิภาพที่ดีในทางบวกต่อการใช้งานซอฟต์แวร์อย่างชัดเจน โดยมีคะแนนเฉลี่ยรวมของตัวอย่างทั้ง 2 กลุ่ม เท่ากับ 5.90 ($M = 5.90$, $SD = 0.76$) ซึ่งสูงกว่าเกณฑ์ที่กำหนด 3.50 และเมื่อพิจารณารายการประเมินแต่ละข้อ พบว่า ประสิทธิภาพการใช้งานซอฟต์แวร์ที่มีผลการประเมินสูงสุด 3 อันดับแรก คือ ข้อที่ 18 (ขาดประสิทธิภาพ - มีประสิทธิภาพ) ซึ่งมีค่าเฉลี่ยรวมอยู่ที่ 6.26 ($M = 6.26$, $SD = 0.89$) ข้อที่ 1 (เข้าใจยาก - เข้าใจง่าย) มีค่าเฉลี่ยรวมอยู่ที่ 6.24 ($M = 6.24$, $SD = 0.71$) และข้อที่ 5 (น่าเบื่อ - น่าสนใจ) มีค่าเฉลี่ยอยู่ที่ 6.20 ($M = 6.20$, $SD = 0.81$) จากผลการประเมินประสิทธิภาพผู้ใช้งานซอฟต์แวร์ทุกรายการซึ่งสูงกว่าเกณฑ์ 3.50 สามารถเชื่อมั่นได้ว่า ซอฟต์แวร์ i-ACT เป็นซอฟต์แวร์ที่ผู้ใช้งานสามารถนำไปใช้วิเคราะห์ข้อสอบแบบอิงเกณฑ์ได้จริงและผู้ใช้งานมีประสิทธิภาพที่ดีต่อการใช้งานซอฟต์แวร์ดังกล่าว

อภิปรายผลการวิจัย

จากผลการวิจัยข้างต้นสามารถนำมาอภิปรายได้ 2 ประเด็น ดังนี้

1. ปัจจุบันหลักสูตรการศึกษาขั้นพื้นฐานที่ใช้ในประเทศไทยเป็นหลักสูตรอิงมาตรฐาน (standard-based curriculum) ซึ่งใช้แนวทางการตัดสินผลการเรียนรู้แบบอิงเกณฑ์ แต่บุคลากรทางการศึกษาในประเทศไทยส่วนใหญ่ยังมีข้อจำกัดในการเข้าถึงการวิเคราะห์คุณภาพข้อสอบแบบอิงเกณฑ์ เนื่องจากซอฟต์แวร์โดยทั่วไปที่มีการเผยแพร่ให้ใช้งานเป็นการวิเคราะห์คุณภาพข้อสอบแบบอิงกลุ่ม ซึ่งอาจนำไปสู่ความคลาดเคลื่อนในการแปลผลการวิเคราะห์คุณภาพข้อสอบที่ไม่สอดคล้องกับหลักสูตร ด้วยเหตุนี้ การพัฒนาซอฟต์แวร์ i-ACT ซึ่งเป็นซอฟต์แวร์คอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ที่สอดคล้องกับหลักสูตรอิงมาตรฐานของการศึกษาขั้นพื้นฐานในประเทศไทย จะช่วยให้บุคลากรทางการศึกษาสามารถวิเคราะห์ข้อสอบแบบอิงเกณฑ์ได้ถูกต้องและสอดคล้องกับหลักสูตรในปัจจุบัน รวมถึงซอฟต์แวร์ที่พัฒนาขึ้นสามารถรองรับระบบปฏิบัติการคอมพิวเตอร์ในปัจจุบัน ซึ่งช่วยแก้ปัญหาการใช้งานซอฟต์แวร์การวิเคราะห์ข้อสอบแบบอิงเกณฑ์ดั้งเดิมที่ไม่รองรับระบบปฏิบัติการรุ่นใหม่ นอกจากนี้ ยังมีการเพิ่มเติมฟังก์ชันการนำเข้าข้อมูล การแปลผลการวิเคราะห์ และการส่งออกรายงานผลการวิเคราะห์ (export) เพื่อนำไปใช้งานต่อในส่วนที่เกี่ยวข้อง จึงเป็นการช่วยลดภาระของบุคลากรทางการศึกษาในการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ด้วยตนเอง สอดคล้องกับข้อค้นพบของ Lee et al. (2024) ที่กล่าวไว้ว่า ข้อจำกัดและความกังวลสำคัญของครูผู้สอนมักอยู่ที่การใช้ข้อมูลเพื่อการประเมินผู้เรียน เนื่องจากกระบวนการที่มีความซับซ้อนทางสถิติ ดังนั้น การเปลี่ยนสูตรการคำนวณที่ซับซ้อนให้อยู่ในรูปแบบสารสนเทศที่เข้าใจ เช่น ตารางและคำอธิบาย จึงเป็นเครื่องมือสำคัญในการขับเคลื่อนแนวคิดการประเมินเพื่อพัฒนาการเรียนรู้ (assessment for learning) ในระดับชั้นเรียน เพื่อให้ครูผู้สอนสามารถนำผลการวิเคราะห์ข้อสอบที่ถูกต้องไปใช้ในการปรับปรุงและพัฒนาการจัดการเรียนการสอนที่ช่วยส่งเสริมและพัฒนาผู้เรียนได้ตรงตามวัตถุประสงค์ในการเรียนรู้อย่างยิ่งขึ้น ขณะเดียวกัน การพัฒนาซอฟต์แวร์สำหรับการวิเคราะห์คุณภาพข้อสอบแบบอิงเกณฑ์ทำให้ผลการวิเคราะห์ข้อสอบมีความสอดคล้องกับโครงสร้างหลักสูตรแกนกลางของประเทศไทยในปัจจุบัน อีกทั้งการใช้งานซอฟต์แวร์ที่ไม่ได้รับการอัปเดตให้ทันสมัยอาจลดทอน

ประสิทธิภาพการทำงานของซอฟต์แวร์ และสร้างช่องโหว่ด้านความปลอดภัยที่อาจส่งผลต่อการบริหารจัดการและวิเคราะห์ข้อมูลที่มีความสำคัญ (Ampatzoglou et al., 2015) การพัฒนาซอฟต์แวร์ i-ACT จึงไม่เพียงช่วยเติมเต็มองค์ความรู้ด้านการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ให้กับบุคลากรทางการศึกษาและผู้ใช้งานทั่วไป แต่ยังเป็นการพัฒนาซอฟต์แวร์ที่ช่วยส่งเสริมการวัดและการประเมินการศึกษาที่มีความถูกต้องและเหมาะสมกับบริบทการศึกษาในไทย

2. ผลการประเมินซอฟต์แวร์ i-ACT จากประสบการณ์ของผู้ใช้งาน พบว่า ผู้ใช้งานมีประสบการณ์ที่ดีในการใช้งานซอฟต์แวร์ในเชิงบวก ซึ่งอาจสะท้อนถึงความพึงพอใจของผู้ใช้งานต่อซอฟต์แวร์ ประเด็นหลักที่ทำให้ผู้ใช้งานเกิดประสบการณ์ที่ดีในการใช้งานซอฟต์แวร์อาจมาจากการออกแบบอินเตอร์เฟซที่เรียบง่ายและให้ความสำคัญกับกระบวนการทำงานของซอฟต์แวร์ในทุกส่วน นอกจากนี้ ซอฟต์แวร์ยังสามารถแปลความหมายของคุณภาพข้อสอบรายข้อได้ครบถ้วน ได้แก่ ค่าความยากและค่าอำนาจจำแนก ซึ่งผลการวิเคราะห์ที่ได้จากซอฟต์แวร์เป็นข้อความในลักษณะเชิงบรรยาย จึงช่วยลดโอกาสการเกิดข้อผิดพลาดในการแปลผลและการนำผลการวิเคราะห์ไปใช้งาน สอดคล้องกับแนวทางการออกแบบซอฟต์แวร์ที่สามารถลดข้อผิดพลาดจากการแปลผลการวิเคราะห์ นอกจากนี้ ยังสามารถแก้ไขข้อจำกัดของซอฟต์แวร์วิเคราะห์คุณภาพข้อสอบแบบดั้งเดิมซึ่งต้องอาศัยความรู้พื้นฐานด้านการวัดและประเมินผลการศึกษาในการแปลผลการวิเคราะห์ การออกแบบดังกล่าวสอดคล้องกับกฎ 8 ข้อในการออกแบบอินเตอร์เฟซที่ Shneiderman (2010) ได้เสนอไว้ว่า การออกแบบที่เรียบง่าย และมีวิธีการใช้งานซึ่งสามารถอำนวยความสะดวกให้กับผู้ใช้งานจะช่วยลดภาระทางความจำระยะสั้น (short-term memory load) และช่วยส่งเสริมให้ผู้ใช้งานสามารถควบคุมการทำงานได้ด้วยตนเอง (internal locus of control) ผ่านระบบการจัดการข้อมูลที่มีแบบแผน และเป็นขั้นเป็นตอน นอกจากนี้ การออกแบบที่อำนวยความสะดวกให้กับผู้ใช้งานยังสามารถช่วยลดช่องว่างด้านการวัดและการประเมินผลให้กับบุคลากรทางการศึกษาและผู้ใช้งานทั่วไปที่ไม่มีพื้นฐานด้านการวัดและประเมินผลสามารถแปลความผลการวิเคราะห์ในประเด็นเฉพาะทางได้ ส่งผลให้ผู้ใช้งานมีแนวโน้มจะนำซอฟต์แวร์ไปใช้งานจริง (Norman, 2013)

สรุปผลการวิจัย

การพัฒนาซอฟต์แวร์วิเคราะห์คุณภาพข้อสอบแบบอิงเกณฑ์ที่มีความเหมาะสมและสอดคล้องกับบริบทการใช้งานของครูผู้สอนในประเทศไทย พบว่า ซอฟต์แวร์ที่พัฒนาขึ้นสามารถทดแทนซอฟต์แวร์ดั้งเดิมซึ่งการแปลผลการวิเคราะห์อาจเข้าใจได้ยากสำหรับผู้ที่ไม่มีความรู้พื้นฐานด้านการวัดและการประเมินผลการศึกษา รวมถึงซอฟต์แวร์ดั้งเดิมไม่รองรับระบบปฏิบัติการคอมพิวเตอร์ใหม่ที่มีการพัฒนาขึ้นอย่างรวดเร็ว ซอฟต์แวร์ที่พัฒนาขึ้นในการวิจัยนี้จึงสามารถช่วยลดข้อจำกัดดังกล่าว ซอฟต์แวร์ที่พัฒนาในการวิจัยครั้งนี้ชื่อว่า i-ACT (Item Analysis for Criterion-referenced Test) เป็นซอฟต์แวร์ที่พัฒนาขึ้นด้วยภาษา Visual Basic สามารถใช้งานได้โดยไม่ต้องติดตั้งลงบนเครื่องคอมพิวเตอร์ (portable software) ผลการวิเคราะห์ข้อสอบที่สำคัญจากซอฟต์แวร์ประกอบด้วย ค่าความยาก ค่าอำนาจจำแนก และค่าความเที่ยงตามแนวคิดการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ ผู้วิจัยใช้กรอบแนวคิดการพัฒนาซอฟต์แวร์แบบวงจรการพัฒนาระบบ (system development life cycle: SDLC) ซึ่งช่วยให้เกิดการพัฒนาระบบ ตรวจสอบ

และปรับปรุงการใช้งานซอฟต์แวร์อย่างเป็นระบบ ก่อนนำซอฟต์แวร์ไปทดลองใช้ได้ดำเนินการประเมินคุณภาพก่อนนำไปใช้จริงด้วยการประเมินแบบฮิวริสติกจากผู้เชี่ยวชาญด้านการวัดและประเมินผล และผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศ เพื่อประเมินความถูกต้องเหมาะสมของระบบ ผลการประเมินพบว่า ความจำเป็นในการปรับปรุงซอฟต์แวร์อยู่ในระดับน้อยถึงไม่มีความจำเป็นต้องปรับปรุง หลังจากนั้นจึงได้ดำเนินการจัดทำคู่มือการใช้งานและส่งซอฟต์แวร์ให้ตัวอย่างวิจัยได้ทดลองใช้งานในการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ ตัวอย่างวิจัยที่ทดลองใช้งานซอฟต์แวร์ประกอบด้วยครูผู้สอนระดับประถมศึกษาและมัธยมศึกษา รวมทั้งสิ้น 50 คน หลังจากทดลองใช้งานซอฟต์แวร์เรียบร้อยแล้วตัวอย่างวิจัยดำเนินการประเมินประสบการณ์ผู้ใช้ ซึ่งในภาพรวมผู้ใช้งานมีประสบการณ์เชิงบวกต่อการใช้งานซอฟต์แวร์ สะท้อนให้เห็นว่าซอฟต์แวร์ที่พัฒนาขึ้นสามารถตอบสนองความต้องการการใช้งานได้อย่างเป็นรูปธรรม

ข้อเสนอแนะของการวิจัย

ข้อเสนอแนะในการนำผลวิจัยไปใช้

1. ควรมีการสนับสนุนและผลักดันในเชิงนโยบายเพื่อนำซอฟต์แวร์ไปใช้ในการพัฒนาชุดความรู้ด้านการประเมิน (Assessment literacy) ของบุคลากรทางการศึกษาและเชื่อมโยงไปยังการประกันคุณภาพการศึกษา (Quality Assurance) ที่สอดคล้องกับการประเมินคุณภาพผู้เรียนในประเทศไทย

2. ควรผลักดันการนำซอฟต์แวร์ไปใช้ร่วมกับระบบนิเวศเทคโนโลยีการศึกษาแบบเปิด (Open-access Educational Technology Ecosystem) เพื่อสร้างชุมชนแห่งการเรียนรู้ในด้านการวิเคราะห์ข้อสอบแบบอิงเกณฑ์ที่สอดคล้องกับหลักสูตรอิงมาตรฐานของประเทศไทย

ข้อเสนอแนะในการวิจัยครั้งถัดไป

1. ซอฟต์แวร์ที่พัฒนาขึ้นปัจจุบันต้องใช้การนำเข้าผลการตอบข้อสอบจากซอฟต์แวร์อื่นด้วยเหตุนี้ งานวิจัยในอนาคตควรมีการเชื่อมต่อกับเครื่องตรวจกระดาษคำตอบอัตโนมัติที่ใช้เทคโนโลยี Optical Mark Recognition (OMR) ซึ่งจะทำให้สามารถนำเข้าผลการตอบข้อสอบและวิเคราะห์ข้อสอบแบบอิงเกณฑ์ได้ทันที เป็นการลดภาระการจัดกระทำข้อมูลและความคลาดเคลื่อนที่อาจเกิดจากการนำเข้าข้อมูล เนื่องจากปัจจุบันเครื่องตรวจกระดาษคำตอบอัตโนมัติส่วนใหญ่ยังคงใช้การวิเคราะห์ข้อสอบแบบอิงกลุ่มซึ่งอาจยังไม่สอดคล้องกับหลักสูตรอิงมาตรฐานของประเทศไทย

2. ด้วยเทคโนโลยีปัญญาประดิษฐ์เข้ามามีบทบาทเป็นอย่างมากในการวิเคราะห์ข้อมูล การพัฒนาซอฟต์แวร์และการวิเคราะห์ข้อมูลแบบดั้งเดิมอาจต้องใช้กระบวนการและเวลาที่ค่อนข้างนาน การนำเทคโนโลยีปัญญาประดิษฐ์มาช่วยในการวิเคราะห์ข้อมูลจะทำให้ผลการวิเคราะห์มีความแม่นยำและรวดเร็วมากยิ่งขึ้น รวมถึงอาจสถานการณ์ให้ข้อมูลป้อนกลับเชิงวินิจฉัย (diagnostic feedback) จากผลการวิเคราะห์ไปยังผู้เรียนเพื่อการพัฒนาศักยภาพผู้เรียนรายบุคคล

บทบาทของผู้เขียน (author contributions) ณภัทร ชัยมงคลเป็นผู้ประพันธ์อันดับแรก (first author) และผู้รับผิดชอบบทความหลัก (corresponding author)

ผลประโยชน์ทับซ้อน (competing interests) ผู้เขียนขอรับรองว่าไม่มีผลประโยชน์ทับซ้อนใด ๆ ไม่ว่าจะเป็นทางการเงินหรือประเด็นอื่นใดที่เกี่ยวข้องกับงานวิจัยนี้

แหล่งทุนสนับสนุน (funding) การวิจัยครั้งนี้ได้รับทุนสนับสนุนบางส่วนจากทุนพัฒนาอาจารย์ใหม่/นักวิจัยใหม่ กองทุนรัชดาภิเษกสมโภช จุฬาลงกรณ์มหาวิทยาลัย (Grant for Development of New Faculty Staff, Ratchadaphiseksomphot Endowment Fund, Chulalongkorn University) เลขที่สัญญา DNS 63_040_27_003_1

กิตติกรรมประกาศ (acknowledgement) ผู้วิจัยขอขอบพระคุณ ศาสตราจารย์กิตติคุณ ดร.ศิริชัย กาญจนวาสี ที่ให้คำแนะนำในการทำวิจัยครั้งนี้ให้สำเร็จลุล่วงไปได้ด้วยดี

การเปิดเผยข้อมูล (data availability) หากผู้อ่านต้องการเข้าถึงชุดข้อมูลผ่านการประมวลผล (processed dataset) หรือเอกสารประกอบเพิ่มเติม สามารถติดต่อผู้รับผิดชอบบทความหลักได้ตามอีเมลที่ระบุ

จริยธรรมในการวิจัย (ethical approval) งานวิจัยนี้ได้รับใบรับรองโครงการวิจัยที่ 110/64 จากคณะกรรมการพิจารณาจริยธรรมการวิจัยในคน กลุ่มสหสถาบัน ชุดที่ 2 สังคมศาสตร์ มนุษยศาสตร์ และศิลปกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การใช้ปัญญาประดิษฐ์ (AI) ในการวิจัย ผู้วิจัยใช้ปัญญาประดิษฐ์เพื่อแปลบทคัดย่อจากภาษาไทยเป็นภาษาอังกฤษ และตรวจสอบความถูกต้องของไวยากรณ์ภาษาอังกฤษเท่านั้น

รายการอ้างอิง

สาคร แสงผึ่ง. (2547). *การวิเคราะห์ข้อสอบเลือกตอบ*.

<http://pn.nfe.go.th/pranee/UserFiles/Pdf/manual%20B-Index.pdf>

สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ.

(2557). *แนวปฏิบัติการวัดและประเมินผลการเรียนรู้ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน*

พุทธศักราช 2551 (พิมพ์ครั้งที่ 4). โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด.

ศิริชัย กาญจนวาสี. (2556). *ทฤษฎีการทดสอบแบบดั้งเดิม (พิมพ์ครั้งที่ 7). สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.*

Ampatzoglou, A., Ampatzoglou, A., Chatzigeorgiou, A., & Avgeriou, P. (2015). The financial aspect of managing technical debt: A systematic literature review. *Information and Software Technology*, 64, 52–73. <https://doi.org/10.1016/j.infsof.2015.04.001>

Conger, S. A. (2011). Software development life cycles and methodologies: Fixing the old and adopting the new. *International Journal of Information Technologies and Systems Approach*, 4(1), 1–22. <https://doi.org/10.4018/jitsa.2011010101>

Galavi, Z., Norouzi, S., & Khajouei, R. (2024). Heuristics used for evaluating the usability of mobile health applications: A systematic literature review. *Digital Health*, 10, Article 20552076241253539. <https://doi.org/10.1177/20552076241253539>

Latif, A. B., Yusof, I. J., Amin, N. F. M., Libunao, W. H., & Yusri, S. S. (2016). Multiple-choice items analysis using classical test theory and Rasch measurement model. *Man in India*, 96(1–2), 173–181.

Lee, J., Alonzo, D., Beswick, K., Abril, J. M. V., Chew, A. W., & Oo, C. Z. (2024). Dimensions of teachers' data literacy: A systematic review of literature from 1990 to 2021. *Educational Assessment, Evaluation and Accountability*, 36, 145–200. <https://doi.org/10.1007/s11092-024-09435-8>

- Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 152–158). Association for Computing Machinery. <https://doi.org/10.1145/191666.191729>
- Norman, D. A. (2013). *The design of everyday things* (Revised and expanded ed.). Basic Books.
- Rao, D. B. J., Kalpana, P., Kumar, G. S. N., & Atcha, N. M. (2025). A comparative analysis of software development models: Waterfall, Agile and DevOps. In V. S. Reddy, V. K. Prasad, J. Wang, & N. M. Rao Dasari (Eds.), *Intelligent systems and sustainable computing* (Vol. 417, pp. 589–599). Springer. https://doi.org/10.1007/978-981-97-8355-7_51
- Shneiderman, B. (2010). *Designing the user interface: Strategies for effective human-computer interaction* (5th ed.). Pearson Education India.

Translated Thai References

- Bureau of Academic Affairs and Educational Standards, Office of the Basic Education Commission, Ministry of Education. (2014). *Guidelines for learning measurement and evaluation according to the Basic Education Core Curriculum B.E. 2551* (4th ed.). The Agricultural Co-operative Federation of Thailand Printing House. [in Thai]
- Kanjanawasee, S. (2013). *Classical test theory* (7th ed.). Chulalongkorn University Press. [in Thai]
- Saengphueng, S. (2004). *Analysis of multiple-choice tests*. <http://lpn.nfe.go.th/pranee/UserFiles/Pdf/manual%20B-Index.pdf> [in Thai]