

การพัฒนาแบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือกสำหรับประเมินความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21 สำหรับนักศึกษาปริญญาตรี

The Development of a Multidimensional Forced-Choice Situational Judgment Test for Assessing Undergraduate Students' Expectations Toward Blended Learning in the 21st Century

นุรซีตา เพอแสะละ^{1*} และ นัฏฐภรณ์ หลาวทอง²

Nurseeta Phoosalae^{1*} and Nuttaporn Lawthong²

(Received: March 20, 2025; Revised: May 6, 2025; Accepted: May 7, 2025)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมายเพื่อ (1) พัฒนาแบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือกสำหรับประเมินความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21 สำหรับนักศึกษาปริญญาตรี และ (2) ตรวจสอบคุณภาพของแบบวัดด้านความสมดุลของผลการตอบ ด้านความตรงด้วยการหาวิเคราะห์ความตรงตามเนื้อหา ความตรงตามเกณฑ์สัมพัทธ์ และความตรงตามทฤษฎีโดยใช้วิธีการวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี และด้านความเที่ยงด้วยวิธีวัดความสอดคล้องภายในและวัดความคงที่โดยใช้วิธีการทดสอบซ้ำ การพัฒนาแบบวัดในครั้งนี้มีการตรวจสอบคุณภาพจากประชากร 4 กลุ่ม โดยแบ่งเป็นกลุ่มเป้าหมายที่มีการคัดเลือกแบบเฉพาะเจาะจง 3 กลุ่ม ได้แก่ ผู้เชี่ยวชาญด้านการวัดและประเมินผล อาจารย์ผู้สอนในระดับอุดมศึกษา และนักศึกษาในระดับปริญญาตรี กลุ่มตัวอย่างที่ได้จากการสุ่มตัวอย่าง 1 กลุ่ม คือ นักศึกษาระดับปริญญาตรี จำนวน 120 คน โดยมีการสุ่มตัวอย่างแบบชั้นภูมิซึ่งมีคณะเป็นชั้นภูมิ ผลการวิจัยพบว่า แบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือกที่มีการใช้ภาพประกอบ จำนวน 25 ข้อ แบ่งเป็น 5 สถานการณ์และวัดความคาดหวังต่อการเรียนรู้แบบผสมผสาน 5 ด้าน ผลการตรวจสอบความสมดุลของผลการตอบไม่พบสัญญาณเบื้องต้นของการมีความลำเอียงของแบบวัด พบความสัมพันธ์ของคะแนนที่ได้จากแบบวัดที่พัฒนาขึ้นกับแบบวัดอื่นจึงเป็นส่วนหนึ่งของหลักฐานที่ยืนยันถึงคุณภาพด้านความตรงตามเกณฑ์สัมพัทธ์ ผลการตรวจสอบความตรงตามทฤษฎีโดยใช้การวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี พบว่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากการวัด

¹ อาจารย์ สาขาการสอนอิสลามศึกษา คณะวิทยาการอิสลาม มหาวิทยาลัยสงขลานครินทร์

² รองศาสตราจารย์ ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹ Lecturer, Teaching Islamic Studies Program, Faculty of Islamic Sciences, Prince of Songkhla University

² Associate Professor, Department of Educational Research and Psychology, Faculty of Education, Chulalongkorn University

งานวิจัยนี้ได้รับทุนสนับสนุนจาก สำนักงานปลัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม

This research project was supported by Office of the Permanent Secretary, Ministry of Higher Education, Science, Research and Innovation

* Corresponding Author E-mail: Nurseeta.p@psu.ac.th

ลักษณะต่างกันด้วยวิธีเดียวกันสูงกว่าสหสัมพันธ์อื่นจึงใช้เป็นหลักฐานสนับสนุนความตรงแบบรู้เข้า ผลการตรวจสอบคุณภาพด้านความเที่ยงโดยการวัดความสอดคล้องภายในอยู่ในระดับที่ยอมรับได้และวิธีการทดสอบซ้ำพบว่าค่าสัมประสิทธิ์สหสัมพันธ์เฉลี่ยทั้งฉบับอยู่ในระดับสูง

คำสำคัญ แบบวัดบังคับเลือก แบบวัดพหุมิติ แบบวัดเชิงสถานการณ์ การเรียนรู้แบบผสมผสาน

Abstract

This study aimed to (1) develop a multidimensional forced-choice situational judgment test to assess undergraduate students' expectations toward 21st-century blended learning and (2) examine the quality of the instrument in terms of response balance, validity, and reliability. Validity evidence was gathered through content validity assessment, criterion-related validation, and construct validity using a multitrait-multimethod (MTMM) matrix. Reliability was assessed using internal consistency and test-retest methods. The instrument development involved four groups: three purposively selected samples, consisting of experts in educational measurement and evaluation, higher education instructors, and undergraduate students, along with one stratified random sample of 120 undergraduate students, with academic faculties serving as strata. The final version of the instrument comprised 25 items incorporating visual stimuli, representing five distinct scenarios, each measuring five dimensions of blended learning expectations. The response balance analysis revealed no initial signs of response bias. Correlations between scores obtained from the developed instrument and external measures provided supportive evidence of criterion-related validity. Construct validity, assessed using the MTMM matrix, demonstrated higher correlations between different traits measured by the same method, thereby supporting convergent validity. Internal consistency was found to be at an acceptable level, and test-retest reliability indicated a high average correlation coefficient across the full instrument.

Keywords: forced-choice measurement, multidimensional measurement, situational judgment test, blended learning

บทนำ

แบบวัดรูปแบบมาตรประมาณค่า (rating scale) เป็นแบบวัดที่ออกแบบให้นำเสนอข้อความแก่ผู้ตอบคราวละหนึ่งข้อความ เพื่อให้พิจารณาและตัดสินใจแต่ละข้อความแบบเชิงสัมบูรณ์ (absolute judgment) ผ่านระดับคะแนนที่กำหนดให้ ซึ่งจากกระบวนการสร้างที่ไม่ซับซ้อน ตอบง่าย สะดวกในการตรวจให้คะแนนและตรวจสอบคุณภาพ จึงเป็นรูปแบบที่ได้รับความนิยมอย่างแพร่หลาย (Schünemann & Ziegler,

2023) อย่างไรก็ตาม งานวิจัยจำนวนมากพบว่าข้อจำกัดสำคัญของรูปแบบมาตรประมาณค่า คือความเปราะบางต่อการบิดเบือนคำตอบที่อาจเกิดขึ้นทั้งโดยตั้งใจและไม่ตั้งใจ เช่น การตอบเกินจริง การตอบตามความคาดหวังทางสังคม การเลือกค่ากลางหรือค่าสุดขั้ว ปัญหาเหล่านี้พบได้ทั้งภายใต้สถานการณ์ที่มีผู้มีส่วนได้ส่วนเสียทั้งสูงและต่ำ (Brown & Maydeu-Olivares, 2011) ปัจจุบันการสร้างแบบวัดในรูปแบบบังคับเลือก (forced choice) จึงเป็นทางเลือกที่ได้รับความสนใจเพิ่มขึ้นอย่างต่อเนื่องในหลายบริบท โดยเฉพาะรูปแบบบังคับเลือกแบบพหุมิติ (multidimensional forced-choice) ผลการวิจัยจำนวนมากพบว่าการใช้รูปแบบบังคับเลือกแบบพหุมิติช่วยแก้ปัญหาการตั้งใจบิดเบือนคำตอบได้ดีและให้ความตรงตามเกณฑ์สัมพัทธ์ที่เท่ากันหรือดีกว่ารูปแบบมาตรประมาณค่า (Lee et al., 2024)

รูปแบบบังคับเลือกแบบพหุมิติเป็นแบบวัดที่อยู่บนพื้นฐานของการให้ผู้ตอบเกิดการตัดสินใจเชิงเปรียบเทียบ (comparative judgments) ซึ่งมีการนำเสนอข้อความที่มีคุณลักษณะทางจิตมิติต่างองค์ประกอบกันตั้งแต่ 2 ตัวเลือกขึ้นไปให้แก่ผู้ตอบ เพื่อให้ผู้ตอบเปรียบเทียบและบังคับให้ผู้ตอบเลือกข้อความตามวิธีการที่กำหนด (Brown & Maydeu-Olivares, 2011) โดยอาจเสนอข้อความที่เป็นตัวเลือกให้แก่ผู้ตอบได้ตามจำนวนที่กำหนด เช่น 2 ข้อความ (pair) 3 ข้อความ (triplet) หรือ 4 ข้อความ (quadruplet) และอาจกำหนดวิธีการให้ผู้ตอบเลือกข้อความที่ตรงกับตัวเองมากที่สุดขึ้นมาเพียง 1 ข้อความ (pick) หรือกำหนดให้ผู้ตอบเลือกขึ้นมา 2 ข้อความที่ตรงกับตนเองมากที่สุดและน้อยที่สุด (MOLE) หรือกำหนดให้ผู้ตอบเรียงลำดับข้อความที่ตรงกับตนเองจากมากที่สุดไปน้อยที่สุด (rank) จากลักษณะดังกล่าวจึงเป็นรูปแบบที่เหมาะสมกับแบบวัดที่มีวัตถุประสงค์หลักเพื่อเปรียบเทียบแต่ละองค์ประกอบภายในบุคคลเดียวกัน (Intra-individual comparison) ที่สามารถใช้การคิดค่าคะแนนได้ในทุกรูปแบบ อย่างไรก็ตามปัจจุบันมีผลการวิจัยที่พบว่าความซับซ้อนของกระบวนการตอบแบบวัดในรูปแบบบังคับเลือกทำให้เกิดการกระตุ้นให้ผู้ตอบเกิดกระบวนการสะท้อนตัวเองตลอดเวลา เพื่อเปรียบเทียบทุกคู่ของข้อความที่นำเสนอในแต่ละข้อ (Schünemann & Ziegler, 2023) จึงเป็นสาเหตุที่ทำให้เพิ่มการใช้ภาระทางปัญญา (cognitive load) ของผู้ตอบและส่งผลกระทบต่อประสิทธิภาพเนื่องไปจากความเหนื่อยล้าจากการทำแบบวัด (test fatigue) โดยเฉพาะการออกแบบที่มีการนำเสนอข้อความที่เป็นคู่เทียบในแต่ละข้อมากเกินไป ตลอดจนการมีจำนวนข้อที่ยาวเกินไป (Lee et al., 2021) ซึ่งอาจเป็นเหตุผลที่ทำให้ผู้ตอบไม่สนใจตอบและมีพฤติกรรมการตอบแบบไม่ใส่ใจ (Ward & Meade, 2023) ซึ่งอาจกระทบต่อความตรงของแบบวัดในที่สุด การวิจัยในครั้งนี้จึงศึกษาเพื่อต่อยอดองค์ความรู้ในการพัฒนาแบบวัดรูปแบบบังคับเลือกที่ลดภาระทางปัญญาของผู้ตอบลง ในขณะที่คุณภาพของแบบวัดยังคงเดิม

การลดภาระทางปัญญาของผู้ตอบแบบวัดรูปแบบบังคับเลือกอาจทำได้หลากหลายแนวทาง เช่น การกำหนดความยาวของแบบวัดอย่างเหมาะสม การกำหนดจำนวนข้อความในแต่ละข้อให้ไม่มากจนเกินไป การลดความซับซ้อนของข้อความที่นำเสนอ (Sass et al., 2020) การกำหนดขั้วของข้อความ (block polarity) ให้อยู่ในขั้วเดียวกันทั้งหมด (Lee et al., 2021) เป็นต้น นอกจากนี้หากมีการนำเสนอข้อความโดยการใช้สถานการณ์ต่าง ๆ ที่ผู้ตอบคุ้นเคยภายใต้บริบทที่ต้องการวัด ยังเป็นอีกหนึ่งปัจจัยที่ช่วยลดการใช้ภาระทางปัญญาของผู้ตอบได้ โดยการใช้สถานการณ์สามารถลดผลกระทบจากการได้รับผลการตอบที่ขึ้นอยู่กับ

ความสามารถในการสะท้อนตนเองของผู้ตอบ (Hooper et al., 2006) และการเพิ่มบริบทสถานการณ์เข้าไปในแบบวัดทำให้เพิ่มความตรงในการทำนายผลลัพธ์ได้ดีขึ้น (Swift & Peterson, 2019)

ด้วยขอบเขตองค์ความรู้ด้านการศึกษาที่ขยายตัวอย่างต่อเนื่อง รูปแบบการจัดการเรียนรู้แนวทางใหม่จึงถูกเสนอขึ้นหลากหลายรูปแบบ หนึ่งในรูปแบบที่สำคัญคือรูปแบบการเรียนรู้แบบผสมผสาน (blended learning) ซึ่งเป็นรูปแบบที่ใช้วิธีการที่หลากหลาย ผสมผสานการเรียนในห้องเรียนแบบเผชิญหน้าระหว่างผู้เรียนและผู้สอน กับการเรียนนอกห้องเรียนผ่านกิจกรรมต่าง ๆ (Yimyam et al., 2015) เป็นรูปแบบการเรียนรู้ที่ยืดหยุ่น ตอบสนองต่อความแตกต่างระหว่างบุคคลทั้งด้านรูปแบบการเรียน ความสนใจ และความสามารถของผู้เรียนแต่ละคน (Bonk et al., 2005) ซึ่งเหมาะสมสำหรับผู้เรียนในศตวรรษที่ 21 อย่างไรก็ตาม การออกแบบการเรียนรู้แบบผสมผสานมีหลากหลายรูปแบบ (Salleh et al., 2017) นอกจากนี้ แนวทางการวัดและประเมินผลการเรียนรู้ ช่องทางการรับข้อมูลย้อนกลับ รายละเอียดของข้อมูลย้อนกลับ ตลอดจนแนวทางการตัดสินผลการเรียนรู้ ยังเป็นองค์ประกอบที่สำคัญที่ส่งผลต่อประสิทธิภาพในการเรียนรู้ ผลการศึกษาจำนวนมากแสดงให้เห็นความสัมพันธ์ระหว่างคุณลักษณะด้านจิตพิสัยกับลักษณะต่างๆ ด้านการศึกษาของบุคคลอย่างต่อเนื่อง การออกแบบการจัดการเรียนรู้ให้เป็นไปอย่างมีประสิทธิภาพ ผู้สอนจึงควรขยายการศึกษาและทำความเข้าใจคุณลักษณะด้านจิตพิสัยของผู้เรียน

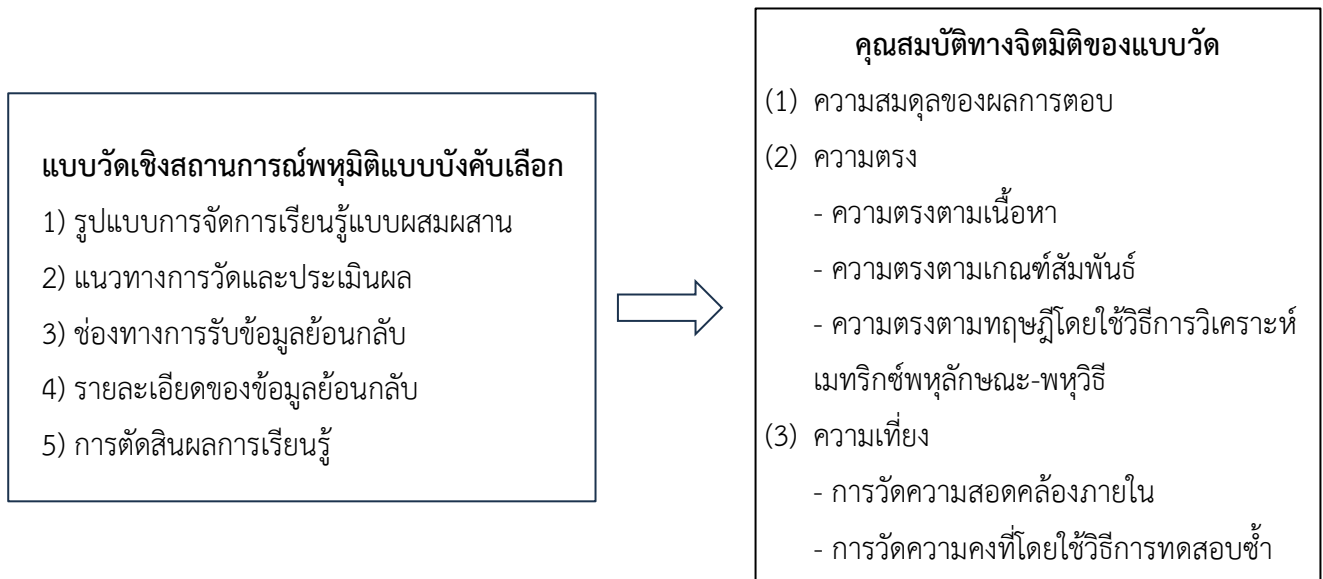
การศึกษาครั้งนี้จึงพัฒนาแบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือกสำหรับประเมินความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21 สำหรับนักศึกษาปริญญาตรี โดยกำหนด 5 สถานการณ์ เพื่อวัดความคาดหวัง 5 ด้าน รวมทั้งสิ้น 25 ข้อ กำหนดวิธีการตอบแบบ MOLE ซึ่งเป็นวิธีการตอบที่กำหนดให้ผู้ตอบเลือก 2 ข้อความที่ตรงกับตนเองมากที่สุดและน้อยที่สุดเพื่อเพิ่มสารสนเทศที่ได้จากการทำแบบวัด (Hontangas et al., 2015) และมีการคิดค่าคะแนนแบบกึ่งอิงตัวเอง (partially ipsative scoring) โดยมีการตรวจสอบคุณภาพของแบบวัดในรูปแบบบังคับเลือกด้วยวิธีการที่หลากหลาย ตั้งแต่การตรวจสอบคุณภาพเบื้องต้นจากผู้เชี่ยวชาญและผู้เกี่ยวข้องจำนวน 3 รอบ ตลอดจนการทดลองใช้เพื่อตรวจสอบความสมดุลของผลการตอบ (balanced response distributions) และหาสัญญาณเบื้องต้นของการมีความลำเอียงของแบบวัด (item bias) รวมถึงการตรวจสอบคุณภาพด้านความตรงและความเที่ยงหลากหลายวิธีการ เพื่อให้ได้มาซึ่งเครื่องมือที่ช่วยวางแผนและออกแบบการจัดการเรียนรู้ให้ตรงตามคาดหวังและเหมาะสมกับผู้เรียนซึ่งจะนำไปสู่การทำให้ผู้เรียนบรรลุผลลัพธ์การเรียนรู้ได้ในที่สุด

ความมุ่งหมายของการวิจัย

1. พัฒนาแบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือกสำหรับประเมินความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21 สำหรับนักศึกษาปริญญาตรี
2. ตรวจสอบคุณภาพของแบบวัดที่พัฒนาขึ้น ด้านความสมดุลของผลการตอบ (Balanced response distributions) ด้านความตรง (Validity) และความเที่ยง (Reliability)

กรอบแนวคิดของการวิจัย

งานวิจัยนี้พัฒนาแบบวัดความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21 สำหรับนักศึกษาปริญญาตรี โดยเป็นแบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือกที่วัดความคาดหวัง 5 ด้าน และทำการตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดด้วยวิธีการที่หลากหลาย โดยมีกรอบแนวคิดในการวิจัยดังนี้



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

การศึกษาครั้งนี้เป็นการพัฒนาเครื่องมือวัดที่มีการตรวจสอบคุณภาพจำนวน 4 รอบ (Schünemann & Ziegler, 2023) จากกลุ่มประชากร 4 กลุ่ม โดยแบ่งเป็นกลุ่มเป้าหมายที่มีการคัดเลือกแบบเฉพาะเจาะจงจำนวน 3 กลุ่ม ได้แก่ (1) ผู้เชี่ยวชาญด้านการวัดและประเมินผลที่มีประสบการณ์ไม่น้อยกว่า 5 ปี จำนวน 5 คน เพื่อตรวจสอบความเหมาะสมของการออกแบบโครงสร้างของแบบวัด (2) อาจารย์ผู้สอนในระดับอุดมศึกษาที่มีอายุงานไม่น้อยกว่า 5 ปี จำนวน 10 คน เพื่อตรวจสอบความเหมาะสมเกี่ยวกับสถานการณ์และข้อความในแบบวัด และ (3) นักศึกษาในระดับปริญญาตรี มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี จำนวน 10 คน เพื่อทดลองใช้เบื้องต้นและให้ข้อเสนอแนะเกี่ยวกับความเหมาะสมของสถานการณ์และข้อความต่าง ๆ ที่ปรากฏในแบบวัด นอกจากนี้ยังมีกลุ่มตัวอย่างที่ได้จากการสุ่มตัวอย่าง จำนวน 1 กลุ่ม ได้แก่ นักศึกษาระดับปริญญาตรี มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี จำนวน 120 คน เพื่อเป็นกลุ่มทดลองใช้ โดยมีการสุ่มตัวอย่างแบบชั้นภูมิ (Stratified sampling) ซึ่งมีคณะเป็นชั้นภูมิ คำนวณขนาดตัวอย่างโดยใช้โปรแกรมสำเร็จรูป G*power สำหรับการวิเคราะห์สหสัมพันธ์ (Correlation) ที่มีการกำหนดค่าอิทธิพล (Effect size) = 0.30 กำหนดค่า $\alpha = 0.05$ และกำหนดค่า $\text{power}(1-\beta) = 0.95$ ได้จำนวนตัวอย่างทั้งสิ้น 111 คน และเพื่อป้องกันปัญหาการได้ข้อมูลไม่ครบถ้วน จึงกำหนดกลุ่มตัวอย่างขึ้น 120 คน

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นแบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือกสำหรับประเมินความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21 สำหรับนักศึกษาปริญญาตรี จำนวน 25 ข้อ จาก 5 สถานการณ์ ประกอบด้วย (1) ผู้เรียนที่มีความคาดหวัง (2) ผู้เรียนที่ประสบปัญหา (3) ผู้แนะนำ (4) ผู้แก้ปัญหา และ (5) ผู้ออกแบบการจัดการเรียนรู้ ในแต่ละสถานการณ์ประกอบด้วยข้อคำถาม 5 ข้อ ได้แก่ (1) ความคาดหวังด้านรูปแบบการเรียน (2) ความคาดหวังด้านการวัดและประเมินผล (3) ความคาดหวังด้านช่องทางการรับข้อมูลย้อนกลับ (4) ความคาดหวังด้านรายละเอียดของข้อมูลย้อนกลับ และ (5) ความคาดหวังด้านการวัดและประเมินผลการเรียนรู้ ในแต่ละข้อปรากฏตัวเลือกที่มาจากแต่ละองค์ประกอบย่อย ซึ่งกำหนดวิธีการตอบแต่ละข้อโดยใช้รูปแบบ MOLE ที่ระบุให้ผู้ตอบพิจารณาเปรียบเทียบข้อความที่เสนอเป็นตัวเลือกทั้งหมดแล้วตัดสินใจเลือกข้อความที่ตรงกับตนเองมากที่สุดขึ้นมา 1 ข้อความ และข้อความที่ตรงกับตนเองน้อยที่สุดขึ้นมา 1 ข้อความ ใช้การคิดค่าคะแนนแบบกึ่งอิงตัวเอง (Partially ipsative scoring) โดยให้คะแนน +1 กับองค์ประกอบที่ผู้ตอบระบุว่าตรงกับตนเองมากที่สุด และให้คะแนน -1 กับองค์ประกอบที่ผู้ตอบระบุว่าตรงกับตนเองน้อยที่สุด โดยมีรายละเอียดองค์ประกอบย่อย ดังนี้

ตาราง 1 นิยามเชิงปฏิบัติการของความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21

ด้าน	องค์ประกอบ	คำอธิบาย
ความคาดหวังด้านรูปแบบการเรียน	(A) Flipped-classroom model	จัดการเรียนรู้แบบเผชิญหน้าทุกคาบเรียน โดยผู้เรียนศึกษาเนื้อหาด้วยตนเองก่อนถึงคาบเรียนตามที่ผู้สอนกำหนด เมื่อถึงคาบเรียนจึงจัดกิจกรรมในชั้นเรียนโดยมีผู้เรียนเป็นศูนย์กลางและผู้สอนเป็นผู้ชี้แนะ
	(B) Flex model	ให้อิสระแก่ผู้เรียนในการบริหารการเรียนรู้ด้วยตนเอง จากทรัพยากรการเรียนรู้ที่ผู้สอนกำหนด ไม่จำเป็นต้องจัดการเรียนรู้แบบเผชิญหน้าทุกคาบเรียน แต่มีช่องทางสนับสนุนในการช่วยแก้ปัญหาให้แก่ผู้เรียนอย่างชัดเจน
	(C) Self-blend model	เปิดโอกาสให้ผู้เรียนผสมผสานเนื้อหาและวิธีการเรียนได้ตามอัธยาศัย โดยสถานศึกษาและผู้สอนทำหน้าที่เป็นแกนกลางในการช่วยควบคุม ให้ข้อเสนอแนะ และเป็นผู้อำนวยความสะดวกในการผสมผสานของผู้เรียนเท่านั้น
	(D) Enriched-virtual model	จัดการเรียนรู้ที่ใช้ระบบเครือข่ายเป็นห้องเรียนเสมือน โดยผู้สอนเป็นผู้จัดการเรียนรู้เป็นหลัก แต่ทั้งหมดถูกจัดขึ้นในรูปแบบออนไลน์ ซึ่งผู้เรียนถูกกำหนดให้มาเข้าห้องเรียนแบบเผชิญหน้าเฉพาะในกิจกรรมที่ผู้สอนกำหนดเท่านั้น
ความคาดหวังด้านการวัดและประเมินผล	(A) Assessment AS learning	ใช้กระบวนการวัดและประเมินผลเป็นเครื่องมือในการขับเคลื่อนการเรียนรู้
	(B) Assessment FOR learning	ใช้กระบวนการวัดและประเมินผลเป็นเครื่องมือในการช่วยพัฒนาการเรียนรู้

ตาราง 1 (ต่อ)

ด้าน	องค์ประกอบ	คำอธิบาย
	(C) Assessment OF learning	ใช้กระบวนการวัดและประเมินผลเป็นเครื่องมือในการสรุปผลการเรียนรู้
ความคาดหวังต่อช่องทางการรับข้อมูลย้อนกลับ	(A) Cohort feedback	ให้ข้อมูลย้อนกลับเป็นหมู่คณะแบบส่วนรวมภายในห้องเรียน โดยเปิดโอกาสให้ผู้สอนและเพื่อนร่วมกันให้ข้อมูลย้อนกลับ
	(B) Individual feedback	ให้ข้อมูลย้อนกลับโดยผู้สอนแบบส่วนตัวระหว่างผู้สอนกับผู้เรียนเป็นรายบุคคล
	(C) Computer-based feedback	ให้ข้อมูลย้อนกลับโดยระบบคอมพิวเตอร์อัตโนมัติแบบส่วนตัวระหว่างผู้เรียนและ ระบบคอมพิวเตอร์
ความคาดหวังต่อรายละเอียดของข้อมูลย้อนกลับ	1. Knowledge of result feedback	ให้ข้อมูลเพียงผลลัพธ์ของการกระทำของผู้เรียนเท่านั้นว่า ผู้เรียนทำผิดหรือถูก โดยไม่มีการบอกข้อมูลอื่นหรือคำอธิบายเพิ่มเติม
	2. Corrective feedback	ให้ข้อมูลย้อนกลับแบบบอกคำตอบที่ถูกต้อง หากผู้เรียนทำถูกจะบอกว่าถูก และถ้าผู้เรียนทำผิดจะบอกว่าผิดพร้อมทั้งบอกคำตอบที่ถูกต้องให้กับผู้เรียน แต่ไม่มีคำอธิบายเพิ่มเติม
	3. Specific and directive elaborated feedback	ให้ข้อมูลย้อนกลับแบบอธิบายรายละเอียดอย่างเจาะจงและชี้แนะ อธิบายรายละเอียดตั้งแต่โจทย์ อธิบายคำตอบ และอธิบายข้อผิดพลาด โดยใช้ข้อมูลจากโจทย์และคำตอบของผู้เรียนเป็นหลัก
	4. General and facilitative elaborated feedback	ให้ข้อมูลย้อนกลับแบบอธิบายรายละเอียดแบบทั่วไปโดยใช้ตัวอย่างช่วย อธิบายโดยยกตัวอย่างในสถานการณ์อื่นที่เทียบเคียงกับโจทย์แล้วจึงแนะแนวทางให้ผู้เรียนประยุกต์ใช้กับโจทย์ข้อนั้นด้วยตนเอง
ความคาดหวังด้านการตัดสินผลการเรียนรู้	1. Criterion-referenced assessment	การตัดสินผลการเรียนรู้แบบอิงเกณฑ์ โดยมีการเปรียบเทียบกับเกณฑ์ที่เป็นมาตรฐาน
	2 Norm-referenced assessment	การตัดสินผลการเรียนรู้แบบอิงกลุ่ม โดยมีการเปรียบเทียบระหว่างบุคคล
	3. Self-referenced assessment	การตัดสินผลการเรียนรู้แบบอิงตนเอง โดยเปรียบเทียบกับพัฒนาการของตัวนักเรียนเอง

3. การเก็บรวบรวมข้อมูล

3.1 ออกแบบ 5 สถานการณ์ และสร้างแบบวัดรูปแบบเชิงสถานการณ์พหุมิติแบบบังคับเลือกเพื่อวัดความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21 พร้อมภาพประกอบ โดยในแต่ละสถานการณ์มีการวัดความคาดหวังทั้งสิ้น 5 ด้าน รวมเป็น 25 ข้อ

3.2 ตรวจสอบคุณภาพของแบบวัดรอบที่ 1 โดยผู้เชี่ยวชาญด้านการวัดและประเมินผลเพื่อตรวจสอบความเหมาะสมของการออกแบบโครงสร้างของแบบวัด ตรวจสอบคุณภาพรอบที่ 2 โดยอาจารย์

ผู้สอนระดับปริญญาตรีเพื่อตรวจสอบความเหมาะสมเกี่ยวกับสถานการณ์และข้อความในแบบวัด ตรวจสอบคุณภาพรอบที่ 3 โดยการทดลองใช้เบื้องต้นจากนักศึกษาระดับปริญญาตรีเพื่อให้ข้อเสนอแนะเกี่ยวกับความเหมาะสมของสถานการณ์และข้อความต่างๆที่ปรากฏในแบบวัด และปรับปรุงข้อความตามข้อเสนอแนะของกลุ่มต่าง ๆ

3.3 นำแบบวัดที่ตรวจสอบและปรับปรุงแล้วไปใช้ในการเก็บข้อมูลจากกลุ่มตัวอย่างจำนวน 120 คน และทำการตรวจสอบคุณภาพด้านความสมดุลของผลการตอบ (Balanced response distributions) โดยการตรวจสอบสัดส่วนของผลการตอบในแต่ละตัวเลือกภายในข้อว่ามีความสมดุลกันหรือไม่ และทำการปรับปรุงข้อความที่ทำให้เกิดความไม่สมดุลของผลการตอบ

3.4 ตรวจสอบคุณภาพด้านความตรง ซึ่งประกอบด้วย ความตรงตามเกณฑ์สัมพัทธ์ โดยทำการเปรียบเทียบผลที่ได้จากแบบวัดที่พัฒนาขึ้นกับผลที่ได้จากแบบวัดที่วัดคุณลักษณะเดียวกันหรือใกล้เคียงกัน โดยการศึกษาครั้งนี้ใช้แบบวัด แบบวัดการรับรู้ต่อประสบการณ์การเรียนรู้แบบห้องเรียนกลับด้าน (perception of flipped learning experience) (Muhlisoh et al., 2020) แบบวัดมโนทัศน์เกี่ยวกับการให้ข้อมูลย้อนกลับ (conceptions of feedback questionnaire-II :ScoF-II) (Peterson & Irving, 2008) และความตรงตามทฤษฎีโดยใช้วิธีการวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี (Multitrait-Multimethod: MTMM) (Campbell & Fiske, 1959) แบบ 3x2 ซึ่งมีการวัดลักษณะด้านความคาดหวังต่อรูปแบบการเรียนรู้แบบผสมผสานที่แตกต่างกัน 3 ลักษณะและวิธีการวัดที่แตกต่างกัน 2 วิธี ได้แก่ วิธีการวัดโดยใช้ส่วนหนึ่งของแบบวัดที่พัฒนาขึ้นและวิธีการวัดโดยใช้แบบวัดอื่นซึ่งปรับปรุงมาจาก แบบวัดการรับรู้ต่อประสบการณ์การเรียนรู้แบบห้องเรียนกลับด้าน (Muhlisoh et al., 2020) แบบวัดมโนทัศน์เกี่ยวกับการให้ข้อมูลย้อนกลับ (Peterson & Irving, 2008) รวมทั้งการตรวจสอบคุณภาพด้านความเที่ยง ซึ่งประกอบด้วย การวัดความสอดคล้องภายใน และการวัดความคงที่โดยใช้วิธีการทดสอบซ้ำจำนวน 2 สัปดาห์ และปรับปรุงแบบวัดตามผลการตรวจสอบคุณภาพเพื่อให้ได้แบบวัดที่มีคุณภาพเพียงพอต่อการนำไปใช้งานในสถานการณ์จริง

4. การวิเคราะห์ข้อมูล

4.1 ตรวจสอบความตรงตามเนื้อหาโดยใช้สูตรดัชนีความสอดคล้องของข้อคำถาม (index of content validity) โดยคัดเลือกเฉพาะข้อความที่มีค่าความสอดคล้อง 0.70 ขึ้นไป (Kanjawasee, 2013)

4.2 การตรวจสอบคุณภาพด้านความสมดุลของผลการตอบโดยใช้สัดส่วนและร้อยละ สำหรับองค์ประกอบที่มี 3 ตัวเลือก กำหนดการแจกแจงที่คาดหวัง $33.33\% \pm 15\%$ ในขณะที่องค์ประกอบที่มี 4 ตัวเลือก กำหนดการแจกแจงที่คาดหวัง $25\% \pm 10\%$

4.3 การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรในการตรวจสอบคุณภาพด้านความตรงตามเกณฑ์สัมพัทธ์ ความตรงตามทฤษฎีโดยใช้วิธีการวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี และความเที่ยงแบบการทดสอบซ้ำ โดยใช้สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's correlation coefficient)

4.4 การวิเคราะห์ความสอดคล้องภายในโดยการวิเคราะห์สัมประสิทธิ์แอลฟาของครอนบาค

ผลการวิจัย

1) ผลการพัฒนาแบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือกสำหรับประเมินความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21 สำหรับนักศึกษาปริญญาตรี

การออกแบบ 5 สถานการณ์ เพื่อวัดความคาดหวัง 5 ด้าน จึงได้แบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือก จำนวน 25 ข้อ ในแต่ละข้อมีการวัดเพียง 1 มิติ โดยในแต่ละมิติมีข้อคำถามที่ใช้ในการวัดจำนวน 5 ข้อ จาก 5 สถานการณ์ ผ่านการตรวจสอบคุณภาพเบื้องต้นเพื่อพัฒนาและปรับปรุง 3 รอบ จากกลุ่มผู้เชี่ยวชาญด้านการวัดและประเมินผล กลุ่มอาจารย์ผู้สอนระดับอุดมศึกษา และกลุ่มนักศึกษาระดับปริญญาตรี จนกระทั่งได้แบบวัดที่มีคุณภาพเป็นไปตามเกณฑ์ที่กำหนด โดยมีดัชนีความสอดคล้องของข้อคำถามทั้งฉบับ 0.92 ก่อนการทดลองใช้กับกลุ่มตัวอย่างจริง เพื่อตรวจสอบคุณภาพในลำดับถัดไป

ตัวอย่าง สถานการณ์ที่ 2 ผู้เรียนที่ประสบปัญหา

จากปัญหาข้างต้น นักศึกษาคาดหวังว่าจะได้รับข้อมูลย้อนกลับในรูปแบบใดมากที่สุดและน้อยที่สุด

ข้อความ	องค์ประกอบ	ตัวอย่างภาพประกอบ
(A) ให้ข้อมูลย้อนกลับแบบเป็นหมู่คณะในห้องเรียน โดยเปิดโอกาสให้อาจารย์และเพื่อนร่วมชั้นเรียนร่วมกันให้ข้อมูลย้อนกลับ	การให้ข้อมูลย้อนกลับแบบหมู่คณะ (cohort feedback)	
(B) อาจารย์ให้ข้อมูลย้อนกลับผู้เรียนเป็นรายบุคคล แบบเผชิญหน้าและเป็นส่วนตัวระหว่างนักศึกษาและอาจารย์	การให้ข้อมูลย้อนกลับรายบุคคล (individual feedback)	
(C) อาจารย์บันทึกข้อมูลย้อนกลับแก่นักศึกษาผ่านระบบคอมพิวเตอร์เพื่อให้นักศึกษาสามารถเรียกดูข้อมูลย้อนกลับได้ตามอัธยาศัย	การให้ข้อมูลย้อนกลับโดยระบบเทคโนโลยีสารสนเทศ (computer-based feedback)	

2) ผลตรวจสอบคุณภาพของแบบวัดที่พัฒนาขึ้น ด้านความสมดุลของผลการตอบ (balanced response distributions) ด้านความตรง (validity) และความเที่ยง (reliability)

■ ผลการตรวจสอบคุณภาพด้านความสมดุลของผลการตอบ

แบบวัดรูปแบบบังคับเลือกที่ดี ทุกข้อความที่ถูกเสนอเป็นตัวเลือกในแต่ละข้อควรดึงดูดให้ผู้ตอบเลือกตอบได้ในระดับเดียวกันหรือใกล้เคียงกัน ด้วยเหตุนี้การตรวจสอบสัดส่วนการเลือกตอบของผู้ตอบแบบวัดว่ามีความสมดุลกันในแต่ละตัวเลือกหรือไม่ จึงเป็นตัวชี้วัดแรกในการสร้างแบบวัดรูปแบบบังคับเลือกที่มีประสิทธิภาพ เนื่องจากสัดส่วนดังกล่าวสามารถสะท้อนได้ว่าแต่ละตัวเลือกที่แบบวัดนำเสนอให้กับผู้ตอบนั้นมีความน่าจะเป็นในการถูกเลือกตอบอย่างสมดุลกันหรือไม่ หากมีบางตัวเลือกมีผู้ตอบมากเกินไป ในขณะที่บางตัวเลือกภายในข้อเดียวกันไม่มีผู้เลือกตอบเลย เป็นสัญญาณเบื้องต้นของการมีความลำเอียงของแบบวัด

(item bias) (Schünemann & Ziegler, 2023) จากแบบวัดที่พัฒนาขึ้น แต่ละข้อกำหนดให้ผู้ตอบเลือก 2 ตัวเลือก คือ ตัวเลือกที่ตรงกับตนเองมากที่สุดและตรงกับตนเองน้อยที่สุด เมื่อพิจารณาความสมดุลของผลการตอบในภาพรวมของด้านที่มี 4 ตัวเลือก พบว่ามีการแจกแจงผลการตอบอยู่ระหว่าง 15.89% ถึง 34.57% ซึ่งแตกต่างจากการแจกแจงที่คาดหวัง 25% ในระดับหนึ่ง แต่ยังอยู่ในขอบเขตที่กำหนด ($25\% \pm 10\%$) ในขณะที่ด้านที่มี 3 ตัวเลือก มีการแจกแจงผลการตอบระหว่าง 29.35% - 37.55% ซึ่งแตกต่างจากการแจกแจงที่คาดหวัง 33.33% เพียงเล็กน้อย และเมื่อพิจารณาการแจกแจงผลการตอบในแต่ละตัวเลือก พบว่าไม่มีตัวเลือกใดในองค์ประกอบใดที่ไม่ถูกเลือกตอบเลย ไม่ว่าจะเป็นการเลือกตอบเพราะตรงกับตนเองมากที่สุดหรือเลือกตอบเพราะตรงกับตนเองน้อยที่สุด โดยอัตราการเลือกตอบต่ำที่สุด คือ การเลือก enriched-virtual model เป็นความคาดหวังด้านรูปแบบการเรียนรู้ที่ตรงกับตนเองมากที่สุด โดยมีการเลือกตอบเพียง 4.85% ของการเลือกตอบทั้งหมดในด้านดังกล่าว

ตาราง 2 การแจกแจงผลการตอบของกลุ่มทดลองใช้

ด้าน	การแจกแจงของผลการตอบ		
	ทั้งหมด	มากที่สุด	น้อยที่สุด
ความคาดหวังด้านรูปแบบการเรียนรู้			
Flipped-classroom model	31.99%	22.82%	9.17%
Flex model	18.99%	12.23%	6.76%
Self-blend model	20.07%	10.14%	9.93%
Enriched-virtual model	28.95%	4.85%	24.10%
รวม	100%	100%	
ความคาดหวังด้านการวัดและประเมินผล			
Assessment as learning	30.86%	15.70%	15.16%
Assessment for learning	36.14%	24.67%	11.47%
Assessment of learning	33.00%	10.63%	22.37%
รวม	100%	100%	
ความคาดหวังต่อช่องทางการรับข้อมูลย้อนกลับ			
Cohort feedback	37.55%	17.01%	20.54%
Individual feedback	30.57%	17.91%	12.66%
Computer-based feedback	31.88%	15.83%	16.05%
รวม	100%	100%	
ความคาดหวังต่อรายละเอียดของข้อมูลย้อนกลับ			
Knowledge of result feedback	34.57%	6.52%	28.06%
Corrective feedback	15.89%	6.89%	9.00%
Specific and directive elaborated	22.14%	16.48%	5.66%
General and facilitative elaborated	27.39%	20.23%	7.16%
รวม	100%	100%	

ตาราง 2 (ต่อ)

ด้าน	การแจกแจงของผลการตอบ		
	ทั้งหมด	มากที่สุด	น้อยที่สุด
ความคาดหวังด้านการตัดสินผลการเรียนรู้			
Criterion-referenced assessment	35.73%	25.28%	10.45%
Norm-referenced assessment	29.35%	6.76%	22.59%
Self-referenced assessment	34.92%	18.11%	16.81%
รวม	100%	100%	

■ ผลการตรวจสอบคุณภาพด้านความตรง

(a) คุณภาพด้านความตรงตามเนื้อหา

แบบวัดที่พัฒนาขึ้นผ่านกระบวนการตรวจสอบคุณภาพด้านความตรงตามเนื้อหาทั้งหมดทั้งสิ้น 3 รอบ โดยรอบที่ 1 ผลการตรวจสอบคุณภาพพบว่าแบบวัดมีค่าดัชนีความสอดคล้องของข้อคำถามกับวัตถุประสงค์ 0.82 – 1.00 และทำการปรับปรุงการเขียนสถานการณ์และข้อคำถามให้สอดคล้องตามโครงสร้างของแบบวัดที่กำหนดไว้ ผลการตรวจสอบคุณภาพรอบที่ 2 พบว่า แบบวัดมีค่าดัชนีความสอดคล้องของข้อคำถามกับวัตถุประสงค์ 0.85 -1.00 และทำการปรับปรุงการเขียนสถานการณ์และข้อความในแบบวัดให้สอดคล้องกับสถานการณ์ที่กำหนด ผลการตรวจสอบคุณภาพรอบที่ 3 พบว่าแบบวัดมีค่าดัชนีความสอดคล้องของข้อคำถามกับวัตถุประสงค์ 0.96 – 1.00 และได้ทำการปรับปรุงการเขียนสถานการณ์และข้อคำถามให้สอดคล้องกับธรรมชาติและความเข้าใจของผู้ใช้แบบวัดมากขึ้น ซึ่งผลการตรวจสอบคุณภาพด้านความตรงตามเนื้อหาทั้ง 3 รอบ พบว่า คุณภาพของแบบวัดเป็นไปตามเกณฑ์ที่กำหนด

(b) ความตรงตามเกณฑ์สัมพัทธ์

การวิเคราะห์สหสัมพันธ์โดยใช้ค่า Pearson's correlation coefficient ระหว่างคะแนนของกลุ่มตัวอย่างที่ได้จากการทำแบบวัดที่พัฒนาขึ้นกับคะแนนที่ได้จากแบบวัดอื่นที่วัดคุณลักษณะสอดคล้องกัน เพื่อตรวจสอบคุณภาพด้านความตรงตามเกณฑ์สัมพัทธ์ พบว่า คะแนนความคาดหวังด้านรูปแบบการเรียนรู้แบบ flipped-classroom model ที่ได้จากแบบวัดที่พัฒนาขึ้น มีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 กับคะแนนที่ได้จากการแบบวัดการรับรู้ต่อประสบการณ์การเรียนรู้แบบห้องเรียนกลับด้าน (Muhlisoh et al., 2020) โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ 0.48 นอกจากนี้ ยังพบว่าคะแนนด้านความคาดหวังต่อช่องทางการรับข้อมูลย้อนกลับแบบ Cohort feedback ที่ได้จากแบบวัดที่พัฒนาขึ้น มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 กับคะแนนที่ได้จากองค์ประกอบ Peer help ในแบบวัดมโนทัศน์เกี่ยวกับการให้ข้อมูลย้อนกลับ (Peterson & Irving, 2008) โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ 0.33 รวมถึงคะแนนด้านความคาดหวังเกี่ยวกับการวัดและประเมินผลแบบ Assessment for learning ที่ได้จากแบบวัดที่พัฒนาขึ้น มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 กับคะแนนที่ได้จากองค์ประกอบ Ignore ในแบบวัดมโนทัศน์เกี่ยวกับการให้ข้อมูลย้อนกลับ (Peterson & Irving, 2008) โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ 0.28 เมื่อมีการกลับทิศของข้อความเชิงลบ จากความสัมพันธ์ของคะแนนที่ได้จากแบบวัดที่พัฒนาขึ้นกับแบบวัดอื่น

ที่วัดคุณลักษณะที่สอดคล้องกัน จึงเป็นส่วนหนึ่งของหลักฐานที่ยืนยันถึงคุณภาพด้านความตรงตามเกณฑ์สัมพันธของแบบวัดที่พัฒนาขึ้น

(ค) ความตรงตามทฤษฎีโดยใช้วิธีการวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี

จากการนำคะแนนความคาดหวังด้านรูปแบบการเรียนรู้แบบ flipped-classroom model ความคาดหวังต่อช่องทางการรับข้อมูลย้อนกลับแบบ Cohort feedback, และความคาดหวังเกี่ยวกับการวัดและประเมินผลแบบ Assessment for learning ที่ได้จากแบบวัดที่พัฒนาขึ้น รวมถึงคะแนนที่ได้จากแบบวัดการรับรู้ต่อประสบการณ์การเรียนรู้แบบห้องเรียนกลับด้าน (Muhlisoh et al., 2020) และแบบวัดมโนทัศน์เกี่ยวกับการให้ข้อมูลย้อนกลับ (Peterson & Irving, 2008) ในองค์ประกอบ Peer help และ Ignore มาสร้างตารางวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี (Multitrait-Multimethod: MTMM) ดังตาราง 3 พบว่าสัมประสิทธิ์สหสัมพันธ์ความตรงมีนัยสำคัญและมีค่าค่อนข้างสูงเมื่อเทียบกับสัมประสิทธิ์สหสัมพันธ์อื่น จึงใช้เป็นหลักฐานสนับสนุนความตรงแบบลู่เข้า (convergent validity) ในขณะที่ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการวัดลักษณะต่างกันด้วยวิธีเดียวกัน (HTMM) กับสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการวัดลักษณะต่างกันด้วยวิธีการต่างกัน (HTMM) มีค่าใกล้เคียงกัน โดยค่า HTMM มีค่าสูงกว่าค่า HTMM เพียงเล็กน้อย จึงเป็นหลักฐานแสดงว่าไม่มีความลำเอียงของวิธีการวัด (method bias) เกิดขึ้น

ตาราง 3 ตารางวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี

ลักษณะ	วิธีวัด	Flipped classroom		Assessment for learning		Cohort feedback	
		A ¹	B ²	A ¹	B ²	A ¹	B ²
Flipped classroom	A ¹	0.66 ³					
	B ²	0.48 ⁴	0.78 ³				
Assessment for learning	A ¹	0.04 ⁵	0.08	0.64 ³			
	B ²	0.13	0.04 ⁵	0.28 ⁴	0.73 ³		
Cohort feedback	A ¹	-0.06 ⁵	0.03	0.11 ⁵	0.04	0.82 ³	
	B ²	0.23	0.16 ⁵	0.09	0.11 ⁵	0.33 ⁴	0.88 ³

¹วัดโดยใช้แบบวัดที่พัฒนาขึ้น, ²วัดโดยใช้แบบวัดอื่นที่วัดคุณลักษณะเดียวกัน, ³สัมประสิทธิ์ความเที่ยงโดยใช้ Cronbach's alpha, ⁴สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการวัดลักษณะต่างกันด้วยวิธีเดียวกัน (HTMM), ⁵ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการวัดลักษณะต่างกันด้วยวิธีเดียวกัน (HTMM)

■ ผลการตรวจสอบคุณภาพด้านความเที่ยง

การตรวจสอบคุณภาพด้านความเที่ยงด้วยวิธีการวัดความสอดคล้องภายในโดยใช้ Cronbach's alpha พบว่าค่าความเที่ยงของแต่ละองค์ประกอบอยู่ในช่วง 0.49 – 0.68 โดยมีค่าความเที่ยงรวมทั้งฉบับอยู่ที่ 0.57 นอกจากนี้ เมื่อตรวจสอบคุณภาพด้านความเที่ยงโดยวิธีการทดสอบซ้ำ (test-retest reliability) ในระยะเวลาห่างกัน 2 สัปดาห์ พบว่าค่า Pearson's correlation coefficient เฉลี่ยทั้งฉบับมีค่า อยู่ในช่วง 0.68 – 0.88 และมีค่ารวมทั้งฉบับอยู่ที่ 0.77 ดังตาราง 4

ตาราง 4 ผลการตรวจสอบคุณภาพด้านความเที่ยง

องค์ประกอบ	วิธีวัดความสอดคล้องภายใน	วิธีการสอบซ้ำ
ความคาดหวังด้านรูปแบบการเรียนรู้	0.53	0.82
ความคาดหวังด้านการวัดและประเมินผล	0.49	0.75
ความคาดหวังต่อช่องทางการรับข้อมูลย้อนกลับ	0.68	0.88
ความคาดหวังต่อรายละเอียดของข้อมูลย้อนกลับ	0.56	0.72
ความคาดหวังด้านการตัดสินผลการเรียนรู้	0.67	0.68
รวม	0.57	0.77

อภิปรายผล

1 การพัฒนาแบบวัดเชิงสถานการณ์พหุมิติแบบบังคับเลือกสำหรับประเมินความคาดหวังต่อการเรียนรู้แบบผสมผสานในศตวรรษที่ 21 สำหรับนักศึกษาปริญญาตรี

แบบวัดรูปแบบเชิงสถานการณ์พหุมิติแบบบังคับเลือกซึ่งมีข้อค้นพบจากงานวิจัยจำนวนมากว่าสามารถช่วยลดปัญหาเกี่ยวกับความแปรปรวนต่อการบิดเบือนคำตอบที่อาจเกิดขึ้นได้ทั้งโดยตั้งใจและไม่ตั้งใจของผู้ตอบแบบวัดในรูปแบบมาตรฐานค่า (Brown & Maydeu-Olivares, 2011) แต่ในอีกทางหนึ่งพบว่า เป็นรูปแบบที่กระตุ้นให้ผู้ตอบเกิดกระบวนการสะท้อนตัวเองตลอดเวลา (Schünemann & Ziegler, 2023) เพื่อเปรียบเทียบทุกคู่ของข้อความที่นำเสนอในแต่ละข้อ จึงเป็นสาเหตุที่ทำให้การใช้แบบวัดในรูปแบบนี้เพิ่มการใช้ภาระทางปัญญา (cognitive load) ของผู้ตอบมากขึ้น และส่งผลต่อเนื่องไปยังความเหนื่อยล้าจากการทำแบบวัด (test fatigue) โดยเฉพาะการออกแบบที่มีการนำเสนอข้อความที่เป็นคู่เทียบในแต่ละข้อ มากจนเกินไป ตลอดจนการมีจำนวนข้อที่ยาวเกินไป (Lee et al., 2021) นำไปสู่การตอบแบบไม่มีคุณภาพในที่สุด

จากผลการตรวจสอบคุณภาพเบื้องต้นของแบบวัดที่พัฒนาขึ้นซึ่งพบว่าแบบวัดมีคุณภาพเบื้องต้นเพียงพอต่อการนำไปทดลองใช้กับกลุ่มตัวอย่างขนาดใหญ่ต่อไป คาดว่าเกิดจากกระบวนการออกแบบแบบวัดที่มุ่งเน้นการลดการใช้ภาระทางปัญญาของผู้ตอบโดยมีการกำหนดสถานการณ์ต่าง ๆ ที่ผู้ตอบคุ้นเคยและใช้รูปภาพประกอบสถานการณ์เพื่อลดการกระตุ้นการใช้ภาระทางปัญญาที่สูงเกินไป ทำให้ผู้ตอบเกิดการสะท้อนตัวเองในปริมาณที่พอเหมาะและเป็นคำตอบที่ตรงกับสิ่งที่มุ่งวัดมากที่สุด อีกทั้งคาดว่า เป็นผลมาจากกระบวนการพัฒนาแบบวัดที่กำหนดให้ผ่านกระบวนการตรวจสอบคุณภาพเบื้องต้นเพื่อการพัฒนาและปรับปรุงจำนวน 3 รอบ จากกลุ่มเป้าหมายที่แตกต่างกัน และมีการตรวจสอบทั้งเชิงปริมาณและเชิงคุณภาพ (Schünemann & Ziegler, 2023) ผลการตรวจสอบคุณภาพเบื้องต้นหลังจากการปรับปรุงจึงพบว่าทุกข้อความที่ปรากฏในแบบวัดมีคุณภาพเป็นไปตามเกณฑ์ที่กำหนดและมีแนวโน้มที่จะมีคุณภาพเพียงพอที่จะนำไปทดลองใช้ต่อไป

2. ตรวจสอบคุณภาพของแบบวัดที่พัฒนาขึ้น ด้านความสมดุลของผลการตอบ (balanced response distributions) ด้านความตรง (validity) และความเที่ยง (reliability)

จากผลตรวจสอบที่พบว่า ในภาพรวมขององค์ประกอบที่มี 3 ตัวเลือก มีการแจกแจงใกล้เคียงกับการแจกแจงที่คาดหวัง 33.33% (29.35% - 37.55%) แสดงให้เห็นถึงความสมดุลของการแจกแจงผลการตอบ

ในขณะที่ในภาพรวมขององค์ประกอบที่มี 4 ตัวเลือก ซึ่งมีการแจกแจงที่คาดหวัง 25% พบว่า ผลการตอบจริงมีความแตกต่างจากการแจกแจงที่คาดหวังในระดับหนึ่ง (15.89% ถึง 34.57%) เมื่อพิจารณาในรายละเอียดโดยการจำแนกผลการเลือกตอบออกเป็นตรงกับตนเองมากที่สุดและตรงกับตนเองน้อยที่สุด พบว่า ตัวเลือกที่มีอัตราการเลือกตอบต่ำที่สุด คือ ตัวเลือกที่ระบุว่า “enriched-virtual model” เป็นความคาดหวังด้านรูปแบบการเรียนรู้ที่ตรงกับตนเองมากที่สุด โดยมีการเลือกเพียง 57 ครั้ง จากจำนวนทั้งสิ้น 1,180 ครั้ง คิดเป็น 4.85% ซึ่งอาจเกิดได้จากหลายสาเหตุ เช่น ปัญหาด้านการเขียนแต่ละข้อความของตัวเลือกที่วัดความคาดหวังรูปแบบการเรียนรู้แบบ enriched-virtual model ในแบบวัด อาจเข้าใจยาก ให้ข้อมูลไม่เพียงพอ หรืออาจกระตุ้นความรู้สึกทางลบที่ทำให้ผู้ตอบไม่รู้สึกลอยเลือกตอบ หรืออาจเป็นสาเหตุที่เกี่ยวกับการสะท้อนถึงพฤติกรรมจริงและความคาดหวังจริงของผู้ตอบแบบวัดที่มีประสบการณ์ที่ไม่ดีเกี่ยวกับการจัดการเรียนรู้ในรูปแบบ enriched-virtual model จริง โดย enriched-virtual model เป็นรูปแบบที่ใช้ระบบเครือข่ายเป็นห้องเรียนเสมือน ซึ่งผู้สอนเป็นผู้จัดการเรียนรู้เป็นหลักและทั้งหมดจะถูกจัดขึ้นในรูปแบบออนไลน์ ทั้งนี้ Lakho et al. (2023) พบว่าการเรียนรู้ในรูปแบบห้องเรียนเสมือนเพียงอย่างเดียวไม่เพียงพอสำหรับนักศึกษาระดับปริญญาตรีเนื่องจากผู้เรียนจำเป็นต้องได้รับการฝึกแบบปฏิบัติด้วย สอดคล้องกับ Prayitno et al. (2023) ที่พบว่า enriched-virtual model มีข้อจำกัดในด้านการสนับสนุนผู้เรียนในกระบวนการลงมือปฏิบัติ จึงจำเป็นต้องมีการกำหนดแนวทางที่ชัดเจนและเหมาะสมในการจัดการเรียนรู้ระหว่างการเรียนแบบเสมือนและแบบพบหน้า เพื่อให้การเรียนรู้มีคุณภาพตามมาตรฐานที่กำหนด ซึ่งจากกระบวนการพัฒนาข้อความในแบบวัดฉบับนี้ที่ผ่านกระบวนการตรวจสอบคุณภาพของข้อความ 3 รอบเพื่อพัฒนาและปรับปรุงข้อความให้สมบูรณ์ จึงคาดว่าผลของความไม่สมดุลของการแจกแจงการตอบในด้านความคาดหวังด้านรูปแบบการเรียนรู้ที่มีการเลือกตอบ enriched-virtual model ต่ำกว่ารูปแบบอื่นๆ อย่างชัดเจนมีแนวโน้มว่าอาจมาจากข้อมูลตามสภาพจริงของกลุ่มตัวอย่าง ซึ่งต้องศึกษาเพิ่มเติมต่อไป

ผลการตรวจสอบคุณภาพด้านความตรงตั้งแต่ความตรงตามเนื้อหาที่ผ่านกระบวนการตรวจสอบคุณภาพตามเกณฑ์ที่กำหนด นอกจากนี้จากค่าสหสัมพันธ์ระหว่างส่วนหนึ่งของแบบวัดกับแบบวัดอื่นที่ผ่านกระบวนการตรวจสอบคุณภาพพบสหสัมพันธ์อย่างมีนัยสำคัญทางสถิติระดับ 0.05 โดยเป็นความสัมพันธ์ระดับกลาง ($r=0.33$) และต่ำ ($r=0.28$) (Cohen, 1988) ซึ่งเป็นหลักฐานที่ยืนยันถึงคุณภาพด้านความตรงตามเกณฑ์สัมพัทธ์ของแบบวัดที่พัฒนาขึ้น อีกทั้งการตรวจสอบความตรงตามทฤษฎีโดยใช้วิธีการวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี พบว่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากการวัดลักษณะต่างกันด้วยวิธีเดียวกัน (HTMM) อยู่ในระดับกลางและต่ำ ($0.28 \leq r \leq 0.48$) (Cohen, 1988) โดยเป็นความสัมพันธ์ที่มีนัยสำคัญทางสถิติที่ระดับ 0.05 และสูงกว่าสัมประสิทธิ์สหสัมพันธ์อื่นจึงใช้เป็นหลักฐานสนับสนุนความตรงแบบลู่เข้า (convergent validity) จากผลการตรวจสอบคุณภาพด้านความตรงทั้ง 3 วิธี จึงสามารถใช้เป็นหลักฐานยืนยันถึงคุณภาพด้านความตรงที่เพียงพอต่อการแบบวัดไปใช้ในสถานการณ์จริงได้ นอกจากนี้ผลการตรวจสอบคุณภาพด้านความเที่ยงซึ่งตรวจสอบโดยใช้วิธีการวัดความสอดคล้องภายในโดยใช้ Cronbach's alpha พบว่าความเที่ยงในภาพรวม ($\alpha=0.57$) อยู่ในระดับที่ยอมรับได้ (George & Mallery, 2003) และตรวจสอบคุณภาพ

ด้านความเที่ยงโดยวิธีการทดสอบซ้ำ (test-retest reliability) ในระยะเวลาห่างกัน 2 สัปดาห์ พบว่าค่าสัมประสิทธิ์สหสัมพันธ์เฉลี่ยทั้งฉบับ ($r=0.77$) อยู่ในระดับสูง (Cohen, 1988)

จากกระบวนการตรวจสอบคุณภาพเบื้องต้นหลายขั้นตอนจากกลุ่มตัวอย่างที่หลากหลายเป็นหนึ่งปัจจัยที่ทำให้แบบวัดที่พัฒนาขึ้นนั้นมีคุณภาพเพียงพอเมื่อนำไปใช้ในสถานการณ์จริง (Schünemann & Ziegler, 2023) ประกอบกับการใช้บริบทสถานการณ์มาช่วย สามารถลดผลกระทบจากการได้รับผลการตอบที่ขึ้นอยู่กับความสามารถในการสะท้อนตนเองของผู้ตอบ (Hooper et al., 2006) และเพิ่มความตรงในการทำนายผลลัพธ์ได้ดีขึ้น (Swift & Peterson, 2019) อย่างไรก็ตาม ถึงแม้ว่าผลการตรวจสอบความตรงและความเที่ยงของแบบวัดฉบับนี้อยู่ในระดับที่ยอมรับได้ แต่พบว่าเป็นค่าคุณภาพที่อยู่ในระดับไม่สูงนัก เช่น ความสัมพันธ์ที่ใช้ตรวจสอบคุณภาพด้านความตรงตามเกณฑ์สัมพันธ์อยู่เพียงในระดับปานกลางและต่ำ รวมถึงการตรวจสอบคุณภาพด้านความเที่ยงด้วยวิธีวัดความสอดคล้องภายในอยู่ในระดับที่ยอมรับได้แต่มีความสอดคล้องอยู่ในระดับต่ำ (George & Mallery, 2003) อาจมีสาเหตุมาจากหลายสาเหตุ เช่น ปัญหากลุ่มตัวอย่างมีความแปรปรวนต่ำเกินไป (Restricted range problem) จึงทำให้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรต่ำกว่าที่ควรจะเป็น (Sackett & Yang, 2000) โดยกลุ่มตัวอย่างในงานวิจัยนี้เป็นนักศึกษาระดับปริญญาตรีในหนึ่งมหาวิทยาลัยซึ่งอาจมีอายุ ภูมิหลังทางวัฒนธรรม และประสบการณ์ในการเรียนรู้ในรูปแบบต่าง ๆ ที่ใกล้เคียงกัน จึงทำให้ผลการตอบแบบวัดที่ใช้ในการศึกษาครั้งนี้ไม่แตกต่างกันนักจึงทำให้มีความแปรปรวนต่ำในที่สุด การศึกษาในครั้งต่อไปอาจเลือกกลุ่มตัวอย่างที่หลากหลายมากขึ้นจากหลากหลายบริบท เช่น นักศึกษาจากหลายระดับ หลายมหาวิทยาลัย รวมถึงประเภทของสถาบันการศึกษาที่หลากหลาย เป็นต้น หรืออาจใช้กระบวนการทางสถิติ range restriction correction (Hunter & Schmidt, 2004) ในการปรับค่าสัมประสิทธิ์สหสัมพันธ์ให้สะท้อนกับความเป็นจริงมากขึ้น

นอกจากนี้อาจเป็นผลมาจากความเหนื่อยล้าจากการทำแบบวัด โดยกระบวนการตอบแบบวัดในรูปแบบบังคับเลือกกระตุ้นให้ผู้ตอบเกิดกระบวนการสะท้อนตัวเอง (Schünemann & Ziegler, 2023) เพื่อเปรียบเทียบทุกคู่ของข้อความที่น่าเสนอในแต่ละข้อ จึงเป็นสาเหตุที่ทำให้เพิ่มการใช้ภาระทางปัญญา (cognitive load) ซึ่งเป็นความสามารถที่มีขีดจำกัดของมนุษย์ และอาจเกิดการใช้ภาระทางปัญญาเกินกำหนด (cognitive overload) ได้ หากการตอบแบบวัดทำให้ผู้ตอบต้องประมวลผลข้อมูลมากเกินไป (Sweller, 1988) และส่งผลต่อเนื่องไปยังความเหนื่อยล้าจากการทำแบบวัด (test fatigue) (Lee et al., 2021) โดย Satisficing Theory (Krosnick, 1987) เสนอว่าโดยทั่วไปแล้ว ผู้ตอบแบบวัดใช้ 2 วิธีในการตอบ ได้แก่ การตอบแบบตั้งใจ (optimizing) ที่ผู้ตอบอ่านและทำความเข้าใจทุกตัวเลือกก่อนเลือกตอบ หรือการตอบแบบลดความพยายาม (satisficing) ที่ผู้ตอบเลือกโดยไม่พยายามคิดอย่างลึกซึ้ง จากแบบวัดฉบับนี้ซึ่งมีการกำหนด 5 สถานการณ์ อาจเป็นสาเหตุที่ทำให้ผู้ตอบใช้ภาระทางปัญญาที่เกิน จึงอาจเป็นสาเหตุที่ทำให้ระดับคุณภาพของแบบวัดลดลง การศึกษาในครั้งต่อไปอาจพิจารณาลดจำนวนสถานการณ์หรือใช้วิธีการต่าง ๆ ที่ช่วยให้ผู้ตอบแบบวัดใช้ภาระทางปัญญาไม่มากและน้อยเกินไป หรืออาจใช้เทคนิคต่าง ๆ นอกเหนือจากการใช้สถานการณ์และภาพประกอบสถานการณ์ในการช่วยลดภาระทางปัญญาของผู้ตอบให้ได้เหมาะสมเพื่อให้แบบวัดในรูปแบบบังคับเลือกสามารถนำไปใช้ได้อย่างมีประสิทธิภาพสูงสุด

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

จากผลการตรวจสอบคุณภาพความตรงและความเที่ยงของแบบวัดฉบับนี้ซึ่งอยู่ในระดับที่ยอมรับได้ แต่พบว่าเป็นระดับคุณภาพที่อยู่ในระดับไม่สูงนัก ซึ่งอาจเป็นผลมาจากแบบวัดที่พัฒนาขึ้นยังไม่สามารถลดการใช้ภาระทางปัญญาของผู้ตอบได้อย่างชัดเจนจึงกระตุ้นให้ผู้ตอบเกิดความเหนื่อยล้าจากการทำแบบวัดในรูปแบบบังคับเลือก การนำแบบวัดฉบับนี้ไปใช้ในสถานการณ์จริงจึงควรจัดสภาพแวดล้อมให้เหมาะสมที่สุด เช่น ลดเสียงรบกวน ลดแรงกดดันด้านเวลา เพื่อป้องกันปัญหาด้านสภาพแวดล้อม (environmental distractions) ที่กระตุ้นให้เกิดความเหนื่อยล้าในการตอบมากขึ้นและส่งผลกระทบต่อคุณภาพของผลการตอบในที่สุด รวมทั้งการแจ้งวัตถุประสงค์และความสำคัญของการวัดในครั้งนี้เพื่อกระตุ้นแรงจูงใจในการตอบ

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 การศึกษาความยาวที่เหมาะสมของแบบวัดในรูปแบบบังคับที่ทำให้การใช้ภาระทางปัญญาของผู้ตอบแบบวัดอยู่ในระดับที่เหมาะสมไม่เกิดการใช้ภาระทางปัญญาเกิน (cognitive overload) และมีรูปแบบการตอบแบบตั้งใจและมีประสิทธิภาพสูงสุดตลอดการทำแบบวัด

2.2 จากการตรวจสอบคุณภาพของแบบวัดที่อยู่ในระดับที่ยอมรับได้แต่เป็นระดับคุณภาพที่ไม่สูงมากนัก การศึกษาในครั้งต่อไปจึงควรศึกษาและทดสอบประสิทธิภาพของแบบวัดรูปแบบบังคับเลือกที่ใช้เทคนิคต่างๆในการช่วยลดภาระทางปัญญาของผู้ตอบให้มากที่สุด นอกเหนือจากการใช้สถานการณ์และภาพประกอบสถานการณ์ เช่น การใช้การเปรียบเทียบข้อความทีละคู่ (paired comparison), แบบวัดบังคับเลือกแบบปรับเหมาะ (adaptive forced-choice test) เป็นต้น เพื่อให้ผลการตอบแบบวัดในรูปแบบบังคับเลือกมีประสิทธิภาพสูงสุด

References

- Bonk, C. J., Graham, C. R., Cross, J., & Moore, M. G. (2005). *The handbook of blended learning: Global perspectives, local designs*. Pfeiffer.
- Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement*, 71(3), 460–502.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Allyn & Bacon.

- Hontangas, P. M., de la Torre, J., Ponsoda, V., Leenen, I., Morillo, D., & Abad, F. J. (2015). Comparing traditional and IRT scoring of forced-choice tests. *Applied Psychological Measurement, 39*(8), 598–612.
- Hooper, A. C., Cullen, M. J., & Sackett, P. R. (2006). Operational threats to the use of SJTs: Faking, coaching, and retesting issues. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 205–232). Lawrence Erlbaum Associates.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). SAGE.
- Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly, 51*(2), 201–219.
- Lakho, S., Jalbani, D., Memon, I., Siraj, S., & Ali, A. (2023). Blended enriched virtual model for the prediction of students' performance using probabilistic-based model. In *Proceedings of the 13th International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 138–144). Springer.
- Lee, P., Joo, S. H., & Stark, S. (2021). Detecting DIF in multidimensional forced-choice measures using the Thurstonian item response theory model. *Organizational Research Methods, 24*(4), 739–771.
- Lee, P., Joo, S., & Son, M. (2024). Detecting careless respondents in multidimensional forced choice data: An application of lz person-fit statistic to the TIRT model. *Journal of Business and Psychology, 39*, 541–564.
- Muhlisoh, E. D., Santihastuti, A., & Wahjuningsih, E. (2020). Students' perceptions of flipped approach in EFL classroom: A survey research. *Journal of Education, Teaching, and Learning, 5*(2), 393–400.
- Peterson, E., & Irving, S. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction, 18*(3), 238–250.
- Prayitno, E. H., Masunah, J., & Milyartini, R. (2023). Implementation of enriched virtual learning model through online guidance to create film production in videography class. In *Proceedings of the Fifth International Conference on Arts and Design Education (ICADE 2022)* (pp. 531–540). Atlantis Press.
- Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology, 85*(1), 112–118.

- Salleh, F. I. M., Baharum, H. I., & Shamsudin, S. (2017). Comparative study between flipped learning and flex model in English as a second language classroom. *Advanced Science Letters*, 23(4), 2663–2666.
- Sass, R., Frick, S., Reips, U.-D., & Wetzel, E. (2020). Taking the test taker's perspective: Response process and test motivation in multidimensional forced-choice versus rating scale instruments. *Assessment*, 27(3), 572–584.
- Schünemann, A. L., & Ziegler, M. (2023). Use the force! Adaptation of response formats: From rating scale to multidimensional forced choice. *Psychological Test Adaptation and Development*, 4(1), 218–234.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285.
- Swift, V., & Peterson, J. B. (2019). Contextualization as a means to improve the predictive validity of personality models. *Personality and Individual Differences*, 144, 153–163.
- Ward, M. K., & Meade, A. W. (2023). Dealing with careless responding in survey data: Prevention, identification, and recommended best practices. *Annual Review of Psychology*, 74(1), 1–20.

Translate Thai References

- Kanjanawasee, S. (2013). *Classical test theory*. Chulalongkorn University. (in Thai)
- Yimyam, S., Jaruwatcharapanichkul, A., Charoensanti, C., Intharangkura Na Ayudhya, A., Chuto, P., & Chaloemsuk, N. (2015). The development of blended learning management to enhance 21st-century learning skills. *Nursing Journal*, 42(Special Issue), 129–140. (in Thai)