

## การเปรียบเทียบคุณภาพเมทริกซ์คิวของแบบสอบวินิจฉัยทางพุทธิปัญญาระหว่าง วิธีการใช้การตัดสินโดยผู้เชี่ยวชาญกับวิธีการใช้ดัชนีพีวีเอเอฟ

### Comparison of Q-Matrix Validations for a Cognitive Diagnostic Test between Expert Judgement Method and PVAF Index Method

วรุทม์ ผิวงาม<sup>1\*</sup> และ ณภัทร ชัยมงคล<sup>2</sup>

Warute Phiwngam<sup>1\*</sup> and Nhabhat Chaimongkol<sup>2</sup>

(Received: April 27, 2021 ; Revised: May 29, 2021 ; Accepted: June 1, 2021)

#### บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อเปรียบเทียบคุณภาพของเมทริกซ์คิวของแบบสอบวินิจฉัยทางพุทธิปัญญาระหว่างวิธีการใช้การตัดสินโดยผู้เชี่ยวชาญกับวิธีการใช้ดัชนีพีวีเอเอฟ (PVAF) โดยแบบสอบวินิจฉัยทางพุทธิปัญญามีทั้งหมด 1 ฉบับ ซึ่งเป็นแบบสอบแบบเลือกตอบ 4 ตัวเลือก ทั้งหมด 40 ข้อ แต่ละข้อมุ่งจำแนกคุณลักษณะจำนวน 1 ถึง 3 คุณลักษณะ และมีตัวอย่างในการวิจัย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 4 การศึกษา 2563 ในกรุงเทพมหานคร จำนวน 225 คน โดยใช้โมเดลจีไดนา (G-DINA Model) เป็นโมเดลการวินิจฉัยทางพุทธิปัญญา และใช้โปรแกรมอาร์ในการวิเคราะห์ผลสำหรับการตรวจสอบคุณภาพเมทริกซ์คิวด้วยวิธีการใช้ดัชนี PVAF

ผลการวิจัยพบว่า เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพด้วยการตัดสินโดยผู้เชี่ยวชาญสามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้นมากกว่าวิธีการใช้ดัชนี PVAF โดยเมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยใช้วิธีการตัดสินโดยผู้เชี่ยวชาญสามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้นทั้งหมด ส่วนเมทริกซ์คิวที่ผ่านการตรวจสอบตรวจสอบคุณภาพโดยใช้ดัชนี PVAF มีข้อสอบ 22 ข้อ (ร้อยละ 55.0) ที่สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้น แต่มีข้อสอบจำนวน 9 ข้อ (ร้อยละ 22.5) ที่จำแนกคุณลักษณะคลาดเคลื่อนไปจากเมทริกซ์คิวที่สร้างขึ้น และมีข้อสอบจำนวน 9 ข้อ (ร้อยละ 22.5) ที่ไม่สามารถจำแนกคุณลักษณะใด ๆ ได้

**คำสำคัญ:** แบบสอบวินิจฉัยทางพุทธิปัญญา เมทริกซ์คิว การตรวจสอบคุณภาพเมทริกซ์คิว ดัชนีพีวีเอเอฟ  
โมเดลจีไดนา

<sup>1</sup> นิสิตระดับปริญญาโท สาขาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

<sup>2</sup> อาจารย์ประจำภาควิชาวิจัยและจิตวิทยาทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

<sup>1</sup> Master's degree student, Educational Measurement and Evaluation Program, Faculty of Education, Chulalongkorn University

<sup>2</sup> Lecturer, Department of Educational Research and Psychology, Faculty of Education, Chulalongkorn University

\* Corresponding Author E-mail: Warute.p@chula.ac.th

## Abstract

The purpose of this research was to compare the quality of two Q-matrix validation methods which were the expert judgement method and the PVAF index method, for a cognitive diagnostic test. The cognitive diagnostic test in this research was a 4-choice test which contained 40 questions. Each item aimed to measure one to three attributes. The sample of this research consisted of 225 grade 10 students in Bangkok, in the academic year 2020. The result of this research was obtained by using G-DINA Model as a cognitive diagnostic model. The R Program was applied to analyze the quality of Q-matrix with PVAF index method.

The result of this research showed that the Q-matrix validated by the expert judgement method could measure the students' attributes more accurately than the PVAF index method, based on the developed Q-matrix, as the expert judgement method could specify Q-matrix without any misspecifications. Meanwhile, the PVAF index method showed only 22 items (55.0%) from 40 items that were correctly specified, compared to the developed Q-matrix, and 9 items (22.5%) were incorrectly specified, compared to the developed Q-matrix, and there were 9 items (22.5%) that attributes could not be specified.

**Keywords:** cognitive diagnostic test, Q-matrix, Q-matrix validation, PVAF Index, G-DINA Model

## บทนำ

ปัจจุบันการศึกษาในประเทศต่าง ๆ ให้ความสำคัญกับการจัดการศึกษาฐานสมรรถนะที่เน้นผู้เรียนเป็นศูนย์กลาง กระตุ้นให้ผู้เรียนได้เรียนรู้ตามสิ่งที่ตนเองถนัด หรือตามความสนใจของตนเอง โดยยึดคุณลักษณะสำคัญในการใช้ชีวิต ผสมเข้ากับทักษะพื้นฐานที่พบในหลักสูตรการศึกษาเดิม ในส่วนของการวัดและประเมินผลในการจัดการศึกษาฐานสมรรถนะจะเน้นการให้ข้อมูลย้อนกลับแก่ผู้เรียนในรูปแบบคุณลักษณะ (attributes) หรือทักษะ (skills) ที่ต้องการมุ่งวัดว่าผู้เรียนผ่านตามเกณฑ์ที่พึงประสงค์หรือไม่ เพื่อนำไปสู่การพัฒนาและปรับปรุงการเรียนรู้ของผู้เรียน (Office of the Education Council, 2019)

วิธีการที่ใช้ในการตรวจสอบคุณลักษณะหรือทักษะของผู้เรียนมีหลายวิธี วิธีการหนึ่งที่สามารถทำได้คือการใช้แบบสอบวินิจฉัยทางพุทธิปัญญา ซึ่งเป็นแบบสอบที่ช่วยให้ข้อมูลย้อนกลับแก่ผู้เรียน โดยช่วยให้ผู้เรียนทราบถึงระดับความรอบรู้ในคุณลักษณะของผู้เรียนว่าคุณลักษณะใดเป็นจุดเด่นและคุณลักษณะใดเป็นข้อจำกัดของตน เพื่อนำไปสู่การปรับปรุงและพัฒนาคุณลักษณะดังกล่าวให้ดียิ่งขึ้น (Jang & Wagner, 2013; Ma & de la Torre, 2020a) นอกจากนั้นแบบสอบวินิจฉัยทางพุทธิปัญญายังสามารถช่วยให้ผู้สอนสามารถวางแผนการสอนให้เหมาะสมกับผู้เรียน เพื่อให้ผู้เรียนได้จัดการกับความรู้อันคลาดเคลื่อน และพัฒนาคุณลักษณะและสมรรถนะที่สำคัญให้บรรลุตามจุดประสงค์ที่มุ่งวัดได้

แบบสอบวินิจฉัยทางพุทธิปัญญาที่มักพบในงานวิจัยเป็นแบบสอบแบบเลือกตอบพหุระดับ (multi-tier diagnostic test) เช่น แบบสอบวินิจฉัยแบบเลือกตอบ 2 ระดับ แบบสอบวินิจฉัยแบบเลือกตอบ 3 ระดับ อย่างไรก็ตาม แบบสอบวินิจฉัยทั้ง 2 รูปแบบยังคงข้อจำกัด ส่งผลให้ไม่สามารถจำแนกผู้สอบที่รอบรู้กับผู้สอบที่เดาคำตอบได้ถูกต้อง อีกทั้งยังใช้เวลาในการทำข้อสอบนาน เนื่องจากข้อสอบ 1 ข้อ มีคำถามย่อยมากกว่า 1 คำถาม อาจส่งผลให้เกิดความเหนื่อยล้าในการทำแบบสอบ (Waugh & Gronlund, 2013) แต่ในปัจจุบันที่เทคโนโลยีทางการศึกษามีการพัฒนาไปอย่างมาก จึงมีผู้ศึกษาทฤษฎีโมเดลทางสถิติที่เรียกว่า โมเดลการวินิจฉัยทางพุทธิปัญญาเข้ามาประยุกต์ใช้กับแบบสอบวินิจฉัยทางพุทธิปัญญา (Meesakul et al., 2015) ที่ช่วยให้การประเมินความรู้ในคุณลักษณะหรือสมรรถนะของผู้เรียนมีความถูกต้องและแม่นยำ (Chaimongkol, 2017) โดยโมเดลการวินิจฉัยทางพุทธิปัญญาจะระบุความสัมพันธ์ระหว่างทักษะทางจิตวิทยาของผู้สอบกับคุณลักษณะที่แบบสอบวินิจฉัยทางพุทธิปัญญามุ่งวัดได้อย่างถูกต้อง แม่นยำ และมีความน่าเชื่อถือทางสถิติ (Nájera et al., 2019) รวมทั้งเป็นโมเดลที่สามารถใช้กับแบบสอบแบบเลือกตอบอย่างง่าย (simple multiple-choice test) หรือแบบสอบที่ 1 ข้อมีเพียงคำถามเพียงข้อเดียว ส่งผลให้ผู้สอบไม่เกิดความเหนื่อยล้าในการทำแบบสอบ อีกทั้งยังมีความเป็นปรนัย ซึ่งนำไปสู่การตรวจสอบความรู้ในคุณลักษณะหรือสมรรถนะของผู้สอบ และการพัฒนาและปรับปรุงการสอนของผู้สอนให้ตรงตามจุดประสงค์ที่ต้องการมุ่งวัดได้

ประเภทของโมเดลการวินิจฉัยทางพุทธิปัญญาที่พบในปัจจุบันมีหลากหลาย เช่น โมเดลจีไดนา (G-DINA), ไดนา (DINA), ไดโน (DINO), ไนดา (NIDA), addictive CDM, loglinear CDM แต่เนื่องจากข้อจำกัดและความซับซ้อนของโมเดล ทำให้นำมาใช้ในทางปฏิบัติได้ยาก ส่งผลให้ในปัจจุบันมีเพียงโมเดลจีไดนา และไดนาที่นิยมนำมาใช้จำแนกคุณลักษณะของผู้สอบในแบบสอบวินิจฉัยทางพุทธิปัญญา (Nájera et al., 2019) เนื่องจากสะดวกในทางปฏิบัติและยังคงสามารถวินิจฉัยคุณลักษณะของผู้สอบได้อย่างถูกต้องและแม่นยำ และจากงานวิจัยของ Meesakul et al. (2015) ที่ทำการเปรียบเทียบประสิทธิภาพในการจำแนกคุณลักษณะระหว่างโมเดลจีไดนากับโมเดลไดนาด้วยค่าดัชนีสอดคล้องของโมเดลเชิงสัมบูรณ์และสัมพัทธ์ พบว่าโมเดลจีไดนามีประสิทธิภาพในการจำแนกคุณลักษณะมากกว่าโมเดลไดนา เนื่องจากโมเดลจีไดนาเป็นโมเดลแบบชดเชย (compensatory model) ที่ยืดหยุ่นกว่าโมเดลไดนา หรืออธิบายได้ว่าเป็นโมเดลที่ยิงผู้สอบที่มีจำนวนคุณลักษณะมากเท่าใด โอกาสในการตอบข้อสอบได้ถูกต้องจะมากกว่าผู้สอบที่มีจำนวนคุณลักษณะน้อยกว่าหรือไม่มีคุณลักษณะใด ๆ เลย ซึ่งทำให้โมเดลจีไดนามีความสอดคล้องกับสภาพความเป็นจริงที่ผู้สอบแต่ละคนมีความสามารถที่หลากหลายมากกว่าโมเดลไดนานั้นเอง ส่วนสิ่งสำคัญอีกประการสำหรับการจำแนกคุณลักษณะของโมเดลการวินิจฉัยทางพุทธิปัญญา คือ เมทริกซ์คิว (Q-matrix) ซึ่งเป็นเมทริกซ์ที่แสดงความสัมพันธ์ระหว่างข้อสอบแต่ละข้อกับคุณลักษณะที่มุ่งวัด เมทริกซ์คิวที่มีคุณภาพจึงเป็นเมทริกซ์ที่ข้อสอบแต่ละข้อสามารถจำแนกคุณลักษณะของผู้สอบได้ตรงกับคุณลักษณะที่ต้องการมุ่งวัด โดยงานวิจัยที่ผ่านมานิยมตรวจสอบคุณภาพของแบบสอบวินิจฉัยด้วยวิธีการตัดสินโดยผู้เชี่ยวชาญพิจารณาความสอดคล้องระหว่างข้อสอบกับคุณลักษณะของเมทริกซ์คิว (Nájera et al., 2019) แต่เนื่องจากวิธีดังกล่าวเป็นวิธีที่ค่อนข้างนามธรรม การตัดสินความสอดคล้องและแม่นยำของเมทริกซ์คิวมาจากบุคคลภายนอก ไม่ใช่ผู้สอบโดยตรง จึงอาจเกิดปัญหาความแตกต่างของระดับคุณลักษณะระหว่างผู้เชี่ยวชาญกับตัวผู้สอบจริง (Johnson et al.,

2009) ส่งผลให้เกิดความคลาดเคลื่อนในการจำแนกคุณลักษณะของผู้สอบได้ (de la Torre & Chiu, 2016; Ma & de la Torre, 2020a) ซึ่ง de la Torre (2011) และ Nájera et al. (2019) ได้พัฒนาวิธีการตรวจสอบคุณภาพเมตริกซ์ควิตด้วยดัชนีพีวีเอเอฟ (Proportion of Variance Accounted For, PVAF) ซึ่งเป็นค่าที่แสดงความสามารถจำแนกคุณลักษณะที่มุ่งวัดว่าข้อสอบแต่ละข้อสามารถจำแนกคุณลักษณะใด โดยใช้ค่าคะแนนจุดตัด Epsilon ( $\epsilon$ , EPS) เป็นเกณฑ์ในการพิจารณาข้อสอบแต่ละข้อว่าสามารถวัดคุณลักษณะใดได้บ้าง ซึ่งข้อสอบที่สามารถวัดคุณลักษณะได้อย่างมีคุณภาพจะต้องมีค่าดัชนี PVAF สูงกว่าค่า EPS ส่งผลให้การตรวจสอบคุณภาพเมตริกซ์ควิตด้วยดัชนี PVAF มีหลักฐานทางสถิติและเป็นหลักฐานเชิงประจักษ์มากกว่าวิธีการใช้ผู้เชี่ยวชาญ

ด้วยเหตุผลที่กล่าวมาข้างต้น ผู้วิจัยจึงสนใจที่จะเปรียบเทียบคุณภาพในการจำแนกคุณลักษณะของเมตริกซ์ควิตของแบบสอบวินิจฉัยทางพุทธิปัญญาระหว่างวิธีการตัดสินโดยผู้เชี่ยวชาญกับวิธีการใช้ดัชนี PVAF โดยเนื้อหาที่นำมาใช้เพื่อเปรียบเทียบคุณภาพเมตริกซ์ควิตครั้งนี้ คือเรื่องพันธะไอออนิก ในรายวิชาเคมี ระดับชั้นมัธยมศึกษาปีที่ 4 ที่ครอบคลุมคุณลักษณะทั้งหมด 3 คุณลักษณะ ได้แก่ 1) การอธิบายการเกิดพันธะไอออนิก 2) การเขียนสูตรและเรียกชื่อสารประกอบไอออนิก และ 3) การระบุปฏิริยาระหว่างสารประกอบไอออนิก เพื่อเป็นแนวทางในการประยุกต์ใช้ค่าดัชนี PVAF ในการตรวจสอบคุณภาพเมตริกซ์ควิตของแบบสอบวินิจฉัยทางพุทธิปัญญา และนำไปสู่การใช้ค่าดัชนี PVAF ในการสนับสนุนผลการตรวจสอบคุณภาพเมตริกซ์ควิตร่วมกับวิธีการใช้การตัดสินใจโดยผู้เชี่ยวชาญตามที่ Ma & de la Torre (2020a) ได้เสนอไว้

### ความมุ่งหมายของการวิจัย

เพื่อเปรียบเทียบคุณภาพของเมตริกซ์ควิตของแบบสอบวินิจฉัยทางพุทธิปัญญาระหว่างวิธีการใช้การตัดสินใจโดยผู้เชี่ยวชาญกับวิธีการใช้ดัชนี PVAF

### วิธีดำเนินการวิจัย

#### 1. ประชากรและกลุ่มตัวอย่าง

ประชากรในการวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 ปีการศึกษา 2563 ในกรุงเทพมหานคร จำนวน 60,770 คน

ตัวอย่างในการวิจัยครั้งนี้ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 4 ปีการศึกษา 2563 ในกรุงเทพมหานคร จำนวน 225 คน ซึ่งมากกว่าจำนวนตัวอย่างวิจัยขั้นต่ำ 200 คน (Nájera et al., 2019) ผ่านการเลือกแบบเจาะจง เนื่องจากตัวอย่างในการวิจัยมีความพร้อมในการให้ความร่วมมือในการเก็บข้อมูลในสถานการณ์การแพร่ระบาดของเชื้อไวรัสโควิด-19 และเพื่อความหลากหลายของรูปแบบคุณลักษณะในการเรียนรู้ ในรายวิชาเคมี ดังตาราง 1

ตาราง 1 ตัวอย่างในการวิจัย

โรงเรียน	ตัวอย่างในการวิจัย (คน)
1. โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม	57
2. โรงเรียนบางมดวิทยา	44
3. โรงเรียนราชวินิต มัธยม	30
4. โรงเรียนทวีธาภิเศก	25
5. โรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ ปทุมวัน	20
6. โรงเรียนชิโนรสวิทยาลัย	18
7. โรงเรียนพรตพิทยพยัต	18
8. โรงเรียนบดินทรเดชา (สิงห์ สิงหเสนี)	13
รวม	225

## 2. เครื่องมือที่ใช้ในการวิจัย

แบบสอบวินิจฉัยทางพุทธิปัญญาเรื่องพันธะไอออนิกแบบเลือกตอบอย่างง่าย (simple multiple-choice test) จำนวน 1 ฉบับ มีข้อสอบทั้งหมด 40 ข้อ กำหนดระยะเวลาในการทำข้อสอบ 50 นาที ข้อสอบแต่ละข้อแบ่งออกเป็น 4 ตัวเลือก ได้แก่ ตัวถูก จำนวน 1 ตัวเลือก พัฒนามาจากมโนทัศน์ที่ถูกต้องเรื่องพันธะไอออนิก และตัวลวง จำนวน 3 ตัวเลือกที่พัฒนามาจากมโนทัศน์ที่คลาดเคลื่อนที่พบในเนื้อหาเรื่องพันธะไอออนิก ดังภาพ 1 แต่เนื่องจากสถานการณ์การแพร่ระบาดของเชื้อไวรัสโควิด-19 ส่งผลให้เกิดความยากลำบากในการเก็บข้อมูลในรูปแบบข้อสอบแบบกระดาษ ผู้วิจัยจึงจำเป็นต้องเก็บข้อมูลในรูปแบบออนไลน์เพิ่มเติมเพื่อให้จำนวนตัวอย่างในการวิจัยเพียงพอต่อจำนวนขั้นต่ำ โดยแบบสอบที่พัฒนาขึ้นมุ่งจำแนกคุณลักษณะทั้งหมด 3 คุณลักษณะ ได้แก่ 1) การอธิบายการเกิดพันธะไอออนิก 2) การเขียนสูตรและเรียกชื่อสารประกอบไอออนิก และ 3) การระบุปฏิกิริยาระหว่างสารประกอบไอออนิก โดยแต่ละข้อจำแนกคุณลักษณะจำนวน 1 ถึง 3 คุณลักษณะ ดังตาราง 2 ซึ่งแสดงเมทริกซ์คิวของแบบสอบวินิจฉัยทางพุทธิปัญญาเรื่องพันธะไอออนิก ค่า 1 หมายถึง ข้อสอบต้องการจำแนกความรู้ในคุณลักษณะดังกล่าวในการตอบข้อนั้น ส่วนค่า 0 ในเมทริกซ์คิว หมายความว่า ข้อสอบไม่ต้องการจำแนกความรู้ในคุณลักษณะดังกล่าวในการตอบข้อนั้น ในเมทริกซ์คิวจะประกอบไปด้วยเวกเตอร์คิว (q-vector) ซึ่งเป็นเวกเตอร์ของข้อสอบแต่ละข้อที่ระบุคุณลักษณะที่มุ่งจำแนก เช่น เวกเตอร์คิวของข้อสอบข้อที่ 17 คือ {010} เป็นข้อสอบที่มุ่งจำแนก 1 คุณลักษณะ ได้แก่คุณลักษณะที่ 2 ส่วนเวกเตอร์คิวของข้อสอบข้อที่ 32 คือ {101} มุ่งจำแนกทั้งหมด 2 คุณลักษณะ ได้แก่ คุณลักษณะที่ 1 และ 3 เป็นต้น

17. ข้อใดจับคู่สูตรเคมีกับชื่อสารประกอบไอออนิกได้ถูกต้อง

1. CsClO - ซีเซียมคลอไรด์
2. Fe<sub>2</sub>O<sub>3</sub> - ไอร์ออนออกไซด์
3. BiF<sub>3</sub> - บิสมัท (III) ฟลูออไรด์
4. Co<sub>2</sub>O<sub>3</sub> - โคบอลต์ไดรอกไซด์

32. ข้อใดถูกต้องเมื่อผสมสารละลาย MnSO<sub>4</sub> เข้ากับสารละลายต่อไปนี้

	สารละลาย	ตะกอนที่เกิดขึ้น	สมบัติของตะกอน
1.	Sr(NO <sub>3</sub> ) <sub>2</sub>	SrSO <sub>4</sub>	ยึดเหนี่ยวกันด้วยพันธะระหว่าง Sr <sup>2+</sup> กับ SO <sub>4</sub> <sup>2-</sup>
2.	Na <sub>2</sub> S	MnS	เป็นโมเลกุลสารประกอบ
3.	LiCl	Li <sub>2</sub> SO <sub>4</sub>	เป็นสารประกอบโคเวเลนต์
4.	Rb <sub>2</sub> O	Rb <sub>2</sub> SO <sub>4</sub>	มีโครงสร้างเป็นเส้นตรง

ภาพ 1 ตัวอย่างเครื่องมือที่ใช้ในการวิจัย

ตาราง 2 เมทริกซ์ควิซสำหรับการวินิจฉัยของแบบสอบวินิจฉัยทางพุทธิปัญญา

ข้อที่	คุณลักษณะ			ข้อที่	คุณลักษณะ			ข้อที่	คุณลักษณะ			ข้อที่	คุณลักษณะ						
	1	2	3		1	2	3		1	2	3		1	2	3				
1	1	0	0	9	0	0	1	17	0	1	0	25	1	1	0	33	0	1	1
2	0	1	0	10	1	0	0	18	0	0	1	26	1	0	1	34	1	1	0
3	0	0	1	11	0	1	0	19	1	0	0	27	0	1	1	35	1	0	1
4	1	0	0	12	0	0	1	20	0	1	0	28	1	1	0	36	0	1	1
5	0	1	0	13	1	0	0	21	0	0	1	29	1	0	1	37	1	1	1
6	0	0	1	14	0	1	0	22	1	1	0	30	0	1	1	38	1	1	1
7	1	0	0	15	0	0	1	23	1	0	1	31	1	1	0	39	1	1	1
8	0	1	0	16	1	0	0	24	0	1	1	32	1	0	1	40	1	1	1

### 3. การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลแบ่งเป็นขั้นตอนดังต่อไปนี้

3.1 สร้างเมทริกซ์ควิซ โดยกำหนดให้ข้อสอบแต่ละข้อวัดคุณลักษณะในเรื่องพันธะไอออนิกจำนวน 1 – 3 คุณลักษณะ รวมทั้งหมด 40 ข้อ

3.2 สร้างแบบสอบแบบเลือกตอบ โดยข้อสอบแต่ละข้อมุ่งจำแนกคุณลักษณะของผู้สอบตามเมทริกซ์ควิซที่สร้างขึ้น จำนวน 40 ข้อ

3.3 นำแบบสอบที่สร้างขึ้นไปตรวจสอบคุณภาพด้วยวิธีการตัดสินโดยผู้เชี่ยวชาญ เพื่อหาความสอดคล้องระหว่างข้อสอบแต่ละข้อกับคุณลักษณะที่มุ่งจำแนก

3.4 นำแบบสอบที่ผ่านการตรวจสอบคุณภาพไปใช้กับตัวอย่างวิจัยจำนวน 225 คน แต่เนื่องจากความยากลำบากในการเก็บรวบรวมข้อมูลในสถานการณ์การแพร่ระบาดของเชื้อไวรัสโควิด-19 ผู้วิจัย

จึงจำเป็นต้องแบ่งการทดสอบเป็น 2 กลุ่ม ได้แก่ กลุ่มที่ใช้แบบสอบแบบกระดาษ จำนวน 189 คน และกลุ่มที่ใช้แบบสอบออนไลน์ จำนวน 36 คน

#### 4. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล มีรายละเอียดดังนี้

4.1 ตรวจสอบคุณภาพของเมทริกซ์ควิในการจำแนกคุณลักษณะด้วยวิธีการตัดสินโดยผู้เชี่ยวชาญ ด้วยค่าดัชนีความสอดคล้องระหว่างข้อสอบแต่ละข้อกับคุณลักษณะ (Item Objective Congruence, IOC) ผ่านผู้เชี่ยวชาญทั้งหมด 5 ท่าน โดยกำหนดค่าคะแนนตามที่ Kanjanawasee. (2013) กำหนดไว้ ดังนี้

- +1 เมื่อ แนใจว่าข้อสอบข้อนั้นมุ่งจำแนกคุณลักษณะได้ตามที่กำหนดไว้
- 0 เมื่อ ไม่แนใจว่าข้อสอบข้อนั้นมุ่งจำแนกคุณลักษณะได้ตามที่กำหนดไว้
- 1 เมื่อ แนใจว่าข้อสอบข้อนั้นไม่ได้มุ่งจำแนกคุณลักษณะตามที่กำหนดไว้

เมื่อนำคะแนนของข้อสอบแต่ละข้อที่ได้จากการตัดสินผู้เชี่ยวชาญ มาคำนวณและนำไปเปรียบเทียบกับเกณฑ์การพิจารณา ตามที่ Kanjanawasee (2013) ระบุไว้ ดังนี้

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

- เมื่อ  $\sum R$  คือ ผลรวมคะแนนการตัดสินของผู้เชี่ยวชาญ
- $N$  คือ จำนวนผู้เชี่ยวชาญทั้งหมด

#### เกณฑ์การพิจารณาค่า IOC

$IOC \geq 0.5$  แสดงว่า ข้อสอบข้อนั้นมุ่งจำแนกคุณลักษณะได้ตรงตามที่กำหนดไว้ในเมทริกซ์ควิ

$IOC < 0.5$  แสดงว่า ข้อสอบข้อนั้นไม่ได้มุ่งจำแนกคุณลักษณะตามที่กำหนดไว้ในเมทริกซ์ควิ

4.2 ตรวจสอบค่าดัชนีอำนาจจำแนก (item discrimination index) ของข้อสอบแต่ละข้อ ด้วยฟังก์ชัน *discrim* ใน GDINA package ด้วยโปรแกรม R (Ma & de la Torre, 2020) ซึ่งเป็นค่าที่คำนวณได้จากผลต่างระหว่างความน่าจะเป็นในการตอบถูกของผู้สอบที่มีความรอบรู้ในทุกคุณลักษณะที่ข้อสอบข้อดังกล่าวมุ่งวัด กับความน่าจะเป็นในการตอบถูกของผู้สอบที่ไม่มีความรอบรู้ในทุกคุณลักษณะที่ข้อสอบข้อดังกล่าวมุ่งวัด โดย Nájera et al. (2019) ระบุการแปลผลค่าดัชนีอำนาจจำแนกไว้ดังตาราง 3

ตาราง 3 การแปลผลค่าดัชนีอำนาจจำแนก

ค่าดัชนีอำนาจจำแนก	การแปลผล
0.800 เป็นต้นไป	อำนาจจำแนกสูง
0.600 ถึง 0.799	อำนาจจำแนกปานกลาง
0.400 ถึง 0.599	อำนาจจำแนกต่ำ แต่พอใช้ได้
ต่ำกว่า 0.400	อำนาจจำแนกใช้ไม่ได้ ควรตัดทิ้ง

4.3 ตรวจสอบจำแนกคุณลักษณะที่ต้องการมุ่งวัดของข้อสอบแต่ละข้อในแบบสอบด้วยวิธีการใช้ดัชนี PVAF ด้วยฟังก์ชัน  $Qval()$  ใน GDINA package ผ่านโปรแกรม R (Ma & de la Torre, 2020b) ซึ่งจะแสดงความสามารถจำแนกคุณลักษณะที่มุ่งวัดว่าข้อสอบแต่ละข้อสามารถจำแนกคุณลักษณะใดได้บ้าง พร้อมทั้งเสนอรูปแบบคุณลักษณะทั้งหมดที่เหมาะสมต่อการใช้จำแนกในข้อสอบแต่ละข้อในรูปของ Mesa plot ที่ในแกนแนวนอนจะแสดงเวกเตอร์คิวิที่แสดงคุณลักษณะที่ข้อสอบข้อดังกล่าวจำแนกได้ และแกนแนวตั้งจะแสดงค่าดัชนี PVAF ของแต่ละเวกเตอร์คิวิพร้อมทั้งระบุค่าดัชนี EPS ที่เป็นเส้นประแสดงค่าขั้นต่ำของดัชนี PVAF ดังตาราง 4 โดยเกณฑ์ในการพิจารณาว่าข้อสอบข้อดังกล่าวจำแนกคุณลักษณะใดได้ถูกต้องคือ ค่าดัชนี PVAF ต้องสูงกว่าค่าดัชนี EPS ซึ่งหมายถึงข้อสอบข้อดังกล่าวสามารถจำแนกคุณลักษณะได้อย่างแม่นยำ โดยค่า EPS เป็นค่าที่แปรผันตามค่าอำนาจจำแนก จำนวนข้อสอบ และจำนวนผู้สอบ Nájera et al. (2019) จึงได้เสนอค่า EPS ขั้นต่ำสำหรับตัวอย่างวิจัยจำนวน 200 คน ไว้ที่ 0.81 ส่วนข้อสอบที่มีเวกเตอร์คิวิที่มีค่า PVAF ผ่านเกณฑ์มากกว่า 1 ค่า เวกเตอร์คิวิที่มีจำนวนคุณลักษณะที่มุ่งจำแนกน้อยที่สุด จะเป็นเวกเตอร์ที่ถูกต้อง ส่วนเวกเตอร์คิวิมากกว่า 1 เวกเตอร์ที่มีค่า PVAF ผ่านเกณฑ์และมีจำนวนคุณลักษณะที่มุ่งจำแนกเท่ากัน เวกเตอร์ที่มีค่า PVAF สูงกว่าจัดเป็นเวกเตอร์ที่จำแนกได้ถูกต้องมากกว่า

4.4 เปรียบเทียบผลการตรวจสอบคุณภาพเมตริกซ์คิวิของแบบสอบวินิจฉัยทางพุทธิปัญญาระหว่างวิธีการตัดสินโดยผู้เชี่ยวชาญกับวิธีการใช้ดัชนี PVAF ว่าสามารถจำแนกคุณลักษณะได้ตรงกับเมตริกซ์คิวิที่สร้างขึ้นหรือไม่ โดยวิธีการตัดสินโดยผู้เชี่ยวชาญพิจารณาจากค่าเกณฑ์ในข้อ 4.1 ส่วนวิธีการตัดสินด้วยดัชนี PVAF พิจารณาจาก Mesa plot ตามเกณฑ์ในข้อ 4.3

## ผลการวิจัย

ผลการวิเคราะห์ข้อมูลเพื่อตอบความมุ่งหมายของการวิจัยมีรายละเอียด ดังนี้

ผลการตรวจสอบคุณภาพของแบบสอบวินิจฉัยทางพุทธิปัญญาเรื่องพันธะไอออนิกด้วยวิธีการตัดสินโดยผู้เชี่ยวชาญ จำนวน 5 ท่าน ด้วยค่าดัชนี IOC พบว่า ข้อสอบแต่ละข้อในแบบสอบวินิจฉัยทางพุทธิปัญญาเรื่องพันธะไอออนิก มีค่าอยู่ในช่วง 0.60 ถึง 1.00 ซึ่งเป็นค่าที่ผ่านเกณฑ์ หรือมีค่า IOC > 0.50 ตามที่ Kanjanawasee (2013) ได้ระบุไว้ จึงสรุปได้ว่า ข้อสอบทุกข้อสามารถจำแนกคุณลักษณะที่มุ่งวัดได้ตรงกับคุณลักษณะที่มุ่งวัดของเมตริกซ์คิวิที่สร้างขึ้น หรือเมตริกซ์คิวิที่สร้างขึ้นมีคุณภาพในการจำแนกคุณลักษณะที่มุ่งวัดนั่นเอง

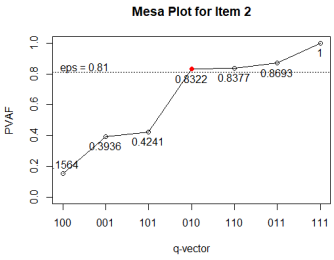
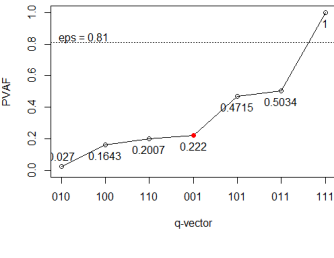
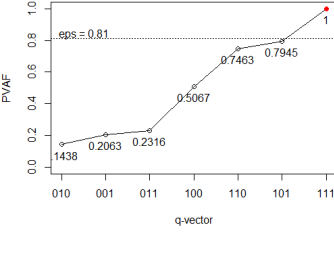
ผลการตรวจสอบค่าดัชนีอำนาจจำแนกของข้อสอบแต่ละข้อในแบบสอบวินิจฉัยทางพุทธิปัญญาเรื่องพันธะไอออนิก ด้วยฟังก์ชัน *discrim* ใน GDINA package ด้วยโปรแกรม R พบว่า ข้อสอบแต่ละข้อมีค่าดัชนีอำนาจจำแนกอยู่ในช่วง -0.064 ถึง 0.849 ซึ่งประกอบด้วยข้อสอบที่มีค่าดัชนีอำนาจจำแนกสูง 2 ข้อ ได้แก่ ข้อ 33 และ 39 โดยมีค่าเท่ากับ 0.806 และ 0.849 ตามลำดับ ส่วนข้อสอบที่มีค่าดัชนีอำนาจจำแนกปานกลาง มี 7 ข้อ ประกอบด้วย ข้อ 5, 8, 23, 25, 26, 27 และ 35 ซึ่งมีค่าเท่ากับ 0.692, 0.602, 0.662, 0.674, 0.738, 0.626 และ 0.754 ตามลำดับ นอกจากนั้นข้อสอบที่มีอำนาจจำแนกต่ำ แต่พอใช้ได้ทั้งหมด 4 ข้อ ประกอบด้วย ข้อ 1, 2, 28 และ 38 ที่มีค่าดัชนีอำนาจจำแนกเท่ากับ 0.491, 0.435, 0.463 และ 0.434



ตามลำดับ จึงสรุปได้ว่า แบบสอบวินิจฉัยทางพุทธิปัญญาเรื่องพันธะไอออนิกมีจำนวนข้อสอบที่มีดัชนีอำนาจจำแนกผ่านเกณฑ์ทั้งหมด 13 ข้อ จากทั้งหมด 40 ข้อ ซึ่งคิดเป็นร้อยละ 32.5 โดยมีรายละเอียด ดังตาราง 5

ผลการตรวจสอบความสามารถในการจำแนกคุณลักษณะของข้อสอบแต่ละข้อในแบบสอบด้วยวิธีการใช้ดัชนี PVAF มีข้อสอบจำนวน 22 ข้อ จากทั้งหมด 40 ข้อ (ร้อยละ 55.0) ที่การตรวจสอบคุณภาพด้วยวิธีการใช้ดัชนี PVAF สามารถจำแนกคุณลักษณะที่มุ่งวัดได้ตรงกับเมทริกซ์ควิที่สร้างขึ้น แต่มีข้อสอบจำนวน 9 ข้อ (ร้อยละ 22.5) ได้แก่ ข้อ 10 11 12 14 18 20 23 27 และ 33 ที่การใช้ดัชนี PVAF จำแนกคุณลักษณะที่มุ่งวัดเกิดความคลาดเคลื่อนหรือไม่สอดคล้องกับเมทริกซ์ควิที่สร้างขึ้น นอกจากนี้ยังมีข้อสอบจำนวน 9 ข้อ (ร้อยละ 22.5) ที่วิธีการใช้ดัชนี PVAF ไม่สามารถจำแนกคุณลักษณะใด ๆ ได้เลย ได้แก่ ข้อ 3 4 9 16 17 19 21 24 และ 34 ดังตาราง 5 โดยตัวอย่างการตรวจสอบคุณภาพเมทริกซ์ควิด้วยดัชนี PVAF แสดงอยู่ในรูปของกราฟ ข้อสอบข้อที่ 2 ที่เมทริกซ์ควิกำหนดคุณลักษณะที่ 2 เป็นคุณลักษณะที่มุ่งวัดในรูปของเวกเตอร์ {010} และมีเวกเตอร์ควิที่มีค่าดัชนี PVAF > EPS ขั้นต่ำ (0.81) อยู่ถึง 4 เวกเตอร์ ได้แก่ {010}, {110}, {011} และ {111} โดยแต่ละเวกเตอร์มีค่าดัชนี PVAF เท่ากับ 0.832, 0.838, 0.869 และ 1.000 ตามลำดับ แต่เวกเตอร์ที่จำแนกที่ถูกต้องจะต้องมีเพียงหนึ่งเดียวและจำแนกจำนวนคุณลักษณะน้อยที่สุด (de la Torre & Chiu, 2016; ;Nájera et al., 2019; Ma & de la Torre, 2020a) ดังนั้นข้อสอบข้อที่ 2 จึงจำแนกคุณลักษณะที่ 2 ได้ถูกต้องตรงกับคุณลักษณะที่กำหนดของเมทริกซ์ควิที่สร้างขึ้น ส่วนตัวอย่างข้อสอบข้อที่ 21 ที่เมทริกซ์ควิกำหนดคุณลักษณะที่ 3 เป็นคุณลักษณะที่มุ่งวัดในรูปของเวกเตอร์ {001} แต่มีเวกเตอร์ควิที่มีค่าดัชนี PVAF > EPS ขั้นต่ำเพียงเวกเตอร์เดียว คือ {111} โดยมีค่าดัชนี PVAF เท่ากับ 1.000 ในขณะที่เวกเตอร์ควิ {001} มีค่าดัชนี PVAF เท่ากับ 0.222 ซึ่งต่ำกว่าเกณฑ์ที่กำหนด เนื่องจากเวกเตอร์ที่จำแนกที่ถูกต้องจะต้องมีเพียงหนึ่งเดียวและจำแนกจำนวนคุณลักษณะน้อยที่สุด ส่งผลให้เวกเตอร์ {111} เป็นเวกเตอร์ที่มีจำนวนคุณลักษณะที่มุ่งวัดมากเกินไปจริงเมื่อเทียบกับเมทริกซ์ควิที่สร้างขึ้น ประกอบกับค่าดัชนี PVAF ของเวกเตอร์ {111} มีสูงกว่าเวกเตอร์ควิ {001} ที่กำหนดไว้อย่างมากจึงอาจก่อให้เกิดความคลาดเคลื่อนของเมทริกซ์ควิได้ (Nájera et al., 2019) ดังนั้นข้อสอบข้อที่ 21 จึงไม่สามารถจำแนกคุณลักษณะใด ๆ ได้ ต่อมาคือตัวอย่างข้อสอบข้อที่ 39 ที่เมทริกซ์ควิกำหนดวัดคุณลักษณะทั้ง 3 คุณลักษณะในรูปของเวกเตอร์ {111} และมีเวกเตอร์ควิที่มีค่าดัชนี PVAF > EPS ขั้นต่ำเพียง 1 เวกเตอร์ คือ {111} ที่มีค่าดัชนี PVAF เท่ากับ 1.000 ซึ่งตรงกับคุณลักษณะที่กำหนดไว้ในเมทริกซ์ควิที่สร้างขึ้น ดังนั้น ข้อสอบข้อที่ 39 จึงสามารถจำแนกคุณลักษณะได้ถูกต้องทั้ง 3 คุณลักษณะ

ตาราง 4 ตัวอย่างการตรวจสอบคุณภาพเมทริกซ์ควิตด้วยดัชนี PVAF

ข้อ ที่	ผลการตรวจสอบคุณภาพด้วยดัชนี PVAF			การแปลผล	
	Mesa plot	คุณลักษณะ			
		1	2		3
2		0	1	0	สามารถจำแนกคุณลักษณะที่ 2 ได้ เนื่องจากเวกเตอร์ {010} เป็นเวกเตอร์ที่ค่าดัชนี PVAF > EPS และมีจำนวนคุณลักษณะที่มุ่งวัดน้อยที่สุด
21		0	0	0	ไม่สามารถจำแนกคุณลักษณะที่มุ่งวัดใด ๆ ได้ เนื่องจากเนื่องจากเวกเตอร์ {111} เป็นเวกเตอร์ที่ค่าดัชนี PVAF > EPS แต่มีจำนวนคุณลักษณะที่มุ่งวัดและค่าดัชนี PVAF มากเกินจริงเมื่อเปรียบเทียบกับเวกเตอร์ {001} ที่เป็นเวกเตอร์ที่มุ่งจำแนกคุณลักษณะตามเมทริกซ์ควิตที่สร้างขึ้น (Ma & de la Torre, 2020b) และมีค่าดัชนี PVAF เพียง 0.222
39		1	1	1	สามารถจำแนกได้ทั้ง 3 คุณลักษณะ เนื่องจากเวกเตอร์ {111} เป็นเวกเตอร์เดียวที่ค่าดัชนี PVAF > EPS และเป็นเวกเตอร์ที่จำแนกคุณลักษณะจำนวนน้อยที่สุด

ผลการเปรียบเทียบการตรวจสอบคุณภาพเมทริกซ์ควิตระหว่างวิธีการตัดสินโดยผู้เชี่ยวชาญกับวิธีการใช้ดัชนี PVAF พบว่า ในการวิจัยครั้งนี้ เมทริกซ์ควิตที่ผ่านการตรวจสอบคุณภาพด้วยการตัดสินโดยผู้เชี่ยวชาญสามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์ควิตที่สร้างขึ้นมากกว่าวิธีการใช้ดัชนี PVAF โดยเมทริกซ์ควิตที่ผ่านการตรวจสอบคุณภาพโดยใช้วิธีการตัดสินโดยผู้เชี่ยวชาญสามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์ควิตที่สร้างขึ้นทั้งหมด ส่วนเมทริกซ์ควิตที่ผ่านการตรวจสอบคุณภาพโดยใช้ดัชนี PVAF มีข้อสอบ 22 ข้อ (ร้อยละ 55.0) ที่สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์ควิตที่สร้างขึ้น แต่มีข้อสอบจำนวน 9 ข้อ (ร้อยละ 22.5) ที่จำแนกคุณลักษณะคลาดเคลื่อนหรือไม่ตรงกับเมทริกซ์ควิตที่สร้างขึ้น และมีข้อสอบจำนวน 9 ข้อ (ร้อยละ 22.5) ที่ไม่สามารถจำแนกคุณลักษณะใด ๆ ได้







ข้อ ที่	อำนาจ จำแนก	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยใช้									ผลการเปรียบเทียบ
		ผู้เชี่ยวชาญ						ดัชนี PVAF			
		IOC	คุณลักษณะ			PVAF ที่จำแนก ได้	คุณลักษณะ				
			1	2	3		1	2	3		
31	0.321	1.0	1	1	0	0.867	1	1	0	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญและดัชนี PVAF สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้น เนื่องจากทั้งสองวิธีมีค่าดัชนี IOC > 0.5 ส่วนค่าดัชนี PVAF > 0.81 และจำแนกคุณลักษณะจำนวนน้อยที่สุด ตามลำดับ	
32	0.330	1.0	1	0	1	0.837	1	0	1	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญและดัชนี PVAF สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้น เนื่องจากทั้งสองวิธีมีค่าดัชนี IOC > 0.5 ส่วนค่าดัชนี PVAF > 0.81 และเป็นค่าในการจำแนก 2 คุณลักษณะที่มากที่สุด	
33	0.806	1.0	0	1	1	0.931	0	0	1	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญสามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้นเนื่องจากมีค่าดัชนี IOC > 0.5 แต่เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพด้วยดัชนี PVAF เกิดความคลาดเคลื่อนในการจำแนกคุณลักษณะเมื่อเทียบกับเมทริกซ์คิวที่สร้างขึ้น เนื่องจากค่าดัชนี PVAF ที่ผ่านเกณฑ์และจำแนกคุณลักษณะน้อยที่สุดคือเวกเตอร์ {001} ไม่ใช่เวกเตอร์คิว {011}	
34	0.314	1.0	1	1	0	1.000	0	0	0	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญสามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้นเนื่องจากมีค่าดัชนี IOC > 0.5 แต่เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพด้วยดัชนี PVAF ไม่สามารถจำแนกคุณลักษณะใด ๆ ได้ เนื่องจากค่าดัชนี PVAF ที่ผ่านเกณฑ์เป็นของเวกเตอร์คิว {111} ซึ่งมีค่าที่มากเกินไปจริง ส่วนค่าดัชนี PVAF ของเวกเตอร์คิว {110} ที่มุ่งจำแนกมีค่าเพียง 0.611	
35	0.754	1.0	1	0	1	0.954	1	0	1	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญและดัชนี PVAF สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้น เนื่องจากทั้งสองวิธีมีค่าดัชนี IOC > 0.5 ส่วนค่าดัชนี PVAF > 0.81 และจำแนกคุณลักษณะจำนวนน้อยที่สุด ตามลำดับ	
36	0.299	1.0	0	1	1	0.958	0	1	1	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญและดัชนี PVAF สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้น เนื่องจากทั้งสองวิธีมีค่าดัชนี IOC > 0.5 ส่วนค่าดัชนี PVAF > 0.81 และจำแนกคุณลักษณะจำนวนน้อยที่สุด ตามลำดับ	
37	0.135	1.0	1	1	1	1.000	1	1	1	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญและดัชนี PVAF สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้น เนื่องจากทั้งสองวิธีมีค่าดัชนี IOC > 0.5 ส่วนค่าดัชนี PVAF ที่ผ่านเกณฑ์มีทั้งหมด 2 ค่าซึ่งเป็นของเวกเตอร์คิว {111} และ {101} แต่เนื่องจากค่าดัชนี PVAF ของเวกเตอร์คิว {111} มีค่าสูงกว่าเวกเตอร์คิว {101} อยู่ 0.163 ซึ่งมีค่าสูงพอที่จะสรุปได้ว่าข้อสอบข้อนี้มุ่งจำแนกทั้ง 3 คุณลักษณะ (Ma & de la Torre, 2020b)	
38	0.434	0.8	1	1	1	1.000	1	1	1	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญและดัชนี PVAF สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้น เนื่องจากทั้งสองวิธีมีค่าดัชนี IOC > 0.5 ส่วนค่าดัชนี PVAF > 0.81 และจำแนกคุณลักษณะจำนวนน้อยที่สุด ตามลำดับ	
39	0.849	1.0	1	1	1	1.000	1	1	1	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญและดัชนี PVAF สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้น เนื่องจากทั้งสองวิธีมีค่าดัชนี IOC > 0.5 ส่วนค่าดัชนี PVAF > 0.81 และจำแนกคุณลักษณะจำนวนน้อยที่สุด ตามลำดับ	
40	0.058	1.0	1	1	1	1.000	1	1	1	เมทริกซ์คิวที่ผ่านการตรวจสอบคุณภาพโดยผู้เชี่ยวชาญและดัชนี PVAF สามารถจำแนกคุณลักษณะได้ตรงกับเมทริกซ์คิวที่สร้างขึ้น เนื่องจากทั้งสองวิธีมีค่าดัชนี IOC > 0.5 ส่วนค่าดัชนี PVAF > 0.81 และจำแนกคุณลักษณะจำนวนน้อยที่สุด ตามลำดับ	

## อภิปรายผล

จากผลการวิจัยสามารถนำมาใช้ในการอภิปรายผลเกี่ยวกับสาเหตุที่ทำให้เมตริกซ์ควิที่ผ่านการตรวจสอบคุณภาพด้วยวิธีการตัดสินโดยผู้เชี่ยวชาญมีความสามารถในการจำแนกคุณลักษณะได้ตรงกับเมตริกซ์ควิที่สร้างขึ้นมากกว่าวิธีการใช้ดัชนี PVAF ได้ว่า ความสามารถในการจำแนกคุณลักษณะของเมตริกซ์ควิจากวิธีการตัดสินโดยผู้เชี่ยวชาญขึ้นอยู่กับแนวคิดของผู้เชี่ยวชาญกับผู้สร้างแบบสอบ (Qin et al., 2020) ซึ่งในการวิจัยครั้งนี้ทั้งผู้เชี่ยวชาญ 5 ท่านและผู้สร้างแบบสอบมีแนวคิดเกี่ยวกับคุณลักษณะในเรื่องพันธะไอออนิกทั้ง 3 คุณลักษณะไปในทิศทางเดียวกัน ส่งผลให้ค่าดัชนี IOC ของข้อสอบทุกข้อมีค่าผ่านเกณฑ์ที่กำหนดทำให้เมตริกซ์ควิสามารถจำแนกได้ตรงกับเมตริกซ์ควิที่สร้างขึ้น ส่วนความสามารถในการจำแนกคุณลักษณะของเมตริกซ์ควิด้วยวิธีการใช้ดัชนี PVAF นั้นขึ้นอยู่กับความหลากหลายในรูปแบบการรับรู้คุณลักษณะหรือความสามารถของผู้สอบ (de la Torre & Chiu, 2016) เมื่อตัวอย่างในการวิจัยครั้งนี้มีคุณลักษณะไม่แตกต่างกัน ทำให้ความหลากหลายของคุณลักษณะต่ำ จึงก่อให้เกิดความคลาดเคลื่อนในการจำแนกคุณลักษณะ หรือไม่สามารจำแนกคุณลักษณะของผู้สอบได้ อีกทั้ง เมื่อคุณลักษณะของผู้สอบขาดความหลากหลาย ส่งผลให้ข้อสอบไม่สามารถจำแนกความแตกต่างของคุณลักษณะของผู้สอบได้ ค่าดัชนีอำนาจจำแนกของข้อสอบจึงมีค่าต่ำ ซึ่งส่งผลโดยตรงต่อคุณภาพของเมตริกซ์ควิที่ผ่านการตรวจสอบคุณภาพด้วยดัชนี PVAF ถึงแม้ว่าผู้วิจัยดำเนินการเลือกตัวอย่างวิจัยโดยคำนึงถึงความหลากหลายทางคุณลักษณะในการเรียนรู้ในรายวิชาเคมี แต่เนื่องจากการทำแบบสอบของตัวอย่างวิจัยขาดแรงจูงใจในการทำแบบสอบ ส่งผลให้ตัวอย่างในวิจัยบางคนละเลย ไม่ตั้งใจทำแบบสอบอย่างเต็มความสามารถ หรือมีการเดาคำตอบเกิดขึ้น จึงส่งผลต่อค่าดัชนีอำนาจจำแนกของข้อสอบ (Wise & Kong, 2005 as cited in Finn, 2015) ส่วนความคลาดเคลื่อนในการจำแนกคุณลักษณะของเมตริกซ์ควิด้วยวิธีการใช้ดัชนี PVAF ในข้อสอบข้อที่ 33 พบว่าเวกเตอร์ควิ {011} ซึ่งเป็นเวกเตอร์ที่มุ่งจำแนกคุณลักษณะมีค่าดัชนี PVAF ต่ำกว่าเวกเตอร์ควิ {101} อยู่เพียง 0.03 ซึ่งความคลาดเคลื่อนดังกล่าวอาจเกิดจากจำนวนตัวอย่างในการวิจัย โดยในการวิจัยครั้งนี้ที่เป็นเพียงจำนวนขั้นต่ำหรือจำนวน 200 คน ตามที่ Nájera et al. (2019) ได้ระบุไว้ ส่งผลให้ความหลากหลายของคุณลักษณะในวิจัยต่ำ ทำให้การตรวจสอบคุณภาพด้วยการใช้ดัชนี PVAF เกิดความคลาดเคลื่อนไปจากเมตริกซ์ควิที่สร้างขึ้น โดยสาเหตุที่ทำให้ผู้วิจัยเก็บข้อมูลกับตัวอย่างในวิจัยได้จำนวนที่จำกัดเป็นเพราะสถานการณ์การแพร่ระบาดของเชื้อไวรัสโคโรนา-19 ก่อให้เกิดความยากลำบากในการเก็บข้อมูลนั่นเอง

## ข้อเสนอแนะ

### 1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

การใช้ดัชนี PVAF ในการตรวจสอบคุณภาพเมตริกซ์ควิควบคู่กับการตรวจสอบคุณภาพเมตริกซ์ควิด้วยวิธีการตัดสินโดยผู้เชี่ยวชาญจะช่วยให้มีหลักฐานทางสถิติและเป็นหลักฐานเชิงประจักษ์ในการสนับสนุนความถูกต้องมากยิ่งขึ้น (de la Torre & Chiu, 2016; Nájera et al., 2019) แต่การใช้ดัชนี PVAF ในการตรวจสอบคุณภาพเมตริกซ์ควิมีปัจจัยสำคัญที่ควรคำนึง คือ จำนวนตัวอย่างในวิจัยที่ควรมีมากเพียงพอที่ก่อให้เกิดความหลากหลายในรูปแบบการรับรู้คุณลักษณะของตัวอย่างในวิจัย โดย Nájera et al. (2019)

ได้เสนอว่า จำนวนตัวอย่างในวิจัยขั้นต่ำ คือ 200 คน ซึ่งจำนวนดังกล่าวจะส่งผลให้ความหลากหลายของรูปแบบความรู้คุณลักษณะต่ำ ส่งผลให้ค่าดัชนี PVAF มีแนวโน้มที่ต่ำจนมีน้อยกว่าค่าเกณฑ์ที่กำหนดหรือค่า EPS แต่หากตัวอย่างในวิจัยมีจำนวน 500 คนขึ้นไป จะทำให้ความหลากหลายในรูปแบบความรู้คุณลักษณะของตัวอย่างในวิจัยสูงขึ้น ส่งผลให้การใช้ดัชนี PVAF ในการตรวจสอบคุณภาพเมทริกซ์ควมมีแนวโน้มสูงขึ้นจนมากกว่าค่าเกณฑ์ที่กำหนดหรือค่า EPS

## 2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

ในการวิจัยครั้งต่อไป ควรมีการตรวจสอบความหลากหลายในรูปแบบความรู้คุณลักษณะของตัวอย่างในวิจัยก่อนการเก็บข้อมูล เพื่อคัดเลือกตัวอย่างที่มีความหลากหลายในรูปแบบความรู้คุณลักษณะ และทำให้ข้อสอบแต่ละข้อสามารถจำแนกคุณลักษณะของตัวอย่างในวิจัยได้มากยิ่งขึ้น ในส่วนของจำนวนตัวอย่างในวิจัย ควรมีจำนวน 500 คน เป็นต้นไป เพื่อให้เกิดความหลากหลายในรูปแบบความรู้คุณลักษณะ เพื่อให้การตรวจสอบคุณภาพเมทริกซ์ควมด้วยดัชนี PVAF มีความถูกต้องและแม่นยำมากยิ่งขึ้น

## References

- de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, 76(2), 179–199. <https://doi.org/10.1007/s11336-011-9207-7>
- de la Torre, J. & Chiu, C. (2016). A General Method of Empirical Q-matrix Validation. *Psychometrika*. 81. 10.1007/s11336-015-9467-8. Finn, B., 2015. Measuring Motivation in Low-Stakes Assessments. *ETS Research Report Series*, 1–17. <https://doi:10.1002/ets2.12067>
- Jang, E.E. & Wagner, M. (2013). Diagnostic Feedback in the Classroom. In *The Companion to Language Assessment*, A.J. Kunnan (Ed.). <https://doi.org/10.1002/9781118411360.wbcla081>
- Johnson, R. L., Penny, J. A., & Gordon, B. (2009). *Assessing performance: Designing, scoring, and validating performance tasks*. Guilford Press.
- Ma, W., & de la Torre, J. (2020a). An empirical Q-matrix validation method for the sequential generalized DINA model. *The British journal of mathematical and statistical psychology*, 73(1), 142–163. <https://doi.org/10.1111/bmsp.12156>
- Ma W, de la Torre J. (2020b). GDINA: An R Package for Cognitive Diagnosis Modeling. *Journal of Statistical Software*, 93(14), 1–26. doi: 10.18637/jss.v093.i14.
- Nájera, P., Sorrel, M. A., & Abad, F. J. (2019). Reconsidering cutoff points in the general method of empirical Q-matrix validation. *Educational and Psychological Measurement*, 79(4), 727–753. <https://doi.org/10.1177/0013164418822700>



- Qin, C., Jia, S., Fang, X., & Yu, X. (2020). Relationship validation among items and attributes. *Journal of Statistical Computation and Simulation*, 90(18), 3360-3375.
- Waugh, C. K., & Gronlund, N. E. (2013). *Assessment of Student Achievement* (10th ed.). Pearson.

### Translated Thai References

- Office of the Education Council. (2019). *Guidelines for Development of Students Competency at the Basic Education Level* (1st ed.). Ministry of Education Thailand, 1-15. (in Thai)
- Kanjanawasee, S. (2013). *Classical Test Theory* (7th ed). Chulalongkorn University. (in Thai)
- Meesakul, S., Naiyapatana, O., Khampalikit ., & Kritkharuehart, S. (2015). The Study for the Efficiency of Diagnostic Classification Models. *Research Methodology & Cognitive Science*, 13(1), 27-37. (in Thai)
- Chaimongkol, N. (2017). Assessment for Cognitive Diagnosis. *Journal of The Social Science Research Association of Thailand*. 4(1), 14-23. (in Thai)