

วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม

Journal of Educational Measurement

Mahasarakham University



วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม

Journal of Educational Measurement Mahasarakham University

ปีที่ 29 ฉบับที่ 1: มกราคม – มิถุนายน 2566

Vol.29 No.1: January – June 2023

ISSN 2985-0614 (Online)

วารสารตีพิมพ์ปีละ 2 ฉบับ โดยมีกำหนดการพิมพ์เผยแพร่
ฉบับที่ 1: มกราคม - มิถุนายน
ฉบับที่ 2: กรกฎาคม - ธันวาคม

ภาควิชาวิจัยและพัฒนาการศึกษา

คณะศึกษาศาสตร์

มหาวิทยาลัยมหาสารคาม

วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม
ปีที่ 29 ฉบับที่ 1: มกราคม – มิถุนายน 2566
Journal of Educational Measurement Mahasarakham University
Vol.29 No.1: January – June 2023

Aim & Scope

กองบรรณาธิการจะรับพิจารณาบทความที่เกี่ยวข้องกับศาสตร์การวัดและประเมินผลการศึกษา การวิจัยทางการศึกษา สถิติและจิตวิทยาการศึกษาหรือที่เกี่ยวข้อง ทั้งนี้ ขึ้นอยู่กับดุลยพินิจของกองบรรณาธิการในการพิจารณา

วัตถุประสงค์

1. เพื่อส่งเสริมและเผยแพร่การศึกษาค้นคว้าการวิจัยและความรู้ใหม่ๆ ในสาขาวิชาการวัดผลและการวิจัยการศึกษา ตลอดจนวิชาอื่น ๆ ที่เกี่ยวข้อง
2. เพื่อเป็นสื่อแลกเปลี่ยนความคิดเห็นทางวิชาการ ทั้งทางทฤษฎีและการปฏิบัติระหว่างอาจารย์นักวิชาการและนิสิต
3. เพื่อส่งเสริมการเสนอผลงานทางวิชาการของอาจารย์และนิสิต
4. เพื่อเป็นเอกสารค้นคว้าอ้างอิงของอาจารย์ นิสิตและบุคลากร ผู้สนใจในความก้าวหน้าทางวิชาการ

ออกแบบปก

นางสาวอรุณภา คำบัวโคตร

ติดต่อวารสาร

ภาควิชาวิจัยและพัฒนาการศึกษา อาคารวิทยพัฒนา ชั้น 8 คณะศึกษาศาสตร์
มหาวิทยาลัยมหาสารคาม ถ.นครสวรรค์ ต.ตลาด อ.เมืองมหาสารคาม จ.มหาสารคาม 44000
โทรศัพท์: 043-754321-40 ต่อ 6272
E-mail: Jem2msu@gmail.com

บทความทุกบทความที่ตีพิมพ์ในวารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม
จะต้องผ่านการตรวจประเมินคุณภาพจากผู้ทรงคุณวุฒิตั้งน้อย 2 ท่าน
ในรูปแบบ Double – Blinded Review

บทบรรณาธิการ

วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม ผ่านการประเมินคุณภาพวารสารวิชาการ
ในฐานข้อมูล TCI รอบที่ 4 พ.ศ. 2563 – 2567 โดยถูกจัดให้เป็นวารสารกลุ่มที่ 1 (จนถึง 31 ธันวาคม 2567)

ในปี พ.ศ. 2566 วารสารได้ยกเลิก ISSN 2651-1444 (Online) และเปลี่ยนเป็น ISSN 2985-0614
(Online) สำหรับวารสารปีที่ 29 ฉบับที่ 1: มกราคม – มิถุนายน 2566 เป็นวารสารที่เผยแพร่บทความ
ในรูปแบบออนไลน์เท่านั้น โดยวารสารฉบับนี้ มีบทความนำเสนอทั้งหมด 20 เรื่อง ซึ่งประกอบด้วยบทความ
วิชาการ 1 เรื่อง และบทความวิจัย 19 เรื่อง จากหน่วยงานทั้งภายในและภายนอก โดยผ่านการพิจารณา
กลั่นกรองจากผู้ทรงคุณวุฒิ 2 - 3 ท่าน ในรูปแบบ Double – Blinded Review

กองบรรณาธิการหวังเป็นอย่างยิ่งว่าวารสารฉบับนี้จะเป็นประโยชน์ทางวิชาการสำหรับผู้อ่านทุกท่าน
หากท่านมีข้อเสนอแนะประการใด กองบรรณาธิการยินดีน้อมรับคำแนะนำเพื่อจะได้นำไปปรับปรุงและพัฒนา
คุณภาพวารสารให้มีมาตรฐาน ยิ่งขึ้นต่อไป

ผู้ช่วยศาสตราจารย์ ดร.ทัศนศิริรินทร์ สว่างบุญ

บรรณาธิการ



ทีมบรรณาธิการ

ที่ปรึกษาทั่วไป

อธิการบดี มหาวิทยาลัยมหาสารคาม
คณบดีคณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม
หัวหน้าภาควิชาวิจัยและพัฒนการศึกษา

ที่ปรึกษากองบรรณาธิการ

ศ.ดร.ศิริชัย กาญจนวาสี
จุฬาลงกรณ์มหาวิทยาลัย
รศ.ดร.บุญชม ศรีสะอาด
มหาวิทยาลัยมหาสารคาม
รศ.ดร.ประวิต เอราวรรณ์
สำนักงานคณะกรรมการข้าราชการครูและ
บุคลากรทางการศึกษา
รศ.สมนึก ภัททิยธนี
ข้าราชการเกษียณ

บรรณาธิการ

ผศ.ดร.ทัศน์ศิริรินทร์ สว่างบุญ มหาวิทยาลัยมหาสารคาม

กองบรรณาธิการ

รศ.ดร.ประเสริฐ เรือนนระการ มหาวิทยาลัยมหาสารคาม
รศ.ดร.ทรงศักดิ์ ภูสีอ่อน มหาวิทยาลัยมหาสารคาม
รศ.ดร.ญาณภัทร สีหะมงคล มหาวิทยาลัยมหาสารคาม

ผศ.ดร.อรนุช วราอัศวปติ ศรีสะอาด มหาวิทยาลัยมหาสารคาม
รศ.ดร.วราพร เอราวรรณ์ มหาวิทยาลัยมหาสารคาม
รศ.ดร.อพนันต์รี พูลพุทธา มหาวิทยาลัยมหาสารคาม
อ.ดร.สาวตรี ราญมิชัย มหาวิทยาลัยมหาสารคาม
รศ.ดร.ณัฐภรณ์ หลาวทอง จุฬาลงกรณ์มหาวิทยาลัย
รศ.ดร.ชนะศึก นิชานนท์ มหาวิทยาลัยสวนดุสิต
รศ.ดร.สมบัติ ท้ายเรือคำ มหาวิทยาลัยเชียงใหม่
รศ.ดร.สุนทรพจน์ ดำรงพานิช มหาวิทยาลัยเชียงใหม่
ผศ.ดร.ศักดิ์สิทธิ์ ฤทธิธิน มหาวิทยาลัยกาฬสินธุ์
ผศ.ดร.สาธิตา สกุลรัตน์กุลชัย มหาวิทยาลัยธรรมศาสตร์
ผศ.ดร.พินดา วราสุนันท์ มหาวิทยาลัยเกษตรศาสตร์
ผศ.ดร.จตุภูมิ เขตจัตุรัส มหาวิทยาลัยขอนแก่น
ผศ.ดร.สุกัญญา รุจิเมธธาส มหาวิทยาลัยราชภัฏอุดรดิตถ์
ผศ.ดร.สรัญญา จันทร์ชูสกุล มหาวิทยาลัยศิลปากร
ผศ.ดร.กมลทิพย์ ศรีหาเศษ มหาวิทยาลัยรามคำแหง
อ.ดร.ชลี ภัทรพิชญธรรม Southern Methodist
University, USA

ผู้ทรงคุณวุฒิตรวจสอบความถูกต้องด้านภาษา

อาจารย์สกล สรเสนา

ฝ่ายประสานงาน

นางสาวเอมอร สิงห์พันธ์

สารบัญ

บทความ	หน้า
บทความวิชาการ.....	1
การประเมินสมรรถนะของผู้เรียนในการเรียนรู้วิทยาศาสตร์ Assessing Learners' Competencies in Science Learning ชนินันท์ พุกฤษ์ประมุล.....	1
บทความวิจัย.....	20
การประเมินหลักสูตรแบบ CIPPI Model : การศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์ และคณิตศาสตร์ คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม CIPPI Model for Curriculum Evaluation: Master of Education in Teaching of Science and Mathematics, Faculty of Education, Mahasarakham University กัญญารัตน์ โคจร และ กันยารัตน์ สอนสุภาพ.....	20
โมเดลสมการโครงสร้างการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงาน ของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 The Structural Equation Modeling of Diversity Generation Management Affecting Performance Effectiveness of the Teachers under the Office of Khon Kaen Primary Education Service Area 4 ณัฐชญา เพ็ญธรรม และ เสาวณี สิริสุขศิลป์.....	37
การวิเคราะห์องค์ประกอบเชิงยืนยันการสร้างสรรค์และนวัตกรรมสำหรับนักศึกษามหาวิทยาลัยราชภัฏ The Confirmatory Factor Analysis of Creativity and Innovation for Rajabhat University Students ธราเทพ เตมีรักษ์ และ อารยา ปิยะกุล.....	53
อิทธิพลพัฒนาการการลดความสิ้นเปลืองในการใช้จ่ายภาครัฐที่มีต่อพัฒนาการความร่วมมือด้าน การวิจัยและพัฒนา ระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา : การวิเคราะห์โมเดลโค้งพัฒนาการคู่ขนาน Effect of Wastefulness of Government Spending Reduction Development on Growth of University-Industry Collaboration in R&D and Quality of the Education System: A Parallel Latent Growth Curve Modeling ธีรยุทธ พิริยะอารยะกุล, ณัฐพล อนันต์ธนสาร และ ชยุตม์ ภิรมย์สมบัติ.....	67

สารบัญ (ต่อ)

บทความ

หน้า

การเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์ รูปแบบผสม ที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน	
A Comparison of the Item Information Function of the Mixed-Format Test for Mathematical Literacy with Different Test Lengths and Proportions of Items	
นุชรากร นันตะสุยะ, จุฑาภรณ์ มาสันเทียะ และ พนิดา พานิชวัฒนะ.....	84
การออกแบบวิธีการตรวจให้คะแนนอัตโนมัติเพื่อวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะ วิทยาศาสตร์ผ่านการเรียนรู้ของเครื่อง	
Designing Automated Scoring Approach to Diagnose Scientific Competency Misconceptions through Machine Learning	
ประไพพิมพ์ ทองเปราะ, พัชร จันทร์เพ็ง, ชัยวัฒน์ ทะวะรุ่งเรือง และ ธนพงศ์ อินทระ.....	99
การพัฒนาแบบวัดการตระหนักรู้สถานการณ์ของนักศึกษาวิสัญญีพยาบาล	
Development of Situation Awareness Scale for Nurse Anesthetist Students	
ปาริชาติ อภิเดชากุล, ศิริเดช สุชีวะ, กมลวรรณ ตังธนกานนท์, พงศ์ธรา วิจิตเวชไพศาล, วิรัตน์ พันชุกกลาง และ ธัชนัน จินตดาวงษ์.....	117
การเปรียบเทียบความเที่ยงและความตรงตามเกณฑ์ในการตรวจให้คะแนนของผู้ตรวจ ที่มีคุณลักษณะต่างกันของแบบสอบัดนัยสำหรับวัดสมรรถนะทางวิทยาศาสตร์ของนักเรียน ชั้นมัธยมศึกษาปีที่ 3	
The Comparison of Reliability and Validity of the Raters' Scoring Criteria with Different Characteristics of Raters of Essay Tests for Measuring Scientific Competence of Grade 9 Students	
พนิดา จังหะ และ ประกฤติยา ทักษิโณ.....	133
การวิเคราะห์ปัจจัยด้านความเหลื่อมล้ำทางการศึกษาที่ส่งผลต่อความฉลาดรู้คณิตศาสตร์ภายใต้ โปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล (PISA)	
Analysis of Educational Inequality Factors Affecting Mathematical Literacy under the International Student Assessment Program (PISA)	
มาริษา นามกุล และ ประกฤติยา ทักษิโณ.....	148
รูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่ การศึกษาประถมศึกษาชัยภูมิเขต 3	
A Model for Development of Digital Learning Culture in Primary Schools under Chaiyaphum Primary Education Service Area Office 3	
รติยากร คชา และ ดาวรรุวรรณ ถวิลการ.....	161

สารบัญ (ต่อ)

บทความ

หน้า

การเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ Enhancing Complex Problem-Solving Skills of Adolescents by Applying Information Processing Theory and Self-Regulation Learning วิหวัศ เพ็ญภู, พูลพงศ์ สุขสว่าง และ พีร วงศ์อุปราช.....	176
การออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ โดยการให้ข้อมูลป้อนกลับแบบอัตโนมัติผ่านการเรียนรู้ของเครื่อง Designing Methods to Diagnose Scientific Misconceptions by using Automated feedback through Machine Learning ศรัณย์ธร อินทศิริ, พัชรี จันทร์เพ็ง, ชัยวัฒน์ ทะวะรุ่งเรือง และ ธนพงศ์ อินทระ.....	193
การออกแบบระบบตรวจให้คะแนนแบบสอบอัตนัย โดยการให้ข้อมูลป้อนกลับอัตโนมัติเพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่านการเรียนรู้ของเครื่อง Designing Automated Scoring System of Open-Ended Test by Providing Automatic Feedback to Diagnose Mathematical Proficiency Levels through Machine Learning ศรารรรณ สุวิชัย, พัชรี จันทร์เพ็ง, ชัยวัฒน์ ทะวะรุ่งเรือง และ ธนพงศ์ อินทระ.....	210
การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 Detecting Differential Item Functioning of Grade 9 Science Literacy Test ศรารุท โพธิ์สิงห์ และ วราพร เอราวรรณ.....	231
การพัฒนาแบบสอบคู่ขนานวิชาคณิตศาสตร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนสังกัดกรุงเทพมหานคร โดยการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ Development of Parallel Tests in Mathematics for Students in Grade 2 in Schools under the Bangkok Metropolitan Administration through an Application of Two-Parameter Item Response Theory สาวินา บุญเสน, จุฑาภรณ์ มาสันเทียะ และ พนิดา พานิชวัฒนะ.....	245
กลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 School Management Strategy to Develop Students' Life Skills According to the Philosophy of Sufficiency Economy by Using the Professional Learning Community Process of Schools under the Office of Udon Thani Primary Education Service Area 2 สุธรรม ธรรมทัศนานนท์.....	259

สารบัญ (ต่อ)

บทความ	หน้า
การเปรียบเทียบคุณภาพของการตรวจให้คะแนนของแบบสอบอัตนัยวัดความสามารถในการ แก้ปัญหาทางคณิตศาสตร์ ที่ใช้การออกแบบการตรวจให้คะแนนที่ต่างกัน: การประยุกต์ใช้ ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือผลของการวัด Quality Comparison of Scoring of Essay Tests to Measure Mathematics Problem Solving Ability Using Different Scoring Designs: Application of Generalizability Theory สุรัชย์ มาตราข และ ประกฤติยา ทักษิโณ.....	274
การศึกษาอิทธิพลของชีวิตการทำงานที่มีต่อภาวะสุขภาพจิต และความเหนื่อยหน่าย ในการทำงานของครู Examining the Influences of Areas of Work-life on Teachers' Mental Health and Job Burnout อารยา ปิยะกุล.....	291
การพัฒนาแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 Development of Technological Innovation Literacy Scale for Grade 12 Students เอกพัฒน์ เสงี่ยมใส่อง, ชำนาญ ปาณาวงษ์, น้ำทิพย์ ่องอาจวาณิชย์ และ นันทิมา นาคาพงศ์ อัครวัชร.....	308
ภาคผนวก.....	326
รายชื่อผู้ทรงคุณวุฒิพิจารณาบทความ วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม ปีที่ 28 ฉบับที่ 2 (กรกฎาคม – ธันวาคม 2565).....	327
หลักเกณฑ์และคำแนะนำสำหรับผู้เขียน.....	328



บทความวิชาการ

Academic Article

การประเมินสมรรถนะของผู้เรียนในการเรียนรู้วิทยาศาสตร์ Assessing Learners' Competencies in Science Learning

ชนินันท์ พฤกษ์ประมูล^{1*}Chaninan Pruekpramool^{1*}

(Received: February 4, 2022; Revised: May 16, 2022; Accepted: May 30, 2022)

บทคัดย่อ

การประเมินฐานสมรรถนะเป็นองค์ประกอบสำคัญของการศึกษาฐานสมรรถนะ ซึ่งเป็นการจัดการเรียนรู้รูปแบบหนึ่งที่เน้นการพัฒนาสมรรถนะผู้เรียนเพื่อแก้ปัญหาในโลกความเป็นจริง การประเมินฐานสมรรถนะมีลักษณะสำคัญห้าประการ ได้แก่ (1) เน้นประเมินการปฏิบัติงานและประเมินตามสภาพจริง (2) ใช้เครื่องมือวัดและเกณฑ์ที่หลากหลาย (3) เชื่อมโยงกับสถานการณ์ในโลกความเป็นจริง และติดตามการเปลี่ยนแปลงสมรรถนะของผู้เรียน (4) เป็นการประเมินรายบุคคล ที่เกิดขึ้นได้ทั้งก่อน ระหว่างและหลังเรียนรู้ และ (5) ให้ผลสะท้อนกลับเพื่อพัฒนากระบวนการเรียนรู้ของผู้เรียน การประเมินสมรรถนะของผู้เรียนในการเรียนรู้วิทยาศาสตร์สอดคล้องกับลักษณะข้างต้น ครูวิทยาศาสตร์ใช้แนวทางและวิธีการที่หลากหลายในชั้นเรียนของตนเอง เช่น การมอบหมายงาน การสร้างแผนผังมโนทัศน์ และการอภิปรายประเด็นทางวิทยาศาสตร์ บทความนี้เสนอการออกแบบการประเมินสมรรถนะของผู้เรียนในชั้นเรียนวิทยาศาสตร์โดยแบ่งเป็นสี่ขั้นตอน ได้แก่ ขั้นที่ 1 การระบุสมรรถนะ ขั้นที่ 2 การเลือกวิธีการและเครื่องมือประเมินสมรรถนะ ขั้นที่ 3 การสร้างเครื่องมือประเมินและรูบริค และ ขั้นที่ 4 การตรวจสอบคุณภาพเครื่องมือประเมินและรูบริค นอกจากนี้ บทความยังได้เสนอตัวอย่างเครื่องมือและรูบริคเพื่อสามารถประยุกต์ใช้กับบริบทที่แตกต่างกันอย่างเหมาะสม

คำสำคัญ: สมรรถนะ การศึกษาฐานสมรรถนะ การประเมินฐานสมรรถนะ การเรียนรู้วิทยาศาสตร์

Abstract

Competency-based assessment is a significant component of competency-based education that focuses on teaching and learning to develop learners' competencies in solving real-world problems. There are 5 key features of competency-based assessment: (1) emphasizing performance assessment and authentic assessment, (2) using various measurement tools and

¹ ผู้ช่วยศาสตราจารย์ ศูนย์วิทยาศาสตร์ศึกษา คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

¹ Assistant Professor, Science Education Center, Faculty of Science, Srinakharinwirot University

*Corresponding Author E-mail: chaninan@g.swu.ac.th

criteria, (3) connecting the assessment to real-world contexts and tracking changes in learners' competencies, (4) assessing learners individually before, during, and after their learning, and (5) providing feedback to improve their learning.

Assessment of learners' competencies in science learning corresponds to the features above. Science teachers use several approaches and methods to assess learning in classroom such as giving assignments, creating concept maps, and discussing science topics. The article presents 4 steps of designing a classroom assessment of learners' competencies, including (1) identifying competencies, (2) choosing the assessment method and tool, (3) creating the tool and rubrics, and (4) validating the tool and rubrics. Moreover, the article presents examples of tools and rubrics to apply to different contexts appropriately.

Keywords: competency, competency-based education, competency-based assessment, science learning

บทนำ

การศึกษาฐานสมรรถนะ (competency-based education) เป็นรูปแบบการศึกษาที่ประเทศไทยโดยสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานกำลังพัฒนาขึ้นเพื่อใช้เป็นกรอบการศึกษาชาติ (national education framework) ปัญหาการศึกษาชาติ ไม่ว่าจะเป็น ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนต่ำและผู้เรียนมีความสามารถที่ไม่สะท้อนสมรรถนะในการทำงาน (Ruchaiphanit, 2016; Office of the Education Council, 2019) เป็นสาเหตุสำคัญของการปรับปรุงหลักสูตร อิงมาตรฐานการเรียนรู้ในปัจจุบันของประเทศไทย (Ministry of Education, 2010) ให้สามารถพัฒนาศักยภาพผู้เรียนให้พร้อมรับมือปัญหาและความท้าทายต่าง ๆ สามารถดำรงชีวิตอย่างเป็นปกติสุข ภายใต้ภาวะการเปลี่ยนแปลง (CBE Thailand, 2021a) นอกจากการปรับเปลี่ยนการจัดการเรียนรู้แล้ว การปรับเปลี่ยนการประเมินการเรียนรู้ของผู้เรียน เป็นภารกิจสำคัญที่ต้องสอดคล้องกับการจัดการเรียนรู้ สำหรับวิชาวิทยาศาสตร์นั้น มีธรรมชาติของรายวิชาที่เน้นการเรียนรู้แบบสืบเสาะเพื่อให้ผู้เรียนสร้างองค์ความรู้ แล้วนำไปใช้อธิบายปรากฏต่าง ๆ ในธรรมชาติสามารถดำรงชีวิตภายใต้สถานการณ์ปัจจุบันสำหรับบทความนี้ผู้เขียนนำเสนอถึง การศึกษาฐานสมรรถนะ ความหมายและธรรมชาติของสมรรถนะ การประเมินสมรรถนะของผู้เรียนในการเรียนวิทยาศาสตร์ รวมทั้งเสนอขั้นตอนการออกแบบการประเมินสมรรถนะของผู้เรียนในชั้นเรียนวิทยาศาสตร์ พร้อมทั้งตัวอย่างเครื่องมือและเกณฑ์ เพื่อเป็นแนวทางในการนำไปปรับใช้ได้ตามความเหมาะสมกับบริบทของตนเองและนำไปสู่การพัฒนาผู้เรียนได้อย่างเต็มศักยภาพต่อไป

การศึกษาฐานสมรรถนะ

การศึกษาฐานสมรรถนะ หรือใช้ตัวย่อว่า CBE มีนิยามและการตีความที่หลากหลาย จึงทำให้การอธิบายการศึกษาฐานสมรรถนะ ใช้คำที่แตกต่างกันด้วย เช่น การเรียนรู้โดยใช้ปัญหาเป็นฐาน (problem-based learning) การเรียนรู้ที่มุ่งสู่ความเชี่ยวชาญ (mastery-based learning) การเรียนรู้เพื่อมุ่งผลลัพธ์ (outcome-based learning) และ การเรียนรู้ที่เน้นการปฏิบัติ (performance-based learning) โดยคำเหล่านี้ถูกนำมาประมวลและสร้างเป็นนิยามของ CBE ว่าเป็นแนวทางการศึกษาที่มุ่งผลลัพธ์ (outcome-based approach) ที่รวมการจัดการเรียนรู้และการประเมินความเชี่ยวชาญของผู้เรียน ในด้านความรู้ เจตคติ ค่านิยม ทักษะและพฤติกรรมเข้าไว้ด้วยกัน (Gervais, 2016)

CBE แตกต่างจากการจัดการศึกษาในชั้นเรียนแบบดั้งเดิมหลายประการ ประการแรก CBE ให้ความสำคัญกับสมรรถนะของผู้เรียนมากกว่าระดับผลการเรียนหรือจำนวนชั่วโมงเรียนรู้ตามที่กำหนดในหลักสูตร ประการที่สอง CBE เน้นเชื่อมโยงการเรียนรู้ในชั้นเรียนเข้ากับบริบทของโลกรจริงและการใช้ความรู้ในสถานการณ์ต่าง ๆ มากกว่าการท่องจำความรู้หรือความเข้าใจเนื้อหาวิชา จึงทำให้ ผู้เรียนเป็นผู้รับผิดชอบการเรียนรู้ของตนเอง ได้ฝึกปฏิบัติเพื่อสร้างความเข้าใจและสร้างความรู้ที่เกี่ยวข้องกับสิ่งที่ได้เรียนรู้ ประการสุดท้าย CBE เชื่อว่า เวลาที่ใช้เรียนรู้ไม่สะท้อนสมรรถนะของผู้เรียน เพราะผู้เรียนแต่ละคนสามารถใช้เวลาเรียนรู้เรื่องเดียวกันไม่เท่ากัน (McClarty & Gaertner, 2015; Pearson's CBE Solutions and Services, 2015; Gervais, 2016)

โมเดลของ CBE มีหลากหลาย โดยทั่วไปประกอบด้วย กรอบของสมรรถนะ (competency framework) และการประเมินสมรรถนะ (competency assessment) (McClarty & Gaertner, 2015) โดยการประเมินสมรรถนะเน้นการประเมินเพื่อพัฒนา (formative assessment) และเป็นการประเมินแบบอิงเกณฑ์ (criterion referenced) เพื่อบรรยายระดับความสามารถของผู้เรียนให้ชัดเจนมากขึ้น (Carraccio et al., 2002) การประเมินสมรรถนะจะสำเร็จเป็นอย่างดี จำเป็นอย่างยิ่งที่ครูต้องเข้าใจสมรรถนะ แนวทางการจัดการเรียนรู้ และขั้นตอนการประเมินสมรรถนะของผู้เรียนอย่างถูกต้อง (Mathelitsch, 2013; Gervais, 2016)

ความหมายและธรรมชาติของสมรรถนะ

สมรรถนะ หรือ competency และสามัตถิยะ หรือ competence มักใช้แทนกันในหลายบริบท อย่างไรก็ตาม Trotter & Ellison (1997) อธิบายความแตกต่างของสองคำนี้ กล่าวคือ competency และ competence เกี่ยวข้องกับทักษะ (skills) และคุณลักษณะ (characteristics) หลายอย่างที่บุคคลนำไปใช้แก้ปัญหาในสถานการณ์ต่าง ๆ แล้วเกิดผลสำเร็จ โดย competence คือ ความสามารถในการทำภารกิจที่เฉพาะเจาะจง (Davies & Ellison, 2003) เป็นความสามารถของบุคคลในการปฏิบัติงานเฉพาะอย่างให้ได้มาตรฐานตามที่ได้ตั้งไว้ โดยมักใช้ในบริบทของการประกอบอาชีพ ที่ต้องใช้ความรู้ ทักษะและเจตคติเฉพาะทาง (Guthrie, 2009) ส่วน competency เป็นคุณลักษณะของบุคคลที่แสดงออกในสถานการณ์หลากหลาย นอกจากนี้

พจนานุกรมศัพท์ศึกษาศาสตร์ ฉบับราชบัณฑิตยสถาน พ.ศ.2555 (Royal Academy, 2012) ให้ความหมายคำว่า สมรรถนะว่า คือ “คุณลักษณะและพฤติกรรมที่บ่งชี้ถึงความสามารถ ความชำนาญในการใช้ความรู้ ความเข้าใจ และทักษะที่มีอยู่อย่างเชี่ยวชาญ” ส่วน สามัตถิยะ คือ “ทักษะ ความสามารถของบุคคลที่นำความรู้ ความเข้าใจ มาแสดงออกหรือกระทำการอย่างใดอย่างหนึ่งได้ถึงระดับมาตรฐานที่กำหนด” ดังนั้น สองคำนี้มีความหมายใกล้เคียงกันในเชิงความสามารถและคุณลักษณะของบุคคล แต่ต่างกันตรงที่ competence เน้นความสามารถเฉพาะทางของบุคคลในการปฏิบัติงานหรือการประกอบอาชีพให้เป็นไปตามมาตรฐานหรือเกณฑ์ที่กำหนด ซึ่งเน้นบริบทที่เฉพาะเจาะจงกว่า competency ส่วนการมีสมรรถนะไม่จำเป็นต้องถึงระดับมาตรฐานที่กำหนด เพราะระดับของสมรรถนะของแต่ละบุคคลไม่จำเป็นต้องเป็นมาตรฐานเดียวกันก็ได้ ในทำนองเดียวกัน สมรรถนะของผู้เรียนแต่ละคนไม่จำเป็นต้องมีมาตรฐานเดียวกัน เพราะผู้เรียนแต่ละคนมีความหลากหลายและมีสมรรถนะเริ่มต้นในระดับที่ต่างกัน การพัฒนาสมรรถนะของผู้เรียนแต่ละคนจึงไม่จำเป็นต้องเท่ากัน

ธรรมชาติของสมรรถนะของผู้เรียนไม่มีความหมายครอบคลุมเพียงแค่ความรู้ (knowledge) หรือทักษะ (skills) (Mathelitsch, 2013; Pearson's CBE Solutions and Services, 2015; OECD, 2018) แต่รวมถึงความตระหนักรู้ (awareness) และการประยุกต์ใช้ความรู้ในบริบทหรือสถานการณ์ใหม่ ๆ ด้วย เนื่องจากสมรรถนะของผู้เรียนเป็นศูนย์กลางของระบบการศึกษานานาชาติ (Mathelitsch, 2013) จึงทำให้สมรรถนะต้องเฉพาะเจาะจงและเป็นสมรรถนะที่จำเป็นในสถานการณ์ที่เกิดขึ้นจริง โดยกำหนดให้การประยุกต์ใช้ความรู้เป็นผลลัพธ์ (outcome) หนึ่งของสมรรถนะ (Pearson's CBE Solutions and Services, 2015) นอกจากนี้ OECD (2018) ให้ความหมายของสมรรถนะ คือ การประสานระหว่างความรู้ ทักษะ เจตคติ (attitude) และค่านิยม (value) ของผู้เรียน เพื่อตอบสนองต่อความต้องการที่ซับซ้อน ผู้เรียนที่สามารถดำรงชีวิตในอนาคตจำเป็นต้องมีทั้งความรู้ทั่วไปและความรู้เฉพาะทางหลากหลาย เช่น การคิดอย่างมีวิจารณญาณ การคิดสร้างสรรค์ ความเห็นอกเห็นใจ การรับรู้ความสามารถของตนเอง การทำงานร่วมกัน (collaboration) และ การใช้อุปกรณ์ทางการสื่อสารและสารสนเทศใหม่ ๆ

สำหรับร่างกรอบหลักสูตรฐานสมรรถนะของประเทศไทย ได้ให้ความหมายของสมรรถนะและกำหนดสมรรถนะหลัก (core competencies) ดังนี้ (CBE Thailand, 2021b)

“สมรรถนะที่กำหนดให้เป็นพื้นฐานที่ผู้เรียนทุกคนต้องได้รับการพัฒนาให้เป็นความสามารถติดตัวเมื่อจบการศึกษา มีลักษณะเป็นสมรรถนะข้ามสาระการเรียนรู้หรือคร่อมวิชา สามารถพัฒนาให้เกิดขึ้นแก่ผู้เรียนได้ในสาระการเรียนรู้ต่าง ๆ ที่หลากหลาย หรือสามารถนำไปประยุกต์ใช้ในการพัฒนาผู้เรียนให้เรียนรู้สาระต่าง ๆ ได้ดีขึ้น...”

โดยสมรรถนะหลักต้องใช้ความรู้เชิงกระบวนการ (procedural knowledge) แบ่งเป็นสมรรถนะหกด้าน ได้แก่ สมรรถนะการจัดการตนเอง (self-management) การคิดขั้นสูง (higher order thinking) การสื่อสาร (communication) การรวมพลังทำงานเป็นทีม (teamwork collaboration) การเป็นพลเมืองที่เข้มแข็ง (civic

literacy) และ การอยู่ร่วมกับธรรมชาติและวิทยาการอย่างยั่งยืน (sustainable coexistence with nature and science)

จะเห็นได้ว่า สมรรถนะเป็นความสามารถ คุณลักษณะและพฤติกรรมของแต่ละบุคคลที่แสดงออก ในสถานการณ์ที่หลากหลาย เป็นการร้อยเรียงกันระหว่างความรู้ ทักษะ ค่านิยมและคุณลักษณะอื่น ๆ ที่ผู้เรียน พึงมีเพื่อสะท้อนถึงผลลัพธ์การเรียนรู้ ไม่ขึ้นกับเนื้อหารายวิชาและใช้ความรู้เชิงกระบวนการเพื่อทำให้ผู้เรียนบรรลุ เป้าหมายการเรียนรู้ที่ตั้งไว้

การประเมินสมรรถนะของผู้เรียน

การประเมินสมรรถนะของผู้เรียน หรือการประเมินฐานสมรรถนะ (competency-based assessment: CBA) ถือเป็นองค์ประกอบสำคัญในการศึกษาฐานสมรรถนะ โดย CBA เชื่อมโยงกับแนวทางการสอน (pedagogical approaches) ที่หลากหลาย เน้นประเมินการปฏิบัติงาน (performance assessment) และ ประเมินตามสภาพจริง (authentic assessment) (Gallardo, 2020) CBA ที่มีประสิทธิภาพต้องสามารถวัดได้ทั้ง ความรู้ ทักษะและความสามารถในการบูรณาการ สังเคราะห์และประยุกต์ใช้ความรู้เพื่ออธิบายและแก้ปัญหา ในบริบทของโลกรจริงหรือสถานการณ์ในชีวิตประจำวัน (Drisko, 2014; Pearson's CBE Solutions and Services, 2015, Gervais, 2016) ผู้เรียนต้องสามารถถ่ายทอดความรู้และใช้ทักษะเพื่อแก้ปัญหา หรือเสนอ แนวคิดใหม่เพื่อแก้ปัญหาจากสถานการณ์ที่กำหนดได้ มุ่งเน้นกระบวนการเรียนรู้ (learning process) ของผู้เรียน มากกว่าผลงานหรือชิ้นงาน (Gallardo, 2020) โดยเป็นการประเมินรายบุคคลที่ไม่เปรียบเทียบกับผู้อื่น การประเมินสามารถเกิดขึ้นได้ทั้งก่อน ระหว่างและหลังเรียน ซึ่งจะช่วยให้ผู้สอนรู้ความต้องการของผู้เรียน รู้สิ่งที่ ต้องเพิ่มเติมเฉพาะเจาะจงกับผู้เรียนแต่ละคน นำไปสู่การปรับเปลี่ยนกิจกรรมการเรียนรู้และเวลา Gervais (2016) เสนอว่าการประเมินที่หลากหลาย มีบทบาทสำคัญต่อกระบวนการเรียนรู้ของผู้เรียน ซึ่งควรประเมินเพื่อพัฒนา (formative assessment) เน้นการให้ผลสะท้อนกลับ (feedback) แก่ผู้เรียนเพื่อปรับปรุงการเรียนรู้ของตนเอง รวมถึงการจัดการเรียนรู้ของผู้สอนเพื่อให้เหมาะสมกับการพัฒนาผู้เรียนด้วย ส่วนการประเมินเพื่อสรุปการเรียนรู้ (summative assessments) เป็นการแสดงให้เห็นถึงระดับสมรรถนะของผู้เรียนเมื่อเสร็จสิ้นการเรียนรู้ โดยผู้เรียน ควรได้รับการประเมินเมื่อพร้อมและควรเป็นการประเมินหลายครั้ง ด้วยเครื่องมือที่หลากหลาย เพื่อให้ผู้เรียน แสดงออกถึงสมรรถนะต่าง ๆ ได้อย่างเต็มที่ (Drisko, 2014; Gervais, 2016)

เครื่องมือประเมินสมรรถนะต้องสามารถวัดได้ตรงกับสิ่งที่ต้องการวัด โดยความรู้สามารถวัดได้ด้วย แบบทดสอบทั้งแบบเลือกตอบและเขียนตอบ แต่ผู้เรียนต้องสามารถสะท้อนพฤติกรรมตามสภาพจริงและควรเป็น การประเมินการปฏิบัติ (Drisko, 2014; Pearson's CBE Solutions and Services, 2015; Gervais, 2016) เครื่องมือที่ประเมินหลาย ๆ สมรรถนะพร้อมกันในบริบทโลกรจริง จะสะท้อนสมรรถนะของผู้เรียนได้ดีกว่าเครื่องมือ ที่ประเมินเพียงสมรรถนะเดียว (Drisko, 2014) และมักจะเป็นเครื่องมือที่ไม่มีคำตอบที่ถูกต้องหรือผิด (Pearson's

CBE Solutions and Services, 2015) การพัฒนาเครื่องมือประเมินสมรรถนะ มีขั้นตอนที่สำคัญสองประการคือ (1) การพัฒนาการประเมินและการตรวจสอบความถูกต้องของคะแนน เพื่อให้เครื่องมือที่สร้างขึ้นสะท้อนระดับความรู้ และทักษะที่สนใจ และ (2) การกำหนดเกณฑ์สมรรถนะที่แยกผู้เรียนที่มีสมรรถนะออกจากไม่มีสมรรถนะได้ โดยอาจใช้เป็นคะแนนจุดตัด (cut scores) เข้ามาช่วยได้ (McClarty & Gaertner, 2015) โดยการประเมินโดยใช้คะแนนจุดตัด จะสามารถจำแนกผู้เรียนกลุ่มเชี่ยวชาญและไม่เชี่ยวชาญออกจากกันได้ ซึ่งมีความสำคัญต่อการตีความและแปลผลการประเมินสมรรถนะ

ครูต้องสร้างเครื่องมือประเมินสมรรถนะอย่างระมัดระวัง โดย McClarty & Gaertner (2015) ได้เสนอว่า CBE ควรกำหนดนิยามของสมรรถนะที่ต้องการส่งเสริมผู้เรียนให้ชัดเจน เชื่อมโยงสมรรถนะที่จะวัดกับการประเมิน เพื่อแสดงให้เห็นว่า พฤติกรรมของนักเรียนหรือกระบวนการคิดสะท้อนสมรรถนะของผู้เรียน เชื่อมโยงการปฏิบัติของผู้เรียนเข้ากับการประเมินสมรรถนะและเทียบเคียงกับเครื่องมือวัดอื่น ๆ มีการติดตามความก้าวหน้าหรือพัฒนาการของผู้เรียน และบอกความสัมพันธ์ระหว่างผลประเมินและผลลัพธ์ที่จะเกิดขึ้นกับผู้เรียนในอนาคต ซึ่งเครื่องมือประเมินที่ดีสามารถวัดสมรรถนะที่ต้องการได้ เป็นการประเมินตามสภาพจริง และเป็นปรนัย (McClarty & Gaertner, 2015; Pearson's CBE Solutions and Services, 2015) กล่าวคือ ผู้เรียนมีผลประเมินสอดคล้องกันหรืออยู่ในระดับสมรรถนะเดียวกัน แม้ว่าจะถูกประเมินจากเครื่องมืออื่น ๆ หรือผู้ประเมินคนอื่น ๆ ก็ตาม

รูบริก (rubrics) สามารถเป็นเครื่องมือ เพื่อทำให้การวัดและประเมินสมรรถนะเป็นไปอย่างถูกต้องและน่าเชื่อถือ (Ab Rahman et al., 2014; Pearson's CBE Solutions and Services, 2015; Ana et al., 2020; Gallardo, 2020) รูบริกเป็นเกณฑ์ที่ผู้สอนกำหนดขึ้นเพื่อให้คะแนนชิ้นงานของผู้เรียน เป็นแนวทางให้ผู้สอนใช้ประเมินความรู้ ทักษะ คุณลักษณะ พฤติกรรมหรือผลงาน (Ritcharoon, 2019) รูบริกที่มีคุณภาพสูง จะประเมินผลการปฏิบัติ เพื่อสะท้อนสมรรถนะได้อย่างแม่นยำ สามารถจำแนกผู้เรียนตามระดับสมรรถนะได้ และมีคำอธิบายชัดเจน ผู้ใช้รูบริกเข้าใจความหมายตรงกัน (Jonsson, 2014; Pearson's CBE Solutions and Services, 2015; Virk et al., 2020) รูบริกที่ดีต้องทำให้ผู้เรียนทราบว่า ตนเองมีสมรรถนะอยู่ในระดับใด และทำให้ผู้สอนรู้ว่า จะพัฒนาให้ผู้เรียนมีสมรรถนะสูงขึ้นได้อย่างไร (Pearson's CBE Solutions and Services, 2015)

การให้ผลสะท้อนกลับกับผู้เรียน สามารถทำได้ทั้งระหว่างเรียน และเมื่อเสร็จสิ้นการเรียนรู้ เพื่อให้ผู้เรียนเรียนรู้และพัฒนาการเรียนรู้ของตนเอง (Ab Rahman et al., 2014; Motsenbocker, 2018; Gallardo, 2020; Ilma et al., 2021) จากการศึกษาค้นคว้าของ Ilma et al. (2021) พบว่า การให้ผลสะท้อนกลับแก่ผู้เรียน ช่วยให้ผู้เรียนได้เข้าใจแนวคิดที่ครูผู้สอนพยายามถ่ายทอด ส่งเสริมให้ผู้เรียนมีความเข้าใจความรู้และมีแนวทางในการปรับปรุงกระบวนการเรียนรู้ของตนเองที่ชัดเจน ทราบข้อผิดพลาดของตนเอง ช่วยให้ผู้เรียนมีแรงจูงใจและความมั่นใจในการที่จะเรียนรู้ต่อ นอกจากนี้ผลสะท้อนกลับยังช่วยแก้ไขความเข้าใจที่คลาดเคลื่อนของผู้เรียนได้อีกด้วยการให้ผลสะท้อนกลับที่ดีที่สุดคือ การให้ผลสะท้อนกลับแบบมีปฏิสัมพันธ์ (interactive feedback) โดยใช้การพูดกับผู้เรียนโดยตรง ไม่ว่าจะมาจากครูหรือเพื่อนร่วมชั้นเรียน

จากการวิเคราะห์ข้างต้น สามารถสรุปลักษณะสำคัญของ CBA ได้แก่ (1) เน้นประเมินการปฏิบัติงานและการประเมินตามสภาพจริง (2) ใช้เครื่องมือที่มีคุณภาพ สามารถวัดสมรรถนะได้หลากหลาย มีรูปวิธีที่ดี มีความเป็นปรนัย (3) ให้สารสนเทศเกี่ยวกับความสามารถในการแก้ปัญหาที่สัมพันธ์กับบริบทโลกจริงของผู้เรียนและสามารถติดตามการเปลี่ยนแปลงสมรรถนะของผู้เรียนได้ (4) เป็นการประเมินรายบุคคล โดยไม่นำผลการประเมินไปเปรียบเทียบกับผู้เรียนคนอื่น สามารถทำได้ทั้งก่อน ระหว่างและหลังเรียน และ (5) ให้ผลสะท้อนกลับระหว่างเรียนและเมื่อเสร็จสิ้นการเรียนรู้เพื่อเป็นแนวทางสำคัญพัฒนาการเรียนรู้ของผู้เรียน

การประเมินสมรรถนะของผู้เรียนในการเรียนวิทยาศาสตร์

การประเมินผู้เรียนในการเรียนวิทยาศาสตร์นั้นควรสอดคล้องกับลักษณะสำคัญของ CBA ข้างต้น อย่างไรก็ตาม จากการศึกษาของ Ab Rahman et al. (2014) พบว่า การประเมินผู้เรียนในการเรียนวิทยาศาสตร์ส่วนใหญ่ยังคงใช้เครื่องมือแบบเดิม คือ เน้นประเมินความรู้มากกว่าประเมินการปฏิบัติและประเมินตามสภาพจริง การออกแบบใบงานปฏิบัติการทางวิทยาศาสตร์ยังขาดการตรวจสอบคุณภาพของใบงานเหล่านั้น กำหนดเกณฑ์หรือมาตรฐานสมรรถนะของผู้เรียนไม่ชัดเจน โดยส่วนใหญ่เป็นเกณฑ์แบบองค์รวม ทำให้ได้สารสนเทศในภาพรวมและส่งผลกระทบต่อความน่าเชื่อถือในการตัดสินความสามารถของผู้เรียน การให้คะแนนตามรูปวิธีมีขอบเขตเนื้อหากว้างเกินไปและไม่สามารถให้ผลสะท้อนกลับกับผู้เรียน ด้วยปัญหาดังกล่าวจึงนำไปสู่การพัฒนาเครื่องมือที่มีความน่าเชื่อถือเพิ่มมากขึ้นเพื่อนำไปใช้ตัดสินระดับสมรรถนะของผู้เรียนได้อย่างเป็นธรรม และผู้เรียนสามารถประเมินตนเองได้ (McClarty & Gaertner, 2015)

Mathelitsch (2013) วิเคราะห์ลักษณะการประเมินสมรรถนะในรายวิชาวิทยาศาสตร์ของประเทศเยอรมนี สวิตเซอร์แลนด์ และออสเตรเลีย พบว่า มีความคล้ายคลึงกัน กล่าวคือ เป็นการประเมินในสามมิติ ได้แก่ มิติของเนื้อหา (contents) มิติของระดับ (levels) และมิติของทักษะ (skills) โดยเฉพาะมิติของระดับมีการแบ่งระดับของความเข้าใจเป็นสามระดับได้แก่ ทำตาม (reproduction) เข้าใจและประยุกต์ใช้ (understanding and application) และ ถ่ายทอด (transfer) โดยผู้เรียนในระดับสุดท้ายสามารถเชื่อมโยงความรู้ทางวิทยาศาสตร์กับสถานการณ์จริงได้ ส่วนมิติของทักษะและมิติของเนื้อหา มีการกำหนดทักษะที่แตกต่างกันในแต่ละประเทศ

สำหรับประเทศไทยเองนั้น ได้กำหนดระดับสมรรถนะหลัก 4 ระดับ ได้แก่ ระดับเริ่มต้น กำลังพัฒนา สามารถและเหนือความคาดหวัง และมีคำบรรยายสมรรถนะที่พึงมีในตามระดับชั้น (CBE Thailand, 2021b) เพื่อให้ผู้สอนสามารถนำไปประยุกต์ใช้ในการเรียนการสอนและการประเมินผู้เรียนของตนเองได้

Ilma et al. (2021) ได้พัฒนาโมเดลการประเมินเพื่อการเรียนรู้ (Assessment for Learning Model หรือ AfL model) สำหรับใช้ประเมินสมรรถนะสำคัญในศตวรรษที่ 21 ในรายวิชาฟิสิกส์ เมื่อพิจารณาโมเดลดังกล่าว ได้เสนอกรอบสมรรถนะและการประเมินสมรรถนะ ซึ่งสอดคล้องกับโมเดลของ CBE ที่เสนอโดย McClarty & Gaertner (2015) ดังตาราง 1

ตาราง 1 กรอบสมรรถนะและการประเมินสมรรถนะสำคัญในศตวรรษที่ 21 ในรายวิชาวิทยาศาสตร์

กรอบสมรรถนะ	วิธีประเมินสมรรถนะ	คำบรรยายลักษณะงาน
การคิดสร้างสรรค์	การใช้ปัญหา (giving problems)	มอบหมายงานในรูปแบบของปัญหาที่ต้องการให้ผู้เรียนแก้ไข ซึ่งปัญหานั้นต้องสะท้อนความคิดสร้างสรรค์เพื่อให้ได้วิธีการหรือแนวทางในการแก้ปัญหาหรือสร้างชิ้นงานใหม่ ๆ สำหรับปัญหานั้น
	การเขียนชิ้นงาน (making a paper)	มอบหมายงานในรูปแบบของงานเขียน เรียงความ บทความ หรือเอกสาร ที่ผู้เรียนต้องใช้ความคิดสร้างสรรค์เพื่อเสนอมุมมองในงานเขียนของตนเอง การเขียนชิ้นงานจะสะท้อนกระบวนการเรียนรู้ของผู้เรียนบนพื้นฐานของความรู้ทางวิทยาศาสตร์ที่ถูกต้อง ซึ่งถือเป็นการพัฒนาด้านพุทธิพิสัย (cognitive domain) จิตพิสัย (affective domain) และทักษะพิสัย (psychomotor domain) ไปพร้อมกันด้วย
	การสร้างแผนผังมโนทัศน์ (concept map creation)	ผู้เรียนออกแบบแผนผังมโนทัศน์ของตน ซึ่งต้องใช้ความคิดสร้างสรรค์และสามารถใช้โปรแกรมหรือแอปพลิเคชันช่วยออกแบบได้ ผู้สอนประเมินความถูกต้องของมโนทัศน์และความคิดสร้างสรรค์ในการออกแบบ และให้ผลสะท้อนกลับเพื่อแก้ไขแนวคิดที่คลาดเคลื่อนของผู้เรียน
	การทำโครงการ/โครงการ (project creation)	โครงการหรือโครงการอาจอยู่ในรูปแบบของการทำวิดีโอ หน่วย/ชุดการเรียนรู้ สื่อการเรียนรู้ นวัตกรรมทางวิทยาศาสตร์ต่าง ๆ และผู้สอนทำการประเมินโดยมีการกำหนดเกณฑ์ให้คะแนนเป็นส่วน ๆ เช่น คุณภาพผลงาน เนื้อหา โครงสร้าง เป็นต้น
	การทำโปสเตอร์ (making a poster)	โปสเตอร์เป็นผลงานที่ผู้เรียนต้องคำนึงหลายอย่าง ได้แก่ เป้าหมาย ข้อความที่ใช้ซึ่งต้องสั้น ตรงประเด็น ชัดเจนและสามารถโน้มน้าวผู้คนได้ ทำให้เป็นผลงานที่ใช้ประเมินได้หลายด้าน ทั้งด้านสุนทรียศาสตร์ (aesthetic aspects) เนื้อหา และภาษา
การสื่อสาร	การอภิปรายในชั้นเรียน (class discussion)	การอภิปรายในชั้นเรียนทำให้ผู้เรียนเกิดสมรรถนะด้านสื่อสาร เกิดความกระตือรือร้นในการเตรียมข้อมูล ผู้เรียนที่มีความกระตือรือร้นและมีวิจรรย์ญาณมากจะยิ่งอภิปรายในชั้นเรียนได้ดี
	การมอบหมายโครงการ/โครงการ (project assignment)	การให้ผู้เรียนได้ทำโครงการ/โครงการ ทำให้ผู้เรียนได้เกิดการถกเถียงโต้แย้ง แลกเปลี่ยนความคิด ทำให้เกิดการสื่อสารระหว่างผู้เรียน
	การใช้คำถาม (questioning)	ผู้สอนใช้คำถามที่เกี่ยวข้องกับเนื้อหาที่เรียนกระตุ้นให้ผู้เรียนตอบ อภิปราย และแสดงถึงการมีส่วนร่วมในชั้นเรียน
	การโต้เถียงออนไลน์ (online debate model)	การโต้เถียงที่ทำให้ผู้เรียนแสดงความคิดเห็น เกิดการอภิปรายในประเด็นที่กำหนด
การคิดอย่างมีวิจรรย์ญาณ	การทดสอบย่อย (giving quizzes)	การทดสอบย่อยอาจทำในรูปแบบการถามคำถามให้ผู้เรียนได้คิดอย่างไตร่ตรองก่อนที่จะตอบ สามารถใช้เมื่อเสร็จสิ้นการบรรยาย เมื่อผู้เรียนทำการทดสอบย่อยแล้ว ผู้สอนทำการเฉลย ให้คะแนน และให้ผลสะท้อนกลับต่อคำตอบของผู้เรียน ทำให้เกิดการสื่อสารระหว่างผู้เรียนและผู้สอนไปพร้อมกันด้วย
	การใช้คำถามขั้นสูง (giving high order thinking skills (HOTS) Questions)	คำถามขั้นสูงมีความสำคัญมากต่อการเรียนรู้วิทยาศาสตร์ เนื่องจากผู้เรียนต้องแก้ปัญหาในบริบทโลกจริง ดังนั้น ผู้เรียนจึงต้องการมากกว่าความรู้ความจำ แต่ต้องใช้ทั้งการคิดอย่างมีวิจรรย์ญาณ การแก้ปัญหาและการคิดขั้นสูงอื่น ๆ เพื่อตอบคำถามขั้นสูง

ตาราง 1 (ต่อ)

กรอบสมรรถนะ	วิธีประเมินสมรรถนะ	คำบรรยายลักษณะงาน
การทำงานแบบ ร่วมมือ	การอภิปรายปัญหา (discussion of problems)	ผู้เรียนสามารถแลกเปลี่ยนความคิด และได้แย้งต่อความคิดเห็นหรือข้อเสนอแนะที่เกี่ยวข้องกับปัญหา
	การเขียนชิ้นงาน (making a paper)	การเขียนต้องอาศัยการคิดอย่างมีวิจารณญาณเป็นอย่างมากเพื่อสังเกตปรากฏการณ์ทางวิทยาศาสตร์ต่าง ๆ และต้องเขียนอธิบายอย่างเป็นเหตุเป็นผล
	การอภิปราย (discussion)	ทั้งผู้เรียนและผู้สอนสามารถตั้งประเด็นการอภิปราย ใช้คำถามกระตุ้นการเรียนรู้และให้ผลสะท้อนกลับเพื่อให้กลุ่มช่วยกันหาแนวทางในการแก้ปัญหาที่ดีที่สุด
	การทำงานกลุ่มโครงการ/โครงการ (group assignments in the form of projects)	สมาชิกในกลุ่มทำงานร่วมกันเพื่อแก้ปัญหา โครงการ หรือโครงการอาจอยู่ในรูปแบบที่หลากหลาย เช่น การสร้างวิดีโอ สื่อการเรียนรู้หรือนวัตกรรมที่ต้องอาศัยความร่วมมือของกลุ่ม จึงจะประสบความสำเร็จ
	การนำเสนอแบบกลุ่ม (group presentations)	การนำเสนอผลงานช่วยสะท้อนการทำงานเป็นกลุ่มแบบร่วมมือของผู้เรียนได้ โดยผู้สอนให้ข้อเสนอแนะ ผลสะท้อนกลับผลงานของผู้เรียนเพื่อให้กลุ่มสามารถปรับปรุงผลงานให้ดีขึ้น

เมื่อพิจารณาสมรรถนะสำคัญในศตวรรษที่ 21 ตามกรอบสมรรถนะของ Ilma et al. (2021) เปรียบเทียบกับสมรรถนะหลักของ (ร่าง) กรอบหลักสูตรการศึกษาขั้นพื้นฐานของประเทศไทย (CBE Thailand, 2021b) พบว่า เป็นชุดสมรรถนะที่มีความสอดคล้องกันเป็นอย่างมาก และจะเห็นได้ว่ามีแนวทางและวิธีการที่หลากหลายที่สามารถประเมินสมรรถนะผู้เรียนในการเรียนวิทยาศาสตร์ โดยส่วนใหญ่เป็นแนวทางที่ครูวิทยาศาสตร์ใช้อยู่ในชั้นเรียนของตน แต่ในการนำไปปฏิบัติจริงอาจจะยังไม่ปรากฏเครื่องมือประเมินและรูปรีดที่ชัดเจน อีกทั้งครูวิทยาศาสตร์อาจจะยังขาดความรู้ในการสร้างเครื่องมือและเกณฑ์ที่จะสามารถใช้ในการประเมินสมรรถนะของผู้เรียน รวมถึงการตรวจสอบคุณภาพของเครื่องมือ ประเด็นเหล่านี้เป็นความจำเป็นที่ต้องเตรียมบุคลากรให้มีความรู้ความเข้าใจเพียงพอเพื่อนำไปใช้ปฏิบัติงาน ผู้เขียนจึงขอเสนอวิธีการออกแบบการประเมินสมรรถนะของผู้เรียนในชั้นเรียนวิทยาศาสตร์ พร้อมทั้งยกตัวอย่างเครื่องมือและเกณฑ์ในการประเมินสมรรถนะของผู้เรียนในบางสมรรถนะเพื่อเป็นทางเลือกหนึ่งให้ครูวิทยาศาสตร์และผู้สนใจสามารถนำไปปรับใช้ในการเรียนการสอนของตนเอง

การออกแบบการประเมินสมรรถนะของผู้เรียนในชั้นเรียนวิทยาศาสตร์

ผู้เขียนเสนอขั้นตอนการออกแบบการประเมินสมรรถนะของผู้เรียนในชั้นเรียนวิทยาศาสตร์ ดังนี้

ขั้นตอนที่ 1 การระบุสมรรถนะ เป็นขั้นตอนที่ผู้สอนกำหนดสมรรถนะ องค์ประกอบและพฤติกรรมที่ต้องการประเมินผู้เรียนให้ชัดเจน โดยองค์ประกอบต้องครอบคลุมพฤติกรรมตามผลลัพธ์การเรียนรู้ และกระบวนการเรียนรู้ ดังตัวอย่างต่อไปนี้

ในการประเมินสมรรถนะการสื่อสารด้านการเขียนในรายวิชาวิทยาศาสตร์ระดับมัธยมศึกษาตอนต้น ผู้เขียนได้กำหนดนิยามของสมรรถนะการสื่อสารด้านการเขียน หมายถึง ความสามารถในการส่งสารผ่านการเล่าเรื่องสรุปความ ตีความและนำเสนอจากประเด็นทางวิทยาศาสตร์หรือที่เกี่ยวข้องกับวิทยาศาสตร์ แบ่งออกเป็น

สิ่งประกอบ ได้แก่ (1) ความสามารถในการเล่าเรื่อง (2) ความสามารถในการสรุปความ (3) ความสามารถในการตีความ และ (4) ความสามารถในการนำเสนอข้อมูล

ขั้นตอนที่ 2 การเลือกวิธีการและเครื่องมือประเมินสมรรถนะ เป็นขั้นตอนที่ผู้สอนพิจารณาเลือกวิธีการและเครื่องมือประเมินที่ใช้สอดคล้องกับองค์ประกอบและพฤติกรรมตามสมรรถนะที่ตั้งใจ โดยอาจใช้วิธีการและเครื่องมือที่นักเรียนได้เรียนได้ตามความเหมาะสม ดังตัวอย่างต่อไปนี้

ผู้สอนอาจเลือกวิธีการประเมินก่อนเรียน เพื่อวินิจฉัยสมรรถนะของผู้เรียนก่อนการเรียนรู้และนำผลการประเมินไปใช้วางแผนจัดการเรียนรู้ให้กับผู้เรียนโดยใช้เครื่องมือเป็นแบบวัดสมรรถนะการสื่อสารด้านการเขียนตามองค์ประกอบข้างต้น สำหรับการประเมินระหว่างเรียนผู้สอนอาจเลือกใช้การสังเกตพฤติกรรมการเรียนรู้ของผู้เรียน อาจจัดกิจกรรมการเรียนรู้หรือมอบหมายงานในรูปแบบของการเขียน เช่น ให้เขียนเล่าเรื่องจากภาพหรือสถานการณ์ทางวิทยาศาสตร์ เขียนสรุปความและตีความจากผลการทดลองทางวิทยาศาสตร์ เป็นต้น ส่วนวิธีการประเมินหลังเรียน ผู้สอนอาจเลือกวิธีการประเมินผู้เรียนด้วยแบบวัดอีกครั้งหรือประเมินจากการนำเสนอข้อมูลผลการทดลองทางวิทยาศาสตร์และการตอบคำถามจากใบงาน

ขั้นตอนที่ 3 การสร้างเครื่องมือประเมินและรูบรีค เป็นขั้นตอนที่ผู้สอนออกแบบเครื่องมือที่นำมาใช้ในการประเมินและเกณฑ์การให้คะแนนแบบรูบรีคตามองค์ประกอบที่กำหนดข้างต้น โดยผู้สอนอาจเลือกใช้เป็นเกณฑ์การให้คะแนนแบบองค์รวม (holistic rubrics) เป็นการให้คะแนนสิ่งที่ต้องการประเมินในภาพรวมของหรือเกณฑ์การให้คะแนนแบบวิเคราะห์ (analytic rubrics) เป็นการให้คะแนนสิ่งที่ต้องการประเมินแยกตามองค์ประกอบที่ต้องการประเมิน (Royal Academy, 2012; Nitko, 2019)

ดังตัวอย่างต่อไปนี้

ในการประเมินสมรรถนะการสื่อสารด้านการเขียนเพื่อวินิจฉัยสมรรถนะของผู้เรียนก่อนเรียน ผู้เขียนขอยกตัวอย่างแบบวัดสมรรถนะการสื่อสารด้านการเขียน ซึ่งผู้สอนสามารถนำไปปรับใช้เป็นแบบฝึกหัดหรือกิจกรรมการเรียนรู้เพื่อประเมินผู้เรียนระหว่างเรียนได้ด้วย ตามองค์ประกอบ ดังนี้

องค์ประกอบที่ 1 ความสามารถในการเล่าเรื่อง ผู้เรียนแสดงพฤติกรรมการเรียนรู้โดยเขียนเล่าเรื่องจากภาพสถานการณ์ทางวิทยาศาสตร์หรือเกี่ยวข้องกับวิทยาศาสตร์ ได้แก่ เล่าเรื่องได้ตรงประเด็น ถูกต้องตามแนวคิดทางวิทยาศาสตร์และใช้ภาษาได้ถูกต้อง ผู้สอนออกแบบข้อคำถามโดยอาจเลือกกำหนดภาพสถานการณ์จากเนื้อหาที่นักเรียนได้เคยเรียนมาแล้วในระดับชั้นอื่น เพื่อให้ผู้เรียนได้เขียนเล่าเรื่องจากภาพสถานการณ์ที่กำหนดให้ ดังตัวอย่างคำถามต่อไปนี้

ตัวอย่างคำถามสมรรถนะการสื่อสารด้านการเขียน

องค์ประกอบที่ 1 ความสามารถในการเล่าเรื่อง

คำถาม จากภาพสถานการณ์ ให้นักเรียนเขียนเล่าเรื่องจากภาพสถานการณ์ที่กำหนดให้



เขียนเล่าเรื่องจากภาพ.....
.....
.....

ที่มาภาพ: <https://www.thinglink.com/scene/575774465934229505>

องค์ประกอบที่ 2 ความสามารถในการสรุปความ ผู้เรียนแสดงพฤติกรรมการเรียนรู้ โดยเขียนสรุปใจความสำคัญจากบทความทางวิทยาศาสตร์หรือที่เกี่ยวข้องกับวิทยาศาสตร์ ได้แก่ สรุปเนื้อหาจากการอ่านได้ตรงประเด็น และใช้ภาษาได้ถูกต้อง และองค์ประกอบที่ 3 ความสามารถในการตีความ ผู้เรียนแสดงพฤติกรรมการเรียนรู้โดยเขียนตีความเจตนาของผู้เขียนบทความทางวิทยาศาสตร์หรือที่เกี่ยวข้องกับวิทยาศาสตร์ ได้แก่ ตีความเจตนาของผู้เขียนบทความตามความคิดของตนเองได้ชัดเจนและใช้ภาษาได้ถูกต้อง ผู้สอนออกแบบข้อคำถามโดยให้ผู้เรียนเขียนสรุปใจความสำคัญและตีความจากบทความที่กำหนดให้ ดังตัวอย่างคำถามต่อไปนี้

ตัวอย่างคำถามสมรรถนะการสื่อสารด้านการเขียน

องค์ประกอบที่ 2 ความสามารถในการสรุปความและองค์ประกอบที่ 3 ความสามารถในการตีความ

คำถาม ให้นักเรียนเขียนสรุปใจความสำคัญและตีความจากบทความที่กำหนดให้ต่อไปนี้

นิโคตินจากบุหรี่สามารถเข้าสู่ร่างกายได้ผ่านผิวหนัง

ผลการศึกษาพบว่า นิโคติน สารเคมีในควันบุหรี่สามารถเข้าสู่ร่างกายทางอื่น ๆ ได้ นอกจากการหายใจ โดยนิโคตินสามารถผ่านเข้าไปทางผิวหนังและเข้าสู่กระแสเลือดจากอากาศหรือเสื้อผ้าที่สัมผัสกับควันบุหรี่ นักวิทยาศาสตร์ได้กล่าวว่าอนุภาคของสารที่ออกมาจากควันบุหรี่จะปนอยู่ในอากาศเป็นเวลานานกว่าชั่วโมง

การค้นพบนี้มีความสำคัญอย่างมากกับกลุ่มเด็กและวัยรุ่น นิโคตินนั้นส่งผลกระทบต่อสมองซึ่ง Charles Weschle นักเคมีจาก Rutgers University ได้กล่าวว่า "ถ้าเราอยู่ในห้องที่มีการสูบบุหรี่ เราจะได้รับควันบุหรี่เหล่านั้นทางผิวหนังรวมถึงทางปอดด้วย" ดังนั้น แผ่นแปะนิโคตินจึงได้ถูกออกแบบให้ส่งสารเคมีนี้ไปทางผิวหนัง ซึ่งเป้าหมายคือต้องการที่จะช่วยให้คนที่ติดบุหรี่นั้นเลิกสูบบุหรี่ แต่การสัมผัสกับนิโคตินทางผิวหนังสามารถก่อให้เกิดอาการป่วยหรือเสียชีวิตได้อีกด้วยถ้าสัมผัสกับมันมากเกินไป เนื่องจากสารเคมีชนิดนี้เป็นพิษ ซึ่งถูกใช้เป็นส่วนผสมในยาแก้ปวดที่รุนแรง

จากการทดลองโดยให้นักวิจัยที่ไม่สูบบุหรี่จำนวนสองคนใช้เวลาสามชั่วโมงในห้องปิดที่เต็มไปด้วยควันบุหรี่ โดยปริมาณของควันบุหรี่ยังเทียบเท่าสถานบันเทิงที่มีการสูบบุหรี่เป็นจำนวนมาก นิโคตินจะเริ่มแตกตัวเมื่อสัมผัสกับเลือด การแตกตัวของสารเคมีนี้จะออกไปยังปัสสาวะ ซึ่งจากการทดลอง พบว่า ปัสสาวะจากนักวิจัยทั้งสองคนแสดงให้เห็นถึงปริมาณของนิโคตินที่เพิ่มขึ้นโดยได้รับการดูดซับของนิโคตินประมาณ 570 ไมโครกรัมผ่านทางผิวหนัง ซึ่งมีค่ามากกว่ากับคนที่สูบบุหรี่ระหว่าง 0.5 ถึง 6 มวน นอกจากนี้ นักวิจัยต้องการตรวจสอบว่าเสื้อที่แขวนอยู่นั้นอาจจะผลการสัมผัสนิโคตินอีกด้วย ดังนั้นจึงทำการแขวนเสื้อยืดเป็นเวลา 5 วันในห้องที่มีการสูบบุหรี่ หลังจากนั้นให้นักวิจัยคนหนึ่งที่ไม่ได้สูบบุหรี่ทำการสวมใส่เสื้อยืดนั้น นักวิทยาศาสตร์ค้นพบว่าเขามีปริมาณของนิโคตินถึงสี่เท่าในเลือดหลังจากที่สวมใส่เสื้อยืดที่มีการสัมผัสกับควันบุหรี่

ที่มา: Journal: G. Bekö et al. Measurements of dermal uptake of nicotine directly from air and clothing. Indoor Air. Published online August 24, 2016. doi: 10.1111/ina.12327.

ดัดแปลงจาก: <http://www.vcharkarn.com/vnews/505756>

1.จากบทความสามารถสรุปใจความสำคัญได้อย่างไร

.....
.....

2. จากบทความ นักเรียนคิดว่าผู้เขียนมีเจตนาอย่างไรในการเสนอบทความเรื่องนี้

.....
.....

องค์ประกอบที่ 4 ความสามารถในการนำเสนอข้อมูล ผู้เรียนแสดงพฤติกรรมการเรียนรู้ โดยเขียนนำเสนอจากข้อมูลทางวิทยาศาสตร์หรือที่เกี่ยวข้องกับวิทยาศาสตร์ เช่น ผลการทดลอง หรือข้อมูลจากฐานข้อมูลของหน่วยงานต่าง ๆ พฤติกรรมที่ประเมินได้แก่ จัดกระทำข้อมูลเพื่อนำเสนอในรูปแบบใดรูปแบบหนึ่ง (เช่น แผนภูมิ กราฟ บรรยาย ฯลฯ) ได้อย่างถูกต้องและสื่อความชัดเจน ดังตัวอย่างคำถามต่อไปนี้

ตัวอย่างคำถามสมรรถนะการสื่อสารด้านการเขียน

องค์ประกอบที่ 4 ความสามารถในการนำเสนอข้อมูล

คำถาม จากข้อมูลสัดส่วนการใช้เชื้อเพลิงการผลิตพลังงานไฟฟ้าของประเทศไทยในปี พ.ศ. 2564 หากนักเรียนต้องนำเสนอข้อมูลดังกล่าวในที่ประชุมวิชาการด้านพลังงานไฟฟ้า ให้นักเรียนจัดระบบข้อมูลเพื่อใช้ในการนำเสนอลงในพื้นที่ที่กำหนดให้

ประเภทเชื้อเพลิง	จำนวน (ล้านกิโลวัตต์-ชั่วโมง)
ก๊าซธรรมชาติ	76,000
ถ่านหิน (รวมลิกไนต์)	30,000
พลังงานหมุนเวียน	23,000
น้ำมันเตา	300
น้ำมันดีเซล	200
อื่น ๆ	1,500

ที่มาข้อมูล: การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (มกราคม - สิงหาคม พ.ศ.2564) โดยทำการปรับเป็นตัวเลขเป็นจำนวนเต็ม
https://www.egat.co.th/index.php?option=com_content&view=article&id=2455&Itemid=116

แนวทางการนำเสนอข้อมูล

ผู้สอนสร้าง rubric เพื่อประเมินคุณภาพคำตอบตามองค์ประกอบและพฤติกรรมที่กำหนดข้างต้น โดยการแบ่งระดับคะแนนต้องสอดคล้องกับระดับความสามารถหรือพฤติกรรมที่แสดงออก โดยในที่นี้ผู้เขียนเลือกออกแบบและสร้างเกณฑ์การให้คะแนนแบบวิเคราะหฺ์ ดังตัวอย่าง rubric และตัวอย่างคำตอบ ต่อไปนี้

ตาราง 2 ตัวอย่าง rubric และตัวอย่างคำตอบของสมรรถนะการสื่อสารด้านการเขียน

องค์ประกอบ	ระดับคะแนน	คำบรรยาย	ตัวอย่างคำตอบ
ความสามารถในการเล่าเรื่อง	2	เล่าเรื่องจากภาพได้ตรงประเด็นเกี่ยวกับวัฏจักรน้ำ มีการยกตัวอย่างการอธิบายเพิ่มเติมจากภาพได้อย่างน้อย 1 ประเด็นที่ถูกต้องตามแนวคิดทางวิทยาศาสตร์ และใช้ภาษาได้ถูกต้อง	“ภาพวัฏจักรน้ำแสดงการหมุนเวียนของน้ำจากแหล่งหนึ่งไปยังอีกแหล่งหนึ่ง โดยอาจเปลี่ยนสถานะหรือไม่เปลี่ยนก็ได้ เช่น น้ำจากแหล่งน้ำระเหยเป็นไอควบแน่นและจับตัวกันเป็นเมฆ ละอองน้ำในเมฆรวมตัวเป็นหยดน้ำกลายเป็นฝน น้ำฝนไหลลงดินและแหล่งน้ำ เป็นต้น”
	1	เล่าเรื่องจากภาพได้ตรงประเด็นเกี่ยวกับวัฏจักรน้ำ มีการยกตัวอย่างการอธิบายเพิ่มเติมจากภาพได้อย่างน้อย 1 ประเด็นที่ถูกต้องตามแนวคิดทางวิทยาศาสตร์ แต่ยังไม่สมบูรณ์ โดยใช้ภาษาได้ถูกต้องหรืออาจสะกดผิดเล็กน้อย	“ภาพวัฏจักรน้ำแสดงการหมุนเวียนของน้ำจากแหล่งต่าง ๆ บนโลก เช่น น้ำจากใต้ดินไหลลงแหล่งน้ำ น้ำฝนไหลลงสู่ดิน เป็นต้น” โดยผู้เรียนอาจไม่ได้กล่าวถึงเรื่องสถานะของน้ำที่เกิดการหมุนเวียนในบริเวณต่าง ๆ หรืออาจกล่าวถึงสถานะของน้ำแต่ใช้คำอธิบายกระบวนการที่คลาดเคลื่อนเช่น “ภาพวัฏจักรน้ำแสดงการหมุนเวียนของน้ำจากแหล่งต่าง ๆ บนโลก เช่น น้ำจากแหล่งน้ำลอยขึ้นไปรวมตัวเป็นเมฆ”
	0	เล่าเรื่องไม่เกี่ยวข้องกับวัฏจักรน้ำหรือ เล่าเรื่องตรงตามประเด็นแต่ไม่อธิบายหรืออธิบายไม่ถูกต้องตามแนวคิดวิทยาศาสตร์หรือไม่เขียนตอบ	“จากภาพแสดงสภาพภูมิประเทศที่แตกต่างกัน” “การไหลของน้ำจากที่สูงลงสู่ที่ต่ำ” “ภาพวัฏจักรน้ำ”
ความสามารถในการสรุปความ	2	สรุปเนื้อหาจากการอ่านได้ตรงประเด็นเกี่ยวกับ นิโคติน อันตรายของนิโคตินและการเข้าสู่ร่างกาย โดยใช้ภาษาได้ถูกต้อง	“นิโคติน เป็นสารเคมีที่ในควันบุหรี่ หากได้รับปริมาณมากเกินไปจะอันตรายถึงชีวิต สามารถเข้าสู่ร่างกายจากการหายใจเอาควันบุหรี่เข้าไปและการดูดซับทางผิวหนังจากการสัมผัส”
	1	สรุปเนื้อหาจากการอ่านได้ตรงประเด็นเกี่ยวกับ นิโคติน โดยกล่าวถึงอันตรายของนิโคตินหรือการเข้าสู่ร่างกายเพียงเรื่องใดเรื่องหนึ่ง โดยใช้ภาษาได้ถูกต้องหรืออาจสะกดผิดเล็กน้อย	“นิโคติน เป็นสารเคมีที่ในควันบุหรี่ หากได้รับปริมาณมากเกินไปจะอันตรายถึงชีวิต” “นิโคติน สามารถเข้าสู่ร่างกายจากการหายใจเอาควันบุหรี่เข้าไปและ/หรือการดูดซับทางผิวหนังจากการสัมผัส”
	0	สรุปเนื้อหาไม่เกี่ยวข้องกับนิโคติน หรือสรุปตรงตามประเด็นแต่ไม่กล่าวถึงอันตรายของนิโคติน หรือการเข้าสู่ร่างกาย หรือไม่เขียนตอบ	“การสูบบุหรี่ไม่ดีต่อสุขภาพ” “นิโคติน เป็นสารเคมีที่ในควันบุหรี่”

ตาราง 2 (ต่อ)

องค์ประกอบ	ระดับคะแนน	คำบรรยาย	ตัวอย่างคำตอบ
ความสามารถในการตีความ	2	ตีความเจตนาของผู้เขียนบทความได้ชัดเจนและใช้ภาษาได้ถูกต้อง	“ต้องการให้ผู้อ่านได้รู้เกี่ยวกับอันตรายของนิโคติน และการเข้าสู่ร่างกายจากการหายใจและการสัมผัสถึงแม้คนนั้นจะไม่สูบบุหรี่ก็ตาม”
	1	ตีความเจตนาของผู้เขียนได้แต่เขียนสื่อความยังไม่ชัดเจน โดยใช้ภาษาได้ถูกต้องหรืออาจสะกดผิดเล็กน้อย	“ต้องการให้ผู้อ่านได้รู้เกี่ยวกับนิโคติน ที่ทำให้เสียชีวิตได้” “ต้องการให้ผู้อ่านได้รู้เกี่ยวกับผลการทดลองของนิโคติน”
	0	ตีความเจตนาของผู้เขียนไม่ได้ หรือเขียนประเด็นอื่นที่ไม่เกี่ยวกับนิโคติน และ/หรือใช้ภาษาไม่ถูกต้อง หรือไม่เขียนตอบ	“ต้องการให้ผู้อ่านได้รู้ว่าการสูบบุหรี่ไม่ดีต่อสุขภาพ” “สามารถใช้แผ่นแปะนิโคตินแทนการสูบบุหรี่ได้”
ความสามารถในการนำเสนอข้อมูล	2	จัดกระทำข้อมูลเพื่อนำเสนอในรูปแบบใดรูปแบบหนึ่ง (แผนภูมิ กราฟ บรรยาย ฯลฯ) ได้อย่างถูกต้องและสื่อความชัดเจน	นำเสนอข้อมูลโดยอาจทำเป็นแผนภูมิแท่งหรือแผนภูมิวงกลม ที่แสดงปริมาณหรือร้อยละของการใช้เชื้อเพลิงประเภทต่างๆ ได้ถูกต้อง พร้อมทั้งเขียนรายละเอียดองค์ประกอบของแผนภูมิ ได้แก่ ชื่อแผนภูมิ คำอธิบายชุดข้อมูล หากเป็นแผนภูมิแท่งต้องระบุชื่อแกนให้ครบถ้วน
	1	จัดกระทำข้อมูลเพื่อนำเสนอในรูปแบบใดรูปแบบหนึ่ง (แผนภูมิ กราฟ บรรยาย ฯลฯ) ได้ถูกต้องแต่สื่อความไม่ชัดเจนหรือข้อมูลไม่ครบ	นำเสนอข้อมูลโดยอาจทำเป็นแผนภูมิแท่งหรือแผนภูมิวงกลม ที่แสดงปริมาณหรือร้อยละของการใช้เชื้อเพลิงประเภทต่างๆ ได้ถูกต้องแต่ไม่เขียนรายละเอียดองค์ประกอบของแผนภูมิ หรือเขียนไม่ครบถ้วน
	0	ไม่สามารถจัดกระทำข้อมูลเพื่อนำเสนอในรูปแบบใดรูปแบบหนึ่งได้ หรือนำเสนอข้อมูลผิด	นำเสนอข้อมูลโดยอาจทำเป็นแผนภูมิแท่งหรือแผนภูมิวงกลม แต่ข้อมูลไม่ถูกต้อง โดยไม่เขียนรายละเอียดองค์ประกอบของแผนภูมิ หรือเขียนไม่ครบถ้วน

นอกจากนี้การกำหนดเกณฑ์การแปลความหมายของคะแนน จะช่วยให้ผู้สอนสามารถระบุกลุ่มของผู้เรียนตามระดับของสมรรถนะ ซึ่งสามารถกำหนดได้ตามช่วงคะแนนที่ได้จากการประเมิน เช่น จากตัวอย่างที่ผ่านมาแต่ละองค์ประกอบมี 2 คะแนน รวมคะแนนเต็ม 8 คะแนน หากผู้สอนต้องการกำหนดการแปลความหมายของคะแนนเป็น 3 ระดับ ได้แก่ ระดับสูง ปานกลาง ควรปรับปรุง สามารถกำหนดได้ดังตัวอย่างต่อไปนี้

- 0 - 2 คะแนน หมายถึง มีสมรรถนะการสื่อสารด้านการเขียนอยู่ในระดับ ควรปรับปรุง
- 3 - 5 คะแนน หมายถึง มีสมรรถนะการสื่อสารด้านการเขียนอยู่ในระดับ ปานกลาง
- 6 - 8 คะแนน หมายถึง มีสมรรถนะการสื่อสารด้านการเขียนอยู่ในระดับ สูง

อย่างไรก็ดี ในการแบ่งกลุ่มของผู้เรียนตามระดับของสมรรถนะ เพียงช่วยให้ผู้สอนเห็นภาพรวมของผู้เรียนตนเองว่ามีสมรรถนะเป็นอย่างไร ถึงแม้ผู้เรียนจะถูกจัดอยู่ในกลุ่มเดียวกันแต่ไม่ได้หมายความว่าผู้เรียนทุกคนในกลุ่มมีการแสดงออกของพฤติกรรมด้านการเขียนสื่อสารที่ตรงกันทุกประการ ทั้งนี้เนื่องมาจากเกณฑ์การแปลความหมายของคะแนน เป็นการแปลระดับสมรรถนะโดยภาพรวมจากทุกองค์ประกอบ ดังนั้น ในกลุ่มนักเรียนที่ได้คะแนนในระดับเดียวกัน อาจมีคะแนนจากองค์ประกอบย่อยแตกต่างกัน หากผู้สอนจะนำผลการประเมินใช้เพื่อ

วินิจฉัยนักเรียนเป็นรายบุคคล ควรพิจารณาคะแนนของผู้เรียนตามองค์ประกอบย่อย โดยครูผู้สอนสามารถให้ผลสะท้อนกลับกับนักเรียนเพื่อการพัฒนาตัวเองเป็นรายบุคคลให้มีสมรรถนะที่สูงขึ้นได้ และช่วยให้ผู้สอนสามารถปรับการเรียนการสอนให้สามารถส่งเสริมสมรรถนะของผู้เรียนได้ดีขึ้น

ทั้งนี้ หากผู้สอนสนใจที่จะนำรูปрикและเกณฑ์การแปลความหมายของคะแนนไปเทียบเคียงกับระดับการพัฒนาของสมรรถนะหลักเป็น 4 ระดับตามที่ CBE Thailand (2021b) ได้กำหนดไว้ ได้แก่ ระดับเริ่มต้น กำลังพัฒนา สามารถและเหนือความคาดหวัง ผู้สอนควรวิเคราะห์คำบรรยายระดับของสมรรถนะตามระดับขั้นของผู้เรียนให้เข้าใจ เพื่อให้สามารถเทียบเคียงระดับการพัฒนาสมรรถนะได้อย่างเหมาะสมและเป็นไปตามวัยของผู้เรียน โดยผู้เขียนขอเสนอแนวคิดในการเทียบเคียงจากตัวอย่างรูปริกสมรรถนะการสื่อสารด้านการเขียนข้างต้น ซึ่งผู้เขียนกำหนดคะแนนเป็น 0 1 และ 2 อาจเทียบเคียงได้กับระดับเริ่มต้น กำลังพัฒนา และสามารถ ตามลำดับ ซึ่งถือว่าระดับ 2 หรือ สามารถ เป็นสมรรถนะที่ผู้เรียนควรมี ส่วนระดับเหนือความคาดหวังผู้สอนสามารถกำหนดระดับคะแนนเป็น 3 และเขียนคำบรรยายที่สะท้อนพฤติกรรมเกินกว่าที่ปรากฏในระดับ 2 เช่น *องค์ประกอบที่ 1 ความสามารถในการเล่าเรื่อง* ถ้าจะได้ 3 คะแนน สามารถเขียนคำบรรยายได้เป็น “เล่าเรื่องจากภาพได้ตรงประเด็นเกี่ยวกับวัฏจักรน้ำ มีการยกตัวอย่างการอธิบายเพิ่มเติมจากภาพได้ 3 ประเด็นขึ้นไปที่ต้องตามแนวคิดทางวิทยาศาสตร์ได้อย่างเชื่อมโยง โดยใช้ภาษาได้ถูกต้อง” ส่วนเกณฑ์การแปลความหมายของคะแนนผู้สอนสามารถระบุกลุ่มของผู้เรียนตามระดับของสมรรถนะเป็น 4 ระดับมา โดยคิดแต่ละองค์ประกอบมี 3 คะแนน รวมคะแนนเต็ม 12 คะแนน สามารถกำหนดได้ดังตัวอย่างต่อไปนี้

0 - 3 คะแนน หมายถึง มีสมรรถนะการสื่อสารด้านการเขียนอยู่ในระดับ เริ่มต้น

4 - 6 คะแนน หมายถึง มีสมรรถนะการสื่อสารด้านการเขียนอยู่ในระดับ กำลังพัฒนา

7 - 9 คะแนน หมายถึง มีสมรรถนะการสื่อสารด้านการเขียนอยู่ในระดับ สามารถ

10 - 12 คะแนน หมายถึง มีสมรรถนะการสื่อสารด้านการเขียนอยู่ในระดับ เหนือความคาดหวัง

นอกจากนี้สำหรับผู้สนใจ อาจสร้างเครื่องมือประเมินหรือออกแบบกิจกรรมการเรียนรู้ที่สามารถดึงสมรรถนะของผู้เรียนให้แสดงพฤติกรรมได้สูงหรือซับซ้อนมากขึ้น ซึ่งผู้เขียนมองว่าสำหรับรายวิชาวิทยาศาสตร์นั้น การใช้คำถามไม่ว่าจะเป็นในเครื่องมือประเมินหรือการจัดการเรียนรู้มีความสำคัญเป็นอย่างมากที่จะส่งเสริมให้ผู้เรียนแสดงพฤติกรรมตามสมรรถนะที่ได้ตั้งเป้าหมายไว้ ยกตัวอย่างจากคำถามสมรรถนะการสื่อสารด้านการเขียนที่สามารถสะท้อนพฤติกรรมหรือการคิดของผู้เรียนได้สูงขึ้น ในองค์ประกอบที่ 3 ความสามารถในการตีความ ผู้สอนอาจมีการใช้คำถามเพิ่มเติมให้ผู้เรียนแสดงผล เช่น “จากบทความ นักเรียนคิดว่าผู้เขียนมีเจตนาอย่างไรในการเสนอบทความเรื่องนี้ เพราะเหตุใดนักเรียนจึงคิดเช่นนั้น” องค์ประกอบที่ 4 ความสามารถในการนำเสนอข้อมูล “เพราะเหตุใดนักเรียนจึงเลือกแนวทางดังกล่าวในการนำเสนอข้อมูล” เป็นต้น ซึ่งการตอบของผู้เรียนจะสะท้อนถึงการคิดอย่างเป็นเหตุเป็นผล เป็นลำดับขั้นตอน และสะท้อนความเข้าใจที่ถูกต้องในการเลือกใช้แนวทางที่เหมาะสมในการนำเสนอข้อมูล

ขั้นตอนที่ 4 การตรวจสอบคุณภาพเครื่องมือประเมินและรูบรีค เป็นขั้นตอนที่ผู้สอนหาคุณภาพของเครื่องมือและและเกณฑ์การประเมินที่ได้สร้างขึ้น จากที่กล่าวในตอนต้นว่าลักษณะสำคัญประการหนึ่งของการประเมินฐานสมรรถนะคือ เครื่องมือที่ใช้เพื่อประเมินสมรรถนะต้องได้รับการพัฒนาและตรวจสอบคุณภาพอย่างเหมาะสมเพื่อให้สามารถได้ผลการประเมินที่ถูกต้องและน่าเชื่อถือ เครื่องมือประเมินมีประสิทธิภาพและมีความเป็นปรนัย ตามมาตรฐานการทดสอบทางการศึกษาและจิตวิทยา ของประเทศสหรัฐอเมริกา (Standards for Educational and Psychological Testing: American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014) โดยมาตรฐานดังกล่าวกำหนดความน่าเชื่อถือของผลประเมิน สามารถพิจารณาได้จากหลักฐานห้าประเภท ได้แก่ (1) หลักฐานเกี่ยวกับเนื้อหาของข้อสอบหรือแบบสอบ (validity evidence based on test content) ที่พิจารณาจากการให้ผู้เชี่ยวชาญพิจารณาเนื้อหาที่ใช้วัดประเมินว่ามีความถูกต้องสอดคล้องกับแนวคิดทฤษฎีหรือไม่ (2) หลักฐานเกี่ยวกับกระบวนการตอบ (validity evidence based on response processes) ที่พิจารณาจากการวิเคราะห์คำตอบของผู้ถูกประเมินเทียบกับในหลายกลุ่มความสามารถ หรือการพิจารณาความสอดคล้องของเกณฑ์การประเมินกับคำตอบโดยผู้ตรวจหรือผู้ประเมิน (3) หลักฐานเกี่ยวกับตามโครงสร้างภายใน (validity evidence based on internal structure) ที่พิจารณาจากความสัมพันธ์ระหว่างข้อสอบหรือรายการประเมินกับกรอบแนวคิด (conceptual framework) ของเครื่องมือ (4) หลักฐานเกี่ยวกับความสัมพันธ์กับตัวแปรอื่น (validity evidence based on relations to other variables) ที่พิจารณาจากการวิเคราะห์ความสัมพันธ์ของคะแนนกับตัวแปรเกณฑ์หรือผลจากเครื่องมืออื่น ๆ และ (5) หลักฐานเกี่ยวกับผลจากการนำเครื่องมือไปใช้ (validity evidence based on testing consequences) ที่พิจารณาจากผลการทดสอบว่าเป็นไปตามจุดประสงค์หรือตีความตามแนวคิดทฤษฎีได้ถูกต้องหรือไม่ โดยในการหลักฐานความตรงของแต่ละเครื่องมือ อาจไม่จำเป็นต้องครบห้าประเภทแต่ควรแสดงหลักฐานให้ได้มากที่สุด เพื่อให้สะท้อนผลการประเมินที่ถูกต้องและน่าเชื่อถือ

จากแบบวัดสมรรถนะการสื่อสารด้านการเขียนและรูบรีค ที่ผู้เขียนยกตัวอย่างข้างต้น ในการตรวจสอบคุณภาพ ผู้สอนสามารถดำเนินการโดยให้ผู้เชี่ยวชาญพิจารณาข้อคำถามและรูบรีคว่าเป็นไปตามนิยามและองค์ประกอบที่กำหนดหรือไม่ เป็นหลักฐานความตรงเกี่ยวกับเนื้อหาของข้อสอบหรือแบบสอบ ซึ่งสามารถดำเนินการโดยใช้การวิเคราะห์ค่าดัชนีความสอดคล้อง (IOC) ส่วนหลักฐานเกี่ยวกับกระบวนการตอบ ผู้สอนสามารถพิจารณาความสอดคล้องของคำตอบของผู้เรียนกับรูบรีค และหากมีผู้ตรวจแบบวัดหลายคน สามารถดำเนินการโดยให้ผู้ตรวจประเมินผู้เรียน โดยใช้เครื่องมือและรูบรีคเดียวกันว่าจะสามารถประเมินผู้เรียนได้ระดับคะแนนเดียวกันหรือไม่จะแสดงถึงความเป็นปรนัยของข้อคำถามและรูบรีคที่ได้สร้างขึ้น

บทสรุป

การประเมินฐานสมรรถนะมีลักษณะสำคัญหลายประการที่ควรยึดถือเป็นสิ่งนำทางในการปฏิบัติ ซึ่งลักษณะสำคัญดังกล่าวสามารถนำมาประยุกต์ใช้ได้ในทุกสายวิชา สำหรับรายวิชาวิทยาศาสตร์นั้นเมื่อพิจารณาจากแนวทางในการประเมินสมรรถนะ พบว่า วิธีการและเครื่องมือที่ใช้ไม่ได้แตกต่างจากที่ครูวิทยาศาสตร์ใช้กันอยู่ในปัจจุบัน เพียงแต่ในการพัฒนาเครื่องมือประเมินและรูบริค การตรวจสอบคุณภาพ และการนำไปปฏิบัติจริง อาจมีข้อจำกัดหลายประการ ได้แก่ ประการแรก ครูวิทยาศาสตร์อาจยังมีความเข้าใจที่ไม่ชัดเจนในเรื่องของสมรรถนะ จึงอาจทำให้ไม่สามารถเลือกใช้เครื่องมือประเมินและสร้างเกณฑ์ได้เหมาะสม ประการที่สอง ขาดความรู้เรื่องการพัฒนาและหาคุณภาพของเครื่องมือเนื่องจากครูวิทยาศาสตร์และครูในสาขาอื่น ๆ ส่วนใหญ่จะสร้างเครื่องมือเพื่อใช้ประเมินผู้เรียนขึ้นเองโดยไม่ได้คำนึงถึงขั้นตอนการพัฒนาเครื่องมือและการหาความตรงและความเที่ยงของผลการวัดและประเมิน อีกทั้งแนวคิดด้านความตรงที่ปรับเปลี่ยนอาจนำมาซึ่งความไม่เข้าใจอย่างลึกซึ้งในการพัฒนาและตรวจสอบคุณภาพเครื่องมือสำหรับประเมินสมรรถนะ และประการสุดท้าย การนำไปปฏิบัติจริง การปรับเปลี่ยนระบบการศึกษาของประเทศไทยเข้าสู่การศึกษาระดับสมรรถนะ ต้องขับเคลื่อนทั้งด้านการจัดการเรียนการสอนและการประเมินฐานสมรรถนะไปพร้อมกันไม่สามารถปรับเปลี่ยนแค่เพียงส่วนใดส่วนหนึ่งได้ ซึ่งบุคคลที่มีส่วนสำคัญมากที่สุดในการขับเคลื่อนคือครูผู้สอน โดยครูต้องมีความเข้าใจสมรรถนะ แนวทางการจัดการเรียนรู้ และการประเมินสมรรถนะของผู้เรียนอย่างถูกต้องชัดเจน หน่วยงานด้านการจัดการศึกษา การวัดประเมินผลของประเทศ สถาบันการผลิตครู รวมถึงผู้มีส่วนเกี่ยวข้องทุกฝ่ายจึงต้องทำงานร่วมกันทำความเข้าใจให้ตรงกันทั้งระบบเพื่อให้สามารถถ่ายทอดความรู้และแนวทางที่ถูกต้องเหมาะสมให้กับครูได้นำไปปฏิบัติได้อย่างชัดเจน ซึ่งผู้เขียนหวังเป็นอย่างยิ่งว่าบทความนี้จะมีส่วนช่วยในการสร้างความเข้าใจและช่วยให้ผู้อ่านสามารถต่อยอดองค์ความรู้และนำไปปรับใช้ในงานของตนเองได้ไม่มากนัก

References

- Ab Rahman, A., Hanafi, N. M., Mukhtar, M. I., & Ahmad, J. (2014). Assessment practices for competency-based education and training in vocational college, Malaysia. *Procedia - Social and Behavioral Sciences*, 112, 1070-1076. <https://doi.org/10.1016/j.sbspro.2014.01.1271>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association. https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf
- Ana, A., Yulia, C., Jubaedah, Y., Muktiarni, M., Dwiyantri, V., & Maosul, A. (2020). Assessment of student competence using electronic rubric. *Journal of Engineering Science and Technology*, 15(6), 3559-3570.

- Carraccio, C., Wolfsthal, S. D., Englander, R., Ferentz, K., & Martin, C. (2002). Shifting paradigms: from flexner to competencies. *Academic Medicine*, 77, 361-367.
- CBE Thailand. (2021a). *Basic Education core curriculum framework (competency-based curriculum)*. <https://cbethailand.com/>
- CBE Thailand. (2021b). *Six core competencies*. <https://cbethailand.com>
- Davies, A., & Ellison, L. (2003). *School leadership for the 21st century: a competency and knowledge approach*. Routledge.
- Drisko, J. W. (2014). Competencies and their assessment. *Journal of Social Work Education*, 50, 414-426. <https://doi.org/10.1080/10437797.2014.917927>
- Gallardo, K. (2020). Competency-based assessment and the use of performance-based evaluation rubrics in higher education: Challenges towards the next decade. *Problems of Education in the 21st Century*, 78(1), 61-79. <https://doi.org/10.33225/pec/20.78.61>
- Gervais, J. (2016). *The operational definition of competency-based education*. Wiley Online Library. <https://doi.org/10.1002/cbe2.1011>.
- Guthrie, H. (2009). *Competence and competency-based training: what the literature says*. National Centre for Vocational Education Research.
- Ilma, A. Z., Adhelacahya, K., & Ekawati, E. Y. (2021). Assessment for learning model in competency assessment of 21st century student assisted by google classroom. *Journal of Physics: Conference Series*. 1085(1), 1-8. <https://doi.org/10.1088/1742-6596/1805/1/012005>
- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment and Evaluation in Higher Education*, 39(7), 840-852. <https://doi.org/10.1080/02602938.2013.875117>
- Mathelitsch, L. (2013). Competencies in science teaching. *Center for Educational Policy Studies Journal*, 3(3), 49-64.
- McClarty, K. L., & Gaertner, M. N. (2015). *Measuring mastery: Best practices for assessment in competency-based education*. Center on Higher Education Reform, American Enterprise Institute.
- Motsenbocker, P. S. (2018). *A comparative analysis of competency-based versus traditional assessment with respect to academic performance and feedback processes* [Doctoral dissertation]. Concordia University. <https://www.proquest.com/openview/9f8171ac8971ad6da6e47f262a333120/1?pq-origsite=gscholar&cbl=18750>

- Nitko, A. J. (2019). *Educational Assessment of Students* (8th Ed.). Pearson.
- OECD. (2018). *The future of education and skills education 2030: The future we walk*. OECD Publishing.
[https://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf)
- Pearson's CBE Solutions and Services. (2015). *Defining competencies and outlining assessment strategies for competency-based education programs*. Always Learning. Pearson Education, Inc. https://www.pearson.com/content/dam/one-dot-com/one-dot-com/ped-blogs/wp-content/pdfs/584G245_CBE-Assessment_WP_WEB_f.pdf
- Trotter, A., & Ellison, L. (1997). *Understanding competence and competency*, in Davies, B. and Ellison, L. (Eds.) *School Leadership for the 21st Century: A Competency and Knowledge Approach*. Routledge.
- Virk, A., Joshi, A., Mahajan, R., & Singh, T. (2020). The power of subjectivity in competency-based assessment. *Journal of Postgraduate Medicine*, 66(4), 200-205.
https://doi.org/10.4103/jpgm.JPGM_591_20
- Wolf, A. (1995). *Competency-based assessment*. Open University Press.

Translate Thai References

- Ministry of Education. (2010). *Basic Education Core Curriculum B.E. 2551 (A.D. 2008)* (2nd Ed.). Agricultural Co-operative Federation of Thailand Limited. (in Thai)
- Office of the Education Council. (2019). *State of Education in Thailand B.E. 2561/2562 (A.D. 2018/2019) Educational Transformation in the Digital Age*. Parbpim Printing. (in Thai)
- Ritcharoon, P. (2019). Scoring Rubrics: The Tool for Teachers to Accurately and Fairly Evaluate Learning Outcomes. *STOU Education Journal*, 12(1), 1-16. (in Thai)
- Royal Academy. (2012). *Education Dictionary Royal Institute Edition*. Aroonkarnpim Limited Partnership. (in Thai)
- Ruchaiphanit, W. (2016). Thailand launched a Study into the Issue of Quality Education. *Journal of Research for Development Social and Community Rajabhat Mahasarakham University*, 2(6), 2-8. (in Thai)



บทความวิจัย

Research Article

**การประเมินหลักสูตรแบบ CIPPI Model : การศึกษามหาบัณฑิต สาขาวิชา
การสอนวิทยาศาสตร์และคณิตศาสตร์ คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม**
**CIPPI Model for Curriculum Evaluation: Master of Education in Teaching of
Science and Mathematics, Faculty of Education, Mahasarakham University**

กัญญารัตน์ โคจร¹ และ กัญญารัตน์ สอนสุภาพ^{2*}
Kanyarat Cojorn¹ and Kanyarat Sonsupap^{2*}

(Received: January 3, 2022; Revised: February 21, 2022; Accepted: March 7, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมาย เพื่อประเมินหลักสูตรการศึกษามหาบัณฑิต สาขาวิชา การสอน วิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม และ เพื่อศึกษาแนวทางการพัฒนาปรับปรุงหลักสูตร การวิจัยครั้งนี้ เป็นการวิจัยประเมินผลโดยใช้รูปแบบ การประเมินแบบยึดการตัดสินใจ CIPPI model ของ Stufflebeam กลุ่มตัวอย่างผู้ให้ข้อมูลจำนวน 113 คน ประกอบไปด้วย นิสิต อาจารย์ผู้รับผิดชอบหลักสูตรและอาจารย์ผู้สอน มหาบัณฑิต และผู้ใช้บัณฑิต เครื่องมือ ที่ใช้ในการเก็บข้อมูลคือ แบบสัมภาษณ์และแบบสอบถาม การวิเคราะห์ข้อมูลใช้สถิติค่าร้อยละ (%) ค่าเฉลี่ย ของประชากร (μ) ค่าความแปรปรวนของประชากร (σ) และการวิเคราะห์เนื้อหา มีผลการวิจัยจำแนกตาม องค์ประกอบการประเมิน ดังนี้

1) ด้านสภาวะแวดล้อม (context) สภาวะแวดล้อมในภาพรวมมีความเหมาะสมอยู่ในระดับมาก เมื่อพิจารณาเป็นองค์ประกอบ พบว่าด้านวัตถุประสงค์หลักสูตรและโครงสร้างหลักสูตรมีความเหมาะสม ระดับมาก ส่วนในด้านสาระรายวิชามีความเหมาะสมมากที่สุด แต่บางรายวิชามีความซ้ำซ้อนในเนื้อหาจึงควร ปรับลดความซ้ำซ้อนของบางรายวิชาที่มีเนื้อหาใกล้เคียงกันและพิจารณาปรับปรุงรายวิชาให้ตรงหรือสอดคล้อง กับความต้องการของผู้เรียนและสถานการณ์ปัจจุบัน

2) ด้านปัจจัยนำเข้า (input) ปัจจัยนำเข้าในภาพรวม มีความเหมาะสมอยู่ในระดับมาก เมื่อพิจารณา เป็นองค์ประกอบ พบว่าด้านการรับนิสิตและสิ่งสนับสนุนการเรียนรู้มีความเหมาะสมระดับมาก ส่วนในด้าน อาจารย์มีความเหมาะสมมากที่สุด มีข้อเสนอแนะสำหรับบางองค์ประกอบย่อย ได้แก่ ควรเพิ่มการแนะนำ การเข้าถึงแหล่งข้อมูลและแหล่งเรียนรู้ต่างๆด้านการศึกษา การทำวิจัยวิทยานิพนธ์และระบบงานอินเทอร์เน็ต

¹ ผู้ช่วยศาสตราจารย์ ภาควิชาหลักสูตรและการสอน คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

² อาจารย์ ภาควิชาหลักสูตรและการสอน คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

¹ Assistant Professor, Department of Curriculum and Instruction, Faculty of Education, Mahasarakham University

² Lecturer, Department of Curriculum and Instruction, Faculty of Education, Mahasarakham University

* Corresponding Author E-mail: kanyarat.s@msu.ac.th

3) ด้านกระบวนการ (process) กระบวนการในภาพรวม มีความเหมาะสมอยู่ในระดับมาก เมื่อพิจารณาเป็นองค์ประกอบ พบว่าด้านกระบวนการจัดการเรียนรู้ การฝึกประสบการณ์วิชาชีพ และกิจกรรมเสริมหลักสูตร มีความเหมาะสมระดับมาก ส่วนในด้านการทำวิทยานิพนธ์มีความเหมาะสมมากที่สุด มีข้อเสนอแนะคือควรมีการจัดโครงการพัฒนานิสิตด้านต่างๆอย่างต่อเนื่องตลอดระยะเวลาที่นิสิตศึกษาในหลักสูตร

4) ด้านผลผลิต (product) ผลผลิตในภาพรวมมีความเหมาะสมอยู่ในระดับมาก เมื่อพิจารณาเป็นองค์ประกอบ พบว่าคุณลักษณะของบัณฑิตด้านคุณธรรม จริยธรรม ด้านทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ และด้านทักษะการวิเคราะห์เชิงตัวเลข การสื่อสารและการใช้เทคโนโลยีสารสนเทศ มีความเหมาะสมระดับมาก ส่วนในด้านความรู้ ความเข้าใจในวิชาชีพและวิชาเอก ด้านทักษะทางปัญญาและด้านทักษะการจัดการเรียนรู้และความสามารถในการวิจัย มีความเหมาะสมมากที่สุด แสดงให้เห็นว่าหลักสูตรได้รับการยอมรับว่าสามารถผลิตบัณฑิตได้เป็นอย่างดี

5) ด้านผลกระทบ (impact) กระทบในภาพรวม มีความเหมาะสมอยู่ในระดับมากที่สุด เมื่อพิจารณาเป็นรายข้อ พบว่า ผู้ใช้บัณฑิตมีความพึงพอใจต่อผลการปฏิบัติงานของบัณฑิตในทุกองค์ประกอบระดับมากที่สุด มีข้อเสนอแนะคือหลักสูตรควรให้ความสำคัญกับประเด็นเกี่ยวกับการให้ความรู้ด้านแนวทางการพัฒนาวิชาชีพให้กับนิสิต

คำสำคัญ: การประเมินหลักสูตร โมเดล CIPPI ระดับอุดมศึกษา

Abstract

This research aimed to evaluate the Master of Education Curriculum in Teaching of Science and Mathematics (Revised Version A.D. 2017) of the Faculty of Education, Mahasarakham University, and to study guidelines for revision of the Master of Education Curriculum in Teaching of Science and Mathematics. This research was evaluation research, using the decision-oriented CIPPI model based on Stufflebeam. The sample consisted of 113 informants who were students, faculty's members responsible for the curriculum and teaching staff, master's degree graduates, and graduate's employers. The instruments for data collection were an interview form and a questionnaire. The analysis of data employed percentage (%), mean (μ), standard deviation (σ) and content analysis. The results were classified according to the components of the evaluation as follows:

1) The context of the curriculum, as a whole, was appropriate at the high level. When each component was considered, it was found that the curriculum objectives and structure were appropriate at the high level. Meanwhile, the course descriptions were appropriate at the highest level. However, some courses had overlapping content that should be reduced.

So, the similar content in courses should be improved and be made consistent with the needs of students and the present situations.

2) The input of the curriculum, as a whole, was appropriate at the high level. When each component was considered, it was found that the course entry process, study resources and materials were appropriate at the high level. Meanwhile, the teaching staff were appropriate at the highest level. Suggestions were made on some sub-components as follows: there should be more advice on accessing data sources and different learning sources, on doing thesis, and on the Internet system.

3) The process of the curriculum, as a whole, was appropriate at the high level. When each component was considered, it was found that the learning management, practicum and extracurricular activities were appropriate at the high level. Meanwhile, doing thesis was appropriate at the highest level. The suggestion was: projects for student development on different aspects should be organized continuously throughout the curriculum.

4) The product of the curriculum, as a whole, was appropriate at the high level. When each component was considered, it was found that the virtuous and ethical characteristics, interpersonal relation skills, numerical analysis skills, communication and information technological skills were appropriate at the high level. Meanwhile, knowledge and understanding of the profession and major, cognitive skills, learning management skills and ability to do research were appropriate at the highest level. All of these findings indicated that the curriculum was accepted as being able to produce graduates well.

5) The impact of the curriculum, as a whole, was appropriate at the highest level. When each component was considered, it was found that the graduate employers were satisfied with the graduates' performance at the highest level in every component. The suggestion was: the curriculum should prioritize issues relating to approaches in students' professional development.

Keywords: curriculum evaluation, CIPPI model, higher education

บทนำ

คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม เล็งเห็นความสำคัญการประเมินผลการดำเนินงานของหลักสูตรซึ่งเป็นกระบวนการที่เกิดขึ้นภายหลังจากการนำหลักสูตรไปใช้ ว่ามีความสำคัญและจำเป็นอย่างยิ่งต่อการพัฒนาคุณภาพของหลักสูตร โดยเฉพาะเมื่อใช้หลักสูตรไปได้ระยะหนึ่งจะต้องมีการประเมินเพื่อพิจารณาทบทวนเกี่ยวกับคุณภาพของหลักสูตรโดยใช้ผลการใช้หลักสูตรในแง่มุมต่าง ๆ การประเมินหลักสูตรในสถาบันอุดมศึกษา นอกจากผู้ใช้หลักสูตรจะประเมินเพื่อพิจารณา ตรวจสอบระบบการบริหารหลักสูตรเพื่อปรับปรุงคุณภาพตามขั้นตอนของการพัฒนาหลักสูตรแล้วยังต้องปฏิบัติเพื่อให้สอดคล้องกับเกณฑ์

มาตรฐานหลักสูตรระดับบัณฑิตศึกษา พ.ศ. 2558 ตามประกาศกระทรวงศึกษาธิการที่กำหนดให้ทุกหลักสูตร มีการพัฒนาหลักสูตรให้ทันสมัย แสดงการปรับปรุงดัชนีมาตรฐานและคุณภาพการศึกษา เป็นระยะอย่างน้อย ทุก 5 ปี และมีการประเมินเพื่อพัฒนาหลักสูตรอย่างต่อเนื่องทุก 5 ปี ทั้งนี้เพื่อให้หลักสูตรมีคุณภาพและ มาตรฐานอย่างต่อเนื่อง

เนื่องจากสาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ ภาควิชาหลักสูตรและการสอน คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม ได้มีการปรับปรุงหลักสูตรมาแล้วหลังจากใช้หลักสูตรมาระยะหนึ่งใน ปีพ.ศ. 2557 และต่อมาได้มีการปรับปรุงหลักสูตรเดิมพร้อมกับดำเนินการใช้หลักสูตรใหม่คือหลักสูตร การศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) เริ่มใช้กับ นิสิตระดับบัณฑิตศึกษารุ่นปีการศึกษา 2560 ซึ่งหลังจากใช้หลักสูตรปรับปรุง พ.ศ. 2560 มาจนถึงปีการศึกษา 2563 ยังไม่ได้ทำการประเมินและติดตามผลผลิตของหลักสูตรปรับปรุงพ.ศ. 2560 เพื่อพัฒนาผู้เรียนให้เป็นไป ตามจุดมุ่งหมายของสถาบัน จึงทำการวิเคราะห์หลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์ และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ประกอบด้วยโครงสร้างหลักสูตรจำนวน 48 หน่วยกิต เกี่ยวข้อง กับรายวิชาต่าง ๆ ที่ เป็นความจำเพาะของวิชาชีพครูและวิชาเฉพาะสาขา ตลอดจนการนำความรู้จาก ภาควิชาปฏิบัติ และการฝึกประสบการณ์วิชาชีพครูในสถานศึกษา ตามกรอบมาตรฐานคุณวุฒิทั้ง 6 ด้าน คือ การปฏิบัติงานในด้านคุณธรรม จริยธรรม ความรู้ความเข้าใจในวิชาชีพและวิชาเอก ความสามารถ ด้านทักษะทางปัญญา ทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ ทักษะการวิเคราะห์เชิงตัวเลข การสื่อสารและการใช้เทคโนโลยีสารสนเทศ และการจัดการเรียนรู้ความสามารถในการวิจัย ซึ่งพบว่า ด้านความพึงพอใจของผู้บังคับบัญชาและผู้ร่วมงานโดยรวมอยู่ในระดับพึงพอใจปานกลางถึงพึงพอใจมาก ดังนั้น คณะผู้วิจัยจึงมีแนวคิดในการดำเนินการประเมินหลักสูตรการศึกษามหาบัณฑิตสาขาวิชาการสอน วิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษาแห่งชาติ และแบบยึดการตัดสินใจ CIPPI Model ของ Stufflebeam et al. (2007) ทั้งในด้านบริบท (context) ด้านปัจจัยนำเข้า (input) ด้านกระบวนการ (process) ด้านผลผลิต (product) และด้านผลกระทบ (impact) ซึ่งผลที่ได้จากการประเมินหลักสูตรในครั้งนี้จะนำมาเป็นข้อมูลและใช้เป็นแนวทางสำหรับผู้ที่เกี่ยวข้องในการ จัดทำหลักสูตรที่จะครบวงจรรอบการปรับปรุงในปีการศึกษา 2565 ให้สามารถนำข้อมูลที่ได้ไปดำเนินการ พิจารณาปรับปรุงและพัฒนาหลักสูตรให้มีประสิทธิภาพยิ่งขึ้นและเหมาะสมสอดคล้องกับความต้องการ ของสังคมในอนาคตต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อประเมินหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) มหาวิทยาลัยมหาสารคาม ใน 5 ด้านได้แก่ ด้านบริบท ด้านปัจจัยนำเข้า ด้านกระบวนการ ด้านผลผลิต และด้านผลกระทบ
2. เพื่อศึกษาแนวทางในการปรับปรุงหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์ และคณิตศาสตร์ มหาวิทยาลัยมหาสารคาม

วิธีดำเนินการวิจัย

1. กลุ่มตัวอย่าง

กลุ่มตัวอย่าง ในการเก็บข้อมูลข้อมูลวิจัยครั้งนี้มีจำนวน 113 คน ประกอบไปด้วย 4 กลุ่มได้แก่

1. นิสิตที่กำลังศึกษาอยู่ในหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ระหว่างปีการศึกษา 2563-2564 จำนวน 37 คน
2. อาจารย์ผู้รับผิดชอบหลักสูตรและอาจารย์ผู้สอน หลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) จำนวน 7 คน
3. มหาวิทยาลัยที่สำเร็จการศึกษาหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ที่จบการศึกษาตั้งแต่ปีการศึกษา 2562 - 2564 จำนวน 59 คน
4. ผู้ใช้บัณฑิตที่สำเร็จการศึกษาหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) จำนวน 10 คน

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือวิจัยประกอบไปด้วยแบบสอบถาม จำนวน 4 ชุด และแบบสัมภาษณ์ จำนวน 2 ชุด ที่ผู้วิจัยพัฒนาขึ้น โดยจำแนกตามกลุ่มผู้ให้ข้อมูล ดังนี้

1. แบบสอบถาม ประกอบด้วย 4 ชุด ดังนี้

ชุดที่ 1 แบบสอบถามความคิดเห็นที่มีต่อหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) สำหรับมหาบัณฑิต โดยผู้วิจัยได้การวิเคราะห์คุณภาพเครื่องมือ โดยการหาค่าความเที่ยงตรงเชิงเนื้อหา (content validity) ของข้อคำถามในแบบสอบถาม โดยผู้เชี่ยวชาญจำนวน 5 ท่าน พบว่ามีค่าอยู่ระหว่าง 0.80-1.00

ชุดที่ 2 แบบสอบถามความคิดเห็นที่มีต่อหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) สำหรับนิสิต โดยผู้วิจัยได้การวิเคราะห์คุณภาพเครื่องมือ โดยการหาค่าความเที่ยงตรงเชิงเนื้อหา (content validity) ของข้อคำถามในแบบสอบถาม โดยผู้เชี่ยวชาญจำนวน 5 ท่าน พบว่ามีค่าอยู่ระหว่าง 0.80-1.00

ชุดที่ 3 แบบสอบถามความคิดเห็นที่มีต่อหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) สำหรับอาจารย์ผู้รับผิดชอบหลักสูตรและอาจารย์ผู้สอน โดยผู้วิจัยได้การวิเคราะห์คุณภาพเครื่องมือ โดยการหาค่าความเที่ยงตรงเชิงเนื้อหา (content validity) ของข้อคำถามในแบบสอบถาม โดยผู้เชี่ยวชาญจำนวน 5 ท่าน พบว่ามีค่าอยู่ระหว่าง 0.80-1.00

ชุดที่ 4 แบบสอบถามความคิดเห็นที่มีต่อหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) สำหรับผู้ใช้บัณฑิต โดยผู้วิจัยได้การวิเคราะห์คุณภาพเครื่องมือ โดยการหาค่าความเที่ยงตรงเชิงเนื้อหา (content validity) ของข้อคำถามในแบบสอบถาม โดยผู้เชี่ยวชาญจำนวน 5 ท่าน พบว่ามีค่าอยู่ระหว่าง 0.80-1.00

2. แบบสัมภาษณ์ ประกอบด้วย 2 ชุด ดังนี้

ชุดที่ 1 แบบสัมภาษณ์ความคิดเห็นที่มีต่อหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) สำหรับอาจารย์ผู้รับผิดชอบหลักสูตรและอาจารย์ผู้สอน โดยผู้วิจัยได้การวิเคราะห์คุณภาพเครื่องมือ โดยการหาค่าความเที่ยงตรงเชิงเนื้อหา (content validity) ของข้อคำถามในสัมภาษณ์ โดยให้ผู้เชี่ยวชาญจำนวน 5 ท่าน พบว่ามีค่าอยู่ระหว่าง 0.80-1.00 และมีค่าความเหมาะสมของข้อคำถาม 4.60-5.00

ชุดที่ 2 แบบสัมภาษณ์ความคิดเห็นที่มีต่อหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุงพ.ศ. 2560) สำหรับมหาบัณฑิต โดยผู้วิจัยได้การวิเคราะห์คุณภาพเครื่องมือ โดยการหาค่าความเที่ยงตรงเชิงเนื้อหา (content validity) ของข้อคำถามในสัมภาษณ์ โดยผู้เชี่ยวชาญจำนวน 5 ท่าน พบว่ามีค่าอยู่ระหว่าง 0.80-1.00 และมีค่าความเหมาะสมของข้อคำถาม 4.40-5.00

3. การเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลในการประเมินผลหลักสูตรครั้งนี้จะดำเนินการเก็บรวบรวมข้อมูลในด้านต่าง ๆ ตาม CIPPI Model จากกลุ่มผู้ให้ข้อมูลที่แตกต่างกัน และมีการเก็บรวบรวมข้อมูลทั้งเชิงปริมาณจากแบบสอบถาม และเชิงคุณภาพจากแบบสัมภาษณ์ รายละเอียดดังตาราง 1

ตาราง 1 ประเด็นการประเมินและวิธีการเก็บรวบรวมข้อมูลจำแนกตามกลุ่มผู้ให้ข้อมูล

ประเด็นที่ประเมิน	ผู้ให้ข้อมูล(แบบสอบถาม)				ผู้ให้ข้อมูล(สัมภาษณ์)			
	ผู้รับผิดชอบ/ อาจารย์ ประจำ หลักสูตร	นิสิต	มหา บัณฑิต	ผู้ใช้ บัณฑิต	ผู้รับผิดชอบ/ อาจารย์ ประจำ หลักสูตร	นิสิต	มหา บัณฑิต	ผู้ใช้ บัณฑิต
บริบท (context)								
วัตถุประสงค์ของหลักสูตร	✓	✓	✓	-	✓	-	✓	-
โครงสร้างรายวิชา	✓	✓	✓	-	✓	-	✓	-
สาระรายวิชารายวิชา	✓	✓	✓	-	✓	-	✓	-
ปัจจัยนำเข้า (input)								
ด้านการรับนิสิต	✓	✓	✓	-	✓	-	✓	-
ด้านอาจารย์	✓	✓	✓	-	✓	-	✓	-
สิ่งสนับสนุนการเรียนรู้	✓	✓	✓	-	✓	-	✓	-
กระบวนการ (process)								
กระบวนการจัดการเรียนรู้	✓	✓	✓	-	✓	-	✓	-
การฝึกประสบการณ์วิชาชีพ	✓	✓	✓	-	✓	-	✓	-
การทำวิทยานิพนธ์	✓	✓	✓	-	✓	-	✓	-
กิจกรรมเสริมหลักสูตร	✓	✓	✓	-	-	-	✓	-

ตาราง 1 (ต่อ)

ประเด็นที่ประเมิน	ผู้ให้ข้อมูล(แบบสอบถาม)				ผู้ให้ข้อมูล(สัมภาษณ์)			
	ผู้รับผิดชอบ/ อาจารย์ ประจำ หลักสูตร	นิสิต	มหา บัณฑิต	ผู้ใช้ บัณฑิต	ผู้รับผิดชอบ/ อาจารย์ ประจำ หลักสูตร	นิสิต	มหา บัณฑิต	ผู้ใช้ บัณฑิต
ผลผลิต (product)								
ด้านคุณธรรม จริยธรรม	✓	✓	✓	-	-	-	-	-
ด้านความรู้ ความเข้าใจในวิชาชีพ และวิชาเอก	✓	✓	✓	-	-	-	-	-
ด้านทักษะทางปัญญา	✓	✓	✓	-	-	-	-	-
ด้านทักษะความสัมพันธ์ระหว่าง บุคคลและความรับผิดชอบ	✓	✓	✓	-	-	-	-	-
ด้านทักษะการวิเคราะห์เชิง ตัวเลข การสื่อสารและการใช้ เทคโนโลยีสารสนเทศ	✓	✓	✓	-	-	-	-	-
ด้านทักษะการจัดการเรียนรู้และ ความสามารถในการวิจัย	✓	✓	✓	-	-	-	-	-
ผลกระทบ (impact)								
ด้านคุณธรรม จริยธรรม	-	-	✓	✓	-	-	-	-
ด้านความรู้ ความเข้าใจในวิชาชีพ และวิชาเอก	-	-	✓	✓	-	-	-	-
ด้านทักษะทางปัญญา	-	-	✓	✓	-	-	-	-
ด้านทักษะความสัมพันธ์ระหว่าง บุคคลและความรับผิดชอบ	-	-	✓	✓	-	-	-	-
ด้านทักษะการวิเคราะห์เชิง ตัวเลข การสื่อสารและการใช้ เทคโนโลยีสารสนเทศ	-	-	✓	✓	-	-	-	-
ด้านทักษะการจัดการเรียนรู้และ ความสามารถในการวิจัย	-	-	✓	✓	✓	-	✓	

ทั้งนี้ดำเนินการเก็บข้อมูล รายละเอียดดังนี้

1. การเก็บรวบรวมข้อมูลจากการสอบถามความคิดเห็น ผู้วิจัยติดต่อขอความร่วมมือจากกลุ่มตัวอย่าง เพื่อขอความอนุเคราะห์ในการเก็บรวบรวมข้อมูลด้วยวิธีการตอบแบบสอบถามผ่านระบบ google form และ เมื่อได้ข้อมูลมาแล้วก็ทำการการวิเคราะห์ข้อมูลเชิงปริมาณ

2. การเก็บรวบรวมข้อมูลจากการสัมภาษณ์ โดยผู้วิจัยติดต่อขอความร่วมมือจากกลุ่มตัวอย่าง และนัดหมายเวลาเพื่อดำเนินการในสัมภาษณ์ออนไลน์ผ่านโปรแกรม google meet โดยผู้วิจัยเตรียมแบบสัมภาษณ์ส่งให้กลุ่มตัวอย่างก่อนการสัมภาษณ์รวมทั้งมีการชี้แจงจุดมุ่งหมายในสัมภาษณ์และเริ่มดำเนินการสัมภาษณ์จากนั้นบันทึกผลสัมภาษณ์ แล้ววิเคราะห์ข้อมูลจากสัมภาษณ์ด้วยการวิเคราะห์เชิงเนื้อหา (content analysis)

4. การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลพื้นฐานของกลุ่มตัวอย่างโดยใช้ค่าความถี่และร้อยละ และการวิเคราะห์ข้อมูลจากแบบสอบถามความคิดเห็นที่มีต่อหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ด้วยการหาค่าเฉลี่ย (μ) และค่าเบี่ยงเบนมาตรฐาน (σ) และแปลความหมายข้อมูล (Srisa-ard, 2010) รายละเอียดดังนี้

ค่าเฉลี่ย 4.51 – 5.00 หมายถึง มีความเหมาะสมอยู่ในระดับมากที่สุด

ค่าเฉลี่ย 3.51 – 4.50 หมายถึง มีความเหมาะสมอยู่ในระดับมาก

ค่าเฉลี่ย 2.51 – 3.50 หมายถึง มีความเหมาะสมอยู่ในระดับปานกลาง

ค่าเฉลี่ย 1.51 – 2.50 หมายถึง มีความเหมาะสมอยู่ในระดับน้อย

ค่าเฉลี่ย 1.00 – 1.50 หมายถึง มีความเหมาะสมอยู่ในระดับที่ต่ำ

ในส่วนการวิเคราะห์ข้อมูลจากการสัมภาษณ์ความคิดเห็นที่มีต่อหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ผู้วิจัยได้นำมาจัดหมวดหมู่และลำดับคำตอบเพื่อนำไปวิเคราะห์เนื้อหา (content analysis) และนำเสนอผลการวิเคราะห์ในลักษณะพรรณนาความ

ผลการวิจัย

ผลจากการประเมินหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม สามารถสรุปเป็นด้านต่าง ๆ ได้ดังนี้

ด้านบริบทของหลักสูตรซึ่งประเมินบริบทของหลักสูตรด้านวัตถุประสงค์ของหลักสูตรโครงสร้างรายวิชา และสาระรายวิชา รายละเอียดดังตาราง 2

ตาราง 2 ผลการวิเคราะห์ข้อมูลองค์ประกอบด้านบริบทของหลักสูตร

ด้านบริบทของหลักสูตร	μ	σ	ระดับความเหมาะสม
วัตถุประสงค์ของหลักสูตร	4.28	0.48	มาก
โครงสร้างรายวิชา	4.35	0.70	มาก
สาระรายวิชา	4.53	0.59	มากที่สุด
ภาพรวมด้านบริบท	4.39	0.59	มาก

จากตาราง 2 พบว่า ในภาพรวมสภาพแวดล้อมด้านบริบทของหลักสูตรมีความเหมาะสมระดับมาก โดยด้านวัตถุประสงค์หลักสูตรและโครงสร้างหลักสูตรมีความเหมาะสมระดับมาก ส่วนในด้านสาระรายวิชา มีความเหมาะสมมากที่สุด นอกจากนี้ยังมีข้อเสนอแนะในประเด็นต่าง ๆ ดังต่อไปนี้ ควรเพิ่มสาระรายวิชา ที่เกี่ยวข้องกับการสร้างเครื่องมือในการวัดผลประเมินผลในตัวแปรต่าง ๆ ทั้งนี้ได้มีการสะท้อนถึงปัญหา อุปสรรคในการศึกษาในหลักสูตร คือระยะเวลาในการศึกษาในหลักสูตรมีน้อยเกินไป ทำให้ขาดความเข้าใจเชิงลึก ทั้งนี้เพราะนิสิตเปลี่ยนสายวิชาจากด้านวิทยาศาสตร์มาด้านศึกษาศาสตร์

จากการสัมภาษณ์และสอบถามความคิดเห็นผู้บริหารและผู้สอน มหาบัณฑิตพบว่าควรเพิ่มรายวิชา ทางด้านการการสอนและการจัดการเรียนรู้เนื่องจากมีความจำเป็นเมื่อต้องนำไปใช้ในการจัดการเรียนการสอน นอกจากนี้ยังมีการเสนอให้เพิ่มการจัดการรู้ สื่อการสอนและนวัตกรรมการเรียนรู้แบบออนไลน์เพิ่มขึ้นเพื่อให้ สอดคล้องกับบริบทในปัจจุบัน

“ควรมีการปรับปรุงรายวิชาให้ทันสมัยโดยเฉพาะรายวิชาที่เกี่ยวกับการจัดการเรียนการสอน ออนไลน์” (อาจารย์ผู้สอน คนที่ 1)

“หลักสูตรควรมีวิชาที่เกี่ยวข้องกับนวัตกรรมการสอนแบบออนไลน์ สื่อการสอน และรูปแบบ การสอนแบบออนไลน์ เพิ่มขึ้น เพื่อให้บัณฑิตได้เตรียมเพื่อไปฝึกประสบการณ์วิชาชีพได้อย่างมั่นใจ” (อาจารย์ผู้สอน คนที่ 3)

นอกจากนี้ยังมีข้อเสนอให้รวมรายวิชาที่เนื้อหาที่มีความซ้ำซ้อนเพื่อลดภาระทางการเรียนให้กับนิสิตและ ลดความตึงเครียดที่เกิดจากการเรียน

“หากมีเนื้อหาบางรายวิชารวมกันได้ก็น่าจะดี” (อาจารย์ผู้สอน คนที่ 2)

อย่างไรก็ตามมีข้อเสนอจากมหาบัณฑิตเสนอให้เพิ่มเวลาในการทดลองสอนเพิ่มขึ้นเพื่อให้บัณฑิต ได้ฝึกปฏิบัติและหาแนวทางการสอนของตนก่อนที่จะต้องนำไปใช้ในการฝึกประสบการณ์วิชาชีพและในการ ประกอบวิชาชีพ

“อยากให้มีการทดลองสอนหลาย ๆ แบบ เพื่อจะได้หาสไตล์การสอนของตัวเอง” (มหาบัณฑิต คนที่ 1)

ความคิดเห็นจากการสนทนากลุ่มเสนอให้มีรายวิชาที่ช่วยให้นิสิตปรับตัวเมื่อไปอยู่ในบริบทของ สถานศึกษาให้เข้ากับระบบ การทำงานและการใช้ชีวิตหลังจบการศึกษา

“อยากให้มียุทธศาสตร์เกี่ยวกับการปรับตัวและบุคลิกภาพ การวางตัวเมื่อเวลาไปสอน” (สนทนากลุ่ม)

ด้านปัจจัยนำเข้าของหลักสูตรซึ่งประเมินบริบทของหลักสูตรด้านการรับนิสิต ด้านอาจารย์ และ สิ่งสนับสนุนการเรียนรู้รายละเอียดดังตาราง 4

ตาราง 4 ผลการวิเคราะห์ข้อมูลองค์ประกอบด้านปัจจัยนำเข้าของหลักสูตร

ด้านปัจจัยนำเข้าของหลักสูตร	μ	σ	ระดับความคิดเห็น
ด้านการรับนิสิต	4.01	0.71	มาก
ด้านอาจารย์	4.64	0.69	มากที่สุด
สิ่งสนับสนุนการเรียนรู้	3.83	0.72	มาก
ภาพรวมด้านปัจจัยนำเข้า	4.16	0.70	มาก

จากตาราง 4 พบว่าในภาพรวมด้านปัจจัยนำเข้าของหลักสูตรมีความเหมาะสมระดับมาก โดยด้านการรับนิสิตและสิ่งสนับสนุนการเรียนรู้มีความเหมาะสมระดับมาก ส่วนในด้านอาจารย์มีความเหมาะสมมากที่สุด นอกจากนี้ยังมีข้อเสนอแนะในประเด็นต่าง ๆ ดังนี้ ด้านระบบสิ่งสนับสนุนการเรียนรู้ มีการเสนอให้ควรมีการสนับสนุนให้บริการด้านอินเทอร์เน็ตให้ครอบคลุมมากขึ้น สนับสนุนการเข้าถึงแหล่งเรียนรู้ต่าง ๆ เช่น การสืบค้นฐานข้อมูลออนไลน์ ฐานข้อมูลงานวิจัย การให้บริการจากสำนักวิทยบริการ ทั้งนี้ได้มีการสะท้อนถึงปัญหา อุปสรรคในการศึกษาในหลักสูตร ดังนี้ แหล่งพื้นที่บริการอินเทอร์เน็ตที่ยังไม่ครอบคลุม การเข้าถึงแหล่งเรียนรู้ต่าง ๆ ด้านการศึกษา การทำวิทยานิพนธ์มีจำกัด และขาด พบว่า การสนับสนุนห้อง/พื้นที่เรียนรู้ด้วยตนเองทั้งในและนอกเวลาราชการ

จากการสัมภาษณ์ พบว่าด้านอาจารย์ผู้สอนมีความสามารถในการสอนตรงตามสาระวิชาและมีความเอาใจใส่ผู้เรียน

“อาจารย์มีความรู้ในวิชาที่สอน” (มหาบัณฑิต คนที่ 3)

“สามารถช่วยเหลือดีมาก ให้คำปรึกษาดี อาจารย์พร้อมที่จะช่วย” (มหาบัณฑิต คนที่ 2)

ในด้าน วัสดุอุปกรณ์ บรรยากาศและสถานที่ จากการสนทนากลุ่มสามารถสรุปได้ว่าห้องเรียนมีความเพียงพอ มีคุณภาพ และความสะดวกตามความต้องการพื้นฐาน

ด้านกระบวนการของหลักสูตรซึ่งประเมินบริบทของหลักสูตรด้านกระบวนการจัดการเรียนรู้การฝึกประสบการณ์วิชาชีพ การทำวิทยานิพนธ์ และกิจกรรมเสริมหลักสูตร รายละเอียดดังตาราง 5

ตาราง 5 ผลการวิเคราะห์ข้อมูลองค์ประกอบด้านกระบวนการของหลักสูตร

ด้านกระบวนการของหลักสูตร	μ	σ	ระดับความคิดเห็น
กระบวนการจัดการเรียนรู้	4.33	0.49	มาก
การฝึกประสบการณ์วิชาชีพ	4.46	0.63	มาก
การทำวิทยานิพนธ์	4.53	0.56	มากที่สุด
กิจกรรมเสริมหลักสูตร	4.03	0.81	มาก
ภาพรวมด้านกระบวนการ	4.34	0.62	มาก

จากตาราง 5 พบว่าในภาพรวมด้านกระบวนการของหลักสูตรมีความเหมาะสมระดับมาก โดยด้านกระบวนการจัดการเรียนรู้ การฝึกประสบการณ์วิชาชีพ และกิจกรรมเสริมหลักสูตร มีความเหมาะสมระดับมาก ส่วนในด้านการทำวิทยานิพนธ์มีความเหมาะสมมากที่สุด นอกจากนี้ยังมีข้อเสนอแนะในประเด็นต่าง ๆ ดังนี้ กิจกรรมการเรียนการสอนควรเพิ่มเติมเน้นการใช้ภาษาอังกฤษมากขึ้น และเนื่องจากกิจกรรมการเรียนการสอนในวิชาชีพครูต้องมีการออกปฏิบัติจริงในสถานศึกษา ซึ่งด้วยช่วงเวลาที่จำกัดจึงเป็นการลำบากในการบริหารจัดการเวลาระหว่างเวลาเรียนและการออกปฏิบัติการในสถานศึกษา จึงอยากให้จัดสรรเวลาให้ชัดเจน เช่น จัดการเรียนการสอนในชั้นเรียน เพื่อออกปฏิบัติกิจกรรมต่างในสถานศึกษา นอกจากนี้ยังเสนอให้ เพิ่มเติมการให้ตัวอย่างนวัตกรรมการเรียนการสอนที่หลากหลายมากขึ้นในการจัดกิจกรรมการเรียนการสอน เพิ่มกิจกรรมโครงการเกี่ยวกับการจัดการเรียนรู้ในศตวรรษที่ 21 และการใช้เทคโนโลยีในการจัดการเรียนการสอน ควรจัดกิจกรรมเพื่อเป็นการเตรียมความพร้อมให้กับนิสิตในการเตรียมตัวสอบใบประกอบวิชาชีพครู และจัดกิจกรรม/โครงการ เกี่ยวกับการทำวิทยานิพนธ์อยากเป็นระบบตั้งแต่เริ่มหัวข้อวิทยานิพนธ์ การเตรียมเค้าโครงวิทยานิพนธ์ การสร้างเครื่องมือ การเขียนรายงานวิทยานิพนธ์ ตลอดจนการเตรียมบทความวิจัยเพื่อเผยแพร่ ตลอดจนวารสารที่เหมาะสมในการเผยแพร่ผลงานวิจัย นอกจากนี้ทั้งนี้ได้มีการสะท้อนถึงปัญหา อุปสรรคในการศึกษาในหลักสูตร ดังนี้ นิสิตมีความกดดันจากการเรียนหลายวิชาพร้อมกัน รวมทั้งเป็นรายวิชาที่ไม่คุ้นเคย เพราะเป็นการเปลี่ยนจากด้านวิทยาศาสตร์มาสู่ด้านศึกษาศาสตร์ และระยะเวลาที่จำกัดจึงทำให้ขาดความรู้ความเข้าใจเชิงลึกในสาระรายวิชา นิสิตยังขาดขาดความรู้ความเข้าใจในการทำวิทยานิพนธ์ในด้านการศึกษา เพราะขาดประสบการณ์ จึงทำให้การทำวิทยานิพนธ์ล่าช้า ส่งผลต่อการจบการศึกษา

ประเด็นด้านการทำวิทยานิพนธ์พบว่านิสิตประสบปัญหาในช่วงแรกของการทำวิทยานิพนธ์โดยเฉพาะช่วงการคิดหาหัวข้อวิทยานิพนธ์

“ช่วงแรกมีปัญหาในการหาหัวข้อวิทยานิพนธ์ เพราะไม่เข้าใจว่าการวิจัยทางด้านศึกษาศาสตร์ควรจะเป็นแบบไหน” (มหาบัณฑิต คนที่ 5)

โดยมีการเสนอความคิดเห็นเพิ่มเติมว่าหลักสูตรควรจะมีระบบที่จัดให้มีรุ่นพี่ให้คำปรึกษากับรุ่นน้องในเรื่องเกี่ยวกับการคิดหาหัวข้อวิทยานิพนธ์และการทำวิทยานิพนธ์

“ควรมีรุ่นพี่ให้คำแนะนำเรื่องการทำหัวข้อวิทยานิพนธ์เหมือนเป็นระบบ mentor เพราะบางทีนิสิตอาจไม่กล้าคุยกับอาจารย์เท่ากับคุยกับรุ่นพี่” (มหาบัณฑิต คนที่ 1)

อย่างไรก็ตามจากการสัมภาษณ์พบว่าระบบกระบวนการจัดระบบอาจารย์ที่ปรึกษาวิทยานิพนธ์ และการกำกับติดตามวิทยานิพนธ์มีระบบที่ดี

“มีระบบการติดตามการทำวิทยานิพนธ์ที่ดี ถ้าทำตามระบบได้จะ perfect มาก” (มหาบัณฑิต คนที่ 2)

ในส่วนของอาจารย์ที่ปรึกษามีเอาใจใส่ พร้อมทั้งจะให้คำแนะนำปรึกษา และติดตามการทำวิทยานิพนธ์ของนิสิตจนกระทั่งนิสิตทำวิทยานิพนธ์เสร็จสมบูรณ์

“ที่ปรึกษามีเวลาให้ เอาใจใส่ และติดตามงานดี” (มหาบัณฑิต คนที่ 1)

“สามารถติดต่อได้หลายช่องทาง” (มหาบัณฑิต คนที่ 2)

“ถ้าต้องเซ็นเอกสารอะไรอาจารย์ก็เซ็นให้ตลอดค่ะ” (มหาบัณฑิต คนที่ 5)

ด้านผลผลิตของหลักสูตรซึ่งประเมินองค์ประกอบด้านคุณธรรม จริยธรรม ด้านความรู้ ความเข้าใจ ในวิชาชีพและวิชาเอก ด้านทักษะทางปัญญา ด้านทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ ด้านทักษะการวิเคราะห์เชิงตัวเลข การสื่อสารและการใช้เทคโนโลยีสารสนเทศ และด้านทักษะการจัดการ การเรียนรู้และความสามารถในการวิจัย รายละเอียดดังตาราง 7

ตาราง 7 ผลการวิเคราะห์ข้อมูลองค์ประกอบด้านผลผลิตของหลักสูตร

ด้านผลผลิตของหลักสูตร	μ	σ	ระดับความคิดเห็น
ด้านคุณธรรม จริยธรรม	4.25	0.62	มาก
ด้านความรู้ ความเข้าใจในวิชาชีพและวิชาเอก	4.56	0.56	มากที่สุด
ด้านทักษะทางปัญญา	4.61	0.42	มากที่สุด
ด้านทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ	4.39	0.63	มาก
ด้านทักษะการวิเคราะห์เชิงตัวเลข การสื่อสารและการใช้เทคโนโลยีสารสนเทศ	4.49	0.52	มาก
ด้านทักษะการจัดการเรียนรู้และความสามารถในการวิจัย	4.51	0.54	มากที่สุด
ภาพรวมด้านผลผลิตหลักสูตร	4.46	0.54	มาก

จากตาราง 7 พบว่าในภาพรวมด้านผลผลิตหลักสูตรมีความเหมาะสมระดับมาก โดยด้านคุณธรรม จริยธรรม ด้านทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ และด้านทักษะการวิเคราะห์เชิงตัวเลข การสื่อสารและการใช้เทคโนโลยีสารสนเทศ มีความเหมาะสมระดับมาก ส่วนในด้านความรู้ ความเข้าใจ ในวิชาชีพและวิชาเอก ด้านทักษะทางปัญญาและด้านทักษะการจัดการเรียนรู้และความสามารถในการวิจัย มีความเหมาะสมมากที่สุด นอกจากนี้ยังมีข้อเสนอแนะในประเด็นต่างๆ ดังนี้ หลักสูตรควรมีการส่งเสริม เรื่องข้อมูลของแหล่งตีพิมพ์บทความ รวมทั้งให้ความรู้ในเรื่องการตีพิมพ์มากยิ่งขึ้นเพื่อนำมาปรับใช้ในวิชาชีพด้วย

อาจารย์ผู้สอนยังเห็นว่า หลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์เป็นหลักสูตรที่มีความสมบูรณ์ทางวิชาการที่ผู้เรียนสามารถเป็นมหาบัณฑิตที่มีคุณภาพได้

“เวลาที่ผมได้ไปโรงเรียนและได้สอบถามกับคุณครูในโรงเรียนก็เห็นว่าทุกคนก็บอกว่านิสิตของเรา มีคุณภาพ สอนดีครับ” (อาจารย์ผู้สอน คนที่2)

ด้านผลกระทบของหลักสูตรประเมินองค์ประกอบด้านคุณธรรม จริยธรรม ด้านความรู้ ความเข้าใจ ในวิชาชีพและวิชาเอก ด้านทักษะทางปัญญา ด้านทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ ด้านทักษะการวิเคราะห์เชิงตัวเลข การสื่อสารและการใช้เทคโนโลยีสารสนเทศ และด้านทักษะการจัดการ การเรียนรู้และความสามารถในการวิจัย รายละเอียดดังตาราง 8

ตาราง 8 ผลการวิเคราะห์ข้อมูลองค์ประกอบด้านผลกระทบของหลักสูตร

ด้านผลกระทบของหลักสูตร	μ	σ	ระดับความคิดเห็น
ด้านคุณธรรม จริยธรรม	4.60	0.41	มากที่สุด
ด้านความรู้ ความเข้าใจในวิชาชีพและวิชาเอก	4.70	0.47	มากที่สุด
ด้านทักษะทางปัญญา	4.80	0.50	มากที่สุด
ด้านทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ	4.60	0.50	มากที่สุด
ด้านทักษะการวิเคราะห์เชิงตัวเลขการสื่อสารและการใช้เทคโนโลยีสารสนเทศ	4.60	0.60	มากที่สุด
ด้านทักษะการจัดการเรียนรู้และความสามารถในการวิจัย	4.70	0.47	มากที่สุด
ภาพรวมด้านผลกระทบ	4.67	0.49	มากที่สุด

จากตาราง 8 พบว่า ในภาพรวม และในทุกองค์ประกอบของผลกระทบของหลักสูตร มีความเหมาะสมระดับมากที่สุด ข้อเสนอแนะในประเด็นต่างๆ ดังนี้คือควรมีการส่งเสริมเรื่องของการทำผลงานวิชาการอย่างต่อเนื่องเพื่อพัฒนาตนเองและวิชาชีพต่อไป

มหาบัณฑิตได้นำความรู้ที่ได้จากการสอนและการทำวิทยานิพนธ์ ไปประยุกต์ใช้ในการปฏิบัติงาน มีความเป็นผู้นำเป็นผู้รับผิดชอบในงานด้านต่างและทำโครงการวิจัยทางด้านการจัดการเรียนการสอน วิทยาศาสตร์และคณิตศาสตร์ รายละเอียดตัวอย่างผลการสัมภาษณ์ที่เป็นประเด็นที่น่าสนใจ มีดังนี้

“ผมยังเอาวิธีสอนในวิทยานิพนธ์ไปใช้ตอนสอนอยู่เลยครับ” (มหาบัณฑิต คนที่ 1)

“เราได้ใช้ความรู้เรื่องการวิจัยไปใช้ในการทำวิจัยตอนสอนด้วย” (มหาบัณฑิต คนที่ 2)

“ได้รับมอบหมายงานหลายหน้าที่” (มหาบัณฑิต คนที่ 4)

อภิปรายผล

1. ผลการประเมินหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ (หลักสูตรปรับปรุง พ.ศ. 2560) คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม โดยภาพรวม และรายด้าน พบว่า นิสิต อาจารย์ผู้สอน มหาบัณฑิตและผู้ใช้บัณฑิต มีความคิดเห็นต่อหลักสูตรสอดคล้องกันอยู่ในระดับมาก และภาพรวมด้านผลกระทบของหลักสูตรอยู่ในระดับมากที่สุด เมื่อพิจารณาองค์ประกอบของการประเมินรายด้าน คือ ด้านบริบทพบว่า นิสิต อาจารย์ผู้สอน และมหาบัณฑิตมีความคิดเห็นต่อหลักสูตรในด้านบริบทอยู่ในระดับมาก ด้านที่อยู่ในระดับมากที่สุดคือด้านสาระรายวิชา ส่วนอาจารย์ผู้สอน เห็นว่าหลักสูตรนี้ได้มาตรฐานตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษา ทั้งนี้อาจเป็นเพราะว่าคณะกรรมการผู้ร่างหลักสูตรซึ่งประกอบด้วยผู้เชี่ยวชาญและผู้ทรงคุณวุฒิด้านการศึกษา มีการศึกษาข้อมูลพื้นฐาน สภาพด้านการการสอนวิทยาศาสตร์และคณิตศาสตร์ ตลอดจนแนวโน้มการศึกษา แผนพัฒนาประเทศและเศรษฐกิจ และความต้องการของสังคม ส่งผลให้สามารถนำมากำหนดเป็นวัตถุประสงค์ของหลักสูตรและรายวิชาที่เหมาะสมกับสภาพปัจจุบันในการสอน วิทยาศาสตร์และคณิตศาสตร์ในประเทศไทย สอดคล้องกับผลการวิจัยของ Jarasrawewas et al. (2020)

ที่พบว่าผลการประเมินหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาหลักสูตรและการสอน (หลักสูตรปรับปรุง พ.ศ. 2559) คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา ด้านบริบทมีความเหมาะสมระดับมาก เนื่องจากหลักสูตรมีการกำหนดปรัชญาของหลักสูตร วัตถุประสงค์ของหลักสูตร กลยุทธ์ของหลักสูตร โครงสร้างของหลักสูตรและรายวิชาที่เปิดสอนตามกระบวนการพัฒนาและปรับปรุงหลักสูตรตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษา รวมถึงการวิเคราะห์ข้อมูลพื้นฐาน มาตรฐานและนโยบายต่างๆ เพื่อกำหนดจุดมุ่งหมายของหลักสูตร ด้านปัจจัยนำเข้า พบว่า นิสิต อาจารย์ผู้สอน และมหาบัณฑิตมีความคิดเห็นต่อหลักสูตรในด้านบริบทอยู่ในระดับมากที่สุดมีความคิดเห็นในระดับมาก ด้านที่มีความเห็นระดับมากที่สุดคือด้านอาจารย์ โดยมีความเห็นว่าอาจารย์ผู้สอนมีคุณภาพและมีความรอบรู้ในเนื้อหาสาระวิชาที่สอนเป็นอย่างดีซึ่งเป็นไปตามคุณสมบัติในกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษาแห่งชาติที่กำหนด สอดคล้องกับงานวิจัยของ Booncherdchoo et al. (2017) ทำการประเมินหลักสูตรศึกษาศาสตรบัณฑิต สาขาวิชาการศึกษาศาสตร์บัณฑิต ด้านปัจจัยนำเข้าด้านพฤติกรรมกรรมการจัดการเรียนการสอนของผู้สอนผลการประเมินในภาพรวมและรายด้านมีความเหมาะสมในระดับมากถึงมากที่สุดเรียงตามลำดับจากมากไปหาน้อยดังนี้มีประมวลการสอนตามรายวิชาที่สอนมีบุคลิกภาพเช่นการแต่งกายกิริยาท่าทางการใช้ภาษาเหมาะสมกับครูมีการเตรียมการสอนล่วงหน้าและมีเทคนิคและความสามารถในการถ่ายทอดความรู้ และสอดคล้องกับงานวิจัยของ Chittranun & Wicharapote (2015) ที่ทำการประเมินหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการบริหารการศึกษา (หลักสูตรปรับปรุง พ.ศ. 2554) คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม พบว่าผลการประเมินด้านปัจจัยนำเข้า โดยรวมเห็นว่ามีความเหมาะสมอยู่ในระดับมากจากผลการวิจัยพบว่าอาจารย์ผู้สอนโดยรวมมีความเหมาะสมในระดับมากแสดงให้เห็นว่าอาจารย์มีความรู้ความเชี่ยวชาญและประสบการณ์ในวิชาที่สอนมีความสามารถและเทคนิควิธีการสอนที่ดี ให้คำ แนะนำปรึกษาแนะนำด้านการศึกษาและการเรียนกิจกรรม และด้านการทำ วิจัย ส่งเสริมให้นิสิตคิดวิเคราะห์และกล้าแสดงความคิดเห็น พร้อมทั้งรับฟังและแลกเปลี่ยนความคิดเห็นกับนิสิต ด้านกระบวนการพบว่านิสิต อาจารย์ผู้สอน และมหาบัณฑิตมีความคิดเห็นต่อหลักสูตรในด้านกระบวนการอยู่ในระดับมาก และมีความคิดเห็นในด้านการทำวิทยานิพนธ์อยู่ในระดับมากที่สุด อาจารย์และนิสิต และมหาบัณฑิตมีความเห็นว่าระบบการกำกับติดตามการทำวิทยานิพนธ์ของหลักสูตรเป็นระบบที่ดี นอกจากนี้อาจารย์ที่ปรึกษาวิทยานิพนธ์ยังมีความเชี่ยวชาญในศาสตร์การสอนวิทยาศาสตร์และคณิตศาสตร์ และเอาใจใส่ในการทำวิทยานิพนธ์ของนิสิตอย่างสม่ำเสมอ สอดคล้องกับงานวิจัยของ Nillapun (2012) ทำการประเมินหลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาหลักสูตรและการนิเทศ คณะศึกษาศาสตร์ มหาวิทยาลัยศิลปากร ด้านปัจจัยนำเข้า พบว่าโดยภาพรวมและรายข้ออยู่ในระดับมาก มีแหล่งสืบค้นแสวงหาความรู้ที่ตอบสนองต่อการสืบค้นหาความรู้ มีทรัพยากรการเรียนรู้ที่เอื้อและสนับสนุนผู้เรียนและตรงกับความต้องการในการพัฒนาการเรียนรู้มีความสะดวกในการเลือกใช้นวัตกรรมและเทคโนโลยีเพื่อนำมาประกอบการเรียนรู้ในห้องเรียน สภาพห้องเรียนมีสื่อและวัสดุที่เอื้อต่อการจัดการเรียนรู้ที่หลากหลายมีวิทยากร ผู้ทรงคุณวุฒิภายนอกเข้ามาร่วมในระบบการเรียนรู้เพื่อพัฒนาความรู้และกระบวนการคิดและมีหน่วยงานองค์กรจากภายนอกให้การสนับสนุนในการจัดการเรียนรู้ อาจารย์มีความรู้และประสบการณ์ในการสอน การวิจัยดูแลเอาใจใส่ติดตามในการทำวิทยานิพนธ์ดีมาก ด้านผลผลิตพบว่า นิสิต อาจารย์ผู้สอน และมหาบัณฑิตมีความคิดเห็นต่อหลักสูตรในด้านผลผลิตของหลักสูตร

อยู่ในระดับมากและมีความคิดเห็นในด้านความรู้ความเข้าใจในวิชาชีพและวิชาเอกด้านทักษะทางปัญญาและด้านทักษะการจัดการเรียนรู้และความสามารถในการวิจัยในระดับมากที่สุด นอกจากนี้อาจารย์ผู้สอนเห็นว่าบัณฑิตที่จบการศึกษาเป็นที่ยอมรับของหน่วยงานและบุคคลในวิชาชีพ สอดคล้องกับผลการวิจัยของ Unaromlert et al. (2016) ที่ว่าผลการประเมินหลักสูตรปรัชญาดุษฎีบัณฑิตสาขาวิชาพัฒนศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศิลปากร ด้านผลผลิตหลักสูตร พบว่า ในภาพรวมมีคุณภาพอยู่ในระดับมาก และด้านคุณธรรมจริยธรรมในการทำงานมีค่าเฉลี่ยสูงที่สุด ผลผลิตระหว่างทางของหลักสูตรด้านคุณลักษณะความรู้ ความสามารถ มีคุณลักษณะในการใฝ่รู้ใฝ่เรียนมีความพร้อมในการเรียนรู้ ทักษะและการนำความรู้ไปใช้ประโยชน์อยู่ในเกณฑ์ที่ดีผลผลิตมีคุณภาพและด้านผลกระทบพบว่ามหาวิทยาลัยและผู้ใช้บัณฑิตมีความคิดเห็นต่อหลักสูตรในด้านผลกระทบของหลักสูตรอยู่ในระดับมากที่สุด มหาวิทยาลัยได้นำความรู้ที่ได้ไปใช้ประยุกต์ใช้เพื่อการจัดการเรียนการสอนวิทยาศาสตร์และคณิตศาสตร์และการปฏิบัติงานในหน้าที่ มีการนำความรู้ด้านการทำวิจัยเพื่อใช้ในการแก้ปัญหาในการจัดการเรียนการสอนและการพัฒนาวิชาชีพ สอดคล้องกับงานวิจัยของ Nillapun et al. (2018) ด้านผลผลิต พบว่า นักศึกษามีคุณธรรมในวิชาชีพ และผู้สอนเห็นว่าบัณฑิตที่จบการศึกษาเป็นที่ยอมรับของหน่วยงาน/ชุมชน และบุคคลในวิชาชีพ ทั้งนี้อาจเป็นเพราะว่า นักศึกษาได้รับการอบรม และแนะนำจากอาจารย์ผู้สอนในแต่ละรายวิชาซึ่งมีแนวคิดในการพัฒนาผู้เรียนเป็นไปในทิศทางเดียวกันในทุกด้าน ทั้งด้านความรู้ ทักษะ และคุณธรรมจริยธรรม ที่จะเป็นพื้นฐานให้นักศึกษาปริญญาเอกก้าวสู่การเป็นผู้นำทางวิชาการและวิชาชีพตามความคาดหวังของสังคมโดยมีอาจารย์ผู้สอนเป็นแบบอย่างและสอดคล้องกับผลการวิจัยของ Prewnim (2018) ทำการประเมินหลักสูตรพยาบาลศาสตร์ คณะพยาบาลศาสตร์ มหาวิทยาลัยกรุงเทพธนบุรี ด้านผลกระทบของหลักสูตรพบว่าผู้สำเร็จการศึกษาได้นำความรู้ที่ได้ไปใช้ประยุกต์ใช้เพื่อการปฏิบัติงานมีการนำเสนอการพัฒนาวิชาชีพอยู่ในรูปการพัฒนางานวิจัยเชิงสำรวจและนวัตกรรมทางการพยาบาลมีความเป็นผู้นำสมรรถนะในด้านสุขภาพผ่านโครงการต่างๆแต่จะต้องใช้เวลาในการปรับตัวเมื่อต้องทำงานร่วมกับสมาชิกในทีมทำงานสมรรถนะในด้านสุขภาพผ่านโครงการต่าง ๆ

2. ผลการศึกษาปัญหาและแนวทางการพัฒนาและปรับปรุงหลักสูตรการศึกษามหาบัณฑิต สาขาวิชา การสอนวิทยาศาสตร์และคณิตศาสตร์ คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคามพบว่าด้านบริบท มีบางรายวิชาที่มีความซ้ำซ้อนในเนื้อหา ควรปรับลดความซ้ำซ้อนของบางรายวิชาที่มีเนื้อหาใกล้เคียงกัน ควรออกแบบกิจกรรมการเรียนการสอนให้นิสิตได้ปฏิบัติการทดลองสอนให้มากขึ้น ควรพิจารณาปรับปรุงรายวิชาให้ตรงหรือสอดคล้องกับความต้องการของผู้เรียนและสถานการณ์ปัจจุบัน เช่น นวัตกรรมการสอนแบบออนไลน์ สื่อการสอน และรูปแบบการสอนแบบออนไลน์ ด้านปัจจัยนำเข้า พบว่า ยังมีอุปสรรคทางพื้นที่บริการอินเทอร์เน็ตที่ยังไม่ครอบคลุมทั้งในและนอกห้องเรียน และการเข้าถึงแหล่งเรียนรู้ต่างๆด้านการศึกษา การทำวิจัยวิทยานิพนธ์ รวมถึงการไม่มีพื้นที่สำหรับการเรียนรู้ด้วยตนเองทั้งในและนอกเวลาราชการ ดังนั้นในการพัฒนาหลักสูตร จึงควรให้ความสำคัญกับการบริหารทรัพยากรการเรียนการสอนให้ทันสมัยและเพียงพอรวมถึงการแนะนำ การเข้าถึงแหล่งข้อมูลและแหล่งเรียนรู้ต่างๆด้านการศึกษาการทำวิจัยวิทยานิพนธ์โดยผ่านการประชุม คณะกรรมการผู้ดูแลหลักสูตรเพื่อทบทวนและประเมินผลการจัดการเรียนการสอนในแต่ละภาคการศึกษา ด้านกระบวนการพบว่าบางรายวิชากิจกรรมการเรียนการสอนที่ต้องฝึกปฏิบัติการสอนในสถานศึกษาทำให้นิสิต

บริหารจัดการเวลาในการเรียนได้ค่อนข้างลำบาก ดังนั้นในการพัฒนาหลักสูตรด้านกระบวนการ ควรต้องสร้างความเข้าใจร่วมกันกับนิสิตเพื่อให้รับทราบแนวทางการจัดการเรียนการสอนของหลักสูตรและมีการจัดโครงการพัฒนานิสิตด้านต่างๆอย่างต่อเนื่องตลอดระยะเวลาที่นิสิตศึกษาในหลักสูตร ด้านผลผลิต พบว่า นิสิตยังมีความต้องการให้หลักสูตรจัดการให้ความรู้ในเรื่องการตีพิมพ์และแหล่งข้อมูลในการตีพิมพ์ซึ่งมีความจำเป็นต่อการพัฒนาวิชาชีพต่อไป ดังนั้นหลักสูตรควรจัดกิจกรรมเหล่านี้เสริมเพื่อตอบสนองต่อความต้องการของผู้เรียนและด้านผลกระทบ พบว่า มีความต้องการให้มีการส่งเสริมเรื่องของการทำผลงานวิชาการอย่างต่อเนื่องเพื่อพัฒนาตนเองและวิชาชีพต่อไป ดังนั้นการพัฒนาหลักสูตร ควรให้ความสำคัญกับประเด็นเกี่ยวกับการให้ความรู้ด้านแนวทางการพัฒนาวิชาชีพให้กับนิสิต

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

คณะกรรมการบริหารหลักสูตรการศึกษามหาบัณฑิต สาขาวิชาการสอนวิทยาศาสตร์และคณิตศาสตร์ควรนำผลการประเมินหลักสูตรเป็นแนวทางในการปรับปรุงและพัฒนาหลักสูตรให้มีประสิทธิภาพและทันต่อการเปลี่ยนแปลงของโลกปัจจุบัน โดยเฉพาะด้านการสนับสนุนการเข้าถึงแหล่งเรียนรู้ออนไลน์ต่าง ๆ และ ระบบอินเทอร์เน็ต ซึ่งมีความสำคัญมากในการจัดการเรียนการสอนในสถานการณ์การแพร่ระบาดของเชื้อไวรัสโควิด 19

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 จากการศึกษาจะพบว่านิสิตจะพบปัญหาในการทำวิทยานิพนธ์ ดังนั้นจึงควรทำวิทยานิพนธ์การศึกษามหาบัณฑิตที่สำเร็จการศึกษาในรอบระยะเวลา ในกรณีที่ประสบความสำเร็จที่เป็นต้นแบบเพื่อเป็นการถอดบทเรียนในการดำเนินการพัฒนาหลักสูตรต่อไป

2.2 ในการประเมินหลักสูตรควรเพิ่มกลุ่มผู้ให้ข้อมูลให้ครอบคลุมถึงครูพี่เลี้ยงในการฝึกประสบการณ์วิชาชีพ ผู้บริหาร ผู้ปกครอง เป็นต้น เพื่อให้ได้ข้อมูลที่หลากหลายแง่มุมมากยิ่งขึ้น

Reference

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models and applications*. Jossey - Bass.

Translate Thai References

Booncherdchoo, S. L., Leetho, M., Atthanuphan, C., Komsan, T., Pechkoom, O., & Songsombat, T. (2017). The Curriculum Evaluation on Bachelor Degree of Education Program in Early Childhood Education Faculty of Education, Silpakorn University. *Veridian E-Journal Silpakorn University*, 9(3), 557-572. (in Thai)

- Chittranun, T., & Wicharapote, J. (2015). An Evaluation of the Master Degree Program in Educational Administration (Revised 2011), Faculty of Education, Mahasarakham University by CIPP Model. *Journal of Administration and Development Mahasarakham University*, 7(1), 95-108. (in Thai)
- Jarasrawewas, S., Tongson, P., Prommas, J., Kitipanusun, R., Chompu, W., & Paksripang, P. (2020). An Evaluation of the Master of Education Program in Curriculum and Instruction (Revised A.D.2016), Faculty of Education, Burapha University. *HRD Journal*, 11(2), 8-17. (in Thai)
- Koolnaphadol, P., Wongteeratorn, D., Srichunnil, C., & Inang, P. (2020). The Evaluation Research of Doctor of Philosophy Curriculum in Counseling Psychology Faculty of Education Burapha University. *Journal of Education Research*, 15(2), 133-143. (in Thai)
- Nillapun, M. (2012). The Evaluation of Master of Education Program in Curriculum and Supervision Faculty of Education Silpakorn University. *Silpakorn Educational Research Journal*, 4(2), 25-40. (in Thai)
- Nillapun, M., Sirisumpun, O., Chatiwat, W., Po-ngern, W., & Markjooy, A. (2018). The Curriculum Evaluation on Doctoral of Philosophy Program in Curriculum and Instruction, Faculty of Education. *Silpakorn University Veridian E-Journal Silpakorn University*, 10(2), 1199-1214. (in Thai)
- Paiwithayasiritham, C. (2007). *A Development of the Curriculum Evaluation Model: An Application of a Meta-evaluation* [Doctoral dissertation]. Srinakharinwirot University (in Thai)
- Prewnim, A. (2018). The Curriculum Evaluation of Bachelor of Nursing Science Program, Bangkok Thonburi University in Accordance with the Standard Criteria of National Higher Education with CIPPI Model. *Journal of Nursing and Health Suan Sunandha Rajabhat University*, 1(2), 90-110. (in Thai)
- Srisa-ard, B. (2010). *Introduction to Research (8th Ed.)*. Suweerivasarn. (in Thai)
- Unaromlert, T., Paiwithayasiritham, C., Weeranawin, L., & Sukjairungwattana, T. (2016). Curriculum Evaluation of Doctor of Philosophy in Development Education, Faculty of Education. *Silpakorn University. Veridian E-Journal Silpakorn University*, 9(3), 489-506. (in Thai)

โมเดลสมการโครงสร้างการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผล
การปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4

The Structural Equation Modeling of Diversity Generation Management
Affecting Performance Effectiveness of the Teachers under the Office of
Khon Kaen Primary Education Service Area 4

ณัฐชญา เฟื่องธรรม^{1*} และ เสาวณี สิริสุขศิลป์²

Natchaya Phengtham^{1*} and Saowanee Sirisooksilp²

(Received: June 13, 2022; Revised: July 25, 2022; Accepted: August 1, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ 1) ศึกษาระดับการจัดการความหลากหลายทางเจเนอเรชันและประสิทธิผลการปฏิบัติงานของครู 2) พัฒนาและตรวจสอบความสอดคล้องกลมกลืนของโมเดลสมการโครงสร้างการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียนที่ผู้วิจัยพัฒนาขึ้นกับข้อมูลเชิงประจักษ์ และ 3) ศึกษาแนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ใช้ระเบียบวิธีวิจัยแบบผสมผสาน (mixed methods research) แบบแผนเชิงอธิบาย (explanatory design) ระยะที่ 1 ใช้วิธีเชิงปริมาณ กลุ่มตัวอย่างจำนวน 200 คน ประกอบด้วยผู้บริหารสถานศึกษาและครู โดยใช้วิธีการสุ่มแบบสองขั้นตอน เครื่องมือที่ใช้ในการวิจัยเป็นแบบสอบถามแบบมาตราส่วนประมาณค่า 5 ระดับ มีความเชื่อมั่น เท่ากับ 0.978 วิเคราะห์ข้อมูลโดยใช้สถิติเชิงพรรณนา ค่าเฉลี่ยส่วนเบี่ยงเบนมาตรฐาน และการวิเคราะห์โมเดลสมการโครงสร้าง (SEM) ระยะที่ 2 ใช้วิธีเชิงคุณภาพ เครื่องมือที่ใช้ในการวิจัยเป็นแบบสัมภาษณ์แบบมีโครงสร้าง กลุ่มเป้าหมายได้มาจากการเลือกแบบเจาะจง วิเคราะห์ข้อมูลด้วยวิธีการวิเคราะห์เนื้อหา ผลการวิจัยพบว่า 1) ระดับการจัดการความหลากหลายทางเจเนอเรชันโดยภาพรวมมีค่าเฉลี่ยอยู่ในระดับมาก ($M = 4.42$, $SD = 0.38$) ระดับประสิทธิผลการปฏิบัติงานของครูโดยภาพรวมมีค่าเฉลี่ยอยู่ในระดับมาก ($M = 4.29$, $SD = 0.24$) 2) โมเดลสมการโครงสร้างการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียนมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ การจัดการความหลากหลายทางเจเนอเรชันมีอิทธิพลทางตรงเชิงบวกต่อ

¹ นักศึกษาปริญญาโท สาขาวิชาการบริหารการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² รองศาสตราจารย์ สาขาวิชาการบริหารการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

¹ Master's degree student, Educational Administration Program, Faculty of Education, Khon Kaen University

² Associate Professor, Educational Administration Program, Faculty of Education, Khon Kaen University

* Corresponding Author E-mail: p.natchaya@kkumail.com

ประสิทธิผลการปฏิบัติงานของครูในโรงเรียน มีขนาดอิทธิพลรวมเท่ากับ 0.796 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยตัวแปรการจัดการความหลากหลายทางเจเนอเรชันในโมเดลสามารถอธิบายความแปรปรวนของประสิทธิผลการปฏิบัติงานของครูในโรงเรียนได้ร้อยละ 63.3 และ 3) แนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน ประกอบด้วย 5 องค์ประกอบ ได้แก่ (1) การพัฒนาระบบพี่เลี้ยงในการสอนงาน ผู้บริหารส่งเสริมให้มีกิจกรรมการสอนงาน จัดระบบคู่พี่เลี้ยงที่มีประสบการณ์ (2) การสื่อสารเพื่อสร้างสัมพันธภาพ มีช่องทางการสื่อสารหลากหลาย นำเทคโนโลยีสารสนเทศมาใช้ในการสื่อสาร เปิดใจและยอมรับความคิดเห็นที่แตกต่าง (3) การธำรงรักษาบุคลากรในองค์กร ผู้บริหารสร้างแรงจูงใจและให้ผลตอบแทนในการปฏิบัติงานที่ดีแก่ครู มอบโอกาสความก้าวหน้าและการใช้ศักยภาพ (4) การส่งเสริมการทำงานเป็นทีม สร้างทีมงานที่หลากหลายที่ประกอบด้วยบุคลากรทุกช่วงวัย มอบหมายให้ครูทุกช่วงวัยได้มีโอกาสเป็นผู้นำทีมงาน และ (5) การสนับสนุนการทำงานแบบยืดหยุ่น ปรับเปลี่ยนวิธีการทำงานให้ยืดหยุ่นสอดคล้องกับลักษณะการทำงานของบุคลากรแต่ละช่วงวัยอย่างเหมาะสม

คำสำคัญ: การจัดการความหลากหลาย เจเนอเรชัน ประสิทธิผลการปฏิบัติงานของครู โมเดลสมการโครงสร้าง

Abstract

The objectives of this research were to 1) study the level of diversity generation management and the level of performance effectiveness of teachers; 2) develop and examine the goodness of fit of the structural equation model of diversity generation management affecting the performance effectiveness of the teachers and the empirical data, that the author had constructed; and 3) study the guideline of diversity generation management affecting the performance effectiveness of the teachers under the Office of Khon Kaen Primary Education Service Area 4. The research employed the Mixed Methods Research: Explanatory Design. Phase 1 used a quantitative method. The sample consisted of 200 educational institution administrators and teachers, derived through two-stage cluster sampling method. The research instrument was a 5-point rating scale questionnaire with the reliability of 0.978. The data analysis employed the mean, standard deviation, and structural equation modeling (SEM). Phase 2 used a qualitative method. The research instrument was a structured interview. The target group consisted of educational institution administrators who were purposively sampled. The data were analyzed using content analysis. The findings showed that 1) the diversity generation management, on the whole, revealed a high level of action ($M = 4.42$, $S.D. = 0.38$). The performance effectiveness of teachers, on the whole, revealed a high level of action ($M = 4.42$, $S.D. = 0.24$). 2) The structural equation model of diversity generation management affecting the performance effectiveness of the teachers under the Office of Khon Kaen Primary

Education Service Area 4, developed by the researcher, had the goodness of fit with the empirical data. The diversity generation management affecting the performance effectiveness of the teachers had direct positive influence on the performance effectiveness of the teachers, with the total influence of 0.796 at the .01 level of statistical significance. The variable of diversity generation management in the model could explain the performance effectiveness of the teachers for 63.3%. And, 3) The guideline of diversity generation management affecting the performance effectiveness of the teachers consists of 5 components: (1) In developing a mentoring system in coaching, the administrator promotes coaching activities, organizes a mentoring system with experienced-mentor pairing; (2) In communication for relationship building, various communication channels are provided, information technology is used in communication, and different opinions are welcomed; (3) In personnel retention, the administrator motivates and rewards teachers for good performance, provides opportunities for advancement and the use of potentials; (4) In encouraging teamwork, teams with age diversity are built, teachers of different ages have the opportunity to be assigned as a team leader; and (5) In supporting flexibility in working, work methods are adjusted to suit the work characteristics of teachers in various ages.

Keywords: diversity management, generation, performance effectiveness of teachers, structural equation model

บทนำ

การเปลี่ยนแปลงอย่างรวดเร็วของสภาพแวดล้อมทางเศรษฐกิจและสังคม รวมทั้งความก้าวหน้าทางด้านเทคโนโลยีได้เข้ามามีอิทธิพลต่อการดำเนินงานขององค์กร ทำให้เกิดการแข่งขันระหว่างองค์กรต่าง ๆ ในระดับสูง ทุกองค์กรจึงต้องมีการเปลี่ยนแปลงและปรับตัวอยู่ตลอดเวลาเพื่อให้เข้ากับสถานการณ์ปัจจุบัน การที่องค์กรจะกระทำการกิจหลักให้บรรลุวัตถุประสงค์และมีประสิทธิภาพนั้น องค์กรจำเป็นจะต้องมีบุคลากรที่มีคุณภาพ มีความเหมาะสมกับงาน (Soisud, 2021) ดังนั้นจึงมีความจำเป็นที่จะต้องเร่งสร้างคนเพื่อยกระดับสมรรถนะบุคลากรให้สามารถตอบสนองต่อสภาพแวดล้อมของสังคมโลกที่เปลี่ยนไป โดยต้องให้โอกาสแก่ความหลากหลายของทรัพยากรมนุษย์อย่างเท่าเทียม ทั้งในด้านวัฒนธรรม ชาติพันธุ์ เชื้อชาติ อายุ ศาสนา เพศ และความเชื่อ (Ardakani et al., 2015) เมื่อบุคลากรในองค์กรมาทำงานร่วมกัน สิ่งหนึ่งที่เกิดขึ้นอย่างหลีกเลี่ยงไม่ได้คือ ความหลากหลาย (diversity) ในการปฏิสัมพันธ์กัน ทั้งในทิศทางเดียวกันหรือตรงกันข้าม ดังนั้นการบริหารจัดการความหลากหลาย (diversity management) ของบุคลากรในองค์กรจึงกลายเป็นกลยุทธ์สำคัญที่จะทำให้องค์กรมีความได้เปรียบหรือเสียเปรียบในการแข่งขันที่ผู้บริหารองค์กรมีอาจจะเลยได้

ในปัจจุบันองค์การหลายแห่งกำลังเผชิญปัญหาช่องว่างระหว่างวัยของพนักงานในองค์การ ซึ่งมีช่วงอายุหรือเรียกว่า เจเนอเรชัน (generation) แตกต่างกัน ทำให้ความคิด ความคาดหวังและมุมมองในด้านต่าง ๆ แตกต่างกันอย่างชัดเจน (Wiriyapipat, 2016) การที่มีความหลากหลายทำให้เกิดข้อได้เปรียบเสียเปรียบในองค์การยุคปัจจุบัน โดยในทางที่ดีนั้นความหลากหลายของบุคลากรถือเป็นคุณค่า มีความสำคัญและเป็นประโยชน์ต่อองค์การ เพราะความหลากหลายนำมาซึ่งความสามารถที่แตกต่าง ความคิดเห็นที่แตกต่าง จนกลายเป็นความคิดสร้างสรรค์ในการทำงาน รวมทั้งการมองเห็นปัญหาและการแก้ไขปัญหาที่หลากหลายสามารถป้องกันความผิดพลาดจากการทำงานได้อย่างแม่นยำ แต่ในทางตรงกันข้ามนั้นหากองค์การไม่สามารถจัดการและปรับตัวกับความหลากหลายที่เกิดขึ้นได้อย่างเหมาะสม ความหลากหลายนี้อาจกลายเป็นความแตกแยก ทำให้การทำงานมีแต่ข้อผิดพลาดหลาย ๆ อย่างเกิดขึ้น ส่งผลโดยตรงต่อประสิทธิภาพและประสิทธิผลขององค์การ (Thumviriyavong, 2016) ผู้บริหารจึงเป็นบุคคลสำคัญที่มีบทบาทในการจัดการความหลากหลายของบุคลากรในองค์การ การเข้าใจทฤษฎีเจเนอเรชันของบุคลากรในองค์การย่อมเป็นเครื่องมือให้ผู้บริหารรู้จักใช้ทรัพยากรมนุษย์ในองค์การได้อย่างมีประสิทธิภาพ

การจัดการความหลากหลายทางเจเนอเรชันของบุคลากรมีความสำคัญอย่างมากต่อองค์การที่ต้องอาศัยการทำงานเป็นทีมหรือมีลักษณะงานที่ต่อเนื่องอย่างองค์การทางการศึกษาในระดับโรงเรียน ซึ่งส่วนใหญ่พบว่าในแต่ละโรงเรียนจะเริ่มรับบุคลากรที่มีอายุตั้งแต่ 20 ปี ขึ้นไปและจะพ้นสภาพการเป็นบุคลากรหรือเกษียณอายุราชการเมื่ออายุครบ 60 ปีบริบูรณ์ เมื่อนำช่วงปี พ.ศ. ที่เกิดมาเทียบกับทฤษฎีเจเนอเรชัน พบว่าบุคลากรในโรงเรียนจะประกอบไปด้วย 3 รุ่น ได้แก่ Baby Boomers (คนที่เกิดระหว่าง พ.ศ. 2489-2507) Generation X (คนที่เกิดระหว่าง พ.ศ. 2508-2522) Generation Y (คนที่เกิดระหว่าง พ.ศ. 2523-2540) (Chueabangkaew, 2014) อายุนับเป็นปัจจัยหลักที่แสดงออกถึงความแตกต่างได้อย่างชัดเจน โดยช่วงวัยที่แตกต่างทำให้มีลักษณะการทำงาน การแสดงความคิดเห็นและการแสดงออกที่ต่างกันไป ดังนั้นถ้าผู้บริหารทราบถึงความแตกต่างในแต่ละกลุ่มอายุ จะช่วยให้สามารถวางแผนการพัฒนาคนให้เหมาะสมกับงานและมอบหมายงานได้ตรงกับความถนัด ในการศึกษาครั้งนี้ ผู้วิจัยได้นำแนวคิดของ Rosado (2006); Chueabangkaew (2014) เป็นหลักในการกำหนดกรอบแนวคิดการวิจัยร่วมกับนักวิชาการท่านอื่น ๆ เพื่อให้ได้องค์ประกอบของการจัดการความหลากหลายทางเจเนอเรชันประกอบด้วย 5 องค์ประกอบ ได้แก่ 1) การส่งเสริมการทำงานเป็นทีม 2) การสื่อสารเพื่อสร้างสัมพันธภาพ 3) การพัฒนาระบบพี่เลี้ยงในการสอนงาน 4) การสนับสนุนการทำงานแบบยืดหยุ่นและ 5) การธำรงรักษามูลค่าบุคลากรในองค์การ

ประสิทธิผลการปฏิบัติงานของครูเป็นผลสำเร็จหรือผลลัพธ์ที่ครูสามารถปฏิบัติงานให้บรรลุตามวัตถุประสงค์และเป้าหมายที่กำหนดไว้อย่างครบถ้วนสมบูรณ์ โดยการใช้ประโยชน์จากทรัพยากรที่มีอยู่ให้เกิดประโยชน์สูงสุด (Barnard, 1968) ซึ่งเป้าหมายนั้นคือคุณภาพทางการศึกษาและคุณลักษณะที่พึงประสงค์ของนักเรียน ทั้งนี้ย่อมแสดงว่าการบริหารงานในสถานศึกษานั้นบังเกิดประสิทธิผลด้วยเช่นกัน ในการศึกษาครั้งนี้ ผู้วิจัยได้นำแนวคิดของ Schermerhorn et al. (2000); Hoy et al. (2013); Office of the Teachers council of Thailand (2013); Luksanapawan & Somprach (2019) เป็นหลักในการกำหนดกรอบแนวคิดการวิจัยร่วมกับนักวิชาการท่านอื่น ๆ เพื่อให้ได้องค์ประกอบของประสิทธิผลการปฏิบัติงานของครู ซึ่งประกอบด้วย

5 องค์ประกอบ ได้แก่ 1) คุณภาพผู้เรียน 2) การจัดการเรียนรู้ 3) ความพึงพอใจในงาน 4) การพัฒนาวิชาชีพ และ 5) ความผูกพันต่อองค์กร ผู้นำที่มีประสิทธิภาพต้องสามารถทำให้ผู้ร่วมงานทำงานได้อย่างมีประสิทธิภาพและเต็มความสามารถ การบริหารจัดการความหลากหลายทางเจเนอเรชันจึงเป็นเครื่องมือสำคัญของผู้บริหารในการพัฒนาและบริหารคุณลักษณะที่แตกต่างกันออกไปของบุคลากรในโรงเรียน องค์การที่มีการจัดการความหลากหลายที่ดีจะสามารถดึงดูดบุคลากรที่มีความสามารถเข้าสู่องค์กร เพิ่มแรงจูงใจ รักษาบุคลากรให้อยู่กับองค์กรได้นานขึ้นและใช้ความหลากหลายให้เกิดประโยชน์แก่องค์กร โดยหากมีการบริหารความหลากหลายเหล่านี้นี้ให้มีความเหมาะสมเป็นที่ยอมรับ ย่อมเกิดพฤติกรรมการทำงานที่อาจก่อให้เกิดประสิทธิผลกับการปฏิบัติงาน (Ninaroon, 2017)

โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ประกอบด้วย โรงเรียนขนาดเล็ก ขนาดกลางและขนาดใหญ่ บุคลากรในโรงเรียนส่วนใหญ่มีความหลากหลายทางเจเนอเรชันและมีความแตกต่างทางด้านมุมมอง ความคิด พฤติกรรมและรูปแบบการปฏิบัติงาน ก่อให้เกิดการแสดงออกด้านความรับผิดชอบ ความสนใจ ความรอบรู้ ความสามารถในการแก้ไขปัญหา ความมีเหตุผล และวุฒิภาวะด้านอื่น ๆ ที่แตกต่างกัน ผลการดำเนินงานการขับเคลื่อนการจัดการศึกษาที่ผ่านมาพบว่าแนวโน้มผลการพัฒนาการศึกษาลดลง (The Office Of Khon Kaen Primary Educational Service Area 4, 2021) โดยเฉพาะปัญหาในการบริหารและการจัดการเรียนการสอนหลายประการ เช่น งบประมาณไม่เพียงพอ ครุยังขาดนวัตกรรมการเรียนการสอนที่เหมาะสม ขาดแรงจูงใจและการติดต่อประสานงานกันภายในโรงเรียนอย่างต่อเนื่อง ซึ่งเป็นส่วนสำคัญที่ทำให้ผลสัมฤทธิ์ทางการเรียนไม่บรรลุเป้าหมาย จากเหตุผลดังกล่าวผู้วิจัยจึงสนใจที่จะศึกษาโมเดลความสัมพันธ์เชิงสาเหตุของตัวแปรแฝงการจัดการความหลากหลายทางเจเนอเรชันและประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 และเพื่อให้ได้แนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 เนื่องจากแนวทางที่ได้จากการวิจัยจะเป็นประโยชน์สำหรับการบริหารจัดการความหลากหลายทางเจเนอเรชันของครูและบุคลากรในโรงเรียนให้เกิดประสิทธิภาพและประสิทธิผลต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาระดับการจัดการความหลากหลายทางเจเนอเรชันและประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4
2. เพื่อพัฒนาและตรวจสอบความสอดคล้องกลมกลืนของโมเดลสมการโครงสร้างการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4
3. เพื่อศึกษาแนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยแบบผสมผสาน (mixed methods research) เป็นแบบแผนเชิงอธิบาย (explanatory design) โดยระยะที่ 1 ใช้วิธีเชิงปริมาณ และระยะที่ 2 ใช้วิธีเชิงคุณภาพ (Creswell & Clark, 2011) โดยได้กำหนดรายละเอียดวิธีการดำเนินการวิจัย ดังนี้

1. ระยะที่ 1 การศึกษาเชิงปริมาณ

1.1 ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย คือ ผู้บริหารสถานศึกษาและครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ปีการศึกษา 2564 จำนวน 175 โรงเรียน ประกอบด้วยผู้บริหารสถานศึกษา 175 คน และครู 1,242 คน รวมทั้งหมดจำนวน 1,417 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ ผู้บริหารสถานศึกษาและครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ปีการศึกษา 2564 จำนวน 100 โรงเรียน โรงเรียนละ 2 คน จำนวน 200 คน เนื่องจากการวิจัยในครั้งนี้ผู้วิจัยใช้สถิติขั้นสูงในการวิเคราะห์ข้อมูลโมเดลสมการโครงสร้าง (structural equation modeling : SEM) จึงกำหนดขนาดตัวอย่างเป็น 20 เท่าของจำนวนตัวแปรสังเกตได้ซึ่งเป็นจำนวนที่เหมาะสมในการวิเคราะห์โมเดล (Hair et al., 1998 as cited in Wiratchai, 1999) ในการวิจัยครั้งนี้มีตัวแปรสังเกตได้จำนวน 10 ตัวแปร จึงได้ขนาดของกลุ่มตัวอย่างเท่ากับ 200 คน ใช้วิธีการสุ่มแบบสองขั้นตอน (two-stage cluster sampling) ขั้นตอนที่ 1 เป็นการสุ่มแบบแบ่งชั้น (stratified random sampling) โดยแบ่งชั้นประชากรตามขนาดโรงเรียน ได้จำนวน 100 โรงเรียน ขั้นตอนที่ 2 ใช้การสุ่มอย่างง่าย (simple random sampling) เพื่อสุ่มผู้ให้ข้อมูลตามสัดส่วนประชากรและกลุ่มตัวอย่างเก็บข้อมูลโดยใช้มุมมองของผู้บริหารสถานศึกษาและครูแต่ละโรงเรียน โรงเรียนละ 2 คน รวมทั้งสิ้น 200 คน

1.2 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย เป็นแบบสอบถามแบบมาตราส่วนประมาณค่า 5 ระดับ (rating scale) เป็นแบบสอบถามที่ผู้วิจัยสร้างขึ้น ตรวจสอบความตรงเชิงเนื้อหาโดยผู้ทรงคุณวุฒิ 5 คน มีค่า IOC อยู่ระหว่าง 0.60-1.00 จากการทดลองใช้ (try out) กับผู้บริหารสถานศึกษาและครูสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 2 จำนวน 40 ฉบับ แบบสอบถามทั้งฉบับมีความเชื่อมั่นสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach, 1970) เท่ากับ 0.978 โดยแบ่งเป็น 3 ตอน ดังนี้

ตอนที่ 1 เป็นแบบสอบถามสถานภาพทั่วไปของผู้ตอบแบบสอบถาม

ตอนที่ 2 เป็นแบบสอบถามระดับการจัดการความหลากหลายทางเจเนอเรชัน ประกอบด้วย 5 ด้าน ได้แก่ 1) การส่งเสริมการทำงานเป็นทีม 2) การสื่อสารเพื่อสร้างสัมพันธภาพ 3) การพัฒนาระบบพี่เลี้ยงในการสอนงาน 4) การสนับสนุนการทำงานแบบยืดหยุ่น และ 5) การจรรีกรักษาบุคลากรในองค์กร จากการทดลองใช้ (try out) มีค่าความเชื่อมั่นสัมประสิทธิ์แอลฟาของครอนบาคเท่ากับ 0.969

ตอนที่ 3 เป็นแบบสอบถามระดับประสิทธิผลการปฏิบัติงานของครู ประกอบด้วย 5 ด้าน ได้แก่ 1) คุณภาพผู้เรียน 2) การจัดการเรียนรู้ 3) ความพึงพอใจในงาน 4) การพัฒนาวิชาชีพ และ 5) ความผูกพันต่อองค์กร จากการทดลองใช้ (try out) มีค่าความเชื่อมั่นสัมประสิทธิ์แอลฟาของครอนบาคเท่ากับ 0.944

1.3 การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลด้วยแบบสอบถาม โดยการส่งไปรษณีย์หนังสือราชการแบบสอบถามในรูปแบบลิงก์หรือ QR Code แบบสอบถามออนไลน์โดยใช้ Google Forms ไปยังโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ที่เป็นกลุ่มตัวอย่าง 110 โรงเรียน ในระหว่างวันที่ 28 มีนาคม ถึง 30 เมษายน 2565 จากการเก็บรวบรวมข้อมูลผู้วิจัยได้รับแบบสอบถามที่มีความสมบูรณ์กลับคืนมาทาง Google forms จำนวน 212 ฉบับ คิดเป็นร้อยละ 96.36 ของจำนวนกลุ่มตัวอย่างทั้งหมด และรวบรวมเพื่อนำไปใช้ในการวิเคราะห์ข้อมูล

1.4 การวิเคราะห์ข้อมูล

ผู้วิจัยได้วิเคราะห์ข้อมูลซึ่งมีรายละเอียดดังนี้ 1. ค่าเฉลี่ยเลขคณิต (M) 2. ส่วนเบี่ยงเบนมาตรฐาน (SD) 3. ค่าสถิติทดสอบความสอดคล้องกลมกลืนของโมเดลสมการโครงสร้างกับข้อมูลเชิงประจักษ์ ตามเกณฑ์ของ Hair et al. (2010) และ 4. ค่าสัมประสิทธิ์น้ำหนักองค์ประกอบในรูปคะแนนมาตรฐานของตัวบ่งชี้ในโมเดลสมการโครงสร้าง (β) โดยใช้โปรแกรมสำเร็จรูป SPSS และ Mplus Editor Version 7

2. ระยะที่ 2 การศึกษาเชิงคุณภาพ

2.1 กลุ่มเป้าหมาย

2.1.1 ผู้ให้ข้อมูลสำหรับการตรวจสอบและยืนยันแนวทาง

ผู้วิจัยใช้วิธีการสัมภาษณ์แบบมีโครงสร้าง (structured interview) จึงใช้วิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง (purposive sampling) โดยเป็นผู้ทรงคุณวุฒิจำนวน 3 คน ได้แก่ ผู้บริหารสถานศึกษา แบ่งตามขนาดโรงเรียน ขนาดเล็ก ขนาดกลางและขนาดใหญ่ ซึ่งมีเกณฑ์ในการพิจารณา คือ 1) ดำรงตำแหน่งผู้บริหารสถานศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 2) มีประสบการณ์ในการบริหารสถานศึกษาไม่ต่ำกว่า 5 ปี และ 3) มีวุฒิการศึกษาไม่ต่ำกว่าปริญญาโท

2.1.2 ผู้ให้ข้อมูลสำหรับการประเมินแนวทาง

ผู้วิจัยใช้แบบประเมินความเหมาะสม ความเป็นไปได้ และความเป็นประโยชน์ของแนวทาง โดยผู้ทรงคุณวุฒิ จำนวน 5 คน ได้แก่ ผู้บริหารการศึกษา จำนวน 1 คน ผู้บริหารสถานศึกษา จำนวน 3 คน และศึกษานิเทศก์ จำนวน 1 คน ซึ่งมีเกณฑ์ในการพิจารณา คือ มีประสบการณ์ในตำแหน่งไม่น้อยกว่า 5 ปี และมีวุฒิการศึกษาไม่ต่ำกว่าปริญญาโท

2.2 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยระยะที่ 2 ประกอบด้วย

2.2.1 แบบสัมภาษณ์แบบมีโครงสร้างเพื่อยืนยันแนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิภาพการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ที่ผู้วิจัยร่างขึ้นจากผลการวิจัยระยะที่ 1 ประกอบด้วย 5 ด้าน ดังนี้ คือ 1) การพัฒนาระบบพี่เลี้ยงในการสอนงาน 2) การสื่อสารเพื่อสร้างสัมพันธภาพ 3) การส่งเสริมการทำงานเป็นทีม 4) การดำรงรักษาบุคลากรในองค์กร และ 5) การสนับสนุนการทำงานแบบยืดหยุ่น นำแบบสัมภาษณ์แบบมีโครงสร้างที่สร้างขึ้นเสนออาจารย์ที่ปรึกษา ตรวจสอบความตรงของเนื้อหาและความเหมาะสมของข้อคำถาม

2.2.2 แบบประเมินความเหมาะสม ความเป็นไปได้ และความเป็นประโยชน์ของแนวทาง ซึ่งมีลักษณะเป็นแบบสอบถามมาตราประมาณค่า (rating scale) 5 ระดับ ตรวจสอบความตรงของเนื้อหาและความเหมาะสมของข้อคำถามโดยอาจารย์ที่ปรึกษา

2.3 การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลด้วยแบบสัมภาษณ์แบบมีโครงสร้าง และแบบประเมินความเหมาะสม ความเป็นไปได้ และความเป็นประโยชน์ของแนวทาง ผู้วิจัยทำการสัมภาษณ์ด้วยตนเอง

2.4 การวิเคราะห์ข้อมูล

ข้อมูลที่ได้จากการสัมภาษณ์เป็นข้อมูลเชิงคุณภาพ ผู้วิจัยจึงใช้วิธีการวิเคราะห์เนื้อหา (content analysis) ข้อมูลที่ได้จากแบบประเมินความเหมาะสม ความเป็นไปได้ และความเป็นประโยชน์ เป็นข้อมูลเชิงปริมาณ ผู้วิจัยจะใช้ค่าเฉลี่ย (M) และส่วนเบี่ยงเบนมาตรฐาน (SD)

ผลการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยแบบผสมผสาน (mixed methods research) โดยระยะที่ 1 ใช้วิธีเชิงปริมาณเพื่อพัฒนาและตรวจสอบความสอดคล้องกลมกลืนของโมเดล และระยะที่ 2 ใช้วิธีเชิงคุณภาพเพื่อศึกษาแนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิภาพการปฏิบัติงานของครูในโรงเรียน ซึ่งผู้วิจัยจะนำเสนอผลการวิเคราะห์ข้อมูลในแต่ละวัตถุประสงค์ ดังนี้

1. ผลการศึกษาระดับการจัดการความหลากหลายทางเจเนอเรชันและประสิทธิภาพการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4

1.1 ผลการวิเคราะห์การจัดการความหลากหลายทางเจเนอเรชันของผู้บริหารสถานศึกษา ในภาพรวม มีค่าเฉลี่ยอยู่ในระดับ “มาก” ($M = 4.42$, $SD = 0.38$) และเมื่อพิจารณาเป็นรายด้าน พบว่า ด้านการดำรงรักษาบุคลากรในองค์กร มีค่าเฉลี่ยสูงสุด ($M = 4.55$, $SD = 0.43$) รองลงมาคือ ด้านการสื่อสารเพื่อสร้างสัมพันธภาพ ($M = 4.51$, $SD = 0.43$) และด้านที่มีค่าเฉลี่ยต่ำสุดคือ การพัฒนาระบบพี่เลี้ยงในการสอนงาน ($M = 4.18$, $SD = 0.40$) แสดงในตาราง 1

ตาราง 1 ค่าเฉลี่ย (M) และส่วนเบี่ยงเบนมาตรฐาน (SD) แสดงระดับการจัดการความหลากหลายทางเจเนอเรชัน โดยภาพรวม

ด้านที่	การจัดการความหลากหลายทางเจเนอเรชัน	ระดับปฏิบัติ		
		M	SD	แปลผล
1	การส่งเสริมการทำงานเป็นทีม	4.50	0.41	มาก
2	การสื่อสารเพื่อสร้างสัมพันธภาพ	4.51	0.43	มากที่สุด
3	การพัฒนาระบบพี่เลี้ยงในการสอนงาน	4.18	0.40	มาก
4	การสนับสนุนการทำงานแบบยืดหยุ่น	4.34	0.45	มาก
5	การดำรงรักษาบุคลากรในองค์กร	4.55	0.43	มากที่สุด
รวม		4.42	0.38	มาก

1.2 ผลการวิเคราะห์ประสิทธิผลการปฏิบัติงานของครูในโรงเรียนในภาพรวมมีค่าเฉลี่ยอยู่ในระดับ “มาก” ($M = 4.29$, $SD = 0.24$) และเมื่อพิจารณาเป็นรายด้าน พบว่า ด้านที่มีค่าเฉลี่ยสูงสุด คือ ด้านการจัดการเรียนรู้ ($M = 4.52$, $SD = 0.31$) รองลงมาคือ ด้านการพัฒนาวิชาชีพ ($M = 4.51$, $SD = 0.35$) ส่วนด้านที่มีค่าเฉลี่ยต่ำสุดคือ ด้านคุณภาพผู้เรียน ($M = 3.69$, $SD = 0.42$) แสดงในตาราง 2

ตาราง 2 ค่าเฉลี่ย (M) และส่วนเบี่ยงเบนมาตรฐาน (SD) แสดงระดับประสิทธิผลการปฏิบัติงานของครู โดยภาพรวม

ด้านที่	ประสิทธิผลการปฏิบัติงานของครู	ระดับปฏิบัติ		
		M	SD	แปลผล
1	คุณภาพผู้เรียน	3.69	0.42	มาก
2	การจัดการเรียนรู้	4.52	0.31	มากที่สุด
3	ความพึงพอใจในงาน	4.42	0.38	มาก
4	การพัฒนาวิชาชีพ	4.51	0.35	มากที่สุด
5	ความผูกพันต่อองค์กร	4.33	0.37	มาก
รวม		4.29	0.24	มาก

2. ผลการพัฒนาและตรวจสอบความสอดคล้องกลมกลืนของโมเดลสมการโครงสร้างการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียนที่ผู้วิจัยพัฒนาขึ้นกับข้อมูลเชิงประจักษ์

2.1 ผลการตรวจสอบความสอดคล้องกลมกลืนโมเดลการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ที่ผู้วิจัยพัฒนาขึ้นกับข้อมูลเชิงประจักษ์ พบว่ามีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ ซึ่งค่าสถิติผ่านตามเกณฑ์แสดงว่าโมเดลการวิจัยสอดคล้องกับข้อมูลเชิงประจักษ์ตามเงื่อนไขระดับการยอมรับตามเกณฑ์ของ Hair et al. (2010) แสดงดังตาราง 3

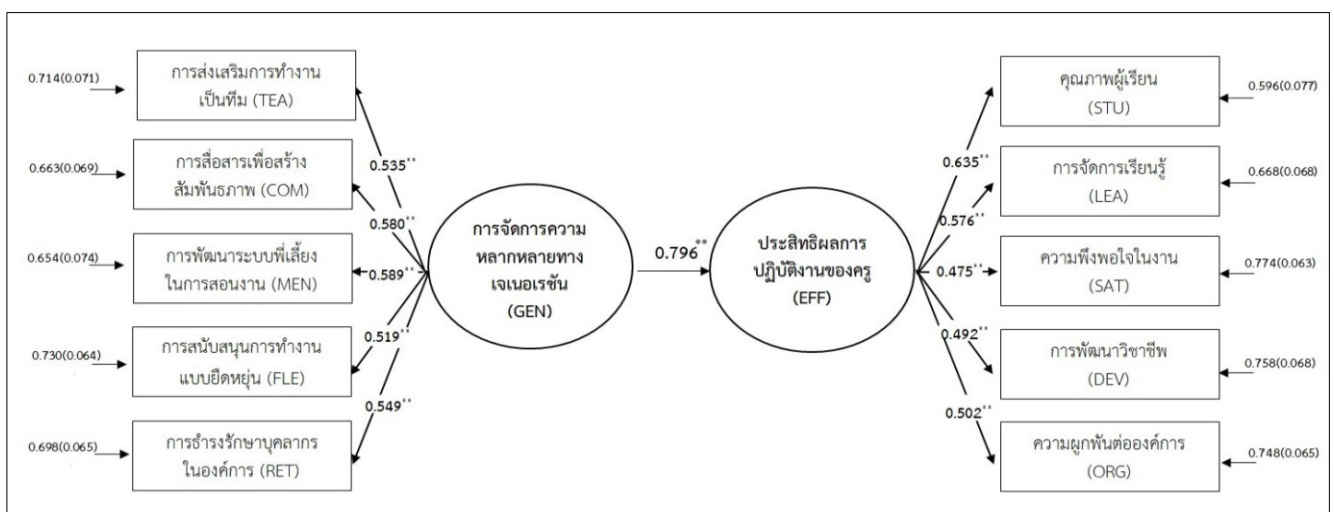
ตาราง 3 ค่าสถิติความสอดคล้องกลมกลืนของการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียนกับข้อมูลเชิงประจักษ์ ตามเกณฑ์การพิจารณา

สถิติที่ใช้วัดความกลมกลืน	เกณฑ์พิจารณา	ค่าสถิติของโมเดล	ผลการพิจารณา
p-value of χ^2	สูงกว่า 0.05	0.068	ผ่านเกณฑ์
χ^2/df	น้อยกว่า 2.00	1.425	ผ่านเกณฑ์
CFI	มากกว่า 0.95	0.970	ผ่านเกณฑ์
TLI	มากกว่า 0.95	0.952	ผ่านเกณฑ์
RMSEA	น้อยกว่า 0.05	0.044	ผ่านเกณฑ์
SRMR	น้อยกว่า 0.08	0.038	ผ่านเกณฑ์
จำนวนการปรับโมเดล		6	-

2.2 ผลการศึกษาอิทธิพลของการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4

2.2.1 ผลการวิเคราะห์ค่าอิทธิพลของการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 พบว่า การจัดการความหลากหลายทางเจเนอเรชัน (GEN) มีอิทธิพลรวมและอิทธิพลทางตรงเชิงบวกเท่ากับ 0.796 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 หมายความว่า ถ้าการจัดการความหลากหลายทางเจเนอเรชัน (GEN) เพิ่มขึ้น 1 หน่วย จะส่งผลให้ประสิทธิผลการปฏิบัติงานของครู (EFF) เพิ่มขึ้น 0.796 หน่วย โดยตัวแปรการจัดการความหลากหลายทางเจเนอเรชันในโมเดลสามารถอธิบายความแปรปรวนในการเพิ่มขึ้นของตัวแปรประสิทธิผลการปฏิบัติงานของครูได้ร้อยละ 63.3 ($R^2 = 0.633$)

2.2.2 ผลการวิเคราะห์ค่าสัมประสิทธิ์น้ำหนักองค์ประกอบในรูปคะแนนมาตรฐานของโมเดลการวัดการจัดการความหลากหลายทางเจเนอเรชัน มีนัยสำคัญทางสถิติที่ระดับ .01 ทุกตัวแสดงให้เห็นว่าองค์ประกอบทั้ง 5 ตัวนี้เป็นองค์ประกอบที่สำคัญที่สามารถบ่งบอกถึงการจัดการความหลากหลายทางเจเนอเรชันของผู้บริหารสถานศึกษาทุกตัว ซึ่งองค์ประกอบที่สามารถอธิบายการจัดการความหลากหลายทางเจเนอเรชันได้ดีที่สุด คือ ด้านการพัฒนาระบบพี่เลี้ยงในการสอนงาน ($\beta = 0.589$) รองลงมาคือ ด้านการสื่อสารเพื่อสร้างสัมพันธ์ภาพ ($\beta = 0.580$) ส่วนด้านที่มีค่าสัมประสิทธิ์น้ำหนักองค์ประกอบต่ำที่สุดคือ ด้านการสนับสนุนการทำงานแบบยืดหยุ่น ($\beta = 0.519$) ค่าสัมประสิทธิ์น้ำหนักองค์ประกอบในรูปคะแนนมาตรฐานของตัวบ่งชี้ในโมเดลการวัดประสิทธิผลการปฏิบัติงานของครูในโรงเรียน มีนัยสำคัญทางสถิติที่ระดับ .01 ทุกตัว แสดงให้เห็นว่าองค์ประกอบทั้ง 5 ตัวนี้เป็นองค์ประกอบที่สำคัญที่สามารถบ่งบอกถึงประสิทธิผลการปฏิบัติงานของครูในโรงเรียนทุกตัว ซึ่งองค์ประกอบที่สามารถอธิบายประสิทธิผลการปฏิบัติงานของครูได้ดีที่สุด คือ ด้านคุณภาพผู้เรียน ($\beta = 0.635$) รองลงมาคือ ด้านการจัดการเรียนรู้ ($\beta = 0.576$) ส่วนด้านที่มีค่าสัมประสิทธิ์น้ำหนักองค์ประกอบต่ำที่สุดคือ ด้านความพึงพอใจในงาน ($\beta = 0.475$) ดังภาพ 1



ภาพ 1 โมเดลสมการโครงสร้างการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผล

การปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4

$$\chi^2 = 39.894, df = 28, p\text{-value} = 0.068, \chi^2/df = 1.425, CFI = 0.970, TLI = 0.952, RMSEA = 0.044, SRMR = 0.038$$

**มีนัยสำคัญทางสถิติที่ระดับ .01

3. ผลการศึกษาแนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิภาพการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4

3.1 แนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิภาพการปฏิบัติงานของครูในโรงเรียน ประกอบไปด้วย 5 ประเด็น ดังนี้ 1) ด้านการพัฒนาระบบพี่เลี้ยงในการสอนงาน ผู้บริหารสถานศึกษาส่งเสริมให้มีกิจกรรมการสอนงาน จัดระบบคู่มือพี่เลี้ยงที่มีประสบการณ์ในการถ่ายทอดความรู้ อย่างหลากหลาย 2) ด้านการสื่อสารเพื่อสร้างสัมพันธภาพ สถานศึกษามีช่องทางการสื่อสารหลากหลาย นำเทคโนโลยีสารสนเทศมาใช้ในการสื่อสาร 3) ด้านการดำรงรักษาบุคลากรในองค์กร ผู้บริหารสถานศึกษาสร้างแรงจูงใจและให้ผลตอบแทนในการปฏิบัติงานที่ดีแก่ครู 4) ด้านการส่งเสริมการทำงานเป็นทีม สร้างทีมงานที่หลากหลายที่ประกอบด้วยบุคลากรทุกช่วงวัย เปิดโอกาสให้ครูทุกช่วงวัยได้มีโอกาสเป็นผู้นำทีมงาน และ 5) ด้านการสนับสนุนการทำงานแบบยืดหยุ่น ผู้บริหารสถานศึกษาปรับเปลี่ยนวิธีการทำงานให้ยืดหยุ่นสอดคล้องกับลักษณะการทำงานของบุคลากรแต่ละช่วงวัยอย่างเหมาะสม

3.2 ผลการประเมินแนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิภาพการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 โดยภาพรวมมีความเหมาะสมอยู่ในระดับมากที่สุด ($M = 4.70$, $SD = 0.40$) มีความเป็นไปได้อยู่ในระดับมาก ($M = 4.48$, $SD = 0.55$) และมีความเป็นประโยชน์อยู่ในระดับมากที่สุด ($M = 4.82$, $SD = 0.26$)

อภิปรายผล

1. ผลการวิจัยระดับการจัดการความหลากหลายทางเจเนอเรชัน พบว่าโดยภาพรวมมีค่าเฉลี่ยอยู่ในระดับมาก ทั้งนี้อาจเป็นเพราะในปัจจุบันโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 กำลังเผชิญกับความหลากหลายทางเจเนอเรชันหรือความหลากหลายด้านช่วงอายุของบุคลากร ทำให้ความคิด ความคาดหวัง ความสามารถในการทำงานและมุมมองในด้านต่าง ๆ แตกต่างกัน ผู้บริหารจึงมีบทบาทในการจัดการความหลากหลายทางเจเนอเรชันของบุคลากรมากขึ้น หากผู้บริหารสามารถนำเอาความหลากหลายเหล่านั้นออกมาบริหารจัดการและใช้ให้เกิดประโยชน์แก่องค์กรได้ ย่อมนำไปสู่การเสริมสร้างประสิทธิภาพและประสิทธิผลการดำเนินงานขององค์กรให้ประสบผลสำเร็จ สอดคล้องกับงานวิจัยของ Thumviriyavong (2016) ที่สรุปว่า การที่มีความหลากหลายทำให้เกิดข้อได้เปรียบเสียเปรียบในองค์กรยุคปัจจุบัน โดยในทางที่ดีนั้นความหลากหลายของบุคลากรถือเป็นคุณค่า มีความสำคัญและเป็นประโยชน์ต่อองค์กร เพราะความหลากหลายนำมาซึ่งความสามารถที่แตกต่าง ความคิดเห็นที่แตกต่าง จนกลายเป็นความคิดสร้างสรรค์ในการทำงาน แต่ในทางตรงกันข้ามนั้นหากองค์กรไม่สามารถจัดการกับความหลากหลายที่เกิดขึ้นได้ ความหลากหลายนี้อาจกลายเป็นความแตกแยก ทำให้การทำงานมีแต่ข้อผิดพลาด สอดคล้องกับงานวิจัยของ Nakprachaya & Sirisooksilp (2016) ที่ได้ศึกษาการบริหารความหลากหลายของบุคลากรในโรงเรียน โดยภาพรวมทั้ง 7 ด้าน อยู่ในระดับมาก

ผลการวิจัยระดับประสิทธิผลการปฏิบัติงานของครูในโรงเรียน โดยภาพรวมมีค่าเฉลี่ยอยู่ในระดับมาก ทั้งนี้อาจเป็นเพราะประสิทธิผลการปฏิบัติงานของครูเป็นผลสำเร็จที่ครูสามารถปฏิบัติงานให้บรรลุตาม

วัตถุประสงค์และเป้าหมายที่กำหนดไว้อย่างครบถ้วนสมบูรณ์ ทั้งนี้ย่อมแสดงว่าการบริหารงานในสถานศึกษานั้นบังเกิดประสิทธิผลด้วยเช่นกัน เมื่อผู้บริหารสถานศึกษามีการบริหารจัดการที่ดี มีทิศทางในการบริหารงานที่ชัดเจนจึงส่งผลให้ครูมีแนวทางในการปฏิบัติงานได้ดีเช่นเดียวกัน สอดคล้องกับ Chapanya (2016) ที่กล่าวว่า การเกิดประสิทธิผลของการทำงานพิจารณาจากผลสำเร็จที่เกิดขึ้นจากการจัดการเรียนรู้ที่โรงเรียนดำเนินการ ซึ่งสอดคล้องกับวิสัยทัศน์ของโรงเรียน โดยผลสำเร็จของงานมีความเกี่ยวข้องกับทั้งคุณภาพของนักเรียนและคุณภาพการทำงานของครูซึ่งเป็นผลจากการที่ครูได้รับการพัฒนาด้านการจัดการเรียนรู้อย่างต่อเนื่อง สอดคล้องกับงานวิจัยของ Labbun (2020) ที่ได้ศึกษาคุณลักษณะส่วนบุคคลของผู้บริหารที่ส่งผลต่อประสิทธิผล การปฏิบัติงานของครูในโรงเรียน ที่พบว่าประสิทธิผลการปฏิบัติงานของครูในภาพรวมอยู่ในระดับมาก

2. ผลการตรวจสอบความสอดคล้องกลมกลืนของโมเดลสมการโครงสร้างการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษา ประถมศึกษาขอนแก่น เขต 4 พบว่า โมเดลการวัดการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียนมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ นั่นคือ องค์ประกอบที่สังเคราะห์ขึ้นมีความสอดคล้องกับบริบทของโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษา ประถมศึกษาขอนแก่น เขต 4 โดยเมื่อพิจารณาค่าน้ำหนักองค์ประกอบ (factor loading) มีข้อค้นพบที่ควรนำมาอภิปรายผล ดังนี้

หากพิจารณาค่าสัมประสิทธิ์น้ำหนักองค์ประกอบในรูปคะแนนมาตรฐานในโมเดลการวัดการจัดการความหลากหลายทางเจเนอเรชัน มีนัยสำคัญทางสถิติที่ระดับ .01 ทุกตัว ซึ่งองค์ประกอบที่สามารถอธิบายการจัดการความหลากหลายทางเจเนอเรชันได้ดีที่สุด คือ ด้านการพัฒนาระบบพี่เลี้ยงในการสอนงาน ทั้งนี้ อาจเป็นเพราะโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 เป็นหน่วยงานหนึ่งที่บุคลากรมีความหลากหลาย มีบุคลากรที่ใกล้จะเกษียณอายุราชการและมีการรับสมัครหรือบรรจุบุคลากรใหม่เพื่อมาเป็นแรงงานหลักในการปฏิบัติงาน การพัฒนาระบบพี่เลี้ยงเพื่อให้เกิดการสอนงาน ปกป้องดูแลกัน ในองค์กรจึงมีความสำคัญอย่างยิ่งและสามารถนำมาวัดการจัดการความหลากหลายทางเจเนอเรชันของผู้บริหารสถานศึกษาได้ดีที่สุด สอดคล้องกับ Wiriyaipat (2016) ที่กล่าวถึงการสอนงานเพื่อวางแผนความต่อเนื่องว่า เนื่องจากพนักงาน Generation Baby Boomers กำลังทยอยเกษียณอายุออกจากองค์กร ซึ่งพนักงานกลุ่มนี้จะสะสมความรู้ ความสามารถและประสบการณ์ที่มีคุณค่าตลอดชีวิตการทำงานทำให้องค์กรหลายแห่งประสบปัญหาไม่สามารถสร้างพนักงานรุ่นใหม่ขึ้นมาทดแทนได้ทัน ผู้บริหารจึงต้องแต่งตั้งพนักงาน Generation Baby Boomers เป็นที่ปรึกษาหรือผู้สอนงานตลอดจนทักษะที่จำเป็นให้แก่พนักงานเจเนอเรชันถัดไป

ค่าสัมประสิทธิ์ น้ำหนักองค์ประกอบในรูปคะแนนมาตรฐานในโมเดลการวัดประสิทธิผล การปฏิบัติงานของครูในโรงเรียน มีนัยสำคัญทางสถิติที่ระดับ .01 ทุกตัว ซึ่งองค์ประกอบที่สามารถอธิบายประสิทธิผลการปฏิบัติงานของครูในได้ดีที่สุด คือ ด้านคุณภาพผู้เรียน ทั้งนี้ อาจเป็นเพราะหน้าที่ในการปฏิบัติงานหลักของครูในโรงเรียน คือ การจัดการเรียนการสอน เนื่องจากครูเป็นกลไกสำคัญที่จะทำให้การเรียนการสอนมีคุณภาพหรือด้อยคุณภาพ ดังนั้นคุณภาพการปฏิบัติงานของครูจึงมีอิทธิพลต่อคุณภาพของผู้เรียน ซึ่งประกอบไปด้วยผลสัมฤทธิ์ทางการเรียนและคุณลักษณะที่พึงประสงค์ของนักเรียน สอดคล้องกับผลการวิจัย

ของ Luksanapawan & Somprach (2019) ที่ได้ศึกษาโมเดลภาวะผู้นำพลังร่วมที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน พบว่า องค์ประกอบประสิทธิผลการปฏิบัติของครูที่มีค่าน้ำหนักองค์ประกอบมากที่สุดคือ ด้านคุณภาพผู้เรียน

ผลการวิเคราะห์ค่าอิทธิพลของการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 มีขนาดอิทธิพลทางตรงเชิงบวกเท่ากับ 0.796 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 หมายความว่า ถ้าการจัดการความหลากหลายทางเจเนอเรชัน (GEN) เพิ่มขึ้น 1 หน่วย จะส่งผลให้ประสิทธิผลการปฏิบัติงานของครู (EFF) เพิ่มขึ้น 0.796 หน่วย โดยตัวแปรการจัดการความหลากหลายทางเจเนอเรชันในโมเดลสามารถอธิบายความแปรปรวนในการเพิ่มขึ้นของตัวแปรประสิทธิผลการปฏิบัติงานของครูได้ร้อยละ 63.3 จากข้อค้นพบดังกล่าวจะเห็นได้ว่าการจัดการความหลากหลายทางเจเนอเรชันซึ่งมีองค์ประกอบอยู่ 5 องค์ประกอบ คือ การส่งเสริมการทำงานเป็นทีม การสื่อสารเพื่อสร้างสัมพันธภาพ การพัฒนาระบบพี่เลี้ยงในการสอนงาน การสนับสนุนการทำงานแบบยืดหยุ่น และการธำรงรักษาบุคลากรในองค์กร เป็นปัจจัยที่สำคัญอย่างยิ่งที่จะส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน เนื่องจากการปฏิบัติงานในองค์กรที่บุคลากรมีความหลากหลาย มีความแตกต่างด้านมุมมอง ความคิดและทัศนคติในการทำงาน หากคาดหวังให้การปฏิบัติงานเกิดประสิทธิภาพและเกิดประสิทธิผลแล้ว ผู้บริหารจำเป็นต้องอย่างยิ่งที่จะต้องเรียนรู้ลักษณะร่วมและแตกต่างของบุคลากรแต่ละช่วงวัย และใช้เทคนิคที่หลากหลายในการบริหารบุคลากรให้สามารถปฏิบัติงานร่วมกันได้อย่างราบรื่นและเต็มศักยภาพ สอดคล้องกับ Dejthongjun (2021) ที่กล่าวว่าองค์การใดที่มีความหลากหลายทางด้านทรัพยากรมนุษย์และทุนมนุษย์ หากองค์กรนั้นสามารถนำเอาความหลากหลายเหล่านั้นออกมาบริหารจัดการและใช้ให้เกิดประโยชน์แก่องค์กรได้ ย่อมนำไปสู่การเสริมสร้างประสิทธิภาพประสิทธิผลการดำเนินงานขององค์กรให้ประสบผลสำเร็จ กล่าวได้ว่า ความสามารถในการจัดการความหลากหลายจะสามารถนำคุณค่าที่หลากหลายของบุคลากรในองค์กรมาสร้างจุดแข็งและความได้เปรียบทางการแข่งขันเพื่อให้บรรลุเป้าหมายขององค์กร ซึ่งการจัดการความหลากหลายนี้มีองค์ประกอบ 3 ประเด็น ได้แก่ 1) การบูรณาการความรู้ความสามารถที่โดดเด่น 2) ความยืดหยุ่นในการประเมินผลการปฏิบัติงาน และ 3) ความสามารถในการปรับตัวตามสภาพแวดล้อม นอกจากนี้ยังสอดคล้องกับงานวิจัยของ Tangjongrach & Thawinkarn (2021) ที่ได้ศึกษารูปแบบการบริหารจัดการความหลากหลายทางเจเนอเรชันของบุคลากรในโรงเรียน พบว่า โมเดลการวัดตัวบ่งชี้การบริหารจัดการความหลากหลายทางเจเนอเรชันของบุคลากรในโรงเรียน มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ มีค่าน้ำหนักองค์ประกอบ (factor loading) ตั้งแต่ 0.783-0.938 ประกอบด้วย 1) การสร้างทีมงานที่หลากหลาย 2) การสอนงานปรึกษาและดูแล 3) การพัฒนาบุคลากรอย่างต่อเนื่อง 4) การธำรงรักษาบุคลากรคุณภาพ และ 5) การสื่อสารเพื่อสร้างสัมพันธภาพ

3. ผลการศึกษาแนวทางการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ประกอบด้วย 5 ประเด็น ได้แก่ การพัฒนาระบบพี่เลี้ยงในการสอนงาน การสื่อสารเพื่อสร้างสัมพันธภาพ การธำรงรักษาบุคลากรในองค์กร การส่งเสริมการทำงานเป็นทีม และการสนับสนุนการทำงานแบบยืดหยุ่น สอดคล้องกับ

งานวิจัยของ Thongkhum (2021) ที่ได้ศึกษาแนวทางการบริหารจัดการความหลากหลายทางเจเนอเรชันของครูในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เลย หนองบัวลำภู พบว่าประกอบด้วย 5 ด้าน ได้แก่ 1) ด้านการพัฒนาวิธีการทำงานที่ยืดหยุ่นตามเจเนอเรชัน 2) ด้านการสร้างวัฒนธรรมการเรียนรู้ 3) ด้านการพัฒนาทักษะการจัดการ Soft Skills 4) ด้านการสร้างบรรยากาศองค์กร และ 5) ด้านการสร้างระบบที่เลี้ยง นอกจากนี้ผู้บริหารควรมีการจัดคนให้เหมาะสมกับลักษณะงาน และบริบทของโรงเรียน โดยให้ครูแต่ละคนได้ใช้ศักยภาพที่โดดเด่นของตนมาทำงานร่วมกันภายในโรงเรียนเพื่อให้เกิดประสิทธิภาพและประสิทธิผลบนความหลากหลายทางเจเนอเรชัน

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 จากผลการวิจัยพบว่า องค์กรประกอบที่สามารถอธิบายการจัดการความหลากหลายทางเจเนอเรชันได้ดีที่สุด คือ ด้านการพัฒนาระบบที่เลี้ยงในการสอนงาน ดังนั้นสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ควรมีการจัดกิจกรรมการสอนงานเพื่อวางแผนความต่อเนื่อง เตรียมบุคลากรเพื่อทดแทนบุคลากร Generation Baby Boomersที่กำลังทยอยเกษียณอายุราชการ และจัดกิจกรรมพัฒนาบุคลากรสู่การเป็นบุคลากรต้นแบบในแต่ละด้าน

1.2 จากผลการวิจัยพบว่า องค์กรประกอบที่สามารถอธิบายประสิทธิผลการปฏิบัติงานของครูได้ดีที่สุด คือ ด้านคุณภาพผู้เรียน ดังนั้น สำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ควรมีนโยบายเพื่อยกระดับคุณภาพผู้เรียนให้ได้รับการพัฒนาในทุกมิติ เป็นคนดี คนเก่ง มีคุณภาพ ส่งเสริมให้โรงเรียนมุ่งเน้นการพัฒนาศักยภาพและคุณภาพผู้เรียนให้มีสมรรถนะตามหลักสูตรและคุณลักษณะในศตวรรษที่ 21

1.3 จากผลการวิจัยพบว่า ค่าอิทธิพลการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลต่อประสิทธิผลการปฏิบัติงานของครูในโรงเรียน มีอิทธิพลทางบวกที่ส่งผลต่อกันอย่างมีนัยสำคัญทางสถิติ ดังนั้นสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 4 ควรส่งเสริมให้ผู้บริหารสถานศึกษามีการจัดการความหลากหลายทางเจเนอเรชันในโรงเรียนทุกองค์ประกอบ เพื่อส่งเสริมให้เกิดการพัฒนาประสิทธิผลการปฏิบัติงานของครูให้ครบทุกมิติต่อไป

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 จากการศึกษาเชิงคุณภาพ ข้อค้นพบที่ได้จากการสัมภาษณ์ผู้ทรงคุณวุฒิ ควรศึกษาตัวแปรคันกลาง (mediator variable) ที่ส่งผ่านจากการจัดการความหลากหลายทางเจเนอเรชันที่ส่งผลไปยังประสิทธิผลการปฏิบัติงานของครูในโรงเรียน เช่น ภาวะผู้นำของผู้บริหาร การสร้างบรรยากาศองค์กร การจัดสรรทรัพยากรสนับสนุนการปฏิบัติงาน เป็นต้น เพื่อให้เห็นมุมมองของการส่งผลที่ชัดเจนมากยิ่งขึ้น

2.2 ควรมีการนำผลการวิจัยไปศึกษาหรือพัฒนาต่อในรูปแบบการวิจัยและพัฒนา (the research and development) เพื่อศึกษาและพัฒนาวิธีการอย่างลึกซึ้ง ในการพัฒนาการจัดการความหลากหลายทางเจเนอเรชัน เพื่อใช้ในการยกระดับประสิทธิผลการปฏิบัติงานของครู

2.3 ควรศึกษาแนวทางการจัดการความหลากหลายทางเจเนอเรชัน โดยทำการศึกษาทฤษฎีฐานราก (grounded theory) ลงพื้นที่เก็บข้อมูลจากสถานศึกษาที่มีการบริหารจัดการความหลากหลายทางเจเนอเรชันของบุคลากรที่โดดเด่น และใช้การเก็บข้อมูลจากการสนทนากลุ่ม (focused group) เพื่อให้เห็นมุมมองและแนวคิดของบุคลากรจากเจเนอเรชันต่าง ๆ

References

- Ardakani, M. S., Abzari, M., Shaemi, A., & Fathi, S. (2015). Diversity management and human resources productivity: Mediating effects of perceived organizational attractiveness Organizational justice and social identity in Isfahan's steel industry. *Iranian Journal of Management Studies*, 9(2), 407-432.
- Barnard, C. L. (1968). *The functions of executive*. Harvard University Press.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods Research* (2nd Ed.). Sage.
- Cronbach, L. J. (1970). *The evolution of research*. Harper Collins.
- Gilbert, J. A., & Ivancevich, J. M. (2000). Valuing Diversity: A Tale of Two Organizations. *The Academy of Management Executive*, 14(1) 93-105.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E., (2010). *Multivariate Data Analysis: A Global Perspective*. Pearson Education.
- Hoy, W. K., Miskel, C. G., & John Tarter, C. (2013). *Educational Administration: Theory Research and Practice*. McGraw- Hill.
- Rosado, Caleb. (2006). *Workforce Diversity, Vol.3: Concepts and Cases*. ICAFAI University.
- Schermerhorn, J. R., Jr.; Hunt, J. G., Osborn, R. N., & Uhl-Bien, M. (2000). *Organizational Behavior*. Wiley.

Translate Thai References

- Chapanya, Y. (2016). *Leadership and Ethics of the Leaders*. Roi Et Rajabhat University. (in Thai)
- Chueabangkaew, P. (2014). Diversity Management: Human Resource Management Strategy in the Creative Economy. *Executive Journal*. 31(1), 153-159. (in Thai)
- Dejthongjun, S. (2021). Human Resource Diversity Management in Organization for Equality and Fairness: a Case Study on Human Capital Development of Employees with Disabilities in Government Sector. *Journal of Social Science and Buddhist Anthropology*, 4(3), 269-284. (in Thai)

- Labun, W. (2020). *Personal Characteristics of School Administrators Affecting Effectiveness of Teacher performance in Schools under Bueng Kan Primary Educational Service Area Office* [Master's thesis]. Sakon Nakhon Rajabhat University. https://gsmis.snru.ac.th/e-thesis/thesis_detail?r=61421229110 (in Thai)
- Luksanapawan, B., & Somprach, K. (2019). Collective Leadership Affecting Effectiveness of Teacher's Performance in Schools under the Secondary Educational Service Office Area 25. *Dhammathas Academic Journal*, 19(4), 1-13. (in Thai)
- Nakprachaya, J., & Sirisooksilp, S. (2016). Workforce Diversity Management in Schools under The Office of Secondary Educational Service Area 25. *College of Asian Scholar Journal*, 4(2), 39-47. (in Thai)
- Ninaroon, P. (2017). *The Causal Relationship and Effect of Migrants Worker Management Capability: an Empirical Evidence of Small and Medium Enterprises in Samutsakorn Provinces* [Doctoral dissertation], Silpakorn University. <http://itthesis-ir.su.ac.th/dspace/bitstream/123456789/1648/1/57604932.pdf> (in Thai)
- Office of the Teachers council of Thailand. (2013). *Guideline for professional practice of Education*. Kurusapa Printing Ladphrao. (in Thai)
- Soisud, V. (2021). Human Resource Management is the Key to Driving Today's Organization. *Journal of Modern Learning Development*, 4(5), 340-350. (in Thai)
- Tangjongrach, P., & Thawinkarn, D. (2021). A Multi-Generation Management Model for School Personnel under the Office of Udonthani Primary Educational Service Area 2. *Phranakhon Rajabhat Research Journal*, 14(1), 140-159. (in Thai)
- The Office of Khon Kaen Primary Educational Service Area 4. (2021). *Annual Action Plan 2021*. The Office of Khon Kaen Primary Educational Service Area 4. (in Thai)
- Thongkhum, A. (2021). *Guidelines for Generational Diversity Management in School Under the Office of Secondary Educational Service Area 19* [Unpublished Master's thesis]. Khon Kaen University. (in Thai)
- Thumviriyavong, P. (2016). Workforce Diversity Management. *Academic Journal of Bangkokthonburi University*, 5 (2), 1-12. (in Thai)
- Wiratchai, N. (1999). *LISREL Model: Analytical Statistics for Research* (3rd Ed.). Chulalongkorn University Press. (in Thai)
- Wiriyapipat, N. (2016). The Role of HRM: Generation Gap Management in Organization. *RMUTP Research Journal Humanities and Social Science*, 1(2), 73-88. (in Thai)

การวิเคราะห์องค์ประกอบเชิงยืนยันการสร้างสรรค์และนวัตกรรม สำหรับนักศึกษามหาวิทยาลัยราชภัฏ

The Confirmatory Factor Analysis of Creativity and Innovation for Rajabhat University Students

ธราเทพ เตมีรักษ์^{1*} และ อารยา ปิยะกุล²

Tharathep Tameru^{1*} and Araya Piyakun²

(Received: March 14, 2022; Revised: April 22, 2022; Accepted: April 25, 2022)

บทคัดย่อ

บทความวิจัยนี้มีความมุ่งหมายเพื่อศึกษาองค์ประกอบการสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏใช้รูปแบบการวิจัยเชิงปริมาณ ตัวอย่างในการศึกษา คือ นักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏ ปีการศึกษา 2564 จำนวน 1,200 คน เครื่องมือวิจัยที่ใช้ในการเก็บข้อมูล คือ แบบวัดการสร้างสรรค์และนวัตกรรม มีค่าความเชื่อมั่นที่ .92 สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ การวิเคราะห์องค์ประกอบเชิงยืนยัน การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน ส่วนเบี่ยงเบนมาตรฐาน ค่าเฉลี่ย และร้อยละ

ผลการวิจัยพบว่า องค์ประกอบการสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ ประกอบด้วย 4 องค์ประกอบ เรียงตามลำดับความสำคัญจากมากไปหาน้อย ได้แก่ ปัญญาสร้างสรรค์ ปฏิสัมพันธ์ทางสังคม จิตใจเชิงสร้างสรรค์ และการสร้างนวัตกรรมตามลำดับ ความสอดคล้องระหว่างองค์ประกอบกับข้อมูลเชิงประจักษ์ในดัชนีกลุ่มสัมบูรณ์พบว่า Chi-square 194.68, df=165, p-value=0.06, ค่า GFI=0.99, ค่า AGFI=0.98, ค่า RMSEA=0.01 เมื่อพิจารณาค่าความสอดคล้องพบว่า ค่า Chi-square ไม่นับนัยสำคัญทางสถิติและดัชนีอื่น ๆ ซึ่งให้เห็นว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ผลของการศึกษาค้นคว้าครั้งนี้เป็นข้อสนเทศที่สำคัญในการพัฒนาการสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ

คำสำคัญ: การวิเคราะห์องค์ประกอบเชิงยืนยัน ปัญญาสร้างสรรค์ จิตใจเชิงสร้างสรรค์ ปฏิสัมพันธ์ทางสังคม
การสร้างสรรค์และนวัตกรรม

¹ นิสิตดุษฎีบัณฑิต สาขาจิตวิทยาการศึกษาและการแนะแนว คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

² อาจารย์ สาขาจิตวิทยาการศึกษาและการแนะแนว คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

¹ Doctoral degree student, Educational Psychology and Guidance, Faculty of Education, Mahasarakham University

² Lecturer, Educational Psychology and Guidance, Faculty of Education, Mahasarakham University

* Corresponding Author E-mail: T0824554953@gmail.com

Abstract

This research article aimed to study the components of creativity and innovation of Rajabhat University students. The study applied a quantitative research method. The sample used in the study consisted of 1,200 undergraduate students studying in Rajabhat Universities in the academic year 2021. The research instrument was a questionnaire inquiring about creativity and innovation of Rajabhat University students, with the reliability of .92. The data were analyzed using confirmatory factor analysis, Pearson Product Moment Correlation, standard deviation (S.D.), mean, and percentage.

The results showed that the creativity and innovation of Rajabhat University students consisted of 4 components, listed in descending order according to their importance, the components are as follows: creative cognition, social interaction, creative mind, and innovation creation respectively. The components had a good fit with the empirical data (Chi-square= 194.68, df=165, p-value=0.06, GFI=0.99, AGFI=0.98, RMSEA=0.01). When the consistency criterion was considered, it was found that the Chi-square value had no statistical significance, and other indexes indicated that the model was consistent with the empirical data. The findings of this study provide an insight into creativity and innovation of Rajabhat University students.

Keywords: confirmatory factor analysis, creative cognition, creative mind, social interaction, creativity and innovation

บทนำ

โลกในปัจจุบันมีความเจริญก้าวหน้าอย่างรวดเร็ว จากการที่มนุษย์ประดิษฐ์คิดค้นสิ่งต่าง ๆ เพื่ออำนวยความสะดวกและแก้ไขปัญหาที่เกิดขึ้นในทุกด้านของชีวิต ไม่มีใครที่สามารถปฏิเสธเทคโนโลยีที่กำลังเข้ามามีบทบาทสำคัญและสร้างผลกระทบต่อทุกแง่มุมของชีวิต มนุษย์จึงต้องเรียนรู้ที่จะปรับตัวเพื่อความอยู่รอด โดยการปรับเปลี่ยนมุมมอง รวมถึงการฝึกฝนทักษะต่าง ๆ เพื่อให้ตนเป็นที่ยอมรับและเป็นสมาชิกในสังคมอย่างปกติสุข การสร้างสรรค์และนวัตกรรม (creativity and innovation) จึงกลายเป็นความสามารถที่มีความสำคัญอย่างยิ่งในปัจจุบันและอนาคต และหากมองในทางจิตวิทยาจะมีความสอดคล้องกับทฤษฎีความคิดสร้างสรรค์เชิงมนุษยนิยม ที่เชื่อว่ามนุษย์มีความคิดสร้างสรรค์ติดตัวมาตั้งแต่กำเนิด โดยผู้ที่สามารถดึงศักยภาพด้านความคิดสร้างสรรค์ออกมาใช้ได้ คือผู้ที่รู้จักตนเอง พอใจในตนเอง และใช้ศักยภาพตนเองอย่างเต็มขั้น (Muangsuriya, 2016) ขณะที่ปัจจุบันการสื่อสารและเทคโนโลยีเจริญขึ้นอย่างรวดเร็วทักษะการคิดสร้างสรรค์เพียงอย่างเดียวจึงไม่เพียงพอสำหรับการพัฒนามนุษย์ในปัจจุบัน ซึ่งความจริงแล้วความสามารถในการสร้างสรรค์และนวัตกรรมเป็นความสามารถที่มีทั้งองค์ประกอบด้านปัญญา (cognition) และด้านจิตใจ (affect) ดังนั้นการสร้างสรรค์และนวัตกรรมจึงเป็นความสามารถใหม่ที่มีความสำคัญเป็นอันดับที่ 1 ขององค์กรทั่วโลก

(Batey, 2011) และเป็นความท้าทายทางจิตวิทยาในการศึกษาเพื่อให้เกิดความเข้าใจ สามารถอธิบาย ทำนาย และควบคุมส่งเสริมให้เกิดความสามารถในการสร้างสรรค์และนวัตกรรมในบุคคล เพื่อนำไปใช้ประโยชน์ได้จริง อันจะส่งผลให้ประเทศเกิดการพัฒนายอย่างสูงสุด (Piyakun, 2016)

ในการพัฒนาประเทศนั้น ระบบการศึกษาถือว่ามีสำคัญอย่างมากในการพัฒนาทรัพยากรมนุษย์ สถาบันอุดมศึกษาซึ่งเป็นภาคผลิตบัณฑิตสู่ตลาดแรงงาน จะต้องจัดเตรียมหลักสูตรการเรียนการสอนให้รองรับกับการเข้าทำงานและอาชีพที่หลากหลาย ปรับปรุงรูปแบบการเรียนรู้ สู่การอุดมศึกษาไทยในยุค 4.0 ที่มีหัวใจสำคัญคือ “การสร้างสรรค์นวัตกรรมและพัฒนาทรัพยากรมนุษย์” (Kachintorn, 2018) มหาวิทยาลัยต้องมีมาตรฐานทักษะการใช้เทคโนโลยีสารสนเทศและการสื่อสารที่มีความสำคัญต่อผู้เรียนในศตวรรษที่ 21 ในระดับมหาวิทยาลัย ซึ่งประกอบด้วย การสร้างสรรค์และนวัตกรรม การสื่อสารและการทำงานร่วมกัน แต่จากผลการวิจัยของ Power (2016) พบว่า การศึกษาในประเทศไทยไม่สามารถพัฒนาผู้เรียนให้ไปถึงเป้าหมายในการพัฒนาประเทศได้ โดยเฉพาะการพัฒนาทรัพยากรมนุษย์ ในด้านการเป็นผู้มีความคิดสร้างสรรค์และสามารถสร้างนวัตกรรมเพื่อการแข่งขันระหว่างประเทศ โดยการศึกษาถึงปัจจัยสาเหตุของปัญหาดังกล่าว พบว่า วัฒนธรรมไทย มีผลกระทบอย่างมีนัยสำคัญต่อความคิดสร้างสรรค์ของผู้เรียนในประเทศไทย การมีบรรทัดฐานทางวัฒนธรรมจำนวนมากถูกมองว่าเป็นอุปสรรคต่อความสามารถของผู้เรียนในการสร้างสรรค์ เช่น วัฒนธรรมทางพุทธศาสนาที่แข็งแกร่ง ด้วยความเคารพต่อผู้สูงอายุและผู้มีอำนาจ นอกจากนี้ผลการวิจัยยังพบว่า กว่าร้อยละ 75 ของตัวอย่างในการทดลองเห็นพ้องกันว่า มหาวิทยาลัยควรทำอะไรให้มากขึ้นเพื่อสนับสนุนความคิดสร้างสรรค์ในการศึกษา ร้อยละ 67 ของกลุ่มตัวอย่างต้องการโอกาสในการสร้างสรรค์และพัฒนาแรงจูงใจในการสร้างสรรค์โดยผู้สอน ปัจจัยอีกประการหนึ่งจากการเข้าร่วมการสนทนากลุ่มสะท้อนพบว่า ภาระงานที่หนักและแรงกดดันเพื่อให้ได้เกรดที่ดีเป็นปัจจัย ที่ทำให้นักศึกษาไม่กล้าเสี่ยงที่จะทำอะไรสร้างสรรค์นอกกรอบ ดังนั้นรัฐบาลและมหาวิทยาลัย จำเป็นต้องปรับทิศทางและนโยบายทางการศึกษาให้ตอบสนองต่อการพัฒนาความคิดสร้างสรรค์ของนักศึกษา (Power, 2016)

ในการวิจัยครั้งนี้ผู้วิจัยมุ่งศึกษากลุ่มมหาวิทยาลัยราชภัฏ เนื่องจากมหาวิทยาลัยราชภัฏต้องสร้างพันธกิจให้สัมพันธ์กับความเป็นอุดมศึกษาเพื่อการพัฒนาท้องถิ่นให้เข้มข้นทั้งด้าน เศรษฐกิจ เทคโนโลยี การเมือง วัฒนธรรม สังคมและท้องถิ่นอย่างแท้จริง บนฐานการบริหารจัดการที่เน้นธรรมาภิบาล ซึ่งมียุทธศาสตร์ที่เกี่ยวข้องกับการพัฒนาการสร้างสรรค์และนวัตกรรมคือ ยุทธศาสตร์ที่ 1 การเสริมสร้างและพัฒนาศักยภาพทุนมนุษย์ของประเทศไทยยังมีปัญหาในด้านคุณภาพคนในแต่ละช่วงวัย โดยผลลัพธ์ทางการศึกษาของเด็กวัยเรียนค่อนข้างต่ำ การพัฒนาความรู้และทักษะของแรงงานไม่ตรงกับตลาดงาน ในขณะที่คนไทยจำนวนไม่น้อยที่ยังไม่สามารถคัดกรองและเลือกรับวัฒนธรรมได้อย่างเหมาะสม ซึ่งส่งผลต่อวิกฤตค่านิยม ทศนคติ และพฤติกรรมในการดำเนินชีวิต การพัฒนาเหล่านี้มหาวิทยาลัยราชภัฏจึงต้องให้ความสำคัญกับการวางรากฐานการพัฒนาคนให้มีความสมบูรณ์ เพื่อให้คนไทยมีทศนคติและพฤติกรรมตามบรรทัดฐานที่ดีของสังคม ได้รับการศึกษาที่มีคุณภาพสูงตามมาตรฐานสากล และสามารถเรียนรู้ด้วยตนเองอย่างต่อเนื่อง มีสุขภาวะที่ดีขึ้น คนทุกช่วงวัยมีทักษะ มีความรู้ และมีความสามารถเพิ่มขึ้น รวมทั้งมีความเข้มแข็งและมีส่วนร่วมในการพัฒนาประเทศ ซึ่งการขับเคลื่อนของมหาวิทยาลัยราชภัฏสู่การปฏิบัติ เพื่อให้เป็นสถาบันอุดมศึกษาในการพัฒนา

ท้องถิ่นอย่างแท้จริง ควรมีหลักการและแนวทางการพัฒนางาน คือ การใช้องค์ความรู้ เทคโนโลยี นวัตกรรม และความคิดสร้างสรรค์ให้เป็นเครื่องมือหลักในการขับเคลื่อนพัฒนาในทุกภาคส่วนในระดับพื้นที่ ท้องถิ่นและชุมชน ผ่านการจัดการเรียนการสอนในแต่ละหลักสูตรทั้งระดับปริญญาตรีและระดับบัณฑิตศึกษา (Chinnawong, 2001)

แม้ว่าการสร้างสรรค์และนวัตกรรม มีที่มาโดยปรากฏเป็นความสามารถหนึ่งของทักษะแห่งศตวรรษที่ 21 ที่พัฒนาโดยองค์กรภาคีเพื่อทักษะแห่งศตวรรษที่ 21 ที่ระบุว่า การสร้างสรรค์และนวัตกรรม เป็นหนึ่งในทักษะด้านการเรียนรู้และนวัตกรรม ซึ่งมีองค์ประกอบของการสร้างสรรค์และนวัตกรรม 3 องค์ประกอบ คือ สร้างสรรค์ (think creatively) ทำงานกับผู้อื่นอย่างสร้างสรรค์ (work creatively with others) และการสร้างนวัตกรรม (implement innovations) (Partnership for 21st Century, 2019) แต่จากการศึกษาผู้วิจัยยังพบว่า ยังมีผู้เชี่ยวชาญอีกกลุ่มหนึ่งที่เสนอว่าแรงจูงใจภายในที่บุคคลต้องการจะสร้างสรรค์สิ่งต่าง ๆ เป็นองค์ประกอบที่สำคัญของการสร้างสรรค์และนวัตกรรม (Amabile, 1988; Morrison, 2010) ดังนั้น จะเห็นได้ว่าองค์ประกอบของการสร้างสรรค์และนวัตกรรมยังมีช่องว่างขององค์ความรู้และส่งผลต่อการพัฒนาการสร้างสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ ผู้วิจัยจึงมีความสนใจที่จะทำการศึกษาวิจัยเรื่อง "การศึกษาองค์ประกอบการสร้างสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ" ในครั้งนี้

ความมุ่งหมายของการวิจัย

เพื่อศึกษาองค์ประกอบการสร้างสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงปริมาณ ผู้วิจัยมีการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องกับตัวแปรที่ศึกษา ประกอบด้วย ความหมายและความสำคัญของการสร้างสรรค์และนวัตกรรม องค์ประกอบและตัวบ่งชี้การสร้างสร้างสรรค์และนวัตกรรม การวัดประเมินผลการสร้างสรรค์และนวัตกรรม และการวิเคราะห์องค์ประกอบเชิงยืนยันการสร้างสร้างสรรค์และนวัตกรรม ซึ่งมีขั้นตอนและรายละเอียดในการดำเนินการวิจัยดังนี้

1. ประชากรและตัวอย่าง

กลุ่มผู้ให้ข้อมูลในการวิจัยเพื่อศึกษานิยามองค์ประกอบ คือ ผู้เชี่ยวชาญทางจิตวิทยาการศึกษาหรือผู้เชี่ยวชาญทางการสร้างสรรค์และนวัตกรรม จำนวน 5 คน สัมภาษณ์เพื่อสร้างนิยาม ความหมายของการสร้างสรรค์และนวัตกรรมสำหรับนักศึกษามหาวิทยาลัยราชภัฏ และผู้เชี่ยวชาญทางจิตวิทยาการศึกษาและการสร้างสรรค์และนวัตกรรม 5 คน เพื่อตรวจสอบความตรงเชิงเนื้อหา

ประชากรในการวิจัยเพื่อวิเคราะห์องค์ประกอบเชิงยืนยัน คือ นักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏ ชั้นปีที่ 1 - 4 ปีการศึกษา 2564 จำนวน 38 แห่ง ขนาดประชากร 50,104 คน (Ministry of Higher Education, Science, Research and Innovation, 2021)

ตัวอย่างในการวิจัยเพื่อวิเคราะห์องค์ประกอบเชิงยืนยัน คือ นักศึกษาระดับปริญญาตรี ชั้นปีที่ 1 - 4 มหาวิทยาลัยราชภัฏ ปีการศึกษา 2564 กำหนดขนาดกลุ่มตัวอย่าง 20 เท่าของพารามิเตอร์ (Lindeman et

al., 1980) การวิจัยครั้งนี้มีทั้งสิ้น 53 พารามิเตอร์ ซึ่งควรได้ตัวอย่างจำนวน 1,060 คน เป็นเกณฑ์ขั้นต่ำ แต่เพื่อให้การวิจัยครั้งนี้เกิดความคลาดเคลื่อนน้อยที่สุด ผู้วิจัยจึงเพิ่มขนาดตัวอย่างเป็น 1,200 คน สุ่มตัวอย่างแบบหลายขั้นตอน (multi-stage sampling) โดยมีขั้นตอนการสุ่มตัวอย่าง คือ การกำหนดแบ่งภาคต่าง ๆ ได้ 4 ภาค คือ ภาคเหนือ ภาคกลาง ภาคตะวันออกเฉียงเหนือ และภาคใต้ จากนั้นสุ่มมหาวิทยาลัยตัวอย่างในแต่ละภาค ด้วยวิธีการสุ่มอย่างง่าย (simple random sampling) โดยการจับสลาก โดยใช้เกณฑ์ร้อยละ 15 (Tayraukham, 2012) จับสลากได้ 6 มหาวิทยาลัยราชภัฏ ดังนี้ ภาคเหนือ ได้แก่ มหาวิทยาลัยราชภัฏเชียงใหม่ มหาวิทยาลัยราชภัฏลำปาง ภาคอีสาน ได้แก่ มหาวิทยาลัยราชภัฏสกลนคร มหาวิทยาลัยราชภัฏเลย ภาคกลาง ได้แก่ มหาวิทยาลัยราชภัฏนครปฐม มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา ภาคใต้ ได้แก่ มหาวิทยาลัยราชภัฏสุราษฎร์ธานี มหาวิทยาลัยราชภัฏสงขลา จากนั้นทำการสุ่มคณะและสาขาวิชาจากมหาวิทยาลัยที่เป็นตัวอย่าง ในชั้นที่ 1 ด้วยวิธีการสุ่มอย่างง่าย ผู้วิจัยทำการเก็บข้อมูลกับนักศึกษาชั้นปีที่ 1 - 4 จากคณะและสาขาวิชาที่เป็นตัวอย่างเพื่อให้ได้นักศึกษาตัวอย่างครบตามจำนวน 1,200 คน

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบถามออนไลน์ ประกอบด้วย

2.1 แบบสัมภาษณ์ผู้เชี่ยวชาญ ในการกำหนดนิยามและองค์ประกอบของตัวบ่งชี้ของการสร้างสรรค์ และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ

2.2 แบบวัดการสร้างสรรค์และนวัตกรรมสำหรับนักศึกษามหาวิทยาลัยราชภัฏ โดยเป็นแบบมาตราส่วนประเมินค่า (rating scale) 5 ระดับ ซึ่งผ่านการพิจารณาความตรงเชิงเนื้อหา (content validity) โดยผู้เชี่ยวชาญที่มีความรู้ความเชี่ยวชาญด้านจิตวิทยาและการสร้างสรรค์และนวัตกรรม เป็นผู้สำเร็จการศึกษาระดับปริญญาเอก มีประสบการณ์ด้านการสอนหรือทำงานในสาขาวิชาที่เชี่ยวชาญไม่น้อยกว่า 5 ปี จำนวน 5 ท่าน ได้ค่าดัชนีความสอดคล้อง (item objective congruence : IOC) มีค่าถามทั้งหมด 24 ข้อ ผู้วิจัยได้นำแบบวัดไปทดลองใช้ (try out) กับนักศึกษาที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 150 คน แล้วนำข้อมูลที่ได้มาวิเคราะห์หาค่าอำนาจจำแนกเป็นรายข้อพบว่าอยู่ระหว่าง 0.38 - 0.65 โดยใช้วิธีการแจกแจงของที (t-distribution) และหาค่าความเชื่อมั่นโดยใช้วิธีสัมประสิทธิ์อัลฟา (Alpha coefficient) ของครอนบาค (Cronbach) ได้ค่าความเชื่อมั่นของแบบวัดดังนี้ องค์ประกอบด้านปัญญาสร้างสรรค์เท่ากับ 0.71 องค์ประกอบด้านจิตใจเชิงสร้างสรรค์เท่ากับ 0.82 องค์ประกอบด้านปฏิสัมพันธ์ทางสังคมเท่ากับ 0.82 องค์ประกอบด้านการสร้างนวัตกรรมเท่ากับ 0.79 ซึ่งแบบวัดมีค่าความเชื่อมั่นทั้งฉบับเท่ากับ 0.92

3. การเก็บรวบรวมข้อมูล

ก่อนดำเนินการเก็บรวบรวมข้อมูล ผู้วิจัยดำเนินการขอรับรองจริยธรรมการวิจัยในมนุษย์ จากคณะกรรมการจริยธรรมการวิจัยฯ ของมหาวิทยาลัยมหาสารคาม ซึ่งผลการพิจารณาจากคณะกรรมการฯ ได้รับรองจริยธรรมการวิจัยฯ เลขที่การรับรอง 219/2563 จากนั้นผู้วิจัยติดต่อมหาวิทยาลัยราชภัฏทั้ง 8 แห่ง ที่เป็นตัวอย่างในการวิจัย เพื่อขอความอนุเคราะห์นำลิงก์ Google form แบบวัดการสร้างสรรค์และนวัตกรรม เพื่อเก็บข้อมูลกับนักศึกษาตามสาขาวิชาที่เป็นตัวอย่างเมื่อได้รับการตอบรับจากอาจารย์ประจำสาขาวิชาของแต่ละมหาวิทยาลัย ผู้วิจัยขอให้อาจารย์ประจำสาขาวิชาให้นักศึกษาในขณะที่อยู่ในเวลาเรียน โดยมีการประชุม

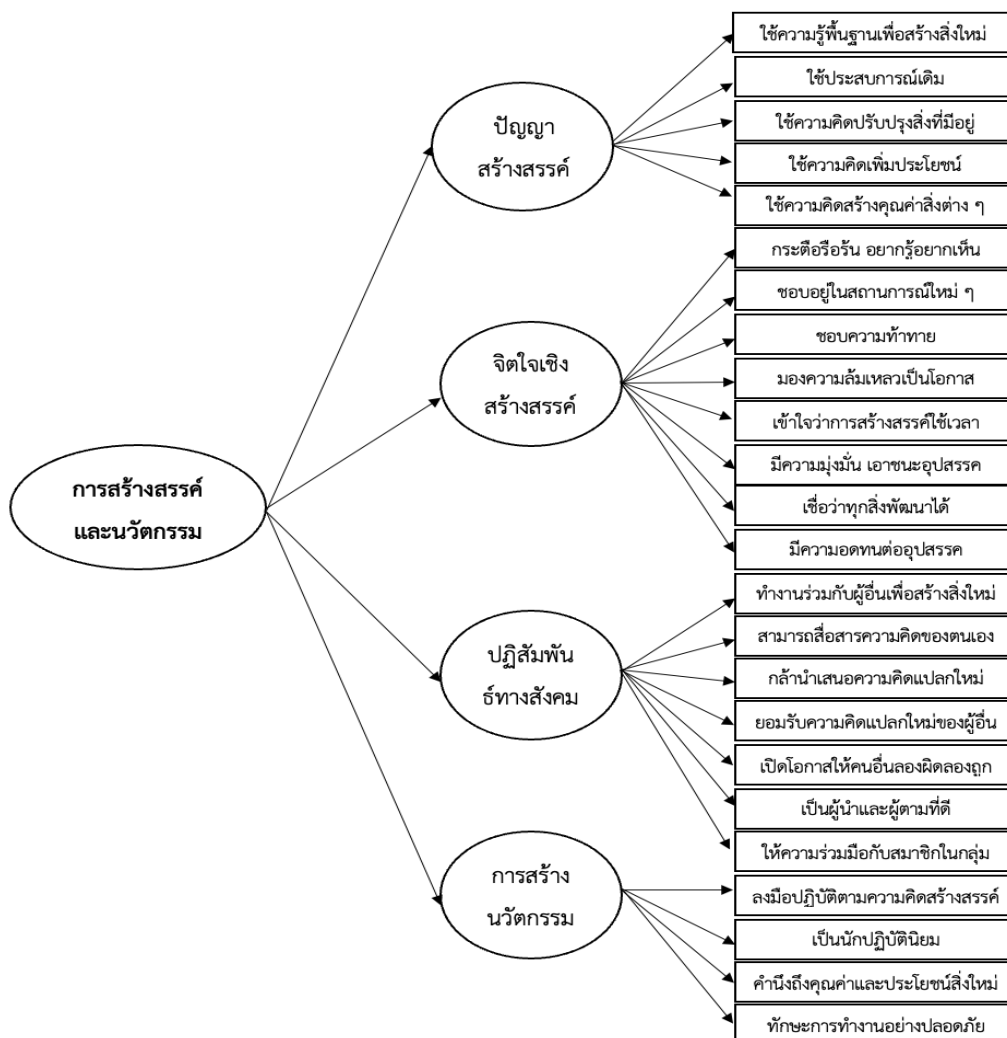
อาจารย์ที่เก็บข้อมูลทุกท่าน เพื่อทำความเข้าใจถึงขั้นตอน และวิธีเก็บข้อมูลเพื่อป้องกันการเกิดตัวแปรแทรก ผู้วิจัยตรวจสอบการตอบแบบวัดผ่านอินเทอร์เน็ตเพื่อตรวจสอบความผิดปกติของแบบวัดที่นักศึกษาตัวอย่างตอบกลับมา จากนั้นนำแบบวัดที่เก็บรวบรวมได้มาตรวจสอบความถูกต้องสมบูรณ์ โดยคัดเลือกแบบสอบถามที่มีคุณภาพตอบครบถ้วน จำนวน 1,200 ชุด คิดเป็นร้อยละ 100.00 ไม่มีข้อมูลสูญหาย

4. การวิเคราะห์ข้อมูล

ผู้วิจัยนำผลข้อมูลมาวิเคราะห์หาค่าความเที่ยงตรงเชิงโครงสร้างและหาค่าความเชื่อมั่นของแบบวัด โดยมีรายละเอียดดังนี้

1) วิเคราะห์ค่าสถิติ ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐานของแบบวัดการสร้างสรรค์และนวัตกรรมของ นักศึกษามหาวิทยาลัยราชภัฏ 2) วิเคราะห์ค่าความเที่ยงตรงของแบบวัดการสร้างสรรค์และนวัตกรรมของ นักศึกษา โดยการหาค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับนิยามเชิงปฏิบัติการ (item objective congruence: IOC) ตรวจสอบความสอดคล้องของข้อคำถามกับนิยาม 3) วิเคราะห์คุณภาพแบบวัดการ สร้างสรรค์และนวัตกรรมของนักศึกษาและหาค่าอำนาจจำแนกรายข้อ 4) วิเคราะห์หาค่าความเชื่อมั่นโดยรวม และรายองค์ประกอบของแบบวัด การสร้างสรรค์และนวัตกรรมของนักศึกษา และ 5) วิเคราะห์ความเที่ยงตรง เชิงโครงสร้าง โดยการวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA)

5. กรอบแนวคิดในการวิจัย



ภาพ 1 กรอบแนวคิดการวิจัย

ผลการวิจัย

ผลการศึกษาคำนำหน้าองค์ประกอบ การสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ พบว่า การสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ ประกอบด้วย 4 องค์ประกอบ ได้แก่ ปัญญาสร้างสรรค์, จิตใจเชิงสร้างสรรค์, ปฏิสัมพันธ์ทางสังคม และการสร้างนวัตกรรม ซึ่งมีรายละเอียดดังตาราง 1

ตาราง 1 แสดงคำนำหน้าองค์ประกอบ นิยามเชิงปฏิบัติการและตัวบ่งชี้การสร้างสรรค์และนวัตกรรมสำหรับ นักศึกษามหาวิทยาลัยราชภัฏ

องค์ประกอบ	นิยามเชิงปฏิบัติการ	ตัวบ่งชี้
1) ด้านปัญญา สร้างสรรค์ (creativity cognitive) คำนำหน้า 0.97	คุณลักษณะในการนำความรู้พื้นฐานหรือ ประสบการณ์เดิมที่มีอยู่มาปรับปรุง สร้างสรรค์เป็น ความคิดที่แปลกใหม่ แตกต่างจากเดิม ปรับปรุงสิ่ง เดิมให้ดีขึ้น มีประโยชน์มีคุณค่ามากขึ้น	1) ใช้ความรู้พื้นฐานเพื่อสร้างสิ่งใหม่ 2) ใช้ ประสบการณ์เดิม 3) ใช้ความคิดในการปรับปรุงให้ดีขึ้น 4) ใช้ความคิดที่ปรับปรุงสิ่งเดิมให้มีประโยชน์ 5) ใช้ความคิดสร้างคุณค่าของสิ่งต่าง ๆ
2) ด้านจิตใจเชิง สร้างสรรค์ (creativity affective) คำนำหน้า 0.90	คุณลักษณะภายในตัวบุคคลที่แสดงถึงความ กระตือรือร้นอยากรู้อยากเห็น กล้าเสี่ยง กล้าทำ หาย มองความล้มเหลวเป็นโอกาสในการเรียนรู้ อดทนต่ออุปสรรค	1) กระตือรือร้น อยากรู้อยากเห็น 2) ชอบอยู่ใน สถานการณ์ใหม่ ๆ 3) ชอบความท้าทาย 4) มอง ความล้มเหลวเป็นโอกาส 5) เข้าใจว่าการสร้างสรรค์ ต้องใช้เวลา 6) มีความมุ่งมั่น เอาชนะอุปสรรค 7) เชื่อว่าทุกสิ่งพัฒนาได้ 8) มีความอดทนต่ออุปสรรค
3) ด้านปฏิสัมพันธ์ ทางสังคม (social interaction) คำนำหน้า 0.92	คุณลักษณะในการสื่อสารความคิดให้ผู้อื่นเข้าใจใน ความคิดของตนเอง ความกล้าในการเสนอความคิด แปลกใหม่ของตนเองในกลุ่มเพื่อนร่วมงาน การ ยอมรับความคิดแปลกใหม่ของผู้อื่นเพื่อก่อให้เกิด การสร้างสรรค์สิ่งใหม่	1) การทำงานร่วมกับผู้อื่นสิ่งใหม่ 2) สื่อสาร ความคิดให้ผู้อื่นเข้าใจ 3) กล้าในการเสนอความคิด แปลกใหม่ 4) ยอมรับความคิดแปลกใหม่ของผู้อื่น 5) เปิดโอกาสให้สมาชิกลองผิดลองถูก 6) เป็นผู้นำ และผู้ตามที่ดี 7) ให้ความร่วมมือในกลุ่ม
4) ด้านการสร้าง นวัตกรรม implement innovations) คำนำหน้า 0.81	คุณลักษณะในการลงมือสร้างชิ้นงานให้เกิดเป็น นวัตกรรมใหม่เป็นรูปธรรม โดยมีความเป็นนัก ปฏิบัตินิยม คือ มีการ ตั้งเป้าหมาย การวางแผน การลงมือปฏิบัติ ตามแผน การนำนวัตกรรมไปใช้ การประเมิน และการปรับปรุงชิ้นงาน โดยคำนึงถึง คุณค่า และประโยชน์ของชิ้นงานใหม่ที่สร้างขึ้น	1) การลงมือปฏิบัติตามความคิดสร้างสรรค์เพื่อ สร้างชิ้นงานให้เกิดเป็นนวัตกรรมใหม่เป็นรูปธรรม 2) การเป็นนักปฏิบัตินิยม 3) คำนึงถึงคุณค่า และ ประโยชน์ของชิ้นงานใหม่ที่สร้างขึ้น 4) ทักษะการ ทำงานอย่างปลอดภัย

จากตาราง 1 สรุปได้ว่า คำนำหน้าองค์ประกอบของการสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏทั้ง 4 องค์ประกอบ มีค่าอยู่ระหว่าง 0.81 ถึง 0.97 โดยองค์ประกอบที่มีคำนำหน้าสูงสุดอันดับแรกคือ องค์ประกอบปัญญาเชิงสร้างสรรค์ ประกอบด้วย 5 ตัวบ่งชี้ มีคำนำหน้าองค์ประกอบเท่ากับ 0.97 รองลงมาคือ องค์ประกอบปฏิสัมพันธ์ทางสังคม ประกอบด้วย 7 ตัวบ่งชี้ มีคำนำหน้าองค์ประกอบเท่ากับ 0.90 และองค์ประกอบจิตใจเชิงสร้างสรรค์ ประกอบด้วย 8 ตัวบ่งชี้ มีคำนำหน้าองค์ประกอบเท่ากับ 0.92 และองค์ประกอบที่มีคำนำหน้าองค์ประกอบที่น้อยที่สุดคือ องค์ประกอบการสร้างนวัตกรรม ประกอบด้วย 4 ตัวบ่งชี้ มีคำนำหน้าองค์ประกอบเท่ากับ 0.81

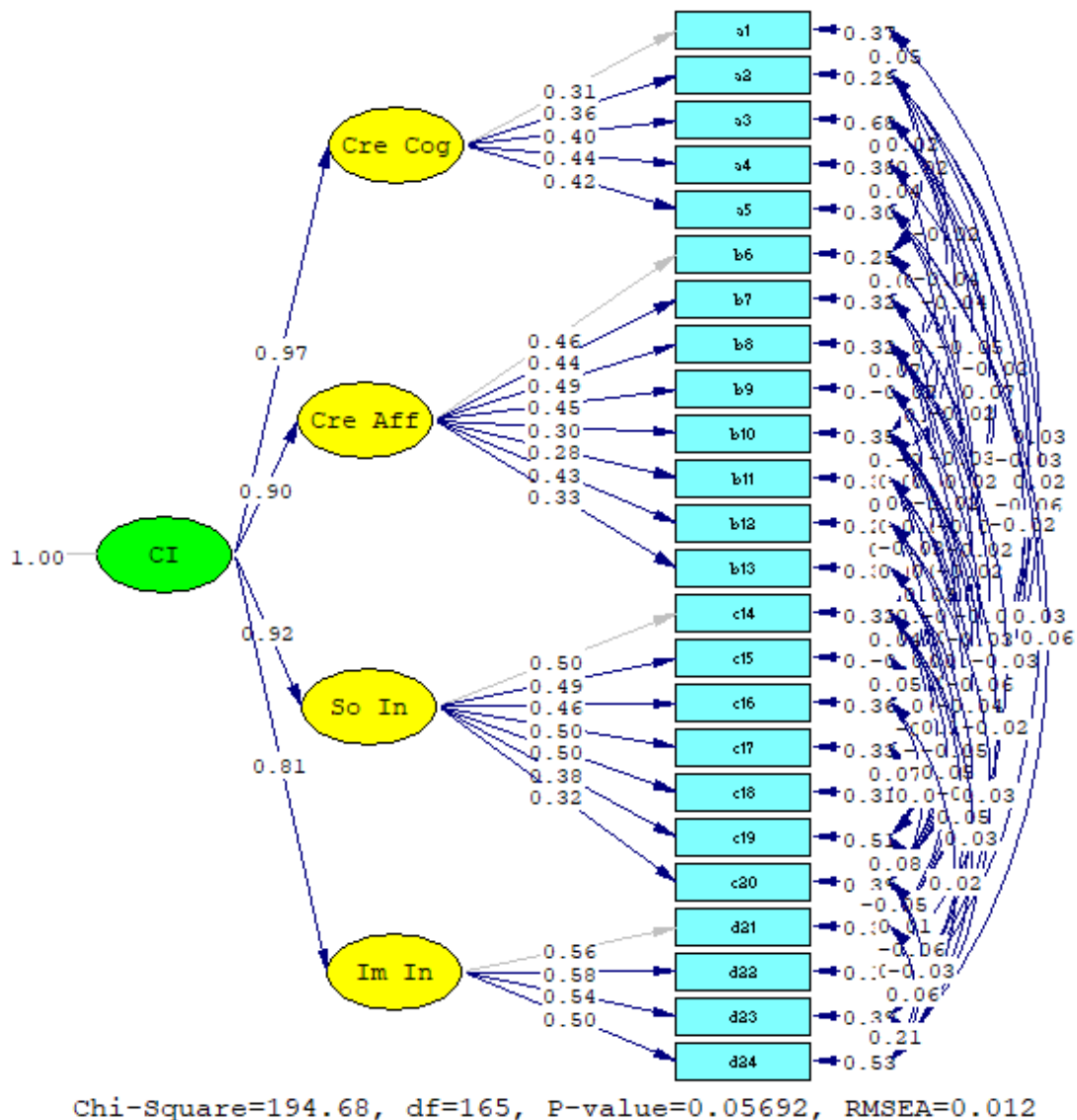
ดัชนีวัดระดับความกลมกลืนระหว่างโมเดลกับข้อมูลเชิงประจักษ์ ได้ค่าไค-สแควร์ (Chi-Square) เท่ากับ 194.68 ค่า df เท่ากับ 165 ค่า p-value เท่ากับ 0.06 ซึ่ง Chi-Square/df เท่ากับ 1.17 ค่า GFI เท่ากับ 0.99 และค่า AGFI เท่ากับ 0.98 ค่า RMSEA เท่ากับ 0.01 ซึ่งทุกตัวผ่านเกณฑ์การยอมรับ (Tayraukham, 2012) แสดงว่าโมเดลที่พัฒนาขึ้นสอดคล้องกับข้อมูลเชิงประจักษ์ในเกณฑ์ดี สามารถใช้เป็นองค์ประกอบ การสร้างสรรค์และนวัตกรรมได้ สามารถแสดงโมเดลผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ได้ดังตาราง 2 และภาพ 2

ตาราง 2 การวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลองค์ประกอบปัจจัยเชิงสาเหตุ

ตัวแปร	น้ำหนักองค์ประกอบ (b)	beta	ความคลาดเคลื่อนของตัวบ่งชี้ (SE)	สปส. การพยากรณ์ (R^2)	สปส. คะแนนองค์ประกอบ (FS)
a1	0.31**	0.45	(---)	0.21	0.09
a2	0.36**	0.56	0.03	0.32	0.18
a3	0.40**	0.44	0.04	0.20	0.10
a4	0.44**	0.58	0.03	0.33	0.14
a5	0.42**	0.61	0.03	0.37	0.17
b6	0.46**	0.68	(---)	0.46	0.11
b7	0.44**	0.62	0.02	0.38	0.17
b8	0.49**	0.65	0.03	0.42	0.25
b9	0.45**	0.62	0.03	0.39	0.16
b10	0.30**	0.46	0.02	0.20	0.11
b11	0.28**	0.42	0.02	0.18	-0.05
b12	0.43**	0.63	0.02	0.40	0.20
b13	0.33**	0.50	0.02	0.25	0.00
c14	0.50**	0.66	(---)	0.44	0.26
c15	0.49**	0.63	0.03	0.41	0.18
c16	0.46**	0.6	0.03	0.35	0.11
c17	0.50**	0.65	0.03	0.42	0.17
c18	0.50**	0.67	0.03	0.44	0.20
c19	0.38**	0.47	0.03	0.24	0.04
c20	0.32**	0.46	0.03	0.21	0.06
d21	0.56**	0.66	(---)	0.44	0.22
d22	0.58**	0.80	0.03	0.66	0.60
d23	0.54**	0.65	0.03	0.42	0.17
d24	0.50**	0.56	0.03	0.32	0.07

$\chi^2 = 194.68$ p-value = 0.06, df = 165, $\chi^2/df = 1.17$, RMSEA = 0.01, GFI = 0.99, AGFI = 0.98

หมายเหตุ : ** p < .01, (---) ไม่รายงานค่า SE และ t เนื่องจากเป็นพารามิเตอร์บังคับ



ภาพ 2 แสดงโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองที่สอดคล้องกับข้อมูลเชิงประจักษ์

อภิปรายผล

จากผลการวิจัย สามารถนำสู่การอภิปราย ดังนี้

การศึกษาองค์ประกอบและตัวบ่งชี้ของการสร้างสรรค์และนวัตกรรมสำหรับนักศึกษามหาวิทยาลัยราชภัฏ พบว่ามี 4 องค์ประกอบ 24 ตัวบ่งชี้ เป็นโมเดลที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์ ค่าสถิติวัดความกลมกลืนผ่านตามเกณฑ์ โดยมีความสอดคล้องในระดับดีทุกตัว (Diamantopoulos et al., 2000) แสดงว่า องค์ประกอบเชิงยืนยันที่พัฒนาขึ้นนี้มีความเที่ยงตรงเชิงโครงสร้างดี สามารถใช้ศึกษาการสร้างสรรค์และนวัตกรรมสำหรับนักศึกษาได้ การที่พบว่า โมเดลมีความเที่ยงตรงดี เนื่องจากองค์ประกอบที่นำมาวิเคราะห์เป็นองค์ประกอบองค์ที่มีความสำคัญในการพัฒนาคุณลักษณะในการสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ โดยเฉพาะอย่างยิ่งองค์ประกอบด้านปัญญาสร้างสรรค์ เป็นองค์ประกอบที่มีความสำคัญ

มากที่สุด ซึ่งแน่นอนว่าในการที่มนุษย์จะสร้างสรรค์สิ่งใดสิ่งหนึ่งที่ขาดไม่ได้ คือ ความคิดสร้างสรรค์ ซึ่งจากการศึกษา เอกสารงานวิจัย และการสัมภาษณ์ผู้เชี่ยวชาญเชิงลึกต่างให้ความสำคัญกับความคิดสร้างสรรค์ ว่าเป็นสิ่งแรกที่จะช่วยพัฒนาการสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ เพราะความคิดสร้างสรรค์เป็นสิ่งสำคัญสำหรับการสร้างนวัตกรรมเพราะเป็นจุดเริ่มต้น หากไม่มีความคิดสร้างสรรค์ก็จะไม่เกิดนวัตกรรม (Amabile, 1988) ดังนั้นความคิดสร้างสรรค์จึงเป็นขั้นตอนแรกและสำคัญที่สุดในการสร้างสรรค์นวัตกรรม ซึ่งในองค์ประกอบนี้ประกอบด้วย 5 ตัวบ่งชี้ โดยตัวบ่งชี้ในองค์ประกอบนี้สะท้อนถึงคุณลักษณะในการนำความรู้พื้นฐานหรือประสบการณ์เดิมที่มีอยู่มาปรับปรุง สร้างสรรค์เป็นความคิดที่แปลกใหม่ แตกต่างจากเดิม ปรับปรุงสิ่งเดิมให้ดีขึ้น มีประโยชน์มีคุณค่ามากขึ้น ซึ่งนำไปสู่การพัฒนาการสร้างสรรค์และนวัตกรรม สอดคล้องกับแนวคิดของ Buanoi (2016) ที่ให้นิยามว่า นวัตกรรมและความคิดสร้างสรรค์จึงมีความหมายที่ใกล้เคียงกัน คือ การปรับปรุง เปลี่ยนแปลงและพัฒนาสิ่งต่าง ๆ “ที่มีอยู่แล้ว ให้ดียิ่งขึ้น” และสอดคล้องกับ Partnership for 21st Century (2019) ที่เสนอว่า คิดสร้างสรรค์ (think creatively) คือ ใช้เทคนิคเพื่อสร้างการคิดที่หลากหลาย (เช่น กระบวน การระดมความคิด) สร้างแนวคิดใหม่ ๆ โดยนึกถึงประโยชน์ จำนวนและคุณภาพของการคิด และความสามารถในการประเมินความคิดของตนเอง วิเคราะห์ ปรับแต่งเปลี่ยนแปลงความคิด คิดในรายละเอียดและเพิ่มขีดความสามารถในการสร้างสรรค์ และสอดคล้องกับ Wongyai & Patphol (2014) ที่อธิบายพฤติกรรมบ่งชี้ของการคิดอย่างสร้างสรรค์ว่า มีลักษณะคิดริเริ่มในสิ่งที่มีประโยชน์ ใช้เทคนิควิธีการคิดหลายอย่าง คิดบนข้อมูลพื้นฐานความรู้ และสอดคล้องกับ Educational Quality Assurance Center King Mongkut's University of Technology North Bangkok (2014) ที่ให้ความหมายของการคิดสร้างสรรค์ (think creativity) ว่าหมายถึง การรู้จักการใช้เทคนิคการสร้างความคิดที่หลากหลาย สร้างแนวคิดใหม่และมีคุณค่า และสามารถอธิบาย ชัดเจน วิเคราะห์และประเมินแนวคิดของตนเองเพื่อปรับปรุงและพัฒนาการสร้างสรรค์ให้มากขึ้น และยังสอดคล้องกับ Somabut (2013) ที่เสนอว่า ความคิดสร้างสรรค์และนวัตกรรม (creativity and innovation) ประกอบด้วย การคิดสร้างสรรค์ (think creativity) เป็นการใช้เทคนิคของการสร้างสรรค์ทางความคิด ที่เปิดกว้าง เช่น การระดมสมอง สร้างสรรค์สิ่งแปลกใหม่และเสริมสร้างคุณค่าทางความคิดและสติปัญญา และมีความละเอียดรอบคอบต่อการคิดวิเคราะห์และประเมินแนวความคิดเพื่อนำไปสู่การปรับปรุงและพัฒนางานในเชิงสร้างสรรค์

องค์ประกอบด้านจิตใจเชิงสร้างสรรค์ ซึ่งเป็นองค์ประกอบที่มีความสำคัญในลำดับที่ 3 แต่ถือว่าเป็นองค์ประกอบที่มีความสำคัญมากเช่นกัน เพราะเนื่องจากเรื่องของแรงจูงใจภายในเพื่อการสร้างสรรค์ ซึ่งเป็นคุณลักษณะภายในตัวบุคคลที่แสดงถึงความกระตือรือร้นอยากรู้อยากเห็น กล้าเสี่ยง กล้าท้าทาย มองความล้มเหลวเป็นโอกาสในการเรียนรู้ อดทนต่ออุปสรรค อันจะสามารถพัฒนาการสร้างสรรค์และนวัตกรรมของนักศึกษาได้ เพราะหากนักศึกษามีความคิดสร้างสรรค์ มีความสามารถในการสร้างสรรค์แต่ขาดแรงจูงใจในการสร้างสรรค์ ก็เป็นไปได้ว่า การสร้างสรรค์ต่าง ๆ อาจไม่เกิดขึ้นตั้งแต่ต้นสอดคล้องกับ Amabile (1988) ที่เสนอว่า แรงจูงใจในตนเอง เกิดขึ้นจากภายในของตนเอง ความกระตือรือร้น ความรู้สึกท้าทายเมื่อเจอปัญหาการทำงานสำคัญ ความเชื่อมั่น ความมุ่งมั่นต่อความคืบหน้าต่อความคิดสร้างสรรค์และนวัตกรรมถึงร้อยละ 40 ความกล้าเสี่ยง กล้าทำอะไรรอบนอกกรอบแนว ชอบความท้าทาย มีผลพัฒนาความคิดสร้างสรรค์และ

นวัตกรรมถึง ร้อยละ 34 และยังกล่าวถึงปัจจัยที่ทำลายความคิดสร้างสรรค์ว่ามี 5 ประการ คือ การขาดแรงจูงใจ ร้อยละ 30 ความไม่ชำนาญ ร้อยละ 24 ความไม่ยืดหยุ่น ร้อยละ 22 การขาดแรงจูงใจภายนอก ร้อยละ 14 และการขาดทักษะทางสังคม ร้อยละ 7 ยังสอดคล้องกับ Morrison (2010) ที่เสนอว่า การสร้างสรรค์และนวัตกรรมต้องประกอบด้วย ความกล้าเสี่ยง คือ กล้าทดลอง คิดและทำสิ่งใหม่ ๆ และยังสอดคล้องกับผลการวิจัยของ Power (2016) ที่ว่าร้อยละร้อยละ 67 ของกลุ่มตัวอย่างต้องการโอกาสในการสร้างสรรค์และพัฒนาแรงจูงใจในการสร้างสรรค์โดยผู้สอน

องค์ประกอบที่ด้านปฏิสัมพันธ์ทางสังคม เป็นองค์ประกอบที่สำคัญของการสร้างสรรค์และนวัตกรรม เป็นลำดับที่ 2 เป็นเพราะว่าคุณลักษณะการทำงานร่วมกับผู้อื่นเพื่อสร้างสรรค์นวัตกรรมที่แปลกใหม่ การสื่อสารความคิดให้ผู้อื่นเข้าใจในความคิดของตนเอง ความกล้าในการเสนอความคิดแปลกใหม่ของตนเองในกลุ่มเพื่อนร่วมงาน การยอมรับความคิดแปลกใหม่ของผู้อื่นเพื่อก่อให้เกิดการสร้างสรรค์สิ่งใหม่ เป็นสิ่งที่ช่วยให้นักศึกษาพัฒนาการสร้างสรรค์และนวัตกรรม ดังที่ Partnership for 21st Century (2019) ที่เสนอว่า การสร้างสรรค์และนวัตกรรม เป็นความสามารถเพื่อการปรับตัวให้เข้ากับสังคม สภาวะแวดล้อม และสภาวะการทำงานที่มีความซับซ้อน มีจุดมุ่งเน้นเพื่อพัฒนาความคิดสร้างสรรค์ การคิดวิเคราะห์ การสื่อสารและการทำงานร่วมกับผู้อื่นเป็นสำคัญ การทำงานกับผู้อื่นอย่างสร้างสรรค์ (work creatively with others) คือ ความสามารถในการนำความคิดของตนเองไป พัฒนา ประยุกต์ใช้และสื่อสารกับผู้อื่นได้อย่างมีประสิทธิภาพ ใจกว้าง ยอมรับสิ่งใหม่ ๆ มีความคิดที่หลากหลาย การเสนอความคิดต่อกลุ่มและกล้าให้ข้อเสนอแนะในการทำงาน กล้าเสนอความคิดริเริ่มแปลกใหม่และสร้างสรรค์ในการทำงาน พร้อมทั้งเข้าใจข้อจำกัดของโลก ในความเป็นจริงในการนำเสนอแนวคิดนั้นเห็นความล้มเหลวเป็นโอกาสในการเรียนรู้และพัฒนาตน เข้าใจว่า การสร้างสรรค์และนวัตกรรมใหม่ ๆ จะสามารถนำไปสู่ความสำเร็จ ไม่กลัวความผิดพลาดเล็ก ๆ น้อย ๆ ยังสอดคล้องกับ Wongyai & Patphol (2014) ที่เสนอว่า องค์ประกอบหนึ่งของการสร้างสรรค์และนวัตกรรม คือ การทำงานร่วมกับบุคคลอื่นอย่างสร้างสรรค์ (work creatively with others) สื่อสารความคิดของตนเองกับผู้อื่นได้อย่างมีประสิทธิภาพ เปิดรับและตอบสนองความคิดเห็นใหม่ ๆ ของบุคคลอื่น แสดงความคิดริเริ่มในการปฏิบัติงานและปรับให้สอดคล้องกับบริบท ทำงานร่วมกับบุคคลอื่นด้วยความร่วมมือร่วมใจ แลกเปลี่ยนเรียนรู้กับบุคคลอื่นเพื่อความสำเร็จของงาน เคารพความคิดของคนอื่นทั้งที่สอดคล้องและไม่สอดคล้องกับความคิดของตน ยังสอดคล้องกับ Educational Quality Assurance Center King Mongkut's University of Technology North Bangkok (2014) ที่ให้องค์ประกอบของการคิดสร้างสรรค์และนวัตกรรม ด้านหนึ่งคือ การทำงานอย่างสร้างสรรค์ร่วมกับผู้อื่น (work creativity with others) สามารถพัฒนานำไปปฏิบัติและสื่อสารความคิดใหม่ ๆ ไปสู่ผู้อื่นได้อย่างมีประสิทธิภาพ เปิดใจกว้างและยอมรับมุมมองหรือโลกทัศน์ใหม่ ๆ ที่มีความหลากหลายให้ความเห็นต่อกลุ่มในการทำงาน แสดงถึงความเป็นผู้นำในการคิดริเริ่มสร้างสรรค์ในการทำงาน และเข้าใจสภาพการณ์ที่อาจเป็นข้อจำกัดในการรับความคิดใหม่ ๆ และสามารถนำความผิดพลาดมาเป็นโอกาสในการเรียนรู้ เข้าใจว่าการสร้างสรรค์และนวัตกรรมเป็นกระบวนการที่ต้องใช้เวลา เป็นกระบวนการที่มีวัฏจักรของความสำเร็จและความผิดพลาด และ Chantago (2019) ที่เสนอว่า ทักษะในการทำงานร่วมกับผู้อื่นอย่างสร้างสรรค์ คือ การพัฒนา ลงมือปฏิบัติ และสื่อสารมุมมองใหม่กับผู้อื่นเสมอ เปิดใจรับและตอบสนองต่อ

มุมมองใหม่ๆ เปิดใจรับข้อคิดเห็นจากกลุ่ม รวมถึงการประเมินผลงานจากกลุ่มเพื่อนำไปปรับปรุง ทำงานด้วยความคิดหรือวิธีการใหม่ๆ และเข้าใจข้อจำกัดของโลกในการยอมรับมุมมองใหม่ๆ ยังสอดคล้องกับ Amabile (1988) ที่ได้สรุปคุณลักษณะของบุคคลที่ที่มีการสร้างสรรค์และนวัตกรรม ว่าคุณภาพของกลุ่ม คือ การทำงานร่วมกัน มองในด้านสติปัญญาและสังคม การทำงานเป็นทีม ส่งผลต่อความคิดสร้างสรรค์ถึงร้อยละ 30

ถึงแม้ว่าองค์ประกอบด้านการสร้างนวัตกรรมจะเป็นองค์ประกอบที่มีความสำคัญน้อยที่สุด แต่เป็นองค์ประกอบที่สามารถตอบคุณลักษณะการสร้างสรรค์และนวัตกรรมได้อย่างชัดเจนเป็นรูปธรรม เพราะหา นักศึกษามีใจที่จะสร้างสรรค์ มีความคิดสร้างสรรค์ และมีทักษะการทำงานร่วมกับผู้อื่นที่ดี แต่ไม่เริ่มที่จะลงมือทำอะไรให้เป็นรูปธรรม การสร้างสรรค์ต่าง ๆ ที่เกิดก่อนหน้าที่ก็อาจไม่มีประโยชน์ใด ๆ ดังนั้นคุณลักษณะในการลงมือปฏิบัติเพื่อประดิษฐ์นวัตกรรมจากแนวคิดในการสร้างสรรค์ การลงมือสร้างชิ้นงานให้เกิดเป็นนวัตกรรมใหม่เป็นรูปธรรม โดยมีความเป็นนักปฏิบัตินิยม คือ มีการตั้งเป้าหมาย การวางแผน การลงมือปฏิบัติตามแผน การนำนวัตกรรมไปใช้ การประเมินและการปรับปรุงชิ้นงาน โดยคำนึงถึงคุณค่าและประโยชน์ของชิ้นงานใหม่ที่สร้างขึ้น เป็นสิ่งที่ช่วยในการพัฒนาการสร้างสรรค์และนวัตกรรมของนักศึกษา ดังที่ Johnson (2010) ที่เสนอว่า ความสร้างสรรค์และนวัตกรรมคือความสามารถในการสร้างหรือประดิษฐ์สิ่งที่เป็นต้นแบบ และสอดคล้องกับ Partnership for 21st Century (2019) ที่เสนอว่า การสร้างนวัตกรรม (implement innovations) คือ สามารถสร้างนวัตกรรมตามสิ่งที่คิดสร้างสรรค์นั้นให้เกิดเป็นรูปธรรมเพื่อส่วนรวมและเกิดประโยชน์ต่อวงการนั้น ๆ ยังสอดคล้องกับ Panit (2013) ที่ให้ความหมายของความคิดสร้างสรรค์และนวัตกรรม หมายถึง การประยุกต์ใช้จินตนาการและการประดิษฐ์ เป็นทักษะที่สอนโดยตรงไม่ได้แต่สามารถจัดกระบวนการให้เกิดการเรียนรู้ได้ ยังสอดคล้องกับ Wongyai & Patphol (2014) ที่เสนอว่า องค์ประกอบหนึ่งของการสร้างสรรค์และนวัตกรรม คือ การสร้างสรรค์นวัตกรรมให้สำเร็จ (implement innovations) เป็นการวางแผน ดำเนินการพัฒนานวัตกรรมที่สืบเนื่องมาจากความคิดสร้างสรรค์ การดำเนินการพัฒนานวัตกรรมตามแผนที่กำหนดไว้อย่างมีประสิทธิภาพการประเมินคุณภาพของนวัตกรรมที่พัฒนาขึ้นโดยใช้ข้อมูลเชิงประจักษ์ และปรับปรุงแก้ไขจุดบกพร่องของนวัตกรรมให้ดีขึ้นอย่างต่อเนื่อง ยังสอดคล้องกับ Educational Quality Assurance Center King Mongkut's University of Technology North Bangkok (2014) ที่ให้องค์ประกอบของการคิดสร้างสรรค์และนวัตกรรมว่า เป็นการนำนวัตกรรมสู่การปฏิบัติ (implement innovations) สามารถดำเนินการให้เกิดการนำนวัตกรรมไปใช้ และสอดคล้องกับ Chantago (2019) ที่เสนอว่า การประยุกต์สู่นวัตกรรม คือ การลงมือปฏิบัติตามความคิดสร้างสรรค์เพื่อนำไปสู่ผลสำเร็จที่เป็นรูปธรรม ยังสอดคล้องกับ Amabile (1988) ที่ได้สรุปคุณลักษณะของบุคคลที่ที่มีการสร้างสรรค์และนวัตกรรม ว่าความเชี่ยวชาญในงานพรสวรรค์ และประสบการณ์ความรู้ที่ได้รับเฉพาะงานส่งผลต่อความคิดสร้างสรรค์ถึงร้อยละ 30

สรุปได้ว่า องค์ประกอบและตัวบ่งชี้การสร้างสรรค์และนวัตกรรมของนักศึกษามหาวิทยาลัยราชภัฏ ประกอบด้วย ปัญญาสร้างสรรค์, จิตใจเชิงสร้างสรรค์, ปฏิสัมพันธ์ทางสังคม และการสร้างนวัตกรรม

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 ผลการวิจัยครั้งนี้พบว่า องค์ประกอบปัญญาสร้างสรรค์มีความสำคัญสูงสุด ดังนั้นอาจารย์ผู้สอน หากต้องการพัฒนานักศึกษามหาวิทยาลัยราชภัฏให้เกิดการสร้างสรรค์และนวัตกรรม ควรให้ความสำคัญในการพัฒนาด้านปัญญาสร้างสรรค์ เช่น การใช้ความรู้พื้นฐานเพื่อสร้างสิ่งใหม่ การใช้ประสบการณ์เดิมเพื่อสร้างสิ่งใหม่ การใช้ความคิดในการปรับปรุงสิ่งที่มีอยู่ให้ดีขึ้น การใช้ความคิดที่ปรับปรุงสิ่งที่มีอยู่เดิมให้มีประโยชน์มากขึ้น และการใช้ความคิดในการสร้างคุณค่าของสิ่งต่าง ๆ ให้เพิ่มขึ้น

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรทำการศึกษากับตัวอย่างในระดับชั้นหรือสถาบันการศึกษาอื่น เช่น ระดับชั้นประถมศึกษา มัธยมศึกษา ประกาศนียบัตรวิชาชีพ การศึกษานอกระบบ และการศึกษาตามอัธยาศัย เป็นต้น เพื่อนำผลการศึกษามาเปรียบเทียบกับมีความเหมือนหรือแตกต่างจากผลการวิจัยในครั้งนี้อย่างไร

2.2 นอกจากการศึกษาเรื่องการสร้างสรรค์และนวัตกรรมด้วยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยัน ในการศึกษารุ่นต่อไป ควรมีการศึกษาด้วยระเบียบวิธีวิจัยอื่น ๆ เช่น การวิจัยและพัฒนา การวิจัยทฤษฎีฐานราก การวิจัยอนาคตภาพ การวิจัยแบบข้ามวัฒนธรรม และการพัฒนารูปแบบ เป็นต้น

References

- Amabile, T. M. (1988). A Model of Creativity and Innovation in Organizations. *Research in Organizational Behavior*, 10(1), 123-167.
- Batey, M. (2011). *Is Creativity the Number 1 Skill for the 21st Century?* *Psychology Today*.
<https://www.psychologytoday.com/us/blog/working-creativity/201102/is-creativity-the-number-1-skill-the-21st-century>.
- Diamantopoulos, A., Siguaw, J. A., & Cadogan, J. W. (2000). Export Performance: The Impact of Cross-Country Export Market Orientation. *In American Marketing Association Conference Proceedings*, 11(1), 177.
- Johnson, M. W. (2010). The Time has come for Business Model Innovation. *Leader to Leader*, 2010(57), 6-10.
- Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). *Introduction to Bivariate and Multivariate Analysis*. Scott Foresman & Company.
- Morrison, M. (2010). *Creativity and Innovation Models - applying in practice*. <https://rapidbi.com/creativity-and-innovation-models-applying-in-prac/>
- Partnership for 21st Century. (2019). *Learning and Innovation Skills-4Cs Key Subjects-3Rs and 21st Century Framework for 21st Century*. 2001: One Massachusetts Avenue, Suite 700. www.P21.org

Power, J. B. (2016). An Investigation into the Factors Affecting Student Creativity in Higher Education in Thailand. *Thammasat Review*, 18(1), 177–198.

Translate Thai References

- Buanoi, D. (2016). Innovation and Creativity. *Journal of Marketing and Management Rajamangala University of Technology Thanyaburi*, 3(1), 1-12. (in Thai)
- Chantago, J. (2019). *Teaching in 21st Century. Teaching Materials*. Mahamakut Buddhist University Isan Campus. Khon Kaen. (in Thai)
- Chinnawong, W. (2001). Integrated Learning and Teaching. *Academic Journal*, 4(10), 140-155. (in Thai)
- Educational Quality Assurance Center King Mongkut's University of Technology North Bangkok. (2014). *21st Century Learning Skills*. KMUTNB, 303(1). (in Thai)
- Kachintorn, U. (2018). *The Solution of Thai Higher Educational Crisis*. Graduate School Development and Research Chiang Mai University. (in Thai)
- Ministry of Higher Education, Science, Research and Innovation. (2021). *University Student Data is divided by educational level and gender*. Ministry of Higher Education, Science, Research and Innovation. https://data.go.th/dataset/univ_std_11_01 (in Thai)
- Muangsurriya, R. (2016). Community Development: Rethinking Issues in Changing Society. *Journal of Social Development*, 18(1), 123–141. (in Thai)
- Panit, W. (2013). *Teachers for Learners to build flipped classroom*. www.scbfoundation.com. (in Thai)
- Piyakun, A. (2016). *Teaching Document for 0502 101 Psychology for Teachers Revised Edition 2016*. Educational Psychology and Guidance. Faculty of Education Mahasarakham University. (in Thai)
- Somabut, A. (2013). *Constructivist Theory*. <https://teacherweekly.wordpress.com/2013/09/25/> (in Thai)
- Tayraukham, S. (2012). *Research Methodology for Humanities and Social Sciences*. Mahasarakham University. (in Thai)
- Wongyai, W., & Patphol, M. (2014). *Coaching for Thinking* (3rd Ed.). Charansanitwong Printing. (in Thai)

อิทธิพลพัฒนาการการลดความสิ้นเปลืองในการใช้จ่ายภาครัฐที่มีต่อพัฒนาการความร่วมมือ
ด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบ

การศึกษา : การวิเคราะห์โมเดลโค้งพัฒนาการคู่ขนาน

Effect of Wastefulness of Government Spending Reduction Development on
Growth of University-Industry Collaboration in R&D and Quality of the
Education System: A Parallel Latent Growth Curve Modeling

ธีรยุทธ พิริยะอารยะกุล^{1*} ณัฐพล อนันต์ธนสาร² และ ชัยุตม์ ภิรมย์สมบัติ³

Theerayuth Phiriyaarayakul ^{1*}, Nattapon Anantanasan ² and Chayut Piromsombat³

(Received: April 16, 2022; Revised: May 19, 2022; Accepted: May 24, 2022)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมาย 1) เพื่อวิเคราะห์พัฒนาการของการลดความสิ้นเปลืองในการใช้จ่ายภาครัฐ ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา และ 2) เพื่อวิเคราะห์อิทธิพลเชิงสาเหตุของการลดความสิ้นเปลืองในการใช้จ่ายภาครัฐที่มีต่อความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัย และภาคอุตสาหกรรมและคุณภาพระบบการศึกษา การวิจัยนี้ นำเสนอผลการวิเคราะห์ข้อมูลทุติยภูมิด้วยโมเดลโค้งพัฒนาการคู่ขนานเพื่อวิเคราะห์วิถีพัฒนาการของ ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา ในช่วงเวลา 10 ปี (2009 - 2018) จากข้อมูลของตัวอย่างวิจัย จำนวน 126 ประเทศทั่วโลก ($N = 126$) การวิเคราะห์ข้อมูล ได้แก่ สถิติแบบบรรยาย และเทคนิคการวิเคราะห์ตัวแปรพหุนาม ประกอบด้วย ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน โมเดลโค้งพัฒนาการ และโมเดลโค้งพัฒนาการคู่ขนาน

ผลการวิจัย พบว่า โมเดลโค้งพัฒนาการของการลดความสิ้นเปลืองในการใช้จ่ายภาครัฐ (χ^2 (3, $N = 126$) = .3.439, $p = .328$, relative chi-square = 1.146, TLI = .997, CFI = .998, RMSEA = .036, SRMR = .018) ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม (χ^2 (4, $N = 126$) = 4.717, $p = .312$, relative chi-square = 1.179, TLI = .997, CFI = .998, RMSEA = .039,

¹ อาจารย์ประจำบัณฑิตวิทยาลัย มหาวิทยาลัยรัตนบัณฑิต

² นักวิจัยประจำสถาบันศึกษานโยบายสาธารณะ

³ อาจารย์ประจำสาขาวิชาวิธีวิทยาการพัฒนานวัตกรรมทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹ Lecturer, Graduate School, Rattana Bundit University

² Researcher, Public Policy Studies Institute

³ Lecturer, Methodology for Innovation Development in Education Division, Faculty of Education, Chulalongkorn University

* Corresponding Author E-mail: theerayuth@outlook.com

SRMR = .077) และคุณภาพระบบการศึกษา (χ^2 (5, $N = 126$) = .920, $p = .314$, relative chi-square = .184, TLI = .997, CFI = .998, RMSEA = .040, SRMR = .058) มีความสอดคล้องกับข้อมูลเชิงประจักษ์

ผลการวิเคราะห์โมเดลโค้งพัฒนาการคู่ขนาน พบว่า ประเทศที่มีคุณภาพระบบการศึกษาสูงอยู่แล้ว มีพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมเพียงเล็กน้อย ($\beta = -.298$, $p < .001$) แต่อย่างไรก็ตาม ประเทศที่มีความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมระดับสูงมีพัฒนาการของคุณภาพระบบการศึกษาเพียงเล็กน้อยเช่นกัน ($\beta = -.382$, $p < .001$) นอกจากนี้ ประเทศที่สามารถลดความสิ้นเปลืองในการใช้จ่ายภาครัฐได้สูงขึ้นกลับทำให้พัฒนาการของคุณภาพระบบการศึกษาและความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมลดลงในระยะยาว ($\beta = -.644$, $p < .001$ และ $\beta = -.155$, $p = .267$ ตามลำดับ)

คำสำคัญ: ดัชนีความสามารถทางการแข่งขันระดับโลก การลดความสิ้นเปลืองในการใช้จ่ายภาครัฐ
ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม
คุณภาพระบบการศึกษา โมเดลโค้งพัฒนาการคู่ขนาน

Abstract

The purposes of this study were 1) to analyze a latent growth curve model of development of wastefulness of government spending reduction, university-industry collaboration in R&D, and quality of the education system, and 2) to analyze a parallel latent growth curve model of wastefulness of government spending reduction development on growth of university-industry collaboration in R&D and quality of the education system. The present study reports the results of a secondary data analysis that used parallel latent growth curve modeling to examine growth trajectories of university-industry collaboration in R&D and quality of the education system over a ten-year period (2009 - 2018) with data from a sample ($N = 126$) of countries across the world. Data analysis was performed by using descriptive, and multivariate analysis technique including the mean, standard deviation, latent growth curve modeling (LGCM), and parallel latent growth curve modeling (Parallel LGCM).

The results suggest that the latent growth curve model of wastefulness of government spending reduction development (χ^2 (3, $N = 126$) = .3.439, $p = .328$, relative chi-square = 1.146, TLI = .997, CFI = .998, RMSEA = .036, SRMR = .018), the university-industry collaboration in R&D (χ^2 (4, $N = 126$) = 4.717, $p = .312$, relative chi-square = 1.179, TLI = .997, CFI = .998, RMSEA = .039, SRMR = .077), and the quality of the education system (χ^2 (5, $N = 126$) = .920, $p = .314$, relative chi-square = .184, TLI = .997, CFI = .998, RMSEA = .040, SRMR = .058) were valid and well fitted to the empirical data.

The parallel latent growth curve model found that countries with high levels of quality of the education system had a low university-industry collaboration in R&D ($\beta = -.298, p < .001$), while countries with high levels of university-industry collaboration in R&D had a low quality of the education system ($\beta = -.382, p < .001$). Furthermore, a negative linear change in wastefulness of government spending reduction reduced the effects of university-industry collaboration in R&D and quality of the education system longitudinally ($\beta = -.644, p < .001$ and $\beta = -.155, p = .267$ respectively).

Keywords: Global Competitiveness Index, wastefulness of government spending reduction, university-industry collaboration in R&D, quality of the education system, parallel latent growth curve modeling

บทนำ

ยุคปัจจุบันการเติบโตทางเศรษฐกิจที่มีการเปลี่ยนแปลงอย่างรวดเร็วและเกิดความผันแปรจากปัจจัยอื่น ๆ ที่ไม่สามารถคาดคิดได้ ชี้ยังทำนายแนวโน้มการเปลี่ยนแปลงที่เกิดขึ้นนั้นได้ยากขึ้น ในขณะที่ นวัตกรรมก็ก้าวเข้ามามีความสำคัญเพิ่มมากขึ้นในระบบเศรษฐกิจ รวมทั้ง นวัตกรรมเป็นตัวขับเคลื่อนทิศทางเศรษฐกิจของประเทศเห็นได้จากกระบวนทัศน์ทางนวัตกรรมในภาคอุตสาหกรรมที่เปลี่ยนไปเพื่อยกระดับความได้เปรียบในการแข่งขันของตตนันผลักดันให้เกิดความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมเพิ่มมากขึ้นเป็นเงาตามตัว โดยทั้งสองภาคส่วนต่างก็ตั้งเป้าหมายในการแลกเปลี่ยนความรู้และทักษะต่าง ๆ เพิ่มขึ้น รวมทั้ง ความร่วมมือลักษณะนี้ทำให้ภาคอุตสาหกรรมลดต้นทุนค่าจ้างแรงงานที่ต้องฝึกอบรมลง พร้อมทั้งมหาวิทยาลัยสามารถผลิตบัณฑิตให้ตรงตามความต้องการของภาคอุตสาหกรรมด้วยเช่นกัน (Milovanovic, 2015)

แต่อย่างไรก็ตาม การพัฒนาการของการลดความสิ้นเปลืองในการใช้จ่ายภาครัฐที่นอกจากจะสะท้อนให้เห็นถึงความโปร่งใสของรัฐแล้วยังเป็นปัจจัยที่มีอิทธิพลเป็นอย่างมากต่อการพัฒนาคุณภาพระบบการศึกษาของประเทศนั้นอีกด้วย (Kokins & Lavrinovics, 2006) กล่าวคือ บทบาทของภาครัฐในการสนับสนุนการศึกษานั้นสังเกตได้จากค่าใช้จ่ายในภาครัฐที่เกี่ยวข้องกับการวิจัยและพัฒนาทั้งการศึกษาและภาคอุตสาหกรรม และนโยบายรัฐในการปรับฐานความรู้ของประชาชนในประเทศ (Olsson & Meek, 2014; Djeflat, 2015) แต่่นโยบายการลดค่าใช้จ่ายยังส่งผลให้รัฐบาลในแต่ละประเทศลดความช่วยเหลือทางการเงินแก่สถาบันอุดมศึกษาลง (Lubienski, 2016; Rizvi, 2016; Carrasco & Gunter, 2019) จากที่กล่าวมาข้างต้นสถาบันอุดมศึกษาจำเป็นต้องหันมาปรับกระบวนทัศน์ทางการศึกษาด้วยการมุ่งเน้นการพึ่งพากันและร่วมกันพัฒนานวัตกรรมกับภาคอุตสาหกรรมเพิ่มมากขึ้น ซึ่งจะยังประโยชน์ให้ทั้งสองฝ่าย

แม้ว่าการปรับปรุงคุณภาพทางการศึกษาช่วยขับเคลื่อนประสิทธิภาพทางเศรษฐกิจของประเทศเติบโตขึ้นอย่างเห็นได้ชัดก็ตาม (Carnoy et al., 2013) แต่ทว่า ระบบการศึกษามักถูกวิพากษ์วิจารณ์ว่า หลักสูตรไม่ทันสมัย

และวิชาการมากเกินไปจนไม่สามารถนำไปปฏิบัติจริงได้ (Aleksejeva, 2016) รวมถึงการเคลื่อนไหวเพื่อปฏิรูปการศึกษาทั่วโลก (Global Educational Reform Movement: GERM) นอกจากช่วยพัฒนาและปรับปรุงคุณภาพการศึกษาแล้ว (Carrasco & Gunter, 2019) ยังส่งผลให้แต่ละประเทศเพิ่มความร่วมมือและแลกเปลี่ยนความรู้กับภาคอุตสาหกรรมมากขึ้นในทุกระดับการศึกษา โดยเฉพาะด้านเทคโนโลยีและเน้นการฝึกทักษะการคิดวิเคราะห์ให้กับผู้เรียนเพิ่มมากขึ้น (Hayes, 2015)

การประชุมผู้นำเศรษฐกิจของโลก (World Economic Forum: WEF) ได้จัดทำดัชนีความสามารถทางการแข่งขันระดับโลก (Global Competitiveness Index: GCI) ที่สามารถชี้ให้เห็นถึงพัฒนาการของประสิทธิภาพในความสามารถทางการแข่งขันของแต่ละประเทศ โดยเฉพาะปัจจัยประสิทธิภาพในการใช้จ่ายภาครัฐ (efficiency of government spending) ต่างส่งผลต่อปัจจัยคุณภาพระบบการศึกษา (quality of the education system) และปัจจัยความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม (university-industry collaboration in R&D) (Schwab, 2018) รวมถึงมีลักษณะการทำงานแบบเชิงระบบที่สามารถบ่งชี้ประสิทธิภาพในการใช้ทรัพยากรในการปรับปรุงคุณภาพระบบการศึกษาและส่งเสริมความร่วมมือระหว่างภาคการศึกษาและภาคอุตสาหกรรมของประเทศต่าง ๆ ในโลกได้ (Olsson & Meek, 2014; Djeflat, 2015)

จากการทบทวนงานวิจัยที่ผ่านมา คณะผู้วิจัยเห็นว่า การศึกษาบทบาทของภาครัฐในการพัฒนาระบบการศึกษานั้นเป็นสิ่งที่สำคัญในการทำความเข้าใจพัฒนาการของช่วงเวลาที่ผ่านมา เช่น งานวิจัยของ Djeflat (2015) ได้ศึกษาเอกสารเกี่ยวกับนโยบายของประเทศโมร็อกโคที่ส่งเสริมการเปลี่ยนแปลงการศึกษาของชาติไปสู่สังคมแห่งการเรียนรู้และนวัตกรรมเพื่อขับเคลื่อนเศรษฐกิจของประเทศ โดยบ่งชี้ถึงนโยบายรัฐที่กำหนดยุทธศาสตร์การวิจัยแห่งชาติให้สามารถขับเคลื่อนนวัตกรรมได้ด้วยเทคโนโลยีในประเทศที่พัฒนาขึ้นมาและการสร้างมูลค่าเพิ่มให้กับนวัตกรรมที่มีอยู่ ในขณะที่งานวิจัยของ Olsson and Meek (2014) ได้ศึกษาประสิทธิภาพของนโยบายรัฐและช่องว่างความรู้ในการวิจัยและนวัตกรรมของกลุ่มประเทศเอเชียตะวันออกเฉียงใต้ด้วยดัชนีวิทยาศาสตร์ เทคโนโลยีและนวัตกรรมขององค์การเพื่อความร่วมมือและการพัฒนาทางเศรษฐกิจ (OECD. STI Scoreboard) โดยชี้ให้เห็นถึงความสำคัญของนโยบายรัฐในการส่งเสริมความร่วมมือด้านการวิจัยและพัฒนาที่มีต่อประสิทธิภาพโดยรวมของคุณภาพการศึกษา จากที่กล่าวมาข้างต้นเห็นได้ว่า นโยบายของภาครัฐสามารถส่งเสริมคุณภาพระบบการศึกษาในภาพรวมได้ โดยอาศัยกระบวนการขับเคลื่อนด้านการวิจัยและพัฒนาทั้งในมหาวิทยาลัยและภาคอุตสาหกรรม

แม้ว่าปรากฏงานวิจัยของ Oturakci (2021) ที่ศึกษา คุณลักษณะที่แตกต่างของแต่ละประเทศที่ทำให้ความสามารถในการแข่งขันแตกต่างกัน และงานของ Sohn et al. (2016) ที่ศึกษาอิทธิพลของปัจจัยต่าง ๆ ที่มีต่อความสามารถในการสร้างสรรค์นวัตกรรมในลักษณะการศึกษาแบบภาคตัดขวาง แต่อย่างไรก็ตามงานวิจัยที่ผ่านมายังขาดหลักฐานสนับสนุนและขาดงานวิจัยที่ศึกษาพัฒนาการและอิทธิพลเชิงสาเหตุของการเปลี่ยนแปลงของปัจจัยต่าง ๆ ที่ส่งผลต่อความร่วมมือด้านการวิจัยและพัฒนา และคุณภาพระบบการศึกษาในระยะยาวได้ คณะผู้วิจัยสนใจศึกษาโมเดลโค้งพัฒนาการ (latent growth curve modeling: LGCM) ที่ใช้ค่าเริ่มต้นของตัวแปรแฝงหนึ่งทำนายพัฒนาการของอีกตัวแปรแฝงหนึ่งมีระดับเพิ่มขึ้นหรือลดลงเมื่อระยะเวลา

ผ่านไป พร้อมทั้งการวิเคราะห์โมเดลโค้งพัฒนาการคู่ขนาน (parallel latent growth curve modeling) ที่สามารถวิเคราะห์ลักษณะพัฒนาการของปัจจัยที่มีลักษณะเป็นตัวแปรแฝงได้ ตลอดจนสามารถศึกษาอิทธิพลของปัจจัยภายในโมเดลได้หลายตัวพร้อมกัน ซึ่งลักษณะการวิเคราะห์ตัวแปรแฝงในการวิจัยนี้แตกต่างจากข้อมูลทุติยภูมิในลักษณะรายปีแบบอนุกรมเวลา (time series) ที่เหมาะกับการวิเคราะห์เศรษฐมิติ (econometric) หรือการวิเคราะห์อนุกรมเวลา (time series analysis) ดังนั้น คณะผู้วิจัยจึงสนใจศึกษาการวิจัยระยะยาวเพื่อวิเคราะห์พัฒนาการและวิเคราะห์อิทธิพลเชิงสาเหตุของปัจจัยที่กล่าวข้างต้น

ความมุ่งหมายของการวิจัย

1. เพื่อวิเคราะห์พัฒนาการของการลดความเสี่ยงในการใช้จ่ายภาครัฐ ความร่วมมือด้านการวิจัย และพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา
2. เพื่อวิเคราะห์อิทธิพลเชิงสาเหตุของการลดความเสี่ยงในการใช้จ่ายภาครัฐที่มีต่อความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ข้อมูลที่ใช้ในการวิเคราะห์ เป็นข้อมูลทุติยภูมิของดัชนีความสามารถทางการแข่งขันระดับโลก จากรายงานดัชนีความสามารถทางการแข่งขันระดับโลก (GCI) ในการประชุมผู้นำเศรษฐกิจของโลก (World Economic Forum: WEF) ของเขตเศรษฐกิจต่าง ๆ จำนวน 152 ประเทศ โดยจัดเก็บข้อมูล 4 ช่วงเวลา (time point) กล่าวคือ ปี ค.ศ. 2009 2012 2015 และ 2018 โดยเว้นปีที่ทำการศึกษา คณะผู้วิจัยคัดเลือกช่วงเวลาดังกล่าวเนื่องจากเป็นช่วงหลังวิกฤตเศรษฐกิจในประเทศสหรัฐอเมริกาที่ส่งผลกระทบต่อเขตเศรษฐกิจอื่น ๆ ทั่วโลก ทำให้เขตเศรษฐกิจทั่วโลกหันมาตระหนักในการสร้างความสามารถในการแข่งขันอย่างยั่งยืน รวมทั้งคณะผู้วิจัยได้คัดดัชนีของประเทศต่าง ๆ ออกจากการวิเคราะห์ครั้งนี้ ได้แก่ ประเทศแองโกลา บาร์เบโดส เบลีซ โบลิเวีย ภูฏาน บรูไน บูร์กินาฟาโซ เคปเวิร์ด คองโก กินี กาบอง เฮติ อิหร่าน ลาว เลบานอน ไลบีเรีย ลิเบีย มอลโดวา เมียนมาร์ รมันดา เซเชลส์ เซียร์ราลีโอน ซูรินาม สวาซิแลนด์ ซีเรีย และเยเมน เนื่องจากมีข้อมูลบางปีสูญหายหรือไม่มีการรายงานดัชนีในปีนั้น ๆ ทำให้ข้อมูลของประเทศนั้น ๆ ไม่เพียงพอต่อการวิเคราะห์ แม้ว่า การวิเคราะห์โมเดลโค้งพัฒนาการแนะนำให้ใช้จำนวนตัวอย่างอย่างน้อย 60 ถึง 200 ตัวอย่างในแต่ละช่วงเวลา และกำหนดช่วงเวลาอย่างน้อย จำนวน 3 ถึง 5 ช่วงเวลา (Felt et al., 2017; Shi et al., 2021) ที่ทำให้ข้อมูลแต่ละประเทศเพียงพอต่อการวิเคราะห์และลดความซับซ้อนในการประมาณค่าโมเดลลง แต่อย่างไรก็ตาม ในงานวิจัยนี้จำเป็นต้องศึกษาแนวโน้มปัจจัยที่ส่งผลกระทบต่อระบบการศึกษาของประเทศต่าง ๆ เมื่อเวลาเปลี่ยนผ่านไป ซึ่งจำนวนประเทศที่มีอยู่เป็นข้อจำกัดของการศึกษานี้ ดังนั้นในการวิจัยนี้ได้จำนวนข้อมูลทั้งสิ้นรวม 126 ประเทศ

2. เครื่องมือที่ใช้ในการวิจัย

ตัวแปรที่ใช้ในการวิจัยนี้ ได้แก่ ความสามารถทางการแข่งขันระดับโลก หมายถึง ชุดข้อมูลของสถาบัน นโยบาย และปัจจัยที่กำหนดระดับผลิตภาพของเขตเศรษฐกิจที่บ่งชี้ระดับความมั่งคั่งของเขตเศรษฐกิจนั้น ๆ ที่สามารถบรรลุผลสำเร็จได้ (Schwab, 2018) โดยทางการประชุมผู้นำเศรษฐกิจของโลกได้เก็บรวบรวมตัวบ่งชี้ความสามารถทางการแข่งขันทั้งหมดจำนวน 114 ตัว ซึ่งได้มีการพัฒนาตัวบ่งชี้มาตั้งแต่ปี ค.ศ. 2005 จำแนกเป็น 12 เสาหลัก (pillar) โดยแต่ละเสาหลัก ประกอบด้วย ปัจจัย 3 หลัก ที่มีลักษณะเป็นระบบกล่าวคือ ปัจจัยพื้นฐาน (basic requirements) คือ ปัจจัยป้อนเข้า ปัจจัยยกระดับประสิทธิภาพ (efficiency enhancers) คือ กระบวนการ และปัจจัยนวัตกรรมและและศักยภาพทางธุรกิจ (innovation and sophistication) คือ ผลลัพธ์ ซึ่งดัชนีนี้ได้มาจากข้อมูลขององค์กรระหว่างประเทศหลายแหล่ง เช่น กองทุนการเงินระหว่างประเทศ (International Monetary Fund: IMF) ธนาคารโลก (World Bank) องค์การเพื่อการศึกษา วิทยาศาสตร์และวัฒนธรรมแห่งสหประชาชาติ (UNESCO) องค์การอนามัยโลก (World Health Organization: WHO) โดยประมวลมาจากผลการดำเนินการของแต่ละประเทศที่มีฐานไม่เท่ากันมาจัดทำคะแนนในรูปดัชนีที่มีค่าตั้งแต่ 1 ถึง 7 คะแนน (1 = ไม่มีประสิทธิภาพอย่างยิ่ง และ 7 = มีประสิทธิภาพอย่างยิ่ง)

คณะผู้วิจัยบันทึกข้อมูลดัชนีด้วยแบบบันทึกดัชนีความสามารถทางการแข่งขันระดับโลก และคัดเลือกตัวแปรที่ใช้ในการวิเคราะห์จาก 3 เสาหลักที่คณะผู้วิจัยสนใจศึกษาดัชนีที่เกี่ยวข้องกับความสามารถด้านนวัตกรรมและคุณภาพการศึกษา ที่มีลักษณะเป็นระบบที่ประกอบด้วย ปัจจัยป้อนเข้า กระบวนการ และผลลัพธ์ โดยนิยามดัชนีตามแนวคิดของ Schwab (2018) ได้แก่ (1) เสาหลักด้านหน่วยงานรัฐ คือ ดัชนีการลดความสิ้นเปลืองในการใช้จ่ายภาครัฐ (wastefulness of government spending: WGS) หมายถึง ตัวบ่งชี้ประสิทธิภาพของการตัดสินใจใช้จ่ายภาครัฐในการพัฒนาความยั่งยืนของระบบเศรษฐกิจ (2) เสาหลักด้านการอุดมศึกษาและการฝึกอบรม คือ ดัชนีคุณภาพระบบการศึกษา (quality of the education system: QES) หมายถึง ตัวบ่งชี้ประสิทธิภาพในการตอบสนองของระบบการศึกษาต่อความต้องการจำเป็นในด้านความสามารถทางการแข่งขันของระบบเศรษฐกิจ (3) เสาหลักด้านนวัตกรรม คือ ดัชนีความร่วมมือในการวิจัยและพัฒนา ระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม (university-industry collaboration in R&D: UIC) หมายถึง ตัวบ่งชี้ประสิทธิภาพของการขยายความร่วมมือในการวิจัยและพัฒนา นวัตกรรมระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม

3. การเก็บรวบรวมข้อมูล

คณะผู้วิจัยเก็บรวบรวมข้อมูลดัชนีความสามารถทางการแข่งขันระดับโลกจากรายงานประจำปีของการประชุมผู้นำเศรษฐกิจของโลก ตั้งแต่ปี ค.ศ. 2009 ถึง 2018 (รายงานโดยการประชุมผู้นำเศรษฐกิจของโลก ในขณะที่ดำเนินการวิจัยนี้ยังไม่มีกรายงานดัชนีในปี ค.ศ. 2019 – 2022 เนื่องจากสถานการณ์ COVID-19) ที่จำแนกเป็นรายดัชนีของแต่ละประเทศ โดยข้อมูลดัชนีความสามารถทางการแข่งขันระดับโลกปี ค.ศ. 2018 นั้นทางการประชุมผู้นำเศรษฐกิจของโลกได้ปรับเปลี่ยนการรายงานดัชนีการลดความสิ้นเปลืองในการใช้จ่ายภาครัฐเป็นประสิทธิภาพในการใช้จ่ายภาครัฐ ดังนั้นคณะผู้วิจัยจึงต้องแปลงค่ากลับเป็นการลดความสิ้นเปลืองในการใช้จ่ายภาครัฐก่อนนำมาวิเคราะห์ข้อมูล

4. การวิเคราะห์ข้อมูล

1. วิเคราะห์ระดับการลดความสับสนในการใช้จ่ายภาครัฐ (WGS) ความร่วมมือในการวิจัยและพัฒนา ระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม (UIC) และคุณภาพระบบการศึกษา (QES) ด้วยสถิติบรรยายเพื่ออธิบายลักษณะของตัวแปรที่ใช้ในการวิจัยได้แก่ ค่าเฉลี่ย (M) ค่าส่วนเบี่ยงเบนมาตรฐาน (SD) คะแนนสูงสุด (MAX) คะแนนต่ำสุด (MIN) ค่าความเบ้ (SK) ค่าความโด่ง (KU) และค่าสัมประสิทธิ์การกระจาย ($C.V.$) ของตัวแปรสังเกตได้ ในการศึกษาลักษณะการกระจายและการแจกแจงของตัวแปรสังเกตได้แต่ละตัวแปร

2. วิเคราะห์พัฒนาการของการลดความสับสนในการใช้จ่ายภาครัฐ ความร่วมมือด้านการวิจัยและพัฒนา ระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา ด้วยการวิเคราะห์โมเดลโค้งพัฒนาการ (latent growth curve modeling: LGCM) โดยใช้โปรแกรม MPLus 8.4 ซึ่งมีผลการวิเคราะห์ที่สำคัญ ได้แก่ ค่าเฉลี่ยของค่าเริ่มต้น ค่าเฉลี่ยของอัตราการเปลี่ยนแปลง และดัชนีความสอดคล้องกลมกลืน (goodness of fit indices)

3. วิเคราะห์อิทธิพลเชิงสาเหตุของการลดความสับสนในการใช้จ่ายภาครัฐที่มีต่อความร่วมมือด้านการวิจัยและพัฒนา ระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา ด้วยการวิเคราะห์โมเดลโค้งพัฒนาการคู่ขนาน (parallel latent growth curve modeling) โดยใช้โปรแกรม MPLus 8.4 ซึ่งมีผลการวิเคราะห์ที่สำคัญ ได้แก่ ค่าอิทธิพล ค่าเฉลี่ยค่าเริ่มต้น ค่าเฉลี่ยของอัตราการเปลี่ยนแปลง และดัชนีความสอดคล้องกลมกลืน (goodness of fit indices) ทั้งค่าไค-สแควร์ (chi-square) และค่าไค-สแควร์สัมพัทธ์ (relative chi-square) ส่วนค่าดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (comparative fit index: CFI) และค่าดัชนีของทักเกอร์และลูอิส (Tucker-Lewis index: TLI) ควรมีค่าเข้าใกล้ 1.000 ในขณะที่ค่าดัชนีรากของกำลังสองเฉลี่ยของเศษเหลือมาตรฐาน (root mean square error of approximation: RMSEA) ควรให้มิต่ำน้อยกว่า .08 เนื่องจากเป็นระดับความคาดเคลื่อนแบบที่สองที่ยอมรับได้ในกรณีโมเดลมีความซับซ้อน (acceptable Type II error rates) (Hu & Bentler, 1999; Marsh et al., 2009) และควรพิจารณาควบคู่กับค่ามาตรฐานของดัชนีรากของกำลังสองเฉลี่ยของเศษ (standardized root mean square residual: SRMR) ที่ควรมีค่าต่ำกว่า .08 เนื่องจาก ค่า SRMR ไม่ได้กำหนดค่า penalty สำหรับโมเดลที่มีความซับซ้อน (Hu & Bentler, 1999; Piromsombat, 2019; Kenny, 2020) ซึ่งดัชนีทั้งสองนี้สามารถสะท้อนให้เห็นถึงโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ได้

ผลการวิจัย

การวิจัยครั้งนี้ใช้ข้อมูลจากรายงานดัชนีความสามารถทางการแข่งขันระดับโลก (GCI) ซึ่งคณะผู้วิจัยใช้ข้อมูล 3 ช่วงเวลา คือ ปี ค.ศ. 2009 2012 2015 และ 2018 ทั้ง 3 ตัวแปร คือ การลดความสับสนในการใช้จ่ายภาครัฐ (WGS) ความร่วมมือด้านการวิจัยและพัฒนา ระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม (UIC) และคุณภาพระบบการศึกษา (QES)

คณะผู้วิจัยจำแนกผลการวิจัยตามวัตถุประสงค์การวิจัยดังนี้ ในการตอบวัตถุประสงค์การวิจัยข้อที่ 1 นำเสนอผลการวิจัยเป็น 2 ตอน ได้แก่ ตอนที่ 1 คือ ผลการวิเคราะห์ค่าสถิติพื้นฐานของตัวแปรที่ใช้ในงานวิจัย

และตอนที่ 2 คือ ผลการวิเคราะห์พัฒนาการของการลดความสับสนเปลืองในการใช้จ่ายภาครัฐ ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัย และภาคอุตสาหกรรมและคุณภาพระบบการศึกษา ส่วนการตอบวัตถุประสงค์การวิจัยข้อที่ 2 นำเสนอเป็นตอนที่ 3 คือ ผลการวิเคราะห์อิทธิพลเชิงสาเหตุของการลดความสับสนเปลืองในการใช้จ่ายภาครัฐที่มีต่อพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัย และภาคอุตสาหกรรม และคุณภาพระบบการศึกษา โดยมีรายละเอียดดังนี้

ตอนที่ 1 ผลการวิเคราะห์ค่าสถิติพื้นฐานของตัวแปรที่ใช้ในงานวิจัย

ผลการวิเคราะห์ค่าสถิติพื้นฐานของตัวแปรที่ใช้ในงานวิจัยนี้ ได้แก่ 1) การลดความสับสนเปลืองในการใช้จ่ายภาครัฐมีคะแนนเฉลี่ยเท่ากับ 3.504 ($SD = 0.878$) ในปี ค.ศ. 2009 และมีแนวโน้มลดลงต่อเนื่องในปี ค.ศ. 2012 ($M = 3.319$, $SD = 0.905$) และปี ค.ศ. 2015 ($M = 3.244$, $SD = 0.930$) แต่ในปี ค.ศ. 2018 คะแนนเฉลี่ยของการลดความสับสนเปลืองในการใช้จ่ายภาครัฐมีค่าสูงขึ้นจากปี ค.ศ. 2015 ($M = 3.353$, $SD = 1.034$) 2) ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมมีคะแนนเฉลี่ยเท่ากับ 3.462 ($SD = 0.934$) ในปี ค.ศ. 2009 และมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่องในปี ค.ศ. 2012 ($M = 3.803$, $SD = 0.875$) และปี ค.ศ. 2015 ($M = 3.822$, $SD = 0.887$) แต่ในปี ค.ศ. 2018 คะแนนเฉลี่ยมีค่าลดลงน้อยกว่าปี ค.ศ. 2012 และปี ค.ศ. 2015 ($M = 3.689$, $SD = 0.867$) 3) คุณภาพระบบการศึกษามีคะแนนเฉลี่ยเท่ากับ 3.774 ($SD = 1.011$) ในปี ค.ศ. 2009 และมีคะแนนเฉลี่ยเพิ่มขึ้นในปี ค.ศ. 2012 ($M = 3.820$, $SD = 0.906$) แต่ในปี ค.ศ. 2015 กลับมีค่าเฉลี่ยลดลงจากปี ค.ศ. 2012 ($M = 3.796$, $SD = 0.914$) และมีค่าใกล้เคียงกันกับปี ค.ศ. 2018 ($M = 3.805$, $SD = 0.931$)

นอกจากนี้เมื่อพิจารณาความสัมพันธ์การกระจาย (C.V.) พบว่า ตัวแปรการลดความสับสนเปลืองในการใช้จ่ายภาครัฐมีค่าอยู่ประมาณร้อยละ 25 ถึง 31 และตัวแปรความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมมีค่าอยู่ประมาณร้อยละ 23 ถึง 26 ส่วนตัวแปรคุณภาพระบบการศึกษามีค่าอยู่ประมาณร้อยละ 23 ถึง 27 แสดงให้เห็นว่า การกระจายตัวของข้อมูลทั้งสามตัวแปรไม่ได้มีความแตกต่างกันมากนัก และไม่มีค่าผิดปกติ (outlier) ในข้อมูล ส่วนค่าความเบ้ (Sk) และความโด่ง (Ku) ของทุกตัวแปรพบว่า ทุกตัวแปรมีค่าความเบ้และความโด่งแตกต่างกัน แต่ไม่มีนัยสำคัญทางสถิติที่ระดับ .05 ดังตาราง 1

ตาราง 1 การวิเคราะห์ค่าสถิติพื้นฐานของตัวแปรที่ใช้ในการศึกษา ($N = 126$)

ตัวแปร/ปี ค.ศ.	MIN	MAX	M	SD	Sk	Ku	C.V. (%)
1. การลดความสับสนเปลืองในการใช้จ่ายภาครัฐ (WGS)							
1) 2009	1.461	6.055	3.504	.878	.319	-.076	25.057
2) 2012	1.658	6.144	3.319	.905	.652	.170	27.267
3) 2015	1.180	6.013	3.244	.930	.662	.472	28.668
4) 2018	1.140	6.243	3.353	1.034	.584	.126	30.838
2. ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม (UIC)							
1) 2009	1.885	5.847	3.462	.934	.760	-.206	26.978
2) 2012	1.969	5.777	3.803	.875	.445	-.453	23.008
3) 2015	2.031	5.968	3.822	.887	.499	-.389	23.207
4) 2018	2.265	5.773	3.689	.867	.715	-.367	23.502

ตาราง 1 (ต่อ)

ตัวแปร/ปี ค.ศ.	MIN	MAX	M	SD	Sk	Ku	C.V. (%)
3. คุณภาพระบบการศึกษา (QES)							
1) 2009	1.967	6.236	3.774	1.011	.514	-.488	26.779
2) 2012	1.991	5.944	3.820	.906	.309	-.520	23.717
3) 2015	2.053	6.132	3.796	.914	.393	-.445	24.078
4) 2018	1.853	6.189	3.805	.931	.375	-.502	24.467

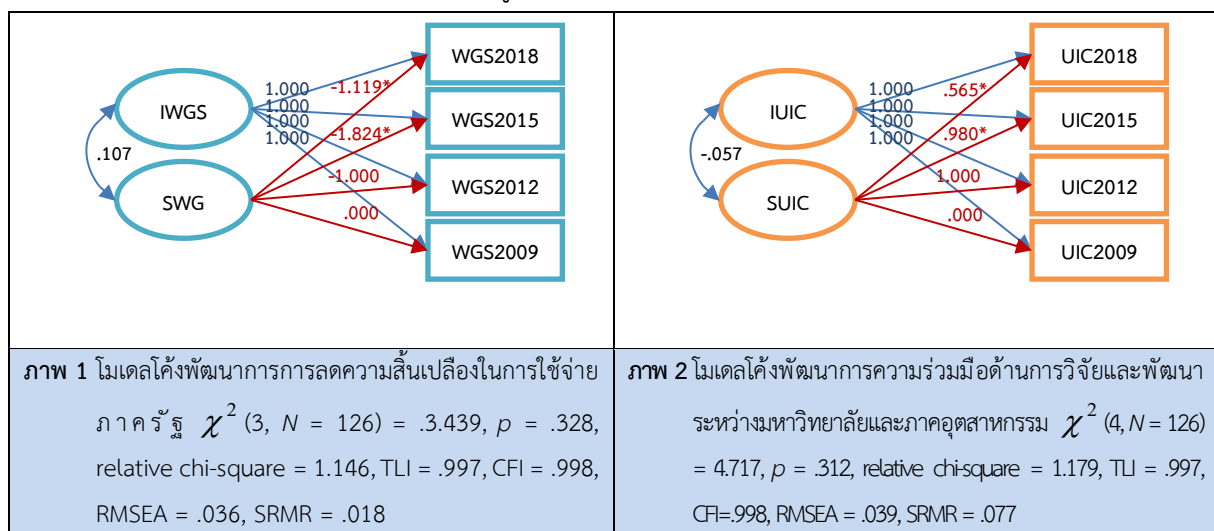
หมายเหตุ : $p < .05$; SE of Sk = .216, SE of Ku = .428

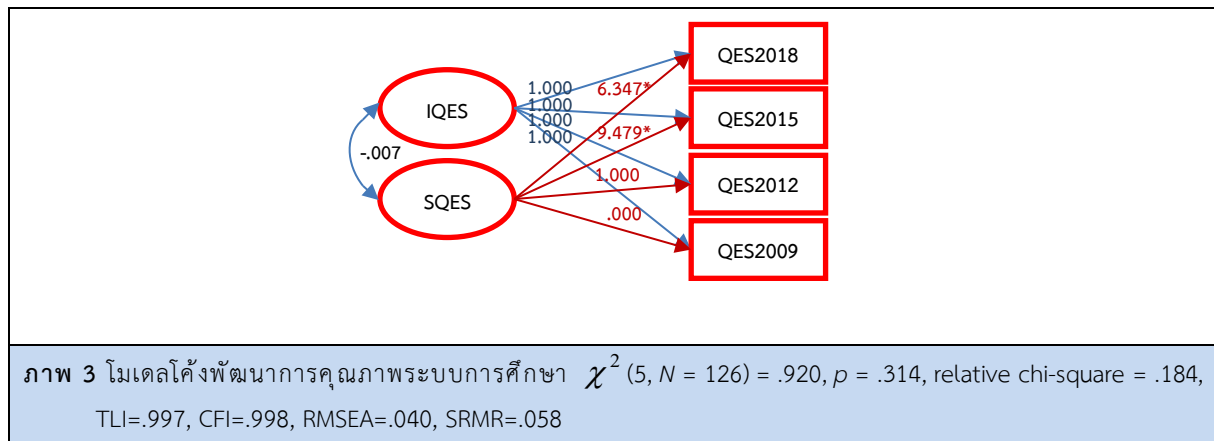
ตอนที่ 2 ผลการวิเคราะห์พัฒนาการของการลดความสับสนในการใช้จ่ายภาครัฐ ความร่วมมือด้านการวิจัยและพัฒนา ระหว่างมหาวิทยาลัย และภาคอุตสาหกรรมและคุณภาพระบบการศึกษา

ในการวิเคราะห์โมเดลโค้งพัฒนาการนี้ คณะผู้วิจัยไม่ได้กำหนดรูปแบบของโมเดลโค้งพัฒนาการ ได้ผลการวิเคราะห์ดังนี้ ผลการวิเคราะห์โมเดลโค้งพัฒนาการของการลดความสับสนในการใช้จ่ายภาครัฐ กับข้อมูลเชิงประจักษ์ พบว่า โมเดลมีค่าไค-สแควร์ (chi-square) มีค่าเท่ากับ $\chi^2(3, N = 126) = .3439$, $p = .328$, relative chi-square = 1.146 มีค่าดัชนี TLI เท่ากับ .997 ค่าดัชนี CFI เท่ากับ .998 ค่าดัชนี RMSEA เท่ากับ .036 และค่าดัชนี SRMR เท่ากับ .018 ซึ่งเป็นไปตามเกณฑ์ที่กำหนดจึงสามารถสรุปได้ว่า โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

ผลการวิเคราะห์โมเดลโค้งพัฒนาการความร่วมมือด้านการวิจัยและพัฒนา ระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมกับข้อมูลเชิงประจักษ์ พบว่า โมเดลมีค่าไค-สแควร์ (chi-square) มีค่าเท่ากับ $\chi^2(4, N = 126) = 4.717$, $p = .312$, relative chi-square = 1.179 มีค่าดัชนี TLI เท่ากับ .997 ค่าดัชนี CFI เท่ากับ .998 ค่าดัชนี RMSEA เท่ากับ .039 และค่าดัชนี SRMR เท่ากับ .077 ซึ่งเป็นไปตามเกณฑ์ที่กำหนดจึงสามารถสรุปได้ว่า โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

ผลการวิเคราะห์โมเดลโค้งพัฒนาการคุณภาพระบบการศึกษากับข้อมูลเชิงประจักษ์ พบว่า โมเดลมีค่าไค-สแควร์ (chi-square) มีค่าเท่ากับ $\chi^2(5, N = 126) = .920$, $p = .314$, relative chi-square = .184 มีค่าดัชนี TLI เท่ากับ .997 ค่าดัชนี CFI เท่ากับ .998 ค่าดัชนี RMSEA เท่ากับ .040 และค่าดัชนี SRMR เท่ากับ .058 แสดงว่า โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ดังภาพ 1 ถึงภาพ 3





ผลประมาณค่าพารามิเตอร์โมเดลโค้งพัฒนาการการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐ พบว่า ค่าเฉลี่ยพัฒนาการ (MS) มีค่าเท่ากับ .393 แสดงให้เห็นว่าประเทศมีพัฒนาการการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐโดยค่าเฉลี่ยพัฒนาการมีค่าเท่ากับ .393 และมีค่าเริ่มต้นเท่ากับ 3.940 ซึ่งสอดคล้องกับค่าเฉลี่ยที่มีแนวโน้มลดลงเมื่อเทียบกับปีฐาน (ค.ศ. 2009)

ผลประมาณค่าพารามิเตอร์โมเดลโค้งพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม พบว่า ค่าเฉลี่ยพัฒนาการ (MS) มีค่าเท่ากับ 1.126 แสดงให้เห็นว่าประเทศมีพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมโดยค่าเฉลี่ยพัฒนาการมีค่าเท่ากับ 1.126 และมีค่าเริ่มต้นเท่ากับ 4.055 ซึ่งสอดคล้องกับค่าเฉลี่ยที่มีแนวโน้มเพิ่มขึ้นเมื่อเทียบกับปีฐาน (ค.ศ. 2009)

ผลประมาณค่าพารามิเตอร์โมเดลโค้งพัฒนาการคุณภาพระบบการศึกษา พบว่า ค่าเฉลี่ยพัฒนาการ (MS) มีค่าเท่ากับ -.085 ไม่มีนัยสำคัญทางสถิติที่ระดับ .05 แสดงให้เห็นว่า ประเทศมีพัฒนาการคุณภาพระบบการศึกษาไม่มากนัก และมีค่าเริ่มต้นเท่ากับ 4.327 ซึ่งสอดคล้องกับค่าเฉลี่ยที่มีค่าใกล้เคียงกันทุกปี ดังตาราง 2

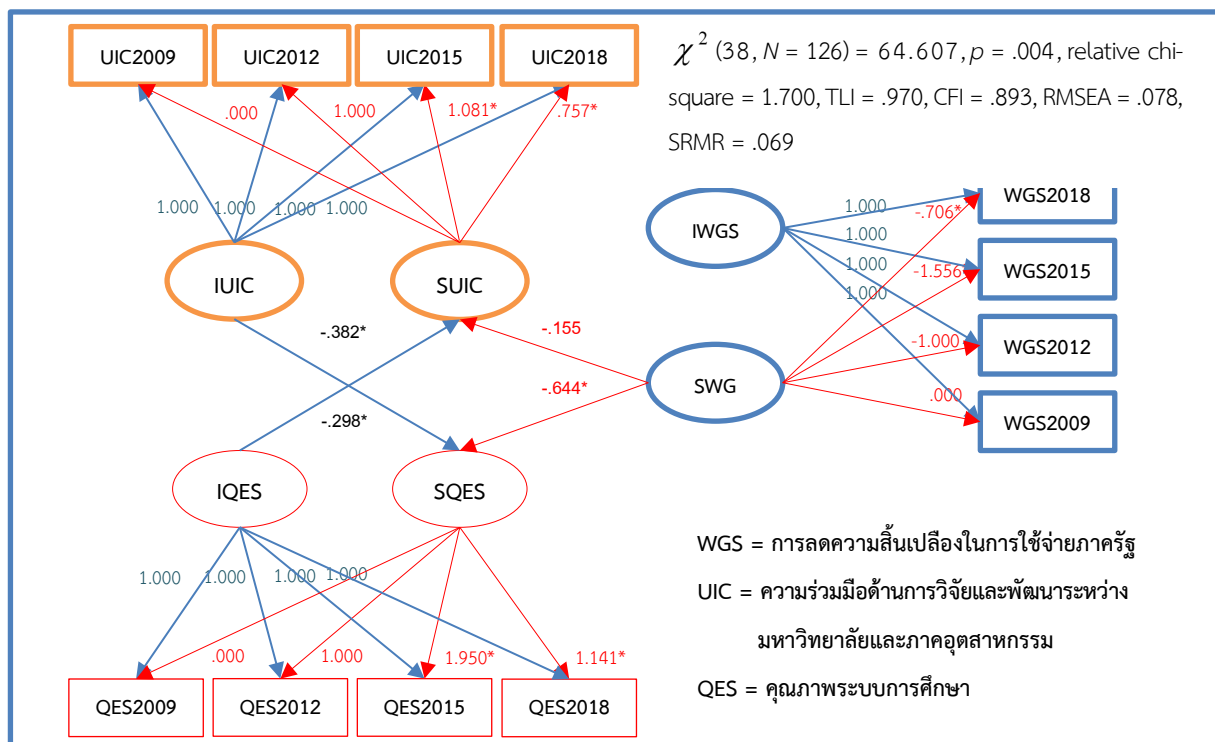
ตาราง 2 ผลการวิเคราะห์พัฒนาการของการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐ ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา

ค่าพารามิเตอร์	Est.	SE	t
การลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐ (WGS)			
Mean Intercept	3.940	.303	13.012*
Means slope	.393	.131	3.000*
ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม (UIC)			
Mean Intercept	4.055	.233	17.400*
Means slope	1.126	.229	4.913*
คุณภาพระบบการศึกษา (QES)			
Mean Intercept	4.327	.256	16.914*
Means slope	-.085	.103	-.822

หมายเหตุ : * = $p < .05$

ตอนที่ 3 ผลการวิเคราะห์อิทธิพลเชิงสาเหตุของการลดความสับสนในการใช้จ่ายภาครัฐที่มีต่อพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา

ผลการวิเคราะห์อิทธิพลเชิงสาเหตุของการลดความสับสนในการใช้จ่ายภาครัฐที่มีต่อพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษากับข้อมูลเชิงประจักษ์ คณะผู้วิจัยได้กำหนดค่าเริ่มต้นเท่ากับ 1.000 และไม่ได้กำหนดค่าพัฒนาการ (free parameter) คณะผู้วิจัยได้ปรับโมเดลโดยยอมให้ความคาดเคลื่อนของตัวแปรสังเกตได้มีความสัมพันธ์กันพบว่า โมเดลมีค่าไค-สแควร์ (chi-square) มีค่าเท่ากับ $\chi^2(38, N = 126) = 64.607, p = .004$, relative chi-square = 1.700 มีค่าดัชนี TLI เท่ากับ .970 ค่าดัชนี CFI เท่ากับ .983 นอกจากนี้ ค่าดัชนี RMSEA เท่ากับ .078 ซึ่งมีค่าน้อยกว่า .08 เป็นระดับความคาดเคลื่อนแบบที่สองที่ยอมรับได้ในกรณีโมเดลมีความซับซ้อน และเมื่อพิจารณาควบคู่กับค่าดัชนี SRMR เท่ากับ .069 ซึ่งมีค่าน้อยกว่า .08 ซึ่งสามารถสรุปได้ว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ดังภาพ 4



ภาพ 4 โมเดลอิทธิพลของการลดความสับสนในการใช้จ่ายภาครัฐที่มีต่อความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา

ผลการวิเคราะห์อิทธิพลของค่าเริ่มต้นของความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมที่มีต่อค่าพัฒนาการของคุณภาพระบบการศึกษา พบว่า ค่าเริ่มต้นของความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมมีอิทธิพลต่อค่าพัฒนาการของคุณภาพระบบการศึกษาเท่ากับ $-.382$ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($\beta = -.382, p < .001$) แสดงว่า ประเทศที่มีค่า

เริ่มต้นของความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมในระดับสูงจะมีพัฒนาการของคุณภาพระบบการศึกษาในระดับต่ำ

ผลการวิเคราะห์อิทธิพลของค่าเริ่มต้นของคุณภาพระบบการศึกษาที่มีต่อค่าพัฒนาการของความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม พบว่า ค่าเริ่มต้นของคุณภาพระบบการศึกษามีอิทธิพลต่อค่าพัฒนาการของความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมเท่ากับ $-.298$ อย่างมีนัยสำคัญทางสถิติที่ระดับ $.05$ ($\beta = -.298, p < .001$) แสดงว่า ประเทศที่มีค่าเริ่มต้นของคุณภาพระบบการศึกษาในระดับสูงมีพัฒนาการของความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมในระดับต่ำ

ผลการวิเคราะห์อิทธิพลของพัฒนาการการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐที่มีต่อค่าพัฒนาการของคุณภาพระบบการศึกษา พบว่า พัฒนาการการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐมีอิทธิพลต่อค่าพัฒนาการของคุณภาพระบบการศึกษาเท่ากับ $-.644$ อย่างมีนัยสำคัญทางสถิติที่ระดับ $.05$ ($\beta = -.644, p < .001$) แสดงว่า ประเทศที่มีค่าพัฒนาการของการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐในระดับสูงจะมีพัฒนาการของคุณภาพระบบการศึกษาต่ำ

ในขณะที่ผลการวิเคราะห์อิทธิพลของพัฒนาการการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐที่มีต่อค่าพัฒนาการของความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม พบว่า พัฒนาการของการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐมีอิทธิพลต่อค่าพัฒนาการของความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมเท่ากับ $-.155$ ($\beta = -.155, p = .267$) แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ $.05$ แสดงว่า ประเทศที่มีค่าพัฒนาการการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐไม่มีอิทธิพลต่อพัฒนาการของความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม ดังตาราง 3

ตาราง 3 ผลการวิเคราะห์อิทธิพลของพัฒนาการการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐที่มีต่อพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา

พารามิเตอร์	Intercept	SE	slope	SE
Loadings on การลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐ (WGS)				
EGS2009	.985*	.002	.000	
EGS2012	.954*	.038	-.345*	.048
EGS2015	.902*	.055	-.508*	.098
EGS2018	.839*	.040	-.214*	.085
Loadings on ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม (UIC)				
UIC2009	.981*	.014	.000	
UIC2012	.962*	.037	.392*	.057
UIC2015	1.005*	.036	.442*	.057
UIC2018	.969*	.030	.299*	.067

ตาราง 3 (ต่อ)

Loadings on คุณภาพระบบการศึกษา (QES)				
QES2009	.947*	.013	.000	
QES2012	.995*	.034	.235*	.055
QES2015	1.071*	.049	.494*	.072
QES2018	1.020*	.035	.275*	.067
โมเดลโครงสร้าง				
พารามิเตอร์	Est.	SE	t	p-value
IUIC -> SQES	-.382	.097	-3.936*	$p < .001$
IQES -> SUIC	-.298	.086	-3.458*	$p < .001$
SEGS -> SUIC	-.155	.14	-1.11	$p = .267$
SEGS -> SQES	-.644	.124	-5.207*	$p < .001$
$\chi^2 (38, N = 126) = 64.607, p = .004$, relative chi-square = 1.700, TLI = .970, CFI = .893, RMSEA = .078, SRMR = .069				

หมายเหตุ : * = $p < .05$

อภิปรายผล

วัตถุประสงค์ในการศึกษาระยะยาวในการวิจัยนี้ชี้ให้เห็นถึงพัฒนาการของปัจจัยต่าง ๆ ที่เกี่ยวข้องกับความสามารถในการสร้างสรรค์นวัตกรรมเมื่อเวลาผ่านไป และศึกษาอิทธิพลของการลดความสับสนเปลืองในการใช้จ่ายภาครัฐที่มีต่อการเปลี่ยนแปลงของระบบการศึกษา โดยใช้การวิเคราะห์โมเดลโค้งพัฒนาการและการวิเคราะห์โมเดลโค้งพัฒนาการคู่ขนานที่เหมาะสมต่อการวิเคราะห์ข้อมูลระยะยาว เนื่องจากผลการวิเคราะห์นี้ได้แบบจำลองที่แสดงให้เห็นถึงวิธีการเปลี่ยนแปลงของการใช้จ่ายภาครัฐและระบบการศึกษาในแต่ละประเทศตามช่วงเวลา พร้อมทั้งผลการวิจัยสามารถพิจารณาวิถีพัฒนาการของปัจจัยที่เกี่ยวข้องกับระบบการศึกษาที่เปลี่ยนแปลงไปมาและอิทธิพลที่เกิดขึ้นจากการลดความสับสนเปลืองในการใช้จ่ายภาครัฐต่อการเปลี่ยนแปลงของระบบการศึกษาในแต่ละประเทศได้ชัดเจนมากขึ้น คณะผู้วิจัยอภิปรายผลการวิจัยตามวัตถุประสงค์การวิจัยกล่าวคือ อภิปรายผลการวิเคราะห์โมเดลโค้งพัฒนาการ และอภิปรายผลการวิเคราะห์อิทธิพลเชิงสาเหตุโดยมีรายละเอียดดังนี้

1. ผลการวิเคราะห์โมเดลโค้งพัฒนาการของการลดความสับสนเปลืองในการใช้จ่ายภาครัฐ ความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา ทั้งสามโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ เมื่อพิจารณาผลประมาณค่าพารามิเตอร์โมเดลโค้งพัฒนาการจากรายงานดัชนีความสามารถทางการแข่งขันระดับโลก (GCI) พบว่า ค่าพารามิเตอร์ของโมเดลโค้งพัฒนาการที่ประมาณค่าได้ทั้งสามโมเดลสอดคล้องกับค่าเฉลี่ย โดยพัฒนาการของการลดความสับสนเปลืองในการใช้จ่ายภาครัฐมีแนวโน้มลดลงเมื่อเทียบกับปีฐาน ส่วนพัฒนาการคุณภาพระบบการศึกษาที่มีค่าใกล้เคียงกันทุกปี และพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมมีแนวโน้มเพิ่มขึ้นเมื่อเทียบกับปีฐาน ซึ่งบ่งชี้ความสามารถในการสร้างสรรค์นวัตกรรมในระยะยาวได้ แม้ภาครัฐจะลดงบประมาณรายจ่ายลงแต่ระบบการศึกษายังคงรักษาคุณภาพได้อยู่พร้อมกับเตรียมความพร้อมพัฒนา

องค์ความรู้และเทคโนโลยีร่วมกับภาคอุตสาหกรรม (Oturakci, 2021; Wang et al., 2021) ผู้มีส่วนเกี่ยวข้องกับการจัดทำนโยบายการศึกษาสามารถนำข้อมูลนี้มาใช้ในการเร่งรัดการสร้างสรรค์นวัตกรรมด้วยการเติมทรัพยากรที่เพิ่มคุณภาพการศึกษาและเอื้ออำนวยให้เกิดความร่วมมือระหว่างสองภาคส่วนมากขึ้น

2. ผลการวิเคราะห์อิทธิพลเชิงสาเหตุของการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐที่มีต่อพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษาพบว่า ประเทศที่มีคุณภาพระบบการศึกษาสูงอยู่แล้วมีพัฒนาการความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมเพียงเล็กน้อย ($\beta = -.298$) ในขณะที่ ประเทศที่มีความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมสูงอยู่แล้วก็มีพัฒนาการของคุณภาพระบบการศึกษาเพียงเล็กน้อยเช่นกัน ($\beta = -.382$) แสดงว่า ประเทศที่ระบบการศึกษามีคุณภาพนั้น มีพัฒนาการความร่วมมือด้านการวิจัยและพัฒนา โดยเห็นได้จากประเทศสหรัฐอเมริกา อังกฤษ ประเทศในกลุ่มยุโรปตะวันตก ซึ่งล้วนเป็นประเทศที่มีคุณภาพระบบการศึกษาและมีความร่วมมือด้านการวิจัยและพัฒนาระหว่างสถาบันทั้งสองอยู่แล้ว (Carrasco & Gunter, 2019) นอกเหนือจากความร่วมมือระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมจะช่วยลดช่องว่างของการนำความรู้สู่การปฏิบัติแล้ว ปัจจัยความร่วมมือกันระหว่างสถาบันยังส่งผลให้ประเทศมีความสามารถในการสร้างสรรค์นวัตกรรมด้วยการเพิ่มมูลค่าความรู้และสามารถสร้างองค์ความรู้ใหม่ที่มีความลึกซึ้งมากขึ้น (Sohn et al., 2016) ซึ่งการพัฒนาวัตกรรมการต้องอาศัยวิธีการขยายความร่วมมือกันในลักษณะเครือข่ายโดยรัฐที่เป็นตัวกลางประสานให้มหาวิทยาลัยและภาคอุตสาหกรรมก็ช่วยเติมเต็มช่องว่างดังกล่าวได้ (Cai et al., 2019)

ส่วนประเทศที่สามารถลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐได้มากขึ้นกลับทำให้พัฒนาการของคุณภาพระบบการศึกษาลดลง ($\beta = -.644$) โดยประเทศที่สามารถลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐได้มากขึ้นก็ไม่ได้ให้การสนับสนุนความร่วมมือด้านการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรมอย่างมีนัยสำคัญทางสถิติ ($\beta = -.155$) ผลการวิจัยนี้ชี้ให้เห็นว่า แม้ว่ารัฐบาลสามารถใช้จ่ายได้มีประสิทธิภาพมากกว่าเดิมก็ไม่ได้จำเป็นมาใช้จ่ายในภาคการศึกษาเท่าที่ควร สังเกตได้จากรัฐบาลมีนโยบายสนับสนุนให้สถาบันการศึกษาออกนอกกระบบงบประมาณเพื่อลดค่าใช้จ่ายของรัฐลง (privatization) มาตั้งแต่ช่วงทศวรรษที่ 1990 ยุครัฐบาลโรนัล เรแกนและมากาเร็ต แทตเชอร์ ต่างส่งผลกระทบต่อสถาบันอุดมศึกษารัฐต้องปรับตัวให้รองรับรายจ่ายที่เพิ่มมากขึ้น รวมทั้ง การเคลื่อนไหวเพื่อปฏิรูปการศึกษาทั่วโลก (GERM) จำเป็นต้องใช้งบประมาณในการพัฒนาคุณภาพระบบการศึกษามากขึ้น (Lubienski, 2016; Rizvi, 2016; Carrasco & Gunter, 2019)

สรุปผลการวิเคราะห์อิทธิพลเชิงสาเหตุของการวิจัยนี้ได้ว่า ประเทศที่มีงบประมาณรายจ่ายแตกต่างกันมีความสามารถในการสร้างสรรค์นวัตกรรมแตกต่างกันทั้งด้านการพัฒนาคุณภาพระบบการศึกษา และความร่วมมือด้านการวิจัยและพัฒนาของสองภาคส่วน ซึ่งผลการวิจัยนี้สอดคล้องกับงานวิจัยของ Oturakci (2021); Wang et al. (2021) นอกเหนือจากผลการวิจัยนี้ช่วยให้นักวิจัยเข้าใจลักษณะความสัมพันธ์ของการลดความสั่นเปลี่ยนในการใช้จ่ายภาครัฐ ความร่วมมือด้านการวิจัยและพัฒนาของมหาวิทยาลัยและภาคอุตสาหกรรม และคุณภาพระบบการศึกษา โดยอาศัยต้นทุนเดิม (ค่าเริ่มต้น) และพัฒนาการในแต่ละปี

(การเปลี่ยนแปลง) ของแต่ละประเทศในเรื่องที่เกี่ยวข้องกับภาพรวมระบบการศึกษาของประเทศต่าง ๆ ทั่วโลก การวิเคราะห์โมเดลคงพัฒนาการยังช่วยให้เห็นปัจจัยเชิงเห็นพัฒนาการที่แตกต่างกันและเห็นวิธีการพัฒนาของปัจจัยที่เกี่ยวกับภาพรวมระบบการศึกษาที่ได้เกิดการเปลี่ยนแปลงในช่วงทศวรรษที่ผ่านมา ซึ่งผู้บริหารการศึกษา และนักนโยบายการศึกษาสามารถนำข้อมูลมาใช้เป็นฐานการตัดสินใจพัฒนาระบบการศึกษาให้มีคุณภาพต่อไป

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1. นักวิจัยสามารถใช้ผลการวิจัยนี้ใช้เป็นตัวบ่งชี้ทิศทางการพัฒนาระบบการศึกษาที่เน้นการขับเคลื่อนเศรษฐกิจของประเทศไปพร้อมกัน และผู้มีส่วนเกี่ยวข้องกับการจัดทำนโยบายการศึกษาใช้เป็นตัวกำหนดนโยบายและแผนงานต่าง ๆ ในการปรับปรุงและพัฒนาระบบการศึกษาอย่างรอบด้าน

2. มหาวิทยาลัยควรเน้นกลยุทธ์การพัฒนาองค์ความรู้ของตนให้ตอบสนองกับการค้าและเชิงพาณิชย์มากขึ้น โดยการขยายความร่วมมือกับภาคอุตสาหกรรม ลงทุนในการเสริมสร้างขีดความสามารถและขยายโอกาสสำหรับการวิจัยและพัฒนาระหว่างมหาวิทยาลัยและภาคอุตสาหกรรม เพื่อส่งเสริมคุณภาพของการศึกษาในภาพรวมของมหาวิทยาลัยให้มากยิ่งขึ้น

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

ในงานวิจัยนี้เป็นการวิเคราะห์ข้อมูลในเชิงปริมาณที่ยังขาดการอธิบายถึงเหตุและปัจจัยที่เกิดจากเหตุการณ์ที่สำคัญที่เกิดขึ้นในอดีตที่อาจเป็นบ่อเกิดของอุปสรรคในการพัฒนาการของปัจจัยที่ส่งผลกระทบต่อระบบการศึกษาในภาพรวมดังกล่าว ดังนั้น ในการวิจัยต่อไปในอนาคตควรมีการศึกษาควบคู่กับเหตุการณ์ที่สำคัญตามสายกาลเวลา (timeline) เช่น งบประมาณทางการศึกษา การเปลี่ยนแปลงด้านนโยบายการศึกษา เพื่อสามารถชี้ชัดไปยังเหตุของการเปลี่ยนแปลงตามช่วงเวลาได้ครอบคลุมและมีความชัดเจนมากยิ่งขึ้น

References

- Aleksejeva, L. (2016). Country's competitiveness and sustainability: Higher education impact. *Journal of security and sustainability issues*, 5(3), 355-363. [https://doi.org/10.9770/jssi.2016.5.3\(4\)](https://doi.org/10.9770/jssi.2016.5.3(4))
- Barboza, G. E., Dominguez, S., & Pinder, J. (2017). Trajectories of post-traumatic stress and externalizing psychopathology among maltreated foster care youth: A parallel process latent growth curve model. *Child abuse & neglect*, 72, 370-382. <https://doi.org/10.1016/j.chiabu.2017.09.007>
- Cai, Y., Ramis Ferrer, B., & Luis Martinez Lastra, J. (2019). Building university-industry co-innovation networks in transnational innovation ecosystems: Towards a transdisciplinary approach of integrating social sciences and artificial intelligence. *Sustainability*, 11(17), 4633. <https://doi.org/10.3390/su11174633>

- Carnoy, M., Loyalka, P., Dobryakova, M., Dossani, R., Isak, D., Kuhns, K., & Wang, R. (2013). BRIC universities as institutions in the process of change. In M., Carnoy, P., Loyalka, M., Dobryakova, R., Dossani, D., Isak, K., Kuhns, & R., Wang (Eds.), *University Expansion in a Changing Global Economy: Triumph of the BRICs?* (pp. 140-178). Stanford University.
- Carrasco, A., & Gunter, H. M. (2019). The “private” in the privatisation of schools: The case of Chile, *Educational Review*, 71(1), 67-80. <https://doi.org/10.1080/00131911.2019.1522035>
- Cheong, J., MacKinnon, D. P., & Khoo, S. T. (2003). Investigation of mediational processes using parallel process latent growth curve modeling. *Structural Equation Modeling*, 10(2), 238-262. https://doi.org/10.1207/S15328007SEM1002_5
- Djefflat, A. (2015). *Morocco's efforts on the knowledge economy. Economic Intelligence Unit, Morocco Country Profile 2012*. cmimarseille.org. <https://www.cmimarseille.org/knowledge-library/morocco-efforts-knowledge-economy>
- Felt, J. M., Depaoli, S., & Tiemensma, J. (2017). Latent growth curve models for biomarkers of the stress response. *Frontiers in Neuroscience*, 11, 315. <https://doi.org/10.3389/fnins.2017.00315>
- Hayes, M. (2015). Headwinds: Growth, democracy, and the middle class in Latin America. *Georgetown Journal of International Affairs*, 16(1), 94-105.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Kenny, D. A. (2020). *Measuring Model Fit*. SEM Fit. <http://www.davidakenny.net/>
- Kim, Y., Lee, J. M., Kim, J., Dhurandhar, E., Soliman, G., Wehbi, N. K., & Canedy, J. (2017). Longitudinal associations between body mass index, physical activity, and healthy dietary behaviors in adults: A parallel latent growth curve modeling approach. *PLOS ONE*, 12(3), e0173986. <https://doi.org/10.1371/journal.pone.0173986>
- Kokins, E., & Lavrinovics, V. (2006). *Latvia: Catching-up towards the world production frontier, an industry level analysis* [Bachelor's thesis], Stockholm School of Economics.
- Lubienski, C. (2016). Sector distinctions and the privatization of public education policymaking. *Theory and Research in Education*, 14(2), 193-212. <https://doi.org/10.1177/1477878516635332>
- Marsh, H. W., Muthen, B., Asparouhov, T., Ludtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, Integrating CFA and EFA: Application to students' evaluation of university teaching. *Structural Equation Modeling*, 16(3), 439-476. <https://doi.org/10.1080/10705510903008220>
- Milovanovic, S. (2015). Balancing differences and similarities within the global economy: Towards a collaborative business strategy. *Procedia Economics and Finance*, 23, 185-190. [https://doi.org/10.1016/S2212-5671\(15\)00369-X](https://doi.org/10.1016/S2212-5671(15)00369-X)

- Mitchell, C. M., Kaufman, C. E., Beals, J., & Pathways of Choice and Healthy Ways Project Team. (2005). Resistive efficacy and multiple sexual partners among American Indian young adults: A parallel-process latent growth curve model. *Applied Developmental Science*, 9(3), 160-171. https://doi.org/10.1207/s1532480xads0903_3
- Olsson, A., & Meek, L. (2014). *Effectiveness of research and innovation management at policy and institutional levels: Cambodia, Malaysia, Thailand and Vietnam*. OECD.
- Oturakci, M. (2021). Comprehensive analysis of the global innovation index statistical and strategic approach. *Technology Analysis & Strategic Management*. <https://doi.org/10.1080/09537325.2021.1980209>
- Piehler, T. F., Bloomquist, M. L., August, G. J., Gewirtz, A. H., Lee, S. S., & Lee, W. S. (2014). Executive functioning as a mediator of conduct problems prevention in children of homeless families residing in temporary supportive housing: A parallel process latent growth modeling approach. *Journal of abnormal child psychology*, 42(5), 681-692. <https://doi.org/10.1007/s10802-013-9816-y>
- Rizvi, F. (2016). *Privatization in education: Trends and consequences*. Education, Research and Foresight: Working Papers, Document code ED-2016/WP/2. <https://unesdoc.unesco.org/ark:/48223/pf0000246485>
- Schwab, K. (Eds.). (2018). *Insight Report: The Global Competitiveness Report 2017–2018*. World Economic Forum. www.weforum.org/gcr
- Shi, D., DiStefano, C., Zheng, X., Liu, R., & Jiang, Z. (2021). Fitting latent growth models with small sample sizes and non-normal missing data. *International Journal of Behavioral Development*, 45(2), 179-192. <https://doi.org/10.1177/0165025420979365>
- Sohn, S. Y., Kim, D. H., & Jeon, S. Y. (2016). Re-evaluation of global innovation index based on a structural equation model. *Technology Analysis & Strategic Management*, 28(4), 492-505. <https://doi.org/10.1080/09537325.2015.1104412>
- Wang, X., Wang, Z., & Jiang, Z. (2021). Configurational differences of national innovation capability: A fuzzy set qualitative comparative analysis approach. *Technology Analysis & Strategic Management*, 33(6), 599-611. <https://doi.org/10.1080/09537325.2020.1832211>

Translate Thai References

- National Science Technology and Innovation Policy Office. (2017). *Review of Global Competitiveness Index for Thailand 2017 - 2018 by World Economic Forum: WEF*. Information News, Science, Technology, and Innovation Information Center. http://www.sti.or.th/sti/news-detail.php?news_type=2&news_id=250& (in Thai)
- Piromsombat, C. (2019). *Handout of Longitudinal Research in Growth and Development (2756716)*. [Unpublished manuscript], Faculty of Education Chulalongkorn University. (in Thai)

การเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์ รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

A Comparison of the Item Information Function of the Mixed-Format Test for Mathematical Literacy with Different Test Lengths and Proportions of Items

นุชรากร นันตะสุยะ^{1*} จุฑาภรณ์ มาสันเทียะ² และ พนิดา พานิชวัฒนะ³

Nucharakorn Nantasuya^{1*}, Chuthaphon Masantiah² and Panida Panidvadana³

(Received: March 19, 2022; Revised: April 5, 2022; Accepted: April 11, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ 1) สร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน และ 2) เปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์ที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน กลุ่มตัวอย่างที่ใช้ในการวิจัย ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 900 คน เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน ได้แก่ (1) โมเดล 1 มีความยาว 30 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% (2) โมเดล 2 มีความยาว 30 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% (3) โมเดล 3 มีความยาว 50 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% และ (4) โมเดล 4 มีความยาว 50 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน การวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA) ผลการวิจัยพบว่า

1. ในการสร้างแบบทดสอบรูปแบบผสมได้แบบทดสอบที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน 4 โมเดล โดยโมเดล 1 และ โมเดล 2 ซึ่งมีความยาว 30 ข้อ มีความเที่ยง 0.842 และ 0.803 ตามลำดับ ส่วนโมเดล 3 และ โมเดล 4 ซึ่งมีความยาว 50 ข้อ มีความเที่ยง 0.876 และ 0.872 ตามลำดับ โดยทั้ง 4 โมเดลที่สร้างขึ้นมีข้อสอบรวมจำนวน 11 ข้อ

¹ นักศึกษาปริญญาโท สาขาวิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

² ผู้ช่วยศาสตราจารย์ สาขาวิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

³ อาจารย์ สาขาวิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

¹ Master's degree student, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

² Assistant Professor, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

³ Lecturer, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

* Corresponding Author E-mail: nucharakorn14@gmail.com

2. ผลการเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดลที่แตกต่างกัน พบว่า มีค่าเฉลี่ยสารสนเทศของข้อสอบแตกต่างกันอย่างน้อย 1 คู่ เมื่อเปรียบเทียบค่าเฉลี่ยรายคู่ พบว่า มีค่าเฉลี่ยสารสนเทศของข้อสอบแตกต่างกัน 5 คู่ โดยโมเดล 1 มีค่าสารสนเทศมากกว่าโมเดล 2, 3 และ 4 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และโมเดล 3 มีค่าเฉลี่ยสารสนเทศของข้อสอบสูงกว่าโมเดล 2 และ 4 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

คำสำคัญ: ค่าสารสนเทศของข้อสอบ ทฤษฎีการตอบสนองข้อสอบ แบบทดสอบรูปแบบผสม

Abstract

The objectives of this research were to 1) construct mixed format of mathematical literacy tests with different test lengths and proportions of items, and 2) compare the item information function of the mixed-format tests for mathematical literacy with different test lengths and proportions of items. The sample consisted of 900 grade 9 students. The research instruments were mixed-format tests for mathematical literacy with different test lengths and proportions of items. These tests consisted of the following models: (1) Model 1, with the length of thirty items and the proportion of multiple-choice items and subjective items of 30:70 percent; (2) Model 2, with the length of thirty items and the proportion of multiple-choice items and subjective items of 50:50 percent; (3) Model 3, with the length of fifty items and the proportion of multiple-choice items and subjective items of 30:70 percent; and (4) Model 4, with the length of fifty items and the proportion of multiple-choice items and subjective items of 50:50 percent. The statistics used in the data analysis were frequency, percentage, mean, standard deviation and one-way ANOVA. The findings were as follows:

1. In the construction of the mixed-format tests, 4 different test lengths and proportions of items were obtained. Model 1 and Model 2, each with the length of 30 items, had the reliability of 0.842 and 0.803 respectively. Model 3 and Model 4, each with the length of 50 items, had the reliability of 0.876 and 0.872 respectively. All 4 models of the tests that were created consisted of a combined exam of 11 items.

2. The result of the comparison of the item information function between different models showed that there was at least one pair of different means of item information function. When each individual mean pair was considered, it was found that there were five pairs of different means of item information function. Model 1 had the item information function higher than model 2, 3, and 4, with statistical significance at the .05 level. And Model

3 had the mean of item information function higher than models 2 and 4, with statistical significance at the .05 level.

Keywords: item information function, Item Response Theory, mixed-format test

บทนำ

คณิตศาสตร์มีบทบาทสำคัญต่อความสำเร็จในการเรียนรู้ในศตวรรษที่ 21 เนื่องจากคณิตศาสตร์เป็นเครื่องมือในการศึกษาด้านวิทยาศาสตร์ เทคโนโลยี และศาสตร์อื่น ๆ อันเป็นรากฐานในการพัฒนาบุคลากรของชาติให้มีคุณภาพ ซึ่งถือว่าเป็นสิ่งจำเป็นสำหรับการพัฒนาและการแข่งขันทางเศรษฐกิจของประเทศ เห็นได้จากการประเมินการรู้เรื่องคณิตศาสตร์ (Mathematical Literacy) ซึ่งเป็นจุดเน้นของโครงการประเมินผลนักเรียนร่วมกับนานาชาติ (Programme for International Student Assessment: PISA) ดำเนินการโดยองค์การเพื่อความร่วมมือทางเศรษฐกิจและการพัฒนา (Organisation for Economic Co-operation and Development: OECD) เพื่อสะท้อนว่าระบบการศึกษาของประเทศได้เตรียมเยาวชนของชาติให้มีความพร้อมสำหรับการใช้ชีวิตและการทำงานในอนาคต จากรายงานผลการประเมินโครงการ PISA ในประเทศไทย ตั้งแต่ PISA 2000 ถึงปัจจุบัน พบว่านักเรียนที่จบการศึกษาภาคบังคับ (นักเรียนอายุ 15 ปี) มีคะแนนเฉลี่ยการรู้เรื่องคณิตศาสตร์ต่ำกว่าคะแนนเฉลี่ยของทุกประเทศรวมกันและต่ำกว่าระดับพื้นฐาน (Institute for the Promotion of Teaching Science and Technology, 2018) ผู้วิจัยจึงสนใจศึกษา รูปแบบข้อสอบของ PISA ซึ่งประกอบด้วยข้อสอบแบบปรนัยและอัตนัยรวมอยู่ในฉบับเดียวกัน

ปัจจุบันการประเมินผู้เรียนมีความพยายามในการวัดทักษะความสามารถการคิดวิเคราะห์ คิดสังเคราะห์ คิดสร้างสรรค์ ของผู้เรียนมากกว่าการวัดความรู้ ความจำ และความเข้าใจในเนื้อหาเท่านั้น แต่การเลือกใช้ข้อสอบรูปแบบเดียวมีข้อจำกัดในการวัด มีจุดแข็งและจุดอ่อน (Kinsey, 2003; Kim et al., 2007; Reshetar & Melican, 2010) การทดสอบจำเป็นต้องใช้แบบทดสอบทั้ง 2 รูปแบบควบคู่กันไป เรียกว่า “แบบทดสอบรูปแบบผสม” (Mixed-format Tests) ประกอบด้วยข้อสอบแบบปรนัยและข้อสอบอัตนัย ซึ่งแบบทดสอบรูปแบบผสมสามารถลดข้อจำกัดของข้อสอบปรนัยและข้อสอบอัตนัย โดยสามารถวัดได้ครอบคลุมเนื้อหาที่ต้องการวัดมากขึ้นและวัดระดับความสามารถของผู้สอบในขั้นสูงได้ นอกจากนี้ยังเปิดโอกาสให้ผู้สอบแสดงความรู้ความสามารถของตนเองและช่วยลดโอกาสในการเดาข้อสอบปรนัยได้ (Kinsey, 2003; Reshetar & Melican, 2010; Srikolchan, 2009)

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบทดสอบรูปแบบผสมในประเด็นเกี่ยวกับความยาว พบว่า มีการสร้างแบบทดสอบที่มีความยาวแตกต่างกันตั้งแต่ 10 ข้อ ถึง 80 ข้อ (Nansombat, 1999; Kinsey, 2003; Shin, 2007; Jurich & Goodman, 2009; Saen-amnuaiphon, 2012) โดยผลการวิจัย พบว่า แบบทดสอบรูปแบบผสมไม่เหมาะสมกับแบบทดสอบที่มีความยาวน้อยกว่า 30 ข้อ เนื่องจากจะทำให้ความเที่ยงและความตรงของแบบทดสอบมีค่าต่ำลง ความยาวที่เหมาะสมจะอยู่ระหว่าง 30-50 ข้อ ซึ่งความเที่ยงและความตรงของแบบทดสอบจะมีค่าใกล้เคียงกัน

เมื่อศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบทดสอบรูปแบบผสมในประเด็นสัดส่วนของข้อสอบพบว่า มีการสร้างแบบทดสอบรูปแบบผสมที่มีสัดส่วนระหว่างข้อสอบปรนัยและอัตนัยที่แตกต่างกัน เช่น (1) 90:10 (2) 80:20 (3) 50:50 (4) 30:70 และ (5) 20:80 (Bastari, 2000; Shin, 2007; Saen-amnuaiphon, 2012) โดยผลการวิเคราะห์ พบว่า แบบทดสอบที่มีสัดส่วนของข้อสอบอัตนัยมากกว่าจะให้ค่าสารสนเทศของข้อสอบสูงกว่าและแบบทดสอบที่มีสัดส่วนข้อสอบปรนัยน้อยกว่าอัตนัย จะให้ประสิทธิภาพของแบบทดสอบคือ ความเที่ยง และความตรงสูงสุด

จากการศึกษาเอกสารและงานวิจัยในอดีตที่มีการตรวจสอบคุณภาพของแบบทดสอบรูปแบบผสม มีการศึกษาคุณสมบัติทางจิตมิติโดยใช้ทฤษฎีการตอบสนองข้อสอบ เช่น ค่าอำนาจจำแนก (a) ค่าความยาก (b) ค่าโอกาสในการเดา (c) ค่าความเที่ยง (reliability) ความตรง (validity) ค่าสารสนเทศของข้อสอบ (Item Information Function: IIF) ค่าสารสนเทศของแบบสอบ (Test Information Function: TIF) ความไม่ลำเอียงของข้อคำถาม และอิทธิพลของวิธีวัด (Chaisuwan, 2012; Sawangsri, 2015) การหาคุณภาพเครื่องมือต้องมีการตรวจสอบความเที่ยงเสมอ ซึ่งความเที่ยงเป็นคุณสมบัติที่มีความสำคัญ ถ้าแบบวัดได้มีค่าความเที่ยงสูง แสดงว่าแบบวัดนั้นมีคุณภาพสูงด้วย งานวิจัยนี้จึงมุ่งสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาว 30 ข้อ และ 50 ข้อ รวมถึงมีสัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 2 แบบ คือ 30% : 70% และ 50% : 50% โดยพิจารณาค่าความเที่ยงรวมถึงค่าสารสนเทศของข้อสอบโดยเปรียบเทียบว่าความยาวและสัดส่วนข้อสอบแบบใดที่จะให้ค่าสารสนเทศของข้อสอบสูงสุด โดยนำองค์ประกอบของการประเมินผลนักเรียนระดับนานาชาติ PISA มาเป็นองค์ประกอบหนึ่งในการสร้างแบบทดสอบในครั้งนี้ด้วย ผลที่ได้จากการวิจัยจะเป็นสารสนเทศสำคัญในการเลือกใช้แบบทดสอบรูปแบบผสมโดยพิจารณาความยาวของแบบทดสอบและสัดส่วนของข้อสอบปรนัยและอัตนัย เพื่อให้ได้แบบทดสอบรูปแบบผสมที่มีความถูกต้องแม่นยำในการประมาณค่า และสามารถประมาณค่าได้ใกล้เคียงกับระดับความสามารถที่แท้จริงของผู้สอบให้มากที่สุด อีกทั้งเป็นแนวทางแก่หน่วยงานที่เกี่ยวข้องใช้สร้างแบบวัดผลสัมฤทธิ์ทางการเรียน แบบทดสอบการรู้เรื่องคณิตศาสตร์ของผู้เรียนที่มีรูปแบบผสมที่มีคุณสมบัติทางจิตมิติที่เหมาะสมต่อไป

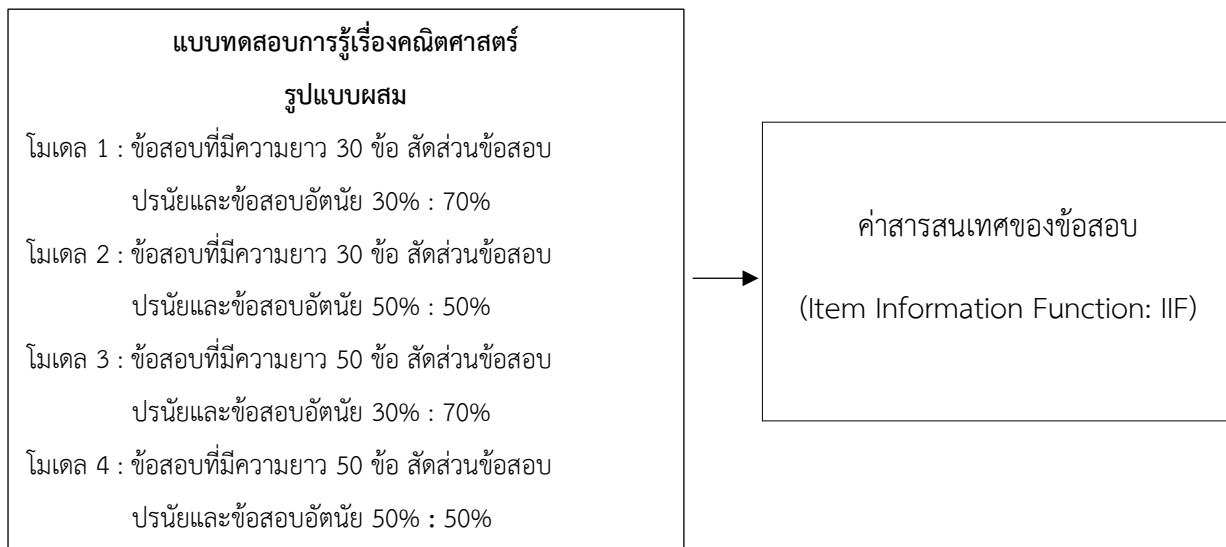
ความมุ่งหมายของการวิจัย

1. เพื่อสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน
2. เพื่อเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มี

ความยาวและสัดส่วนแตกต่างกัน

กรอบแนวคิดในการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องดังกล่าว สามารถกำหนดเป็นกรอบแนวคิดการวิจัย การเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน ได้ดังภาพ 1



ภาพ 1 กรอบแนวคิดในการวิจัย

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2563 ที่กำลังศึกษาอยู่ในจังหวัดนครนายก สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 7 ในจังหวัดนครนายก จำนวน 11 โรงเรียน จำนวน 1,849 คน

กลุ่มตัวอย่างในงานวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 7 ในจังหวัดนครนายก ปีการศึกษา 2563 ได้กลุ่มตัวอย่างจำนวน 1,020 คน เป็นกลุ่มทดลองใช้ จำนวน 120 คน และกลุ่มใช้จริง จำนวน 900 คน แบ่งเป็น 4 กลุ่มย่อย ๆ ละ 225 คน จำแนกตามขนาดโรงเรียน พบว่า เป็นโรงเรียนขนาดใหญ่พิเศษคิดเป็นร้อยละ 62.22 ขนาดใหญ่ 17.78 ขนาดกลาง 17.78 ขนาดเล็ก 2.22 และเมื่อจำแนกตามเพศพบว่าส่วนใหญ่เป็นเพศหญิงคิดเป็นร้อยละ 64.78 ซึ่งจำนวนกลุ่มตัวอย่างที่เหมาะสมสำหรับการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบและแบบทดสอบแบบ 1 พารามิเตอร์ควรมีจำนวน 200 คน (Kinsey, 2003; Srikolchan, 2009; Saen-amnuaiophon, 2012) เนื่องจากงานวิจัยนี้มีค่าพารามิเตอร์ของข้อสอบและแบบสอบแบบ 1 พารามิเตอร์ จำนวน 4 โมเดล จึงได้กลุ่มตัวอย่าง 800 คน และเพื่อป้องกันการสูญหายจึงปรับเพิ่มกลุ่มตัวอย่างเป็นโมเดลละ 225 คน ได้กลุ่มตัวอย่างทั้งหมด 900 คน ตัวอย่างที่ได้มาจากการสุ่มแบบหลายขั้นตอน (Multi-stage random sampling)

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน ซึ่งประกอบด้วยข้อสอบ 2 รูปแบบ คือ ข้อสอบปรนัย เป็นข้อสอบแบบเลือกตอบ 4 ตัวเลือก และข้อสอบอัตนัย เป็นข้อสอบแบบเขียนตอบ แบบตอบสั้น จำนวน 80 ข้อ คัดเลือกเหลือ 67 ข้อ ซึ่งข้อสอบปรนัยและอัตนัยจะเป็นแบบทดสอบที่ตรวจให้คะแนนสองค่า (ตอบถูกต้อง 1 คะแนน และตอบผิดหรือไม่ตอบได้ 0 คะแนน) รวมอยู่ในฉบับเดียวกัน จำนวน 4 โมเดล ได้แก่ (1) โมเดล 1 ข้อสอบที่มีความยาว

30 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% (2) โมเดล 2 ข้อสอบที่มีความยาว 30 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% (3) โมเดล 3 ข้อสอบที่มีความยาว 50 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% และ (4) โมเดล 4 ข้อสอบที่มีความยาว 50 ข้อ สัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% โดยแบบทดสอบนี้สร้างขึ้นจากองค์ประกอบของการรู้เรื่องคณิตศาสตร์ ประกอบด้วย (1) สถานการณ์หรือบริบทของปัญหา ได้แก่ บริบทส่วนตัว บริบทสังคม บริบทการงานอาชีพ และบริบทวิทยาศาสตร์ (2) เนื้อหาคณิตศาสตร์ ได้แก่ เนื้อหาปริภูมิและรูปทรงสามมิติ เนื้อหาการเปลี่ยนแปลงและความสัมพันธ์ เนื้อหาปริมาณ และ เนื้อหาความไม่แน่นอน และ (3) กลุ่มสมรรถนะทางคณิตศาสตร์ ได้แก่ สมรรถนะการทำให้ใหม่ สมรรถนะการต่อเชื่อม และ สมรรถนะการสะท้อนและการสื่อสาร ในที่นี้ผู้วิจัยได้นำเสนอข้อสอบรวมที่มีค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) ที่เหมาะสม รายละเอียดตัวอย่างข้อสอบดังตาราง 1

ตาราง 1 ตัวอย่างข้อสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

ข้อสอบปรนัย (ข้อสอบรวม)	ข้อสอบอัตนัย (ข้อสอบรวม)
บริบท : สังคม	บริบท : การงานอาชีพ
เนื้อหา : การเปลี่ยนแปลงและความสัมพันธ์	เนื้อหา : ปริมาณ
สมรรถนะ : การสะท้อนและการสื่อสาร	สมรรถนะ : การทำให้ใหม่
$p = .65$	$p = .53$
$r = .30$	$r = .72$
การเดินทาง การเดินทางจากนครนายก ไปกรุงเทพฯ สามารถเลือกเดินทางได้ 2 วิธี ดังนี้ วิธีที่ 1 เดินทางโดยรถประจำทาง รถมามาทุกๆ 10 นาที และใช้เวลาในการเดินทางไปถึงกรุงเทพฯ ๖ เป็นเวลา 80 นาที วิธีที่ 2 เดินทางโดยรถตู้ รถมามาทุกๆ 15 นาที และใช้เวลาในการเดินทางไปถึงกรุงเทพฯ ๖ เป็นเวลา 75 นาที ถ้าต้องการเดินทางจากนครนายก ไปยังกรุงเทพฯ ต้องเลือกโดยสารรถแบบเดียวเท่านั้น และรถจะออกเที่ยวแรกพร้อมกันในเวลา 08.00 น 1. พิมพ์ต้องเดินทางจากนครนายกเพื่อไปเรียนในกรุงเทพฯ ๖ เธอมาถึงป้ายรถประจำทาง เวลา 09.10 น. เธอจะเลือกขึ้นรถประจำทาง หรือ รถตู้ จึงจะใช้เวลาในการเดินทางน้อยที่สุด และเมื่อรวมระยะเวลาในการรอรถคิดเป็นระยะเวลาเท่าใด ก. รถประจำทาง ใช้เวลา 85 นาที ข. รถประจำทาง ใช้เวลา 80 นาที ค. รถตู้ ใช้เวลา 75 นาที ง. รถประจำทาง หรือ รถตู้ ก็ได้ ใช้เวลา 80 นาที เท่ากัน	ร้านเบเกอรี่ ทนายต้องการจะเปิดร้านเบเกอรี่ โดยภายในร้านมีผลิตภัณฑ์ต่าง ๆ เช่น ขนมเค้ก ขนมปัง คุกกี้และโดนัต ทนายต้องตัดสินใจเข้าสถานที่เพื่อเปิดร้านเบเกอรี่ โดยมีข้อมูลสถานที่ทั้ง 3 แห่ง ดังนี้ ● ร้าน ก คิดค่าเช่าเดือนละ 5,000 บาท จ่ายค่าเช่าล่วงหน้า 3 เดือน ค่าประกันร้านค้า 1,500 บาท ● ร้าน ข คิดค่าเช่าปีละ 50,000 บาท จ่ายค่าเช่าก่อนครบปีหนึ่งเดือน ค่าประกันร้านค้า 10,000 บาท ● ร้าน ค คิดค่าเช่าเดือนละ 8,000 บาท จ่ายค่าเช่าทุกสิ้นเดือน ไม่คิดค่าประกันร้านค้า 2. หากคำนึงถึงค่าใช้จ่ายในการเข้าสถานที่เท่านั้น ทนายควรจะทำสัญญาเช่ากับร้านใดจึงจะประหยัดค่าใช้จ่ายได้มากที่สุด และจ่ายไปปีละเท่าใด ตอบ.....

3. การเก็บรวบรวมข้อมูล

3.1 นำแบบทดสอบที่ผ่านการคัดเลือกข้อคำถามที่มีค่า IOC ตั้งแต่ .50 ขึ้นไป จำนวน 80 ข้อ ไปทดสอบกับกลุ่มทดลองใช้ (try out) จำนวน 120 คน เพื่อหาคุณภาพของแบบทดสอบ ซึ่งมีข้อสอบที่ผ่านเกณฑ์ 67 ข้อ

3.2 นำแบบทดสอบที่ผ่านคุณภาพมาจัดชุดข้อสอบจำนวน 4 ฉบับ ไปทดสอบกับกลุ่มตัวอย่าง 900 คน ซึ่งในการทดสอบใช้กลุ่มตัวอย่างโมเดลละ 225 คน โดยกำหนดให้ใช้ระยะเวลาในการสอบ 60 นาที

3.3 นำแบบทดสอบมาตรวจให้คะแนน โดยมีเกณฑ์สำหรับการให้คะแนน คือ ตอบถูกให้ 1 คะแนน และตอบผิดหรือไม่ตอบให้ 0 คะแนน

3.4 นำผลที่ได้จากการทดสอบมาวิเคราะห์หาคุณภาพของข้อสอบและค่าสารสนเทศของข้อสอบ

4. การวิเคราะห์ข้อมูล

4.1 การวิเคราะห์คุณภาพของแบบทดสอบ

คุณภาพของแบบทดสอบ พิจารณาจาก 1) ค่าดัชนีความสอดคล้อง (IOC) ด้านความเที่ยงตรงเชิงเนื้อหา (content validity) จากผู้เชี่ยวชาญ 3 ท่าน มีค่า IOC ตั้งแต่ .67 ขึ้นไป ถือว่าผ่านเกณฑ์และได้คัดเลือกไว้ 80 ข้อ คิดเป็นร้อยละ 100.00 2) ค่าความยาก (p) และค่าอำนาจจำแนก (r) เมื่อนำไปทดสอบกับกลุ่มทดลองใช้ (try out) โดยพิจารณาคัดเลือกข้อสอบที่มีค่าความยาก (p) อยู่ระหว่าง .20 ถึง .80 ค่าอำนาจจำแนก (r) ตั้งแต่ .20 ขึ้นไป ซึ่งมีข้อสอบที่ผ่านเกณฑ์ทั้งหมด 67 ข้อ คิดเป็นร้อยละ 83.75 และ 3) ค่าความเที่ยง (reliability) ของแบบทดสอบ ใช้วิธีของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson method) KR_{20} วิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูป

4.2 การวิเคราะห์คุณภาพของข้อสอบ

คุณภาพของข้อสอบ พิจารณาจาก ค่าสารสนเทศของข้อสอบ (item information function) โดยใช้โปรแกรม MULTILOG 7.03 การวิเคราะห์ค่าสารสนเทศของข้อสอบ (Karnjanawasee, 2007) ดังสูตร

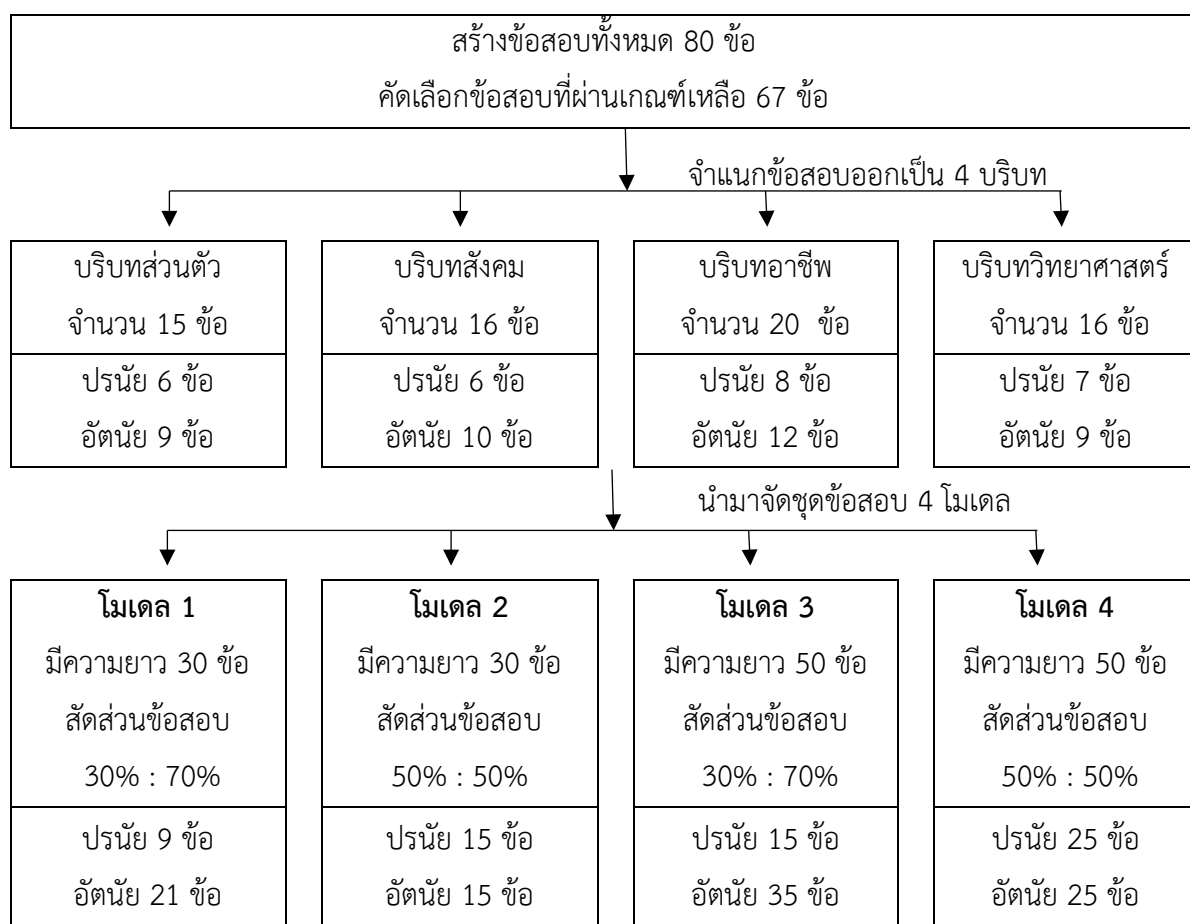
$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)} \quad , i = 1, 2, 3, \dots, k$$

เมื่อ $I_i(\theta)$	แทน	ค่าสารสนเทศ หรือค่าสารสนเทศที่ได้รับจากข้อสอบข้อที่ i สำหรับผู้ตอบที่มีความสามารถ θ
$P'_i(\theta) = P'_i$	แทน	ความชันของฟังก์ชันการตอบข้อสอบข้อที่ i ณ ตำแหน่งความสามารถ θ
$P_i(\theta) = P_i$	แทน	ความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง
$Q_i(\theta) = Q_i = 1 - P_i(\theta)$		

ผลการวิจัย

1. ผลการสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

1.1 ผลการจัดชุดข้อสอบแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมผู้วิจัยได้คัดเลือกข้อสอบที่ผ่านเกณฑ์การหาคุณภาพข้อสอบ จำนวน 67 ข้อ จากทั้งหมด 80 ข้อ นำมาจัดชุดข้อสอบจำแนกตามสถานการณ์และบริบทที่มีความยาวและสัดส่วนข้อสอบแตกต่างกันจำนวน 4 โมเดล ซึ่งแต่ละโมเดลได้กำหนดให้มีข้อสอบรวมจำนวน 11 ข้อ คิดเป็นร้อยละ 36.67 และร้อยละ 22.00 ตามที่ Srikolchan (2009) ระบุว่า แบบทดสอบรูปแบบผสมควรมีข้อสอบรวมอย่างน้อยร้อยละ 20 ของข้อสอบทั้งหมด ข้อสอบรวมดังกล่าวมีค่าความยาก (p) อยู่ระหว่าง .40 ถึง .69 และค่าอำนาจจำแนก (r) อยู่ระหว่าง .30 ถึง .72 ซึ่งถือว่าอยู่ในเกณฑ์คุณภาพที่เหมาะสม โดยผลการสร้างแบบทดสอบสรุปได้ดังภาพ 2

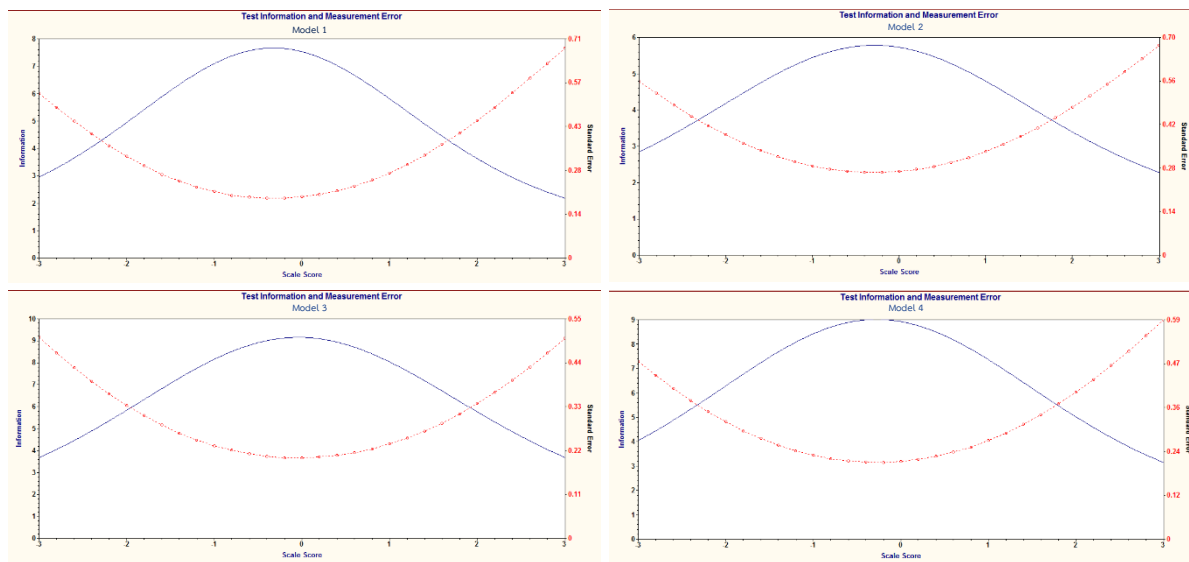


ภาพ 2 การจัดชุดข้อสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

การจัดชุดข้อสอบทั้ง 4 โมเดล มีข้อสอบรวม จำนวน 11 ข้อ เป็นข้อสอบรวมในบริบทส่วนตัว 4 ข้อ บริบทสังคม 4 ข้อ บริบทอาชีพ 1 ข้อ และบริบทวิทยาศาสตร์ 2 ข้อ

2. ผลการเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน

2.1 ค่าสารสนเทศของแบบทดสอบ (test information function) การรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน จำนวน 4 โมเดล ณ ช่วงความสามารถอยู่ระหว่าง -3 ถึง 3 พบว่า (1) โมเดล 1 มีค่าสูงสุด เท่ากับ 7.652 ณ ระดับความสามารถเท่ากับ -.40 (2) โมเดล 2 มีค่าสูงสุด เท่ากับ 5.774 ณ ระดับความสามารถ เท่ากับ -.20 (3) โมเดล 3 มีค่าสูงสุด เท่ากับ 9.150 ณ ระดับความสามารถ เท่ากับ .00 และ (4) โมเดล 4 มีค่าสารสนเทศของแบบทดสอบสูงสุด เท่ากับ 8.999 ณ ระดับความสามารถ เท่ากับ -.20 โดยมีผลการวิเคราะห์ค่าสารสนเทศของแบบทดสอบ ดังภาพ 3



ภาพ 3 ค่าสารสนเทศของแบบทดสอบทั้ง 4 โมเดล

2.2 การวิเคราะห์และเปรียบเทียบค่าสารสนเทศของข้อสอบ (item information function) ระหว่างโมเดลที่แตกต่างกัน ซึ่งนำความยาวและสัดส่วนข้อสอบมาพิจารณาพร้อมกันได้ 4 โมเดล โดยการวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA) ผลการวิเคราะห์ ดังตาราง 2

ตาราง 2 ผลการวิเคราะห์และเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดลที่แตกต่างกัน

แหล่งความแปรปรวน	SS	df	MS	F	p
ระหว่างกลุ่ม	.210	3	.070	19.457	.000*
ภายในกลุ่ม	.562	156	.004		
รวม	.772	159			

*p < .05

จากตาราง 2 ผลการวิเคราะห์และเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดลที่แตกต่างกัน พบว่า มีค่าเฉลี่ยสารสนเทศของข้อสอบของโมเดลแตกต่างกันอย่างน้อย 1 คู่ อย่างมีนัยสำคัญทางสถิติ .05 ($F = 19.457, p = .000$) ดังนั้นต้องทำการวิเคราะห์ Post-hoc เพื่อเปรียบเทียบค่าเฉลี่ยรายคู่โดยใช้วิธีของ Tamhane เพราะผลการตรวจสอบของ Levene's test พบว่า มีความแปรปรวนแตกต่างกันอย่างมีนัยสำคัญทางสถิติ .05 (Levene = 6.251, $p = .000$) ดังตาราง 3

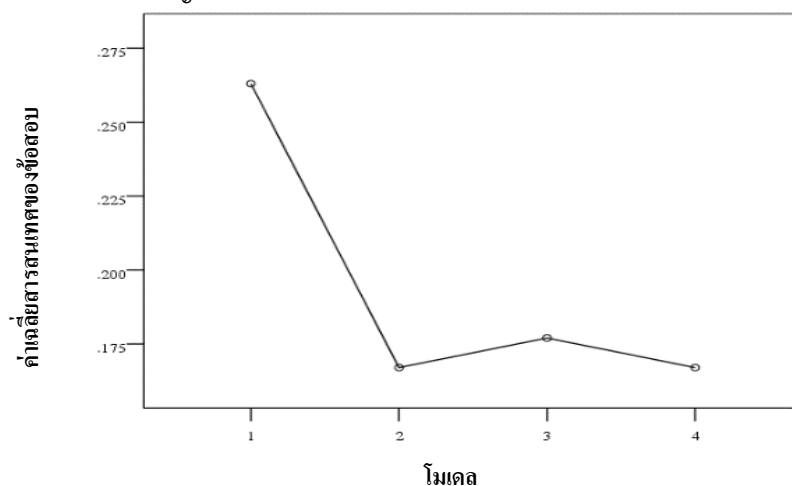
ตาราง 3 ผลการเปรียบเทียบรายคู่ (Post-hoc test) โดยวิธีของ Tamhane ในการเปรียบเทียบ

ค่าสารสนเทศของข้อสอบ (item information function) ระหว่างแต่ละโมเดลที่แตกต่างกัน

โมเดล	$\bar{x}$	โมเดล 1	โมเดล 2	โมเดล 3	โมเดล 4
		.2630	.1670	.1770	.1670
โมเดล 1	.2630	-	.0960*	.0860*	.0960*
			(.004)	(.012)	(.004)
โมเดล 2	.1670		-	-.0100*	.0000
				(.000)	(1.000)
โมเดล 3	.1770			-	.0100*
					(.000)
โมเดล 4	.1670				-

ตัวเลขในวงเล็บ คือ ค่า p-value * มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 3 ผลการเปรียบเทียบรายคู่ (Post-hoc test) โดยวิธีของ Tamhane ในการเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดล 4 แบบ พบว่า มีค่าเฉลี่ยของสารสนเทศของข้อสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 5 คู่ ซึ่งสรุปได้ว่า โมเดล 1 มีค่าเฉลี่ยสารสนเทศของข้อสอบสูงกว่าโมเดล 2, 3 และ 4 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และโมเดล 3 มีค่าเฉลี่ยสารสนเทศของข้อสอบสูงกว่าโมเดล 2 และ 4 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับผลในภาพ 4



ภาพ 4 ผลการเปรียบเทียบค่าเฉลี่ยสารสนเทศของข้อสอบระหว่างโมเดลที่แตกต่างกัน

อภิปรายผล

1. อภิปรายผลการสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน

จากผลการวิเคราะห์ความตรงเชิงเนื้อหาของแบบสอบรูปผสมวัดการรู้เรื่องคณิตศาสตร์ จำนวน 80 ข้อ ผ่านเกณฑ์การพิจารณาความตรงเชิงเนื้อหาจากผู้เชี่ยวชาญ โดยมีค่าดัชนีความสอดคล้องอยู่ระหว่าง 0.67 ถึง 1.00 คัดเลือกไว้จำนวน 80 ข้อ ซึ่งแสดงให้เห็นว่า ข้อสอบที่สร้างขึ้นมีความตรงเชิงเนื้อหา สอดคล้องกับ Saiyod & Saiyod (2000); Pasiphol (2016) ที่ได้กล่าวไว้ว่า ดัชนีความสอดคล้องที่คำนวณได้ต้องมีค่ามากกว่า 0.50 จึงจะมีความตรงเชิงเนื้อหา ที่เป็นเช่นนี้เพราะว่าผู้วิจัยได้ศึกษากรอบแนวคิดของ PISA เพื่อกำหนดกรอบโครงสร้าง (Test Blueprint) ที่ชัดเจน ศึกษางานวิจัยที่เกี่ยวข้องกับการสร้างแบบทดสอบรูปผสมของข้อสอบปรนัยและข้อสอบอัตนัย และตรวจสอบความตรงเชิงเนื้อหา โดยผู้เชี่ยวชาญพิจารณาความสอดคล้องระหว่างข้อคำถามกับเนื้อหาและจุดประสงค์เชิงพฤติกรรม 2 ครั้ง แล้วปรับปรุงแก้ไขตามที่ผู้เชี่ยวชาญเสนอแนะ ส่งผลให้แบบทดสอบที่ผู้วิจัยสร้างขึ้นเป็นแบบทดสอบที่สามารถนำไปใช้ได้

จากผลการวิเคราะห์ความเที่ยงของแบบทดสอบที่แตกต่างกันใน 4 โมเดล พบว่ามีค่าความเที่ยงใกล้เคียงกัน ดังนี้ โมเดล 1 และโมเดล 2 ซึ่งมีความยาว 30 ข้อ มีค่าความเที่ยง 0.842 และ 0.803 ตามลำดับ ขณะที่โมเดล 3 และโมเดล 4 ซึ่งมีความยาว 50 ข้อ มีค่าความเที่ยง 0.876 และ 0.872 ตามลำดับ จะเห็นว่าจำนวนข้อสอบที่มากขึ้นส่งผลให้แบบทดสอบมีค่าความเที่ยงสูงขึ้น สอดคล้องกับ Karnjanawasee (2007) ที่กล่าวว่าความยาวของแบบทดสอบที่ส่งผลต่อความเที่ยงของแบบทดสอบ แบบทดสอบที่มีความยาวมากจะมีผลให้ค่าความเที่ยงของแบบทดสอบสูงกว่าแบบทดสอบที่มีความยาวน้อยกว่า นอกจากนี้ยังสอดคล้องกับ Bastari (2000); Saen-amnuaiophon (2012) ที่กล่าวว่า แบบทดสอบควรมีความยาวอยู่ระหว่าง 30-50 ข้อ เพราะจะมีค่าความเที่ยงและค่าความตรงของแบบทดสอบใกล้เคียงกัน แต่แบบทดสอบที่มีความยาว 50 ข้อ มีแนวโน้มว่ามีประสิทธิภาพสูงกว่าความยาว 30 ข้อและ 10 ข้อ

2. อภิปรายผลการเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนแตกต่างกัน

จากผลการวิเคราะห์ค่าสารสนเทศของแบบทดสอบ (test information function) พบว่าแบบทดสอบทั้ง 4 โมเดลมีค่าสารสนเทศของแบบทดสอบระหว่าง 5.774 ถึง 9.150 ณ ความสามารถช่วง -.40 ถึง .00 สะท้อนให้เห็นว่าแบบทดสอบที่สร้างขึ้นเหมาะสมสำหรับนำไปใช้กับผู้สอบที่มีความสามารถปานกลางไปจนถึงความสามารถต่ำ จะให้ค่าสารสนเทศของข้อสอบที่เพียงพอ สอดคล้องกับงานวิจัยของ Kawmanee (2002) ที่เปรียบเทียบค่าสารสนเทศของแบบทดสอบเลือกตอบวิชาคณิตศาสตร์ ซึ่งผลการวิจัย พบว่า แบบทดสอบที่สร้างขึ้นเหมาะสมสำหรับผู้สอบที่มีความสามารถปานกลางไปจนถึงต่ำ ที่ความสามารถระหว่าง -.40 ถึง .30 แต่อย่างไรก็ดีเมื่อพิจารณาค่าสารสนเทศของแบบทดสอบ พบว่ามีค่าต่ำกว่า 10 ซึ่ง Karnjanawasee (2007) ได้กล่าวว่า ค่าสารสนเทศของแบบทดสอบควรมีค่ามากกว่า 10 ดังนั้นอาจเกิดจากลักษณะของรูปแบบข้อสอบรวมถึงลักษณะของตัวเลือกข้อสอบ ซึ่งสอดคล้องกับ Kawmanee (2002) ที่พบว่า ลักษณะของตัวเลือก

ในข้อสอบปรนัยมีผลต่อค่าสารสนเทศของแบบทดสอบโดยเฉพาะตัวเลือกแบบปลายเปิดส่งผลให้ค่าสารสนเทศของแบบทดสอบสูงกว่าตัวเลือกแบบปกติ

ผลการวิเคราะห์และเปรียบเทียบค่าสารสนเทศของข้อสอบระหว่างโมเดลที่ต่างกันพบว่า มีค่าเฉลี่ยสารสนเทศของข้อสอบแตกต่างกันอย่างน้อย 1 คู่ เมื่อเปรียบเทียบค่าเฉลี่ยรายคู่ของโมเดลทั้ง 4 พบว่าโมเดล 1 คือความยาว 30 ข้อ สัดส่วนระหว่างข้อสอบปรนัยและอัตนัย 30% : 70% มีค่าสารสนเทศสูงกว่าโมเดล 2, 3 และ 4 และโมเดล 3 คือ ความยาว 50 ข้อ สัดส่วนระหว่างข้อสอบปรนัยและอัตนัย 30% : 70% มีค่าสารสนเทศสูงกว่าโมเดล 2 และ 4 สะท้อนให้เห็นว่า ในกรณีแบบทดสอบที่มีความยาว 30 ข้อ และมีสัดส่วนข้อสอบอัตนัยมากกว่าข้อสอบปรนัย มีค่าสารสนเทศของข้อสอบสูงกว่า อาจเป็นเพราะข้อสอบอัตนัยในการวิจัยครั้งนี้แม้จะเป็นแบบเติมคำตอบสั้น แต่ก็เดาได้ยากกว่าข้อสอบปรนัย จึงสะท้อนความสามารถของผู้สอบได้ดีกว่าปรนัย ซึ่งมีโอกาสในการเดา จึงส่งผลให้ค่าสารสนเทศเมื่อมีข้อสอบอัตนัยสูงกว่าปรนัยด้วย สอดคล้องกับ Nidehah (2000); Kim et al. (2007) ที่กล่าวว่า ถ้ามีการเดาลดลงจะทำให้ค่าสารสนเทศของข้อสอบเพิ่มขึ้น

เมื่อพิจารณาความยาวของข้อสอบ จะเห็นว่า ในงานวิจัยนี้แบบทดสอบที่มีความยาว 30 ข้อ มีค่าสารสนเทศของข้อสอบสูงกว่าแบบทดสอบที่มีความยาว 50 ข้อ ซึ่งให้ผลแตกต่างจากการศึกษาของ Kinsey (2003); Saen-amnuaiophon (2012) ที่พบว่า แบบทดสอบที่มีความยาว 50 ข้อ มีค่าสารสนเทศสูงกว่าแบบทดสอบที่มีความยาว 30 ข้อ ทั้งนี้อาจเนื่องมาจาก ลักษณะของรูปแบบข้อสอบ และอีกประเด็นหนึ่งตามทฤษฎีการตอบสนองข้อสอบ ต้องเป็นการสอบที่ไม่แข่งขันด้านเวลา ความเร็วในการตอบจะต้องไม่มีอิทธิพลต่อผลการตอบ การจัดสอบจึงต้องไม่อยู่ในสถานการณ์ที่สอบแข่งขันกันด้วยเวลา แต่ในสภาพการทดสอบจริงไม่สามารถจัดการสอบที่ไม่แข่งขันด้านเวลาได้ในทุกครั้ง ซึ่งอาจมีผลต่อค่าพารามิเตอร์ของข้อสอบ สอดคล้องกับ Karnjanawasee (2007) ที่กล่าวว่า ค่าสารสนเทศของข้อสอบมีความสอดคล้องกับค่าความยากง่าย กล่าวคือ ค่าสารสนเทศของข้อสอบจะสูงขึ้น เมื่อค่าความสามารถสูงสุดใกล้เคียงกับค่าความยากง่าย

เมื่อพิจารณาสัดส่วนของข้อสอบ พบว่า แบบทดสอบที่มีสัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 30% : 70% มีค่าสารสนเทศของข้อสอบสูงกว่าแบบทดสอบที่มีสัดส่วนข้อสอบปรนัยและข้อสอบอัตนัย 50%:50% สอดคล้องกับ Kim et al. (2007); Reshetar & Melican (2010) ที่กล่าวถึงจุดเด่นของข้อสอบอัตนัยที่ช่วยลดโอกาสในการเดาคำตอบถูกของผู้สอบ หรือเมื่อไม่มีโอกาสในการเดา ค่าสารสนเทศของข้อสอบจะสูงขึ้น

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 จากผลการวิจัยที่พบว่า แบบทดสอบที่มีความยาว 50 ข้อ มีค่าความเที่ยงสูงกว่าความยาว 30 ข้อ ดังนั้นในบริบทของการนำแบบทดสอบไปใช้ ครูหรือผู้ที่มีส่วนเกี่ยวข้องหากนำไปใช้ในการวิจัยหรือการวัดผลมีระยะเวลาในการทดสอบที่เพียงพอ ควรเลือกใช้แบบทดสอบที่มีความยาว 50 ข้อ เนื่องจากให้ค่าความเที่ยงที่สูงกว่า 30 ข้อ

1.2 จากผลการวิจัยจำแนกตามโมเดลของแบบทดสอบ พบว่า โมเดล 1 ที่มีความยาว 30 ข้อ มีสัดส่วนระหว่างปรนัยและอัตนัย 30% : 70% มีค่าสารสนเทศของข้อสอบสูงกว่าสัดส่วนระหว่างข้อสอบปรนัยและข้อสอบอัตนัย 50% : 50% ดังนั้นในสถานการณ์ที่มีระยะเวลาจำกัด ครูอาจประหยัดเวลาโดยสร้างข้อสอบเพียง 30 ข้อได้ แต่ควรมีสัดส่วนข้อสอบอัตนัยสูงกว่าข้อสอบปรนัยจึงจะให้ค่าสารสนเทศของข้อสอบสูง

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรมีการศึกษาคุณสมบัติทางจิตมิติของแบบทดสอบรูปแบบผสมที่มีความยาวและสัดส่วนของข้อสอบที่แตกต่างกัน โดยการประยุกต์ใช้โมเดลตอบสนองข้อสอบ 2 พารามิเตอร์และ 3 พารามิเตอร์ เนื่องจากให้สารสนเทศที่เป็นประโยชน์ในการสร้างคลังข้อสอบของแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสม ซึ่งประกอบด้วยค่าอำนาจจำแนก (a) และค่าโอกาสในการเดา (c) ต้องทำการศึกษาเกี่ยวกับกลุ่มตัวอย่างที่มีขนาดใหญ่ขึ้น เพื่อลดความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ข้อสอบและความสามารถของผู้สอบ

2.2 ควรมีการศึกษการสร้างแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมโดยประยุกต์ใช้ทฤษฎีตอบสนองข้อสอบ ที่มีรูปแบบการตรวจให้คะแนนมากกว่าสองค่า หรือการตรวจให้คะแนนบางส่วน เนื่องจากเนื้อหาวิชาคณิตศาสตร์บางสาระจำเป็นที่จะต้องให้ผู้เรียนเขียนแสดงวิธีคิดซึ่งสามารถตรวจให้คะแนนได้หลายค่า

กิตติกรรมประกาศ

บทความวิจัยเรื่อง การเปรียบเทียบค่าสารสนเทศของข้อสอบในแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน เป็นส่วนหนึ่งของวิทยานิพนธ์ระดับปริญญาโทเรื่อง การเปรียบเทียบคุณสมบัติทางจิตมิติของแบบทดสอบการรู้เรื่องคณิตศาสตร์รูปแบบผสมที่มีความยาวและสัดส่วนข้อสอบแตกต่างกัน ซึ่งเป็นส่วนหนึ่งของการสำเร็จการศึกษาในหลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาวิชาการประเมินและการวิจัยการศึกษา มหาวิทยาลัยรามคำแหง โดยมี ผู้ช่วยศาสตราจารย์ ดร.จุฑาภรณ์ มาสันเทียะ เป็นที่ปรึกษาวิทยานิพนธ์หลัก และอาจารย์ ดร.พนิดา พานิชวัฒน์ เป็นที่ปรึกษาวิทยานิพนธ์ร่วม

References

- Bastari, B. (2000). *Linking Multiple – Choice and Constructed – Response Items to a common Proficiency Scale* [Doctoral dissertation]. Massachusetts Amherst University.
- Jurich, D., & Goodman, J. (2009). *A Comparison of IRT Parameter Recovery in Mixed Format Examinations Using PARSCALE and ICL* [Poster presented]. The Annual meeting of Northeastern Educational Research Association.

- Kim, S., Walker, M. E., & McHale, F. (2007). *Equating of Mixed-Format Tests in Large Scale Assessments* [Paper Presented]. The Annual Meeting of the American Educational Research Association (AERA) and The National Council on Measurement in Education (NCME).
- Kinsey, T.L. (2003). *Comparison of IRT and Rasch Procedures in A Mixed- Item Format Test* [Doctoral dissertation]. University of North Texas.
- Reshetar, R., & Melican, G. J. (2010). *Design and Evaluation of Mixed-Format Large Scale Assessments for the Advanced Placement Program (AP)* [Paper presented]. The Annual Meeting of the American Education Research Association in Denver.
- Shin, D. (2007). *A Comparison of Methods of Estimating Subscale Scores for Mixed Format Tests*. Pearson Educational Measurement.

Translate Thai References

- Chaisuwan, K. (2012). *A Comparison of Psychometric Properties among the Short Forms Developed Using Different Methods: An Adversity Quotient Test Study* [Master's thesis]. Chulalongkorn University. (in Thai)
- Institute for the Promotion of Teaching Science and Technology. (2018). *PISA 2015 Assessment Results Science Reading and Math Educational Excellence and Equality*. Success Public Company Limited. (in Thai)
- Karnjanawasee, S. (2007) *New Exam Theory*. (3th Ed.). Chulalongkorn University Press. (in Thai)
- Kawmanee, K. (2002) *A Comparison of Test Information Function of Mathematics Multiple Choice Test with Different Model Options* [Master's thesis]. Thaksin University. (in Thai)
- Nansombat, U. (1999). *A Study of Item Analysis According to Item Response Theory Calculated from Various Sample Sizes and Test Lengths* [Master's thesis]. Mahasarakham University. (in Thai)
- Nidehah, N. (2000). *A Comparison of Test Information Function of Mathematics Multiple Choice Test with Different Type of Choices* [Master's thesis]. Songkla University. (in Thai)
- Pasiphol, S. (2016). *Learning Measurement and Evaluation*. Chulalongkorn University Press. (in Thai)

- Saen-amnuaiphon, R. (2012). *The Efficiency of Mixed-Format Achievement Test: An Application of Partial Credit Model and Generalized Partial Credit Model* [Doctoral dissertation]. Khon Kaen University. (in Thai)
- Saiyod, L., & Saiyod, U. (2000). Techniques for measuring learning outcomes. Suweeriyasan. (in Thai)
- Sawangsri, P. (2015). *Comparison of Psychometric Properties Integrating Indicators of Reading Analytical Thinking and Writing Abilities and Indicators of Content Science of Nine Grade using The Different Item Review Methods* [Master's thesis]. Chulalongkorn University. (in Thai)
- Srikolchan, A. (2009). *The Quality of Equating for Mixed-Format Test: An Application of IRT Equating with Characteristic Curve and Concurrent Calibration Methods* [Doctoral dissertation]. Chulalongkorn University. (in Thai)

การออกแบบวิธีการตรวจให้คะแนนอัตโนมัติเพื่อวินิจฉัยมโนทัศน์ ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ผ่านการเรียนรู้ของเครื่อง

Designing Automated Scoring Approach to Diagnose Scientific Competency Misconceptions through Machine Learning

ประไพพิมพ์ ทองเปราะ^{1*} พัทธี จันทรเพ็ง² ชัยวัฒน์ ทะวะรุ่งเรือง³ และ ธนพงศ์ อินทะ⁴

Prapaipim Thongprowh^{1*}, Putcharee Junpeng², Chaiwat Tawarungruang³ and Thanapong Intharah⁴

(Received: March 31, 2022; Revised: May 23, 2022; Accepted: May 27, 2022)

บทคัดย่อ

การออกแบบวิธีการตรวจให้คะแนนอัตโนมัติเพื่อวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ผ่านการเรียนรู้ของเครื่อง (machine learning) เป็นการออกแบบวิธีการตรวจให้คะแนนแบบสอบวินิจฉัยรูปแบบผสมเพื่อประเมินระดับสมรรถนะทางวิทยาศาสตร์โดยใช้เครื่องตรวจคะแนนในขั้นตอนเดียวนวัตกรรมนี้สามารถรายงานผลและให้ข้อมูลป้อนกลับแก่ผู้เรียนแบบเรียลไทม์ ดังนั้นการวิจัยครั้งนี้ได้กำหนดความมุ่งหมายเพื่อนำไปสู่การออกแบบนวัตกรรมต้นแบบดังนี้ (1) เพื่อวิเคราะห์ผลการตอบของผู้เรียนสู่การออกแบบนวัตกรรมต้นแบบ ประกอบด้วยรูปแบบและระดับของมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ การกำหนดจุดตัดจากผลการตอบแบบวินิจฉัยสมรรถนะวิทยาศาสตร์ และการกำหนดตัวแปรเพื่อสร้างสมการทำนายจากข้อมูลพื้นฐานของนักเรียน และ (2) เพื่อออกแบบวิธีการตรวจให้คะแนนอัตโนมัติในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ผ่านการเรียนรู้ของเครื่อง การศึกษาครั้งนี้ใช้ข้อมูลพฤติกรรมจากกลุ่มผู้สอบคือนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 จำนวน 847 คน วิเคราะห์ข้อมูลตามโมเดลพหุมิติสัมพันธ์การสุ่มแบบโลจิก (MRCMLM) และกำหนดจุดตัดบน wright map ผลการศึกษาพบว่า

¹ นักศึกษาปริญญาโท สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² รองศาสตราจารย์ สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

³ นักวิจัย ศูนย์จัดการข้อมูลและวิเคราะห์ทางสถิติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

⁴ อาจารย์ ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

งานวิจัยนี้ได้รับทุนอรรถวิญญานักวิจัยรุ่นกลาง ประจำปี 2565 จากสำนักงานการวิจัยแห่งชาติ (วช.) รหัสโครงการวิจัย N42A650236

¹ Master's degree student, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

² Associate Professor, Department of Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

³ Researcher, Data Management and Statistical Analysis Center, Faculty of Public Health, Khon Kaen University

⁴ Lecturer, Department of Statistics, Faculty of Science, Khon Kaen University

The research received a research grant for Talented Mid-Career Researchers 2022 from the National Research Council of Thailand (NRCT) Project No N42A650236

* Corresponding Author E-mail: prapaipim_th@kkumail.com

(1) เกิดมโนทัศน์ที่คลาดเคลื่อนขึ้น 5 รูปแบบ และการหาจุดตัดทางสมรรถนะวิทยาศาสตร์ทั้ง 3 มิติ พบว่า มิติด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ มีจุดตัดที่ -0.49 0.27 และ 0.83 มิติด้านการประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ มีจุดตัดที่ -0.09 0.54 และ 1.72 ตามลำดับ และมิติด้านการแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์มีจุดตัดที่ -0.78 0.47 และ 2.98 ตามลำดับ นอกจากนี้พบตัวแปรต้นที่สำคัญ คือ (1) เกรดเฉลี่ยรวมเทอมที่ผ่านมา (2) เกรดเฉลี่ยวิชาวิทยาศาสตร์เทอมที่ผ่านมา และ (3) จำนวนชั่วโมงที่ศึกษาวิชาวิทยาศาสตร์ด้วยตนเอง/สัปดาห์

(2) การออกแบบวิธีการตรวจให้คะแนนอัตโนมัติเพื่อวินิจฉัยสมรรถนะวิทยาศาสตร์ในข้อสอบอัตนัย ใช้วิธีการจับคำ (keywords) และใช้วิธีการแปลงคำตอบจากการสังเคราะห์ผลการตอบในข้อมูลชุดเดียวให้เป็นตัวเลือกเพื่อช่วยให้การตรวจให้คะแนนและประมวลผลระดับสมรรถนะและการรายงานระดับมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์สามารถทำได้แบบอัตโนมัติและเรียลไทม์ สำหรับผลการประเมินคุณภาพการออกแบบวิธีการตรวจให้คะแนนจากผู้เชี่ยวชาญจำนวน 5 คน ครอบคลุมทั้ง 5 มาตรฐาน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสมขอรับรอง ด้านความถูกต้อง และด้านความรับผิดชอบการประเมิน

คำสำคัญ: การเรียนรู้ของเครื่อง วิธีการตรวจให้คะแนนอัตโนมัติ มโนทัศน์ที่คลาดเคลื่อน สมรรถนะวิทยาศาสตร์ โมเดลพหุมิติสัมประสิทธิ์การสุ่มแบบโลจิท

Abstract

Designing automated scoring to diagnose misconceptions in scientific competencies through machine learning is a scoring method for a mixed-format diagnostic test to assess scientific competency levels through a scoring machine in a single step. This innovation can report results and provide real-time feedback to learners. Therefore, to design the prototypical innovation, the objectives of the present study were defined as follows: (1) to analyze the students' response patterns to the design of the prototypical innovation, including identifying types and levels of misconceptions in scientific competencies, defining cut scores based on response patterns in the scientific competency diagnostic test, and determining variables to create the predictive equation based on students' background data; and (2) to design automated scoring for diagnosis of misconceptions in scientific competencies through machine learning. This study drew on secondary data of a sample of 847 grade 7 students. The data were analyzed using a multidimensional random coefficients multinomial logit model (MRCMLM) while cut scores on the Wright map were defined. The results are reported below.

(1) There were 5 types of misconceptions in scientific competencies. Besides, on determining cut scores of 3 scientific competencies, it was found that the competency to explain

phenomena scientifically had the cut scores of -0.49, 0.27 and 0.83; the cut scores of the competency to evaluate and design scientific enquiry were -0.09, 0.54 and 1.72 respectively; and those of the competency to interpret data and use evidence scientifically were -0.78, 0.47 and 2.98 respectively. In addition, the important independent variables were discovered as follows: (1) the total GPA of the previous semester; (2) the GPA in the science subject of the previous semester; and (3) the number of hours of self-study on a science subject per week.

(2) The designed scoring to diagnose scientific competencies in subjective tests employed identified keywords and converted answers from the synthesis of response patterns in the secondary data into choices to allow for scoring, processing scientific competency levels, and reporting the levels of misconceptions automatically and in real time. The results also showed that, based on five experts' evaluation, the design of automated scoring was found to cover 5 standards, namely utility, feasibility, propriety, accuracy and accountability.

Keywords: machine learning, automated scoring, misconception, scientific competency, multidimensional random coefficients multinomial logit model

บทนำ

การทดสอบสมรรถนะวิทยาศาสตร์ระดับนานาชาติด้วยโปรแกรมประเมินสมรรถนะนักเรียนตามมาตรฐานสากล (programmed for international student assessment; PISA) เน้นความฉลาดรู้ด้านวิทยาศาสตร์ (scientific literacy) ภายใต้กรอบการประเมินสมรรถนะทางวิทยาศาสตร์ (scientific competencies) ประกอบด้วย (1) การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ (2) การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และ (3) การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ (Organization for Economic Co-operation and Development, 2019) ซึ่งผลการประเมินนักเรียนไทยที่ผ่านพบว่ามีคะแนนระดับสมรรถนะต่ำกว่าค่าเฉลี่ยที่กำหนด เมื่อพิจารณาเป็นรายสาระพบว่าสาระที่ 2 มีคะแนนต่ำกว่าค่าเฉลี่ยรวมของวิชาวิทยาศาสตร์ โดยเฉพาะอย่างยิ่งในเนื้อหา เรื่อง สารและสมบัติของสาร สอดคล้องกับการเรียนวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ที่ครูผู้สอนพบว่าในสาระดังกล่าวนักเรียนมักเกิดมโนทัศน์ที่คลาดเคลื่อนต่อการเรียนรู้ที่ส่งผลต่อระดับความสามารถทางวิทยาศาสตร์ของนักเรียนโดยตรง โดยการวัดระดับความรู้ในโรงเรียนและ PISA ต่างมุ่งเน้นการประเมินด้านวิทยาศาสตร์ของนักเรียน สำหรับแนวทางที่จะช่วยลดการเกิดมโนทัศน์ที่คลาดเคลื่อนคือ “การวินิจฉัย” เพื่อหาความคลาดเคลื่อนจากการเรียนรู้ของนักเรียน (Khawsuk & Srihaset, 2016)

การวินิจฉัย (diagnosis) เป็นเครื่องมือประเมินผลการเรียนรู้ที่ครูใช้วัดและประเมินผลนักเรียน เพราะการวินิจฉัยสามารถชี้ให้เห็นข้อบกพร่องของนักเรียนเป็นรายบุคคล (Insawat, 2016) จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องพบว่าแบบทดสอบเพื่อการวินิจฉัยมีหลายรูปแบบ เช่น ปรนัย อัตนัย และผสม

โดยเฉพาะแบบผสมนั้นสามารถลดข้อจำกัดของปรนัยและอัตนัย และมีจุดเด่นคือสามารถวัดทักษะของผู้เรียนทุกด้านในฉบับเดียว (Boonjun, 2019) ปัจจุบันทฤษฎีการตอบสนองข้อสอบแบบพหุมิติเป็นทฤษฎีที่ใช้อย่างกว้างขวางเพื่อสร้างแบบวินิจัย โดยมีโมเดลที่มุ่งไปที่ปฏิสัมพันธ์ระหว่างผู้สอบและข้อสอบและพิจารณาจากผลการตอบของผู้สอบเพื่อประมาณค่าพารามิเตอร์ของผู้สอบในหลายมิติ (Junpeng, 2018) ซึ่งคุณลักษณะของข้อสอบพหุมิติมีอยู่ 2 ลักษณะ คือ ความเป็นพหุมิติระหว่างข้อคำถามและความเป็นพหุมิติภายในข้อคำถาม (Chianchana, 2009; Rawanprakhon & Tayraukham, 2017) สำหรับโมเดลพหุมิติสัมพันธ์ การสุ่มแบบโลจิต (MRCMLM) เป็นหนึ่งในโมเดลของทฤษฎีนี้ที่นำมาประยุกต์ใช้ในการวิเคราะห์โครงสร้างข้อสอบเพื่อตรวจสอบคุณภาพของเครื่องมือทั้งคุณภาพรายข้อ และคุณภาพของเครื่องมือทั้งฉบับ ทั้งสามารถตรวจสอบความตรงเชิงโครงสร้างภายในโดยสะท้อนออกมาเป็น wright map ที่เป็นหัวใจสำคัญของการตรวจสอบคุณภาพซึ่งนำไปสู่การตีความของนักเรียนรายบุคคล โดยมีงานวิจัยที่โมเดลนี้สร้างเครื่องมือวัดเพื่อประเมินผู้เรียนในหลายด้าน เช่น การพัฒนาแบบวัดพหุมิติทักษะการแก้ปัญหาแบบร่วมมือสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น (Wongphummueang & Piyakun, 2021) และการพัฒนาข้อสอบพหุมิติเพื่อวัดสมรรถนะทางวิทยาศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 (Intasoi, 2020)

งานวิจัยที่ผ่านมาข้อสอบรูปแบบผสมพบในเครื่องมือวัดแบบออนไลน์มากขึ้นแต่การตรวจให้คะแนนในเครื่องมือออนไลน์ยังมีจุดอ่อนคือในข้อสอบอัตนัยต้องรอให้คนตรวจจึงสามารถรายงานผลการประเมินได้ ดังนั้นปัญหาที่เกิดขึ้นของการตรวจให้คะแนนในส่วนข้ออัตนัยคือไม่สามารถป้อนข้อมูลกลับสู่ผู้สอบได้ทันที (Fiothong et al., 2022) สำหรับบริบทของสมรรถนะทางวิทยาศาสตร์พบว่ามุ่งเน้นไปที่การประเมินระดับสมรรถนะทางวิทยาศาสตร์แต่ขาดการวินิจัยมิติที่คลาดเคลื่อนที่เกิดขึ้น เช่น งานวิจัยของ Intasoi et al. (2020); Intasoi (2020) ที่พัฒนาข้อสอบแบบพหุมิติผ่านเครื่องมือดิจิทัลในบทเรียนเรื่อง สารบริสุทธิ์ ครอบคลุมการวัดระดับสมรรถนะทางวิทยาศาสตร์ใน 3 มิติ ภายใต้กรอบของ PISA โดยประยุกต์ใช้โมเดลเชิงโครงสร้างและใช้เทคนิคการวิเคราะห์ข้อมูลตามโมเดลการตอบสนองข้อสอบแบบพหุมิติในการวัดสมรรถนะ งานวิจัยดังกล่าวได้ให้ข้อมูลเพื่อประเมินระดับสมรรถนะของนักเรียนแต่มีข้อจำกัดคือ แยกขั้นตอนเป็น 3 ส่วน คือ การทำแบบทดสอบของผู้เรียนผ่าน google form (2) การตรวจให้คะแนน และ (3) การรายงานผลการประเมินกลับแต่ในข้อที่ตรวจให้คะแนนหลายค่ายังต้องรอให้ผู้ประเมินมาตรวจทำให้ไม่สามารถวินิจัยและให้ข้อมูลป้อนกลับได้ทันทีประกอบกับยังขาดการให้สารสนเทศเพื่อวิเคราะห์มิติที่คลาดเคลื่อนทางสมรรถนะทางวิทยาศาสตร์ ดังนั้นจึงต้องหาวิธีการที่สามารถดำเนินการวินิจัยระดับสมรรถนะและรายงานระดับมิติที่คลาดเคลื่อนในขั้นตอนเดียว

ปัจจุบันการนำปัญญาประดิษฐ์ (artificial intelligence, AI) มาช่วยพัฒนาการตรวจให้คะแนนนั้นเรียกว่า “ระบบการตรวจให้คะแนนอัตโนมัติ” หรือ “automatic scoring system” (Sudjai & Mungsing, 2021) และหนึ่งในวิธีการตรวจแบบอัตโนมัตินี้คือการนำการเรียนรู้ของเครื่อง (Machine learning, ML) ที่เป็นส่วนย่อยของปัญญาประดิษฐ์เข้ามาช่วยพัฒนาโดยใช้อัลกอริทึมมาคาดการณ์ซึ่งอาศัยหลักการเรียนรู้จากกลไกของตัวเครื่อง ในด้านการศึกษาได้นำการเรียนรู้ของเครื่องสร้างเครื่องมือวัดที่สามารถตรวจให้คะแนนอัตโนมัติและรายงานผลการเรียนรู้สู่ผู้เรียนทันที เช่น การให้คะแนนโดยใช้ระบบตอบกลับอัตโนมัติ (Lee et al., 2021) และการตรวจให้คะแนนข้อสอบพหุมิติในวิชาเคมีและฟิสิกส์ (Maestrales et al., 2021) ซึ่งหลักการ

ตรวจให้คะแนนอัตโนมัติผ่านการเรียนรู้ของเครื่องจะอาศัยหลักการจำแนก (classification) และหลักการทำนาย (prognosis) ที่ต้องอาศัยตัวแปรเพื่อนำมาสร้างสมการทำนายปรากฏการณ์ที่จะเกิดขึ้น (Lintean et al., 2012) ทำให้ผลการวิเคราะห์โดยการเรียนรู้ของเครื่องถูกต้องและไม่ต่างจากการใช้คนตรวจ มีความแม่นยำสามารถลดขั้นตอนระยะเวลา และให้ผลการวินิจฉัยได้แบบเรียลไทม์

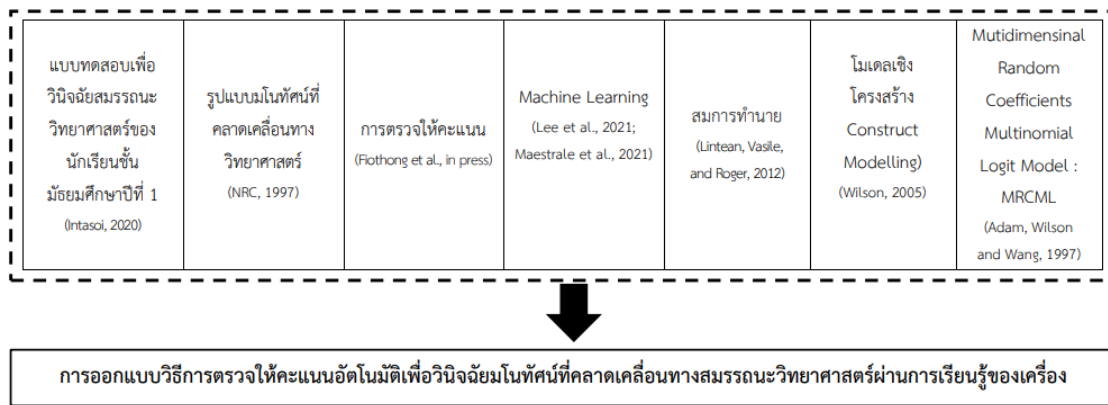
ดังนั้นเพื่อต่อยอดงานวิจัยของ Intasoi et al. (2020); Intasoi (2020) ที่มีจุดอ่อนในเรื่องการตรวจให้คะแนนเนื่องจากไม่สามารถตรวจอัตโนมัติแบบเรียลไทม์และเพื่อหาแนวทางสำคัญในการพัฒนาแพลตฟอร์มการตรวจให้คะแนนอัตโนมัติในขั้นตอนเดียวที่ถูกต้องและที่มีความคงเส้นคงวาในการตรวจให้คะแนนจากการศึกษาที่ผ่านมาพบว่าหากนำทฤษฎีของ MRCMLM และหลักการทำงานของการทำงานของเครื่องมาสร้างเครื่องมือวัด โดยใช้ทฤษฎี MRCMLM เป็นฐานในการวิเคราะห์ชุดข้อมูลขนาดใหญ่นำไปสู่การวินิจฉัยระดับความสามารถของนักเรียน โดยอาศัยวิธีการจำแนกกับหลักการทำนายของการเรียนรู้ของเครื่องในการประมวลผลเพื่อการตรวจให้คะแนนอัตโนมัติ จากที่กล่าวมาหากนำมาประยุกต์เพื่อสร้างระบบการตรวจให้คะแนนผ่านการเรียนรู้ของเครื่องจะทำให้การตรวจให้คะแนนสามารถทำได้ไม่ต่างจากการใช้คนตรวจและยังรายงานผลการวินิจฉัยได้แบบเรียลไทม์ ด้วยเหตุนี้คณะผู้วิจัยจึงสนใจออกแบบวิธีการตรวจให้คะแนนอัตโนมัติเพื่อการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ผ่านการเรียนรู้ของเครื่อง โดยนำข้อมูลทฤษฎีจากงานวิจัยของ Intasoi (2020) มาวิเคราะห์ซึ่งประกอบด้วย รูปแบบและระดับของมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ การกำหนดจุดตัดจากผลการตอบของนักเรียนเพื่อใช้เป็นเกณฑ์ในการวินิจฉัยระดับความสามารถของนักเรียน การวิเคราะห์ข้อมูลพื้นฐานของนักเรียนเพื่อกำหนดตัวแปรต้นและคะแนนในแต่ละมิติเป็นตัวแปรตามในการสร้างสมการทำนาย รวมถึงการออกแบบวิธีการตรวจให้คะแนนอัตโนมัติผ่านการเรียนรู้ของเครื่อง ซึ่งงานวิจัยนี้จะเป็นประโยชน์อย่างยิ่งต่อการตรวจให้คะแนนเพื่อรายงานผลการวินิจฉัยระดับการเรียนรู้แก่นักเรียนเป็นรายบุคคล เพราะช่วยลดระยะเวลาในการตรวจและลดทรัพยากรบุคคลในการทำงานเนื่องจากมีระบบตรวจให้คะแนนที่ตรงและแม่นยำเทียบเท่าบุคคลตรวจ ทั้งยังเพิ่มความสะดวกในการใช้งานเพราะเป็นระบบออนไลน์ที่ครูและนักเรียนสามารถเข้าถึงได้ง่าย

ความมุ่งหมายของการวิจัย

1. เพื่อวิเคราะห์ผลการตอบของนักเรียนสู่การออกแบบนวัตกรรมต้นแบบ ประกอบด้วย รูปแบบและระดับของมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ การกำหนดจุดตัดจากผลการตอบแบบวินิจฉัยสมรรถนะวิทยาศาสตร์ และการกำหนดตัวแปรเพื่อสร้างสมการทำนายจากข้อมูลพื้นฐานของนักเรียน
2. เพื่อออกแบบวิธีการตรวจให้คะแนนอัตโนมัติในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ผ่านการเรียนรู้ของเครื่อง

กรอบแนวคิดในการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ซึ่งสามารถกำหนดกรอบแนวคิดการวิจัยได้ดังภาพ 1



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยการออกแบบ (Design Research: DR) (Reeves, 2006; Wongwanit, 2020) มีขั้นตอนในการดำเนินการ 4 ระยะ คือระยะที่ 1 วิเคราะห์ข้อมูลชุดข้อมูล ระยะที่ 2 ออกแบบวิธีการตรวจให้คะแนนผ่านการเรียนรู้ของเครื่อง ระยะที่ 3 นำนวัตกรรมใหม่ไปทดลองใช้ และระยะที่ 4 นำเสนอหลักการออกแบบระบบการตรวจให้คะแนนอัตโนมัติผ่านการเรียนรู้ของเครื่อง ซึ่งบทความวิจัยนี้เป็นการศึกษานำร่องของ 2 ระยะแรก โดยมีรายละเอียดทั้ง 2 ระยะ ดังนี้

ระยะที่ 1 วิเคราะห์ข้อมูลชุดข้อมูลจากผลการตอบแบบวินิจฉัยสมรรถนะวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1

มุ่งเน้นการวิเคราะห์ข้อมูลชุดข้อมูลจากการตอบแบบวินิจฉัยของนักเรียน โดยมีรายละเอียดดังนี้

1. ประชากรและตัวอย่างวิจัย

ประชากรและตัวอย่างวิจัยของงานวิจัยครั้งนี้มาจากข้อมูลชุดข้อมูลจากงานวิจัยเรื่องการพัฒนาเครื่องมือดิจิทัลเพื่อประเมินสมรรถนะทางวิทยาศาสตร์แบบพหุมิติของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ปีพุทธศักราช 2562 (Intasoi et al., 2020; Intasoi, 2020) ประชากรคือนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2562 จำนวน 5,685 คน ใช้วิธีกำหนดตัวอย่างวิจัยขั้นต่ำของ Wright & Stone (1979) เพื่อกระจายความสามารถของตัวอย่าง โดยการสุ่มแบบหลายขั้นตอนจำนวน 847 คน

2. เครื่องมือที่ใช้ในการวิจัย

1) แบบทดสอบวินิจฉัยพหุมิติแบบผสมเพื่อวัดสมรรถนะวิทยาศาสตร์ที่พัฒนาโดย Intasoi (2020) จำนวน 22 ข้อ โดยมีมิติการวัด 3 มิติ ภายใต้กรอบของ PISA ประกอบด้วย มิติการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ จำนวน 10 ข้อ มิติการประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ จำนวน 4 ข้อ และมิติการแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ จำนวน 8 ข้อ มีค่าดัชนีความตรงเชิงเนื้อหารายข้อ (I-CVI) อยู่ระหว่าง 0.50 – 1.00 และค่าดัชนีความตรงเชิงเนื้อหาทั้งฉบับ (S-CVI) เท่ากับ 0.90 ค่าสัมประสิทธิ์ความเที่ยงแบบ EAP/PV ในมิติที่ 1 2 และ 3 มีค่าเท่ากับ 0.71 0.61 และ 0.69 ตามลำดับ ค่าความเหมาะสมรายข้อ และค่า INFIT MNSQ อยู่ระหว่าง 0.82 - 1.28

2) แบบฟอร์มวิเคราะห์หมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์จาก National Research Council (1997) ร่วมกับหลักแนวคิดความเข้าใจทางวิทยาศาสตร์ (Desstya et al., 2019)

3.การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูล

ผู้วิจัยดำเนินการศึกษาข้อมูลที่ใช้ประกอบการสร้างจุดตัดระดับสมรรถนะทางวิทยาศาสตร์ มีรายละเอียดดังขั้นตอนต่อไปนี้

1) ศึกษาข้อมูลทุติยภูมิเพื่อกำหนดรูปแบบและระดับของหมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ และวิเคราะห์แผนที่โครงสร้างระดับผลลัพธ์การเรียนรู้ ดังภาพ 2

ระดับ	มิติการอธิบายปรากฏการณ์ ในเชิงวิทยาศาสตร์	มิติการประเมินและออกแบบกระบวนการสืบเสาะหา ความรู้ทางวิทยาศาสตร์	มิติการแปลความหมายข้อมูลและการใช้ ประจักษ์พยานในเชิงวิทยาศาสตร์
4  สูง (High Level)	- สามารถเขียนหรืออธิบายระดับขอบเขตความสัมพันธ์ของแนวคิด และโมเดลทางวิทยาศาสตร์ เพื่อสร้างคำอธิบายของปรากฏการณ์ เหตุการณ์และกระบวนการ - ใช้ความรู้ที่เหนือกว่าหลักสูตรวิทยาศาสตร์และแสดงให้เห็นถึงวิธีการได้มาซึ่งความรู้ที่เหมาะสม	- สามารถประเมินการออกแบบการแข่งขันที่ซับซ้อน ศึกษาภาคสนาม หรือสถานการณ์จำลอง และการปรับตัวเลือก	- สามารถประเมินความแข็งแกร่งจากข้อมูลจากสมมติฐานและปรับข้อสรุปโดยใช้โมเดลทางวิทยาศาสตร์ที่เป็นนามธรรม - สามารถแยกความแตกต่างระหว่างข้อมูลที่เกี่ยวข้องและไม่เกี่ยวข้อง และสร้างคำอธิบายได้
3 ปานกลาง (Intermediate Level)	- จัดจำ หรือใช้แนวคิดทางวิทยาศาสตร์ที่ได้รับเพื่อสร้างคำอธิบายของเหตุการณ์และกระบวนการที่ค่อนข้างซับซ้อน - ใช้แนวคิดหรือโมเดลทางวิทยาศาสตร์ที่เป็นนามธรรมเพื่ออธิบาย	- ทำการทดลองเกี่ยวกับตัวแปรอิสระสองตัวหรือมากกว่า โดยแสดงขั้นตอน และสามารถตีความข้อมูล และตั้งข้อสรุปที่เหมาะสม - ใช้ข้อมูลจากบริบทที่มีความคุ้นเคยน้อย เพื่อระบุแนวโน้มและทำนาย - ประเมินการออกแบบหรือการตีความข้อมูล และการปรับตัวเลือก ระบุข้อจำกัดของการตีความของชุดข้อมูล	- สามารถตีความ และจัดการกับชุดข้อมูลที่มีความซับซ้อนในระดับปานกลางเพื่อเลือกหรือแสดงข้อสรุปที่เหมาะสม - สามารถแยกความแตกต่างระหว่างปัญหาทางวิทยาศาสตร์และสังคมหรือส่วนบุคคล - สามารถตีความชุดข้อมูลที่มีความซับซ้อนในระดับปานกลางโดยใช้โมเดลทางวิทยาศาสตร์
2 พื้นฐาน (Basic Level)	- จัดจำ และใช้ข้อเท็จจริงและแนวคิดทางวิทยาศาสตร์อย่างง่าย หรือเลือกคำอธิบายวิทยาศาสตร์อย่างง่าย - เขียนข้อเท็จจริงและแนวคิดทางวิทยาศาสตร์ ในการระบุหรือสร้างคำอธิบายง่าย ๆ ที่คุ้นเคยอย่างเหมาะสม	- เขียนขั้นตอน ในการออกแบบการทดลองอย่างง่าย - รวบรวมและตีความข้อมูล แยกความแตกต่างระหว่างคำถามที่ไม่ใช่วิทยาศาสตร์และวิทยาศาสตร์ - เขียนขั้นตอนในการออกแบบ และปรับมุมมองของการออกแบบการทดลองสามารถแยกแยะระหว่างประเด็นทางวิทยาศาสตร์ และไม่ใช่วิทยาศาสตร์	- สามารถระบุ ข้อมูลที่สนับสนุนคำตอบเพื่อเลือกคำอธิบายที่เกี่ยวข้อง - สามารถตีความและแปลงข้อมูลเพื่อสนับสนุนคำตอบหรือข้อสรุปทางวิทยาศาสตร์
1 ต่ำ (Low Level)	- รับรู้ศัพท์ทางวิทยาศาสตร์ และใช้ข้อเท็จจริงที่ใกล้เคียงกับประสบการณ์ส่วนตัว - เลือกตัวอย่างที่เหมาะสมของโมเดลทางวิทยาศาสตร์อย่างง่ายหรือระบุคำอธิบายทางวิทยาศาสตร์ที่เหมาะสมซึ่งสอดคล้องกับข้อมูลที่ได้รับ	- สามารถทำตามคำแนะนำขั้นตอนทางวิทยาศาสตร์ - ทำการทดลองง่าย ๆ ที่เกี่ยวข้องกับตัวแปรต้นหนึ่งตัว และตัวแปรตามหนึ่งตัว	- ระบุนข้อมูลในรูปแบบอย่างง่าย - สามารถทำการตีความชุดข้อมูลอย่างง่าย

ภาพ 2 แผนที่โครงสร้างเพื่อประเมินสมรรถนะทางวิทยาศาสตร์

ที่มา Intasoi (2020)

2) วิเคราะห์โครงสร้างของข้อสอบ รูปแบบแนวทางการให้คะแนนในแต่ละข้อตามระดับแผนที่โครงสร้าง ข้อสอบมีลักษณะเป็นพหุมิติระหว่างข้อคำถาม (between items multidimensionality) ตามโมเดลการตอบสนองข้อสอบแบบพหุมิติ (MRCMLM) แบ่งเป็น 3 รูปแบบ ได้แก่ (1) แบบเชิงซ้อน จำนวน 8 ข้อ (2) แบบเลือกตอบ จำนวน 6 ข้อ และ (3) แบบเขียนตอบ จำนวน 8 ข้อ

3) กำหนดจุดตัดด้วยการกำหนดเกณฑ์พื้นที่บน wright map โปรแกรม acer conquest version 2.0 (Wu et al., 2007) และประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบด้วยวิธี MLE พร้อมนำเสนอเป็น wright map ดังภาพ 3 การกำหนดเกณฑ์พื้นที่บน wright map พิจารณาจากค่าเฉลี่ยของความยากหรือระดับ threshold เดียวกันหารด้วยจำนวนข้อสอบและแยกตามมิติที่ศึกษา ผลการคำนวณจุดตัดปรากฏดังตาราง 1

ตาราง 1 ค่าความยากและค่า threshold ของข้อคำถาม ในแต่ละมิติเพื่อนำไปสู่การกำหนดจุดตัด

มิติด้านการอธิบายปรากฏการณ์ ในเชิงวิทยาศาสตร์					มิติด้านการประเมินและออกแบบกระบวนการ สืบเสาะหาความรู้ทางวิทยาศาสตร์					มิติด้านการแปลความหมายข้อมูลและการใช้ ประจักษ์พยานในเชิงวิทยาศาสตร์				
ข้อ	ค่าความยาก (b)	Threshold			ข้อ	ค่าความ ยาก (b)	Threshold			ข้อ	ค่าความยาก (b)	Threshold		
		1	2	3			1	2	3			1	2	3
1	0.932			0.93	5	-0.094	-0.09			8	1.530		0.08	2.98
2	0.053		0.05		7	0.099		0.10		10	-0.605	-0.60		
3	-0.722	-0.72			9	0.864	-0.08	0.98	1.72	11	-1.014	-1.02		
4	-0.914	-0.91			16	0.531		0.53		12	-1.252	-1.25		
6	0.096	-0.22	0.41							13	-0.333	-0.34		
15	-0.114	-0.12								14	-0.713	-0.71		
17	1.017			1.02						20	-0.114		0.77	
18	0.748			0.75						21	0.549		0.55	
19	0.623			0.63										
22	0.348		0.35											
ค่าเฉลี่ย		-0.49	0.27	0.83			-0.09	0.54	1.72			-0.78	0.47	2.98

5) นำคะแนนจุดตัดที่กำหนดในพื้นที่บน wright map และเปลี่ยนให้เป็นคะแนนสเกล (θ_d) เพื่อใช้ในการนำไปใช้ โดยใช้สูตร คะแนนสเกล = $50 + 10(\theta_d)$ เมื่อ θ_d แทนค่าความสามารถของผู้สอบซึ่งประมาณได้จากโมเดล MRCML ในมิติที่ทำการวินิจฉัย กำหนดให้มีค่าเฉลี่ยเท่ากับ 50 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 ซึ่งคะแนนสเกลจะใช้ในการตีความเพื่อนำไปสู่ผลการวินิจฉัยได้อย่างมีความหมาย (Junpeng, 2018)

6) วิเคราะห์ข้อคำถามและตีความระดับมโนทัศน์ในแต่ละมิติและกำหนดตัวแปรสำหรับสร้างสมการทำนายจากข้อมูลพื้นฐานของนักเรียน

ระยะที่ 2 พัฒนาวิธีการตรวจให้คะแนนแบบอัตโนมัติผ่านการเรียนรู้ของเครื่อง

ในระยะนี้เป็นการออกแบบเพื่อพัฒนานวัตกรรมการตรวจให้คะแนนอัตโนมัติ ระหว่างผู้วิจัยและทีมเทคโนโลยี โดยใช้ผลการวิเคราะห์ในระยะที่ 1 มาเป็นข้อมูลพื้นฐานในการออกแบบวิธีการตรวจให้คะแนนอัตโนมัติดังนี้

1. ผู้ให้ข้อมูล

ผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา ผู้เชี่ยวชาญในด้านการใช้เทคโนโลยี วิศวกรคอมพิวเตอร์และครูวิทยาศาสตร์ชั้นมัธยมศึกษาปีที่ 1 จำนวน 5 คน เพื่อทำการสนทนากลุ่ม (focus group)

2. เครื่องมือ

แบบประเมินแบบสำรวจรายการตามมาตรฐานการประเมินคุณภาพของนวัตกรรมต้นแบบตาม Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2010; Varasunun, 2011)

3. การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูล

นำผลการกำหนดจุดตัด ระดับมโนทัศน์ที่คลาดเคลื่อน และตัวแปรที่ใช้ในการสร้างสมการทำนาย ใช้ออกแบบนวัตกรรมต้นแบบที่ใช้ในการตรวจให้คะแนนอัตโนมัติผ่านการเรียนรู้ของเครื่อง พิจารณาคุณภาพของนวัตกรรมต้นแบบที่พัฒนาขึ้นก่อนนำไปทดลองใช้โดยผู้เชี่ยวชาญที่เกี่ยวข้อง และร่วมกันสะท้อนความคิดและปรับปรุงก่อนนำไปทดลองใช้ในลำดับต่อไป

ผลการวิจัย

1. ผลการวิเคราะห์ผลการตอบของนักเรียนสู่การออกแบบนวัตกรรมต้นแบบ ประกอบด้วยรูปแบบและระดับของมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ การกำหนดจุดตัดจากผลการตอบแบบวินิจฉัยสมรรถนะวิทยาศาสตร์ และการกำหนดตัวแปรเพื่อสร้างสมการทำนายจากข้อมูลพื้นฐานของนักเรียน

1.1 ผลการวิเคราะห์รูปแบบและระดับมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์

จากการวิเคราะห์ข้อมูลหัตถ์ภูมิสามารถจำแนกรูปแบบของมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ ได้ 5 รูปแบบตามลักษณะข้อคำถามจากแบบทดสอบจำนวน 22 ข้อ โดยผลการวิเคราะห์รูปแบบมโนทัศน์ที่คลาดเคลื่อน แสดงในตาราง 2

ตาราง 2 มโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์

รูปแบบมโนทัศน์ที่คลาดเคลื่อน	ลักษณะของข้อคำถาม	เกณฑ์การตอบ	ข้อคำถาม
อุปาทาน (preconceived notion)	ข้อคำถามที่รูปแบบของคำตอบต้องอาศัยความคิดเห็นส่วนบุคคลควบคู่กับพื้นฐานความรู้วิทยาศาสตร์	คำตอบไม่ถูกต้องเพราะเกิดจากการนึกคิดและใช้ความเห็นส่วนตัวโดยไม่มีหลักฐานเพียงพอ	1,2,3,4,5,6,7,8,9,10,11,12,13,14, 15,16,17,18, 19,20 และ 21
ความเชื่อที่ไม่เป็นวิทยาศาสตร์ (nonscientific belief)	ข้อคำถามที่รูปแบบของคำตอบต้องสอดคล้องกับหลักการและทฤษฎีทางวิทยาศาสตร์ และอาศัยแหล่งข้อมูลมาพิจารณาก่อนตอบคำถาม โดยแหล่งข้อมูลนั้นคือ ตาราง แผนภูมิ รูปภาพ และผลการทดลอง	คำตอบไม่ถูกต้องตามกฎและหลักการที่เป็นบรรทัดฐานทางวิทยาศาสตร์	ทุกข้อคำถาม
ความเข้าใจผิดเกี่ยวกับมโนทัศน์ (conceptual misunderstanding)	ข้อคำถามที่รูปแบบของคำตอบต้องใช้ความรู้ที่ถูกต้องเท่านั้น ประกอบด้วยเนื้อหา ดังนี้ (1) สมบัติของธาตุ (2) โลหะ กึ่งโลหะ อโลหะ (3) ความดันบรรยากาศ (4) อุณหภูมิของสาร (5) ความหนาแน่น (6) สารเนื้อเดียว (7) โครงสร้างอะตอมของธาตุ (8) อุณหภูมิกับการเปลี่ยนสถานะของสาร และ (9) ความรู้ทั่วไป	1.คำตอบไม่ถูกต้องตามหลักและเนื้อหาที่ใช้ตั้งข้อคำถาม 2.อธิบายคำตอบได้แต่การอธิบายนั้นแสดงความคลาดเคลื่อนจากการเรียนรู้ที่ผ่านมา 3.ลักษณะของคำตอบกำกวมและนักเรียนไม่มั่นใจในคำตอบ	ทุกข้อคำถาม
มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา (vernacular misconception)	ข้อคำถามที่ต้องใช้การแปลความหมายจาก (1) สูตรเคมี (2) ตัวเลข (3) การคำนวณ (4) อุปกรณ์วิทยาศาสตร์ (5) สัญลักษณ์นิวเคลียร์ (6) แบบจำลองอนุภาคของสาร และ (7) บทความ	1.คำตอบแสดงความคลาดเคลื่อนจากการแปลความหมาย หรือตีความ จากลักษณะของข้อคำถาม	1,2,3,4,5,6,7,8,9,10,11,12,13,14, 15,16,17,18, 19,20 และ 21
มโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับข้อเท็จจริง (factual misconception)	ข้อคำถามที่ต้องวิเคราะห์ข้อเท็จจริงจากบทเรียนเพื่อให้ได้คำตอบที่ถูกต้อง โดยพิจารณาจากข้อมูลตามรูปภาพ ตารางและสถานการณ์ที่กำหนด	1.อธิบายไร้เหตุผล 2.อธิบายคำตอบโดยแสดงความเข้าใจที่ผิดจากการเรียนรู้ที่ผ่านมา	1,2,3,4,6,8,9,11,12,13,14,15,16,17,18 และ 19

จากตาราง 2 พบรูปแบบความเชื่อที่ไม่เป็นวิทยาศาสตร์ (nonscientific belief) และความเข้าใจผิดเกี่ยวกับมโนทัศน์ (conceptual misunderstanding) เกิดในทุกข้อคำถาม โดยลักษณะของข้อคำถามนั้นมักเป็นข้อคำถามที่ต้องหาคำตอบโดยสามารถอธิบายจากหลักการและทฤษฎีทางวิทยาศาสตร์ได้ โดยการใช้ความรู้ที่ถูกต้องเท่านั้นในการตอบคำถาม และจากการวิเคราะห์รูปแบบของมโนทัศน์ที่คลาดเคลื่อนที่เกิดขึ้นสามารถจัดระดับมโนทัศน์ที่คลาดเคลื่อนจากการวิเคราะห์ผลการตอบของนักเรียนได้ โดยอิงเกณฑ์การจัดระดับตามคำอธิบายของแผนที่โครงสร้างทั้ง 4 ระดับ คลอบคลุมสมรรถนะทางวิทยาศาสตร์ใน 3 มิติ (Intasoi, 2020) ระดับมโนทัศน์ที่คลาดเคลื่อนแสดงในตาราง 3

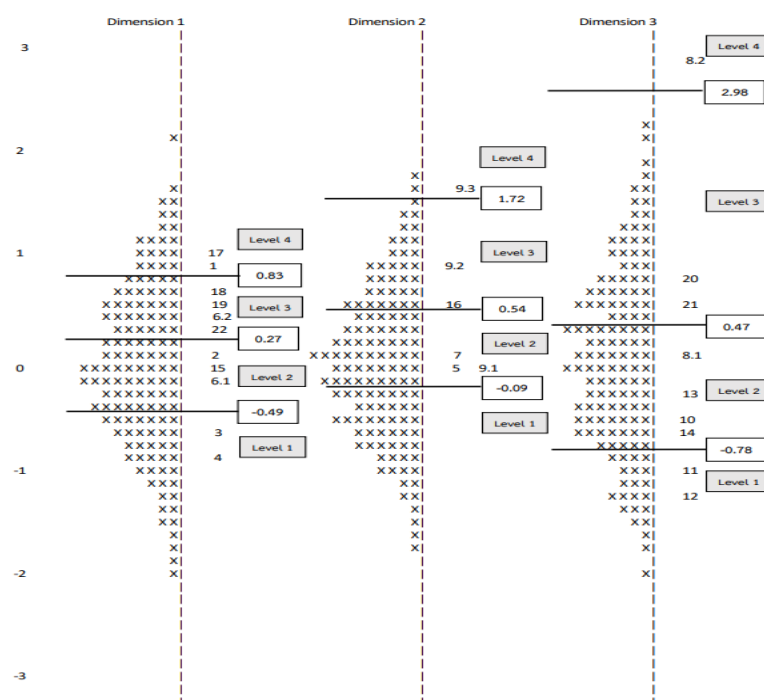
ตาราง 3 ระดับมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์

สมรรถนะทาง วิทยาศาสตร์	มโนทัศน์ที่คลาดเคลื่อน		
	มิติที่ 1	มิติที่ 2	มิติที่ 3
ระดับสูง (High Level)	-	-	-
ระดับปานกลาง (Intermediate Level)	- ความเชื่อที่ไม่เป็นวิทยาศาสตร์ - ความเข้าใจผิดเกี่ยวกับมโนทัศน์	- ความเข้าใจผิดเกี่ยวกับมโนทัศน์ - มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา	- อุปาทาน - ความเข้าใจผิดเกี่ยวกับมโนทัศน์ - มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา
ระดับพื้นฐาน (Basic Level)	- อุปาทาน - ความเชื่อที่ไม่เป็นวิทยาศาสตร์ - มโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับข้อเท็จจริง	- อุปาทาน - ความเข้าใจผิดเกี่ยวกับมโนทัศน์ - มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา - มโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับข้อเท็จจริง	- ความเชื่อที่ไม่เป็นวิทยาศาสตร์ - ความเข้าใจผิดเกี่ยวกับมโนทัศน์ - มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา - มโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับข้อเท็จจริง
ระดับต่ำ (Low Level)	- อุปาทาน - ความเข้าใจผิดเกี่ยวกับมโนทัศน์ - ความเชื่อที่ไม่เป็นวิทยาศาสตร์ - มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา - มโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับข้อเท็จจริง	- อุปาทาน - ความเข้าใจผิดเกี่ยวกับมโนทัศน์ - ความเชื่อที่ไม่เป็นวิทยาศาสตร์ - มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา	- อุปาทาน - ความเข้าใจผิดเกี่ยวกับมโนทัศน์ - ความเชื่อที่ไม่เป็นวิทยาศาสตร์ - มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา

จากตาราง 3 พบว่ารูปแบบมโนทัศน์ที่คลาดเคลื่อนที่พบมากในสมรรถนะระดับปานกลาง คือ ความเข้าใจผิดเกี่ยวกับมโนทัศน์ และระดับพื้นฐาน คือ มโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับข้อเท็จจริง ส่วนรูปแบบมโนทัศน์ที่คลาดเคลื่อนที่ในระดับสมรรถนะต่ำมักพบ 4 รูปแบบ คือ อุปาทาน ความเข้าใจผิดเกี่ยวกับมโนทัศน์ ความเชื่อที่ไม่เป็นวิทยาศาสตร์ และมโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา

1.2 ผลการวิเคราะห์ผลการตอบของผู้เรียนเพื่อพัฒนาและกำหนดจุดตัดในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์

ผลการกำหนดจุดตัดบน wright map ของสมรรถนะทางวิทยาศาสตร์ทั้ง 3 มิติ แบ่งจุดตัดได้ 4 ระดับสมรรถนะ ดังแสดงในภาพ 3



ภาพ 3 จุดตัดบน wright map ของสมรรถนะทางวิทยาศาสตร์ ในมิติการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์

(ซ้าย) มิติการประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ (กลาง) และมิติการแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ (ขวา)

จากภาพ 3 พบว่ามีมิติด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์แบ่งได้ 4 ระดับ 3 จุดตัด คือ -0.49 0.27 และ 0.83 ตามลำดับ มิติด้านการประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์แบ่งได้ 4 ระดับ 3 จุดตัด คือ -0.09 0.54 และ 1.72 ตามลำดับ และมิติด้านการแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์แบ่งได้ 4 ระดับ 3 จุดตัด คือ -0.78 0.47 และ 2.98 ตามลำดับ จาก 3 จุดตัดในแต่ละมิติดังกล่าวข้างต้น นำไปสู่การกำหนดช่วงการประเมินระดับสมรรถนะทางวิทยาศาสตร์ในแต่ละมิติ เป็น 4 ช่วง ที่สามารถแปลงเป็นช่วงคะแนนสเกล และช่วงคะแนนดิบ เพื่อนำไปใช้ในการวินิจฉัยผู้เรียนเป็นรายบุคคล ได้ดังตาราง 4 เช่น ผู้สอบที่ประมาณค่าพารามิเตอร์ระดับความสามารถด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ตั้งแต่ 0.83 ($\theta \geq 0.83$) แสดงว่า ผู้สอบจะได้รับการวินิจฉัยระดับสมรรถนะทางวิทยาศาสตร์อยู่ในระดับที่ 4 โดยมีคะแนนสเกลเท่ากับ 58.30 ขึ้นไป และเมื่อนำไปเทียบกับคะแนนดิบจะมีค่าเท่ากับ 8 คะแนน จากคะแนนเต็ม 11 คะแนน เป็นต้น ดังตาราง 4

ตาราง 4 เกณฑ์จุดเปลี่ยนผ่านระดับสมรรถนะทางวิทยาศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1

มิติการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์					
ระดับ	ชื่อระดับ	จุดตัด	ช่วงระดับสมรรถนะทางวิทยาศาสตร์ (θ)	คะแนนสเกล	คะแนนดิบ (11 คะแนน)
4	ระดับสูง	$\theta = 0.83$	$\theta = 0.83$ ขึ้นไป	58.30 ขึ้นไป	8-11
3	ระดับปานกลาง	$\theta = 0.27$	$\theta = 0.27$ ถึง 0.82	52.70 ถึง 58.29	6-7
2	ระดับพื้นฐาน	$\theta = -0.49$	$\theta = -0.49$ ถึง 0.26	45.10 ถึง 52.69	4-5
1	ระดับต่ำ		ต่ำกว่า $\theta = -0.49$	ต่ำกว่า 45.10	0-3
มิติการประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์					
ระดับ	ชื่อระดับ	จุดตัด	ช่วงระดับสมรรถนะทางวิทยาศาสตร์ (θ)	คะแนนสเกล	คะแนนดิบ (6 คะแนน)
4	ระดับสูง	$\theta = 1.72$	$\theta = 1.72$ ขึ้นไป	67.20 ขึ้นไป	5-6
3	ระดับปานกลาง	$\theta = 0.54$	$\theta = 0.01$ ถึง 1.71	55.40 ถึง 67.19	3-4
2	ระดับพื้นฐาน	$\theta = -0.09$	$\theta = -0.09$ ถึง 0.00	49.10 ถึง 55.39	2
1	ระดับต่ำ		ต่ำกว่า $\theta = -0.09$	ต่ำกว่า 49.10	0-1
มิติการแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์					
ระดับ	ชื่อระดับ	จุดตัด	ช่วงระดับสมรรถนะทางวิทยาศาสตร์ (θ)	คะแนนสเกล	คะแนนดิบ (9 คะแนน)
4	ระดับสูง	$\theta = 2.98$	$\theta = 2.98$ ขึ้นไป	79.80 ขึ้นไป	+9
3	ระดับปานกลาง	$\theta = 0.47$	$\theta = 0.47$ ถึง 2.97	54.70 ถึง 79.79	6-8
2	ระดับพื้นฐาน	$\theta = -0.78$	$\theta = -0.78$ ถึง 0.46	42.20 ถึง 54.69	4-5
1	ระดับต่ำ		ต่ำกว่า $\theta = -0.78$	ต่ำกว่า 42.20	0-3

2.2 ผลการวิเคราะห์ตัวแปรที่จะใช้ในการสร้างสมการทำนายจากข้อมูลชุดเดียว

การวิเคราะห์ข้อมูลพื้นฐานของนักเรียนเพื่อทำนาย (prognosis) ระดับสมรรถนะทางวิทยาศาสตร์ต้องกำหนดตัวแปรเพื่อสร้างสมการทำนายโดยการพิจารณาถึงตัวแปรที่จะส่งผลต่อระดับสมรรถนะทางวิทยาศาสตร์ ซึ่งได้กำหนดตัวแปรที่เหมาะสมดังตาราง 5

ตาราง 5 ตัวแปรอิสระ ตัวแปรตาม คำนิยามและสัญลักษณ์ของตัวแปร

ชนิดของตัวแปร	ข้อมูลทฤษฎี	คำนิยาม	สัญลักษณ์ตัวแปร
ตัวแปรต้น	เกรดเฉลี่ยรวมเทอมที่ผ่านมา	เกรดเฉลี่ยรวมทุกวิชาของนักเรียนจากตลอดเทอมการศึกษาที่ผ่านมา	X_1
	เกรดเฉลี่ยวิชาวิทยาศาสตร์เทอมที่ผ่านมา	เกรดเฉลี่ยวิชาวิทยาศาสตร์ของนักเรียนจากตลอดเทอมการศึกษาที่ผ่านมา โดยมีระดับเกรดจากที่โรงเรียนกำหนด	X_2
	จำนวนชั่วโมงที่ศึกษาวิชาวิทยาศาสตร์ด้วยตนเอง/สัปดาห์	จำนวนชั่วโมงที่นักเรียนใช้ในการศึกษาวิชาวิทยาศาสตร์ด้วยตนเอง ซึ่งมีเกณฑ์ดังนี้ (1) น้อยกว่า 1 ชั่วโมง (2) 1-2 ชั่วโมง (3) 2-3 ชั่วโมง (4) 3-4 ชั่วโมง (5) 4-5 ชั่วโมง (6) ตั้งแต่ 5 ชั่วโมงขึ้นไป	X_3
ตัวแปรตาม	คะแนนมิติ 1	มิติการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์	Y_1
ตัวแปรตาม	คะแนนมิติ 2	มิติการประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์	Y_2
ตัวแปรตาม	คะแนนมิติ 3	มิติการแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์	Y_3

จากตาราง 5 พบว่าตัวแปรต้นที่จะส่งผลต่อการสร้างสมการทำนาย ได้แก่ เกรดเฉลี่ยรวมเทอมที่ผ่านมา เกรดเฉลี่ยวิชาวิทยาศาสตร์เทอมที่ผ่านมา และจำนวนชั่วโมงที่ศึกษาวิชาวิทยาศาสตร์ด้วยตนเอง/สัปดาห์ ซึ่งการกำหนดตัวแปรขึ้นเพื่อสร้างสมการทำนายนี้จะช่วยสนับสนุนและยืนยันแนวโน้มของผลการประเมินสมรรถนะทางวิทยาศาสตร์ผ่านแบบทดสอบวินิจฉัยได้เป็นอย่างดี

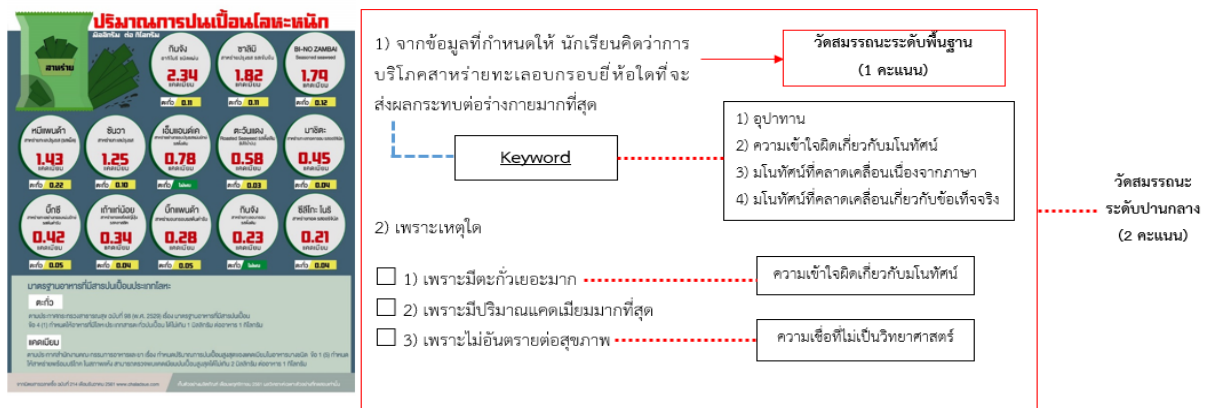
2. ผลการออกแบบวิธีการตรวจให้คะแนนอัตโนมัติในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์ผ่านการเรียนรู้ของเครื่อง

จากผลการวิเคราะห์รูปแบบ ระดับมโนทัศน์ที่คลาดเคลื่อน การกำหนดจุดตัดและการกำหนดตัวแปรข้อมูลพื้นฐานของนักเรียนเพื่อการสร้างสมการทำนายนั้น นำไปสู่การออกแบบการตรวจให้คะแนนอัตโนมัติผ่านการเรียนรู้ของเครื่องเพื่อตีความระดับสมรรถนะวิทยาศาสตร์ โดยนำเสนอผลการออกแบบดังนี้

2.1 ผลการออกแบบวิธีการตรวจให้คะแนนผ่านการเรียนรู้ของเครื่อง

ผลการออกแบบรูปแบบการตรวจให้คะแนนใช้โมเดลเชิงโครงสร้าง (construct modeling) เป็นแนวทางในการพัฒนาการตรวจให้คะแนน โดยมีส่วนข้อมูลป้อนกลับแก่ผู้เรียนคือระดับมโนทัศน์ที่คลาดเคลื่อน และส่วนรายงานผลการประเมินคือระดับสมรรถนะวิทยาศาสตร์ของนักเรียน ตัวอย่างการตรวจแสดงในภาพ 4

6) ศูนย์ทดสอบฉลาดซื้อ มูลนิธิเพื่อผู้บริโภค (มพบ.) เผยผลทดสอบ สหรัยทะเลบกรอบ ดังภาพ



ภาพ 4 รูปแบบการตรวจให้คะแนนผ่านระบบ

จากภาพ 4 ตัวอย่างข้อสอบเป็นรูปแบบอัตนัยที่วัดสมรรถนะระดับปานกลาง (2 คะแนน) ในการออกแบบการตรวจให้คะแนนได้ออกแบบให้มีการตรวจใน 2 ลักษณะ คือ (1) ใช้วิธีการตรวจจับคำ (keyword) และ (2) ใช้วิธีการแปลงคำตอบจากข้อมูลหัตถ์เขียนให้เป็นตัวเลือก และจากการออกแบบวิธีการตรวจให้คะแนนนี้ นำไปสู่การร่างระบบตรวจให้คะแนน ดังภาพ 5



ภาพ 5 ระบบการตรวจให้คะแนนอัตโนมัติผ่านการเรียนรู้ของเครื่อง

รูปแบบระบบการตรวจให้คะแนนนั้นประกอบด้วย 5 ส่วนสำคัญ คือ (1) ส่วนข้อมูลนำเข้า ประกอบด้วยช่วงคะแนนดิบและคะแนนสเกล คำอธิบายระดับความสามารถกำหนดเป็นเกณฑ์การให้คะแนน (2) ส่วนประมวลผล เน้นไปที่กระบวนการทำงานของอัลกอริทึมในการตรวจให้คะแนนในแบบสอบปรนัยและอัตนัย (3) ส่วนแสดงผล จะรายงานระดับสมรรถนะวิทยาศาสตร์รวมถึงระดับมโนทัศน์ที่คลาดเคลื่อนที่เกิดขึ้นจากการเรียนรู้ (4) ส่วนให้ข้อมูลป้อนกลับแก่ผู้เรียน จะให้ข้อเสนอแนะเพื่อให้พัฒนาตนเองและรายงานระดับมโนทัศน์ที่คลาดเคลื่อน และ (5) ส่วนรายงานผลการประเมิน เป็นนำผลการประเมินมารายงานให้ผู้ใช้งานระบบเข้าใจได้ง่ายมากยิ่งขึ้น

การประเมินคุณภาพเบื้องต้นของผลการออกแบบการตรวจให้คะแนนอัตโนมัติ ด้วยวิธีการสนทนากลุ่ม (focus group) ได้ผลการประเมินเชิงคุณภาพจากผู้เชี่ยวชาญจำนวน 5 คน ซึ่งผู้เชี่ยวชาญได้ให้ข้อเสนอแนะในการออกแบบระบบ โดยครอบคลุม 5 มาตรฐาน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสมชอบธรรม ด้านความถูกต้อง และด้านความรับผิดชอบการประเมิน และจากการสนทนาพบว่า การออกแบบเพื่อสร้างนวัตกรรมการตรวจให้คะแนนอัตโนมัติเป็นเครื่องมือวัดที่ดีเพราะเป็นการนำเทคโนโลยีสมัยใหม่เข้ามาช่วยในการออกแบบเครื่องมือวัดที่เหมาะสมกับสถานการณ์ในยุคปัจจุบัน เนื่องจากสามารถตรวจให้คะแนนแทนคนตรวจได้และรายงานผลการเรียนรู้ให้แก่ผู้เรียนได้แบบเรียลไทม์

อภิปรายผล

1. ผลวิเคราะห์รูปแบบมโนทัศน์ที่คลาดเคลื่อนทางสมรรถนะวิทยาศาสตร์จากข้อมูลหัตถ์เขียนในเนื้อหาเรื่อง สารบริสุทธิ์ พบรูปแบบมโนทัศน์ที่คลาดเคลื่อนขึ้น 5 รูปแบบ คือ อุปาทาน ความเชื่อที่ไม่เป็นวิทยาศาสตร์ ความเข้าใจผิดเกี่ยวกับมโนทัศน์ มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา และมโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับข้อเท็จจริง ทั้งนี้อาจเนื่องมาจากการประเมินสมรรถนะวิทยาศาสตร์ภายใต้บทเรียนเรื่อง สารบริสุทธิ์ นั้นต้องพิจารณาข้อมูลต่างๆ เช่น รูปภาพ กราฟ และสถานการณ์ร่วมด้วย ประกอบกับการทำแบบทดสอบต้องใช้ความเข้าใจที่แท้จริงเพื่อหาคำตอบเป็นเหตุให้นักเรียนที่ไม่มีพื้นฐานหรือ

เข้าใจไม่ลึกซึ้งเกิดแนวความคิดคลาดเคลื่อน ไม่สามารถสรุปข้อเท็จจริงในเนื้อหาสาระได้ซึ่งสอดคล้องกับสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยีที่ระบุว่านักเรียนมีแนวความคิดคลาดเคลื่อนเนื่องจากไม่มีความรู้เดิมหรือความรู้พื้นฐานทำให้เกิดมโนทัศน์ที่คลาดเคลื่อนในบทเรียนเรื่องสารบริสุทธิ์ (Institute for the promotion of technology science and technology, 2017) และจากผลการวิเคราะห์รูปแบบมโนทัศน์ที่เกิดสอดคล้องกับแนวคิดมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ของสภาวิจัยแห่งชาติ National Research Council (1997) และการศึกษาของ Kamtet (2017)

จากการศึกษาพบว่าเกิดระดับมโนทัศน์ที่คลาดเคลื่อนตามระดับสมรรถนะทางวิทยาศาสตร์แบ่งออกเป็น 4 ระดับ ได้แก่ สูง ปานกลาง พื้นฐาน และต่ำ ซึ่งพบว่าสอดคล้องกับงานวิจัยของ Intasoi (2020) โดยในระดับสมรรถนะปานกลางและพื้นฐาน พบรูปแบบมโนทัศน์ที่คลาดเคลื่อนที่พบมากที่สุดคือความเข้าใจผิดเกี่ยวกับมโนทัศน์ และมโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับข้อเท็จจริง ตามลำดับ ซึ่งสาเหตุดังกล่าวมาจากการที่นักเรียนไม่มีความรู้เบื้องต้นหรือขาดการคิดเชื่อมโยง และการให้เหตุผลที่ไม่สมบูรณ์ รวมถึงขาดความสนใจในการเรียนวิทยาศาสตร์ ซึ่งผลเหล่านี้แสดงให้เห็นว่ามโนทัศน์เป็นอุปสรรคต่อการเรียนรู้และเป็นปัจจัยสำคัญที่ขัดขวางไม่ให้เกิดการเรียนรู้วิทยาศาสตร์อย่างถาวร การรายงานรูปแบบและระดับมโนทัศน์แก่นักเรียนช่วยให้นักเรียนพัฒนาได้ตรงจุดมากขึ้นเพราะการเกิดมโนทัศน์นั้นแสดงว่านักเรียนมีความรู้บางส่วนแต่เกิดการคลาดเคลื่อนระหว่างการเรียนรู้ ดังนั้นนักเรียนต้องได้รับข้อมูลย้อนกลับที่แสดงถึงจุดบกพร่องที่ต้องพัฒนาหลังจากการเรียนรู้ สอดคล้องกับ Anuntasawat (2017) พบว่าการรายงานผลพร้อมให้ข้อมูลย้อนกลับถึงจุดเด่นจุดด้อยทำให้นักเรียนทราบความสามารถของตนเอง ช่วยเสริมแรงบันดาลใจในการเรียน และทราบจุดที่ควรพัฒนาไปสู่เป้าหมายที่ตั้งไว้

2. จากผลการวิจัยพบว่าจุดตัดที่ได้จากการใช้ข้อมูลชุดข้อมูลมีส่งผลต่อการประมาณค่าในการตรวจให้คะแนน เพราะแสดงค่าความสามารถของนักเรียนในแต่ละมิติและกระจายได้ครอบคลุมทุกระดับความสามารถสอดคล้องกับแผนที่โครงสร้างที่ได้กำหนดไว้ เพราะการศึกษาในครั้งนี้มีการวิเคราะห์ผลการตอบที่ได้มาจากฐานข้อมูลขนาดใหญ่ที่มีตัวแทนในทุกระดับความสามารถของนักเรียน ซึ่งการคัดเลือกกลุ่มผู้สอบนั้นใช้เทคนิคการกำหนดตัวอย่างวิจัยขั้นต่ำของ Wright & Stone (1979) เพื่อให้ได้การกระจายความสามารถของกลุ่มตัวอย่างทำให้สารสนเทศมีคุณภาพ สอดคล้องกับ Junpeng et al. (2019); Junpeng et al. (2020) ที่พบว่าถ้าการกระจายความสามารถของผู้สอบไม่เหมาะสมจะส่งผลต่อการกำหนดจุดตัดเพื่อใช้ในการวินิจฉัยระดับความสามารถของผู้เรียน ส่งผลให้การตีความเพื่อวินิจฉัยขาดความถูกต้องแม่นยำในการประมาณค่า และส่งผลกระทบต่อการพิจารณาในระดับของมโนทัศน์ที่คลาดเคลื่อนที่เกิดขึ้นของผู้เรียน

3. กระบวนการสำคัญที่นำไปสู่การตรวจให้คะแนนอัตโนมัติผ่านการเรียนรู้ของเครื่อง คือการวิเคราะห์รูปแบบและระดับมโนทัศน์ การกำหนดจุดตัด และการกำหนดตัวแปรจากการสังเคราะห์ข้อมูลพื้นฐานเพื่อสร้างสมการทำนายที่เหมาะสมเพราะสมการทำนายจะช่วยทำนายผลการวินิจฉัยระดับความสามารถได้อย่างถูกต้องและแม่นยำ ซึ่งงานวิจัยของ Kladchuen & Sanrach (2018) พบว่า การเลือกใช้ตัวแปรเพื่อสร้างสมการที่เหมาะสมในการทำนายผลสัมฤทธิ์ทางการเรียนของนักศึกษาระดับอาชีวศึกษาช่วยให้การทำนายผลสัมฤทธิ์ทางการเรียนแม่นยำ นอกจากนี้ในข้อสอบอัตนัยได้ออกแบบให้ การตอบอยู่ในลักษณะ

ของการเติมคำ หรือวลีสั้นๆ เพื่อให้ระบบการตรวจที่จะพัฒนาขึ้นตรวจในลักษณะ การจับคำ (keywords) เพราะระบบจะตรวจจับคำได้ง่ายสอดคล้องกับงานวิจัยที่ผ่านมาได้นำการเรียนรู้ของเครื่องมาตรวจจับคำและวิธีการตอบของนักเรียนในวิชาเคมีและฟิสิกส์เพื่อประเมินระดับความสามารถของนักเรียน (Maestrales et al., 2021)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1. ครูผู้สอนสามารถนำรูปแบบโมเดลที่คลาดเคลื่อนเกิดขึ้นทั้ง 5 รูปแบบ ไปวางแผนการจัดการเรียนรู้ที่เน้นการส่งเสริมสมรรถนะทางวิทยาศาสตร์อันครอบคลุมทั้ง 3 ด้าน ได้แก่ ด้านการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และการแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ เพื่อลดและแก้ไขปัญหาการเกิดโมเดลที่คลาดเคลื่อนในสมรรถนะวิทยาศาสตร์ได้ตรงจุด

2. การตรวจให้คะแนนผ่านการเรียนรู้ของเครื่องมีความจำเป็นอย่างยิ่งที่ควรให้ความสำคัญกับการตรวจให้คะแนนตามแผนที่โครงสร้างในแต่ละมิติ โดยเฉพาะอย่างยิ่งในส่วนนำเข้าข้อมูลที่จำเป็นต้องได้จากการวิเคราะห์ข้อมูลทฤษฎีซึ่งเป็นส่วนสำคัญในการนำไปสู่ส่วนการให้ข้อมูลป้อนกลับและส่วนรายงานผลในลำดับถัดไป

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. งานวิจัยนี้เป็นการออกแบบวิธีการตรวจให้คะแนนอัตโนมัติผ่านการเรียนรู้ของเครื่อง โดยในระยะเวลาการศึกษานี้ได้กำหนดตัวแปรไว้สร้างสมการทำนายเพื่อเป็นอัลกอริทึมในการเรียนรู้ของเครื่อง ดังนั้นเพื่อให้งานวิจัยมีความสมบูรณ์มากยิ่งขึ้นขั้นตอนต่อไปควรนำตัวแปรที่กำหนดมาสร้างสมการทำนายเพื่อป้อนสู่ระบบการตรวจให้คะแนนอัตโนมัติผ่านการเรียนรู้ของเครื่อง

2. ในการทำวิจัยครั้งต่อไปเมื่อพัฒนาระบบตรวจให้คะแนนผ่านการเรียนรู้ของเครื่องเรียบร้อยแล้วควรตรวจสอบคุณภาพของระบบก่อนโดยการพิจารณาความเที่ยงของผู้ประเมินระหว่างการตรวจของเครื่องและใช้คนตรวจ เพื่อประเมินประสิทธิภาพในการตรวจให้คะแนนอัตโนมัติและค้นหาข้อบกพร่องเพื่อปรับปรุงระบบต่อไป

References

Adams, R. A., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21(1), 1-23.

Desstya, A., Prasetyo, Z.K., Suyanta., Susila, I., & Irwanto. (2019). Developing an Instrument to Detect Science Misconception of an Elementary School Teacher. *International Journal of Instruction*, 12(3), 201-218.

- Intasoi, S., Jungpeng P., Ngang Tang K., Ketchatturat J., Zhang Y., & Wilson M. (2020). Developing an assessment framework of multidimensional scientific competencies. *International Journal of Evaluation and Research in Education*, 9(4), 963-970.
- Junpeng, P., Krotha, J., Chanayota, K., Tang, K. N., & Wilson, M. (2019). Constructing Progress Maps of Digital Technology for Diagnosing Mathematical Proficiency. *Journal of Education and Learning*, 8(6), 90-102.
- Junpeng, P., Marwiang, M., Chiajunthuk, S., Suwannatrai, P., Chanayota, K., Pongboriboon, K., Tang, K. N., & Wilson, M. (2020). Validation of a digital tool for diagnosing mathematical proficiency. *International Journal of Evaluation and Research in Education*, 9(3), 665-674.
- Lee, H. S., Gweon, G. H., Lord, T., Paessel, N., Pallant, A., & Pryputniewicz, S. (2021). Machine Learning-Enabled Automated Feedback: Supporting Students' Revision of Scientific Arguments Based on Data Drawn from Simulation. *Journal of Science Education and Technology*, 30(2), 168-192.
- Lintean, M., Vasile, R., & Roger, A. (2012). Automatic Detection of Student Mental Models Based on Natural Language Student Input during Metacognitive Skill Training. *International Journal of Artificial Intelligence in Education*, 21(3), 169-190.
- Maestrales, S., Zhai, X., Touitou, I., Baker, Q., Schneider, B., & Krajcik, J. (2021). Using Machine Learning to Score Multi-Dimensional Assessments of Chemistry and Physics. *Journal of Science Education and Technology*, 30(2), 239-254.
- National Research Council. (1997). *Science teaching reconsidered: A handbook*. National Academies Press.
- Organisation for Economic Co-operation and Development. (2019). *PISA 2018 Assessment and Analytical Framework*. PISA, OECD Publishing.
- Reeves, T. (2006). *Design research from a technology perspective*. In J.V.D. Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.). *Educational design research* (pp. 52–66). Routledge.
- Varasunun, P. (2011). The Program Evaluation Standards (3rd Ed.) in 2010. *Journal of Research Methodology*, 24(2), 273-278.
- Wilson, M. (2005). *Constructing Measures: an item response modeling approach*. Lawrence Erlbaum Associates.
- Wright, B.D., & Stone, M. H. (1979). *Best Test Design: Rasch Measurement*. Mesa Press.

Wu, M. L., Adums, R., Wilson, M., & Haldane, S. A. (2007). *ConQuest Version 2: Generalised Item Response Modelling Software*. ACER Press.

Yarbrough, D.B., Shulha, L.M., Hopson, R.K., and Caruthers, F.A. (2010). *The program Evaluation standards: A guide for evaluators and evaluation users (3rd Ed.)*. Sage.

Translate Thai References

Anuntasawat, S. (2017). *Development of a Misconception Diagnostic system in Using a Three-Tier Diagnostic Test with Computer Based Reflective Feedback for Tenth Grade students* [Doctoral dissertation]. Chulalongkorn University. (in Thai)

Chianchana, C. (2009). Multidimensional Analysis. *Journal of Education Khon Kaen University*, 32(4), 13-22. (in Thai)

Insawat, M. (2016). *A comparison of the quality of three-tier diagnostic test of mathematical misconception using different levels of confidence* [Master's thesis]. Chulalongkorn University. (in Thai)

Institute for the promotion of technology science and technology. (2017). *Teacher Guide Basic Science Volume 1*. Bangkok: office of the welfare promotion commission for teachers and educational personnel press. (in Thai)

Intasoi, S. (2020). *Developing a digital tool for assessing multidimensional scientific competencies of the seventh-grade students* [Master's thesis]. Khon Kaen University. (in Thai)

Junpeng, P. (2018). *Applications of the multidimensional item response theory of research* (4th Ed.). Khon Kaen University press. (in Thai)

Kamtet, W. (2017). Scientific Misconception: Type and Assessment Tool. *STOU Education Journal*. 10(2), 54-64. (in Thai)

Khawsuk, J., & Srihasat, K. (2016). The Development of a Mathematics Diagnostic Test for Polynomials at the Matthayom Sueksa One Level. *Veridian E-Journal, Silpakorn University*, 9(3), 1206-1220. (in Thai)

Kladchuen, R., & Sanrach, J. (2018). An Efficiency Comparison of Algorithms and Feature Selection Methods to Predict the Learning Achievement of Vocational Students. *Research Journal Rajamangala University of Technology Thanyaburi*, 17(1), 1-10. (in Thai)

- Rawanprakhon, P., & Tayraukham, S. (2017). The Development of Multidimensional Essay Test. *Journal of Community Development Research (Humanities and Social Sciences)*, 10(4), 13-23. (in Thai)
- Sudjai, M., & Mungsing, S. (2021). Development of an Automated subjective answer scoring System with full-text search technique and PHP text comparison function *Journal of Information Science and Technology*, 11(1), 8-17. (in Thai)
- Wongphummueang, L., & Piyakun, A. (2021). Development of Multidimensional Test of Collaborative Problem-Solving Skills for Lower Secondary School Students. *Journal of Educational Measurment, Mahasarakham University*, 27(1), 232-243. (in Thai)
- Wongwanit, S. (2020). *Design Research in Education*. Chulalongkorn University press. (in Thai)
- Boonjun, S. (2019). *A Comparison of Differential Item Funtioning for Mixed Format Tests by Tree Dffferentail Methods* [Doctoral dissertation]. Burapha University. (in Thai)
- Fiothong, A., Junpeng, P., Suwannatrai, P., Chinjunthuk, S., & Tawarungruang, C. (2022). Designing Open-Ended Question Scoring for Assessment of student Mathematical Proficiency Levels through Digital Technology. *Journal of Educational Measurement Mahasarakham University*, 28(1), 346-362. (in Thai)

การพัฒนาแบบวัดการตระหนักรู้สถานการณ์ของนักศึกษาวิสัญญีพยาบาล Development of Situation Awareness Scale for Nurse Anesthetist Students

ปาริชาติ อภิเดชากุล¹ ศิริเดช สุชีวะ^{2*} กมลวรรณ ตังธนากานนท์²
พงศ์ธรา วิจิตเวชไพศาล³ วิรันตริ พันชุกกลาง¹ และ ธัชนัน จินตดาวงษ์¹
Parichad Apidechakul¹, Siridej Sujiva^{2*}, Kamonwan Tangthanakanon², Phongthara Vichitvejpaisal¹,
Wiruntri Punchuklang¹ and Thatchanan Jintadawong¹

(Received: March 2, 2022; Revised: April 9, 2022; Accepted: April 11, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อสำรวจประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์และพัฒนาแบบวัดการตระหนักรู้สถานการณ์การระงับความรู้สึกทั่วไปเพื่อการผ่าตัดคลอด ผู้ร่วมวิจัยเป็นผู้ทรงคุณวุฒิสาขาวิสัญญีพยาบาลจำนวน 14 ท่าน และนักศึกษาวิสัญญีพยาบาล 136 คน ผลการสำรวจเดลฟายต่อประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์จำนวน 3 รอบ พบว่า ICC ≥ 0.74 ในรอบที่ 2 และ 3 ประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์มีจำนวน 54 ประเด็น แบบวัดการตระหนักรู้สถานการณ์มีจำนวน 48 ข้อ ข้อมูลที่ได้จากการวัดมีความสอดคล้องกับโมเดล 2PL ผลการประมาณค่าพารามิเตอร์อำนาจจำแนก และพารามิเตอร์ความง่ายส่วนใหญ่อยู่ในเกณฑ์ที่น่าไปใช้ได้ แบบวัดการตระหนักรู้สถานการณ์ที่พัฒนาจากประเด็นที่ควรตระหนักรู้ในสถานการณ์ สามารถนำไปใช้ประเมินการตระหนักรู้สถานการณ์ทั่วไปของนักศึกษาวิสัญญีพยาบาลได้ต่อไป

คำสำคัญ: แบบวัดการตระหนักรู้สถานการณ์ นักศึกษาวิสัญญีพยาบาล

¹ วิสัญญีพยาบาล ภาควิชาวิสัญญีวิทยา คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล

² รองศาสตราจารย์ สาขาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

³ ศาสตราจารย์ ภาควิชาวิสัญญีวิทยา คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล

¹ Nurse Anesthetist, Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Mahidol University

² Associate Professor, Measurement and Evaluation, Faculty of Education, Chulalongkorn University

³ Professor, Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Mahidol University
งานวิจัยนี้ได้รับทุนอุดหนุนการวิจัย จากสำนักงานคณะกรรมการวิจัยแห่งชาติปีงบประมาณ 2564

This research was funded by the National Research Council of Thailand, Fiscal Year 2021.

*Corresponding Author E-mail: ssiridej@chula.ac.th

Abstract

The objectives of this study were to survey the situation awareness requirement items and to develop situation awareness scale in obstetric anesthesia. Fourteen scholars in anesthesiology and 136 nurse anesthetist students joined the study. The situation awareness requirement items were determined by Delphi technique in three consecutive rounds with $ICC \geq 0.74$ for the last two rounds. The situation awareness requirement consisted of 54 items. The situation awareness scale comprised 48 queries and it fit 2PL model. Most queries showed satisfying discrimination and easiness properties. This instrument had been developed from situation awareness requirement items and it can be further used by nurse anesthetist students for situation awareness global assessment technique.

Keywords: Situation Awareness Scale, nurse anesthetist student

บทนำ

สืบเนื่องจากยุทธศาสตร์หลักเพื่อพัฒนาบุคลากรทางสาธารณสุขให้มีศักยภาพและเกิดความตระหนักรู้ในเรื่องความปลอดภัยตามนโยบาย (Patient and Personnel Safety, 2P Safety Goals) นโยบายดังกล่าวจึงถูกนำไปสู่การปฏิบัติ โดยเฉพาะงานทางวิสัญญีวิทยาในการบริหารยาระงับความรู้สึกที่ควรเกิดภาวะแทรกซ้อนต่อผู้ป่วยน้อยที่สุด ราชวิทยาลัยวิสัญญีแห่งประเทศไทยจึงกำหนดวัตถุประสงค์การเรียนการสอนเพื่อบริหารจัดการความเสี่ยงสำหรับผู้เข้ารับการฝึกอบรมวิสัญญีพยาบาล โดยมีจุดเน้นเรื่องทักษะที่ไม่ใช่ทางเทคนิคด้านการตระหนักรู้สถานการณ์ นอกจากความรู้และทักษะเทคนิคทางวิสัญญีแล้ว ผู้เข้ารับการฝึกอบรมวิสัญญีพยาบาลควรมีการตระหนักรู้สถานการณ์ที่ไม่แน่นอนขณะทำการระงับความรู้สึกให้กับผู้ป่วย

การตระหนักรู้สถานการณ์เป็นคุณลักษณะทางปัญญาที่สำคัญของบุคลากรทางวิสัญญี ประกอบด้วย 3 องค์ประกอบคือ การรับรู้ การเข้าใจ และการคาดการณ์ และในเบื้องต้นมี 6 ตัวบ่งชี้ คือ การรวบรวมข้อมูล การตรวจความปกติ การตีความ การสรุปประเด็นปัญหา การคาดคะเนเหตุการณ์ และการคาดคะเนการปฏิบัติ (Sarter & Woods, 1991; Vidulich et al., 1994; Billings, 1995; Endsley, 1995a; Endsley, 1995b; Smith & Hancock, 1995; Goss & Murray, 1996; Endsley, 1988; Endsley & Garland, 2000; Bedny et al., 2004; McKenna et al., 2014) การตระหนักรู้สถานการณ์ถูกนำมาใช้เพื่อรับรู้ต่อสภาพปัญหาที่อาจเกิดขึ้นขณะปฏิบัติงานภายใต้สภาพแวดล้อมที่เป็นพลวัตตามแนวคิดของเอนสลีย์ ซึ่งจะเกิดขึ้นก่อนการตัดสินใจและการลงมือปฏิบัติในบุคลากรทุกระดับ (Wright, 2015; Taylor, 2017) โดยเฉพาะนักศึกษาวิสัญญีพยาบาลที่มีเป้าหมายมุ่งเน้นในกระบวนการคิด การเฝ้าระวัง การตัดสินใจ และการปฏิบัติ เพื่อให้การให้บริการยาระงับความรู้สึกแบบทั่วไปโดยการใส่ท่อช่วยหายใจด้วยตนเองตามมาตรฐานการพยาบาล ซึ่งต้องปฏิบัติงานภายใต้การกำกับดูแลของแพทย์ผู้เชี่ยวชาญ แต่สิ่งนี้อาจเป็นสาเหตุสำคัญที่ทำให้นักศึกษาขาดความ

เชื่อมั่นในการตระหนักรู้ในสถานการณ์เฉพาะหน้าและตัดสินใจด้วยตนเอง โดยนักศึกษาประมาณร้อยละ 70 ไม่ผ่านเกณฑ์การสอบภาคปฏิบัติในสถานการณ์จำลอง (Nurse anesthetist training project, 2019)

อุปสรรคส่วนใหญ่ทางวิสัญญีที่พบได้ในรายงานทั้งในและต่างประเทศมีความเกี่ยวข้องกับระบบทางเดินหายใจของผู้ป่วย (Cheney et al., 2006; Davies et al., 2009; Roh et al., 2015) ภาระงานหลักของวิสัญญีพยาบาลคือการให้การระงับความรู้สึกแบบทั่วไป ซึ่งสถานการณ์การดมยาสลบที่มีระดับความเสี่ยงสูง เช่น การระงับความรู้สึกทั่วไปในผู้ป่วยตั้งครรภ์ที่มารับผ่าตัดคลอด เนื่องจากการตามีสรีรวิทยาที่เปลี่ยนแปลงไปขณะตั้งครรภ์ส่งผลให้เพิ่มความเสี่ยงต่อการใส่ท่อช่วยหายใจยากมากกว่าผู้ป่วยทั่วไปถึง 8 เท่า (Munnur & Suresh 2004; Hawkins et al., 2011) สถานการณ์การผ่าตัดคลอดกรณีเร่งด่วนจากภาวะคุกคามต่อชีวิตของมารดาหรือทารกที่ต้องใช้วิธีการระงับความรู้สึกทั่วไป จึงเป็นสถานการณ์ที่สร้างความท้าทายต่อบุคลากรทางวิสัญญีทุกระดับ

ผลจากการสังเคราะห์เอกสารงานวิจัยและตำราข้อมูลจาก Mushambi (2015) พบว่ามีมีโนทัศน์เบื้องต้นเกี่ยวกับการระงับความรู้สึกแบบทั่วไปแก่ผู้ป่วยผ่าตัดคลอดที่มีจำนวน 10 ประเด็นหลักที่สำคัญใน 3 ระยะ คือระยะเตรียมการ ระยะนำสลบ และระยะใส่ท่อช่วยหายใจ ผู้วิจัยนำประเด็นหลักหรือจุดเน้นที่สำคัญในสถานการณ์ที่ต้องการการตัดสินใจที่ถูกต้องเพื่อให้เกิดการระงับความรู้สึกที่ปลอดภัย บนพื้นฐานของการตระหนักรู้สถานการณ์ที่ดี นำไปสู่การตัดสินใจถูกต้อง และส่งผลให้เกิดการปฏิบัติที่ดีทั้ง 3 องค์ประกอบคือ ด้านการรับรู้ การเข้าใจ และการคาดการณ์ สถานการณ์การระงับความรู้สึกทั่วไปในผู้ป่วยผ่าตัดคลอดมีความเหมาะสมในการนำมาใช้เพื่อการวัดการตระหนักรู้สถานการณ์ในนักศึกษาวิสัญญีพยาบาล เนื่องจากสามารถอ้างอิงกับแนวปฏิบัติสากลที่ได้รับการยอมรับจากองค์กรวิสัญญีชั้นนำระดับโลก

เทคนิคการประเมินการตระหนักรู้สถานการณ์ทางตรงแบบปรนัย situation awareness global assessment technique (SAGAT) ของเอนสลีย์ได้รับการตีพิมพ์และเป็นที่รู้จักอย่างแพร่หลาย (Endsley, 1995a) โดยใช้วิธีวัดการตระหนักรู้สถานการณ์ทางตรง ซึ่งพยายามวัดปริมาณของการตระหนักรู้สถานการณ์ของบุคคลโดยตรง โดยประเมินจากการรับรู้ของแต่ละบุคคลเกี่ยวกับลำดับของสถานการณ์จำลองด้วยข้อคำถามวัดการตระหนักรู้สถานการณ์แบบเลือกตอบ ประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์ได้มาจากการวิเคราะห์แยกแยะจุดเน้นที่อ้างอิงจากแนวปฏิบัติสากลและต้องการการตัดสินใจหรือจากมุมมองของผู้เชี่ยวชาญสาขาวิชาเรียกว่าการวิเคราะห์ระบบงาน (Goal directed task analysis หรือ GDTA) ซึ่งเป็นขั้นตอนที่สำคัญที่สุดของเทคนิคการประเมินการตระหนักรู้สถานการณ์ทางตรงแบบปรนัย (SAGAT) เพื่อช่วยกำหนดขอบเขตของสถานการณ์ที่นำมาใช้ในการประเมิน จากการสำรวจ สัมภาษณ์ หรือเทคนิคเดลฟายต่อความคิดเห็นของผู้เชี่ยวชาญสาขาวิชา (Endsley & Jones, 2004) โดยประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์จะถูกนำไปสร้างเป็นข้อคำถามวัดการตระหนักรู้สถานการณ์ต่อไป

ผู้วิจัยยังไม่พบการวิเคราะห์ระบบงานเพื่อนำไปใช้วัดการตระหนักรู้สถานการณ์และประเมินการตระหนักรู้สถานการณ์ทางตรงแบบปรนัย situation awareness global assessment technique (SAGAT) ในศาสตร์ด้านวิสัญญีวิทยาในประเทศไทย งานวิจัยนี้มีเป้าประสงค์เพื่อค้นหาประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์วิกฤตทางวิสัญญีเพื่อการผ่าตัดคลอด และเพื่อพัฒนาแบบวัดการตระหนักรู้สถานการณ์วิกฤต

ทางวิสัญญีเพื่อการผ่าตัดคลอดในรูปแบบข้อคำถามเชิงสถานการณ์ตามกรอบแนวคิดของเอนสลีย์และตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อสำรวจประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์การระงับความรู้สึกทั่วไปเพื่อการผ่าตัดคลอด
2. เพื่อพัฒนาและตรวจสอบคุณภาพแบบวัดการตระหนักรู้สถานการณ์การระงับความรู้สึกทั่วไปเพื่อการผ่าตัดคลอดของนักศึกษาวิสัญญีพยาบาล

วิธีดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยและพัฒนา (research and development)

1. ประชากรและกลุ่มตัวอย่าง

กลุ่มเป้าหมายที่ 1 เป็นผู้เชี่ยวชาญสาขาวิชาวิสัญญีวิทยาเพื่อวิเคราะห์ระบบงาน (goal directed task analysis) ด้วยการสำรวจเดลฟายต่อประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์การระงับความรู้สึกแบบทั่วไปในผู้ป่วยผ่าตัดคลอดในลักษณะมาตรประมาณค่า 5 ระดับจำนวน 3 รอบ ซึ่งเป็นอาจารย์แพทย์วิสัญญีที่มีคุณวุฒิทางการศึกษาเฉพาะทางวิสัญญีวิทยา และมีประสบการณ์การทำงานมากกว่า 10 ปี จำนวน 7 ท่าน ซึ่งการสำรวจเดลฟายสามารถใช้ผู้ทรงคุณวุฒิที่มีขนาดตั้งแต่ 4-3000 คน (Linstone, 1985; Thangaratinam & Redman, 2005) โดย Linstone (1985) แนะนำว่าจำนวนผู้ทรงคุณวุฒิจำนวนน้อยที่สุดที่ใช้ได้ด้วยวิธีเดลฟายคือ 7 ท่าน

ประชากรเป้าหมายเพื่อทดลองใช้แบบวัดการตระหนักรู้สถานการณ์เป็นนักศึกษาวิสัญญีพยาบาลที่กำลังศึกษาในหลักสูตรฝึกอบรมวิสัญญีพยาบาลปีการศึกษา 2563-2564 ภายในสถาบันฝึกอบรมที่ผ่านการรับรองจากราชวิทยาลัยวิสัญญีแห่งประเทศไทยจำนวน 267 คน เพื่อตรวจสอบคุณสมบัติทางจิตมิติด้านความเที่ยงทั้งฉบับ และวิเคราะห์พารามิเตอร์ข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบหลังการทดลองใช้กับนักศึกษาวิสัญญีพยาบาลที่กำลังศึกษาในหลักสูตรฝึกอบรมวิสัญญีพยาบาลปีการศึกษา 2563-2564 ซึ่งการกำหนดขนาดตัวอย่างที่มีขนาดเล็กที่น้อยกว่าหรือเท่ากับ 100 คน สามารถทำการตรวจสอบคุณภาพรายข้อโดยใช้การวิเคราะห์พารามิเตอร์ข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ (dichotomous IRT) (Ding et al., 1995; Tinsley & Tinsley, 1987; Anderson & Gerbing, 1988; Tabachnick and Fidell, 2007) ผู้วิจัยจึงกำหนดจำนวนตัวอย่างที่เป็นนักศึกษาในการทดลองใช้แบบวัดฯ จำนวน 136 คน โดยกลุ่มตัวอย่างจากสถาบันที่ยินยอมเข้าร่วมการวิจัยจำนวน 7 สถาบันและนักศึกษาทุกคนเข้าร่วมการวิจัย คิดเป็นกลุ่มตัวอย่างสถาบันละ 12-32 คน โดยไม่มีการแบ่งกลุ่มตัวอย่าง ประกอบด้วย 1) คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี 2) คณะแพทยศาสตร์มหาวิทยาลัยสงขลานครินทร์ 3) โรงพยาบาลพระมงกุฎเกล้า 4) โรงพยาบาลราชวิถี 5) โรงพยาบาลชลบุรี 6) โรงพยาบาลพุทธชินราช และ 7) โรงพยาบาลสรรพสิทธิประสงค์

2. เครื่องมือที่ใช้ในการวิจัย

2.1 แบบสำรวจเดลฟายต่อประเด็นจำเป็นที่ควรตระหนักรู้ 3 องค์ประกอบใน 3 สถานการณ์ ได้แก่ ระเบียบเตรียมการ ระยชนนำสลบ และระยชนใส่ท่อช่วยหายใจ ในรูปแบบมาตรประมาณค่า 5 ระดับ จำนวน 55 ข้อรายการ

2.1 แบบวัดการตระหนักรู้สถานการณ์ของนักศึกษาวิสัญญีพยาบาลมีลักษณะเป็นข้อคำถามแบบเลือกตอบคำตอบเดียวจาก 4 ตัวเลือกที่สามารถใช้วัดการตระหนักรู้สถานการณ์ในแต่ละตัวบ่งชี้ รวมทั้งหมดประมาณ 48 ข้อ ได้แก่ ระเบียบเตรียมการ ระยชนนำสลบ และระยชนใส่ท่อช่วยหายใจ

3. การเก็บรวบรวมข้อมูล

3.1 ผู้วิจัยนำมโนทัศน์เกี่ยวกับการระงับความรู้สึกแบบทั่วไปแก่ผู้ป่วยผ่าตัดคลอดที่สังเคราะห์ได้มาจัดวางไว้ตามองค์ประกอบของการวิเคราะห์ระบบงาน ได้แก่ เป้าหมายหลัก เป้าหมายย่อย จุดที่ต้องการการตัดสินใจ เพื่อแตกย่อยเป็นประเด็นการตระหนักรู้สถานการณ์ซึ่งเป็นจุดต่างๆในสถานการณ์ที่บุคคลควรรับรู้ เข้าใจ และคาดการณ์ และสำรวจเดลฟายต่อประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์ในรูปแบบมาตรประมาณค่า 5 ระดับจำนวน 3 รอบ ซึ่งดำเนินการตามแนวทางมาตรฐานเพื่อประเมินการตระหนักรู้สถานการณ์แบบองค์รวม (SAGAT) ของ Endsley (1995a) โดยใช้สถานการณ์ที่คัดเลือกไว้จำนวน 3 เรื่องเกี่ยวกับการระงับความรู้สึกทั่วไปเพื่อการผ่าตัดคลอดซึ่งมีผู้ตัดสินใจหลักในสถานการณ์คือนักศึกษาวิสัญญีพยาบาล

3.2 ผู้วิจัยได้ประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์จากการสำรวจเดลฟาย จึงนำมาพัฒนาเป็นข้อคำถามการตระหนักรู้สถานการณ์ตามตัวบ่งชี้ในโมเดลการวัดการตระหนักรู้สถานการณ์ ผู้วิจัยใช้แนวคิดการตระหนักรู้สถานการณ์ของ Endsley (Endsley, 1995b; Endsley, 2015) เป็นพื้นฐานของโครงสร้างการวัดและนิยามตัวแปรเชิงปฏิบัติการและตัวบ่งชี้การวัดตัวแปร โดยกำหนดแผนผังการสร้างข้อสอบ สร้างโจทย์สถานการณ์ที่มีความเหมาะสมในการวัดการตระหนักรู้สถานการณ์สำหรับนักศึกษาวิสัญญีพยาบาลแบบวัดการตระหนักรู้สถานการณ์ของนักศึกษาวิสัญญีพยาบาลมีลักษณะเป็นแบบวัดเชิงสถานการณ์ต่อเนื่อง (เตรียมการ นำสลบ ใส่ท่อช่วยหายใจ) และตรวจสอบคุณภาพของแบบวัดด้านความตรงเชิงเนื้อหาโดยผู้ทรงคุณวุฒิจำนวน 7 ท่าน โดยใช้ค่าดัชนีความสอดคล้อง (Index of Item – Objective Congruence)

3.3 ผู้วิจัยนำหนังสือขออนุมัติเก็บข้อมูล เพื่อเข้าเก็บข้อมูล ผู้วิจัยเป็นผู้เก็บข้อมูลร่วมกับผู้ช่วยวิจัย 1 ท่าน ภายหลังจากโครงการวิจัยผ่านการอนุมัติจากคณะกรรมการจริยธรรมและวิจัยในคน กลุ่มสหสถาบัน

3.4 ผู้วิจัยประชาสัมพันธ์เพื่อเชิญชวนอาสาสมัคร และจัดประชุมเพื่อชี้แจงวัตถุประสงค์การวิจัย โดยเปิดโอกาสให้นักศึกษาซักถามจนเข้าใจ รวมถึงชี้แจงสิทธิในการตอบรับหรือปฏิเสธเข้าร่วมการวิจัย โดยอาสาสมัครจะต้องผ่านการคัดกรองก่อนจึงจะเป็นผู้มีส่วนร่วมในการวิจัยได้โดยต้องเป็นผู้มีคุณสมบัติตามเกณฑ์การคัดเข้า-คัดออก ผู้วิจัยมอบเอกสารข้อมูลสำหรับผู้มีส่วนร่วมในการวิจัยและหนังสือแสดงความยินยอมเข้าร่วมงานวิจัย และนัดหมายเพื่อเก็บข้อมูล

3.5 ผู้วิจัยจัดการทดลองใช้แบบวัดจำนวน 48 ข้อที่ได้ปรับปรุงตามข้อเสนอแนะของผู้ทรงคุณวุฒิกับสถาบันการศึกษาทางวิสัญญีที่ยินดีให้ความร่วมมือในการทดลอง มีจำนวน นักศึกษา 136 คน ใช้เวลาทดลองประมาณ 60 นาที จำนวน 1 ครั้ง ภายในสถาบันฝึกอบรมของนักศึกษาวิสัญญีพยาบาล

3.6 ผู้วิจัยตรวจสอบคุณภาพของแบบวัดการตระหนักรู้สถานการณ์ โดยตรวจสอบคุณสมบัติทางจิตมิติ ความเที่ยงทั้งฉบับ และวิเคราะห์พารามิเตอร์ข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ

4. การวิเคราะห์ข้อมูล

4.1 วิเคราะห์เพื่อนำเสนอข้อมูลในรอบต่างๆ จากการสำรวจเดลฟาย โดย กำหนดการผ่านเกณฑ์ ดังนี้ ค่าเฉลี่ย (Mean) ± 1.64 ค่ามัธยฐาน (Median) ≥ 3.5 และค่าพิสัยระหว่างควอไทล์ (interquartile range) $IQR \leq 1$ ในทุกรายการของประเด็นจำเป็นที่ควรตระหนักรู้จากผู้ทรงคุณวุฒิสาขาวิชาทุกท่าน (Rayens & Hahn, 2000) วิเคราะห์ฉันทามติด้วยสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (intra-class correlation Coefficient) เพื่อประเมินความสอดคล้อง (consistency) ของการตอบสำหรับมาตรฐานค่า 5 ระดับ (Brender, et al., 2006)

4.2 วิเคราะห์คุณภาพรายข้อของแบบวัดการตระหนักรู้สถานการณ์ ได้แก่ พารามิเตอร์ความง่ายและพารามิเตอร์อำนาจจำแนกเพื่อได้สารสนเทศของข้อสอบรายข้อตามแนวทางทฤษฎีการตอบสนองข้อสอบ ด้วยโมเดลการตอบสนองข้อสอบแบบตรงให้คะแนน 2 ค่า โดยผู้วิจัยใช้การประมาณค่าของ Bayesian IRT ในโปรแกรม R ด้วย package brms (Bürkner, 2017a; Bürkner, 2017b) ร่วมกับโปรแกรมภาษา Stan (Carpenter et al., 2017) โดย package brms ใช้ NUTS sampler (Hoffman & Gelman, 2014) ซึ่งสอดคล้องกับโมเดลที่ซับซ้อน และกลุ่มตัวอย่างที่มีขนาดน้อยกว่า 1000 คน โดยการคำนวณรอบต่าง ๆ ค่อนข้างใช้เวลานาน เพื่อให้การสุ่มตัวอย่างที่มีประสิทธิภาพในแต่ละรอบและทำให้กลุ่มตัวอย่างมีขนาดใหญ่ขึ้น

4.3 วิเคราะห์ความเที่ยงของแบบวัดการตระหนักรู้สถานการณ์ ด้วยการหาค่าความเที่ยงแบบความสอดคล้องภายใน (Cronbach's Alpha Coefficient) ด้วยโปรแกรม SPSS Version 22 for Window

ผลการวิจัย

1. ผลการสำรวจประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์การระงับความรู้สึกทั่วไปเพื่อการผ่าตัดคลอดของนักศึกษาวิสัญญีพยาบาล

ผลจากการสำรวจเดลฟายต่อประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์ในรอบต่างๆ จำนวน 55 ประเด็น พบว่าการวิเคราะห์ ค่าเฉลี่ย (M) ของการสำรวจรอบที่ 1-3 มีค่าระหว่าง 4.00-5.00 ส่วนเบี่ยงเบนมาตรฐาน (SD) ของ การสำรวจรอบที่ 1-3 มีค่าระหว่าง 0.00-1.46 ค่ามัธยฐาน (Mdn) ของการสำรวจรอบที่ 1-3 มีค่า ระหว่าง 4-5 และค่าพิสัยระหว่างควอไทล์ (Interquartile Range) ของการสำรวจรอบที่ 1-3 มีค่า ระหว่าง 0-1.5 และเมื่อเสร็จสิ้นการสำรวจรอบที่ 3 ผู้วิจัยตัดประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์ที่ไม่ผ่านเกณฑ์ที่กำหนดออก และผลการวิเคราะห์ฉันทามติของผู้เชี่ยวชาญสาขาวิชาด้วยสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (intra-class correlation coefficient) ต่อประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์ในการ

สำรวจเดลฟายรอบที่ 2 และ 3 มีค่า ICC ≥ 0.74 ซึ่งสัมพันธ์กับความสัมพันธ์ภายในชั้นมีค่าระหว่าง 0.75 – 1.00 และยุติการสำรวจเดลฟายเมื่อฉันทามติของผู้เชี่ยวชาญ สาขาวิชาอยู่ที่ร้อยละ 70 คงเหลือประเด็นจำเป็นที่ควรตระหนักไว้ในสถานการณ์ จำนวน 54 ประเด็นเพื่อนำไปเป็นจุดตั้งต้นในการสร้างข้อคำถามเพื่อวัดการตระหนักรู้สถานการณ์ในลำดับต่อไป

ตาราง 1 ผลการสำรวจเดลฟายต่อประเด็นจำเป็นที่ควรตระหนักไว้ในสถานการณ์รอบที่ 1-3

เรื่องที่ 1 การเตรียมการเพื่อระงับความรู้สึกทั่วไปเพื่อการผ่าตัดคลอด

ประเด็นที่ต้องการการตัดสินใจคือด้านระดับความพร้อมของผู้ป่วย ทารกในครรภ์ ทีมวิสัญญี สำหรับการระงับความรู้สึกทั่วไปกรณีเร่งด่วนเป็นอย่างไร

ประเด็นจำเป็นที่ควรตระหนักไว้ในสถานการณ์	รอบที่ 1				รอบที่ 2				รอบที่ 3			
	M	SD	IQR	Mdn	M	SD	IQR	Mdn	M	SD	IQR	Mdn
1.1 รวบรวมข้อมูล	4.29	0.95	1.5	5	4.71	0.49	0.5	5	4.86	0.38	0	5
1.2 ตรวจสอบผิดปกติ	4.57	0.79	0.5	5	4.71	0.49	0.5	5	4.71	0.49	0.5	5
1.3 ตีความ	4.00	1.15	1.5	4	4.57	0.79	0.5	5	4.71	0.49	0.5	5
1.4 สรุปประเด็นปัญหา	4.71	0.76	0	5	4.86	0.38	0	5	4.86	0.38	0	5
1.5 สรุปประเด็นปัญหา	4.57	0.79	0.5	5	4.71	0.76	0	5	4.71	0.76	0	5
1.6 คาดคะเนการปฏิบัติ	4.57	0.79	0.5	5	4.57	0.79	0.5	5	4.71	0.76	0	5
1.7 คาดคะเนเหตุการณ์	4.43	0.79	1	5	4.43	0.79	1	5	4.43	0.79	1	5
ประเด็นที่ต้องการการตัดสินใจคืออุปกรณ์ที่จำเป็นในการระงับความรู้สึกทั่วไปกรณีเร่งด่วนของการผ่าตัดคลอดรายนี้เป็นอย่างไร												
1.8 รวบรวมข้อมูล	4.71	0.49	0.5	5	4.86	0.38	0	5	4.86	0.38	0	5
1.9 คาดคะเนเหตุการณ์	4.43	0.79	1	5	4.43	0.79	1	5	4.57	0.53	1	5
ประเด็นที่ต้องการการตัดสินใจคือระดับค่า BMI ที่สัมพันธ์กับภาวะการใส่ท่อหายใจยากของผู้ป่วยรายนี้เป็นอย่างไร												
1.10 รวบรวมข้อมูล	4.71	0.49	0.5	5	4.71	0.49	0.5	5	4.71	0.49	0.5	5
1.11 ตรวจสอบผิดปกติ	4.86	0.38	0	5	4.86	0.38	0	5	4.86	0.38	0	5
1.12 สรุปประเด็นปัญหา	5.00	0.00	0	5	5.00	0.00	0	5	5.00	0.00	0	5
1.13 คาดคะเนการปฏิบัติ	4.86	0.38	0	5	5.00	0.00	0	5	5.00	0.00	0	5
1.14 รวบรวมข้อมูล	4.71	0.49	0.5	5	4.86	0.38	0	5	5.00	0.00	0	5
1.15 สรุปประเด็นปัญหา	4.43	0.79	1	5	4.29	1.11	1	5	4.57	0.53	1	5
1.16 สรุปประเด็นปัญหา	4.57	0.53	1	5	4.57	0.79	0.5	5	4.71	0.49	0.5	5
ประเด็นที่ต้องการการตัดสินใจคือทำการนอนของผู้ป่วยเพื่อป้องกันความดันโลหิตลดลง และช่วยในการใส่ท่อหายใจเป็นอย่างไร												
1.17 ตีความ	5.00	0.00	0	5	5.00	0.00	0	5	5.00	0.00	0	5
1.18 คาดคะเนเหตุการณ์	4.43	0.79	1	5	4.43	0.98	1	5	4.57	0.79	0.5	5

ประเด็นจำเป็นที่ควร ตระหนักรู้ ในสถานการณ์	รอบที่ 1				รอบที่ 2				รอบที่ 3			
	M	SD	IQR	Mdn	M	SD	IQR	Mdn	M	SD	IQR	Mdn
เรื่องที่ 2 การนำเสนอ												
ประเด็นที่ต้องการการตัดสินใจคือประเด็นที่สำคัญที่วิสัญญีควรทราบเกี่ยวกับผู้ป่วยก่อนการนำเสนอคืออะไร												
2.1 รวบรวมข้อมูล	4.29	1.11	1	5	4.57	0.79	0.5	5	4.71	0.49	0.5	5
2.2 ตรวจสอบความผิดปกติ	4.57	0.79	0.5	5	4.71	0.49	0.5	5	4.86	0.38	0	5
2.3 ติความ	4.29	1.11	1	5	4.14	0.69	0.5	4	4.29	0.49	0.5	4
2.4 คาดคะเนเหตุการณ์	4.29	1.25	1	5	4.14	1.46	1.5	5	4.43	0.98	1	5
ประเด็นที่ต้องการการตัดสินใจคือควรใช้ชนิด และขนาดของยา และวิธีในการนำเสนออย่างไร												
2.5 รวบรวมข้อมูล	4.71	0.49	0.5	5	4.71	0.49	0.5	5	4.86	0.38	0	5
2.6 คาดคะเนการปฏิบัติ	4.71	0.49	0.5	5	4.86	0.38	0	5	4.86	0.38	0	5
2.7 คาดคะเนเหตุการณ์	4.43	0.79	1	5	4.71	0.49	0.5	5	4.71	0.49	0.5	5
2.8 ติความ	4.43	0.79	1	5	4.57	0.79	0.5	5	4.71	0.49	0.5	5
ประเด็นที่ต้องการการตัดสินใจคือก่อนการนำเสนอควรสำรองออกซิเจนในปอดผู้ป่วยรายนี้อย่างไร												
2.9 รวบรวมข้อมูล	4.43	0.98	1	5	4.43	0.98	1	5	4.57	0.79	0.5	5
2.10 ตรวจสอบความผิดปกติ	4.43	0.98	1	5	4.29	0.95	1.5	5	4.43	0.79	1	5
2.11 ติความ	4.71	0.76	0	5	4.86	0.38	0	4	5.00	0.00	0	5
2.12 สรุปประเด็นปัญหา	4.71	0.76	0	5	4.86	0.38	0	5	5.00	0.00	0	5
2.13 คาดคะเนเหตุการณ์	4.71	0.49	0.5	5	4.71	0.49	0.5	5	4.86	0.38	0	5
2.14 คาดคะเนการปฏิบัติ	4.86	0.38	0	5	5.00	0.00	0	5	5.00	0.00	0	5
ประเด็นที่ต้องการการตัดสินใจคือก่อนและหลังการนำเสนอควรปฏิบัติกรกดคอผู้ป่วยรายนี้อย่างไร												
2.15 คาดคะเนเหตุการณ์	4.43	0.98	1	5	4.43	0.79	1	5	4.43	0.79	1	5
2.16 สรุปประเด็นปัญหา	4.57	0.79	0.5	5	4.43	0.79	1	5	4.57	0.79	0.5	5
2.17 คาดคะเนการปฏิบัติ	4.86	0.38	0	5	4.86	0.38	0	5	4.86	0.38	0	5
2.18 คาดคะเนเหตุการณ์	4.86	0.38	0	5	5.00	0.00	0	5	5.00	0.00	0	5
เรื่องที่ 3 การใส่ท่อช่วยหายใจ												
ประเด็นที่ต้องการการตัดสินใจคือการใส่ท่อช่วยหายใจที่สำเร็จและมีประสิทธิภาพควรใช้วิธีการและอุปกรณ์ที่ช่วยหายใจอย่างไร												
3.1 รวบรวมข้อมูล	4.43	0.79	1	5	4.57	0.53	1	5	4.71	0.49	0.5	5
3.2 ตรวจสอบความผิดปกติ	4.57	0.53	1	5	4.57	0.53	1	5	4.57	0.53	1	5
3.3 ติความ	4.29	1.11	1	4	4.29	1.25	1	5	4.57	0.79	0.5	5
3.4 สรุปประเด็นปัญหา	4.29	1.11	1	5	4.00	1.15	1.5	4	4.43	0.79	1	5
3.5 ติความ	4.57	0.79	0.5	5	4.57	0.79	0.5	5	4.57	0.79	0.5	5
3.6 คาดคะเนการปฏิบัติ	4.57	0.53	1	5	4.71	0.49	0.5	5	4.86	0.38	0	5
3.7 ติความ	4.29	1.25	1	5	4.14	1.07	2	5	4.14	1.07	2	5
3.8 ติความ	4.71	0.49	0.5	5	4.86	0.38	0	5	5.00	0.00	0	5

ประเด็นจำเป็นที่ควร ตระหนักรู้ ในสถานการณ์	รอบที่ 1				รอบที่ 2				รอบที่ 3			
	M	SD	IQR	Mdn	M	SD	IQR	Mdn	M	SD	IQR	Mdn
3.9 ดีความ	4.57	0.79	0.5	5	4.71	0.49	0.5	5	4.86	0.38	0	5
3.10 สรุปประเด็นปัญหา	4.71	0.76	0	5	4.86	0.38	0	5	5.00	0.00	0	5
ประเด็นที่ต้องการการตัดสินใจคือการบาดเจ็บที่เกิดกับทางเดินหายใจส่วนบนของผู้ป่วยหลังใส่ท่อหายใจไม่สำเร็จเป็นอย่างไร												
3.11 รวบรวมข้อมูล	4.43	0.79	1	5	4.43	0.79	1	5	4.43	0.79	1	5
3.12 ตรวจสอบผิดปกติ	4.29	1.11	1	5	4.29	0.95	1.5	5	4.29	0.95	1.5	5
3.13 สรุปประเด็นปัญหา	4.43	0.79	1	5	4.43	0.53	1	4	4.57	0.53	1	5
3.14 คัดคะแนนการปฏิบัติ	4.43	0.98	1	5	4.43	0.79	1	5	4.57	0.79	0.5	5
ประเด็นที่ต้องการการตัดสินใจคือผู้เชี่ยวชาญเพื่อช่วยหายใจครั้งที่ 3 ควรเป็นใครในสถานการณ์นี้												
3.15 คัดคะแนนการปฏิบัติ	5.00	0.00	0	5	5.00	0.00	0	5	5.00	0.00	0	5
ประเด็นที่ต้องการการตัดสินใจคือควรปลุกผู้ป่วยหรือการเริ่มการผ่าตัดคลอดต่อไปหรือไม่อย่างไร												
3.16 ดีความ	4.71	0.49	0.5	5	4.71	0.49	0.5	5	4.71	0.49	0.5	5
3.17 คัดคะแนนเหตุการณ์	4.29	1.25	1	5	3.71	1.25	2	3	4.00	1.29	2	5
3.18 คัดคะแนนเหตุการณ์	4.43	1.13	0.5	5	4.29	1.11	1	5	4.29	1.11	1	5
3.19 คัดคะแนนการปฏิบัติ	4.57	0.79	0.5	5	4.43	0.98	1	5	4.43	0.98	1	5

2. ผลการพัฒนาแบบวัดการตระหนักรู้สถานการณ์การระงับความรู้สึกทั่วไปเพื่อการผ่าตัดคลอดของนักศึกษาวิสัญญีพยาบาล

ผลการตรวจสอบความตรงเชิงเนื้อหาของแบบวัดฯ โดยผู้ทรงคุณวุฒิจำนวน 7 ท่าน ข้อคำถามและตัวเลือกคำตอบมีค่า IOC (Item-Objective Congruence) อยู่ระหว่าง 0.14-0.86 โดยค่า IOC ตั้งแต่ 0.50 ขึ้นไปถือว่าผ่านเกณฑ์ (Kanjanasri, 2013) พบว่าข้อคำถามที่ไม่ผ่านเกณฑ์จำนวน 7 ข้อ ซึ่งเป็นข้อคำถามที่ไม่สอดคล้องกับหลักฐานเชิงประจักษ์ และขัดแย้งเหตุการณ์จริง ผู้วิจัยจึงตัดข้อคำถามดังกล่าว คงเหลือข้อคำถามเพื่อวัดการตระหนักรู้สถานการณ์ที่พัฒนาขึ้นจำนวน 48 ข้อ เพื่อนำไปทดลองใช้กับกลุ่มตัวอย่างเพื่อตรวจสอบคุณภาพเครื่องมือต่อไป

2.1 หลังการทดลองใช้จึงนำมาวิเคราะห์เพื่อได้สารสนเทศของข้อสอบรายข้อตามทฤษฎีการตอบสนองข้อสอบ ด้วยโมเดลการตอบสนองข้อสอบแบบตรงให้คะแนน 2 ค่า โดยผู้วิจัยใช้การประมาณค่าของ Bayesian IRT Package brms โปรแกรม R ผู้วิจัยตรวจสอบความสอดคล้องของข้อมูลเชิงประจักษ์กับโมเดลและใช้การเปรียบเทียบโมเดลจากการประมาณ leave-one-out cross validation (LOO-CV) โดย LOOs ที่มีค่าน้อยกว่าหมายถึงโมเดลนั้นมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่า (Gelfand et al., 1992) และพิจารณาความแตกต่างของ LOOIC สำหรับการประมาณค่า (LOO-CV)

พบว่าโมเดล 2PL มีค่า (LOOIC= 6297.2; SE=84.0) ซึ่งเป็นค่าที่น้อยที่สุดเมื่อเปรียบเทียบกับโมเดล 1PL (LOOIC= 6976.2; SE=71.8) และ 3PL (LOOIC= 7548.0; SE=44.9) เมื่อเปรียบเทียบความแตกต่าง

ระหว่างโมเดล 1PL และ 2PL พบว่าค่า Δ LOOIC แตกต่างกันอย่างน้อยกว่า โมเดล 2PL และ 3PL ดังนั้นจึงเลือกใช้โมเดลที่มีความซับซ้อนน้อย และมีค่า LOOIC น้อยที่สุดคือโมเดล 2PL ข้อมูลจากแบบวัดการตระหนักรู้สถานการณ์จึงมีความสอดคล้องกับโมเดล 2PL โดยมีค่า Rhat =1.00 ที่แสดงถึงการเข้าสู่ของโมเดลเพื่อนำเสนอพารามิเตอร์ของข้อสอบในค่าอำนาจจำแนก ซึ่งใช้การประมาณค่าของค่าเฉลี่ยด้านหลังเพื่อนำเสนอค่าพารามิเตอร์ความง่าย (easiness) ที่แปลผลตรงกันข้ามกับค่าความยาก โดยผู้วิจัยใช้เกณฑ์ในการพิจารณาข้อสอบที่มีความเหมาะสม ได้แก่ ค่าความยากของข้อสอบควรมีค่าอยู่ระหว่าง +2.50 ถึง -2.50 ค่าอำนาจจำแนกของข้อสอบควรมีค่ามากกว่า +0.50 และค่าโอกาสในการเดา ควรมีค่าไม่เกิน 0.30 (Embretson & Reise, 2013; Kanjanawasri, 2013) จากการประมาณค่าพารามิเตอร์ด้วยโมเดล 2 PL อำนาจจำแนกมีค่าระหว่าง -1.28-2.06 ส่วนพารามิเตอร์ความง่ายมีค่าระหว่าง -1.87-3.91 แบบวัดการตระหนักรู้สถานการณ์ของนักศึกษาวิทยาลัยพยาบาลมีจำนวน 48 ข้อ โดยมีข้อสอบที่เหมาะสมจำนวน 41 ข้อ และมีข้อคำถามที่ควรตัดทิ้งจำนวน 7 ข้อดังตาราง 2

ตาราง 2 ผลการวิเคราะห์ค่าความง่าย อำนาจจำแนกของแบบวัดการตระหนักรู้สถานการณ์ของนักศึกษาวิทยาลัยพยาบาล

Item	Easiness (SE)	Discriminant (SE)	แปลผล	Item	Easiness (SE)	Discriminant (SE)	แปลผล
ข้อที่ 1.1	2.38 (1.01)	1.56 (0.61)	ใช้ได้	ข้อที่ 2.9	2.69 (3.26)	-1.28 (0.51)	ใช้ไม่ได้
ข้อที่ 1.2	2.18 (1.76)	1.12 (0.48)	ใช้ได้	ข้อที่ 2.10	2.69 (1.03)	-0.88 (0.48)	ใช้ไม่ได้
ข้อที่ 1.3	2.05 (0.54)	0.75 (0.52)	ใช้ได้	ข้อที่ 2.11	1.93 (1.42)	1.22 (0.50)	ใช้ได้
ข้อที่ 1.4	2.02 (0.91)	0.74 (0.51)	ใช้ได้	ข้อที่ 2.12	3.24 (1.61)	-0.21 (0.52)	ใช้ไม่ได้
ข้อที่ 1.5	2.07 (1.89)	-0.39 (0.51)	ใช้ไม่ได้	ข้อที่ 2.13	2.20 (0.23)	0.63 (0.54)	ใช้ได้
ข้อที่ 1.6	2.07 (0.41)	0.55 (0.54)	ใช้ได้	ข้อที่ 2.14	2.03 (0.81)	1.17 (0.57)	ใช้ได้
ข้อที่ 1.7	2.08 (0.28)	0.59 (0.54)	ใช้ได้	ข้อที่ 2.15	3.13 (0.64)	-0.47 (0.53)	ใช้ไม่ได้
ข้อที่ 1.8	-0.79 (0.67)	0.69 (0.53)	ใช้ได้	ข้อที่ 3.1	0.50 (0.79)	2.06 (0.67)	ใช้ได้
ข้อที่ 1.9	-2.35 (0.66)	1.73 (0.56)	ใช้ได้	ข้อที่ 3.2	1.80 (0.29)	1.08 (0.49)	ใช้ได้
ข้อที่ 1.10	1.59 (0.56)	1.59 (0.55)	ใช้ได้	ข้อที่ 3.3	2.32 (1.29)	1.02 (0.54)	ใช้ได้

Item	Easiness (SE)	Discriminant (SE)	แปลผล	Item	Easiness (SE)	Discriminant (SE)	แปลผล
ข้อที่ 1.11	1.85 (0.60)	1.47 (0.53)	ใช้ได้	ข้อที่ 3.4	2.32 (1.64)	-0.78 (0.51)	ใช้ไม่ได้
ข้อที่ 1.12	1.72 (0.93)	0.21 (0.54)	ใช้ได้	ข้อที่ 3.5	1.86 (0.72)	1.84 (0.64)	ใช้ได้
ข้อที่ 1.13	1.87 (0.30)	1.38 (0.57)	ใช้ได้	ข้อที่ 3.6	1.08 (1.16)	0.54 (0.55)	ใช้ได้
ข้อที่ 1.14	-0.44 (0.41)	1.68 (0.64)	ใช้ได้	ข้อที่ 3.7	-0.03 (1.18)	1.57 (0.63)	ใช้ได้
ข้อที่ 1.15	2.11 (0.55)	0.19 (0.53)	ใช้ได้	ข้อที่ 3.8	-0.62 (1.76)	1.58 (0.64)	ใช้ได้
ข้อที่ 1.16	2.39 (0.33)	0.60 (0.55)	ใช้ได้	ข้อที่ 3.9	2.46 (2.52)	1.48 (0.59)	ใช้ได้
ข้อที่ 2.1	2.06 (0.03)	1.81 (0.63)	ใช้ได้	ข้อที่ 3.10	1.55 (1.31)	1.51 (0.54)	ใช้ได้
ข้อที่ 2.2	1.67 (0.62)	1.02 (0.49)	ใช้ได้	ข้อที่ 3.11	-0.92 (1.04)	1.66 (0.62)	ใช้ได้
ข้อที่ 2.3	2.19 (1.47)	1.55 (0.56)	ใช้ได้	ข้อที่ 3.12	-1.87 (0.88)	0.70 (0.50)	ใช้ได้
ข้อที่ 2.4	2.22 (1.23)	0.65 (0.55)	ใช้ได้	ข้อที่ 3.13	1.87 (1.56)	0.74 (0.54)	ใช้ได้
ข้อที่ 2.5	0.40 (0.50)	0.59 (0.54)	ใช้ได้	ข้อที่ 3.14	2.32 (0.70)	0.51 (0.52)	ใช้ได้
ข้อที่ 2.6	2.14 (1.11)	0.64 (0.54)	ใช้ได้	ข้อที่ 3.15	1.09 (0.21)	0.55 (0.54)	ใช้ได้
ข้อที่ 2.7	1.78 (0.55)	1.24 (0.53)	ใช้ได้	ข้อที่ 3.16	2.04 (1.08)	0.83 (0.54)	ใช้ได้
ข้อที่ 2.8	3.91 (0.34)	-1.01 (0.49)	ใช้ไม่ได้	ข้อที่ 3.17	2.50 (1.21)	0.56 (0.51)	ใช้ได้

2.2 ผลการวิเคราะห์ความเที่ยงของแบบวัดด้านความสอดคล้องภายใน พบว่าค่าความเที่ยงของแบบวัดสถานการณ์ที่ 1 เท่ากับ 0.75 ค่าความเที่ยงของแบบวัดสถานการณ์ที่ 2 เท่ากับ 0.74 ค่าความเที่ยงของแบบวัดสถานการณ์ที่ 3 เท่ากับ 0.8 ค่าความเที่ยงของแบบวัดทั้งฉบับเท่ากับ 0.75 แสดงว่าแบบวัดที่พัฒนาขึ้นมีคุณภาพอยู่ในระดับดี

อภิปรายผล

1. ผลการสำรวจประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์

การสำรวจเดลฟายในรอบที่ 1-3 มีการกำหนดเกณฑ์ผ่านในการสำรวจเดลฟายในรอบต่าง ๆ รวมถึงเกณฑ์ผ่านของฉันทมติที่ชัดเจน โดยประเด็นการตระหนักรู้สถานการณ์ในรอบที่ 1-3 จำนวน 55 ประเด็น เมื่อยุติการสำรวจพบว่าประเด็นที่ไม่ผ่านเกณฑ์และตัดออกจำนวน 1 ประเด็น ผู้วิจัยยุติการสำรวจเดลฟายเมื่อฉันทมติของผู้ทรงคุณวุฒิสาขาวิชาอยู่ที่ร้อยละ 70 ซึ่งค่า ICC ของการสำรวจเดลฟายในรอบที่ 2-3 มีค่า $ICC \geq 0.74$ ซึ่งสัมพันธ์สัมพันธ์ภายในชั้นที่มีค่าระหว่าง 0.75 – 1.00 หมายถึงความเที่ยงระหว่างผู้ประเมินมีความสอดคล้องกันในระดับดีถึงดีมาก (Roh et al., 2015) การที่ผู้วิจัยเลือกใช้สถานการณ์ทางวิสัญญีวิทยาที่มีโครงสร้างของสถานการณ์ที่ชัดเจน และสามารถอ้างอิงได้ตามแนวปฏิบัติมาตรฐานเพื่อสะท้อนระยะต่างๆของการระงับความรู้สึกทั่วไป จึงเป็นส่วนหนึ่งที่ทำให้ผู้เชี่ยวชาญสาขาวิชามองเห็นภาพเดียวกัน จากการบรรยายและภาพประกอบในสถานการณ์ส่งผลให้การสำรวจเดลฟายสามารถยุติได้ในรอบที่ 3 และผู้วิจัยนำผลจากการสำรวจโดยนำประเด็นการตระหนักรู้สถานการณ์ไปสร้างเป็นตัวอย่างข้อคำถามโดยสร้างขึ้นตามคำนิยามโมเดลการวัดการตระหนักรู้สถานการณ์ที่พัฒนา ซึ่งสอดคล้องกับ Vincent et al., 2018 ที่ใช้ข้อคำถามการตระหนักรู้สถานการณ์ที่สร้างจากประเด็นที่ได้จากการวิเคราะห์ระบบงาน และตรวจสอบความตรงจากการนำไปใช้ของเครื่องมือวัดการตระหนักรู้สถานการณ์ระหว่าง SAGAT และ SART ต่อคลาดเคลื่อนของทีม ICU ในระหว่างการฝึกสถานการณ์จำลอง โดยทีมสหสาขาวิชาที่มีพยาบาล 2 คน และแพทย์ฝึกหัด รวมถึงสถานการณ์จำลองมาตรฐานในภาวะฉุกเฉิน

2. ผลการพัฒนาแบบวัดการตระหนักรู้สถานการณ์การระงับความรู้สึกทั่วไปเพื่อการผ่าตัดคลอดของนักศึกษาวิสัญญีพยาบาล

แบบวัดมีการตรวจสอบความตรงเชิงเนื้อหา พบว่ามีค่า IOC อยู่ระหว่าง 0.14-0.86 โดยข้อคำถามที่ไม่ผ่านเกณฑ์เป็นข้อคำถามที่ไม่สอดคล้องกับหลักฐานเชิงประจักษ์และแนวปฏิบัติที่ดี หรือขัดแย้งกับเหตุการณ์จริง

จากการวิเคราะห์คุณภาพของข้อสอบรายข้อด้วยทฤษฎีการตอบสนองข้อสอบข้อมูลเชิงประจักษ์จากแบบวัดการตระหนักรู้สถานการณ์มีความสอดคล้องกับโมเดล 2PL โดยมีค่า $Rhat = 1.00$ จากการประมาณค่าพารามิเตอร์ด้วยโมเดล 2PL อำนาจจำแนกมีค่าระหว่าง -1.28-2.06 ส่วนพารามิเตอร์ความง่ายมีค่าระหว่าง -1.87-3.91 แบบวัดการตระหนักรู้สถานการณ์ของนักศึกษาวิสัญญีพยาบาลมีจำนวน 48 ข้อ โดยมีข้อสอบที่เหมาะสมจำนวน 41 ข้อ สอดคล้องกับงานวิจัยของ Lavoie et al. (2016) ทดสอบการตัดสินใจทางคลินิกของนักศึกษาพยาบาลต่อการเปลี่ยนแปลงของผู้ป่วยสถานการณ์จำลอง และพัฒนาเครื่องมือวัดการตระหนักรู้สถานการณ์ กลุ่มตัวอย่างเป็นนักศึกษาพยาบาลระดับปริญญาตรีโดยใช้เทคนิคการประเมินการตระหนักรู้สถานการณ์ทั่วไป (SAGAT) ผลการประมาณค่าความยาก อำนาจจำแนก ดัชนีความเสมือนจริงในการพัฒนาข้อคำถามเพื่อวัดการตระหนักรู้สถานการณ์จำนวน 31 ข้อ ซึ่งข้อคำถามส่วนใหญ่มีความตรงเชิงเนื้อหาในระดับสูง และมีค่าความยากและอำนาจจำแนกรวมถึงดัชนีความเสมือนจริงอยู่ในเกณฑ์ที่ดี

ผลการวิเคราะห์ความเที่ยงของแบบวัดแบบความสอดคล้องภายใน (Cronbach's Alpha) พบว่าค่าความเที่ยงของแบบวัดเรื่องที่ 1-3 ทั้งฉบับเท่ากับ 0.74-0.8 ตามลำดับ แสดงว่า แบบวัดที่พัฒนาขึ้นมีคุณภาพอยู่ในระดับดี ซึ่งสอดคล้องกับ Hultin et al. (2019) ศึกษาวิจัยโดยมีวัตถุประสงค์เพื่อทำการตรวจสอบรายการที่ใช้ในเพื่อประเมินอาการที่แสดงความบกพร่องของระบบทางเดินหายใจ การไหลเวียนโดยใช้เทคนิคการประเมินการตระหนักรู้สถานการณ์ (SAGAT) เพื่อตรวจสอบความเที่ยงแบบสอดคล้องภายในสำหรับการประเมินความบกพร่องของระบบทางเดินหายใจและระบบไหลเวียนตาม PSQ รายการใน SAGAT มีความสัมพันธ์กับสถานการณ์ 96% ของ อัลฟาของ Cronbach สำหรับ SAGAT / TSAGAT สถานการณ์ที่ 1 เท่ากับ 0.80 / 0.83 สถานการณ์ที่ 2 เท่ากับ 0.62 / 0.76

กิตติกรรมประกาศ (Acknowledgements)

งานวิจัยฉบับนี้ได้รับทุนสนับสนุนการวิจัย “ทุนพัฒนานวัตกรรมศึกษา” จากสำนักงานคณะกรรมการวิจัยแห่งชาติปีงบประมาณ 2564

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

ผลที่ได้จากการวิจัยต่อประเด็นจำเป็นที่ควรตระหนักรู้ในสถานการณ์จากการสำรวจเดลฟายนำไปใช้เป็นจุดตั้งต้นในการสร้างแบบวัดการตระหนักรู้สถานการณ์การระงับความรู้สึกทั่วไปเพื่อการผ่าตัดคลอดที่สามารถเพิ่มจุดเน้นของการตระหนักรู้สถานการณ์ในระยะต่างๆของการระงับความรู้สึก เช่น การเสียเลือดมดลูกหดตัวไม่ดี ภาวะวิกฤตที่เกิดกับมารดา ก่อนคลอด เป็นต้น ซึ่งจะช่วยให้การวัดการตระหนักรู้สถานการณ์มีความหลากหลายและน่าสนใจของสถานการณ์ทดสอบมากขึ้น ส่งผลต่อการคัดเลือกสถานการณ์และข้อคำถามที่ดีเพื่อพัฒนาต่อเป็นรูปแบบการประเมินในลักษณะของสถานการณ์จำลองเสมือนจริงด้วย virtual reality ได้ต่อไป

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

ในการพัฒนางานวิจัยครั้งต่อไปอาจพิจารณาคัดเลือกสถานการณ์วิกฤตที่มีความหลากหลาย โดยพิจารณาว่าระดับการตระหนักรู้สถานการณ์แบบกลุ่มหรือรายบุคคล อาจพิจารณาใช้วิธีการวัดหรือเครื่องมือมาตรฐานเพื่อช่วยตรวจสอบความตรงตามเกณฑ์สัมพัทธ์ ความตรงเชิงโครงสร้างสำหรับแบบวัดการตระหนักรู้สถานการณ์ต่อไป หรือพัฒนาโมเดลข้อสอบซึ่งเป็นเทคนิคการเขียนข้อสอบที่ช่วยให้ผู้เขียนข้อสอบออกข้อสอบคุณภาพได้จำนวนมากซึ่งช่วยเพิ่มคุณค่าให้การสร้างข้อสอบจากแบบวัดต่อไป

References

- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, 103(3), 411.
- Bedny, G. Z., Karwowski, W., & Jeng, O. J. (2004). The situational reflection of reality in activity theory and the concept of situation awareness in cognitive psychology. *Theoretical Issues in Ergonomics Science*, 5(4), 275-296.
- Billings, C. E. (1995, November). Situation awareness measurement and analysis: A commentary. In *Proceedings of the International Conference on Experimental Analysis and Measurement of Situation Awareness (Vol. 1)*. Embry-Riddle Aeronautical University Press.
- Brender, J., Ammenwerth, E., Nykänen, P., & Talmon, J. (2006). Factors influencing success and failure of health informatics systems. *Methods of information in medicine*, 45(01), 125-136.
- Bürkner, P. C. (2017a). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80, 1-28.
- Bürkner, P. C. (2017b). *Advanced Bayesian multilevel modeling with the R package brms*. https://cran.r-project.org/web/packages/brms/vignettes/brms_multilevel.pdf
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A., (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1).
- Cheney, F. W., Posner, K. L., Lee, L. A., Caplan, R. A., & Domino, K. B. (2006). Trends in anesthesia-related death and brain damage: a closed claims analysis. *The Journal of the American Society of Anesthesiologists*, 105(6), 1081-1086.
- Davies, J. M., Posner, K. L., Lee, L. A., Cheney, F. W., & Domino, K. B. (2009). Liability associated with obstetric anesthesia: a closed claims analysis. *The Journal of the American Society of Anesthesiologists*, 110(1), 131-139.
- Ding, L., Velicer, W. F., & Harlow, L. L. (1995). Effects of estimation methods, number of indicators per factor, and improper solutions on structural equation modeling fit indices. *Structural Equation Modeling: A Multidisciplinary Journal*, 2(2), 119-143.
- Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press.
- Endsley, M. R. (1988, October). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society annual meeting*, 32(2), 97-101.

- Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human factors*, 37(1), 65-84.
- Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human factors*, 37(1), 32-64.
- Endsley, M. R. (2015). Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making*, 9(1), 4-32.
- Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *The International Journal of Aviation Psychology*, 4(3), 241-264.
- Endsley, M. R., & Garland, D. J. (Eds.). (2000). *Situation awareness analysis and measurement*. CRC Press.
- Endsley, M. R., & Jones, D. G. (2004). *Chapter 7 Confidence and Uncertainty in situation awareness and decision making-Designing for situation awareness*.
- Gelfand, A. E., Dey, D. K., & Chang, H. (1992). *Model determination using predictive distributions with implementation via sampling-based methods*. Stanford University.
- Goss, S., & Murray, G. (1996). *Artificial Intelligence Applications in Aircraft Systems*. Defence Science and Technology.
- Hawkins, J. L., Chang, J., Palmer, S. K., Gibbs, C. P., & Callaghan, W. M. (2011). Anesthesia related maternal mortality in the United States: 1979–2002. *Obstetrics & Gynecology*, 117(1), 69-74.
- Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(1), 1593-1623.
- Hultin, M., Jonsson, K., Härgestam, M., Lindkvist, M., & Brulin, C. (2019). Reliability of instruments that measure situation awareness, team performance and task performance in a simulation setting with medical students. *BMJ open*, 9(9), e029412.
- Kanjanawasri, S. (2013). *Classical test theory*. 7th ed. Chulalongkorn University Press.
- Lavoie, P., Cossette, S., & Pepin, J. (2016). Testing nursing students' clinical judgment in a patient deterioration simulation scenario: development of a situation awareness instrument. *Nurse Education Today*, 38, 61-67.
- Linstone, H. A. (1985). *The delphi technique*. *Environmental Impact Assessment*. Springer.
- McKenna, L., Missen, K., Cooper, S., Bogossian, F., Bucknall, T., & Cant, R. (2014). Situation awareness in undergraduate nursing students managing simulated patient deterioration. *Nurse Education Today*, 34(6), e27-e31.

- Munnur, U., & Suresh, M. S. (2004). Airway problems in pregnancy. *Critical care clinics*, 20(4), 617-642.
- Mushambi, M. C., Kinsella, S. M., Popat, M., Swales, H., Ramaswamy, K. K., Winton, A. L., & Quinn, A. C. (2015). Obstetric Anaesthetists' Association and Difficult Airway Society guidelines for the management of difficult and failed tracheal intubation in obstetrics. *Anaesthesia*, 70(11), 1286-1306. <https://doi.org/10.1111/anae.13260>
- Nurse anesthetist training project. (2019). *Performance assessment of nurse anesthetist students 50th report* [Report document]. Faculty of Medicine, Mahidol university.
- Rayens, M. K., & Hahn, E. J. (2000). Building consensus using the policy Delphi method. *Policy, politics, & nursing practice*, 1(4), 308-315.
- Roh, W. S., Kim, D. K., Jeon, Y. H., Kim, S. H., Lee, S. C., Ko, Y. K., ... & Lee, G. H. (2015). Analysis of anesthesia-related medical disputes in the 2009-2014 period using the Korean Society of Anesthesiologists database. *Journal of korean medical science*, 30(2), 207-213.
- Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology*, 1(1), 45-57.
- Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human factors*, 37(1), 137-148.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics Boston*. Allyn and Bacon.
- Taylor, R. M. (2017). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational awareness* (pp. 111-128). Routledge.
- Thangaratnam, S., & Redman, C. W. (2005). The delphi technique. *The obstetrician & gynaecologist*, 7(2), 120-125.
- Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of counseling psychology*, 34(4), 414.
- Vidulich, M., Dominguez, C., Vogel, E., & McMillan, G. (1994). *Situation awareness: Papers and annotated bibliography*. Armstrong Laboratory.
- Vincent, L., Higham, H., Greig, P., Warren, R., Chantler, J., McKechnie, S., Venes, T., Young, D. and Vincent, C. (2018). SC25 Using the situation awareness global assessment tool (SAGAT) and the situation awareness rating technique (SART) using standardized scenarios in simulation training for intensive care teams. *BMJ Simulation & Technology Enhanced Learning*, 4(2), A26.
- Wright, S. M. (2015). Patient safety in anesthesia: learning from the culture of high-reliability organizations. *Critical Care Nursing Clinics*, 27(1), 1-16.

การเปรียบเทียบความเที่ยงและความตรงตามเกณฑ์ในการตรวจให้คะแนนของผู้ตรวจ
ที่มีคุณลักษณะต่างกันของแบบสอบอัตนัยสำหรับวัดสมรรถนะ
ทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

The Comparison of Reliability and Validity of the Raters' Scoring Criteria
with Different Characteristics of Raters of Essay Tests for Measuring Scientific
Competence of Grade 9 Students

พนิดา จังหะ^{1*} และ ประกฤติยา ทักซิโน²

Phanida Changwa^{1*} and Prakittiya Tuksino²

(Received: April 8, 2022; Revised: June 14, 2022; Accepted: June 21, 2022)

บทคัดย่อ

การวิจัยในครั้งนี้มีความมุ่งหมายของการวิจัย คือ 1) เพื่อศึกษาค่าความสอดคล้องของผลการตรวจให้คะแนนแบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ ภายใต้คุณลักษณะของผู้ตรวจที่แตกต่างกัน โดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) 2) เพื่อเปรียบเทียบความตรงตามเกณฑ์การให้คะแนนแบบองค์รวม ของแบบสอบอัตนัยสำหรับวัดสมรรถนะทางวิทยาศาสตร์ โดยพิจารณาจากความเห็นพ้อง (rater agreement) ภายใต้คุณลักษณะของผู้ตรวจที่แตกต่างกัน 3. เพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) ภายใต้คุณลักษณะของผู้ตรวจที่แตกต่างกัน ที่มีรูปแบบการตรวจ 2 รูปแบบ ได้แก่ รูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$(p \times i : r)$] และ รูปแบบการตรวจทุกข้อของผู้สอบทุกคน [$(p \times i \times r)$] กลุ่มตัวอย่างแบ่งออกเป็น 2 กลุ่ม คือ กลุ่มนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 100 คน และกลุ่มผู้ตรวจให้คะแนน ซึ่งแบ่งออกเป็นกลุ่มผู้ตรวจให้คะแนนที่เป็นผู้ตรวจที่มีคุณวุฒิตรงเอก จำนวน 3 คน และกลุ่มผู้ตรวจให้คะแนนที่เป็นผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก จำนวน 3 คน เครื่องมือที่ใช้ในการวิจัย คือ 1) แบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 3 สถานการณ์ รวม 9 ข้อคำถาม 2) เกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) วิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงโดยใช้โปรแกรม EduG ผลการวิจัยพบว่า 1) การศึกษาความสอดคล้องของผลการตรวจให้คะแนนระหว่างผู้ตรวจ พิจารณาโดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) ที่มีความสอดคล้องกัน

¹ นักศึกษาปริญญาโท สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² ผู้ช่วยศาสตราจารย์ สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

¹ Master's degree student, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

² Assistant Professor, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

* Corresponding Author E-mail: Care_phanida@kkumail.com *

โดยเฉลี่ย พบว่าความสอดคล้องของผลการตรวจให้คะแนน ทั้ง 3 สถานการณ์ 9 ข้อ อยู่ในระดับสอดคล้องต่ำไปจนถึงระดับดีมาก ทั้งกลุ่มผู้ตรวจให้คะแนนที่เป็นกลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก และกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก 2) ค่าความตรงตามเกณฑ์การให้ตรวจให้คะแนนโดยพิจารณาจากความเห็นพ้อง (rater agreement) ของผลการตรวจให้คะแนนของผู้ตรวจ (x) กับเกณฑ์มาตรฐาน (y) พบว่าในกลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก มีค่าความเห็นพ้อง ตั้งแต่ร้อยละ 14 ถึง ร้อยละ 84 และในกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอกมีค่าความเห็นพ้อง ตั้งแต่ ร้อยละ 27 ถึง ร้อยละ 89 3) ค่าสัมประสิทธิ์การสรุปอ้างอิงของรูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$p \times (i : r)$] มีค่าสูงกว่า รูปแบบการตรวจทุกข้อของผู้สอบทุกคน [$p \times i \times r$] ทั้งกลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก และกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก

คำสำคัญ: เกณฑ์การตรวจให้คะแนนแบบองค์รวม ดัชนีความเห็นพ้องของการตรวจ ทฤษฎีการสรุปอ้างอิง
ค่าความเชื่อมั่นของผลการตรวจ

Abstract

This research aimed 1) to study the inter-rater reliability of the raters' scoring criteria of the essay test for measuring scientific competence by Intra-Class Correlation: ICC under different raters' characteristics. 2) to compare the validity of the raters' scoring criteria with the holistic scoring rubric of the essay test by Rater Agreement under different raters' characteristics. And 3) to compare the G-coefficient, under different raters' characteristics, for Cross design [$p \times i \times r$] and Nested design [$p \times (i : r)$]. The sample was divided into 2 groups: the group of 100 Grade 9 students and the group of raters. The group of raters comprised 3 raters who were science majors and another 3 raters who were non-science majors. The research instruments were 1) an essay test to measure the scientific competency of grade 9 students in 3 situations, containing 9 questions; and 2) a holistic scoring rubric. Generalizability Coefficient scores were analyzed by EduG. The research findings were 1) the reliability of the scoring results for each item, analyzed by the Intra-Class Correlation (ICC) statistics, were found to be from low to very good for all of the raters in the group, both science majors and non-science majors. 2) the validity of the raters' scoring criteria for each item analyzed by Rater Agreement between the raters' scores (x) and standard scores (y) revealed that the 3 raters who were science majors had the agreement index from 14 percent to 84 percent, and the 3 raters who were non-science majors had the agreement index from 27 percent to 89 percent. 3) The Generalizability Coefficient scores of the $p \times (i : r)$ design was higher than the $p \times i \times r$ design for all raters in the group.

Keywords: holistic scoring rubric, rater agreement, G Theory, inter-rater reliability

บทนำ

การนำประเทศเข้าสู่สังคมโลกในศตวรรษที่ 21 นั้น จำเป็นอย่างยิ่งที่จะต้องเตรียมความพร้อมในเรื่องของทรัพยากรมนุษย์ ให้มีทั้งความรู้ ความสามารถ และสมรรถนะที่จำเป็นให้สามารถปรับตัว และรู้เท่าทันต่อกระแสความเปลี่ยนแปลงของโลก (Ministry of Education, 2017) การจัดการศึกษาจะต้องพัฒนาและสามารถสร้างเสริมทักษะสำคัญในศตวรรษที่ 21 ซึ่งหนึ่งในสิ่งที่ผู้เรียนจำเป็นต้องมี คือ การรู้เรื่องทางวิทยาศาสตร์ (scientific literacy) The Institute for the Promotion of Teaching Science and Technology, (2020) ได้กล่าวถึง บุคคลที่ได้ชื่อว่าฉลาดรู้ด้านวิทยาศาสตร์ (scientifically literate person) จะต้องเป็นผู้ที่สามารถใช้สมรรถนะทางวิทยาศาสตร์ ได้แก่ 1) การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ (explain phenomena scientifically) 2) การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ (evaluate and design scientific enquiry) และ 3) การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ (interpret data and evidence scientifically) ดังนั้น ในการจัดการเรียนรู้ จึงจำเป็นที่จะต้องมีการวัดและประเมินผล เพื่อให้ทราบว่าผู้เรียนเกิดการเรียนรู้ตามวัตถุประสงค์ของการจัดกิจกรรมการเรียนรู้หรือไม่ (Tuksino, 2013)

แบบสอบที่จะวัดสมรรถนะทางวิทยาศาสตร์ของผู้เรียนได้ดี ควรจะเลือกใช้แบบสอบที่กระตุ้นให้ผู้เรียนได้แสดงความคิดความเข้าใจ และความรู้ความสามารถของผู้เรียนผ่านการเขียนแสดงวิธีการคิดและหาคำตอบ จึงจำเป็นต้องมีการสร้างแบบสอบอัตนัย (essay test) เพื่อเปิดโอกาสให้ผู้เรียนได้แสดงออกอย่างหลากหลายทั้งในด้านของวิธีการแก้ปัญหาการสื่อสารวิธีการแก้ปัญหาของผู้เรียน การให้เหตุผล รวมทั้งการแสดงความคิดสร้างสรรค์ แต่แบบสอบอัตนัย (essay test) ยังคงมีปัญหาในเรื่องของความเที่ยงต่ำ สิ้นเปลืองเวลา แรงงาน และค่าใช้จ่ายในการตรวจมาก (Coffman, 1971; Chiu & Wolfe, 2002) และจุดอ่อนที่สำคัญที่สุดคือเรื่องของการตรวจให้คะแนน (Pinyoanuntapong, 2004) การตรวจให้คะแนนจึงไม่ควรใช้ผู้ตรวจคนเดียว เพราะจะทำให้ขาดความเที่ยงในการให้คะแนน ดังที่ผลการวิจัย พบว่า เมื่อเพิ่มจำนวนผู้ตรวจ จะทำให้ความเที่ยงสูงขึ้น (Brennan & Johnson, 1995; Swartz et al., 1999) และควรควบคุมแหล่งความคลาดเคลื่อนอื่น เช่น ข้อสอบ ผู้ตรวจ (Phadungphon, 2017) การวัดและประเมินผู้เรียนด้วยแบบสอบอัตนัย (essay test) ยังคงมีปัญหาในเรื่องของความเที่ยงของคะแนนที่ได้ ซึ่งจะขึ้นอยู่กับผู้ตรวจเป็นหลัก และปัจจุบันปัญหาของระบบการศึกษาไทย ยังคงมีปัญหาในเรื่องของอัตราค่าจ้างการบรรจุครู ทำให้โรงเรียนขนาดเล็กจำนวนมาก มีปัญหาในเรื่องของการขาดแคลนครู ส่งผลให้ภายในโรงเรียนนั้น มีครูที่สอนไม่ครบชั้น และสอนไม่ตรงตามคุณวุฒิ หรือวิชาเอกที่จบมา ทำให้ในการวัดประเมินสมรรถนะทางวิทยาศาสตร์ ยังคงมีคุณครูที่สอนไม่ตรงตามคุณวุฒิ หรือวิชาเอกที่จบมา มาเป็นผู้ตรวจแบบสอบ และทำให้ค่าความเที่ยงของการให้คะแนนต่ำ (Phusing, 2020) ดังนั้น เพื่อให้การตรวจมีความเที่ยงเพิ่มมากขึ้น จะต้องมีการออกแบบวิธีการตรวจให้คะแนน และเกณฑ์การให้คะแนนที่เหมาะสม เพื่อให้เกิดความเป็นปรนัยในการตรวจ และมีความเที่ยงเพิ่มมากขึ้น (Umnacil, 2014) และนำเอาทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดมาช่วยในการประเมินค่าสัมประสิทธิ์การสรุปอ้างอิงคะแนนของแบบสอบ (Kanjawasee, 2007) โดยวิธีการออกแบบฟาเซตซึ่งฟาเซตที่นิยมนำมาออกแบบ คือ 2 ฟาเซต ได้แก่ จำนวนข้อสอบ (i) และจำนวนผู้ตรวจ (r) และ

มีวิธีการออกแบบการตรวจให้คะแนน 2 ลักษณะ คือ ผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน (two-facet crossed design) หรือ $p \times i \times r$ และ ผู้ตรวจตรวจข้อสอบบางข้อของผู้สอบทุกคน (two-facet nested design) หรือ $p \times (i : r)$ ซึ่งผลการวิจัยพบว่า เมื่อเพิ่มจำนวนผู้ตรวจจะทำให้ความเที่ยงสูงขึ้น (Brennan & Johnson, 1995; Swartz et al., 1999; Intanate, 2011; Sanguanwai, 2015)

ดังนั้น ผู้วิจัยจึงได้สนใจที่จะพัฒนาเกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) เพื่อใช้ในการตรวจให้คะแนนแบบสอบอัตนัยสำหรับการสมรรถนะทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และศึกษาเกี่ยวกับการเปรียบเทียบความเที่ยงและความตรงตามเกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) ของแบบสอบอัตนัยสำหรับการสมรรถนะทางวิทยาศาสตร์ ภายใต้คุณลักษณะของผู้ตรวจที่แตกต่างกัน คือ กลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก เป็นครูผู้สอนในรายวิชาวิทยาศาสตร์ และกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก ที่เป็นครูผู้สอนในรายวิชาอื่น ๆ ได้แก่ ครูผู้สอนรายวิชาภาษาไทย ครูผู้สอนรายวิชาคณิตศาสตร์ และครูผู้สอนรายวิชาสังคมศึกษา โดยประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ซึ่งเป็นทฤษฎีทางสถิติสำหรับวิเคราะห์ความน่าเชื่อถือของผลการวัดในสถานการณ์การวัดในลักษณะต่าง ๆ ที่เป็นเป้าหมายของการนำเครื่องมือไปใช้ เพื่อนำผลที่ได้ไปใช้ตัดสินใจการสรุปอ้างอิงการตรวจแบบสอบอัตนัยสำหรับวัดสมรรถนะทางวิทยาศาสตร์ ว่าควรผู้ตรวจที่มีคุณลักษณะที่แตกต่างกัน เมื่อใช้เกณฑ์การตรวจให้คะแนนแบบองค์รวม ในตรวจให้คะแนนแบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ ควรใช้รูปแบบการตรวจให้คะแนนแบบใด จึงจะมีค่าความเที่ยงของการให้คะแนนที่เหมาะสม และเพื่อให้ได้แบบสอบและเกณฑ์การให้คะแนนของแบบสอบอัตนัยสำหรับการวัดสมรรถนะทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีมาตรฐาน สามารถนำไปใช้วัดสมรรถนะทางวิทยาศาสตร์ของผู้เรียนได้จริง

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาค่าความสอดคล้องของผลการตรวจให้คะแนนแบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ ภายใต้คุณลักษณะของผู้ตรวจที่แตกต่างกัน
2. เพื่อเปรียบเทียบความตรงตามเกณฑ์การให้คะแนนแบบองค์รวม ของแบบสอบอัตนัยสำหรับวัดสมรรถนะทางวิทยาศาสตร์ ภายใต้คุณลักษณะของผู้ตรวจที่แตกต่างกัน
3. เพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง ภายใต้คุณลักษณะของผู้ตรวจที่แตกต่างกัน ที่มีรูปแบบการตรวจ 2 รูปแบบ ได้แก่ รูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน $[(p \times (i : r))]$ และรูปแบบการตรวจทุกข้อของผู้สอบทุกคน $[p \times i \times r]$

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการศึกษาแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มผู้สอบและกลุ่มผู้ตรวจให้คะแนน

- 1) กลุ่มผู้สอบ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2564 สังกัดสำนักงานเขตพื้นที่การศึกษา มัธยมศึกษาขอนแก่น จำนวน 84 โรงเรียน มีจำนวนนักเรียนทั้งสิ้น 10,575 คน
- 2) กลุ่มผู้ตรวจให้คะแนน

เป็นคุณครูผู้ที่มีใบประกอบวิชาชีพครูและมีวุฒิการศึกษาระดับปริญญาตรี มีจำนวนทั้งสิ้น 9 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัยแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มผู้สอบและกลุ่มผู้ตรวจให้คะแนนแบบสอบอัตนัย 1) กลุ่มผู้สอบ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2564 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาขอนแก่น ได้มาจากการใช้การสุ่มหลายขั้นตอน (multistage random sampling) กล่าวคือ ขั้นที่ 1. แบ่งโรงเรียนโดยใช้ขนาดของโรงเรียนเป็นเกณฑ์ ได้แก่ โรงเรียนขนาดเล็ก กลาง ใหญ่ และขนาดใหญ่พิเศษ ขั้นที่ 2 สุ่มโรงเรียน โดยการสุ่มอย่างง่าย (simple random sampling) เลือกตัวแทนโรงเรียนแต่ละขนาด จะได้ตัวแทนจำนวน 4 โรงเรียน และขั้นที่ 3 ใช้การสุ่มผู้เรียนโดยการสุ่มอย่างง่าย (simple random sampling) โรงเรียนละ 25 คน รวมเป็นนักเรียนจำนวน 100 คน 2) ผู้ตรวจให้คะแนนแบบสอบอัตนัย จะแบ่งออกเป็น 2 กลุ่ม คือ 2.1) กลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก คือ คุณครูที่สอนในรายวิชาวิทยาศาสตร์ จำนวน 3 คน โดยกำหนดคุณสมบัติ คือ เป็นผู้ที่มีใบประกอบวิชาชีพครูและมีวุฒิการศึกษาระดับปริญญาตรีในสาขาของวิทยาศาสตร์ ได้แก่ คณะศึกษาศาสตร์ คณะครุศาสตร์ และคณะวิทยาศาสตร์ สาขาวิชาฟิสิกส์ เคมี ชีววิทยา และวิทยาศาสตร์ทั่วไป 2.2) กลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก คือ คุณครูที่สอนในรายวิชาอื่น ๆ จำนวน 3 คน โดยกำหนดคุณสมบัติ คือ เป็นผู้ที่มีใบประกอบวิชาชีพครูและมีวุฒิการศึกษาระดับปริญญาตรีในสาขาอื่น ๆ ที่ไม่ใช่สาขาของวิทยาศาสตร์ ซึ่งในงานวิจัยนี้ มีคุณครูที่สอนในรายวิชาภาษาไทย คณิตศาสตร์ และสังคมศึกษา โดยใช้วิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง (purposive sampling) เนื่องจากงานวิจัยในครั้งนี้ต้องใช้ครูผู้สอนทั้งที่มีคุณวุฒิตรงเอกและคุณวุฒิไม่ตรงเอกในการตรวจให้คะแนนแบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ โดยใช้เกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) ซึ่งต้องใช้เวลา และความเต็มใจในการตรวจให้คะแนนของนักเรียน

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีจำนวน 3 สถานการณ์ 9 ข้อ โดยประกอบด้วย 1) สถานการณ์ผลกระทบของแผ่นดินไหว ที่ส่งผลกระทบต่อระบบนิเวศ 2) สถานการณ์ของการดำรงชีวิตของพืชและสิ่งแวดล้อม 3) สถานการณ์ของประโยชน์และโทษของพันธุวิศวกรรม โดยทั้ง 3 สถานการณ์จะมีข้อย่อย จำนวน 3 ข้อย่อย ที่กระตุ้นให้ผู้เรียนแสดงออกถึงความรู้ด้านเนื้อหาวิทยาศาสตร์ตามกรอบการประเมินสมรรถนะ 3 ด้าน ของ PISA ได้แก่ การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์, การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และการแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ จากนั้นนำไปตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ และนำแบบสอบดังกล่าวไปทดลองใช้ (try out) กับกลุ่มทดลองซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 50 คน และตรวจให้คะแนนโดยผู้ตรวจให้คะแนน จำนวน 3 คน เพื่อตรวจสอบคุณภาพของแบบสอบ พบว่า ค่าความตรงเชิงเนื้อหา (IOC) มีค่าตั้งแต่ 0.80 ถึง 1.00 ดัชนีค่าความยากง่าย (p) มีค่าตั้งแต่ 0.42 ถึง 0.74 และมีค่าอำนาจจำแนก (r) มีค่าตั้งแต่ 0.50 ถึง 0.86 และ เกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) โดยมีการพัฒนาและกำหนดประเด็นการประเมิน 3 ประเด็น ตามกรอบการประเมินสมรรถนะของ PISA ได้แก่ 1. การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ 2. การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และ 3. การแปลความหมาย

ข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ และมีคะแนนประเด็นละ 3 คะแนน รวมข้อสอบทั้งฉบับมีคะแนนเต็ม 27 คะแนน และจากนั้นนำไปตรวจสอบความตรงเชิงเนื้อหา (IOC) พบว่า มีค่าตั้งแต่ 0.80 ถึง 1.00

3. การเก็บรวบรวมข้อมูล

ผู้วิจัยนำแบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ผ่านการปรับปรุงแก้ไขแล้ว ไปเก็บข้อมูลกับกลุ่มตัวอย่าง จำนวน 100 คน แล้วนำผลการตอบมาตรวจให้คะแนนโดยใช้เกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) และใช้กลุ่มผู้ตรวจให้คะแนนจำนวน 2 กลุ่ม ได้แก่ กลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก จำนวน 3 คน และกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก จำนวน 3 คน จากนั้นนำผลการตรวจให้คะแนนมาวิเคราะห์ข้อมูล ได้แก่ 1. วิเคราะห์ค่าความสอดคล้องของผลการตรวจให้คะแนน โดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) 2. วิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิง และ 3. วิเคราะห์ค่าความตรงตามเกณฑ์การให้ตรวจให้คะแนนโดยพิจารณาจากความเห็นพ้อง (rater agreement) ของผลการตรวจให้คะแนนของผู้ตรวจ (x) กับเกณฑ์มาตรฐาน (y) โดยใช้โปรแกรมสำเร็จรูป SPSS โดยนำผลการตรวจให้คะแนนโดยใช้เกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) มาวิเคราะห์หาค่าความตรงของเกณฑ์การให้ตรวจให้คะแนนโดยพิจารณาจากความเห็นพ้อง (rater agreement) ของผลการตรวจให้คะแนนของผู้ตรวจ (x) กับเกณฑ์มาตรฐาน (y) โดยคะแนนเกณฑ์มาตรฐาน (y) ได้มาจากคะแนนชั้นตามติ จากผู้เชี่ยวชาญที่มีคุณสมบัติ คือ ครูที่มีวิทยฐานะชำนาญการพิเศษ และจบการศึกษาในระดับปริญญาโท สาขาทางวิทยาศาสตร์ และสาขาการวัดและประเมินผล จำนวน 5 คน และตรวจให้คะแนนโดยใช้เกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) และ 3. วิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ โดยใช้โปรแกรม EduG ในการตรวจให้คะแนนของผู้ตรวจที่มีคุณวุฒิตรงเอก และผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก ที่มีรูปแบบการตรวจ 2 รูปแบบ ได้แก่ รูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน $[p \times (i : r)]$ และ รูปแบบการตรวจทุกข้อของผู้สอบทุกคน $[p \times i \times r]$

4. การวิเคราะห์ข้อมูล

1) การวิเคราะห์ค่าความสอดคล้องของผลการตรวจให้คะแนนแบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ โดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) โดยใช้โปรแกรมสำเร็จรูป SPSS

2) การวิเคราะห์ค่าความตรงตามเกณฑ์การให้ตรวจให้คะแนนโดยพิจารณาจากความเห็นพ้อง (rater agreement) ของผลการตรวจให้คะแนนของผู้ตรวจ (x) กับเกณฑ์มาตรฐาน (y) โดยใช้โปรแกรมสำเร็จรูป SPSS

3) การวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ โดยใช้โปรแกรม EduG ในการตรวจให้คะแนนของผู้ตรวจที่มีคุณวุฒิตรงเอก และผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก ที่มีรูปแบบการตรวจ 2 รูปแบบ ได้แก่ รูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน $[p \times (i : r)]$ และ รูปแบบการตรวจทุกข้อของผู้สอบทุกคน $[p \times i \times r]$

ผลการวิจัย

1. ผลการวิเคราะห์ค่าความสอดคล้องของผลการตรวจให้คะแนนแบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ โดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC)

ผู้วิจัยได้ทำการวิเคราะห์ค่าความสอดคล้องของการตรวจให้คะแนน เพื่อวิเคราะห์หาความเที่ยงของการตรวจให้คะแนนระหว่างผู้ตรวจ (inter rater reliability) ตั้งแต่ 2 คน ขึ้นไป โดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) โมเดลอิทธิพลผสมแบบสองทาง (two-way mixed effects model) ที่ช่วงความเชื่อมั่น 95% (confidence interval) โดยใช้ผลการตรวจให้คะแนนจากกลุ่มผู้ตรวจที่มีคุณสมบัติตรงเอก และกลุ่มผู้ตรวจที่มีคุณสมบัติไม่ตรงเอก ที่ตรวจด้วยรูปแบบการตรวจทุกข้อของผู้สอบทุกคน มาวิเคราะห์ด้วยโปรแกรมสำเร็จรูป SPSS เพื่อพิจารณาโดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) และใช้เกณฑ์การแปลความหมายของ Koo & Li 2016 ซึ่งผลการวิเคราะห์ปรากฏดังตาราง 1 และตาราง 2

ตาราง 1 ค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) ของการตรวจให้คะแนนแบบสอบโดยกลุ่มผู้ตรวจที่มีคุณสมบัติตรงเอก

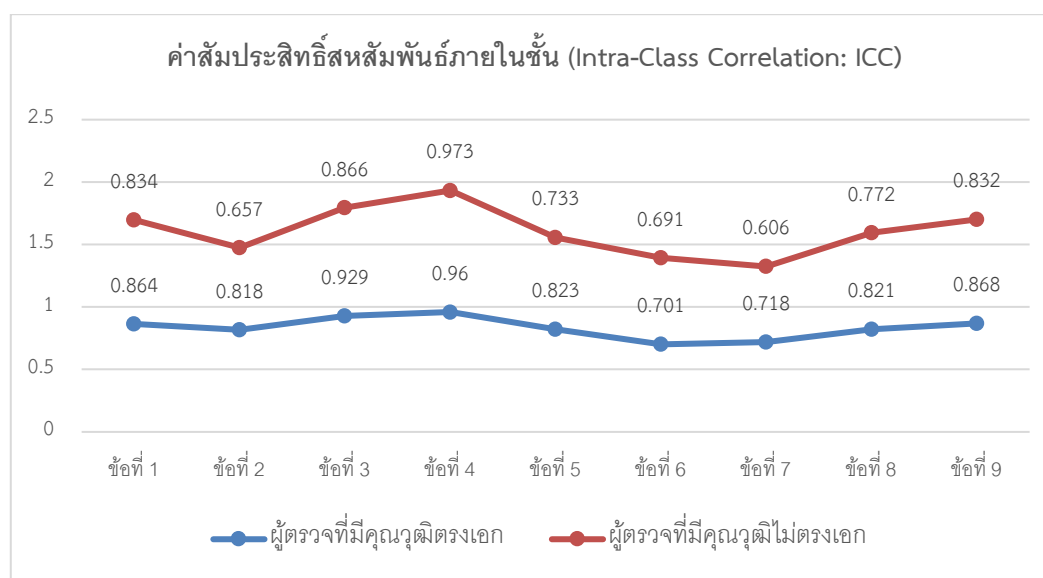
สถานการณ์	ข้อ	รายบุคคล		โดยเฉลี่ย	
		ICC (95% CI)	ระดับเกณฑ์	ICC (95% CI)	ระดับเกณฑ์
1	1	0.613 (0.523-0.699)	พอใช้	0.864 (0.814-0.903)	ดี
	2	0.529 (0.431-0.626)	พอใช้	0.818 (0.752-0.870)	ดี
	3	0.767 (0.701-0.824)	ดี	0.929 (0.904-0.949)	ดีมาก
2	4	0.857 (0.812-0.894)	ดี	0.960 (0.945-0.971)	ดีมาก
	5	0.538 (0.441-0.634)	พอใช้	0.823 (0.759-0.874)	ดี
	6	0.369 (0.266-0.479)	ต่ำ	0.701 (0.592-0.786)	พอใช้
3	7	0.389 (0.285-0.497)	ต่ำ	0.718 (0.615-0.798)	พอใช้
	8	0.534 (0.436-0.630)	พอใช้	0.821 (0.756-0.872)	ดี
	9	0.621 (0.532-0.706)	พอใช้	0.868 (0.820-0.906)	ดี

จากตาราง 1 พบว่า ค่าความสอดคล้องในการตรวจให้คะแนนระหว่างผู้ตรวจ กลุ่มผู้ตรวจที่มีคุณสมบัติตรงเอก มีความสอดคล้องกันรายบุคคล อยู่ในระดับต่ำ พอใช้ และดี โดยพบว่า สถานการณ์ที่ 2 ข้อย่อยที่ 1 มีค่า ICC_(2,1) มากที่สุด คือ มีค่าตั้งแต่ 0.812 ถึง 0.894 และ สถานการณ์ที่ 2 ข้อย่อยที่ 3 มีค่า ICC_(2,3) ต่ำที่สุด คือ มีค่าตั้งแต่ 0.266 ถึง 0.479 ส่วนค่าความสอดคล้องในการตรวจให้คะแนนระหว่างผู้ตรวจ กลุ่มผู้ตรวจที่มีคุณสมบัติตรงเอก มีความสอดคล้องกันโดยเฉลี่ย อยู่ในระดับพอใช้ ดี และดีมาก โดยพบว่า สถานการณ์ที่ 2 ข้อย่อยที่ 1 มีค่า ICC_(2,1) มากที่สุด คือ มีค่าตั้งแต่ 0.945 ถึง 0.971 และ สถานการณ์ที่ 2 ข้อย่อยที่ 3 มีค่า ICC_(2,3) ต่ำที่สุด คือ มีค่าตั้งแต่ 0.592 ถึง 0.786

ตาราง 2 ค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) ของการตรวจให้คะแนนแบบสอบโดยกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก

สถานการณ์	ข้อ	รายบุคคล		โดยเฉลี่ย	
		ICC (95% CI)	ระดับเกณฑ์	ICC (95% CI)	ระดับเกณฑ์
1	1	0.557 (0.461-0.651)	พอใช้	0.834 (0.774-0.882)	ดี
	2	0.324 (0.222-0.436)	ต่ำ	0.657 (0.533-0.755)	พอใช้
	3	0.618 (0.529-0.703)	พอใช้	0.866 (0.818-0.905)	ดี
2	4	0.900 (0.867-0.927)	ดีมาก	0.973 (0.963-0.981)	ดีมาก
	5	0.407 (0.304-0.514)	ต่ำ	0.733 (0.636-0.809)	พอใช้
	6	0.359 (0.256-0.469)	ต่ำ	0.691 (0.579-0.779)	พอใช้
3	7	0.278 (0.177-0.390)	ต่ำ	0.606 (0.463-0.719)	พอใช้
	8	0.458 (0.356-0.562)	ต่ำ	0.772 (0.689-0.837)	ดี
	9	0.554 (0.457-0.647)	พอใช้	0.832 (0.771-0.880)	ดี

จากตาราง 2 พบว่า ค่าความสอดคล้องในการตรวจให้คะแนนระหว่างผู้ตรวจ กลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก มีความสอดคล้องกันรายบุคคล อยู่ในระดับต่ำ พอใช้ และดีมาก โดยพบว่า สถานการณ์ที่ 2 ข้อย่อยที่ 1 มีค่า ICC_(2,1) มากที่สุด คือ มีค่าตั้งแต่ 0.867 ถึง 0.927 และ สถานการณ์ที่ 3 ข้อย่อยที่ 1 มีค่า ICC_(3,1) ต่ำที่สุด คือ มีค่าตั้งแต่ 0.177 ถึง 0.390 ส่วนค่าความสอดคล้องในการตรวจให้คะแนนระหว่างผู้ตรวจกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก มีความสอดคล้องกันโดยเฉลี่ย อยู่ในระดับพอใช้ ดี และดีมาก โดยพบว่า สถานการณ์ที่ 2 ข้อย่อยที่ 1 มีค่า ICC_(2,1) มากที่สุด คือ มีค่าตั้งแต่ 0.963 ถึง 0.981 และ สถานการณ์ที่ 3 ข้อย่อยที่ 1 มีค่า ICC_(3,1) ต่ำที่สุด คือ มีค่าตั้งแต่ 0.463 ถึง 0.719



ภาพ 1 เปรียบเทียบค่าความสอดคล้องของผลการตรวจให้คะแนน โดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) ของกลุ่มผู้ตรวจที่มีคุณลักษณะที่แตกต่างกัน

จากภาพ 1 เมื่อพิจารณาค่าความสอดคล้องของผลการตรวจให้คะแนน โดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) ของผู้ตรวจที่มีคุณวุฒิตรงเอก พบว่ามีค่าระหว่าง 0.701 – 0.960 และค่า ICC ของผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก พบว่ามีค่าอยู่ระหว่าง 0.606 – 0.973

2. ผลการวิเคราะห์ค่าความตรงตามเกณฑ์การให้ตรวจให้คะแนนโดยพิจารณาจากความเห็นพ้อง (rater agreement) ของผลการตรวจให้คะแนนของผู้ตรวจ (x) กับเกณฑ์มาตรฐาน (y)

ผู้วิจัยได้ใช้โปรแกรม SPSS วิเคราะห์หาค่าความตรงตามเกณฑ์การให้ตรวจให้คะแนนโดยพิจารณาจากความเห็นพ้อง (rater agreement) ของผลการตรวจให้คะแนนของผู้ตรวจที่มีคุณลักษณะที่แตกต่างกัน (x) กับเกณฑ์มาตรฐาน (y) โดยคะแนนเกณฑ์มาตรฐาน (y) ได้มาจากคะแนนฉันทามติ จากผู้เชี่ยวชาญที่มีคุณสมบัติ คือ ครูที่มีวิทยฐานะชำนาญการพิเศษ และจบการศึกษาในระดับปริญญาโท สาขาทางวิทยาศาสตร์ และสาขาการวัดและประเมินผล จำนวน 5 คน และตรวจให้คะแนนโดยใช้เกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) จากนั้นนำคะแนนที่ได้จากผู้เชี่ยวชาญมาคำนวณหาค่าความเห็นพ้อง (rater agreement) กับคะแนนจากกลุ่มผู้ตรวจให้คะแนน จากสูตร ดังนี้ (Gronlund, 2003 cited in Tuksino, 2013)

$$\text{ร้อยละความเห็นพ้อง} = \frac{\text{จำนวนผู้สอบที่ได้คะแนนเท่ากันจากผู้ตรวจ 2 คน}}{\text{จำนวนผู้สอบทั้งหมด}} \times 100 =$$

ซึ่งผลการวิเคราะห์ปรากฏดังตาราง 3

ตาราง 3 ค่าความตรงตามเกณฑ์การให้ตรวจให้คะแนนโดยพิจารณาจากความเห็นพ้อง (rater agreement) ของผลการตรวจให้คะแนนของผู้ตรวจที่มีคุณลักษณะที่แตกต่างกัน (x) กับเกณฑ์มาตรฐาน (y) (ร้อยละ)

สถานการณ์	ข้อ	ผู้ตรวจที่มีคุณวุฒิตรงเอก			ผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก		
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 1	คนที่ 2	คนที่ 3
1	1	50	73	60	35	52	27
	2	44	63	49	38	42	38
	3	73	81	53	64	75	59
2	4	81	84	70	84	88	89
	5	33	65	55	63	60	71
	6	18	71	39	63	52	74
3	7	66	63	50	75	75	69
	8	29	62	40	70	42	63
	9	46	55	40	38	41	46

จากตาราง 3 ค่าความเห็นพ้อง (rater agreement) ของผู้ตรวจ (x) กับเกณฑ์มาตรฐาน (y) พบว่า คุณลักษณะของผู้ตรวจที่แตกต่างกัน ส่งผลให้ค่าความเห็นพ้องในแต่ละข้อคำถามแตกต่างกัน โดยผู้ตรวจที่มี คุณวุฒิตรงเอก คนที่ 2 มีค่าความเห็นพ้องกับเกณฑ์มาตรฐาน (y) สูงกว่าผู้ตรวจคนอื่น ๆ ในข้อคำถามย่อยที่ 1, 2, 3, 5, และ 9 รองลงมาคือผู้ตรวจที่มีคุณวุฒิที่ไม่ตรงเอก คนที่ 3 มีค่าความเห็นพ้องกับเกณฑ์มาตรฐาน (y) สูงกว่าผู้ตรวจคนอื่น ๆ ในข้อคำถามย่อยที่ 4 และ 6 ส่วนของผู้ตรวจที่มีค่าความเห็นพ้องกับเกณฑ์ มาตรฐาน (y) ต่ำ ได้แก่ ผู้ตรวจที่มีคุณวุฒิตรงเอก คนที่ 1 โดยมีค่าความเห็นพ้องต่ำในข้อคำถามย่อยที่ 5, 6 และ 8 ผู้ตรวจที่มีคุณวุฒิตรงเอก คนที่ 3 โดยมีค่าความเห็นพ้องต่ำในข้อคำถามย่อยที่ 3, 4 และ 7 ผู้ตรวจที่มี คุณวุฒิไม่ตรงเอก คนที่ 1 โดยมีค่าความเห็นพ้องต่ำในข้อคำถามย่อยที่ 2 และ 9 และผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก คนที่ 3 โดยมีค่าความเห็นพ้องต่ำในข้อคำถามย่อยที่ 1 และ 2

3. ผลการวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงโดยใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือ ของผลการวัด (G-theory) โดยใช้โปรแกรม EduG

ผู้วิจัยได้นำผลการตรวจให้คะแนนแบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ของนักเรียน ชั้นมัธยมศึกษาปีที่ 3 โดยใช้เกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) โดยตรวจให้คะแนน จากกลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก และกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก มาเปรียบเทียบค่าความเที่ยง (สัมประสิทธิ์การสรุปอ้างอิง) ของการตรวจให้คะแนนแบบสอบ ผ่านรูปแบบการตรวจ 2 รูปแบบ ได้แก่ รูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$p \times (i:r)$] และ รูปแบบการตรวจทุกข้อของผู้สอบทุกคน [$p \times i \times r$] มีผลการวิเคราะห์ปรากฏดังตาราง 4

ตาราง 4 ผลการวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงโดยใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของ ผลการวัด (G-theory)

รูปแบบการตรวจ (design)	คุณลักษณะของผู้ตรวจ	
	ผู้ตรวจที่มีคุณวุฒิตรงเอก	ผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก
$p \times i \times r$ Design	ρ_{δ}^2	0.86
	ρ_{Δ}^2	0.80
$p \times (i:r)$ Design	ρ_{δ}^2	0.87
	ρ_{Δ}^2	0.40

จากตาราง 4 การเปรียบเทียบค่าความเที่ยง (สัมประสิทธิ์การสรุปอ้างอิง) ของการตรวจให้คะแนน แบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้เกณฑ์การให้คะแนน แบบองค์รวม (holistic scoring rubric) พบว่า รูปแบบการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบทุกข้อ ของผู้สอบทุกคน [$p \times i \times r$] เมื่อกำหนดคุณลักษณะของผู้ตรวจที่แตกต่างกัน กลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก จะมีค่าความเที่ยงของการตรวจให้คะแนนสูงกว่ากลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก โดยมีค่าสัมประสิทธิ์การสรุป อ้างอิงสำหรับตัดสินใจสัมพัทธ์ (ρ_{δ}^2) เท่ากับ 0.86 และมีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับตัดสินใจสมบูรณ์

(ρ_{Δ}^2) เท่ากับ 0.80 และหากตรวจให้คะแนนโดยรูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$(p \times i : r)$] กลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอกจะมีค่าความเที่ยงของการตรวจให้คะแนนใกล้เคียงกับกลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก โดยมีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับตัดสินใจสัมพัทธ์ (ρ_{δ}^2) เท่ากับ 0.87 และมีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับตัดสินใจสมบูรณ์ (ρ_{Δ}^2) เท่ากับ 0.40

อภิปรายผล

1. ผลการวิเคราะห์ค่าความสอดคล้องของผลการตรวจให้คะแนนแบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ โดยใช้เกณฑ์การให้คะแนนแบบองค์รวม (holistic scoring rubric) โดยใช้สถิติสหสัมพันธ์ภายในชั้น (Intra-Class Correlation: ICC) พบว่า ความสอดคล้องของผลการตรวจให้คะแนนแบบสอบอัตนัยทั้ง 3 สถานการณ์ 9 ข้อนั้น อยู่ในระดับสอดคล้องต่ำไปจนถึงระดับดีมาก แสดงให้เห็นว่า ผู้ตรวจให้คะแนนยังมีความเห็นไปในทิศทางที่ต่างกันบางสถานการณ์และข้อคำถาม โดยกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก มีค่าความสอดคล้องของผลการตรวจให้คะแนนโดยเฉลี่ย สูงกว่ากลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก ซึ่งคุณลักษณะที่แตกต่างกันในกลุ่มของผู้ตรวจให้คะแนน อาจเป็นสาเหตุที่ทำให้เกิดความคลาดเคลื่อนของการตรวจให้คะแนนสอดคล้องกับ Aphaikawi (2019) ที่กล่าวว่า สาเหตุของการให้คะแนนที่แตกต่างกัน อาจจะมาจากการคุณลักษณะของผู้ตรวจ เช่น ประสบการณ์ในการทำงาน การศึกษาตรงตามสาขาวิชา เป็นต้น ดังนั้นเพื่อลดความคลาดเคลื่อนที่อาจจะเกิดขึ้น จึงควรควบคุมแหล่งความคลาดเคลื่อนต่าง ๆ เช่น เพิ่มการฝึกอบรม เพื่อเพิ่มทักษะให้กับผู้ตรวจ คัดเลือกผู้ตรวจที่มีลักษณะความเข้มงวด ใจดี ใกล้เคียงกัน เพื่อให้ผลการตรวจมีค่าความสอดคล้องที่สูงขึ้น (Hopkins & Antes, 1990)

2. ผลการวิเคราะห์และเปรียบเทียบค่าความตรงของเกณฑ์การให้ตรวจให้คะแนน โดยพิจารณาจากความเห็นพ้อง (rater agreement) ของผลการตรวจให้คะแนนของผู้ตรวจ (x) กับเกณฑ์มาตรฐาน (y) ของกลุ่มผู้ตรวจที่มีคุณลักษณะที่แตกต่างกัน พบว่า ผู้ตรวจที่มีคุณวุฒิตรงเอก มีค่าความเห็นพ้องกับเกณฑ์มาตรฐาน (y) สูงกว่าผู้ตรวจคนอื่น ๆ ซึ่งแสดงให้เห็นว่า ผู้ตรวจที่มีคุณวุฒิตรงเอก มีความเชี่ยวชาญและมีประสบการณ์เดิมในด้านของเนื้อหาและการประเมินสมรรถนะในรายวิชาวิทยาศาสตร์ มากกว่าผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก จึงทำให้ผลของการตรวจให้คะแนนมีค่าความเห็นพ้องใกล้เคียงกับเกณฑ์มาตรฐาน (y) ดังนั้นหากต้องการให้ค่าความเห็นพ้องของกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอกมีค่าสูงมากขึ้น จึงควรมีเกณฑ์การให้คะแนนที่เหมาะสมกับผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก หรือต้องมีการจัดอบรมการใช้เกณฑ์การให้คะแนนเพื่อพัฒนาทักษะให้กับผู้ตรวจ (Aphaikawi, 2019)

3. ผลการวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงโดยใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของการวัด (G-theory) พบว่า กลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก เมื่อใช้รูปแบบการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน [$p \times i \times r$] มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับตัดสินใจสัมพัทธ์ (ρ_{δ}^2) และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับตัดสินใจสมบูรณ์ (ρ_{Δ}^2) เท่ากับ 0.00 ซึ่งถือว่าไม่มีค่าความเที่ยงเลย ทั้งนี้ เนื่องจากกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก มาจากครูผู้สอนรายวิชาภาษาไทย คณิตศาสตร์ และสังคมศึกษา ซึ่งมีธรรมชาติของรายวิชาที่แตกต่างกัน เมื่อตรวจให้คะแนนแบบสอบอัตนัยในรายวิชาวิทยาศาสตร์

ซึ่งผู้ตรวจไม่มีประสบการณ์ในด้านของการสอนรายวิชาวิทยาศาสตร์เลย จึงทำให้ความรู้พื้นฐานด้านเนื้อหาและความเห็นในเรื่องของการลงคะแนน มีความแตกต่างกัน หรือคุณลักษณะที่แฝงอยู่กับผู้ตรวจ เช่น การเข้มงวด/ใจดี ของผู้ตรวจแตกต่างกัน จึงส่งผลต่อการให้คะแนนมีความแตกต่างกัน ซึ่งสอดคล้องกับงานวิจัยของ Taoto (2016) ที่พบว่า ความแปรปรวนด้านผู้ตรวจให้คะแนนแบบสอบจะมีค่าน้อยสุด เนื่องมาจากมีการคัดเลือกผู้ตรวจที่มีคุณสมบัติเหมาะสม มีความรู้ความสามารถให้ด้านการสอนและตัวชี้วัดทางวิทยาศาสตร์ นอกจากนี้ยังมีการอบรมและฝึกปฏิบัติจริง ตลอดจนมีคู่มือการดำเนินเพื่อเพิ่มความชัดเจนในแนวทางการดำเนินงาน ดังนั้น หากต้องการให้กลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก มีผลการตรวจให้คะแนนที่มีค่าสัมประสิทธิ์การสรูปอ้างอิงที่สูงขึ้น จึงจำเป็นต้องมีการ คัดเลือกผู้ตรวจที่มีลักษณะความเข้มงวด ใจดี ใกล้เคียงกัน เพื่อให้ผลการตรวจมีค่าความเที่ยงของการตรวจให้คะแนนที่สูงขึ้น (Hopkins & Antes, 1990) และต้องมีการฝึกอบรม และฝึกปฏิบัติจริง เพื่อพัฒนาทักษะให้กับผู้ตรวจ (Aphaikawi, 2019) รวมทั้ง ต้องมีคู่มือเกณฑ์การให้คะแนนที่มีการกำหนดแนวทาง และวิธีการให้คะแนนอย่างชัดเจน (Welk et al., 2004)

4. ผลการวิเคราะห์ค่าสัมประสิทธิ์การสรูปอ้างอิงโดยใช้ทฤษฎีการสรูปอ้างอิงความน่าเชื่อถือของผลการวัด (G-theory) พบว่า รูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน $[(p \times (i : r))]$ มีค่าสัมประสิทธิ์การสรูปอ้างอิงสูงกว่า รูปแบบการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน $[p \times i \times r]$ ทั้งในกลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก และกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก เนื่องจากการตรวจข้อสอบทุกข้อของผู้สอบทุกคน อาจจะทำให้ผู้ตรวจเกิดอาการเมื่อยล้า หรืออาจจะมีข้อคำถามบางข้อ หรือบางสถานการณ์ที่ไม่ถนัด แต่หากเป็นการตรวจในรูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน ผู้ตรวจจะได้รับผิดชอบเฉพาะบางข้อเท่านั้นซึ่งจะทำให้ผู้ตรวจมีความเชี่ยวชาญในเนื้อหา และเกณฑ์การให้คะแนนมากยิ่งขึ้น สอดคล้องกับผลการวิจัยของ Kwanja, (2013); Sanguanwai, (2015); Saosin, (2018) ที่พบว่า แบบสอบที่มีการออกแบบการตรวจให้คะแนนแบบที่ผู้ตรวจ ตรวจให้คะแนนเฉพาะเหตุการณ์ของผู้สอบทุกคน มีค่าความเที่ยงสูงกว่าการที่ผู้ตรวจให้คะแนนเหตุการณ์ทุกเหตุการณ์ของผู้สอบทุกคน

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1. แบบสอบอัตนัยวัดสมรรถนะทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ผู้วิจัยพัฒนาขึ้น สามารถนำไปใช้สำหรับวัดสมรรถนะทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ได้ โดยในการตรวจให้คะแนนแบบสอบอัตนัยโดยใช้เกณฑ์การตรวจให้คะแนนแบบองค์รวม (holistic scoring rubric) ที่ผู้วิจัยพัฒนาขึ้นนั้น สำหรับกลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก สามารถตรวจให้คะแนนได้ทั้งรูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน $[(p \times (r : i))]$ และ รูปแบบการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน $[p \times i \times r]$ เนื่องจากให้ค่าความเที่ยงของการให้คะแนนอยู่ในระดับที่เหมาะสม คือ มากกว่า 0.7 ทั้งสองรูปแบบการตรวจ แต่เพื่อให้ประหยัดเวลาในการตรวจให้คะแนน และลดความเหนื่อยล้าของผู้ตรวจ ควรตรวจให้คะแนนรูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน $[(p \times (r : i))]$ แต่ในกรณีที่ผู้ตรวจให้คะแนนเป็นกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก ให้ใช้รูปแบบการตรวจให้คะแนนที่เป็น

รูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน $[(p \times (i : r))]$ เท่านั้น เนื่องจากจะให้ค่าความเที่ยงของการตรวจให้คะแนนอยู่ในระดับสูง

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. จากการศึกษางานวิจัยในครั้งนี้อย่างถี่ถ้วน พบว่ากลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอก มีค่าความสอดคล้องของการตรวจให้คะแนนสูง ซึ่งกลุ่มผู้ตรวจที่มีคุณวุฒิตรงเอกเป็นคุณครูที่จบจากสาขา/วิชาเอกวิทยาศาสตร์เหมือนกัน แต่กลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก มีคุณครูที่จบจากสาขา/วิชาเอกคณิตศาสตร์ ภาษาไทย และสังคมศึกษา ซึ่งมีความแตกต่างกัน ดังนั้นในการศึกษาครั้งต่อไป ควรศึกษากลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอกที่เป็นคุณครูที่จบจากสาขา/วิชาเอกเดียวกัน แล้วนำผลการตรวจให้คะแนนมาเปรียบเทียบกับว่าส่งผลต่อคุณภาพของการตรวจให้คะแนนที่แตกต่างหรือไม่

2. ในการศึกษาค่าความสอดคล้องของผลการตรวจให้คะแนนในงานวิจัยนี้จะเห็นว่าค่าความสอดคล้องของผลการตรวจให้คะแนนในบางข้อยังอยู่ในระดับต่ำ ดังนั้นในการวิจัยครั้งต่อไป จึงควรมีการคัดเลือกคุณลักษณะของกลุ่มผู้ตรวจ เช่น ความใจดี ความเข้มงวด หรือ ควรมีการอบรมการใช้เกณฑ์การตรวจให้คะแนน และนำผลการตรวจให้คะแนนก่อนและหลังการอบรม มาเปรียบเทียบกับ เพื่อศึกษาหาความแปรปรวนของการตรวจให้คะแนนในครั้งต่อไป และควบคุมแหล่งความแปรปรวนที่อาจจะส่งผลต่อการตรวจให้คะแนนเหล่านั้น

References

- Brennan, R. L., & Johnson, E. G. (1995). Generalizability of Performance Assessments. *Journal of Educational Measurement*, 14(4), 9-12.
- Chiu, C., & Wolfe, E. (2002). A Method for Analyzing Sparse Data Matrices in the Generalizability Theory Framework. *SAGE Journal*, 26(3), 321-338.
- Coffman, W. E. (1971). On the Reliability of Ratings of Essay Examinations in English. *JSTOR Journal*, 5(1), 24-36.
- Hopkins, C. D., & Antes, R. L. (1990). *Classroom Measurement and Evaluation*. Peacock Press.
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155-163.
- Swartz, C. W., Hooper, S. R., Montgomery, J. W., Wakely, M. B., Kruif, R. E. L., Reed, M., Brown, T. T., Levine, M. D., & White, K. P. (1999). Using Generalizability Theory to Estimate the Reliability of Writing Scores Derived from Holistic and Analytical Scoring Methods. *Sage Journal*, 59(3), 492-506.
- Welk et al., (2004). Reliability of accelerometry-based activity monitors: a generalizability study. *Ovid Journal*, 36(9), 1637-1645.

Translate Thai References

- Aphaikawi, D. (2019). Scoring results of subjective exams when different groups of inspectors and examination patterns. *The 27th Thailand Measurement Evaluation and Research*, 108-124. (in Thai)
- Intanate, N. (2011). *Characteristic of the open-ended mathematics test scores for different numbers of raters and scoring patterns using generalizability model and many-facet Rasch model* [Doctoral dissertation]. Srinakharinwirot University. (in Thai)
- Kanjanawasee S. (2007). *Modern test theories*. Chulalongkorn University Press. (in Thai)
- Kwanja, N. (2013). *Comparison of summaries reference coefficients of the process skills scale Grade 4 science with different scoring patterns* [Master's thesis]. Mahasarakham University. (in Thai)
- Ministry of Education. (2017). *Thailand Education Plan B.E. 2560 - 2579 (A.D. 2017 – 2036)*. Office of the Education Council Press. (in Thai)
- Phadungphon, S. (2017). Comparison of reliability of modified essay question test for measuring the abilities in using scientific method in physic under different numbers of event and rater: an application of generalizability theory. *Educational Electronic Journal*, 12(4), 381-393. (in Thai)
- Phusing N. (2020). Scienceteacher development model throughstem education for the schools with non-science majoringteachers (nsmt). *Journal of MCU Ubon Review*, 5(3), 439-454. (in Thai)
- Pinyoanuntapong, B. (2004). *Measurement and evaluation*. Srinakharinwirot University Press. (in Thai)
- Sanguanwai, C. (2015). *Comparison of test reliability for Measuring Mathematical Creative problem-solving ability: Application of Generalizability theory* [Master's thesis]. Chulalongkorn University. (in Thai)
- Saosin, K. (2019). Comparison of Reliability of Math Problem Solving Proficiency Test with Sub-analytical Scoring At the lower secondary level: application of summary theory referring to the reliability of measurement results. *Educational Electronic Journal*, 13(3), 423-438. (in Thai)
- Taoto, J. (2016). A study of the confidence values of students' math subjective test scores. Secondary school with different number of examiners and scoring patterns using the theory of summaries. Reference. *Hat Yai Academic Journal*, 14(1), 1-14. (in Thai)

- The Institute for the Promotion of Teaching Science and Technology, (2020). *Scientific Literacy*. <https://pisathailand.ipst.ac.th/about-pisa/scientific-literacy/> (in Thai)
- Tuksino, P. (2013). *Teaching documentation educational research methodology*. Khon Kaen University Press. (in Thai)
- Umnacil, M. (2014). *Comparison of reliability of modified essay question test for measuring scientific problem-solving ability using different scoring methods under different number of events: an application of generalizability theory* [Master's thesis]. Chulalongkorn University. (in Thai)

การวิเคราะห์ปัจจัยด้านความเหลื่อมล้ำทางการศึกษาที่ส่งผลต่อความฉลาดรู้คณิตศาสตร์ ภายใต้โปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล (PISA) Analysis of Educational Inequality Factors Affecting Mathematical Literacy under the International Student Assessment Program (PISA)

มาริษา นามกุล^{1*} และ ประกฤติยา ทักชีโน²
Marisa Namkul^{1*} and Prakritiya Tuksino²

(Received: April 19, 2022; Revised: July 23, 2022; Accepted: August 1, 2022)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมายเพื่อวิเคราะห์ปัจจัยที่มีผลต่อความฉลาดรู้ทางคณิตศาสตร์ จากข้อมูลของโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากลหรือ PISA ปี 2018 ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิ ซึ่งเป็นฐานข้อมูลของโครงการประเมินผล PISA 2018 กลุ่มตัวอย่างได้แก่ นักเรียนที่มีอายุ 15 ปี ซึ่งเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 – 3 จำนวน 8,662 คน ในสถานศึกษา 290 แห่ง โดยวิเคราะห์คะแนนจากแบบประเมินความฉลาดรู้คณิตศาสตร์ ในโครงการประเมินผล PISA 2018 วิเคราะห์ข้อมูลด้วยการวิเคราะห์พหุระดับด้วยโปรแกรม HLM version 8.2

ผลการวิจัยพบว่า ปัจจัยที่มีผลต่อความฉลาดรู้คณิตศาสตร์ เมื่อนำตัวแปรมาจัดกลุ่มเชื่อมโยงกับฐานข้อมูลของ PISA 2018 สรุปตัวแปร 2 ระดับ ประกอบด้วย (1) ระดับสถานศึกษา ได้แก่ อัตราส่วนจำนวนนักเรียนต่อครู ขนาดห้องเรียน ดัชนีการขาดแคลนครู งบประมาณจากรัฐบาล ดัชนีทรัพยากรการเรียนรู้ สื่อการเรียนการสอน ประเภทสถานศึกษา และขนาดชุมชนที่ตั้งของสถานศึกษา (2) ระดับนักเรียน ได้แก่ เศรษฐฐานะของครอบครัว ดัชนีความมั่งคั่งของครอบครัว เจตคติในการเรียน จากการวิเคราะห์พหุระดับพบว่า ตัวแปรระดับสถานศึกษา ที่มีค่าสัมประสิทธิ์ถดถอยอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ ขนาดห้องเรียนที่มีนักเรียนจำนวน 21-25 คน ($\beta = 73.342$) งบประมาณจากรัฐบาล ($\beta = -0.747$) อัตราส่วนจำนวนนักเรียนต่อครู ($\beta = -1.791$) ประเภทของสถานศึกษาเอกชน ($\beta = -40.087$) ตัวแปรระดับนักเรียน ที่มีค่าสัมประสิทธิ์ถดถอยอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 คือ ทรัพยากรในบ้าน ($\beta = 16.318$) ความมั่งคั่งของครอบครัว ($\beta = -13.866$) การศึกษาสูงสุดของผู้ปกครองระดับหลังมัธยมศึกษาตอนอุดมศึกษา ($\beta = -15.934$)

คำสำคัญ: ความเหลื่อมล้ำทางการศึกษา ความฉลาดรู้ทางคณิตศาสตร์ คุณภาพการศึกษา

¹ นิสิตระดับปริญญาโท สาขาวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² ผู้ช่วยศาสตราจารย์ สาขาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

¹ Master's degree student, Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

² Assistant Professor, Department of Educational Measurement and Evaluation, Faculty of Education, Khon Kaen University

* Corresponding Author E-mail: asiram@kkumail.com

Abstract

The purpose of this research was to analyze factors affecting mathematical literacy. Based on data from the 2018 International Student Assessment Program (PISA), the data used in this research were secondary data in the database of the PISA 2018 Assessment Project. The sample consisted of 8,662 students 15 years of age and were in grade 8 and 9 in 290 educational institutions. In the PISA 2018 Assessment Program, the data were analyzed with multilevel statistical analysis using HLM version 8.2.

The results showed that the factors affecting mathematical literacy, when the variables were grouped and linked to the PISA 2018 database, 2-level variables were summarized: (1) at the educational institution level, they were: the ratio of the number of students to teachers, classroom size, teacher shortage Index, government budget, learning resource index, teaching and learning materials, type of the educational institution and the size of the community where the educational institution was situated; (2) at the student level, they were: family socioeconomic status, family wealth index, and attitude in learning. From the multilevel analysis of factors that caused educational inequality, it was found that the school-level variables which had statistical significance effect at the .05 level were: classroom size with 21-25 students ($\beta = 73.342$), government budget ($\beta = -0.747$), student-teacher ratio ($\beta = -1.791$), type of private educational institutions ($\beta = -40.087$); the student-level variables with a statistical significance effect at the .01 level were: property in the home ($\beta = 16.318$), family wealth ($\beta = -13.866$), the highest education of parents at the post-secondary level before tertiary education ($\beta = -15.934$).

Keywords: educational inequality, mathematical literacy, quality of education

บทนำ

โปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล (Programme for International Student Assessment หรือ PISA) ริเริ่มโดยองค์การเพื่อความร่วมมือทางเศรษฐกิจและการพัฒนา (Organization for Economic Co-operation and Development หรือ OECD) มีวัตถุประสงค์เพื่อประเมินคุณภาพของระบบการศึกษาของประเทศต่าง ๆ ในการเตรียมความพร้อมให้เยาวชนมีศักยภาพหรือความสามารถพื้นฐานที่จำเป็นต่อการดำรงชีวิตในโลกที่มีการเปลี่ยนแปลง โดย PISA เน้นการประเมินสมรรถนะของนักเรียนเกี่ยวกับการใช้ความรู้และทักษะในชีวิตจริงมากกว่าการเรียนรู้ตามหลักสูตรในโรงเรียน ปัจจุบันนี้มีประเทศจากทั่วโลกเข้าร่วม PISA มากกว่า 80 ประเทศ (OECD, 2019) โดยเฉพาะการวัดความฉลาดรู้ทางคณิตศาสตร์เป็นสิ่งที่จำเป็นอย่างมากเนื่องจากความสามารถทางคณิตศาสตร์เป็นสิ่งที่สะท้อนให้เห็นว่าผู้เรียนมีความรู้และทักษะที่เพียงพอหรือไม่

ต่อการดำรงชีวิตในศตวรรษที่ 21 ดังนั้นการเตรียมผู้เรียนให้พร้อมให้มีความสามารถทางคณิตศาสตร์ จึงเป็นสิ่งสำคัญซึ่งส่งผลให้ผู้เรียนรู้เท่าทันการเปลี่ยนแปลงของระบบเศรษฐกิจ สังคม วัฒนธรรม และสภาพแวดล้อม สามารถอยู่ร่วมกับประชาคมโลกได้ (Ministry of Education, 2017) และการประเมินความสามารถทางคณิตศาสตร์ถือเป็นสิ่งสำคัญในการขับเคลื่อนการศึกษา โดยเฉพาะอย่างยิ่ง การประเมินความสามารถทางคณิตศาสตร์ที่ไม่ได้มุ่งเน้นเพียงคำตอบที่ถูกต้องเท่านั้น แต่ยังมุ่งเน้นไปที่กระบวนการคิดทางคณิตศาสตร์ด้วยเช่นกัน (Chinjunthuk & Junpeng, 2020)

อีกทั้งแบบประเมินความฉลาดรู้ทางคณิตศาสตร์ที่วัดได้จากโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล (PISA) เป็นการสอบเพื่อประเมินคุณภาพของระบบการศึกษาของประเทศ ในการเตรียมความพร้อมให้เยาวชนมีศักยภาพหรือความสามารถพื้นฐานที่จำเป็นต่อการดำรงชีวิตในโลกที่มีการเปลี่ยนแปลงในอนาคต แต่ปัญหาการศึกษาและการทดสอบต่าง ๆ ของไทยมุ่งเน้นเพียงผลสัมฤทธิ์ของผู้เรียนเท่านั้น เพื่อที่จะบ่งชี้คุณภาพของการศึกษาไทยกับนานาชาติ ควรหาวิธีในการพัฒนา แก้ไข ปรับปรุง ให้การศึกษาไทยเกิดความก้าวหน้า จากการสอบวัดระดับความรู้นานาชาติ โปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล หรือ PISA โดยผลการประเมินความฉลาดรู้ทางคณิตศาสตร์ PISA ของประเทศไทยหลายสิบปีที่ผ่านมา ผลคะแนนยังต่ำกว่าคะแนนเฉลี่ยของประเทศในกลุ่ม OECD เช่นเดียวกันกับในปี 2018 ในการสอบวัดความฉลาดรู้ทางคณิตศาสตร์ ที่ยังมีผลคะแนนเฉลี่ยต่ำกว่าประเทศในกลุ่ม OECD และยังพบว่านักเรียนที่ต่างพื้นที่ ต่างกลุ่มโรงเรียน มีผลประเมินที่แตกต่างกัน นักเรียนจากพื้นที่กรุงเทพฯ และปริมณฑล มีคะแนนคณิตศาสตร์สูงกว่านักเรียนจากพื้นที่อื่น ๆ โดยมีคะแนนสูงกว่าพื้นที่ภาคเหนือตอนบน ภาคใต้ภาคตะวันตก และภาคอีสานตอนล่าง (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี) เนื่องจากความแตกต่างเหล่านี้นำให้เกิดความไม่เท่าเทียมจนอาจเกิดเป็นความเหลื่อมล้ำ ดังมีปรากฏอยู่ในทุกเรื่อง ทุกพื้นที่ ทุกภาคส่วน และทุกเวลา โดยเฉพาะในด้านของความเหลื่อมล้ำทางการศึกษา ที่เกิดจากความไม่เท่าเทียมกันของคุณภาพและมาตรฐานของการจัดการศึกษา มีผลต่อการได้รับโอกาสทางการศึกษา (Pharcharuen et al., 2021) อันมีสาเหตุมาจากสภาพเศรษฐกิจ ปัจจัยภูมิหลังทางครอบครัวและสังคมของผู้เรียน รวมถึงคุณภาพและมาตรฐานที่แตกต่างกันของสถานศึกษาต่าง ๆ

อย่างไรก็ตามรายงานผลการประเมินความฉลาดรู้ทางคณิตศาสตร์ในโครงการดังกล่าว เป็นเพียงผลคะแนนที่ได้จากการทำแบบสอบ และไม่ได้คำนึงถึงความแตกต่างของคุณลักษณะพื้นฐานในระดับนักเรียนระดับสถานศึกษา ตลอดจนถึงในระดับประเทศ ในการวิจัยครั้งนี้จึงคำนึงถึงข้อมูลที่เป็นโครงสร้างแบบลดหลั่นด้วยการวิเคราะห์ทุกระดับ (HLM) พิจารณาถึงปัจจัยอื่นที่นอกเหนือการควบคุมของสถานศึกษา ที่อาจก่อให้เกิดความแตกต่างต่อผลคะแนน เพื่อเป็นแนวทางของการประเมินคุณภาพการจัดการศึกษา (Tuksino, 2009) โดยคำนึงถึงปัจจัยพื้นฐานที่อาจจะมีแตกต่างกันในแต่ละสถานศึกษา อาจมีปัจจัยที่เกิดจากตัวแปรที่หลากหลายที่น่าสนใจมาวิเคราะห์ โดยจากการศึกษาและสังเคราะห์เอกสารงานวิจัย 30 เรื่อง ที่ศึกษาปัจจัยที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน ซึ่งเป็นตัวบ่งชี้ถึงคุณภาพการจัดการศึกษาของสถานศึกษาและเชื่อมโยงกับข้อมูลตัวแปร ของ PISA 2018 ที่มีการเก็บพื้นฐานข้อมูล (OECD, 2019) เมื่อนำมาพิจารณาตัวแปรในระดับสถานศึกษา ได้แก่ อัตราส่วนจำนวนนักเรียนต่อครู ขนาดห้องเรียน ดัชนีการขาดแคลนครู จำนวนงบประมาณ ดัชนีทรัพยากรการเรียน สื่อการเรียนการสอน ประเภทสถานศึกษา และขนาดชุมชนที่ตั้ง

ของสถานศึกษา ตัวแปรในระดับนักเรียน ได้แก่ เศรษฐฐานะของครอบครัว ดัชนีความมั่งคั่งของครอบครัว และ เจตคติในการเรียน

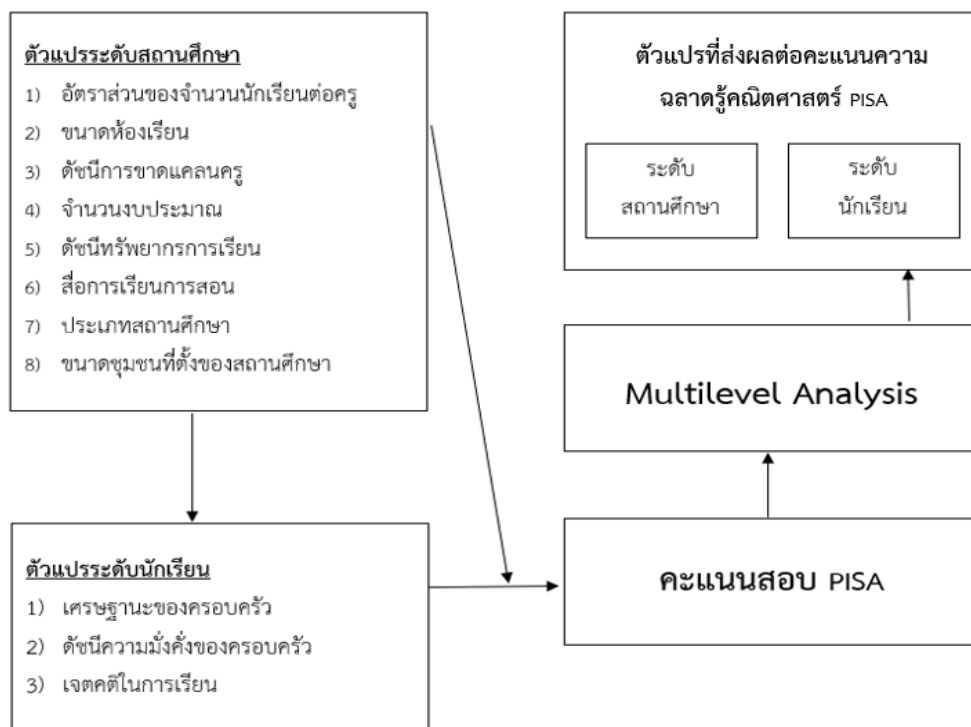
จากการศึกษาข้อมูลการรวบรวมข้อมูลของโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล (PISA) เป็นการรวบรวมข้อมูลจากนักเรียนจำนวนมาก ทั้งคุณลักษณะพื้นฐานที่แตกต่างกันในนักเรียนและสถานศึกษา ข้อมูลส่วนบุคคล ข้อมูลครอบครัว รวมทั้งข้อมูลของโรงเรียนที่เข้าร่วม โดยเป็นฐานข้อมูลที่มีความละเอียด อาจส่งผลนำมาซึ่งความแตกต่างในคะแนนผลการประเมินผลของนักเรียน ผู้วิจัยจึงมีความสนใจในการนำผลโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล PISA คะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ (Mathematical Literacy) มาใช้ในการวิเคราะห์ งานวิจัยนี้จึงต้องการวิเคราะห์ปัจจัยที่ส่งผลต่อคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ จากการวิเคราะห์พหุระดับ โดยผู้วิจัยคาดหวังว่าวิจัยฉบับนี้จะเป็นประโยชน์ต่อการกำหนดนโยบายทางการศึกษาของภาครัฐ ผู้บริหาร และหน่วยงานต่าง ๆ ที่เกี่ยวข้องในการพัฒนาการศึกษาของประเทศไทย เพื่อหาแนวทางการส่งเสริมให้ผู้เรียนมีการเรียนรู้ในระดับที่สูงขึ้น อีกทั้งเป็นข้อเสนอแนะเพื่อนำไปสู่การวางแผนการพัฒนาแก้ไขปัญหาดังกล่าวต่าง ๆ ทางการศึกษาต่อไป

ความมุ่งหมายของการวิจัย

การวิจัยนี้มีจุดประสงค์เพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อความฉลาดรู้คณิตศาสตร์ (mathematical literacy) โดยใช้การวิเคราะห์พหุระดับ (multilevel analysis) จากข้อมูลของโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล PISA ปี 2018

กรอบแนวคิดการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยสามารถสรุปกรอบแนวคิดการวิจัยได้ดังนี้



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

1. กลุ่มตัวอย่าง

กลุ่มตัวอย่างของโครงการเป็นนักเรียนที่มีอายุ 15 ปี ซึ่งเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 – 3 จำนวน 8,662 คน ในสถานศึกษา 290 แห่ง ที่ครอบคลุมสถานศึกษาในสังกัดของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน สำนักงานการศึกษาเอกชน สังกัดของกรุงเทพมหานคร สำนักงานคณะกรรมการการอุดมศึกษา สำนักบริหารการศึกษาท้องถิ่น และสำนักงานการอาชีวศึกษา

2. ข้อมูลที่ใช้ในการวิจัย

เป็นข้อมูลจากคะแนนจากแบบสอบที่ใช้ในการศึกษาเป็นคะแนนจากแบบประเมินความฉลาดรู้คณิตศาสตร์ (Mathematical Literacy) ซึ่งเป็นข้อมูลหุติยภูมิจากฐานข้อมูล (data base) ของโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล หรือ PISA ปี 2018

3. การเก็บรวบรวมข้อมูล

1) ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลหุติยภูมิ ซึ่งเป็นฐานข้อมูล (data base) ของโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล ปี 2561 PISA 2018 (IPST, 2021) โดยกลุ่มตัวอย่างการประเมินเป็นนักเรียนของประเทศไทยที่มีอายุ 15 ปี ซึ่งเป็นวัยที่จบการศึกษาภาคบังคับ โดยในทางปฏิบัติใช้นักเรียนที่มีช่วงอายุ 15 ปี 3 เดือน จนถึง 16 ปี 2 เดือน และสำหรับประเทศไทยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 – 3

ฐานข้อมูลของโครงการประเมินผล PISA ประกอบด้วยสถานศึกษาที่เป็นกลุ่มตัวอย่าง 290 แห่ง จำแนกตามสังกัดของสถานศึกษา และจะสุ่มนักเรียนในแต่ละสถานศึกษา ได้นักเรียนที่สุ่มได้ทั้งหมด 9,078 คน มีจำนวนนักเรียนที่เข้าสอบจริง 8,662 คน

2) ปัจจัยหรือตัวแปรนำมาใช้ในการวิจัยตัวแปรที่นำมาใช้ในการวิจัยครั้งนี้ ผู้วิจัยพิจารณาถึงโครงสร้างของข้อมูลในระบบการจัดการศึกษา ซึ่งลักษณะข้อมูลมีโครงสร้างแบบลดหลั่น นั่นคือนักเรียนสอดแทรกอยู่ในสถานศึกษา ดังนั้นจึงแบ่งตัวแปรเป็นระดับสถานศึกษาและระดับนักเรียน โดยตัวแปรที่สามารถเข้าถึงฐานข้อมูลในโครงการประเมินผล PISA 2018 และนำมาใช้ในการวิจัยมีดังนี้

ตัวแปรระดับสถานศึกษา ได้แก่ สัดส่วนของจำนวนนักเรียนต่อครู ขนาดห้องเรียน ดัชนีการขาดแคลนครู จำนวนงบประมาณ ดัชนีทรัพยากรการเรียนรู้ สื่อการเรียนการสอน ประเภทสถานศึกษา และขนาดชุมชนที่ตั้งของสถานศึกษา

ตัวแปรระดับนักเรียน ได้แก่ เศรษฐฐานะของครอบครัว ดัชนีความมั่งคั่งของครอบครัว เจตคติในการเรียน

4. การวิเคราะห์ข้อมูล

ผู้วิจัยมีขั้นตอนหลัก ขึ้นดังนี้

ขั้นที่ 1 โดยจากการศึกษาและสังเคราะห์เอกสารงานวิจัยที่ศึกษาเกี่ยวกับปัจจัยที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน ซึ่งเป็นตัวบ่งชี้ถึงคุณภาพการจัดการศึกษาของสถานศึกษาและเชื่อมโยงกับข้อมูลตัวแปรของ PISA 2018 ที่มีการเก็บในฐานข้อมูล โดยมีขั้นตอนในการสังเคราะห์ปัจจัย ดังนี้

- 1) ศึกษาเอกสารงานวิจัยที่ศึกษาเกี่ยวกับปัจจัยที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน
- 2) วิเคราะห์ตัวแปรโดยพิจารณาตามระดับตัวแปร คือ ระดับสถานศึกษา และระดับนักเรียน
- 3) สังเคราะห์ปัจจัยและนำมาจัดกลุ่มตัวแปรเชื่อมโยงกับข้อมูลในการเก็บข้อมูลของโครงการของ PISA 2018 โดยแบ่งตามตัวแปรระดับสถานศึกษาและระดับนักเรียน

ขั้นที่ 2 วิเคราะห์ปัจจัยที่ส่งผลต่อคะแนน PISA จากข้อมูลของโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากลหรือ PISA ปี 2018 ด้วยการวิเคราะห์พหุระดับ

1) จัดเตรียมไฟล์ข้อมูล ซึ่งเป็นฐานข้อมูล (data base) ของโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล ปี 2561 PISA 2018 (OECD) ด้วยโปรแกรม SPSS for Windows V.26 และใช้การวิเคราะห์พหุระดับ โดยใช้โมเดลเชิงเส้นตรงแบบลดหลั่น (HLM) ด้วยการวิเคราะห์ 2 ระดับ ด้วยโปรแกรม HLM for Windows Version 8.2

2) การเตรียมระบบฐานข้อมูลสำหรับการวิเคราะห์ โดยการวิเคราะห์ข้อมูลครั้งนี้ผู้วิจัยใช้การวิเคราะห์พหุระดับ โดยใช้โมเดลเชิงเส้นตรงแบบลดหลั่น (HLM) ผู้วิจัยจัดเตรียมข้อมูลโดยใช้โปรแกรม SPSS for Windows V.26 แบ่ง ตามระดับของการวิเคราะห์ จำนวน 2 ไฟล์ ดังนี้

ไฟล์ที่ 1: ระดับนักเรียน เป็นไฟล์ที่รวมนักเรียนทั้งหมดที่เป็นกลุ่มตัวอย่างจากการทำแบบสอบ เข้าอยู่ในไฟล์เดียวกัน ประกอบด้วยรหัสสถานศึกษา รหัสนักเรียน และตัวแปรระดับนักเรียน โดยรหัสสถานศึกษา จะเรียงตามแนวแถว ตั้งแต่สถานศึกษาแรกไปถึงสถานศึกษาสุดท้าย รหัส นักเรียน จะเรียงตามแนวแถว ตั้งแต่นักเรียนคนแรกไปถึงคนสุดท้ายที่อยู่ในสถานศึกษานั้น และตัวแปรระดับนักเรียน

ไฟล์ที่ 2: ระดับสถานศึกษา ประกอบด้วยรหัสสถานศึกษา และตัวแปรระดับสถานศึกษา โดยรหัสสถานศึกษา จะเรียงตามแนวแถว ตั้งแต่สถานศึกษาแรกไปถึงสถานศึกษาสุดท้าย รหัส นักเรียน จะเรียงตามแนว แถว ตั้งแต่นักเรียนคนแรกไปถึง คนสุดท้ายที่อยู่ในสถานศึกษานั้น

3) นำเข้าข้อมูลที่จัดเรียงทั้งหมด และใช้การวิเคราะห์พหุระดับ โดยใช้โมเดลเชิงเส้นตรงแบบลดหลั่น (HLM) โดยใช้โปรแกรม HLM for Windows Version 8.2

ผลการวิจัย

จากการศึกษาสามารถสรุปผลการวิจัยได้ดังนี้ ผลการวิเคราะห์ปัจจัยด้านความเหลื่อมล้ำทางการศึกษาที่ส่งผลต่อความฉลาดรู้คณิตศาสตร์ (mathematical literacy) จากข้อมูลของโครงการประเมินผลนักเรียนระดับนานาชาติหรือ PISA ปี 2018 พบว่า ตัวแปรระดับสถานศึกษา ตัวแปรระดับสถานศึกษา ที่มีผลต่อคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ ได้แก่ ขนาดห้องเรียนที่มีนักเรียนจำนวน 21-25 คน จำนวนงบประมาณจากรัฐบาล อัตราส่วนจำนวนนักเรียนต่อครู และประเภทของสถานศึกษา ตัวแปรระดับนักเรียนที่มีผลต่อคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ ได้แก่ เศรษฐฐานะครอบครัวยุติธรรมในบ้านของนักเรียน ความมั่งคั่งของครอบครัวของนักเรียน ตัวแปรการศึกษาสูงสุดของผู้ปกครองของนักเรียนระดับหลังมัธยมศึกษา ก่อนอุดมศึกษา (ที่ไม่ใช่ระดับมหาวิทยาลัย) โดยสรุปดังตาราง 1

ตาราง 1 อิทธิพลคงที่ (fixed effect) และค่าอิทธิพลสุ่ม (random effect) ของการวิเคราะห์โมเดลสมมติฐาน (hypothetical model)

Fixed Effect	Coefficient	Standard Error	T-ratio	df	p - value
ค่าคงที่, γ_{00}	405.594**	5.255	77.166	261	<0.001
ตัวแปรระดับสถานศึกษา					
อัตราส่วนจำนวนนักเรียนต่อครู, γ_{01}	-1.791*	0.420	-4.263	261	<0.001
ขนาดห้องเรียน, γ_{02}	-0.300	0.638	-0.471	261	0.638
ดัชนีการขาดแคลนครู, γ_{03}	4.993	3.196	1.563	261	0.119
จำนวนงบประมาณจากรัฐบาล, γ_{04}	-0.747*	0.124	-5.975	261	<0.001
ดัชนีทรัพยากรทางการเรียน, γ_{05}	5.222	10.921	0.478	261	0.633
ขนาดห้องเรียน 16-20 คน, γ_{06}	27.297	15.643	1.745	261	0.082
ขนาดห้องเรียน 21-25 คน, γ_{07}	73.342*	13.410	5.469	261	<0.001
ขนาดห้องเรียน 26-30 คน, γ_{08}	-0.611	14.155	-0.043	261	0.966
ขนาดห้องเรียน 31-35 คน, γ_{09}	0.076	14.969	0.005	261	0.996
ขนาดห้องเรียน 36-40 คน, γ_{010}	15.915	16.430	0.969	261	0.334
ขนาดห้องเรียน 41-45 คน, γ_{011}	21.451	19.624	1.093	261	0.275
ขนาดห้องเรียน 46-50 คน, γ_{012}	45.264	25.270	1.791	261	0.074
ขาดสื่อการเรียนการสอน(น้อย), γ_{013}	-22.464	12.479	-1.800	261	0.073
ขาดสื่อการเรียนการสอน (บางส่วน), γ_{014}	-25.142	15.856	-1.586	261	0.114
ขาดสื่อการเรียนการสอน(มาก), γ_{015}	-44.012	22.889	-1.923	261	0.056
สื่อการเรียนการสอนคุณภาพต่ำ (น้อย), γ_{016}	-18.464	12.450	-1.483	261	0.139
สื่อการเรียนการสอนคุณภาพต่ำ (บางส่วน), γ_{017}	-36.031	17.666	-2.039	261	0.042
สื่อการเรียนการสอนคุณภาพต่ำ (มาก), γ_{018}	-25.816	25.704	-1.004	261	0.316
ประเภทสถานศึกษาเอกชน, γ_{019}	-40.087*	9.650	-4.154	261	<0.001
ชุมชนที่ตั้งของโรงเรียนเมืองเล็ก, γ_{020}	13.496	9.186	1.469	261	0.143
ชุมชนที่ตั้งของโรงเรียนเมือง, γ_{021}	11.040	9.317	1.185	261	0.237

ตาราง 1 (ต่อ)

Fixed Effect	Coefficient	Standard Error	T-ratio	df	p - value
ชุมชนที่ตั้งของโรงเรียนเมืองใหญ่, γ_{022}	15.696	10.104	1.553	261	0.122
ชุมชนที่ตั้งของโรงเรียนเมืองใหญ่มาก, γ_{023}	37.097	14.520	2.555	261	0.011
ตัวแปรระดับนักเรียน					
ความมั่งคั่งของครอบครัว, γ_{10}	-13.866**	2.129	-6.512	8148	<0.001
ทรัพย์สินในบ้าน, γ_{20}	16.318**	2.541	6.423	8148	<0.001
ทรัพยากรการศึกษาที่บ้าน, γ_{30}	1.090	0.949	1.149	8148	0.251
การศึกษาสูงสุดของผู้ปกครองระดับประถมศึกษา, γ_{40}	-13.360	4.614	-2.896	8148	0.004
การศึกษาสูงสุดของผู้ปกครองระดับมัธยมศึกษาตอนต้น, γ_{50}	-3.220	4.620	-0.697	8148	0.486
การศึกษาสูงสุดของผู้ปกครองระดับมัธยมศึกษาตอนปลาย, γ_{60}	-6.875	5.443	-1.263	8148	0.206
การศึกษาสูงสุดของผู้ปกครองระดับหลังมัธยมศึกษาตอนปลาย (ไม่ใช่ระดับมหาวิทยาลัย), γ_{70}	-15.934**	4.516	-3.527	8148	<0.001
การศึกษาสูงสุดของผู้ปกครองระดับปริญญาตรี โท เอก, γ_{80}	-8.178	4.611	-1.773	8148	0.076
Random Effect	Standard Deviation	Variance Component	df	χ^2	p-value
INTRCPT1, u_0	43.683	1908.315	261	5610.45	<0.001
level-1, r_{ij}	52.544	2760.995		419	
Random Level-1 coefficient			Reliability estimate		
INTRCPT1, β_0			0.909		

*p < .05 , **p < .01, Deviance = 91461.228, Number of estimated parameters = 2

จากตาราง 1 อิทธิพลคงที่ (Fixed Effect) ค่าเฉลี่ยรวมของคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ มีค่าเท่ากับ 405.594 ค่าความคลาดเคลื่อนมาตรฐานเท่ากับ 5.255 เมื่อทดสอบค่าสถิติที่ พบว่า $T = 77.166$ และ $p < 0.001$ ตัวแปรที่มีอิทธิพลต่อคะแนนแบบประเมินการรู้เรื่องคณิตศาสตร์เรียงลำดับในแต่ละระดับเป็นดังนี้

ระดับสถานศึกษา ตัวแปรระดับสถานศึกษา ที่มีผลต่อคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ ขนาดห้องเรียนที่มีนักเรียนจำนวน 21-25 คน จำนวนงบประมาณจากรัฐบาล อัตราส่วนจำนวนนักเรียนต่อครู ประเภทของสถานศึกษา เมื่อพิจารณาจากขนาดอิทธิพลหรือค่าสัมประสิทธิ์ถดถอย พบว่า ตัวแปรที่มีค่าสัมประสิทธิ์ถดถอยที่เป็นบวกสูงสุด คือ ขนาดห้องเรียนที่มีนักเรียนจำนวน 21-25 คน ($\beta = 73.342$) ส่วนตัวแปรที่มีค่าสัมประสิทธิ์ถดถอยที่เป็นลบสูงสุด คือ ประเภทของสถานศึกษา ($\beta = -40.087$) อัตราส่วนจำนวนนักเรียนต่อครู ($\beta = -1.791$)

ระดับนักเรียน ตัวแปรระดับนักเรียนที่มีผลต่อคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ได้แก่ เศรษฐฐานะครอบครัวทรัพย์สินในบ้านของนักเรียน ความมั่งคั่งของครอบครัวของนักเรียน ตัวแปรการศึกษาสูงสุดของผู้ปกครองของนักเรียนระดับหลังมัธยมศึกษา ก่อนอุดมศึกษา (ที่ไม่ใช่ระดับมหาวิทยาลัย) เมื่อพิจารณาขนาดอิทธิพลหรือค่าสัมประสิทธิ์ถดถอย พบว่า ตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวกสูงสุด คือ เศรษฐฐานะครอบครัวทรัพย์สินในบ้านของนักเรียน ($\beta = 16.318$) ส่วนตัวแปรระดับนักเรียนที่มีอิทธิพลทางลบสูงสุด คือ การศึกษาสูงสุดของผู้ปกครองของนักเรียนระดับหลังมัธยมศึกษา ก่อนอุดมศึกษา (ที่ไม่ใช่ระดับมหาวิทยาลัย) ($\beta = -15.934$) ความมั่งคั่งของครอบครัวของนักเรียน ($\beta = -13.866$)

การประมาณค่าความเชื่อมั่นของสัมประสิทธิ์ตัวแปรสุ่มในระดับนักเรียน (Level -1 model) ในโมเดลตามสมมติฐาน มีค่าเท่ากับ 0.909 แสดงว่าค่าเฉลี่ยคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ของกลุ่มตัวอย่างมีความเชื่อมั่นสูง แสดงถึงความเป็นค่าเฉลี่ยคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ได้

อภิปรายผล

งานวิจัยการวิเคราะห์ปัจจัยด้านความเหลื่อมล้ำทางการศึกษาที่ส่งผลต่อความฉลาดรู้คณิตศาสตร์ ภายใต้โปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล (PISA) อภิปรายในประเด็นสำคัญ ดังนี้

จากการวิเคราะห์ปัจจัยที่ส่งผลต่อความฉลาดรู้คณิตศาสตร์ ภายใต้โปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล (PISA) วิเคราะห์หาค่าความสัมพันธ์ระหว่างตัวแปร ศึกษาอิทธิพลหลักและอิทธิพลปฏิสัมพันธ์ของตัวแปรต้นที่จะศึกษาในระดับที่ต่างกัน Tuksino (2009) โดยใช้โมเดลเชิงเส้นตรงแบบลดหลั่น (HLM) วิเคราะห์ตัวแปรการสังเคราะห์ตัวแปรจากข้อมูลของโครงการประเมินผล PISA 2018 ตัวแปรที่ส่งผลต่อคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ (mathematical literacy)

ตัวแปรระดับสถานศึกษา จากผลการวิเคราะห์ พบว่าตัวแปรที่มีอิทธิพลต่อคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ 4 ตัวแปร ตัวแปรที่มีอิทธิพลทางบวกจากค่าสัมประสิทธิ์สูงสุด คือ ตัวแปรขนาดห้องเรียนที่มีนักเรียนจำนวน 21-25 คน และตัวแปรที่มีอิทธิพลทางลบเรียงลำดับจากค่าสัมประสิทธิ์สูงไปต่ำ ได้แก่ ประเภทของสถานศึกษาเอกชน อัตราส่วนจำนวนนักเรียนต่อครู และจำนวนงบประมาณจากรัฐบาล ส่วนตัวแปรระดับนักเรียน ตัวแปรที่มีอิทธิพลต่อคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ 3 ตัวแปร ตัวแปรที่มีอิทธิพลทางบวกจากค่าสัมประสิทธิ์สูงสุด คือ เศรษฐฐานะครอบครัวทรัพย์สินในบ้านของนักเรียน และตัวแปรที่มีอิทธิพลทางลบเรียงลำดับจากค่าสัมประสิทธิ์สูงไปต่ำ ได้แก่ การศึกษาสูงสุดของผู้ปกครองของ

นักเรียนระดับหลังมัธยมศึกษาตอนอุดมศึกษา (ที่ไม่ใช่ระดับมหาวิทยาลัย) และความมั่งคั่งของครอบครัวของนักเรียน

จากตัวแปรที่มีอิทธิพลต่อคะแนนแบบประเมินความฉลาดรู้คณิตศาสตร์ มีตัวแปรที่น่าสนใจหลายลักษณะดังนี้

ประการที่หนึ่ง ลักษณะของตัวแปรระดับสถานศึกษาที่มีอิทธิพลทางบวก คือ ตัวแปรขนาดห้องเรียนที่มีนักเรียนจำนวน 21-25 คน นั่นคือสถานศึกษาที่มีคุณลักษณะการจัดห้องเรียนที่มีขนาด 21-25 คน จะทำให้ผลการประเมินสูงกว่าสถานศึกษาที่จัดขนาดห้องเรียนในลักษณะอื่น สอดคล้องกับงานวิจัยของ Malakhonburi (2017) พบว่า การจัดการเรียนการสอนผู้สอนควรเลือกใช้ทฤษฎีให้เหมาะสมกับขนาดชั้นเรียนนั้น ๆ เพราะจะส่งผลต่อคุณภาพการจัดการเรียนการสอน สอดคล้องกับ Mueller (2011) กล่าวว่า ขนาดชั้นเรียนมีผลต่อประสิทธิภาพ อีกทั้งควรจัดสรรครูตามประสบการณ์จัดครูให้เหมาะสมกับขนาดของชั้นเรียน

ประการที่สอง ลักษณะของตัวแปรระดับสถานศึกษาที่มีอิทธิพลทางลบ ได้แก่ ประเภทของสถานศึกษาเอกชน อัตราส่วนจำนวนนักเรียนต่อครู และจำนวนงบประมาณจากรัฐบาล นั่นคือสถานศึกษาที่มีคุณลักษณะประเภทของสถานศึกษาเอกชน จะทำให้ผลคะแนนประเมินต่ำกว่าสถานศึกษาลักษณะอื่น ทั้งนี้สอดคล้องกับผลสรุปการประเมิน PISA 2018 ที่พบว่านักเรียนจากกลุ่มโรงเรียนที่เน้นทางด้านวิทยาศาสตร์และคณิตศาสตร์มีคะแนนสูงที่สุด ซึ่งอยู่ในระดับเดียวกับคะแนนเฉลี่ยของกลุ่มประเทศ/เศรษฐกิจที่มีคะแนนสูงสุดห้าอันดับแรก รองลงมาคือ กลุ่มโรงเรียนสามัญของมหาวิทยาลัย ส่วนกลุ่มโรงเรียนอื่น ๆ รวมถึงกลุ่มสถานศึกษาเอกชน ยังคงมีคะแนนต่ำกว่าค่าเฉลี่ยของประเทศสมาชิก OECD (OECD, 2019) ส่วนอัตราส่วนของจำนวนนักเรียนต่อครูสูง จะทำให้ผลคะแนนการประเมินต่ำกว่าสถานศึกษาลักษณะอื่น สอดคล้องกับ Kittikajorn (2012) กล่าวว่า การจัดสรรกำลังคนทางการศึกษาไม่ทำให้เกิดความเสมอภาค โดยเฉพาะสถานศึกษาขนาดเล็กในเขตยากจนโดยเฉพาะในชนบทจะมีครูอาจารย์ตามวุฒิน้อย ทำให้มีผลสัมฤทธิ์ทางการเรียนต่ำ อีกทั้งสอดคล้องกับ Ministry of Education (2017) ได้รายงานการศึกษาไทยในเรื่องแนวนโยบายการพัฒนาวิชาชีพและบุคลากรทางการศึกษา มุ่งเน้นที่การปรับปรุงประเด็นการขจัดปัญหาการขาดแคลนอัตราจำนวนครูเนื่องจากเป็นสิ่งที่ส่งผลต่อคุณภาพนักเรียน และนอกจากนี้ตัวแปรที่น่าสนใจอีกตัวแปรหนึ่งคือ จำนวนงบประมาณจากรัฐบาล ซึ่งพบว่ามีอิทธิพลทางลบนั่นแสดงว่าการเพิ่มงบประมาณทางการศึกษา ไม่ได้เป็นหลักประกันว่าคุณภาพการศึกษาจะสูงขึ้น ซึ่งสอดคล้องกับ Ferreira & Gignoux (2011) พบว่า โรงเรียนเกิดความสัมพันธ์เชิงลบกับส่วนแบ่งการใช้จ่ายด้านการศึกษาที่ส่วนกลางจัดสรรให้แก่โรงเรียน ดังนั้นการจัดสรรงบประมาณอย่างมีประสิทธิภาพย่อมส่งผลต่อคุณภาพของนักเรียนทั้งทางด้านผลสัมฤทธิ์ทางการเรียน และด้านคุณลักษณะอันพึงประสงค์ (Daduang, 2018)

ประการที่สาม ลักษณะของตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวก คือ เศรษฐฐานะครอบครัวทรัพย์สินในบ้านของนักเรียน ซึ่งมีนักเรียนหลายท่านได้ศึกษาตัวแปรที่ส่งผลต่อการเรียนรู้ เช่น Tuksino (2009) พบว่า เศรษฐฐานะทางครอบครัว เป็นตัวแปรระดับนักเรียนที่มีอิทธิพลทางบวกต่อคะแนนประเมินวิชาวิทยาศาสตร์ Farooq et al. (2011) สรุปผลว่า เศรษฐกิจและฐานะทางสังคมส่งผลต่อความสำเร็จในการเรียนของนักเรียนอย่างมีนัยสำคัญทางสถิติ

ประการที่สี่ ลักษณะของตัวแปรระดับนักเรียนที่มีอิทธิพลทางลบ ได้แก่ การศึกษาสูงสุดของผู้ปกครองของนักเรียนระดับหลังมัธยมศึกษาตอนต้น (ที่ไม่ใช่ระดับมหาวิทยาลัย) Farooq et al. (2011) ได้ศึกษาปัจจัยที่ส่งผลต่อความสำเร็จในการเรียนของนักเรียน พบว่า การศึกษาของผู้ปกครองมีผลกระทบอย่างมีนัยสำคัญทางสถิติ และความมั่งคั่งของครอบครัวของนักเรียน อาจจะเป็นข้อมูลส่วนหนึ่งที่ชี้ให้เห็นว่านักเรียนที่มาจากครอบครัวที่มีฐานะการเป็นอยู่ที่สะดวกสบายไม่จำเป็นว่าจะได้เปรียบกว่านักเรียนที่มีฐานะการเป็นอยู่ที่ยากแค้นกว่า (Tuksino, 2009)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 ระดับสถานศึกษา ผลการวิจัยที่บ่งชี้ถึงตัวแปรระดับสถานศึกษาที่ทำให้ผลคะแนนสูงขึ้น ได้แก่ ขนาดห้องเรียนที่มีนักเรียนจำนวน 21-25 คน เป็นสิ่งที่ควรได้รับการพัฒนาหรือสนับสนุนให้เกิดขึ้น นอกจากนี้ จำนวนงบประมาณจากรัฐบาล อัตราส่วนจำนวนนักเรียนต่อครู ประเภทของสถานศึกษา ล้วนแต่เป็นตัวแปรที่มีอิทธิพลทางลบต่อผลคะแนน จึงควรพัฒนาหรือใช้ปัจจัยในด้านอื่น ๆ เพิ่มขึ้นเพื่อชดเชยปัจจัยที่สูญเสียไป เช่น การจัดแหล่งทรัพยากรทางการเรียนอื่น ๆ มาทดแทนโอกาสในการศึกษาหาความรู้โดยการใช้ อินเทอร์เน็ต หรือการใช้กลยุทธ์เพื่อให้ผู้ปกครองเข้ามามีบทบาทมากขึ้นต่อการปฏิบัติงานของโรงเรียน เพื่อให้ผู้ปกครองสร้างแรงกดดันต่อการปฏิบัติงานของสถานศึกษามากขึ้น ดังนั้นสถานศึกษาควรมีการวางแผนพัฒนาตัวแปรดังกล่าวให้มีคุณภาพสูงยิ่งขึ้น

1.2 ระดับนักเรียน จากผลการวิจัยที่พบว่า การมีเศรษฐกิจในครอบครัว ทรัพย์สินในบ้านพร้อม จะส่งผลให้ผลคะแนนของนักเรียนสูงขึ้น ดังนั้น ผู้ปกครองหรือครอบครัว สามารถให้การสนับสนุน ส่งเสริมสถานศึกษา สร้างสภาวะกระตุ้นให้สถานศึกษาเอาใจใส่ต่อผู้รับบริการ หรือสร้างแรงกระตุ้นให้บุตรหลานโดยไม่จำเป็นต้องจ่ายเป็นตัวแทนเงินเสมอไป สนับสนุนเอาใจใส่ในส่วนที่บุตรหลานขาด ให้คำแนะนำในการเรียนให้มากที่สุด เพื่อเป็นการยกระดับผลสัมฤทธิ์ทางการเรียนของนักเรียนได้ดียิ่งขึ้น นอกจากนี้ผลการวิจัยยังพบว่า ความมั่งคั่งของครอบครัวไม่ได้ส่งเสริมให้นักเรียนมีคะแนนที่เพิ่มขึ้น เนื่องจากความมั่งคั่งของครอบครัวเป็นสิ่งที่ไม่ได้มีผลต่อการสนับสนุนทางการศึกษา แต่นักเรียนที่มีความมั่งคั่งน้อยทำให้มีแรงกระตุ้นในการศึกษาเป็นสิ่งที่ส่งเสริมให้คะแนนสูงขึ้น ดังนั้นผู้ปกครองจึงควรตระหนักถึงการใช้เวลาให้กับบุตรหลาน และปรับเปลี่ยนค่านิยมเกี่ยวกับการให้สิ่งที่ไม่จำเป็นต่อการศึกษา แต่ควรส่งเสริมหรือกระตุ้นให้นักเรียนมีแรงบันดาลใจ และเห็นประโยชน์จากการศึกษามากขึ้น

อย่างไรก็ตาม ปัจจัยหรือตัวแปรที่มีอิทธิพลทางบวกต่อการจัดการศึกษา ไม่ได้แสดงอิทธิพลได้เพียงตัวแปรเดียว แต่เป็นตัวแปรที่มีความสัมพันธ์เกี่ยวข้องกัน ดังนั้น การจัดการศึกษาจึงควรวางแผนและทำให้เกิดความร่วมมือกัน ทั้งสถานศึกษาและครอบครัวควรมีการวางแผนพัฒนาตัวแปรที่เกี่ยวข้องให้มีคุณภาพสูงยิ่งขึ้น ตระหนักถึงปัจจัยที่จะส่งเสริมให้นักเรียนใช้เวลาในการศึกษา อย่างเป็นระบบ และควรใช้สารสนเทศที่ได้จากการวิเคราะห์ข้อมูล เพื่อเพิ่มคุณภาพการจัดการศึกษาให้มากยิ่งขึ้น

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรมีการใช้ข้อมูลของ PISA ในปีต่าง ๆ อีกทั้งประเทศไทย ยังขาดการนำข้อมูลในระดับนานาชาติมาวิเคราะห์รอบสอง (second analysis) เพื่อให้ได้สารสนเทศเพิ่มเติมต่อการนำไปพัฒนาการจัดการศึกษา โดยทำการวิจัยเพื่อเปรียบเทียบผลในระยะยาว หรือเปรียบเทียบผลภายในระยะเวลาเดียวกัน ทั้งนี้ อาจจะศึกษาในรูปแบบของวิชาเดียวกัน ปีที่แตกต่างกัน หรือปีเดียวกันโดยใช้ข้อมูลในหลายวิชา เช่น วิเคราะห์ความฉลาดรู้ด้านการอ่าน ความฉลาดรู้ด้านวิทยาศาสตร์ เพื่อให้ได้ข้อสารสนเทศต่อการนำไปพัฒนาเพิ่มคุณภาพการจัดการศึกษาได้มากยิ่งขึ้น

2.2 ควรมีการวิจัยเชิงคุณภาพ เพื่อศึกษาถึงสาเหตุจากปัจจัยที่ส่งผลต่อคุณภาพทางการศึกษาในระดับสถานศึกษา ระดับนักเรียน อีกทั้งควรนำผลในการวิเคราะห์ทางสถิติไปดำเนินการต่อ เพื่อให้เห็นสาเหตุและแนวทางในการนำไปใช้เพื่อประโยชน์ในการออกแบบการจัดการศึกษาต่อไป

References

- Farooq, M., Chaudhry, A., Shafiq, M., & Berhanu, G. (2011). Factors affecting students' quality of academic performance: a case of secondary school level. *Journal of Quality and Technology Management*, 7(2), 1-14.
- Ferreira, F. H. G., & Gignoux, J. (2011). The Measurement of Educational Inequality: Achievement and Opportunity. *IZA Discussion Papers*, 6161. <https://docs.iza.org/dp6161.pdf>
- Mueller, S. (2011). Teacher experience and the class size effect - experimental evidence. *IWQW Discussion Papers*, 2011(7).
- OECD. (2012). *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*. OECD Publishing.
- OECD. (2019). *PISA 2018 Results (Volume III): What School Life Means for Students' Lives*, PISA. OECD Publishing.

Translate Thai References

- Chamamporn, C., Sriharun, B., & Boutngern, C. (2014). Multi-level factors affecting the effectiveness of primary schools under office of the basic education commission. *Valayalongkorn Research and Development Journal*, 9(1), 166-175. (In Thai)
- Chinjunthuk, S., & Junpeng, P. (2020). Assessment Guidelines for Student's Personalized Mathematical Proficiency Development. *Journal of Educational Measurement Mahasarakham University*, 26(1), 47- 64. (in Thai)

- Daduang A. (2018). Approaches for budget allocation improvement affecting student quality of schools under the secondary educational service area office 1. *An Online Journal of Education OJED*, 13(4), 80-93. (In Thai)
- Institute for the promotion of technology science and technology: (IPST). (2021). *PISA Assessment Results 2018 Reading, Mathematics and Science*. Institute for the promotion of technology science and technology. (In Thai)
- Kittikajorn, P. (2012). "Overhauling Thai Education to Reduce Inequality" Paper presented at the meeting: *Educational inequality*. Graduate Business University. (In Thai)
- Malakhonburi, C. (2017). The Improvement of Classroom Efficiency: A Comparison between Large, Middle and Small Classroom in College of Politics and Governance, Mahasarakham University. *Journal of Politics and Governance*, 7(3), 198-210. (In Thai)
- Ministry of Education. (2017). *Indicators and subjects of learning the core Learning group Mathematics according to the basic education core curriculum (Revised version 2017)*. Agricultural Co-operative Federation of Thailand Limited. (In Thai)
- Pharcharuen, W., Uthaphan, P., Koklang, K., & Phutchu, S. (2021). Educational disparities of thai citizens studying age with the online learning situation in the viral age covid-19. *Pannapanithan Journal*, 6(1), 1-14. (In Thai)
- Ponjang, N., Tulmethakaan, M., & Ketchatturat, J. (2017). Quality education study of Mathematics at the elementary level in school under the office of Loei educational service area i: an application of value-added model. *Graduate Journal Valaya Alongkorn Rajabhat University under the Royal Patronage*, 11(Special Issue), 73-84. (In Thai)
- Sukjaroen, N., & Yoonisil, W. (2014). Meta-Analysis and Meta-Synthesis. *Journal of Humanities and Social Sciences Rajabhat Mahasarakham University*, 8(3). (In Thai)
- Tuksino, P. (2009). *Assessment of Science Instructional Management in Basic Education Schools: An Application of a Value-added Model and Differential Item Functioning* [Doctoral dissertation]. Chulalongkorn University. (In Thai)

รูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา

สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3

A Model for Development of Digital Learning Culture in Primary Schools

under Chaiyaphum Primary Education Service Area Office 3

รติยากร คชา^{1*} และ ดาวรรุณ ถวิลการ²

Ratiyakorn Khacha^{1*} and Dawruwan Thawinkarn²

(Received: April 28, 2022; Revised: June 1, 2022; Accepted: June 7, 2022)

บทคัดย่อ

การวิจัยครั้งนี้เป็นการวิจัยแบบผสมผสานวิธี (mixed methods) โดยมีความมุ่งหมายดังนี้ 1) เพื่อศึกษาองค์ประกอบและตัวบ่งชี้ของวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 2) เพื่อตรวจสอบความสอดคล้องกลมกลืนของโมเดลตัวบ่งชี้ของวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 ที่ผู้วิจัยพัฒนาขึ้นกับข้อมูลเชิงประจักษ์ 3) เพื่อสร้างและประเมินรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3 กลุ่มตัวอย่าง ได้แก่ ผู้บริหารและครู จำนวน 270 คน ได้มาโดยวิธีการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือแบบสอบถามแบบมาตรประมาณค่า 5 ระดับ มีค่าความตรงเชิงเนื้อหาอยู่ระหว่าง 0.80 – 1.00 มีค่าความเที่ยงทั้งฉบับเท่ากับ 0.967 วิเคราะห์ค่าสถิติเชิงพรรณนาและวิเคราะห์สถิติอ้างอิงด้วยโปรแกรม SPSS สำหรับการวิเคราะห์สถิติพื้นฐาน และวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรม Mplus ร่างรูปแบบโดยการสัมภาษณ์ และประเมินความเหมาะสม ความเป็นไปได้และความเป็นประโยชน์ของรูปแบบ

โดยมีผลการวิจัย ดังนี้

1. วัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา มีจำนวน 5 องค์ประกอบหลัก ได้แก่
 - 1) การฝึกอบรมบุคลากรแบบผสมผสานทางดิจิทัล มี 2 ตัวบ่งชี้ 5 พฤติกรรมบ่งชี้ 2) ค่านิยมการทำงานร่วมกันผ่านสังคมออนไลน์ มี 2 ตัวบ่งชี้ 5 พฤติกรรมบ่งชี้ 3) สมรรถนะเชิงของบุคลากร มี 3 ตัวบ่งชี้ 8 พฤติกรรมบ่งชี้ 4) ภาวะผู้นำดิจิทัล มี 3 ตัวบ่งชี้ 8 พฤติกรรมบ่งชี้ 5) สภาพแวดล้อมการเรียนรู้เชิงดิจิทัล มี 3 ตัวบ่งชี้ 7 พฤติกรรมบ่งชี้ รวมองค์ประกอบของวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3 มีจำนวน 5 องค์ประกอบ 13 ตัวบ่งชี้ 33 พฤติกรรมบ่งชี้

¹ นักศึกษาปริญญาโท สาขาวิชาการบริหารการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² ผู้ช่วยศาสตราจารย์ สาขาวิชาการบริหารการศึกษา ศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

¹ Master's degree student, Educational Administration Program, Faculty of Education, Khon Kaen University

² Assistant Professor Dr., Educational Administration Program, Faculty of Education, Khon Kaen University

* E-mail: khacha25011995@gmail.com

2. ผลการทดสอบความสอดคล้องกลมกลืนของโมเดลตัวบ่งชี้วัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา กับข้อมูลเชิงประจักษ์ พบว่า โมเดลที่ผู้วิจัยพัฒนาขึ้นมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ ซึ่งมีค่า $\chi^2 = 56.948$, $df = 43$, $\chi^2 / df = 1.324$, P-Value = 0.0754, RMSEA = 0.035, SRMR = 0.023, CFI = 0.996, TLI = 0.992

3. ผลการสร้างรูปแบบการพัฒนาวัฒนธรรม การเรียนรู้ เชิงดิจิทัลในโรงเรียนประถมศึกษา ประกอบด้วย ชื่อรูปแบบ หลักการแนวคิด วัตถุประสงค์ ค่าเป้าหมาย สู่ความสำเร็จ และกลไกขับเคลื่อนสู่ความสำเร็จ พบว่า มีผลการประเมินความเหมาะสม ความเป็นไปได้ และความเป็นประโยชน์อยู่ในระดับมากที่สุด

คำสำคัญ: รูปแบบ วัฒนธรรมการเรียนรู้ วัฒนธรรมการเรียนรู้เชิงดิจิทัล

Abstract

The research employed mixed methods, with the following objectives: 1) to study the components and indicators of digital learning culture in primary schools under Chaiyaphum Primary Education Service Area Office 3; 2) to examine the consistency of the indicator model of the digital learning culture in primary schools under Chaiyaphum Primary Education Service Area Office 3, that the author had constructed, with empirical data; 3) to create and evaluate a model for development of digital learning culture in primary schools under Chaiyaphum Primary Education Service Area Office 3. The sample of 270 administrators and teachers was obtained by multistage sampling. The tool used for data collection was a 5-level rating scale questionnaire with the content validity between 0.80 to 1.00 and the total reliability of 0.967. The analysis of data employed descriptive and inferential statistics and SPSS program. Confirmatory factor analysis (CFA) was performed with the Mplus program. Drafting of the model was carried out by interviewing, followed by assessing the suitability, possibility and usefulness of the model. The research results were as follows:

1. The digital learning culture in primary schools consisted of 5 elements: 1) digital blended personnel training, having 2 indicators with 5 behavioral indicators; 2) on-line social collaboration values, having 2 indicators with 5 behavioral indicators; (3) digital competency, having 3 indicators with 8 behavioral indicators; (4) digital leadership, having 3 indicators with 8 behavioral indicators; (5) digital learning environment, having 3 indicators with 7 behavioral indicators. Totally, the digital learning culture in primary schools under Chaiyaphum Primary Education Service Area Office 3 had 5 components, 13 indicators and 33 behavioral indicators.

2. The test results of the model of the digital learning culture in primary schools, developed by the author, were consistent with the empirical data as having the values of

$\chi^2 = 56.948$, $df = 43$, $\chi^2 / df = 1.324$, P-Value = 0.0754, RMSEA = 0.035, SRMR = 0.023, CFI = 0.996, TLI = 0.992

3.The evaluation results of the model for development of digital learning culture in primary schools revealed that the model consists of the name, principle and concept, objective, achievement goal, and mechanism for success. The result of the model evaluation indicated that the model was suitable, possible and useful at the highest level.

Keywords: model, learning culture, digital learning culture

บทนำ

ด้วยกระแสการเปลี่ยนแปลงทางเทคโนโลยีดิจิทัลที่เกิดขึ้นในสังคมปัจจุบันนี้จะเห็นได้ว่าการพัฒนาศักยภาพของประชาชนหรือทุนมนุษย์เป็นสิ่งสำคัญ เทคโนโลยีดิจิทัลจะเป็นเครื่องมือในการสร้างศักยภาพยกระดับประชาชนไปสู่สังคมฐานความรู้ ทำให้มีความสามารถขยับไปสู่การผลิตที่ใช้เทคโนโลยีหรือรู้จักใช้เทคโนโลยีและข้อมูลข่าวสาร ในการประกอบอาชีพมากขึ้น (Ministry of Information and Communication Technology, 2016) จากแผนการศึกษาแห่งชาติ (พ.ศ. 2560-2579) ประเทศไทยได้ให้ความสำคัญด้านการศึกษาในฐานะกลไกหลักในการพัฒนาประเทศมาโดยตลอด ความท้าทายที่เป็นพลวัตของโลกศตวรรษที่ 21 ทั้งในส่วนที่เป็นแรงกดดันภายนอก ได้แก่ การเปลี่ยนแปลงของบริบทเศรษฐกิจและสังคมโลก อันเนื่องจากการปฏิวัติดิจิทัล (digital revolution) รวมทั้ง ผลกระทบของการเป็นประชาคมอาเซียน ความต้องการกำลังคนที่มีทักษะในศตวรรษที่ 21 ส่งผลกระทบต่อระบบการศึกษา ที่ต้องปรับเปลี่ยนให้ตอบสนองและรองรับความท้าทายดังกล่าว จึงมีความจำเป็นที่ประเทศไทยต้องปฏิรูปการศึกษา เพื่อให้ระบบการศึกษาเป็นกลไกหลักของการขับเคลื่อนประเทศ

ในการปฏิรูปการศึกษา โรงเรียนเป็นหน่วยงานหลักที่จะทำให้การปฏิรูปการศึกษาเกิดความสำเร็จ เนื่องจากโรงเรียนเป็นหน่วยงานสำคัญในการพัฒนาศักยภาพมนุษย์ เมื่อโรงเรียนเป็นหน่วยงานหลัก ดังนั้น ครูจึงเป็นผู้มีบทบาทสำคัญต่อการเรียนรู้และพัฒนาการทุก ๆ ด้านของผู้เรียน ครูต้องมีการพัฒนาตนเองและร่วมกันปรับเปลี่ยนวัฒนธรรมองค์การให้เอื้อต่อการพัฒนาบุคลากร เพราะวัฒนธรรมองค์การคือสิ่งที่เป็แรงผลักดันให้ไปองค์การสู่ความสำเร็จ โดยการพัฒนาตนเองครูจะต้องตระหนักว่าไม่ใช่ผู้รู้ทุกเรื่องแต่ต้องเรียนรู้ไปพร้อมกับเด็กและบุคลากรที่ทำงานร่วมกัน ซึ่งเกิดเป็นวัฒนธรรมการเรียนรู้ร่วมกันโดยวัฒนธรรมการเรียนรู้จะมีผลต่อการเปลี่ยนแปลงองค์การเป็นเครื่องมือทางการบริหารและการพัฒนาองค์การชนิดหนึ่งที่เป็นตัวกำหนดทิศทางการดำเนินการขององค์การมากกว่าการใช้กฎระเบียบเพื่อสนับสนุนให้องค์การสามารถปรับตัวตอบสนองต่อการเปลี่ยนแปลงในทุกมิติในสังคม (Office of the Permanent Secretary for Ministry of Agriculture and Cooperatives, 2012)

ในยุคของการได้รับอิทธิพลจากเทคโนโลยีดิจิทัล วัฒนธรรมการเรียนรู้ภายในโรงเรียนต้องปรับเปลี่ยนให้ทันกับยุคสมัยเพื่อให้สถานศึกษาพัฒนาไปสู่ความสำเร็จตามเป้าหมายที่ตั้งไว้ จึงต้องมีการสร้าง

วัฒนธรรมการเรียนรู้เชิงดิจิทัลร่วมกันภายในโรงเรียน ซึ่งหมายถึง การสร้างแบบแผนพฤติกรรม ทักษะ ค่านิยมของบุคลากรในสถานศึกษาให้เกิดการยอมรับและสนับสนุนต่อการนำเทคโนโลยีดิจิทัลเข้ามาใช้ในสถานศึกษา โดยมีการประยุกต์ใช้เทคโนโลยีดิจิทัลเข้ามามีส่วนร่วมในการทำงานการจัดการเรียนการสอน การนำมาใช้อำนวยความสะดวกในการพัฒนาตนเองของบุคลากรซึ่งทำให้บุคลากรสามารถเรียนรู้ได้ทุกที่ ทุกเวลา อีกทั้งยังสามารถแลกเปลี่ยนเรียนรู้ร่วมกันแบบออนไลน์ของบุคลากรในสถานศึกษา ผู้ปกครอง และชุมชน โดยในสถานศึกษาจะต้องมีการจัดสภาพแวดล้อมการเรียนรู้เชิงดิจิทัลที่มีอุปกรณ์เทคโนโลยีที่ทันสมัยเพียงพอแก่บุคลากรและผู้เรียน มีการจัดเก็บข้อมูลสารสนเทศของสถานศึกษาที่เป็นระบบสามารถเรียกใช้ข้อมูลได้อย่างรวดเร็วและเป็นปัจจุบัน โดยผู้บริหารซึ่งเป็นส่วนสำคัญหลักในสถานศึกษาซึ่งต้องฝึกทักษะให้เหมาะสมกับยุคดิจิทัลมีภาวะผู้นำดิจิทัลที่แสดงออกถึงความเข้าใจในเทคโนโลยี การมีวิสัยทัศน์ ความรู้ และความสามารถในการนำกระบวนการทางดิจิทัลมาใช้อย่างเหมาะสม (Ward, 2020)

จากที่กล่าวมาจะเห็นได้ว่าการนำวัฒนธรรมการเรียนรู้เชิงดิจิทัลเข้ามาใช้ในโรงเรียนเป็นสิ่งสำคัญ ดังนั้นการดำเนินงานของสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 ซึ่งเป็นหน่วยงานที่กำกับ ดูแล ส่งเสริม สนับสนุน และประสานงานการจัดการศึกษาของสถานศึกษาให้เป็นไปตามนโยบายและมาตรฐาน การศึกษาจึงต้องมีการให้ความสำคัญในการพัฒนาทรัพยากรบุคคลด้านเทคโนโลยีดิจิทัลมากขึ้น ติดตามการนำ เทคโนโลยีดิจิทัลเข้าไปใช้ในสถานศึกษาให้มีความเป็นระบบ โดยเฉพาะในโรงเรียนประถมศึกษาที่ยังมีปัญหา ในด้านความพร้อมของห้องปฏิบัติการเกี่ยวกับเทคโนโลยีสารสนเทศและวัสดุครุภัณฑ์สำนักงาน ปัญหาในเรื่อง ความพร้อมของระบบเครือข่ายอินเทอร์เน็ต ตลอดจนสื่อซอฟต์แวร์ที่จะใช้งาน อีกทั้งยังพบว่าครูและผู้บริหาร บางส่วนยังปฏิเสธการใช้เทคโนโลยีดิจิทัลในทุกรูปแบบ จึงไม่สามารถปรับเปลี่ยนการเรียนการสอนรวมถึง การทำงานร่วมกันในสถานศึกษาให้ทันสมัยเหมาะสมสำหรับยุคดิจิทัลได้เท่าที่ควร จากปัญหาที่พบของสภาพ บริบทโรงเรียนประถมศึกษา สำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 จึงได้มีการจัดทำ แผนพัฒนาการศึกษา ระยะ 5 ปี (พ.ศ.2564 – 2568) ที่มีทิศทางการพัฒนาแผนตามยุทธศาสตร์ชาติ แผนการ ศึกษาแห่งชาติ และสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน มีเป้าหมายสำคัญในการพัฒนาศักยภาพและ ชีตความสามารถของทรัพยากรมนุษย์ให้มีทักษะ ความรู้ ความสามารถ และสมรรถนะที่พร้อมรับ การเปลี่ยนแปลงในโลกยุคดิจิทัล และการแข่งขันอย่างเสรีแบบไร้พรมแดนในยุคเศรษฐกิจและสังคม 4.0 (Chaiyaphum Primary Educational Service Area Office 3, 2021)

จากความเป็นมาและความสำคัญดังกล่าวข้างต้น ผู้วิจัยจึงได้เล็งเห็นความสำคัญที่จะมุ่งหวังให้ โรงเรียนประถมศึกษาได้พัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลให้สอดคล้องกับการเปลี่ยนแปลงทางเทคโนโลยี ดิจิทัล บุคลากรมีทัศนคติทางบวกต่อการนำเทคโนโลยีดิจิทัลเข้ามาใช้ในสถานศึกษาอย่างเต็มรูปแบบ โดยการศึกษาวรรณกรรมการเรียนรู้เชิงดิจิทัลเพื่อให้ทราบถึงองค์ประกอบและตัวบ่งชี้ของวัฒนธรรมการเรียนรู้ เชิงดิจิทัลที่จะนำมาเป็นกลไกในการขับเคลื่อนสถานศึกษาสู่ความสำเร็จ โดยคาดหวังผลประโยชน์เพื่อให้ โรงเรียนประถมศึกษาในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3 ได้มีแนวทางจาก ผลการศึกษาครั้งนี้มาเป็นข้อมูลในการพัฒนาโรงเรียนให้เกิดประสิทธิภาพและประสิทธิผล

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาองค์ประกอบและตัวบ่งชี้ของวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษาสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3
2. เพื่อตรวจสอบความสอดคล้องกลมกลืนของโมเดลตัวบ่งชี้ของวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3 ที่ผู้วิจัยพัฒนาขึ้นกับข้อมูลเชิงประจักษ์
3. เพื่อสร้างและประเมินรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3

วิธีดำเนินการวิจัย

การวิจัยรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3 ผู้วิจัยใช้ระเบียบวิธีวิจัยแบบผสมผสาน (mixed methods research) ด้วยแบบแผนเชิงอธิบาย (explanatory design) (Buason, 2013)

1. ประชากรและกลุ่มตัวอย่าง

การวิจัยระยะที่ 1: เป็นการศึกษาองค์ประกอบและตัวบ่งชี้ของวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา เพื่อพัฒนาองค์ประกอบ ตัวบ่งชี้ของวัฒนธรรมการเรียนรู้เชิงดิจิทัล และตรวจสอบความสอดคล้องกลมกลืนของโมเดลตัวบ่งชี้ในโรงเรียนประถมศึกษา ประชากรที่ใช้ในการวิจัยระยะนี้ คือ ผู้บริหาร และครูในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3 ปีการศึกษา 2564 ประกอบด้วยผู้บริหาร 150 คน และครู 872 คน รวมทั้งหมดจำนวน 1,022 คน กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ ผู้บริหาร และครูในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3 ปีการศึกษา 2564 ประกอบด้วย ผู้บริหาร 39 คนและครู 231 คน รวมทั้งหมด 270 คน โดยใช้เกณฑ์สำหรับกำหนดขนาดของกลุ่มตัวอย่างเป็นฟังก์ชันของจำนวนพารามิเตอร์ในการวิเคราะห์องค์ประกอบ (factor analysis) (Hair et al., 2010) ซึ่งการวิจัยครั้งนี้กำหนดขนาดของกลุ่มตัวอย่างประมาณ 15 คน ต่อ 1 พารามิเตอร์ (มีจำนวน 18 พารามิเตอร์ใช้อัตราส่วน 15 :1) ได้ขนาดของกลุ่มตัวอย่างจำนวน 270 คน ซึ่งได้มาโดยวิธีการสุ่มแบบหลายขั้นตอน (multi-stage sampling) โดยใช้การสุ่มแบบกลุ่ม (cluster sampling) ให้ได้กลุ่มตัวอย่างของแต่ละขนาดโรงเรียน และสุ่มอย่างง่าย (simple random sampling) ให้ได้กลุ่มตัวอย่างที่มีความเป็นตัวแทน โดยกำหนดขนาดกลุ่มตัวอย่างย่อยตามสัดส่วนที่เป็นกลุ่มตัวอย่างของโรงเรียน

การวิจัยระยะที่ 2: เป็นการวิจัยเชิงคุณภาพและเชิงปริมาณ โดยการวิจัยเชิงคุณภาพ เพื่อให้ได้แนวทางในการพัฒนารูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา กลุ่มเป้าหมายที่ใช้ในการวิจัยระยะนี้ โดยการเลือกแบบเจาะจงจำนวน 3 คน ประกอบด้วย ผู้ทรงคุณวุฒิด้านการบริหารการศึกษาในระดับเขตพื้นที่ขึ้นไป จำนวน 2 คน และผู้บริหารสถานศึกษาซึ่งเป็นบุคคลที่มีความรู้ ความเชี่ยวชาญการบริหารที่มีความรู้และประสบการณ์ด้านการบริหาร 10 ปีขึ้นไป จำนวน 1 คน และการวิจัย

เชิงปริมาณ เพื่อให้ได้ระดับความเหมาะสม ความเป็นไปได้ และความเป็นประโยชน์ของรูปแบบการพัฒนา
วัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา กลุ่มเป้าหมายที่ใช้ในการวิจัยระยะนี้ โดยการเลือก
แบบเจาะจง (purposive sampling) จำนวน 5 คน ได้แก่ ผู้ทรงคุณวุฒิที่มีส่วนเกี่ยวข้องกับการวางนโยบาย
ในระดับเขตพื้นที่การศึกษาขึ้นไป จำนวน 3 คน ผู้บริหารสถานศึกษาซึ่งเป็นบุคคลที่มีความรู้ความเชี่ยวชาญ
การบริหารที่มีความรู้และประสบการณ์ด้านการบริหาร 10 ปีขึ้นไป จำนวน 2 คน

2. เครื่องมือที่ใช้ในการวิจัย

การวิจัยในระยะที่ 1 เป็นแบบสอบถามแบบมาตราประมาณค่า (rating scale) 5 ระดับ จำนวน
1 ชุด มี 2 ตอน ประกอบด้วย 5 องค์ประกอบ 13 ตัวบ่งชี้ ได้ข้อคำถามทั้งสิ้น 38 ข้อ แบ่งเป็น ตอนที่ 1
แบบสอบถามเกี่ยวกับสถานภาพ ของผู้ตอบ ได้แก่ เพศ อายุ วุฒิการศึกษา ตำแหน่งปัจจุบันและประสบการณ์
ในการทำงาน มีลักษณะเป็นแบบตรวจสอบรายการ (checklist) จำนวน 5 ข้อ และตอนที่ 2 สอบถามระดับ
พฤติกรรมวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา ซึ่งมีลักษณะเป็นมาตราประมาณค่า 5 ระดับ
จำนวน 33 ข้อ การตรวจสอบคุณภาพของเครื่องมือ มีการตรวจสอบความตรงเชิงเนื้อหา (content validity)
จากผู้เชี่ยวชาญ จำนวน 5 ท่าน ได้ข้อคำถามที่มีค่า IOC อยู่ระหว่าง 0.80 ถึง 1.00 นำไปทดลองใช้ (try out)
กับผู้บริหารและครู สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 2 จำนวน 30 คน ซึ่งไม่ได้ถูก
สุ่มเป็นกลุ่มตัวอย่างจริง จากนั้นนำข้อมูลมาวิเคราะห์หาความเที่ยงโดยใช้ค่าสัมประสิทธิ์อัลฟาของครอนบาค
(Cronbach's alpha coefficient) ซึ่งได้ค่าความเที่ยงทั้งฉบับเท่ากับ 0.967

การวิจัยในระยะที่ 2 แบบสัมภาษณ์กึ่งโครงสร้าง (semi-structured interview) เพื่อร่าง
รูปแบบ และแบบประเมิน มีลักษณะเป็นแบบมาตราประมาณค่า 5 ระดับ สร้างเพื่อประเมินความเหมาะสม
ความเป็นไปได้และความเป็นประโยชน์ของรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียน
ประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3

3. การเก็บรวบรวมข้อมูล

การวิจัยในระยะที่ 1 ดำเนินการเก็บข้อมูลโดยการส่งแบบสอบถามให้กลุ่มตัวอย่างในโรงเรียน
ด้วยตนเอง ส่งแบบออนไลน์ และส่งทางกล่องจดหมายโรงเรียนของสำนักงานเขตพื้นที่การศึกษาประถมศึกษา
ชัยภูมิเขต 3 จำนวน 270 ชุด และขอความอนุเคราะห์ให้ตอบกลับภายใน 2 สัปดาห์ แล้วจึงไปขอรับ
แบบสอบถามคืนโดยขอความร่วมมือจาก ผู้บริหารโรงเรียนหรือผู้ประสานงานที่เป็นผู้รวบรวม เครื่องมือและ
ติดต่อรับคืนด้วยตนเองจากผู้ประสานงาน

การวิจัยในระยะที่ 2 ดำเนินการเก็บข้อมูลการสัมภาษณ์โดยขออนุญาตหมายผู้ทรงคุณวุฒิล่วงหน้า 1
สัปดาห์ พร้อมทั้งสอบถามถึงวิธีการในการให้สัมภาษณ์ที่ผู้ทรงคุณวุฒิสะดวก ซึ่งอาจจะเป็นการสัมภาษณ์
ออนไลน์หรือออนไลน์ขึ้นอยู่กับผู้ทรงคุณวุฒิเป็นหลัก โดยใช้เวลาในการสัมภาษณ์ 1 ชั่วโมง ต่อผู้ทรงคุณวุฒิ
1 ท่าน แล้วนำบทสัมภาษณ์มารวบรวมอย่างเป็นระบบเพื่อสะดวกต่อการวิเคราะห์ข้อมูลต่อไป และดำเนินการ
เก็บข้อมูลจากผู้ทรงคุณวุฒิในการประเมินรูปแบบโดยการส่งแบบประเมินรูปแบบให้ผู้ทรงคุณวุฒิตามเวลาและ
ช่องทางที่ได้นัดหมายไว้กับผู้ทรงคุณวุฒิตามที่แต่ละท่านสะดวก

4. การวิเคราะห์ข้อมูล

การวิจัยในระยะที่ 1 ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลเชิงปริมาณ (quantitative data) โดยใช้คอมพิวเตอร์ในการจัดกระทำข้อมูลด้วยโปรแกรมสำเร็จรูปทางสถิติ (SPSS) เพื่อหาค่าสถิติต่าง ๆ ดังนี้

1) วิเคราะห์ข้อมูลเบื้องต้นของกลุ่มตัวอย่างในแบบสอบถามตอนที่ 1 เพศ อายุ ระดับการศึกษา ตำแหน่งปัจจุบัน และประสบการณ์การทำงานในตำแหน่ง โดยคำนวณหาค่าความถี่ (frequency) และร้อยละ (percentage)

2) การวิเคราะห์ข้อมูลเกี่ยวกับความเหมาะสมของตัวบ่งชี้วัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา โดยวิเคราะห์ค่าเฉลี่ย โดยใช้เกณฑ์ในการคัดเลือกตัวบ่งชี้เพื่อกำหนดในรูปแบบที่จะนำไปทดสอบดังนี้ คือ ค่าเฉลี่ยเท่ากับหรือมากกว่า 3.00 และค่าสัมประสิทธิ์การกระจาย เท่ากับหรือต่ำกว่า 20% (Khonkan, 2004) เพื่อคัดสรรกำหนดไว้ในโมเดลความสัมพันธ์โครงสร้างเพื่อการวิเคราะห์องค์ประกอบเชิงยืนยันต่อไป

3) การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) โดยการทดสอบความสอดคล้องกลมกลืนของโมเดลตัวบ่งชี้วัฒนธรรมการเรียนรู้เชิงดิจิทัลกับข้อมูลเชิงประจักษ์ ซึ่งได้จากการวิเคราะห์ข้อมูลจากแบบสอบถามด้วยโปรแกรม Mplus เพื่อหาน้ำหนักตัวแปรย่อยที่ใช้สร้างตัวบ่งชี้ และทำการตรวจสอบความสอดคล้องกลมกลืนของโมเดลตัวบ่งชี้ที่ผู้วิจัยสร้างขึ้นด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สองกับข้อมูลเชิงประจักษ์

การวิจัยในระยะที่ 2 ผู้วิจัยนำข้อมูลที่ได้จากการสัมภาษณ์มาวิเคราะห์เชิงเนื้อหา (content analysis) แล้วนำผลที่ได้มาประกอบการร่างรูปแบบการพัฒนามาตรฐานการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา และนำร่างรูปแบบที่ได้ไปประเมินรูปแบบวิเคราะห์ข้อมูล ด้านความเหมาะสม เพื่อเป็นข้อเสนอแนะทางการพัฒนามาตรฐานการเรียนรู้เชิงดิจิทัล ให้เกิดความชอบธรรม จรรยาบรรณ และสิทธิมนุษยชน ผู้มีส่วนได้ส่วนเสีย ด้านความเป็นไปได้ เพื่อเป็นข้อเสนอที่เป็นแนวทางการพัฒนามาตรฐานการเรียนรู้เชิงดิจิทัลในการปฏิบัติ ให้เกิดประสิทธิภาพขึ้นจริงได้ และด้านความเป็นประโยชน์ เพื่อเป็นข้อเสนอที่เป็นแนวทางการพัฒนารูปแบบการพัฒนามาตรฐานการเรียนรู้เชิงดิจิทัล เกี่ยวกับประโยชน์ของสารสนเทศ ความคุ้มค่า และตอบสนองความต้องการในการนำไปใช้จริง โดยใช้สถิติพรรณนา (descriptive statistics) ได้แก่ ค่าเฉลี่ย ($\bar{X}$) และค่าเบี่ยงเบนมาตรฐาน (S.D.) เพื่อวัดการกระจายของข้อมูล โดยใช้เกณฑ์การแปลความหมายค่าเฉลี่ยของแบบสอบถามแบบมาตราส่วนประมาณค่าที่กำหนดไว้ดังนี้ (Srisa-ard, 2017)

ค่าเฉลี่ย 4.51 – 5.00	หมายความว่า	รูปแบบมีความเหมาะสม ความเป็นไปได้ ความเป็นประโยชน์ อยู่ในระดับมากที่สุด
ค่าเฉลี่ย 3.51 – 4.49	หมายความว่า	รูปแบบมีความเหมาะสม ความเป็นไปได้ ความเป็นประโยชน์ อยู่ในระดับมาก
ค่าเฉลี่ย 2.51 – 3.49	หมายความว่า	รูปแบบมีความเหมาะสม ความเป็นไปได้ ความเป็นประโยชน์ อยู่ในระดับปานกลาง
ค่าเฉลี่ย 1.51 – 2.49	หมายความว่า	รูปแบบมีความเหมาะสม ความเป็นไปได้ ความเป็นประโยชน์ อยู่ในระดับน้อย
ค่าเฉลี่ย 1.00 – 1.50	หมายความว่า	รูปแบบมีความเหมาะสม ความเป็นไปได้ ความเป็นประโยชน์ อยู่ในระดับน้อยที่สุด

พิจารณาค่าเฉลี่ยตั้งแต่ 3.51ขึ้นไป (มีระดับ ความเหมาะสม ความเป็นไปได้และความเป็นประโยชน์ อยู่ในระดับมากขึ้นไป) จึงจะถือว่ามีความเหมาะสม ผ่านเกณฑ์

ผลการวิจัย

1. ผลการศึกษาองค์ประกอบและตัวบ่งชี้วัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3 พบว่า มีจำนวน 5 องค์ประกอบ 13 ตัวบ่งชี้ ได้แก่ 1) การฝึกอบรมบุคลากรแบบผสมผสานทางดิจิทัล ประกอบด้วย 1.1) การฝึกอบรมแบบออนไลน์ 1.2) การฝึกอบรมแบบออฟไลน์ 2) ค่านิยมการทำงานร่วมกันผ่านสังคมออนไลน์ ประกอบด้วย 2.1) การสื่อสารผ่านเทคโนโลยีดิจิทัล 2.2) การเข้าถึงข้อมูลโดยใช้เทคโนโลยีสารสนเทศ 3) สมรรถนะเชิงดิจิทัลของบุคลากร ประกอบด้วย 3.1) ทักษะและการรู้ดิจิทัล 3.2) การมีส่วนร่วมผ่านเทคโนโลยีดิจิทัล และ 3.3) มุมมองและประสบการณ์ด้านเทคโนโลยี 4) ภาวะผู้นำดิจิทัล ประกอบด้วย 4.1) การพัฒนาวิชาชีพ 4.2) วิสัยทัศน์เชิงดิจิทัล และ 4.3) การส่งเสริมการใช้เทคโนโลยีดิจิทัล 5) สภาพแวดล้อมการเรียนรู้เชิงดิจิทัล ประกอบด้วย 5.1) บุคลากรเข้าถึงเทคโนโลยีที่ทันสมัย 5.2) การระดมความคิดร่วมกันผ่านเครือข่ายสังคมออนไลน์ และ 5.3) ความคล่องตัวในการทำงาน

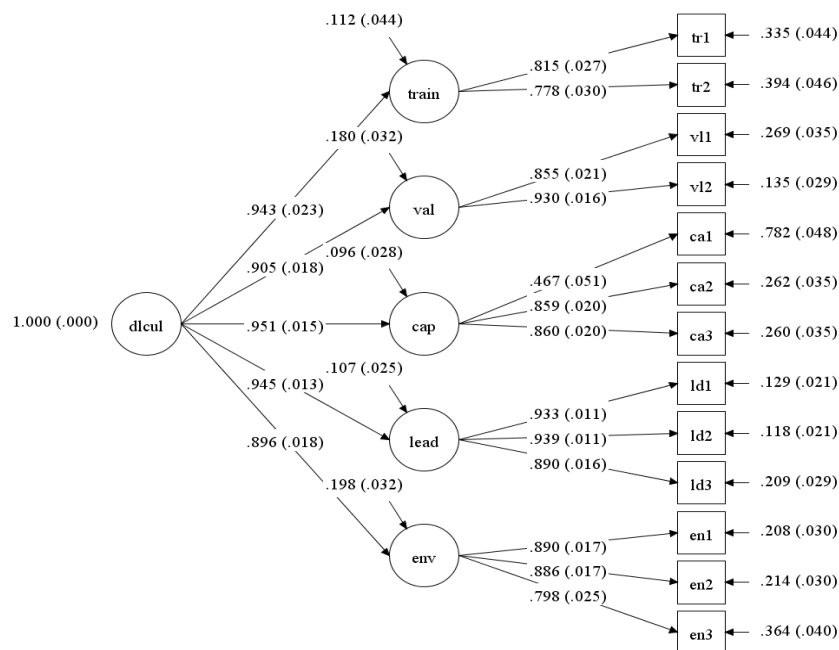
2. ผลการตรวจสอบความสอดคล้องกลมกลืนของโมเดลวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3 กับข้อมูลเชิงประจักษ์

2.1 ผู้วิจัยตรวจสอบความสอดคล้องกลมกลืนของโมเดลตามเกณฑ์ที่กำหนดของ (Hu & Bentler, 1999 as cited in Chutcham, 2004) พบว่า ค่าดัชนีความสอดคล้องกลมกลืนของโมเดลการวัดตัวบ่งชี้ของวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 เป็นไปตามเกณฑ์ที่กำหนดมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ แสดงได้ดังตาราง 1

ตาราง 1 ค่าดัชนีความสอดคล้องกลมกลืนของโมเดลวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิ เขต 3

ค่าดัชนีความกลมกลืน	เกณฑ์	ผลการวิเคราะห์	ผลการพิจารณา
χ^2 - Test	$P > 0.05$	$\chi^2 = 56.948, df = 43, p = 0.0754$	ผ่าน
χ^2/df	< 2.00	1.324	กลมกลืนดี
RMSEA	≤ 0.05	0.035	กลมกลืนดี
SRMR	≤ 0.08	0.023	สอดคล้องกับข้อมูล
CFI	≥ 0.95	0.996	กลมกลืนดี
TLI	≥ 0.95	0.992	กลมกลืนดี

2.2 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันขององค์ประกอบวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา พบว่า องค์ประกอบในรูปคะแนนมาตรฐานที่มีค่าน้ำหนักองค์ประกอบ (factor loading) ลำดับแรก ได้แก่ สมรรถนะเชิงดิจิทัลของบุคลากร ($\beta = 0.951$) รองลงมาได้แก่ ภาวะผู้นำดิจิทัล ($\beta = 0.945$) สำหรับ องค์ประกอบในรูปคะแนนมาตรฐานที่มีค่าน้ำหนักองค์ประกอบ ลำดับสุดท้าย ได้แก่ สภาพแวดล้อมการเรียนรู้เชิงดิจิทัล ($\beta = 0.896$) ดังภาพ 1 และตาราง 2



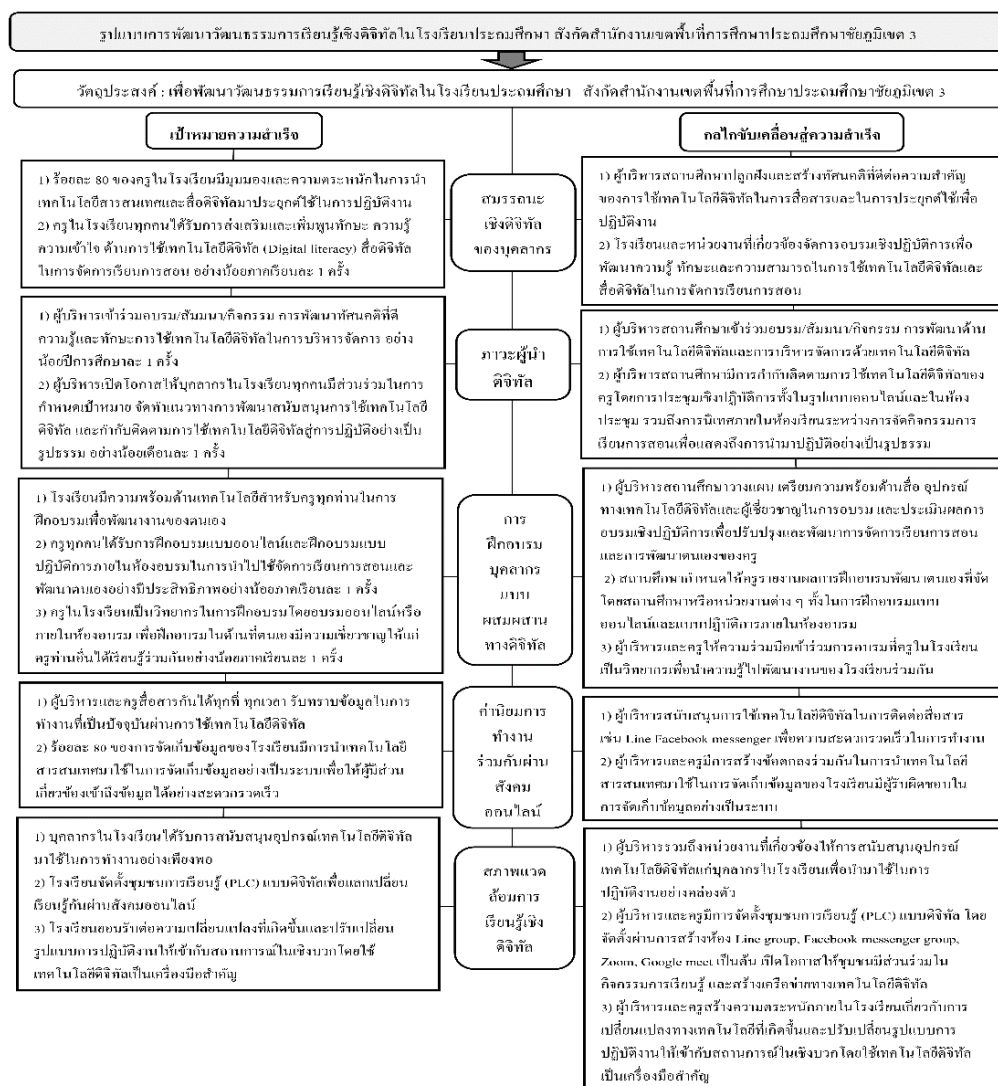
ภาพ 1 โมเดลการวัดองค์ประกอบวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา

สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3

ตาราง 2 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันขององค์ประกอบหลักของวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3

องค์ประกอบหลัก	น้ำหนักองค์ประกอบ			ค่าสัมประสิทธิ์การพยากรณ์ (R^2)	สัมประสิทธิ์ของคะแนนองค์ประกอบ (FS)
	β	S.E.	t		
1. การฝึกอบรมบุคลากรแบบผสมผสานทางดิจิทัล	0.943	0.023	40.534	0.888	0.027
2. ค่านิยมการทำงานร่วมกันผ่านสังคมออนไลน์	0.905	0.018	51.500	0.820	0.023
3. สมรรถนะเชิงดิจิทัลของบุคลากร	0.951	0.015	63.941	0.904	0.006
4. ภาวะผู้นำดิจิทัล	0.945	0.013	70.226	0.893	0.018
5. สภาพแวดล้อมการเรียนรู้เชิงดิจิทัล	0.896	0.018	49.639	0.802	0.035

3. ผลการพัฒนาและหาคุณภาพรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา ผู้วิจัยได้พัฒนารูปแบบขึ้นภายใต้ผลการศึกษาในระยะที่ 1 ซึ่งได้องค์ประกอบและตัวบ่งชี้วัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษาที่มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ นั่นคือ องค์ประกอบที่สังเคราะห์ขึ้นสามารถนำไปใช้ในการอธิบายสภาพความเป็นจริงในการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษาได้ ผู้วิจัยจึงนำผลการศึกษาดังกล่าวมาเป็นข้อมูลพื้นฐานในการพัฒนารูปแบบ ซึ่งการยกร่างรูปแบบนั้น จะเรียงลำดับความสำคัญตามค่าน้ำหนักองค์ประกอบ ผสมผสานกับการศึกษาหลักการ แนวคิด และทฤษฎีที่เกี่ยวข้องเกี่ยวกับการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา บริบทของสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 มาเป็นแนวคิดในการยกร่างรูปแบบให้มีคุณภาพและสามารถนำไปปฏิบัติได้จริง และผลการหาคุณภาพรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 พบว่า โดยภาพรวมมีความเหมาะสมอยู่ในระดับมากที่สุด ($\bar{X} = 4.93$) มีความเป็นไปได้อยู่ในระดับมากที่สุด ($\bar{X} = 4.83$.) และมีความเป็นประโยชน์อยู่ในระดับมากที่สุด ($\bar{X} = 4.93$) ซึ่งรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา มีรายละเอียดดังภาพ 2



ภาพ 2 รูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา
สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3

อภิปรายผล

1. ผลการศึกษาองค์ประกอบวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3

ผลการศึกษาองค์ประกอบวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 มีจำนวน 5 องค์ประกอบหลัก โดยผู้วิจัยได้ศึกษาแนวคิดและทฤษฎีที่เกี่ยวข้องกับวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษาอย่างหลากหลายและรัดกุมจนได้องค์ประกอบที่มีความสอดคล้องกับวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา ซึ่งสอดคล้องกับแนวคิดของ Ward (2020) ได้กล่าวถึงการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลของบุคลากรในองค์กร โดยทั่วไปซึ่งประกอบด้วยองค์ประกอบ 5 องค์ประกอบ ได้แก่ 1) กิจกรรมการฝึกอบรมบุคลากร เป็นสิ่งสำคัญในการสร้างวัฒนธรรมการเรียนรู้ มีกิจกรรมการฝึกอบรมที่ผสมผสานทางเทคโนโลยีดิจิทัลและมีความเสมือนจริงกับสถานการณ์ที่พบเจอในชีวิตประจำวัน 2) การสร้างค่านิยมการมีส่วนร่วมทางการเรียนรู้ การเรียนรู้ทางสังคมผ่านการมีส่วนร่วมกับเพื่อนร่วมงาน บุคลากรที่มีประสบการณ์มากขึ้นสามารถทำหน้าที่เป็นพี่เลี้ยงให้กับบุคลากรใหม่หรือบุคลากรที่มีผลงานไม่ดี 3) การฝึกอบรมภาวะผู้นำ ซึ่งเป็นองค์ประกอบสำคัญของการส่งเสริมวัฒนธรรมการเรียนรู้เป็นต้นแบบค่านิยมการเรียนรู้ การมีวิสัยทัศน์ในการนำเทคโนโลยีมาใช้ในองค์กร 4) วัฒนธรรมการเรียนรู้ผ่านการสื่อสาร ส่งเสริมการส่งข้อมูลข่าวสารผ่านช่องทางเทคโนโลยีดิจิทัล สร้างสภาพแวดล้อมทางการเรียนรู้ด้วยเทคโนโลยีที่เหมาะสมกับการเปลี่ยนแปลง และ 5) บุคลากรในองค์กรต้องมีทักษะที่จำเป็นในยุคดิจิทัล

2. ผลการทดสอบความสอดคล้องกลมกลืนของโมเดลการวัดตัวบ่งชี้วัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 กับข้อมูลเชิงประจักษ์ มีข้อค้นพบที่ควรนำมาอภิปรายดังนี้

จากผลการวิจัยพบว่า โมเดลการวัดตัวบ่งชี้วัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ นั่นคือองค์ประกอบที่สังเคราะห์ขึ้นมีความสอดคล้องกับบริบทของโรงเรียนประถมศึกษา โดยเมื่อเรียงตามค่าน้ำหนักองค์ประกอบ (factor loading) มีข้อค้นพบที่ควรนำมาอภิปรายผล ดังนี้

2.1 สมรรถนะเชิงดิจิทัลของบุคลากรมีค่าน้ำหนักองค์ประกอบมากที่สุด แสดงให้เห็นว่าสมรรถนะเชิงดิจิทัลของบุคลากรเป็นองค์ประกอบที่สำคัญที่สุดสำหรับการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา ทั้งนี้เนื่องจากสมรรถนะเชิงดิจิทัลของบุคลากร คือการที่บุคลากรมีความรู้ ทักษะ ทักษะคิด ต่อการนำเทคโนโลยีดิจิทัลมาประยุกต์ใช้ในการปฏิบัติงาน สร้างมุมมองและประสบการณ์ด้านเทคโนโลยีดิจิทัลให้แก่บุคลากรอย่างต่อเนื่อง พัฒนาระบบการทำงาน หรือระบบงานภายในโรงเรียนให้มีความสะดวก ทันสมัย และมีประสิทธิภาพซึ่งเป็นถือเป็นสิ่งสำคัญที่บุคลากรจะต้องมีเมื่อนำเทคโนโลยีเข้ามาใช้ ซึ่งสอดคล้องกับงานวิจัยของ Jarupoom (2016) ที่ได้ทำการศึกษาวิจัยเรื่องการศึกษาสมรรถนะของบุคลากรด้านเทคโนโลยีสารสนเทศเพื่อรองรับเศรษฐกิจดิจิทัลของหน่วยงานราชการ ได้กล่าวถึงการพัฒนาสมรรถนะบุคลากรด้านเทคโนโลยีสารสนเทศ ควรจะต้องมีการส่งเสริมในเรื่อง digital mindset ทักษะคิด

ทางบวกต่อการยอมรับการนำเทคโนโลยีดิจิทัลเข้ามาใช้ มีความสามารถในการใช้เทคโนโลยีซึ่งเป็นสิ่งสำคัญในการทำงาน การค้นหาข้อมูล รวมถึงการวิเคราะห์ข้อมูลที่มีคุณภาพจากข้อมูลที่มีอยู่มากมาย การเรียนรู้เกี่ยวกับทักษะการใช้เทคโนโลยีสารสนเทศที่ถูกต้อง ทักษะในการทำงานรูปแบบใหม่ในการใช้เทคโนโลยีเพื่อให้สามารถทำงานร่วมกันได้ในสถานที่ต่าง ๆ ทักษะการปกป้องข้อมูลรวมถึงจรรยาบรรณในการใช้งาน

2.2 องค์ประกอบหลักด้านภาวะผู้นำดิจิทัล พบว่า ตัวบ่งชี้ที่มีค่าน้ำหนักมากที่สุด คือ วิสัยทัศน์เชิงดิจิทัล เนื่องจาก การเป็นผู้นำสิ่งสำคัญคือการมีวิสัยทัศน์ที่กว้างไกลสามารถนำบุคลากรในองค์กรให้ก้าวหน้าเกิดการพัฒนายอย่างต่อเนื่อง โดยในยุคดิจิทัลผู้บริหารสถานศึกษาจะต้องมีวิสัยทัศน์ที่ถ่ายทอดออกมาแสดงถึงการนำเทคโนโลยีดิจิทัล สื่อดิจิทัล และนวัตกรรมดิจิทัลมาใช้ในการพัฒนาสถานศึกษาให้ทันต่อการเปลี่ยนแปลงของสังคมในยุคดิจิทัล ซึ่งสอดคล้องกับแนวคิดของ Thawinkarn (2021) ที่ได้กล่าวว่า ภาวะผู้นำดิจิทัลของผู้บริหารสถานศึกษา คือ การมีวิสัยทัศน์และเป็นผู้ทำให้ครูและบุคลากรที่เกี่ยวข้องนำเทคโนโลยีมาบูรณาการใช้ในการจัดการศึกษาและการทำงานภายในสถานศึกษาได้อย่างสอดคล้องกับบริบทและความต้องการของผู้เรียน ครู และผู้ที่มีส่วนเกี่ยวข้อง

2.3 องค์ประกอบหลักด้านการฝึกอบรมบุคลากรแบบผสมผสานทางดิจิทัล พบว่า ตัวบ่งชี้ที่มีค่าน้ำหนักมากที่สุด คือ การฝึกอบรมแบบออนไลน์ เนื่องจาก การฝึกอบรมแบบออนไลน์มีความยืดหยุ่นสามารถใช้เทคโนโลยีเข้ามาเป็นเครื่องมือในการสนับสนุนการฝึกอบรมได้ทุกที่ ทุกเวลา ตามความพร้อมของบุคลากร โดยบุคลากรสามารถทบทวนความรู้ได้ตลอดเวลา ซึ่งสอดคล้องกับแนวคิดของ Lakhong (2019) ที่ได้กล่าวถึง การฝึกอบรมผ่านเครือข่ายอินเทอร์เน็ต เพื่อตอบสนองความแตกต่างของผู้เข้ารับการอบรมโดยมีจุดมุ่งหมายเพื่อให้ผู้เข้ารับการอบรมทุกคนบรรลุจุดมุ่งหมายของการฝึกอบรมที่ตั้งไว้

2.4 องค์ประกอบหลักด้านค่านิยมการทำงานร่วมกันผ่านสังคมออนไลน์ พบว่า ตัวบ่งชี้ที่มีค่าน้ำหนักมากที่สุด คือ การเข้าถึงข้อมูลโดยใช้เทคโนโลยีสารสนเทศ เนื่องจาก การเข้าถึงข้อมูลโดยใช้เทคโนโลยีสารสนเทศ จะทำให้บุคลากรสามารถนำข้อมูลจากแหล่งเดียวกันมาแลกเปลี่ยนเรียนรู้กันได้อย่างตรงประเด็น สามารถค้นหาข้อมูลได้สะดวกเกิดการทำงานร่วมกันได้อย่างรวดเร็ว ได้รับข้อมูลที่เป็นปัจจุบันมีความคล่องตัวในการทำงาน ซึ่งสอดคล้องกับแนวคิดของ Sathukitchai (2020) ได้กล่าวถึง ค่านิยมการทำงานร่วมกันผ่านสังคมออนไลน์ คือ การมีพฤติกรรมการนำเทคโนโลยีมาใช้ในการปรับรูปแบบการทำงานเพิ่มประสิทธิภาพในการทำงาน มีแพลตฟอร์มในการทำเอกสารออนไลน์โดยบุคลากรทุกคนสามารถทำงานอยู่บนเอกสารเดียวกันได้แม้จะอยู่ต่างสถานที่กัน มีการแบ่งปันงานให้แก่กันผ่านแพลตฟอร์มการแชร์ไฟล์งานด้วยเทคโนโลยีในรูปแบบต่าง ๆ

2.5 องค์ประกอบหลักด้านสภาพแวดล้อมการเรียนรู้เชิงดิจิทัล พบว่า ตัวบ่งชี้ที่มีค่าน้ำหนักมากที่สุด คือ บุคลากรเข้าถึงเทคโนโลยีที่ทันสมัย เนื่องจาก เมื่อบุคลากรเข้าถึงเทคโนโลยีที่ทันสมัยจะทำให้เกิดการเรียนรู้และพัฒนางาน เกิดนวัตกรรมใหม่ ๆ ขึ้นในสถานศึกษา บุคลากรสามารถเข้าถึงองค์ความรู้จากที่ต่าง ๆ ได้ง่ายขึ้น พัฒนาทักษะได้ตรงตามงานและการจัดการเรียนการสอนของตนเอง ซึ่งสอดคล้องกับแนวคิดของ Wit & Dompsele (2017) ได้กล่าวถึงสภาพแวดล้อมการเรียนรู้เชิงดิจิทัล หมายถึง บุคลากร

สามารถเข้าถึงเทคโนโลยีที่ทันสมัย มีการนำเทคโนโลยีเข้ามาเป็นส่วนสำคัญในการเรียนรู้และการทำงานของบุคลากร มีเทคโนโลยีเป็นสื่อกลางในการแลกเปลี่ยนข้อมูลกับองค์กรภายนอก

3. ผลการสร้างและประเมินรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3

รูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา ประกอบไปด้วย 1) ชื่อรูปแบบ 2) หลักการและแนวคิดพื้นฐานของรูปแบบ 3) วัตถุประสงค์ของรูปแบบ 4) เป้าหมายสู่ความสำเร็จ และ 5) กลไกขับเคลื่อนสู่ความสำเร็จ ซึ่งรูปแบบมีความเหมาะสม ความเป็นไปได้และเป็นประโยชน์อยู่ในระดับมากที่สุดทุกด้าน ทั้งนี้เป็นเพราะว่า กระบวนการสร้างรูปแบบได้ผ่านขั้นตอนและวิธีการ โดยมีผู้เชี่ยวชาญให้สัมภาษณ์เพื่อยกร่างรูปแบบและตรวจสอบประเมินรูปแบบอย่างถูกต้องตามหลักวิชาการ จึงทำให้รูปแบบผ่านการประเมิน โดยอิงแนวคิดทฤษฎีที่ว่าด้วยรูปแบบ (model) ของ Keeves (1997) ซึ่งการประเมินรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา มีการเลือกใช้การประเมินรูปแบบตามแนวคิดของ The Joint Committee on Standard for Education Evaluation (1994) ซึ่งผู้วิจัยได้ปรับใช้ 3 มาตรฐาน ดังนี้ 1. มาตรฐานความเป็นไปได้ (feasibility standards) เป็นการประเมินความเป็นไปได้ในการนำไปปฏิบัติจริง จากผลการศึกษาพบว่าความเป็นไปได้ของรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา นั้นอยู่ในระดับมากที่สุด แสดงให้เห็นว่ารูปแบบที่ผู้วิจัยได้สร้างขึ้นมีความเป็นไปได้ในการนำไปปฏิบัติจริงในบริบทของโรงเรียนประถมศึกษาสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 ในระดับมากที่สุด 2. มาตรฐานด้านความเป็นประโยชน์ (utility standards) เป็นการประเมินการตอบสนองต่อความต้องการของผู้ใช้รูปแบบ จากผลการศึกษาพบว่าความเป็นประโยชน์ของรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา นั้นอยู่ในระดับมากที่สุด แสดงให้เห็นว่ารูปแบบที่ผู้วิจัยได้สร้างขึ้นสามารถตอบสนองต่อความต้องการของผู้ใช้รูปแบบไปใช้ นั่นก็คือผู้บริหารสถานศึกษาและครูในโรงเรียนประถมศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 ในระดับมากที่สุด และ 3. มาตรฐานด้านความเหมาะสม (propriety standards) เป็นการประเมินความเหมาะสมทั้งในด้านกฎหมายและศีลธรรมจรรยา จากผลการศึกษาพบว่าความเหมาะสมของรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา นั้นอยู่ในระดับมากที่สุด แสดงให้เห็นว่ารูปแบบที่ผู้วิจัยได้สร้างขึ้นมีความเหมาะสมทั้งในด้านกฎหมายและศีลธรรมจรรยา ในระดับมากที่สุด

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 สำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 สามารถนำรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษา ไปใช้เป็นแนวทางกำหนดนโยบาย ยุทธศาสตร์ หรือติดตามภารกิจ เพื่อประกอบการตัดสินใจ และประเมินผลการดำเนินงานในการพัฒนาองค์กร

1.2 โรงเรียนควรนำองค์ประกอบที่มีค่าน้ำหนักองค์ประกอบมากที่สุดมาพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัล เรียงตามลำดับ ได้แก่ สมรรถนะเชิงดิจิทัลของบุคลากร ภาวะผู้นำดิจิทัล การฝึกอบรม

บุคลากรแบบผสมผสานทางดิจิทัล คำนึงการทำงานร่วมกันผ่านสังคมออนไลน์ และสภาพแวดล้อมการเรียนรู้เชิงดิจิทัล

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรมีการนำรูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลในโรงเรียนประถมศึกษาสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาชัยภูมิเขต 3 ไปทำวิจัยและพัฒนา (R&D) เพื่อการวิจัยต่อยอดหรือพัฒนาให้สมบูรณ์ยิ่งขึ้นต่อไป

2.2 ควรมีการศึกษาวิจัยเกี่ยวกับวัฒนธรรมการเรียนรู้เชิงดิจิทัลตามขนาดโรงเรียนระดับต่าง ๆ หรือในบริบทโรงเรียนเอกชนเพื่อให้ได้รูปแบบการพัฒนาวัฒนธรรมการเรียนรู้เชิงดิจิทัลที่เหมาะสมตามบริบทของขนาดโรงเรียน

References

- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis*. (7th Ed.). Prentice Hall.
- Keeves, J. P. (1997). *Models and model building*. In Keeves, J.P. (Eds.). *Educational research, methodology and measurement: An International Handbook* (2nd Ed.). Oxford Peraman Press.
- The Joint Committee on Standard for Education Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs*. Sage.
- Ward, G. (2020). *How To Build A Digital Learning Culture In Your Organization*.
<https://elearningindustry.com/build-digital-learning-culture-organization>
- Wit, M. D. & Dompsele, H. V. (2017). *How to create a digital learning environment consisting of various components and acting as a whole?*.
https://www.eunis.org/download/2017/EUNIS_2017_paper_16.pdf

Translate Thai References

- Buason, R. (2013). *Qualitative research in education* (4th Ed.). Chulalongkorn University. (in Thai)
- Chaiyaphum Primary Educational Service Area Office 3. (2021). *The 5-year Educational Development Plan B.E. 2564 - 2568 (A.D. 2021-2025)*.
https://web.chaiyaphum3.go.th/files/pat_plan65.pdf (in Thai)
- Chutcham, S. (2004). A Confirmatory Factor Analysis. *Research Educational Measurement and Evaluation journal*, 2(1), 15-42.

- Jarupoom, A. (2016). *A Study of the Competency of Information Technology Personnel to Support the Digital Economy of Government Agencies: A Case Study of the Center for Information and Communication Technology, Office of the Permanent Secretary, Ministry of Finance* [Master's Independent Research]. Thammasat University. (in Thai)
- Khonkan, S. (2004). *The Development of Indicators of State Higher Education Institutions*. [Doctoral dissertation]. Srinakharinwirot University. (in Thai)
- Lakthong, S. (2019). *Developing an integrated training model based on self-leading principles to enhance digital literacy capabilities For teachers under the Office of the Basic Education Commission* [Doctor's dissertation]. Mahasarakham University. (in Thai)
- Ministry of Information and Communication Technology. (2016). *Digital Development Plan for Economy and Society*. Ministry of Information and Communication Technology. (in Thai)
- Office of the Permanent Secretary for Ministry of Agriculture and Cooperatives. (2012). *Organizational Culture Analysis*. https://www.opsmoac.go.th/km-km_org_center-files-391991791962 (in Thai)
- Sathukitchai, k. (2020). *10 ways to cope with Work From Home with technology*. <https://www.smethailandclub.com/startups-5578-id.html> (in Thai)
- Srisa-ard, B. (2017). *Preliminary research* (10th Ed.). Suwiriyasan. (in Thai)
- Thawinkarn, D. (2021). *Digital Leadership* (2nd Ed.). Teaching Innovation Center Khon Kaen University. (in Thai)

การเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้

Enhancing Complex Problem-Solving Skills of Adolescents by Applying Information Processing Theory and Self-Regulation Learning

วิหวัธ เพ็ญภู^{1*} พูลพงศ์ สุขสว่าง^{2*} และ พีร วงศ์อุปราช³

Withawat Penphu^{1*}, Poonpong Suksawang^{2*} and Peera Wongupparaj³

(Received: May 30, 2022; Revised: July 22, 2022; Accepted: August 1, 2022)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมายเพื่อพัฒนาและศึกษาผลการใช้โปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ และศึกษาผลของโปรแกรมต่อการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่น กลุ่มตัวอย่างเป็นวัยรุ่นอายุ 13 – 17 ปี จำนวน 60 คน สุ่มเข้ากลุ่มทดลอง 30 คน และกลุ่มควบคุม 30 คน เก็บรวบรวมข้อมูลจากแบบทดสอบทักษะการแก้ปัญหาที่ซับซ้อน COMPRO วิเคราะห์ข้อมูลด้วยสถิติ Pair – sample t-test ANOVA MANOVA และ MANCOVA ผลการวิจัยปรากฏว่า 1) โปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นแบ่งออกเป็น 3 ระยะ ได้แก่ ระยะที่ 1 การวางแผนกำหนดเป้าหมายและการวางกลยุทธ์ ระยะที่ 2 การตรวจสอบประสิทธิภาพ การประเมินกลยุทธ์

¹ นิสิตระดับดุษฎีบัณฑิต สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา หน่วยวิจัยวิทยาการปัญญาและนวัตกรรม วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

² รองศาสตราจารย์ สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา หน่วยวิจัยวิทยาการปัญญาและนวัตกรรม วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

³ อาจารย์ สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา หน่วยวิจัยวิทยาการปัญญาและนวัตกรรม วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

ได้รับการสนับสนุนทุนวิจัยและนวัตกรรม ทุนพัฒนาบัณฑิตศึกษา ประจำปีงบประมาณ 2564 จากสำนักงานคณะกรรมการวิจัยแห่งชาติ (วช.)

ได้รับการสนับสนุนทุนอุดหนุนการวิจัยระดับบัณฑิตศึกษา มหาวิทยาลัยบูรพา ประจำปีงบประมาณ 2563

¹ Doctoral degree student, Research and Statistics in Cognitive Science, Cognitive Science and Innovation Research Unit, College of Research Methodology and Cognitive Science, Burapha University

² Associate Professor, Research and Statistics in Cognitive Science, Cognitive Science and Innovation Research Unit, College of Research Methodology and Cognitive Science, Burapha University

³ Lecturer, Research and Statistics in Cognitive Science, Cognitive Science and Innovation Research Unit, College of Research Methodology and Cognitive Science, Burapha University

The research received a grant for research and innovation from the Graduate Study Development Fund, fiscal Year 2021, National Research Council of Thailand (NRCT)

The research also received a graduate research grant from Burapha University, fiscal year 2020.

*Corresponding Author E-mail: psuksawang@gmail.com

และปรับปรุงกลยุทธ์ ระยะที่ 3 การประยุกต์ใช้ความรู้ โดยโปรแกรมมีความเหมาะสม ในการนำไปใช้เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่น (S-CVI=.90) และ 2) ทักษะการแก้ปัญหาที่ซับซ้อนในกลุ่มทดลอง ภายหลังการฝึกสูงกว่ากลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($F_{Wilk'Lambda}=6.26, p=.00$)

คำสำคัญ: การแก้ปัญหาที่ซับซ้อน ทฤษฎีประมวลสารสนเทศ การกำกับตนเองในการเรียนรู้

Abstract

This research aimed to develop and investigate the effects of a computer program for enhancing complex problem-solving skills of adolescents by applying information processing theory and self-regulation learning. The sample consisted of 60 adolescents 13 – 17 years of age. Thirty adolescents were randomly assigned into the experimental group and the other 30 into the control group. Data of complex problem-solving skills were collected, using COMPRO. The analysis of data employed pair-sample t-test, ANOVA, MANOVA and MANCOVA. The research results revealed that: 1) the computer program developed consisted of 3 phases: Phase 1 dealt with planning, setting goals and strategies. Phase 2 dealt with examining efficiency, assessing strategies and improve strategies. Phase 3 dealt with application of knowledge. The computer program was suitable for enhancing complex problem-solving skills in adolescence (S-CVI=.90); and 2) the complex problem-solving scores of the experimental group, after training, were higher than those of the control group, with statistical significance at the .05 level ($F_{Wilk'Lambda}=6.26, p=.00$).

Keywords: complex problem solving, Information Processing Theory, self-regulation learning

บทนำ

ทุกวันนี้โลกของเราขับเคลื่อนไปข้างหน้าอย่างรวดเร็ว การปรับตัวเพื่อเพิ่มทักษะที่จำเป็นในศตวรรษที่ 21 กลายเป็นหัวใจสำคัญของความสำเร็จที่คนรุ่นใหม่ต้องใส่ใจและพัฒนาอย่างรวดเร็ว ซึ่งจำเป็นต้องมีศักยภาพในหลาย ๆ ด้าน เช่น ความคิดสร้างสรรค์ การเรียนรู้ตลอดชีวิต และทักษะในการทำงานแบบองค์รวม (collaboration skills) (Griffin & Care, 2014) และความสามารถอย่างหนึ่งที่มีความสำคัญอย่างยิ่ง คือความสามารถในการแก้ปัญหาที่ซับซ้อน (complex problems solving skill) (Mainzer, 2009) World Economic Forum (2016) รายงานว่า โลกสมัยใหม่ที่เรอาศัยอยู่นี้ปัญหาต่างๆ ที่ต้องพบเจอล้วนไม่ใช่เพียงแค่ปัญหาที่มีสาเหตุหรือปัจจัยเพียงหนึ่งเดียว แต่ต้องเจอกับปัญหาที่มีความซับซ้อนเพิ่มมากขึ้น ส่งผลให้เกิดความเสี่ยงในการชีวิตในสังคม ดังนั้นทักษะที่จำเป็นมากสำหรับโลกในอนาคตจึงเป็นทักษะการแก้ปัญหาที่ซับซ้อน

Frensch & Funke (1995) ได้นิยามปัญหาที่ซับซ้อนคืออุปสรรคที่ขวางอยู่ระหว่างช่วงสถานะการเริ่มต้น (start state) กับสถานะเป้าหมายที่ตั้งไว้ (goal state) การแก้ปัญหาที่ซับซ้อนต้องอาศัยกิจกรรมทางปัญญา (cognitive activities) และพฤติกรรม (behavior) ถ้าเป็นการแก้ปัญหาแบบธรรมดา คือการแก้ปัญหาที่รู้ว่าจะไปถึงเป้าหมายต้องเจออุปสรรคอะไรบ้าง แต่ถ้าเป็นการแก้ปัญหาที่ซับซ้อนจะไม่สามารถรู้ถึงอุปสรรคระหว่างทางได้ และอุปสรรคนั้นยังสามารถเปลี่ยนแปลงไปตามช่วงเวลาได้ การแก้ปัญหาที่ซับซ้อนต้องใช้ทั้งกระบวนการทางปัญญา อารมณ์ และ ความรู้พื้นฐานของแต่ละบุคคล หรืออีกนัยหนึ่งกล่าวได้ว่า ปัญหาแบบธรรมดาคือปัญหาที่ถูกกำหนดคำตอบของปัญหาไว้อย่างชัดเจน แต่การแก้ปัญหาแบบซับซ้อนคำตอบของปัญหาจะไม่ได้ถูกกำหนดไว้ชัดเจนตายตัว ซึ่งความแตกต่างระหว่างปัญหาที่ถูกกำหนดไว้อย่างชัดเจนกับปัญหาที่ไม่ได้กำหนดไว้อย่างชัดเจนคือ ปัญหาที่ถูกกำหนดไว้อย่างชัดเจน (well-defined problems) มีแนวทางอย่างชัดเจนและแม่นยำในการไปให้ถึงเป้าหมายที่กำหนดไว้ เช่น การแก้ปัญหาทางคณิตศาสตร์ที่ถูกกำหนดคำตอบไว้อย่างชัดเจน แต่ปัญหาที่ไม่ได้กำหนดไว้อย่างชัดเจน (ill-defined problems) เป้าหมายจะไม่ได้ถูกกำหนดไว้อย่างชัดเจน ตัวอย่างเช่น การแก้ไขปัญหาคความขัดแย้งทางการเมืองแบบนี้อาจมีทางออกของปัญหาได้หลายทาง ปัญหาประเภทนี้เรียกว่าปัญหาที่ซับซ้อน (complex problem) การแก้ไขปัญหที่ต้องอาศัยกระบวนการทางจิตภายในแต่ละบุคคลและการจัดการกับปัญหาที่ไม่มีความชัดเจนอยู่ภายใต้คำจำกัดความของการแก้ไขปัญหาที่ซับซ้อน (complex problem solving)

การพัฒนาบุคลากรให้สอดคล้องกับความต้องการของภาคอุตสาหกรรมต้องให้ความสำคัญกับการพัฒนาทักษะที่จำเป็นในอนาคต (future skills development) จากผลสำรวจ Future of Jobs Report 2018 ของ World Economic Forum ในประเด็นความต้องการของทักษะอนาคต ในปี ค.ศ. 2022 พบว่ามากกว่า 1 ใน 3 หรือ 36% ของงานในทุกภาคอุตสาหกรรม จำเป็นต้องใช้ทักษะการแก้ไขปัญหาที่ซับซ้อน (World Economic Forum, 2018)

การแก้ไขปัญหที่ซับซ้อนจะเน้นถึงการแก้ปัญหาในโลกแห่งความเป็นจริง ต้องมีการรวบรวมกระบวนการทางจิตวิทยาที่ใช้ในการควบคุมตนเอง (self-regulated) และกิจกรรมที่จำเป็นในสภาพแวดล้อมที่เปลี่ยนแปลงไปตลอดเวลา (dynamic environments) เพื่อให้สามารถบรรลุเป้าหมายที่ไม่สามารถกำหนดคำตอบของปัญหาได้อย่างชัดเจนที่เรียกว่า “ill-defined” การผสมผสานความคิดสร้างสรรค์และกลยุทธ์ต่าง ๆ เป็นสิ่งจำเป็นการแก้ปัญหาจึงจะมีความสมบูรณ์ กระบวนการแก้ปัญหาประกอบด้วยกระบวนการทางปัญญา (cognitive) อารมณ์ (emotional) และการสร้างแรงบันดาลใจ (motivational aspects) ตลอดจนตัวแปรที่ไม่เกี่ยวข้องกับการกระบวนการทางปัญญา (non-cognitive) ยิ่งในสถานการณ์ที่มีการเติมพ้นสูงก็ยิ่งส่งผลต่อการแก้ปัญหาอีกด้วย (Dörner & Funke, 2017)

วัยรุ่นเป็นกำลังสำคัญที่จะต้องเติบโตขึ้นไปเป็นผู้ที่ทำหน้าที่ร่วมพัฒนาประเทศชาติบ้านเมืองต่อไปในอนาคต การวางแผนพัฒนาเพื่อวางรากฐานให้กลุ่มวัยรุ่นมีคุณลักษณะอันพึงประสงค์เพื่อเติบโตเป็นผู้ใหญ่ที่ดีของประเทศชาติโดยสมบูรณ์เป็นสิ่งที่จำเป็น โดยเฉพาะในยุคแห่งการเปลี่ยนแปลงทางสังคมที่เกิดขึ้นในศตวรรษที่ 21 ที่ส่งผลกระทบต่อวิถีชีวิตของคนในสังคมอย่างทั่วถึง วัยรุ่นควรมีทักษะการคิดขั้นสูง เทคโนโลยีมีล้นเต็ม ความรู้พื้นฐานเชิงพหุ และความสามารถในการแก้ปัญหาที่ซับซ้อน ประเทศไทยได้มีการ

กำหนดยุทธศาสตร์ชาติ 20 ปี (พ.ศ. 2561-2580) เป็นยุทธศาสตร์ชาติฉบับแรกของประเทศไทย ตามรัฐธรรมนูญแห่งราชอาณาจักรไทย ซึ่งจะต้องนำไปสู่การปฏิบัติเพื่อให้ประเทศไทยบรรลุวิสัยทัศน์ โดยเฉพาะในยุทธศาสตร์ชาติด้านการพัฒนาและเสริมสร้างศักยภาพทรัพยากรมนุษย์มีเป้าหมายการพัฒนาที่สำคัญเพื่อพัฒนาคนในทุกมิติและในทุกช่วงวัยให้เป็นคนดี เก่ง และมีคุณภาพ มีทักษะที่จำเป็นในศตวรรษที่ 21 โดยเฉพาะอย่างยิ่งทักษะที่สำคัญที่สุดคือการแก้ปัญหาที่ซับซ้อน

จากการศึกษาการวิธีการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อน พบว่า มีอยู่ด้วยกัน 2 วิธี ได้แก่ 1) วิธีการฝึกอบรม และ 2) วิธีการฝึกการแก้ปัญหาด้วยโปรแกรมคอมพิวเตอร์ วิธีการฝึกอบรมนั้นยังพบข้อจำกัดอยู่ เนื่องจากวิธีการฝึกอบรมใช้ทรัพยากรค่อนข้างมากในการจัดแต่ละครั้ง วิธีที่น่าสนใจกว่าคือการฝึกด้วยโปรแกรมคอมพิวเตอร์ซึ่งมีความสะดวกสบายมากกว่า แต่จากการศึกษางานวิจัยที่เกี่ยวข้องกับการฝึกการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนด้วยคอมพิวเตอร์นั้น พบว่ามีโปรแกรมเพียงไม่กี่โปรแกรมที่มีผลงานวิจัยรองรับ และโปรแกรมห้างมีวัตถุประสงค์ในการสร้างขึ้นมาเพื่อทดสอบทักษะการแก้ปัญหาที่ซับซ้อนเท่านั้น ยังไม่พบโปรแกรมคอมพิวเตอร์ที่สร้างขึ้นมาเพื่อใช้ฝึกทักษะการแก้ปัญหาที่ซับซ้อนโดยตรง (Kretzschmar & Süß, 2015)

จากการศึกษาทฤษฎีที่มีความสัมพันธ์เกี่ยวกับการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนคือทฤษฎีประมวลสารสนเทศ (information processing theory) ซึ่งมีความเกี่ยวข้องกับการวางแผนกลยุทธ์ การแก้ปัญหาที่ซับซ้อน (strategies) การติดตามงาน (monitoring) และการประเมินงาน (evaluation) การที่จะทำให้เกิดการวางแผนกลยุทธ์ในการแก้ปัญหาได้ดี สิ่งสำคัญคือต้องมีการให้ความรู้ ซึ่งทำให้เกิดเนื้อหาความจำ (memory contents) ทำให้เกิดสร้างองค์ความรู้ในการแก้ปัญหาที่ซับซ้อน (Klausmeier, 1985) และการแก้ไขปัญหาก็จะเน้นถึงการแก้ปัญหาในโลกแห่งความเป็นจริง ต้องใช้กระบวนการทางจิตวิทยาในการกำกับตนเอง (self-regulated) (Dörner & Funke, 2017) จากการศึกษาวิจัยของ ElAdl & Polpol (2020) ที่ศึกษาผลของการใช้กลยุทธ์การกำกับตนเองในการเรียนรู้ต่อทักษะการแก้ปัญหา พบว่า เด็กนักเรียนที่ได้รับการฝึกโดยใช้กลยุทธ์การกำกับตนเองในการเรียนรู้เป็นเวลา 12 สัปดาห์ มีทักษะการแก้ปัญหา ความคิดสร้างสรรค์และการรับรู้ความสามารถของตนเอง (self-efficacy) มากกว่ากลุ่มที่ไม่ได้รับการฝึก ดังนั้นผู้วิจัยจึงนำทฤษฎีการกำกับตนเองในการเรียนรู้ (self-regulation learning) (Zimmerman, 2002) มาประยุกต์ใช้สำหรับการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อน

จากที่กล่าวมาข้างต้น ผู้วิจัยจึงมีความสนใจพัฒนาโปรแกรมคอมพิวเตอร์สำหรับเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยใช้ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ ทั้งนี้เพื่อให้ได้องค์ความรู้ใหม่และวิธีเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นอย่างมีประสิทธิภาพ

ความมุ่งหมายของการวิจัย

1. เพื่อพัฒนาโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้
2. เพื่อศึกษาผลการใช้โปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้

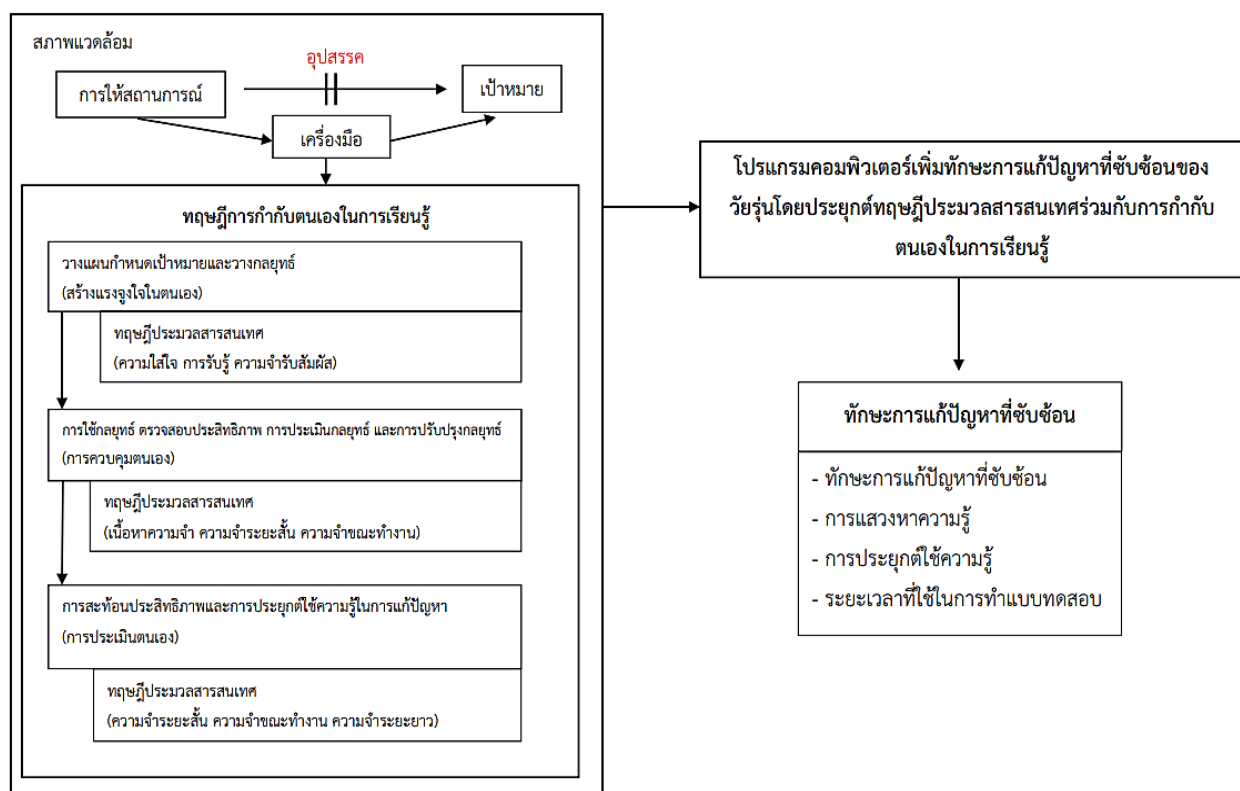
กรอบแนวคิดการวิจัย

การพัฒนาโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ มีแนวคิดในการสร้างโปรแกรมโดยการให้สถานการณ์ (situation) และต้องเผชิญกับอุปสรรค (barriers) ที่มีอยู่ระหว่างการแก้ปัญหาที่กำหนด เพื่อไปถึงเป้าหมาย (goals) โดยอุปสรรคจะมีความซับซ้อนเปลี่ยนแปลงอยู่ตลอดเวลา โดยต้องอยู่ในสภาพแวดล้อม (environment) เช่น ความคาดหวัง ความร่วมมือ ความกดดันจากเพื่อน หรือ สิ่งรบกวนอื่นๆ ผู้แก้ปัญหามustใช้เครื่องมือ (tools) เพื่อทำให้บรรลุเป้าหมายที่กำหนด (Frensch & Funke, 1995) ซึ่งเครื่องมือของผู้แก้ปัญหานั้นใช้องค์ประกอบของทฤษฎีประมวลสารสนเทศ (information processing theory) ซึ่งเกี่ยวข้องกับ การวางแผนกลยุทธ์ (strategies) การติดตาม (monitoring) และการประเมิน (evaluation) การที่จะทำให้ เกิดการวางแผนกลยุทธ์ได้ดี มีการติดตามงานเพื่อให้เป็นไปตามกลยุทธ์และสามารถประเมินงานได้นั้น สิ่งสำคัญคือต้องมีการให้เนื้อหาความจำ (memory contents) เพื่อทำให้เกิดองค์ความรู้ การทำให้เกิดความจำ อธิบายได้ด้วยทฤษฎีประมวลสารสนเทศ เริ่มจากการรับสิ่งเร้าผ่านความจำรับสัมผัส (sensory memory) โดยต้องมีความใส่ใจ (attention) และ การรับรู้ (perception) เป็นปัจจัยสำคัญในการรับข้อมูล เพื่อเก็บข้อมูล ไว้ในความจำระยะสั้น (short-term memory) และในความจำระยะสั้นนี้เนื่องจากการประมวลผลข้อมูลจะมีการดึงความจำชั่วขณะหนึ่งออกมาใช้งาน เรียกว่าความจำขณะทำงาน (working memory) ก่อนที่สมองจะ บันทึกข้อมูลในความจำระยะยาว (long-term memory) ผ่านการทำซ้ำ (rehearsal) หรือ การเข้ารหัส (encoding) และเมื่อมีการเรียกคืน (retrieval) ข้อมูลในความจำระยะยาวเพื่อนำออกมาใช้ ข้อมูลจะถูก ส่งกลับไปยังความจำขณะทำงานเพื่อตอบสนอง กระบวนการนี้มีความสัมพันธ์กับการเข้ารหัส หากการเข้ารหัส ทำให้เกิดการเก็บจำได้ดีมีประสิทธิภาพการเรียกคืนก็จะมีประสิทธิภาพตามไปด้วย (Klausmeier, 1985)

การแก้ไขปัญหที่ซับซ้อนจะเน้นถึงการแก้ปัญหาในโลกแห่งความเป็นจริง ต้องมีการรวบรวม กระบวนการทางจิตวิทยาที่ใช้ในการกำกับตนเอง (self-regulated) (Dörner & Funke, 2017) เพื่อใช้ในการ ควบคุมกำกับให้สามารถทำตามกระบวนการประมวลสารสนเทศได้อย่างมีประสิทธิภาพ โดยงานวิจัยนี้ได้ใช้ ทฤษฎีการกำกับตนเองในการเรียนรู้ (self-regulated learning) โดยใช้หลักวงจรการกำกับตนเองในการ เรียนรู้ เริ่มจาก 1) การวางแผนกำหนดเป้าหมายและวางแผนกลยุทธ์ การทำแผนก่อนเริ่มทำงานจะช่วยให้พวกเขา วางแผนได้ตั้งแต่เริ่มต้น กระตุ้นให้เกิดความพยายามและแรงจูงใจในตนเอง (self-motivation) 2) การใช้ กลยุทธ์ ตรวจสอบประสิทธิภาพ การประเมินกลยุทธ์ และการปรับปรุงกลยุทธ์ ในขั้นตอนนี้คือการทำตามแผน ที่ระบุไว้ในขั้นตอนแรก จะทำให้เกิดความมั่นใจเกิดการควบคุมตนเอง (self-control) ให้เป็นไปตามแผน กลยุทธ์ที่กำหนด การตรวจสอบประสิทธิภาพช่วยให้เกิดการสังเกตตนเอง (self-observation) กระตุ้นให้ยึด ติดกับกลวิธีที่ตั้งเป้าหมายไว้และเป็นการตรวจสอบตนเองเกี่ยวกับความก้าวหน้าในการดำเนินการตามกลยุทธ์ และ 3) การสะท้อนประสิทธิภาพและการประยุกต์ใช้ความรู้ในการแก้ปัญหา ช่วยให้เกิดการประเมิน ประสิทธิภาพของตนเองและผลลัพธ์ที่ได้ สะท้อนประสิทธิผลของกลยุทธ์ที่ใช้ เช่นเลือกกลยุทธ์ที่เหมาะสม หรือไม่ ปฏิบัติตามด้วยกลยุทธ์ที่เลือกไว้หรือไม่ ช่วยจัดการอารมณ์ของตนเอง ทำให้เกิดการประเมินตนเอง (self-assessment) นำไปสู่แนวความคิดที่มีประสิทธิผลเกี่ยวกับการปรับปรุงประสิทธิภาพการทำงานของ

ตัวเองได้ ส่วนสำคัญของกระบวนการนี้คือ เพื่อวางแผนปรับปรุงกลยุทธ์ให้มีประสิทธิภาพสำหรับงานต่อ ๆ ไป (Zimmerman, 2002)

จากแนวคิดของ Fischer et al., (2012) กล่าวว่าทักษะการแก้ปัญหาที่ซับซ้อนมุ่งเน้นไปที่สองด้านที่เป็นศูนย์กลางและมีการเชื่อมโยงกัน ซึ่งได้แก่ การแสวงหาความรู้ (knowledge acquisition) และการประยุกต์ใช้ความรู้ (knowledge application) การแสวงหาความรู้เป็นทักษะที่ช่วยให้ผู้คนสามารถสำรวจและเข้าใจการทำงานของระบบที่ซับซ้อนในขณะที่การประยุกต์ใช้ความรู้เป็นสิ่งจำเป็นเพื่อให้สามารถเข้าใจและควบคุมสภาพแวดล้อมในการแก้ปัญหาได้ ตลอดจนระยะเวลาที่ใช้ในการทำแบบทดสอบการแก้ปัญหาที่ซับซ้อนซึ่งถือเป็นตัวบ่งชี้หนึ่งของความสามารถในการแก้ปัญหาของมนุษย์ ซึ่งวัดได้จากแบบทดสอบทักษะการแก้ปัญหาที่ซับซ้อนมาตรฐาน COMPRO: Complex Problem Solving Test (Greiff & Wüstenberg, 2015) ดังภาพ 1



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษา โรงเรียนสาธิตพิบูลบำเพ็ญ มหาวิทยาลัยบูรพา จังหวัดชลบุรี ปีการศึกษา 2564 เพศชายและเพศหญิง อายุ 13-17 ปี ที่มีคุณสมบัติตามเกณฑ์การคัดเลือกเข้า (inclusion criteria) ได้แก่ สนใจเข้าร่วมการทดลอง และไม่มีความพิการทางด้านร่างกายที่ส่งผลกระทบต่อ การทดลอง จำนวน 60 คน สุ่มเข้ากลุ่มทดลองและกลุ่มควบคุม กลุ่มละ 30 คน โดยงานวิจัยนี้ได้ผ่านการรับรอง

จากคณะกรรมการพิจารณาจริยธรรมการวิจัยในมนุษย์ของมหาวิทยาลัยบูรพา และกลุ่มตัวอย่างทุกคนลงนามยินยอมการเข้าร่วมโครงการวิจัย

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยนี้แบ่งเครื่องมือที่ใช้ในการวิจัยออกเป็น 3 ประเภท ประกอบด้วย 1) เครื่องมือที่ใช้คัดกรองผู้เข้าร่วมการวิจัย 2) เครื่องมือที่ใช้ในการทดลอง 3) เครื่องมือที่ใช้ในการวัดตัวแปรตาม และ มีรายละเอียด ดังนี้

1. เครื่องมือที่ใช้คัดกรองผู้เข้าร่วมการวิจัย ได้แก่

1.1 แบบสอบถามข้อมูลส่วนบุคคล ประกอบด้วย เพศ อายุ ระดับการศึกษา เกรดเฉลี่ยรวม เทอมล่าสุด เกรดเฉลี่ยวิชาคณิตศาสตร์เทอมล่าสุด และประวัติการเจ็บป่วยทางกาย

1.2 แบบทดสอบมาตรฐานความฉลาดทางด้านสติปัญญา (IQ) Advance Progressive Matrices (APM) เป็นแบบทดสอบอวัจนภาษา จำนวน 36 ข้อ มีค่าความเที่ยงอยู่ที่ 0.80 ทดสอบเป็นรายบุคคล ใช้เวลาทดสอบประมาณ 15-40 นาที คิดคะแนนโดยรวมคะแนนตอบถูกต้องทั้งหมดและนำคะแนนดิบที่ได้มาเทียบกับตารางแปลผลคะแนน IQ โดยบุคคลปกติทั่วไป มีค่า IQ มาตรฐานอยู่ที่ 90 ถึง 110 (Raven, 2008)

2. เครื่องมือที่ใช้ในการทดลอง

2.1 เครื่องมือที่ใช้ในกลุ่มทดลอง คือ โปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้

2.2 เครื่องมือที่ใช้ในกลุ่มควบคุม คือ เกมคอมพิวเตอร์ซูโดกุ (Sudoku) เนื่องจากการศึกษา งานวิจัยพบว่า เกมดังกล่าวสามารถฝึกทักษะการแก้ปัญหาได้ (Pelánek, 2011) ผู้วิจัยจึงสนใจนำมาใช้ในกลุ่มควบคุมเพื่อเปรียบเทียบกับกลุ่มทดลองที่ใช้โปรแกรมคอมพิวเตอร์ที่ผู้วิจัยพัฒนาขึ้น และยังเป็นเกมที่มีการเผยแพร่ทั่วไปให้ใช้ฟรีอีกด้วย โดยเกมคอมพิวเตอร์ซูโดกุ คือเกมปริศนาตัวเลข ที่ผู้เล่นต้องเลือกใส่ หมายเลขตั้งแต่ เลข 1 ถึงเลข 9 โดยมีเงื่อนไขว่าในแต่ละแถวและแต่ละหลักตัวเลขต้องไม่ซ้ำกัน ตารางซูโดกุจะมี 9×9 ช่อง ซึ่งประกอบจากตารางย่อย 9 ตาราง ในลักษณะ 3×3 แบ่งแยกกันโดยเส้นหนา และในแต่ละตารางย่อยจะต้องมีตัวเลข 1 ถึง 9 เช่นเดียวกัน เมื่อเริ่มเกมจะมีตัวเลขบางส่วนให้มาเป็นคำใบ้ และผู้เล่นจะต้องใส่ทุกช่องที่เหลือให้ครบ โดยตามเงื่อนไขว่าแต่ละตัวเลขในแต่ละแถวห้ามซ้ำกัน (Jana et al., 2022)

3. เครื่องมือที่ใช้ในการวัดตัวแปรตาม

3.1 แบบทดสอบทักษะการแก้ปัญหาที่ซับซ้อนมาตรฐาน COMPRO: Complex Problem Solving Test ใช้สำหรับประเมินความสามารถในการแก้ปัญหาที่ซับซ้อน เหมาะสำหรับนักจิตวิทยาทั่วไป และ นักจิตวิทยาการศึกษา แบบทดสอบทักษะการแก้ปัญหาที่ซับซ้อนมาตรฐาน COMPRO มีค่าความเที่ยง (reliability) อยู่ที่ 0.92 แบบทดสอบประกอบไปด้วยการจำลองสถานการณ์การแก้ปัญหาที่ซับซ้อน จำนวน 11 ข้อ ใช้เวลาทำแบบทดสอบประมาณ 50 นาที (Greiff & Wüstenberg, 2015) ซึ่งแบบทดสอบแบ่งเป็นสองส่วน ได้แก่ ส่วนแรกคือการทดสอบการแสวงหาความรู้ (knowledge acquisition) และส่วนที่สองคือการทดสอบการประยุกต์ใช้ความรู้ (knowledge application) ซึ่งการคิดคะแนนพิจารณาได้ดังนี้ 1) คะแนนความสามารถ

ในการแก้ปัญหาที่ซับซ้อน พิจารณาจากผลรวมของคะแนนการแสวงหาความรู้และการประยุกต์ใช้ความรู้ มีคะแนนเต็ม 100 คะแนน 2) คะแนนการแสวงหาความรู้ พิจารณาจากผลรวมของคะแนนการตอบถูกที่ได้จาก ส่วนแรกของแบบทดสอบ มีคะแนนเต็ม 100 คะแนน 3) คะแนนการประยุกต์ใช้ความรู้ พิจารณาจากผลรวม ของคะแนนการตอบถูกที่ได้จากส่วนที่สองของแบบทดสอบ มีคะแนนเต็ม 100 คะแนน 4) ระยะเวลาที่ใช้ ในการทำแบบทดสอบ พิจารณาจากความเร็วที่ใช้ในการทำแบบทดสอบมีหน่วยเป็นนาที (Krieger et al., 2021)

3. การเก็บรวบรวมข้อมูล

การวิจัยนี้แบ่งวิธีดำเนินการทดลองออกเป็น 2 ระยะ ประกอบด้วย ระยะก่อนการทดลองและ ระยะการทดลอง โดยมีรายละเอียด ดังนี้

3.1 ระยะก่อนการทดลอง

3.1.1 จัดทำใบสมัครเข้าร่วมโครงการออนไลน์ด้วยโปรแกรม Google form โดยกำหนด รายละเอียดที่ต้องกรอกในสมัครได้แก่ ชื่อ นามสกุล เพศ อายุ ระดับการศึกษา เบอร์โทรศัพท์ เกรดเฉลี่ยรวม ภาคการศึกษาล่าสุด เกรดเฉลี่ยวิชาคณิตศาสตร์เทอมล่าสุด นอกจากนี้ผู้วิจัยได้ระบุวัตถุประสงค์ในการทำวิจัย และการพิทักษ์สิทธิ์กลุ่มตัวอย่างไว้ในใบสมัครด้วย และคัดเลือกกลุ่มตัวอย่าง โดยพิจารณาคุณสมบัติตามเกณฑ์ คัดเลือกเข้าและเกณฑ์คัดออก และได้กลุ่มตัวอย่างที่มีคุณสมบัติตามเกณฑ์ จำนวน 60 คน จากนั้นสุ่มตัวอย่าง เข้าสู่การทดลองโดยการสุ่มอย่างง่าย (random assignment) ได้กลุ่มทดลองและกลุ่มควบคุม กลุ่มละ 30 คน

3.1.2 ผู้วิจัยชี้แจงขั้นตอนการดำเนินการทั้งหมด การเตรียมตัวก่อนการทดลอง ระหว่าง การทดลอง และหลังการทดลอง ขอความร่วมมืองดเข้าร่วมกิจกรรมที่มีลักษณะใกล้เคียงกับการเพิ่มทักษะ การแก้ปัญหา และควบคุมความตรงภายใน (internal validity) และเรื่องการรั่วไหลของข้อมูลการทดลองของ กลุ่มทดลองทั้ง 2 กลุ่มไปยังกลุ่มทดลองอื่น โดยผู้วิจัยชี้แจงรายละเอียดให้ทราบเป็นรายบุคคลทางโทรศัพท์ พร้อมนัดหมายวันและเวลาในการเข้าร่วมการทดลองของแต่ละบุคคลเพื่อให้เข้ามาทำการทดลอง ณ ห้องปฏิบัติการ ศูนย์ความเป็นเลิศด้านวิทยาการปัญญา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

3.2 ระยะการทดลอง

3.2.1 ผู้วิจัยดำเนินการชี้แจงกระบวนการวิจัย จากนั้นให้ผู้ปกครองและผู้เข้าร่วมการทดลอง ลงนามในแบบฟอร์มยินยอมการเข้าร่วมโครงการวิจัย ทดสอบ IQ ของกลุ่มตัวอย่างเป็นรายบุคคล โดยใช้เวลา ประมาณ 45 นาที และทำกิจกรรมทดสอบทักษะการแก้ปัญหาที่ซับซ้อนก่อนการทดลอง (pre-test) ของกลุ่ม ทดลองและกลุ่มควบคุม ใช้เวลาประมาณคนละ 45 นาที

3.2.2 ดำเนินการฝึกโปรแกรมทักษะการแก้ปัญหาที่ซับซ้อนให้กับกลุ่มทดลองและกลุ่มควบคุม ดังนี้

3.2.2.1 กลุ่มทดลอง ผู้วิจัยให้ฝึกโปรแกรมเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่น โดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองให้การเรียนรู้ เนื่องจากระยะเวลาขณะทำการวิจัย เป็นช่วงที่มีการระบาดของไวรัสติดเชื้อโคโรนา (COVID-19) ผู้วิจัยจึงได้พัฒนาโปรแกรมให้เป็นแบบออนไลน์ เพื่อสะดวกในการใช้งานเป็นรายบุคคลและลดการรวมกลุ่ม โดยผู้วิจัยอธิบายวิธีใช้งานโปรแกรมและแจกคู่มือ การฝึกให้กับกลุ่มทดลองกลับไปทำการฝึกที่บ้าน โดยฝึกสัปดาห์ละ 2 ครั้ง ครั้งละ 45 นาที เป็นจำนวน 10 ครั้ง

3.2.2.2 กลุ่มควบคุม ผู้วิจัยให้ฝึกเล่นเกมฝึกสมองซูโดกุ เนื่องจากระยะเวลาขณะทำการวิจัยเป็นช่วงที่มีการระบาดของไวรัสติดเชื้อโคโรนา (COVID-19) ผู้วิจัยจึงให้ฝึกเล่นเกมเป็นแบบออนไลน์ เพื่อสะดวกในการใช้งานเป็นรายบุคคลและลดการรวมกลุ่ม โดยผู้วิจัยอธิบายวิธีการเล่นเกมและแจกคู่มือการฝึกให้กับกลุ่มทดลองกลับไปทำการฝึกที่บ้าน โดยฝึกสัปดาห์ละ 2 ครั้ง ครั้งละ 45 นาที เป็นจำนวน 10 ครั้ง

3.2.3 ดำเนินการทำการกิจกรรมทดสอบทักษะการแก้ปัญหาที่ซับซ้อนหลังการทดลอง (post-test) ของกลุ่มทดลองและกลุ่มควบคุม ใช้เวลาประมาณคนละ 45 นาที

4. การวิเคราะห์ข้อมูล

การวิจัยนี้ใช้วิธีการทางสถิติเพื่อการวิเคราะห์ข้อมูล การวิเคราะห์ข้อมูลพื้นฐานทั่วไป โดยใช้สถิติเชิงบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน วิเคราะห์คุณภาพเครื่องมือด้วยดัชนีความตรงเชิงเนื้อหา (Content Validity Index: CVI) วิเคราะห์เปรียบเทียบค่าเฉลี่ยคะแนนความฉลาดทางด้านสติปัญญา (IQ) เกรดเฉลี่ยวิชาคณิตศาสตร์และเกรดเฉลี่ยรวม ด้วยสถิติ ANOVA เปรียบเทียบคะแนนทักษะการแก้ปัญหาที่ซับซ้อนในกลุ่มทดลองระยะก่อนกับหลังการทดลองด้วยสถิติ pair – sample t-test และเปรียบเทียบค่าเฉลี่ยคะแนนทักษะการแก้ปัญหาที่ซับซ้อนระหว่างกลุ่มควบคุมกับกลุ่มทดลองระยะก่อนการทดลองด้วยสถิติ MANOVA และระยะหลังทดลองด้วยสถิติ MANCOVA เนื่องจากพบว่าตัวแปรคะแนนทักษะการแก้ปัญหาที่ซับซ้อนก่อนเข้าร่วมกิจกรรมเป็นตัวแปรแทรกซ้อนในการวิจัยนี้

ผลการวิจัย

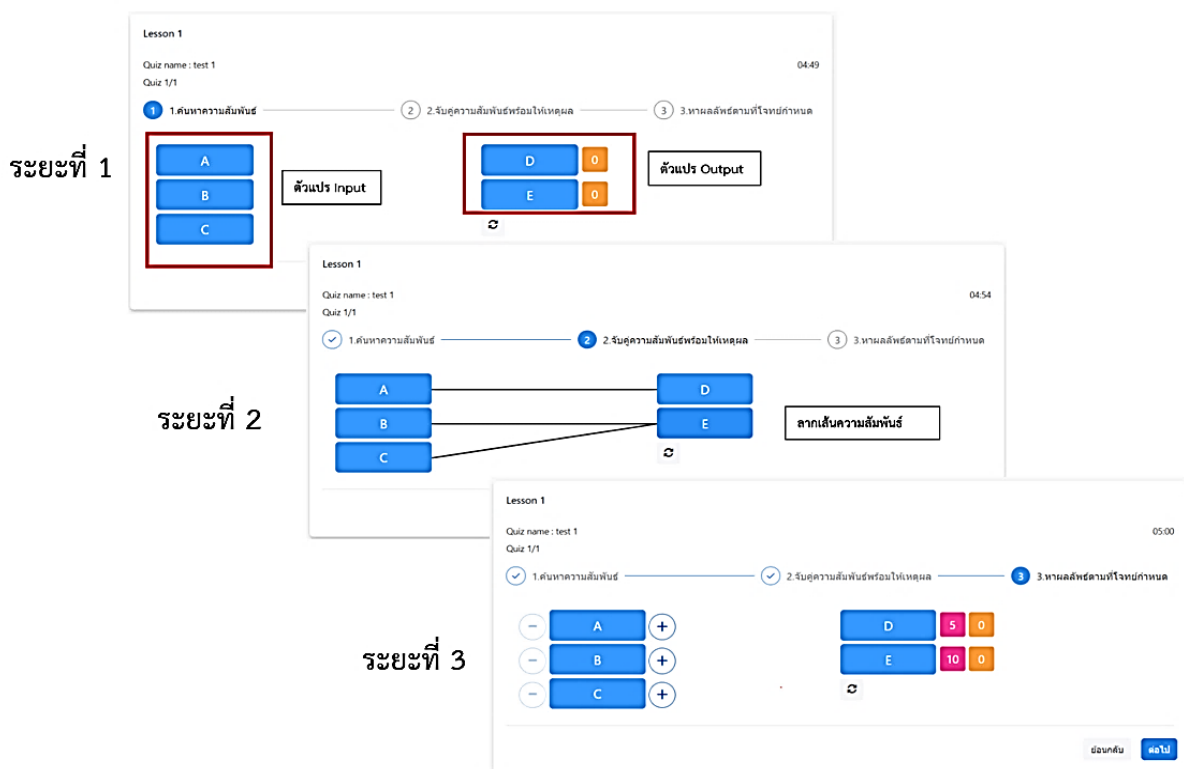
ผู้วิจัยได้ออกแบบโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนให้สอดคล้องกับทฤษฎีประมวลสารสนเทศและทฤษฎีการกำกับตนเองในการเรียนรู้ โดยการกำหนดสถานการณ์การแก้ปัญหาและใช้กระบวนการแก้ปัญหาที่แบ่งกิจกรรมออกเป็น 3 ระยะ คือ ระยะที่ 1 การวางแผนกำหนดเป้าหมายและการวางกลยุทธ์ ระยะที่ 2 การตรวจสอบประสิทธิภาพ การประเมินกลยุทธ์ การปรับปรุงกลยุทธ์ และระยะที่ 3 การประยุกต์ใช้ความรู้

ระยะที่ 1 การวางแผนกำหนดเป้าหมายและการวางกลยุทธ์ ซึ่งในโปรแกรมจะใช้ชื่อว่า “ค้นหาความสัมพันธ์” มีขั้นตอนดังนี้ 1) กดที่ปุ่มตัวอักษร A B หรือ C ซึ่งจะเป็นตัวแปรป้อนเข้า (Input) เมื่อกดแล้วจะทำให้ค่าของ D และ E เปลี่ยนไปซึ่งเป็นตัวแปรผลลัพธ์ (Output) โดยค่าที่เปลี่ยนไปจะแสดงเป็นตัวเลข ผู้เข้าร่วมการทดลองจะต้องค้นหาให้ได้ว่า เมื่อกดที่ตัวอักษร A B หรือ C แล้ว ค่าของตัวแปร D และ E จะมีค่าเพิ่มขึ้นหรือลดลงเท่าไร โดยให้ใช้เวลา 5 นาที 2) เมื่อผู้เข้าร่วมการทดลองทราบความสัมพันธ์ระหว่างตัวแปรป้อนเข้าและตัวแปรผลลัพธ์ และพอใจกับผลลัพธ์แล้ว ให้กดปุ่ม “ถัดไป”

ระยะที่ 2 การตรวจสอบประสิทธิภาพ การประเมินกลยุทธ์ การปรับปรุงกลยุทธ์ ซึ่งในโปรแกรมจะใช้ชื่อว่า “จับคู่ความสัมพันธ์พร้อมให้เหตุผล” มีขั้นตอนดังนี้ 1) ให้ผู้เข้าร่วมการทดลองหาความสัมพันธ์ระหว่างตัวแปรป้อนเข้าและตัวแปรผลลัพธ์ โดยให้ลากเส้น ระหว่างตัวแปรที่คิดว่ามีความสัมพันธ์กันเชิงบวกเท่านั้น เช่น กด A แล้ว D เพิ่มขึ้นแสดงว่า A และ D มีความสัมพันธ์กันเชิงบวก เป็นต้น วิธีการลากเส้นคือ ให้คลิกที่ตัวแปรฝั่งซ้าย 1 ครั้ง และ คลิกที่ตัวแปรฝั่งขวา 1 ครั้ง เพื่อกำหนดเส้นความสัมพันธ์ 2) เมื่อลากเส้น

ความสัมพันธ์ได้แล้วให้ผู้เข้าร่วมการทดลองกดที่ปุ่ม “วิธิตด” จากนั้นหน้าต่างให้ระบุวิธีคิดจะปรากฏขึ้นมาให้ผู้เข้ารับการทดลองพิมพ์ข้อความให้เหตุผลลงไปว่าทำไมถึงได้ลากเส้นความสัมพันธ์แบบนั้น มีวิธีคิดอย่างไรแล้วแต่ผู้ทดลองจะระบุ เช่น กด A แล้ว D ค่าเพิ่มขึ้น 1 แสดงว่ามีความสัมพันธ์กันเชิงบวก เมื่อให้เหตุผลเสร็จแล้วจากนั้นกดที่ปุ่ม “ตกลง” 3) โปรแกรมจะทำการตรวจคำตอบการลากเส้นความสัมพันธ์ระหว่างตัวแปรป้อนเข้าและตัวแปรผลลัพธ์ว่าถูกต้องหรือไม่ ถ้าถูกต้องโปรแกรมจะดำเนินการต่อไป แต่ถ้าผิดโปรแกรมจะให้กลับไปลากเส้นความสัมพันธ์ใหม่ พร้อมกับให้เหตุผลใหม่ โดยขั้นตอนในระยะที่ 2 นี้ทั้งหมดใช้เวลา 5 นาที

ระยะที่ 3 การประยุกต์ใช้ความรู้ ซึ่งในโปรแกรมจะใช้ชื่อว่า “หาผลลัพธ์ตามโจทย์ที่กำหนด” มีขั้นตอนดังนี้ 1) โปรแกรมจะกำหนดค่าเป้าหมายของตัวแปรผลลัพธ์มาให้ 2) ผู้เข้าร่วมการทดลองจะต้องปรับค่าของตัวแปรผลลัพธ์ ให้ได้ตามเป้าหมายที่โปรแกรมกำหนดมาให้ โดยการกดปุ่มเครื่องหมาย “+” หรือ “-” ของตัวแปรป้อนเข้า และต้องทำให้ได้ตามเวลาที่กำหนด ให้เวลาในขั้นตอนนี้ 5 นาที แสดงดังภาพ 2 ตรวจสอบความเหมาะสมของโปรแกรมด้วยดัชนีความตรงตามเนื้อหา (Content Validity Index: CVI) ได้ค่าเท่ากับ 0.90



ภาพ 2 โปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้

ผลการวิเคราะห์ข้อมูลพื้นฐานลักษณะทั่วไปของกลุ่มตัวอย่าง โดยใช้สถิติบรรยาย ผลปรากฏว่ากลุ่มตัวอย่างทั้งหมดจำนวน 60 คน เพศหญิง จำนวน 41 คน (ร้อยละ 68.30) เพศชาย จำนวน 19 คน (ร้อยละ 31.70) มีอายุอยู่ระหว่าง 13–17 ปี ระดับชั้นมัธยมศึกษาปีที่ 1-6 กลุ่มตัวอย่างทุกคนมีสุขภาพแข็งแรง

เนื่องจากงานวิจัยนี้ศึกษาทักษะการแก้ปัญหาที่ซับซ้อนซึ่งจากการทบทวนวรรณกรรม พบว่า ความฉลาดทางด้านสติปัญญา (IQ) เกรดเฉลี่ยวิชาคณิตศาสตร์ และ เกรดเฉลี่ยรวมของกลุ่มอย่าง ส่งผลต่อ ทักษะการแก้ปัญหาที่ซับซ้อน (Fischer et al., 2012) ผู้วิจัยจึงได้มีการวัดผลและนำมาวิเคราะห์เปรียบเทียบ เพื่อแสดงให้เห็นถึงความเท่าเทียมกันของกลุ่มทดลองและกลุ่มควบคุม ผลวิเคราะห์เปรียบเทียบคะแนน ความฉลาดทางด้านสติปัญญา (IQ) เกรดเฉลี่ยวิชาคณิตศาสตร์และเกรดเฉลี่ยรวม ด้วยสถิติ ANOVA ปรากฏว่า ความฉลาดทางด้านสติปัญญา ($F=2.79, p=.10$) เกรดเฉลี่ยวิชาคณิตศาสตร์ ($F=1.28, p=.26$) และ เกรดเฉลี่ยรวม ($F=0.69, p=.41$) ระหว่างกลุ่มทดลองและกลุ่มควบคุม ไม่มีความแตกต่างกัน

ผลการเปรียบเทียบทักษะการแก้ปัญหาที่ซับซ้อนระหว่างก่อนกับหลังการฝึกด้วยโปรแกรม คอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎี การกำกับตนเองในการเรียนรู้ในกลุ่มทดลองด้วยสถิติ pair – sample t-test ปรากฏว่าคะแนนทักษะ การแก้ปัญหาที่ซับซ้อน ($t=4.23, p=.00$) คะแนนการแสวงหาความรู้ ($t=2.42, p=.02$) คะแนนการประยุกต์ใช้ ความรู้ ($t=4.30, p=.00$) และของกลุ่มทดลองหลังการฝึกสูงกว่าก่อนการฝึกและมีระยะเวลาที่ใช้ในการทำ แบบทดสอบหลังการฝึกน้อยกว่าก่อนการฝึก ($t= 5.88, p=.00$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงดัง ตาราง 1

ตาราง 1 ผลการเปรียบเทียบทักษะการแก้ปัญหาที่ซับซ้อนระหว่างก่อนกับหลังการฝึกด้วยโปรแกรม คอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศ ร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้ในกลุ่มทดลอง

กลุ่มทดลอง (n=30)	ก่อนการฝึก		หลังการฝึก		df	t	p
	M	SD	M	SD			
คะแนนทักษะการแก้ปัญหาที่ซับซ้อน	72.56	15.58	80.63	13.62	29	4.23**	.00
คะแนนการแสวงหาความรู้	78.23	13.62	84.00	11.57	29	2.42*	.02
คะแนนการประยุกต์ใช้ความรู้	58.03	26.88	71.40	29.99	29	4.30**	.00
ระยะเวลาที่ใช้ในการทำแบบทดสอบ	27.06	3.82	21.00	4.05	29	5.88**	.00

* $p<.05$, ** $p<.01$

ผลการเปรียบเทียบทักษะการแก้ปัญหาที่ซับซ้อนก่อนการฝึกด้วยโปรแกรมคอมพิวเตอร์เพิ่มทักษะ การแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎีการกำกับตนเองในการ เรียนรู้ระหว่างกลุ่มทดลองกับกลุ่มควบคุม ผลการตรวจสอบข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวน ร่วมพหุแบบทางเดียว (MANOVA) โดยการทดสอบเมทริกซ์ความแปรปรวนร่วมของทักษะการแก้ปัญหา ที่ซับซ้อน ก่อนการฝึก ระหว่างกลุ่มทดลองและกลุ่มควบคุม ได้ค่าสถิติทดสอบ Box's M เท่ากับ 17.01 และ ค่าสถิติทดสอบ $F_{Box's M} = 1.57$ โดยมีค่าความน่าจะเป็นทางสถิติ $p=.10$ แสดงให้เห็นว่า เมทริกซ์ ความแปรปรวนร่วมของกลุ่มทดลองและกลุ่มควบคุมไม่แตกต่างกัน ผลการทดสอบความแปรปรวนของทักษะ

การแก้ปัญหาที่ซับซ้อนหลังการทดลองระหว่างกลุ่มทดลองและกลุ่มควบคุมจำแนกรายด้านด้วยสถิติ Levene's Test ปรากฏว่า คะแนนทักษะการแก้ปัญหาที่ซับซ้อน ($F_{Levene's} = 2.19, p = .14$) คะแนนการแสวงหาความรู้ ($F_{Levene's test} = .23, p = .63$) คะแนนการประยุกต์ใช้ความรู้ ($F_{Levene's} = 3.74, p = .05$) ระยะเวลาที่ใช้ในการทำแบบทดสอบ ($F_{Levene's test} = 1.59, p = .22$) มีความแปรปรวนของความคลาดเคลื่อนระหว่างกลุ่มไม่แตกต่างกัน

ผลการเปรียบเทียบทักษะการแก้ปัญหาที่ซับซ้อนก่อนการฝึกด้วยโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้ระหว่างกลุ่มทดลองกับกลุ่มควบคุม ปรากฏว่าคะแนนทักษะการแก้ปัญหาที่ซับซ้อน ($F = 2.80, p = .09$) คะแนนการประยุกต์ใช้ความรู้ ($F = 0.79, p = .37$) และระยะเวลาที่ใช้ในการทำแบบทดสอบ ($F = 1.92, p = .17$) ไม่แตกต่างกัน ในขณะที่คะแนนการแสวงหาความรู้ ($F = 5.32, p = .02$) ในกลุ่มทดลองสูงกว่ากลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ซึ่งจากความแตกต่างกันของคะแนนการแสวงหาความรู้ในระยะก่อนการทดลองที่เกิดขึ้น ผู้วิจัยจึงนำมาใช้เป็นตัวแปรร่วม (covariate) ในการวิเคราะห์เปรียบเทียบทักษะการแก้ปัญหาที่ซับซ้อนหลังการฝึกด้วยโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้ระหว่างกลุ่มทดลองกับกลุ่มควบคุมด้วยสถิติทดสอบ MANCOVA

ผลการตรวจสอบข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนร่วมพหุแบบทางเดียว (MANCOVA) โดยการทดสอบเมทริกซ์ความแปรปรวนร่วมของทักษะการแก้ปัญหาที่ซับซ้อนหลังการฝึกระหว่างกลุ่มทดลองและกลุ่มควบคุม ได้ค่าสถิติทดสอบ Box's $M = 52.42$ และค่าสถิติทดสอบ $F_{Box's M} = 4.85$ โดยมีค่าความน่าจะเป็นทางสถิติ $p = .00$ แสดงให้เห็นว่า เมทริกซ์ความแปรปรวนร่วมของ ระหว่างกลุ่มทดลองและกลุ่มควบคุมแตกต่างกัน แต่เนื่องจากงานวิจัยนี้ผู้วิจัยได้กำหนดการสุ่มเข้ากลุ่มทดลองและกลุ่มควบคุมโดยมีความเท่าเทียมกันตามเกณฑ์การคัดกรองที่ผู้วิจัยกำหนดขึ้น ทำให้สามารถวิเคราะห์ความแปรปรวนร่วมพหุแบบทางเดียวได้ (Hair et al., 2010)

ผลการทดสอบความแปรปรวนของทักษะการแก้ปัญหาที่ซับซ้อนหลังการฝึกระหว่างกลุ่มทดลองกับกลุ่มควบคุม จำแนกรายด้านด้วยสถิติ Levene's Test ปรากฏว่า คะแนนทักษะการแก้ปัญหาที่ซับซ้อน ($F_{Levene's test} = .93, p = .33$) คะแนนการแสวงหาความรู้ ($F_{Levene's test} = 2.13, p = .15$) คะแนนการประยุกต์ใช้ความรู้ ($F_{Levene's test} = 1.18, p = .28$) ระยะเวลาที่ใช้ในการทำแบบทดสอบ ($F_{Levene's test} = 3.11, p = .08$) มีความแปรปรวนของความคลาดเคลื่อนระหว่างกลุ่มไม่แตกต่างกัน

ผลการเปรียบเทียบทักษะการแก้ปัญหาที่ซับซ้อนหลังการใช้โปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้ ระหว่างกลุ่มทดลองกับกลุ่มควบคุม ($F_{Wilk's Lambda} = 6.26, p = .00$) ปรากฏว่า คะแนนทักษะการแก้ปัญหาที่ซับซ้อน ($F = 8.12, p = .00$) คะแนนการแสวงหาความรู้ ($F = 32.6, p = .00$) คะแนนการประยุกต์ใช้ความรู้ ($F = 8.03, p = .00$) ในกลุ่มทดลองสูงกว่ากลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติ ในขณะที่ระยะเวลาที่ใช้ในการทำแบบทดสอบ ($F = 2.64, p = .11$) ไม่แตกต่างกัน ดังตาราง 2

ตาราง 2 ผลการเปรียบเทียบทักษะการแก้ปัญหาที่ซับซ้อน หลังการฝึกด้วยโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้ ระหว่างกลุ่มทดลองกับกลุ่มควบคุม

ผลการเปรียบเทียบ	<i>n</i>	<i>M</i>	<i>SD</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η^2
คะแนนทักษะการแก้ปัญหาที่ซับซ้อน									
กลุ่มทดลอง	30	78.12	13.62	861.89	1	861.89	8.12**	.00	.12
กลุ่มควบคุม	30	70.20	12.76						
คะแนนการแสวงหาความรู้									
กลุ่มทดลอง	30	82.43	11.57	5602.42	1	5602.42	32.6**	.00	.27
กลุ่มควบคุม	30	62.23	16.09						
คะแนนการประยุกต์ใช้ความรู้									
กลุ่มทดลอง	30	69.42	21.99	3510.09	1	3510.09	8.03**	.00	.12
กลุ่มควบคุม	30	53.44	21.50						
ระยะเวลาที่ใช้ในการทำแบบทดสอบ									
กลุ่มทดลอง	30	21.23	4.05	31.93	1	31.93	2.64	.11	.04
กลุ่มควบคุม	30	22.76	2.93						

^a $F_{Wilk'Lambda}=6.26, p<.01$, ^bCovariate คือ คะแนนการแสวงหาความรู้ก่อนการทดลอง, ** $p<.01$

อภิปรายผล

โปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ สามารถนำไปใช้สำหรับเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นได้ เนื่องจากโปรแกรมถูกออกแบบตามแนวคิดของ Frensch & Funke (1995) โดยการให้สถานการณ์ (situation) และต้องเผชิญกับอุปสรรค (barriers) ที่มีอยู่ระหว่างการแก้ปัญหาที่กำหนด เพื่อไปถึงเป้าหมาย (goals) โดยอุปสรรคจะมีความซับซ้อนเปลี่ยนแปลงอยู่ตลอดเวลา โดยต้องอยู่ในสภาพแวดล้อม (environment) ที่มีการรบกวนอื่นๆ ผู้แก้ปัญหามustใช้เครื่องมือ (tools) เพื่อทำให้บรรลุเป้าหมายที่กำหนด และโปรแกรมที่ผู้วิจัยพัฒนาขึ้นใช้การประยุกต์แนวคิดทฤษฎีประมวลสารสนเทศของ Klausmeier (1985) ร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้ของ Zimmerman (2002) โดยแบ่งกิจกรรมออกเป็น 3 ระยะ คือ ระยะที่ 1 การวางแผนกำหนดเป้าหมายและการวางกลยุทธ์ ระยะที่ 2 การตรวจสอบประสิทธิภาพ การประเมินกลยุทธ์ การปรับปรุงกลยุทธ์ และระยะที่ 3 การประยุกต์ใช้ความรู้ในการแก้ปัญหา โดยระยะที่ 1 การวางแผนกำหนดเป้าหมายและการวางกลยุทธ์ มีความสอดคล้องกับแนวคิดทฤษฎีประมวลสารสนเทศของ Klausmeier (1985) ที่เกี่ยวกับการรับข้อมูล การเข้ารหัสความจำจากการสัมผัส (sensory memory) ทำให้เกิดการเรียนรู้ปัญหาที่เกิดขึ้น และสอดคล้องกับทฤษฎีการกำกับตนเองในการเรียนรู้ที่ Zimmerman (2002) กล่าวว่า การที่จะแก้ปัญหาได้ดีต้องเริ่มจากการวางแผนกำหนดเป้าหมายและวางกลยุทธ์ การทำแผนก่อนเริ่ม

ทำงานจะช่วยให้วางแผนได้ตั้งแต่เริ่มต้น กระตุ้นให้เกิดความพยายามและแรงจูงใจในตนเอง (self-motivation) ในระยะที่ 2 การตรวจสอบประสิทธิภาพ การประเมินกลยุทธ์ การปรับปรุงกลยุทธ์ สอดคล้องกับทฤษฎีประมวลสารสนเทศของ Klausmeier (1985) ที่เกี่ยวข้องกับความจำระยะสั้น (short-term memory) และความจำขณะทำงาน (working memory) โดยใช้ความจำที่ถูกเข้ารหัสไว้ในระยะแรก และเรียกคืนมาใช้ในการทบทวนปัญหา เพื่อประเมินกลยุทธ์ที่ใช้ในการแก้ปัญหา และสอดคล้องกับทฤษฎีการกำกับตนเองในการเรียนรู้ของ Zimmerman (2002) ที่เกี่ยวกับการใช้กลยุทธ์และตรวจสอบประสิทธิภาพ ในขั้นตอนนี้การทำตามแผนที่ระบุไว้ในขั้นตอนการวางแผน ทำให้เกิดความมั่นใจ เกิดการควบคุมตนเอง (self-control) ให้เป็นไปตามแผนกลยุทธ์ที่กำหนด ในระยะที่ 3 การประยุกต์ใช้ความรู้ในการแก้ปัญหา สอดคล้องกับทฤษฎีประมวลสารสนเทศของ Klausmeier (1985) ที่เกี่ยวข้องกับการดึงข้อมูล ความจำระยะสั้น (short-term Memory) ความจำขณะทำงาน (working memory) ความจำระยะยาว (long-term Memory) ที่ถูกบันทึกไว้ในกิจกรรมระยะที่ 1 และ 2 มาใช้ในการแก้ปัญหา และสอดคล้องกับทฤษฎีการกำกับตนเองในการเรียนรู้ของ Zimmerman (2002) ที่เกี่ยวข้องกับการสะท้อนประสิทธิภาพ ประเมินประสิทธิภาพของตนเองและผลลัพธ์ที่ได้ สะท้อนประสิทธิผลของกลยุทธ์ที่ใช้ เช่น เลือกกลยุทธ์ที่เหมาะสมหรือไม่ ปฏิบัติตามด้วยกลยุทธ์ที่เลือกไว้หรือไม่ ช่วยให้เกิดการปรับปรุงกลยุทธ์วิธีการแก้ปัญหาในครั้งต่อไปได้

นอกจากนี้โปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ ถูกพัฒนาให้มีลักษณะให้เป็นสถานการณ์การแก้ปัญหาที่มีการเปลี่ยนแปลงไปตามช่วงเวลาหรือเรียกได้ว่าเป็นสถานการณ์การแก้ปัญหาแบบพลวัตและมีแนวทางการพัฒนาจากโปรแกรม MicroDYN สอดคล้องกับของวิจัยของ Schoppek & Fischer (2017) ที่พบว่า การกำหนดสถานการณ์การแก้ปัญหาที่ซับซ้อนแบบพลวัตและมีตัวแปรที่หลากหลายและประยุกต์ใช้โปรแกรม MicroDYN เป็นองค์ประกอบที่สำคัญในการฝึกทักษะการแก้ปัญหาที่ซับซ้อนให้มีประสิทธิภาพ และสอดคล้องกับงานของ Kretzschmar & Süß (2015) พบว่าการสร้างสถานการณ์จำลองการแก้ปัญหาด้วยโปรแกรมคอมพิวเตอร์มีผลต่อกระบวนการของการแก้ปัญหาที่ซับซ้อนโดยเฉพาะการแสวงหาความรู้ (knowledge acquisition) และการประยุกต์ใช้ความรู้ (knowledge application)

ผลการใช้โปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้ ปรากฏว่าการฝึกด้วยโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้สัปดาห์ละครั้ง 2 ครั้ง ครั้งละ 45 นาที เป็นจำนวน 10 ครั้ง ส่งผลให้กลุ่มที่ได้รับการฝึกมีค่าเฉลี่ยคะแนนทักษะการแก้ปัญหาที่ซับซ้อน ค่าเฉลี่ยคะแนนการแสวงหาความรู้ ค่าเฉลี่ยคะแนนการประยุกต์ใช้ความรู้ หลังการฝึกสูงกว่าก่อนได้รับการฝึก และมีระยะเวลาที่ใช้ในการทำแบบทดสอบหลังการฝึกน้อยกว่าก่อนการฝึก เมื่อพิจารณาระยะหลังการทดลองระหว่างกลุ่มที่ได้รับการฝึกด้วยโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับทฤษฎีการกำกับตนเองในการเรียนรู้ และกลุ่มที่เล่นเกมคอมพิวเตอร์ชุดๆ พบว่า กลุ่มที่ได้รับการฝึก มีค่าเฉลี่ยคะแนนทักษะการแก้ปัญหาที่ซับซ้อน ค่าเฉลี่ยคะแนนการแสวงหาความรู้ และค่าเฉลี่ยคะแนนการประยุกต์ใช้ความรู้ สูงกว่า

กลุ่มที่เล่นเกมซูโดกุ จากผลเชิงพหุติกรรมดังกล่าวชี้ให้เห็นว่า ผู้ที่ได้รับการฝึกด้วยโปรแกรมที่พัฒนาขึ้นนั้นได้รับการฝึกด้วยกระบวนการที่มีความซับซ้อนในสถานการณ์การแก้ปัญหาที่ไม่ตายตัวหรือมีการเปลี่ยนแปลงพลวัตสอดคล้องกับการศึกษาของ Schoppek & Fischer (2017) ที่กล่าวว่ากำหนัดสถานการณ์การแก้ปัญหาที่ซับซ้อนแบบพลวัตและมีตัวแปรที่หลากหลายเป็นองค์ประกอบที่สำคัญในการฝึกทักษะการแก้ปัญหาที่ซับซ้อนให้มีประสิทธิภาพ และการฝึกด้วยโปรแกรมที่ผู้วิจัยพัฒนาขึ้นใช้ทฤษฎีประมวลสารสนเทศ (Klausmeier, 1985) ซึ่งอธิบายกระบวนการทางปัญญาขั้นสูงและทฤษฎีการกำกับตนเองในการเรียนรู้ (Zimmerman, 2002) ที่ใช้กระบวนการทางจิตวิทยาในการกำกับตนเองมาช่วยเพิ่มประสิทธิภาพสำหรับการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อน สอดคล้องกับ Dörner & Funke (2017) ได้กล่าวว่า การแก้ไขปัญหที่ซับซ้อนเน้นถึงการให้สถานการณ์ความเป็นจริงที่มีความซับซ้อนสูงนอกจากต้องใช้ความสามารถทางปัญญาแล้ว จำเป็นต้องมีการรวบรวมกระบวนการทางจิตวิทยาที่ใช้ในการกำกับตนเอง (self-regulated) สอดคล้องกับศึกษาวิจัยของ ElAdl & Polpol (2020) ที่ศึกษาผลของการใช้กลยุทธ์การกำกับตนเองในการเรียนรู้ต่อทักษะการแก้ปัญหา พบว่าเด็กนักเรียนที่ได้รับการฝึกโดยใช้กลยุทธ์การกำกับตนเองในการเรียนรู้เป็นเวลา 12 สัปดาห์ มีทักษะการแก้ปัญหา มากกว่ากลุ่มที่ไม่ได้รับการฝึก และสอดคล้องกับงานวิจัยของ Süß & Kretschmar (2018) ได้ศึกษาทักษะการแก้ปัญหาที่ซับซ้อนของกลุ่มนักเรียน 2 กลุ่ม กลุ่มแรกมีนักเรียน 137 คน ได้ใช้การฝึกด้วยโปรแกรม Tailorshop ซึ่งเป็นการฝึกที่เน้นกระบวนการแก้ปัญหาโดยการให้เหตุผลเชิงตัวเลข (numerical reasoning) และกลุ่มที่สองจำนวน 152 คน ฝึกด้วยโปรแกรมที่เน้นการแก้ปัญหาในชีวิตจริงเป็นการจำลองสถานการณ์การแก้ปัญหาที่ซับซ้อนแบบพลวัต ใช้กระบวนการทางปัญญา (cognitive processes) ร่วมกับกระบวนการทางจิตวิทยา ผลการวิจัยสรุปว่า นักเรียนในกลุ่มที่สองมีความสามารถในการแก้ปัญหาที่ซับซ้อนดีกว่ากลุ่มแรก เนื่องจากเนื่องการฝึกโดยใช้ความสามารถทางปัญญาและกระบวนการทางจิตวิทยาที่ส่งผลต่อสภาพแวดล้อมความเป็นจริงมีผลต่อประสิทธิภาพของความสามารถในการแก้ปัญหาที่ซับซ้อน

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 จากผลการวิจัยแสดงให้เห็นว่าหลักการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่น โดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ สามารถนำมาพัฒนาเป็นโปรแกรมเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นได้ โรงเรียน สถาบันการศึกษา หรือผู้ที่เกี่ยวข้องกับการพัฒนาศักยภาพมนุษย์ในช่วงวัยรุ่น สามารถนำหลักการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ ไปประยุกต์ใช้สำหรับสร้างกิจกรรมหรือโปรแกรมเพื่อเป็นทางเลือกสำหรับการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่น

1.2 นักวิจัยหรือผู้ที่สนใจสามารถนำหลักการหรือโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ ไปใช้สำหรับพัฒนาความสามารถทางปัญญาหรือความสามารถในการแก้ปัญหาในรูปแบบอื่น ๆ ได้ เช่น การเรียน การตัดสินใจ การแก้ปัญหาเชิงสร้างสรรค์ เป็นต้น

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 การเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนในการศึกษาวิจัยครั้งนี้เป็นการประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ ซึ่งการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนมีทฤษฎีที่เกี่ยวข้องกับการเพิ่มความสามารถทางปัญญาที่สามารถนำมาประยุกต์ใช้ได้อีกหลากหลายทฤษฎี หรือเพิ่มตัวแปรที่เกี่ยวข้องกับการแก้ปัญหาที่ซับซ้อนให้มากขึ้น เพื่อพัฒนาการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนได้ดีขึ้นได้อีกในอนาคตตลอดจนพัฒนาให้สามารถเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนได้ในทุกช่วงวัย

2. การศึกษาครั้งนี้ เป็นการศึกษาเฉพาะกลุ่มวัยรุ่นที่มีอายุอยู่ระหว่าง 13 – 17 ปี เท่านั้น จึงควรมีการศึกษาผลของโปรแกรมคอมพิวเตอร์เพิ่มทักษะการแก้ปัญหาที่ซับซ้อนที่ผู้วิจัยพัฒนามาจากหลักการเพิ่มทักษะการแก้ปัญหาที่ซับซ้อนของวัยรุ่นโดยประยุกต์ทฤษฎีประมวลสารสนเทศร่วมกับการกำกับตนเองในการเรียนรู้ ในประชากรกลุ่มอื่น ๆ เพื่อตรวจสอบศักยภาพของโปรแกรม

References

- Dörner, D., & Funke, J. (2017). Complex problem solving: what it is and what it is not. *Frontiers in psychology*, 8, 1153.
- ElAdl, A. M., & Polpol, Y. S. (2020). The Effect of Self-Regulated Learning Strategies on Developing Creative Problem Solving and Academic Self-Efficacy among Intellectually Superior High School Students. *International Journal of Psycho-Educational Sciences*, 9(1), 97-106.
- Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *Journal of Problem Solving*, 4(1), 19 – 42.
- Frensch, P. A., & Funke, J. (Eds.). (1995). *Complex Problem Solving: The European Perspective*. Erlbaum.
- Griffin, P., & Care, E. (Eds.). (2014). *Assessment and teaching of 21st century skills: Methods and approach*. Springer.
- Greiff, S., & Wüstenberg, S. (2015). *Komplexer Problemlösetest COMPRO [Complex problem-solving test COMPRO]*. Schuhfried.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*. (7th Ed.). Pearson.
- Jana, S., Dey, A., Maji, A. K., & Pal, R. K. (2022). Solving Sudoku Using Neighbourhood-Based Mutation Approach of Genetic Algorithm. In *Advanced Computing and Systems for Security*. 13, 153-167.
- Klausmeier, H. J. (1985). *Educational psychology*. Harpercollins College Division.

- Kretzschmar, A., & Süß, H. M. (2015). A study on the training of complex problem solving competence. *Journal of Dynamic Decision Making*, 1, 4-4.
- Krieger, F., Stadler, M., Bühner, M., Fischer, F., & Greiff, S. (2021). Assessing complex problem-solving skills in under 20 minutes. *Psychological Test Adaptation and Development*. <http://dx.doi.org/10.1027/2698-1866/a000009>.
- Mainzer, K. (2009). Challenges of Complexity in the 21st Century. *Advanced Science Letters*, 1(1), 145-150.
- Pelánek, R. (2011). *Human problem solving: Sudoku case study*. Technical Report FIMURS-2011-01, Masaryk University Brno.
- Raven, J. (2008). The Raven progressive matrices tests: their theoretical basis and measurement model. *Uses and abuses of Intelligence. Studies advancing Spearman and Raven's quest for non-arbitrary metrics*, 17-68.
- Schoppek, W., & Fischer, A. (2017). Common process demands of two complex dynamic control tasks: transfer is mediated by comprehensive strategies. *Frontiers in psychology*, 8, 2145.
- Süß, H. M., & Kretzschmar, A. (2018). Impact of cognitive abilities and prior knowledge on complex problem solving performance—empirical results and a plea for ecologically valid microworlds. *Frontiers in psychology*, 9, 626.
- World Economic Forum. (2016). *Global Risks 2016: Insight Report* (11th Ed.). World Economic Forum.
- World Economic Forum. (2018). *The Future of Jobs Report 2018*. World Economic Forum.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64-70.

การออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์
โดยการให้ข้อมูลป้อนกลับแบบอัตโนมัติผ่านการเรียนรู้ของเครื่อง
Designing Methods to Diagnose Scientific Misconceptions
by using Automated feedback through Machine Learning

ศรัณย์ธร อินทศิริ^{1*} พัทธี จันท์เพ็ง² ชัยวัฒน์ ทะวะรุ่งเรือง³ และ ธนพงศ์ อินทะระ⁴

Sarunthorn Inthasiri^{1*}, Putharee Junpeng², Chaiwat Tawarungruang³ and Thanapong Intharah⁴

(Received: March 31, 2022; Revised: May 23, 2022; Accepted: May 27, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมาย (1) เพื่อวิเคราะห์ผลการตอบของผู้เรียนในระดับแผนที่โครงสร้างมโนทัศน์ที่คลาดเคลื่อน การออกแบบตัวเลือกแบบเลือกตอบ การกำหนดจุดตัด และข้อมูลพื้นฐานที่ใช้ในการทำนายระดับความสามารถของผู้เรียน และ (2) เพื่อออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ของผู้เรียน จากการให้ข้อมูลป้อนกลับอัตโนมัติผ่านการเรียนรู้ของเครื่อง โดยใช้การวิจัยการออกแบบ กลุ่มผู้สอบ คือผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 4 จำนวน 713 คน ใช้แบบทดสอบวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์แบบพหุมิติ ประกอบด้วยมิติความรู้และมิติการให้เหตุผล โดยใช้โมเดล MRCML วิเคราะห์ผลการตอบของผู้เรียน และออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์

ผลการวิจัยปรากฏ ดังนี้

1. ผลการวิเคราะห์ผลการตอบของผู้เรียน พบว่า (1) ระดับความสามารถของข้อสอบตามแผนที่โครงสร้างสามารถวัดระดับความสามารถของผู้เรียนได้ตามระดับแผนที่โครงสร้าง และมีบางข้อที่ผู้เรียนมีความสามารถเกินความสามารถรายข้อที่กำหนดไว้ (2) ผู้เรียนส่วนมากมีรูปแบบมโนทัศน์ที่คลาดเคลื่อนรายข้อ

¹ นักศึกษาปริญญาโท สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² รองศาสตราจารย์ สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

³ นักวิจัย ศูนย์จัดการข้อมูลและวิเคราะห์ทางสถิติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

⁴ อาจารย์ ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

งานวิจัยนี้ได้รับทุนอรรถริยาพนกวิจัยรุ่นกลาง ประจำปี 2565 จากสำนักงานวิจัยแห่งชาติ (วช.) รหัสโครงการวิจัย N42A650236

¹ Master's degree student, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

² Associate Professor, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

³ Researcher, Data Management and Statistical Analysis Center, Faculty of Public Health, Khon Kaen University

⁴ Lecturer, Department of Statistics, Faculty of Science, Khon Kaen University

The research received a research grant for Talented Mid-Career Researchers 2022 from the National Research Council of Thailand (NRCT) Project No. N42A650236

* Corresponding Author E-mail: Sarunthorni@kkumail.com

ในรูปแบบความเข้าใจผิดเกี่ยวกับมโนทัศน์และภาษา และทั้งฉบับในมิติด้านความรู้ในระดับความไม่เข้าใจ และ มิติการให้เหตุผลในระดับที่ถูกต้องแต่ไม่สมบูรณ์ (3) ออกแบบตัวเลือกแบบเลือกตอบจากคำตอบของผู้เรียน โดยตัวเลือกมีคะแนนหลายค่าตามระดับความสามารถของตัวเลือก (4) กำหนดจุดตัดโดยการกำหนดพื้นที่บน Wright map ในมิติความรู้ได้ 5 ระดับ 4 จุดตัดที่ -0.50 0.50 0.42 และ 1.19 ตามลำดับ และมิติการให้เหตุผลได้ 4 ระดับ 3 จุดตัดที่ -0.19 2.29 และ 5.49 ตามลำดับ และ (5) ข้อมูลพื้นฐานที่สามารถทำนายระดับความสามารถของผู้เรียนได้ คือ เกรดเฉลี่ยรวม เกรดเฉลี่ยวิชาวิทยาศาสตร์ และจำนวนชั่วโมงที่ศึกษาวิชาวิทยาศาสตร์

2. ผลการออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ โดยการให้ข้อมูลป้อนกลับ อัตโนมัตินผ่านการเรียนรู้ของเครื่องประกอบด้วย 4 ส่วน คือ (1) ส่วนรับข้อมูล (2) ส่วนประมวลผล (3) การให้ข้อมูลป้อนกลับอัตโนมัติ และ (4) ส่วนรายงานผลการประเมิน

คำสำคัญ: การวินิจฉัย ความเข้าใจที่คลาดเคลื่อน การป้อนกลับอัตโนมัติ การเรียนรู้ของเครื่อง
โมเดลพหุมิติสัมประสิทธิ์การสุ่มแบบโลจิท

Abstract

This research had the objectives to: (1) analyze student response patterns regarding construct map, misconceptions, objective test transformation, cut scores, and the students' baseline data, to predict their level; and (2) design diagnostic methods for scientific misconceptions using the automated feedback system through machine learning. The research was conducted under design research model. The participants comprised 713 grade 10 students as respondents. The instrument was a diagnosing multidimensional scientific misconceptions test that included knowledge dimension and rational thinking dimension. The analysis of the student response patterns was based on the multidimensional model under MRCML model. Then, the diagnostic methods for scientific misconceptions were developed.

The results were as follows:

1. On the respondent patterns analysis, it was found that (1) the test was capable of measuring the student's ability according to the construct map, accompanied with superior level evaluation in some particular segments; (2) most students had their misconceptions on each item in conceptual misunderstanding and vernacular misconceptions. Their misconceptions on the whole test in the knowledge dimension were at the no understanding level (NU) and in the rational thinking dimension were at the partial understanding (PU) level. (3) Objective test transformation derived from the student responses contained multiple choices with varied scores elicited on individual choice capability. (4) The setting cut scores on Wright map in the knowledge dimension were divided into five levels and four cut scores, at -0.50, 0.50, 0.42 and

1.19 respectively. The rational thinking dimension featured four levels with three cut scores at -0.19, 2.29 and 5.49 respectively. And (5) the students' baseline data to predict the level of the students' ability included the overall GPA, science GPA and hours per week of self-study in science.

2. The design of the diagnostic methods for scientific misconceptions using the automated feedback system through machine learning was developed on the basis of 4 elements: 1. input, 2. processing, 3. automated feedback, and 4. assessment reporting.

Keywords: diagnosis, misconceptions, automated feedback, machine Learning, MRCML

บทนำ

มโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์เป็นอุปสรรคในการเรียนวิทยาศาสตร์ของผู้เรียน โดยจำกัดการพัฒนาความคิดที่ถูกต้องซึ่งเป็นพื้นฐานที่สำคัญสำหรับการเรียนรู้ชั้นสูง (National Research Council, 1997) ปัจจัยบางประการที่ทำให้นักเรียนมีมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ เช่น ประสบการณ์ในชีวิตประจำวัน หนังสือเรียน ครูผู้สอน และภาษาที่ใช้ หรืออาจเกิดจากลักษณะของมโนทัศน์ที่มีความเป็นนามธรรมและมีความซับซ้อน (Ling, 2017; Widiyatmoko & Shimizu, 2018) เช่น มโนทัศน์ในเรื่อง พันธุศาสตร์ เป็นต้น นักวิชาการหลายกลุ่มจึงมีความพยายามพัฒนามโนทัศน์ที่คลาดเคลื่อนของผู้เรียนด้วยวิธีการที่หลากหลาย เช่น การสร้างเครื่องมือเพื่อวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน (Kalas et al., 2013) การนำผลการประเมินที่ได้มาออกแบบกิจกรรมการเรียนรู้ (Thakho & Termtachatipongsa, 2015) เป็นต้น การวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนจึงจำเป็นอย่างยิ่งที่เป็นแนวทางสำคัญในการพัฒนาผู้เรียน ให้มีมโนทัศน์ที่ถูกต้องและสมบูรณ์มากยิ่งขึ้นต่อไป

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องพบว่า วิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนซึ่งเป็นหัวใจสำคัญนั้นยังไม่มีที่แพร่หลายในทางปฏิบัติมากนัก ส่วนมากมุ่งเน้นการพัฒนาเครื่องมือและพิจารณาความถี่ของมโนทัศน์ที่คลาดเคลื่อน (Gurel et al., 2015; Soeharto et al., 2019) ประกอบกับยังไม่มี การป้อนกลับผลการตอบของผู้เรียนอัตโนมัติอย่างทันที แต่มีความพยายามในการประเมินผู้เรียนจากแบบทดสอบ เพื่อลดความเข้าใจมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน เช่น มีการพัฒนาเครื่องมือดิจิทัลเพื่อวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ เรื่อง พันธุศาสตร์ ร่วมกับการใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ มาวินิจฉัยผลการตอบและรายงานผลผู้เรียนรายบุคคลผ่านระบบ web application แต่ยังพบข้อจำกัดที่ต้องตรวจแบบทดสอบด้วยตนเอง ไม่สามารถรายงานผลได้ทันที และจำเป็นต้องใช้กับกลุ่มผู้เรียนขนาดใหญ่เพื่อความถูกต้องและคงเส้นคงวา (Kantahan et al., 2020; Kantahan, 2020) ซึ่งหากสามารถป้อนกลับข้อมูลสู่ผู้เรียนได้ทันที จะทำให้ผู้เรียนสามารถจดจำได้ดียิ่งขึ้น แต่ในทางตรงกันข้ามหากกระงับการให้ข้อมูลป้อนกลับ จะมีการจดจำในเรื่องที่เรียนลดลงและเกิดการจดจำในสิ่งที่ผิดพลาด (Zhang et al., 2016)

การออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน จำเป็นต้องอาศัยเทคโนโลยีทางการศึกษาเข้ามามีส่วนร่วมในการเก็บรวบรวมข้อมูลและวินิจฉัยอย่างอัตโนมัติ ซึ่งปัจจุบันการประเมินประเภทต่าง ๆ มีการนำ การเรียนรู้ของเครื่อง (machine learning) เข้ามามีบทบาทมากขึ้น เนื่องจากเป็นอัลกอริธึมของคอมพิวเตอร์

ที่เรียนรู้และปรับปรุงได้อัตโนมัติผ่านประสบการณ์และการใช้งานของข้อมูล (Mitchell, 1997) เช่น มีการใช้การเรียนรู้ของเครื่องในการจัดประเภทข้อความเพื่อทำนายผลการเรียนของนักเรียนและระบุนักเรียนที่เสี่ยงต่อการไม่จบหลักสูตร (Wu et al., 2018) ในการวิจัยครั้งนี้คณะผู้วิจัยสนใจศึกษาการเรียนรู้แบบมีผู้สอน (supervised learning) (Mohammed et al., 2016) โดยอัลกอริทึมจำเป็นต้องใช้ข้อมูลชุดเรียนรู้ (training data) ซึ่งได้จากชุดข้อมูล (data Set) และส่วนที่รับกลับมาจากผู้เรียน (feedback) เพื่อปรับปรุงและเรียนรู้ความสัมพันธ์ระหว่างข้อมูลนำเข้ากับข้อมูลที่ส่งออกในการสร้างสมการการทำนายเพื่อทำนายระดับความสามารถของผู้เรียนจากข้อมูลพื้นฐาน

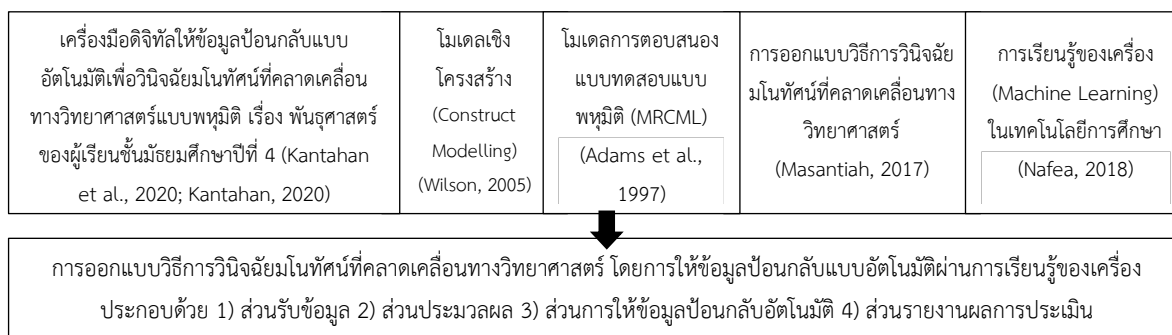
ดังนั้น ผู้วิจัยจึงสนใจที่จะออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ โดยการให้ข้อมูลป้อนกลับแบบอัตโนมัติผ่านการเรียนรู้ของเครื่อง โดยใช้ข้อมูลทฤษฎีภูมิจากงานวิจัย เรื่อง การพัฒนาเครื่องมือดิจิทัลเพื่อวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ในเนื้อหาพันธุศาสตร์ (Kantahan et al., 2020; Kantahan, 2020) ซึ่งเครื่องมือที่ใช้เป็นแบบทดสอบแบบพหุมิติที่มีการวัดความสามารถของผู้เรียนในมิติความรู้และมิติการให้เหตุผลทางวิทยาศาสตร์จำแนกตามแผนที่โครงสร้างในแต่ละมิติ ด้วยเป็นแบบทดสอบแบบพหุมิติจึงจำเป็นต้องใช้โดยใช้โมเดลพหุมิติสัมประสิทธิ์การสุ่มแบบโลจิท (Multidimensional Random Coefficients Multinomial Logit; MRCML) (Adams et al., 1997; Junpeng, 2018) ที่เป็นโมเดลที่ช่วยวินิจฉัยมโนทัศน์ กำหนดจุดตัด และกำหนดพื้นที่ของแต่ละระดับความสามารถบน Wright Map เพื่อสร้างเป็นนวัตกรรมต้นแบบในการประเมินผลและป้อนกลับข้อมูลอัตโนมัติได้อย่างทันที สำหรับผู้เรียน ครู ผู้ปกครองตลอดจนบุคลากรทางการศึกษา สะท้อนถึงสิ่งที่คุณครูรู้ ระดับมโนทัศน์ที่คลาดเคลื่อนและสิ่งที่คุณครูควรพัฒนาได้อย่างชัดเจน

ความมุ่งหมายของการวิจัย

1. เพื่อวิเคราะห์ผลการตอบของผู้เรียนในระดับแผนที่โครงสร้าง มโนทัศน์ที่คลาดเคลื่อน การออกแบบตัวเลือกแบบเลือกตอบ การกำหนดจุดตัด และข้อมูลพื้นฐานที่ใช้ในการทำนายระดับความสามารถของผู้เรียน
2. เพื่อออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ของผู้เรียน โดยการให้ข้อมูลป้อนกลับผ่านการเรียนรู้ของเครื่อง

กรอบแนวคิดการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง สามารถกำหนดกรอบแนวคิดการวิจัยได้ดังภาพ 1



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ใช้ระเบียบวิธีวิจัยการออกแบบ (design research) (Wongwanich, 2017) โดยแบ่งเป็น 4 ระยะ คือ (1) การวิเคราะห์ผลการตอบของผู้เรียนเพื่อนำไปสู่การออกแบบนวัตกรรมต้นแบบ (2) การออกแบบวิธีการวินิจฉัยโมทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ โดยการให้ข้อมูลป้อนกลับผ่านการเรียนรู้ของเครื่อง (3) การนำนวัตกรรมต้นแบบไปทดลองใช้เพื่อนำข้อมูลที่ได้มาวิเคราะห์ ปรับปรุง พัฒนา และรายงานคุณภาพของนวัตกรรม และ (4) เสนอหลักการออกแบบใหม่ในการพัฒนานวัตกรรมต้นแบบ โดยการนำเสนอในบทความครั้งนี้ เป็นการศึกษาในระยะที่ 1 และระยะที่ 2 ดังนี้

ระยะที่ 1 วิเคราะห์ผลการตอบของผู้เรียนเพื่อนำไปสู่การออกแบบนวัตกรรมต้นแบบ

ระยะนี้มุ่งเน้นการวิเคราะห์ผลการตอบของผู้เรียนซึ่งเป็นข้อมูลทุติยภูมิ เพื่อนำไปสู่การออกแบบนวัตกรรมต้นแบบ มีรายละเอียดดังนี้

1. ประชากรและตัวอย่าง

ประชากรและตัวอย่างของงานวิจัยครั้งนี้มาจากข้อมูลทุติยภูมิจากงานวิจัยเรื่อง การพัฒนาเครื่องมือดิจิทัลเพื่อวินิจฉัยโมทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ เรื่อง พันธุศาสตร์ (Kantahan et al., 2020; Kantahan, 2020) ประชากรเป็นผู้เรียนชั้นมัธยมศึกษาปีที่ 4 ภาคเรียนที่ 2 ปีการศึกษา 2562 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 31 จำนวน 10,362 คน กลุ่มตัวอย่างจำนวน 713 คน ซึ่งได้จากการสุ่มแบบหลายขั้นตอน (multi stage random sampling) ที่ครอบคลุมทุกระดับความสามารถ ประกอบด้วย เก่ง ปานกลาง และอ่อน

2. เครื่องมือที่ใช้ในการวิจัย

(1) แบบทดสอบวินิจฉัยโมทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์แบบพหุมิติ เรื่อง พันธุศาสตร์ พัฒนาโดย Kantahan et al. (2020); Kantahan (2020) เป็นแบบทดสอบปรนัยจำนวน 20 ข้อ มิติความรู้จำนวน 11 ข้อและมิติการให้เหตุผลจำนวน 9 ข้อ ผ่านระบบการทดสอบออนไลน์ เรียกว่า “eMAT-Testing” บน Web Application ที่สามารถตอบสนองต่อผู้เรียนได้โดยตรงและทันที (Junpeng et al., 2020) คุณภาพของเครื่องมือเป็นไปตามเกณฑ์ที่ยอมรับได้ตามมาตรฐานการทดสอบทางการศึกษาและจิตวิทยา (AERA et al., 2014) คือ (1) หลักฐานความตรง สะท้อนให้เห็นถึงความตรงเชิงโครงสร้างภายใน พบว่า โมเดลพหุมิติมีความเหมาะสมและสอดคล้องมากกว่าโมเดลเอกมิติ โดยพิจารณาจากค่า G^2 (Deviance statistic) ของโมเดลพหุมิติมีค่าน้อยกว่าการศึกษาแบบเอกมิตินอกจากนี้เมื่อพิจารณาเปรียบเทียบความแตกต่างของค่า G^2 มีค่าไคว์สแควร์ เท่ากับ 192.34 และ $df=2$ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ ระดับ .01 ($\chi^2 = 192.34, df = 2, p = 0.01$) แสดงว่าการวิเคราะห์โดยใช้โมเดลพหุมิติ มีความถูกต้อง น่าเชื่อถือมากกว่า และเหมาะสำหรับการตรวจให้คะแนน ตามแผนที่โครงสร้างที่กำหนดไว้ (2) หลักฐานความเที่ยงแบบ Expected A Posteriori (EAP/PV) มีค่าความเที่ยงของโมเดลพหุมิติเท่ากับ 0.86 และมีค่า SEM อยู่ระหว่าง 0.38 – 0.56 แสดงว่า การประมาณค่ามีความคลาดเคลื่อนปานกลางถึงสูง สะท้อนให้เห็นว่าเครื่องมือดังกล่าวมีความคงเส้น คงวา และ (3) ความเหมาะสมรายข้อ พบว่า มีค่า INFIT MNSQ อยู่ระหว่าง 0.87 – 1.26

ซึ่งมีค่าตามเกณฑ์ที่ 0.75-1.33 แสดงว่าข้อคำถามทุกข้อมีคุณภาพด้านความตรงเชิงโครงสร้าง โดยสอดคล้องกับแผนที่โครงสร้างทั้งสองมิติ

(2) แบบฟอร์มการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ เป็นตารางอธิบายมโนทัศน์ที่คลาดเคลื่อนจากผลการตอบของผู้เรียนรายข้อ จัดรูปแบบมโนทัศน์ที่คลาดเคลื่อน 5 รูปแบบตามนิยามของ National Research Council (1997) และผลการตอบของผู้เรียนทั้งฉบับแบ่งระดับความเข้าใจมโนทัศน์ 5 ระดับตามนิยามของ Westbrook & Marek (1992)

3. การเก็บรวบรวมและวิเคราะห์ข้อมูล

นำข้อมูลหัตถ์ภูมิ (Kantahan et al., 2020; Kantahan, 2020) ที่ได้จากการตอบแบบทดสอบมาวิเคราะห์ผล เพื่อนำไปสู่การออกแบบนวัตกรรมต้นแบบ มีรายละเอียดดังนี้

3.1 ศึกษาแผนที่โครงสร้างของระดับผลลัพธ์การเรียนรู้ ภาพ 2 และกำหนดแนวทางในการตรวจให้คะแนนตามแผนที่โครงสร้างประกอบด้วย (1) มิติความรู้ แบ่งระดับความสามารถของผู้เรียน 5 ระดับ มีการตรวจให้คะแนนสองค่า คือ 0 และ 1 และ (2) มิติการให้เหตุผล แบ่งระดับความสามารถของผู้เรียน 4 ระดับ มีการตรวจให้คะแนนหลายค่าตั้งแต่ 0-3 คะแนน

ระดับสูง RE	มิติที่ 1 ความรู้ทางวิทยาศาสตร์	ระดับสูง RE	มิติที่ 2 การให้เหตุผลทางวิทยาศาสตร์
↑	5 การคิดเชิงขยาย (Extended Thinking)	↑	4 การเชื่อมโยงแนวคิดวิทยาศาสตร์ได้สมบูรณ์ (Complete Link and Science Idea)
	4 การคิดเชิงกลยุทธ์/มียุทธวิธีในการคิด (Strategic Thinking)		3 การเชื่อมโยงในการคิด (Relating and Link)
	3 ทักษะ/ความคิดรวบยอด (Skill/Concept)		2 การอธิบาย (Skill and Representing)
	2 จดจำได้/เรียกคืน (Recall)		1 ตอบไม่สอดคล้อง/ ไม่เกี่ยวข้อง (Non-recall)
↓	1 ตอบไม่สอดคล้อง/ ไม่เกี่ยวข้อง (Non-recall)	↓	
ระดับต่ำ RE		ระดับต่ำ RE	

ภาพ 2 แผนที่โครงสร้างมิติความรู้และมิติการให้เหตุผล (Kantahan et al., 2020; Kantahan, 2020)

3.2 นำข้อมูลหัตถ์ภูมิจากงานวิจัยของ Kantahan et al. (2020); Kantahan (2020) มาตรวจให้คะแนนและจัดกลุ่มตามระดับแผนที่โครงสร้าง เพื่อเปรียบเทียบความสามารถของนักเรียนกับความสามารถของแบบทดสอบ และวิเคราะห์คำตอบของผู้เรียนในแต่ละระดับแผนที่โครงสร้าง รูปแบบมโนทัศน์ที่คลาดเคลื่อนที่เป็นไปได้ทั้งหมด เพื่อเป็นแนวทางในการออกแบบตัวเลือกแบบเลือกตอบในมิติการให้เหตุผลและการป้อนกลับมโนทัศน์ที่คลาดเคลื่อนในแต่ละมิติ

3.3 สร้างจุดวัดระดับความสามารถของผู้เรียนโดยใช้โปรแกรม ACER ConQuest Version 2.0 (Wu et al., 2007) ด้วยวิธี MLE เนื่องจากมีความคงเส้นคงวาในการประมาณค่าและง่ายต่อการนำไปใช้จริง (Paneaw et al., 2021; Paneaw, 2021) ในการประมาณค่าพารามิเตอร์ของแบบทดสอบและผู้เรียน พร้อมนำเสนอเป็น Wright map โดยนำเสนอแยกตามมิติ ซึ่งประกอบด้วย มิติความรู้และมิติการให้เหตุผล

3.4 นำค่าชั้นของความคลาดเคลื่อน (threshold) ในแต่ละชั้นของการตอบแบบเดียวกัน รวมกันแล้วหาค่าเฉลี่ยในแต่ละมิติ เพื่อกำหนดเป็นจุดตัดในการแบ่งระดับความสามารถ และนำค่าของจุดตัด เปรียบเทียบกับ Wright map เพื่อกำหนดพื้นที่ของแต่ละระดับความสามารถ

3.5 นำค่าจุดตัดซึ่งเป็นพารามิเตอร์ความสามารถที่เหมาะสมที่ปรากฏใน Wright map แปลงเป็นคะแนนสเกล (θ_d) ที่เป็นคะแนนมาตรฐานในการวินิจฉัย โดยใช้สูตรดังสมการ คะแนนสเกล = $50 + 10(\theta_d)$ เมื่อ θ_d แทน ค่าความสามารถของผู้สอบซึ่งประมาณได้จากโมเดล MRCML ในมิติที่ทำการวินิจฉัย

3.6 กำหนดตัวแปรจากข้อมูลพื้นฐานและผลการตอบของผู้เรียน เพื่อพิจารณาว่าตัวแปรใดบ้าง ที่ส่งผลต่อระดับความสามารถของผู้เรียน เพื่อนำไปสู่การสร้างสมการการทำนายในการเรียนรู้ของเครื่อง

ระยะที่ 2 ออกแบบนวัตกรรมต้นแบบ

ในระยะนี้เป็นการออกแบบนวัตกรรมต้นแบบ โดยการสนทนากลุ่ม (focus group) ร่วมกับผู้เชี่ยวชาญ ซึ่งใช้ผลการวิเคราะห์ระยะที่ 1 มาออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ โดยการให้ ข้อมูลป้อนกลับแบบอัตโนมัติผ่านการเรียนรู้ของเครื่อง มีรายละเอียดดังนี้

1. ผู้ให้ข้อมูล

ผู้ให้ข้อมูลในการวิจัยระยะที่ 2 เป็นผู้เชี่ยวชาญจำนวนทั้งสิ้น 6 คน ได้แก่ ครูวิทยาศาสตร์ 2 คน วิศวกรคอมพิวเตอร์ 2 คน และผู้เชี่ยวชาญด้านการวัดและประเมินผล 2 คน

2. เครื่องมือ

แบบประเมินแบบสำรวจรายการตามมาตรฐานการประเมินคุณภาพของนวัตกรรมต้นแบบตาม Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2010; Varasunun, 2011) ประกอบด้วย 5 มาตรฐาน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสม ด้านความถูกต้อง และด้านความรับผิดชอบการประเมิน

3. การเก็บรวบรวมและวิเคราะห์ข้อมูล

3.1 จัดเตรียมข้อมูลที่ได้จากการวิเคราะห์การตอบของผู้เรียนระยะที่ 1

3.2 สนทนากลุ่มร่วมกับผู้เชี่ยวชาญ ในการออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ โดยการให้ข้อมูลป้อนกลับผ่านการเรียนรู้ของเครื่อง ประกอบด้วย ขั้นตอนการทำงานของระบบ รายละเอียดแต่ละขั้นตอน วิธีการในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนและการรายงานผลสู่ผู้เรียน

3.3 กำหนดวิธีการในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ของผู้เรียน และขั้นตอนการให้ข้อมูลป้อนกลับผ่านแพลตฟอร์มการเรียนรู้ดิจิทัล (work flow) อย่างเป็นขั้นตอน

3.4 วิเคราะห์และพัฒนาคุณภาพของนวัตกรรมต้นแบบโดยการสนทนากลุ่ม พบว่า ผลการประเมินคุณภาพที่ได้เป็นไปตามมาตรฐานการประเมินคุณภาพของนวัตกรรมต้นแบบใน 5 มาตรฐาน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสม ด้านความถูกต้อง และด้านความรับผิดชอบ การประเมิน โดยนวัตกรรมต้นแบบมีความทันสมัย สามารถวินิจฉัยถูกต้องตรงตามความสามารถของผู้เรียน และให้การป้อนกลับผู้เรียนได้อย่างทันที ใช้งานได้สะดวกผ่านทางคอมพิวเตอร์หรือสมาร์ทโฟน

ผลการวิจัย

จากการวิเคราะห์ข้อมูลการตอบของผู้เรียนและการร่วมกันออกแบบนวัตกรรมต้นแบบ ผู้วิจัย
ขอเสนอผลการวิจัยในระยะที่ 1 และระยะที่ 2 ดังนี้

1. ผลการวิเคราะห์การตอบของผู้เรียนจากข้อมูลทฤษฎี

1.1 ผลการวิเคราะห์ผลการตอบของผู้เรียนในระดับแผนที่โครงสร้าง จากข้อมูลทฤษฎี
(Kantahan et al., 2020; Kantahan, 2020) ด้วยแบบทดสอบวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนแบบพหุมิติ
ประกอบด้วยมิติความรู้และมิติการให้เหตุผล ซึ่งมีรายละเอียดดังนี้ มิติความรู้ พบว่าผู้เรียนมีความสามารถ
ในมิติความรู้เฉลี่ยร้อยละ 48.72 ในแต่ละระดับความสามารถตามแผนที่โครงสร้างของข้อสอบรายข้อนั้น ๆ
พบว่า ที่ระดับ 2 ระดับ 3 ระดับ 4 และระดับ 5 ผู้เรียนตอบถูกเฉลี่ยคิดเป็นร้อยละ 48.95, 48.67, 40.12 และ
52.03 ตามลำดับ และมิติการให้เหตุผลทางวิทยาศาสตร์ พบว่าผู้เรียนมีความสามารถในการให้เหตุผลเฉลี่ย
ร้อยละ 54.76 โดยระดับความสามารถของข้อสอบรายข้อตามแผนที่โครงสร้าง สามารถวัดความสามารถของ
ผู้เรียนได้หลายค่า (0-3 คะแนน) ขึ้นอยู่กับการตอบของผู้เรียน และมี 6 ข้อ ที่ผู้เรียนสามารถตอบแบบทดสอบ
เกินระดับความสามารถที่กำหนดไว้เฉลี่ยร้อยละ 2.23 ในข้อที่ 5 , 8 , 10 , 12 , 14 และ 16 ดังตาราง 1

ตัวอย่างเช่น ข้อที่ 14. “ให้ผู้เรียนอธิบายความแตกต่างระหว่าง DNA และ RNA” ระดับความสามารถ
ของข้อสอบอยู่ในระดับที่ 2 การอธิบาย แต่มีผู้เรียนที่สามารถตอบได้เกินระดับความสามารถของข้อสอบร้อยละ
19.07 ซึ่งอยู่ในระดับที่ 3 การเชื่อมโยงในการคิด แสดงว่าข้อสอบสามารถวัดความสามารถของผู้เรียนได้ถึง
ระดับที่ 3 การเชื่อมโยงในการคิดตามแผนที่โครงสร้าง

ตาราง 1 ร้อยละคะแนนที่ได้จากการวิเคราะห์ผลการตอบของผู้เรียนในมิติความรู้และการให้เหตุผล

ข้อ	ระดับข้อสอบ	คะแนนมิติความรู้		ข้อ	ระดับข้อสอบ	คะแนนมิติการให้เหตุผล			
		0	1			0	1	2	3
1	2	37.59	62.41	2	2	69.71	30.29		
3	3	57.92	42.08	5	2	55.82	44.04	0.14	
4	3	44.88	55.12	8	3	33.66	54.42	11.50	0.42
6	4	46.84	53.16	10	3	40.53	44.88	14.45	0.14
7	3	61.43	38.57	12	2	53.16	46.14	0.56	0.14
9	2	47.97	52.03	14	2	29.03	51.89	19.07	
11	3	55.54	44.46	15	3	34.78	43.90	21.32	
13	4	72.93	27.07	16	3	54.42	20.90	24.54	0.14
17	3	36.89	63.11	18	4	36.04	46.98	14.03	2.95
19	5	47.97	52.03						
20	2	54.14	45.86						

1.2 ผลการวิเคราะห์มโนทัศน์ที่คลาดเคลื่อนของผู้เรียน ประกอบด้วยมโนทัศน์ที่คลาดเคลื่อนรายข้อ
และมโนทัศน์ที่คลาดเคลื่อนทั้งฉบับ ซึ่งมีรายละเอียดดังนี้ การวิเคราะห์มโนทัศน์ที่คลาดเคลื่อนของผู้เรียน
รายข้อ มีการจัดรูปแบบของมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนตามที่ National Research Council (1997)
กำหนดไว้ ประกอบด้วย 5 ประเภท ได้แก่ อุปาทาน ความเชื่อที่ไม่เป็นวิทยาศาสตร์ ความเข้าใจผิดเกี่ยวกับ
มโนทัศน์ มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา และมโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับข้อเท็จจริง พบว่า ผู้เรียน

ส่วนมากมีรูปแบบมโนทัศน์ที่คลาดเคลื่อนในรูปแบบความเข้าใจผิดเกี่ยวกับมโนทัศน์และภาษา ตัวอย่างเช่น ข้อที่ 1. “เมนเดลเลือกถั่วลันเตามาใช้ในการทดลอง เนื่องจากมีคุณสมบัติข้อหนึ่งคือเป็นพืชที่ผสมตัวเอง (self-fertilized) ข้อใดให้ความหมายของการผสมตัวเองได้ถูกต้องที่สุด” พบว่าผู้เรียนเลือกคำตอบไม่ถูกต้อง อาจอันเนื่องมาจากผู้เรียนมีมโนทัศน์ที่คลาดเคลื่อนในรูปแบบ (3) ความเข้าใจผิดเกี่ยวกับมโนทัศน์ หรือ (4) มโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา กล่าวคือผู้เรียนอาจจะไม่รู้จักคำศัพท์คำว่า “ผสมตัวเอง” ว่ามีกระบวนการอย่างไร หรือผู้เรียนอาจมีความเข้าใจที่คลาดเคลื่อนของการผสมตัวเอง

การวิเคราะห์มโนทัศน์ที่คลาดเคลื่อนของผู้เรียนทั้งฉบับ เมื่อผู้เรียนทำแบบทดสอบทั้งฉบับเรียบร้อยแล้ว จะมีการรายงานผลมโนทัศน์ที่คลาดเคลื่อนทั้งฉบับ มีการจัดระดับความเข้าใจมโนทัศน์ 5 ระดับตามที่ Westbrook & Marek (1992) กำหนดไว้ ประกอบด้วย 5 ประเภท (1) ระดับความเข้าใจมโนทัศน์ในระดับความไม่เข้าใจ (NU) (2) ระดับความเข้าใจมโนทัศน์ในระดับที่ถูกต้องแต่ไม่สมบูรณ์ (PU) (3) ระดับที่คลาดเคลื่อนบางส่วน (PS) (4) ระดับความเข้าใจมโนทัศน์ในระดับที่สมบูรณ์ (CU) และ (5) ระดับความเข้าใจมโนทัศน์ระดับที่คลาดเคลื่อน (AC) ดังตาราง 2

ตาราง 2 การกำหนดจุดตัด ระดับความสามารถ คะแนนสเกล คะแนนดิบและระดับมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนทั้งฉบับในมิติความรู้และมิติการให้เหตุผล

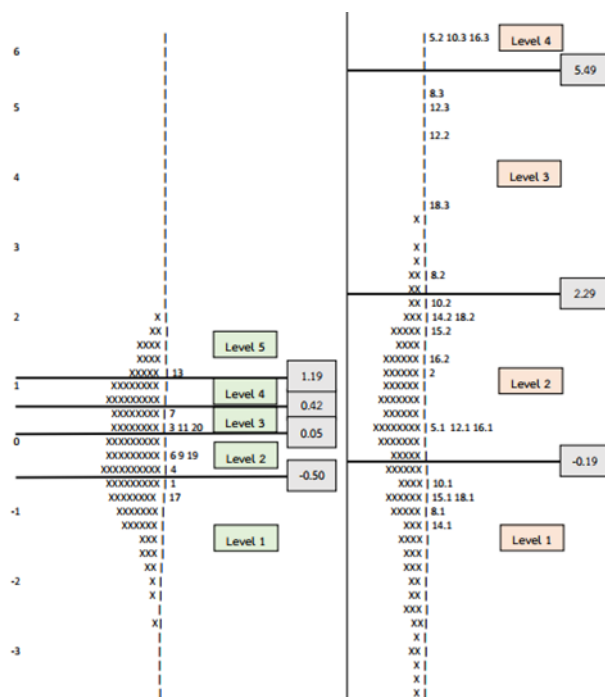
มิติความรู้ทางวิทยาศาสตร์ (คะแนนเต็ม (11 คะแนน))					
จุดตัด	ระดับ ความสามารถ	ช่วงระดับ ความสามารถ (θ)	คะแนนสเกล	ช่วง คะแนนดิบ	ระดับความเข้าใจ มโนทัศน์
$\theta = 1.19$	ระดับ 5	$\theta = 1.19$ ขึ้นไป	61.9 ขึ้นไป	9-11	CU
$\theta = 0.42$	ระดับ 4	$\theta = 0.42$ ถึง 1.18	54.2 – 61.8	7-8	PU
$\theta = 0.05$	ระดับ 3	$\theta = 0.05$ ถึง 0.41	50.5 – 54.1	6	PS
$\theta = -0.50$	ระดับ 2	$\theta = -0.50$ ถึง 0.04	45 – 50.4	4-5	AC
	ระดับ 1	ต่ำกว่า $\theta = -0.50$	ต่ำกว่า 45	0-3	NU
มิติการให้เหตุผลทางวิทยาศาสตร์ (คะแนนเต็ม (22 คะแนน))					
$\theta = 5.49$	ระดับ 4	$\theta = 5.49$ ขึ้นไป	104.8 ขึ้นไป	19-22	CU
$\theta = 2.29$	ระดับ 3	$\theta = 2.19$ ถึง 5.48	71.9 – 104.8	12-18	PU
$\theta = -0.19$	ระดับ 2	$\theta = -0.19$ ถึง 2.18	48.1 – 71.8	5-11	PS
	ระดับ 1	ต่ำกว่า $\theta = -0.19$	ต่ำกว่า 48.1	0-4	NU

1.3 ผลการออกแบบตัวเลือกแบบเลือกตอบ เพื่อการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนแบบอัตโนมติและทันที ผู้วิจัยจึงนำผลการตอบแบบทดสอบมิติการให้เหตุผลซึ่งเป็นคำตอบแบบอัตนัย จากข้อมูลทฤษฎีของ Kantahan et al. (2020); Kantahan (2020) นำมาวิเคราะห์ตามแผนที่โครงสร้าง และออกแบบตัวเลือกแบบเลือกตอบ พบว่า แต่ละตัวเลือกจะมีคะแนนที่แตกต่างกันและสอดคล้องกับระดับความสามารถของแบบทดสอบข้อนั้น ๆ ดังตาราง 3 ตัวอย่างเช่น ข้อที่ 14. “ให้ผู้เรียนอธิบายความแตกต่างระหว่าง DNA และ RNA” ผู้เรียนสามารถตอบได้เกินระดับความสามารถของข้อสอบที่อยู่ในระดับที่ 3 การเชื่อมโยงในการคิด

จึงออกแบบตัวเลือกแบบเลือกตอบที่สามารถวัดถึงระดับที่ 3 โดยตรวจให้คะแนนหลายค่า 0-2 คะแนน และสร้างตัวเลือกแบบเลือกตอบได้ 3 ระดับความสามารถ ได้แก่ ระดับ 1 ตอบไม่สอดคล้อง/ไม่เกี่ยวข้อง 0 คะแนน ระดับ 2 การอธิบาย 1 คะแนน และระดับ 3 การเชื่อมโยงในการคิด 2 คะแนน

1.4 ผลการพัฒนาจุดตัดเพื่อวินิจฉัยนิเทศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ จากการวิเคราะห์ผลการตอบของผู้เรียน โดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ MRCML โดยวิเคราะห์ความยากของแบบทดสอบ การกำหนดจุดตัด และการกำหนดพื้นที่ของแต่ละระดับความสามารถบน Wright map พบว่า มิติความรู้ มีระดับความสามารถของผู้เรียน 5 ระดับ 4 จุดตัดที่ -0.50 0.05 0.42 และ 1.19 และ มิติการให้เหตุผล มีระดับความสามารถของผู้เรียน 4 ระดับ 3 จุดตัดที่ -0.19 2.29 และ 5.49 นำไปสู่การกำหนดช่วงการประเมินระดับความสามารถ ดังภาพ 3 และแปลงค่าช่วงคะแนนสเกลและช่วงคะแนนติดดังตาราง 2 เพื่อนำไปใช้ในการประเมินผลผู้เรียนของนวัตกรรมต้นแบบ

1.5 ข้อมูลพื้นฐานที่ใช้ในการทำนายระดับความสามารถของผู้เรียน เป็นข้อมูลพื้นฐานของผู้เรียนก่อนทำแบบทดสอบที่ได้มาจากข้อมูลทฤษฎีของ Kantahan et al. (2020); Kantahan (2020) นำข้อมูลพื้นฐานของผู้เรียนมาพิจารณาว่าตัวแปรสอดคล้องกับระดับความสามารถของผู้เรียน พบว่า สามารถกำหนดตัวแปรได้ 2 กลุ่ม ประกอบด้วย (1) ตัวแปรต้น ได้แก่ เกรดเฉลี่ยรวม เกรดเฉลี่ยวิชาวิทยาศาสตร์ จำนวนชั่วโมงที่ศึกษาวิทยาศาสตร์ด้วยตนเอง/สัปดาห์ และ (2) ตัวแปรตาม ได้แก่ ความสามารถของผู้เรียนในมิติความรู้และมิติเหตุผล



ภาพ 3 การกำหนดเกณฑ์จุดเปลี่ยนผ่านบน Wright Map ในมิติความรู้ (ด้านซ้าย) และมิติการให้เหตุผล (ด้านขวา)

ตาราง 3 ตัวอย่างการวิเคราะห์ผลการตอบของผู้เรียนและออกแบบตัวเลือกแบบเลือกตอบ

มิติการให้เหตุผลทางวิทยาศาสตร์				
ตัวอย่างข้อสอบ 14. ให้ผู้เรียนอธิบายความแตกต่างระหว่าง DNA และ RNA มิติด้านการให้เหตุผล (ระดับการอธิบาย)				
ระดับความสามารถ	คำบรรยายความสามารถ	ตัวอย่างคำตอบของผู้เรียน	ตัวเลือกแบบเลือกตอบ	การป้อนกลับ
ระดับ 3 การเชื่อมโยงในการคิด	ผู้เรียนสามารถระบุคำตอบได้ถูกต้อง สอดคล้องกับเนื้อหา แสดงถึงการใช้เทคนิควิธีการ ทักษะที่หลากหลายในการวิเคราะห์และให้เหตุผลทางพันธุศาสตร์ที่เกี่ยวข้องสัมพันธ์กับคำตอบ	- DNA เป็นสายคู่ ส่วน RNA เป็นสายเดี่ยว ไนโตรจีนัสเบสของ DNA คือ A ,T ,C,G ของ RNA คือ A,U,C,G น้ำตาลของ DNA เป็นดีออกซีไรโบสและของ RNA เป็น ไรโบส - DNA เป็นสองสาย ที่มีไนโตรจีนัสเบส ATCG / RNA เป็นสายเดี่ยว AUCG	ง. DNA เป็นสายนิวคลีโอไทด์ 2 สาย แต่ RNA เป็นสายนิวคลีโอไทด์ 1 สาย /เบสที่พบใน DNA และ RNA จะต่างกันว่า T,U โดย DNA มี T และ RNA มี U (คะแนน 2)	0 คะแนน PDF + อธิบายโมโนที่คลาดเคลื่อน ก. สิ่งที่แตกต่างกันของการทำงานของชนิดเบาคูสม น้ำตาลที่ใช้ และแหล่งที่พบคืออะไร ข. DNA และ RNA มีไนโตรจีนัสเบส
ระดับ 2 การอธิบาย	ผู้เรียนสามารถระบุคำตอบได้ถูกต้อง สอดคล้องกับเนื้อหา แสดงถึงความเข้าใจในเหตุผลที่มาสสนับสนุนคำตอบ สามารถให้เหตุผลที่สอดคล้องกับคำตอบได้บ้างหรือใกล้เคียง อาจมีเหตุผลบางส่วนไม่ถูกต้อง	- DNA มีสภาพเป็นสายคู่ RNA มีสภาพเป็นสายเดี่ยว / DNA มีไนโตรจีนัสเบส RNA มีเบส U เบสเข้ามาแทนเบส T - DNA เป็นหน่วยควบคุมลักษณะทางพันธุกรรม / RNA เป็นตัวนำข้อมูลจาก DNA ส่งไปยังไรโบโซม - RNA มีน้ำตาลดีออกซีไรโบส มีเบส AUGC / DNA มีน้ำตาลไรโบส มีเบส ATGC	ข. DNA มีสภาพเป็นสายคู่ RNA มีสภาพเป็นสายเดี่ยว / DNA มีไนโตรจีนัสเบส RNA มีเบส U เบสเข้ามาแทนเบส T (คะแนน1) ค. DNA เป็นหน่วยควบคุมลักษณะทางพันธุกรรม / RNA เป็นตัวนำข้อมูลจาก DNA ส่งไปยังไรโบโซม (คะแนน 1)	อะไรบ้าง 1 คะแนน PDF + อธิบายโมโนที่คลาดเคลื่อน ค. นอกจากหน้าที่ที่แตกต่างกันแล้ว DNA และ RNA มีลักษณะใดอีกที่ต่างกัน 2 คะแนน PDF ง. ผู้เรียน DNA และ RNA สามารถเชื่อมโยงไปสู่สถานการณ์ปัญหาอื่น ๆ ทางพันธุศาสตร์ในเรื่องนี้ได้
ระดับ 1 ไม่ตอบ/ตอบไม่สอดคล้อง	ผู้เรียนไม่ระบุคำตอบ หรือมีการระบุคำตอบที่ไม่ถูกต้อง หรือไม่ใกล้เคียง เป็นเพียงการตอบที่มาจากกาจดจำได้ แต่ขาดการอธิบายเหตุผลสนับสนุน	- การทำงานต่างกัน ชนิดเบสคู่สมต่างกัน น้ำตาลที่ใช้ต่างกัน แหล่งที่พบมากต่างกัน	ก. DNA และ RNA มีการทำงานที่ต่างกัน ชนิดเบสคู่สมต่างกัน น้ำตาลที่ใช้ต่างกัน แหล่งที่พบมากต่างกัน (คะแนน 0)	

2. ผลการออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ของผู้เรียน โดยการให้ข้อมูล ป้อนกลับผ่านการเรียนรู้ของเครื่อง (Machine Learning)

จากการวิเคราะห์ผลการตอบในระยะเวลาที่ 1 นำมาสู่การสนทนากลุ่มเพื่อออกแบบนวัตกรรมต้นแบบ
ผ่านการเรียนรู้ของเครื่อง ซึ่งประกอบด้วย ขั้นตอนการทำงานของระบบและรายละเอียดแต่ละขั้นตอน โดยมี
รายละเอียดดังนี้

2.1 ขั้นตอนการทำงานของระบบ ดำเนินการออกแบบนวัตกรรมต้นแบบ ที่สามารถให้การวินิจฉัย
มโนทัศน์ที่คลาดเคลื่อนของผู้เรียนได้อย่างอัตโนมัติและมีการป้อนกลับข้อมูลไปยังผู้เรียนได้อย่างทันที
ผ่าน web application ที่มีชื่อว่า “eMAT-Testing” ซึ่งพัฒนาต่อจาก Junpeng et al., (2020) รองรับระบบ
iOS Android และ Window ประกอบด้วย 4 ส่วน คือ (1) ส่วนรับข้อมูล (2) ส่วนประมวลผล (3) ส่วนให้ข้อมูล
ป้อนกลับอัตโนมัติ และ (4) ส่วนรายงานผลการประเมิน สรุปดังภาพ 4 แสดงรายละเอียดแต่ละขั้นตอนดังนี้

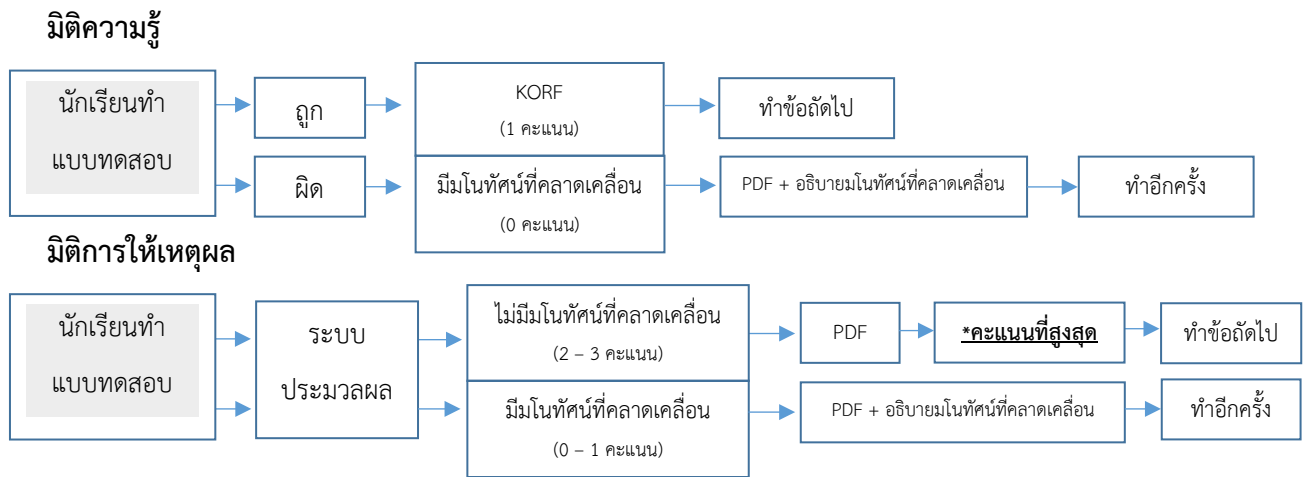


ภาพ 4 ระบบการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ผ่านการเรียนรู้ของเครื่อง

1.1.1 ส่วนรับข้อมูล เป็นการรับข้อมูลจากผู้เรียนเข้าสู่ระบบประกอบด้วย (1) ส่วนข้อมูล
พื้นฐานของผู้เรียน เพื่อเป็นข้อมูลเบื้องต้นในการเรียนรู้ของเครื่อง (2) การทำแบบทดสอบ โดยเป็นแบบทดสอบ
ปรนัยจำนวน 20 ข้อ มิตติความรู้จำนวน 11 ข้อและมิตติการให้เหตุผลจำนวน 9 ข้อ

1.1.2 ส่วนประมวลผล เป็นขั้นตอนการตรวจให้คะแนนแบบทดสอบอัตโนมัติ โดยนำส่วน
รับข้อมูลมาประมวลผลร่วมกับผลจากการวิเคราะห์ในระยะเวลาที่ 1 ประกอบด้วย (1) ส่วนวินิจฉัยที่ให้คะแนน
ข้อสอบ โดยมีการให้คะแนนแบบทดสอบรายข้อและเมื่อทำแบบทดสอบแล้วเสร็จ ระบบจะทำการรวมคะแนน
ดิบของข้อสอบในแต่ละมิติเทียบกับคะแนนดิบที่กำหนดไว้จากตาราง 2 จากนั้นระบบจะนำไปประมวลผล
เทียบกับคะแนนสเกล และนำไปสู่การวินิจฉัยระดับความสามารถทางวิทยาศาสตร์ในมิตินั้น ๆ ว่าอยู่ในระดับใด
(2) ส่วนวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน โดยมีการวินิจฉัยรายข้อจากการเลือกตอบแต่ละข้อของผู้เรียนและ
การวินิจฉัยทั้งฉบับที่เทียบคะแนนของผู้เรียนกับตาราง 2

1.3 ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ จากการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน
จะมีการให้ข้อมูลป้อนกลับจากการทำแบบทดสอบของผู้เรียน ประกอบด้วย (1) มิตติความรู้ ถ้านักเรียนตอบถูก
จะได้รับการป้อนกลับแบบบอกผลการตอบ (Knowledge of Result Feedback: KORF) ซึ่งบอกเพียงถูกหรือผิด
หากนักเรียนตอบผิดจะได้รับการป้อนกลับแบบบางส่วนโดยใช้การชี้แนะ (Partial Directive Feedback: PDF)
ร่วมกับการอธิบายมโนทัศน์ที่คลาดเคลื่อน และ (2) มิตติการให้เหตุผล หากนักเรียนตอบได้คะแนนสูงสุดหรือ
อยู่ในช่วง 2 – 3 คะแนน จะได้รับการป้อนกลับแบบบางส่วนโดยใช้การชี้แนะ และถ้าได้คะแนนน้อยกว่า 2
จะได้รับการป้อนกลับแบบบางส่วนโดยใช้การชี้แนะร่วมกับการอธิบายมโนทัศน์ที่คลาดเคลื่อน ดังภาพ 5



ภาพ 5 ขั้นตอนการให้ข้อมูลป้อนกลับในมิตินความรู้และมิตินการให้เหตุผล

1.4 ส่วนรายงานผลการประเมิน เป็นการรายงานผลผู้เรียนรายบุคคล จากการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนจะมีการรายงานระดับความสามารถของผู้เรียนในแต่ละมิติตามแผนที่โครงสร้างซึ่งเป็นสิ่งที่ผู้เรียนทำได้ ความเข้าใจที่คลาดเคลื่อนทั้งฉบับ ตลอดจนสิ่งที่ผู้เรียนควรพัฒนาต่อไป ประกอบด้วย (1) ข้อมูลส่วนบุคคล (2) ข้อมูลสารสนเทศและสถิติของการวินิจฉัย (3) ข้อมูลจากการวินิจฉัยระดับความสามารถของผู้เรียนในแต่ละมิติตามแผนที่โครงสร้างซึ่งเป็นสิ่งที่ผู้เรียนทำได้ ความเข้าใจที่คลาดเคลื่อนทั้งฉบับ และสิ่งที่ผู้เรียนควรพัฒนาต่อไป

อภิปรายผล

1. จากผลการวิเคราะห์การตอบของผู้เรียนในระดับแผนที่โครงสร้าง พบว่า มีความสอดคล้องกับระดับแผนที่โครงสร้าง โดยแผนที่โครงสร้างสามารถแยกระดับความสามารถของผู้เรียนได้อย่างชัดเจน และในมิตินการให้เหตุผลผู้เรียนสามารถตอบได้เกินระดับความสามารถของข้อสอบบางข้อ อาจเนื่องมาจากการให้เหตุผลเป็นการแสดงออกทางความคิดปลายเปิด หรือมีตัวอย่างที่ใกล้เคียงกับชีวิตประจำวัน เมื่อวิเคราะห์มโนทัศน์ที่คลาดเคลื่อนของผู้เรียน พบว่าส่วนมากผู้เรียนมีรูปแบบมโนทัศน์ที่คลาดเคลื่อนรายชื่อในรูปแบบมีความเข้าใจผิดเกี่ยวกับมโนทัศน์และมโนทัศน์ที่คลาดเคลื่อนเนื่องจากภาษา เนื่องจากธรรมชาติของรายวิชาที่มีลักษณะที่เข้าใจยาก เป็นนามธรรม และมีเนื้อหาและคำศัพท์เฉพาะทางวิทยาศาสตร์ (Kerdkliang et al., 2019) นอกจากนี้ข้อมูลพื้นฐานที่สามารถทำนายระดับความสามารถของผู้เรียน ได้แก่ เกรดเฉลี่ยรวม เกรดเฉลี่ยวิชาวิทยาศาสตร์ จำนวนชั่วโมงที่ศึกษาวิชาวิทยาศาสตร์ด้วยตนเอง/สัปดาห์ สอดคล้องกับ Gadzella et al. (2002) กล่าวว่า จากการคำนวณสมการถดถอยตัวทำนายเกรดเฉลี่ยที่ดีที่สุดคือคะแนนการทดสอบจากการวัดและประเมินผล

2. จากผลการออกแบบวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนทางวิทยาศาสตร์ โดยการให้ข้อมูลป้อนกลับแบบอัตโนมัติผ่านการเรียนรู้ของเครื่อง ซึ่งวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนจากข้อมูลนำเข้า อีกทั้งยังรวบรวมข้อมูลของผู้เรียนเก็บเป็นคลังข้อมูลเพื่อเพิ่มประสิทธิภาพในการทำนายมากยิ่งขึ้น โดยการ

รวบรวมฐานข้อมูลขนาดใหญ่ เปรียบเป็นเหมือนสมองให้กับการศึกษา (Nafea, 2018) นวัตกรรมต้นแบบประกอบด้วย 4 ส่วน คือ (1) ส่วนรับข้อมูล (2) ส่วนประมวลผล (3) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ และ (4) ส่วนรายงานผลการประเมิน ซึ่งส่วนที่มีบทบาทที่สำคัญในการวิธีการวินิจฉัยมากที่สุด ได้แก่ (2) ส่วนประมวลผล และ (3) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ เนื่องจากส่วนประมวลผล เป็นในการตรวจให้คะแนนแบบทดสอบ หากส่วนนี้ผิดพลาดจะทำให้การวินิจฉัยมโนทัศน์ของผู้เรียนคลาดเคลื่อน ส่งผลต่อการให้ข้อมูลป้อนกลับแก่ผู้เรียนและรายงานผลต่อผู้เรียนต่อไป และส่วนให้ข้อมูลป้อนกลับอัตโนมัติเป็นอีกหนึ่งส่วนที่สำคัญ โดยการป้อนกลับแบบปรับเหมาะตามระดับความสามารถของผู้เรียนนี้ (adaptive feedback framework) จะสอดคล้องกับคุณลักษณะและความต้องการส่วนบุคคลของผู้เรียนมากขึ้น อีกทั้งทำให้ผู้เรียนได้ทบทวนและพิจารณาความเข้าใจที่คลาดเคลื่อนของตนเอง Gouli et al. (2008)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 จากผลการวิเคราะห์มโนทัศน์ของผู้เรียน พบว่ารูปแบบมโนทัศน์ที่คลาดเคลื่อนที่พบมาก ได้แก่ รูปแบบความเข้าใจผิดเกี่ยวกับมโนทัศน์และเนื่องจากภาษา ดังนั้น ผู้นำผลวิจัยไปใช้ควรเน้นการสร้างคำถามตัวเลือกแบบเลือกตอบหรือการป้อนกลับผู้เรียนในส่วนของรูปแบบมโนทัศน์ที่คลาดเคลื่อนดังกล่าว เพื่อแก้ไขหรือพัฒนามโนทัศน์ที่คลาดเคลื่อนเรื่อง พันธุศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 4

1.2 จากการผลการออกแบบนวัตกรรมต้นแบบพบว่า ส่วนที่มีบทบาทที่สำคัญในการวิธีการวินิจฉัยมากที่สุด ได้แก่ (2) ส่วนประมวลผล และ (3) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ ดังนั้น ผู้นำผลวิจัยไปใช้ควรทำความเข้าใจกับการกำหนดจุดตัด ระดับความสามารถ คะแนนสเกล คะแนนดิบและระดับมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนเป็นข้อมูลเบื้องต้น เพื่อเป็นแนวทางที่สำคัญในการวินิจฉัยผลการตอบของผู้เรียน อีกทั้งในส่วนให้ข้อมูลป้อนกลับอัตโนมัติ เป็นส่วนที่สำคัญในการแก้ไขมโนทัศน์ของผู้เรียน การกำหนดรูปแบบการป้อนกลับต่างๆ ควรคำนึงความสัมพันธ์กับแผนที่โครงสร้างและมโนทัศน์ที่คลาดเคลื่อนที่เกิดขึ้นในแต่ละข้อที่แตกต่างกัน อีกทั้งยังมีขั้นตอนในการให้ข้อมูลป้อนกลับแตกต่างกันในมิติความรู้และการให้เหตุผลด้วย

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 การวิจัยครั้งนี้มีการให้ข้อมูลป้อนกลับแบบรายข้อที่สามารถวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนจากการเลือกตัวเลือกในแต่ละข้อได้ จึงควรมีการศึกษาเปรียบเทียบระหว่างกลุ่มควบคุมและกลุ่มทดลองเพื่อศึกษาและเปรียบเทียบคุณภาพของนวัตกรรมต้นแบบที่พัฒนาขึ้น

2.2 เนื่องจากวิธีการวินิจฉัยครั้งนี้ มีการตอบสนองให้แก่ผู้เรียนฝ่ายเดียว ในการวิจัยครั้งต่อไปอาจมีการออกแบบให้สามารถรับข้อมูลจากผู้เรียนเพิ่มเติมได้ เช่น ข้อเสนอแนะ การเสนอตัวเลือกแต่ละข้อเพิ่มเติมเพื่อให้ระบบหรือผู้พัฒนาสามารถพัฒนาวิธีการวินิจฉัยให้ดียิ่งขึ้นต่อไป

References

- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Standards for Educational and Psychological Testing. The American Educational Research Association
- Adams, R.J., Wilson, M., & Wang, W.C. (1997). The multidimensional random coefficients multinomial logic model. *Applied Psychological Measurement*, 21(1), 1-23.
- Gadzella, B. M., Baloglu, M., & Stephens, R. (2002). Prediction of GPA with Educational psychology grades and critical thinking scores. *Project Innovation (Alabama)*, 122(3), 618-623.
- Gouli, E., Gogoulou, A., & Grigoriadou, M. (2008). Supporting Self- Peer- and Collaborative- Assessment in E-Learning: The Case of the PEer and Collaborative ASSESSment Environment (PECASSE). *Journal of Interactive Learning Research*, 19(4), 615-647.
- Gurel, K. G, Eryilmaz, A., & McDermott, L. C. (2015). A Review and Comparison of Diagnostic Instruments to Identify Students' Misconceptions in Science. *Eurasia Journal of Mathematics. Science & Technology Education*, 11(5).
<https://doi.org/10.12973/eurasia.2015.1369a>
- Kantahan, S., Junpeng, P., Punturat, S., Tang, K. N., P. Gochyyev., & Wilson, M. (2020). Designing and verifying a tool for diagnosing multidimensional scientific misconceptions in genetics topic. *International Journal of Evaluation and Research in Education*, 9(3), 564-571. <https://doi.org/10.11591/ijere.v9i3.20544>
- Kalas, P., O'Neill, A., Pollock, C., & Birol, G. (2013). Development of a meiosis concept inventory. *CBE Life Sciences Education*, 12(4), 655-664. <https://doi.org/10.1187/cbe.12-10-0174>
- Ling, T.W. (2017). *Fostering Understanding and Reducing Misconceptions about Image Formation by a Plane Mirror Using Constructivist-Based Hands-On Activities*. Springer.
- Mitchell, T. M. (1997). Does Machine Learning Really Work?. *AI Magazine*, 18(3), 11.
<https://doi.org/10.1609/aimag.v18i3.1303>
- Mohammed, M., Khan, M.B., and Bashier, E.B.M. (2016). *Machine learning: algorithms and applications*. CRC Press. <https://doi.org/10.1201/9781315371658>

- Nafea, I. (2018). Machine Learning in Educational Technology. *Machine Learning – Advanced Techniques and Emerging Applications*, 175-183. <https://doi.org/10.5772/intechopen.72906>
- National Research Council. (1997). *Science teaching reconsidered*. National Academies Press.
- Paneaw, S., Junpeng, P., & Tang, K. N. (2021). Designing Standards-Setting for Levels of Mathematical Proficiency in Measurement and Geometry: Multidimensional Item Response Model. *Journal of Education and Learning*, 10(6), 103-111. <https://doi.org/10.5539/jel.v10n6p103>
- Soeharto, S., Csapó, B., Sarimanah, E., Dewi, F.I., and Sabri, T. (2019). A Review of Students' Common Misconceptionss in Science and Their Diagnostic Assessment Tools. *Journal Pendidikan IPA Indonesia*, 8(2), 247-266. <https://doi.org/10.15294/jpii.v8i2.18649>
- Westbrook, S. L., & Marek, E. A. (1992). "A cross-age study of student understanding of the concept of diffusion". *Journal of Research in Science Teaching*, 29(1), 51-61.
- Widiyatmoko, A., & Shimizu, K. (2018). Literature Review of Factors Contributing to Students' Misconceptions in Light and Optical Instruments. *International Journal of Environmental and Science Education*, 13(10), 853-863.
- Wilson, M. (2005). Constructing measures: An item response modeling approach. *Evaluation and Program Planning*, 28(4), 433-434. <https://doi.org/10.1016/j.evalprogplan.2005.07.008>
- Wu, M. L., Adams, R. J., & Wilson, M. R. (2007). *ACER Conquest: Generalized item response modelling software*. ACER.
- Wu, J. Y., Hsiao, Y. C., & Nian, M. W. (2018). Using supervised machine learning on large-scale online forums to classify course-related Facebook messages in predicting learning achievement within the personal learning environment. *Interactive Learning Environments*, 28(1), 1-16. <https://doi.org/10.1080/10494820.2018.1515085>
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program Evaluation standards: A guide for evaluators and evaluation users* (3rd Ed.). Sage.
- Zhang, F., Zhang, X., Luo, M., & Geng, H. (2016). *The Effects of Feedback on Memory Strategies of Younger and Older Adults*. PLOS ONE. <https://doi.org/10.1371/journal.pone.0168896>

Translate Thai References

- Junpeng, P. (2018). *Application of Multidimensional Item Response Theory to Research*. Khon Kaen University Press. (in Thai)
- Junpeng, P., Marwiang, M., Chinjunthuk, S., Suwannatrai, P., Grotha, J., Chanayotha, K., Tawarungruang, C., Thuanman, J., Ngang, T.K., & Wilson, M. R. (2020). *Developing Students' Mathematical Proficiency Level Diagnostic Tools through Information Technology in Assessment for Learning Report*. Thailand Science Research and Innovation. (in Thai)
- Kantahan, S. (2020). *Designing and verifying a tool for diagnosing multidimensional scientific misconceptions in genetics topic* [Master's thesis]. Khon Kaen University (in Thai)
- Kerdkliang, S., Intasuwan, S., & Wannasuth, N. (2019). Effect of Inquiry Learning Cycle (7E) and Jigsaw Technique on Plants Reproduction and Development Topic to Develop Biological Learning Achievement and Social Skill for Mattayomseuksa 5 Student. *Research and Innovation for Sustainability Development*, 29(29), 1943-1950. (in Thai)
- Masantiah, J. (2017). *Development of A Computer-Based Testing System with Immediate Feedback for Different Student's Ability Levels: Application of Rasch Sirt Model* [Master's thesis]. Chulalongkorn University. (in Thai)
- Paneaw, S. (2021). *Standard Setting of Mathematical Proficiency Level for the Mixed Format Test through Digital Technology* [Master's thesis]. Khon Kaen University. (in Thai)
- Thakho, P., & Termtachatipongsa, P. (2015). The Development of Grade 11 Students' Conception of Photosynthesis using Inquiry-Based Approach Associated with Scaffolding Strategies. *Journal of Education Khon Kaen University*, 38(3), 94-104. (in Thai)
- Varasunun, P. (2011). The Program Evaluation Standards (3rd Ed.) in 2010. *Journal of Research Methodology*, 24(2), 273-278. (in Thai)
- Wongwanich, S. (2017). *Design Research in Education*. Chulalongkorn University Press. (in Thai)

การออกแบบระบบตรวจให้คะแนนแบบสอบอัตนัย โดยการให้ข้อมูลป้อนกลับอัตโนมัติ เพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่านการเรียนรู้ของเครื่อง Designing Automated Scoring System of Open-Ended Test by Providing Automatic Feedback to Diagnose Mathematical Proficiency Levels through Machine Learning

ศราวรรณ สุขวิชัย^{1*} พัทธี จันทร์เพ็ง² ชัยวัฒน์ ทะวะรุ่งเรือง³ และ ธนพงศ์ อินทะ⁴
Sarawan Sukwichai^{1*}, Putcharee Junpeng², Chaiwat Tawarungruang³ and Thanapong Intharah⁴

(Received: March 31, 2022; Revised: May 23, 2022; Accepted: May 27, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมาย (1) เพื่อวิเคราะห์ผลการตอบของผู้เรียนสู่การออกแบบนวัตกรรมต้นแบบ ประกอบด้วย มโนทัศน์ที่คลาดเคลื่อนของผู้เรียน การกำหนดจุดตัดระดับความสามารถทางคณิตศาสตร์ การให้ข้อมูลป้อนกลับตามมโนทัศน์ที่คลาดเคลื่อนและตัวแปรที่ใช้ในการทำนายผ่านการเรียนรู้ของเครื่อง (machine learning) และ (2) เพื่อออกแบบระบบตรวจให้คะแนนแบบสอบอัตนัย ของผู้เรียนผ่านการเรียนรู้ของเครื่อง (machine learning) โดยใช้ระเบียบวิธีวิจัยการออกแบบ กลุ่มผู้สอบ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 จำนวนนักเรียนทั้งหมด 495 คน เครื่องมือที่ใช้ คือ แบบทดสอบอัตนัย สารสถิติและความน่าจะเป็น ผ่านชุดเครื่องมือวินิจฉัยในระบบการทดสอบออนไลน์ “eMAT-Testing” วิเคราะห์ข้อมูลโดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ MRCML ผลการวิจัยปรากฏดังนี้

¹ นิสิตระดับปริญญาโท หลักสูตรวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² รองศาสตราจารย์ หลักสูตรการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

³ นักวิจัย ศูนย์จัดการข้อมูลและวิเคราะห์ทางสถิติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

⁴ อาจารย์ ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

งานวิจัยนี้ได้รับทุนอำนวยการวิทยานิพนธ์ทุนกลาง ประจำปี 2565 จากสำนักงานการวิจัยแห่งชาติ (วช.) รหัสโครงการวิจัย N42A650236

¹ Master's degree student, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

² Associate Professor Department of Educational Measurement and Evaluation Faculty of Education Khon Kaen University

³ Researcher, Data Management and Statistical Analysis Center, Faculty of Public Health, Khon Kaen University

⁴ Lecturer, Department of Statistics, Faculty of Science, Khon Kaen University

The research received a research grant for Talented Mid-Career Researchers 2022 from the National Research Council of Thailand (NRCT) Project No. N42A650236

* Corresponding Author E-mail: sarawansukwichai@kkumail.com

1. ผลการวิเคราะห์ผลการตอบของผู้เรียนจากข้อมูลหัตถ์นิยม พบว่า (1) ผู้เรียนมีมีโนทัศน์ทั้งหมด 4 รูปแบบ คือ 1) การใช้ข้อมูล 2) การบิดเบือนทฤษฎีบท กฎ สูตร บทนิยาม และสมบัติ 3) การคำนวณและขาดการตรวจสอบในระหว่างการแก้ปัญหา และ 4) ความผิดพลาดในการตีความด้านภาษาและสัญลักษณ์ทางคณิตศาสตร์ โดยทั้งในมิติกระบวนการทางคณิตศาสตร์ (MAP) และ มิติโครงสร้างความคิดรวบยอด (SLO) นักเรียนมีมีโนทัศน์ที่คลาดเคลื่อนมากที่สุดคือ ด้านการใช้ข้อมูล (2) ผลการกำหนดจุดตัด โดยการกำหนดเกณฑ์พื้นที่บน Wright Map พบว่า มิติกระบวนการทางคณิตศาสตร์ แบ่งได้ 5 ระดับ 4 จุดตัด จากระดับต่ำสุดไปสูงสุด ที่ -0.99 0.26 0.44 และ 0.61 ตามลำดับ และมิติโครงสร้างความคิดรวบยอด แบ่งได้ 5 ระดับ 4 จุดตัด จากระดับต่ำสุดไปสูงสุด ที่ -0.70 0.15 0.76 และ 1.39 ตามลำดับ ซึ่งจุดตัดดังกล่าวสามารถนำไปสู่การกำหนดช่วงระดับความสามารถ คะแนนสเกล และคะแนนดิบเพื่อนำไปใช้เป็นเกณฑ์การประเมินระดับความสามารถทางคณิตศาสตร์ในแต่ละมิติ และ (3) ตัวแปรจากข้อมูลหัตถ์นิยม ที่นำไปใช้ในการสร้างสมการทำนายระดับความสามารถทางคณิตศาสตร์ผ่านการเรียนรู้ของเครื่อง (machine learning) พบว่า ตัวแปรต้นที่จะส่งผลต่อการสร้างสมการทำนาย คือ เกรดคณิตเทอมที่ผ่านมา เกรดเฉลี่ยรวมเทอมที่ผ่านมา ชั่วโมงเรียนพิเศษและชั่วโมงศึกษาคณิตศาสตร์ด้วยตนเอง และตัวแปรตามประกอบด้วย คะแนนดิบมิติกระบวนการทางคณิตศาสตร์ คะแนนดิบมิติโครงสร้างความคิดรวบยอด และคะแนนรวม 2 มิติ

2. ผลการออกแบบการระบบตรวจให้คะแนนผ่าน machine learning สามารถแบ่งออกเป็น 5 ส่วน ประกอบด้วย (1) ส่วนรับข้อมูล (2) ส่วนประมวลผล (3) ส่วนแสดงผล (4) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ และ (5) ส่วนรายงานผลการประเมิน โดยผลการประเมินคุณภาพของการออกแบบระบบได้ผ่านผู้เชี่ยวชาญจำนวน 5 คน ด้วยวิธีการสนทนากลุ่ม ครอบคลุมทั้ง 5 มาตรฐาน ประกอบด้วย ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสมขอค่าธรรมเนียม ด้านความถูกต้อง และด้านความรับผิดชอบการประเมิน

คำสำคัญ: การตรวจให้คะแนนแบบสอบอัตโนมัติ การให้ข้อมูลป้อนกลับอัตโนมัติ วิทยาการออกแบบ

โมเดลการตอบสนองข้อสอบพหุมิติ การเรียนรู้ของเครื่อง

Abstract

The study aimed to (1) analyze students' response patterns for the design of the prototypical innovation, including students' misconceptions, determining cut scores for mathematical proficiency levels, and providing the automated feedback based on misconceptions and variables used to make predictions through machine learning; and (2) design the scoring system for the open-ended question test through machine learning. Design research was applied in this study. The sample comprised 495 grade 7 students. The research instrument was an open-ended test on the topic of Statistics and Probability through diagnostic tools in an online testing system, "eMAT-Testing." The collected data were

analyzed using a multidimensional random coefficients multinomial model (MRCMLM). The results are reported below.

1. Regarding the analysis of the students' response patterns based on the secondary data, it was found that (1) the students had 4 types of misconceptions: 1) misused data; 2) distorted theorems, rules, formulas, definitions and properties; 3) unverified solutions; and 4) misinterpreted language and mathematical symbols. In both the mathematical process dimension (MAP) and the conceptual structures dimension (SLO), the students' most common misconception among the 4 types was the misused data. (2) On determining cut scores with defined criterion zones on the Wright map for diagnosing mathematical proficiency levels through Machine learning, it was found that the mathematic proficiency of MAP was divided into five levels based on four cut scores, ranging from the lowest to the highest as follows: -0.99, 0.26, 0.44 and 0.61 respectively. In the same manner, that of SLO was distinguished into five levels according to four cut scores, including -0.70, 0.15, 0.76 and 1.39 respectively. Such cut scores can be employed to determine proficiency level ranges, scale scores, and raw scores as criteria for assessment of mathematical proficiency levels in each dimension. (3) And regarding the variables from secondary data used to generate the equation to predict mathematical proficiency levels through machine learning, the results showed that the independent variables that would affect the generation of the predictive equation were the mathematics GPA in the previous semester, the total GPA of the previous semester, the number of hours of tuition, and that of hours of self-study on mathematics. In addition, the dependent variables included raw scores of MAP, raw scores of SLO, and the total score of both dimensions.

2. The results also showed that the design of the scoring system for the open-ended question test through machine learning featured 5 parts, namely input, processing, output, automated feedback, and assessment reporting. Based on the qualitative assessment of the system design conducted by 5 experts through a focus group discussion, it covered 5 standards, including utility, feasibility, propriety, accuracy, and evaluation accountability.

Keywords: open-ended question scoring system, automated feedback, design research, multidimensional item response model, machine learning

บทนำ

ความสามารถทางคณิตศาสตร์เป็นสิ่งที่สะท้อนให้เห็นว่าผู้เรียนมีความรู้และทักษะที่เพียงพอหรือไม่ต่อการดำรงชีวิตในศตวรรษที่ 21 ดังนั้นการเตรียมผู้เรียนให้มีความสามารถทางคณิตศาสตร์ จึงเป็นสิ่งสำคัญ ซึ่งส่งผลให้ผู้เรียนรู้เท่าทันการเปลี่ยนแปลงของระบบเศรษฐกิจ สังคม วัฒนธรรม และสภาพแวดล้อม สามารถอยู่ร่วมกับประชาคมโลกได้ (Ministry of Education, 2017) และการประเมินความสามารถทางคณิตศาสตร์ถือเป็นสิ่งสำคัญในการขับเคลื่อนการศึกษา โดยเฉพาะอย่างยิ่ง การประเมินความสามารถทางคณิตศาสตร์ที่ไม่ได้มุ่งเน้นเพียงคำตอบที่ถูกต้องเท่านั้น แต่ยังมุ่งเน้นไปที่กระบวนการคิดทางคณิตศาสตร์ด้วยเช่นกัน (Chinjunthuk & Junpeng, 2020)

จากผลการประเมินการเรียนรู้โครงการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ของนักเรียนไทยเทียบกับนานาชาติ (TIMSS) โดยผลการประเมินของประเทศไทยด้านคณิตศาสตร์ในระดับชั้นมัธยมศึกษาปีที่ 2 พบว่า วิชาคณิตศาสตร์ ในปี ค.ศ. 1999 2007 2011 และ 2015 มีคะแนนเฉลี่ยแนวโน้มน้อยลงอย่างต่อเนื่อง (Chinjunthuk & Junpeng, 2020) ประกอบกับผลการทดสอบทางการศึกษาระดับชาตินักเรียนขั้นพื้นฐาน (O-NET) บ่งชี้เช่นเดียวกันว่าผู้เรียนมีคะแนนเฉลี่ยของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ต่ำกว่าร้อยละ 50 ซึ่งเป็นมาตรฐานขั้นต่ำ โดยพบว่า ผลการสอบ O-NET ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2563 สารที่ได้คะแนนน้อยที่สุด คือ สารที่ 3 สถิติและความน่าจะเป็น มีคะแนนเฉลี่ย 23.85 คะแนน (The National Institute of Educational Testing Service, 2021) ซึ่งการประเมินทั้งสองรูปแบบนี้ต่างมุ่งเน้นการประเมินความสามารถทางคณิตศาสตร์ของผู้เรียน จากผลการประเมินทั้งในระดับนานาชาติและระดับชาติข้างต้น สะท้อนให้เห็นว่านักเรียนไทยมีปัญหาในการเรียนรู้คณิตศาสตร์

เครื่องมือที่นับว่าสำคัญ และมีประโยชน์ในการค้นหาข้อบกพร่อง ตลอดจนปัญหาของผู้เรียนแต่ละคน คือ แบบสอบวินิจฉัย (diagnostic test) ผลการตอบสามารถระบุได้ว่านักเรียนบกพร่องในทักษะใด (Adulpongpaisal, 2009) จากการศึกษาที่ผ่านมาพบว่า มีงานวิจัยที่พัฒนานวัตกรรมระบบการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนทั้งในบริบทของต่างประเทศและในประเทศ เช่น งานวิจัยของ Junpeng et al. (2020) ได้พยายามพัฒนาระบบการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ผ่านเทคโนโลยีดิจิทัล ในสาระสถิติและความน่าจะเป็น ของนักเรียนระดับมัธยมศึกษาปีที่ 1 อย่างไรก็ตามนวัตกรรมการประเมินดังกล่าวยังมีข้อจำกัด คือ ไม่สามารถตรวจให้คะแนนแบบสอบอัตโนมัติในการที่มีการให้คะแนนหลายค่าและไม่สามารถให้ข้อมูลป้อนกลับแบบอัตโนมัติเพื่อวินิจฉัยความสามารถผู้เรียนได้ ในกรณีที่มุ่งเน้นไปที่กระบวนการแก้ปัญหาของผู้เรียน โดยสามารถตรวจอัตโนมัติได้เฉพาะในกรณีที่มุ่งเน้นไปที่คำตอบที่ถูกของผู้เรียน และนำผลการตอบที่ได้ไปเทียบกับเกณฑ์ให้คะแนนที่สอดคล้องกับแผนที่เชิงโครงสร้างในแต่ละระดับ จึงสามารถให้คำแนะนำหรือแหล่งในการเรียนรู้เพิ่มเติม เพื่อให้ผู้เรียนปรับปรุงได้ในทันที หากต้องการพิจารณาถึงกระบวนการแสดงวิธีทำของผู้เรียนเพื่อสะท้อนให้เห็นถึงกระบวนการแก้ปัญหา จำเป็นต้องอาศัยระยะเวลาในการให้ครูหรือผู้ประเมิน ตรวจให้คะแนน แล้วจึงนำผลการตรวจให้คะแนนที่ได้ไปสู่การรายงานผลการวินิจฉัยเป็นรายบุคคล และภาพรวมในลำดับต่อไป

จากที่กล่าวมาจะเห็นได้ว่า ระบบดังกล่าวจำเป็นต้องอาศัยระยะเวลาให้ผู้เรียนดำเนินการทดสอบแล้วเสร็จ จึงสามารถประมวลผลและวินิจฉัยระดับความสามารถทางคณิตศาสตร์ได้ จึงถือเป็นจุดอ่อนที่สำคัญ ซึ่งในการวิจัยครั้งนี้ คณะผู้วิจัยจึงสนใจนำการเรียนรู้ของเครื่อง (machine learning) มาเป็นเครื่องมือในการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียน เนื่องจากเป็นเครื่องมือที่กำลังแพร่หลาย สำหรับการวิเคราะห์ข้อมูลจำนวนมาก สามารถเรียนรู้ได้จากข้อมูลที่ซับซ้อนเพื่อสร้างแบบจำลอง (Artificial Intelligence: AI) เป็นกระบวนการที่ทำให้คอมพิวเตอร์มีความสามารถในการเรียนรู้ หลักการทำงานด้วยตนเอง สามารถเลียนแบบความฉลาดและความสามารถของมนุษย์ ผ่านทางชุดคำสั่งในภาษาคอมพิวเตอร์ หรือ อัลกอริทึม สามารถเรียนรู้และปรับตัวตามข้อมูลที่ได้รับ โดยอาศัยแบบจำลอง (model) ที่สร้างจากชุดข้อมูลที่มีความสัมพันธ์กัน และสามารถนำไปทำนาย หรือใช้ในการตัดสินใจภายหลัง (Thamcharoen, 2019; Shalev-Shwartz & Ben-David, 2014) จัดเป็นเทคนิคและเครื่องมือที่ช่วยในการวินิจฉัย และพยากรณ์ได้ว่า ผู้เรียนเป็นอย่างไร มีความสามารถทางคณิตศาสตร์ระดับใด ควรได้รับการปรับปรุงพัฒนาอย่างไร โดยผู้เรียนจะทราบผลการวินิจฉัยได้ ต้องมีการให้ข้อมูลป้อนกลับ การให้ข้อมูลของครูที่เกี่ยวข้องโดยตรงกับเป้าหมายของการเรียนรู้ ช่วยให้ผู้เรียนได้รับข้อมูลสารสนเทศเกี่ยวกับความคืบหน้าหรือจุดบกพร่องของผู้เรียนต่อเป้าหมายดังกล่าว และแนะนำวิธีการที่จะแก้ไขข้อบกพร่องนั้นเพื่อให้ผู้เรียนสามารถทำได้ใกล้เคียงกับเป้าหมายที่คาดหวังไว้ (Brookhart, 2008) โดยในการวิจัยครั้งนี้ให้ข้อมูลป้อนกลับ 3 ลักษณะที่สำคัญ (Hattie & Timperley, 2007) คือ (1) การให้ข้อมูลกระตุ้นการเรียนรู้ (feed up) (2) การให้ข้อมูลป้อนกลับ (feedback) และ (3) การให้ข้อมูลเพื่อการเรียนรู้ต่อยอด (feed forward)

ดังนั้นคณะผู้วิจัย มีความสนใจในการนำ machine learning มาช่วยในการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 1 สาระที่ 3 สถิติและความน่าจะเป็น โดยต่อยอดจากงานวิจัยของ Junpeng et al. (2020) ซึ่งจะทำการวินิจฉัยระดับความสามารถใน 2 มิติ คือ มิติกระบวนการทางคณิตศาสตร์ และมิติโครงสร้างความคิดรวบยอด โดยใช้เทคนิคการวิเคราะห์ข้อมูลตามโมเดลการตอบสนองข้อสอบแบบพหุมิติ (MRCML) (Adams et al., 1997) และโมเดลเชิงโครงสร้าง (Wilson, 2005) ที่เน้นการตีความอย่างมีความหมาย เพื่อนำไปสู่การวินิจฉัยและปรับปรุงพัฒนานักเรียนได้อย่างชัดเจนมากขึ้น (Junpeng, 2018) และสามารถทำนายความสามารถทางคณิตศาสตร์ของผู้เรียนโดยผ่านการเรียนรู้ของเครื่อง (machine learning) สำหรับตัวแปรที่ใช้ในการทำนาย ตัวแปรอิสระ คือ เกรดเฉลี่ยรวมเทอมที่ผ่านมา เกรดคณิตศาสตร์เทอมที่ผ่านมา จำนวนชั่วโมงศึกษาคณิตศาสตร์ และชั่วโมงการเรียนรู้พิเศษ ผลการทำนายนำไปสู่การพัฒนาวัตกรรมการวัดแบบในการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่าน machine learning สร้างองค์ความรู้ด้านวัดและประเมินผลผู้เรียนโดยใช้เทคโนโลยี นำไปสู่การพัฒนาความสามารถทางคณิตศาสตร์ของผู้เรียนจากระดับหนึ่งไปยังอีกระดับที่สูงขึ้น จะเป็นประโยชน์อย่างยิ่งต่อผู้เรียนหากได้ระบบการตรวจแบบสอบอัตโนมัติ โดยให้ข้อมูลป้อนกลับอัตโนมัติ ที่ผู้เรียนสามารถเข้าถึงการทดสอบและวินิจฉัยระดับความสามารถของผู้เรียนได้ตลอดเวลา เป็นประโยชน์ต่อครูผู้สอน ผู้ปกครอง และผู้ที่เกี่ยวข้อง ในการนำผลการวินิจฉัยความสามารถทางคณิตศาสตร์ ไปสู่การปรับปรุง พัฒนาผู้เรียนให้อยู่ในระดับที่สูงขึ้น นอกจากนี้ยังสามารถวินิจฉัยและตรวจให้คะแนนแบบสอบได้จำนวนหลายคนพร้อมกัน

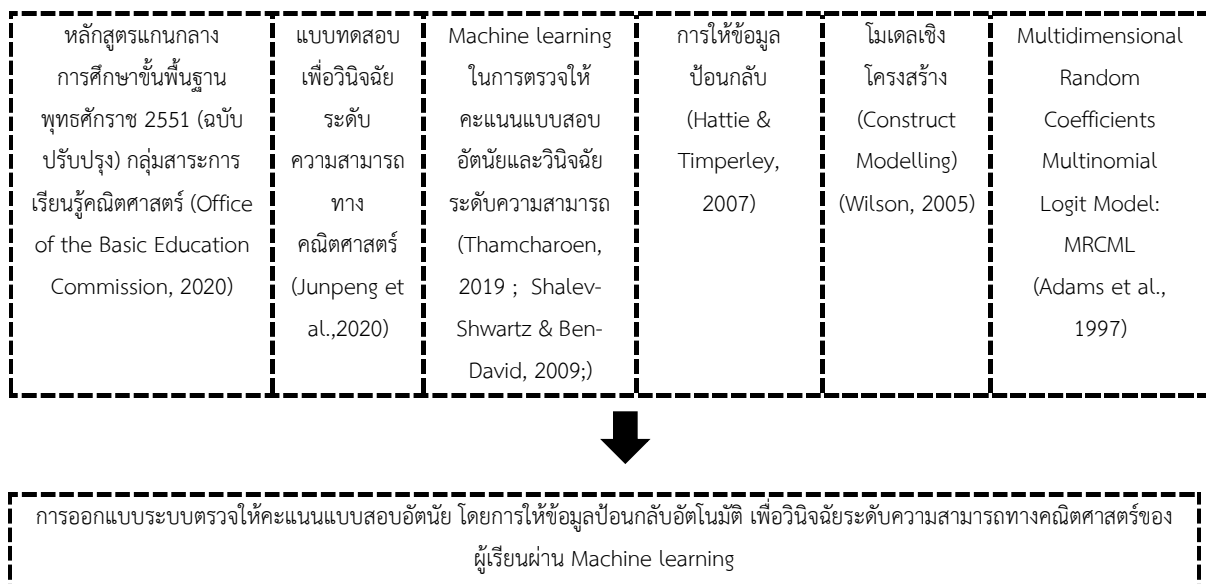
ช่วยลดทรัพยากรบุคคลในการตรวจแบบสอบ และช่วยลดระยะเวลา ขั้นตอน งบประมาณและข้อจำกัดต่าง ๆ ในการตรวจ นำไปใช้ได้ ในสถานศึกษาหลายแห่ง โดยใช้แบบสอบที่เป็นมาตรฐานการตรวจเดียวกัน ได้อย่างแม่นยำและมีประสิทธิภาพต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อวิเคราะห์ผลการตอบของผู้เรียนสู่การออกแบบนวัตกรรมต้นแบบ ประกอบด้วย มโนทัศน์ที่คลาดเคลื่อนของผู้เรียน การกำหนดจุดตัดระดับความสามารถทางคณิตศาสตร์ การให้ข้อมูลป้อนกลับตามมโนทัศน์ที่คลาดเคลื่อนและตัวแปรที่ใช้ในการทำนาย machine learning
2. เพื่อออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ ของผู้เรียนผ่านการเรียนรู้ของ machine learning

กรอบแนวคิดการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง สามารถกำหนดกรอบแนวคิดการวิจัยได้ดังภาพ 1



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ใช้ระเบียบวิธีวิจัยการออกแบบ (design research) (Wongwanit, 2020) โดยแบ่งเป็น 4 ระยะ คือ (1) วิเคราะห์ผลการตอบของผู้เรียนสู่การออกแบบนวัตกรรมต้นแบบ (2) ออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ ของผู้เรียนผ่านการเรียนรู้ของ Machine learning (3) นำนวัตกรรมต้นแบบไปทดลองใช้เพื่อนำข้อมูลที่ได้มาวิเคราะห์ ปรับปรุงพัฒนาระบบให้สมบูรณ์ และรายงานคุณภาพของนวัตกรรม และ (4) เสนอหลักการออกแบบใหม่ในการพัฒนานวัตกรรมต้นแบบ โดยการนำเสนอในบทความครั้งนี้เป็นการศึกษานำร่องซึ่งจะนำเสนอสองระยะแรก มีรายละเอียดทั้งสองระยะดังนี้

ระยะที่ 1 วิเคราะห์ผลการตอบของผู้เรียนเพื่อนำไปสู่การออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ

คณะผู้วิจัยร่วมกับครุคณิตศาสตร์ วิศวกรรมคอมพิวเตอร์ ผู้เชี่ยวชาญในด้านการวัดและประเมินผล และผู้เชี่ยวชาญในด้านการใช้เทคโนโลยี จำนวน 5 คน เพื่อร่วมกันวิเคราะห์คำตอบทั้งหมดที่ผู้เรียนตอบ ข้อคิดเห็นและวิธีคิดต่าง ๆ ของผู้เรียนที่มีต่อเรื่อง สถิติและความน่าจะเป็น

1. กลุ่มผู้สอบ

กลุ่มผู้สอบที่ใช้วิเคราะห์ผลการตอบ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2562 จำนวนนักเรียนทั้งหมด 495 คน จากข้อมูลทุติยภูมิในโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ของโรงเรียนใน 4 ภูมิภาค คือ ภาคเหนือ ภาคตะวันออกเฉียงเหนือ ภาคกลางและภาคใต้ จากทั้งหมด 23 โรงเรียนโดยให้ครอบคลุมทุกระดับความสามารถของผู้เรียนทั้งเก่ง ปานกลาง และอ่อน (Junpeng et al., 2020) เนื่องจากการศึกษาในระยะนี้ใช้โมเดล MRCML (Adams et al., 1997) ในการวิเคราะห์ข้อมูลและการกำหนดจุดตัด จึงให้ความสำคัญกับการกำหนดขนาดตัวอย่างและการได้มาของตัวอย่างที่ครอบคลุมเพียงพอในทุกระดับความสามารถในแต่ละมิติที่ทำการศึกษา

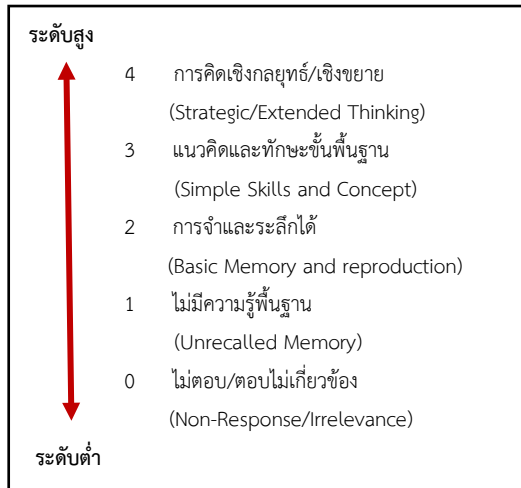
2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นเครื่องมือทุติยภูมิ มาจากโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ ผ่านชุดเครื่องมือวินิจฉัยเรียกว่า “eMAT-Testing” (Junpeng et al., 2020) ซึ่ง “eMAT-Testing” เป็นชุดเครื่องมือในระบบออนไลน์แบบเรียลไทม์ผ่าน Web Application แบบตอบสนองโต้ตอบได้โดยตรง (interactive) รองรับระบบ iOS Android และ Window ซึ่งเป็นแบบทดสอบอัตโนมัติ ตามมาตรฐานและตัวชี้วัดหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง 2560) กลุ่มสาระการเรียนรู้คณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 ในสาระที่ 3 สถิติและความน่าจะเป็น ประกอบด้วย 6 สถานการณ์ 10 ข้อย่อย แบ่งเป็นมิติกระบวนการทางคณิตศาสตร์ จำนวน 5 ข้อ และมิติโครงสร้างความคิดรวบยอด จำนวน 5 ข้อ โดยมีค่าดัชนีความตรงเชิงเนื้อหา รายข้อ (I-CVI) อยู่ระหว่าง 0.50 – 1.00 และ ค่าสัมประสิทธิ์ความเที่ยงแบบ EAP/PV ในมิติที่ 1 และ 2 มีค่าเท่ากับ 0.85 และ 0.83 ตามลำดับ และค่า INFIT MNSQ อยู่ระหว่าง 0.84 - 1.27

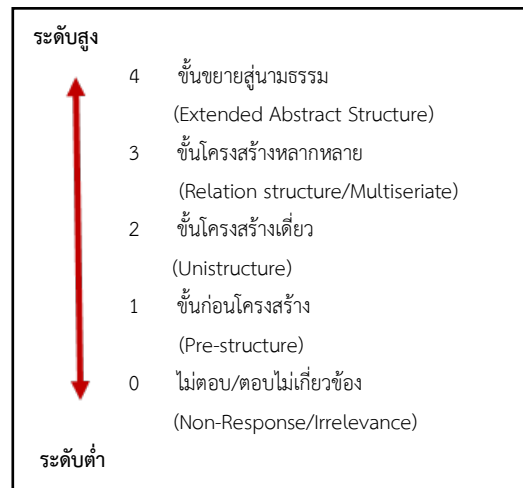
3. การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูล

1. ศึกษาแผนที่โครงสร้าง ระดับความสามารถทางคณิตศาสตร์ของโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ โดยทำการสร้างแผนที่โครงสร้างของระดับผลลัพธ์การเรียนรู้เพื่อใช้ในการประเมินระดับความสามารถทางคณิตศาสตร์ 2 มิติ (Junpeng et al., 2020) คือ (1) ด้านกระบวนการทางคณิตศาสตร์ และ (2) ด้านโครงสร้างความคิดรวบยอด เพื่อใช้เป็นแนวทางในการออกแบบระบบตรวจให้คะแนนอัตโนมัติ ดังภาพ 2

มิติที่ 1 ด้านกระบวนการทางคณิตศาสตร์



มิติที่ 2 ด้านโครงสร้างความคิดรวบยอด



ภาพ 2 construct map ในแต่ละด้านเพื่อประเมินความสามารถทางคณิตศาสตร์

(Junpeng et al., 2020)

2. กำหนดจุดตัดเพื่อนำไปสู่การวินิจฉัยระดับความสามารถทางคณิตศาสตร์ในระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการพิจารณากำหนดพื้นที่บน Wright map แต่ละระดับความสามารถ โดยวิเคราะห์ผลการตอบจากข้อมูลหุติภูมิ โดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ MRCML ด้วยโปรแกรม ACER ConQuest Version 2.0 (Wu et al., 2007) ในการประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบด้วยวิธี MLE พร้อมนำเสนอเป็น Wright map โดย Wright map ทางซ้ายมือนำเสนอค่าความสามารถของผู้เรียนแต่ละข้อ ส่วน Wright map ทางขวามือนำเสนอค่าระดับความสามารถทางคณิตศาสตร์แต่ละขั้นของการตอบ

3. คำนวณจุดตัดร่วมกับการพิจารณาการกำหนดเกณฑ์พื้นที่บน Wright map แต่ละระดับความสามารถทางคณิตศาสตร์ ซึ่งพิจารณาจากค่าเฉลี่ยของความยากหรือระดับ threshold เดียวกัน หาดด้วยจำนวนข้อสอบที่นำมาใช้คำนวณ แยกตามมิติที่ทำการศึกษา และนำคะแนนจุดตัดที่ได้เทียบกับพื้นที่บน Wright map แบ่งเป็นช่วงระดับความสามารถตามที่กำหนดไว้ในโมเดลเชิงโครงสร้างในแต่ละมิติ

4. นำค่าจุดตัดซึ่งเป็นพารามิเตอร์ความสามารถที่เหมาะสมที่ปรากฏใน Wright map และแปลงเป็นคะแนนสเกล (θ_d) ที่เป็นคะแนนมาตรฐานในการวินิจฉัยโดยใช้สูตรดัดสมการคะแนนสเกล $= 50 + 10(\theta_d)$

5. วิเคราะห์ผลการตอบของผู้เรียนทั้งหมด เพื่อหาแนวโน้มที่คลาดเคลื่อนรูปแบบต่าง ๆ และนำไปสู่การออกแบบรูปแบบการให้ข้อมูลป้อนกลับตามแนวโน้มที่คลาดเคลื่อน และวิเคราะห์ผลการตอบของผู้เรียน เพื่อนำรูปแบบคำตอบที่เป็นไปได้ทั้งหมดไปเป็นข้อมูลต้นแบบในการพัฒนาคำถามและแสดงวิธีหาคำตอบในระบบตรวจให้คะแนนอัตโนมัติ ที่สอดคล้องกับระดับความสามารถตามแผนที่เชิงโครงสร้าง และคำอธิบายในแต่ละระดับความสามารถ

6. วิเคราะห์รูปแบบการให้ข้อมูลป้อนกลับเพื่อพัฒนาผู้เรียนตามแนวโน้มที่คลาดเคลื่อน ในการพัฒนาระดับความสามารถทางคณิตศาสตร์ของผู้เรียน

7. พัฒนาคำถามตามแผนที่โครงสร้าง เป็นการออกแบบข้อคำถามแบบอัตนัยเพื่อให้ได้ข้อมูลตามที่บรรยายไว้ในแผนที่โครงสร้าง โดยได้ออกแบบจากการวิเคราะห์ผลการตอบของผู้เรียน และข้อสอบแต่ละข้อเป็นการตรวจให้คะแนนแบบหลายค่า (polytomously scoring)

8. จัดเตรียมข้อมูลพื้นฐานของผู้เรียนเพื่อเข้าสู่การทำนายระดับความสามารถทางคณิตศาสตร์ของผู้เรียน โดย machine learning

ระยะที่ 2 ออกแบบนวัตกรรมต้นแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติผ่าน Machine learning

นำข้อมูลผลการตอบ ที่ได้จากการวิเคราะห์ด้วยการสนทนากลุ่ม (focus group) ระหว่างผู้วิจัยผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา วิศวกรคอมพิวเตอร์ และผู้เชี่ยวชาญด้านเทคโนโลยี ฯ จำนวน 5 คน จากข้อมูลทฤษฎีในโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ (Junpeng et al., 2020) มาออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ โดยใช้แบบสอบอัตโนมัติจากข้อมูลทฤษฎีจำนวน 10 ข้อ แบ่งเป็นมิติกระบวนการทางคณิตศาสตร์ 5 ข้อ และมิติโครงสร้างความคิดรวบยอด 5 ข้อ ก่อนการนำไปเก็บข้อมูลจริง

1. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือคือแบบประเมินการกำหนดมาตรฐานด้วยแบบอิงมาตรฐาน โดยคุณภาพเครื่องมือตามมาตรฐานการประเมินคุณภาพของนวัตกรรมต้นแบบตาม Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2010; Varasunun, 2011) ประกอบด้วย 5 มาตรฐาน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสมขอธรรม ด้านความถูกต้อง และด้านความรับผิดชอบการประเมิน

2. การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูล

1. รวบรวมข้อมูลประกอบด้วย ผลการกำหนดจุดตัดจากระยะที่ 1 ข้อคำถามเพื่อวิเคราะห์มโนทัศน์ที่คลาดเคลื่อนของผู้เรียนรายข้อและภาพรวม การให้ข้อมูลป้อนกลับในแต่ละรูปแบบตามมโนทัศน์ที่คลาดเคลื่อน และชุดข้อมูลปัจจัยนำเข้าในสมการทำนาย

2. ร่วมกันออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ และการให้ข้อมูลป้อนกลับอัตโนมัติตามมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน

3. กำหนดองค์ประกอบของระบบการให้ข้อมูลป้อนกลับอัตโนมัติ แผนผังการทำงานของนวัตกรรมผ่าน machine learning (work flow) พร้อมทั้งหน้าที่ในแต่ละองค์ประกอบ

4. พิจารณาคุณภาพของนวัตกรรมต้นแบบที่พัฒนาขึ้นก่อนนำไปทดลองใช้ โดยผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา ด้านการใช้เทคโนโลยีและครุคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 1 เพื่อพิจารณาความถูกต้อง ความเหมาะสม และความเป็นไปได้ในการนำไปใช้ โดยผลการตรวจสอบคุณภาพของระบบใน 2 ด้าน คือ (1) การประเมินระบบด้วยแบบฮิวริสติก และ (2) การประเมินแบบอิงมาตรฐาน

5. นำผลการประเมินที่ได้มาร่วมกันสะท้อนคิดและปรับปรุงแก้ไขเป็นนวัตกรรมการประเมินต้นแบบก่อนนำไปทดลองใช้ในลำดับถัดไป

ผลการวิจัย

การวิจัยเรื่อง การออกแบบระบบตรวจให้คะแนนแบบสออัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ เพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่านการเรียนรู้ของเครื่อง สรุปผลการวิจัยดังนี้

1. ผลการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนจากผลการตอบของผู้เรียน จุดตัดระดับความสามารถทางคณิตศาสตร์ รูปแบบการให้ข้อมูลป้อนกลับตามมโนทัศน์ที่คลาดเคลื่อนและตัวแปรที่ใช้ในการทำนาย Machine learning

1.1 ผลการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน ในการแก้โจทย์ปัญหาในสถานการณ์เกี่ยวกับเรื่องสถิติ และความน่าจะเป็น ผู้วิจัยได้ทำการวิเคราะห์ข้อสอบรายข้อ เพื่อหารูปแบบมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน ซึ่งเป็นเครื่องมือจากโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียน โดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ (Junpeng et al., 2020) พบว่า

1) มิติกระบวนการทางคณิตศาสตร์ (MAP) นักเรียนมีมโนทัศน์ที่คลาดเคลื่อนมากที่สุดคือ ประเภท ด้านการใช้ข้อมูล ละเลยการใช้ข้อมูลที่จำเป็นในการแก้ปัญหา ทำผิดพลาด หาคำตอบในสิ่งที่ไม่ต้องการ ร้อยละ 25.45 และมโนทัศน์ที่คลาดเคลื่อนน้อยที่สุดคือ ด้านการคำนวณและขาดการตรวจสอบในระหว่างการแก้ปัญหา ร้อยละ 5.58 สามารถแสดงมโนทัศน์ของผู้เรียนรายข้อ ดังแสดงในตาราง 1

ตาราง 1 ผลการวิเคราะห์มโนทัศน์ที่คลาดเคลื่อนมิติกระบวนการทางคณิตศาสตร์

ข้อที่	สถานการณ์	ด้านการใช้ข้อมูล (ร้อยละ)	ด้านการ บิตเบื่อนทฤษฎี บท กฎ สูตร บทนิยาม และ สมบัติ (ร้อยละ)	ด้านการคำนวณ และขาดการ ตรวจสอบใน ระหว่างการ แก้ปัญหา (ร้อยละ)	ผิดพลาดในการตีความ ด้านภาษาและ สัญลักษณ์ทาง คณิตศาสตร์ (ร้อยละ)	ไม่มี มโนทัศน์ที่ คลาดเคลื่อน (ร้อยละ)	ไม่ตอบ (ร้อยละ)
1	1.1	33.33	1.62	3.03	29.70	16.16	21.41
2	1.2	21.01	1.21	1.82	14.55	42.22	23.84
4	3	35.35	9.90	13.54	26.26	1.41	23.84
5	4.1	16.36	12.73	5.86	7.07	28.08	34.14
6	4.2	21.21	13.54	3.64	8.69	19.8	37.98
รวม		25.45	7.80	5.58	17.25	21.54	21.53

2) มิติโครงสร้างความคิดรวบยอด (SLO) นักเรียนมีมโนทัศน์ที่คลาดเคลื่อนมากที่สุด คือ ด้าน การใช้ข้อมูล ละเลยการใช้ข้อมูลที่จำเป็นในการแก้ปัญหา ทำผิดพลาด หาคำตอบในสิ่งที่ไม่ต้องการ ร้อยละ 9.05 และมโนทัศน์ที่คลาดเคลื่อนน้อยที่สุด คือ ด้านผิดพลาดในการตีความด้านภาษาและสัญลักษณ์ทางคณิตศาสตร์ร้อยละ 4.40 สามารถแสดงมโนทัศน์ของผู้เรียนรายข้อ ดังแสดงในตาราง 2

ตาราง 2 ผลการวิเคราะห์หมโนทัศน์ที่คลาดเคลื่อนมิติโครงสร้างความคิดรวบยอด

ข้อที่	สถานการณ์	ด้านการใช้ข้อมูล (ร้อยละ)	ด้าน การบิดเบือน ทฤษฎีบท กฎ สูตร บทนิยาม และสมบัติ (ร้อยละ)	ด้านการคำนวณ และขาด การตรวจสอบ ในระหว่างการ แก้ปัญหา (ร้อยละ)	ผิดพลาดใน การตีความด้าน ภาษาและ สัญลักษณ์ทาง คณิตศาสตร์ (ร้อยละ)	ไม่มี หมโนทัศน์ที่ คลาดเคลื่อน (ร้อยละ)	ไม่ตอบ (ร้อยละ)
3	2	10.71	12.32	3.84	4.65	54.34	17.17
7	5.1	3.64	-	7.68	3.43	73.74	14.95
8	5.2	3.84	-	6.06	1.82	73.54	18.18
9	5.3	7.88	14.55	8.89	8.08	23.43	37.98
10	6	19.19	13.94	5.05	4.04	24.65	34.34
รวม		9.05	8.16	6.30	4.40	49.94	49.93

1.2 การพัฒนาความสามารถทางคณิตศาสตร์ของผู้เรียน ในการออกแบบระบบการให้ข้อมูล ป้อนกลับแบบอัตโนมัติ ตามหมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน พบว่า สามารถกำหนดขอบข่ายของการให้ ข้อมูลป้อนกลับอัตโนมัติ โดยเน้นการให้ข้อมูลป้อนกลับใน 4 รูปแบบ คือ (1) แบบชี้แนะบางส่วน (2) แบบยกตัวอย่าง (3) แบบอธิบายรายละเอียด และ (4) แบบอธิบายรายละเอียดและบอกคำตอบ หลักใน การเลือกใช้แต่ละรูปแบบ ในแต่ละข้ออาจใช้แบบใดแบบหนึ่งหรือหลายแบบบูรณาการรวมกันได้ ขึ้นอยู่กับ หมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน ซึ่งได้จากการวิเคราะห์ผลการตอบผู้เรียน เพื่อนำไปสู่การกำหนดข้อความ ป้อนกลับอัตโนมัติในระบบ ตัวอย่างดังปรากฏในตาราง 3

ตาราง 3 ตัวอย่างการกำหนดขอบข่ายของการให้ข้อมูลป้อนกลับเพื่อนำไปสร้างข้อความอัตโนมัติในระบบ

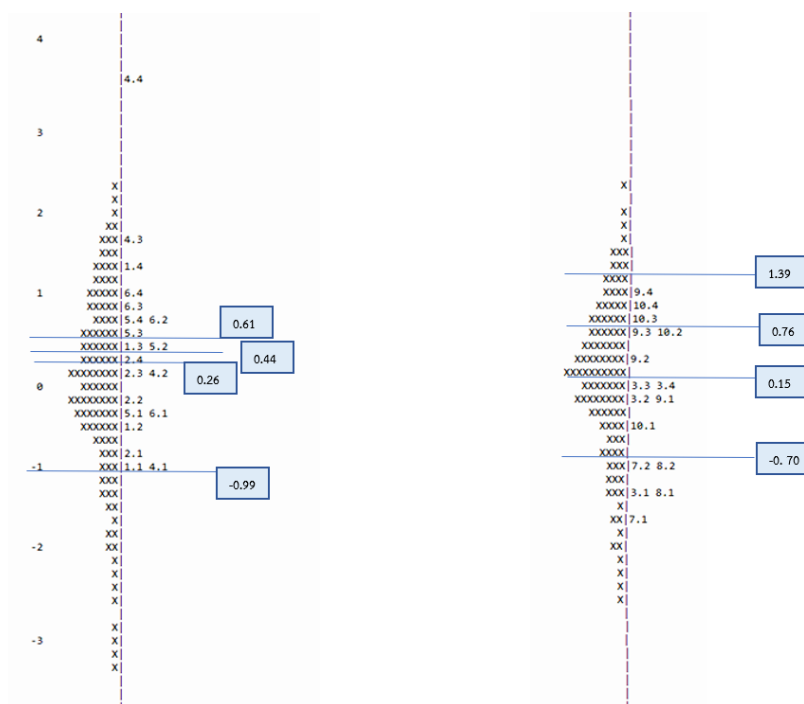
รูปแบบหมโนทัศน์ที่ คลาดเคลื่อน	รูปแบบการให้ ข้อมูลป้อนกลับ	ข้อความป้อนกลับที่ปรากฏในระบบ
ตัวอย่างแบบสอบอัตนัย สถานการณ์ที่ 6 : ครูนำตัวอักษรจากคำว่า FATHER มาทำเป็นบัตรคำให้นักเรียนสุ่มหยิบขึ้นมา 1 ครั้ง นักเรียนจะมีโอกาสหยิบได้บัตรที่เป็นสระของภาษาอังกฤษ คิดเป็นเท่าไร?		
ด้านการใช้ข้อมูล	แบบชี้แนะบางส่วน	แนวคิด สระในภาษาอังกฤษ คือ A E I O U ดังนั้นคำว่า FATHER มีสระ _ ตัว
ผิดพลาดใน การตีความด้านภาษา และสัญลักษณ์ทาง คณิตศาสตร์	แบบอธิบาย รายละเอียด	แนวคิด จากคำว่า FATHER มีอักษรทั้งหมด _ ตัว และมีสระทั้งหมด _ ตัว ดังนั้น โอกาสหรือความน่าจะเป็น เท่ากับ อัตราส่วนของจำนวนเหตุการณ์ที่เราสนใจ ต่อ จำนวนผลลัพธ์ทั้งหมดที่อาจเกิดขึ้นได้ $P(E) = \frac{n}{N}$
ด้านการคำนวณและ ขาดการตรวจสอบใน ระหว่างการแก้ปัญหา	ยกตัวอย่าง	แนวคิด ยกตัวอย่าง คำว่า MONDAY มีอักษรที่เป็นสระ 2 ตัว คือ O และ A จากสระทั้งหมด 6 ตัว ดังนั้น โอกาสหยิบได้บัตรที่เป็นสระของภาษาอังกฤษ คิดเป็น $\frac{2}{6}$
ด้านการบิดเบือน ทฤษฎีบท กฎ สูตร บทนิยาม และสมบัติ	แบบอธิบาย รายละเอียดและ ยกตัวอย่าง	แนวคิด จากคำว่า FATHER มีอักษรทั้งหมด _ ตัว และมีสระทั้งหมด _ ตัว โอกาสหรือความน่าจะเป็น เท่ากับ อัตราส่วนของจำนวนเหตุการณ์ที่เราสนใจ(n) ต่อ จำนวน ผลลัพธ์ทั้งหมดที่อาจเกิดขึ้นได้(N) ซึ่งมีสูตรในการคิดคำนวณดังนี้ $P(E) = \frac{n}{N}$ ยกตัวอย่าง คำว่า MONDAY มีอักษร 6 ตัว มีสระ 2 ตัว คือ O และ A ดังนั้น โอกาสหยิบได้บัตรที่เป็นสระของภาษาอังกฤษ คิดเป็น $\frac{2}{6}$

1.3 ผลการคำนวณจุดตัดร่วมกับการพิจารณาการกำหนดเกณฑ์พื้นที่บน Wright map แต่ละระดับความสามารถทางคณิตศาสตร์ ซึ่งพิจารณาจากค่าเฉลี่ยของความยากหรือระดับ threshold เดียวกัน หารด้วยจำนวนข้อสอบที่นำมาใช้คำนวณ แยกตามมิติที่ทำการศึกษา ดังตาราง 4

ตาราง 4 ค่าความยากและค่า threshold ของข้อคำถาม ในแต่ละมิติเพื่อนำไปสู่การคำนวณจุดตัด

มิติกระบวนการทางคณิตศาสตร์						มิติโครงสร้างความคิดรวบยอด							
ข้อที่	ค่า	ระดับชั้นความยาก (threshold)				ข้อที่	ค่า	ระดับชั้นความยาก (threshold)					
		ความ ยาก	1	2	3			4	ความ ยาก	1	2	3	4
3	-0.497	-1.34	-0.27	-0.07	-0.07	1	0.133	-0.97	-0.46	0.48	1.44		
7	-1.281	-1.55	-1.02			2	-0.176	-0.80	-0.12	0.11	0.26		
8	-1.127	-1.30	-0.95			4	1.096	-0.99	0.10	1.69	3.58		
9	0.397	-0.27	0.29	0.62	0.99	5	0.293	-0.43	0.49	0.64	0.69		
10	0.425	-0.50	0.66	0.78	0.91	6	0.533	-0.32	0.75	0.89	0.99		
Mean		-0.99	0.26	0.44	0.61	Mean		-0.70	0.15	0.76	1.39		

ผลการกำหนดจุดตัดเพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ร่วมกับการกำหนดเกณฑ์พื้นที่บน Wright map พบว่า มิติกระบวนการทางคณิตศาสตร์ แบ่งได้ 5 ระดับ 4 จุดตัด จากระดับต่ำสุดไปสูงสุด ที่ -0.99 0.26 0.44 และ 0.61 ตามลำดับ และมิติโครงสร้างความคิดรวบยอด แบ่งได้ 5 ระดับ 4 จุดตัด จากระดับต่ำสุดไปสูงสุดที่ -0.70 0.15 0.76 และ 1.39 ตามลำดับ ปรากฏดังภาพ 3



ภาพ 3 การกำหนดจุดตัดในมิติด้านกระบวนการทางคณิตศาสตร์(ซ้าย) และมิติด้านโครงสร้างความคิด รวบยอด (ขวา)

จาก 4 จุดตัดในแต่ละมิติดังกล่าวข้างต้น นำไปสู่การกำหนดช่วงการประเมินระดับความสามารถของผู้เรียนในมิติกระบวนการทางคณิตศาสตร์ และ มิติโครงสร้างทางความคิดรวบยอด เป็น 5 ช่วง ตามแผนที่โครงสร้าง โดยมีคะแนนจุดตัดแบ่งระดับความสามารถของผู้เรียน ออกเป็น 5 ระดับ ที่สามารถแปลงเป็นช่วงคะแนนสเกล และช่วงคะแนนดิบ เพื่อนำไปใช้ในการวินิจฉัยผู้เรียนเป็นรายบุคคล โดยแสดงรายละเอียดดังตาราง 5 และตาราง 6

ตาราง 5 เกณฑ์จุดตัดระดับความสามารถทางคณิตศาสตร์ในมิติกระบวนการทางคณิตศาสตร์ (MAP)

จุดตัด	มิติกระบวนการทางคณิตศาสตร์	ช่วงระดับความสามารถ (θ)	คะแนนสเกล	ช่วงคะแนนดิบ (คะแนนเต็ม = 17 คะแนน)
$\theta = 0.61$	ระดับ 5	$\theta = 0.61$ ขึ้นไป	56.10 ขึ้นไป	14 ขึ้นไป
$\theta = 0.44$	ระดับ 4	$\theta = 0.44$ ถึง 0.60	54.4 – 56	9-13
$\theta = 0.26$	ระดับ 3	$\theta = 0.26$ ถึง 0.43	52.60 – 54.30	4-8
$\theta = -0.99$	ระดับ 2	$\theta = -0.99$ ถึง 0.25	40.10 – 52.50	2-3
	ระดับ 1	ต่ำกว่า $\theta = -0.99$	ต่ำกว่า 40.10	0-1

จากตาราง 5 มิติกระบวนการทางคณิตศาสตร์ (MAP) พบว่า มีจุดตัด 4 จุดตัด นำไปสู่การกำหนดช่วงการประเมินระดับความสามารถทางคณิตศาสตร์ในแต่ละมิติ เป็น 5 ช่วง ที่สามารถแปลงเป็นช่วงคะแนนสเกล และช่วงคะแนนดิบ เพื่อนำไปใช้ในการวินิจฉัยผู้เรียนเป็นรายบุคคล เช่น ผู้สอบที่ประมาณค่าพารามิเตอร์ระดับความสามารถด้านกระบวนการทางคณิตศาสตร์ตั้งแต่ 0.61 (θ) ≥ 0.61 แสดงว่า ผู้สอบจะได้รับการวินิจฉัยระดับความสามารถทางคณิตศาสตร์อยู่ในระดับที่ 5 โดยมีคะแนนสเกลเท่ากับ 56.10 ขึ้นไป และเมื่อนำไปเทียบกับคะแนนดิบจะมีค่าเท่ากับ 14 คะแนนขึ้นไป จากคะแนนเต็ม 17 คะแนน

ตาราง 6 เกณฑ์จุดตัดระดับความสามารถทางคณิตศาสตร์ในมิติโครงสร้างความคิดรวบยอด (SLO)

จุดตัด	มิติโครงสร้างความคิดรวบยอด	ช่วงระดับความสามารถ (θ)	คะแนนสเกล	ช่วงคะแนนดิบ (คะแนนเต็ม = 21 คะแนน)
$\theta = 1.39$	ระดับ 5	$\theta = 1.39$ ขึ้นไป	63.90 ขึ้นไป	14 ขึ้นไป
$\theta = 0.76$	ระดับ 4	$\theta = 0.76$ ถึง 1.38	57.6 – 63.80	10-13
$\theta = 0.15$	ระดับ 3	$\theta = 0.15$ ถึง 0.75	51.50 – 57.50	4-9
$\theta = -0.70$	ระดับ 2	$\theta = -0.70$ ถึง 0.14	43 – 51.40	2-3
	ระดับ 1	ต่ำกว่า $\theta = -0.70$	ต่ำกว่า 43	0-1

จากตาราง 6 พบว่า มิติโครงสร้างความคิดรวบยอด มีจุดตัด 4 จุดตัด นำไปสู่การกำหนดช่วงการประเมินระดับความสามารถทางคณิตศาสตร์ในแต่ละมิติ เป็น 5 ช่วง ที่สามารถแปลงเป็นช่วงคะแนนสเกล และช่วงคะแนนดิบ เพื่อนำไปใช้ในการวินิจฉัยผู้เรียนเป็นรายบุคคล เช่น ผู้สอบที่ประมาณค่าพารามิเตอร์ระดับความสามารถด้านกระบวนการทางคณิตศาสตร์ตั้งแต่ 1.39 (θ) ≥ 1.39 แสดงว่า ผู้สอบจะได้รับการ

การวินิจฉัยระดับความสามารถทางคณิตศาสตร์อยู่ในระดับที่ 5 โดยมีคะแนนสเกลเท่ากับ 63.90 ขึ้นไป และเมื่อนำไปเทียบกับคะแนนดิบจะมีค่าเท่ากับ 14 คะแนนขึ้นไป จากคะแนนเต็ม 21 คะแนน

1.4 ผลการกำหนดแนวทางการให้คะแนนแบบอัตโนมัติในระบบการตรวจ โดยใช้แผนที่โครงสร้างของระดับผลลัพธ์การเรียนรู้ เป็นเกณฑ์ในการตรวจให้คะแนนผู้เรียนในระบบรายข้อ และนำเสนอภาพรวมประกอบด้วย 2 มิติ คือ (1) กระบวนการทางคณิตศาสตร์ และ (2) โครงสร้างความคิดรวบยอด ในแต่ละมิติจะแบ่งระดับคุณลักษณะแบ่งใน 5 ระดับ และมีความเข้มข้นของการคิดเชิงคณิตศาสตร์ลดหลั่นกันจากน้อยไปมากตั้งแต่ 0 - 4 ดังตาราง 7

ตาราง 7 ตัวอย่างการตรวจให้คะแนนแบบหลายค่าตามแผนที่โครงสร้าง

ระดับ	คะแนน	ชื่อ	คำอธิบายระดับความสามารถ
เกณฑ์การให้คะแนนและตัวอย่างคำตอบของข้อคำถามกระบวนการทางคณิตศาสตร์			
ดีมาก	4	การคิดเชิงกลยุทธ์/เชิงขยาย (Strategic/Extended Thinking)	สามารถแสดงการแก้ปัญหาได้อย่างหลากหลายและสามารถสรุปเป็นวิธีการที่สามารถใช้ในระดับที่สูงขึ้นได้ พร้อมแสดงที่มาของคำตอบได้ชัดเจนโดยมีกระบวนการคิดที่เห็นเชิงประจักษ์ เป็นขั้นตอนที่ครบถ้วน
ดี	3	แนวคิดและทักษะขั้นพื้นฐาน (Simple Skills and Concept)	สามารถใช้ความเข้าใจในกระบวนการ หลักการ ทฤษฎี ทางคณิตศาสตร์มาแก้ปัญหาได้อย่างหลากหลาย สามารถแก้ไขปัญหาที่ซับซ้อนได้อย่างมาก 1 ขั้นตอนได้ พร้อมแสดงที่มาของคำตอบได้ชัดเจนโดยมีกระบวนการคิดที่เห็นเชิงประจักษ์ เป็นขั้นตอนบางส่วน
พอใช้	2	การจำและระลึกได้ (Basic Memory and Reproduction)	สามารถใช้ความเข้าใจในกระบวนการพื้นฐานทางคณิตศาสตร์ อย่างง่ายมาประกอบการหาคำตอบ ซึ่งทำให้ได้คำตอบที่ยังคงไม่สมบูรณ์ ขาดการประยุกต์ใช้
ปรับปรุง	1	ไม่มีความรู้พื้นฐาน (Unrecalled Memory)	การตอบคำถามเป็นการนำความรู้บางส่วนมาตอบแต่ผลการตอบสะท้อนให้เห็นว่าผู้เรียนไม่สามารถจดจำเนื้อหาที่สำคัญหรือจำเป็นมาเป็นพื้นฐานในการแก้ปัญหาได้
ปรับปรุง เร่งด่วน	0	ไม่ตอบ (Non- Response/ Irrelevance)	ไม่ตอบหรือตอบไม่เกี่ยวข้องกับคำถาม
เกณฑ์การให้คะแนนและตัวอย่างคำตอบของข้อคำถามทักษะด้านความคิดรวบยอด			
ดีมาก	4	ขั้นขยายสู่นามธรรม (Extended abstract structure)	- นำความคิดรวบยอดเชื่อมโยงและสรุปอ้างอิง - เป็นองค์ความรู้เชิงนามธรรมได้ - สร้างข้อสรุปจากความคิดรวบยอดได้
ดี	3	ขั้นความสัมพันธ์เชิงโครงสร้าง ความคิดรวบยอด (Relation structure)	- เชื่อมโยงความคิดรวบยอดที่ซับซ้อนได้ - จำแนก/ อธิบาย/ บรรยายแต่ละส่วนเข้าด้วยกันได้ - แต่ไม่สามารถสรุปเป็นองค์ความรู้เชิงนามธรรมได้
พอใช้	2	ขั้นโครงสร้างเดี่ยว (Unistructure)	- เชื่อมโยงความคิดรวบยอดอย่างง่ายได้ คิดรวบยอดซับซ้อนไม่ได้ - ระบุสิ่งที่เรียนรู้ในลักษณะของความจำ
ปรับปรุง	1	ขั้นก่อนโครงสร้าง (pre-structure)	- ไม่สามารถสร้างความคิดรวบยอด - ไม่สามารถเชื่อมโยงความคิดรวบยอด - มีความเข้าใจคลาดเคลื่อนในการคิดรวบยอดในเรื่องที่เรียน
ปรับปรุง เร่งด่วน	0	ไม่ตอบ (Non- Response/ Irrelevance)	ไม่ตอบหรือตอบไม่เกี่ยวข้องกับคำถาม

1.5 ผลการวิเคราะห์ตัวแปรที่จะใช้ในการสร้างสมการทำนายจากข้อมูลทฤษฎี

ข้อมูลพื้นฐานของผู้เรียนเพื่อใช้ในการทำนายความสามารถทางคณิตศาสตร์ของผู้เรียน โดย machine learning ได้จากข้อมูลทฤษฎี (Junpeng et al., 2020) นำมาพิจารณาว่าตัวแปรใดบ้างที่ส่งผลต่อระดับความสามารถทางคณิตศาสตร์ของผู้เรียนในแต่ละมิติ ดังแสดงในตาราง 8

ตาราง 8 ตารางแสดงตัวแปรอิสระและตัวแปรตามเพื่อนำไปใช้สร้างสมการทำนาย

ประเภทตัวแปร	ข้อมูลทฤษฎี	ตัวแปร
ตัวแปรอิสระ	เกรดคณิตเทอมที่ผ่านมา	X_1
	เกรดเฉลี่ยรวม	X_2
	ชั่วโมงเรียนพิเศษ	X_3
	ชั่วโมงศึกษาคณิตศาสตร์ด้วยตนเอง	X_4
ตัวแปรตาม	รวมคะแนนดิบมิติกระบวนการคณิตศาสตร์	Y_1
	รวมคะแนนดิบมิติโครงสร้างความคิดรวบยอด	Y_2
	คะแนนรวม 2 มิติ	Y_3

จากตาราง 8 พบว่าตัวแปรต้นที่จะส่งผลต่อการสร้างสมการทำนาย ได้แก่ เกรดคณิตเทอมที่ผ่านมา เกรดเฉลี่ยรวมเทอมที่ผ่านมา ชั่วโมงเรียนพิเศษและชั่วโมงศึกษาคณิตศาสตร์ด้วยตนเอง และตัวแปรตามประกอบด้วยคะแนนดิบมิติกระบวนการคณิตศาสตร์ คะแนนดิบมิติโครงสร้างความคิดรวบยอด และคะแนนรวม 2 มิติ

ผลการวิเคราะห์หมโนทัศน์ที่คลาดเคลื่อน การออกแบบรูปแบบการให้ข้อมูลป้อนกลับตามหมโนทัศน์ที่คลาดเคลื่อน ผลการกำหนดจุดตัด และตัวแปรที่ใช้ในการทำนาย ข้อมูลทั้งหมดนี้เป็นข้อมูลป้อนเข้าในการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติผ่าน machine learning

2. ผลการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติเพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่านการเรียนรู้ของ Machine learning

จากการวิเคราะห์ผลการตอบ รูปแบบการให้ข้อมูลป้อนกลับและผลการกำหนดจุดตัดในระยะแรก นำมาสู่การออกแบบนวัตกรรมต้นแบบผ่าน Machine learning เพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์และหมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนทั้ง 2 มิติ โดยระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ ประกอบด้วย 5 ส่วน คือ (1) ส่วนรับข้อมูล (2) ส่วนประมวลผล (3) ส่วนแสดงผล (4) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ และ (5) ส่วนรายงานผลการประเมิน ดังภาพ 4



ภาพ 4 แสดงการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติผ่าน Machine learning

รายละเอียดขั้นตอนของการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ ผ่าน machine learning ดังนี้

(1) ส่วนรับข้อมูล เป็นการนำเข้าสู่ข้อมูล เพื่อใช้วินิจฉัยโมเดลที่คลาดเคลื่อนทางคณิตศาสตร์ และรายงานระดับความสามารถทางคณิตศาสตร์ในแต่ละมิติ ประกอบด้วย ผลการกำหนดจุดตัด คะแนนสเกลและคะแนนดิบ ข้อคำถามเพื่อใช้ในการวิเคราะห์โมเดลที่คลาดเคลื่อนของผู้เรียน รูปแบบคำตอบที่ถูกต้องที่เป็นไปได้ทั้งหมดของแบบสอบอัตโนมัติ ข้อมูลป้อนกลับทั้ง 4 รูปแบบในรายข้อกรณีตอบไม่ถูกต้อง คำอธิบายระดับความสามารถในแต่ละมิติ และตัวแปรเพื่อใช้ในการทำนายความสามารถทางคณิตศาสตร์ผ่าน machine learning เช่น เกรดเฉลี่ยวิชาคณิตศาสตร์ การเรียนพิเศษ ชั่วโมงการทบทวนความรู้ด้วยตนเอง และต้องอาศัยข้อมูลของโรงเรียนที่ใช้ทำการทดสอบ ประกอบด้วย รหัสโรงเรียน รายชื่อโรงเรียน รหัสประจำตัว เพื่อให้ผู้เรียน login username และ password ก่อนดำเนินการสอบ และเมื่อผู้เรียนดำเนินการสอบถือเป็นขั้นตอนสำคัญในการแสดงวิธีหาคำตอบ ลงในสถานการณ์ปลายเปิดที่เป็นแบบทดสอบอัตโนมัติที่กำหนดไว้

(2) ส่วนประมวลผล เป็นการตรวจให้คะแนนผู้เรียนรายข้อผ่าน machine learning หากผู้เรียนตอบผิดหรือไม่ตอบ จะมีการให้ข้อมูลป้อนกลับในแต่ละส่วน ตามโมเดลที่คลาดเคลื่อนที่กำหนดไว้ในข้อนั้น ๆ เมื่อผู้เรียนทำแบบทดสอบครบ 10 ข้อ ระบบจะตรวจให้คะแนนประมวลผลความสามารถทางคณิตศาสตร์ของผู้เรียนใน 2 มิติ คือมิติกระบวนการทางคณิตศาสตร์ และมิติโครงสร้างความคิดรวบยอด และรายงานความสามารถในแต่ละมิติ โดยแบ่งระดับความสามารถในแต่ละมิติจากจุดตัดที่กำหนดไว้ และระบบสามารถประมวลผลและทำนายระดับความสามารถทางคณิตศาสตร์ของผู้เรียนก่อนเริ่มทำการทดสอบ โดยใช้ machine learning จากข้อมูลส่วนตัว เช่น เพศ อายุ เกรดเฉลี่ย การเรียนพิเศษ ชั่วโมงการทบทวนความรู้ และอื่นๆ จากสมการทำนาย

(3) ส่วนแสดงผล เป็นการให้ข้อมูลที่ได้รับจากส่วนนำเข้ามาสู่กระบวนการประมวลผล โดยเน้นไปที่กระบวนการงานของโปรแกรม (algorithm) เช่น การแสดงผลการตรวจคำตอบแต่ละข้อ และแสดงข้อมูลป้อนกลับแก่ผู้เรียนกรณีตอบผิด และในขั้นตอนสุดท้ายจะแสดงผลระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยระบบจะทำการรวมคะแนนดิบของข้อสอบทุกข้อในแต่ละมิติ ไปเทียบกับคะแนนดิบที่กำหนดไว้ จากนั้นระบบจะนำไปประมวลผลเทียบกับคะแนนสเกล และนำไปสู่การวินิจฉัยระดับความสามารถทางคณิตศาสตร์ ในมิตินั้น ๆ ว่าอยู่ในระดับใด โดยจะแสดงคำอธิบายระดับความสามารถที่กำหนดไว้ในแผนที่โครงสร้าง ซึ่งถือเป็นสิ่งที่ผู้เรียนทำได้ สิ่งที่คุณเรียนควรพัฒนาต่อ

(4) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ ระบบจะให้ข้อมูลป้อนกลับรายข้อ โดยระบบจะแสดงข้อความคำอธิบายเพิ่มเติมในแต่ละข้อเมื่อผู้เรียนไม่ตอบหรือตอบผิด ตามโมเดลที่คลาดเคลื่อนแต่ละรูปแบบ เพื่อเป็นแนวทางในการตอบคำถามเพื่อพัฒนาความสามารถทางคณิตศาสตร์

(5) ส่วนรายงานผลการประเมิน เป็นการรายงานผลการทดสอบทั้งหมดเพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ เป็นการนำผลจากการประมวลผลมารายงานให้ผู้ใช้งานเข้าใจได้ง่ายมากยิ่งขึ้น เป็นการรายงานผลรายบุคคล (individual report) การรายงานในส่วนนี้จะเป็นการรายงานผลจากการประเมินและจากการวิเคราะห์ของระบบ โดยจะแสดงผลเป็นรายบุคคล สำหรับผู้เรียนในการรายงาน

จะประกอบด้วย 1) ข้อมูลส่วนบุคคล 2) ข้อมูลสารสนเทศและสถิติของการประเมิน 3) ข้อมูลจากการประเมินผลระดับความสามารถทางคณิตศาสตร์

การประเมินคุณภาพของผลการออกแบบระบบในครั้งนี้ได้ผ่านการประเมินโดยผู้เชี่ยวชาญ จำนวน 5 คน ด้วยวิธีการสนทนากลุ่ม (focus group) ครอบคลุม 5 มาตรฐาน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสมขอรรถธรรม ด้านความถูกต้อง และด้านความรับผิดชอบการประเมิน โดยพบว่ากระบวนการออกแบบสะดวกต่อการใช้งาน มีความเหมาะสม และง่ายต่อการเรียนรู้ มีความเป็นขั้นตอนในการนำเข้า ทำให้เข้าใจง่าย รวมถึงสามารถนำไปใช้ได้จริง ซึ่งจะเป็นประโยชน์อย่างมากต่อผู้เรียน หากมีการพัฒนาให้ครบวงจรได้ โดยไม่มีกระบวนการเตรียมข้อมูล ระบบจะมีความสมบูรณ์มากยิ่งขึ้น

อภิปรายผล

1. จากผลการกำหนดจุดตัดพบว่า มีความเหมาะสมที่สามารถวินิจฉัยความสามารถทางคณิตศาสตร์ได้ ดังจะเห็นได้จากความสามารถของผู้เรียนในแต่ละมิติ มีการกระจายความสามารถตั้งแต่ -4 ถึง 4 ซึ่งสอดคล้องกับแผนที่โครงสร้างที่ได้กำหนดไว้ เนื่องจากผลการตอบที่ใช้ในการวิเคราะห์ข้อมูล เป็นกลุ่มตัวอย่างที่ครอบคลุมทุกระดับความสามารถ จาก 4 ภูมิภาคของประเทศไทย สอดคล้องกับแนวคิดของ Demars (2010) ที่ให้ความสำคัญกับการกำหนดตัวอย่างและการได้มาของตัวอย่างในการกำหนดจุดตัดที่มีคุณลักษณะของระดับความสามารถที่ครอบคลุม เพียงพอในทุกระดับความสามารถในแต่ละมิติที่ทำการศึกษา มากกว่ามุ่งเน้นการสุ่มตัวอย่างให้มีความเป็นตัวแทนของประชากรเหมือนการออกแบบการวิจัยที่มุ่งเน้นการสรุปอ้างอิงนัยทั่วไปจากกลุ่มผู้สอบไปยังประชากร (Demars, 2010) แต่เมื่อพิจารณาในมิติกระบวนการทางคณิตศาสตร์ มีจุดตัดที่ใกล้เคียงกันอย่างเห็นได้ชัด โดยมีจุดตัดจากระดับที่ 2 ไประดับที่ 3 ระดับที่ 3 ไประดับที่ 4 ซึ่งจุดตัดที่ใกล้กันจะส่งผลต่อเกณฑ์การให้คะแนน โดยการกำหนดค่าอธิบายระดับความสามารถให้มีความแตกต่างและชัดเจนมากขึ้นกว่าเดิม เพื่อให้ข้อสอบที่พัฒนาขึ้นแสดงให้เห็นถึงมวลเนื้อหาหรือสิ่งที่มุ่งวัดในแต่ละชุดครอบคลุมผู้สอบหรือไม่ หากครอบคลุมจะสะท้อนให้เห็นถึงหลักฐานความตรงเชิงเนื้อหาของรายการประเมิน (Wilson, 2005)

2. ผลการวิเคราะห์หม็อตศน์ที่คลาดเคลื่อน ในการแก้โจทย์ปัญหาในสถานการณ์เกี่ยวกับเรื่องสถิติ และความน่าจะเป็น พบว่า ทั้งมิติกระบวนการทางคณิตศาสตร์ (MAP) และมิติโครงสร้างความคิดรวบยอดนักเรียนมีหม็อตศน์ที่คลาดเคลื่อนมากที่สุด คือ ด้านการใช้ข้อมูล ละเลยการใช้ข้อมูลที่จำเป็นในการแก้ปัญหา ทำผิดพลาด หาคำตอบในสิ่งที่ไม่ต้องการ อาจเนื่องมาจากผู้เรียนเลือกใช้หน่วยของตัวแปรผิด หรือการลอกรายละเอียดเกี่ยวกับโจทย์ผิด โดยข้อผิดพลาดอาจเกิดขึ้นตั้งแต่เริ่มต้น หรืออาจเกิดขึ้นในช่วงระหว่าง การดำเนินการกับข้อมูล สอดคล้องกับผลการวิจัยของ Sukkrom (2015) ที่พบว่า ข้อผิดพลาดนักเรียนที่ได้จากการทำแบบวัดหม็อตศน์และข้อผิดพลาดทางคณิตศาสตร์ แบบอัตนัย นักเรียนมีข้อผิดพลาดมากที่สุด คือ การใช้ข้อมูลผิด ไม่ทำตามทีโจทย์ระบุอย่างชัดเจน ใช้ข้อมูลอื่นที่ไม่สอดคล้องต่อการใช้แก้ปัญหา ซึ่งผลจากการวิเคราะห์หม็อตศน์ที่คลาดเคลื่อน จะใช้เป็นข้อมูลในการสะท้อนหม็อตศน์ที่คลาดเคลื่อนของผู้เรียนในระบบ การตรวจให้คะแนน โดยพิจารณาจากการตอบคำถาม ในข้อคำถามที่ผู้วิจัยได้ออกแบบเพื่อวินิจฉัยหม็อตศน์

ที่คลาดเคลื่อนไว้แล้ว ซึ่งการละเลยการใช้ข้อมูลที่จำเป็นในการแก้ปัญหา จะส่งผลกระทบต่อคะแนนเป็นอย่างมาก เพราะถ้าหากใช้ข้อมูลผิด จะส่งผลต่อการหาคำตอบที่ผิดในขั้นต่อไป

3. การให้ข้อมูลป้อนกลับ ถือเป็นรูปแบบกระบวนการหรือวิธีการตรวจสอบผู้เรียนโดยอาศัยทั้งข้อมูลและสารสนเทศ เพื่อทำการประเมินการเรียนรู้ผู้เรียน ทั้งที่เกิดขึ้นทั้งภายในและภายนอกเป็นรายบุคคล เพื่อที่จะนำไปสู่การปรับเปลี่ยน วางแผนกระบวนการจัดการเรียนรู้ที่เหมาะสมกับผู้เรียนแต่ละบุคคล (Fisher & Frey, 2011) ซึ่งในระบบตรวจให้คะแนนแบบสอบอัตนัยครั้งนี้ เน้นการให้ข้อมูลป้อนกลับใน 4 รูปแบบ คือ (1) แบบชี้แนะบางส่วน (2) ยกตัวอย่าง (3) แบบอธิบายรายละเอียด และ (4) แบบอธิบายรายละเอียดและยกตัวอย่าง ในแต่ละข้ออาจใช้แบบใดแบบหนึ่งหรือหลายแบบบูรณาการรวมกันได้ขึ้นอยู่กับมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนในแต่ละข้อ ซึ่งแต่ละรูปแบบได้ให้ข้อมูลที่จำเป็น บอกรายละเอียดในแก้โจทย์ปัญหา พร้อมยกตัวอย่างประกอบใกล้เคียงกับโจทย์ที่ผู้เรียนกำลังเผชิญอยู่ เพื่อกระตุ้นความคิดของผู้เรียน มากกว่าการบอกว่าถูกหรือผิดเพียงอย่างเดียว จึงเป็นประโยชน์ต่อการนำไปใช้พัฒนาระดับความสามารถทางคณิตศาสตร์ให้สูงขึ้น ซึ่งสอดคล้องกับผลการวิเคราะห์ห่อภิมาณของ Van Der Kleij et al. (2015) ที่พบว่า ข้อมูลย้อนกลับแบบอธิบายรายละเอียดมีขนาดอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนมากกว่าข้อมูลย้อนกลับแบบบอกความถูกต้อง และข้อมูลย้อนกลับแบบบอกผลการกระทำ ซึ่งการให้ข้อมูลป้อนกลับช่วยกระตุ้นให้นักเรียนได้คิดแก้ปัญหา สามารถเชื่อมโยงความรู้เดิมเข้ามาปรับใช้กับสิ่งที่ตนเองยังไม่ทราบหรือยังมีมโนทัศน์ที่ไม่แม่นยำ

4. ผลการออกแบบระบบตรวจให้คะแนนแบบสอบอัตนัย โดยการให้ข้อมูลป้อนกลับอัตโนมัติ เพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่านการเรียนรู้ของ machine learning ประกอบด้วย 5 ส่วน คือ (1) ส่วนรับข้อมูล (2) ส่วนประมวลผล (3) ส่วนแสดงผล (4) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ และ (5) ส่วนรายงานผลการประเมิน ซึ่งทุกส่วนล้วนมีความสำคัญในการตรวจให้คะแนนผู้เรียนให้ถูกต้องและแม่นยำ แต่ควรมุ่งเน้นไปที่ส่วนการรับข้อมูลซึ่งเป็นส่วนแรกของการทำงานของระบบ เช่น หากข้อมูลในส่วนจุดตัดไม่ถูกต้อง ก็จะส่งผลต่อการอธิบายความสามารถของผู้เรียนในแต่ละมิติได้ไม่ถูกต้อง หากคำตอบที่นำเข้ามาไม่ครอบคลุมคำตอบที่เป็นไปได้ทั้งหมดที่ผู้เรียนตอบ ก็จะส่งผลต่อการตรวจคำตอบของระบบและการให้คะแนนไม่ถูกต้อง ทั้งนี้เพื่อให้ได้ระบบการตรวจให้คะแนนอัตโนมัติที่มีมาตรฐาน ควรมีผู้เชี่ยวชาญตรวจสอบผลการออกแบบทุกขั้นตอน ซึ่งสอดคล้องกับงานวิจัยของ (Yadav et al., 2017) ที่พบว่าการปรับปรุงสมรรถนะของระบบหรือแพลตฟอร์มดิจิทัลใดๆ ต้องใช้กระบวนการมาตรฐานและควรมีผู้เชี่ยวชาญตรวจสอบผลการออกแบบและพัฒนารวมถึงการประเมินผลในทุกขั้นตอน โดยจะสามารถสร้างแพลตฟอร์มที่สามารถตอบสนองต่อความต้องการของผู้ใช้และได้มาตรฐานในระดับสากล

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 จากผลการวิเคราะห์จะพบว่าในมิติกระบวนการทางคณิตศาสตร์ มีจุดตัดที่ใกล้เคียงกัน อย่างเห็นได้ชัดจากระดับที่ 2 ไประดับที่ 3 ระดับที่ 3 ไประดับที่ 4 ดังนั้นในการกำหนดคำอธิบายความสามารถ

ในมิติกระบวนการทางคณิตศาสตร์ ควรมีการเขียนเขียนคำอธิบายให้แตกต่างและชัดเจนเพื่อนำไปสู่การตรวจให้คะแนนในระบบได้อย่างถูกต้อง

1.2 ผลการวิเคราะห์หมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนใน 2 มิติพบว่า ทั้ง 2 มิติ ผู้เรียนมีหมโนทัศน์ที่คลาดเคลื่อนมากที่สุดคือ ด้านการใช้ข้อมูล ดังนั้นในการออกแบบระบบการให้ข้อมูลป้อนกลับ ควรอธิบายรายละเอียดข้อมูลที่โจทย์กำหนด เพื่อให้ผู้เรียนใช้ข้อมูลได้อย่างถูกต้อง

1.3 การออกแบบระบบควรมุ่งเน้นไปที่ส่วนการรับข้อมูล ซึ่งเป็นส่วนแรกของการทำงานของระบบ เช่น หากข้อมูลในส่วนจุดตัดไม่ถูกต้อง ก็จะส่งผลต่อการอธิบายความสามารถของผู้เรียนในแต่ละมิติได้ไม่ถูกต้อง หากคำตอบที่นำเข้าไปไม่ครอบคลุมคำตอบที่เป็นไปได้ทั้งหมดที่ผู้เรียนตอบ ก็จะส่งผลต่อการตรวจคำตอบของระบบและการให้คะแนนไม่ถูกต้อง ทั้งนี้เพื่อให้ได้ระบบการตรวจให้คะแนนอัตโนมัติที่มีมาตรฐาน ควรมีผู้เชี่ยวชาญตรวจสอบผลการออกแบบทุกขั้นตอน

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 เนื่องจากการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติโดยการให้ข้อมูลป้อนกลับผ่าน machine learning มีความจำเป็นต้องสร้างผลการตอบให้สอดคล้องกับการทำงานของโมเดล machine learning ดังนั้นในการวิจัยครั้งต่อไป ควรมีการเปรียบเทียบโมเดลที่เหมาะสมในการตรวจให้คะแนนแบบสอบอัตโนมัติแบบอัตโนมัติ

2.2 การวิจัยครั้งนี้สามารถตรวจให้คะแนนแบบสอบอัตโนมัติ โดยกำหนดรูปแบบในการตอบตามกรอบที่ผู้วิจัยกำหนดไว้ เพื่อให้วินิจฉัยหมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนได้อย่างถูกต้อง ในการวิจัยครั้งต่อไป ควรพัฒนาการตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ ผ่าน machine learning ที่ผู้เรียนสามารถเขียนแสดงกระบวนการแก้ปัญหาด้วยตนเองทั้งหมด โดยไม่มีกรอบในการตอบ เพื่อให้ผู้เรียนสามารถแสดงแนวคิดและการแก้ปัญหาได้อย่างหลากหลาย

References

- Adams, R. J., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21(1), 1-23.
- Brookhart, S. M., Moss, C. M., & Long, B. A. (2008). *Teacher inquire into formative assessment practices among Title I reading teachers and literacy coaches*.
http://www.castl.duq.edu/Castl_TechReports.htm
- Demars, C. (2010). *Item Response Theory: Understanding Statistics Measurement*. Oxford University Press.
- Finnegan, D., Kauppinen, A., & Wärensby, A. (2015). *Automated Feedback in a Blended Learning Environment: Student Experience and Development*. In G. Rijlaarsdam (Eds.), M. Deane & T. Guasch. (Vol. Eds.), *Studies in Writing: Vol. x, Learning and teaching writing online: Strategies for success* (pp. 31-45). Brill.

- Fisher, D., Grant, M., Frey, N., & Johnson, C. (2007). Taking formative assessments schoolwide. *Educational Leadership*, 65(4), 64–68.
- Fisher, D., & Frey, N. (2011). *The formative assessment action plan*. ASCD.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81-112.
- Junpeng, P., Krotha, J., Chanayota, K., Tang, K. N., & Wilson, M. (2019). Constructing Progress Maps of Digital Technology for Diagnosing Mathematical Proficiency. *Journal of Education and Learning*, 8(6), 90-102
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine learning: From theory to algorithms*. Cambridge university press
- Stiggins, R. J., & Chappuis, J. (2011). *An Introduction to Student-Involved Assessment FOR Learning* (6th Ed.). Pearson.
- Van Der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research*, 85(4), 475-511.
- Varasunun, P. (2011). The Program Evaluation Standards (3rd Ed.) in 2010. *Journal of Research Methodology*, 24(2), 273-278.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Routledge.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (2007). *ACER Conquest: Generalized item response modelling software*. ACER Press.
- Yadav, R., Tiruwa, A., & Suri, P. K. 2017. Internet Based Learning (IBL) in Higher Education: A Literature Review. *Journal of International Education in Business*, 10(2), 102-129.
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program Evaluation standards: A guide for evaluators and evaluation users* (3rd Ed.). Sage.

Translate Thai References

- Adulpongpaial, S. (2009). *Construction of Mathematics Diagnostic Test on Relations and Functions for MathayomSuksa 5 Students* [Master's thesis]. Chiang Mai University. (in Thai)
- Chinjunthuk, S., & Junpeng, P. (2020). Assessment Guidelines for Student's Personalized Mathematical Proficiency Development. *Journal of Educational Measurement, Mahasarakham University*, 26(1), 47- 64. (in Thai)

- Junpeng, P., Marwiang, M., Chinjunthuk, S., Suwannatrai, P., Krotha, J., Chanayota, K., Tawarungruang, C., Thuanman, J., Tang K. N., & Wilson M. (2020). *Developing Students' Mathematical Proficiency Level Diagnostic Tools through Information Technology in Assessment for Learning Report*. The Thailand Research Fund and Khon Kaen University. (in Thai)
- Junpeng, P. (2018). *Applications of the multidimensional item response theory of research* (4th Ed.). Khon Kaen University Printing. (in Thai)
- Ministry of Education. (2017). *Indicators and subjects of learning the core Learning group Mathematics according to the basic education core curriculum (Revised version 2017)*. Agricultural Co-operative Federation of Thailand Limited. (in Thai)
- Office of the Basic Education Commission. (2020). *Instruction manual learning mathematics (revised version 2020) According to the core curriculum of basic education version 2017*. Agricultural Co-operative Federation of Thailand Limited. (in Thai)
- Sukkrom, P. & Mekanong, A. (2015). A Study of Misconceptions and Mistakes in Mathematics of Eleventh Grade Students in Schools Under the Secondary Educational Service Area Office 1 and 2. *OJED*, 10(4), 599-611. (in Thai)
- Thamcharoen, T. (2019). *Fault Detection and Diagnosis of Machine Downtime in Slider Attachment Process using Machine learning Technique* [Master's thesis]. Suranaree University of Technology. (in Thai)
- The National Institute of Educational Testing Service. (2021). *The result of the national basic educational testing (O-NET) Mathayom 3 Year 2020*. <http://www.niet.or.th> (in Thai)
- Wongwanit, S. (2020). *Design Research in Education*. Chulalongkorn University. (in Thai)

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

Detecting Differential Item Functioning of Grade 9 Science Literacy Test

ศราวุธ โพธิ์สิงห์^{1*} และ วราพร เอรารรณ์²

Sarawut Posing^{1*} and Waraporn Erawan²

(Received: January 26, 2022; Revised: March 6, 2022; Accepted: March 14, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ 1) สร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 และ 2) วิเคราะห์การทำหน้าที่ต่างกัน (DIF) ของข้อสอบการรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC โดยใช้โปรแกรมประยุกต์ Mplus จำแนกตามเพศ กลุ่มตัวอย่างในงานวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 ในโรงเรียนสังกัดจังหวัดชัยภูมิ จำนวน 689 คน จากการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบการรู้เรื่องวิทยาศาสตร์ จำนวน 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ ผลการวิจัยพบว่า

1. การสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ จำนวน 54 ข้อ มีค่าดัชนีความสอดคล้องตั้งแต่ 0.60 ถึง 1 เมื่อพิจารณาคุณภาพของข้อสอบผ่านเกณฑ์การคัดเลือก จำนวน 45 ข้อ

1.1 คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) พบว่า มีค่าความยาก (p) ตั้งแต่ 0.22 ถึง 0.79 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.21 ถึง 0.77 และค่าความเชื่อมั่นทั้งฉบับ มีค่าเท่ากับ 0.92

1.2 คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory: IRT) โดยใช้โมเดลโลจิสติกแบบ 2 พารามิเตอร์ (two-parameter model) พบว่า ค่าความยากของข้อสอบ (b-parameter) มีค่าตั้งแต่ -0.537 ถึง 4.000 และค่าอำนาจจำแนกของข้อสอบ (a-parameter) มีค่าตั้งแต่ 0.313 ถึง 1.382

¹ นิสิตปริญญาโท สาขาวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

² ผู้ช่วยศาสตราจารย์ สาขาวิชาวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

¹ Master's degree student, Educational Research and Evaluation Program, Faculty of Education, Mahasarakham University

² Assistant Professor, Educational Research and Evaluation Program, Faculty of Education, Mahasarakham University

* Corresponding Author E-mail: nang49010211152@gmail.com

2. การตรวจสอบการทำหน้าที่ต่างกัน (DIF) ของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ด้วยวิธี MIMIC โดยใช้โปรแกรมประยุกต์ Mplus จำแนกตามเพศ พบว่าเพศชาย มีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 18 ข้อ และเพศหญิงมีโอกาตอบถูกมากกว่าเพศชาย จำนวน 7 ข้อ

คำสำคัญ: การรู้เรื่องวิทยาศาสตร์ การตรวจสอบการทำหน้าที่ต่างกัน วิธี MIMIC

Abstract

The aims of this research were 1) to construct and develop a science literacy test for grade 9, and 2) to analyze the differential item functioning (DIF) of the test by using Multiple Indicators and Multiple Causes (MIMIC) of the science literacy test through Mplus Program, classified by the gender. The sample consisted of 689 grade 9 students in the academic year 2018 from schools in Chaiyaphum Province, obtained through multistage random sampling. The research instrument was a science literacy test with 54 items dealing with 3 parts of knowledge content: physical systems, in 18 items, living systems, in 18 items, and Earth and space systems, in 18 items. The research findings showed that:

1. The creation and finding the quality of the 54-item science literacy test revealed the consistency index ranging from 0.60 to 1. When the quality of the test items was considered, 45 items passed the selection criteria.

1.1 The quality of the science literacy test according to the Classical Test Theory (CTT) was found to be as follows: the difficulty (p) ranged from 0.22 to 0.79; the discrimination (r) ranged from 0.21 to 0.77 and the total reliability was 0.92.

1.2 The quality of the science literacy test, according to the Item Response Theory (IRT) by using a two-parameter logistic model (Two-Parameter Model) was found to be as follows: the difficulty of the test (b -parameter) ranged from -0.537 to 4.000 and the discrimination of the test (a -parameter) ranged from 0.313 to 1.382.

2. The analysis of the differential item functioning (DIF) of the test by using Multiple Indicators and Multiple Causes (MIMIC) of the science literacy test through Mplus Program, classified by the gender, revealed that males had a higher possibility of answering correctly than females in 18 items, and females had a higher possibility of answering correctly than males in 7 items.

Keywords: science literacy test, differential item functioning (DIF), Multiple Indicators and Multiple Causes (MIMIC)

บทนำ

ความรู้และความเข้าใจทางวิทยาศาสตร์และเทคโนโลยีเป็นเครื่องมือสำคัญอย่างยิ่งในการเตรียมเยาวชนให้สามารถดำเนินชีวิตในโลกปัจจุบันที่มีวิทยาศาสตร์และเทคโนโลยีเป็นพื้นฐานและส่งผลกระทบต่อทุกชีวิตในทุกระดับ ทั้งตัวบุคคล ในอาชีพการงานและในสังคมวัฒนธรรมของทุก ๆ ชีวิต ทำให้บุคคลสามารถรับรู้และตัดสินใจประเด็นปัญหาของสังคมที่เกิดจากผลกระทบของวิทยาศาสตร์และเทคโนโลยีอย่างมีความรู้ความเข้าใจ มีส่วนร่วมในสังคมระดับชุมชน ระดับประเทศ และระดับโลกอย่างเต็มภาคภูมิ เป้าหมายของการศึกษาวิทยาศาสตร์ คือ การทำให้นักเรียนทุกคนมีการรู้เรื่องวิทยาศาสตร์ (scientific literacy) ซึ่งรวมถึงความรู้มิติต่าง ๆ ทางวิทยาศาสตร์ ได้แก่ ความรู้ความสามารถทางสติปัญญา กระบวนการแสวงหาความรู้ทางวิทยาศาสตร์ และความสัมพันธ์ระหว่างวิทยาศาสตร์กับเทคโนโลยีด้วย ในการประเมินการรู้วิทยาศาสตร์ PISA ต้องการหาตัวชี้วัดว่านักเรียนเรียนรู้ทฤษฎีและแนวคิดพื้นฐานทางวิทยาศาสตร์มากน้อยเพียงใด มีสมรรถนะทางวิทยาศาสตร์ และสามารถแก้ปัญหาในชีวิตจริงที่เกี่ยวข้องกับวิทยาศาสตร์และเทคโนโลยีได้ดีเพียงใด ทั้งนี้เพื่อหาคำตอบให้กับระดับนโยบายและระดับปฏิบัติในการปรับปรุงการเรียนการสอนวิทยาศาสตร์ ประเด็นหลักที่ PISA ให้ความสำคัญในการประเมินผลวิทยาศาสตร์ คือ การประเมินผลต้องให้ความชัดเจนกับสมรรถนะที่เหมาะสมกับนักเรียนวัย 15 ปี กล่าวคือ ต้องตั้งอยู่บนพื้นฐานที่ว่านักเรียนวัย 15 ปี ควรทำอะไร ทำอะไรได้ และให้คุณค่ากับอะไรบ้าง ภายในขอบเขตของบริบทส่วนตัว บริบทของสังคม และบริบทของโลก ซึ่งประเด็นนี้อาจแตกต่างจากการเรียนการสอนและการประเมินผลตามปกติในโรงเรียน (The Institute for the Promotion of Teaching Science and Technology, 2017a)

โครงการประเมินผลนักเรียนร่วมกับนานาชาติ (Programme for International Student Assessment หรือ PISA) ริเริ่มโดยองค์การเพื่อความร่วมมือทางเศรษฐกิจและการพัฒนา (Organization for Economic Co-operation and Development หรือ OECD) มีวัตถุประสงค์เพื่อประเมินคุณภาพของระบบการศึกษาทั้งประเทศสมาชิก OECD และประเทศที่ไม่ใช่สมาชิก OECD (partner countries) ในการเตรียมความพร้อมให้เยาวชนมีศักยภาพสำหรับการแข่งขันในอนาคต ใช้ความร่วมมือและความเชี่ยวชาญของนานาชาติในการวางกรอบโครงสร้างการประเมิน การสร้างเครื่องมือ และการศึกษาวิจัย เพื่อประกันคุณภาพของการศึกษาวิจัยให้สามารถเปรียบเทียบกันได้ในระดับนานาชาติ และข้อมูลที่ได้สามารถชี้บอกถึงคุณภาพการศึกษาของประเทศต่าง ๆ ได้ ปัจจุบันมีประเทศเข้าร่วมโครงการ PISA มากกว่า 70 ประเทศทั่วโลก โครงการ PISA ประเมินผลนักเรียนในระดับโรงเรียนและดำเนินการอย่างต่อเนื่องทุก ๆ 3 ปี กลุ่มตัวอย่างของ PISA คือนักเรียนกลุ่มอายุ 15 ปี ซึ่งโดยสากลถือว่าเป็นวัยจบการศึกษาภาคบังคับ การประเมินของ PISA เน้นการประเมินความสามารถของนักเรียนในการใช้ความรู้และทักษะเพื่อแก้ปัญหาในชีวิตจริง หรือที่เรียกว่า “การรู้เรื่อง” (literacy) ในสามด้าน ได้แก่ การรู้เรื่องการอ่าน (reading literacy) การรู้เรื่องคณิตศาสตร์ (mathematical literacy) และการรู้เรื่องวิทยาศาสตร์ (scientific literacy) ซึ่งการรู้เรื่องทั้งสามด้านนี้ ถือว่าเป็นสิ่งจำเป็นสำหรับการเรียนรู้ตลอดชีวิต และเป็นสิ่งที่ประชากรจำเป็นต้องมีเพื่อการพัฒนาและการแข่งขันทางเศรษฐกิจของประเทศ สำหรับประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี ทำหน้าที่เป็นศูนย์แห่งชาติดำเนินการวิจัยโครงการ PISA โดยได้เข้าร่วมโครงการตั้งแต่ พ.ศ. 2543 หรือใน PISA 2000

ซึ่งเป็นรอบแรกของการประเมินและได้เข้าร่วมโครงการอย่างต่อเนื่อง ในแต่ละรอบการประเมิน PISA จะสอบสามวิชาพื้นฐานดังกล่าว โดยมีสัดส่วนของข้อสอบแต่ละวิชาแตกต่างกัน ขึ้นอยู่กับว่าในรอบการประเมินนั้นเน้นการประเมินวิชาใด สำหรับรอบการประเมิน PISA 2015 เน้นการประเมินวิชาวิทยาศาสตร์ จึงมีสัดส่วนของข้อสอบวิทยาศาสตร์ประมาณ 60% ส่วนวิชาคณิตศาสตร์และการอ่าน มีสัดส่วนของข้อสอบประมาณวิชาละ 20% ซึ่งใน PISA 2015 ประเมินวิทยาศาสตร์เป็นวิชาหลักนับเป็นครั้งที่สองต่อจาก PISA 2006 (The Institute for the Promotion of Teaching Science and Technology, 2017a)

ผลการประเมิน PISA 2015 ของประเทศไทย แนวโน้มจากการประเมิน PISA 2000 ถึง PISA 2015 พบว่า ผลการประเมินทั้งสามด้านมีแนวโน้มลดลง ผลการประเมินการรู้เรื่องวิทยาศาสตร์ คะแนนเฉลี่ย OECD ของวิทยาศาสตร์ใน PISA 2015 เป็นคะแนนมาตรฐานที่ 493 คะแนน คะแนนเฉลี่ยวิทยาศาสตร์ของนักเรียนไทยคือ 421 คะแนน ซึ่งต่ำกว่าค่าเฉลี่ย OECD แนวโน้มคะแนนวิทยาศาสตร์ของนักเรียนไทยโดยรวมลดต่ำลงจาก PISA 2012 ถึง PISA 2015 คะแนนวิทยาศาสตร์ลดลงอย่างมีนัยสำคัญ (23 คะแนน) และคะแนนลดลงจนเท่ากับการประเมินรอบ PISA 2006 ที่วิทยาศาสตร์เป็นวิชาหลักและในภาพรวมระดับนานาชาติ ความแตกต่างระหว่างเพศ ด้านวิทยาศาสตร์ พบว่า นักเรียนชายมีคะแนนวิทยาศาสตร์สูงกว่านักเรียนหญิงอย่างมีนัยสำคัญทางสถิติ (4 คะแนน) (The Institute for the Promotion of Teaching Science and Technology, 2017b) ซึ่งสอดคล้องกับ Maier & Casselman (1970) กล่าวว่าเพศชายจะมีความสามารถในการแก้ปัญหาทางคณิตศาสตร์และวิชาวิทยาศาสตร์เก่งกว่าเพศหญิง ส่วนเพศหญิงจะมีความสามารถในทางภาษาและสังคมศาสตร์เก่งกว่าเพศชาย

การทดสอบแต่ละครั้งผู้สอบอาจจะมีลักษณะแตกต่างกัน เช่น เชื้อชาติ ศาสนา วัฒนธรรม ภูมิลำเนา สังคม เพศ ภาษา อายุ สภาพทางเศรษฐกิจ และประสบการณ์ เป็นต้น ทำให้ผู้สอบไม่ได้รับความยุติธรรมในการทำข้อสอบ โดยข้อสอบบางข้ออาจมีความลำเอียงเข้าข้างผู้สอบกลุ่มย่อยบางกลุ่มของผู้เข้าสอบทั้งหมด ทำให้เกิดการได้เปรียบหรือเสียเปรียบระหว่างกลุ่มผู้สอบด้วยกัน ทั้ง ๆ ที่สอบด้วยข้อสอบข้อเดียวกันหรือแบบทดสอบฉบับเดียวกัน แสดงว่าแบบทดสอบหรือข้อสอบฉบับนั้นขาดความตรง คือ ไม่ได้วัดความสามารถหลักที่ต้องการวัด (target ability) เพียงอย่างเดียว แต่ยังวัดความสามารถแทรกซ้อนที่ไม่ต้องการวัด (nuisance ability) อีกด้วย เช่น แบบทดสอบวัดความสามารถด้านคำศัพท์ภาษาไทยฉบับหนึ่ง ข้อสอบบางข้ออาจถามความรู้สำหรับผู้ชายเป็นพิเศษ เช่น ความรู้เรื่องกีฬา ในบางข้ออาจถามความรู้สำหรับผู้หญิงเป็นพิเศษ เช่น ความรู้เรื่องการตัดเย็บ จากสถานการณ์นี้แบบทดสอบวัดความสามารถคำศัพท์ในวิชาภาษาไทยเป็นความสามารถหลักที่ต้องการวัด ส่วนความสามารถด้านกีฬาและด้านการตัดเย็บเป็นความสามารถแทรกซ้อน ทำให้การตอบข้อสอบกลุ่มย่อยมีโอกาสการตอบถูกไม่เท่ากันขึ้นอยู่กับกลุ่มใดมีความสามารถแทรกซ้อนสูงกว่ากัน ทั้ง ๆ ที่ระดับความสามารถหลักที่ต้องการวัดเท่ากัน จึงทำให้ข้อสอบนั้นทำหน้าที่ต่างกัน สอดคล้องกับงานวิจัยของ Le (2009) ได้ศึกษาการทำหน้าที่ต่างกันของข้อสอบในแบบสอบการประเมินผลนักเรียนนานาชาติ (PISA) ในเนื้อหาประเมินการรู้เรื่องทางวิทยาศาสตร์ (science literacy) พบว่า ข้อสอบทำหน้าที่ต่างกันซึ่งส่วนหนึ่งมาจากรูปแบบข้อสอบและเนื้อหาที่ใช้ออกข้อสอบ และสอดคล้องกับงานวิจัยของ Sukaek (2017) กล่าวว่า ลักษณะข้อสอบที่เกิดจากการทำหน้าที่ต่างกัน โดยจำแนกตามเพศ เนื่องจากข้อสอบสามารถตอบได้หลาย

คำตอบตามความคิดเห็นของแต่ละบุคคล และข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ อาจมีคำศัพท์ที่เกี่ยวข้องกับเพศชายหรือเพศหญิง จึงทำให้ข้อสอบเข้าข้างกลุ่มนั้น ๆ และทั้งนี้อาจเป็นเพราะความสามารถที่แตกต่างกันระหว่างเพศชายหรือเพศหญิงที่มีความสามารถ ความถนัด และความสนใจในเรื่องนั้นๆ ต่างกัน

การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning: DIF) เป็นการตรวจสอบคุณภาพด้านความตรง โดยเป็นการตรวจสอบในประเด็นของความยุติธรรมของข้อสอบและแบบทดสอบ (item and test unfairness) แต่เดิมใช้คำว่า ความลำเอียงของข้อสอบ (item bias) หรือความลำเอียงของแบบทดสอบ (test bias) ซึ่งต่อมามีการเปลี่ยนมาใช้คำที่เหมาะสมกว่าเป็นการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) จากผลการตรวจข้อสอบของผู้สอบกลุ่มต่าง ๆ ในประชากรมีมานานแล้ว แต่การศึกษาคูณภาพด้านความยุติธรรมของข้อสอบหรือแบบทดสอบระหว่างผู้สอบกลุ่มต่าง ๆ ปลายปี ค.ศ. 1960 มีการเสนอวิธีการต่าง ๆ เพื่อตรวจสอบความลำเอียงของข้อสอบ (item bias) ความลำเอียงของแบบทดสอบ (test bias) และความลำเอียงในการคัดเลือก (selection bias) โดยนิยามความลำเอียงว่าเป็นความคลาดเคลื่อนอย่างเป็นระบบ (systematic error) ซึ่งเป็นการจัดข้อสอบที่ทำให้เกิดปัญหาความไม่ยุติธรรมระหว่างกลุ่มข้อสอบกลุ่มต่าง ๆ ที่มีลักษณะบางอย่างแตกต่างกัน (Kanjawasee, 2012)

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เป็นการเปรียบเทียบผลการตอบข้อสอบระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่มขึ้นไป กลุ่มแรก เรียกว่า กลุ่มเปรียบเทียบ (focal group: F) เป็นกลุ่มที่ผู้วิจัยสนใจศึกษา และคาดว่าจะได้ประโยชน์ในการตอบข้อสอบ และกลุ่มที่สองเรียกว่า กลุ่มอ้างอิง (reference group: R) เป็นกลุ่มที่คาดว่าจะได้ประโยชน์จากการตอบข้อสอบได้ถูกต้อง วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่สำคัญ ๆ ได้แก่ การวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) วิธีการวิเคราะห์การถดถอยโลจิสติก (Logistic Regression: LR) วิธีแปลงค่าความยากของข้อสอบ (Transformed Item Difficulty: TID) วิธีวัดพื้นที่ความแตกต่างระหว่างโค้งการตอบสนองข้อสอบ (Item Response Theory-D2: IRT-D2) วิธีแมนเทลแฮนส์เซล (Mantel-Haenszel: MH) วิธีไคสแควร์ของลอร์ด (Lord's chi square) วิธีอัตราส่วนไลค์ลิฮูด ลอกลิเนียร์ (Loglinear Likelihood Ratio) วิธี SIBTEST และวิธีมิมิค (Multiple Indicators and Multiple Causes: MIMIC) (Kanjawasee, 2012)

วิธี MIMIC เป็นวิธีหนึ่งที่ใช้ในโปรแกรม Mplus สำหรับการวิเคราะห์คุณภาพของข้อสอบ ตามทฤษฎีการตอบสนองข้อสอบ (IRT) ซึ่งวิธี MIMIC เป็นโมเดลลิสม์ที่มีตัวแปรแฝงเป็นตัวแปรเดียว โดยที่ตัวแปรแฝงนั้นได้รับอิทธิพลจากตัวแปรภายนอกสังเกตได้หลายตัวแปรและส่งอิทธิพลไปยังตัวแปรภายในสังเกตได้หลายตัวแปร กล่าวอีกอย่างหนึ่งคือเป็นโมเดลลิสม์เรลของคุณลักษณะแฝงที่มีหลายสาเหตุและวัดได้จากตัวบ่งชี้หลายตัว ลักษณะโมเดลจะเห็นว่าการวัดตัวแปรภายนอกสังเกตได้ต้องมีข้อตกลงข้างต้นว่าไม่มีความคลาดเคลื่อนในการวัด วิธี MIMIC นี้เป็นประโยชน์มากในการตรวจสอบความเป็นเอกมิติ (unidimensionality) การวัดผลการศึกษาสามารถวิเคราะห์ค่าพารามิเตอร์คุณลักษณะข้อสอบ และค่าความสามารถของผู้สอบไม่สามารถสังเกตโดยตรง จึงต้องประมาณจากการตอบข้อสอบ การประมาณค่าพารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ

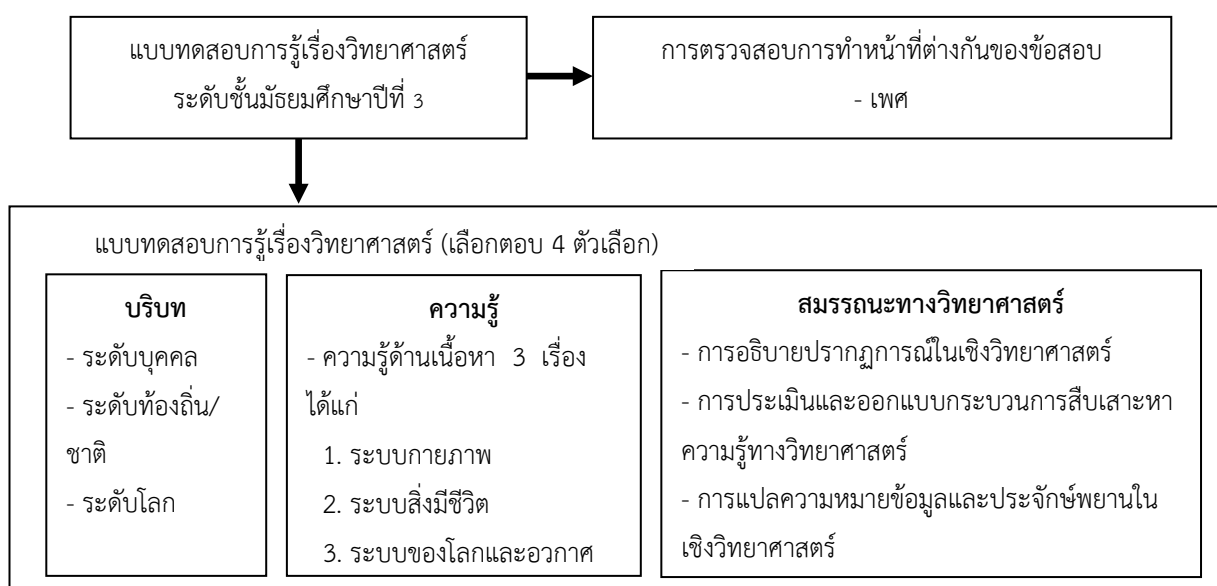
การวิจัยครั้งนี้ ผู้วิจัยมุ่งที่จะพัฒนาแบบทดสอบการรู้เรื่องวิทยาศาสตร์ เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธี MIMIC โดยใช้โปรแกรมประยุกต์ Mplus จำแนกตามเพศ โดยใช้ข้อมูลในการสร้าง

แบบทดสอบตามกรอบโครงสร้างการประเมินการรู้เรื่องวิทยาศาสตร์ โครงการประเมินผลนักเรียนนานาชาติ (PISA) ค.ศ. 2015 ได้แก่ สถานการณ์ บริบท สมรรถนะ ความรู้ทางวิทยาศาสตร์: แนวคิดและเนื้อหาที่ครอบคลุม ซึ่งเป็นแบบทดสอบที่เป็นแบบเลือกตอบ 4 ตัวเลือก โดยไม่คำนึงความแตกต่างของคุณลักษณะพื้นฐานในระดับนักเรียน เพื่อเตรียมความพร้อมให้กับนักเรียนได้ฝึกการคิดวิเคราะห์สถานการณ์ สามารถนำความรู้ทางวิทยาศาสตร์มาแก้ปัญหาตามสถานการณ์ที่กำหนด ซึ่งเป็นการพัฒนาองค์ความรู้ นำเสนอสารสนเทศที่เป็นประโยชน์ทางด้านวัดผลการศึกษา และการนำความรู้ไปใช้ปฏิบัติจริง

ความมุ่งหมายของการวิจัย

1. เพื่อสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3
2. เพื่อตรวจสอบการทำหน้าที่ต่างกัน (DIF) ของข้อสอบการรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3

กรอบแนวคิดในการวิจัย



ภาพ 1 กรอบแนวคิดในการวิจัยการสร้างแบบทดสอบการรู้เรื่องวิทยาศาสตร์ที่มี
การทำหน้าที่ต่างกันของข้อสอบ

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

1.1 ประชากร

นักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 ในโรงเรียนสังกัดจังหวัดชัยภูมิ จำนวน 6,890 คน จาก 66 โรงเรียน

1.2 กลุ่มตัวอย่าง

กลุ่มตัวอย่างในงานวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 ในจังหวัดชัยภูมิ จำนวน 689 คน โดยใช้เกณฑ์ ร้อยละ 10 ของจำนวนประชากรทั้งหมด (Srisa-ard, 2010)

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลในงานวิจัยครั้งนี้ ได้แก่ แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 เป็นแบบทดสอบเชิงสถานการณ์ 4 ตัวเลือก จำนวน 54 ข้อ ตามกรอบโครงสร้างการประเมินการรู้เรื่องวิทยาศาสตร์ ซึ่งประกอบด้วย สมรรถนะทางวิทยาศาสตร์ บริบทของวิทยาศาสตร์ ระดับบริบทของวิทยาศาสตร์ และความรู้ทางวิทยาศาสตร์ (ความรู้ด้านเนื้อหา) 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ ซึ่งเป็นเนื้อหาที่ครอบคลุมตามกรอบเนื้อหาของ PISA 2015

3. การเก็บรวบรวมข้อมูล

3.1 ทำหนังสือขอความอนุเคราะห์ในการรวบรวมข้อมูลในการวิจัยถึงผู้บริหารสถานศึกษาที่ได้รับการสุ่มเป็นกลุ่มตัวอย่างในการวิจัย

3.2 ประสานงานชี้แจงทำความเข้าใจกับผู้บริหารและครูรวบรวมข้อมูลเกี่ยวกับจุดมุ่งหมายการวิจัยเครื่องมือในการรวบรวมข้อมูล วิธีการเก็บรวบรวมข้อมูล ระยะเวลาการเก็บรวบรวมข้อมูล และกำหนดนัดหมายสำหรับการรับคืนแบบทดสอบ

3.3 สถานศึกษาทำหน้าที่สุ่มห้องเรียนและทำแบบทดสอบของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ตามระยะเวลาที่กำหนด

3.4 ประสานงานทางโทรศัพท์เพื่อยืนยันความพร้อมของสถานศึกษาที่จะให้เก็บรวบรวมข้อมูล และการยืนยันระยะเวลาการรวบรวมข้อมูลในแต่ละครั้ง รวมถึงการปรับระยะเวลาการเก็บรวบรวมข้อมูล ในกรณีการรวบรวมข้อมูลในสถานศึกษาบางแห่งไม่เป็นไปตามที่กำหนด

3.5 ข้อมูลที่ได้รับกลับ ผู้วิจัยทำการลงรหัสรายชื่อนักเรียน สถานศึกษาเป็นรหัสที่สามารถแยกแยะได้ว่าผู้ตอบเป็นนักเรียนในโรงเรียนใด ทั้งนี้เพื่อเป็นการเชื่อมโยงข้อมูลการวัดแต่ละครั้งเข้าด้วยกันได้ง่าย

4. การวิเคราะห์ข้อมูล

4.1 ตรวจสอบความเที่ยงตรงของข้อสอบที่พัฒนาขึ้น ตามกรอบโครงสร้างตามแนว PISA 2015 โดยพิจารณาค่าเฉลี่ยของความคิดเห็นของผู้เชี่ยวชาญ จำนวน 5 ท่าน

4.2 วิเคราะห์หาคุณภาพของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (CTT) เพื่อวิเคราะห์ค่าความยาก (p) ค่าอำนาจจำแนก (r) และค่าความเชื่อมั่น กับกลุ่มทดลองใช้ (try Out) ที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 100 คน

4.3 วิเคราะห์ค่าพารามิเตอร์ของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ตามแนวทฤษฎีการตอบสนองของข้อสอบ (IRT) เพื่อวิเคราะห์ค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a) ของข้อสอบ กับกลุ่มตัวอย่าง จำนวน 689 คน โดยผู้วิจัยได้เขียนคำสั่งจากโปรแกรม Mplus

4.4 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธี MIMIC (multiple indicators and multiple causes) เมื่อจำแนกตามเพศ โดยประยุกต์ใช้โปรแกรม Mplus มีขั้นตอนดังต่อไปนี้

4.4.1. วิเคราะห์โมเดล CFA โดยไม่มีตัวแปรทำนาย ในการวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองของข้อสอบ (IRT) เป็นการศึกษาคุณสมบัติอยู่บนพื้นฐานข้อตกลงเอกมิติ (unidimensional) ของตัวแปรแฝง

4.4.2. การเพิ่มตัวแปรทำนาย แต่ไม่มีอิทธิพลทางตรงต่อข้อสอบ การวิจัยครั้งนี้มีตัวแปร คือ เพศ ซึ่งผู้วิจัยกำหนดให้ กลุ่มอ้างอิง (reference group) ได้แก่ เพศชาย และกลุ่มเปรียบเทียบ (focal group) ได้แก่ เพศหญิง

4.4.3. เพิ่มอิทธิพลทางตรงต่อข้อสอบบังคับค่าให้เป็น 0 เพื่อกำหนดให้ตัวแปรต้นไม่มีผลทางตรงต่อข้อสอบ

4.4.4. ประเมินความกลมกลืนของโมเดล และตรวจสอบผลทางตรงที่มีนัยสำคัญทางสถิติ ถ้าตรวจพบว่ามีผลทางตรง และมีนัยสำคัญทางสถิติ $p < 0.05$ แสดงว่าข้อสอบข้อนั้น คือ ข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF)

ผลการวิจัย

1. การสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

1.1 ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม

1) ความเที่ยงตรง (validity)

ผู้วิจัยได้สร้างแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 1 ฉบับ 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ แล้วนำไปให้ผู้เชี่ยวชาญ จำนวน 5 คน พิจารณาความสอดคล้อง พบว่า ข้อสอบผ่านเกณฑ์ทุกข้อ มีค่าตั้งแต่ 0.60 ถึง 1.00 จึงสามารถนำข้อสอบทั้ง 54 ข้อ ไปทดลองกลุ่มตัวอย่าง จำนวน 100 คน เพื่อหาคุณภาพเป็นรายข้อและทั้งฉบับ

2) ค่าความยาก (difficulty) ค่าอำนาจจำแนก (discrimination) และค่าความเชื่อมั่น (reliability)

การวิเคราะห์ค่าความยาก และค่าอำนาจจำแนก ของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง พบว่า ระบบทางกายภาพ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ รวมข้อสอบที่ผ่านเกณฑ์ทั้งหมด จำนวน 45 ข้อ มีค่าความยาก (p) ตั้งแต่ 0.22 ถึง 0.79 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.21 ถึง 0.77 และค่าความเชื่อมั่นทั้งฉบับ มีค่าเท่ากับ 0.92

1.2 ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการตอบสนองของข้อสอบ

การวิเคราะห์แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 15 ข้อ ระบบสิ่งมีชีวิต จำนวน 15 ข้อ

และระบบของโลกและอวกาศ จำนวน 15 ข้อ ที่เป็นกลุ่มตัวอย่าง จำนวน 689 คน พบว่า ค่าความยากของข้อสอบ (b-parameter) มีค่าตั้งแต่ -0.537 ถึง 4.000 และค่าอำนาจจำแนกของข้อสอบ (a-parameter) มีค่าตั้งแต่ 0.313 ถึง 1.382

2. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

2.1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC จำนวน 45 ข้อ กับกลุ่มตัวอย่าง จำนวน 689 คน พบว่า

ระบบทางกายภาพ จำนวน 15 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 2 ข้อ ได้แก่ 5 และ 8 คิดเป็นร้อยละ 13.33 ของข้อสอบทั้งหมด

ระบบสิ่งมีชีวิต จำนวน 15 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 14 ข้อ ได้แก่ 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29 และ 30 คิดเป็นร้อยละ 93.33 ของข้อสอบทั้งหมด

ระบบของโลกและอวกาศ จำนวน 15 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 9 ข้อ ได้แก่ 31, 32, 33, 34, 36, 37, 41, 43 และ 44 คิดเป็นร้อยละ 60.00 ของข้อสอบทั้งหมด

2.2 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ด้วยวิธี MIMIC จำแนกตามเพศ โดยใช้โปรแกรมประยุกต์ Mplus

ตาราง 1 สรุปผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC จำแนกตามเพศ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธี MIMIC จำแนกตามเพศ	
ชาย (Reference)	หญิง (Focal)
5, 8, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 34, 37	31, 32, 33, 36, 41, 43, 44

หมายเหตุ เพศชาย = 1 (Reference Group), เพศหญิง = 2 (Focal Group)

จากตาราง 1 พบว่า ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ด้วยวิธี MIMIC จำแนกตามเพศ โดยใช้โปรแกรมประยุกต์ Mplus พบว่า เพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 18 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 7 ข้อ

อภิปรายผล

1. ผลการสร้างและหาคุณภาพของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 1 ฉบับ 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ ผู้วิจัยได้สร้าง วิเคราะห์ และตรวจสอบคุณภาพของแบบทดสอบด้วยเทคนิคและวิธีการวิเคราะห์ข้อสอบตามแนวทฤษฎีการทดสอบแบบดั้งเดิม และทฤษฎีการตอบสนองของข้อสอบ ซึ่งทั้งสองแนวคิดต่างก็มีข้อดีและข้อเสียต่างกัน แต่เมื่อนำ

สองทฤษฎีมาใช้เป็นแนวทางในการวิเคราะห์ข้อสอบก็สามารถลดข้อจำกัดลงได้ จึงทำให้ได้แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ที่มีคุณภาพ รายละเอียดดังต่อไปนี้

1.1 ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม

1) ค่าความเที่ยงตรง (validity)

การหาความเที่ยงตรงเชิงเนื้อหาด้วยวิธีหาค่าเฉลี่ยความสอดคล้องความคิดเห็นของผู้เชี่ยวชาญ จำนวน 5 ท่าน พบว่า ดัชนีความสอดคล้องมีค่าตั้งแต่ 0.60 ถึง 1.00 แสดงให้เห็นว่าแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ข้อสอบแต่ละข้อมีความสอดคล้องกับนิยามที่ต้องการวัด โดยจุดประสงค์เชิงพฤติกรรมมีความครอบคลุมเนื้อหา จึงสรุปได้ว่าข้อสอบที่สร้างขึ้นมีความเที่ยงตรงเชิงเนื้อหา ซึ่งผู้วิจัยได้ดำเนินการสร้างข้อสอบให้สอดคล้องครอบคลุมตามกรอบเนื้อหาของ PISA 2015 และพฤติกรรมที่ต้องการวัด จึงทำให้ข้อสอบที่สร้างขึ้นมีความตรงเชิงเนื้อหา สอดคล้องกับแนวคิดของ Ahmann and Glock (1967) กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยไว้ว่าเน้นความเที่ยงตรงเชิงเนื้อหา (content validity) สอดคล้องกับแนวคิดของ Worakham (2015) กล่าวว่า ควรเลือกข้อคำถามที่มีค่าดัชนีความสอดคล้องตั้งแต่ 0.50 ขึ้นไป หากต่ำกว่า 0.50 ถือว่าใช้ไม่ได้ สอดคล้องกับงานวิจัยของ Weerasin (2018) พบว่า แบบทดสอบความคิดสร้างสรรค์เชิงตัวเลขสำหรับนักเรียนระดับชั้นประกาศนียบัตรวิชาชีพชั้นปีที่ 1 มีค่าดัชนีความสอดคล้องตั้งแต่ 0.80 ถึง 1.00 สอดคล้องกับงานวิจัยของ Thanakhwang (2010) พบว่า แบบวัดความคิดสร้างสรรค์ทางคณิตศาสตร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 4 มีค่าดัชนีความสอดคล้องตั้งแต่ 0.71 ถึง 1.00

2) ค่าความยาก (difficulty) ค่าอำนาจจำแนก (discrimination)

ค่าความยากและค่าอำนาจจำแนก จากการทดสอบกับกลุ่มตัวอย่าง จำนวน 100 คน ของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง พบว่า ระบบทางกายภาพ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ รวมข้อสอบที่ผ่านเกณฑ์ทั้งหมด จำนวน 45 ข้อ มีค่าความยาก ตั้งแต่ 0.22 ถึง 0.79 ค่าอำนาจจำแนก ตั้งแต่ 0.21 ถึง 0.77 ข้อสอบส่วนใหญ่มีค่าความยากและค่าอำนาจจำแนกรายข้ออยู่ในเกณฑ์ดี แต่มีบางข้อมีค่าความยากต่ำเกินไป แสดงว่าข้อสอบนั้นยาก ผู้วิจัยจึงคัดเลือกเฉพาะข้อสอบที่มีนัยสำคัญทางสถิติซึ่งมีค่าอำนาจจำแนกเข้าเกณฑ์ทุกข้อ จึงเชื่อได้ว่าแบบวัดที่ละข้อมีอำนาจจำแนกอย่างแท้จริง และค่าอำนาจจำแนกนี้ได้แสดงให้เห็นว่าสามารถจำแนกนักเรียนกลุ่มสูงและกลุ่มต่ำออกจากกันได้ สอดคล้องกับ Saiyot & Saiyot (2000) ซึ่งกล่าวไว้ว่า ค่าความยากของข้อสอบควรอยู่ในเกณฑ์ 0.20 ถึง 0.80 สอดคล้องกับ Phatthiyathani (2008) ซึ่งกล่าวไว้ว่า ค่าความยากของข้อสอบควรอยู่ในเกณฑ์ 0.20 ถึง 0.80 และค่าอำนาจจำแนกของข้อสอบควรมีค่าตั้งแต่ 0.20 ถึง 1.00 และยังสอดคล้องกับงานวิจัยของ Chantawan (2019) พบว่า แบบทดสอบความสามารถในการให้เหตุผลทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5 ในจังหวัดกาฬสินธุ์ มีค่าความยากรายข้อ ตั้งแต่ 0.34 ถึง 0.80 และมีค่าอำนาจจำแนกรายข้อตั้งแต่ 0.22 ถึง 0.65 แสดงว่า มีค่าตามเกณฑ์ทุกข้อ

3) ความเชื่อมั่นของแบบทดสอบ (reliability)

การหาค่าความเชื่อมั่นของแบบทดสอบ ผู้วิจัยใช้วิธีของคูเดอร์-ริชาร์ดสัน โดยใช้แบบทดสอบฉบับเดียวดำเนินการสอบเพียงครั้งเดียว ใช้สูตร KR-20 (Phatthiyathani, 2010) ค่าความเชื่อมั่นของแบบทดสอบ เท่ากับ 0.92 จะเห็นว่าค่าความเชื่อมั่นของแบบทดสอบวัดที่ผู้วิจัยสร้างขึ้นมีค่าสูงพอที่จะเชื่อถือได้ว่าเป็นแบบทดสอบวัดที่มีคุณภาพเหมาะสม สาเหตุที่เป็นเช่นนี้เนื่องมาจากแบบทดสอบวัดได้ผ่านการคัดเลือกข้อที่เข้าเกณฑ์ปรับปรุงข้อคำถามให้เหมาะสมยิ่งขึ้นสอดคล้องกับแนวคิดของ Saiyot & Saiyot (2000) ซึ่งกล่าวว่า ค่าความเชื่อมั่นจะมีค่าอยู่ระหว่าง -1 ถึง +1 และจะพิจารณาเฉพาะค่าที่เป็นบวกเท่านั้น ซึ่งควรมีค่ามากกว่า 0.70 จึงจะเป็นแบบทดสอบที่มีความเชื่อมั่น แต่หากพิจารณาเป็นรายด้านจะพบว่าค่าความเชื่อมั่นไม่ค่อยสูงนัก เนื่องจากจำนวนข้อของแต่ละด้านมีจำนวนน้อย ทำให้ค่าความเชื่อมั่นต่ำกว่าความเชื่อมั่นรวมทั้งฉบับ เนื่องจากจำนวนข้อสอบมีผลต่อค่าความเชื่อมั่น สอดคล้องกับงานวิจัยของ Sumlertrum (2015) พบว่า แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนรายวิชาวิทยาศาสตร์ สำหรับชั้นมัธยมศึกษาปีที่ 1 ตามแนวข้อสอบในโครงการประเมินผลนักเรียนร่วมกับนานาชาติ (PISA) จากการทดลองใช้ทั้ง 3 ครั้ง มีค่าความเชื่อมั่น 0.82, 0.85 และ 0.83 ตามลำดับ

1.2 ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการตอบสนองของข้อสอบ

1) ค่าความยาก (b-parameter) ของข้อสอบ มีค่าระหว่าง -0.537 ถึง 4.000 หมายความว่า ค่าความยากของข้อสอบทุกข้ออยู่ในเกณฑ์ตามทฤษฎีกำหนด แสดงว่าแบบทดสอบนี้มีคุณภาพอยู่ในเกณฑ์ที่ใช้ได้ ซึ่ง Kanjanawasee (2012) กล่าวว่า ตามทฤษฎีการตอบสนองของข้อสอบ มีค่าความยากตั้งแต่ -2.50 ถึง +2.50 ค่า b ที่อยู่ใกล้ -2.50 แสดงว่าเป็นข้อสอบที่ง่าย ส่วนค่า b ที่อยู่ใกล้ +2.50 เป็นข้อสอบที่ยาก

2) ค่าอำนาจจำแนก (a-parameter) ของข้อสอบ มีค่าอยู่ระหว่าง 0.307 ถึง 0.864 แสดงให้เห็นว่าข้อสอบมีค่าอยู่เกณฑ์คุณภาพที่ใช้ได้ ทั้งนี้ อาจเนื่องมาจากข้อสอบได้ผ่านการวิเคราะห์และปรับปรุงคุณภาพมาแล้วครั้งหนึ่ง จึงทำให้ข้อสอบที่ได้มีค่าพารามิเตอร์อยู่ในเกณฑ์ที่ดี ซึ่ง Kanjanawasee (2012) กล่าวว่าตามทฤษฎีการตอบสนองของข้อสอบ ค่าอำนาจจำแนกอยู่ระหว่าง +0.5 ถึง +2.5

2. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) การรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ด้วยวิธี MIMIC โดยใช้โปรแกรมประยุกต์ Mplus จำแนกตามเพศ พบว่า เพศชายมีโอกาสดอกถูกมากกว่าเพศหญิง จำนวน 18 ข้อ และเพศหญิงมีโอกาสดอกถูกมากกว่าเพศชาย จำนวน 7 ข้อ ซึ่งสอดคล้องกับแนวคิดของ Maier and Casselman (1970) กล่าวว่าเพศชายจะมีความสามารถในการแก้ปัญหาทางคณิตศาสตร์และวิทยาศาสตร์เก่งกว่าเพศหญิง ส่วนเพศหญิงจะมีความสามารถในทางภาษาและสังคมศาสตร์เก่งกว่าเพศชาย สอดคล้องกับแนวคิดของ Kanjanawasee (2012) กล่าวว่า การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เป็นการเปรียบเทียบผลการตอบข้อสอบเป็นรายข้อระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่ม มีความสามารถหลัก (primary ability) ที่มุ่งวัดเท่ากัน แต่คาดว่าจะมีความได้เปรียบเสียเปรียบกัน โดยกลุ่มหนึ่งถือเป็น กลุ่มอ้างอิง (reference group) ซึ่งคาดว่าจะได้เปรียบในการตอบข้อสอบข้อนั้น หรือมีโอกาสดอกข้อสอบได้ถูกต้องมากกว่า ส่วนอีกกลุ่มคือ กลุ่มเปรียบเทียบ (focal group) ซึ่งเป็นกลุ่มที่สนใจศึกษา

และคาดว่าจะจะเป็นกลุ่มที่เสียเปรียบ สอดคล้องกับงานวิจัยของ Sukaek (2017) กล่าวว่า ลักษณะข้อสอบที่เกิดจากการทำหน้าที่ต่างกัน โดยจำแนกตามเพศ เนื่องจากข้อสอบสามารถตอบได้หลายคำตอบตามความคิดเห็นของแต่ละบุคคล และข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ อาจมีคำศัพท์ที่เกี่ยวข้องกับเพศชายหรือเพศหญิง จึงทำให้ข้อสอบเข้าข้างกลุ่มนั้น ๆ และทั้งนี้อาจเป็นเพราะความสามารถที่แตกต่างกันระหว่างเพศชายหรือเพศหญิงที่มีความสามารถ ความถนัด และความสนใจในเรื่องนั้น ๆ ต่างกัน สอดคล้องกับงานวิจัยของ Le (2009) ได้ศึกษาการทำหน้าที่ต่างกันของข้อสอบในแบบสอบการประเมินผลนักเรียนนานาชาติ (PISA) ในเนื้อหาประเมินการรู้เรื่องทางวิทยาศาสตร์ (Science Literacy) พบว่า ข้อสอบทำหน้าที่ต่างกันซึ่งส่วนหนึ่งมาจากรูปแบบข้อสอบและเนื้อหาที่ใช้ข้อสอบ และสอดคล้องกับผลการประเมิน PISA 2015 ยังพบความแตกต่างของคะแนนเฉลี่ยวิทยาศาสตร์ระหว่างนักเรียนชายและนักเรียนหญิง โดยเฉลี่ยในประเทศสมาชิก OECD นักเรียนชายมีคะแนนวิทยาศาสตร์สูงกว่านักเรียนหญิง 4 คะแนน แม้ความแตกต่างจะมีช่องว่างไม่มากนักแต่ก็มีความสำคัญทางสถิติ (The Institute for the Promotion of Teaching Science and Technology, 2018)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 หน่วยงานที่เกี่ยวข้องในการประเมินการรู้เรื่องวิทยาศาสตร์ของนักเรียน สามารถนำแนวคิดในการจัดความไม่ยุติธรรมของข้อสอบที่เอนเอียงไปยังนักเรียนกลุ่มใดกลุ่มหนึ่งนั้นจำเป็นอย่างยิ่งต่อคะแนนที่แท้จริงของนักเรียนแต่ละคน โดยภาพรวมจะต้องมีการตรวจสอบการทำหน้าที่ต่างกันของแบบสอบร่วมกับการตรวจสอบการทำหน้าที่ต่างกันออกไปซึ่งจะช่วยเพิ่มความตรงของข้อสอบให้เหมาะสมกับกลุ่มนักเรียนมากขึ้น และเพื่อเป็นแนวทางในการพัฒนาคุณภาพข้อสอบต่อไป

1.2 ผู้ที่เกี่ยวข้องในการออกข้อสอบ ควรนำผลการวิจัยนี้ไปใช้เป็นข้อมูลในการพิจารณาทบทวนข้อสอบ กำจัดความเป็นไปได้ของสาเหตุที่จะทำให้เกิดการทำหน้าที่ต่างกันของข้อสอบ เช่น คำศัพท์ที่เกี่ยวข้องกับเพศชายหรือเพศหญิง ภาษาที่ใช้ในการเขียนข้อสอบ ลักษณะเนื้อหาที่นำมาสร้างข้อสอบ ลักษณะเฉพาะของกลุ่มผู้สอบ เป็นต้น

1.3 นักวิจัยและนักวัดผลทางการศึกษาที่มีความสนใจเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบสามารถเลือกใช้ขนาดกลุ่มตัวอย่างที่มีความเหมาะสมที่จะทำให้มีประสิทธิภาพ ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบดีที่สุด กลุ่มตัวอย่างขนาดใหญ่จะทำให้สามารถตรวจพบข้อสอบที่ทำหน้าที่ต่างกันได้ดีกว่ากลุ่มตัวอย่างขนาดเล็ก โดยขนาดกลุ่มตัวอย่างสำหรับการวิเคราะห์โมเดลสมการโครงสร้างต้องไม่ต่ำกว่า 100 ตัวอย่าง และมีสัดส่วนจำนวนเท่าของขนาดตัวอย่างต่อจำนวนพารามิเตอร์ที่ต้องการประมาณค่าเป็น 10-12 ตัวอย่าง ต่อ 1 พารามิเตอร์

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 วิธี MIMIC มีประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบที่มีการให้คะแนนแบบ 2 ค่า จึงควรมีการนำไปเปรียบเทียบกับวิธีการตรวจสอบอื่น ๆ และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบที่มีการให้คะแนนมากกว่า 2 ค่า

2.2 ควรมีการศึกษาตัวแปรกลุ่มอื่น ๆ ในการเปรียบเทียบการทำหน้าที่ต่างกันของข้อสอบ เช่น ภูมิฐานะ ศาสนา เชื้อชาติ ภาษา อายุ เป็นต้นว่าจะส่งผลให้ข้อสอบเกิดการการทำหน้าที่ต่างกันหรือไม่

References

- Ahman, S.J., & Glock, M.D. (1967). *Evaluating Pupil Growth Principle of Tests and Measurement* (3rd Ed.). Allyn & Bacon.
- Le, L. T. (2009). Investigating Gender Differential Item Functioning Across Countries and Test Languages for PISA Science Items. *International Journal of Testing*, 9(2), 122-133.
- Maier, N. R. F., & Casselman, G. G. (1970). Locating the difficulty in insight problems: Individual and sex differences. *Psychological Rep*, 26, 103-107.

Translate Thai References

- Chantawan, S. (2019). *Development of Mathematics Reasoning Abilities Assessment for Matthayomsueksa 5 Students in Kalasin Province* [Maste's thesis]. Mahasarakham University. (in Thai)
- Kanjanawasee, S. (2012). *Modern Test Theories* (4th Ed.). Chulalongkorn University Press. (in Thai)
- Ministry of Education. (2009). *Basic Education Core Curriculum B.E. 2551 (2008)*, Office of the Basic Education Commission, Ministry of Education. Agricultural Co-operative Association of Thailand Company Limited. (in Thai)
- Phatthiyathani, S. (2008). *Educational measurement* (6th Ed.). Prasarn Press. (in Thai)
- Phatthiyathani, S. (2010). *Educational measurement* (7th Ed.). Prasarn Press. (in Thai)
- Saiyot, L., & Saiyot, A. (2000). *Learning Measurement Techniques* (2nd Ed). Suweerivasarn Press. (in Thai)
- Srisa-ard, B. (2010). *Preliminary Research* (8th Ed). Suweerivasarn Press. (in Thai)
- Sukaek, P. (2017). *Detecting Differential Item Functioning Using Multiple Indicators and Multiple Causes (MIMIC)* [Maste's thesis]. Prince of Songkla University. (in Thai)

- Sumlertrum, Y. (2015). The Construction of a Science Academic Achievement Test at the Matthayom Sueksa One Level as Based on the Programme for International Student Assessment (PISA). *Journal of Humanities and Social Sciences Surin Rajabhat University*, 19(2), 21-34. (in Thai)
- Thanakhwang, T. (2010). *Construction of creative thinking test in mathematics for Prathom Suksa 4 students* [Maste's thesis]. Chiang Mai University. (in Thai)
- The Institute for the Promotion of Teaching Science and Technology. (2017a). *PISA 2015 Assessment Framework*. Aroonkarnpim Press. (in Thai)
- The Institute for the Promotion of Teaching Science and Technology. (2017b). *Summary of preliminary data PISA 2015*. Aroonkarnpim Press. (in Thai)
- The Institute for the Promotion of Teaching Science and Technology. (2018). *PISA 2015 results Snapshot of performance in science, reading and mathematics*. Success Publication Company Limited. (in Thai)
- Weerasin, N. (2018). *Construction of numerical creative thinking test for vocational students 1 at Intrachai Commercial College of Vocational Education Commission* [Maste's thesis]. Burapha University. (in Thai)
- Worakham, P. (2015). *Educational research* (7th Ed.). Takkasila press. (in Thai)

การพัฒนาแบบสอบคู่ขนานวิชาคณิตศาสตร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2
โรงเรียนสังกัดกรุงเทพมหานคร โดยการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ
แบบ 2 พารามิเตอร์

Development of Parallel Tests in Mathematics for Students in Grade 2 in
Schools under the Bangkok Metropolitan Administration through an Application
of Two-Parameter Item Response Theory

สาวินา บุญเสน^{1*} จุฑาภรณ์ มาสันเทียะ² และ พนิดา ปานิดวัฒนะ³

Savina Boonsen^{1*}, Chuthaphon Masantiah² and Panida Panidvadana³

(Received: January 1, 2022; Revised: February 28, 2022; Accepted: March 10, 2022)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมายเพื่อ 1) พัฒนาแบบสอบคู่ขนานวิชาคณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 2 สาระการวัด และ 2) ตรวจสอบความเป็นคู่ขนานของแบบสอบคู่ขนานวิชาคณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 2 สาระการวัด กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 2 ปีการศึกษา 2562 จำนวน 800 คน จากโรงเรียน สังกัดกรุงเทพมหานคร ในเขตหนองจอก จำนวน 13 โรงเรียน โดยการสุ่มแบบแบ่งชั้น (stratified random sampling) เครื่องมือที่ใช้ คือ แบบสอบแบบปรนัยเรื่องความยาว (1A, 1B) และเรื่องน้ำหนัก (2A, 2B) ซึ่งเป็น ข้อสอบแบบเลือกตอบฉบับละ 20 ข้อ การวิเคราะห์ข้อมูล ได้แก่ ความตรงตามเนื้อหา ความเที่ยง ค่าฟังก์ชัน สารสนเทศของแบบสอบ ค่าความยาก ค่าอำนาจจำแนก ค่าฟังก์ชันสารสนเทศของข้อสอบ และการตรวจสอบ ความเป็นคู่ขนานของข้อสอบและแบบสอบโดยใช้ Root mean square deviation of item information (RMSD_{IF}) และ Root mean square deviation of test information (RMSD_{TF}) ตามลำดับ

ผลการวิจัยพบว่า

1) แบบสอบคู่ขนานวิชาคณิตศาสตร์ทั้ง 4 ฉบับ มีค่า IOC ซึ่งเป็นดัชนีความตรงตามเนื้อหาของแบบสอบ ทุกฉบับ ตั้งแต่ 0.67-1.00 ความเที่ยงมีค่า 0.711, 0.652, 0.718 และ 0.635 ตามลำดับ ค่าฟังก์ชันสารสนเทศของ แบบสอบ 1A และ 1B ให้สารสนเทศสูงสุดสำหรับผู้ที่มีระดับความสามารถระหว่าง -0.60 ถึง -0.80 และแบบสอบ

¹ นักศึกษาปริญญาโท สาขาวิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

² ผู้ช่วยศาสตราจารย์ สาขาวิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

³ อาจารย์ สาขาวิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

¹ Master's degree student, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

² Assistant Professor, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

³ Lecturer, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

* Corresponding Author E-mail: savina_jack@hotmail.com

ฉบับ 2A และ 2B ให้สารสนเทศสูงสุดสำหรับผู้ที่มีระดับความสามารถระหว่าง -0.40 ถึง -0.60 ขณะที่ค่าความยาก (b) อยู่ระหว่าง -2.98 ถึง +0.57 ค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.22 ถึง 1.13 และค่าฟังก์ชันสารสนเทศของข้อสอบทั้ง 4 ฉบับมีค่าอยู่ระหว่าง 0.013 ถึง 0.320

2) ผลการวิเคราะห์ความเป็นคู่ขนานรายข้อและทั้งฉบับของแบบสอบคู่ขนาน เรื่อง ความยาว (1A, 1B) มีค่า $RMSD_{IF}$ ตั้งแต่ 0.008 ถึง 0.167 และค่า $RMSD_{TIF}$ เท่ากับ 0.484 ขณะที่แบบสอบคู่ขนาน เรื่อง น้ำหนัก (2A, 2B) มีค่า $RMSD_{IF}$ ตั้งแต่ 0.020 – 0.110 และค่า $RMSD_{TIF}$ เท่ากับ 0.492

คำสำคัญ: ทฤษฎีการตอบสนองข้อสอบ แบบสอบคู่ขนาน ฟังก์ชันสารสนเทศของข้อสอบ
ฟังก์ชันสารสนเทศของแบบสอบ

Abstract

The purposes of this research were to 1) develop parallel tests in mathematics for grade 2, in the strand of measurement, and 2) verify the parallelism of the parallel tests in mathematics for grade 2, in the strand of measurement. The sample consisted of 800 grade 2 students in the academic year 2019 from 13 schools in Nong Chok District, under the Bangkok Metropolitan Administration, selected by stratified random sampling. The research instruments were 4 objective tests, on length (1A, 1B) and weight (2A, 2B). Each test comprised 20 multiple-choice items. The data were analyzed for content validity, reliability, test information function, item difficulty, discrimination, item information function, and the parallelism of the parallel tests and items using the root mean square deviation of item information ($RMSD_{IF}$) and root mean square deviation of test information ($RMSD_{TIF}$) respectively.

The research findings revealed that:

1) All the 4 parallel tests in mathematics had the index of item objective congruence (IOC), which was the index of content validity, ranging from 0.67 to 1.00; the reliability of 0.711, 0.652, 0.718, and 0.635 respectively; the maximum value of the test information function for the proficiency level was from -0.60 to -0.80 for the 1A and 1B tests and -0.40 to -0.60 for the 2A and 2B tests; the item difficulty (b) was between -2.98 and 0.57; the discrimination (a) was between 0.22 and 1.13, and the item information function of all the 4 tests was between 0.013 and 0.320.

2) The results of the parallelism analysis of the items and tests on length (1A, 1B) revealed that the $RMSD_{IF}$ ranged from 0.008 to 0.167, and $RMSD_{TIF}$ was 0.484; while the parallelism analysis of the items and tests on weight (2A, 2B) revealed that the $RMSD_{IF}$ ranged from 0.020 – 0.110, and $RMSD_{TIF}$ was 0.492.

Keywords: Item Response Theory, parallel test, item information function, test information function

บทนำ

การวัดผลและประเมินผลการเรียนรู้เป็นกระบวนการที่ผู้สอนใช้พัฒนาคุณภาพผู้เรียนเพราะผลจากการวัดจะทำให้ได้สารสนเทศที่แสดงถึงพัฒนาการความก้าวหน้าและความสำเร็จทางการเรียนของผู้เรียน โดยเครื่องมือที่ใช้ในการวัดผลการศึกษาที่มีหลายชนิด ซึ่งเครื่องมือที่นิยมใช้ในการวัดผลสัมฤทธิ์ทางการเรียน ได้แก่ แบบสอบที่ใช้สำหรับวัดความสามารถทางสมองหรือด้านสติปัญญาของผู้เรียนว่ามีความสามารถมากน้อยเพียงใด จากแนวทางการจัดการกระบวนการเรียนรู้ของสังกัดกรุงเทพมหานครที่มุ่งเน้นการจัดการศึกษาที่มีคุณภาพ และพัฒนาให้เยาวชนในพื้นที่กรุงเทพฯ มีความรู้ มีจิตสำนึก รักเมืองกรุงเทพฯ ห่วงแค้นและร่วมสืบทอดวัฒนธรรมและอัตลักษณ์ของเมืองกรุงเทพฯ มีคุณธรรม มีทักษะในการดำรงชีวิตและประกอบอาชีพสุจริต (Education Strategy Office Education Bureau, 2016) ครูผู้สอนจึงนำแนวคิดนี้มาพัฒนานักเรียนให้มีความสามารถด้านความรู้ ทักษะ และคุณลักษณะอันพึงประสงค์ ตลอดจนครูต้องมีการวัดและประเมินผลการเรียนรู้ให้สอดคล้องตามสิ่งที่มุ่งหวัง

จากสภาพการวัดและประเมินผลของโรงเรียนสุเหร่าหะยีมิยา ผู้วิจัยในฐานะครูผู้สอนวิชาคณิตศาสตร์ พบปัญหาที่สำคัญ คือ กระบวนการสร้างแบบสอบที่ขาดมาตรฐาน เช่น ขาดการตรวจสอบคุณภาพของข้อสอบ และแบบสอบที่ครูสร้างขึ้น รวมถึงการใช้ข้อสอบซ้ำโดยการนำข้อสอบในปีที่ผ่านมาไปใช้สอบซ้ำ จึงทำให้ข้อสอบมีการเผยแพร่ไปสู่รุ่นต่อไปได้ สอดคล้องกับงานวิจัย Sawangboon (2011) ได้ศึกษาการสร้างแบบทดสอบคู่ขนานตามรูปแบบพาเซทวิชาคณิตศาสตร์ เรื่องการแยกตัวประกอบพหุนามดีกรีสองขั้นมัธยมศึกษาปีที่ 2 โดยใช้แบบทดสอบคู่ขนานเพื่อแก้ปัญหาการจำข้อสอบของผู้เรียน จึงสร้างแบบทดสอบที่มีลักษณะเหมือนกันทั้งในเนื้อหาและค่าสถิติเพื่อนำคะแนนแต่ละครั้งมาเปรียบเทียบกันได้ ผู้สอนมีความสะดวกในการที่จะนำแบบทดสอบฉบับใด ๆ ไปทดสอบกับนักเรียนก่อนก็ได้ และสอดคล้องกับงานวิจัย Kanchanapenkul (2011) ได้ศึกษาการสร้างข้อสอบคู่ขนานอิงโดเมนโดยใช้รูปแบบพาเซท วิชาคณิตศาสตร์ เรื่องวงกลม ชั้นมัธยมศึกษาปีที่ 3 ได้สร้างแบบทดสอบคู่ขนานเพื่อนำมาสลับกันใช้ในแต่ละครั้งที่ดำเนินการสอบไม่ต้องออกข้อสอบใหม่ ทำให้ทุ่นเวลาและป้องกันข้อสอบรั่วได้ด้วย จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง พบว่า วิธีการที่สามารถแก้ปัญหาข้างต้นได้ คือ การสร้างแบบสอบคู่ขนาน (Parallel Test) ที่เป็นแบบสอบตั้งแต่ 2 ฉบับขึ้นไป วัดและประเมินในเรื่องเดียวกัน มีค่าความยากง่าย อำนาจจำแนก ความเชื่อมั่น สัมประสิทธิ์สหสัมพันธ์ภายใน ค่าเฉลี่ยและความแปรปรวนของคะแนนใกล้เคียงกัน ตลอดจนมีไค้ลักษณะของข้อสอบ และฟังก์ชันสารสนเทศของข้อสอบและแบบสอบเหมือนกัน (Karnjanawasee, 2012; Pasiphol, 2016) ทั้งนี้จากปัญหาของโรงเรียนที่เกิดขึ้นผู้วิจัยจึงสนใจศึกษาวิธีการแก้ปัญหาการใช้แบบสอบซ้ำโดยการสร้างแบบสอบคู่ขนานซึ่งเป็นแบบสอบที่วัดในเนื้อหาสาระของข้อสอบคุณลักษณะตามเกณฑ์มาตรฐาน

การสร้างแบบสอบคู่ขนาน สามารถแบ่งตามการบริหารการสอบได้เป็น 2 รูปแบบ คือ (1) การสร้างแบบสอบคู่ขนานหลังการสอบ ข้อดีของรูปแบบนี้ คือ ทำให้มีความน่าเชื่อถือและมีความเป็นคู่ขนานของข้อสอบและแบบสอบตามเกณฑ์ความเป็นคู่ขนานที่ชัดเจน (Lin, 2008) และ (2) การสร้างแบบสอบคู่ขนานก่อนการสอบ ข้อดีของรูปแบบนี้ คือ สามารถสร้างข้อสอบในสภาพจริงได้ง่ายเนื่องจากใช้ผู้เชี่ยวชาญที่มีอยู่ในสาขานั้น ๆ และสร้างข้อสอบใหม่ได้ตามโครงสร้างและจุดประสงค์ที่กำหนดโดยไม่ต้องอาศัยค่าสถิติหรือพารามิเตอร์ของข้อสอบเดิมจากคลังข้อสอบ แต่ข้อเสียของวิธีนี้ คือ การพิจารณาคุณภาพของแบบสอบโดยผู้เชี่ยวชาญ ขึ้นอยู่กับระดับ

ความสามารถในการตัดสินใจรายบุคคลของผู้เชี่ยวชาญที่แตกต่างกัน (Pasiphol, 2016) ด้วยเหตุนี้ผู้วิจัยจึงเลือกใช้รูปแบบที่ 1 การสร้างแบบสอบคู่ขนานหลังการสอบ

เทคนิคในการสร้างข้อสอบคู่ขนานมี 5 เทคนิค คือ (1) สลับคำตอบมาเป็นคำถาม เป็นการนำคำตอบของฟอร์ม 1 มาเป็นคำถามของฟอร์ม 2 (2) เปลี่ยนคำถามตรงกันข้ามหรือเป็นปฏิเสธ โดยสร้างคำถามของฟอร์ม 1 เป็นข้อความทางบวก แต่คำถามของฟอร์ม 2 เป็นทางลบ (3) ใช้คำถามเหมือนกันแต่ตัวเลือกแตกต่างกัน โดยสร้างคำถามที่เหมือนกันทั้งฟอร์ม 1 และฟอร์ม 2 แต่ตัวเลือกมีความแตกต่างกันทั้งหมด (4) ใช้คำถามต่างกันแต่ตัวเลือกเหมือนกัน และ (5) เปลี่ยนถ้อยคำในคำถามและคำตอบแต่ถามในประเด็นเดียวกัน ซึ่งใช้ข้อความที่ต่างกัน แต่มีนัยเดียวกันในฟอร์ม 1 และฟอร์ม 2 และตัวเลือกก็มีความแตกต่างกันทั้ง 2 ฟอร์ม (Pasiphol, 2020) จาก 5 เทคนิคนี้ ผู้วิจัยสนใจที่จะประยุกต์ใช้เทคนิคในการสร้างแบบสอบคู่ขนาน 3 เทคนิคคือ (1) สลับคำตอบมาเป็นคำถาม (2) ใช้คำถามเหมือนกันแต่ตัวเลือกแตกต่างกัน และ (3) ใช้คำถามต่างกันแต่ตัวเลือกเหมือนกัน เพราะหากใช้เปลี่ยนคำถามตรงกันข้ามหรือเป็นปฏิเสธ และเปลี่ยนถ้อยคำในคำถามและคำตอบแต่ถามในประเด็นเดียวกัน จะทำให้ข้อสอบมีความยากเกินไปสำหรับผู้เรียนในระดับชั้นประถมศึกษาปีที่ 2

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการตรวจสอบความเป็นคู่ขนานของแบบสอบที่สร้างขึ้น สามารถสรุปได้ 3 วิธี คือ (1) การตรวจสอบความเป็นคู่ขนานโดยอาศัยผู้เชี่ยวชาญ (2) การตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยข้อตกลงเบื้องต้นของทฤษฎีการทดสอบแบบดั้งเดิม กำหนดให้แบบสอบคู่ขนาน 2 ฉบับจะต้องมีคะแนนจริงของผู้สอบแต่ละคนและค่าความคลาดเคลื่อนมาตรฐานของประชากรที่สอบทั้ง 2 ฉบับเท่ากัน (Lord & Novick, 1968; Lord, 1980) โดยคะแนนจริงมีค่าเท่ากับคะแนนเฉลี่ยของคะแนนที่ได้จากการวัดซ้ำหลาย ๆ ครั้ง ซึ่งเป็นการพิจารณาค่าสถิติหลังจากการสอบ หากจะนำแบบสอบคู่ขนานนี้ไปใช้ในการสอบครั้งต่อไปต้องใช้แบบสอบซ้ำและความเที่ยงของแบบสอบคู่ขนานอาจเปลี่ยนไปตามประชากรที่แตกต่างกัน ซึ่งขาดสารสนเทศในการพิจารณาความเป็นคู่ขนานของข้อสอบรายข้อด้วย (Dorans et al., 2007) และ (3) การตรวจสอบความเป็นคู่ขนานโดยใช้ทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้ในการสร้างแบบสอบคู่ขนาน ซึ่งจากการศึกษางานวิจัยในอดีตที่ผ่านมา พบว่า มีการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบในการวิเคราะห์ความเป็นคู่ขนาน จำนวน 3 โมเดล ได้แก่ โมเดล 1 พารามิเตอร์ โมเดล 2 พารามิเตอร์ และ โมเดล 3 พารามิเตอร์ และเมื่อพิจารณาจากขนาดของกลุ่มตัวอย่างของนักเรียนที่อยู่ในสังกัดกรุงเทพมหานคร ในเขตหนองจอก จำนวน 800 คน ซึ่งเป็นจำนวนของกลุ่มตัวอย่างที่เหมาะสมจะวิเคราะห์โดยใช้โมเดล 2 พารามิเตอร์ในการตรวจสอบความเป็นคู่ขนาน เนื่องจาก Foley (2010) ได้กล่าวว่า ขนาดกลุ่มตัวอย่างที่เหมาะสม ลักษณะของตัวเลือก และจำนวนข้อสอบ สำหรับการวิเคราะห์โดยใช้โมเดลแบบ 2 พารามิเตอร์คือ 600 ถึง 1,200 คน

ดังนั้นจากสภาพปัญหาที่ได้กล่าวมาข้างต้น ผู้วิจัยจึงสนใจที่ศึกษาวิธีการแก้ปัญหาการใช้แบบสอบซ้ำ โดยการสร้างแบบสอบคู่ขนานวิชาคณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 2 สาระการวัด เนื่องจากเป็นสาระที่มีผลสัมฤทธิ์ทางการเรียนค่อนข้างต่ำ และเหตุผลในการเลือกเรื่องความยาว และน้ำหนัก เพราะนักเรียนเกิดความสับสนในหน่วยการวัด โดยรูปแบบที่ใช้ในการวิจัยเป็นการสร้างแบบสอบคู่ขนานหลังการสอบ โดยใช้เทคนิคในการสร้าง 3 เทคนิค (1) สลับคำตอบมาเป็นคำถาม (2) ใช้คำถามเหมือนกันแต่ตัวเลือกแตกต่างกัน และ (3) ใช้คำถามต่างกันแต่ตัวเลือกเหมือนกัน ส่วนการตรวจสอบคุณภาพของแบบสอบคู่ขนานประยุกต์ใช้โมเดลการตอบสนอง

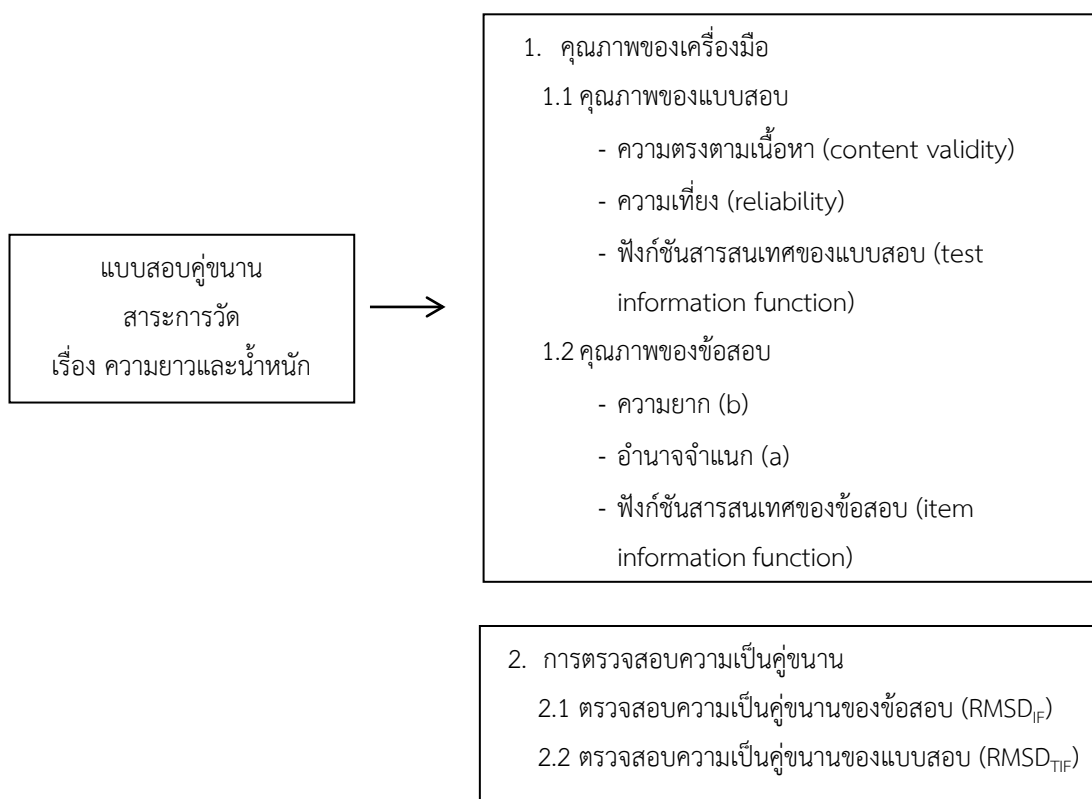
ข้อสอบ 2 พารามิเตอร์ ซึ่งประกอบไปด้วยการวิเคราะห์ คุณภาพของแบบสอบ คุณภาพของข้อสอบ ความเป็นคู่ขนานของข้อสอบ (RMSD_{IF}) และความเป็นคู่ขนานของแบบสอบ (RMSD_{TIF})

ความมุ่งหมายของการวิจัย

1. เพื่อพัฒนาแบบสอบคู่ขนานวิชาคณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 2 สาระการวัด
2. เพื่อตรวจสอบความเป็นคู่ขนานของแบบสอบคู่ขนานวิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 2 สาระการวัด

กรอบแนวคิดในการวิจัย

จากการศึกษาแนวคิด หลักการและงานวิจัยที่เกี่ยวข้องดังกล่าว สามารถกำหนดเป็นกรอบแนวคิด การวิจัย การพัฒนาแบบสอบคู่ขนานวิชาคณิตศาสตร์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนสังกัด กรุงเทพมหานครโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ ได้ดังภาพ 1



ภาพ 1 กรอบแนวคิดในการวิจัย

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 2 ปีการศึกษา 2562 จากโรงเรียน สังกัด กรุงเทพมหานคร ในเขตหนองจอก 37 โรงเรียน จำนวน 1,853 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนชั้นประถมศึกษาปีที่ 2 ปีการศึกษา 2562 จากโรงเรียน สังกัด กรุงเทพมหานคร ในเขตหนองจอก 13 โรงเรียน จำนวน 800 คน จำแนกตามขนาดโรงเรียนพบว่าโรงเรียนขนาดใหญ่พิเศษคิดเป็นร้อยละ 24.63 ขนาดใหญ่ 23.12 ขนาดกลาง 22.13 ขนาดเล็ก 30.12 และเมื่อจำแนกตามเพศพบว่าส่วนใหญ่เป็นเพศหญิงคิดเป็นร้อยละ 63.37 มากกว่าเพศชายคิดเป็นร้อยละ 36.63 ซึ่งขนาดตัวอย่างที่เหมาะสมสำหรับการวิเคราะห์โดยใช้โมเดลแบบ 2 พารามิเตอร์ คือ 600 ถึง 1,200 คน (Foley, 2010) โดยตัวอย่างที่ได้มาจากการสุ่มแบบแบ่งชั้น (stratified random sampling)

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยเป็นแบบสอบถามแบบเลือกตอบ 3 ตัวเลือกที่คู่ขนานกันในรายวิชาคณิตศาสตร์ สารการวัดที่มีการให้คะแนนแบบ 2 ค่า หรือ 0,1 ซึ่งแบบสอบถามในการวิจัยครั้งนี้ประกอบด้วยแบบสอบถามจำนวน 4 ฉบับ คือ 1A, 1B, 2A และ 2B โดยจำแนกตามเนื้อหา คือ (1) ความยาว จำนวน 2 ฉบับ คือ 1A และ 1B ฉบับละ 40 ข้อ และคัดเลือกเหลือฉบับละ 20 ข้อ และ (2) น้ำหนัก จำนวน 2 ฉบับ คือ 2A และ 2B ฉบับละ 40 ข้อ และคัดเลือกเหลือฉบับละ 20 ข้อ รายละเอียดดังตาราง 1

ตาราง 1 เครื่องมือที่ใช้ในการวิจัยจำแนกตามเนื้อหา

เนื้อหา	สร้าง (ข้อ)		คัดเลือก (ข้อ)	
	ชุดA	ชุดB	ชุดA	ชุดB
1.ความยาว (1A,1B)	40	40	20	20
2.น้ำหนัก (2A,2B)	40	40	20	20
รวม	80	80	40	40
รวมทั้งหมด	160		80	

3. การเก็บรวบรวมข้อมูล

1. นำแบบสอบถามคู่ขนาน เรื่องการวัด ทั้ง 4 ฉบับ ไปทดสอบกับกลุ่มตัวอย่างทั้ง 800 คน โดยอธิบายให้นักเรียนได้ทราบถึงวัตถุประสงค์ของการสอบในครั้งนี้และขอความร่วมมือในการทำแบบสอบ เพื่อให้นักเรียนได้เข้าใจถึงวิธีการทำแบบสอบ วิธีการตอบ ก่อนที่จะให้ทุกคนเริ่มต้นทำแบบสอบ โดยทำการสอบ ชุด 1A และ 2A ก่อน และผ่านไปหนึ่งสัปดาห์ แล้วสอบชุด 1B และ 2B

2. ตรวจให้คะแนน โดยมีเกณฑ์สำหรับการให้คะแนน คือ ตอบถูกให้ 1 คะแนน และตอบผิดหรือไม่ตอบให้ 0 คะแนน

3. นำผลที่ได้จากการทดสอบมาวิเคราะห์หาคุณภาพของข้อสอบ แบบสอบ และความเป็นคู่ขนานของข้อสอบและแบบสอบ

4. การวิเคราะห์ข้อมูล

4.1 การวิเคราะห์คุณภาพของแบบสอบ

คุณภาพของแบบสอบ พิจารณาจาก ความตรงตามเนื้อหา (content validity) โดยผู้เชี่ยวชาญจำนวน 3 คน เป็นผู้เชี่ยวชาญด้านการศึกษาคณิตศาสตร์จำนวน 2 คน และด้านการวัดและประเมินผล 1 คน ความเที่ยง

(reliability) ของแบบสอบ ใช้วิธีของ Kuder-Richardson (KR-20) และฟังก์ชันสารสนเทศของแบบสอบ (test information function) โดยใช้โปรแกรม MULTILOG 7.03

4.2 การวิเคราะห์คุณภาพของข้อสอบ

คุณภาพของข้อสอบ พิจารณาจาก ความยาก (b) อำนาจจำแนก (a) ฟังก์ชันสารสนเทศของข้อสอบ (item information function) โดยใช้โปรแกรม MULTILOG 7.03

4.3 การวิเคราะห์ความเป็นคู่ขนาน

การวิเคราะห์ความเป็นคู่ขนานรายข้อ พิจารณาจากค่าส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของข้อสอบ (Root mean square deviation of item information ; $RMSD_{IF}$) โดยจะถือว่าข้อสอบคู่ขนานกัน ค่า $RMSD_{IF}$ ควรมีค่าไม่เกิน 0.50 และการวิเคราะห์ความเป็นคู่ขนานของแบบสอบ พิจารณาจากค่าส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของแบบสอบ (Root mean square deviation of test information; $RMSD_{TIF}$) โดยจะถือว่าข้อสอบคู่ขนานกัน ค่า $RMSD_{TIF}$ ควรมีค่าไม่เกิน 0.50 มีสูตรการคำนวณดังนี้

การวิเคราะห์ความเป็นคู่ขนานรายข้อ

$$RMSD_{IF} = \sqrt{\frac{\sum_{j=1}^q [i(\theta_j) - i'(\theta_j)]^2}{q}}$$

การวิเคราะห์ความเป็นคู่ขนานของแบบสอบ

$$RMSD_{TIF} = \sqrt{\frac{\sum_{j=1}^q [I(\theta_j) - I'(\theta_j)]^2}{q}}$$

เมื่อ q แทน จำนวนระดับความสามารถที่ใช้ในการตรวจสอบ

$i(\theta)$ แทน ค่าสารสนเทศของข้อสอบฉบับ A ที่มีระดับความสามารถ θ

$i'(\theta)$ แทน ค่าสารสนเทศของข้อสอบฉบับ B ที่มีระดับความสามารถ θ

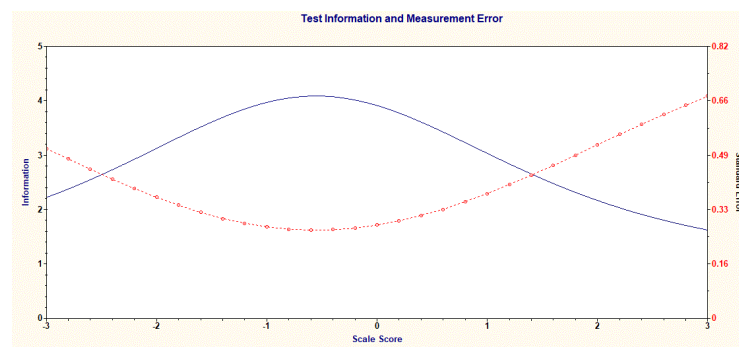
ผลการวิจัย

1. ผลการพัฒนาแบบสอบคู่ขนานวิชาคณิตศาสตร์เมื่อพิจารณาคุณภาพของแบบสอบมีดังนี้

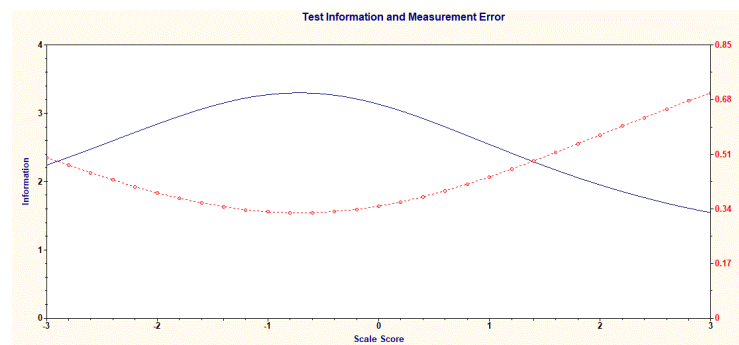
1.1 ความตรงตามเนื้อหาของแบบสอบคู่ขนาน ทั้ง 4 ฉบับ จำนวน 160 ข้อ มีค่า IOC (index of item objective congruence) ตั้งแต่ 0.67 ขึ้นไปถือว่าผ่านเกณฑ์ และได้คัดเลือกให้เหลือ 80 ข้อ

1.2 ความเที่ยงของแบบสอบทั้ง 4 ฉบับที่ผ่านการคัดเลือกเหลือ 80 ข้อ พบว่า ฉบับ 1A มีค่า 0.711, ฉบับ 1B มีค่า 0.652, ฉบับ 2A มีค่า 0.718 และ ฉบับ 2B มีค่า 0.635 ซึ่งถือว่ามีความเที่ยงในระดับปานกลาง

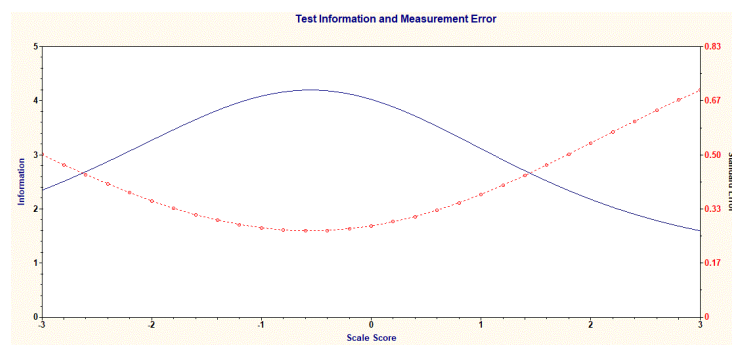
1.3 ค่าฟังก์ชันสารสนเทศของแบบสอบ (Test Information Function) ณ ช่วงความสามารถอยู่ระหว่าง -3 ถึง 3 พบว่า (1) แบบสอบฉบับ 1A มีค่าสูงสุด เท่ากับ 4.083 ณ ระดับความสามารถ เท่ากับ -0.6 (2) แบบสอบฉบับ 1B มีค่าสูงสุด เท่ากับ 3.293 ณ ระดับความสามารถ เท่ากับ -0.8 (3) แบบสอบฉบับ 2A มีค่าสูงสุด เท่ากับ 4.106 ณ ระดับความสามารถ เท่ากับ -0.4 และ (4) แบบสอบฉบับ 2B มีค่าสูงสุด เท่ากับ 3.498 ณ ระดับความสามารถ เท่ากับ -0.6 โดยมีผลการวิเคราะห์ฟังก์ชันสารสนเทศของแบบสอบ ดังภาพ 2 - 5



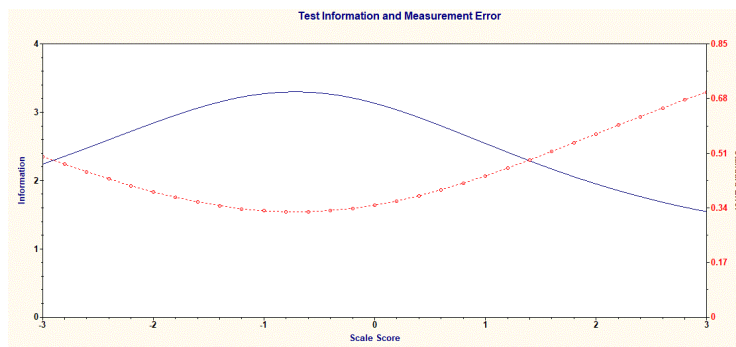
ภาพ 2 ฟังก์ชันสารสนเทศของแบบทดสอบ ฉบับ 1A



ภาพ 3 ฟังก์ชันสารสนเทศของแบบทดสอบ ฉบับ 1B



ภาพ 4 ฟังก์ชันสารสนเทศของแบบทดสอบ ฉบับ 2A



ภาพ 5 ฟังก์ชันสารสนเทศของแบบทดสอบ ฉบับ 2B

1.2 ผลการพัฒนาแบบสอบคู่ขนานวิชาคณิตศาสตร์เมื่อพิจารณาคุณภาพของข้อสอบมีดังนี้

1.2.1 ผลการวิเคราะห์คุณภาพรายข้อ เรื่องความยาว ฉบับ 1A มีค่าความยาก (b) อยู่ระหว่าง -2.16 ถึง 0.57 ค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.30 ถึง 1.13 แปลผลได้ว่า ข้อสอบผ่านเกณฑ์ทุกข้อ ส่วนแบบสอบ ฉบับ 1B มีค่าความยาก (b) อยู่ระหว่าง -2.98 ถึง -0.29 ค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.23 ถึง 1.06 ซึ่งมีข้อสอบ ไม่ผ่านเกณฑ์ 2 ข้อ คือ ข้อที่ 1 มีค่าความยากเท่ากับ -2.98 ข้อที่ 20 มีค่าอำนาจจำแนกน้อยกว่า 0.3

1.2.2 ผลการวิเคราะห์คุณภาพรายข้อ เรื่องน้ำหนัก ฉบับ 2A มีค่าความยาก (b) อยู่ระหว่าง -1.81 ถึง 0.00 ค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.30 ถึง 1.08 แปลผลได้ว่า ข้อสอบผ่านเกณฑ์ทุกข้อ ส่วนแบบสอบ ฉบับ 2B มีค่าความยาก (b) อยู่ระหว่าง -2.90 ถึง -0.73 ค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.22 ถึง 0.86 ซึ่งมีข้อสอบที่ไม่ผ่านเกณฑ์จำนวน 5 ข้อ คือ ข้อที่ 1, 4, 9, 19 และ 20 โดยข้อที่ 1 และข้อที่ 20 มีค่าความยาก เท่ากับ -2.90 และ -2.77 ตามลำดับ ส่วนข้อที่ 4, 9, 19 และ 20 มีค่าอำนาจจำแนกน้อยกว่า 0.3

1.2.3 ผลการวิเคราะห์ค่าฟังก์ชันสารสนเทศของข้อสอบ (item information function) ณ ช่วง ความสามารถอยู่ระหว่าง -3 ถึง 3

แบบสอบฉบับ 1A จำนวน 20 ข้อ โดยข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบมากที่สุดคือ ข้อ 10 มีค่า 0.320 และข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบน้อยที่สุด คือ ข้อ 4 มีค่า 0.023

แบบสอบฉบับ 1B โดยข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบมากที่สุดคือ ข้อ 17 มีค่า 0.281 และ ข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบน้อยที่สุด ข้อ 20 มีค่า 0.013

แบบสอบฉบับ 2A โดยข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบมากที่สุดคือ ข้อ 14 มีค่า 0.291 และ ข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบน้อยที่สุด คือ ข้อ 20 มีค่า 0.022

แบบสอบฉบับ 2B โดยข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบมากที่สุดคือ ข้อ 5 มีค่า 0.186 และ ข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบน้อยที่สุด คือ ข้อ 4 มีค่า 0.013

2. ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบคู่ขนานวิชาคณิตศาสตร์ระดับชั้นประถมศึกษา ปีที่ 2 สาระการวัด

2.1 ผลการวิเคราะห์ความเป็นคู่ขนานรายข้อและทั้งฉบับของแบบสอบเรื่อง ความยาว จำนวน 2 ฉบับ (1A และ 1B) ฉบับละ 20 ข้อ พบว่า ผลการตรวจสอบความเป็นคู่ขนานรายข้อเมื่อพิจารณาจากค่า $RMSD_{IF}$ มีค่า

ตั้งแต่ 0.008 ถึง 0.167 โดยค่า $RMSD_{IF}$ ไม่เกิน 0.50 สะท้อนให้เห็นว่าข้อสอบทั้ง 20 ข้อ มีความเป็นคู่ขนานกัน ขณะที่ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบทั้งฉบับเมื่อพิจารณาจากค่า $RMSD_{TIF}$ มีค่าเท่ากับ 0.484 โดยค่า $RMSD_{TIF}$ มีค่าไม่เกิน 0.50 สะท้อนให้เห็นว่าแบบสอบฉบับ 1A และ 1B มีความเป็นคู่ขนานกัน

2.2 ผลการวิเคราะห์ความเป็นคู่ขนานรายข้อและทั้งฉบับของแบบสอบเรื่อง น้ำหนัก จำนวน 2 ฉบับ (2A และ 2B) ฉบับละ 20 ข้อ พบว่า ผลการตรวจสอบความเป็นคู่ขนานรายข้อเมื่อพิจารณาจากค่า $RMSD_{IF}$ มีค่าตั้งแต่ 0.020 ถึง 0.110 โดยค่า $RMSD_{IF}$ ไม่เกิน 0.50 สะท้อนให้เห็นว่าข้อสอบทั้ง 20 ข้อ มีความเป็นคู่ขนานกันขณะที่ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบทั้งฉบับเมื่อพิจารณาจากค่า $RMSD_{TIF}$ มีค่าเท่ากับ 0.492 โดยค่า $RMSD_{TIF}$ มีค่าไม่เกิน 0.50 สะท้อนให้เห็นว่าแบบสอบฉบับ 2A และ 2B มีความเป็นคู่ขนานกัน รายละเอียดดังตาราง 2

ตาราง 2 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบคู่ขนานวิชาคณิตศาสตร์ระดับชั้นประถมศึกษา
ปีที่ 2 สาระ การวัด

การตรวจสอบความเป็นคู่ขนาน	รายข้อ ($RMSD_{IF}$)	ทั้งฉบับ ($RMSD_{TIF}$)	การแปลผล
แบบสอบเรื่อง ความยาว (ฉบับ 1A, 1B)	0.008 – 0.167	0.484	มีความคู่ขนานทั้งรายข้อและทั้งฉบับ
แบบสอบเรื่อง น้ำหนัก (ฉบับ 2A, 2B)	0.020 – 0.110	0.492	มีความคู่ขนานทั้งรายข้อและทั้งฉบับ

อภิปรายผล

1. อภิปรายผลการพัฒนาแบบสอบคู่ขนานวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 2 สาระการวัด

จากผลการวิจัยที่พบว่า ความตรงตามเนื้อหาของแบบสอบคู่ขนาน ทั้ง 4 ฉบับ จำนวน 160 ข้อ มีค่า IOC ตั้งแต่ 0.67 สะท้อนให้เห็นว่าแบบสอบทั้ง 4 ฉบับ ประกอบด้วยข้อสอบที่มีความตรงตามเนื้อหา ทั้งนี้เพราะผู้วิจัยได้ศึกษาโครงสร้างและการสร้างกรอบการประเมินที่ชัดเจน เมื่อนำมาให้ผู้เชี่ยวชาญพิจารณาประเมิน เครื่องมือจึงมีความตรงตามเกณฑ์ที่กำหนด ซึ่งค่า IOC ที่ได้สอดคล้องกับเกณฑ์ที่ใช้ในการตัดสินความตรงตามเนื้อหาควรมีค่าตั้งแต่ 0.50 ขึ้นไป (Karnjanawasee, 2013)

จากผลการวิเคราะห์ค่าความเที่ยงของแบบสอบทั้ง 4 ฉบับที่ผ่านการคัดเลือกเหลือ 80 ข้อ พบว่า ฉบับ 1A มีค่า 0.711, ฉบับ 1B มีค่า 0.652, ฉบับ 2A มีค่า 0.718 และ ฉบับ 2B มีค่า 0.635 ซึ่งมีความเที่ยงในระดับปานกลาง สอดคล้องกับ Katsa & Amonrattanasak (2008); Karnjanawasee (2012) ที่กล่าวว่า แบบสอบที่ใช้ในการทดสอบ ควรมีค่าสัมประสิทธิ์ความเที่ยงสูงที่สุดเท่าที่จะเป็นไปได้ภายใต้สภาพการณ์นั้น อย่างน้อยที่สุดควรมีค่าไม่ต่ำกว่า 0.5 แต่อย่างไรก็ดีแม้ว่าค่าความเที่ยงของแบบสอบในการวิจัยครั้งนี้จะผ่านเกณฑ์แต่ก็ยังไม่สูงมากนัก ทั้งนี้อาจเป็นผลมาจากความยาวของข้อสอบซึ่งประกอบไปด้วยข้อสอบฉบับละ 20 ข้อ ซึ่งความยาวของข้อสอบ จำนวนตัวเลือก 3 ตัวเลือก และข้อสอบชุด B จะมีความเที่ยงต่ำกว่าชุด A อาจเป็นเพราะชุด B ถูกจัดให้ทำหลังชุด A กลุ่มตัวอย่างที่ทำข้อสอบอาจจะเกิดความเหนื่อยล้าในการทำแบบสอบ ซึ่งน่าจะเป็นเป็นปัจจัยที่ส่งผลต่อค่าความเที่ยง

จากผลการวิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบสอบ (test information function) ณ ช่วงความสามารถอยู่ระหว่าง -3 ถึง 3 พบว่า แบบสอบฉบับ 1A มีค่าสูงสุด เท่ากับ 4.083 ณ ระดับความสามารถ เท่ากับ -0.60 และแบบสอบฉบับ 1B มีค่าสูงสุด เท่ากับ 3.293 ณ ระดับความสามารถ เท่ากับ -0.80 สะท้อนให้เห็นว่าแบบสอบ 1A และ 1B ให้สารสนเทศสูงสุดสำหรับผู้ที่มีระดับความสามารถระหว่าง -0.60 ถึง -0.80 หรือแบบสอบทั้ง 2 ฉบับนี้เหมาะสำหรับผู้สอบกลุ่มที่มีความสามารถปานกลางไปจนถึงค่อนข้างอ่อน ขณะที่แบบสอบฉบับ 2A มีค่าสูงสุด เท่ากับ 4.106 ณ ระดับความสามารถ เท่ากับ -0.40 และแบบสอบฉบับ 2B มีค่าสูงสุด เท่ากับ 3.498 ณ ระดับความสามารถ เท่ากับ -0.60 สะท้อนให้เห็นว่าแบบสอบ 2A และ 2B ให้สารสนเทศสูงสุดสำหรับผู้ที่มีระดับความสามารถระหว่าง -0.40 ถึง -0.60 หรือแบบสอบทั้ง 2 ฉบับนี้เหมาะสำหรับผู้สอบกลุ่มที่มีความสามารถปานกลางไปจนถึงค่อนข้างอ่อน ซึ่งในการวิจัยครั้งนี้จะสังเกตได้ว่า ค่าฟังก์ชันสารสนเทศของแบบสอบทั้ง 2 ฉบับมีค่าใกล้เคียงกัน เกิดจากขั้นตอนการสร้างแบบสอบได้มีการศึกษาการสร้างตามโครงสร้างจำแนกระดับพฤติกรรมตรงตามตัวชี้วัดในเนื้อหาการวัดของนักเรียนระดับชั้นประถมศึกษาปีที่ 2 สะท้อนให้เห็นว่าการสร้างข้อสอบครั้งนี้ที่เป็นคู่ขนานทั้งเนื้อหา ค่าความยาก อำนาจจำแนกที่ใกล้เคียงกัน สอดคล้องกับงานวิจัยของ Kantivong (2017); Ali & Rijn (2016) ที่กล่าวว่าค่าสารสนเทศของแบบสอบเกิดจากประสบการณ์และความเชี่ยวชาญในการออกข้อสอบเกิดจากการสร้างแบบสอบให้มีคุณภาพความตรงความถูกต้องสูงขึ้น

จากผลการวิเคราะห์ค่าความยาก (b) ของข้อสอบคู่ขนานวิชาคณิตศาสตร์ สารการวัด ทั้ง 4 ฉบับจำนวน 80 ข้อ พบว่า ค่าความยาก มีค่าอยู่ระหว่าง -2.98 ถึง 0.00 ซึ่งไม่ผ่านเกณฑ์จำนวน 3 ข้อ จากข้อคำถามชุด B เพราะเป็นคำถามที่วัดพฤติกรรมด้านความจำ ดังนั้นต้องสร้างแบบสอบที่วัดพฤติกรรมที่สูงกว่าความจำ ส่วนข้อสอบที่มีความยากผ่านเกณฑ์แต่มีแนวโน้มว่าข้อสอบยากง่ายปานกลางค่อนข้างไปทางง่าย โดยสอดคล้องกับ Karnjanawasee (2013) ที่กล่าวว่า ข้อสอบที่ดีควรมีค่าความยากง่ายอยู่ในระดับปานกลางหรืออยู่ในช่วง -2.50 ถึง 2.50

จากผลการวิเคราะห์ค่าอำนาจจำแนก (a) ของข้อสอบคู่ขนานวิชาคณิตศาสตร์ สารการวัด ทั้ง 4 ฉบับจำนวน 80 ข้อ พบว่า ค่าอำนาจจำแนก มีค่าอยู่ระหว่าง 0.22 ถึง 1.13 จากข้อสอบทั้งหมดที่ได้มีการวิเคราะห์แล้วนั้น มีข้อสอบจำนวน 5 ข้อที่มีค่าอำนาจจำแนกน้อยกว่า 0.3 สอดคล้องกับงานวิจัยของ Sawangsri (2015) ที่พบว่า ค่าอำนาจจำแนกที่ผ่านเกณฑ์ควรมากกว่า 0.3 เพราะข้อสอบที่มีค่าน้อยกว่านี้แสดงว่าข้อสอบนั้นไม่ดีและควรตัดออก ส่วนข้อสอบที่มีค่าอำนาจจำแนกผ่านเกณฑ์ 0.3 ขึ้นไปจำนวน 75 ข้อ สอดคล้องกับ Reckase (2009) ที่กล่าวว่าค่าอำนาจจำแนกควรมีค่าอยู่ระหว่าง 0.30 ถึง 2.00 และสอดคล้องกับ Karnjanawasee (2012) ที่ระบุว่าค่าอำนาจจำแนก (a) ควรมีค่าเป็นบวกและในทางปฏิบัติควรมีค่าระหว่าง 0.50 ถึง 2.50

จากผลการวิเคราะห์ค่าฟังก์ชันสารสนเทศของข้อสอบ (item information function) ณ ช่วงความสามารถอยู่ระหว่าง -3 ถึง 3 แบบสอบฉบับ 1A ข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบมากที่สุดคือ ข้อ 10 และน้อยที่สุดคือ ข้อ 4 แบบสอบฉบับ 1B ข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบมากที่สุดคือ ข้อ 17 และน้อยที่สุด คือ ข้อ 20 แบบสอบฉบับ 2A ข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบมากที่สุดคือ ข้อ 14 และน้อยที่สุด คือ ข้อ 20 และแบบสอบฉบับ 2B ข้อที่ค่าฟังก์ชันสารสนเทศของข้อสอบมากที่สุดคือ ข้อ 5 และน้อยที่สุด คือ ข้อ 4 ทั้งนี้ข้อที่มีค่าฟังก์ชันสารสนเทศของข้อสอบน้อยที่สุดนั้นจะสัมพันธ์กับค่าความยากและอำนาจจำแนกที่ไม่ผ่านเกณฑ์ด้วย เช่น ข้อ 20 ของฉบับ 1B ที่พบว่าค่าฟังก์ชันสารสนเทศของข้อสอบน้อยที่สุด ปรากฏว่าเป็นข้อที่มีอำนาจจำแนก

น้อยกว่า 0.3 ด้วย ทั้งนี้สอดคล้องกับ Karnjanawasee (2012) ที่ได้กล่าวว่าฟังก์ชันสารสนเทศข้อสอบเป็นดัชนีผสมที่สร้างจากดัชนีคุณลักษณะของข้อสอบหลายลักษณะประกอบด้วยค่าพารามิเตอร์ความยาก อำนาจจำแนก และค่าความแปรปรวนของคะแนนรายข้อเพื่อบ่งชี้คุณภาพข้อสอบ สอดคล้องกับงานวิจัยของ Kantivong (2017); Sawangsri (2015) ซึ่งประยุกต์ใช้โมเดลแบบ 2 พารามิเตอร์ ที่มีการพิจารณาค่าความยาก และค่าอำนาจจำแนกของข้อสอบซึ่งสัมพันธ์กับค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ

2.อภิปรายผลการตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบ วิชาคณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 2 สาระการวัด

จากผลการวิเคราะห์ความเป็นคู่ขนานรายข้อและทั้งฉบับของแบบสอบเรื่อง ความยาว จำนวน 2 ฉบับ (1A และ 1B) ที่พบว่า ข้อสอบทั้ง 20 ข้อ มีความเป็นคู่ขนานกัน และแบบสอบทั้ง 2 ฉบับ คือ ฉบับ 1A และ 1B คู่ขนานกัน และความเป็นคู่ขนานรายข้อและทั้งฉบับของแบบสอบเรื่อง น้ำหนัก จำนวน 2 ฉบับ (2A และ 2B) ที่พบว่า ข้อสอบทั้ง 20 ข้อ มีความเป็นคู่ขนานกัน และแบบสอบทั้ง 2 ฉบับ คือ ฉบับ 2A และ 2B คู่ขนานกัน สะท้อนให้เห็นว่า ข้อสอบที่ผู้วิจัยสร้างขึ้นนั้นมีความเป็นคู่ขนานในรายข้อส่งผลให้เกิดความคู่ขนานในภาพรวม คือความเป็นคู่ขนานของแบบสอบทั้งฉบับ สอดคล้องกับงานวิจัยของ Kantivong (2017); Debeer et al. (2017) ที่ให้ความสำคัญต่อความเป็นคู่ขนานรายข้อ เพราะจะนำไปสู่ผลรวมของความเป็นคู่ขนานของแบบสอบทั้งฉบับ ซึ่งหากแบบสอบ 2 ฉบับที่คู่ขนานกันในรายข้อมีโอกาที่แบบสอบคู่นั้นจะคู่ขนานกันสูงขึ้น

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 ได้แบบสอบคู่ขนานวิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 2 สาระการวัด ที่ผ่านการตรวจสอบความเที่ยง ความตรงตามเนื้อหา ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบที่ผ่านเกณฑ์จำนวน 75 ข้อ ซึ่งสามารถนำไปใช้ในการจัดสร้างคลังข้อสอบวิชาคณิตศาสตร์ สาระการวัด เรื่องความยาวและน้ำหนัก สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 เพื่อใช้ในการวัดและประเมินผู้เรียนก่อนเรียนและหลังเรียนได้

1.2 จากผลการวิเคราะห์ค่าความเที่ยงของแบบสอบทั้ง 4 ฉบับ ซึ่งมีค่า 0.711, 0.652, 0.718 และ 0.635 โดยมีค่าความเที่ยงในระดับปานกลาง ซึ่งอาจเป็นผลมาจากความยาวของข้อสอบซึ่งประกอบไปด้วยข้อสอบฉบับละ 20 ข้อ ซึ่งความยาวของข้อสอบ เป็นปัจจัยหนึ่งที่ส่งผลต่อค่าความเที่ยง ดังนั้นในการนำผลวิจัยไปใช้อาจพิจารณาเรื่องจำนวนของข้อสอบ เพื่อให้เป็นแนวทางการสร้างแบบสอบคู่ขนานให้มีความเที่ยงมากยิ่งขึ้น

2. ข้อเสนอแนะสำหรับการทำวิจัยครั้งต่อไป

2.1 การวิจัยครั้งนี้ ผู้วิจัยได้นำแบบสอบคู่ขนานไปใช้ในบริบทของการทำแบบสอบกับกระดาษ (paper-based testing) ซึ่งผลการวิจัยที่ได้อาจจะแตกต่างจากการทดสอบแบบออนไลน์ ดังนั้นในการวิจัยครั้งต่อไป ควรมีการสร้างแบบสอบในรูปแบบออนไลน์ เพื่อให้สอดคล้องกับบริบทและสถานการณ์การจัดการเรียนการสอนของยุคปกติใหม่

2.2 การวิจัยครั้งนี้ ผู้วิจัยวิเคราะห์ค่าพารามิเตอร์โดยใช้ทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ ในการวิจัยครั้งต่อไปควรมีการวิเคราะห์ค่าโอกาสในการเดา (c) โดยการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ

แบบ 3 พารามิเตอร์ เนื่องจากจะได้ข้อสอบแต่ละข้อที่มีความแตกต่างกันทั้งพารามิเตอร์ a, b และ c เพื่อลดความคลาดเคลื่อนของข้อสอบให้ได้มากที่สุด ซึ่งทำให้มีข้อสอบที่มีคุณภาพมากยิ่งขึ้น

2.3 การวิจัยในครั้งนี้เป็นการสร้างและพัฒนาแบบสอบคู่ขนานที่เป็นข้อสอบปรนัยทั้งหมด ดังนั้น ในการวิจัยครั้งต่อไปควรมีการสร้างและพัฒนาแบบสอบคู่ขนานในรูปแบบอัตนัยและแบบรูปผสม เพื่อให้ได้คลังข้อสอบของแบบสอบคู่ขนานที่มีความหลากหลายและมีคุณภาพสำหรับนำไปใช้ในโอกาสที่แตกต่างกัน

2.4 การวิจัยครั้งนี้ใช้เทคนิคในการสร้างข้อสอบ 3 เทคนิค ในการวิจัยครั้งต่อไป ควรมีการศึกษาเพิ่มเติมว่าเทคนิคการสร้างข้อสอบเทคนิคใดมีความเหมาะสมเพื่อยกระดับความเป็นคู่ขนานของแบบสอบ

References

- Ali, U. S., & Rijn, P. W. (2016). An evaluation of different statistical target for assembling parallel form in item response theory. *Applied psychological measurement*, 40(3), 163-179.
- Debeer, D., Ali, U. S., & Rijn, P. W. (2017). Evaluating Statistical Targets for Assembling Parallel Mixed-Format Test Forms. *Journal of Educational Measurement*, 54(2), 218-242.
- Dorans, N. J., Pommerich, M., & Holland, P. W. (2007). *Linking and aligning scores and scales*. Springer-Verlag.
- Foley, B. F. (2010). *Improving IRT item parameter estimates with small sample sizes: Evaluating the efficacy of a new data augmentation technique* [Doctoral dissertation]. University of Nebraska.
- Lin, C. J. (2008). Comparisons between classical theory and item response theory in automated assembly of parallel test forms. *The journal of technology, learning and assessment*.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Erlbaum.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley.
- Reckase, M. (2009). *Multidimensional item response theory*. Springer.
- Rupp, A., & Templin, J. (2008). *Effects of Q-matrix misspecification on parameter estimates and misclassification rates in the DINA model*. Educational and Psychological Measurement.

Translate Thai References

- Education Strategy Office Education Bureau. (2016). *Bangkok Basic Education Development Plan issue 2 (2017-2020)*. Agricultural Co-operative Federation of Thailand Press. (in Thai)
- Kanchanapenkul, S. (2011). *A Construction of Parallel Domainre-referenced test on "Circles" by Using Facet Design for Matthayousuksa 3* [Master's thesis]. Srinakharinwirot University. (in Thai)

- Kantivong, A. (2017). *Comparison of Item and Test Parallel Index Between Test Forms that Selected by Different Characteristics of Experts: Application of Formal Cultural Consensus Theory* [Master's thesis]. Chulalongkorn University. (in Thai)
- Karnjanawasee, S. (2012). *New Exam Theory* (4th Ed.). Chulalongkorn University Press. (in Thai)
- Karnjanawasee, S. (2013). *Classical Test Theory* (7th Ed.). Chulalongkorn University Press. (in Thai)
- Katsa, T., & Amonrattanasak, S. (2008). *Educational evaluation*. Ramkhamhaeng University. (in Thai)
- Pasiphol, S. (2016). *Learning Measurement and Evaluation*. Chulalongkorn University Press. (in Thai)
- Pasiphol, S. (2020). *The Development of Parallel Tests: Creation and Parallelism analysis*. Chulalongkorn University Press. (in Thai)
- Sawangboon, N. (2011). *A Construction of Mathematics Parallel Test Using Facet Design on "Factoring of Degree-2 Polynomials" for Matthayomsuksa 2 Students* [Master's thesis]. Mahasarakham University. (in Thai)
- Sawangsri, P. (2015). *Comparison of Psychometric Properties Integrating of Reading Analytical Thinking and Writing Abilities and Indicators of Content Science of Nine Grade Using The Different Item Review Methods* [Master's thesis]. Chulalongkorn University. (in Thai)

กลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญา
ของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับ
โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2
School Management Strategy to Develop Students' Life Skills According to
the Philosophy of Sufficiency Economy by Using the Professional Learning
Community Process of Schools under the Office of Udon Thani Primary
Education Service Area 2

สุธรรม ธรรมทัศนานนท์^{1*}

Sutum Thummatasananon¹

(Received: March 22, 2022; Revised: May 13, 2022; Accepted: May 20, 2022)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมายเพื่อ 1) ศึกษาสภาพปัจจุบัน สภาพที่พึงประสงค์ และความต้องการจำเป็นของการบริหารสถานศึกษา เพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 กลุ่มตัวอย่าง ได้แก่ ผู้บริหารสถานศึกษาและครูหัวหน้าฝ่ายวิชาการ จำนวน 178 คน โดยการสุ่มแบบแบ่งชั้น 2) กำหนดกลยุทธ์ฯ กลุ่มผู้ให้ข้อมูล จำนวน 9 คน 3) ประเมินกลยุทธ์ฯ กลุ่มผู้ให้ข้อมูล ได้แก่ ผู้ทรงคุณวุฒิ จำนวน 9 คน เลือกแบบเจาะจง เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือ แบบสอบถาม มีค่าความเที่ยงตรง ตั้งแต่ 0.80-1.00 มีค่าความเชื่อมั่นของแบบสอบถามสภาพปัจจุบัน เท่ากับ 0.97 มีค่าความเชื่อมั่นของแบบสอบถามสภาพที่พึงประสงค์ เท่ากับ 0.98 แบบสัมภาษณ์ แบบบันทึกการสนทนากลุ่ม และแบบประเมินกลยุทธ์ฯ มีค่าความเที่ยงตรง ตั้งแต่ 0.80-1.00 ด้านความเหมาะสมมีค่าความเชื่อมั่น เท่ากับ 0.95 ด้านความเป็นไปได้ มีค่าความเชื่อมั่น เท่ากับ 0.97 สถิติที่ใช้ในการวิจัย ได้แก่ การแจกแจงความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และวิเคราะห์ลำดับความสำคัญของความต้องการจำเป็นด้วยค่าดัชนี

PNI_{Modified}

ผลการวิจัย พบว่า

1. สภาพปัจจุบันของการบริหารสถานศึกษา เพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ โดยรวมอยู่ในระดับปานกลาง สภาพที่พึงประสงค์โดยรวมอยู่ในระดับมาก ความต้องการจำเป็นในการกำหนดกลยุทธ์ฯ เรียงลำดับ ดังนี้ ด้านการ

¹ รองศาสตราจารย์ สาขาวิชาการบริหารการศึกษา มหาวิทยาลัยศรีปทุม วิทยาเขตขอนแก่น

¹ Associate Professor, Educational Administration Program, Sripatum University, Khon Kaen Campus

* Corresponding Author Email: Sutum011728@gmail.com

บริหารสถานศึกษาตามหลักปรัชญาของเศรษฐกิจพอเพียง ด้านการพัฒนาบุคลากรของสถานศึกษา ด้านหลักสูตรการพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง ด้านการจัดกิจกรรมพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง และด้านความสำเร็จ ตามลำดับ

2. กลยุทธ์ที่กำหนดขึ้น ประกอบด้วย 4 กลยุทธ์หลัก ดังนี้ 1) กลยุทธ์พัฒนาความสามารถทางวิชาการ 2) กลยุทธ์ประยุกต์ใช้ภูมิปัญญาท้องถิ่น 3) กลยุทธ์บูรณาการใช้แหล่งเรียนรู้ภายในและภายนอกโรงเรียน และ 4) กลยุทธ์เสริมสร้างประสบการณ์
3. ผลการประเมินกลยุทธ์ มีความเหมาะสมและความเป็นไปได้โดยรวมอยู่ในระดับมาก

คำสำคัญ: กลยุทธ์ การบริหารสถานศึกษา ปรัชญาของเศรษฐกิจพอเพียง ชุมชนการเรียนรู้ทางวิชาชีพ

Abstract

The objectives of this research were: 1) to study the current condition, desirable condition and needs of school management to develop students' life skills according to the Philosophy of Sufficiency Economy by using the professional learning community process, for the schools under the Office of Udon Thani Primary Education Service Area 2. The sample consisted of 178 administrators and academic head teachers, obtained through stratified random sampling; 2) to set strategies, having 9 informants; 3) to verify appropriateness of the strategies, having 9 experts, selected through purposive sampling.

The tools used for data collection comprised 2 questionnaires, having the validity from 0.80 to 1.00. One questionnaire was for the current condition, with the reliability of 0.97. The other was for the desirable condition, with the reliability of 0.98. The tools also included an interview form, a focus group discussion record, and a strategy assessment form, having the validity from 0.80 to 1.00. For the suitability aspect, the reliability was 0.95. For the possibility, the reliability was 0.97. The analysis of data employed descriptive statistics: frequency distribution, percentage, mean, standard deviation. The analysis of priorities of the needs employed the PNI ^{Modified}.

The results were as follows:

1. The current condition of school management to develop students' life skills according to the Philosophy of Sufficiency Economy by using the professional learning community process, on the whole, was at the moderate level. The overall desirable condition was at the high level. The needs in setting strategies were arranged in the following order: school administration according to the Philosophy of Sufficiency Economy, educational institution's personnel development, curriculum for developing students' life skills according

to the Philosophy of Sufficiency Economy, activities to develop students' life skills according to the Philosophy of Sufficiency Economy, and success, respectively.

2. The four strategies set for school management consisted of 4 main strategies as follows: 1) strategy for developing academic ability, 2) strategy for applying local wisdom, 3) strategy for integrating learning resources inside and outside the school, and 4) strategy for enhancing experiences.

3. The strategy evaluation results revealed that the suitability and possibility, on the whole, were at the high level.

Keywords: strategy, educational institution management, Philosophy of Sufficiency Economy, professional learning community

บทนำ

การเปลี่ยนแปลงทางสังคมสู่สังคมฐานเศรษฐกิจในยุคปัจจุบันก่อให้เกิดผลกระทบต่อ สภาพเศรษฐกิจ การเมือง ความเสื่อมของวัฒนธรรมและศีลธรรม ระบบการศึกษาที่มีในปัจจุบัน ขาดบทบาทเพียงพอในการ เตรียมคนและพัฒนาศักยภาพของคนให้พร้อมรับกับการเปลี่ยนแปลง สู่สังคมเศรษฐกิจได้ประชากร ขาดทักษะ ในการคิดและไม่สามารถเชื่อมโยงประยุกต์ใช้ความรู้เพื่อการดำเนินชีวิตได้อย่างมีประสิทธิภาพ สิ่งเหล่านี้ ล้วนก่อให้เกิดปัญหาทางสังคม ตั้งแต่ปัญหาส่วนตัวไปจนถึงครอบครัว ชุมชน และประเทศ ซึ่งเป็นปัญหาที่มี ระดับธรรมดาไปจนถึงขั้นซับซ้อน ทั้งนี้การพัฒนาประชากรนั้นจำเป็นต้องใช้ความรู้ ความคิด ทักษะจากหลาย สาขามาช่วยในการแก้ไขปัญหาและพัฒนาเพื่อให้มีทักษะชีวิตที่สมบูรณ์มากยิ่งขึ้น (Ministry of Education, 2005; Office of the Basic Education Commission, 2004)

การริเริ่มพัฒนาและนำการจัดการศึกษาทักษะชีวิตเข้าสู่โรงเรียนนั้นได้นำไปใช้ใน หลายประเทศ ทั่วโลกโดยความต้องการจำเป็นของการจัดการศึกษานั้นเป็นไปเพื่อการเผชิญความ ท้าทายในชีวิตประจำวัน เช่น ปัญหาโรคเอดส์ การตั้งครรภ์ในวัยรุ่น ความรุนแรง เป็นต้น ในหลาย ประเทศได้มีการพิจารณาการ พัฒนาการจัดการศึกษาทักษะชีวิตในการปฏิรูปการศึกษาโดยสนอง ต่อสังคมสมัยใหม่และชีวิตยุคเศรษฐกิจ การจัดการศึกษาเพื่อพัฒนาทักษะชีวิตนั้นได้ออกแบบเพื่ออำนวยความสะดวกในการปฏิบัติและได้มีการ เสริมแรงทักษะทางจิตวิทยาในวิถีทางวัฒนธรรม และด้านการพัฒนาอย่างเหมาะสมตามบริบท โดยมีการ ส่งเสริมการพัฒนาบุคคลและสังคม การป้องกันสุขภาพของตนเองและปัญหาทางสังคมและการปกป้องสิทธิ มนุษยชน ทั้งนี้เพื่อเป็นการ เตรียมบุคคลเพื่อให้ประสบความสำเร็จในชีวิตทั้งในบ้าน ชุมชน และสถานที่ที่ ทำงาน โดยใช้หลักการเรียนรู้ (learning to know) เรียนรู้ที่จะทำ (learning to do) และเรียนรู้ที่จะอยู่กับ ผู้อื่น (learning to live together and with others) (World Health Organization, 1997; The United Nations Educational, Scientific and Cultural Organization, 2005)

การปฏิรูปการศึกษานั้นได้มีการพัฒนาไปสู่รูปแบบการเปลี่ยนผ่านการศึกษาซึ่งเป็นการจัดการศึกษาที่สามารถพัฒนาบุคคลให้เป็นผู้มีศักยภาพ เป็นทรัพยากรบุคคลที่เน้นความรู้คู่ คุณธรรมและความพอเพียง รวมทั้งมีทักษะในการประกอบอาชีพ เป็นแนวคิดในการนำความรู้มาสร้างเป็นผลผลิตที่พัฒนาขึ้นโดยเกิดจากความรู้เป็นฐานหลัก (Sinlarat, 2007) การจัดการศึกษาเพื่อการประกอบอาชีพ (career education) เป็นอีกกรอบความคิดหนึ่ง ที่มีบทบาทในการพัฒนาหลักสูตรและการจัดการเรียนการสอนสำหรับผู้เรียนโดยเฉพาะในระดับการศึกษาขั้นพื้นฐาน (Office of the Basic Education Commission, 2002; Ministry of Education, 2009; Office for National Education Standards and Quality Assessment, 2009) โดยมีวัตถุประสงค์เพื่อให้ผู้เรียนได้มีความคิดที่เป็นเหตุเป็นผลในการทำงาน สังสมความรู้และทักษะที่จำเป็นในการทำงาน รู้จักการหาโอกาสและช่องทางในการทำงาน และสำรวจการก้าวเข้าสู่โลกแห่งการทำงาน ซึ่งระยะเวลาที่ได้ศึกษาหาความรู้ในสถานศึกษาขั้นพื้นฐานนั้นจำเป็นอย่างยิ่งที่จะต้องพัฒนาความรู้ทักษะและเจตคติอันดีผ่านการศึกษา การฝึกอบรม และการปฏิบัติงานเพื่อช่วยให้ผู้เรียนมีประสบการณ์ที่จะตัดสินใจในการศึกษาต่อ และการทำงานต่อไปได้ในอนาคต โดยผู้สอนจำเป็นต้องใช้ชุมชนแห่งการเรียนรู้เชิงวิชาชีพ (PLC: Professional Learning Community) เป็นแนวทางการปฏิบัติเพื่อปรับปรุงคุณภาพของนักเรียน ครู และโรงเรียน เกิดจากแนวทางการดำเนินงานเกี่ยวข้องกับการเปลี่ยนแปลงวัฒนธรรมของโรงเรียนจากเดิมที่ส่วนใหญ่ครูมักเป็นผู้รับ (passive teacher) มาเป็นครูผู้เรียนรู้แบบกระตือรือร้น (active teacher) ที่มีเป้าหมายในการพัฒนาและยกระดับคุณภาพตนเอง เพื่อนครู ผู้เรียน และโรงเรียนร่วมกัน ซึ่งจะนำไปสู่การพัฒนาปรับปรุงอย่างยั่งยืนไม่สิ้นสุด (Kitrongreuang & Sirisampan, 2017) สำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 มีวิสัยทัศน์ “บริหารจัดการเป็นเลิศ สถานศึกษาได้มาตรฐาน นักเรียนมีคุณภาพ สู่สังคมอนาคตที่ยั่งยืน” มุ่งส่งเสริมสนับสนุนการจัดการศึกษาเพื่อพัฒนาคุณภาพชีวิตที่เป็นมิตรกับสิ่งแวดล้อมตามหลักปรัชญาของเศรษฐกิจพอเพียง ส่งเสริม พัฒนาระบบการบริหารจัดการศึกษาทุกระดับ และจัดการศึกษาโดยใช้ เทคโนโลยีดิจิทัล (digital technology) เพื่อพัฒนาไปสู่ศตวรรษที่ 21 มีเป้าประสงค์ให้ผู้เรียน เป็นบุคคลแห่งการเรียนรู้ คิดริเริ่มและสร้างสรรค์สื่อ นวัตกรรม มีความรู้ มีทักษะมีสมรรถนะตามหลักสูตร และคุณลักษณะของผู้เรียนในศตวรรษที่ 21 มีสุขภาวะ ที่เหมาะสมตามวัยมีความสามารถในการพึ่งพาตนเอง ตามหลักปรัชญาของเศรษฐกิจพอเพียงและการเป็นพลเมือง พลโลกที่ดี (global citizen) พร้อมก้าวสู่มาตรฐานสากล และจำเป็นต้องมีการใช้นวัตกรรมในการขับเคลื่อนการพัฒนาทักษะชีวิตของนักเรียนตามหลักปรัชญาของเศรษฐกิจพอเพียงเพื่อให้มีประสิทธิภาพยิ่งขึ้น (The Primary Educational Service Area Office Udon Thani 2, 2020)

จากปัญหาดังกล่าวผู้วิจัยจึงมีความสนใจกำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 ให้เหมาะสมกับบริบทของแต่ละสถานศึกษาสามารถนำไปใช้ในการบริหารจัดการพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาสภาพปัจจุบัน สภาพที่พึงประสงค์ และความต้องการจำเป็นในการกำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

2. เพื่อกำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

3. เพื่อประเมินกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ระยะที่ 1 ศึกษาสภาพปัจจุบัน สภาพที่พึงประสงค์ และความต้องการจำเป็นในการกำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

1) ประชากร ได้แก่ ผู้บริหารสถานศึกษา และครูหัวหน้าฝ่ายวิชาการ โรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 ปีการศึกษา 2564 จำนวน 167 โรงเรียน จำนวนทั้งสิ้น 334 คน

2) กลุ่มตัวอย่าง ได้แก่ ผู้บริหารสถานศึกษา จำนวน 89 คน และครูหัวหน้าฝ่ายวิชาการ จำนวน 89 คน โรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 ปีการศึกษา 2564 รวมจำนวน 178 คน เปิดตารางสำเร็จรูปของเคจซีมอร์แกน (Krejcie Morgan) และใช้เทคนิคการสุ่มแบบแบ่งชั้นภูมิ (stratified random sampling) โดยใช้หน่วยอำเภอในการสุ่ม

ระยะที่ 2 การกำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

กลุ่มผู้ให้ข้อมูล ได้แก่ ผู้ทรงคุณวุฒิ จำนวน 9 คน โดยการใช้แบบเจาะจง (purposive sampling)

ระยะที่ 3 การประเมินกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

กลุ่มผู้ให้ข้อมูล ได้แก่ ผู้ทรงคุณวุฒิ จำนวน 9 คน โดยการใช้แบบเจาะจง (purposive sampling)

2. เครื่องมือที่ใช้ในการวิจัย

ระยะที่ 1 การศึกษาสภาพปัจจุบัน สภาพที่พึงประสงค์ และความต้องการจำเป็นการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

แบบสอบถามสภาพปัจจุบันและสภาพที่พึงประสงค์ มีลักษณะเป็นแบบตรวจสอบรายการ (check list) และแบบมาตราส่วนประมาณค่า (Rating Scale) 5 ระดับ ผ่านการตรวจสอบความเที่ยงตรง (validity) ได้ค่า IOC ตั้งแต่ 0.80–1.00 นำไปลองใช้ (try out) มีค่าอำนาจจำแนก (discrimination) อยู่ระหว่าง 0.450–0.898 และค่าความเชื่อมั่น (reliability) ของแบบสอบถามสภาพปัจจุบัน อยู่ในระดับ 0.97 แบบสอบถามสภาพที่พึงประสงค์ อยู่ในระดับ 0.98

ระยะที่ 2 การกำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 แบ่งออกเป็น 2 ตอน ดังนี้

ตอนที่ 1 การศึกษาสภาพแวดล้อม จุดแข็ง จุดอ่อน โอกาส ภาวะคุกคาม จำนวน 3 โรงเรียน

แบบสัมภาษณ์กึ่งโครงสร้าง นำข้อมูลจากระยะที่ 1 มาวิเคราะห์ เรียงลำดับความสำคัญของข้อมูล modified priority needs index: PNI modified และนำมาสร้างแบบสัมภาษณ์ในการศึกษาสภาพแวดล้อม จุดแข็ง จุดอ่อน โอกาส ภาวะคุกคาม

ตอนที่ 2 กำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

แบบบันทึกการสนทนากลุ่ม ผ่านการตรวจสอบความเที่ยงตรง ค่า IOC ตั้งแต่ 0.80–1.00

ระยะที่ 3 การประเมินกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

แบบประเมินกลยุทธ์ มีลักษณะเป็นแบบตรวจสอบรายการ และแบบมาตราส่วนประมาณค่า 5 ระดับ ผ่านการตรวจสอบความเที่ยงตรง ค่า IOC ตั้งแต่ 0.80–1.00 นำไปลองใช้ มีค่าอำนาจจำแนก อยู่ระหว่าง 0.550–0.878 และค่าความเชื่อมั่น ของแบบสอบถามความเหมาะสม อยู่ในระดับ 0.95 แบบสอบถามความเป็นไปได้ อยู่ในระดับ 0.97

3. การเก็บรวบรวมข้อมูล

3.1 ผู้วิจัยดำเนินการขอหนังสือมหาวิทยาลัยศรีปทุม วิทยาเขตขอนแก่น ถึงผู้อำนวยการสถานศึกษา และครูหัวหน้าฝ่ายวิชาการ โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 เพื่อขออนุญาตในการจัดเก็บข้อมูลจากกลุ่มตัวอย่าง โดยการติดตามของผู้วิจัยโดยได้รับแบบสอบถามคืน 162 ฉบับ คิดเป็นร้อยละ 91.01

3.2 เก็บรวบรวมข้อมูลการศึกษาสภาพแวดล้อม จุดแข็ง จุดอ่อน โอกาส ภาวะคุกคาม จากการสัมภาษณ์ ผู้บริหารและครูหัวหน้าฝ่ายวิชาการ จำนวน 3 โรงเรียน กลุ่มผู้ให้ข้อมูล จำนวน 6 คน โดยผู้วิจัยและผู้ช่วยผู้วิจัยดำเนินการเก็บข้อมูลจากกลุ่มผู้ให้ข้อมูล

3.3 เก็บข้อมูลจากการสนทนากลุ่ม (Focus Group Discussion) โดยผู้วิจัยดำเนินการเก็บข้อมูลด้วยตนเอง เพื่อตรวจสอบยืนยันข้อมูลด้วย ฉบับร่าง

3.4 ประเมินความเหมาะสมและความเป็นไปได้ของกลยุทธ์ฯ โดยผู้วิจัยดำเนินการเก็บข้อมูลด้วยตนเองหลังจากการสนทนากลุ่ม (Focus Group Discussion) ได้รับแบบสอบถามประเมินกลยุทธ์ฯ คืนจำนวน 9 ฉบับ คิดเป็นร้อยละ 100

4. การวิเคราะห์ข้อมูล

4.1 การวิเคราะห์ข้อมูลเชิงปริมาณ การวิเคราะห์สภาพปัจจุบันและสภาพที่พึงประสงค์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 การวิเคราะห์แบบสอบถามประเมินกลยุทธ์ฯ ในประเด็นความเหมาะสมความเป็นไปได้ โดยการหาค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (standard deviation) การวิเคราะห์ดัชนีประเมินความต้องการจำเป็นใช้สูตร (modified priority needs index: PNI modified)

4.2 การวิเคราะห์ข้อมูลเชิงคุณภาพจากแบบสัมภาษณ์ และแบบบันทึกการสนทนากลุ่ม ด้วยการวิเคราะห์เนื้อหา (content analysis)

ผลการวิจัย

1. ผลการวิเคราะห์สภาพปัจจุบัน สภาพที่พึงประสงค์ และความต้องการจำเป็นในการกำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 ดังตาราง 1

ตาราง 1 ผลการวิเคราะห์สภาพปัจจุบัน สภาพที่พึงประสงค์ และความต้องการจำเป็นการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง	สภาพปัจจุบัน			สภาพที่พึงประสงค์			PNI	ลำดับ
	$\bar{X}$	S.D.	ระดับการดำเนินงาน	$\bar{X}$	S.D.	ระดับการดำเนินงาน		
1. ด้านการบริหารสถานศึกษาตามหลักปรัชญาของเศรษฐกิจพอเพียง	3.30	0.77	ปานกลาง	4.48	0.51	มาก	0.358	1

ตาราง 1 (ต่อ)

การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง	สภาพปัจจุบัน			สภาพที่พึงประสงค์			PNI	ลำดับ
	$\bar{X}$	S.D.	ระดับการดำเนินงาน	$\bar{X}$	S.D.	ระดับการดำเนินงาน		
2. ด้านหลักสูตรการพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง	3.35	0.73	ปานกลาง	4.48	0.51	มาก	0.337	3
3. ด้านการจัดกิจกรรมพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง	3.38	0.67	ปานกลาง	4.49	0.50	มาก	0.328	4
4. ด้านการพัฒนาบุคลากรของสถานศึกษา	3.34	0.62	ปานกลาง	4.49	0.50	มาก	0.344	2
5. ด้านความสำเร็จ	3.42	0.71	ปานกลาง	4.50	0.50	มาก	0.316	5
โดยรวม	3.36	0.70	ปานกลาง	4.49	0.50	มาก	0.337	

จากตาราง 1 พบว่า สภาพปัจจุบันของการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 โดยรวมอยู่ในระดับปานกลาง สภาพที่พึงประสงค์โดยรวมอยู่ในระดับมาก ความต้องการจำเป็นเรียงลำดับจากค่าเฉลี่ยมากไปหาน้อย ดังนี้ ด้านการบริหารสถานศึกษาตามหลักปรัชญาของเศรษฐกิจพอเพียง ด้านการพัฒนาบุคลากรของสถานศึกษา ด้านหลักสูตรการพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง ด้านการจัดกิจกรรมพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง และด้านความสำเร็จ ตามลำดับ

2. จุดแข็งของการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 คือ ผู้บริหารมีประสบการณ์ในการบริหารจัดการสถานศึกษาโน้มนำหลักปรัชญาของเศรษฐกิจพอเพียงมาปฏิบัติ จุดอ่อน คือ ครูและบุคลากรทางการศึกษาขาดความรู้ความเข้าใจในการบูรณาการหลักปรัชญาของเศรษฐกิจพอเพียงในการพัฒนาทักษะชีวิตนักเรียน โอกาส คือ ได้รับการสนับสนุนการดำเนินงานจากชุมชน และภาวะคุกคาม คือ เทคโนโลยีเข้ามามีบทบาทในการเรียนการสอนทำให้นักเรียนขาดความสนใจหลักปรัชญาของเศรษฐกิจพอเพียง และนำมากำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 ประกอบด้วย 4 กลยุทธ์หลัก ดังนี้ 1) กลยุทธ์พัฒนาความสามารถทางวิชาการ 2) กลยุทธ์ประยุกต์ใช้ภูมิปัญญาท้องถิ่น 3) กลยุทธ์บูรณาการใช้แหล่งเรียนรู้ภายในและภายนอกโรงเรียน และ 4) กลยุทธ์เสริมสร้างประสบการณ์ ภายใต้กรอบของหลักปรัชญาของเศรษฐกิจพอเพียง คือ พอประมาณ มีเหตุผล มีภูมิคุ้มกัน มีเงื่อนไขความรู้และคุณธรรมเป็นตัวกำกับ เมื่อนำกลยุทธ์ฯ ไปใช้จะส่งผลให้ผู้เรียนมีทักษะชีวิตภายใต้เงื่อนไขของหลักปรัชญาของเศรษฐกิจพอเพียง คือ มีความรู้และคุณธรรม ดังภาพประกอบ 1



ภาพ 1 กลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

3. ผลการประเมินกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 ดังตาราง 2

ตาราง 2 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเหมาะสมและความเป็นไปได้ของกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

กลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ	ความเหมาะสม			ความเป็นไปได้		
	$\bar{X}$	S.D.	ความเหมาะสม	$\bar{X}$	S.D.	ความเป็นไปได้
1. กลยุทธ์พัฒนาความสามารถทางวิชาการ	4.48	0.44	มาก	4.50	0.50	มาก
2. กลยุทธ์ประยุกต์ใช้ภูมิปัญญาท้องถิ่น	4.50	0.50	มาก	4.48	0.50	มาก
3. กลยุทธ์บูรณาการใช้แหล่งเรียนรู้ภายในและภายนอกโรงเรียน	4.45	0.53	มาก	4.48	0.53	มาก
4. กลยุทธ์เสริมสร้างประสบการณ์	4.48	0.53	มาก	4.50	0.50	มาก
โดยรวม	4.48	0.50	มาก	4.49	0.51	มาก

จากตาราง 2 พบว่า กลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 มีความเหมาะสมและความเป็นไปได้โดยรวมอยู่ในระดับมาก

อภิปรายผล

1. การศึกษาสภาพปัจจุบัน สภาพที่พึงประสงค์ และความต้องการจำเป็นการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2

1.1 สภาพปัจจุบัน การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 โดยรวมอยู่ในระดับปานกลาง ทั้งนี้เป็นเพราะว่า สถานศึกษาในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 มีการการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงแต่ยังต้องการแนวทางการดำเนินการที่เป็นรูปธรรมเพื่อให้สอดคล้องกับการเปลี่ยนแปลงของสังคมในมิติต่าง ๆ ดังที่ Ministry of Education (2005); Office of the Basic Education Commission (2004) ได้นำเสนอว่า การเปลี่ยนแปลงทางสังคมสู่สังคมฐานเศรษฐกิจในยุคปัจจุบันก่อให้เกิดผลกระทบต่อ สภาพเศรษฐกิจ การเมือง ความเสื่อมของวัฒนธรรมและศีลธรรม ระบบการศึกษาที่มีในปัจจุบัน ขาดบทบาทเพียงพอในการเตรียมคนและพัฒนาศักยภาพของคนให้พร้อมรับกับการเปลี่ยนแปลง สู่สังคมเศรษฐกิจได้ประชากรขาดทักษะในการคิดและไม่สามารถเชื่อมโยงประยุกต์ใช้ความรู้เพื่อการดำเนินชีวิตได้อย่างมีประสิทธิภาพ สอดคล้องกับงานวิจัยของ Thummatasananon (2019) ได้วิจัยเรื่อง การพัฒนารูปแบบการบริหารจัดการหลักสูตรและการจัดกิจกรรมการเรียนรู้ตามหลักปรัชญาของเศรษฐกิจพอเพียง สำหรับโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 25 ผลการวิจัยพบว่า สภาพปัจจุบันการบริหารจัดการหลักสูตรและการจัดกิจกรรมการเรียนรู้ตามหลักปรัชญาของเศรษฐกิจพอเพียง สำหรับโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 25 โดยรวม พบว่า อยู่ในระดับมาก และสอดคล้องกับงานวิจัยของ Boontan (2019) ได้วิจัยเรื่อง การพัฒนาแนวทางการบริหารตามหลักปรัชญาของเศรษฐกิจพอเพียง (โรงเรียนพอเพียงท้องถิ่น) สำหรับโรงเรียนสังกัดเทศบาลกลุ่มจังหวัดการศึกษาท้องถิ่นที่ 12 ผลการวิจัยพบว่า สภาพปัจจุบันของการบริหารตามหลักปรัชญาของเศรษฐกิจพอเพียง (โรงเรียนพอเพียงท้องถิ่น) สำหรับโรงเรียนสังกัดเทศบาลกลุ่มจังหวัดการศึกษาท้องถิ่นที่ 12 โดยรวมอยู่ในระดับปานกลาง

1.2 สภาพที่พึงประสงค์ การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 อยู่ในระดับมาก ทั้งนี้เป็นเพราะว่า สถานศึกษาในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 มีความต้องการในการการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชน

การเรียนรู้ทางวิชาชีพ เป็นการเสริมสร้างทักษะชีวิตให้นักเรียนมุ่งเน้นให้มีความรู้คู่คุณธรรมและนำหลักปรัชญาของเศรษฐกิจพอเพียงไปใช้ในการดำเนินชีวิตสู่การประกอบอาชีพในอนาคต ดังที่ Sinlarat (2007) ได้นำเสนอว่า การปฏิรูปการศึกษานั้นได้มีการพัฒนาไปสู่รูปแบบการเปลี่ยนผ่านการศึกษาซึ่งเป็นการจัดการศึกษาที่สามารถพัฒนาบุคคลให้เป็นผู้มีศักยภาพ เป็นทรัพยากรบุคคลที่เน้นความรู้คู่ คุณธรรมและความพอเพียง รวมทั้งมีทักษะในการประกอบอาชีพ เป็นแนวคิดในการนำความรู้มาสร้างเป็นผลผลิตที่พัฒนาขึ้นโดยเกิดจากความรู้เป็นฐานหลัก สอดคล้องกับงานวิจัยของ Thummatasananon (2019) ได้วิจัยเรื่อง การพัฒนารูปแบบการบริหารจัดการหลักสูตรและการจัดกิจกรรมการเรียนรู้ตามหลักปรัชญาของเศรษฐกิจพอเพียง สำหรับโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 25 ผลการวิจัยพบว่า สภาพที่พึงประสงค์การบริหารจัดการหลักสูตรและการจัดกิจกรรมการเรียนรู้ตามหลักปรัชญาของเศรษฐกิจพอเพียง สำหรับโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 25 โดยรวม พบว่า อยู่ในระดับมาก และสอดคล้องกับงานวิจัยของ Boontan (2019) ได้วิจัยเรื่อง การพัฒนาแนวทางการบริหารตามหลักปรัชญาของเศรษฐกิจพอเพียง (โรงเรียนพอเพียงท้องถิ่น) สำหรับโรงเรียนสังกัดเทศบาลกลุ่มจังหวัดการศึกษาท้องถิ่นที่ 12 ผลการวิจัยพบว่า สภาพที่พึงประสงค์ของการบริหารตามหลักปรัชญาของเศรษฐกิจพอเพียง (โรงเรียนพอเพียงท้องถิ่น) สำหรับโรงเรียนสังกัดเทศบาลกลุ่มจังหวัดการศึกษาท้องถิ่นที่ 12 โดยรวมอยู่ในระดับมากที่สุด

1.3 ความต้องการจำเป็นการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 เรียงลำดับจากค่าเฉลี่ยมากไปหาน้อย ดังนี้ ด้านการบริหารสถานศึกษาตามหลักปรัชญาของเศรษฐกิจพอเพียง ด้านการพัฒนาบุคลากรของสถานศึกษา ด้านหลักสูตรการพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง ด้านการจัดกิจกรรมพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง และด้านความสำเร็จ ตามลำดับทั้งนี้เป็นเพราะว่า การเปลี่ยนแปลงทางสังคมมีผลกระทบต่อการศึกษาและการดำเนินชีวิตจำเป็นต้องมีการนำปรัชญาของเศรษฐกิจพอเพียงมาใช้ในการศึกษาและการดำเนินชีวิต ดังที่ Ministry of Education (2005); Office of the Basic Education Commission (2004) ได้นำเสนอว่า การเปลี่ยนแปลงทางสังคมสู่สังคมฐานเศรษฐกิจในยุคปัจจุบันก่อให้เกิดผลกระทบต่อ สภาพเศรษฐกิจ การเมือง ความเสื่อมของวัฒนธรรมและศีลธรรม ระบบการศึกษาที่มีในปัจจุบัน ทั้งนี้การพัฒนาประชารัฐนั้นจำเป็นต้องใช้ความรู้ความคิด ทักษะจากหลายสาขามาช่วยในการแก้ไขปัญหาและพัฒนาเพื่อให้มีทักษะชีวิตที่สมบูรณ์มากยิ่งขึ้น สอดคล้องกับงานวิจัยของ Chutrirat (2018) ได้วิจัยเรื่อง การพัฒนาโปรแกรมพัฒนาครูด้านการจัดการเรียนรู้โดยบูรณาการตามหลักปรัชญาของเศรษฐกิจพอเพียง กลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนา และวัฒนธรรม สำหรับสถานศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 25 ผลการวิจัยพบว่า ลำดับความต้องการจำเป็นในการพัฒนาโปรแกรมพัฒนาครูโดยเรียงลำดับความต้องการจำเป็นจากมาก 1) การประเมินผล การเรียนรู้ 2) การวางแผนการสอน 3) การดำเนินการสอน ตามลำดับ และสอดคล้องกับงานวิจัยของ Boontan (2019) ได้วิจัยเรื่อง การพัฒนาแนวทางการบริหารตามหลักปรัชญาของเศรษฐกิจพอเพียง (โรงเรียน

พอเพียงท้องถิ่น) สำหรับโรงเรียนสังกัดเทศบาลกลุ่มจังหวัดการศึกษาท้องถิ่นที่ 12 ผลการวิจัยพบว่า ลำดับความต้องการจำเป็นในการพัฒนาแนวทางการบริหารตามหลักปรัชญาของเศรษฐกิจพอเพียง (โรงเรียนพอเพียงท้องถิ่น) สำหรับโรงเรียนสังกัดเทศบาลกลุ่มจังหวัดการศึกษาท้องถิ่นที่ 12 พบว่า ด้านการบริหารจัดการสถานศึกษา มีความจำเป็นมากที่สุด รองลงมาคือ ด้านผลลัพธ์/ภาพความสำเร็จ ด้านการพัฒนาบุคลากรของสถานศึกษา ด้านหลักสูตรและการจัดกิจกรรมการเรียนรู้และด้านการจัดกิจกรรมพัฒนาผู้เรียนตามลำดับ

2. การกำหนดกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 ประกอบด้วย 4 กลยุทธ์หลัก ดังนี้ 1) กลยุทธ์พัฒนาความสามารถทางวิชาการ 2) กลยุทธ์ประยุกต์ใช้ภูมิปัญญาท้องถิ่น 3) กลยุทธ์บูรณาการใช้แหล่งเรียนรู้ภายในและภายนอกโรงเรียน และ 4) กลยุทธ์เสริมสร้างประสบการณ์ทั้งนี้เป็นเพราะว่าในการวิจัยครั้งนี้ผู้วิจัยได้ศึกษาข้อมูลเชิงปริมาณและข้อมูลเชิงคุณภาพเพื่อให้ได้มาซึ่งข้อมูลเชิงประจักษ์ที่ถูกต้องชัดเจนสามารถนำมากำหนดกลยุทธ์ได้อย่างถูกต้องตามหลักทฤษฎีและสามารถอธิบายเชิงเหตุผลของเรื่องที่ศึกษาตามกระบวนการกำหนดกลยุทธ์การบริหารสถานศึกษา Siribanpitak (2009); Koontz & Weihrich (1990); Brian (2002); Wheelen & Hunger (2012) ประกอบด้วย 1) การวิเคราะห์สภาพแวดล้อมภายนอก 2) การวิเคราะห์สภาพแวดล้อมภายใน และ 3) การกำหนดกลยุทธ์และวิธีปฏิบัติ สอดคล้องกับงานวิจัยของ Sanitklang (2018) ได้วิจัยเรื่องกลยุทธ์การบริหารโรงเรียนมัธยมศึกษาตามแนวคิดการพัฒนานักเรียนให้มีภาวะผู้นำที่ยอดเยี่ยม ผลการวิจัยพบว่า กลยุทธ์การบริหารโรงเรียนมัธยมศึกษาตามแนวคิดการพัฒนานักเรียนให้มีภาวะผู้นำที่ยอดเยี่ยมมีกลยุทธ์หลัก 5 กลยุทธ์ ได้แก่ 1) ปฏิรูปหลักสูตรสถานศึกษาเพื่อเสริมสร้างนักเรียนให้มีภาวะผู้นำที่ยอดเยี่ยม 2) ปรับปรุงการจัดกิจกรรมการเรียนการสอน เพื่อพัฒนานักเรียนให้มีภาวะผู้นำที่ยอดเยี่ยม 3) ปรับเปลี่ยนระบบการวัดผลและประเมินผลที่มีประสิทธิภาพ เพื่อพัฒนานักเรียนให้มีภาวะผู้นำที่ยอดเยี่ยม 4) ยกระดับการดำเนินงานสถานักเรียนยุคใหม่ที่มุ่งส่งเสริมภาวะผู้นำที่ยอดเยี่ยมในนักเรียน และ 5) พัฒนาขีดความสามารถในการจัดกิจกรรมเสริมหลักสูตรให้นักเรียนมีภาวะผู้นำที่ยอดเยี่ยม จุดประกายฝันมุ่งมั่นทำความดีสร้างชีวิตที่พอเพียง มีกลยุทธ์รอง 10 กลยุทธ์และมีวิธีดำเนินการ 30 วิธี และสอดคล้องกับงานวิจัยของ Kaniyam (2018) ได้วิจัยเรื่อง กลยุทธ์การบริหารโรงเรียนประจำตามแนวคิดความเป็นพลเมืองในศตวรรษที่ 21 ผลการวิจัยพบว่า กลยุทธ์การบริหารโรงเรียนประจำตามแนวคิดความเป็นพลเมืองในศตวรรษที่ 21 ประกอบไปด้วย 5 กลยุทธ์หลัก ได้แก่ (1) พลิกโฉมหลักสูตรโรงเรียนประจำมุ่งพัฒนาผู้เรียนด้านความเป็นพลเมืองในศตวรรษที่ 21 (2) ปรับการเรียนเปลี่ยนการสอนมุ่งพัฒนาผู้เรียนด้านความเป็นพลเมืองในศตวรรษที่ 21 (3) ปฏิรูปการวัดและประเมินผลมุ่งพัฒนาผู้เรียนด้านความเป็นพลเมืองในศตวรรษที่ 21 (4) ยกระดับการบริหารกิจการนักเรียนมุ่งพัฒนาผู้เรียนด้านความเป็นพลเมืองในศตวรรษที่ 21 (5) ปฏิรูปการบริหารหอพักมุ่งพัฒนาผู้เรียนด้านความเป็นพลเมืองในศตวรรษที่ 21

3. การประเมินกลยุทธ์การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่

การศึกษาประเมินศึกษาอุดรธานี เขต 2 มีความเหมาะสมและความเป็นไปได้โดยรวมอยู่ในระดับมาก ทั้งนี้ เป็นเพราะว่าผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องนำข้อมูลมาศึกษาสภาพแวดล้อมและกำหนดกลยุทธ์ นำเข้าสู่กระบวนการสนทนากลุ่มให้ผู้ทรงคุณวุฒิได้ตรวจสอบ เป็นการนำประเด็นท้าทายในสังคมมากำหนด กลยุทธ์เพื่อนำสู่การปฏิบัติอย่างเป็นรูปธรรม ดังที่ World Health Organization (1997); The United Nations Educational, Scientific and Cultural Organization (2005) ได้นำเสนอว่า การริเริ่มพัฒนาและ นำการจัดการศึกษาทักษะชีวิตเข้าสู่โรงเรียนนั้นได้นำไปใช้ใน หลายประเทศทั่วโลกโดยความต้องการจำเป็น ของการจัดการศึกษานั้นเป็นไปเพื่อการเผชิญความ ท้าทายในชีวิตประจำวัน เช่น ปัญหาโรคเอดส์ การตั้งครรภ์ ในวัยรุ่น ความรุนแรง เป็นต้น ในหลายประเทศได้มีการพิจารณาการพัฒนาการจัดการศึกษาทักษะชีวิต ในการปฏิรูปการศึกษาโดยสนอง ต่อสังคมสมัยใหม่และชีวิตยุคเศรษฐกิจการจัดการศึกษาเพื่อพัฒนาทักษะ ชีวิตนั้น ได้ออกแบบเพื่ออำนวยความสะดวกในการปฏิบัติและได้มีการเสริมแรงทักษะทางจิตวิทยาในวิถี ทางวัฒนธรรม และด้านการพัฒนาอย่างเหมาะสมตามบริบท สอดคล้องกับงานวิจัยของ Thovicha (2019) ได้วิจัยเรื่อง การพัฒนากลยุทธ์การบริหารสถานศึกษาสมรรถนะสูง สังกัดสำนักงานเขตพื้นที่การศึกษา มัธยมศึกษา ผลการวิจัยพบว่า มีความถูกต้องและมีความเหมาะสมอยู่ในระดับมากที่สุด และสอดคล้องกับ งานวิจัยของ Tanapatyotin (2021) ได้วิจัยเรื่อง การพัฒนากลยุทธ์การบริหารโดยใช้โรงเรียนเป็นฐาน เพื่อพัฒนาท้องถิ่น สำหรับสถานศึกษาสังกัดองค์กรปกครองส่วนท้องถิ่น ผลการวิจัยพบว่า ผลการประเมินกลยุทธ์ ด้านความเหมาะสมและความเป็นไปได้โดยรวมอยู่ในระดับมาก

ข้อเสนอแนะ

1. ข้อเสนอแนะเชิงนโยบาย

ผลจากการศึกษาสภาพปัจจุบันของการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียน ตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ สำหรับโรงเรียน ในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุดรธานี เขต 2 พบว่า ด้านการบริหารสถานศึกษาตามหลัก ปรัชญาของเศรษฐกิจพอเพียง มีค่าเฉลี่ยน้อยที่สุด และเมื่อนำมาจัดลำดับความต้องการจำเป็นปรากฏว่า มีความต้องการจำเป็นอยู่อันดับ 1 แสดงให้เห็นว่า การบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียน ตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง การบริหารสถานศึกษามีความสำคัญอย่างยิ่งต่อการพัฒนาทักษะ ชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียง เพราะกระบวนการบริหารเป็นการขับเคลื่อนงาน ทุกด้านให้เกิดประสิทธิภาพและประสิทธิผล ผู้บริหารระดับสำนักงานเขตพื้นที่การศึกษาจึงควรกำหนด นโยบายพัฒนาการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจ พอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ และมีการติดตามการดำเนินการอย่างต่อเนื่อง

2. ข้อเสนอแนะเชิงปฏิบัติ

ผู้บริหารสถานศึกษาควรนำนโยบายพัฒนาการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียน ตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ มาปฏิบัติได้อย่าง

เป็นรูปธรรมมีการมอบหมายงานตามโครงสร้างการบริหารสถานศึกษา และส่งเสริม สนับสนุน นิเทศและติดตามอย่างต่อเนื่องโดยทุกฝ่ายมีส่วนร่วมในการดำเนินการ

3. ข้อเสนอแนะในการวิจัยครั้งต่อไป

ผลจากการศึกษาสภาพปัจจุบันของการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ พบว่า ด้านการบริหารสถานศึกษาตามหลักปรัชญาของเศรษฐกิจพอเพียง มีค่าเฉลี่ยน้อยที่สุด และเมื่อนำมาจัดลำดับความต้องการจำเป็นปรากฏว่ามีความต้องการจำเป็นอยู่อันดับ 1 ควรมีการวิจัยในครั้งต่อไปลงเชิงลึกเพื่อพัฒนาการบริหารสถานศึกษาเพื่อพัฒนาทักษะชีวิตของนักเรียนตามแนวคิดปรัชญาของเศรษฐกิจพอเพียงโดยใช้กระบวนการชุมชนการเรียนรู้ทางวิชาชีพ ด้านดังกล่าวให้มีการพัฒนาในทางที่ดีขึ้น ในลักษณะของการวิจัยการพัฒนา

References

- Brian, F. (2002). *Strategic Management for School Development*. London: Chapman Publishing.
- Koontz, H., & Weihrich, H. (1990). *Essentials of Management*. McGraw-Hill.
- The United Nations Educational, Scientific and Cultural Organization. (2005). *Contributing to More Sustainable Future: Quality Education, Life Skills and Education for sustainable Development*. Division for the Promotion of Quality Education.
- Wheelen, T. L., & Hunger, J. D. (2012). *Strategic Management and Business Policy*. Pearson Prentice Hall.
- World Health Organization. (1997). *Life skills education for children and adolescents in schools*. World Health Organization.

Translate Thai Reference

- Boontan, T. (2019). *Developing Guidelines for Administration of the Philosophy of Sufficiency Economy (Local Sufficiency School) for Municipal Schools in the Provincial Group of Education 12* [Master's thesis]. Mahasarakham University. (in Thai)
- Chutrirat, P. (2018). *Development of a Program to Strengthen Teachers' Learning Management Focusing on Sufficiency Economy Integration in the Department of Social Studies, Religion and Culture for Schools under the Jurisdiction of the Secondary Educational Service Area Office 25* [Master's thesis]. Mahasarakham University. (in Thai)
- Kaniam, D. (2018). *Management Strategy for Boarding Schools Based on the 21st* [Doctoral dissertation]. Chulalongkorn University. (in Thai)

- Kitrongreuang, P., & Sirisampan, O. (2017). Proactive Learning Management Strategy To Develop Thinking and Enhance the Quality of Education for the 21st Century. (12nd Ed.). Petchkasem Printing Group Press. (in Thai)
- Ministry of Education. (2005). *Active Learning*. Ministry of Education. (in Thai)
- Ministry of Education. (2009). *Basic Education Core Curriculum 2008*. Teachers' council of Ladprao Press. (in Thai)
- Office for National Education Standards and Quality Assessment. (2009). *Second round of external quality assessment situation proposed to the Commission on Higher Education*. Education Standards and Quality Assessment. (in Thai)
- Office of the Basic Education Commission. (2002). *National Education Plan 2002-2016*. Prigwhan Graphic Press. (in Thai)
- Office of the Basic Education Commission. (2004). *Learn Integrate*. Teachers' council of Ladprao Press. (in Thai)
- Sanitklang, D. (2018). *Secondary School Management Strategies According to the Concept of Developing Transcendental Leadership in Student* [Doctoral dissertation]. Chulalongkorn University. (in Thai)
- Sinlarat, P. (2007). *The transition of education to the knowledge economy era Research Papers Included: Sila for Transition Education*. Chulalongkorn University. (in Thai)
- Siribanpitak, P. (2009). *Administration and Management of Education for a Small World*. Prigwhan Graphic Press. (in Thai)
- Tanapatyotin, K. (2021). *The Development of Strategic SBM Model Management School under Department of Local Administration* [Doctoral dissertation]. Mahasarakham University. (in Thai)
- The Primary Educational Service Area Office Udon Thani 2. (2020). *Government action plan for the fiscal year 2021*. The Primary Educational Service Area Office Udon Thani 2. (in Thai)
- Thovicha, S. (2019). *Development of administrative strategy for High Performance Schools under Secondary Educational Service Area Office*. [Doctoral dissertation]. Valaya Alongkorn Rajabhat University. (in Thai)
- Thummatasananon, S. (2019). Development of models for Curriculum Management and for Organizing Learning Activities According to the Philosophy of Sufficiency Economy for Schools under the Office of the Secondary Education Service Area 25. *Journal of Educational Measurement Mahasarakham University*, 25(2), 225-269. (in Thai)

การเปรียบเทียบคุณภาพของการตรวจให้คะแนนของแบบสอบอัตนัยวัดความสามารถในการ แก้ปัญหาทางคณิตศาสตร์ ที่ใช้การออกแบบการตรวจให้คะแนนที่ต่างกัน:

การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือผลของการวัด

Quality Comparison of Scoring of Essay Tests to Measure Mathematics

Problem Solving Ability Using Different Scoring Designs:

Application of Generalizability Theory

สุรัชย์ มาตราช^{1*} และ ประกฤติยา ทักซิโน²

Surachai Matharach^{1*} and Prakittiya Tuksino²

(Received: April 8, 2022; Revised: June 12, 2022; Accepted: June 21, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมาย 1. เพื่อศึกษาคุณภาพการตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของผู้ตรวจให้คะแนนภายใต้รูปแบบการตรวจให้คะแนนต่างกัน 2. เพื่อศึกษาขนาดของแหล่งความแปรปรวนขององค์ประกอบและค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ เมื่อจำนวนผู้ตรวจ และรูปแบบการตรวจให้คะแนนแตกต่างกัน 3. เพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ภายใต้รูปแบบการตรวจให้คะแนนต่างกัน และจำนวนผู้ตรวจต่างกัน กลุ่มตัวอย่างในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 132 คน ได้จากการสุ่มแบบ 2 ขั้นตอน และกลุ่มผู้ตรวจให้คะแนนซึ่งเป็นครูผู้สอนวิชาคณิตศาสตร์ จำนวน 3 คน ได้จากการเลือกแบบเจาะจงเครื่องมือที่ใช้ในการวิจัยคือ แบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ประกอบไปด้วยข้อสอบจำนวน 3 ข้อ มีค่าความยากตั้งแต่ 0.22 – 0.69 และค่าอำนาจจำแนกตั้งแต่ 0.56 – 0.80 ผลการวิจัยพบว่า 1) คุณภาพของการตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ผู้ตรวจให้คะแนนคนที่ 1 มีความเห็นพ้องในการตรวจแบบสอบเมื่อเทียบกับเกณฑ์การตรวจให้คะแนนข้อที่ 3 มากที่สุด ผู้ตรวจให้คะแนนคนที่ 2 มีความเห็นพ้องในการตรวจแบบสอบเมื่อเทียบกับเกณฑ์การตรวจให้คะแนนข้อที่ 1 และ 2 มากที่สุด ส่วนผู้ตรวจให้คะแนนคนที่ 3 มีความเห็นพ้องในการตรวจแบบสอบเมื่อเทียบกับเกณฑ์การตรวจ

¹ นักศึกษาปริญญาโท สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์มหาวิทยาลัยขอนแก่น

² ผู้ช่วยศาสตราจารย์ สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

¹ Master's degree student, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

² Assistant Professor, Department of Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

* Corresponding Author E-mail: Surachai_m@kkumail.com

ให้คะแนนทั้ง 3 ข้อน้อยที่สุด และความเที่ยงในการตรวจให้คะแนนระหว่างผู้ตรวจ (inter rater reliability) มีความสอดคล้องกันในระดับดีมาก 2) แหล่งความแปรปรวนในแต่ละรูปแบบการตรวจให้คะแนน รูปแบบการตรวจให้คะแนนแบบ $P \times I \times R$ แหล่งที่มีความแปรปรวนมากที่สุดคือ PI และแหล่งความแปรปรวนที่มีค่าน้อยที่สุดคือ PR และรูปแบบการตรวจให้คะแนนแบบ $P \times (I : R)$ แหล่งที่มีความแปรปรวนมากที่สุดคือ PI : R และแหล่งความแปรปรวนที่มีค่าน้อยที่สุดคือ PR 3) รูปแบบการตรวจให้คะแนนที่มีค่าสัมประสิทธิ์การสรุปอ้างอิงสูงที่สุดคือรูปแบบ $P \times I \times R$ และจำนวนผู้ตรวจ 3 คน ผลการตรวจมีค่าความเที่ยงมากกว่าผู้ตรวจ 2 คน

คำสำคัญ: ดัชนีความเห็นพ้องของผลการตรวจ สัมประสิทธิ์สหสัมพันธ์ภายในชั้น ทฤษฎีการสรุปอ้างอิง
ความน่าเชื่อถือของผลการวัด

Abstract

The objectives of this research were: 1) to study the quality of scoring the essay test measuring the mathematical problem-solving ability of the scorers under different scoring models; 2) to study the magnitude of the source of variance of the components and the reference summation coefficient of the essay test scores on mathematical problem solving when the number of scorers and scoring formats are different; 3) to compare the reference summation coefficients. (G-Coefficient) of the essay test that measures the ability to solve mathematical problems under different scoring formats and different number of scorers.

The sample consisted of 132 grade 9 students, obtained from two-stage cluster sampling, and a group of 3 scorers who were mathematics teachers, obtained through purposive sampling. The research instrument was an essay test measuring the ability to solve mathematical problems, with 3 items. The difficulty ranged from 0.22 to 0.69 and the discrimination ranged from 0.56 to 0.80. The results were as follows: 1) Regarding the quality of scoring of the essay test measuring the ability to solve mathematical problems, scorer 1 had the most consensus on the test, compared to the scoring criterion of item 3. Scorer 2 had the most consensus on the test, compared to the scoring criterion of item 1 and 2, while scorer 3 had the least consensus on the test, compared to the scoring criteria of all 3 items; and the reliability of the scoring between the scorers (Inter Rater Reliability) was consistent at a very good level.

2) Regarding the source of variance in each scoring format, in the $P \times I \times R$ scoring format, the source with the greatest variance was the PI, and the source with the lowest variance was the PR; and in the $P \times (I : R)$ scoring format, the source with the greatest variance was the PI : R, and the source with the lowest variance was then PR; 3) The scoring format with

the highest referencing coefficient of summation was the P x I x R format, and the results of 3 raters were more reliable than 2 raters.

Keywords: rater agreement index, intraclass correlation coefficient, Generalizability Theory

บทนำ

คณิตศาสตร์มีบทบาทสำคัญยิ่งต่อความสำเร็จในการเรียนรู้ในศตวรรษที่ 21 ซึ่งมีรายละเอียดตัวชี้วัดและสาระการเรียนรู้แกนกลาง กลุ่มสาระการเรียนรู้คณิตศาสตร์ (ฉบับปรับปรุง พ.ศ.2560) ระบุไว้หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 โดยหลักสูตรได้คำนึงถึงการส่งเสริมให้ผู้เรียนมีทักษะที่จำเป็นสำหรับการเรียนรู้ในศตวรรษที่ 21 เป็นสำคัญ และ Programme for International Student Assessment (PISA) ได้นิยามความฉลาดรู้ด้านคณิตศาสตร์ของ PISA 2021 ไว้ว่า ความสามารถของแต่ละบุคคลในการให้เหตุผลทางคณิตศาสตร์ และสามารถแปลงปัญหา ใช้คณิตศาสตร์ และตีความผลลัพธ์ทางคณิตศาสตร์เพื่อใช้ในการแก้ปัญหาในบริบทของโลกรับใช้จริง รวมถึงการใช้แนวคิด กระบวนการ ข้อเท็จจริง และเครื่องมือทางคณิตศาสตร์เพื่อบรรยาย อธิบาย และคาดการณ์ปรากฏการณ์ต่าง ๆ การประเมินความฉลาดรู้ด้านคณิตศาสตร์ของ PISA 2021 มีลักษณะสำคัญคือเป็นการใช้คณิตศาสตร์ในการแก้ปัญหาจากสถานการณ์และบริบทต่าง ๆ ที่เกิดขึ้นจริงในโลก โดยผนวกการให้เหตุผลเชิงคณิตศาสตร์กับการแก้ปัญหาโดยใช้หลักการ กระบวนการ ข้อเท็จจริง และเครื่องมือทางคณิตศาสตร์ไว้ด้วยกัน เพื่อนำไปใช้ในการอธิบาย การให้เหตุผล และการคาดการณ์ปรากฏการณ์ต่าง ๆ นอกจากนี้การประเมิน PISA 2021 เป็นการประเมินที่เน้นความฉลาดรู้ด้านคณิตศาสตร์ (mathematical literacy) เพื่อให้นักเรียนอายุ 15 ปี ได้เห็นความสำคัญของคณิตศาสตร์และเป็นการสร้างประสบการณ์ให้นักเรียนได้พบเจอสถานการณ์และบริบทที่เกิดขึ้นในชีวิตจริง โดยวัดกระบวนการคิดทางคณิตศาสตร์ของนักเรียนผ่านสถานการณ์และบริบทในชีวิตจริง ซึ่งแนวคิดดังกล่าวยังเป็นแกนหลักในการวัดความฉลาดรู้ด้านคณิตศาสตร์

PISA ยังคำนึงถึงการเปลี่ยนแปลงอย่างรวดเร็วของเทคโนโลยีใหม่ ๆ ที่เข้ามามีบทบาทในชีวิตมากขึ้น และการพัฒนาทักษะที่จำเป็นสำหรับศตวรรษที่ 21 ด้วยเหตุนี้จึงทำให้มีการเน้นความสามารถในการให้เหตุผลทางคณิตศาสตร์และการคิดเชิงคำนวณ (computational thinking) ซึ่งเป็นกระบวนการในการแก้ปัญหา การคิดวิเคราะห์อย่างมีเหตุผลเป็นขั้นตอน รวมถึงบริบทที่สอดคล้องกับทักษะแห่งศตวรรษที่ 21 เข้ามาเป็นส่วนหนึ่งของกรอบการประเมินนี้ โดยกรอบโครงสร้างการประเมินผลการรู้เรื่องคณิตศาสตร์ PISA 2021 เป็นการสรุปแนวทางในการประเมินและแนวทางการสร้างข้อสอบ ซึ่งจะเชื่อมโยงกับ 3 องค์ประกอบที่เกี่ยวข้องกัน ได้แก่ 1) การให้เหตุผลทางคณิตศาสตร์ 2) เนื้อหาคณิตศาสตร์ และ 3) บริบทที่ใช้ในแบบทดสอบซึ่งสัมพันธ์กับทักษะแห่งศตวรรษที่ 21 (Institute for the Promotion of Teaching Science and Technology, 2020) ดังนั้นข้อสอบจึงสร้างขึ้นบนพื้นฐานของบริบทที่เกี่ยวข้องกับปัญหาที่ตั้งขึ้น และออกแบบให้ครอบคลุมสาระตามแนวคณิตศาสตร์เป็นเรื่อง ๆ นักเรียนนำคณิตศาสตร์เข้าไปมีส่วนร่วมในการแก้ปัญหาหรือสร้างความท้าทายที่เกิดขึ้นในโลกที่เป็นบริบทของชีวิตจริง ซึ่งในแต่ละกระบวนการแก้ปัญหาต้องอาศัยการให้เหตุผล

ทางคณิตศาสตร์ทั้งสิ้น ดังนั้นในการตรวจให้คะแนนข้อสอบนี้จึงต้องสร้างเกณฑ์ไว้ให้ดี เพื่อให้ผู้ตรวจจะได้มีแนวการตรวจที่ตรงกัน โดยอาศัยเกณฑ์การตรวจให้คะแนน (Rubric Scoring) ซึ่งเกณฑ์การตรวจให้คะแนน (rubric scoring) เป็นรูปแบบของการให้คะแนนที่เกิดจากการการรวมกันระหว่างเกณฑ์การให้คะแนนกับมาตราประมาณค่าหรือระดับคะแนน เพื่อระบุถึงความแตกต่างของคำตอบจากการทำแบบทดสอบ ซึ่งสามารถช่วยลดปัญหาเกี่ยวกับเวลาที่ใช้ในการตรวจลงได้ โดยเกณฑ์การให้คะแนนมี 2 รูปแบบที่นิยมและเป็นที่รู้จัก คือ เกณฑ์การตรวจให้คะแนนแบบภาพรวมหรือแบบองค์รวม (holistic rubric) และเกณฑ์การประเมินแบบแยกส่วนหรือแบบแยกองค์ประกอบ (analytic rubric) ข้อดีสำหรับการตรวจให้คะแนนโดยใช้เกณฑ์แบบองค์รวม (holistic scoring rubric) คือใช้เวลาน้อยกว่าการใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) (Tangdhanakanond, 2014) และเป็นการตรวจให้คะแนนโดยพิจารณาภาพรวมทั้งหมดของการตอบคำถามของผู้สอบ เปิดโอกาสให้มีการพิจารณาการตอบของผู้เข้าสอบได้อย่างรวดเร็ว ใช้เวลาน้อยในการตรวจ เหมาะกับการตรวจและประเมินคำตอบของผู้สอบที่ใช้ความคิดสร้างสรรค์มากกว่าการมีคำตอบที่ถูกต้อง ชัดเจน แน่นนอน แต่ข้อด้อยคือ ไม่สามารถวิเคราะห์ จุดแข็ง-จุดอ่อนของผู้เรียนได้ ซึ่งวิธีการที่จะลดความคลาดเคลื่อนของผลการตรวจให้คะแนนจากแบบสอบอัตนัย นอกจากวิธีการฝึกอบรมและชี้แจงเกณฑ์แก่ผู้ตรวจให้คะแนน และติดตามผลการอบรมและชี้แจงเกณฑ์ก่อนการตรวจให้คะแนน พบว่ามีการนำวิธีการทางสถิติมาช่วยในการตรวจสอบแหล่งความคลาดเคลื่อน หรือความแปรปรวนจากแหล่งต่างๆ ที่เกิดจากการวัดผลด้วย โดยเฉพาะความคลาดเคลื่อนจากข้อสอบและผู้ตรวจ ซึ่งการออกแบบการวัดในทฤษฎี G-Theory เป็นการออกแบบสถานการณ์ในการวัดให้มีเงื่อนไขหรือพาเซท ตามที่ผู้ประเมินสนใจ และคำนึงถึงแหล่งความคลาดเคลื่อนที่เป็นระบบอย่างน้อย 1 แหล่ง โดยระดับความน่าเชื่อถือของผลการวัดที่ได้มีชื่อเรียกเฉพาะว่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (generalizability coefficient)

ความรู้ทฤษฎี G-Theory สามารถบอกได้ว่า ความแปรปรวนหรือความคลาดเคลื่อนของการวัดเกิดจากแหล่งใดมากน้อย โดยแหล่งของความคลาดเคลื่อนจะแตกต่างกันออกไปตามการออกแบบของผู้วิจัย ทั้งรูปแบบไขว้ (crossed) แบบแฝง (nested) หรือแบบแฝงบางส่วน (partially nested) ทำให้สามารถควบคุมได้ตรงจุด และประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ที่เทียบเท่าสัมประสิทธิ์ความเชื่อมั่นหรือสัมประสิทธิ์ความเที่ยง (reliability) ในทฤษฎี CTT โดยมุ่งเน้นไปที่องค์ประกอบความแปรปรวนที่ส่งผลต่อความคลาดเคลื่อนในการวัด โดยแยกพิจารณาเพื่อตัดสินใจได้ใน 2 ลักษณะ คือการตัดสินใจเชิงสัมบูรณ์ (absolute decision) และการตัดสินใจเชิงสัมพัทธ์ (relative decision) ขึ้นอยู่กับการวางแผนการทดสอบว่าอิงกลุ่มหรืออิงเกณฑ์ แยกการวิเคราะห์เป็น 2 ขั้นตอนคือ ขั้นที่ 1 การศึกษาเพื่อการสรุปอ้างอิง (G-study) และขั้นที่ 2 การศึกษาเพื่อการตัดสินใจ (D-study) เป็นการใช้ข้อมูลจาก G-study ออกแบบการวัดที่ทำให้ความคลาดเคลื่อนน้อยที่สุดและให้ค่าสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) สูง เหมาะสม คำนวณค่าทั้งเวลาและค่าใช้จ่าย ตลอดจนเป็นไปได้กับสถานการณ์จริง (Kanjawasee, 2012) โดยการนำองค์ความรู้เกี่ยวกับองค์ประกอบความแปรปรวน รูปแบบความสัมพันธ์ และสัมประสิทธิ์การสรุปอ้างอิง ในทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด มาออกแบบรูปแบบการตรวจให้คะแนนแบบสอบอัตนัยวิชาคณิตศาสตร์ เพื่อควบคุมแหล่งความคลาดเคลื่อน ค้นหาแบบการตรวจที่มีความเหมาะสม และมีความเป็นไปได้ในทางปฏิบัติ

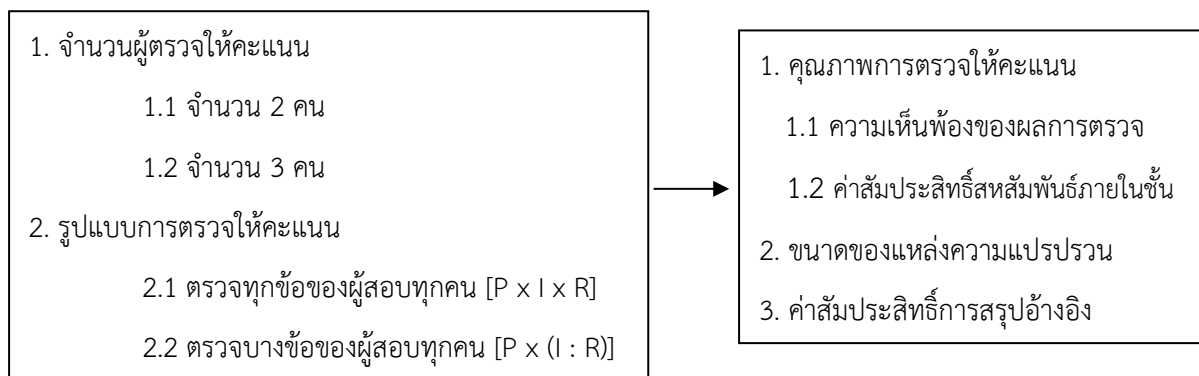
จากที่กล่าวมาข้างต้น ผู้วิจัยพบว่าแหล่งความคลาดเคลื่อนที่เกิดจากการตรวจแบบสอบการเขียนนั้นเกิดจากหลายแหล่ง โดยผู้วิจัยสนใจที่จะศึกษาขนาดของความแปรปรวนขององค์ประกอบเพื่อที่จะสามารถใช้ควบคุมแหล่งความคลาดเคลื่อนดังกล่าว โดยมีการกำหนดปัจจัยต่าง ๆ สำหรับเงื่อนไขและสถานการณ์การวัด ได้แก่ จำนวนผู้ตรวจ และรูปแบบการตรวจ เพื่อที่จะศึกษาคุณภาพการตรวจให้คะแนนและผลการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงเมื่อมีการควบคุมปัจจัยดังกล่าว ซึ่งก่อให้เกิดประโยชน์ต่อผู้บริหาร ครู ตลอดจนผู้ที่เกี่ยวข้องที่จะได้นำชุดเครื่องมือไปใช้วัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียน และนำไปพัฒนาชุดเครื่องมือวัดทักษะกระบวนการด้านอื่น ๆ ต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาคุณภาพการตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของผู้ตรวจให้คะแนนภายใต้รูปแบบการตรวจให้คะแนนต่างกัน
2. เพื่อศึกษาขนาดของแหล่งความแปรปรวนขององค์ประกอบและค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ เมื่อจำนวนผู้ตรวจ และรูปแบบการตรวจให้คะแนนแตกต่างกัน
3. เพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ภายใต้รูปแบบการตรวจให้คะแนนต่างกัน และจำนวนผู้ตรวจต่างกัน

กรอบแนวคิดการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง สามารถกำหนดกรอบแนวคิดการวิจัยได้ดังภาพ 1



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง แบ่งเป็น 2 กลุ่ม ได้แก่

1.1 ประชากรและกลุ่มตัวอย่าง ที่เป็นนักเรียนที่กำลังศึกษาระดับชั้นมัธยมศึกษาปีที่ 3 สำนักงานเขตพื้นที่การศึกษามัธยมศึกษาขอนแก่น ได้มาจากการสุ่ม 2 ขั้นตอน คือ ขั้นตอนที่ 1 ทำการสุ่มโรงเรียน จากขนาดเล็ก กลาง และใหญ่ มาขนาดละ 1 โรงเรียน และขั้นตอนที่ 2 ทำการสุ่มห้องเรียน จึงได้กลุ่มตัวอย่างจากโรงเรียน 4 ขนาด ได้นักเรียนที่เป็นกลุ่มตัวอย่างโรงเรียนละ 33 คน รวมจำนวน 132 คน

1.2 ประชากรและกลุ่มตัวอย่าง ที่เป็นผู้ตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ซึ่งเป็นครูผู้สอนวิชาคณิตศาสตร์ ที่มีคุณวุฒิในการจบทางด้านการสอนคณิตศาสตร์โดยตรง ซึ่งมีประสบการณ์การทำงานไม่น้อยกว่า 3 ปี ในโรงเรียนในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน ซึ่งผู้วิจัยทำการสุ่มครูผู้ตรวจโดยใช้วิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง (purposive sampling) จำนวน 3 คน

โดยขนาดของกลุ่มตัวอย่างที่สุ่มมาสำหรับใช้ในการศึกษามีขนาด $n_p \times n_i \times n_r$ เท่ากับ $132 \times 3 \times 3$ ได้ค่าสังเกต 1,188 ค่า ผ่านเกณฑ์การกำหนดขนาดของกลุ่มตัวอย่างอย่างน้อย 1,080 ค่าสังเกต ตามข้อเสนอของ Smith (1978, as cited in Buasiripan, 2000)

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย เป็นแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ประกอบไปด้วยข้อสอบจำนวน 3 ข้อ การสร้างข้อคำถามจะสอดคล้องกับเนื้อหาวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560) ที่มีบริบทหรือสถานการณ์ปัญหาที่สอดคล้องกับการเปลี่ยนแปลงในปัจจุบัน และสะท้อนความสามารถให้เหตุผลทางคณิตศาสตร์ของผู้สอบผ่านกระบวนการแก้ปัญหา แนวทางการสร้างข้อสอบตามกรอบการประเมินความฉลาดรู้ด้านคณิตศาสตร์ของ PISA จะเชื่อมโยงกับ 3 องค์ประกอบที่เกี่ยวข้องกัน ได้แก่ 1) การให้เหตุผลทางคณิตศาสตร์ 2) เนื้อหาคณิตศาสตร์ และ 3) บริบทที่ใช้ในแบบทดสอบ ซึ่งสัมพันธ์กับทักษะแห่งศตวรรษที่ 21 ดังนั้นข้อสอบจึงมีรูปแบบเป็นข้อคำถามที่มีคำตอบถูกเพียงคำตอบเดียว แต่อาจมีหลากหลายวิธีการได้มาซึ่งคำตอบ สร้างขึ้นบนพื้นฐานของบริบทที่เกี่ยวข้องกับปัญหาที่ตั้งขึ้น และออกแบบให้ครอบคลุมสาระตามแนวคณิตศาสตร์เป็นเรื่อง ๆ นักเรียนนำคณิตศาสตร์เข้าไปมีส่วนร่วมในการแก้ปัญหาหรือสร้างความท้าทายที่เกิดขึ้นในโลกที่เป็นบริบทของชีวิตจริง ซึ่งในแต่ละกระบวนการแก้ปัญหาต้องอาศัยการให้เหตุผลทางคณิตศาสตร์ กระบวนการหรือเทคนิคเฉพาะที่นักเรียนต้องอาศัยความรู้ มโนคติ การคิดวิเคราะห์ ประสบการณ์ และทักษะพื้นฐานต่าง ๆ ที่มีอยู่ไปประยุกต์ใช้ ตลอดจนการคิดหาแนวทางปฏิบัติเพื่อให้ปัญหาได้บรรลุตามจุดมุ่งหมายหรือคำตอบของปัญหานั้น ๆ โดยให้นักเรียนเขียนอธิบาย แสดงเหตุผลหรือแสดงวิธีการหาคำตอบที่มีวิธีการหรือขั้นตอนที่ชัดเจน ซึ่งมีค่าความตรงเชิงเนื้อหา (IOC) 0.80 – 1.00 มีค่าตั้งแต่ ค่าความยากตั้งแต่ 0.22 – 0.69 และค่าอำนาจจำแนกตั้งแต่ 0.56 – 0.80 และผลการตรวจสอบความสอดคล้องของความสอดคล้องของแบบสอบกับเกณฑ์การตรวจให้คะแนนแบบองค์รวม (holistic method) โดยผู้เชี่ยวชาญ พบว่า แนวคำตอบในแต่ละข้อคำถามของแต่ละข้อ และวิธีการให้คะแนนตามเกณฑ์การตรวจให้คะแนนแบบองค์รวม (holistic method) ที่ผู้วิจัยพัฒนาขึ้นทั้ง 3 ข้อ มีระดับความสอดคล้องกับข้อคำถามเมื่อเทียบกับเกณฑ์มีค่ามากกว่า 0.5 แสดงว่าเกณฑ์การตรวจให้คะแนนแบบองค์รวม (holistic method) สามารถนำมาใช้ตรวจให้คะแนนได้

ตาราง 1 แผนผังข้อสอบ (test blueprint) แบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์
ระดับชั้นมัธยมศึกษาปีที่ 3

สาระ	มาตรฐาน	ตัวชี้วัด	สาระการเรียนรู้แกนกลาง	สถานการณ์/ คำถาม	ค่าความ ยาก	ค่าอำนาจ จำแนก
สาระที่ 1 จำนวน และ พีชคณิต	ค 1.2 เข้าใจและ วิเคราะห์แบบรูป ความสัมพันธ์ ฟังก์ชัน ลำดับและ อนุกรม และนำไปใช้	เข้าใจและใช้ ความรู้เกี่ยวกับ ฟังก์ชันกำลังสอง ในการแก้ปัญหา คณิตศาสตร์	ฟังก์ชันกำลังสอง ■ กราฟของฟังก์ชันกำลังสอง ■ การนำความรู้เกี่ยวกับ ฟังก์ชันกำลังสองไปใช้ในการแก้ ปัญหา	1) กล่องใส่พิชชา	0.69	0.80
สาระที่ 2 การวัด และ เรขาคณิต	มาตรฐาน ค 2.1 เข้าใจพื้นฐาน เกี่ยวกับการวัด วัด และคาดคะเนขนาด ของสิ่งที่ต้องการวัด และนำไปใช้	ประยุกต์ใช้ความรู้ เรื่องปริมาตรของ พีระมิด กรวย และ ทรงกลม ในการ แก้ปัญหา คณิตศาสตร์และ ปัญหาในชีวิตจริง	ปริมาตร ■ การหาปริมาตรของพีระมิด กรวย และทรงกลม ■ การนำความรู้เกี่ยวกับ ปริมาตรของพีระมิด กรวย และ ทรงกลม ไปใช้ในการแก้ปัญหา	2) รัศมีเป็น เท่าไร	0.27	0.56
สาระที่ 3 สถิติและ ความ น่าจะเป็น	มาตรฐาน ค 3.2 เข้าใจหลักการ นับเบื้องต้น ความ น่าจะเป็น และนำไป ใช้	เข้าใจเกี่ยวกับการ ทดลองสุ่มและนำ ผลที่ได้ไปหาความ น่าจะเป็นของ เหตุการณ์	ความน่าจะเป็น ■ เหตุการณ์จากการทดลองสุ่ม ■ ความน่าจะเป็น ■ การนำความรู้เกี่ยวกับ ความน่าจะเป็นไปใช้ในชีวิตจริง	3) ลง หรือ ไม่ลง	0.22	0.64

ตัวอย่าง สถานการณ์และคำถาม : รัศมีเป็นเท่าไร

แม่ค้าขายขนมบัวลอยขนาดแป้งและส่วนผสมอื่น ๆ เข้าด้วยกัน ได้ขนาดเป็นทรงกลมตันรัศมี 6 เซนติเมตร หากแบ่งแป้งที่ผสมเรียบร้อยแล้วเป็นก้อนกลม 3 ลูกต่างขนาดกัน โดยลูกที่หนึ่งเป็นทรงกลมรัศมี 5 เซนติเมตร ลูกที่สองเป็นทรงกลมรัศมี 3 เซนติเมตร อยากทราบว่าลูกที่สามจะมีรัศมียาวกี่เซนติเมตร

ตาราง 2 ตัวอย่างเกณฑ์การตรวจให้คะแนนแบบองค์รวม

คะแนน	รูปแบบการตอบ
0	- ไม่เขียนข้อความหรือสัญลักษณ์ใด ๆ
1	- เขียนข้อความหรือสัญลักษณ์แต่ไม่เกี่ยวข้องกับประเด็นปัญหาหรือสถานการณ์ที่กำหนด
2	- มีการเขียนข้อความเพียงเล็กน้อย ไม่เพียงพอที่จะทำให้เข้าใจ และมีแนวโน้มไม่นำไปสู่คำตอบ
3	- มีการแสดงเหตุผลที่เลือกตัดสินใจ แต่ไม่มีการเขียนแสดงกระบวนการแก้ปัญหาและการให้เหตุผลทางคณิตศาสตร์ และไม่มีการสรุปคำตอบที่ชัดเจน
4	- มีการเขียนสรุปคำตอบที่ชัดเจน ถูกต้อง แต่ไม่มีสอดคล้องกับเหตุผลที่เลือกใช้ในกระบวนการแก้ปัญหาและการให้เหตุผลทางคณิตศาสตร์ที่ชัดเจน
5	- มีการเขียนสรุปคำตอบที่ชัดเจน ถูกต้อง แสดงเหตุผลที่เลือกตัดสินใจ แต่ไม่มีการเขียนแสดงกระบวนการแก้ปัญหาและการให้เหตุผลทางคณิตศาสตร์ที่ยังไม่มีความชัดเจน

ตาราง 2 (ต่อ)

คะแนน	รูปแบบการตอบ
6	- มีการเขียนคำตอบที่ชัดเจน ถูกต้อง พร้อมแสดงผลที่เลือกตัดสินใจ มีการเขียนแสดงกระบวนการแก้ปัญหา แต่การให้เหตุผลทางคณิตศาสตร์ยังไม่มีชัดเจน
7	- มีการเขียนคำตอบที่ชัดเจน ถูกต้อง พร้อมแสดงผลที่เลือกตัดสินใจ มีการเขียนแสดงกระบวนการแก้ปัญหา การให้เหตุผลทางคณิตศาสตร์มีความชัดเจน แนวทางการตอบแบบที่ 1 สูตรการหาพื้นที่รูปสี่เหลี่ยมจัตุรัส คือ ความยาวด้าน x ความยาวด้าน จะได้ ความยาวด้าน x ความยาวด้าน = 256 ความยาวด้าน x ความยาวด้าน = 16 x 16 ความยาวด้าน = 16 เซนติเมตร ดังนั้น กล่องใส่พิชชา ควรมีเส้นผ่านศูนย์กลางไม่เกิน 16 นิ้ว แนวทางการตอบแบบที่ 2 พื้นที่กล่องเป็นรูปสี่เหลี่ยมจัตุรัส = (x)(x) $256 = x^2$ $\sqrt{256} = x$ $x = 16$ ดังนั้น กล่องใส่พิชชา ควรมีเส้นผ่านศูนย์กลางไม่เกิน 16 นิ้ว

3. การเก็บรวบรวมข้อมูล

1. ผู้วิจัยนำเป็นแบบสอบถามวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ที่สร้างขึ้น และขอความอนุเคราะห์โรงเรียนที่เป็นกลุ่มตัวอย่าง ดำเนินการสอบตามวัน และเวลาที่กำหนด

2. การตรวจให้คะแนน ผู้วิจัยได้สร้างเกณฑ์การตรวจให้คะแนนแบบวิธีแบบองค์รวม (holistic rubric) โดยผู้ตรวจให้คะแนนจะแบ่งคะแนนของแต่ละข้อออกเป็นส่วนๆ มีรูปแบบผู้ตรวจให้คะแนน จำนวน 2 คน และ 3 คน โดยก่อนที่ผู้ตรวจจะทำการตรวจแบบสอบ ผู้วิจัยได้อธิบายการตรวจและเกณฑ์การตรวจให้คะแนน รวมถึงอธิบายตัวอย่างหรือแนวคำตอบในการตรวจให้คะแนน เพื่อลดความคลาดเคลื่อนอันเนื่องมาจากการให้คะแนน พร้อมทั้งให้คำแนะนำในการตรวจให้คะแนนและเมื่อผู้ตรวจตรวจให้คะแนนจะต้องบันทึกคะแนนลงในแบบบันทึกคะแนนที่ผู้วิจัยจัดทำขึ้น

4. การวิเคราะห์ข้อมูล

1. เพื่อวิเคราะห์คุณภาพการตรวจให้คะแนนแบบสอบวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของผู้ตรวจให้คะแนนภายใต้รูปแบบการตรวจให้คะแนนต่างกัน ได้แก่ 1.1 การวิเคราะห์ร้อยละความเห็นพ้อง (rater agreement) ของผลการตรวจระหว่างผู้ตรวจเทียบกับคะแนนเกณฑ์มาตรฐาน หาได้จากการสร้างตารางแจกแจงความถี่แบบสองทางโดยโปรแกรม SPSS ผ่านเมนู analyze -> descriptive statistics -> crosstabs ทำการตรวจสอบความถี่ของผลการตรวจให้คะแนนที่ตรงกันและทำการคำนวณในรูปแบบของร้อยละ แปลความหมายของค่าสถิติร้อยละ และ 1.2 การวิเคราะห์สถิติสหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) โดยแปลผลตามเกณฑ์ของ Koo & Li (2016)

2. เพื่อวิเคราะห์ขนาดของแหล่งความแปรปรวนขององค์ประกอบและค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ เมื่อจำนวนผู้ตรวจ และรูปแบบการตรวจให้คะแนนแตกต่างกัน โดยการวิเคราะห์หาขนาดของแหล่งความแปรปรวน ได้แก่ ผู้สอบ ข้อสอบ และผู้ตรวจ โดยใช้โปรแกรม EduG

3. เพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ภายใต้รูปแบบการตรวจให้คะแนนต่างกัน และจำนวนผู้ตรวจต่างกัน โดยทำการวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ ภายใต้เงื่อนไขในการตรวจให้คะแนนระหว่างผู้ตรวจ 2 คน และ 3 คน ที่มีรูปแบบการตรวจ 2 รูปแบบ ได้แก่ 1. ผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน [$P \times I \times R$ Design] 2. ผู้ตรวจตรวจข้อสอบบางข้อของผู้สอบทุกคน [$P \times (I : R)$ Design] โดยใช้โปรแกรม EduG

ผลการวิจัย

1. คุณภาพของการตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์

1.1 ผลการวิเคราะห์ความเห็นพ้อง (rater agreement) ของผลการตรวจระหว่างผู้ตรวจ เทียบกับคะแนนเกณฑ์มาตรฐานที่ได้จากฉันทามติในการตรวจให้คะแนนจากผู้เชี่ยวชาญ แสดงได้ดังตาราง 3

ตาราง 3 ร้อยละความเห็นพ้องของผลการตรวจระหว่างผู้ตรวจเทียบกับคะแนนเกณฑ์มาตรฐาน

ข้อสอบ	ผู้ตรวจ	ร้อยละความเห็นพ้อง
ข้อ 1	คนที่ 1	54.55
	คนที่ 2	55.30
	คนที่ 3	26.52
ข้อ 2	คนที่ 1	44.70
	คนที่ 2	55.30
	คนที่ 3	32.58
ข้อ 3	คนที่ 1	46.21
	คนที่ 2	45.45
	คนที่ 3	34.09

จากตาราง 3 พบว่า ร้อยละความเห็นพ้องของผลการตรวจผู้ตรวจให้คะแนนคนที่ 1 มีความเห็นพ้องกับเกณฑ์การตรวจข้อที่ 1 ข้อที่ 2 และ ข้อที่ 3 เท่ากับ ร้อยละ 54.55, 44.70, 46.21 ตามลำดับ ร้อยละความเห็นพ้อง ของผลการตรวจผู้ตรวจให้คะแนนคนที่ 2 มีความเห็นพ้องกับเกณฑ์การตรวจข้อที่ 1 ข้อที่ 2 และ ข้อที่ 3 เท่ากับ ร้อยละ 55.30, 55.30, 45.45 ตามลำดับ และร้อยละความเห็นพ้องของผลการตรวจผู้ตรวจให้คะแนนคนที่ 3 มีความเห็นพ้องกับเกณฑ์การตรวจข้อที่ 1 ข้อที่ 2 และ ข้อที่ 3 เท่ากับ ร้อยละ 26.52, 32.58, 34.09 ตามลำดับ ซึ่งผู้ตรวจให้คะแนนคนที่ 1 มีความเห็นพ้องในการตรวจแบบสอบเมื่อเทียบกับเกณฑ์การตรวจให้คะแนนข้อที่ 3 มากที่สุด ผู้ตรวจให้คะแนนคนที่ 2 มีความเห็นพ้องในการตรวจแบบสอบเมื่อเทียบ

กับเกณฑ์การตรวจให้คะแนนข้อที่ 1 และ 2 มากที่สุด ส่วนผู้ตรวจให้คะแนนคนที่ 3 มีความเห็นพ้องในการตรวจแบบสอบเมื่อเทียบกับเกณฑ์การตรวจให้คะแนนทั้ง 3 ข้อ น้อยที่สุด

1.2 ผลการวิเคราะห์การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) แสดงได้ดังตาราง 4

ตาราง 4 ค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC)

ข้อสอบ	วิธีคำนวณ	Intraclass Correlation Coefficient: ICC (95% CI)	ระดับ
ข้อ 1	รายบุคคล	0.86 (0.82 - 0.89)	ดี
	โดยเฉลี่ย	0.96 (0.95 - 0.97)	ดีมาก
ข้อ 2	รายบุคคล	0.78 (0.73 - 0.83)	ดี
	โดยเฉลี่ย	0.94 (0.92 - 0.95)	ดีมาก
ข้อ 3	รายบุคคล	0.95 (0.93 - 0.96)	ดีมาก
	โดยเฉลี่ย	0.99 (0.98 - 0.99)	ดีมาก

หมายเหตุ : 95% CI = 95% Confidence Interval

จากตาราง 4 พบว่า ค่าความเที่ยงในการตรวจให้คะแนนค่าความสอดคล้องระหว่างผู้ตรวจให้คะแนนจากค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) ที่มีผู้ตรวจให้คะแนน 3 คน ที่ระดับความเชื่อมั่น 95% จำแนกเป็นรายข้อซึ่งค่า ICC จะอยู่ระหว่าง 0 ถึง 1 โดยค่า ICC มากกว่าหรือเท่ากับ 0.90 แต่ไม่เกิน 1.00 ถือว่ามีความน่าเชื่อถืออยู่ในระดับดีมาก ค่า ICC มากกว่าหรือเท่ากับ 0.75 ถือว่ามีความน่าเชื่อถืออยู่ในระดับดี ค่า ICC มากกว่าหรือเท่ากับ 0.50 แต่น้อยกว่า 0.75 ถือว่ามีความน่าเชื่อถืออยู่ในระดับพอใช้ และค่า ICC น้อยกว่า 0.50 ถือว่ามีความน่าเชื่อถืออยู่ในระดับต่ำ ตามเกณฑ์ของ Koo & Li (2016) ผลการวิเคราะห์พบว่า ความสอดคล้องระหว่างผู้ตรวจให้คะแนนแบ่งตามข้อคำถาม โดยข้อที่ 1 โดยเฉลี่ยมีค่า ICC = 0.96 (0.95 - 0.97) ข้อที่ 2 โดยเฉลี่ยมีค่า ICC = 0.94 (0.92 - 0.95) และ ข้อที่ 3 โดยเฉลี่ยมีค่า ICC = 0.99 (0.98 - 0.99) ซึ่งทั้ง 3 ข้อคำถามถือว่ามีความน่าเชื่อถืออยู่ในระดับดีมาก

2. ผลการวิเคราะห์หาขนาดของแหล่งความแปรปรวน ได้แก่ ผู้สอบ ข้อสอบ และผู้ตรวจ ดังตาราง 5

ตาราง 5 ความแปรปรวนและร้อยละของแหล่งความแปรปรวนในแต่ละรูปแบบการตรวจให้คะแนน

	P x I x R			P x (I : R)		
	DF	Var.	%	DF	Var.	%
P	131	0.313	4.1	131	1.538	18.9
I	2	2.139	28.1			
R	2	0.699	9.2	2	-0.014	28.3
PI	262	3.674	48.2			
PR	262	-0.008	0.0	262	-1.233	0.0
IR	4	0.166	2.2			
PIR	524	0.627	8.2			
I : R	-			6	2.305	28.3
PI : R	-			786	4.301	52.8
ผลรวม	1187			1187		

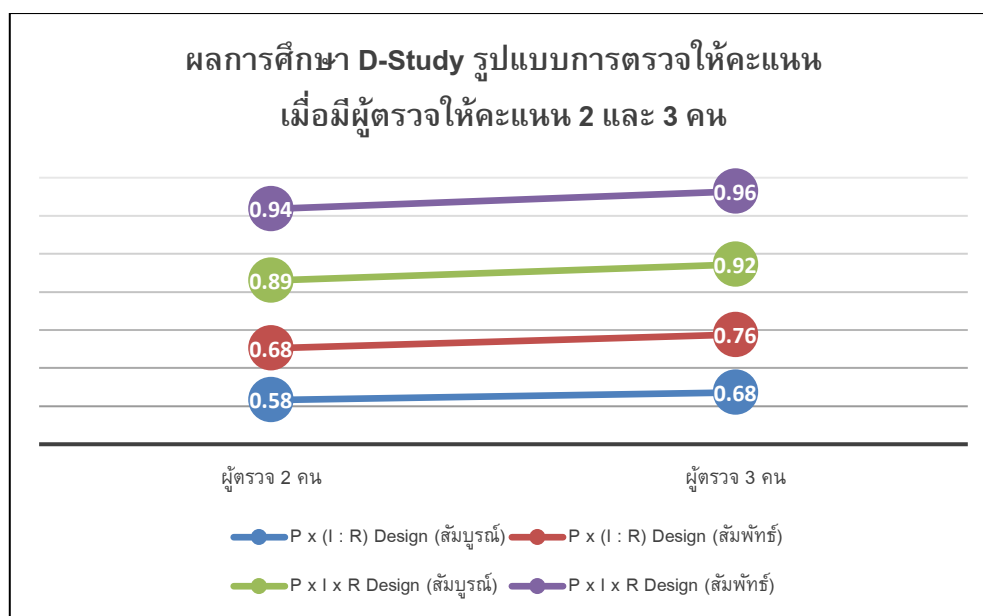
จากตาราง 5 การวิเคราะห์หาขนาดของแหล่งความแปรปรวนในแต่ละรูปแบบการตรวจให้คะแนน ผลการวิเคราะห์พบว่า รูปแบบการตรวจให้คะแนนแบบ $P \times I \times R$ แหล่งที่มีความแปรปรวนมากที่สุดคือ PI มีค่าเท่ากับ 3.674 และแหล่งความแปรปรวนที่มีค่าน้อยที่สุดคือ PR มีค่าเท่ากับ -0.008 และรูปแบบการตรวจให้คะแนนแบบ $P \times (I : R)$ แหล่งที่มีความแปรปรวนมากที่สุดคือ PI : R มีค่า 4.301 และแหล่งความแปรปรวนที่มีค่าน้อยที่สุดคือ PR มีค่าเท่ากับ -1.233

3. ผลการวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ ระหว่างผู้ตรวจ 2 คนและ 3 คน ที่มีรูปแบบการตรวจ 2 รูปแบบ ได้แก่ 1. ผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน [$P \times I \times R$ Design] 2. ผู้ตรวจตรวจข้อสอบบางข้อของผู้สอบทุกคน [$P \times (I : R)$ Design] มีผลดังตาราง 6

ตาราง 6 ผลการศึกษา D-Study รูปแบบการตรวจให้คะแนนเมื่อมีผู้ตรวจให้คะแนน 2 และ 3 คน

รูปแบบการตรวจ (Design)		จำนวนผู้ตรวจ (n_r)	
		2	3
$P \times I \times R$ Design	ρ_δ^2 (สัมพัทธ์)	0.94	0.96
	ρ_Δ^2 (สัมบูรณ์)	0.89	0.92
$P \times (I : R)$ Design	ρ_δ^2 (สัมพัทธ์)	0.68	0.76
	ρ_Δ^2 (สัมบูรณ์)	0.58	0.68

เพื่อแสดงให้เห็นการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนสำหรับนำไปใช้ตัดสินใจเชิงสัมบูรณ์ (G – coefficient for absolute decisions) และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (relative coefficient) จึงได้ของคะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ตามรายละเอียดที่กล่าวมาแล้วข้างต้น ผู้วิจัยนำเสนอการเปรียบเทียบผลการศึกษา D-Study รูปแบบการตรวจให้คะแนนเป็นแผนภูมิแท่ง ดังแสดงในภาพ 2



ภาพ 2 ผลการศึกษา D-Study รูปแบบการตรวจให้คะแนน เมื่อมีผู้ตรวจให้คะแนน 2 และ 3 คน

จากตาราง 6 และภาพ 2 ผลการศึกษา D-study รูปแบบการตรวจให้คะแนนแบบ $P \times I \times R$ พบว่า เมื่อมีผู้ตรวจให้คะแนน 2 และ 3 คน มีสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (relative coefficient) คือ 0.94 , 0.96 ตามลำดับ และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (absolute coefficient) คือ 0.89, 0.92 ตามลำดับ รูปแบบการตรวจให้คะแนนแบบ $P \times (I : R)$ พบว่า เมื่อมีผู้ตรวจให้คะแนน 2 และ 3 คน มีสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (relative coefficient) คือ 0.68, 0.76 ตามลำดับ และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (absolute coefficient) คือ 0.58, 0.68 ตามลำดับ

อภิปรายผล

จากผลการวิจัย ผู้วิจัยนำเสนอประเด็นการอภิปราย ดังนี้

1. คุณภาพของการตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์

1.1 ผู้ตรวจให้คะแนนคนที่ 1 มีความเห็นพ้องในการตรวจแบบสอบเมื่อเทียบกับคะแนนเกณฑ์มาตรฐานข้อที่ 3 มากที่สุด ผู้ตรวจให้คะแนนคนที่ 2 มีความเห็นพ้องในการตรวจแบบสอบเมื่อเทียบกับคะแนนเกณฑ์มาตรฐานข้อที่ 1 และ 2 มากที่สุด ส่วนผู้ตรวจให้คะแนนคนที่ 3 มีความเห็นพ้องในการตรวจแบบสอบเมื่อเทียบกับคะแนนเกณฑ์มาตรฐานทั้ง 3 ข้อ น้อยที่สุด เนื่องจากผู้ตรวจมีความเห็นไปในทิศทางที่ต่างกันในบางสถานการณ์คำถาม การชี้แจงเกณฑ์การตรวจให้คะแนนอาจยังไม่เพียงพอที่จะให้ผู้ตรวจให้คะแนนได้ใกล้เคียงหรือสอดคล้องกับคะแนนเกณฑ์มาตรฐาน สอดคล้องกับ Aphaikawee (2019) ที่กล่าวว่า สาเหตุของการแบบแผนการให้คะแนนที่แตกต่างกันนี้อาจมาจากคุณลักษณะของผู้ตรวจ เช่น ประสบการณ์การทำงาน การจบการศึกษาตรงตามสาขาวิชา การได้รับการฝึกอบรม เป็นต้น หากมีการควบคุมปัจจัยดังกล่าวข้างต้นจะทำให้ผู้ตรวจมีแบบแผนในการตรวจให้คะแนนและมีความสอดคล้องกับเกณฑ์มากขึ้น

1.2 ผลการวิเคราะห์ความสอดคล้องของผลการตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ด้วยค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) พบว่า ความเที่ยงในการตรวจให้คะแนนระหว่างผู้ตรวจ (inter rater reliability) มีความสอดคล้องกันในระดับดีมาก แสดงให้เห็นว่าผู้ตรวจให้คะแนนสอดคล้องไปในทิศทางเดียวกัน เนื่องมาจากการอธิบายการตรวจและเกณฑ์การตรวจให้คะแนน รวมถึงอธิบายตัวอย่างหรือแนวคำตอบในการตรวจให้คะแนน เพื่อลดความคลาดเคลื่อนอันเนื่องมาจากการให้คะแนน พร้อมทั้งให้คำแนะนำในการตรวจให้คะแนน ซึ่งสอดคล้องกับ Prasertsinponma & Tuksino (2021) ที่พบว่า ค่าความเที่ยงในการตรวจให้คะแนนค่าความสอดคล้องระหว่างผู้ตรวจให้คะแนนจากค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) ที่มีผู้ตรวจให้คะแนน 3 คน ที่ระดับความเชื่อมั่น 95% จำแนกเป็นรายข้อ โดยข้อที่ 1 โดยเฉลี่ยมีค่า $ICC = 0.96$ (0.92-0.98) ข้อที่ 2 โดยเฉลี่ยมีค่า $ICC = 0.97$ (0.96-0.98) และ ข้อที่ 3 โดยเฉลี่ยมีค่า $ICC = 0.97$ (0.95-0.98) ซึ่งทั้ง 3 ข้อคำถามถือว่ามีความน่าเชื่อถืออยู่ในระดับดีมาก

2. ผลการวิเคราะห์แหล่งความแปรปรวนของแต่ละองค์ประกอบที่มีผลต่อการสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) ผลการประมาณค่าความแปรปรวนของทั้ง 3 องค์ประกอบ คือ ผู้สอบ (P) ข้อสอบ (I)

และผู้ตรวจ (R) เมื่อใช้รูปแบบการตรวจข้อสอบทุกข้อของผู้สอบทุกคน [$P \times I \times R$] พบว่า แหล่งความแปรปรวนที่มากที่สุด คือ แหล่งปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบ (PI) แสดงว่าความสามารถของนักเรียนมีความแตกต่างกันและมีความถนัดในข้อสอบแต่ละข้อต่างกัน ซึ่งสอดคล้องกับ Bhu-iam (2007) ซึ่งกล่าวว่า ปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบมีผลต่อนักเรียนเนื่องจากนักเรียนแต่ละคนมีความสามารถต่างกันทำให้เกิดความแปรปรวน และแหล่งที่ความแปรปรวนน้อยที่สุดคือ แหล่งปฏิสัมพันธ์ระหว่างผู้สอบและผู้ตรวจ (PR) รูปแบบการตรวจให้คะแนนแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$P \times (I : R)$] พบว่าแหล่งแปรปรวนมากที่สุดคือ $PI : R$ แสดงว่า มีความแปรปรวนคงเหลือที่เกิดจากปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบจากการตรวจของผู้สอบบางคนและยังไม่สามารถอธิบายได้ และแหล่งความแปรปรวนที่น้อยที่สุดคือปฏิสัมพันธ์ระหว่างผู้สอบกับผู้ตรวจ (PR) แสดงว่าผู้ตรวจสามารถตรวจให้คะแนนได้เหมาะสมกับกลุ่มตัวอย่างที่เป็นผู้สอบ และรูปแบบการตรวจให้คะแนนแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$P \times (I : R)$] มีความแปรปรวนที่มาจากผู้ตรวจมากกว่าความแปรปรวนจากผู้สอบ จะเห็นได้ว่า ซึ่งสอดคล้องกับ ผลการวิจัยของ Inthanate (2011); Chuithong (2011); Bhu-iam (2007) พบว่า ผู้ตรวจให้คะแนนเป็นแหล่งความคลาดเคลื่อนที่สำคัญที่ส่งผลต่อความเที่ยงของคะแนน และมีอิทธิพลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง

3. ผลการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ที่มีการรูปแบบการตรวจให้คะแนน 2 ลักษณะ คือ ผู้ตรวจตรวจทุกข้อของผู้สอบทุกคน [$P \times I \times R$] และผู้ตรวจตรวจบางข้อของผู้สอบทุกคน [$P \times (I : R)$] สรุปผลได้ ดังนี้

3.1 รูปแบบการตรวจแบบผู้ตรวจตรวจทุกข้อของผู้สอบทุกคน [$P \times I \times R$] มีค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมพัทธ์ (relative coefficient) เท่ากับ 0.94 และ 0.96 เมื่อผู้ตรวจตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ 3 ข้อ จำนวน 2 และ 3 คน ตามลำดับ และมีค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) เท่ากับ 0.89 และ 0.92 เมื่อผู้ตรวจตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ 3 ข้อ จำนวน 2 และ 3 คน ตามลำดับ

3.2 รูปแบบการตรวจแบบผู้ตรวจตรวจบางข้อของผู้สอบทุกคน [$P \times (I : R)$] มีค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมพัทธ์ (relative coefficient) เท่ากับ 0.68 และ 0.76 เมื่อผู้ตรวจตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ 3 ข้อ จำนวน 2 และ 3 คน ตามลำดับ และมีค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) เท่ากับ 0.58 และ 0.68 เมื่อผู้ตรวจตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ 3 ข้อ จำนวน 2 และ 3 คน ตามลำดับ

จากการศึกษาเพื่อการตัดสินใจสรุปอ้างอิง (D-study) พบว่า 1) เจอนไขผู้ตรวจให้คะแนน 3 คน ส่งผลต่อสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (relative coefficient) และ สัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (absolute coefficient) สูงกว่าผู้ตรวจให้คะแนน 2 คน เนื่องจากความเห็นพ้องกับคะแนนเกณฑ์มาตรฐานที่ต่างกัน และผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (ICC) พบว่า ความสอดคล้องของผู้ตรวจให้คะแนนรายบุคคล ผู้ตรวจบางคนอยู่ในระดับดี แต่ความสอดคล้องของผู้ตรวจ

ให้คะแนนโดยเฉลี่ยอยู่ในระดับดีมาก ผลการตรวจให้คะแนนจากผู้ตรวจ 2 คน จึงมีระดับความน่าเชื่อถือของผู้ตรวจให้คะแนนน้อยกว่าผลการตรวจให้คะแนนจากผู้ตรวจ 3 คน สอดคล้องกับ Taoto et al. (2016) กล่าวว่าในแต่ละรูปแบบการตรวจให้คะแนนที่เหมือนกัน เมื่อใช้จำนวนผู้ตรวจให้คะแนนมากขึ้น ค่าความเชื่อมั่นจะมีแนวโน้มเพิ่มมากขึ้น สอดคล้องกับผลการวิจัยของ Aphaikawi (2019) พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับตัดสินใจเชิงสมบูรณ์ของผลการให้คะแนน มีค่าเพิ่มขึ้นเมื่อผู้ตรวจตั้งแต่ 3 คนหรือจำนวนข้อสอบเพิ่มขึ้น 3 ข้อขึ้นไป จะทำให้ผลการตรวจยังมีความเที่ยงเพิ่มมากขึ้น และสอดคล้องกับ Thitikanpodchana & Tuksino (2021). ที่กล่าวว่าเมื่อจำนวนผู้ตรวจเพิ่มขึ้น ส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (relative coefficient) และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสมบูรณ์ (absolute coefficient) และ 2) ผลการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ที่มีรูปแบบการตรวจที่ต่างกัน ภายใต้จำนวนข้อสอบและจำนวนผู้ตรวจต่างกันตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของการวัด จะได้ว่าเมื่อตรวจให้คะแนนด้วยรูปแบบการตรวจข้อสอบทุกข้อของผู้สอบทุกคน [$P \times I \times R$] มีค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสัมพัทธ์ (relative coefficient) และค่าสัมประสิทธิ์การสรุปอ้างอิงการตัดสินใจเชิงสมบูรณ์ (absolute coefficient) สูงกว่ารูปแบบการตรวจให้คะแนนแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$P \times (I : R)$] เนื่องจากรูปแบบการตรวจข้อสอบทุกข้อของผู้สอบทุกคน [$P \times I \times R$] รูปแบบนี้ผู้สอบทุกคนได้รับเงื่อนไขในการตรวจให้คะแนนจากผู้ตรวจคนเดียวกันเหมือนกันทุกคนอย่างเท่าเทียมกัน ส่วนรูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$P \times (I : R)$] รูปแบบนี้ผู้สอบแต่ละคนได้รับเงื่อนไขในการตรวจให้คะแนนจากผู้ตรวจต่างกัน โดยไม่ได้รับการตรวจจากผู้ตรวจทุกคน และอาจจะไม่ได้มีความถนัดในการตอบข้อสอบที่ได้รับการตรวจให้คะแนน รวมถึงขนาดของกลุ่มตัวอย่างที่เป็นผู้ตรวจให้คะแนนถือว่าค่อนข้างน้อย ทำให้มีความเห็นในการตรวจให้คะแนนต่างกัน สอดคล้องกับผลการวิเคราะห์ความเห็นพ้อง (rater agreement) ของผลการตรวจระหว่างผู้ตรวจ เทียบกับคะแนนเกณฑ์มาตรฐานที่ได้จากฉันทามติในการตรวจให้คะแนนจากผู้เชี่ยวชาญที่แสดงให้เห็นว่าผู้ตรวจให้คะแนนมีความเห็นพ้องที่ต่างกันไปใน การตรวจข้อสอบแต่ละข้อ และผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (ICC) พบว่า ความสอดคล้องระหว่างผู้ตรวจให้คะแนนแบ่งตามข้อคำถามทั้ง 3 ข้อคำถามถือว่ามีความน่าเชื่อถืออยู่ในระดับดีมาก ดังนั้น การตรวจให้คะแนนการทำแบบสอบจากกลุ่มตัวอย่างที่เป็นผู้สอบ ควรเลือกใช้รูปแบบการตรวจให้คะแนนแบบตรวจข้อสอบทุกข้อของผู้สอบทุกคน [$P \times I \times R$] จะให้ค่าว่าค่าสัมประสิทธิ์การสรุปอ้างอิงสูงกว่ารูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$P \times (I : R)$] เพราะเป็นสถานการณ์การตรวจให้คะแนนในระดับชั้นเรียน ซึ่งเป็นสถานการณ์ที่มีจำนวนครูผู้สอนหรือผู้ตรวจที่จำกัดและมีจำนวนนักเรียนไม่มากนัก จึงควรเลือกใช้รูปแบบการตรวจข้อสอบทุกข้อของผู้สอบทุกคน [$P \times I \times R$] ถึงแม้จะเป็นรูปแบบที่ค่อนข้างใช้เวลาในการตรวจนานเมื่อเปรียบเทียบกับรูปแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน [$P \times (I : R)$] ซึ่งสามารถลดเวลาและภาระในการตรวจให้คะแนนได้มากกว่า แต่ก็เป็นรูปแบบที่สอดคล้องกับสถานการณ์จริงและไม่จำเป็นต้องใช้ผู้ตรวจมากนักเพื่อลดต้นทุนการพัฒนาผู้ตรวจ

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1) ผลการวิเคราะห์ความเห็นพ้อง (rater agreement) ของผลการตรวจระหว่างผู้ตรวจ เทียบกับคะแนนเกณฑ์มาตรฐาน พบว่า ผู้ตรวจให้คะแนนคนที่ 3 มีผลการตรวจให้คะแนนที่มีความเห็นพ้องกับคะแนนเกณฑ์มาตรฐานน้อยกว่าผู้ตรวจคนอื่นในทุกข้อ ดังนั้น ในการคัดเลือกผู้ตรวจให้คะแนน ควรมีการเลือกผู้ตรวจที่มีคุณลักษณะเหมือนกับผู้ตรวจคนที่ 1 และ 2 ที่มีการยอมรับเกณฑ์การตรวจให้คะแนนมากกว่า ซึ่งจะช่วยให้ผลการวิเคราะห์ร้อยละความเห็นพ้องสูงกว่า ส่งผลให้ผลการตรวจให้คะแนนมีความสอดคล้องกับคะแนนเกณฑ์มาตรฐาน มีความน่าเชื่อถือมากขึ้น และลดความคลาดเคลื่อนในการตรวจให้คะแนนลงได้ ดังนั้นเพื่อลดความคลาดเคลื่อนอันเนื่องมาจากการให้คะแนน ก่อนการตรวจให้คะแนนควรมีการอบรมหรือชี้แจงเกณฑ์การตรวจให้คะแนนพร้อมทั้งให้คำแนะนำในการตรวจให้คะแนนแก่ผู้ตรวจให้คะแนน รวมถึงอธิบายตัวอย่างหรือแนวคำตอบในการตรวจให้คะแนน ให้ผู้ตรวจให้คะแนนยอมรับ และมีแนวทางการตรวจให้คะแนนไปในแนวเดียวกัน

2) ผลการวิเคราะห์แหล่งความแปรปรวนของแต่ละองค์ประกอบที่มีผลต่อการสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) ผลการประมาณค่าความแปรปรวนของทั้ง 3 องค์ประกอบ คือ ผู้สอบ (P) ข้อสอบ (I) และผู้ตรวจ (R) เมื่อใช้รูปแบบการตรวจข้อสอบทั้งสองรูปแบบ พบว่า มีความแปรปรวนที่มาจากผู้ตรวจมากกว่าความแปรปรวนจากผู้สอบและข้อสอบ ซึ่งอาจเกิดจากเกณฑ์การตรวจให้คะแนนที่เข้าใจยาก ดังนั้นควรใช้เวลาในการการอบรมหรือชี้แจงเกณฑ์การตรวจให้คะแนนมากขึ้น

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

1) ผลการวิจัยในครั้งนี้ ศึกษาผลการตรวจให้คะแนนโดยใช้เกณฑ์การตรวจให้คะแนนแบบองค์รวม (Holistic Rubric) พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงในรูปแบบการตรวจด้วยรูปแบบการตรวจข้อสอบทุกข้อของผู้สอบทุกคน $[P \times I \times R]$ มีค่าสูงกว่า การตรวจข้อสอบบางข้อของผู้สอบทุกคน $[P \times (I : R)]$ จึงควรเปรียบเทียบกับผลการตรวจให้คะแนนแบบวิธีวิเคราะห์ย่อย (Analytic Rubric) ซึ่งเป็นเกณฑ์ที่มีความละเอียด มีรายละเอียดของแต่ละสเกลชัดเจน เพื่อใช้ในการตัดสินใจเลือกใช้เกณฑ์และรูปแบบการตรวจให้คะแนนให้เหมาะสมในการตรวจให้คะแนนครั้งต่อไป

2) การวิจัยครั้งนี้ผู้วิจัยได้เลือกกลุ่มตัวอย่างที่เป็นผู้ตรวจให้คะแนนแบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ซึ่งเป็นครูผู้สอนวิชาคณิตศาสตร์ ที่มีคุณวุฒิในการจบทางด้าน การสอนคณิตศาสตร์โดยตรง ในการออกแบบหรือศึกษาการตรวจให้คะแนนครั้งต่อไป ควรศึกษาผลการตรวจให้คะแนนโดยกลุ่มผู้ตรวจที่มีคุณวุฒิไม่ตรงเอก เพื่อเปรียบเทียบผลการตรวจให้คะแนนว่าแตกต่างกันหรือไม่

3) การคัดเลือกกลุ่มตัวอย่างที่เป็นผู้ตรวจควรตรวจสอบคุณลักษณะของผู้ตรวจว่ามีค่าความใจดี เป็นกลาง และเข้มงวดอย่างไร และคัดเลือกผู้ตรวจที่มีความเป็นกลางมาตรวจให้คะแนน จึงจะเห็นความสัมพันธ์ในการตรวจให้คะแนนมากยิ่งขึ้น

References

- Koo, T. K., & Li, M. Y. (2016). A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163.
- Miller, M. D., & Legg, S. M. (1993). Alternative assessment in a high-stakes environment. *Educational Measurement: Issues and Practice*, 12(2), 9-15.

Translate Thai References

- Aphaikawee, D. (2019). *Improving the effectiveness of the subjective test scoring* [Doctoral dissertation]. Khon Kaen University. (in Thai)
- Bhu-iam, W. (2007). *A comparison of generalizability coefficient and error variance between traditional and two-tier diagnostic tests In mathematics* [Master's thesis]. Chulalongkorn University. (in Thai)
- Buasiripan, U. (2000). *Comparison of reference sum coefficients of existing math tests. The methods of examination, the number of examiners, and the experience of the inspectors differ* [Master's thesis]. Srinakharinwirot University. (in Thai)
- Chuthong, S. (2011). *Model quality examination of the learning assessment toolkit based on actual conditions, mathematics learning subject group Mathayom 1 by applying the theory of reference summary* [Master's thesis]. Srinakharinwirot University. (in Thai)
- Institute for the Promotion of Teaching Science and Technology. (2020). *Mathematical Literacy*. <https://pisathailand.ipst.ac.th/about-pisa/mathematical-literacy/> (in Thai)
- Inthanate, N. (2011). *Characteristics of the open-ended mathematics test scores for Different numbers of raters and scoring patterns using Generalizability model and many-facet rasch model* [Doctoral dissertation]. Srinakharinwirot University. (in Thai)
- Kanjanawasee, S. (2012). *New test theory* (4th Ed.). Chulalongkorn University Press. (in Thai)
- Prasertsinponma, D., & Tuksino, P. (2021). Comparison of results of subjective test for measuring mathematical competencies of primary 6: application of generalizability theory. *The 22nd National Graduate Research Conference (Online Conference)*, 22(1), 398-408. (in Thai)
- Saenpluem, B. (2013). *Using the Characteristics Testing Method and the Proportion of the Number of Reviewers on the Accuracy of the Writing Ability Measurement of Grade 3 Students* [Doctoral dissertation]. Srinakharinwirot University. (in Thai)

- Sa-nguanwai, C. (2015). *Comparison of the reliability of problem-solving ability tests mathematical creativity The application of summation theory references the reliability of measurement results* [Master's thesis]. Chulalongkorn University. (in Thai)
- Tangdhanakanond, K. (2014). *Measuring and evaluating practical skills*. Chulalongkorn University Press. (in Thai)
- Taoto, J., Jarernvongrayab, A., & Baikularb, P. (2016). A study of the reliability of mathematics essay test score for matayomsuksa 2 students: The different number of raters and scoring patterns using generalizability theory. *Hatyai Journal*, 14(1), 1-14. (in Thai)
- Thitikanpodchana, W., & Tuksino, P. (2021). Comparisons of the Results of Essay Test by Different Scoring Designs: Application of Generalizability Theory. *The 22nd National Graduate Research Conference (Online Conference)*, 22(1), 409-418. (in Thai)
- Tuksino, P. (2013). *Teaching documentation Educational research methodology*. Faculty of Education, Khon Kaen University. (in Thai)
- Umnacil, M. (2014). *Comparison of reliability of modified essay question test for measuring Scientific problem solving ability using different scoring methods Under different numbers of event: An application of generalizability theory* [Master's thesis]. Chulalongkorn University. (in Thai)

การศึกษาอิทธิพลของชีวิตการทำงานที่มีต่อภาวะสุขภาพจิต
และความเหนื่อยหน่ายในการทำงานของครู
Examining the Influences of Areas of Work-life on
Teachers' Mental Health and Job Burnout

อารยา ปิยะกุล^{1*}Araya Piyakun^{1*}

(Received: January 21, 2022; Revised: February 18, 2022; Accepted: February 28, 2022)

บทคัดย่อ

การศึกษานี้มีความมุ่งหมายเพื่อศึกษาชีวิตการทำงาน ภาวะสุขภาพจิต และความเหนื่อยหน่ายในการทำงาน และเพื่อศึกษาอิทธิพลของชีวิตการทำงานที่มีต่อภาวะสุขภาพจิตและความเหนื่อยหน่ายของครู ตัวอย่างในการศึกษานี้เป็นครูในโรงเรียนประถมศึกษาและมัธยมศึกษาจำนวน 210 คนที่อาสาเข้าร่วมตอบแบบสอบถามออนไลน์ เครื่องมือในการวิจัย ได้แก่ แบบประเมินภาวะซึมเศร้า ความวิตกกังวล ความเครียด 21 (DASS-21) แบบสำรวจชีวิตการทำงาน (AWS) และ แบบประเมินความเหนื่อยหน่ายสำหรับครูของ Maslach (MBI-ES) สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ความถี่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน และการวิเคราะห์ถดถอยพหุคูณวิธีแบบ Stepwise ผลการวิเคราะห์ข้อมูล พบว่าชีวิตการทำงาน 6 ด้าน มีคะแนนเฉลี่ยอยู่ในระดับปานกลางถึงมาก ภาวะสุขภาพจิตของครูด้านความเครียดและภาวะซึมเศร้าอยู่ในระดับปกติ ในขณะที่ความวิตกกังวลอยู่ในระดับสูงเล็กน้อย สำหรับความเหนื่อยหน่ายในการทำงาน พบว่า ค่าเฉลี่ยด้านความอ่อนล้าทางอารมณ์และความเฉา ไม่ใส่ใจอยู่ในระดับต่ำ ในขณะที่ค่าเฉลี่ยด้านประสิทธิภาพในการทำงานอยู่ในระดับสูง ผลจากการวิเคราะห์ถดถอยพหุคูณ พบว่า ปัจจัยชีวิตการทำงานด้านการควบคุมงานและภาระงานสามารถพยากรณ์สุขภาพจิตด้านความเครียดได้ร้อยละ 19.0 ชุมชนในที่ทำงานและการควบคุมงานสามารถพยากรณ์ความวิตกกังวลได้ร้อยละ 17.5 การควบคุมงานและรางวัลสามารถร่วมทำนายภาวะซึมเศร้าได้ร้อยละ 25.4 การควบคุมงาน ความยุติธรรม ภาระงานและค่านิยมสามารถพยากรณ์ความอ่อนล้าทางอารมณ์ได้ร้อยละ 34.2 ชุมชนในที่ทำงานและค่านิยมสามารถร่วมทำนายความเฉา ไม่ใส่ใจได้ร้อยละ 19.5 และชุมชนในที่ทำงานและการควบคุมงานสามารถร่วมทำนายประสิทธิภาพในการทำงานได้ร้อยละ 23.4 ผลของการศึกษานี้เป็นข้อสนเทศที่สำคัญในการตระหนักถึงความสำคัญของการปัจจัยด้านการทำงานที่มีต่อสุขภาพของครู

คำสำคัญ: ภาวะสุขภาพจิต ชีวิตการทำงาน ความเหนื่อยหน่ายในการทำงาน

¹ อาจารย์ สาขาจิตวิทยาการศึกษาและการแนะแนว คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

¹ Lecturer, Educational Psychology and Guidance, Faculty of Education, Mahasarakham University

* Corresponding Author E-mail: araya.p@msu.ac.th

Abstract

This study aimed to examine teachers' areas of work-life, mental health and job burnout, and to investigate the influences of work-life areas on mental health and burnout symptoms. The participants consisted of 210 primary and secondary school teachers who voluntarily responded to the online questionnaire. The instruments for data collection included the Depression Anxiety Stress Scales – 21 (DASS–21), Areas of Work-life Survey (AWS), and Maslach Burnout Inventory – Educators Survey (MBI – ES). The statistics used to analyze the data included frequency, mean, standard deviation, Pearson correlation coefficient, and stepwise multiple regression analysis. The results showed that the mean scores of the six areas of work-life were ranging from average to high levels. For the mental health, the participants' stress and depression were found to be at a normal level, whereas their anxiety was classified a mild level. The results of the job burnout indicated that the participants' emotional exhaustion and depersonalization mean scores were at low level, while personal achievement mean score was at high level. The findings from the multiple regression analyses revealed that control and workload could predict stress about 19%, community and control anticipated anxiety about 17.5%, control and reward predicted depression about 25.4%, control, fairness, workload and values predicted emotional exhaustion about 34.2 %; community and values predicted depersonalization about 19.5%, and, lastly, community and control could predict efficacy about 23.4%. The findings of this study provide an insight into the roles of workplace factors on teachers' mental well-being.

Keywords: mental health status, areas of work-life, job burnout

บทนำ

ครูผู้สอนเป็นกลไกที่สำคัญของระบบการศึกษาและเป็นองค์ประกอบที่มีอิทธิพลต่อการเรียนรู้ของผู้เรียน ประสิทธิภาพในการทำงานของครูจึงเป็นสิ่งสำคัญในการสนับสนุนให้ผู้เรียนสามารถพัฒนาตนเองได้อย่างเต็มศักยภาพ อย่างไรก็ตาม สภาพความเป็นจริงที่เกิดขึ้นในด้านคุณภาพชีวิตการทำงานของครู พบว่า ปัญหาเกี่ยวกับครูมีอยู่หลายมิติ โดยเฉพาะคาดหวังจากสังคมที่มีต่อครู ตัวอย่างเช่น ครูต้องเป็นตัวแทนที่ดีให้ผู้เรียน เป็นผู้แนะนำ สั่งสอนให้ผู้เรียนเกิดการเรียนรู้และมีเจตคติที่ดีตามหลักประชาธิปไตย มีสัมพันธภาพที่ดีกับผู้เรียน ผู้บริหาร เพื่อนร่วมงาน และผู้ปกครอง ซึ่งการเติมเต็มความคาดหวังเหล่านี้ ครูต้องใช้ความพยายามในการจัดการอารมณ์ของตนเองเมื่อต้องเผชิญกับสถานการณ์ที่ไม่พึงประสงค์ในการทำงาน กระบวนการนี้เรียกว่า แรงงานทางอารมณ์ (emotional labor) ที่อาจส่งผลกระทบต่อทางลบ ไม่ว่าจะเป็นความรู้สึกไม่มั่นใจในความสามารถของตนเองในการตอบสนองความคาดหวังทั้งหมดได้อย่างมีประสิทธิภาพ จนเกิดความอ่อนล้า

ทางจิตใจ และอาการเหนื่อยหน่ายจากการทำงาน (Yilmaz et al., 2015; Mahmoodi-Shahreababaki, 2018; Ratanasiripong et al., 2021) กอปรกับสถานการณ์ปัจจุบันที่มีการแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) ทำให้เกิดความไม่แน่นอนในนโยบายทางการศึกษา โรงเรียนส่วนใหญ่ในประเทศไทย กำหนดให้จัดการเรียนการสอนผ่านระบบออนไลน์เป็นหลักตามแนวทางวิถีปกติใหม่เพื่อป้องกันการแพร่ระบาดของเชื้อไวรัส ครูจำเป็นต้องปรับเปลี่ยนชุดพฤติกรรมการสอนของตนเองแบบฉับพลัน จึงทำให้ปฏิเสธไม่ได้ว่า ปัญหาต่าง ๆ เหล่านี้ก่อให้เกิดความเครียดและความวิตกกังวลซึ่งเป็นอันตรายต่อการทำงานเพราะส่งผลกระทบต่อภาวะทางสุขภาพจิตและศักยภาพในการปฏิบัติงานของครูอย่างหลีกเลี่ยงไม่ได้

ปัญหาจากการทำงานของครูที่เกิดขึ้นอย่างรวดเร็วยาวนานส่งผลให้ครูต้องเผชิญกับปัญหาทางสุขภาพจิต ไม่ว่าจะเป็นภาวะซึมเศร้า (depression) ความวิตกกังวล (anxiety) ความเครียด (stress) และความเหนื่อยหน่าย (burnout) (Martinez-Monteagudo et al., 2019) ภาวะซึมเศร้าเป็นความรู้สึกเศร้า ไม่มีความสุข หดหู่ ไร้ค่า รังเกียจตนเอง ดูถูกตนเอง ไม่สนใจหรือไม่อยากมีส่วนร่วมในกิจกรรมต่าง ๆ ขาดความคิดริเริ่ม ส่วนของความวิตกกังวลเป็นความรู้สึกกังวล มีอาการตัวสั่น ปากแห้ง หายใจยาก รู้สึกกังวล ในสิ่งที่ควบคุมไม่ได้ สำหรับความเครียดเป็นความรู้สึกที่ไม่สามารถผ่อนคลายได้ เสียใจได้ง่าย หงุดหงิด อยู่ไม่สุข ไม่อดทนต่อสิ่งที่มาทรมานใจหรือความล่าช้าที่เกิดขึ้น (depression Anxiety Stress Scales, 2018) ความเครียดที่เรื้อรังจากการทำงานนำไปสู่ปัญหาความเหนื่อยหน่าย (Jones-Rincon & Howard, 2019) ซึ่งองค์การอนามัยโลก (World Health Organization, 2019) ได้บรรจุความเหนื่อยหน่ายไว้ในบัญชีจำแนกโรคระหว่างประเทศ ฉบับปรับปรุงที่ 11 (the 11th Revision of the International Classification of Diseases: ICD – 11) ให้เป็นภาวะผิดปกติที่เกิดจากการทำงาน (occupational phenomenon) ที่ไม่ได้เป็นโรคหรือเงื่อนไขทางการแพทย์ (medical condition) ครูที่มีความเหนื่อยหน่ายในการทำงานมีอาการ 3 ลักษณะ ดังนี้ 1) รู้สึกอ่อนล้าทางจิตใจ (exhaustion) ทำให้ครูไม่สามารถอุทิศตนในงานที่ทำ 2) รู้สึกเฉยชา (cynicism) ที่จะตอบสนองความต้องการของผู้เรียน มีเจตคติทางลบต่อผู้เรียน ไม่ต้องการมีปฏิสัมพันธ์กับผู้เรียน และ 3) รู้สึกขาดศักยภาพในการทำงาน (inefficacy) เพื่อส่งเสริมให้ผู้เรียนเกิดการเรียนรู้และเติบโต (Leiter & Maslach, 2004; Maslach et al., 2018) ครูที่มีคุณภาพด้านสุขภาพจิตต่ำ มีระดับความเครียดสูง มีภาวะซึมเศร้า เป็นโรคตื่นตระหนก และเป็นโรควิตกกังวล มีแนวโน้มในการลาออกจากงานสูง (Mack et al., 2019)

นักวิจัยทางจิตวิทยาและทางการศึกษามีความพยายามในการสืบค้นถึงเหตุปัจจัยของปัญหาสุขภาพจิตและความเหนื่อยหน่ายในการทำงานของครู อย่างไรก็ตาม จากการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง พบว่า ปัจจัยที่ส่งผลต่อสุขภาพจิตและความเหนื่อยหน่ายมีมากมายหลายปัจจัย ตัวอย่างเช่น Kabito & Wami (2020) ศึกษาพบว่า ประสบการณ์การสอน (น้อยกว่า 5 ปี) ระดับการศึกษา ข้อเรียกร้องจากงานที่สูง และสัมพันธภาพที่ไม่ดี มีความสัมพันธ์กับความเครียดในการทำงานของครู นอกจากนี้ Ratanasiripong et al. (2021) ที่ศึกษาปัจจัยที่ส่งผลต่อสุขภาพจิตและความเหนื่อยหน่ายของครูในประเทศไทย พบว่า สถานะทางเศรษฐกิจของครอบครัว คุณภาพของความสัมพันธ์ ความเข้มแข็งทางใจ ขนาดของห้องเรียน การนอนหลับ และเพศ มีอิทธิพลต่อสุขภาพจิตของครู ในขณะที่คุณภาพของความสัมพันธ์และอายุเป็นตัวแปรพยากรณ์ที่สำคัญของความเหนื่อยหน่าย ซึ่งความหลากหลายในตัวแปรพยากรณ์นี้ทำให้ยากต่อการ

วางแผนและออกแบบนโยบายหรือแนวทางในการแก้ไขปัญหาของครู ดังนั้น การศึกษาปัจจัยทางตรงโดยเฉพาะบริบทของการทำงานที่ส่งผลต่อปัญหาสุขภาพจิตและความเหนื่อยหน่ายในการทำงานของครูจึงมีความสำคัญและจำเป็นอย่างยิ่ง

หลากหลายปีที่ผ่านมา นักจิตวิทยาให้ความสนใจเกี่ยวกับความสอดคล้องระหว่างบุคคลกับงาน (job-person fit) และผลกระทบต่อสุขภาพจิตในการทำงาน โดยงานวิจัยในอดีตพบว่า ความสอดคล้องระดับสูงสามารถพยากรณ์การปรับตัวและระดับความเครียดที่ลดลง (Leiter & Maslach, 2011) ดังนั้น รูปแบบของความสอดคล้อง-ไม่สอดคล้องระหว่างบุคคลกับงานจึงเป็นประเด็นสำคัญเพื่อความเข้าใจปัญหาสุขภาพจิตและความเหนื่อยหน่ายของครู จากการศึกษาค้นคว้าของ Maslach & Leiter (1997) ได้มีการสรุปลักษณะของความไม่สอดคล้องระหว่างบุคคลกับงานหรือที่เรียกว่า ชีวิตการทำงาน (areas of work life) ใน 6 ด้าน ดังนี้ 1) ภาระงานที่มากเกินไปจนไม่มีเวลาพัก ระยะเวลาในการทำงานที่ยาวนานเกินไป และความซับซ้อนของงานที่ได้รับ 2) การไม่สามารถควบคุมในการใช้แหล่งทรัพยากรและไม่สามารถตัดสินใจในการทำงาน 3) การได้รับรางวัลที่ไม่เพียงพอไม่ว่าจะเป็นเงิน การได้รับการยกย่อง และความมั่นคงในงาน 4) ปัญหาความสัมพันธ์ระหว่างบุคคลในชุมชนงาน ปัญหาในการทำงานเป็นทีม และการทำงานแบบต่างคนต่างทำ 5) ความไม่ยุติธรรมซึ่งเป็นปัญหาด้านความเท่าเทียม การเปิดเผย การไว้วางใจ และความเคารพกัน และ 6) ค่านิยมส่วนบุคคลกับค่านิยมขององค์กรที่ขัดแย้งกันส่งผลกระทบต่อความผูกพันที่บุคคลมีต่อองค์กร จากการศึกษาของ Maslach (2019) เกี่ยวกับความสัมพันธ์ระหว่างชีวิตการทำงานกับความเหนื่อยหน่าย พบว่า คนที่มีระดับความอ่อนล้าสูงมักจะมีปัญหาภาระงานที่มากเกินไป ในขณะที่คนที่มีระดับความเฉื่อยชาสูงจะแยกตัวโดดเดี่ยว ในขณะที่คนที่ขาดประสิทธิภาพในการทำงานจะไม่มีอาการอ่อนล้าหรือเฉื่อยชา แต่จะทำงานอย่างขาดคุณภาพ

แม้ว่าปัญหาสุขภาพจิตและความเหนื่อยหน่ายของครูเป็นประเด็นสำคัญที่ต้องศึกษา แต่งานวิจัยในประเทศไทยยังมีน้อยมาก (Ratanasiripong et al., 2021) โดยเฉพาะการศึกษาเกี่ยวกับปัจจัยเหตุของปัญหาสุขภาพจิตของครู (Ploytabtim et al., 2021) ทำให้ขาดความชัดเจนเกี่ยวกับปัจจัยเชิงองค์กรในแง่ของความสอดคล้องระหว่างบุคคลและงานที่มีผลต่อปัญหาสุขภาพจิตของครูในประเทศไทย นอกจากนี้ ผู้วิจัยพบว่า งานวิจัยที่ศึกษาเกี่ยวกับสุขภาพจิตและความเหนื่อยหน่ายของครูในประเทศไทยมีขอบเขตของการศึกษาจำกัดอยู่เฉพาะกลุ่ม เช่น กลุ่มโรงเรียน เขตพื้นที่ สังกัด ระดับชั้นที่สอน เป็นต้น ความไม่เพียงพอของข้อมูลนี้ส่งผลให้นักจิตวิทยาและนักศึกษายังขาดองค์ความรู้ที่สำคัญเกี่ยวกับความเชื่อมโยงระหว่างสภาพการทำงานในองค์กรกับปัญหาสุขภาพจิตของครู

ด้วยเหตุผลดังกล่าว งานวิจัยครั้งนี้จึงมีความมุ่งหมายที่จะศึกษาชีวิตการทำงานทั้ง 6 ด้าน (ภาระงาน การควบคุมงาน รางวัล ชุมชนในที่ทำงาน ความยุติธรรม และค่านิยม) สุขภาพจิตทั้ง 3 ด้าน (ความเครียด ความวิตกกังวล และภาวะซึมเศร้า) และความเหนื่อยหน่ายทั้ง 3 ด้าน (ความอ่อนล้าทางอารมณ์ ความเฉื่อยชา ไม่ใส่ใจ และประสิทธิภาพในการทำงาน) ของครูในโรงเรียนประถมศึกษาและมัธยมศึกษา รวมทั้งศึกษาอิทธิพลของชีวิตการทำงานในแต่ละด้านที่มีต่อคุณภาพทางสุขภาพจิตของครู เพื่อเติมเต็มช่องว่างขององค์ความรู้เกี่ยวกับปัจจัยเชิงองค์กรที่ส่งผลต่อสุขภาพจิตของครู และเพื่อให้ได้มาซึ่งข้อเสนอแนะสำหรับครู ผู้บริหารและบุคคลที่เกี่ยวข้องกับการพัฒนาบุคลากรทางการศึกษาของประเทศได้ตระหนักและเข้าใจถึงเหตุปัจจัยทางจิตวิทยาและองค์กร

ที่มีผลกระทบต่อประสิทธิภาพในการทำงาน สามารถนำไปใช้ในการกำหนดนโยบายการส่งเสริมและพัฒนาครูให้มีสุขภาวะทางจิต มีความพร้อมทั้งด้านร่างกาย ความรู้ และจิตใจในการปฏิบัติงานในโรงเรียนได้อย่างมีประสิทธิภาพ

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาชีวิตการทำงาน ภาวะสุขภาพจิต และความเหนื่อยหน่ายในการทำงานของครู
2. เพื่อศึกษาอิทธิพลของชีวิตการทำงานที่มีต่อภาวะสุขภาพจิตและความเหนื่อยหน่ายในการทำงานของครู

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงปริมาณ ผู้วิจัยมีการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องกับตัวแปรที่ศึกษา ประกอบด้วย ชีวิตการทำงาน (ภาระงาน การควบคุมงาน ชุมชนในที่ทำงาน รางวัล ความยุติธรรม และค่านิยม) สุขภาพจิต (ความเครียด ความวิตกกังวล และภาวะซึมเศร้า) และความเหนื่อยหน่ายในการทำงาน (ความอ่อนล้าทางอารมณ์ ความเฉยชา ไม่ใส่ใจ และประสิทธิภาพในการทำงาน) โดยดำเนินการสอบถามผ่านระบบออนไลน์

1. ประชากรและตัวอย่าง

ประชากรในการวิจัยครั้งนี้ เป็นครูผู้สอนที่ปฏิบัติงานในโรงเรียนระดับประถมศึกษาและระดับมัธยมศึกษาที่สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน สังกัดองค์กรปกครองส่วนท้องถิ่น และโรงเรียนเอกชนในประเทศไทย

ตัวอย่างในการวิจัยเป็นครูผู้สอนที่อาสาเข้าร่วมตอบแบบสอบถามออนไลน์ ผู้วิจัยกำหนดขนาดตัวอย่างโดยการพิจารณาจากความมุ่งหมายของการวิจัยเพื่อทราบถึงอิทธิพลของชีวิตการทำงานที่มีต่อภาวะสุขภาพจิตและความเหนื่อยหน่ายในการทำงาน โดยใช้การวิเคราะห์การถดถอยพหุคูณ ซึ่ง Steven (1996, as cited in Phusee-orn, 2015) ได้เสนอให้มีการกำหนดขนาดของตัวอย่างเป็น 15 เท่าของจำนวนตัวแปรพยากรณ์ โดยงานวิจัยครั้งนี้มีตัวแปรพยากรณ์ จำนวน 6 ตัว จึงใช้ตัวอย่างอย่างน้อย 90 คน เนื่องจากการวิจัยครั้งนี้ ผู้วิจัยมีข้อจำกัดของงบประมาณในการขอลิขสิทธิ์ใช้แบบสำรวจชีวิตการทำงานและแบบประเมินความเหนื่อยหน่ายในการทำงาน ผู้วิจัยจึงกำหนดขนาดตัวอย่าง 210 คนซึ่งเป็นจำนวนที่เป็นไปตามหลักการสำหรับวิธีการได้มาซึ่งตัวอย่างเป็นแบบตามสะดวก (convenience sampling)

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบถามออนไลน์ ประกอบด้วย

2.1 แบบประเมินสุขภาพจิต เป็นข้อคำถามเกี่ยวกับสภาวะอารมณ์เชิงลบ ประกอบด้วย ภาวะซึมเศร้า ความวิตกกังวล และความเครียด โดยใช้ Depression Anxiety Stress Scales – 21 (DASS-21) พัฒนาโดย Lovibond & Lovibond (1995) ผู้วิจัยนำแบบประเมินฉบับแปลภาษาไทยซึ่งเป็นแบบมาตราส่วนประมาณค่า 4 ระดับ ไปทดลองใช้กับครูที่ไม่ใช่ตัวอย่างจำนวน 40 คน เพื่อหาค่าอำนาจจำแนกตามวิธี item

total correlation และค่าความเชื่อมั่นโดยใช้สัมประสิทธิ์แอลฟาของ Cronbach การแปลผลการประเมิน DASS-21 จะต้องนำคะแนนรวมแต่ละด้านคูณด้วย 2 และจะต้องไม่นำผลทั้ง 3 ด้านมารวมกัน (Lovibond & Lovibond, 1995) การหาค่าความเชื่อมั่นจึงเป็นการวิเคราะห์ข้อมูลรายด้าน จากผลการวิเคราะห์พบว่า ภาวะซึมเศร้าจำนวน 7 ข้อ มีค่าอำนาจจำแนก 0.447-0.793 และค่าความเชื่อมั่น 0.886 ความวิตกกังวลจำนวน 7 ข้อ มีค่าอำนาจจำแนก 0.484-0.829 และค่าความความเชื่อมั่น 0.839 และความเครียดจำนวน 7 ข้อ มีค่าอำนาจจำแนก 0.399-0.790 และค่าความเชื่อมั่น 0.836

2.2 แบบสำรวจชีวิตการทำงาน เป็นข้อคำถามเกี่ยวกับความสอดคล้องระหว่างบุคคลกับงาน 6 ด้าน ผู้วิจัยใช้ Areas of Worklife Survey (AWS) ของ Leiter และ Maslach เนื่องด้วย Leiter และ Maslach พัฒนา AWS โดยใช้ตัวอย่างจากหลากหลายอาชีพรวมถึงวิชาชีพครู จึงทำให้ AWS ได้รับการยอมรับและถูกนำมาใช้ในงานวิจัยที่ศึกษาชีวิตการทำงานของครูในหลายประเทศ (Mojsa-Kaja et al., 2015; Berquam, 2020) ลิขสิทธิ์การใช้ AWS ดูแลโดย Mind Garden ซึ่งให้สิทธิ์ผู้วิจัยในการใช้แบบสำรวจผ่านระบบออนไลน์ ภายใต้ข้อกำหนดในจำนวนการตอบรับ ผู้วิจัยนำแบบสำรวจภาษาอังกฤษมาแปลเป็นภาษาไทย และใช้เทคนิคการแปลกลับ (back-translation technique) ต่อมา ผู้วิจัยนำแบบสำรวจภาษาไทยซึ่งเป็นแบบมาตราส่วนประมาณค่า 5 ระดับ ไปทดลองใช้กับครูที่ไม่ใช่ตัวอย่างจำนวน 40 คน เพื่อหาค่าอำนาจจำแนกและค่าความเชื่อมั่น เนื่องด้วยการแปลผลคะแนนจะเป็นการคำนวณในแต่ละด้านไม่สามารถนำผลจาก 6 ด้านมารวมกัน (Leiter & Maslach, 2011) จากผลการวิเคราะห์ข้อมูลแต่ละด้าน พบว่า ภาระงาน จำนวน 4 ข้อ มีค่าอำนาจจำแนก 0.554-0.834 และค่าความความเชื่อมั่น 0.838 การควบคุมงานจำนวน 4 ข้อ มีค่าอำนาจจำแนก 0.724-0.812 และค่าความเชื่อมั่น 0.890 รางวัลจำนวน 3 ข้อ มีค่าอำนาจจำแนก 0.498-0.650 และค่าความความเชื่อมั่น 0.738 ชุมชนในที่ทำงานจำนวน 5 ข้อ มีค่าอำนาจจำแนก 0.533-0.768 และค่าความความเชื่อมั่น 0.806 ความยุติธรรมจำนวน 5 ข้อ มีค่าอำนาจจำแนก 0.658-0.844 และค่าความเชื่อมั่น 0.885 และค่านิยมจำนวน 4 ข้อ มีค่าอำนาจจำแนก 0.643-0.803 และค่าความความเชื่อมั่น 0.865

2.3 แบบประเมินความเหนื่อยหน่ายสำหรับครูผู้สอนที่ใช้ในการวิจัยครั้งนี้ คือ Maslach Burnout Inventory – Educators Survey (MBI – ES) พัฒนาโดย Maslach, Jackson และ Schwab ที่มี Mind Garden เป็นผู้ดูแลลิขสิทธิ์ และให้สิทธิ์ผู้วิจัยในการใช้แบบประเมินภายใต้เงื่อนไขเดียวกับแบบสำรวจชีวิตการทำงาน ผู้วิจัยนำแบบประเมินภาษาอังกฤษมาแปลเป็นภาษาไทยด้วยวิธีการแปลกลับตามขั้นตอนเดียวกับแบบสำรวจชีวิตการทำงาน แบบประเมินความเหนื่อยหน่ายเป็นแบบมาตราส่วนประมาณค่า 7 ระดับ เนื่องด้วยการแปลความหมายของคะแนนควรนำเสนอแต่ละด้านและไม่ควรนำเสนอคะแนนรวมของความเหนื่อยหน่าย (Maslach et al., 2018) ผลการทดลองใช้แบบประเมินกับครูที่ไม่ใช่ตัวอย่างจำนวน 40 คน เพื่อหาค่าอำนาจจำแนกและค่าความเชื่อมั่น พบว่า ความอ่อนล้าทางอารมณ์จำนวน 9 ข้อ มีค่าอำนาจจำแนก 0.620-0.894 และค่าความความเชื่อมั่น 0.951 ความเฉยชา ไม่ใส่ใจจำนวน 5 ข้อ มีค่าอำนาจจำแนก 0.500-0.806 และค่าความความเชื่อมั่น 0.837 และประสิทธิภาพในการทำงานจำนวน 8 ข้อ มีค่าอำนาจจำแนก 0.274-0.578 และค่าความความเชื่อมั่น 0.760

3. การเก็บรวบรวมข้อมูล

ก่อนดำเนินการเก็บรวบรวมข้อมูล ผู้วิจัยดำเนินการขอรับรองจริยธรรมการวิจัยในมนุษย์จากคณะกรรมการจริยธรรมการวิจัยฯ ของมหาวิทยาลัยมหาสารคาม ซึ่งผลการพิจารณาจากคณะกรรมการฯ ได้รับรองจริยธรรมการวิจัยฯ (เลขที่การรับรอง 005-333/2564) การวิจัยครั้งนี้ ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลผ่านระบบออนไลน์ และเลือกวิธีการได้มาซึ่งตัวอย่างแบบตามสะดวก ซึ่งเป็นวิธีการที่มักถูกใช้ในการศึกษาทางจิตวิทยา โดยเฉพาะงานวิจัยที่ไม่สามารถเข้าถึงกลุ่มคนบางกลุ่มที่สำคัญได้ หรือเป็นกลุ่มบุคคลที่อยู่ห่างไกล (Morling, 2018) เนื่องด้วยการวิจัยครั้งนี้มีวัตถุประสงค์ในการศึกษาชีวิตการทำงาน ภาวะสุขภาพจิตและความเหนื่อยหน่ายในการทำงานของครูที่ครอบคลุมทั่วประเทศ โดยเก็บข้อมูลในช่วงเดือนพฤศจิกายน ถึง ธันวาคม พ.ศ.2564 ซึ่งมีการแพร่ระบาดของเชื้อไวรัสโคโรนา 2019 ทำให้การเข้าถึงครูในทุกภูมิภาคมีข้อจำกัด ผู้วิจัยจึงดำเนินการเก็บข้อมูลผ่านระบบออนไลน์เพื่อให้ครูในทุกภูมิภาคสามารถตอบแบบสอบถามได้ตามสะดวก ผู้วิจัยจึงดำเนินการเผยแพร่ข้อความประชาสัมพันธ์ในสื่อสังคมออนไลน์ เพื่อขออาสาสมัครตอบแบบสอบถามผ่าน Google form โดยกำหนดผู้ตอบแบบสอบถาม คือ ครูผู้สอนระดับประถมศึกษาและระดับมัธยมศึกษาไม่จำกัดสังกัดของหน่วยงาน เมื่อจำนวนตอบรับของตัวอย่างครบ 210 คน ตามจำนวนที่ผู้วิจัยได้รับสิทธิ์ในการใช้แบบสำรวจชีวิตการทำงานและแบบประเมินความเหนื่อยหน่ายสำหรับครู ผู้วิจัยทำการปิดรับคำตอบในแบบสอบถามออนไลน์ โดยแบบสอบถามที่ได้รับการตอบกลับทั้งหมดมีความครบถ้วนสมบูรณ์ และไม่พบข้อมูลขาดหาย

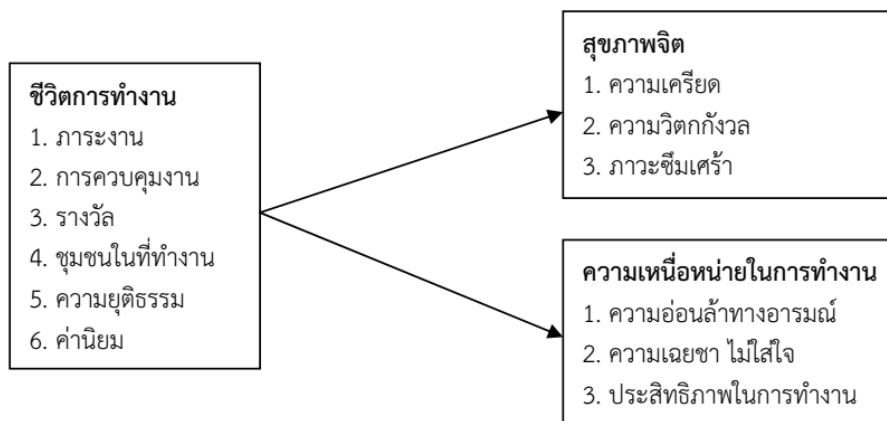
4. การวิเคราะห์ข้อมูล

ผู้วิจัยนำเครื่องมือฉบับภาษาไทยไปทดลองใช้กับครูที่ไม่ใช่ตัวอย่างจำนวน 40 คน เพื่อหาค่าอำนาจจำแนกตามวิธี item total correlation และค่าความเชื่อมั่นโดยใช้สัมประสิทธิ์แอลฟาของ Cronbach จากนั้น ผู้วิจัยวิเคราะห์ข้อมูลเพื่อตอบความมุ่งหมายของการวิจัยข้อที่ 1 โดยใช้สถิติเชิงพรรณนา ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และวิเคราะห์ข้อมูลโดยใช้สถิติเชิงอนุมาน เพื่อตอบความมุ่งหมายของการวิจัยข้อที่ 2 ได้แก่ การวิเคราะห์ถดถอยพหุคูณ โดยวิธีแบบ Forward

5. กรอบแนวคิดในการวิจัย

กรอบแนวคิดในการวิจัยครั้งนี้ มาจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับชีวิตการทำงาน สุขภาพจิต และความเหนื่อยหน่ายในการทำงานของครู โดยชีวิตการทำงานเป็นความสอดคล้องระหว่างบุคคลกับงานใน 6 ด้าน ได้แก่ ภาระงาน (workload) ที่เกี่ยวข้องกับความสอดคล้องระหว่างปริมาณของงานที่ได้รับกับความสามารถในการจัดการงานของบุคคล การควบคุมงาน (control) ที่เกี่ยวข้องโอกาสให้ครูตัดสินใจแก้ปัญหาและรับผิดชอบงาน รางวัล (reward) ที่เกี่ยวข้องกับความสอดคล้องระหว่างความคาดหวังของบุคคลกับการยอมรับและการสนับสนุนทางการเงินที่ได้รับ ชุมชนในที่ทำงาน (community) ที่เกี่ยวข้องกับความผูกพันทางสังคมทั้งด้านการทำงาน ทีม การได้รับการสนับสนุนและความรู้สึกเชิงบวกกับเพื่อนร่วมงาน หัวหน้างาน ความยุติธรรม (fairness) เกี่ยวข้องกับบทบาทหน้าที่ที่บุคคลได้รับอย่างเท่าเทียม และค่านิยม (values) เป็นความสอดคล้องระหว่างค่านิยมขององค์กรกับตัวบุคคล (Leiter & Maslach, 2011) ความสอดคล้องระหว่างบุคคลกับงานทั้ง 6 ด้านนี้ส่งผลต่อภาวะสุขภาพจิต ไม่ว่าจะเป็นความเครียด ความวิตกกังวล และภาวะ

ซึมเศร้า นอกจากนี้ ปัญหาสุขภาพจิตที่เรื้อรังก็ส่งผลให้บุคคลเกิดความอ่อนล้าทางอารมณ์ ความรู้สึกเฉยชา ไม่อยากใส่ใจ และขาดประสิทธิภาพในการทำงาน



ภาพ 1 กรอบแนวคิดในการวิจัย

ผลการวิจัย

จากการวิเคราะห์ข้อมูล สามารถสรุปผลการวิจัยได้ดังนี้

1. ผลการศึกษาชีวิตการทำงาน ภาวะสุขภาพจิต และความเหนื่อยหน่ายในการทำงานของครู

ตาราง 1 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของชีวิตการทำงาน ภาวะสุขภาพจิต และความเหนื่อยหน่ายในการทำงานของครู (n = 210)

ตัวแปร	M	SD	แปลผล
ชีวิตการทำงาน			
ภาระงาน	3.18	0.89	ปานกลาง
การควบคุมงาน	3.78	0.80	มาก
รางวัล	3.51	0.92	มาก
ชุมชนในที่ทำงาน	3.90	0.59	มาก
ความยุติธรรม	3.14	0.75	ปานกลาง
ค่านิยม	3.84	0.67	มาก
ภาวะสุขภาพจิต			
ความเครียด	12.37	7.76	ปกติ
ความวิตกกังวล	7.16	7.22	สูงเล็กน้อย
ภาวะซึมเศร้า	5.85	6.96	ปกติ
ความเหนื่อยหน่าย			
ความอ่อนล้าทางอารมณ์	1.97	1.45	ต่ำ
ความเฉยชา ไม่ใส่ใจ	0.85	1.07	ต่ำ
ประสิทธิภาพในการทำงาน	4.45	1.05	สูง

จากตาราง 1 พบว่า ชีวิตการทำงานในทุกด้านมีคะแนนเฉลี่ยสูงกว่าค่ากึ่งกลาง (midpoint) ที่ 3.00 แปลความหมายตาม AWS manual ได้ว่า ชีวิตการทำงานของครูในทุกด้านมีความสอดคล้องระหว่างบุคคลกับองค์กร (Leiter & Maslach, 2011) โดยด้านชุมชนในที่ทำงานมีคะแนนเฉลี่ยสูงสุด ($M = 3.90$, $SD = 0.59$) รองลงมา คือ ค่านิยม ($M = 3.84$, $SD = 0.67$) ในขณะที่ด้านความยุติธรรมและภาระงานคะแนนเฉลี่ยน้อยที่สุด ($M = 3.14$, $SD = 0.75$ และ $M = 3.18$, $SD = 0.89$ ตามลำดับ) สำหรับภาวะสุขภาพจิต ผลการแปลความหมายของคะแนนเฉลี่ยเมื่อเทียบกับตารางแปลความหมายในคู่มือมาตรวัดภาวะซึมเศร้า ความวิตกกังวล และความเครียด (Lovibond & Lovibond, 1995) พบว่า ครูมีภาวะสุขภาพจิตด้านความเครียดและภาวะซึมเศร้าในระดับปกติ (normal) ในขณะที่ความวิตกกังวลอยู่ในระดับสูงเล็กน้อย (mild) สำหรับความเหนื่อยหน่ายในการทำงาน พบว่า ครูมีค่าเฉลี่ยด้านความเฉยชา ไม่ใส่ใจใญ่่น้อยที่สุด ($M = 0.85$, $SD = 1.07$) รองลงมา คือ ด้านความอ่อนล้าทางอารมณ์ ($M = 1.97$, $SD = 1.45$) ซึ่งแสดงถึงความเหนื่อยหน่ายของครูอยู่ในระดับต่ำ ในขณะที่ด้านประสิทธิภาพในการทำงานอยู่ในระดับสูง ($M = 4.45$, $SD = 1.05$) ซึ่งเรียกรูปแบบนี้ว่า การมีส่วนร่วม (engaged profile)

2. ผลการวิเคราะห์การถดถอยพหุคูณแบบ Stepwise เพื่อศึกษาอิทธิพลของชีวิตการทำงานของครูที่มีต่อภาวะสุขภาพจิต และความเหนื่อยหน่าย

ตาราง 2 การวิเคราะห์การถดถอยพหุคูณของปัจจัยชีวิตการทำงานที่มีอิทธิพลต่อความเครียด (stress)

ตัวแปรพยากรณ์	b	SE	Beta	t	Sig.	Tolerance	VIF
(Constant)	30.283	2.615		11.580	.000		
Control	-3.227	.630	-.333	-5.123	.000	.927	1.078
Workload	-1.799	.568	-.206	-3.169	.002	.927	1.078
R = .436 R ² = .190 R ² _{adj} = .182 SE _{est} = 7.01447							
F = 24.288 (.000)							

จากตาราง 2 พบว่า ปัจจัยชีวิตการทำงานมี 2 ตัวแปร ได้แก่ การควบคุมงาน (control) และภาระงาน (workload) มีอิทธิพลต่อความเครียดอย่างมีนัยสำคัญทางสถิติที่ .01 โดยมีค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณ (R) เท่ากับ .436 ค่าสัมประสิทธิ์การพยากรณ์ (R²) เท่ากับ .190 หรือร้อยละ 19.00 มีค่าสัมประสิทธิ์ของตัวแปรพยากรณ์ในรูปคะแนนดิบ (b) เท่ากับ -3.227 และ -1.799 ตามลำดับ มีค่าสัมประสิทธิ์ของตัวแปรพยากรณ์ในรูปคะแนนมาตรฐาน (Beta) เท่ากับ -.333 และ -.206 ตามลำดับ และมีค่าคงที่ของสมการพยากรณ์ในรูปคะแนนดิบ เท่ากับ 30.283 โดยเขียนเป็นสมการได้ ดังนี้

สมการพยากรณ์ในรูปคะแนนดิบ

$$\text{stress} = 30.283 - 3.227\text{control} - 1.799\text{workload}$$

สมการพยากรณ์ในรูปคะแนนมาตรฐาน

$$Z_{\text{Stress}} = -.333Z_{\text{Control}} - .206Z_{\text{Workload}}$$

ตาราง 3 การวิเคราะห์การถดถอยพหุคูณของปัจจัยชีวิตการทำงานที่มีอิทธิพลต่อความวิตกกังวล (anxiety)

ตัวแปรพยากรณ์	b	SE	Beta	t	Sig.	Tolerance	VIF
(Constant)	26.984	3.150		8.565	.000		
Community	-2.985	.902	-.242	-3.310	.001	.744	1.344
Control	-2.165	.660	-.240	-3.279	.001	.744	1.344
R = .418 R ² = .175 R ² _{adj} = .167 SE _{est} = 6.58597							
F = 21.970 (.000)							

จากตาราง 3 พบว่า ปัจจัยชีวิตการทำงานมี 2 ตัวแปร ได้แก่ ชุมชนในที่ทำงาน (community) และการควบคุมงาน (control) มีอิทธิพลต่อความวิตกกังวลอย่างมีนัยสำคัญทางสถิติที่ .01 โดยมีค่า R เท่ากับ .418 ค่า R² เท่ากับ .175 หรือร้อยละ 17.50 มีค่า b เท่ากับ -2.985 และ -2.165 ตามลำดับ มีค่า Beta เท่ากับ -.242 และ -.240 ตามลำดับ และมีค่าคงที่ของสมการพยากรณ์ในรูปคะแนนดิบ เท่ากับ 26.984 โดยเขียนเป็นสมการได้ ดังนี้

สมการพยากรณ์ในรูปคะแนนดิบ

$$\text{anxiety} = 26.984 - 2.985\text{community} - 2.165\text{control}$$

สมการพยากรณ์ในรูปคะแนนมาตรฐาน

$$Z_{\text{Anxiety}} = -.242Z_{\text{Community}} - .240Z_{\text{Control}}$$

ตาราง 4 การวิเคราะห์การถดถอยพหุคูณของปัจจัยชีวิตการทำงานที่มีอิทธิพลต่อภาวะซึมเศร้า (depression)

ตัวแปรพยากรณ์	b	SE	Beta	t	Sig.	Tolerance	VIF
(Constant)	24.900	2.309		10.786	.000		
Control	-3.122	.540	-.359	-5.779	.000	.934	1.071
Reward	-2.069	.470	-.273	-4.400	.000	.934	1.071
R = .504 R ² = .254 R ² _{adj} = .247 SE _{est} = 6.03775							
F = 35.237 (.000)							

จากตาราง 4 พบว่า ปัจจัยชีวิตการทำงานมี 2 ตัวแปร ได้แก่ การควบคุมงาน (control) และรางวัล (reward) มีอิทธิพลต่อภาวะซึมเศร้าอย่างมีนัยสำคัญทางสถิติที่ .01 โดยมีค่า R เท่ากับ .504 ค่า R² เท่ากับ .254 หรือร้อยละ 25.40 มีค่า b เท่ากับ -3.122 และ -2.069 ตามลำดับ มีค่า Beta เท่ากับ -.359 และ -.273 ตามลำดับ และมีค่าคงที่ของสมการพยากรณ์ในรูปคะแนนดิบ เท่ากับ 24.900 โดยเขียนเป็นสมการได้ดังนี้

สมการพยากรณ์ในรูปคะแนนดิบ

$$\text{depression} = 24.900 - 3.122\text{control} - 2.069\text{reward}$$

สมการพยากรณ์ในรูปคะแนนมาตรฐาน

$$Z_{\text{Depression}} = -.359Z_{\text{Control}} - .273Z_{\text{Reward}}$$

ตาราง 5 การวิเคราะห์การถดถอยพหุคูณของปัจจัยชีวิตการทำงานที่มีอิทธิพลต่อความอ่อนล้าทางอารมณ์ (EE)

ตัวแปรพยากรณ์	b	SE	Beta	t	Sig.	Tolerance	VIF
(Constant)	7.283	.548		13.302	.000		
Control	-.454	.122	-.251	-3.722	.000	.704	1.421
Fairness	-.333	.134	-.172	-2.473	.014	.660	1.516
Workload	-.403	.105	-.248	-3.853	.000	.777	1.286
Values	-.331	.156	-.154	-2.123	.035	.610	1.641
R = .585 R ² = .342 R ² _{adj} = .329 SE _{est} = 1.18456							
F = 26.609 (.000)							

จากตาราง 5 พบว่า ปัจจัยชีวิตการทำงานมี 4 ตัวแปร ได้แก่ การควบคุมงาน (control) ความยุติธรรม (fairness) ภาระงาน (workload) และค่านิยม (values) มีอิทธิพลต่อความอ่อนล้าทางอารมณ์ อย่างมีนัยสำคัญทางสถิติที่ .05 โดยมีค่า R เท่ากับ .585 ค่า R² เท่ากับ .342 หรือร้อยละ 34.20 มีค่า b เท่ากับ -.454, -.333, -.403 และ -.331 ตามลำดับ มีค่า Beta เท่ากับ -.251, -.172, -.248 และ -.154 ตามลำดับ และมีค่าคงที่ของสมการพยากรณ์ในรูปคะแนนดิบ เท่ากับ 7.283 โดยเขียนเป็นสมการได้ ดังนี้

สมการพยากรณ์ในรูปคะแนนดิบ

$$EE = 7.283 - .454\text{control} - .333\text{fairness} - .403\text{workload} - .331\text{values}$$

สมการพยากรณ์ในรูปคะแนนมาตรฐาน

$$Z_{EE} = -.251Z_{\text{Control}} - .172Z_{\text{Fairness}} - .248Z_{\text{Workload}} - .154Z_{\text{Values}}$$

ตาราง 6 การวิเคราะห์การถดถอยพหุคูณของปัจจัยชีวิตการทำงานที่มีอิทธิพลต่อความเฉชา ไม่ใส่ใจ (DP)

ตัวแปรพยากรณ์	b	SE	Beta	t	Sig.	Tolerance	VIF
(Constant)	4.284	.490		8.751	.000		
Community	-.595	.130	-.325	-4.567	.000	.768	1.302
Values	-.290	.113	-.182	-2.553	.011	.768	1.302
R = .442 R ² = .195 R ² _{adj} = .188 SE _{est} = .69724							
F = 25.123 (.000)							

จากตาราง 6 พบว่า ปัจจัยชีวิตการทำงานมี 2 ตัวแปร ได้แก่ ชุมชนในที่ทำงาน (community) และค่านิยม (values) มีอิทธิพลต่อความเฉชา ไม่ใส่ใจอย่างมีนัยสำคัญทางสถิติที่ .05 โดยมีค่า R เท่ากับ .442 ค่า R² เท่ากับ .195 หรือร้อยละ 19.50 มีค่า b เท่ากับ -.595 และ -.290 ตามลำดับ มีค่า Beta เท่ากับ -.325 และ -.182 ตามลำดับ และมีค่าคงที่ของสมการพยากรณ์ในรูปคะแนนดิบ เท่ากับ 4.284 โดยเขียนเป็นสมการได้ ดังนี้

สมการพยากรณ์ในรูปคะแนนดิบ

$$DP = 4.284 - .595\text{community} - .290\text{values}$$

สมการพยากรณ์ในรูปคะแนนมาตรฐาน

$$Z_{DP} = -.325Z_{\text{Community}} - .182Z_{\text{Values}}$$

ตาราง 7 การวิเคราะห์การถดถอยพหุคูณของปัจจัยชีวิตการทำงานที่มีอิทธิพลต่อประสิทธิภาพในการทำงาน (PA)

ตัวแปรพยากรณ์	b	SE	Beta	t	Sig.	Tolerance	VIF
(Constant)	1.077	.444		2.427			
Community	.529	.127	.294	4.163	.000	.744	1.344
Control	.347	.093	.263	3.731	.000	.744	1.344
R = .483 R ² = .234 R ² _{adj} = .226 SE _{est} = 0.92765							
F = 31.557 (.000)							

จากตาราง 7 พบว่า ปัจจัยชีวิตการทำงานมี 2 ตัวแปร ได้แก่ ชุมชนในที่ทำงาน (community) และการควบคุมงาน (control) มีอิทธิพลต่อประสิทธิภาพในการทำงานอย่างมีนัยสำคัญทางสถิติที่ .01 โดยมีค่า R เท่ากับ .483 ค่า R² เท่ากับ .234 หรือร้อยละ 23.4 มีค่า b เท่ากับ .539 และ .347 ตามลำดับ มีค่า Beta เท่ากับ .294 และ .263 ตามลำดับ และมีค่าคงที่ของสมการพยากรณ์ในรูปคะแนนดิบ เท่ากับ 1.077 โดยเขียนเป็นสมการได้ ดังนี้

สมการพยากรณ์ในรูปคะแนนดิบ

$$PA = 1.077 + .529community + .347control$$

สมการพยากรณ์ในรูปคะแนนมาตรฐาน

$$Z_{PA} = .294Z_{Community} + .263Z_{Control}$$

อภิปรายผล

จากผลการวิจัย สามารถนำสู่การอภิปราย ดังนี้

1. ผลการวิเคราะห์ข้อมูลชีวิตการทำงาน ภาวะสุขภาพจิต และความเหนื่อยหน่าย พบว่า ชีวิตการทำงานของครูทั้ง 6 ด้านมีความสอดคล้องระหว่างความคาดหวังหรือการรับรู้ส่วนบุคคลกับสภาพแวดล้อมของงาน อาจจะเนื่องมาจากผู้ประกอบวิชาชีพครูได้ผ่านกระบวนการผลิตและพัฒนาครูที่มีการสร้างความรู้ พัฒนาประสบการณ์วิชาชีพและการปฏิบัติงานตามมาตรฐานวิชาชีพ ซึ่งกระบวนการนี้หล่อหลอมให้บุคคลมีสมรรถนะตามบทบาทหน้าที่และความรับผิดชอบในฐานะครู ตลอดจนระบบการทำงานในโรงเรียน ทำให้มีความพร้อมและสร้างความคาดหวังได้สอดคล้องกับสภาพแวดล้อมการทำงานที่แท้จริง และอาจจะส่งผลให้ครูมีสุขภาพจิตด้านความเครียด และภาวะซึมเศร้าอยู่ในระดับปกติ อย่างไรก็ตาม ผลการวิเคราะห์ข้อมูลพบว่า ความวิตกกังวลของครูอยู่ในระดับสูงเล็กน้อย อาจจะเนื่องมาจากธรรมชาติของวิชาชีพครูที่ต้องเผชิญกับปัญหาและความกดดันในหลากหลายด้าน เช่น การขาดแคลนทรัพยากรด้านการเรียนการสอน ความกดดันที่ต้องรับผิดชอบหลายบทบาท ปัญหาความสัมพันธ์ระหว่างบุคคล ความคาดหวังต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน (Ratanasiripong et al., 2020) กอรปกับช่วงเวลาที่ผู้วิจัยดำเนินการเก็บข้อมูลอยู่ในช่วงของการแพร่ระบาดของเชื้อไวรัสโคโรนา 2019 ส่งผลให้เกิดความไม่แน่นอนในชีวิตและการทำงานได้ ครูอาจจะเกิดสภาวะที่เรียกว่าการไม่อดทนต่อความไม่แน่นอน (intolerance of uncertainty) ซึ่งเป็นความเชื่อเชิงลบเกี่ยวกับความไม่แน่นอน

บุคคลจะประมวลข้อมูลที่คลุมเครือด้วยความลำเอียง มีการตีความสถานการณ์ที่ไม่แน่นอนว่าเป็นภัยคุกคาม (Jacoby, 2020) ความคิดและพฤติกรรมเหล่านี้ส่งผลให้บุคคลเกิดความกังวลใจ มีความวิตกกังวลและรับรู้ถึงภัยคุกคามเพิ่มมากขึ้น (Carleton et al., 2010) สำหรับความเหนื่อยหน่ายในการทำงานของครูด้านความอ่อนล้าทางอารมณ์ และด้านความเฉยชา ไม่ใส่ใจ มีค่าเฉลี่ยอยู่ในระดับต่ำ ในขณะที่ด้านประสิทธิภาพในการทำงานมีค่าเฉลี่ยอยู่ในระดับสูง สะท้อนให้เห็นว่า ครูสามารถจัดการอารมณ์และความรู้สึกทางลบได้ สามารถปฏิบัติงานได้อย่างมีประสิทธิภาพ อย่างไรก็ตาม แม้ว่าภาพรวมจะพบว่า ภาวะสุขภาพจิตของครูอยู่ในระดับปกติ แต่ไม่สามารถปฏิเสธได้ว่า ครูจำนวนหนึ่งมีความเครียด ความวิตกกังวล และภาวะซึมเศร้าอยู่ในระดับรุนแรงจนถึงรุนแรงมาก ดังที่ Ratanasiripong et al. (2021) ได้นำเสนอข้อมูลภาวะสุขภาพจิตและความเหนื่อยหน่ายของครูในโรงเรียนรัฐฯ 5 แห่ง พบว่า ครูส่วนใหญ่มีภาวะสุขภาพจิตในระดับปกติ แต่ครูที่มีความวิตกกังวล ความเครียด และภาวะซึมเศร้าระดับรุนแรงจนถึงรุนแรงมากมีถึงร้อยละ 11.2, 6.0 และ 3.2 ตามลำดับ ด้วยเหตุนี้ ความรู้และความเข้าใจเกี่ยวกับปัญหาสุขภาพจิตจึงเป็นประเด็นสำคัญที่ต้องนำมาพิจารณาในกระบวนการพัฒนาผู้ที่ประกอบวิชาชีพครู

2. ผลจากการวิเคราะห์การถดถอยพหุคูณ พบว่า ชีวิตการทำงานสามารถพยากรณ์ภาวะสุขภาพจิต ซึ่งมีรายละเอียดดังนี้ การควบคุมงานและภาระงานสามารถร่วมกันทำนายความเครียด การควบคุมงานและรางวัลมีอิทธิพลต่อภาวะซึมเศร้า และชุมชนในที่ทำงานและการควบคุมงานมีอิทธิพลต่อความวิตกกังวล ผลการวิจัยนี้สามารถอธิบายได้จากแนวคิดทางจิตวิทยาสังคม 2 แนวคิด โดยแนวคิดแรก คือ รูปแบบความเครียดจากงาน (job strain model) ที่พัฒนาโดย Karasek และ Theorell ซึ่งอธิบายถึงความสัมพันธ์ระหว่างปัญหาสุขภาพจิตกับลักษณะของงานที่ไม่พึงประสงค์ในหลายมิติ ได้แก่ ข้อเรียกร้องของงาน อำนาจในการควบคุมงาน และการสนับสนุนทางสังคม และแนวคิดที่สอง คือ รูปแบบความไม่สมดุลระหว่างความพยายามและรางวัล (effort-reward imbalance model) ที่พัฒนาโดย Siegrist และคณะ ที่อธิบายผลของความไม่สมดุลที่เกิดขึ้นจากการทำงานหนักแต่ได้รับการตอบแทนที่ต่ำทั้งในแง่ของเงิน โอกาสการทำงาน และความมั่นคงในงาน (Pikhart et al., 2004) กล่าวโดยสรุป ชีวิตการทำงานที่เกี่ยวข้องกับปริมาณงานที่ครูได้รับ อำนาจในการควบคุมงาน การได้รับรางวัลตอบแทน ความสัมพันธ์ในที่ทำงานตลอดจนการได้รับการสนับสนุนทางสังคม จึงเป็นปัจจัยสำคัญของสุขภาพจิตของครู ดังนั้น เพื่อป้องกันปัญหาสุขภาพจิตของครูในการทำงานจึงควรมุ่งเน้นในการสร้างความสอดคล้องและสมดุลระหว่างบุคคลกับงานโดยเฉพาะด้านภาระงาน การควบคุมงาน รางวัล และชุมชนในที่ทำงาน

สำหรับความเหนื่อยหน่ายในการทำงาน พบว่า การควบคุมงาน ความยุติธรรม ภาระงานและค่านิยมสามารถร่วมกันทำนายความอ่อนล้าทางอารมณ์ ชุมชนในที่ทำงานและค่านิยมมีอิทธิพลต่อความเฉยชา ไม่ใส่ใจ และชุมชนในที่ทำงานและการควบคุมงานมีอิทธิพลต่อประสิทธิภาพในการทำงาน ผลการวิจัยครั้งนี้สามารถอธิบายได้ตามรูปแบบความเหนื่อยหน่ายของครู (A working model of teacher burnout) นำเสนอโดย Maslach & Leither (1999) ที่อธิบายปัจจัยและผลกระทบของความเหนื่อยหน่ายในการทำงาน โดยปัจจัยที่ส่งผลต่อความเหนื่อยหน่ายมาจากหลายปัจจัย ได้แก่ สภาพแวดล้อมทางสังคม ลักษณะของโรงเรียน ธรรมชาติของงาน และคุณลักษณะส่วนบุคคลของครู นอกจากนี้ องค์ประกอบของความเหนื่อยหน่ายทั้ง 3 ด้านมีลักษณะ

เป็นคู่ขนาน (parallel) ไม่ได้เกิดขึ้นเป็นลำดับ (sequence) ดังนั้น ปัจจัยจากสภาพแวดล้อมของการทำงาน จึงส่งผลต่อความเหนื่อยหน่ายในแต่ละด้านแตกต่างกัน ซึ่ง Leiter & Maslach (2011) ได้อธิบายความสัมพันธ์ระหว่างปัจจัยด้านชีวิตการทำงานแต่ละด้านกับความเหนื่อยหน่ายแต่ละด้าน ดังนี้ ภาระงานที่มากเกินไป บุคคลจะบริหารจัดการได้เป็นปัจจัยสำคัญของความอ่อนล้าทางอารมณ์ และปัจจัยด้านการควบคุมงานที่เกิดจากความไม่ชัดเจนหรือความสับสนในบทบาทหน้าที่จากความขัดแย้งระหว่างบุคคลกับผู้มีอำนาจ และความไม่สอดคล้องระหว่างค่านิยมที่บุคคลไม่สามารถควบคุมได้ ซึ่งปัจจัยเหล่านี้มีความสัมพันธ์สูงกับความอ่อนล้า นอกจากนี้ ในด้านของประสิทธิภาพของการทำงาน หากบุคคลขาดการควบคุมงาน หรือไม่รับรู้ถึงความเป็นอิสระในการตัดสินใจหรือแก้ปัญหาในการทำงาน บุคคลจะทำงานอย่างมีประสิทธิภาพได้ยาก การสนับสนุนจากเพื่อนร่วมงานซึ่งเป็นชีวิตการทำงานด้านชุมชนในที่ทำงาน มีความสัมพันธ์อย่างใกล้ชิดกับ การทำงานที่มีประสิทธิภาพ

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 ผลการวิจัยครั้งนี้สามารถนำไปใช้เป็นสารสนเทศสำหรับผู้บริหารสถานศึกษาตลอดจนองค์กรที่มีส่วนเกี่ยวข้องในการส่งเสริมและพัฒนาครูเพื่อออกแบบและกำหนดนโยบายและแผนการบริหารจัดการ ให้สภาพแวดล้อมการทำงานมีความสอดคล้องระหว่างความต้องการขององค์กรและความต้องการหรือการรับรู้ของบุคคล เพื่อป้องกันปัญหาทางสุขภาพจิต และเพื่อส่งเสริมให้ครูสามารถพัฒนาตนเองและการทำงานอย่างมีประสิทธิภาพ

1.2 ผลการวิจัยครั้งนี้สามารถนำไปใช้เป็นสารสนเทศสำหรับนักจิตวิทยาการศึกษา หรือสาขาที่เกี่ยวข้องในการเข้าใจถึงสาเหตุที่เป็นปัจจัยเชิงองค์กรที่มีผลต่อปัญหาสุขภาพจิตและความเหนื่อยหน่ายของครู เพื่อออกแบบโปรแกรมฝึกอบรมให้ครูตระหนักถึงเหตุของปัญหาและหาแนวทางในการสร้างสมดุลระหว่างชีวิตการทำงานแต่ละด้านที่มีอิทธิพลต่อปัญหาทางสุขภาพจิต

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 การวิจัยครั้งนี้เป็นการสำรวจครูในทุกระดับและทุกสังกัด จึงทำให้ข้อมูลที่ปรากฏเป็นภาพรวมของปัจจัยที่ส่งผลกระทบต่อสุขภาพจิตของครู หากมีการศึกษาปัจจัยด้านชีวิตการทำงานที่มีต่อสุขภาพจิตของครูในแต่ละระดับชั้นที่สอน (ระดับประถมศึกษา และระดับมัธยมศึกษา) หรือจำแนกตามสังกัด อาจจะสะท้อนปัญหาที่แตกต่างกันตามธรรมชาติของสภาพแวดล้อมการทำงานของครูแต่ละระดับชั้นและสังกัดของหน่วยงาน

2.2 การวิจัยครั้งนี้มีขีดจำกัดภายใต้ลิขสิทธิ์ของการใช้เครื่องมือวิจัย ดังนั้น การวิจัยครั้งต่อไปอาจจะใช้เครื่องมืออื่นที่มีคุณภาพใกล้เคียง หรือพัฒนาเครื่องมือเพื่อเก็บข้อมูลด้านชีวิตการทำงานและความเหนื่อยหน่ายในการทำงานของครู

กิตติกรรมประกาศ งานวิจัยครั้งนี้เป็นส่วนหนึ่งของงานวิจัยที่ได้รับทุนสนับสนุนการวิจัยงบประมาณรายได้ คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

References

- Baroud, T., EL Masri, J., Shammaa, H., Chanbour, H., & Salameh, P. (2021). The impact of Covid19 pandemic, financial wellness, and online teaching on Lebanese schoolteachers' mental health. *Clinical Case Reports and Clinical Study*, 5(4).
<https://doi.org/09.2021/1.108>
- Berquam, J. (2020). *The Relationship Between Job-Person Fit and Use of Emotional Labor in Teachers* [Doctoral dissertation]. Grand Canyon University.
- Carleton, R. N., Gosselin, P., & Asmundson, G. J. G. (2010). The intolerance of uncertainty index: Replication and extension with an English sample. *Psychological Assessment*, 22(2), 396-406. <https://doi.org/10.1037/a0019230>
- Depression Anxiety Stress Scales. (2018). *Overview of the DASS and its uses*.
<http://www2.psy.unsw.edu.au/dass//over.htm>
- Hassan, O., & Ibourk, A. (2021). Burnout, self-efficacy and job satisfaction among primary school teachers in morocco. *Social Sciences & Humanities Open*, 4(1), 100148.
- Jacoby, R. J. (2020). Intolerance of uncertainty. In *Clinical handbook of fear and anxiety: Maintenance processes and treatment mechanisms* (pp. 45-63). American Psychological Association.
- Jones-Rincon, A., & Howard, K. J. (2019). Anxiety in the workplace: A comprehensive occupational health evaluation of anxiety disorder in public school teachers. *Journal of Applied Biobehavioral Research*, 24(1), e12133.
- Kabito, G. G., & Wami, S. D. (2020). Perceived work-related stress and its associated factors among public secondary school teachers in Gondar city: a cross-sectional study from Ethiopia. *BMC research notes*, 13(1), 1-7.
- Leiter, M. P. (1999). Burnout among teachers as a crisis in psychological contracts. In *Understanding and preventing teacher burnout: A sourcebook of international research and practice* (pp. 202-210). Cambridge University Press.
- Leiter, M. P., & Maslach, C. (2004). Areas of work life: A structured approach to organizational predictors of job burnout. In *Research in occupational stress and well being: Vol. 3. Emotional and physiological processes and positive intervention strategies* (pp. 91-134). Elsevier
- Leiter, M. P. & Maslach, C. (2011), *Areas of work life survey manual* (5th Ed.). Mind Garden.
- Leiter, M. P., & Maslach, C. (2016). Latent burnout profiles: A new approach to understanding the burnout experience. *Burnout Research*, 3(4), 89-100.

- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety & stress scales* (2nd Ed.). Psychology Foundation.
- Mack, J. C., Johnson, A., Jones-Rincon, A., Tsatenawa, V., & Howard, K. (2019). Why do teachers leave? A comprehensive occupational health study evaluating intent-to-quit in public school teachers. *Journal of Applied Biobehavioral Research*, 24(1), e12160.
- Mahmoodi-Shahrehabaki, M. (2018). Teacher burnout. *The TESOL encyclopedia of English language teaching*, 1-8. <https://doi.org/10.1002/9781118784235.eelt0964>
- Maslach, C. (1999). Progress in understanding teacher burnout. In *Understanding and preventing teacher burnout: A sourcebook of international research and practice* (pp. 211 – 222). Cambridge University Press.
- Maslach, C. (2019). A message from the Maslach Burnout Inventory Authors. *Mind Garden*. <https://www.mindgarden.com/blog/post/44-a-message-from-the-maslach-burnout-inventory-authors>
- Maslach, C., Jackson, S. E., Leiter, M. P., Schaufeli, W. B., & Schwab, R. L. (2018). *Maslach burnout inventory manual* (4th Ed.). Mind Garden.
- Maslach, C., & Leiter, M. P. (1997). *The truth about burnout: How organizations cause personal stress and what to do about it*. John Wiley & Sons.
- Maslach, C., & Leiter, M. P. (1999). Teacher burnout: a research agenda. In *Understanding and preventing teacher burnout* (pp. 295-303). Cambridge University Press.
- Maslach, C., Leiter, M. P., & Schaufeli, W. (2009). Measuring burnout. In *The oxford handbook of organizational well-being* (pp. 86-108). Oxford University Press.
- Martínez-Monteagudo, M. C., Inglés, C. J., Granados, L., Aparisi, D., & García-Fernández, J. M. (2019). Trait emotional intelligence profiles, burnout, anxiety, depression, and stress in secondary education teachers. *Personality and Individual Differences*, 142, 53-61. <https://doi.org/10.1016/j.paid.2019.01.036>
- Mojša-Kaja, J., Golonka, K., & Marek, T. (2015). Job burnout and engagement among teachers – work life areas and personality traits as predictors of relationships with work. *International Journal of Occupational Medicine and Environmental Health*, 28(1), 102-119.
- Morling, B. (2018). *Research methods in psychology: Evaluating a world of information* (3rd Ed.). W. W. Norton & Company.

- Pikhart, H., Bobak, M., Pajak, A., Malyutina, S., Kubinova, R., Topor, R., Sebakova, H., Nikitin, Y., & Marmot, M. (2004). Psychosocial factors at work and depression in three countries of Central and Eastern Europe. *Social Science & Medicine*, 58(8), 1475-1482.
- Ratanasiripong, P., China, T., Ratanasiripong, N. T., & Toyama, S. (2020). Resiliency and mental health of school teachers in Okinawa. *Journal of Health Research*.
<https://doi.org/10.1108/JHR-11-2019-0248>
- Ratanasiripong, P., Ratanasiripong, N. T., Nungdanjark, W., Thongthammarat, Y., & Toyama, S. (2021). Mental health and burnout among teachers in Thailand. *Journal of Health Research*. <https://doi.org/10.1108/JHR-05-2020-0181>
- World Health Organization. (2019). *Burn-out an "occupational phenomenon": International Classification of Diseases*. <https://www.who.int/news/item/28-05-2019-burn-out-an-occupational-phenomenon-international-classification-of-diseases>
- Yilmaz, K., Altinkurt, Y., Guner, M., & Sen, B. (2015). The relationship between teachers' emotional labor and burnout level. *Eurasian Journal of Educational Research*, 59, 75-90. <http://dx.doi.org/10.14689/ejer.2015.59.5>

Translated Thai References

- Phusee-orn, S. (2015). *Applying SPSS for analyzing research data* (7th Ed.). Taksila publishing. (in Thai)
- Piyaphotchanakorn, S., Kaewkungwal, S., Sompakdee, S., & Yimsook, N. (2017). Self-efficacy, job demands-job resources and burnout with perception of organizational justice as a moderating variable. *Journal of Social Work*, 25(1), 91-118. (in Thai)
- Ploytabtim, K., Buathong, N., Prasartpornsirichoke, J., & Chandarasiri, P. (2021). Prevalence and associated factors of anxiety and depression among teachers of one private all-girls boarding school. *Journal of the Psychiatric Association of Thailand*, 66(1), 25-38. (in Thai)
- Tongying, C. (2016). *Teacher Problems: Problems to be reformed*.
<https://dl.parliament.go.th/backoffice/viewer2300/web/previewer.php> (in Thai)

การพัฒนาแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 Development of Technological Innovation Literacy Scale for Grade 12 Students

เอกพัฒน์ เอื้องใส่อง^{1*} ชำนาญ ปาณาวงษ์² นันทิพย์ องอาจวานิชย์²
และ นันทิมา นาคพงศ์ อัสวรักษ์²

Ekkapad Huengsaiong^{1*}, Chamnan Panawong², Namthip Ongardwanich²
and Nanthima Nakaphong Asvaraksha²

(Received: March 1, 2022; Revised: April 3, 2022; Accepted: April 4, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายดังนี้ 1) เพื่อสร้างแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 และ 2) เพื่อตรวจสอบความตรงเชิงโครงสร้างของแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ระยะที่ 1 แหล่งข้อมูลในการสร้างแบบวัดคือผู้ทรงคุณวุฒิจำนวนทั้งหมด 18 คน ที่ได้จากการเลือกแบบเจาะจง และนักเรียนชั้นมัธยมศึกษาปีที่ 6 โรงเรียนวิทยาศาสตร์จุฬารณราชมหาวิทยาลัยเลยจำนวน 30 คน ระยะที่ 2 กลุ่มตัวอย่างที่ใช้ในการตรวจสอบความตรงเชิงโครงสร้างคือนักเรียนชั้นมัธยมศึกษาปีที่ 6 โรงเรียนในกลุ่มโรงเรียนวิทยาศาสตร์จุฬารณราชมหาวิทยาลัยจำนวน 254 คน ที่ได้จากการสุ่มแบบหลายขั้นตอน และวิเคราะห์ความตรงเชิงโครงสร้างด้วยเทคนิควิธีการวิเคราะห์องค์ประกอบเชิงยืนยัน

ผลการวิจัยพบว่า

1. แบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาชั้นปีที่ 6 ที่สร้างขึ้นประกอบด้วยข้อคำถามแบบปรนัย 4 ตัวเลือกจำนวน 18 ข้อ ข้อคำถามเชิงสถานการณ์จำนวน 4 ข้อ และข้อสอบอัตนัยจำนวน 1 ข้อ รวมข้อคำถามในแบบวัดทั้งหมดจำนวน 23 ข้อ ข้อคำถามทุกข้อมีความตรงเชิงเนื้อหา (IOC = 0.60 - 1.00) ค่าความยากง่ายมีค่าระหว่างค่อนข้างง่ายถึงค่อนข้างยาก (0.23 - 0.67) ค่าอำนาจจำแนกมีค่าอยู่ระหว่างจำแนกได้พอใช้ถึงจำแนกได้ดีมาก (0.29 - 0.91) และความเที่ยงทั้งฉบับมีค่าค่อนข้างสูง ($\alpha = 0.71$)

2. แบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาชั้นปีที่ 6 ที่สร้างขึ้นมีความตรงเชิงโครงสร้าง ($\chi^2 = 141.71$, $df = 164$, $p = 0.89515$ GFI = 0.96, AGFI = 0.92, RMSEA = 0.00,

¹ นิสิตปริญญาเอก สาขาวิชาการวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร

² ผู้ช่วยศาสตราจารย์ สาขาวิชาการวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร

¹ Doctoral degree student, Educational Research and Evaluation, Faculty of Education, Naresuan University

² Assistant Professor, Educational Research and Evaluation, Faculty of Education, Naresuan University

* Corresponding Author E-mail: Ekkapad.h@gmail.com

RMR = 0.094, NFI = 0.99, NNFI = 1.00 and CFI = 1.00) โดยประกอบด้วย 2 องค์ประกอบหลัก ได้แก่ 1) ความรู้เกี่ยวกับนวัตกรรมด้านเทคโนโลยี (knowledge of technology innovation) ประกอบด้วย 3 ตัวบ่งชี้ ได้แก่ 1.1) ความรู้เนื้อหา นวัตกรรมด้านเทคโนโลยี (knowledge of technology innovation content) วัดได้จากข้อคำถามที่ 1 ถึง 4 1.2) ความรู้กระบวนการ นวัตกรรมด้านเทคโนโลยี (knowledge of technology innovation process) วัดได้จากข้อคำถามที่ 5 ถึง 8 และ 1.3) ความรู้รวบยอด นวัตกรรมด้านเทคโนโลยี (knowledge of technology innovation concept) วัดได้จากข้อคำถามที่ 9 ถึง 12 2) และคุณลักษณะ นวัตกรรม (innovator attribute) ประกอบด้วย 3 ตัวบ่งชี้ ได้แก่ 2.1) การจัดการและสร้างความรู้ (management and knowledge creation) วัดได้จากข้อคำถามที่ 13 ถึง 16 2.2) ความคิดสร้างสรรค์ (creative thinking) วัดได้จากข้อคำถามที่ 23 และ 2.3) ภาวะผู้นำ (leadership) วัดได้จากข้อคำถามที่ 17 ถึง 22

คำสำคัญ: การรู้เรื่องนวัตกรรมด้านเทคโนโลยี แบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยี
การพัฒนาแบบทดสอบและแบบวัด

Abstract

The purposes of this research were 1) to create a technological innovation literacy scale for grade 12 students, and 2) to validate the construct validity of the technological innovation literacy scale for grade 12 students. In Phase 1, 18 experts and 30 grade 12 students from Princess Chulabhorn Science High School Loei, selected through purposive sampling, were the sources of data for the scale construction. In Phase 2, 254 grade 12 students from the schools in the Princess Chulabhorn Science High School Group, selected through multistage random sampling, validated the construct validity of the scale. The analysis of the construct validity of the scale employed confirmatory factor analysis.

The results were as follows:

1. The technological innovation literacy scale for grade 12 students that had been constructed consisted of 18 4-choice items, 4 situational questions and 1 subjective question, totaling 23 items. The content validity of the items ranged from 0.60 to 1.00, the difficulty ranged from rather easy to rather difficult (0.23 to 0.67), the discrimination ranged from fairly discriminating to very discriminating (0.29 to 0.91), and the total reliability was rather high ($\alpha = 0.71$).

2. The technological innovation literacy scale for grade 12 students that had been constructed had the construct validity ($\chi^2 = 141.71$, $df = 164$, $p = 0.89515$ GFI = 0.96, AGFI = 0.92, RMSEA = 0.00, RMR = 0.094, NFI = 0.99, NNFI = 1.00 and CFI = 1.00). It had 2 main components: 1) knowledge of technological innovation which consisted of 3 indicators: 1.1) knowledge of technological innovation content, measured with items 1 to 4; 1.2) knowledge

of technological innovation process, measured with items 5 to 8; and 1.3) knowledge of technological innovation concept, measured with items 9 to 12; 2) the innovator attribute consisted of 3 indicators: 2.1) management and knowledge creation, measured with items 13 to 16; 2.2) creative thinking, measured with item 23; and 2.3) leadership, measured with items 17 to 22.

Keywords: technological innovation literacy, technological innovation literacy scale, development of tests and scales

บทนำ

การเปลี่ยนแปลงโครงสร้างของเศรษฐกิจและสังคมโลกในทุกยุคทุกสมัยที่เป็นไปอย่างก้าวกระโดด จวบจนถึงปัจจุบันที่กำลังจะก้าวเข้าสู่ศตวรรษที่ 21 ที่ทุกประเทศต่างมีความกระตือรือร้นและแข่งขันกันอย่างสูงในทุกๆ ด้าน สืบเนื่องมาจากความเจริญก้าวหน้า ด้านวิทยาศาสตร์ เทคโนโลยี และนวัตกรรม โดยเฉพาะอย่างยิ่ง นวัตกรรมด้านเทคโนโลยี (technology innovation) ควบคู่กับสภาวะวิกฤตทางเศรษฐกิจและสังคม โดยเฉพาะอย่างยิ่งปัญหาความขัดแย้งทางการเมืองภายในประเทศมีอิทธิพลทั้งทางตรงและทางอ้อมต่อการชะลอตัวของ การเจริญเติบโตของประเทศไทยอย่างชัดเจนจากการติดอยู่ในกลุ่มประเทศที่มีรายได้ระดับปานกลาง (middle income countries) เป็นระยะเวลายาวนานกว่า 40 ปี เนื่องจากการขาดความสามารถในการเปลี่ยนแปลง ทรัพยากรที่มีให้เป็นนวัตกรรมดังจะเห็นภาพเชิงประจักษ์ได้จากสินค้าส่งออกสำคัญที่สร้างรายได้ให้แก่ประเทศ ที่เป็นผลผลิตด้านการเกษตร ในทางกลับกันสินค้านำเข้าสำคัญที่เป็นรายจ่ายของประเทศเป็นวัสดุอุปกรณ์ หรือ ผลิตภัณฑ์ที่เป็นเทคโนโลยีและนวัตกรรม (Information and Communication Technology Center, 2017) จากผลการประเมินสภาพแวดล้อมในการพัฒนาประเทศ พบว่า ประเทศไทยเป็นประเทศที่อยู่ในกลุ่มประเทศ ที่มีโครงสร้างพื้นฐานด้านวิทยาศาสตร์ เทคโนโลยี และนวัตกรรมในระดับต่ำโดยสาเหตุของปัญหาดังกล่าว เกิดจากการขาดการพัฒนาบุคลากรด้านวิทยาศาสตร์ เทคโนโลยี วิจัย และนวัตกรรมที่ยังไม่มีคุณภาพเทียบเท่าสากลและไม่เพียงพอต่อความต้องการของตลาดแรงงาน (Ministry of industry Thailand, 2016; Cornell University, INSEAD & the World Intellectual Property Organization, 2019) แสดงให้เห็นว่าประเทศไทย ควรที่จะเร่งพัฒนาบุคลากรภายในชาติให้มีความรู้ความสามารถในการสร้างหรือพัฒนาเทคโนโลยีและ นวัตกรรมเพื่อนำมาซึ่งบุคลากรด้านวิทยาศาสตร์ วิจัย และนวัตกรรมในชาติที่มีคุณภาพ เทียบเท่าสากลและเพียงพอต่อความต้องการของตลาดแรงงานในอนาคต เพื่อให้เป็นไปตามที่ระบุไว้ในรัฐธรรมนูญแห่ง ราชอาณาจักรไทย พุทธศักราช 2560 ว่ามีวัตถุประสงค์มุ่งหวังที่จะเสริมสร้างศักยภาพบุคลากรในชาติ ให้มีความรู้ความสามารถด้านวิทยาศาสตร์ เทคโนโลยี วิจัย และนวัตกรรมให้มีคุณภาพและมีประสิทธิภาพ เพียงพอต่อความต้องการของตลาดแรงงานในอนาคตเพื่อเตรียมความพร้อมในการเข้าสู่สังคมโลกในศตวรรษที่ 21 (The Secretariat of the Cabinet, 2019)

“การรู้เรื่องนวัตกรรมด้านเทคโนโลยี” หรือ “technology innovation literacy” เป็นคำใหม่ ที่เกิดขึ้นจากการที่ผู้วิจัยรวมคำ 2 คำ เข้าด้วยกัน ซึ่งประกอบด้วยคำว่า “innovation literacy” ตัวแปรใหม่

ที่ผู้วิจัยสนใจศึกษาและให้คำนิยามเป็นภาษาไทยว่า “การรู้เรื่องนวัตกรรม” ถูกกล่าวถึงครั้งแรกในประเทศสหรัฐอเมริกาโดย Michael J. Gelb และ Sarah Miller Caldicott ในปี ค.ศ. 2007 ในหนังสือชื่อ “Innovative Like Edison: The Five - Step System for Breakthrough Business Success” (Gelb & Caldicott, 2007) โดย Niyaza Endogan และ Mehmet Sencer Corlu ในการศึกษาเรื่อง “Investigating an innovative course to improve innovation literacy” และ Niyaza Endogan Mehmet Sencer Corlu ในปี ค.ศ. 20011 และ Niyaza Endogan Mehmet Sencer Corlu และ Robert M. Capraro ในปี ค.ศ. 2013 ในการศึกษาเรื่อง “Definition Innovation Literacy: Do Robotics Program Help Students Develop Innovation Literacy skill?” (Erdogan et al, 2013) ในประเทศแคนาดา โดย George Brown College ในงานเปิดตัวหลักสูตรการเรียนการสอนสาขาวิชาการวิจัยเชิงพานิชย์และนวัตกรรมของวิทยาลัย (Luke, 2008) ในบทความออนไลน์ และ Martin & Florida (2009) ในการศึกษาโครงการ “Creative Age” (โครงการที่มุ่งเน้นส่งเสริมให้ใช้ความคิดสร้างสรรค์ในการขับเคลื่อนเศรษฐกิจในรัฐออนแทรีโอ) (Martin & Florida, 2009) ในประเทศญี่ปุ่นโดย Thoshihiko Yamakami ในปี ค.ศ. 2012 ในบทความเรื่อง “Innovation Literacy: Implication from a shift toward dynamic multidisciplinary engineering” (Yamakami, 2012) ในประเทศฝรั่งเศสโดย Organization for Economic Co-operation and Development: OECD หรือองค์การเพื่อความร่วมมือและการพัฒนาทางเศรษฐกิจ ในปี ค.ศ. 2015 ที่กล่าวว่าเป็นความรู้และทักษะที่จำเป็น และประเทศสหรัฐอเมริกาหับเอมิเรตส์ โดย Myint Swe Khine ในปี ค.ศ. 2017 ในหนังสือ “Robotics in STEM Education: Redesigning the Learning Experiences” ตามลำดับ (Khine, 2017) และคำว่า “technology” หรือ “เทคโนโลยี” ที่ตามแนวคิดของกลุ่มนักวิทยาศาสตร์กายภาพ (physic science) ในชาติตะวันตกเชื่อกันว่าเป็นผลผลิตจากความรู้ทางวิทยาศาสตร์และวิศวกรรม (Sukhothai Thammathirat open university, 2012) ยกตัวอย่างเช่น หม้อหุงข้าวไฟฟ้า รถยนต์ไฟฟ้า และเคมีภัณฑ์ต่าง ๆ เป็นต้น กล่าวคือ เชื่อว่าเทคโนโลยีจะต้องอาศัยกระบวนการหรือขั้นตอนที่เป็นระบบในการสร้าง/การผลิต โดยเนื้อหาใจความที่พบทั้งหมดยังอยู่ในขั้นเริ่มต้นคือเป็นเพียงการกล่าวถึง การชี้ให้เห็นความสำคัญ การพยายามให้คำนิยาม และภาพรวมในแง่ของแนวคิดและความหมายที่เกี่ยวข้องตามมุมมองความเชื่อและวิธีการที่แต่ละบุคคลสนใจและศึกษาเท่านั้นสำหรับการรู้เรื่องนวัตกรรม ส่วนการรู้เรื่องนวัตกรรมด้านเทคโนโลยีนั้นทั้งในไทยและต่างประเทศยังไม่มีมีการทำการศึกษาแต่อย่างใด ยังคงคลุมเครือ (ambiguous) เป็นนามธรรม (abstract) ไม่สามารถวัดหรือสังเกตได้โดยตรง และยังไม่มีเครื่องมือที่จะใช้ในการวัด ดังนั้นจึงควรอย่างยิ่งที่จะต้องทำให้นิยามความหมายให้ชัดเจนและโดยเฉพาะอย่างยิ่งการสร้างเครื่องมือสำหรับวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีอันจะนำมาซึ่งสารสนเทศในการพัฒนา (Medin & Ross, 1992) เนื่องจากแบบวัด (scales) เป็นเครื่องมือในการทดสอบ (testing) อย่างหนึ่งในการวัดหรือการทดสอบความรู้ ความสามารถหรือทักษะ หรือพฤติกรรมของมนุษย์ อันจะนำมาซึ่งสารสนเทศที่สำคัญและเป็นประโยชน์ในการบรรยาย การอธิบาย การทำนาย และการควบคุม เพื่อการพัฒนาสิ่งที่ต้องการวัดหรือพฤติกรรมที่พึงปรารถนานั้นๆ ในลำดับต่อไป (Angoff, 1984; Kerlinger, 1986; Gregory, 1992; Gipps, 1994)

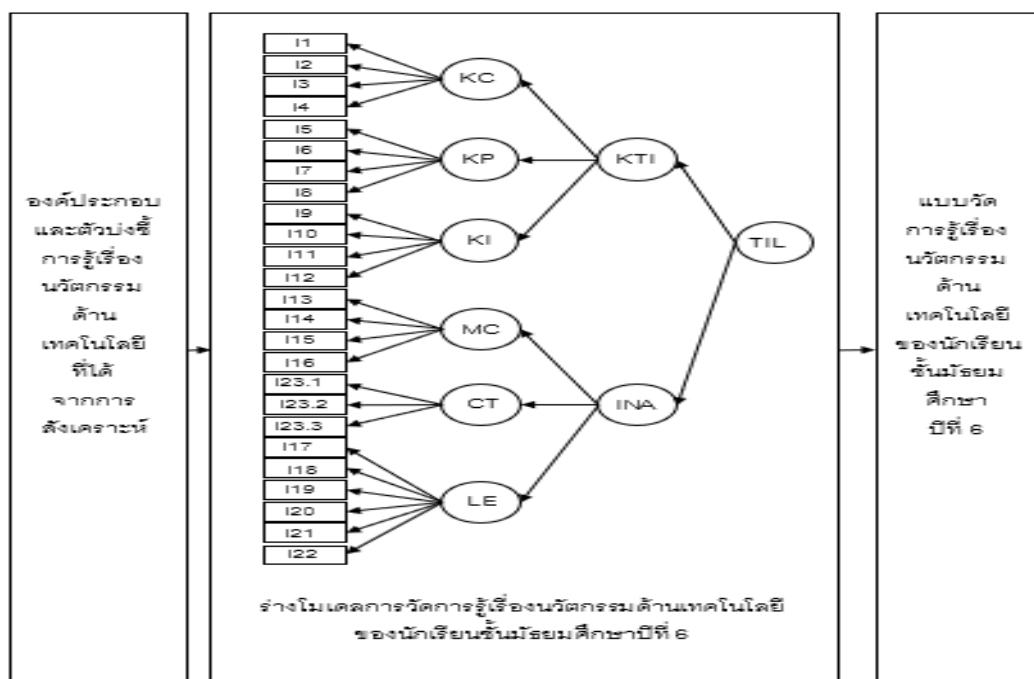
ด้วยเหตุนี้เพื่อเป็นส่วนหนึ่งในการพัฒนาบุคลากรในชาติให้มีความรู้ความสามารถในการสร้างหรือพัฒนาเทคโนโลยีและนวัตกรรมอันจะนำมาซึ่งบุคลากรด้านวิทยาศาสตร์ วิจัย และนวัตกรรมในชาติที่มีคุณภาพเทียบเท่าสากลและเพียงพอต่อความต้องการของตลาดแรงงานในอนาคต ผู้วิจัยจึงสนใจศึกษาค้นคว้าเรื่องการพัฒนาแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ซึ่งอยู่ในช่วงวัยที่มีความพร้อมทั้งร่างกาย สติปัญญา และอารมณ์ความรู้สึกในการมองโลกในบริบทชีวิตจริง (Weiner, 2003; Kalat, 2017) กล่าวคือ เป็นเยาวชนที่กำลังจะเป็นบุคลากรที่สำคัญด้านวิทยาศาสตร์ เทคโนโลยี วิจัย และนวัตกรรมของชาติในอนาคต โดยผลการวิจัยในครั้งนี้จะนำมาซึ่งเครื่องมือวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ที่มีคุณภาพ และเป็นสารสนเทศพื้นฐานสำหรับผู้ที่มีส่วนเกี่ยวข้องและผู้สนใจศึกษาค้นคว้าเพื่อการวางแผนปรับปรุงและพัฒนาการจัดการศึกษาอันจะเป็นประโยชน์และนำมาซึ่งบุคลากรด้านวิทยาศาสตร์ เทคโนโลยี วิจัย และนวัตกรรมสืบไป

ความมุ่งหมายของการวิจัย

1. เพื่อสร้างแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6
2. เพื่อตรวจสอบความตรงเชิงโครงสร้างของแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6

กรอบแนวคิดในการวิจัย

จากการศึกษาแนวคิด ทฤษฎี และเอกสารงานวิจัยที่เกี่ยวข้องกับการรู้เรื่องนวัตกรรมด้านเทคโนโลยี (technology innovation literacy) สามารถสรุปเป็นกรอบแนวคิดในการวิจัยได้ดังภาพ 1



ภาพ 1 กรอบแนวคิดในการวิจัย

วิธีดำเนินการวิจัย

ระยะที่ 1 สร้างแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีรายละเอียดในการดำเนินการดังนี้

1.1 สังเคราะห์องค์ประกอบและตัวบ่งชี้ของการรู้เรื่องนวัตกรรมด้านเทคโนโลยี มีรายละเอียดในการดำเนินการดังนี้

1.1.1 แหล่งข้อมูล คือ เอกสารและงานวิจัยที่มีความเกี่ยวข้องกับแนวคิดการรู้เรื่องนวัตกรรมด้านเทคโนโลยี (technology innovation literacy)

1.1.2 เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือ ตารางสังเคราะห์องค์ประกอบและตัวบ่งชี้ของการรู้เรื่องนวัตกรรมด้านเทคโนโลยี (technology innovation literacy)

1.1.3 ขั้นตอนในการดำเนินการ มีรายละเอียดขั้นตอนในการดำเนินการที่สำคัญอันได้แก่ 1) ศึกษาค้นคว้าเอกสารแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง 2) สังเคราะห์องค์ประกอบและตัวบ่งชี้ 3) ตรวจสอบความตรงและความครอบคลุมขององค์ประกอบและตัวบ่งชี้ และ 4) ตรวจสอบความตรงและความครอบคลุมขององค์ประกอบและตัวบ่งชี้โดยผู้ทรงคุณวุฒิ

1.2 สร้างแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีรายละเอียดในการดำเนินการดังนี้

1.2.1 แหล่งข้อมูล คือ องค์ประกอบและตัวบ่งชี้ที่ได้จากการสังเคราะห์ในข้อ 1

1.2.2 เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือ ข้อคำถามที่ได้จากการสร้างตามกรอบในการสร้างเครื่องมือ

1.2.3 ขั้นตอนในการดำเนินการ มีรายละเอียดขั้นตอนในการดำเนินการที่สำคัญอันได้แก่ 1) ศึกษาค้นคว้าเอกสารแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องกับตัวบ่งชี้ที่ได้จากผลการศึกษาในข้อ 1 2) กำหนดกรอบนิยามพฤติกรรมและจำนวนข้อคำถามที่จะใช้ในการสร้างเครื่องมือ 3) สร้างข้อคำถาม 4) ตรวจสอบความตรงและความครอบคลุมของเนื้อหาข้อคำถามที่สร้างขึ้น โดยคณะกรรมการที่ปรึกษาวิทยานิพนธ์ และปรับแก้ไขให้สมบูรณ์ยิ่งขึ้น และ 5) ตรวจสอบความตรงและความครอบคลุมของเนื้อหาข้อคำถามโดยผู้ทรงคุณวุฒิ 6) นำไปทดลองใช้ 7) วิเคราะห์ค่าความยากง่าย วิเคราะห์ค่าอำนาจจำแนก และวิเคราะห์ความเที่ยง และ 8) คัดเลือกข้อที่ผ่านเกณฑ์ และจัดพิมพ์แบบวัดฉบับสมบูรณ์

1.3 ดำเนินการเก็บรวบรวม

1.4 การวิเคราะห์ข้อมูล ใช้โปรแกรม TAP และ SPSS

1.5 สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าเฉลี่ย ($\bar{X}$) ส่วนเบี่ยงเบนมาตรฐาน (SD) ค่าความตรงเชิงเนื้อหาโดยใช้ดัชนีความสอดคล้อง (IOC) ค่าความยากง่าย (P) ค่าอำนาจจำแนก (r) และค่าความเที่ยงใช้สูตรสัมประสิทธิ์แอลฟาของครอนบาค (α)

ระยะที่ 2 ตรวจสอบความตรงเชิงโครงสร้างของแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยี ของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีรายละเอียดในการดำเนินการดังนี้

2.1 แหล่งข้อมูล

2.1.1 ประชากร คือ นักเรียนชั้นมัธยมศึกษาปีที่ 6 โรงเรียนในสังกัดสำนักงานการศึกษา
ขั้นพื้นฐาน

2.1.2 กลุ่มตัวอย่าง คือ นักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปีที่ 6 ภาคเรียนที่ 1
ประจำปีการศึกษา 2564 โรงเรียนในกลุ่มโรงเรียนวิทยาศาสตร์จุฬาราชวิทยาลัยเพราะเป็นกลุ่มโรงเรียน
ที่จัดตั้งขึ้นเพื่อมุ่งส่งเสริมนักเรียนที่มีความรู้ความสามารถพิเศษเฉพาะด้านวิทยาศาสตร์อันเป็นพื้นฐานที่สำคัญ
ในการผลิตเทคโนโลยี โดยกำหนดจำนวนขนาดกลุ่มตัวอย่างตามกฎแห่งความชัดเจน (rule of thumb)
กำหนดให้มีขนาดกลุ่มตัวอย่าง 10 - 20 เท่า ต่อ 1 ตัวแปร หรือขั้นต่ำสุดจำนวน 100 ตัวอย่าง (Schumacker
& Lomax, 2016) ซึ่งการวิจัยในครั้งนี้มีจำนวนตัวแปรสังเกตได้ทั้งหมด 25 ตัวแปร ดังนั้นกลุ่มตัวอย่างขั้นต่ำ
อย่างน้อยควรมีจำนวน 100 หรือ 250 - 500 คน เพื่อให้โมเดลมีความคงทน (robustness) โดยในการวิจัยครั้งนี้
ผู้วิจัยใช้กลุ่มตัวอย่างทั้งหมด จำนวน 254 คน จากการสุ่มแบบหลายขั้นตอน (multi - stage random
sampling) โดยมีรายละเอียดขั้นตอนในการสุ่มอันได้แก่ ขั้นที่ 1 ใช้ภูมิภาคทั้งหมด 6 ภูมิภาคตามหลัก
วัตถุประสงค์ทางภูมิศาสตร์และวิทยาศาสตร์เป็นหน่วยในการสุ่ม โดยทำการสุ่มอย่างง่ายได้ภูมิภาค จำนวน 4
ภูมิภาค คิดเป็นร้อยละ 66 (>60%) (Angoff, 1984) ขั้นที่ 2 ใช้โรงเรียนเป็นหน่วยในการสุ่ม โดยทำการสุ่ม
อย่างง่ายมาภูมิภาคละ 1 โรงเรียน ได้โรงเรียนทั้งหมดจำนวน 4 โรงเรียน ขั้นที่ 3 ใช้ห้องเรียนเป็นหน่วยในการสุ่ม
โดยสุ่มอย่างง่ายมาโรงเรียนละ 4 ห้องเรียน ได้ห้องเรียนทั้งหมดจำนวน 16 ห้องเรียน และขั้นที่ 4 ใช้นักเรียน
เป็นหน่วยในการสุ่ม โดยใช้นักเรียนทั้งหมดในห้องเรียนเป็นกลุ่มตัวอย่าง

2.2 เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือ แบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยี
ของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่ได้จากการศึกษาในระยะที่ 1

2.3 ดำเนินการเก็บรวบรวมข้อมูล

2.4 การวิเคราะห์ข้อมูล ใช้โปรแกรม SPSS และ LISREL for Student

2.5 สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าสถิติทดสอบไค-สแควร์ (χ^2) ค่าดัชนีวัดระดับ
ความสอดคล้อง (*GFI*) ดัชนีวัดความสอดคล้องที่ปรับแก้แล้ว (*AGFI*) ดัชนีรากกำลังสองเฉลี่ยของ
ความแตกต่างโดยประมาณ (*RMSEA*) ดัชนีรากของกำลังสองเฉลี่ยของเศษเหลือ (*RMR*) ดัชนีเอ็นเอฟไอ
(*NFI*) ดัชนีเอ็นเอ็นเอฟไอ (*NNFI*) และดัชนีซีเอฟไอ (*CFI*)

ผลการวิจัย

1. ผลการสร้างแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6

1.1 การตรวจสอบความตรงขององค์ประกอบและตัวบ่งชี้โดยผู้ทรงคุณวุฒิจำนวน 14 ท่าน ได้แก่

1) ผู้ทรงคุณวุฒิด้านวิทยาศาสตร์ 2) ผู้ทรงคุณวุฒิด้านเทคโนโลยี และธุรกิจและการสื่อสาร 3) ผู้ทรงคุณวุฒิ
ด้านการสอนวิชาเทคโนโลยีและนวัตกรรม 4) ผู้ทรงคุณวุฒิด้านการวัดและประเมินผลการศึกษา และ

5) ผู้ทรงคุณวุฒิด้านการสอนวิทยาศาสตร์ระดับมัธยมศึกษา โดยผลจากการตรวจสอบ พบว่าองค์ประกอบและตัวบ่งชี้ที่สร้างขึ้นทั้งหมดผ่านเกณฑ์ที่กำหนด โดยมีค่าอยู่ระหว่าง 0.71 ถึง 0.93

1.2 แบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่ได้จากการวิจัยประกอบด้วยข้อคำถามทั้งหมดจำนวน 23 ข้อ แบ่งเป็นข้อสอบปรนัยจำนวน 18 ข้อ ข้อสอบเชิงสถานการณ์จำนวน 4 ข้อ และข้อสอบอัตนัยจำนวน 1 ข้อ เมื่อทำการตรวจสอบคุณภาพ พบว่า ข้อคำถามทั้งหมดผ่านเกณฑ์คุณภาพโดยมีรายละเอียดผลการวิเคราะห์สรุปได้ดังตาราง 1

ตาราง 1 ผลการตรวจสอบคุณภาพของข้อคำถามที่ใช้ในแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6

ข้อ	IOC	p	r	ข้อ	IOC	p	r
1	0.80	0.43	0.78	14	1.00	0.27	0.78
2	1.00	0.67	0.46	15	1.00	0.40	0.29
3	1.00	0.40	1.00	16	1.00	0.43	0.33
4	0.80	0.30	0.72	23.1	1.00	0.65	0.56
5	0.60	0.27	0.54	23.2	1.00	0.61	0.56
6	0.80	0.27	0.63	23.3	1.00	0.62	0.56
7	1.00	0.23	0.33	17	1.00	คำถามเชิงสถานการณ์	
8	1.00	0.23	0.36	18	0.80	คำถามเชิงสถานการณ์	
9	1.00	0.27	0.67	19	1.00	0.23	0.40
10	1.00	0.27	0.50	20	0.80	0.43	0.91
11	1.00	0.43	0.90	21	1.00	คำถามเชิงสถานการณ์	
12	1.00	0.27	0.72	22	0.60	คำถามเชิงสถานการณ์	
13	1.00	0.27	0.73	α = 0.71			

หมายเหตุ ค่า IOC ในตารางได้จากการตรวจสอบโดยผู้ทรงคุณวุฒิจำนวน 5 ท่าน ได้แก่ 1) ผู้ทรงคุณวุฒิด้าน

การสอนวิทยาศาสตร์ระดับมัธยมศึกษา (สายสามัญ) และ 2) ผู้ทรงคุณวุฒิด้านการสอนเทคโนโลยี (สายวิชาชีพ)

กล่าวโดยสรุป คำถามทั้ง 23 ข้อ ที่ใช้ในแบบแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 เป็นข้อคำถามที่มีความถูกต้องและเหมาะสมในการวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6

2. การตรวจสอบความตรงเชิงโครงสร้างแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6

2.1 ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรสังเกตได้ที่ใช้ในการวิเคราะห์ความตรงเชิงโครงสร้างแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 พบว่า ค่าความสัมพันธ์ของตัวแปรสังเกตได้มีค่าอยู่ระหว่าง 0.11 ถึง 0.94 และเป็นความสัมพันธ์ทางบวกทั้งหมด เมื่อพิจารณาผลการคำนวณค่าดัชนี Kaiser - Meryer - Olkin (KMO) พบว่า มีค่าเท่ากับ 0.909 และจากผลการคำนวณค่าสถิติ Bartlett's test of Sphericity พบว่า มีค่าเท่ากับ 7265.272 ($p < 0.01$) แสดงให้เห็นว่าจำนวนกลุ่มตัวอย่างที่ใช้

มีความเหมาะสมในการวิเคราะห์องค์ประกอบของข้อมูลชุดนี้ และตัวแปรแต่ละตัวมีความสัมพันธ์กันและความเหมาะสมที่จะนำไปใช้ในการพัฒนาและตรวจสอบความตรงเชิงโครงสร้างของโมเดลต่อไป

2.2 ผลการวิเคราะห์ความตรงเชิงโครงสร้างแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ก่อนปรับโมเดลเชิงทฤษฎีให้มีความสอดคล้องกับข้อมูลเชิงประจักษ์เพื่อตรวจสอบความตรงเชิงโครงสร้าง พบว่า $\chi^2 = 1489.16$, $df = 268$, $p = 0.00$, $GFI = 0.61$, $AGFI = 0.68$, $RMSEA = 0.134$, $RMR = 0.21$, $NFI = 0.91$, และ $CFI = 0.93$ และหลังปรับโมเดลพบว่า $\chi^2 = 141.71$, $df = 164$, $p = 0.89515$, $GFI = 0.96$, $AGFI = 0.92$, $RMSEA = 0.00$, $RMR = 0.094$, $NFI = 0.99$, และ $CFI = 1.00$ โดยผลจากการวิเคราะห์มีรายละเอียดสามารถสรุปได้ดังตาราง 2

ตาราง 2 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันขององค์ประกอบและตัวบ่งชี้แบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6

องค์ประกอบหลัก	องค์ประกอบย่อย	ตัวบ่งชี้	β	se	t	r^2
ความรู้เกี่ยวกับ นวัตกรรมด้านเทคโนโลยี (Knowledge of Technology Innovation: KTI)	Fixed - 1.00	-	1.00	-	-	-
	ความรู้เนื้อหา นวัตกรรมด้านเทคโนโลยี (Knowledge of Technology Innovation Content: KC)	-	0.68**	0.08	12.16	0.46
		I1	0.96**	0.08	12.00	0.92
		I2	0.63**	0.07	9.02	0.40
		I3	0.71**	0.07	9.87	0.51
		I4	0.33**	0.08	4.64	0.11
	ความรู้กระบวนการ นวัตกรรมด้านเทคโนโลยี (Knowledge of Technology Innovation Process: KP)	-	0.99**	0.07	19.46	0.97
		I5	0.94**	0.11	8.54	0.89
		I6	0.96**	0.03	34.24	0.93
		I7	0.98**	0.03	31.74	0.96
		I8	0.49**	0.06	8.84	0.24
	ความรู้รวบยอด นวัตกรรมด้านเทคโนโลยี (Knowledge of Technology Innovation concept: KI)	-	1.00**	0.07	20.26	0.99
		I9	0.95**	0.10	9.50	0.90
		I10	0.98**	0.02	43.01	0.97
		I11	0.98**	0.02	43.00	0.97
		I12	0.59**	0.05	11.36	0.35
ลักษณะนวัตกรรม (Innovator Attribute: INA)	Fixed - 1.00	-	1.00	-	-	-
	การจัดการและสร้างความรู้ (Management and Knowledge Creation : MC)	-	0.92**	0.09	13.57	0.84
		I13	0.86**	0.08	10.75	0.74
		I14	0.86**	0.06	16.42	0.74
		I15	0.81**	0.08	12.16	0.65
		I16	0.86**	0.08	12.48	0.73
	ความคิดสร้างสรรค์ (Creative Thinking: CT)	-	0.62**	0.06	9.60	0.39
		I23.1	0.82**	0.06	13.67	0.68
		I23.2	0.95**	0.06	16.42	0.91
		I23.3	0.97**	0.05	21.67	0.94

ตาราง 2 (ต่อ)

องค์ประกอบหลัก	องค์ประกอบย่อย	ตัวบ่งชี้	β	se	t	r^2
	ภาวะผู้นำ (Leadership: LE)	-	0.71**	0.07	7.97	0.50
		I17	0.63**	0.07	9.00	0.40
		I18	0.62**	0.08	13.50	0.38
		I19	0.53**	0.18	5.66	0.28
		I20	0.86**	0.15	9.14	0.74
		I21	0.80**	0.12	9.62	0.64
		I22	0.72**	0.12	8.98	0.52

Chi-Square = 141.71, df = 164, p = 0.89515

GFI = 0.96, AGFI = 0.92, RMSEA = 0.00, RMR = 0.094, NFI = 0.99 and CFI = 1.00

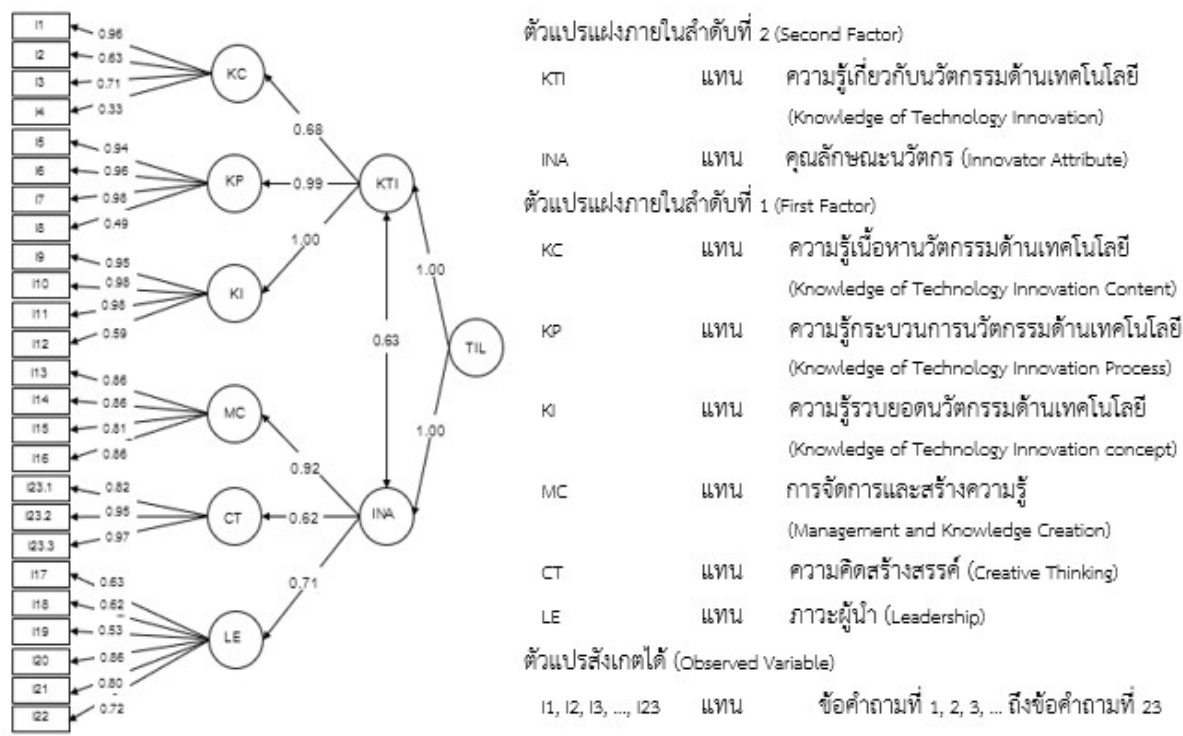
หมายเหตุ ** คือ นัยสำคัญทางสถิติ 0.01

จากตาราง 2 ผลการตรวจสอบความตรงเชิงโครงสร้างของโมเดลฯ ที่สร้างมาจากโครงสร้างความคิดหรือภาวะสันนิษฐาน พบว่า แบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีความเชิงโครงสร้างหรือความตรงเชิงทฤษฎี โดยเมื่อพิจารณาองค์ประกอบที่ 1 พบว่า ความรู้เกี่ยวกับนวัตกรรมด้านเทคโนโลยี (KTI) เป็นองค์ประกอบของการรู้เรื่องนวัตกรรมด้านเทคโนโลยี (TIL) สอดคล้องกับแนวคิดของ Gelb & Caldicott (2007); Yamakami (2012) อ้างอิงใน (Khine, 2017) ภายใต้องค์ประกอบเมื่อพิจารณาตัวบ่งชี้ที่ 1 พบว่า ความรู้เนื้อหาวัตกรรมการด้านเทคโนโลยี (KC) เป็นองค์ประกอบของความรู้เกี่ยวกับนวัตกรรมด้านเทคโนโลยี (KTI) และภายใต้ตัวบ่งชี้ที่ 1 ประกอบด้วยพฤติกรรมที่ต้องการวัด 2 พฤติกรรม โดยผลการวิจัยพบว่าพฤติกรรมที่ 1 ความสามารถในการระบุชิ้นงานหรือผลิตภัณฑ์ที่มาจากความรู้ทางวิทยาศาสตร์ (I1,I2) เป็นองค์ประกอบของความรู้เนื้อหาวัตกรรมการด้านเทคโนโลยี (KC) และพฤติกรรมที่ 2 ความสามารถในการระบุแผนการหรือกลยุทธ์ของชิ้นงานหรือผลิตภัณฑ์ที่มาจากความรู้ทางวิทยาศาสตร์ (I3,I4) เป็นองค์ประกอบของความรู้เนื้อหาวัตกรรมการด้านเทคโนโลยี ซึ่งสอดคล้องกับแนวคิดของ Philip Drucker (1956, as cited in Advanced Workplace Associate, 2015; Woods et al., 1987; Davenport et al., 2006; Shane, 2009; Katz, 2011; Trott, 2017; Kabir, 2019; Costello, 2020) และงานวิจัยของ Leber et al. (2015) ที่พบว่า ความรู้เกี่ยวกับผลิตภัณฑ์เป็นความรู้ที่จะนำมาซึ่งแนวทางผลิตภัณฑ์ หรือนวัตกรรมใหม่ เมื่อพิจารณาตัวบ่งชี้ที่ 2 พบว่า ความรู้กระบวนการนวัตกรรมด้านเทคโนโลยี (KP) เป็นองค์ประกอบของความรู้เกี่ยวกับนวัตกรรมด้านเทคโนโลยี (KTI) และภายใต้ตัวบ่งชี้ที่ 2 ประกอบด้วยพฤติกรรมที่ต้องการวัด 2 พฤติกรรม โดยผลการวิจัยพบว่าพฤติกรรมที่ 1 ความสามารถในการอธิบายขั้นตอน การออกแบบชิ้นงานหรือผลิตภัณฑ์ที่มาจากความรู้ทางวิทยาศาสตร์ (I5,I6) เป็นองค์ประกอบของความรู้กระบวนการนวัตกรรมด้านเทคโนโลยี และพฤติกรรมที่ 2 ความสามารถในการอธิบายขั้นตอนการสร้างชิ้นงานหรือผลิตภัณฑ์ที่มาจากความรู้ทางวิทยาศาสตร์ (I7,I8) เป็นองค์ประกอบของความรู้กระบวนการนวัตกรรมด้านเทคโนโลยี (KP) ซึ่งสอดคล้องกับแนวคิดของ Philip Drucker (1956, as cited in Advanced Workplace Associate, 2015) และ Woods et al. (1987, as cited in Shavinina, 2004; Shane, 2009; Carayannis et al., 2015; Trott, 2017; Costello, 2020) และงานวิจัยของ Leber et al. (2015) ที่พบว่า ความรู้เกี่ยวกับกระบวนการผลิตและวิศวกรรม

(process and engineering of product) เป็นความรู้ที่จะนำมาซึ่งแนวทางผลิตภัณฑ์หรือนวัตกรรมใหม่ เมื่อพิจารณาตัวบ่งชี้ที่ 3 พบว่า ความรู้รวบยอดนวัตกรรมด้านเทคโนโลยี (KI) เป็นตัวบ่งชี้ของความรู้เกี่ยวกับนวัตกรรมด้านเทคโนโลยี (KTI) และภายใต้ตัวบ่งชี้ที่ 3 ประกอบด้วยพฤติกรรมที่ต้องการวัด 2 พฤติกรรม โดยผลการวิจัยพบว่าพฤติกรรมที่ 1 ความสามารถในการอธิบายแนวคิดที่ทำให้เกิดชิ้นงานหรือผลิตภัณฑ์ที่สร้างมาจากความรู้ทางวิทยาศาสตร์ (I9,I10) เป็นองค์ประกอบของความรู้รวบยอดนวัตกรรมด้านเทคโนโลยี (KI) และพฤติกรรมที่ 2 ความสามารถในการสรุปแนวคิดในภาพรวมที่ทำให้เกิดชิ้นงานหรือผลิตภัณฑ์ที่สร้างมาจากความรู้ทางวิทยาศาสตร์ (I11,I12) เป็นองค์ประกอบของความรู้รวบยอดนวัตกรรมด้านเทคโนโลยี (KI) ซึ่งสอดคล้องกับแนวคิดของ Christensen (1997, as cited in Shavinina, 2004; Henry & Mayle, 2006; Vlok, 2012; Wessner, 2013; Songkram, 2013; Carayannis et al., 2015) และงานวิจัยของ Leber et al. (2015) ที่พบว่า ผลิตภัณฑ์หรือนวัตกรรมภายในองค์กรคือความรู้ในแง่มุมมองต่างๆ ที่ผ่านกระบวนการคิด การผสมผสาน และการสร้างความรู้ใหม่จนทำให้เกิดเป็นองค์ความรู้เฉพาะภายในองค์กรที่ไม่เหมือนใคร กล่าวคือ ความคิดรวบยอดที่จะนำมาซึ่งแนวทางผลิตภัณฑ์หรือนวัตกรรมใหม่ เมื่อพิจารณาองค์ประกอบที่ 2 พบว่า คุณลักษณะนวัตกรรม (INA) เป็นองค์ประกอบของการรู้เรื่องนวัตกรรมด้านเทคโนโลยี (TIL) ซึ่งสอดคล้องกับแนวคิดของ Angoff (1984); Gelb & Caldicott (2007); Luke (2008); Luke (2009); Yamakami (2012) อ้างอิงใน Khine (2017) ภายใต้องค์ประกอบเมื่อพิจารณาตัวบ่งชี้ที่ 4 พบว่า การจัดการและสร้างความรู้ (MC) เป็นองค์ประกอบของคุณลักษณะนวัตกรรม (INA) ภายใต้ตัวบ่งชี้ที่ 4 ประกอบด้วยพฤติกรรมที่ต้องการวัด 2 พฤติกรรม โดยผลการวิจัยพบว่าพฤติกรรมที่ 1 ความสามารถในการรวบรวมข้อมูลเกี่ยวกับชิ้นงานหรือผลิตภัณฑ์ที่สร้างมาจากความรู้ทางวิทยาศาสตร์ (I13,I14) เป็นองค์ประกอบของการจัดการและสร้างความรู้ (MC) และพฤติกรรมที่ 2 ความสามารถในการสร้างความรู้เกี่ยวกับชิ้นงานหรือผลิตภัณฑ์ที่สร้างมาจากความรู้ทางวิทยาศาสตร์ (I15,I16) เป็นองค์ประกอบของการจัดการและสร้างความรู้ (MC) ซึ่งสอดคล้องกับแนวคิดของ Christensen (1997); Davenport et al. (2006); Shane (2009); Katz (2011); Vlok (2012); Wessner (2013); Songkram (2013); Carayannis et al. (2015); Trott (2017), Edmondson (2019); Kabir (2019) และงานวิจัยของ Yuyuen et al. (2017) ที่พบว่า ปัจจัยการเป็นนวัตกรรม สามารถแบ่งออกเป็น 4 ประการ ได้แก่ 1) มีความคิดที่ขัดแย้ง 2) มีความคิดริเริ่มสร้างสรรค์ 3) การเปิดกว้างทางความคิด และ 4) การมีความคิดเชื่อมโยง โดยปัจจัยทั้ง 4 ประการ มีความเชื่อมโยงกันจากปัจจัยหนึ่งสู่อีกปัจจัยหนึ่งและส่งผลให้เกิดเป็นนวัตกรรม นอกจากนี้งานวิจัยของ Leber et al. (2015) ยังพบว่า ผลิตภัณฑ์หรือนวัตกรรมภายในองค์กรคือความรู้ในแง่มุมมองต่างๆ ที่ผ่านกระบวนการคิด การผสมผสาน และการสร้างความรู้ใหม่ เมื่อพิจารณาตัวบ่งชี้ที่ 5 พบว่า ความคิดสร้างสรรค์ (CT) เป็นองค์ประกอบของคุณลักษณะนวัตกรรม (INA) ภายใต้ตัวบ่งชี้ที่ 5 ประกอบด้วยพฤติกรรมที่ต้องการวัด 3 พฤติกรรม โดยผลการวิจัยพบว่าพฤติกรรมที่ 1 ความสามารถในการนำความรู้มาคิดต่อยอดให้เป็นชิ้นงานหรือผลิตภัณฑ์ใหม่ที่สร้างมาจากความรู้ทางวิทยาศาสตร์ (I23.1) เป็นองค์ประกอบของความคิดสร้างสรรค์ (CT) พฤติกรรมที่ 2 ความสามารถในการคิดออกแบบชิ้นงานหรือผลิตภัณฑ์ใหม่ที่สร้างมาจากความรู้ทางวิทยาศาสตร์ (I23.2) เป็นองค์ประกอบของความคิดสร้างสรรค์ (CT) และพฤติกรรมที่ 3 ความสามารถในการคิดสร้างชิ้นงานหรือผลิตภัณฑ์ใหม่ที่มาจากความรู้ทางวิทยาศาสตร์

(I23.3) เป็นองค์ประกอบของความคิดสร้างสรรค์ (CT) ซึ่งสอดคล้องกับแนวคิดของ Shane (2009); Katz (2011); Songkram (2013) อ้างอิงใน Shalley et al. (2015), Carayannis, Samara, & Bakouros (2015), Trott (2017); Costello (2020) และงานวิจัยของ Dorin Maier (2018) ที่พบว่า ในระดับตัวบุคคลคือความคิดที่แปลกใหม่ (novelty or new idea) ปัจจัยสำคัญที่ทำให้ผลิตภัณฑ์หรือนวัตกรรมใหม่เข้าสู่ตลาดได้อย่างยั่งยืน และเมื่อพิจารณาตัวบ่งชี้ที่ 6 พบว่า ภาวะผู้นำ (LE) เป็นองค์ประกอบของคุณลักษณะนวัตกรรม (INA) ภายใต้ตัวบ่งชี้ที่ 6 ประกอบด้วยพฤติกรรมที่ต้องการวัด 3 พฤติกรรม โดยผลการวิจัยพบว่าพฤติกรรมที่ 1 ความสามารถในการกำหนดกลยุทธ์หรือวัตถุประสงค์ให้ตรงเป้าหมายเกี่ยวกับชิ้นงานหรือผลิตภัณฑ์ที่มาจากความรู้ทางวิทยาศาสตร์ (I17,I18) เป็นองค์ประกอบของภาวะผู้นำ (LE) พฤติกรรมที่ 2 ความสามารถในการพิจารณาตัดสินใจข้อมูลอย่างรอบด้านเกี่ยวกับชิ้นงานหรือผลิตภัณฑ์ที่มาจากความรู้ทางวิทยาศาสตร์ (I19,I20) เป็นองค์ประกอบของภาวะผู้นำ (LE) และพฤติกรรมที่ 3 ความสามารถในการสร้างความร่วมมือระหว่างบุคคลภายในทีมให้ไปสู่เป้าหมายเดียวกันเกี่ยวกับชิ้นงานหรือผลิตภัณฑ์ที่มาจากความรู้ทางวิทยาศาสตร์ (I21,I22) เป็นองค์ประกอบของภาวะผู้นำ (LE) ซึ่งสอดคล้องกับแนวคิดของ Davenport et al. (2006); Katz (2011); Songkram (2013) อ้างอิงใน Shalley et al. (2015); Trott (2017); Edmondson (2019) และงานวิจัยของ Srimanee (2019) ที่พบว่า ภาวะผู้นำการเปลี่ยนแปลงเป็นปัจจัยต่อพฤติกรรมกรรมการสร้างสรรค์นวัตกรรมระดับบุคคล

กล่าวโดยสรุป แบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่สร้างขึ้นมีความตรงเชิงทฤษฎีหรือความเชิงโครงสร้าง (construct validity) ในการวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ซึ่งจากที่กล่าวมาสามารถสรุปได้ดังภาพ 2



ภาพ 2 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันขององค์ประกอบและตัวบ่งชี้แบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6

อภิปรายผล

ประเด็นที่ 1 : ความถูกต้อง เหมาะสม และน่าเชื่อถือของแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่สร้างขึ้น

เนื่องจากผลการวิเคราะห์พบว่าแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่ผู้วิจัยสร้างขึ้นมีความถูกต้อง เหมาะสม และน่าเชื่อถือในการวัด ทั้งนี้สืบเนื่องมาจากกระบวนการในการทบทวนวรรณกรรมที่มีประสิทธิภาพ สอดคล้องกับแนวคิดของ Angoff W.H. และ Cronbach L.J. ที่กล่าวว่า ค่าดัชนีความสอดคล้องของผู้ทรงคุณวุฒิจะมีค่าสูงตามความถูกต้องและความครอบคลุมของเนื้อหาที่ได้จากการทบทวนวรรณกรรมอันจะนำมาซึ่งเครื่องมือที่มีคุณภาพในที่สุด (Angoff, 1984; Cronbach, 1970)

ประเด็นที่ 2 : ค่าความยากง่ายของข้อคำถามที่ใช้ในแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 บางข้อที่มีค่าค่อนข้างยาก

เนื่องจากผลการวิเคราะห์พบว่าค่าความยากง่ายของข้อคำถามที่ใช้ในแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 บางข้อที่มีค่าค่อนข้างยาก (ค่าต่ำสุดคือ 0.23) ทั้งนี้สืบเนื่องมาจากเนื้อหาที่ใช้ในการวัดบางข้ออยู่นอกเหนือกรอบสาระมาตรฐานและตัวชี้วัดกลุ่มสาระการเรียนรู้วิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ตามหลักสูตรแกนกลางฯ ยกตัวอย่างเช่น พื้นฐานความรู้เรื่องวงจรผลิตภัณฑ์ (Product Life Cycle) ที่เป็นเนื้อหาเกี่ยวกับวิชาการตลาดที่ใช้ความรู้ในขั้นความจำก็สามารถตอบคำถามในข้อนี้ได้ แต่เมื่อนำมาสอบจริงนักเรียนกลับทำไม่ได้เนื่องจากนักเรียนไม่ได้เรียนรู้เนื้อหาดังกล่าวมาก่อน จึงกลายเป็นว่าข้อคำถามค่อนข้างยาก เป็นต้น ซึ่งตามหลักพื้นฐานในการสร้างเครื่องมือทดสอบหรือวัดโดยปกติต้องวัดความรู้ในเรื่องที่ผู้รับการทดสอบหรือวัดได้เรียนรู้เนื้อหาในเรื่องนั้นมาแล้วจึงจะทำให้ได้ค่าที่มีความเหมาะสมอย่างแท้จริง (Angoff, 1984; Cronbach, 1970)

ประเด็นที่ 3 : ค่าอำนาจจำแนกของข้อคำถามที่ใช้ในแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 บางข้อที่มีค่าอำนาจจำแนกได้ค่อนข้างน้อย

เนื่องจากผลการวิเคราะห์พบว่าค่าอำนาจจำแนกของข้อคำถามที่ใช้ในแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 บางข้อที่มีค่าค่อนข้างต่ำหรือพอใช้ (ค่าต่ำสุดคือ 0.29) ทั้งนี้สืบเนื่องมาจากเนื้อหาที่ใช้ในการวัดบางข้ออยู่นอกเหนือกรอบสาระมาตรฐานดังที่ได้กล่าวมาแล้วในประเด็นของความยากง่าย ด้วยเหตุนี้นักเรียนส่วนใหญ่จึงตอบผิดและทำให้ค่าอำนาจจำแนกในบางข้อค่อนข้างต่ำ ซึ่งตามหลักข้อสอบจะจำแนกได้ดีก็ต่อเมื่อกลุ่มสูงตอบข้อยากได้ถูกต้องและกลุ่มต่ำตอบข้อยากผิด (Angoff, 1984; Cronbach, 1970)

ประเด็นที่ 4 : ความตรงเชิงทฤษฎีหรือเชิงโครงสร้างของแบบวัดการรู้เรื่องนวัตกรรมของนักเรียนชั้นมัธยมศึกษาปีที่ 6

เนื่องจากผลการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่าโมเดลการวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่สร้างขึ้นมีความตรงเชิงทฤษฎีหรือความตรงเชิงโครงสร้าง ทั้งนี้สืบเนื่องมาจากกระบวนการในการทบทวนวรรณกรรมที่มีประสิทธิภาพ สอดคล้องกับแนวคิดของ Thorndike

R.M. และ David E.L. ที่กล่าวว่า ความครอบคลุมของการศึกษา (review) แนวคิด (concept) จะนำมาซึ่งความตรงเชิงทฤษฎีหรือความตรงเชิงโครงสร้างของเครื่องมือที่สร้าง (Thorndike & David, 1990)

ประเด็นที่ 5 : ค่า MI มีค่าสูงกว่าค่าวิกฤตของไค-สแควร์ ($141.71 > 3.84$)

เนื่องจากผลการวิเคราะห์องค์ประกอบเชิงยืนยัน พบว่า หลังจากการปรับโมเดลให้มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ ค่าสถิติทดสอบไค-สแควร์มีค่าเท่ากับ 141.71 ค่าองศาอิสระ (degree of freedom: df) มีค่าเท่ากับ 164 ค่า กล่าวคือค่า MI มีค่ามากกว่าค่าวิกฤตของไค - สแควร์ (3.84) สืบเนื่องมาจากการที่ผู้วิจัยหยุดปรับโมเดลจากการพิจารณาเมตริกซ์ความคลาดเคลื่อนมาตรฐาน (standardized residual) ที่มีค่าไม่เกิน 2.58 ทั้งนี้ผลการวิเคราะห์ก็ไม่ได้มีความผิดพลาดแต่ประการใด เพราะ P - Value มีค่าเท่ากับ 0.89515 (Diamantopoulos & Sigauw, 2000 as cited in Angsuchoti et al, 2011) ซึ่งสอดคล้องกับการศึกษาค่า MI พบว่า การพิจารณาโมเดลจากค่า MI เพียงอย่างเดียวอาจเกิดความผิดพลาดของผลการวิเคราะห์ได้เพราะค่า MI มักไม่ค้นหาโมเดลที่ถูกต้อง (Hox & Bechger, 2000 as cited in Angsuchoti et al., 2011)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 เนื่องจากผลการวิเคราะห์องค์ประกอบเชิงยืนยัน พบว่า โมเดลการวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่สร้างมาจากโครงสร้างความคิดหรือภาวะสันนิษฐาน มีความสอดคล้องกับข้อมูลเชิงประจักษ์ แสดงให้เห็นว่า องค์ประกอบหลักทั้งหมดจำนวน 2 องค์ประกอบ ตัวบ่งชี้ทั้งหมดจำนวน 6 ตัวบ่งชี้ และพฤติกรรมทั้งหมดจำนวน 14 พฤติกรรม ที่อยู่ในโมเดลการวัดย่อมมีความสำคัญต่อการพัฒนาการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ดังนั้นสถานศึกษาและหน่วยงานที่เกี่ยวข้องควรให้ความสำคัญกับการพัฒนานักเรียนให้มีความรู้ ความสามารถหรือทักษะในเรื่องดังกล่าวตามความเหมาะสม

1.2 เนื่องจากผลการวิเคราะห์องค์ประกอบเชิงยืนยัน พบว่า โมเดลการวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่สร้างมาจากโครงสร้างความคิดหรือภาวะสันนิษฐาน มีความสอดคล้องกับข้อมูลเชิงประจักษ์ แสดงให้เห็นว่า องค์ประกอบ ตัวบ่งชี้ และพฤติกรรมที่อยู่ในโมเดลการวัดย่อมมีความสำคัญต่อการพัฒนาการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ซึ่งเมื่อพิจารณากลุ่มสาระการเรียนรู้และเนื้อหาในรายวิชาที่นักเรียนจะได้เรียนรู้แล้วจะพบว่านักเรียนยังไม่มีรายวิชาที่จัดประสบการณ์การเรียนรู้เกี่ยวกับความรู้ด้านธุรกิจและการตลาดซึ่งเป็นพฤติกรรมที่สำคัญในการรู้เรื่องนวัตกรรมด้านเทคโนโลยี ดังนั้นสถานศึกษาและหน่วยงานที่เกี่ยวข้องควรให้ความสำคัญในการปรับเพิ่มหรือพัฒนาให้มียุทธศาสตร์การสอนให้นักเรียนมีความรู้ความสามารถด้านธุรกิจและการตลาด

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 เนื่องจากแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่ผู้วิจัยสร้างขึ้นอยู่บนพื้นฐานของทฤษฎีการทดสอบแบบดั้งเดิม (classical test theory) ดังนั้น เพื่อความน่าเชื่อถือ

ที่มากยิ่งขึ้นควรพิจารณาใช้ทฤษฎีการทดสอบแบบใหม่ (modern test theory) ในการทำการวิจัยหรือสร้างแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ในสังกัดอื่น ๆ ในครั้งถัดไป

2.2 เนื่องจากผลการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่า โมเดลการวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่สร้างมาจากโครงสร้างความคิดหรือภาวะสันนิษฐานมีความสอดคล้องกับข้อมูลเชิงประจักษ์ แสดงให้เห็นว่า องค์ประกอบหลักทั้งหมดจำนวน 2 องค์ประกอบ ตัวบ่งชี้ทั้งหมดจำนวน 6 ตัวบ่งชี้ และพฤติกรรมทั้งหมดจำนวน 14 พฤติกรรม ที่อยู่ในโมเดลการวัดย่อมมีความสำคัญต่อการพัฒนาการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ดังนั้นเพื่อประโยชน์สูงสุดในการพัฒนาพัฒนาการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 จึงควรมีการทำการวิจัยตัวแปรต่างๆ ที่อยู่ในโมเดลโมเดลการวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ทั้งให้ลุ่มลึกหรือกว้างขวางมากยิ่งขึ้น เช่น พัฒนาแบบวัดการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่ใช้ข้อคำถามในรูปแบบพหุมิติ การค้นหาสาเหตุหรือปัจจัยที่ส่งผลหรือมีอิทธิพลต่อการรู้เรื่องนวัตกรรมด้านเทคโนโลยีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 เป็นต้น

References

- Angoff, W. H. (1984). *Scales, Norms, and Equivalent Scores*. Educational Testing Service.
- Advanced Workplace Associate. (2015). *The 6 Factor of Knowledge Worker Productivity*. Advanced Workplace Associate.
- Carayannis, E. G., Samara, E. T., & Bakouros, Y. L. (2015). *Innovation, Technology, and Knowledge Management: Theory, Policy and Practice*. Springer.
- Christensen, C. M. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business School Press.
- Cornell University, INSEAD, & the World Intellectual Property Organization. (2019). *Global Innovation Index 2019. Creating Healthy Lives—The Future of Medical Innovation*. https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2019.pdf.
- Costello, G. J. (2020). *The Teaching of Design and Innovation: Principles and Practices*. Springer.
- Cronbach, L. J. (1970). *Essentials of Psychological Testing: Third Edition*. Harper & Row Publishers.
- Davenport, T. H., Leibold M., & Voelpel, A. S. (2006). *Strategic Management in the Innovation Economy: Strategy Approaches and Tools for Dynamic Innovation Capabilities*. John Wiley & Sons, Inc.
- Dorin, M. (2018). *Product and Process Innovation: A New Perspective on the Organizational Development*. Technical University of Cluj-Napoca.

- Edmondson, A. C. (2019). *The fearless organization Creating Psychological Safety in the Workplace for Learning, Innovation, and Growth*. John Wiley & Sons, Inc.
- Erdogan, N., Corlu, M. S., & Capraro, R. M. (2013). Defining Innovation Literacy: Do Robotics Program Help Students Develop Innovation Literacy Skills. *International Online Journal of Educational Sciences*, 5(1), 1-9. <https://iojes.net>.
- Gelb, M. J., & Caldicott, S. M. (2007). *Innovate Like Edison: The Five – Step System for Breakthrough Business Success*. Springer.
- Gipps, C. V. (1994). *Beyond Testing: Towards a theory of educational assessment*. The Falmer Press.
- Gregory, R. J. (1992). *Psychological Testing, History, Principles, and Applications*. Allyn & Bacon A Division of Simon & Schuster, Inc.
- Henry, J., & Mayle, D. (2006). *Innovation in practice*. Creativity, innovation and change media book.
- Katz, R. (2011). *Managing Creativity and Innovation*. Expernet limited.
- Kalat, J. W. (2017). *Introduction to Psychology*. Cengage Learning.
- Kabir, M. N. (2019). *Knowledge - Based Social Entrepreneurship: Understanding Knowledge Economy, Innovation, and the Future of Social Entrepreneurship*. This Palgrave Macmillan.
- Kerlinger, F. N. (1986). *Foundations of Behavioral Research Third Edition*. Holt, Rinehart and Winston, Inc.
- Khine, M. S. (2017). *Robotics in STEM Education: Redesigning the Learning Experiences*. Springer.
- Leber, M., Buchmeister, B., & Ivanisevic, A. (2015). “IMPACT OF KNOWLEDGE ON INNOVATION PROCESS”. <http://estatedocbox.com/Architects/79398719-Daaam-international-vienna.html>
- Luke, R. (2008). *George Brown College Launches Post – Graduate Certificate in Research Commercialization and Innovation*. <http://applied-research.blogspot.com>
- Luke, R. (2009). *Innovation Literacy*. <http://applied-research.blogspot.com>
- Martin, R., & Florida, R. (2009). *Ontario in the creative age*. University of Toronto.
- Medin, D. L., & Ross, B. H. (1992). *Cognitive Psychology*. Harcourt Brace Jovanovich, Inc.
- Schumacker, R. E., & Lomax, R. G. (2016). *A BEGINNER’S GUIDE TO STRUCTURAL EQUATION MODELING: Fourth Edition*. Taylor & Francis.

- Shalley, C. E., Hitt, M. A., & Zhou, J. (2015). *The Oxford Handbook of Creativity, Innovation, and Entrepreneurship*. Oxford University Press.
- Shane, S. (2009). *Handbook of Technology and Innovation Management*. John Wiley & Sons, Inc.
- Shavinina, L. V. (2004). *Silicon Valley North: A High-Tech Cluster of Innovation and Entrepreneurship (Technology, Innovation, Entrepreneurship and Competitive Strategy)*. Elsevier.
- Thorndike, R. M., & David, E. L. (1990). *A Century of Ability Testing*. The Riverside Publishing Company.
- Trott, P. (2017). *Innovation Management and New Product Development*. Pearson Education.
- Vlok, A. (2012). *A leadership competency profile for innovation leaders in a science-based research and innovation organization in South Africa*. Procedia—Social and Behavioral Sciences.
- Weiner, I. B. (2003). *Handbook of Psychology (Vol. 7)*. John Wiley & Sons, Inc.
- Wessner, C. W. (2013). *Best Practice in State and Regional Innovation Initiatives*. The National Academies Press.
- Woods, D. P., Obrien J. F., & Hanes L. F. (1987). *Human Factors challenges in process control: The case of nuclear power plants*. John Wiley & Sons, Inc.
- Yamakami, T. (2012). *Innovation Literacy: Implication from a shift toward dynamic multidisciplinary engineering*. 2012 8th International conference on information science and digital content technology, Japan. <https://www.semanticscholar.org>.

Translated Thai References

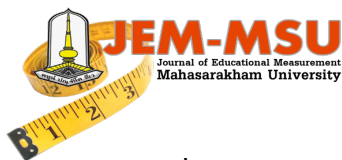
- Angsuchoti, S., Wijitwanna, S., & Pinyopanuwat, R. (2011). *Statistical analysis for social and behavioral science: technique for LISREL*. Charoendee Mankong Printing Company. (in Thai)
- Information and Communication Technology Center. (2017). *Important export products of Thailand according to the structure of world export products*. <http://tradereport.moc.go.th/Tradethai.aspx?chk=0> (in Thai)
- Ministry of industry Thailand. (2016). *Strategy for the development of Thai industry (in the past 20 years)*. <https://www.nstda.or.th/th/nstda-doc-archives/thailand-40/11623-oie-thailand-4> (in Thai)

- Songkram, N. (2013). *Innovation Creation: Turning students into innovators*. Chulalongkorn University Press. (in Thai)
- Srimanee, P. (2019). *Factors Affecting Individual Innovative Behavior: The Case of Support Staff in Prince of Songkla University (Hat Yai Campus)* [Master's thesis]. Prince of Songkla University. (in Thai)
- Sukhothai Thammathirat open university. (2012). *Educational Technology and Communications*. The Office of the University Press Sukhothai Thammathirat open university. (in Thai)
- The Secretariat of the Cabinet. (2019). *The Constitution of the Kingdom of Thailand*. <http://www.ratchakitcha.soc.go.th/DATA/PDF/2560/A/040/1.PDF> (in Thai)
- Yuyuen, P., Bhumiotar, A., & Sriyothin, S. (2017). *Factors that influence Innovator: case study of PUNW*. http://www.mis.ms.su.ac.th/MISMS01/PDF01/1844_20190611_p_296.pdf (in Thai)



ภาคผนวก

- รายชื่อผู้ทรงคุณวุฒิพิจารณาบทความ วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม ปีที่ 29 ฉบับที่ 1 (มกราคม - มิถุนายน 2566)
- หลักเกณฑ์และคำแนะนำสำหรับผู้เขียน



รายชื่อผู้ทรงคุณวุฒิพิจารณาบทความวารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม
ประจำปี 29 ฉบับที่ 1 (มกราคม - มิถุนายน 2566)

1. ผศ.ดร.อรนุช วราอัศวปติ ศรีสะอาด	มหาวิทยาลัยมหาสารคาม
2. รศ.ดร.วราพร เอราวรรณ	มหาวิทยาลัยมหาสารคาม
3. รศ.ดร.อพันธ์ พูลพุดธา	มหาวิทยาลัยมหาสารคาม
4. ผศ.ดร.สรัญญา จันทร์ชูสกุล	มหาวิทยาลัยศิลปากร
5. ผศ.ดร.กมลทิพย์ ศรีหาเศษ	มหาวิทยาลัยรามคำแหง
6. ผศ.ดร.จุฑาภรณ์ มาสันเทียะ	มหาวิทยาลัยรามคำแหง
7. ผศ.ดร.พินดา วราสุนันท์	มหาวิทยาลัยเกษตรศาสตร์
8. รศ.ดร.พูลพงศ์ สุขสว่าง	มหาวิทยาลัยบูรพา
9. อ.ดร.จริยา ชื่นศิริมงคล	มหาวิทยาลัยนวมินทราธิราช
10. รศ.ดร.อัจจรา ประเสริฐสิน	มหาวิทยาลัยศรีนครินทรวิโรฒ
11. อ.ดร.พนัส จันทรเปล่ง	มหาวิทยาลัยศรีนครินทรวิโรฒ
12. อ.ดร.อัญชลี ศรีกลชาญ	มหาวิทยาลัยศรีนครินทรวิโรฒ
13. อ.ดร.วีรภัทร์ สุขศิริ	มหาวิทยาลัยศรีนครินทรวิโรฒ
14. รศ.ดร.ชนะศึก นิขานนท์	มหาวิทยาลัยสวนดุสิต
15. รศ.ดร.สุนทรพจน์ ดำรงค์พานิช	มหาวิทยาลัยเชียงใหม่
16. ผศ.ดร.ณัฐพล แจ้งอักษร	มหาวิทยาลัยเชียงใหม่
17. รศ.ดร.พัชรี จันทรเพ็ง	มหาวิทยาลัยขอนแก่น
18. อ.ดร.ชนิตา พิมพ์ศรี	มหาวิทยาลัยขอนแก่น
19. ผศ.ดร.สาธิตา สกุรัตน์กุลชัย	มหาวิทยาลัยธรรมศาสตร์
20. ผศ.ดร.สิรินธร สิ้นจินดาวงศ์	มหาวิทยาลัยศรีปทุม
21. รศ.ดร.ชัยวิชิต เขียวชนะ	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
22. น.ต.หญิง ดร.หญทัย อาจปฐ	กองวิทยาการ ศูนย์วิทยาการ กรมแพทยทหารเรือ
23. ดร.พรพิมล ะวันประโคน	สำนักงานเขตพื้นที่การศึกษาประถมศึกษาบุรีรัมย์ เขต 2
24. อ.ดร.ถนอมรัตน์ ศิริภาพ	โรงเรียนโพธิสารศึกษา
25. ดร.วรัญญา รุมแสง	บริษัท อบรมและทดสอบ พีทีเอส จำกัด

หลักเกณฑ์และคำแนะนำสำหรับผู้เขียน (Principles and Guideline for Author)

ประเภทผลงานวิชาการ

วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม (Journal of Educational Measurement Mahasarakham University : JEM-MSU) จะรับพิจารณาผลงานวิชาการใน 3 ลักษณะ ได้แก่ บทความวิชาการ บทความวิจัย/วิทยานิพนธ์ และบทวิจารณ์หนังสือ โดยผลงานทางวิชาการดังกล่าวจะต้องไม่เคยถูกนำไปตีพิมพ์เผยแพร่ในวารสารหรือเอกสารอื่นใดมาก่อน ซึ่งผู้เขียนต้องเป็นผู้รับผิดชอบในการตรวจสอบประเด็นดังกล่าวอย่างเคร่งครัด

การเตรียมต้นฉบับ

1. ภาษา สามารถนำเสนอได้ทั้งภาษาไทยหรือภาษาอังกฤษโดยจะต้องมีบทคัดย่อทั้งภาษาไทยและภาษาอังกฤษในผลงานวิชาการนั้น ๆ

2. หัวข้อในการเขียนบทความ

2.1 บทความวิจัย มีหัวข้อดังนี้

- ชื่อเรื่อง มีทั้งภาษาไทยและภาษาอังกฤษ
- ชื่อ-สกุล/สถานภาพผู้เขียน มีทั้งภาษาไทยและภาษาอังกฤษ
- บทคัดย่อ มีทั้งภาษาไทยและภาษาอังกฤษ
- คำสำคัญ ภาษาไทยและภาษาอังกฤษ
- บทนำ
- ความมุ่งหมายของการวิจัย
- วิธีดำเนินการวิจัย
- ผลการวิจัย
- อภิปรายผล
- ข้อเสนอแนะ
- เอกสารอ้างอิง (อ้างอิงภาษาไทยต้องแปลเป็นภาษาอังกฤษทั้งหมด ตามรูปแบบ APA)

2.2 บทความวิชาการ มีหัวข้อดังนี้

- ชื่อเรื่อง มีทั้งภาษาไทยและภาษาอังกฤษ
- ชื่อ-สกุล/สถานภาพผู้เขียน มีทั้งภาษาไทยและภาษาอังกฤษ
- บทคัดย่อ มีทั้งภาษาไทยและภาษาอังกฤษ
- คำสำคัญ ภาษาไทยและภาษาอังกฤษ
- บทนำ
- เนื้อหา
- สรุป
- เอกสารอ้างอิง (อ้างอิงภาษาไทยต้องแปลเป็นภาษาอังกฤษทั้งหมด ตามรูปแบบ APA)

2.3 บทความวิจารณ์หนังสือ มีหัวข้อดังนี้

- ชื่อเรื่อง (ชื่อหนังสือ) มีทั้งภาษาไทยและภาษาอังกฤษ
- ชื่อ-สกุลผู้เขียนหนังสือ ภาษาไทยหรือภาษาอังกฤษ ขึ้นกับภาษาที่ใช้ในหนังสือ
- รายละเอียดเกี่ยวกับหนังสือที่นำมาวิจารณ์ เช่น สำนักพิมพ์ ภาษาที่ใช้ จำนวนหน้า ภาพปก
- ชื่อ-สกุล/สถานภาพผู้วิจารณ์ มีทั้งภาษาไทยและภาษาอังกฤษ
- เนื้อหาการวิจารณ์
- เอกสารอ้างอิง (อ้างอิงภาษาไทยต้องแปลเป็นภาษาอังกฤษทั้งหมด ตามรูปแบบ APA)

3. การพิมพ์

- 3.1 ให้จัดพิมพ์ด้วยโปรแกรม Microsoft Word โดยกำหนดขนาด A4 ตั้งค่าหน้ากระดาษสำหรับการพิมพ์ห่างจากขอบกระดาษทุกด้าน ๆ ละ 1 นิ้ว (2.54 เซนติเมตร) และจัดรูปแบบเป็นหนึ่งคอลัมน์เท่านั้น
- 3.2 ตัวอักษร TH SarabunPSK ขนาด 18 pt สำหรับชื่อเรื่อง จัดกลางหน้ากระดาษ
- 3.3 ตัวอักษร TH SarabunPSK ขนาด 16 pt สำหรับเนื้อหา จัดชิดขอบ
- 3.4 ตัวอักษร TH SarabunPSK ขนาด 15 pt สำหรับ FootNote จัดชิดซ้ายหน้ากระดาษ
- 3.5 ช่องว่างระหว่างคำให้ตั้งค่าเป็น Normal ไม่บีบหรือขยายตัวอักษร ยกเว้นในตารางหรือแผนภาพ
- 3.6 ช่องว่างระหว่างบรรทัดให้ตั้งค่าเป็น single space
- 3.7 การเว้นวรรค 1 เคาะเท่านั้น
- 3.8 ตาราง และ ภาพ ใช้คำว่า ตาราง 1 และ ภาพ 1
4. จำนวนหน้าของผลงาน ไม่ควรเกิน 15 หน้า (รวมเอกสารอ้างอิง หลังตัดอ้างอิงภาษาไทยออกแล้ว)
5. เอกสารอ้างอิง มีรายการอ้างอิงครบถ้วน สมบูรณ์ ตามหลักการเขียนการอ้างอิงเอกสารแบบ APA

การเขียนเอกสารอ้างอิง

ด้วยวารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม มีความพยายามที่จะพัฒนาคุณภาพวารสารสู่มาตรฐานสากล ดังนั้น กองบรรณาธิการวารสารฯ จึงกำหนดรูปแบบการเขียนเอกสารดังนี้

- กำหนดให้ใช้รูปแบบการอ้างอิงแบบ APA (7th edition)
- เอกสารอ้างอิงภาษาไทยต้องแปลเป็นภาษาอังกฤษทุกรายการและเติมคำว่า (in Thai) ต่อท้าย
- กรณีส่งบทความให้วารสารพิจารณา ยังคงเอกสารอ้างอิงภาษาไทยเดิมไว้ด้วย เขียนจัดเรียงคู่กัน เพื่อให้วารสารฯ สามารถตรวจสอบความถูกต้องของการเขียนอ้างอิงในรูปแบบ Translate Thai Reference และอ้างอิงภาษาไทยจะถูกตัดออกในขั้นตอนของการตีพิมพ์เผยแพร่
- การเรียงลำดับอ้างอิง กรณีเอกสารอ้างอิงที่แปลจากภาษาไทยเป็นภาษาอังกฤษให้ยึดตัวอักษรภาษาอังกฤษในการเรียงลำดับ

*การเขียนรายการอ้างอิงภาษาไทยเป็นภาษาอังกฤษ มีวิธีการเขียนที่คล้ายคลึงกับการเขียนรายการอ้างอิงตามรูปแบบของภาษาอังกฤษในแต่ละประเภท มีรายละเอียดดังนี้

- 1) ชื่อผู้เขียน ขึ้นต้นด้วยนามสกุลของผู้เขียนและต่อด้วยชื่อย่อ
- 2) ชื่อสำนักงานหรือกระทรวง เขียนเป็นชื่อภาษาอังกฤษ
- 3) ปีที่พิมพ์ เปลี่ยนจาก พ.ศ. เป็น ค.ศ.
- 4) ชื่อเรื่อง ชื่อวิทยานิพนธ์ ชื่อบทความ ชื่อวารสาร และชื่องานประชุมวิชาการ แปลเป็นภาษาอังกฤษ

ตัวอย่างการเขียนอ้างอิงจากเอกสารต่าง ๆ

วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม กำหนดให้ผู้เขียนอ้างอิงเป็นภาษาอังกฤษทั้งหมด กรณีอ้างอิงภาษาไทยให้แปลเป็นภาษาอังกฤษ อย่างไรก็ตาม เพื่อให้ผู้เขียนเห็นถึงที่มาของการแปล ทางวารสารฯ จึงนำเสนอตัวอย่างทั้งสองภาษา ดังนี้

1. การอ้างอิงจากหนังสือ

1) ผู้แต่ง 1 คน

ไทย	ชื่อ/สกุล/(ปีที่พิมพ์)/ชื่อเรื่อง/(พิมพ์ครั้งที่)/สำนักพิมพ์.
อังกฤษ	สกุล,/อักษรย่อชื่อ/(ปีที่พิมพ์)/ชื่อเรื่อง/(พิมพ์ครั้งที่)/สำนักพิมพ์.

ประวิต เอรารวรรณ์. (2562). *การออกแบบโครงการทางการศึกษาและการประเมินแบบเร่งด่วน*. สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

Joint Committee. (1994). *The program evaluation standards: How to assess evaluations of educational programs*. SAGE Publication.

2) ผู้แต่ง 2 คน

ไทย	ชื่อ/สกุลและชื่อ/สกุล/(ปีที่พิมพ์)/ชื่อเรื่อง/(พิมพ์ครั้งที่)/สำนักพิมพ์.
อังกฤษ	สกุล,/อักษรย่อชื่อ./&/สกุล,/ชื่อ/(ปีที่พิมพ์)/ชื่อเรื่อง/(พิมพ์ครั้งที่)/สำนักพิมพ์.

นงลักษณ์ วิรัชชัย และ สุวิมล ว่องวาณิช. (2550). *แนวทางการให้คำปรึกษาวิทยานิพนธ์*. ศูนย์ตำราและเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

Watt, S., & Darling, N. (2017). *Teaching visual methods in the social sciences*. Routledge.

3) ผู้แต่ง 3-20 คน ให้ใส่ชื่อทุกคน

ไทย	ชื่อ ¹ /สกุล ¹ ,/ชื่อ ² /สกุล ² ,/ชื่อ ³ /สกุล ³ ,/ชื่อ ⁴ /สกุล ⁴ ,/ชื่อ ⁵ /สกุล ⁵ ,/ชื่อ ⁶ /สกุล ⁶ ,/ชื่อ ⁷ /สกุล ⁷ /(ปีที่พิมพ์)/ชื่อเรื่อง/(พิมพ์ครั้งที่)/สำนักพิมพ์.
อังกฤษ	สกุล ¹ ,/อักษรย่อชื่อ ¹ ./สกุล ² ,/อักษรย่อชื่อ ² ./สกุล ³ ,/อักษรย่อชื่อ ³ ./สกุล ⁴ ,/อักษรย่อชื่อ ⁴ ./สกุล ⁵ ,/อักษรย่อชื่อ ⁵ ./สกุล ⁶ ,/อักษรย่อชื่อ ⁶ ./&/สกุล ⁷ ,/อักษรย่อชื่อ ⁷ /(ปีที่พิมพ์)/ชื่อเรื่อง/(พิมพ์ครั้งที่)/สำนักพิมพ์.

พชร สันทัด, สุรัชชาติ ณ หนองคาย, สมาน งามสนธิ, ฐนันดรศักดิ์ บวรนนท์กุล, เชษฐรัชดา พรธนาธิกุล, ฤกษ์ สติธยพัฒนานนท์, และ ประยงค์ เต็มขวลา. (2557). *ศาสตร์และศิลป์การบริหารการพัฒนาทรัพยากรมนุษย์*. สมาคมส่งเสริมคุณธรรม.

Ary, D., Jacobs, L. C., Irvine, A. K. S., & Walker, D. A. (2018). *Introduction to research in education*. Wadsworth.

* กรณีไม่ปรากฏปีที่พิมพ์ให้ใส่ (ม.ป.ป.) หรือ (n.d.)

** กรณีพิมพ์ครั้งที่ 1 ไม่ต้องระบุครั้งที่พิมพ์

*** กรณีไม่ปรากฏสำนักพิมพ์ให้ใส่ (ม.ป.ท.) หรือ (n.p.)

2. การอ้างอิงบทความจากวารสารแบบรูปเล่ม/อิเล็กทรอนิกส์ที่ไม่มี DOI

ไทย	ชื่อ/สกุล/(ปีที่พิมพ์)/ชื่อบทความ/ชื่อวารสาร,/ปีที่(ฉบับที่),/เลขหน้า.
อังกฤษ	สกุล,/อักษรย่อชื่อ/(ปีที่พิมพ์)/ชื่อบทความ/ชื่อวารสาร,/ปีที่(ฉบับที่),/เลขหน้า.

สุนันท์ สีสาย และ ไพฑูรย์ สีนลารัตน์. (2561). เปลี่ยนผ่านการศึกษาไทยสู่ การศึกษา 4.0. *วารสารการวัดผล การศึกษา มหาวิทยาลัยมหาสารคาม*, 24(2), 13-27.

Wang, W., Song, L., Chen, P., Meng, Y., & Ding, S. (2015). Attribute level and pattern level classification consistency and accuracy indices for cognitive diagnostic assessment. *Journal of Educational Measurement*, 52(4), 457-476.

3. การอ้างอิงบทความจากวารสารอิเล็กทรอนิกส์ (มี DOI)

ไทย	ชื่อ/สกุล/(ปีพิมพ์)/ชื่อบทความ/ชื่อวารสาร,/เลขของปีที่(เลขของฉบับที่), / /////เลขหน้า./https://doi.org/เลขdoi
อังกฤษ	สกุล,/อักษรย่อชื่อ/(ปีพิมพ์)/ชื่อบทความ/ชื่อวารสาร,/เลขของปีที่(เลขของฉบับที่),/ /////เลขหน้า./https://doi.org/doi

Rubright, J. D. (2018). Impact of both local item dependencies and cut-point locations on examinee classifications. *Educational Measurement: Issues and Practice*, 37(3), 40-45.
<https://doi.org/10.1111/emip.12183>

4. การอ้างอิงจากวิทยานิพนธ์ (รูปเล่ม)

ไทย	ชื่อ/สกุล/(ปีที่เผยแพร่)/ชื่อวิทยานิพนธ์/[วิทยานิพนธ์ปริญญาคุณวุฒิปริญญาตรี หรือ /////วิทยานิพนธ์ปริญญาโท/ปริญญาเอก]./ชื่อมหาวิทยาลัย.
อังกฤษ	สกุล,/อักษรย่อชื่อ/(ปีที่เผยแพร่)/ชื่อวิทยานิพนธ์/[Doctoral dissertation or /////Master's thesis]./ชื่อมหาวิทยาลัย.

สุจิตรา โชฬิมา. (2559). การพัฒนาตัวบ่งชี้สมรรถนะเชิงวัฒนธรรมของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย
สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 20 (อุดรธานี) [วิทยานิพนธ์ปริญญาโท].
มหาวิทยาลัยมหาสารคาม.

Ata, A. (2015). *Factor effecting teacher-child communication skill & self- efficacy beliefs: An
investigation on preschool teachers* [Master's thesis]. Middle East Technical University.

5. การอ้างอิงวิทยานิพนธ์จากเว็บไซต์

ไทย	ชื่อ/สกุล/(ปีที่เผยแพร่)/ชื่อวิทยานิพนธ์/[วิทยานิพนธ์ปริญญาโทหรือ //////วิทยานิพนธ์ปริญญาโทหรือปริญญาเอก/ชื่อเว็บไซต์./URL
อังกฤษ	สกุล/อักษรย่อชื่อ/(ปีที่เผยแพร่)/ชื่อวิทยานิพนธ์/[Doctoral dissertation or //////Master's thesis]/ชื่อเว็บไซต์./URL

สุจิตรา โชฬิมา. (2559). การพัฒนาตัวบ่งชี้สมรรถนะเชิงวัฒนธรรมของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย
สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 20 (อุดรธานี) [วิทยานิพนธ์ปริญญาโท].
Academic Resource Center Mahasarakham University. [http://khoon.msu.ac.th/_dir/
fulltext/fulltextman/full4/sujitta11402/titlepage.pdf](http://khoon.msu.ac.th/_dir/fulltext/fulltextman/full4/sujitta11402/titlepage.pdf)

Ata, A. (2015). *Factor effecting teacher-child communication skill & self- efficacy beliefs: An
investigation on preschool teachers* [Master's thesis]. Middle East Technical University
Library. <http://etd.lib.metu.edu.tr/upload/12619101/index.pdf>

6. การอ้างอิงจากเว็บไซต์

ไทย	ชื่อ/สกุลผู้เขียน/(ปี,/วัน/เดือนที่เผยแพร่)/ชื่อบทความ./ชื่อเว็บไซต์./URL
อังกฤษ	สกุล/อักษรย่อชื่อผู้เขียน/(ปี, เดือน/วันที่เผยแพร่)/ชื่อบทความ./ชื่อเว็บไซต์./URL
* กรณีที่ไม่มีวันที่เผยแพร่ปรากฏ ให้ใส่ (ม.ป.ป.) หรือ (n.d.)	
* กรณีที่มีปรากฏเฉพาะ พ.ศ. หรือ ค.ศ. ให้ใส่แค่ (ปี) เท่านั้น	
* กรณีชื่อผู้เขียนและชื่อเว็บไซต์เป็นชื่อเดียวกัน ให้ตัดชื่อเว็บไซต์ออก	

เกรียงศักดิ์ เจริญวงศ์ศักดิ์. (2552). *บทเรียนสหรัฐอเมริกา เปิดหลักสูตรสายอาชีพตามความต้องการ
ตลาดแรงงาน*. <http://www.kriengsak.com/>

การอ้างอิงในเนื้อหา

การอ้างอิงในเนื้อหาให้ใช้ระบบนาม-ปี โดยต้องแปลเป็นภาษาอังกฤษทั้งหมด โดยมีรูปแบบการเขียนดังนี้

1. กรณีผู้แต่ง 1 คน

สกุล/(ปี) หรือ (สกุล, ปี)

Parker (2019) พบว่า.....

.....การใช้สื่อสังคมออนไลน์และความเครียดของวัยรุ่นมีความสัมพันธ์กัน (Parker, 2019)

2. กรณีผู้แต่ง 2 คน ให้เชื่อมด้วย “and” หรือ “&” ดังนี้

สกุล/and/สกุล/(ปี) หรือ (สกุล/&/สกุล, ปี)

Sawangboon and Erawan (2020) กล่าวว่า

....รวมถึงยกระดับระบบการประเมินคุณภาพให้ได้มาตรฐาน (Sawangboon & Erawan, 2020)

3. กรณีผู้แต่ง 2 คน ขึ้นไป ให้เขียนนามสกุลเฉพาะคนแรกแล้วตามด้วย et al

สกุล/et/al./(ปี) หรือ (สกุล/et/al., ปี)

Pearson et al. (1984) กล่าวว่า.....

.....(Pearson et al, 1984)

4. กรณีการอ้างอิงมากกว่าสองแหล่งข้อมูล ให้เขียนเรียงตามลำดับอักษร และค้นแต่ละแหล่งด้วยเครื่องหมายอัฒภาค (;)

Embretson & Reise (2000); Marvelde et al. (2006) ได้ข้อค้นพบว่าสอดคล้องกันว่า.....

..... (Embretson & Reise, 2000; Marvelde et al., 2006)

5. กรณีการอ้างอิงที่ไม่ได้อ้างจากต้นฉบับแต่เป็นการอ้างอิงต่อให้ใช้คำว่า “as cited in”

งานของผู้เขียน (1) ถูกอ้างอยู่ในงานของผู้เขียน (2)

สกุล⁽¹⁾/(ปี,/as cited in/สกุล⁽²⁾,/ปี).....หรือ(สกุล⁽¹⁾,/ปี/as cited in/สกุล⁽²⁾,/ปี)

Arnett (2000, as cited in Claiborne & Drewery, 2010) ให้ข้อเสนอแนะว่า.....

..... (Arnett, 2000 as cited in Claiborne & Drewery, 2010)

การเขียนรายการอ้างอิงภาษาไทยเป็นภาษาอังกฤษ (Translate Thai Reference)

การเขียนรายการอ้างอิงภาษาไทยเป็นภาษาอังกฤษ มีวิธีการเขียนที่คล้ายคลึงกับการเขียนรายการอ้างอิงตามรูปแบบของภาษาอังกฤษในแต่ละประเภทของสารสนเทศ มีรายละเอียดดังนี้

1. ชื่อผู้เขียน ขึ้นต้นด้วยนามสกุลของผู้เขียนและต่อด้วยชื่อย่อ
2. ชื่อสำนักงานหรือกระทรวง เขียนเป็นชื่อภาษาอังกฤษ
3. ปีที่พิมพ์ เปลี่ยนจาก พ.ศ. เป็น ค.ศ.
4. ชื่อเรื่อง ชื่อวิทยานิพนธ์ ชื่อบทความ ชื่อวารสาร และชื่องานประชุมวิชาการ แปลเป็นภาษาอังกฤษ

ตัวอย่างการอ้างอิงในเนื้อหา

ทัศนศิริรินทร์ สว่างบุญ (2563) ได้ศึกษาวิจัยเรื่อง.....

Sawangboon (2020) ได้ศึกษาวิจัยเรื่อง.....

....รวมถึงยกระดับระบบการประเมินคุณภาพให้ได้มาตรฐาน (ทัศนศิริรินทร์ สว่างบุญ, 2563)

....รวมถึงยกระดับระบบการประเมินคุณภาพให้ได้มาตรฐาน (Sawangboon, 2020)

รายการอ้างอิง (Reference)

ประวิต เอรารวรรณ์. (2562). *การออกแบบโครงการทางการศึกษาและการประเมินแบบเร่งด่วน*. สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

Erawan, P. (2019). *Designing education program and rapid appraisal*. Chulalongkorn University Press. (in Thai)

พิสิษฐ์ พจนจารุวิทย์. (2559). การสร้างแบรนด์องค์กรของมหาวิทยาลัยราชภัฏในประเทศไทย. *วารสารวิทยาลัยพาณิชยศาสตร์บูรพาปริทัศน์*, 11(1), 17 – 49.

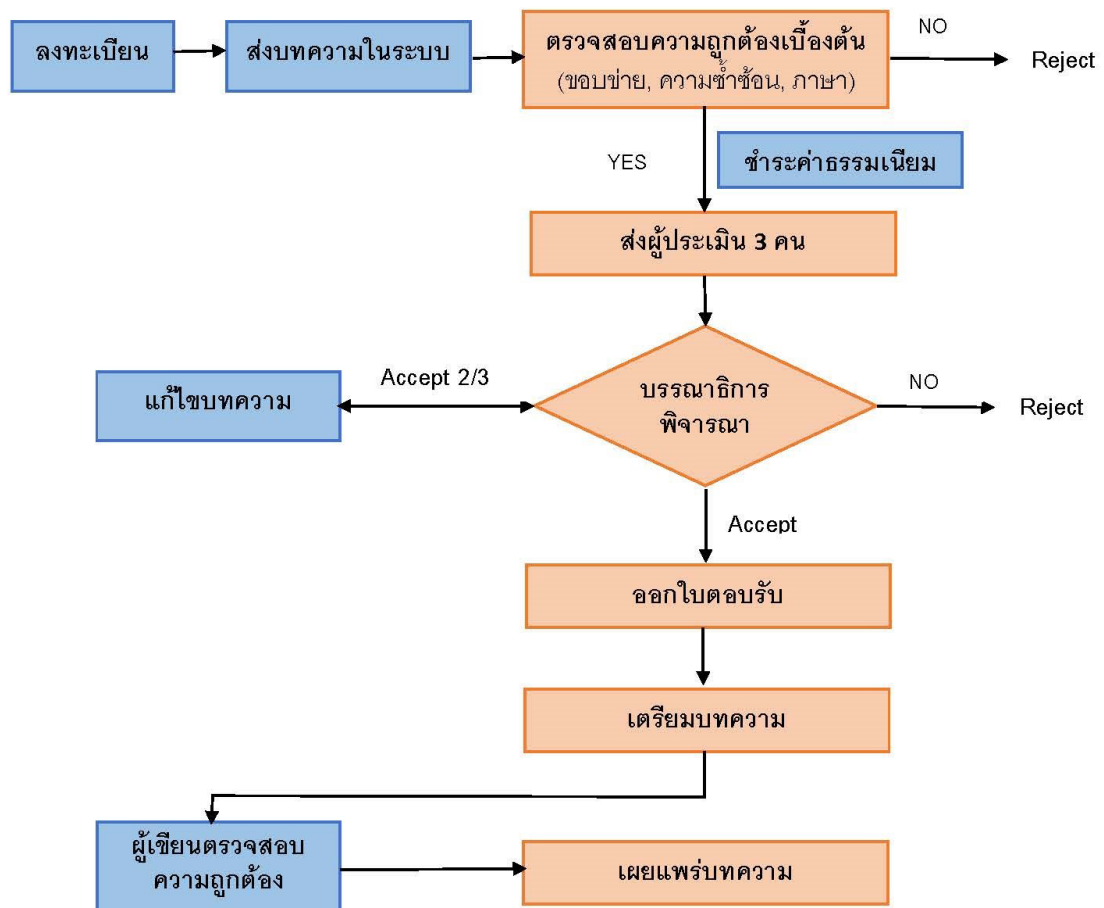
Potjanajruwit, P. (2016). The Corporate Branding Strategy of Rajabhat University in Thailand. *Journal of Graduate School of Commerce Burapha Review*, 11(1), 17 – 49. (in Thai)

สุจิตรา โชฬิมาย. (2559). *การพัฒนาตัวบ่งชี้สมรรถนะเชิงวัฒนธรรมของนักเรียนระดับชั้นมัธยมศึกษาตอนปลายสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 20 (อุดรธานี) [วิทยานิพนธ์ปริญญาโท]*. มหาวิทยาลัยมหาสารคาม.

Sopimai. S. (2016). *Developing Indicators of Cultural Competence for High School Students under The Secondary Education Service Area 20 (Udon Thani)* [Master's thesis]. Mahasarakham University. (in Thai)

ขั้นตอนการยื่นบทความ

ขั้นตอนการยื่นบทความให้ผู้พิมพ์ (Author) สมัครในระบบ TCI-Thaijo โดยมีกระบวนการดังต่อไปนี้ (ศึกษารายละเอียดเพิ่มเติมจากเว็บไซต์)



อัตราค่าพิจารณาบทความ (Rate)

1. กรณีผู้พิมพ์เป็นคณาจารย์ และบุคลากร คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม จ่ายค่าพิจารณาบทความในอัตรา 3,000 บาท/1 บทความ
2. กรณีผู้พิมพ์เป็นบุคคลทั่วไป จ่ายค่าพิจารณาบทความในอัตรา 4,000 บาท/1 บทความ

การชำระค่าธรรมเนียม

โปรดชำระค่าธรรมเนียมผ่านทาง

ธนาคารทหารไทยธนชาต สาขาเสริมไทยคอมเพล็กซ์ มหาสารคาม

ชื่อบัญชี วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม (JEM MSU)

ประเภท ออมทรัพย์ บัญชีเลขที่ 438-2-70054-2



วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม
Journal of Educational Measurement
Mahasarakham University



Website: <https://so02.tci-thaijo.org/index.php/jemmsu/index>

ภาควิชาวิจัยและพัฒนาศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

สำนักงาน: ชั้น 8 อาคารวิทยพัฒนา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม
ต.ตลาด อ.เมือง จ.มหาสารคาม 44000

โทรศัพท์: 043-754321-40 ต่อ 6272

E-mail : jem2msu@gmail.com