

วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม

Journal of Educational Measurement

Mahasarakham University



วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม

Journal of Educational Measurement Mahasarakham University

ปีที่ 31 ฉบับที่ 1: มกราคม – มิถุนายน 2568

Vol. 31 No. 1: January – June 2025

ISSN 2985-0614 (Online)

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

ภาควิชาวิจัยและพัฒนาการศึกษา

คณะศึกษาศาสตร์

มหาวิทยาลัยมหาสารคาม

วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม
ปีที่ 31 ฉบับที่ 1: มกราคม – มิถุนายน 2568
Journal of Educational Measurement Mahasarakham University
Vol. 31 No. 1: January – June 2025

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
ผ่านการตรวจประเมินคุณภาพจากผู้ทรงคุณวุฒิจำนวน 3 ท่าน
ในรูปแบบ Double – Blinded Review



บทบรรณาธิการ

วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม ISSN 2985-0614 (Online) ผ่านการประเมินคุณภาพวารสารวิชาการในฐานข้อมูล TCI รอบที่ 5 พ.ศ. 2568 – 2572 โดยถูกจัดให้เป็นวารสารกลุ่มที่ 2 ทั้งนี้วารสารฯ ยังคงรักษาคุณภาพและมาตรฐานเพื่อให้ผู้อ่านได้รับประโยชน์สูงสุด

สำหรับวารสารปีที่ 31 ฉบับที่ 1: มกราคม – มิถุนายน 2568 มีบทความวิจัยเผยแพร่ทั้งสิ้น 15 เรื่อง โดยแบ่งเป็นบทความภาษาอังกฤษ 1 เรื่อง และบทความภาษาไทย 14 เรื่อง โดยบทความทุกเรื่องผ่านการพิจารณากลั่นกรองจากผู้ทรงคุณวุฒิจำนวน 3 ท่าน จากหลากหลายสถาบัน ในรูปแบบ Double – Blinded Review และทุกกระบวนการอยู่ภายใต้การควบคุมจรรยาบรรณอย่างเคร่งครัด

กองบรรณาธิการหวังเป็นอย่างยิ่งว่าวารสารฉบับนี้จะเป็นประโยชน์ทางวิชาการสำหรับผู้อ่านทุกท่าน หากท่านมีข้อเสนอแนะประการใด กองบรรณาธิการยินดีน้อมรับคำแนะนำเพื่อจะได้นำไปปรับปรุงและพัฒนาคุณภาพวารสารให้มีมาตรฐาน 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 เอราวรรณ
สำนักงานเลขาธิการสภาการศึกษา

รศ.สมนึก ภัททิยธนี
ข้าราชการเกษียณ

บรรณาธิการ

ผศ.ดร.ทัศน์ศิริรินทร์ สว่างบุญ
มหาวิทยาลัยมหาสารคาม

ผู้ช่วยบรรณาธิการ

อ.ดร.ศุภานดา ด้านกุลประเสริฐ
มหาวิทยาลัยมหาสารคาม

กองบรรณาธิการ

รศ.ดร.ญาณภัทร สีหะมงคล	มหาวิทยาลัยมหาสารคาม
รศ.ดร.ทรงศักดิ์ ภูสีอ่อน	มหาวิทยาลัยมหาสารคาม
รศ.ดร.ประเสริฐ เรือนนงการ	มหาวิทยาลัยมหาสารคาม
รศ.ดร.วราพร เอราวรรณ	มหาวิทยาลัยมหาสารคาม
รศ.ดร.อพนันต์ พูลพุทธา	มหาวิทยาลัยมหาสารคาม
อ.ดร.สวิตรี 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 ฤทธิสัน	มหาวิทยาลัยกาฬสินธุ์
ผศ.ดร.สาธิตา สกุลรัตน์กุลชัย	มหาวิทยาลัยธรรมศาสตร์
ผศ.ดร.พินดา วราสุนันท์	มหาวิทยาลัยเกษตรศาสตร์
รศ.ดร.จตุภูมิ เขตจัตุรัส	มหาวิทยาลัยขอนแก่น
ผศ.ดร.สุกัญญา รุจิเมธธาส	มหาวิทยาลัยราชภัฏอุดรดิตถ์
ผศ.ดร.สรัญญา จันทร์ชูสกุล	มหาวิทยาลัยศิลปากร
รศ.ดร.กมลทิพย์ ศรีหาเศษ	มหาวิทยาลัยรามคำแหง
อ.ดร.ชลี ภัทรพิชญธรรม Southern Methodist University, USA	
ผศ.ดร.อรนุช วราอัศวปติ ศรีสะอาด ข้าราชการเกษียณ	

ฝ่ายประสานงาน

นางสาวเอมอร สิงห์พันธ์

สารบัญ

บทความ

หน้า

<i>Development of Multicultural Competency in Pre-Service Teacher Students: Evidence from a University in Thailand (English)</i>	
Bovornpot Choompunuch, Khanika Kamdee and Prakittiya Taksino.....	1
การพัฒนาแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6	
<i>The Development of a Critical Thinking Skills Measurement Tool in Science and Technology for Grade 6 Students</i>	
ปิยาภรณ์ ประถมวงษ์ และ ประกฤติยา ทักซิโน.....	18
การพัฒนาแบบวัดโกรว์ธมายด์เซตสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้นด้วยการประยุกต์ใช้แองเคอริงวินเยตต์	
<i>The Development of a Scale to Measure Growth Mindsets for Lower Secondary School Students with the Application of Anchoring Vignettes</i>	
สิริวัฒน์ หลูโน, สัจวรรณ ังดกระโทก และ ปุริมปรัชญ์ คณินพศุทธิ์.....	34
การพัฒนาชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเลิร์ฟสำหรับนักกีฬา เซปักตะกร้อระดับเยาวชน	
<i>Development of an Assessment Package for Physical Fitness and Serving Ability in Youth Sepak Takraw Athletes</i>	
กฤษฎา พุ่มพิน, ศุภวรรณ วงศ์สร้างทรัพย์ และ นิรอมลี มะกาเจ.....	53
การพัฒนาหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทางสู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ	
<i>Development of a Training Curriculum to Enhance Teaching Standards Aligned with Formative and Summative Indicators for Quality School Curriculum Development</i>	
กิตติรัตน์ เกษตรสุนทร, จักรชัย ตระกูลโอสถ, รักษ์ทวี เถาโต, วชิรญาณ สาดล่าง, วัชรุฒม์ อยู่ในคิล และ กวีเชษฐ เปี้ย.....	69
การพัฒนาโมเดลการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนมัธยมศึกษาหลังวิกฤต COVID-19: การวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง	
<i>Development of New Normal Learning Adaptation Model for High School Students Post-COVID-19 Crisis: A Second-Order Confirmatory Factor Analysis</i>	
ธนัพร พาสุริยันต์ และ สุนทรพจน์ ดำรงพานิช.....	87

สารบัญ (ต่อ)

บทความ	หน้า
การสร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียน ชั้นมัธยมศึกษาตอนปลาย : การประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น <i>Construction of Career Decision-Making Self-Efficacy Scale in the Digital Age of High School Students: Application of the Graded Response Model</i> สุลักษณ์ อ่อนนวล, จุฑาภรณ์ มาสันเทียะ และ ดารุณี ทิพย์กุลไพโรจน์.....	102
การพัฒนาแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์เรื่อง แก๊ส สำหรับนักเรียนระดับชั้น มัธยมศึกษาตอนปลาย <i>Development of a Scientific Explanation Test on the topic of Gases for Upper Secondary School Students</i> จักรพงศ์ ปินชัย และ ผนวรา สีที.....	121
การเปรียบเทียบคุณภาพของรูปแบบการวัดและประเมินผลตามแผนการศึกษาของ มหาวิทยาลัยสุโขทัยธรรมมาธิราช <i>A Comparison of the Quality of Measurement and Evaluation Models under the Study Plans of Sukhothai Thammathirat Open University</i> วินิตา แก้วเกื้อ และ สฤณีเทพ สุขแก้ว.....	138
การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกเพื่อการสื่อสารระหว่างนิสิตครูกับนักเรียนที่ มีความต้องการพิเศษ: การประยุกต์โมเดลซีเอ็นเอ็นร่วมกับอาร์เอ็นเอ็น <i>Development of Thai sign language recognition model using deep learning for communication between student teachers and student with special needs: application of CNN and RNN models</i> ธีรวิษณุ คะรานรัมย์ และ สิวะโชติ ศรีสุทธียากร.....	152
การพัฒนาแบบทดสอบสมรรถนะดิจิทัลในระดับปริญญาตรี โดยใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ <i>Development of the Digital Competency Test for Bachelor's Degree Using the Multidimensional Item Response Theory</i> อัศนีย์ ทองศิลป์ และ กมลวรรณ ตังธนกานนท์.....	173
การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ สำหรับการศึกษาทางไกล <i>The Development of Automatic Parallel Multiple-Choice Question Generation System Using Automatic Item Generation for Distance Education</i> รัชกฤษ ธนพัฒน์.....	189

สารบัญ (ต่อ)

บทความ	หน้า
การพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรีด้วยการจัดการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับ <i>The Development of Learning Design Competency for Academic Support Staff of King Mongkut's University of Technology Thonburi through Feedback Processes in Experiential Learning</i> วรรณชรร ไชยเดช และ วารุณี สักนโชคดี.....	209
ผลการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการที่มีต่อความสามารถในการเขียน ความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6: การทดลองแบบอนุกรมเวลา <i>The Effects of the Process Writing Approach on Thai Essay Writing Ability of Grade 6 Students: A Time Series Experiment</i> จิรัฐติกา สาระบัณฑิตกุล, สุวิมล กฤษศยาสา และ เด่นดาว ชลวิทย์.....	228
การพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติในรูปแบบแอปพลิเคชันสำหรับนักเรียน ชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ <i>Development of a Multidimensional Mathematical Literacy Test in an Application Platform for Grade 9 Students with Feedback Integration</i> ปราณี แก้วมา, น้ำทิพย์ งามอจจาณิษฐ์ และ สำราญ มีแจ้ง.....	244
ภาคผนวก.....	260
รายชื่อผู้ทรงคุณวุฒิพิจารณาบทความ วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม ปีที่ 31 ฉบับที่ 1 (มกราคม – มิถุนายน 2568).....	261
หลักเกณฑ์และคำแนะนำสำหรับผู้เขียน.....	262

Development of Multicultural Competency in Pre-Service Teacher Students: Evidence from a University in Thailand

Bovornpot Choompunuch¹ Khanika Kamdee² and Prakittiya Taksino^{3*}

(Received: July 10, 2024; Revised: August 22, 2024; Accepted: August 27, 2024)

Abstract

In today's globalized world, the diverse cultural dynamics within classrooms necessitate that educators possess strong multicultural competencies. This study sought to enhance these competencies through a structured Multicultural Competency Training Program (MCTP). The research aimed to assess the effectiveness of the MCTP in improving the multicultural competencies of these students. This quasi-experimental study involved 60 third-year pre-service teacher students, divided into an intervention group and a control group, each with 30 students. The intervention group underwent 14 sessions of the MCTP over seven weeks, while the control group continued their regular curriculum. Multicultural competency was measured using the Multicultural Competency for Pre-Service Teacher Students Test with reliability values .972 before the intervention, immediately after, and one month later. Data were analyzed using descriptive statistics and repeated measures ANOVA. The results showed that multicultural competence scores of the intervention group after participating in the Multicultural Competency Training Program (MCTP) and at follow-up was higher than before the experiment with statistical significance at the level of $p < .001$, and the Multicultural competence scores of the intervention group after participating in the MCTP and at follow-up was higher than that of the control group with statistical significance at the level of $p < .001$. In conclusion, the MCTP effectively enhances the multicultural awareness, knowledge, and personal skills of pre-service teachers, which are essential for their professional preparation in diverse educational environments.

Keywords: Multicultural competency, Pre-service teacher students, Developing of multicultural competency, Multicultural competency training program

¹ Ph.D. student, Multicultural Psychology Program, Faculty of Education, Khon Kaen University

² Assistant Professor, Division of Humanities, Faculty of Humanities and Social Sciences, Khon Kaen University

³ Assistant Professor, Division of Educational Measurement and Evaluation, Faculty of Education, Khon Kaen University

* Corresponding Author E-mail: prakittiya.t@kku.ac.th

Introduction

In the rapidly changing field of education, the roles and responsibilities of teachers are undergoing significant shifts (Ingersoll et al., 2021; Reinius et al., 2022). Historically, educators were perceived as the primary sources of knowledge, responsible for the transmission of societal wisdom to students (Akiba, 2011; Liu et al., 2023; Tenzin, 2023). However, this conventional model is being reshaped by technological advancements and demographic shifts, which demand that teachers not only disseminate knowledge but also foster an inclusive and dynamic learning environment (Kim & Maloney, 2020; Rock et al., 2016; Samsari et al., 2024). As digital technology makes information more accessible, the role of teachers increasingly involves facilitating learning, mentoring, and guiding students (Borko, 2004; Haleem et al., 2022).

This paradigm shifts towards a more student-centered learning approach underscores the necessity for teachers to develop a wide range of competencies, particularly in handling the diverse cultural backgrounds of their students. Multicultural competence, essential in today's global education systems, involves the ability to effectively engage with students from various cultural, ethnic, and linguistic backgrounds (Chun & Evans, 2016). Multicultural competency, a crucial element in global education systems, involves the ability to effectively engage with students from various cultural, ethnic, and linguistic backgrounds (Ganesan & Morales, 2024; Lee et al., 2023; Woolfolk, 1998). This set of skills ensures equitable learning opportunities for all students and is becoming increasingly vital given the global movement of populations and the resulting cultural diversity in classrooms (Darling-Hammond, 2012; Landsman & Lewis, 2023; Liddicoat, 2022; Parkhouse et al., 2019).

Multicultural competence encompasses three foundational dimensions: cultural awareness, cultural knowledge, and personal skills (Banks, 1994; Sue & Sue, 2008). Cultural awareness involves understanding and appreciating different cultural backgrounds, including one's own. It requires recognizing personal biases and stereotypes and developing sensitivity to how these perceptions affect interactions with others. Cultural knowledge extends this awareness to include a substantive understanding of the norms, values, and beliefs of different cultural groups, enabling educators to implement culturally responsive teaching strategies (Banks, 1994; Chiu & Hong, 2013; Sue & Sue, 2008). Personal skills related to multicultural competence, such as effective communication, empathy, conflict resolution, and adaptability, are crucial for creating supportive and inclusive classroom environments (Adili et al., 2023).

These skills help teachers navigate the complexities of multicultural classrooms, thereby enhancing student engagement and learning outcomes (Zakaria, 2000).

The increasing diversity of student populations presents challenges in preparing teachers who can effectively work in multicultural settings (Gay & Howard, 2000; Noordhoff & Kleinfeld, 1993; Wee et al., 2024). Research indicates that well-designed teacher training programs focusing on multicultural competencies can significantly improve educational experiences and outcomes for students from diverse backgrounds. Such training helps teachers recognize and overcome cultural barriers to learning, which is essential for closing achievement gaps among different student groups (Burnett et al., 2004; Keengwe, 2010). Furthermore, integrating the concepts of Banks (1994) and Sue & Sue (2008) with theories from learning, educational psychology, and counseling psychology, along with group process techniques, training methodologies, and technology, can create an innovative and comprehensive training program.

Additionally, the role of education in promoting social cohesion highlights the importance of multicultural training in teacher education. Teachers proficient in multicultural competencies can foster understanding and respect among students from various cultural backgrounds, contributing to a more inclusive society. This is particularly crucial in contexts where ethnic and cultural tensions need to be addressed. This aligns with the 20-Year National Strategy (2018-2037) of Thailand, which emphasizes the importance of preparing individuals with multicultural competencies to support cultural diversity (Office of the Secretary of the Board of Directors National Strategy, 2018). The 20-Year National Education Plan (2017-2036) of Thailand also underscores education that fosters citizenship and promotes living together in a multicultural society (Office of the Education Council Secretariat, 2017).

Furthermore, the role of education in fostering social cohesion underscores the importance of multicultural training in teacher education. Educators proficient in multicultural competencies can promote understanding and respect among students from different cultural backgrounds, thereby contributing to a more inclusive society. This is particularly vital in contexts where ethnic and cultural tensions need to be addressed.

The increasing diversity in Thai classrooms, coupled with the country's participation in global educational initiatives, underscores the critical need for multicultural competency training in pre-service teacher education. This study addresses this gap by evaluates the effectiveness of a Multicultural Competency Training Program (MCTP) designed for pre-service teachers at a major university in Thailand. By enhancing participants' cultural awareness,

cultural knowledge, and personal skills, the program aims to prepare them to manage diverse classrooms effectively. Comparing the multicultural competencies of teachers who have undergone the MCTP with those who have not, this research seeks to determine the impact of structured multicultural training on teacher preparedness.

Objectives

The study aimed to assess the effectiveness of the MCTP in boosting the multicultural competencies of pre-service teacher students in Thailand.

Method

1. Research Design and Sampling

This quasi-experimental study: the two-group pretest–posttest design, aimed to develop multicultural competency among pre-service teacher students in Thailand. We proposed two hypotheses:

Hypothesis 1: Multicultural competence of the intervention group after participating in the Multicultural Competency Training Program (MCTP) and at follow-up will be higher than before the experiment.

Hypothesis 2: Multicultural competence of the intervention group after participating in the MCTP and at follow-up will be higher than that of the control group.

The sample size was determined using the G* Power program. The following data was used to calculate the desired sample size: test family: F tests, statistical test: ANOVA: repeated measures, within-between interaction, effect size: .27, power Analysis: .80 (Cohen, 1977), α : .05, number of measurements: 3, resulting in 24 pre-service teacher students for each group. We considered the potential dropout rates and incomplete responses and added an additional 20%, which resulted in a sample size of 30 pre-service teacher students for each group (intervention and control), totaling 60 participants. Selection criteria included being over 18 years old, enrolled as third-year students at the Faculty of Education, Mahasarakham University, and fluent in the Thai language. Participants unwilling to participate were excluded from the study.

The study utilized purposive sampling to select areas and participants, who were then equally divided into two groups: control and intervention. To minimize extraneous variance, the matching pair method was employed, using similar scores.

The intervention group underwent the MCTC program, while the control group maintained their regular schedule. Both groups were assessed using the Multicultural Competency for Pre-Service Teacher Students Test (MCPSTST) to measure changes in multicultural competency before the intervention, immediately afterward, and at one month follow-up.

2. Intervention

Following approval from the Institutional Review Board of Khon Kaen University (protocol code HE653150 and dated 9 August 2022), we initiated the study by meeting with participants to explain the objectives, procedures, and protections for participant rights. All participants provided signed informed consent before participating. The Multicultural Competency Training Program (MCTP) covered three core areas of multicultural competency: cultural awareness, cultural knowledge, and personal skills (Banks, 1994; Sue & Sue, 2008).

The intervention consisted of 14 sessions held over seven weeks, with meetings on Mondays and Fridays. Each session lasted 120 minutes, focusing on interactive and participatory learning methods such as lectures, small group discussions, whole-class practices, case studies, games, role plays, and other group activities (Benuto et al., 2018). The training encompassed 12 topics, distributed as follows: the first were orientation and relationship building, two through four sessions were dedicated to developing cultural awareness, sessions five through seven targeted the enhancement of cultural knowledge, the remaining sessions from eight to thirteen were focused on building personal skills, and fourteen sessions were post-orientation.

Learning outcomes that are designed to build upon one another, ensuring a holistic development of multicultural competency necessary for effective and inclusive teaching. By the end of the program, participants are expected to have a comprehensive toolkit of skills and knowledge to bring into their professional practice, enabling them to contribute positively to the diverse educational landscapes they will encounter.

This comprehensive approach to multicultural training, underpinned by robust psychological, cultural, and educational theories, highlights the importance of structured intervention in developing effective multicultural competency among pre-service teachers.

3. Research Instrument

3.1 Multicultural Competency for Pre-Service Teacher Students Test (MCPSTST)

The MCPSTST, developed by researchers and based on frameworks by Banks (1994) and Sue & Sue (2008), measures the multicultural competency of pre-service teacher students

in Thailand. The instrument comprises 104 items across three components: cultural awareness, cultural knowledge, and personal skills. Responses are recorded on a five-point Likert scale ranging from 0 (least) to 5 (very high). The scores are categorized into five levels: lowest (1.00-1.50), low (1.51-2.50), moderate (2.51-3.50), high (3.51-4.50), and highest (4.51-5.00) (Best, 1981). The reliability of the scale is confirmed with a Cronbach's alpha of 0.972.

3.2 The Multicultural Competency Training Program (MCTP)

This program was designed by researchers to enhance the multicultural competencies of pre-service teachers in Thailand, drawing upon a comprehensive literature review. The program covers all three multicultural competency elements: cultural awareness, cultural knowledge, and personal skills. The 14-day training program includes lectures, small group sessions, whole-class practices, case studies, role-plays, group activities, and discussions. Five experts specializing in cultural psychology, training psychology, and evaluation and measurement validated the content of the program. Each activity within the program was assessed for appropriateness of content of the program and scored between 1.00 and 5.00, the mean scores were divided into appropriateness five levels: lowest (1.00-1.50), low (1.51-2.50), moderate (2.51-3.50), high (3.51-4.50) and highest (4.51-5.00) (Best, 1981), with average score of each item was between 4.20 – 4.73 and the average of every activity was at 4.46, indicating a high level of training program appropriateness.

3.3 Sociodemographic Variables

The study collected socioeconomic characteristics through eight items, incorporating both multiple-choice and open-ended questions. These variables included gender, age, religion, field of study, year of study, region of institution, and experience traveling abroad, with the tool development informed by literature reviews.

4. Data analysis

Data were analyzed using the Statistical Package for the Social Sciences (SPSS), Version 21 (SPSS, Chicago, IL). Participant characteristics were described using descriptive statistics, including frequencies (percentages), means and standard deviation. Repeated measures ANOVA analyses tested the associations between multicultural competency among pre-service teacher students at different times. The analysis confirmed that the assumptions of ANOVA (normality and homogeneity of variance) were met. The primary objective of the analysis was to determine if there were significant differences in outcomes before, immediately after, and one month following the intervention between the intervention and control groups. Statistical significance was established at a p-value .05.

Results

The study involved 30 pre-service teacher students in each of the intervention and control groups. Among these participants, 46.67% ($n = 14$) in the intervention group were female with an average age of 20 years ($SD \pm 0.50$). Similarly, 46.67% in the control group were male, also with an average age of 20 years ($SD \pm 0.50$). All participants were Buddhist, in their third year of study, and from institutions in the Northeast region of Thailand. In terms of academic disciplines, 16.67% ($n = 5$) in both groups were studying Thai and English. Regarding international exposure, 50% ($n = 15$) of the intervention group participants had experience traveling abroad, with the remaining 50% having no such experience. In contrast, 53.33% ($n = 16$) of the control group had traveled abroad.

As depicted in Figure 1, the intervention group's multicultural competency initially scored at a low level (1.80 ± 0.15) but showed significant improvement in the post-test (3.93 ± 0.38) and follow-up (4.20 ± 0.20). In contrast, the control group started at a similarly low level (1.74 ± 0.21) and only reached a moderate level in both the post-test (2.62 ± 0.14) and follow-up (2.86 ± 0.11).

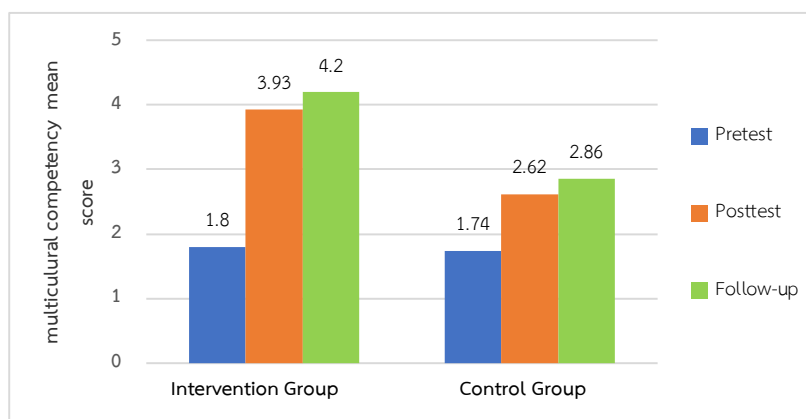


Figure 1. Distribution of multicultural competency among Pre-Service Teacher Students of intervention group and control group

As presented in Table 1, significant differences were observed in multicultural competency scores across the three time points: before the intervention, immediately after, and one month post-intervention. The mean scores of the intervention group were significantly higher than those of the control group ($p < .001$). Moreover, a statistically significant interaction effect was found between the intervention method and time ($p < .001$), indicating a sustained and progressive improvement in the intervention group compared to the control group over the study period.

Table 1. The variance of multicultural competency of pre-service teacher students between the Intervention Method & the Duration of the Experiment

Source	Sum of Squares	df	Mean Square	F	p
Between Groups					
Intervention – Control	1435.985	1	1435.985	25792.163	<.001
Error	1.615	29	.056		
Within Groups					
Duration of Experiment	97.336	1.709	56.946	2260.801	<.001
Duration of Experiment X Method	15.946	1.641	9.715	488.862	<.001
Error	.946	47.600	.020		

Discussion

The Multicultural Competency Training Program (MCTP) has proven to be significantly effective in enhancing the multicultural competencies of pre-service teachers in Thailand, thereby supporting the initial research hypothesis. The intervention group, which participated in the MCTP, demonstrated marked improvements in their ability to comprehend and engage across various cultural contexts when compared to the control group. This enhancement is vital in today's global educational settings, where teachers frequently encounter culturally diverse student populations. The observed improvements in the intervention group's multicultural competencies can be attributed to the comprehensive design of the MCTP, which included targeted training on cultural awareness, cultural knowledge, and personal intercultural skills. Diaz-Rico (2019) underscores the necessity of extensive cultural competency training within teacher education programs to adequately prepare educators for the diversity present in contemporary classrooms. This assertion aligns with recent studies indicating that such programs are effective in enhancing pre-service teachers' readiness for diverse educational environments (Jones & Lee, 2021). The significant post-intervention improvements underscore the program's efficacy in bridging the gap between theoretical knowledge and practical application, which is essential for the professional preparedness of pre-service teachers.

The first hypothesis anticipated that the multicultural competence of the intervention group, both immediately after participating in the MCTP and at a follow-up, would be higher than prior to the intervention. This hypothesis was confirmed by the substantial improvements observed post-intervention. In contrast, the control group, which did not

undergo the MCTP, exhibited no significant advancements in multicultural competency. This disparity highlights the importance of structured and systematic multicultural education within teacher training programs. Gay (2018) notes that without intentional and structured interventions like the MCTP, pre-service teachers may not develop the necessary depth of cultural competencies required in today's diverse educational landscapes. The clear difference in post-training competencies between the two groups strongly advocates for the incorporation of formal multicultural training modules into teacher education curricula (Banks, 2015; Buysse et al., 2009; Buysse et al., 2012).

Recent studies corroborate the effectiveness of such structured interventions. Christensen and Knezek found that reducing bias through simulation experiences significantly enhances the effectiveness of culturally responsive teaching strategies (Christensen & Knezek, 2022). Additionally, Walker (2023) highlighted that pre-service teacher engaged in culturally responsive teaching practices demonstrated improved cultural competence and equity in education. Muñiz (2019) conducted a 50-state survey, revealing the critical need for culturally responsive teaching standards to ensure educators are well-prepared to address the cultural dynamics of their classrooms. Similarly, Caingcoy et al. (2022) emphasized the importance of culturally responsive teaching in enhancing teachers' ability to effectively manage classroom diversity. Implementing such programs can bridge cultural gaps that often hinder the educational process, thereby improving both teacher confidence and student performance.

The second hypothesis posited that the multicultural competence of the intervention group, both immediately after participating in the MCTP and at a follow-up, would be higher than that of the control group. This hypothesis was validated by the statistically significant improvement in the intervention group's competency scores compared to the control group. The MCTP's design is grounded in various theoretical principles. These frameworks facilitate the development of cultural awareness, knowledge of culture, and personal skills, and solution-focused approaches to address presenting issues. Recent studies have further supported the importance of integrating these frameworks into multicultural competency training. The development of culturally competent pre-service teachers is crucial for fostering inclusive classroom environments (Lewis Chiu et al., 2017). Such environments encourage discussions about cultural differences and help build authentic relationships with students from diverse backgrounds (Moore, 2021).

The integration of cultural competency, diversity, equity, and inclusion (CCDEI) standards into teacher education programs has been shown to enhance teachers' ability to

manage and appreciate classroom diversity (Professional Educator Standards Board, 2021). Furthermore, research underscores the impact of comprehensive cultural competence training, noting that such programs significantly improve teachers' cultural awareness and effectiveness in diverse educational settings (Jones & Lee, 2021). The inclusion of role plays, case discussions, and lectures in the MCTP is particularly effective in developing practical counseling skills among academic advisors. These activities provide a dynamic platform for pre-service teachers to engage in real-world scenarios, thereby enhancing their readiness to address cultural issues in classrooms (Buysse et al., 2009). By incorporating these diverse theoretical approaches and practical activities, the MCTP ensures that academic advisors are well-equipped to foster inclusive and supportive learning environments.

Implications of the Findings

The findings from the Multicultural Competency Training Program (MCTP) hold significant implications for educational practice and policy, particularly within the realm of teacher education. As demonstrated in the study, the enhanced multicultural competencies among pre-service teachers not only prepare them to effectively manage diverse classrooms but also contribute to broader educational goals such as equity and inclusivity.

For educational practice, the successful implementation and outcomes of the MCTP suggest that integrating structured multicultural competency training within pre-service teacher programs can profoundly impact the effectiveness of future educators in diverse environments. Teachers equipped with these competencies are better prepared to address the varied cultural backgrounds and needs of their students, thereby promoting a more inclusive and supportive learning environment. This aligns with the assertion by Banks (2015) that multicultural education is not merely an additive to teacher education but a necessary component to address the realities of contemporary classrooms.

Moreover, the significant improvements in cultural awareness, cultural knowledge, and interpersonal skills among the intervention group underscore the need for such programs to be comprehensive, touching on various aspects of cultural competency. Teacher education should encompass not only the theoretical understanding of cultural diversity but also practical strategies for managing it. This can include training on cultural sensitivity, conflict resolution, and effective communication across cultural boundaries, all of which are essential for fostering an environment conducive to learning for all students.

From a policy perspective, the study's findings advocate for the systemic integration of multicultural training programs like the MCTP across all teacher education in-situations.

Policymakers should consider mandating multicultural competency as a core component of teacher certification and ongoing professional development. Such policy initiatives could ensure that all teachers have the requisite skills to manage and leverage the benefits of classroom diversity.

Furthermore, policies supporting such programs could encourage higher education institutions to allocate resources towards developing and updating their multi-cultural training curricula. This could involve collaborations with multicultural experts to ensure that the training remains relevant and reflective of current social dynamics and educational needs.

Strategically, the implications of integrating a program like the MCTP into teacher education extend beyond the classroom. Educators trained in multicultural competencies can act as catalysts for change, promoting social cohesion and intercultural understanding among the broader community. As society becomes increasingly globalized, the ability of teachers to model and teach inclusive practices and cultural sensitivity becomes more critical. Therefore, by investing in comprehensive multicultural training for teachers, educational systems not only enhance the quality of education but also contribute to the building of more harmonious and integrated communities.

Limitations of the Study

The study assessing the effectiveness of the MCTP for pre-service teachers in Thailand presents several limitations that are crucial to consider when interpreting the findings and considering their applicability. The sample size, while adequate for initial findings, is relatively small and restricted to a single educational institution. This limitation may impact the generalizability of the results to other contexts or regions, potentially limiting the applicability of the findings to different educational settings or cultural backgrounds. Moreover, the duration of the follow-up period only one month after the intervention may not be sufficient to determine the long-term retention of the multicultural competencies acquired through the MCTP. Future studies could benefit from a longer follow-up duration to assess the enduring impact of the training.

Overall, while this study provides valuable insights into the potential benefits of multicultural training programs in teacher education, the limitations noted call for cautious

interpretation of the results and suggest areas for further research to validate and extend these findings.

Recommendations for Future Research of the Study

The findings from this study on the MCTP offer several directions for future re-search, aimed at enhancing the development of multicultural competencies in pre-service teachers and examining the broader impacts of such training.

Firstly, future studies should aim to assess the long-term effects of multicultural competency training. While the current study indicates significant improvements shortly after the intervention and one month later, it is crucial to determine whether these competencies are sustained over a longer period. Researchers could conduct longitudinal studies that follow participants for a year or more to evaluate the persistence of multicultural skills and their practical application in diverse classroom settings. This would provide deeper insights into the training's effectiveness and inform continuous improvement of the program. Additionally, extending the research to include diverse educational settings and geographical locations would help to validate and generalize the findings of this study. Comparing how multicultural training programs perform in different cultural contexts or across various educational levels could reveal essential insights into the adaptability and scalability of the MCTP. It would also address the limitation related to the small and region-specific sample size of the current study. Moreover, future research should explore the potential for enhancing the training modules within the MCTP. This could involve integrating technology-enhanced learning tools, such as virtual reality or online collaborative platforms, to simulate more realistic and diverse intercultural interactions. Researchers might also consider developing advanced modules that cover broader aspects of cultural competencies and multi-dimensional interventions on psychological or social factors, including implicit bias training and advanced conflict resolution techniques. The educational program should facilitate a comprehensive understanding of life course theory and the temporal experiences associated with individual life trajectories, particularly within the context of contemporary late modern societies.

Conclusions

This study conclusively demonstrates the effectiveness of the MCTP in enhancing the multicultural competencies of pre-service teacher students in Thailand. The intervention group, which received structured training in cultural awareness, cultural knowledge, and

personal skills, showed significant improvements compared to the control group. The findings underscore the necessity of integrating targeted multicultural training into teacher education programs to prepare educators for diverse classrooms. The substantial improvements observed highlight the MCTP's role in fostering inclusive teaching practices, which likely enhance educational outcomes by creating a learning environment that respects and values diversity. Statistical analysis confirmed the program's impact, showing significant differences in multicultural competency between the intervention and control groups post-training and at one-month follow-up. These results validate the MCTP's effectiveness and support its continued use and expansion within teacher training curricula. Given the global nature of contemporary education, prioritizing multicultural competence in teacher training is imperative. This study confirms that structured interventions like the MCTP can effectively develop these competencies. Future research should explore the long-term impacts, scalability, and implementation across different cultural settings to further refine the program components. In conclusion, comprehensive multicultural training is essential in modern teacher education, preparing teachers to contribute positively to diverse educational landscapes.

Acknowledgments

We thank the participants of the study for their valuable contributions.

Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Khon Kaen University (protocol code HE653150 and dated 9 August 2022).

References

- Adili, B., Ameti, M., & Zenki-Dalipi, A. (2023). Teachers' attitudes towards desirable intercultural competences necessary for working in multicultural classrooms. *Journal of Educational Technology and Instruction*, 2(2), 32-46. <https://ijeti-edu.org/index.php/ijeti/article/view/74>
- Akiba, M. (2011). Identifying program characteristics for preparing pre-service teachers for diversity. *Teachers College Record*, 113(3), 658-697. <https://doi.org/10.1177/0161468111111300304>

- Banks, J. (1994). *An introduction to multicultural education*. Allyn & Bacon.
- Banks, J. A. (2015). Multicultural education. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 18–21). Elsevier.
<https://doi.org/10.1016/B978-0-08-097086-8.92097-X>
- Benuto, L. T., Casas, J., & O'Donohue, W. T. (2018). Training culturally competent psychologists: A systematic review of the training outcome literature. *Training and Education in Professional Psychology, 12*(3), 125-134.
<https://doi.org/10.1037/tep0000190>
- Best, J. W. (1981). *Research in Education*. Prentice-Hall.
- Professional Educator Standards Board. (2021). *Cultural competency, diversity, equity, and inclusion (CCDEI) standards*. <https://www.pesb.wa.gov/innovation-policy/ccdei/>
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher, 33*(8), 3-15. <https://doi.org/10.3102/0013189X033008003>
- Burnett, J. A., Hamel, D., & Long, L. L. (2004). Service learning in graduate counselor education: Developing multicultural counseling competency. *Journal of Multicultural Counseling and Development, 32*(3), 180-191. <https://doi.org/10.1002/j.2161-1912.2004.tb00370.x>
- Buyse, V., Winton, P. J., & Rous, B. (2009). Reaching consensus on a definition of professional development for the early childhood field. *Topics in early childhood special education, 28*(4), 235-243.
- Buyse, V., Winton, P. J., Rous, B. S., Epstein, D. J., & Lim, C.-I. (2012). Evidence-Based Practice: Foundation for CONNECT 5-Step Learning Cycle[™] in Professional Development. *Zero to Three, 32*, 25-29.
- Caingcoy, M., Vivian Irish, L., Lorenzo, M., Ramirez, I. A., Libertad, C., Pabiona, R., Ruffie, M., & Mier, M. R. (2022). Assessing practice teachers' culturally responsive teaching: The role of gender and degree programs in competence development. *IAFOR Journal of Cultural Studies, 7*, 21-35. <https://doi.org/10.22492/ijcs.7.1.02>
- Chiu, C. Y., & Hong, Y. Y. (2013). *Social psychology of culture*. Psychology Press.
- Christensen, R., & Knezek, G. (2022). Impacting Culturally Responsive Teaching Strategies by Decreasing Bias Through Simulation Experiences. *Excellence and Innovation in Learning and Teaching, 7*(2), 39-56. <https://doi.org/10.3280/exioa2-2022oa15077>

- Chun, E., & Evans, A. (2016). *Rethinking cultural competence in higher education: An ecological framework for student development (ASHE higher education report, 42[4])*. John Wiley & Sons.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press.
- Darling-Hammond, L. (2012). *Powerful Teacher Education: Lessons from Exemplary Programs*. John Wiley & Sons.
- Diaz-Rico, L. T. (2019). *Multicultural education in a pluralistic society*. Pearson.
- Ganesan, U., & Morales, A. R. (2024). A science teacher's experiences when fostering intercultural competence among students in multilingual classrooms: a narrative study. *Cultural Studies of Science Education*, 19(1), 189-208. <https://doi.org/10.1007/s11422-023-10206-y>
- Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.
- Gay, G., & Howard, T. C. (2000). Multicultural teacher education for the 21st century. *The Teacher Educator*, 36(1), 1-16. <https://doi.org/10.1080/08878730009555246>
- Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, 3, 275-285.
- Ingersoll, R., Merrill, E., Stuckey, D., Collins, G., & Harrison, B. (2021). The Demographic Transformation of the Teaching Force in the United States. *Education Sciences*, 11(5), 234. <https://www.mdpi.com/2227-7102/11/5/234>
- Jones, J. M., & Lee, L. H. (2021). Multicultural competency building: A multi-year study of trainee self-perceptions of cultural competence. *Contemporary School Psychology*, 25(3), 288-298. <https://doi.org/10.1007/s40688-020-00339-0>
- Keengwe, J. (2010). Fostering cross cultural competence in preservice teachers through multicultural education experiences. *Early Childhood Education Journal*, 38(3), 197-204. <https://doi.org/10.1007/s10643-010-0401-5>
- Kim, J., & Maloney, E. J. (2020). *Learning innovation and the future of higher education*. JHU Press.
- Landsman, J., & Lewis, C. W. (2023). *White teachers/diverse classrooms: Creating inclusive schools, building on students' diversity, and providing true educational equity*. Taylor & Francis.

- Lee, S., Cha, Y. K., & Ham, S. H. (2023). The global institutionalization of multicultural education as an academic discourse. *Societies*, 13(8), 191.
<https://www.mdpi.com/2075-4698/13/8/191>
- Lewis Chiu, C., Sayman, D., Carrero, K. M., Gibbon, T., Zolkoski, S. M., & Lusk, M. E. (2017). Developing culturally competent preservice teachers. *Multicultural Perspectives*, 19(1), 47-52. <https://doi.org/10.1080/15210960.2017.1267515>
- Liddicoat, A. J. (2022). Teacher education for diversity: afterword. *Language and Education*, 36(2), 188-194. <https://doi.org/10.1080/09500782.2021.1981929>
- Liu, X., Gao, W., & Chen, L. (2023). Does pre-service teacher preparation affect students' academic performance? evidence from China. *Education Sciences*, 13(1), 69.
- Moore, S. (2021). *Cultural Competence in Education*.
<https://www.teachhub.com/professional-development/2021/06/cultural-competence-in-education/>
- Muñiz, J. (2019). Culturally responsive teaching: A 50-state survey of teaching sandards. *New America*.
- Noordhoff, K., & Kleinfeld, J. (1993). Preparing teachers for multicultural classrooms. *Teaching and Teacher Education*, 9(1), 27-39.
- Parkhouse, H., Lu, C. Y., & Massaro, V. R. (2019). Multicultural education professional development: A review of the literature. *Review of Educational Research*, 89(3), 416-458. <https://doi.org/10.3102/0034654319840359>
- Reinius, H., Kaukinen, I., Korhonen, T., Juuti, K., & Hakkarainen, K. (2022). Teachers as transformative agents in changing school culture. *Teaching and Teacher Education*, 120, 103888. <https://doi.org/10.1016/j.tate.2022.103888>
- Rock, M. L., Spooner, F., Nagro, S., Vasquez, E., Dunn, C., Leko, M., Luckner, J., Bausch, M., Donehower, C., & Jones, J. L. (2016). 21st century change drivers: Considerations for constructing transformative models of special education teacher development. *Teacher Education and Special Education*, 39(2), 98-120.
- Samsari, E., Palaiologou, N., & Nikolaou, G. (2024). Exploring narratives of teachers working with culturally diverse students: Any insights about Inclusion? *Societies*, 14(4), 55.
<https://www.mdpi.com/2075-4698/14/4/55>
- Sue, D. W., & Sue, D. (2008). *Counseling the culturally diverse: Theory and practice*. John Wiley & Sons.

- Tenzin, W. (2023). *The journey of Bhutanese school curriculum: Evolution towards competency-based education* [Curriculum Specialist and Dean, DCPD, MoE (Retired)]. <https://doi.org/10.13140/RG.2.2.27993.90728>
- Walker, A. (2023). Transformative potential of culturally responsive teaching: Examining preservice teachers' collaboration practices centering refugee youth. *Education Sciences*, 13(6), 621. <https://www.mdpi.com/2227-7102/13/6/621>
- Wee, S. J., Son, M., Ly-Hoang, K., & Zambrano, L. (2024). Culture is where I come from: An analysis of cultural competence of student teachers of color in early childhood education. *Journal of Early Childhood Teacher Education*, 45(2), 157-176. <https://doi.org/10.1080/10901027.2023.2258369>
- Woolfolk, A. E. (1998). *Educational Psychology*. Allyn & Bacon.
- Zakaria, N. (2000). The effects of cross- cultural training on the acculturation process of the global workforce. *International Journal of Manpower*, 21, 492-510. <https://doi.org/10.1108/01437720010377837>

Translate Thai References

- Office of the Education Council Secretariat. (2017). The National Education Plan 2017-2036. Prikwan Graphic. (in Thai)
- Office of the Secretary of the Board of Directors National Strategy. (2018). National Strategy: 2018–2037 (Summary). Office of the National Economic and Social Development Board. (in Thai)

การพัฒนาแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี ระดับชั้นประถมศึกษาปีที่ 6

The Development of a Critical Thinking Skills Measurement Tool in Science and Technology for Grade 6 Students

ปิยาภรณ์ ประถมวงษ์^{1*} และ ประกฤติยา ทักชีโน²

Piyaporn Prathomwong^{1*} and Prakittiya Tuksino²

(Received: April 19, 2024; Revised: August 26, 2024; Accepted: August 27, 2024)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ 1) ศึกษาองค์ประกอบ และพฤติกรรมบ่งชี้ทักษะการคิดอย่างมี
วิจารณญาณของนักเรียนชั้นประถมศึกษาปีที่ 6 2) พัฒนาและตรวจสอบคุณภาพแบบวัดทักษะการคิดอย่างมี
วิจารณญาณของนักเรียนชั้นประถมศึกษาปีที่ 6 และ 3) พัฒนาจุดตัดคะแนนทักษะการคิดอย่างมีวิจารณญาณ
ของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยใช้การวิจัยเชิงออกแบบ กลุ่มตัวอย่างที่ใช้ในการเก็บข้อมูลจริง
เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ที่กำหนดโดยวิธีการสุ่มแบบหลายขั้นตอน จำนวน 500 คน เครื่องมือที่ใช้
ในการวิจัยครั้งนี้เป็นแบบวัดทักษะการคิดอย่างมีวิจารณญาณ ชนิดเลือกตอบ 4 ตัวเลือก ตอบถูกได้ 1 คะแนน
ตอบผิดได้ 0 คะแนน การวิเคราะห์ข้อมูลตามทฤษฎีการทดสอบแบบดั้งเดิม และทฤษฎีการตอบสนองข้อสอบ
ผลการวิจัยพบว่า

1. การสังเคราะห์องค์ประกอบในการวัดทักษะการคิดอย่างมีวิจารณญาณ ได้องค์ประกอบของทักษะ
การคิดอย่างมีวิจารณญาณ 5 องค์ประกอบ ได้แก่ 1) ด้านการระบุปัญหา 2) ด้านการระบุสมมติฐาน 3) ด้าน
การรวบรวมข้อมูล 4) ด้านการวิเคราะห์ข้อมูล และ 5) ด้านการลงข้อสรุป

2. คุณภาพของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี
ระดับชั้นประถมศึกษาปีที่ 6 พบว่าการตรวจสอบคุณภาพก่อนนำไปใช้ด้วยทฤษฎีการทดสอบแบบดั้งเดิม
พบว่ามีค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับจุดประสงค์ เท่ากับ 1 จำนวน 40 ข้อ และเมื่อตรวจสอบ
คุณภาพหลังจากนำไปใช้ตามทฤษฎีการตอบสนองข้อสอบ พบว่า ค่าพารามิเตอร์อำนาจจำแนก (a_i) มีค่าอยู่
ระหว่าง 0.39 – 5.03 ค่าพารามิเตอร์ความยากของข้อสอบ (b_i) มีค่าอยู่ระหว่าง -1.29 - 5.14 และสัมประสิทธิ์
การเดา (c_i) มีค่าอยู่ระหว่าง 0.13 - 0.27

¹ นักศึกษาปริญญาโท สาขาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² ผู้ช่วยศาสตราจารย์ สาขาการวัดและประเมินผลการศึกษา 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: piyaporn.p@kkumail.com

3. ผลการพัฒนาจุดตัดคะแนนจากการพิจารณาความกลมกลืนของโมเดล พบว่า โมเดลที่มีจำนวนกลุ่มแฝง 3 กลุ่ม ได้แก่ กลุ่มสูง กลุ่มปานกลาง และกลุ่มต่ำ เป็นโมเดลที่มีความเหมาะสมและมีคุณสมบัติตรงตามเกณฑ์การพิจารณามากที่สุด (Likelihood = -11858.198, $df = 378$, AIC = 23960.395, BIC = 24474.578, Adj.BIC = 24087.342, $E_k = 0.846$) โดยมีคะแนนจุดตัด เท่ากับ 23.7 (แบ่งกลุ่มสูงและกลุ่มปานกลางออกจากกัน) และ 16.1 (แบ่งกลุ่มปานกลางและกลุ่มต่ำออกจากกัน)

คำสำคัญ ทักษะการคิดอย่างมีวิจารณญาณ การวิเคราะห์กลุ่มแฝง แบบวัด

Abstract

This research aimed to: 1) identify the components and behavioral indicators of critical thinking skills in sixth-grade students; 2) develop and examine the quality of a critical thinking skills assessment tool for sixth-grade students; and 3) establish cutoff scores for critical thinking skills using a design-based research approach. The sample consisted of 500 sixth-grade students selected through multi-stage random sampling for the main phase of data collection. The instrument was a multiple-choice test with four options per item; a correct response received 1 point, while an incorrect response received 0 points. Data were analyzed using classical test theory (CTT) and item response theory (IRT). The findings as follows:

1. The synthesis of critical thinking components resulted in five key elements: 1) problem identification, 2) hypothesis formulation, 3) data collection, 4) data analysis, and 5) conclusion drawing.

2. The quality assessment of the instrument revealed that the item-objective congruence index was 1.00 for all 40 items. The quality after using with item response theory, it was found that the discrimination parameter (a_i) was between 0.39 and 5.03, the difficulty parameter (b_i) was between -1.29 and 5.14, and the guessing parameter (c_i) was between 0.13 and 0.27.

3. Model fit indices suggested that a three-class latent model—comprising high, medium, and low critical thinking skill groups—was the most appropriate (Likelihood = -11858.198, $df = 378$, AIC = 23960.395, BIC = 24474.578, Adj.BIC = 24087.342, $E_k = 0.846$). The cutoff scores were 23.7 (between high and medium groups) and 16.1 (between medium and low groups).

Keywords: critical thinking skills, latent class analysis, measurement tool

บทนำ

วิทยาศาสตร์และเทคโนโลยีเข้ามามีบทบาทในสังคม การดำเนินชีวิตและข้าวของเครื่องใช้ต่าง ๆ ล้วนเป็นผลที่ขึ้นเกิดจากความรู้ด้านวิทยาศาสตร์ผสมผสานกับศาสตร์อื่นๆ ความรู้ด้านวิทยาศาสตร์ช่วยให้เกิดความรู้และความเข้าใจในปรากฏการณ์ธรรมชาติมากมาย ส่งผลทำให้มีการพัฒนาเทคโนโลยีต่าง ๆ ขึ้นเป็นอย่างมาก ทุกคนจึงจำเป็นต้องได้รับการพัฒนาให้รู้วิทยาศาสตร์ เพื่อนำมาใช้ในการพัฒนาคุณภาพชีวิตที่ดีขึ้น ช่วยให้มีความรู้ความเข้าใจที่ถูกต้องเกี่ยวกับการใช้ประโยชน์การดูแลรักษา และพัฒนาสิ่งแวดล้อม โดยการจัดกิจกรรมการเรียนรู้จึงต้องจัดให้ผู้เรียน คิดและเรียนรู้ได้ด้วยตนเอง Kammanee (2018) ได้กล่าวว่า การสอนคิดเป็นเรื่องสำคัญในการจัดการศึกษาให้มีคุณภาพสูง เพื่อพัฒนาผู้เรียนทั้งด้าน สติปัญญา คุณธรรมและการเป็นพลเมืองที่ดีของประเทศ กระบวนการคิดเป็นตัวบ่งชี้สำคัญที่ส่งผลให้ผู้เรียนเกิดการเรียนรู้ที่มีความหมาย สามารถนำความรู้ไปประกอบการตัดสินใจในการเผชิญสถานการณ์ได้อย่างมีประสิทธิภาพ โดยเฉพาะอย่างยิ่งการคิดอย่างมีวิจารณญาณ (critical thinking) เป็นการคิดวิเคราะห์ให้ไตร่ตรองประมวลข้อมูลปัญหาเรื่องราวต่างๆ ก่อนที่จะตัดสินใจเชื่อหรือ กระทำสิ่งต่างๆ เพื่อให้ได้ความคิดที่รอบคอบ สมเหตุสมผล เพื่อเป็นแนวทางในการเลือกตัดสินใจให้ดีที่สุด (Kammanee, 2018)

การคิดอย่างมีวิจารณญาณของนักเรียนเป็นสิ่งสำคัญอย่างยิ่งต่อการจัดการศึกษาไทยโดยเฉพาะกลุ่มสาระการเรียนรู้วิทยาศาสตร์และเทคโนโลยีซึ่งสอดคล้องกับแนวคิดของ Schraer & Stoltze (Schraer & Stoltze, 1968 as cited in Wonggom, 2009) ได้กล่าวว่าในการเรียนการสอนวิทยาศาสตร์นักเรียนจะสามารถพัฒนาความรู้ได้ก็ต่อเมื่อได้รับการพัฒนาการคิดอย่างมีวิจารณญาณซึ่งจัดว่าเป็นทักษะการคิดขั้นสูง โดยมีนักการศึกษาหลายท่านได้สร้างเครื่องมือที่ใช้วัดทักษะการคิดอย่างมีวิจารณญาณไว้อย่างหลากหลาย เช่น Watson & Glaser (1980) ได้สร้างแบบสอบที่ใช้วัดทักษะการคิดอย่างมีวิจารณญาณ ส่วน Cornell Critical Thinking Test พัฒนาโดย Ennis & Millman (1985) พัฒนาขึ้นมาโดยยึดทฤษฎีของ Ennis เป็นหลัก ในปัจจุบันได้มีการนำแบบทดสอบประเมินทักษะการคิดมาใช้เป็นส่วนหนึ่งของการรับสมัครการประมวลผลสำหรับการเข้าสู่ระดับปริญญาตรีหลักสูตรบางส่วนของมหาวิทยาลัยเคมบริดจ์ มหาวิทยาลัยออกฟอร์ด และมหาวิทยาลัยคอลเลจลอนดอน โดยแบบทดสอบแบ่งเป็น 2 ส่วน ได้แก่ ส่วนที่ 1 เป็นข้อสอบแบบปรนัยจำนวน 50 ข้อ ทดสอบการแก้ปัญหาและทักษะการคิดอย่างมีวิจารณญาณ ส่วนที่ 2 เป็นคำถามเรียงความหนึ่งคำถามจากตัวเลือกสี่ข้อ เป็นการทดสอบความสามารถในการจัดระเบียบความคิดในลักษณะที่ชัดเจนและรัดกุม และสื่อสารอย่างมีประสิทธิภาพเป็นลายลักษณ์อักษร

ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง 2560) ในรายวิชาวิทยาศาสตร์และเทคโนโลยี ได้มีเป้าหมายของการจัดการเรียนการสอนวิทยาศาสตร์เพื่อพัฒนาทักษะที่จำเป็นต่อผู้เรียน ได้แก่ ทักษะกระบวนการทางวิทยาศาสตร์ และทักษะแห่งศตวรรษที่ 21 ซึ่งทักษะการคิดอย่างมีวิจารณญาณเป็นทักษะแห่งศตวรรษที่ 21 ที่นักเรียนจะต้องได้รับการพัฒนา เมื่อมีการจัดการเรียนการสอนเรียบร้อยแล้ว จะต้องมีการวัดการประเมินที่เหมาะสม และต้องมีการกำหนดเกณฑ์การประเมินให้สอดคล้องกับเกณฑ์การประเมินนั้น ๆ โดยการกำหนดจุดตัดหรือการกำหนดเกณฑ์การประเมินนั้นสามารถทำได้หลายวิธี เช่น การกำหนดเกณฑ์มาตรฐาน (Millman, 1973) การกำหนดจุดตัดด้วยวิธีการตัดสินใจโดยการพิจารณา

จากความสามารถต่ำสุดของผู้สอบ (Glass, 1978) ฯลฯ เนื่องจากปัจจุบันนี้ถึงแม้ว่าประเทศไทยจะมีการจัดการเรียนการสอนเพื่อพัฒนาทักษะการคิดอย่างมีวิจารณญาณมาระยะหนึ่งแล้ว การจะกำหนดเกณฑ์ของทักษะการคิดอย่างมีวิจารณญาณ โดยใช้กลุ่มรูขั้วนั้นอาจจะไม่เหมาะสมเท่าที่ควร เพราะยังไม่มีหลักฐานที่แสดงเด่นชัดว่าบุคคลใดมีทักษะการคิดอย่างมีวิจารณญาณมากหรือน้อยเพียงใด การวิจัยในครั้งนี้จึงเลือกใช้วิธีการกำหนดจุดตัดโดยวิธีการวิเคราะห์กลุ่มแฝง (Latent Class Analysis) มาเป็นตัวช่วยในการกำหนดเกณฑ์ในการประเมินทักษะการคิดอย่างมีวิจารณญาณของนักเรียนชั้นประถมศึกษาปีที่ 6 เนื่องจากการวิเคราะห์กลุ่มแฝงสามารถบ่งชี้ลักษณะแฝงของกลุ่มย่อยทักษะการคิดอย่างมีวิจารณญาณที่เป็นอยู่ และยังให้ข้อมูลว่าแต่ละกลุ่มย่อยนั้นมีระดับทักษะการคิดอย่างมีวิจารณญาณอยู่ในระดับใด ซึ่งข้อมูลที่ได้นี้จะตรงกับลักษณะกลุ่มแฝงที่เป็นอยู่ของนักเรียนชั้นประถมศึกษาปีที่ 6

ด้วยเหตุผลที่กล่าวมาข้างต้น ผู้วิจัยจึงมีความสนใจที่จะพัฒนาเครื่องมือวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 โดยมีการตรวจสอบคุณภาพของแบบวัดทักษะการคิดอย่างมีวิจารณญาณตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ประยุกต์ใช้ร่วมกับทฤษฎีการตอบสนองข้อสอบ (IRT) โดยใช้วิธีการวิเคราะห์กลุ่มแฝงในการกำหนดจุดตัดคะแนน

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาองค์ประกอบ และพฤติกรรมบ่งชี้ทักษะการคิดอย่างมีวิจารณญาณของนักเรียนชั้นประถมศึกษาปีที่ 6
2. เพื่อพัฒนาและตรวจสอบคุณภาพแบบวัดทักษะการคิดอย่างมีวิจารณญาณของนักเรียนชั้นประถมศึกษาปีที่ 6
3. เพื่อพัฒนาจุดตัดคะแนนทักษะการคิดอย่างมีวิจารณญาณของนักเรียนชั้นประถมศึกษาปีที่ 6

วิธีดำเนินการวิจัย

งานวิจัยผ่านการรับรองจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยขอนแก่น วันที่ 18 ธันวาคม พ.ศ.2566 หมายเลขใบรับรอง เลขที่ : HE663364 ให้สามารถเก็บข้อมูลได้กับกลุ่มเป้าหมายและตัวอย่างที่ระบุ โดยการวิจัยในครั้งนี้ประยุกต์ใช้การวิจัยอิงการออกแบบเพื่อพัฒนาแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6 วิธีการดำเนินการวิจัยแบ่งออกเป็น 3 ระยะ ดังต่อไปนี้

ระยะที่ 1 การวิเคราะห์และสำรวจ

เป็นการดำเนินการเพื่อศึกษาข้อมูล และรวบรวมข้อมูลที่เกี่ยวข้องกับการวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6 ประกอบไปด้วยขั้นตอนดังต่อไปนี้

- 1.1 ขั้นตอนการวิเคราะห์องค์ประกอบ โดยขั้นนี้ผู้วิจัยได้กำหนดจุดมุ่งหมายในการสร้างและพัฒนาแบบวัดทักษะการคิดอย่างมีวิจารณญาณสำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 จากนั้นศึกษาเอกสารและ

วรรณกรรมที่เกี่ยวข้องกับทักษะการคิดอย่างมีวิจารณญาณ และสังเคราะห์องค์ประกอบของทักษะการคิดอย่างมีวิจารณญาณ โดยใช้ตารางสังเคราะห์องค์ประกอบ ได้เป็น 5 องค์ประกอบ คือ 1) ด้านการระบุปัญหา 2) ด้านการระบุสมมติฐาน 3) ด้านการรวบรวมข้อมูล 4) ด้านการวิเคราะห์ข้อมูล และ 5) ด้านการลงข้อสรุป แล้วจึงนำตารางการสังเคราะห์องค์ประกอบและตัวชี้วัดที่ได้ไปให้ผู้เชี่ยวชาญจำนวน 5 ท่าน ตรวจสอบความตรงเชิงทฤษฎี (construct validity)

1.2 การกำหนดผู้ใช้ (user) ผู้ใช้ในการวิจัยครั้งนี้ คือ ครูผู้สอนในรายวิชาวิทยาศาสตร์และเทคโนโลยี ระดับชั้นประถมศึกษาปีที่ 6

ระยะที่ 2 การออกแบบและพัฒนาต้นแบบแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6

เป็นการดำเนินงานเพื่อออกแบบและพัฒนาต้นแบบแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยมีขั้นตอนการดำเนินงานที่ประกอบไปด้วย

2.1 การกำหนดหลักการออกแบบต้นแบบแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยการวิเคราะห์องค์ประกอบและกำหนดรูปแบบของแบบทดสอบ

2.2 การพัฒนาต้นแบบแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6 ตามหลักการออกแบบที่กำหนด โดยสร้างเป็นแบบทดสอบวัดทักษะการคิดอย่างมีวิจารณญาณชนิดเลือกตอบ 4 ตัวเลือก ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน โดยข้อคำถามแต่ละข้อจะวัดองค์ประกอบของทักษะการคิดอย่างมีวิจารณญาณทั้งหมด 5 องค์ประกอบ จากนั้นตรวจสอบความเที่ยงตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญจำนวน 5 ท่าน และปรับปรุงข้อคำถามตามคำแนะนำแล้วจึงนำแบบวัดทักษะการคิดอย่างมีวิจารณญาณไปทดลองใช้กับนักเรียนที่เป็นกลุ่มตัวอย่างจำนวน 60 คน

2.3 การประเมินต้นแบบแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6 ขั้นต้น ซึ่งมีรายละเอียดและขั้นตอนในการสร้างดังนี้

2.3.1 พัฒนาและตรวจสอบคุณภาพแบบวัดทักษะการคิดอย่างมีวิจารณญาณ

2.3.2 วิเคราะห์หาคุณภาพของเครื่องมือ โดยใช้การวิเคราะห์แบบ CTT ดังนี้ 1) ค่าอำนาจจำแนก (discrimination) 2) ค่าความยากง่าย (difficulty) 3) ความตรงเชิงทฤษฎี (construct validity) และ 4) ความเที่ยง (reliability)

2.3.3 ปรับปรุงข้อคำถามที่ที่คุณภาพต่ำกว่าเกณฑ์คุณภาพของเครื่องมือที่ยอมรับได้

ระยะที่ 3 การประเมินและการสะท้อนคิด

เป็นการดำเนินการเพื่อประเมินและสะท้อนผลการทดลองใช้แบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยมีขั้นตอนการดำเนินงานที่ประกอบไปด้วย

3.1 กลุ่มตัวอย่างที่ใช้ในการเก็บข้อมูลจริง เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ที่เรียนอยู่ในปีการศึกษา 2566 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 1 – 5 ที่กำหนดโดยวิธีการสุ่มแบบหลายขั้นตอน (Multi - Stage Random Sampling) จำนวน 500 คน (ในการวิเคราะห์ข้อมูลผู้วิจัยใช้ การวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (dichotomous item response theory) เพื่อให้ได้กลุ่มตัวอย่างมีจำนวนตัวอย่างสมดุลกันและมีความน่าเชื่อถือในการสรุปอ้างอิง โดยในการวิจัยในครั้งนี้ได้คำนวณขนาดกลุ่มตัวอย่าง จากจำนวนประชากรนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่นเขต 1-5 จำนวน 13,717 คน ดังนั้น จึงคำนวณหาประชากรตามหลักของทาโรยามาเน่ (Yamane, 1973) ที่ระดับความน่าเชื่อถือร้อยละ 95 ความคลาดเคลื่อนที่ร้อยละ ± 5 ดังนี้

$$n = \frac{13,717}{1 + 13,717(0.05)^2}$$

ดังนั้น $n = 388.67$
หรือ กลุ่มตัวอย่าง = 389 คน

เนื่องจากในการวิเคราะห์ข้อมูลผู้วิจัยใช้ การวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า เพื่อให้ได้กลุ่มตัวอย่างมีจำนวนตัวอย่างสมดุลกันและมีความน่าเชื่อถือในการสรุปอ้างอิง จึงได้กำหนดกลุ่มตัวอย่าง 500 คน (Kanjawasee, 2020) ในการสุ่มตัวอย่าง โดยใช้การสุ่มแบบ 2 ขั้นตอน ดังนี้

ขั้นที่ 1 สุ่มโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 1 – 5 โดยการสุ่มอย่างง่าย โดยสุ่มจากโรงเรียนที่มีจำนวนนักเรียนชั้นประถมศึกษาปีที่ 6 มากกว่า 50 คนขึ้นไป จำนวน 10 โรงเรียน

ขั้นที่ 2 ทำการสุ่มจำนวนนักเรียน โรงเรียนละ 50 คน โดยใช้วิธีการสุ่มอย่างง่าย

3.2 ตัวแปร

3.2.1 ตัวแปรจัดกระทำ คือ แบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6

3.2.2 ตัวแปรตาม คือ คุณภาพของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 500 คน

3.3 เครื่องมือวิจัย เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นแบบวัดทักษะการคิดอย่างมีวิจารณญาณ ชนิดเลือกตอบ 4 ตัวเลือก ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน โดยข้อคำถามแต่ละข้อจะวัดองค์ประกอบของทักษะการคิดอย่างมีวิจารณญาณทั้งหมด 5 องค์ประกอบ ได้แก่ 1) ด้านการระบุปัญหา 2) ด้านการระบุสมมติฐาน 3) ด้านการรวบรวมข้อมูล 4) ด้านการวิเคราะห์ข้อมูล และ 5) ด้านการลงข้อสรุป ที่มีการปรับปรุงข้อคำถามเรียบร้อยแล้ว จำนวน 40 ข้อ

3.4 การเก็บรวบรวมข้อมูล

3.4.1 ทำหนังสือจากคณะต้นสังกัดไปถึงผู้อำนวยการโรงเรียนเพื่อขออนุญาตดำเนินการวิจัยและเข้าถึงกลุ่มเป้าหมายที่เป็นนักเรียน ชี้แจงรายละเอียดของโครงการวิจัยกับนักเรียน และฝากเอกสารข้อมูลฯ และ

หนังสือแสดงความยินยอมจากผู้ปกครอง หากผู้ปกครองยินยอมให้เด็กในปกครองเข้าร่วมการวิจัย ขอให้ส่งหนังสือแสดงความยินยอมถึงผู้วิจัยโดยตรง จากนั้นผู้วิจัยนัดหมายวันเวลาเก็บรวบรวมข้อมูลโดยผ่านทางผู้อำนวยการสถานศึกษา

3.4.2 ดำเนินการสอบตามวัน เวลาที่นัด โดยก่อนดำเนินการสอบผู้วิจัยอธิบายถึงวัตถุประสงค์ในการทดสอบ ความสำคัญของการทำแบบวัด ประโยชน์ที่นักเรียนจะได้รับ และวิธีการทำ

3.5 การวิเคราะห์ข้อมูล

3.5.1 การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สองของตัวชี้วัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6

3.5.2 วิเคราะห์คุณภาพเครื่องมือ โดยใช้การวิเคราะห์ ดังนี้ 1) ค่าอำนาจจำแนก (discrimination) 2) ค่าความยากง่าย (difficulty) 3) ความตรงเชิงทฤษฎี (construct validity) 4) ความเที่ยง (reliability) และ 5) ฟังก์ชันสารสนเทศของข้อสอบ

3.5.3 การวิเคราะห์จุดตัดคะแนนของทักษะการคิดอย่างมีวิจารณญาณ โดยใช้การวิเคราะห์กลุ่มแฝง

ผลการวิจัย

ผลการดำเนินการวิจัย เรื่อง การพัฒนาแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 ผู้วิจัยนำเสนอผลการวิเคราะห์ข้อมูลโดยแบ่งเป็น 3 ตอน คือ

ตอนที่ 1 ผลการสังเคราะห์องค์ประกอบในการวัดทักษะการคิดอย่างมีวิจารณญาณ

1.1 ผลการวิเคราะห์องค์ประกอบและพัฒนาตัวชี้วัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6

1.1.1 ผลการพัฒนาพัฒนาตัวชี้วัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 จากการสังเคราะห์ข้อมูล พบว่าพัฒนาตัวชี้วัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 มี 5 องค์ประกอบ 16 ตัวชี้วัด จากนั้นให้ผู้เชี่ยวชาญจำนวน 5 ท่านพิจารณาความสอดคล้องระหว่างองค์ประกอบและตัวชี้วัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 พบว่ามีค่า IOC เท่ากับ 1.00 จำนวน 13 ข้อ และมีค่า IOC เท่ากับ 0.80 จำนวน 3 ข้อ ทั้งนี้ผู้เชี่ยวชาญยังได้ให้ข้อเสนอแนะในการพัฒนาตัวชี้วัดเพิ่มเติม ผู้วิจัยจึงนำข้อเสนอแนะจากความคิดเห็นของผู้เชี่ยวชาญมาดำเนินการปรับปรุงและพัฒนาตัวชี้วัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี

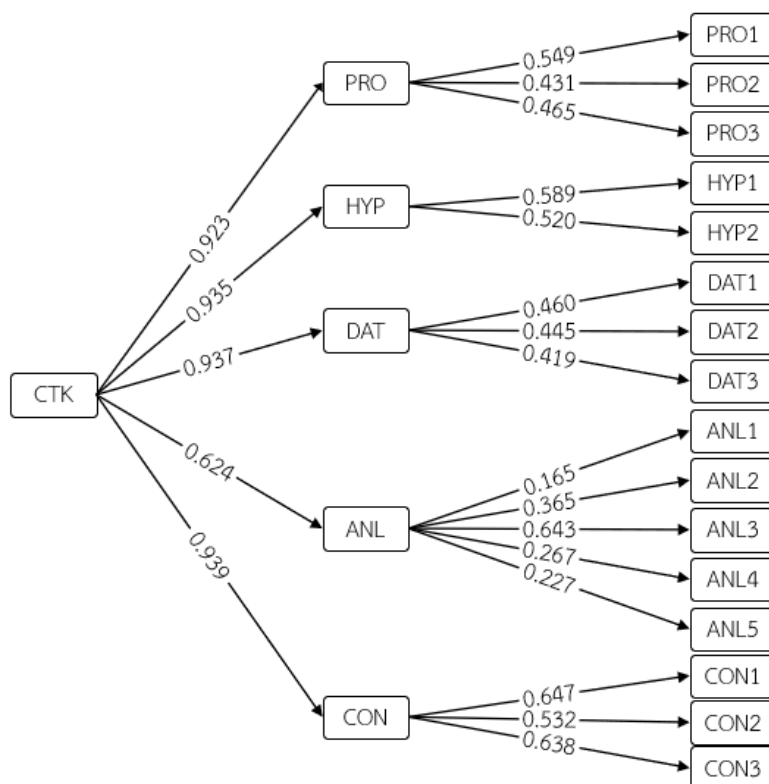
1.1.2 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สองของตัวชี้วัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 พบว่ามีค่า $\chi^2 = 128.412$, $df = 100$, $p\text{-value} = 0.0231$ แสดงว่าค่า χ^2 ไม่มีนัยสำคัญทางสถิติ โดยมีค่า CFI = 0.987 และค่า TLI = 0.992 ซึ่งมีค่ามากกว่า 0.90 ตามเกณฑ์ (Hooper et al., 2008) ค่า SRMR = 0.031 และค่า RMSEA = 0.026 มีค่าน้อยกว่า 0.05 ตามเกณฑ์ (Schumacker & Lomax, 2004) ดังนั้นโมเดลองค์ประกอบทักษะการคิดอย่างมีวิจารณญาณมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ผลแสดงได้ดังตาราง 1 และภาพ 1

ตาราง 1 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สองของโมเดลองค์ประกอบทักษะการคิดอย่างมี
 วิจารณญาณ

ตัวแปร		β	S.E.	R^2
การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่หนึ่ง				
การระบุปัญหา (PRO)	PRO1	0.549	0.039	0.698
	PRO2	0.431	0.042	0.185
	PRO3	0.465	0.043	0.216
การระบุสมมติฐาน (HYP)	HYP1	0.589	0.042	0.347
	HYP2	0.520	0.042	0.271
การรวบรวมข้อมูล (DAT)	DAT1	0.460	0.044	0.211
	DAT2	0.445	0.044	0.198
	DAT3	0.419	0.045	0.176
การวิเคราะห์ข้อมูล (ANL)	ANL1	0.165	0.049	0.027
	ANL2	0.365	0.046	0.133
	ANL3	0.643	0.043	0.414
	ANL4	0.267	0.045	0.071
	ANL5	0.227	0.047	0.052
การลงข้อสรุป (CON)	CON1	0.647	0.032	0.418
	CON2	0.532	0.036	0.283
	CON3	0.638	0.032	0.407
การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สอง				
ทักษะการคิดอย่างมีวิจารณญาณ (CTK)	PRO	0.923	0.047	
	HYP	0.935	0.056	
	DAT	0.937	0.058	
	ANL	0.624	0.053	
	CON	0.939	0.028	

จากตาราง 1 และภาพ 1 พบว่าค่าน้ำหนักองค์ประกอบในรูปคะแนนมาตรฐาน (standardized coefficient : β) ของโมเดลองค์ประกอบทักษะการคิดอย่างมีวิจารณญาณในอันดับที่หนึ่ง (first-order) จำนวน 16 ค่า มีค่าตั้งแต่ 0.165 ถึง 0.647 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ทั้ง 16 ค่า โดยตัวบ่งชี้แต่ละตัวมีค่าความเชื่อมั่น (Reliability: R^2) ตั้งแต่ 0.027 ถึง 0.698 แสดงว่าตัวบ่งชี้ทั้ง 16 ตัวนี้เป็นตัวบ่งชี้ที่สำคัญต่อทักษะการคิดอย่างมีวิจารณญาณในแต่ละองค์ประกอบย่อย

เมื่อพิจารณาค่าน้ำหนักองค์ประกอบในรูปคะแนนมาตรฐานของโมเดลองค์ประกอบสมรรถนะของครูด้านการประเมินการเรียนรู้ในอันดับที่สอง (second-order) จำนวน 5 ค่า พบว่า มีค่าตั้งแต่ 0.624 ถึง 0.939 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ทั้ง 5 ค่า



ภาพ 1 โมเดลองค์ประกอบทักษะการคิดอย่างมีวิจารณญาณตามตัวชี้วัดที่พัฒนาผู้วิจัยได้กำหนดสัญลักษณ์และความหมายของสัญลักษณ์ ดังนี้

CTK	แทน ทักษะการคิดอย่างมีวิจารณญาณ
PRO	แทน องค์ประกอบหลักด้านการระบุปัญหา
HYP	แทน องค์ประกอบหลักด้านการระบุมมติฐาน
DAT	แทน องค์ประกอบหลักด้านการรวบรวมข้อมูล
ANL	แทน องค์ประกอบหลักด้านการวิเคราะห์ข้อมูล
CON	แทน องค์ประกอบหลักด้านการลงข้อสรุป
PRO1	แทน ตัวบ่งชี้ด้านการเข้าใจประเด็นปัญหา ข้อคำถาม ข้ออ้าง หรือข้อโต้แย้ง
PRO2	แทน ตัวบ่งชี้ด้านการพิจารณาข้อมูลหรือสถานการณ์ที่ปรากฏ รวมทั้งความหมายของคำ หรือความชัดเจนของข้อความ
PRO3	แทน ตัวบ่งชี้ด้านการกำหนดประเด็นข้อสงสัยและประเด็นหลักที่ควรพิจารณา และการแสวงหาคำตอบได้
HYP1	แทน ตัวบ่งชี้ด้านการพิจารณาแนวทางการสรุปอ้างอิงของปัญหาโดยการนำข้อมูลที่ได้จากการจัดระบบแล้วมาพิจารณาหาความสัมพันธ์ ความสอดคล้องของข้อมูล
HYP2	แทน ตัวบ่งชี้ด้านการพิจารณาความเชื่อมโยงของข้อมูล เพื่อระบุแนวทางการสรุปที่น่าจะเป็นไปได้โดยพิจารณาเลือกแนวทางที่เป็นไปได้ที่สอดคล้องมากที่สุด
DAT1	แทน ตัวบ่งชี้ด้านการรวบรวมข้อมูลทั้งทางตรงและทางอ้อมจากแหล่งข้อมูลต่าง ๆ รวมถึงการรวมข้อมูลจากประสบการณ์เดิมที่มีอยู่ ซึ่งได้จากการคิด การพูดคุย การสังเกตที่เกิดขึ้นจากตนเองและผู้อื่น

- DAT2 แทน ตัวบ่งชี้ด้านการพิจารณา ประเมิน ตรวจสอบ ตัดสินข้อมูลทั้งในเชิงปริมาณและคุณภาพ
- DAT3 แทน ตัวบ่งชี้ด้านการพิจารณาถึงที่มาของข้อมูลสถิติ และหลักฐานที่ปรากฏ รวมทั้งความเพียงพอของข้อมูลในแง่มุมต่าง ๆ ที่จะนำไปสู่การลงข้อสรุปอย่างมีเหตุผล หากยังไม่เกี่ยวข้องที่จะใช้พิจารณาข้อสรุป ก็จะต้องรวบรวมข้อมูลเพิ่มเติม
- ANL1 แทน ตัวบ่งชี้ด้านการจำแนกประเภทของข้อมูล ระบุแนวคิดที่อยู่เบื้องต้นหลังข้อมูลที่ปรากฏ
- ANL2 แทน ตัวบ่งชี้ด้านการพิจารณา แยกแยะ เปรียบเทียบความแตกต่างของข้อมูล ตีความข้อมูล ประเมินว่าข้อมูลใดเป็น ข้อเท็จจริง ข้อมูลใดเป็นข้อคิดเห็น
- ANL3 แทน ตัวบ่งชี้ด้านการพิจารณาความน่าเชื่อถือของแหล่งข้อมูล
- ANL4 แทน ตัวบ่งชี้ด้านการระบุข้อสันนิษฐานหรือข้อตกลงเบื้องต้นที่อยู่เบื้องต้นหลังข้อมูลที่ปรากฏ เป็นการนำความรู้ไปใช้ในสถานการณ์ใหม่ที่อาศัยข้อมูลจากประสบการณ์เดิมมาร่วมพิจารณา
- ANL5 แทน ตัวบ่งชี้ด้านการสังเคราะห์จัดกลุ่มและจัดลำดับความสำเร็จของข้อมูลเพื่อใช้เป็นแนวทางสำหรับการแก้ไขปัญหา
- CON1 แทน ตัวบ่งชี้ด้านการพิจารณาเลือกแนวทางที่สมเหตุสมผลที่สุดจากข้อมูลที่ได้ทำการรวบรวมมาในเบื้องต้น
- CON2 แทน ตัวบ่งชี้ด้านการประเมินความถูกต้อง พิจารณาการกระทำหรือข้อมูลต่าง ๆ ที่เชื่อมโยงเกี่ยวข้องกัน ซึ่งต้องอาศัยความสามารถในการวิเคราะห์และประเมินอย่างรอบคอบ
- CON3 แทน ตัวบ่งชี้ด้านการพิจารณาความสมเหตุสมผลจากข้อมูลที่มีอยู่ และสรุปเพื่อตัดสินใจ

ตอนที่ 2 ผลการพัฒนาและตรวจสอบคุณภาพของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6

2.1 การพัฒนาแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6

แบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยพัฒนาขึ้นมี 5 องค์ประกอบ คือ 1) ด้านการระบุปัญหา 2) ด้านการระบุสมมติฐาน 3) ด้านการรวบรวมข้อมูล 4) ด้านการวิเคราะห์ข้อมูล และ 5) ด้านการลงข้อสรุป จำนวนด้านละ 8 ข้อ รวมเป็น 40 ข้อ

2.2. ผลการตรวจสอบคุณภาพของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6

ระยะที่ 1 ผลการตรวจสอบคุณภาพของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 โดยผู้เชี่ยวชาญและจากการเก็บข้อมูลกับกลุ่มทดลองใช้ (try out) คือ นักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 60 คน โดยวิเคราะห์หาคุณภาพของเครื่องมือ โดยใช้การวิเคราะห์แบบ CTT ดังนั้น ค่าดัชนีความสอดคล้องระหว่างองค์ประกอบทักษะการคิดอย่างมีวิจารณญาณกับตัวชี้วัดองค์ประกอบทักษะการคิดอย่างมีวิจารณญาณโดยผู้เชี่ยวชาญ 5 ท่าน พบว่า ข้อสอบมีค่า IOC เท่ากับ 1 ทั้งหมด 40 ข้อ วิเคราะห์หาค่าความเที่ยง (reliability) ของแบบวัดทักษะการคิดอย่างมีวิจารณญาณ

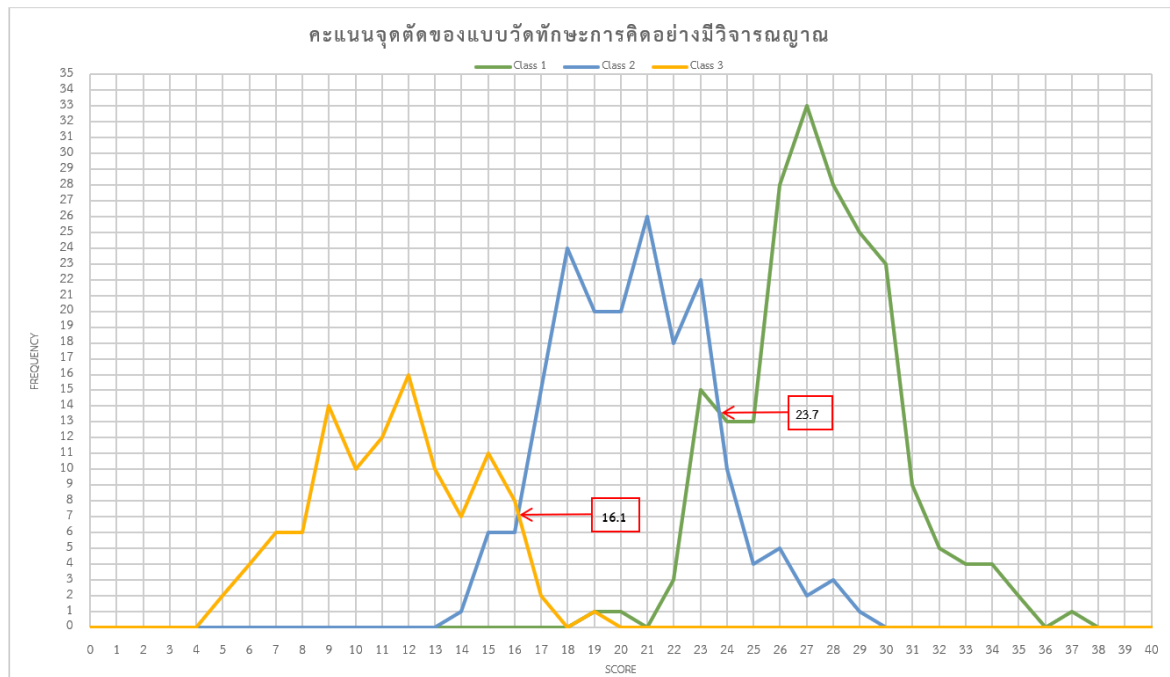
ในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 รายองค์ประกอบย่อยและทั้งฉบับโดยการหาสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Alpha coefficient) โดยใช้โปรแกรม SPSS พบว่าค่าความเที่ยงรายองค์ประกอบมีค่าระหว่าง 0.497 - 0.661 ค่าความเที่ยงของข้อสอบทั้งฉบับ มีค่าเท่ากับ 0.825 การวิเคราะห์ค่าอำนาจจำแนกตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ของการวิเคราะห์ความสัมพันธ์ของคะแนนรายข้อ มีค่าอยู่ระหว่าง 0.275 - 0.530 และการวิเคราะห์ค่าความยากง่ายตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) โดยใช้ดัชนีความยากง่าย มีค่าอยู่ระหว่าง 0.350 - 0.730

ระยะที่ 2 การตรวจสอบคุณภาพของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 จากการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่าง จำนวน 500 คน โดยใช้การวิเคราะห์แบบ IRT พบว่า ค่าพารามิเตอร์อำนาจจำแนก (a_i) มีค่าอยู่ระหว่าง 0.39 - 5.03 และเมื่อพิจารณาคัดเลือกข้อคำถามที่มีคุณภาพที่ยอมรับได้ จำนวน 32 ข้อ ค่าพารามิเตอร์ความยากของข้อสอบ (b_i) มีค่าอยู่ระหว่าง -1.29 ถึง 5.14 และเมื่อพิจารณาคัดเลือกข้อคำถามที่มีคุณภาพที่ยอมรับได้ จำนวน 37 ข้อ และสัมประสิทธิ์การเดา (c_i) มีค่าอยู่ระหว่าง 0.13 ถึง 0.27 และเมื่อพิจารณาคัดเลือกข้อคำถามที่มีคุณภาพที่ยอมรับได้ทั้งหมดจำนวน 40 ข้อ

ตอนที่ 3 ผลการพัฒนาจุดตัดคะแนนทักษะการคิดอย่างมีวิจารณญาณของนักเรียนชั้นประถมศึกษาปีที่ 6

สำหรับการวิจัยครั้งนี้ได้ทำการจัดกลุ่มนักเรียนตามทักษะการคิดอย่างมีวิจารณญาณโดยการวิเคราะห์กลุ่มแฝง ตามองค์ประกอบทั้ง 5 ด้าน ซึ่งประกอบด้วย การระบุปัญหา (PRO) การระบุสมมติฐาน (HYP) การรวบรวมข้อมูล (DAT) การวิเคราะห์ข้อมูล (ANL) และการลงข้อสรุป (CON) จากการพิจารณาความกลมกลืนของโมเดลในการวิเคราะห์กลุ่มแฝงจากดัชนี AIC, BIC, ABIC และพิจารณาค่าดัชนีบ่งชี้คุณภาพการจำแนกกลุ่มแฝง (Entropy : E_k) พบว่า โมเดลที่มีจำนวนกลุ่มแฝง 3 กลุ่ม เป็นโมเดลที่มีความเหมาะสมและมีคุณสมบัติตรงตามเกณฑ์การพิจารณามากที่สุด (Likelihood = -11858.198 , df = 378 , AIC = 23960.395 , BIC = 24474.578 , Adj.BIC = 24087.342 , E_k = 0.846) และเมื่อพิจารณาค่า Lo-Mendell-Rubin Adjusted LRT Test (LMR) พบว่ากลุ่มแฝง 3 กลุ่มและกลุ่มแฝง 4 กลุ่มมีค่า Likelihood Ratio (L^2) ไม่แตกต่างกัน (p-value = 0.6181) ดังนั้นจำนวนกลุ่มแฝงของนักเรียนตามทักษะการคิดอย่างมีวิจารณญาณที่เหมาะสมจึงควรมี 3 กลุ่ม ผู้วิจัยได้วิเคราะห์กำหนดคะแนนจุดตัดของแบบวัดของกลุ่มแฝงทักษะการคิดอย่างมีวิจารณญาณของนักเรียนระดับชั้นประถมศึกษาปีที่ 6 ซึ่งพิจารณาจากคะแนนรวมและความถี่ของกลุ่มแฝงทักษะการคิดอย่างมีวิจารณญาณที่ได้จากวิธีการวิเคราะห์กลุ่มแฝง (Latent class analysis) จำนวน 3 กลุ่ม ได้แก่ กลุ่มแฝงทักษะการคิดอย่างมีวิจารณญาณ 1 (กลุ่มสูง) กลุ่มแฝงทักษะการคิดอย่างมีวิจารณญาณ 2 (กลุ่มปานกลาง) และกลุ่มแฝงทักษะการคิดอย่างมีวิจารณญาณ 3 (กลุ่มต่ำ) เมื่อพิจารณาคะแนนจุดตัดของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี ระดับชั้นประถมศึกษาปีที่ 6 พบว่า ณ จุดคะแนนเท่ากับ 23.7 เป็นคะแนนจุดตัดที่สามารถแบ่งกลุ่มแฝงทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี 1 (กลุ่มสูง) และกลุ่มแฝงทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี 2 (กลุ่มปานกลาง) ออกจากกันตามระดับคะแนนทักษะการคิดอย่างมีวิจารณญาณ

ในรายวิชาวิทยาศาสตร์และเทคโนโลยี และ ณ จุดคะแนนเท่ากับ 16.1 เป็นคะแนนจุดตัดที่สามารถแบ่งกลุ่มแฝงทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี 2 (กลุ่มปานกลาง) และกลุ่มแฝงทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี 3 (กลุ่มต่ำ) ออกจากกันตามระดับคะแนนทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี



ภาพ 2 คะแนนจุดตัดของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี ระดับชั้นประถมศึกษาปีที่ 6

อภิปรายผล

ประเด็นที่ 1 การสังเคราะห์องค์ประกอบในการวัดทักษะการคิดอย่างมีวิจารณญาณ

จากผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สองของตัวชี้วัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 พบว่ามีค่า $\chi^2 = 131.422$, $df = 99$, $p\text{-value} = 0.0163$ แสดงว่าค่า χ^2 ไม่มีนัยสำคัญทางสถิติ โดยมีค่า CFI = 0.977 และค่า TLI = 0.972 ซึ่งมีค่ามากกว่า 0.90 ตามเกณฑ์ (Hooper et al., 2008) ค่า SRMR = 0.032 และค่า RMSEA = 0.026 มีค่าน้อยกว่า 0.05 ตามเกณฑ์ (Schumacker & Lomax, 2004) และพบว่าองค์ประกอบที่ 2 ด้านการระบุสมมติฐาน (Hypothesis Formulation) มีน้ำหนักองค์ประกอบในรูปคะแนนมาตรฐานสูงที่สุด มีค่าเท่ากับ 1.075 ซึ่งชี้ให้เห็นว่าทักษะในการระบุสมมติฐานมีความสำคัญมากที่สุด เพราะสมมติฐานที่ดีจะเป็นแนวทางสู่การออกแบบแนวทางการหาคำตอบที่ดีเพราะจะทำให้สามารถกำหนดขอบเขตของการทดลองได้ มุ่งไปสู่การหาคำตอบได้อย่างแม่นยำ ไม่ออกนอกประเด็น และในท้ายที่สุดก็จะนำไปสู่การตรวจสอบกับสมมติฐานตั้งต้น

ประเด็นที่ 2 คุณภาพของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6

จากผลการวิเคราะห์คุณภาพของแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 ที่ได้พัฒนาขึ้น มีคุณภาพตามเกณฑ์ที่ยอมรับได้ โดยมีรายละเอียดดังนี้

1.1 ความตรงเชิงทฤษฎี (construct validity) พบว่า แบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 มีค่าดัชนีความสอดคล้องระหว่างองค์ประกอบทักษะการคิดอย่างมีวิจารณญาณกับตัวชี้วัดองค์ประกอบทักษะการคิดอย่างมีวิจารณญาณ โดยผู้เชี่ยวชาญ 5 ท่าน เท่ากับ 1 ทั้งหมด 40 ข้อ ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ คือ 0.8 ขึ้นไป (Davis, 1992)

1.2 ความเที่ยง (Reliability) พบว่าแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 มีค่าความเที่ยงตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) รายองค์ประกอบมีค่าระหว่าง 0.497-0.661 ซึ่งอยู่ในเกณฑ์ค่อนข้างต่ำ ซึ่งเกณฑ์ที่ยอมรับได้คือ มากกว่า 0.5 (Kanjawasee, 2013) แต่เมื่อพิจารณาทั้งฉบับ มีค่าเท่ากับ 0.825 ซึ่งอาจจะเกิดจากปัจจัยหลายอย่าง เช่น ความยาวของข้อสอบ ระยะเวลาในการทำข้อสอบ หรือจำนวนข้อสอบที่มากจนเกินไป ดังนั้นหากจะนำไปใช้จริง ควรปรับปรุงแบบวัดและตรวจสอบคุณภาพข้อสอบใหม่อีกครั้ง

1.3 ค่าอำนาจจำแนก (discrimination power) พบว่าแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 มีค่าอำนาจจำแนกตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) มีค่าอยู่ระหว่าง 0.2746-0.5300 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ คือ มีค่าอยู่ระหว่าง 0.2-0.8 (Kanjawasee, 2013)

1.4 ค่าความยากง่าย (difficulty) พบว่าแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 มีดัชนีความยากง่ายตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) อยู่ระหว่าง 0.35-0.73 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ คือ มีค่าอยู่ระหว่าง 0.2-0.8 (Kanjawasee, 2013)

1.5 ค่าพารามิเตอร์อำนาจจำแนก (a_i) พบว่าแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 มีค่าพารามิเตอร์อำนาจจำแนก (a_i) ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) อยู่ระหว่าง 0.39 – 5.03 และเมื่อพิจารณาคัดเลือกข้อคำถามที่มีคุณภาพที่ยอมรับได้ จำนวน 32 ข้อ ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ คือ มีค่าอยู่ระหว่าง +0.50 ถึง +2.5 (Kanjawasee, 2020) มี 8 ข้อที่ควรปรับปรุงคุณภาพของข้อสอบ

1.6 ค่าพารามิเตอร์ความยากของข้อสอบ (b_i) พบว่าแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 มีค่าพารามิเตอร์ความยากของข้อสอบ (b_i) ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) อยู่ระหว่าง -1.29 ถึง 5.14 และเมื่อพิจารณาคัดเลือกข้อคำถามที่มีคุณภาพที่ยอมรับได้ จำนวน 37 ข้อ ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ คือ มีค่าอยู่ระหว่าง -2.5 ถึง +2.5 (Kanjawasee, 2020) มี 3 ข้อที่ควรปรับปรุงคุณภาพของข้อสอบ

1.7 สัมประสิทธิ์การเดา (c_i) พบว่าแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 มีค่าสัมประสิทธิ์การเดา (c_i) ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) อยู่ระหว่าง 0.13 ถึง 0.27 และเมื่อพิจารณาคัดเลือกข้อคำถามที่มีคุณภาพที่ยอมรับได้ทั้งหมดจำนวน 40 ข้อ ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ คือมีค่าไม่เกิน 0.30 (Kanjanaawasee, 2020)

ประเด็นที่ 3 การพัฒนาจุดตัดคะแนนทักษะการคิดอย่างมีวิจารณญาณของนักเรียนชั้นประถมศึกษาปีที่ 6

จากผลการพัฒนาจุดตัดคะแนนทักษะการคิดอย่างมีวิจารณญาณของนักเรียนชั้นประถมศึกษาปีที่ 6 พบว่า สามารถจำแนกกลุ่มแบ่งได้เป็น 3 กลุ่ม ได้แก่ กลุ่มแบ่งทักษะการคิดอย่างมีวิจารณญาณ 1 คือ กลุ่มที่มีคะแนนทักษะการคิดอย่างมีวิจารณญาณอยู่ในระดับสูง กลุ่มแบ่งทักษะการคิดอย่างมีวิจารณญาณ 2 คือ กลุ่มที่มีคะแนนทักษะการคิดอย่างมีวิจารณญาณอยู่ในระดับปานกลาง และกลุ่มแบ่งทักษะการคิดอย่างมีวิจารณญาณ 3 คือ กลุ่มที่มีคะแนนทักษะการคิดอย่างมีวิจารณญาณอยู่ในระดับต่ำ

การพิจารณากำหนดคะแนนจุดตัด พบว่า ณ จุดคะแนนเท่ากับ 23.7 เป็นคะแนนจุดตัดที่สามารถแบ่งกลุ่มแบ่งทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี 1 (กลุ่มสูง) และกลุ่มแบ่งทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี 2 (กลุ่มปานกลาง) ออกจากกันตามระดับคะแนนทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี และ ณ จุดคะแนนเท่ากับ 16.1 เป็นคะแนนจุดตัดที่สามารถแบ่งกลุ่มแบ่งทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี 2 (กลุ่มปานกลาง) และกลุ่มแบ่งทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี 3 (กลุ่มต่ำ) ออกจากกันตามระดับคะแนนทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยี โดยอ้างอิงแนวทางการกำหนดจุดตัดคะแนนจาก Malinagam (2015) ซึ่งผลการแบ่งกลุ่มระดับความสามารถในการคิดอย่างมีวิจารณญาณสอดคล้องกับแบบสอบ Cornell Critical Thinking Test ที่พัฒนาโดย Ennis & Millman (1985) ได้แบ่งกลุ่มระดับความสามารถในการคิดอย่างมีวิจารณญาณ ออกเป็น 3 กลุ่มได้แก่ ระดับสูง ระดับปานกลาง และระดับปรับปรุง

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 แบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 ฉบับนี้ พัฒนาขึ้นจากตัวอย่างของนักเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 1-5 เท่านั้น หากมีการนำไปใช้กับนักเรียนในบริบทที่แตกต่างกันควรมีการตรวจสอบคุณภาพในเบื้องต้นก่อนนำไปใช้

1.2 การนำแบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 ไปใช้กับกลุ่มตัวอย่างที่เป็นกลุ่มเดียวกันกับประชากรในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาขอนแก่น เขต 1-5 ให้แปลผลคะแนนทักษะการคิดอย่างมีวิจารณญาณตามการวิเคราะห์กลุ่มแบ่งทักษะการคิดอย่างมีวิจารณญาณเป็น 3 กลุ่ม ได้แก่ กลุ่มสูง กลุ่มปานกลาง และกลุ่มต่ำ

1.3 องค์ประกอบของทักษะการคิดอย่างมีวิจารณญาณ ผู้วิจัยได้สังเคราะห์จากงานเอกสารและงานวิจัยที่เกี่ยวข้อง แต่พบว่าไม่มีบางองค์ประกอบที่ความถี่ยังน้อยอยู่ ดังนั้นผู้ที่มีความสนใจจะนำไปใช้ควรพิจารณาถึงตัวอย่างที่จะนำไปใช้ด้วย

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 เนื่องจากการทำวิจัยครั้งนี้ใช้การวิเคราะห์กลุ่มแฝงในการวิเคราะห์เพื่อจัดกลุ่มนักเรียน ดังนั้นในการวิจัยครั้งต่อไปควรศึกษาเปรียบเทียบวิธีการวิเคราะห์คะแนนจุดตัดเทียบกับวิธีการวิเคราะห์ในรูปแบบอื่นเพื่อทำให้การวิจัยมีความชัดเจนมากยิ่งขึ้น

2.2 เนื่องจากในงานวิจัยนี้แบบวัดทักษะการคิดอย่างมีวิจารณญาณในรายวิชาวิทยาศาสตร์และเทคโนโลยีระดับชั้นประถมศึกษาปีที่ 6 ใช้รูปแบบการให้คะแนนตามทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (dichotomous item response theory) ในการวิจัยครั้งต่อไปควรมีการเปรียบเทียบกับรูปแบบการให้คะแนนตามทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (polytomous item response theory) เพื่อให้เห็นคุณภาพของเครื่องมือทั้ง 2 รูปแบบ

References

- Davis, L. L. (1992). Instrument review: Getting the most from your panel of experts. *Applied Nursing Research*, 5(4), 194–197. [https://doi.org/10.1016/S0897-1897\(05\)80008-4](https://doi.org/10.1016/S0897-1897(05)80008-4)
- Ennis, R. H., & Millman, J. (1985). *Cornell Critical Thinking Test, Level X*. Midwest Publications.
- Glass, G. V. (1978). Standards and criteria. *Journal of Educational Measurement*, 15(4), 243–257.
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53–60.
- Millman, J. (1973). Passing scores and test length for domain-referenced measures. *Review of Educational Research*, 43(2), 205–216. <https://doi.org/10.3102/00346543043002205>
- Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Lawrence Erlbaum Associates.
- Watson, G., & Glaser, E. M. (1980). *Watson-Glaser Critical Thinking Appraisal, Forms A and B manual*. The Psychological Corporation.
- Yamane, T. (1973). *Statistics: An introductory analysis* (3rd ed.). Harper & Row.

Translate Thai References

- Kammanee, T. (2018). *Teaching theory: Knowledge for organizing effective learning processes*. Chulalongkorn University. (in Thai)

Kanjanawasee, S. (2013). *Classical test theory* (7th ed.). Chulalongkorn University Press.

(in Thai)

Kanjanawasee, S. (2020). *Modern test theory* (5th ed.). Chulalongkorn University Press.

(in Thai)

Malingam, C. (2015). *A development of scientific mild scale for lower secondary school students: Application of latent class analysis for establishing cut scores* [Master's thesis]. Chulalongkorn University. (in Thai)

Wonggom, J. (2009). *Critical thinking ability and learning achievement in science subject for grade 7 students on the property and separation of substances by inquiry cycle* [Master's thesis]. Khon Kaen University. (in Thai)

การพัฒนาแบบวัดโกรว์ธมายด์เซตสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น
ด้วยการประยุกต์ใช้แองเคอริงวินเยตต์

The Development of a Scale to Measure Growth Mindsets for Lower
Secondary School Students with the Application of Anchoring Vignettes

สิริวัฒน์ หลูโน^{1*} สัງวรรณ งัดกระโทก² และ ปุริมปรัชญ์ คณินพศุตย์³
Siriwat Luno^{1*}, Sungworn Ngudgratoke² and Purimpratch Khaninphasut³
(Received: May 5, 2024; Revised: June 7, 2024; Accepted: June 11, 2024)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมายเพื่อ 1) พัฒนาแบบวัดโกรว์ธมายด์เซตและแองเคอริงวินเยตต์สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น และ 2) ตรวจสอบคุณภาพของแบบวัดโกรว์ธมายด์เซตสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้นด้วยการประยุกต์ใช้แองเคอริงวินเยตต์เพื่อช่วยให้นักเรียนสามารถตีความและตอบมาตรประมาณค่าได้อย่างมีประสิทธิภาพ กลุ่มตัวอย่างประกอบด้วยนักเรียนมัธยมศึกษาตอนต้นที่ศึกษาอยู่ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาลำปาง เขต 1 ถึง 3 จำนวน 300 คน โดยใช้วิธีการสุ่มตัวอย่างแบบหลายขั้นตอนเพื่อให้มั่นใจในการเป็นตัวแทนจากโรงเรียนต่าง ๆ การวิจัยนี้ใช้มาตรประมาณค่าร่วมกับแองเคอริงวินเยตต์เป็นเครื่องมือในการเก็บข้อมูล การวิเคราะห์ข้อมูลใช้ค่าสัมประสิทธิ์แอลฟาของครอนบาคเพื่อประเมินความสอดคล้องภายใน ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเพื่อสำรวจความสัมพันธ์ระหว่างตัวแปร การวิเคราะห์องค์ประกอบเชิงยืนยันเพื่อยืนยันโครงสร้างของแบบวัด และทฤษฎีการตอบสนองข้อสอบเพื่อประเมินประสิทธิภาพและลักษณะการทำงานของข้อคำถาม ผลการวิจัยพบว่า แบบวัดโกรว์ธมายด์เซตที่พัฒนาขึ้นมีคุณภาพด้านความตรงเชิงเนื้อหาโดยมีค่าดัชนีความตรงเชิงเนื้อหาระหว่าง .60 ถึง 1.00 ค่าสัมประสิทธิ์แอลฟาของครอนบาคสูงถึง .91 บ่งชี้ได้ว่าแบบวัดมีความเที่ยงแบบสอดคล้องภายในสูง นอกจากนี้ แองเคอริงวินเยตต์ที่พัฒนาขึ้นมีคุณสมบัติตรงตามสมมติฐานของความเท่าเทียมกันของวินเยตต์ เนื่องจากผู้ตอบส่วนใหญ่รับรู้ลำดับของวินเยตต์ในทิศทางเดียวกัน ประการที่สอง การใช้แองเคอริง

¹ นักศึกษาปริญญาโท สาขาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช

² รองศาสตราจารย์ สาขาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช

³ ผู้ช่วยศาสตราจารย์ สาขาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช

¹ Master's degree student, Educational Measurement and Evaluation Program, Faculty of Education, Sukhothai Thammathirat Open University

² Associate Professor, Educational Measurement and Evaluation Program, Faculty of Education, Sukhothai Thammathirat Open University

³ Assistant Professor, Educational Measurement and Evaluation Program, Faculty of Education, Sukhothai Thammathirat Open University

* Corresponding Author E-mail: siriwat.luno@gmail.com

วินยัตต์เพื่อลดอคติในการตอบ ส่งผลให้แบบวัดมีคุณภาพดีขึ้นจากหลักฐานที่แสดงถึงการเพิ่มขึ้นของค่าสัมประสิทธิ์ความเที่ยง ค่าอำนาจจำแนก โครงสร้างขององค์ประกอบ และความตรงเชิงโครงสร้าง นอกจากนี้ การวิเคราะห์ข้อมูลโดยใช้ทฤษฎีการตอบสนองข้อสอบยังแสดงให้เห็นว่าการใช้มาตรประมาณค่าร่วมกับแองเคอร์วินยัตต์ในการวัดโกรว์ธมายด์เซตช่วยเพิ่มความแม่นยำในการวัดระดับความสามารถของผู้ตอบที่หลากหลาย

คำสำคัญ โกรว์ธมายด์เซต แองเคอร์วินยัตต์ ทฤษฎีการทดสอบแบบดั้งเดิม ทฤษฎีการตอบสนองข้อสอบ

Abstract

This research aimed to: 1) develop a scale and a set of anchoring vignettes to measure growth mindsets for lower secondary school students, and 2) validate the developed scale by applying the anchoring vignettes as supplementary tools to aid students in interpreting and responding to the scale items effectively. The sample for this study comprised 300 lower secondary school students enrolled in the LamPang Primary Educational Service Area Offices 1 to 3. They were selected using a multi-stage random sampling method to ensure representation across the schools. The research utilized a combination of a growth mindsets scale and anchoring vignettes as its primary data collection tools. Data analysis was conducted using Cronbach's alpha coefficient to assess internal consistency, Pearson's correlation coefficient to explore relationships between variables, confirmatory factor analysis to validate the scale structure, and item response theory to evaluate item performance and characteristics. The results of the study revealed two main findings. Firstly, the developed growth mindset scale exhibited acceptable content validity, with a content validity index ranging from .60 to 1.00. Additionally, the high Cronbach's alpha coefficients of .91 indicated a high level of internal consistency within the scale. Furthermore, the analysis confirmed that the developed vignettes met the assumption of vignette equivalence, as respondents perceived the situations described in the vignettes uniformly. Secondly, the utilization of anchoring vignettes to mitigate response biases resulted in an improved quality of growth mindset measurement. This improvement was evidenced by enhancements in Cronbach's alpha coefficients, item discrimination, factor structure, and construct validity. Moreover, the data analysis employing item response theory demonstrated that the incorporation of anchoring vignettes in measuring growth mindsets led to enhanced measurement precision across a wide range of proficiency levels.

Keywords: growth mindsets, anchoring vignettes, classical test theory, item response theory

บทนำ

โกรว์ธมายด์เซตเป็นความเชื่อของบุคคลเกี่ยวกับการปรับปรุงตัวเองด้วยความพยายามและถือเป็นทักษะที่สำคัญสำหรับความสำเร็จ (critical success skill) Dweck ได้กำหนดคำที่ใช้อธิบายความเชื่อที่ว่าสติปัญญาไม่สามารถพัฒนาได้ว่า ฟิกซ์มายด์เซต (fixed mindset) ซึ่งมีลักษณะตรงข้ามกับความเชื่อที่ว่าสติปัญญาสามารถพัฒนาได้หรือที่เรียกว่า โกรว์ธมายด์เซต (growth mindset) บุคคลที่มีโกรว์ธมายด์เซตจะยอมรับความท้าทาย ไม่ย่อท้อต่อความล้มเหลว ตระหนักในความพยายามเรียนรู้ที่นำไปสู่ความสำเร็จ เรียนรู้จากคำวิจารณ์ ค้นหาบทเรียนและแรงบันดาลใจจากความสำเร็จของผู้อื่น (Dweck, 2006) การศึกษาเกี่ยวกับโกรว์ธมายด์เซตในปัจจุบันมีการใช้คำที่หลากหลาย เช่น การใช้คำศัพท์ภาษาอังกฤษ “Growth Mindset” (Wongyai & Patphol, 2020) การใช้คำทับศัพท์ “โกรว์ธมายด์เซต” (Suriyachanhom & Piyakul, 2018) และการแปลความหมายโดยใช้คำว่า “กรอบความคิดแบบเติบโต” (Tippawon, 2019) อย่างไรก็ตามในการศึกษาวิจัยนี้ผู้วิจัยใช้คำทับศัพท์ว่า “โกรว์ธมายด์เซต” เป็นหลักเพื่อความสอดคล้องและชัดเจนในการนำเสนอแนวคิดการวิจัย

โกรว์ธมายด์เซตมีความสำคัญต่อการพัฒนาคุณลักษณะของนักเรียนโดยเฉพาะนักเรียนชั้นมัธยมศึกษาตอนต้น เนื่องจากอยู่ในช่วงวัยที่มีจากการเปลี่ยนแปลงทั้งด้านร่างกาย และฮอร์โมนซึ่งส่งผลต่อการพัฒนาการทางอารมณ์ การใช้เหตุผล และการปรับตัว (Absuksiri, 2022) นอกจากนี้ยังมีความสำคัญต่อการจัดการเรียนการสอนด้านการพัฒนาศักยภาพผู้เรียน (Wongyai & Patphol, 2019) นอกจากนี้โกรว์ธมายด์เซตมีความสำคัญต่อการพัฒนาผลสัมฤทธิ์ทางการเรียนของนักเรียนโดยสามารถทำนายผลสัมฤทธิ์ทางการเรียน (Sutthisomboon, 2020; Stankov et al., 2009; Stankov et al., 2012) และมีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียน (Limsantitham et al., 2019) หากนักเรียนไม่ได้รับการพัฒนาโกรว์ธมายด์เซตเมื่อเขาเผชิญกับความล้มเหลว เขาจะมีความเสี่ยงสูงที่จะเลิกเรียนกลางคันและไม่สามารถพัฒนาศักยภาพของตัวเอง (Rissanen et al., 2019) ถึงแม้ว่าโกรว์ธมายด์เซตจะมีความสำคัญมากแต่ Haimovitz และ Dweck กล่าวว่าการศึกษาในปัจจุบันส่งเสริมฟิกซ์มายด์เซตมากกว่าโกรว์ธมายด์เซตซึ่งอาจส่งผลให้นักเรียนมีภาวะซึมเศร้า เสียกำลังใจ ไม่สนุกกับการเรียนรู้และออกกลางคัน (Westby, 2020)

ด้วยความสำคัญของโกรว์ธมายด์เซตที่กล่าวข้างต้น การวัดโกรว์ธมายด์เซตจึงมีความจำเป็นเพราะจะทำให้ได้ข้อมูลสำหรับพัฒนานักเรียน การสร้างและพัฒนาเครื่องมือวัดโกรว์ธมายด์เซตที่ผ่านมานักวิจัยหลายท่านใช้รูปแบบเครื่องมือที่มีลักษณะเป็นมาตรประมาณค่า เช่น Dweck (2000 as cited in Midkiff et al., 2018) สร้างแบบวัดโกรว์ธมายด์เซตด้วยข้อคำถาม 8 ข้อ ที่มีลักษณะเป็นมาตรประมาณค่า 6 ระดับ Yilmaz (2022) สร้างมาตรวัดโกรว์ธมายด์เซตที่มีลักษณะเป็นมาตรประมาณค่า 4 ระดับ ประกอบด้วย ข้อคำถามโกรว์ธมายด์เซตจำนวน 6 ข้อ และข้อคำถามฟิกซ์มายด์เซต จำนวน 7 ข้อ Midkiff et al., (2018) ศึกษาความเหมาะสมของแบบวัดโกรว์ธมายด์และฟิกซ์มายด์เซตที่มีลักษณะเป็นมาตรประมาณค่า 5 ระดับ ด้วยทฤษฎีการตอบสนองข้อสอบพบว่า แบบวัดโกรว์ธมายด์เซตมีความแม่นยำในการวัดบางช่วงของความสามารถ สำหรับในประเทศไทย Suriyachanhom & Piyakul (2018) สร้างแบบวัดโกรว์ธมายด์เซตด้วยมาตรประมาณค่า 6 ระดับ Tippawon (2019) พัฒนาตัวบ่งชี้กรอบความคิดแบบเติบโตสำหรับนักเรียน

ชั้นประถมศึกษาปีที่ 6 ด้วยมาตรฐานค่า 6 ระดับ Limsantitham et al., (2019) ศึกษาโกรว์ธมายด์เซตในการเรียนวิทยาศาสตร์ด้วยมาตรฐานค่า 6 ระดับ

ถึงแม้ว่างานวิจัยที่ผ่านมาทั้งในประเทศและต่างประเทศจะใช้มาตรฐานค่าในการพัฒนาเครื่องมือวัดการพัฒนาเครื่องมือวัดโกรว์ธมายด์เซต แต่มาตรฐานค่าก็มีข้อจำกัดโดยเฉพาะอาจมีความคลาดเคลื่อนของการวัดเนื่องจากผู้ตอบอาจตีความหมายของมาตรฐานค่าแตกต่างกัน (von Davier, 2010, as cited in von Davier et al., 2018) ทำให้ผลการวัดไม่สามารถนำไปเปรียบเทียบกันได้ เพื่อแก้ไขข้อจำกัดนี้แองเคอริงวินเยตต์จึงถูกพัฒนาขึ้นเพื่อแก้ไขปัญหาการตีความมาตรฐานค่าที่ต่างกันของผู้ตอบ (King et al., 2009) นักวิจัยได้ศึกษาประสิทธิภาพของแองเคอริงวินเยตต์และพบว่า แองเคอริงวินเยตต์ช่วยเพิ่มคุณภาพของแบบวัด (Hopkins & King, 2010; Kyllonen, 2012) แก้ปัญหาความเป็นอัตนัยของแบบวัด (Kyllonen, 2012) และสามารถลดอคติของค่าสัมประสิทธิ์สหสัมพันธ์ซึ่งส่งผลให้โมเดลการวัดมีความเที่ยงและความตรงมากขึ้น (Untong et al., 2018)

ดังนั้น เพื่อให้สามารถวัดโกรว์ธมายด์ได้อย่างน่าเชื่อถือ ผู้วิจัยจึงมุ่งพัฒนาแบบวัดโกรว์ธมายด์เซตโดยสร้างมาตรฐานค่าคู่กับแองเคอริงวินเยตต์เพื่อใช้เป็นจุดยึดในการปรับคะแนนประเมินตนเองเพื่อลดความลำเอียงในการตอบและเปรียบเทียบคุณภาพในมุมมองของทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory, CTT) ที่ยังคงอยู่ในสถานการณ์ของการทดสอบมาเป็นระยะเวลายาวนาน ซึ่งมีข้อจำกัดเกี่ยวกับคะแนนที่มีความคลาดเคลื่อน ค่าพารามิเตอร์ของข้อสอบและแบบสอบที่มีลักษณะขึ้นอยู่กับกลุ่มของผู้ตอบ และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) ที่แก้ไขข้อบกพร่องเบื้องต้นของทฤษฎีการทดสอบแบบดั้งเดิมที่เชื่อว่าความน่าจะเป็นของการตอบข้อสอบได้ถูกต้องขึ้นอยู่กับความสามารถจริงของผู้ตอบ เพื่อหาข้อสรุปเชิงประจักษ์เกี่ยวกับความตรงและการแก้ปัญหาความเป็นอัตนัยของการตอบ (subjective) (Kyllonen, 2012) ในการวัดโกรว์ธมายด์เซตโดยการประยุกต์ใช้แองเคอริงวินเยตต์

ความมุ่งหมายของการวิจัย

การวิจัยนี้มีความมุ่งหมายเพื่อ 1) พัฒนาแบบวัดโกรว์ธมายด์เซตและแองเคอริงวินเยตต์สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น และ 2) ตรวจสอบคุณภาพของแบบวัดโกรว์ธมายด์เซตสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้นด้วยแองเคอริงวินเยตต์

วิธีดำเนินการวิจัย

การดำเนินการวิจัยในครั้งนี้ใช้รูปแบบการวิจัยและพัฒนา เพื่อสร้างแบบวัดโกรว์ธมายด์เซตคู่กับแองเคอริงวินเยตต์โดยใช้แองเคอริงวินเยตต์ปรับความเที่ยงของการวัดภายใต้องค์ประกอบของโกรว์ธมายด์เซต หมายถึง ความเชื่อว่าสติปัญญาสามารถพัฒนาได้ ซึ่งเกิดจากการสั่งสมประสบการณ์ และการฝึกฝน มี 5 องค์ประกอบ (Dweck, 2006) ดังนี้

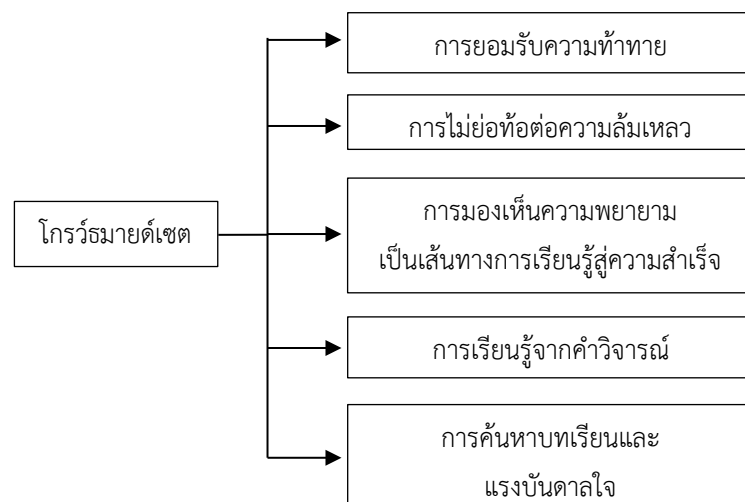
1) การยอมรับความท้าทาย หมายถึง ความพร้อมที่จะเรียนรู้ในสิ่งที่ยากขึ้นและพร้อมเรียนรู้สิ่งใหม่ที่เข้ามาในทุกโอกาส

2) การไม่ย่อท้อต่อความล้มเหลว หมายถึง การไม่ยอมแพ้ต่ออุปสรรคและความล้มเหลวโดยเปลี่ยนความท้อและความล้มเหลวเป็นแรงผลักดันในการพยายามเรียนรู้ ปรับปรุงแก้ไขปัญหา ข้อผิดพลาดที่เกิดขึ้นเพื่อให้ตนเองเกิดการเรียนรู้และทักษะให้มากที่สุด

3) การตระหนักในความพยายามเรียนรู้ที่นำไปสู่ความสำเร็จ หมายถึง การตระหนักและมองหาหนทางที่จะสั่งสมความรู้ ฝึกฝนประสบการณ์ ตลอดจนแสวงหาหนทางและวิธีการที่จะทำสิ่งหนึ่ง ๆ ให้ประสบความสำเร็จเพื่อเป็นการพัฒนาความรู้และศักยภาพของตนเอง

4) การเรียนรู้จากคำวิจารณ์ หมายถึง การพร้อมรับฟังคำแนะนำ ข้อเสนอแนะ รวมถึงคำติชมต่าง ๆ โดยนำข้อมูลที่เกี่ยวข้องกับการพัฒนาตนเอง มาเป็นข้อมูลในการประมวลผลเพื่อแก้ไขและปรับปรุงพัฒนา ศักยภาพของตนเองให้ดียิ่งขึ้น

5) การค้นหาบทเรียนและแรงบันดาลใจจากความสำเร็จของผู้อื่น หมายถึง การเชื่อว่าการประสบความสำเร็จของผู้อื่นแฝงด้วยประสบการณ์ และกระบวนการที่น่าสนใจ และเก็บมาสร้างเป็นแรงบันดาลใจให้กับตนเอง



ภาพ 1 กรอบแนวคิดการวิจัย

ระยะที่ 1 การพัฒนาแบบวัดโกรว์ธมายด์เซตและแองเคอรีંગวินเยตต์ ผู้วิจัยมีวิธีการดำเนินการดังนี้

1. การสร้างแบบวัดโกรว์ธมายด์เซตและแองเคอรี้งวินเยตต์

1.1 กลุ่มเป้าหมายคือ ผู้ทรงคุณวุฒิจำนวน 5 ท่าน ประกอบด้วย ผู้ทรงคุณวุฒิที่มีความสำเร็จในองค์ความรู้และมีประสบการณ์ในการสร้างเครื่องมือโดยประยุกต์ใช้แองเคอรี้งวินเยตต์ 2 ท่าน ผู้สำเร็จในองค์ความรู้และมีประสบการณ์ในการสร้างเครื่องมือวัดโกรว์ธมายด์เซต 2 ท่าน และผู้ที่มีคุณวุฒิทางด้านการวัดและประเมินผลการศึกษา 1 ท่าน

1.2 เครื่องมือที่ใช้ ประกอบด้วย แบบตรวจสอบความตรงเชิงเนื้อหาของแบบวัดโกรว์ธมายด์เซต 5 องค์ประกอบ จำนวน 25 ข้อ ด้วยมาตรประมาณค่า 3 ระดับ (1= เห็นด้วย, 0 = ไม่แน่ใจ, -1 = ไม่เห็นด้วย) และแบบตรวจสอบความสอดคล้องกับนิยามและการเรียงลำดับของวินเยตต์ 5 องค์ประกอบ รวม 15 วินเยตต์ ด้วยมาตรประมาณค่า 3 ระดับ (1 = ต่ำ, 2 = ปานกลาง, 3 = สูง)

1.3 การเก็บรวบรวมข้อมูลโดยนำแบบตรวจสอบความตรงเชิงเนื้อหาของมาตรฐานค่าและ
 แอ่งเคอร์ริงวินเยตต์ส่งให้ผู้ทรงคุณวุฒิพิจารณาทางไปรษณีย์

1.4 การวิเคราะห์ข้อมูลด้วย Index of Item Objective Cognitive Congruence (IOC) และ
 content validity Index (CVI)

2. ตรวจสอบคุณภาพของแบบวัดโกรว์ธมายด์เซตและแอ่งเคอร์ริงวินเยตต์

2.1 ผู้วิจัยดำเนินการตรวจสอบคุณภาพของแบบวัดกับนักเรียนชั้นมัธยมศึกษาตอนต้นในโรงเรียน
 ที่มีลักษณะเป็นโรงเรียนขยายโอกาสสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาและไม่ใช่กลุ่มตัวอย่าง
 จำนวน 2 โรงเรียน โรงเรียนละ 30 คน เพื่อทำการทดลองใช้เครื่องมือจำนวน 2 ครั้ง ซึ่งแต่ละครั้งผู้วิจัยเก็บ
 ข้อมูลจากนักเรียนชั้นมัธยมศึกษาปีที่ 1 ถึง 3 ชั้นละ 10 คน รวม 30 คน

2.2 เครื่องมือที่ใช้ในการวิจัยประกอบด้วย แบบวัดโกรว์ธมายด์เซต จำนวน 25 ข้อ ด้วยมาตร
 ปรมาณค่า 4 ระดับ (1=ไม่เห็นด้วยอย่างยิ่ง, 2=ไม่เห็นด้วย, 3=เห็นด้วย, 4=เห็นด้วยอย่างยิ่ง) และแอ่งเคอร์ริง
 วินเยตต์ 5 ชุดคำถามวินเยตต์ รวม 15 วินเยตต์ แต่ละวินเยตต์ประเมินด้วยมาตรปรมาณค่า 4 ระดับ
 (1=ไม่เห็นด้วยอย่างยิ่ง, 2=ไม่เห็นด้วย, 3=เห็นด้วย, 4=เห็นด้วยอย่างยิ่ง)

2.3 การเก็บรวบรวมข้อมูลโดยขอหนังสือขอความอนุเคราะห์ทดลองใช้เครื่องมือวิจัยจาก
 มหาวิทยาลัยสุโขทัยธรรมาธิราชไปยังโรงเรียนที่ไม่ใช่กลุ่มตัวอย่างเพื่อทดลองใช้เครื่องมือวิจัยครั้งที่ 1 และ 2

2.4 การวิเคราะห์ข้อมูล ประกอบด้วย ค่าสัมประสิทธิ์แอลฟาของครอนบาค และสัดส่วน
 การลำดับวินเยตต์โดยใช้โปรแกรม R เวอร์ชัน R 4.2.2 แพคเกจ mokken, mirt และ ltm

ระยะที่ 2 การตรวจสอบคุณภาพของแบบวัดก่อนและหลังปรับคะแนนโดยใช้แอ่งเคอร์ริงวินเยตต์

1. ประชากรคือ นักเรียนชั้นมัธยมศึกษาตอนต้น สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษา
 จังหวัดลำปาง จำนวน 2,650 คน Joseph et al., (2018) เสนอว่า ตัวอย่างมากกว่า 250 คน เพียงพอต่อการ
 ทดสอบความสอดคล้องกลมกลืนของโมเดลที่มี 25 ตัวแปร ดังนั้นเพื่อให้ตัวอย่างมีความเป็นตัวแทนของแต่ละ
 เขตพื้นที่อย่างเท่าเทียม ผู้วิจัยจึงกำหนดตัวอย่างเขตพื้นที่ละ 100 คน รวม 300 คน ได้จากการสุ่มตัวอย่าง
 แบบหลายขั้นตอนดังนี้

1.1 แบ่งเขตพื้นที่การศึกษาโดยใช้เขตพื้นที่เป็นเกณฑ์ได้ 3 เขตพื้นที่ ประกอบด้วย สำนักงาน
 เขตพื้นที่การศึกษาประถมศึกษาลำปางเขต 1 สำนักงานเขตพื้นที่การศึกษาประถมศึกษาลำปางเขต 2 และ
 สำนักงานเขตพื้นที่การศึกษาประถมศึกษาลำปางเขต 3

1.2 เลือกตัวอย่างแบบง่าย (sample random sampling) โดยการจับสลากได้โรงเรียนจากแต่ละ
 เขตพื้นที่ละ 5 โรงเรียน รวม 15 โรงเรียน

1.3 เลือกตัวอย่างแบบไม่เป็นสัดส่วน (non-proportional sampling) เนื่องจากนักเรียนในแต่ละ
 ชั้นของแต่ละโรงเรียนมีจำนวนที่แตกต่างกัน บางกลุ่มตัวอย่างในประชากรมีขนาดเล็กมาก ผู้วิจัยจึงใช้การ
 สุ่มตัวอย่างแบบไม่เป็นสัดส่วนซึ่งจะทำให้มั่นใจได้ว่าตัวอย่างมีความเป็นตัวแทนที่เพียงพอสำหรับการวิเคราะห์
 ข้อมูลจึงกำหนดตัวอย่างโรงเรียนละ 20 คน

1.4 ครูผู้ประสานงานของโรงเรียนเป็นผู้เลือกตัวอย่างตามความสะดวก (convenience
 sampling) ประกอบด้วย นักเรียนชั้นมัธยมศึกษาปีที่ 1 ถึง 3 รวม 20 คน

2. เครื่องมือที่ใช้ในการวิจัยประกอบด้วย แบบวัดโกรว์ธมายด์เซตแสดงดังภาพ 2 และแองเคอริง วินเยตต์แสดงดังภาพ 3

ข้อ	ข้อความ	ความคิดเห็น			
		(1) ไม่เห็นด้วย อย่างยิ่ง	(2) ไม่เห็นด้วย	(3) เห็นด้วย	(4) เห็นด้วย อย่างยิ่ง
1	เมื่อเรียนเนื้อหาที่ยากขึ้น ฉันคิดวิธีการต่าง ๆ ที่จะทำให้เข้าใจมากยิ่งขึ้น				
2	เมื่อครูสอนเนื้อหาที่ยากขึ้น ฉันตั้งใจฟังมากกว่าเดิม				
3	เมื่อมีโอกาสเข้ามา ฉันพร้อมเรียนรู้สิ่งใหม่ๆ				
4	เมื่อครูมอบหมายงานที่ยากและท้าทายความสามารถ ฉันจะกระตือรือร้นทำงานจนสำเร็จ				
5	ฉันสนุกเมื่อได้เรียนรู้สิ่งใหม่ ๆ แม้จะเป็นเรื่องยากก็ตาม				

ภาพ 2 ตัวอย่างข้อคำถามของแบบวัดโกรว์ธมายด์เซต องค์ประกอบที่ 1 การยอมรับความท้าทาย

สถานการณ์ชุดที่ 1

ก) มานะเห็นว่าการเขียนโค้ดคำสั่งควบคุมการทำงานของหุ่นยนต์ที่ครูสอนเป็นเรื่องที่ยาก เขาจึงขอให้ครูสอนเขียนโค้ดอย่างอื่นที่ง่ายกว่านี้และทำงานตามที่ได้รับมอบหมาย

นักเรียนเห็นด้วยหรือไม่ว่ามานะเป็นคนที่ยอมรับความท้าทาย/ความยากลำบาก

☐ ไม่เห็นด้วยอย่างยิ่ง ☐ ไม่เห็นด้วย ☐ เห็นด้วย ☐ เห็นด้วยอย่างยิ่ง

ข) มานะเห็นว่าการเขียนโค้ดคำสั่งควบคุมการทำงานของหุ่นยนต์ที่ครูสอนเป็นเรื่องที่ยากและไม่ได้ใช้ในชีวิตประจำวัน เขาจึงไม่ตั้งใจและไม่ตั้งใจทำงานที่ครูมอบหมาย

นักเรียนเห็นด้วยหรือไม่ว่ามานะเป็นคนที่ยอมรับความท้าทาย/ความยากลำบาก

☐ ไม่เห็นด้วยอย่างยิ่ง ☐ ไม่เห็นด้วย ☐ เห็นด้วย ☐ เห็นด้วยอย่างยิ่ง

ค) มานพเห็นว่าการเขียนโค้ดคำสั่งควบคุมการทำงานของหุ่นยนต์ที่ครูสอนเป็นเรื่องที่ยาก แต่เป็นความรู้ใหม่ที่น่าสนใจและอาจจะนำความรู้ไปใช้ในอนาคตได้ เขาจึงตั้งใจทำงานที่ครูมอบหมายและซักถามเมื่อเกิดข้อสงสัย

นักเรียนเห็นด้วยหรือไม่ว่ามานพเป็นคนที่ยอมรับความท้าทาย/ความยากลำบาก

☐ ไม่เห็นด้วยอย่างยิ่ง ☐ ไม่เห็นด้วย ☐ เห็นด้วย ☐ เห็นด้วยอย่างยิ่ง

ภาพ 3 ตัวอย่างแองเคอริงวินเยตต์ ชุดที่ 1 การยอมรับความท้าทาย

3. การเก็บรวบรวมข้อมูลโดยขอหนังสือขอความอนุเคราะห์ใช้เครื่องมือวิจัยจากสาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราชไปยังโรงเรียนกลุ่มตัวอย่างโดยเดินทางไปยังโรงเรียนด้วยตนเอง และจัดส่งเครื่องมือทางไปรษณีย์ในกรณีร่นมีการเดินทางเกิน 100 กิโลเมตร จากจำนวนตัวอย่าง 300 คน ผู้วิจัยได้ข้อมูลตอบกลับจำนวน 300 คน คิดเป็นร้อยละ 100

4. การวิเคราะห์ข้อมูล ประกอบด้วย ค่าอำนาจจำแนก ค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's alpha) ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson's correlation coefficient) การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis : EFA) การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) และทฤษฎีการตอบสนองข้อสอบด้วย Graded response model (GRM)

ผลการวิจัย

1. ผลการพัฒนาแบบวัดโครงข่ายมาดซ์เซตและแองเคอร์ริงวินเยตต์ พบว่า แบบวัดมีค่าสัมประสิทธิ์ความเที่ยงทั้งฉบับเท่ากับ .91 รายองค์ประกอบระหว่าง .68 ถึง .78 ค่าอำนาจจำแนกรายข้อระหว่าง .36 ถึง .59 ผลการพัฒนาแองเคอร์ริงวินเยตต์พบว่า ความเป็นลำดับของวินเยตต์ของแต่ละชุดคำถามมีค่าเป็นบวก เมื่อพิจารณาการตอบลำดับของวินเยตต์พบว่า ผู้ตอบ 97% ถึง 98% ของผู้ตอบทั้งหมดตอบลำดับตามธรรมชาติของวินเยตต์ โดยชุดคำถามที่ตอบตามตามธรรมชาติของวินเยตต์มากที่สุดคือ การค้นหาบทเรียนและแรงบันดาลใจจากความสำเร็จของผู้อื่น 98% รองลงมา คือ การมองเห็นความพยายามเป็นเส้นทางสู่ความเชี่ยวชาญ 97% การยอมรับความท้าทาย 97.66% การไม่ย่อท้อต่อความล้มเหลว 97.33% และการมองเห็นความพยายามเป็นเส้นทางสู่ความเชี่ยวชาญ 97% ดังตาราง 1

ตาราง 1 ผลการวิเคราะห์การตอบลำดับของวินเยตต์ และความเป็นลำดับของวินเยตต์ ($n = 300$)

ลำดับของวินเยตต์		การตอบลำดับของวินเยตต์			ความเป็นลำดับของวินเยตต์			ร้อยละ การไม่ละเมิดลำดับ ของวินเยตต์
		1 (ต่ำ)	2 (ปานกลาง)	3 (สูง)	1	2	3	
GM1	1 (ต่ำ)	-	.753	.967	-	.747	.963	.97
	2 (ปานกลาง)	.007	-	.797	.240	-	.780	
	3 (สูง)	.003	.017	-	.030	.187	-	
GM2	1 (ต่ำ)	-	.547	.963	-	.523	.960	.93
	2 (ปานกลาง)	.023	-	.887	.430	-	.883	
	3 (สูง)	.003	.003	-	.033	.110	-	
GM3	1 (ต่ำ)	-	.967	.970	-	.926	.970	.97
	2 (ปานกลาง)	0	-	.233	.033	-	.193	
	3 (สูง)	0	.030	-	.030	.747	-	
GM4	1 (ต่ำ)	-	.543	.917	-	.500	.907	.94
	2 (ปานกลาง)	.043	-	.847	.413	-	.833	
	3 (สูง)	.010	.013	-	.073	.140	-	
GM5	1 (ต่ำ)	-	.877	.947	-	.863	.947	.98
	2 (ปานกลาง)	.013	-	.593	.110	-	.587	
	3 (สูง)	0	.007	-	.053	.400	-	

หมายเหตุ ร้อยละการไม่ละเมิดลำดับของวินเยตต์พิจารณาค่า Nviolation เท่ากับ 0 (ศูนย์)

จากการตอบวินเยตต์ในแต่ละลำดับของวินเยตต์พบว่า ผู้ตอบส่วนใหญ่ให้คะแนนของวินเยตต์สูงกว่าเมื่อวินเยตต์มีพฤติกรรมที่เข้มข้นเมื่อพิจารณาอย่างละเอียดพบว่า การมองเห็นความพยายามเป็นเส้นทางสู่ความเชี่ยวชาญ (GM3) ผู้ตอบส่วนใหญ่ 74% ให้คะแนนวินเยตต์ปานกลางเท่ากับวินเยตต์สูง

2. ผลการตรวจสอบคุณภาพของแบบวัดก่อนและหลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์

ก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ ผู้วิจัยปรับชุดคำตอบที่ละเมิดลำดับของวินเยตต์ให้มีลำดับตามธรรมชาติของวินเยตต์โดยใช้วิธีปรับคะแนนของ PISA 2012 จำนวน 47 ชุดคำตอบ โดยชุดคำตอบ

วินยัตต์ที่มีการปรับลำดับของวินยัตต์มากที่สุดคือ การเรียนรู้จากคำวิจารณ์จำนวน 17 ชุดคำตอบ รองลงมา คือ การมองเห็นความพยายามเป็นเส้นทางสู่ความสำเร็จ จำนวน 9 ชุดคำตอบ การไม่ย่อท้อต่อความล้มเหลว จำนวน 8 ชุดคำตอบ การยอมรับความท้าทายจำนวน 7 ชุดคำตอบ และการค้นหาบทเรียนและแรงบันดาลใจจากความสำเร็จของผู้อื่นจำนวน 6 ชุดคำตอบ ตามลำดับ จากนั้นปรับคะแนนการประเมินตนเองโดยใช้แองเคอร์ริงวินยัตต์ โดยนำเสนอผลการวิเคราะห์ตรวจสอบคุณภาพด้วยทฤษฎีการทดสอบแบบดั้งเดิม และทฤษฎีการตอบสนองข้อสอบ ดังนี้

2.1 ผลการตรวจสอบคุณภาพด้วยทฤษฎีการทดสอบแบบดั้งเดิม

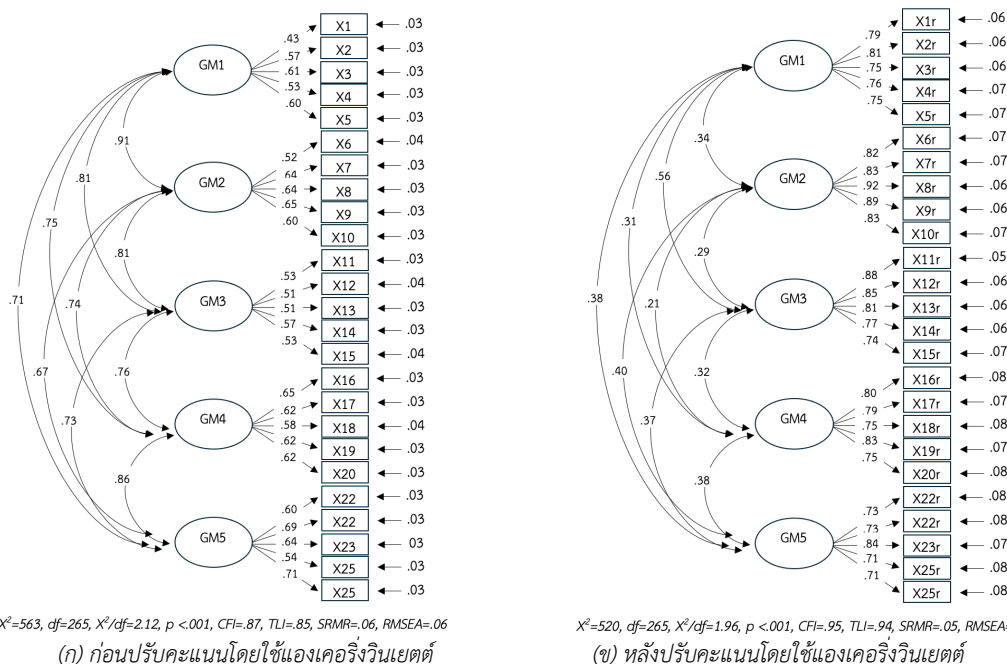
2.1.1 คุณภาพด้านความตรงและความเที่ยงก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินยัตต์ ค่าดัชนีความเที่ยงทั้งฉบับเท่ากับ .91 รายองค์ประกอบระหว่าง .66 ถึง .78 ค่าอำนาจจำแนกรายข้อระหว่าง .36 ถึง .66 เมื่อปรับคะแนนโดยใช้แองเคอร์ริงวินยัตต์ค่าดัชนีความเที่ยงทั้งฉบับเท่ากับ .91 รายองค์ประกอบระหว่าง .86 ถึง .93 และค่าอำนาจจำแนกรายข้อระหว่าง .45 ถึง .58 แสดงว่าการปรับคะแนนโดยใช้แองเคอร์ริงวินยัตต์ทำให้ค่าดัชนีความเที่ยงองค์ประกอบและค่าอำนาจจำแนกสูงขึ้น

2.1.2 ผลการทดสอบความสัมพันธ์ระหว่างข้อคำถามที่อยู่ในองค์ประกอบเดียวกันและต่างองค์ประกอบด้วยสูตรสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันพบว่า แบบวัดก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินยัตต์มีขนาดความสัมพันธ์ของตัวแปรระหว่าง .071 ถึง .508 ขนาดความสัมพันธ์ของตัวแปรภายในองค์ประกอบมีค่าระหว่าง .14 ถึง .50 ซึ่งตัวแปรส่วนใหญ่มีขนาดความสัมพันธ์ภายในองค์ประกอบต่ำกว่าความสัมพันธ์กับตัวแปรต่างองค์ประกอบ เมื่อปรับคะแนนโดยใช้แองเคอร์ริงวินยัตต์พบว่า ขนาดความสัมพันธ์ระหว่างตัวแปรมีค่าเพิ่มขึ้นโดยมีค่าระหว่าง .075 ถึง .831 ขนาดความสัมพันธ์ของตัวแปรภายในองค์ประกอบมีค่าระหว่าง .45 ถึง .83 ซึ่งตัวแปรทุกคู่ที่อยู่ในองค์ประกอบเดียวกันมีขนาดความสัมพันธ์สูงกว่าตัวแปรต่างองค์ประกอบ แสดงว่า การปรับคะแนนโดยใช้แองเคอร์ริงวินยัตต์ทำให้ตัวแปรภายในองค์ประกอบมีขนาดความสัมพันธ์เพิ่มมากขึ้น

2.1.3 คุณภาพด้านความตรงเชิงโครงสร้างหลังจากเก็บข้อมูลกับนักเรียนจำนวน 300 คน ก่อนวิเคราะห์องค์ประกอบเชิงยืนยัน ผู้วิจัยวิเคราะห์องค์ประกอบเชิงสำรวจ ด้วยการวิเคราะห์ปัจจัยร่วม (Common Factor Analysis) เพื่อตรวจสอบการจัดองค์ประกอบของตัวแปรตามโครงสร้างที่ผู้วิจัยคาดไว้ พบว่าแบบวัดก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินยัตต์สามารถจัดองค์ประกอบได้ 3 องค์ประกอบ ค่าน้ำหนักองค์ประกอบระหว่าง .417 ถึง .695 ซึ่งโครงสร้างของแบบวัดไม่เป็นไปตามที่ผู้วิจัยคาดไว้ หลังปรับคะแนนโดยใช้แองเคอร์ริงวินยัตต์สามารถจัดองค์ประกอบได้ 5 องค์ประกอบ ค่าน้ำหนักองค์ประกอบระหว่าง .652 ถึง .887 ซึ่งเป็นไปตามโครงสร้างที่ผู้วิจัยคาดไว้ แสดงว่าการปรับคะแนนโดยใช้แองเคอร์ริงวินยัตต์ทำให้การจัดองค์ประกอบของตัวแปรมีความชัดเจนมากขึ้นโดยสามารถจัดองค์ประกอบได้ตรงตามที่ผู้วิจัยคาดไว้

ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งเพื่อพิจารณาความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ น้ำหนักองค์ประกอบ ความตรงเชิงลู่เข้า (convergent validity, AVE) และความตรงเชิงจำแนก (discriminant validity) พบว่า โมเดลการวัดหลังปรับคะแนนด้วยแองเคอร์ริงวินยัตต์มีความสอดคล้องกับข้อมูลเชิงประจักษ์ ค่าน้ำหนักองค์ประกอบของข้อคำถามระหว่าง .71 ถึง .92 χ^2 เท่ากับ 520 df เท่ากับ

265 χ^2/df เท่ากับ 1.96 CFI เท่ากับ .95 TLI เท่ากับ .94 SRMR เท่ากับ .05 RMSEA เท่ากับ .06 ส่วนโมเดลการวัดของแบบวัดโกรว์ธมายด์เซตก่อนปรับคะแนนโดยใช้แองเคอรี่งวินเยตต์ไม่สอดคล้องกับข้อมูลเชิงประจักษ์ ค่าน้ำหนักองค์ประกอบของข้อคำถามระหว่าง .43 ถึง .69 ค่า χ^2 เท่ากับ 563 df เท่ากับ 265 χ^2/df เท่ากับ 2.12 CFI เท่ากับ .87 TLI เท่ากับ .85 SRMR เท่ากับ .06 RMSEA เท่ากับ .06 แสดงว่าโมเดลหลังปรับคะแนนโดยใช้แองเคอรี่งวินเยตต์สอดคล้องกับข้อมูลเชิงประจักษ์ ค่าพารามิเตอร์ตรงกับทฤษฎีโดยไม่ปรับโมเดลสามารถอธิบายได้ว่า ตัวแปรสังเกตได้ ทั้ง 25 ตัวแปรสามารถนำมาใช้วัดคุณลักษณะโกรว์ธมายด์เซตได้ แสดงดังภาพ 4



ภาพ 4 โมเดลการวิเคราะห์องค์ประกอบก่อนและหลังปรับคะแนนโดยใช้แองเคอรี่งวินเยตต์

ตาราง 2 ความตรงเชิงลู่เข้า ความตรงเชิงจำแนก และค่าสัมประสิทธิ์ความเที่ยงขององค์ประกอบ

	ก่อนปรับคะแนนโดยใช้แองเคอรี่งวินเยตต์						หลังปรับคะแนนโดยใช้แองเคอรี่งวินเยตต์					
	GM1	GM2	GM3	GM4	GM5	CR	GM1	GM2	GM3	GM4	GM5	CR
GM1	.30					.98	.60					.99
GM2	.82	.37				.98	.11	.74				.99
GM3	.65	.65	.28			.97	.32	.08	.66			.99
GM4	.60	.55	.58	.38		.98	.09	.04	.10	.62		.99
GM5	.52	.44	.53	.74	.41	.98	.14	.16	.14	.14	.56	.99

หมายเหตุ ค่าในแนวทแยง คือ ค่า AVE ค่านอกแนวทแยง คือ ค่าสัมประสิทธิ์สหสัมพันธ์ภายในองค์ประกอบ

จากตาราง 2 พบว่า ค่าความตรงเชิงลู่เข้าขององค์ประกอบก่อนปรับคะแนนโดยใช้แองเคอรี่งวินเยตต์มีค่าน้อยกว่าเกณฑ์ของความตรงเชิงลู่เข้า ($AVE \geq .50$) ส่วนองค์ประกอบหลังปรับคะแนนโดยใช้แองเคอรี่งวินเยตต์มีค่า AVE มากกว่า .50 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้หรือเพียงพอที่จะประเมินว่ามีความตรงเชิงลู่เข้า เมื่อพิจารณาความตรงเชิงจำแนก โดยการเปรียบเทียบค่า AVE ของตัวแปรแฝงกับค่ากำลังสองของ

ความสัมพันธ์ของตัวแปรแฝงนั้นกับตัวแปรแฝงอื่น พบว่า องค์ประกอบก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ไม่ผ่านเกณฑ์ที่ยอมรับได้ (ค่าในแนวทแยงน้อยกว่าค่านอกแนวทแยง) ส่วนองค์ประกอบหลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์มีค่าความตรงเชิงจำแนกผ่านเกณฑ์ที่ยอมรับได้ (ค่าในแนวทแยงมากกว่าค่านอกแนวทแยง) และเมื่อพิจารณาความเที่ยงของตัวแปรแฝงโดยใช้เกณฑ์ที่ยอมรับได้ที่มากกว่า .60 พบว่า แบบวัดก่อนและหลังปรับคะแนนมีค่าผ่านเกณฑ์ที่ยอมรับได้

2.2 ผลการตรวจสอบคุณภาพด้วยทฤษฎีการตอบสนองข้อสอบ

ผลการทดสอบความเป็นเอกมิติจากการพิจารณาค่าความแปรปรวนขององค์ประกอบแรก อธิบายความแปรปรวนทั้งหมดได้อย่างน้อย 20% (Reckase, 1979) พบว่า แบบวัดก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์มีค่าความแปรปรวนขององค์ประกอบแรกอธิบายความแปรปรวนได้เท่ากับ 31.3% ซึ่งมีค่าน้อยกว่าหลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์โดยค่าความแปรปรวนขององค์ประกอบแรกอธิบายความแปรปรวนได้เท่ากับ 32.4% ซึ่งค่าความแปรปรวนขององค์ประกอบแรกหลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์มีค่ามากกว่าก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ 1.2% แสดงว่าหลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์มีความเป็นเอกมิติมากกว่าก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์อย่างมีนัยสำคัญทางสถิติ เมื่อตรวจสอบความเป็นอิสระต่อกันของข้อคำถาม พบว่า ค่า χ^2/df ระหว่าง .82 ถึง 1.87 และค่า $RMSEA$ ระหว่าง .00 ถึง .05 แสดงถึงข้อคำถามไม่มีความเป็นอิสระต่อกัน เนื่องจากค่า $RMSEA$ ควรน้อยกว่า .05 ส่วนแบบวัดหลังปรับค่าคะแนนโดยใช้แองเคอร์ริงวินเยตต์มีค่า χ^2/df ระหว่าง .88 ถึง 1.49 และค่า $RMSEA$ ระหว่าง .00 ถึง .04 แสดงว่าการปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ทำให้ข้อคำถามมีความเป็นอิสระต่อกันมากขึ้น

ผลการวิเคราะห์ด้วยโมเดล GRM พบว่า แบบวัดก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ค่าความชันร่วมของข้อคำถามระหว่าง .89 ถึง 1.98 ค่าพารามิเตอร์เทรซโฮลด์ของแต่ละรายการคำตอบระหว่าง -6.049 ถึง 1.770 หลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์พบว่า ค่าความชันร่วมของข้อคำถามระหว่าง .930 ถึง 1.579 ค่าพารามิเตอร์เทรซโฮลด์ของแต่ละรายการคำตอบระหว่าง -6.019 ถึง 3.998 ดังตาราง 3

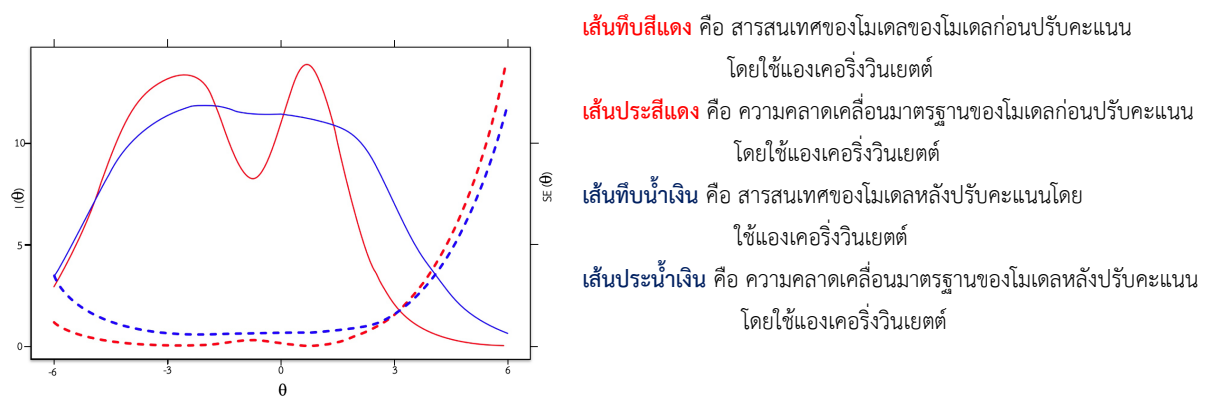
ตาราง 3 ค่าความชันร่วมของข้อคำถาม(α) และพารามิเตอร์เทรซโฮลด์ของแต่ละรายการคำตอบ(β)

ข้อที่	ก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์				หลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์						
	α	β_1	β_2	β_3	α	β_1	β_2	β_3	β_4	β_5	β_6
X1	0.89	-6.049	-3.82	1.77	.93	-6.019	-3.271	-2.887	-.231	1.033	3.998
X2	1.38	-4.759	-3.34	.45	1.492	-4.814	-2.608	-2.376	-.532	.179	2.053
X3	1.42	-4.756	-3.35	.51	1.316	-4.661	-2.902	-2.441	-.554	.171	2.172
X4	1.22	-4.081	-2.14	1.48	1.108	-4.009	-2.512	-1.967	-.188	.94	3.295
X5	1.71	-4.268	-1.92	.99	1.4	-4.177	-2.274	-1.92	-.155	.686	2.521
X6	1.06	-5.271	-3.07	.48	.99	-4.497	-2.829	-2.426	-1.434	-.369	1.984
X7	1.42	-4.757	-2.62	.47	1.047	-4.535	-2.896	-2.509	-1.389	-.292	2.075
X8	1.37	-4.325	-2.74	.83	1.179	-3.729	-2.602	-2.29	-1.286	-.03	2.005
X9	1.98	-3.95	-2.35	.45	1.3	-3.778	-2.468	-2.171	-1.239	-.202	1.74
X10	1.88	-1.88	.97		1.416	-3.671	-2.089	-1.716	-.959	.12	2.264
X11	1.11	-3.512	.61		1.181	-5.587	-2.973	-2.827	-2.433	-.699	1.631
X12	1.02	-2.826	.93		1.206	-4.914	-2.709	-2.591	-1.933	-.383	1.649

ตาราง 3 (ต่อ)

ข้อที่	ก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์				หลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์						
	α	β_1	β_2	β_3	α	β_1	β_2	β_3	β_4	β_5	β_6
X13	1.03	-6.05	-3.54	.93	1.119	-5.178	-2.974	-2.839	-2.46	-.517	1.623
X14	1.42	-3.06	.52		1.489	-4.771	-2.506	-2.387	-2.061	-.566	1.351
X15	1.40	-1.902	1.11		1.576	-3.856	-1.961	-1.892	-1.55	-.125	1.531
X16	1.67	-3.838	-1.92	.85	1.102	-3.499	-1.694	-1.457	-.11	.564	3.334
X17	1.69	-2.666	.58		1.089	-4.028	-2.117	-1.839	-.357	.304	3.553
X18	1.48	-3.785	-1.34	1.28	1.18	-2.96	-1.175	-.965	.203	.941	3.578
X19	1.84	-2.479	.93		1.101	-3.866	-1.959	-1.709	-.124	.772	3.342
X20	1.55	-4.516	-2.70	.62	1.041	-3.801	-2.327	-1.943	-.437	.375	3.038
X21	1.30	-3.971	-2.19	1.14	1.282	-3.893	-2.067	-1.705	.588	.893	2.181
X22	1.70	-3.791	-1.81	.93	1.366	-3.775	-1.967	-1.725	.594	.83	2.059
X23	1.50	-3.82	-1.78	1.06	1.516	-3.193	-1.856	-1.498	.554	.771	1.86
X24	1.39	-4.261	-2.12	1.06	1.351	-3.609	-2.064	-1.662	.557	.767	2.139
X25	2.40	-3.224	-1.83	.52	1.373	-3.584	-2.163	-1.772	.303	.473	1.795
-2Log Likelihood = 10565.176					-2Log Likelihood = 20782.298						

จากตาราง 3 ค่าถามของแบบวัดโกรว์ธมายด์เซตก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ข้อที่ 10, 11, 12, 14, 15, 17 และ 19 มีค่าพารามิเตอร์ทรซโซล์ของรายการคำตอบ 2 ตำแหน่งคือ β_1 และ β_2 แสดงว่าไม่สามารถวัดความสามารถของผู้สอบตำแหน่ง β_3 ส่งผลให้สารสนเทศของแบบวัดช่วงความสามารถ - 2 ถึง .5 ลดต่ำลงและความคลาดเคลื่อนมาตรฐานช่วงดังกล่าวเพิ่มสูงขึ้น ดังภาพ 5



ภาพ 5 สารสนเทศและความคลาดเคลื่อนมาตรฐานของโมเดล

ผลการวิเคราะห์ค่าสารสนเทศของแบบวัดพบว่า แบบวัดก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ให้สารสนเทศในช่วง -4 ถึง -2 และช่วง .5 ถึง 1 ส่วนแบบวัดหลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ให้สารสนเทศในช่วง -4 ถึง 2 แสดงว่าหลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ทำให้ผลการวัดมีความแม่นยำในกลุ่มผู้สอบที่มีความสามารถหลากหลายกว่าก่อนปรับคะแนน

อภิปรายผล

การวิจัยในครั้งนี้มุ่งศึกษาคุณภาพของการวัดของแบบวัดโกรว์ธมายด์เซตโดยใช้แองเคอร์ริงวินเยตต์ ปรับคะแนนการประเมินตนเองของนักเรียนชั้นมัธยมศึกษาตอนต้น-ผลการวิจัยชี้ให้เห็นว่าการใช้แองเคอร์ริงวินเยตต์ปรับคะแนนประเมินตนเองส่งผลให้แบบวัดมีคุณภาพมากขึ้น โดยช่วยเพิ่มความตรงและความเที่ยงของการวัด ซึ่งผู้วิจัยลำดับการอภิปรายผลตามประเด็นสำคัญที่ค้นพบตามวัตถุประสงค์การวิจัย ดังนี้

1. ผลการพัฒนาแบบวัดโกรว์ธมายด์เซตและแองเคอร์ริงวินเยตต์

แองเคอร์ริงวินเยตต์ที่พัฒนาขึ้นพบว่า ผู้ตอบ 97% ถึง 98% ของผู้ตอบทั้งหมดตอบลำดับตามธรรมชาติของวินเยตต์ โดยชุดคำถามที่ตอบตามตามธรรมชาติของวินเยตต์มากที่สุดคือ การค้นหาบทเรียนและแรงบันดาลใจจากความสำเร็จของผู้อื่น 98% รองลงมา คือ การมองเห็นความพยายามเป็นเส้นทางสู่ความสำเร็จ 97% การยอมรับความท้าทาย 97.66% การไม่ย่อท้อต่อความล้มเหลว 97.33% และการมองเห็นความพยายามเป็นเส้นทางสู่ความสำเร็จ 97% ซึ่งบ่งบอกถึงการเข้าใจความเข้มของวินเยตต์ที่ชัดเจน สอดคล้องกับฐานสมมติฐานของการสร้างแองเคอร์ริงวินเยตต์ที่ผู้ตอบแบบสอบถามเข้าใจลำดับของวินเยตต์อย่างเข้มงวดในลักษณะเดียวกัน (von Davier et al., 2018) และความเท่าเทียมกันของวินเยตต์ (King et al., 2004) อย่างไรก็ตามสัดส่วนการตอบในด้านการมองเห็นความพยายามเป็นเส้นทางสู่ความสำเร็จพบว่า ผู้ตอบส่วนใหญ่รับรู้ว่ามีวินเยตต์ปานกลางมีความเข้มเท่ากับวินเยตต์สูงซึ่งอาจบ่งบอกถึงการใช้ภาษาที่ผู้วิจัยควรระมัดระวังในการเขียนวินเยตต์ให้สามารถอธิบายให้เห็นถึงระดับที่แท้จริงของแต่ละวินเยตต์ (King et al., 2004) อย่างไรก็ตามการพิจารณาคุณภาพของแองเคอร์ริงวินเยตต์ไม่ได้ขึ้นอยู่กับสัดส่วนการตอบลำดับของวินเยตต์ที่ถูกต้องแต่เป็นสัดส่วนความเข้าใจวินเยตต์ของผู้ตอบโดยเฉลี่ยในลักษณะเดียวกันถึงแม้ผู้ตอบจะมีรูปแบบการตอบที่ต่างกัน (King et al., 2004) ทั้งนี้การตรวจสอบความเข้าใจลำดับวินเยตต์ของผู้ตอบอย่างเท่าเทียมกันยังคงมีความสำคัญเพื่อให้แน่ใจว่าผู้ตอบส่วนใหญ่ตีความวินเยตต์ได้อย่างถูกต้องโดยการตรวจสอบซ้ำและปรับปรุงการแสดงออกทางภาษาช่วยลดสัดส่วนการจัดลำดับวินเยตต์ที่ผิดพลาดโดยผลการปรับปรุงความเข้มของวินเยตต์ส่งผลให้ผู้ตอบสามารถจัดลำดับวินเยตต์ได้ถูกต้องมากขึ้นและสะท้อนถึงวิธีการที่สามารถแก้ปัญหาเกี่ยวกับความเข้าใจลำดับวินเยตต์ของผู้ตอบที่ชัดเจนมากยิ่งขึ้น (Wand et al., 2011)

ถึงแม้ว่าการใช้แองเคอร์ริงวินเยตต์ในการพัฒนาแบบวัดโกรว์ธมายด์เซตในงานวิจัยนี้จะช่วยให้การวัดมีคุณภาพมากขึ้น การนำแองเคอร์ริงวินเยตต์ไปใช้ในการพัฒนาเครื่องมือวัดก็ยังคงมีความท้าทายและจำเป็นต้องพัฒนาแองเคอร์ริงวินเยตต์ให้มีคุณภาพ การพัฒนาแองเคอร์ริงวินเยตต์ให้มีคุณภาพสูงและความเท่าเทียมกันเป็นประเด็นที่ยากและผู้วิจัยต่างๆ ให้ความสำคัญ การศึกษาต่าง ๆ ก่อนหน้านี้แสดงให้เห็นว่าสัดส่วนการตอบที่ไม่สอดคล้องกับลำดับของวินเยตต์ที่ค่อนข้างสูง ซึ่งบ่งบอกถึงความจำเป็นในการปรับปรุงการเขียนและการนำเสนอวินเยตต์ให้มีความชัดเจน โดยพบว่า 25% ของรูปแบบการตอบใน PISA (Kyllonen & Bertling, 2014), 18% ถึง 32% ของผู้ตอบแบบสอบถามในการวัดบุคลิกภาพห้าด้าน (Primi et al., 2016) และ 8% ถึง 35% ของผู้ตอบแบบสอบถามเรื่องความมีสติระดับนานาชาติของ Möttus et al. (2012) ไม่ได้ให้คะแนนที่สอดคล้องกับลำดับของวินเยตต์ และพบว่ารูปแบบการตอบแองเคอร์ริงวินเยตต์ของ Janprasert (2019) ผู้ตอบละเมิดลำดับของวินเยตต์ 8.5% ถึง 15.6% ซึ่งแสดงให้เห็นว่า การพัฒนาแองเคอร์ริงวินเยตต์ให้คุณภาพสูงอาจ

เป็นเรื่องยากพอ ๆ กับการพัฒนาคำถามประเมินตนเองให้มีคุณภาพ (King et al., 2004) ดังนั้นการปรับปรุงวินยัตต์โดยการตรวจสอบซ้ำทำให้ผู้ตอบเข้าใจลำดับวินยัตต์ได้ถูกต้องและมีความสอดคล้องกันมากขึ้น จะแสดงถึงการเพิ่มประสิทธิภาพของแองเคอร์วินยัตต์ (Wand et al., 2011)

2. ผลการตรวจสอบคุณภาพของแบบวัดก่อนและหลังปรับคะแนนโดยใช้แองเคอร์วินยัตต์

ผลการตรวจสอบคุณภาพของแบบวัดก่อนและหลังปรับคะแนนโดยใช้แองเคอร์วินยัตต์ ในมุมมองทฤษฎีการทดสอบแบบดั้งเดิมโดยการวิเคราะห์ค่าสัมประสิทธิ์ความเที่ยง ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน การวิเคราะห์องค์ประกอบเชิงสำรวจ การวิเคราะห์องค์ประกอบเชิงยืนยัน พบว่า แบบวัดที่ปรับคะแนนโดยใช้แองเคอร์วินยัตต์มีความเที่ยงทั้งฉบับเท่ากับแบบวัดก่อนปรับแก้คะแนน แต่แบบวัดหลังปรับคะแนนมีค่าสัมประสิทธิ์ความเที่ยงองค์ประกอบสูงกว่าแบบวัดก่อนปรับคะแนนอย่างมีนัยสำคัญ แสดงให้เห็นว่าการปรับแก้คะแนนช่วยให้คะแนนที่ได้มีความเที่ยงในการวัดมากขึ้น ทั้งนี้อาจเป็นผลมาจากวินยัตต์มีความชัดเจนและมีเป็นลำดับของวินยัตต์สูง เมื่อปรับแก้คะแนนจึงมีความเที่ยงตรงในการวัดได้ดีขึ้นแสดงถึงประสิทธิภาพในการแก้ปัญหาลักษณะของผู้ตอบที่มีความแตกต่างกันในด้านบริบทและวัฒนธรรม (von Davier, 2010 as cited in von Davier et al., 2018, p.1) สอดคล้องกับผลการศึกษาคุณภาพของแบบวัดที่ใช้แองเคอร์วินยัตต์ในการปรับคะแนนของ Choothong (2014), Janprasert (2019), Kaewkua (2020) ที่พบว่า การปรับแก้คะแนนด้วยวิธีแองเคอร์วินยัตต์สามารถเพิ่มความเที่ยงของการวัดในทุก ๆ ด้าน เช่น ค่าสัมประสิทธิ์ความเที่ยง ค่าสัมประสิทธิ์สหสัมพันธ์ และความตรงเชิงโครงสร้าง

นอกจากนี้การปรับคะแนนด้วยแองเคอร์วินยัตต์ยังช่วยปรับปรุงความสัมพันธ์ของตัวแปรภายในองค์ประกอบโดยพบว่า การปรับคะแนนโดยใช้แองเคอร์วินยัตต์ทำให้ตัวแปรภายในองค์ประกอบทั้ง 5 องค์ประกอบมีขนาดความสัมพันธ์เพิ่มขึ้น ตัวอย่างเช่น การปรับปรุงความสัมพันธ์ของตัวแปรสังเกตได้ที่ใช้วัดคุณลักษณะแฝงด้านการยอมรับความท้าทายทั้ง 5 ตัวแปร ก่อนปรับคะแนนด้วยแองเคอร์วินยัตต์มีความสัมพันธ์ระหว่างตัวแปรในระดับต่ำ แต่หลังจากปรับคะแนนด้วยแองเคอร์วินยัตต์ความสัมพันธ์ระหว่างตัวแปรมีขนาดเพิ่มขึ้นเป็นระดับปานกลางถึงสูง แสดงให้เห็นถึงการเพิ่มคุณสมบัติเฉพาะในการวัดด้านการพร้อมที่จะเรียนรู้ในสิ่งที่ยากขึ้นและพร้อมเรียนรู้สิ่งใหม่ ๆ ที่เข้ามาในทุกโอกาสได้ชัดเจนขึ้น ในทางตรงกันข้ามกับตัวแปรสังเกตได้ที่อยู่ต่างองค์ประกอบจะมีความสัมพันธ์กันน้อยกว่าตัวแปรภายในองค์ประกอบ เนื่องจากตัวแปรแต่ละองค์ประกอบถูกออกแบบให้มุ่งวัดคุณลักษณะแฝงที่ต่างกัน เช่น ตัวแปรภายในองค์ประกอบการยอมรับความท้าทายมุ่งวัดเกี่ยวกับการพร้อมที่จะเรียนรู้ในสิ่งที่ยากขึ้นและพร้อมเรียนรู้สิ่งใหม่ ๆ ที่เข้ามาในทุกโอกาส ในขณะที่องค์ประกอบการไม่ย่อท้อต่อความล้มเหลวมุ่งวัดเกี่ยวกับการไม่ยอมแพ้ต่ออุปสรรคและความล้มเหลวโดยเปลี่ยนความท้อและความล้มเหลวเป็นแรงผลักดันในการพยายามเรียนรู้ปรับปรุงแก้ไขปัญหา ข้อผิดพลาดที่เกิดขึ้นเพื่อให้ตนเองเกิดการเรียนรู้และทักษะให้มากที่สุด ผลการวิจัยชี้ให้เห็นว่าการปรับคะแนนด้วยแองเคอร์วินยัตต์ช่วยปรับปรุงความสัมพันธ์ของตัวแปรภายในองค์ประกอบให้เพิ่มขึ้นสะท้อนถึงความสามารถในการลอคติและความคลาดเคลื่อนทำให้ความสัมพันธ์ระหว่างตัวแปรมีความชัดเจน (Untong et al., 2018)

การวิเคราะห์องค์ประกอบเชิงสำรวจมีประโยชน์ในกรณีที่ไม่มีทฤษฎีรองรับที่ชัดเจนซึ่งมีความสำคัญเป็นอย่างมากในการหาความสัมพันธ์ของตัวแปรเพื่อจัดองค์ประกอบ (Steven, 2009) แต่ทั้งนี้การวิเคราะห์องค์ประกอบเชิงสำรวจช่วยสะท้อนให้เห็นถึงโครงสร้างในการจัดองค์ประกอบของตัวแปรสังเกตได้ โดยความสัมพันธ์ของตัวแปรสังเกตได้ก่อนปรับคะแนนด้วยแองเคอร์ริงวินเยตต์สามารถจัดองค์ประกอบได้เพียง 3 องค์ประกอบ บ่งชี้ว่าตัวแปรสังเกตได้ไม่ได้จัดองค์ประกอบตามที่คาดหวัง อาจเกิดจากอคติของผู้ตอบที่ส่งผลต่อค่าความสัมพันธ์ของตัวแปร (Untong et al., 2018) เมื่อปรับคะแนนด้วยแองเคอร์ริงวินเยตต์พบว่าความสัมพันธ์ของตัวแปรสังเกตได้สามารถจัดองค์ประกอบได้ดีขึ้น และสามารถจัดองค์ประกอบได้ตรงตามโครงสร้างที่คาดหวัง 5 องค์ประกอบ แสดงให้เห็นถึงว่า การปรับคะแนนด้วยแองเคอร์ริงวินเยตต์สามารถลดอคติของผู้ตอบและช่วยเพิ่มความแม่นยำในการวัดได้ตรงตามโครงสร้างของตัวแปรแฝงที่ใช้วัดคุณลักษณะโกรวีธมายด์เซต

ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับหนึ่งพบว่า โมเดลก่อนปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ไม่สอดคล้องกับข้อมูลเชิงประจักษ์ ในขณะที่โมเดลหลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์สอดคล้องกับข้อมูลเชิงประจักษ์ ตัวอย่างเช่น โมเดลมี ค่า CFI ที่สูงกว่า .9 โดยไม่ปรับแก้โมเดล แสดงถึงความสอดคล้องที่ดีมากในโมเดลการวัด ขณะที่ค่า TLI ที่เพิ่มขึ้นแสดงให้เห็นว่าการปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ทำให้โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากขึ้นโดยไม่ต้องเพิ่มความซับซ้อนใด ๆ ในโมเดล (Joseph et al., 2018) นอกจากนี้การใช้แองเคอร์ริงวินเยตต์เพื่อปรับแก้คะแนนไม่เพียงแต่ปรับปรุงความสอดคล้องของโมเดลเท่านั้น ยังช่วยให้ตัวแปรสังเกตได้โมเดลสามารถวัดตัวแปรแฝงได้ตรงตามทฤษฎีที่คาดหวัง ทั้งนี้อาจเป็นผลมาจากความสัมพันธ์ระหว่างข้อคำถามที่เพิ่มขึ้น โดยลดอคติของค่าสัมประสิทธิ์สหสัมพันธ์ส่งผลให้โมเดลสามารถวัดได้ตรงตามทฤษฎีมากขึ้น Untong et al., (2018) สอดคล้องกับผลการศึกษาก่อนหน้านี้ของ Choothong (2014) และ Janprasert (2019)

ผลการตรวจสอบคุณภาพของแบบวัดก่อนและหลังปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ในมุมมองของทฤษฎีการตอบสนองข้อสอบสามารถทำนายการตอบของผู้ตอบได้อย่างแม่นยำโดยความสัมพันธ์ของข้อมูลก่อนปรับคะแนนมีลักษณะของข้อมูลที่ไม่กระจายตัวซึ่งอาจเกิดจากความเคยชินและอคติในการตอบ (Organisation for Economic Co-operation and Development, 2014) ทำให้มีผลต่อการคำนวณความน่าจะเป็นของคำถามข้อที่ 10, 11, 12, 14, 15, 17 และ 19 ที่ไม่สามารถคำนวณความน่าจะเป็นของรายการคำตอบ “1” กับรายการคำตอบ “2,3,4” ได้ ถึงแม้ความสามารถในการแยกแยะระดับความสามารถของผู้ตอบของแบบวัดก่อนปรับคะแนนจะสูงกว่าหลังปรับคะแนน แต่ประสิทธิภาพในการวัดความสามารถของผู้ตอบจากสารสนเทศที่มีลักษณะลดต่ำลงในช่วงของความสามารถชี้ให้เห็นถึงข้อบกพร่องในการวัดบางระดับความสามารถ สอดคล้องกับผลการวิจัยของ Rungrotrangsarn (2017) ที่พบข้อบกพร่องในการวัดบางระดับความสามารถ โดยสารสนเทศของแบบวัดลดต่ำลงเล็กน้อย 2 ช่วง ซึ่งการใช้แองเคอร์ริงวินเยตต์สามารถปรับแก้ข้อบกพร่องดังกล่าวได้ ทั้งนี้อาจเป็นเพราะการปรับคะแนนมีผลต่อความแปรปรวน และลักษณะการกระจายตัวของข้อมูล เมื่อกำหนดความน่าจะเป็นของรายการตอบจึงทำให้สามารถคำนวณค่าพารามิเตอร์ของแต่ละรายการคำตอบได้ครบทุกตำแหน่ง สะท้อนให้เห็นถึงการปรับปรุงประสิทธิภาพการทำนายความสามารถของ

ผู้ตอบได้แม่นยำและสามารถตีความคุณลักษณะโกรว์มายด์เซตได้ชัดเจนมากยิ่งขึ้นโดยพิจารณาจากคุณสมบัติของข้อคำถามและความสามารถในการวัดได้ครอบคลุมความสามารถของผู้ตอบ (Kanjawasee, 2020) สอดคล้องกับการศึกษาประสิทธิภาพในการปรับคะแนนโดยใช้แองเคอร์ริงวินเยตต์ที่ส่งผลให้แบบวัดมีประสิทธิภาพในการวัดได้แม่นยำขึ้นของ Hopkins & King (2010, as cited in Kyllonen, 2012)

ทั้งนี้จากผลการวิจัยพบว่า การประยุกต์ใช้แองเคอร์ริงวินเยตต์นอกจากจะช่วยเพิ่มประสิทธิภาพของการวัดแล้ว ยังเสริมความเข้าใจที่ดีขึ้นเกี่ยวกับการสร้างเครื่องมือวัดทางจิตวิทยาที่สามารถแก้ปัญหาความเป็นอัตวิสัยของการตอบ (Kyllonen, 2012) โดยเฉพาะอย่างยิ่งในการวัดโกรว์มายด์เซตที่ปรับคะแนนด้วยแองเคอร์ริงวินเยตต์จะช่วยให้ครูผู้สอนทราบพฤติกรรมแฝงที่แท้จริงของนักเรียนด้านการยอมรับความท้าทาย การไม่ย่อท้อต่อความล้มเหลว การตระหนักในความพยายามเรียนรู้ที่นำไปสู่ความสำเร็จ การเรียนรู้จากคำวิจารณ์ และการค้นหาบทเรียนและแรงบันดาลใจจากความสำเร็จของผู้อื่น ซึ่งผลการวิจัยทำให้มั่นใจได้ว่า ข้อมูลที่ได้จากเครื่องมือวัดนี้มีความเที่ยงตรงและเป็นธรรมเพื่อใช้ออกแบบและปรับปรุงกระบวนการเรียนการสอนให้สอดคล้องกับความต้องการของผู้เรียนรายบุคคล นอกจากนี้ยังสามารถใช้ทำนายและพัฒนาผลสัมฤทธิ์ทางการเรียน (Limsantitham et al., 2019; Stankov et al., 2009; Stankov et al., 2012; Sutthisomboon, 2020) และลดความเสี่ยงในการเลิกเรียนกลางคันได้ (Rissanen et al., 2019)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 แบบวัดโกรว์มายด์เซตที่พัฒนาขึ้นมีความเที่ยงและความตรงสูง ซึ่งบ่งชี้ถึงคุณภาพของแบบวัดที่สามารถนำไปใช้ได้จริง โรงเรียนสามารถใช้แบบวัดนี้ในการประเมินและติดตามการพัฒนาของนักเรียนในด้านโกรว์มายด์เซตอย่างต่อเนื่อง ครูผู้สอนสามารถใช้ผลการวัดโกรว์มายด์เซตเพื่อวางแผนการจัดการเรียนการสอนและจัดกิจกรรมที่เหมาะสมกับความต้องการของแต่ละคนเพื่อช่วยเพิ่มความสามารถในการเรียนรู้และการปรับตัวของนักเรียน ลดอัตราการเลิกเรียนกลางคัน และส่งเสริมการพัฒนาทักษะที่สำคัญต่อความสำเร็จในอนาคต

1.2 ผลการวิจัยชี้ให้เห็นว่า แบบวัดที่ใช้แองเคอร์ริงวินเยตต์ควบคู่กับแบบประเมินตนเองช่วยเพิ่มคุณภาพการวัดให้ดีขึ้นทั้งด้านความตรงและความเที่ยง ดังนั้นนักวิจัยและครูสามารถนำแองเคอร์ริงวินเยตต์มาใช้ในการพัฒนาและประเมินเจตคติของนักเรียนในด้านต่าง ๆ เพื่อแก้ปัญหาความเป็นอัตวิสัยของการตอบและทำให้ได้ข้อมูลที่มีความแม่นยำและยุติธรรมในการใช้เปรียบเทียบระหว่างบุคคล

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

การพัฒนาแบบวัดโกรว์มายด์เซตโดยประยุกต์ใช้แองเคอร์ริงวินเยตต์ครั้งต่อไป ผู้วิจัยควรใช้แบบวัดที่พัฒนาขึ้นในการศึกษาความสัมพันธ์ระหว่างผลการประเมินโกรว์มายด์เซตกับผลสัมฤทธิ์ทางการเรียน เนื่องจากโกรว์มายด์เซตเป็นความเชื่อว่าสติปัญญาสามารถพัฒนาได้ ดังนั้นหากจะตรวจสอบความน่าเชื่อถือของเครื่องมือวัดโกรว์มายด์เซตแล้ว ควรมีการศึกษาความสัมพันธ์ของคะแนนที่ได้จากการวัดระดับโกรว์มายด์เซตกับผลสัมฤทธิ์ทางการเรียนของนักเรียน Limsantitham et al. (2019) กล่าวว่าความเชื่อว่า

สติปัญญาสามารถพัฒนาได้มีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียน หากพบความสัมพันธ์ของตัวแปรเหล่านี้จะเป็นหลักฐานของความตรงประเภทความสัมพันธ์กับตัวแปรอื่น (relationship to other variables) ตามกรอบการตรวจสอบความตรงของ Messick (1995) ซึ่งจะช่วยเพิ่มความน่าเชื่อถือของการวัดไกรว์ธมาดค์เซตได้มากขึ้น และควรสร้างเกณฑ์ที่เป็นมาตรฐานในการจำแนกนักเรียนเพื่อใช้เป็นเกณฑ์มาตรฐานการการวัดเพื่อใช้พัฒนาไกรว์ธมาดค์เซตของนักเรียนให้เหมาะสมระดับความสามารถ

References

- Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hopkins, D. J., & King, G. (2010). Improving anchoring vignettes: Designing surveys to correct interpersonal incomparability. *Public Opinion Quarterly*, 74(2), 201–222. <https://doi.org/10.1093/poq/nfq011>
- King, G., Murray, C. J. L., Salomon, J. A., & Tandon, A. (2009). *Anchoring vignettes: Frequently asked questions*. Retrieved January 15, 2011, from <https://gking.harvard.edu/files/gking/files/vfaq.pdf>
- King, G., Murray, C. J., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(1), 191–207. <https://doi.org/10.1017/S000305540400108X>
- Kyllonen, P. C. (2012, May). *Measurement of 21st century skills within the Common Core State Standards*. Paper presented at the Invitational Research Symposium on Technology Enhanced Assessments, 7–8.
- Kyllonen, P. C., & Bertling, J. J. (2014). Innovative questionnaire assessment methods to increase cross-country comparability. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 277–286). CRC Press.
- Messick, S. (1995). Validity of psychological assessment: Validity of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>

- Midkiff, B., Langer, M., Demetriou, C., & Panter, A. T. (2018). An IRT analysis of the growth mindset scale. In V. Cherkassky, V. Vovk, & H. Wang (Eds.), *Quantitative psychology* (pp. 203–214). Springer. https://doi.org/10.1007/978-3-319-77249-3_14
- Möttus, R., Allik, J., Realo, A., Rossier, J., Zecca, G., Ah-Kion, J., Amoussou-Yéyé, D., Bäckström, M., Barkauskiene, R., Barry, O., Bhowon, U., Björklund, F., Bochaver, A., Bochaver, K., de Bruin, G., Cabrera, H. F., Chen, S. X., Church, A. T., Cissé, D. D., Dahourou, D., ..., Johnson, W. (2012). The effect of response style on self-reported conscientiousness across 20 countries. *Personality and Social Psychology Bulletin*, 38(11), 1423–1436. <https://doi.org/10.1177/0146167212451275>
- Organisation for Economic Co-operation and Development. (2014). *PISA 2012 technical report*. <https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf>
- Primi, R., Zanon, C., Santos, D., De Fruyt, F., & John, O. (2016). Anchoring vignettes. *European Journal of Psychological Assessment*, 32(1), 39-51. <https://doi.org/10.1027/1015-5759/a000336>
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4(3), 207-230. <https://doi.org/10.3102/10769986004003207>
- Rissanen, I., Kuusisto, E., Tuominen, M., & Tirri, K. (2019). In search of a growth mindset pedagogy: A case study of one teacher's classroom practices in a Finnish elementary school. *Teaching and Teacher Education*, 77, 204-213.
- Stankov, L., Lee, J., & Paek, I. (2009). Realism of confidence judgments. *European Journal of Psychological Assessment*, 25, 123-130.
- Stankov, L., Lee, J., Luo, W., & Hogan, D. J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept, and anxiety?. *Learning and Individual Differences*, 22(6), 747-758. <https://doi.org/10.1016/j.lindif.2012.05.013>
- Steven, J. P. (2009). *Applied multivariate statistics for the social sciences* (p. 345). Routledge.
- von Davier, M. (2010). Why sum scores may not tell us all about test takers. *Newborn and Infant Nursing Reviews*, 10(1), 27–36. <https://doi.org/10.1053/j.nainr.2009.12.011>
- von Davier, M., Shin, H. J., Khorramdel, L., & Stankov, L. (2018). The effects of vignette scoring on reliability and validity of self-reports. *Applied Psychological Measurement*, 42(4), 291-306. <https://doi.org/10.1177/0146621617730389>
- Wand, J., King, G., & Lau, O. (2011). Anchors: Software for anchoring vignette data. *Journal of Statistical Software*, 42, 1-25.
- Westby, C. (2020). Growth mindsets: Ideas from Carol Dweck. *Word of Mouth*, 31(5), 1-3.

Yilmaz, E. (2022). Development of Mindset Theory Scale (Growth and Fixed Mindset): A validity and reliability study. *Research on Education and Psychology (REP)*, 6, 1-24.

Translate Thai References

Absuksiri, A. (2022). *Psychology for teachers (7th ed.)*. Chulalongkorn University Press. (in Thai)

Choothong, M. (2014). *A development of students' achievement motivation assessment tool using anchoring vignette* [Master's thesis]. Chulalongkorn University. (in Thai)

Janprasert, B. (2019). A development of study burnout measurement tool using anchoring vignette. *Veridian E-Journal Silpakorn University*, 12(2), 825–845. (in Thai)

Kaewkua, V. (2020). *A development of the goodness scale to evaluate growing good program: Score adjustment by using anchoring vignettes method* [Doctoral dissertation]. Chulalongkorn University. (in Thai)

Kanjanawasee, S. (2020). *Classical Test Theory (4th ed.)*. Chulalongkorn University Press. (in Thai)

Limsantitham, N., Khumwong, P., Phonpoke, N., & Dahsah, C. (2019). Growth mindset in science learning of lower secondary students in schools under the Bangkok Metropolitan Administration. *Srinakharinwirot Science Journal*, 35(2), 149-165. (in Thai)

Rungrotrangsarn, K. (2017). A development of the sportsmanship scale of athletes in the Institute of Physical Education [Master's thesis]. Songkla University. (in Thai)

Suriyachanhom, T., & Piyakul, A. (2018). The factors and the indicators of growth mindset for undergraduate teacher professional students. *Journal of Graduate School, Pitchayatat*, 14(1), 81-90. (in Thai)

Sutthisomboon, S. L. (2020). *Growth mindset intervention in teaching and learning: Meta-analysis and content analysis* [Doctoral dissertation]. Rangsit University. (in Thai)

Tippawon, C. (2019). *Development indicators of growth mindset for sixth grade students* [Master's thesis]. Mahasarakham University. (in Thai)

Untong, A., Guntawongwan, K., & Piromsombat, K. (2018). Correcting scale perception bias in measuring local residents' perception in tourism development impact using anchoring vignettes. *Journal of Research Methodology*, 30(2), 225–250. (in Thai)

Wongyai, W., & Patphol, M. (2019). Learning management for enhancing growth mindset. *Sikkha Journal of Education*, 6(1), 52–60. (in Thai)

Wongyai, W., & Patphol, M. (2020). *Learning based on growth mindset*. Graduate School, Srinakharinwirot University. (in Thai)

การพัฒนาชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬา เซปักตะกร้อระดับเยาวชน

Development of an Assessment Package for Physical Fitness and Serving Ability in Youth Sepak Takraw Athletes

กฤษฎา พุ่มพิน^{1*} ศุภวรรณ วงศ์สร้างทรัพย์² และ นิรอมลี มะกาเจ³

Kritchaya Poompin^{1*}, Suppawan Vongsangsap² and Niromlee Makaje³

(Received: May 7, 2024; Revised: June 10, 2024; Accepted: June 12, 2024)

บทคัดย่อ

งานวิจัยนี้มีความมุ่งหมายเพื่อ 1) ศึกษาองค์ประกอบในการประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน และ 2) ตรวจสอบคุณภาพของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน การวิจัยจำแนกเป็น 2 ระยะ ได้แก่ ระยะที่ 1 การศึกษาองค์ประกอบในการประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟ เก็บข้อมูลกับกลุ่มผู้ให้ข้อมูลสำคัญ ได้แก่ ผู้ฝึกสอน ผู้ช่วยผู้ฝึกสอน ผู้จัดการทีม และนักกีฬาเซปักตะกร้อหญิงตำแหน่งเสิร์ฟ รวมจำนวน 50 คน โดยใช้แบบสอบถาม และวิเคราะห์ข้อมูลโดยใช้สถิติบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ระยะที่ 2 การตรวจสอบคุณภาพของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟ เก็บข้อมูลกับนักกีฬาเซปักตะกร้อหญิงโรงเรียนกีฬากรุงเทพมหานคร จำนวน 30 คน และวิเคราะห์ข้อมูลโดยใช้สถิติบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน และตรวจสอบคุณภาพของชุดแบบประเมินโดยการวิเคราะห์ความตรงเชิงเนื้อหา ความเชื่อมั่นด้วยวิธีทดสอบซ้ำโดยใช้สูตรของสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันโปรดักโมเมนต์และการวิเคราะห์ความเป็นปรนัยด้วยการวิเคราะห์สหสัมพันธ์ภายในชั้น ผลการวิจัยพบว่า

1. องค์ประกอบของสมรรถภาพทางกายที่สัมพันธ์กับการเสิร์ฟที่มีความสำคัญมากที่สุด ใน 4 ลำดับแรก ได้แก่ 1) ความแข็งแรงของกล้ามเนื้อ ($M = 4.88$, $SD = 0.33$) 2) ความอ่อนตัว ($M = 4.86$, $SD = 0.35$) 3) การทรงตัว ($M = 4.72$, $SD = 0.61$) และ 4) พลังกล้ามเนื้อ ($M = 4.64$, $SD = 0.53$) และความสามารถ

¹ นิสิตดุษฎีบัณฑิต สาขาวิชาพลศึกษา คณะศึกษาศาสตร์และพัฒนศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน

² ผู้ช่วยศาสตราจารย์ คณะศึกษาศาสตร์และพัฒนศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน

³ ผู้ช่วยศาสตราจารย์ คณะวิทยาศาสตร์การกีฬา มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน

¹ Doctoral Degree's student, Physical Education Program, Faculty of Education and Development Sciences, Kasetsart University, Kamphaeng Saen Campus

² Assistant Professor, Faculty of Education and Development Sciences, Kasetsart University, Kamphaeng Saen Campus

³ Assistant Professor, Faculty of Sports Science, Kasetsart University, Kamphaeng Saen Campus

* Corresponding Author E-mail: kritchaya.p@ku.th

ในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชนที่สะท้อนให้เห็นถึงประสิทธิภาพในการเสิร์ฟอันได้แก่ ลูกเสิร์ฟที่เร็ว แรงและแม่นยำ สามารถประเมินจาก 3 องค์ประกอบ ได้แก่ 1) การจัดระเบียบร่างกาย พบว่า เป็นการประเมินจากความสมบูรณ์ของวงสวิงในการเสิร์ฟ 3 ช่วงที่ต่อเนื่องกัน คือ ช่วงเตรียมพร้อมและเหวี่ยงแขนขาไปด้านหลัง ช่วงเตะลูก (กระทบสัมผัสลูก) และช่วงส่งแรงตามลูก 2) ความแม่นยำในการเสิร์ฟ พบว่า เป็นการประเมินความแม่นยำใน 2 รูปแบบ ได้แก่ การเสิร์ฟได้ลงตรงตามตำแหน่งในสนามที่มีคะแนนแตกต่างกัน และการเสิร์ฟได้ลงตรงตามจุดที่กำหนด และ 3) ความเร็วของลูกตะกร้อในการเสิร์ฟ เป็นการคัดเลือกลูกเสิร์ฟที่มีความเร็วสูงสุดจากการเสิร์ฟทั้งหมด 5 ครั้ง

2. ผลการประเมินความตรงเชิงเนื้อหาของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน พบว่า ค่า CVI ของรายการประเมินมีค่า 1.00 ทุกรายการ ความเชื่อมั่นของรายการประเมินมีค่าความเชื่อมั่นระหว่าง .72 ถึง .99 และความเป็นปรนัยของรายการประเมินมีค่าสหสัมพันธ์ภายในชั้น ระหว่าง .96 ถึง 1.00

คำสำคัญ ความสามารถในการเสิร์ฟ ชุดแบบประเมิน นักกีฬาเซปักตะกร้อระดับเยาวชน สมรรถภาพทางกาย

ABSTRACT

This research aimed to 1) investigate the components of physical fitness and serving ability assessments for youth sepak takraw athletes and 2) examine the quality of the assessment tools for physical fitness and serving ability. The study was conducted in two phases. Phase 1 investigated the components of physical fitness and serving ability. Data were collected from 50 key informants—coaches/assistant coaches, team managers, and female sepak takraw athletes in the serving position—using questionnaires. The data were analyzed using descriptive statistics, including frequency, percentage, mean, and standard deviation. Phase 2 examined the quality of the assessment tools. Data were collected from 30 female sepak takraw athletes from Bangkok Sports School and analyzed using descriptive statistics, content validity analysis via the Content Validity Index (CVI), and test-retest reliability analysis using Pearson's correlation coefficient and the intraclass correlation coefficient (ICC). The research findings revealed the following:

1. The top four physical fitness components most significantly associated with serving were: 1) muscular strength ($M = 4.88$, $SD = 0.33$), 2) flexibility ($M = 4.86$, $SD = 0.35$), 3) balance ($M = 4.72$, $SD = 0.61$), and 4) muscular power ($M = 4.64$, $SD = 0.53$). These components reflect the serving performance of youth sepak takraw athletes in terms of speed, power, and accuracy. Serving ability can be assessed through three elements: 1) body coordination, which evaluates the completeness of the serving motion in three phases—preparation and backswing, ball contact, and force transmission; 2) serving accuracy, assessed either by the ability to hit specific

target zones on the court or by hitting predefined target points; and 3) serving speed, measured by selecting the fastest attempts from five total serves.

2. The content validity index (CVI) of the assessment package for physical fitness and serving ability was 1.00 for each item. The reliability coefficients ranged from .72 to .99, and the consistency, measured by ICC, ranged from .96 to 1.00.

Keywords: serving ability, assessment package, youth sepak takraw athletes, physical fitness

บทนำ

กีฬาเซปักตะกร้อเป็นกีฬาที่ต้องใช้ทักษะที่สำคัญต่าง ๆ ในการเล่น เช่น การโยน การเสิร์ฟ การเปิดลูก การชง การเตะ และการบล็อก เป็นต้น สำหรับ“การเสิร์ฟ” นับได้ว่าเป็นการรุกที่สำคัญอันดับแรก เพราะถ้าเสิร์ฟดีทีมอาจจะได้คะแนนทันทีจากการเสิร์ฟ หรือทีมตรงข้ามได้กลับเล่นลูกกลับมาง่าย ๆ ทำให้ได้เปรียบในการเล่นเกมรุก หรืออาจทำให้ได้เปรียบเสียเปรียบในการทำคะแนน ผู้ฝึกสอนและนักกีฬาก็จำเป็นต้องฝึกซ้อมวิธีการเสิร์ฟที่มีประสิทธิภาพสูงในการทำคะแนน ซึ่งต้องอาศัยจังหวะความสัมพันธ์ระหว่างผู้โยนกับผู้เสิร์ฟในการโยนส่งลูกตะกร้อและต้องอาศัยสมรรถภาพทางกายที่ดีของผู้เสิร์ฟ และภาวะจิตใจที่มั่นคง (Kaewsawang, 2005) โดย Pranee (2008) ได้กล่าวว่าการเสิร์ฟนั้นเป็นส่วนที่สามารถทำคะแนนได้ง่ายที่สุด หากมีความชำนาญตามขบวนการฝึกที่ถูกต้อง โดยประสิทธิภาพในการเสิร์ฟตะกร้อเกิดจากปัจจัยในหลายประการ ได้แก่ 1) ปัจจัยด้านชีวกลศาสตร์ 2) ปัจจัยด้านสมรรถภาพทางกายของผู้เล่น 3) ปัจจัยด้านรูปแบบการฝึก โปรแกรมการฝึก และ 4) ปัจจัยด้านการโยน (Hamdan et al., 2012; Mounpool, 2018; Office of Sport Development and Recreation, 2008; Pranee, 2008; Punjamanut & Komaratad, 2015)

จากปัจจัยที่เกี่ยวข้องกับประสิทธิภาพในการเสิร์ฟที่ได้กล่าวมาข้างต้นสะท้อนให้เห็นว่า ปัจจัยด้านสมรรถภาพทางกายของผู้เล่นเป็นสิ่งที่เกี่ยวข้องความสามารถในการเล่นเกมรุกด้วยการเสิร์ฟ โดย Watchayakarn (2017) ได้กล่าวว่า องค์ประกอบของสมรรถภาพทางกายในการเล่นกีฬาเซปักตะกร้อประกอบด้วย 10 องค์ประกอบ ได้แก่ ความแข็งแรงของกล้ามเนื้อ พลังกล้ามเนื้อ ความทนทานของกล้ามเนื้อ ความทนทานของหัวใจและปอด ความเร็ว ความคล่องตัว ความอ่อนตัว การทรงตัว ปฏิกริยาตอบสนอง การประสานสัมพันธ์ของประสาทและกล้ามเนื้อ ซึ่งในอดีตที่ผ่านมาพบว่า มีการวิจัยเพื่อวัดและประเมินสมรรถภาพทางกายของนักกีฬาเซปักตะกร้อในภาพรวมโดยใช้เกณฑ์มาตรฐานของการกีฬาแห่งประเทศไทยและกรมพลศึกษาในการวิเคราะห์สมรรถภาพของนักกีฬา แต่ด้วยการใช้พลังงานและกล้ามเนื้อของนักกีฬาเซปักตะกร้อแต่ละตำแหน่งมีความแตกต่างกัน การวัดและประเมินสมรรถภาพทางกายเฉพาะตำแหน่งของนักกีฬาเซปักตะกร้อก็ควรมีความแตกต่างกันด้วย ดังเช่นงานวิจัยของ Ngammoo & Gomaratut (2016) ที่ได้ศึกษาความต้องการจำเป็นด้านสมรรถภาพทางกายของนักกีฬาเซปักตะกร้อในภาพรวม จำแนกเป็น 2 องค์ประกอบ ได้แก่ 1) สมรรถภาพทางกายด้านกล้ามเนื้อ และ 2) สมรรถภาพทางกายด้านระบบพลังงาน โดยผลการวิจัยพบว่า เมื่อจำแนกตามตำแหน่งผู้เล่นกีฬาเซปักตะกร้อในตำแหน่งตัวเสิร์ฟ สมรรถภาพทางกายด้านกล้ามเนื้อที่มีความต้องการจำเป็นในระดับสูงมีจำนวน 8 ข้อ ได้แก่ ความอ่อนตัว พลังของกล้ามเนื้อ การทำงานประสานกันระหว่างระบบประสาทและ

กล้ามเนื้อ เวลาปฏิกิริยา ความไว การทรงตัว ความแข็งแรงของกล้ามเนื้อ ความคล่องแคล่วว่องไว และระดับกลางถึงสูง 2 ข้อ ได้แก่ ความทนทานของกล้ามเนื้อ และความเร็ว แต่จากงานวิจัยในอดีตพบว่ามีการศึกษาสมรรถภาพทางกายของนักกีฬาเซปักตะกร้อของนักกีฬาทุกตำแหน่งด้วยวิธีการประเมินเดียวกัน ซึ่งการประเมินสมรรถภาพทางกายในแต่ละตำแหน่งอาจมีรูปแบบ วิธีการและเครื่องมือที่แตกต่างกัน (Tongkampanit, 2009)

ขณะที่งานวิจัยส่วนใหญ่มุ่งวัดและประเมินทักษะและความสามารถในการเล่นเซปักตะกร้อไม่ว่าจะเป็นการวัดและประเมินการเสิร์ฟ การตั้งรับ การเตะ การโหม่ง (Irawan et al., 2021; Leesakul, 2004) หรือการเล่นลูกข้างเท้าด้านใน การเล่นลูกหลังเท้า การเล่นลูกเข่า และการเล่นลูกศีรษะ (Klangpimai & Yamngamluer, 2022) นอกจากนี้ยังมีการศึกษาเกี่ยวกับปัจจัยด้านชีวกลศาสตร์ เช่น คิเนมาติกส์ของการเสิร์ฟ ได้แก่ ความเร็วสูงสุดของลูกเตะกร้อ ความเร็วสูงสุดของข้อเท้า ความเร่งของข้อเท้าในทุกช่วงของการเสิร์ฟ หรือการเคลื่อนไหวของร่างกายขณะเล่นกีฬาเซปักตะกร้อ ซึ่งเป็นปัจจัยที่เกี่ยวข้องกับประสิทธิภาพในการเสิร์ฟ (Hamdan et al., 2012; Moungpool, 2018; Udomtaku & Konham, 2020) ดังนั้นเห็นว่าความสามารถในการเสิร์ฟนั้นย่อมเป็นผลมาจากการจัดระเบียบร่างกายที่ดี อันส่งผลให้ลูกเสิร์ฟมีความเร็ว แรงและแม่นยำ

จากที่กล่าวมาข้างต้นจึงสะท้อนให้เห็นว่ายังขาดการวัดและประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟของนักกีฬาเซปักตะกร้อหญิงระดับเยาวชนเฉพาะตำแหน่งเสิร์ฟ ทั้งนี้การมีทักษะและการเคลื่อนไหวที่มีประสิทธิภาพได้นั้นย่อมต้องเป็นผลมาจากการมีสมรรถภาพทางกายที่ดีด้วย ดังนั้นจะเป็นประโยชน์อย่างยิ่งหากมีการศึกษาองค์ประกอบการประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟ เพื่อพัฒนาชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อหญิงระดับเยาวชน และนำผลที่ได้มาใช้ในการพัฒนาความสามารถในการเสิร์ฟตามความสามารถของแต่ละบุคคล ตลอดจนเป็นแนวทางการพัฒนาชุดแบบประเมินสมรรถภาพทางกายสำหรับนักกีฬาเซปักตะกร้อหญิงระดับเยาวชนเพื่อพัฒนาความสามารถในทักษะต่าง ๆ ของนักกีฬาเซปักตะกร้อจำแนกตามตำแหน่งต่าง ๆ และแบบทดสอบสมรรถภาพทางกายเฉพาะกีฬาตะกร้อต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาองค์ประกอบในการประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน
2. เพื่อตรวจสอบคุณภาพของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

วิธีดำเนินการวิจัย

ข้อเสนอการวิจัยได้รับรองจริยธรรมการวิจัยจากคณะกรรมการจริยธรรมการวิจัยที่เกี่ยวข้องกับมนุษย์จากคณะกรรมการจริยธรรมและจรรยาบรรณการวิจัย สถาบันวิจัยและพัฒนา มหาวิทยาลัยรามคำแหง เมื่อวันที่ 24 มกราคม 2567 และสิ้นสุดการรับรองวันที่ 23 มกราคม 2568 หมายเลขใบรับรอง RU-HRE 67/0032 หมายเลขอ้างอิงโครงการวิจัย อว 0601.24/2072 สำหรับการวิจัยในครั้งนี้มีการดำเนินการวิจัยจำแนกออกเป็น 2 ระยะ ตามวัตถุประสงค์ของการวิจัย โดยมีรายละเอียดดังต่อไปนี้

ระยะที่ 1 การศึกษาองค์ประกอบในการประเมินสมรรถภาพทางกายและความสามารถในการเลิร์ฟสำหรับ นักกีฬาเซปักตะกร้อระดับเยาวชน

1. ผู้ให้ข้อมูลสำคัญ

ผู้ให้ข้อมูลสำคัญ ได้แก่ ผู้มีส่วนได้ส่วนเสีย (Stakeholder) ในการใช้ชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเลิร์ฟ จำนวน 50 คน ประกอบด้วย ผู้มีส่วนได้ส่วนเสียจำนวน 3 กลุ่มดังนี้

- ผู้ฝึกสอน ผู้ช่วยผู้ฝึกสอนกีฬาเซปักตะกร้อหญิงระดับเยาวชน จำนวน 30 คน
- นักกีฬาเซปักตะกร้อหญิงตำแหน่งเลิร์ฟ จำนวน 10 คน ที่มีประสบการณ์ในการแข่งขัน 10 ปี ขึ้นไป
- ผู้จัดการทีม หรือผู้ที่มีส่วนเกี่ยวข้องกับการทดสอบสมรรถภาพทางกายของนักกีฬาเซปักตะกร้อหญิงระดับเยาวชน จำนวน 10 คน

2. เครื่องมือในการวิจัย

แบบสอบถามเรื่อง “องค์ประกอบในการประเมินสมรรถภาพทางกายและความสามารถในการเลิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน” โดยประยุกต์ใช้แบบสอบถามการประเมินความต้องการจำเป็นของสมรรถภาพทางกายในกีฬาเซปักตะกร้อของ Ngammoo & Gomaratut (2016) โดยแบบสอบถามจำแนกเป็น 2 ตอนได้แก่

ตอนที่ 1 ข้อมูลพื้นฐานของผู้ตอบแบบสอบถาม เป็นแบบตรวจสอบรายการ (Checklist)

ตอนที่ 2 ความคิดเห็นที่มีต่อองค์ประกอบของการประเมินสมรรถภาพทางกายสำหรับนักกีฬาเซปักตะกร้อหญิงระดับเยาวชน เป็นแบบมาตราประมาณค่า 5 ระดับ จำนวน 18 รายการประเมิน จำแนกเป็น ด้านสมรรถภาพทางกาย จำนวน 10 รายการ และความสามารถในการเลิร์ฟ จำนวน 8 รายการ

การสร้างและตรวจสอบคุณภาพเครื่องมือ

การตรวจสอบคุณภาพของแบบสอบถามเรื่อง “องค์ประกอบในการประเมินสมรรถภาพทางกายและความสามารถในการเลิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน” ประกอบด้วยตรวจสอบเครื่องมือใน 2 ด้านได้แก่

1) การตรวจสอบความตรงเชิงเนื้อหาโดยวิเคราะห์จากดัชนีความสอดคล้องระหว่างวัตถุประสงค์กับข้อคำถาม (IOC) โดยผู้เชี่ยวชาญด้านกีฬาเซปักตะกร้อ จำนวน 5 คน ผลการวิเคราะห์ พบว่ารายการคำถามมีค่า IOC ตั้งแต่ 0.60 – 1.00 ถือว่ามีความตรงเชิงเนื้อหา สามารถนำไปใช้ได้

2) การตรวจสอบความเชื่อมั่นโดยวิเคราะห์ค่าความเชื่อมั่นแบบสอดคล้องภายในโดยวิเคราะห์จากค่าสัมประสิทธิ์ของครอนบาคแอลฟา ซึ่งทดลองใช้เครื่องมือกับกลุ่มทดลองที่มีลักษณะใกล้เคียงกับผู้ให้ข้อมูลสำคัญรวมจำนวน 30 คน ซึ่งพบว่ามีค่าความเชื่อมั่นของแบบประเมิน 0.94 ถือว่ามีค่าความเชื่อมั่นสูงนำไปใช้ได้

3. การเก็บรวบรวมข้อมูล

3.1 ประสานกับกลุ่มผู้ให้ข้อมูลสำคัญเพื่อขอความอนุเคราะห์ในการส่งแบบสอบถาม

3.2 ดำเนินการชี้แจงและส่งแบบสอบถามเรื่อง “องค์ประกอบในการประเมินสมรรถภาพทางกายและความสามารถในการเลิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน” ไปยังกลุ่มผู้ให้ข้อมูลสำคัญที่กำหนดไว้ โดยดำเนินการตามหลักการทำวิจัยในมนุษย์

4. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในระยะที่ 1 เป็นการวิเคราะห์ข้อมูลเชิงปริมาณ ได้แก่ ข้อมูลพื้นฐานของผู้ตอบแบบสอบถาม และความคิดเห็นที่มีต่อองค์ประกอบสมรรถภาพทางกายและความสามารถในการเสิร์ฟ โดยใช้สถิติบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ซึ่งวิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูป

ระยะที่ 2 การตรวจสอบคุณภาพของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

1. กลุ่มเป้าหมาย

1.1 การวิเคราะห์ความตรงเชิงเนื้อหา กลุ่มเป้าหมาย ได้แก่ ผู้ฝึกสอนกีฬาเซปักตะกร้อหญิงระดับเยาวชน จำนวน 5 คน

1.2 การวิเคราะห์ความเชื่อมั่น กลุ่มเป้าหมาย ได้แก่

1.2.1 ผู้ประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟ (assessor) ได้แก่ ผู้ฝึกสอนนักกีฬาเซปักตะกร้อหญิงโรงเรียนกีฬากรุงเทพมหานคร จำนวน 1 คน ที่มีประสบการณ์ในการเป็นผู้ฝึกสอน 10 ปีขึ้นไป

1.2.2 ผู้ถูกประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟ (assessee) ได้แก่ นักกีฬาเซปักตะกร้อหญิงโรงเรียนกีฬากรุงเทพมหานคร จำนวน 30 คน โดยเป็นนักเรียนที่กำลังศึกษาอยู่ในระดับมัธยมศึกษาตอนต้นและตอนปลายที่เป็นนักกีฬาเซปักตะกร้อหญิง

1.3 การวิเคราะห์ความเป็นปรนัย กลุ่มเป้าหมาย ได้แก่

1.3.1 ผู้ประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟ (Assessor) ได้แก่ ผู้ฝึกสอนนักกีฬาเซปักตะกร้อหญิงโรงเรียนกีฬากรุงเทพมหานคร จำนวน 3 คน ที่มีประสบการณ์ในการเป็นผู้ฝึกสอน 10 ปีขึ้นไป ซึ่ง 1 ใน 3 คน เป็นกลุ่มเดียวกันกับการวิเคราะห์ความเชื่อมั่น

1.3.2 ผู้ถูกประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟ (Assessee) ได้แก่ นักกีฬาเซปักตะกร้อหญิงโรงเรียนกีฬากรุงเทพมหานคร จำนวน 30 คน ซึ่งเป็นกลุ่มเดียวกันกับการวิเคราะห์ความเชื่อมั่น

2. เครื่องมือในการวิจัย

สำหรับเครื่องมือที่ใช้ในการวิจัยระยะที่ 2 เป็นการออกแบบจากผลการวิจัยในระยะที่ 1 ที่พบว่าสมรรถภาพทางกายที่ส่งผลต่อประสิทธิภาพในการเสิร์ฟ ประกอบด้วย 1) ความแข็งแรง 2) ความอ่อนตัว 3) การทรงตัว และ 4) พลังกล้ามเนื้อ ซึ่งในการออกแบบการทดสอบสมรรถภาพทางกายที่สอดคล้องกับการเสิร์ฟในภาคสนาม มีดังนี้ 1) การทดสอบความแข็งแรงของกล้ามเนื้อขาและพลังกล้ามเนื้อขา โดยการทดสอบการกระโดดขาเดียวจำนวน 3 ก้าว (Single Leg Triple Hop for Distance Test; SLTH) เนื่องจาก Hamilton et al. (2008) พบว่า ระยะทางที่สามารถกระโดดเป็นตัวแปรที่ใช้ในการทำนายความแข็งแรงและพลังของกล้ามเนื้ออย่างคร่าวๆของร่างกายได้ดีมาก 2) การทดสอบความอ่อนตัวโดยการทดสอบการแยกขา ซึ่งการวิจัยครั้งนี้ ได้ประยุกต์แนวทางในการวัดความอ่อนตัวของร่างกายส่วนล่างจากกีฬายิมนาสติกใน 2 รูปแบบ ได้แก่ 1) การแยกขาไปด้านหน้า (Front Split Test) แล้ววัดระยะห่างระหว่างแกนกลางลำตัวกับพื้น (separation distance) (Irrutia et al., 2010; Sands et al., 2008) และ 2) การแยกขาออกด้านข้าง (side split test) และ

วัดองศาที่สามารถแยกขาออกจากกันได้สูงสุด (angle of lateral split) (Iruetia et al., 2010) และ 3) การทดสอบการทรงตัว ในการวิจัยครั้งนี้เป็นการทดสอบการทรงตัวแบบอยู่นิ่ง (static balance) ที่ต้องอาศัยความแข็งแรงของขาซึ่งมีความสอดคล้องกับลักษณะของการเสิร์ฟที่ใช้ขาซ้ายเป็นขาหลักในการยืนซึ่งต้องขยับหรือบิดเท้าหาจังหวะในการเสิร์ฟ ทั้งนี้การทดสอบโดยการยืนขาเดียวบนโบลจะประเมินว่านักกีฬาสามารถรักษาสมดุลของขาในขณะที่เสิร์ฟได้สมดุลมากน้อยเพียงใด ซึ่งมีวิธีการในการทดสอบ ได้แก่ การยืนทรงตัวด้วยขาข้างเดียว (One-Leg Stance Test; OLS) หรือ Single Leg Stance Test ที่เป็นการประเมินการยืนทรงตัวด้วยขาข้างเดียวแล้วจับเวลา (Sanyawong, 2016) โดยจากงานวิจัยในอดีตพบว่ามีวิธีการวัดความสามารถในการทรงตัวที่หลากหลายรูปแบบ เช่น Bădău et al. (2019) ได้ประยุกต์ใช้ Bosu Ball ซึ่งเป็นลูกบอลที่นิยมใช้ในการฝึกความแข็งแรงของกล้ามเนื้อและฝึกการทรงตัวมาใช้เป็นเครื่องมือในการทดสอบ โดยวิธีการประเมินและอุปกรณ์ที่ใช้ประเมิน สรุปได้ดังตาราง 1

ตาราง 1 วิธีการประเมินและอุปกรณ์ที่ใช้ประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟฯ

การประเมิน	การทดสอบ	วิธีการประเมิน	อุปกรณ์	การตรวจสอบคุณภาพของชุดแบบประเมิน
1. สมรรถภาพทางกายที่สัมพันธ์กับการเสิร์ฟ				1. ความตรงเชิงเนื้อหา (Content Validity) 2. ความเชื่อมั่น (Reliability)
1.1 ความแข็งแรงและพลังของกล้ามเนื้อขา (Leg muscle strength and power)	การทดสอบการกระโดดขาเดียวจำนวน 3 ก้าว (Single Leg Triple Hop for Distance Test; SLTH)	1. ยืนด้วยขาขวาแล้วกระโดดต่อเนื่องจำนวน 3 ครั้ง (ชม.) 2. ยืนด้วยขาซ้ายแล้วกระโดดต่อเนื่องจำนวน 3 ครั้ง (ชม.)	1. ตลับเมตร 2. กรวย	
1.2 ความอ่อนตัว (Flexibility)	1. การแยกขาไปด้านหน้า (Front Split Test)	ระยะห่างระหว่างแกนกลางลำตัวกับพื้น (ชม.)	ไม้บรรทัด	
	2. การแยกขาออกด้านข้าง (Side Split Test)	ความสามารถในการแยกขาออกด้านข้าง (องศา)	โกนีโอมิเตอร์	
1.3 การทรงตัว (Balance)	การยืนทรงตัวด้วยขาข้างเดียว (Single Leg Stance Test)	1. การทรงตัวด้วยการยืนขาเดียวบนโบลด้วยขาขวา (วินาที) 2. การทรงตัวด้วยการยืนขาเดียวบนโบลด้วยขาซ้าย (วินาที)	1. Bosu ball 2. นาฬิกาจับเวลา	
2. ความสามารถในการเสิร์ฟ				3. ความเป็นปรนัย (Objectivity)
2.1 การจัดระเบียบร่างกาย	การทดสอบความสมบูรณ์ของวงสวิง (Serving Motion Test)	การประเมินวงสวิงจำนวน 5 ครั้ง	1. อุปกรณ์แขวนลูกตะกร้อ 2. ตลับเมตร	
2.2 ความแม่นยำในการเสิร์ฟ	การทดสอบความแม่นยำในการเสิร์ฟ (Serve Accuracy Test)	การประเมินความแม่นยำในการเสิร์ฟจำนวน 5 ครั้ง	3. กรวย	
2.3 ความเร็วของลูกตะกร้อในการเสิร์ฟ	การทดสอบความเร็วของลูกตะกร้อ (Speed Ball Test)	ความเร็วของลูกตะกร้อที่เกิดจากการเสิร์ฟด้วยรูปแบบที่ถนัด (กม./ชม.)	1. อุปกรณ์แขวนลูกตะกร้อ 2. ปืนเรดาร์วัดความเร็ว	

3. การเก็บรวบรวมข้อมูล

3.1 ดำเนินการส่งจดหมายขอความอนุเคราะห์ในการเก็บข้อมูลเพื่อตรวจสอบคุณภาพของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสริมสำหรับนักกีฬาเซปักตะกร้อหญิงระดับเยาวชน

3.2 ดำเนินการร่างชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสริมสำหรับนักกีฬาฯ

3.3 ศึกษาคุณภาพของชุดแบบประเมินสมรรถภาพทางกายสำหรับนักกีฬาเซปักตะกร้อหญิงระดับเยาวชน ตามข้อมูลที่ได้จากการศึกษาของประกอบในการประเมินสมรรถภาพทางกายและความสามารถในการเสริมฯ โดยวิเคราะห์ความตรงเชิงเนื้อหา (CVI) ซึ่งเก็บข้อมูลกับผู้ฝึกสอนกีฬาเซปักตะกร้อหญิงระดับเยาวชนจำนวน 5 คน

3.4 นำชุดแบบประเมินสมรรถภาพทางกายสำหรับนักกีฬาเซปักตะกร้อหญิงระดับเยาวชน ไปทดลองใช้เพื่อเก็บข้อมูล ดังนี้ 1) การศึกษาความเชื่อมั่น ผู้วิจัยได้เก็บข้อมูลจากผู้ประเมินที่เป็นผู้ฝึกสอนนักกีฬาเซปักตะกร้อหญิงโรงเรียนกีฬากรุงเทพมหานคร จำนวน 1 คน โดยประเมินนักกีฬาเซปักตะกร้อหญิงโรงเรียนกีฬากรุงเทพมหานคร จำนวน 30 คน ที่มีการประเมิน 2 ครั้ง โดยเว้นระยะในการประเมินซ้ำกับกลุ่มเดิม 1 สัปดาห์ และ 2) การศึกษาความเป็นปรนัย ผู้วิจัยได้เก็บข้อมูลจากผู้ประเมินที่เป็นผู้ฝึกสอนนักกีฬาเซปักตะกร้อหญิงโรงเรียนกีฬากรุงเทพมหานคร จำนวน 3 คน โดยประเมินนักกีฬาเซปักตะกร้อหญิงโรงเรียนกีฬากรุงเทพมหานคร จำนวน 30 คน ไปพร้อมกัน

4. การวิเคราะห์ข้อมูล

4.1 สถิติที่ใช้ในการวิเคราะห์ข้อมูลพื้นฐานของนักกีฬาระดับเยาวชน ได้แก่ ความถี่ และร้อยละ

4.2 สถิติที่ใช้วิเคราะห์คุณภาพของชุดแบบประเมินใช้สถิติบรรยายเพื่อวิเคราะห์ 1) ความตรงเชิงเนื้อหา จากการพิจารณาค่า CVI 2) ความเชื่อมั่นแบบวัดซ้ำโดยใช้สูตรของสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน โปรดักโมเมนต์ และ 3) ความเป็นปรนัยด้วยการวิเคราะห์สหสัมพันธ์ภายในชั้น (ICC) ซึ่งวิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูป

ผลการวิจัย

1. ผลการศึกษาองค์ประกอบในการประเมินสมรรถภาพทางกายและความสามารถในการเสริมสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

1.1 ผลการศึกษาองค์ประกอบในการประเมินสมรรถภาพทางกายที่สัมพันธ์การเสริมสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

การศึกษาคำตอบของการประเมินสมรรถภาพทางกายที่สัมพันธ์กับการเสริมสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน ในการวิจัยครั้งนี้ได้วิเคราะห์จากผลการตอบแบบสอบถามของกลุ่มผู้ให้ข้อมูลสำคัญจำนวน 50 คน ซึ่งเป็นข้อมูลเชิงปริมาณที่ประเมินลำดับความสำคัญของสมรรถภาพทางกายที่สัมพันธ์กับการเสริมสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชนจากทั้งสิ้น 10 องค์ประกอบ ซึ่งจากการศึกษาพบว่า สมรรถภาพทางกายที่มีความสำคัญมากที่สุด 4 ลำดับแรก ได้แก่ 1) ความแข็งแรงของกล้ามเนื้อ ($M = 4.88, SD = 0.33$) 2) ความอ่อนตัว ($M = 4.86, SD = 0.35$) 3) การทรงตัว ($M = 4.72, SD = 0.61$) และ 4) พลังกล้ามเนื้อ ($M = 4.64, SD = 0.53$)

1.2 ผลการศึกษาองค์ประกอบในการประเมินความสามารถในการเลิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

ความสามารถในการเลิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชนสำหรับการวิจัยในครั้งนี้พิจารณาจาก 1) การจัดระเบียบร่างกาย 2) ความแม่นยำในการเลิร์ฟ และ 3) ความเร็วของลูกตะกร้อในการเลิร์ฟ ดังนี้

“การจัดระเบียบร่างกาย” ผลการวิจัยพบว่า ผู้ตอบแบบสอบถามส่วนใหญ่มีความคิดเห็นว่า จำเป็นต้องประเมินการจัดระเบียบร่างกายในการเลิร์ฟ ได้แก่ ช่วงเตรียมพร้อม ช่วงเหวี่ยงขาไปด้านหลัง ช่วงเตะลูก และช่วงส่งแรงตามลูก และผู้ทรงคุณวุฒิได้ให้ข้อเสนอแนะว่า “ควรมีการพิจารณาช่วงของการเหวี่ยงแขน บิดไหล่ เพื่อเพิ่มเทคนิคในการจัดระเบียบร่างกาย” ดังนั้นสามารถสรุปได้ว่าการจัดระเบียบร่างกาย เป็นการพิจารณาจากความสมบูรณ์ของวงสวิงในการเลิร์ฟใน 3 ช่วงที่ต่อเนื่องกันคือ 1) ช่วงเตรียมพร้อมและเหวี่ยงแขนขาไปด้านหลัง 2) ช่วงเตะลูก (กระทบสัมผัสลูก) และ 3) ช่วงส่งแรงตามลูก

“ความแม่นยำในการเลิร์ฟ” ซึ่งจำแนกออกเป็น 2 รูปแบบ ได้แก่ 1) การเลิร์ฟได้ลงตรงตามตำแหน่งในสนามที่มีคะแนนแตกต่างกัน และ 2) การเลิร์ฟได้ลงตรงตามจุดที่กำหนด ผลการศึกษาพบว่าผู้ตอบแบบสอบถามทุกคน มีความคิดเห็นว่าควรประเมินทั้ง 2 แบบ จำนวน 50 คน คิดเป็นร้อยละ 100.00 ซึ่งจากผลการศึกษา รูปแบบการให้คะแนนความแม่นยำในการเลิร์ฟข้างต้น ผู้วิจัยจึงได้สร้างแบบทดสอบความแม่นยำในการเลิร์ฟตะกร้อที่ประกอบด้วย 2 สถานการณ์ ได้แก่ สถานการณ์ที่ 1 คือ การเลิร์ฟได้ลงตรงตามตำแหน่งในสนามที่มีคะแนนแตกต่างกัน จำนวน 25 ช่อง ซึ่งมีโอกาสเลิร์ฟ 4 ครั้ง แต่เลือกลูกที่มีคะแนนสูงสุดได้ 2 ลูก และสถานการณ์ที่ 2 คือ การเลิร์ฟได้ลงตรงตามจุดที่กำหนด 3 จุด โดยแต่ละจุดมีโอกาสเลิร์ฟได้ทั้งหมด 2 ครั้ง รวมทั้งหมด 6 ครั้ง แล้วเลือกลูกที่ดีที่สุด 3 ลูก ดังนั้นผู้ถูกประเมินจะถูกทดสอบรวม 10 ครั้งแต่คัดเลือกเหลือ 5 ครั้ง ซึ่งจะถูกนำมาคิดเป็นคะแนนโดยมีคะแนนเต็ม 25 คะแนน

“ความเร็วของลูกตะกร้อในการเลิร์ฟ” พบว่า ผู้ตอบแบบสอบถามทุกคน มีความคิดเห็นว่าควรประเมินจากลูกเลิร์ฟที่มีความเร็วสูงสุดจากการเลิร์ฟทั้งหมด 5 ครั้ง มาใช้ในการวิเคราะห์ จำนวน 50 คน คิดเป็นร้อยละ 100.00

2. ผลการตรวจสอบคุณภาพของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเลิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

ผลการประเมินความตรงเชิงเนื้อหา ของชุดแบบประเมินฯ พบว่า ค่า CVI ของรายการประเมินมีค่า 1.00 ทุกรายการ และมีข้อเสนอแนะในการปรับปรุง คือ 1) การทดสอบความสามารถในการเลิร์ฟ ควรทดสอบการเลิร์ฟจาก 13 ครั้ง ลดเหลือ 5 ครั้ง และ 2) ควรวัดความเร็วของลูกตะกร้อจากการเลิร์ฟด้วยรูปแบบที่ถนัด

ผลการวิเคราะห์ความเชื่อมั่นของรายการประเมินฯ พบว่า รายการประเมิน มีค่าความเชื่อมั่นระหว่าง .72 ถึง .99 ซึ่งมีค่าความเชื่อมั่นมากกว่า .70 สะท้อนให้เห็นว่า รายการประเมินในชุดแบบประเมินเหมาะสมที่นำไปใช้เก็บรวบรวมข้อมูล

ผลการวิเคราะห์ความเป็นปรนัยของรายการประเมินฯ จากผู้ประเมินทั้ง 3 ท่านโดยใช้การวิเคราะห์ Intraclass Correlation Coefficient (ICC) พบว่า รายการประเมิน มีค่า ICC ระหว่าง .96 ถึง 1.00 ซึ่ง มากกว่า .90 สะท้อนให้เห็นว่า รายการประเมินมีความเป็นปรนัยนำไปใช้ได้ รายละเอียดดังตาราง 2

ตาราง 2 ผลการวิเคราะห์ความเป็นปรนัยของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเลิร์ฟ

การทดสอบสมรรถภาพทางกายและความสามารถในการเลิร์ฟ	จำนวน (คน)	ผู้ประเมิน 1 (SD)	ผู้ประเมิน 2 (SD)	ผู้ประเมิน 3 (SD)	Inter-rater reliability		
					ICC	95% CI Low	95% CI High
1.ด้านความอ่อนตัว							
1.1 การทดสอบการแยกขาออกด้านข้าง (Side Split Test) (องศา)	30	171.17 (13.56)	171.80 (10.68)	173.33 (10.03)	.96	.93	.98
1.2 การทดสอบการแยกขาออกด้านหน้า (Front Split Test) (เซนติเมตร)	30	5.83 (8.72)	5.83 (8.72)	5.67 (8.78)	1.00	1.00	1.00
2.ด้านการทรงตัว							
2.1 การทรงตัวด้วยการยืนขาเดียวบนโบลด้วยขาขวา (วินาที)	30	52.30 (35.95)	52.30 (35.95)	52.30 (35.95)	1.00	1.00	1.00
2.2 การทรงตัวด้วยการยืนขาเดียวบนโบลด้วยขาซ้าย (วินาที)	30	45.33 (19.41)	45.33 (19.41)	45.33 (19.41)	1.00	1.00	1.00
3.ด้านความแข็งแรงและพลังกล้ามเนื้อขา							
3.1 ยืนด้วยขาขวาแล้วกระโดดต่อเนื่อง จำนวน 3 ครั้ง (เซนติเมตร)	30	473.20 (74.31)	473.20 (74.31)	473.20 (74.31)	1.00	1.00	1.00
3.2 ยืนด้วยขาซ้ายแล้วกระโดดต่อเนื่อง จำนวน 3 ครั้ง (เซนติเมตร)	30	465.20 (81.40)	465.20 (81.40)	465.20 (81.40)	1.00	1.00	1.00
4. ด้านการจัดระเบียบร่างกาย							
4.1 คะแนนวงสวิง (คะแนนเต็ม 5 คะแนน)	30	3.03 (1.40)	3.37 (1.19)	3.33 (1.12)	.96	.93	.98
5.ด้านความแม่นยำในการเลิร์ฟ							
5.1 คะแนนความแม่นยำในการเลิร์ฟ (คะแนนเต็ม 25 คะแนน)	30	7.87 (4.51)	7.57 (4.34)	7.80 (4.66)	.99	.98	1.00
6. ด้านความเร็วของลูกตะกร้อ							
6.1 ความเร็วลูกสูงที่สุดลำดับที่ 1 (กิโลเมตร/ชั่วโมง)	30	41.54 (8.73)	41.54 (8.73)	41.54 (8.73)	1.00	1.00	1.00

อภิปรายผล

1. อภิปรายผลการศึกษาแนวทางในการประเมินสมรรถภาพทางกายและความสามารถในการเลิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

1.1 อภิปรายผลการศึกษาองค์ประกอบในการประเมินสมรรถภาพทางกายที่สัมพันธ์การเลิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

จากผลการศึกษาในการวิจัยที่พบว่า “**ความแข็งแรงของกล้ามเนื้อ**” เป็นสมรรถภาพทางกายที่สัมพันธ์กับการเลิร์ฟอันดับ 1 เพราะความแข็งแรงของกล้ามเนื้อ ซึ่งมุ่งเน้นไปที่ความแข็งแรงของกล้ามเนื้อขา สำหรับการเลิร์ฟจำเป็นต้องใช้ความสามารถของกลุ่มกล้ามเนื้อต้นขาด้านหน้า กล้ามเนื้อด้านใน กล้ามเนื้อด้านนอก กล้ามเนื้อด้านหลังและกล้ามเนื้อน่อง รวมทั้งกล้ามเนื้อกลุ่มสนับสนุนของนักกีฬาเซปักตะกร้อหญิงระดับเยาวชนในการออกแรง

หรือแรงระเบิดสูงสุด ในกรณีที่นักกีฬาในตำแหน่งเสิร์ฟมีความถนัดในการใช้ขาขวาในการเสิร์ฟ ดังนั้นขณะที่ยืนทำการเสิร์ฟ นักกีฬาจะใช้ขาซ้ายเป็นขาที่ต้องยืนข้างเดียว ขณะที่ขาข้างขวาเป็นส่วนที่ต้องวางค้ำเป็นวงสวิงหาจังหวะกระทุบลูก ดังนั้นความแข็งแรงของกล้ามเนื้อขาในขณะที่ยืนเตรียมเสิร์ฟและขณะเสิร์ฟจึงเป็นสมรรถภาพทางกายที่มีความสัมพันธ์ต่อการเสิร์ฟ สอดคล้องกับงานวิจัยในอดีตที่พบว่า สมรรถภาพทางกายที่เกี่ยวข้องกับทักษะซึ่งเป็นองค์ประกอบสำคัญของทักษะการเสิร์ฟในกีฬาเซปักตะกร้อ คือ ความแข็งแรงของกล้ามเนื้อ (Anggraeni et al., 2023; Khampaengsri & Tingsabhat, 2018; Poca & Tingsabhat, 2014) โดยเฉพาะอย่างยิ่งความแข็งแรงของกล้ามเนื้อขาเป็นสมรรถภาพทางกายที่มีความสำคัญมากโดยเฉพาะตำแหน่งตัวเสิร์ฟ (Tongkampanit, 2009) ขณะที่ **“พลังกล้ามเนื้อ”** เป็นสมรรถภาพทางกายที่สัมพันธ์กับการเสิร์ฟอันดับ 4 เนื่องจากพลังกล้ามเนื้อ คือ ความสามารถของกลุ่มกล้ามเนื้อต้นขาด้านหน้า กล้ามเนื้อด้านใน กล้ามเนื้อด้านนอก กล้ามเนื้อด้านหลังและกล้ามเนื้อน่อง รวมทั้งกล้ามเนื้อกลุ่มสนับสนุนของนักกีฬาเซปักตะกร้อหญิงระดับเยาวชนในการหดตัวสูงสุดที่ทำงานแบบไม่ใช้ออกซิเจน ซึ่งการหดตัวมีความสำคัญสำหรับการเคลื่อนไหวแบบแรงระเบิดของความเร็วในช่วงเวลาสั้นๆ โดย Ngammoo & Gomaratut (2016) กล่าวว่า สมรรถภาพทางกายด้านพลังของกล้ามเนื้อมีระดับความต้องการจำเป็นอยู่ในระดับสูง เพราะรูปแบบการเล่นของกีฬาเซปักตะกร้อมีการเคลื่อนที่อย่างรวดเร็วในระยะเวลาสั้น ๆ สอดคล้องกับ Tohidin et al. (2021) ที่พบว่า พลังกล้ามเนื้อขาเป็นปัจจัยที่ส่งผลต่อความสามารถในการเสิร์ฟอย่างมีนัยสำคัญทางสถิติ โดย Aji (2018) พบว่า การฝึกฝนพลังกล้ามเนื้อขาจะช่วยให้เพิ่มความเร็วในการเสิร์ฟและการกำหนดเป้าหมายในการเสิร์ฟ ซึ่งในการวิจัยครั้งนี้จึงออกแบบทดสอบ **“ความแข็งแรงและพลังกล้ามเนื้อขา”** โดยใช้การทดสอบการกระโดดขาเดียวจำนวน 3 ก้าว (Single Leg Triple Hop for Distance Test; SLTH) โดย Hamilton et al. (2008) พบว่า ระยะทางที่สามารถกระโดดเป็นตัวแปรที่ใช้ในการทำนายความแข็งแรงและพลังของกล้ามเนื้ออย่างคร่าวๆ ส่วนล่างของร่างกายได้ดีมาก

“ด้านความอ่อนตัว” เป็นสมรรถภาพทางกายที่สัมพันธ์กับการเสิร์ฟอันดับ 2 เนื่องจากในขณะที่นักกีฬาทำการเสิร์ฟ เมื่อมีการยกขาขึ้นเป็นวงสวิงหาจังหวะกระทุบลูกได้แม่นยำมากน้อยเพียงใดนั้น เป็นผลจากความสามารถในการยืดเหยียดของกล้ามเนื้อ เอ็น และข้อต่อต่าง ๆ ของสะโพกและรยางค์ส่วนล่าง (lower limb) ในร่างกาย ซึ่งได้แก่ ต้นขา เข่า และปลายเท้า ซึ่งสมรรถภาพด้านความอ่อนตัวเป็นสมรรถภาพที่มีความสัมพันธ์กับความแข็งแรงของกล้ามเนื้อ (Krabuanrat, 2009) โดยวิจัยในอดีตของ Tohidin et al. (2021) พบว่า ความอ่อนตัวเป็นสมรรถภาพทางกายที่ส่งผลโดยตรงต่อความสามารถในการเสิร์ฟอย่างมีนัยสำคัญทางสถิติ เพราะความอ่อนตัวช่วยให้ร่างกายมีการจัดระเบียบที่ดีในระหว่างการเสิร์ฟและยกขาได้สูงสุดเมื่อมีการเสิร์ฟ ซึ่งจากผลงานวิจัยในอดีตก็สอดคล้องกับผลการศึกษาในการวิจัยครั้งนี้โดยสามารถสรุปได้ว่า สมรรถภาพทางกายที่เกี่ยวข้องกับทักษะซึ่งเป็นองค์ประกอบสำคัญของทักษะการเสิร์ฟในกีฬาเซปักตะกร้อ คือ ความอ่อนตัว (Anggraeni et al., 2023; Gani et al., 2024; Poca & Tingsabhat, 2014; Tongkampanit, 2009)

“ด้านการทรงตัว” เป็นสมรรถภาพทางกายที่สัมพันธ์กับการเสิร์ฟอันดับ 3 ที่เป็นเช่นนี้เนื่องจากในขณะที่ทำการเสิร์ฟ เช่น ในกรณีที่นักกีฬาในตำแหน่งเสิร์ฟมีความถนัดในการใช้ขาขวาในการเสิร์ฟ ดังนั้นขณะที่ยืนทำการเสิร์ฟ นักกีฬาจะใช้ขาซ้ายเป็นขาที่ต้องยืนข้างเดียวซึ่งการทรงตัวของขาซ้ายที่ดีและมั่นคงในขณะที่นักกีฬายกขาขึ้นเป็นวงสวิงหาจังหวะกระทุบลูกนั้น ย่อมส่งผลให้นักกีฬาสามารถเสิร์ฟไปยังทิศทางและเป้าหมายได้อย่างแม่นยำ สอดคล้องกับ Anggraeni et al. (2023) ที่ได้ส่งเคราะห์งานวิจัยในอดีตแล้วพบว่า สมรรถภาพทางกาย

ด้านการทรงตัวเป็นปัจจัยที่มีอิทธิพลอย่างมากต่อความสามารถในการเสิร์ฟ สอดคล้องกับ Ramli et al. (2023) ที่พบว่า การทรงตัวและการทำงานประสานกันระหว่างตาและเท้า ส่งผลร่วมกันต่อความสามารถในการเสิร์ฟ

1.2 อภิปรายผลการศึกษาคู่ประกอบในการประเมินความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

“การจัดระเบียบร่างกาย” จากผลการวิจัยที่พบว่า การจัดระเบียบร่างกาย เป็นการพิจารณาจากความสมบูรณ์ของวงสวิงในการเสิร์ฟ 3 ช่วงที่ต่อเนื่องกันคือ 1) ช่วงเตรียมพร้อมและช่วงเหวี่ยงแขนขาไปด้านหลัง 2) ช่วงเตะลูก (กระทบสัมผัสลูก) และ 3) ช่วงส่งแรงตามลูก สอดคล้องกับ Hamdan et al. (2012) ที่กล่าวว่า ระยะห่างของเท้าในขณะยืนเสิร์ฟ ความสูงของเท้าในขณะสัมผัสลูก และระยะห่างระหว่างบอลกับเท้า คือ การเคลื่อนไหวของร่างกายที่ส่งผลต่อความเร็วของตะกร้อในการเสิร์ฟ

“ความแม่นยำในการเสิร์ฟ” จากผลการวิจัยที่พบว่า ความแม่นยำในการเสิร์ฟ เป็นการประเมินจากผลจากการเสิร์ฟได้ลงตรงตามตำแหน่งในสนามที่มีคะแนนแตกต่างกัน และการเสิร์ฟได้ลงตรงตามจุดที่กำหนดจำนวน 5 ครั้งซึ่งเสิร์ฟด้วยรูปแบบที่ถนัดคือ ข้างเท้าด้านใน หลังเท้า และฝ่าเท้า สอดคล้องกับ Mounpool (2018) ที่ศึกษาตัวแปรทางคิเนมาติกส์ของการเสิร์ฟลูกตะกร้อด้วยหลังเท้าของนักกีฬามืออาชีพและสมัครเล่นจำนวน 20 ครั้งและกำหนดให้มีการเสิร์ฟลงในบริเวณที่กำหนดอย่างน้อย 5 ครั้ง และ Irawan et al. (2021) ที่พัฒนาแบบทดสอบความแม่นยำในการเสิร์ฟเซปักตะกร้อที่มีการให้คะแนนแตกต่างกันไปตามโอกาสในการได้แต้มในแต่ละบริเวณของสนามที่ยากง่ายต่างกัน

“ความเร็วของลูกตะกร้อในการเสิร์ฟ” เป็นการใช้สมรรถภาพทางกายแบบผสมผสาน โดยจากผลการวิจัยพบว่า ความเร็วของตะกร้อในการเสิร์ฟเป็นการคัดเลือกลูกเสิร์ฟที่มีความเร็วสูงสุดจากการเสิร์ฟทั้งหมด 5 ครั้งที่เป็นเช่นนี้เนื่องจากประสิทธิภาพในการเสิร์ฟนั้นคือ ลูกเสิร์ฟที่มีความเร็ว แรง และแม่นยำ ดังนั้นการพิจารณาจากความเร็วของลูกตะกร้อในการเสิร์ฟจึงจำเป็นต้องพิจารณาจากลูกเสิร์ฟที่มีความแม่นยำ คือ ได้แต้มจากการเสิร์ฟลงในบริเวณที่กำหนดแล้วจึงคัดเลือกลูกที่มีความเร็วสูงสุดสอดคล้องกับ Mounpool (2018) ที่กำหนดให้ทำการเสิร์ฟด้วยหลังเท้า 20 ครั้ง ต้องลงในบริเวณที่กำหนดอย่างน้อย 5 ครั้งแล้วคัดเลือกลูกเสิร์ฟที่ลงในบริเวณที่กำหนดและมีความเร็วสูงสุดมาใช้ในการวิเคราะห์โดย Aji (2018) พบว่า การฝึกฝนพลังกล้ามเนื้อขาจะช่วยให้เพิ่มความเร็วในการเสิร์ฟและการกำหนดเป้าหมายในการเสิร์ฟ

2. อภิปรายผลการตรวจสอบคุณภาพของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน

ความตรงเชิงเนื้อหาของชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬาเซปักตะกร้อระดับเยาวชน พบว่า ค่า CVI ของรายการประเมินมีค่า 1.00 ทุกรายการ สะท้อนให้เห็นว่า ชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟฯ มีความตรงเชิงเนื้อหาสามารถนำไปใช้ในการวัดได้ตรงตามสิ่งที่ต้องการวัดที่เป็นเช่นนี้เนื่องจากรายการประเมินในสมรรถภาพทางกายและความสามารถในการเสิร์ฟได้มาจากการศึกษาของคู่ประกอบในระยะที่ 1 ซึ่งมาจากผู้ทรงคุณวุฒิที่มีความเชี่ยวชาญและประสบการณ์ในเรื่องของการเสิร์ฟตะกร้อจึงทำให้ได้ข้อมูลที่ถูกต้องและสอดคล้องกับบริบทการทดสอบสมรรถภาพทางกายและความสามารถในการเสิร์ฟระดับเยาวชน โดยเครื่องมือวิจัยที่มีคุณภาพควรมีดัชนีความตรงเชิงเนื้อหาตั้งแต่ .75 ขึ้นไป

(Yaghmale, 2003) และไม่ควรถ้ากว่า .70 Dalawi et al. (2023) สอดคล้องกับ Tongkampanit & Kanchanapradit (2021) ที่สร้างชุดอุปกรณ์และโปรแกรมการฝึกที่พัฒนาความคล่องแคล่วว่องไวสำหรับนักกีฬาเซปักตะกร้อโดยพบว่า ทั้งอุปกรณ์การฝึกและโปรแกรมการฝึกสำหรับนักกีฬาเซปักตะกร้อมีค่า CVI 1.00 แสดงให้เห็นว่าสามารถนำไปใช้ฝึกความคล่องแคล่วว่องไวสำหรับนักกีฬาเซปักตะกร้อได้

ผลการวิเคราะห์ความเชื่อมั่นแบบวัดซ้ำโดยใช้สูตรของสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันโปรดักโมเมนต์ของรายการประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟ พบว่า รายการประเมินในชุดแบบประเมินมีค่าความเชื่อมั่นระหว่าง .72 ถึง .99 ซึ่งทุกรายการประเมินมีค่าความเชื่อมั่นมากกว่า .70 ซึ่ง Smits-Engelsman et al. (2021) ได้มีการศึกษาคุณภาพของชุดแบบประเมินสมรรถภาพทางกายด้านการเคลื่อนไหวในเด็กอายุระหว่าง 5-12 ปี โดยใช้การวิเคราะห์ความเชื่อมั่นแบบวัดซ้ำ ซึ่งมีค่าความเชื่อมั่นตั้งแต่ .80 ขึ้นไปสะท้อนให้เห็นว่าผลการประเมินที่ได้จากชุดแบบประเมินของงานวิจัยในครั้งนี้มีความคงที่ในการวัดซ้ำ ที่เป็นผลเช่นนี้เนื่องจากชุดแบบประเมินในงานวิจัยในครั้งนี้มีค่าความตรงเชิงเนื้อหาที่สูงซึ่งส่งผลให้การตีความของรายการประเมินมีความชัดเจนไม่ว่าจะประเมินกี่ครั้งผลการวัดจึงมีความคงที่ สอดคล้องกับ Kanjanawasee (2013) ที่กล่าวว่าความเที่ยงหรือความเชื่อมั่นมีความสัมพันธ์กับความตรง

ผลการวิเคราะห์ความเป็นปรนัยของรายการประเมินฯ จากผู้ประเมินทั้ง 3 ท่านโดยใช้การวิเคราะห์ Intraclass Correlation Coefficient (ICC) พบว่า รายการประเมิน มีค่า ICC ระหว่าง .96 ถึง 1.00 ซึ่งทุกรายการประเมิน (item) มีค่า ICC มากกว่า .90 สะท้อนให้เห็นว่า ผู้ประเมินทั้ง 3 คน ซึ่งเป็นผู้ประเมินในการวิจัยครั้งนี้ต่างก็มีประสบการณ์และคุณวุฒิใกล้เคียงกัน มีความเข้าใจรายการและวิธีการประเมินของชุดแบบประเมินไปในทิศทางเดียวกัน ที่เป็นเช่นนี้เนื่องจากชุดแบบประเมินในงานวิจัยในครั้งนี้มีค่าความตรงเชิงเนื้อหาที่สูงซึ่งส่งผลให้ผู้ประเมินแม้จะเป็นคนละคนกัน แต่การตีความและการประเมินก็ให้ผลที่สอดคล้องกัน สอดคล้องกับ Changwa & Tuksino (2023) ที่กล่าวว่า คุณลักษณะที่แตกต่างกันในกลุ่มของผู้ตรวจหรือผู้ประเมินส่งผลต่อคะแนน

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 จากผลการวิจัยที่พบว่า สมรรถภาพทางกายที่สัมพันธ์กับการเสิร์ฟที่มีความสำคัญ 4 ลำดับแรกได้แก่ความแข็งแรงของกล้ามเนื้อ ความอ่อนตัว การทรงตัว และพลังกล้ามเนื้อ ดังนั้นครูและผู้ฝึกสอนนักกีฬาคควรให้ความสำคัญกับสมรรถภาพทางกายดังกล่าวในการฝึกฝนและจัดโปรแกรมการฝึกซ้อมให้เหมาะสมกับนักกีฬาหญิงระดับเยาวชนในตำแหน่งเสิร์ฟอันจะส่งผลให้การเสิร์ฟมีประสิทธิภาพ มีความเร็ว แรงและแม่นยำมากยิ่งขึ้น

1.2 จากผลการวิจัยที่พบว่า ชุดแบบประเมินสมรรถภาพทางกายที่พัฒนาขึ้นมีความตรงเชิงเนื้อหา ความเชื่อมั่น และความเป็นปรนัย ดังนั้นครูและผู้ฝึกสอนนักกีฬาสามารถนำชุดแบบประเมินไปใช้เพื่อวัดและประเมินสมรรถภาพทางกายสำหรับการเตรียมความพร้อม และพัฒนาความสามารถในการเสิร์ฟของนักกีฬาเซปักตะกร้อหญิงระดับเยาวชนตำแหน่งเสิร์ฟ ในระดับรายการแข่งขันกีฬาเยาวชนแห่งชาติต่อไป

2. ข้อเสนอแนะสำหรับการทำวิจัยครั้งต่อไป

2.1 ควรมีการศึกษาแนวทางในการประเมินสมรรถภาพทางกายและความสามารถเฉพาะตำแหน่งของนักกีฬาเซปักตะกร้อหญิงระดับเยาวชน เนื่องจากในการวิจัยครั้งนี้เป็นการศึกษาเฉพาะนักกีฬาเซปักตะกร้อหญิง

ระดับเยาวชนในตำแหน่งหลังที่เป็นรูปแบบเกมรุกด้วยการเสิร์ฟ ซึ่งในการเล่นตะกร้อมีลักษณะการเล่นเป็นทีมที่ต้องอาศัยเกมรุกและรับที่มีความสัมพันธ์กันของผู้เล่นในตำแหน่งอื่น ๆ ที่มีสมรรถภาพทางกายแต่ละด้านแตกต่างกัน

2.2 ควรมีการศึกษาแนวทางในการพัฒนาสมรรถภาพทางกายเฉพาะกีฬาเซปักตะกร้อในระดับต่าง ๆ

กิตติกรรมประกาศ

ผลการวิจัยในบทความวิจัยฉบับนี้เป็นส่วนหนึ่งผลการศึกษาในดัชนีพนธ์เรื่อง “การพัฒนาชุดแบบประเมินสมรรถภาพทางกายและความสามารถในการเสิร์ฟสำหรับนักกีฬา เซปักตะกร้อระดับเยาวชน” ในหลักสูตรศึกษาศาสตรดุษฎีบัณฑิต สาขาวิชาพลศึกษา คณะศึกษาศาสตร์และพัฒนศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน จังหวัดนครปฐม โดย ผศ.ดร.ศุภวรรณ วงศ์สร้างทรัพย์ เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก และผศ.ดร.นิรอมล มะกาเจ เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

References

- Aji, T. (2018). Effect of training leg muscle power on the accuracy and speed of service as takraw in Central Java man players. *Journal of Indonesian Physical Education and Sport*, 4(1), 46–56.
- Anggraeni, E., Tomolius, & Sukamti, E. R. (2023). The effect of leg muscle strength, leg flexibility and balance on the ability to service Sepaktakraw: Literature review. *International Journal of Multidisciplinary and Analysis*, 6(7), 2945–2949.
- Bădău, A., Bădău, D., & Enoiu, R. S. (2019). Evaluation of stable balance capacity by using Bosu ball surfaces on different pressure levels. *Materiale Plastice*, 56(1), 216–219.
<https://doi.org/10.37358/mp.19.1.5154>
- Dalawi, I., Isa, M. R., Chen, X. W., Azhar, Z. I., & Aimran, N. (2023). Development of the Malay language of understanding, attitude, practice and health literacy questionnaire on COVID-19 (MUAPHQ C-19): Content validity & face validity analysis. *BMC Public Health*, 23, 1131.
<https://doi.org/10.1186/s12889-023-16044-5>
- Gani, A., Resmana, D., & Anugrah, S. M. (2024). Physical attributes and performance differences: Server vs. spiker players in men’s doubles sepak takraw. *Journal Sport Area*, 9(1), 151–159.
- Hamdan, N., Suwarganda, E., & Wilson, B. (2012). Factors correlated with sepak takraw serve speed. In E. Bradshaw & A. Burnett (Eds.), *Proceedings of the 30th International Conference on Biomechanics in Sports* (pp. 413–416). International Society of Biomechanics in Sports.
- Hamilton, R. T., Shultz, S. J., Schmitz, R. J., & Perrin, D. H. (2008). Triple-hop distance as a valid predictor of lower limb strength and power. *Journal of Athletic Training*, 43(2), 144–151.
<https://doi.org/10.4085/1062-6050-43.2.144>

- Irawan, R., Padli, V., Purba, R., & Susanti, S. (2021). Developing of top serve accuracy test on Sepak takraw sport game. *Journal of Human Sport and Exercise*, 16, Article 48. <https://doi.org/10.14198/jhse.2021.16.Proc3.48>
- Irurtia, A., Busquets, A., Carrasco, M., Ferrer, B., & Marina, M. (2010). Flexibility testing in young competing gymnasts using a trigonometric method: One-year follow-up. *Apunts Med Esport*, 45(168), 235-242.
- Ramli, R., Aji, T., Suwardi, S., Yanti, N., & Hanafi, M. (2023). Eye-foot coordination and balance with serving ability: A correlation study in sepak takraw game *Jurnal SPORTIF: Jurnal Penelitian Pembelajaran*, 9(3), 483–495. https://doi.org/10.29407/js_unpgri.v9i3.21217
- Sands, W. A., McNeal, J. R., Stone, M. H., Kimmel, W. L., Haff, G. G., & Jemni, M. (2008). The effect of vibration on active and passive range of motion in elite female synchronized swimmers. *European Journal of Sport Science*, 8(4), 217–223.
- Smits-Engelsman, B. C. M., Smit, E., Doe-Asinyo, R. X., Lawerteh, S. E., Aertssen, W., Ferguson, G., & Jelsma, D. L. (2021). Inter-rater reliability and test-retest reliability of the Performance and Fitness (PERF-FIT) test battery for children: A test for motor skill related fitness. *BMC Pediatrics*, 21, 1-11.
- Tohidin, D., Sepdanius, E., & Putra, A. A. (2021). Study of Service Ability in Sepaktakraw (Effect Leg Muscle Power, Flexibility and Self Confidence of Service Abilities). *International Journal of Human Movement and Sports Sciences*, 9(5), 905-911.
- Udomtaku, K., & Konharn, K. (2020). Energy Expenditure and Movement Activity Analysis of Sepaktakraw Players in the Thailand League. *Journal of Exercise Science & Fitness*, 18, 136-146.
- Yaghmale, F. (2003). Content validity and its estimation. *Journal of Medical Education*, 3(1), 25–27.

Translate Thai References

- Changwa, P., & Tuksino, P. (2023). 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. *Journal of Educational Measurement, Mahasarakham University*, 29(1), 133 – 147. (in Thai)
- Kaewsawang, S. (2005). *Handbook of Sepak Takraw Training*. Sports Authority of Thailand. (in Thai)
- Kanjanawasee, S. (2013). *Classical test theory* (7th ed.). Chulalongkorn University Press. (in Thai)

- Khampaengsri, S., & Tingsabhat, S. (2018). Effect of Using Speak Takraw with Yoga Training on Takraw Serving Ability of Lower Secondary School Student. *An Online Journal of Education*, 13(4), 67-79. (in Thai)
- Klangpimai, N., & Yamngamluer, V. (2022). Scoring Rubrics of Takraw Skill for Mattayomsuksa 3 Students of Sripruetta School. *Narkbhutparitat Journal*, 14(3), 24-33. (in Thai)
- Krabuanrat, C. (2009). *Stretching Exercise*. Faculty of Education, Kasetsart University. (in Thai)
- Leesakul, P. (2004). *A Construction of Speak Takraw Skill Test for Male Students in Levels 3 in Lopburi Province* [Master's thesis]. Srinakharinwirot University. (in Thai)
- Moungpool, B. (2018). *Study of Kinematic Variables of Back-Foot Speak Takraw Serving in Professional and Amateur Players* [Master's thesis]. Chulalongkorn University. (in Thai)
- Ngammoo, J., & Gomaratut, C. (2016). Physical Fitness Needs of Sepaktakraw Players. *Journal of Sports Science and Health*, 17(3), 1-15. (in Thai)
- Office of Sport Development and Recreation. (2008). *Handbook on Using Sports Science to Develop the Potential of Thai Student Athletes*. Cooperative Printing House, Thai Agricultural Cooperatives Limited. (in Thai)
- Poca, S., & Tingsabhat, S. (2014). Effect of a Sepak Takraw Serving Skill Training Program with Balance, Flexibility and Muscular Strength Training on Speak Takraw Serving Skills of the Lower Secondary School Students. *An Online Journal of Education*, 9(2), 129-143. (in Thai)
- Pranee, S. (2008). *Sepak Takraw: A Handbook of Sports Skill Training Strategies*. Odeonstore Press. (in Thai)
- Punjamanut, J., & Komaratad, R. (2015). The Analysis of Service Pattern of Sepaktakraw in the 26 th King's Cup. *Academic Journal Institute of Physical Education*, 7(2), 131 – 143. (in Thai)
- Sanyawong, N. (2016). *Effects of Patterned and Randomized Balance Training on Balancing Performance and Kicking Time of Taekwondo Players Age 8-12 Years* [Master's thesis]. Chulalongkorn University. (in Thai)
- Tongkampanit, P. (2009). *Physical Fitness of Sepaktakraw Players in Lampang Sports School* [Master's Independent study]. Chiang Mai University. (in Thai)
- Tongkampanit, P., & Kanchanapradit, K. (2021). A Construction of Equipment Package and Sport Specific Training Program for Improving Agility in Sepak Takraw Players. *Journal of Sports Science and Health*, 22(3), 303 – 317. (in Thai)
- Watchayakarn, K. (2017). *Skill and Instructional Sepaktakraw Sport*. Chulalongkorn University Press. (in Thai)

การพัฒนาหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัด
ระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ
Development of a Training Curriculum to Enhance Teaching Standards
Aligned with Formative and Summative Indicators for Quality School
Curriculum Development

กิตติรัตน์ เกษตรสุนทร^{1*} จักรชัย ตระกูลโอสถ² รัชชวี เถาโต³ วชิรญาณ สาดสง⁴

วัยวุฒม์ อยู่โนศล⁵ และ กวีเชษฐ์ เปีย⁶

Kittirat Kasatsuntorn^{1*}, Jakchai Trakoonosot², Raktawee Taoto³, Wachiraya Sadsang⁴,

Waiyawut Yoonisil⁵ and Kaweechate Pia⁶

(Received: June 17, 2024; Revised: September 10, 2024; Accepted: September 16, 2024)

บทคัดย่อ

การวิจัยครั้งนี้ มีความมุ่งหมายเพื่อ 1) ศึกษาความต้องการยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ 2) พัฒนาหลักสูตรฝึกอบรม 3) ทดลองใช้หลักสูตรฯ และ 4) ประเมินและปรับปรุงหลักสูตรฯ โดยใช้รูปแบบการวิจัยและพัฒนา (R&D) กลุ่มตัวอย่างจำนวนทั้งสิ้น 80 คน ใช้วิธีการเลือกแบบเจาะจง เครื่องมือ ได้แก่ แบบสัมภาษณ์ ประเด็นสนทนากลุ่ม แบบทดสอบความรู้และแบบประเมินผลการใช้หลักสูตร การวิเคราะห์ข้อมูลเชิงคุณภาพวิเคราะห์เชิงเนื้อหา และข้อมูลเชิงปริมาณ วิเคราะห์ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และ t-test โดยสรุปผลการวิจัยได้ดังนี้

¹ อาจารย์ คณะศึกษาศาสตร์ มหาวิทยาลัยทักษิณ

² เจ้าหน้าที่บริหารงานทั่วไป มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา

³ อาจารย์ คณะศึกษาศาสตร์ มหาวิทยาลัยมหามกุฏราชวิทยาลัย

⁴ อาจารย์ กลุ่มสาระการเรียนรู้คณิตศาสตร์ โรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร (ฝ่ายประถม)

⁵ ผู้ช่วยศาสตราจารย์ บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ

⁶ ครู กลุ่มสาระการเรียนรู้ภาษาต่างประเทศ โรงเรียนบ้านคังวังวัว

¹ Lecturer, Faculty of Education, Thaksin University

² General administration officer, Bansomdejchaopraya Rajabhat University

³ Lecturer, Faculty of Education, Mahamakut Buddhist University

⁴ Lecturer, Mathematics, Srinakharinwirot University: Prasarnmit Demonstration School (Elementary)

⁵ Assistant Professor, The Graduate School, Srinakharinwirot University

⁶ Teacher, Foreign Language Department, Bankungwangwua School

* Corresponding Author E-mail: torn.facnlub@gmail.com

1. ผลการศึกษาความต้องการฯ ได้ 3 ประเด็น ดังนี้ 1) ควรเน้นการสร้าง/ออกแบบหลักสูตรให้สอดคล้องกับบริบทในสถานศึกษา 2) ควรเน้นเกี่ยวกับการวิเคราะห์ตัวชี้วัดต้นทาง ระหว่างทาง และปลายทางให้สอดคล้องกับบริบทในสถานศึกษา และ 3) การนำหลักสูตรสถานศึกษาไปใช้ในห้องเรียนประกอบด้วยตัวชี้วัดการเรียนรู้ แผนการจัดการเรียนรู้ วิธีการจัดการเรียนรู้ เครื่องมือที่ใช้การวัดและประเมินผลการเรียนรู้ และวิธีการวัดและประเมินผลการเรียนรู้

2. ผลพัฒนาหลักสูตรฝึกอบรมฯ พบว่า องค์ประกอบของหลักสูตรออกเป็น 5 หมวด ได้แก่ หมวดที่ 1 ข้อมูลทั่วไป หมวดที่ 2 ข้อมูลเฉพาะของหลักสูตรฝึกอบรม หมวดที่ 3 การดำเนินการ และโครงสร้างของหลักสูตรฝึกอบรม หมวดที่ 4 หลักเกณฑ์ในการประเมินผล และหมวดที่ 5 การประเมินและปรับปรุงกระบวนการดำเนินการของหลักสูตรฝึกอบรม ส่วนการวิพากษ์ความเหมาะสมของหลักสูตรฉบับร่าง พบว่ามีความสอดคล้องเหมาะสม ลักษณะของหลักสูตรเป็นการบรรยายให้ความรู้ และฝึกปฏิบัติโดยการวิพากษ์ตัวชี้วัดในระดับต่าง ๆ พร้อมกับส่งเสริมความรู้ความเข้าใจเกี่ยวกับ การยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพที่แท้จริง

3. การทดลองใช้หลักสูตรฝึกอบรมฯ พบว่า ผู้เข้าอบรมมีค่าเฉลี่ยของผลสัมฤทธิ์หลังการเข้ารับการฝึกอบรมสูงกว่าก่อนการเข้ารับการฝึกอบรม

4. การประเมินผลและปรับปรุงหลักสูตรฝึกอบรมฯ พบว่า ผลการประเมินการใช้หลักสูตรฝึกอบรมฯ ในภาพรวมอยู่ในระดับมากที่สุด

คำสำคัญ หลักสูตรฝึกอบรม ตัวชี้วัด หลักสูตรสถานศึกษา

Abstract

The aim of this research was to: (1) study the needs related to raising teaching standards in alignment with formative and summative indicators toward the development of quality school curricula; (2) develop a training curriculum; (3) implement a pilot trial of the curriculum; and (4) evaluate and improve the curriculum. This research employed a research and development (R&D) model. The sample consisted of 80 individuals selected through purposive sampling. Data collection tools included interview protocols, focus group guidelines, a knowledge test, and curriculum evaluation forms. Data analysis involved qualitative content analysis and quantitative statistics, including frequency, percentage, mean, standard deviation, and t-tests. The findings were as follows:

1. The needs assessment revealed three key aspects: (1) the importance of designing curricula aligned with the context of the educational institution; (2) the necessity of analyzing learning indicators—initial, intermediate, and final—to ensure contextual relevance; and (3)

the implementation of school curricula in classrooms, including learning indicators, lesson plans, teaching methods, assessment tools, and evaluation procedures.

2. The developed training curriculum consisted of five major components: (1) general information, (2) specific details about the training program, (3) curriculum operations and structure, (4) evaluation criteria, and (5) procedures for evaluating and improving the implementation process. Expert reviews found the draft curriculum to be appropriate and consistent. The curriculum combined lectures with practical exercises that focused on analyzing indicators at different levels while promoting understanding of how to elevate teaching standards in alignment with both formative and summative indicators for the development of quality school curricula.

3. The curriculum trial showed that participants' post-training scores were significantly higher than their pre-training scores.

4. The evaluation and revision phase indicated that the overall quality of the training curriculum was rated at the highest level.

Keywords: training curriculum, indicators, school curriculum

บทนำ

การปฏิรูปการศึกษาในศตวรรษที่ 21 มีหลักการในการจัดการศึกษาที่สำคัญในเรื่อง “การกำหนดนโยบายให้มีเอกภาพ และให้มีความหลากหลายในการปฏิบัติ” โดยใช้ยุทธศาสตร์สำคัญ คือ การกระจายอำนาจ และการมีส่วนร่วมของทุกภาคส่วนเพื่อให้การกระจายอำนาจ และการมีส่วนร่วมของทุกภาคส่วนจึงจำเป็นต้องมีการกำหนดมาตรฐานทางการศึกษาและการประกันคุณภาพการศึกษา ทั้งนี้เพื่อให้การศึกษามีคุณภาพ สถานศึกษามีคุณภาพใกล้เคียงกัน และบุคคลได้รับสิทธิและโอกาสเสมอกันในการเข้าทางการศึกษาที่รัฐจัดให้ (Thailand Development Partners Forum Project, 2014) ดังนั้นเพื่อให้บุคคลได้รับความเสมอภาคทางการศึกษาทางรัฐบาลไทยจึงได้พัฒนาหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานขึ้นมาเพื่อเป็นแกนกลางในสถานศึกษานำไปปฏิบัติและยึดเป็นต้นแบบในการจัดการเรียนการสอน โดยมีการพัฒนาอย่างต่อเนื่องเพื่อยกระดับหลักสูตรให้ไปในแนวทางเดียวกัน (Sakulpramote, 2014) ในการจัดการศึกษาตามหลักสูตรแกนกลางเป็นการจัดการศึกษาให้ผู้เรียนได้รับการพัฒนาอย่างเต็มศักยภาพ สถานศึกษาจึงมีบทบาทสำคัญในการจัดทำหลักสูตรสถานศึกษา เพื่อนำสู่การจัดการเรียนการสอนในชั้นเรียนอย่างมีประสิทธิภาพ สร้างความมั่นใจต่อพ่อแม่ผู้ปกครองและชุมชนว่าผู้เรียนมีคุณภาพตามหลักสูตรของชาติและความต้องการของท้องถิ่น เพื่อให้บรรลุเจตนารมณ์ดังกล่าว สถานศึกษาจึงต้องมีการออกแบบหลักสูตรให้ครอบคลุมส่วนที่เป็นหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานและสอดคล้องกับสภาพปัญหาและความต้องการของชุมชนและท้องถิ่น ซึ่งสถานศึกษามีภารกิจที่สำคัญ ดังนี้ จัดทำหลักสูตรสถานศึกษา ส่งเสริมสนับสนุนในการพัฒนาบุคลากร

จัดสรรงบประมาณ ทรัพยากร ติดตามการใช้หลักสูตร กำกับดูแล นิเทศภายในให้การจัดการเรียนการสอนมีประสิทธิภาพ วิจัยและพัฒนาการใช้หลักสูตร ตลอดจนมีการประกันคุณภาพภายในของสถานศึกษา (Ratchakhom, 2014)

จากการสัมภาษณ์คณะผู้บริหารโรงเรียนบ้านสะพานเคียนเกี่ยวกับสภาพและปัญหาในการนำหลักสูตรแกนกลาง หลักสูตรสถานศึกษา และการนำตัวชี้วัดระหว่างทาง และตัวชี้วัดปลายทางสู่การปฏิบัติพบว่า บุคลากรภายในโรงเรียนยังขาดความรู้ ความเข้าใจเกี่ยวกับหลักสูตรสถานศึกษา การนำตัวชี้วัดระหว่างทางและตัวชี้วัดปลายทางสู่การปฏิบัติ เนื่องจากตัวชี้วัดแต่ละเรื่อง มีความเป็นเอกลักษณ์เฉพาะตัว หากไม่มีการศึกษารายละเอียดให้รอบคอบ จะทำให้เกิดปัญหาในการนำไปใช้ ซึ่งตัวชี้วัดมีความสอดคล้องกับการจัดการเรียนการสอน รวมไปถึงการวัดและประเมินผลของตัวชี้วัด หากบุคลากรไม่มีความรู้ ความเข้าใจในเรื่องดังกล่าว จะส่งผลให้เกิดปัญหาอย่างมากในการนำตัวชี้วัดระหว่างทางและตัวชี้วัดปลายของหลักสูตรสถานศึกษาสู่การปฏิบัติ ซึ่งสอดคล้องกับสภาพและปัญหาของ Nutrawong (2010) จากผลการวิจัยนำร่องการใช้หลักสูตร ซึ่งได้ทำการเก็บรวบรวมข้อมูลด้วยวิธีการที่หลากหลาย สามารถสรุปได้ตามประเด็นสำคัญ ดังนี้ 1) กระบวนการจัดทำกรอบหลักสูตรระดับท้องถิ่น/หลักสูตรสถานศึกษามีกระบวนการจัดทำหลักสูตรสถานศึกษาไม่ครบขั้นตอน และขั้นตอนที่สถานศึกษาส่วนใหญ่ไม่ได้ดำเนินการคือการศึกษาข้อมูลพื้นฐานของโรงเรียน และการศึกษากรอบหลักสูตรระดับท้องถิ่น 2) การบรรจุสาระการเรียนรู้ท้องถิ่นในหลักสูตรสถานศึกษาส่วนใหญ่สะท้อนให้เห็นถึงความไม่มั่นใจในคุณภาพของกรอบหลักสูตรระดับท้องถิ่นที่จัดทำขึ้น เพราะยังขาดความเข้าใจที่ชัดเจนในเรื่องดังกล่าว ประกอบกับระยะเวลาในการจัดทำมีจำกัดในส่วนของการนำสาระการเรียนรู้ท้องถิ่นสู่การเรียนการสอน และ 3) ปัญหาอุปสรรคในการจัดทำหลักสูตรสถานศึกษา ส่วนใหญ่บุคลากรทั้งในระดับเขตพื้นที่การศึกษา และระดับสถานศึกษายังขาดความรู้ ความเข้าใจเกี่ยวกับหลักสูตรใหม่ที่เพียงพอที่จะทำให้การขับเคลื่อนการใช้หลักสูตรจากส่วนกลางสู่ระดับเขตพื้นที่การศึกษา และสถานศึกษาเป็นไปอย่างมีประสิทธิภาพ นอกจากนี้ นักวิจัยทีดีอาร์ไอ (Wongkitrungrueang, 2011) ได้ศึกษาการบริหารองค์ความรู้ในหลักสูตร พบว่า ยังมีหลายส่วนไม่สอดคล้องกัน โดยเฉพาะการกำหนดตัวชี้วัด และเกณฑ์การวัดและประเมินผลตามตัวชี้วัดในมาตรฐานการเรียนรู้ทั้ง 8 กลุ่มสาระการเรียนรู้ มีความละเอียดและซ้ำซ้อน ส่งผลให้เนื้อหาหลักสูตรที่เรียนในแต่ละโรงเรียนไม่แตกต่างกัน เพราะจะต้องปฏิบัติตามตัวชี้วัดที่กำหนดไว้อย่างเคร่งครัด

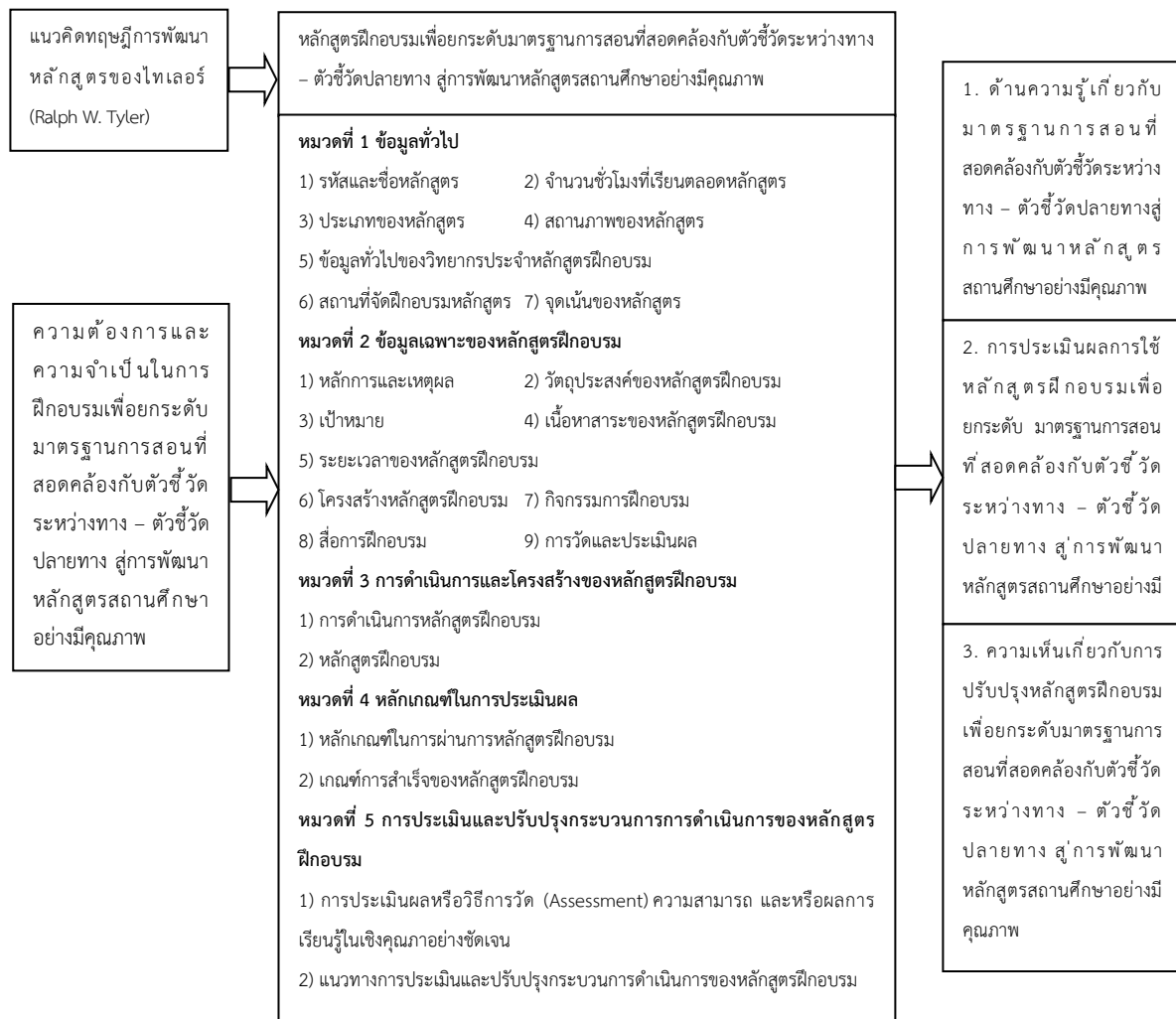
จากสภาพปัจจุบันเกี่ยวกับการยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ ตามที่คณะกรรมการการศึกษาขั้นพื้นฐาน (กพฐ.) ได้มีการประกาศปรับปรุงหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 ฉบับปรับปรุงแก้ไข พ.ศ. 2566 โดยปรับตัวชี้วัดทั้งหมด 2,056 ตัวชี้วัด เป็นตัวชี้วัดระหว่างทาง 1,285 ตัวชี้วัด และตัวชี้วัดปลายทาง 771 ตัวชี้วัดในกลุ่มสาระการเรียนรู้ต่าง ๆ ดังนั้น เพื่อให้เกิดประสิทธิภาพในการบริหารสถานศึกษาด้านการบริหารงานวิชาการและหลักสูตรสถานศึกษาและประสิทธิผลด้านการใช้หลักสูตร การจัดการเรียนการสอน การวัดและการประเมินผล รวมถึงผลลัพธ์ การเรียนรู้ของผู้เรียนตามตัวชี้วัดของหลักสูตรที่กำหนดไว้ ดังนั้น ผู้วิจัยจึงสนใจที่จะศึกษาการพัฒนาหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัด

ระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ จากปัญหาดังกล่าวที่ผู้วิจัย จึงได้ดำเนินการศึกษารูปแบบในการพัฒนาหลักสูตร จากการสังเคราะห์เอกสารเกี่ยวกับรูปแบบการพัฒนา หลักสูตร พบว่า รูปแบบการพัฒนาหลักสูตรที่เหมาะสมกับกับปัญหาดังกล่าว คือ รูปแบบการพัฒนาหลักสูตร ของไทเลอร์ (Ralph W. Tyler) หลักการสำหรับการพัฒนาหลักสูตรของไทเลอร์ มีองค์ประกอบ คือ 1) จุดมุ่งหมายทางการศึกษาที่โรงเรียนควรบรรลุคืออะไร 2) มีประสบการณ์ทางการศึกษาอะไรที่สามารถทำให้บรรลุจุดมุ่งหมาย 3) ประสบการณ์ทางการศึกษาเหล่านี้ สามารถจัดให้มีประสิทธิภาพอย่างไร และ 4) เราจะสามารถทราบได้อย่างไรว่าบรรลุจุดมุ่งหมายที่กำหนดไว้ โดยไทเลอร์ระบุว่านักพัฒนาหลักสูตรควร ตอบคำถามทั้ง 4 ข้อนี้ให้ชัดเจนตามลำดับ เพื่อให้การดำเนินการพัฒนาหลักสูตรมีประสิทธิภาพ ซึ่งการกำหนด จุดมุ่งหมายทางการศึกษาที่ชัดเจน จะทำให้ง่ายในการคัดเลือกประสบการณ์การเรียนรู้ที่ตอบสนองจุดมุ่งหมาย รวมทั้งส่งผลต่อการจัดประสบการณ์การเรียนรู้และการประเมินผลการเรียนรู้ (Patphol, 2024) ผู้วิจัยจึงนำ หลักการองค์ประกอบการพัฒนาหลักสูตรตามแนวคิดของไทเลอร์มาเป็นฐานในการพัฒนาหลักสูตรฝึกอบรม และแนวทางในการบุคลากรเพื่อให้เกิดความรู้ ความเข้าใจในเรื่องการนำตัวชี้วัดระหว่างทาง – การนำตัวชี้วัด ปลายทางไปสู่การปฏิบัติ

ความมุ่งหมายในการวิจัย

1. เพื่อศึกษาความต้องการยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัด ปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ
2. เพื่อการพัฒนาหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ
3. เพื่อการทดลองใช้หลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัด ระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ
4. เพื่อประเมินและปรับปรุงหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับ ตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ

กรอบแนวคิดการวิจัย



ภาพ 1 กรอบแนวคิดในการวิจัยพัฒนาหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ ผู้วิจัยใช้ระเบียบวิธีการวิจัยและพัฒนา (Research and Development : R & D) โดยมีขั้นตอนการดำเนินการวิจัย 4 ขั้นตอน ทั้ง 4 ขั้นตอนของการวิจัยและพัฒนา ผู้วิจัยได้นำหลักการองค์ประกอบทั้ง 4 องค์ประกอบของไทเลอร์มาเป็นฐานในการพัฒนาหลักสูตร โดยมีกระบวนการดังนี้

ขั้นตอนที่ 1 การวิจัย (R_1 : Research) ศึกษาความต้องการยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ

1. กลุ่มเป้าหมาย

กลุ่มเป้าหมายที่ใช้ในการวิจัย ได้แก่ ผู้อำนวยการโรงเรียนบ้านสะพานเคียน โรงเรียนลำลอง และโรงเรียนบ้านพรุหาว จำนวน 3 คน โดยใช้วิธีการเลือกแบบเจาะจง (purposive sampling) เนื่องจากทั้งสาม

โรงเรียนอยู่ในพื้นที่บริบทสภาพแวดล้อม แหล่งเรียนรู้ในพื้นที่เดียวกัน และทั้งสามโรงเรียนได้มีการทำบันทึกความเข้าใจร่วมกันระหว่างหน่วยงานภายใต้หลักเกณฑ์แนวปฏิบัติเดียวในเรื่องยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ

2. เครื่องมือที่ใช้ในการวิจัย

แบบสัมภาษณ์แบบมีโครงสร้าง (structured interview) เพื่อศึกษาความต้องการยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทางสู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพโดยการตรวจสอบคุณภาพของเครื่องมือโดยผู้เชี่ยวชาญ จำนวน 3 คน ได้ข้อคำถามที่มีค่า IOC เท่ากับ 1.00 และจัดทำแบบสัมภาษณ์แบบมีโครงสร้างฉบับสมบูรณ์ เพื่อนำไปเก็บข้อมูลกับกลุ่มตัวอย่างในการวิจัยครั้งนี้

3. การเก็บรวบรวมข้อมูล ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลด้วยตนเองกับผู้อำนวยการโรงเรียนบ้านสะพานเคียน โรงเรียนลำลอง และโรงเรียนบ้านพรุหาว จำนวน 3 คน โดยการสัมภาษณ์อย่างเป็นทางการ เริ่มจากผู้วิจัยแนะนำตนเอง สร้างบรรยากาศที่เป็นกันเอง ใช้คำถามที่ชัดเจน หลีกเลี่ยงคำถามที่ขึ้นคำตอบหรือเป็นผลกระทบต่ออารมณ์ของผู้ให้สัมภาษณ์ จดบันทึก และขออนุญาต หากมีการบันทึกเสียง หรือถ่ายรูป

4. การวิเคราะห์ข้อมูล

ข้อมูลเชิงคุณภาพที่ได้จากการสัมภาษณ์ วิเคราะห์ข้อมูลด้วยวิธีการวิเคราะห์เชิงเนื้อหา (content analysis)

ขั้นตอนที่ 2 การพัฒนา (D₁ : Development) การพัฒนาหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ

1. กลุ่มเป้าหมาย

กลุ่มเป้าหมายแบ่งออกเป็น 2 ส่วน คือ การพัฒนาหลักสูตร ได้แก่ ผู้อำนวยการโรงเรียน 3 คน ตามข้อมูลในขั้นตอนที่ 1 และ การวิพากษ์หลักสูตร ได้แก่ ผู้เชี่ยวชาญด้านการพัฒนาหลักสูตร จำนวน 3 คน ประกอบไปด้วย อาจารย์ผู้มีความเชี่ยวชาญด้านการพัฒนาหลักสูตร ระดับตำแหน่งทางวิชาการ (ผู้ช่วยศาสตราจารย์) จำนวน 1 คน ศึกษานิเทศก์ จำนวน 2 คน โดยใช้วิธีการเลือกแบบเจาะจง (purposive sampling)

2. เครื่องมือที่ใช้ในการวิจัย

ประเด็นสนทนากลุ่ม (focus group discussion) โดยสนทนาในประเด็นเกี่ยวกับความเหมาะสมและความสอดคล้องในองค์ประกอบของหลักสูตรฝึกอบรม โดยผ่านการตรวจสอบคุณภาพของเครื่องมือจากผู้เชี่ยวชาญ จำนวน 3 คน ได้ข้อคำถามที่มีค่า IOC เท่ากับ 1.00 และจัดทำประเด็นสนทนากลุ่มฉบับสมบูรณ์ เพื่อนำไปเก็บข้อมูลกับกลุ่มตัวอย่างในการวิจัยครั้งนี้

3. การเก็บรวบรวมข้อมูล

ผู้วิจัยทำหน้าที่เป็นผู้ดำเนินการสนทนา (moderator) กำหนดประเด็นสนทนากลุ่ม (focus group discussion) โดยสนทนาในประเด็นเกี่ยวกับความเหมาะสม และความสอดคล้องในองค์ประกอบของหลักสูตรฝึกอบรม และข้อเสนอแนะเพิ่มเติมในการพัฒนาหลักสูตร

4. การวิเคราะห์ข้อมูล

ข้อมูลเชิงคุณภาพที่ได้จากการสนทนากลุ่ม และความเหมาะสมของหลักสูตร ผู้วิจัยได้ทำการวิเคราะห์ข้อมูลด้วยวิธีการวิเคราะห์เชิงเนื้อหา (content analysis)

ขั้นตอนที่ 3 การวิจัย (R₂ : Research) การทดลองใช้หลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ

1. กลุ่มเป้าหมาย

กลุ่มเป้าหมาย ได้แก่ ผู้บริหารสถานศึกษา ครูผู้สอน และบุคลากรทางการศึกษา จากโรงเรียนบ้านสะพานเคียน โรงเรียนบ้านลำลอง โรงเรียนบ้านพรหาวา ซึ่งได้มาโดยวิธีการเลือกแบบเจาะจง (purposive sampling) จำนวนทั้งสิ้น 80 คน ตามข้อมูลในขั้นตอนที่ 1

2. เครื่องมือที่ใช้ในการวิจัย

แบบทดสอบความรู้ เรื่อง ยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทางสู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ จำนวน 1 ชุด จำนวน 20 ข้อ เป็นแบบทดสอบชนิดปรนัยลักษณะรูปแบบเลือกตอบ โดยวัดองค์ประกอบความรู้ในเรื่องตัวชี้วัดของสถานศึกษา และการยกระดับมาตรฐานการเรียนการสอนในชั้นเรียน แบบทดสอบความรู้ได้ผ่านการตรวจสอบคุณภาพของเครื่องมือจากผู้เชี่ยวชาญ จำนวน 3 คน ได้ข้อคำถามที่มีค่า IOC เท่ากับ 1.00 นำไป (try out) กับครูโรงเรียนนาทวีวิทยาคม จำนวน 30 คน เนื่องจากโรงเรียนดังกล่าวอยู่ในพื้นที่ใกล้เคียงกับโรงเรียนที่ใช้เป็นกลุ่มเป้าหมาย ผู้วิจัยจึงเร่งเห็นถึงความสอดคล้องในบริบทดังกล่าวจึงใช้โรงเรียนดังกล่าวเป็นกลุ่ม (try out) โดยมีค่าความยากระหว่าง 0.28 - 0.74 มีค่าอำนาจจำแนก ระหว่าง 0.36 - 0.89 และมีค่าความเชื่อมั่นเท่ากับ 0.79 และจัดทำแบบทดสอบฉบับสมบูรณ์เพื่อนำไปเก็บข้อมูลกับกลุ่มตัวอย่างในการวิจัยครั้งนี้

3. การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลในขั้นตอนการเตรียมการทดลองใช้หลักสูตรได้มีการเตรียมความพร้อม ในส่วนเอกสารประกอบการฝึกอบรม พร้อมประสานงาน และแจ้งกำหนดการฝึกอบรมให้แก่ผู้เข้ารับการฝึกอบรมทราบ ลำดับถัดไป คือขั้นตอนการดำเนินการทดลองใช้หลักสูตร มีการทดสอบความรู้ของผู้เข้าอบรม (pre - test) และดำเนินการฝึกอบรมตามหลักสูตรที่กำหนดไว้ และขั้นตอนสุดท้าย คือการประเมินผลหลังการฝึกอบรมด้วยวิธีการทดสอบความรู้ของผู้เข้าอบรม (post - test) เพื่อนำผลการวัดและประเมินผลการเรียนรู้ มาเปรียบเทียบกับองค์ความรู้ที่ได้รับหลังการฝึกอบรม จากนั้นตรวจสอบข้อมูลที่ได้จากการทดลองใช้หลักสูตรฝึกอบรมให้สมบูรณ์เพื่อใช้ในการวิเคราะห์ข้อมูล ต่อไป

4. การวิเคราะห์ข้อมูล

ผู้วิจัยนำข้อมูลที่ได้จากแบบทดสอบความรู้ มาวิเคราะห์โดยการเปรียบเทียบคะแนนก่อนและหลังการฝึกอบรม โดยการหาค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ในกลุ่มเดียวกัน

ขั้นตอนที่ 4 การพัฒนา (D₂ : Development) การประเมินผลและปรับปรุงหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ

1. กลุ่มเป้าหมาย

กลุ่มเป้าหมาย ได้แก่ ผู้บริหารสถานศึกษา ครูผู้สอน และบุคลากรทางการศึกษา จากโรงเรียนบ้านสะพานเคียน โรงเรียนบ้านลำลอง โรงเรียนบ้านพรุหว่า ซึ่งได้มาโดยวิธีการเลือกแบบเจาะจง (purposive sampling) จำนวนทั้งสิ้น 80 คน ตามข้อมูลในขั้นตอนที่ 1

2. เครื่องมือที่ใช้ในการวิจัย

แบบประเมินผลการใช้หลักสูตรฝึกอบรมฯ เป็นแบบสอบถามแบบมาตราประมาณค่า (rating scale) 5 ระดับ จำนวน 1 ชุด ประกอบด้วยข้อคำถาม จำนวน 5 ด้าน 20 ข้อ โดยผ่านการตรวจสอบคุณภาพของเครื่องมือ มีการตรวจสอบความตรงเชิงเนื้อหา (content validity) จากผู้เชี่ยวชาญ จำนวน 3 ท่าน ได้ข้อคำถามที่มีค่า IOC อยู่ระหว่าง 0.60 ถึง 1.00 นำไปทดลองใช้ (try out) กับครูโรงเรียนนาทวีวิทยาคม จำนวน 30 คน ตามข้อมูลในขั้นตอนที่ 3 จากนั้นนำข้อมูลมาวิเคราะห์หาความเที่ยงโดยใช้ค่าสัมประสิทธิ์อัลฟาของครอนบาค (Cronbach's alpha coefficient) ซึ่งได้ค่าความเที่ยงทั้งฉบับเท่ากับ 0.961 และจัดทำแบบทดสอบฉบับสมบูรณ์เพื่อนำไปเก็บข้อมูลกับกลุ่มตัวอย่างในการวิจัยครั้งนี้

3. การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล โดยเดินทางไปเก็บรวบรวมข้อมูลจากผู้เข้าร่วมอบรม จำนวน 80 คน โดยการใช้แบบประเมินความพึงพอใจ เริ่มจากผู้วิจัยแนะนำตนเอง สร้างบรรยากาศที่เป็นกันเอง ใช้แบบประเมินกับกลุ่มเป้าหมาย จากนั้นจัดเตรียมข้อมูลที่ได้จากการประเมินความพึงพอใจในการพัฒนาโปรแกรมการจัดทำแบบบันทึกผลการเรียนประจำรายวิชา เพื่อใช้ในการวิเคราะห์ข้อมูลต่อไป

4. การวิเคราะห์ข้อมูล

ผู้วิจัยนำข้อมูลที่ได้จากแบบประเมินความพึงพอใจ โดยใช้สถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ($\bar{X}$) และส่วนเบี่ยงเบนมาตรฐาน (S.D.)

การวิเคราะห์ข้อมูลผลการประเมินการใช้หลักสูตรฝึกอบรมฯ ผลการประเมินพิจารณาค่าเฉลี่ยคะแนนความพึงพอใจ และนำมาแปลความหมายตามเกณฑ์ ดังนี้ (Nilphan, 2015)

ค่าเฉลี่ยคะแนนอยู่ระหว่าง 4.50 - 5.00 หมายถึง มีความพึงพอใจระดับมากที่สุด

ค่าเฉลี่ยคะแนนอยู่ระหว่าง 3.50 - 4.49 หมายถึง มีความพึงพอใจระดับมาก

ค่าเฉลี่ยคะแนนอยู่ระหว่าง 2.50 - 3.49 หมายถึง มีความพึงพอใจระดับปานกลาง

ค่าเฉลี่ยคะแนนอยู่ระหว่าง 1.50 - 2.49 หมายถึง มีความพึงพอใจระดับน้อย

ค่าเฉลี่ยคะแนนอยู่ระหว่าง 1.00 - 1.49 หมายถึง มีความพึงพอใจระดับน้อยที่สุด

5. การปรับปรุงหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง

– ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ ดำเนินการหลังจากการทดลองใช้หลักสูตร โดยผู้วิจัยได้สังเกตผลระหว่างการใช้หลักสูตรจากการสังเกตพฤติกรรม การตอบคำถาม ข้อเสนอแนะต่าง ๆ และข้อบกพร่องที่พบขณะทดลองใช้หลักสูตร ผลจากการทำแบบทดสอบความรู้ และผลการประเมินผลการใช้หลักสูตรฯ นำมาวิเคราะห์เพื่อปรับปรุงแก้ไขให้หลักสูตรมีความสมบูรณ์ยิ่งขึ้น เพื่อประโยชน์ต่อการพัฒนาหลักสูตรฯ ที่ส่งต่อการยกระดับความรู้ด้านตัวชี้วัดต่อไป

ผลการวิจัย

ผลการวิจัย พบว่า ผู้วิจัยใช้ระเบียบวิธีของการวิจัยและพัฒนา (Research and Development : R & D) โดยมีขั้นตอนการดำเนินการวิจัย 4 ขั้นตอน ผู้วิจัยได้นำเสนอผลการวิจัยตามวัตถุประสงค์ ดังต่อไปนี้

1. ผลการศึกษาความต้องการฯ โดยการสัมภาษณ์คณะผู้บริหารโรงเรียนบ้านสะพานเคียน จำนวน 3 คน พบว่า ความต้องการในการพัฒนาฯ สรุปได้จำนวน 3 ประเด็น ได้แก่ ประเด็นที่ 1 ควรเน้นการสร้าง/ออกแบบหลักสูตรให้สอดคล้องกับบริบทในสถานศึกษา เพื่อให้บุคลากรทางการศึกษา มีความรู้ ความเข้าใจ เกี่ยวกับการสร้าง/ออกแบบหลักสูตรให้สอดคล้องกับความต้องการของผู้เรียน และให้ผู้เรียนมีความสุขในการเรียนตามการออกแบบหลักสูตรของสถานศึกษา ประเด็นที่ 2 ควรเน้นเกี่ยวกับการวิเคราะห์ตัวชี้วัดต้นทาง ระหว่างทาง และปลายทางให้สอดคล้องกับบริบทในสถานศึกษา เพื่อให้ครูผู้สอนแต่ละกลุ่มสาระการเรียนรู้ เข้าถึงและเข้าใจในส่วนประกอบของตัวชี้วัด และสามารถแยกตัวชี้วัดในระยะต่าง ๆ ควรใช้เครื่องมือวัดและประเมินผล และสร้างเนื้อหาให้สอดคล้องกับบริบทตามที่ตัวชี้วัดได้กำหนด เข้าใจถึงวัตถุประสงค์ของบริบท แต่ละตัวชี้วัดเป็นอย่างไร และเมื่อวิเคราะห์ตัวชี้วัดแล้วนั้น ผลที่ได้จากการวิเคราะห์ตัวชี้วัดจะทำให้ทราบถึง แนวทางการวัดและประเมินผลการเรียนรู้ รวมไปถึงแนวทางในการจัดการเรียนการสอน และประเด็นที่ 3 การนำหลักสูตรสถานศึกษาไปใช้ในห้องเรียน ประกอบด้วยตัวชี้วัดการเรียนรู้ แผนการจัดการเรียนรู้ วิธีการจัดการเรียนรู้ เครื่องมือที่ใช้การวัดและประเมินผลการเรียนรู้ และวิธีการวัดและประเมินผลการเรียนรู้ ควรเน้น การนำหลักสูตรหรือเขียนแผนการจัดการเรียนรู้ที่สามารถนำไปใช้ปฏิบัติได้จริงตามบริบทแต่ละโรงเรียน โดยการนำแผนการจัดการเรียนรู้/แผนการจัดการเรียนรู้นำไปสู่เชิงปฏิบัติได้จริง และสามารถใช้เครื่องมือ การวัดและประเมินผลการเรียนรู้ที่สอดคล้องกับแผนการจัดการเรียนรู้ สอดคล้องกับบริบทของสถานศึกษา ได้อย่างแท้จริง

2. ผลการพัฒนาหลักสูตรฝึกอบรมฯ โดยลักษณะของหลักสูตรเป็นการบรรยายให้ความรู้ และฝึกปฏิบัติ โดยการวิพากษ์ตัวชี้วัดในระดับต่าง ๆ พร้อมทั้งแลกเปลี่ยนองค์ความรู้ระหว่างวิทยากรและผู้เข้ารับการอบรม และมีการส่งเสริมความรู้ความเข้าใจเกี่ยวกับการยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง ผู้พัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพที่แท้จริง เพื่อยกระดับองค์ความรู้ของผู้เข้ารับการอบรมให้เป็นไปตามมาตรฐานที่กำหนดไว้ในหลักสูตร ซึ่งในการพัฒนาหลักสูตรฝึกอบรมฯ ผู้วิจัยได้ร่วมพัฒนาหลักสูตรกับผู้อำนวยการโรงเรียนบ้านสะพานเคียน โรงเรียนลำลอง และโรงเรียนบ้านพรุหาว จำนวน 3 คน ซึ่งได้วิเคราะห์องค์ประกอบของหลักสูตรออกมาเป็น 5 หมวด ได้แก่ หมวดที่ 1 ข้อมูลทั่วไป หมวดที่ 2 ข้อมูลเฉพาะของหลักสูตร หมวดที่ 3 การดำเนินการและโครงสร้างของหลักสูตรฝึกอบรม หมวดที่ 4 หลักเกณฑ์ในการประเมินผล และหมวดที่ 5 การประเมินและปรับปรุงกระบวนการดำเนินการของหลักสูตรฝึกอบรม โดยมีองค์ประกอบในแต่ละหมวด ดังนี้

หมวดที่ 1 ข้อมูลทั่วไป ประกอบด้วย 1) รหัสและชื่อหลักสูตร 2) จำนวนชั่วโมงที่เรียนตลอดหลักสูตร 3) ประเภทของหลักสูตร 4) สถานภาพของหลักสูตร 5) ข้อมูลทั่วไปของวิทยากรประจำหลักสูตรฝึกอบรม 6) สถานที่จัดฝึกอบรมหลักสูตร และ 7) จุดเน้นของหลักสูตร

หมวดที่ 2 ข้อมูลเฉพาะของหลักสูตรฝึกอบรม ประกอบด้วย 1) หลักการและเหตุผล 2) วัตถุประสงค์ของหลักสูตรฝึกอบรม 3) เป้าหมาย 4) เนื้อหาสาระของหลักสูตรฝึกอบรม 5) ระยะเวลาของหลักสูตรฝึกอบรม 6) โครงสร้างหลักสูตรฝึกอบรม 7) กิจกรรมการฝึกอบรม 8) สื่อการฝึกอบรม และ 9) การวัดและประเมินผล

หมวดที่ 3 การดำเนินการและโครงสร้างของหลักสูตรฝึกอบรม ประกอบด้วย 1) การดำเนินการหลักสูตรฝึกอบรม และ 2) หลักสูตรฝึกอบรม

หมวดที่ 4 หลักเกณฑ์ในการประเมินผล ประกอบด้วย 1) หลักเกณฑ์ในการผ่านการหลักสูตรฝึกอบรม และ 2) เกณฑ์การสำเร็จของหลักสูตรฝึกอบรม

หมวดที่ 5 การประเมินและปรับปรุงกระบวนการดำเนินการของหลักสูตรฝึกอบรม ประกอบด้วย 1) การประเมินผลหรือวิธีการวัด (Assessment) ความสามารถ และหรือผลการเรียนรู้ในเชิงคุณภาพอย่างชัดเจน และ 2) แนวทางการประเมินและปรับปรุงกระบวนการดำเนินการของหลักสูตรฝึกอบรม

3. ผลการทดลองใช้หลักสูตรฝึกอบรมฯ ผู้วิจัยได้มีเป้าหมาย ดังนี้ 1) ประเมินผลการเรียนรู้ระหว่างอบรมเพื่อเป็นข้อมูลในการนำผลการประเมินไปพัฒนาผู้เข้ารับการอบรมให้บรรลุตามเป้าหมายของแต่ละหน่วยการเรียนรู้ โดยกำหนดแนวทางการวัดและประเมินผล ดังนี้ ระหว่างการจัดกิจกรรมของแต่ละหน่วยการเรียนรู้จะทำการวัดและประเมินการเรียนรู้โดยใช้การสังเกตและการโต้ตอบคำถามเพื่อใช้เป็นข้อมูลในการพัฒนาผู้เข้ารับการอบรมให้เป็นไปตามจุดประสงค์การเรียนรู้ในแต่ละหน่วย ปรากฏว่า ผู้เข้าอบรมได้มีการโต้ตอบและแลกเปลี่ยนเรียนรู้กับวิทยากรในทั้ง 3 หน่วยการเรียนรู้ ในเรื่อง 1) ความเป็นมาพัฒนาการหลักสูตรสถานศึกษา 2) การวิเคราะห์พฤติกรรมกรรมการเรียนรู้ของตัวชี้วัดระหว่างทางและตัวชี้วัดปลายทาง ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 ฉบับปรับปรุงแก้ไข พ.ศ. 2566 และ 3) การนำหลักสูตรสถานศึกษาสู่แผนการจัดการเรียนรู้ในชั้นเรียน ทำให้ผู้เข้าอบรมเกิดการเรียนรู้ แลกเปลี่ยนกระบวนการความคิดเห็นต่างๆ ในองค์ความรู้ที่ตนเองมี วิทยากรมีการกระตุ้นด้วยการถามตอบเพื่อให้ผู้เข้าอบรมเกิดความสนใจ กระตือรือร้นที่จะเข้าอบรม ซึ่งในการบรรยายเป็นการเปิดโอกาสให้ผู้เข้าอบรมได้แสดงความคิดเห็นแลกเปลี่ยนความรู้ทุกช่วงระยะเวลาในการดำเนินกิจกรรม ทำให้ผู้เข้าอบรมเกิดองค์ความรู้ หรือแนวทางใหม่ ๆ ในระหว่างการบรรยาย เป็นการเพิ่มเติมความรู้ที่ขาดหายไปของตนเอง และ 2) เพื่อประเมินผลหลักสูตรฝึกอบรมฯ โดยกำหนดแนวทางการวัดและประเมินผล ดังนี้ เมื่อดำเนินการจัดกิจกรรมครบทุกหน่วยการเรียนรู้ จะดำเนินการประเมินผลหลักสูตรฝึกอบรมฯ โดยใช้แบบทดสอบและแบบประเมินผลการใช้หลักสูตรที่มีความสอดคล้องกับวัตถุประสงค์ของหลักสูตรฝึกอบรมฯ ซึ่งมีการกำหนดเกณฑ์เพื่อวัดประสิทธิภาพของหลักสูตรฝึกอบรมฯ จากการทดสอบความรู้ของผู้เข้าอบรมมีผลสัมฤทธิ์เกี่ยวกับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพหลังการเข้ารับการฝึกอบรมโดยมีค่า ($\bar{X}$ = 11.90) สูงกว่าก่อนการเข้ารับการฝึกอบรมโดยมีค่าเฉลี่ย ($\bar{X}$ = 9.26) ดังปรากฏในตาราง 1

ตาราง 1 ผลการเปรียบเทียบผลสัมฤทธิ์ เรื่อง การยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทางสู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ

ความรู้เกี่ยวกับมาตรฐานการสอนฯ	n	คะแนนเต็ม	คะแนนได้	
			$\bar{X}$	S.D.
ก่อนการอบรม	80	20	9.26	2.57
หลังการอบรม	80	20	11.90	2.63

4. ผลการประเมินและปรับปรุงหลักสูตรฝึกอบรมฯ พบว่า ผลการประเมินผลการใช้หลักสูตรฝึกอบรมฯ พบว่า ผลการประเมินการใช้หลักสูตรฝึกอบรมฯ ในภาพรวมอยู่ในระดับ มากที่สุด เมื่อพิจารณาเป็นรายด้าน พบว่า ผลการประเมินการใช้หลักสูตรฝึกอบรมฯ ในภาพรวมอยู่ในระดับมากที่สุด ดังปรากฏในตาราง 2

ตาราง 2 ผลการประเมินการใช้หลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชี้วัดระหว่างทาง – ตัวชี้วัดปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ (n = 80)

รายการ	ระดับความพึงพอใจ		
	$\bar{X}$	S.D.	ความหมาย
1. ด้านกระบวนการ/ ขั้นตอนการจัดกิจกรรม	4.57	0.51	มากที่สุด
1.1) รูปแบบในการจัดกิจกรรม	4.58	0.56	มากที่สุด
1.2) ลำดับขั้นตอนในการจัดกิจกรรม	4.56	0.59	มากที่สุด
1.3) ระยะเวลาที่ใช้ในการจัดกิจกรรม	4.55	0.57	มากที่สุด
1.4) สื่อที่ใช้ในการจัดกิจกรรม	4.58	0.54	มากที่สุด
2. ด้านเนื้อหาที่ใช้ในการจัดกิจกรรม	4.65	0.43	มากที่สุด
2.1) เนื้อหาสอดคล้องกับวัตถุประสงค์ของโครงการ	4.62	0.48	มากที่สุด
2.2) เนื้อหาที่บรรยายสามารถนำไปบูรณาการได้	4.63	0.50	มากที่สุด
2.3) เนื้อหาที่บรรยายมีความเข้าใจง่าย	4.62	0.48	มากที่สุด
2.4) เนื้อหาที่บรรยายมีความทันสมัย	4.78	0.44	มากที่สุด
3. ด้านวิทยากร	4.68	0.42	มากที่สุด
3.1) ประสบการณ์ในเรื่องที่บรรยาย	4.65	0.47	มากที่สุด
3.2) ความรู้ในเรื่องที่บรรยาย	4.70	0.48	มากที่สุด
3.3) เทคนิคในการถ่ายทอดความรู้	4.67	0.49	มากที่สุด
3.4) การใช้ภาษาที่เข้าใจง่าย	4.72	0.44	มากที่สุด
4. ด้านสิ่งอำนวยความสะดวก	4.62	0.47	มากที่สุด
4.1) การประชาสัมพันธ์ของโครงการฯ	4.25	0.55	มากที่สุด
4.2) ระบบการเข้าถึงไฟล์ข้อมูลต่าง ๆ ของการจัดกิจกรรม	4.63	0.50	มากที่สุด
4.3) ความชัดเจนของภาพ เสียงในระบบออนไลน์	4.65	0.55	มากที่สุด
4.4) การควบคุมสื่อ โสตทัศนูปกรณ์ผ่านระบบออนไลน์	4.67	0.49	มากที่สุด

ตาราง 2 (ต่อ)

รายการ	ระดับความพึงพอใจ		
	$\bar{X}$	S.D.	ความหมาย
5. ด้านคณะกรรมการดำเนินโครงการฯ	4.66	0.46	มากที่สุด
5.1) การติดต่อประสานงานคณะกรรมการดำเนินโครงการฯ	4.62	0.53	มากที่สุด
5.2) การให้บริการที่เป็นธรรม (ไม่เลือกปฏิบัติ)	4.67	0.49	มากที่สุด
5.3) ความกระตือรือร้นในการให้ความช่วยเหลือ	4.68	0.46	มากที่สุด
5.4) ความสุภาพของคณะกรรมการดำเนินโครงการฯ	4.67	0.52	มากที่สุด
รวม	4.64	0.43	มากที่สุด

ส่วนการปรับปรุงหลักสูตรฝึกอบรมฯ ดำเนินการหลังจากการทดลองใช้หลักสูตร โดยผู้วิจัยได้บันทึกผลระหว่างการใช้หลักสูตรจากการสังเกตพฤติกรรมการตอบคำถาม ข้อเสนอแนะต่าง ๆ และข้อบกพร่องที่พบขณะทดลองใช้หลักสูตร ผลจากการทำแบบทดสอบความรู้ และผลการประเมินผลการใช้หลักสูตรฯ นำมาวิเคราะห์ พบว่า เนื่องจากจำนวนชั่วโมงในการจัดกิจกรรม มีเพียงแค่จำนวน 6 ชั่วโมง ซึ่งค่อนข้างจำกัดในการจัดกิจกรรม และกิจกรรมส่วนใหญ่เป็นกิจกรรมในลักษณะฟังบรรยายในรูปแบบออนไลน์ จึงทำให้ตัวกิจกรรมไม่น่าดึงดูดเท่าที่ควร และกิจกรรมไม่ได้นำไปสู่ภาคการปฏิบัติ ซึ่งด้วยระยะเวลา ที่มีค่อนข้างจำกัด จึงไม่สามารถนำผู้เข้าอบรมสู่การปฏิบัติได้ และด้วยอุปสรรคในการจัดโครงการในรูปแบบออนไลน์ ไม่ได้เอื้อต่อการนำไปสู่การปฏิบัติอย่างแท้จริง ถึงแม้ทางทีมวิทยากรและผู้จัดโครงการ ได้พยายามออกแบบรูปแบบการจัดโครงการให้เอื้อต่อการปฏิบัติ แต่ก็ยังไม่ประสบความสำเร็จเท่าที่ควรเพราะระยะเวลาที่มีอย่างจำกัดกับสถานการณ์ในการจัดที่ไม่ค่อยเอื้อต่อการปฏิบัติ ควรกำหนดจุดเน้นของหลักสูตร และมุ่งเน้นอธิบายขยายความที่เกี่ยวข้องกับการจัดการเรียนรู้ที่สอดคล้องกับตัวชีวิต และควรมีการระบุจุดประสงค์ กิจกรรมในการพัฒนาเครื่องมือการวัด สื่อที่ใช้ ตั้งแต่กระบวนการพัฒนาหลักสูตรจนถึงการประเมิน ติดตามความสำเร็จของหลักสูตรเป็นระยะ เพื่อดูความสำเร็จของหลักสูตร

อภิปรายผล

จากการพัฒนาหลักสูตรฝึกอบรมเพื่อยกระดับมาตรฐานการสอนที่สอดคล้องกับตัวชีวิตระหว่างทาง – ตัวชีวิตปลายทาง สู่การพัฒนาหลักสูตรสถานศึกษาอย่างมีคุณภาพ ผู้วิจัยได้อภิปรายผลตามวัตถุประสงค์ ดังนี้

1. จากผลการศึกษาความต้องการฯ ได้ 3 ประเด็น ดังนี้ 1. ควรเน้นการสร้าง/ออกแบบหลักสูตรให้สอดคล้องกับบริบทในสถานศึกษา 2. ควรเน้นเกี่ยวกับการวิเคราะห์ตัวชีวิตต้นทาง ระหว่างทาง และปลายทาง ให้สอดคล้องกับบริบทในสถานศึกษา และ 3. การนำหลักสูตรสถานศึกษาไปใช้ในห้องเรียน ประกอบด้วย ตัวชีวิตการเรียนรู้ แผนการจัดการเรียนรู้ วิธีการจัดการเรียนรู้ เครื่องมือที่ใช้การวัดและประเมินผลการเรียนรู้ และวิธีการวัดและประเมินผลการเรียนรู้ ซึ่งพบว่า คณะผู้บริหารต้องการให้บุคลากรของโรงเรียนเกิดความรู้ความเข้าใจเกี่ยวกับการสร้าง/ออกแบบหลักสูตรให้สอดคล้องกับบริบทในสถานศึกษามีความรู้ ความเข้าใจเกี่ยวกับการวิเคราะห์ตัวชีวิตต้นทาง ระหว่างทาง และปลายทาง ให้สอดคล้องกับบริบทในสถานศึกษา เพื่อให้ครูผู้สอนแต่ละกลุ่มสาระการเรียนรู้เข้าถึงและเข้าใจในส่วนประกอบของตัวชีวิต และสามารถแยกตัวชีวิต

ในระยะต่าง ๆ รวมไปถึงแนวทางในการจัดการเรียนการสอน และการนำหลักสูตรสถานศึกษาไปใช้ในห้องเรียน สอดคล้องกับบริบทของสถานศึกษาได้อย่างแท้จริง สอดคล้องกับกระบวนการพัฒนาหลักสูตรที่เหมาะสมกับปัญหาดังกล่าว คือ รูปแบบการพัฒนาหลักสูตรของไทเลอร์ (Ralph W. Tyler) ในองค์ประกอบที่ 1 ว่าด้วยการกำหนด จุดมุ่งหมายทางการศึกษาที่ควรบรรลุให้ชัดเจน โดยการสำรวจความต้องการหรือวิเคราะห์ข้อมูลเกี่ยวกับหลักสูตรจากการรวบรวมจากบุคลากร หรือสังเคราะห์จากเอกสารต่าง ๆ ที่เกี่ยวข้องกับการกำหนด จุดมุ่งหมายทางการศึกษา (Patphol, 2024) โดยสอดคล้องกับงานวิจัยของ Saengthong (2021) และ Somkratok (2020) ได้พัฒนาหลักสูตรสถานศึกษา และพัฒนาหลักสูตรฝึกอบรมสมรรถนะหลักสถานศึกษา พบว่า มีผลการสำรวจความต้องการที่คล้ายคลึงกัน คือ สถานศึกษาสร้างความตระหนักให้ครูและบุคลากรที่เกี่ยวข้องเห็นควรตระหนักและเห็นความสำคัญของการพัฒนาหลักสูตร มีการวิเคราะห์บริบทตัวชี้วัดและจุดเน้นของสถานศึกษา รวมไปถึงความต้องการและ ความจำเป็นต่อฐานสมรรถนะของหลักสูตร เพื่อกำหนดแนวทางในการพัฒนาหลักสูตรสถานศึกษา จึงทำให้หลักสูตรมีความสอดคล้องกับบริบทของสถานศึกษา และเกิดประสิทธิภาพในการดำเนินการจัดการเรียนการสอน

2. จากการพัฒนาหลักสูตรฝึกอบรมฯ นั้นคือ ขั้นตอนการสังเคราะห์เพื่อพัฒนาองค์ประกอบของหลักสูตรฝึกอบรมฯ พบว่า การสังเคราะห์เอกสารและงานวิจัยที่เกี่ยวข้อง รวมไปถึงการกำหนดประเด็นหารือกับคณะผู้บริหารโรงเรียน สามารถวิเคราะห์องค์ประกอบของหลักสูตร ออกเป็น 5 หมวด ได้แก่ หมวดที่ 1 ข้อมูลทั่วไป ได้แก่ 1) รหัสและชื่อหลักสูตร 2) จำนวนชั่วโมงที่เรียนตลอดหลักสูตร 3) ประเภทของหลักสูตร 4) สถานภาพของหลักสูตร 5) ข้อมูลทั่วไปของวิทยากรประจำหลักสูตรฝึกอบรม 6) สถานที่จัดฝึกอบรมหลักสูตร และ 7) จุดเน้นของหลักสูตร หมวดที่ 2 ข้อมูลเฉพาะของหลักสูตรฝึกอบรม ได้แก่ 1) หลักการและเหตุผล 2) วัตถุประสงค์ของหลักสูตรฝึกอบรม 3) เป้าหมาย 4) เนื้อหาสาระของหลักสูตรฝึกอบรม 5) ระยะเวลาของหลักสูตรฝึกอบรม 6) โครงสร้างหลักสูตรฝึกอบรม 7) กิจกรรมการฝึกอบรม 8) สื่อการฝึกอบรม และ 9) การวัดและประเมินผล หมวดที่ 3 การดำเนินการและโครงสร้างของหลักสูตรฝึกอบรม ได้แก่ 1) การดำเนินการหลักสูตรฝึกอบรม และ 2) หลักสูตรฝึกอบรม หมวดที่ 4 หลักเกณฑ์ในการประเมินผล ได้แก่ 1) หลักเกณฑ์ในการผ่านการหลักสูตรฝึกอบรม และ 2) เกณฑ์การสำเร็จของหลักสูตรฝึกอบรม และหมวดที่ 5 การประเมินและปรับปรุงกระบวนการดำเนินการของหลักสูตรฝึกอบรม ได้แก่ 1) การประเมินผลหรือวิธีการวัด (assessment) ความสามารถ และหรือผลการเรียนรู้ในเชิงคุณภาพอย่างชัดเจน และ 2) แนวทางการประเมินและปรับปรุงกระบวนการดำเนินการของหลักสูตรฝึกอบรมเมื่อนำไปวิพากษ์หลักสูตร โดยผู้เชี่ยวชาญ ผลการประเมินหลักสูตรฉบับร่าง พบว่า มีความสอดคล้องเหมาะสม เป็นไปตามเงื่อนไขของการพัฒนาหลักสูตรของไทเลอร์ (Ralph W. Tyler) ในองค์ประกอบที่ 2 องค์ประกอบของหลักสูตรต้องมีความพร้อมที่จะจัดในการประสบการณ์การเรียนรู้ที่สามารถทำให้บรรลุจุดมุ่งหมายของหลักสูตรได้ โดยมีขั้นตอนการพัฒนาหลักสูตร ได้แก่ 1) การศึกษาวิเคราะห์บริบท 2) การกำหนดจุดมุ่งหมายชั่วคราว 3) การกลั่นกรองจุดมุ่งหมายทางการศึกษา 4) การกำหนดจุดมุ่งหมายที่แท้จริง 5) การจัดกิจกรรมและประสบการณ์ และ 6) การประเมินผลการศึกษา (Patphol, 2024) และสอดคล้องกับงานวิจัยของ Mekdee & Sripatta (2022) และ Ratchada & Chulasuwan (2023) ได้ศึกษาแนวทางในการพัฒนาบริหารและพัฒนาหลักสูตรสถานศึกษา

พบว่า ประกอบด้วย 5 องค์ประกอบ ได้แก่ 1) หลักการ 2) จุดมุ่งหมาย 3) กลไกการดำเนินงาน 4) แนวทางการพัฒนาหลักสูตรสถานศึกษาโดยใช้การมีส่วนร่วม จำนวน 4 ด้าน 39 แนวทาง ดังนี้ 4.1) การเตรียมความพร้อมของสถานศึกษา จำนวน 6 แนวทาง 4.2) การจัดทำหลักสูตรสถานศึกษา จำนวน 10 แนวทาง 4.3) การนำหลักสูตรไปใช้ จำนวน 14 แนวทาง และ 4.4) การกำกับดูแลคุณภาพ จำนวน 9 แนวทาง และ 5) เงื่อนไขความสำเร็จ โดยผลการประเมินของผู้ทรงคุณวุฒิ พบว่า โดยภาพรวม มีความเหมาะสม สอดคล้องกับ (Surachot, 2020) ได้ศึกษาการพัฒนาหลักสูตรการจัดการเรียนรู้ในศตวรรษที่ 21 พบว่า ซึ่งมีองค์ประกอบของหลักสูตรประกอบด้วย เหตุผลและความสำคัญของหลักสูตร หลักการของหลักสูตร จุดมุ่งหมายของหลักสูตร เนื้อหาสาระของหลักสูตร โครงสร้างของหลักสูตร ระยะเวลาของหลักสูตร แผนฝึกอบรม กิจกรรมการฝึกอบรม สื่อการฝึกอบรม วิธีการวัดและประเมินผล ตารางฝึกอบรม และเอกสารประกอบหลักสูตร โดยมีผลการประเมินจากผู้เชี่ยวชาญว่ามีประสิทธิภาพตามเกณฑ์ คือ ทุกองค์ประกอบของหลักสูตรมีความเหมาะสมในระดับมาก

3. จากการทดลองใช้หลักสูตรฝึกอบรมฯ ได้มีเป้าหมาย ดังนี้ 1) ผู้เข้าอบรมได้มีการโต้ตอบและแลกเปลี่ยนเรียนรู้กับวิทยากรในทั้ง 3 หน่วยการเรียนรู้ ในเรื่อง 1) ความเป็นมาพัฒนาหลักสูตรสถานศึกษา 2) การวิเคราะห์พฤติกรรมการณ์การเรียนรู้ของตัวชี้วัดระหว่างทางและตัวชี้วัดปลายทาง ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 ฉบับปรับปรุงแก้ไข พ.ศ. 2566 และ 3) การนำหลักสูตรสถานศึกษาสู่แผนการจัดการเรียนรู้ในชั้นเรียน ทำให้ผู้เข้าอบรมเกิดการเรียนรู้ แลกเปลี่ยนกระบวนการความคิดเห็นต่าง ๆ ในองค์ความรู้ที่ตนเองมี วิทยากรมีการกระตุ้นด้วยการถามตอบเพื่อให้ผู้เข้าอบรมเกิดความสนใจ กระตือรือร้นที่จะเข้าอบรม ซึ่งในการบรรยายเป็นการเปิดโอกาสให้ผู้เข้าอบรมได้แสดงความคิดเห็นแลกเปลี่ยนความรู้ทุกช่วงระยะเวลาในการดำเนินกิจกรรม ทำให้ผู้เข้าอบรมเกิดองค์ความรู้ หรือแนวทางใหม่ ๆ ในระหว่างการบรรยายเป็นการเพิ่มเติมความรู้ที่ขาดหายไปของตนเอง และ 2) เพื่อประเมินผลหลักสูตรฝึกอบรมฯ โดยกำหนดแนวทางการวัดและประเมินผล ดังนี้ เมื่อดำเนินการจัดกิจกรรมครบทุกหน่วยการเรียนรู้ จะดำเนินการประเมินผลหลักสูตรฝึกอบรมฯ โดยใช้แบบทดสอบและแบบประเมินผลการใช้หลักสูตรที่มีความสอดคล้องกับวัตถุประสงค์ของหลักสูตรฝึกอบรมฯ ซึ่งมีการกำหนดเกณฑ์เพื่อวัดประสิทธิภาพของหลักสูตรฝึกอบรมฯ ได้สอดคล้องกับการพัฒนาหลักสูตรของไทเลอร์ (Ralph W. Tyler) ในองค์ประกอบที่ 3 คือ การจัดกิจกรรมของหลักสูตรจะต้องมีการ จัดกิจกรรมและประสบการณ์การเรียนรู้อย่างต่อเนื่องทั้งในด้านเวลาต่อเวลา เนื้อหาต่อเนื้อหา และยึดหยุ่นต่อผู้เรียน โดยใช้กระบวนการวัดและประเมินอยู่เป็นระยะ เพื่อเป็นสารสนเทศในการปรับปรุงและพัฒนากิจกรรมในตัวของหลักสูตร (Patphol, 2024) และเมื่อเปรียบเทียบด้านความรู้พบว่า ผู้เข้าอบรมมีค่าเฉลี่ยของผลสัมฤทธิ์หลังการเข้ารับการฝึกอบรมสูงกว่าก่อนการเข้ารับการฝึกอบรม สอดคล้องกับงานวิจัยของ Kitjapipat (2022); Saringkar (2021); Suttaluang (2022) ทั้ง 3 คน ได้ศึกษาเรื่องการพัฒนาหลักสูตรฝึกอบรมของสถานศึกษา พบว่า ค่าเฉลี่ยของคะแนนด้านความรู้หลังการฝึกอบรมสูงกว่าค่าเฉลี่ยคะแนนก่อนการฝึกอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทั้งนี้อาจจะเนื่องจากการจัดการเรียนรู้ของกิจกรรมมีเป้าหมายที่ชัดเจนและสามารถทำให้ผู้เข้าอบรมเกิดทักษะการเรียนรู้ตามที่เป้าหมาย

ของหลักสูตรฝึกอบรมได้กำหนด รวมไปถึงการได้รับการฝึกฝนให้นำความรู้ทักษะดังกล่าวในหลักสูตรไปใช้ หรือแก้ปัญหาในสถานการณ์ต่าง ๆ ในชีวิตจริง

4. ผลการประเมินและปรับปรุงหลักสูตรฝึกอบรมฯ ผลการประเมินผลการใช้หลักสูตรฝึกอบรมฯ ในภาพรวมอยู่ในระดับมากที่สุด จากการผลสังเกตระหว่างการใช้หลักสูตรจากการสังเกตพฤติกรรม การตอบคำถาม ข้อเสนอแนะต่าง ๆ และข้อบกพร่องที่พบขณะทดลองใช้หลักสูตรผลจากการทำแบบทดสอบความรู้ และผลการประเมินผลการใช้หลักสูตรฯ พบว่า กิจกรรมส่วนใหญ่เป็นกิจกรรมในลักษณะฟังบรรยายในรูปแบบออนไลน์ จึงทำให้ตัวกิจกรรมไม่น่าดึงดูดเท่าที่ควร และกิจกรรมไม่ได้นำไปสู่ภาคการปฏิบัติ ซึ่งด้วยระยะเวลาที่มีค่อนข้างจำกัด จึงไม่สามารถนำผู้เข้าอบรมสู่การปฏิบัติได้ และด้วยอุปสรรคในการจัดโครงการในรูปแบบออนไลน์ ไม่ได้เอื้อต่อการนำไปสู่การปฏิบัติอย่างแท้จริง ถึงแม้ทางทีมวิทยากรและผู้จัดโครงการ ได้พยายามออกแบบรูปแบบการจัดโครงการให้เอื้อต่อการปฏิบัติ แต่ก็ยังไม่ประสบความสำเร็จเท่าที่ควร ในส่วนขององค์ความรู้ กระบวนการจัดกิจกรรม วัดและประเมินผลการเรียนรู้ รวมไปถึงเกณฑ์การประเมินของหลักสูตรฝึกอบรมฯ มีความชัดเจนทุกในตามกระบวนการออกแบบหลักสูตร ซึ่งสอดคล้องกับองค์ประกอบที่ 4 ของการพัฒนาหลักสูตรของไทเลอร์ (Ralph W. Tyler) การจัดทำหลักสูตรต้องมีการวัดและประเมินผลที่ชัดเจน แล้วนำผลดังกล่าวไปใช้ในการปรับปรุงพัฒนาหลักสูตรให้มีบรรลุจุดมุ่งหมายที่กำหนด และลดปัญหาในการใช้งานในครั้งถัด (Patphol, 2024) และสอดคล้องกับ Kitjapipat (2022); Lertwattanawanich (2017); Saengthong (2021) พบว่า ผลการประเมินการใช้หลักสูตรการส่งเสริมทักษะอาชีพ หลักสูตรฝึกอบรมเสริมสร้างการออกแบบการเรียนรู้ และหลักสูตรสถานศึกษา ทั้ง 3 หลักสูตรมีความพอใจในระดับมากที่สุด และมีผลการปรับปรุงที่สอดคล้องกันทั้ง 3 หลักสูตร เนื่องจากทั้ง 3 หลักสูตรมีเวลาที่ค่อนข้างจำกัดในการจัดกิจกรรมให้แก่ผู้เรียนจึงทำให้ผู้เรียนไม่เข้าใจในเนื้อหากิจกรรมของหลักสูตรเท่าที่ควร ซึ่งกิจกรรมส่วนใหญ่เป็นการจัดกิจกรรมในรายวิชาชุมชนไม่มีการบูรณาการรายวิชาอื่นเข้ามามีส่วนร่วมเท่าที่ควร ทำให้ตัวกิจกรรมที่ส่งผลถึงการเรียนรู้ของผู้เรียน ไม่สามารถเกิดได้อย่างสมบูรณ์ เนื่องจากหมดเวลาเสียก่อน และเนื่องจากเวลาที่ใช้ในการจัดกิจกรรมค่อนข้างจำกัด ทำให้บางกิจกรรมไม่สามารถจัดกิจกรรมจบได้ภายในชั่วโมงนั้น ส่วนการวัดและประเมินผลชิ้นงานของนักเรียนในเรื่องของเกณฑ์การประเมินชิ้นงานที่อาจจะใช้ภาษาที่ยากเกินไปทำให้เกิดการไม่เข้าใจในการวัดและประเมินชิ้นงานของนักเรียน ซึ่งได้ทำการปรับเกณฑ์ใหม่ให้เข้าใจง่ายขึ้นใช้ภาษาที่ง่ายขึ้น มีความกระชับ เพื่อให้สามารถวัดและประเมินผลชิ้นงานของนักเรียนได้ตรงตามวัตถุประสงค์ที่ได้กำหนดไว้

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 ควรสร้างและออกแบบหลักสูตรให้มีความสอดคล้องกับบริบทของสถานศึกษาตามการออกแบบหลักสูตรของสถานศึกษา

1.2 ควรพัฒนาแนวทางการจัดการเรียนการสอน และการเลือกใช้เครื่องมือในการวัดและประเมินผลการเรียนรู้ให้สอดคล้องและเหมาะสมกับบริบทตามที่ตัวชี้วัดได้กำหนด

1.3 ควรพัฒนาแนวทางการนำหลักสูตรและการเขียนแผนการจัดการเรียนรู้ที่สามารถนำไปใช้ปฏิบัติได้จริงตามบริบท และมีความสอดคล้องของแต่ละโรงเรียนอย่างแท้จริง

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรเพิ่มกิจกรรมในหลักสูตรฝึกอบรมที่ส่งเสริมให้ผู้เข้ารับการอบรมเกิดทักษะเชิงสร้างสรรค์ให้เกิดชิ้นงานในการเรียนรู้ร่วมกันตามกระบวนการชุมชนแห่งการเรียนรู้เชิงวิชาชีพ (PLC)

2.2 ควรศึกษาวิจัยปัญหาการบริหารหลักสูตรสถานศึกษาในเชิงลึก เพื่อเป็นแนวทางในการบริหารหลักสูตรให้มีประสิทธิภาพ โดยศึกษาจากประชากรหรือกลุ่มตัวอย่างที่หลากหลาย เช่น ครู คณะกรรมการสถานศึกษา และผู้นำชุมชน เป็นต้น

2.3 ควรศึกษาวิจัยเปรียบเทียบการบริหารหลักสูตรสถานศึกษา ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ตามขนาดของโรงเรียน

Translate Thai References

- Kitjapipat, T. (2022). *Development of training courses to strengthen design abilities. Proactive learning management for primary school teachers In the Bhumtawanwittaya Foundation School* [Master's Thesis]. Dhurakij Pundit University. (in Thai)
- Lertwattanawanich, S. (2017). *Development of project-based learner development activities to promote career skills and career characteristics for secondary school students.* [Doctoral dissertation]. Silpakorn University. (in Thai)
- Mekdee, C., & Sripatta, T. (2022). *Study of conditions and guidelines for administration of educational institution curricula. Under the jurisdiction of the Secondary Educational Service Area Office In Phitsanulok Province. Journal of Roi Kaensarn Academi, 7(1), 652 – 665.* (in Thai)
- Nilphan, M. (2015). *Educational research methods* (5th ed.). Silpakorn University Printing House. (in Thai)
- Nutrawong, R. (2010). *Summary of research results from pilot use of curriculum. Basic Education Curriculum, B.E. 2008* [Research report]. Ministry of Education. (in Thai)
- Patphol, M. (2024). *Contemporary curriculum development model* (2nd ed.). Graduate School Srinakharinwirot University. (in Thai)
- Ratchada, Y., & Chulasuwan, S. (2023). Guidelines for developing school curriculum using participation for schools to expand educational opportunities. Under the jurisdiction of the Maha Sarakham Primary Educational Service Area Office, Area 3. *Buddhist Philosophy Wiwat Journal, 7(4), 160 – 665.* (in Thai)

- Ratchakhom, P. (2014). Educational institution curriculum management. According to the 2008 Basic Education Core Curriculum of educational institutions under the Chiang Rai Primary Educational Service Area Office. *Journal of Education Khon Kaen University*, 37(4), 94 – 102. (in Thai)
- Saengthong, A. (2021). *Development of the educational curriculum of Wat Chandrawat School (Sukprasarat)* [Master's Independent research]. Silpakorn University. (in Thai)
- Sakulpramote, S. (2014). *Reform of the basic education curriculum*. Secretariat Printing House House of Representatives. (in Thai)
- Saringkar, N. (2021). *Development of a training curriculum to enhance competency in teaching Dhamma studies according to the concept of Active Learning (Active Learning) for monks teaching morality* [Doctoral dissertation]. Silpakorn University. (in Thai)
- Somkratok, C. (2020). *Developing a training curriculum by applying the concept of competency-based learning management in collaboration with a professional learning community to strengthen the ability to design learning management on Life safety of health education teachers* [Doctoral dissertation]. Chulalongkorn University. (in Thai)
- Surachot, C. (2020). *Development of learning management curriculum in the 21st century using the professional learning community (PLC) process* [Doctoral dissertation]. Burapha University. (in Thai)
- Suttaluang, S. (2022). *Development of a training curriculum to enhance competency in active learning management using guidance and mentoring processes. For primary school teachers* [Doctoral dissertation]. Naresuan University. (in Thai)
- Thailand Development Partners Forum Project. (2014). *Reform of the learning system and education management*. Sahamit Printing and Publishing Company Limited. (in Thai)
- Wongkitrungrueang, W. (2011). *New Future Skills: Education for the 21st Century*. Open Worlds. (in Thai)

การพัฒนาโมเดลการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนมัธยมศึกษาหลังวิกฤต

COVID-19: การวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง

Development of New Normal Learning Adaptation Model for High School Students Post-COVID-19 Crisis: A Second-Order Confirmatory Factor Analysis

ธนัชร พาสุริยันต์¹ และ สุนทรพจน์ ดำรงค์พานิช^{2*}Thanutchaphorn Pasuriyan¹ and Suntorapot Damrongpanit^{2*}

(Received: June 19, 2024; Revised: August 2, 2024; Accepted: August 5, 2024)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อพัฒนาและตรวจสอบความตรงของโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา หลังสถานการณ์แพร่ระบาดของโรค COVID-19 ในโรงเรียนภาคเหนือตอนบน สังกัดสำนักงานเขตพื้นที่การศึกษาชั้นพื้นฐาน จำนวน 1,295 คน เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือ แบบสอบถามการปรับตัวทางการเรียนรู้ จำนวน 54 ข้อ โดยมีค่าความเชื่อมั่นรายด้านระหว่าง 0.846 - 0.905 และวิเคราะห์ข้อมูลด้วยสถิติเชิงบรรยาย การวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ (correlation analysis) และการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง (second-order confirmatory factor analysis) ผลการวิจัย พบว่า ผลการพัฒนาโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา หลังสถานการณ์แพร่ระบาดของโรค COVID-19 พบว่า มี 4 องค์ประกอบ 15 ตัวบ่งชี้ ได้แก่ 1) การดูแลร่างกายในการเรียน มี 2 ตัวบ่งชี้ คือ การลดโอกาสการเกิดโรคในการเรียน และการเสริมสร้างความแข็งแรงให้กับร่างกายในการเรียน 2) ความรู้สึกนึกคิดของตนเองต่อการเรียน มี 4 ตัวบ่งชี้ คือ การควบคุมอารมณ์ของตนเองในการเรียน ยึดหยุ่นต่อสถานการณ์ที่เกิดขึ้นในการเรียน การมีสมาธิในการเรียน และการแก้ไขปัญหาทางการเรียนด้วยตนเอง 3) บทบาทหน้าที่ในการเรียน มี 5 ตัวบ่งชี้ คือ เรียนรู้เทคโนโลยีในการเรียน ความกระตือรือร้นในการเรียนการพัฒนาตนเองในการเรียน การใช้สื่อและแอปพลิเคชันช่วยเรียน และการแบ่งเวลาระหว่างการเรียนและการทำกิจกรรมอื่น 4) การมีปฏิสัมพันธ์ระหว่างการเรียนรู้ มี 4 ตัวบ่งชี้ คือ การยอมรับความแตกต่างระหว่างนักเรียนด้วยกัน การแสดงออกที่เหมาะสมของนักเรียน การสอนเพื่อนด้วยความเต็มใจ และการใช้ศักยภาพในการเรียนอย่างเต็มที่ และผลการตรวจสอบความตรงของโมเดล พบว่า โมเดลการวัดมีความสอดคล้องกลมกลืน โดยพิจารณาจากค่าไคสแควร์ (χ^2) มีค่าเท่ากับ 34.402 ที่องศาอิสระ

¹ นักศึกษาปริญญาโท สาขาวิชาประเมินผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเชียงใหม่² รองศาสตราจารย์ ภาควิชาพื้นฐานและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเชียงใหม่¹ Master's degree student, Educational Evaluation and Research Program, Faculty of Education, Chiang Mai University² Assistant Professor, Department of Educational Foundations and Development, Faculty of Education, Chiang Mai University

* Corresponding Author E-mail: suntonrapot.d@cmu.ac.th

(df) เท่ากับ 94 ระดับนัยสำคัญเท่ากับ 0.7571 ค่าวัดความกลมกลืน (CFI) เท่ากับ 1.000 ค่าดัชนี TLI เท่ากับ 1.002 ค่าดัชนีรากของกำลังสองเฉลี่ยของค่าความแตกต่างโดยประมาณ (RMSEA) เท่ากับ 0.000 และค่าดัชนีรากที่สองของค่าเฉลี่ยกำลังสองของเศษเหลือในรูปแบบคะแนนมาตรฐาน (SRMR) เท่ากับ 0.011 ค่าน้ำหนักองค์ประกอบอันดับหนึ่งมีค่าอยู่ระหว่าง 0.649 - 0.898 ค่าน้ำหนักองค์ประกอบอันดับสองอยู่ระหว่าง 0.752-0.888 และมีค่าความแปรปรวนของการปรับตัวทางการเรียนรู้ของนักเรียน เท่ากับ 0.152

คำสำคัญ วิธีใหม่ โมเดลการวัด การปรับตัวทางการเรียนรู้ของนักเรียน

สถานการณ์แพร่ระบาดของโรค COVID-19 การวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง

Abstract

This research aimed to develop and validate a measurement model of students' learning adjustment in the new normal after the COVID-19 pandemic situation among 1,295 high school students in the upper northern region of Thailand under the Office of Basic Education Commission. The research instrument was a 54-item adaptive learning questionnaire with reliability values ranging from 0.846 to 0.905. Data analysis encompassed descriptive statistics, correlation analysis, and second-order confirmatory factor analysis. Findings revealed that the developed model for measuring students' adaptive learning in the new normal following the outbreak of the COVID-19 pandemic comprised 4 components with 15 indicators: 1) Health Care in Learning, consisting of two indicators: reducing disease risk and enhancing physical resilience. 2) Self-Sentiment in Learning, consisting of four indicators: self-regulation of emotions, adaptability to learning situations, practicing mindfulness, and self-solving learning issues. 3) Roles of Learning, consisting of five indicators: learning technology, enthusiasm, self-development, utilizing media and applications, and time allocation between learning and other activities. 4) Interaction in Learning, consisting of four indicators: accepting differences among students, appropriate student behavior, teaching peers willingly, and utilizing full potential in learning. The measurement model demonstrates good fit, as evidenced by the following indices: The chi-square (χ^2) value is 34.402 with 94 degrees of freedom (df), yielding a p-value of 0.7571. The Comparative Fit Index (CFI) is 1.000, and the Tucker-Lewis Index (TLI) is 1.002. The Root Mean Square Error of Approximation (RMSEA) is 0.000, and the Standardized Root Mean Square Residual (SRMR) is 0.011. The model's validity was confirmed, with first-order loading values ranging from 0.649 to 0.898, second-order loading values ranging from 0.752 to 0.888, and a variance of student adaptive learning of 0.152.

Keywords: new normal, measurement model, students' adaptive learning, the post-COVID-19 era, second-order confirmatory factor analysis

บทนำ

สถานการณ์การแพร่ระบาดของโรค COVID-19 ส่งผลกระทบอย่างรุนแรงต่อระบบการศึกษาทั่วโลก นำไปสู่การปิดสถานศึกษาเพื่อควบคุมการแพร่เชื้อ (Caasi & Pentang, 2022) ส่งผลให้เกิดการเปลี่ยนแปลงรูปแบบการจัดการเรียนรู้อย่างมีนัยสำคัญ จากการเรียนรู้แบบผสมผสานระหว่างห้องเรียนและออนไลน์สู่การเรียนออนไลน์แบบเต็มรูปแบบ แม้สถานการณ์จะคลี่คลาย แต่ผลกระทบต่อการเรียนรู้ของนักเรียนยังคงอยู่ โดยเฉพาะในด้านการปรับตัวเข้าสู่วิถีใหม่ การเข้าใจกระบวนการปรับตัวในบริบทนี้จึงมีความสำคัญ เนื่องจากจะช่วยในการพัฒนาแนวทางสนับสนุนที่มีประสิทธิภาพในระยะยาว ด้วยเหตุนี้ การศึกษาการปรับตัวทางการเรียน (learning adaptability) หลังสถานการณ์การแพร่ระบาดของโรค COVID-19 จึงมีความจำเป็นอย่างยิ่ง ทั้งนี้ นักพฤติกรรมศาสตร์และสังคมศาสตร์ได้ให้ความสนใจต่อศึกษาการปรับตัวของผู้เรียนมาอย่างต่อเนื่อง โดยมุ่งเน้นให้มนุษย์เป็นศูนย์กลางของการเรียนรู้จากประสบการณ์ (Rogers, 1967) นอกจากนี้แนวคิดเกี่ยวกับการปรับตัวทางการเรียนมีองค์ประกอบหลากหลายมิติ ทั้งการปรับตัวต่อสภาพแวดล้อม สังคม บทบาทหน้าที่ (Andrews & Roy, 1991) การมีปฏิสัมพันธ์กับผู้อื่น ตลอดจนคุณลักษณะภายในตัวบุคคล โดยนักเรียนที่มีความสามารถในการปรับตัวได้ดีมีแนวโน้มที่จะประสบความสำเร็จและมีพัฒนาการทางการเรียนรู้ที่ดีขึ้นอย่างต่อเนื่อง (Gopal et al., 2021)

คุณลักษณะการปรับตัวที่ดีของนักเรียน มีองค์ประกอบหลักที่สำคัญหลายประการ เช่น ภาวะต่อรู้อารมณ์ในการเรียนรู้และแสวงหาความรู้ใหม่ ๆ ความเชื่อมั่นในตนเองและศักยภาพของตน (Utomo et al., 2021) การบริหารจัดการเวลาได้อย่างมีประสิทธิภาพ การตั้งเป้าหมายอย่างชัดเจนและมีแรงบันดาลใจ ทักษะการทำงานเป็นทีมและประสานงานกับผู้อื่น (Basaure et al., 2021; Elfirdoussi et al., 2020) ความสามารถในการแก้ไขปัญหาและรับมือกับความเครียดได้อย่างเหมาะสม (Ayres, 2021; Lee, 2021; Longis et al., 2021) ตลอดจนความมุ่งมั่นและตั้งใจอย่างแน่วแน่ที่จะบรรลุเป้าหมายที่วางไว้ (Burhan et al., 2021; Lassoued et al., 2020) การมีคุณลักษณะเหล่านี้จะช่วยให้นักเรียนสามารถปรับตัวและเผชิญกับความท้าทายต่าง ๆ ในบริบทการศึกษาได้อย่างราบรื่น อันจะนำไปสู่ความสำเร็จในการศึกษาในที่สุด จึงมีงานวิจัยจำนวนมากที่มุ่งเน้นสำรวจและทำความเข้าใจพฤติกรรมปรับตัวของนักเรียนในมิติต่าง ๆ เช่น การใช้เทคโนโลยีและแพลตฟอร์มออนไลน์ในการเรียน กระบวนการสอนออนไลน์ ปฏิสัมพันธ์ในการทำงานกลุ่ม การวางแผนการจัดการเวลา (Asghari et al., 2022; Zobeida et al., 2022) รวมถึงทักษะการจัดการความเครียดที่ส่งผลต่อสุขภาพจิต (Gaurav et al., 2021; Hammami et al., 2022; Singh et al., 2021) ข้อมูลเชิงประจักษ์เกี่ยวกับพฤติกรรมปรับตัวเหล่านี้มีความสำคัญต่อการพัฒนาโมเดลการวัดการปรับตัวทางการเรียนของนักเรียนให้มีความสอดคล้องกับบริบทการเรียนรู้วิถีใหม่ในยุคหลังการแพร่ระบาด

อย่างไรก็ตาม จากการทบทวนวรรณกรรมและการสังเคราะห์งานวิจัยที่ผ่านมา พบว่า ยังขาดสารสนเทศเกี่ยวกับองค์ประกอบและตัวบ่งชี้ที่เหมาะสมสำหรับการวัดการปรับตัวทางการเรียนรู้ในบริบทวิถีใหม่หลังสถานการณ์การแพร่ระบาดของโรค COVID-19 ที่มีลักษณะเชิงประจักษ์อย่างแน่ชัด พฤติกรรมบ่งชี้การปรับตัวที่ได้จากการสังเคราะห์ มีความคาบเกี่ยวกันระหว่างพฤติกรรมก่อนและระหว่างสถานการณ์การแพร่ระบาด ทำให้มีความจำเป็นต้องพัฒนาโมเดลการวัดที่สามารถประเมินระดับการปรับตัวทางการเรียนรู้

ของนักเรียนเรียนให้สอดคล้องกับกระบวนการเรียนรู้วิถีใหม่ได้อย่างเหมาะสม ด้วยเหตุนี้ งานวิจัยนี้จึงมุ่งเน้นพัฒนาโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียน หลังสถานการณ์การแพร่ระบาด COVID-19 เพื่อสะท้อนถึงคุณลักษณะการปรับตัวของนักเรียนได้อย่างครอบคลุม เป็นประโยชน์ต่อครูและสถานศึกษาในการประเมินศักยภาพของนักเรียน และการจัดกระบวนการเรียนรู้ที่ส่งเสริมให้นักเรียนเกิดปรับตัวทางการเรียนรู้ในบริบทวิถีใหม่ มุ่งสู่ความสำเร็จทางการศึกษาในอนาคตได้ นอกจากนี้ได้ทราบองค์ความรู้ที่เป็นประโยชน์ต่อนักเรียนในการเตรียมความพร้อม การตอบสนองต่อสถานการณ์ การบริหารจัดการตนเอง การสร้างความตระหนักรู้ในการดูแลสุขภาพจิตให้ดียิ่งขึ้น และช่วยให้เข้าใจทักษะที่จำเป็นต่อการพัฒนาตนเอง เช่น การเรียนรู้ด้วยตนเอง การทำงานเป็นทีม และการใช้เทคโนโลยีสารสนเทศและการสื่อสาร ซึ่งจะนำไปสู่การยกระดับคุณภาพการศึกษาและการพัฒนาผู้เรียนให้มีศักยภาพพร้อมรับมือกับความท้าทายใหม่ ๆ ในสังคมโลกอย่างยั่งยืนต่อไป

ความมุ่งหมายของการวิจัย

เพื่อพัฒนาและตรวจสอบโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา หลังสถานการณ์แพร่ระบาดของโรค COVID-19

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากร เป็นนักเรียนระดับมัธยมศึกษา ในโรงเรียนภาคเหนือตอนบน สังกัดสำนักงานเขตพื้นที่การศึกษาขั้นพื้นฐาน มุ่งเน้นที่นักเรียนระดับมัธยมศึกษาเนื่องจากอยู่ในช่วงเปลี่ยนผ่านระหว่างวัยเด็กและวัยรุ่น ซึ่งในช่วงนี้ทักษะการเรียนรู้เชิงปรับตัวมีความสำคัญอย่างยิ่งต่อความสำเร็จทางการเรียนและโอกาสในอนาคต

ในการกำหนดขนาดของกลุ่มตัวอย่างสำหรับการวิเคราะห์การวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง ผู้วิจัยได้พิจารณาตามหลักการของ Rule of thumb หรือกฎแห่งความชัดเจน ซึ่งแนะนำว่าควรมีตัวอย่างอย่างน้อย 10 คนต่อ 1 พารามิเตอร์ที่ต้องการประมาณค่า (Bentler & Chou, 1987; Brown, 2015; Kline, 2015) จึงได้ดำเนินการสุ่มแบบหลายขั้นตอน (multi-stage random sampling) (Singh, 2014) ดังนี้

1) ระบุจังหวัดในภาคเหนือตอนบน ได้แก่ เชียงใหม่ เชียงราย ลำปาง ลำพูน แม่ฮ่องสอน พะเยา แพร่ และน่าน

2) สุ่มเลือกโรงเรียนร้อยละ 25 ตามขนาดโรงเรียน (เล็ก กลาง ใหญ่ ใหญ่พิเศษ) ของแต่ละจังหวัด โดยใช้วิธีการสุ่มอย่างง่าย (simple random sampling)

3) สุ่มแบบแบ่งชั้น (stratified random sampling) โดยแบ่งนักเรียนในโรงเรียนที่ถูกเลือกเป็นชั้นภูมิ (strata) ตามระดับชั้น (มัธยมศึกษาปีที่ 1 - 6) คำนวณสัดส่วนนักเรียนในแต่ละระดับชั้น และกำหนดจำนวนตัวอย่างที่ต้องการจากแต่ละระดับชั้นตามสัดส่วน จากนั้นสุ่มเลือกนักเรียนจากแต่ละระดับชั้นตามจำนวนที่กำหนดโดยวิธีการสุ่มอย่างง่าย (simple random sampling)

4) ดำเนินการเก็บรวบรวมข้อมูล ซึ่งได้จำนวนนักเรียน 1,295 คน แบ่งเป็นนักเรียนชาย 423 คน และนักเรียนหญิง 872 คน ซึ่งเป็นขนาดที่เพียงพอและเหมาะสมสำหรับการวิเคราะห์ข้อมูล

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบสอบถามการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียน หลังสถานการณ์การแพร่ระบาดของโรค COVID-19 มีข้อคำถามจำนวนทั้งหมด 54 ข้อ มีลักษณะการตอบแบบ มาตราส่วนประมาณค่า (rating scale) 5 ระดับ (1 = เห็นด้วย/ปฏิบัติตนตรงกับรายการนั้นในระดับน้อยที่สุด, 5 = เห็นด้วย/ปฏิบัติตนตรงกับรายการนั้นในระดับมากที่สุด) ขอบเขตของการวัดการปรับตัวทางการเรียนรู้ของ นักเรียนที่ศึกษา มีดังนี้ ค่าดัชนีความเที่ยงตรงของแบบสอบถาม (Index of item Objective Congruence: IOC) จากผู้เชี่ยวชาญ 5 ท่าน ค่าอำนาจจำแนก (discrimination) ด้วยวิธีการพิจารณาความสัมพันธ์ระหว่าง คะแนนรายข้อและคะแนนรวม (item-total correlation) การวิเคราะห์ค่าความเชื่อมั่นในการวัดตัวบ่งชี้ โดยใช้ค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's alpha coefficient)

การดูแลร่างกายในการเรียน (Health Care in Learning; SHCL) หมายถึง การป้องกันภาวะ ความเจ็บป่วยจากการแพร่ระบาดของเชื้อโรค การลดโอกาสการเกิดโรคในการเรียน การเสริมสร้างความ แข็งแรงให้กับร่างกาย มี 2 ตัวบ่งชี้ ได้แก่ การลดโอกาสการเกิดโรคในการเรียน และการเสริมสร้างความ แข็งแรงให้กับร่างกายในการเรียน มีทั้งหมด 9 ข้อ โดยการทดสอบ t มีค่าอยู่ระหว่าง 6.900 - 10.890 ข้อคำถามมีค่าดัชนีความสอดคล้อง (IOC) ระหว่าง 0.8 - 1.0 ค่าอำนาจจำแนกด้วยวิธีการพิจารณาความสัมพันธ์ ระหว่างคะแนนรายข้อและคะแนนรวม มีค่าอยู่ระหว่าง 0.372 - 0.625 และวิเคราะห์ค่าความเชื่อมั่นในการวัด ตัวบ่งชี้ โดยใช้ค่าสัมประสิทธิ์แอลฟาของครอนบาค ของทั้งฉบับ มีค่าอยู่ระหว่าง 0.656 - 0.832 ตัวอย่าง ข้อคำถาม ได้แก่ “ฉันล้างมือด้วยสบู่หรือเจลแอลกอฮอล์ หลังสัมผัสจุดเสี่ยง เช่น ราวบันได โต๊ะ เก้าอี้ ในโรงเรียน” “ฉันทำความสะอาดอุปกรณ์การเรียนทุกวันเพื่อออกไปเรียนที่โรงเรียน”

ความรู้สึกนึกคิดของตนเองต่อการเรียน (Self-Sentiment in Learning; SSSL) หมายถึง การรู้จักควบคุมอารมณ์และความต้องการของตนเองมีความยืดหยุ่นต่อสถานการณ์ที่เกิดขึ้นในการเรียน การยอมรับตนเองตามความเป็นจริง มีความสามารถในการใช้เทคโนโลยี สามารถทำงานที่ได้รับมอบหมาย มี 4 ตัวบ่งชี้ ได้แก่ การควบคุมอารมณ์ของตนเองในการเรียน ยืดหยุ่นต่อสถานการณ์ที่เกิดขึ้นในการเรียน การมีสมาธิในการเรียน และการแก้ไขปัญหาทางการเรียนด้วยตนเอง โดยการทดสอบ t มีค่าอยู่ระหว่าง 4.380 - 10.541 ข้อคำถามมีค่าดัชนีความสอดคล้อง (IOC) ระหว่าง 0.6 - 1.0 ค่าอำนาจจำแนกด้วยวิธีการพิจารณา ความสัมพันธ์ระหว่างคะแนนรายข้อและคะแนนรวม มีค่าอยู่ระหว่าง 0.354 - 0.695 และวิเคราะห์ค่าความ เชื่อมั่นในการวัดตัวบ่งชี้ โดยใช้ค่าสัมประสิทธิ์แอลฟาของครอนบาค ของทั้งฉบับ มีค่าอยู่ระหว่าง 0.592 - 0.843 ตัวอย่างข้อคำถาม ได้แก่ “เมื่อฉันเครียดจากการเรียนออนไลน์ ฉันจะทำกิจกรรมผ่อนคลายหรือพูดคุย กับผู้อื่นให้สบายใจก่อนที่จะทำงาน”, “ฉันเชื่อว่าตนเองสามารถบริหารจัดการตนเองในสถานการณ์ ที่เปลี่ยนแปลงได้ดี”

บทบาทหน้าที่ในการเรียน (Roles of Learning; SROL) หมายถึง ความอยากรู้อยากเห็น ที่จะสนุกกับการเรียนรู้ในรูปแบบแบบการสอนที่มีการปรับเปลี่ยนใหม่ ความกระตือรือร้น ตื่นตัว ไม่หยุดนิ่ง

ต่อการเรียนรู้มีการเตรียมความพร้อมในการเรียนการมีสมาธิในการเรียน ตั้งใจเรียน รู้จักแบ่งเวลาในการเรียน และการทำงานกิจกรรมอื่น มี 5 ตัวบ่งชี้ ได้แก่ เรียนรู้เทคโนโลยีในการเรียน ความกระตือรือร้นในการเรียน การพัฒนาตนเองในการเรียน การใช้สื่อและแอปพลิเคชันช่วยเรียน และการแบ่งเวลาระหว่างการเรียนและการทำงานกิจกรรมอื่น โดยการทดสอบ t มีค่าอยู่ระหว่าง 7.510 - 11.820 ข้อคำถามมีค่าดัชนีความสอดคล้อง (IOC) ระหว่าง 0.6 - 1.0 ค่าอำนาจจำแนกด้วยวิธีการพิจารณาความสัมพันธ์ระหว่างคะแนนรายข้อและคะแนนรวม มีค่าอยู่ระหว่าง 0.550 - 0.758 และวิเคราะห์ค่าความเชื่อมั่นในการวัดตัวบ่งชี้ โดยใช้ค่าสัมประสิทธิ์แอลฟาของครอนบาค ของทั้งฉบับ มีค่าอยู่ระหว่าง 0.762 - 0.827 ตัวอย่างข้อคำถาม ได้แก่ “ฉันชอบรับฟังข้อมูลข่าวสารด้านเทคโนโลยีการเรียน การเปลี่ยนแปลงของเหตุการณ์ในปัจจุบันตลอดเวลา”, “ฉันสามารถใช้เทคโนโลยีเพื่อการเรียนรู้แทนการเรียนรู้ในห้องได้โดยไม่เกิดปัญหาทางการเรียน”

การมีปฏิสัมพันธ์ระหว่างการเรียน (Interaction in Learning; SINL) หมายถึง การยอมรับความแตกต่างระหว่างเพื่อนนักเรียนและครูที่มีพื้นฐานที่แตกต่างกัน การแสดงออกที่เหมาะสม ความเต็มใจในการเรียน การใช้ศักยภาพในการเรียนอย่างเต็มที่ในการสร้างองค์ความรู้ที่มีประสิทธิภาพต่อนักเรียน มี 4 ตัวบ่งชี้ ได้แก่ การยอมรับความแตกต่างระหว่างนักเรียนด้วยกัน การแสดงออกที่เหมาะสมของนักเรียน การสอนเพื่อนด้วยความเต็มใจและการใช้ศักยภาพในการเรียนอย่างเต็มที่ โดยการทดสอบ t มีค่าอยู่ระหว่าง 5.980 - 9.920 ข้อคำถามมีค่าดัชนีความสอดคล้อง (IOC) ระหว่าง 0.6 - 1.0 ค่าอำนาจจำแนกด้วยวิธีการพิจารณาความสัมพันธ์ระหว่างคะแนนรายข้อและคะแนนรวม มีค่าอยู่ระหว่าง 0.504 - 0.674 และวิเคราะห์ค่าความเชื่อมั่นในการวัดตัวบ่งชี้ โดยใช้ค่าสัมประสิทธิ์แอลฟาของครอนบาค ของทั้งฉบับ มีค่าอยู่ระหว่าง 0.672 - 0.818 ตัวอย่างข้อคำถาม ได้แก่ “ฉันคิดว่าการเข้าใจความรู้ของนักเรียนเป็นเรื่องปกติไม่ว่าจะเป็นการเรียนในห้องเรียนหรือการเรียนแบบออนไลน์”, “ฉันจะแนะนำการใช้เทคโนโลยีหรือวิธีการแก้ไขปัญหาการใช้เทคโนโลยีที่ฉันรู้จักให้กับครูหรือเพื่อน”

3. การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลสำหรับงานวิจัยนี้ ผู้วิจัยได้ทำหนังสือบันทึกขอความอนุเคราะห์ในการเก็บข้อมูลไปยังโรงเรียนระดับมัธยมศึกษาจำนวน 43 แห่ง โดยกลุ่มเป้าหมายเป็นนักเรียนในระดับชั้นมัธยมศึกษาปีที่ 1-6 ระดับมัธยมศึกษาในภาคเหนือตอนบน สำนักงานเขตพื้นที่การศึกษาชั้นพื้นฐาน เนื่องจากการเก็บข้อมูลอยู่ในช่วงการแพร่ระบาดของโรค COVID-19 เพื่อความสะดวกและความปลอดภัยของผู้ตอบแบบสอบถาม การตอบแบบสอบถามจึงอยู่ในรูปแบบออนไลน์ผ่าน Google Form โดยผู้วิจัยได้จัดทำ QR-code เพื่อให้ นักเรียนสามารถเข้าถึงแบบสอบถามได้อย่างง่าย ระยะเวลาในการเก็บรวบรวมข้อมูลทั้งสิ้น 3 เดือน

ทั้งนี้ งานวิจัยครั้งนี้ได้ผ่านการรับรองจากคณะกรรมการจริยธรรมการวิจัยในคนมหาวิทยาลัยเชียงใหม่ เลขที่ CMUREC No. 65/005 โดยผู้วิจัยให้ความสำคัญกับการพิทักษ์สิทธิและรักษาความลับของกลุ่มตัวอย่างเป็นอย่างยิ่ง มีการชี้แจงรายละเอียดเกี่ยวกับสิทธิของผู้เข้าร่วมวิจัยอย่างเป็นลายลักษณ์อักษร โดยนักเรียนที่เป็นกลุ่มตัวอย่างสามารถยุติการให้ข้อมูลได้ตลอดเวลาหากรู้สึกไม่สะดวกหรือไม่สบายใจในการตอบคำถาม นอกจากนี้ ข้อมูลทั้งหมดที่ได้จากแบบสอบถามจะถูกเก็บรักษาไว้เป็นความลับ และจะทำลายทิ้งเมื่อสิ้นสุด

การวิจัย การวิเคราะห์ข้อมูลและการรายงานผลจะเป็นไปในภาพรวม โดยไม่มีการเชื่อมโยงถึงตัวบุคคล หน่วยงาน หรือองค์กรใด ๆ

ผลจากการเก็บรวบรวมข้อมูล พบว่า ผู้วิจัยได้แจกแบบสอบถามออนไลน์ไปทั้งสิ้นจำนวน 1,500 ชุด และมีนักเรียนตอบกลับมา 1,453 ชุด โดยไม่มีนักเรียนทำนโดยวิธีการให้ข้อมูลระหว่างการตอบแบบสอบถาม อย่างไรก็ตาม พบว่ามีแบบสอบถามที่กรอกข้อมูลไม่สมบูรณ์จำนวน 158 ชุด ทำให้เหลือข้อมูลที่มีความสมบูรณ์ และสามารถนำเข้าสู่การวิเคราะห์ทางสถิติได้จำนวน 1,295 ชุด คิดเป็นร้อยละ 86.33 ของจำนวนแบบสอบถามที่แจกไปทั้งหมด ซึ่งถือเป็นอัตราการตอบกลับที่อยู่ในเกณฑ์ที่ดีและเพียงพอสำหรับการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง

4. การวิเคราะห์ข้อมูล

ผลการดำเนินการตรวจสอบคุณภาพข้อมูลและข้อตกลงเบื้องต้นทางสถิติ การวิเคราะห์ด้วยสถิติเชิงบรรยาย ได้แก่ ค่าเฉลี่ย (M) ส่วนเบี่ยงเบนมาตรฐาน (SD) การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's product moment correlation) ของตัวบ่งชี้และองค์ประกอบการวิเคราะห์ค่าดัชนีไกเซอร์-ไมเยอร์-ออลกิน (Kaiser-Meyer-Olkin Measure of Sampling Adequacy: KMO) โดยค่าควรอยู่ระหว่าง 0-1 และค่าสถิติของบาร์ตเลตต์ (Bartlett's test of Sphericity) เพื่อทดสอบตัวบ่งชี้ทั้งหมดเป็นเมทริกซ์เอกลักษณะ ผลการวิเคราะห์ควรมีนัยสำคัญทางสถิติที่ .05 (Bartlett, 1954)

ในการตรวจสอบความตรงของโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา โดยใช้การวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง พิจารณาดัชนีความสอดคล้องกลมกลืนของโมเดลกับข้อมูลเชิงประจักษ์ ประกอบไปด้วยค่าไคสแควร์ (χ^2) ผลการนำเสนอค่าไคสแควร์ไม่มีนัยสำคัญทางสถิติที่ระดับ .05 ค่าดัชนีรากของกำลังสองเฉลี่ยของค่าความแตกต่างโดยประมาณ (RMSEA) <0.05 ค่าดัชนีรากที่สองของค่าเฉลี่ยกำลังสองของเศษเหลือในรูปแบบคะแนนมาตรฐาน (SRMR) <0.05 ค่าวัดความกลมกลืน (CFI) >0.95 และค่าดัชนี Tucker-Lewis (TLI) >0.95 (Byrne, 2010; Kline, 2015)

ผลการวิจัย

1. การตรวจสอบคุณภาพข้อมูลและข้อตกลงเบื้องต้นทางสถิติ ได้แก่ การวิเคราะห์สถิติเชิงบรรยาย พบว่า ตัวบ่งชี้มีค่าเฉลี่ย 3.617-4.133 ตัวบ่งชี้เฉลี่ยสูงสุด คือ การยอมรับความแตกต่างระหว่างนักเรียนด้วยกัน (SI1) (M=4.133) ส่วนตัวบ่งชี้ที่มีค่าน้อยที่สุด คือ การแบ่งเวลาระหว่างการเรียนรู้และการทำกิจกรรมอื่น (SR5) (M=3.617) การกระจายของข้อมูล พบว่า มีค่าเบี่ยงเบนมาตรฐาน 0.393-0.588 ตัวบ่งชี้ที่มีส่วนเบี่ยงเบนมาตรฐานสูงสุด คือ การแบ่งเวลาระหว่างการเรียนรู้และการทำกิจกรรมอื่น (SR5) (S.D.=0.588) ส่วนตัวบ่งชี้ที่มีการกระจายข้อมูลน้อยที่สุด คือ การเสริมสร้างความแข็งแรงให้กับร่างกายในการเรียน (SH2) (SD= 0.393) สำหรับการแจกแจงข้อมูล พบว่า การพิจารณาลักษณะการแจกแจงข้อมูลในแต่ละตัวบ่งชี้จากความเบ้ (Skewness: Sk) มีค่าเป็นลบทุกตัว โดยมีค่าอยู่ระหว่าง -0.658 ถึง -0.197 การแจกแจงจึงมีลักษณะการแจกแจงปกติอย่างมีนัยสำคัญทางสถิติ เนื่องจากอยู่ในช่วง -2 ถึง 2 และความโด่ง (Kurtosis: Kur) มีค่าระหว่าง -0.334 ถึง 0.469 แสดงว่าตัวบ่งชี้มีลักษณะการแจกแจงเข้าใกล้โค้งปกติ

2. ผลการพัฒนาโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา

และผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร จากเอกสารและงานวิจัยที่เกี่ยวข้อง นำเสนอตัวบ่งชี้ต่อผู้เชี่ยวชาญเพื่อพิจารณาความเหมาะสม พบว่า โมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา ประกอบด้วย 4 องค์ประกอบ 15 ตัวบ่งชี้ ได้แก่ องค์ประกอบที่ 1 การดูแลร่างกายในการเรียน (SHCL) (2 ตัวบ่งชี้) องค์ประกอบที่ 2 ความรู้สึกนึกคิดของตนเองต่อการเรียน (SSSL) (4 ตัวบ่งชี้) องค์ประกอบที่ 3 บทบาทหน้าที่ในการเรียน (SROL) (5 ตัวบ่งชี้) องค์ประกอบที่ 4 การมีปฏิสัมพันธ์ระหว่างการเรียน (SINL) (4 ตัวบ่งชี้) โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน พบว่า ทุกตัวบ่งชี้มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 มีลักษณะความสัมพันธ์ทางบวก ขนาดของความสัมพันธ์มีค่าตั้งแต่ 0.297-0.745 ซึ่งตัวบ่งชี้การสอนเพื่อนด้วยความเต็มใจ (SI3) กับการแสดงออกที่เหมาะสมของนักเรียน (SI2) มีความสัมพันธ์มากที่สุด และการยอมรับความแตกต่างระหว่างนักเรียนด้วยกัน (SI1) กับการยอมรับความแตกต่างระหว่างนักเรียนด้วยกัน (SI1) มีความสัมพันธ์น้อยที่สุด

ทั้งนี้การหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวบ่งชี้ทั้งหมดมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 105 คู่ โดยตัวบ่งชี้มีความสัมพันธ์ทางบวกทั้งหมด มีค่าดัชนีไกเซอร์-ไมเยอร์-ออลคิน (Kaiser-Meyer-Olkin Measure of Sampling Adequacy: KMO) เท่ากับ 0.943 และค่าสถิติของบาร์ตเลตต์ (Bartlett's test of Sphericity) เท่ากับ 14083 ที่ระดับนัยสำคัญ 0.000 สรุปว่าตัวบ่งชี้มีความสัมพันธ์กันและมีความเหมาะสมที่จะนำไปวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สอง ดังตาราง 1

ตาราง 1 ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันของตัวบ่งชี้ ในโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา หลังสถานการณ์แพร่ระบาดของโรค COVID-19 (n=1,295)

ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันของตัวบ่งชี้															
	SH1	SH2	SL1	SL2	SL3	SL4	SR1	SR2	SR3	SR4	SR5	SI1	SI2	SI3	SI4
SH1	1.000														
SH2	.663**	1.000													
SL1	.443**	.529**	1.000												
SL2	.464**	.472**	.604**	1.000											
SL3	.475**	.457**	.508**	.690**	1.000										
SL4	.425**	.457**	.519**	.715**	.745**	1.000									
SR1	.444**	.459**	.478**	.552**	.562**	.606**	1.000								
SR2	.469**	.488**	.505**	.572**	.589**	.622**	.694**	1.000							
SR3	.511**	.479**	.499**	.589**	.671**	.628**	.645**	.650**	1.000						
SR4	.401**	.484**	.518**	.549**	.564**	.566**	.669**	.718**	.673**	1.000					
SR5	.492**	.457**	.467**	.543**	.691**	.577**	.586**	.608**	.742**	.590**	1.000				
SI1	.297**	.400**	.411**	.418**	.349**	.411**	.427**	.504**	.366**	.516**	.317**	1.000			
SI2	.417**	.478**	.501**	.490**	.450**	.481**	.512**	.583**	.472**	.622**	.423**	.722**	1.000		
SI3	.398**	.445**	.492**	.469**	.476**	.470**	.510**	.532**	.534**	.579**	.471**	.568**	.745**	1.000	
SI4	.387**	.444**	.463**	.522**	.505**	.548**	.578**	.615**	.527**	.626**	.496**	.589**	.713**	.668**	1.000
Mean	3.960	4.110	3.850	3.770	3.670	3.760	3.800	3.930	3.630	3.950	3.620	4.130	4.090	3.930	3.970
SD	.643	.627	.696	.683	.747	.750	.746	.683	.746	.670	.767	.759	.636	.749	.708

Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) = 0.943, Bartlett's Test of Sphericity Approx. Chi-Square = 14083,

df = 105, p-value = 0.000

** p-value < .01

3. ผลการตรวจสอบความตรงของโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา พบว่า ก่อนปรับมีค่าดัชนีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ เท่ากับ $\chi^2 = 696.522$, $df = 49$, $p\text{-value} = 0.000$, $CFI = 0.938$, $TLI = 0.925$, $RMSEA = 0.074$, $SRMR = 0.045$ แต่หลังจากการปรับโมเดล โดยได้ค่าดัชนีในการพิจารณาความสอดคล้องกลมกลืน เท่ากับ $\chi^2 = 34.402$, $df = 94$, $p\text{-value} = 0.757$, $CFI = 1.000$, $TLI = 1.002$, $RMSEA = 0.000$, $SRMR = 0.011$ นั่นคือยอมรับสมมติฐานศูนย์ว่าเมตริกซ์ความสัมพันธ์ของโมเดลสมมติฐานไม่แตกต่างจากเมตริกซ์ของโมเดลเชิงประจักษ์ แสดงว่า โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์

หลังจากการตรวจสอบโมเดลตามสมมติฐานมีความสอดคล้องกับข้อมูลเชิงประจักษ์ จึงได้พิจารณาค่าพารามิเตอร์ค่าน้ำหนักองค์ประกอบของตัวบ่งชี้ได้ในโมเดล พบว่า ผลการประมาณค่าค่าน้ำหนักองค์ประกอบลำดับขั้นที่ 1 ค่าน้ำหนักองค์ประกอบของตัวบ่งชี้ในรูปคะแนนมาตรฐานอยู่ระหว่าง 0.649-0.898 และมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกตัว โดยตัวแปรที่มีค่าน้ำหนักองค์ประกอบมาตรฐานมากที่สุด คือ การใช้ศักยภาพในการเรียนอย่างเต็มที่ (SI4) ($\beta=0.898$) รองลงมา คือ การแก้ไขปัญหาทางการเรียนด้วยตนเอง (SL4) ($\beta=0.881$) และการมีสมาธิในการเรียน (SL3) ($\beta=0.842$) ตามลำดับ ตัวแปรที่มีค่าน้ำหนักองค์ประกอบมาตรฐานน้อยที่สุด คือ การยอมรับความแตกต่างระหว่างนักเรียนด้วยกัน (SI1) ($\beta=0.649$) โดยการวัดแต่ละตัวแปรมีค่าความเชื่อมั่นขั้นต่ำอยู่ระหว่าง 0.422 (SI1) - 0.776 (SL4)

สำหรับผลการประมาณค่าค่าน้ำหนักองค์ประกอบลำดับขั้นที่ 2 พบว่าทุกองค์ประกอบมีค่าน้ำหนักองค์ประกอบในรูปคะแนนมาตรฐานในระดับสูง ได้แก่ องค์ประกอบความรู้สึกรู้สึกนึกคิดของตนเองต่อการเรียน (SSSL) ($\beta=0.888$) รองลงมา คือ บทบาทหน้าที่ในการเรียน (SROL) ($\beta=0.871$) การมีปฏิสัมพันธ์ระหว่างการเรียน (SINL) ($\beta=0.795$) และการดูแลร่างกายในการเรียน (SHCL) ($\beta=0.752$) ตามลำดับ โดยการวัดแต่ละตัวแปรมีค่าความเชื่อมั่นขั้นต่ำอยู่ระหว่าง 0.566 (SHCL) - 0.943 (SROL) ดังรายละเอียดตาราง 2 และภาพ 1

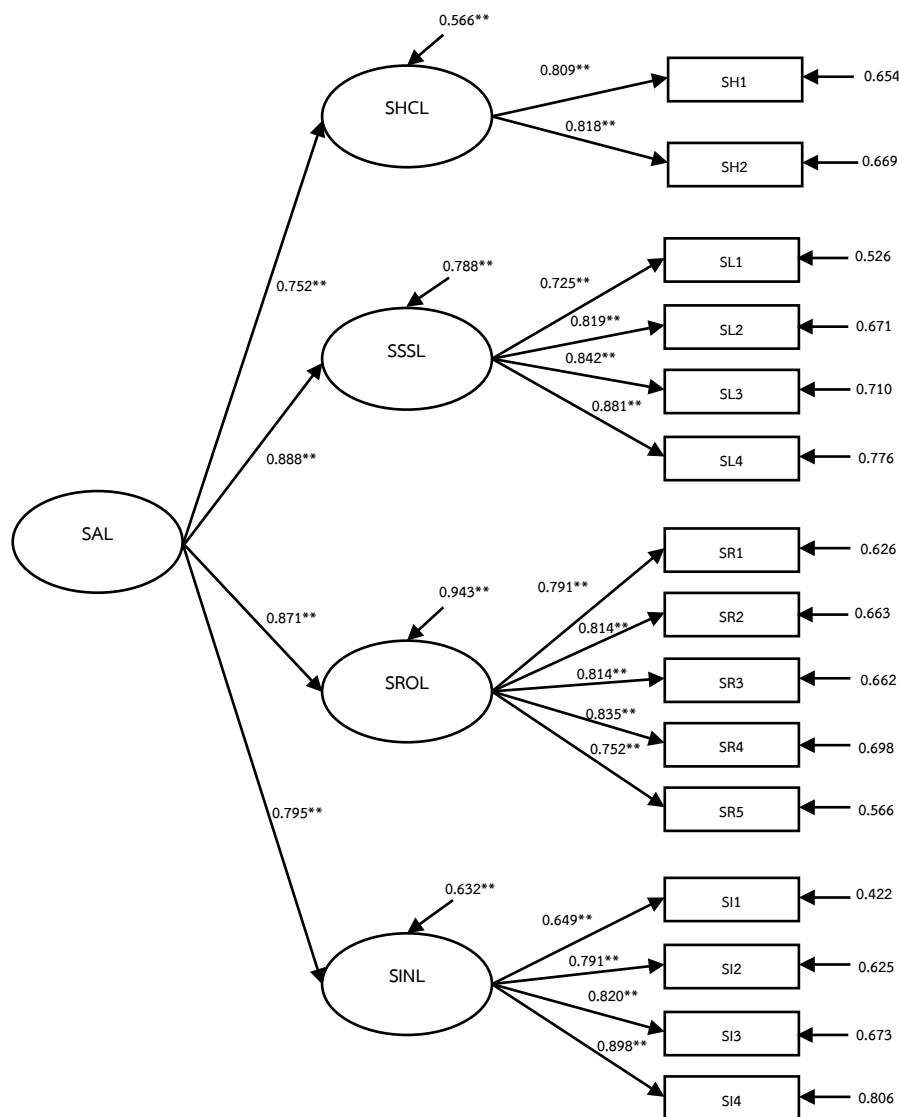
ตาราง 2 ผลการประมาณค่าพารามิเตอร์ในโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา หลังสถานการณ์แพร่ระบาดของโรค COVID-19 (n=1,295)

องค์ประกอบ	ตัวบ่งชี้	b	β	S.E.	Z-value	p-value	R ²
ลำดับขั้นที่ 1 (First Order)							
1. การดูแลร่างกายในการเรียน (SHCL)	SH1	1.000	0.809	0.020	41.084	0.000	0.654
	SH2	0.989	0.818	0.018	44.950	0.000	0.669
2. ความรู้สึกนึกคิดของตนเองต่อการเรียน (SSSL)	SL1	1.000	0.725	0.020	36.313	0.000	0.526
	SL2	1.110	0.819	0.013	65.460	0.000	0.671
	SL3	1.246	0.842	0.011	77.170	0.000	0.710
	SL4	1.313	0.881	0.011	80.826	0.000	0.776
3. บทบาทหน้าที่ในการเรียน (SROL)	SR1	1.000	0.791	0.015	53.190	0.000	0.626
	SR2	0.942	0.814	0.014	59.400	0.000	0.663
	SR3	1.028	0.814	0.013	63.065	0.000	0.662
	SR4	0.948	0.835	0.016	53.306	0.000	0.698
	SR5	0.978	0.752	0.017	44.584	0.000	0.566

ตาราง 2 (ต่อ)

องค์ประกอบ	ตัวบ่งชี้	b	β	S.E.	Z-value	p-value	R ²
4. การมี	SI1	1.000	0.649	0.023	27.801	0.000	0.422
ปฏิสัมพันธ์	SI2	1.021	0.791	0.017	45.306	0.000	0.625
ระหว่างการเรียนรู้	SI3	1.251	0.820	0.025	33.232	0.000	0.673
(SINL)	SI4	1.291	0.898	0.017	54.319	0.000	0.806
ลำดับขั้นที่ 2 (Second Order)							
การปรับตัว	SHCL	1.000	0.752	0.023	32.935	0.000	0.566
ทางการเรียนรู้	SSSL	1.147	0.888	0.014	63.647	0.000	0.788
ของนักเรียน	SROL	1.468	0.871	0.013	72.412	0.000	0.943
(SAL)	SINL	1.001	0.795	0.019	41.870	0.000	0.632

หมายเหตุ b = ค่าสัมประสิทธิ์, β = รูปคะแนนมาตรฐาน, S.E. = ความคลาดเคลื่อนมาตรฐาน, Z-value = ค่าสถิติทดสอบมาตรฐาน, p-value = ระดับนัยสำคัญทางสถิติ, R² = ค่าโคสแควร์



ภาพ 1 โมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา

หลังสถานการณ์การแพร่ระบาดของโรค COVID-19

อภิปรายผล

จากผลการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สองของโมเดลการวัดการปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา หลังสถานการณ์การแพร่ระบาดของโรค COVID-19 ในโรงเรียนภาคเหนือตอนบน สังกัดสำนักงานเขตพื้นที่การศึกษาขั้นพื้นฐาน พบว่า โมเดลประกอบด้วย 4 องค์ประกอบหลัก ได้แก่ ความสามารถในการดูแลร่างกาย ความรู้สึกนึกคิดของตนเองต่อการเรียน บทบาทหน้าที่ในการเรียน และการมีปฏิสัมพันธ์ระหว่างการเรียนรู้ โดยแต่ละองค์ประกอบประกอบด้วยตัวบ่งชี้พฤติกรรม รวมทั้งสิ้น 15 ตัวบ่งชี้ ผลการวิเคราะห์แสดงให้เห็นว่าค่าสถิติจากการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สองมีนัยสำคัญทางสถิติทุกองค์ประกอบและทุกตัวบ่งชี้ นอกจากนี้โมเดลการวัดที่พัฒนาขึ้นมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ในระดับที่ดี เช่นเดียวกับผลการศึกษาของ Kaplan (2023) ความสอดคล้องระหว่างโมเดลการวัดและข้อมูลเชิงประจักษ์นี้แสดงให้เห็นถึงความเที่ยงตรงของโมเดลในการวัดการปรับตัวทางการเรียนรู้ของนักเรียนระดับมัธยมศึกษาในบริบทหลังสถานการณ์การแพร่ระบาดของโรค COVID-19 โดยครอบคลุมมิติที่สำคัญทั้งด้านร่างกาย จิตใจ หน้าที่ และปฏิสัมพันธ์ทางสังคม สะท้อนถึงความซับซ้อนของกระบวนการปรับตัวในสถานการณ์ที่เปลี่ยนแปลงไป สามารถนำไปใช้ในการประเมินการปรับตัวทางการเรียนรู้วิถีใหม่ of นักเรียนระดับมัธยมศึกษาได้อย่างมีประสิทธิภาพ นอกจากนี้ผู้วิจัยยังพบประเด็นที่น่าสนใจจากการวิจัย 2 ประเด็น ดังนี้

1. ผลการวิเคราะห์ค่าน้ำหนักองค์ประกอบ (Factor Loading) อันดับแรกของตัวบ่งชี้พฤติกรรมต่าง ๆ พบว่า ทุกตัวบ่งชี้มีความแปรปรวนร่วมกับองค์ประกอบหลักในระดับสูง โดยตัวบ่งชี้ที่มีค่าน้ำหนักสูงที่สุด ได้แก่ การใช้ศักยภาพของตนเองในการเรียนอย่างเต็มที่ และการคิดวิเคราะห์แก้ไขปัญหาด้วยตนเอง ซึ่งสอดคล้องกับลักษณะของการเรียนรู้แบบออนไลน์ในช่วงการแพร่ระบาดของโรค COVID-19 ที่นักเรียนจำเป็นต้องมีความรับผิดชอบและความสามารถในการจัดการปัญหาด้วยตนเองมากขึ้น (Asghari et al., 2022; Basaure et al., 2021; Elfirdoussi et al., 2020) โดยเฉพาะอย่างยิ่ง พบว่า นักเรียนที่สามารถแก้ไขปัญหาและเรียนรู้ได้ด้วยตนเองจะมีความสามารถในการปรับตัวและผลสัมฤทธิ์ทางการเรียนที่สูงกว่านักเรียนที่พึ่งพาผู้อื่น ในทางกลับกัน ตัวบ่งชี้ที่มีค่าน้ำหนักน้อยที่สุด คือ การยอมรับความแตกต่างระหว่างนักเรียน ทั้งนี้ อาจเนื่องมาจากมาตรการเว้นระยะห่างทางสังคมที่ทำให้การสื่อสารและปฏิสัมพันธ์ระหว่างนักเรียนลดลง ส่งผลให้นักเรียนไม่ค่อยมีโอกาสดูแลหรือสร้างความคุ้นเคยกับความแตกต่างหลากหลายของเพื่อนร่วมชั้นเรียนเท่าที่ควร (Ayes, 2021; English & English, 1958; Hammami et al., 2022; Lee, 2021; Zobeida et al., 2022) อย่างไรก็ตาม ทักษะการยอมรับความแตกต่างยังคงเป็นทักษะที่มีความสำคัญในการอยู่ร่วมกันในสังคมและการทำงานในอนาคต

2. ผลการวิเคราะห์ค่าน้ำหนักองค์ประกอบอันดับสอง ซึ่งให้ห้เห็นว่าองค์ประกอบด้านความรู้สึกนึกคิดของตนเองต่อการเรียนและบทบาทหน้าที่ในการเรียนมีความสำคัญมากที่สุดต่อการปรับตัวในภาพรวม เนื่องจากการเปลี่ยนแปลงของสภาพแวดล้อมและรูปแบบการเรียนรู้ในช่วงการแพร่ระบาดของโรค COVID-19 ส่งผลโดยตรงต่อสภาวะทางจิตใจ อารมณ์ความรู้สึก รวมถึงบทบาทหน้าที่ของนักเรียนค่อนข้างมาก (Asghari et al., 2022; Burhan et al., 2021; Elfirdoussi et al., 2020; Lassoued et al., 2020; Utomo et al., 2021) ซึ่งสอดคล้องกับงานวิจัยของ Yehuda (2021) ที่พบว่าสภาวะทางจิตใจ ความเชื่อมั่นในตนเอง และ

ทักษะการจัดการตนเอง เป็นปัจจัยที่สำคัญที่สุดต่อการปรับตัวและการเผชิญความเครียดของนักเรียนในยุค COVID-19 ในขณะที่องค์ประกอบด้านการมีปฏิสัมพันธ์ระหว่างการเรียนและการดูแลสุขภาพร่างกายมีความสำคัญน้อยที่สุดในบริบทนี้ ซึ่งเป็นผลมาจากการเรียนรู้แบบออนไลน์ที่มุ่งเน้นการใช้เทคโนโลยีเป็นหลัก (Asghari et al., 2022; Burhan et al., 2021; Lassoued et al., 2020; Utomo et al., 2021) ประกอบกับความยืดหยุ่นของเวลาและสถานที่เรียนที่ให้นักเรียนให้ความสำคัญกับการเรียนและพื้นที่ส่วนตัวมากขึ้น (Caasi & Pentang, 2022; Longis et al., 2021) อย่างไรก็ตาม Zander (2022) ได้ให้ข้อสังเกตว่า การมุ่งเน้นแต่การเรียนและละเลยการดูแลสุขภาพอาจส่งผลเสียต่อการเรียนรู้ในระยะยาวได้

โดยสรุป ผลการวิจัยครั้งนี้ได้ชี้ให้เห็นถึงโครงสร้างและความสัมพันธ์ขององค์ประกอบและตัวบ่งชี้การปรับตัวทางการเรียนรู้วิถีใหม่ของนักเรียนระดับมัธยมศึกษา หลังสถานการณ์การแพร่ระบาดของโรค COVID-19 สะท้อนให้เห็นถึงการเปลี่ยนแปลงของพฤติกรรมและการให้ความสำคัญในการดำรงชีวิตของนักเรียนที่ปรับเปลี่ยนไปตามบริบทของการเรียนรู้ที่เปลี่ยนแปลงไป องค์ความรู้ที่ได้จะเป็นประโยชน์ในการพัฒนากระบวนการจัดการเรียนรู้และการส่งเสริมให้นักเรียนให้สามารถปรับตัวและเรียนรู้ได้อย่างมีประสิทธิภาพ

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

ผู้วิจัยมีข้อเสนอแนะในการนำผลการวิจัยไปประยุกต์ใช้ได้ 3 ระดับ คือ

1.1 ระดับชั้นเรียน ครูสามารถใช้องค์ประกอบที่ค้นพบจากการวิจัยเป็นกรอบในการสังเกตและติดตามการปรับตัวของนักเรียน เพื่อออกแบบกิจกรรมการเรียนการสอนที่ตอบสนองความต้องการของผู้เรียนได้อย่างเหมาะสม ทั้งนี้ ควรใช้ร่วมกับเครื่องมือประเมินอื่น ๆ และการสังเกตอย่างต่อเนื่อง

1.2 ระดับโรงเรียน ผู้บริหารสถานศึกษาสามารถใช้ผลการศึกษาเป็นข้อมูลประกอบการวางแผนนโยบายและกิจกรรมที่ส่งเสริมการปรับตัวของนักเรียนรวมทั้ง 4 ด้าน ได้แก่ ด้านร่างกาย จิตใจ บทบาทหน้าที่ และปฏิสัมพันธ์ โดยคำนึงถึงบริบทเฉพาะของแต่ละโรงเรียน

1.3 ระดับครอบครัว ผู้ปกครองสามารถใช้ความเข้าใจเกี่ยวกับองค์ประกอบการปรับตัวที่สำคัญเป็นแนวทางในการสนับสนุนบุตรหลาน โดยเน้นการสร้างสภาพแวดล้อมที่เอื้อต่อการเรียนรู้และการพัฒนาทักษะการปรับตัวในด้านต่างๆ

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 การศึกษาเปรียบเทียบการปรับตัวทางการเรียนรู้ของนักเรียนในบริบทที่แตกต่างกัน เช่น ในเมืองและชนบท โรงเรียนขนาดเล็กและขนาดใหญ่ เพื่อทำความเข้าใจความแตกต่างของปัญหาและอุปสรรคในการปรับตัวทางการเรียนรู้ของนักเรียน

2.2 การวิจัยในอนาคตควรศึกษาปัจจัยเชิงสาเหตุ พัฒนานวัตกรรมการสอน และประยุกต์ใช้โมเดลในการวางแผนนโยบาย

2.3 เพิ่มการเก็บข้อมูลจากการสังเกตนอกเหนือจากแบบสอบถาม อีกทั้งศึกษาอิทธิพลระหว่างองค์ประกอบ และทำการวิจัยระยะยาวเพื่อติดตามพัฒนาการการปรับตัวของนักเรียน จะช่วยให้ได้ข้อมูลเชิงลึกมากขึ้น

References

- Alessandri, G., De Longis, E., Golfieri, F., & Crocetti, E. (2020). Can Self-Concept Clarity Protect against A Pandemic? A Daily Study on Self-Concept Clarity and Negative Affect during the COVID-19 Outbreak. *Identity*, 21(1), 6–19.
<https://doi.org/10.1080/15283488.2020.1846538>
- Andrews, M., & Roy, S. (1991). The adaptive self: Personal continuity and intentional self-development. In H. Leitner, J. C. Dill, & N. J. Thomas (Eds.), *Advances in personal construct psychology* (Vol. 1, pp. 1–31). Praeger Publishers.
- Asghari, M., Dehghan, M., Shahinfar, S., & Manjiri, E. A. (2022). The role of self-concept, emotion regulation, and coronavirus anxiety in predicting self-care behaviors related with COVID-19 disease. *Journal of Research in Psychological Health*, 15(4), 39–44.
- Ayres, J. S. (2020). The biology of physiological health. *Cell*, 181(2), 250–269.
<https://doi.org/10.1016/j.cell.2020.03.036>
- Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 16(2), 296–298.
- Basaure, M., Joignant, A., & Mascareño, A. (2021). Between distancing and interdependence: The conflict of solidarities in the COVID-19 pandemic. *SAGE Journals*, 64(5), 706–725.
<https://doi.org/10.1177/07311214211005492>
- Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, 16(1), 78–117. <https://doi.org/10.1177/0049124187016001004>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford Publications.
- Burhan, O., Orchida, F., & Hala, I. (2021). Student perspectives of online teaching and learning during the COVID-19 pandemic. *Online Learning Journal*, 25(4), 461–485.
<https://doi.org/10.24059/olj.v25i4.2523>
- Byrne, B. M. (2010). *Structural equation modelling with AMOS: Basic concepts, applications, and programming*. Routledge.
- Caasi, N., & Pentang, J. (2022). Parental factors related to students' self-concept and academic performance amid COVID-19 and distance learning. *Universal Journal of Educational Research*, 1(4), 202–209.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4295472
- Elfirdoussi, S., Lachgar, M., Kabaili, H., Rochdi, A., Goujdami, D., & El Firdoussi, L. (2020). Assessing distance learning in higher education during the COVID-19 pandemic. *Education Research International*, 1–13. <https://doi.org/10.1155/2020/8890633>

- English, H. B., & English, A. C. (1958). *A comprehensive dictionary of psychological and psychoanalytical terms: A guide to usage*. David McKay Company.
- Gaurav, K., Sharma, A., & Chaudhary, R. (2021). Impact of COVID-19 on education: An Indian perspective. *Journal of Ethnic and Cultural Studies*, 8(1), 105-121.
<https://doi.org/10.1177/21501327211056800>
- Gopal, R., Singh, V., & Aggarwal, A. (2021). Impact of online classes on the satisfaction and performance of students during the pandemic period of COVID-19. *Education and Information Technologies*, 26, 6923–6947. <https://doi.org/10.1007/s10639-021-10523-1>
- Hammami, A., Harrabi, B., Mohr, M., & Krustup, P. (2020). Physical activity and coronavirus disease 2019 (COVID-19): specific recommendations for home-based physical training. *Managing Sport and Leisure*, 27(1–2), 26–31.
<https://doi.org/10.1080/23750472.2020.1757494>
- Kilday, J. E., & Ryan, A. M. (2022). The intersection of the peer ecology and teacher practices for student motivation in the classroom. *Educational Psychology Review*, 34(4), 2095–2127. <https://doi.org/10.1007/s10648-022-09712-2>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Lassoued, Z., Alhendawi, M., & Bashitalshaaer, R. (2020). An Exploratory Study of the Obstacles for Achieving Quality in Distance Learning during the COVID-19 Pandemic. *Education Sciences*, 10(9), 232. <https://doi.org/10.3390/educsci10090232>
- Lee, D. J. (2021). Relationships among the Degree of Participation in Physical Activity, Self-Concept Clarity, and COVID-19 Stress in Adolescents. *Healthcare*, 9(4), 482.
<https://doi.org/10.3390/healthcare9040482>
- Rogers, C. R. (1967). *On becoming a person: A therapist's view of psychotherapy*. Houghton Mifflin.
- Singh, D. (1968). Estimates in Successive Sampling Using a Multi-Stage Design. *Journal of the American Statistical Association*, 63(321), 99–112.
<https://doi.org/10.1080/01621459.1968.11009225>
- Singh, S., Roy, D., Sinha, K., Parveen, S., Sharma, G., & Joshi, G. (2020). Impact of COVID-19 and lockdown on mental health of children and adolescents: A narrative review with recommendations. *Psychiatry Research*, 293, 113429.
<https://doi.org/10.1016/j.psychres.2020.113429>

- Utomo, M. N. Y., Sudaryanto, M., & Saddhono, K. (2020). Tools and strategy for distance learning to respond COVID-19 pandemic in Indonesia. *Ingénierie des Systèmes d'Information*, 25(3), 383–390. <https://doi.org/10.18280/isi.250314>
- Yehuda, R. (2021). Stress and mental health in times of COVID-19. *Nature Human Behaviour*, 5(10), 1303-1304.
- Zander, L. (2022). Student well-being and learning in times of COVID-19: What have we learnt and where do we go from here?. *European Journal of Education*, 57(3), 333-339. <https://doi.org/10.1111/ejed.12531>
- Zobeida, S., Pilco, S., Yang, Y., & Zhang, Z. (2022). Student engagement in online learning in Latin American higher education during the COVID-19 pandemic: A systematic review. *British Journal of Educational Technology*, 53(3), 593-619. <https://doi.org/10.1111/bjet.13190>

การสร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล
ของนักเรียนชั้นมัธยมศึกษาตอนปลาย : การประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น
Construction of Career Decision-Making Self-Efficacy Scale in the Digital Age
of High School Students: Application of the Graded Response Model

สุวลักษณ์ อ่อนนวล^{1*} จุฑาภรณ์ มาสันเทียะ² และ ดารุณี ทิพยกุลไพโรจน์³
Suwaluck Onnual^{1*}, Chuthaphon Masantiah² and Darunee Tippayakulpairoj³

(Received: August 7, 2024; Revised: September 30, 2024; Accepted: October 4, 2024)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ (1) สร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ที่มีความยาวแตกต่างกัน ได้แก่ 25 ข้อ และ 35 ข้อ (2) เปรียบเทียบคุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลที่มีความยาวต่างกัน โดยประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น กลุ่มตัวอย่าง ได้แก่ นักเรียนชั้นมัธยมศึกษาตอนปลายโรงเรียนวิทยาศาสตร์จุฬาราชวิทยาลัย บุรีรัมย์ และ เพชรบุรี จำนวน 572 คน คน โดยใช้การสุ่มตัวอย่างแบบแบ่งกลุ่ม เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล ความยาว 25 ข้อ และ 35 ข้อ สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และการวิเคราะห์โมเดลการให้คะแนนแบบลำดับขั้น (Graded Response Model) โดยใช้โปรแกรม MULTILOG 7.03 ผลการวิจัยพบว่า

1. แบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ใช้แนวคิดของ Taylor และ Betz ประกอบด้วย 5 องค์ประกอบ 10 พฤติกรรม บ่งชี้ ได้แก่ (1) การประเมินตนเองได้ถูกต้อง (2) การรวบรวมข้อมูลเกี่ยวกับอาชีพ (3) การเลือกเป้าหมาย (4) การวางแผนในอนาคต (5) การแก้ปัญหา โดยมีความยาว จำนวน 25 ข้อ 1 ฉบับ และ 35 ข้อ 1 ฉบับ

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: 6512681004@rumail.ru.ac.th

2. คุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล พบว่า แบบวัดที่มีความยาว 25 ข้อ มีค่าเฉลี่ยความชันร่วม (α) เท่ากับ 2.088 ค่าเฉลี่ยเทรชโฮลด์ (threshold) ของแต่ละรายการคำตอบ (β_i) พบว่า β_1 , β_2 , β_3 และ β_4 มีค่าเฉลี่ยเท่ากับ -2.944, -1.526, 0.171 และ 1.787 ตามลำดับ ค่าสารสนเทศของแบบวัด (TIF) มีค่าเท่ากับ 35.021 ที่ความสามารถระดับ 0.0 และค่าความเที่ยง เท่ากับ 0.966 และแบบวัดที่มีความยาว 35 ข้อ มีค่าเฉลี่ยความชันร่วม เท่ากับ 1.655 ค่าเฉลี่ยเทรชโฮลด์ (threshold) ของแต่ละรายการคำตอบ (β_i) พบว่า β_1 , β_2 , β_3 และ β_4 มีค่าเฉลี่ยเท่ากับ -2.921, -4.733, -0.020 และ 1.526 ตามลำดับ ค่าสารสนเทศของแบบวัด (TIF) มีค่าเท่ากับ 32.492 ที่ความสามารถระดับ -1.4 และค่าความเที่ยง เท่ากับ 0.965 สรุปได้ว่า แบบวัดที่มีความยาว 25 ข้อ มีค่าเฉลี่ยความชันร่วม (α) ค่าเฉลี่ยเทรชโฮลด์ (threshold) ค่าสารสนเทศของแบบวัด (TIF) และค่าความเที่ยง สูงกว่าแบบวัดที่มีความยาว 35 ข้อ

คำสำคัญ การตัดสินใจเลือกอาชีพ แบบวัดการรับรู้ความสามารถของตนเอง
โมเดลการให้คะแนนแบบลำดับขั้น ยุคดิจิทัล

Abstract

The objectives of this research were to 1) construct a career decision-making self-efficacy scale for high school students in the digital age with two different test lengths: 25 items and 35 items, and 2) compare the psychometric properties of the scales using the graded response model. The sample consisted of 572 high school students from Princess Chulabhorn Science High School Buriram and Phetchaburi, selected through cluster random sampling. The research instruments were the 25-item and 35-item versions of the Career Decision-Making Self-Efficacy Scale in the digital age. Data were analyzed using frequency, percentage, mean, and the graded response model via the MULTILOG 7.03 program. The findings were as follows:

1. The Career Decision-Making Self-Efficacy Scale, based on the framework of Taylor & Betz, comprised five factors and ten indicators: (1) accurate self-appraisal, (2) gathering occupational information, (3) goal selection, (4) planning for the future, and (5) problem-solving. These factors were measured using both the 25-item and 35-item versions of the scale.

2. Regarding psychometric properties, the 25-item version yielded a higher mean slope parameter ($\alpha = 2.088$) and average threshold parameters ($\beta_1 = -2.944$, $\beta_2 = -1.526$, $\beta_3 = 0.171$, $\beta_4 = 1.787$). The test information function (TIF) was 35.021 at $\theta = 0.0$, and the reliability was 0.9655. In contrast, the 35-item version had a lower mean slope parameter ($\alpha = 1.655$) and threshold parameters ($\beta_1 = -2.921$, $\beta_2 = -4.733$, $\beta_3 = -0.020$, $\beta_4 = 1.526$). The TIF was 32.492 at $\theta = -1.4$, and the reliability was 0.9650. In conclusion, the 25-item version

demonstrated superior psychometric qualities—including higher slope parameters, threshold values, test information, and reliability—compared to the 35-item version.

Keywords: career decision-making, self-efficacy scale, Graded Response Model, digital age

บทนำ

การประกอบอาชีพเป็นสิ่งสำคัญในวิถีชีวิตของมนุษย์ เนื่องจากการสร้างรายได้เพื่อเลี้ยงชีพและครอบครัว และเนื่องจากอาชีพมีความหลากหลายแตกต่างกัน โดยแต่ละอาชีพจะมีความเหมาะสมกับบุคคลที่แตกต่างกันออกไป โดยมีแนวทางในการเลือกอาชีพ 3 ประการ ได้แก่ การรู้จักตนเอง การรู้จักอาชีพ และการรู้จักตลาดแรงงาน (Department of Employment, 2016) และเมื่อใดก็ตามที่บุคคลมีการรับรู้ว่าคุณมีความสามารถอะไร สิ่งนั้นจะช่วยกำหนดทางเลือกในการทำสิ่งต่าง ๆ เพื่อประสบความสำเร็จได้ รวมถึงความคาดหวังต่อการรับรู้ความสามารถของตนเอง หากบุคคลมีความคาดหวังต่อการรู้ความสามารถของตนเองต่ำ บุคคลนั้นก็จะมีแนวโน้มที่จะหลีกเลี่ยงงานหรือพฤติกรรมนั้น ๆ และส่งผลให้เกิดความไม่แน่ใจต่อความสามารถของตนเอง (Taylor & Betz, 1983; Lent et al., 1994) ท่ามกลางการเปลี่ยนแปลงของเศรษฐกิจ สังคม และเทคโนโลยี ส่งผลให้มนุษย์เกิดการปรับตัวในด้านต่าง ๆ รวมถึงอนาคตของการทำงานที่เปลี่ยนไป เช่น บางอาชีพอาจสูญหายไปเนื่องจากมีระบบอัตโนมัติและหุ่นยนต์ เป็นต้น (Future Tales LAB, National Innovation Agency and Innovation Foresight Institute, 2021) ซึ่งสอดคล้องกับนโยบายและแผนระดับชาติว่าด้วย การพัฒนาดิจิทัลเพื่อเศรษฐกิจและสังคม (พ.ศ. 2561–2580) เป้าหมายที่ 3 เกี่ยวกับการเตรียมความพร้อม ให้บุคลากรทุกกลุ่ม มีความรู้และทักษะที่เหมาะสมต่อการดำเนินชีวิต และการประกอบอาชีพในยุคดิจิทัลและยุทธศาสตร์ที่ 5 การพัฒนากำลังคนให้พร้อมเข้าสู่ยุคเศรษฐกิจและสังคมดิจิทัล (Ministry of Digital Economy and Society, 2023) และสอดคล้องกับงานวิจัยของ Suwannual et al. (2022) ที่พบว่า กลุ่มทำงาน วัย Generation Y เติบโตมาในยุคที่มีอัตราการเจริญเติบโตทางเศรษฐกิจสูงและมีความเจริญก้าวหน้าทางเทคโนโลยี อย่างมาก ทำให้ชอบการทำงานเป็นทีม หางานที่ถูกใจทำโดยต้องใช้ชีวิตสบายไปพร้อม ๆ กับค่าตอบแทนสูง เป็นต้น ดังนั้น หากมีการส่งเสริมในด้านการศึกษา โดยให้นักเรียนได้รับรู้ความสามารถของตนเอง ก็จะทำให้รู้จักตนเอง รู้จักอาชีพ และรู้จักตลาดแรงงาน ส่งผลให้นักเรียนทราบความสามารถของตนเอง มีข้อมูลที่เท่าทันเศรษฐกิจและความเจริญก้าวหน้า เพื่อประกอบการศึกษาต่อและตัดสินใจเลือกอาชีพในอนาคตได้

ปัจจุบันการจัดการศึกษาในประเทศไทยมีการจัดการเรียนการสอนที่เน้นคณิตศาสตร์และวิทยาศาสตร์ เพื่อเตรียมความพร้อมผู้เรียนให้เท่าทันกับการเปลี่ยนแปลงของเศรษฐกิจ สังคม และเทคโนโลยี โดยเฉพาะอย่างยิ่งการจัดการเรียนการสอนของโรงเรียนวิทยาศาสตร์จุฬาราชวิทยาลัยที่มุ่งเน้นความเป็นเลิศด้านคณิตศาสตร์และด้านวิทยาศาสตร์ สำหรับผู้ที่มีความสามารถด้านคณิตศาสตร์และวิทยาศาสตร์ทั้งระดับมัธยมศึกษาตอนต้นและมัธยมศึกษาตอนปลาย โดยมีการศึกษาต่อในระดับอุดมศึกษาของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ทั้ง 12 แห่ง ในปีการศึกษา 2562 จำนวน 1,703 คน พบว่า นักเรียนเข้าศึกษาต่อในด้านการผลิตและอุตสาหกรรม คิดเป็นร้อยละ 57 (Excellence in Science Education Bureau, 2024) ซึ่งเป็นด้านที่มีผู้เข้าเรียนมากที่สุด และมีความสอดคล้องกับนโยบายและแผนระดับชาติว่าด้วยการพัฒนา

ดิจิทัลเพื่อเศรษฐกิจและสังคม (พ.ศ. 2561 – 2580) จึงมีความเป็นไปได้ว่านักเรียนที่สำเร็จการศึกษาในระดับชั้นมัธยมศึกษาปีที่ 6 มีการรับรู้ความสามารถของตนเอง และมีความมุ่งมั่นตั้งใจที่จะเข้าศึกษาต่อในระดับอุดมศึกษา ดังนั้น ในช่วงที่นักเรียนกำลังศึกษาในระดับมัธยมศึกษาตอนปลายจึงเป็นช่วงเวลาที่สำคัญที่ควรมีการส่งเสริมให้นักเรียนได้เรียนรู้ในสิ่งที่ตนเองสนใจ สำรวจความสามารถ ความถนัด ความต้องการของตนเอง ค้นหาข้อมูลเกี่ยวกับอาชีพที่เหมาะสมกับตนเอง รวมถึงความต้องการของตลาดแรงงานที่มีการเปลี่ยนแปลงไปตามเศรษฐกิจ เพื่อความเจริญก้าวหน้าในทางหน้าที่การงานของตนเองในอนาคต ซึ่งสอดคล้องกับวัตถุประสงค์ในการจัดกิจกรรมแนะแนวของนักเรียนระดับมัธยมศึกษาตอนปลาย ที่ต้องการให้ผู้เรียนได้รู้จัก เข้าใจ รักและเห็นคุณค่าในตนเองและผู้อื่น สามารถวางแผน การเรียน อาชีพ รวมทั้งการดำเนินชีวิต และสังคมได้ รวมถึงสามารถปรับตัวได้อย่างเหมาะสม และอยู่ร่วมกับผู้อื่นอย่างมีความสุข (Office of the Basic Education Commission, 2012) และทฤษฎีการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพ ของ Taylor & Betz (1983) ประกอบด้วย 5 องค์ประกอบ ได้แก่ การประเมินตนเองได้ถูกต้อง การรวบรวมข้อมูลเกี่ยวกับอาชีพ การเลือกเป้าหมาย การวางแผนในอนาคต และการแก้ปัญหา

ทฤษฎีการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพ ของ Taylor & Betz (1983) เป็นแนวคิดที่ประยุกต์ใช้องค์ความรู้และความคาดหวังในการรับรู้ความสามารถของตนเองกับการขาดความมั่นใจในการตัดสินใจด้านอาชีพ จึงกำหนดนิยามใหม่ได้ว่า ความคาดหวังต่อการรับรู้ความสามารถของตนเองต่ำ จะมีแนวโน้มที่จะหลีกเลี่ยงงานหรือพฤติกรรมนั้น ๆ และส่งผลให้เกิดความไม่แน่ใจ จึงนำไปสู่การสร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพ (Career Decision-Making Self-Efficacy Scale : CDMSE) ที่ประกอบด้วย 5 องค์ประกอบตามที่กล่าวไปข้างต้น และจากการศึกษางานวิจัยที่เกี่ยวข้องกับการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพ พบว่า งานวิจัยส่วนมากใช้แบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพ ของ Taylor & Betz (1983) ความยาว 50 ข้อ และแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพ แบบสั้น (short form) ของ Betz et al. (1996) ความยาว 25 ข้อ เป็นเครื่องมือในการวิจัย โดยมีลักษณะเป็นแบบวัด (แบบมาตรประมาณค่า) ซึ่งจากงานวิจัยของ Nilsson et al. (2002) ที่ได้ทบทวนวรรณกรรมเพื่อสำรวจความเที่ยงของแบบวัด ทั้งฉบับของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพ ความยาว 50 ข้อ (จำนวน 11 เรื่อง) และ ความยาว 25 ข้อ (จำนวน 11 เรื่อง) พบว่า แบบวัดความยาว 50 ข้อ มีค่าความเที่ยงอยู่ระหว่าง 0.83 – 0.97 ($M = 0.95$, $SD = 0.4$) ในขณะที่แบบวัดความยาว 25 ข้อ มีค่าความเที่ยงอยู่ระหว่าง 0.92 – 0.97 ($M = 0.94$, $SD = 0.1$) จากผลการวิจัยแม้ค่าเฉลี่ยของความเที่ยงของแบบวัดความยาว 50 ข้อ จะสูงกว่า 25 ข้อ แต่ความแปรปรวน ของแบบวัดความยาว 25 ข้อนั้นน้อยกว่า และมีการตรวจสอบคุณภาพของเครื่องมือโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ซึ่งมีข้อจำกัดที่สำคัญ คือ การที่พารามิเตอร์ของข้อสอบผันแปรตามกลุ่มผู้สอบและคะแนนที่สังเกตได้ หรือค่าประมาณความสามารถของผู้ตอบไม่เป็นอิสระหรือขึ้นอยู่กับข้อสอบและแบบสอบที่นำมาใช้ (Kanjana-wasee, 2013) ดังนั้น ในงานวิจัยนี้ผู้วิจัยจึงสนใจสร้างแบบวัดที่มีความยาว 35 ข้อ และ 25 ข้อ เนื่องจากความยาวของแบบวัดมีผลต่อทั้งความแปรปรวนของคะแนนจริงและความแปรปรวนของคะแนนที่สังเกตได้ ผู้วิจัยจึงสร้างแบบวัดที่มีความยาว 35 ข้อ ที่ไม่สั้นและยาวจนเกินไป และเลือกใช้ทฤษฎีการตอบสนองแนวโน้มใหม่ (IRT) เนื่องจากทฤษฎีการตอบสนองข้อสอบมีพื้นฐานความเชื่อว่า พฤติกรรมการตอบสนองต่อข้อสอบของผู้สอบ

ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิดจะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่อยู่ภายในตัวบุคคล ทำให้ผลการวัดชัดเจนตรงประเด็นมากขึ้น สามารถนำแบบวัดหรือแบบสอบถามไปใช้ในสถานการณ์ต่าง ๆ ได้อย่างมีประสิทธิภาพ (Kanjanaawasee, 2013; Kanjanaawasee, 2020)

จากเหตุผลที่กล่าวมาข้างต้น เพื่อเป็นการเตรียมรับมือกับการเปลี่ยนแปลงของเศรษฐกิจ สังคม และเทคโนโลยี และตอบสนองต่อนโยบายและแผนระดับชาติว่าด้วยการพัฒนาดิจิทัลเพื่อเศรษฐกิจและสังคม (พ.ศ. 2561 – 2580) ผู้วิจัยจึงสนใจสร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลที่มีความยาวแตกต่างกัน และเปรียบเทียบคุณสมบัติทางจิตมิติของแบบวัดที่มีความยาวต่างกัน โดยใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุวิภาค เนื่องจากงานวิจัยที่ผ่านมามีการตรวจสอบคุณภาพของแบบวัดหรือแบบสอบถาม โดยทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ผู้วิจัยจึงใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุวิภาค เพื่อให้ได้แบบวัดที่มีประสิทธิภาพ และจากการศึกษางานวิจัยที่ผ่านมายังพบว่าการพัฒนาเครื่องมือ (มาตรวัด/แบบวัด/แบบสอบถาม) ในรูปแบบของมาตรประมาณค่า ซึ่งมีการใช้โมเดลที่หลากหลาย เช่น Rash Model (RM) Graded Response Model (GRM) และ Generalized Partial Credit Model (G-PCM) เป็นต้น เพื่อตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด ซึ่งโมเดลที่นิยมใช้มากที่สุด คือ Graded Response Model (GRM) (Am et al., 2023; Cahyawulan & Yundianto, 2022; Herwin et al., 2023; Kassakul, 2020; Liu et al., 2022; Miguel et al., 2013; Montgomery et al., 2023; Molek-Winiarska & Zięba, 2019; Rungrotrangsarn, 2017; Yaşar, 2019)

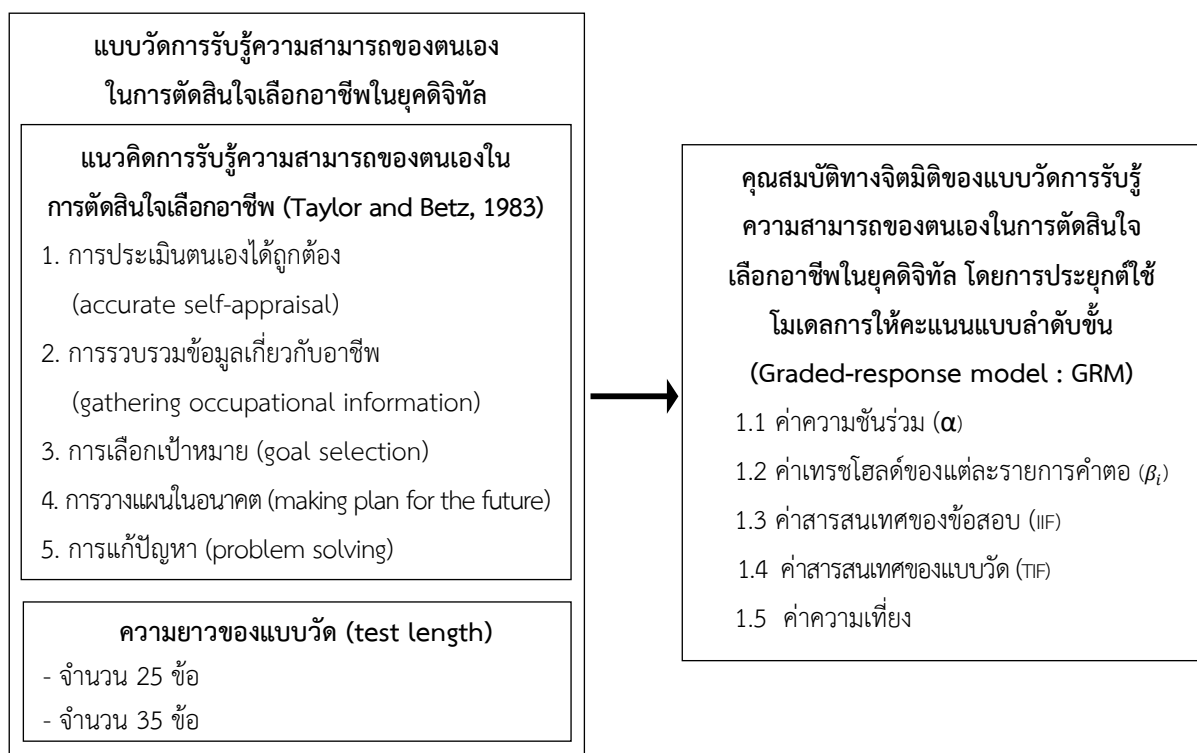
ในงานวิจัยครั้งนี้ผู้วิจัยสนใจสร้างแบบวัดที่มีการตรวจให้คะแนนมากกว่า 2 ค่า จึงประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น (Graded Response Model : GRM) เพื่อตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด โดยมีการวิเคราะห์ ค่าความชันร่วม (α) ค่าเทรชโฮลด์ (threshold) ของแต่ละรายการคำตอบ (β_i) ค่าสารสนเทศของข้อสอบ (IIF) ค่าสารสนเทศของแบบวัด (TIF) และค่าความเที่ยง ซึ่งผลจากการวัดที่ได้จากเครื่องมือที่มีคุณสมบัติทางจิตมิติที่ดีย่อมส่งผลให้ข้อมูลมีความน่าเชื่อถือในการนำไปใช้ สะท้อนการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพและผลจากการวัดระดับการรับรู้ ทำให้ครูผู้สอน ตลอดจนผู้ที่เกี่ยวข้อง สามารถนำไปใช้เป็นสารสนเทศในการวางแผนการจัดการเรียนการสอนหรือโปรแกรม กิจกรรมแนะแนวต่าง ๆ เพื่อให้ผู้เรียนสามารถตัดสินใจเลือกอาชีพได้เหมาะสมกับตนเองภายใต้บริบทของยุคดิจิทัลต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อสร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ที่มีความยาวแตกต่างกัน ได้แก่ 25 ข้อ และ 35 ข้อ
2. เพื่อเปรียบเทียบคุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ที่มีความยาวต่างกัน โดยประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น

กรอบแนวคิดในการวิจัย

ผู้วิจัยได้ศึกษาเอกสาร แนวคิด และทฤษฎีที่เกี่ยวกับการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล โดยมีรายละเอียดดังนี้ (1) สังเคราะห์นิยามและองค์ประกอบของการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล โดยใช้แนวคิดของ (Taylor & Betz, 1983) ประกอบด้วย 5 องค์ประกอบ ได้แก่ การประเมินตนเองได้ถูกต้อง การรวบรวมข้อมูลเกี่ยวกับอาชีพ การเลือกเป้าหมาย การวางแผนในอนาคต และการแก้ปัญหา (2) กำหนดความยาวของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล มี 2 แบบ ได้แก่ จำนวน 25 ข้อ และ จำนวน 35 ข้อ (3) ตรวจสอบคุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล โดยใช้โมเดลการให้คะแนนแบบลำดับขั้น (Graded Response Model : GRM) ได้แก่ ค่าความชันร่วม (α) ค่าเทรชโฮลด์ (threshold) ของแต่ละรายการคำตอบ (β_i) ค่าสารสนเทศของข้อสอบ (IIF) ค่าสารสนเทศของแบบวัด (TIF) และค่าความเที่ยง และสามารถสรุปเป็นกรอบแนวคิดในการวิจัยได้ ดังภาพ 1



ภาพ 1 กรอบแนวคิดในการวิจัย

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย คือ นักเรียนชั้นมัธยมศึกษาตอนปลายโรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย จำนวน 12 โรงเรียน จำนวน 5,205 คนที่กำลังศึกษาในภาคเรียนที่ 1 ปีการศึกษา 2567

กลุ่มตัวอย่างที่ใช้ในการวิจัย นักเรียนชั้นมัธยมศึกษาตอนปลายของโรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย เนื่องจากเป็นนักเรียนที่มีคุณลักษณะที่คล้ายคลึงกัน ผู้วิจัยจึงใช้การสุ่มกลุ่มตัวอย่างแบบแบ่งกลุ่ม

(cluster random sampling) ตามเขตพื้นที่บริการของโรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย จำนวน 2 โรงเรียน ได้แก่ (1) โรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย บุรีรัมย์ นักเรียนจำนวน 303 คน และ (2) โรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย เพชรบุรี นักเรียนจำนวน 269 คน รวมทั้งหมด 572 คน และเนื่องจากการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) โดยประมาณค่าพารามิเตอร์ที่มีการตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous IRT Models) โมเดลการให้คะแนนแบบลำดับขั้น (Graded Response Model : GRM) จึงควรใช้กลุ่มตัวอย่าง อย่างน้อย 250 คน และหากต้องการให้ผลดีกลุ่มตัวอย่างนั้นไม่ควรต่ำกว่า 500 คน (Reise & Yu, 1990, as cited in Kanjanawasee, 2020) ดังนั้นเพื่อความแม่นยำของงานวิจัย และพิจารณาบนพื้นฐานของความเป็นไปได้ในการเก็บรวบรวมข้อมูล ผู้วิจัยจึงปรับขนาดกลุ่มตัวอย่างเป็น 572 คน เพื่อให้มีความครอบคลุม และเป็นตัวแทนที่ดีของประชากร และเมื่อแบ่งออกเป็น 2 กลุ่ม ผู้วิจัยได้คำนวณขนาดกลุ่มตัวอย่างทั้ง 2 กลุ่ม ด้วยสูตรของ Yamane (1973, as cited in Thirakanan, 2014) พบว่า กลุ่มตัวอย่างที่ 1 มีความคลาดเคลื่อนที่เกิดขึ้นและยอมรับได้ เท่ากับ 0.05 และกลุ่มตัวอย่างที่ 2 มีความคลาดเคลื่อนที่เกิดขึ้นและยอมรับได้ เท่ากับ 0.05 เช่นกัน

2. เครื่องมือที่ใช้ในการวิจัย

แบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย จำนวน 2 ฉบับ ได้แก่ ฉบับที่ 1 มีข้อความจำนวน 25 ข้อ และฉบับที่ 2 มีข้อความจำนวน 35 ข้อ โดยประยุกต์ใช้ทฤษฎีของ Taylor & Betz (1983) ประกอบด้วย (1) การประเมินตนเองได้ถูกต้อง (2) การรวบรวมข้อมูลเกี่ยวกับอาชีพ (3) การเลือกเป้าหมาย (4) การวางแผนในอนาคต และ (5) การแก้ปัญหา ซึ่งในแต่ละฉบับจะมีจำนวนข้อความแตกต่างกันออกไปขึ้นอยู่กับพฤติกรรมบ่งชี้ในแต่ละองค์ประกอบ กล่าวคือ (1) แบบวัดความยาว 25 ข้อ ประกอบด้วย องค์ประกอบที่ 1 (5 พฤติกรรมบ่งชี้ 10 ข้อคำถาม) องค์ประกอบที่ 2 (2 พฤติกรรมบ่งชี้ 6 ข้อคำถาม) องค์ประกอบที่ 3 (1 พฤติกรรมบ่งชี้ 3 ข้อคำถาม) องค์ประกอบที่ 4 (1 พฤติกรรมบ่งชี้ 3 ข้อคำถาม) และองค์ประกอบที่ 5 (1 พฤติกรรมบ่งชี้ 3 ข้อคำถาม) (2) แบบวัดความยาว 35 ข้อ องค์ประกอบที่ 1 (5 พฤติกรรมบ่งชี้ 15 ข้อคำถาม) องค์ประกอบที่ 2 (2 พฤติกรรมบ่งชี้ 8 ข้อคำถาม) องค์ประกอบที่ 3 (1 พฤติกรรมบ่งชี้ 4 ข้อคำถาม) องค์ประกอบที่ 4 (1 พฤติกรรมบ่งชี้ 4 ข้อคำถาม) และองค์ประกอบที่ 5 (1 พฤติกรรมบ่งชี้ 4 ข้อคำถาม) โดยลักษณะของเครื่องมือ เป็นแบบมาตราประมาณค่า และให้เลือกระดับความมั่นใจในการรับรู้ความสามารถของตนเอง 5 ระดับ คือ 5 4 3 2 และ 1 มีความหมายว่าฉันมีความมั่นใจมากที่สุด มาก ปานกลาง น้อยน้อยที่สุด ตามลำดับ ผู้วิจัยนำเสนอตัวอย่างข้อคำถามดังตาราง 1

ตาราง 1 ตัวอย่างข้อคำถามในแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย

องค์ประกอบ	พฤติกรรมบ่งชี้	ตัวอย่างข้อคำถาม
1	1.1	ฉันทำแบบทดสอบเพื่อวัดความสามารถของตนเอง เช่น แบบทดสอบในวิชาแนะแนว เว็บไซต์ แบบทดสอบค้นหาตนเอง
	1.2	ฉันมักจะติดตามบัญชีบนโลกออนไลน์ซึ่งมีเนื้อหาที่ตนเองสนใจ เช่น ช่องยูทูบทำอาหาร ช่องติ๊กต็อกสอนเต้น เพจเฟซบุ๊กสัตว์เลี้ยง

ตาราง 1 (ต่อ)

องค์ประกอบ	พฤติกรรมบ่งชี้	ตัวอย่างข้อคำถาม
2	1.3	ฉันมักจะศึกษารูปแบบการใช้ชีวิตจากสื่อบนโลกออนไลน์เพื่อเป็นแนวทางให้กับตนเอง เช่น ยูทูบ เฟสบุ๊ก ดิกด็อก
	1.4	ฉันคาดเดาได้ว่าอาชีพที่มีความเหมาะสมกับตนเองมากที่สุดคืออาชีพใด เช่น ฉันตัดต่อวิดีโอได้เป็นอย่างดี จึงคิดจะทำอาชีพนักตัดต่อวิดีโอ
	1.5	ฉันมั่นใจว่าตนเองจะทำงานได้เป็นอย่างดี แม้ว่าจะเกิดการเปลี่ยนแปลงอย่างกะทันหัน
	2.1	ฉันบอกได้ว่าอาชีพที่ตรงกับความสนใจและความสามารถของตนเองเป็นอย่างไรจากการสอบถาม/ค้นคว้าข้อมูลเพิ่มเติมจากแหล่งอื่น
	2.2	ฉันสอบถามข้อมูลเพิ่มเติมจากองค์กร/หน่วยงาน/รุ่นพี่/ครู ผ่านช่องทางออนไลน์เพื่อสอบถามข้อมูลเกี่ยวกับแผนการเรียน/สถานศึกษา
3	3.1	ฉันจะคำนึงถึงความต้องการของตลาดแรงงานก่อนการตัดสินใจเลือกอาชีพ เช่น อาชีพด้านไอที อาชีพด้านวิศวกร อาชีพด้านขนส่งและโลจิสติกส์
4	4.1	ฉันเรียงลำดับความสำคัญสิ่งที่ต้องทำในชีวิตประจำวันได้เป็นอย่างดี
5	5.1	หากฉันต้องอ่านหนังสือที่มีเนื้อหาค่อนข้างเยาะและยาก ฉันสามารถเปลี่ยนวิธีการอ่านหนังสือ เช่น ฟังสรุปเนื้อหาบนยูทูบ วิดีโอคอลกับเพื่อนเพื่อแลกเปลี่ยนความรู้

3. การเก็บรวบรวมข้อมูล

ผู้วิจัยเก็บรวบรวมข้อมูลกับนักเรียนที่เป็นกลุ่มตัวอย่าง โดยมีขั้นตอนดังนี้

1. ทำหนังสือขอความร่วมมือในการเก็บรวบรวมข้อมูลจากคณะศึกษาศาสตร์ถึงผู้บริหารโรงเรียนที่เป็นกลุ่มตัวอย่างทั้ง 2 โรงเรียน เพื่อขออนุญาตและขอความร่วมมือในการเก็บรวบรวมข้อมูล
2. ดำเนินการส่งหนังสือขอความร่วมมือในการเก็บรวบรวมข้อมูลและแบบวัดให้กับผู้ประสานงานของโรงเรียนทั้ง 2 โรงเรียน ซึ่งนักเรียนแต่ละคนจะได้ทำแบบวัดคนละ 1 ฉบับ โดยมีรายละเอียดดังนี้
 - 2.1 ส่งหนังสือขอความร่วมมือในการเก็บรวบรวมข้อมูลและแบบลิงก์ Google Form แบบวัดที่มีความยาว จำนวน 25 ข้อ ให้กับโรงเรียนวิทยาศาสตร์จุฬารณราชวิทยาลัย บุรีรัมย์
 - 2.2 ส่งหนังสือขอความร่วมมือในการเก็บรวบรวมข้อมูลและแบบลิงก์ Google Form แบบวัดที่มีความยาว จำนวน 35 ข้อ ให้กับโรงเรียนวิทยาศาสตร์จุฬารณราชวิทยาลัย เพชรบุรี
3. กำกับติดตามแบบวัดคืนโดยการติดต่อกับผู้ประสานงานของแต่ละโรงเรียนอย่างสม่ำเสมอ
4. ตรวจสอบจำนวนและความสมบูรณ์ของการตอบกลับและวิเคราะห์ข้อมูลต่อไป

4. การวิเคราะห์ข้อมูล

4.1 การสร้างแบบวัดที่มีความยาวต่างกัน โดยมีขั้นตอนดังต่อไปนี้

1. ผู้วิจัยสร้างข้อคำถามจำนวน 50 ข้อ โดยแต่ละพฤติกรรมบ่งชี้จะมีข้อคำถามแบ่งเป็น 3 ชุด ได้แก่ ชุดที่ 1 (A) ชุดที่ 2 (B) และ ชุดที่ 3 (C) เพื่อให้ข้อคำถามในแบบวัดครอบคลุมกับเนื้อหาและมีความสอดคล้องกัน เมื่อแยกเป็น 2 ฉบับ จากนั้นตรวจสอบค่าดัชนีความสอดคล้องระหว่างวัตถุประสงค์และข้อคำถาม (IOC) จากผู้เชี่ยวชาญ 3 ท่าน โดยค่า IOC ตั้งแต่ 0.50 ขึ้นไป ถือว่าผ่านเกณฑ์ (Thirakanan,

2008) ซึ่งพบว่า มีข้อคำถามที่ผ่านเกณฑ์ จำนวน 47 ข้อ และปรับปรุงข้อคำถามให้มีความเหมาะสมตามข้อเสนอแนะของผู้เชี่ยวชาญ

2. คัดเลือกข้อคำถามเพียง 35 ข้อ โดยพิจารณาจากข้อคำถามที่ผ่านเกณฑ์การวิเคราะห์ค่าดัชนีความสอดคล้องระหว่างวัตถุประสงค์และข้อคำถาม (IOC) จากผู้เชี่ยวชาญทั้ง 3 ท่าน โดยพิจารณาค่า IOC ในแต่ละชุดข้อคำถาม (ชุด A, ชุด B, ชุด C) ตั้งแต่ 0.50 ขึ้นไป ของแต่ละพฤติกรรมบ่งชี้ ซึ่งข้อคำถามที่คัดเลือกส่วนมากจะมีค่า IOC เท่ากับ 1.00 แต่ในพฤติกรรมบ่งชี้ 1.1, 1.3, 1.4, 1.5 และ 3.1 ผู้วิจัยคัดเลือกข้อคำถามที่มีค่า IOC เท่ากับ 0.66 โดยคำนึงถึงความเหมาะสมของข้อคำถาม

3. จากนั้นนำแบบวัดความยาว 35 ข้อ ไปทดสอบกับกลุ่มทดลองใช้ (try out) กับกลุ่มทดลองที่มีลักษณะใกล้เคียงกับกลุ่มตัวอย่าง เพื่อตรวจสอบค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม (corrected item-total correlation) ซึ่งพบว่า ข้อคำถามทั้ง 35 ข้อ มีค่าอยู่ระหว่าง 0.57 ถึง 0.87 โดยมีค่าตั้งแต่ 0.2 ขึ้นไปถึงว่าผ่านเกณฑ์ และค่าความเที่ยง โดยใช้สัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Alpha coefficient; α) พบว่า มีค่าความเที่ยง เท่ากับ 0.981 ซึ่งควรมีค่ามากกว่า 0.7 ขึ้นไป พบว่า มีค่าความเที่ยง เท่ากับ 0.985

4. นำผลที่ได้ในข้างต้นมาพิจารณาเพื่อคัดเลือกข้อคำถาม จำนวน 25 ข้อ โดยพิจารณาคัดเลือกจากข้อคำถามในแต่ละชุดที่มีค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม (corrected item-total correlation) สูงที่สุดตามลำดับ ซึ่งแม้ว่าจะตัดข้ออื่น ๆ ออกแต่ก็ยังคงเนื้อหาที่ครอบคลุมตามองค์ประกอบ จึงสามารถสรุปแผนผังของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล ความยาว 25 ข้อ และ 35 ข้อ โดยมีรายละเอียด ดังตาราง 2

4.2 ตรวจสอบข้อตกลงเบื้องต้นของ IRT โดยการตรวจสอบความเป็นเอกมิติ (unidimensional) ผู้วิจัยพิจารณาจากอัตราส่วนค่าไอเกนองค์ประกอบที่ 1 กับค่าไอเกนองค์ประกอบที่ 2 (E1/E2) ที่ควรมีค่ามากกว่า 3.00 เพื่อความเป็นเอกมิติของแบบวัด (unidimensional)

4.3 การเปรียบเทียบคุณสมบัติทางจิตมิติของแบบวัด

พิจารณาจากค่าพารามิเตอร์ความชันร่วม (α) ค่าพารามิเตอร์เทรชโฮลด์ (threshold: β) ค่าสารสนเทศรายข้อ (item information) ตรวจสอบคุณภาพทั้งฉบับจากสารสนเทศของแบบวัด (test information) และค่าความเที่ยง เพื่อตรวจสอบคุณสมบัติทางจิตมิติของแบบวัด โดยใช้โปรแกรม MULTILOG 7.03 (Kanjawasee, 2020) ดังสูตร

$$P_{ix}^*(\theta) = \frac{\exp [\alpha_i(\theta) - \beta_{ij}]}{1 + \exp [\alpha_i(\theta) - \beta_{ij}]}$$

เมื่อ $X = j = 1, \dots, m_i$

$P_{ix}^*(\theta)$ หมายถึง ความน่าจะเป็นที่ผู้ตอบซึ่งมีคุณลักษณะระดับ θ จะตอบข้อ i ด้วยการเลือกรายการคำตอบที่ X เมื่อ $X = 1, 2, \dots, m_i$

α_i หมายถึง ค่าพารามิเตอร์ความชันร่วม (slope parameter) ของข้อที่ i

β_{ij} หมายถึง ค่าพารามิเตอร์ threshold ของแต่ละรายการคำตอบ (Threshold parameter) ของข้อที่ i

ตรวจสอบคุณภาพรายข้อจากค่าความชันร่วม (α) โดยพิจารณาจากค่าอำนาจจำแนกตามเกณฑ์ ดังนี้ (Baker, 2001)

0.01 - 0.34	หมายถึง	ค่าอำนาจจำแนกต่ำมาก
0.35 - 0.64	หมายถึง	ค่าอำนาจจำแนกต่ำ
0.65 - 1.34	หมายถึง	ค่าอำนาจจำแนกปานกลาง
1.35 - 1.69	หมายถึง	ค่าอำนาจจำแนกสูง
ตั้งแต่ 1.70 ขึ้นไป	หมายถึง	ค่าอำนาจจำแนกสูงมาก

ตรวจสอบค่าพารามิเตอร์เทอร์สโฮลด์ (Threshold) ของแต่ละรายการคำตอบ ($\beta_1 < \beta_2 < \beta_3 < \beta_4$) โดยที่ค่า β เป็นค่าเรียงจากน้อยไปมาก ควรมีค่าตั้งแต่ -2.50 ถึง +2.50 (Kanjawasee, 2020)

ตรวจสอบค่าสารสนเทศรายข้อ (Item Information) โดยพิจารณาพารามิเตอร์ a, b, และ c ของข้อสอบ ซึ่งสรุปเป็นแนวทางได้ดังนี้ (Kanjawasee, 2020) (โดยในบทความวิจัยนี้ ผู้วิจัยนำเสนอค่าสารสนเทศของข้อคำถามที่มีค่าสูงที่สุดในแต่ละฉบับ)

- ค่าสารสนเทศของข้อสอบจะสูงขึ้น สำหรับผู้สอบมีความสามารถใกล้เคียงกับค่าพารามิเตอร์ b ของข้อสอบ และค่าสารสนเทศจะลดลงสำหรับผู้สอบที่ความสามารถไกลจากค่าพารามิเตอร์ b ของข้อสอบ
- ค่าสารสนเทศของข้อสอบโดยทั่วไปจะมีค่าสูงขึ้น ถ้าค่าพารามิเตอร์ a ของข้อสอบมีค่ามากขึ้น
- ค่าสารสนเทศของข้อสอบมีค่าสูงขึ้น เมื่อค่าพารามิเตอร์ c ของข้อสอบเข้าใกล้ 0

ตรวจสอบคุณภาพทั้งฉบับจากสารสนเทศของแบบวัด (test information) โดยที่ค่าสารสนเทศมีความสัมพันธ์ผกผันกับความคลาดเคลื่อนมาตรฐานของการประมาณค่า ดังนั้น ถ้าค่าสารสนเทศของแบบสอบมีค่า θ ไต่ก็จะมีค่าถูกต้องแม่นยำสูงในการประมาณค่าความสามารถของผู้ตอบในช่วง θ นั้น ๆ โดยมีความคลาดเคลื่อนมาตรฐานของการประมาณค่าต่ำ (Kanjawasee, 2020)

ตรวจสอบค่าความเที่ยงของแบบวัด โดยใช้สัมประสิทธิ์แอลฟา (Alpha coefficient) ของครอนบาค (Cronbach) ซึ่งความเที่ยงของแบบวัด ควรมีค่ามากกว่า 0.7 ขึ้นไป (Nunnally, 1978 as cited in Thirakanan, 2008)

ผลการวิจัย

1. ผลการสร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดีดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ที่มีความยาวแตกต่างกัน

แบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดีดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ใช้แนวคิดของ (Taylor & Betz, 1983) ประกอบด้วย 5 องค์ประกอบ และ 10 พฤติกรรมบ่งชี้ ได้แก่ (1) การประเมินตนเองได้ถูกต้อง (2) การรวบรวมข้อมูลเกี่ยวกับอาชีพ (3) การเลือกเป้าหมาย (4) การวางแผนในอนาคต และ (5) การแก้ปัญหา โดยหลังจากการสร้างข้อคำถามจำนวน 50 ข้อ จากนั้นนำไปตรวจสอบคุณภาพของแบบวัดจากผู้เชี่ยวชาญทั้ง 3 ท่าน และทดลองใช้ (try out) กับกลุ่มทดลอง พบว่า ข้อคำถามทั้ง 35 ข้อ มีค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม (corrected

item-total correlation) อยู่ระหว่าง 0.57 ถึง 0.88 ซึ่งข้อคำถามผ่านเกณฑ์ทั้ง 35 ข้อ และมีค่าความเที่ยงเท่ากับ 0.985 จากนั้นคัดเลือกข้อคำถาม จำนวน 25 ข้อ โดยพิจารณาคัดเลือกจากข้อคำถามในแต่ละชุดที่มีค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมสูงที่สุดตามลำดับ ซึ่งแม้ว่าจะตัดข้ออื่น ๆ ออกแต่ก็ยังคงเนื้อหาที่ครอบคลุมตามองค์ประกอบ สามารถสรุปแผนผังของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพ ในยุคดิจิทัล ความยาว 25 ข้อ และ 35 ข้อ โดยมีรายละเอียดดังตาราง 2

ตาราง 2 แผนผังของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ความยาว 25 ข้อ และ 35 ข้อ

องค์ประกอบ	พฤติกรรม บ่งชี้	ข้อคำถาม			ความยาวของแบบวัด	
		ชุดที่ 1 (A)	ชุดที่ 2 (B)	ชุดที่ 3 (C)	35 ข้อ	25 ข้อ
1. การประเมินตนเอง ได้ถูกต้อง	1.1	1A, 2A	1B, 2B	-	3 ข้อ (2A, 1B, 2B)	2 ข้อ (2A,1B)
	1.2	1A, 2A	1B, 2B	-	3 ข้อ (1A, 2A, 1B)	2 ข้อ (2A,1B)
	1.3	1A, 2A	1B, 2B	-	3 ข้อ (1A,2A,2B)	2 ข้อ (2A,2B)
	1.4	1A, 2A	1B, 2B	-	3 ข้อ (A, 1B, 2B)	2 ข้อ (2A,1B)
	1.5	1A, 2A	1B, 2B	-	3 ข้อ (1A, 2A, 1B)	2 ข้อ (1A,1B)
2. การรวบรวมข้อมูล เกี่ยวกับอาชีพ	2.1	1A, 2A	1B, 2B	1C, 2C	4 ข้อ (1A,1B, 2B,2C)	3 ข้อ (1A,1B,2C)
	2.2	1A, 2A	1B, 2B	1C, 2C	4 ข้อ (1A,1B, 2B,1C)	3 ข้อ (1A,1B,1C)
3. การเลือกเป้าหมาย	3.1	1A, 2A	1B, 2B	1C, 2C	4 ข้อ (1A,2A, 1B,1C)	3 ข้อ (2A,1B,1C)
4. การวางแผนในอนาคต	4.1	1A, 2A	1B, 2B	1C, 2C	4 ข้อ (1A,2A, 1B,1C)	3 ข้อ (2A,1B,1C)
5. การแก้ปัญหา	5.1	1A, 2A	1B, 2B	1C, 2C	4 ข้อ (2A,2B,1C,2C)	3 ข้อ (2A,2B,2C)
รวม		20	20	10	35	25

จากตาราง 2 แผนผังของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ความยาว 25 ข้อ และ 35 ข้อ ประกอบด้วย 5 องค์ประกอบ 3 ชุดข้อคำถาม ได้แก่ ชุด A, ชุด B, และ ชุด C เพื่อให้ครอบคลุมเนื้อหาตามพฤติกรรมบ่งชี้เมื่อต้องคัดเลือกข้อคำถามออกเป็น 2 ฉบับ โดยการคัดเลือกข้อคำถามในแบบวัดความยาว 25 ข้อ เกิดจากการพิจารณาค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม (corrected item-total correlation) ในแบบวัดความยาว 35 ข้อ ที่มีค่าสูงที่สุด ในแต่ละชุดข้อคำถาม (A,B,C) ของแต่ละพฤติกรรมบ่งชี้ ตัวอย่างเช่น แบบวัดความยาว 35 ข้อ องค์ประกอบที่ 1 การประเมินตนเองได้ถูกต้อง พฤติกรรมบ่งชี้ที่ 1.1 ระบุความสามารถของตนเองได้โดยใช้วิธีการที่หลากหลายในการสำรวจตนเอง มีข้อคำถาม จำนวน 3 ข้อ ได้แก่ 2A, 1B, และ 2B ซึ่งทุกข้อนั้นมีค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมผ่านเกณฑ์ทั้งหมด แต่เมื่อต้องคัดเลือกข้อคำถามเพื่อสร้างแบบวัดความยาว 25 ข้อ จะพิจารณาค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมที่สูงที่สุดในแต่ละชุด ดังนั้น ในแบบวัดความยาว 25 ข้อ องค์ประกอบที่ 1 พฤติกรรมบ่งชี้ 1.1 จึงประกอบด้วยข้อคำถาม 2 ข้อ ได้แก่ 2A และ 1B จึงสะท้อนให้เห็นถึงความเท่าเทียมของแบบวัดทั้งสองฉบับ

2. ผลการเปรียบเทียบคุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ที่มีความยาวต่างกัน โดยประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น (Graded Response Model: GRM)

2.1 ผลการตรวจสอบข้อตกลงเบื้องต้นของ IRT โดยการตรวจสอบความเป็นเอกมิติ (Unidimensional) ซึ่งผู้วิจัยพิจารณาจากอัตราส่วนค่าไอเกนองค์ประกอบที่ 1 กับอัตราส่วนค่าไอเกนองค์ประกอบที่ 2 (E1/E2) ซึ่งควรมีค่ามากกว่า 3.00 พบว่า แบบวัดความยาว 25 ข้อ มีอัตราส่วนค่าไอเกนเท่ากับ 7.842 และแบบวัดความยาว 35 ข้อ มีอัตราส่วนค่าไอเกน เท่ากับ 5.842 ซึ่งมีความสูงผ่านเกณฑ์การตรวจสอบข้อตกลงเบื้องต้น สะท้อนให้เห็นว่าแบบวัดทั้ง 2 ฉบับ มีความเป็นเอกมิติ (unidimensional)

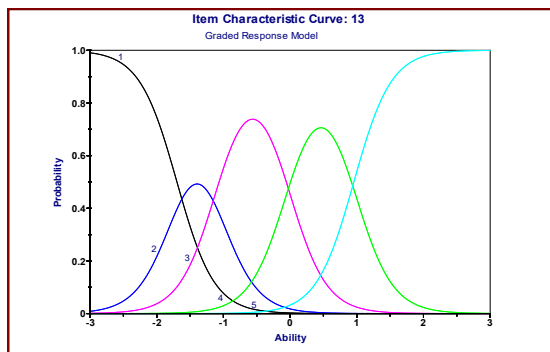
2.2 ผลการวิเคราะห์คุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ความยาว 25 ข้อ

เมื่อพิจารณาค่าพารามิเตอร์ Threshold ของแต่ละรายการคำตอบ (β_i) พบว่า β_1 มีค่าอยู่ระหว่าง -9.28 ถึง -1.52 β_2 มีค่าอยู่ระหว่าง -5.90 ถึง -0.84 β_3 มีค่าอยู่ระหว่าง -0.47 ถึง 4.70 และ β_4 มีค่าอยู่ระหว่าง 0.41 ถึง 14.78 โดยข้อคำถามทุกข้อ มีค่า $\beta_1 < \beta_2 < \beta_3 < \beta_4$ อาจกล่าวได้ว่าผู้ที่มีความสามารถสูงมีโอกาเลือกตอบรายการคำตอบระดับ 5 มากกว่า รายการคำตอบระดับ 1, 2, 3, และ 4

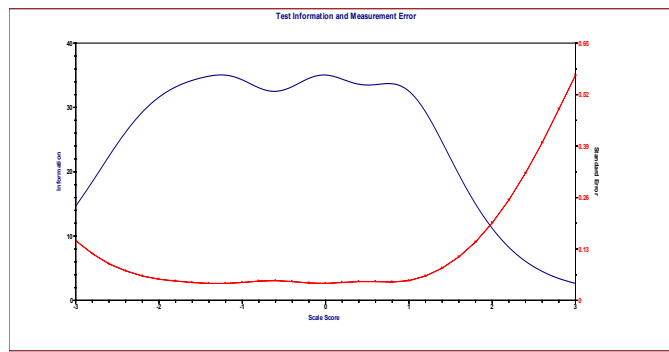
เมื่อพิจารณาค่าพารามิเตอร์ความชันร่วม (α) มีค่าอยู่ระหว่าง 0.16 ถึง 3.56 โดยข้อคำถามที่มีค่าพารามิเตอร์ความชันร่วมสูงสุด คือ ข้อที่ 13 มีค่าเท่ากับ 3.56 และข้อคำถามที่มีค่าพารามิเตอร์ต่ำสุด คือ ข้อที่ 16 มีค่าเท่ากับ 0.16 อาจกล่าวได้ว่า ข้อคำถามที่มีค่าพารามิเตอร์ความชันร่วมสูงกว่าข้ออื่น แสดงว่ามีค่าอำนาจจำแนกสูงกว่าข้ออื่น

เมื่อพิจารณาค่าสารสนเทศรายข้อ (IIF) พบว่า แบบวัดสามารถประมาณค่าความสามารถของผู้ตอบได้อย่างแม่นยำในระดับต่างกันออกไป ตามระดับค่าฟังก์ชันสารสนเทศของข้อสอบ โดยผู้วิจัยจะนำเสนอสารสนเทศของข้อคำถามที่มีค่าสูงที่สุด ได้แก่ ข้อที่ 13 มีค่า IIF เท่ากับ 3.592 ณ ระดับความสามารถที่ -1.4 ซึ่งเป็นข้อคำถามในองค์ประกอบด้านการประเมินตนเองได้ถูกต้อง พฤติกรรมบ่งชี้ที่ 1.4 ระบุอาชีพที่คาดว่าตนเองจะสามารถทำได้ในอนาคตได้โดยใช้วิธีการที่หลากหลายในการสำรวจตนเอง ดังภาพ 2

เมื่อพิจารณาค่าสารสนเทศของแบบวัด (TIF) พบว่า แบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ความยาว 25 ข้อ ให้สารสนเทศสูงสุดที่ความสามารถของผู้ตอบเท่ากับ 0.0 โดยมีค่าสารสนเทศเท่ากับ 35.021 แสดงว่าแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ความยาว 25 ข้อ มีความแม่นยำในการประมาณค่าสูงสุด เมื่อนำไปใช้กับผู้ตอบที่มีการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล อยู่ในระดับปานกลาง ดังภาพ 3



ภาพ 2 โค้งลักษณะแบบวัด (ICC) ข้อที่ 13



ภาพ 3 โค้งสารสนเทศของแบบวัด ความยาว 25 ข้อ

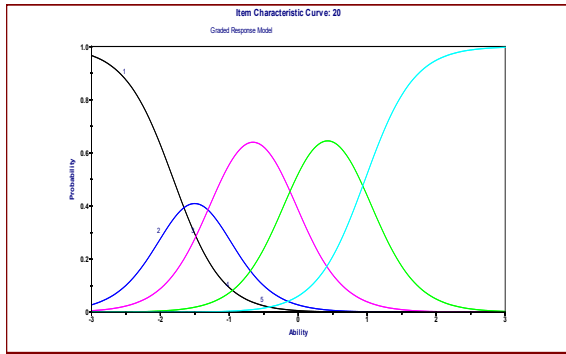
2.3 ผลการวิเคราะห์คุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเอง ในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ความยาว 35 ข้อ

เมื่อพิจารณาค่าพารามิเตอร์ Threshold ของแต่ละรายการคำตอบ (β_i) พบว่า β_1 มีค่าอยู่ระหว่าง -8.72 ถึง -1.59 β_2 มีค่าอยู่ระหว่าง -3.74 ถึง -0.77 β_3 มีค่าอยู่ระหว่าง -0.63 ถึง 7.94 และ β_4 มีค่าอยู่ระหว่าง 0.14 ถึง 14.07 โดยข้อคำถามทุกข้อ มีค่า $\beta_1 < \beta_2 < \beta_3 < \beta_4$ อาจกล่าวได้ว่าผู้ที่ความสามารถสูงมีโอกาเลือกตอบรายการคำตอบระดับ 5 มากกว่า รายการคำตอบระดับ 1, 2, 3, และ 4

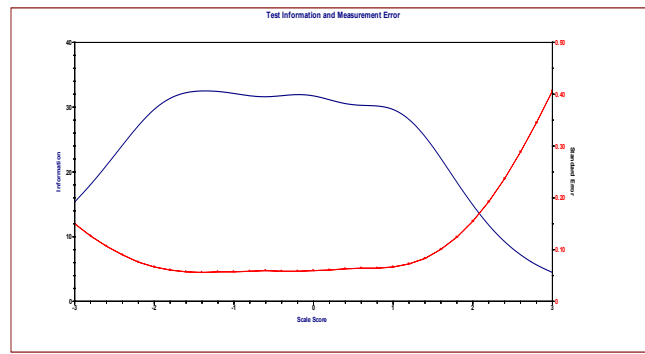
เมื่อพิจารณาค่าพารามิเตอร์ความชันร่วม (α) มีค่าอยู่ระหว่าง 0.16 ถึง 2.82 โดยข้อคำถามที่มีค่าพารามิเตอร์ความชันร่วมสูงสุด คือ ข้อ 20 มีค่าเท่ากับ 2.82 และข้อคำถามที่มีค่าพารามิเตอร์ต่ำสุด คือ ข้อ 23 มีค่าเท่ากับ 0.16 อาจกล่าวได้ว่า ข้อคำถามที่มีค่าพารามิเตอร์ความชันร่วมสูงกว่าข้ออื่น แสดงว่ามีค่าอำนาจจำแนกสูงกว่าข้ออื่น

เมื่อพิจารณาค่าสารสนเทศรายข้อ (IIF) พบว่า แบบวัดสามารถประมาณค่าความสามารถของผู้ตอบได้อย่างแม่นยำในระดับต่างกันออกไป ตามระดับค่าฟังก์ชันสารสนเทศของข้อสอบ โดยผู้วิจัยจะนำเสนอสารสนเทศของข้อคำถามที่มีค่าสูงที่สุด ได้แก่ ข้อที่ 20 มีค่า IIF เท่ากับ 2.338 ณ ระดับความสามารถที่ -1.4 ($\theta = -1.4$) ซึ่งเป็นข้อคำถามในองค์ประกอบด้านการรวบรวมข้อมูลเกี่ยวกับอาชีพ พฤติกรรมบ่งชี้ที่ 2.2 ค้นคว้าข้อมูลเกี่ยวกับแนวทางในการศึกษาต่อที่สอดคล้องกับความสนใจและความสามารถของตนเองได้ โดยสืบค้นจากแหล่งข้อมูลที่หลากหลายมีความถูกต้องและน่าเชื่อถือ ดังภาพ 4

เมื่อพิจารณาค่าสารสนเทศของแบบวัด (TIF) พบว่า แบบวัดการรับรู้ความสามารถของตนเอง ในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ความยาว 35 ข้อ ให้สารสนเทศสูงสุดที่ความสามารถของผู้ตอบเท่ากับ -1.4 โดยมีค่าสารสนเทศเท่ากับ 32.492 แสดงว่าแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ความยาว 35 ข้อ มีความแม่นยำในการประมาณค่าสูงสุด เมื่อนำไปใช้กับผู้ที่ตอบที่มีการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล อยู่ในระดับปานกลาง ดังภาพ 5



ภาพ 4 โค้งลักษณะแบบวัด (ICC) ข้อที่ 20



ภาพ 5 โค้งสารสนเทศของแบบวัด ความยาว 35 ข้อ

2.4 ผลการเปรียบเทียบคุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลายที่มีความยาวต่างกัน โดยประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น (Graded Response Model : GRM) เมื่อพิจารณาแบบวัดที่มีความยาว 25 ข้อ พบว่า ค่าเฉลี่ยความชันร่วม เท่ากับ 2.088 ค่าเฉลี่ย Threshold ของแต่ละรายการคำตอบ (β_i) พบว่า β_1 มีค่าเฉลี่ยเท่ากับ -2.944 β_2 มีค่าเฉลี่ยเท่ากับ -1.526 β_3 มีค่าเฉลี่ยเท่ากับ 0.171 และ β_4 มีค่าเฉลี่ยเท่ากับ 1.787 ค่าสารสนเทศของแบบวัด (TIF) มีค่าเท่ากับ 35.021 ที่ความสามารถ (θ) ระดับ 0.0 และค่าความเที่ยง เท่ากับ 0.965 และเมื่อพิจารณาแบบวัดที่มีความยาว 35 ข้อ พบว่า ค่าเฉลี่ยความชันร่วม เท่ากับ 1.655 ค่าเฉลี่ย Threshold ของแต่ละรายการคำตอบ (β_i) พบว่า β_1 มีค่าเฉลี่ยเท่ากับ -2.921 β_2 มีค่าเฉลี่ยเท่ากับ -4.733 β_3 มีค่าเฉลี่ยเท่ากับ -0.020 และ β_4 มีค่าเฉลี่ยเท่ากับ 1.526 ค่าสารสนเทศของแบบวัด (TIF) มีค่าเท่ากับ 32.492 ที่ความสามารถ (θ) ระดับ -1.4 และค่าความเที่ยงเท่ากับ 0.965 ดังตาราง 2

ตาราง 3 ผลการเปรียบเทียบคุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลายที่มีความยาวต่างกัน โดยประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น (Graded Response Model : GRM)

ความยาว ของแบบวัด	ค่าเฉลี่ย ความชันร่วม(α)	ค่าเฉลี่ย Threshold				TIF	ความเที่ยง
		β_1	β_2	β_3	β_4		
25 ข้อ	2.088	-2.944	-1.526	0.171	1.787	35.021 ($\theta = 0.0$)	0.966
35 ข้อ	1.655	-2.921	-4.733	-0.020	1.526	32.492 ($\theta = -1.4$)	0.965

หมายเหตุ ค่าเฉลี่ยเทอร์ชโฮลด์ (Threshold) ในตารางที่ 3 เป็นเพียงการนำเสนอภาพรวมของค่าพารามิเตอร์เทอร์ชโฮลด์ (Threshold) เพื่อให้เห็นแนวโน้มว่า $\beta_1 < \beta_2 < \beta_3 < \beta_4$ ทั้ง 2 ฉบับ และเมื่อพิจารณาค่าพารามิเตอร์เทอร์ชโฮลด์ (Threshold) แยกเป็นรายข้อและรายฉบับก็มีแนวโน้ม $\beta_1 < \beta_2 < \beta_3 < \beta_4$ ซึ่งเรียงจากน้อยไปมากทุกข้อทั้ง 2 ฉบับเช่นกัน

อภิปรายผล

1. อภิปรายผลการสร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ที่มีความยาวแตกต่างกัน

สามารถสร้างแบบวัดได้ครอบคลุมตามองค์ประกอบและแผนผังข้อแบบวัด (test blueprint) ที่เป็นเช่นนี้ เนื่องจากก่อนการจัดชุดของแบบวัดผู้วิจัยได้วางแผนการสร้างแบบวัดให้มีความสอดคล้องตามองค์ประกอบของการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล และดำเนินการสร้างข้อคำถามตามนิยามและองค์ประกอบได้ครอบคลุม รวมถึงภายหลังจากการวิเคราะห์ความตรงเชิงเนื้อหาพบว่า ผู้ทรงคุณวุฒิได้ให้ข้อเสนอแนะเพื่อแก้ไขปรับปรุงข้อคำถามให้ถูกต้องเหมาะสมมากยิ่งขึ้น โดยภายหลังจากการปรับปรุง มีการนำไปทดลองใช้ พบว่า ข้อคำถามในแบบวัด 35 ข้อ มีคุณภาพทุกข้อ เนื่องจากข้อคำถามมีค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมผ่านเกณฑ์ทุกข้อ และคัดเลือกข้อคำถาม 25 ข้อ โดยพิจารณาคัดเลือกจากข้อคำถามในแต่ละชุดที่มีค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม สูงที่สุดตามลำดับ ซึ่งแม้ว่าจะตัดรายการข้อคำถามบางข้อออกไปก็ยังคงเนื้อหาที่ครอบคลุมตามองค์ประกอบอยู่ จึงส่งผลให้ไม่เกิดความซ้ำซ้อนในขณะที่จัดชุดแบบวัด และสามารถวัดองค์ประกอบการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลได้อย่างครอบคลุม ซึ่งสอดคล้องกับ Thirakanan, (2008) ที่กล่าวว่า เครื่องมือที่มีคุณภาพเป็นเครื่องมือเก็บรวบรวมข้อมูลที่นักวิจัยต้องการ ซึ่งต้องมีการกำหนดขั้นตอนการสร้างเครื่องมือเพื่อลดความคลาดเคลื่อนให้มากที่สุด โดยเริ่มจากการศึกษาแนวคิด ทฤษฎี เอกสารที่เกี่ยวข้อง กำหนดนิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการ พฤติกรรมบ่งชี้ เพื่อเป็นแนวทางในการสร้างแผนผังของแบบวัด (Test Blueprint) และสร้างข้อคำถามเกี่ยวกับการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ซึ่งส่งผลให้แบบวัด มีความตรงและความเที่ยงที่ผ่านเกณฑ์

2. อภิปรายผลการเปรียบเทียบคุณสมบัติทางจิตมิติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ที่มีความยาวต่างกัน โดยประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น

จากการตรวจสอบข้อตกลงเบื้องต้นของ IRT ที่พบว่าแบบวัดทั้ง 2 ฉบับมีอัตราส่วนค่าไอเกนสูงผ่านเกณฑ์การตรวจสอบข้อตกลงเบื้องต้น ซึ่งสอดคล้องกับ Kanjanawasee, (2020) ที่กล่าวว่า การวิเคราะห์องค์ประกอบ (factor analysis) เพื่อคำนวณค่าไอเกน (Eigen value) สำหรับอัตราส่วนระหว่างค่าไอเกนของตัวประกอบแรกกับตัวประกอบถัดไป ถ้ามีอัตราส่วนที่สูงแสดงถึงเครื่องมือหรือแบบสอบวัดคุณลักษณะเด่นเดียว (single dominant factor) ดังนั้น จึงสะท้อนให้เห็นว่าแบบวัดทั้ง 2 ฉบับนี้ มีความเป็นเอกมิติ (unidimensional)

จากผลการวิจัยที่พบว่า แบบวัดความยาว 25 ข้อ มีค่าเฉลี่ยความชันร่วม เท่ากับ 2.088 ค่าสารสนเทศของแบบวัด (TIF) มีค่าเท่ากับ 35.021 ที่ความสามารถ (θ) ระดับ 0.0 และค่าความเที่ยง เท่ากับ 0.966 ขณะที่แบบวัดความยาว 35 ข้อ มีค่าเฉลี่ยความชันร่วม เท่ากับ 1.655 ค่าสารสนเทศของแบบวัด (TIF) มีค่าเท่ากับ 32.492 ที่ความสามารถ (θ) ระดับ -1.4 และค่าความเที่ยง เท่ากับ 0.965 ที่เป็นเช่นนี้เนื่องจากแบบวัดที่มีความยาว 25 ข้อ ประกอบไปด้วยข้อคำถามที่มีคุณสมบัติทางจิตมิติผ่านเกณฑ์เป็นส่วนใหญ่

ซึ่งมีเพียงข้อคำถามที่ 16 ที่พบว่ามีค่าความชันร่วม (α) ต่ำมาก จำนวน 1 ข้อ ได้แก่ “ฉันบอกได้ว่าแนวทางการศึกษาต่อเกี่ยวกับอาชีพที่ตนเองสนใจเป็นอย่างไร โดยไม่จำเป็นต้องสอบถาม/ค้นคว้าข้อมูลเพิ่มเติมจากแหล่งอื่น (-)” ในขณะที่แบบวัดที่มีความยาวจำนวน 35 ข้อ มีข้อคำถามที่มีความชันร่วม (α) ต่ำมาก จำนวน 1 ข้อเช่นกัน ได้แก่ ข้อที่ 23 ซึ่งเป็นข้อคำถามที่ตรงกันกับในแบบวัดที่มีความยาว 25 ข้อ ดังนั้นจึงส่งผลให้ในภาพรวมค่าเฉลี่ยความชันร่วมของแบบวัด 25 ข้อ มีค่าสูงกว่า ตลอดจนให้สารสนเทศสูงกว่าแบบวัดที่มี 35 ข้อ ซึ่งลักษณะของค่าความชันร่วมที่ต่ำกว่าบ่งบอกถึงผู้ตอบหลายคนเลือกคำตอบแบบสุดขั้ว (Kanjana-wasee, 2013) และเมื่อค่าความชันร่วมมีค่าสูง ก็ย่อมส่งผลต่อคุณภาพของแบบวัดทั้งฉบับ ซึ่งพบว่าแบบวัดที่มีความยาว 25 ข้อ มีค่าความเที่ยงที่สูงกว่า 35 ข้อ ซึ่งสอดคล้องกับ Winingsih et al. (2023) ที่ได้ศึกษาเกี่ยวกับการตรวจสอบคุณภาพของแบบวัดทางจิตวิทยาโดยตรวจสอบความตรงและความเที่ยงของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพ แบบสั้น (CDSE-SF) ของ (Betz et al., 1996) จำนวน 25 ข้อ พบว่า น้ำหนักองค์ประกอบมีค่ามากกว่า 0.5 ค่าความเที่ยงของแต่ละองค์ประกอบมีค่ามากกว่า 0.85 ซึ่งบ่งชี้ว่าแบบวัดนี้มีความน่าเชื่อถือ และยังสอดคล้องกับ Montgomery (2023) ที่ได้ศึกษาการใช้ทฤษฎีการตอบสนองข้อสอบเพื่อพัฒนาแบบวัดสิ่งแวดล้อมการปฏิบัติงานของพยาบาล ผลการวิจัย แบบวัดสิ่งแวดล้อมการปฏิบัติงานของพยาบาล ฉบับ 20 ข้อ มีความตรงและเหมาะสมมากกว่าฉบับ 31 ข้อ ซึ่งช่วยลดภาระเวลาในการเก็บข้อมูล แต่ยังคงรักษาองค์ประกอบได้เหมือนกับแบบวัดต้นฉบับ ดังนั้น จากผลการวิจัยครั้งนี้ แบบวัดความยาว 25 ข้อ จึงมีความเหมาะสมที่สุดกับการนำไปใช้วัดระดับการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนชั้นมัธยมศึกษาตอนปลาย เนื่องจากมีความน่าเชื่อถือและไม่ทำให้เกิดความเหนื่อยล้าในขณะที่ทำแบบวัด

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 จากผลการวิจัยที่พบว่า แบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ประกอบด้วย 5 องค์ประกอบ ได้แก่ (1) การประเมินตนเองได้ถูกต้อง (2) การรวบรวมข้อมูลเกี่ยวกับอาชีพ (3) การเลือกเป้าหมาย (4) การวางแผนในอนาคต (5) การแก้ปัญหา ดังนั้นในการนำผลวิจัยไปใช้ นักเรียน ครูผู้สอน ตลอดจนผู้ที่เกี่ยวข้องสามารถนำองค์ประกอบเหล่านี้ไปใช้ในการพัฒนาและส่งเสริมรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลได้ เช่น ครูผู้สอนจัดกิจกรรมที่ส่งเสริมให้นักเรียนได้ศึกษาค้นคว้าข้อมูลเกี่ยวกับอาชีพที่ตนเองสนใจจากแหล่งข้อมูลที่น่าเชื่อถือ นักเรียนสังเกตความสามารถ ความถนัดของตนเองผ่านการทำกิจกรรมหรือสิ่งที่ชอบทำเป็นกิจวัตรประจำวัน เป็นต้น

1.2 จากผลการวิจัยที่พบว่า แบบวัดที่มีความยาว 25 ข้อ มีค่าความชันร่วมเฉลี่ย ค่าสารสนเทศแบบสอบ (TIF) และความเที่ยง สูงกว่าความยาว 35 ข้อ ดังนั้นในการนำผลวิจัยไปใช้ ครูผู้สอน ตลอดจนผู้เกี่ยวข้องสามารถนำแบบวัดฉบับ 25 ข้อ ซึ่งเป็นแบบวัดฉบับสั้นไปใช้ในการประเมินการรับรู้ความสามารถ

ของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของผู้เรียนได้ ซึ่งเป็นการประหยัดเวลาและลดความล่าช้าในการประเมินตนเองได้ ไม่แตกต่างจาก ฉบับ 35 ข้อ

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ในการวิจัยครั้งต่อไปควรสร้างแบบวัดในลักษณะอื่น ๆ เช่น แบบวัดเชิงสถานการณ์ (situational test) ซึ่งช่วยในการกระตุ้นให้ผู้ตอบสามารถแสดงความคิดความรู้สึกออกมาอย่างเด่นชัด และเหมาะสมหากจะนำไปสร้างเพื่อใช้ประเมินผู้เรียนในระดับมัธยมศึกษาตอนปลาย

2.2 ในการวิจัยครั้งต่อไปควรมีการศึกษาหรือเปรียบเทียบคุณสมบัติของแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัล ของนักเรียนชั้นมัธยมศึกษาตอนปลาย ที่มีการให้คะแนนแบบ 2 ค่า (dichotomous) และแบบหลายค่า (polytomous)

2.3 ในการวิจัยครั้งต่อไปควรศึกษาความตรงเชิงโครงสร้างของแบบวัดทั้งก่อนและหลังการตัดรายการข้อคำถามออก

กิตติกรรมประกาศ

บทความวิจัยเรื่อง การสร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนชั้นมัธยมศึกษาตอนปลาย : การประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น เป็นส่วนหนึ่งของวิทยานิพนธ์ระดับปริญญาโทเรื่อง การสร้างแบบวัดการรับรู้ความสามารถของตนเองในการตัดสินใจเลือกอาชีพในยุคดิจิทัลของนักเรียนชั้นมัธยมศึกษาตอนปลาย : การประยุกต์ใช้โมเดลการให้คะแนนแบบลำดับขั้น ซึ่งเป็นส่วนหนึ่งของการสำเร็จการศึกษาในหลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาการประเมินและการวิจัยการศึกษา มหาวิทยาลัยรามคำแหง โดยมี ผู้ช่วยศาสตราจารย์ ดร. จุฑามรณ มาสันเทียะ เป็นที่ปรึกษาวิทยานิพนธ์หลัก และอาจารย์ ดร. ดารุณี ทิพยกุลไพโรจน์ เป็นที่ปรึกษาร่วม

References

- Am, M. A., Setiawati, F. A., Hadi, S., & Istiyono, E. (2023). Psychometric properties career of commitment instrument using classical test theory and graded response model. *Journal of Pedagogical Sociology and Psychology*, 5(2), 26-40.
- Betz, N. E., Klein, K. L., & Taylor, K. M. (1996). Evaluation of a short form of the Career Decision-Making Self-Efficacy Scale. *Journal of Career Assessment*, 4(1), 47-57. <https://doi.org/10.1177/106907279600400103>
- Cahyawulan, W., & Yundianto, D. (2022). Career Decision Self-Efficacy Scale-Short Form with Indonesian University Graduates: A Rasch Model Approach. *Indonesian Journal of Learning Education and Counseling*, 4(2), 113-124. <https://doi.org/10.31960/ijolec.v4i2.1527>

- Herwin, H., Nurhayati, R., Lidyasari, A. T., & da Costa, A. (2023). Graded Response Models on the Curiosity Measurement of Elementary School Students. *International Journal of Educational Methodology*, 9(1), 53-62. <https://doi.org/10.12973/ijem.9.1.53>
- Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45(1), 79–122. <https://doi.org/10.1006/jvbe.1994.1027>
- Liu, F., Zhang, Z., Lin, B., Ping, Z., & Mei, Y. (2022). Assessing the psychometric properties of the Chinese return-to-work self-efficacy questionnaire using Rasch model analysis. *Health and Quality of Life Outcomes*, 20(1), 27. <https://doi.org/10.1186/s12955-022-01929-7>
- Miguel, J. P., Silva, J. T., & Prieto, G. (2013). Career Decision Self-Efficacy Scale — Short Form: A Rasch analysis of the Portuguese version. *Journal of Vocational Behavior*, 82(2), 116–123. <https://doi.org/10.1016/j.jvb.2012.12.001>
- Molek-Winiarska, D., & Zięba, A. (2019). The use of the graded response model (GRM) in the process of the assessment of the psychological content of work among employees. *Argumenta Oeconomica*, 42(1), 295–316. <https://doi.org/10.15611/aoe.2019.1.12>
- Montgomery, A. P., Campbell, C. M., Azuero, A., Swiger, P. A., & Patrician, P. A. (2023). Using item response theory to develop a shortened practice environment scale of the nursing work index. *Research in Nursing & Health*, 46(4), 400-410. <https://doi.org/10.1002/nur.22324>
- Nilsson, J. E., Schmidt, C. K., & Meek, W. D. (2002). Reliability Generalization: An Examination of the Career Decision-Making Self-Efficacy Scale. *Educational and Psychological Measurement*, 62(4), 647-658. <https://doi.org/10.1177/0013164402062004007>
- Reise, S., & Yu, J. (1990). Parameter recovery in the graded response model using MULTILOG. *Journal of Educational Measurement*, 27(2), 133–144. <https://doi.org/10.1111/j.1745-3984.1990.tb00738.x>
- Taylor, M. K., & Betz N. E. (1983). Applications of Self-Efficacy Theory to the Understanding and Treatment of Career Indecision. *Journal of Vocational Behavior*, 22(1), 63 - 81. [https://doi.org/10.1016/0001-8791\(83\)90006-4](https://doi.org/10.1016/0001-8791(83)90006-4)
- Winingsih, E., Hambali, I., & Hidayah, N. (2023). Psychological measurement of the validity and reliability of the short form career decision self-efficacy scale. *Journal for Re Attach Therapy and Developmental Diversities*, 6(5), 449-456. <https://www.jrtdd.com/index.php/journal/article/view/588/419>
- Yamane, T. (1973). *Statistics: An introductory analysis* (3rd ed.). Harper & Row.

Yaşar, M. (2019). Development of a "Perceived Stress Scale" based on classical test theory and graded response model. *International Journal of Assessment Tools in Education*, 6(4), 522–538. <https://doi.org/10.21449/ijate.626053>

Translate Thai References

- Department of Employment. (2016). *Handbook of preparing before to labour market*. Department of Employment. <https://www.doe.go.th/vgnew> (in Thai)
- Excellence in Science Education Bureau. (2024). *Educational achievement of Princess Chulabhorn Science school*. http://www.eseb.obec.go.th/?page_id=2260 (in Thai)
- Future Tales LAB, National Innovation Agency, and Innovation Foresight Institute. (2021). *Future of Work*. Tonklar publication. <https://ifi.nia.or.th/wp-content/uploads/2022/01/NIA-x-FTL-2021-Future-of-Work.pdf> (in Thai)
- Kanjanawasee, S. (2013). *Classical test theory* (7th ed.). Chulalongkorn University Press. (in Thai)
- Kanjanawasee, S. (2020). *Modern test theories* (5th ed.). Chulalongkorn University Press. (in Thai)
- Kassakul, B. (2020). *A construction of employability skills scale for lower secondary school students* [Master's thesis]. Srinakharinwirot University. <http://ir-thesis.swu.ac.th/dspace/bitstream/123456789/1297/1/g601130127.pdf> (in Thai)
- Ministry of Digital Economy and Society (2023). *Ministry of digital economy and society action plan for fiscal year 2023 (6-month revised plan)*. <https://www.mdes.go.th/law/download/6980> (in Thai)
- Office of the Basic Education Commission. (2012). *Guidelines for organizing guidance activities according to the Basic Education Core Curriculum, B.E. 200*. Agricultural Co-operative Federation of Thailand, Ltd. http://www.esbuy.net/site/download-file.php?doc_id=4518 (in Thai)
- Rungrotrangsarn, K. (2017). *A development of the sportsmanship scale of athletes in the institute of physical education* [Doctoral dissertation]. Prince of Songkla University. <https://kb.psu.ac.th/psukb/handle/2016/11747> (in Thai)
- Suwannal, P., Sopa, S., Suwannachan, P. C., & Katapanyo, P. N. (2022). How to deal with generation Y workers not changing work. *Journal of MCU Nakhondhat*, 9(9), 228–241. <https://so03.tci-thaijo.org/index.php/JMND/article/view/263887> (in Thai)
- Thirakanan, S. (2008). *Construction of measurement instruments for social science research: A guideline to practice* (2nd ed.). Chulalongkorn University Press. (in Thai)
- Thirakanan, S. (2014). *Research methodology in social science: A guideline to practice* (12th ed.). Chulalongkorn University Press. (in Thai)

การพัฒนาแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส
สำหรับนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย

Development of a Scientific Explanation Test on the topic of Gases
for Upper Secondary School Students

จักรพงษ์ ปินชัย^{1*} และ ณวรา สีสี่²

Jakkrapong Pinchai^{1*} and Navara Seetee²

(Received: September 5, 2024; Revised: October 5, 2024; Accepted: October 8, 2024)

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส สำหรับนักเรียนระดับชั้นมัธยมศึกษาตอนปลายให้มีคุณภาพ มีขั้นตอนในการพัฒนาแบบวัด ดังนี้ 1) กำหนดนิยามเชิงปฏิบัติการของการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส 2) กำหนดโครงสร้างของแบบวัดและรายละเอียดของข้อคำถาม 3) สร้างข้อคำถามและเกณฑ์การให้คะแนนแบบจำเพาะ 4) ประเมินความเที่ยงตรงและความเหมาะสมต่อการนำไปใช้ของแบบวัด 5) ปรับปรุงแบบวัดแล้วนำไปทดลองใช้ 6) หาคุณภาพในด้านความยากง่าย อำนาจจำแนก ความเชื่อมั่น และคุณภาพของเกณฑ์การให้คะแนน ผลการวิจัย พบว่า แบบวัดอัตนัยแบบความเรียงที่ไม่จำกัดคำตอบที่สร้างขึ้นมีความเที่ยงตรงเชิงเนื้อหาและมีความเหมาะสมต่อการนำไปใช้ เมื่อนำไปทดลองใช้กับกลุ่มตัวอย่างซึ่งเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 5 ที่เรียน เรื่อง แก๊ส มาแล้ว แต่ยังไม่เคยฝึกการสร้างคำอธิบายทางวิทยาศาสตร์ จำนวน 41 คน จากโรงเรียนขนาดใหญ่พิเศษแห่งหนึ่งในกรุงเทพมหานคร พบว่า ข้อคำถามมีความยากง่ายและอำนาจจำแนกเหมาะสม ($P_i = .33 - .57$, $R_i = .22 - .42$) มีค่าความเชื่อมั่นอยู่ในระดับสูง (ค่าสัมประสิทธิ์แอลฟาของครอนบาค เท่ากับ .732) และเกณฑ์การให้คะแนนมีคุณภาพ ผู้ตรวจให้คะแนน ประเมินได้สอดคล้องกัน ($ICC = .862$) ดังนั้น แบบวัดที่สร้างขึ้นสามารถนำไปใช้ประเมินการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส ของนักเรียนได้อย่างมีประสิทธิภาพ

คำสำคัญ การสร้างคำอธิบายทางวิทยาศาสตร์ กรอบแนวคิด CER การประเมิน เคมี่

¹ นิสิตปริญญาโท สาขาวิทยาศาสตร์ศึกษา ศูนย์วิทยาศาสตร์ศึกษา คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

² อาจารย์ ศูนย์วิทยาศาสตร์ศึกษา คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

¹ Master Degree's Student, Science Education Program, Science Education Center, Faculty of Science, Srinakharinwirot University

² Lecturer, Science Education Center, Faculty of Science, Srinakharinwirot University

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

Abstract

This research aimed to develop a quality scientific explanation test on the topic of gases for upper secondary school students. The development of the test followed these steps: 1) defining the operational definition of scientific explanation in the context of gases, 2) creating a test blueprint and item specifications, 3) constructing test items and their corresponding analytic rubrics, 4) examining content validity and overall suitability, 5) finalizing the test and conducting a pilot study, and 6) evaluating the finalized test in terms of item difficulty, discrimination, reliability, and rubric quality. The results indicated that the extended-response, constructed-response test demonstrated appropriate content validity and alignment with the intended construct. Following a pilot study with 41 eleventh-grade students from a large public school in Bangkok—students who had studied the topic of gases but had not received explicit instruction in constructing scientific explanations—the test showed acceptable item difficulty ($P_i = .33-.57$) and discrimination indices ($R_i = .22-.42$). The test demonstrated satisfactory reliability (Cronbach's $\alpha = .732$). In addition, the analytic rubrics were of high quality, with strong inter-rater consistency ($ICC = .862$). Therefore, the developed test is valid and reliable for accurately assessing students' scientific explanation skills related to the topic of gases.

Keywords: scientific explanation, CER framework, assessment, chemistry

บทนำ

การจัดการเรียนรู้วิทยาศาสตร์มีเป้าหมายสำคัญเพื่อให้นักเรียนเข้าใจหลักการ ทฤษฎี และกฎที่เป็นพื้นฐานของธรรมชาติ (OBEC, 2017) การสร้างคำอธิบายทางวิทยาศาสตร์ (construction of scientific explanation) เป็นพฤติกรรมหนึ่งที่สามารถสะท้อนความเข้าใจในองค์ความรู้ทางวิทยาศาสตร์ของนักเรียนได้ (Promyod, 2019) ช่วยให้ครูวินิจฉัยได้ว่านักเรียนมีความเข้าใจในหลักการ ทฤษฎี และกฎที่เป็นพื้นฐานของวิชาวิทยาศาสตร์มากน้อยเพียงใด ดังเช่นในมาตรฐานการเรียนรู้วิทยาศาสตร์ยุคใหม่ของประเทศสหรัฐอเมริกา (Next Generation Science Standards; NGSS) ได้กำหนดให้การสร้างคำอธิบายทางวิทยาศาสตร์เป็นหนึ่งในแนวปฏิบัติทางวิทยาศาสตร์และวิศวกรรม (science and engineering practices) โดยกำหนดไว้อย่างชัดเจนว่า นักเรียนจะต้องสร้างคำอธิบายเพื่ออธิบายปรากฏการณ์ที่เกิดขึ้นโดยประยุกต์ใช้ความรู้ที่ได้เรียนมา ซึ่งจะมีระดับความซับซ้อนที่แตกต่างกันไปตามระดับชั้น (National Research Council, 2012) นอกจากนี้ การสร้างคำอธิบายทางวิทยาศาสตร์ยังเป็นหนึ่งในสมรรถนะของการเป็นผู้มีความฉลาดรู้ทางวิทยาศาสตร์ (scientific literacy) ของโครงการประเมินผลนักเรียนนานาชาติ (Program for International Student Assessment; PISA) (IPST, 2017)

การสร้างคำอธิบายทางวิทยาศาสตร์ หมายถึง การลงข้อสรุปหรือการให้คำตอบที่จำเพาะกับสถานการณ์หรือคำถามทางวิทยาศาสตร์ที่สนใจ โดยมีข้อมูลสนับสนุนที่น่าเชื่อถือ ภายใต้กรอบของหลักการหรือความรู้ทางวิทยาศาสตร์ที่จำเพาะกับสิ่งที่กำลังสนใจ อาจเป็นการบอกสาเหตุหรือทำนายผลที่จะเกิดขึ้น (McNeill & Krajcik, 2008; Tang, 2015; Yao & Guo, 2018) ประกอบด้วย 3 องค์ประกอบ (McNeill & Krajcik, 2008) ได้แก่ 1) ข้อกล่าวอ้าง (claim) หมายถึง ข้อสรุปหรือคำตอบของประเด็นทางวิทยาศาสตร์ที่สนใจศึกษา 2) หลักฐาน (evidence) เป็นข้อมูลเชิงประจักษ์ที่ถูกเลือกนำมาสนับสนุนข้อกล่าวอ้าง ต้องมีความเหมาะสม สอดคล้องกับสถานการณ์และมีจำนวนมากเพียงพอที่จะทำให้ผู้อื่นคล้อยตามและยอมรับได้ และ 3) การให้เหตุผล (reasoning) คือ การใช้หลักการหรือความรู้ทางวิทยาศาสตร์ที่จำเพาะกับสถานการณ์นั้น ๆ เชื่อมโยงหลักฐานให้เข้ากับข้อกล่าวอ้างอย่างสมเหตุสมผล เมื่อนำแต่ละองค์ประกอบมารวมกัน

การส่งเสริมให้นักเรียนระดับชั้นมัธยมศึกษาตอนปลายได้สร้างคำอธิบายทางวิทยาศาสตร์มีความจำเป็น เพราะการฝึกสร้างคำอธิบายทางวิทยาศาสตร์อย่างต่อเนื่อง จะช่วยให้นักเรียนพัฒนาตนเองจนกลายเป็นผู้มีความฉลาดรู้ทางวิทยาศาสตร์ ซึ่งเป็นคุณลักษณะที่สำคัญของพลเมืองไทยในยุคปัจจุบันและอนาคต (IPST, 2017) นักเรียนจึงควรได้ฝึกสร้างคำอธิบายทางวิทยาศาสตร์อย่างสม่ำเสมอผ่านการส่งเสริมและสนับสนุนของครู แต่การที่ครูจะส่งเสริมได้อย่างเหมาะสมนั้น ครูจำเป็นต้องทราบเกี่ยวกับการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียนก่อนว่ามีลักษณะเป็นอย่างไร จึงจะสามารถช่วยเหลือและสนับสนุนให้นักเรียนสร้างคำอธิบายทางวิทยาศาสตร์ที่มีความสมบูรณ์มากขึ้นได้ กระบวนการที่จะช่วยให้ครูทราบเกี่ยวกับลักษณะการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียนได้ คือ การวัดและประเมินการสร้างคำอธิบายทางวิทยาศาสตร์ ที่ผ่านมา มีการใช้เครื่องมือวัดและประเมินผลในลักษณะที่ต่างกันไป เช่น แบบทดสอบปรนัย สี่ตัวเลือกที่เป็นการถามแยกองค์ประกอบ (IPST, 2017) หรือแบบทดสอบอัตนัยแบบความเรียง (Aniwatanawong & Jantrasri, 2019; Hirio et al., 2020; Juthanaruepakij et al., 2018; McNeill & Krajcik, 2008; Tang, 2015; Yao & Guo, 2018) รูปแบบของเครื่องมือที่นิยมใช้มากที่สุด คือ แบบทดสอบอัตนัยแบบความเรียง (Singwee, 2020) เพราะเป็นแบบวัดที่นักเรียนได้ลงมือสร้างแต่ละองค์ประกอบของคำอธิบายทางวิทยาศาสตร์ด้วยตนเอง แบบวัดในลักษณะนี้จึงสะท้อนคุณลักษณะของนักเรียนอย่างแท้จริง ต่างจากแบบทดสอบปรนัย ที่มีตัวเลือกเป็นองค์ประกอบของคำอธิบายทางวิทยาศาสตร์ส่วนต่าง ๆ แล้วให้นักเรียนพิจารณาเลือกว่าคำตอบใดคือคำตอบที่เหมาะสมที่สุดเท่านั้น นอกจากนี้ แบบทดสอบอัตนัยยังแบ่งเป็นแบบทดสอบอัตนัยที่จำกัดคำตอบ คือ ระบุให้นักเรียนเขียนคำอธิบายทางวิทยาศาสตร์แยกตามแต่ละองค์ประกอบอย่างชัดเจน และแบบทดสอบอัตนัยที่ไม่จำกัดคำตอบ คือ ระบุให้นักเรียนเขียนคำอธิบายทางวิทยาศาสตร์เท่านั้น ไม่มีการแจกแจงองค์ประกอบ ซึ่งแบบทดสอบอัตนัยที่ไม่จำกัดคำตอบมีประสิทธิภาพในการประเมินสูงกว่าแบบทดสอบอัตนัยที่จำกัดคำตอบ (Wannathai & Pruekpramool, 2021)

การเรียนรู้ในวิชาเคมี เป็นการเรียนรู้เพื่ออธิบายเกี่ยวกับการเปลี่ยนแปลงของสารที่เกิดขึ้นในปรากฏการณ์ต่าง ๆ รอบตัว เช่น การนำไฟฟ้าของโลหะ การระเบิดของถังบรรจุแก๊ส การเปลี่ยนสีของเครื่องดื่ม การเรียนรู้ในวิชาดังกล่าวมีความเหมาะสมที่จะส่งเสริมให้นักเรียนได้สร้างคำอธิบายทางวิทยาศาสตร์ โดยเฉพาะอย่างยิ่งระดับชั้นมัธยมศึกษาปีที่ 5 ที่มีจำนวนผลการเรียนรู้เน้นให้นักเรียนต้อง “อธิบาย” มากที่สุด

ตัวอย่างเช่น อธิบายการประยุกต์ใช้ความรู้เกี่ยวกับสมบัติและกฎต่าง ๆ ของแก๊สในการอธิบายปรากฏการณ์หรือแก้ปัญหาในชีวิตประจำวันและในอุตสาหกรรม (OBEC, 2017)

จากการศึกษางานวิจัยที่ผ่านมา พบว่า มีแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ในเนื้อหาเคมีระดับชั้นมัธยมศึกษาปีที่ 5 ซึ่งส่วนใหญ่เป็นแบบทดสอบอัตนัยที่จำกัดคำตอบ เช่น อัตราการเกิดปฏิกิริยาเคมี (Nimdam et al., 2020) สมดุลเคมี (Tanang, 2019) สารละลายอิเล็กโทรไลต์และนอนอิเล็กโทรไลต์ (Aniwatanawong & Jantrasri, 2019) และเคมีไฟฟ้า (Supasatwong et al., 2020) แต่ยังไม่พบแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์แบบไม่จำกัดคำตอบในเนื้อหาเรื่อง แก๊ส

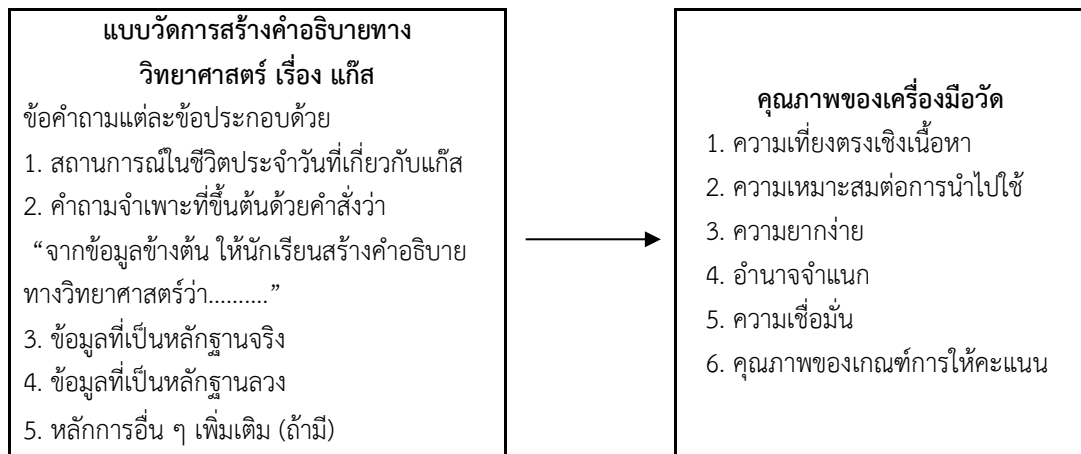
ดังนั้น การวิจัยนี้จึงมุ่งพัฒนาแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส สำหรับนักเรียนระดับชั้นมัธยมศึกษาตอนปลายให้มีคุณภาพ ในลักษณะของแบบทดสอบอัตนัยที่ไม่จำกัดคำตอบ ซึ่งจะเป็นประโยชน์ต่อครูผู้สอนเคมี สามารถนำไปใช้เป็นเครื่องมือเพื่อประเมินการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียนที่มีองค์ประกอบครบถ้วน และนำไปสู่การส่งเสริมการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียนในเรื่องนี้ได้อย่างเต็มความสามารถ

ความมุ่งหมายของการวิจัย

เพื่อพัฒนาและตรวจสอบคุณภาพแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์เรื่อง แก๊ส สำหรับนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย

กรอบแนวคิดการวิจัย

การสร้างคำอธิบายทางวิทยาศาสตร์ตามกรอบแนวคิดข้อกล่าวอ้าง-หลักฐาน-การให้เหตุผลของ McNeill & Krajcik (2008) นักเรียนจะต้องพิจารณาข้อมูลก่อนจะตัดสินใจสร้างข้อกล่าวอ้าง และเลือกใช้ข้อมูลที่พิจารณาแล้วมาเป็นหลักฐานมาสนับสนุนข้อกล่าวอ้าง โดยมีกฎหรือหลักการมาแสดงการเชื่อมโยงอย่างสมเหตุสมผลว่าหลักฐานนั้นสนับสนุนข้อกล่าวอ้าง เพื่อให้เครื่องมือที่พัฒนาขึ้นสามารถวัดการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส ของนักเรียนได้ครอบคลุมทุกองค์ประกอบอย่างมีประสิทธิภาพ ดังนั้นในการวิจัยนี้ ผู้วิจัยจึงดำเนินการสร้างแบบวัดโดยใช้ข้อคำถามที่มีสถานการณ์หลากหลายที่เกี่ยวข้องกับ แก๊ส มีข้อมูลประกอบทั้งที่เป็นหลักฐานจริง และหลักฐานลวง รวมทั้งให้ข้อมูลเพิ่มเติมเกี่ยวกับหลักการอื่น ๆ ที่ไม่ใช่กฎของแก๊ส ในกรณีที่เป็นสถานการณ์ที่มีความซับซ้อน เพื่อให้ให้นักเรียนได้พิจารณาเลือกใช้ แล้วสร้างเป็นคำอธิบายทางวิทยาศาสตร์ เมื่อสร้างข้อคำถามในแบบวัดได้ครบถ้วนตามรายละเอียดข้างต้น จะนำไปสู่การศึกษาคุณภาพของเครื่องมือวัดที่สร้างขึ้น ทำให้กรอบแนวคิดของการวิจัยนี้ แสดงดังภาพ 1



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

การวิจัยนี้เป็นการพัฒนาเครื่องมือประเมินการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียนในเนื้อหาเรื่อง แก๊ส ระดับชั้นมัธยมศึกษาตอนปลาย ซึ่งมีรายละเอียดดังนี้

1. ประชากรและกลุ่มตัวอย่าง

ประชากร คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 5 แผนการเรียนวิทยาศาสตร์-คณิตศาสตร์ จำนวน 266 คน ของโรงเรียนมัธยมสหศึกษาขนาดใหญ่พิเศษแห่งหนึ่ง ในจังหวัดกรุงเทพมหานคร

กลุ่มตัวอย่าง คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 5 แผนการเรียนวิทยาศาสตร์-คณิตศาสตร์ จำนวน 41 คน ที่เรียนเนื้อหาเรื่อง แก๊ส มาเรียบร้อยแล้ว ได้มาจากการเลือกตามความสะดวก (convenience sampling) เพราะโรงเรียนให้ความอนุเคราะห์ในการเข้าไปเก็บข้อมูลในห้องเรียนปกติ ที่มีการละความสามารถของผู้เรียน อย่างไรก็ตาม นักเรียนกลุ่มนี้ยังไม่เคยได้รับการฝึกหรือมีประสบการณ์เกี่ยวกับการสร้างคำอธิบายทางวิทยาศาสตร์มาก่อน

2. เครื่องมือที่ใช้ในการวิจัย

การวิจัยนี้มุ่งพัฒนาแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ให้มีคุณภาพ ซึ่งมีขั้นตอนในการพัฒนาเครื่องมือ ดังนี้

2.1 ระบุผลการเรียนรู้ที่ต้องการจะวัดและกำหนดพฤติกรรมการเรียนรู้ จากการวิเคราะห์ตัวชี้วัดและสาระการเรียนรู้แกนกลาง กลุ่มสาระการเรียนรู้วิทยาศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) จุดมุ่งหมายของการวิจัยนี้สอดคล้องกับผลการเรียนรู้ในสาระเคมี ข้อ 1 ระดับชั้นมัธยมศึกษาปีที่ 5 ข้อย่อย 6 คือ อธิบายการประยุกต์ใช้ความรู้เกี่ยวกับสมบัติและกฎต่าง ๆ ของแก๊สในการอธิบายปรากฏการณ์หรือแก้ปัญหาในชีวิตประจำวันและในอุตสาหกรรม พฤติกรรมที่ต้องการให้นักเรียนแสดงออก คือ เขียนคำอธิบายทางวิทยาศาสตร์ของสถานการณ์หรือปรากฏการณ์ในชีวิตประจำวันที่เกี่ยวข้องกับเรื่อง แก๊ส

2.2 กำหนดนิยามเชิงปฏิบัติการ จากการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องนำไปสู่การกำหนดนิยามเชิงปฏิบัติการของคำว่า การสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส หมายถึง การให้ข้อสรุปแก่ประเด็นที่เกี่ยวข้องกับปรากฏการณ์รอบตัว เรื่อง แก๊ส อย่างสมเหตุสมผลและน่าเชื่อถือ มีหลักฐานประกอบอย่างเหมาะสมและเพียงพอ สามารถวัดได้ด้วยแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์

2.3 กำหนดโครงสร้างและรายละเอียดของแบบวัด จากการสังเคราะห์งานวิจัยเกี่ยวกับการส่งเสริมการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียน แล้วพบว่าแบบทดสอบอัตนัยแบบความเรียงที่ไม่จำกัดคำตอบมีประสิทธิภาพในการประเมินสูงกว่า ในการวิจัยนี้จึงใช้แบบทดสอบอัตนัยดังกล่าว พร้อมกับเกณฑ์การให้คะแนนแบบแยกประเด็นตามองค์ประกอบของคำอธิบายทางวิทยาศาสตร์ ประกอบด้วยข้อคำถามทั้งหมด 4 ข้อ ใช้เวลาประมาณ 50 นาที (Singwee, 2020) คำถามที่ใช้มีสองลักษณะ คือ 1) ถามให้คาดคะเนหรือทำนาย และ 2) ถามหาความสัมพันธ์ ทุกข้อใช้ประโยคคำสั่งขึ้นต้นเหมือนกัน คือ “จากข้อมูลข้างต้น ให้นักเรียนสร้างคำอธิบายทางวิทยาศาสตร์ว่า.....” แล้วจึงระบุคำถามที่จำเพาะของข้อนั้น ๆ แบบวัดแต่ละข้อจะมีข้อมูลประกอบสถานการณ์ที่เป็นทั้งหลักฐานจริง และหลักฐานลวง เพื่อประเมินว่า นักเรียนสามารถใช้กฎของแก๊สในการพิจารณาเลือกหลักฐานที่เกี่ยวข้องได้อย่างสมเหตุสมผลหรือไม่ นอกจากนั้น หากสถานการณ์มีความซับซ้อน จำเป็นที่จะต้องใช้หลักการอื่น ๆ ที่ไม่ใช่กฎของแก๊สเพิ่มเติม จะระบุเข้าไปในข้อมูลประกอบสถานการณ์ด้วย เพื่อให้ นักเรียนมีข้อมูลเพียงพอที่จะสร้างคำอธิบายทางวิทยาศาสตร์ รายละเอียดของข้อคำถามแต่ละข้อ แสดงดังตาราง 1

ตาราง 1 โครงสร้างแบบวัดรายข้อ

สถานการณ์	คำถามจำเพาะ	หลักฐานจริง	หลักฐานลวง	กฎของแก๊ส	หลักการเพิ่มเติม
ถุงลมปอด	ถุงลมปอดหมายเลข 4 ที่บรรจุแก๊สไนโตรเจนในเทอร์ต 9.76 กรัม ควรมีเส้นรอบวงประมาณเท่าใด (คาดคะเน)	1. สมการการสลายตัวของแก๊สไนโตรเจนในเทอร์ต 2. เส้นรอบวงของถุงลมปอดที่เพิ่มขึ้น	1. มวลแก๊สไนโตรเจนในเทอร์ต 2. ปริมาตรของถุงลมปอดที่เพิ่มขึ้น	กฎของอาโวกาโดร	ไม่มี
เสื้อชูชีพ	เสื้อชูชีพพองได้ที่มีขนาดเล็กที่สุดควรมีกระบอกบรรจุแก๊สคาร์บอนไดออกไซด์หมายเลขใดเป็นส่วนประกอบบนตัวเสื้อ (ความสัมพันธ์)	1. มวลของแก๊สคาร์บอนไดออกไซด์ที่บรรจุไว้ในกระบอก 2. แรงดันของเสื้อชูชีพแต่ละขนาดสร้างได้	1. วัสดุที่ใช้ทำเสื้อชูชีพ 2. วัสดุที่เหมาะสมกับขนาดของเสื้อชูชีพ 3. ภาพแสดงส่วนประกอบของเสื้อชูชีพ	กฎของอาโวกาโดร	แรงดันของแก๊สแปรผันตรงกับขนาดของเสื้อชูชีพ
ฟองอากาศ	ฟองอากาศที่คัลเลนสังเกตเห็น ในบริเวณใกล้ผิวน้ำจะมีขนาดเป็นอย่างไร เมื่อเปรียบเทียบกับขนาดที่สังเกตได้ในตอนแรก ณ บริเวณใต้น้ำ (ทำนาย)	1. กราฟแสดงค่าความดันใต้น้ำที่เพิ่มขึ้นตามระดับความลึก 2. ภาพฟองอากาศที่มีขนาดเปลี่ยนแปลงตามระดับความลึก	อุณหภูมิของน้ำทะเล	กฎของบอยล์	ฟองอากาศจะปรับความดันให้เท่ากับภายนอกเพื่อให้คงรูปร่างอยู่ได้

ตาราง 1 (ต่อ)

สถานการณ์	คำถามจำเพาะ	หลักฐานจริง	หลักฐานลง	กฎของแก๊ส	หลักการเพิ่มเติม
ขวดน้ำดื่มแช่เย็น	ขวดน้ำดื่มที่วินนี้ดื่มไม่หมด แล้วนำไปแช่ตู้เย็น ที่มีอุณหภูมิ 4°C จะมีลักษณะที่เปลี่ยนแปลงไปอย่างไร (ทำนาย)	1. ปริมาตรของแก๊สไนโตรเจนที่ลดลงตามอุณหภูมิ 2. ขวดน้ำดื่มทำมาจาก Polyethylene terephthalate	องค์ประกอบของอากาศ	กฎของชาร์ล	Polyethylene terephthalate มีความยืดหยุ่น
แป้งพืชมะพร้าว	แป้งพืชมะพร้าวอุณหภูมิเท่าใด จะมีความนุ่มมากที่สุด (คาดคะเน)	1. ปริมาตรของแก๊สมีค่าเพิ่มขึ้นตามอุณหภูมิ 2. แก๊สมีปริมาตรสูงสุดที่ 270 °C	ภาพแสดงลักษณะของแป้งพืชมะพร้าวในช่วงเวลาต่าง ๆ	กฎของชาร์ล	ความนุ่มของพืชมะพร้าวขึ้นกับขนาดช่องว่างภายในเนื้อแป้ง
หลอดโซเดียมความดันต่ำ	ช่วงเวลาใด ที่หลอดโซเดียมความดันต่ำ จะมีความดันรวมของแก๊สภายในหลอด สูงที่สุด (คาดคะเน)	ชนิดของแก๊สที่พบภายในหลอด	1. สีของหลอดไฟ 2. ภาพหลอดไฟ ณ ช่วงเวลาต่าง ๆ 3. สมการแสดงการเปลี่ยนแปลงของโซเดียม	กฎความดันของดอลตัน	ไม่มี
แก๊สปิโตรเลียม	หลังปรับอุณหภูมิและความดัน ได้ตามที่ต้องการแล้ว แก๊สปิโตรเลียมชนิดใด จะเคลื่อนที่ไปถึงหอควบแน่นหมายเลข 1 เป็นลำดับสุดท้าย (ทำนาย)	มวลต่อโมลของแก๊ส บิวเทนมีค่าสูงที่สุด	1. สูตรโมเลกุลของแก๊สปิโตรเลียม 2. จุดเดือดของแก๊สปิโตรเลียม	กฎการแพร่ผ่านของเกรแฮม	ไม่มี

2.4 สร้างแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ตามโครงสร้างที่กำหนด พร้อมกับเกณฑ์การให้คะแนนแบบแยกประเด็นที่จำเพาะกับข้อคำถามนั้น ๆ ในหนึ่งเกณฑ์จะประเมินตามองค์ประกอบ 3 ด้าน คือ ข้อกล่าวอ้าง (C) หลักฐาน (E) และการให้เหตุผล (R) แต่ละด้านแบ่งเป็นการให้คะแนน 3 ระดับ คือ ได้ 2 คะแนน เมื่อสร้างองค์ประกอบนั้นได้ถูกต้องครบถ้วนสมบูรณ์ ได้ 1 คะแนน เมื่อสร้างได้ถูกต้องและสมบูรณ์บางส่วน และได้ 0 คะแนน เมื่อสร้างได้ไม่ถูกต้อง หรือ ไม่สร้างองค์ประกอบนั้นเลย ตัวอย่างข้อคำถามและเกณฑ์การให้คะแนน แสดงดังภาพ 2 และตาราง 2 ตามลำดับ

สถานการณ์ที่ 3 ฟองอากาศใต้ท้องทะเล



ภาพที่ 1 : เต่าหายใจ

ที่มา : <https://www.pinterest.com/pin/bubble-ring-blowing-at-its-finest-ig-user-charlottepiho-31666003621196500/>

เมื่อกลับมาถึงที่พัก คัลเลนก็ได้เล่าสิ่งที่ตนเองสังเกตเห็นให้กับจู้ดีฟัง จู้ดีอธิบายว่า ฟองอากาศจะคงรูปอยู่ได้ เมื่อความดันภายในฟองอากาศมีค่าเท่ากับความดันภายนอก การที่คัลเลนสังเกตเห็นว่าฟองอากาศไม่แตก แสดงว่าฟองอากาศมีการปรับความดันในขณะที่กำลังลอยขึ้นสู่ผิวน้ำ ซึ่งความดันบรรยากาศใต้น้ำที่ระดับความลึกต่างกัน จะมีการเปลี่ยนแปลงดังกราฟที่ 1 และจู้ดีได้เปิดภาพเต่าหายใจให้กับคัลเลน ดังภาพที่ 1 ในภาพจะสังเกตเห็นว่า ฟองอากาศที่ออกมาจากปากของเต่าจะมีขนาดเปลี่ยนแปลงไปจากเดิมเมื่อลอยขึ้นสู่ผิวน้ำ

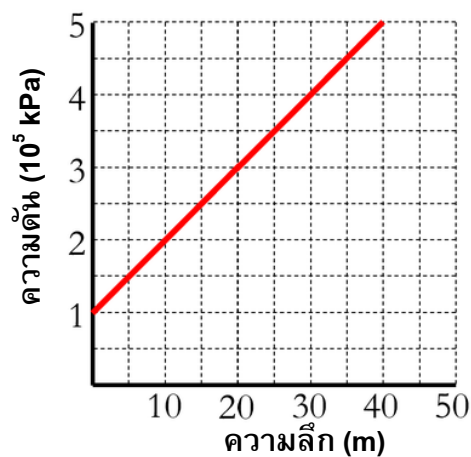
จากข้อมูลข้างต้น ให้นักเรียนสร้างคำอธิบายทางวิทยาศาสตร์ว่า ฟองอากาศที่คัลเลนสังเกตเห็นในบริเวณใกล้ผิวน้ำจะมีขนาดเป็นอย่างไร เมื่อเปรียบเทียบกับขนาดที่สังเกตได้ในตอนแรก ณ บริเวณใต้น้ำ

.....

.....

ในวันหยุดมาชมบูชาที่ผ่านมา คัลเลนกับพี่จ้องได้ไปดำน้ำแบบ snorkeling ที่เกาะแก้วพิสดาร เนื่องจากเป็นสถานที่ที่มีปะการังน้ำตื้นสวยงาม ติดอันดับโลก ในช่วงท้ายของการดำน้ำ คัลเลนสังเกตเห็นฟองอากาศลอยออกมาจากท่อหายใจของพี่จ้อง จึงว่ายน้ำตามฟองอากาศฟองเดิมจนถึงบริเวณผิวน้ำ โดยไม่ทำให้ฟองอากาศแตก ซึ่งคัลเลนพบว่าฟองอากาศมีขนาดต่างไปจากเดิม เขาเก็บความสงสัยนี้เอาไว้เพื่อรอไปถามจู้ดี ผู้ซึ่งสามารถอธิบายได้ทุกอย่าง ในวันดังกล่าว น้ำทะเลมีอุณหภูมิประมาณ 27°C และอุณหภูมิของน้ำมีการเปลี่ยนแปลงน้อยมาก เมื่อระดับความลึกของน้ำเปลี่ยนไป

กราฟที่ 1 ความดันบรรยากาศที่ระดับความลึกต่าง ๆ



ที่มา : <https://study.com/skill/learn/how-to-calculate-the-density-of-a-fluid-from-a-linear-pressure-versus-depth-plot-explanation.html>

ภาพ 2 ตัวอย่างข้อคำถามและข้อมูลประกอบสถานการณ์ในแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์

ตาราง 2 เกณฑ์การให้คะแนนเฉพาะ สถานการณ์ที่ 3 พองอากาศใต้ท้องทะเล

องค์ประกอบ	ระดับคะแนน		
	0	1	2
ข้อกล่าวอ้าง	เขียนข้อความที่ระบุว่า พองอากาศที่คลื่นสังเกตเห็นในบริเวณใกล้ผิวน้ำ <u>จะมีขนาดลดลง หรือ มีขนาดเท่าเดิม</u>	เขียนข้อความที่บ่งบอกว่า พองอากาศที่คลื่นสังเกตเห็นในบริเวณใกล้ผิวน้ำ <u>จะมีขนาดเปลี่ยนแปลงไปจากเดิม</u>	เขียนข้อความที่สื่อความหมายได้ว่า พองอากาศที่คลื่นสังเกตเห็นในบริเวณใกล้ผิวน้ำ <u>จะมีขนาดใหญ่ขึ้น</u>
หลักฐาน	ระบุว่าสังเกตจากข้อมูลอื่น ๆ <u>นอกเหนือจาก</u> (1) ความดันบรรยากาศใต้น้ำที่มีค่าลดลง เมื่อระดับความลึกลดลง และ (2) พองอากาศที่ออกมาจากปากของเต่ามีขนาดใหญ่ขึ้น เมื่อลอยสูงจากตัวเต่ามากขึ้น	ระบุว่าสังเกตจาก (1) ความดันบรรยากาศใต้น้ำที่มีค่าลดลง เมื่อระดับความลึกลดลง และ (2) พองอากาศที่ออกมาจากปากของเต่ามีขนาดใหญ่ขึ้น เมื่อลอยสูงจากตัวเต่ามากขึ้น <u>เพียงอย่างเดียวหนึ่ง</u>	ระบุว่า สังเกตจาก (1) ความดันบรรยากาศใต้น้ำที่มีค่าลดลง เมื่อระดับความลึกลดลง และ (2) พองอากาศที่ออกมาจากปากของเต่ามีขนาดใหญ่ขึ้น เมื่อลอยสูงจาก ตัวเต่ามากขึ้น
การให้เหตุผล	ใช้หลักการทางวิทยาศาสตร์อื่น ๆ เชื่อมโยง <u>หลักฐาน</u> (1) ความดันบรรยากาศใต้น้ำที่มีค่าลดลง เมื่อระดับความลึกลดลง และ (2) พองอากาศที่ออกมาจากปากของเต่ามีขนาดใหญ่ขึ้น เมื่อลอยสูงจากตัวเต่ามากขึ้น ให้เข้ากับ <u>ข้อกล่าวอ้าง</u> ที่ว่า พองอากาศที่คลื่นสังเกตเห็นในบริเวณใกล้ผิวน้ำ จะมีขนาดใหญ่ขึ้น เมื่อเปรียบเทียบกับพองอากาศที่สังเกตเห็นที่บริเวณใต้น้ำ <u>หรือ</u> ใช้หลักการทางวิทยาศาสตร์อื่น ๆ ในการเชื่อมโยงหลักฐานอื่น ๆ ให้เข้ากับข้อกล่าวอ้างที่ต่างออกไป <u>หรือ</u> ไม่ใช้หลักการทางวิทยาศาสตร์ใด ๆ ในการเชื่อมโยงหลักฐานให้เข้ากับข้อกล่าวอ้าง	ใช้ กฎของบอยล์ หรือ หลักการปรับความดันของพองอากาศ <u>เพียงอย่างเดียวหนึ่ง</u> เชื่อมโยง <u>หลักฐาน</u> (1) ความดันบรรยากาศใต้น้ำที่มีค่าลดลง เมื่อระดับความลึกลดลง และ (2) พองอากาศที่ออกมาจากปากของเต่ามีขนาดใหญ่ขึ้น เมื่อลอยสูงจากตัวเต่ามากขึ้น <u>เพียงอย่างเดียวหนึ่ง</u> ให้เข้ากับข้อกล่าวอ้างที่ว่า พองอากาศที่คลื่นสังเกตเห็นในบริเวณใกล้ผิวน้ำ จะมีขนาดใหญ่ขึ้น เมื่อเปรียบเทียบกับพองอากาศที่สังเกตเห็นที่บริเวณใต้น้ำ	ใช้ทั้ง กฎของบอยล์ และ หลักการปรับความดันของพองอากาศ เชื่อมโยง <u>หลักฐาน</u> (1) ความดันบรรยากาศใต้น้ำที่มีค่าลดลง เมื่อระดับความลึกลดลง และ (2) พองอากาศที่ออกมาจากปากของเต่ามีขนาดใหญ่ขึ้น เมื่อลอยสูงจากตัวเต่ามากขึ้น ให้เข้ากับ <u>ข้อกล่าวอ้าง</u> ที่ว่าพองอากาศที่คลื่นสังเกตเห็นในบริเวณใกล้ผิวน้ำ จะมีขนาดใหญ่ขึ้น เมื่อเปรียบเทียบกับพองอากาศที่สังเกตเห็นที่บริเวณใต้น้ำ

2.5 นำแบบวัดที่สร้างขึ้นเสนอต่อผู้เชี่ยวชาญ ประกอบด้วยผู้เชี่ยวชาญในเนื้อหาเรื่องแก๊สที่เป็นอาจารย์สังกัดภาควิชาเคมี 1 ท่าน ผู้เชี่ยวชาญด้านการสร้างคำอธิบายทางวิทยาศาสตร์ที่เป็นอาจารย์สาขาวิชาวิทยาศาสตร์ศึกษา 1 ท่าน และครูผู้สอนวิชาวิทยาศาสตร์ระดับมัธยมศึกษาที่เคยทำวิจัยในระดับปริญญาเอกเกี่ยวกับการสร้างคำอธิบายทางวิทยาศาสตร์ 1 ท่าน รวมจำนวน 3 ท่าน ประเมินคุณภาพด้านความเที่ยงตรงเชิงเนื้อหาและประเมินความเหมาะสมต่อการนำไปใช้ด้วยมาตรประมาณค่า 5 ระดับ

2.6 นำแบบวัดไปทดลองใช้กับกลุ่มตัวอย่างที่เรียนเนื้อหาเรื่อง แก๊ส มาเรียบร้อยแล้ว

2.7 หาคุณภาพของแบบวัด โดยนำข้อมูลมาวิเคราะห์ทางสถิติเพื่อรายงานคุณภาพของแบบวัด ได้แก่ ความยากง่าย อำนาจจำแนก และ ความเชื่อมั่นของแบบวัดทั้งฉบับ

2.8 หาความน่าเชื่อถือของเกณฑ์การให้คะแนน โดยผู้วิจัยได้เลือกคำตอบของนักเรียน 5 คนที่เป็นตัวแทนของคำอธิบายทางวิทยาศาสตร์ในแต่ละลักษณะที่แตกต่างกัน ไปให้ครูอีกสองคนซึ่งไม่เคยมีประสบการณ์ในการสร้างหรือประเมินคำอธิบายทางวิทยาศาสตร์มาก่อน ตรวจให้คะแนนในแต่ละองค์ประกอบของทุกข้อคำถามด้วยเกณฑ์เดียวกัน แล้วนำคะแนนที่ได้ไปคำนวณหาค่าสหสัมพันธ์ภายในชั้น

3. การเก็บรวบรวมข้อมูล

ผู้วิจัยทำหนังสือขออนุญาตเข้าไปเก็บข้อมูลภายในโรงเรียน เมื่อได้รับอนุญาตจากโรงเรียนแล้ว จึงนำแบบวัดที่ผ่านการตรวจสอบคุณภาพและแก้ไขตามข้อเสนอแนะของผู้เชี่ยวชาญเรียบร้อยแล้ว ไปทดลองใช้กับกลุ่มตัวอย่าง ในภาคเรียนที่ 2 ปีการศึกษา 2566 ช่วงเดือนกุมภาพันธ์ พ.ศ. 2567 ก่อนลงมือเก็บข้อมูล ผู้วิจัยซึ่งมีได้เกี่ยวข้องหรือมีอิทธิพลต่อการตัดสินใจของนักเรียนได้ชี้แจงให้นักเรียนทราบเกี่ยวกับการวิจัยและรายละเอียดต่าง ๆ เช่น ข้อมูลที่รวบรวมในครั้งนี้เป็นส่วนหนึ่งของการพัฒนาเครื่องมือ ถ้านักเรียนยินยอมเข้าร่วมการวิจัยจะต้องทำแบบวัดโดยเขียนตอบจำนวน 7 ข้อ นักเรียนสามารถใช้เวลาได้อย่างเต็มที่ นักเรียนสามารถเข้าร่วมและถอนตัวจากการเก็บรวบรวมข้อมูลได้อย่างอิสระ และจะปกปิดข้อมูลส่วนบุคคลของนักเรียนเป็นความลับ ซึ่งพบว่านักเรียนใช้เวลาในการทำแบบวัดประมาณ 50 นาที โดยผู้วิจัยดำเนินการเก็บรวบรวมแบบวัดด้วยตนเอง

4. การวิเคราะห์ข้อมูล

การวิจัยครั้งนี้ มีการวิเคราะห์ข้อมูลสองช่วง ช่วงแรก คือ ก่อนนำแบบวัดไปทดลองใช้ ผู้วิจัยได้วิเคราะห์ข้อมูลดังนี้

1) ความเที่ยงตรงเชิงเนื้อหา ผู้วิจัยนำความคิดเห็นของผู้เชี่ยวชาญทั้ง 3 ท่าน มาคำนวณหาค่าดัชนีความสอดคล้อง (IOC) โดยข้อคำถามที่สอดคล้องกับนิยามเชิงปฏิบัติการควรมีค่า IOC ตั้งแต่ .50 ขึ้นไป

2) ความเหมาะสมต่อการนำไปใช้ ผู้วิจัยนำความคิดเห็นจากรายการประเมินทั้งสี่ประเด็นในข้อคำถามแต่ละข้อ คือ สถานการณ์ ข้อคำถาม เกณฑ์การให้คะแนน และภาษาที่ใช้ มาคำนวณหาค่าเฉลี่ยของความเหมาะสมต่อการนำไปใช้ และแปลความหมายของความเหมาะสมต่อการนำไปใช้ของข้อคำถามแต่ละข้อ ตามเกณฑ์ดังนี้ 4.50 – 5.00 หมายถึง เหมาะสมมากที่สุด 3.50 – 4.49 หมายถึง เหมาะสมมาก 2.50 – 3.49 หมายถึง เหมาะสมปานกลาง 1.50 – 2.49 หมายถึง เหมาะสมน้อย และ 1.00 – 1.49 หมายถึง เหมาะสมน้อยที่สุด (Ritcharoon, 2021)

ช่วงที่สอง หลังจากที้นำแบบวัดไปทดลองใช้ ผู้วิจัยนำแบบวัดที่นักเรียนทำเรียบร้อยแล้วมาตรวจให้คะแนนตามเกณฑ์การให้คะแนนเฉพาะ จากนั้นวิเคราะห์หาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนในแต่ละข้อ และวิเคราะห์ข้อมูลเพื่อรายงานคุณภาพของแบบวัด ดังนี้

1) ความยากง่ายและอำนาจจำแนก ใช้เทคนิค 50% แบ่งคะแนนที่ได้จากการทำแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียนแต่ละข้อเป็นกลุ่มสูงและต่ำ แล้วคำนวณหาค่าความยากง่าย (P_i)

และค่าอำนาจจำแนก (R_i) โดยใช้โปรแกรม Microsoft Excel ข้อคำถามควรมีค่าความยากง่ายอยู่ระหว่าง .20 - .80 และมีค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป (Kanjawasee, 2013)

2) ความเชื่อมั่นของแบบวัดทั้งฉบับ นำคะแนนที่ได้จากการทำแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียนในข้อที่มีความยากง่ายและอำนาจจำแนกเหมาะสมจำนวน 4 ข้อ ไปคำนวณหาค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's " α " Coefficient) ของแบบวัดทั้งฉบับ ผ่านโปรแกรม IBM SPSS Statistics 22 แบบวัดที่มีความเชื่อมั่นสูง ควรมีค่าสัมประสิทธิ์ตั้งแต่ .70 ขึ้นไป (Worakham, 2023)

3) ความน่าเชื่อถือของเกณฑ์การให้คะแนน แสดงผ่านค่าความเชื่อมั่นระหว่างผู้ให้คะแนน การวิจัยครั้งนี้ มีผู้ตรวจให้คะแนนจำนวน 3 คน โดยผู้วิจัยคัดเลือกคำอธิบายทางวิทยาศาสตร์ของนักเรียนที่มีลักษณะแตกต่างกัน จำนวน 5 คน แล้วนำคะแนนที่ได้จากการตรวจซ้ำของนักเรียนห้าคน ซึ่งจัดเป็นข้อมูลในระดับอันตรภาค (interval scale) มาหาค่าสหสัมพันธ์ภายในชั้น (Intra-class correlation coefficient; ICC) เกณฑ์การให้คะแนนที่มีคุณภาพและน่าเชื่อถือ ควรมีความเชื่อมั่นระหว่างผู้ตรวจให้คะแนน .85 ขึ้นไป (Worakham, 2023)

ผลการวิจัย

แบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส ที่ได้สร้างขึ้น มีค่าดัชนีความสอดคล้อง (IOC) และค่าความเหมาะสมต่อการนำไปใช้เฉลี่ย ดังตาราง 3

ตาราง 3 ค่าดัชนีความสอดคล้องและความเหมาะสมต่อการนำไปใช้ของข้อคำถามในแบบวัด

ข้อ ที่	สถานการณ์	IOC	การแปลความหมาย	ความเหมาะสม ต่อการนำไปใช้เฉลี่ย	การแปลความหมาย
1	ถุงลมปอด	1.00	สอดคล้อง	4.50	เหมาะสมมากที่สุด
2	เลือดซึบพองได้	0.67	สอดคล้อง	4.00	เหมาะสมมาก
3	ฟองอากาศใต้ทะเล	1.00	สอดคล้อง	4.33	เหมาะสมมาก
4	ขวดน้ำดื่มแช่เย็น	1.00	สอดคล้อง	4.50	เหมาะสมมากที่สุด
5	แป้งพืชมะพร้าว	0.67	สอดคล้อง	4.08	เหมาะสมมาก
6	หลอดโซเดียมความดันต่ำ	0.67	สอดคล้อง	4.08	เหมาะสมมาก
7	แก๊สปิโตรเลียม	1.00	สอดคล้อง	4.42	เหมาะสมมาก

จากตาราง 3 พบว่า ข้อคำถามทุกข้อมีค่า IOC อยู่ระหว่าง .67 – 1.00 ซึ่งมากกว่า .50 แสดงว่า ข้อคำถามทุกข้อวัดได้ตรงตามนิยามเชิงปฏิบัติการ (Ritcharoon, 2021) และทุกข้อมีความเหมาะสมต่อการนำไปใช้เฉลี่ยอยู่ระหว่าง 4.00 – 4.50 จัดอยู่ในระดับมากถึงมากที่สุด โดยผู้เชี่ยวชาญมีความคิดเห็นเพิ่มเติม ดังนี้ ในด้านสถานการณ์ มีความคิดเห็นตรงกัน คือ สถานการณ์น่าสนใจ นำไปสู่การสร้างคำอธิบายทางวิทยาศาสตร์ได้ แต่ควรเพิ่มหลักฐานลงให้ครบทุกข้อ จะได้ประเมินการเลือกหลักฐานของนักเรียนได้อย่างชัดเจน ผู้วิจัยจึงได้เพิ่มข้อมูลในข้อคำถาม จนมีรายละเอียดดังตาราง 1 ด้านเกณฑ์การให้คะแนน ควรเพิ่ม

รายละเอียดของเกณฑ์การให้คะแนนที่ชัดเจนมากยิ่งขึ้น ซึ่งผู้วิจัยจึงได้แจกแจงกรณีเพิ่มเติม เช่น ด้านการให้เหตุผล เดิมระบุเพียงว่า ได้ 0 คะแนน เมื่อใช้หลักการทางวิทยาศาสตร์เชื่อมโยงหลักฐานให้เข้ากับข้อกล่าวอ้างไม่ถูกต้อง ปรับเป็นดังตัวอย่างในตาราง 2 ที่มีกรณีที่เป็นไปได้สามกรณีที่แตกต่างกัน

เมื่อผู้วิจัยตรวจให้คะแนนตามเกณฑ์การให้คะแนนเฉพาะเสร็จเรียบร้อยแล้ว นำคะแนนแยกตามองค์ประกอบและคะแนนรวม ไปหาค่าเฉลี่ย ($\bar{X}$) และส่วนเบี่ยงเบนมาตรฐาน (S.D.) ได้ข้อมูลดังตาราง 4

ตาราง 4 คะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

ข้อที่	สถานการณ์	ข้อกล่าวอ้าง (2 คะแนน)		หลักฐาน (2 คะแนน)		การให้เหตุผล (2 คะแนน)		คะแนนรวม (6 คะแนน)	
		$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
1	ถูกลมนิรภัย	1.28	0.58	0.59	0.50	0.00	0.00	1.88	0.94
2	เสื้อชูชีพพองได้	1.78	0.61	0.34	0.55	0.19	0.40	2.31	1.06
3	พองอากาศได้ทะเล	1.88	0.42	0.97	0.59	0.56	0.50	3.41	0.80
4	ขวดน้ำดื่มแช่เย็น	1.56	0.62	0.41	0.61	0.13	0.34	2.09	0.82
5	แปรงพิชชาหลังอบ	1.87	0.43	0.42	0.56	0.45	0.51	2.74	0.96
6	หลอดโซเดียมความดันต่ำ	1.94	0.25	0.47	0.51	0.06	0.35	2.47	0.67
7	แก๊สปิโตรเลียม	0.91	0.69	0.66	0.60	0.44	0.72	2.00	1.57

จากตาราง 4 เมื่อพิจารณาคะแนนรวมจะเห็นว่า คำถามข้อที่ 3 มีคะแนนเฉลี่ยสูงที่สุด เท่ากับ 3.41 จากคะแนนเต็ม 6 คะแนน และคำถามข้อที่ 1 มีคะแนนเฉลี่ยต่ำที่สุด คือ 1.88 คำถามข้อที่ 7 มีส่วนเบี่ยงเบนมาตรฐานสูงที่สุด คือ 1.57 และคำถามข้อที่ 6 มีส่วนเบี่ยงเบนมาตรฐาน ต่ำที่สุด คือ 0.67

เมื่อผู้วิจัยนำคะแนนรวมของนักเรียนแต่ละคนมาเรียงลำดับจากสูงไปต่ำ แล้วใช้เทคนิค 50% แบ่งนักเรียนออกเป็นสองกลุ่ม คือ กลุ่มสูงและกลุ่มต่ำที่มีจำนวนเท่า ๆ กัน จากนั้นคำนวณหาความยากง่ายและอำนาจจำแนกของข้อคำถามแต่ละข้อ ได้ข้อมูลดังตาราง 5

ตาราง 5 ความยากง่าย อำนาจจำแนก และผลการพิจารณา

ข้อที่	สถานการณ์	ความยากง่าย	ความหมาย	อำนาจจำแนก	ความหมาย	ผลการพิจารณา
1	ถูกลมนิรภัย	.31	ค่อนข้างยาก	.27	จำแนกได้พอใช้	นำไปใช้ได้
2	เสื้อชูชีพพองได้	.39	ค่อนข้างยาก	.23	จำแนกได้พอใช้	นำไปใช้ได้*
3	พองอากาศได้ทะเล	.57	ปานกลาง	.22	จำแนกได้พอใช้	นำไปใช้ได้*
4	ขวดน้ำดื่มแช่เย็น	.35	ค่อนข้างยาก	.18	จำแนกได้ค่อนข้างต่ำ	ตัดออก
5	แปรงพิชชาหลังอบ	.44	ปานกลาง	.24	จำแนกได้พอใช้	นำไปใช้ได้*
6	หลอดโซเดียมความดันต่ำ	.41	ปานกลาง	.19	จำแนกได้ค่อนข้างต่ำ	ตัดออก
7	แก๊สปิโตรเลียม	.33	ค่อนข้างยาก	.42	จำแนกได้ดี	นำไปใช้ได้*

* ข้อที่ถูกคัดเลือก

จากตาราง 5 จะเห็นว่า คำถามทุกข้อมีค่าความยากง่ายตั้งแต่ระดับปานกลางถึงค่อนข้างยาก คำถามข้อที่ 1, 2, 3, 5 และ 7 มีอำนาจจำแนกตั้งแต่ระดับพอใช้ถึงดี แต่คำถามข้อที่ 4 และ 6 มีอำนาจจำแนกต่ำกว่า .20 แปลว่า จำแนกได้ค่อนข้างต่ำ ผู้วิจัยจึงพิจารณาตัดข้อคำถามทั้งสองข้อออก เนื่องจากเนื้อหาเรื่อง แก๊สประกอบด้วย 6 หัวข้อ ได้แก่ กฎของอาโวกาโดร กฎของบอยล์ กฎของชาร์ล กฎของเกย์-ลูสแซก กฎความดันย่อยของดอลตัน และกฎการแพร่ผ่านของเกรแฮม แต่ข้อคำถามในแบบวัดมีได้เพียง 4 ข้อตามที่กำหนดไว้ และควรมีสถานการณ์ในชีวิตประจำวันที่ครอบคลุมกฎต่าง ๆ ของแก๊สให้ได้มากที่สุด ข้อที่ 3 ใช้กฎของบอยล์ ข้อที่ 5 ใช้กฎของชาร์ล และข้อที่ 7 ใช้กฎการแพร่ผ่านของเกรแฮม ส่วนคำถามข้อที่ 1 และ 2 ใช้กฎเดียวกันคือกฎของอาโวกาโดร ทั้งสองข้อมีค่าความยากง่ายอยู่ในระดับค่อนข้างยากเหมือนกัน และมีอำนาจจำแนกอยู่ในระดับพอใช้เท่ากัน ผู้วิจัยจึงพิจารณาที่ลักษณะคำตอบของนักเรียนจากการนำไปทดลองใช้ ซึ่งพบว่า คำถามข้อที่ 1 มีนักเรียนบางส่วนเข้าใจผิดเขียนแสดงวิธีการคำนวณหาค่าเส้นรอบวงของทรงกลมนิรภัยจากมวลของแก๊ส แทนการเขียนคำอธิบายทางวิทยาศาสตร์เพื่อตอบคำถามดังกล่าว ส่วนคำถามข้อที่ 2 นักเรียนทุกคนเขียนตอบในลักษณะที่เป็นคำอธิบายทางวิทยาศาสตร์ แต่อาจจะมียกประกอบไม่ครบถ้วนสมบูรณ์ ดังนั้นในการวิจัยนี้คำถามข้อที่ 2, 3, 5 และ 7 จึงเป็นข้อที่ถูกคัดเลือกไว้

ผู้วิจัยนำคะแนนของนักเรียนที่ได้จากแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ในข้อที่ 2, 3, 5 และ 7 ไปคำนวณหาค่าความเชื่อมั่นของแบบวัดทั้งฉบับด้วยค่าสัมประสิทธิ์แอลฟาของครอนบาค พบว่ามีค่าเท่ากับ .732 แสดงว่า แบบวัดดังกล่าวมีความเชื่อมั่นในระดับสูง สามารถนำไปใช้วัดและประเมินผลการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส ของนักเรียนระดับชั้นมัธยมศึกษาตอนปลายได้

ด้านความน่าเชื่อถือของเกณฑ์การให้คะแนน ผู้วิจัยนำคะแนนที่ได้จากการตรวจให้คะแนนโดยใช้เกณฑ์การให้คะแนนเดียวกันจากผู้ตรวจทั้งหมด 3 คน ไปวิเคราะห์หาความเชื่อมั่นระหว่างผู้ตรวจให้คะแนนด้วยค่าสหสัมพันธ์ภายในชั้น (ICC) พบว่า มีค่าเท่ากับ .862 ซึ่งสูงกว่าระดับที่กำหนดไว้ แสดงว่า เกณฑ์การให้คะแนนมีความละเอียด ชัดเจน ครูที่ไม่เคยมีประสบการณ์เกี่ยวกับการสร้างคำอธิบายทางวิทยาศาสตร์มาก่อนก็สามารถตรวจให้คะแนนได้ไม่แตกต่างกัน

อภิปรายผล

การวิจัยนี้มุ่งพัฒนาแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย เรื่อง แก๊ส ให้มีคุณภาพ โดยแบบวัดที่พัฒนาขึ้นใช้สถานการณ์ที่เกิดขึ้นรอบตัวและเกิดขึ้นได้ในโลกแห่งความจริงเกี่ยวกับแก๊ส พร้อมให้ข้อมูลและหลักการเพิ่มเติม (ถ้ามี) และถามให้สร้างคำอธิบายทางวิทยาศาสตร์

จากผลการวิจัยพบว่า แบบวัดที่ได้สร้างขึ้นสามารถวัดพฤติกรรมด้านการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส ของนักเรียนระดับชั้นมัธยมศึกษาตอนปลายได้ตรงตามนิยามเชิงปฏิบัติการ มีความเหมาะสมต่อการนำไปใช้ในระดับมากถึงมากที่สุด มีความยากง่ายตั้งแต่ระดับปานกลางถึงค่อนข้างยาก มีอำนาจจำแนกพอใช้ถึงดี มีความเชื่อมั่นในระดับสูง มีความน่าเชื่อถือของเกณฑ์การให้คะแนน สามารถอภิปรายผลการวิจัยในประเด็นต่าง ๆ ได้ ดังนี้

ข้อคำถามทุกข้อในแบบวัดที่สร้างขึ้นมีความเที่ยงตรงเชิงเนื้อหา ผู้วิจัยใช้สถานการณ์ที่เกี่ยวข้องกับการประยุกต์ใช้กฎของแก๊สในการอธิบาย แก๊ส เช่น หลอดโซเดียมความดันต่ำที่ใช้ให้แสงสว่างบนท้องถนน การนำขวดน้ำดื่มเข้าไปแช่ในตู้เย็น การพองของถุงลมนิรภัย และต่างก็เป็นสถานการณ์ที่นักเรียนสามารถพบได้ในชีวิตประจำวัน สอดคล้องกับลักษณะของแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ที่ดีที่ Kamtet et al. (2020) กล่าวไว้ คือ ควรเป็นสถานการณ์ที่เกิดขึ้นจริงในชีวิตประจำวัน หรือสถานการณ์ที่มีบริบทจำเพาะเพื่อช่วยกระตุ้นให้นักเรียนพิจารณาข้อมูลประกอบสถานการณ์ที่ให้ อีกทั้งยังใช้ประโยคคำสั่งขึ้นต้นเหมือนกันในทุกข้อคำถาม คือ “จากข้อมูลข้างต้น ให้นักเรียนสร้างคำอธิบายทางวิทยาศาสตร์ว่า.....” แล้วจึงระบุคำถามที่จำเพาะกับสถานการณ์ ทำให้ทุกข้อคำถามมีความชัดเจนว่าต้องการให้นักเรียนทำอะไร ส่งผลให้ข้อคำถามทุกข้อมีค่าดัชนีความสอดคล้อง (IOC) อยู่ระหว่าง .67 – 1.00 อยู่ในระดับยอมรับได้ และ มีค่าใกล้เคียงกับแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ก่อนหน้า (Nimdam et al., 2020; Tanang, 2019)

ความเหมาะสมต่อการนำไปใช้โดยเฉลี่ยของแบบวัด มีค่าอยู่ระหว่าง 4.00 – 4.50 ซึ่งอยู่ในระดับมากถึงมากที่สุด เนื่องจากลักษณะของคำถามในแบบวัดไม่ได้เป็นการถามความรู้ความจำ (Janto et al., 2020) แต่เป็นการถามตามรายละเอียดของสถานการณ์ที่ให้ ดังที่ Wetworanan et al. (2019) ได้ระบุลักษณะคำถามที่ใช้เพื่อให้นักเรียนสร้างคำอธิบายทางวิทยาศาสตร์ไว้ว่า รูปแบบคำถาม จะต้องเป็นคำถามที่กระตุ้นให้หาแนวโน้มที่เป็นไปได้มากที่สุด ผ่านการคิดพิจารณาตามเงื่อนไขที่มีในข้อคำถาม

ข้อคำถามที่คัดเลือกไว้ 4 ข้อในแบบวัด มีค่าความยากง่ายตั้งแต่ .33 - .57 จัดอยู่ในระดับปานกลางถึงค่อนข้างยาก สอดคล้องกันกับการวิจัยก่อนหน้าของ Wannathai & Pruekpramool (2021) ที่มีความยากอยู่ในระดับปานกลางถึงค่อนข้างยาก ซึ่งอาจเป็นผลมาจากสถานการณ์ที่ใช้ในข้อคำถามบางข้อมีการใช้หลักการอื่น ๆ เพิ่มเติม นอกเหนือจากกฎของแก๊ส ซึ่งต่างจากงานวิจัยของ Janto et al. (2020) ที่ได้ให้ข้อเสนอแนะเอาไว้ว่า แบบวัดหนึ่งข้อ ควรมีหลักการทางวิทยาศาสตร์ที่เกี่ยวข้องเพียงหนึ่งแนวคิด เพราะโดยทั่วไปนักเรียนจะร่างคำตอบขึ้นมาตามแนวคิดหรือหลักการทางวิทยาศาสตร์ก่อน แล้วจึงค่อยเลือกหลักฐานมาสนับสนุน (Wetworanan et al., 2019) เมื่อมีแนวคิดที่มากขึ้น จึงอาจทำให้นักเรียนเกิดความสับสนได้ แต่สถานการณ์ที่เลือกมานั้น บางสถานการณ์ไม่สามารถอธิบายได้ด้วยกฎของแก๊สเพียงข้อเดียว เพราะเป็นสถานการณ์ในชีวิตประจำวันที่มีความซับซ้อน จึงมีความจำเป็นที่จะต้องใช้หลักการเพิ่ม และได้ระบุลงไป ข้อมูลประกอบสถานการณ์ด้วย และในส่วนของการให้เหตุผล นักเรียนส่วนใหญ่ไม่ได้เขียนตอบในส่วนดังกล่าว (Hirio et al., 2020) ทำให้ได้คะแนนเป็นศูนย์ จึงทำให้ค่าที่คำนวณออกมามีค่าค่อนข้างน้อย ส่งผลให้ความยากง่ายของแบบวัดอยู่ในระดับปานกลางถึงค่อนข้างยาก

ข้อคำถามที่ได้คัดเลือกไว้ 4 ข้อของแบบวัดมีค่าอำนาจจำแนกอยู่ระหว่าง .22 - .42 คือ จำแนกได้พอใช้ถึงจำแนกได้ดี ส่วนใหญ่อยู่ในระดับพอใช้ ซึ่งต่างจากงานวิจัยก่อนหน้าที่ข้อคำถามมีอำนาจจำแนกอยู่ในระดับดีมาก (Wannathai & Pruekpramool, 2021) โดยมีข้อสังเกตคือ ข้อคำถามที่มีการกระจายตัวของคะแนนมาก จะมีอำนาจจำแนกที่สูง เช่น ข้อที่ 7 ที่มีส่วนเบี่ยงเบนมาตรฐานสูงที่สุดเท่ากับ 1.57 มีอำนาจจำแนกเป็น .42 ซึ่งมีค่าสูงที่สุด และจัดอยู่ในระดับดี

ความเชื่อมั่นของแบบวัดฉบับนี้ มีค่าสัมประสิทธิ์แอลฟาของครอนบาคเท่ากับ .732 จัดอยู่ในระดับสูง ซึ่งไม่แตกต่างจากแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ก่อนหน้านี้ (Wannathai & Pruekpramool, 2021) เป็นแบบวัดที่มีคุณภาพและเหมาะสมที่จะนำไปใช้ประเมินการสร้างคำอธิบายทางวิทยาศาสตร์ เรื่อง แก๊ส ของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย การที่แบบวัดมีความเชื่อมั่นในระดับสูง สอดคล้องกับค่าความเที่ยงตรงเชิงเนื้อหาที่บ่งบอกว่าข้อคำถามทุกข้อในแบบวัดสามารถวัดได้ตรงตามนิยามเชิงปฏิบัติการ

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 แบบวัดฉบับนี้เป็นแบบวัดการสร้างคำอธิบายทางวิทยาศาสตร์ที่มีความจำเพาะต่อเนื้อหา เรื่อง แก๊ส สำหรับนักเรียนระดับชั้นมัธยมศึกษาตอนปลายแผนการเรียนวิทยาศาสตร์-คณิตศาสตร์เท่านั้น การนำแบบวัดไปใช้จะต้องมั่นใจว่า นักเรียนจะต้องได้เรียนรู้เกี่ยวกับกฎต่าง ๆ ของแก๊สมาเรียบร้อยแล้ว เพื่อไม่ให้ความรู้เรื่องแก๊ส เป็นอุปสรรคต่อการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียน โดยอาจจะนำไปประเมินในตอนท้ายของบทเรียน เช่นเดียวกันกับการสอบเก็บคะแนน

1.2 ข้อคำถามในแบบวัดมีความง่ายอยู่ในระดับปานกลางถึงค่อนข้างยาก อาจจะต้องพิจารณาถึงระดับความสามารถของนักเรียนก่อนนำแบบวัดนี้ไปใช้

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ผลการวิจัยในครั้งนี้ ได้มาจากกลุ่มตัวอย่างที่เป็นนักเรียนจากโรงเรียนมัธยมศึกษาขนาดใหญ่พิเศษ ในจังหวัดกรุงเทพมหานครเพียงโรงเรียนเดียว ซึ่งอาจจะมีบริบทในการเรียนที่แตกต่างจากโรงเรียนในลักษณะอื่น ๆ เช่น โรงเรียนหญิงล้วน โรงเรียนขนาดเล็กในต่างจังหวัด ในการวิจัยครั้งต่อไป จึงควรนำแบบวัดฉบับเดียวกันนี้ไปทดลองใช้กับนักเรียนในโรงเรียนอื่น ๆ ที่มีบริบทแตกต่างไปจากเดิม

2.2 แบบวัดนี้สามารถบอกได้เพียงระดับคะแนนในแต่ละองค์ประกอบของการสร้างคำอธิบายทางวิทยาศาสตร์เท่านั้น แต่ยังไม่สามารถให้รายละเอียดเชิงลึกได้ ดังนั้น การวิจัยครั้งต่อไปควรออกแบบวิธีการวัดและประเมินที่แสดงให้เห็นถึงรายละเอียดขั้นตอนหรือกระบวนการสร้างคำอธิบายทางวิทยาศาสตร์ของนักเรียน เช่น การใช้ Think aloud การสัมภาษณ์ เพื่อที่ครูจะได้นำผลการประเมินไปช่วยเหลือนักเรียนในการสร้างองค์ประกอบด้านหลักฐานและการให้เหตุผลได้ตรงจุดมากยิ่งขึ้น

2.3 ควรศึกษาความสัมพันธ์ระหว่างลักษณะของข้อคำถาม และลักษณะของข้อมูลประกอบสถานการณ์ว่า ลักษณะข้อคำถาม และลักษณะของข้อมูลประกอบสถานการณ์แบบใด ที่ช่วยกระตุ้นหรือส่งเสริมให้นักเรียนสามารถสร้างคำอธิบายทางวิทยาศาสตร์ได้ ซึ่งจะเป็นประโยชน์ต่อการพัฒนาสื่อประกอบการจัดการเรียนรู้ที่ช่วยส่งเสริมการสร้างคำอธิบายทางวิทยาศาสตร์ในอนาคตได้

References

- McNeill, K. L., & Krajcik, J. S. (2008). Assessing middle school students' content knowledge and reasoning through written scientific explanations. In J. Coffey, R. Douglas, & C. Stearns (Eds.), *Assessing science learning: Perspectives from research and practice* (pp. 101–116). NSTA Press.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press.
<https://doi.org/10.17226/13165>
- Tang, K. S. (2015). The PRO instructional strategy in the construction of scientific explanations. *Teaching Science*, 61(4), 14-21.
- Yao, J. X., & Guo, Y. Y. (2018). Validity evidence for a learning progression of scientific explanation. *Journal of Research in Science Teaching*, 55(2), 299-317.

Translate Thai References

- Aniwatanawong P., & Jantrasri, R. (2019). Evaluating the impact of model-based inquiry with model-evidence link on grade 11 students' scientific explanation of electrolyte and non-electrolyte solutions. *Walailak Journal of Learning Innovations*, 5(1), 65–83. (in Thai)
- Hirio, S., Pitiporntapin, S., Chumnannpuen, P., & Phongpajit, P. (2020). The development of 11th grade students' scientific explanation construction in learning unit of circulatory system using inquiry-based mobile learning. *Proceedings of the 21st National Graduate Research Conference (Online), Khon Kaen University*, 518–528. (in Thai)
- Janto, N., Kijkuakul, S., & Chapoo, S. (2020). The Development of the instruction argumentation using forensic science issue for promoting the scientific explanation ability of grade 10 students on human homeostasis. *Journal of Education and Innovation*, 24(1), 176–187. (in Thai)
- Juthanaruepakij, K., Chanthrakhanthi, E., & Chiangga, S. (2018). Implementing scientific inquiry approaches to develop grade 10th students' scientific explanation in the topic of circular motion. *Proceedings of the 13th RSU National Graduate Research Conference*, 13, 728–740. (in Thai)
- Kamtet, W., Wannagatesiri, T., & Kruea-In, N. (2020). Scientific Explanation Ability Assessment Based on Context-Rich Problems for Biology Student Teachers. *Journal of Education, Prince of Songkla University, Pattani Campus*, 31(3), 45–59. (in Thai)

- Kanjanawasee, S. (2013). *Classical Test Theory* (7th ed.). Chulalongkorn University Press. (in Thai)
- Nimdam, S., Chaowakeratipong, N., & Suwanjinda, D. (2020). The Effects of Inquiry Learning Management Together with Science Writing Heuristic in the Topic of Reaction Rate on Science Learning Achievement and Scientific Explanation of MathayomSuksa V at Small School in Chumphon Campus Area II of Chumphon Province. *Journal of Kasetsart Educational Review*, 35(2), 96 – 109. (in Thai)
- Office of the Basic Education Commission (OBEC). (2017). *Learning Standards and Indicators in Science (revised 2017) according to Basic Education Core Curriculum B.E. 2551 (A.D. 2008)*. The Agricultural co-operative federation of Thailand Press. (in Thai)
- Promyod, N. (2019). Construction Scientific Explanation through CER (Claim, Evidence, and Reasoning) Learning Process. *IPST Magazine*, 47(219), 11-15. (in Thai)
- Ritcharoon, P. (2021). *Classroom research Techniques*. Chatuporn design. (in Thai)
- Singwee, K. (2020). Guidelines for Science Instructional Management to Promote Ability in Scientific Explanation making of Students in Basic Educational Level. *Journal of Education Mahasarakham University*, 14(3), 18-31. (in Thai)
- Supasatwong, P., & Sangsuwan, A. (2022). The Development of Scientific Explanation Ability with Scientific Model on Electrochemistry of Grade 11 Students Using Model-Based Inquiry Combining with Animation. *Journal of Inclusive and Innovative Education*, 6(3), 101–115. (in Thai)
- Tanang, M. (2019). *The development of scientific explanation ability in chemical equilibrium of grade 11 students using predict observe explain (POE)* [Master's thesis]. Rangsit University. (in Thai)
- The Institute for the Promotion of Teaching Science and Technology (IPST). (2017). PISA 2015 Assessment and Analytical Framework. The Institute for the Promotion of Teaching Science and Technology (IPST). (In Thai)
- Wannathai, P., & Pruekpramool, C. (2021). Comparing the efficiency between ability in constructing scientific explanation restricted – and extended – response subjective test for mathayom suksa 4 students. *Journal of graduate research*, 12(2), 103 - 118. (in Thai)
- Wetworanan, R., Tanak, A., Faikhamta, C. & Sritha, S. (2019). Developing Grade 10 Students' Scientific Explanation Skill using Game-based Inquiry. *Proceedings of the 20th National Graduate Research Conference, Khon Kaen University*, 1664-1674. (in Thai)
- Worakham, P. (2023). *Educational Research* (14th ed.). Taksila press. (in Thai)

การเปรียบเทียบคุณภาพของรูปแบบการวัดและประเมินผลตามแผนการศึกษา ของมหาวิทยาลัยสุโขทัยธรรมมาธิราช

A Comparison of the Quality of Measurement and Evaluation Models under the Study Plans of Sukhothai Thammathirat Open University

วินิตา แก้วแก้ว¹ และ สฤษฎ์เทพ สุขแก้ว^{1*}

Vinita Kaewkua¹ and Sarhistthep Sukkaew^{1*}

(Received: September 5, 2024; Revised: October 5, 2024; Accepted: October 8, 2024)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมายเพื่อ 1) เปรียบเทียบคะแนนรวมที่ได้มาจากคะแนนการสอบและคะแนนที่ได้จากการทำกิจกรรมทั้งหมดของแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3 2) เปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่านโดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3 และ 3) เปรียบเทียบความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3 การวิจัยครั้งนี้มีวิธีดำเนินการวิจัย 3 ระยะ คือ ระยะที่ 1 การปรับเทียบคะแนนรวมที่ได้มาจากคะแนนการสอบและคะแนนที่ได้จากการทำกิจกรรมทั้งหมด ระยะที่ 2 การเปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน ด้วยสถิติ Mann – Whitney U test และ ระยะที่ 3 การเปรียบเทียบความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนน โดยใช้สถิติสัมประสิทธิ์แคปปา

ผลการวิจัย พบว่า 1) ผลการปรับเทียบคะแนนในแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3 ไม่แตกต่างกัน 2) ผลการเปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่านโดยพิจารณาจากผลการตัดเกรดในแต่ละแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3 ไม่แตกต่างกัน 3) ผลการเปรียบเทียบความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก2 มีระดับความสอดคล้องมาก มีค่าเท่ากับ 0.893 และ ผลการเปรียบเทียบความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก3 มีระดับความสอดคล้องมาก มีค่าเท่ากับ 0.830

คำสำคัญ การเปรียบเทียบคุณภาพ รูปแบบการวัดและประเมินผล แผนการศึกษา มหาวิทยาลัยเปิด

¹ อาจารย์ สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมมาธิราช

¹ Lecturer, Office of Registration Records and Evaluation, Sukhothai Thammathirat Open University

ทุนอุดหนุนการวิจัยสถาบันประจำปี 2566 (ประเภทโครงการวิจัยสถาบันเชิงนโยบาย) มหาวิทยาลัยสุโขทัยธรรมมาธิราช

Institutional Research Grant 2023 funds (Policy Institutional Research Project), Sukhothai Thammathirat Open University

* Corresponding Author E-mail: sarhistthep.suk@stou.ac.th

Abstract

This research aimed to: (1) compare the total scores obtained from examinations and activity-based assessments in Educational Plans A1 and A2 with those in Plans A1 and A3 ; (2) compare the differences in the proportion of students who passed the examination based on the grading results of Plans A1 and A2, and Plans A1 and A3; and (3) compare the concordance between grading results using raw scores (before equating) and grading results using equated scores in Plans A1 and A2, and Plans A1 and A3. This research consisted of three phases. Phase 1 involved equating the total scores from examination and activity-based assessments. Phase 2 involved comparing the differences in the proportion of students who passed the examination using the Mann–Whitney U test. Phase 3 involved comparing the consistency of grading results before and after score equating using Kappa coefficient statistics.

The research findings were as follows: (1) There was no difference in the equated scores between Plans A1 and A2 and Plans A1 and A3; (2) There was no difference in the proportion of students who passed the examination under Plans A1 and A2, and Plans A1 and A3; and (3) The comparison of grading concordance before and after score equating showed strong agreement. The Kappa value for Plans A1 and A2 was 0.893, and for Plans A1 and A3 was 0.830.

Keywords: comparison of the quality, measurement and evaluation model, study plans, open university

บทนำ

มหาวิทยาลัยสุโขทัยธรรมมาธิราชเป็นมหาวิทยาลัยเปิดที่ใช้ระบบการเรียนการสอนทางไกลมายาวนานกว่า 40 ปี และได้เผชิญกับความเปลี่ยนแปลงของภาวะการแข่งขันทางการจัดการศึกษาในยุคดิจิทัล ตลอดจนปัญหาการลดลงของจำนวนนักศึกษาที่เข้าศึกษาในมหาวิทยาลัย ซึ่งเป็นวิกฤตที่มหาวิทยาลัยต้องเผชิญ มหาวิทยาลัยสุโขทัยธรรมมาธิราชจึงได้มีนโยบายการปฏิรูปมหาวิทยาลัยสุโขทัยธรรมมาธิราชเมื่อ พ.ศ. 2561 เพื่อมุ่งปรับปรุงระบบการศึกษาทางไกลและระบบที่เกี่ยวข้อง เพื่อการพัฒนาเทคโนโลยีการศึกษาทางไกลที่สนับสนุนทุกขั้นตอน เพื่อการจัดการศึกษาให้สอดคล้องกับรูปแบบการจัดการเรียนการสอนที่เปลี่ยนแปลงไป และให้เหมาะสมกับความต้องการของนักศึกษา โดยมีการนำแบบจำลองระบบการศึกษาทางไกล มสธ. 2561 เป็นแผนแม่บทของระบบการสอนทางไกล มหาวิทยาลัยสุโขทัยธรรมมาธิราช และใช้อำนาจในการจัดการศึกษาของมหาวิทยาลัยในระดับปริญญาตรี โดยพัฒนาแผนการจัดการศึกษาแบบเรียนรู้ด้วยตนเองในระดับปริญญาตรีออกเป็น 3 แผน คือ แผนการศึกษา ก1 นักศึกษาศึกษาด้วยตนเองจากเอกสารการสอนชุดวิชา หรือการทำกิจกรรมการเรียนการสอนใน e – Learning มีการวัดและประเมินผลจากการสอบ

ประจำภาคการศึกษา 100 คะแนน แผนการศึกษา ก2 นักศึกษาศึกษาด้วยตนเองจากเอกสารการสอนชุดวิชา ร่วมกับการเข้าร่วมกิจกรรมสอนเสริม เพื่อเสริมความเข้าใจในเนื้อหา เพิ่มเติมจากการศึกษาด้วยตนเอง โดยเน้นการสอนแบบตัวตามหน่วยการสอน ทั้งการสอนแบบเผชิญหน้าและการสอนแบบเสมือนจริง แบบใช้เทคโนโลยี มีการวัดและประเมินผลจากกิจกรรมสอนเสริม 40 คะแนน และการสอบประจำภาคการศึกษา 60 คะแนน แผนการศึกษา ก3 ที่นักศึกษาศึกษาด้วยตนเองจากเอกสารการสอนชุดวิชา ร่วมกับการเข้าร่วม กิจกรรมร่วมเรียนรู้ เพื่อประยุกต์ใช้องค์ความรู้และเพิ่มพูนประสบการณ์จากการทำกิจกรรม ทั้งการสอน แบบเผชิญหน้าและการสอนแบบเสมือนจริง แบบใช้เทคโนโลยี มีการวัดและประเมินผลจากกิจกรรมร่วมเรียนรู้ 40 คะแนน และการสอบประจำภาคการศึกษา 60 คะแนน (แบ่งการสอบเป็นสองช่วง สอบกลางภาค 30 คะแนน และสอบปลายภาค 30 คะแนน) (Sukhothai Thammathirat Open University, n.d.)

จากแผนการศึกษา ก1 ก2 และ ก3 ดังกล่าวที่มีการจัดการเรียนการสอนและการให้คะแนนแตกต่างกัน ก่อนที่จะนำคะแนนในแต่ละส่วน คือ คะแนนจากการสอบ และคะแนนจากการทำกิจกรรมที่จัดขึ้นในแต่ละ แผนการศึกษา มารวมกันและตัดเกรด ทำให้ผู้วิจัยในฐานะของอาจารย์นักวัดผลการศึกษา สำนักทะเบียนและ วัดผลจึงเกิดคำถามขึ้นว่าแผนการศึกษา ก2 และ ก3 ซึ่งเป็นแผนการศึกษาที่ใช้รูปแบบการวัดและประเมินผล รูปแบบใหม่จะช่วยให้มีจำนวนนักศึกษาที่สอบผ่านโดยเฉลี่ยมากขึ้นหรือไม่ เมื่อเปรียบเทียบกับแผนการศึกษา ก1 คะแนนเก็บร้อยละ 40 ของแผนการศึกษา ก2 และ ก3 จะเป็นคะแนนช่วยให้มีจำนวนนักศึกษาที่สอบผ่าน โดยเฉลี่ยมากขึ้นหรือไม่ เมื่อเปรียบเทียบกับแผนการศึกษา ก1 และคะแนนจากการสอบกลางภาคในแผน การศึกษา ก3 จะช่วยให้มีจำนวนนักศึกษาที่สอบผ่านโดยเฉลี่ยมากขึ้นหรือไม่ และผู้วิจัยต้องการศึกษา ความแตกต่างของคุณภาพแบบสอบ โดยในที่นี้คุณภาพของแบบสอบพิจารณาจาก 3 ประเด็น คือ 1) ความ สอดคล้องของผลการเปรียบเทียบคะแนนรวมที่ได้มาจากคะแนนการสอบและคะแนนที่ได้จากการทำกิจกรรม ทั้งหมดของแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3 2) ความไม่แตกต่างกันระหว่างสัดส่วน นักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และ สัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก2 กับ ความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และ สัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผล การตัดเกรดของแผนการศึกษา ก3 และ 3) ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการ เปรียบเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการเปรียบเทียบคะแนนของแผนการศึกษา ก1 และ ก2 กับ ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการเปรียบเทียบคะแนนกับผลการตัด เกรดจากการใช้คะแนนหลังจากการเปรียบเทียบคะแนนของแผนการศึกษา ก1 และ ก3 ที่ใช้ในการสอบนักศึกษา ที่ลงทะเบียนเรียนในแผนการศึกษา ก1 และ ก3 ซึ่งเป็นแบบสอบต่างฉบับกัน เพราะแผนการศึกษา ก3 มีการ ใช้แบบสอบกลางภาคซึ่งต้องจัดการสอบก่อนแผนการศึกษา ก1 และ ก2 ที่มีเฉพาะการสอบปลายภาคเท่านั้น จึงจำเป็นต้องสร้างแบบสอบขึ้นใหม่เฉพาะหน่วยการเรียนรู้ที่ 1 – 7 ซึ่งต้องทำให้มีความแตกต่างจากเนื้อหา ของแบบสอบในหน่วยการเรียนรู้ที่ 1 – 7 ที่ใช้ในแผนการศึกษา ก1 และ ก2 ด้วยเหตุนี้ผู้วิจัยจึงต้องใช้ การเปรียบเทียบคะแนนซึ่งเป็นวิธีการทางสถิติในการนำคะแนนจากแบบทดสอบต่างฉบับกันแต่วัดในสิ่งเดียวกัน ไปแปลงให้มีหน่วยวัดเดียวกัน เพื่อให้คะแนนจากแบบสอบต่างฉบับกันสามารถเปรียบเทียบกันได้โดยตรง

(Angoff, 1984; Kanchanawasi, 1998; Lord, 1980; Petersen et al., 1982) และเลือกใช้วิธีการปรับเทียบคะแนนแบบอิกวิเปอร์เซ็นต์ไทล์ (Equipercentile equating) ซึ่งเป็นวิธีการปรับเทียบที่ยึดหลักการว่า การแจกแจงของคะแนนจากแบบสอบ X และแบบสอบ Y มีลักษณะคล้ายกัน การเทียบหาคะแนนสมมูลทำได้โดยใช้คะแนน ณ ตำแหน่งเปอร์เซ็นต์ไทล์เดียวกันของคะแนน 2 ชุดนั้นและเนื่องจาก ในงานวิจัยนี้มีนักศึกษาที่เข้าสอบจำนวน 5,598 คน ซึ่งเป็นตามข้อแนะนำในการเลือกใช้วิธีการปรับเทียบคะแนนแบบอิกวิเปอร์เซ็นต์ไทล์ของ Angoff (1984) ที่กล่าวว่าควรใช้วิธีการนี้กับกลุ่มผู้สอบขนาดใหญ่ที่มีความสามารถค่อนข้างกระจายและกระจายพอ ๆ กัน นอกจากนี้วิธีการปรับเทียบคะแนนแบบอิกวิเปอร์เซ็นต์ไทล์ถูกนำมาใช้ในงานวิจัยเพื่อปรับเทียบคะแนนสอบจากแบบทดสอบต่างฉบับกัน เช่น งานวิจัยของ Chanchusakul et al. (2017) ที่ทำการศึกษาเปรียบเทียบคุณภาพของการปรับเทียบคะแนนสอบภาษาอังกฤษระหว่างวิธีอิกวิเปอร์เซ็นต์ไทล์วิธีเชิงเส้นตรง และวิธีสมการถดถอยกับนักศึกษาปริญญาตรี มหาวิทยาลัยศิลปากร จำนวน 40 คน โดยใช้แบบสอบ Silpakorn English Proficiency Test (SEPT) และแบบสอบ Test of English for International Communication (TOEIC) และงานวิจัยของ Sun & Kim (2021) ที่ทำการศึกษาเรื่องการประเมินทกแนวทางในการจัดการกับคะแนนความถึเป็นศูนย์ โดยใช้วิธีการปรับเทียบคะแนนแบบอิกวิเปอร์เซ็นต์ไทล์ เป็นต้น ทั้งนี้ผู้วิจัยคาดหวังว่าคำตอบที่ได้จากการวิจัยจะเป็นประโยชน์ต่อหน่วยงานที่เกี่ยวข้องของมหาวิทยาลัยสุโขทัยธรรมาธิราช เช่น สำนักวิชาการ สามารถนำผลการวิจัยที่ได้ไปใช้เป็นแนวทางในการพัฒนาและปรับปรุงรูปแบบในการจัดกิจกรรมการเรียนการสอนให้สอดคล้องกับความต้องการของนักศึกษา และสำนักทะเบียนและวัดผลสามารถนำสารสนเทศที่ได้ไปใช้ในการพัฒนาและปรับปรุงรูปแบบและวิธีการวัดและประเมินผลที่สะท้อนสมรรถนะที่แท้จริงของนักศึกษา อันจะนำไปสู่การพัฒนานักศึกษาได้เต็มตามศักยภาพของนักศึกษาอย่างี่ควรจะเป็น และเป็นแนวทางในการวิจัยในด้านการเปรียบเทียบคุณภาพของรูปแบบการวัดและประเมินผลจากการใช้แบบสอบและการให้คะแนนที่แตกต่างกันให้กับงานวิจัยอื่น ๆ ต่อไป

ความมุ่งหมายของการวิจัย

- 1) เพื่อปรับเทียบคะแนนรวมที่ได้มาจากคะแนนการสอบและคะแนนที่ได้จากการทำกิจกรรมทั้งหมดของแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3
- 2) เพื่อเปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก2 กับ เปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก3
- 3) เพื่อเปรียบเทียบความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ และ ก3

ข้อมูลที่ใช้ในการวิจัย

ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิที่ได้มีการเก็บรวบรวมไว้แล้ว โดยได้รับการอนุเคราะห์ข้อมูลจากฝ่ายวัดผลการศึกษา สำนักทะเบียนและวัดผล และสำนักคอมพิวเตอร์ โดยข้อมูลที่ใช้ในการวิเคราะห์เป็นคะแนนกิจกรรม คะแนนสอบ และผลการศึกษาของนักศึกษาที่เข้าสอบในชุดวิชาที่ทำการวิจัย ในภาคเรียนที่ 1 ปีการศึกษา 2565 ของแผนการศึกษา ก1 ก2 และ ก3 ซึ่งพิจารณาเฉพาะชุดวิชาที่ผ่านตามเกณฑ์ในการคัดเลือกชุดวิชาที่กำหนดขึ้นโดยผู้วิจัย

เกณฑ์ในการคัดเลือกชุดวิชา ประกอบด้วย

- 1) เป็นชุดวิชาที่มีนักศึกษาเข้าสอบเป็นจำนวนมากกว่า 200 คนขึ้นไป และต้องเป็นชุดวิชาที่มีข้อมูลครบถ้วน สมบูรณ์
- 2) เป็นชุดวิชาที่เปิดให้ลงทะเบียนเรียนและมีการจัดสอบทั้งในแผนการศึกษา ก1 และ ก2 กับ ชุดวิชาที่เปิดให้ลงทะเบียนเรียนและมีการจัดสอบทั้งในแผนการศึกษา ก1 และ ก3
- 3) เป็นชุดวิชาที่เปิดให้ลงทะเบียนเรียนและมีการจัดสอบในแผนการศึกษา ก1 และต้องมีกิจกรรมประจำชุดวิชาและมีการตรวจให้คะแนนกิจกรรม สำหรับแผนการศึกษา ก2 จะต้องมีความเห็นจากกิจกรรมสอนเสริมแบบเก็บคะแนน และแผนการศึกษา ก3 จะต้องมีความเห็นจากกิจกรรมร่วมเรียนรู้ในชุดวิชา
- 4) เป็นชุดวิชาที่ผู้เข้าสอบมีจำนวนสูงสุด

จากเกณฑ์ทั้ง 4 ข้อข้างต้น ผู้วิจัยจึงได้ชุดวิชาที่นำมาวิเคราะห์เป็นชุดวิชาในกลุ่มสาขาวิชาทางสังคมศาสตร์ เช่น วิทยาการจัดการ ศิลปศาสตร์ เศรษฐศาสตร์ นิติศาสตร์ และรัฐศาสตร์ ทั้งนี้ชุดวิชาที่นำมาวิเคราะห์ดังกล่าว จำแนกตามแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3 ที่เปิดสอบร่วมกัน (อย่างละ 3 ชุดวิชา) แบบสอบของแต่ละชุดวิชาที่นำมาวิเคราะห์มีจำนวน 120 ข้อ 100 คะแนน มีนักศึกษารวมจำนวน 5,598 คน มาจากชุดวิชาจำแนกตามแผนการศึกษา ก1 และ ก2 ในชุดวิชา A สาขาวิชาศิลปศาสตร์ จำนวน 272 คน ชุดวิชา B สาขาวิชาวิทยาการจัดการ จำนวน 734 คน ชุดวิชา C สาขาวิชาเศรษฐศาสตร์ จำนวน 767 คน และมาจากชุดวิชาจำแนกตามแผนการศึกษา ก1 และ ก3 ในชุดวิชา D สาขาวิชารัฐศาสตร์ จำนวน 1,341 คน ชุดวิชา E สาขาวิชานิติศาสตร์ จำนวน 1,963 คน และ ชุดวิชา F สาขาวิชารัฐศาสตร์ จำนวน 521 คน

ขั้นตอนการดำเนินการวิจัย

1) ผู้วิจัยติดต่อขอรับข้อมูลจากฝ่ายวัดผลการศึกษา สำนักทะเบียนและวัดผล และสำนักคอมพิวเตอร์ โดยข้อมูลที่ได้รับ คือ คะแนนกิจกรรม คะแนนสอบและผลการศึกษาของนักศึกษาที่เข้าสอบในชุดวิชาที่ทำการวิจัย ในภาคเรียนที่ 1 ปีการศึกษา 2565 ของแผนการศึกษา ก1 ก2 และ ก3 โดยจำแนกตามชุดวิชาที่เปิดสอบตามแผนการศึกษา ก1 และ ก2 กับ ชุดวิชาที่เปิดสอบตามแผนการศึกษา ก1 และ ก3 แล้วจึงทำการตรวจสอบว่าข้อมูลที่ได้รับมีความครบถ้วน ถูกต้อง สมบูรณ์ หรือไม่

2) ผู้วิจัยทำการพิจารณาคัดเลือกข้อมูลคะแนนกิจกรรม คะแนนสอบ และผลการศึกษาของนักศึกษาที่เข้าสอบในชุดวิชาที่ทำการวิจัย ในภาคเรียนที่ 1 ปีการศึกษา 2565 ของแผนการศึกษา ก1 ก2 และ ก3 โดยจำแนกตามชุดวิชาที่เปิดสอบตามแผนการศึกษา ก1 และ ก2 กับ ชุดวิชาที่เปิดสอบตามแผนการศึกษา ก1 และ ก3 โดยพิจารณาเฉพาะชุดวิชาตามเกณฑ์ในการคัดเลือกชุดวิชาที่กำหนดขึ้นโดยผู้วิจัย

3) ผู้วิจัยนำข้อมูลที่ได้จากข้อ 2 มาจัดเตรียมข้อมูลให้อยู่ในรูปแบบไฟล์ Excel เพื่อใช้ในการวิเคราะห์เปรียบเทียบคะแนนรวมทั้งหมดของแผนการศึกษา ก1 และ ก2 กับ ก1 และ ก3 ด้วยวิธีการเปรียบเทียบคะแนนอิควิเปอร์เซ็นไทล์ (Equipercetile Equating) ตามข้อแนะนำในการเลือกใช้วิธีการเปรียบเทียบคะแนนแบบอิควิเปอร์เซ็นไทล์ของ Angoff (1984) ที่กล่าวว่าควรใช้วิธีการนี้กับกลุ่มผู้สอบขนาดใหญ่ที่มีความสามารถค่อนข้างกระจายและกระจายพอ ๆ กัน โดยใช้โปรแกรม Jamovi

4) ผู้วิจัยนำผลการเปรียบเทียบคะแนนที่ได้จากการวิเคราะห์ด้วยโปรแกรม Jamovi ในข้อ 3 ไปวิเคราะห์เปรียบเทียบสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และ ก2 กับ ก1 และ ก3 ด้วยสถิติ Mann – Whitney U test โดยใช้โปรแกรม SPSS

5) ผู้วิจัยนำผลการเปรียบเทียบคะแนนที่ได้จากข้อ 4 มาตัดเกรดตามเกณฑ์ของมหาวิทยาลัย และนำผลการวิเคราะห์ที่ได้เฉพาะกลุ่มนักศึกษาที่สอบผ่าน ได้เกรด S และ H มาใช้ในการวิเคราะห์เปรียบเทียบความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับคะแนนหลังจากการปรับเทียบคะแนนตามแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3 ด้วยสถิติสัมประสิทธิ์แคปปา โดยใช้โปรแกรม SPSS

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการวิจัยนี้เป็นการวิเคราะห์ข้อมูลเชิงปริมาณ ได้แก่ การปรับเทียบคะแนนอิควิเปอร์เซ็นไทล์ (Equipercetile Equating) การทดสอบ Mann – Whitney U test และ การทดสอบความสอดคล้อง ตามรายละเอียดดังนี้

1. วิเคราะห์คะแนนรวมทั้งหมดที่ได้จากคะแนนสอบและคะแนนกิจกรรมทั้งหมดของแผนการศึกษา ก1 ก2 และ ก3 โดยวิเคราะห์แยกเป็นรายชุดวิชาที่เปิดสอบพร้อมกันตามแผนการศึกษา ก1 และ ก2 กับ ก1 และ ก3 โดยเลือกชุดวิชาที่มีจำนวนนักศึกษาเข้าสอบมากที่สุด แผนการศึกษาละ 3 ชุดวิชา รวม 6 ชุดวิชา และพิจารณาเฉพาะชุดวิชาที่ผ่านตามเกณฑ์ในการคัดเลือกชุดวิชาที่กำหนดขึ้นโดยผู้วิจัย จากนั้นจึงวิเคราะห์ด้วยการปรับเทียบคะแนนอิควิเปอร์เซ็นไทล์ โดยใช้โปรแกรม Jamovi

2. วิเคราะห์เปรียบเทียบสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 ก2 และ ก3 ด้วยการทดสอบ Mann – Whitney U test โดยใช้โปรแกรม SPSS เพื่อทดสอบสมมติฐานข้อที่ 1 : สัดส่วนนักศึกษาที่ได้ผลการตัดเกรดว่าสอบผ่านจากแผนการศึกษา ก1 แตกต่างกับสัดส่วนนักศึกษาที่ได้ผลการตัดเกรดว่าสอบผ่านจากแผนการศึกษา ก2 และเพื่อทดสอบสมมติฐานข้อที่ 2: สัดส่วนนักศึกษาที่ได้ผลการตัดเกรดว่าสอบผ่านจากแผนการศึกษา ก1 แตกต่างกับสัดส่วนนักศึกษาที่ได้ผลการตัดเกรดว่าสอบผ่านจากแผนการศึกษา ก3

3. วิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3 ด้วยสถิติสัมประสิทธิ์แคปปา (Kappa Coefficient) และใช้เกณฑ์การพิจารณาระดับความสอดคล้องของสถิติแคปปาของ McHugh (2012, as cited in Pasunon, 2015) ดังตาราง 1

ตาราง 1 ระดับความสอดคล้องของสถิติแคปปาตามแนวทางของ McHugh (2012)

ค่าสถิติ Kappa	ระดับความสอดคล้องระหว่างผู้ประเมิน	ร้อยละของความเชื่อมั่น
0.00 – 0.19	ไม่มีความสอดคล้อง (none)	0 – 3
0.20 – 0.39	ความสอดคล้องน้อยมาก (minimal)	4 – 15
0.40 – 0.59	ความสอดคล้องน้อย (weak)	16 – 35
0.60 – 0.79	ความสอดคล้องปานกลาง (moderate)	36 – 63
0.80 – 0.89	ความสอดคล้องมาก (strong)	64 – 80
0.90 – 1.00	ความสอดคล้องมากที่สุด (almost perfect)	81 – 100

โดยใช้โปรแกรม SPSS เพื่อทดสอบสมมติฐานข้อที่ 3 : ผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก2 มีความสอดคล้องกัน และทดสอบสมมติฐานข้อที่ 4 : ผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก3 มีความสอดคล้องกัน

ผลการวิจัย

1. ผลการปรับเทียบคะแนนรวมที่ได้มาจากคะแนนการสอบและคะแนนที่ได้จากการทำกิจกรรมทั้งหมดของแผนการศึกษา ก1 และ ก2 กับ แผนการศึกษา ก1 และ ก3

1.1 ผลการปรับเทียบคะแนนรวมที่ได้มาจากคะแนนการสอบและคะแนนที่ได้จากการทำกิจกรรมทั้งหมดของแผนการศึกษา ก1 และ ก2 พบว่า ชุดวิชา A สาขาวิชาศิลปศาสตร์ ชุดวิชา B สาขาวิชาวิทยาการจัดการ และ ชุดวิชา C สาขาวิชาเศรษฐศาสตร์ และผลการปรับเทียบคะแนนในภาพรวมทั้ง 3 ชุดวิชา จาก 3 สาขาวิชา เห็นได้ว่าให้ผลคะแนนที่สอดคล้องกันและเป็นไปในทิศทางเดียวกัน ซึ่งผลการปรับเทียบคะแนนมีความสอดคล้องกับลักษณะของแบบสอบจากแผนการศึกษา ก1 และ ก2 ที่มหาวิทยาลัยสุโขทัยธรรมาธิราชได้กำหนดไว้ว่า แบบสอบจากแผนการศึกษา ก1 และ ก2 เป็นแบบสอบฉบับเดียวกัน ผลที่ได้จากการปรับเทียบค่าคะแนนจึงมีค่าใกล้เคียงกัน หรือแตกต่างกันไม่มาก ทั้งในส่วนของคะแนนต่ำสุด 3 อันดับแรก และคะแนนสูงสุด 3 อันดับแรก และในช่วงกลางของช่วงคะแนนทั้งหมดที่นำมาปรับเทียบ ดังตัวอย่างผลการวิเคราะห์จากชุดวิชา C สาขาวิชาเศรษฐศาสตร์ ตาราง 2

ตาราง 2 ผลการเปรียบเทียบคะแนนรวมที่ได้มาจากคะแนนการสอบและคะแนนที่ได้จากการทำกิจกรรมทั้งหมดของชุดวิชา C สาขาวิชาเศรษฐศาสตร์ ในแผนการศึกษา ก1 และ ก2

ความสอดคล้องของคะแนน (Concordance)			ความสอดคล้องของคะแนน (Concordance)		
คะแนนจากแผนการศึกษา ก2	คะแนนจากแผนการศึกษา ก2 ที่เปรียบเทียบกับคะแนนจากแผนการศึกษา ก1	คะแนนจากแผนการศึกษา ก1	คะแนนจากแผนการศึกษา ก2	คะแนนจากแผนการศึกษา ก2 ที่เปรียบเทียบกับคะแนนจากแผนการศึกษา ก1	คะแนนจากแผนการศึกษา ก1
18	30	28	61	41.6	42
20	30.5	29	63	41.9	42
43	31.4	30	65	42.4	42
45	33.3	31	67	44.8	45
48	33.8	33	67	45.7	46
50	34.3	34	68	46.7	47
56	36.2	34	68	47.6	48
57	38.1	38	68	50.5	50
59	38.4	38	76	55.2	55
59	38.7	39	76	60	60
59	41	39	80	61	61
60	41.3	41	82	65.7	66

1.2 ผลการเปรียบเทียบคะแนนรวมที่ได้มาจากคะแนนการสอบและคะแนนที่ได้จากการทำกิจกรรมทั้งหมดของแผนการศึกษา ก1 และ ก3 ใน 3 ชุดวิชา จาก 3 สาขาวิชา ประกอบด้วย ชุดวิชา D สาขาวิชา รัฐศาสตร์ ชุดวิชา E สาขาวิชานิติศาสตร์ และชุดวิชา F สาขาวิชา รัฐศาสตร์ เห็นได้ว่าให้ผลคะแนนที่สอดคล้องกันและเป็นไปในทิศทางเดียวกัน คือผลการเปรียบเทียบคะแนนมีความสอดคล้องกับลักษณะของแบบสอบจากแผนการศึกษา ก1 และ ก3 ที่มหาวิทยาลัยสุโขทัยธรรมาธิราช ได้กำหนดไว้ว่า แบบสอบจากแผนการศึกษา ก1 และ ก3 มีการใช้ข้อสอบร่วมกันในการสอบปลายภาค (หน่วยที่ 8 – 15) แต่ข้อสอบจะแตกต่างกันในส่วน of แบบสอบกลางภาคของแผนการศึกษา ก3 ที่ออกข้อสอบจากหน่วยที่ 1 – 7 แต่แบบสอบของแผนการศึกษา ก1 ในหน่วยที่ 1 – 7 จะเป็นข้อสอบอีกชุดหนึ่งที่น่าสนใจในการสอบปลายภาคร่วมกับหน่วยที่ 8 – 15 รวมเป็น 15 หน่วย คะแนนจากการเปรียบเทียบจึงมีค่าแตกต่างกันมากในบางช่วง และมีค่าใกล้เคียงกันในบางช่วง ทั้งในส่วน of คะแนนต่ำสุด 3 อันดับแรก และคะแนนสูงสุด 3 อันดับแรก และในช่วงกลางของช่วงคะแนนทั้งหมดที่นำมาเปรียบเทียบ ดังตัวอย่างผลการวิเคราะห์จากชุดวิชา F สาขาวิชา รัฐศาสตร์ ตาราง 3

ตาราง 3 ผลการวิเคราะห์เปรียบเทียบคะแนนรวมที่ได้มาจากคะแนนการสอบและคะแนนที่ได้จากการทำ
กิจกรรมทั้งหมดของชุดวิชา F สาขาวิชารัฐศาสตร์ ของแผนการศึกษา ก1 และ ก3

ความสอดคล้องของคะแนน (Concordance)			ความสอดคล้องของคะแนน (Concordance)		
คะแนนจาก แผนการ ศึกษา ก3	คะแนนจาก แผนการศึกษา ก3 ที่ปรับเทียบกับ คะแนนจาก แผนการศึกษา ก1	คะแนนจาก แผนการ ศึกษา ก1	คะแนนจาก แผนการ ศึกษา ก3	คะแนนจาก แผนการศึกษา ก3 ที่ปรับเทียบกับ คะแนนจาก แผนการศึกษา ก1	คะแนนจาก แผนการศึกษา ก1
15	27.5	21	66	36.7	37
18	28.3	23	67	39.2	39
29	30	26	68	42.5	43
33	30.4	28	68	43.3	43
37	30.8	28	69	44.2	45
42	31.3	28	69	44.2	45
47	31.7	30	70	45	45
54	32.1	31	71	45.4	46
56	33.1	31	73	46.7	47
57	33.6	32	76	54.2	54
61	35	35	78	55	55
62	36.3	36	85	76.7	77

2. ผลการเปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และ สัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก2 กับ ความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และ สัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก3

2.1 ผลการเปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และ สัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก2 พบว่า สัดส่วนของนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของชุดวิชา A สาขาวิชาศิลปศาสตร์ ชุดวิชา B สาขาวิชาวิทยาการจัดการ และชุดวิชา C สาขาวิชาเศรษฐศาสตร์ ตามแผนการศึกษา ก1 และ ก2 หลังจากการปรับเทียบคะแนนแล้ว มีสัดส่วนการสอบผ่านไม่แตกต่างกัน แสดงว่าจำนวนนักศึกษาที่สอบผ่านจากแผนการศึกษา ก1 และ ก2 มีจำนวนใกล้เคียงกัน ดังตัวอย่างผลการวิเคราะห์จากชุดวิชา A สาขาวิชาศิลปศาสตร์ ตาราง 4

ตาราง 4 ผลการเปรียบเทียบสัดส่วนนักศึกษาที่สอบผ่านโดยพิจารณาจากผลการตัดเกรดของชุดวิชา A
สาขาวิชาศิลปศาสตร์ ของแผนการศึกษา ก1 และ ก2

	Mann – Whitney U test	Sig. (2 – tailed)
นักศึกษาที่สอบผ่าน (เกรด S และ H)	43.000	1.000

2.2 ผลการเปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และ สัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก3 พบว่า สัดส่วนของนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของชุดวิชา D สาขาวิชารัฐศาสตร์ ชุดวิชา E สาขาวิชานิติศาสตร์ และชุดวิชา F สาขาวิชารัฐศาสตร์ ตามแผนการศึกษา ก1 และ ก3 หลังจากการปรับเทียบคะแนนแล้ว มีสัดส่วนการสอบผ่านไม่แตกต่างกัน แสดงว่า จำนวนนักศึกษาที่สอบผ่านจากแผนการศึกษา ก1 และ ก3 มีจำนวนใกล้เคียงกันดังตัวอย่างผลการวิเคราะห์จากชุดวิชา D สาขาวิชารัฐศาสตร์ ตาราง 5

ตาราง 5 ผลการเปรียบเทียบสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของชุดวิชา D สาขาวิชารัฐศาสตร์ ของแผนการศึกษา ก1 และ ก3

	Mann – Whitney U test	Sig. (2 – tailed)
นักศึกษาที่สอบผ่าน (เกรด S และ H)	69.000	1.000

3. ผลการวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนน กับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก2 กับความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก3

3.1 ผลการวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก2 ของชุดวิชา B สาขาวิชาวิทยาการจัดการ พบว่า ค่า Sig. มีค่าเท่ากับ 0.003 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญ 0.05 แสดงว่าผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนมีความสอดคล้องกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนน ซึ่งสอดคล้องกับค่าสถิติ Kappa ที่มีค่าเท่ากับ 0.893 ซึ่งมีค่าอยู่ในช่วง 0.80 – 0.89 แสดงว่าผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนมีความสอดคล้องกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนน อยู่ในระดับมาก และมีร้อยละของความเชื่อมั่นอยู่ระหว่าง 64 – 80 ดังตัวอย่างผลการวิเคราะห์จากชุดวิชา B สาขาวิชาวิทยาการจัดการ ตาราง 6

ตาราง 6 ผลการวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนน กับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนน ของชุดวิชา B สาขาวิชาวิทยาการจัดการ ของแผนการศึกษา ก1 และ ก2

จำนวนนักศึกษาที่สอบผ่านที่ใช้ในการประเมิน (คน)	Kappa	Sig.
184	0.893	0.003

3.2 ผลการวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนนของแผนการศึกษา ก1 และ ก3 ของชุดวิชา E สาขาวิชานิติศาสตร์ พบว่า ค่า Sig. มีค่าเท่ากับ 0.033 ซึ่งมีความน้อยกว่าระดับนัยสำคัญ 0.05 แสดงว่าผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนมีความสอดคล้องกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนน ซึ่งสอดคล้องกับค่าสถิติ Kappa ที่มีค่าเท่ากับ 0.830 ซึ่งมีค่าอยู่ในช่วง 0.80 – 0.89 แสดงว่าผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนมีความสอดคล้องกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนน อยู่ในระดับมาก และมีร้อยละของความเชื่อมั่นอยู่ระหว่าง 64 – 80 ดังตัวอย่างผลการวิเคราะห์จากชุดวิชา E สาขาวิชานิติศาสตร์ ตาราง 7

ตาราง 7 ผลการวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับผลการตัดเกรดจากการใช้คะแนนหลังจากการปรับเทียบคะแนน ของชุดวิชา E สาขาวิชานิติศาสตร์ ของแผนการศึกษา ก1 และ ก3

จำนวนนักศึกษาที่สอบผ่านที่ใช้ในการประเมิน (คน)	Kappa	Sig.
145	0.830	0.033

อภิปรายผล

การอภิปรายผลการวิจัยในเบื้องต้น แบ่งออกเป็น 2 ประเด็น ดังนี้ 1) การปรับเทียบคะแนนของแบบสอบ และ 2) การวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับคะแนนหลังจากการปรับเทียบคะแนน

1) การปรับเทียบคะแนนของแบบสอบ สำหรับการวิจัยครั้งนี้ใช้วิธีการปรับเทียบคะแนนด้วยวิธีอิควิเปอร์เซ็นต์ไทล์ (Equipercentile Equating) ซึ่งเป็นวิธีการที่อยู่ในรูปแบบการปรับเทียบคะแนนโดยใช้ทฤษฎีดั้งเดิม วิธีนี้ควรใช้ในกรณีที่กลุ่มผู้สอบมีขนาดใหญ่ ซึ่งเหมาะสมกับงานวิจัยนี้ที่มีจำนวนนักศึกษาที่เข้าสอบจำนวน 5,598 คน และควรเป็นกลุ่มที่มีความสามารถค่อนข้างกระจายและกระจายพอ ๆ กัน ซึ่งสอดคล้องกับลักษณะของกลุ่มผู้สอบซึ่งเป็นนักศึกษาระดับปริญญาตรีของมหาวิทยาลัยสุโขทัยธรรมมาธิราชที่เป็นกลุ่มผู้สอบที่มีขนาดใหญ่ ความสามารถของนักศึกษามีลักษณะกระจาย และกระจายพอ ๆ กัน และยังสอดคล้องกับ Meng (2012); Hanson & Béguin (2002, as cited in Meng, 2012) ที่กล่าวว่า ถ้ากลุ่มตัวอย่างในการวิเคราะห์มีขนาดใหญ่ ย่อมเป็นตัวแทนประชากรที่ดีกว่าและทำให้การประมาณค่าได้ดีกว่า เมื่อพิจารณาผลคะแนนที่ได้จากการปรับเทียบคะแนนพบว่ามีความสอดคล้องกับลักษณะของแบบสอบจากแผนการศึกษา ก1 และ ก2 ที่มหาวิทยาลัยสุโขทัยธรรมมาธิราช ได้กำหนดไว้ว่า แบบทดสอบจากแผนการศึกษา ก1 และ ก2 เป็นแบบสอบฉบับเดียวกัน ผลที่ได้จากการปรับค่าคะแนนจึงมีค่าแตกต่างกันไม่มาก ซึ่งมีความแตกต่างกับผลการปรับเทียบคะแนนของแบบสอบที่ใช้กับการสอบตามแผนการศึกษา ก1 และ ก3 ที่มีการใช้ข้อสอบร่วมกันในการสอบปลายภาค (หน่วยที่ 8 – 15) แต่ข้อสอบจะแตกต่างกันในส่วน of แบบสอบกลางภาคของแผนการศึกษา ก3 ที่ออกข้อสอบจากหน่วยที่ 1 – 7 โดยคะแนนจากการปรับเทียบจะมีค่าแตกต่างกันมากในบางช่วง และมีค่าใกล้เคียงกันในบางช่วง

2) การวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับคะแนนหลังจากการปรับเทียบคะแนน ผลการวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับคะแนนหลังจากการปรับเทียบคะแนนตามแผนการศึกษา ก2 ของแต่ละชุดวิชาที่มาจากต่างสาขาวิชา พบว่ามีความสอดคล้องอยู่ในระดับมากและระดับมากที่สุด เช่นเดียวกับผลการวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับคะแนนหลังจากการปรับเทียบคะแนนตามแผนการศึกษา ก3 ที่มีความสอดคล้องอยู่ในระดับปานกลางถึงระดับมากที่สุด สอดคล้องกับ Chutinantakul et al. (2018) ที่กล่าวว่าการหาความสัมพันธ์ของการตัดเกรดหลังการปรับเทียบคะแนนทั้ง 3 ระดับ และ 8 ระดับ ด้วยวิธีเคอเนลและวิธี IRT 2 พารามิเตอร์ พบว่ามีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติทุกเงื่อนไข โดยที่การตัดเกรด 3 ระดับ หลังการปรับเทียบคะแนนด้วยวิธีเคอเนลและวิธี IRT 2 พารามิเตอร์ ด้วยเงื่อนไขข้อสอบรวมมีความยากอย่างสม่ำเสมอ กับทุกขนาดตัวอย่างมีความสัมพันธ์กันในระดับดีมาก ส่วนเงื่อนไขข้อสอบรวมมีความยากอย่างสม่ำเสมอและตัดข้อสอบที่ไม่มีคุณภาพทิ้ง มีความสัมพันธ์อยู่ในระดับดี ส่วนการตัดเกรด 8 ระดับ หลังการปรับเทียบคะแนนด้วยวิธีเคอเนลและวิธี IRT 2 พารามิเตอร์ มีเพียงเงื่อนไขข้อสอบรวมมีความยากอย่างสม่ำเสมอเพียงเงื่อนไขเดียว ที่การตัดเกรดหลังการปรับเทียบคะแนนของทั้งสองวิธีมีความสัมพันธ์กันในระดับดี ทั้งนี้อาจเป็นไปได้ว่าบางชุดวิชาที่เป็นวิชาที่ค่อนข้างยาก เช่น ชุดวิชาที่มีการคำนวณ ชุดวิชาที่เป็นภาษาอังกฤษ เป็นต้นลักษณะเนื้อหาของชุดวิชาเหล่านี้เป็นชุดวิชาที่มีความยากสำหรับนักศึกษา ถ้าในชุดวิชาหนึ่ง ๆ หากเกรดที่ได้จากคะแนนก่อนการปรับเทียบน้อย เกรดที่ได้จากคะแนนหลังจากการปรับเทียบจะมีค่าน้อยเช่นกัน เมื่อพิจารณาความสอดคล้องแล้วก็จะพบว่าเกรดที่ได้ก่อนการปรับเทียบคะแนนและหลังการปรับเทียบคะแนนมีความสอดคล้องกันสูง สำหรับผลการวิเคราะห์ความสอดคล้องของผลการตัดเกรดจากการใช้คะแนนก่อนการปรับเทียบคะแนนกับคะแนนหลังจากการปรับเทียบคะแนนตามแผนการศึกษา ก3 ในบางชุดวิชาที่มีค่าความสอดคล้องอยู่ในระดับปานกลางอาจเป็นไปได้ว่าในชุดวิชาดังกล่าว นั้น มีนักศึกษาลงทะเบียนเรียนในแผนการศึกษาดังกล่าวค่อนข้างน้อย เมื่อพิจารณาค่าความสอดคล้องจึงอยู่ในระดับปานกลาง หากนักศึกษาลงทะเบียนเรียนในแผนการศึกษา ก3 มากขึ้น หรือนักศึกษามีการส่งกิจกรรมครบทุกคน จะทำให้ทุกคนมีคะแนนกิจกรรมเพิ่มขึ้น เมื่อพิจารณาความสอดคล้องของเกรดที่ได้ อาจทำให้ค่าความสอดคล้องอยู่ในระดับที่สูงกว่าระดับปานกลาง

3) ผลการเปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก2 กับ เปรียบเทียบความแตกต่างระหว่างสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก1 และสัดส่วนนักศึกษาที่สอบผ่าน โดยพิจารณาจากผลการตัดเกรดของแผนการศึกษา ก3 พบว่า มีจำนวนนักศึกษาที่สอบผ่านไม่แตกต่างกัน ทั้งนี้ น่าจะมีสาเหตุมาจากความยากของข้อสอบในแต่ละแผนการศึกษามีความยากเฉลี่ยใกล้เคียงกัน ซึ่งสอดคล้องกับงานวิจัยของ Juimounsri & Wjitwanna (2019) ที่กล่าวว่าระดับความยากของแบบทดสอบส่งผลต่อการสอบผ่านของนักศึกษา แม้ว่าความยากของแบบทดสอบอาจมีความยากในบางภาคการศึกษา แต่โดยภาพรวมแบบทดสอบแต่ละฉบับมีความยากใกล้เคียงกัน เพราะมีการดำเนินการทำฉบับแบบทดสอบมีการดำเนินการในรูปของคณะกรรมการ โดยประธานชุดเป็นผู้พิจารณา

ความถูกต้องตามเนื้อหาของข้อสอบ และนักวัดผลเป็นผู้ตรวจสอบตามเทคนิคด้านการวัดผล แบบทดสอบจึงมีคุณภาพและมาตรฐาน ซึ่งสะท้อนให้เห็นว่าเป็นเรื่องที่ดีเพราะแสดงให้เห็นว่าแบบทดสอบที่นำมาใช้ในการสอบนั้นมีคุณภาพใกล้เคียงกัน

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 เนื่องจากความแตกต่างระหว่างจำนวนนักศึกษาที่สอบผ่านในแต่ละแผนการศึกษาไม่แตกต่างกัน สะท้อนให้เห็นว่าแผนการศึกษาในแต่ละแผนมีคุณภาพเท่าเทียมกัน ดังนั้นมหาวิทยาลัยสุโขทัยธรรมมาธิราชสามารถนำผลการวิจัยไปใช้เป็นข้อมูลในการประชาสัมพันธ์ให้แก่บุคคลทั่วไปที่สนใจจะสมัครเรียนและเลือกแผนการศึกษากับทางมหาวิทยาลัยสุโขทัยธรรมมาธิราชและนักศึกษาปัจจุบันที่กำลังตัดสินใจจะเลือกแผนการศึกษาใหม่ สามารถเลือกศึกษาในแผนการศึกษาใดก็ได้ตามความสนใจ

1.2 การวิจัยครั้งนี้ใช้ข้อมูลคะแนนสอบและผลการสอบของนักศึกษาเพียงบางชุดวิชา จากบางสาขาวิชาใช้ในการศึกษาเท่านั้น ซึ่งเป็นชุดวิชาทางสังคมศาสตร์ที่มีนักศึกษาเข้าสอบเป็นจำนวนมาก แต่อาจจะไม่เป็นตัวแทนของทุกชุดวิชาหรือสาขาวิชาอื่น ๆ ที่ไม่ได้นำมาวิเคราะห์ ดังนั้นในการนำผลการวิจัยไปใช้ในการอ้างอิงจึงควรพิจารณาในกรณีนี้ด้วย

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรใช้วิธีการปรับเทียบคะแนนวิธีอื่น เช่น Linear Equating, Rasch Model หรือวิธีการอื่นใดที่มีความเหมาะสมกับลักษณะของแบบทดสอบที่จะนำมาใช้ในงานวิจัย ตลอดจนเงื่อนไขของแต่ละวิธี

2.2 ควรได้มีการตรวจสอบคุณภาพของข้อสอบที่จะนำมาใช้ในการประเมินทุกครั้ง ทั้งในเรื่องของความตรง ความเที่ยง ความยากง่าย และอำนาจจำแนก ของแบบทดสอบทุกครั้งก่อนนำไปใช้ เพื่อเป็นการสร้างความเชื่อมั่นในเรื่องของผลการประเมินว่ามีความถูกต้องและสามารถนำมาเปรียบเทียบกันได้

References

- Angoff, W. H. (1984). *Scales, Norms, and Equivalent Scores*. Educational Testing Service.
- Hanson, B. A., & Béguin, A. A. (2002). Obtaining a common scale for item response theory item parameters using separate versus concurrent estimation in the common-item equating design. *Applied Psychological Measurement*, 26(1), 3–24.
<https://doi.org/10.1177/0146621602026001001>
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Routledge.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282. <https://doi.org/10.11613/BM.2012.031>

- Meng, Y. (2012). *Comparison of Kernel equating and item response theory equating methods* [Doctoral dissertation]. University of Massachusetts.
<https://search.proquest.com/docview/1033227222>
- Petersen, N. S., Marco, C. L., & Stewart, E. E. (1982). *A test of the adequacy of linear Score Equating Models*. Academic Press.
- Sun, T. & Kim, S. Y. (2021). Evaluating Six Approaches to Handling Zero-Frequency Scores under Equipercentile Equating. *Measurement: Interdisciplinary Research and Perspectives*, 19(4), 213 – 235.

Translate Thai References

- Chanchusakul, S., Paiwittayasiritham, C., Phonphanthin, Y., & Suphanophap, P. (2017). Comparing the quality of English language test scores equating between the equipercentile method straight lines method and regression equation method. *Academic journal Veridian E-Journal, Silpakorn University Humanities Social Sciences and Arts field*, 10(2), 2444–2455. (in Thai)
- Chutinantakul, S., Wongnam, P., & Panhun, S. (2018). Results of score calibration using the kernel method and IRT method under different conditions. *Journal of Education Sukhothai Thammathirat Open University*, 11(1), 294–306. (in Thai)
- Juimounsri, S., & Wijitwanna, S. (2019). *Factors Affecting Test Scores of STOU Students Who Take Walk-in Examination via Computer* [Research report]. Institute of Research and Development Sukhothai Thammathirat Open University. (in Thai)
- Kanchanawasi, S. (1998). *Score comparison between tests (Test Equating)*. Textbook Center and academic documents Faculty of Education Chulalongkorn University. (in Thai)
- Pasunon, P. (2015). Assessment of confidence between evaluators using Kappa coefficient statistics. *Journal of Applied Liberal Arts King Mongkut's University of Technology North Bangkok*, 8(1), 2–20. (in Thai)
- Sukhothai Thammathirat Open University. (n.d.). *Teaching Management*.
<https://www.stou.ac.th/main/StouPlan.html> (in Thai)

การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกเพื่อการสื่อสารระหว่างนิสิตครู
กับนักเรียนที่มีความต้องการพิเศษ: การประยุกต์โมเดลซีเอ็นเอ็นร่วมกับอาร์เอ็นเอ็น
Development of Thai sign language recognition model using deep learning for
communication between student teachers and student with special needs:
application of CNN and RNN models

ธีรวิทย์ คะรานรัมย์^{1*} และ สิวะโชติ ศรีสุทธิยากร²

Teerawit Karanram^{1*} and Siwachoat Srisuttiyakorn²

(Received: October 17, 2024; Revised: November 29, 2024; Accepted: December 2, 2024)

บทคัดย่อ

การวิจัยนี้มีจุดมุ่งหมาย 1) เพื่อวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน และ 2) เพื่อพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกตัวอย่างวิจัย ได้แก่ ครูการศึกษาพิเศษในโรงเรียนสอนคนหูหนวก และอาสาสมัครที่มีความสามารถในการใช้ภาษามือ เครื่องมือวิจัย ได้แก่ แบบสอบถามความถี่การใช้งานประโยคภาษามือ การวิเคราะห์ข้อมูลแบ่งออกเป็น 2 ตอน ตอนที่ 1 การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียนจากแบบสอบถามโดยใช้ค่าเฉลี่ยและร้อยละ ตอนที่ 2 การพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก ผู้วิจัยเลือกใช้โมเดลการเรียนรู้เชิงลึก convolutional neural network (CNN) ร่วมกับ recurrent neural network (RNN) จำนวน 4 โมเดล ในการพัฒนาโมเดลรู้จำภาษามือไทยจำนวน 20 ประโยค โดยใช้ข้อมูลจากวิดีโอที่ถ่ายทำกับอาสาสมัคร จากนั้นเปรียบเทียบประสิทธิภาพของโมเดลโดยพิจารณาจากความถูกต้อง ขนาดของโมเดล และเวลาที่ใช้ในการฝึกสอนโมเดล เพื่อนำโมเดลที่มีประสิทธิภาพสูงสุดไปใช้ประเมินประสิทธิภาพของโมเดลในการจำแนกประโยคภาษามือในภาพรวมและในแต่ละประโยคโดยใช้ค่าความถูกต้อง ค่าความแม่นยำ ค่าการเรียกคืน และคะแนน F1

ผลการวิจัยพบว่า 1) ภาษามือที่ใช้ในชั้นเรียนมักเป็นประโยคที่สั้นและกะทัดรัดมากกว่าภาษาพูด และภาษามือที่ใช้มากที่สุดจัดอยู่ในหมวดการเสริมแรง 2) การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้โมเดล EfficientNetB0 ร่วมกับ GRU สามารถจำแนกประโยคภาษามือได้มีประสิทธิภาพสูงสุดทั้งในด้านความถูกต้อง

¹ นิสิตปริญญาโท สาขาวิชาวิธีวิทยาการพัฒนานวัตกรรมทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

² ผู้ช่วยศาสตราจารย์ ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹ Master's degree student, Methodology for Innovation Development in Education Program, Faculty of Education, Chulalongkorn University

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

* Corresponding Author E-mail: teerawit.karanram@gmail.com

ขนาดของโมเดล และเวลาที่ใช้ในการฝึกสอนโมเดล และ 3) การประเมินประสิทธิภาพของโมเดลรู้จำภาษามือไทยโดยใช้ข้อมูลชุดทดสอบ โมเดลมีความถูกต้องร้อยละ 73 โดยประโยคที่โมเดลสามารถจำแนกได้ดี ได้แก่ ภูมิใจ พร้อมยัง ดูภาพ และเข้าใจ และประโยคที่โมเดลจำแนกได้ไม่ดี ได้แก่ งานสวย ทำได้ไหม และถูกหรือผิด

คำสำคัญ การรู้จำภาษามือ การเรียนรู้เชิงลึก ภาษามือไทย

Abstract

This research aims to 1) analyze the Thai sign language used by teachers to communicate with students with special needs in the classroom, and 2) develop and evaluate the performance of a Thai sign language recognition model using deep learning. The research subjects include special education teachers in schools for the deaf and volunteers proficient in sign language. The research tools include a questionnaire on the frequency of sign language sentences. Data analysis is divided into two parts: Part 1 involves analyzing the Thai sign language used by teachers to communicate with students with special needs in the classroom using mean and percentage from the questionnaire data. Part 2 involves developing and evaluating the performance of a Thai sign language recognition model using deep learning. The researcher selected four deep learning models combining CNN and RNN to develop a model that recognizes 20 Thai sign language sentences using video data recorded with volunteers. The models' performance was compared in terms of accuracy, model size, and training time, to identify the most effective model for evaluating the overall and individual sentence recognition using accuracy, precision, recall, and F1-score.

The research findings revealed that 1) the sign language used in the classroom tends to be shorter and more concise than spoken language, with reinforcement phrases being the most commonly used; 2) the development of a Thai sign language recognition model using the EfficientNetB0 combined with GRU achieved the highest performance in terms of accuracy, model size, and training time; and 3) the evaluation of the Thai sign language recognition model's performance using a test dataset showed that the model achieved an accuracy of 73%. Sentences that the model recognized well included "proud," "are you ready?" "look at the picture," and "understand," while sentences it struggled with included "beautiful work," "can you do it?" and "right or wrong."

Keywords: sign language recognition, deep learning, Thai sign language

บทนำ

ภาษามือเป็นภาษาที่ใช้ในการสื่อสารระหว่างคนหูหนวกด้วยกันเองและคนที่มีความสามารถในการแปลภาษามือ ซึ่งอาศัยองค์ประกอบของภาษามือ ได้แก่ ท่ามือ ตำแหน่งของมือ การเคลื่อนไหวของมือ การหันฝ่ามือ และการแสดงออกทางสีหน้า เพื่อใช้ในการสื่อสารแทนเสียงพูด อย่างไรก็ตามภาษามือของแต่ละชุมชนหรือท้องถิ่นจะมีลักษณะแตกต่างกันไปตามวัฒนธรรม ขนบธรรมเนียมประเพณี และลักษณะทางภูมิศาสตร์ เช่นเดียวกับภาษาพูดที่มีคำศัพท์ความหมายเหมือนกันแตกต่างกันในแต่ละท้องถิ่น (Boonya, 2008; Buangamdee, 2017; Petchporee, 2005)

ภาษามือนอกจากจะเป็นเครื่องมือการสื่อสารของคนหูหนวกแล้ว ภาษามือยังเป็นเครื่องมือสำคัญในการเรียนรู้ของคนหูหนวก การสื่อสารระหว่างครูและนักเรียนหูหนวกเป็นปัจจัยสำคัญต่อผลลัพธ์การเรียนรู้ โดยเฉพาะการจัดการเรียนรู้แบบเรียนรวม (inclusive education) การขาดครูการศึกษาพิเศษหรือครูที่มีความรู้ด้านภาษามืออาจทำให้การสื่อสารระหว่างครูกับนักเรียนไม่มีประสิทธิภาพเท่าที่ควร และทำให้การจัดการเรียนการสอนขาดคุณภาพ (Kelly et al., 2020) ซึ่งครูการศึกษาพิเศษในประเทศไทยมีจำนวนที่จำกัด ถ้าหากมีเครื่องมือหรือตัวช่วยในการรู้จำภาษามือให้กับครูหรือนิสิตนักศึกษาครูที่ไม่มีพื้นฐานภาษามือก็จะช่วยลดช่องว่างในการสื่อสารระหว่างครูกับนักเรียนหูหนวก และช่วยให้นักเรียนมีผลลัพธ์การเรียนรู้ที่ดีขึ้น

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับภาษามือพบว่า มีการนำเทคโนโลยีเกี่ยวกับการรู้จำภาษามือ (sign language recognition) เข้ามาใช้เป็นเครื่องมือที่ช่วยในการให้ความหมายของท่าภาษามือเป็นข้อความหรือคำพูด โดยรูปแบบของภาษามือที่ถูกใช้ในการศึกษามืออยู่สองรูปแบบ รูปแบบแรกคือการสะกดนิ้วมือ (finger spelling) เป็นท่ามือที่แทนพยัญชนะ สระ และวรรณยุกต์ รูปแบบที่สองคือท่าภาษามือที่แสดงถึงคำศัพท์ วลี หรือประโยค (Buangamdee, 2017) ทั้งนี้ ในการวิจัยนี้จะเป็นการศึกษาและพัฒนาโมเดลในส่วนของการรู้จำประโยคภาษามือไทยพื้นฐานที่ใช้ในห้องเรียน เพื่อให้ได้โมเดลที่สามารถนำไปประยุกต์ใช้เป็นเครื่องมือให้ครูสามารถเรียนรู้เกี่ยวกับภาษามือและสื่อสารกับนักเรียนหูหนวกในชั้นเรียนของตนเองได้

การรู้จำภาษามือต้องอาศัยข้อมูลที่สะท้อนถึงท่าทางของผู้ทำภาษามือ จากการศึกษพบว่าวิธีการที่ใช้ในการรับข้อมูลดังกล่าวอยู่ 2 วิธี (Sharma & Singh, 2023) วิธีแรกคือการใช้เซนเซอร์ (sensor-based approaches) ซึ่งเซนเซอร์นี้จะอยู่ในอุปกรณ์ที่สวมใส่ไปบนร่างกายของผู้ทำภาษามือ ซึ่งในแต่ละการวิจัยเกี่ยวกับการรู้จำภาษามือไทยแต่ละงานจะมีการใช้อุปกรณ์ต่างกัน เช่น Jaruwawat (2007); Saengsri et al. (2012) ใช้ถุงมืออิเล็กทรอนิกส์ซึ่งเป็นอุปกรณ์ที่จะส่งสัญญาณไฟฟ้าเมื่อเกิดการเคลื่อนไหว ส่วน Amatanon et al. (2014); Vachirapipop et al. (2017) ใช้เครื่องวัดสัญญาณไฟฟ้ากล้ามเนื้อ (electromyogram) ซึ่งจะรับสัญญาณโดยตรงจากกล้ามเนื้อเมื่อเกิดการเคลื่อนไหว ทั้งนี้การใช้อุปกรณ์สวมใส่ที่กล่าวมาข้างต้นมีข้อดีคือเซนเซอร์สามารถรับข้อมูลการเคลื่อนไหวได้อย่างแม่นยำ และไม่จำเป็นต้องอาศัยการเตรียมข้อมูลที่ซับซ้อนก่อนนำไปพัฒนาโมเดล แต่อย่างไรก็ตามอุปกรณ์ดังกล่าวมีต้นทุนที่สูง และการสวมใส่อุปกรณ์จะทำให้ความยืดหยุ่นในการทำภาษามือลดลงด้วย ซึ่งจากข้อจำกัดนี้ทำให้ในเวลาต่อมามากไม่ค่อยพบการศึกษาเกี่ยวกับการรู้จำภาษามือโดยใช้อุปกรณ์ดังกล่าว จากการศึกษางานวิจัยเกี่ยวกับการรู้จำภาษามือไทยในเวลาต่อมา (Chaikaew, 2022; Pariwat & Seresangtakul, 2021; Plianwong & Kangkachit, 2021; Sriboonruang et al., 2022; Vijitkunsawat et al., 2023) พบว่าวิธีการรับข้อมูลจากการทำภาษามืออีกวิธีที่เป็นที่นิยม

มากกว่า คือ การใช้ภาพ (image-based approaches) ซึ่งมีทั้งในรูปแบบของภาพนิ่งและวิดีโอขึ้นอยู่กับภาษา มือที่ถูกนำมาใช้ในการศึกษา โดยวิธีนี้มีข้อได้เปรียบคือมีต้นทุนที่ต่ำกว่า และผู้ทำภาษามือสามารถทำท่าทางต่าง ๆ ได้อย่างอิสระ แต่วิธีการนี้ต้องอาศัยเทคนิคในการเตรียมข้อมูลที่มีความซับซ้อน ทั้งนี้ วิธีการที่ผู้วิจัยเลือกมาใช้ในการวิจัยนี้คือการใช้ภาพซึ่งจะใช้ข้อมูลในรูปแบบวิดีโอ เนื่องจากครูและนิสิตนักศึกษาครูที่นำโมเดลไปใช้สามารถใช้กล้องในการรับข้อมูลเพื่อรู้จำภาษามือได้ ซึ่งมีความสะดวกในการใช้งานและประหยัดต้นทุนกว่าการใช้อุปกรณ์สวมใส่

การพัฒนาโมเดลรู้จำภาษามือเพื่อใช้ในการแปลภาษามือนั้นเป็นการประยุกต์ใช้โมเดลการเรียนรู้ของเครื่อง (machine learning) ในการจำแนกท่าทางของผู้ทำภาษามือ จากการศึกษางานวิจัยที่เกี่ยวข้องกับการรู้จำภาษามือไทยโดยการใช้ภาพ พบว่าโมเดลการเรียนรู้ของเครื่องที่ถูกนำมาใช้มีความหลากหลาย โดยแบ่งเป็น 2 กลุ่ม กลุ่มแรกคือโมเดลการเรียนรู้ของเครื่องแบบดั้งเดิม เช่น hidden Markov model (Sakulsujirapa, 2011) decision tree (Veerasakulthong, 2007) KNN (Klomsae et al., 2017; Pariwat & Seresangtakul, 2021) อย่างไรก็ตามในกรณีข้อมูลที่เป็นรูปภาพหรือวิดีโอ การพัฒนาโมเดลจะมีเตรียมข้อมูลให้มีความเหมาะสมโดยการสร้างคุณลักษณะ (feature engineering) ซึ่งในการใช้โมเดลการเรียนรู้ของเครื่องแบบดั้งเดิมนั้นจะมีขั้นตอนการสร้างคุณลักษณะที่ยาก และมีข้อจำกัดในการเรียนรู้ที่มีความซับซ้อน จากปัญหาดังกล่าวจึงมีการนำโมเดลกลุ่มที่สองคือโมเดลการเรียนรู้เชิงลึก (deep learning) มาใช้แทนโมเดลการเรียนรู้ของเครื่องแบบดั้งเดิม เช่น artificial neural networks หรือ ANN (Amatanon et al., 2014; Chansri & Srinonchat, 2016) เป็นโมเดลที่ได้แรงบันดาลใจมาจากการทำงานของเซลล์ประสาทในสมองมนุษย์ในการแก้ปัญหาที่มีความซับซ้อน convolutional neural networks หรือ CNN (Gedkhaw, 2022; Pariwat & Seresangtakul, 2021; Sripairajthikoon & Harnsomburana, 2019) เป็นโมเดลที่สามารถดึงคุณลักษณะจากข้อมูลรูปภาพได้ recurrent neural networks หรือ RNN (Chaikaew, 2022; Vijitkunsawat et al., 2023; Sriboonruang et al., 2022) เป็นโมเดลที่สามารถเรียนรู้กับข้อมูลที่เป็นลำดับได้ และนอกจากนี้ยังมีโมเดลการเรียนรู้เชิงลึกอื่นอีกหลายโมเดล โดยการศึกษาวิจัยนี้ผู้วิจัยเลือกใช้โมเดล CNN ร่วมกับ RNN เนื่องจากโมเดล CNN มีจุดเด่นในการดึงคุณลักษณะ (feature extraction) จากข้อมูลรูปภาพ สามารถนำมาใช้ดึงคุณลักษณะจากภาพในวิดีโอภาษามือได้ โดยโมเดลในกลุ่ม CNN นั้นมีอยู่หลายโมเดล การวิจัยนี้ผู้วิจัยเลือกใช้โมเดล CNN จำนวน 2 โมเดล ได้แก่ DenseNet121 และ EfficientNetB0 เนื่องจากเป็นโมเดลที่มีขนาดเล็ก เป็นโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) บนชุดข้อมูลขนาดใหญ่ ทำให้มีประสิทธิภาพในการทำนาย สำหรับโมเดล RNN มีจุดเด่นในการเรียนรู้ข้อมูลที่เป็นลำดับ (sequence learning) ซึ่งสามารถนำมาใช้กับข้อมูลวิดีโอภาษามือซึ่งเกิดจากรูปภาพที่นำมาต่อกันตามลำดับเวลา โดยโมเดลในกลุ่ม RNN นั้นมีอยู่หลายโมเดล การวิจัยนี้ผู้วิจัยเลือกใช้โมเดล RNN จำนวน 2 โมเดล ได้แก่ long-short term memory และ gated recurrent unit (GRU) เนื่องจากเป็นโมเดลที่นิยมใช้ในการเรียนรู้ข้อมูลที่เป็นลำดับเวลา ดังนั้นการใช้โมเดล CNN ร่วมกับ RNN จึงมีความเหมาะสมในการนำไปพัฒนาโมเดลการรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกจากข้อมูลที่เป็นวิดีโอ

จากที่กล่าวมาการพัฒนาโมเดลรู้จำภาษามือโดยใช้การเรียนรู้เชิงลึกจึงเป็นทางเลือกที่เหมาะสมสำหรับการแก้ปัญหาการสื่อสารระหว่างครูกับนักเรียนหูหนวก โดยโมเดลที่พัฒนาขึ้นนี้สามารถนำไปใช้เป็น

เครื่องมือในการแปลภาษามือไทยพื้นฐานที่ใช้ในห้องเรียน ช่วยให้ครูหรือนิสิตนักศึกษาครูที่ไม่มีความรู้พื้นฐานภาษามือ สามารถเข้าใจและสื่อสารกับนักเรียนหูหนวกได้อย่างมีประสิทธิภาพและมีความยืดหยุ่นมากขึ้น นอกจากนี้โมเดลที่พัฒนาขึ้นยังสามารถนำไปใช้เป็นโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) สำหรับงานวิจัยอื่น ๆ ซึ่งช่วยต่อยอดการพัฒนาโมเดลรู้จำภาษามือ และช่วยให้การฝึกสอนโมเดลในงานใหม่เป็นไปได้อย่างรวดเร็วและมีประสิทธิภาพมากขึ้นโดยไม่จำเป็นต้องเริ่มฝึกสอนโมเดลใหม่ตั้งแต่ต้น

ความมุ่งหมายของการวิจัย

1. เพื่อวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน
2. เพื่อพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

วิธีดำเนินการวิจัย

การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกในการวิจัยนี้ มีรายละเอียดวิธีดำเนินการวิจัย ได้แก่ ตัวอย่างวิจัย เครื่องมือที่ใช้ในการวิจัย การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล โดยมีการวิจัย 2 ระยะ ได้แก่ การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน และการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

1. ตัวอย่างวิจัย

ตัวอย่างวิจัยที่ใช้การวิจัยนี้แบ่งออกเป็น 2 ระยะ ดังนี้

ระยะที่ 1 การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน ตัวอย่างวิจัย คือ ครูการศึกษาพิเศษ จำนวน 13 คน โดยการเลือกแบบเจาะจง มีเกณฑ์ในการคัดเลือก ได้แก่ 1) เป็นครูหรืออาจารย์ที่ทำการสอนประจำในโรงเรียนสอนคนหูหนวก 2) มีประสบการณ์อย่างน้อยหนึ่งปีในการสอนนักเรียนหูหนวกในระดับการศึกษาขั้นพื้นฐาน และ 3) มีความรู้ความเข้าใจในการใช้ภาษามือไทยเพื่อการสื่อสารกับนักเรียนหูหนวก

ระยะที่ 2 การพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก ตัวอย่างวิจัย คือ อาสาสมัครที่มีความสามารถในการใช้ภาษามือ จำนวน 4 คน โดยการเลือกแบบเจาะจง มีเกณฑ์ในการคัดเลือกได้แก่ 1) เป็นผู้ที่มีการใช้ภาษามือสื่อสารในโรงเรียนสอนคนหูหนวก 2) มีทักษะในการใช้ภาษามือไทยในการสื่อสารอยู่ในระดับดี และ 3) สามารถทำภาษามือตามประโยคที่ใช้ในการวิจัยได้ ทั้งนี้จำนวนตัวอย่างที่ใช้ในการวิจัยมีขนาดเพียงพอต่อการพัฒนาโมเดลรู้จำภาษามือโดยใช้การเรียนรู้เชิงลึก เนื่องจากผู้วิจัยจะทำการเพิ่มจำนวนข้อมูล (data augmentation) และใช้โมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) มาต่อยอดในการพัฒนาโมเดลรู้จำภาษามือ ทำให้สามารถใช้ตัวอย่างที่มีขนาดเล็กได้

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยนี้ คือ แบบสอบถามความถี่การใช้งานประโยคภาษามือ ซึ่งจะนำมาใช้ในส่วนของการวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน โดยแบบสอบถามนี้เป็นแบบสอบถามแบบมาตราประมาณค่า 4 ระดับ โดยที่ระดับ 4 หมายถึง ใช้ทุกคาบเรียน ระดับ 3 หมายถึง ใช้ในคาบเรียนส่วนใหญ่ ระดับ 2 หมายถึง ใช้ในบางคาบเรียน และระดับ 1 หมายถึง ไม่เคยใช้เลย

แบบสอบถามมีจำนวน 58 ข้อ ประกอบด้วยประโยคภาษามือที่แบ่งออกเป็น 5 หมวด ได้แก่ การนำเข้าสู่บทเรียน การตรวจสอบความเข้าใจ การออกคำสั่ง การเสริมแรง และการโต้ตอบของนักเรียน ทั้งนี้ ประโยคภาษามือที่ใช้ในแบบสอบถามนี้ได้มีการตรวจสอบความเหมาะสมและปรับแก้ตามคำแนะนำของผู้ทรงคุณวุฒิ ด้านการศึกษาพิเศษก่อนนำไปใช้เก็บรวบรวมข้อมูล

3. การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในการวิจัยนี้ผู้วิจัยแบ่งออกเป็น 2 ระยะ ได้แก่ การเก็บรวบรวมข้อมูลภาษามือที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน เพื่อนำมาวิเคราะห์หาประโยคภาษามือที่ใช้บ่อยในชั้นเรียนสำหรับนำไปใช้พัฒนาโมเดลรู้จำภาษามือ และการเก็บรวบรวมข้อมูลสำหรับการฝึกสอนและประเมินประสิทธิภาพของโมเดลรู้จำภาษามือ

ระยะที่ 1 การเก็บรวบรวมข้อมูลภาษามือที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน เป็นการเก็บรวบรวมข้อมูลเพื่อวิเคราะห์ประโยคภาษามือที่ครูมักใช้ในสื่อสารชั้นเรียนเพื่อนำประโยคดังกล่าวมาใช้ในการพัฒนาโมเดลรู้จำภาษามือ โดยผู้วิจัยได้ศึกษาประโยคที่ครูใช้ในชั้นเรียนจากนั้นนำไปปรึกษาผู้ทรงคุณวุฒิเพื่อตรวจสอบความเหมาะสมของประโยคที่ใช้ในชั้นเรียนของโรงเรียนสอนคนหูหนวก และปรับแก้ตามคำแนะนำของผู้ทรงคุณวุฒิ จากนั้นจึงจัดทำแบบสอบถามความถี่การใช้งานประโยคภาษามือและส่งแบบสอบถามไปยังครูการศึกษาพิเศษที่ทำการสอนในโรงเรียนสอนคนหูหนวก

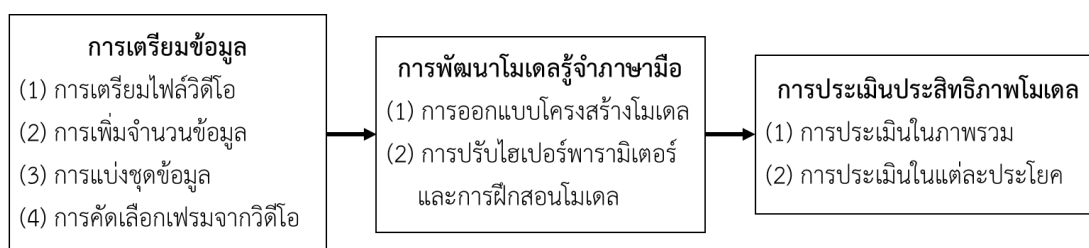
ระยะที่ 2 การเก็บรวบรวมข้อมูลสำหรับการฝึกสอนและประเมินประสิทธิภาพของโมเดลรู้จำภาษามือ เป็นการเก็บรวบรวมข้อมูลในรูปแบบวิดีโอการทำภาษามือไทย โดยนำประโยคภาษามือที่มีการใช้สื่อสารในชั้นเรียนจำนวน 20 ประโยค มาถ่ายทำกับอาสาสมัครจำนวน 4 คน และมีการถ่ายซ้ำจำนวน 4 ครั้ง ทำให้ได้วิดีโอสำหรับการพัฒนาโมเดลการรู้จำภาษามือไทยรวมทั้งสิ้น 320 วิดีโอ โดยหนึ่งวิดีโอจะแทนหนึ่งประโยค และแต่ละประโยคจะมีจำนวนวิดีโอเท่ากัน ทั้งนี้การเก็บรวบรวมข้อมูลในการวิจัยนี้ได้รับความอนุเคราะห์จากโรงเรียนเศรษฐเสถียร ในพระราชูปถัมภ์ โรงเรียนเฉพาะทางสำหรับนักเรียนหูหนวก ซึ่งมีการใช้ภาษามือในการจัดการเรียนการสอนเป็นหลัก และระหว่างการถ่ายวิดีโอมีผู้ทรงคุณวุฒิด้านภาษามือ 2 ท่านคอยตรวจสอบความถูกต้องในการทำภาษามือของอาสาสมัคร

4. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการวิจัยนี้แบ่งออกเป็น 2 ระยะ ได้แก่ การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียนเพื่อให้ได้ภาษามือที่จะใช้ในการพัฒนาโมเดล และการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกเพื่อหาโมเดลที่มีประสิทธิภาพสูงที่สุดและวิเคราะห์ความสามารถในการรู้จำภาษามือในภาพรวมและในแต่ละประโยค

ระยะที่ 1 การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน เป็นการวิเคราะห์ข้อมูลจากแบบสอบถามการใช้งานประโยคภาษามือ เพื่อคัดเลือกประโยคภาษามือที่มีการใช้สื่อสารบ่อยในชั้นเรียนจำนวน 20 ประโยค โดยใช้ค่าเฉลี่ยของคะแนนความถี่ในการใช้งานประโยคภาษามือเพื่อให้ได้ประโยคภาษามือที่มีคะแนนความถี่ในการใช้งานมากที่สุด 20 อันดับแรก สำหรับนำมาใช้ในการพัฒนาโมเดลรู้จำภาษามือ และใช้คำร้อยละเพื่อเปรียบเทียบจำนวนประโยคภาษามือที่ใช้ในแต่ละหมวด

ระยะที่ 2 การพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษาไทยโดยใช้การเรียนรู้เชิงลึก ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลโดยใช้ Google Colab ในการประมวลผลบน TPU ด้วยภาษา Python ผ่าน Jupyter notebook ในการเขียนคำสั่ง และสร้างโมเดลการเรียนรู้เชิงลึกผ่าน Tensorflow และ Keras API โดยการพัฒนาโมเดลประกอบด้วย 3 ขั้นตอน ดังภาพ 1 ได้แก่ การเตรียมข้อมูลเพื่อให้ได้ข้อมูลในรูปแบบที่สามารถนำไปใช้ฝึกสอนโมเดลการเรียนรู้เชิงลึกได้ การพัฒนาโมเดลเพื่อออกแบบโครงสร้างโมเดลและฝึกสอนโมเดล และการประเมินประสิทธิภาพของโมเดลเพื่อวิเคราะห์ความสามารถในการรู้จำภาษาไทยของโมเดล ในภาพรวมและในแต่ละประโยค



ภาพ 1 ขั้นตอนการวิเคราะห์ข้อมูลในส่วนของการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษาไทย

1. การเตรียมข้อมูล สำหรับการพัฒนาโมเดลรู้จำภาษาไทยในการวิจัยนี้แบ่งเป็น 4 ขั้นตอน ประกอบด้วย การเตรียมไฟล์วิดีโอเพื่อให้ได้วิดีโอเฉพาะส่วนที่มีสารสนเทศสำหรับการพัฒนาโมเดล การเพิ่มจำนวนข้อมูลเพื่อให้โมเดลสามารถเรียนรู้ได้มีประสิทธิภาพมากขึ้น การแบ่งชุดข้อมูลสำหรับฝึกสอน ตรวจสอบและทดสอบประสิทธิภาพของโมเดล และการคัดเลือกเฟรมจากวิดีโอเพื่อให้ได้เฟรมรูปภาพในจำนวนที่เหมาะสมกับการพัฒนาโมเดล

(1) การเตรียมไฟล์วิดีโอ เนื่องจากแต่ละวิดีโอที่ผู้วิจัยบันทึกไว้มีส่วนที่ไม่ได้มีการสารสนเทศเกี่ยวกับภาษามือ ได้แก่ ช่วงที่อาสาสมัครอยู่นิ่ง และภาพพื้นหลัง ดังนั้นผู้วิจัยจึงได้ใช้โปรแกรม CapCut ตัดวิดีโอให้เหลือเฉพาะช่วงที่มีการเคลื่อนไหวของมือ และลบพื้นหลังของวิดีโอออกเป็นสีดำให้เหลือเฉพาะผู้ทำภาษามือ เพื่อให้ได้เฉพาะวิดีโอในส่วนที่มีสารสนเทศที่จำเป็นในการนำไปสร้างโมเดลรู้จำภาษามือ ทั้งนี้ ชื่อไฟล์วิดีโอจะถูกตั้งให้ตามความหมายของประโยคภาษามือในแต่ละวิดีโอ

(2) การเพิ่มจำนวนข้อมูล (data augmentation) เนื่องจากวิดีโอที่ผู้วิจัยรวบรวมมาได้ในแต่ละประโยคมีจำนวนที่น้อย แต่ในการพัฒนาโมเดลการเรียนรู้เชิงลึกหากมีข้อมูลให้โมเดลได้เรียนรู้จำนวนมาก จะช่วยให้โมเดลได้เรียนรู้จากข้อมูลที่มีความหลากหลายและทำให้โมเดลมีประสิทธิภาพมากขึ้น ดังนั้นในงานวิจัยนี้ผู้วิจัยจึงได้ทำการเพิ่มจำนวนข้อมูลของวิดีโอภาษามือ โดยนำวิดีโอที่มีอยู่เดิมมาทำการดัดแปลงให้มีความหลากหลายมากขึ้นผ่านไลบรารี OpenCV ด้วยวิธีการเพิ่มจำนวนข้อมูล 7 วิธี ได้แก่ การกลับด้าน (flip) การหมุน (rotate) การเพิ่มจลรบกวน (noise) การเบลอ (blur) การปรับเป็นสีขาวดำ (greyscale) การเพิ่มความเร็ว (speed up) และการลดความเร็ว (slow down) ทำให้ได้วิดีโอใหม่อีก 2,240 วิดีโอ ซึ่งรวมกับวิดีโอเดิมที่มีอยู่ 320 วิดีโอ ทำให้ได้วิดีโอสำหรับนำไปพัฒนาโมเดลจำนวนทั้งสิ้น 2,560 วิดีโอ

(3) การแบ่งชุดข้อมูล การพัฒนาโมเดลการเรียนรู้เชิงลึกในการวิจัยครั้งนี้มีได้นำข้อมูลที่เก็บรวบรวมได้และนำไปเพิ่มจำนวนวิดีโอจนได้เป็น 960 วิดีโอ มาแบ่งออกเป็น 3 ชุดข้อมูล เพื่อใช้สำหรับฝึกสอน

ตรวจสอบ และทดสอบประสิทธิภาพของโมเดล ดังนี้ 1) ชุดข้อมูลฝึกสอน (training dataset) ใช้ข้อมูลร้อยละ 75 ของข้อมูลทั้งหมด เป็นชุดข้อมูลที่นำมาใช้ฝึกสอนโมเดลการเรียนรู้เชิงลึก โดยให้โมเดลแต่ละโมเดลเรียนรู้จากข้อมูลชุดนี้ 2) ชุดข้อมูลตรวจสอบ (validation dataset) ใช้ข้อมูลร้อยละ 12.5 ของข้อมูลทั้งหมด เป็นชุดข้อมูลที่นำมาใช้ตรวจสอบประสิทธิภาพโมเดลการเรียนรู้เชิงลึกขณะที่กำลังฝึกสอนโมเดล และใช้เปรียบเทียบประสิทธิภาพของแต่ละโมเดลที่ใช้ในการวิจัยครั้งนี้ และ 3) ชุดข้อมูลทดสอบ (test dataset) ใช้ข้อมูลร้อยละ 12.5 ของข้อมูลทั้งหมด เป็นชุดข้อมูลที่นำมาใช้ทดสอบประสิทธิภาพของโมเดลการเรียนรู้เชิงลึกที่มีความแม่นยำมากที่สุด และใช้ประเมินประสิทธิภาพการจำแนกของแต่ละประโยคที่ใช้ในโมเดลการเรียนรู้เชิงลึก ทั้งนี้ ชุดข้อมูลฝึกสอนจะใช้ข้อมูลวิดีโอจากอาสาสมัคร 3 คน เพื่อให้มีข้อมูลจำนวนมากพอสำหรับการฝึกสอนโมเดล ส่วนชุดข้อมูลตรวจสอบและชุดข้อมูลทดสอบจะใช้ข้อมูลวิดีโอจากอาสาสมัคร 1 คน แล้วแบ่งข้อมูลวิดีโอออกเป็นสองส่วนสำหรับใช้ในสองชุดข้อมูลดังกล่าว

(4) การคัดเลือกเฟรมจากวิดีโอ วิดีโอภาษาไทยที่ได้ทำการแบ่งชุดข้อมูลเรียบร้อยแล้วจะถูกนำมาคัดเลือกเฟรมและปรับขนาดให้มีความเหมาะสมเพื่อนำไปใช้สำหรับพัฒนาโมเดลการเรียนรู้เชิงลึก เนื่องจากหากใช้ทุกเฟรมจากวิดีโอหรือภาพในแต่ละเฟรมเป็นภาพขนาดใหญ่ จะทำให้โมเดลมีขนาดใหญ่และใช้เวลานานในการพัฒนาโมเดล ดังนั้นผู้วิจัยจึงได้กำหนดจำนวนเฟรมที่คัดเลือกจากแต่ละวิดีโอไว้ที่จำนวน 20 เฟรม และกำหนดขนาดภาพในแต่ละเฟรมให้อยู่ที่ 224×224 พิกเซล เนื่องจากเป็นขนาดที่ถูกกำหนดให้ใช้ในการพัฒนาโมเดลการเรียนรู้เชิงลึกด้วยโมเดลที่ผู้วิจัยเลือกใช้ในงานวิจัยนี้

2. การพัฒนาโมเดล แบ่งเป็น 2 ส่วน คือ การออกแบบโครงสร้างของโมเดลสำหรับการรู้จำภาษาไทย และการปรับไฮเปอร์พารามิเตอร์และการฝึกสอนโมเดลเพื่อหาโมเดลที่มีประสิทธิภาพดีที่สุด

(1) การออกแบบโครงสร้างของโมเดล

การวิจัยนี้เป็นการพัฒนาโมเดลรู้จำภาษาไทยจากวิดีโอภาษาไทยแต่ละประโยค โดยหนึ่งวิดีโอแทนหนึ่งประโยคภาษาไทย โมเดลที่ใช้ในการรู้จำภาษาไทยจึงเป็นโมเดลจำแนกแบบหลายกลุ่ม (multi-class classification model) และเนื่องจากวิดีโอเป็นข้อมูลที่เกิดจากรูปภาพที่นำต่อกันตามลำดับเวลา ซึ่งโมเดลการเรียนรู้เชิงลึกที่มีประสิทธิภาพในการทำงานกับข้อมูลรูปภาพคือโมเดลในกลุ่ม convolutional neural network (CNN) และโมเดลการเรียนรู้เชิงลึกโมเดลการเรียนรู้เชิงลึกที่มีประสิทธิภาพในการทำงานกับข้อมูลที่เป็นลำดับคือโมเดลในกลุ่ม recurrent neural network (RNN) ดังนั้นในการพัฒนาโมเดลการรู้จำภาษาไทยโดยใช้การเรียนรู้เชิงลึก ผู้วิจัยได้พัฒนาโมเดลการรู้จำภาษาไทยโดยประยุกต์ใช้ CNN ร่วมกับ RNN

โมเดล CNN และ RNN ที่ใช้ในการวิจัยนี้ โมเดลกลุ่ม CNN ผู้วิจัยเลือกใช้โมเดล DenseNet121 และ EfficientNetB0 เนื่องจากเป็นโมเดลขนาดเล็ก และเป็นโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) กับชุดข้อมูลรูปภาพขนาดใหญ่ ImageNet สามารถนำมาต่อยอดในการดึงคุณลักษณะจากเฟรมภาพในวิดีโอภาษาไทยได้ ในส่วนของโมเดลกลุ่ม RNN ผู้วิจัยเลือกใช้โมเดล long-short term memory (LSTM) และโมเดล gated recurrent unit (GRU) เนื่องจากเป็นโมเดลที่นิยมใช้ในการจำแนกข้อมูลที่มีลักษณะเป็นลำดับ ดังนั้นจากโมเดลที่กล่าวมาข้างต้น การพัฒนาโมเดลการรู้จำภาษาไทยในการวิจัยนี้มีโมเดลที่ถูกใช้ทั้งหมด 4 โมเดล ได้แก่ 1) DenseNet121 ร่วมกับ LSTM 2) DenseNet121 ร่วมกับ GRU 3) EfficientNetB0 ร่วมกับ LSTM และ 4) EfficientNetB0 ร่วมกับ GRU

(2) การปรับไฮเปอร์พารามิเตอร์ของโมเดลและการฝึกสอนโมเดล

ในส่วนของการปรับไฮเปอร์พารามิเตอร์ (hyperparameter) ซึ่งเป็นพารามิเตอร์ที่กำหนดไว้ก่อนการฝึกสอน ผู้วิจัยเลือกใช้วิธีการปรับการพารามิเตอร์ (optimizer) แบบ adam และใช้ฟังก์ชันค่าสูญเสีย (loss function) สำหรับวัดความแตกต่างของค่าจริงกับค่าที่โมเดลทำนายเป็น sparse categorical cross entropy เนื่องจากผู้วิจัยได้กำหนดผลลัพธ์การทำนายภาษามือ 20 ประโยคหรือคลาสสำหรับการทำนายในรูปแบบตัวเลข 0 ถึง 19 ในการฝึกสอนโมเดล จากนั้นกำหนดดัชนีที่ใช้ประเมินประสิทธิภาพของโมเดล (metrics) ระหว่างการฝึกสอนโมเดลแต่ละรอบเป็นค่าความถูกต้อง (accuracy) เนื่องจากโมเดลในการวิจัยนี้เป็นโมเดลจำแนกแบบหลายกลุ่ม ต่อมากำหนดจำนวนตัวอย่างข้อมูลในการปรับค่าพารามิเตอร์หนึ่งครั้ง (batch size) เท่ากับ 32 จำนวนรอบในการพัฒนาโมเดล (epoch) เท่ากับ 20 รอบ และกำหนดให้มีการหยุดฝึกสอนโมเดลก่อนครบจำนวนรอบ (early stopping) เมื่อค่าความสูญเสียบนข้อมูลชุดตรวจสอบ (validation loss) ไม่ลดลงติดต่อกัน 10 รอบ เพื่อป้องกันการเรียนรู้บนชุดข้อมูลฝึกสอนมากเกินไป (overfitting)

ผู้วิจัยดำเนินการฝึกสอนโมเดลโดยนำเข้าข้อมูลเฟรมภาพที่ได้จากการเตรียมข้อมูล ซึ่งเป็นข้อมูลที่ครอบคลุมในทุกท่ามือและการเคลื่อนไหวของภาษามือในแต่ละประโยค ในระหว่างการฝึกสอนโมเดลผู้วิจัยจะดำเนินการติดตามค่าความสูญเสีย (loss) ซึ่งเป็นค่าความแตกต่างระหว่างค่าจริงกับค่าที่โมเดลทำนาย และค่าความถูกต้อง (accuracy) ซึ่งเป็นสัดส่วนข้อมูลตัวอย่างที่โมเดลทำนายได้ถูกต้อง โดยจะติดตามค่าดังกล่าวในชุดข้อมูลฝึกสอนและชุดข้อมูลตรวจสอบ เพื่อตรวจสอบความถูกต้องและความผิดพลาดของแต่ละโมเดล จากนั้นเมื่อฝึกสอนเสร็จสิ้นครบทุกโมเดลแล้ว ผู้วิจัยจะนำแต่ละโมเดลมาเปรียบเทียบประสิทธิภาพโดยเปรียบเทียบจากค่าความสูญเสียบนข้อมูลชุดตรวจสอบ (validation loss) ค่าความถูกต้องบนชุดข้อมูลตรวจสอบ (validation accuracy) ขนาดโมเดล และเวลาเฉลี่ยที่ใช้ในการฝึกสอนโมเดลต่อรอบ ทั้งนี้ โมเดลที่มีประสิทธิภาพดีที่สุดจะถูกนำไปใช้ประเมินประสิทธิภาพโมเดลกับข้อมูลชุดทดสอบ

(3) การประเมินประสิทธิภาพของโมเดลรู้จำภาษามือไทย

การประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทย ผู้วิจัยนำโมเดลที่มีประสิทธิภาพสูงที่สุดจากขั้นตอนการพัฒนาโมเดลมาประเมินประสิทธิภาพกับชุดข้อมูลทดสอบ ทั้งนี้ เนื่องจากโมเดลการแปลภาษามือไทยโดยใช้การเรียนรู้เชิงลึกเป็นโมเดลจำแนกแบบหลายกลุ่ม (multi-class classification model) ดังนั้นการประเมินประสิทธิภาพของโมเดลจึงใช้คอนฟิวชันเมทริกซ์ (confusion matrix) แบบหลายกลุ่ม (multi-class) ซึ่งเป็นเมทริกซ์ที่แสดงจำนวนตัวอย่างข้อมูลในประโยคที่โมเดลทำนายเทียบกับประโยคจริงว่าโมเดลทำนายได้ถูกต้องมากน้อยเพียงใด มีการทำนายผิดพลาดอย่างไรบ้าง และนำจำนวนตัวอย่างข้อมูลที่ได้จากคอนฟิวชันเมทริกซ์มาประเมินประสิทธิภาพของโมเดลรู้จำภาษามือในภาพรวมและประสิทธิภาพการรู้จำในแต่ละประโยคภาษามือโดยใช้ค่าต่อไปนี้

1) ค่าความถูกต้อง (accuracy) เป็นค่าที่แสดงถึงความถูกต้องในการทำนายของโมเดลในภาพรวมทุกประโยค คำนวณจากสมการ
$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

2) ค่าความแม่นยำ (precision) เป็นค่าที่แสดงถึงความแม่นยำในการทำนายในแต่ละประโยค คำนวณจากสมการ
$$\text{precision} = \frac{TP}{TP + FP}$$

3) ค่าการเรียกคืน (recall) เป็นค่าที่แสดงถึงการไม่พลาดการทำนายในแต่ละประโยค

$$\text{คำนวณจากสมการ recall} = \frac{TP}{TP + FN}$$

4) คะแนน F1 (F1-score) เป็นค่าเฉลี่ยแบบฮาร์โมนิกของค่า precision และ recall คำนวณ

$$\text{จากสมการ F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

โดยถ้ากำหนดให้ positive เป็นคลาสหรือประโยคที่สนใจ และคลาสอื่นหรือประโยคอื่นเป็น negative จะได้ว่า true positive (TP) คือ จำนวนตัวอย่างที่โมเดลทำนายว่าเป็นคลาส positive และอยู่ในคลาส positive จริง true negative (TN) คือ จำนวนตัวอย่างที่โมเดลทำนายว่าเป็นคลาส negative และอยู่ในคลาส negative จริง false positive (FP) คือ จำนวนตัวอย่างที่โมเดลทำนายว่าเป็นคลาส positive แต่ความจริงไม่ได้อยู่ในคลาส positive จริง และ false negative (FN) คือ จำนวนตัวอย่างที่โมเดลทำนายว่าเป็นคลาส negative แต่ความจริงไม่ได้อยู่ในคลาส negative

ผลการวิจัย

ผลการวิจัยแบ่งออกได้เป็น 2 ตอน ได้แก่ ผลการวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน และผลการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

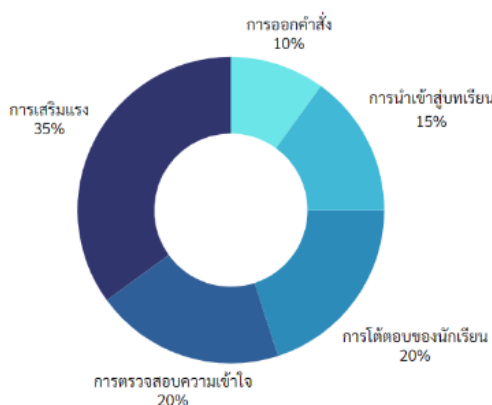
ตอนที่ 1 ผลการวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน

จากการวิเคราะห์ค่าเฉลี่ยของคะแนนความถี่ประโยคภาษามือที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษจากแบบสอบถามความถี่การใช้งานประโยคภาษามือในชั้นเรียน โดยคะแนนความถี่จะอยู่ในช่วง 1 ถึง 4 คะแนน คะแนนที่มีค่ามากจะหมายถึงการใช้ประโยคภาษามือในชั้นเรียนบ่อยครั้ง ผลการวิเคราะห์พบว่าประโยคภาษามือที่มีการใช้งานบ่อยที่สุด 20 อันดับแรก ได้แก่ สวัสดี พร้อมยัง ถูกหรือผิด เก่งมาก เรียบร้อย ตั้งใจเรียน เข้าใจ ถูกไหม ดูภาพ ทำได้ไหม ใครตอบได้ งานสวย ภูมิใจ รู้ไหม ทำใบงาน เก็บอุปกรณ์ เยี่ยม น่ารัก ผิดไหม และทำอย่างไร โดยแต่ละประโยคมีค่าเฉลี่ยคะแนนความถี่จำแนกตามหมวด ดังตาราง 1

ตาราง 1 ค่าเฉลี่ยคะแนนความถี่ของประโยคภาษามือ 20 ประโยคที่ใช้งานบ่อยที่สุดในแต่ละหมวด

หมวด	ประโยค	ค่าเฉลี่ยคะแนนความถี่	หมวด	ประโยค	ค่าเฉลี่ยคะแนนความถี่
1. การนำเข้าสู่บทเรียน	สวัสดี	4.00	4. การเสริมแรง	เก่งมาก	3.92
	พร้อมยัง	3.92		เรียบร้อย	3.92
	ดูภาพ	3.85		ตั้งใจเรียน	3.92
2. การตรวจสอบความเข้าใจ	ถูกหรือผิด	3.92		งานสวย	3.85
	ทำได้ไหม	3.85		ภูมิใจ	3.85
	ใครตอบได้	3.85		เยี่ยม	3.77
	รู้ไหม	3.77		น่ารัก	3.77
3. การออกคำสั่ง	ทำใบงาน	3.77	5. การโต้ตอบของนักเรียน	เข้าใจ	3.92
	เก็บอุปกรณ์	3.77		ถูกไหม	3.92
				ผิดไหม	3.77
				ทำอย่างไร	3.69

จากตาราง 1 พบว่าประโยคภาษามือทั้ง 20 ประโยคมีการใช้งานบ่อยในชั้นเรียนเนื่องจากมีค่าคะแนนความถี่ใกล้เคียงกับ 4 ซึ่งเป็นค่าสูงสุด และเมื่อเปรียบเทียบจำนวนประโยคภาษามือในที่ถูกคัดเลือกแต่ละหมวดมาพบว่าหมวดการเสริมแรงมีจำนวนประโยคมากที่สุด (ร้อยละ 35) รองลงมาคือหมวดการตรวจสอบความเข้าใจ (ร้อยละ 20) หมวดการโต้ตอบของนักเรียน (ร้อยละ 20) หมวดการนำเข้าสู่บทเรียน (ร้อยละ 15) และหมวดการออกคำสั่งมีจำนวนประโยคน้อยที่สุด (ร้อยละ 10) ดังภาพ 2 โดยประโยคภาษามือที่ได้มาในขั้นตอนนี้จะถูกนำไปใช้ในการพัฒนาโมเดลการเรียนรู้จำภาษามือโดยใช้การเรียนรู้เชิงลึกในขั้นตอนต่อไป



ภาพ 2 แผนภูมิแสดงจำนวนประโยคภาษามือที่ใช้บ่อยจำแนกตามหมวด

ตอนที่ 2 ผลการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

ผลการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก แบ่งออกเป็น 3 ส่วน ได้แก่ ข้อมูลที่ใช้ในการพัฒนาโมเดล ผลการพัฒนาโมเดลรู้จำภาษามือไทย และผลการประเมินประสิทธิภาพของโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

1. ข้อมูลที่ใช้ในการพัฒนาโมเดล

ข้อมูลที่ใช้ในการพัฒนาโมเดลรู้จำภาษามือไทยเป็นวิดีโอท่าภาษามือที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษจำนวน 20 ประโยค โดยมีลักษณะท่ามือและการเคลื่อนไหว รวมไปถึงความยาวของวิดีโอที่ใช้ในการพัฒนาโมเดลแสดงได้ดังตาราง 2

ตาราง 2 ลักษณะท่ามือและความยาววิดีโอในแต่ละประโยคภาษามือที่ใช้ในการวิจัย

ประโยค	ท่ามือและการเคลื่อนไหว	ความยาววิดีโอ (วินาที)
พร้อมยัง	(1) คว่ำฝ่ามือลงทั้งสองข้างไว้บริเวณกลางลำตัวและมือท่างกัน เคลื่อนฝ่ามือซ้ายไปขวา (2) มือข้างหนึ่งชูนิ้วโป้งและนิ้วก้อย จากนั้นขยับมือกลับไปกลับมาซ้ำ ๆ	2.4
งานสวย	(1) กำมือสองข้างข้างหนึ่งอยู่ด้านบน อีกข้างอยู่ด้านล่าง เคลื่อนที่เข้ามาหากันบริเวณหน้าอก (2) แขนมือแล้วเคลื่อนที่วนรอบบริเวณใบหน้า (3) ชูนิ้วโป้งไว้ที่บริเวณด้านข้างศีรษะ	1.9
ทำได้ไหม	(1) กำมือสองข้างข้างหนึ่งอยู่ด้านบน อีกข้างอยู่ด้านล่าง เคลื่อนที่เข้ามาหากันบริเวณกลางลำตัว และทำซ้ำอีกครั้ง (2) ชูนิ้วโป้งบริเวณด้านหน้า (3) ใช้นิ้วชี้เคลื่อนที่เป็นรูปเครื่องหมายคำถาม	1.7
รู้ไหม	(1) ชี้นิ้วชี้และนิ้วกลางไปยังบริเวณข้างศีรษะ (2) ใช้นิ้วชี้เคลื่อนที่เป็นรูปเครื่องหมายคำถาม	1.3
ทำไปงาน	(1) กำมือสองข้างข้างหนึ่งอยู่ด้านบน อีกข้างอยู่ด้านล่าง เคลื่อนที่เข้ามาหากันบริเวณกลางลำตัว และทำซ้ำอีกครั้ง (2) ตั้งฝ่ามือสองข้างไว้ที่ข้างศีรษะ หันออกมาด้านหน้า ขยับมือซ้ายขวาเล็กน้อย	2.1

ตาราง 2 (ต่อ)

ประโยค	ท่ามือและการเคลื่อนไหว	ความยาววิดีโอ (วินาที)
	พร้อมเคลื่อนมือลงมาด้านล่าง (3) กำมือหนึ่งข้างไว้ด้านโดยไขว้กับมืออีกข้าง จากนั้นขยับมือ ด้านบนไปข้างหน้าสองจังหวะ	
เยียม	(1) ชี้นิ้วชี้ของมือข้างหนึ่งไปด้านหน้า (2) ชูนิ้วโป้งโดยใช้มือข้างเดียวกัน	0.9
สวัสดี	(1) แบมือและหันฝ่ามือไปแตะที่ศีรษะแล้วเคลื่อนฝ่ามือออกมาด้านหน้า	1.0
ทำอะไร	(1) กำมือสองข้างข้างหนึ่งอยู่ด้านบน อีกข้างอยู่ด้านล่าง เคลื่อนที่เข้ามาหากันบริเวณกลางลำตัว (2) แบมือไว้บริเวณกลางลำตัว แล้วม้วนฝ่ามือออกไปด้านนอก	1.6
ถูกไหม	(1) มือข้างหนึ่งชูนิ้วก้อยแล้วเคลื่อนที่ขึ้นบริเวณมืออีกข้างหนึ่งที่กำลึงแบอยู่ (2) ใช้นิ้วชี้เคลื่อนที่เป็นรูปเครื่องหมายคำถาม	1.7
ผิดไหม	(1) มือข้างหนึ่งชูนิ้วก้อยแล้วเคลื่อนที่ปิดลงบริเวณมืออีกข้างหนึ่งที่กำลึงแบอยู่ (2) ใช้นิ้วชี้เคลื่อนที่เป็นรูปเครื่องหมายคำถาม	1.5
ดูภาพ	(1) ตั้งฝ่ามือสองข้าง ข้างหนึ่งอยู่บริเวณศีรษะ ข้างหนึ่งอยู่บริเวณไหล่ (2) เคลื่อนมือด้านบนมาหามือที่อยู่ด้านล่าง โดยนิ้วหัวแม่มือและนิ้วทั้งสี่ขยับเข้าหากัน (3) มือข้างหนึ่งใช้นิ้วชี้และนิ้วกลางชี้ไปยังมือที่แบอยู่ แล้วเคลื่อนมือลง	2.3
เรียบร้อย	(1) มือข้างหนึ่งหงายวางไว้ด้านล่าง มืออีกข้างหนึ่งคว่ำไว้ด้านบนบริเวณกลางลำตัว แล้วเคลื่อนมือ ออกจากกันไปด้านข้าง	1.6
น่ารัก	(1) ขยับนิ้วโป้งและนิ้วทั้งสี่เข้าและออกจากกันบริเวณข้างใบหน้า	1.1
ภูมิใจ	(1) ตั้งฝ่ามือข้างหนึ่งไว้กลางลำตัว แล้วเคลื่อนฝ่ามือขึ้นไปบริเวณหน้าอก	1.8
ถูกหรือผิด	(1) มือข้างหนึ่งชูนิ้วก้อยแล้วเคลื่อนที่ขึ้นบริเวณมืออีกข้างหนึ่งที่กำลึงแบอยู่ (2) มือข้างที่ชูนิ้วก้อย เคลื่อนที่ลงบริเวณมืออีกข้างหนึ่งที่กำลึงแบอยู่ (3) ใช้นิ้วชี้เคลื่อนที่ซ้ายขวากลับไปกลับมา	2.6
เก็บอุปกรณ์	(1) กางนิ้วหัวแม่มือ นิ้วชี้ และนิ้วกลางของทั้งสองมือให้อยู่ในแนวตั้งฉากกัน มือข้างหนึ่งอยู่ด้านบน อีกข้างอยู่ด้านล่าง แล้วเคลื่อนที่เข้าหากัน (2) หายฝ่ามือโดยอนิ้วเข้าหากัน ขยับมือขึ้นสลับลงเล็กน้อยพร้อมกับเคลื่อนมือไปด้านข้าง	2.0
ตั้งใจเรียน	(1) ตั้งฝ่ามือสองข้างไว้ที่บริเวณข้างศีรษะแล้วเคลื่อนที่ลงมา (2) แบมือข้างหนึ่งไว้บริเวณหน้าอก มืออีกข้างนิ้วทั้งห้าแตะกันและเคลื่อนที่ขึ้นไปบริเวณศีรษะ	1.7
เข้าใจ	(1) กำมือไว้บริเวณข้างศีรษะ (2) ชูนิ้วชี้บริเวณข้างศีรษะจากนั้นขยับนิ้วชี้ลง	1.4
เก่งมาก	(1) แบมือข้างหนึ่งไว้ด้านล่าง มืออีกข้างอยู่ด้านบนชูนิ้วชี้กับนิ้วกลางแล้วเคลื่อนที่ลงมาหามือ ด้านล่าง (2) ตั้งฝ่ามือสองข้างแล้วเคลื่อนฝ่ามือออกจากกัน	1.7
ใครตอบได้	(1) นิ้วชี้ข้างหนึ่งเคลื่อนที่วนรอบบริเวณใบหน้า (2) ชูนิ้วทั้งสองมือไปบริเวณปากและปิดออกมา ด้านหน้า (3) กำมือข้างหนึ่งและชูนิ้วโป้ง	2.2

จากตารางพบว่าภาษามือในแต่ละประโยคจะมีท่ามือและการเคลื่อนไหวแตกต่างกันทำให้ความยาววิดีโอที่ได้แตกต่างกันออกไปตามความซับซ้อนของท่าภาษามือ เมื่อพิจารณาท่ามือและการเคลื่อนไหวของแต่ละประโยคภาษามือ พบว่า ประโยคภาษามือที่ใช้ในการวิจัยนี้มีจำนวนท่ามือ 1-3 ท่า และยังพบว่าบางประโยคมีท่ามือและการเคลื่อนไหวแตกต่างกันอย่างชัดเจน ในขณะที่บางประโยคมีท่ามือที่มีลักษณะคล้ายกันซึ่งอาจส่งผลต่อประสิทธิภาพของโมเดลรู้จำภาษามือไทย

2. ผลการพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

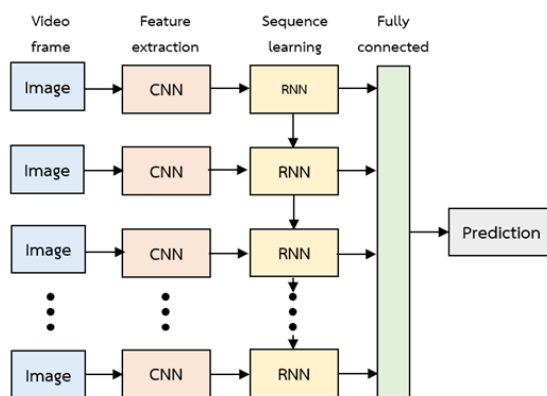
การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกในการวิจัยนี้เป็นการการประยุกต์ใช้โมเดล CNN ร่วมกับ RNN โดยการออกแบบโครงสร้างของโมเดลแสดงได้ดังภาพ 3 ซึ่งผู้วิจัยได้กำหนดชั้น (layer) ของโมเดลไว้ 4 ส่วน แต่ละส่วนมีรายละเอียด ดังนี้

(1) rescaling layer เป็นชั้นที่นำข้อมูลนำเข้า (input) จากภาพในแต่ละเฟรมของวิดีโอซึ่งมีค่าตั้งแต่ 0 ถึง 255 มาจัดกระทำให้มีค่าอยู่ระหว่าง 0 ถึง 1 เพื่อให้โมเดลสามารถเรียนรู้ได้มีประสิทธิภาพมากยิ่งขึ้น

(2) layer ของโมเดลในกลุ่ม CNN สำหรับดึงคุณลักษณะ (feature extraction) ได้แก่ time distributed layer ของโมเดล EfficientNetB0 และ DenseNet121 ซึ่งโดยทั่วไปโมเดลทั้งสองนี้เป็นโมเดลสำหรับข้อมูลที่เป็นรูปภาพ หากนำมาครอบด้วย time distributed layer จะทำให้โมเดลสามารถนำไปใช้กับข้อมูลที่เป็นวิดีโอได้ โดย time distributed layer จะแบ่งเฟรมภาพของวิดีโอแต่ละเฟรมเข้าไปในโมเดล EfficientNetB0 และ DenseNet121 และผู้วิจัยนำชั้นดังกล่าวมาเชื่อมต่อกับ time distributed global average pooling 2D เพื่อลดขนาดของข้อมูลสำหรับนำไปใช้ในส่วนต่อไป ทั้งนี้ในส่วนฐานของโมเดล CNN ทั้งสองนี้จะใช้ค่าพารามิเตอร์จากโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) บนชุดข้อมูลขนาดใหญ่ ImageNet โดยจะตัดในส่วนการจำแนก (classifier) ของโมเดลเดิมออก แล้วแทนที่ด้วยชั้นของโมเดลส่วนต่อไปที่ใช้ในการวิจัยนี้เพื่อให้สามารถจำแนกประโยคภาษามือได้อย่างมีประสิทธิภาพ

(3) layer ของโมเดลในกลุ่ม RNN สำหรับเรียนรู้ข้อมูลที่เป็นลำดับ (sequence learning) ได้แก่ โมเดล LSTM และ GRU โดยผู้วิจัยกำหนดจำนวนหน่วยประมวลผล (neuron) ไว้เท่ากับ 64

(4) dense layer หรือ fully connected layer เป็นชั้นสุดท้ายสำหรับการจำแนกวิดีโอประโยคภาษามือ 20 ประโยค หรือ 20 คลาส ผู้วิจัยจึงกำหนดจำนวนหน่วยประมวลผล (neuron) ไว้เท่ากับ 20



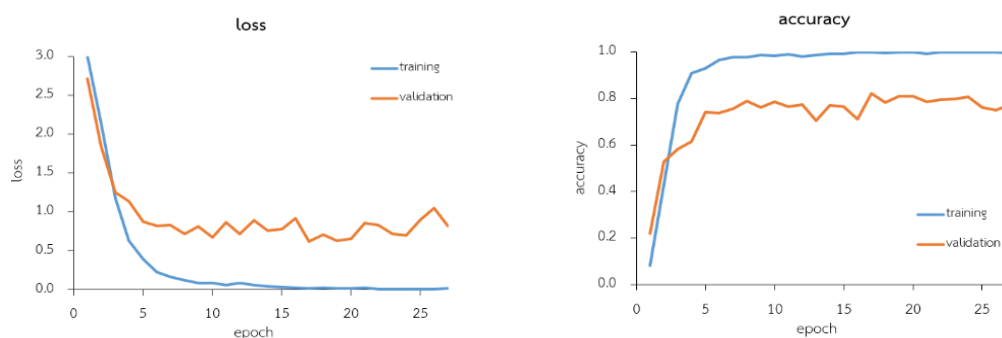
ภาพ 3 โครงสร้างของโมเดลการเรียนรู้เชิงลึกที่ใช้ในการวิจัย

โมเดลการเรียนรู้เชิงลึกสำหรับรู้จำภาษามือไทยที่มีโครงสร้างดังที่กล่าวมาข้างต้นเมื่อนำมาฝึกสอนและนำประสิทธิภาพของโมเดลมาเปรียบเทียบกับดังตาราง 3 พบว่า การใช้โมเดล EfficientNetB0 ร่วมกับ GRU ทำให้โมเดลการรู้จำภาษามือมีประสิทธิภาพมากที่สุดทั้งในด้านความถูกต้อง ขนาดของโมเดล และเวลาที่ใช้ในการฝึกสอนโมเดล รองลงมาคือ โมเดล EfficientNetB0 ร่วมกับ LSTM โมเดล DenseNet121 ร่วมกับ GRU และโมเดล DenseNet121 ร่วมกับ LSTM ตามลำดับ ซึ่งแสดงให้เห็นว่าโมเดลที่มีประสิทธิภาพในการทำนาย นั้นไม่จำเป็นต้องมีขนาดใหญ่และใช้เวลาฝึกสอนมาก นอกจากนี้เมื่อพิจารณาการฝึกสอนโมเดลจากกราฟในภาพ 4 ซึ่งแสดงค่าความสูญเสีย (loss) และค่าความถูกต้อง (accuracy) ในการฝึกสอนโมเดล EfficientNetB0 ร่วมกับ GRU ซึ่งพบว่าค่าความสูญเสียบนชุดข้อมูลฝึกสอน (training loss) และค่าความ

ถูกต้องบนชุดข้อมูลฝึกสอน (training accuracy) มีแนวโน้มลู่เข้าสู่ค่าคงที่ แสดงให้เห็นว่าโมเดลสามารถเรียนรู้จากชุดข้อมูลฝึกสอนได้ ในขณะที่ค่าความสูญเสียบนชุดข้อมูลตรวจสอบ (validation loss) และค่าความถูกต้องบนชุดข้อมูลตรวจสอบ (validation accuracy) มีการเพิ่มขึ้นสลับลดลงและมีแนวโน้มคงที่ แสดงให้เห็นว่าโมเดลยังไม่เกิดการเรียนรู้มากเกินไป (overfitting)

ตาราง 3 การเปรียบเทียบโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

model	validation loss	validation accuracy	size (MB)	average training time per epoch (s)
EfficientNetB0+GRU	0.8238	0.7750	16.44	541
EfficientNetB0+LSTM	1.2220	0.6969	16.77	549
DenseNet121+GRU	1.5096	0.5969	27.65	944
DenseNet121+LSTM	2.1585	0.4719	27.91	952

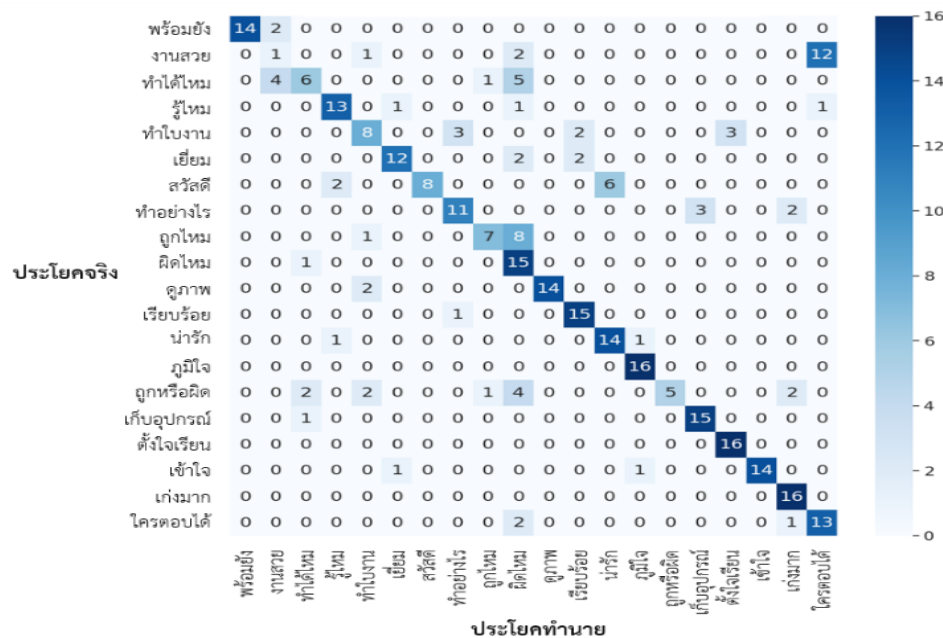


ภาพ 4 กราฟแสดงค่าความสูญเสียและค่าความถูกต้องของการฝึกสอนโมเดล EfficientNetB0 ร่วมกับ GRU

3. ผลการประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

โมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกที่ผู้วิจัยพัฒนาขึ้น เป็นโมเดลจำแนกแบบหลายกลุ่ม โดยใช้โมเดล CNN ร่วมกับ RNN จำนวนทั้งสิ้น 4 โมเดล และสามารถจำแนกประโยคภาษามือไทยได้จำนวน 20 ประโยค จากการพัฒนาโมเดลรู้จำภาษามือไทยพบว่าโมเดลที่มีประสิทธิภาพสูงที่สุดคือโมเดล EfficientNetB0 ร่วมกับ GRU ผู้วิจัยจึงนำโมเดลดังกล่าวมาประเมินประสิทธิภาพโดยใช้ชุดข้อมูลทดสอบ โดยวิเคราะห์ข้อมูลจากคอนฟิวชันเมทริกซ์ (confusion matrix) ซึ่งแสดงจำนวนตัวอย่างข้อมูลในประโยคที่โมเดลทำนายเทียบกับประโยคจริงว่าโมเดลทำนายได้ถูกต้องมากน้อยเพียงใด และมีการทำนายพลาดอย่างไรบ้าง

เมื่อให้โมเดลรู้จำภาษามือไทยจำแนกประโยคภาษามือในชุดข้อมูลทดสอบ โดยในแต่ละประโยคในชุดข้อมูลทดสอบจะมีข้อมูลจากวิดีโอจำนวนเท่ากันคือ 16 วิดีโอ ผลการจำแนกเมื่อเปรียบเทียบระหว่างประโยคจริงกับประโยคที่โมเดลทำนาย แสดงได้ด้วยคอนฟิวชันเมทริกซ์ขนาด 20×20 ดังภาพ 5 พบว่าประโยคที่โมเดลทำนายได้ถูกต้องทั้งหมดมีจำนวน 4 ประโยค ได้แก่ พร้อมยัง ภูมิใจ ตั้งใจเรียน และเก่งมาก ส่วนประโยคที่โมเดลทำนายได้ถูกต้องน้อยที่สุด คือ งานสวย ซึ่งมักมีการทำนายพลาดเป็น ใครตอบได้ นอกจากนี้ยังมีประโยคอื่น ๆ ที่โมเดลมักทำนายพลาดเกินครึ่งหนึ่งของข้อมูลตัวอย่างในประโยคนั้น ได้แก่ ทำได้ไหม ถูกไหม และถูกหรือผิด



ภาพ 5 คอนฟิวชันเมทริกซ์ของโมเดลรู้จำภาษาไทย

เมื่อนำค่าที่ได้จากคอนฟิวชันเมทริกซ์มาคำนวณค่าความถูกต้อง (accuracy) เพื่อประเมินประสิทธิภาพการจำแนกของโมเดลรู้จำภาษาไทยในภาพรวม ค่าความแม่นยำ (precision) ค่าการเรียกคืน (recall) และคะแนน F1 (F1-score) เพื่อประเมินประสิทธิภาพการจำแนกของโมเดลรู้จำภาษาไทยในแต่ละประโยค พบว่า โมเดลมีค่าความถูกต้องเท่ากับ 0.73 ค่าความแม่นยำเฉลี่ยเท่ากับ 0.76 ค่าการเรียกคืนเฉลี่ยเท่ากับ 0.73 และคะแนน F1 เฉลี่ยเท่ากับ 0.72 ดังตาราง 4 ซึ่งแสดงให้เห็นว่าโมเดลรู้จำภาษาไทยโดยใช้การเรียนรู้เชิงลึกที่พัฒนาขึ้นสามารถจำแนกประโยคภาษาไทยได้โดยมีประสิทธิภาพอยู่ในระดับปานกลาง

เมื่อประเมินประสิทธิภาพของโมเดลรู้จำภาษาไทยในแต่ละประโยค โดยพิจารณาประโยคภาษาไทยที่มีค่าความแม่นยำ ค่าการเรียกคืน และคะแนน F1 สูงสุดและต่ำสุด 3 อันดับแรก ดังตาราง 4 และภาพ 6 พบว่า

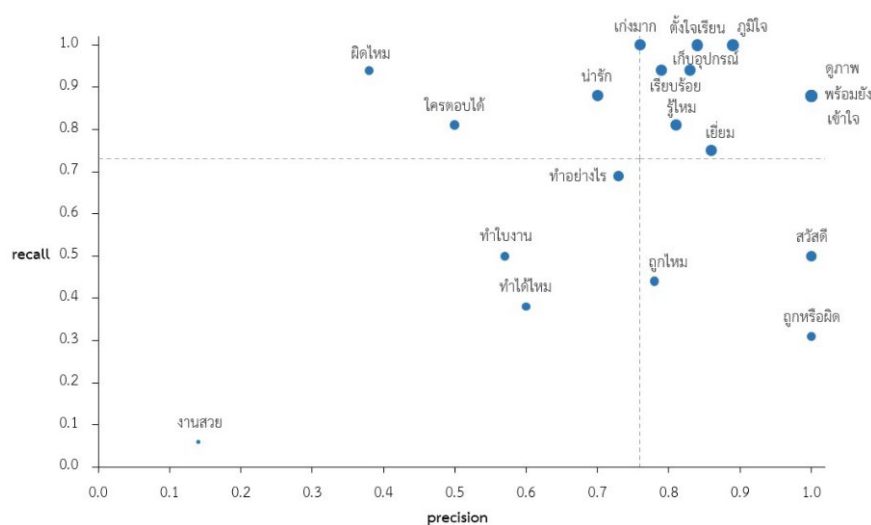
(1) ประโยคที่มีค่าความแม่นยำ (precision) มากที่สุด คือ พร้อมยัง ดูภาพ เข้าใจ สวัสดี และถูกหรือผิด มีค่าความแม่นยำเท่ากันคือ 1.00 เป็นประโยคที่โมเดลสามารถทำนายได้แม่นยำ ส่วนประโยคที่มีค่าความแม่นยำต่ำที่สุด คือ งานสวย ผิดไหม และทำใบงาน เป็นประโยคที่โมเดลสามารถทำนายได้ไม่แม่นยำ

(2) ประโยคที่มีค่าการเรียกคืน (recall) มากที่สุด คือ ภูมิใจ ตั้งใจเรียน และเก่งมาก มีค่าการเรียกคืนเท่ากันคือ 1.00 เป็นประโยคที่โมเดลไม่ค่อยพลาดการทำนาย ส่วนประโยคที่มีค่าการเรียกคืนต่ำที่สุดคือ “งานสวย” “ถูกหรือผิด” และ “ทำได้ไหม” เป็นประโยคที่โมเดลพลาดการทำนายบ่อยครั้ง

(3) ประโยคที่มีคะแนน F1 (F1-score) มากที่สุด คือ พร้อมยัง ดูภาพ และเข้าใจ เป็นประโยคที่โมเดลสามารถจำแนกได้ดี ส่วนประโยคที่มีค่าคะแนน F1 ต่ำที่สุดคือ คือ งานสวย ทำได้ไหม และถูกหรือผิด เป็นประโยคที่โมเดลยังไม่สามารถจำแนกได้ดี

ตาราง 4 ประสิทธิภาพของโมเดลรู้จำภาษาไทย

ประโยค	precision	recall	F1 – score
ภูมิใจ	0.89	1.00	0.94
พร้อมยัง	1.00	0.88	0.93
ดูภาพ	1.00	0.88	0.93
เข้าใจ	1.00	0.88	0.93
ตั้งใจเรียน	0.84	1.00	0.91
เก็บอุปกรณ์	0.83	0.94	0.88
เรียบร้อย	0.79	0.94	0.86
เก่งมาก	0.76	1.00	0.86
รู้ไหม	0.81	0.81	0.81
เยี่ยม	0.86	0.75	0.80
น่ารัก	0.70	0.88	0.78
ทำอะไร	0.73	0.69	0.71
สวัสดี	1.00	0.50	0.67
ใครตอบได้	0.50	0.81	0.62
ถูกไหม	0.78	0.44	0.56
ผิดไหม	0.38	0.94	0.55
ทำใบงาน	0.57	0.50	0.53
ถูกหรือผิด	1.00	0.31	0.48
ทำได้ไหม	0.60	0.38	0.46
งานสวย	0.14	0.06	0.09
เฉลี่ย	0.76	0.73	0.72
accuracy = 0.73			



ภาพ 6 แผนภาพการกระจายค่า precision และ recall ของแต่ละประโยคภาษาไทย โดยที่ขนาดจุดแทนค่า F1-score

อภิปรายผล

1. การศึกษาภาษามือที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษ พบว่า ภาษามือที่ใช้มักเป็นประโยคที่สั้นและกะทัดรัดมากกว่าภาษาพูด เนื่องจากภาษามือไม่ได้มีคำศัพท์หลากหลายเท่ากับภาษาพูด และประโยคที่สั้นและกะทัดรัดนี้จะช่วยให้ครูและนักเรียนที่มีความต้องการพิเศษสามารถสื่อสารและทำความเข้าใจกันได้อย่างสะดวกและรวดเร็ว นอกจากนี้จากการสอบถามครูที่ใช้ภาษามือในการจัดการเรียนการสอนพบว่า ภาษามือส่วนใหญ่ที่ครูใช้จัดอยู่ในหมวดการเสริมแรง ซึ่งเป็นการกล่าวชื่นชมและให้กำลังใจแก่นักเรียน ทั้งนี้เป็นเพราะว่าการเสริมแรงเป็นเทคนิคในการสร้างกำลังใจให้กับนักเรียนที่สามารถใช้ได้กับการสอนในทุกรายวิชา โดยประโยคที่ใช้ในการเสริมแรงแต่ละรายวิชามักจะเป็นประโยคในลักษณะเดียวกันทำให้ประโยคภาษามือในหมวดนี้มีการใช้งานมากที่สุด ในขณะที่ประโยคภาษามือในหมวดอื่นนั้นอาจมีการใช้งานที่แตกต่างกันออกไปตามแต่ละรายวิชา ดังนั้นในการเรียนรู้และฝึกฝนเกี่ยวกับการใช้ภาษามือเพื่อการสื่อสารของครูหรือนิสิตนักศึกษาครูที่ไม่ได้มีพื้นฐานเกี่ยวกับภาษามือสามารถเริ่มศึกษาได้จากภาษามือในหมวดการเสริมแรงที่สามารถใช้ได้กับทุกรายวิชา จากนั้นจึงศึกษาภาษามือในหมวดอื่นเพิ่มเติมที่แตกต่างกันไปในแต่ละรายวิชา

2. การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก พบว่า โมเดลรู้จำภาษามือไทยที่พัฒนาขึ้นโดยใช้โมเดล EfficientNetB0 ร่วมกับ GRU สามารถจำแนกประโยคภาษามือได้มีประสิทธิภาพสูงสุดในด้านความถูกต้อง เนื่องจากเป็นการใช้กลุ่มโมเดล CNN ที่สามารถดึงคุณลักษณะจากรูปภาพในแต่ละเฟรมของวิดีโอ และโมเดลกลุ่ม RNN ที่สามารถเรียนรู้ข้อมูลที่เป็นลำดับจากแต่ละเฟรมของวิดีโอ สอดคล้องกับงานวิจัยของ Puranik et al. (2022) ที่มีการใช้โมเดล VGG16 ร่วมกับ LSTM ในการรู้จำภาษามืออเมริกัน และงานวิจัย Huang & Chouvatut (2024) ที่มีการใช้โมเดล ResNet ร่วมกับ LSTM ในการรู้จำภาษามืออาร์เจนตินา ทำให้สามารถรู้จำภาษามือได้อย่างมีประสิทธิภาพ นอกจากนี้การใช้โมเดล EfficientNetB0 เป็นโมเดลที่มีผ่านการฝึกมาแล้ว (pretrained model) กับชุดข้อมูลขนาดใหญ่ และโมเดล EfficientNetB0 กับ GRU เป็นโมเดลที่มีขนาดเล็ก ทำให้ใช้เวลาในการฝึกสอนโมเดลน้อยกว่าโมเดลอื่น ๆ อย่างไรก็ตามโมเดลในการวิจัยนี้จะสามารถรู้จำภาษามือได้ดีในระดับหนึ่ง แต่ถ้าหากมีขนาดตัวอย่างที่ใช้ในการพัฒนาโมเดลจำนวนมากขึ้นก็จะช่วยให้โมเดลมีประสิทธิภาพสูงขึ้นได้อีก

3. การประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทย พบว่า โมเดลรู้จำภาษามือไทยที่มีประสิทธิภาพสูงสุดในการวิจัยนี้สามารถจำแนกประโยคภาษามือจำนวน 20 ประโยคได้ถูกต้องร้อยละ 73 จากประโยคภาษามือทั้งหมด เมื่อพิจารณาประสิทธิภาพในการจำแนกในแต่ละประโยคภาษามือโดยใช้ค่าคะแนน F1 พบว่า ประโยคที่โมเดลสามารถจำแนกได้ดี ได้แก่ ภูมิใจ พร้อมยัง ดูภาพ และเข้าใจ อาจเป็นเพราะประโยคดังกล่าวเป็นประโยคที่มีท่ามือเพียง 1-2 ท่า ไม่มีการเคลื่อนไหวของมือที่ซับซ้อน เป็นเพียงการเคลื่อนไหวแค่ 1-2 จังหวะ และท่ามือกับการเคลื่อนไหวของมือมีความแตกต่างจากประโยคอื่น ๆ ในขณะที่ประโยคที่โมเดลจำแนกได้ไม่ดี ได้แก่ งานสวย ทำได้ไหม ถูกหรือผิด อาจเป็นเพราะประโยคดังกล่าวมีการเปลี่ยนท่ามือถึง 3 ท่า มีการเคลื่อนไหวของมือ 2-3 จังหวะ และท่ามือกับการเคลื่อนไหวของมือมีความคล้ายกับประโยคอื่น ๆ เช่นเดียวกับงานวิจัยของ Balaha et al. (2022) ที่พบว่าโมเดลรู้จำภาษามืออาหรับจำแนกภาษามือที่มีการเคลื่อนไหวคล้ายกันได้ไม่ดี สำหรับในการวิจัยนี้ประโยคที่โมเดลจำแนกได้ไม่ดี คือ “งานสวย” มักถูกทำนายผิดเป็น “ใคร

ตอบได้” เนื่องจากในท่าภาษามือของประโยคทั้งสองมีการเคลื่อนที่ของมือที่คล้ายกันคือมีการเคลื่อนที่ของมือวนรอบบริเวณใบหน้า นอกจากนี้คำว่า “งาน” กับ “ทำ” มีท่าทางและการเคลื่อนไหวที่คล้ายกันคือมีการกำมือและเคลื่อนมือเข้าหากัน และประโยคที่มีท่าทางคล้ายกันอีกคือ “ถูกไหม” “ผิดไหม” และ “ถูกหรือผิด” เนื่องจากประโยคทั้งสามมีคำที่เหมือนกันทำให้ท่าทางคล้ายกัน ซึ่งท่าทางที่คล้ายกันนี้อาจเป็นสาเหตุที่ทำให้โมเดลทำนายพลาดไป ทั้งนี้ สาเหตุที่โมเดลจำแนกประโยคที่มีท่ามือซับซ้อนและมีท่าคล้ายประโยคอื่น ๆ ได้ไม่ดีนั้น อาจเกิดขึ้นได้จากการกำหนดจำนวนเฟรมที่ดึงจากแต่ละวิดีโอมีจำนวนน้อย ทำให้การดึงรายละเอียดการเคลื่อนไหวของท่ามืออาจได้มาไม่ครบถ้วน ดังนั้นในการพัฒนาโมเดลครั้งต่อไปหากมีการกำหนดจำนวนเฟรมในแต่ละวิดีโอที่เหมาะสมในการพัฒนาโมเดล ก็จะช่วยให้โมเดลสามารถเรียนรู้รายละเอียดต่าง ๆ ของการเคลื่อนไหวแต่ละท่ามือได้มากขึ้น และทำให้โมเดลมีประสิทธิภาพสูงขึ้นได้

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1) ภาษามือที่ครูใช้ในการสื่อสารกับนักเรียนที่มีความต้องการพิเศษในการวิจัยนี้ ซึ่งประกอบด้วยหมวดการนำเข้าสู่บทเรียน การตรวจสอบความเข้าใจ การออกคำสั่ง การเสริมแรง และการโต้ตอบของนักเรียนสามารถนำไปใช้เป็นภาษามือพื้นฐานสำหรับครูและนิสิตนักศึกษาครูในการเรียนรู้และฝึกฝนการใช้ภาษามือในการสื่อสารกับนักเรียนที่มีความต้องการพิเศษ และภาษามือเหล่านี้สามารถนำไปพัฒนาโมเดลรู้จำภาษามือในชั้นเรียนที่มีประสิทธิภาพมากขึ้นได้

2) โมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกที่มีการใช้โมเดล EfficientNetB0 ร่วมกับ GRU ซึ่งในการวิจัยนี้พบว่ามีประสิทธิภาพสูงสุด สามารถนำไปพัฒนาเป็นโมเดลในแอปพลิเคชันสำหรับให้ครูและนิสิตครูสามารถเรียนรู้ประโยคภาษามือไทยพื้นฐานในชั้นเรียนและฝึกสอนการทำภาษามือ เพื่อให้สามารถสื่อสารกับนักเรียนที่มีความต้องการพิเศษได้อย่างมีประสิทธิภาพ

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

1) การศึกษาภาษามือที่ครูใช้ในการสื่อสารกับนักเรียนที่มีความต้องการพิเศษ สามารถศึกษาภาษามือที่ใช้ในแต่ละรายวิชา เนื่องจากแต่ละรายวิชามีคำศัพท์และประโยคที่เฉพาะตามเนื้อหาหรือวิธีการจัดการเรียนการสอนในรายวิชานั้น ๆ ซึ่งจะทำให้ภาษามือที่ใช้ในการพัฒนาโมเดลมีความหลากหลายมากยิ่งขึ้น ซึ่งสามารถนำโมเดลรู้จำภาษามือจากการวิจัยนี้ไปเป็นโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) สำหรับพัฒนาโมเดลรู้จำภาษามือครั้งต่อไปได้โดยไม่ต้องเริ่มใหม่ตั้งแต่ต้น

2) การประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกจากไฟล์วิดีโอ พบว่าโมเดลมีข้อจำกัดในการจำแนกบางประโยคภาษามือได้อย่างแม่นยำ ปัญหานี้เกิดจากความซับซ้อนของการเคลื่อนไหวในวิดีโอ เช่น การเปลี่ยนตำแหน่งของมือ ความสัมพันธ์เชิงเวลาของแต่ละเฟรมวิดีโอ และการเคลื่อนไหวที่คล้ายกันของประโยคที่ภาษามือแตกต่างกัน ส่งผลให้โมเดลมีแนวโน้มที่จะสับสนระหว่างประโยคที่มีลักษณะการเคลื่อนไหวคล้ายกัน ดังนั้นเพื่อปรับปรุงประสิทธิภาพของโมเดลสำหรับการใช้งานจริงในงานคอมพิวเตอร์วิทัศน์ (computer vision) โดยข้อมูลนำเข้าเป็นไฟล์วิดีโอ สามารถดำเนินการได้ดังนี้

(1) ขยายและปรับปรุงชุดข้อมูล เช่น เพิ่มตัวอย่างวิดีโอในกลุ่มประโยคที่มีการซับซ้อนร่วมกับการเพิ่มจำนวนข้อมูล (data augmentation) หรือเพิ่มจำนวนเฟรมของวิดีโอ และ (2) ปรับปรุงสถาปัตยกรรมของโมเดล โดยอาจเลือกใช้โมเดลในกลุ่ม attention หรือ transformer เช่น vision transformer (ViT) และ temporal attention mechanism ที่สามารถจับความสัมพันธ์ระหว่างเฟรมได้ดี ซึ่งอาจช่วยให้โมเดลมีประสิทธิภาพดีขึ้นได้

References

- Amatanon, V., Chanhang, S., Naiyanetr, P., & Thongpang, S. (2014). Sign language-Thai alphabet conversion based on Electromyogram (EMG). In *2014 7th Biomedical Engineering International Conference (BMEiCON)* (pp. 1–4). IEEE.
<https://doi.org/10.1109/BMEiCON.2014.7017398>
- Balaha, M. M., El-Kady, S., Balaha, H. M., Salama, M., Emad, E., Hassan, M., & Saafan, M. M. (2022). A vision-based deep learning approach for independent-users Arabic sign language interpretation. *Multimedia Tools and Applications*, 82(5), 6807–6826.
<https://doi.org/10.1007/s11042-022-13423-9>
- Chaikaew, A. (2022). An Applied Holistic Landmark with Deep Learning for Thai Sign Language Recognition. In *2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)* (pp. 1046–1049). IEEE.
<https://doi.org/10.1109/ITC-CSCC55581.2022.9895052>
- Chansri, C., & Srinonchat, J. (2016). Hand gesture recognition for Thai sign language in complex background using fusion of depth and color video. *Procedia Computer Science*, 86, 257–260. <https://doi.org/10.1016/j.procs.2016.05.113>
- Gedkhaw, E. (2022). The Performance of Thai Sign Language Recognition with 2D Convolutional Neural Network Based on NVIDIA Jetson Nano Developer Kit. *TEM Journal*, 11(1), 411–419. <https://doi.org/10.18421/TEM111-52>
- Huang, J., & Chouvatut V. (2024). Video-Based Sign Language Recognition via ResNet and LSTM Network. *Journal of Imaging*, 10(6), 149.
<https://doi.org/10.3390/jimaging10060149>
- Kelly, J. F., McKinney E. L., & Swift, O. (2020). Strengthening teacher education to support deaf learners. *International Journal of Inclusive Education*, 26(13), 1289–1307.
<https://doi.org/10.1080/13603116.2020.1806366>
- Klomsae, A., Auephanwiriyaikul, S., & Theera-Umpon, N. (2017). A Novel String Grammar Unsupervised Possibilistic C-Medians Algorithm for Sign Language Translation Systems. *Symmetry*, 9(12), 321. <https://doi.org/10.3390/sym9120321>

- Pariwat, T., & Seresangtakul, P. (2021). Multi-stroke thai finger-spelling sign language recognition system with deep learning. *Symmetry*, 13(2), 262.
<https://doi.org/10.3390/sym13020262>
- Puranik, V., Gawande, V., Gujarathi, J., Patani, A., & Rane, T. (2022). Video-based Sign Language Recognition using Recurrent Neural Networks. In *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)* (pp. 1-6). IEEE.
<https://doi.org/10.1109/ASIANCON55314.2022.9909061>
- Saengsri, S., Niennattrakul, V., & Ratanamahatana, C. A. (2012). TFRS: Thai finger-spelling sign language recognition system. In *2012 second international conference on digital information and communication technology and it's applications (DICTAP)* (pp. 457-462). IEEE. <https://doi.org/10.1109/DICTAP.2012.6215407>
- Sharma, S., & Singh, S. (2023). ISL recognition system using integrated mobile-net and transfer learning method. *Expert Systems with Applications*, 221, 119772.
<https://doi.org/10.1016/j.eswa.2023.119772>
- Sripairojthikoon, N., & Harnsomburana, J. (2019). Thai Sign Language Recognition Using 3D Convolutional Neural Networks. In *Proceedings of the 2019 7th International Conference on Computer and Communications Management* (pp. 186-189). Association for Computing Machinery. <https://doi.org/10.1145/3348445.3348452>
- Vachirapipop, M., Soymat, S., Tiraronnakul, W., & Hnoohom, N. (2017, November). Sign Translation with Myo Armbands. In *2017 21st International Computer Science and Engineering Conference (ICSEC)* (pp. 1-5). IEEE.
<https://doi.org/10.1109/ICSEC.2017.8443836>
- Vijitkunsawat, W., Racharak, T., Nguyen, C., & Le Minh, N. (2023). Video-Based Sign Language Digit Recognition for the Thai Language: A New Dataset and Method Comparisons. In *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods* (pp. 775-782). SciTePress.
<https://doi.org/10.5220/0011643700003411>

Translate Thai References

- Boonya, R. (2008). Sign Language: Language of the Deaf. *Journal of Ratchasuda Institute*, 4(1), 77-94. (in Thai)
- Buangamdee, C. (2017) *A Translation of Thai Language to Thai Sign Language with Rule-based Technique* [Master's thesis]. Naresuan University. (in Thai)

- Jaruwanawat, A. (2007). *Continuous Thai Sign Language Recognition Using Hidden Markov Models* [Master's thesis]. King Mongkut Institute of Technology Ladkrabang. (in Thai)
- Petchporee, S. (2005). *A Study on Sign Language Vocabulary Memory of Hearing Impaired Children in Prathomsuksa 1 by Minor Game* [Master's thesis]. Srinakarinwirot University. (in Thai)
- Plianwong, N., & Kangkachit, T. (2021). Thai Sign Language Translation System Through Few Shot Learning. *DPU Graduate Studies Journal*, 10(2), 119-132. (in Thai)
- Sakulsujirapa, B. (2011). *Development of a Thai Sign Language Recognition System using Hidden Markov Models* [Master's thesis]. Prince of Songkla University. (in Thai)
- Sriboonruang, P., Tanachai, P., Khaggathog, K. Vijitkunsawat, W., & Anunvrapong, P. (2022). Software Sign Language Translator to Text and Speech by Using the Landmarks Technique of MediaPipe. *The Journal of Industrial Technology : Suan Sunandha Rajabhat University*, 10(2), 66-76. (in Thai)
- Veerasakulthong, W. (2007). *Thai Hand Sign Recognition* [Master's thesis]. National Institute of Development Administration. (in Thai)

การพัฒนาแบบสอบสมรรถนะดิจิทัลในระดับปริญญาตรี
โดยใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ
Development of the Digital Competency Test for Bachelor's Degree
Using the Multidimensional Item Response Theory

อัศนีย์ ทองศิลป์^{1*} และ กมลวรรณ ตังชันกานนท์²

Asanee Tongsilp^{1*} and Kamonwan Tangdhanakanond²

(Received: October 30, 2024; Revised: December 5, 2024; Accepted: December 9, 2024)

บทคัดย่อ

งานวิจัยนี้มีความมุ่งหมายเพื่อ 1) พัฒนาโมเดลสมรรถนะดิจิทัลสำหรับคุณวุฒิระดับปริญญาตรี 2) พัฒนาแบบสอบสมรรถนะดิจิทัลสำหรับคุณวุฒิระดับปริญญาตรี โดยใช้ทฤษฎีการตอบสนองแบบพหุมิติ และ 3) ศึกษาสมรรถนะดิจิทัลของนักศึกษาในระดับปริญญาตรี งานวิจัยนี้ใช้ระเบียบวิธีวิจัยแบบวิจัยและพัฒนา ตัวอย่างวิจัย คือ นักศึกษาระดับปริญญาตรี ในสถาบันอุดมศึกษา จำนวน 1,256 คน ใช้วิธีการสุ่มตัวอย่างแบบหลายขั้นตอน เครื่องมือวิจัยเป็นแบบสอบวัดสมรรถนะดิจิทัล แบบเลือกตอบชนิด 4 ตัวเลือก จำนวน 30 ตัวชี้วัด 240 ข้อ สถิติที่ใช้ในการวิจัย ได้แก่ สถิติเชิงบรรยาย การวิเคราะห์คุณภาพข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ และการวิเคราะห์องค์ประกอบเชิงยืนยัน ผลการวิจัยพบว่า สมรรถนะดิจิทัลสำหรับปริญญาตรี ประกอบด้วย 1. การรู้เท่าทันสื่อสารสนเทศและดิจิทัล 2. ทักษะการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัล และ 3. การปรับตัวสู่การเปลี่ยนแปลงทางดิจิทัล แบบทดสอบสอดคล้องกับโมเดลการตอบสนองข้อสอบแบบพหุมิติชนิดแยกตามองค์ประกอบแบบ 3 พารามิเตอร์ มีข้อสอบผ่านเกณฑ์ 211 ข้อ (ร้อยละ 87.92) มีค่าความยากเฉลี่ย 1.44 ค่าอำนาจจำแนกเฉลี่ย 0.24 และค่าการเดาเฉลี่ย 0.13 มีค่าดัชนีความเหมาะสมรายข้อ OUTFIT MNSQ อยู่ระหว่าง 0.62 ถึง 1.05 และ INFIT MNSQ อยู่ระหว่าง 0.74 ถึง 1.14 แบบสอบมีค่าสัมประสิทธิ์ความเที่ยงแบบ EAP ทั้งฉบับ เท่ากับ 0.865 และโมเดลสมรรถนะดิจิทัลระดับปริญญาตรีมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ ($\chi^2 = 233.011$, p-value = 0.217, $\chi^2/df = 1.074$, AGFI = 0.975, RMSEA = 0.008) ผลการศึกษาสมรรถนะดิจิทัลพบว่า นักศึกษามีคะแนนมากที่สุดคือ ทักษะการสื่อสารในยุคดิจิทัล และน้อยที่สุดคือ การคิดเชิงคำนวณ ($M = 5.99$ และ 2.40 ตามลำดับ)

คำสำคัญ สมรรถนะดิจิทัล แบบสอบสมรรถนะดิจิทัล การตอบสนองข้อสอบแบบพหุมิติ

¹ผู้ช่วยศาสตราจารย์ สำนักวิจัย มหาวิทยาลัยเกษมบัณฑิต

²รองศาสตราจารย์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹Assistant Professor, Research Center, Kasem Bundit University

²Associate Professor, Faculty of Education, Chulalongkorn University

ได้รับทุนอุดหนุนการวิจัยและนวัตกรรมจากสำนักงานการวิจัยแห่งชาติ (วช.): N42A660894

This project is funded by National Research Council of Thailand (NRCT): N42A660894

* Corresponding Author E-mail: asanee.ton@kbu.ac.th

Abstract

This research aimed to: 1) develop a digital competency model for bachelor's degree students, 2) develop a digital competency assessment using the multidimensional item response theory (MIRT), and 3) examine the digital competency levels of undergraduate students. The study employed a research and development methodology. The sample consisted of 1,200 undergraduate students from higher education institutions, selected through a multi-stage sampling technique. The research instrument was a four-option multiple-choice digital competency test comprising 30 indicators and 240 items. Data were analyzed using descriptive statistics, item quality analysis based on MIRT, and confirmatory factor analysis (CFA). The findings revealed that the digital competency model for bachelor's degree students consisted of three components: 1) Media, Information, and Digital Literacy, 2) Skills in Using, Developing, and Solving Problems with Digital Tools, and 3) Adaptive Digital Transformation. The test was consistent with empirical data and aligned with the between-item multidimensional 3-parameter model. A total of 211 items (87.92%) met the quality criteria, with an average difficulty of 1.44, average discrimination power of 0.24, and an average guessing parameter of 0.13. Item-fit indices ranged from 0.62 to 1.05 for OUTFIT MNSQ and 0.74 to 1.14 for INFIT MNSQ. The test demonstrated high reliability, with an EAP reliability coefficient of 0.865. The digital competency model exhibited good fit with the empirical data ($\chi^2 = 233.011$, $p = .217$, $\chi^2/df = 1.074$, AGFI = 0.975, RMSEA = 0.008). Results from the competency assessment indicated that students scored highest in Digital Communication Skills and the lowest score in Computational Thinking ($M = 5.99$ and 2.40 respectively).

Keywords: digital competency, digital competency test, multidimensional item response

บทนำ

ปัจจุบันเทคโนโลยีดิจิทัลมีความสำคัญต่อการดำเนินชีวิตและการทำงานมากขึ้น อีกทั้งยังมีบทบาทสำคัญต่อการขับเคลื่อนและพัฒนาโครงสร้างของประเทศไทยให้ก้าวเข้าสู่ยุค Thailand 4.0 สอดคล้องกับนโยบายชาติ (20 ปี) ว่าด้วยการพัฒนาดิจิทัลเพื่อเศรษฐกิจและสังคม พ.ศ. 2561–2580 มีเป้าหมายเพื่อเพิ่มขีดความสามารถในการแข่งขันทางเศรษฐกิจของประเทศด้วยการใช้นวัตกรรมและเทคโนโลยีดิจิทัล เตรียมความพร้อมให้บุคลากรทุกกลุ่มมีความรู้และทักษะที่เหมาะสมต่อการดำเนินชีวิต และการประกอบอาชีพในยุคดิจิทัล (Jarupoom, 2016) แต่อย่างไรก็ตามเยาวชนยังคงมีปัญหาด้านความปลอดภัยจากการใช้สื่อออนไลน์ ในหลายด้าน เช่น การถูกขโมยอัตลักษณ์ การเสพติดเกมและการพนันออนไลน์ การใช้ประทุษวาจา (hate speech) ปัญหาข่าวปลอม โดยเฉพาะการรังแกในพื้นที่ไซเบอร์ ทั้งนี้พบว่าเยาวชนเป็นผู้กระทำร้อยละ 43 เป็นผู้ถูกกระทำร้อยละ 49 ซึ่งเป็นผลมาจากการขาดความรู้เท่าทันสื่อออนไลน์ (Samae et al., 2021)

ทั้งนี้เพื่อส่งเสริมให้เยาวชนมีความรู้เท่าทันทางดิจิทัลมากขึ้น คณะกรรมการมาตรฐานการอุดมศึกษา จึงมีประกาศคณะกรรมการมาตรฐานการอุดมศึกษา ลงวันที่ 9 กันยายน 2565 เรื่อง รายละเอียดผลลัพธ์ การเรียนรู้ ตามมาตรฐานคุณวุฒิระดับอุดมศึกษา พ.ศ. 2565 กล่าวถึงผลลัพธ์การเรียนรู้ตามคุณวุฒิแต่ละระดับต้อง สอดคล้องกับอัตลักษณ์ของหลักสูตรโดยนักศึกษาต้องมีความรู้ และทักษะ เพื่อการประกอบอาชีพ ดำรงชีวิต อยู่ร่วมกันในสังคม และพัฒนาอย่างยั่งยืนสำหรับการดำรงชีวิตในยุคดิจิทัล โดยสถาบันการศึกษาจะต้อง ส่งเสริมทักษะด้านดิจิทัลและประเมินความรู้เมื่อนักศึกษาขอสำเร็จการศึกษา ดังนั้นเครื่องมือที่ใช้ควรเป็นแบบสอบ ที่สร้างขึ้นตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษา พ.ศ. 2565 อ้างอิงสมรรถนะจากแนวคิดทักษะการเรียนรู้ ในศตวรรษที่ 21 และ Top 10 Skills ของ World Economic Forum (Office of the Higher Education Commission, 2022)

การสร้างแบบสอบเพื่อวัดระดับสมรรถนะดิจิทัลพบว่า ปัจจุบันยังคงใช้แนวคิดวิธีการวัดและ ประเมินผลตามวิธีการทดสอบแบบดั้งเดิมที่มีข้อจำกัดในการประมาณค่าความสามารถที่แท้จริง ทั้งนี้มีแนวคิด ทฤษฎีการตอบสนองข้อสอบ (item response theory) ที่สามารถประมาณค่าความสามารถที่แท้จริงของ ผู้สอบได้ด้วยอาศัยโมเดลทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็นในการเลือกคำตอบถูก กับคุณลักษณะแฝงที่มีอิทธิพลอยู่เบื้องหลังการเลือกคำตอบนั้น ซึ่งการประมาณค่าความสามารถของผู้สอบ จะให้ผลได้แม่นยำมากกว่าวิธีการทดสอบแบบดั้งเดิม (Brossman & Guille, 2014) โมเดลการตอบสนอง ข้อสอบ ประกอบด้วยโมเดลเอกมิติ (unidimensional model) เป็นชุดของข้อคำถามที่มุ่งวัดคุณลักษณะแฝง เพียงคุณลักษณะเดียว และโมเดลพหุมิติ (multidimensional model) เป็นการวัดคุณลักษณะที่มี ความสัมพันธ์กัน 2 ลักษณะคือ ความเป็นพหุมิติระหว่างข้อคำถาม (between-items multidimensionality) และความเป็นพหุมิติภายในข้อคำถาม (within-items multidimensionality) (Chianchana, 2009)

งานวิจัยนี้นำโมเดลการตอบสนองข้อสอบแบบพหุมิติมาประยุกต์ใช้พัฒนาแบบสอบสมรรถนะดิจิทัล ในระดับปริญญาตรี ตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษาแห่งชาติ พ.ศ. 2565 เพื่อใช้ประเมินสมรรถนะ ดิจิทัลให้นักศึกษา ซึ่งเป็นการเตรียมความพร้อมให้มีความรู้ ทักษะต่อการดำเนินชีวิต และการประกอบ อาชีพในยุคดิจิทัล สมรรถนะดิจิทัลมีคุณลักษณะหลายมิติ เช่น การเข้าใจดิจิทัล การใช้ดิจิทัล การแก้ปัญหา ด้วยเครื่องมือดิจิทัล และการปรับตัวสู่การเปลี่ยนแปลงดิจิทัล (National Digital Economy and Society Commission, 2019) ฉะนั้นการใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติจึงเหมาะสมสำหรับนำมาใช้พัฒนา แบบสอบสมรรถนะดิจิทัลนี้ ซึ่งงานวิจัยนี้มุ่งวัดคุณลักษณะความเป็นพหุมิติระหว่างข้อคำถาม (between-items multidimensionality) ของสมรรถนะดิจิทัลในหลายมิติ โดยศึกษาความสัมพันธ์คุณลักษณะของ ข้อสอบและความสามารถของผู้สอบต่อความน่าจะเป็นของการตอบถูก (Chianchana, 2009)

ความมุ่งหมายของการวิจัย

1. เพื่อพัฒนาโมเดลการวัดสมรรถนะดิจิทัล ตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษา ในระดับ ปริญญาตรี
2. เพื่อพัฒนาแบบสอบสมรรถนะดิจิทัล ในระดับปริญญาตรี โดยใช้ทฤษฎีการตอบสนองข้อสอบ แบบพหุมิติ
3. เพื่อศึกษาสมรรถนะดิจิทัลของนักศึกษาในระดับปริญญาตรี

วิธีดำเนินการวิจัย

การวิจัยนี้ใช้ระเบียบวิธีวิจัยแบบการวิจัยและพัฒนา มีรายละเอียดดังนี้

1. ประชากรและกลุ่มตัวอย่างวิจัย

ประชากรวิจัย คือ นักศึกษาที่กำลังศึกษาระดับปริญญาตรี ปีการศึกษา 2566 ในสถาบันอุดมศึกษา ภายใต้การกำกับดูแลของกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม จำนวน 155 แห่ง แบ่งเป็น มหาวิทยาลัยในกำกับ 26 แห่ง มหาวิทยาลัยรัฐ 57 แห่ง และมหาวิทยาลัยเอกชน 72 แห่ง จำนวน 1,021,317 คน (Ministry of Higher Education, Science, Research and Innovation, 2023)

ตัวอย่างวิจัย คือ นักศึกษาที่กำลังศึกษาระดับปริญญาตรี ในสถาบันอุดมศึกษา ภายใต้การกำกับดูแลของกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม กำหนดขนาดตัวอย่างใช้เกณฑ์ของการวิเคราะห์ห้อยประกอบ ไม่ควรมีน้อยกว่า 200 คน อาจทำให้ผลการวิเคราะห์มีความคลาดเคลื่อนจากการวัดสูง ถ้ามีจำนวน 500 คน ถือว่าดี (Comrey & Lee, 2013) ดังนั้นผู้วิจัยจึงกำหนดขนาดตัวอย่างเป็น 1,256 คน ใช้วิธีการสุ่มตัวอย่างแบบหลายขั้นตอน ขั้นที่ 1 สุ่มสถาบันการศึกษา แบ่งเป็น มหาวิทยาลัยในกำกับ 1 แห่ง มหาวิทยาลัยรัฐ 3 แห่ง และมหาวิทยาลัยเอกชน 4 แห่ง ขั้นที่ 2 สุ่มนักศึกษาจากสาขาวิทยาศาสตร์สุขภาพ สาขาวิทยาศาสตร์-เทคโนโลยี และสาขาสังคมศาสตร์-มนุษยศาสตร์ จำนวน 3 สาขา ๆ ละ 50 คน จาก 8 แห่ง

2. เครื่องมือที่ใช้ในการวิจัย

แบบทดสอบสมรรถนะดิจิทัล ตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษาแห่งชาติ พ.ศ. 2565 สำหรับคุณวุฒิระดับปริญญาตรี เป็นแบบสอบเลือกตอบชนิด 4 ตัวเลือก ให้คะแนนแบบ 2 ค่า (ตอบถูกได้ 1 คะแนน และตอบผิดได้ 0 คะแนน) มี 30 ตัวชี้วัด แต่ละตัวชี้วัดมีข้อสอบ 8 ข้อ รวม 240 ข้อ เนื่องจากข้อสอบมีจำนวนมาก ผู้วิจัยจึงจัดข้อสอบออกเป็นฉบับย่อย ๆ จำนวน 7 ฉบับ แต่ละฉบับมีข้อสอบรวมและข้อสอบเฉพาะอย่างละ 30 ข้อ รวมฉบับละ 60 ข้อ 60 คะแนน ซึ่งข้อสอบแต่ละฉบับสร้างขึ้นตามนิยามจากตัวชี้วัดเดียวกันโดยมีข้อสอบรวมเป็นตัวเชื่อมโยงอยู่ในทุกฉบับ ทั้งนี้แสดงถึงความเป็นคู่ขนานกันในโครงสร้าง รูปแบบของข้อสอบ และเวลาสอบเหมือนกัน การเปรียบเทียบข้อสอบใช้วิธีการเปรียบเทียบพร้อมกัน ผลการเปรียบเทียบแสดงให้เห็นว่าข้อสอบทุกข้อมีค่าพารามิเตอร์อยู่บนสเกลเดียวกันซึ่งค่าพารามิเตอร์ความสามารถของผู้สอบที่ทำแบบสอบต่างฉบับกันจะอยู่บนสเกลเดียวกัน โดยใช้ชุดคำสั่ง mirt ในโปรแกรม R

3. การเก็บรวบรวมข้อมูล

การพัฒนาแบบทดสอบสมรรถนะดิจิทัล มีขั้นตอนในการดำเนินการวิจัยเป็น 2 ระยะ ดังนี้

ระยะที่ 1 การพัฒนาโมเดลการวัดสมรรถนะดิจิทัลสำหรับคุณวุฒิระดับปริญญาตรี

กำหนดจุดมุ่งหมายและกรอบการวัดของโมเดลการวัดสมรรถนะดิจิทัล เพื่อนำไปใช้สร้างแบบสอบวัดสมรรถนะดิจิทัล ในระดับปริญญาตรี ผู้วิจัยพัฒนาโมเดลการวัดจากการศึกษามาตรฐานคุณวุฒิระดับอุดมศึกษา พ.ศ. 2565 ทักษะการเรียนรู้ในศตวรรษที่ 21 และ Top 10 Skills ของ World Economic Forum หลักสูตรการเข้าใจดิจิทัล พ.ศ. 2562 ของสำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ เป้าหมายการพัฒนาที่ยั่งยืน (SDGs) นโยบายและแผนระดับชาติว่าด้วยการพัฒนาดิจิทัลเพื่อ

เศรษฐกิจและสังคม (พ.ศ. 2561–2580) แล้วจึงนำมาสังเคราะห์เพื่อระบุงค์ประกอบ ตัวชี้วัด และนิยามเชิงพฤติกรรมของตัวชี้วัด ด้วยวิธีการสนทนากลุ่มจากผู้ทรงคุณวุฒิด้านเทคโนโลยีสารสนเทศ จำนวน 10 คน

ระยะที่ 2 การพัฒนาแบบสอบถามสมรรถนะดิจิทัล สำหรับคุณวุฒิระดับปริญญาตรี

2.1 กำหนดจุดมุ่งหมาย ขอบเขตของแบบสอบ ตามกรอบคุณวุฒิระดับปริญญาตรี พ.ศ. 2565 และหลักสูตรการเข้าใจดิจิทัล พ.ศ. 2562 และแผนพัฒนาด้านดิจิทัลของประเทศไทย มีโครงสร้างและรูปแบบของแบบสอบ โดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงยืนยัน (Confirmatory MIRT Model) กำหนดเนื้อหา ได้แก่ (1) การรู้เท่าทันสื่อสารสนเทศและดิจิทัล (2) ทักษะการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัล และ (3) การปรับตัวสู่การเปลี่ยนแปลงดิจิทัล แสดงผังโครงสร้างข้อสอบ ดังตาราง 1

ตาราง 1 ผังโครงสร้างการออกข้อสอบ

สมรรถนะดิจิทัล ระดับปริญญาตรี					
1.การรู้เท่าทันสื่อสารสนเทศ	จำนวน (80 ข้อ)	2. ทักษะการใช้/พัฒนาฯ	จำนวน (112 ข้อ)	3. การปรับตัวฯ	จำนวน (48 ข้อ)
1.ความรู้พื้นฐานคอมพิวเตอร์	8	11. การใช้งานคอมพิวเตอร์	8	25. ยืดหยุ่น ปรับตัว	8
2. สิทธิและความรับผิดชอบ	8	12. การใช้อินเทอร์เน็ต	8	26. ทำงานร่วมในสังคม	8
3. การเข้าถึงสื่อดิจิทัล	8	13. โปรแกรมประมวลผลคำ	8	27. คิดริเริ่มเรียนรู้ด้วยตนเอง	8
4. การสื่อสารยุคดิจิทัล	8	14. โปรแกรมตารางวิเคราะห์	8	28. สร้าง/เป็นผู้ประกอบการ	8
5. ความปลอดภัยในยุคดิจิทัล	8	15. โปรแกรมนำเสนอ	8	29. เป็นผู้นำทางเทคโนโลยี	8
6. ความเข้าใจสื่อดิจิทัล	8	16. ความมั่นคงคอมพิวเตอร์	8	30. ติดตามเทคโนโลยีใหม่	8
7. จริยธรรมในสังคมดิจิทัล	8	17. ออกแบบ สร้างสรรค์ดิจิทัล	8		
8. สุขภาพยุคดิจิทัล	8	18. เขียนคำสั่งโปรแกรม	8		
9. ดิจิทัลคอมเมอร์ซ	8	19. ติดตั้ง ซ่อมบำรุงอุปกรณ์	8		
10. กฎหมายดิจิทัล	8	20. แก้ปัญหาทางเทคนิค	8		
		21. ปรับเปลี่ยนในยุคดิจิทัล	8		
		22. จัดการสิ่งแวดล้อมดิจิทัล	8		
		23. ใช้เทคโนโลยีสร้างสรรค์	8		
		24. การคิดเชิงคำนวณ	8		

2.2 การพัฒนาข้อสอบ ผู้วิจัยดำเนินการสร้างข้อสอบตามนิยามตัวชี้วัด โดยเริ่มจากสร้างข้อสอบรวมตามตัวชี้วัด 30 ตัว ตัวชี้วัดละ 1 ข้อ จำนวน 30 ข้อ ให้ความยากอยู่ในระดับง่าย ปานกลาง และยาก จากนั้นตรวจสอบระดับความยากจากผู้ทรงคุณวุฒิด้านเทคโนโลยีสารสนเทศ 1 คน แล้วจึงนำข้อสอบไปทดลองใช้กับนักศึกษาคณะจิตวิทยา 139 คน ใช้เวลาสอบ 50 นาที เพื่อคัดเลือกข้อสอบที่มีคุณภาพ โดยพิจารณาค่าอำนาจจำแนกอยู่ระหว่าง 0.50 – 2.50 และมีค่าความยากอยู่ระหว่าง -3 ถึง +3 (Kanjawasee, 2012) พบว่า มีข้อสอบที่มีคุณภาพผ่านเกณฑ์ทุกข้อ โดยมีค่าความยากอยู่ระหว่าง -1.720 ถึง 2.403 และค่าอำนาจจำแนกอยู่ระหว่าง 0.283 ถึง 1.082 จากนั้นผู้วิจัยสร้างข้อสอบเฉพาะเพื่อวัดสมรรถนะดิจิทัลตามโครงสร้างข้อสอบตามนิยามของตัวชี้วัด 30 ตัว ตัวชี้วัดละ 7 ข้อ จำนวน 210 ข้อ ให้ความยากอยู่ในระดับง่าย ปานกลาง และยาก ตามข้อสอบรวม ดังนั้นจะมีข้อสอบทั้งหมด จำนวน 240 ข้อ ผู้วิจัยจัดแบ่งเป็นข้อสอบ

เป็นฉบับย่อย 7 ฉบับ โดยแต่ละฉบับมีข้อสอบ จำนวน 60 ข้อ ประกอบด้วย ข้อสอบรวม จำนวน 30 ข้อ และข้อสอบเฉพาะ จำนวน 30 ข้อ โดยข้อสอบเฉพาะของแต่ละฉบับแตกต่างกัน แต่วัดในตัวชี้วัดที่เหมือนกัน ตามโครงสร้างและรูปแบบเดียวกัน จึงถือได้ว่าแบบสอบทั้ง 7 ฉบับมีลักษณะความเป็นคู่ขนานกัน ทั้งนี้ความยาวของข้อสอบรวมมีความเหมาะสม เนื่องจากจำนวนข้อสอบรวมคิดเป็นร้อยละ 20 ของข้อสอบแต่ละฉบับ (Thorndike, 1971) ทั้งนี้ข้อมูลผลคำตอบจากแบบสอบ 7 ฉบับ จะนำมารวมกัน จากการปรับเทียบแบบสอบ ด้วยวิธีการปรับเทียบพร้อมกัน ซึ่งมีข้อสอบรวม 30 ข้อ ทำหน้าที่เชื่อมโยงข้อสอบเฉพาะอีก 210 ข้อ เข้าด้วยกัน เพื่อให้พารามิเตอร์ของข้อสอบที่มาจากต่างฉบับและค่าพารามิเตอร์ความสามารถผู้สอบ ไม่เหมือนกัน จะถูกประมาณค่าพร้อมกันด้วยการวิเคราะห์เพียงครั้งเดียว (Lord, 1980)

2.3 การตรวจสอบคุณภาพ ได้แก่ การตรวจสอบความตรงตามเนื้อหาจากผู้ทรงคุณวุฒิด้าน สารสนเทศคอมพิวเตอร์ จำนวน 3 คน และด้านการวัดและประเมินผล จำนวน 2 คน พิจารณาดัชนีความ สอดคล้องระหว่างข้อสอบกับตัวชี้วัด ข้อสอบกับพฤติกรรมการเรียนรู้ และโจทย์กับคำตอบ ด้วยวิธีการหาค่า ดัชนีความสอดคล้อง (IOC) แล้วคัดเลือกข้อสอบที่มีค่า IOC ตั้งแต่ 0.50 ขึ้นไป (Rovinelli & Hambleton, 1977) จากนั้นผู้วิจัยขอรับรองจริยธรรมการวิจัยในมนุษย์ จากมหาวิทยาลัยเกษมบัณฑิต หนังสือรับรองเลขที่ R001/67P วันที่รับรอง 31 มกราคม 2567 วันหมดอายุ 31 มกราคม 2569 แล้วจึงนำแบบสอบไปทดสอบกับ นักศึกษาในสถาบันอุดมศึกษาทั้ง 8 แห่ง จำนวน 1,256 คน แต่ละแห่งจะได้แบบสอบทั้ง 7 ฉบับเหมือนกัน นักศึกษาสามารถเลือกทำฉบับใดก็ได้ เนื่องจากเป็นแบบสอบคู่ขนานตามโครงสร้าง รูปแบบ และใช้เวลาทดสอบ 50 นาที จากนั้นนำมาตรวจสอบคุณภาพข้อสอบรายข้อตามทฤษฎีการตอบสนองข้อสอบ โดยตรวจสอบความ ตรงเชิงโครงสร้างด้วยวิธีการวิเคราะห์พหุมิติ ตรวจสอบความเหมาะสมรายข้อ (item fit) ตรวจสอบความเที่ยง และตรวจสอบความตรงเชิงโครงสร้างของโมเดลสมรรถนะดิจิทัล ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน

4. การวิเคราะห์ข้อมูล

4.1 สถิติเชิงบรรยาย ได้แก่ จำนวน ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

4.2 การตรวจสอบคุณภาพของแบบสอบสมรรถนะดิจิทัล ได้แก่ 2.1 ตรวจสอบความตรง เชิงโครงสร้างด้วยวิธีการวิเคราะห์พหุมิติ เป็นการตรวจสอบความถูกต้องแม่นยำของการวัดโครงสร้างที่มี ลักษณะการวัดพหุมิติ แสดงหลักฐานด้วยวิธีการประเมินเปรียบเทียบโมเดลระหว่างโมเดลโครงสร้างแบบ พหุมิติ (multidimensional approach) กับโมเดลโครงสร้างเอกมิติ (unidimensional approach) โดยพิจารณาจากค่า Deviance Statistic (G^2) ค่า Akaike information criterion (AIC) และ ค่า Bayesian information criterion (BIC) ที่ค่าต่ำที่สุด (Yao & Schwarz, 2006) 2.2 ตรวจสอบความเหมาะสมรายข้อ (item fit) ตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ ด้วย ค่าสถิติ OUTFIT MNSQ (OUTFIT Mean Square หรือ Unweighted Mean Square) และค่าสถิติ INFIT MNSQ (INFIT Mean Square หรือ weighted fit) ซึ่งเกณฑ์ในการพิจารณาความเหมาะสมของทั้งสองค่านี้ควรมีค่าอยู่ระหว่าง 0.60 – 1.50 (Lunz et al., 1990) 2.3 ตรวจสอบความเที่ยงของแบบสอบด้วยวิธีการวิเคราะห์พหุมิติ พิจารณาค่า EAP reliability ควรมีค่ามากกว่า 0.7 ขึ้นไป (Nunnally & Bernstein, 1994)

3. การตรวจสอบความตรงเชิงโครงสร้างของโมเดล ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน โดยพิจารณาความสอดคล้องกลมกลืนของโมเดลกับข้อมูลเชิงประจักษ์ ใช้เกณฑ์ค่าไคสแควร์ (χ^2) มีค่า $p > .05$ ค่าไคสแควร์สัมพัทธ์ (χ^2/df) > 2.00 ดัชนีระดับความกลมกลืน (GFI) และ ดัชนีวัดระดับความกลมกลืนที่ปรับค่า (AGFI) $> .90$ และความคลาดเคลื่อนในการประมาณค่า (RMSEA) $< .05$ (Schumacker & Lomax, 2010)

ผลการวิจัย

1. ผลการพัฒนาโมเดลการวัดสมรรถนะดิจิทัล สำหรับคุณวุฒิระดับปริญญาตรี จากการสังเคราะห์ผลลัพธ์การเรียนรู้ตามมาตรฐานคุณวุฒิระดับอุดมศึกษา พ.ศ. 2565 ทักษะการเรียนรู้ในศตวรรษที่ 21 และ Top 10 Skills ของ World Economic Forum หลักสูตรการเข้าใจดิจิทัล พ.ศ. 2562 เป้าหมายการพัฒนาที่ยั่งยืน (SDGs) และแผนระดับชาติว่าด้วยการพัฒนาดิจิทัลเพื่อเศรษฐกิจและสังคม (พ.ศ. 2561–2580) ร่วมกับการสนทนากลุ่มจากผู้ทรงคุณวุฒิด้านเทคโนโลยีสารสนเทศจากสถาบันอุดมศึกษา จำนวน 10 คน พบว่าสมรรถนะดิจิทัลในระดับปริญญาตรี ประกอบด้วย (1) การรู้เท่าทันสื่อสารสนเทศและดิจิทัล มี 10 ตัวชี้วัด มีเนื้อหา 20 ประเด็น (2) ทักษะการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัลมี 14 ตัวชี้วัดมีเนื้อหา 42 ประเด็น และ (3) การปรับตัวสู่การเปลี่ยนแปลงดิจิทัลมี 6 ตัวชี้วัด มีเนื้อหา 18 ประเด็น แสดงดังตาราง 2

ตาราง 2 องค์ประกอบของสมรรถนะดิจิทัล ในระดับปริญญาตรี

สมรรถนะดิจิทัล	รายละเอียด	จำนวน เนื้อหา	จำนวน ตัวชี้วัด
1. การรู้เท่าทัน สื่อสารสนเทศและ ดิจิทัล	ความรู้พื้นฐานเกี่ยวกับคอมพิวเตอร์ สิทธิและความรับผิดชอบในยุค ดิจิทัล การเข้าถึงสื่อ การสื่อสาร ความปลอดภัยในยุคดิจิทัล ความเข้าใจ สื่อดิจิทัล แนวปฏิบัติและจริยธรรมในสังคมดิจิทัล สุขภาพดีในยุคดิจิทัล ดิจิทัลคอมเมิร์ซ และกฎหมายดิจิทัล	20	10
2. ทักษะการใช้ การพัฒนา และ การแก้ปัญหา	การใช้คอมพิวเตอร์และการจัดการไฟล์ข้อมูล การใช้อินเทอร์เน็ต การใช้ โปรแกรมประมวลผลคำ โปรแกรมตาราง โปรแกรมนำเสนอ การใช้งาน เพื่อความมั่นคงคอมพิวเตอร์ การพัฒนาออกแบบผลิตภัณฑ์ดิจิทัล การ เขียนคำสั่งโปรแกรมคอมพิวเตอร์ การติดตั้ง ซ่อมบำรุงอุปกรณ์ต่อพ่วง ไทร์เวอร์ การแก้ปัญหาทางเทคนิคของการใช้เทคโนโลยี การปรับเปลี่ยน ทักษะในยุคดิจิทัล การจัดการสิ่งแวดล้อมดิจิทัล การใช้เทคโนโลยีอย่าง สร้างสรรค์ และการคิดเชิงคำนวณ	42	14
3. การปรับตัวสู่ การเปลี่ยนแปลง ดิจิทัล	การยืดหยุ่นและปรับตัวต่อเทคโนโลยี การทำงานร่วมในสังคมและ วัฒนธรรมดิจิทัล การคิดริเริ่มและเรียนรู้ด้วยตนเอง การสร้างผลผลิต และการเป็นผู้ประกอบการ การเป็นผู้นำทางเทคโนโลยี และติดตาม ความก้าวหน้าเทคโนโลยีอุบัติใหม่	18	6
รวม		80	30

2. ผลการพัฒนาแบบสอบสมรรถนะดิจิทัลระดับปริญญาตรี โดยโมเดลตอบสนองข้อสอบแบบพหุมิติ

2.1 ผลการพัฒนาแบบสอบสมรรถนะดิจิทัล ตามกรอบคุณวุฒิระดับปริญญาตรี พ.ศ. 2565

แบบสอบสมรรถนะดิจิทัล ในระดับปริญญาตรี ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ เชิงยืนยัน (confirmatory MIRT model) กำหนดพฤติกรรม (มิติ) ที่ต้องการวัด คือ มิติด้านสมรรถนะดิจิทัล จากกรอบตามมาตรฐานคุณวุฒิระดับอุดมศึกษา พ.ศ. 2565 โดยวัดพฤติกรรมการเรียนรู้ ตามแนวคิดของ Anderson and Krathwohl's Taxonomy 2001 (Wilson, 2016) จำนวน 240 ข้อ แบ่งเป็น (1) การรู้เท่าทันสื่อ สารสนเทศและดิจิทัล (2) ทักษะการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัล และ (3) การปรับตัวสู่การเปลี่ยนแปลงดิจิทัล โดยส่วนใหญ่วัดพฤติกรรมในระดับรู้จำ จำนวน 178 ข้อ รองลงมา ได้แก่ ระดับเข้าใจ ระดับนำไปใช้ และระดับวิเคราะห์ (จำนวน 43 ข้อ, 13 ข้อ และ 6 ข้อตามลำดับ) ดังตาราง 3

ตาราง 3 จำนวนข้อสอบตามพฤติกรรมการเรียนรู้ในแต่ละองค์ประกอบของสมรรถนะดิจิทัล

สมรรถนะดิจิทัล	รู้จำ	เข้าใจ	นำไปใช้	วิเคราะห์	ประเมินค่า	สร้างสรรค์	รวม
1. การรู้เท่าทันสื่อสารสนเทศและดิจิทัล	52	23	3	2	-	-	80
2. ทักษะการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัล	96	12	2	2	-	-	112
3. การปรับตัวสู่การเปลี่ยนแปลงดิจิทัล	38	5	3	2	-	-	48
รวม	178	43	13	6	-	-	240

2.2 ผลการตรวจสอบคุณภาพของข้อสอบและแบบสอบสมรรถนะดิจิทัล

2.2.1 ผลการตรวจสอบความตรงตามเนื้อหา พิจารณาความตรงตามเนื้อหาด้วยวิธีการวิเคราะห์ดัชนีความสอดคล้อง (IOC) จากผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศ และด้านการวัดและประเมินผล จำนวน 5 คน พบว่า การพิจารณาดัชนีความสอดคล้องระหว่างข้อสอบกับตัวชี้วัด ระหว่างข้อสอบกับพฤติกรรมการเรียนรู้ และระหว่างโจทย์กับคำตอบ มีความสอดคล้องกันโดยมีค่า IOC อยู่ระหว่าง 0.67 – 1.00 ซึ่งเป็นไปตามเกณฑ์มากกว่า 0.5 ขึ้นไป (Rovinelli & Hambleton, 1977)

2.2.2 ผลการตรวจสอบคุณภาพข้อสอบรายข้อตามทฤษฎีการตอบสนองข้อสอบ ในการเปรียบเทียบความสอดคล้องกลมกลืนของโมเดลการตอบสนองข้อสอบแบบพหุมิติ พบว่า โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ เป็นโมเดลที่มีความสอดคล้องกลมกลืนกับข้อมูลคำตอบของแบบสอบมากที่สุด เนื่องจากมีค่าดัชนี deviance (G^2) ค่า AIC และ BIC มีค่าต่ำที่สุด (Briggs & Wilson, 2003) แสดงดังตาราง 4

ตาราง 4 ผลเปรียบเทียบโมเดลการตอบสนองข้อสอบแบบพหุมิติ

โมเดลพหุมิติ	n	AIC	BIC	Log.Lik	G^2
Rasch model	139	4975.83	5066.79	-2456.91	3542.04
2PL model	139	4961.79	5137.86	-2420.89	3470.00
3PL model	139	4931.13	5045.23	-2400.57	3429.35

2.2.3 ผลการคัดเลือกข้อสอบ จากแบบสอบสมรรถนะดิจิทัล 240 ข้อ พบว่า มีข้อสอบผ่านเกณฑ์ 211 ข้อ (ร้อยละ 87.92) แบ่งเป็น (1) ด้านการรู้เท่าทันสื่อสารสนเทศและดิจิทัล 70 ข้อ (2) ด้านทักษะการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัล 102 ข้อ และ(3) ด้านปรับตัวสู่การเปลี่ยนแปลงดิจิทัล 39 ข้อ คัดเลือกข้อสอบจากโมเดลตอบสนองข้อสอบแบบ 3 PL ใช้เกณฑ์ของ Baker (2017) ให้อำนาจจำแนก (a) > 0.5 ความยาก (b) ระหว่าง -1 ถึง 1 และการเดา (c) < 0.3 พบว่า ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.44 ความยากเฉลี่ยเท่ากับ 0.24 และการเดาเฉลี่ยเท่ากับ 0.13 พิจารณารายองค์ประกอบ พบว่า ด้านที่ 1 มีค่าอำนาจจำแนกระหว่าง 0.5 ถึง 2.44 มีความยากระหว่าง -3 ถึง 3 ด้านที่ 2 มีค่าอำนาจจำแนกระหว่าง 0.5 ถึง 5.19 มีความยากระหว่าง -2.72 ถึง 2.97 และด้านที่ 3 มีค่าอำนาจจำแนกระหว่าง 0.4 ถึง 2.48 มีความยากระหว่าง -2.04 ถึง 1.92 โดยทั้ง 3 ด้าน มีค่าการเดาไม่เกิน 0.3 แสดงดังตาราง 5

ตาราง 5 การวิเคราะห์คุณภาพข้อสอบรายข้อตามทฤษฎีการตอบสนองข้อสอบ

องค์ประกอบ	ค่าอำนาจจำแนก (a>0.5)	ค่าความยาก (-3<b<3)			ค่าการเดา (C<0.3)	ข้อสอบ ผ่านเกณฑ์
		ง่าย (-3≤b<-1)	ปานกลาง (-1≤b<1)	ยาก (1≤b<3)		
ด้านที่ 1	0.5 ถึง 2.44	-3.00 ถึง -1.04	-0.99 ถึง 0.85	1.03 ถึง 3.00	0 ถึง 0.3	70
ด้านที่ 2	0.5 ถึง 5.19	-2.72 ถึง -1.03	-0.89 ถึง 0.89	1.04 ถึง 2.97	0 ถึง 0.3	102
ด้านที่ 3	0.4 ถึง 2.48	-2.04 ถึง -1.07	-0.38 ถึง 0.94	1.06 ถึง 1.92	0 ถึง 0.3	39
	aเฉลี่ย = 1.44		bเฉลี่ย = 0.24		Cเฉลี่ย = 0.13	211

2.3 ผลการตรวจสอบความตรงเชิงโครงสร้างด้วยวิธีการวิเคราะห์พหุมิติ โดยการเปรียบเทียบโมเดล (competing models) ระหว่างโมเดลแบบเอกมิติ โมเดลเอกมิติแยกตามมิติ และโมเดลแบบพหุมิติ พบว่า โมเดลพหุมิติระหว่างข้อสอบเป็นโมเดลที่ดีที่สุด เนื่องจากมีค่า deviance index (G^2), ค่า AIC และค่า BIC น้อยกว่าโมเดลเอกมิติ และโมเดลเอกมิติแยกตามมิติ (Briggs & Wilson, 2003) เมื่อทดสอบความแตกต่างของค่า G^2 ระหว่างโมเดลเอกมิติกับพหุมิติ และเอกมิติแยกตามมิติกับพหุมิติ พบว่ามีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($\chi^2 = 64.89$, $p < .05$ และ $\chi^2 = 102.77$, $p < .05$ ตามลำดับ) ดังตาราง 6

ตาราง 6 เปรียบเทียบโมเดลเพื่อตรวจสอบความตรงเชิงโครงสร้างด้วยการวิเคราะห์พหุมิติ

โมเดลพหุมิติ	Deviance: G^2	n	จำนวนพารามิเตอร์	AIC	BIC
1.เอกมิติ	44078.87	1,256	31	44140.87	44174.94
2.เอกมิติแยกตามมิติ	42581.74	1,256	34	42649.74	42687.11
3.พหุมิติระหว่างข้อสอบ	40834.06	1,256	36	40906.06	40945.63
ผลต่างของค่า G^2 ระหว่างโมเดล 1 กับ 3, ค่า $\chi^2 = 64.89$, $p < .05$					
ผลต่างของค่า G^2 ระหว่างโมเดล 2 กับ 3, ค่า $\chi^2 = 102.77$, $p < .05$					

การตรวจสอบความเหมาะสมของข้อสอบรายข้อตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ โดยพิจารณาค่าดัชนีวัดความเหมาะสม ผลการวิเคราะห์ พบว่า มีค่าดัชนีความเหมาะสมรายข้อ OUTFIT MNSQ อยู่ระหว่าง 0.62 ถึง 1.05 และ INFIT MNSQ อยู่ระหว่าง 0.74 ถึง 1.14 ซึ่งอยู่ในเกณฑ์ที่กำหนด Lunz et al. (1990) แสดงว่าโมเดลมีความเหมาะสมกับข้อสอบรายข้อ

2.4 ผลการตรวจสอบความเที่ยงของแบบสอบสมรรถนะดิจิทัล วิเคราะห์โดยค่าสัมประสิทธิ์ความเที่ยงแบบ EAP พบว่า แบบสอบทั้งฉบับ มีค่าความเที่ยง 0.865 เป็นไปตามเกณฑ์ (Kanjawasee, 2012) แสดงดังตาราง 7

ตาราง 7 ค่าสัมประสิทธิ์ความเที่ยงของแบบสอบ

มิติที่วัด	ค่าสัมประสิทธิ์ความเที่ยงแบบ EAP
มิติด้านการรู้เท่าทันสื่อ สารสนเทศ และดิจิทัล	0.816
มิติด้านการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัล	0.827
มิติด้านการปรับตัวสู่การเปลี่ยนแปลงดิจิทัล	0.708
ทั้งฉบับ	0.865

2.5 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลการวัดสมรรถนะดิจิทัล ตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษาแห่งชาติ ในระดับปริญญาตรี ประจำปี 2565 ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) ผลการทดสอบค่าสหสัมพันธ์แบบเพียร์สัน พบว่าตัวแปรสังเกตได้ส่วนใหญ่มีความสัมพันธ์กัน อย่างมีนัยสำคัญที่ .05 และ .01 จำนวน 340 คู่ ร้อยละ 78.16 สำหรับการทดสอบสมมติฐานความเป็นเมทริกซ์เอกลักษณะจากค่าสถิติ Bartlett's test of sphericity พบว่า มีค่าเท่ากับ 5610.478 ($p < .000$) แสดงว่าเมทริกซ์สหสัมพันธ์ระหว่างตัวชี้วัดแตกต่างจากเมทริกซ์เอกลักษณะอย่างมีนัยสำคัญทางสถิติสอดคล้องกับผลการวิเคราะห์ค่าดัชนีไคเซอร์-เมเยอร์-ออลคิน (KMO) มีค่าเท่ากับ 0.790 ซึ่งปกติจะมีค่ามากกว่า 0.50 ผลการทดสอบนี้แสดงให้เห็นว่าตัวชี้วัดให้ข้อมูลมีความสัมพันธ์กันและมีความเหมาะสมที่จะนำมาวิเคราะห์องค์ประกอบได้ ผลการตรวจสอบความตรงของโมเดลพบว่า โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ ($\chi^2 = 233.011$, $p < .05$, $df = 217$, $\chi^2/df > 2$) โดยมีค่าดัชนีวัดระดับความกลมกลืน > 0.9 (GFI = 0.987, AGFI = 0.975) และค่าความคลาดเคลื่อนไม่เกิน 0.05 (RMSEA = 0.008, SRMR = 0.029) น้ำหนักองค์ประกอบมาตรฐาน (β) อยู่ระหว่าง 0.31 ถึง 0.89 ค่าน้ำหนักองค์ประกอบทุกตัวมีนัยสำคัญทางสถิติที่ระดับ .05 ($p < .05$) และมีสัดส่วนความแปรปรวนที่อธิบายได้ด้วย องค์ประกอบสมรรถนะดิจิทัลไม่เกินร้อยละ 78 แสดงดังตาราง 8

ตาราง 8 การวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลสมรรถนะดิจิทัลตามกรอบมาตรฐานคุณวุฒิ

สมรรถนะดิจิทัล	ตัวชี้วัด	น้ำหนักองค์ประกอบ			t	R ²	สปส.คะแนน องค์ประกอบ
		b	SE	β			
1. ด้านการรู้เท่าทันสื่อ สารสนเทศ และดิจิทัล (LIT)	A1	0.38	0.012	0.79	31.76	0.62	0.38
	A2	0.33	0.012	0.72	27.93	0.52	0.48
	A3	0.32	0.013	0.64	24.02	0.42	0.58
	A4	0.28	0.011	0.66	25.13	0.44	0.56
	A5	0.26	0.013	0.56	20.03	0.31	0.69
	A6	0.26	0.013	0.56	20.26	0.31	0.69
	A7	0.23	0.012	0.53	19.11	0.28	0.72
	A8	0.28	0.013	0.59	21.59	0.35	0.65
	A9	0.30	0.013	0.63	23.66	0.40	0.60
	A10	0.35	0.013	0.70	26.80	0.49	0.51
2. ด้านการใช้ การ พัฒนา และการ แก้ปัญหาด้วยเครื่องมือ ดิจิทัล (SKI)	B11	0.23	0.014	0.47	17.12	0.22	0.78
	B12	0.24	0.013	0.49	18.05	0.24	0.76
	B13	0.29	0.013	0.57	21.52	0.33	0.67
	B14	0.39	0.014	0.37	13.27	0.24	0.86
	B15	0.34	0.013	0.31	11.01	0.29	0.90
	B16	0.37	0.014	0.33	11.74	0.31	0.89
	B17	0.28	0.013	0.57	21.31	0.32	0.68
	B18	0.44	0.011	0.89	40.05	0.80	0.20
	B19	0.27	0.013	0.54	20.08	0.29	0.71
	B20	0.22	0.013	0.45	16.40	0.21	0.79
	B21	0.43	0.011	0.87	38.41	0.76	0.24
	B22	0.21	0.014	0.42	15.02	0.18	0.82
	B23	0.43	0.011	0.88	38.67	0.77	0.23
	B24	0.28	0.013	0.57	21.60	0.33	0.67
3. ด้านการปรับตัวสู่ การเปลี่ยนแปลงดิจิทัล (ADA)	C25	0.29	0.014	0.58	20.43	0.34	0.66
	C26	0.31	0.015	0.42	14.27	0.48	0.82
	C27	0.45	0.015	0.49	17.00	0.34	0.76
	C28	0.43	0.013	0.86	32.76	0.75	0.25
	C29	0.39	0.014	0.58	20.40	0.34	0.66
	C30	0.36	0.014	0.53	18.41	0.38	0.72

Chi-square = 233.01, df = 217, p-value = .217, χ^2/df = 1.074, GFI = 0.987, AGFI = 0.975, RMSEA = 0.008

3. ผลคะแนนสอบสมรรถนะดิจิทัล

ผลคะแนนสอบสมรรถนะดิจิทัลสำหรับปริญญาตรี จำนวน 1,256 คน จากแบบสอบคู่ขนาน 7 ชุด ผลการทดสอบพบว่า แบบสอบฉบับที่ 1 มีคะแนนเฉลี่ยมากที่สุด รองลงมาได้แก่ แบบสอบฉบับที่ 3, 5, 6, 2, 7 และ 4 ($M = 32.1, 32.0, 31.9, 31.5, 30.6, 29.0$ และ 28.8 ตามลำดับ) ดังตาราง 9

ตาราง 9 จำนวนผู้สอบ และค่าสถิติเบื้องต้นจากแบบสอบสมรรถนะดิจิทัล

สถาบัน	จำนวน (คน)	ฉบับที่ 1		ฉบับที่ 2		ฉบับที่ 3		ฉบับที่ 4		ฉบับที่ 5		ฉบับที่ 6		ฉบับที่ 7	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
จุฬาลงกรณ์มหาวิทยาลัย	80	38.5	4.21	38.4	5.77	29.2	6.12	31.8	10.0	36.3	5.24	33.3	5.23	40.8	5.06
มรภ.จันทระเกษม	77	31.7	5.09	34.7	3.20	30.3	7.92	25.0	7.05	34.7	6.25	27.9	7.13	32.3	6.80
มรภ.นครปฐม	82	30.5	8.74	20.8	7.73	33.8	8.67	25.9	7.93	28.4	10.1	33.8	8.67	20.8	7.73
มรภ.อุดรดิตถ์	182	31.1	8.17	27.0	6.79	33.5	10.4	25.4	9.67	29.2	7.68	33.5	10.3	27.0	6.78
ม. เกษมบัณฑิต	257	33.2	10.4	27.3	6.30	30.7	7.44	27.1	8.71	32.9	10.6	30.7	7.44	27.3	6.30
ม. รังสิต	236	32.3	7.16	33.1	9.58	32.9	9.41	24.9	7.86	31.7	6.98	28.8	8.97	27.8	8.47
ม. สยาม	227	25.0	7.56	33.2	7.85	30.0	8.41	31.1	8.46	28.2	8.56	28.1	8.20	26.1	7.78
ม.เซนต์อีส์ท์บางกอก	115	34.5	7.47	30.1	8.07	35.9	8.11	38.8	4.98	33.7	7.32	35.9	8.11	30.1	8.06
รวม	1,256	32.1	8.52	30.6	8.43	32.0	8.63	28.8	9.14	31.9	8.41	31.5	8.60	29.0	7.87

เมื่อพิจารณาคะแนนเฉลี่ยตามองค์ประกอบสมรรถนะดิจิทัล พบว่า คะแนนเฉลี่ยมากที่สุด คือ ด้านที่ 1 การรู้เท่าทันสื่อสารสนเทศและดิจิทัลเกี่ยวกับการสื่อสารยุคดิจิทัล รองลงมาได้แก่ ด้านที่ 2 ทักษะการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัล และด้านที่ 3 การปรับตัวสู่การเปลี่ยนแปลงดิจิทัล ($M = 4.49$ 3.98 และ 3.52 ตามลำดับ) โดยนักศึกษามีคะแนนเรื่อง การสื่อสารยุคดิจิทัลมากที่สุด ($M = 5.99$) สำหรับคะแนนน้อยที่สุดคือเรื่อง การคิดเชิงคำนวณ ($M = 2.40$) แสดงดังตาราง 10

ตาราง 10 สถิติเชิงบรรยายของคะแนนสอบสมรรถนะดิจิทัลสำหรับปริญญาตรี

สมรรถนะดิจิทัล ในระดับปริญญาตรี								
สมรรถนะที่ 1	M	SD	สมรรถนะที่ 2	M	SD	สมรรถนะที่ 3	M	SD
1. ความรู้พื้นฐานคอมฯ	3.72	1.46	11. การจัดการไฟล์	3.74	1.47	25. ยืดหยุ่นปรับตัว	2.68	1.28
2. สิทธิ ความรับผิดชอบ	4.97	1.17	12. ใช้อินเทอร์เน็ตได้	4.67	1.44	26. ทำงานสังคมดิจิทัล	4.02	1.48
3. เข้าถึงสื่อดิจิทัล	3.98	1.40	13. โปรแกรมประมวลผล	4.61	1.25	27. คิดริเริ่มเรียนรู้	3.89	1.46
4. การสื่อสารยุคดิจิทัล	5.99	1.32	14. โปรแกรมตาราง	4.34	1.64	28. สร้างผู้ประกอบการ	3.30	1.36
5. ปลอดภัยในยุคดิจิทัล	5.04	1.32	15. โปรแกรมนำเสนอ	5.13	1.44	29. ผู้นำทางเทคโนโลยี	4.25	1.38
6. เข้าใจสื่อดิจิทัล	4.30	1.09	16. ใช้งานความมั่นคง	3.97	1.48	30. เทคโนโลยีอุบัติใหม่	2.95	1.39
7. จริยธรรมดิจิทัล	4.59	1.02	17. พัฒนาออกแบบ	3.58	1.34			
8. สุขภาพดียุคดิจิทัล	4.79	1.15	18. เขียนโปรแกรม	2.73	1.41			
9. ดิจิทัลคอมเมิร์ซ	3.27	1.34	19. ติดตั้งบำรุงรักษา	4.46	1.46			
10. กฎหมายดิจิทัล	4.21	1.21	20. แก้ปัญหาเทคนิค	4.96	1.31			
			21. ปรับเปลี่ยนทักษะ	2.97	1.44			
			22. สิ่งแวดล้อมดิจิทัล	3.46	1.32			
			23. เทคโนโลยีสร้างสรรค์	4.66	1.33			
			24. คิดเชิงคำนวณ	2.40	1.26			
รวม	4.49	1.21	รวม	3.98	1.35	รวม	3.52	1.33

อภิปรายผล

1. การพัฒนาโมเดลสมรรถนะดิจิทัล สำหรับคุณวุฒิระดับปริญญาตรี ตามมาตรฐานคุณวุฒิระดับอุดมศึกษาแห่งชาติ พ.ศ. 2565 ประกอบด้วย 3 สมรรถนะ 30 ตัวชี้วัด ได้แก่ ด้านที่ 1 การรู้เท่าทันสื่อสารสนเทศและดิจิทัล 10 ตัวชี้วัด ได้แก่ ความรู้พื้นฐานคอมพิวเตอร์ สิทธิและความรับผิดชอบ การเข้าถึงสื่อดิจิทัล การสื่อสารยุคดิจิทัล ความปลอดภัยในยุคดิจิทัลความเข้าใจสื่อดิจิทัล จริยธรรมในสังคมดิจิทัล สุขภาพในยุคดิจิทัลดิจิทัลคอมเมิร์ซ และกฎหมายดิจิทัล ด้านที่ 2 ทักษะการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัล 14 ตัวชี้วัด ได้แก่ การใช้คอมพิวเตอร์จัดการไฟล์ การใช้อินเทอร์เน็ต การใช้โปรแกรมประมวลผลคำ การใช้โปรแกรมตารางวิเคราะห์ การใช้โปรแกรมนำเสนอ การใช้งานเพื่อความมั่นคง การพัฒนาออกแบบ ผลงาน การเขียนโปรแกรมคอมพิวเตอร์ การติดตั้ง บำรุงรักษาโปรแกรม การแก้ปัญหาทางเทคนิค การปรับเปลี่ยนทักษะยุคดิจิทัล การจัดการสิ่งแวดล้อมดิจิทัล การใช้เทคโนโลยีอย่างสร้างสรรค์ และการคิดเชิงคำนวณ และด้านที่ 3 การปรับตัวสู่การเปลี่ยนแปลงดิจิทัล 6 ตัวชี้วัด ได้แก่ การยืดหยุ่นปรับตัวต่อเทคโนโลยี การทำงานในสังคมดิจิทัล การคิดริเริ่มเรียนรู้ด้วยตนเอง การสร้างผลผลิตเป็นผู้ประกอบการ การเป็นผู้นำทางเทคโนโลยี และการติดตามเทคโนโลยีอุบัติใหม่ สอดคล้องกับเป้าหมายของการประชุมเอเปคด้านการศึกษานปี ค.ศ. 2022 กล่าวว่า สมรรถนะดิจิทัล ประกอบด้วย การเข้าใจดิจิทัล การใช้ดิจิทัล การผลิตและการติดต่อสื่อสารด้วยเครื่องมือดิจิทัล การแก้ปัญหาด้วยเครื่องมือดิจิทัล การประเมินข้อมูลสารสนเทศดิจิทัล และ การใช้เทคโนโลยีดิจิทัลอย่างมีจริยธรรม (Busabong, 2023) และสอดคล้องกับ กรอบแนวคิดสมรรถนะดิจิทัลของสำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ พ.ศ. 2562 กรอบมาตรฐานคุณวุฒิระดับอุดมศึกษา พ.ศ. 2565 ซึ่งอ้างอิงทักษะจากกรอบแนวคิดทักษะการเรียนรู้ในศตวรรษที่ 21 และ Top 10 Skills ของ World Economic Forum โดยเฉพาะทักษะการใช้ ควบคุม ดูแลเทคโนโลยี และทักษะการออกแบบเทคโนโลยี และเขียนโปรแกรม ซึ่งเป็นทักษะการใช้และพัฒนาเทคโนโลยี

2. การพัฒนาแบบสอบสมรรถนะดิจิทัลระดับปริญญาตรี วัดพฤติกรรมการเรียนรู้ตามแนวคิดของบลูมและคณะปี ค.ศ. 2001 พบว่ามีข้อคำถามครอบคลุมตั้งแต่ระดับรู้จำไปจนถึงการวิเคราะห์ ทั้งนี้ ข้อสอบส่วนใหญ่เป็นการวัดในระดับรู้จำ เนื่องจากจุดมุ่งหมายของการสอบเพื่อประเมินความรู้ก่อนสำเร็จการศึกษา (Exit-Examination) ที่มุ่งเน้นความเข้าใจในหลักการและความหมายในยุคดิจิทัล เมื่อเข้าใจแล้วจึงสามารถนำไปใช้ได้อย่างถูกต้อง สอดคล้องกับ Duangprakesa (2018) กล่าวว่า พฤติกรรมการเรียนรู้ควรเป็นไปตามลำดับขั้น โดยเริ่มจากความรู้ความเข้าใจแล้วจึงไปสู่คำถามในระดับที่สูงขึ้น สำหรับคุณภาพของแบบสอบพบว่า มีความสอดคล้องระหว่างข้อสอบกับตัวชี้วัด ข้อสอบกับพฤติกรรมการเรียนรู้ และโจทย์กับคำตอบเป็นไปตามเกณฑ์ใช้ได้ (Rovinelli & Hambleton, 1977) และมีโมเดลสอดคล้องกับแบบความเป็นพหุมิติระหว่างข้อสอบ (Between-items Multidimensionality) พิจารณาจากค่า deviance index (G^2), ค่า AIC และ ค่า BIC ที่น้อยกว่าโมเดลเอกมิติ และโมเดลเอกมิติแยกตามมิติ (Briggs & Wilson, 2003) และจากค่าความสอดคล้องระหว่างโมเดลรายข้อมีค่าสถิติ OUTFIT MNSQ และค่า INFIT MNSQ ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ (Lunz et al., 1990) แสดงว่า แบบสอบนี้สามารถนำไปใช้วัดสมรรถนะดิจิทัลในระดับปริญญาตรีได้อย่างแม่นยำ และเชื่อถือได้

3. คะแนนสอบสมรรถนะดิจิทัล สำหรับปริญญาตรี พบว่า นักศึกษามีคะแนนสอบจากแบบสอบทั้ง 7 ฉบับ ใกล้เคียงกันเนื่องจาก ผู้สอบในแต่ละฉบับมีลักษณะใกล้เคียงกันกล่าวคือ มาจากหลายสาขาวิชาเหมือนกันจึงทำให้แบบสอบแต่ละฉบับมีความแปรปรวนของคะแนนน้อย จึงทำให้คะแนนสอบใกล้เคียงกัน อีกทั้งแบบสอบยังมีความเป็นคู่ขนานกันเนื่องจาก ข้อสอบแต่ละฉบับสร้างขึ้นตามนิยามจากตัวชี้วัดเดียวกัน มีรูปแบบของการเก็บข้อมูลสำหรับการเปรียบเทียบเป็นรูปแบบ ผู้สอบกลุ่มไม่เท่าเทียมกันในหลายสาขาวิชา โดยใช้ข้อสอบร่วมเป็นตัวเชื่อมโยงอยู่ในทุกฉบับ ทั้งนี้แสดงถึงความเป็นคู่ขนานกันในโครงสร้าง รูปแบบของข้อสอบ และเวลาสอบเหมือนกัน ทั้งนี้พบว่าคะแนนมากที่สุดคือ การรู้เท่าทันสื่อสารสนเทศและดิจิทัล โดยเฉพาะเรื่องการสื่อสารในยุคดิจิทัล แสดงให้เห็นว่าเยาวชนไทยมีความรู้สามารถสื่อสารในยุคดิจิทัลได้ดี เช่น การใช้ การสืบค้น การโต้ตอบ เป็นต้น สอดคล้อง Techataweewan & Prasertsin (2016) กล่าวว่า นักศึกษาระดับปริญญาตรีมีการรู้ดิจิทัลในด้านการตระหนักรู้ ทักษะการร่วมมือ และทักษะการคิด ตามลำดับ สอดคล้องกับ Suwanroj et al. (2020) พบว่า สมรรถนะดิจิทัลที่จำเป็นสำหรับนักศึกษาระดับปริญญาตรี ได้แก่ ความรู้พื้นฐานคอมพิวเตอร์ การเข้าถึง การใช้ การผลิตและสร้างสรรค์สื่อ การจัดการ และการประเมินค่า ซึ่งอยู่ในด้านการรู้เท่าทันสื่อสารสนเทศและดิจิทัล สอดคล้องกับงานวิจัยของ Sirisak (2016) พบว่าสมรรถนะดิจิทัลของครูที่สำคัญที่สุดคือ ทักษะการใช้สื่อและเทคโนโลยีดิจิทัล สำหรับคะแนนเฉลี่ยน้อยที่สุดคือ สมรรถนะที่ 2 ด้านทักษะการใช้ การพัฒนา และการแก้ปัญหาด้วยเครื่องมือดิจิทัลคือ ทักษะการคิดเชิงคำนวณ ซึ่งกล่าวได้ว่าเยาวชนไทยมีปัญหาในเรื่องการคิดเชิงคำนวณซึ่งเป็นทักษะทางวิทยาศาสตร์ สอดคล้องกับผลการทดสอบระดับนานาชาติด้านทักษะการคิดเชิงคำนวณปี พ.ศ. 2563 ประเทศไทยอยู่ในอันดับที่ 41 ของโลก สอดคล้องกับการทดสอบนานาชาติปีบราส ประเทศไทย 2561 เด็กไทยผ่านเพียงร้อยละ 1.13 (Puechsing, 2021)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1. ผู้บริหารสถานศึกษา สามารถนำแบบสอบสมรรถนะดิจิทัลนี้ ซึ่งมีคุณภาพทั้งความตรงเชิงเนื้อหา ความตรงเชิงโครงสร้าง ความเที่ยง และเป็นข้อสอบที่มีค่าพารามิเตอร์ข้อสอบ ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าการเดาที่เหมาะสมกับการนำไปใช้ในการวัดสมรรถนะดิจิทัลของนักศึกษาในระดับปริญญาตรี เพื่อให้ให้นักศึกษาไปฝึกฝน และเตรียมความพร้อมก่อนสอบจริงเมื่อจบการศึกษา ตามเงื่อนไขนโยบายของคณะกรรมการการอุดมศึกษา

2. อาจารย์ผู้สอน สามารถนำแบบสอบสมรรถนะดิจิทัลนี้ไปสอบวัดระดับสมรรถนะดิจิทัล นักศึกษาเพื่อเป็นแนวทางในการพัฒนาผู้เรียนให้มีคุณลักษณะที่พึงประสงค์ตามแผนกลยุทธ์ด้านดิจิทัลของสถาบันการศึกษาได้ เช่น การพัฒนาหลักสูตร การจัดกิจกรรมเพื่อส่งเสริมสมรรถนะดิจิทัลที่ต้องโดดเด่น และปรับปรุงสมรรถนะที่ต้องแก้ไข

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

นักวิชาการ อาจารย์ผู้สอน และผู้ที่เกี่ยวข้อง สามารถนำแบบสอบสมรรถนะดิจิทัลนี้ไปพัฒนาเป็นคลังข้อสอบอิเล็กทรอนิกส์แบ่งเป็น ระดับง่าย ปานกลาง และยาก และสามารถต่อยอดไปสู่ระบบการ

ทดสอบแบบออนไลน์ตามทฤษฎีการวัดและประเมินผลแบบใหม่เพื่อประหยัดแรงงานในการคุมสอบ ประหยัดเวลา ประหยัดค่าใช้จ่าย เช่น การพัฒนาระบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ การพัฒนาระบบทดสอบออนไลน์เพื่อวินิจฉัยสมรรถนะดิจิทัล เป็นต้น

References

- Baker, F. B. (2017). *The basics of item response theory using R*. Springer Nature.
- Briggs, D. C., & Wilson, M. (2003). An introduction to multidimensional measurement using Rasch models. *Journal of Applied Measurement*, 4(1), 87-100.
- Brossman, B. G., & Guille, R. A. (2014). A comparison of multi-stage and liner test designs for medium-size licensure and certification examinations. *Journal of Computerized Adaptive Testing*, 2(2), 18-36.
- Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Erlbaum.
- Lunz, M. E., Wright, B. D., & Linacre, J. M. (1990). Measuring the impact of judge severity on examination scores. *Applied measurement in education*, 3(4), 331-345.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory*. McGraw-Hill.
- Rovinelli, R. J., & Hambleton, R. K. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Tijdschrift voor Onderwijsresearch*, 2(2), 49-60.
- Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). Routledge.
- Thorndike, R. L. (Ed.). (1971). *Educational measurement* (2nd ed.). American Council on Education.
- Wilson, L. O. (2016). Anderson and Krathwohl Bloom's taxonomy revised understanding the new version of Bloom's taxonomy. *The Second Principle*, 1(1), 1-8.
- Yao, L., & Schwarz, R.D. (2006). A multidimensional partial credit model with associated item and test statistics: An application to mixed-format tests. *Applied Psychological Measurement*, 30(6), 469-492.

Translate Thai References

- Busabong, C. (2023). Digital Competency: A Sustainable Learning Skill in APEC Education Goals 2022. *ECT Education and Communication Technology Journal*, 18(24), 70-85. (in Thai)

- Chianchana, C. (2009). *Multidimensional Analysis. Journal of Education Khon Kaen University*, 32(4), 13-22. (in Thai)
- Duangprakesa, N. (2018) Learning Management within the Framework of Bloom's Taxonomy Questioning Method. *Axademic Journal of Pheichaburi Rajabhat University*, 8(3), 130-138. (in Thai)
- Jarupoom, A. (2016) *A Study of Information Technology Competency for Government Readiness on Digital Economy: Case Study at Information and Communication Technology Centre-office of the Permanent Secretary, Ministry of Finacne* [Master's Independent Study]. Thammasat University. (in Thai)
- Kanjanawasee, S. (2012). *New Test Theory* (4th ed.). Chulalongkorn University Printing House. (in Thai)
- Ministry of Higher Education, Science, Research and Innovation. (2023). *Download additional higher education statistics*. <https://info.mhesi.go.th/> (in Thai)
- National Digital Economy and Society Commission. (2019). *25 Elements Digital Competency*. <https://www.dlbaseline.org/digital-competency>.(in Thai)
- Office of the Higher Education Commission. (2022). *Announcement of higher education standards on details of learning outcomes of educational qualifications, B.E. 2565*. Government Gazette, 139(Special Issue 212), 35–36. (in Thai)
- Puechsing, Y. (2021). *The development of Computational Thinking Skills Using Problem Based Learning and Social Network for Eighth Grade Students* [Master's thesis]. Mahasarakham University. (in Thai)
- Samae, N., Guadamuz, T., & Warachwarawan, W. (2021). *Digital resilience: immunity in the digital world*. Mahidol University. <https://ebookservicepro.com/showcase/DigitalResilience/> (in Thai)
- Sirisak, K. (2016) *Curriculum Research on Teacher Education Program for Developing Digital Competence Enhancement Guidance* [Master's thesis]. Chulalongkorn University. (in Thai)
- Suwanroj, T., LeeKejwattana, P., Saeung, O., & Siripan, A. (2020). The Essential digital competency for undergraduate students in Thai higher education institutions: Academic documents analysis. *Narkkhabut Paritat Journal*, 12(2), 88-106. (in Thai)
- Techataweewan, W., & Prasertsin, U. (2016). Digital Literacy Assessment of the Undergraduate Students to the Universities in Bangkok and Its Vicinity. *Journal of Information Science Research and Practice*, 34(4), 1-28. (in Thai)

การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบ อัตโนมัติสำหรับการศึกษาทางไกล

The Development of Automatic Parallel Multiple-Choice Question Generation System Using Automatic Item Generation for Distance Education

รัชกฤษ ธนพัฒน์ดล^{1*}Ratchakrit Tanapattanadol^{1*}

(Received: December 2, 2024; Revised: January 8, 2025; Accepted: January 14, 2025)

บทคัดย่อ

การวิจัยนี้มีความมุ่งหมายเพื่อ 1) สร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี 2) พัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ และ 3) ตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยแบ่งวิธีดำเนินการวิจัยออกเป็น 3 ระยะ ได้แก่ ระยะที่ 1 การสร้างและประเมินคุณภาพโมเดลข้อสอบ จำนวนทั้งสิ้น 42 โมเดล พร้อมทั้งออกแบบคลังข้อมูลสำหรับจัดเก็บค่าตัวแปรของโมเดลข้อสอบที่สร้างขึ้น เพื่อให้ทุกโมเดลข้อสอบสามารถใช้ข้อมูลที่จัดเก็บร่วมกันได้ ระยะที่ 2 การพัฒนาระบบในรูปแบบเว็บแอปพลิเคชัน ให้สามารถสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติได้ พร้อมทั้งประเมินคุณภาพของระบบที่พัฒนาขึ้น ระยะที่ 3 การตรวจสอบความเป็นคู่ขนานของแบบทดสอบ โดยนำแบบทดสอบต้นแบบและแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น ส่งให้ผู้เชี่ยวชาญประเมินความเป็นคู่ขนานของข้อสอบรายข้อ และนำไปทดลองสอบกับนักศึกษาสาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช ที่ลงทะเบียนชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 30 คน

ผลการวิจัยพบว่า 1) โมเดลข้อสอบชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยีที่สร้างขึ้นมีคุณภาพผ่านเกณฑ์คุณภาพที่กำหนดทุกโมเดล 2) ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติที่พัฒนาขึ้น สามารถสร้างข้อสอบคู่ขนานอัตโนมัติ ได้สอดคล้องกับวัตถุประสงค์ของการพัฒนาระบบ โดยมีคุณภาพของระบบในภาพรวมอยู่ใน ระดับมาก 3) ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบรายข้อโดยผู้เชี่ยวชาญ พบว่าข้อสอบทุกข้อของแบบทดสอบทั้ง 2 ฉบับ สอดคล้องกับเกณฑ์ความเป็นคู่ขนานที่กำหนด และผลการวิเคราะห์ข้อมูลจากการทดลองสอบ พบว่า แบบทดสอบต้นแบบ มีค่าความยากเฉลี่ยทั้งฉบับเท่ากับ 0.52 ค่าอำนาจจำแนกเฉลี่ยทั้งฉบับเท่ากับ 0.41 ค่าความแปรปรวน

¹ อาจารย์ สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช

¹ Lecturer, Office of Registration Records and Evaluation, Sukhothai Thammathirat Open University
ทุนวิจัยการศึกษาทางไกล มหาวิทยาลัยสุโขทัยธรรมาธิราช

Research grants for distance education, Sukhothai Thammathirat Open University

* Corresponding Author E-mail: ratchakrit.tan@stou.ac.th

ทั้งฉบับเท่ากับ 0.24 ค่าเฉลี่ยของคะแนนแบบทดสอบเท่ากับ 21.90 ค่าความเชื่อมั่นเท่ากับ 0.89 และค่าความคลาดเคลื่อนมาตรฐานของการวัดเท่ากับ 2.98 แบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์มีค่าความยากเฉลี่ยทั้งฉบับเท่ากับ 0.51 ค่าอำนาจจำแนกเฉลี่ยทั้งฉบับเท่ากับ 0.40 ค่าความแปรปรวนทั้งฉบับเท่ากับ 0.25 ค่าเฉลี่ยของคะแนนแบบทดสอบเท่ากับ 21.00 ค่าความเชื่อมั่นเท่ากับ 0.90 และค่าความคลาดเคลื่อนมาตรฐานของการวัดเท่ากับ 2.84 เมื่อทำการทดสอบความแตกต่างข้อมูลดังกล่าว ของแบบทดสอบทั้ง 2 ฉบับปรากฏว่าไม่แตกต่างกัน จากผลการตรวจสอบความเป็นคู่ขนานโดยผู้เชี่ยวชาญและผลจากการทดลองสอบจึงสรุปได้ว่า แบบทดสอบต้นแบบและแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์มีความเป็นคู่ขนานกัน

คำสำคัญ โมเดลข้อสอบ การสร้างข้อสอบอัตโนมัติ ข้อสอบคู่ขนาน การศึกษาทางไกล

Abstract

The aims of this research were to: 1) create item models and design a data warehouse for mathematics and statistics for science and technology; 2) develop an automatic parallel multiple-choice question generation system using automatic item generation; and 3) investigate parallelism between the prototype test and computer-generated test. The research methodology was divided into three phases: Phase 1 involved item model creation and evaluation to create 42 models, alongside the design of a data warehouse to store parameter values of the created item models for all item models to share stored data. Phase 2 focused on system development in web application form to create automatic parallel multiple-choice questions, together with quality evaluation of the developed system. Phase 3 consisted of parallelism investigation by having experts evaluate the parallelism of each item between the prototype test and computer-generated test, followed by experimental testing with 30 students enrolled in Mathematics and Statistics for Science and Technology at the School of Science and Technology, Sukhothai Thammathirat Open University.

The results revealed that: 1) All created item models of mathematics and statistics for science and technology passed the required criteria. 2) The automatic parallel multiple-choice question generation system successfully created parallel questions that aligned with the system's design objectives, with the overall system quality at a high level. 3) Expert evaluation confirmed that all items in both tests (the prototype and computer-generated tests) met the parallelism criteria. The experimental results showed that: (1) The prototype test had an average difficulty of 0.52, average discrimination of 0.41, variance of 0.24, mean score of 21.90, reliability of 0.89, and standard error of measurement of 2.98. (2) The computer-generated test had an average difficulty of 0.51, average discrimination of 0.40, variance of 0.25, mean

score of 21.00, reliability of 0.90, and standard error of measurement of 2.84. Statistical analysis indicated no significant differences between the two test versions. Based on both expert evaluation and experimental results, it was concluded that the prototype test and computer-generated test were parallel.

Keywords: item model, automatic item generation, parallel test, distance education

บทนำ

การสอบเป็นวิธีการวัดและประเมินผลการศึกษาที่สามารถใช้ได้กับผู้เรียนในทุกระดับ และเป็นวิธีการวัดและประเมินผลที่ใช้กันอย่างแพร่หลายในสถาบันการศึกษา โดยเฉพาะในระบบการศึกษาทางไกล อย่างมหาวิทยาลัยสุโขทัยธรรมมาธิราช ซึ่งมีรูปแบบการเรียนการสอนผ่านทางไกลเป็นรูปแบบหลัก มีการเผชิญหน้าระหว่างผู้เรียนกับผู้สอนเป็นส่วนน้อย วิธีการวัดและประเมินผลเพื่อตัดสินผลการเรียนรู้ที่สอดคล้องกับรูปแบบการเรียนการสอนดังกล่าว จึงเป็นการสอบโดยการจัดการแบบรวมศูนย์เพื่อให้เป็นมาตรฐานเดียวกันทั้งมหาวิทยาลัย ซึ่งเครื่องมือหลักที่ใช้ในการวัดและประเมินผลคือ แบบทดสอบ ประกอบด้วยข้อสอบแบบเลือกตอบ (5 ตัวเลือก) หรือข้อสอบแบบเลือกตอบผสมแบบเขียนตอบ หรือเป็นข้อสอบแบบเขียนตอบล้วน ทั้งนี้ขึ้นอยู่กับข้อกำหนดของแต่ละสาขา และระดับการศึกษาของผู้สอบ

ระบบคลังข้อสอบมาตรฐานที่ใช้สำหรับการวัดและประเมินผลการเรียนรู้ในระบบการศึกษาทางไกลของมหาวิทยาลัยสุโขทัยธรรมมาธิราช หรือเรียกสั้น ๆ ว่าระบบคลังข้อสอบ นับเป็นกลไกสำคัญในการทำหน้าที่เก็บรวบรวมข้อสอบสำหรับสร้างเป็นแบบทดสอบดังกล่าว เพื่อใช้ในการสอบในรูปแบบต่างๆ ของมหาวิทยาลัย นอกจากนี้ระบบคลังข้อสอบยังมีวัตถุประสงค์เพื่อพัฒนาคุณภาพแบบทดสอบ ลดภาระในการออกข้อสอบใหม่ รวมถึงการลดค่าใช้จ่ายในการสร้างข้อสอบใหม่อีกด้วย นอกจากบทบาทสำคัญในการสร้าง และพัฒนาคุณภาพแบบทดสอบดังกล่าวแล้ว ระบบคลังข้อสอบยังมีข้อควรระวังที่สำคัญประการหนึ่ง ซึ่งอาจกระทบต่อประสิทธิภาพของการวัดประเมินผลการศึกษาได้ คือถ้าหากระบบคลังข้อสอบมีจำนวนข้อสอบไม่มากพอ หรือมีจำนวนไม่สอดคล้องกับความต้องการใช้งานในปัจจุบัน หรืออนาคต อาจส่งผลให้ข้อสอบในคลังข้อสอบมีโอกาสถูกนำมาใช้งานซ้ำบ่อยครั้งขึ้น และการที่ข้อสอบถูกนำไปใช้ในการทดสอบบ่อยครั้งเกินไป ย่อมทำให้เกิดผลกระทบต่อความปลอดภัยของการทดสอบ กล่าวคือเมื่อมีข้อสอบที่ถูกนำไปใช้ในการทดสอบบ่อยครั้งเกินไป จนรู้กันในกลุ่มผู้สอบ เป็นผลทำให้ผู้สอบในรุ่นถัดมาสามารถคาดการณ์ล่วงหน้าได้ว่าตนเองจะได้ข้อสอบข้อใด ทำให้สามารถตอบข้อสอบข้อดังกล่าวได้โดยไม่ต้องใช้ความสามารถของตนเอง ส่งผลให้คะแนนสอบมีความคลาดเคลื่อน ผู้สอบสามารถทำคะแนนได้มากเกินไปความสามารถของตนเอง กระทั่งคุณภาพของการวัดและประเมินผลการศึกษา (Premthongsuk et al., 2017) ดังนั้นการมีข้อสอบที่มีคุณภาพจำนวนมากในระบบคลังข้อสอบ หรือการสร้างข้อสอบใหม่ที่มีคุณภาพเทียบเท่ากับข้อสอบที่อยู่ในระบบคลังข้อสอบได้จำนวนมาก ในระยะเวลารวดเร็ว และทันเวลาต่อความต้องการใช้งาน จึงเป็นสิ่งสำคัญและจำเป็น ในขณะที่การสร้างข้อสอบเพื่อเพิ่มจำนวนข้อสอบลงในระบบคลังข้อสอบ โดยผ่านขั้นตอนการจัดทำคลังข้อสอบที่ใช้อยู่ในปัจจุบันของ

มหาวิทยาลัยนั้น ไม่เอื้อให้สามารถสร้างข้อสอบจำนวนมาก ๆ ในระยะเวลาที่รวดเร็วได้ เนื่องจากต้องใช้ผู้เชี่ยวชาญด้านเนื้อหาในการออกข้อสอบ รวมถึงการดำเนินการในแต่ละขั้นตอนมีรายละเอียดที่ต้องใช้ระยะเวลาค่อนข้างนานในการดำเนินการ รวมทั้งมีผู้เกี่ยวข้องหลายฝ่าย จากข้อจำกัดดังกล่าวย่อมเป็นอุปสรรคต่อการสร้างข้อสอบจำนวนมากที่พร้อมใช้งานได้ในระยะเวลาอันสั้น จากปัญหาข้างต้นหากมีวิธีที่สามารถสร้างข้อสอบแบบอัตโนมัติเพื่อบรรจุไว้ในคลังข้อสอบได้จำนวนมาก รวดเร็ว และมีความเป็นคู่ขนานกับข้อสอบที่อยู่ในระบบคลังข้อสอบ หรือข้อสอบต้นฉบับ ย่อมทำให้ระบบคลังข้อสอบสามารถสร้างแบบทดสอบที่วัดผลการเรียนรู้เดียวกันได้หลากหลาย ส่งผลให้การวัดและประเมินผลการศึกษาของมหาวิทยาลัยมีประสิทธิภาพมากยิ่งขึ้น นอกจากนี้ยังช่วยลดภาระหรือเวลาในการออกข้อสอบใหม่ของผู้เชี่ยวชาญด้านเนื้อหา และลดค่าใช้จ่ายในการสร้างข้อสอบใหม่อีกด้วย

ปัจจุบันมีการพัฒนาเทคโนโลยีคอมพิวเตอร์มาช่วยในการสร้างข้อสอบที่เรียกว่าการสร้างข้อสอบอัตโนมัติ (Automatic Item Generation : AIG) ซึ่งการสร้างข้อสอบอัตโนมัติดังกล่าว เป็นกระบวนการสร้างข้อสอบภายใต้องค์ประกอบของโมเดลข้อสอบที่สร้างขึ้นตามหลักการของ Gierl et al. (2008) ที่ช่วยให้สามารถสร้างข้อสอบได้อย่างรวดเร็ว เมื่อเปรียบเทียบกับวิธีการอื่น (Gierl & Haladyna, 2013) รวมทั้งมีงานวิจัยอีกหลายเรื่องที่เป็นต้นแบบ ในการนำหลักการสร้างข้อสอบอัตโนมัติดังกล่าว ไปใช้ดำเนินการวิจัย เช่น งานวิจัยของ Chantharamaha & Chadcham (2018); Gierl et al. (2012); Gierl & Lai (2012; 2013); Pradujprom & Pantong (2019); Raksombat et al. (2022) เป็นต้น ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะพัฒนาระบบสร้างข้อสอบคู่ขนานอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ ภายใต้หลักการการสร้างโมเดลข้อสอบของ Gierl et al. (2008) เพื่อเป็นอีกหนึ่งแนวทางที่สามารถสร้างข้อสอบคู่ขนานได้จำนวนมาก รวดเร็ว และข้อสอบมีคุณภาพเท่าเทียมกับข้อสอบต้นแบบ (Chantharamaha & Chadcham, 2018) โดยความเท่าเทียมดังกล่าวจะอยู่ในบริบทของความเป็นคู่ขนาน คือข้อสอบมีเนื้อหาเดียวกัน หรือวัดวัดคุณสมบัติและระดับพฤติกรรมเดียวกัน มีความยากง่ายใกล้เคียงกัน เพื่อให้สามารถนำข้อสอบของแต่ละวัดคุณสมบัติการเรียนรู้ของชุดวิชาที่สร้างจากระบบที่พัฒนาขึ้นไปใช้ในการทดสอบรูปแบบต่าง ๆ หรือจัดเก็บเพิ่มเติมไว้ในระบบคลังข้อสอบได้ทันที โดยการพัฒนาระบบสร้างข้อสอบคู่ขนานอัตโนมัติดังกล่าว ผู้วิจัยได้มุ่งเน้นพัฒนาระบบคอมพิวเตอร์ให้มีความสามารถในการสร้างข้อสอบคู่ขนานแบบเลือกตอบโดยเฉพาะ เนื่องจากข้อสอบแบบเลือกตอบดังกล่าว ได้ถูกใช้เป็นเครื่องมือหลักสำหรับวัดและประเมินผลการศึกษาในระบบการศึกษาทางไกลของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ในระดับปริญญาตรีซึ่งเป็นนักศึกษาส่วนใหญ่ของมหาวิทยาลัย

สำหรับการวิจัยเพื่อพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติที่นำเสนอ นั้น จะดำเนินการภายใต้ชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี ซึ่งเป็นชุดวิชาที่ใช้ในการเรียนการสอนในระบบการศึกษาทางไกลของมหาวิทยาลัยสุโขทัยธรรมมาธิราช เนื่องจากเป็นชุดวิชาที่มีเนื้อหาเอื้อต่อการนำข้อสอบมาสร้างโมเดลข้อสอบแล้วคาดว่ามีแนวโน้มที่จะได้โมเดลข้อสอบหลากหลายรูปแบบ มากกว่าการดำเนินการภายใต้ชุดวิชาอื่น ๆ ที่อาจมีเนื้อหาของข้อสอบประกอบด้วยส่วนที่แปรเปลี่ยนได้น้อยกว่า ส่งผลให้ได้โมเดลข้อสอบรูปแบบซ้ำ ๆ ไม่หลากหลาย ดังนั้นการมีรูปแบบของโมเดลข้อสอบที่หลากหลายดังกล่าว ย่อมสามารถใช้เป็นแนวทางในการสร้างโมเดลข้อสอบของชุด

วิชาอื่น ๆ ที่เนื้อหาชุดวิชามีบริบทใกล้เคียงกับกับชุดวิชาคณิตศาสตร์ และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยีดังกล่าวต่อไปในอนาคตได้ นอกจากนี้ข้อสอบที่นำมาใช้เป็นข้อสอบต้นแบบ เพื่อนำไปใช้ในกระบวนการสร้างข้อสอบอัตโนมัติตามวิธีที่นำเสนอดังกล่าว ผู้วิจัยจะดำเนินการสร้างข้อสอบต้นแบบตามหลักการสร้างและตรวจสอบคุณภาพข้อสอบ โดยเนื้อหาข้อสอบต้นแบบที่สร้างเป็นของชุดวิชาดังข้างต้น แทนการใช้ข้อสอบจริงที่อยู่ในคลังข้อสอบ ด้วยเหตุผลเรื่องความปลอดภัยของข้อสอบในระหว่างการทำวิจัย และการเผยแพร่เนื้อหาข้อสอบในรายงานการวิจัย เมื่อดำเนินการวิจัยเสร็จสิ้น ระบบสร้างข้อสอบที่พัฒนาขึ้นจากงานวิจัยนี้ ยังสามารถขยายผลให้กับข้อสอบจริงที่จัดเก็บอยู่ในระบบคลังข้อสอบมาตรฐานของมหาวิทยาลัย เพื่อใช้สำหรับวัดและประเมินผลการศึกษาในระบบการศึกษาทางไกลได้ต่อไป นอกจากนี้สถาบันการศึกษาอื่น ๆ ยังสามารถนำวิธีดำเนินการวิจัย ไปประยุกต์ใช้เป็นแนวทางในการพัฒนาระบบสร้างข้อสอบคู่ขนานอัตโนมัติสำหรับใช้วัดและประเมินผลการศึกษาในสถาบันของตนเองได้เช่นเดียวกัน

ความมุ่งหมายของการวิจัย

1. เพื่อสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี
2. เพื่อพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ
3. เพื่อตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์

วิธีดำเนินการวิจัย

การวิจัยนี้แบ่งวิธีการดำเนินการวิจัยออกเป็น 3 ระยะ ดังนี้

ระยะที่ 1 การสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี

การวิจัยในระยะนี้มีวัตถุประสงค์เพื่อสร้างโมเดลข้อสอบ และออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี มีวิธีดำเนินการวิจัยที่เกี่ยวข้องดังนี้

1. การสร้างข้อสอบต้นแบบเพื่อนำไปสร้างโมเดลข้อสอบมีขั้นตอนดังนี้

1) กำหนดวัตถุประสงค์การเรียนรู้ ระดับพฤติกรรม และเนื้อหา สำหรับใช้สร้างข้อสอบแบบเลือกตอบ 5 ตัวเลือก ซึ่งเนื้อหาดังกล่าวเป็นส่วนหนึ่งของเนื้อหาชุดวิชา 96102 คณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 6 เรื่องหลัก ได้แก่ (1) อนุพันธ์ (2) เซต (3) สมการ (4) อสมการ (5) เรขาคณิตวิเคราะห์ และ (6) เมทริกซ์และดีเทอร์มิแนนต์

2) เขียนลักษณะเฉพาะของข้อสอบ (Item Specification) ของเนื้อหาข้อสอบทุกเรื่องที่ถูกกำหนดไว้สำหรับนำมาใช้สร้างข้อสอบ เรื่องละ 7 ข้อ รวมทั้งสิ้น 42 ข้อ ซึ่งรายละเอียดของลักษณะเฉพาะของข้อสอบประกอบด้วย ชื่อเนื้อหา วัตถุประสงค์ ระดับพฤติกรรม ลักษณะคำถาม ลักษณะตัวถูก และตัวอย่างข้อสอบ โดยไม่เขียนลักษณะตัวถูก

3) สร้างข้อสอบเฉพาะคำถาม และตัวถูก ตามลักษณะเฉพาะของข้อสอบที่กำหนด พร้อมทั้งนำข้อสอบเฉพาะคำถาม และคำตอบ ที่สร้างขึ้นไปสัมภาษณ์ผู้สอนวิชาคณิตศาสตร์ จำนวน 10 คน จากหลากหลายสถาบัน โดยสัมภาษณ์ในประเด็นเกี่ยวกับความเข้าใจคลาดเคลื่อน หรือความผิดพลาดของผู้เรียนในการทำข้อสอบจากประสบการณ์ที่ผ่านมาของผู้สอน เพื่อนำไปใช้เป็นข้อมูลส่วนหนึ่งในการสร้างตัวถูก

4) เขียนลักษณะเฉพาะของข้อสอบ (item specification) ของเนื้อหาข้อสอบแต่ละเรื่องที่ถูกกำหนดไว้สำหรับนำมาใช้สร้างข้อสอบให้สมบูรณ์ ซึ่งประกอบด้วย ชื่อเนื้อหา วัตถุประสงค์ ระดับพฤติกรรม ลักษณะคำถาม ลักษณะตัวถูก ลักษณะตัวถูก และตัวอย่างข้อสอบ พร้อมทั้งสร้างข้อสอบตามลักษณะเฉพาะของข้อสอบ (item specification) จำนวน 60 ข้อ และนำข้อสอบที่สร้างขึ้นไปสัมภาษณ์ผู้สอนวิชาคณิตศาสตร์ จำนวน 10 คน อีกครั้ง ในประเด็นความเหมาะสมของ คำถาม ตัวถูก และตัวถูกที่สร้างขึ้น พร้อมทั้งปรับข้อสอบตามข้อเสนอแนะจากการสัมภาษณ์

5) ตรวจสอบคุณภาพด้านความตรงเชิงเนื้อหา (content validity) ของข้อสอบที่สร้างขึ้น โดยผู้เชี่ยวชาญทางด้านคณิตศาสตร์ จำนวน 3 คน เพื่อพิจารณาความสอดคล้องของข้อสอบกับวัตถุประสงค์การเรียนรู้ และความสอดคล้องของข้อสอบกับลักษณะเฉพาะของข้อสอบ จากนั้นนำไปทดลองใช้กับกลุ่มตัวอย่าง ได้แก่ นักศึกษามหาวิทยาลัยสุโขทัยธรรมาธิราช ที่ลงทะเบียนชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 30 คน และทำการวิเคราะห์คุณภาพข้อสอบรายข้อ และรายฉบับ พร้อมทั้งคัดเลือกข้อสอบที่มีคุณภาพจำนวน 42 ข้อ เพื่อนำไปใช้เป็นข้อสอบต้นแบบสำหรับสร้างโมเดลข้อสอบต่อไป

2. การสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยีมีขั้นตอนการดำเนินการดังนี้

1) นำข้อสอบต้นแบบและลักษณะเฉพาะของข้อสอบที่สร้างขึ้น จำนวน 42 ข้อ ไปใช้เป็นข้อมูลสำหรับสร้างโมเดลข้อสอบ โดยเลือกรูปแบบโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) ที่ต้องการสร้าง โดยวิเคราะห์เนื้อหาของข้อสอบต้นแบบ และลักษณะเฉพาะของข้อสอบแต่ละข้อว่า มีความสอดคล้องกับประเภทของโมเดลข้อสอบโมเดลใดใน 10 โมเดล (โมเดลที่ 1 – โมเดลที่ 10) ดังตาราง 1 จากนั้นจึงดำเนินการสร้างโมเดลข้อสอบตามโครงสร้างของโมเดลนั้น ๆ

ตาราง 1 การจัดหมวดหมู่โมเดลข้อสอบ (item model taxonomy)

ตัวเลือก (Options)	โจทย์ (Stem)			
	อิสระ (independent)	ไม่อิสระ (dependent)	ผสม (mixed)	คงที่ (fixed)
รูปแบบการสุ่ม (randomly selected)	โมเดลที่ 1	โมเดลที่ 4	โมเดลที่ 7	โมเดลที่ 10
รูปแบบจำกัด (constrained)	โมเดลที่ 2	โมเดลที่ 5	โมเดลที่ 8	N/A
รูปแบบคงที่ (fixed)	โมเดลที่ 3	โมเดลที่ 6	โมเดลที่ 9	N/A

จากตาราง 1 โจทย์ของโมเดลข้อสอบ แบ่งออกเป็น 4 รูปแบบ ได้แก่ (1) องค์ประกอบที่แปรเปลี่ยนได้อย่างอิสระ (Independent Element) เมื่อองค์ประกอบใดในโจทย์มีการเปลี่ยนแปลงจะไม่ส่งผลกระทบต่อองค์ประกอบอื่น ๆ (2) องค์ประกอบที่แปรเปลี่ยนได้อย่างไม่อิสระ (dependent element) เมื่อองค์ประกอบใด

องค์ประกอบหนึ่งในโจทย์มีการเปลี่ยนแปลงจะส่งผลให้องค์ประกอบอื่นๆ ที่สัมพันธ์กันเปลี่ยนแปลงตามไปด้วย (3) การผสมกันระหว่างองค์ประกอบที่แปรเปลี่ยนได้อย่างอิสระ และแปรเปลี่ยนได้อย่างไม่อิสระ (mixed independent/dependent element) โจทย์ประกอบไปด้วยองค์ประกอบที่แปรเปลี่ยนได้อย่างอิสระ ซึ่งเมื่อองค์ประกอบใดในโจทย์มีการเปลี่ยนแปลงจะไม่ส่งผลกระทบต่อองค์ประกอบอื่น ๆ แต่ในขณะเดียวกัน โจทย์ดังกล่าวก็ยังมีองค์ประกอบที่แปรเปลี่ยนได้อย่างไม่อิสระ เมื่อองค์ประกอบใดขององค์ประกอบหนึ่งในโจทย์ มีการเปลี่ยนแปลงจะส่งผลให้องค์ประกอบอื่นๆ ที่สัมพันธ์กันเปลี่ยนแปลงตามไปด้วย และ (4) องค์ประกอบคงที่ (fixed element) ลักษณะของโจทย์คงที่ไม่มีการเปลี่ยนแปลง ในขณะที่ตัวเลือกของโมเดลข้อสอบ แบ่งออกเป็น 3 รูปแบบ ได้แก่ (1) ตัวเลือกที่ได้จากการสุ่ม (randomly selected options) คือรูปแบบที่หาคำตอบที่ถูกต้องและตัวลวงถูกเลือกแบบสุ่มจากขอบเขตข้อมูลที่มีเนื้อหาสอดคล้องกัน (2) ตัวเลือกแบบจำกัด (constrained options) คือรูปแบบที่ทั้งคำตอบที่ถูกต้องและตัวลวงถูกสร้างตามข้อกำหนด เช่น สูตร หรือรูปแบบการคำนวณ เป็นต้น และ (3) ตัวเลือกแบบคงที่ (fixed options) คือรูปแบบที่ทั้งคำตอบที่ถูกต้องและตัวลวงไม่มีการเปลี่ยนแปลง

2) สร้างโมเดลข้อสอบโดยการเชื่อมโยงเนื้อหาข้อสอบต้นแบบเข้ากับองค์ประกอบทั้ง 5 ส่วนของโมเดลข้อสอบ แต่ละประเภทที่เลือกไว้ พร้อมทั้งกำหนดรายละเอียดเนื้อหาของโมเดลข้อสอบทั้ง 5 ส่วน ดังนี้ (1) โจทย์ (stem): นำเนื้อหาโจทย์ของข้อสอบต้นแบบมาวิเคราะห์เพื่อกำหนดส่วนคงที่ และส่วนที่สามารถเปลี่ยนค่าได้ (2) ส่วนประกอบ (elements): กำหนดรายละเอียดเนื้อหาโจทย์ของโมเดลข้อสอบ ในส่วนที่สามารถเปลี่ยนแปลงค่าได้ 3 ส่วน ได้แก่ 1. ชื่อตัวแปร และชนิดของตัวแปร 2. ค่าตัวแปร ขอบเขตของค่าตัวแปร และค่ายกเว้นตัวแปร 3. ค่าเงื่อนไขระหว่างตัวแปร (3) ตัวเลือก (options): กำหนดรายละเอียดเนื้อหาตัวเลือกทั้ง 5 ตัวเลือก ของแต่ละโมเดล ซึ่งตัวเลือกดังกล่าวจะมีเนื้อหาในรูปแบบใดขึ้นอยู่กับเนื้อหาของข้อสอบต้นแบบ และประเภทของโมเดลข้อสอบที่สอดคล้องกับข้อสอบต้นแบบที่เลือกไว้ (4) ข้อมูลเสริม (auxiliary information): ข้อมูลเสริมสามารถกำหนดได้ที่ตำแหน่งโจทย์และตัวเลือกของโมเดลข้อสอบ ซึ่งแต่ละโมเดลข้อสอบจะมีข้อมูลเสริมหรือไม่ หรือมีในรูปแบบใดขึ้นอยู่กับเนื้อหาของข้อสอบต้นแบบ โดยข้อมูลเสริมดังกล่าว ได้แก่ ภาพ และตาราง และ (5) เฉลย (key): การกำหนดคำตอบที่ถูกต้องของข้อสอบข้อนั้น ๆ ตามที่โจทย์กำหนด

3) ตรวจสอบคุณภาพโมเดลข้อสอบที่สร้างขึ้น โดยดำเนินการดังนี้ (1) สร้างแบบประเมินคุณภาพโมเดลข้อสอบ โดยประยุกต์ใช้ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ (Index of Item Objective Congruence : IOC) มาเป็นเกณฑ์การประเมินคุณภาพโมเดลข้อสอบของแต่ละรายการประเมิน โดยคะแนน -1 หมายถึง ไม่สอดคล้อง คะแนน 0 หมายถึง ไม่แน่ใจ และคะแนน 1 หมายถึง สอดคล้อง พร้อมทั้งนำแบบประเมินคุณภาพโมเดลข้อสอบ ส่งให้ผู้เชี่ยวชาญทางด้านการวัดและประเมินผลจำนวน 3 คน ตรวจสอบคุณภาพของแบบประเมินก่อนนำไปใช้ประเมินจริง (2) ประเมินคุณภาพของโมเดลข้อสอบ โดยให้ผู้เชี่ยวชาญซึ่งประกอบด้วย 1) ประธานชุดวิชาหรืออาจารย์ผู้สอนที่เป็นผู้แทนคณะกรรมการบริหารชุดวิชา คณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 1 คน 2) ผู้เชี่ยวชาญทางด้านคณิตศาสตร์ จำนวน 1 คน และ 3) ผู้เชี่ยวชาญทางด้านการวัดและประเมินผล ซึ่งนอกจากจะมีคุณวุฒิทางด้านการวัดและประเมินผลแล้ว ยังต้องมีคุณวุฒิด้านคณิตศาสตร์ด้วย จำนวน 1 คน รวมทั้งหมด 3 คน ซึ่งเป็นจำนวน

ผู้เชี่ยวชาญที่เหมาะสมในการประเมินความสอดคล้องตามคำแนะนำของ Srisa-ard (1995) ทำการตรวจสอบคุณภาพของโมเดลข้อสอบพร้อมทั้งทำแบบประเมินคุณภาพโมเดลข้อสอบ ซึ่งรายการประเมินคุณภาพโมเดลข้อสอบ มีประเด็นพิจารณาทั้งสิ้น 5 ประเด็น ได้แก่ ประเด็นที่ 1 ความสอดคล้องระหว่างประเภทของโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) ที่กำหนดกับโมเดลข้อสอบที่สร้างขึ้น ประเด็นที่ 2 ความสอดคล้องระหว่างเนื้อหาของข้อสอบต้นแบบกับเนื้อหาของโมเดลข้อสอบที่สร้างขึ้น ประเด็นที่ 3 ความสอดคล้องระหว่างวัตถุประสงค์การเรียนรู้ที่นำมาสร้างข้อสอบต้นแบบกับโมเดลข้อสอบที่สร้างขึ้น ประเด็นที่ 4 ความสอดคล้องระหว่างระดับพฤติกรรมการเรียนรู้ที่นำมาสร้างข้อสอบต้นแบบกับโมเดลข้อสอบที่สร้างขึ้น และประเด็นที่ 5 ความสอดคล้องระหว่างความยากของข้อสอบต้นแบบกับความยากของโมเดลข้อสอบที่สร้างขึ้น และ (3) วิเคราะห์ผลการประเมินคุณภาพโมเดลข้อสอบจากผู้เชี่ยวชาญ โดยการคำนวณค่าดัชนีความสอดคล้อง ซึ่งโมเดลข้อสอบทุกข้อหรือทุกโมเดล ต้องมีค่าดัชนีความสอดคล้องของทุกรายการประเมินเป็นไปตามเกณฑ์ที่กำหนดคือ ตั้งแต่ 0.5 ขึ้นไป พร้อมทั้งสรุปผลการประเมินคุณภาพโมเดลข้อสอบ

3. การออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี

เป็นการออกแบบพื้นที่จัดเก็บข้อมูลกลาง สำหรับนำไปสร้างข้อสอบคู่ขนานของชุดวิชาดังกล่าวผ่านโครงสร้างของโมเดลข้อสอบ ซึ่งคลังข้อมูลจะถูกออกแบบให้เป็นองค์ประกอบหนึ่งของการสร้างโมเดลข้อสอบ ผ่านระบบคอมพิวเตอร์ที่พัฒนาขึ้น โดยคลังข้อมูลจะเก็บข้อมูลในรูปแบบของค่าตัวแปรชนิดตัวอักษร (character) หรือตัวแปรชนิดข้อความ (string) ที่นำไปใช้สร้างโมเดลข้อสอบ ไว้ร่วมกันที่ส่วนกลาง เพื่อบริการให้ทุกโมเดลข้อสอบสามารถใช้ข้อมูลที่ถูกจัดเก็บไว้ในคลังข้อมูลดังกล่าวร่วมกันได้ โดยที่ผู้สร้างโมเดลข้อสอบไม่ต้องเสียเวลาในการสร้างค่าตัวแปรของข้อมูลใหม่ทุกครั้งให้กับโมเดลข้อสอบที่กำลังสร้างขึ้น หากตัวแปรใดถูกสร้างไว้ในคลังข้อมูลแล้ว ผู้ใช้งานก็สามารถนำไปใช้เป็นองค์ประกอบของโมเดลข้อสอบนั้น ๆ ได้ทันที เพื่อให้คลังข้อมูลมีคุณสมบัติดังกล่าวข้างต้น การออกแบบคลังข้อมูลสามารถดำเนินการได้ดังนี้ 1) ศึกษาองค์ประกอบข้อมูลและรูปแบบการจัดเก็บข้อมูลในคลังข้อมูลข้อสอบ 2) ศึกษาวิธีการเชื่อมโยงระหว่างข้อมูลในคลังข้อมูลกับข้อมูลตัวแปรสำหรับใช้สร้างโมเดลข้อสอบ 3) ออกแบบองค์ประกอบของคลังข้อมูล 4) ออกแบบรูปแบบการจัดเก็บ การลบ การแก้ไข ข้อมูลในคลังข้อมูล 5) ออกแบบขั้นตอนการสร้างคลังข้อมูลผ่านระบบสร้างข้อสอบคู่ขนานที่พัฒนาขึ้น และ 6) ออกแบบขั้นตอนการนำข้อมูลจากคลังข้อมูลไปใช้งาน อย่างไรก็ตาม การออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยีดังกล่าวเป็นเพียงการออกแบบเพื่อบริการการสร้างโมเดลข้อสอบในระบบสร้างข้อสอบคู่ขนานอัตโนมัติให้มีความสะดวกในการสร้างโมเดลข้อสอบมากยิ่งขึ้นเท่านั้น ซึ่งผู้ใช้งานอาจจะใช้งานข้อมูลที่มีในคลังข้อมูลหรือไม่ก็ได้ ทั้งนี้ขึ้นอยู่กับความประสงค์ และบริบทของเนื้อหาข้อสอบที่นำมาสร้างโมเดลข้อสอบนั้น ๆ ด้วย

ระยะที่ 2 การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ

การวิจัยในระยะนี้มีวัตถุประสงค์เพื่อพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ ซึ่งระบบดังกล่าวจะถูกพัฒนาขึ้นสำหรับใช้สร้างข้อสอบคู่ขนานแบบเลือกตอบ

อัตโนมัติ ภายใต้หลักการสร้างข้อสอบอัตโนมัติและการสร้างโมเดลข้อสอบของ Gierl et al. (2008) ทั้ง 10 ประเภท และนำโมเดลข้อสอบที่สร้างในการวิจัยระยะที่ 1 จำนวน 42 ข้อ รวมทั้งคลังข้อมูลข้อสอบที่ได้ ออกแบบไว้ มาเป็นข้อมูลในการพัฒนาและทดสอบระบบไปด้วย ซึ่งมีวิธีดำเนินการวิจัยดังนี้

1. วิเคราะห์และออกแบบระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ

การวิเคราะห์และออกแบบระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ มีขั้นตอนการดำเนินการดังนี้ 1) วิเคราะห์และออกแบบระบบ ด้วยการศึกษางานวิจัย และเทคโนโลยีที่เกี่ยวข้อง เพื่อให้ระบบที่พัฒนาขึ้นสามารถรองรับการทำงานหลักได้แก่ (1) สามารถสร้างโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) ได้ครบทั้ง 10 ประเภท (2) สามารถสร้างคลังข้อมูลข้อสอบเพื่อใช้จัดเก็บข้อมูลค่าตัวแปรต่าง ๆ ของโมเดลข้อสอบได้ และ (3) สามารถสร้างข้อสอบคู่ขนานอัตโนมัติจากโมเดลข้อสอบและเงื่อนไขที่กำหนดได้ 2) ออกแบบระบบฐานข้อมูล เพื่อรองรับหรือจัดเก็บข้อมูลที่เกี่ยวข้องกับระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ และ 3) ออกแบบส่วนติดต่อกับผู้ใช้งาน เพื่อให้ผู้ใช้งานสามารถโต้ตอบหรือสื่อสารกับระบบที่พัฒนาขึ้นได้อย่างสะดวก

2. พัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ

การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ มีขั้นตอนการดำเนินการที่เกี่ยวข้องดังนี้ 1) พัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติในรูปแบบเว็บแอปพลิเคชัน (web application) ให้สอดคล้องกับผลการวิเคราะห์และออกแบบระบบ 2) ทดสอบและแก้ไขข้อผิดพลาดทุกฟังก์ชันการทำงานของระบบที่พัฒนาขึ้นโดยใช้เทคนิคการทดสอบแบบกล่องดำ (black box testing) และ 3) จัดทำคู่มือการใช้งาน

3. ตรวจสอบคุณภาพของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติที่พัฒนาขึ้น

การตรวจสอบคุณภาพของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติที่พัฒนาขึ้น มีขั้นตอนการดำเนินการดังนี้ 1) สร้างแบบประเมินคุณภาพระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ ที่มีรูปแบบเป็นมาตรประเมินค่า (rating scale) 5 ระดับ 2) นำแบบประเมินคุณภาพระบบที่สร้างขึ้น ส่งให้ผู้เชี่ยวชาญทางด้านการวัดและประเมินผลจำนวน 3 คน ตรวจสอบคุณภาพของแบบประเมินก่อนนำไปใช้ประเมินจริง 3) ประเมินคุณภาพระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น โดยให้ผู้เชี่ยวชาญจำนวน 5 คน ประกอบด้วย (1) ผู้เชี่ยวชาญด้านคอมพิวเตอร์ จำนวน 3 คน (2) ผู้เชี่ยวชาญด้านการวัดและประเมินผล จำนวน 1 คน และ (3) ประธานชุดวิชาหรืออาจารย์ผู้สอนที่เป็นผู้แทนคณะกรรมการบริหารชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 1 คน ทำการทดสอบการใช้งานระบบที่พัฒนาขึ้น พร้อมทั้งทำแบบประเมินคุณภาพระบบ ซึ่งมีประเด็นพิจารณาทั้งสิ้น 4 ด้าน ได้แก่ (1) ด้านการออกแบบระบบ (2) ด้านการทำงานของระบบ (3) ด้านการใช้งานระบบ และ (4) ด้านคู่มือการใช้งานระบบ 4) วิเคราะห์ผลการประเมินคุณภาพระบบ จากผู้เชี่ยวชาญโดยการคำนวณค่าเฉลี่ย (mean) และความเบี่ยงเบนมาตรฐาน (standard deviation) จากนั้นนำค่าเฉลี่ยดังกล่าว ไปเทียบ

กับเกณฑ์ระดับคุณภาพว่ารายการประเมินนั้นมีคุณภาพอยู่ที่ระดับใด ซึ่งระบบที่พัฒนาขึ้นต้องมีผลการประเมินคุณภาพ ตั้งแต่ ระดับมากขึ้นไป ทุกรายการประเมินจึงจะถือว่าผ่านเกณฑ์การประเมิน 5) สรุปผลการประเมินคุณภาพระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้นพร้อมทั้งแก้ไขตามข้อเสนอแนะของผู้เชี่ยวชาญ และ 6) นำระบบไปให้ผู้ใช้งานที่เกี่ยวข้องทดลองใช้งาน

ระยะที่ 3 การตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์

การวิจัยในระยะนี้มีวัตถุประสงค์เพื่อตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบจำนวน 2 ฉบับ ได้แก่ (1) แบบทดสอบต้นแบบ และ (2) แบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ (ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่ผู้วิจัยพัฒนาขึ้น) ซึ่งมีวิธีดำเนินการวิจัยดังนี้

1. สร้างแบบทดสอบด้วยระบบคอมพิวเตอร์ที่พัฒนาขึ้น

การสร้างแบบทดสอบด้วยระบบคอมพิวเตอร์ที่พัฒนาขึ้น ผู้วิจัยดำเนินการโดยกำหนดให้ระบบคอมพิวเตอร์สร้างข้อสอบจำนวน 42 ข้อ ตามรูปแบบของโมเดลข้อสอบที่ได้สร้างไว้แล้วจากข้อสอบต้นแบบทุกข้อในการวิจัยระยะที่ 1 โดยกำหนดเงื่อนไขให้โมเดลข้อสอบแต่ละข้อ สามารถนำไปสร้างข้อสอบได้จำนวน 1 ข้อต่อหนึ่งโมเดลข้อสอบ เพื่อใช้เป็นตัวแทนของข้อสอบทั้งหมดที่โมเดลข้อสอบนั้น ๆ สามารถสร้างได้ ซึ่งตัวแทนข้อสอบของแต่ละโมเดลข้อสอบดังกล่าว ระบบจะสร้างด้วยวิธีการสุ่มค่าตัวแปรต่าง ๆ ภายใต้ออบเจกต์ตัวแปรที่กำหนดไว้ในแต่ละโมเดล จากนั้นนำข้อสอบทั้ง 42 ข้อ มาจัดฉบับเป็นแบบทดสอบเพื่อนำไปตรวจสอบความเป็นคู่ขนานเทียบกับข้อสอบต้นแบบต่อไป

2. ตรวจสอบความเป็นคู่ขนานของข้อสอบรายข้อระหว่างข้อสอบจากแบบทดสอบต้นแบบกับข้อสอบจากแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยผู้เชี่ยวชาญ

การตรวจสอบความเป็นคู่ขนานระหว่างข้อสอบจากแบบทดสอบต้นแบบ (ฉบับ A) กับข้อสอบจากแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น (ฉบับ B) ผู้วิจัยดำเนินการในรูปแบบให้ผู้เชี่ยวชาญประเมินความเป็นคู่ขนานของข้อสอบเป็นรายข้อ จากแบบทดสอบทั้ง 2 ฉบับดังกล่าว โดยมีขั้นตอนการดำเนินการดังนี้ 1) สร้างแบบประเมินความเป็นคู่ขนานของข้อสอบ โดยประยุกต์ใช้ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ มาเป็นเกณฑ์การประเมินคุณภาพโมเดลข้อสอบของแต่ละรายการประเมิน พร้อมทั้งนำแบบประเมินความเป็นคู่ขนานของข้อสอบ ส่งให้ผู้เชี่ยวชาญทางด้านการวัดและประเมินผลจำนวน 3 คน ตรวจสอบคุณภาพของแบบประเมินก่อนนำไปใช้ประเมินจริง 2) ประเมินความเป็นคู่ขนานของข้อสอบเป็นรายข้อ โดยให้ผู้เชี่ยวชาญกลุ่มเดียวกันกับที่ประเมินคุณภาพโมเดลข้อสอบ ทำการตรวจสอบความเป็นคู่ขนานของแบบทดสอบทั้ง 2 ฉบับ พร้อมทั้งทำแบบประเมินความเป็นคู่ขนานของข้อสอบ ซึ่งมีประเด็นพิจารณาทั้งสิ้น 4 ประเด็น ได้แก่ (1) ความสอดคล้องของเนื้อหาข้อสอบ (2) ความสอดคล้องของวัตถุประสงค์การเรียนรู้ (3) ความสอดคล้องของระดับพฤติกรรมการเรียนรู้ และ (4) ความสอดคล้องของความยาก และ 3) วิเคราะห์ผลการประเมินความเป็นคู่ขนานของข้อสอบจากผู้เชี่ยวชาญ โดยการคำนวณค่าดัชนีความสอดคล้อง ซึ่งข้อสอบที่มีความเป็นคู่ขนานกัน ต้องมีค่าดัชนีความสอดคล้องของทุกรายการประเมิน เป็นไปตามเกณฑ์ที่กำหนดคือ ตั้งแต่ 0.5 ขึ้นไป และสรุปผลการประเมินความเป็นคู่ขนานของข้อสอบ

3. ตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น โดยการทดลองสอบ

การตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น โดยการทดลองสอบ มีขั้นตอนการดำเนินการดังนี้ 1) นำแบบทดสอบต้นแบบ จำนวน 1 ฉบับ และแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น จำนวน 1 ฉบับ ให้นักศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช ที่ลงทะเบียนชุดวิชาคณิตศาสตร์และสถิติสำหรับวิทยาศาสตร์และเทคโนโลยี จำนวน 30 คน ทำการทดสอบ โดยใช้เวลาทดสอบฉบับละ 2 ชั่วโมง 2) ตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น โดยใช้ผลจากการทำแบบทดสอบทั้ง 2 ฉบับ ของนักศึกษา มาดำเนินการตรวจสอบความเป็นคู่ขนานของแบบทดสอบทั้ง 2 ฉบับ ดังกล่าว ในด้านต่าง ๆ ดังนี้ (1) ความแตกต่างของค่าความยากรายข้อและค่าความยากเฉลี่ยรายฉบับ ทดสอบด้วยสถิติ Dependent t-test (2) ความแตกต่างของค่าอำนาจจำแนกเฉลี่ยรายฉบับ ทดสอบด้วยสถิติ Dependent t-test (3) ความแตกต่างของค่าความแปรปรวนรายข้อและรายฉบับ ทดสอบด้วยสถิติ F-test (4) ความแตกต่างของค่าเฉลี่ยของคะแนนแบบทดสอบ ทดสอบด้วยสถิติ dependent t-test (5) ค่าความเชื่อมั่นแบบสมมูลของแบบทดสอบ โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนแบบทดสอบ (6) ค่าความเชื่อมั่นของแบบทดสอบ โดยใช้สูตรของ Kuder-Richardson 20 (KR_{20}) และ (7) ค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบ และ 3) สรุปผลการตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น

ผลการวิจัย

1. ผลการสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์ฯ

ผู้วิจัยแบ่งผลการวิจัยออกเป็น 3 ส่วนได้แก่ (1) ผลการสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ จำนวน 42 ข้อ (2) ผลการประเมินคุณภาพโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ และ (3) ผลการออกแบบคลังข้อมูลข้อสอบชุดวิชาคณิตศาสตร์ฯ โดยแต่ละส่วนมีรายละเอียดดังนี้

1.1 ผลการสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ แสดงตัวอย่างดังตาราง 2

ตาราง 2 ตัวอย่างผลการสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ เรื่อง อนุพันธ์

ข้อสอบต้นแบบ: ข้อที่ 1	
อนุพันธ์ของฟังก์ชัน $f(x) = 2x^5 + 5x^3 - 8$ มีค่าเท่ากับข้อใด	
1. $10x^4 + 15x^2$	4. $10x^4 + 15x^2 - 8$
2. $10x^5 + 15x^3$	5. $10x^6 + 15x^4 - 8$
3. $10x^6 + 15x^4$	
เฉลย : 1	
โมเดลข้อสอบ : ข้อที่ 1	

ตาราง 2 (ต่อ)

ประเภทโมเดลข้อสอบ : โมเดลที่ 5

โจทย์ : อนุพันธ์ของฟังก์ชัน $f(x) = I_1x^{I_4} + I_2x^{I_5} + I_3$ มีค่าเท่ากับข้อใด

ส่วนประกอบ :

I_1, I_2, I_3 : มีขอบเขตข้อมูลเป็นจำนวนเต็มในช่วง -9 ถึง 9 และ $I_1, I_2 \neq 0, 1, -1$ และ $I_1 \neq I_2 \neq I_3$

I_4, I_5 : มีขอบเขตข้อมูลเป็นจำนวนเต็มบวกในช่วง 2 ถึง 9 และ $I_4 > I_5$

ตัวเลือก :

- | | |
|--|--|
| 1. $(I_1I_4)x^{(I_4-1)} + (I_2I_5)x^{(I_5-1)}$ | 4. $(I_1I_4)x^{(I_4-1)} + (I_2I_5)x^{(I_5-1)} + I_3$ |
| 2. $(I_1I_4)x^{I_4} + (I_2I_5)x^{I_5}$ | 5. $(I_1I_4)x^{(I_4+1)} + (I_2I_5)x^{(I_5+1)} + I_3$ |
| 3. $(I_1I_4)x^{(I_4+1)} + (I_2I_5)x^{(I_5+1)}$ | |

ข้อมูลเสริม : ไม่มี

เฉลย : 1

1.2 ผลการตรวจสอบคุณภาพโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ โดยผู้เชี่ยวชาญดังกล่าว พบว่า โมเดลข้อสอบที่สร้างขึ้นจำนวน 42 ข้อ ทุกข้อผ่านเกณฑ์คุณภาพที่กำหนดไว้ คือโมเดลข้อสอบทุกข้อหรือทุกโมเดล ต้องมีค่าดัชนีความสอดคล้องของทุกรายการประเมิน เป็นไปตามเกณฑ์ที่กำหนดคือ ตั้งแต่ 0.5 ขึ้นไป

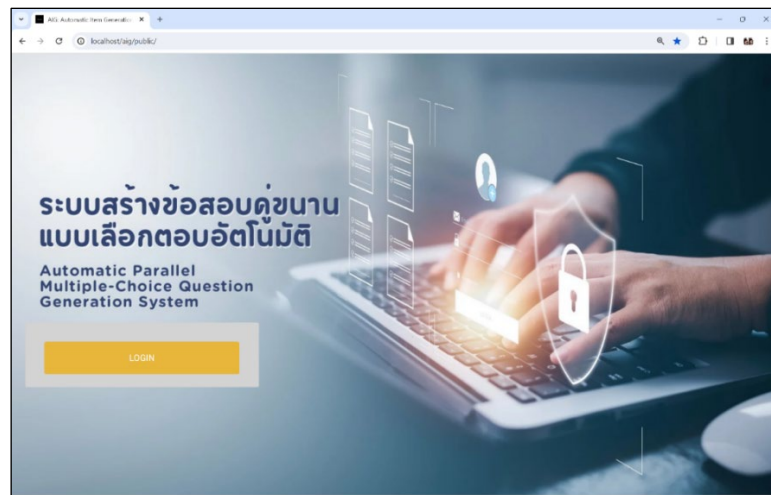
1.3 ผลการออกแบบคลังข้อสอบชุดวิชาคณิตศาสตร์ฯ มีดังนี้ 1) องค์ประกอบของคลังข้อสอบที่ออกแบบขึ้น ประกอบด้วยองค์ประกอบของข้อมูล แบ่งเป็น 2 กลุ่ม ได้แก่ หมวดข้อมูล และรายการข้อมูลที่อยู่ภายใต้หมวดข้อมูลที่สร้างขึ้น โดยข้อมูลที่จัดเก็บไว้ในคลังข้อมูลต้องเป็นข้อมูลชนิดตัวอักษร หรือข้อมูลชนิดข้อความเท่านั้น 2) ขั้นตอนการสร้างคลังข้อสอบผ่านระบบสร้างข้อสอบคู่ขนานที่พัฒนาขึ้น ดำเนินการได้ดังนี้ (1) กำหนดข้อมูลที่ต้องการจัดเก็บในคลังข้อมูล (2) เลือกเมนูหรือฟังก์ชันจัดการคลังข้อมูลข้อสอบ (3) ระบบแสดงหน้าจอให้กรอกข้อมูล ชื่อหมวดข้อมูล และรายการข้อมูลภายใต้หมวดข้อมูลแต่ละหมวด (4) กรอกชื่อหมวดข้อมูล พร้อมทั้งกดปุ่มเพิ่มหมวดข้อมูล จากนั้นให้กรอกรายการข้อมูลภายใต้หมวดข้อมูลที่สร้างขึ้นแต่ละหมวด พร้อมทั้งกดปุ่มเพิ่มรายการข้อมูล เพื่อบันทึกข้อมูลคลังข้อมูลข้อสอบ (5) ระบบแสดงรายการข้อมูลที่ถูกจัดเก็บไว้ในคลังข้อมูลข้อสอบ (6) กรณีที่ต้องการ ลบ/แก้ไขข้อมูลคลังข้อมูลข้อสอบ สามารถกดปุ่ม ลบ/แก้ไขได้ที่ส่วนแสดงข้อมูลรายการข้อมูลในคลังข้อมูลข้อสอบที่ได้บันทึกไว้ในระบบ และ 3) ผลการออกแบบขั้นตอนการนำข้อมูลจากคลังข้อมูลข้อสอบไปใช้งาน เป็นดังนี้คือ เมื่อผู้ใช้งานมีการกำหนดค่าตัวแปรเป็นชนิดตัวอักษร หรือกำหนดค่าตัวแปรเป็นชนิดข้อความ ให้กับตัวแปรใดตัวแปรหนึ่งของโมเดลข้อสอบ ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น จะต้องแสดงข้อมูลในคลังข้อมูลให้ผู้ใช้งานพิจารณาเลือกนำไปใช้ กำหนดค่าตัวแปรให้กับโมเดลข้อสอบทุกครั้ง จากผลการออกแบบคลังข้อมูลข้อสอบชุดวิชาคณิตศาสตร์ฯ ดังกล่าวข้างต้น ได้ถูกนำไปใช้เป็นข้อมูลสำหรับพัฒนาค้นคลังข้อมูลขึ้นใช้งานจริง เพื่อเป็นฟังก์ชันการทำงาน ฟังก์ชันหนึ่งของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น

2. ผลการพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ

ผู้วิจัยแบ่งการนำเสนอผลการวิจัยออกเป็น 3 ส่วน ได้แก่ (1) ผลการออกแบบระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ (2) ผลการพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ และ (3) ผลการตรวจสอบคุณภาพของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ มีรายละเอียดดังนี้

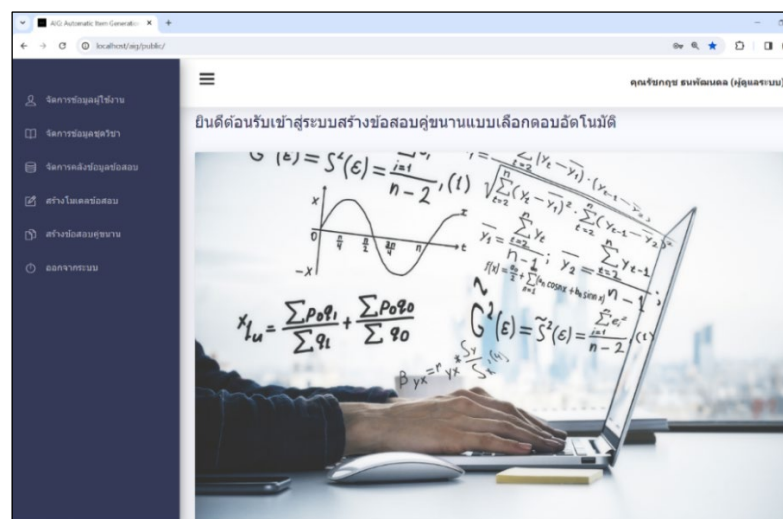
2.1 ผลการออกแบบระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ มีรายละเอียด ประกอบด้วย ฟังก์ชันการทำงานทั้งหมดจำนวน 6 ฟังก์ชัน ได้แก่ (1) ฟังก์ชันการเข้าสู่ระบบ (2) ฟังก์ชันจัดการข้อมูลผู้ใช้งาน (3) ฟังก์ชันจัดการข้อมูลชุดวิชา (4) ฟังก์ชันจัดการคลังข้อมูลข้อสอบ (5) ฟังก์ชันสร้างโมเดลข้อสอบ และ (6) ฟังก์ชันสร้างข้อสอบคู่ขนาน

2.2 ผลการพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ แสดงตัวอย่างหน้าจอของบางฟังก์ชันการทำงาน ได้ดังภาพ 1-4



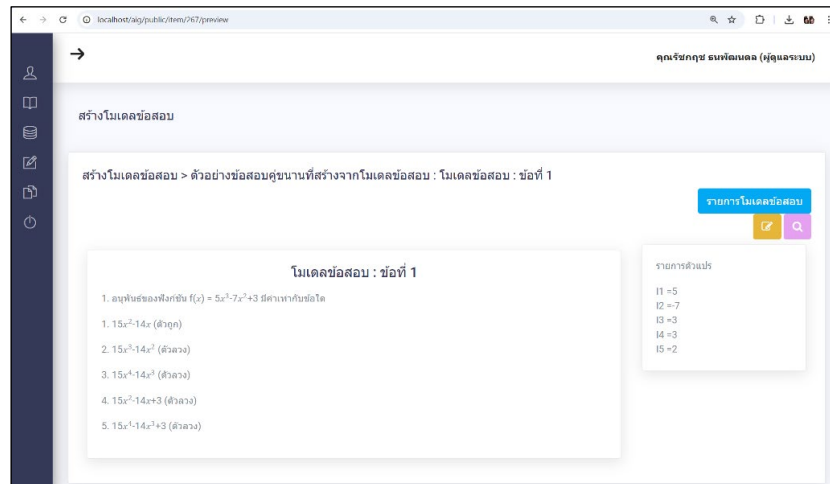
ภาพ 1 หน้าจอแรกเมื่อผู้ใช้งานเรียกใช้งานระบบ

จากภาพ 1 หน้าจอแรกเมื่อผู้ใช้งานเรียกใช้งานระบบ ประกอบด้วย (1) ชื่อระบบ และ (2) ปุ่ม Login เพื่อเชื่อมโยงไปยังหน้าจอเข้าสู่ระบบ สำหรับให้ผู้ใช้งานกรอกรหัสผู้ใช้และรหัสผ่านเพื่อเข้าสู่ระบบ



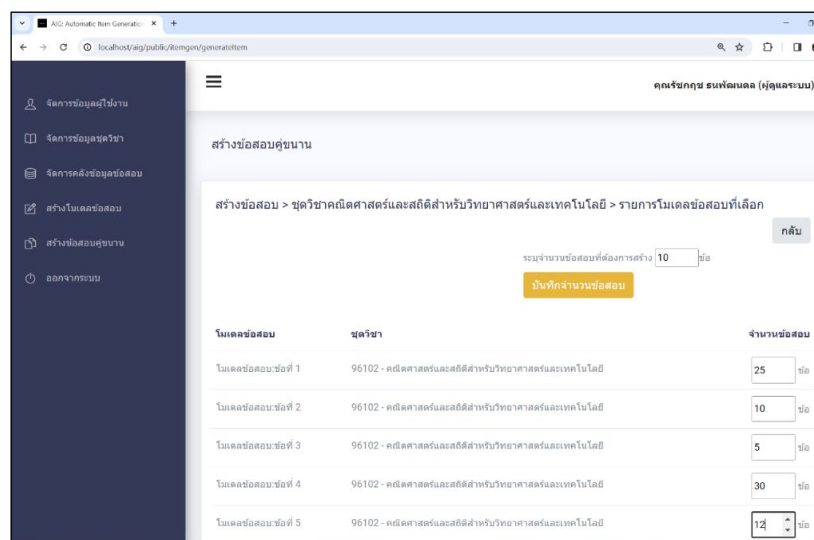
ภาพ 2 หน้าจอหลักของระบบ

จากภาพ 2 หน้าจอหลักของระบบ ประกอบด้วย (1) แถบเมนูให้คลิกเลือกฟังก์ชันการทำงานของระบบ ได้แก่ ฟังก์ชันจัดการข้อมูลผู้ใช้งาน ฟังก์ชันจัดการข้อมูลชุดวิชา ฟังก์ชันจัดการคลังข้อมูลข้อสอบ ฟังก์ชันสร้างโมเดลข้อสอบ ฟังก์ชันสร้างข้อสอบคู่ขนาน และออกจากระบบ (2) ข้อความแสดง ชื่อ – นามสกุล ผู้ใช้งานระบบ และ (3) ส่วนแสดงผลข้อมูลเมื่อผู้ใช้งานคลิกเลือกฟังก์ชันการทำงานของฟังก์ชันใดฟังก์ชันหนึ่งในแถบเมนู ซึ่งหน้าจอดังกล่าวพัฒนาขึ้นสำหรับให้ผู้ใช้เลือกเมนูที่ต้องการใช้งานระบบ



ภาพ 3 หน้าจอแสดงตัวอย่างข้อสอบคู่ขนานที่สร้างจากโมเดลข้อสอบ

จากภาพ 3 หน้าจอแสดงตัวอย่างข้อสอบคู่ขนานที่สร้างจากโมเดลข้อสอบข้อที่ 1 ในตาราง 2 ซึ่งหน้าจอประกอบด้วยส่วนสำคัญหลัก 2 ส่วน ได้แก่ (1) ส่วนแสดงเนื้อหาข้อสอบที่สร้างจากโมเดลข้อสอบ (2) ส่วนแสดงรายการตัวแปรและค่าตัวแปรของโมเดลข้อสอบที่ระบบสุ่มค่าได้ ซึ่งหน้าจอดังกล่าวพัฒนาขึ้นสำหรับ ใช้ตรวจสอบความถูกต้องโครงสร้าง และเนื้อหาของข้อสอบคู่ขนานที่สร้างจากโมเดลข้อสอบก่อนนำไปสร้างข้อสอบคู่ขนาน



ภาพ 4 หน้าจอระบุจำนวนข้อสอบคู่ขนานที่ต้องการสร้างจากโมเดลข้อสอบ

จากภาพ 4 หน้าจอสำหรับให้ผู้ใช้งานระบุจำนวนข้อสอบคู่ขนานที่ต้องการสร้างจากโมเดลข้อสอบที่สร้างไว้ สามารถกำหนดให้ทุกโมเดลข้อสอบสร้างข้อสอบจำนวนข้อเท่ากัน หรือต่างกันก็ได้ เมื่อระบบสร้างข้อสอบคู่ขนานครบตามจำนวนที่ผู้ใช้งานกำหนดเสร็จสิ้นแล้ว จะทำการบันทึกไฟล์ข้อสอบคู่ขนานที่ถูกสร้างลงในเครื่องคอมพิวเตอร์ที่ใช้งานระบบในรูปแบบของไฟล์ word (.docx)

2.3 ผลการตรวจสอบคุณภาพในภาพรวม ของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติโดยใช้วิธีการสร้างข้อสอบอัตโนมัติ โดยผู้เชี่ยวชาญ แสดงรายละเอียดดังตาราง 3

ตาราง 3 ผลการตรวจสอบคุณภาพในภาพรวมของระบบโดยผู้เชี่ยวชาญ

รายการ	<i>M</i>	<i>SD</i>	ระดับคุณภาพ
ภาพรวมของระบบ			
1. การออกแบบระบบ	4.37	0.55	มาก
2. การทำงานของระบบ	4.56	0.54	มากที่สุด
3. การใช้งานระบบ	4.53	0.51	มากที่สุด
4. คู่มือการใช้งานระบบ	4.10	0.55	มาก
เฉลี่ยรวม	4.39	0.56	มาก

จากตาราง 3 ผลการวิเคราะห์ข้อมูลการตรวจสอบคุณภาพในภาพรวม ของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติฯ โดยผู้เชี่ยวชาญ พบว่าระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ มีคุณภาพเฉลี่ยในภาพรวมอยู่ในระดับมาก และนอกจากนี้ระบบที่พัฒนาขึ้นยังมีผลประเมินการใช้งานระบบจากผู้ใช้งานเฉลี่ยในภาพรวมอยู่ในระดับมากอีกด้วย

3. ผลการตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์

3.1 ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบรายข้อระหว่างข้อสอบจากแบบทดสอบต้นแบบกับข้อสอบจากแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยผู้เชี่ยวชาญ จำนวน 42 ข้อ พบว่า ข้อสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้นมีความเป็นคู่ขนานกับข้อสอบต้นแบบทุกข้อ เนื่องจากผ่านเกณฑ์ความเป็นคู่ขนานของข้อสอบที่กำหนดไว้ คือข้อสอบที่มีความเป็นคู่ขนานกัน ต้องมีค่าดัชนีความสอดคล้องของทุกรายการประเมิน เป็นไปตามเกณฑ์ที่กำหนดคือ ตั้งแต่ 0.5 ขึ้นไป

3.2 ผลการตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยการทดลองสอบ มีผลการวิเคราะห์ข้อมูลที่เกี่ยวข้องกับการตรวจสอบความเป็นคู่ขนานของแบบทดสอบ ดังนี้

1) ค่าความยากและผลการทดสอบความแตกต่างของค่าความยาก พบว่าค่าความยากรายข้อของแบบทดสอบต้นแบบ (p_A) อยู่ระหว่าง 0.33-0.77 ค่าความยากเฉลี่ยทั้งฉบับเท่ากับ 0.52 ค่าความยากรายข้อของแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น (p_B) อยู่ระหว่าง 0.27-0.77 ค่าความยากเฉลี่ยทั้งฉบับเท่ากับ 0.51 และผลการทดสอบความแตกต่างของค่าความยากเฉลี่ยทั้งฉบับ แสดงดังตาราง 4

ตาราง 4 ผลการทดสอบความแตกต่างของค่าความยากเฉลี่ยทั้งฉบับของแบบทดสอบ 2 ฉบับ

ค่าความยาก	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i> – value
p_A	42	0.52	0.12	1.25	41	0.24
p_B	42	0.51	0.13			

จากตาราง 4 เมื่อพิจารณาจากค่าสถิติที่ใช้ทดสอบความแตกต่างของข้อมูลแล้ว สามารถสรุปได้ว่าค่าความยากเฉลี่ยทั้งฉบับของแบบทดสอบทั้ง 2 ฉบับ ไม่แตกต่างกัน

2) ค่าอำนาจจำแนกและผลการทดสอบความแตกต่างของค่าอำนาจจำแนก พบว่าค่าอำนาจจำแนกรายข้อของแบบทดสอบต้นแบบ (r_A) อยู่ระหว่าง 0.20-0.80 และค่าอำนาจจำแนกรายข้อของแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น (r_B) อยู่ระหว่าง 0.20-0.73 ผลการทดสอบความแตกต่างของค่าอำนาจจำแนกเฉลี่ยทั้งฉบับแสดงดังตาราง 5

ตาราง 5 ผลการทดสอบความแตกต่างของค่าอำนาจจำแนกเฉลี่ยทั้งฉบับของแบบทดสอบ 2 ฉบับ

ค่าอำนาจจำแนก	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i> – value
r_A	42	0.41	0.19	0.10	41	0.91
r_B	42	0.40	0.15			

จากตาราง 5 เมื่อพิจารณาจากค่าสถิติที่ใช้ทดสอบความแตกต่างของข้อมูลแล้ว สามารถสรุปได้ว่าค่าอำนาจจำแนกเฉลี่ยทั้งฉบับของแบบทดสอบทั้ง 2 ฉบับ ไม่แตกต่างกัน

3) ค่าความแปรปรวนและผลการทดสอบความแตกต่างของค่าความแปรปรวน พบว่าค่าความแปรปรวนรายข้อของแบบทดสอบต้นแบบ อยู่ระหว่าง 0.19-0.26 ค่าความแปรปรวนทั้งฉบับเท่ากับ 0.24 ค่าความแปรปรวนรายข้อของแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น อยู่ระหว่าง 0.19-0.26 ค่าความแปรปรวนทั้งฉบับเท่ากับ 0.25 และเมื่อพิจารณาจากค่าสถิติที่ใช้ทดสอบความแตกต่างของข้อมูลที่มีค่า $F = 1.04$ ($\alpha = 0.05$, $df_1 = 29$, $df_2 = 29$, $F = 1.85$) สามารถสรุปได้ว่าค่าความแปรปรวนทั้งฉบับของแบบทดสอบทั้ง 2 ฉบับไม่แตกต่างกัน

4) ค่าเฉลี่ยและผลการทดสอบความแตกต่างของค่าเฉลี่ยของคะแนนแบบทดสอบระหว่างแบบทดสอบต้นแบบ ($\bar{T}_A$) และแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้น ($\bar{T}_B$) แสดงดังตาราง 6

ตาราง 6 ค่าเฉลี่ยและผลการทดสอบความแตกต่างของค่าเฉลี่ยของคะแนนแบบทดสอบ 2 ฉบับ

คะแนนแบบทดสอบ	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i> – value
$\bar{T}_A$	30	21.9	9.26	0.52	29	0.60
$\bar{T}_B$	30	21.0	9.28			

จากตาราง 6 เมื่อนำแบบทดสอบไปใช้ทดสอบกับนักศึกษาจำนวน 30 คน พบว่าค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบต้นแบบและแบบทดสอบที่สร้างจากระบบคอมพิวเตอร์ที่พัฒนาขึ้นเท่ากับ $M=21.9$, $SD=9.26$ และ $M=21.0$, $SD=9.28$ ตามลำดับ เมื่อพิจารณาจากค่าสถิติที่ใช้ทดสอบความแตกต่างของข้อมูลแล้ว สามารถสรุปได้ว่าค่าเฉลี่ยของคะแนนแบบทดสอบทั้ง 2 ฉบับ ไม่แตกต่างกัน

5) ค่าความเชื่อมั่นแบบสมมูลของแบบทดสอบ แสดงดังตาราง 7

ตาราง 7 ผลการหาค่าความเชื่อมั่นแบบสมมูลของแบบทดสอบ 2 ฉบับ

คะแนนแบบทดสอบ	<i>n</i>	Correlation	<i>p</i> – value
แบบทดสอบต้นแบบ	30	0.98	0.00
แบบทดสอบสร้างจากระบบคอมพิวเตอร์			

จากตาราง 7 พบว่าค่าความเชื่อมั่นแบบสมมูลของแบบทดสอบโดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนแบบทดสอบต้นแบบและแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้นเท่ากับ 0.98 ดังนั้นคะแนนแบบทดสอบต้นแบบและแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ที่พัฒนาขึ้นจึงมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

6) ค่าความเชื่อมั่นของแบบทดสอบ (r_{tt}) โดยใช้สูตรของ Kuder-Richardson 20 (KR_{20}) และค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบ (SE) แสดงได้ดังตาราง 8

ตาราง 8 ผลการหาค่าความเชื่อมั่นของแบบทดสอบ 2 ฉบับ

แบบทดสอบ	k (จำนวนข้อสอบ)	n	r_{tt}	SE
แบบทดสอบต้นแบบ	42	30	0.89	2.98
แบบทดสอบสร้างจากระบบคอมพิวเตอร์	42	30	0.90	2.84

จากตาราง 8 พบว่าแบบทดสอบทั้ง 2 ฉบับมีค่าความเชื่อมั่นและความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบใกล้เคียงกัน ซึ่งเป็นสิ่งหนึ่งที่บ่งบอกได้ว่าแบบทดสอบทั้ง 2 ฉบับมีความเป็นคู่ขนานกัน

อภิปรายผล

ผลการสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์ฯ การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติฯ และการตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ มีประเด็นอภิปรายดังนี้

1. การสร้างโมเดลข้อสอบและออกแบบคลังข้อมูลชุดวิชาคณิตศาสตร์ฯ อภิปรายผลการวิจัยได้ดังนี้
1) ประเภทของโมเดลข้อสอบที่สามารถสร้างขึ้นในงานวิจัยนี้มีจำนวนทั้งสิ้น 2 ประเภท ได้แก่ โมเดลที่ 5 และโมเดลที่ 8 โดยไม่สามารถสร้างโมเดลข้อสอบภายใต้หลักการการสร้างโมเดลข้อสอบของ Gierl et al. (2008) ได้ครบทั้ง 10 ประเภท เนื่องจากข้อจำกัดทางด้านเนื้อหาของข้อสอบต้นแบบที่ไม่สามารถจัดเข้ากับโครงสร้างของโมเดลข้อสอบอีก 8 ประเภท ที่เหลือได้ซึ่งสอดคล้องกับงานวิจัยของ Chantharamaha & Chadcham (2018); Jantana & Pradujprom (2020); Pradujprom & Pantong (2019) ที่ไม่สามารถสร้างโมเดลข้อสอบวิชาคณิตศาสตร์ได้ครบทั้ง 10 โมเดล เนื่องจากข้อจำกัดทางด้านเนื้อหาของข้อสอบต้นแบบดังกล่าว อย่างไรก็ตามการสร้างโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) ก็ได้มีข้อกำหนดไว้ว่าต้องสร้างโมเดลข้อสอบให้ครบทั้ง 10 ประเภทเช่นกัน และ 2) คลังข้อมูล หรือคลังข้อมูลข้อสอบ เป็นคลังข้อมูลที่ออกแบบขึ้น

เพื่อพัฒนาเป็นฟังก์ชันการทำงานฟังก์ชันหนึ่งของระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติ ที่สามารถรองรับการสร้างโมเดลข้อสอบในระบบสร้างข้อสอบคู่ขนานอัตโนมัติ ให้มีความสะดวกในการสร้างโมเดลข้อสอบมากยิ่งขึ้นเท่านั้น ซึ่งผู้ใช้งานอาจจะใช้งานข้อมูลที่มีในคลังข้อมูลหรือไม่ก็ได้ ทั้งนี้ขึ้นอยู่กับความประสงค์ และบริบทของเนื้อหาข้อสอบที่นำมาสร้างโมเดลข้อสอบนั้น ๆ ด้วยว่าถูกกำหนดให้มีตัวแปรชนิดตัวอักษร หรือตัวแปรชนิดข้อความหรือไม่ และตัวแปรที่ต้องการใช้งานถูกสร้างไว้ในคลังข้อมูลแล้วหรือไม่

2. การพัฒนาระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติฯ อภิปรายผลการวิจัยได้ดังนี้

1) ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น สามารถสร้างข้อสอบคู่ขนานอัตโนมัติได้เป็นผลจากการบูรณาการปัจจัยดังต่อไปนี้ (1) หลักการสร้างข้อสอบอัตโนมัติผ่านโมเดลข้อสอบของ Gierl et al. (2008) (2) การสร้างโมเดลข้อสอบชุดวิชาคณิตศาสตร์ฯ ที่คำนึงถึงหลักการความเป็นคู่ขนานของข้อสอบ และ (3) หลักการพัฒนาระบบตามวงจรพัฒนาระบบ SDLC ตามเอกสารของ Preechapanich (2014). และหลักการพัฒนาโปรแกรม นอกจากนี้ระบบสร้างข้อสอบคู่ขนานที่ดังกล่าว ยังถูกพัฒนาให้สามารถรองรับการสร้างข้อสอบคู่ขนาน ของข้อสอบวิชาคณิตศาสตร์ เรื่องอื่น ๆ ที่นอกเหนือจากที่แสดงในงานวิจัยนี้ได้ รวมทั้งสามารถรองรับเนื้อหาข้อสอบวิชาอื่น ๆ ที่บริบทเนื้อหาวิชาดังกล่าวสามารถประยุกต์เข้ากับหลักการสร้างโมเดลข้อสอบของ Gierl et al. (2008) ได้ 2) ระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น ผู้วิจัยสามารถออกแบบและพัฒนาระบบ ให้ผู้ใช้งานสร้างโมเดลข้อสอบได้ภายใต้ฟังก์ชันสร้างโมเดลข้อสอบที่พัฒนาขึ้น ตามหลักการของ Gierl et al. (2008) ได้ครบทั้ง 10 โมเดล ซึ่งแตกต่างจากระบบสร้างข้อสอบอัตโนมัติในงานวิจัยอื่น เช่นงานวิจัยของ Chantharamaha & Chadcham (2018); Jantana & Pradujprom (2020); Pradujprom & Pantong (2019); Raksombat et al. (2022) ที่มีข้อจำกัดที่ผู้ใช้งานไม่สามารถสร้างโมเดลข้อสอบเพิ่มเติมหรือแก้ไขโมเดลข้อสอบได้ แต่อย่างไรก็ตามฟังก์ชันสร้างโมเดลข้อสอบที่พัฒนาขึ้นดังกล่าว ยังมีข้อจำกัดที่ผู้ใช้งานต้องเป็นผู้สร้างโมเดลข้อสอบในระบบเอง ยังไม่สามารถพัฒนาให้ระบบสามารถสร้างโมเดลข้อสอบได้อย่างอัตโนมัติ

3. การตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ อภิปรายผลการวิจัยได้ว่า ข้อสอบจากแบบทดสอบต้นแบบ และข้อสอบจากแบบทดสอบสร้างขึ้นจากระบบคอมพิวเตอร์ทั้ง 42 ข้อ มีความเป็นคู่ขนานกัน ทั้งเชิงเนื้อหาและเชิงสถิติ อาจเป็นผลจากในกระบวนการสร้างโมเดลข้อสอบ ได้นำข้อมูลของลักษณะเฉพาะของข้อสอบ และข้อสอบต้นแบบ มาเป็นขอบเขตข้อมูลในการสร้างโมเดลข้อสอบ จากนั้นจึงนำโมเดลข้อสอบที่สร้างขึ้นไปใช้สร้างข้อสอบคู่ขนานผ่านระบบที่พัฒนาขึ้นต่อไป จึงส่งผลให้ข้อสอบคู่ขนานที่สร้างขึ้น มีรูปแบบเนื้อหา และคุณภาพของข้อสอบใกล้เคียงกับข้อสอบต้นแบบหรือข้อสอบต้นฉบับ ทำให้การวัดและประเมินผลการเรียนรู้ด้วยแบบทดสอบคู่ขนานดังกล่าวมีความเท่าเทียมกัน หรือมีมาตรฐานเดียวกัน และมีความน่าเชื่อถือ นอกจากนี้การตรวจสอบตรวจสอบความเป็นคู่ขนานระหว่างแบบทดสอบต้นแบบกับแบบทดสอบที่สร้างขึ้นจากระบบคอมพิวเตอร์ โดยอาศัยดุลยพินิจของผู้เชี่ยวชาญจำนวน 3 คน นั้นมีความเพียงพอที่ทำให้ผลการทดสอบแสดงให้เห็นว่าแบบทดสอบทั้ง 2 ฉบับ มีความเป็นคู่ขนานกัน เนื่องจากแบบทดสอบที่นำมาให้ผู้เชี่ยวชาญตรวจสอบความเป็นคู่ขนานนั้น ได้สร้างจากโมเดลข้อสอบตามหลักการของ Gierl et al. (2008) จึงส่งผลให้ข้อสอบที่ได้มีแนวโน้มของความเป็นคู่ขนานกันอยู่ก่อนแล้ว ซึ่งสอดคล้องกับ Chantharamaha & Chadcham (2018)

และด้วยเหตุผลดังกล่าวยังส่งผลให้การทดลองสอบมีค่าสถิติที่บ่งบอกถึงความเป็นคู่ขนานของแบบทดสอบ โดยมีความสอดคล้องกับงานวิจัยของ Charuthat (2003); Kanchanapenkul (2011); Sawangboon et al. (2012) อีกด้วย

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 อาจารย์ผู้สอนสามารถนำระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น ไปปรับใช้สร้างข้อสอบคู่ขนานอัตโนมัติของวิชาอื่น ๆ ที่มีบริบทเนื้อหาใกล้เคียง หรือที่มีบริบทเนื้อหาที่สามารถประยุกต์เข้ากับหลักการสร้างโมเดลข้อสอบที่ใช้ในงานวิจัยนี้ได้ เช่น วิชาคณิตศาสตร์ระดับชั้นต่าง ๆ วิชาฟิสิกส์ วิชาเคมี เป็นต้น

1.2 อาจารย์ผู้สอนสามารถนำระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น ไปปรับใช้สร้างแบบฝึกหัด หรือข้อคำถาม เพื่อนำไปใช้วัดและประเมินผลการเรียนรู้ของผู้เรียนในรูปแบบต่าง ๆ ได้

1.3 ผู้ดูแลระบบคลังข้อสอบมาตรฐานสามารถนำข้อสอบคู่ขนานที่สร้างจากระบบสร้างข้อสอบคู่ขนานแบบเลือกตอบอัตโนมัติที่พัฒนาขึ้น ไปจัดเก็บไว้ในระบบคลังข้อสอบมาตรฐานเพื่อใช้วัดและประเมินผลการเรียนรู้ของผู้เรียน ในภาคการศึกษาต่าง ๆ ได้ต่อไป

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 จากวิธีการวิเคราะห์หรือตรวจสอบความเป็นคู่ขนานของแบบทดสอบในครั้งนี้อาศัยดุลยพินิจของผู้เชี่ยวชาญและการตรวจสอบความเป็นคู่ขนานโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม ดังนั้นในการวิจัยครั้งต่อไปอาจประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบมาตรวจสอบความเป็นคู่ขนานของข้อสอบ

2.2 จากชนิดของข้อสอบที่สร้างขึ้นในครั้งนี้เป็นแบบเลือกตอบ 100% ดังนั้นในการวิจัยครั้งต่อไปควรมีการศึกษาเพื่อพัฒนาระบบสร้างข้อสอบคู่ขนานรูปผสม หรือแบบอัตนัยอัตโนมัติ เพื่อให้สามารถใช้ในเนื้อหาวิชาอื่นได้ต่อไป

References

- Gierl, M. J., & Haladyna, T. (Eds.) (2013). *Automatic Item Generation: Theory and Practice*. Routledge.
- Gierl, M. J., & Lai, H. (2012). The Role of Item Models in Automatic Item Generation. *International Journal of Testing*, 12(3), 273-298.
- Gierl, M. J., & Lai, H. (2013). Instructional Topics in Educational Measurement (ITEMS) Module: Using Automated Processes to Generate Test Items. *Educational Measurement: Issues and Practice*, 32(3), 36-50.
- Gierl, M. J., Lai, H., & Turner, S. (2012). Using Automatic Item Generation to Create Multiple-Choice Items for Assessments in Medical Education. *Medical Education*, 46(8), 757-765.

Gierl, M.J., Zhou, J., & Alves, C. (2008). Developing a Taxonomy of Item Model Types to Promote Assessment Engineering. *Journal of Technology, Learning, and Assessment*, 7(2), 1-51.

Translate Thai Reference

Chantharamaha, S., & Chadcham, S. (2018). Automatic Item Generation in Mathematics for Grade Six Students Using Computer Software. *Research Methodology & Cognitive Science*, 16(1), 138-149. (in Thai)

Charuthat, S. (2003). Study of Parallel of Prathomsuksa 6 Mathematics Achievement Test Constructed with Facet Design [Master's thesis]. Chiang Mai University. (in Thai)

Jantana, S., & Pradujprom. P. (2020). Development of Automatic Item Generation Program in Mathematics for Grade Nine Level. *Journal of Educational Measurement*, 37(102), 162-177. (in Thai)

Kanchanapenkul, S. (2011) A Construction of Parallel Domain-Referenced Test by Using Facet Design on "Circles" for Matthayomsuksa 3. *Sripatum Chonburi Academic Journal*, 9(4), 56-60. (in Thai)

Pradujprom, P., & Panthong, K. (2019). A Computer Program for Developing Multiple Choice Test Items Using Automatic Item Generation Method. *Journal of Southern Technology*, 12(2), 74-87. (in Thai)

Preechapanich, O. (2014). *Guide book: System Analysis and Design, Complete version*. IDC Premier. (in Thai)

Premthongsuk, P., Chadcham, S., & Pradujprom, P. (2017). Development of the Next Item Selection Procedure Using Hurwicz Criterion and Item Exposure Control for Computerized Adaptive Testing. *Veridian E-Journal, Science and Technology Silpakorn University*, 4(6), 32-50. (in Thai)

Raksombat, S., Pradujphom, P., & Panthong, K. (2022). Development of an Automatic Item Generation System for Item Bank Classified by Content and Difficulty Levels: Application of Assessment Engineering Concepts. *Journal of Educational Measurement, Mahasarakham University*, 28(1), 311-330. (in Thai)

Sawangboon, N., Phattiyathani, S., & Wongrattana, C. (2012). A Construction of Mathematics Parallel Tests Using Facet Design on "Factoring of Degree-2 Polynomials" for Student in Matthayomsuksa 2. *Journal of Educational Measurement, Mahasarakham University*, 17(1), 369-378. (in Thai)

Srisa-ard, O. (1995). Verification of the validity of measurement tools by experts. *Journal of Educational Measurement, Mahasarakham University*, 1(1), 45-49. (in Thai)

การพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้
ของบุคลากรสายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
ด้วยการจัดการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับ
The Development of Learning Design Competency for Academic Support Staff
of King Mongkut's University of Technology Thonburi through Feedback
Processes in Experiential Learning

วรรณชร ไชยเดช^{1*} และ วารุณี ลักนโชคดี²

Wannachorn Chaidet^{1*} and Warunee Lapanachokdee²

(Received: December 4, 2024; Revised: January 5, 2025; Accepted: January 14, 2025)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อพัฒนากิจกรรมการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับสำหรับการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ และเพื่อศึกษาผลการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ที่ได้รับการจัดกิจกรรมการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับของบุคลากรสายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี กลุ่มตัวอย่าง คือ บุคลากรสายสนับสนุนวิชาการ จำนวน 31 คน ได้มาจากการเลือกแบบเจาะจง เครื่องมือที่ใช้ คือ ใบงานกิจกรรมและแบบตรวจสอบรายการพฤติกรรม วิเคราะห์ข้อมูลโดยใช้ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน การวิเคราะห์ความแปรปรวนแบบวัดซ้ำทางเดียว และการวิเคราะห์ค่าขนาดอิทธิพล ผลการวิจัยพบว่า กิจกรรมการเรียนรู้ประกอบด้วย 3 ส่วน ได้แก่ การเรียนรู้เนื้อหาสาระสำคัญด้วยตนเอง การเรียนรู้ผ่านการทดลองฝึกปฏิบัติ และการเรียนรู้ผ่านโจทย์สถานการณ์จริง ผลจากการจัดกิจกรรมช่วยให้บุคลากรสายสนับสนุนวิชาการเกิดการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้เพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ซึ่งจากการดำเนินการประเมิน 3 ครั้งในระหว่างการเรียนรู้พบว่า คะแนนสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ $F(1.483, 44.487) = 898.729, p < .001$ โดยสมรรถนะดังกล่าวเพิ่มขึ้นในทุกรอบของการประเมิน

คำสำคัญ สมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ การเรียนรู้ผ่านประสบการณ์
การประเมินเพื่อการพัฒนา กลยุทธ์การให้ข้อมูลย้อนกลับ

¹ นักศึกษาปริญญาเอก สาขาการวิจัยและประเมินทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

² ผู้ช่วยศาสตราจารย์ ภาควิชาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

¹ Doctoral Degree's student, Educational Research and Evaluation Program, Faculty of Education, Kasetsart University

² Assistant Professor Department of Education Faculty of Education Kasetsart University

* Corresponding Author E-mail: wannachorn.cha@kmutt.ac.th

Abstract

This research aimed to develop experiential learning activities integrated with feedback to enhance the learning design competency of academic support staff at King Mongkut's University of Technology Thonburi, and to investigate the competency development of those who participated in the program. The sample consisted of 31 academic support staff selected through purposive sampling. The research instruments included activity worksheets and behavioral checklist forms. Data were analyzed using mean, standard deviation, one-way repeated measures ANOVA, and effect size. The developed learning activities comprised three components: (1) self-directed learning of core concepts, (2) practice-based learning, and (3) real-world problem-based scenarios. The implementation significantly enhanced participants' learning design competency ($p < .001$). Repeated measures ANOVA showed a statistically significant difference in competency scores across three time points, $F(1.483, 44.487) = 898.729, p < .001$. Competency scores increased in each successive assessment.

Keywords: learning design competency, experiential learning, formative assessment, feedback strategies

บทนำ

สมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้เป็นหนึ่งในสมรรถนะเฉพาะงานของบุคลากรที่รับผิดชอบด้านการพัฒนาการศึกษา (Adelsberger et al., 2008; Branch, 2009) ครอบคลุมทั้งการจัดการศึกษาในระดับหลักสูตร รายวิชา ตลอดจนโปรแกรมอบรมต่าง ๆ (Spector et al., 2009) โดยการออกแบบการจัดการเรียนรู้ต้องคำนึงถึงความสอดคล้องกับความต้องการของผู้มีส่วนได้ส่วนเสียเป็นสำคัญ (Ritzhaupt et al., 2021) ปัจจุบันแนวคิดการจัดการศึกษาแบบมุ่งผลลัพธ์ (Outcome-based Education, OBE) เป็นแนวคิดที่ได้รับความนิยมและถูกนำมาใช้เป็นกรอบในการวางแผนการจัดการศึกษาโดยเฉพาะในระดับอุดมศึกษา ทั้งในสหรัฐอเมริกา ยุโรป และเอเชีย (Rao, 2020; Spady, 1988) โดย Spady (1988) กล่าวว่า การจัดการศึกษาแบบมุ่งผลลัพธ์ให้ความสำคัญกับความสามารถของผู้เรียนที่จะสามารถปฏิบัติได้หลังสำเร็จการเรียนรู้ จากการศึกษาเอกสารและงานวิจัยสามารถสรุปสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ออกได้เป็น 4 สมรรถนะหลัก ได้แก่ 1) การวิเคราะห์ความต้องการของผู้มีส่วนได้ส่วนเสีย 2) การกำหนดและเขียนผลลัพธ์การเรียนรู้ที่คาดหวัง 3) การออกแบบกิจกรรมการเรียนรู้ และ 4) การออกแบบกระบวนการวัดและประเมิน โดยคำนึงถึงความสอดคล้องไปในทิศทางเดียวกันขององค์ประกอบในข้างต้นอย่างเป็นระบบ (Biggs & Tang, 2011; Dick & Carey, 1996; Spady, 1988)

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรีให้ความสำคัญกับการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้แก่บุคลากรสายสนับสนุนวิชาการ ซึ่งเป็นผู้เรียนวัยผู้ใหญ่ (adult learner)

โดย Knowles et al. (2005) กล่าวว่า ในการจัดการเรียนรู้สำหรับผู้เรียนวัยผู้ใหญ่ต้องคำนึงถึงความสะดวกคล่องระหว่างสิ่งที่เรีเรียนรู้กับลักษณะงานที่รับผิดชอบ การได้ลงมือปฏิบัติจริง และควรให้ข้อเสนอแนะในระหว่างการเรียนรู้ ซึ่งเป็นลักษณะสำคัญของการเรียนรู้ผ่านประสบการณ์ โดย Knowles (1973); Sisselman-Borgia & Torino (2017) ได้เสนอว่าการเรียนรู้ผ่านประสบการณ์เป็นรูปแบบการเรียนรู้ที่มีความเหมาะสมสำหรับผู้เรียนวัยผู้ใหญ่ ส่วน Hattie & Timperley (2007) เสนอว่าต้องมีกลไกการให้ข้อมูลย้อนกลับแก่ผู้เรียนในระหว่างการเรียนรู้เพื่อช่วยปิดช่องว่างความสามารถที่เป็นอยู่กับที่คาดหวัง สอดคล้องกับ Bin Mubayrik (2020); Black & Wiliam (2009); Schildkamp et al. (2020) ที่กล่าวว่า การให้ข้อมูลย้อนกลับช่วยสนับสนุนให้ผู้เรียนบรรลุเป้าหมายการเรียนรู้ ในขณะที่ Shute (2008) กล่าวว่า การให้ข้อมูลย้อนกลับกับการประเมินต้องดำเนินควบคู่กันอย่างเป็นระบบ ทั้งนี้ Fuentes-Cimma et al. (2024) กล่าวว่า ในการให้ข้อมูลย้อนกลับควรคำนึงถึงความสะดวกคล่องกับเป้าหมายการเรียนรู้ที่กำหนดไว้ ส่วน Lipnevich & Panadero (2021) ได้สังเคราะห์งานวิจัยเกี่ยวกับการให้ข้อมูลย้อนกลับ โดยผลการศึกษายืนยันว่าการให้ผู้สอนให้ข้อมูลย้อนกลับแก่ผู้เรียน จำนวน 3 รอบ “Three-cycle model feedback” ช่วยให้ผู้เรียนเกิดการพัฒนาความสามารถเพิ่มขึ้นอย่างต่อเนื่อง ซึ่ง Sadler (1989, as cited in Lipnevich & Panadero, 2021) ได้เสนอว่าผู้สอนต้องให้ข้อมูลอย่างเพียงพอและพอเหมาะสำหรับการปิดช่องว่างความสามารถ สอดคล้องกับผลการศึกษาของ Bin Mubayrik (2020) ที่กล่าวว่า การให้ข้อมูลย้อนกลับช่วยทำให้คะแนนของผู้เรียนเพิ่มสูงขึ้นจนบรรลุตามผลลัพธ์และเกณฑ์ที่กำหนด ในขณะที่ผลการศึกษาของ Wisniewski et al. (2019) พบว่าการให้ข้อมูลย้อนกลับมีอิทธิพลต่อการพัฒนาผลลัพธ์การเรียนรู้โดยเฉพาะด้านความรู้และด้านทักษะ

Hu et al. (2021); Molloy & Boud (2014); Narciss (2013); Shute (2008) กล่าวว่า การให้ข้อมูลย้อนกลับในระหว่างการเรียนรู้มีอิทธิพลต่อการพัฒนาผู้เรียนให้บรรลุตามผลลัพธ์การเรียนรู้ที่คาดหวัง โดยได้เสนอกลยุทธ์สำหรับการให้ข้อมูลย้อนกลับ ดังนี้ 1) การตรวจสอบความถูกต้อง 2) การระบุคำตอบที่ถูกต้อง 3) การทดลองทำซ้ำ 4) การแจ้งข้อผิดพลาด 5) การขยายความเพิ่มเติม 6) การให้ข้อเสนอแนะในประเด็นที่เกี่ยวข้อง 7) การให้ข้อเสนอแนะที่มีความเฉพาะเจาะจง 8) การให้คำแนะนำ/ ข้อเสนอแนะ/ ข้อความแจ้ง 9) การระบุข้อบกพร่อง/ ความเข้าใจผิด ส่วน Mason & Bruning (2001, as cited in Shute, 2008) กล่าวว่า หากโจทย์หรือแบบฝึกหัดมีความซับซ้อนและผู้เรียนมีประสบการณ์ก่อนหน้ามาบ้างแล้ว ผู้สอนสามารถเลือกใช้วิธีการให้ผู้เรียนได้ทดลองทำซ้ำและการให้ข้อมูลย้อนกลับแบบเฉพาะเจาะจงในประเด็นที่เกี่ยวข้องควบคู่กันได้ ในขณะเดียวกันหากผู้เรียนยังไม่เคยมีประสบการณ์อาจจะต้องใช้กลยุทธ์หลากหลายวิธีการร่วมกัน ส่วน Morris & Chikwa (2016) ได้ศึกษาวิจัยเกี่ยวกับช่องทางการให้ข้อมูลย้อนกลับพบว่าผู้เรียนมีความพึงพอใจต่อการให้ข้อมูลย้อนกลับด้วยวิธีการอัดเสียง (audio) มากกว่าวิธีการเขียน (written) แต่ไม่พบความแตกต่างของคะแนนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในขณะที่ผลการศึกษาของ O'Donovan et al. (2016, as cited in Carless & Young, 2024) กล่าวว่า การให้ข้อมูลย้อนกลับแบบรายบุคคลด้วยวิธีการพูดคุย (oral feedback to individuals) เกิดประสิทธิภาพต่อผู้เรียนมากกว่าวิธีการเขียนเพียงอย่างเดียวอย่างมีนัยสำคัญทางสถิติ

ปัจจุบันบุคลากรสายสนับสนุนวิชาการมีบทบาทเพิ่มมากขึ้นในการสนับสนุนการจัดการเรียนการสอน ซึ่งส่วนใหญ่ยังขาดสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนการสอน มากกว่านั้นกระบวนการจัดการศึกษาในยุค 4.0 มีความแตกต่างไปจากรูปแบบหรือวิธีการแบบดั้งเดิม ดังนั้นจึงมีความจำเป็นอย่างเร่งด่วนในการยกระดับสมรรถนะในข้างต้นเพื่อเพิ่มประสิทธิภาพในการปฏิบัติงานสนับสนุนด้านวิชาการ อย่างไรก็ตามจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องพบว่าการจัดการเรียนรู้สำหรับผู้เรียนวัยผู้ใหญ่ควรมีรูปแบบและวิธีการเรียนรู้ที่หลากหลาย ไม่เฉพาะแค่การบรรยาย มีส่งเสริมให้ผู้เรียนได้ฝึกลงมือปฏิบัติจริง สามารถบรรลุตามผลลัพธ์การเรียนรู้ในระยะเวลาที่เหมาะสม ตอบโจทย์ความต้องการของผู้เรียน ตลอดจนควรนำกระบวนการประเมินเพื่อการพัฒนาไปใช้เป็นส่วนหนึ่งของกระบวนการจัดการเรียนรู้ มีกระบวนการให้ข้อมูลย้อนกลับในระหว่างการเรียนรู้อย่างเป็นระบบ นอกจากนี้ จากการศึกษาเอกสารและงานวิจัยยังพบว่าถึงแม้ว่าจะมีงานวิจัยที่กล่าวถึงประโยชน์และความสำคัญของการให้ข้อมูลย้อนกลับและถูกนำมาใช้เป็นส่วนหนึ่งของการจัดการศึกษา แต่ยังไม่ปรากฏว่ามีการนำกระบวนการให้ข้อมูลย้อนกลับมาใช้ในการจัดการศึกษา โดยเฉพาะสำหรับผู้เรียนวัยผู้ใหญ่เป็นรูปธรรมมากนัก อีกทั้งยังไม่ทราบแน่ชัดว่ากลยุทธ์การให้ข้อมูลย้อนกลับใดบ้างที่มีความเหมาะสมและสามารถยกระดับสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ไปสู่ระดับที่คาดหวังตามเกณฑ์หรือมาตรฐานที่กำหนดไว้ จึงเป็นที่น่าสนใจเป็นอย่างยิ่งว่าภายใต้การพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้แก่ผู้เรียนวัยใหญ่นั้นควรดำเนินการอย่างไร ตลอดจนกลยุทธ์การให้ข้อมูลย้อนกลับใดบ้างที่สามารถช่วยยกระดับความสามารถให้บรรลุตามผลลัพธ์การเรียนรู้ที่คาดหวังได้

ความมุ่งหมายของการวิจัย

1. เพื่อพัฒนากิจกรรมการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับสำหรับการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
2. เพื่อศึกษาผลการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรีที่ได้รับการจัดกิจกรรมการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับ

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย คือ บุคลากรสายสนับสนุนวิชาการของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ตำแหน่งงานนักพัฒนาการศึกษา นักเทคโนโลยีการศึกษา และนักบริการการศึกษา จำนวน 71 คน

กลุ่มตัวอย่างในการวิจัย คือ ได้มาจากการเลือกแบบเฉพาะเจาะจง โดยพิจารณาจากประสบการณ์ทำงานที่ไม่ต่ำกว่า 3 ปี มีความสมัครใจเข้าร่วมโครงการ และอยู่ในระหว่างการเตรียมความพร้อมเพื่อเลื่อนระดับตำแหน่งงาน จำนวน 31 คน โดยเป็นนักพัฒนาการศึกษา จำนวน 22 คน นักเทคโนโลยีการศึกษา จำนวน 8 คน และนักบริการการศึกษา จำนวน 1 คน

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย คือ ชุดใบงานกิจกรรมสำหรับการออกแบบกระบวนการจัดการเรียนรู้ในระดับหน่วยการเรียนรู้ และแบบตรวจสอบรายการพฤติกรรมในระหว่างการเรียนรู้ ดังนี้

2.1 ชุดใบงานกิจกรรม คือ ชุดใบงานสำหรับฝึกออกแบบกระบวนการจัดการเรียนรู้เพื่อมุ่งประเมินในส่วนของความรู้และทักษะของสมรรถนะการออกแบบกิจกรรมการเรียนรู้ 4 สมรรถนะหลัก ได้แก่ 1) การวิเคราะห์ความต้องการของผู้มีส่วนได้ส่วนเสีย 2) การกำหนดผลลัพธ์การเรียนรู้ 3) การออกแบบกิจกรรมการเรียนรู้ และ 4) การออกแบบกระบวนการวัดและประเมิน โดยมีการกำหนดเกณฑ์การให้คะแนนแบบรูบริก (scoring rubrics)

2.2 แบบตรวจสอบรายการพฤติกรรม คือ แบบประเมินเพื่อตรวจสอบพฤติกรรมการแสดงออกในระหว่างการเรียนรู้ผ่านประสบการณ์ เพื่อมุ่งประเมินในส่วนของทัศนคติของสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ 4 สมรรถนะหลัก ได้แก่ 1) การวิเคราะห์ความต้องการของผู้มีส่วนได้ส่วนเสีย 2) การกำหนดผลลัพธ์การเรียนรู้ 3) การออกแบบกิจกรรมการเรียนรู้ และ 4) การออกแบบกระบวนการวัดและประเมิน โดยมีการประเมิน 2 ระดับ คือ ปฏิบัติ/ แสดงให้เห็น กับไม่ได้ปฏิบัติ/ ไม่แสดงให้เห็น

ทั้งนี้ ผู้เชี่ยวชาญ 3 ท่าน ได้กำหนดค่าน้ำหนักคะแนนในการประเมินสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ตามองค์ประกอบด้านความรู้ ด้านทักษะ และด้านทัศนคติ ดังแสดงในตาราง 1

ตาราง 1 ค่าน้ำหนักคะแนนจำแนกตามสมรรถนะหลักและองค์ประกอบด้านความรู้ ทักษะ และทัศนคติ

สมรรถนะ ด้านการออกแบบกระบวนการจัดการเรียนรู้	คะแนน (100)	คะแนนจำแนกตามองค์ประกอบสมรรถนะ		
		ด้านความรู้	ด้านทักษะ	ด้านทัศนคติ
		20%	50%	30%
1. การวิเคราะห์ความต้องการของผู้มีส่วนได้ส่วนเสีย	30	6	15	9
2. การกำหนดผลลัพธ์การเรียนรู้	10	2	5	3
3. การออกแบบกิจกรรมการเรียนรู้	30	6	15	9
4. การออกแบบกระบวนการวัดและประเมิน	30	6	15	9

เกณฑ์การประเมินเพื่อตัดสินผลการพัฒนาสมรรถนะ ผู้วิจัยได้ออกแบบและพัฒนาเกณฑ์การประเมินสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ โดยมีลักษณะเป็นรูบริกแบบองค์รวม (Holistic Rubrics) ที่ประกอบด้วย ระดับสมรรถนะ ช่วงคะแนน และคำอธิบายระดับสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้เพื่อใช้เป็นเกณฑ์ในการประเมินและตัดสินสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ ดังแสดงในตาราง 2

ตาราง 2 ระดับสมรรถนะ ช่วงคะแนน และคำอธิบายระดับสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้

ระดับ	คะแนน	คำอธิบายระดับสมรรถนะ
A	80-100	สามารถออกแบบองค์ประกอบของหน่วยการเรียนรู้ได้ถูกต้องและครบถ้วนตามแนวคิดและหลักการสำคัญที่เกี่ยวข้องในทุกองค์ประกอบของหน่วยการเรียนรู้ได้ด้วยตนเองเป็นหลัก ตลอดจนแสดงให้เห็นถึงการตระหนักถึงผู้เรียนกลุ่มเป้าหมายหรือผู้มีส่วนได้ส่วนเสียของหน่วยการเรียนรู้
B*	70-79	สามารถออกแบบหน่วยการเรียนรู้ได้ด้วยตนเองเป็นหลัก แต่ยังขาดความถูกต้องและไม่สอดคล้องกับแนวคิด หลักการ ทฤษฎีที่เกี่ยวข้องเพียงเล็กน้อยในบางองค์ประกอบ ตระหนักถึงผู้เรียนกลุ่มเป้าหมายหรือผู้มีส่วนได้ส่วนเสียของหน่วยการเรียนรู้ ต้องได้รับคำแนะนำจากโค้ชที่เลี้ยงเพียงเล็กน้อย 1-2 ประเด็น
C	60-69	สามารถออกแบบหน่วยการเรียนรู้ได้ แต่ยังขาดความถูกต้องและไม่สอดคล้องกับแนวคิด หลักการที่เกี่ยวข้อง ขาดการตระหนักถึงผู้เรียนกลุ่มเป้าหมายหรือผู้มีส่วนได้ส่วนเสียของหน่วยการเรียนรู้ ต้องได้รับคำแนะนำจากโค้ชที่เลี้ยงมากกว่า 3-4 ประเด็น
F	<59	สามารถออกแบบหน่วยการเรียนรู้ได้ แต่ยังขาดความถูกต้องและไม่สอดคล้องกับแนวคิด หลักการที่เกี่ยวข้องในทุกองค์ประกอบของการออกแบบการจัดการเรียนรู้ ตลอดจนขาดการตระหนักถึงผู้เรียนกลุ่มเป้าหมายหรือผู้มีส่วนได้ส่วนเสียของหน่วยการเรียนรู้ ต้องได้รับคำแนะนำจากโค้ชที่เลี้ยงอย่างใกล้ชิดในทุกขั้นตอน

หมายเหตุ: * คือ เกณฑ์มาตรฐานที่คาดหวัง

ผู้วิจัยนำเครื่องมือไปวิเคราะห์คุณภาพ ดังนี้

1. ชุดใบงานกิจกรรม ดำเนินการหาคุณภาพชุดใบงานกิจกรรม ดังนี้

1.1) การตรวจสอบความตรง ใช้การวิเคราะห์ความตรงเชิงเนื้อหา (content validity) โดยให้ผู้เชี่ยวชาญ 3 ท่าน ซึ่งเป็นผู้เชี่ยวชาญด้านการออกแบบกิจกรรมการเรียนรู้ตามแนวคิดการจัดการศึกษาแบบมุ่งผลลัพธ์ (OBE) ตลอดจนด้านการวัดและประเมิน พิจารณาความตรงเชิงเนื้อหาของใบงานกิจกรรม ประกอบด้วย เนื้อหาสาระสำคัญตามแนวคิด หลักการ และทฤษฎีที่เกี่ยวข้อง คำนวณน้ำหนักคะแนน คำอธิบายชุดใบงานกิจกรรม ลำดับขั้นการเรียนรู้ เกณฑ์การประเมิน ขั้นตอนการประเมิน ช่วงเวลาการประเมิน และการรายงานผลการประเมิน โดยผลการพิจารณาพบว่าชุดใบงานกิจกรรมมีค่า IOC ระหว่าง 0.67 – 1.00 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้โดยมีค่ามากกว่า 0.5 (Turner & Carlson, 2003)

1.2) การตรวจสอบคุณภาพเชิงประจักษ์ โดยให้บุคลากรสายสนับสนุนวิชาการจำนวน 34 คนที่ไม่ใช่กลุ่มตัวอย่างในการวิจัยทดลองใช้ชุดใบงานกิจกรรม แล้วนำผลการทดลองใช้มาวิเคราะห์ ดังนี้

1.2.1) ความยากของชุดใบงานกิจกรรม (difficulty) ทำการแบ่งคะแนนออกเป็น 2 กลุ่ม คือ กลุ่มเก่ง (กลุ่มสูง) และกลุ่มอ่อน (กลุ่มต่ำ) โดยใช้เทคนิค 25% โดยพบว่าชุดใบงานกิจกรรมมีค่าความยากระหว่าง 0.38 – 0.54 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ โดยมีค่าระหว่าง 0.20 – 0.80 (Whitney & Sabers, 1970)

1.2.2) อำนาจจำแนก (discrimination) ผลการวิเคราะห์ค่าอำนาจจำแนกพบว่าชุดใบงานกิจกรรมมีค่าอำนาจจำแนกระหว่าง 0.21–0.43 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้โดยมีค่า 0.20 ขึ้นไป (Whitney & Sabers, 1970)

1.2.3) ความเชื่อมั่น (reliability) วิเคราะห์ความเชื่อมั่นของชุดใบงานกิจกรรมโดยใช้วิธีการทดสอบซ้ำ (test-retest method) โดยทิ้งช่วงระยะเวลาของการสอบครั้งที่ 1 และ 2 จำนวน 30 วัน โดยใช้ผู้ตรวจซึ่งมีความเชี่ยวชาญด้านการออกแบบกระบวนการจัดการเรียนรู้ จำนวน 4 ท่าน แล้วนำคะแนนที่ได้จากการสอบทั้ง 2 ครั้ง มาหาความเชื่อมั่นโดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน พบว่ามีค่าความเชื่อมั่นเท่ากับ 0.94 โดยมีค่ามากกว่า 0.70 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ (Taber, 2017)

2. แบบตรวจสอบรายการพฤติกรรม ดำเนินการหาคุณภาพแบบตรวจสอบรายการพฤติกรรมด้วยการตรวจสอบความตรงเชิงเนื้อหา (content validity) โดยให้ผู้เชี่ยวชาญ 3 ท่านวิเคราะห์ความตรงเชิงโครงสร้างโดยผู้เชี่ยวชาญจำนวน 3 ท่าน ตรวจสอบเนื้อหาของแบบตรวจสอบรายการพฤติกรรมว่ามีความครอบคลุมเนื้อหาหรือไม่ ซึ่งพบว่ามีค่า IOC ระหว่าง 0.67 – 1.00 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้เนื่องจากมีค่า IOC มากกว่า 0.5 (Turner & Carlson, 2003)

3. การเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการประชุมเชิงปฏิบัติการสำหรับเตรียมความพร้อมในการทำหน้าที่เป็นผู้ประเมินแก่โค้ชที่เลี้ยงประจำกลุ่ม (coaches) จำนวน 4 ท่าน โดยเป็นผู้ที่มีความรู้ความสามารถด้านการออกแบบกระบวนการจัดการเรียนการสอน ซึ่งทั้งหมดผ่านการรับรองความสามารถมาตรฐานวิชาชีพเพื่อการสอนและการสนับสนุนการเรียนรู้ที่มีคุณภาพของประเทศไทยในระดับ senior fellow โดยดำเนินการจัดอบรมเชิงปฏิบัติการ จำนวน 2 ครั้ง เพื่อชี้แจงและทำความเข้าใจเกี่ยวกับรายละเอียดของคู่มือ ขั้นตอนการประเมิน การใช้เครื่องมือ เกณฑ์การประเมิน (scoring rubrics) ในระหว่างการอบรมฯ ผู้วิจัยได้นำตัวอย่างผลงานมาให้โค้ชทดลองประเมิน ผลปรากฏว่าโค้ชทั้ง 4 คน ประเมินคุณภาพของชิ้นงานสอดคล้องไปในทิศทางเดียวกัน ส่วนผู้สนับสนุนการเรียนรู้ประจำกลุ่ม (facilitator) ผู้วิจัยได้ประชุมชี้แจงและทำความเข้าใจเกี่ยวกับกระบวนการบันทึกข้อมูลต่าง ๆ ในขณะการเก็บข้อมูลนั้นผู้วิจัยสร้างโฟลเดอร์แบบออนไลน์สำหรับการรับส่งชุดใบงานกิจกรรม ใน google share drive โดยมีการแชร์โฟลเดอร์เอกสารประกอบการประเมิน ได้แก่ คู่มือการประเมิน ตารางสำเร็จรูปสำหรับบันทึกคะแนนแก่โค้ชที่เลี้ยงและผู้ช่วยสนับสนุนการเรียนรู้ประจำกลุ่ม ซึ่งดำเนินการเก็บรวบรวมผลการประเมินสมรรถนะ จำนวน 3 ครั้ง เป็นไปตามรอบการประเมิน ทั้งนี้ในระหว่างนำเสนองานและติดตามความก้าวหน้าระหว่างโค้ชที่เลี้ยงประจำกลุ่มกับบุคลากรสายสนับสนุนวิชาการทั้งแบบออนไลน์และแบบเผชิญหน้า (online & face-to-face) พร้อมทั้งบันทึกช่วงเวลา ช่องทาง และกลยุทธ์การให้ข้อมูลย้อนกลับในระหว่างการเรียนรู้

4. การวิเคราะห์ข้อมูล

วิเคราะห์ข้อมูลโดยใช้ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน การวิเคราะห์ความแปรปรวนแบบวัดซ้ำทางเดียว (One-way repeated measure ANOVA) และการวิเคราะห์ขนาดอิทธิพล (effect size) สำหรับเปรียบเทียบผลการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้จำแนกตามรอบการประเมิน และการให้ข้อมูลย้อนกลับ

ผลการวิจัย

การนำเสนอผลการวิจัยครั้งนี้จะเสนอตามลำดับดังต่อไปนี้

1. รูปแบบกิจกรรมการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับ

การออกแบบและพัฒนาแบบการจัดกิจกรรมการเรียนรู้ผ่านประสบการณ์ตามแนวคิดการเรียนรู้แบบ 10: 20: 70 สามารถสรุปรายละเอียดได้ดังนี้

1.1 การศึกษาเรียนรู้ด้วยตนเอง (self-paced Learning: 10%) เป็นการศึกษาเรียนรู้เนื้อหาสาระสำคัญ หลักการ แนวคิด และทฤษฎีที่เกี่ยวข้องกับการวิเคราะห์ความต้องการของผู้มีส่วนได้ส่วนเสียของหน่วยการเรียนรู้ การกำหนดผลลัพธ์การเรียนรู้ การออกแบบกิจกรรมการเรียนรู้ การวัดและประเมินผล ผ่านสื่อวีดิทัศน์และเอกสารสรุป ในระบบการจัดการเรียนรู้ของมหาวิทยาลัย “KMUTT LEB2” ระยะเวลา 45 วัน พร้อมทั้งการทำแบบทดสอบและแบบประเมินท้ายบทเรียน

1.2 การศึกษาเรียนรู้ผ่านการอบรมเชิงปฏิบัติการ (workshop: 20%) เป็นการศึกษาเรียนรู้ผ่านการทดลองฝึกปฏิบัติออกแบบหน่วยการเรียนรู้จากโจทย์สถานการณ์จำลอง โดยใช้ระยะเวลาในการเรียนรู้ 2 เดือน (2 สัปดาห์/ สมรรถนะหลัก) โดยมีกระบวนการและขั้นตอนการเรียนรู้ ดังนี้

1.2.1) การดำเนินกิจกรรมการอบรมเชิงปฏิบัติการ ใช้ระยะเวลาในการเรียนรู้รวม 84 ชั่วโมง ประกอบด้วย ระยะเวลาเรียนรู้และฝึกปฏิบัติในห้องอบรมเชิงปฏิบัติการจำนวน 6 ชั่วโมง/ สมรรถนะหลัก รวมระยะเวลาทั้งหมด 24 ชั่วโมง จากนั้นบุคลากรสายสนับสนุนวิชาการดำเนินการออกแบบและพัฒนาหน่วยการเรียนรู้จากโจทย์สถานการณ์จำลองในชุดใบงานกิจกรรม โดยใช้ระยะเวลาประมาณ 15 ชั่วโมง/ สมรรถนะหลัก รวมระยะเวลา 60 ชั่วโมง พร้อมทั้งส่งชิ้นงานในระบบบริหารการเรียนรู้ (KMUTT LEB2) เพื่อให้โค้ชพี่เลี้ยงประจำกลุ่มประเมินตามเกณฑ์ที่กำหนดไว้ พร้อมทั้งให้ข้อมูลย้อนกลับตามกระบวนการที่กำหนดไว้จนบรรลุตามผลลัพธ์การเรียนรู้ที่คาดหวังหรือเกณฑ์มาตรฐานที่กำหนดไว้ และการทำแบบสำรวจความคิดเห็น

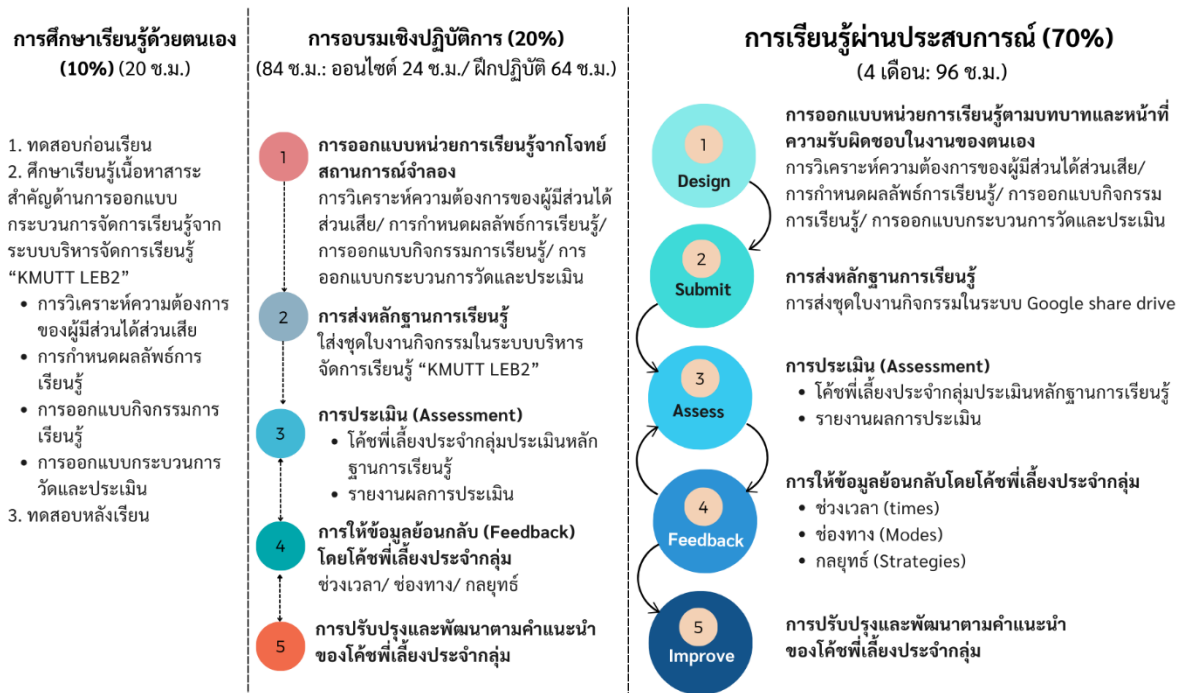
1.2.2) การรวบรวมและวิเคราะห์ข้อมูลการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้พร้อมทั้งรายงานผลการพัฒนาไปยังบุคลากรสายสนับสนุนวิชาการและโค้ชพี่เลี้ยงประจำกลุ่มเพื่อติดตามความก้าวหน้า

1.3 การศึกษาเรียนรู้ผ่านประสบการณ์จริง (experiential learning: 70%) เป็นการศึกษาเรียนรู้ผ่านการปฏิบัติจริงด้วยการออกแบบและพัฒนาหน่วยการเรียนรู้ที่สอดคล้องกับภาระงาน และบทบาทหน้าที่ของบุคลากรสายสนับสนุนวิชาการ โดยใช้ระยะเวลาในการเรียนรู้ จำนวน 4 เดือน (1 เดือน/ สมรรถนะหลัก) โดยมีกระบวนการและขั้นตอนการเรียนรู้ ดังนี้

1.3.1) บุคลากรสายสนับสนุนวิชาการดำเนินการออกแบบและพัฒนาหน่วยการเรียนรู้ โดยใช้ระยะเวลาในการออกแบบและพัฒนาองค์ประกอบของหน่วยการเรียนรู้ เป็นระยะเวลา 4 เดือน ครอบคลุมการพัฒนาทั้ง 4 สมรรถนะหลัก ดังกล่าวในข้างต้น

1.3.2) บุคลากรสายสนับสนุนวิชาการดำเนินการส่งชุดใบงานกิจกรรมใน google share drive นำเสนอความก้าวหน้า โค้ชพี่เลี้ยงประจำกลุ่มประเมิน พร้อมทั้งให้ข้อมูลย้อนกลับตามกระบวนการที่กำหนดไว้

1.3.3) การรวบรวมและวิเคราะห์ข้อมูลการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้พร้อมทั้งรายงานผลการพัฒนาไปยังบุคลากรสายสนับสนุนวิชาการและโค้ชพี่เลี้ยง เพื่อติดตามความก้าวหน้า ดังแสดงในภาพ 1



ภาพ 1 แสดงรูปแบบการเรียนรู้สำหรับการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้

2. ผลการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรีโดยใช้การเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับ

2.1 คะแนนสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรีโดยใช้การเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับ

ในช่วงการเรียนรู้ผ่านประสบการณ์โค้ชพี่เลี้ยงดำเนินการประเมินสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ จำนวน 3 ครั้ง และให้ข้อมูลย้อนกลับ จำนวน 2 รอบ สามารถสรุปได้ดังนี้

ผลการศึกษาในภาพรวม พบว่าบุคลากรสายสนับสนุนวิชาการได้รับคะแนนสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้เฉลี่ยเพิ่มขึ้นทุกรอบของการประเมิน ($\bar{x}_{t1} = 32.85$, $SD = 1.788$; $\bar{x}_{t2} = 58.98$, $SD = 2.225$; $\bar{x}_{t3} = 88.25$, $SD = 1.192$) หากพิจารณารายสมรรถนะหลักพบว่าบุคลากรสายสนับสนุนวิชาการมีพัฒนาการในทุกสมรรถนะหลักเพิ่มขึ้นในทุกรอบของการประเมิน สรุปได้ดังนี้ 1) ด้านการวิเคราะห์ความต้องการของผู้มีส่วนได้ส่วนเสียของหน่วยการเรียนรู้ ($\bar{x}_{t1} = 11.16$, $SD = 2.207$; $\bar{x}_{t2} = 16.51$, $SD = 2.953$; $\bar{x}_{t3} = 26.16$, $SD = 2.114$) 2) ด้านการกำหนดผลลัพธ์การเรียนรู้ของหน่วยการเรียนรู้ ($\bar{x}_{t1} = 3.22$, $SD = 0.616$; $\bar{x}_{t2} = 5.35$, $SD = 0.660$; $\bar{x}_{t3} = 8.45$, $SD = 8.50$) 3) การออกแบบกิจกรรมการเรียนรู้ของหน่วย

การเรียนรู้ ($\bar{x}_{t1} = 9.54$, SD = 2.126; $\bar{x}_{t2} = 16.58$, SD = 2.884; $\bar{x}_{t3} = 26.61$, SD = 2.231) และ 4) ด้านการออกแบบกระบวนการวัดและประเมินผลของหน่วยการเรียนรู้ ($\bar{x}_{t1} = 8.93$, SD = 2.205; $\bar{x}_{t2} = 20.54$, SD = 2.406; $\bar{x}_{t3} = 27.03$, SD = 2.456)

นอกจากนั้น เมื่อพิจารณาระดับสมรรถนะพบว่า การประเมินในครั้งที่ 1 บุคลากรสายสนับสนุนวิชาการมีสมรรถนะอยู่ในระดับ C ในทุกสมรรถนะ ในขณะที่ การประเมินในครั้งที่ 2 บุคลากรสายสนับสนุนวิชาการ มีสมรรถนะอยู่ในระดับ B เพียงสมรรถนะเดียว คือ การออกแบบการวัดและประเมินผล ส่วนสมรรถนะด้านอื่นมีผลการประเมินอยู่ในระดับ C ในขณะที่ผลการประเมินในครั้งที่ 3 บุคลากรสายสนับสนุนวิชาการมีการพัฒนาสมรรถนะที่ดีขึ้นจนอยู่ในระดับ A ทุกสมรรถนะ ดังแสดงในตาราง 3

ตาราง 3 คะแนนสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้จำแนกตามสมรรถนะหลักและรอบ

การเรียนรู้		การประเมิน								
Learning Design Competency	Full Marks	Assessment Time								
		Time 1			Time 2			Time 3		
		$\bar{x}$ (K+S+A)	SD	Level	$\bar{x}$ (K+S+A)	SD	Level	$\bar{x}$ (K+S+A)	SD	Level
1. SNA	30	11.16	2.207	C	16.51	2.953	C	26.16	2.114	A
2. IELO	10	3.22	0.616	C	5.35	0.660	C	8.45	0.850	A
3. LAD	30	9.54	2.126	C	16.58	2.884	C	26.61	2.231	A
4. APD	30	8.93	2.205	C	20.54	2.406	B	27.03	2.456	A
Total	100	32.85	1.788	C	58.98	2.225	C	88.25	1.912	A

Note: ระดับที่คาดหวัง คือ B
SNA = การวิเคราะห์ความต้องการของผู้มีส่วนได้ส่วนเสีย, IELO = การกำหนดผลลัพธ์การเรียนรู้ที่คาดหวัง, LAD = การออกแบบกิจกรรมการเรียนรู้, APD = การออกแบบกระบวนการวัดและประเมิน

2.2 ผลการเปรียบเทียบผลการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี โดยใช้การเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับจำแนกตามรอบการประเมิน

ผลการศึกษาในภาพรวมจากการประเมินซ้ำ 3 ครั้ง พบว่าบุคลากรสายสนับสนุนวิชาการมีการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้อย่างต่อเนื่อง โดยผลจากการทดสอบของ Huynh-Feldt = .741 ซึ่งแสดงว่าสมรรถนะจากการประเมิน ทั้ง 3 ครั้ง มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ $F(1.483, 44.487) = 898.729$, $p < .001$ ทั้งนี้ เมื่อเปรียบเทียบคะแนนรายคู่ด้วยวิธีของ Bonferroni พบว่า 1) คะแนนครั้งที่ 2 > ครั้งที่ 1 ($d, t_2 - t_1 = 26.12$) 2) คะแนนครั้งที่ 3 > ครั้งที่ 1 ($d, t_3 - t_1 = 55.38$) และ 3) คะแนนครั้งที่ 3 > ครั้งที่ 2 ($d, t_3 - t_2 = 29.25$) โดยคะแนนครั้งที่ 1 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 3.63 คะแนนครั้งที่ 3 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 5.92 และคะแนนครั้งที่ 3 กับ ครั้งที่ 1 มีค่าขนาดอิทธิพล = 6.05 ซึ่งแสดงให้เห็นว่าสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้เพิ่มขึ้นอย่างต่อเนื่อง ดังแสดงในตาราง 4

หากพิจารณารายสมรรถนะหลัก สามารถสรุปได้ดังนี้

1) ด้านการวิเคราะห์ความต้องการของผู้มีส่วนได้ส่วนเสียของหน่วยการเรียนรู้ โดยผลจากการทดสอบของ Huynh-Feldt = 1 ซึ่งแสดงว่าสมรรถนะจากการประเมิน ทั้ง 3 ครั้ง มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ $F(2, 60) = 608.144, p < .001$ ทั้งนี้ เมื่อเปรียบเทียบคะแนนรายคู่ด้วยวิธีของ Bonferroni พบว่า 1) คะแนนครั้งที่ 2 > ครั้งที่ 1 ($d, t_2 - t_1 = 5.53$) 2) คะแนนครั้งที่ 3 > ครั้งที่ 1 ($d, t_3 - t_1 = 15.00$) และ 3) คะแนนครั้งที่ 3 > ครั้งที่ 2 ($d, t_3 - t_2 = 9.64$) โดยคะแนนครั้งที่ 1 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 2.28 คะแนนครั้งที่ 3 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 4.37 และคะแนนครั้งที่ 3 กับ ครั้งที่ 1 มีค่าขนาดอิทธิพล = 5.71 ซึ่งแสดงให้เห็นว่าสมรรถนะด้านการวิเคราะห์ความต้องการของผู้มีส่วนได้ส่วนเสียของหน่วยการเรียนรู้เพิ่มขึ้นอย่างต่อเนื่อง

2) ด้านการกำหนดผลลัพธ์การเรียนรู้ของหน่วยการเรียนรู้ โดยผลจากการทดสอบของ Huynh-Feldt = .699 ซึ่งแสดงว่าสมรรถนะจากการประเมิน ทั้ง 3 ครั้ง มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ $F(1.398, 41.932) = 925.953, p < .001$ ทั้งนี้ เมื่อเปรียบเทียบรายคู่ด้วยวิธีของ Bonferroni พบว่า 1) คะแนนครั้งที่ 2 > ครั้งที่ 1 ($d, t_2 - t_1 = 2.12$) 2) คะแนนครั้งที่ 3 > ครั้งที่ 1 ($d, t_3 - t_1 = 5.22$) และ 3) คะแนนครั้งที่ 3 > ครั้งที่ 2 ($d, t_3 - t_2 = 3.09$) โดยคะแนนครั้งที่ 1 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 3.17 คะแนนครั้งที่ 3 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 7.07 และคะแนนครั้งที่ 3 กับ ครั้งที่ 1 มีค่าขนาดอิทธิพล = 6.49 ซึ่งแสดงให้เห็นว่าสมรรถนะด้านการกำหนดผลลัพธ์การเรียนรู้ของหน่วยการเรียนรู้เพิ่มขึ้นอย่างต่อเนื่อง

3) ด้านการออกแบบกิจกรรมการเรียนรู้ของหน่วยการเรียนรู้ โดยผลจากการทดสอบของ Greenhouse-Geisser = .850 ซึ่งแสดงว่าสมรรถนะจากการประเมิน ทั้ง 3 ครั้ง มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ $F(1.356, 40.681) = 727.920, p < .001$ ทั้งนี้ เมื่อเปรียบเทียบรายคู่ด้วยวิธีของ Bonferroni พบว่า 1) คะแนนครั้งที่ 2 > ครั้งที่ 1 ($d, t_2 - t_1 = 7.03$) 2) คะแนนครั้งที่ 3 > ครั้งที่ 1 ($d, t_3 - t_1 = 17.06$) และ 3) คะแนนครั้งที่ 3 > ครั้งที่ 2 ($d, t_3 - t_2 = 10.03$) โดยคะแนนครั้งที่ 1 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 2.94 คะแนนครั้งที่ 3 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 5.94 และคะแนนครั้งที่ 3 กับ ครั้งที่ 1 มีค่าขนาดอิทธิพล = 5.77 ซึ่งแสดงให้เห็นว่าสมรรถนะด้านการออกแบบกิจกรรมการเรียนรู้ของหน่วยการเรียนรู้เพิ่มขึ้นอย่างต่อเนื่อง

4) ด้านการออกแบบกระบวนการวัดและประเมินผลของหน่วยการเรียนรู้ โดยผลจากการทดสอบของ Greenhouse-Geisser = .850 ซึ่งแสดงว่าสมรรถนะจากการประเมิน ทั้ง 3 ครั้ง มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ $F(1.700, 51.002) = 648.082, p < .001$ ทั้งนี้ เมื่อเปรียบเทียบรายคู่ด้วยวิธีของ Bonferroni พบว่า 1) คะแนนครั้งที่ 2 > ครั้งที่ 1 ($d, t_2 - t_1 = 11.61$) 2) คะแนนครั้งที่ 3 > ครั้งที่ 1 ($d, t_3 - t_1 = 18.09$) และ 3) คะแนนครั้งที่ 3 > ครั้งที่ 2 ($d, t_3 - t_2 = 6.48$) โดยคะแนนครั้งที่ 1 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 3.87 คะแนนครั้งที่ 3 กับ ครั้งที่ 2 มีค่าขนาดอิทธิพล = 2.86 และคะแนนครั้งที่ 3 กับ ครั้งที่ 1 มีค่าขนาดอิทธิพล = 5.46 ซึ่งแสดงให้เห็นว่าสมรรถนะด้านการออกแบบกระบวนการวัดและประเมินผลของหน่วยการเรียนรู้เพิ่มขึ้นอย่างต่อเนื่อง

ตาราง 4 การเปรียบเทียบผลการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากร
สายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี โดยการเรียนรู้ผ่าน
ประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับ จำแนกตามรอบการประเมินในภาพรวม

Mauchly's Test of Sphericity ^b							
Measure: MEASURE_1							
Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
LDC	.603	14.693	2	.001	.716	.741	.500
Tests of Within-Subjects Effects							
Measure: MEASURE_1							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Learning Design Competency	Sphericity	47600.409	2	23800.204	898.729	.000	.968
	Assumed						
	Greenhouse-Geisser	47600.409	1.431	33260.715	898.729	.000	.968
	Huynh-Feldt	47600.409	1.483	32099.348	898.729	.000	.968
	Lower-bound	47600.409	1.000	47600.409	898.729	.000	.968
Error (LDC)	Sphericity	1588.925	60				
	Assumed						
	Greenhouse-Geisser	1588.925	42.934				
	Huynh-Feldt	1588.925	44.487				
	Lower-bound	1588.925	30.000				
Post Hoc Comparisons – Learning Design Competency (LDC)							
Comparison							
(I) LDC	(J) LDC	Mean Dif.	SE	df	t	p_{bonferro ni}	Effect size
Time 1	Time 2	-26.129*	1.280	30.0	-20.0	< .001	3.63
	Time 3	-55.387*	1.635	30.0	-33.4	< .001	5.92
Time 2	Time 3	-29.258*	.902	30.0	-32.7	< .001	6.05

3. ผลการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี โดยการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับจำแนกตามชุดกลยุทธ์การให้ข้อมูลย้อนกลับ

ผลการศึกษาพบว่าโค้ชพี่เลี้ยงประจำกลุ่มทั้งหมด 4 คน เลือกใช้วิธีการให้ข้อมูลย้อนกลับแบบทั้งช่วงเวลา ส่วนช่องทางในการให้ข้อมูลย้อนกลับพบว่ารอบที่ 1 เลือกใช้การสนทนาแบบเผชิญหน้าทั้งหมด ในขณะที่รอบที่ 2 ใช้วิธีการแบบผสมผสานระหว่างการเขียนและการสนทนาแบบเผชิญหน้า คิดเป็นร้อยละ

64.5 และใช้วิธีการการสนทนาแบบเผชิญหน้า ร้อยละ 34.5 ตามลำดับ ผลการศึกษาพบว่าหลังจากได้รับข้อมูลย้อนกลับจากโค้ชที่เลี้ยงครั้งที่ 1 บุคลากรสายสนับสนุนวิชาการได้รับคะแนนเฉลี่ย 59 คะแนน ในขณะที่หลังจากได้รับการให้ข้อมูลย้อนกลับครั้งที่ 2 ได้รับคะแนนเฉลี่ย 88.26 คะแนน หากพิจารณาการเลือกใช้กลยุทธ์ในการให้ข้อมูลย้อนกลับ พบว่าครั้งที่ 1 ใช้ชุดกลยุทธ์ ACEFGHI ร้อยละ 39 ($\bar{x} = 63.42$, $SD = 9.423$) รองลงมา ABCDEFGHI ร้อยละ 13 ($\bar{x} = 53.75$, $SD = 3.500$) ลำดับที่สาม ACDEFGH ร้อยละ 13 ($\bar{x} = 55.25$, $SD = 2.217$) ACEFGH ร้อยละ 13 ($\bar{x} = 58.00$, $SD = 8.679$) ACEFHI ร้อยละ 13 ($\bar{x} = 57.50$, $SD = 7.853$) และ ACDEFGHI ร้อยละ 10 ($\bar{x} = 56.67$, $SD = 6.429$) ตามลำดับ ในขณะที่ครั้งที่ 2 ใช้ชุดกลยุทธ์ ACEFGHI เป็นลำดับแรก ร้อยละ 55 ($\bar{x} = 88.18$, $SD = 7.393$) รองลงมา ACDEFGHI ร้อยละ 26 ($\bar{x} = 90.25$, $SD = 7.066$) และ ACDFGHI ร้อยละ 19 ($\bar{x} = 85.83$, $SD = 6.940$) ตามลำดับ ดังแสดงในตารางที่ 4

จากผลการศึกษาข้างต้นแสดงให้เห็นว่าการให้ข้อมูลย้อนกลับในรอบที่ 1 ที่โค้ชที่เลี้ยงใช้ชุดกลยุทธ์การให้ข้อมูลย้อนกลับที่แตกต่างกันนั้นพบว่าทุกชุดกลยุทธ์ยังไม่สามารถช่วยยกระดับสมรรถนะให้ผ่านตามเกณฑ์ที่กำหนดได้ แต่เมื่อดำเนินการให้ข้อมูลย้อนกลับซ้ำในรอบที่ 2 พบว่าบุคลากรสายสนับสนุนวิชาการมีสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ที่ดีขึ้นจนบรรลุตามระดับที่คาดหวัง (> 70 คะแนน) โดยมีข้อสังเกตที่น่าสนใจ คือ รอบที่ 1 โค้ชที่เลี้ยงใช้ชุดกลยุทธ์การให้ข้อมูลย้อนกลับที่แตกต่างกัน จำนวน 6 ชุดกลยุทธ์ ในขณะที่รอบที่ 2 โค้ชที่เลี้ยงใช้ชุดกลยุทธ์การให้ข้อมูลย้อนกลับที่แตกต่างกันเพียง 3 ชุดกลยุทธ์ เท่านั้น

ตาราง 4 คะแนนความสามารถด้านการออกแบบกระบวนการจัดการเรียนรู้จำแนกตามชุดกลยุทธ์การให้ข้อมูลย้อนกลับ

Feedback Strategy (รอบที่ 1)	N	%	Mean	SD	Feedback Strategy (รอบที่ 2)	N	%	Mean	SD
ABCDEFGHI	4	13	53.75	3.500	ACDEFGHI	8	26	90.25	7.066
ACDEFGH	4	13	55.25	2.217	ACDFGHI	6	19	85.83	6.940
ACDEFGHI	3	10	56.67	6.429	ACEFGHI	17	55	88.18	7.393
ACEFGH	4	13	58.00	8.679					
ACEFGHI	12	39	63.42	9.424					
ACEFHI	4	13	57.50	7.853					
รวม	31	100	59.00	8.066	รวม	31	100	88.26	7.146
หมายเหตุ: การตรวจสอบความถูกต้อง (A) การระบุคำตอบที่ถูกต้อง (B) การให้ผู้เรียนทดลองทำอีกครั้ง (C) การแจ้งข้อผิดพลาด (D) การขยายความเพิ่มเติม (E) การให้ข้อเสนอแนะในประเด็นที่เกี่ยวข้อง (F) การให้ข้อเสนอแนะที่มีความเฉพาะเจาะจง (G) การให้คำแนะนำ/ ข้อเสนอแนะ/ ข้อความแจ้ง (H) การระบุข้อบกพร่อง/ ความเข้าใจผิด (I)									

นอกจากนี้ พบว่าในแต่ละชุดกลยุทธ์การให้ข้อมูลย้อนกลับประกอบด้วยกลยุทธ์ร่วม 7 กลยุทธ์ ได้แก่ 1) การตรวจสอบความถูกต้อง (A) 2) การให้ผู้เรียนทดลองทำซ้ำ (C) 3) การขยายความเพิ่มเติม (E) 4) การให้ข้อเสนอแนะในประเด็นที่เกี่ยวข้อง (F) 5) การให้ข้อเสนอแนะที่มีความเฉพาะเจาะจง และ 6) การให้คำแนะนำ/

ข้อเสนอแนะ/ ข้อความแจ้ง (H) และ 7) การระบุข้อบกพร่อง/ ความเข้าใจผิด (I) ดังแสดงในตาราง 4 อย่างไรก็ตามเมื่อเปรียบเทียบคะแนนสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการที่ได้รับกลยุทธ์การให้ข้อมูลย้อนกลับแตกต่างวิธีการ พบว่าสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการไม่แตกต่างกันตามชุดกลยุทธ์ กล่าวคือ การใช้ชุดกลยุทธ์ที่แตกต่างกันไม่มีผลทำให้สมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการมีความแตกต่างกัน ดังแสดงในตาราง 5

ตาราง 5 ผลการเปรียบเทียบคะแนนสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการที่ได้รับการให้ข้อมูลย้อนกลับด้วยชุดกลยุทธ์การให้ข้อมูลย้อนกลับต่างวิธีการ

Feedback Strategy round 1					
ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	429.917	5	85.983	1.412	.254
Within Groups	1522.083	25	60.883		
Total	1952	30			
Feedback Strategy round 2					
ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	429.917	2	85.983	.642	.534
Within Groups	1522.083	28	60.883		
Total	1952	30			

Note: * = significant at the 0.05 level

อภิปรายผล

จากผลการศึกษาพบว่า การจัดกิจกรรมการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับสามารถยกระดับสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการให้บรรลุตามผลลัพธ์การเรียนรู้ที่คาดหวัง ซึ่งจุดเด่นของการเรียนรู้ผ่านประสบการณ์ คือ การเรียนรู้ผ่านการลงมือปฏิบัติ ใช้โจทย์สถานการณ์จริงที่เกี่ยวข้องกับบทบาทหน้าที่ความรับผิดชอบ มีความยืดหยุ่น เป็นประโยชน์ต่อการประยุกต์ใช้ได้จริงในการปฏิบัติงาน สอดคล้องกับ Knowles et al. (2005) ที่กล่าวว่าการจัดการเรียนรู้สำหรับผู้เรียนวัยผู้ใหญ่ควรใช้การเรียนรู้ผ่านประสบการณ์เป็นหลักซึ่งเป็นการเชื่อมโยงระหว่างการเรียนรู้กับการทำงานเข้าด้วยกัน นอกจากนี้จากผลการศึกษาพบว่าบุคลากรสายสนับสนุนวิชาการมีสมรรถนะเพิ่มขึ้นอย่างต่อเนื่องหลังจากได้รับข้อมูลย้อนกลับจากโค้ชพี่เลี้ยง โดยคะแนนทั้งสามครั้งมีแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ซึ่งสามารถบรรลุตามผลลัพธ์การเรียนรู้ที่คาดหวังหลังจากได้รับการให้ข้อมูลย้อนกลับในรอบที่ 2 ผลการศึกษาในข้างต้นสอดคล้องกับ Black & Wiliam (2009); Hattie & Timperley (2007); Knowles et al. (2005); Schildkamp et al. (2020); Shute (2008) ที่กล่าวว่าการให้ข้อมูลย้อนกลับ

ในระหว่างการเรียนรู้ที่ทันตแพทย์ที่สามารถยกระดับความสามารถของผู้เรียน สอดคล้องกับผลการศึกษาของ Karaman (2021); Lipnevich & Panadero (2021); Voinea (2018) ที่พบว่า การให้ข้อมูลย้อนกลับด้วยการให้ข้อมูลที่เหมาะสมและเพียงพอส่งผลกระทบต่อพัฒนาความสามารถของผู้เรียน ในขณะที่ผลการศึกษาของ Bin Mubayrik (2020); Wisniewski et al. (2019) พบว่าการให้ข้อมูลย้อนกลับมีอิทธิพลต่อการพัฒนาโดยเฉพาะด้านความรู้และด้านทักษะ อย่างไรก็ตาม Asregid et al. (2023) เสนอว่าการให้ข้อมูลย้อนกลับควรอยู่ในขอบเขตเกณฑ์ที่กำหนดและควรอธิบายเกณฑ์การประเมินให้ผู้เรียนทราบ

จากผลการศึกษาพบว่าโค้ชที่เลี้ยงประจำกลุ่มเลือกใช้ชุดกลยุทธ์สำหรับการให้ข้อมูลย้อนกลับ ซึ่งในแต่ละชุดประกอบด้วยกลยุทธ์รวม ได้แก่ 1) การตรวจสอบความถูกต้อง 2) การให้ผู้เรียนทดลองทำซ้ำ 3) การขยายความเพิ่มเติม 4) การให้ข้อเสนอแนะในประเด็นที่เกี่ยวข้อง 5) การให้ข้อเสนอแนะที่มีความเฉพาะเจาะจง 6) การให้คำแนะนำ/ ข้อเสนอแนะ/ ข้อความแจ้ง และ 7) การระบุข้อบกพร่อง/ ความเข้าใจผิด สอดคล้องกับผลการศึกษาของ Dawson et al. (2020); Lipnevich & Panadero (2021) ที่กล่าวว่าการให้ข้อมูลย้อนกลับด้วยการระบุสิ่งที่ผู้เรียนต้องปรับปรุงและพัฒนาเพิ่มเติมให้ชัดเจน การให้ข้อมูลที่มีความเฉพาะเจาะจง การให้ข้อมูลที่มีความแม่นยำ ถูกต้องและน่าเชื่อถือ สามารถปิดช่องว่างความสามารถของผู้เรียนได้ สอดคล้องกับผลการศึกษาของ Morris et al. (2021) ที่พบว่า การให้ข้อมูลย้อนกลับในลักษณะของการให้คำแนะนำ (provision of feedback) ส่งผลกระทบต่อผู้เรียนโดยได้รับคะแนนเฉลี่ยเพิ่มขึ้น 13% จากส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบ นอกจากนี้ สอดคล้องกับผลการศึกษาของ Hu et al. (2021) ที่พบว่ากลยุทธ์การให้ข้อมูลย้อนกลับแบบเฉพาะเจาะจง (specific feedback strategy) ถูกนำมาใช้มากที่สุดในการพัฒนาและยกระดับความสามารถของผู้เรียนในรายวิชาวิทยาศาสตร์ โดยผลการศึกษายืนยันว่าการให้ข้อมูลย้อนกลับช่วยให้ผู้เรียนพัฒนาความสามารถบรรลุตามเกณฑ์ที่กำหนด สอดคล้องกับ Mason & Bruning (2001, as cited in Shute, 2008) ที่เสนอว่าหากเป็นเรื่องที่มีความซับซ้อนและผู้เรียนยังไม่เคยมีประสบการณ์มาก่อน ผู้สอนอาจจะต้องใช้หลายกลยุทธ์ในการยกระดับความสามารถของผู้เรียน เช่น การตรวจสอบความถูกต้อง การทิ้งช่วงเวลาสักครู่สำหรับการระบุคำตอบที่ใกล้เคียงแก่ผู้เรียน หรือการให้ข้อเสนอแนะที่มีความเฉพาะเจาะจงในประเด็นที่เกี่ยวข้อง เป็นต้น สอดคล้องกับ McKimm & Barrow (2009) ที่กล่าวว่าควรเลือกใช้กลยุทธ์การให้ข้อมูลย้อนกลับที่ช่วยปิดจุดอ่อนและส่งเสริมให้ผู้เรียนบรรลุตามเป้าหมายที่กำหนดไว้

จากผลการศึกษาพบว่าโค้ชที่เลี้ยงประจำกลุ่มทั้งหมดเลือกใช้วิธีการให้ข้อมูลย้อนกลับแบบทั้งช่วงเวลาทั้ง 2 รอบการประเมิน ส่วนช่องทางในการให้ข้อมูลย้อนกลับพบว่าครั้งที่ 1 โค้ชที่เลี้ยงเลือกใช้ช่องทางการสนทนาแบบเผชิญหน้าทั้งหมด ในขณะที่ครั้งที่ 2 เลือกใช้วิธีการแบบผสมผสานระหว่างการเขียนและการสนทนาแบบเผชิญหน้า คิดเป็นร้อยละ 64.5 และใช้วิธีการการสนทนาแบบเผชิญหน้า คิดเป็นร้อยละ 34.5 ตามลำดับ สอดคล้องกับผลการศึกษาของ O'Donovan et al. (2016, as cited in Carless & Young, 2024) ที่พบว่า การให้ข้อมูลย้อนกลับแก่ผู้เรียนเป็นรายบุคคลด้วยวิธีการพูดคุย (Oral feedback to individuals) จะเกิดประสิทธิภาพต่อผู้เรียนมากกว่าวิธีการเขียน (Written) เพียงอย่างเดียว สอดคล้องกับผลการศึกษาของ Tantiphlachiva et al. (2023) ซึ่งพบว่าคะแนนของผู้เรียนที่ได้รับการให้ข้อมูลย้อนกลับแบบเผชิญหน้า (F2F feedback) สูงกว่าผู้เรียนที่ได้รับข้อมูลย้อนกลับผ่านวิดีโออย่างมีนัยสำคัญทางสถิติ ในขณะที่ผลการศึกษาของ

Morris & Chikwa (2016) พบว่าผู้เรียนมีความพึงพอใจต่อการให้ข้อมูลย้อนกลับด้วยเสียง (audio feedback) มากกว่าวิธีการเขียน ถึงแม้ว่าจะไม่พบความแตกต่างของคะแนนอย่างมีนัยสำคัญทางสถิติ ซึ่งสามารถกล่าวสรุปได้ว่าการให้ข้อมูลย้อนกลับด้วยวิธีการแบบเผชิญหน้าเป็นวิธีการที่มีความเหมาะสม เนื่องจากผู้เรียนสามารถสอบถามหรือโต้ตอบกับผู้สอนในประเด็นต่าง ๆ ได้ทันทีที่ นอกจากนี้ เมื่อพิจารณาการเลือกใช้ชุดกลยุทธ์สำหรับให้ข้อมูลย้อนกลับของโค้ชพี่เลี้ยงประจำกลุ่มพบว่ารอบที่ 1 ใช้กลยุทธ์จำนวน 7 ชุดกลยุทธ์ ในขณะที่รอบที่ 2 ใช้เพียง 3 ชุดกลยุทธ์ ซึ่งสามารถกล่าวได้ว่าโค้ชพี่เลี้ยงประจำกลุ่มมีประสบการณ์ในการเลือกใช้ชุดกลยุทธ์ที่มีความเหมาะสมสำหรับการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 หน่วยงานที่เกี่ยวข้องกับการพัฒนาทรัพยากรบุคคล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี สามารถนำรูปแบบการจัดกิจกรรมการเรียนรู้ผ่านประสบการณ์ร่วมกับการให้ข้อมูลย้อนกลับไปใช้จริงสำหรับการวางแผนพัฒนาสมรรถนะเฉพาะงานของบุคลากรสายสนับสนุนวิชาการอย่างเป็นรูปธรรมในมิติของการเตรียมความพร้อมสำหรับการจัดกิจกรรมการเรียนรู้ ระบบการวัดและประเมิน ตลอดจนการจัดทำเป็นคู่มือสำหรับการพัฒนาสมรรถนะด้านการออกแบบกระบวนการจัดการเรียนรู้ของบุคลากรสายสนับสนุนวิชาการ

1.2 มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี สามารถนำเสนอสารสนเทศที่ได้จากการประเมิน เช่น ช่องทางการให้ข้อมูลย้อนกลับและชุดกลยุทธ์การให้ข้อมูลย้อนกลับที่ส่งผลต่อการพัฒนาสมรรถนะของบุคลากรสายสนับสนุนวิชาการไปใช้สำหรับการประเมินผลการปฏิบัติงานประจำปีของบุคลากรสายสนับสนุนวิชาการ โดยเฉพาะตำแหน่งงานนักพัฒนาการศึกษาและเทคโนโลยีการศึกษา

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ปัจจุบันมีการใช้ปัญญาประดิษฐ์แบบรู้สร้าง (AI) เข้ามาช่วยสนับสนุนกระบวนการจัดการเรียนรู้เพิ่มมากขึ้นและมีแนวโน้มถูกนำมาใช้ในการจัดการเรียนการสอนเพิ่มขึ้นอย่างต่อเนื่อง ถึงแม้ว่าผลการศึกษายืนยันชัดเจนว่าการให้ข้อมูลย้อนกลับแบบเผชิญหน้า (Face-to-face) จะส่งผลต่อการบรรลุตามผลลัพธ์การเรียนรู้ที่คาดหวังโดยประจักษ์ ดังนั้นจึงเป็นที่น่าสนใจเป็นอย่างยิ่งว่าหากมีการดำเนินการวิจัยเกี่ยวกับใช้กระบวนการให้ข้อมูลย้อนกลับด้วยวิธีการผสมผสานระหว่างโค้ชพี่เลี้ยงควบคู่กับการใช้เทคโนโลยีปัญญาประดิษฐ์แบบรู้สร้างจะส่งผลกระทบต่อพัฒนาสมรรถนะของบุคลากรสายสนับสนุนวิชาการเป็นเช่นไร และควรมีรูปแบบการดำเนินการอย่างไร

2.2 เป็นที่น่าสนใจเป็นอย่างยิ่งว่าหากผู้เรียนมีคุณลักษณะ มีความรู้ความสามารถในเรื่องที่จะพัฒนาแล้วบ้าง หรือหากสมรรถนะที่ต้องการพัฒนามีความยากและมีความซับซ้อนควรจะปรับสัดส่วนรูปแบบการจัดการเรียนรู้ให้มีความเหมาะสมเป็นอย่างไรบ้างที่ทำให้ผู้เรียนบรรลุตามผลลัพธ์การเรียนรู้ที่คาดหวัง มากกว่านั้นควรมีจุดเชื่อมโยงในแต่ละส่วนของการเรียนรู้อย่างไร

References

- Adelsberger, H. H., Kinshuk, Pawlowski, J. M., & Sampson, D. G. (2008). *Handbook on Information Technologies for Education and Training*. Springer.
- Asregid, D., Mihiretie, D. M., & Kassa, S. A. (2023). Teacher educators' use of feedback to facilitate reflective practice among pre-service teachers during microteaching. *Cogent Education*, 10(2). <https://doi.org/10.1080/2331186X.2023.2257121>
- Bice, M. R., & Sedani, L. A. (2021). The competencies for instructional designers in higher education. *Journal of Applied Instructional Design*, 10(2), 1–16. <https://doi.org/10.37074/jaid.2021.10.2.3>
- Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University: What the Student Does*. McGraw Hill.
- Bin Mubayrik, H. F. (2020). New trends in formative-summative evaluations for adult education. *SAGE Open*, 10(3). <https://doi.org/10.1177/2158244020941006>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. Springer.
- Carless, D., & Young, S. (2024). Feedback seeking and student reflective feedback literacy: A sociocultural discourse analysis. *Higher Education*, 88, 857–873. <https://doi.org/10.1007/s10734-023-01146-1>
- Dawson, P., Carless, D., & Lee, P. P. W. (2020). Authentic feedback: Supporting learners to engage in disciplinary feedback practices. *Assessment & Evaluation in Higher Education*, 45(3), 395–408. <https://doi.org/10.1080/02602938.2020.1769022>
- Dick, W., & Carey, L. (1996). *The systematic design of instruction*. HarperCollins College.
- Fuentes-Cimma, J., Sluijsmans, D., Riquelme, A., Jahouh, M., & Schuwirth, L. (2024). Designing feedback processes in the workplace-based learning of undergraduate health professions education: A scoping review. *BMC Medical Education*, 24(1), 1-25. <https://doi.org/10.1186/s12909-024-05439-6>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hu, B. Y., Guan, L., Li, Y., Roberts, S. K., & Zhang, X. (2021). Feedback matters: Examining the use of feedback strategies by Chinese preschool teachers in science lessons. *Early Childhood Education Journal*, 50(8), 1355–1371. <https://doi.org/10.1007/s10643-021-01265-3>

- Karaman, P. (2021). The effect of formative assessment practices on student learning: A meta-analysis study. *International Journal of Assessment Tools in Education*, 8(4), 801–817. <https://doi.org/10.21449/ijate.870300>
- Knowles, M. S. (1973). *The adult learner: A neglected species*. Gulf Publishing.
- Knowles, M. S., Holton, E. F., & Swanson, R. A. (2005). *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. Elsevier.
- Lipnevich, A., & Panadero, E. (2021). A review of feedback models and theories: Descriptions, definitions, and conclusions. *Frontiers in Education*, 6, 1-29. <https://doi.org/10.3389/feduc.2021.720195>
- Mason, B. J., & Bruning, R. (2001). *Providing feedback in computer-based instruction: What the research tells us* [Report No. 14]. Center for Instructional Innovation, University of Nebraska–Lincoln. <http://dwb.unl.edu/Edit/MB/MasonBruning.html>
- McKimm, J., & Barrow, M. (2009). Curriculum and course design. *British Journal of Hospital Medicine*, 70(12), 714–717. <https://doi.org/10.12968/hmed.2009.70.12.45510>
- Molloy, E. K., & Boud, D. (2014). Feedback models for learning, teaching and performance. In J. M. Spector, D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th ed., pp. 413–424). Springer. https://doi.org/10.1007/978-1-4614-3185-5_33
- Morris, C., & Chikwa, G. (2016). Audio versus written feedback: Exploring learners’ preference and the impact of feedback format on students’ academic performance. *Active Learning in Higher Education*, 17(2), 125–137. <https://doi.org/10.1177/1469787416637482>
- Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education*, 9(3), 1–26. <https://doi.org/10.1002/rev3.3292>
- Narciss, D. (2013). Designing and Evaluating Tutoring Feedback Strategies for digital learning environments on the basis of the Interactive Tutoring Feedback Model. *Digital Education Review*, 23(1), 7-23
- O’Donovan, B. (2017). How student beliefs about knowledge and knowing influence their satisfaction with assessment and feedback. *Higher Education*, 74(4), 617–633. <https://doi.org/10.1007/s10734-016-0068-y>
- Rao, N. J. (2020). Outcome-based education: An outline. *Higher Education for the Future*, 7(1), 5–11. <https://doi.org/10.1177/2347631119886418>

- Ritzhaupt, A., Kumar, S., & Martin, F. (2021). *The competencies for instructional designers in higher education*. Edtech Books.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144. <https://doi.org/10.1007/BF00117714>
- Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research*, 103, 1–16. <https://doi.org/10.1016/j.ijer.2020.101602>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Sisselman-Borgia, A. G., & Torino, G. C. (2017). Innovations in Experiential Learning for Adult Learners. *Journal of Applied Learning in Higher Education*, 7, 3-13.
- Spady, W. G. (1988). Organizing for Results-The Basis of Authentic Restructuring and Reform. *Educational Leadership*, 42(2), 4-8.
- Spector, J. M., Merrill, M. D., Elen, J., & Bishop, M. J. (2009). *Handbook of Research on Educational Communications and Technology*. Springer.
- Taber, K. S. (2017). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tantiphlachiva, K., Iramaneerat, C., & Lertbunnaphong, T. (2023). Improving surgical skills with feedback: Directly-observed versus video-recorded practice. *BMC Medical Education*, 23(1), 660. <https://doi.org/10.1186/s12909-023-04635-0>
- Turner, R. C., & Carlson, L. (2003). Indexes of item-objective congruence for multidimensional items. *International Journal of Testing*, 3(2), 163–171. https://doi.org/10.1207/s15327574ijt0302_5
- Voinea, L. (2018). Formative assessment as assessment for learning development. *Revista de Pedagogie - Journal of Pedagogy*, LXVI(1), 7–23. <https://doi.org/10.26755/RevPed/2018.1/7>
- Whitney, D. R., & Sabers, D. L. (1970). *Improving essay examinations III: Use of item analysis*, Technical Bulletin II [Mimeographed]. University Evaluation and Examination Service.
- Wisniewski, B., Zierer, K., & Hattie, J. (2019). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10, 3087. <https://doi.org/10.3389/fpsyg.2019.03087>

ผลการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการที่มีต่อความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6: การทดลองแบบอนุกรมเวลา

The Effects of the Process Writing Approach on Thai Essay Writing Ability of Grade 6 Students: A Time Series Experiment

จิรัฏฐิติกา สารบัณฑิตกุล^{1*} สุวิมล กฤษศยาห์² และ เด่นดาว ชลวิทย์³

Jirattitikan Sarabunditkun^{1*}, Suwimon Kritkharuehart² and Dendow Choolawit³

(Received: December 13, 2024; Revised: December 16, 2024; Accepted: December 22, 2024)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ (1) เปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 ก่อนและหลังการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการ (2) เปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 ระหว่างนักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการ กับนักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติ และ (3) ศึกษาความคงทนของความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 ในกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการ การวิจัยครั้งนี้เป็นการวิจัยกึ่งทดลองแบบอนุกรมเวลา กลุ่มตัวอย่าง คือ นักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 69 คน เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ (1) แผนการจัดกิจกรรมการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการและแบบปกติ (2) คู่มือการตรวจแก้ไขงานเขียน คู่มือการให้ข้อมูลย้อนกลับแบบให้คำชี้แนะสำหรับการตรวจงานเขียนความเรียงภาษาไทย (3) แบบวัดความสามารถในการเขียนความเรียงภาษาไทย (4) เกณฑ์การประเมินการเขียนความเรียงภาษาไทย สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน สถิติทดสอบทีแบบไม่เป็นอิสระต่อกัน (t-test dependent) และการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (repeated measures ANOVA) ผลการวิจัย สรุปได้ดังนี้ (1) ความสามารถในการ

¹ นักศึกษาปริญญาโท สาขาวิชาการประเมินและการวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

² รองศาสตราจารย์ สาขาวิชาการประเมินและการวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

³ ผู้ช่วยศาสตราจารย์ สาขาวิชาหลักสูตรและการสอน คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

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

² Associate Professor, Educational Evaluation and Research Program, Faculty of Education, Ramkhamhaeng University

³ Assistant Professor, Curriculum and Instruction Program, Faculty of Education, Ramkhamhaeng University

* Corresponding Author E-mail: s.jirattitikan@gmail.com

เขียนความเรียงภาษาไทยของนักเรียนหลังการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ สูงกว่าก่อนการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (2) นักเรียนกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ มีความสามารถในการเขียนความเรียงภาษาไทย สูงกว่านักเรียนกลุ่มที่ได้รับการจัดการเรียนรู้แบบปกติ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (3) นักเรียนในกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ พบว่ามีความคงทนของความสามารถในการเขียนความเรียงภาษาไทย อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

คำสำคัญ ความสามารถในการเขียนความเรียงภาษาไทย การเขียนแบบเน้นกระบวนการ
การทดลองแบบอนุกรมเวลา การให้ข้อมูลย้อนกลับแบบให้คำชี้แนะ

Abstract

This research aims to: (1) compare the Thai essay writing ability of Grade 6 students before and after instruction using the process writing approach; (2) compare the Thai essay writing ability between students taught using the process writing approach and those taught using the traditional approach; and (3) examine the retention of Thai essay writing ability among students taught using the process writing approach. This study employed a time series quasi-experimental design. The sample consisted of sixty-nine Grade 6 students. The research instruments included: (1) lesson plans based on the process writing approach and the traditional approach; (2) a feedback manual for evaluating Thai essay writing; (3) a Thai essay writing assessment form; and (4) an assessment rubric for evaluating Thai essay writing. Data were analyzed using mean, standard deviation, dependent samples t-test, and repeated measures ANOVA. The findings were as follows: (1) The Thai essay writing ability of students significantly improved after being taught using the process writing approach ($p < .05$). (2) Students taught using the process writing approach demonstrated significantly higher essay writing ability than those taught using the traditional approach ($p < .05$). (3) Students who learned through the process writing approach retained their Thai essay writing ability over time at a statistically significant level ($p < .05$).

Keywords: essay writing ability, process writing approach, time series experiment, hints feedback

บทนำ

การใช้ภาษาไทยให้ถูกต้อง เป็นทักษะอย่างหนึ่งที่ทำให้มนุษย์สามารถติดต่อสื่อสารกันให้ไปในทิศทางเดียวกันได้ แต่หากจะให้การสื่อสารมีประสิทธิภาพแล้ว การฝึกฝนทักษะในการใช้ภาษา ทั้งทักษะการฟัง การพูด การอ่าน และการเขียน จึงมีความจำเป็นอย่างมากเพราะจะทำให้การสื่อสารมีประสิทธิภาพมากขึ้น ดังที่ ดวงใจ ไทยอุบล (Thaiuboon, 2013) ได้กล่าวไว้ว่า การเขียนเป็นทักษะหนึ่งของการสื่อสารที่คงหลักฐานอ้างอิงไว้ได้ และสามารถใช้ทักษะการเขียนควบคู่กับทักษะการฟัง การอ่าน การพูด ได้อีกด้วย ความสำคัญในการใช้ทักษะการเขียนนอกจากเป็นการแสดงออกซึ่งประสบการณ์และความรู้ของผู้เขียนแล้ว การจัดการเรียนการสอนโดยให้ความสำคัญด้านการเขียนจึงมีความสำคัญอย่างยิ่งกับผู้เรียน แต่ในทางปฏิบัติกลับพบว่า งานเขียนเป็นทักษะที่มีความยาก การที่ผู้เรียนจะเริ่มงานเขียนได้ก็ต่อเมื่อผู้เรียนมีความรู้ ประสบการณ์ และจินตนาการที่มากพอในการสร้างงานเขียนขึ้นมา (Sumniengngam, 2021) อีกทั้งการเรียนการสอนส่วนใหญ่ไม่ส่งเสริมความสามารถด้านการเขียนทั้งที่การเขียนเป็นการสื่อสารที่สำคัญ และจากการสอนรายวิชาภาษาไทยในระดับประถมศึกษาชั้นปีที่ 6 พบว่า ผู้เรียนส่วนใหญ่ยังประสบปัญหาในด้านการเขียน (Thipyotha, 2023) ทั้งยังมีผลการทดสอบทางการศึกษาระดับชาตินั้นพื้นฐาน (O-NET) ประจำปีการศึกษา 2566 มีจุดที่ควรพัฒนาเกี่ยวกับการใช้ภาษาไทยจากรูปแบบการเขียนและสรุปปัญหาการใช้ภาษาไทยในภาพรวม 4 ด้าน ได้แก่ 1) ด้านการเรียบเรียงความ กลวิธีการสรุปความ การเสนอแนวคิดสำคัญ และการคัดลอกข้อความ 2) ด้านการสะกดคำ 3) ด้านการใช้คำ และ 4) ด้านการเขียนประโยค (National Institute of Educational Testing Service, 2024) อีกทั้งผลการทดสอบระดับชาติ O-NET วิชาภาษาไทย ชั้นประถมศึกษาปีที่ 6 ตั้งแต่ปีการศึกษา 2564 - 2566 นั้นมีคะแนนเฉลี่ยเท่ากับ 50.38, 53.89 และ 57.30 ตามลำดับ (National Institute of Educational Testing Service, 2024) ครูผู้สอนจึงควรมีวิธีการที่จะทำให้ผู้เรียนเกิดประสิทธิภาพ และเกิดความคงทนของความสามารถในการเขียนความเรียงภาษาไทยให้สูงขึ้น นั่นคือการที่นักเรียนได้ฝึกเขียนด้วยตนเองบ่อยครั้ง มีผู้เชี่ยวชาญคอยให้คำแนะนำ และคอยเสริมเทคนิคการเขียน จนผู้เรียนสามารถเรียนรู้ด้วยตนเองและเกิดเป็นความชำนาญ

สำหรับแนวคิดการเขียนแบบเน้นกระบวนการ เป็นแนวคิดที่ใช้ในการสอนเขียนที่ให้ความสำคัญกับกระบวนการคิดและเน้นการฝึกปฏิบัติอย่างเป็นขั้นตอน ด้วยการแลกเปลี่ยนความคิดเห็น การวางแผนร่วมกัน และการให้ข้อเสนอแนะระหว่างสมาชิกในกลุ่ม (Brown & Hood, 1989) รวมทั้งการให้ข้อมูลย้อนกลับ สอดคล้องกับ Bureau of Academic Affairs and Educational Standards, Office of the Basic Education Commission (2017) ได้กล่าวว่า การให้ข้อมูลย้อนกลับแก่ผู้เรียน เป็นการพัฒนาการเรียนรู้ให้ผู้เรียนตระหนักในการเรียนรู้อย่างต่อเนื่อง สอดคล้องกับการประเมินผู้เรียนในศตวรรษที่ 21 รวมทั้ง van den Bergh et al. (2016) ได้กล่าวว่ากระบวนการเขียนเป็นกิจกรรมที่ค่อนข้างซับซ้อน และต้องมีการโต้ตอบระหว่างหลายปัจจัยทั้งในด้านของความคิด อารมณ์ และสถานการณ์ต่าง ๆ ที่เกี่ยวข้องกับการเขียน เช่น การสร้างไอเดีย การวางแผน การจัดระเบียบข้อมูล และปรับกระบวนการเหล่านี้ตามข้อมูลย้อนกลับที่ได้จากงานเขียน การเขียนแบบเน้นกระบวนการจะเน้นให้ผู้เรียนวิเคราะห์ปัญหาและวางแผนการแก้ปัญหาตลอดกระบวนการเขียน ตั้งแต่การวางแผน การเขียน ขั้นการเขียน และการปรับปรุงแก้ไขงานเขียน (White & Arndt, 1991) เน้นการใช้กลวิธีในการเขียน

หลากหลายรูปแบบและฝึกให้ผู้เรียนได้ปรับปรุงงานเขียนของตนเอง กระบวนการจัดการเรียนการสอนโดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ เป็นวิธีการเขียนที่นักเรียนมุ่งเน้นไปที่กระบวนการในการสร้างงานเขียนของตนมากกว่าที่จะมุ่งเน้นไปที่ชิ้นงานที่สำเร็จ (Onozawa, 2010) แนวคิดสำคัญและเทคนิคในการสอนเขียนแบบเน้นกระบวนการ มีจุดประสงค์เพื่อช่วยให้ครูผู้สอนสามารถออกแบบกิจกรรมการเรียนการสอนในการพัฒนาทักษะการเขียนของนักเรียน (Raimes, 1983) การจัดการเรียนการสอนเขียนแบบเน้นกระบวนการจะสามารถส่งเสริมความสามารถในการเขียนความเรียงให้กับผู้เรียนได้อย่างมีประสิทธิภาพ การเปลี่ยนแปลงที่เกิดขึ้นจากการเขียนไม่เพียงแต่เป็นกระบวนการที่เกิดขึ้นของผู้เขียน แต่ยังรวมถึงการทำกิจกรรมร่วมกันหรือการมีปฏิสัมพันธ์กับผู้อื่นเพื่อแสดงให้เห็นความสำคัญของการรับรู้ถึงความหลากหลายในการเขียน (Atkinson, 2003) การศึกษาแนวคิดการเขียนแบบเน้นกระบวนการ จึงเป็นสิ่งที่ทำให้ผู้เรียนได้ฝึกปฏิบัติและมีหลักในการเขียน ผู้เรียนสามารถบรรลุวัตถุประสงค์ในการเขียนได้ถูกต้อง ฝึกให้ผู้เรียนมีการวางแผนในการคิดก่อนที่จะทำการเขียน และที่สำคัญในแนวคิดการเขียนแบบเน้นกระบวนการยังมีการให้ข้อมูลย้อนกลับเพื่อให้ผู้เรียนได้แก้ไขงานเขียนของตนเองได้ถูกต้องตามหลักการเขียน

แม้การเขียนจะมีความสำคัญ แต่หากครูผู้สอนสอนเพียงอย่างเดียวโดยไม่ใส่เทคนิคการสอนด้านอื่นที่จะส่งเสริมความสามารถในการเขียนของผู้เรียน การเขียนความเรียงก็จะประสบปัญหาแบบเดิม ๆ การให้ข้อมูลย้อนกลับในการเขียนจึงมีส่วนสำคัญที่จะทำให้ผู้เรียนแสดงออกถึงพัฒนาการด้านการเรียนรู้ของตนเอง และนำข้อมูลที่สะท้อนกลับเหล่านั้นไปปรับแก้และปรับปรุงงานเขียนของตนเองในครั้งต่อไปได้ดียิ่งขึ้น และนักเรียนมีความคงทนของทักษะการเขียนจากการให้ข้อมูลย้อนกลับของครู (Mason & Bruning, 2001) จากความสำคัญของปัญหา และแนวคิดดังกล่าว ผู้วิจัยจึงสนใจที่จะศึกษาและวิเคราะห์ผลของการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการที่มีต่อความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 ด้วยการทดลองแบบอนุกรมเวลา เพื่อเป็นประโยชน์ในการนำวิธีการให้จัดการเรียนการสอนมาทดลอง และช่วยส่งเสริมให้ผู้เรียนได้พัฒนาความสามารถในการเขียนความเรียงให้ดีที่สุด รวมทั้งการพัฒนาทักษะทางภาษาไทยด้านการเขียนเพื่อการสื่อสารในรายวิชาภาษาไทย สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ให้มีความคงทนและเกิดประสิทธิภาพสูงสุดในการเขียน

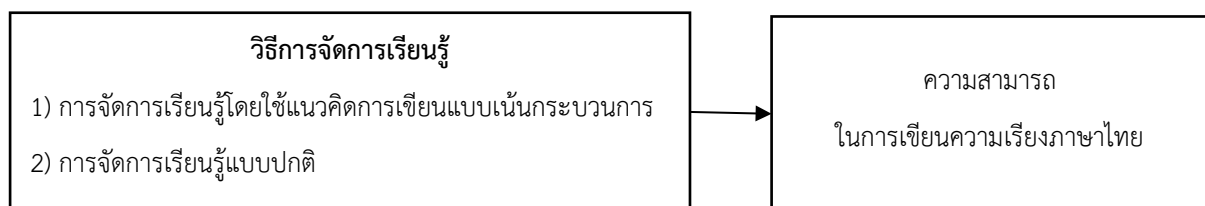
ความมุ่งหมายของการวิจัย

1. เพื่อเปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 ก่อนและหลังการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ
2. เพื่อเปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 ระหว่างนักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ กับกลุ่มนักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติ
3. เพื่อศึกษาความคงทนของความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 ในกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ

สมมติฐานการวิจัย

1. นักเรียนชั้นประถมศึกษาปีที่ 6 ที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ มีความสามารถในการเขียนความเรียงภาษาไทย สูงกว่าก่อนการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ
2. นักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ จะมีความสามารถในการเขียนความเรียงภาษาไทย หลังการทดลองสูงกว่ากลุ่มนักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติ
3. นักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ มีความคงทนของความสามารถในการเขียนความเรียงภาษาไทย

กรอบแนวคิดการวิจัย



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ที่ศึกษาในโรงเรียนขนาดใหญ่พิเศษ สังกัดสำนักงานเขตดินแดง กรุงเทพมหานคร จำนวนทั้งหมด 3 โรงเรียน ได้แก่ โรงเรียนสามเสนนอก (ประชาราษฎร์อ่อนกุล) ซึ่งมีนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 10 ห้องเรียน และจำนวนนักเรียนรวม 353 คน โรงเรียนวิสุทธิศ ซึ่งมีชั้นประถมศึกษาปีที่ 6 จำนวน 4 ห้องเรียน และจำนวนนักเรียนรวม 131 คน และโรงเรียนวิชากร ซึ่งมีชั้นประถมศึกษาปีที่ 6 จำนวน 4 ห้องเรียน และจำนวนนักเรียนรวม 135 คน มีนักเรียนชั้นประถมศึกษาปีที่ 6 ทั้ง 3 โรงเรียน จำนวนนักเรียนรวม 619 คน โดยทำการทดสอบความแปรปรวนความสามารถด้านภาษาไทย ของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่เป็นประชากรในการวิจัย พบว่า ไม่มีความแตกต่างกัน ดังนั้น ผู้วิจัยจึงดำเนินการสุ่มโรงเรียน ด้วยการสุ่มแบบกลุ่มครั้งที่ 1 โดยใช้โรงเรียนเป็นหน่วยในการสุ่ม ผลการสุ่มพบว่า คือโรงเรียนวิชากร สังกัดสำนักงานเขตดินแดง กรุงเทพมหานคร จำนวน 4 ห้องเรียน ซึ่งมีการจัดห้องเรียนแบบคละความสามารถ

กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ นักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนวิชากร สังกัดสำนักงานเขตดินแดง จำนวน 2 ห้องเรียน เป็นนักเรียนทั้งหมดจำนวน 69 คน ซึ่งกำหนดขนาดตัวอย่างด้วยโปรแกรม G*Power 3.1.9.4 (Faul et al., 2007) ที่วิเคราะห์ด้วย ANOVA: Repeated measures, within-between interaction กำหนดค่า Effect size = 0.20, ค่า α err prob = .05, ค่า Power ($1-\beta$ err prob) = 0.95, ค่า Number of groups = 2 และค่า Number of measurements = 4 เมื่อกตคำนวณจึงได้ค่าขนาดของ

กลุ่มตัวอย่าง จำนวน 56 คน โดยทำการทดสอบความแปรปรวนความสามารถด้านภาษาไทย ของนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนวิชาการ พบว่า นักเรียนแต่ละห้องมีความสามารถด้านภาษาไทยไม่แตกต่างกัน ผู้วิจัยจึงดำเนินการสุ่มห้องเรียนด้วยการสุ่มแบบกลุ่มครั้งที่ 2 โดยใช้ห้องเรียนเป็นหน่วยในการสุ่ม ผลการสุ่มพบว่า คือ นักเรียนชั้นประถมศึกษาปีที่ 6/3 และนักเรียนชั้นประถมศึกษาปีที่ 6/4 โรงเรียนวิชาการ ที่ได้มาโดยวิธีการสุ่มแบบกลุ่มสองขั้นตอน (two-stage cluster sampling) และสุ่มตัวอย่างเข้ากลุ่มทดลองและกลุ่มควบคุม ด้วยวิธีการจับสลาก ได้แก่ (1) กลุ่มทดลอง คือ นักเรียนชั้นประถมศึกษาปีที่ 6/4 จำนวน 35 คน เป็นนักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ (2) กลุ่มควบคุม คือ นักเรียนชั้นประถมศึกษาปีที่ 6/3 จำนวน 34 คน เป็นนักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติ

แบบแผนการวิจัย

การวิจัยในครั้งนี้ใช้แบบแผนการวิจัยกึ่งทดลอง (quasi-experimental designs) ศึกษาสองกลุ่มวัดหลายครั้งแบบอนุกรมเวลา (interrupted time series with control-group) (Patrawiwat, 2016; Shadish et al., 2002) แบ่งเป็น กลุ่มทดลอง 1 กลุ่ม และกลุ่มควบคุม 1 กลุ่ม จากนั้นจัดกิจกรรมการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ ทำการวัดความสามารถในการเขียนความเรียงภาษาไทย เพื่อศึกษาความคงทนของความสามารถในการเขียนความเรียงภาษาไทยหลังการจัดกิจกรรมการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการแล้วสัปดาห์ที่ 1 และสัปดาห์ที่ 2 ตามลำดับ เพื่อพัฒนาความสามารถในการเขียนความเรียงภาษาไทยของผู้เรียนต่อไป ซึ่งมีแบบแผนการวิจัย ดังภาพ 2

E	O ₁	O ₂	X	O ₃	O ₄
C	O ₁	O ₂		O ₃	O ₄

ภาพ 2 แบบแผนการวิจัย

เมื่อ E หมายถึง นักเรียนกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ

C หมายถึง นักเรียนกลุ่มที่ได้รับการจัดการเรียนรู้แบบปกติ

O₁ หมายถึง การวัดความสามารถในการเขียนความเรียงภาษาไทยก่อนการทดลองครั้งที่ 1

O₂ หมายถึง การวัดความสามารถในการเขียนความเรียงภาษาไทยก่อนการทดลองครั้งที่ 2

O₃ หมายถึง การวัดความสามารถในการเขียนความเรียงภาษาไทยหลังการทดลองครั้งที่ 1

O₄ หมายถึง การวัดความสามารถในการเขียนความเรียงภาษาไทยหลังการทดลองครั้งที่ 2

X หมายถึง การจัดกิจกรรมการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ

2. เครื่องมือที่ใช้ในการวิจัย

2.1 แผนการจัดกิจกรรมการเรียนรู้ ชั้นประถมศึกษาปีที่ 6 จำนวน 5 แผน 10 ชั่วโมง คือ

- (1) แผนการจัดกิจกรรมการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการสำหรับกลุ่มทดลอง
- (2) แผนการจัดกิจกรรมการเรียนรู้แบบปกติสำหรับกลุ่มควบคุม ผลการพิจารณาความตรงเชิงเนื้อหาจากผู้เชี่ยวชาญ จำนวน 5 ท่าน มีค่า IOC อยู่ระหว่าง .60 ถึง 1.00 และนำแผนการจัดกิจกรรมการเรียนรู้

โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ จำนวน 1 แผน ไปทดลองใช้กับนักเรียนที่มีลักษณะคล้ายคลึงกัน กับกลุ่มตัวอย่าง จำนวน 33 คน เพื่อตรวจสอบความเหมาะสมด้านภาษาและเวลาที่ใช้ นำแผนที่ปรับปรุงไปใช้เป็นเครื่องมือวิจัย ดังตาราง 1

ตาราง 1 ตัวอย่างการจัดกิจกรรมการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการและการจัดการเรียนรู้แบบปกติ

กิจกรรม	วิธีการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ	วิธีการจัดการเรียนรู้แบบปกติ
ชั่วโมงที่ 1 (60 นาที)		
ขั้นก่อนการเขียน (Pre-writing)	นักเรียนใช้กระบวนการกลุ่มเพื่อพูดคุยแลกเปลี่ยนความคิดเห็นตามประเด็นที่กำหนด โดยทุกคนในกลุ่มร่วมกันแสดงความคิดเห็นเพื่อฝึกเขียนตามกลวิธีการเขียนที่เป็นกระบวนการอย่างต่อเนื่อง ตามแนวทางการจัดกิจกรรม 4 รูปแบบ ได้แก่ • การระดมความคิด (Brainstorming idea) เป็นการใช้กระบวนการกลุ่มในการออกแบบความรู้ที่นักเรียนได้ตามหัวข้อที่กำหนด โดยให้นักเรียนออกแบบในรูปของแผนที่ความคิด • การเลือกความคิด (Selecting ideas) นำส่วนจากการระดมความคิดมาฝึกเขียนเป็นงานฉบับร่าง • การวางแผน (Planning) นักเรียนแต่ละกลุ่มร่วมกันทำใบงานที่ 1 และร่วมกันตรวจสอบข้อผิดพลาดของงานเขียน • การเรียบเรียงเนื้อหา (Organizing) นักเรียนเขียนงานฉบับร่างด้วยตนเอง โดยมีเพื่อนในกลุ่มให้ความช่วยเหลือซึ่งกันและกัน	ครูกำหนดหัวข้อให้นักเรียนเขียนงานของตนเอง ตามแนวทางการจัดกิจกรรม 3 รูปแบบ ได้แก่ • การเลือกความคิด (Selecting ideas) โดยนักเรียนแต่ละคนทราบประเด็นที่กำหนด โดยนักเรียนต่างคนต่างต้องฝึกงานเขียนฉบับร่างของตนเอง • การวางแผน (Planning) นักเรียนนำงานที่ฝึกเขียนฉบับร่างมาทำใบงานที่ 1 • การเรียบเรียงเนื้อหา (Organizing) นักเรียนนำงานเขียนฉบับร่างเดิมมาตรวจแก้ไขพร้อมทั้งตรวจสอบข้อผิดพลาดของตนเอง
ชั่วโมงที่ 2 (60 นาที)		
ขั้นการเขียน (Writing)	1) จากงานในกลุ่มที่นักเรียนร่วมกันฝึกการเขียนในใบงานที่ 1 ของกลุ่ม นำสู่ขั้นการเขียนในใบงานที่ 2 โดยใช้ใบงานที่ 1 มาเป็นแนวทางในการเขียน 2) นักเรียนทำใบงานที่ 2 เป็นรายบุคคล เพื่อฝึกการเขียน โดยยังมีเพื่อนในกลุ่มช่วยเหลือ 3) นักเรียนนำงานส่งครูเพื่อตรวจสอบความถูกต้อง โดยตรวจสอบข้อผิดพลาดเบื้องต้นว่าตรงกับประเด็นการเขียน หรือไม่	1) จากใบงานที่ 1 ที่นักเรียนทำรายบุคคล นำสู่ขั้นการเขียนในใบงานที่ 2 โดยใช้ใบงานที่ 1 มาเป็นแนวทางในการเขียน 2) นักเรียนนำงานส่งครูเพื่อตรวจสอบความถูกต้อง โดยตรวจสอบข้อผิดพลาดเบื้องต้นว่าตรงกับประเด็นการเขียน หรือไม่
ขั้นการปรับปรุง (Revising)	1) นำใบงานที่ 2 ที่ตรวจแล้วมาแก้ไขตามคำแนะนำเบื้องต้น 2) ครูตรวจงานโดยใช้คู่มือการให้ข้อมูลย้อนกลับแบบให้คำชี้แนะ พร้อมทั้งปรับปรุงงานเขียนของตนเองอีกครั้ง	1) นำใบงานที่ 2 ที่ตรวจแล้วมาแก้ไขตามคำแนะนำเบื้องต้น 2) ครูตรวจสอบความถูกต้องของการสะกดคำ และการใช้เครื่องหมายวรรคตอนโดยไม่ใช้คู่มือการให้ข้อมูลย้อนกลับแบบให้คำชี้แนะ
ขั้นการแก้ไข (Editing)	1) นักเรียนนำงานส่งครูอีกครั้งเพื่อตรวจสอบ เรื่องของไวยากรณ์ การสะกดคำและเครื่องหมายวรรคตอน 2) นักเรียนแก้ไขงานเพิ่มเติม และส่งงานครู เพื่อให้ครูตรวจให้คะแนนการเขียนเพื่อนำไปประเมินต่อไป	1) นักเรียนแก้ไขงาน และส่งงานครู เพื่อให้ครูตรวจให้คะแนนการเขียนเพื่อนำไปประเมินต่อไป

2.2 คู่มือการตรวจแก้ไขงานเขียน คู่มือการให้ข้อมูลย้อนกลับแบบให้คำชี้แนะ สำหรับการตรวจงานเขียนความเรียงภาษาไทย เป็นคู่มือที่ผู้วิจัยสร้างขึ้น เพื่อใช้เป็นแนวทางในการตรวจงานและการให้ข้อมูลย้อนกลับแบบให้คำชี้แนะ เป็นแนวทางให้ครูผู้สอนนำไปใช้เป็นคู่มือในการตรวจแก้ไขงานเขียนให้เป็นไปในทางเดียวกัน ข้อมูลย้อนกลับที่ให้เกี่ยวข้องกับลักษณะเด่นที่พบในงานเขียนของนักเรียน ใน 4 ด้านใหญ่ ๆ (ตามเกณฑ์การประเมินการเขียนความเรียงภาษาไทยที่ผู้วิจัยสร้างขึ้น) ได้แก่ การนำเสนอเนื้อหา การใช้ภาษา การเรียบเรียงความคิด และการใช้กลไกการเขียน (Rueangkitchanan, 2016) ผลการพิจารณาความตรงเชิงเนื้อหาจากผู้เชี่ยวชาญ จำนวน 5 ท่าน มีค่า IOC เท่ากับ 1.00 หลังจากนั้นผู้วิจัยได้นำข้อเสนอแนะมาปรับปรุงคู่มือการตรวจแก้ไขงานเขียนความเรียงวิชาภาษาไทยและแก้ไขก่อนนำไปทดลองใช้

2.3 แบบวัดความสามารถในการเขียนความเรียงวิชาภาษาไทย เป็นแบบทดสอบอัตนัย 4 ฉบับ เพื่อใช้ในการวัดความสามารถในการเขียนความเรียงภาษาไทยก่อนและหลังการทดลอง โดยให้นักเรียนเขียนความเรียงจากเรื่องที่กำหนด และเขียนให้ตรงตามวัตถุประสงค์ให้ถูกต้องตามเรื่อง ซึ่งเป็นเรื่องที่นักเรียนพบเจอในชีวิตประจำวัน ได้แก่ (1) เด็กดีของโรงเรียน (2) การปฏิบัติตามระเบียบของโรงเรียน (3) การมีส่วนร่วมในการทำกิจกรรมของโรงเรียน และ (4) จิตอาสาต่าง ๆ ทำได้ที่โรงเรียน ความยาว 7-10 บรรทัด คะแนนเต็ม 13 คะแนน โดยให้คะแนนตามเกณฑ์การประเมินการเขียนความเรียงภาษาไทยที่ผู้วิจัยสร้างขึ้น กำหนดเวลาในการเขียน 60 นาที ผลการพิจารณาความตรงเชิงเนื้อหาจากผู้เชี่ยวชาญ จำนวน 5 ท่าน มีค่า IOC เท่ากับ 1.00 ทั้ง 4 ฉบับ ผู้วิจัยได้นำข้อเสนอแนะมาปรับปรุงแบบวัดความสามารถและแก้ไขก่อนนำไปทดลองใช้กับนักเรียนที่มีลักษณะคล้ายคลึงกันกับกลุ่มตัวอย่าง ซึ่งเป็นนักเรียนชั้นประถมศึกษาปีที่ 6/1 โรงเรียนวิชาการ สังกัดสำนักงานเขตดินแดง กรุงเทพมหานคร จำนวน 33 คน วิเคราะห์หาคุณภาพของแบบวัดความสามารถในการเขียนความเรียงภาษาไทยฉบับที่ 1 โดยแบ่งกลุ่มนักเรียนออกเป็น 2 กลุ่ม คือ กลุ่มเก่งและกลุ่มอ่อน โดยใช้เทคนิค 25 % ของจำนวนนักเรียนที่สอบทั้งหมด 33 คน คำนวณจากสูตรของ Whitney & Sabers (1970) วิเคราะห์หาค่าความยาก (P_E) และค่าอำนาจจำแนก (D) ผลการวิเคราะห์ พบว่า แบบวัดความสามารถในการเขียนความเรียงภาษาไทย มีค่าความยาก เท่ากับ 0.45 และค่าอำนาจจำแนก เท่ากับ 0.64 วิเคราะห์หาค่าความเที่ยงของแบบวัดความสามารถในการเขียนความเรียงภาษาไทย โดยใช้สูตรสัมประสิทธิ์แอลฟาของครอนบาค พบว่า แบบวัดความสามารถในการเขียนความเรียงภาษาไทย มีค่าความเที่ยง เท่ากับ 0.80

2.4 เกณฑ์การประเมินการเขียนความเรียงภาษาไทย ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง นำข้อมูลมาสร้างเกณฑ์การประเมินงานเขียนความเรียงวิชาภาษาไทย (Rueangkitchanan, 2016) คือ (1) ด้านการนำเสนอเนื้อหา 4 คะแนน ได้แก่ ความสามารถในการนำเสนอเนื้อหาได้อย่างสอดคล้องกับหัวข้อ และวัตถุประสงค์ของงานเขียน (2) ด้านการใช้ภาษา 4 คะแนน ได้แก่ ความสามารถในการเลือกใช้คำให้ถูกต้อง ไม่ใช้คำซ้ำ (3) ด้านการเรียบเรียงความคิด 3 คะแนน ได้แก่ ความสมบูรณ์ของประโยค เนื้อเรื่องมีการเชื่อมโยงกันตลอดทั้งเรื่อง และ (4) ด้านการใช้รูปแบบและกลไกการเขียน 2 คะแนน ได้แก่ การใช้เครื่องหมายวรรคตอนได้อย่างเหมาะสม ผลการพิจารณาความตรงเชิงเนื้อหาจากผู้เชี่ยวชาญ จำนวน 5 ท่าน มีค่า IOC เท่ากับ 1.00 หลังจากนั้นผู้วิจัยได้นำข้อเสนอแนะมาปรับปรุงและแก้ไขก่อนนำไปทดลองใช้

3. การเก็บรวบรวมข้อมูล ผู้วิจัยมีการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างวิจัยเป็นระยะเวลา 6 สัปดาห์ ดังนี้

3.1 ก่อนการทดลองสัปดาห์ที่ 1 – 2 ผู้วิจัยวัดความสามารถในการเขียนความเรียงภาษาไทย จำนวน 2 ครั้ง โดยใช้แบบวัดความสามารถในการเขียนความเรียงภาษาไทย ฉบับที่ 1 และ 2 ตามลำดับ

3.2 ดำเนินการทดลองสัปดาห์ที่ 3 - 4 จัดกิจกรรมการเรียนการสอนวิชาภาษาไทย และใช้แผนการจัดการเรียนรู้ทั้งหมด 5 แผน 10 ชั่วโมง คือ (1) แผนการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ จำนวน 5 แผน แต่ละแผนใช้เวลา 2 คาบ จำนวน 120 นาที แต่ละแผนประกอบด้วยประเด็นที่จะใช้ในการพัฒนาการเขียน สำหรับกลุ่มทดลอง รวมทั้งสิ้น 5 ประเด็นการเขียน (2) แผนการจัดกิจกรรมการเรียนรู้แบบปกติ จำนวน 5 แผน แต่ละแผนใช้เวลา 2 คาบ จำนวน 120 นาที สำหรับกลุ่มควบคุม แต่ละแผนประกอบด้วยประเด็นที่จะใช้ในการพัฒนาการเขียน รวมทั้งสิ้น 5 ประเด็นการเขียน

3.3 หลังการทดลองสัปดาห์ที่ 5 – 6 ผู้วิจัยวัดความสามารถในการเขียนความเรียงภาษาไทย จำนวน 2 ครั้ง โดยใช้แบบวัดความสามารถในการเขียนความเรียงภาษาไทย ฉบับที่ 3 และ 4

4. การวิเคราะห์ข้อมูล

4.1 เปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนก่อนและหลังการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการด้วยการวิเคราะห์สถิติทดสอบทีแบบไม่เป็นอิสระต่อกัน (t-test dependent)

4.2 เปรียบเทียบความสามารถเปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยระหว่างนักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการและนักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติ ด้วยการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (repeated measures ANOVA)

4.3 เพื่อศึกษาความคงทนของความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนกลุ่มทดลองที่ใช้แนวคิดการเขียนแบบเน้นกระบวนการ ด้วยการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (repeated measures ANOVA)

ผลการวิจัย

1. ผลการเปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทย จากผลครั้งที่ 2 ก่อนการทดลอง (O_2) และผลครั้งแรกหลังการทดลอง (O_3) หลังการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ นำเสนอดังตาราง 1

ตาราง 1 ผลการเปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนกลุ่มทดลองหลังการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ

ความสามารถในการเขียนความเรียงภาษาไทย	M	SD	t	df	p
ผลครั้งที่ 2 ก่อนการทดลอง (O_2)	6.99	1.88	-12.88	68	.000
ผลครั้งแรกหลังการทดลอง (O_3)	9.04	1.97			

จากตาราง 1 พบว่า ความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนหลังการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการสูงกว่าก่อนได้รับการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2. ผลการเปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยระหว่างนักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการและนักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติ นำเสนอดังตาราง 2

ตาราง 2 ผลการทดสอบความแตกต่างของความสามารถในการเขียนความเรียงภาษาไทยระหว่างกลุ่มทดลองและกลุ่มควบคุม

Source	Type III Sum of Squares	df	Mean Square	F	p
Intercept	17538.067	1	17538.067	1851.345	.000
Group	47.169	1	47.169	4.979	.029
Error	634.701	67	9.473		

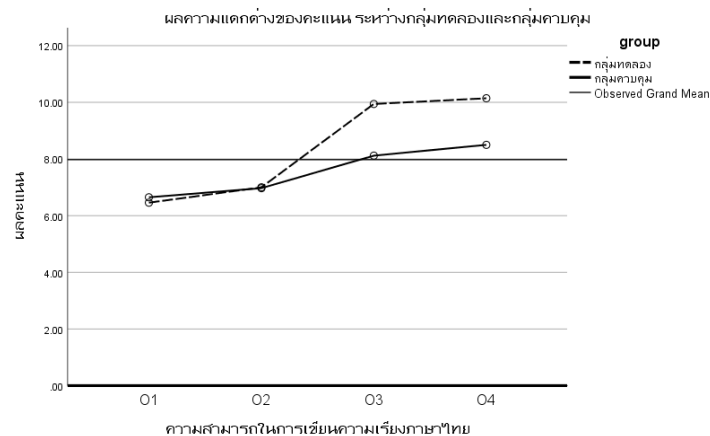
จากตาราง 2 ผลการทดสอบความสามารถในการเขียนความเรียงภาษาไทยจากการวัดต่างกัน ผลการวิเคราะห์พบว่า มีค่า $p = .029$ นั่นคือกลุ่มทดลองกับกลุ่มควบคุมมีคะแนนการทดสอบที่แตกต่างกันอย่างน้อย 1 คู่ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p = 0.00$) เมื่อพิจารณาความแตกต่างระหว่างกลุ่ม และความแตกต่างภายในกลุ่มแล้ว ผู้วิจัยจึงทำการเปรียบเทียบค่าเฉลี่ยรายคู่โดยใช้วิธี Bonferroni ดังตาราง 3

ตาราง 3 ผลการเปรียบเทียบรายคู่ของกลุ่มทดลองและกลุ่มควบคุม

(I) time	(J) time	Mean Difference (I-J)	Std. Error	p
ผลครั้งแรกก่อนการทดลอง (O_1)	2	-.433	.246	.498
	3	-2.478*	.239	.000
	4	-2.769*	.222	.000
ผลครั้งที่ 2 ก่อนการทดลอง (O_2)	1	.433	.246	.498
	3	-2.045*	.118	.000
	4	-2.336*	.129	.000
ผลครั้งแรกหลังการทดลอง (O_3)	1	2.478*	.239	.000
	2	2.045*	.118	.000
	4	-.291*	.088	.009
ผลครั้งที่ 2 หลังการทดลอง (O_4)	1	2.769*	.222	.000
	2	2.336*	.129	.000
	3	.291*	.088	.009

* $p < .05$

จากตาราง 3 ความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนกลุ่มทดลองและกลุ่มควบคุม ซึ่งทำการเปรียบเทียบกัน พบว่าผลครั้งแรกก่อนการทดลอง (O_1) กับผลครั้งที่ 2 ก่อนการทดลอง (O_2) มีความแตกต่างของคะแนน (mean difference) เท่ากับ $-.433$ คะแนน เมื่อทดสอบทางสถิติแล้ว พบว่าคะแนนของทั้งสองกลุ่มไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ $.05$ ($p = .498$) และความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนกลุ่มทดลองและกลุ่มควบคุมจากผลครั้งแรกหลังการทดลอง (O_3) และ ผลครั้งที่ 2 หลังการทดลอง (O_4) มีผลความแตกต่างของคะแนน (mean difference) เท่ากับ $-.291$ คะแนน เมื่อทดสอบทางสถิติแล้วพบว่าคะแนนของทั้งสองกลุ่มแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ $.05$ ($p = .009$) ดังภาพ 3



ภาพ 3 ความสามารถในการเขียนความเรียงภาษาไทยของกลุ่มทดลองและกลุ่มควบคุม

3. ผลการศึกษาความคงทนของความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนในกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ จากผลครั้งแรกหลังการทดลอง (O_3) และ ผลครั้งที่ 2 หลังการทดลอง (O_4) ดังตาราง 4

ตาราง 4 ผลการเปรียบเทียบรายคู่ของกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ

(I) time	(J) time	Mean Difference (I-J)	Std. Error	p
ผลครั้งแรกก่อนการทดลอง (O_1)	2	-.543	.344	.742
	3	-3.543*	.316	.000
	4	-3.686*	.330	.000
ผลครั้งที่ 2 ก่อนการทดลอง (O_2)	1	.543	.344	.742
	3	-3.000*	.071	.000
	4	-3.143*	.117	.000
ผลครั้งแรกหลังการทดลอง (O_3)	1	3.543*	.316	.000
	2	3.000*	.071	.000
	4	-.143	.093	.802
ผลครั้งที่ 2 หลังการทดลอง (O_4)	1	3.686*	.330	.000
	2	3.143*	.117	.000
	3	.143	.093	.802

* $p < .05$

จากตาราง 4 ผลการเปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยรายคู่ของกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ พบว่า ผลครั้งแรกหลังการทดลอง (O_3) และผลครั้งที่ 2 หลังการทดลอง (O_4) มีความแตกต่างของคะแนน (mean difference) เท่ากับ .143 คะแนน เมื่อทดสอบทางสถิติแล้ว พบว่า คะแนนของทั้งสองกลุ่มไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p = .802$) แสดงว่า ผลการศึกษาความคงทนของความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการมีความคงทน

อภิปรายผล

1. ผลการเปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 ก่อนและหลังการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ พบว่านักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการมีความสามารถในการเขียนความเรียงภาษาไทยสูงกว่าก่อนเรียน ซึ่งเป็นไปตามสมมติฐานที่ตั้งไว้ และสอดคล้องกับผลการวิจัยของ Chumjairak (2024) ที่ได้ศึกษาผลของการจัดการเรียนรู้โดยใช้แนวการสอนเขียนแบบเน้นกระบวนการร่วมกับเทคนิค SCAMPER ที่มีต่อความสามารถในการเขียนเชิงสร้างสรรค์ของนักเรียนชั้นประถมศึกษาปีที่ 6 ผลการวิจัยพบว่า ความสามารถในการเขียนเชิงสร้างสรรค์ของนักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวการสอนเขียนเน้นกระบวนการร่วมกับเทคนิค SCAMPER มีคะแนนหลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทั้งนี้อาจเป็นเพราะการสอนเขียนแบบเน้นกระบวนการเป็นการสอนเขียนที่ไม่ได้เน้นให้ผู้เรียนเกิดทักษะการเขียนอย่างเดียว แต่ผู้เรียนเกิดความสามารถในการนำเนื้อหามาเขียนเพื่อนำเสนอความคิดเพื่อสร้างสรรค์งานเขียนได้อย่างมีคุณภาพ เป็นตัวบ่งชี้ความสำเร็จของผู้เรียนถึงสามารถในการเขียนความเรียงภาษาไทย ซึ่งสอดคล้องกับที่ van den Bergh et al. (2016) ที่เสนอว่าแนวทางในการศึกษากระบวนการเขียนจะมุ่งเน้นการทำงานร่วมกันระหว่างทักษะต่าง ๆ เนื่องจากการเขียนเป็นกระบวนการที่มีความซับซ้อนและหลากหลายรูปแบบ ทั้งการคิด การวางแผน และการปรับปรุงผลงานเขียน ที่เกิดขึ้นอย่างต่อเนื่องและสัมพันธ์กัน

2. ผลการเปรียบเทียบความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 ระหว่างนักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการกับกลุ่มนักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติ พบว่านักเรียนที่ได้รับการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการ มีความสามารถในการเขียนความเรียงภาษาไทยสูงกว่ากลุ่มนักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติ ตามสมมติฐานที่ตั้งไว้ ที่สอดคล้องกับผลการวิจัยของ Yusri (2022) ที่ได้ศึกษาผลการจัดการเรียนรู้แบบการสอนเขียนที่เน้นกระบวนการร่วมกับเทคนิค CIRC ที่มีต่อผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย และความสามารถในการเขียนสื่อความของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ผลการวิจัยพบว่า ความสามารถในการเขียนสื่อความและผลสัมฤทธิ์ทางการเรียนสูงขึ้นหลังได้รับการจัดการเรียนรู้แบบการสอนเขียนที่เน้นกระบวนการร่วมกับเทคนิค CIRC ทั้งนี้เพราะการสอนเขียนแบบเน้นกระบวนการเป็นการสอนเขียนที่ให้ผู้เรียนได้ลงมือปฏิบัติและมีส่วนร่วมในการเขียนในทุกกระบวนการของการเขียน ผู้สอนจะเป็นผู้ให้คำแนะนำและช่วยเหลือในระหว่างการจัดการเรียนรู้ ซึ่งผู้เรียนต้องเขียนโครงร่าง และทำการตรวจสอบ แก้ไข ปรับปรุง

งานเขียนด้วยตนเองผ่านกิจกรรมกลุ่มเพื่อเป็นการแลกเปลี่ยนงานเขียนกับสมาชิกภายในกลุ่ม ซึ่งสอดคล้องกับ Hattie & Timperley (2007) ที่เสนอแนวทางในการให้ข้อมูลย้อนกลับแก่ผู้เรียนไว้ว่า เพื่อมุ่งอธิบายข้อดีหรือข้อเสียของนักเรียนและมุ่งระบุถึงประเด็นที่นักเรียนควรแก้ไขปรับปรุง ไม่ใช่การให้ข้อมูลย้อนกลับที่เป็นการตัดสินนักเรียนซึ่งในส่วนนี้จะทำให้นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่ำไม่มีแรงจูงใจในการเรียน

3. จากการศึกษาความคงทนของความสามารถในการเขียนความเรียงภาษาไทยของนักเรียนในกลุ่มที่ได้รับการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการ พบว่าความคงทนของนักเรียนกลุ่มทดลองมีแนวโน้มลดลงเพียงเล็กน้อยในการวัดครั้งที่ 4 ซึ่งมีผลคงที่ แสดงว่าทักษะการเขียนของกลุ่มทดลองหลังการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการมีความคงทน เป็นเพราะการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการช่วยฝึกให้ผู้เรียนเกิดความคิดในการวางแผนการวิเคราะห์เนื้อหาที่จะนำมาเขียน ซึ่งสอดคล้องกับ Atkinson & Shiffin (1968) ที่กล่าวว่า ความคงทนในการจำนั้นจะใช้เวลาประมาณ 14 วัน หลังจากที่ได้ผ่านการเรียนรู้ไปแล้ว ดังนั้น การสร้างความเคยชินในการทำงานเดิมซ้ำกันจะเป็นการส่งเสริมให้เกิดการเรียนรู้ในระยะยาวและจะช่วยให้ความจำถาวรมากยิ่งขึ้น

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 การจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการ ครูผู้สอนมีหน้าที่แนะนำสนับสนุน ช่วยเหลือให้ผู้เรียนเกิดอิสระทางความคิด และตัดสินใจด้วยตนเอง อีกทั้งกระบวนการสอนเขียนแบบเน้นกระบวนการต้องใช้เวลาที่ต่อเนื่องและค่อนข้างนานจึงจะส่งผลให้ผู้เรียนเกิดการพัฒนาความสามารถในการเขียนความเรียง เนื่องจากในการวิจัยครั้งนี้ผู้วิจัยได้กำหนดแผนการจัดการเรียนรู้จำนวน 5 แผน 10 ชั่วโมง ซึ่งใน 1 แผนจะใช้เวลา 2 ชั่วโมง แต่ในทางปฏิบัติบางสถานศึกษาอาจไม่สามารถออกแบบการสอนที่ต่อเนื่องได้ ดังนั้นครูผู้สอนที่สนใจวิธีการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการไปใช้ควรปรับขั้นตอนการจัดการเรียนการสอนให้เหมาะสมกับบริบทที่จะนำไปใช้ต่อไป

1.2 กระบวนการกลุ่ม มีความสำคัญอย่างยิ่งในการจัดการเรียนรู้โดยใช้แนวทางการเขียนแบบเน้นกระบวนการ เพราะในขั้นก่อนการเขียน ผู้เรียนจะต้องร่วมกันเสนอแนวคิดและวางแผนสิ่งที่จะต้องเขียนลงไป ผู้เรียนแต่ละกลุ่มจะต้องช่วยกันหาเทคนิคมาใช้เพื่อช่วยกันร่างงานเขียนให้สมบูรณ์ที่สุด ซึ่งจะไปสู่ชิ้นงานที่สำเร็จต่อไปได้ ดังนั้น กระบวนการกลุ่มในการฝึกเขียน จะช่วยให้ผู้เรียนพัฒนาทักษะการเขียนและการสื่อสารที่ดีขึ้น จากการร่วมกันคิดและวางแผนในการเขียน เพื่อให้ผู้เรียนค้นพบรูปแบบการเขียนที่เหมาะสมกับตนเอง ช่วยให้ผู้เรียนมีทักษะการเขียนที่ดีขึ้น

1.3 การให้ข้อมูลย้อนกลับแบบให้คำชี้แนะ (hints feedback) แก่ผู้เรียนที่จะเกิดขึ้นในการสอนเขียนแบบเน้นกระบวนการในขั้นการปรับปรุงไปจนถึงการแก้ไขงานเขียนนั้น จะทำให้ผู้เรียนทราบข้อผิดพลาดในงานเขียนของตนเอง และหากครูผู้สอนสนใจนำคู่มือการให้ข้อมูลย้อนกลับ คู่มือการตรวจแก้ไขงานเขียนแบบให้คำชี้แนะสำหรับการตรวจงานเขียนความเรียงภาษาไทยที่ผู้วิจัยสร้างขึ้นไปใช้ ควรเตรียมความพร้อมผู้เรียน และแนะนำการใช้คู่มือเพื่อให้ผู้เรียนเกิดความเข้าใจการใช้คู่มือที่ถูกต้องในการให้ข้อมูลย้อนกลับ

เพื่อสร้างความเคยชินในการทำงานเดิมซ้ำกันจนเป็นการส่งเสริมให้เกิดการเรียนรู้ในระยะยาว ผู้เรียนเกิดความรู้ที่คงทนถาวร

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 จากการศึกษาผลการจัดการเรียนรู้โดยใช้แนวคิดการเขียนแบบเน้นกระบวนการที่มีต่อความสามารถในการเขียนความเรียงภาษาไทย ควรนำไปศึกษาต่อในด้านการส่งผลต่อความสามารถในการเขียนรูปแบบอื่น ๆ เช่น การเขียนเชิงวิชาการ การเขียนเรียงความ การเขียนเชิงสร้างสรรค์ การเขียนตามจินตนาการ เป็นต้น เพื่อให้ผู้เรียนได้ฝึกทักษะการเขียนในหลากหลายรูปแบบมากขึ้น

2.2 การให้ข้อมูลย้อนกลับในกระบวนการจัดการเรียนรู้ มีส่วนสำคัญอย่างมากที่จะทำให้ผู้เรียนมีความเข้าใจมากยิ่งขึ้นถึงข้อผิดพลาดของตนเอง ดังนั้นการให้ข้อมูลย้อนกลับแบบอธิบายผลของคำตอบที่ถูกต้อง (response contingent feedback) จึงเป็นอีกหนึ่งวิธีในการให้ข้อมูลย้อนกลับในลักษณะอธิบายรายละเอียด ที่ควรนำมาศึกษาเพิ่มเติม เพราะการให้ข้อมูลย้อนกลับที่มีประสิทธิภาพจะช่วยพัฒนาผู้เรียนในหลายด้าน

References

- Atkinson, R. C., & Shiffrin, R. M. (1968). *Human Memory: A Proposed System and Its Control Processes*. Academic Press.
- Atkinson, D. (2003). L2 writing in the post-process era: Introduction. *Journal of Second Language Writing*, 12, 3-15.
- Brown, K., & Hood, S. (1989). *Writing matters: writing skills and strategies for students of English*. Cambridge University Press.
- Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences. *Behavior Research Methods*, 39, 175-191.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Mason, B. J., & Bruning, R. H. (2001). Providing Feedback in Computer-based Instruction: What the Research Tells Us. *CLASS Research Report*, 9, 1-20.
- Onozawa, C. (2010). A Study of the Process Writing Approach: A Suggestion for an Electric Writing Approach. *Research note*, 10, 153-163.
- Raimes, A. (1983). *Techniques in Teaching Writing*. Oxford University Press.
- Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

- van den Bergh, H., Rijlaarsdam, G., & van Steendam, E. (2016). Writing process theory: A functional dynamic approach. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 57–71). The Guilford Press.
- White, R., & Arndt, V. (1991). *Process Writing*. Harlow Longman.
- Whitney, D. R., & Sabers, D. L. (1970). *Improving essay examinations III: Use of item analysis*, Technical Bulletin II [Mimeographed]. University Evaluation and Examination Service.

Translate Thai References

- Bureau of Academic Affairs and Educational Standards, Office of the Basic Education Commission. (2017). *Indicators and core curriculum content, Science group (Revised edition BE 2560) according to the core curriculum for basic education, BE 2551*. Chumchon Sahakhon Community Printing and Agricultural Cooperative Co., Ltd. (in Thai)
- Chumjairak, J. (2024). *Effects of using the Process Writing Approach Along with The SCAMPER Technique on Creative Writing Ability for Prathomsuksa six students* [Master's thesis]. Srinakharinwirot University. (in Thai)
- National Institute of Educational Testing Service. (2024). *Ordinary National Educational. Test: O-NET*. <https://newonetestresult.niets.or.th/AnnouncementWeb/Login.aspx> (in Thai)
- National Institute of Educational Testing Service. (2024). *Report of Findings: The Use of Thai Language by Students from the O-NET Test with Essay-Type Questions in the Academic Year 2023, Grade 6*. <https://www.niets.or.th/th/catalog/view/497/70> (in Thai)
- Pattrawiwat, K. (2016). Interrupted Time Series Design: Evaluation of the Intervention for Applications in Behavioral Science Research. *Journal of Behavioral Science*, 22(2), 1-16. (in Thai)
- Rueangkitchanan, A. (2016). *Development of an instructional model by integrating cognitive apprenticeship approach and process writing approach for enhancing expository writing ability and reflective thinking ability of undergraduate students* [Doctoral dissertation]. Chulalongkorn University. (in Thai)
- Sumniengngam, S. (2021). *Thai Language for Humanities*. Department of Thai, Faculty of Arts, Silpakorn University. (in Thai)
- Thaiuboon, D. (2013). *Skills in Thai writing* (8nd ed.). Darn Sutha Press Company Limited. (in Thai)

- Thipyotha, P. (2023). *The Development of Learning Activity Management for Creative Writing by Using Concept Map, Thai Language Learning Substance, Pratomsuksa 6* [Master's thesis]. Mahasarakam Rajabhat University. (in Thai)
- Yusri, K. (2022). *Effects of Writing Process Approach with CIRC Technique on Thai Language Learning Achievements and Ability on Communication Writing of grade 7 students* [Master's thesis]. Valaya Alongkorn Rajabhat University Under the Royal Patronage Pathum Thani. (in Thai)

การพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติในรูปแบบแอปพลิเคชัน สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ

Development of a Multidimensional Mathematical Literacy Test in an Application Platform for Grade 9 Students with Feedback Integration

ปราณี แก้วมา^{1*} น้าทิพย์ องอาจวานิชย์² และ สำราญ มีแจ้ง²

Pranee Kaewma^{1*}, Namthip Ongardwanich² and Samran Mejang²

(Received: December 14, 2024; Revised: February 2, 2025; Accepted: February 10, 2025)

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติในรูปแบบแอปพลิเคชัน สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ กลุ่มผู้ให้ข้อมูล คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 483 คน สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาตาก จังหวัดตาก ได้มาโดยการสุ่มแบบหลายขั้นตอน ผลการพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ พบว่า ข้อสอบที่ได้มีคุณภาพในด้านความตรงเชิงเนื้อหาโดยให้ผู้เชี่ยวชาญ 5 คน พิจารณาความสอดคล้องของเนื้อหาและความเหมาะสม คัดเลือกข้อสอบที่ผ่านเกณฑ์ ไว้ 30 ข้อ มีค่า IOC .60 ถึง 1.00 วิเคราะห์ข้อมูล โดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติใช้โปรแกรม NOHARM วิเคราะห์พารามิเตอร์ของข้อสอบด้วยโมเดลปกติสะสมแบบพหุมิติ ผลการวิเคราะห์ พบว่า ข้อสอบมีค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.247 ถึง 3.029 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -1.023 ถึง 2.830 และมีค่าฟังก์ชันสารสนเทศ (IRT PRO) ของข้อสอบ ส่วนใหญ่มีค่าสารสนเทศสูง ณ ระดับความสามารถปานกลาง

คำสำคัญ ความฉลาดรู้ด้านคณิตศาสตร์ แบบวัดพหุมิติ แอปพลิเคชันการศึกษา การให้ข้อมูลย้อนกลับ

¹ นิสิตปริญญาเอก สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้ คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร

² รองศาสตราจารย์ สาขาวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร

¹ Doctoral Degree's student, Innovation in Learning Measurement Program, Faculty of Education, Naresuan University

² Associate Professor, Education Research and Evaluation Program, Faculty of Education, Naresuan University

* Corresponding Author E-mail: praneek65@nu.ac.th

Abstract

This research aimed to develop a Multidimensional Mathematical Literacy Test in an application platform for Grade 9 students with integrated feedback. The sample consisted of 483 students from the Secondary Educational Service Area Office in Tak Province, selected through a multi-stage sampling method. The results from developing the test using Multidimensional Item Response Theory (MIRT) indicated that the assessment demonstrated high quality in terms of validity. A panel of five experts evaluated the content for coherence and relevance, with 30 items meeting the established criteria and achieving Item-Objective Congruence (IOC) values ranging from .60 to 1.00. The data were analyzed using MIRT with the NOHARM software, applying a multivariate cumulative normal model. The analysis revealed that the test exhibited multidimensional discrimination (MDISC) values ranging from 0.247 to 3.029, and difficulty (MDIFF) values ranging from -1.023 to 2.830. The item information function indicated that the test provided the most information for students within the moderate ability range.

Keywords: mathematical literacy, multidimensional assessment, educational application, feedback

บทนำ

องค์การเพื่อความร่วมมือทางเศรษฐกิจและการพัฒนา (OECD, 2018) กล่าวถึงความฉลาดรู้ทางคณิตศาสตร์ (mathematical literacy) ว่าเป็น ความสามารถของแต่ละบุคคลในการให้เหตุผลทางคณิตศาสตร์ และสามารถแปลงปัญหา ใช้คณิตศาสตร์และตีความผลลัพธ์ทางคณิตศาสตร์เพื่อใช้ในการแก้ปัญหาในบริบทของโลกชีวิตจริง รวมถึงการใช้แนวคิด กระบวนการ ข้อเท็จจริง และเครื่องมือทางคณิตศาสตร์เพื่อบรรยาย อธิบาย และคาดการณ์ปรากฏการณ์ต่าง ๆ โดยสิ่งเหล่านี้จะช่วยให้แต่ละบุคคลทราบถึงบทบาทของคณิตศาสตร์ที่มีต่อโลกนี้และสร้างพื้นฐานที่ดีในการลงข้อสรุปสำหรับพลเมืองในศตวรรษที่ 21 นอกจากนี้ยังกล่าวถึงการประเมิน PISA ที่คำนึงถึงการเปลี่ยนแปลงอย่างรวดเร็วของเทคโนโลยีใหม่ ๆ ที่เข้ามามีบทบาทในชีวิตมากขึ้นและการพัฒนาทักษะที่จำเป็นสำหรับศตวรรษที่ 21 ด้วยเหตุนี้จึงทำให้มีการเน้นความสามารถในการให้ความฉลาดรู้ด้านคณิตศาสตร์ การคิดวิเคราะห์อย่างมีเหตุผลเป็นขั้นตอน รวมถึงบริบทที่สอดคล้องกับทักษะแห่งศตวรรษที่ 21 เข้ามาเป็นส่วนหนึ่งของกรอบการประเมิน

จากทางสถาบันวิจัยเพื่อการพัฒนาประเทศไทย (TDRI) ซึ่ง ผลสอบ PISA เด็กไทยมีประสิทธิภาพต่ำลงทุกทักษะ สะท้อนวิกฤตการศึกษา ต้องเร่งแก้ปัญหาให้ถูกจุด โดยผลสอบระดับโปรแกรมประเมินสมรรถนะนักเรียนมาตรฐานสากล หรือ PISA ประจำปี 2022 ของเด็กไทยที่มีคะแนนสอบต่ำสุดในรอบ 20 ปีในทุกทักษะ ทั้งด้านการอ่าน คณิตศาสตร์ และวิทยาศาสตร์ พร้อมทั้งชี้ว่าผลสอบ PISA ครั้งนี้กำลังสะท้อนสัญญาณเตือน

ถึงวิกฤตทางการศึกษาของไทยที่ต้องเร่งแก้ปัญหาให้ถูกจุด ภาพรวมผล PISA เด็กไทยในช่วง 10 ปีล่าสุดมีส่วนต่างคะแนน PISA ของเด็กไทย เมื่อเทียบกับกลุ่มประเทศ OECD ห่างจากกลุ่มประเทศ OECD แล้วเกือบ 100 คะแนน และมีแนวโน้มห่างจากคะแนนเฉลี่ยมากขึ้นเรื่อย ๆ (Manochanpen, 2022) และจากผลการประเมินสมรรถนะนักเรียนมาตรฐานสากลของเด็กไทย PISA 2022 พบว่า นักเรียนไทยมีคะแนนเฉลี่ยด้านคณิตศาสตร์ 394 คะแนน วิทยาศาสตร์ 409 คะแนน และการอ่าน 379 คะแนน ซึ่งเมื่อเทียบกับ PISA 2018 พบว่า คะแนนเฉลี่ยของประเทศไทยทั้งสามด้านลดลง โดยคณิตศาสตร์มีคะแนนเฉลี่ยลดลง 25 คะแนน วิทยาศาสตร์และการอ่าน มีคะแนนเฉลี่ยลดลง 17 คะแนน และ 14 คะแนน ตามลำดับ

จากการวัดคุณลักษณะภายในที่เกี่ยวกับความสามารถของบุคคลเป็นการวัดที่ต้องอาศัยข้อสอบ หรือแบบสอบเป็นสิ่งเร้าเพื่อกระตุ้นให้แสดงพฤติกรรมที่ต้องการออกมาด้วยการตอบสนองต่อข้อสอบ ผลการตอบจะออกมาในรูปของคะแนนแล้วอ้างอิงไปอธิบายหรือทำนายความสามารถที่แท้จริงของบุคคล การที่จะสามารถอธิบายหรือทำนายความสามารถของบุคคลได้ดีเพียงใดนั้นขึ้นอยู่กับคุณภาพของข้อสอบหรือแบบสอบว่าให้ผลการวัดที่น่าเชื่อถือและครอบคลุมคุณลักษณะที่ต้องการวัดมากน้อยเพียงใด ทั้งนี้ข้อสอบหรือแบบสอบที่มีคุณภาพถือว่าเป็นเครื่องมือที่ดีในการวัดพฤติกรรมที่เป็นตัวแทนของมวลพฤติกรรมที่ต้องการศึกษาทั้งหมด ดังนั้นจากอดีตจนถึงปัจจุบันจึงได้มีการพัฒนาคุณภาพของข้อสอบและแบบสอบอย่างต่อเนื่อง ซึ่งการพัฒนาดังกล่าวสามารถแยกตามโครงสร้างของแบบสอบ (test structure) 2 กลุ่มใหญ่ คือ โครงสร้างแบบสอบแบบเอกมิติ (unidimensional test structure) และโครงสร้างแบบสอบแบบพหุมิติ (multidimensional test structure) ซึ่งการพัฒนาข้อสอบหรือแบบสอบตามโครงสร้างทั้งสองแบบนี้ ดำเนินการอยู่ภายใต้แนวคิดทฤษฎีการทดสอบอันเป็นองค์ความรู้ที่เป็นประโยชน์เกี่ยวกับการทดสอบ วิธีการ แก้ปัญหาการทดสอบและการพัฒนาเครื่องมือการทดสอบ ซึ่งจะสามารถช่วยให้นักวัดผลสามารถทำการสร้าง และพัฒนาแบบสอบให้มีคุณภาพ สามารถแปลความหมายผลการวัดได้ถูกต้องและสามารถนำเสนอสารสนเทศไปใช้ สำหรับการตัดสินใจได้อย่างเหมาะสม (Kanjawasee, 2009) จะเห็นว่าการพัฒนาคุณภาพของข้อสอบและแบบสอบนั้นต้องอาศัยทฤษฎีทางการทดสอบเพื่อเป็นแนวทางในการพัฒนา โดยทฤษฎีการทดสอบที่นำมาใช้ในการพัฒนาข้อสอบและแบบสอบ ได้แก่ ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT) ทฤษฎีการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional Item Response Theory : UIRT) และ ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory : MIRT) ซึ่งแต่ละทฤษฎีมีคุณลักษณะที่ต่างกันไป โดย CTT มีแนวคิดในการพัฒนาที่ เข้าใจง่ายไม่ซับซ้อนแต่สามารถนำไปใช้ได้เฉพาะกลุ่ม ส่วน UIRT สามารถประมาณความสามารถที่แท้จริงของผู้สอบได้ แต่อยู่ภายใต้เงื่อนไขของความเป็นเอกมิติของข้อสอบ และ MIRT เป็นทฤษฎีที่ต้องการศึกษาประเด็น การประเมินที่มีความซับซ้อน

จะเห็นว่าทฤษฎีการทดสอบที่พัฒนามาก่อนจะส่งผลให้ข้อสอบหรือแบบสอบมี คุณภาพต่ำกว่า ทฤษฎีการทดสอบที่พัฒนาขึ้นมาทีหลังเพื่อผ่อนคลายข้อตกลงเบื้องต้นของทฤษฎีเดิม แต่อย่างไรก็ตามในปัจจุบันยังมีการนำเอาแนวคิดทฤษฎีทั้งเก่าและใหม่มาใช้ในการพัฒนาข้อสอบและแบบสอบเพื่อใช้ประโยชน์ในทางการวัดผลและประเมินผลให้เห็นอยู่อย่างต่อเนื่อง ดังนั้นถ้าต้องการพัฒนาข้อสอบหรือแบบสอบที่ให้สารสนเทศ ไม่ว่าจะเป็นค่าความยาก ค่าอำนาจจำแนก หรือ ค่าพารามิเตอร์ที่แยกตามมิติ ในแต่ละข้อนักวิจัย

หรือนักวัดผล จึงควรพัฒนาแบบทดสอบพหุมิติที่ให้ผลการวัดที่ดีกว่า (Tayraukhum, n.d) ดังนั้นผู้วิจัยจึงเห็นว่าแบบทดสอบแบบพหุมิติให้ผลการวัดที่ดีกว่า และด้วยยุคที่เทคโนโลยีเข้ามามีบทบาทในชีวิตประจำวัน ทำให้การเรียนการสอนมีการนำเอาเทคโนโลยีเข้ามาเกี่ยวข้องในรูปแบบของสื่อต่าง ๆ ทั้งวีดิทัศน์ คอมพิวเตอร์ ช่วยสอน รวมไปถึงแอปพลิเคชันโดยการใช้สื่อให้เหมาะสมกับผู้เรียนจะช่วยให้ผู้เรียนสามารถเรียนรู้ได้อย่างมีประสิทธิภาพ ซึ่งทำให้เกิดการพัฒนาสื่อการสอนในรูปแบบแอปพลิเคชันบนสมาร์ตโฟน หรือแท็บเล็ต โดยเป็นอีกช่องทางที่ทำให้การเรียนรู้ไม่จำกัดอยู่แต่ภายในห้องเรียน อีกทั้งยังสะดวกสามารถทบทวน บทเรียนได้ทุกที่ตามต้องการ จึงมีความจำเป็นที่จะต้องใช้สื่อการเรียนรู้ที่เหมาะสมแก่บุคคลเหล่านี้ (Khlambunkrong, 2012)

รวมไปถึงกับความสำคัญของข้อมูลย้อนกลับ ซึ่งข้อมูลย้อนกลับ (feedback) เป็นสิ่งที่มีอิทธิพลต่อการเรียนรู้และผลสัมฤทธิ์ทางการเรียนแต่ขณะเดียวกันข้อมูลย้อนกลับก็มีผลกระทบในด้านบวกและด้านลบด้วย ในบริบทของการจัดการเรียนการสอนเพื่อให้ประสิทธิภาพในการจัดการเรียนการสอนควรมีการให้ข้อมูลย้อนกลับด้วย โดยการให้ข้อมูลย้อนกลับอาจจะให้แก่ผู้เรียนหลังจากที่ได้ตอบสนองต่อการเรียนรู้ในครั้งแรก เมื่อได้ให้ข้อมูลย้อนกลับแก่ผู้เรียนในมิติความรู้ทักษะต่าง ๆ ที่ได้เรียนไปแล้วโดยเฉพาะอย่างยิ่งหากเป็นการให้ข้อมูลย้อนกลับเพื่อชี้แจงความผิดพลาดที่เกิดขึ้นจะทำให้ผู้เรียนเกิดความรู้ความเข้าใจที่ดียิ่งขึ้น (Masantiah, 2017) จากการศึกษาข้อมูล เอกสาร และงานวิจัยต่างๆ ที่ได้กล่าวมาข้างต้นผู้วิจัยได้เล็งเห็นถึงปัญหาและความสำคัญดังกล่าวข้างต้นจึงมีความสนใจในการพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติในรูปแบบแอปพลิเคชันสำหรับนักเรียนชั้นมัธยมศึกษาชั้นปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ มาพัฒนาเพิ่มในจุดที่ยังต้องพัฒนา โดยทางผู้วิจัยได้นำรายวิชาคณิตศาสตร์ ในการทำการทดสอบประสิทธิภาพระบบแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ ในรูปแบบแอปพลิเคชันเพื่อให้ระบบสามารถตอบสนองผู้ที่มีความน่าสนใจ ใช้งานง่าย ให้มีเสถียรภาพ เชื่อถือได้และมีประสิทธิภาพมากยิ่งขึ้น ซึ่งมีประโยชน์แก่สถาบันมากมาย ผู้เรียนสามารถทราบถึงผลคะแนน และผลการให้ข้อมูลย้อนกลับด้านความฉลาดรู้ทางคณิตศาสตร์ สามารถรองรับรูปแบบข้อสอบที่มีอยู่อย่างหลากหลายในปัจจุบัน เพื่อที่ระบบที่พัฒนาขึ้นจะได้เป็นมาตรฐานในการเพิ่มประสิทธิภาพการทดสอบเพื่อวัดผลประเมินผลการเรียนต่อไป

ความมุ่งหมายในการวิจัย

เพื่อพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ ในรูปแบบแอปพลิเคชัน สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติในรูปแบบแอปพลิเคชันสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ ผู้วิจัยได้ดำเนินการโดยแบ่ง เป็น 2 ระยะ มีรายละเอียด ดังนี้

ระยะที่ 1 ศึกษาองค์ประกอบและตัวบ่งชี้ความฉลาดรู้ด้านคณิตศาสตร์สำหรับนักเรียน ชั้นมัธยมศึกษาปีที่ 3

ขั้นที่ 1.1 การศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง องค์ประกอบและตัวบ่งชี้ความฉลาดรู้ด้านคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยสังเคราะห์เอกสารที่เกี่ยวข้อง ดำเนินการ ดังนี้

1.1.1 แหล่งข้อมูล ได้แก่ เอกสารและงานวิจัยที่เกี่ยวข้อง เพื่อกำหนดเป็นกรอบแนวคิดในการวิจัย โดยการศึกษาหลักการ องค์ประกอบ แนวคิดและสาระสำคัญเกี่ยวกับความฉลาดรู้ด้านคณิตศาสตร์ และแนวคิดเกี่ยวกับการพัฒนาความฉลาดรู้ด้านคณิตศาสตร์ของนักเรียน

1.1.2 ประเด็นที่ศึกษา คือ องค์ประกอบและตัวบ่งชี้ความฉลาดรู้ด้านคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

ขั้นที่ 1.2 การศึกษาแนวทางการศึกษาองค์ประกอบและตัวบ่งชี้ความฉลาดรู้ด้านคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ดำเนินการ ดังนี้

1.2.1 กลุ่มผู้ให้ข้อมูล ได้แก่ ผู้เชี่ยวชาญด้านวัดและประเมินผล ผู้เชี่ยวชาญที่เกี่ยวข้องและผู้เชี่ยวชาญด้านเนื้อหา รวม จำนวน 5 คน ได้มาโดยการเลือกแบบเจาะจง (purposive sampling)

1.2.2 ประเด็นที่ศึกษา ได้แก่ องค์ประกอบและตัวบ่งชี้ความฉลาดรู้ด้านคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3

ระยะที่ 2 การพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติในรูปแบบแอปพลิเคชันสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ

ขั้นที่ 2.1 การพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ ในรูปแบบแอปพลิเคชันสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ ดำเนินการดังนี้

2.1.1 ผู้วิจัยนำข้อมูลจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง และผลการศึกษาความฉลาดรู้ด้านคณิตศาสตร์และศึกษาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์ แบบพหุมิติในรูปแบบแอปพลิเคชันแบบให้ข้อมูลย้อนกลับ เพื่อส่งเสริมสมรรถนะความฉลาดรู้ด้านคณิตศาสตร์ที่จำเป็นของนักเรียนในสถานศึกษา ผู้วิจัยนำร่างรูปแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ ในรูปแบบแอปพลิเคชันสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ ที่สร้างขึ้นเสนออาจารย์ที่ปรึกษาเพื่อพิจารณาความเหมาะสมและการใช้ภาษา

2.1.2 ผู้วิจัยนำข้อเสนอแนะจากอาจารย์ที่ปรึกษาไปดำเนินการปรับปรุงแก้ไข จัดทำเป็นร่างแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติในรูปแบบแอปพลิเคชันสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ

2.1.3 ผู้วิจัยจัดทำร่างแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ ในรูปแบบแอปพลิเคชันสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ

2.1.4 สร้างเมตริกซ์คิว (Q-matrix) โดยระบุมิติที่ต้องการวัดตามองค์ประกอบของความฉลาดรู้ด้านคณิตศาสตร์

ขั้นที่ 2.2 การตรวจสอบคุณภาพแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติในรูปแบบแอปพลิเคชันสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ

2.2.1 นำแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ ให้ผู้เชี่ยวชาญด้านวัดและประเมินผล ผู้เชี่ยวชาญที่เกี่ยวข้อง และผู้เชี่ยวชาญด้านเนื้อหา รวม จำนวน 5 คน พิจารณาความสอดคล้องของเนื้อหาและความเหมาะสมสร้างแบบทดสอบแบบพหุมิติตามองค์ประกอบของความฉลาดรู้ด้านคณิตศาสตร์ ประกอบไปด้วยมิติสถานการณ์และบริบท มิติเนื้อหาของคณิตศาสตร์ และมิติความฉลาดรู้ด้านคณิตศาสตร์ โดยผู้วิจัยเป็นผู้สร้างข้อสอบ 40 ข้อ เพื่อคัดข้อสอบที่ผ่านเกณฑ์ ให้ได้ 30 ข้อ ดังตาราง 1

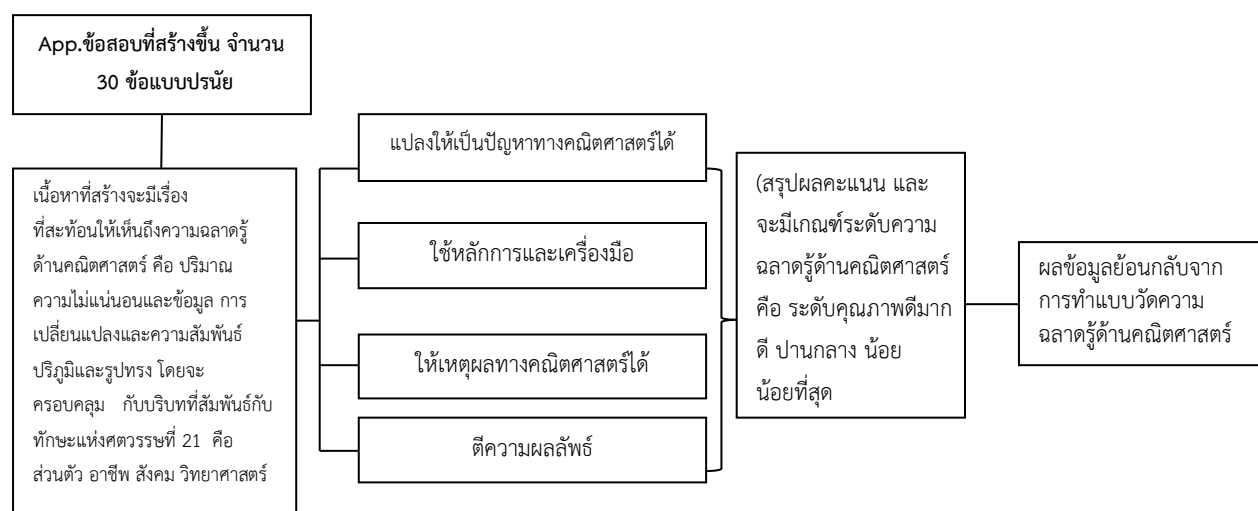
ตาราง 1 ตารางวิเคราะห์แบบทดสอบวัดความฉลาดรู้ด้านคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3

สาระ	มาตรฐาน	ตัวชี้วัด	องค์ประกอบและตัวบ่งชี้ความฉลาดรู้ด้านคณิตศาสตร์ (จำนวนข้อ)																	
			1. ความฉลาดรู้ด้านคณิตศาสตร์				2. เนื้อหาคณิตศาสตร์ที่ใช้ในการแก้ปัญหา				3. บริบทที่สัมพันธ์กับทักษะแห่งศตวรรษที่ 21				รวม					
			การคิด/แปลงปัญหา	การใช้คณิตศาสตร์	การตีความและประเมิน	การให้เหตุผลทางคณิตศาสตร์	ปริมาณ	ความไม่แน่นอนและข้อมูล	การเปลี่ยนแปลงและความสัมพันธ์	ปริภูมิและรูปทรง	ส่วนตัว	อาชีพ	สังคม	วิทยาศาสตร์						
จำนวนและพีชคณิต	ค 1.2	1.2.1	5	4	2	2	4	1	2	1	ใน 1 ข้อ มีหลายมิติ มิติบริบท มิติเนื้อหา คณิตศาสตร์ และมิติความฉลาดรู้ด้านคณิตศาสตร์									
		1.2.2																		
	ค 1.3	1.3.1																		
		1.3.2	5	4	2	2	5	1	2	1										
		1.3.3																		
การวัดและเรขาคณิต	ค 2.1	2.1.1	5	4	3	2	6	1	1	1										
		2.1.2																		
	ค 2.2	2.2.1	5	3	3	2	4	1	2	1										
		2.2.2																		
สถิติและความน่าจะเป็น	ค 3.1	3.1.1	5	3	2	2	5	1	2	1										
	ค 3.2	3.2.1	1	4	2	2	4	1	2	1										
รวมจำนวนข้อสอบทั้งหมด			30	22	14	12	28	6	11	6	8	9	6	7	40					

2.2.2 ให้ผู้เชี่ยวชาญด้านวัดและประเมินผล ผู้เชี่ยวชาญที่เกี่ยวข้อง และผู้เชี่ยวชาญด้านเนื้อหา ตรวจสอบความตรง (validity) และตรวจสอบความเหมาะสมของการระบุมิติตามองค์ประกอบความฉลาดรู้ด้านคณิตศาสตร์ และคัดเลือกข้อสอบที่มีค่า IOC ตั้งแต่ 0.50 ขึ้นไป โดยใช้แบบประเมินความคิดเห็นของผู้เชี่ยวชาญ เกี่ยวกับความตรงเชิงเนื้อหา ความเหมาะสมของการระบุ เนื้อหาคณิตศาสตร์ของข้อสอบเนื้อหาความฉลาดรู้ด้านคณิตศาสตร์ ในการเก็บรวบรวมข้อมูลตรวจสอบความเหมาะสมของ Q-matrix คัดข้อที่มีค่าอำนาจจำแนก .20 ขึ้นไป ค่าความยาก .20 - .80 ค่าความเที่ยง (reliability) .75 (Tayraukhum, 2010)

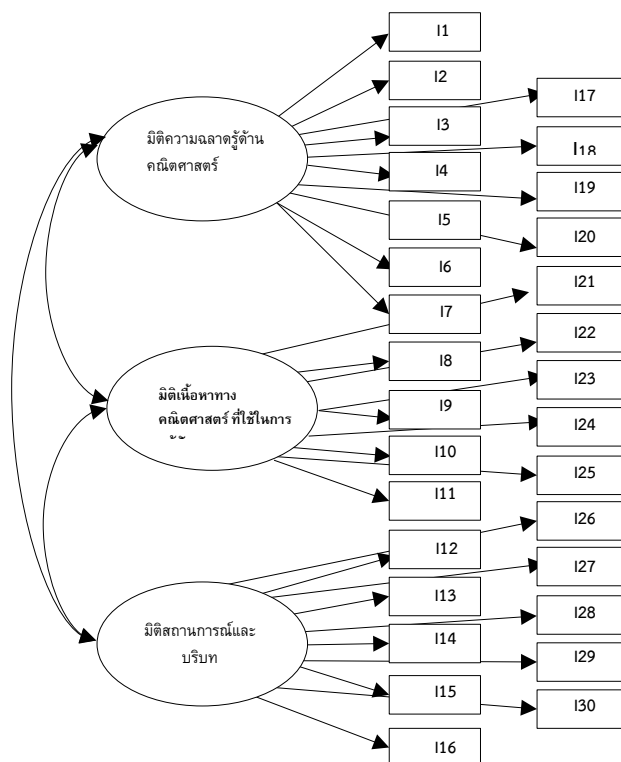
2.2.3 กลุ่มผู้ให้ข้อมูล ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 483 คน ได้มาโดยการสุ่มแบบหลายขั้นตอน (multi-stage sampling) โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาตาก จังหวัดตาก จำนวน 16 โรงเรียน คือ โรงเรียนตากพิทยาคม โรงเรียนผดุงปัญญา โรงเรียนวังประจวบวิทยาคม โรงเรียนบ้านตาก “ประชาวิทยาการ” โรงเรียนยกระบัตริวิทยาคม โรงเรียนนาโบสถ์พิทยาคม โรงเรียนแม่ปะวิทยาคม โรงเรียนอุ้มผางวิทยาคม โรงเรียนสรรพวิทยาคม โรงเรียนทุ่งฟ้าวิทยาคม โรงเรียนวังหินกิตติวิทยาคม โรงเรียนวังเจ้าวิทยาคม โรงเรียนถนนอมราษฎรบำรุง โรงเรียนสามเงาวิทยาคม โรงเรียนพบพระวิทยาคม และโรงเรียนแม่กุวิทยาคม

2.2.4 นำแบบทดสอบที่สร้างขึ้นไปทดลอง กับกลุ่มผู้ให้ข้อมูล จำนวน 483 คน และนำข้อมูลที่ได้อามาวิเคราะห์ข้อสอบรายข้อเพื่อหาคุณภาพของแบบทดสอบ โดยทำการหาคุณภาพของแบบทดสอบตามการวัดโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติเพื่อหาความเที่ยงความยาก และค่าอำนาจจำแนกตามเกณฑ์ที่กำหนด ตามตารางการวิเคราะห์คุณภาพของแบบทดสอบจากนั้นคัดเลือกข้อสอบที่เป็นไปตามเกณฑ์และนำไปพัฒนาในรูปแบบของแอปพลิเคชัน ด้วยการเขียนโปรแกรมด้วยภาษา PHP เพื่อออกแบบสร้างแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติในรูปแบบแอปพลิเคชันสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ ซึ่งเป็นแบบปรนัย ดังภาพ 1 เพื่อนำไปใช้ในการทดสอบกับกลุ่มตัวอย่าง



ภาพ 1 การสร้างแบบวัดความฉลาดรู้ด้านคณิตศาสตร์ด้วยแพลตฟอร์ม PHP

โดยในการวิจัยครั้งนี้ผู้วิจัยได้เลือกที่จะใช้วิธีการวิเคราะห์โมเดลการตอบสนองข้อสอบพหุมิติภายในข้อสอบ (within-items multidimensional IRT Model) มีลักษณะโครงสร้างเป็นองค์ประกอบเชิงซ้อนที่จะช่วยลดข้อคำถามให้น้อยลงแต่ถามได้ ครอบคลุมมากขึ้น โดยความสามารถแต่ละด้านถูกวัดภายในข้อสอบชุดเดียวกัน (within-items) แสดงดังภาพ 2



ภาพ 2 โมเดลการวัดระดับความรู้ด้านคณิตศาสตร์แบบพหุมิติระหว่างข้อสอบ

จากภาพ 2 เป็นโมเดลการตอบสนองข้อสอบแบบ 3 มิติ ซึ่งความสามารถทั้ง 3 ด้าน มีความเกี่ยวข้องสัมพันธ์กันโดยความสามารถแต่ละด้านถูกวัดภายในข้อสอบชุดเดียวกัน (within-items) ในงานวิจัยนี้จะวัด 3 มิติ คือ 1. มิติความรู้ด้านคณิตศาสตร์ 2. มิติเนื้อหาทางคณิตศาสตร์ที่ใช้ในการแก้ปัญหา และ 3. มิติสถานการณ์และบริบท

ผลการวิจัย

การพัฒนาแบบวัดความรู้ด้านคณิตศาสตร์แบบพหุมิติ พบว่า

1.1 ความตรงของแบบวัดความรู้ด้านคณิตศาสตร์

ผลการพิจารณาความสอดคล้องของข้อสอบกับเนื้อหาและความเหมาะสมของการระบุมิตติกับข้อสอบโดยผู้เชี่ยวชาญ ได้มาจากการที่ผู้วิจัย สร้างข้อสอบเพื่อพัฒนาแบบวัดความรู้ด้านคณิตศาสตร์ โดยสร้างแบบวัดแบบเลือกตอบ 4 ตัวเลือก จำนวน 40 ข้อ และให้ผู้เชี่ยวชาญจำนวน 5 คน พิจารณาความสอดคล้องของข้อสอบ ผลการพิจารณาความสอดคล้องกับเนื้อหา และความเหมาะสมในการระบุมิติความรู้ด้านคณิตศาสตร์ พบว่า แบบวัดความรู้ด้านคณิตศาสตร์แบบพหุมิติจำนวน 40 ข้อ มีความสอดคล้องกับเนื้อหาและความเหมาะสมในการระบุมิติ ผ่านเกณฑ์ไว้เพื่อจัดทำแบบทดสอบ จำนวน 30 ข้อ มีค่า IOC เท่ากับ .60 -1.00

1.2 ค่าอำนาจจำแนก ค่าความยากง่าย พบว่า คุณภาพของแบบทดสอบที่วิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยโมเดลปกติสะสม แบบพหุมิติ ด้วยโปรแกรม NOHARM (Normal Ogive by Harmonic Analysis Robust Method) ผลค่าพารามิเตอร์ ค่าอำนาจจำแนกและค่าความยากง่ายของแบบวัดความฉลาดรู้ด้านคณิตศาสตร์ ดังตาราง 2

ตาราง 2 แสดงค่าพารามิเตอร์ของแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ ค่าจุดตัดความยากแบบพหุมิติ ค่าอำนาจจำแนกแบบพหุมิติ และค่าความยากแบบพหุมิติ

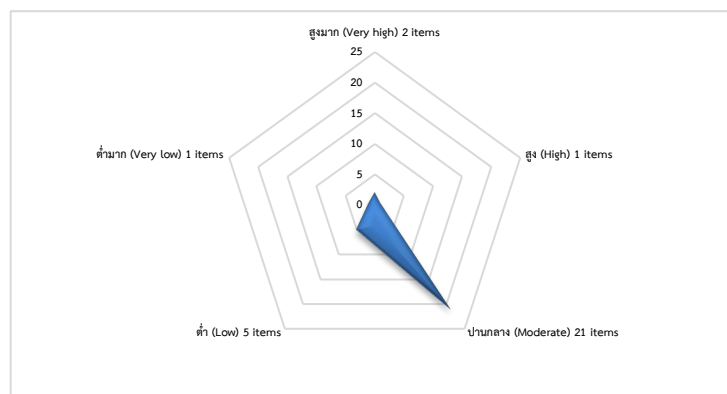
ITEMS	มิติของข้อสอบ			d threshold	MDISC	MDIFF
	a_1	a_2	a_3			
1	0.400	-0.247	0.283	0.817	0.549	1.488
2	-0.027	0.236	-0.066	0.699	0.247	2.830
3	0.424	-0.619	0.523	0.271	0.915	0.296
4	-0.172	0.345	-0.580	0.817	0.696	1.174
5	0.724	-0.285	0.251	1.023	0.818	1.251
6	0.839	-0.807	1.142	0.454	1.631	0.278
7	-0.920	1.282	-1.850	0.244	2.432	0.100
8	1.570	-1.352	2.209	0.065	3.029	0.021
9	0.602	0.353	-0.193	0.641	0.724	0.885
10	0.555	0.170	0.0	0.212	0.580	0.366
11	0.391	0.449	-0.058	-0.244	0.598	-0.408
12	0.770	0.019	0.518	-0.265	0.928	-0.286
13	0.621	0.462	-0.143	-0.244	0.787	-0.310
14	1.024	0.361	0.237	-0.107	1.111	-0.096
15	0.715	0.468	0.023	-0.308	0.855	-0.360
16	0.810	0.278	0.506	-0.471	0.995	-0.473
17	0.306	0.847	-0.174	-0.425	0.917	-0.463
18	0.475	0.633	0.280	-0.590	0.839	-0.703
19	0.421	0.754	0.106	-0.615	0.870	-0.707
20	0.452	0.639	0.119	-0.524	0.792	-0.662
21	0.301	0.819	0.0	-0.767	0.873	-0.879
22	0.146	0.758	0.0	-0.719	0.772	-0.931
23	0.322	0.723	0.264	-0.609	0.834	-0.730
24	0.075	0.907	-0.037	-0.739	0.911	-0.811
25	0.299	0.425	0.545	-0.679	0.753	-0.902
26	0.085	0.655	0.245	-0.609	0.704	-0.865
27	-0.320	0.979	-0.269	-0.590	1.065	-0.554
28	0.152	0.438	0.382	-0.615	0.601	-1.023
29	-0.104	0.691	0.112	-0.431	0.708	-0.609
30	0.009	0.434	0.192	-0.336	0.475	-0.707

จากตาราง 2 พบว่าค่าอำนาจจำแนกแต่ละมิติ มีค่าอยู่ระหว่าง -1.850 ถึง 2.209 ค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -0.767 ถึง 0.817 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.247 ถึง 3.029 ค่าความยากแบบ พหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -1.023 ถึง 2.830 ซึ่งสอดคล้องกับเกณฑ์ที่กำหนดไว้ทุกข้อ โดยมีเกณฑ์พิจารณา ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) จะต้องไม่มีค่าติดลบ ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -4.00 ถึง 4.00 และไม่มีค่าอำนาจจำแนกในแต่ละมิติ แตกต่างกันมากเกินไป

ตาราง 3 สรุปผลค่าพารามิเตอร์ของแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ ค่าจุดตัดความยากแบบพหุมิติ ค่าอำนาจจำแนกแบบพหุมิติ และค่าความยากแบบพหุมิติ

ข้อสอบ	ค่าพารามิเตอร์	มิติของข้อสอบ			d	MDISC	MDIFF
		a1	a2	a3			
30 ข้อ	MEAN	0.365	0.361	0.152	-0.155	0.934	-0.126
	SD	0.458	0.558	0.620	0.542	0.548	0.886
	MIN	-0.920	-1.352	-1.850	-0.767	0.247	-1.023
	MAX	1.570	1.282	2.209	1.023	3.029	2.830

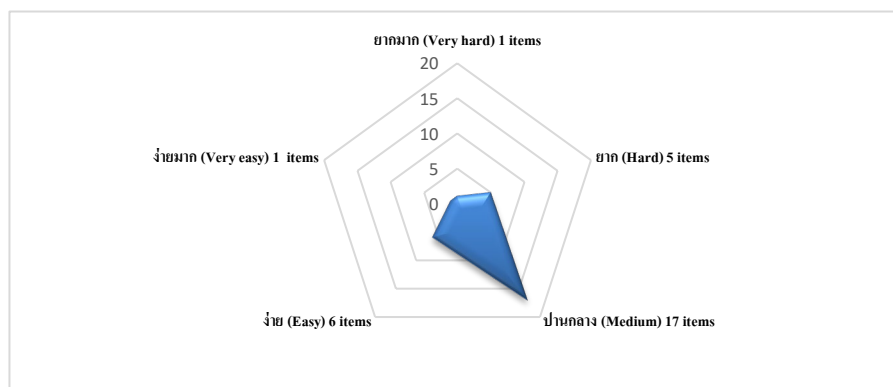
จากตาราง 3 พบว่าค่าอำนาจจำแนกในมิติที่ 1 มีค่าอยู่ระหว่าง -0.920 ถึง 1.570 ค่าเฉลี่ยมีค่าเท่ากับ 0.365 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.458 ค่าอำนาจจำแนกในมิติที่ 2 มีค่าอยู่ระหว่าง -1.352 ถึง 1.282 ค่าเฉลี่ยมีค่าเท่ากับ 0.361 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.558 ค่าอำนาจจำแนกในมิติที่ 3 มีค่าอยู่ระหว่าง -1.850 ถึง 2.209 ค่าเฉลี่ยมีค่าเท่ากับ 0.152 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.620 ความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -1.023 ถึง 2.830 ค่าเฉลี่ยมีค่าเท่ากับ -0.126 และส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 0.886 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.247 ถึง 3.029 ค่าเฉลี่ย มีค่าเท่ากับ 0.934 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.548 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -0.767 ถึง 1.023 ค่าเฉลี่ยมีค่าเท่ากับ -0.155 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.542 ค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (Multidimensional Discrimination : MDISC) มีค่าเฉลี่ยเท่ากับ 0.934 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.548 และมีค่าอยู่ระหว่าง 0.247 ถึง 3.029 เมื่อนำมาทำเป็นแผนภูมิเรดาร์ (Radar Charts) แสดงลักษณะแนวโน้ม โดยรวมของค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ ดังภาพ 3



ภาพ 3 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มิติแบบวัดความฉลาดรู้ด้านคณิตศาสตร์

จากภาพ 3 แนวโน้มโดยรวมของค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) เมื่อพิจารณาจากแผนภูมิพบว่าค่าอำนาจจำแนกแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบ ที่มีค่าอำนาจปานกลาง (moderate) จำนวน 21 ข้อ รองลงมา คือ มีค่าอำนาจจำแนกต่ำ (Low) จำนวน 5 ข้อ ค่าอำนาจจำแนกสูงมาก (very high) จำนวน 2 ข้อ และมีค่าอำนาจจำแนกสูง (High) และค่าอำนาจจำแนกต่ำมาก (very low) จำนวนอย่างละ 1 ข้อตามลำดับ

ค่าความยากแบบพหุมิติของข้อสอบ (Multidimensional Difficulties : MDIFF) มีค่าเฉลี่ยเท่ากับ -0.126 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.886 และมีค่าความยากอยู่ระหว่าง -1.023 ถึง 2.830 เมื่อนำมาทำเป็นแผนภูมิเรดาร์ (radar charts) แสดงลักษณะแนวโน้มโดยรวมของค่าความยากแบบพหุมิติของข้อสอบ ดังภาพ 4



ภาพ 4 ค่าความยากแบบพหุมิติ (MDIFF) มิติแบบวัดความฉลาดรู้ด้านคณิตศาสตร์

จากภาพ 4 แนวโน้มโดยรวมของค่าความยากแบบพหุมิติของข้อสอบ เมื่อพิจารณาจากแผนภูมิพบว่าค่าความยากแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบมีค่าความยากปานกลาง (medium) จำนวน 17 ข้อ รองลงมาข้อสอบง่าย (easy) จำนวน 6 ข้อ ข้อสอบยาก (hard) จำนวน 5 ข้อ และยากมาก (very hard) กับง่ายมากจำนวนอย่างละ 1 ข้อ ตามลำดับ

1.3 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ผู้วิจัยได้ใช้โปรแกรม IRTPRO คำนวณค่าฟังก์ชันสารสนเทศของแบบวัดความฉลาดรู้ด้านคณิตศาสตร์ เนื่องจากค่าฟังก์ชันสารสนเทศของแบบวัดความฉลาดรู้เป็นดัชนีในการพิจารณาความถูกต้อง แม่นยำในการประมาณความสามารถของผู้ตอบ จากการศึกษาค่าฟังก์ชันสารสนเทศของข้อสอบ (item information)

ตาราง 4 ค่าฟังก์ชันสารสนเทศแบบทดสอบ ตามแนวคิดทฤษฎีการตอบสนองข้อสอบ (item response theory)

ข้อ ที่	ค่าพารามิเตอร์							
	a	b	c	-6	-3	0	3	6
1	-1.435	-1.77176	0.766	1.104	5.220	2.928	6.677	1.285
2	0.886	1.548212	0.689	5.400	1.013	1.073	3.455	4.474
3	-1.236	-0.8156	0.437	1.408	4.833	7.542	1.445	9.398
4	-0.013	90.08118	0.134	2.471	2.521	2.572	2.622	2.670
5	-0.057	28.05048	0.136	0.000	0.000	0.000	0.004	0.000
6	-0.122	4.624583	0.115	0.002	0.002	0.002	0.002	0.002
7	5.782	1.568886	0.578	1.104	2.234	3.688	3.230	1.050
8	0.067	-1.45365	0.007	0.001	0.001	0.001	0.001	0.001
9	0.748	-1.59431	0.001	0.018	0.106	0.099	0.016	0.001
10	0.519	-0.75653	0.001	0.015	0.048	0.064	0.029	0.007
11	0.939	0.486727	0.000	0.001	0.029	0.209	0.068	0.004
12	0.827	0.568067	0.000	0.002	0.031	0.162	0.070	0.007
13	1.079	0.411	0.000	0.001	0.027	0.278	0.062	0.002
14	1.540	0.168	7.24E	9.014	1.709	5.842	2.875	2.984
15	1.378	0.472	0.000	0.000	0.015	0.430	0.053	0.000
16	1.740	0.646	4.92E	4.638	4.983	5.659	4.734	2.725
17	2.340	0.513	0.000	6.487	4.881	9.864	1.549	1.452
18	2.751	0.680	0.000	6.497	6.681	8.953	1.213	3.328
19	3.269	0.717	0.021	4.091	1.200	7.085	5.658	3.307
20	2.712	0.707	0.041	2.680	2.833	6.321	1.335	4.120
21	3.186	0.894	0.022	3.589	6.432	3.843	1.139	8.530
22	2.587	0.971	0.042	3.193	1.609	3.012	3.177	1.429
23	4.474	0.702	0.035	4.845	1.880	4.507	6.121	9.825
24	2.964	0.825	0.000	7.025	3.236	6.616	1.316	1.912
25	2.844	0.999	0.075	5.035	1.175	1.782	2.383	4.965
26	2.548	0.826	0.043	1.110	4.423	4.574	2.315	1.168
27	2.789	0.857	0.064	2.729	4.573	3.406	1.747	4.290
28	1.447	0.972	0.012	2.763	1.245	3.140	9.679	1.427
29	1.595	0.618	0.000	1.231	7.361	5.079	5.299	4.741
30	0.962	0.681	0.010	0.000	0.017	0.202	0.079	0.005

จากตาราง 4 ค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information) มีค่าความสามารถสูงสุดอยู่ระหว่าง -6 ถึง 6 สรุปภาพรวมคุณลักษณะของข้อสอบทั้ง 30 ข้อ ตามพารามิเตอร์สำคัญ ค่าอำนาจจำแนก (a) มีค่าเฉลี่ยประมาณ 1.5 ข้อสอบส่วนใหญ่มีอำนาจจำแนกปานกลาง มีข้อสอบที่โดดเด่น 3 ข้อ คือ ข้อ 7

($a = 5.782$) ข้อ 23 ($a = 4.474$) ข้อ 19 ($a = 3.269$) มีข้อสอบที่มีอำนาจจำแนกต่ำ เช่น ข้อ 8 ($a = 0.067$) ค่าความยากง่าย (b) มีค่าเฉลี่ยกระจายระหว่าง -2 ถึง 2 มีความหลากหลายค่อนข้างมาก มีข้อสอบที่มีความยากแตกต่างกันชัดเจน ข้อสอบที่มีความยากสูงสุด คือ ข้อ 4 ($b = 90.081$) ข้อสอบที่มีความยากต่ำสุด คือ ข้อ 1 ($b = -1.771$) ค่าการเดา (c) มีค่าเฉลี่ย ประมาณ 0.05-0.1 ส่วนใหญ่มีค่าการเดาต่ำมาก มีข้อสอบที่มีค่าการเดาสูงสุด คือ ข้อ 25 ($c = 0.075$) การกระจายค่าสารสนเทศข้อสอบส่วนใหญ่มีค่าสารสนเทศสูงบริเวณระดับความสามารถปานกลาง (0) ข้อสอบที่มีค่าสารสนเทศสูงสุด คือ ข้อ 23 สรุปภาพรวมแบบทดสอบ มีคุณภาพค่อนข้างดี มีความหลากหลายของระดับความยาก มีประสิทธิภาพในการวัดความสามารถของผู้สอบที่ระดับปานกลาง

สรุปผลการวิจัย

การพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติ ในรูปแบบแอปพลิเคชัน สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีการให้ข้อมูลย้อนกลับ มีวัตถุประสงค์การวิจัยเพื่อพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์แบบพหุมิติและวิเคราะห์คุณภาพแบบทดสอบโดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ (MIRT) โดยมีกลุ่มผู้ให้ข้อมูล คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จาก 16 โรงเรียน สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาตาก จำนวนกลุ่มทดลอง 483 คน โดยมีขั้นตอนการดำเนินการวิจัย ดังนี้ 1. สร้างแบบทดสอบ 40 ข้อ 2. ผู้เชี่ยวชาญ 5 คนพิจารณาความสอดคล้อง ของเนื้อหาและความเหมาะสม คัดเลือกข้อสอบที่ผ่านเกณฑ์ ไว้ 30 ข้อ มีค่า IOC .60 ถึง 1.00 วิเคราะห์ข้อมูล โดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) ใช้โปรแกรม NOHARM วิเคราะห์พารามิเตอร์ของข้อสอบด้วยโมเดลปกติสะสมแบบพหุมิติ ผลการวิเคราะห์ พบว่า ข้อสอบมีค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.247 ถึง 3.029 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -1.023 ถึง 2.830 และมีค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information) ส่วนใหญ่มีค่าสารสนเทศสูงบริเวณระดับความสามารถปานกลาง สรุปคุณภาพแบบทดสอบ พบว่า มีความเที่ยงตรงและน่าเชื่อถือ มีประสิทธิภาพในการวัดความสามารถของผู้สอบ สามารถจำแนกความแตกต่างของผู้สอบได้ดี และสามารถนำแบบทดสอบไปใช้วัดความฉลาดรู้ด้านคณิตศาสตร์ แต่ควรมีการพัฒนาและปรับปรุงอย่างต่อเนื่อง และนำไปขยายผลการใช้งานในบริบทอื่นๆ ได้ และงานวิจัยนี้เป็นก้าวสำคัญในการพัฒนาเครื่องมือวัดความฉลาดรู้คณิตศาสตร์ที่มีประสิทธิภาพ เหมาะสมกับบริบทการศึกษาในปัจจุบัน

อภิปรายผล

การพัฒนาแบบวัดความฉลาดรู้ด้านคณิตศาสตร์ เริ่มจากการสร้างแบบวัดจำนวน 40 ข้อ และผ่านการพิจารณาจากผู้เชี่ยวชาญ 5 คน มีค่าดัชนีความสอดคล้องตั้งแต่ 0.60 ถึง 1 ที่คัดเลือกไว้ 30 ข้อ ซึ่งแสดงให้เห็นว่าข้อสอบที่สร้างขึ้นมีความตรงตามเนื้อหาทั้งนี้เนื่องจากการออกข้อสอบในครั้งนี้ได้มีการสร้าง Q-matrix เพื่อเป็นการระบุมิติที่ต้องการวัดของข้อสอบแต่ละข้อตามองค์ประกอบของความฉลาดรู้ด้านคณิตศาสตร์ โดยการระบุมิติที่ต้องการวัดของข้อสอบตาม Q-matrix ในแถวของ Q-matrix ที่แทนด้วยตัวเลข

“1” จะสะท้อนถึงโครงสร้างน้ำหนักองค์ประกอบเชิงซ้อนของข้อคำถาม แต่ละข้อซึ่งข้อคำถามแต่ละข้อจะประกอบด้วยมิติหลายมิติ จึงทำให้ข้อคำถามแต่ละข้อมีโครงสร้างน้ำหนักองค์ประกอบในลักษณะเชิงซ้อนหรือมีความหลากหลาย (Rupp & Templin, 2008) จึงทำให้แบบทดสอบที่สร้างขึ้นตามการกำหนดด้วย Q-matrix มีความตรงตามเนื้อหา 30 ข้อ ที่มีคุณภาพ ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ (MIRT) ในการวิเคราะห์ ทำให้ได้คุณลักษณะของแบบทดสอบ ที่มีค่าอำนาจจำแนกส่วนใหญ่อยู่ในระดับปานกลาง มีข้อสอบที่โดดเด่น 3 ข้อ ได้แก่ ข้อ 7, 23 และ 19 ข้อสอบมีความหลากหลายของระดับความยาก และมีค่าการเดาต่ำ (ประมาณ 0.05-0.1)

จากผลการวิเคราะห์ข้อมูลจะเห็นได้ว่า ค่าอำนาจจำแนกจากการทดสอบ มีค่าสูงในบางข้อ อาจเนื่องมาจากข้อสอบบางข้อมีความสามารถในการวัดบางมิติสูงกว่ามิติอื่น ๆ อีกประเด็น หนึ่งตามทฤษฎีการตอบสนองข้อสอบต้องเป็นการสอบที่ไม่แข่งขันด้านเวลาความเร็วในการตอบ จะต้องไม่มีอิทธิพลต่อผลการตอบการจับสอบ จึงต้องไม่อยู่ในสถานการณ์ที่สอบแข่งขันกันด้วยเวลา แต่ในสภาพการทดสอบจริงไม่สามารถจัดการสอบที่ไม่แข่งขันด้านเวลาได้ในทุกครั้ง ซึ่งอาจมีผลต่อ ค่าพารามิเตอร์ของข้อสอบ ในการสร้างข้อสอบในการวิจัยครั้งนี้ ผู้วิจัยสร้างข้อสอบที่สามารถวัด องค์ประกอบของความฉลาดรู้ด้านคณิตศาสตร์ของนักเรียนได้มากกว่า 1 องค์ประกอบ ดังนั้นข้อสอบ 1 ข้อ จึงมีค่าอำนาจจำแนกตั้งแต่ 2 ค่าขึ้นไป (a_1 , a_2 และ a_3) และในการประมาณค่าพารามิเตอร์ของ ข้อสอบพบว่าค่าอำนาจจำแนกของข้อสอบมีค่าเป็นลบในบางมิติ ทั้งนี้เนื่องมาจากผลการวิเคราะห์ เมทริกซ์ความแปรปรวน ความแปรปรวนร่วมขององค์ประกอบของความฉลาดรู้ด้านคณิตศาสตร์ ที่ผู้วิจัยนำมาออกข้อสอบมีจำนวนทั้งสิ้น 3 มิติ พบว่า มีความสัมพันธ์กันค่อนข้างซับซ้อนคือมีความสัมพันธ์กันทั้งทางบวกและลบ และบางกระบวนการก็มีความสัมพันธ์กันค่อนข้างน้อยหรือแทบไม่มีความสัมพันธ์กันเลย ซึ่งสอดคล้องกับแนวคิดของโครงการ PISA ประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) ที่ว่าการประเมินความฉลาดรู้ด้านคณิตศาสตร์เน้นความชัดเจนที่ต้องการให้นักเรียนเผชิญหน้ากับปัญหาทางคณิตศาสตร์ที่มีอยู่ในชีวิตในกระบวนการนี้ต้องการสมรรถนะหลายอย่าง ซึ่งนักเรียนต้องใช้ทักษะและสมรรถนะต่าง ๆ ที่หลากหลายมารวมกันหรือทับซ้อนหรือคาบเกี่ยวกัน (Klanin et al., 2008) สำหรับค่าฟังก์ชันสารสนเทศของแบบวัดความฉลาดรู้ด้านคณิตศาสตร์ ซึ่งสะท้อนความถูกต้องแม่นยำในการประมาณความสามารถของผู้ตอบ เมื่อนำมารวมกันเป็น สารสนเทศของแบบทดสอบจะเป็นดัชนีชี้ว่าจะได้แบบทดสอบตามเป้าหมายที่ต้องการ (Kanjana-wasee, 2009) จากค่าฟังก์ชันสารสนเทศของข้อสอบ (item information) มีค่าความสามารถสูงสุดอยู่ระหว่าง -6 ถึง 6 ค่าฟังก์ชันสารสนเทศของข้อสอบทุกข้อมีความสอดคล้องกับค่าความยากง่ายและค่าอำนาจจำแนก กล่าวคือค่าฟังก์ชันสารสนเทศของข้อสอบจะสูงขึ้นเมื่อค่าความสามารถสูงสุดใกล้กับค่าความยากง่ายและค่าสารสนเทศของข้อสอบโดยทั่วไปมีค่าสูงขึ้น เมื่อค่าอำนาจจำแนกมีค่ามากขึ้น จากลักษณะดังกล่าวแสดงให้เห็นว่าข้อสอบมีคุณภาพตามทฤษฎีการตอบสนองข้อสอบ (Kanjana-wasee, 2007) เช่นกันกับ งานวิจัยของ (Kaewmanee, 2009) ที่มีผลการเปรียบเทียบฟังก์ชันสารสนเทศของข้อสอบและฟังก์ชัน สารสนเทศของแบบสอบ มีความสอดคล้องกับระดับความสามารถสูงสุดค่าความยากง่ายค่าอำนาจจำแนก

ข้อเสนอแนะ

1. ข้อเสนอแนะสำหรับการนำไปใช้

1.1 สามารถนำแอปพลิเคชันนี้ไปใช้วัดความฉลาดรู้ด้านคณิตศาสตร์ของนักเรียนมัธยมศึกษาปีที่ 3 และเป็นเครื่องมือที่ให้ข้อมูลย้อนกลับแก่นักเรียนได้ และนักเรียนสามารถใช้ได้กับ smartphone, computer, ipad และ Notebook

1.2 ครูสามารถนำแอปพลิเคชันนี้ไปใช้วัดและประเมินความฉลาดรู้ด้านคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ทั้งในชั้นเรียนและช่วงกิจกรรมการเรียนรู้เสริม

2. ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

2.1 ควรพัฒนาแอปพลิเคชันให้ครอบคลุมทักษะทางคณิตศาสตร์มากยิ่งขึ้น อาจขยายขอบเขตการวิจัยไปยังระดับชั้นอื่นหรือพื้นที่อื่น และควรมีการวิจัยเพื่อติดตามผลการใช้แอปพลิเคชันอย่างต่อเนื่อง

2.2 ควรมีการวิจัยในระยะยาวเพื่อติดตามผลการใช้แอปพลิเคชัน เช่น ผลต่อทักษะทางคณิตศาสตร์ของนักเรียนหลังการใช้งานในระยะเวลา 6 เดือน หรือ 1 ปี และการวิเคราะห์พฤติกรรมการเรียนรู้และความเปลี่ยนแปลงด้านแรงจูงใจของนักเรียน

References

- OECD. (2018). *PISA 2022 Mathematics Framework (draft)*. <https://pisa2022-maths.oecd.org>
- Rupp, A. A., & Templin, J. (2008). The effects of Q-matrix misspecification on parameter estimates and classification accuracy in the DINA model. *Educational and Psychological Measurement*, 68(1), 78–96. <https://doi.org/10.1177/0013164407301545>

Translate Thai References

- Kaewmanee, K. (2009). *The development of electronic books to enhance reading comprehension skills Thai Language Learning Group: Local Set for Phitsanulok Residents for 6th Grade Elementary Students* [Master's thesis]. Naresuan University. (in Thai)
- Kanjanawasee, S. (2007). *Modern test theory* (3th ed.). Chulalongkorn University Press. (in Thai)
- Kanjanawasee, S. (2009). *Classical test theory* (6th ed.). Chulalongkorn University Press. (in Thai)
- Khlangbunkrong, K. (2012). A web application to assist mothers in monitoring children's health and recording their development [Bachelor of Science's thesis], King Mongkut's University of Technology Thonburi. (in Thai)
- Klanin, S., Dechsri, P., & Pramojanee, A. (2008). *Knowledge and Scientific Competence for Tomorrow's World*. Seven Printing Group. (in Thai)

- Manochanpen, N. (2022). *TDRI points out that Thai students' PISA scores have declined in all competencies, reflecting an educational crisis that must be urgently addressed at its root causes*. <https://thestandard.co/tdri-pisa-test-results/> (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* [Doctoral dissertation]. Chulalongkorn University Press. (in Thai)
- Tayraukhum, S. (2010). *Statistical Methodologies for Empirical Research*. Department of Educational Research and Development, Faculty of Education, Mahasarakham University. (in Thai)
- Tayraukhum, S. (n.d). *Teaching materials on the development of multidimensional tests, Department of Research and Educational Development*. Mahasarakham University. (in Thai)



ภาคผนวก

- รายชื่อผู้ทรงคุณวุฒิพิจารณาบทความ วารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม ปีที่ 31 ฉบับที่ 1: มกราคม - มิถุนายน 2568
- หลักเกณฑ์และคำแนะนำสำหรับผู้เขียน



รายชื่อผู้ทรงคุณวุฒิพิจารณาบทความวารสารการวัดผลการศึกษา มหาวิทยาลัยมหาสารคาม

ประจำปี 31 ฉบับที่ 1: มกราคม - มิถุนายน 2568

- | | |
|------------------------------------|---|
| 1. รศ.ดร.วราพร เอรารวรรณ์ | มหาวิทยาลัยมหาสารคาม |
| 2. รศ.ดร.กัญญารัตน์ โคจร | มหาวิทยาลัยมหาสารคาม |
| 3. รศ.ดร.สุชาติ คุ่มมะณี | มหาวิทยาลัยมหาสารคาม |
| 4. ผศ.ดร.ฉัตรเกล้า เจริญผล | มหาวิทยาลัยมหาสารคาม |
| 5. ผศ. ดร.ณภัทรวรรณ ธนาพงษ์อ่อนนัท | มหาวิทยาลัยมหาสารคาม |
| 6. ผศ.ดร.สมทรง สติธิ | มหาวิทยาลัยมหาสารคาม (ข้าราชการเกษียณ) |
| 7. รศ.ดร.พูลพงศ์ สุขสว่าง | มหาวิทยาลัยบูรพา |
| 8. รศ.ดร.ภัทราวดี มากมี | มหาวิทยาลัยบูรพา |
| 9. รศ.ดร.พัชรี จันทรเพ็ง | มหาวิทยาลัยขอนแก่น |
| 10. ผศ.ดร.ชัยยุทธ กลีบบัว | มหาวิทยาลัยศรีนครินทรวิโรฒ |
| 11. รศ.ดร.ชัยวิชิต เขียวชนะ | มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ |
| 12. ผศ.ดร.มิ่งขวัญ ภาคสัณไชย | มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี |
| 13. ผศ.ดร.เรืองเดช ศิริกิจ | มหาวิทยาลัยศรีนครินทรวิโรฒ |
| 14. ดร.พนัส จันทรเปล่ง | มหาวิทยาลัยศรีนครินทรวิโรฒ |
| 15. ผศ.ดร.สรัญญา จันทรชูสกุล | มหาวิทยาลัยศิลปากร |
| 16. รศ.ดร.เมษา นวลศรี | มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ ในพระบรมราชูปถัมภ์ |
| 17. รศ.ดร.สุนทรพจน์ ดำรงค์พานิช | มหาวิทยาลัยเชียงใหม่ |
| 18. รศ.ดร.กมลทิพย์ ศรีหาเศษ | มหาวิทยาลัยรามคำแหง |
| 19. ผศ.ดร.จุฑาภรณ์ มาสันเทียะ | มหาวิทยาลัยรามคำแหง |
| 20. ผศ.ดร.สัระโชติ ศรีสุทธิยากร | จุฬาลงกรณ์มหาวิทยาลัย |
| 21. ผศ.ดร.ณภัทร ชัยมงคล | จุฬาลงกรณ์มหาวิทยาลัย |
| 22. ผศ.ดร.สุกัญญา รุจิเมธภาส | มหาวิทยาลัยราชภัฏอุดรดิตถ์ |
| 23. ผศ.ดร.ศจี จิระโร | มหาวิทยาลัยสุโขทัยธรรมมาธิราช |
| 24. ผศ.ดร.พินดา วราสุนันท์ | มหาวิทยาลัยเกษตรศาสตร์ |
| 25. ผศ.ดร.สิรินธร สิ้นจินดาวงศ์ | มหาวิทยาลัยศรีปทุม |
| 26. ผศ.ดร.ศิวะพร ภูพันธ์ | มหาวิทยาลัยมหิดล |
| 27. อ.ดร.ณมรัตน์ ศิริภาพ | โรงเรียนโพธิสารศึกษา |

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 ตั้งค่าน้ำกระดาษสำหรับการพิมพ์ห่างจากขอบกระดาษทุกด้าน 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, ฤทธิสฤติพัฒนานนท์, 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 สี่พาย 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/[วิทยานิพนธ์ปริญญาโทหรือปริญญาเอก] /////////วิทยานิพนธ์ปริญญาโทหรือปริญญาเอก/ชื่อเว็บไซต์./URL
อังกฤษ	สกุล/อักษรย่อชื่อ/(ปีที่เผยแพร่)/ชื่อวิทยานิพนธ์/[Doctoral dissertation or Master's thesis] /////////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

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 เปิดหลักสูตรสายอาชีพตามความต้องการตลาดแรงงาน*. <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 และคั่นแต่ละแหล่งด้วยเครื่องหมายอัฒภาค (;)

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.

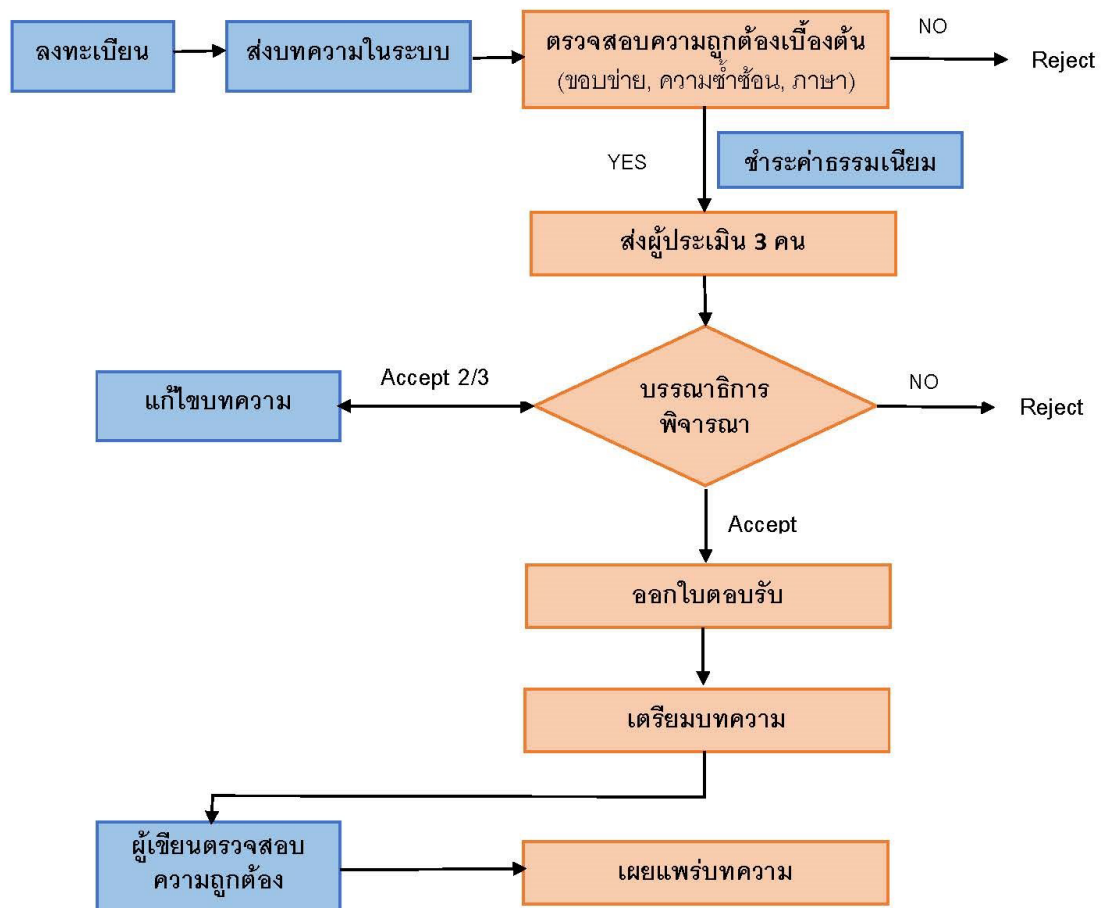
Potjanajruwit, P. (2016). The Corporate Branding Strategy of Rajabhat University in Thailand. *Journal of Graduate School of Commerce Burapha Review*, 11(1), 17 – 49. (in Thai)

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. บทความภาษาไทย 3,000 บาท/1 บทความ
2. บทความภาษาอังกฤษ 5,000 บาท/1 บทความ

*ชำระค่าธรรมเนียมเมื่อผ่านการตรวจสอบคุณภาพเบื้องต้น ก่อนนำเสนอต่อผู้ทรงคุณวุฒิ

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