

การพัฒนาแบบวัดวินิจฉัยสำหรับสมรรถนะแห่งศตวรรษที่ 21

The Development of Diagnostic Assessment for 21st Century Competencies

ณัฐพงษ์ อิมหิรัญ¹ และสำราญ มีแจ้ง²

Natthaphong Imhiran and Samran Mejang

Received: November 26, 2025

Revised: February 06, 2026

Accepted: May 05, 2026

บทคัดย่อ

บทความวิชาการนี้มีวัตถุประสงค์เพื่อนำเสนอกรอบแนวคิดในการพัฒนาแบบวัดวินิจฉัยสมรรถนะ โดยเป็นการศึกษาเชิงสังเคราะห์วรรณกรรมสากลและงานวิจัยที่เกี่ยวข้อง พร้อมบูรณาการเข้ากับบริบทการศึกษาไทย ผลการศึกษา พบว่า การศึกษาในศตวรรษที่ 21 จำเป็นต้องเปลี่ยนแปลงรูปแบบการประเมินที่ต้องก้าวข้ามการวัดความรู้ไปสู่การวัดสมรรถนะที่ซับซ้อนและเป็นนามธรรม โดยชี้ให้เห็นถึงความจำเป็นในการแตกโครงสร้างสมรรถนะให้เป็นมิติน้อยที่สามารถวินิจฉัยได้จากการวิเคราะห์ ถึงแม้จะพบว่าการประเมินเชิงปฏิบัติการใช้เทคโนโลยี เช่น สถานการณ์จำลอง การวิเคราะห์ข้อมูลบันทึกพฤติกรรม และแบบจำลองการวินิจฉัยเชิงปริมาณ (DCMs) จะมีศักยภาพสูงในการวิเคราะห์เชิงลึก แต่ก็มีข้อจำกัดสำคัญด้านความเที่ยงตรงเชิงการตีความ ซึ่งทำให้ผลลัพธ์ที่ได้มีความซับซ้อนและครุ่นไปใช้ได้ยาก โดยเฉพาะในบริบทของประเทศไทยที่ยังมีความพร้อมทางเทคโนโลยีและความเข้าใจในทฤษฎีการวัดขั้นสูงที่แตกต่างกันระหว่างสถานศึกษา เพื่อแก้ปัญหานี้ จึงได้เสนอวงจรวินิจฉัยเพื่อการพัฒนาที่เน้นการให้ข้อมูลย้อนกลับที่ชัดเจนและนำไปปฏิบัติได้ พร้อมทั้งเสนอให้ใช้การบูรณาการข้อมูลแบบสามเส้า โดยผสมผสานข้อมูลเชิงปริมาณจากแบบวัดเข้ากับการสังเกตเชิงคุณภาพของครู รวมทั้งเสนอแนะแนวทางการนำไปปฏิบัติที่เหมาะสมกับบริบทไทย เพื่อให้ครูสามารถนำไปออกแบบกิจกรรมการเรียนรู้ที่มุ่งเป้า อันเป็นการเชื่อมโยงช่องว่างระหว่างทฤษฎีการวัดขั้นสูงและการปฏิบัติจริงในห้องเรียนไทย

¹⁻² มหาวิทยาลัยนเรศวร; Naresuan University

Corresponding author, e-mail: natthaphongi67@nu.ac.th

คำสำคัญ: แบบวัดวินิจฉัย, การประเมินสมรรถนะ, แบบจำลองการวินิจฉัยการเรียนรู้

Abstract

This academic article presents a conceptual framework for developing diagnostic competency assessments. It synthesizes international literature and relevant research and integrates the findings into the Thai educational context. The synthesis indicates that assessment in twenty-first-century education must move beyond measuring knowledge to assessing complex and abstract competencies. This requires breaking competencies down into subdimensions that allow for diagnostic analysis. Technology-supported performance assessments, such as simulations and behavioral log analysis, together with diagnostic classification models (DCMs), offer considerable potential for in-depth analysis. However, challenges concerning the validity of assessment interpretations can make the results complex and difficult for teachers to apply. These challenges are particularly relevant in Thailand, where technological readiness and understanding of advanced measurement theories vary across educational institutions. To address these issues, the article proposes a diagnostic cycle for improvement that emphasizes clear, actionable feedback. It also recommends data triangulation by combining quantitative assessment data with teachers' qualitative observations. Practical guidelines tailored to the Thai context are proposed to help teachers design targeted learning activities, thereby bridging the gap between advanced measurement theory and classroom practice in Thailand.

Keywords: Diagnostic Assessment Instruments, Competency Assessment, Cognitive Diagnostic Models

บทนำ

ในยุคที่โลกขับเคลื่อนด้วยข้อมูล นวัตกรรม และเทคโนโลยีปัญญาประดิษฐ์ กระบวนทัศน์ (Paradigm) ของการจัดการศึกษาได้เปลี่ยนแปลงไปอย่างสิ้นเชิง (Griffin & Care, 2015) ทักษะและความรู้ที่เคยเพียงพอในศตวรรษที่ 20 กำลังถูกท้าทายด้วยความต้องการทรัพยากรมนุษย์

ที่มีสมรรถนะ (Competencies) ที่ซับซ้อนมากขึ้น ส่งผลให้เป้าหมายของการศึกษาจำเป็นต้องเปลี่ยนผ่านจากการเน้นการวัดความรู้ (Knowledge Measurement) ที่ผู้เรียนจดจำได้ ไปสู่การวัดสมรรถนะ (Competency Measurement) หรือความสามารถในการประยุกต์ใช้ความรู้ ทักษะ และคุณลักษณะต่าง ๆ เพื่อแก้ปัญหาในสถานการณ์จริงที่ไม่คุ้นเคย (Glaser, Chudowsky & Pellegrin, 2001)

กระทรวงศึกษาธิการได้ประกาศใช้หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 (ปรับปรุง พ.ศ. 2560) ซึ่งเน้นการพัฒนาผู้เรียนให้มีสมรรถนะสำคัญ 5 ประการ ได้แก่ ความสามารถในการสื่อสาร ความสามารถในการคิด ความสามารถในการแก้ปัญหา ความสามารถในการใช้ทักษะชีวิต และความสามารถในการใช้เทคโนโลยี (กระทรวงศึกษาธิการ, 2560) สอดคล้องกับกรอบสมรรถนะแห่งศตวรรษที่ 21 อย่างชัดเจน นอกจากนี้ แผนการศึกษาแห่งชาติ พ.ศ. 2560-2579 ยังได้กำหนดให้มีการพัฒนาผู้เรียนให้มีทักษะและสมรรถนะที่จำเป็นต่อการดำรงชีวิตและการทำงานในศตวรรษที่ 21 (สำนักงานเลขาธิการสภาการศึกษา, 2560) อย่างไรก็ตาม ความท้าทายสำคัญที่ยังคงมีอยู่คือการขาดเครื่องมือวัดและประเมินที่สามารถวินิจฉัยสมรรถนะเหล่านี้ได้อย่างเป็นระบบและลึกซึ้ง การเปลี่ยนผ่านดังกล่าวได้สร้างความท้าทายที่สำคัญอย่างยิ่งในแวดวงการวัดและประเมินผลการศึกษา เนื่องจากสมรรถนะหลักแห่งศตวรรษที่ 21 เช่น การคิดวิเคราะห์ การสื่อสาร นวัตกรรม หรือการทำงานร่วมกันมีลักษณะเป็นนามธรรม (Abstract Constructs) ที่ซับซ้อนและแฝงเร้น (Latent) การวัดสมรรถนะเหล่านี้จึงไม่สามารถกระทำได้อย่างตรงไปตรงมาเหมือนการวัดความรู้เชิงเนื้อหา (Partnership for 21st Century Skills [P21], 2009) เครื่องมือวัดแบบดั้งเดิม เช่น แบบทดสอบเลือกตอบที่เน้นความจำ มักประสบข้อจำกัดเชิงประจักษ์ในการสะท้อนความสามารถที่แท้จริงของผู้เรียนในมิติที่ซับซ้อนเหล่านี้ ถึงแม้หลักสูตรจะเน้นสมรรถนะแห่งศตวรรษที่ 21 แต่การประเมินส่วนใหญ่ยังคงใช้แบบทดสอบแบบเดิมที่วัดความรู้เชิงเนื้อหาเป็นหลัก โดยเฉพาะในโรงเรียนขนาดเล็กและโรงเรียนในพื้นที่ห่างไกล ซึ่งขาดทรัพยากรและความรู้ในการพัฒนาเครื่องมือประเมินสมรรถนะที่ทันสมัย ทำให้ข้อมูลการประเมินที่ได้ไม่สามารถนำมาใช้ในการพัฒนาผู้เรียนได้อย่างมีประสิทธิภาพ

ดังนั้นแบบวัดวินิจฉัย (Diagnostic Assessment) จึงมีความสำคัญขึ้นอย่างยิ่งยวด ในขณะที่การประเมินผลสัมฤทธิ์ทั่วไป (Summative Assessment) มักให้ข้อมูลเพียงคะแนนรวมที่ระบุว่าผู้เรียนทำได้ หรือทำไม่ได้ซึ่งไม่เพียงพอต่อการพัฒนาการเรียนรู้ แต่แบบวัดวินิจฉัยถูกออกแบบมาโดยมีวัตถุประสงค์เพื่อการวินิจฉัยอย่างลึกซึ้ง กล่าวคือ เพื่อระบุช่องว่างเชิงสมรรถนะ (Competency Gaps) หรือจุดแข็งและจุดอ่อนในองค์ประกอบย่อย (Sub-skills or Attributes)

ของสมรรถนะนั้น ๆ ข้อมูลที่ได้จากแบบวัดวินิจฉัยจะช่วยให้ผู้สอนสามารถให้ข้อมูลย้อนกลับเชิงลึก (In-depth Feedback) และออกแบบการสอนซ่อมเสริมหรือการพัฒนาที่ตรงจุด (Targeted Intervention) ได้อย่างมีประสิทธิภาพ มากกว่าการวัดผลเพื่อตัดสินเกรดเพียงอย่างเดียว (Rupp, Templin, & Henson, 2010)

บทความวิชาการนี้จึงมีวัตถุประสงค์เพื่อสังเคราะห์แนวคิด ทฤษฎี และความท้าทายร่วมสมัยในการวัดสมรรถนะแห่งศตวรรษที่ 21 พร้อมทั้งวิเคราะห์ข้อจำกัดของเครื่องมือวัดในปัจจุบัน และนำเสนอกรอบแนวคิด (Conceptual Framework) ในการพัฒนาแบบวัดวินิจฉัยสมรรถนะ โดยมุ่งเน้นการบูรณาการแนวคิดการประเมินเชิงสถานการณ์ (Situation-based Assessment) เข้ากับแบบจำลองการวินิจฉัยการรู้ (Cognitive Diagnostic Models: CDMs) และปรับให้เหมาะสมกับบริบทการศึกษาไทย เพื่อเป็นแนวทางสำหรับนักวิชาการ นักวัดผล ผู้บริหารสถานศึกษา และผู้กำหนดนโยบายในการสร้างเครื่องมือประเมินยุคใหม่ที่สามารถวินิจฉัยสมรรถนะของผู้เรียนได้อย่างเที่ยงตรงและมีความหมาย

การสังเคราะห์กรอบแนวคิดสมรรถนะและการวินิจฉัย

การพัฒนาแบบวัดวินิจฉัยสำหรับสมรรถนะแห่งศตวรรษที่ 21 จำเป็นต้องอาศัยการสังเคราะห์องค์ความรู้จากกรอบแนวคิดสมรรถนะ (Competency Frameworks) และทฤษฎี การวัดและการวินิจฉัย (Diagnostic Measurement Theories)

การจำแนกประเภทสมรรถนะแห่งศตวรรษที่ 21

สมรรถนะแห่งศตวรรษที่ 21 (21st Century Skills) เป็นคำที่ใช้อธิบายถึงชุดทักษะความรู้ และคุณลักษณะที่จำเป็นต่อการดำรงชีวิตและการทำงานในโลกยุคใหม่ การศึกษาวรรณกรรมและงานวิจัยที่เกี่ยวข้องได้มีการจัดกลุ่มสมรรถนะเหล่านี้ไว้หลากหลายกรอบแนวคิด ที่ได้รับการยอมรับอย่างกว้างขวางที่สุดคือกรอบแนวคิดของ Partnership for 21st Century Skills (P21) (2009) ซึ่งจำแนกสมรรถนะหลักออกเป็น 3 ด้าน โดยมีด้านที่โดดเด่น คือ ทักษะการเรียนรู้และนวัตกรรม (Learning and Innovation Skills) หรือที่รู้จักกันในนาม “4Cs” ได้แก่ การคิดเชิงวิพากษ์ (Critical Thinking) การสื่อสาร (Communication) การทำงานร่วมกัน (Collaboration) และการสร้างสรรค์ (Creativity)

นอกจากนี้ยังมีกรอบแนวคิดอื่น เช่น สมรรถนะสากล (Global Competency) ที่เสนอโดย OECD (2018) ซึ่งมุ่งเน้นความสามารถในการวิเคราะห์ประเด็นปัญหาระดับโลก การเข้าใจและเคารพมุมมองที่แตกต่าง และการมีส่วนร่วมอย่างเหมาะสมในบริบทพหุวัฒนธรรม จากการศึกษา

เกี่ยวกับการพัฒนากรอบสมรรถนะสำหรับผู้เรียนไทยในศตวรรษที่ 21 โดยบูรณาการกับบริบทวัฒนธรรมและค่านิยมไทยของนักการศึกษาหลายท่าน พบว่า นอกจากทักษะ 4Cs แล้ว ยังจำเป็นต้องเน้นคุณลักษณะด้านจริยธรรม ความมีวินัย และความรับผิดชอบต่อสังคม ซึ่งสอดคล้องกับแนวคิดคนดี คนเก่ง ตามปรัชญาการศึกษาไทย การศึกษานี้ชี้ให้เห็นว่า การนำกรอบสมรรถนะสากลมาใช้ในบริบทไทยจำเป็นต้องมีการปรับเพื่อให้สอดคล้องกับคุณค่าและวัฒนธรรมไทย และสำนักงานเลขาธิการสภาการศึกษา (2560) ได้เสนอกรอบทักษะและสมรรถนะของคนไทยในศตวรรษที่ 21 ที่ครอบคลุม 3 กลุ่มหลัก คือ 1) ทักษะการเรียนรู้และนวัตกรรม ประกอบด้วย ความคิดสร้างสรรค์และนวัตกรรม การคิดอย่างมีวิจารณญาณและการแก้ปัญหา การสื่อสารและการร่วมมือ 2) ทักษะชีวิตและอาชีพ ประกอบด้วย ความยืดหยุ่นและการปรับตัว ความคิดริเริ่มและการกำกับตนเอง การมีปฏิสัมพันธ์ทางสังคมและข้ามวัฒนธรรม ความเป็นผู้นำและความรับผิดชอบต่อสังคม และ 3) ทักษะสารสนเทศ สื่อ และเทคโนโลยี กรอบแนวคิดนี้ได้รับการนำไปใช้เป็นแนวทางในการจัดการศึกษาในหลายสถานศึกษาทั่วประเทศ แม้ว่ากรอบแนวคิดเหล่านี้จะมีความหลากหลายทั้งในระดับสากลและระดับชาติ แต่มีจุดร่วมที่สำคัญคือการเน้นย้ำสมรรถนะที่เป็นกระบวนการคิดขั้นสูง (Higher-order Thinking) และทักษะด้านสังคม ซึ่งท้าทายต่อการวัดและประเมินด้วยเครื่องมือแบบดั้งเดิม

แนวคิดพื้นฐานของแบบวัดวินิจฉัย

แนวคิดหลักของแบบวัดวินิจฉัย คือ การให้ข้อมูลย้อนกลับที่ละเอียดและมีความหมาย เพื่อสนับสนุนการปรับปรุงการเรียนรู้ (Black & Wiliam, 1998) ซึ่งแตกต่างจากการประเมินผลสัมฤทธิ์ที่มักให้เพียงคะแนนรวม การวินิจฉัยในทางการศึกษาสามารถจำแนกออกเป็นรูปแบบดังนี้

1. การวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน (Misconception Diagnosis) มุ่งเน้นการค้นหาคำเข้าใจผิดหรือรูปแบบการให้เหตุผลที่ผิดพลาดของผู้เรียน มักใช้ในสาขาวิชาที่มีโครงสร้างความรู้ชัดเจน เช่น คณิตศาสตร์ หรือวิทยาศาสตร์
2. การวินิจฉัยจุดบกพร่องเชิงทักษะ (Skill Deficit Diagnosis) มุ่งเน้นการระบุว่าผู้เรียนมีหรือขาดทักษะย่อย (Sub-skills) หรือองค์ประกอบย่อย (Attributes) ใดที่เป็นองค์ประกอบของความสามารถที่ซับซ้อน
3. การวินิจฉัยตามลำดับขั้นพัฒนาการ (Developmental Diagnosis) มุ่งเน้นการระบุว่าผู้เรียนอยู่ในขั้นตอนใดของลำดับขั้นการเรียนรู้ที่กำหนดไว้

สำหรับการวัดสมรรถนะแห่งศตวรรษที่ 21 ซึ่งเป็นความสามารถที่ซับซ้อน การวินิจฉัยในรูปแบบที่ 2 (Skill Deficit Diagnosis) หรือที่เรียกว่าการประเมินเชิงคุณลักษณะ (Attribute-based Assessment) ถือว่ามีความเหมาะสมและท้าทายที่สุด ถึงแม้งานวิจัยเกี่ยวกับการประเมินวินิจฉัยยังมีจำนวนจำกัด แม้จะมีการนำแบบทดสอบวินิจฉัยมาใช้ในบางสาขาวิชา เช่น คณิตศาสตร์และวิทยาศาสตร์ แต่ส่วนใหญ่มุ่งเน้นการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนมากกว่าการวินิจฉัยสมรรถนะที่ซับซ้อน การขาดแคลนเครื่องมือวินิจฉัยสมรรถนะที่มีคุณภาพทำให้ครูไทยส่วนใหญ่ยังคงพึ่งพาการสังเกตแบบไม่เป็นทางการและประสบการณ์ส่วนตัวในการประเมินสมรรถนะของผู้เรียน ซึ่งอาจขาดความเป็นระบบและความเที่ยงตรง

การเชื่อมโยงสมรรถนะสู่มิติที่วินิจฉัยได้

ความท้าทายสูงสุดในการวินิจฉัยสมรรถนะ คือ การแตกโครงสร้าง (Deconstruction) สมรรถนะที่เป็นนามธรรม เช่น การทำงานร่วมกันให้กลายเป็นมิติที่สามารถวินิจฉัยได้ (Diagnosable Dimensions) ซึ่งหมายถึง องค์ประกอบย่อยที่สามารถสังเกต วัด และระบุจุดแข็งจุดอ่อนได้อย่างชัดเจน (Diagnosable Dimensions) ที่เป็นรูปธรรมและสังเกตได้ หากปราศจากการแตกโครงสร้างนี้ การประเมินจะทำได้เพียงการให้คะแนนแบบภาพรวม (Holistic Scoring) ซึ่งไม่เพียงพอต่อการวินิจฉัย งานวิจัยเกี่ยวกับการประเมินสมรรถนะ เช่น Griffin & Care (2015); Hesse, Care, Buder, Sassenberg & Griffin (2015) ที่ชี้ให้เห็นว่า การจะวินิจฉัยสมรรถนะการทำงานร่วมกันได้นั้น ผู้ประเมินไม่สามารถวัดผลลัพธ์สุดท้ายของงานกลุ่มเพียงอย่างเดียว แต่ต้องสามารถระบุความบกพร่องในองค์ประกอบเชิงกระบวนการ (Process Components) ได้ เช่น สมรรถนะการทำงานร่วมกัน (Collaboration) สามารถแตกโครงสร้างออกเป็นองค์ประกอบย่อยที่สามารถวินิจฉัยได้ ดังนี้

1. การสร้างความเข้าใจร่วมกัน (Shared Understanding) ความสามารถในการเจรจา ทำความเข้าใจเป้าหมาย และสร้างองค์ความรู้ร่วมกัน
2. การเจรจาต่อรองและแบ่งปันบทบาท (Negotiation and Role Allocation) ความสามารถในการจัดการทรัพยากรและกระจายหน้าที่ความรับผิดชอบ
3. การสื่อสารและการจัดการความขัดแย้ง (Communication and Conflict Management) ความสามารถในการแลกเปลี่ยนข้อมูลอย่างมีประสิทธิภาพ และการจัดการความเห็นพ้องต้องกันอย่างสร้างสรรค์

4. การกำกับติดตามและการประเมินผลกลุ่ม (Monitoring and Group Evaluation) ความสามารถในการตรวจสอบความก้าวหน้าของกลุ่มและให้ข้อมูลย้อนกลับระหว่างกัน

เมื่อสามารถระบุองค์ประกอบย่อยเหล่านี้ได้ ผู้ออกแบบแบบวัดจึงจะสามารถสร้างสถานการณ์หรือภาระงาน (Tasks) ที่กระตุ้นให้ผู้เรียนแสดงออกถึงองค์ประกอบเหล่านี้ เพื่อนำไปสู่การวินิจฉัยจุดแข็งหรือจุดบกพร่องในแต่ละมิติได้

การวิเคราะห์แนวทางการพัฒนาแบบวัดวินิจฉัยสมรรถนะในยุคดิจิทัล

การเข้ามาของเทคโนโลยีดิจิทัลได้ปฏิวัติวงการการวัดและประเมินผล โดยเฉพาะการประเมินสมรรถนะที่ซับซ้อนซึ่งยากต่อการวัดด้วยแบบทดสอบแบบดั้งเดิม เทคโนโลยีได้เปิดพรมแดนใหม่ในการเก็บรวบรวมข้อมูลและวิเคราะห์ผลลัพธ์ อย่างไรก็ตาม แนวทางเหล่านี้ต่างก็มีจุดแข็งและข้อจำกัดที่ต้องวิพากษ์อย่างรอบคอบ

1. การประเมินเชิงปฏิบัติการที่ใช้เทคโนโลยี (Technology-enhanced Performance Assessment) แนวทางที่ทรงพลังที่สุดในการวัดสมรรถนะ คือ การให้ผู้เรียนลงมือปฏิบัติในสถานการณ์ที่ซับซ้อนและสมจริงในยุคดิจิทัล สิ่งนี้สามารถทำได้ผ่านสถานการณ์จำลอง และเกมการเรียนรู้ ซึ่งข้อดีสถานการณ์จำลองและเกมสามารถสร้างภาระงานที่ซับซ้อน จะช่วยกระตุ้นให้ผู้เรียนต้องแสดงออกถึงสมรรถนะ เช่น การทำงานร่วมกัน หรือการแก้ปัญหา ในลักษณะที่เป็นกระบวนการมากกว่าการตอบคำถามสั้น ๆ จุดแข็งที่สำคัญที่สุดคือเทคโนโลยีเหล่านี้เปิดโอกาสให้เกิดการประเมินแบบซ่อนเร้น (Stealth Assessment) ในส่วนของบทบาทของข้อมูลบันทึกพฤติกรรมแทนที่จะประเมินเฉพาะคำตอบสุดท้าย เทคโนโลยีสามารถบันทึกข้อมูลบันทึกพฤติกรรม (Log Data) หรือข้อมูลกระบวนการ (Process Data) ได้อย่างละเอียด ข้อมูลเหล่านี้ เช่น การติดตามการคลิก (Clickstream) ลำดับการเข้าถึงข้อมูล หรือเวลาที่ใช้ในแต่ละขั้นตอน (Time-on-task) ได้กลายเป็นหลักฐานเชิงประจักษ์ชิ้นใหม่สำหรับการวินิจฉัยกระบวนการคิด (Cognitive Processes) เช่น การที่ผู้เรียนใช้เวลาในหน้าวิเคราะห์ข้อมูลนาน อาจสะท้อนถึงความลังเล (Hesitation) หรือกำลังใช้กลยุทธ์การแก้ปัญหาที่ลึกลับ ซึ่งขณะที่การคลิกข้ามอย่างรวดเร็วอาจสะท้อนถึงการขาดความไตร่ตรอง แต่อย่างไรก็ตามการพัฒนาสถานการณ์จำลองที่มีคุณภาพสูงก็มีข้อจำกัดในด้านของการมีต้นทุนที่สูงมาก ทั้งยังมีความท้าทายในการวิเคราะห์ข้อมูล Log Data ซึ่งเป็นข้อมูลขนาดใหญ่ (Big Data) ที่ต้องอาศัยเทคนิคขั้นสูงในการสกัดความหมายทางการศึกษา (Educational Data Mining) ซึ่งในประเทศไทย สถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) ได้พัฒนาข้อสอบ PISA-

Based Test for Schools ที่เน้นการวัดความสามารถในการอ่าน คณิตศาสตร์ และวิทยาศาสตร์ในเชิงสมรรถนะ (สถาบันทดสอบทางการศึกษาแห่งชาติ, 2562) โดยมีการใช้สถานการณ์จำลองและคำถามแบบโต้ตอบ อย่างไรก็ตาม การนำไปใช้ในระดับห้องเรียนยังมีข้อจำกัด โดยเฉพาะในโรงเรียนขนาดเล็กที่ขาดทรัพยากร ถึงแม้เกมจะสามารถดึงดูดความสนใจของผู้เรียนได้ดี แต่ความท้าทายอยู่ที่การแปลงข้อมูล Log Data ให้เป็นข้อมูลการวินิจฉัยที่ครูเข้าใจและนำไปใช้ได้ ซึ่งต้องอาศัยความเชี่ยวชาญทั้งด้านเทคโนโลยีและการวัดผล

2. การใช้โมเดลวินิจฉัยเชิงปริมาณขั้นสูง (Advanced Quantitative Diagnostic Models) สิ่งที่ทำให้ควบคู่ไปกับการออกแบบสถานการณ์ คือการพัฒนาโมเดลทางสถิติเพื่อการวินิจฉัยสำหรับสมรรถนะซึ่งมีหลายองค์ประกอบย่อย (Multidimensional) แบบจำลองการวินิจฉัยการรู้ (Cognitive Diagnostic Models: CDMs) หรือที่เรียกในวงกว้างว่า Diagnostic Classification Models (DCMs) จึงทำให้มีศักยภาพสูง ซึ่งสามารถอธิบายได้ดังนี้

2.1 ด้านศักยภาพ DCMs ถูกออกแบบมาเพื่อก้าวข้ามข้อจำกัดของคะแนนรวม (Total Score) โดยโมเดลเหล่านี้สามารถระบุรูปแบบ (Profiles) ของการเรียนรู้ที่แตกต่างกันของผู้เรียน โดยจำแนกตามการมีหรือไม่มี (Mastery/Non-mastery) มิติสมรรถนะย่อย ๆ (Attributes) เช่น โมเดลอาจวินิจฉัยได้ว่า ผู้เรียน A บกพร่องที่การระบุปัญหา แต่เชี่ยวชาญการเสนอทางเลือก ในขณะที่ผู้เรียน B เชี่ยวชาญการระบุปัญหา แต่บกพร่องการประเมินทางเลือก

2.2 ด้านความท้าทายเชิงวิพากษ์ (Critical Challenge) คือ ความยากในการแปลผล (Interpretability) โมเดล DCMs มีความซับซ้อนทางสถิติสูง ผลลัพธ์ที่ได้มักอยู่ในรูปของคะแนนจะเป็นของการมีคุณลักษณะย่อย ซึ่งเป็นข้อมูลที่ครูผู้สอนทั่วไปอาจนำไปใช้ในบริบทจริงหน้าชั้นเรียนได้ยาก หากขาดการออกแบบรายงานผล (Reporting) ที่มีประสิทธิภาพและเข้าใจง่าย

จากการศึกษางานวิจัยของนักการศึกษาพบว่า แม้แบบวัดจะมีคุณภาพทางจิตมิติที่ดี และโมเดลสามารถวินิจฉัยจุดบกพร่องได้อย่างละเอียด แต่เมื่อนำไปทดลองใช้กับครูจริง ครูส่วนใหญ่ยังไม่สามารถตีความและนำผลไปใช้ได้อย่างมีประสิทธิภาพ เนื่องจากขาดความเข้าใจในเชิงสถิติขั้นสูง และรายงานผลที่ได้มีรูปแบบที่ซับซ้อนเกินไป การศึกษานี้เน้นย้ำถึงความจำเป็นที่จะต้องมีการสื่อสารผลการวินิจฉัย (Diagnostic Communication) ที่เข้าถึงได้ง่าย โดยอาจต้องมีการพัฒนา Dashboard หรือระบบรายงานผลแบบภาพ (Visual Report) ที่แปลงข้อมูลทางสถิติให้เป็นข้อมูลที่ครูเข้าใจและสามารถนำไปปฏิบัติได้จริง

3. ความเที่ยงตรง (Validity) และการตีความผลการวินิจฉัย ประเด็นข้างต้นนำมาสู่การวิพากษ์ซึ่งเป็นหัวใจสำคัญของแบบวัดวินิจฉัย นั่นคือความเที่ยงตรง (Validity) ในการประเมินสมัยใหม่ การประเมินไม่ได้หยุดอยู่แค่ความเที่ยงตรงเชิงเนื้อหา (Content Validity) หรือเชิงโครงสร้าง (Construct Validity) เท่านั้น สำหรับแบบวัดวินิจฉัยซึ่งมีเป้าหมายเพื่อการพัฒนาผู้เรียน (Formative Use) ความเที่ยงตรงรูปแบบใหม่ที่ต้องให้ความสำคัญสูงสุด คือ ความเที่ยงตรงเชิงการตีความและการใช้ผล (Interpretive Validity and Use) (Kane, 2006; Messick, 1989) ความท้าทายของแบบวัดวินิจฉัยสมัยใหม่จึงไม่ได้อยู่ที่การสร้างโมเดล DCMs ที่ซับซ้อนที่สุด หรือการเก็บ Log Data ที่ละเอียดที่สุด แต่อยู่ที่ว่า ผลการวินิจฉัยนั้นสามารถสื่อสารไปยังผู้ใช้ เช่น ครู หรือผู้เรียน ได้อย่างชัดเจนและถูกต้องหรือไม่ และผู้ใช้สามารถนำผลการวินิจฉัยนั้นไปใช้ในการตัดสินใจเพื่อปรับปรุงการสอนหรือการเรียนรู้ได้จริงหรือไม่ หากผลการวินิจฉัย มีความซับซ้อนทางเทคนิคมากเกินไปที่ครูจะตีความได้ แบบวัดนั้น แม้จะมีความเที่ยงตรงเชิงสถิติสูงเพียงใด ก็ถือว่าล้มเหลวในฐานะเครื่องมือวินิจฉัยเพื่อการพัฒนา

ดังนั้น การออกแบบแบบวัดวินิจฉัยจึงต้องให้ความสำคัญกับการออกแบบรายงานผล (Report Design) และการสนับสนุนการใช้งานของครู (Teacher Support) ควบคู่กันไปอย่างแยกไม่ออก ซึ่งปัจจัยสำคัญที่ทำให้ครูนำแบบวัดวินิจฉัยไปใช้ได้อย่างมีประสิทธิภาพ ได้แก่ 1) การมีรายงานผลที่เข้าใจง่าย มีภาพประกอบชัดเจน 2) การมีคำแนะนำเชิงปฏิบัติ (Actionable Recommendations) ที่ระบุว่าจะควรทำอย่างไรต่อกับผู้เรียนที่มีจุดบกพร่องในแต่ละมิติ และ 3) การได้รับการอบรมและการสนับสนุนอย่างต่อเนื่อง ไม่ใช่แค่การมอบแบบวัดให้ครูเพียงอย่างเดียว

ข้อเสนอแนะแนวคิดสำหรับการออกแบบแบบวัดวินิจฉัย

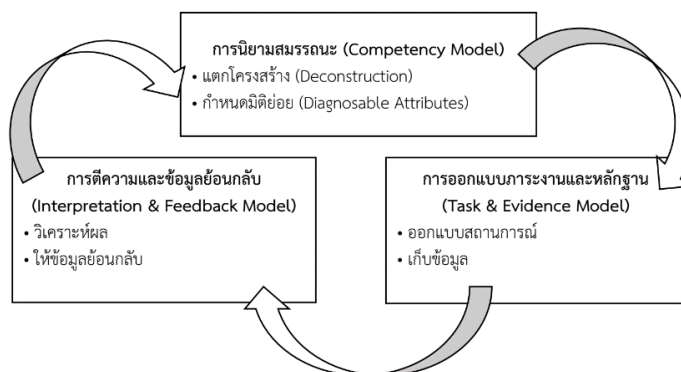
บทความนี้ขอเสนอแนวคิดในการออกแบบแบบวัดวินิจฉัยสมรรถนะแห่งศตวรรษที่ 21 ที่มุ่งเน้น “การนำไปใช้ได้จริง” (Actionability) และ “ความสมบูรณ์ของข้อมูล” (Data Richness) โดยมีองค์ประกอบสำคัญดังนี้

1. โมเดลการออกแบบวงจรวินิจฉัยเพื่อการพัฒนา (The Formative Diagnostic Cycle Model) แทนที่จะมุ่งเน้นเพียงการสร้างเครื่องมือวัด เราจำเป็นต้องเปลี่ยนกระบวนทัศน์ไปสู่การออกแบบกระบวนการวินิจฉัยที่ครบวงจร บทความนี้ขอเสนอโมเดลเชิงแนวคิดที่เชื่อมโยง 3 องค์ประกอบหลัก ซึ่งสอดคล้องกับกรอบแนวคิด Evidence-Centered Design ของ Mislevy, et al. (2003) ดังนี้

1.1 การนิยามสมรรถนะเป้าหมาย (Competency Model) เริ่มต้นจากการแตกโครงสร้างสมรรถนะ (Deconstruction) ที่เป็นนามธรรม เช่น การทำงานร่วมกัน ให้เป็นมิตีย่อยที่สามารถวินิจฉัยได้ (Diagnosable Attributes) ที่เป็นรูปธรรมและสังเกตได้ ซึ่งเป็นหัวใจสำคัญของแบบจำลองผู้เรียน (Student Model) การทำขั้นตอนนี้ต้องอาศัย 1) การทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้อง 2) การวิเคราะห์งาน (Task Analysis) จากผู้เชี่ยวชาญ 3) การปรับให้เหมาะสมกับบริบท โดยคำนึงถึงวัฒนธรรม ค่านิยม และสภาพแวดล้อมการเรียนรู้ของผู้เรียน

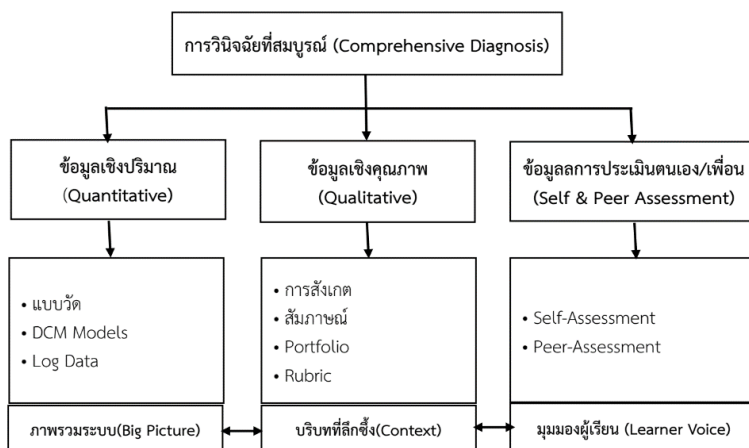
1.2 การออกแบบภาระงานและหลักฐาน (Task & Evidence Model) พัฒนาสถานการณ์ (Scenarios) หรือภาระงานเชิงปฏิบัติ (Performance Tasks) ที่ออกแบบมาโดยเจาะจง (Intentionally Designed) เพื่อกระตุ้นให้ผู้เรียนต้องใช้มิตีย่อยเหล่านั้น ภาระงานต้องสามารถสร้างหลักฐาน (Evidence) ที่ชัดเจน เช่น ข้อมูล Log Data หรือคำตอบ ที่สามารถเชื่อมโยงกลับไปยังมิตีย่อยที่ต้องการวัดได้ หลักการสำคัญในการออกแบบภาระงาน คือ 1) ความเป็นจริง (Authenticity) ภาระงานควรสะท้อนสถานการณ์ในชีวิตจริงหรือบริบทวิชาชีพ 2) ความซับซ้อนที่เหมาะสม (Appropriate Complexity) ไม่ง่ายจนเกินไปหรือยากจนเกินไป ต้องท้าทายพอที่จะกระตุ้นให้ผู้เรียนแสดงสมรรถนะที่แท้จริง และ 3) ความหลากหลาย (Variability) ควรมีภาระงานหลายรูปแบบเพื่อวัดมิติเดียวกัน เพื่อเพิ่มความน่าเชื่อถือ

1.3 การตีความและข้อมูลย้อนกลับ (Interpretation & Feedback Model) นี่เป็นส่วนที่สำคัญที่สุดสำหรับการวินิจฉัย โมเดลนี้ต้องระบุว่าข้อมูลเชิงปริมาณ เช่น ผลจาก DCMs หรือข้อมูลเชิงกระบวนการ เช่น Log Data จะถูกสังเคราะห์และแปลผลอย่างไร เพื่อสร้างเป็น “ข้อมูลย้อนกลับที่นำไปปฏิบัติได้” (Actionable Feedback) ผลลัพธ์ต้องไม่ใช่แค่คะแนน แต่เป็นโปรไฟล์ที่ระบุจุดแข็ง/จุดอ่อนในแต่ละมิตีย่อย เพื่อตอบคำถามว่า ผู้เรียนบกพร่องที่จุดใด และเพราะเหตุใด องค์ประกอบสำคัญของข้อมูลย้อนกลับที่ดี คือ 1) ความชัดเจน (Clarity) ใช้ภาษาที่เข้าใจง่าย หลีกเลี่ยงศัพท์เทคนิคที่ซับซ้อน 2) ความเฉพาะเจาะจง (Specificity) ระบุชัดเจนว่าจุดบกพร่องคืออะไร 3) คำแนะนำเชิงปฏิบัติ (Actionable Suggestions) เสนอวิธีการพัฒนาหรือกิจกรรมที่เหมาะสม 4) การเชื่อมโยงกับบริบท (Contextualization) เชื่อมโยงผลการวินิจฉัยกับเป้าหมายการเรียนรู้และหลักสูตร



ภาพที่ 1 วงจรวินิจฉัยเพื่อการพัฒนา (Formative Diagnostic Cycle Model)

2. การบูรณาการข้อมูลแบบผสมผสาน (Triangulation of Diagnostic Data) การพึ่งพาข้อมูลจากแบบวัดเพียงแหล่งเดียว ไม่ว่าจะเป็น DCMs หรือ Log Data อาจนำไปสู่การวินิจฉัยที่คลาดเคลื่อน ดังนั้นการบูรณาการข้อมูลแบบสามเส้า (Data Triangulation) จะช่วยเพิ่มความเที่ยงตรงเชิงการตีความ (Interpretive Validity) แนวทางนี้เสนอให้มองผลการวินิจฉัยจากแบบวัดว่าเป็นข้อมูลตั้งต้น (Flagging Data) มากกว่าคำตัดสินสุดท้าย (Verdict) เพราะข้อมูลเชิงปริมาณจากแบบวัดให้ภาพมุมกว้างที่รวดเร็วและเป็นระบบ ว่าผู้เรียนคนใดมีแนวโน้มบกพร่องในมิติใด เช่น ขาดทักษะการจัดการความขัดแย้ง และมีความเป็นมาตรฐานและสามารถเปรียบเทียบได้ ส่วนข้อมูลเชิงคุณภาพจากครู เมื่อครูได้รับสัญญาณ (Flag) นี้ ครูสามารถใช้การสังเกตในชั้นเรียนแบบมุ่งเป้า (Targeted Observation) หรือการสัมภาษณ์สั้น ๆ (Brief Interview) เพื่อยืนยันหรือทำความเข้าใจบริบทของความบกพร่องนั้นในสถานการณ์จริง ให้ข้อมูลเชิงลึกเกี่ยวกับสาเหตุและบริบท พร้อมทั้งสะท้อนพัฒนาการในชีวิตประจำวัน และข้อมูลจากการประเมินตนเองและเพื่อน (Self & Peer Assessment) ช่วยให้ได้มุมมองจากผู้เรียนเอง ช่วยพัฒนาทักษะการไตร่ตรอง (Metacognition) และเพิ่มความรับผิดชอบต่อการเรียนรู้ของตนเอง การบูรณาการข้อมูลลักษณะนี้ เช่น การใช้ข้อมูลเชิงปริมาณนำ และใช้ข้อมูลเชิงคุณภาพตามเพื่ออธิบาย จะช่วยลดจุดบอดของข้อมูลแต่ละประเภท และทำให้ได้ภาพการวินิจฉัยที่สมบูรณ์ (Holistic Diagnosis) และน่าเชื่อถือยิ่งขึ้นก่อนนำไปวางแผนการสอน



ภาพที่ 2 การบูรณาการข้อมูลแบบสามเสา

3. ความสำคัญในการปฏิบัติและการพัฒนาหลักสูตร เป้าหมายสูงสุดของแบบวัดวินิจฉัย คือ การปรับปรุงการเรียนการสอน (Instructional Improvement) ผลลัพธ์ที่ได้จากกระบวนการวินิจฉัยตามข้อเสนอข้างต้น ควรนำไปสู่การปฏิบัติ ดังนี้

3.1 การออกแบบกิจกรรมการเรียนรู้ที่มุ่งเป้า (Targeted Instruction) หากผลการวินิจฉัยระบุว่านักเรียนกลุ่มหนึ่งมีความบกพร่องในมิตีย่อยเฉพาะ เช่น การประเมินความน่าเชื่อถือของข้อมูล ซึ่งเป็นส่วนหนึ่งของ การคิดเชิงวิพากษ์ ครูไม่จำเป็นต้องสอนซ้ำทั้งสมรรถนะ แต่สามารถออกแบบกิจกรรมแทรกแซง (Intervention) หรือการเสริมต่อการเรียนรู้ (Scaffolding) ที่มุ่งเน้นแก้ไขจุดบกพร่องนั้นโดยตรง

3.2 การพัฒนาหลักสูตรโดยใช้ข้อมูล (Data-Driven Curriculum Development) หากข้อมูลการวินิจฉัยในภาพรวม เช่น ระดับโรงเรียน แสดงให้เห็นว่าผู้เรียนส่วนใหญ่มีจุดอ่อนร่วมกันในมิติใดมิติหนึ่งอย่างต่อเนื่อง ข้อมูลนี้จะกลายเป็นหลักฐานเชิงประจักษ์ที่สำคัญสำหรับผู้พัฒนาหลักสูตรในการทบทวนและปรับปรุงลำดับการสอน หรือเพิ่มการเน้นย้ำในมิตินั้น ๆ ในหลักสูตรสถานศึกษาต่อไป

ตารางที่ 1 เปรียบเทียบแนวทางการวัดสมรรถนะ

แนวทาง	จุดแข็ง	ข้อจำกัด	ความเหมาะสมกับบริบทไทย
แบบทดสอบแบบดั้งเดิม	- ใช้ง่าย - ต้นทุนต่ำ - คำนวณง่าย	- วัดความรู้เท่านั้น - ไม่สามารถวินิจฉัยเชิงลึก - ไม่เหมาะสมกับสมรรถนะซับซ้อน	ระดับพื้นฐาน เหมาะกับการคัดกรอง
สถานการณ์จำลองดิจิทัล	- วัดสมรรถนะได้เป็นกระบวนการ - เก็บ Log Data - น่าสนใจ	- ต้นทุนสูง - ต้องการโครงสร้างพื้นฐาน - ซับซ้อนในการพัฒนา	โรงเรียนขนาดใหญ่ มีงบประมาณและเทคโนโลยี
Performance Task (Paper-based)	- วัดสมรรถนะจริง - ไม่ต้องพึ่งเทคโนโลยี - ปรับใช้ง่าย	- ใช้เวลานาน - การให้คะแนนอาจไม่สม่ำเสมอ - ต้องการ Rubric ที่ดี	ทุกระดับ: เหมาะสมที่สุดสำหรับบริบทไทย
DCM Models	- วินิจฉัยเชิงลึก - แม่นยำทางสถิติ - ระบุมิติย่อยได้	- ซับซ้อนในการตีความ - ต้องการความเชี่ยวชาญ - ครูใช้ยาก	ระดับสูง: ต้องมีการสนับสนุนและฝึกอบรม
การบูรณาการข้อมูล (Triangulation)	- ข้อมูลสมบูรณ์ - ความเที่ยงตรงสูง - ครูมีส่วนร่วม	- ใช้เวลามาก - ต้องการการประสานงาน	แนะนำสูง: เหมาะกับการตัดสินใจสำคัญ

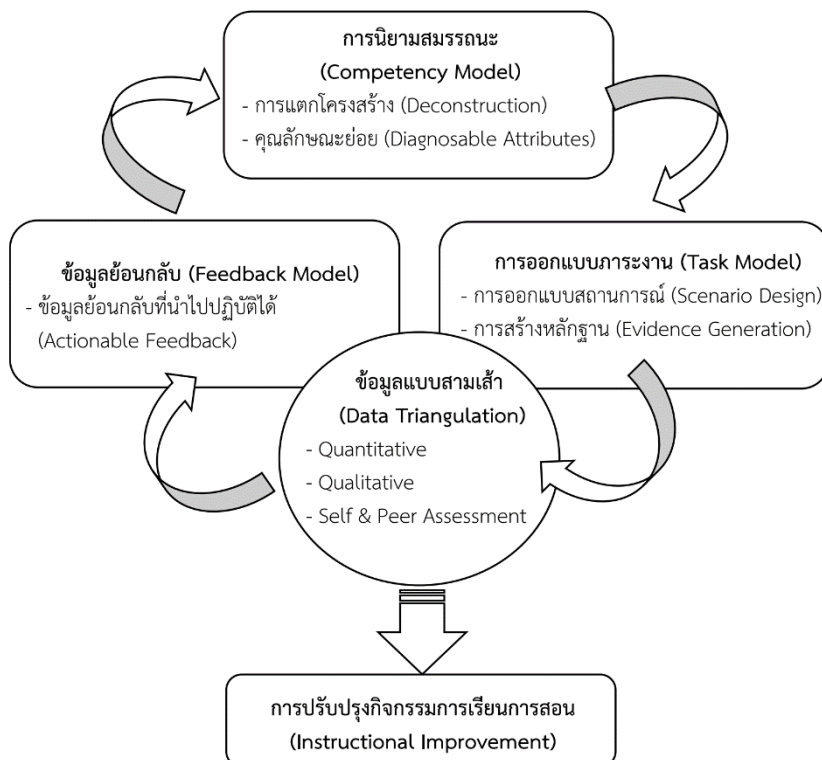
บทสรุป

การเปลี่ยนผ่านกระบวนการศึกษาศึกษาไปสู่การพัฒนาสมรรถนะแห่งศตวรรษที่ 21 ได้สร้างความท้าทายอันยิ่งใหญ่ต่อแวดวงการศึกษาและประเมินผลการศึกษา ภาระงานการประเมินแบบดั้งเดิมที่มุ่งเน้นการวัดความรู้เชิงเนื้อหาไม่เพียงพออีกต่อไปต่อการวัดสมรรถนะที่ซับซ้อนและเป็นนามธรรม บทความวิชาการนี้ได้อภิปรายและสังเคราะห์แนวทางสำคัญในการก้าวข้ามความท้าทายนี้ โดยเน้นย้ำถึงความจำเป็นเร่งด่วนของแบบวัดวินิจฉัย (Diagnostic Assessment) ที่สามารถให้ข้อมูลเชิงลึกเพื่อพัฒนาผู้เรียนอย่างแท้จริง โดยเริ่มตั้งแต่การวิเคราะห์ความจำเป็นในการแตกโครงสร้าง (Deconstruction) สมรรถนะที่เป็นนามธรรมให้กลายเป็นมิติย่อยที่สามารถวินิจฉัยได้ (Diagnosable Attributes) ซึ่งเป็นรากฐานสำคัญของการประเมินเชิงวินิจฉัยที่มีความหมาย

จากนั้นได้วิเคราะห์และวิพากษ์แนวทางการประเมินสมัยใหม่ที่ใช้เทคโนโลยีเป็นฐาน เช่น สถานการณ์จำลอง การวิเคราะห์ Log Data และการใช้แบบจำลองการวินิจฉัยเชิงปริมาณ (DCMs) โดยชี้ให้เห็นทั้งศักยภาพอันงดงามและข้อจำกัดที่สำคัญของเครื่องมือเหล่านี้ ข้อค้นพบสำคัญ คือ แม้เครื่องมือและทฤษฎีการวัดจะมีความก้าวหน้าอย่างมากในทางเทคนิค แต่กลับเผชิญความท้าทายอย่างยิ่งในมิติของความเที่ยงตรงเชิงการตีความ (Interpretive Validity) และการนำผลไปใช้ได้จริง ในชั้นเรียน โดยเฉพาะในบริบทของประเทศไทยที่ยังมีช่องว่างอย่างมากระหว่างทฤษฎีการวัดขั้นสูง กับการพร้อมและความเข้าใจของครูในสนาม ปัญหานี้ไม่ได้เกิดจากขาดเทคโนโลยีหรือทฤษฎีที่ดี แต่เกิดจากการขาดการเชื่อมโยงระหว่างความซับซ้อนทางวิชาการ กับการเรียบง่ายในการปฏิบัติ จึงได้เสนอแนวคิดการออกแบบวงจรวินิจฉัยเพื่อการพัฒนา (Formative Diagnostic Cycle) ที่เชื่อมโยงสมรรถนะ ภาระงาน และการให้ข้อมูลย้อนกลับที่นำไปปฏิบัติได้ (Actionable Feedback) อย่างเป็นระบบ พร้อมทั้งเสนอแนวคิดการบูรณาการข้อมูลแบบสามเส้า (Triangulation) ระหว่าง ข้อมูลเชิงปริมาณจากแบบวัด ข้อมูลเชิงคุณภาพจากการสังเกตของครู และข้อมูลจากการประเมินตนเองและเพื่อน เพื่อให้ได้ภาพการวินิจฉัยที่สมบูรณ์และน่าเชื่อถือ สำหรับประเทศไทยการปรับใช้ ควรคำนึงถึงความพร้อมที่หลากหลายของสถานศึกษา ได้แก่ 1) การพัฒนาเครื่องมือที่ไม่ต้องพึ่งพาเทคโนโลยีสูงเพื่อให้โรงเรียนทุกขนาดสามารถเข้าถึงได้ 2) การสร้างชุมชนการเรียนรู้วิชาชีพครู (PLC) เพื่อพัฒนาความสามารถของครูในการใช้และตีความผลการวินิจฉัยอย่างต่อเนื่อง และ 3) การบูรณาการผลการวินิจฉัยเข้ากับระบบประกันคุณภาพภายในสถานศึกษา (IQA) เพื่อให้เกิดการใช้ข้อมูลเชิงประจักษ์ในการพัฒนาอย่างเป็นระบบ แนวทางเหล่านี้ไม่ได้เป็นเพียงแนวคิดทางทฤษฎี แต่ได้รับการสนับสนุนด้วยหลักฐานเชิงประจักษ์จากงานวิจัยในบริบทไทยที่แสดงให้เห็นถึงความเป็นไปได้และประสิทธิผล ท้ายที่สุด หัวใจของแบบวัดวินิจฉัยไม่ได้อยู่ที่ความแม่นยำของตัวเลข ไม่ได้อยู่ที่ความซับซ้อนของโมเดลทางสถิติ และไม่ได้อยู่ที่ความทันสมัยของเทคโนโลยี แต่อยู่ที่ความสามารถในการจุดประกายการพัฒนาการเรียนรู้ของผู้เรียนได้อย่างตรงจุด โดยให้ข้อมูลที่ชัดเจน เข้าใจง่าย และนำไปปฏิบัติได้จริงแก่ครูผู้ทุ่มเทในแนวหน้า เพื่อให้ทุกคนสามารถร่วมกัน สร้างระบบการศึกษาที่พัฒนาผู้เรียนให้พร้อมรับมือกับความท้าทายในศตวรรษที่ 21 อย่างแท้จริง

องค์ความรู้ใหม่

บทความวิชาการนี้ได้เสนอแนวคิดใหม่ที่เน้นการนำไปใช้ได้จริง (Actionability) ดังนี้



ภาพที่ 3 แผนผังองค์ความรู้การพัฒนาแบบวัฏจักรวินิจฉัยสำหรับสมรรถนะแห่งศตวรรษที่ 21

จากแผนผังองค์ความรู้ สามารถอธิบายได้ว่า องค์ความรู้ใหม่ถูกแบ่งออกเป็น 2 ส่วนที่ทำงานร่วมกัน ได้แก่ วงจรการดำเนินงาน และกลไกการตรวจสอบข้อมูล เพื่อเป้าหมายสูงสุด คือ การพัฒนาผู้เรียน โดยมีรายละเอียดดังนี้

1. วงจรวินิจฉัยเพื่อการพัฒนา (Formative Diagnostic Cycle Model) เป็นส่วนที่แสดงถึงลำดับขั้นตอนการทำงานที่เป็นวงจรต่อเนื่องเพื่อเปลี่ยนจากการวัดผลแบบเดิมมาเป็นการวินิจฉัยเพื่อพัฒนา โดยมีขั้นตอนดังนี้

ขั้นที่ 1 การนิยามสมรรถนะ (Competency Model) เป็นการเริ่มต้นจากการนำสมรรถนะที่เป็นนามธรรมมาแตกโครงสร้าง (Deconstruction) ให้กลายเป็นมิตีย่อยที่วัดและสังเกตได้จริง (Diagnosable Attributes)

ขั้นที่ 2 การออกแบบภาระงานและหลักฐาน (Task & Evidence Model) เป็นการสร้างสถานการณ์จำลองหรือภาระงาน (Scenario Design) ที่ออกแบบมาโดยเฉพาะ เพื่อกระตุ้นให้ผู้เรียนแสดงมิตีย่อยนั้น ๆ ออกมาเป็นหลักฐานเชิงประจักษ์ (Evidence Generation)

ขั้นที่ 3 การตีความและข้อมูลย้อนกลับ (Interpretation & Feedback Model) เป็นการนำหลักฐานมาวิเคราะห์เพื่อสร้างข้อมูลย้อนกลับที่นำไปปฏิบัติได้ (Actionable Feedback) โดยขั้นตอนนี้จะได้รับแรงสนับสนุนสำคัญจากกลไกที่สอง

2. การบูรณาการข้อมูลแบบสามเส้า (Data Triangulation) เป็นส่วนที่แสดงถึงการปฏิสัมพันธ์ของข้อมูล เพื่อเพิ่มความเที่ยงตรงในการวินิจฉัย โดยแบ่งเป็น

2.1 ข้อมูลเชิงปริมาณ (Quantitative) เช่น ผลจากแบบจำลอง DCMs หรือ Log Data ทำหน้าที่เป็นสัญญาณแจ้งเตือน (Flagging Data) เพื่อชี้เป้าในภาพรวมว่าใครมีจุดบกพร่องในมิติใด

2.2 ข้อมูลเชิงคุณภาพ (Qualitative) เป็นข้อมูลที่ได้จากการสังเกตหรือสัมภาษณ์ของครูเพื่อเข้าไปดูบริบทเชิงลึกของสัญญาณแจ้งเตือนนั้น เพื่อยืนยันว่าสิ่งที่ระบบตรวจพบมีสาเหตุมาจากอะไร

2.3 ข้อมูลจากผู้เรียน (Self & Peer Assessment) เป็นข้อมูลที่ได้จากผู้เรียนเพื่อให้เห็นถึงการรับรู้ตนเองและการไตร่ตรอง ซึ่งช่วยให้การวินิจฉัยมีความสมบูรณ์รอบด้าน (Holistic Diagnosis)

ความเชื่อมโยงและการปฏิสัมพันธ์ (Interactions) ของจุดที่สำคัญที่สุด คือ การที่ทั้งสองส่วนนี้ เชื่อมโยงกันที่การตีความของผลลัพธ์จากตัวเลข สู่การปฏิบัติ ซึ่งเป็นข้อมูลเชิงปริมาณที่ซับซ้อน เช่น DCMs ซึ่งจะถูกลดทอนความยุ่งยากด้วยการนำมาบูรณาการกับความเห็นของครู ทำให้ครูสามารถตัดสินใจออกแบบกิจกรรมการเรียนรู้ที่มุ่งเป้า (Targeted Instruction) ได้ทันที นอกจากนี้ ปฏิสัมพันธ์นี้ยังช่วยแก้ปัญหาสำคัญในบริบทไทยที่ครูอาจไม่เชี่ยวชาญสถิติขั้นสูงหรือทฤษฎี แต่มีความเชี่ยวชาญในการสังเกตผู้เรียน แผนผังองค์ความรู้นี้จะช่วยเปลี่ยนความซับซ้อนทางทฤษฎีให้กลายเป็นความเรียบง่ายในการปฏิบัติในห้องเรียน อันเป็นการเชื่อมช่องว่างระหว่างทฤษฎีและการปฏิบัติได้จริง

เอกสารอ้างอิง

- กระทรวงศึกษาธิการ. (2560). *หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560)*. โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย.
- สถาบันทดสอบทางการศึกษาแห่งชาติ. (2562). *รายงานผลการทดสอบ PISA-Based Test for Schools ของโรงเรียนไทย ปี 2562*. สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน).
- สำนักงานเลขาธิการสภาการศึกษา. (2560). *แผนการศึกษาแห่งชาติ พ.ศ. 2560-2579*. กระทรวงศึกษาธิการ.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7- 74. <http://dx.doi.org/10.1080/0969595980050102>
- Griffin, P. & Care, E. (Eds.). (2015). *Assessment and teaching of 21st century skills: Methods and approach*. Springer.
- Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative Problem-solving skills. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills: Methods and approach*. Springer.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement*. Praeger.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement*. American Council on Education/Macmillan.
- Mislevy, R. J., Steinberg, L. S. & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1(1), 3–62.
- OECD. (2018). *Preparing our youth for an inclusive and sustainable world: The OECD PISA global competence framework*. OECD Publishing.
- Partnership for 21st Century Skills (P21). (2009). *Framework for 21st century learning. Creative Education*, 6(2) . <https://www.scirp.org/reference/referencespapers?referenceid=1412323>



- Glaser, R., Chudowsky, N., & Pellegrino, J.W. (Eds.). (2001) *Knowing What Students Know. The Science and Design of Educational Assessment*. National Academies Press.
- Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications*. Guilford Press.