



Academic Article

COGNITIVE DIAGNOSTIC ASSESSMENT: CONCEPTS FOR APPLICATION
IN DIAGNOSIS READING COMPREHENSION ABILITY

การประเมินวินิจฉัยทางปัญญา: แนวคิดเพื่อประยุกต์ใช้ในการวินิจฉัย
ความสามารถด้านการอ่านเพื่อความเข้าใจ

Received: December 14, 2023

Revised: March 19, 2024

Accepted: March 21, 2024

Wilaiorn Khamman^{1*} and Samran Meejang²

วิไลภรณ์ คำนั่น^{1*} และสำราญ มีแจ้ง²

^{1,2}Faculty of Education, Naresuan University, Phitsanulok 65000, Thailand

^{1,2}คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร

*Corresponding Author, E-mail: nellynetchanok@gmail.com

Abstract

Cognitive diagnostic assessments have two important purposes: 1) to identify areas in which students are not yet well-versed and 2) to provide suggestions about the causes or possible reasons for the student's lack of knowledge. Teachers can use the results of the diagnosis to develop students in the areas where they are lacking. There is a method for synthesizing diagnostic information with statistical analysis using Cognitive Diagnostic Models (CDMs). Analyzing respondents' responses to exams according to IRT Response Theory through complex statistical models characterized as models. A latent layer that can classify test takers into the latent layer. It specifies those who are knowledgeable and non-knowledgeable in each ability attribute. A cognitive diagnostic assessment is a type of assessment used to explore the relationship between mental processes and the strategies test takers use to solve problems or answer questions on a test. The information obtained is feedback on the strengths and weaknesses of each test taker's sub-skills or characteristics. Therefore, it is appropriate in the context of assessing reading, which is a brain process according to cognitive theory. There are many ways to apply cognitive diagnostic assessments to language abilities, but one that provides high construct validity is by designing assessments using a cognitive design system (CDS). The relationship between questions and ability characteristics is used to create a cognitive diagnostic test used in cognitive assessment.

Keywords: Cognitive Diagnostic Assessment, Reading Comprehension

บทคัดย่อ

การประเมินวินิจฉัยทางปัญญามีวัตถุประสงค์สำคัญ 2 ประการ คือ 1) เพื่อระบุประเด็นที่นักเรียนยังไม่รอบรู้ และ 2) เพื่อให้ข้อเสนอแนะเกี่ยวกับสาเหตุหรือเหตุผลที่เป็นไปได้เกี่ยวกับการไม่รอบรู้ในทักษะย่อยหรือคุณลักษณะนั้นๆ ครูผู้สอนสามารถนำผลการวินิจฉัยไปใช้ประโยชน์ในการพัฒนาผู้เรียนได้ตรงจุดที่บกพร่อง โดยมีวิธีการสังเคราะห์สารสนเทศผลการวินิจฉัยด้วยการวิเคราะห์ทางสถิติโดยใช้โมเดลวินิจฉัยทางปัญญา (CDMs) วิเคราะห์การตอบสนองข้อสอบของผู้ตอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) ผ่านโมเดลทางสถิติที่ซับซ้อนมีลักษณะเป็นโมเดลชั้นแฝงที่สามารถจำแนกผู้สอบเข้าชั้นแฝง ซึ่งระบุผู้รอบรู้และไม่รอบรู้ในแต่ละคุณลักษณะของความสามารถ การประเมินวินิจฉัยทางปัญญาเป็นรูปแบบหนึ่งของการประเมินที่ใช้ในการสำรวจความสัมพันธ์ระหว่างกระบวนการทางจิตและกลวิธีที่ผู้สอบใช้ในการแก้ปัญหาหรือตอบข้อสอบข้อหนึ่งๆ โดยสารสนเทศที่ได้จะเป็นข้อมูลย้อนกลับเกี่ยวกับจุดแข็งและจุดอ่อนของทักษะย่อยหรือคุณลักษณะของผู้สอบรายบุคคล ดังนั้นจึงมีความเหมาะสมกับบริบทของการประเมินการอ่านซึ่งเป็นกระบวนการทางสมองตามทฤษฎีทางปัญญา การประยุกต์ใช้การประเมินวินิจฉัยทางปัญญาในด้านความสามารถทางภาษาทำได้หลายวิธีแต่วิธีที่ให้ความตรงเชิงโครงสร้างสูงคือการออกแบบการประเมินตามระบบการออกแบบทางปัญญา (CDS) ให้ความสำคัญกับ Q-matrix ซึ่งเป็นเมตริกซ์ความสัมพันธ์ระหว่างข้อคำถามและคุณลักษณะของความสามารถเพื่อนำมาสร้างแบบทดสอบวินิจฉัยทางปัญญาที่ใช้ในการประเมินวินิจฉัยทางปัญญา

คำสำคัญ: การประเมินวินิจฉัยทางปัญญา การอ่านเพื่อความเข้าใจ

บทนำ (Introduction)

การประเมินวินิจฉัยทางปัญญา (Cognitive Diagnostic Assessment: CDA) เป็นรูปแบบการประเมินได้รับการออกแบบมาเพื่อวัดโครงสร้างความรู้ของผู้เข้าสอบและทักษะการประมวลผล โครงสร้างความรู้ประกอบด้วยข้อเท็จจริงและข้อมูลที่ผู้สอบรับสาร ทักษะการประมวลผลคือกระบวนการแปลความหมายข้อมูลและกลยุทธ์ที่จำเป็นในการจัดการข้อมูลนี้ออกมาเป็นสารสนเทศเกี่ยวกับผลการสอบทั้งในแง่ของผลการประเมินโดยสรุปในภาพรวม และผลการประเมินซึ่งเป็นข้อวินิจฉัยเกี่ยวกับจุดแข็งและจุดอ่อนของผู้เรียนในสิ่งที่ประเมินได้เป็นรายบุคคล (Leighton & Gierl, 2007) ซึ่งแตกต่างจากการประเมินในรูปแบบเดิมที่ใช้กันโดยทั่วไป ซึ่งให้สารสนเทศเกี่ยวกับผู้สอบได้เฉพาะในภาพรวม อาทิ “ผ่าน” “ไม่ผ่าน” “ดีมาก” “พอใช้” “ปรับปรุง” หรือให้เป็นคะแนนจากการทำแบบทดสอบ เป็นต้น การประเมินวินิจฉัยทางปัญญามีวัตถุประสงค์สำคัญสองประการ คือ เพื่อระบุประเด็นที่นักเรียนยังไม่รอบรู้ และให้ข้อเสนอแนะเกี่ยวกับสาเหตุ หรือเหตุผลที่เป็นไปได้ว่าเพราะเหตุใดนักเรียนถึงยังไม่รอบรู้ในประเด็นนั้นๆ นำไปใช้ประโยชน์ในการพัฒนานักเรียนได้ตรงจุดที่บกพร่อง โดยใช้โมเดลทางสถิติในการวิเคราะห์ข้อมูล

โมเดลวินิจฉัยทางปัญญา (Cognitive Diagnostic Models: CDMs) เป็นวิธีการแบบสหวิทยาการ (Interdisciplinary Approach) โดยมีการเชื่อมโยงระหว่างจิตวิทยาพุทธิปัญญาและการวิเคราะห์ทางสถิติไว้ด้วยกัน CDMs เป็นโมเดลทางสถิติที่สามารถสังเคราะห์สารสนเทศเชิงวินิจฉัยจากข้อมูล นำสารสนเทศที่ได้มาพัฒนาหรือกำหนดข้อสอบที่วัดองค์ความรู้และทักษะที่เฉพาะเจาะจง และใช้โมเดลนี้กับการวิเคราะห์ทางจิตของรูปแบบการตอบสนองต่อข้อสอบของผู้ตอบ เพื่อส่งเสริมการวินิจฉัยที่เฉพาะเจาะจง (Leighton & Gierl, 2007) มีนักการศึกษาจำนวนมากไม่นำการประเมินเพื่อวินิจฉัยทางปัญญามาใช้ประเมินทักษะความสามารถทางวิชาการกันอย่างแพร่หลาย รวมทั้งการประเมินความสามารถทางภาษาอังกฤษ (Sun & Hwang, 2023)

การประเมินความสามารถทางภาษาอังกฤษด้านการอ่านเพื่อความเข้าใจภาษาอังกฤษ ตามทฤษฎีทางปัญญาวินิจฉัยทางปัญญาเป็นการศึกษากระบวนการที่ผู้อ่านใช้คุณลักษณะหรือทักษะย่อย ๆ ร่วมกันในการทำความเข้าใจสิ่งที่อ่าน นักวิชาการพิจารณาความสามารถทางภาษาจากผลการทำแบบทดสอบทางภาษา เพื่อได้ข้อมูลเชิงลึกเกี่ยวกับสถานะความรู้และระดับทักษะ

ความเชี่ยวชาญทางภาษาของผู้สอบ โดยเครื่องมือที่ใช้วัดอาจเป็นแบบทดสอบวินิจฉัยที่เผยแพร่ทั่วไป แบบทดสอบวินิจฉัยที่ครูผู้สอนหรือนักวิจัยสร้างขึ้น หรือเทคนิคการสังเกต เป็นต้น (Kanjanaawasee, 2009)

ภูมิหลัง

การประเมินวินิจฉัยทางปัญญา (Cognitive Diagnostic Assessment: CDA) มีจุดเริ่มต้นจากยุคแห่งการทดสอบวัดผลสัมฤทธิ์ทางการเรียนของผู้เรียนที่เป็นมาตรฐาน ซึ่งผลการสอบมีความสำคัญและมีผลกระทบต่อผู้ที่เกี่ยวข้อง อาทิ การสอบคัดเลือกเข้าเรียน การจบการศึกษา ได้เพิ่มความกดดัน โดยเฉพาะการประเมินผลสัมฤทธิ์ทางการเรียนของผู้เรียนในระดับมหภาค (Large-scale assessment) หลังจากสหรัฐอเมริกาได้ประกาศพระราชบัญญัติทางการศึกษา No Child Left Behind ปี 2001 (NCLB) หมายถึง ไม่มีเด็กคนใดถูกทิ้งไว้เบื้องหลัง ทุกคนจะได้รับการศึกษาอย่างเท่าเทียมกัน มีวัตถุประสงค์เพื่อสร้างระบบการศึกษาที่ครอบคลุม ตอบสนอง และมีความเป็นธรรมมากขึ้น โดยบัตริยาของนักเรียนจะประกอบด้วยข้อมูลผลสัมฤทธิ์ทางการเรียน แยกตามเชื้อชาติ ชาติพันธุ์ เพศ ความสามารถทางภาษาอังกฤษ สถานะผู้อพยพ ความทุพพลภาพ เด็กด้อยโอกาส และฐานะครอบครัว ส่งผลกระทบต่อระบบการวัดประเมินผลการเรียน โดยสาระสำคัญของพระราชบัญญัตินี้ระบุว่า “การทดสอบเพื่อประเมินผลการเรียนของผู้เรียนจะต้องไม่ได้เป็นเพียงแค่การประเมินผลการเรียนรู้ของผู้เรียน แต่ยังเป็นการประเมินผลการจัดการเรียนการสอนด้วย การประเมินผลการเรียนของผู้เรียนควรมีการให้สารสนเทศย้อนกลับในเชิงการวินิจฉัยจุดเด่น และข้อบกพร่องของผู้เรียนเป็นรายบุคคล และรายงานไปยังผู้เรียน ผู้ปกครอง ครู และผู้ที่มีหน้าที่เกี่ยวข้อง (Nolen & Duignan, 2024)

ในช่วงปี 1983 – 1995 นับได้ว่าเป็นช่วงเวลาแห่งการรณรงค์ของการวินิจฉัยทางปัญญา เมื่อ Embretson ได้ตีพิมพ์ผลงานในวารสาร Psychological Bulletin ในปีค.ศ. 1983 แสดงให้เห็นถึงการรวมกันของจิตวิทยาทางปัญญาและความตรงเชิงโครงสร้าง หมายถึง ความสัมพันธ์ระหว่างโครงสร้างหรือคุณลักษณะที่เป็นตัวแทนความสามารถที่มุ่งศึกษากับการตอบสนองทางปัญญาออกมาในรูปแบบการตอบข้อสอบ โดยอาศัยความรู้เดิม กลยุทธ์ และกระบวนการที่ผู้สอบใช้ในการทำข้อสอบข้อหนึ่งๆ ซึ่งเป็นความสามารถที่แสดงออกมา (Performance) ผลงานของ Embretson เป็นการประยุกต์การพัฒนาจากจิตวิทยาทางปัญญาไปสู่ทฤษฎีการวัด และในหนังสือ Educational Measurement ที่ตีพิมพ์ในปีเดียวกันนี้ Snow and Lohman (1989) ได้แสดงให้เห็นอย่างชัดเจนในการนำแนวคิดทฤษฎีทางปัญญาเข้ามาใช้ในการวัดทางการศึกษาในหัวข้อนัยยะของจิตวิทยาพุทธิปัญญาสำหรับการวัดทางการศึกษา โดยชี้ให้เห็นว่าแนวคิดทฤษฎี และวิธีการทางปัญญาเป็นประโยชน์สำหรับการวัดทางการศึกษาในแง่มุมต่างๆ ดังนี้ สามารถให้ข้อมูลการวิเคราะห์ข้อสอบที่อธิบายถึงโครงสร้างของความรู้ ผลการประเมินให้สารสนเทศเกี่ยวกับจุดแข็งและจุดอ่อนของผู้สอบ สร้างจุดมุ่งหมายการประเมินที่ชัดเจน และสามารถเพิ่มความแกร่งทางทฤษฎีที่เกี่ยวข้อง (Leighton & Gierl, 2007)

มโนทัศน์เบื้องต้นของการประเมินวินิจฉัยทางปัญญา

การประเมินวินิจฉัยทางปัญญา (CDA) หมายถึง วิธีหรือกระบวนการที่นำมาใช้เพื่อวิเคราะห์จุดแข็งและจุดอ่อนของผู้เรียน ตลอดจนปัญหาอุปสรรค และสาเหตุที่เกี่ยวข้องกับกระบวนการเรียนรู้ของผู้เรียน ที่เกิดขึ้นระหว่างการเรียนการสอน เพื่อนำสารสนเทศที่ได้มาใช้ในการออกแบบการจัดการเรียนรู้ ปรับปรุง แก้ไขอย่างตรงประเด็นด้วยวิธีการทดสอบ ซึ่งเครื่องมือที่ใช้วัดอาจเป็นแบบทดสอบวินิจฉัยที่เผยแพร่ทั่วไป (Published Diagnostic Tests) แบบทดสอบวินิจฉัยที่ครูผู้สอนสร้างเอง (Teacher-made Diagnostic Tests) หรือเทคนิคการสังเกต (Observation Techniques) เป็นต้น (Kanjanaawasee, 2009)

การประเมินวินิจฉัยทางปัญญา (Cognitive Diagnostic Assessment: CDA) เป็นรูปแบบการประเมินความรู้และคุณลักษณะหรือทักษะย่อยของความสามารถที่เฉพาะเจาะจงโดยใช้โมเดลวินิจฉัยทางปัญญา (Cognitive diagnostic Models:

CDMs) กล่าวคือ การประเมินวินิจฉัยทางปัญญาเป็นการศึกษาเกี่ยวกับโครงสร้างคุณลักษณะทางปัญญาของผู้เรียนที่ใช้โมเดลวินิจฉัยทางปัญญามาใช้ในการประเมินเพื่อให้ได้สารสนเทศในการวินิจฉัยความรอบรู้ของผู้เรียนในแต่ละคุณลักษณะหรือทักษะย่อยของความสามารถที่มุ่งศึกษา เริ่มต้นจากการสร้างข้อสอบที่วัดความรู้และคุณลักษณะหรือทักษะย่อยที่เฉพาะเจาะจงและใช้โมเดลนี้โดยตรงกับการวิเคราะห์รูปแบบการตอบสนองต่อข้อสอบของผู้เรียนตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เพื่อส่งเสริมการวินิจฉัยที่เฉพาะเจาะจง โดยใช้โมเดลทางสถิติในการวิเคราะห์ที่สามารถสกัดสารสนเทศเชิงวินิจฉัยจากข้อมูลได้ เพื่อให้ได้มาซึ่งข้อมูลดังกล่าว โดยทั่วไปแล้ว การประเมินจะได้รับการออกแบบมาเพื่อวัดชุดของทักษะหลายทักษะที่แยกย่อยออกไป เรียกว่า คุณลักษณะ (Attributes) ซึ่งเป็นตัวแปรแฝงที่สามารถเป็นได้ 2 ค่า (Binary latent variables) ที่มีผลลัพธ์ (outcome) เท่ากับ “1” สำหรับความเป็นผู้รอบรู้ (Mastery) และผลลัพธ์เป็น “0” สำหรับความไม่รอบรู้ (Non-mastery) สามารถศึกษารายละเอียดและตัวอย่างได้ในหัวข้อโมเดลวินิจฉัยทางปัญญา เป้าหมายหลักของการวิเคราะห์ด้วยโมเดลวินิจฉัยทางปัญญา (CDMs) คือการคาดคะเนความน่าจะเป็นรูปแบบความรอบรู้ของผู้เรียนจากการตอบข้อสอบในแบบทดสอบที่ถูกสร้างเพื่อการวินิจฉัยทางปัญญา (Ma & Torre, 2020) ดังนั้นการประเมินวินิจฉัยทางปัญญาจึงจำเป็นที่จะต้องมียละเอียดที่ชัดเจนเพื่ออธิบายกระบวนการที่ผู้เรียนดึงกระบวนการทางปัญญาออกมาใช้ในการตอบข้อสอบ

ระบบการออกแบบทางปัญญา แบบทดสอบวินิจฉัย และโมเดลวินิจฉัยทางปัญญา

การประเมินวินิจฉัยทางปัญญาในยุคปัจจุบันดำเนินไปในรูปของการนำโครงสร้างทางพุทธิปัญญาเข้ามาใช้ในการออกแบบข้อสอบ การให้คะแนน การพัฒนาโมเดล การรายงานผลการวินิจฉัย โดยใช้ทักษะทางปัญญาเป็นฐานเพื่อนำมาใช้ในการประเมินวินิจฉัยทางปัญญาที่ในปัจจุบันเรียกว่าโมเดลวินิจฉัยทางปัญญา (Cognitive Diagnostic Models: CDMs) (Huff & Goodman, 2007) จากการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง พบว่า ระบบการออกแบบทางปัญญาของ Embretson ได้รับความสนใจจากนักวิชาการจำนวนมาก โดยหลักการสำคัญของระบบการออกแบบทางปัญญาดังกล่าวเป็นการเชื่อมโยงระหว่างทฤษฎีพื้นฐานกับโครงสร้าง (Construct) ของการประเมินวินิจฉัยเพื่อที่จะแปลผลจากการประเมินวินิจฉัยได้มากที่สุด โดยประเด็นที่เกี่ยวข้องกับการเสนอแนวคิดเพื่อการประยุกต์ใช้การประเมินวินิจฉัยทางปัญญามี 3 ประเด็น ได้แก่ ระบบการออกแบบทางปัญญา แบบทดสอบวินิจฉัยทางปัญญา โมเดลวินิจฉัยทางปัญญา โดยผู้เขียนจะนำเสนอรายละเอียดในแต่ละประเด็น ดังนี้

1. ระบบการออกแบบทางปัญญา (Cognitive Design System: CDS) ถูกนำเสนอโดย Susan Embretson เมื่อปี ค.ศ. 1994 ทั้งนี้เพื่อตอบข้อวิพากษ์ที่ว่า การสร้างข้อสอบวินิจฉัยที่ไม่อิงทฤษฎีทำให้มีจุดอ่อนในความตรงเชิงโครงสร้างของการแปลผลคะแนน ระบบการออกแบบทางปัญญาจึงถูกออกแบบโดยใช้ทฤษฎีทางปัญญา (Cognitive theory) ในการพัฒนาข้อสอบเพื่อแก้ไขจุดอ่อนดังกล่าวให้การแปลผลคะแนนความสามารถมีความตรงมากขึ้น ระบบการออกแบบทางปัญญามี 10 ขั้นตอน (Embretson, 1994) ดังนี้ 1) กำหนดจุดมุ่งหมายของการวัดให้ชัดเจน (Specify general goals of measurement) 2) ระบุนคุณลักษณะที่เกี่ยวข้องกับสิ่งที่ต้องการวินิจฉัย (Identify design features in task domain) 3) พัฒนาโมเดลทางปัญญา (Develop a cognitive model) 4) ประเมินโมเดลทางปัญญาที่ใช้สำหรับศักยภาพการวัดทางจิตวิทยา (Evaluate cognitive model for psychometric potential) 5) ระบุการแจกแจงความน่าจะเป็นของข้อสอบเกี่ยวกับความซับซ้อนทางปัญญา (Specify item distributions on cognitive complexity) 6) สร้างข้อสอบตามลักษณะเฉพาะของข้อสอบ (Generate items to fit specifications) ลักษณะเฉพาะของข้อสอบ หรือ Test Specifications ประกอบด้วยวัตถุประสงค์ของการทดสอบ ขอบเขตของแบบทดสอบ โครงสร้างและรูปแบบของแบบทดสอบ รวมทั้งขอบเขตเนื้อหาของแบบทดสอบที่สอดคล้องกับคุณภาพผู้เรียน สาระและมาตรฐาน การเรียนรู้ตัวชี้วัด และสาระการเรียนรู้แกนกลางของหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 และลักษณะเฉพาะของข้อสอบ (Item Specification) 7) ประเมินคุณสมบัติของแบบทดสอบฉบับแก้ไขสำหรับการวัดทางจิตวิทยาทางปัญญา (Evaluate cognitive and psychometric properties for revised test) 8) ปรับพารามิเตอร์ข้อสอบและ

การแจกแจงความน่าจะเป็นของความสามารถ (Calibrate final item parameters and ability distributions) 9) จัดทำคลังข้อสอบตามแบบทดสอบจากการรวมกันของพารามิเตอร์ที่มีลักษณะซับซ้อนทางปัญญาสร้างขึ้นเพื่อเป็นตัวแทน (Bank items by cognitive complexity parameters assemble test forms to Represent Specifications) 10) การตรวจสอบความตรง (Validation nomothetic span) โดยเป็นการตรวจสอบความตรงตามสภาพ (Concurrent validity) หมายถึง ในการระบุถึงโครงสร้างอื่นที่เกี่ยวข้องในเชิงทฤษฎีกับสิ่งที่มุ่งวัด (Nomothetic span) ข้อดีคือเป็นวิธีที่ให้ความเที่ยงตรงเชิงโครงสร้างของคุณลักษณะที่ทำการวัดซึ่งสามารถทำความเข้าใจได้ดีกว่าและสามารถแปลผลคะแนนในเชิงการวินิจฉัยได้ง่ายกว่าวิธีการอื่น

ตัวอย่างการสร้าง Q-matrix ซึ่งเป็นเมตริกซ์ที่แสดงความสัมพันธ์ระหว่างข้อคำถาม (items) กับคุณลักษณะ (Attribute) ที่ เป็นความสามารถที่ใช้ในการหาคำตอบของข้อคำถามนั้น ๆ ในภาพที่ 1 แสดงตัวอย่าง Q-matrix สำหรับสถานการณ์จำลองที่มีคุณลักษณะ 6 รายการสำหรับข้อสอบ 1 ข้อของแบบทดสอบการอ่านเพื่อความเข้าใจ ประกอบด้วย 6 คุณลักษณะ ผู้เขียนกำหนดสัญลักษณ์ A1, A2, A3, A4, A5 และ A6 อักษรตัว A มาจากคำว่า “Attribute” มีความหมายว่า “คุณลักษณะ”

Figure 1

An example of a Q-matrix for a simulated scenario with six attributes for a single reading comprehension test item

แสดงตัวอย่าง Q-matrix สำหรับสถานการณ์จำลองที่มีคุณลักษณะ 6 รายการสำหรับข้อสอบ 1 ข้อ ของแบบทดสอบการอ่านเพื่อความเข้าใจ

Items	Attributes					
	A1	A2	A3	A4	A5	A6
According to the passage, which is refer to the word “items”?	1	1	0	0	0	0
1) Clothes						
2) Cookbooks						
3) Course meals						
4) Comfortable shoes						

อธิบายได้ว่า เมื่อกำหนดให้มีค่าคะแนนเท่ากับ “1” เมื่อผู้สอบต้องใช้คุณลักษณะนั้นในการหาคำตอบ และกำหนดให้มีค่าเท่ากับ “0” เมื่อผู้สอบไม่ใช้คุณลักษณะนั้นในการหาคำตอบ หมายความว่า การที่ผู้สอบจะตอบข้อสอบถูก ผู้สอบต้องมีคุณลักษณะ A1 และ A2 ผู้เขียนนำเสนอตัวอย่างการสร้าง Q-matrix สำหรับข้อสอบที่มีจำนวนมาก ดังภาพที่ 2

Figure 2

An example of a Q-matrix for a simulated scenario with three attributes for 15 reading comprehension test items

แสดงตัวอย่าง Q-matrix สำหรับสถานการณ์จำลองที่มีคุณลักษณะ 3 รายการสำหรับข้อสอบ 15 ข้อ ของแบบทดสอบการอ่านเพื่อความเข้าใจ

Item No.	Set No.	Item No. Within Set	Attributes/Skills		
			Attribute 1 (Understanding Details)	Attribute 2 (Connecting Ideas)	Attribute 3 (Inferencing)
1	1	1	1	0	0
2	1	2	0	1	0
3	1	3	0	1	0
4	1	4	1	0	0
5	1	5	0	1	1
6	2	1	1	1	0
7	2	2	0	1	0
8	2	3	0	1	0
9	2	4	1	1	0
10	2	5	0	0	1
11	3	1	1	0	0
12	3	2	0	1	0
13	3	3	0	1	0
14	3	4	1	1	0
15	3	5	0	0	1

ที่มา: Lee and Sawaki (2009)

จาก Figure 2 แสดงตัวอย่าง Q-matrix สำหรับสถานการณ์จำลองที่มีคุณลักษณะ 3 รายการสำหรับข้อสอบ 15 ข้อ ของแบบทดสอบการอ่านเพื่อความเข้าใจ ทั้ง 3 คอลัมน์ทางด้านขวาแสดงถึงคุณลักษณะเป้าหมายในการทดสอบ ในขณะที่แถวแสดงถึงข้อสอบ เช่น ผู้สอบจะตอบข้อสอบข้อ 1 ผู้สอบต้องมีความรอบรู้คุณลักษณะที่ 1 ความเข้าใจรายละเอียดของเรื่องที่อ่าน (Understanding Detail) การที่ผู้สอบจะตอบข้อสอบข้อที่ 2 ได้ถูก ผู้สอบต้องมีคุณลักษณะเชื่อมโยงความคิดหลายๆ ความคิดเข้าด้วยกัน (Connecting Ideas) หรือผู้สอบจะตอบข้อสอบข้อ 14 ได้ถูก ผู้สอบต้องมีความรอบรู้ใน 2 คุณลักษณะ คือ คุณลักษณะที่ 1 และคุณลักษณะที่ 2 หมายความว่า ผู้สอบต้องมีความเข้าใจรายละเอียดของเรื่องที่อ่าน และสามารถเชื่อมโยงความคิดหลายๆ ความคิดเข้าด้วยกันได้ เป็นต้น

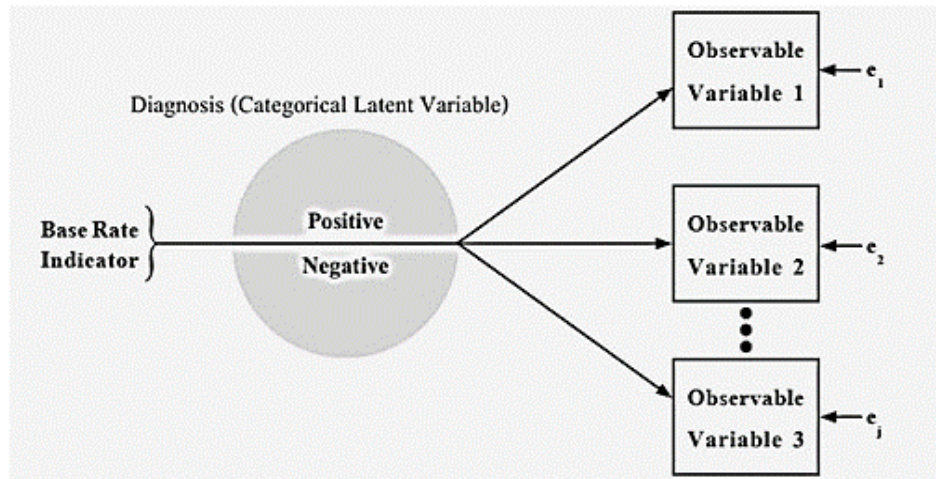
2. แบบทดสอบวินิจฉัยทางปัญญา (Cognitive Diagnostic Test) สิ่งสำคัญประการหนึ่งของการประเมินวินิจฉัยทางปัญญา คือ การพัฒนาแบบทดสอบวินิจฉัย อีกทั้งยังเป็นขั้นตอนหนึ่งในการออกแบบการวินิจฉัยทางปัญญาของ Embretson แบบทดสอบวินิจฉัยทางปัญญา (Cognitive Diagnostic Test) หมายถึง แบบทดสอบที่สร้างขึ้นตามโครงสร้างความรู้และทักษะเฉพาะด้าน ด้วยวิธีการสร้างแบบทดสอบวินิจฉัยที่เป็นระบบ เริ่มต้นจากการกำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ วิเคราะห์เนื้อหาวิชาอย่างละเอียด แล้วแบ่งเป็นเนื้อหาย่อย เขียนวัตถุประสงค์การเรียนรู้ จากนั้นสร้างข้อสอบให้สอดคล้องกับจุดประสงค์การเรียนรู้แล้วนำไปทดลองใช้ ปรับปรุงและตรวจคุณภาพของแบบทดสอบ และจัดทำคู่มือ เพื่อวัดคุณลักษณะทางปัญญาในด้านนั้นๆ อย่างเฉพาะเจาะจงโดยให้สารสนเทศเป็นรายบุคคลเกี่ยวกับผลการเรียนรู้ ผลการประเมินการจัดกิจกรรม การเรียนรู้ จุดแข็งและจุดอ่อนทางปัญญาของนักเรียนจากการคิดและการเรียนรู้ โดยการอนุมานจากคะแนนสอบที่นักเรียนทำได้และให้ข้อมูลย้อนกลับไปในลักษณะของโปรไฟล์ความรู้ในแต่ละคุณลักษณะของนักเรียน รูปแบบของแบบทดสอบที่ใช้ในการประเมินผลความเข้าใจในการอ่านภาษาอังกฤษที่นิยมใช้มี 2 แบบ (Finocchiaro & Sako, 1983) ได้แก่ 1) แบบทดสอบอัตนัย (Subjective test) หมายถึง แบบทดสอบความเรียงที่ให้ผู้เรียนตอบคำถามจากเรื่องที่อ่าน โดยเขียนคำตอบเป็นประโยคหรือข้อความยาวๆ และ

2) แบบทดสอบแบบปรนัย (Objective test) หมายถึง แบบทดสอบแบบเลือกตอบแบบถูกผิด แบบจับคู่ และแบบเติมคำ เป็นต้น ซึ่งในทางปฏิบัติมักนิยมใช้แบบทดสอบปรนัยชนิดเลือกตอบ (Multiple Choice) ในการประเมินผลความเข้าใจในการอ่านภาษาอังกฤษเนื่องจากการตรวจให้คะแนนมีความสะดวก รวดเร็ว เทียบตรงกล่าวคือบุคคลใดบุคคลหนึ่งสามารถเป็นผู้ตรวจได้โดยไม่มีข้อจำกัด รวมทั้งเป็นแบบทดสอบที่กำหนดเนื้อหาในการวัดผลประเมินผลได้อย่างครอบคลุม และสามารถวัดผลประเมินผลกับผู้สอบทุกระดับ ตัวอย่างแบบทดสอบมาตรฐาน 4 แบบ EF SET, CEFR, TOEF และ IELTS ที่มีการประเมินวินิจฉัยทางปัญญาความสามารถด้านการอ่านเพื่อความเข้าใจเป็นส่วนหนึ่งของแบบทดสอบทั้งฉบับได้แก่ Part Reading แบบทดสอบ EF SET เป็นแบบทดสอบความสามารถทางภาษาที่ไม่เสียค่าใช้จ่าย ประเมินครอบคลุมทั้ง 4 ทักษะ ได้แก่ การฟัง การพูด การอ่าน และการเขียน โดยคะแนนของ EF SET จะถูกใช้เพื่อเป็นฐานข้อมูลให้กับหน่วยงาน EF English Proficiency Index (EF EPI) รายงานการจัดอันดับความสามารถทางด้านภาษาอังกฤษรายปีของประเทศ อีกแบบที่รู้จักกันแพร่หลายได้แก่แบบทดสอบ CEFR เป็นแบบทดสอบที่วัดความสามารถทางภาษาตามกรอบมาตรฐานสากลเพื่ออธิบายระดับของความสามารถทางภาษาอังกฤษ 4 ทักษะ คือ การฟัง การพูด การอ่าน และการเขียน สำหรับแบบทดสอบ TOEFL เป็นการทดสอบโดยใช้คอมพิวเตอร์เป็นหลัก รวมทั้งการสอบพูด ส่วนการสอบการพูดของแบบทดสอบ IELTS เป็นการสอบแบบตัวต่อตัวกับบุคคลจริง แบบทดสอบ IELTS มีทั้ง 2 รูปแบบคือกระดาษและคอมพิวเตอร์ โดยแบบทดสอบ TOEFL และ IELTS เน้นนำผลสอบไปใช้ในการเข้าศึกษาต่อในระดับมหาวิทยาลัยของต่างประเทศ จากที่กล่าวมาเห็นได้ว่าการประเมินความสามารถทางปัญญาด้านภาษาอังกฤษเป็นการประเมินโดยใช้แบบทดสอบและเริ่มมีการนำแบบทดสอบวินิจฉัยมาใช้แพร่หลายมากขึ้น

3. โมเดลวินิจฉัยทางปัญญา (Cognitive Diagnostic Models: CDMs) เป็นโมเดลการวัดทางจิตวิทยาที่ใช้ประเมินจุดแข็งและจุดอ่อนของนักเรียน อาทิ การศึกษาของ Torre and Ma (2020) ในงานวิจัยบางเล่มเรียกว่าโมเดลวินิจฉัยจัดประเภท (Diagnostic classification models: DCMs) เช่น การศึกษาของ Rupp et al. (2010) เป็นวิธีการในการคาดคะเนหรือพยากรณ์ความน่าจะเป็นของความรู้ความเข้าใจโดยจัดผู้สอบแยกเข้าชั้นแฝงตามความรู้ทักษะที่มี ผลลัพธ์ของการวิเคราะห์ DCMs สำหรับแต่ละบุคคล คือโปรไฟล์คุณลักษณะความน่าจะเป็นที่บ่งชี้ว่าผู้ตอบแบบสอบถามแต่ละคนมีคุณลักษณะหนึ่งอย่างหรือมากกว่านั้นตรงตามเกณฑ์สำหรับการวินิจฉัยหรือไม่ ภาพ 3 แสดงให้เห็นโครงสร้างโมเดลวินิจฉัยจัดประเภท โดยทางซ้ายมือแสดงถึงตัวแปรแฝง (Diagnosis Latent Variable) มีสัญลักษณ์เป็นวงกลมโดยมีเส้นแบ่งกลางวงกลมหมายถึงการจัดตัวแปรแฝงเป็น 2 ประเภทโดยมีตัวบ่งชี้พื้นฐาน (Base rate indicator) เป็นเกณฑ์การแบ่งการวินิจฉัยตำแหน่งเหนือเส้นตัวบ่งชี้พื้นฐานคือ “positive” หรือกลุ่มที่อยู่สูงกว่าเกณฑ์การแบ่ง ตำแหน่งใต้เส้นตัวบ่งชี้พื้นฐานคือ “negative” หรือกลุ่มที่อยู่ตำแหน่งต่ำกว่าเกณฑ์การแบ่งสี่เหลี่ยมขวามือแสดงถึงตัวแปรสังเกตได้ซึ่งเป็นตัวบ่งชี้ของตัวแปรแฝง โดยมี e เป็นส่วนของความคลาดเคลื่อน (Error term) คล้ายโมเดลการวิเคราะห์องค์ประกอบ (Factor analysis)

Figure 3*Discriminant Diagnostic Model Structure*

โครงสร้างโมเดลวินิจฉัยจัดประเภท



ที่มา: Templin (2011)

บทความในส่วนต่อจากนี้ ผู้เขียนใช้ชื่อโมเดลวินิจฉัยทางปัญญา หรือ CDMs เนื่องจากเป็นชื่อที่สอดคล้องกับการประเมินวินิจฉัยทางปัญญา ประเภทของโมเดลวินิจฉัยทางปัญญาสามารถจำแนกออกเป็น 2 ประเภทหลัก ได้แก่ โมเดลชั้นแฝง (Latent Class Model) และโมเดลคุณลักษณะแฝง (Latent Trait Model) โดยโมเดลชั้นแฝงเป็นโมเดลที่พัฒนาขึ้นเพื่อจำแนกผู้สอบตามทักษะ (Skill) เป็นความรอบรู้ (Mastery) หรือความไม่รอบรู้ (Non-mastery) โมเดลนี้จึงให้รูปแบบของความรอบรู้หรือความน่าจะเป็นของความรอบรู้เกี่ยวกับทักษะของผู้สอบ สำหรับโมเดลคุณลักษณะแฝงเป็นการขยายแนวคิดหรือสรุปอ้างอิงจากโมเดลการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional IRT Model) เกี่ยวกับการจัดตำแหน่งการประมาณค่าความสามารถของผู้สอบในแต่ละทักษะ (Lee & Sawaki, 2009) โมเดลชั้นแฝง (Latent Class Model) จำแนกผู้ตอบแบบทดสอบเป็นชั้นแฝง (Latent Classes) ขึ้นอยู่กับการตอบสนองข้อสอบของผู้ตอบโดยชั้นแฝงนั้น อีกทั้งเป็นชั้นแฝงที่มีจำนวนจำกัด (Restricted Latent Classes) ขึ้นอยู่กับจำนวนของคุณลักษณะ (Attributes) ที่ใช้ในการตอบข้อสอบของแบบทดสอบ หากแบบทดสอบนั้นมีคุณลักษณะที่ต้องใช้จำนวน K คุณลักษณะ ผู้ตอบจะถูกจำแนกเป็น 2^K ชั้นแฝง โดยเลข 2 หมายถึงจำนวนของผลลัพธ์ที่เป็นไปได้สำหรับแต่ละคุณลักษณะ ซึ่งหมายถึงการรอบรู้ (Mastery) หรือไม่รอบรู้ (Non-mastery) ในคุณลักษณะนั้นๆ การคำนวณชั้นแฝงจะทำให้ทราบจำนวนโปรไฟล์แบบแผนการตอบที่เป็นไปได้ทั้งหมด (ตัวอย่างการคำนวณชั้นแฝง และรายละเอียดเกี่ยวกับรูปแบบความรอบรู้ผู้เขียนนำเสนอเพิ่มเติมในหัวข้อการประเมินวินิจฉัยทางปัญญาในการประเมินความสามารถด้านการอ่านเพื่อความเข้าใจ)

ตัวอย่างโมเดลวินิจฉัยทางพุทธิปัญญาที่เป็นโมเดลชั้นแฝง อาทิ โมเดล RSM, AHM, DINA, FM, G-DINA เป็นต้น โมเดลชั้นแฝงที่ยกตัวอย่างมีข้อดีและข้อจำกัด ดังนี้ โมเดล RSM (Rule space model) (Tatsuoka, 1983) มีข้อดีคือสามารถจำแนกแบบแผนการตอบข้อสอบไปสู่แบบแผนความรอบรู้ มีข้อจำกัดคือผลการวินิจฉัยอาจไม่สมบูรณ์ เนื่องจากวินิจฉัยหลังจากสร้างแบบสอบเรียบร้อยแล้วทำให้การจำแนกคุณลักษณะที่ต้องใช้ในการตอบข้อสอบแต่ละข้อ ไม่ครอบคลุมคุณลักษณะที่ผู้สอบจะต้องมีสำหรับโมเดล AHM (Attribute Hierarchy Method: AHM) (Leighton et al., 2004) มีข้อดีคือจำแนกคุณลักษณะก่อนการพัฒนาแบบสอบช่วยให้สามารถวัดความรู้ได้ครบทุกคุณลักษณะ ส่วนข้อจำกัดคือการจำแนกแบบแผนการตอบข้อสอบที่สังเกต

ได้อาจไม่ชัดเจนและไม่สามารถวิเคราะห์สาเหตุของมโนทัศน์ที่คลาดเคลื่อนได้ โมเดล FM (Fusion Model) มีข้อดีคือเพื่อสามารถแก้ไขข้อจำกัดของ DINA Model และ NIDA Model ที่บางพารามิเตอร์ในโมเดลไม่สามารถประมาณค่าด้วยวิธีการทางสถิติ เนื่องจากสมมติฐานที่ว่าโอกาสที่ผู้สอบจะตอบข้อสอบได้ถูกต้องนั้น ผู้สอบจะต้องมีความรอบรู้ในทุกคุณลักษณะ โดยรวมการวินิจฉัยทักษะ (Skill Diagnosis) และการประมาณค่าพารามิเตอร์ในโมเดลเข้าไว้ด้วยกัน แต่มีข้อจำกัดคือต้องใช้ Software ที่มีความเฉพาะในการประมาณค่าของ พารามิเตอร์ อีกโมเดลที่มีการนำมาใช้อย่างแพร่หลายได้แก่โมเดล DINA (Deterministic-Input, Noisy-and-Gate Model) มีข้อดีคือโมเดลสอดคล้องกับข้อมูลเป็นโมเดลที่ประหยัด สามารถแปลผลจาก 2 พารามิเตอร์ในแต่ละข้อด้วยโปรแกรมการวิเคราะห์ทางสถิติที่มีอยู่ทั่วไป มีข้อจำกัดคือเป็นโมเดลแบบไม่ชัดเจน จึงมีข้อดกของเบื้องต้น ว่าผู้สอบต้องมีทุกทักษะ จะขาดทักษะใดไปไม่ได้ ทำให้การจัดประเภททักษะ หรือความสามารถของผู้สอบ มีความเฉพาะกับแบบแผนความสามารถของผู้สอบ (Ueabunpradit, 2020) ต่อมา de la Torre, J. (2011) พัฒนาโมเดล G-DINA เพื่อแก้ไขข้อจำกัดของโมเดล DINA โดย G-DINA มีลักษณะชดเชยความสามารถได้ (Compensatory Model) การประมาณค่าพารามิเตอร์ทำได้ง่าย ใช้เวลาน้อย โมเดลมีความเหมาะสมสอดคล้องกับข้อมูล (Good Model Fit) ดีกว่าโมเดลอื่นๆ สามารถประมาณค่าพารามิเตอร์ได้อย่างถูกต้อง และมีข้อจำกัดน้อยกว่าสามารถประเมินประสิทธิภาพของโมเดลโดยพิจารณาจากค่าสถิติความสอดคล้องเชิงสัมพัทธ์ โมเดล DINA และโมเดล G-DINA มีความคล้ายกันคือสามารถวิเคราะห์หาค่าสถิติแบบทดสอบวินิจฉัยเชิงพหุปัญญาในโปรแกรม R ได้เป็นอย่างดี และเป็นโมเดลที่ประหยัด (Parsimonious) สามารถแปลผลได้จาก 2 พารามิเตอร์ในแต่ละข้อ คือ พารามิเตอร์ความสะเพร่า (Slipping) และพารามิเตอร์การเดาข้อสอบถูก (Guessing) ซึ่งจะทำให้การประมาณค่าความสามารถและการจัดกลุ่มผลการตอบของผู้เรียนมีความถูกต้องแม่นยำมากยิ่งขึ้น (Rupp et al., 2010)

การประยุกต์ใช้การประเมินวินิจฉัยทางปัญญาศาสนาความสามารถด้านการอ่านเพื่อความเข้าใจ

การประเมินเพื่อวินิจฉัยทางปัญญาเป็นรูปแบบการประเมินที่สามารถให้สารสนเทศเกี่ยวกับผลการสอบได้ทั้งผลการประเมินโดยสรุปในภาพรวม และผลการประเมินซึ่งเป็นข้อวินิจฉัยเกี่ยวกับจุดแข็งและจุดอ่อนของผู้เรียนเป็นรายบุคคล ซึ่งครูผู้สอนสามารถนำไปออกแบบกิจกรรมการเรียนรู้เพื่อพัฒนาผู้เรียนแตกต่างจากการประเมินในรูปแบบเดิมที่ใช้กันโดยทั่วไป ซึ่งให้สารสนเทศเกี่ยวกับผู้สอบได้เฉพาะในภาพรวมหรือในรูปแบบคะแนน ดังนั้นจึงมีนักวิชาการให้ความสนใจศึกษาเกี่ยวกับการประเมินความสามารถทางภาษาตามแนวทางของการประเมินวินิจฉัยทางปัญญา ผู้เขียนนำเสนอประเด็นสำคัญที่เกี่ยวข้อง 3 ประเด็น ได้แก่ ความสามารถด้านการอ่านเพื่อความเข้าใจ รูปแบบการประยุกต์ใช้การประเมินวินิจฉัยทางปัญญากับการประเมินความสามารถด้านการอ่านเพื่อความเข้าใจของนักวิชาการในช่วงปี ค.ศ. 2006 ถึง 2023 และแนวโน้มและทิศทางของการประเมินวินิจฉัยทางปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจ โดยมีรายละเอียด ดังนี้

ความสามารถด้านการอ่านเพื่อความเข้าใจ (Reading Comprehension) เป็นทักษะทางปัญญาซึ่งมีความเกี่ยวข้องกับการเรียนรู้ตามทัศนะของนักจิตวิทยากลุ่มคอนสตรัคติวิสต์เชิงปัญญา หรือเรียกว่าทฤษฎีการเรียนรู้ปัญญานิยม (Cognitive Constructivism) ซึ่งเชื่อว่าการอ่านคือกระบวนการทางปัญญาที่บุคคลใช้ในการค้นหาความหมายของสิ่งที่อ่านโดยอาศัยความรู้พื้นฐานหรือโครงสร้างความรู้ที่สะสมไว้หรือมวลประสบการณ์เพื่อความเข้าใจและเกิดการเรียนรู้ โดยที่กระบวนการดังกล่าวประกอบด้วย 1) การถอดรหัสความหมายของภาษา (Decoding) ซึ่งเป็นการวิเคราะห์ตัวอักษร คำ กลุ่มคำ ประโยคของเนื้อหาที่อ่าน และ 2) ความเข้าใจ (Comprehension) ซึ่งเป็นการสร้างความหมายจากการถอดรหัสตัวอักษร คำ กลุ่มคำ ประโยคของเนื้อหาที่อ่านโดยอาศัยพื้นฐานความรู้เดิม (Background Knowledge) สอดคล้องกับทฤษฎีโครงสร้างความรู้ (Schema Theory) เชื่อว่าความรู้ของบุคคลจะถูกรวบรวมเป็นหน่วยๆ ในโครงสร้างที่จัดไว้อย่างเป็นระเบียบเรียกว่า โครงสร้างความรู้ (Schemata) เป็นหน่วยความรู้ที่เกี่ยวกับผู้คน สิ่งของ สถานที่เหตุการณ์ หากความรู้ที่ได้รับมาใหม่มีลักษณะผสมกลมกลืนจะถูกนำมาเพิ่มเข้ากับโครงสร้างเดิมที่มีอยู่ แต่ถ้าไม่อยู่ในลักษณะผสมกลมกลืนจะถูกสร้างเป็นโครงสร้างใหม่เรียงซ้อนกัน

คล้ายบล็อกซึ่งจะมีความซับซ้อนมากขึ้น ดังนั้น ความสามารถในการอ่านเพื่อความเข้าใจ จึงมีความเกี่ยวข้องกับการเรียนรู้ตามทัศนะของนักจิตวิทยากลุ่มคอนสตรัคติวิสต์เชิงปัญญา เนื่องจากการอ่านเป็นกระบวนการทางปัญญาที่บุคคลใช้ในการค้นหาความหมายของสิ่งที่อ่านโดยอาศัยความรู้พื้นฐานหรือโครงสร้างความรู้ที่สะสมไว้หรือมวลประสบการณ์เพื่อความเข้าใจและเกิดการเรียนรู้

รูปแบบการประยุกต์ใช้การประเมินวินิจฉัยทางปัญญากับการประเมินความสามารถด้านการอ่านเพื่อความเข้าใจของนักวิชาการในช่วงปี ค.ศ. 2006 ถึง 2023

กลวิธีตรวจสอบความเข้าใจของการประเมินวินิจฉัยทางปัญญาเป็นวิธีการในการประเมินความสามารถการอ่านของผู้เรียน รวมทั้งเทคนิคหรือวิธีการที่ผู้เรียนใช้ในการทำความเข้าใจหรือช่วยให้เข้าใจเนื้อหาที่อ่านง่ายขึ้น โดยลักษณะต่างๆ ที่ผู้เรียนแสดงออกมาเมื่อเกิดกระบวนการเรียนรู้ ได้แก่ ทักษะ (Skill) หรือคุณลักษณะ (Attribute) ให้ปรากฏ โดยมีวิธีการประเมินความสามารถของผู้เรียน อาทิ การทำแบบทดสอบ เป็นต้น งานวิจัยในช่วงต่อมานักวิชาการด้านภาษาจึงให้ความสำคัญกับการศึกษาเกี่ยวกับการประเมินวินิจฉัยทางปัญญาความสามารถทางภาษากันอย่างแพร่หลายมากขึ้น Table 1 แสดงรูปแบบการประเมินวินิจฉัยทางปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักวิชาการในช่วงปี ค.ศ. 2006 ถึง 2023

Table 1
Application Formats
รูปแบบในการประยุกต์ใช้

นักวิจัย	สร้างแบบทดสอบ วินิจฉัย	การวินิจฉัยทาง พุทธิปัญญา	โมเดลวินิจฉัยทาง พุทธิปัญญา	ชื่อโมเดล
Gao (2006)	✓	✓	✓	TBR
Templin & Henson (2006)	✓	✓	✓	DINA
Jang (2009)	✓	✓	✓	Fusion
Li (2011)	✓	✓	✓	Fusion
Ravand (2012)	✓	✓	✓	DINA
Khamyoo (2014)	✓	✓	✓	AHM
Chen and Chen (2015)	-	✓	✓	G-DINA
Zhan et al. (2015)	✓	✓	✓	DINA
Somrat (2020)	✓	✓	✓	G-DINA
Tonekaboni (2021)	✓	✓	✓	G-DINA
Moradi et al. (2022)	✓	✓	-	Rasch DIF
Shahmirzadi and Marashi (2023)	✓	✓	✓	G-DINA

เห็นได้ชัดว่า แนวทางหนึ่งของการประเมินวินิจฉัยทางปัญญาความสามารถทางภาษา คือการสร้างเครื่องมือซึ่งเป็นแบบทดสอบวินิจฉัย ในการวัดโครงสร้างความรู้ ทักษะเฉพาะ และวัดคุณลักษณะทางพุทธิปัญญาของนักเรียนอย่างเฉพาะเจาะจงเป็นรายบุคคล วิเคราะห์ข้อมูลด้วยโมเดลวินิจฉัยทางปัญญาเพื่อได้สารสนเทศเกี่ยวกับผลการเรียนรู้ จุดแข็งและจุดอ่อนของนักเรียนจากแบบแผนการคิด โดยการอนุมานจากคะแนนสอบที่นักเรียนทำได้และให้ข้อมูลย้อนกลับไปในลักษณะของโปรไฟล์ความรู้ในแต่ละคุณลักษณะของนักเรียน โมเดลวินิจฉัยทางปัญญาแต่ละโมเดลมีข้อดีและข้อจำกัดแตกต่างกัน ดังนั้นการเลือกใช้

โมเดลสำหรับการพัฒนาแบบทดสอบวินิจฉัยจึงขึ้นอยู่กับบริบทของเรื่องที่ต้องการศึกษา จากการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องพบว่า การประเมินวินิจฉัยทางปัญญาความสามารถทางภาษาอังกฤษเป็นการใช้โมเดลวินิจฉัยแบบโมเดลชั้นแฝงซึ่งทำการจำแนกผู้ตอบแบบทดสอบเป็นชั้นแฝง (Latent Classes) ขึ้นอยู่กับการตอบสนองข้อสอบของผู้ตอบโดยชั้นแฝงนั้น อีกทั้งเป็นชั้นแฝงที่มีจำนวนจำกัด (Restricted Latent Classes) ขึ้นอยู่กับจำนวนของคุณลักษณะ (Attributes) ที่ใช้ในการตอบข้อสอบของแบบทดสอบ หากแบบทดสอบนั้นมียคุณลักษณะที่ต้องใช้จำนวน K คุณลักษณะ ผู้ตอบจะถูกจำแนกเป็น 2^K ชั้นแฝง โดยเลข 2 หมายถึงจำนวนของผลลัพธ์ที่เป็นไปได้สำหรับแต่ละคุณลักษณะ ซึ่งหมายถึงการรอบรู้ (Mastery) หรือไม่รอบรู้ (Non-mastery) ในคุณลักษณะนั้นๆ ตัวอย่างการคำนวณชั้นแฝง และรายละเอียดเกี่ยวกับรูปแบบความรอบรู้ ดังต่อไปนี้

ตัวอย่างที่ 1 รูปแบบความรอบรู้ แบบทดสอบชุดหนึ่งต้องใช้คุณลักษณะในการตอบแบบทดสอบจำนวน 3 คุณลักษณะ 2^K แทนค่า ($K=3$) จะได้ $2^3 = 8$ นั่นคือ ผู้ตอบจะถูกจำแนกเป็น 8 ชั้นแฝง กำหนดให้ “1” หมายถึง รอบรู้ และ “0” หมายถึง ไม่รอบรู้ ดังนั้น จำนวนชั้นแฝงของรูปแบบการรอบรู้ $\alpha = [(0,0,0) (0,0,1) (0,1,0) (1,0,0) (0,1,1) (1,0,1) (1,1,0) (1,1,1)]$ แสดงให้เห็นจำนวนชั้นแฝงของรูปแบบการเรียนรู้ที่อาจเกิดขึ้นจากการวินิจฉัยซึ่งผู้สอบจะถูกจัดเข้าชั้นแฝงชุดใดชั้นหนึ่ง ขึ้นอยู่กับความรอบรู้ของผู้สอบว่ามีความสามารถตามคุณลักษณะที่วัดจากแบบทดสอบชุดนั้นๆหรือไม่ (Deenang, 2016) Table 1 แสดงชั้นแฝงที่ผู้สอบจะถูกจำแนกเข้าไปในรูปแบบที่คาดหวัง กรณีแบบทดสอบมีคุณลักษณะที่ต้องใช้ในการตอบแบบทดสอบจำนวน 3 คุณลักษณะ

Table 2

Latent Classes for Test-Takers Based on Three Attributes Required for Responding to the Test

ชั้นแฝงที่ผู้สอบจะถูกจำแนกในกรณีมี 3 คุณลักษณะที่ต้องใช้ในการตอบแบบทดสอบ

ชั้นแฝงที่(รูปแบบ)	คุณลักษณะที่ 1	คุณลักษณะที่ 2	คุณลักษณะที่ 3	ความหมาย
1 (0,0,0)	ไม่รอบรู้	ไม่รอบรู้	ไม่รอบรู้	ผู้สอบไม่รอบรู้ทุกคุณลักษณะ
2 (1,0,0)	รอบรู้	ไม่รอบรู้	ไม่รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะ 1
3 (0,1,0)	ไม่รอบรู้	รอบรู้	ไม่รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะ 2
4 (0,0,1)	ไม่รอบรู้	ไม่รอบรู้	รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะ 3
5 (1,1,0)	รอบรู้	รอบรู้	ไม่รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะ 1, 2
6 (1,0,1)	รอบรู้	ไม่รอบรู้	รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะ 1, 3
7 (0,1,1)	ไม่รอบรู้	รอบรู้	รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะ 2, 3
8 (1,1,1)	รอบรู้	รอบรู้	รอบรู้	ผู้สอบรอบรู้ในทุกคุณลักษณะ

จาก Table 2 อธิบายได้ว่า ชั้นแฝงที่ 1 มีรูปแบบ (0,0,0) หมายถึง ผู้สอบไม่รอบรู้ทุกคุณลักษณะ ชั้นแฝงที่ 2 (1,0,0) หมายถึง ผู้สอบรอบรู้ในคุณลักษณะ 1 เท่านั้น ชั้นแฝงที่ 3 (0,1,0) หมายถึง ผู้สอบรอบรู้ในคุณลักษณะ 2 เท่านั้น ชั้นแฝงที่ 4 (0,0,1) หมายถึง ผู้สอบรอบรู้ในคุณลักษณะ 3 เท่านั้น ชั้นแฝงที่ 5 (1,1,0) หมายถึง ผู้สอบรอบรู้ในคุณลักษณะที่ 1 และ 2 เท่านั้น

Figure 4 แสดงตัวอย่างผลการวินิจฉัยความสามารถด้านการอ่านเพื่อความเข้าใจ (Chen & Chen, 2015) ประกอบด้วย 5 คุณลักษณะ อธิบายได้ดังนี้ คุณลักษณะ Identifying Explicit Information มีความน่าจะเป็นที่ผู้สอบจะมีความรอบรู้คิดเป็นร้อยละ 67.80 คุณลักษณะ Generalising Main Idea มีความน่าจะเป็นที่ผู้สอบจะมีความรอบรู้คิดเป็นร้อยละ 64.48 คุณลักษณะ Interpretation and Explanation มีความน่าจะเป็นที่ผู้สอบจะมีความรอบรู้คิดเป็นร้อยละ 54.65

คุณลักษณะ Getting Inference มีความน่าจะเป็นที่ผู้สอบจะมีความรอบรู้คิดเป็นร้อยละ 53.00 คุณลักษณะ Evaluation and Comment มีความน่าจะเป็นที่ผู้สอบจะมีความรอบรู้คิดเป็นร้อยละ 52.62

Figure 4

An example of diagnostic results for reading comprehension ability

แสดงตัวอย่างผลการวินิจฉัยความสามารถด้านการอ่านเพื่อความเข้าใจ

Cognitive attribute	Mastery probability
Identifying Explicit Information	.6780
Generalising Main Idea	.6448
Interpretation and Explanation	.5465
Getting Inference	.5300
Evaluation and Comment	.5262

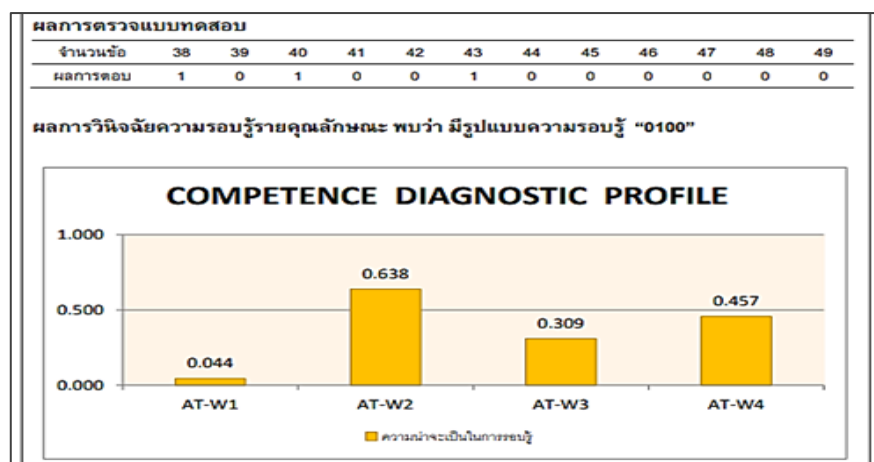
ที่มา: Chen and Chen (2015)

ตัวอย่างการให้ข้อมูลย้อนกลับในรูปแบบ Profile จาก Figure 4 แสดงตัวอย่างผลการวินิจฉัยโดยมีการให้ข้อมูลย้อนกลับของคุณลักษณะความสามารถการสื่อสารภาษาอังกฤษ 4 ทักษะ คือ การฟัง การพูด การอ่าน และการเขียน

Figure 5

An example of feedback provision

แสดงตัวอย่างการให้ข้อมูลย้อนกลับ



ที่มา: Somrat (2023)

จาก Figure 5 พบว่า ผู้สอบได้คะแนนสอบ 3 คะแนน มีรูปแบบความรอบรู้เท่ากับ "0100" มีความน่าจะเป็นในการรอบรู้คุณลักษณะที่ 1 คิดเป็นร้อยละ 0.44 หมายถึง ไม่รอบรู้ มีความน่าจะเป็นในการรอบรู้คุณลักษณะที่ 2 คิดเป็นร้อยละ

63.80 หมายถึง รอบรู้ มีความน่าจะเป็นในการรอบรู้คุณลักษณะที่ 3 คิดเป็นร้อยละ 30.90 หมายถึง ไม่รอบรู้ (เนื่องจากไม่ถึงร้อยละ 50) และมีความน่าจะเป็นในการรอบรู้คุณลักษณะที่ 4 คิดเป็นร้อยละ 45.70 หมายถึง ไม่รอบรู้

แนวโน้มและทิศทางของการประเมินวินิจัยทางปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจ

ในปัจจุบันมีการตื่นตัวในการนำการวินิจัยทางพุทธิปัญญามาใช้ในการประเมินทางการศึกษามากขึ้น แต่ก็นับได้ว่ายังคงอยู่ในช่วงเริ่มต้น ทั้งที่ต้นกำเนิดของการประเมินในรูปแบบนี้มีมาเป็นเวลานาน ด้วยเหตุดังกล่าว นักวิชาการจึงทำการสำรวจแนวโน้มทิศทางของการประเมินวินิจัยทางปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจใน 2 วิธีได้แก่ การทบทวนงานวิจัยที่เกี่ยวข้อง และการใช้เทคนิคที่ทันสมัย มีรายละเอียด ดังนี้

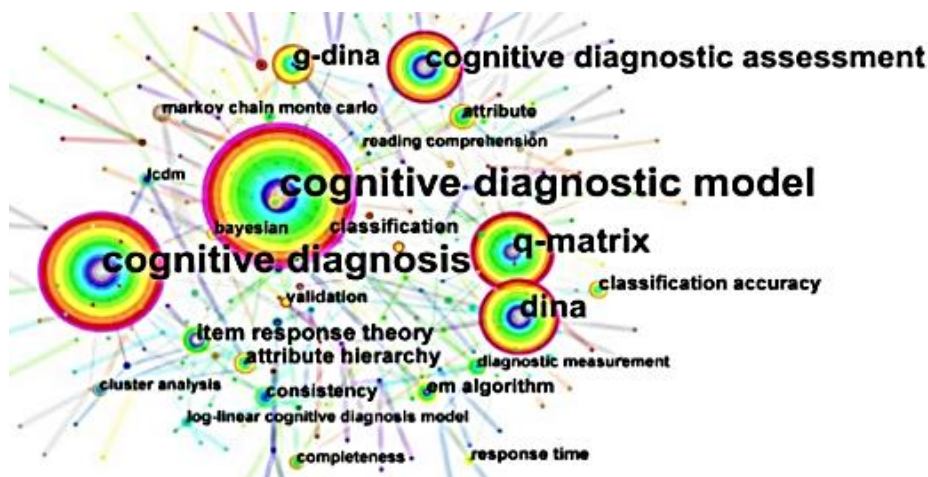
การทบทวนงานวิจัยที่เกี่ยวข้องกับการประเมินความสามารถการอ่านเพื่อความเข้าใจตั้งแต่ปี ค.ศ. 1995 ถึง ค.ศ. 2011 (McKee, 2012) พบว่า วรรณกรรมที่ศึกษาเป็นการวิจัยเกี่ยวกับการสร้างแบบทดสอบที่ใช้วินิจัยทางพุทธิปัญญาให้มีรูปแบบที่หลากหลายเพื่อสามารถประเมินได้ครอบคลุมระดับขั้นความสามารถทางสมอง 6 ระดับของ Bloom เรียงจากระดับ 1 ได้แก่ ความรู้ ความเข้าใจ การประยุกต์ การวิเคราะห์ การสังเคราะห์ และการประเมิน โดยกำหนดวัตถุประสงค์การประเมินเพื่อระบุจุดแข็งและจุดอ่อนของผู้สอบ และการประเมินระดับความรู้และทักษะโดยศึกษาความสัมพันธ์ระหว่างผลการสอบกับตัวแปรประเภทการทดสอบและรูปแบบการทดสอบการ อาทิ แบบเลือกตอบซึ่งสามารถตรวจให้คะแนนได้อย่างสะดวก แบบเขียนเรียงความซึ่งผู้สอบต้องใช้ทักษะการคิดและการวิเคราะห์มากขึ้น แบบโคลซ (Cloze test) ซึ่งมีประสิทธิภาพในการประเมินความรู้ คำศัพท์ หรือการใช้แบบทดสอบมาตรฐานที่มีอยู่ โดยศึกษาปัจจัยที่เกี่ยวข้องกับการทดสอบและการอ่านเพื่อความเข้าใจ อีก 2 ประการคือความยาวของบทอ่าน (Passage) และข้อจำกัดด้านเวลาส่งผลให้ผู้สอบเกิดความเครียด จึงมีการนำโมเดลทางสถิติมาใช้ในการวินิจัยแพร่หลายมากขึ้น

การใช้เทคนิคที่ทันสมัย อาทิ การวิเคราะห์ด้วยภาพ (Visual Analytics) การนำเสนอข้อมูลเชิงกราฟิก (Data visualization) เป็นการแสดงข้อมูลในรูปแบบแผนผังโครงข่ายคำสำคัญที่พบร่วมกันของบทความและงานวิจัย (The Keyword Co-occurrence Network) รวมถึงคุณลักษณะหรือตัวแปรสำหรับหน่วยของข้อมูล (Tinnamas, 2020) Figure 2 แสดงผลข้อมูลเครือข่ายของชื่อเรื่องงานวิจัยเมื่อเลือกแสดงผลด้วย Keyword ซึ่งเป็นวรรณกรรมจาก Web of Science ในช่วงเวลาระหว่างปี 2000 ถึง 2022 ใช้คำค้น "cognitive* diagnos*" (ทางปัญญา* การวินิจัย*) ในชื่อเรื่อง บทคัดย่อ และคำสำคัญ ได้ผลลัพธ์จำนวน 927 ฉบับ ผ่านโปรแกรม Cite Space ที่พัฒนาขึ้นโดย Chen et al. (2016) เพื่อนำเสนอภาพรวมทิศทางการศึกษาวิจัยของ CDA โดยการบูรณาการข้อมูลนามธรรมจำนวนมหาศาลให้เป็นการแสดงผลเชิงโต้ตอบให้มองเห็นภาพของแนวโน้มในพื้นที่ความรู้เฉพาะทาง

Figure 6

Visualization of a Co-occurrence Network of Keywords in Cognitive Diagnostic Assessment

การแสดงความคล้ายคลึงร่วมของการประเมินวินิจฉัยทางปัญญา



ที่มา: Sun and Hwang (2023)

โครงข่ายคำสำคัญร่วม (Key words Co-occurrence) ของงานวิจัยที่เกี่ยวข้องกับการประเมินวินิจฉัยทางปัญญา ซึ่งถูกพบมากที่สุด ได้แก่ คำว่า “โมเดลการวินิจฉัยทางพุทธิปัญญา” (Cognitive diagnostic model) รองลงไปคือคำว่า “การวินิจฉัยทางพุทธิปัญญา” (Cognitive diagnosis) “เมทริกซ์คิว” (Q-matrix) “DINA model” “การประเมินวินิจฉัยทางปัญญา” (Cognitive diagnostic assessment) “G-DINA model” “คุณลักษณะ”(Attribute) “การอ่านเพื่อความเข้าใจ” (Reading Comprehension) ตามลำดับ สิ่งนี้บ่งชี้ว่าทิศทางการศึกษาวิจัยในปัจจุบันมุ่งเน้นไปที่การสร้างทฤษฎีการประเมินวินิจฉัยทางปัญญาให้มีความแกร่งเป็นหลัก นักวิจัยได้พัฒนา ประเมิน และเปรียบเทียบประสิทธิภาพโมเดลการวินิจฉัยทางพุทธิปัญญา อาทิ โมเดล DINA, G-DINA และมีการสำรวจสังเคราะห์แนวคิดทฤษฎีที่มาของคุณลักษณะหรือทักษะย่อยของความสามารถเพื่อสร้างเมทริกซ์คิว มีการสร้างดัชนีและอัลกอริธึมใหม่ๆ ขึ้นมาเพื่อเพิ่มความแม่นยำของการประเมินวินิจฉัยทางปัญญา นอกจากนี้ นักวิจัยยังได้นำข้อค้นพบทางทฤษฎีมาประยุกต์ใช้กับการประเมินความสามารถการอ่านเพื่อความเข้าใจ ภาษาอังกฤษ นับว่าเป็นการขยายฐานรากทางทฤษฎีของการประเมินวินิจฉัยทางปัญญา (Sun & Hwang, 2023)

จากที่กล่าวมา จะเห็นว่าการสำรวจแนวโน้มและทิศทางของการประเมินวินิจฉัยทางปัญญาด้านความสามารถการอ่าน เพื่อความเข้าใจด้วยการทบทวนงานวิจัยที่เกี่ยวข้องและการใช้เทคโนโลยีการวิเคราะห์ด้วยภาพ (Visual Analytics) มีผลการศึกษาที่สอดคล้องกันในแง่ของการนำการประเมินวินิจฉัยทางปัญญา (CDA) มาประยุกต์ใช้ในด้านการศึกษาอย่างแพร่หลาย โดยมุ่งพัฒนาแบบทดสอบวินิจฉัยที่สามารถจำแนกผู้สอบออกเป็นผู้รอบรู้และไม่รอบรู้ในคุณลักษณะหรือทักษะย่อยของความสามารถการอ่านเพื่อความเข้าใจ โดยให้สารสนเทศที่ระบุจุดแข็งและจุดอ่อนที่ครูและผู้สอบจะนำไปเป็นแนวทางการพัฒนาทักษะความสามารถของผู้เรียน และการพัฒนาโมเดลทางสถิติในการวินิจฉัยทางพุทธิปัญญา

บทสรุป (Conclusion)

การประเมินวินิจฉัยทางปัญญา (CDA) ถูกนำมาใช้เพื่อวิเคราะห์จุดแข็งและจุดอ่อนของผู้เรียนที่เกิดขึ้นระหว่างการเรียนการสอนด้วยวิธีการทดสอบ โดยเครื่องมือที่ใช้วัดเป็นแบบทดสอบวินิจฉัยที่ถูกพัฒนาขึ้นตามแนวคิดทฤษฎีการรู้คิดตามแนวคิดระบบการออกแบบทางปัญญาซึ่งมีความตรงเชิงโครงสร้างสูง โดยใช้โมเดลทางปัญญา (CDMs) มาพัฒนาหรือกำหนดข้อสอบที่วัดองค์ความรู้และทักษะคุณลักษณะของความสามารถที่มุ่งประเมิน โมเดลวินิจฉัยทางพุทธิปัญญาเป็นวิธีการแบบสหวิทยาการที่มีการเชื่อมโยงระหว่างจิตวิทยาพุทธิปัญญา และการวิเคราะห์ทางสถิติ จึงสามารถสกัดสารสนเทศเชิงวินิจฉัยเพื่อใช้สำรวจความสัมพันธ์ระหว่างกระบวนการทางจิตและกลวิธีที่ใช้ในการแก้ปัญหาหรือตอบข้อสอบข้อหนึ่งๆ ให้สารสนเทศที่เป็นประโยชน์เกี่ยวกับจุดเด่นและจุดด้อยของผู้สอบในแต่ละคุณลักษณะเป็นรายบุคคล โมเดลวินิจฉัยทางปัญญาแต่ละโมเดลมีข้อดีและข้อจำกัดแตกต่างกัน จึงควรศึกษาเอกสารงานวิจัยที่เกี่ยวข้องเพื่อการตัดสินใจเลือกใช้โมเดลสอดคล้องกับบริบทของงานวิจัย ในการวิเคราะห์ข้อมูลทางสถิติมี software ที่ช่วยให้การวิเคราะห์ได้สะดวกรวดเร็วและมีความแม่นยำการพิจารณาโปรแกรมที่เป็น Free ware คือไม่เสียค่าใช้จ่าย เมื่อพิจารณาแนวโน้มและทิศทางของการประเมินวินิจฉัยทางปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจ ด้วยการทบทวนเอกสารงานวิจัยที่เกี่ยวข้องและผลการศึกษาโครงข่ายคำสำคัญร่วมของชื่อเรื่องงานวิจัยโดยใช้โปรแกรมที่ให้ผลการวิเคราะห์ด้วยภาพ (Visual Analytics) นำเสนอข้อมูลเชิงกราฟิก (Data visualization) พบว่าทั้ง 2 วิธีมีความสอดคล้องกัน สรุปได้ว่า นักวิจัยพัฒนาแบบทดสอบวินิจฉัยทางปัญญาเพื่อใช้ในการวินิจฉัยความสามารถทางภาษาอังกฤษเพื่อวัตถุประสงค์ของการประเมินวินิจฉัย 2 ประการคือ เพื่อระบุประเด็นที่นักเรียนยังไม่รอบรู้ เป็นการตรวจสอบความก้าวหน้าทางการเรียนในวิชาดังกล่าว และเพื่อให้ข้อเสนอแนะเกี่ยวกับสาเหตุหรือเหตุผลที่เป็นไปได้เกี่ยวกับการไม่รอบรู้ในทักษะย่อยหรือคุณลักษณะนั้นๆ ความสัมพันธ์ระหว่างคุณลักษณะและข้อสอบ และส่วนที่ 2 การเลือกใช้โมเดลการวัดทางจิตวิทยาซึ่งเป็นโมเดลทางสถิติที่เหมาะสมกับข้อสอบและคุณลักษณะที่สนใจ เพื่อใช้ในการค้นหาระดับความรู้หรือระดับความชำนาญของผู้เรียนในทักษะทางพุทธิปัญญาที่ระบุใน Q-matrix ดูได้จากรูปแบบการตอบข้อสอบ (Item response pattern) ของผู้เรียน

References

- Chen, H., & Chen, J. (2015). Exploring reading comprehension skill relationships through the G-DINA model. *Educational Psychology*, 36, 1049-1064. <https://doi.org/10.1080/01443410.2015.1076764>
- Chen, C. (2016). *Cite Space: A practical guide for mapping scientific literature*. Nova Science Publishers.
- Deenang, E. (2016). Cognitive Diagnostic Models. *Udonthani Rajabhat University Journal of Humanities & Social Sciences*, 5(1), 1-9.
- de la Torre, J. (2011). The Generalized DINA Model Framework. *Psychometrika*, 76, 179-199.
DOI: 10.1007/s11336-011-9207-7
- Embretson, S. E. (1994). *Cognitive Assessment: A multidisciplinary perspective*. New York: Plenum Press.
- Finocchiaro, M., & Sako, S. (1983). *Foreign Language Testing: A Practical Approach*. New York: Regents.
- Gao, L. (2006). Toward a cognitive processing model of MELAB reading test item performance. In J. S. Johnson (Ed.), *Spaan fellow working papers in second or foreign language assessment* (vol. 4, pp. 1-39). Ann Arbor, MI: University of Michigan.

- Huff, K., & Goodman, D. P. (2007). The demand for cognitive diagnostic assessment. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 19–60). Cambridge University Press.
- Jang, E. E. (2009). Cognitive diagnostic assessment of L2 reading comprehension ability: Validity arguments for Fusion Model application to Language Assessment. *Language Testing*, 26(1), 031-073. DOI: 10.1177/0265532208097336
- Kanjanawasee, S. (2009). *Modern Test Theory*. Bangkok: Chulalongkorn University Press.
- Khamyoo, T. (2014). Diagnosis of English reading skills by applying the trait hierarchy model & computer adaptive testing. *Journal of Research Methodology & Cognitive Science*, 10(2), 55-70.
- Lee, Y., & Sawaki, Y. (2009). Application of Three Cognitive Diagnosis Models to ESL Reading and Listening Assessments. *Language Assessment Quarterly*, 6, 239-263. DOI: 10.1080/15434300903079562
- Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoaka's rule-space approach. *Journal of Educational Measurement*, 41(3), 205-237.
- Leighton J. P., & Gierl M. J. (2007). *Cognitive Diagnostic Assessment for Education: Theory & Applications*. US: Cambridge University Press.
- Li, H. (2011). A Cognitive Diagnostic Analysis of the MELAB Reading Test. *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 9(1), 17-46.
- Ma, W., & de la Torre, J. (2020). GDINA: An R Package for Cognitive Diagnosis Modeling. *Journal of Statistical Software*, 93(14), 1–26. DOI: 10.18637/jss.v093.i14.
- Mckee, S. (2012). Reading Comprehension, What We Know: A Review of Research 1995 to 2011. *Language Testing in Asia*, 2, 45. <https://doi.org/10.1186/2229-0443-2-1-45>
- Moradi, E., Ghabanchi, Z., & Pishghadam, R. (2022). Reading comprehension test fairness across gender and mode of learning: insights from IRT-based differential item functioning analysis. *Lang Test Asia*, 12, 39. <https://doi.org/10.1186/s40468-022-00192-3>
- Ravand (2012). Exploring diagnostic capacity of a high stakes reading comprehension test: A pedagogical demonstration. *Iranian Journal of Language Testing*, 3(1), 11-37.
- Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications*. Guilford Press.
- Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 263–331). Macmillan Publishing; American Council on Education.
- Somrat, K. (2023). *The Development of English competence diagnostic tests for preservice teachers by applying the G-DINA model* (Dissertation dissertation). Phitsanulok: Naresuan University.
- Sun, H., & Hwang, Y. (2023). Enhancing EFL Learners' English Reading Performance through the Diagnostic Feedback of CDA. *Korean Journal of English Language & Linguistics*, 23(1), 535-553. DOI: 10.15738/kjell.23.202307.535

- Shahmirzadi, N., & Marashi, H. (2023) Cognitive diagnostic assessment of reading comprehension for high-stakes tests: Using GDINA model. *Language Testing in Focus: An International Journal*, 8, 1-16.
DOI: 10.32038/ltf.2023.08.01
- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 345-355. DOI: 10.1111/j.1745-3984.1983.tb00212.x
- Templin, J., & Henson, R. A. (2006). Measurement of Psychological Disorders Using Cognitive Diagnosis Models. *Psychological Methods*, 11(1), 287-305.
- Tinnamas, S. (2020). Biometric analysis of research articles written by Naresuan University professors in Scopus & Web of Science databases. *The 10th PULINET National Conference* (pp. 260-270). Chiang Rai: University of Mae Fah Luang.
- Tonekaboni, F. R., Ravand, H., & Rezvani, R. (2021). The Construction and Validation of a Q-matrix for a High-stakes Reading Comprehension Test: A G-DINA Study. *International Journal of Language Testing*, 11(1), 58-87.
- Ueabunpradit, P. (2020). *Development of Mathematical literacy diagnostic test with feedback for ninth grade students using Attribute Hierarchy Method* (Doctoral dissertation). Bangkok: Chulalongkorn University.
- Zhan, P., Li, X., Wang, W. C., Bian, Y., & Wang, L. (2015). The multidimensional test let-effect cognitive diagnostic models. *Acta Psychologica Sinica*, 47(1), 689–701. DOI: 10.3724/SP.J.1041.2015.00689.
- Nolen, Jeannette L., & Duignan, Brian. (2024). “No Child Left Behind”. *Encyclopedia Britannica*. Retrieved from <https://www.britannica.com/topic/No-Child-Left-Behind-Act>
- Templin, J. (2011). *Conceptual foundations of diagnostic measurement*. Retrieved from <https://www.jonathantemplin.com>