การพัฒนาแบบสอบวินิจฉัยโดยแนวคิดด้านพุทธิปัญญา

DIAGNOSTIC TEST DEVELOPMENT BY USING COGNITIVE DIAGNOSTIC MODEL

ผู้วิจัย พิศิษฐ ตัณฑวณิช[1]

บทคัดย่อ

การพัฒนาแบบสอบเพื่อการวินิจฉัยโดยแนวคิดด้านพุทธิปัญญา หรือภาษาอังกฤษใช้คำว่า Cognitive Diagnostic Model เป็นแนวคิดของการพัฒนาแบบสอบ เพื่อการชี้ระบุจุดอ่อนและจุดแข็งของผู้เรียนในวิชาต่าง ๆ ในหัวข้อบทเรียนหนึ่ง ๆ แนวคิดดังกล่าวมีการพัฒนาอย่างมากในช่วงยี่สิบปีที่ผ่านมา แต่ในแวดวงวิชาการด้านการวัดผลประเมินผลในประเทศไทย การกล่าวถึงสาระส่วนนี้แทบยังไม่ปรากฏ ผู้เขียนจึงได้นำเสนอแนวคิดนี้ โดยเริ่มต้นตั้งแต่แนวคิดการสร้างเมตริกความสัมพันธ์ระหว่างข้อสอบกับคุณลักษณะ (Attribute) ต่าง ๆ ที่มุ่งวัด เรียกว่า “Q–Matrix” จากนั้นได้กล่าวถึง แนวคิดในด้านรูปแบบ (Model) ทางคณิตศาสตร์ที่นักวิชาการ ใช้เป็นหลักในการพัฒนาสมการเพื่อนำไปสู่การหาค่าประจำข้อสอบ ประกอบด้วย The DINA Model The G–DINA Model The RUM Model และ The AHM Model จากนั้นผู้เขียนได้นำเสนอทัศนะที่มีต่อการพัฒนาแบบสอบวินิจฉัยดังกล่าว ทั้งในส่วนที่เป็นข้อดีและข้อจำกัด และท้ายสุดเห็นว่า ขณะนี้เป็นช่วงเวลาที่นักวิชาการไทยควรจะได้เข้ามามีบทบาทหรือมีส่วนร่วมในการวิจัยและพัฒนาศาสตร์ในส่วนนี้เพื่อก่อให้เกิดความก้าวหน้าทั้งทางวิชาการและประโยชน์ด้านการจัดการศึกษาของประเทศไทย

คำสำคัญ แบบสอบวินิจฉัย แนวคิดด้านพุทธิปัญญา

---

[1] รองศาสตราจารย์ ดร. สาขาวิชาทดสอบการประเมินผลและวิจัยการศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏสุรินทร์

ABSTRACT

New trend of test construction for diagnosis of students' strength and weakness was developed. It was named "Cognitive Diagnosis Model." (CDM). from the past two decades a lot of articles were printed, various mathematical models were proposed. But in Thai academic atmosphere, there were very few articles mentioned this progression, so in this article concepts about CDM were presented, it consisted of Q–matrix idea, various models about CDM which were the DINA Model, the G–DINA Model, the RUM model, and the AHM Model. After that the author presented some ideas about the goodness and some obstructions of CDA. Finally it was suggested that now it was the best opportunity for Thai scholars to take part in the research about this topics for both academic progression and the development of Thai education.

KEYWORDS : Diagnostic test development, Cognitive diagnostic model

บทนำ

การพัฒนาแบบสอบเพื่อการวินิจฉัยโดยแนวคิดด้านพุทธิปัญญาเป็นคำที่ผู้เขียนต้องการใช้ให้ได้ความหมายอย่างกระจ่างให้สอดรับกับคำในภาษาอังกฤษว่า "Cognitive Diagnosis Model" (ตัวย่อ CDM) การพัฒนาแบบสอบดังกล่าวมีปรากฏในวารสารที่เป็นภาษาอังกฤษอย่างมากมายในช่วงยี่สิบปีที่ผ่านมา แต่เท่าที่ผู้เขียนได้อ่านบทความทางด้านการวัดและประเมินผลทางการศึกษา ยังไม่พบว่ามีการเขียนบทความถ่ายทอดแนวคิดดังกล่าวเป็นภาษาไทย บทความนี้จึงมุ่งนำเสนอแนวคิดขั้นพื้นฐานและพัฒนาการที่มีมาของการพัฒนาแบบสอบในแนวดังกล่าว โดยมีหัวข้อหลักที่จะนำเสนอดังนี้

1. ความเป็นมาและแนวคิดขั้นพื้นฐาน
2. โมเดลของการพัฒนาแบบสอบ
   1) แนวคิดพื้นฐานการวิเคราะห์
   2) The DINA Mode และโมเดลใกล้เคียง
   3) The G–DINA Model
   4) The RUM Model
   5) The AHM Model
3. ข้อดีและข้อจำกัดของการพัฒนาแบบสอบเพื่อการวินิจฉัยโดยแนวคิดด้านพุทธิปัญญา แต่ละหัวข้อมีสาระดังนี้

1) ความเป็นมาและแนวคิดขั้นพื้นฐาน

บุคคลที่ได้รับการยกย่องว่าเป็นผู้เริ่มต้นแนวคิดการพัฒนาแบบสอบในแนวคิดนี้ได้แก่ Tatsuoka, K.K. (1983 และ 1985) และ Tatsuoka, K.K. และ Tatsuoka, M.M. (1992) โดยบุคคลทั้งสองมุ่งสนใจศึกษาหาข้อบกพร่องในการเรียนรู้ของนักเรียน และได้ปรากฏแนวคิดตัวอย่างการนำเสนอแนวทางการวิเคราะห์ภาระงานในด้านการบวกและลบเศษส่วนของนักเรียนในลักษณะที่ว่า ข้อคำถามหนึ่ง ๆ ในด้านการบวก หรือการลบเศษส่วน นักเรียน

จะทำได้ถูกต้องหรือไม่นั้น นักเรียนจะต้องมีความรู้ความสามารถในคุณลักษณะ(Attributes) ต่าง ๆ อะไรบ้าง ตัวอย่าง เช่นการออกแบบข้อสอบในแนวทางพุทธิปัญญาด้านการบวกเศษส่วน 2 จำนวน (ที่ใช้หลักการดึงจำนวนเต็มออกมาบวกนอกเศษส่วนกรณีที่เป็นเศษคละ)

จากปัญหาดังกล่าว เมื่อวิเคราะห์ในรายละเอียดถึงคุณลักษณะที่ผู้เรียนจะต้องมีเพื่อที่จะได้ตอบข้อคำถามหนึ่ง ๆ ได้ถูกต้อง สามารถนำเสนอได้เป็นตารางความสัมพันธ์ดังนี้

ตารางที่ 1 แสดงความเกี่ยวข้องของข้อสอบแต่ละข้อกับคุณลักษณะที่จำเป็นต้องใช้ เพื่อการบวกเศษส่วนสอง จำนวนได้ถูกต้อง

ข้อสอบ	คุณลักษณะประจำข้อสอบ (Attributes) เพื่อการทำข้อสอบได้ถูกต้อง					
	K1	K2	K3	K4	K5	K6
$\frac{1}{5}+\frac{3}{5}$	1	0	0	0	0	0
$\frac{1}{3}+\frac{2}{3}$	1	0	0	0	0	1
$\frac{1}{4}+\frac{1}{2}$	1	1	0	0	0	0
$\frac{5}{4}+\frac{1}{2}$	1	1	0	1	1	0
$\frac{1}{3}+\frac{2}{7}$	1	0	1	0	0	0
$1\frac{1}{2}+\frac{7}{9}$	1	0	1	0	1	1
$\frac{6}{5}+\frac{1}{5}$	1	0	0	1	1	0
$1\frac{1}{3}+1\frac{1}{6}$	1	1	0	0	1	1
$1\frac{1}{3}+1\frac{4}{5}$	1	0	1	0	1	1
$\frac{8}{3}+\frac{3}{4}$	1	0	1	1	1	1

(ปรับจากข้อเขียน Tatsuoka, K.K. และ Tatsuoka, M.M. (1992)

ความหมายของคุณลักษณะในตาราง

K1. เมื่อส่วนเท่ากัน ใช้วิธีการนำเอาเศษมาบวกกัน ส่วนคงเดิม

K2. ทำส่วนให้เท่ากัน เมื่อส่วนของจำนวนหนึ่งเป็นสัดส่วนของส่วนอีกจำนวนหนึ่ง

K3. ทำส่วนให้เท่ากัน เมื่อใช้หลักตัวคูณร่วม

K4. การปรับค่าเศษเกินเป็นเศษคละ

K5. การดึงจำนวนเต็มออกบวกกันและบวกกับเศษส่วนที่มีตามมา

K6. การปรับคำตอบให้เป็นเศษคละ จำนวนเต็ม หรือเป็นเศษส่วนอย่างต่ำตามแต่กรณี

โดย 0 หมายถึง ข้อคำถามไม่เกี่ยวข้องกับคุณลักษณะดังกล่าวนั้น

และ 1 หมายถึง ข้อคำถามนั้น ๆ เกี่ยวข้องกับคุณลักษณะดังกล่าวนั้น

จากตารางที่ 1 มีความหมายว่า การบวกเศษส่วนสองจำนวนโดยวิธีการดึงจำนวนเต็มออกมาบวกกันนั้น มีคุณลักษณะที่ผู้เรียนจะต้องใช้วิเคราะห์ ได้เป็น 6 คุณลักษณะ โดยข้อสอบแต่ละข้อจะมีความสามารถในการวัดคุณลักษณะดังกล่าวได้ต่างกันออกไป ทั้งนี้ในขั้นของการวิเคราะห์ผลการตอบของนักเรียนจะไม่สนใจในคะแนนรวมที่นักเรียนทำได้จากข้อสอบทุกข้อ แต่จะเน้นผลการวิเคราะห์ลงไปว่า ผู้เรียนแต่ละคนเรียนรู้และถือได้ว่าเป็นผู้รอบรู้ (Master) แล้วในคุณลักษณะใดบ้าง และมีคุณลักษณะใดบ้าง ที่นักเรียนยังไม่รอบรู้และยังกระทำไม่ได้ (Non–master) ซึ่งผลจากการวิเคราะห์ข้อมูลในส่วนนี้จะเน้นหนักในด้านการนำผลไปใช้พัฒนาการเรียนรู้ของนักเรียนในแต่ละคนให้สอดคล้องกับพื้นความรู้เดิมที่มีต่อไป

เมตริกซึ่งพัฒนาขึ้นโดยแนวคิดนี้มีชื่อเรียกเฉพาะในช่วงต่อมาว่า “Q–Matrix” และเรียกการมุ่งเน้นศึกษาผลเพื่อพัฒนานักเรียนตามส่วนของคุณลักษณะย่อย ๆ นี้ว่าเป็นการพัฒนาแบบ “Grain Size”

แนวคิดทั้งหมดนี้เป็นความหวังของนักพัฒนาแบบสอบที่คาดว่าจะช่วยให้การแก้ไขพัฒนานักเรียนระหว่างเรียนเป็นสิ่งที่กระทำได้อย่างแท้จริง นับจาก ช่วงปี 1990 เป็นต้นมา แนวคิดส่วนนี้ได้รับการกล่าวถึง มีการเขียนเป็นบทความวิชาการ และมีการพัฒนาโมเดลทางคณิตศาสตร์ขึ้นมาใช้ประกอบแนวคิดหลายโมเดล

2) โมเดลของการพัฒนาแบบสอบ

จากแนวคิดขั้นพื้นฐานในข้อ 1 ได้มีนักวิชาการพัฒนาโมเดลของการวิเคราะห์ข้อสอบนำเสนอขึ้นมาอย่างหลากหลาย จนอาจกล่าวได้ว่าเป็นยุคความเจริญรุ่งเรืองของแนวคิดการพัฒนาโมเดลแบบสอบเพื่อการวินิจฉัย

ในที่นี้ผู้เขียนจะขอกล่าวแนะนำแนวคิดของการพัฒนาโมเดลที่มีการกล่าวถึงค่อนข้างมาก โดยมีหัวข้อย่อยดังนี้

2.1) แนวคิดพื้นฐานการวิเคราะห์

2.2) The DINA Model และโมเดลใกล้เคียง

2.3) The G–DINA Model

2.4) The RUM Model

2.5) The AHM Model

แต่ละส่วนมีสาระดังนี้

2.1) แนวคิดพื้นฐานการวิเคราะห์

ค่าประจำข้อสอบในแนวคิด CDM มีขั้นตอนการวิเคราะห์ที่ซับซ้อนพอควร และต้องใช้โปรแกรมคอมพิวเตอร์เข้ามาช่วยวิเคราะห์ โดยนักวิเคราะห์ในแนวคิดดังกล่าวให้คำอธิบายเบื้องต้นไว้ดังนี้

ก. เนื่องจากจุดสนใจในผลการวิเคราะห์คือการพิจารณาจากคุณลักษณะต่างๆ ของผู้สอบที่สร้างไว้ใน Q–Matrix เหตุนี้คะแนนที่ได้จากการสอบของข้อสอบแต่ละข้อจึงไม่สามารถที่จะนำมารวมกันได้คะแนนในโมเดลดังกล่าวเหล่านี้ มีลักษณะเฉพาะที่ถือว่าเป็น Probabilistic, Multidimensional Latent–variable Models รวมทั้งมีลักษณะเป็น Categorical predictor variables (Li, 2011)

ข. การวิเคราะห์ความสามารถของผู้สอบ จะใช้หลักวิชาที่เรียกว่า “Latent Class Model” (Haertel, 1984) มาจัดกลุ่มและประมาณค่าความสามารถของกลุ่มผู้สอบ โดยอาศัยหลักการที่ว่าจากคำตอบของกลุ่มบุคคลหนึ่ง ๆ ในการทำข้อสอบทุกข้อ สามารถที่จะแบ่งผู้สอบออกเป็นกลุ่มย่อย ๆ ได้ทั้งหมด 2^k กลุ่ม เมื่อ k หมายถึง จำนวนคุณลักษณะที่มุ่งสนใจศึกษาใน Q–Matrix ทั้งนี้ เมื่อมีการกำหนดระบบให้คะแนนเป็น 1 กับ 0

สมมุติเช่น ในการสอบครั้งหนึ่ง มุ่งวัดจำนวน 3 คุณลักษณะ ย่อมแบ่งกลุ่มผู้สอบตามความสามารถเป็น $2^3 = 8$ กลุ่มย่อย คือ กลุ่มที่ทำคะแนนได้เป็น 000 100 010 001 100 101 011 และกลุ่ม 111 ตามลำดับ (000 หมายถึงกลุ่มบุคคลที่ทำข้อสอบผิดหมดทุกคุณลักษณะ คุณลักษณะ 100 หมายถึงคนกลุ่มที่ทำข้อสอบถูก ในคุณลักษณะที่ 1 เท่านั้น กลุ่มอื่นก็อ่านผลเช่นเดียวกัน)

นักวิชาการในกลุ่มนี้นำเสนอการวิเคราะห์ข้อมูลโดยใช้วิธีการสถิติที่เป็นการประมาณค่าพารามิเตอร์ต่าง ๆ เช่น วิธีการทางด้าน Bayesian Procedures หรือใช้วิธีการที่เรียกว่า Marchov Chain Monte Carlo เป็นต้น เป็นแนวทาง ซึ่งสถิติในกลุ่มนี้มุ่งเน้นการประมาณค่าพารามิเตอร์ของข้อสอบโดยมีการประมาณค่าก่อนดำเนินการ แล้วใช้วิธีการวงจรหมุนวนทางโปรแกรมคอมพิวเตอร์ (การประมาณค่าในวงจร Loop) เพื่อประมาณหาค่าพารามิเตอร์ประจำข้อสอบ ที่ควรจะเป็นมากที่สุด (Maximum Likelihood)

นอกจากนี้ นักคิดในแนวคิดของการพัฒนาแบบสอบดังกล่าวได้ย้ำเตือนว่าสิ่งที่สำคัญที่สุดของขั้นตอนการพัฒนา คือ การที่จะต้องมี Q–Matrix ที่ผ่านการวิเคราะห์มาเป็นอย่างดีจนเชื่อถือได้ว่าผลการวิเคราะห์คุณลักษณะในขอบเขตที่ต้องการศึกษาได้รับการพิจาณาอย่างถูกต้อง ครอบคลุมทุกประเด็นที่มุ่งศึกษา (De Carlo, 2011) และสิ่งหนึ่งที่นักวิเคราะห์ เพื่อสร้าง Q–Matrix จะต้องตระหนักถึง คือว่ากลยุทธ์ ในการคิดหรือการให้ได้มาซึ่งคำตอบของข้อถามต่าง ๆ นั้น นักเรียนต่างคนกันอาจใช้วิธีคิดหรือกลยุทธ์ต่างกันออกไปได้ซึ่งอาจมีผลทำให้การวิเคราะห์จุดที่เรียนรู้แล้ว และยังไม่เรียนรู้ของนักเรียนผิดไปได้ เช่น การบวกเศษส่วนสองจำนวน ผู้สอบสามารถที่

จะใช้วิธีการคิดในการหาคำตอบที่ถูกต้องต่างกันได้ ดังตัวอย่างเช่นวิธีการบวกเศษส่วนต่อไปนี้

วิธีคิดแบบที่ 1 ใช้วิธีการทำให้เป็นเศษเกินก่อน

$$1\frac{1}{2}+\frac{3}{4}=\frac{3}{2}+\frac{3}{4}=\frac{6}{4}+\frac{3}{4}=\frac{9}{4}=2\frac{1}{4}$$

วิธีคิดแบบที่ 2 ใช้วิธีดึงจำนวนเต็มออกมาคิดคำนวณ

$$1\frac{1}{2}+\frac{3}{4}=1+\frac{1}{2}+\frac{3}{4}=1+\frac{2}{4}+\frac{3}{4}=1\frac{5}{4}=2\frac{1}{4}$$

ซึ่งแนวคิดวิธีการบวกที่ต่างกันเช่นนี้ ทำให้ผู้สร้างข้อสอบเพื่อการวินิจฉัยในแนวพุทธิปัญญาต้องคำนึงถึงอย่างมากในแนวการกำหนดคุณลักษณะเพื่อการสร้าง Q–Matrix ในแต่ละกรณีของการศึกษา

2.2) The DINA Model และโมเดลข้างเคียง

รูปแบบการวิเคราะห์ในแนวคิดของ The DINA Model (DINA ย่อมาจากคำว่า Deterministic inputs, noisy “and” gate) แนวคิดนี้มีข้อตกลงเบื้องต้นว่า การที่บุคคลหนึ่งจะตอบข้อสอบข้อหนึ่ง ๆ ได้ถูกต้องก็ต่อเมื่อ บุคคลนั้นมีความรอบรู้ครบในทุก ๆ คุณลักษณะที่ข้อสอบนั้นต้องการ ซึ่งเรียกคุณสมบัติของข้อสอบในลักษณะนี้ว่าเป็นแบบ Conjunctive Model หรือเรียกอีกชื่อว่าเป็น Non–compensatory Model การวิเคราะห์ข้อสอบในแนว The DINA Model มุ่งหาค่าประจำข้อสอบ 2 ค่าคือ

Slip Parameters หมายถึงโอกาสในการตอบข้อสอบผิดของผู้ที่รอบรู้ในทุกคุณลักษณะที่ข้อสอบข้อนั้นมุ่งวัด แทนด้วยสัญลักษณ์ “ s_j ” ซึ่งทำให้โอกาสในการตอบถูกของผู้สอบที่รอบรู้ในคุณลักษณะที่มุ่งวัดมีค่า 1 – s_j)

Guessing Parameters หมายถึงโอกาสในการตอบข้อสอบถูกของผู้ที่ไม่รอบรู้ในคุณลักษณะที่ข้อสอบข้อนั้นมุ่งวัด แทนด้วยสัญลักษณ์ “ g_j ”

โดยขยายความได้เพิ่มเติมดังนี้

กำหนดให้ i=1,2.....I แทนลำดับที่ของคนที่เข้าสอบ

กำหนดให้ j=1,2.....J แทนลำดับที่ของข้อสอบทั้งหมดในแบบสอบ

กำหนดให้ k=1,2...K แทนลำดับที่ของคุณลักษณะที่มุ่งศึกษา

ดังนี้จะได้ว่า

$$\eta_{ij}=\prod_{k=1}^{k}\alpha_{ij}^{q_{jk}}$$

η_{ij} หมายถึงสภาพการรอบรู้ (1) หรือไม่รอบรู้ (0) ของผู้เข้าสอบคนที่ i ในข้อสอบข้อที่ j

α_{ik} หมายถึงสถานภาพความรู้ของผู้สอบคนที่ i ในคุณลักษณะต่าง ๆ ของ k

q_{jk} หมายถึงคุณลักษณะที่ k ที่จำเป็นต้องใช้ในการตอบข้อสอบข้อที่ j ให้ถูกต้อง

η_{ij} จากสมการดังกล่าวจะมีค่าเป็น 1 ในกรณีที่ผู้สอบมีความรอบรู้ในทุกๆ คุณลักษณะที่ต้องการใช้ในการตอบข้อสอบข้อนั้น แต่ถ้าหากผู้เข้าสอบขาดความรอบรู้ในคุณลักษณะใดลักษณะหนึ่งที่จำเป็นต้องใช้เขาจะตอบข้อสอบข้อนี้ไม่ได้ ค่าดังกล่าวจะเป็น 0

ดังนี้โอกาสความน่าจะเป็นที่ผู้เข้าสอบคนที่ i จะตอบข้อสอบข้อที่ j ได้ถูกต้องเขียนเป็นสมการหลักได้ดังนี้

$$P(X_{ij}=1|\eta_{ij})=g_j^{1-\eta_{ij}}(1-S_j)^{\eta_{ij}}$$

(de la Torre, 2009 : 118)

ทั้งนี้

$$P(X_{ij}=1|\eta_{ij}=1)=1-S_j$$

ซึ่งหมายถึงโอกาสที่ผู้สอบที่รู้จริงจะทำข้อสอบข้อนี้ได้ถูกต้องและ

$$P(X_{ij}=1|\eta_{ij}=0)=g_j$$

ซึ่งหมายถึงโอกาสที่ผู้ไม่รอบรู้จะเดาตอบข้อสอบข้อนี้ได้ถูกต้อง

โมเดลใกล้เคียงในที่นี้หมายถึง โมเดลที่มีแนวคิดส่วนหนึ่งคล้ายคลึงกับ The DINA Model ที่มีนักวิชาการได้เสนอความคิดและพัฒนาในลำดับต่อมา เช่น

–The DINO Model (Deterministic inputs, noisy "or" gate) เป็นโมเดลที่มีแนวคิดส่วนหนึ่งสอดคล้องกับแนวคิดของ The DINA Model มีความแตกต่างตรงที่ว่า The DINO Model ยึดหลักคิดว่า การที่บุคคลหนึ่งจะตอบข้อสอบข้อหนึ่ง ๆ ได้ถูกต้องนั้น ผู้ตอบอาจแค่มีความรู้ในคุณลักษณะที่สำคัญ บางคุณลักษณะโดยไม่จำเป็นต้องครบทุกคุณลักษณะก็สามารถเอื้อให้ผู้ตอบ ตอบคำถามนั้น ๆ ถูกต้องได้ ซึ่งเรียกว่าเป็น Disjunctive Model หรือเรียกอีกชื่อว่าเป็น Compensatory Model

–The NIDA Model (Noisy inputs, deterministic "and" gate) เป็นโมเดลในแนวคิด Conjunctive Model เช่นเดียวกันกับ The DINA Model แต่การประมาณค่า Slipping Parameters และ Guessing Parameters มุ่งดำเนินการในระดับของ attribute แต่ละตัว ไม่ใช่การวิเคราะห์ระดับรายข้อ

–The NIDO Model (Noisy inputs, deterministic "or" gate) เป็นโมเดลในลักษณะเดียวกันกับ The NIDA Model แต่มีลักษณะเป็น Disjunctive Model (Roussos, Templin, & Henson, 2007)

2.3) The G–DINA Model (Generalized DINA Model)

The G Dina Model ถือได้ว่าเป็นบทขยายของ The DINA Model แนวคิดนี้ได้รับการพัฒนาขึ้นโดย de la Torre, J (2011) ลักษณะของการวิเคราะห์ข้อสอบจะเป็นการหาว่า เมื่อแยกพิจารณาข้อสอบแต่ละข้อตามคุณลักษณะต่าง ๆ ที่ผู้สอบจำเป็นต้องใช้เพื่อการตอบข้อสอบแต่ละข้อได้ถูกต้อง โอกาสการตอบถูกของผู้สอบที่มีความรอบรู้ในคุณลักษณะที่ต่างกันออกไป ควรมีค่าความน่าจะเป็นอยู่เท่าใด ยกตัวอย่างเช่น

ข้อสอบที่ 1 ซึ่งมุ่งวัดคุณลักษณะเพียงคุณลักษณะเดียว ผลที่ได้จากการวิเคราะห์จะเป็นทำนองดังเช่น

คุณลักษณะการรอบรู้	0 ผู้สอบไม่รอบรู้	1 ผู้สอบเป็นผู้รอบรู้
โอกาสตอบถูก	0.11	0.84

(โอกาสตอบถูกเป็นค่าที่ได้จากการประมาณทางสถิติ ผลรวมไม่จำเป็นที่จะต้องเป็น "1")

ข้อสอบที่ 2 ซึ่งมุ่งวัดคุณลักษณะอยู่ 2 คุณลักษณะ ผลที่ได้จากการวิเคราะห์จะเป็นทำนองดังเช่น

คุณลักษณะ การรอบรู้	00	01	10	11
โอกาสตอบถูก	0.05	0.11	0.15	0.74

00 หมายถึงโอกาสที่จะตอบถูกของผู้สอบที่ไม่รอบรู้ในคุณลักษณะทั้งสอง

01 และ 10 หมายถึงโอกาสที่จะตอบถูกของผู้สอบที่รอบรู้ในคุณลักษณะใดเพียง 1 คุณลักษณะ

11 หมายถึงโอกาสที่จะตอบถูกของผู้สอบที่รอบรู้ในคุณลักษณะทั้งสองที่มุ่งวัด

ข้อสอบที่มุ่งวัดในคุณลักษณะ 3 ประการ ก็จะกระจายออกไปในทำนองเดียวกันนี้

2.4) The RUM Model RUM เป็นตัวย่อของคำว่า The Reparameterized Unified Model ในตอนแรก ของการพัฒนาโมเดลนี้มีชื่อว่า The Unified Model โดยผู้ร่วมพัฒนาคือ Di Bello, L.V.; Stout, W.F.; and Roussos, L.A. แต่ในตอนแรกของการพัฒนาไม่สามารถพัฒนาสมการไปถึงขั้นการหาค่าพารามิเตอร์ทั้งหมดได้ จนกระทั่งในปี 2002 Hartz, S.M. ได้ทำวิทยานิพนธ์ปริญญาเอกและได้ลดค่าพารามิเตอร์บางส่วนในโมเดลลง จนสามารถประมาณค่าพารามิเตอร์ต่าง ๆ ได้ เกิดเป็นโมเดลที่ใช้ชื่อว่า The RUM Model ดังกล่าว รูปแบบของสมการหลักในโมเดลคือ

$$P(X_{ij}=1|\underline{\alpha}_j,\eta_j)=\pi_i^*\prod_{k=1}^{k} r_{ik}^{*(1-\alpha_{jk})\times q_{ik}} P_{c_i}(\eta_j)$$

(Roussos, Dibello, Stout, Hartz, Henson & Templin in Leighton & Gierl, 2007)

จากสมการหลักดังกล่าว โมเดลนี้มุ่งหาค่าประจำข้อสอบ 2 ค่า คือ

1. ค่า π_i^* หมายถึงค่าโอกาสความน่าจะเป็นในการตอบข้อสอบข้อที่ i ได้ถูกต้องเมื่อผู้สอบมีความรอบรู้ในทุก ๆ คุณลักษณะที่จำเป็นที่ใช้ในการตอบข้อสอบข้อนั้น ๆ (หรือ คือ เงื่อนไข $\prod_{k=1}^{k}\pi_{ik}^{q_{ik}}$) ซึ่งจะมีค่าอยู่ในช่วง 0 ถึง 1

2. ค่า $r_{ik}^{*}=\frac{P(Y_{ijk}=1|\alpha_{jk}=0)}{P(Y_{ijk}=1|\alpha_{jk}=1)}=\frac{r_{ik}}{\pi_{ik}}$

ค่า r_{ik}^{*} นั้น สามารถเขียนได้อีกรูปหนึ่งดังนี้

$$r_{ik}^{*}=\frac{g_{ik}}{1-s_{ik}}$$

(Su, 2013 : 28)

ค่าดังกล่าวคือค่าสัดส่วนระหว่างโอกาสในการตอบข้อสอบถูกกรณีที่ผู้สอบไม่มีความรอบรู้ในคุณลักษณะหนึ่ง ๆ ที่ใช้ในการตอบข้อสอบข้อนั้น ๆ กับโอกาสในการตอบถูกของผู้สอบที่มีความรอบรู้ในคุณลักษณะที่ใช้ในการตอบข้อสอบข้อนั้น ค่าดังกล่าวนี้จะดีก็ต่อเมื่อตัวเศษมีค่าเข้าใกล้หรือ เป็น 0 ค่าดังกล่าวจึงสะท้อนถึงความจำเป็นของคุณลักษณะที่ใช้ในการตอบข้อสอบนั้นว่ามีมากในระดับใด โดยหากค่าเข้าใกล้ศูนย์ก็จะยิ่งแสดงว่าคุณลักษณะดังกล่าว มีความจำเป็น มาก หากเข้าใกล้ 1 ก็แสดงว่าคุณลักษณะดังกล่าวไม่มีความจำเป็นที่จะใช้ในการตอบข้อสอบ

ข้อนั้น ทั้งนี้ในการวิเคราะห์ข้อสอบแต่ละข้อจะได้จำนวนค่า r_{ik}^{*} เท่ากับจำนวนคุณลักษณะที่กำหนดไว้ใน Q-Matrix

สำหรับค่า c ในส่วนของเทอม $P_{c_i}(\eta_j)$ นั้นเป็นตัวชี้ระบุว่า Q-Matrix ที่กำหนดไว้นั้นยังคงมีปัจจัยหรือคุณลักษณะภายนอกใด ๆ ที่ยังไม่ได้ระบุไว้ใน Q-Matrix เข้ามามีอิทธิพลต่อการตอบข้อสอบข้อดังกล่าวถูกอีกหรือไม่ โดยหาก c มีค่าสูงกว่า 3 ขึ้นไป ถือว่าไม่มีคุณลักษณะหรือปัจจัยภายนอกใดๆ เข้ามามีอิทธิพลต่อการตอบข้อสอบข้อนั้นอีก (ในการวิเคราะห์ข้อสอบ หากผู้วิเคราะห์ไม่ต้องการศึกษาค่าดังกล่าวนี้ สามารถกำหนดให้ $c = \emptyset$ ซึ่งโปรแกรมจะงดนำค่าดังกล่าวเข้าไปร่วมในสมการคำนวณ)

(ซึ่งจะเห็นได้ว่า หากไม่พิจารณาถึงความยากในการวิเคราะห์หาค่าประจำข้อสอบแล้ว โมเดลนี้ให้ค่าประจำข้อสอบที่ใกล้เคียงกับวิธีการวิเคราะห์ข้อสอบแบบเดิม ๆ ที่มีมาก่อนมากที่สุด โดยในทางปฏิบัติ มีการเสนอว่าค่า r_{ik}^{*} ไม่ควรจะมีค่าเกิน 0.90 และค่า π_i^{*} ไม่ควรจะต่ำกว่า 0.60 (Gu, 2011)

2.5) The AHM Model

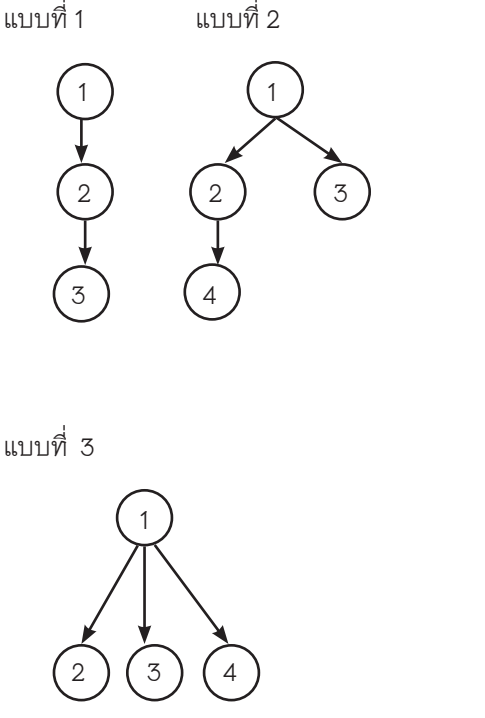
AHM ย่อมาจากคำว่า Attribute Hierarchical Model นักวิชาที่มุ่งสนใจและพัฒนางานวิจัยในด้านนี้มากที่สุดคือ Gierl, M.J. โมเดลนี้นำเสนอความคิดว่า คุณลักษณะที่มุ่งศึกษานั้นควรจะมีคุณสมบัติสำคัญคือ การเรียงตัวจากคุณลักษณะที่ง่ายไปหาคุณลักษณะที่ยากและซับซ้อนยิ่งขึ้นตามลำดับ เช่นข้อสอบเศษส่วน 3 ข้อต่อไปนี้ควรจะเรียงจากง่ายไปหายากได้ดังเช่น

ข้อ 1 $\frac{1}{5}+\frac{2}{5}=?$

ข้อ 2 $\frac{1}{2}+\frac{2}{5}=?$

ข้อ 3 $1\frac{1}{4}+\frac{7}{3}=?$

หน้าที่ของผู้สร้างแบบข้อสอบคือจะต้องวิเคราะห์ภารกิจของการเรียนในหัวข้อเรื่องที่ต้องการสร้างแบบสอบว่ามีลักษณะการเรียนรู้ที่เรียงตัวจากง่ายไปหายากในลักษณะใด ทั้งนี้รูปแบบการเรียงตัวจากง่ายไปหายาก อาจจะมีรูปแบบที่ต่างกันได้อย่างหลากหลายรูปแบบ ยกตัวอย่างบางส่วนที่เป็นไปได้ เช่น



ภาพที่ 1 รูปแบบการเรียงตัวของลักษณะการเรียนรู้

แนวคิดนี้นำเสนอว่า ค่าประจำข้อสอบสามารถใช้ค่าประจำข้อสอบตามแนวทฤษฎีการตอบสนองต่อข้อสอบ (Item Response Theory) เป็นแนวทางการวิเคราะห์ (หมายถึงการหาค่า a, b และ c) โดยค่าความสามารถของผู้สอบ (ค่า θ) และค่าสารสนเทศประจำข้อสอบ (Item Information Function) จะเป็นตัวชี้นำถึงคุณลักษณะที่ผู้เรียนรอบรู้แล้ว และคุณลักษณะที่ผู้เรียนยังไม่รอบรู้ สาระที่สำคัญที่สุดคือในตอนการสร้างข้อสอบ ผู้สร้างจะต้องวิเคราะห์ลักษณะเรียงลำดับความยากของข้อสอบจากง่ายไปหายากว่าควรมีลักษณะเป็นเช่นไรให้ถูกต้อง ก่อนเป็นประเด็นสำคัญที่สุด

3. ข้อดีและข้อจำกัดของการพัฒนาแบบสอบเพื่อการวินิจฉัยโดยแนวคิดด้านพุทธปัญญาแนวคิดการพัฒนาแบบสอบเพื่อการวินัจโดยแนวคิดด้านพุทธิปัญญาที่กล่าวมานั้น ผู้เขียนเห็นว่ามีจุดดีและข้อจำกัดน่าพิจารณาดังนี้

3.1) ข้อดี

ก. นำเสนอแนวคิดการสร้างแบบสอบเพื่อการวินิจฉัยมุ่งเน้นลงลึกถึงรายละเอียดของการเรียนรู้ในส่วนย่อย (ที่เรียกว่า Attribute) ซึ่งสามารถที่จะชี้บ่งถึงสิ่งที่ผู้เรียนทำได้และทำไม่ได้อย่างลึกซึ้ง ที่จะนำไปสู่การแก้ไขพัฒนาผู้เรียนได้อย่างตรงประเด็นเป็นรายบุคคล ถือได้ว่าเป็นแนวคิดใหม่ที่ท้าทายความสามารถนักวิชาการ โดยเฉพาะการสร้างวิเคราะห์และสร้าง Q–Matrix ที่มีคุณภาพและการนำผลที่ได้ไปใช้ในการชี้ระบุความสามารถของผู้เรียนเป็นรายบุคคล

ข. นักวัดผลประเมินผลในประเทศไทยสามารถใช้แนวคิดที่มีมาใช้เป็นแนวทางสร้างแบบสอบเพื่อวินิจฉัยผลการเรียนของนักเรียนในการเรียนในระดับชั้นเรียนและในรายวิชาต่าง ๆ ได้อย่างตรงประเด็น ซึ่งจะเป็นแนวทางการแก้ไขปัญหา โดยเฉพาะการวิเคราะห์ภารกิจที่นักเรียนจะต้องเรียนรู้และจัดระบบให้ปรากฏในรูปของ Q–Matrix แล้วนำมาแก้ไขพัฒนานักเรียนมากกว่าที่จะกล่าวกันเพียงว่า นักเรียนไทยมีผลสัมฤทธิ์ทางการเรียนต่ำ

ค. เป็นโอกาสดีที่นักวิจัยและผู้ที่จัดการเรียนการสอนในระดับบัณฑิตศึกษาที่จะได้มีประเด็นใหม่ ๆ เพื่อศึกษาค้นคว้า และขยายต่อยอดประเด็นที่ควรศึกษาในรูปของหัวข้อวิทยานิพนธ์ทั้งในระดับปริญญามหาบัณฑิตและระดับปริญญาดุษฎีบัณฑิต โดยเฉพาะแนววิธีการดำเนินการสร้างและพัฒนา Q–Matrix ให้มีคุณภาพและชี้ระบุจุดดี จุดด้อยของผู้สอบได้อย่างแท้จริงในสาระการเรียนรู้ในรายวิชาต่าง ๆ ซึ่งแนวคิดและเทคนิควิธีส่วนนี้ยังคงรอการพัฒนาอยู่อีกมาก

3.2) ข้อจำกัด

การพัฒนาแบบสอบเพื่อการวินิจฉัยโดยแนวคิดด้านพุทธิปัญญาเป็นหัวข้อทางวิชาการที่ผู้ศึกษาอาจพบความยากในการศึกษาค้นคว้าและทำวิจัยต่อยอดในประเด็นต่างๆ คือ

ก. ความเข้าใจได้ยากในระดับการใช้สถิติเพื่อการวิเคราะห์หาค่าประจำข้อสอบและการขาดโปรแกรมคอมพิวเตอร์เพื่อการวิเคราะห์ข้อสอบการวิเคราะห์ตามแนวทฤษฎีนี้ใช้วิธีการวิเคราะห์หาค่าประจำข้อสอบ

ที่นักวิจัยทั่วไปไม่ค่อยคุ้นเคย และขั้นตอนการวิเคราะห์ก็ซับซ้อนเกินกว่าที่คนทั่วไปจะสามารถวิเคราะห์ได้ด้วยตนเอง ต้องอาศัยโปรแกรมคอมพิวเตอร์สำเร็จรูปช่วย

ซึ่งตัวโปรแกรมที่จะใช้ก็หาได้ไม่ง่ายนัก ทำให้การรับแนวคิดมาปรับใช้อยู่ในวงจำกัด แม้นักวิชาการบางท่าน ที่ผลิตความรู้ในโมเดลดังกล่าว ก็ให้ความเห็นไปในทิศทางเดียวกัน (de la Torre, 2009)

แต่อย่างไรก็ดี โปรแกรมเพื่อการวิเคราะห์ข้อสอบบางส่วนก็สามารถขอได้จากนักวิชาการเจ้าของโดยตรง เช่น โปรแกรมการวิเคราะห์ข้อสอบในแนว The G-DINA Model สามารถติดต่อขอได้จากเจ้าของแนวคิด (de la Torre, 2011) โปรแกรมการวิเคราะห์ค่าประจำข้อสอบในแนว The RUM Model สามารถติดต่อขอได้จากเจ้าของทฤษฎีเช่นกัน (DiBello and Stout, 2010)

ข. ความยากในการวิเคราะห์และสร้าง Q-Matrix การวิเคราะห์และการสร้าง Q-Matrix ในสาระเนื้อหาที่ต้องการสร้างข้อสอบนั้นกระทำได้ไม่ง่ายนักและคำแนะนำที่ชัดเจนในการพัฒนาเมตริกดังกล่าวยังคงมีอยู่ในระดับที่น้อยมาก และทั้งยังมีความเป็นไปได้ว่า ในสาระการเรียนรู้หนึ่งๆ วิธีการคิดของนักเรียนกับสิ่งที่วิเคราะห์ไว้ใน Q-Matrix ก็อาจจะไม่สอดรับกัน ซึ่งจะนำไปสู่การปรับแก้ Q-Matrix หรือการออกแบบ Q-Matrix ที่ยุ่งยากขึ้น แต่ในขณะเดียวกันก็เป็นแนวคิดหลักวิชาที่ท้าทายความสามารถของนักวิชาการด้วยเช่นกัน

ค. ความเป็นไปได้ในการปรับใช้แนวคิดของทฤษฎีดังกล่าวกับการเรียนรู้ในวิชาต่าง ๆ

จากการที่ผู้เขียนได้ศึกษาแนวคิดนี้จากวารสารที่ปรากฏในต่างประเทศ พบว่าเนื้อหารายวิชาที่ได้รับการกล่าวถึงและยกเป็นตัวอย่าง คือวิชาคณิตศาสตร์ในส่วนของการบวกลบเศษส่วน ซึ่งพัฒนามาจากงานของ Tatsuoka ที่กล่าวถึงแล้ว จากนั้นปรากฏวิชาในภาษาอังกฤษ และวิชาที่จัดสอบในลักษณะของแบบสอบมาตรฐานซึ่งมีผู้สอบเป็นจำนวนมาก โดยมีลักษณะเป็นการศึกษาย้อนหลัง คือ นำเอาแบบสอบที่มีอยู่ มาวิเคราะห์ย้อนว่า Q-Matrix ควรจะมีลักษณะเป็นเช่นไร การพัฒนาแบบสอบในแนวคิดทั้งหมดนี้ โดยสร้างแบบสอบขึ้นใหม่ นำไปใช้สอบ และนำผลมาวิเคราะห์เพื่อการใช้กับระดับชั้นเรียนต่างๆ ยังแทบหาไม่ได้ สภาพเช่นนี้อาจจะเกิดจากปัญหาประการหนึ่ง คือการวิเคราะห์ข้อสอบตามโมเดลต่าง ๆ ในแนวคิดด้านพุทธิปัญญา ต้องใช้กลุ่มตัวอย่างจำนวนมาก การนำเอาทฤษฎีดังกล่าวมาใช้จึงยังไม่แพร่หลาย ยังคงรอเวลาและอาศัยการวิจัยและพัฒนาในด้านการปรับใช้ในวิชาต่าง ๆ อย่างหลากหลายต่อไปอีกระยะหนึ่ง

บทสรุป

ท้ายสุดผู้เขียนมีทัศนะว่าการสร้างแบบสอบ วินิจฉัยในแนวคิดข้างต้นนี้ เป็นศาสตร์ที่เกิดขึ้นใหม่ การเติบโตยังไม่สมบูรณ์เท่าที่ควร นักวิชาการควรจะช่วยกันวิจารณ์ติชม และร่วมกันทำวิจัยพัฒนาต่อยอดความรู้ในส่วนที่ตนเองสามารถจะสามารถกระทำได้เพื่อให้ศาสตร์ดังกล่าวเจริญได้อย่างเต็มที่ ซึ่งการกระทำดังกล่าวจะเป็นการช่วยกันทำให้ข้อจำกัดต่างๆ หมดไป และผู้ใช้สามารถนำมาใช้ประโยชน์ได้อย่างหลากหลาย ณ ช่วงเวลานี้จึงเป็นโอกาสอันดีที่นักวิชาการในประเทศไทยที่มีความสามารถสูงจะได้เข้าไปมีส่วนร่วมในการพัฒนาต่อยอดความรู้ในส่วนนี้ ทั้งเพื่อความก้าวหน้าทางวิชาการและเพื่อการพัฒนาระบบการศึกษาของไทยสืบไป

เอกสารอ้างอิง

De la Torre, J. (2009), 'DINA Model and Parameter Estimation: a Didactic', **Journal of Educational and Behavioral Statistics.** Vol. 34, no. 1, pp. 115–130.

De la Torre J. (2011), 'The Generalized DINA Model Framework', **Psychometrika**. Vol. 76, no. 2, pp. 179 – 199.

De Carlo, L. T. (2011), 'On the Analysis of Fraction Subtraction Data: The DINA Model, Classification, Latent Class Sizes, and the Q–Matrix', **Applied Psychological Measurement.** Vol. 35, no. 1, pp. 8–26.

Di Bello, L.A., & Stout, W.F. (2010), **'Arpeggio Documentation and Analyst Manual'**, [Computer Software Version 3.1.003.] Applied Informative Assessment Research Enterprises.

Di Bello, L.V., Roussos, L.A. & Stout, W.F. (2007), 'Review of Cognitively Diagnosis Assessment and a Summary of Psychometric Models', **Handbook of Statistics**. Vol. 26. pp. 1–52.

Gierl, M.J.; Leighton, J.P., & Hunka, S. (2007), 'Using the Attribute Hierarchical Methodto Make Diagnostic Inferences About Examinees' Cognitive Skills', In Leighton, J.P. and Gierl, M.J. (Eds.), '**Cognitive Diagnostic Assessment for Education: Theory and Applications',** (pp. 242–274). Cambridge, UK : Cambridge University Press.

Gu, Z. (2011), '**Maximizing the Potential of Multiple–Choice Items for Cognitive Diagnostic Assessment**', Unpublished doctoral dissertation. University of Toronto. สืบค้นเมื่อวันที่ 14 สิงหาคม 2557 จาก"https://tspace.library.utoronto.ca/ bitstream /1807/31770/1/Gu_Zhimei_201111_PhD_thesis.pdf"

Haertel, E. (1984), 'An Application of Latent Class Models to Assessment Data', **Applied Psychological Measurement,** Vol. 8, no. 3, pp. 333–346.

Li, H. (2011), '**A Cognitive Diagnosis Analysis of the MELAB Reading Test**. Span Fellow Working Papers in Second or Foreign Language Assessment', Vol. 9, pp. 17– 46. สืบค้นวันที่ 4 มกราคม 2558 จาก http://www.cambridgemichigan.org/Sites/default/files/resources/SpaanPapers/Spaan v9Li.pdf."

Leighton, J.P., Gierl, M.J. and Hunka, S.M. (2004), 'The Attribute Hierarchy Method for Cognitive Assessment : A Variation on Tatsuoka's Rule–Space Approach', **Journal of Educational Measurement**. Vol. 41, no. 3, pp. 205–237.

Roussos, L.A., Templin, J.L. and Henson, R.A. (2007), 'Skills Diagnosis Using IRT–Based Latent Class Models', **Journal of Educational Measurement**. Vol. 44, no. 4, pp. 293–311.

Roussos, L.A., DiBello,L.V., Stout, W., Hartz, S.M., Henson, R.A. and Templin, J.L. (2007), 'The Fusion Models Skills Diagnosis System', In Leighton,J.P. and Gierl, M.J. (Eds.) **Cognitive diagnosis assessment for education : Theory and Applications**. (pp. 275–318). Cambridge, UK : Cambridge University Press.

Su, Y. (2013), **Cognitive Diagnostic Analysis using Hierarchically Structured Skills**. Unpublished doctoral dissertation. University of IOWA. สืบค้นเมื่อวันที่ 14 สิงหาคม 2557 จาก "http://ir.uiowa.edu/cgi/viewcontent.cgi?article=4769&context=etd""

Tatsuoka, K.K. (1983), 'Rule Space: An Approach for Dealing with Misconceptions Based on Item Response Theory', Journal of Educational Measurement. Vol. 20, no. 4, pp. 345–354.

Tatsuoka, K.K. (1985), 'A Probabilistic Model for Diagnosing Misconceptions by the Pattern Classification Approach', **Journal of Educational Statistics**. Vol. 10, no. 1, pp. 55–73.

Tatsuoka, K.K. and Tatsuoka, M.M. (1992), '**A Psychometrically Sound Cognitive Diagnosis Model: Effect of Remedial As Empirical Validity'**, Educational Testing Service, Princeton, New Jersey. สืบค้นเมื่อวันที่ 27 มกราคม 2557 จาก http://dtic.mil/dtic/tr/fulltest/u2/a255926.pdf