

## บทที่ 2

### เอกสารและงานวิจัยที่เกี่ยวข้อง

#### 2.1 แนวคิดเกี่ยวกับการวินิจฉัย

##### 2.1.1 ความหมายของการวินิจฉัย

การวินิจฉัย หมายถึง การไตร่ตรอง การใคร่ครวญหรือการตัดสินใจ (ราชบัณฑิตยสถาน, 2546) ซึ่งตรงกับภาษาอังกฤษที่ว่า diagnostic หรือ diagnosis บางครั้งอาจใช้คำว่า การประเมินวินิจฉัยหรือการตัดสินใจ โดยทั่วไปเราสามารถพบคำดังกล่าวในทางการแพทย์และทางพฤติกรรมศาสตร์ เพื่อศึกษาอาการหรือแก้ไขปรับปรุงพฤติกรรมเบี่ยงเบนนั้นๆ สำหรับทางการศึกษามีการใช้คำว่า “วินิจฉัย” อย่างแพร่หลายมากขึ้น ทั้งในบริบทของการศึกษาสำหรับบุคคลทั่วไปและบุคคลที่มีความต้องการพิเศษ เช่น ความบกพร่องทางการเรียนรู้ (learning disability)

Cohen, Swerdik และ Smith (1992) ได้ให้ความหมายไว้ว่า การวินิจฉัยเป็นการกระทำหรือกระบวนการในการจำแนกหรือกำหนดความผิดปกติ โดยการตรวจพิจารณาหรือการทดสอบ

Gregory (1992) ได้ให้ความหมายไว้ว่า การวินิจฉัยเป็นการกำหนดลักษณะที่เป็นอยู่หรือแหล่งของพฤติกรรมที่ไม่ปกติเพื่อจำแนกแบบแผนของพฤติกรรมตามระบบการวินิจฉัยที่ได้รับการยอมรับ

Gipp (1995) ได้ให้ความหมายไว้ว่า การวินิจฉัยเป็นการระบุปัญหาเฉพาะที่มีต่อความคิดหรือทักษะของนักเรียนแต่ละบุคคล

สรุปได้ว่า การวินิจฉัย หมายถึง การระบุข้อบกพร่องของพฤติกรรม ความคิดหรือสิ่งที่ผิดปกติในตัวบุคคลด้วยวิธีการทดสอบที่เหมาะสม เพื่อแก้ไขหรือปรับปรุงพฤติกรรมที่ผิดปกติ

##### 2.1.2 ความหมายของการประเมินเชิงวินิจฉัย

นอกจากคำว่า วินิจฉัย (diagnostic) แล้วยังมีอีกคำหนึ่งที่เกี่ยวข้องกับการวินิจฉัย คือการประเมินเชิงวินิจฉัย (diagnostic assessment) ซึ่งมีนักวิชาการได้ให้ความหมายไว้ดังนี้

Hopkins และ Antes (1990) ได้ให้ความหมายไว้ว่า การประเมินเชิงวินิจฉัยเป็นการตัดสินใจก่อนและจุดแข็งของนักเรียนหรือข้อบกพร่องของนักเรียนในทักษะและกระบวนการเรียนรู้

Nitko (2007) ได้ให้ความหมายว่า การประเมินเชิงวินิจฉัยเป็นการตรวจสอบข้อบกพร่องของผู้เรียนเป็นรายบุคคลด้วยเครื่องมือที่สร้างขึ้น เช่น แบบสอบวินิจฉัย ซึ่งประกอบด้วย กลุ่มข้อสอบที่เกิดจากการวิเคราะห์คำตอบของนักเรียนเป็นรายข้อ แล้วรวบรวมข้อสอบที่เป็นปัญหาในการตอบของนักเรียนจำนวนมากมาใช้ประโยชน์เพื่อการปรับปรุงการสอนของครู และหาวิธีการแก้ไขโมฆะที่คลาดเคลื่อนของผู้เรียนเป็นรายบุคคล โดยมีเป้าหมายเพื่อบ่งชี้ถึงลักษณะที่ไม่รอบรู้ของผู้เรียน พร้อมทั้งระบุสาเหตุหรือเหตุผลที่ทำให้ผู้เรียนไม่เกิดการเรียนรู้

Miller, Linn และ Gronlund (2009) ได้ให้ความหมายไว้ว่า การประเมินเชิงวินิจฉัยเป็นการระบุอุปสรรคการเรียนรู้ของนักเรียนในระหว่างที่มีการเรียนการสอน

โดยสรุป การประเมินเชิงวินิจฉัย หมายถึง การระบุจุดแข็งจุดอ่อนของนักเรียน มีเป้าหมายเพื่อหาสาเหตุที่ทำให้นักเรียนมีมีโนทัศน์ที่คลาดเคลื่อนหรือไม่รอบรู้ในเนื้อหาหรือทักษะนั้นๆ โดยใช้เครื่องมือที่สร้างขึ้น เช่น แบบสอบวินิจฉัย ซึ่งข้อมูลที่ได้จากการวินิจฉัยจะมาใช้ในการปรับปรุงการสอนของครู และแก้ไขข้อบกพร่องของนักเรียน

### 2.1.3 ประเภทของการวินิจฉัย

Underhill (1972) ได้จำแนกประเภทของการวินิจฉัยตามระดับความละเอียดในการวินิจฉัย เป็น 3 ประเภท ดังนี้

1) การวินิจฉัยแบบทั่วไป (general diagnosis) เป็นการวินิจฉัยระดับทั่วไป (general level) ซึ่งเป็นขั้นการสำรวจเพื่อทราบระดับความสามารถทั่วไปของนักเรียนทั้งรายบุคคลและรายกลุ่ม เครื่องมือส่วนใหญ่ที่ใช้เป็นแบบสอบทั้งแบบสอบวัดผลสัมฤทธิ์ที่ครูสร้างขึ้นและแบบสอบมาตรฐาน อาจใช้การวินิจฉัยโดยการเปรียบเทียบกับเกณฑ์ปกติ (norm) ซึ่งจะทำให้ทราบข้อบกพร่องในด้านต่างๆ ของนักเรียนอย่างกว้างๆ

2) การวินิจฉัยแบบวิเคราะห์ (analytical diagnosis) เป็นการวินิจฉัยระดับเฉพาะ (specific level) โดยการเก็บรวบรวมข้อมูลอย่างละเอียดเกี่ยวกับความสามารถของนักเรียนเพื่อทราบข้อบกพร่องของนักเรียนในเรื่องใดเรื่องหนึ่ง เครื่องมือส่วนใหญ่ที่ใช้เป็นแบบทดสอบที่มีเนื้อหาเฉพาะเจาะจงในเรื่องใดเรื่องหนึ่ง การวินิจฉัยแบบวิเคราะห์มักดำเนินการภายหลังจากการวินิจฉัยแบบทั่วไปเพื่อให้ได้ข้อมูลที่มีรายละเอียดที่ลึกซึ้งมากขึ้น

3) การวินิจฉัยแบบคลินิก (clinical diagnostic) การวินิจฉัยแบบคลินิกเป็นการวินิจฉัยระดับละเอียดลึกซึ้ง (intensive level) เพื่อให้เห็นถึงสาเหตุของปัญหาหรือข้อบกพร่องที่พบในตัวนักเรียนซึ่งมีความซับซ้อน ดังนั้นการใช้ข้อมูลจากแหล่งใดแหล่งหนึ่งเพียงแหล่งเดียวจึงไม่เพียงพอ จำเป็นต้องใช้ข้อมูลจากหลายแหล่งประกอบกัน เช่น การใช้ข้อมูลจากแบบทดสอบ ประกอบการสัมภาษณ์ การสังเกต การศึกษาสภาพครอบครัว เป็นต้น

### 2.1.4 รูปแบบและวิธีการที่ใช้ในการวินิจฉัยทางการศึกษา

ดวงเดือน อ่อนน่วม (2533) และศิริเดช สุชีวะ (2538) ได้จำแนกรูปแบบการวินิจฉัยทางการศึกษาไว้ 2 รูปแบบ ดังนี้

#### 1) การวินิจฉัยที่ไม่เป็นทางการ (informal diagnosis)

การวินิจฉัยที่ไม่เป็นทางการเป็นการค้นหาข้อบกพร่องของนักเรียนด้วยวิธีการง่ายๆ ไม่มีแบบแผนที่แน่นอน วิธีการที่ใช้ ได้แก่ การสังเกต การสอบถาม การสัมภาษณ์ การตรวจผลงาน/แบบฝึกหัดหรือการแสดงวิธีทำของนักเรียน เป็นต้น

## 2) การวินิจฉัยที่เป็นทางการ (formal diagnosis)

การวินิจฉัยที่เป็นทางการเป็นการหาข้อบกพร่องของนักเรียนโดยใช้วิธีการที่สร้างขึ้นมาอย่างเป็นระบบและมีแบบแผน วิธีการที่ใช้ได้แก่ การใช้แบบทดสอบเชิงวินิจฉัยหรือการวินิจฉัยจากการตอบข้อสอบ โดยใช้ดัชนีชี้วัดความผิดปกติของแบบแผนการตอบข้อสอบ การใช้วิธีการ Rule Space ของ Tatsuoka การย้อนรอยกระบวนการคิดของศิริเดช สุชีวะ การวินิจฉัยด้วยวิธีลำดับชั้นของคุณลักษณะ (attribute hierarchy method) ของ Leighton และคณะ หรือการวินิจฉัยทางพุทธิปัญญาด้วยโมเดล DINA ของ Haertel เป็นต้น

### 2.2 โมเดลวินิจฉัยทางพุทธิปัญญา: deterministic-input, noisy-and-gate model (DINA)

การประเมินเชิงวินิจฉัย (diagnostic assessment) เป็นการรวบรวมข้อมูลเพื่อตัดสินจุดเด่น จุดด้อย ปัญหา อุปสรรค รวมทั้งสาเหตุที่ทำให้เกิดปัญหาอุปสรรคเหล่านั้น (Rupp & Templin, 2008) โดยมีเป้าหมายเพื่อปรับปรุงแก้ไข ซึ่งนิยมนำมาใช้ในทางการศึกษากันอย่างแพร่หลายสำหรับผู้ที่ต้องการมีความต้องการพิเศษ เช่น ความบกพร่องทางการเรียนรู้ หรือสำหรับบุคคลทั่วไปที่เกี่ยวข้องกับการเรียนรู้โดยพยายามชี้ถึงจุดเด่น จุดด้อย ข้อบกพร่อง ปัญหาหรืออุปสรรคที่นักเรียนไม่สามารถบรรลุวัตถุประสงค์การเรียนรู้ (Gipps, 1995; Hopkins & Antes, 1990) โดยข้อมูลที่ได้จากการวินิจฉัยจะเป็นแนวทางให้ผู้สอนสามารถจัดการเรียนการสอนได้เหมาะสมกับนักเรียนหรือจัดสอนซ่อมเสริมให้กับนักเรียนที่มีปัญหาในการเรียนเรื่องนั้น เพื่อให้นักเรียนประสบความสำเร็จในการเรียนรู้และพัฒนาตนเองได้อย่างมีประสิทธิภาพ (Nichols, 2009) ในอดีตส่วนใหญ่การวินิจฉัยนักเรียนจะทำในระดับห้องเรียน แต่ในปัจจุบันนี้การวินิจฉัยได้มุ่งไปสู่การทดสอบขนาดใหญ่ (large-scale testing) หรือการสอบที่มีความเสี่ยงสูง (high-stakes testing) เนื่องจากการสอบในอดีตจะมุ่งตัดสินนักเรียนว่ามีผลสัมฤทธิ์ผ่านตามเกณฑ์ที่กำหนดไว้หรือไม่ แต่ไม่ได้ชี้แจงรายละเอียดว่านักเรียนยังมีข้อบกพร่องที่ต้องนำไปปรับปรุงในทักษะใดบ้าง ด้วยเหตุนี้ นักวัดผลจึงได้ให้ความสำคัญกับการประเมินเชิงวินิจฉัยทางพุทธิปัญญา ดังจะเห็นได้จากโมเดลวินิจฉัยทางพุทธิปัญญาที่พัฒนาขึ้นมาอย่างหลากหลาย (DiBello, Stout, & Roussos, 2007; Rupp & Templin, 2008) โดยโมเดลวินิจฉัยทางพุทธิปัญญาที่นิยมนำมาใช้กันอย่างมาก คือ โมเดล DINA (Deterministic-input, noisy-and-gate model)

โมเดล DINA เป็นโมเดลคุณลักษณะแฝง (latent trait model) ประเภทหนึ่ง ซึ่งโมเดลคุณลักษณะแฝงหรือชั้นแฝงพัฒนาขึ้นเพื่อจำแนกผู้สอบตามทักษะ (skill) เช่น ความรอบรู้ (mastery) หรือความไม่รอบรู้ (non-mastery) ในแต่ละทักษะ โดยแสดงออกมาในรูปแบบของความรอบรู้หรือความน่าจะเป็นของความรอบรู้ในแต่ละทักษะของผู้สอบ โมเดลชั้นแฝงแบ่งออกเป็น 2 ประเภท คือ compensatory (โมเดลชดเชยได้) และ non-compensatory (โมเดลชดเชยไม่ได้) ซึ่งโมเดล DINA เป็นโมเดลประเภท non-compensatory หมายถึง ตัวแปรแฝงที่มีค่าสูงกว่าจะไม่สามารถไปชดเชยตัวแปรแฝงที่มีค่าต่ำกว่าในโมเดลได้ หรือความสามารถอื่นไม่สามารถชดเชยความสามารถที่ขาดหายไปได้ เช่น ข้อสอบความเข้าใจในการอ่านวิชาภาษาอังกฤษต้องใช้ความสามารถด้านคำศัพท์และการสังเคราะห์ทั้งสองด้านในการตีความหมายของประโยค ดังนั้นถ้าขาดความสามารถด้านคำศัพท์ก็จะไม่สามารถตีความหมายของประโยคได้ กล่าวคือ ความสามารถด้านการสังเคราะห์ไม่สามารถชดเชยความสามารถด้านคำศัพท์ได้ นอกจากนี้ยังกล่าวอีกนัยหนึ่งว่า โมเดล

วินิจฉัย DINA เป็นโมเดลประเภท conjunctive หมายถึง การที่ผู้สอบสามารถตอบข้อสอบได้ถูกต้องจะเกิดขึ้นเมื่อ ผู้สอบต้องมีความรอบรู้ในทักษะทุกทักษะของข้อสอบข้อนั้น ถ้าขาดทักษะใดคุณทักษะหนึ่งไป ผู้สอบก็ควรจะตอบข้อสอบข้อนั้นผิด (Rupp Templin & Henson, 2010)

โมเดล DINA ใช้การประมาณค่าเพื่อวินิจฉัยคุณลักษณะของผู้สอบที่ใช้ในการแก้โจทย์ที่อยู่ในข้อสอบ โดยใช้กับข้อสอบที่มีการตรวจให้คะแนนแบบสองค่า (0, 1) โดยนำแบบแผนการตอบข้อสอบ (response pattern) และ Q-matrix มาวิเคราะห์ร่วมกัน ซึ่งผลการวิเคราะห์ที่ได้จะแสดงในรูปแบบโปรไฟล์ 0, 1 โดยที่ 0 หมายถึง ผู้เรียนไม่รอบรู้ในทักษะนั้น และ 1 หมายถึง ผู้เรียนรอบรู้ในทักษะนั้น

Q-matrix เป็นเมทริกซ์ที่แสดงคุณลักษณะที่ต้องการวัดของข้อสอบแต่ละข้อ ซึ่งจะแสดงเป็นตัวเลข 0 (ไม่มีคุณลักษณะในข้อนั้น) และ 1 (มีคุณลักษณะในข้อนั้น) โดยในแนวแถว (row) แทน ข้อสอบแต่ละข้อ และแนวหลัก (column) แทน คุณลักษณะที่ต้องการวัด ดังตารางที่ 2.1

ตารางที่ 2.1 Q-matrix ของข้อสอบคณิตศาสตร์

ข้อสอบ	ทักษะ			
	การบวก	การลบ	การคูณ	การหาร
1) $3+4-2 = \square$	1	1	0	0
2) $6/3 = \square$	0	0	0	1
3) $6 \times 3 - 5 = \square$	0	1	1	0
4) $7+14 = \square$	1	0	0	0

ที่มา: ดัดแปลงจาก Rupp, Templin, Henson (2010)

ฟังก์ชันการประมาณค่าของโมเดล DINA คือ

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

โดยที่  $P_j(\alpha_i)$  คือ ความน่าจะเป็นของผู้สอบคนที่  $i$  ที่ถูกจำแนกหรือวินิจฉัยอยู่ในกลุ่มที่มีโปรไฟล์  $\alpha_i$  จะตอบข้อสอบข้อที่  $j$  ได้ถูกต้อง

$\alpha_i$  คือ โปรไฟล์หรือเวกเตอร์ทักษะของผู้สอบคนที่  $i$

$s_j$  คือ พารามิเตอร์สะเพร่า หรือการที่ผู้สอบคนที่  $i$  มีความรอบรู้แต่ตอบข้อสอบข้อที่  $j$  ผิด

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

$g_j$  คือ พารามิเตอร์การเดา คือ การที่ผู้สอบคนที่  $i$  ตอบข้อสอบข้อที่  $j$  ถูกถึงแม้จะไม่รอบรู้

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

$\eta_{ij} = 1$  หมายถึง ผู้สอบคนที่  $i$  รอบรู้ทุกทักษะที่วินิจฉัยในข้อสอบข้อที่  $j$

$\eta_{ij} = 0$  หมายถึง ผู้สอบคนที่  $i$  ไม่รอบรู้อย่างน้อย 1 ทักษะที่วินิจฉัยในข้อสอบข้อที่  $j$

## 2.3 งานวิจัยที่เกี่ยวข้องกับการวินิจฉัยทางพุทธิปัญญา

de la Torre (2009) ศึกษาการใช้โมเดล DINA ในการประเมินเชิงวินิจฉัยทางพุทธิปัญญา เรื่อง การลบเศษส่วน โดยใช้ข้อมูลทฤษฎีของ Tatsuoka (1990) ซึ่งเป็นแบบทดสอบ จำนวน 15 ข้อ และนำมาเขียนเป็น Q-matrix โดยมีคุณลักษณะ จำนวน 5 คุณลักษณะ ดังนี้ คุณลักษณะที่ 1 การลบเศษส่วนที่มีตัวส่วนเท่ากัน (subtract basic fractions) คุณลักษณะที่ 2 การลดรูปและทำให้เศษส่วนอยู่ในรูปอย่างง่าย (reduce and simplify) คุณลักษณะที่ 3 การแยกจำนวนเต็มออกจากเศษส่วน (separate whole from fraction) คุณลักษณะที่ 4 การยืมตัวเลขจากจำนวนเต็ม (borrow from whole) และ คุณลักษณะที่ 5 การแปลงจำนวนเต็มให้เป็นเศษส่วน (convert whole to fraction) สำหรับ Q-matrix แสดงดังตารางต่อไปนี้

ข้อที่	คุณลักษณะ				
	1	2	3	4	5
1	1	0	0	0	0
2	1	1	1	1	0
3	1	0	0	0	0
4	1	1	1	1	1
5	0	0	1	0	0
6	1	1	1	1	0
7	1	1	1	1	0
8	1	1	0	0	0
9	1	0	1	0	0
10	1	0	1	1	1
11	1	0	1	0	0
12	1	0	1	1	0
13	1	1	1	1	0
14	1	1	1	1	1
15	1	1	1	1	0

งานวิจัยเรื่องนี้ได้เปรียบเทียบวิธีการประมาณค่าพารามิเตอร์การเดา (s) และพารามิเตอร์สะเพร่า (g) จากโมเดล DINA และโมเดล HO-DINA (higher-order DINA) โดยที่โมเดล DINA ใช้วิธีการประมาณค่าพารามิเตอร์แบบ Expectation-Maximization (EM) ส่วนโมเดล HO-DINA ใช้วิธี Markov Chain Monte Carlo algorithms (MCMC) ผลการวิจัยพบว่า การประมาณค่าทั้งสองวิธีให้ค่าพารามิเตอร์ใกล้เคียงกันมาก

ยกเว้นข้อที่ 5 ที่ค่า  $g$  แตกต่างกัน และโมเดล HO-DINA ให้ผลการประมาณค่าที่สมเหตุสมผลกับข้อมูลจริงมากกว่าโมเดล DINA

Jurich Gierl, Alves และ Majeau (2010) ได้ศึกษาการใช้วิธีลำดับชั้นของคุณลักษณะ (attribute hierarchy method: AHM) เพื่อสร้างการสรุปอ้างอิงเชิงวินิจฉัยเกี่ยวกับความรู้ (knowledge) และทักษะ (skills) ในวิชาคณิตศาสตร์ มีวัตถุประสงค์ของการวิจัย เพื่ออธิบายและประเมินวิธีลำดับชั้นของคุณลักษณะ (AHM) ที่ประยุกต์ใช้กับการประเมินเชิงวินิจฉัยทางพุทธิปัญญาในวิชาคณิตศาสตร์ ตัวอย่างที่ใช้ในการศึกษาครั้งนี้แบ่งเป็น 2 กลุ่ม คือ นักเรียนประถมศึกษาปีที่ 3 จำนวน 3,338 คน และนักเรียนประถมศึกษาปีที่ 6 จำนวน 6,184 คน โดยนักเรียนทุกคนมีความสมัครใจที่จะเข้าร่วมการทดสอบ เครื่องมือที่ใช้ในการวิจัยคือข้อสอบวินิจฉัยแบบปรนัย 4 ตัวเลือกที่สร้างโดยผู้เชี่ยวชาญด้านเนื้อหาวิชาคณิตศาสตร์ และใช้ระบบการทดสอบแบบคอมพิวเตอร์ออนไลน์ (online computer-based testing) การดำเนินการวิจัยแบ่งเป็น 2 ระยะ คือ ระยะที่ 1 การสร้างโมเดลทางพุทธิปัญญา (cognitive model representation) ที่พัฒนาขึ้นแบ่งเป็น 2 โมเดล ได้แก่ โมเดลทางพุทธิปัญญาที่ใช้กับระดับชั้นประถมศึกษาปีที่ 3 และ 6 ซึ่งประกอบด้วยเนื้อหาวิชาคณิตศาสตร์จำนวน 4 เรื่อง คือ จำนวน (number) แบบรูป และความสัมพันธ์ (pattern and relations) รูปร่างและปริภูมิ (shape and space) และ สถิติและความน่าจะเป็น (statistics and probability) โดยทั้ง 2 โมเดลได้ใช้โครงสร้างของลำดับชั้นแบบเชิงเส้น (linear heirarchy) และระยะที่ 2 การพัฒนาข้อสอบ (item development) สำหรับระดับชั้นประถมศึกษาปีที่ 3 จะมีข้อสอบคู่ขนาน (parallel items) จำนวน 3 ข้อในแต่ละคุณลักษณะ และระดับชั้นประถมศึกษาปีที่ 6 จะมีข้อสอบคู่ขนาน (parallel items) จำนวน 2-3 ข้อในแต่ละคุณลักษณะ ผลการวิจัยสรุปได้ว่า

1) ความสอดคล้องของข้อมูลกับโมเดล (model-data fit) เป็นการตรวจสอบความสอดคล้องระหว่างรูปแบบการตอบข้อสอบที่คาดหวัง (expected response patterns) กับรูปแบบการตอบข้อสอบจริง (actual item response patterns) ด้วยการใช้ดัชนี  $HCI$ , ซึ่งเป็นดัชนีที่ใช้วัดระดับของรูปแบบการตอบข้อสอบสังเกตได้ (observed examinee response pattern) ที่สอดคล้องกับลำดับชั้นของคุณลักษณะ สำหรับโมเดลทางพุทธิปัญญาในระดับชั้นประถมศึกษาปีที่ 3 มีค่าเท่ากับ 0.66 และโมเดลทางพุทธิปัญญาในระดับชั้นประถมศึกษาปีที่ 6 มีค่าเท่ากับ 0.67 ซึ่งทั้งสองโมเดลนี้มีความสอดคล้องของข้อมูลกับโมเดลอยู่ในระดับปานกลาง

2) การให้คะแนนโดยใช้คุณลักษณะเป็นฐาน (attribute-based scoring) และการรายงานเชิงวินิจฉัย (diagnostic reporting) ค่าเฉลี่ย (mean) ของระดับความน่าจะเป็นของคุณลักษณะ (attribute probability level) ในระดับชั้นประถมศึกษาปีที่ 3 มีค่าอยู่ระหว่าง 0.66-0.97 ระดับชั้นประถมศึกษาปีที่ 6 มีค่าอยู่ระหว่าง 0.47-0.98 ส่วนค่าความน่าจะเป็นของคุณลักษณะ (attribute probability) จะแสดงการลู่เข้า (convergence) และการจำแนก (discriminant) ในระดับชั้นประถมศึกษาปีที่ 3 การลู่เข้ามีค่าสัมประสิทธิ์สหสัมพันธ์ต่ำสุดเท่ากับ 0.79 และค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดเท่ากับ 0.96 การลู่ออกมีค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดคือ 0.79 ไปจนถึงต่ำสุดคือ 0.38 โดยเป็นความสัมพันธ์ระหว่าง A1 ถึง A8 ซึ่งมีค่าสัมประสิทธิ์สหสัมพันธ์ลดลงตามลำดับ และระดับชั้นประถมศึกษาปีที่ 6 การลู่เข้ามีค่าสัมประสิทธิ์สหสัมพันธ์ต่ำสุดเท่ากับ

0.69 และค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดเท่ากับ 0.82 การลู่ออกมีค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดคือ 0.65 ไปจนถึงต่ำสุดคือ 0.11 โดยเป็นความสัมพันธ์ระหว่าง A1 ถึง A5 ซึ่งมีค่าสัมประสิทธิ์สหสัมพันธ์ลดลงตามลำดับ

สำหรับการรายงานคะแนนเชิงวินิจฉัยเป็นรายบุคคลจะรายงานด้วยค่าความน่าจะเป็นของคุณลักษณะ (attribute probability values) โดยวิธีลำดับขั้นของคุณลักษณะจะให้ข้อมูลที่เป็นรายละเอียดเกี่ยวกับคุณลักษณะที่ใช้วัดในแบบสอบและระดับความรู้ของคุณลักษณะทางพุทธิปัญญาของผู้สอบ ซึ่งข้อมูลเชิงวินิจฉัยที่ได้นี้จะเกี่ยวข้องโดยตรงกับการบรรยายคุณลักษณะรายบุคคลของนักเรียน และนำไปสู่การตัดสินใจในการจัดการเรียนการสอนที่เหมาะสมกับนักเรียน รวมทั้งเสนอเกณฑ์สำหรับการพิจารณาระดับความสามารถของคุณลักษณะ ได้แก่ ระดับความน่าจะเป็นของคุณลักษณะต่ำกว่า 0.50 หมายถึง “ควรปรับปรุง” ระดับความน่าจะเป็นของคุณลักษณะอยู่ระหว่าง 0.50-0.80 หมายถึง “ยอมรับได้/พอใช้” ระดับความน่าจะเป็นของคุณลักษณะสูงกว่า 0.80 หมายถึง “ดีเยี่ยม” และการประมาณค่าความน่าจะเป็นของคุณลักษณะของผู้สอบที่เป็นข้อมูลจริง

3) ความเที่ยงของคุณลักษณะ (attribute reliability) สำหรับระดับชั้นประถมศึกษาปีที่ 3 ค่าความเที่ยงอยู่ระหว่าง 0.72 (คุณลักษณะ A8) ถึง 0.91 (คุณลักษณะ A1) สำหรับระดับชั้นประถมศึกษาปีที่ 6 ค่าความเที่ยงอยู่ระหว่าง 0.47 (คุณลักษณะ A5) ถึง 0.72 (คุณลักษณะ A1)

ปรารณา พลอภิชชาติ (2556) ได้พัฒนาคู่มือการสร้างแบบสอบวินิจฉัยการแก้โจทย์ปัญหาคณิตศาสตร์สำหรับครูประถมศึกษาโดยใช้โมเดลข้อสอบและวิธีลำดับขั้นของคุณลักษณะ การวิจัยนี้มีวัตถุประสงค์ 1) เพื่อวิเคราะห์ข้อบกพร่องของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่พบในการแก้โจทย์ปัญหาคณิตศาสตร์ 2) เพื่อสร้างแบบสอบวินิจฉัยโจทย์ปัญหาคณิตศาสตร์ใช้โมเดลข้อสอบและวิธีลำดับขั้นของคุณลักษณะ 3) เพื่อวิเคราะห์จุดแข็งและจุดอ่อนของนักเรียนชั้นประถมศึกษาปีที่ 6 ในการแก้โจทย์ปัญหาคณิตศาสตร์ เรื่อง การบวกและการลบเศษส่วน และ 4) เพื่อสร้างและตรวจสอบคุณภาพของคู่มือการสร้างแบบสอบวินิจฉัยการแก้โจทย์ปัญหาคณิตศาสตร์สำหรับครูประถมศึกษาโดยใช้โมเดลข้อสอบและวิธีลำดับขั้นของคุณลักษณะ ตัวอย่างวิจัยแบ่งเป็น 2 กลุ่ม ได้แก่ 1) นักเรียนชั้นประถมศึกษาปีที่ 6 ปีที่ 6 ที่ได้จากการสุ่มอย่างง่าย จำนวน 1,252 คน และ 2) ครูคณิตศาสตร์ระดับชั้นประถมศึกษาที่ได้จากการเลือกแบบเจาะจงจำนวน 13 คน เครื่องมือที่ใช้เก็บรวบรวมข้อมูล ประกอบด้วย แบบบันทึกการสัมภาษณ์ แบบสอบวินิจฉัยการแก้โจทย์ปัญหาคณิตศาสตร์ วิเคราะห์ข้อมูลด้วยสถิติเชิงบรรยาย การวิเคราะห์เนื้อหา และคำนวณคะแนนเชิงวินิจฉัยโดยการประยุกต์ใช้ทฤษฎีเบย์ การตรวจสอบคุณภาพของแบบสอบวินิจฉัยใช้การวิเคราะห์ค่าพหามิเตอร์ความยากและอำนาจจำแนกโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบ 2 พหามิเตอร์ ความตรงเชิงเนื้อหา ความเที่ยงความเที่ยงแบบความสอดคล้องภายในด้วยวิธีของฮอยท์ และความเที่ยงระหว่างผู้ตรวจ ผลการวิจัยสรุปได้ว่า

1. นักเรียนมีข้อบกพร่องที่พบในการแก้โจทย์ปัญหาคณิตศาสตร์ คือ 1) ความเข้าใจในการอ่าน ไม่สามารถระบุความหมายของ “สิ่งที่โจทย์ต้องการให้ตอบ” และ “สิ่งที่โจทย์กำหนดให้” 2) การตีความหมาย

ของคำสำคัญในโจทย์ที่แปลงเป็นการดำเนินการทางคณิตศาสตร์ และ 3) ปัญหาการคำนวณ โดยหน่วยการเรียนรู้ที่นักเรียนบกพร่องมากที่สุด คือ เศษส่วน

2. การสร้างแบบสอบวินิจฉัยด้วยวิธีลำดับชั้นของคุณลักษณะ (attribute hierarchy method) และโมเดลข้อสอบ (item model) มีขั้นตอนหลัก ดังนี้ 1) กำหนดและเรียงลำดับคุณลักษณะ 2) สร้างโมเดลลำดับชั้นของคุณลักษณะ (attribute hierarchy model) 3) สร้างเมทริกซ์กำหนดคุณลักษณะของข้อสอบ (Q-matrix) 4) สร้างแบบสอบตาม Q-matrix และ 5) คำนวณคะแนนเชิงวินิจฉัยเป็นรายคุณลักษณะ และ 6) สร้างแบบสอบวินิจฉัยโดยประยุกต์ใช้โมเดลข้อสอบ คุณภาพของแบบสอบวินิจฉัยที่สร้างขึ้นมีค่าความยากระหว่าง (-2.14) - 1.06 และค่าอำนาจจำแนกระหว่าง 0.77 - 3.07 ความตรงเชิงเนื้อหาเท่ากับ 1.00 ค่าความเที่ยงแบบความสอดคล้องภายในด้วยวิธีของฮอยท์เท่ากับ 0.84 และค่าความเที่ยงระหว่างผู้ตรวจเท่ากับ 0.90 สำหรับคุณลักษณะที่ใช้ในการวินิจฉัยการแก้โจทย์ปัญหาคณิตศาสตร์ เรื่อง การบวกลบเศษส่วนและจำนวนคละ มีทั้งหมด 8 คุณลักษณะ ได้แก่ คุณลักษณะที่ 1 (A1) ความเข้าใจในการอ่านโจทย์ปัญหาเศษส่วน คุณลักษณะที่ 2 (A2) การตีความหมายของคำสำคัญในโจทย์ปัญหาเศษส่วน คุณลักษณะที่ 3 (A3) การบวกเศษส่วนที่มีตัวส่วนเท่ากัน คุณลักษณะที่ 4 (A4) การลบเศษส่วนที่มีตัวส่วนเท่ากัน คุณลักษณะที่ 5 (A5) การบวกเศษส่วนที่มีตัวส่วนไม่เท่ากัน คุณลักษณะที่ 6 (A6) การลบเศษส่วนที่มีตัวส่วนไม่เท่ากัน คุณลักษณะที่ 7 (A7) การบวกจำนวนคละที่มีตัวส่วนไม่เท่ากัน และคุณลักษณะที่ 8 (A8) การลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน และสามารถนำมาเขียนเป็นโมเดลลำดับชั้นของคุณลักษณะได้ดังนี้





3. ผลการวิจัยข้อบกพร่องส่วนใหญ่พบว่า นักเรียนมีความรอบรู้อย่างชัดเจนในคุณลักษณะด้านการอ่านและการตีความคำสำคัญในโจทย์ปัญหา แต่ขาดความรอบรู้ในคุณลักษณะด้านการบวกจำนวนคละและคุณลักษณะด้านการลบจำนวนคละ

4. ผลการใช้คู่มือพบว่า ครูสามารถสร้างแบบสอบวินิจฉัยตามวิธีลำดับขั้นของคุณลักษณะและโมเดลข้อสอบได้ แต่ต้องใช้เวลาในการศึกษาคู่มือนาน เนื่องจากไม่คุ้นเคยกับวิธีการสร้างแบบสอบวินิจฉัยแนวใหม่และศัพท์เทคนิค เช่น คุณลักษณะ Q-matrix

สุปราณี บุระ (2557) ได้พัฒนาแบบสอบวินิจฉัยทางพุทธิปัญญา เรื่อง การดำเนินการเลขคณิตพื้นฐาน โดยใช้คอมพิวเตอร์: การประยุกต์โมเดลดีไอเอเอ็นเอ มีวัตถุประสงค์เพื่อพัฒนาแบบสอบเชิงวินิจฉัยทางพุทธิปัญญา เรื่อง การดำเนินการเลขคณิตพื้นฐาน ตรวจสอบคุณภาพของแบบสอบ และพัฒนาเป็นโปรแกรมการทดสอบ โดยประยุกต์ใช้โมเดลดีไอเอเอ็นเอ วิธีดำเนินการวิจัยแบ่งเป็น 3 ระยะ คือ ระยะที่ 1 เป็นการสร้างแบบสอบด้วยการกำหนดโมเดลพุทธิปัญญา เรื่องการดำเนินการเลขคณิตพื้นฐาน และตรวจสอบความถูกต้องและเหมาะสมกับผู้เชี่ยวชาญ จากนั้นสร้างแบบสอบตามแผนผังการออกข้อสอบจากโมเดลพุทธิปัญญาที่กำหนดขึ้น ระยะที่ 2 เป็นการตรวจสอบคุณภาพของแบบสอบ ด้วยการนำแบบสอบที่พัฒนาขึ้นไปให้ผู้เชี่ยวชาญพิจารณาและทดลองใช้กับนักเรียนระดับชั้นประถมศึกษาปีที่ 3 – 6 เพื่อตรวจสอบคุณภาพของแบบสอบรายข้อด้วยโมเดลดีไอเอเอ็นเอ และตรวจสอบคุณภาพของแบบสอบทั้งฉบับ ประกอบด้วย ความเที่ยง (Reliability) ความตรงเชิงเนื้อหา (Content Validity) และความตรงตามสภาพ (Concurrent Validity) และระยะที่ 3 เป็นการพัฒนาโปรแกรมการทดสอบ โดยวิเคราะห์ผลการวิจัยด้วยโมเดลดีไอเอเอ็นเอ และตรวจสอบความถูกต้องและเหมาะสมของโปรแกรมการทดสอบกับผู้เชี่ยวชาญ จากนั้นทดลองใช้โปรแกรมการทดสอบกับครูผู้สอนวิชาคณิตศาสตร์ระดับประถมศึกษาและนักเรียนระดับชั้นประถมศึกษาปีที่ 3 – 6 เพื่อสอบถามความคิดเห็นต่อการใช้งาน และปรับปรุงแก้ไขให้สมบูรณ์ ผลการวิจัยสรุปได้ดังนี้ 1) แบบสอบเชิงวินิจฉัยทางพุทธิปัญญาเรื่องการดำเนินการเลขคณิตพื้นฐาน ประกอบด้วย แบบสอบทั้งสิ้น 8 ฉบับ ได้แก่ แบบสอบเชิงวินิจฉัยทางพุทธิปัญญาเรื่องการบวก การลบ การคูณ และการหารจำนวนนับ ระดับที่ 1 และระดับที่ 2 ที่พัฒนาจากโมเดลพุทธิปัญญา ซึ่งเป็นโครงสร้างคุณลักษณะของความรู้และทักษะที่นักเรียนใช้ในการแก้ปัญหาโจทย์คณิตศาสตร์ เรื่อง การดำเนินการเลขคณิตพื้นฐานได้อย่างถูกต้อง โดยจัดเรียงตามลักษณะความสัมพันธ์เชิงลำดับขั้นการเรียนรู้ จากคุณลักษณะขั้นพื้นฐานไปยังคุณลักษณะที่สูงกว่า เป็นข้อสอบแบบเติมคำตอบหลายคำตอบในแต่ละข้อ ถ้าตอบถูกทุกคำตอบในแต่ละข้อได้ 1 คะแนน แต่ถ้าตอบผิดแม้คำตอบเดียวในแต่ละข้อได้ 0 คะแนน 2) การตรวจสอบคุณภาพของแบบสอบรายข้อ มีค่าพารามิเตอร์การเดาข้อสอบถูก ( $g_j$ ) อยู่ระหว่าง 0.000 – 0.191 ค่าพารามิเตอร์ความสะเพร่า ( $s_j$ ) อยู่ระหว่าง 0.000 – 0.200 และดัชนีอำนาจจำแนกรายข้อ (Item Discrimination Index) อยู่ระหว่าง 0.631 – 1.000 ในส่วนของการตรวจสอบคุณภาพของแบบสอบทั้งฉบับ มีค่าความเที่ยงด้วยการคำนวณจากวิธีของลิวิตัน (Livingston Method) อยู่ระหว่าง 0.873 – 0.986 และวิธีของโลเวท (Lovett Method) อยู่ระหว่าง 0.816 – 0.985 ความตรงเชิงเนื้อหา แต่ละฉบับมีค่า IOC อยู่ระหว่าง 0.80 – 1.00 และความตรงตามสภาพ แต่ละฉบับมีค่าสัดส่วนของความสอดคล้องและค่าสถิติคัป

ปา (K) ระหว่างผลการวินิจฉัยคุณลักษณะทางพุทธิปัญญาด้วยโมเดลตีโอเอ็นเอและผลการวินิจฉัยด้วยเทคนิคการคิดออกเสียง อยู่ระหว่าง 0.758 – 0.882 และ 0.569 – 0.824 ตามลำดับ 3) โปรแกรมการทดสอบเชิงวินิจฉัยทางพุทธิปัญญาเรื่องการดำเนินการเลขคณิตพื้นฐาน สร้างด้วยโปรแกรม Microsoft Visual Studio.NET 2010 วิเคราะห์ผลการวินิจฉัยด้วยโมเดลตีโอเอ็นเอ โดยใช้โปรแกรม R และรายงานผลการวินิจฉัยด้วยโปรแกรม Report Viewer 2010 โดยออกแบบให้มีการทำงาน 2 ส่วน คือ (1) โปรแกรมการทดสอบสำหรับครู เป็นโปรแกรมควบคุมการทดสอบและวิเคราะห์ผลการวินิจฉัย และ (2) โปรแกรมการทดสอบสำหรับนักเรียน เป็นโปรแกรมแบบสอบทั้ง 8 ฉบับ สำหรับผลการประเมินจากผู้เชี่ยวชาญโดยเฉลี่ยในเรื่องของการทำงาน การออกแบบ ลักษณะทั่วไป และคู่มือการใช้อยู่ในระดับดีมาก ในส่วนของการทดลองใช้ครูและนักเรียนมีความคิดเห็นต่อการใช้โปรแกรม คู่มือการใช้ การออกแบบ และความรู้สึกต่อโปรแกรม อยู่ในระดับมากถึงระดับมากที่สุด

สุมาลี มีสกุล (2558) ได้ทำการวิจัยเรื่อง การประเมินเพื่อวินิจฉัยความสามารถในการแก้ปัญหาเรื่องสมการเชิงเส้นตัวแปรเดียวของนักเรียนระดับชั้นมัธยมศึกษาตอนต้น โดยการประยุกต์ใช้โมเดลการวินิจฉัยเชิงจำแนก โดยมีวัตถุประสงค์การวิจัยเพื่อ 1) พัฒนาแบบทดสอบสำหรับใช้ในการประเมินเพื่อวินิจฉัยความสามารถของนักเรียนระดับชั้นมัธยมศึกษาตอนต้นในการแก้ปัญหา เรื่อง สมการเชิงเส้นตัวแปรเดียว 2) วินิจฉัยความสามารถของนักเรียนระดับชั้นมัธยมศึกษาตอนต้นในการแก้ปัญหา เรื่อง สมการเชิงเส้นตัวแปรเดียว และ 3) ศึกษาและเปรียบเทียบประสิทธิภาพของโมเดลที่นำมาใช้ในการวินิจฉัยความสามารถในการแก้ปัญหา เรื่อง สมการเชิงเส้นตัวแปรเดียวของนักเรียนระดับชั้นมัธยมศึกษาตอนต้นที่มีระดับชั้น เพศ และแผนการเรียนแตกต่างกัน กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ นักเรียนระดับชั้นมัธยมศึกษาตอนต้นที่กำลังศึกษาในภาคเรียนที่ 2 ปีการศึกษา 2557 ได้มาจากการสุ่มตัวอย่างแบบหลายขั้นตอน (multi-stage random sampling) จำนวน 1,214 คน และวิเคราะห์ข้อมูลในเชิงการวินิจฉัยโดยใช้โมเดลการวินิจฉัยเชิงจำแนก DINA model และ G-DINA model โดยกำหนดคุณลักษณะที่นำมาสร้าง Q-matrix จำนวน 5 คุณลักษณะ ได้แก่ คุณลักษณะที่ 1 (S1) การใช้สมบัติการเท่ากันในระบบจำนวนจริง เพื่อหาคำตอบของสมการในรูป  $ax + b = c$  เมื่อ  $a$ ,  $b$  และ  $c$  เป็นจำนวนจริงใดๆ คุณลักษณะที่ 2 (S2) การจัดรูปสมการให้อยู่ในรูป  $ax + b = c$  เมื่อ  $a$ ,  $b$  และ  $c$  เป็นจำนวนจริงใดๆ โดยใช้สมบัติการแจกแจง คุณลักษณะที่ 3 (S3) การจัดรูปสมการให้อยู่ในรูป  $ax + b = c$  เมื่อ  $a$ ,  $b$  และ  $c$  เป็นจำนวนจริงใดๆ โดยใช้การรวมพจน์ที่คล้ายกัน คุณลักษณะที่ 4 (S4) การจัดรูปสมการที่อยู่ในรูปเศษส่วนให้อยู่ในรูป  $ax + b = c$  เมื่อ  $a$ ,  $b$  และ  $c$  เป็นจำนวนจริงใดๆ และคุณลักษณะที่ 5 (S5) การแปลงโจทย์ปัญหาให้อยู่ในรูปของสมการเพื่อนหาคำตอบ ผลการวิจัยสรุปได้ว่า

1. ความสามารถซึ่งเป็นคุณลักษณะที่นักเรียนใช้ในการแก้ปัญหาเรื่องสมการเชิงเส้นตัวแปรเดียวมีทั้งหมดจำนวน 5 คุณลักษณะ ซึ่งสามารถนำมาสร้างแบบแผนการคิดที่ผู้เรียนใช้ในการแก้ปัญหาได้ทั้งหมด 16 แบบแผน และนำมาสร้างข้อคำถามในแบบทดสอบวัดความสามารถในการแก้ปัญหา เรื่อง สมการเชิงเส้นตัวแปรเดียวตามวิธีการพัฒนาแบบทดสอบมาตรฐานเพื่อวินิจฉัยทางพุทธิปัญญาของ Embretson (Embretson's Cognitive Design System Approach: CDS framework) ซึ่งเป็นแบบทดสอบแบบปรนัย

4 ตัวเลือกที่มีการตรวจให้คะแนนแบบสองค่าได้ทั้งหมดจำนวน 32 ข้อ ที่มีค่าเฉลี่ยของค่าอำนาจจำแนกในเชิงการวินิจฉัยของข้อคำถามเท่ากับ 0.5641 ค่าความเชื่อมั่นของแบบทดสอบ (KR-20) เท่ากับ 0.807 และแบบทดสอบให้ค่าสารสนเทศที่ได้จากการสอบสูงสุดเมื่อนำมาใช้กับกลุ่มผู้สอบที่มีระดับความสามารถ ( $\theta$ ) เท่ากับ 1.25

2. นักเรียนระดับชั้นมัธยมศึกษาตอนต้นมีความรอบรู้ในความสามารถซึ่งเป็นคุณลักษณะที่ใช้ในการแก้ปัญหาเรื่องสมการเชิงเส้นตัวแปรเดียวเรียงลำดับจากมากไปน้อย คือ การใช้สมบัติการเท่ากันในระบบจำนวนจริงเพื่อหาคำตอบของสมการ การจัดรูปสมการโดยใช้สมบัติการแจกแจง การจัดรูปสมการโดยใช้การรวมพจน์ที่คล้ายกัน การจัดรูปสมการที่อยู่ในรูปเศษส่วน และการแปลงโจทย์ปัญหาให้อยู่ในรูปของสมการเพื่อหาคำตอบ ส่วนผลการประเมินเพื่อวินิจฉัยจำแนกตามระดับชั้น พบว่า นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 มีความรอบรู้ในคุณลักษณะที่ใช้ในการแก้ปัญหาสูงกว่า นักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 และ 2 ทุกคุณลักษณะ ผลการประเมินเพื่อวินิจฉัยจำแนกตามเพศ พบว่า นักเรียนชายมีความรอบรู้ในคุณลักษณะที่ใช้ในการแก้ปัญหาสูงกว่านักเรียนหญิง จำนวน 2 คุณลักษณะ และผลการประเมินเพื่อวินิจฉัยจำแนกตามแผนการเรียน พบว่า นักเรียนแผนการเรียนวิทยาศาสตร์-คณิตศาสตร์มีความรอบรู้ในคุณลักษณะที่ใช้ในการแก้ปัญหาสูงกว่านักเรียนแผนการเรียนทั่วไปทุกคุณลักษณะ

3. G-DINA model และ DINA model ให้ประสิทธิภาพในการวิเคราะห์ข้อมูลดีกว่า เมื่อนำมาใช้วิเคราะห์ข้อมูลกับกลุ่มนักเรียนที่มีความรอบรู้ในคุณลักษณะที่ใช้ในการแก้ปัญหามากกว่า

สังวรณ ังตกระโทก และคณะ (2560) ได้ทำการวิจัยเรื่อง การวินิจฉัยผลการทดสอบระดับชาติ และข้อเสนอเชิงนโยบายในการใช้ผลการทดสอบระดับชาติเพื่อพัฒนาผู้เรียน โดยมีวัตถุประสงค์เพื่อ 1) วิเคราะห์คุณภาพข้อสอบระดับชาติขั้นพื้นฐานย้อนหลัง 5 ปี และจัดทำแนวทางการพัฒนาคุณภาพข้อสอบ 2) วินิจฉัยผลการทดสอบระดับชาติ จำแนกตามสาระและมาตรฐานการเรียนรู้ของแต่ละวิชา ภูมิหลังของโรงเรียน 3) พัฒนารูปแบบการรายงานผลการวินิจฉัยผลการทดสอบระดับชาติแก่สถานศึกษาและหน่วยงานที่เกี่ยวข้อง และ 4) จัดทำข้อเสนอแนะเชิงนโยบายสำหรับสถานศึกษาและหน่วยงานที่เกี่ยวข้องในการใช้ผลการทดสอบระดับชาติเพื่อพัฒนาผู้เรียน การวิจัยแบ่งออกเป็น 4 ระยะตามวัตถุประสงค์การวิจัย ประกอบด้วย ระยะที่ 1 วิเคราะห์คุณภาพข้อสอบระดับชาติขั้นพื้นฐานปี พ.ศ. 2554 - 2558 ด้วยทฤษฎีการทดสอบแบบดั้งเดิม และทฤษฎีการตอบข้อสอบ และจัดทำแนวทางการพัฒนาคุณภาพข้อสอบจากผลการวิเคราะห์ข้อสอบ ระยะที่ 2 วินิจฉัยผลการทดสอบระดับชาติ จำแนกตามสาระและมาตรฐานการเรียนรู้ของแต่ละวิชาด้วยโมเดลวินิจฉัยเชิงพุทธิปัญญาด้วยโมเดล DINA และจำแนกผลการวินิจฉัยตามขนาดโรงเรียน และภูมิภาคของโรงเรียน ระยะที่ 3 พัฒนารูปแบบการรายงานผลการวินิจฉัยผลการทดสอบระดับชาติแก่สถานศึกษาและหน่วยงานที่เกี่ยวข้องตามแนวทางการรายงานผลการวินิจฉัยที่เน้นการให้ข้อมูลย้อนกลับเกี่ยวกับจุดอ่อนและจุดแข็งของนักเรียน ระยะที่ 4 จัดทำข้อเสนอแนะเชิงนโยบายสำหรับสถานศึกษาและหน่วยงานที่เกี่ยวข้องในการใช้ผลการทดสอบระดับชาติเพื่อพัฒนาผู้เรียน โดยนำข้อเสนอแนะจากผู้ที่มีส่วนได้เสียและผลการวินิจฉัยที่แสดงระดับความรู้ของนักเรียนที่ยังไม่บรรลุมาตรฐานการเรียนรู้มาใช้กำหนดข้อเสนอแนะในการพัฒนาผู้เรียน ผลการวิจัยพบว่า

1. ข้อสอบโอเน็ตส่วนใหญ่มีค่าสถิติข้อสอบผ่านเกณฑ์ด้านค่าความยาก อำนาจจำแนก โดยเฉพาะค่าการเดาจากการวิเคราะห์ด้วยโมเดล IRT นั้นมีค่าค่อนข้างต่ำ หรือนักเรียนเดาไม่ค่อยถูก ค่าความเที่ยงส่วนใหญ่อยู่ระหว่าง 0.6-0.8 ซึ่งถือว่าอยู่ในระดับที่ยอมรับได้ แต่บางปีมีค่าความเที่ยงต่ำ เช่น วิชาภาษาไทย ระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2557 มีค่าความเที่ยง 0.37 ส่วนระดับชั้นประถมศึกษาปีที่ 6 มีค่าความเที่ยงเท่ากับ 0.51 ซึ่งเป็นแบบทดสอบที่มีจำนวนข้อน้อย ข้อสอบบางวิชา เช่น วิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 มีข้อสอบจำนวนน้อย ทำให้อาจมีประเด็นสงสัยเรื่องความตรงและความสอดคล้องกับมาตรฐาน แนวทางการพัฒนาข้อสอบในประเด็นนี้ ควรเพิ่มจำนวนข้อสอบให้มีประมาณ 7 ข้อต่อมาตรฐานการเรียนรู้ และควรมีการตรวจสอบว่าข้อสอบวัดได้ตรงตามมาตรฐานหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานหรือไม่ โดยนำวิธีการตรวจสอบการวัดความสอดคล้องของแบบทดสอบกับมาตรฐาน (alignment) มาใช้

2. นักเรียนส่วนใหญ่ (ประมาณ 60%) ยังมีความรู้ไม่ผ่านสาระในมาตรฐานหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน มีนักเรียนจำนวนเพียงคือ ประมาณ 10,000 คน จากนักเรียนที่เข้าสอบประมาณ 500,000 – 700,000 คน ที่มีความสามารถผ่านทุกสาระในมาตรฐานหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน ผลการวินิจฉัยพบว่า นักเรียนโรงเรียนขนาดกลางและขนาดเล็กไม่ผ่านการประเมินมากที่สุด และเมื่อวิเคราะห์เปรียบเทียบเป็นรายภาคพบว่า โรงเรียนในภาคตะวันออกเฉียงเหนือ และภาคใต้ไม่ผ่านมากที่สุด

3. รูปแบบการรายงานผลการวินิจฉัยผลการทดสอบระดับชาติแก่สถานศึกษาและหน่วยงานที่เกี่ยวข้องควรเป็นรูปแบบการรายงานผลผ่านระบบออนไลน์และเครือข่ายคอมพิวเตอร์ มีข้อมูลเกี่ยวกับจุดอ่อนจุดแข็งของผู้เรียน ผู้ที่เกี่ยวข้องสามารถเข้าถึงการรายงานผลโดยให้รหัสผ่าน หรือหมายเลขประจำตัวผู้เข้าสอบ โดยองค์ประกอบของการรายงานประกอบด้วย 1) ผลการตอบรายข้อของผู้เข้าสอบ 2) คะแนนสอบที่ได้ (คะแนนดิบ) 3) ผลการวิเคราะห์การวินิจฉัยความสามารถของผู้สอบในแต่ละมาตรฐาน และ 4) ข้อสรุปผลการวินิจฉัย โดยรายงานสาระที่เป็นจุดแข็ง และจุดอ่อนของนักเรียน ระบบการรายงานผลการวินิจฉัยควรเป็นการรายงานผ่านอิเล็กทรอนิกส์เพราะมีข้อมูลมาก ไม่สะดวกในการรายงานด้วยรูปแบบกระดาษ ผู้ใช้สามารถเข้าสู่ระบบทางเครือข่ายอินเทอร์เน็ตและโทรศัพท์มือถือได้อย่างสะดวก

4. ผลการวินิจฉัยที่พบว่านักเรียนจำนวนมากยังไม่ผ่าน ผลการวินิจฉัยมีข้อเสนอแนะให้มีการปรับปรุงการเรียนการสอนในสาระต่อไปนี้

4.1 วิชาภาษาไทย ระดับชั้นประถมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ วรรณคดีและวรรณกรรม หลักการใช้ภาษา การฟัง การดู การพูด การเขียน และการอ่าน

4.2 วิชาภาษาไทย ระดับชั้นมัธยมศึกษาปีที่ 3 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ สาระการฟัง การดู และการพูด หลักการใช้ภาษา การอ่าน วรรณคดีและวรรณกรรม และการเขียน

4.3 วิชาภาษาไทย ระดับชั้นมัธยมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ การเขียน การฟัง การดู และการพูด การอ่าน หลักการใช้ภาษา

4.4 วิชาสังคมศึกษา ระดับชั้นประถมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ ภูมิศาสตร์ เศรษฐศาสตร์ และประวัติศาสตร์

4.5 วิชาสังคมศึกษา ระดับชั้นมัธยมศึกษาปีที่ 3 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ สารหน้าที่พลเมือง วัฒนธรรม และการดำเนินชีวิตในสังคม เศรษฐศาสตร์ ภูมิศาสตร์ และประวัติศาสตร์

4.6 วิชาสังคมศึกษา ระดับชั้นมัธยมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ สารภูมิศาสตร์ เศรษฐศาสตร์ ศาสนา ศิลธรรม จริยธรรม ประวัติศาสตร์ และหน้าที่พลเมือง วัฒนธรรม และการดำเนินชีวิตในสังคม

4.7 วิชาภาษาอังกฤษ ระดับชั้นประถมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ ภาษากับความสัมพันธ์กับชุมชนและโลก ภาษาเพื่อการสื่อสาร และภาษาและวัฒนธรรม

4.8 วิชาภาษาอังกฤษ ระดับชั้นมัธยมศึกษาปีที่ 3 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ ภาษาเพื่อการสื่อสาร และภาษาและวัฒนธรรม

4.9 วิชาภาษาอังกฤษ ระดับชั้นมัธยมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ ภาษากับความสัมพันธ์กับกลุ่มสาระการเรียนรู้อื่น ภาษาและวัฒนธรรม ภาษากับความสัมพันธ์กับชุมชนและโลก และภาษาเพื่อการสื่อสาร

4.10 วิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ การวิเคราะห์ข้อมูลและความน่าจะเป็น การวัด พีชคณิต จำนวนและการดำเนินการ และเรขาคณิต

4.11 วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ สารจำนวนและการดำเนินการ การวิเคราะห์ข้อมูลและความน่าจะเป็น ทักษะและกระบวนการทางคณิตศาสตร์ การวัด พีชคณิต และ เรขาคณิต

4.12 วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ สารจำนวนและการดำเนินการ พีชคณิต การวิเคราะห์ข้อมูลและความน่าจะเป็น และการวัด

4.13 วิชาวิทยาศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ ชีวิตกับสิ่งแวดล้อม ดาราศาสตร์และอวกาศ แรงและการเคลื่อนที่ สิ่งมีชีวิตกับกระบวนการดำรงชีวิต สารและสมบัติของสาร พลังงาน และกระบวนการเปลี่ยนแปลงของโลก

4.14 วิชาวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ สิ่งมีชีวิตกับกระบวนการดำรงชีวิต พลังงาน แรงและการเคลื่อนที่ สารและสมบัติของสาร ดาราศาสตร์และอวกาศ และชีวิตกับสิ่งแวดล้อม

4.15 วิชาวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 6 มีสาระที่ควรปรับปรุงการเรียนการสอนเรียงตามลำดับ คือ สารและสมบัติของสาร สิ่งมีชีวิตกับกระบวนการดำรงชีวิต กระบวนการเปลี่ยนแปลงของโลก ดาราศาสตร์และอวกาศ พลังงาน แรงและการเคลื่อนที่ ชีวิตกับสิ่งแวดล้อม และธรรมชาติของวิทยาศาสตร์และเทคโนโลยี

ศุภามณ จันทร์สกุล และสุกัญญา บุญศรี (2560) ได้ศึกษาการใช้โมเดล DINA และ G-DINA เพื่อวินิจฉัยความรอบรู้ทางการพยาบาล และเปรียบเทียบประสิทธิภาพของโมเดล DINA และ G-DINA โดยนำส่วนหนึ่งของเนื้อหาในรายวิชาการพยาบาลมารดาและทารกที่นักศึกษาพยาบาลต้องรอบรู้มากำหนดคุณลักษณะ (attribute: A) และแบบแผนการตอบ Q-matrix ซึ่งได้กำหนดไว้ 4 คุณลักษณะ ได้แก่ Attribute 1 (A1) ความก้าวหน้าในการคลอด Attribute 2 (A2) อาการแสดงในแต่ละระยะของการคลอด Attribute 3 (A3) การปรับตัวทางร่างกายและจิตสังคมของผู้คลอด และ Attribute 4 (A4) การเลือกสรรการพยาบาลที่เหมาะสม ตัวอย่างข้อสอบที่นำมาใช้ครั้งนี้มีจำนวน 10 ข้อ เป็นข้อสอบปรนัย 4 ตัวเลือก (item 1 ถึง item 10) จากนั้นสร้าง Q-matrix แต่ละข้อ สำหรับข้อมูลที่น่ามาวิเคราะห์คือ ผลการตอบข้อสอบของนักศึกษาพยาบาล จำนวน 150 คน ผลการวิจัยสรุปว่า ผลการตรวจสอบคุณภาพข้อคำถามเชิงการวินิจฉัยด้วยด้วยโมเดล DINA พบว่า ค่าสถิติทดสอบความสอดคล้องของข้อสอบ (test-level fit statistics) ได้แก่ -2LL มีค่าเท่ากับ 1838.5266 AIC มีค่าเท่ากับ 1908.5266 และ BIC มีค่าเท่ากับ 2013.8988 และค่าสถิติทดสอบความสอดคล้องของข้อสอบรายข้อ (item-level fit statistics) พิจารณาค่าดัชนีความสอดคล้องของโมเดลในเชิงสัมบูรณ์จากค่าสัดส่วนการตอบถูก (Proportion Correct--Prob.) ค่าสหสัมพันธ์การแปลงมาตรฐาน (Z--Transformed Correlation-Z(Corr)) และค่า Log-Odds Ratio มีค่าเฉลี่ย ค่าสูงสุด และค่าความคลาดเคลื่อนมาตรฐานอยู่ในระดับต่ำ แสดงว่า โมเดล DINA เป็นโมเดลที่มีความเหมาะสมในการนำมาวิเคราะห์หาคุณภาพในเชิงการวินิจฉัยของข้อคำถาม ส่วนการตรวจสอบคุณภาพข้อคำถามเชิงการวินิจฉัยด้วยด้วยโมเดล G-DINA พบว่า ค่าสถิติทดสอบความสอดคล้องของข้อสอบ (test-level fit statistics) ได้แก่ -2LL มีค่าเท่ากับ 1684.0390 AIC มีค่าเท่ากับ 1846.0390 และ BIC มีค่าเท่ากับ 2089.9005 และค่าสถิติทดสอบความสอดคล้องของข้อสอบรายข้อ (item-level fit statistics) พิจารณาค่าดัชนีความสอดคล้องของโมเดลในเชิงสัมบูรณ์จากค่าสัดส่วนการตอบถูก (Proportion Correct--Prob.) ค่าสหสัมพันธ์การแปลงมาตรฐาน (Z--Transformed Correlation-Z(Corr)) และค่า Log-Odds Ratio มีค่าเฉลี่ย ค่าสูงสุด และค่าความคลาดเคลื่อนมาตรฐานอยู่ในระดับต่ำ แสดงว่า โมเดล G-DINA เป็นโมเดลที่มีความเหมาะสมในการนำมาวิเคราะห์หาคุณภาพในเชิงการวินิจฉัยของข้อคำถาม ดังตารางที่ 2.2

ตารางที่ 2.2 ค่าดัชนีความสอดคล้องของโมเดลในเชิงสัมบูรณ์จากการวิเคราะห์ด้วยโมเดล DINA และ G-DINA

ค่าดัชนีความสอดคล้อง ของโมเดลเชิงสัมบูรณ์ (absolute fit indices)	ค่าสัดส่วนการตอบ ถูก: Prob.		ค่าสหสัมพันธ์การ แปลงมาตรฐาน: Z (Corr)		Log-Odds Ratio: Log(OR)	
	DINA	G-DINA	DINA	G-DINA	DINA	G-DINA
ค่าเฉลี่ย (Mean Abs. Dev.)	0.0072	0.0068*	0.1129*	0.1179	0.5180*	0.5474
ค่าสูงสุด (Max Abs. Dev.)	0.0275	0.0179*	0.4361	0.4230*	2.3129*	2.3873
ค่าความคลาดเคลื่อนมาตรฐาน (SE)	0.0364*	0.0377	0.0825	0.0825	0.3683*	0.3741

\* หมายถึง แสดงค่าที่น้อยกว่า

ผลการตรวจสอบคุณภาพของข้อคำถามในเชิงการวินิจฉัยด้วยโมเดล DINA ประกอบด้วย พารามิเตอร์การเดา (guessing parameter:  $g_i$ ) พารามิเตอร์การตอบผิดพลาดหรือพารามิเตอร์สะเพร่า (slipping parameter:  $s_i$ ) และดัชนีค่าอำนาจจำแนกในเชิงการวินิจฉัย ( $d_{i,DINA} = (1-s_i) - g_i$ ) (Rupp, Templin, Henson, 2010) ข้อคำถามทั้ง 10 ข้อ มีค่าพารามิเตอร์การเดาอยู่ระหว่าง 0.0000 – 0.6215 ค่าพารามิเตอร์สะเพร่าอยู่ระหว่าง 0.0000 – 0.3939 และดัชนีค่าอำนาจจำแนกในเชิงการวินิจฉัยอยู่ระหว่าง 0.1022 – 0.9954 โดยข้อคำถามที่มีดัชนีค่าอำนาจจำแนกมากกว่า 0.20 มีจำนวน 7 ข้อ เมื่อพิจารณาโปรไฟล์ความรอบรู้ของผู้ตอบมีจำนวน 16 ชั้นแฝง ซึ่งคำนวณจากสูตร  $2^k$  เมื่อ k แทน จำนวนคุณลักษณะ ซึ่งในที่นี้มี 4 คุณลักษณะ ( $k = 4$ ) โดยโมเดล DINA พบว่า โปรไฟล์ที่พบมากที่สุด คือ 1111 หมายถึง ผู้ตอบรอบรู้ทุกคุณลักษณะ คิดเป็นร้อยละ 38 และโปรไฟล์ที่มีสัดส่วนผู้ตอบน้อยที่สุดมี 8 รูปแบบ ได้แก่ 0000 0100 0010 0001 0110 0101 0011 และ 0111 เมื่อพิจารณาร้อยละของความรอบรู้รายคุณลักษณะ พบว่า ผู้ตอบมีความรอบรู้ในคุณลักษณะที่ 1 (A1) ความก้าวหน้าในการคลอด ร้อยละ 98.08 รองลงมาคือ รอบรู้ในคุณลักษณะที่ 2 (A2) อาการแสดงในแต่ละระยะของการคลอด ร้อยละ 93.70 รอบรู้ในคุณลักษณะที่ 4 (A4) การเลือกสรรการพยาบาลที่เหมาะสม ร้อยละ 71.08 และรอบรู้ในคุณลักษณะที่ 3 (A3) การปรับตัวทางร่างกายและจิตสังคมของผู้คลอดน้อยที่สุด ร้อยละ 54.00 สำหรับโมเดล G-DINA พบว่า โปรไฟล์ที่พบมากที่สุดคือ 1111 หมายถึง ผู้ตอบรอบรู้ทุกคุณลักษณะ คิดเป็นร้อยละ 23.51 และโปรไฟล์ที่มีสัดส่วนผู้ตอบน้อยที่สุดคือ 1100 ร้อยละ 2 เมื่อพิจารณาร้อยละของความรอบรู้รายคุณลักษณะ พบว่า ผู้ตอบรอบรู้ในคุณลักษณะที่ 1 (A1) ความก้าวหน้าในการคลอด ร้อยละ 81.92 รองลงมาคือ รอบรู้ในคุณลักษณะที่ 3 (A3) การปรับตัวทางร่างกายและจิตสังคมของผู้คลอด ร้อยละ 63.42 รอบรู้ในคุณลักษณะที่ 2 (A2) อาการแสดงในแต่ละระยะของการคลอด ร้อยละ 53.22 และรอบรู้ในคุณลักษณะที่ 4 (A4) การเลือกสรรการพยาบาลที่เหมาะสมน้อยที่สุด ร้อยละ 46.30 ดังตารางที่ 2.3

ตารางที่ 2.3 ร้อยละของความรอบรู้ในแต่ละคุณลักษณะของโมเดล DINA และ G-DINA

คุณลักษณะ (attribute)	ร้อยละของความรอบรู้ในแต่ละคุณลักษณะ	
	DINA	G-DINA
(A1) ความก้าวหน้าในการคลอด	98.08	81.92
(A2) อาการแสดงในแต่ละระยะของการคลอด	93.70	53.22
(A3) การปรับตัวทางร่างกายและจิตสังคมของผู้คลอด	54.44	63.42
(A4) การเลือกสรรการพยาบาลที่เหมาะสม	71.08	46.30