

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการสร้างและพัฒนาแบบทดสอบวินิจฉัยวิชาคณิตศาสตร์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ผู้วิจัยได้ศึกษา ตำรา เอกสารและงานวิจัยที่เกี่ยวข้อง ดังนี้

ความหมายและความสำคัญของวิชาคณิตศาสตร์

การวินิจฉัยการเรียนรู้

1. ความหมายของการวินิจฉัยการเรียนรู้
2. ระดับ วิธีการและขั้นตอนของการวินิจฉัยการเรียนรู้

แบบทดสอบวินิจฉัยการเรียนรู้

1. ความหมายของแบบทดสอบวินิจฉัยการเรียนรู้
2. ลักษณะของแบบทดสอบวินิจฉัยการเรียนรู้
3. เทคนิคการสร้างแบบทดสอบวินิจฉัยการเรียนรู้
4. ข้อดีและประโยชน์ของการใช้แบบทดสอบวินิจฉัยการเรียนรู้

เกณฑ์ปกติ (Norms)

1. การสร้างเกณฑ์ปกติ
2. ชนิดของเกณฑ์ปกติ

คะแนนที่ปกติ (Normalized T - Score)

ทฤษฎีการตอบสนองข้อสอบ

1. แนวคิดพื้นฐานของทฤษฎีการตอบสนองข้อสอบ
2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ
3. รูปแบบที่ใช้ในทฤษฎีการตอบสนองข้อสอบ
4. รูปแบบโลจิสติก
5. ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบทดสอบ
6. การประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ
7. สรุปแนวความคิดทฤษฎีการตอบสนองข้อสอบ

งานวิจัยที่เกี่ยวข้องกับการสร้างแบบทดสอบวินิจฉัยการเรียนรู้

งานวิจัยที่เกี่ยวข้องกับการใช้ทฤษฎีการตอบสนองข้อสอบในการพัฒนาแบบทดสอบ

ความหมายและความสำคัญของวิชาคณิตศาสตร์

ความหมายของคณิตศาสตร์

คณิตศาสตร์ (Mathematics) ตามพจนานุกรมฉบับราชบัณฑิตยสถาน พุทธศักราช 2525 ให้ความหมายว่า คณิตศาสตร์เป็นวิชาที่ว่าด้วยการคำนวณ (ราชบัณฑิตยสถาน, 2538 : 164)

เวบสเตอร์ (Webster, 1980 : 1110) อธิบายว่าคณิตศาสตร์ หมายถึงกลุ่มของวิชาต่าง ๆ ได้แก่ เลขคณิต เรขาคณิต แคลคูลัส ฯลฯ เกี่ยวกับปริมาณ (Quantitie) ขนาด (Sizes) รูปร่าง (Forms) และความสัมพันธ์ (Relation) โดยการใช้จำนวน (Number) และสัญลักษณ์ (Symbols) เป็นเครื่องช่วย

มีผู้ให้ความหมายเกี่ยวกับวิชาคณิตศาสตร์ไว้หลากหลายดังนี้ (ฉวีวรรณ เสวตมาลัย, 2545 : 15)

- ตำราหรือวิชาที่ว่าด้วยการคำนวณ
 - ศาสตร์ของการวัดปริมาณและคุณภาพ
 - ศาสตร์ของจำนวนและปริภูมิ (space)
 - ศาสตร์ของการคำนวณ
 - ศาสตร์ของสิ่งที่เป็นนามธรรม
 - ศาสตร์ของการให้เหตุผลทางตรรกะ
 - วิชาที่เกี่ยวกับจำนวน มีการบวก ลบ คูณ หาร
 - วิชาคำนวณที่แก้ปัญหาต่าง ๆ ได้
 - การพิสูจน์ทฤษฎีบทที่เกี่ยวกับรูปทรงต่าง ๆ
 - วิชาที่นักคณิตศาสตร์และนักโหราศาสตร์ใช้ในการคำนวณหาตำแหน่งของดวงดาว
 - การใช้สมการในรูปแบบต่าง ๆ
 - เรื่องที่เกี่ยวกับเซต
- ฯลฯ

ดังนั้นความหมายของคณิตศาสตร์จึงพอสรุปได้ว่า เป็นกลุ่มของวิชาต่าง ๆ ที่ว่าด้วยการคำนวณโดยอาศัยตัวเลขและสัญลักษณ์เป็นเครื่องมือ

ความสำคัญของคณิตศาสตร์

วิชาคณิตศาสตร์มีความจำเป็นและมีบทบาทสำคัญยิ่งในการดำเนินชีวิตของมนุษย์ได้มีผู้ที่กล่าวถึงความสำคัญของวิชาคณิตศาสตร์ ดังนี้

วรรณิ โสมประยูร (2526 : 228-230) ได้สรุปความสำคัญของคณิตศาสตร์ไว้ ดังนี้

1. คณิตศาสตร์มีประโยชน์ในชีวิตประจำวัน กิจกรรมต่าง ๆ ในชีวิตประจำวัน เช่น การซื้อขาย การดูเวลา การนับจำนวน ล้วนต้องอาศัยความรู้ทางคณิตศาสตร์ทั้งสิ้น

2. คณิตศาสตร์ช่วยให้เข้าใจโลก คณิตศาสตร์ช่วยให้เข้าใจ และรู้จักปรากฏการณ์ต่าง ๆ เช่น ทิศทางลม ฤดูกาล แรงดึงดูดของโลก โดยการอธิบายและการคำนวณทางคณิตศาสตร์

3. คณิตศาสตร์ช่วยสร้างทัศนคติที่ถูกต้องทางการศึกษา คณิตศาสตร์ช่วยให้ผู้เรียนคิดอย่างมีเหตุผลด้วยตนเอง รู้จักแก้ไขให้ถูกต้องเมื่อพบสิ่งที่ผิด และรู้จักนำความรู้ไปใช้ให้เป็นประโยชน์

4. คณิตศาสตร์เป็นสิ่งจำเป็นสำหรับการเรียนวิชาวิทยาศาสตร์ เนื่องจากการเรียนวิทยาศาสตร์ต้องมีความรู้ทางคณิตศาสตร์อย่างแท้จริง เพราะต้องอาศัยความรู้ในการสังเกตอย่างถี่ถ้วน การวัดที่ระมัดระวัง และคิดเลขที่ถูกต้อง

5. คณิตศาสตร์เป็นมรดกทางวัฒนธรรม ที่คนรุ่นก่อนคิดสร้างสรรค์ และถ่ายทอดมาสู่คนรุ่นหลัง การศึกษาคณิตศาสตร์จึงเป็นการศึกษาวัฒนธรรมและความก้าวหน้าของมนุษย์

สิริพร ทิพย์คง (2545 : 1) ได้กล่าวถึงความสำคัญของวิชาคณิตศาสตร์ไว้ว่า วิชาคณิตศาสตร์เป็นวิชาที่ช่วยก่อให้เกิดความเจริญก้าวหน้าทั้งทางด้านวิทยาศาสตร์และเทคโนโลยี โลกในปัจจุบันเจริญขึ้นเพราะการคิดค้นทางวิทยาศาสตร์และเทคโนโลยี ซึ่งต้องอาศัยความรู้ทางคณิตศาสตร์ ดังคำกล่าวของ คาร์ล ฟรีดริค เกาส์ (Carl Friedrich Gauss) ซึ่งเป็นนักคณิตศาสตร์ชาวเยอรมันที่มีชื่อเสียงในคริสต์ศตวรรษที่ 19 ว่า “คณิตศาสตร์เป็นราชินีของวิทยาศาสตร์และเลขคณิตเป็นราชินีของคณิตศาสตร์” (Mathematics is the queen of sciences and arithmetic is the queen of Mathematics) นอกจากนี้คณิตศาสตร์ยังช่วยพัฒนาให้แต่ละบุคคลเป็นคนที่สมบูรณ์ เป็นพลเมืองดี เพราะคณิตศาสตร์ช่วยเสริมสร้างความมีเหตุผล ความเป็นคนช่างคิด ช่างริเริ่ม สร้างสรรค์ มีระบบระเบียบในการคิด มีการวางแผนในการทำงาน มีความสามารถในการตัดสินใจ มีความรับผิดชอบต่อกิจการงานที่ได้รับมอบหมาย ตลอดจนมีลักษณะความเป็นผู้นำในสังคม

จะเห็นได้ว่าวิชาคณิตศาสตร์มีความสำคัญต่อมนุษย์เรามาก เพราะวิชาคณิตศาสตร์เป็นเครื่องมือในการเรียนรู้สิ่งต่าง ๆ และเป็นปัจจัยสำคัญในการพัฒนาคุณภาพมนุษย์ นอกจากนี้ ความรู้ในวิชาคณิตศาสตร์ ยังเป็นสิ่งจำเป็นในชีวิตประจำวัน เป็นฐานที่สำคัญแก่ประชาชน เพื่อ ความเป็นทรัพยากรมนุษย์ที่มีคุณค่า สามารถดำรงชีวิตได้อย่างมีประสิทธิภาพในสังคมปัจจุบัน และเป็นกำลังใจในการพัฒนาประเทศชาติต่อไป

การวินิจฉัยการเรียนรู้

1. ความหมายของการวินิจฉัยการเรียนรู้

นักการศึกษาหลายท่านได้ให้ความหมายของการวินิจฉัยไว้ดังนี้

โกวิท ประวาลพุกษ์ (2532 : 4) ได้ให้ความหมายคำว่า “วินิจฉัย” หมายถึงการค้นหา ลักษณะ ข้อบกพร่องของนักเรียนในการทำงานเรื่องหนึ่ง ๆ ว่าการที่นักเรียนทำไม่ได้เนื่องมาจาก มีความบกพร่องเกี่ยวกับความสามารถในประเด็นใดบ้าง เพื่อการแก้ไขต่อไป

วัลลพ กันทรพิชัย (2532 : 4) ให้ความเห็นว่าการวินิจฉัยว่าเป็นเทคนิคของการศึกษาสภาพ ปัญหาทั้งสาเหตุทางตรง และสาเหตุทางอ้อม เพื่อจะได้หาทางเลือกในการแก้ปัญหาได้ตรงจุด

สงบ ลักษณะ (2532 : 4) เห็นว่าการวินิจฉัยมีจุดมุ่งหมาย 3 ประการ กล่าวคือ ประการแรกเป็นการหาข้อบกพร่องในวิชาต่าง ๆ ประการที่สองเพื่อค้นหาสาเหตุข้อบกพร่องและประการที่สามเพื่อหาวิธีการพัฒนาแก้ไขข้อบกพร่อง

กู๊ด (Good, 1973 : 178, อ้างถึงใน ขจรพงษ์ หนูทอง, 2546 : 10) ได้ให้ความหมายของการ วินิจฉัยทางการเรียนไว้ว่า การวินิจฉัยการเรียนรู้ หมายถึง การค้นหาสาเหตุที่เป็นอุปสรรคและ ข้อบกพร่องทางการเรียน

สิงห์ (Singha, 1974 : 187, อ้างถึงใน ขจรพงษ์ หนูทอง, 2546 : 10) กล่าวว่า “การวินิจฉัย การเรียนหมายถึง การสืบค้นกระบวนการที่ไม่ประสบความสำเร็จทางการศึกษาหรือทางการเรียน เพื่อหาสาเหตุแล้วพยายามหาทางแก้ไข”

จากความหมายข้างต้นสรุปได้ว่า การวินิจฉัย หมายถึง การค้นหาข้อบกพร่องในวิชา ต่าง ๆ ของนักเรียนเพื่อตัดสินว่านักเรียนมีข้อบกพร่องเกี่ยวกับความสามารถตรงไหนเรื่องใด อันจะนำไปสู่การพัฒนาการเรียนการสอน ดังนั้น การวินิจฉัยข้อบกพร่องทางการเรียนวิชา คณิตศาสตร์ หมายถึง การค้นหาข้อผิดพลาดที่เป็นปัญหาหรือเป็นอุปสรรคที่ทำให้การเรียน คณิตศาสตร์ไม่ประสบความสำเร็จ

2. ระดับ ขั้นตอน และวิธีการของการวินิจฉัยการเรียนรู้

2.1 ระดับของการวินิจฉัยการเรียนรู้

วัชรีย์ บุรณสิงห์ (2526 : 422-433) ได้แบ่งการวินิจฉัยการเรียนรู้ออกเป็น 3 ระดับ คือ

1. ขั้นสำรวจ เป็นการวินิจฉัยขั้นต้นอย่างหยาบ ๆ เพื่อคว้านักเรียนคนไหนไม่รู้เรื่องเรียนไม่ทันเพื่อน แล้วนำไปวิเคราะห์เพิ่มเติม ซึ่งอาจทำได้โดยการสังเกต การทดสอบผลสัมฤทธิ์ทางการเรียน หรือการทดสอบความสามารถทางสมองเป็นกลุ่ม การวินิจฉัยการเรียนรู้ขั้นนี้ จะให้ผลใกล้เคียงความจริงมากขึ้น ถ้านำผลการสังเกตและการทดลองมากประกอบกัน

2. ขั้นเฉพาะ เป็นการวินิจฉัยเพื่อคัดแยกรายละเอียด จุดเด่น จุดด้อย หรือข้อบกพร่องทางการเรียน โดยใช้แบบทดสอบที่สามารถให้รายละเอียดของทักษะในการเรียนของนักเรียนแต่ละคนได้ ซึ่งมักทำเป็นรายบุคคลหรือกลุ่มย่อย

3. ขั้นละเอียด การวินิจฉัยขั้นนี้ มักใช้มักใช้ในบางกรณี สำหรับนักเรียนที่อ่อนหรือนักเรียนที่อุปสรรคทางการเรียนอย่างมาก

ดวงเดือน อ่อนน่วม (2533 : 34) ได้แบ่งการวินิจฉัยเป็น 3 ระดับคือ ระดับทั่วไป ระดับเฉพาะ ระดับละเอียด

1. การวินิจฉัยระดับทั่วไป เป็นการวินิจฉัยอย่างหยาบ จึงมักเรียกขั้นตอนนี้ว่าเป็นขั้นการสำรวจเพราะเป็นเพียงการสำรวจเพื่อให้ทราบระดับความสามารถทั่วไปของเด็กทั้งกลุ่มและรายบุคคล เมื่อเทียบกับกลุ่มใหญ่เช่นเทียบกับเกณฑ์ปกติ

2. การวินิจฉัยระดับเฉพาะ เป็นการเก็บรวบรวมข้อมูลอย่างละเอียดเกี่ยวกับความสามารถทางคณิตศาสตร์ของนักเรียน เพื่อให้ทราบว่านักเรียนมีข้อบกพร่องที่ใด เครื่องมือที่นิยมใช้ในการวินิจฉัยระดับนี้คือแบบสอบเช่นเดียวกันกับการวินิจฉัยระดับทั่วไป แต่แบบสอบที่ใช้ในการวินิจฉัยระดับนี้ต้องเป็นแบบสอบที่วัดความสามารถทางคณิตศาสตร์ในวงแคบ คือ เฉพาะเจาะจงไปที่เรื่องใดเรื่องหนึ่ง หรือย่อยลงไปถึงความคิดรวบยอดใดความคิดรวบยอดหนึ่ง หรือทักษะใดทักษะหนึ่ง

3. การวินิจฉัยระดับละเอียด เป็นการศึกษารายละเอียดอย่างลึกซึ้งเกี่ยวกับสมรรถภาพทางคณิตศาสตร์ของเด็กที่มีข้อมูลแสดงให้เห็นชัดเจนว่ามีปัญหาซับซ้อน การวินิจฉัยเพียงอย่างเดียวยังได้ข้อมูลไม่เพียงพอ จำเป็นต้องหาข้อมูลเพิ่มเติมด้วยวิธีอื่น เช่น จากการสังเกต จากการสัมภาษณ์ ข้อมูลที่ต้องการเพิ่มเติมเช่นชีวิตครอบครัว ทักษะคิดต่อคณิตศาสตร์ สุขภาพ อารมณ์ บุคลิกภาพ ในกรณีที่มีปัญหาซับซ้อนมาก เกินความรู้ความสามารถของครูสอนคณิตศาสตร์ หรือต้องใช้เวลา

ในการวินิจฉัยมากจนครูที่สอนในชั้นเรียนปกติไม่สามารถทำได้ ก็อาจจำเป็นต้องใช้ผู้ที่ได้รับการฝึกฝนในด้านนี้มาแล้วโดยเฉพาะ

สำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ (2539 : 2-3) ได้แบ่งระดับของการวินิจฉัยออกเป็น 3 ระดับ ได้แก่ ระดับทั่วไป ระดับเฉพาะ ระดับละเอียด

1. ระดับทั่วไป (general level) เป็นการวินิจฉัยอย่างหยาบ เพราะเป็นขั้นตอนการสำรวจเพื่อหาระดับความสามารถทั่ว ๆ ไปของนักเรียน แบบสอบที่จะใช้วัดในระดับนี้ ถ้าหากในต่างประเทศมักใช้แบบสอบมาตรฐาน แต่ในเมืองไทยการใช้แบบสอบวัดผลสัมฤทธิ์ที่ใช้สอบวัดปลายภาคเรียน หรือปลายปีก็สามารถนำมาใช้ได้ ทั้งนี้ในการตรวจให้คะแนนมิใช่ดูที่คะแนนรวม แต่จะพิจารณาคะแนนเป็นรายสมรรถภาพในแต่ละสมรรถภาพ หรือคะแนนในแต่ละโดเมน (Domain) หรือแต่ละ Subdomain ว่านักเรียนไม่บรรลุผลการเรียนใน Subdomain ใดบ้าง ก็คน

2. ระดับเฉพาะ (specific level) เป็นระดับที่ต้องการทราบว่านักเรียนมีข้อบกพร่องในเรื่องใด ณ จุดใด เป็นการวัดความสามารถเฉพาะเจาะจงลงไปในเรื่องใดเรื่องหนึ่ง เช่น การบวก การลบ และในการบวกยังแตกแยกย่อยลงไปอีกว่าบวกจำนวนเต็ม บวกเศษส่วน บวกทศนิยม ดังนี้ เป็นต้น แบบสอบที่ใช้วัดในระดับนี้คือ แบบทดสอบวินิจฉัย ในต่างประเทศมีแบบสอบประเภทนี้ให้เลือกใช้มากมาย แต่ในเมืองไทยเกือบพูดได้ว่าไม่มีเลย ยกเว้นงานวินิจฉัยของนิสิตปริญญาโทของมหาวิทยาลัยต่าง ๆ พอมีบ้าง แต่ก็มีน้อยมาก แบบสอบในลักษณะนี้ครูผู้สอนที่มีความตั้งใจ สนใจก็สามารถสร้าง และพัฒนาให้เป็นแบบสอบวินิจฉัยที่มีคุณภาพได้ ทั้งนี้ต้องมีความรู้ความเข้าใจในหลักการ วิธีการสร้างรู้ลักษณะของแบบทดสอบ และกระบวนการสร้างแบบสอบตลอดจนการแปลผลของคะแนน

3. ระดับละเอียด (intensive level) เป็นการวินิจฉัยอย่างละเอียดลึกซึ้งเป็นการหาข้อมูล อาจใช้วิธีการสังเกต การสัมภาษณ์ ข้อมูลไม่ได้ใช้เฉพาะผลสัมฤทธิ์จากการตอบแบบสอบ แต่อาจวัดเจตคติ บุคลิกภาพ สุขภาพ หรืออื่น ๆ แหล่งข้อมูลมิใช่จำกัดจากนักเรียนเพียงอย่างเดียว อาจต้องสอบถามจากผู้ปกครองครูที่สอนวิชาอื่น ๆ เพื่อนสนิท ผู้วินิจฉัยไม่ใช่เพียงแต่ครูผู้สอนวิชานั้น ๆ อาจประกอบด้วยผู้เชี่ยวชาญทางด้านวัดผล นักจิตวิทยา หรือครูแนะแนวด้วยก็ได้

การวินิจฉัยการเรียนขั้นละเอียดนั้นเป็นหน้าที่ของผู้เชี่ยวชาญในแต่ละสาขาโดยเฉพาะ ส่วนหน้าที่ของครูผู้สอนที่ทำอย่างต่อเนื่องคือ การวินิจฉัยการเรียน ขั้นสำรวจ และขั้นเฉพาะเท่านั้น อย่างไรก็ตาม การวินิจฉัยในแต่ละวิธีหรือแต่ละระดับอาจต้องใช้เครื่องมือบางชนิดเข้ามาช่วย เช่น แบบทดสอบวัดผลสัมฤทธิ์ (Achievement Test) แบบทดสอบวัดระดับสติปัญญา (Intelligence Test) แบบทดสอบความถนัด (Aptitude Test) หรือแบบทดสอบวินิจฉัยการเรียน (Diagnostic Test) (Benjamin S. Bloom, 1971 : 116)

2.2 ขั้นตอนการวินิจฉัย

กรอนลันด์ (Gronlund, 1981 : Unpaged, อ้างถึงใน วลี เฉลยสมัย, 2538 : 28) ได้เสนอขั้นตอนการวินิจฉัยและแก้ไขข้อบกพร่องไว้ 4 ขั้น คือ การระบุตัวนักเรียนที่มีข้อบกพร่อง การระบุข้อบกพร่อง การระบุองค์ประกอบที่เป็นสาเหตุของการมีข้อบกพร่อง และการแก้ไขข้อบกพร่อง ซึ่งมีรายละเอียดดังนี้

(1) การระบุตัวนักเรียนที่มีข้อบกพร่อง

การระบุตัวนักเรียนที่มีข้อบกพร่องสามารถทำได้หลายวิธี เช่น ใช้แบบสอบถาม ใช้การสังเกต เป็นต้น ครูไม่ควรมองแต่ปัญหาด้านเนื้อหาวิชาเท่านั้น ควรมองปัญหาอื่นด้วย เช่น ด้านการปรับตัว ด้านอารมณ์ เพราะปัญหาเหล่านี้อาจมีผลกระทบต่อปัญหาด้านการเรียนของนักเรียน

(2) การระบุข้อบกพร่อง

ปัญหาของนักเรียนมีหลายระดับ ในบางครั้งการวินิจฉัยเพียงระดับทั่วไปอาจให้ข้อมูลพอเพียงสำหรับการแก้ไข บางกรณีต้องวินิจฉัยถึงระดับวิเคราะห์ และบางกรณีต้องวินิจฉัยถึงระดับคลินิกจึงจะสามารถหาข้อแก้ไขได้ วิธีการที่ควรให้ความสำคัญไปกับการใช้แบบทดสอบก็คือให้นักเรียนคิดออกเสียง ซึ่งช่วยให้ครูทราบกระบวนการคิดของนักเรียน

(3) การระบุองค์ประกอบที่เป็นสาเหตุของการมีข้อบกพร่อง

องค์ประกอบที่ควรพิจารณา เช่น สถิติปัญญา ทักษะการเรียน สุขภาพ การปรับตัว ด้านอารมณ์และสิ่งแวดล้อมทางบ้าน เพราะสิ่งเหล่านี้อาจเป็นสาเหตุของปัญหาในการเรียน

(4) การแก้ไขข้อบกพร่อง

การแก้ไขข้อบกพร่องไม่มีรูปแบบตายตัว ขึ้นอยู่กับธรรมชาติของข้อบกพร่องแต่ละอย่าง บางกรณีอาจแก้ไขด้วยการทบทวนหรือสอนใหม่ บางกรณีต้องใช้ในการสร้างแรงจูงใจแก้ไขปัญหาด้านอารมณ์ หรือแก้ไขทักษะการทำงาน

2.3. วิธีการวินิจฉัยข้อบกพร่อง

วลี เฉลยสมัย (2538 : 28-30) ได้สรุปวิธีการวินิจฉัยข้อบกพร่อง จากการศึกษาเครื่องมือและวิธีการที่ใช้สำหรับวินิจฉัยข้อบกพร่องในการเรียนวิชาคณิตศาสตร์นั้นสามารถสรุปได้เป็น 2 ประเภท คือ การวินิจฉัยอย่างไม่เป็นทางการ และการวินิจฉัยอย่างเป็นทางการซึ่งรายละเอียดของการวินิจฉัยในแต่ละประเภทนั้นได้นำเสนอไว้ดังนี้

1. การวินิจฉัยอย่างไม่เป็นทางการ (Informal diagnosis) เป็นการใช้เทคนิควิธีการต่าง ๆ ในการหาข้อมูลเกี่ยวกับข้อบกพร่องของนักเรียน วิธีการที่ใช้ เช่น การสังเกต (Observation) และสอบถามนักเรียนเป็นรายบุคคล และการตรวจแบบฝึกหัด

2. การวินิจฉัยอย่างเป็นทางการ (formal diagnosis) เป็นการใช้วิธีการหาข้อมูลเกี่ยวกับข้อบกพร่องของนักเรียน ด้วยวิธีการที่สร้างขึ้นมาอย่างเป็นระบบ และมีแบบแผน วิธีการที่ใช้ เช่น การวินิจฉัยข้อบกพร่องด้วยแบบทดสอบวินิจฉัย การวินิจฉัยข้อบกพร่องด้วยดัชนีชี้วัดความผิดปกติของแบบแผนคะแนนการตอบ การวินิจฉัยด้วยวิธีการของทาทชูโอกะ และการวินิจฉัยข้อบกพร่องแบบย้อนรอยกระบวนการคิดด้วยวิธีการของศิริเดช สุชีวะ

นอกจากนี้ โอคีย์ (Okey, 1977 : 222–223, อ้างถึงใน สมศักดิ์ ฉันทานุรักษ์, 2529 : 7) ได้สร้างคำถาม 5 ข้อ เพื่อเป็นแนวคิดสำหรับครูผู้สอนที่จะทำการวินิจฉัยการเรียนรู้เพื่อเพิ่มผลสัมฤทธิ์ทางการเรียนของนักเรียนให้ดีกว่าเดิม ดังนี้

1. นักเรียนเรียนรู้ได้ตามที่ครูพยายามสอนหรือไม่
2. นักเรียนมีพื้นฐานความรู้เท่าไรก่อนที่ครูสอน
3. นักเรียนมีความรู้ตามที่ครูคิดว่าพวกเขาควรจะรู้หรือไม่
4. นักเรียนรับรู้สิ่งที่พวกเขาเรียนไปแล้วมากน้อยเพียงใด
5. นักเรียนชอบสิ่งที่พวกเขาได้เรียนหรือไม่

การหาข้อมูลเพื่อตอบคำถามทั้ง 5 ข้อนี้ ควรทำทั้งก่อนเรียน ขณะเรียนและหลังเรียน โดยศึกษาทั้งตัวนักเรียนและครูผู้สอน ซึ่งอาจต้องใช้วิธีการวินิจฉัยหลาย ๆ แบบ เข้ามาช่วยในการตอบคำถามเหล่านี้ ดังที่ พันทิพา อุทัยสุข (2523 : 14) ได้กล่าวไว้โดยสรุปได้ว่าการวินิจฉัยข้อบกพร่องทางการเรียน อาจต้องใช้วิธีการหลาย ๆ แบบประกอบกันดังนี้

1. การสังเกตการสอน เป็นการพิจารณาว่า นักเรียนมีความสนใจ และมีสมาธิในการเรียนหรือไม่
2. การศึกษาเป็นรายกรณี เป็นการศึกษาเรื่องทั่ว ๆ ไปของนักเรียนบางคน que คิดว่าอาจมี ปัญหา
3. การทดสอบปกติ เป็นการดูผลการเรียนที่ได้จากการสอนและความก้าวหน้าของนักเรียน
4. การทดสอบอย่างละเอียด เป็นการค้นหาข้อบกพร่องทางการเรียนของนักเรียนได้ตรงจุดจริง ๆ ว่าส่วนใดต้องแก้ไข โดยพยายามออกข้อสอบให้ได้คำตอบอย่างชัดเจนถึงข้อบกพร่องของนักเรียน
5. จากการสัมภาษณ์ผู้ปกครอง เป็นการปรึกษาหารือเกี่ยวกับปัญหาต่าง ๆ ของนักเรียนทั้งด้านการเรียน และด้านอื่น ๆ

วิธีการวินิจฉัยข้อบกพร่องทางการเรียนเหล่านี้ อาจจะนำมาวินิจฉัย จุดเด่นของนักเรียนเก่ง โดยปรับปรุงวิธีการวินิจฉัยการเรียนให้เหมาะสม

จะเห็นว่า การวินิจฉัยการเรียนควรเริ่มจาก ความต้องการของครูว่าต้องการข้อมูลไปทำอะไร แล้วจึงเลือกใช้เครื่องมือให้เหมาะสมกับวิธีการตามระดับของวินิจฉัยการเรียนที่สอดคล้องกับวัตถุประสงค์นั้น

แบบทดสอบวินิจฉัยการเรียน

1. ความหมายของแบบทดสอบวินิจฉัย

แบบทดสอบวินิจฉัย เป็นเครื่องมือที่ใช้ตรวจค้นข้อบกพร่องทางการเรียนของนักเรียนได้ละเอียดและมีประสิทธิภาพ จึงนับว่าเป็นแบบทดสอบที่ให้ประโยชน์สูงในด้านการศึกษา ได้มีผู้ให้ความหมายของแบบทดสอบวินิจฉัยไว้ดังนี้

บุญชม ศรีสะอาด (2523 : 10) กล่าวว่า แบบทดสอบวินิจฉัยเป็นแบบทดสอบที่สร้างขึ้นเพื่อให้เห็นจุดบกพร่อง จุดที่เป็นปัญหาหรืออุปสรรค (Difficulty) ในการเรียนเรื่องหนึ่ง ๆ ของนักเรียนแต่ละคน ทั้งนี้เพื่อที่จะได้หาทางแก้ไขได้ตรงจุดยิ่งขึ้น อันจะทำให้สามารถช่วยเหลือนักเรียนที่มีปัญหาหรืออุปสรรคในการเรียนบรรลุจุดประสงค์ในการเรียนหรือเกิดการเรียนรู้ได้เหมือนคนอื่น ๆ

วิรัช นิยมเยี่ยม (2525 : 13) กล่าวว่า แบบทดสอบวินิจฉัย เป็นแบบทดสอบที่ใช้ค้นหาข้อบกพร่องในการเรียนวิชาต่าง ๆ และหาสาเหตุของข้อบกพร่องนั้น ๆ เพื่อจัดให้มีการสอนซ่อมเสริมต่อไป

สมเกียรติ ปติฐพร (2526 : 9) กล่าวว่า แบบทดสอบวินิจฉัย เป็นแบบทดสอบที่สร้างขึ้นเพื่อค้นหาความบกพร่องของนักเรียนในการเรียนวิชาต่าง ๆ เป็นรายบุคคลพร้อมกับหาสาเหตุของความบกพร่องของนักเรียนในการเรียนวิชาต่าง ๆ เป็นรายบุคคลพร้อมกับหาสาเหตุของความบกพร่องนั้น เพื่อนำไปสู่การแก้ไขการเรียน และซ่อมเสริมความบกพร่องนั้นต่อไป

สุเทพ สันติวรานนท์ (2533 : 69) ได้ให้ความหมายของแบบทดสอบวินิจฉัย คือ แบบทดสอบที่สร้างขึ้น เพื่อชี้ให้เห็นจุดบกพร่องของนักเรียนที่เกิดขึ้นในการเรียนเนื้อหาวิชานั้น ๆ อีกทั้งช่วยให้ทราบสาเหตุของความบกพร่อง อันจะเป็นประโยชน์ต่อการเรียนการสอนในแง่ของการเป็นแนวทางนำไปสู่การสอนซ่อมเสริมได้ตรงจุด การปรับปรุงแก้ไขนักเรียนให้เกิดการเรียนรู้

เพิ่มขึ้น ปรับปรุงการสอนของครูให้มีประสิทธิภาพมากขึ้น ซึ่งแบบทดสอบประเภทนี้จะใช้สอบกับนักเรียนหลังทำการสอนจบทั้งรายบุคคลหรือกลุ่ม

เบญจา เขียวสม (2534 : 7) กล่าวว่าแบบทดสอบวินิจฉัยเป็นเครื่องมือตรวจค้นข้อบกพร่องในการเรียนของนักเรียนพร้อมทั้งวิเคราะห์สาเหตุที่ทำให้เกิดข้อบกพร่องนั้น ๆ เพื่อให้ครูผู้สอนนำไปใช้เป็นแนวทางในการพัฒนาความสามารถของนักเรียน

ในต่างประเทศได้มีผู้ให้ความหมายของแบบทดสอบวินิจฉัยไว้ดังนี้

อดัมส์ และทอร์เกอร์สัน (Adams and Torgerson, 1964 : 39 – 40, อ้างถึงใน สุภาพร ละอองวิจิตร, 2543 : 18) กล่าวว่าแบบทดสอบวินิจฉัยสร้างขึ้น เพื่อใช้ชี้ให้เห็นจุดบกพร่อง และสาเหตุของการบกพร่องในการเรียนแบบทดสอบวินิจฉัยไม่ได้เน้นความสำคัญที่คะแนนรวม แต่เน้นรูปแบบของคำตอบ (Pattern of Responses) เป็นสำคัญ ผลของการทดสอบทำให้ทราบได้ว่านักเรียนคนใดบกพร่องหรือมีปัญหาการเรียน

อีเบล (Ebel, 1965 : 449 อ้างถึงใน ขจรพงษ์ หนูทอง, 2546 : 18) ได้ให้ความหมายของแบบทดสอบวินิจฉัยว่าเป็นแบบทดสอบที่สร้างขึ้นเพื่อ ใช้ค้นหาจุดอ่อน หรือข้อบกพร่องในการเรียนวิชาต่าง ๆ ของนักเรียน เช่น การอ่านและเลขคณิต แบบทดสอบวินิจฉัยข้อบกพร่องในการเรียนจะสนใจคะแนนแต่ละข้อหรือสนใจคะแนนของกลุ่มเล็ก ๆ ที่ทดสอบในแบบทดสอบที่คล้าย ๆ กัน

อาห์แมนและกล็อก (Ahmann and Glock, 1967 : 18, อ้างถึงใน ขจรพงษ์ หนูทอง, 2546 : 18) กล่าวว่าแบบทดสอบวินิจฉัย คือ แบบทดสอบที่ใช้หลังจากการเรียนการสอนแล้ว เพื่อให้ทราบถึงข้อบกพร่องเฉพาะที่เป็นพื้นฐานที่อยู่เบื้องหลังของนักเรียน

อนาสตาซี (Anastasi, 1968 : 404, อ้างถึงใน สุภาพร ละอองวิจิตร, 2543 : 8) ได้ให้ความหมายของแบบทดสอบวินิจฉัยไว้ว่าเป็นแบบทดสอบที่สร้างขึ้นเพื่อวิเคราะห์ความเก่ง-อ่อนเป็นรายบุคคล และเป็นการบอกสาเหตุของความอ่อนนั้นด้วย

เพย์ (Payne, 1968 : 167, อ้างถึงใน ขจรพงษ์ หนูทอง, 2546 : 18) กล่าวถึง แบบสอบวินิจฉัยว่าเป็นแบบสอบที่สอดคล้องกับจุดมุ่งหมายของหลักสูตร และวัตถุประสงค์ของการสอน เป็นแบบทดสอบที่ใช้ทำการทดสอบหลังจากการเรียนเสร็จสิ้นลง อาจจะจัดเป็นการทดสอบรายบุคคลหรือเป็นกลุ่ม เพื่อชี้ให้เห็นจุดบกพร่องของการเรียนรู้แต่ละตอน อันจะเป็นประโยชน์ต่อการเรียนการสอน

กล่าวโดยสรุปจากแนวความคิดข้างต้นที่เกี่ยวกับแบบทดสอบวินิจฉัย พอสรุปความหมายได้ว่า แบบทดสอบวินิจฉัยเป็นแบบทดสอบที่สร้างขึ้นเพื่อค้นหาความบกพร่องและสาเหตุของความบกพร่องในการเรียนของนักเรียน เพื่อให้ครูผู้สอนนำไปใช้ปรับปรุงการเรียนการสอน และเพื่อแก้ไขความบกพร่องของนักเรียนเป็นรายบุคคลได้ถูกต้องและตรงจุด

2. ลักษณะของแบบทดสอบวินิจฉัย

นักการศึกษาหลายท่านทั้งในประเทศและต่างประเทศได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยไว้ดังนี้

บุญชม ศรีสะอาด (2523 : 9-11) กล่าวถึงลักษณะทั่วไปของแบบทดสอบวินิจฉัยไว้ดังนี้

1. จะมุ่งวัดเป็นเรื่อง ๆ ไป หรือด้าน ๆ ไป ถ้าต้องอาศัยทักษะย่อยหลายทักษะก็อาจแบ่งเป็นแบบทดสอบย่อย (Subtest) วัดตามทักษะย่อย ๆ นั้น

2. มีคะแนนของแต่ละด้านแต่ละตอน เพราะมุ่งค้นหาจุดบกพร่องในแต่ละด้าน ดังนั้นคะแนนรวมของห้องหรือคะแนนรวมของแต่ละคนจะไม่มีประโยชน์

3. มีข้อสอบหลาย ๆ ข้อ วัดมโนภาพหรือทักษะเดียวกัน ซึ่งจะทำให้เพิ่มโอกาสการทำผิดพลาดมากขึ้น อันจะช่วยให้สามารถจำแนกนักเรียนที่มีความบกพร่องในการเรียนเรื่องนั้น ๆ ได้อย่างเพียงพอ

4. มักเป็นแบบไม่เร่งรัดเวลาในการทำ (Power Test) โดยเริ่มจากข้อที่ง่าย ๆ แล้วค่อยเพิ่มความยากขึ้น และโดยส่วนรวมแล้วจะมีแนวโน้มค่อนข้างง่ายกว่าแบบทดสอบที่มุ่งสำรวจ (Survey Test)

5. การสร้างแบบทดสอบชนิดนี้ จะสร้างจากรากฐานการวิเคราะห์ทักษะเฉพาะที่ส่งผลให้เรียนได้สำเร็จ และจากการศึกษาข้อผิดพลาดหรือความบกพร่องที่มักเกิดขึ้นกับนักเรียน

6. ความเป็นมาตรฐานของแบบทดสอบวินิจฉัยจะอยู่ในรูปที่ว่า เครื่องมือที่ใช้ดำเนินการสอบภายใต้สภาพเดียวกัน และการให้คะแนนมีความเป็นปรนัย

ปราณี ทองคำ และ นพพร เจริญทอง (2536 : 8) ได้สรุปลักษณะแบบทดสอบเพื่อการวินิจฉัยดังนี้

1. แบบทดสอบวินิจฉัย มีรากฐานมาจากการวิเคราะห์ทักษะเฉพาะอย่าง ที่ส่งผลให้การเรียนสำเร็จได้

2. แบ่งออกเป็นแบบทดสอบย่อย ๆ หลายฉบับ แต่ละฉบับใช้วัดทักษะเฉพาะอย่าง

3. ในแต่ละฉบับประกอบด้วยข้อสอบที่มีจำนวนข้อมากและค่อนข้างง่าย

4. เน้นความตรงเชิงเนื้อหา
5. ข้อสอบแต่ละข้อสามารถบ่งบอกสาเหตุของการตอบผิดได้
6. แบบทดสอบย่อยแต่ละฉบับต้องกำหนดเกณฑ์ขั้นต่ำสำหรับการวินิจฉัยที่เหมาะสมกับความบกพร่องแต่ละชนิด
7. ไม่มีเกณฑ์ปกติ เพราะมีจุดมุ่งหมายเพื่อค้นหาข้อบกพร่องในการเรียนของนักเรียนเป็นรายบุคคล

อดัมส์และทอร์เกอร์สัน (Adams and Torgerson, 1964 : 472, อ้างถึงใน สุภาพร ละอองวิจิตร, 2543 : 11) ได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยไว้ ดังนี้

1. แบบทดสอบวินิจฉัย จะแยกออกเป็นแบบทดสอบย่อย (Subtests) เพื่อวัดทักษะเฉพาะอย่างของการเรียนวิชาต่าง ๆ และจะต้องกำหนดเกณฑ์ขั้นต่ำในการวินิจฉัยที่เหมาะสมกับความบกพร่องแต่ละชนิด

2. แบบทดสอบย่อยแต่ละฉบับ จะต้องมีความยาวพอที่จะวัดความสามารถของแต่ละบุคคลได้อย่างมีความเชื่อมั่น

3. ปกติแบบทดสอบวินิจฉัยจะใช้กับนักเรียนที่มีผลสัมฤทธิ์ต่ำ ดังนั้นข้อสอบจึงมักมีจำนวนข้อมาก ๆ และเป็นข้อสอบที่ค่อนข้างง่าย

4. เกณฑ์ปกติ (Norms) ไม่มีความสำคัญในแบบทดสอบวินิจฉัยเพราะว่าจุดประสงค์ที่สำคัญที่สุดของแบบทดสอบวินิจฉัย คือ เพื่อที่จะค้นหาว่าสิ่งใดที่นักเรียนไม่สามารถจะทำได้และสาเหตุใดมากกว่าที่จะเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของนักเรียน

อาห์แมนและกล็อก (Ahmann and Glock, 1967 : 364-365, อ้างถึงใน ขจรพงษ์ หนูทอง, 2546 : 24) ได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยไว้ ดังนี้

1. แบบทดสอบวินิจฉัยเน้นความตรงเชิงเนื้อหา (Content Validity) เป็นสำคัญ
2. เกณฑ์ปกติไม่มีความสำคัญในแบบทดสอบวินิจฉัย
3. แบบทดสอบวินิจฉัย ประกอบด้วยกลุ่มข้อสอบที่เกิดจากการวิเคราะห์คำตอบของนักเรียนเป็นรายข้อ แล้วรวบรวมข้อสอบและคำตอบที่เป็นปัญหา ซึ่งเกิดขึ้นกับนักเรียนจำนวนมากไว้เพื่อค้นหาจุดบกพร่องต่อไป
4. แบบทดสอบวินิจฉัย มักใช้เพื่อแก้ไขปัญหาทางการเรียนให้กับนักเรียนที่มีคะแนนต่ำสร้างจากการตอบแบบทดสอบเพื่อสำรวจ (Survey Test)

บลูม (Bloom, 1971 : 91–92, อ้างถึงใน สุภาพร ละอองวิจิตร, 2543 : 10) ได้กล่าวถึงแบบทดสอบวินิจัย ไว้ดังนี้

1. เป็นแบบทดสอบ เพื่อหาจุดที่บกพร่องของนักเรียนเกี่ยวกับทักษะพื้นฐาน เพื่อหา ระดับของการเรียนรู้ เพื่อคัดแยกเด็ก เพื่อปรับปรุงวิธีสอนและเพื่อหาว่านักเรียนคนใดต้องสอนซ้ำ
2. ใช้ทดสอบระหว่างการเรียนการสอนเมื่อนักเรียนได้รับการสอนจากวิธีสอนแบบปกติ พอควรแล้ว
3. ใช้ประเมินผลได้ทั้งพฤติกรรม ด้านพุทธิพิสัย (Cognitive Domain) ด้านจิตพิสัย (Affective Domain) และด้านทักษะพิสัย (Psychomotor Domain)
4. มีทั้งแบบทดสอบเพื่อการวินิจฉัยมาตรฐานและแบบทดสอบที่ครูสร้างขึ้นเอง
5. มีจำนวนมากข้อ และแต่ละข้อมีค่าความยากง่าย ตั้งแต่ 0.65 ขึ้นไป
6. การประเมินคะแนนจากแบบทดสอบ อาจใช้ได้ทั้งแบบอิงเกณฑ์และอิงกลุ่ม
7. วิธีรายงานคะแนนจากแบบทดสอบ ทำได้โดยการเขียนเส้นภาพ (Profile) ของแต่ละคนในแต่ละทักษะย่อย

จากลักษณะของแบบทดสอบวินิจัยที่กล่าวมาแล้ว สามารถสรุปลักษณะของแบบทดสอบวินิจัยดังนี้ คือ

1. มักแยกข้อสอบแต่ละวิชาออกเป็นฉบับย่อย ๆ หลายฉบับ เพื่อที่จะวัดความรู้ความสามารถของนักเรียนเป็นด้าน ๆ ไป
2. มีจำนวนข้อมาก ๆ
3. ไม่คำนึงถึงคะแนนรวม จะพิจารณาเฉพาะคำตอบของข้อสอบแต่ละข้อหรือกลุ่มของข้อสอบ
4. เป็นแบบทดสอบที่ไม่จำกัดเวลา
5. ไม่ได้สร้างขึ้นเพื่อเลื่อนชั้นนักเรียน แต่เพื่อค้นหาอุปสรรคในการเรียน
6. มักจะใช้กับนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่ำ
7. ไม่มีการสร้างเกณฑ์ปกติ
8. เป็นแบบทดสอบที่เน้นความตรงเชิงเนื้อหาเป็นสำคัญ (Content Validity)
9. เนื้อหาที่ต้องวัดจะต้องสอดคล้องกับจุดมุ่งหมายที่กำหนดไว้ในหลักสูตร
10. ใช้ทดสอบระหว่างการเรียนการสอน
11. ประเมินได้ทั้งพฤติกรรมด้านพุทธิพิสัย จิตพิสัย และทักษะพิสัย
12. มีทั้งเป็นแบบทดสอบวินิจัยมาตรฐาน และแบบที่ครูสร้างขึ้น
13. คำถามมักเป็นคำถามที่ค่อนข้างง่าย

นอกจากนี้ ได้มีนักการศึกษากล่าวถึงความตรงของแบบทดสอบวินิจฉัยไว้ดังนี้

ชวาล แพร์ตกุล (2518 : 5-6) ได้กล่าวถึง ลักษณะของแบบทดสอบวินิจฉัยที่เกี่ยวกับ ความตรงของแบบทดสอบไว้ว่า แบบทดสอบวินิจฉัยนิยมที่จะแยกข้อสอบแต่ละวิชาออกเป็นฉบับย่อย ๆ หลายฉบับ โดยมีเป้าหมายที่จะวัดความรู้ความสามารถของนักเรียนเป็นขั้น ๆ ไปสอบ เพื่อให้ครอบคลุมเนื้อหาและพฤติกรรมที่สำคัญตามที่กำหนดไว้ในหลักสูตร

เอื้อ งานทอง (2527 : 15) ได้สรุปลักษณะของแบบทดสอบวินิจฉัยเกี่ยวกับความตรงของแบบทดสอบไว้ว่า แบบทดสอบวินิจฉัย แบ่งข้อสอบออกเป็นแบบทดสอบย่อย ๆ หลายฉบับในแต่ละจุดประสงค์การเรียนรู้ และเน้นความตรงเชิงเนื้อหาเป็นสำคัญ

สุเทพ สันติวรานนท์ (2533 : 69-70) ได้กล่าวถึง ลักษณะของแบบทดสอบวินิจฉัยที่เกี่ยวกับความตรงของแบบทดสอบไว้ว่า แบบทดสอบวินิจฉัยนั้นต้องสร้างให้สอดคล้องกับจุดประสงค์ของหลักสูตรและจุดประสงค์การสอน มีการวิเคราะห์และการสุ่มเนื้อหาอย่างรอบคอบ เน้นความตรงเชิงเนื้อหา (Content Validity) เป็นหลัก

อาห์แมนและกล็อก (Ahmann and Glock, 1967 : 364-365, อ้างถึงใน นงนุช สุภวรรณ์, 2546 : 20) ได้กล่าวถึง ลักษณะของแบบทดสอบวินิจฉัยที่เกี่ยวกับความตรงของแบบทดสอบไว้ว่า แบบทดสอบวินิจฉัยเน้นความตรงเชิงเนื้อหา (Content Validity) เป็นสำคัญ

เพย์ (Payne, 1968 : 167, อ้างถึงใน นงนุช สุภวรรณ์ , 2546 : 20) ได้กล่าวถึง ลักษณะของแบบทดสอบวินิจฉัยที่เกี่ยวกับความตรงของแบบทดสอบวินิจฉัยจะต้องสร้างให้สอดคล้องกับจุดมุ่งหมายของหลักสูตร และจุดประสงค์ของการสอน ประกอบด้วยข้อสอบซึ่งเกิดจากการวิเคราะห์รายละเอียดขององค์ประกอบที่เกี่ยวกับการเรียนรู้เนื้อหานั้น ๆ

จากที่กล่าวมาข้างต้นของนักการศึกษาสรุปได้ว่า ความตรงของแบบทดสอบวินิจฉัยนั้นจะเน้นความตรงเชิงเนื้อหาเป็นสำคัญ

3. เทคนิคการสร้างแบบทดสอบวินิจฉัย

เครื่องมือที่มีคุณค่ายิ่งในการวินิจฉัยข้อบกพร่องของนักเรียนคือแบบทดสอบวินิจฉัยทั้งนี้ เพราะสามารถสร้างได้ตรงตามจุดประสงค์เชิงพฤติกรรม ฉะนั้นในการสร้างแบบทดสอบแต่ละประเภท ย่อมมีเทคนิคและวิธีการสร้างที่แตกต่างกันออกไป ทั้งนี้ขึ้นอยู่กับจุดมุ่งหมายในการนำแบบทดสอบไปใช้ แบบทดสอบวินิจฉัยก็เช่นเดียวกันย่อมแตกต่างจากแบบทดสอบวัดผลสัมฤทธิ์ประเภทอื่น ๆ ดังที่นักการศึกษาหลายท่านได้กล่าวถึงเทคนิคและวิธีการสร้างไว้ต่าง ๆ กันดังนี้

สุเทพ สันติวรานนท์ (2533 : 71) ได้สรุปขั้นตอนในการสร้างแบบทดสอบวินิจฉัยไว้ดังนี้

1. กำหนดจุดมุ่งหมายและวางแผนในการดำเนินการสร้างแบบทดสอบ
2. วิเคราะห์ทักษะที่จำเป็นและเนื้อหาวิชาการอย่างละเอียด แล้วแบ่งออกเป็นองค์

ประกอบย่อย ๆ

3. เขียนจุดประสงค์เชิงพฤติกรรม (behavioral Objective) ให้ครอบคลุมเนื้อหาที่กำหนดไว้

4. เขียนข้อสอบให้สอดคล้องกับจุดประสงค์เชิงพฤติกรรม ในข้อสอบจะกำหนดให้นักเรียนหาคำตอบและสาเหตุของการเลือกตอบ ซึ่งขั้นนี้ถือว่าเป็นขั้นของการสร้างแบบทดสอบเพื่อสำรวจหาสาเหตุของการเลือกตอบ

5. นำไปสอบกับนักเรียนที่เรียนเนื้อหานั้นผ่านมาแล้ว

6. วิเคราะห์คำตอบและหาสาเหตุของการไม่สัมฤทธิ์ผลตามจุดประสงค์เชิงพฤติกรรมจากแบบทดสอบเพื่อสำรวจ ทั้งนี้เพื่อนำผลการวิเคราะห์มากำหนดสร้างตัวเลือกของแบบทดสอบวินิจฉัยต่อไป

7. เขียนข้อสอบ โดยสร้างตัวเลือกจากสาเหตุของการเลือกตอบของนักเรียน

8. นำข้อสอบในขั้นที่ 7 มารวบรวมเป็นแบบทดสอบวินิจฉัย แล้วนำไปทดลองใช้และพัฒนาปรับปรุงคุณภาพให้ดีขึ้น

9. เขียนคู่มือในการใช้แบบทดสอบ และกำหนดแนวทางที่เหมาะสมเพื่อให้สามารถบ่งชี้ถึงข้อบกพร่องและค้นหาสาเหตุของข้อบกพร่องในแต่ละทักษะนั้นได้

สุมาลี โชคสมุทร (2535 : 12) กล่าวถึงขั้นตอนการสร้างแบบทดสอบวินิจฉัยไว้ดังนี้

1. วางแผนในการสร้างข้อสอบ
2. วิเคราะห์เนื้อหาอย่างละเอียด
3. วิเคราะห์ทักษะและความรู้พื้นฐานที่ต้องการวัดออกเป็นองค์ประกอบ ๆ ให้ชัดเจน

4. เขียนข้อสอบให้สอดคล้องกับจุดประสงค์เชิงพฤติกรรม

5. วิเคราะห์สาเหตุที่นักเรียนไม่รอบรู้ตามจุดประสงค์เชิงพฤติกรรมนั้น

6. นำแบบทดสอบไปทดลองใช้และปรับปรุงแบบทดสอบ ซึ่งประกอบด้วยขั้นตอนการทดสอบเพื่อสร้างตัววาง ทดสอบเพื่อวิเคราะห์รายข้อและทดสอบเพื่อหาคุณภาพที่สร้างขึ้น

7. วิเคราะห์จุดบกพร่องในการตอบแบบทดสอบของกลุ่มตัวอย่าง

8. จัดพิมพ์แบบทดสอบและคู่มือดำเนินการสอบ

พร้อมพรรณ อุดมสิน (2538 : 94) กล่าวถึงลำดับขั้นการสร้างแบบทดสอบวินิจฉัยการเรียนรู้ไว้ดังนี้

1. ศึกษาและวิเคราะห์เนื้อหาที่ต้องการวินิจฉัยอย่างละเอียด แล้วแบ่งออกเป็นเนื้อหาย่อย ๆ เป็นตอน ๆ ไป
2. วิเคราะห์ทักษะที่ต้องการวัดเป็นองค์ประกอบย่อยในเนื้อหาของแต่ละตอน
3. ศึกษาและรวบรวมสาเหตุของข้อบกพร่องทางการเรียนในเนื้อหาแต่ละทักษะย่อย
4. เขียนข้อสอบให้สามารถวัดทักษะย่อยเหล่านั้น โดยให้มีจำนวนมากข้อที่จะบ่งชี้ถึงความบกพร่องในแต่ละจุด
5. ข้อสอบในแต่ละทักษะย่อยนั้นควรเป็นข้อสอบที่ง่าย และอาจจะแบ่งข้อสอบออกเป็นแบบทดสอบย่อยตามเนื้อหาแต่ละตอน
6. ตรวจสอบความถูกต้องของแบบทดสอบ แล้วนำไปทดลองใช้ เพื่อปรับปรุงแก้ไขแบบทดสอบให้มีประสิทธิภาพ

7. เขียนคู่มือการใช้และแบบแผนการวินิจฉัย

สหชาติ เหล็กชาย (2538 : 20) ได้สรุปลำดับขั้นในการสร้างแบบทดสอบวินิจฉัยไว้ดังนี้

1. วางแผนในการสร้างแบบทดสอบ
2. วิเคราะห์เนื้อหาอย่างละเอียดและแบ่งออกเป็นเนื้อหาย่อย ๆ
3. วิเคราะห์ทักษะที่จะวัดออกเป็นองค์ประกอบย่อย ๆ ให้ชัดเจน
4. เขียนข้อสอบให้สอดคล้องกับวัตถุประสงค์เชิงพฤติกรรม
5. วิเคราะห์สาเหตุที่นักเรียนไม่รอบรู้ตามวัตถุประสงค์เชิงพฤติกรรมนั้น
6. นำไปทดลองใช้และปรับปรุงแบบทดสอบ
7. เขียนคู่มือการใช้แบบทดสอบ

สำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ (2539 : 7) สรุปวิธีการสร้างและพัฒนาแบบทดสอบวินิจฉัยไว้ว่า

1. วิเคราะห์เนื้อหา กำหนดขอบเขตเนื้อหา และระดับพฤติกรรมอย่างละเอียด
2. สร้างตารางวิเคราะห์โครงสร้างของวิชา / รายวิชา
3. สร้างแบบทดสอบเพื่อสำรวจ
4. เขียนจุดประสงค์การเรียนรู้ / สมรรถภาพ / สมรรถภาพย่อย
5. หาแบบผิด หรือข้อบกพร่องที่คิดว่าน่าจะเกิดในขณะที่นักเรียนทำกิจกรรมหรือแบบฝึกหัดในแต่ละจุดประสงค์การเรียนรู้
6. เขียน Script ของข้อสอบ หรือเขียนลักษณะเฉพาะ (Item Specification)

7. เขียนข้อสอบตาม Script หรือ Item Specification
8. ตรวจสอบคุณภาพข้อสอบรายชื่อ
 - ค่าความหมายรายชื่อ (IOC)
 - ค่าความลำเอียง (Bias)
9. ทดลองสอบ หาค่าสถิติ ปรับปรุงคุณภาพข้อสอบ
10. จัดฉบับแบบทดสอบ ทดลองสอบ หาคคุณภาพของแบบทดสอบ
11. เขียนคู่มือการสร้าง และพัฒนาแบบทดสอบ คู่มือการใช้แบบทดสอบ การแปลความหมายของคะแนน และคู่มือในการวินิจฉัย

นอลล (Noll, 1957 : 363) ได้กล่าวถึงลำดับขั้นในการสร้างแบบทดสอบวินิจฉัยดังนี้

1. การสร้างแบบทดสอบวินิจฉัย จะต้องมีการวิเคราะห์กฎ (Rules) หลักการ (Principle) ความรู้ (Knowledges) หรือทักษะ (Skills) ซึ่งเกี่ยวข้องกับสิ่งที่ต้องการวัดอย่างละเอียด

2. แบบทดสอบวินิจฉัยที่ดีจะต้องสร้างให้ครอบคลุม กฎ ทฤษฎี และหลักการต่าง ๆ

3. ข้อสอบควรจัดเรียงเป็นกลุ่ม ตามลักษณะที่ต้องการวินิจฉัย

ลินด์ควิสต์ (Lindquist, 1966 : 37-38) ได้ให้หลักเกณฑ์ในการสร้างแบบทดสอบวินิจฉัยไว้ดังนี้

1. เขียนจุดประสงค์ที่ต้องการทดสอบให้มีความชัดเจนและให้สัมพันธ์กับหลักสูตร
2. สร้างคำถามให้สามารถวัดได้ตรงตามจุดประสงค์ที่ต้องการทดสอบ
3. วิเคราะห์ข้อสอบอย่างละเอียด โดยอาศัยการทดสอบเพื่อหาอุปสรรคหรือความไม่เข้าใจในการเรียนเป็นหลัก
4. แบบทดสอบ ต้องสามารถวัดพฤติกรรมทางสมองของนักเรียนได้เพียงพอและต้องใช้ค้นหาจุดบกพร่องทางการเรียนได้
5. แบบทดสอบ ต้องเสนอแนะจุดบกพร่องในแต่ละองค์ประกอบทางการเรียนเพื่อที่ทำการวัดได้อย่างถูกต้อง
6. แบบทดสอบต้องมีความครอบคลุมกฎเกณฑ์ทางการเรียนอย่างทั่วถึง
7. สามารถใช้แบบทดสอบความบกพร่องนั้นจากเนื้อหาแต่ละฉบับที่ทำการทดสอบได้
8. ความก้าวหน้าของนักเรียนแต่ละคน แสดงให้ทราบได้จากคำตอบที่ได้จากแบบทดสอบ

บราวน์ (Brown, 1970 : 303) กล่าวถึงแบบทดสอบวินิจฉัยไว้ว่า

1. แบ่งทักษะที่ต้องการออกเป็นองค์ประกอบย่อยได้อย่างชัดเจน
2. แบบทดสอบย่อยแต่ละฉบับสามารถวัดองค์ประกอบของทักษะนั้นได้เพียง

องค์ประกอบเดียว

3. แบบทดสอบทุกฉบับต้องสามารถวัดทักษะย่อยที่ต้องการจะวัดได้จริง ๆ
4. คะแนนจากแบบทดสอบย่อยจะต้องกำหนดแนวทางที่เหมาะสม เพื่อให้สามารถหา

แนวทางสอนซ่อมเสริมได้ตรงจุด

กรอปเปอร์ (Groppe, 1974 : 145) กล่าวว่า การสร้างแบบทดสอบวินิจฉัยมีขั้นตอนดังนี้

1. วางแผนสร้างแบบทดสอบ
2. เขียนข้อสอบโดยใช้จุดประสงค์เชิงพฤติกรรมเป็นเกณฑ์
3. หาสาเหตุที่ไม่สัมฤทธิ์ผลตามจุดมุ่งหมายเชิงพฤติกรรมนั้น
4. นำแบบทดสอบไปทดลองใช้และปรับปรุงแบบทดสอบ

สิงห์ (Singha, 1974 : 200-204) กล่าวถึงการสร้างแบบทดสอบวินิจฉัยว่า

1. ไม่จำเป็นต้องสร้างตารางวิเคราะห์หลักสูตร (Blue-Print) เนื่องจากไม่ได้หา

ความสัมพันธ์ทางเนื้อหาวิชาการกับวิธีการ

2. ไม่จำเป็นต้องสร้างเกณฑ์ปกติในการวิจัย เพราะจุดมุ่งหมายของแบบทดสอบวินิจฉัย เพื่อที่จะค้นหาข้อบกพร่องและสาเหตุของข้อบกพร่องนั้น

3. หากสร้างแบบทดสอบวินิจฉัยเป็นแบบปรนัยชนิดเลือกตอบหรือแบบตอบสั้น ๆ ก็ควรมีจำนวนข้อไม่น้อยกว่า 3 ข้อ ในแต่ละเนื้อหาย่อย

4. แบบทดสอบวินิจฉัยจะเรียงข้อสอบตามเนื้อหาคือ เอาข้อความที่อยู่ในเนื้อหาเดียวกัน เข้าไว้ด้วยกัน โดยไม่จำเป็นต้องคำนึงถึงความยาก

5. การสร้างแบบทดสอบวินิจฉัยอาจสร้างเป็นแบบมาตรฐาน (Standardized Test) หรือแบบที่ครูสร้างขึ้น (Teacher-made Test) แต่ว่าแบบทดสอบที่ครูสร้างขึ้นจะคุ้มค่ามากกว่าเพราะประหยัดเวลาและกำลังงานกว่าแบบทดสอบมาตรฐาน

จากวิธีการสร้างแบบทดสอบวินิจฉัยที่ได้กล่าวมาแล้วข้างต้น สามารถสรุปเป็นขั้นตอนในการสร้างแบบทดสอบวินิจฉัย ได้ดังนี้

1. กำหนดจุดมุ่งหมายและวางแผนในการดำเนินการสร้างแบบทดสอบ
2. ศึกษาหลักสูตร คู่มือครู หนังสือเรียน เพื่อวิเคราะห์เนื้อหา แล้วเขียนจุดประสงค์เชิงพฤติกรรม
3. สร้างตารางวิเคราะห์หลักสูตร

4. กำหนดจำนวนข้อคำถามลงในตารางวิเคราะห์หลักสูตร
5. สร้างแบบทดสอบวินิจฉัย นำไปทดลองใช้ และพัฒนาแบบทดสอบให้มีคุณภาพที่ดีขึ้น
6. เขียนคู่มือการใช้แบบทดสอบ

4. ข้อดีและประโยชน์ของการใช้แบบทดสอบวินิจฉัยการเรียนรู้

ในบรรดาเครื่องมือหลายชนิดที่ใช้เพื่อการวินิจฉัยนั้น แบบทดสอบวินิจฉัยการเรียนรู้ที่สร้างขึ้นโดยเฉพาะ สามารถให้ประโยชน์ในการวินิจฉัยได้มากที่สุด เพราะมีข้อดีอยู่หลายประการ ดังรายละเอียดต่อไปนี้

ทองหล่อ วิภาวีน (2521 : 50 - 51) กล่าวว่า การใช้แบบทดสอบวินิจฉัยการเรียนรู้ให้ประโยชน์ทั้งครูและนักเรียนหลายประการดังนี้

สำหรับนักเรียน

1. เมื่อผู้เรียนทราบล่วงหน้าว่าจบบทเรียนจะมีการทดสอบวินิจฉัยการเรียนรู้ นักเรียนจะกลัวความล้มเหลว ซึ่งจะเป็นตัวกระตุ้นให้เกิดแรงจูงใจ ทั้งยังมีผลทางด้านจิตวิทยาทำให้นักเรียนเรียนดีขึ้น
2. ผลการสอบจะทำให้นักเรียนประเมินตัวเองได้ว่า เขาได้เรียนรู้เรื่องนี้มากน้อยเพียงใด และควรปรับปรุงอะไรบ้าง ทำให้นักเรียนรู้จักความสามารถของตนเอง ซึ่งเป็นจุดมุ่งหมายที่สำคัญอย่างหนึ่งของหลักสูตร

3. แบบทดสอบวินิจฉัยการเรียนรู้จะเป็นเครื่องช่วยตัดสินใจว่านักเรียนเข้าใจในเนื้อหาหรือทักษะในเรื่องเหล่านั้น ตลอดจนมีความพร้อมที่จะเรียนเรื่องต่อไปหรือยัง

สำหรับครู

1. ช่วยปรับปรุงการเรียนการสอนของครู เพื่อให้รู้ว่า ครูควรสอนเรื่องอะไร และหัวข้อใดที่นักเรียนมีข้อบกพร่อง
2. ช่วยให้ครูเตรียมบทเรียนได้ตามความต้องการของนักเรียน โดยใช้เทคนิคที่เหมาะสมกับ นักเรียนในเนื้อหาแต่ละตอน

พร้อมพรรณ อุดมสิน (2538 : 93-94) กล่าวว่า การใช้แบบทดสอบวินิจฉัยการเรียนรู้เป็นเครื่องมือที่สำคัญและมีคุณค่ามากสำหรับการประเมินผลการเรียนระดับมัธยมศึกษาในปัจจุบัน เพราะให้ประโยชน์ทั้งครูและนักเรียนดังนี้

สำหรับครู

1. ช่วยปรับปรุงการสอนของครู เพื่อจะได้รู้ว่าครูควรสอนเรื่องอะไร และหัวข้อใดที่นักเรียนยังมีข้อบกพร่องอยู่เพราะเนื้อหาบางเรื่องต้องใช้ความรู้พื้นฐานเก่า หากว่าครูยังไม่แก้ไขข้อบกพร่องเก่า ๆ จะทำให้การเรียนรู้เนื้อหาต่อไปไม่ประสบความสำเร็จได้
2. ช่วยให้ครูเตรียมบทเรียนได้ตรงตามความต้องการของนักเรียน โดยใช้เทคนิคได้อย่างเหมาะสมกับนักเรียนในเนื้อหาแต่ละตอน เพราะเนื้อหาแต่ละตอนมีความยากไม่เหมือนกันหากว่าครูได้ทราบว่าเนื้อหาตอนใดเป็นปัญหามากต่อนักเรียน ครูก็ควรจะต้องพุ่งเป้าเป็นพิเศษในเนื้อหาตอนนั้น ๆ และหาวิธีสอนที่จะทำให้ให้นักเรียนประสบความสำเร็จได้
3. ประหยัดเวลาและแรงงานของครูในการวินิจฉัย ทำให้มีเวลาสอนซ่อมเสริมเป็นรายบุคคลมากขึ้น

สำหรับนักเรียน

1. ผลการสอบจากแบบสอบวินิจฉัยการเรียนจะทำให้ให้นักเรียนได้ประเมินตนเองได้ว่าตนเองมีจุดประสงค์ใดที่ยังบกพร่องอยู่ สมควรจะได้รับแก้ไข ทำให้นักเรียนรู้ความสามารถของตนเองซึ่งเป็นจุดมุ่งหมายสำคัญอย่างหนึ่งของหลักสูตร
2. จากการทำแบบสอบวินิจฉัยการเรียน ผลที่ได้จะเป็นเครื่องช่วยตัดสินใจว่านักเรียนมีความเข้าใจ เนื้อหาหรือมีทักษะในเรื่องนั้นหรือไม่ ตลอดจนมีความพร้อมที่จะเรียนต่อไปได้หรือไม่
3. จะเป็นแรงจูงใจในการเรียน ให้นักเรียนเตรียมพร้อมในการเรียนอยู่เสมอเพราะถ้านักเรียนทราบว่าจบบทเรียนแล้วจะมีการทดสอบเพื่อวินิจฉัยการเรียน นักเรียนจะกลัวความล้มเหลว จะทำให้สนใจการเรียน

ลินด์ควิสต์ (Lindquist, 1966 : 37, อ้างถึงใน นงนุช สุภวรรณ์ , 2546 : 26) กล่าวถึงข้อดีของแบบทดสอบวินิจฉัยการเรียนสรุปได้ว่า

1. สามารถวิเคราะห์ข้อบกพร่องทางการเรียนของนักเรียนได้ละเอียดกว่าวิธีการอื่น ๆ
2. ช่วยให้ครูได้ทราบถึงองค์ประกอบที่สำคัญ ลำดับขั้นที่จำเป็น ตลอดจนอุปสรรคในการเรียนการสอน
3. ประหยัดเวลาและแรงงานของครูในการวินิจฉัย ทำให้มีเวลาจัดซ่อมเสริมเป็นรายบุคคลได้มากขึ้น
4. ช่วยให้นักเรียนทราบว่าตนเองต้องเรียนอะไรเพิ่มเติมจากการปรับปรุงข้อบกพร่องที่สำคัญของตนเอง

5. ประหยัดเวลาของครูในการจัดหรือเตรียมการซ่อมเสริมและช่วยในการปรับปรุงการเรียนการสอน

กรอนลินด์ (Gronlund, 1981 : 322 อ้างถึงใน ขจรพงษ์ หนูทอง, 2546 : 37) กล่าวว่า การเลือกและการใช้แบบทดสอบวินิจฉัยการเรียนให้เกิดประโยชน์ต้องคำนึงว่า

1. แบบทดสอบวินิจฉัยการเรียนแต่ละฉบับ สะท้อนถึงมโนคติเกี่ยวกับเรื่องที่จะวัดของผู้สร้าง และข้อคิดของผู้เรียนในการวินิจฉัย

2. แบบทดสอบวินิจฉัยการเรียนสร้างขึ้นสำหรับนักเรียนที่มีผลการเรียนต่ำ จึงเหมาะสำหรับพิจารณาข้อบกพร่องทางการเรียน แต่ไม่เหมาะสำหรับการพิจารณาระดับความชำนาญ

3. แบบทดสอบวินิจฉัยการเรียน จะเป็นตัวบ่งประเภทของข้อบกพร่องของนักเรียนแต่ไม่บอกสาเหตุของข้อบกพร่องนั้น แม้ว่าบางครั้งจะสามารถบอกสาเหตุจากประเภทของข้อบกพร่อง หรือจากการอธิบายคำตอบของนักเรียนได้ แต่ข้อบกพร่องบางชนิดอาจเกิดขึ้นจากหลายสาเหตุหรือเกี่ยวข้องกันในลักษณะที่ซับซ้อน

4. แบบทดสอบวินิจฉัยการเรียนที่ทำการวินิจฉัยอุปสรรคทางการเรียนของนักเรียนเพียงส่วนเดียว ต้องพิจารณาความสัมพันธ์ที่มีต่อส่วนประกอบนั้นด้วย

5. ผลที่ได้จากแบบทดสอบย่อยหรือกลุ่มของข้อสอบ ในการวินิจฉัยการเรียนอาจเชื่อถือได้น้อย เพราะอาจมีบางหัวข้อเท่านั้นที่วัดทักษะเฉพาะ ดังนั้นการหาข้อเด่น-ด้อยทางการเรียนควรศึกษาจากการสังเกตในห้องเรียนประกอบด้วย

จากประโยชน์ที่กล่าวมาข้างต้นจะเห็นว่า การนำแบบทดสอบวินิจฉัยการเรียนไปใช้ก่อให้เกิดประโยชน์ต่อการจัดการเรียนการสอนอย่างมาก อย่างไรก็ตาม แบบทดสอบวินิจฉัยการเรียนยังไม่ใช่เป็นเครื่องมือแก้ไขข้อบกพร่องโดยตรง แต่จะเป็นเครื่องมือค้นหาข้อบกพร่องเพื่อหาวิธีแก้ไขต่อไป

เกณฑ์ปกติ (Norms)

เกณฑ์ปกติ หมายถึง ข้อเท็จจริงทางสถิติที่บรรยายการแจกแจงของคะแนนจากประชากรที่นิยามไว้อย่างดีแล้ว และเป็นคะแนนที่จะบอกระดับความสามารถของผู้สอบว่าอยู่ระดับใดของกลุ่มประชากร (ลิวิน สายยศ และอังคณา สายยศ , 2543 : 313-314)

1. การสร้างเกณฑ์ปกติ

ในการสร้างเกณฑ์ปกติ (Norms) ต้องคำนึงถึงหลัก 3 ประการดังนี้ (สมนึก ภัททิยธนี, 2541 : 31-32)

1.1 ความเป็นตัวแทนที่ดี การสุ่มตัวอย่างของประชากรโดยอาศัยความน่าจะเป็นทำได้หลายวิธี เช่น สุ่มแบบธรรมดา สุ่มแบบแบ่งชั้น สุ่มแบบเป็นระบบ หรือสุ่มแบบแบ่งกลุ่ม ทั้งนี้ต้องเลือกสุ่มตามความเหมาะสมโดยพิจารณาประชากรเป็นสำคัญ ถ้าประชากรมีลักษณะเป็นอันหนึ่งอันเดียวกันไม่มีคุณสมบัติอะไรแตกต่างกัน ใช้วิธีการสุ่มแบบธรรมดา (Simple Random Sampling) แต่ถ้าระหว่างประชากรกับกลุ่มย่อยมีลักษณะแตกต่างกัน เช่น ขนาดของโรงเรียนต่างกัน จะต้องใช้วิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) คือสุ่มมาจากประชากรทุกกลุ่มย่อย ในทางตรงข้าม ถ้าระหว่างประชากรกลุ่มย่อยมีลักษณะเหมือนกัน เช่น นักเรียนแต่ละห้องเรียนซึ่งแบ่งคละระหว่างเด็กเก่ง ปานกลาง และอ่อน การสุ่มแบบนี้มีลักษณะการสุ่มแบบแบ่งกลุ่ม (Cluster Random Sampling) คือ สุ่มแบบแบ่งกลุ่มจากประชากรกลุ่มย่อย ทั้ง 3 วิธีนี้ใช้ในการสุ่มเพื่อสร้างเกณฑ์ปกติมากที่สุด

1.2 มีความเที่ยงตรง ในที่นี้หมายถึงการนำคะแนนดิบไปเทียบกับเกณฑ์ปกติที่ทำไว้แล้วสามารถแปลความหมายได้ตรงกับความเป็นจริง เช่น นักเรียนคนหนึ่งสอบวิชาคณิตศาสตร์ได้ 20 คะแนน ตรงกับคะแนนที่ (T) 50 แปลว่า มีความสามารถปานกลาง ความเป็นจริงจะเป็นเช่นนั้นจริงหรือไม่ ดังนั้นความสอดคล้องของคะแนนการสอบกับเกณฑ์ปกติตามความเป็นจริง จึงถือว่าเป็นสิ่งสำคัญมาก

1.3 มีความทันสมัย เกณฑ์ปกตินั้นขึ้นอยู่กับความสามารถของประชากรกลุ่มนั้นการพัฒนาคนมีอยู่ตลอดเวลา เทคโนโลยี สภาพแวดล้อม อาหารการกิน สิ่งเหล่านี้จะช่วยให้คนเก่งหรืออ่อนได้ ดังนั้นเกณฑ์ปกติที่เคยศึกษาไว้นานแล้วหลายปี อาจมีความผิดพลาดจากความเป็นจริง จำเป็นต้องสร้างขึ้นใหม่ให้ทันสมัย โดยทั่วไปแล้วเกณฑ์ปกติควรเปลี่ยนทุก ๆ 5 ปี

2. ชนิดของเกณฑ์ปกติ

ล้วน สายยศ และอังคณา สายยศ (2543 : 315-317) แบ่งชนิดของเกณฑ์ปกติเป็น 2 ลักษณะใหญ่ ๆ

2.1 เกณฑ์ปกติตามลักษณะของประชากร

2.1.1 เกณฑ์ปกติระดับชาติ (National norms) การสร้างเกณฑ์ปกติระดับชาตินั้นใช้ประชากรที่นิยามไว้มากมายทั่วประเทศ เช่น หาเกณฑ์ปกติของวิชาเลขคณิตระดับชั้น

ประถมศึกษาปีที่ 4 ระดับชาติ ก็ต้องสอบนักเรียนชั้นประถมศึกษาปีที่ 4 ทั่วประเทศ หรือสุ่มตัวอย่างให้ครอบคลุมทั่วประเทศ จำนวนนักเรียนที่จะต้องสอบจึงมีมากมาย เพื่อให้รู้ว่าสร้างเมื่อปี พ.ศ. ไດ ก็ต้องกำหนดวัน เดือน ปี การสร้างไว้ด้วย เพื่อคนใช้เกณฑ์ปกติจะรู้ว่าทันสมัยหรือไม่

2.1.2 เกณฑ์ปกติระดับท้องถิ่น (Local norms) เป็นการสร้างเกณฑ์ปกติระดับเล็กลงมา เช่น ระดับจังหวัด หรือระดับอำเภอ การสร้างเกณฑ์ปกติระดับนี้ค่าใช้จ่ายจะน้อยลงและเป็นประโยชน์ในการเปรียบเทียบคะแนนของผู้สอบกับคนทั้งจังหวัดหรืออำเภอ ในการจัดการศึกษา บางครั้งจังหวัดแต่ละจังหวัด อาจเน้นการประมง วิชาที่มีภาระเน้นแตกต่างกัน การสร้างเกณฑ์ปกติระดับท้องถิ่นจะมีประโยชน์มาก แต่วิชาพื้นฐานอื่น ๆ ก็สามารถหาเกณฑ์ปกติระดับท้องถิ่นได้เหมือนกัน เพื่อประโยชน์ในการเปรียบเทียบความสามารถในวิชาการของนักเรียนคนหนึ่งกับคนทั้งจังหวัดหรืออำเภอว่าเด็กคนนั้นสอบแล้วจะอยู่ในระดับใด เก่งหรืออ่อนกว่าคนอื่นเพียงใด จะได้หาทางปรับปรุง แก้ไขทัน ถ้าไม่มีการเปรียบเทียบก็ไม่สามารถพัฒนาได้ถูกต้อง

2.1.3 เกณฑ์ปกติของโรงเรียน (School norms) โรงเรียนบางแห่งมีขนาดใหญ่ นักเรียนแต่ละชั้นมีจำนวนมาก เวลาสร้างข้อสอบแต่ละวิชาแต่ละระดับชั้น ได้ดีมีมาตรฐานแล้วจะสร้างเกณฑ์ปกติของโรงเรียนตัวเองก็ได้ กรณีสร้างเกณฑ์ปกติของโรงเรียนเดียวหรือกลุ่มโรงเรียนในเครือ เรียกว่าเกณฑ์ปกติของโรงเรียน ใช้ประเมินเปรียบเทียบนักเรียนแต่ละคนกับนักเรียนส่วนรวมของโรงเรียน และใช้ประเมินการพัฒนาของโรงเรียนได้ด้วย โดยดูได้จากการศึกษาแต่ละปีว่าเด่นหรือด้อยกว่าปีที่สร้างเกณฑ์ปกติเอาไว้

2.2 แบ่งเกณฑ์ปกติตามลักษณะของการใช้สถิติเปรียบเทียบ

2.2.1 เกณฑ์ปกติเปอร์เซ็นต์ไทล์ (Percentile norms) เกณฑ์แบบนี้สร้างจากคะแนนดิบที่มาจากประชากร หรือกลุ่มตัวอย่างที่เป็นตัวแทนที่ดี แล้วดำเนินการตามวิธีการสร้างเกณฑ์ปกติ และนำคะแนนมาหาค่าเปอร์เซ็นต์ไทล์ เกณฑ์ปกติแบบนี้เป็นเพียงการจัดอันดับคะแนนเท่านั้น ไม่สามารถจะนำไปบวกลบกันได้ แต่สามารถเปรียบเทียบและแปลความหมายได้ เช่น เด็กคนหนึ่งสอบได้ 25 คะแนน ไปเทียบกับเกณฑ์ปกติตรงกับตำแหน่งเปอร์เซ็นต์ไทล์ที่ 80 % เกณฑ์ปกติเปอร์เซ็นต์ไทล์ใช้ควบคู่กับเกณฑ์ปกติคะแนนมาตรฐานอื่น ๆ อยู่เสมอ เพราะแปลผลได้ง่าย เข้าใจได้ทุกคน ไม่สลับซับซ้อนมากนัก

2.2.2 เกณฑ์ปกติคะแนนที (T-score norms) นิยมใช้กันมากเพราะเป็นคะแนนมาตรฐานสามารถนำมาบวกลบและเฉลี่ยได้ มีค่าเหมาะสมในการแปลความหมาย คือมีค่า ตั้งแต่ 0 ถึง 100 มีคะแนนเฉลี่ย 50 ความเบี่ยงเบนมาตรฐาน 10

2.2.3 เกณฑ์ปกติสเตรน (Standard nine points) ค่าตั้งแต่ 1 ถึง 9 คะแนนเฉลี่ยอยู่ที่คะแนน 5 มีคะแนนเบี่ยงเบนมาตรฐานประมาณ 2 คะแนน วิธีการหามักจะเทียบจากเปอร์เซ็นต์ของความถี่ที่คะแนนเรียงตามค่าจะสะดวกกว่า

2.2.4 เกณฑ์ปกติตามอายุ (Age norms) แบบทดสอบมาตรฐานบางอย่างหาเกณฑ์ตามอายุเพื่อพัฒนาการในเรื่องเดียวกันว่า อายุต่างกันจะมีพัฒนาการอย่างไร โดยมากจะเป็นแบบทดสอบเชาว์ปัญญาและความถนัดจะหาเกณฑ์ปกติโดยวิธีนี้ ส่วนแบบทดสอบวัดผลสัมฤทธิ์จะหาเฉพาะแบบทดสอบที่เป็นวิชาพื้นฐาน เช่น ภาษา และคณิตศาสตร์ โดยเนื้อหาวิชาจะต้องไม่มีผลด้วยภาษา เช่น คำศัพท์ สามารถหาได้ตั้งแต่อายุ 5 ปี ถึง 20 ปี ความสามารถในการบวก ลบ คูณหาร ก็สามารถหาได้ในช่วงอายุดังกล่าวเหมือนกัน ทั้งนี้เพื่อจะดูว่าคำศัพท์ที่กำหนดไว้จำนวนหนึ่งนั้น ถ้านักเรียนคนหนึ่งอายุ 10 ปี สอบได้จำนวนหนึ่ง ลองไปเทียบกับเกณฑ์ปกติดูว่าน่าจะเป็นความสามารถคำศัพท์เท่าอายุเท่าไรแน่ อาจจะเท่ากับเด็กอายุ 8 ปี 10 ปี หรือ 15 ปี ก็ต้องเปรียบเทียบดูเกณฑ์แบบนี้วัดผลสัมฤทธิ์ใช้น้อยมากแต่จะทำให้เปรียบเทียบก็เป็นประโยชน์สำหรับการศึกษาได้

2.2.5 เกณฑ์ปกติตามระดับชั้น (Grade norms) เป็นการหาเกณฑ์ปกติตามระดับชั้นว่าคะแนนเท่าไรควรอยู่ระดับชั้นไหนจึงจะเหมาะสม แบบทดสอบที่จะทำเกณฑ์ปกติชนิดนี้ได้ก็ต้องเป็นเนื้อหาเดียวกัน ดังนั้นการวัดที่มีเนื้อหาแตกต่างกันตามระดับชั้นไม่สามารถที่จะหาเกณฑ์ปกติแบบนี้ได้ เพราะไม่สามารถจะอธิบายการแปลผลเปรียบเทียบได้ ดังนั้นวิชาที่ทำจึงเป็นวิชาพื้นฐานดังกล่าวแล้ว ในการสร้างเกณฑ์ปกติอายุนั้นเอง เช่น คำศัพท์ คณิตศาสตร์เบื้องต้นแบบทดสอบก็ต้องออกความรู้ความสามารถที่กว้าง ๆ เช่น คำศัพท์ก็ให้ครอบคลุมตั้งแต่ชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 แล้วศึกษาดูว่าระดับชั้นประถมศึกษาปีที่ 1 จะได้กี่คะแนน ปีที่ 2 จะได้กี่คะแนน ไปเรื่อย ๆ จนถึงชั้นมัธยมศึกษาปีที่ 6 จะได้กี่คะแนน โดยมากแต่ละระดับชั้นก็จะเป็นช่วงถือการแจกแจงของคะแนนจะซ้อนทับกันไปเป็นระยะไป แต่เมื่อสร้างเสร็จแล้ว ถ้าเด็กคนหนึ่งมาสอบแบบทดสอบฉบับนี้ได้คะแนน 20 คะแนน และกำลังเรียนชั้นมัธยมปีที่ 2 แต่เทียบแล้วเท่ากับนักเรียนชั้นประถมศึกษาปีที่ 6 จะได้นำไปพัฒนาต่อ

คะแนนที่ปกติ (Normalized T - score)

ภัทรานิคมานนท์ (2532 : 187-190) กล่าวว่า คะแนน T ปกติ เป็นคะแนนมาตรฐานชนิดหนึ่งที่เกิดมาจากคะแนนดิบ มีการแจกแจงความถี่ของคะแนนเป็นโค้งปกติ คะแนน T มีค่าเฉลี่ย (\bar{X}) เป็น 50 ความเบี่ยงเบนมาตรฐาน (Standard deviation) เป็น 10

วิธีการแปลงคะแนนดิบ ให้เป็นคะแนน T ปกติ หาได้โดยการแปลงคะแนนดิบให้เป็นตำแหน่งเปอร์เซ็นต์ไทล์เสียก่อน แล้วจึงเทียบเปอร์เซ็นต์ไทล์ให้เป็นคะแนน T ปกติ โดยดูว่าเปอร์เซ็นต์ไทล์นั้นเท่ากันหรือใกล้เคียงที่สุดกับค่าเปอร์เซ็นต์ไทล์ใด การแปลงคะแนนดิบเป็นคะแนนที่ปกติ มีความสะดวกกว่าการแปลงเป็นคะแนนที่แนวเส้น ($T = 10Z + 50$) เพราะไม่ต้องการคำนวณหาค่าเฉลี่ย (\bar{X}) ความเบี่ยงเบนมาตรฐาน (S.D) คะแนนซี (Z) เราสามารถแปลงคะแนนได้ครั้งละจำนวนมากไม่เหมือนคะแนนที่แนวเส้น ซึ่งต้องแปลทีละจำนวน

บุญเรือง ขจรศิลป์ (2527 : 112-114) กล่าวว่า คะแนนมาตรฐาน T ปกติ (Normalized T - score) คือ คะแนนมาตรฐานที่ได้จากข้อมูล ที่มีการกระจายเป็นโค้งปกติ หรือโดยการปรับให้เป็นโค้งปกติ การเปลี่ยนคะแนนดิบ ให้เป็นคะแนนมาตรฐาน T โดยใช้สูตร $T = 50 + 10Z$ นั้น เป็นการเปลี่ยนในลักษณะ ที่เรียกว่า Linear transformation ซึ่งการเปลี่ยนแปลงคะแนนแบบนี้ไม่ทำให้ลักษณะของการกระจายคะแนนเดิมเปลี่ยนไป เมื่อเปลี่ยนเป็น T - score แล้ว ลักษณะของการกระจายจะยังคงเหมือนกับการกระจายของคะแนนดิบ แต่การเปลี่ยนคะแนนดิบ ให้เป็นคะแนนมาตรฐานปกติ นั้น เป็นการเปลี่ยนในลักษณะที่เรียกว่า Area transformation ซึ่งการเปลี่ยนแปลงคะแนนแบบนี้ เมื่อเปลี่ยนมาเป็น คะแนน T ปกติแล้ว จะทำให้การกระจายมีลักษณะเป็นโค้งปกติ

สำหรับงานวิจัยครั้งนี้ได้สร้างเกณฑ์ปกติระดับท้องถิ่น (Local norms) ซึ่งเป็นเกณฑ์ปกติระดับจังหวัด เพื่ออธิบายถึง ข้อเท็จจริงทางสถิติที่บรรยายการแจกแจงของคะแนนของกลุ่มตัวอย่าง ซึ่งได้จากแบบทดสอบวินิจัย และเป็นคะแนนตัวแทนที่บอกระดับการสอบของผู้สอบว่าอยู่ในระดับใดของกลุ่มประชากร โดยแสดงลงในตารางซึ่งบอกความสัมพันธ์ระหว่างคะแนนดิบกับคะแนนที่ปกติ (Normalized T - score) ที่แปลงมาจากคะแนนดิบ

ทฤษฎีการตอบสนองข้อสอบ

1. แนวคิดพื้นฐานของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างความสามารถที่มีอยู่ภายในบุคคล (Latent trait or ability) กับผลการตอบข้อสอบหรือข้อคำถาม โดยใช้โค้งลักษณะข้อสอบ (Item Characteristic Curve ; ICC) ซึ่งมีการกำหนดคัลักษณะของข้อสอบด้วยพารามิเตอร์ความยาก (b) อำนาจจำแนก (a) และโอกาสการเดาข้อสอบถูก (c) IRT จึงอยู่บนฐานความคิด 2 ประการ 1) ผลการตอบข้อสอบหรือข้อคำถามของผู้ตอบ สามารถอธิบาย

ได้ด้วยความสามารถที่มีอยู่ภายในของผู้ตอบ และ 2) ความสัมพันธ์ระหว่างผลการตอบข้อสอบ กับความสามารถที่มีอยู่ภายใน สามารถอธิบายได้ด้วยฟังก์ชันลักษณะข้อสอบ หรือโค้งลักษณะข้อสอบ (ICC) อันมีลักษณะเป็นฟังก์ชันทางคณิตศาสตร์ เรียกว่าฟังก์ชันโลจิส (Logistic function) หรือ โค้งเดียวกับฟังก์ชันปกติสะสม (Normal ogive function)

ทฤษฎีการตอบสนองข้อสอบ พยายามอธิบายความสัมพันธ์ระหว่างคุณลักษณะภายใน หรือความสามารถที่มีอยู่ภายในตัวบุคคล กับพฤติกรรมตอบสนองข้อสอบของบุคคลนั้นว่ามีโอกาสตอบข้อสอบถูกมากน้อยเพียงไร ทฤษฎีนี้มีพื้นฐานความเชื่อว่า พฤติกรรมตอบสนองต่อของข้อสอบของผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่มีอยู่ภายในตัวบุคคล ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง ทฤษฎีนี้ได้อธิบายความสัมพันธ์ดังกล่าวในรูปของฟังก์ชันคณิตศาสตร์ หรือโมเดลที่แสดงความสัมพันธ์ระหว่างระดับความสามารถ คุณลักษณะของข้อสอบ และโอกาสของการตอบข้อสอบได้ถูก ที่เรียกว่า ฟังก์ชันการตอบสนองข้อสอบ ซึ่งมีลักษณะความสัมพันธ์เป็นแบบฟังก์ชันโลจิสหรือฟังก์ชันปกติสะสม

ฟังก์ชันการตอบสนองข้อสอบสามารถนำมาใช้ศึกษาความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบแต่ละข้อได้ถูก $[P_i(\theta)]$ กับระดับความสามารถของผู้สอบที่วัดได้โดยแบบสอบฉบับนั้น (θ) เมื่อนำมาเขียนเป็นกราฟจะได้โค้งลักษณะข้อสอบ (Item Characteristic Curve ; ICC) โค้งลักษณะข้อสอบมีหลายลักษณะ ขึ้นอยู่กับโมเดล (Model) หรือแบบจำลองที่ใช้อธิบายความสัมพันธ์ ดังกล่าว โมเดลที่นิยมใช้กันคือ โมเดลแบบหนึ่งพารามิเตอร์ (One-Parameter Model) โมเดลแบบสองพารามิเตอร์ (Two-Parameter Model) โมเดลแบบสามพารามิเตอร์ (Three-Parameter Model) (ศิริชัย กาญจนวาสี, 2545 : 45-46)

2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีข้อตกลงเบื้องต้น ดังนี้

2.1 แบบทดสอบมิติเดียว (Unidimensionality test) หมายความว่าข้อสอบแต่ละข้อในแบบทดสอบจะต้องวัดความสามารถหรือคุณลักษณะ (Trait) ได้เพียงคุณลักษณะเดียวเท่านั้น หรือมีความเป็นเอกพันธ์กัน

2.2 ข้อสอบแต่ละข้อจะต้องเป็นอิสระจากกัน (Local Independence) หมายความว่า พฤติกรรมตอบสนองข้อสอบข้อใดข้อหนึ่งของบุคคลใดบุคคลหนึ่งนั้น ต้องมีความเป็นอิสระเชิงสถิติ (Statistically Independence) กล่าวคือ การตอบข้อสอบข้อหนึ่งไม่มีผลต่อการตอบสนองข้อสอบ

ข้ออื่น ๆ ในแบบสอบ เช่น เนื้อหาของคำถามข้อหนึ่ง ต้องไม่ชี้แนะคำตอบให้แก่ข้ออื่น ๆ มีความเป็นอิสระจากตำแหน่ง (Uncorrelated Independence) กล่าวคือ ข้อสอบแต่ละข้อปรากฏอยู่ในตำแหน่งใดของแบบทดสอบก็ได้ จะไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ

2.3 โค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) ของแต่ละโมเดลเป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบข้อนั้นอยู่กับระดับความสามารถที่วัดได้โดยข้อสอบข้อนั้น โค้งลักษณะข้อสอบมีหลายรูปแบบทั้งนี้ขึ้นอยู่กับความเชื่อในแบบจำลองที่ใช้อธิบายความสัมพันธ์ดังกล่าว

3. รูปแบบที่ใช้ในทฤษฎีการตอบสนองข้อสอบ

เนื่องจากข้อมูลที่ได้จากผู้สอบมีหลายลักษณะ ดังนั้นจึงได้มีการพัฒนารูปแบบที่แสดงความสัมพันธ์ตามแนวคิดทฤษฎีการตอบสนองข้อสอบในหลายลักษณะต่าง ๆ กันเพื่อให้แต่ละรูปแบบมีความเหมาะสมกับข้อมูลที่ได้ ซึ่งสามารถจำแนกได้ดังนี้

3.1 แบบการตอบสนองข้อสอบที่ใช้กับการตรวจให้คะแนนรายข้อแบบ 2 ค่า (Dichotomous) คือข้อสอบที่ตรวจคะแนนเป็น 0 เมื่อตอบผิด ให้คะแนนเป็น 1 เมื่อตอบถูก ในระยะเริ่มแรกมีการพัฒนารูปแบบประเภทนี้ ซึ่งได้แก่ รูปแบบของกัตแมนสมบูรณ์ (Guttman Perfect Scale) รูปแบบระยะห่างแฝง (Latent Distance Model) รูปแบบเชิงเส้นตรง (Linear Model) เป็นต้น ในระยะต่อมาได้มีการพัฒนารูปแบบประเภทนี้เพิ่มขึ้นเช่นรูปแบบปกติสะสม (Normal Ogive Model) แบบ 1, 2, 3 พารามิเตอร์ และรูปแบบโลจิสติก (Logistic Model) แบบ 1, 2, 3, 4 พารามิเตอร์ เป็นต้น ซึ่งเป็นรูปแบบที่ได้รับการปรับปรุงให้มีความสอดคล้องกับข้อมูลจากการสอบในสถานการณ์จริงมากขึ้น สามารถนำไปปฏิบัติได้ง่ายและมีผู้สนใจนำไปใช้อย่างกว้างขวาง

3.2 รูปแบบการตอบสนองข้อสอบที่ใช้การตรวจให้คะแนนรายข้อแบบมากกว่า 2 ค่า (Polytomous) ได้แก่รูปแบบการตอบสนองข้อสอบแบบนามบัญญัติ (Normal Response Model : PCM) รูปแบบการให้คะแนนบางส่วน (Partial Credit Model : PCM) ซึ่งรูปแบบประเภทนี้มีจุดมุ่งหมายเพื่อเพิ่มความแม่นยำในการประมาณค่าความสามารถของผู้สอบ โดยใช้ประโยชน์จากค่าสารสนเทศทั้งจากการตอบข้อสอบถูกหรือผิด

3.3 รูปแบบการตอบสนองข้อสอบที่ใช้การตรวจให้คะแนนรายข้อแบบต่อเนื่อง (Contineous) ได้แก่รูปแบบการตอบสนองข้อสอบแบบต่อเนื่อง (Contineous Response Model : CRM) ซึ่งเป็นประโยชน์ต่อการประมาณค่าความสามารถของผู้เข้าสอบทางจิตวิทยา หรือผู้สนใจศึกษาทางด้านที่เกี่ยวกับทัศนคติ (Attitude) ซึ่งการตอบสนองข้อสอบเป็นมาตราที่ต่อเนื่อง (Continecus Scale)

โดยสามารถจะสรุปรูปแบบต่าง ๆ ที่มีความเหมาะสมกับลักษณะข้อมูลดังกล่าวข้างต้น และผู้ที่นำเสนอรูปแบบนั้น ๆ ได้ตามตารางต่อไปนี้

ตาราง 2 ลักษณะของข้อมูล รูปแบบ (Model) ที่ใช้วิเคราะห์ และผู้นำเสนอรูปแบบ

ลักษณะของข้อมูล	รูปแบบที่ใช้วิเคราะห์	ผู้นำเสนอรูปแบบ
Dichotomous	Latent Linear	Lazarsfeld & Henry (1968)
	Latent Distance	
	Perfact scale	
	One, Two, Three Parameter Normal Ogive	Guttman (1944)
	One, Two, Three Parameter Logistic	Lord (1952)
	Four Parameter Logistic	Birnbaum(1975,1958a, 1958b,1968) LORD& Novick (1968) Wright & Stone (1979) McDonald (1967)
Polytomous	Nomal Response	Barton & Lord (1981)
	Graded Response	Bock (1972)
	Partial Credit Model	Samejima (1969)
Contineous	Contineous Response	Master (1982)
		Samejima (1972)

ในการวิจัยครั้งนี้ ผู้วิจัยจะกล่าวถึงเฉพาะรูปแบบโลจิสติก (Logistic Model) เท่านั้น

4. รูปแบบโลจิสติก (Logistic Model)

รูปแบบโลจิสติก เป็นรูปแบบที่มีลักษณะใกล้เคียงกับรูปแบบปกติของอองไวร์ (Normal Ogive Model) แต่มีความสะดวกในการประมาณค่ามากกว่า และยังได้พัฒนาให้มีความเหมาะสมกับสถานการณ์จริงในการทดสอบอีกด้วย เช่น ในกรณีที่ตอบข้อสอบได้ด้วยการเดา ปลายโค้งลักษณะข้อสอบ (ICC) ไม่เป็นศูนย์ แสดงว่า การทำข้อสอบย่อมมีการเดา และในกรณีที่ผู้ทำข้อสอบด้วยความสุ่ม เพราะ รูปแบบนี้จะเพิ่มค่าความรอบคอบขึ้นอีกหนึ่งค่า ซึ่งสามารถพิจารณาได้จากปลายสูงสุดของโค้งลักษณะข้อสอบมีค่าน้อยกว่า 1

รูปแบบโลจิสติกแบ่งเป็น 3 รูปแบบตามจำนวนของพารามิเตอร์ โค้งลักษณะข้อสอบดังนี้ (Hambleton and Swaminathan, 1987 : 34-49, อ้างถึงใน สิริภพ กาฬสุวรรณ, 2546 : 22)

4.1 รูปแบบโลจิสติกที่มี 1 พารามิเตอร์ (One Parameter Logistic Model)

เบิร์นบอม (Birnbaum) ได้พัฒนารูปแบบนี้ในปี ค.ศ. 1968 เป็นรูปแบบที่อธิบายคุณลักษณะของข้อสอบด้วยค่าพารามิเตอร์เพียงตัวเดียว คือ ค่าความยากของข้อสอบ (b) โดยรูปแบบนี้เชื่อค่าความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบได้ถูกต้องหรือไม่ ขึ้นอยู่กับระดับความสามารถของผู้สอบ (θ) และค่าความยากของข้อสอบ โดยค่าการเดาของข้อสอบ (c) มีค่าน้อยที่สุดซึ่งถือว่ามีค่าเป็นศูนย์ และค่าอำนาจจำแนกของข้อสอบ (a) มีค่าคงที่ที่บังคับ ซึ่งรูปแบบนี้มีความเหมือนกับรูปแบบของราส์ช (Rasch Model) เขียนเป็นสมการได้ดังนี้

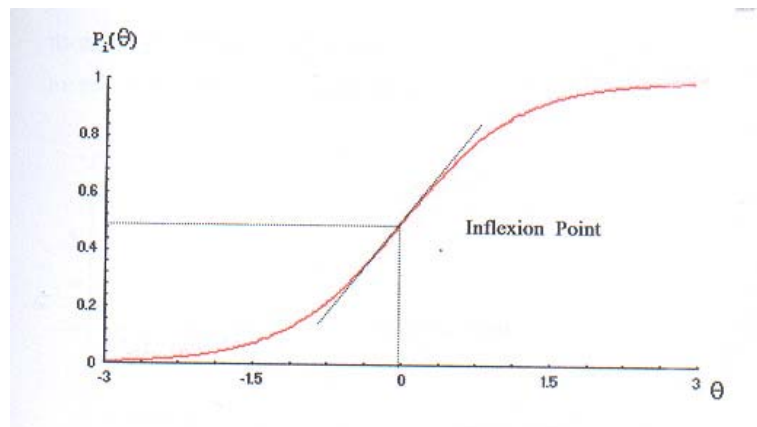
$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} ; i = 1, 2, 3, \dots, n$$

เมื่อ $P_i(\theta)$ หมายถึง ระดับความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ θ จะทำข้อสอบที่ i ถูก

θ หมายถึง ระดับความสามารถที่แท้จริงของผู้สอบ

b_i หมายถึง ค่าความยากของข้อสอบข้อที่ i

e หมายถึง ค่าคงที่มีค่าเท่ากับ 2.7182818



ภาพประกอบ 1 แสดงความหมายของค่า Item Parameter (1 พารามิเตอร์) ของข้อสอบ

4.2 รูปแบบโลจิสติกที่มี 2 พารามิเตอร์ (Two-Parameter Logistic Model)

เบียร์นบอม (Birnbaum) ได้พัฒนารูปแบบนี้ขึ้นมาจาก Normal Ogive Model ซึ่งรูปแบบใหม่มีรูปแบบไม่แตกต่างจากเดิมมากนัก จะเปลี่ยนแต่เพียงการคิดคำนวณเท่านั้น และกำหนดให้ทุกข้อไม่มีการเดาเกิดขึ้น คือ c_i มีค่าเป็นศูนย์ทุกข้อ กล่าวคือ ผู้สอบที่มีความสามารถต่ำสุดไม่มีโอกาสที่จะทำข้อสอบที่มีค่าความยากสูง ซึ่งเบียร์นบอม ได้เสนอรูปของสมการดังนี้

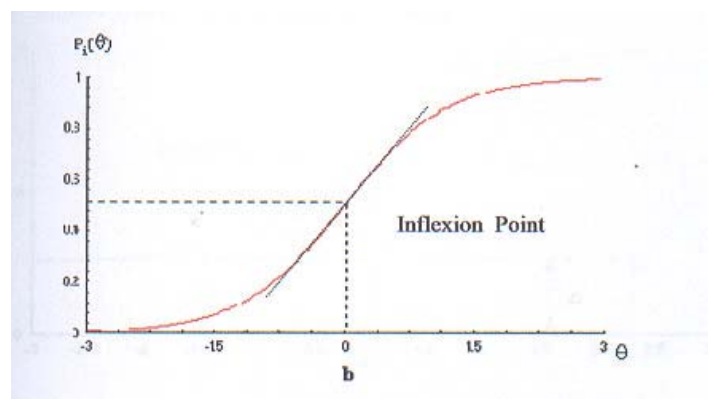
$$P_i(\theta) = \frac{e^{D_{a_i}(\theta - b_i)}}{1 + e^{D_{a_i}(\theta - b_i)}} ; i = 1, 2, 3, \dots, n$$

เมื่อ $P_i(\theta)$ หมายถึง ความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ θ จะทำข้อสอบข้อที่ i ถูก

θ หมายถึง ระดับความสามารถที่แท้จริงของผู้สอบ

a_i หมายถึง ค่าอำนาจจำแนกของข้อสอบข้อที่ i ซึ่งสัดส่วนโดยตรงกับความชันของโค้ง ณ จุดเปลี่ยนโค้งและมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ตามปกติมีค่าไม่เกิน +2.50 ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่า a_i อยู่ระหว่าง +0.50 ถึง +2.50

- b_i หมายถึง ค่าความยากของข้อสอบข้อที่ i มีค่าเท่ากับระดับความสามารถของผู้สอบ ณ จุดเปลี่ยน โคน และมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่า b_i อยู่ระหว่าง -2.50 ถึง 2.50 ค่า b_i ที่อยู่ใกล้ -2.50 แสดงว่าเป็นข้อสอบที่ง่าย ส่วนค่า b_i ที่อยู่ใกล้ $+2.50$ แสดงว่าเป็นข้อสอบที่ยาก
- e หมายถึง ค่าคงที่มีค่าเท่ากับ 2.71828818
- D หมายถึง ค่าคงที่ โดยปกติมีค่า 1.702 หรือ 1.7



ภาพประกอบ 2 แสดงความหมายของค่า Item Parameter (2 พารามิเตอร์) ของข้อสอบ

4.3 รูปแบบโลจิสติกที่มี 3 พารามิเตอร์ (Three-Parameter Logistic Model)

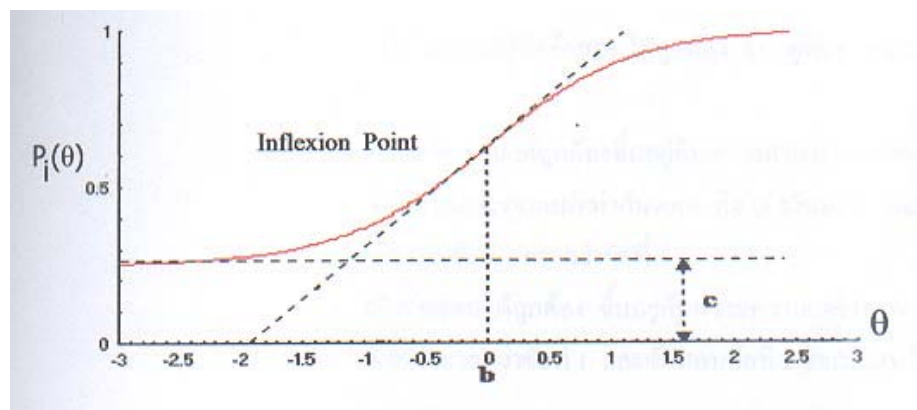
เป็นรูปแบบที่พัฒนามาจาก Two-Parameter Logistic Model เพื่อเหมาะกับแบบสอนที่มีอิทธิพลจากการเดาเข้ามาแฝงอยู่ด้วย และเป็นโค้งลักษณะข้อสอบที่แสดงถึงลักษณะข้อสอบที่มีค่าพารามิเตอร์ของข้อสอบ 3 ตัว ซึ่งเบิร์ตบอม ได้เสนอรูปแบบของสมการดังนี้

$$P_i(\theta) = C_i + (1 - C_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} ; i = 1, 2, 3, \dots, n$$

เมื่อ $P_i(\theta)$ หมายถึง ความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ θ จะทำข้อสอบข้อที่ i ถูก

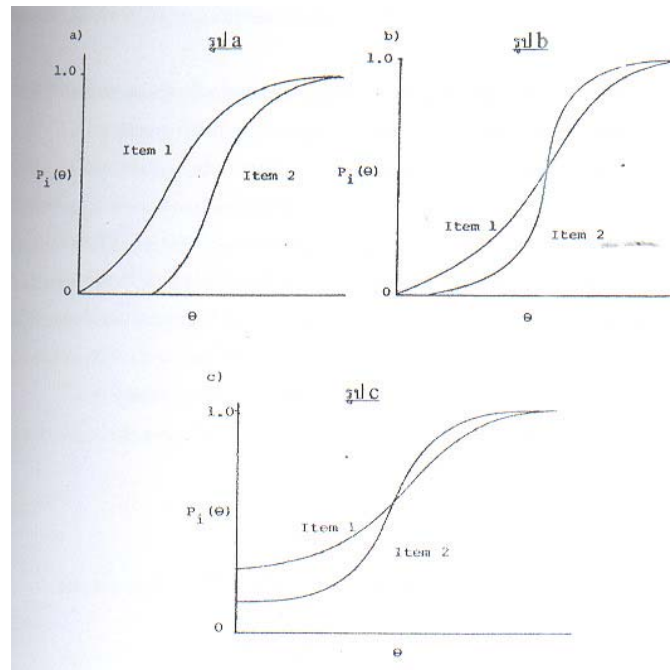
θ หมายถึง ระดับความสามารถที่แท้จริงของผู้สอบ

- a_i หมายถึง ค่าอำนาจจำแนกของข้อสอบข้อที่ i ซึ่งสัดส่วนโดยตรงกับความชันของโค้ง θ เปลี่ยนโค้งและมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ตามปกติมีค่าไม่เกิน +2.50 ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่า a_i อยู่ระหว่าง +0.50 ถึง +2.50
- b_i หมายถึง ค่าความยากของข้อสอบข้อที่ i มีค่าเท่ากับระดับความสามารถของผู้สอบ θ จุดเปลี่ยนโค้ง และมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่า b_i อยู่ระหว่าง -2.50 ถึง 2.50 ค่า b_i ที่อยู่ใกล้ -2.50 แสดงว่าเป็นข้อสอบที่ง่าย ส่วนค่า b_i ที่อยู่ใกล้ +2.50 แสดงว่าเป็นข้อสอบที่ยาก
- c_i หมายถึง ค่าการเดาของข้อสอบที่ i เป็นความน่าจะเป็นหรือโอกาสของคนที่มีความสามารถต่ำจะตอบข้อสอบถูกมีค่าอยู่ระหว่าง 0 ถึง 1 โดยทั่วไปแล้วข้อสอบที่ดีจะต้องมีค่าการเดาต่ำกว่า 0.30
- e หมายถึง ค่าคงที่มีค่าเท่ากับ 2.7182818
- D หมายถึง ค่าคงที่ โดยปกติมีค่า 1.702 หรือ 1.7



ภาพประกอบ 3 แสดงความหมายของค่า Item Parameter (3 พารามิเตอร์) ของข้อสอบ

จากรูปแบบโลจิสติกทั้ง 3 รูปแบบ สามารถเขียนแสดงได้ดังรูปต่อไปนี้
(Hambleton and Cook, 1977 : 79, อ้างถึงใน ประดิษฐ์ เรื่องตระกูล, 2529 : 21-22)



ภาพประกอบ 4 แสดงโค้งลักษณะข้อสอบของรูปแบบโลจิสติก 1,2,3 พารามิเตอร์

จากภาพประกอบ 4 แสดงถึงโอกาสที่ผู้ทำข้อสอบได้ถูกต้อง ขึ้นอยู่กับค่าพารามิเตอร์ได้แก่

รูป a แสดงถึงโอกาสที่จะตอบข้อสอบได้ถูกต้องขึ้นอยู่กับความสามารถของข้อสอบเพียงอย่างเดียว โดยถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันหมด คือ เท่ากับหนึ่ง และค่าการเดาเท่ากับศูนย์ ในรูปแสดงให้เห็นว่าข้อสอบที่ 2 ยากกว่าข้อที่ 1

รูป b แสดงถึงโอกาสที่จะทำข้อสอบได้ถูกต้อง ขึ้นอยู่กับความยากและค่าอำนาจจำแนก ในรูปแสดงให้เห็นว่าข้อสอบข้อที่ 2 ยากกว่าข้อที่ 1 และข้อสอบข้อที่ 2 จำแนกคนได้ดีกว่าข้อที่หนึ่ง

รูป c แสดงถึงโอกาสที่จะทำข้อสอบได้ถูกต้องขึ้นอยู่กับอำนาจจำแนกค่าความยากและค่าการเดา ในรูปแสดงให้เห็นว่าข้อสอบที่ 1 ยากกว่าข้อ 2 และข้อสอบข้อที่ 2 จำแนกคนได้ดีกว่าและมีค่าการเดาน้อยกว่าข้อที่ 1

5. ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบทดสอบ (Test and Item Information Functions)

ในการประมาณค่าความสามารถของผู้สอบด้วยวิธีแมกซิมัมไลคิฮูด (Maximum likelihood) นั้นความแน่นอนของการประมาณค่าความสามารถแสดงได้ในเทอมของค่าฟังก์ชันสารสนเทศ (Information Functions) โดยที่ในทฤษฎีการทดสอบดั้งเดิมนั้นเราศึกษาเรื่องความเที่ยง (Reliability) ของคะแนนและความคลาดเคลื่อนมาตรฐานของการวัด (Standard Error Measurement) ซึ่งค่าที่จะได้แปรเปลี่ยนไปตามกลุ่มผู้สอบดังเป็นจุดอ่อนประการหนึ่ง แต่ในทฤษฎีการตอบสนองข้อสอบจะศึกษาถึงค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Functions) แทนการหาค่าความเที่ยง (Hambleton, 1979 : 64)

สำหรับ โมเดลโลจิสติก 3 พารามิเตอร์ จะสามารถหาค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information Functions) และค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Functions) ในแต่ละระดับความสามารถ (θ) ได้ดังนี้

5.1 ค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information Functions) ในแต่ละระดับความสามารถคำนวณได้จากสมการต่อไปนี้ (Birbaum, 1968 quoted in Hambleton, Swaminathan and Rogers, 1991 : 91)

$$I_i(\theta) = \frac{(P_i'(\theta))^2}{P_i(\theta)Q_i(\theta)} ; I = 1, 2, \dots, n$$

เมื่อ $I_i(\theta)$	คือ	ค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information Functions)
$P_i'(\theta)$	คือ	ความชันของ ICC ที่ระดับความสามารถ θ
$P_i(\theta)$	คือ	ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะต้องตอบข้อสอบข้อที่ i ถูก
$Q_i(\theta)$	คือ	$1 - P_i(\theta)$

หรือสามารถหาฟังก์ชันสารสนเทศของสมการได้จากสมการดังนี้

$$I_i(\theta) = \frac{2.89a_i^2(1-c_i)}{[c_i + e^{1.7a_i(\theta-b_i)}][1 + e^{1.7a_i(\theta-b_i)}]}$$

5.2 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Functions) สามารถหาได้โดยการหาผลรวมของค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information Functions) แต่ละข้อคำนวณได้จากสมการต่อไปนี้ (Birnbaum, 1968 quoted in Hambleton, Swaminathan and Rogers, 1991 : 94)

$$I(\theta) = \sum_{i=1}^n I_i(\theta)$$

เมื่อ $I(\theta)$ คือ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Functions)
 n คือ จำนวนข้อสอบในแบบทดสอบ

ค่าฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อจะขึ้นอยู่กับโค้งลักษณะของข้อสอบ (ICC) และความแปรปรวนของการตอบข้อสอบถูกในแต่ละข้อของแต่ละระดับความสามารถ และยังความชันของโค้งลักษณะของข้อสอบมีค่ามาก ๆ ประกอบกับความแปรปรวนมีค่าน้อย ๆ โค้งฟังก์ชันสารสนเทศของข้อสอบที่ระดับความสามารถนั้นจะยิ่งสูงขึ้น สำหรับโค้งฟังก์ชันสารสนเทศของข้อสอบ (Item Information Functions) ที่มีค่าสูงสุด ณ ระดับความสามารถใดก็จะจำแนกระดับความสามารถของผู้สอบได้ดี ณ ระดับความสามารถนั้น (Hambleton, 1979 : 66)

6. การประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ

ปัจจุบันได้มีการนำเอาทฤษฎีการตอบสนองข้อสอบ โดยเฉพาะรูปแบบของราสค์ (Rasch Model) มาใช้อย่างกว้างขวางในการวัดผลการทดสอบโดยทั่วไปซึ่งไรต์ (Wright, 1890 : 194-196 อ้างถึงใน ประดิษฐ์ เรื่องตระกูล, 2529 : 29) ได้รวบรวมประโยชน์ถึงการนำรูปแบบของราสค์ (Rasch Model) มาใช้ไว้มากมาย แต่อย่างไรก็ตามในงานที่ (Rasch Model) สามารถทำได้นั้น โมเดลโลจิสติก 3 พารามิเตอร์ย่อมทำได้และอาจทำได้ดีกว่า ในบางกรณีดังนั้นการนำรูปแบบโลจิสติก 3 พารามิเตอร์ไปใช้ประโยชน์จึงพอสรุปได้ดังนี้

6.1 ใช้ในการสร้างคลังข้อสอบ (Item Bank) เนื่องจากข้อสอบที่วิเคราะห์แล้วค่าพารามิเตอร์มีลักษณะคงที่ ดังนั้น ข้อสอบเหล่านี้จึงนำมาใช้สร้างข้อสอบชุดใหม่ตามเกณฑ์ที่ต้องการวัด

6.2 ใช้วิเคราะห์ข้อสอบ (Item Analysis) การวิเคราะห์ข้อสอบด้วยรูปแบบโลจิสติก 3 พารามิเตอร์ จะสามารถแก้ปัญหาที่รูปแบบทฤษฎีการทดสอบแบบดั้งเดิมแก้ไม่ได้ (Hambleton, 1979 : 14-15) อันได้แก่

6.2.1 ค่าสถิติและค่าพารามิเตอร์ต่างขึ้นอยู่กับสภาพกลุ่มตัวอย่างที่ทำการสอบ

6.2.2 การเปรียบเทียบความสามารถหรือคุณลักษณะ (Trait) ใด ๆ ในแต่ละบุคคล จะเปรียบเทียบกันได้ก็ต่อเมื่อต้องสอบด้วยแบบทดสอบฉบับเดียวกัน ความคลาดเคลื่อนมาตรฐานในการวัดของแต่ละคนเท่ากัน ซึ่งลอร์ด และ โนวิค (Lord and Novick, 1968 : 15) ได้แสดงว่า ความคลาดเคลื่อนมาตรฐานในการวัดสำหรับคะแนนปานกลางจะมีขนาดเล็กกว่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนสูงหรือคะแนนต่ำ

6.3 ใช้ในการกำหนดเกณฑ์ความสามารถของแบบทดสอบแบบอิงเกณฑ์ ผลของการวิเคราะห์ข้อสอบถูกคัดแปลงให้เป็นคะแนนความสามารถของแต่ละกลุ่มบุคคลจึงสามารถเปรียบเทียบกับคะแนนความสามารถซึ่งเป็นเกณฑ์คงที่ได้ ทำให้เราทราบได้ว่าระดับความสามารถอิงเกณฑ์แต่ละชุด (Minimum Mastery Level) ของข้อสอบแต่ละข้อควรเป็นอย่างไร

6.4 ใช้วินิจฉัยความสามารถของผู้สอบ (Diagnostic) ในกรณีที่ ICC ของข้อสอบไม่เหมาะสมของโค้งของโมเดล แสดงว่าบางสิ่งบางอย่างผิดปกติในตัวผู้สอบที่เราควรสนใจและแก้ไข

6.5 ใช้ในการค้นหาความเป็นอคติของข้อสอบ (Item Bias) เมื่อเกิดเหตุดังกล่าวในข้อ 6.3 แสดงว่ามีความสามารถอื่นแฝงเข้ามาในความสามารถที่ต้องการวัด เราก็สามารถที่จะทำการตรวจสอบความลำเอียงของข้อสอบได้

6.6 ใช้ในการวัดระดับความสามารถของแต่ละบุคคล (Tailoring Test) เราอาจหุ้มข้อสอบที่วิเคราะห์แล้วและมีความยากแยกตามลำดับมาสอบวัดระดับความสามารถของแต่ละบุคคลได้

6.7 ใช้ในการจัดชั้นเรียน (Grade-Placement Tailoring) ค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบอาจใช้ในการจัดชั้นเรียนที่เหมาะสมกับระดับความสามารถของผู้เรียนเป็นกลุ่ม ๆ ได้

6.8 ใช้ในการสร้างแบบทดสอบที่ดีที่สุด (Best Test Design) ผลจากการวิเคราะห์ข้อสอบ โดย Latent Trait Models สามารถนำไปใช้ในการสร้างข้อสอบที่มีคุณลักษณะต่าง ๆ ตามที่ต้องการได้

6.9 ใช้ในการเปรียบเทียบคะแนนต่างชุด (Equating Score) ใช้แบบสอบที่วิเคราะห์แล้ว 2 ชุดที่ต่างกัน แต่วัดในสิ่งเดียวกัน จะสามารถนำคะแนนของผู้สอบในแบบสอบฉบับหนึ่งไปเปรียบเทียบกับคะแนนของผู้สอบในแบบสอบอีกฉบับหนึ่งได้ทั้งนี้เพราะคะแนนแต่ละชุดของทฤษฎี IRT นั้น ถูกแปลงเป็นคะแนนมาตรฐานที่สามารถเปรียบเทียบกันได้

6.10 ใช้หาค่าการเดา (Guessing) ในทฤษฎีการทดสอบดั้งเดิมได้มีการใช้สูตรแก้การเดา แต่ในทางปฏิบัติจริงแล้วถ้าผู้สอบมีเวลาเพียงพอ มีตัวเลือกที่มีประสิทธิภาพและผู้สอบได้แสดงความสามารถอย่างเต็มที่ สำหรับในทฤษฎีการตอบสนองข้อสอบจะไม่ใช้สูตรการเดาแต่จะใช้การพิจารณาค่าการเดา ซึ่งหาได้จากการนำรูปแบบทางคณิตศาสตร์เข้ามาช่วยสนับสนุนทำให้เกิดหลักการพิจารณาค่าการเดามีความสมเหตุสมผลยิ่งขึ้น

7. สรุปแนวความคิดทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นทฤษฎีการทดสอบแนวใหม่ ซึ่งเป็นที่ยอมรับและเป็นแนวคิดกระแสหลักของศาสตร์ทางการวัดผลในปัจจุบัน ทฤษฎีการตอบสนองข้อสอบสามารถจำแนกได้เป็น 2 ประเภท ได้แก่ ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Dichotomous IRT) และทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous IRT)

ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า อธิบายความน่าจะเป็นของการตอบข้อสอบได้ถูกต้อง ขึ้นอยู่กับคุณลักษณะภายในหรือความสามารถของผู้ตอบ (θ) และคุณลักษณะของข้อสอบ ที่แสดงได้ในรูปของฟังก์ชันการตอบสนองข้อสอบหรือโค้งลักษณะข้อสอบ (ICC) ซึ่งมีได้หลายลักษณะ ขึ้นอยู่กับโมเดลการตอบสนองข้อสอบ ได้แก่ โมเดลการตอบสนองข้อสอบแบบ 1-พารามิเตอร์ [ข้อสอบแตกต่างกันเฉพาะค่าความยาก (พารามิเตอร์ b)] โมเดลการตอบสนองข้อสอบ 2-พารามิเตอร์ [ข้อสอบมีคุณลักษณะแตกต่างกันตามค่าอำนาจจำแนก (พารามิเตอร์ a) และค่าความยาก (พารามิเตอร์ b)] และโมเดลการตอบสนองข้อสอบแบบ 3-พารามิเตอร์ [ข้อสอบมีคุณลักษณะแตกต่างกันตามค่าอำนาจจำแนก (พารามิเตอร์ a) ค่าความยาก (พารามิเตอร์ b) และโอกาสการเดาข้อสอบถูก (พารามิเตอร์ c)] ซึ่งนักพัฒนาแบบสอบจะต้องเลือกใช้ให้สอดคล้องกับสถานการณ์ของการทดสอบและธรรมชาติของข้อมูล

ค่าพารามิเตอร์ของข้อสอบในทฤษฎีการตอบสนองข้อสอบแต่ละตัวมีความหมายดังนี้

θ : คือระดับความสามารถของผู้สอบในคะแนนสเกลมาตรฐาน อยู่ระหว่าง -3 ถึง +3 โดยประมาณ ผู้ที่มีความสามารถปานกลางจะมีค่า θ อยู่ระหว่าง -1.0 ถึง +1.0 ถ้าค่า θ มากกว่า 1.0 แสดงว่าค่าความสามารถค่อนข้างสูง ถ้าน้อยกว่า -1.0 แสดงว่าค่าความสามารถค่อนข้างต่ำ

a : คือค่าความชัน (Slope) หรือ อำนาจจำแนก (Item Discriminating) ของสมการที่จับขอบโค้งที่ตรงกับค่าโอกาสที่จะทำข้อสอบถูกเท่ากับ 0.5 (ค่าแกนตั้งที่ 0.5) เป็นค่าพารามิเตอร์ที่แสดงถึงความสามารถของข้อสอบในการจำแนกบุคคลที่มีความสามารถต่างไปออกได้มากน้อยเพียงใด

ถ้าค่าความชันสูงคนที่มีความสามารถต่างกันเพียงเล็กน้อย ก็ส่งผลให้โอกาสที่ทำข้อสอบถูกต่างกันมาก ถ้าข้อสอบใดมีค่า a มากกว่า 0.5 หมายความว่าข้อสอบข้อนั้นมีการจำแนกอยู่ในระดับที่ดี ถ้าค่า a มีค่าตั้งแต่ 0.3 ถึง 0.5 หมายความว่าข้อสอบข้อนั้นมีการจำแนกได้ในระดับปานกลาง ถ้าค่า a มีค่าน้อยกว่า 0.3 หมายความว่า ข้อสอบข้อนั้นมีการจำแนกไม่ดี

b : คือค่าความยาก (Item Difficulty) ก็คือตำแหน่งบนสเกลของค่าความสามารถ (แกนนอน) ที่ตรงกับ ค่าโอกาสที่จะทำข้อสอบถูกเท่ากับ 0.5 นั้น ค่าความยากก็คือ ค่าความสามารถของผู้สอบที่มีโอกาสที่จะทำข้อสอบข้อนั้นถูกร้อยละ 50 ค่าความยากเป็นค่าที่บอกถึงความเหมาะสมในการที่จะเอาข้อสอบข้อนี้ไปใช้ ถ้าค่าความยากสูงกว่า 1.0 ก็ควรนำไปใช้กับผู้มีความสามารถสูง (กลุ่มเก่ง) เช่น นำไปใช้ในการสอบแข่งขัน ถ้าค่าความยากต่ำกว่า -1.0 ก็ควรนำไปใช้กับผู้เรียนที่มีความสามารถต่ำ (กลุ่มอ่อน) เช่น นำไปใช้กับนักเรียนอ่อนเพื่อวินิจฉัยข้อบกพร่องในการเรียนวิชานั้น ส่วนข้อที่มีค่า b อยู่ระหว่าง -1.0 ถึง +1.0 เหมาะสำหรับไปใช้กับ นักเรียนที่มีความสามารถปานกลาง เช่น ใช้ในการเรียนการสอนในชั้นปกติ หรือเรียกว่า วัตถุประสงค์นั้นเอง

c : คือค่าจุดตัดบนแกนตั้ง (ค่าโอกาสที่จะทำข้อสอบข้อนั้นถูก) เรียกค่าโอกาสการเดา (Item Guessing) เพราะเป็น โอกาสที่จะทำข้อสอบนั้นถูกโดยไม่มีความสามารถ มีค่าที่เป็นไปได้ อยู่ระหว่าง 0 ถึง 1 ถ้าข้อสอบนั้นไม่สามารถที่จะตอบถูกด้วยการเดา $c_i = 0$ และปกติควรจะมีค่าน้อยกว่า 0.3 ซึ่งความหมายของค่าการเดามีดังนี้ ถ้า $c < 0.2$ แสดงว่าข้อสอบนั้นมีตัวเลือกดี มาก ถ้า c มีค่าในช่วงตั้งแต่ 0.2 ถึง 0.3 แสดงว่าข้อสอบนั้นมีตัวเลือกดี ถ้า $c > 0.3$ แสดงว่า ข้อสอบนั้น มีตัวเลือกที่เดาถูกได้ง่าย

การใช้โมเดลการตอบสนองข้อสอบที่สอดคล้องกับข้อมูล (Model-Data Fit) จะทำให้ค่าพารามิเตอร์ที่ได้มีลักษณะไม่แปรเปลี่ยน (Invariance) ทั้งความไม่แปรเปลี่ยนของค่าประมาณความสามารถและความไม่แปรเปลี่ยนของค่าประมาณพารามิเตอร์ของข้อสอบ จึงทำให้การเลือกข้อสอบแต่ละข้อ เพื่อสร้างชุดข้อสอบเป็นแบบสอบกระทำได้อย่างมีประสิทธิภาพ นอกจากนี้ ปริมาณสารสนเทศที่ได้จากข้อสอบ เมื่อนำมารวมกัน เป็นสารสนเทศของแบบสอบจะสะท้อนความถูกต้องแม่นยำของการประมาณค่าความสามารถ (θ) จึงทำให้ได้หลักประกันการพัฒนาแบบสอบเพื่อให้ได้แบบสอบตามเป้าหมายที่สนองต่อการนำไปใช้อย่างมีประสิทธิภาพและประสิทธิผล (ศิริชัย กาญจนวาสี, 2545 : 73-74)

งานวิจัยที่เกี่ยวข้องกับการสร้างแบบทดสอบวินิจฉัยการเรียนรู้

นิพนธ์ ชัยพิชิต (2536 : 84-89) ได้สร้างแบบทดสอบวินิจฉัยการเรียนรู้วิชาคณิตศาสตร์ เรื่อง ความยาว พื้นที่ และปริมาตร สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 กลุ่มตัวอย่างที่ใช้เพื่อการวิเคราะห์หาคุณภาพของเครื่องมือ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 1 ในโรงเรียนขยายโอกาสทางการศึกษา สังกัดสำนักงานการประถมศึกษาจังหวัดแพร่ ปีการศึกษา 2535 โดยใช้กลุ่มตัวอย่างจำนวน 286 คน จากการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการวิจัย คือ แบบทดสอบวินิจฉัยวิชาคณิตศาสตร์ เรื่อง ความยาว พื้นที่ และปริมาตร จำนวน 8 ฉบับ ได้แก่

1. การหาความยาวจากการวัดและการประมาณ
2. การเขียนแผนผังโดยใช้มาตราส่วน
3. การเปลี่ยนความยาวจากแผนผังเป็นความยาวจริง
4. การหาพื้นที่รูปสี่เหลี่ยมมุมฉาก รูปสี่เหลี่ยมคางหมู รูปสามเหลี่ยม และรูปวงกลม
5. โจทย์ปัญหาเกี่ยวกับพื้นที่รูปสี่เหลี่ยมมุมฉาก รูปสี่เหลี่ยมคางหมู รูปสามเหลี่ยม และรูปวงกลม
6. ความเข้าใจพื้นฐานเกี่ยวกับปริมาตรรูปทรงสี่เหลี่ยมมุมฉาก
7. การหาปริมาตรรูปทรงสี่เหลี่ยมมุมฉาก
8. โจทย์ปัญหาเกี่ยวกับปริมาตรรูปทรงสี่เหลี่ยมมุมฉาก

วิเคราะห์ข้อมูลโดยหาค่าความยากด้วยวิธีของเบรนนัน ความตรงเชิงเนื้อหาด้วยวิธีของโรวินลีสและแฮมเบิลตันค่าความเที่ยงแบบอิงเกณฑ์โดยใช้สูตร คูเดอร์-ริชาร์ดสัน 20 และปรับแก้โดยใช้วิธีของลิฟวิงสตันและวิเคราะห์หาสาเหตุความบกพร่องโดยผู้วิจัยและผู้เชี่ยวชาญเนื้อหา ร่วมกันพิจารณา

ผลปรากฏว่า แบบทดสอบทั้ง 8 ฉบับ จำนวน 87 ข้อ มีความยากตั้งแต่ 0.50 ถึง 0.96 ค่าอำนาจจำแนกตั้งแต่ 0.08 ถึง 0.72 ค่าความเที่ยงของแบบทดสอบแต่ละฉบับมีค่าเท่ากับ 0.71, 0.75, 0.71, 0.81, 0.73, 0.64, 0.87 และ 0.78 ตามลำดับ ค่าความเที่ยงของแบบทดสอบรวมทั้ง 8 ฉบับมีค่าเท่ากับ 0.93 ผลการพิจารณาความตรงเชิงเนื้อหาของแบบทดสอบวินิจฉัยทั้ง 8 ฉบับ มีความตรง 0.56 ถึง 1.00 แสดงว่า เป็นแบบทดสอบที่สามารถใช้วัดเรื่องความยาว พื้นที่ และปริมาตรได้

สุภาวดี กิตติวิศิษฐ์ (2537 : บทคัดย่อ) ได้สร้างแบบทดสอบวินิจฉัยวิชาคณิตศาสตร์ เรื่อง สมการ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในจังหวัดนครราชสีมา กลุ่มตัวอย่างที่ใช้ จำนวน 420 คน แบบทดสอบวินิจฉัยที่สร้างขึ้นมี 2 ฉบับคือ ฉบับที่ 1 การแก้สมการ มีความยากง่าย ตั้งแต่ 0.51-0.91 มีค่าอำนาจจำแนกตั้งแต่ 0.71-0.81 และมีความเชื่อมั่นเท่ากับ 0.9244 และฉบับที่ 2 โจทย์สมการมีความยากง่ายตั้งแต่ 0.51-0.83 มีค่าอำนาจจำแนกตั้งแต่ 0.81-0.53 และมีความเชื่อมั่นเท่ากับ 0.9545 แบบทดสอบวินิจฉัยทั้ง 2 ฉบับมีความตรงเชิงเนื้อหาจริง

อาภรณ์ เวียงวิเศษ (2540 : บทคัดย่อ) ได้สร้างแบบทดสอบวินิจฉัยการเรียนเรื่อง สมการ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 แบบทดสอบที่สร้างขึ้นมี 4 ฉบับ คือ แบบทดสอบความรู้ พื้นฐานเรื่องสมการ กลุ่มตัวอย่างที่ใช้ในกระบวนการสร้างแบบทดสอบครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2539 ของโรงเรียนสังกัดกรมสามัญศึกษาในจังหวัดบุรีรัมย์ จำนวน 1,200 คน ซึ่งได้จากการสุ่มแบบแบ่งชั้น ผลการวิจัยพบว่าแบบทดสอบวินิจฉัย

แบบทดสอบวินิจฉัยฉบับที่ 1 มีความยากตั้งแต่ 0.28-0.72 ค่าอำนาจจำแนก ตั้งแต่ 0.39-0.83 ค่าความเชื่อมั่น เท่ากับ 0.839

แบบทดสอบวินิจฉัยฉบับที่ 2 มีความยากตั้งแต่ 0.40-0.70 ค่าอำนาจจำแนก ตั้งแต่ 0.46-0.77 ค่าความเชื่อมั่น 0.892

แบบทดสอบวินิจฉัยฉบับที่ 3 มีความยากตั้งแต่ 0.32-0.67 ค่าอำนาจจำแนก ตั้งแต่ 0.53-0.85 ค่าความเชื่อมั่น 0.838

แบบทดสอบวินิจฉัยฉบับที่ 4 มีความยากตั้งแต่ 0.48-0.67 ค่าอำนาจจำแนก ตั้งแต่ 0.37-0.60 ค่าความเชื่อมั่น 0.677

ค่าความเที่ยงตรงเชิงเนื้อหา หาโดยให้ผู้เชี่ยวชาญทางด้านเนื้อหาเป็นผู้พิจารณา ตรวจสอบ ปรากฏว่าแบบทดสอบวินิจฉัยทั้งสี่ฉบับสามารถวัดเรื่องสมการได้จริง

สุพรรณิ ภิรมย์ภักดี (2541 : 71-75) ได้สร้างแบบทดสอบวินิจฉัยข้อบกพร่องทางการเรียน วิชาคณิตศาสตร์ เรื่องฟังก์ชัน สำหรับนักเรียนมัธยมศึกษาปีที่ 4 แบบทดสอบที่สร้างขึ้นมี 5 ฉบับคือ

- แบบทดสอบฉบับที่ 1 ความหมายของฟังก์ชัน
- แบบทดสอบฉบับที่ 2 ตัวอย่างฟังก์ชันที่ควรรู้จัก
- แบบทดสอบฉบับที่ 3 ฟังก์ชันคอมโพสิท
- แบบทดสอบฉบับที่ 4 ฟังก์ชันอินเวอร์ส
- แบบทดสอบฉบับที่ 5 พิษคณิตของฟังก์ชัน

ผลการศึกษาพบว่า แบบทดสอบวินิจฉัยทั้ง 5 ฉบับ มีค่าความยากตั้งแต่ 0.5114-0.9029 ค่าอำนาจจำแนกตั้งแต่ 0.2192-0.8403 ซึ่งคำนวณโดยใช้สูตร ดัชนีอำนาจจำแนก บี (Discrimination Index B) ของเบรนนัน คะแนนจุดตัดของแบบทดสอบทั้ง 5 ฉบับ ใช้วิธีกำหนดเกณฑ์ร้อยละ 80 ของจำนวนข้อสอบที่นักเรียนทำได้ในแต่ละจุดประสงค์เป็น 20, 4, 6, 9 และ 4 ตามลำดับ ส่วนค่าความเชื่อมั่นของแบบทดสอบทั้ง 5 ฉบับ ซึ่งคำนวณโดยใช้สูตรไบโนเมียล (Binomial) ของโลเวทท์ (Lovett) มีค่า 0.9647, 0.9913, 0.9881, 0.9844 และ 0.9925 ตามลำดับ สำหรับค่าความเที่ยงตรงของแบบทดสอบทั้ง 5 ฉบับ ให้ผู้เชี่ยวชาญเป็นผู้พิจารณาได้ค่าดัชนีความสอดคล้องตั้งแต่ 0.60-1.00 แสดงว่าแบบทดสอบวินิจฉัยที่ผู้วิจัยสร้างขึ้นมีความเที่ยงตรงเชิงเนื้อหาที่ต้องการวัดจริง

ผลการวิเคราะห์จุดบกพร่องที่นักเรียนเลือกตอบผิดในแบบทดสอบวินิจฉัยทั้ง 5 ฉบับ ปรากฏว่า แบบทดสอบแต่ละฉบับ มีความบกพร่องดังนี้

แบบทดสอบฉบับที่ 1 เรื่องความหมายของฟังก์ชัน มีข้อบกพร่องคือ ยังสรุปความคิดรวบยอดไม่ได้ ไม่เข้าใจการตรวจสอบฟังก์ชันจากกราฟ ใช้สัญลักษณ์ของช่วงไม่ถูกต้อง

แบบทดสอบฉบับที่ 2 เรื่องตัวอย่างฟังก์ชันที่ควรรู้จัก มีข้อบกพร่องคือ สร้างความสัมพันธ์ สมการไม่ได้ เข้าใจคำถามไม่ตรงกับที่โจทย์ต้องการ

แบบทดสอบฉบับที่ 3 เรื่องฟังก์ชันคอมโพสิท มีข้อบกพร่องคือ ยังสรุปความคิดรวบยอดของฟังก์ชันคอมโพสิทไม่ได้ ยังไม่เข้าใจในการแทนค่าฟังก์ชัน

แบบทดสอบฉบับที่ 4 เรื่องฟังก์ชันอินเวอร์ส มีข้อบกพร่องคือ ยังไม่เข้าใจในเรื่องของช่วง ยังหาโอเปอเรชันกันเซตไม่ได้ ยังใช้แกนสมมาตรของกราฟของฟังก์ชันกับกราฟฟังก์ชันอินเวอร์สไม่ถูกต้อง และยังสรุปนิยามของฟังก์ชันอินเวอร์สไม่ได้

แบบทดสอบฉบับที่ 5 เรื่องพีชคณิตของฟังก์ชัน มีข้อบกพร่องคือ ไม่เข้าใจในการบวก ลบ คูณ และ หารกันของฟังก์ชันโดยนำโดเมนไปปฏิบัติการด้วย ไม่เข้าใจเรื่องการบวก ลบ คูณ และหารกันของพหุนาม

สุชาติ สิริมินนนท์ (2542 :131-141) ได้สร้างแบบทดสอบวินิจฉัยจุดบกพร่องในการเรียนวิชาคณิตศาสตร์ เรื่องภาคตัดกรวยที่มีคุณภาพจำนวน 6 ฉบับ คือ การเลื่อนแกนทางขนาน วงกลม พาราโบลา วงรี ไฮเพอร์โบลา และชนิดของกราฟ เพื่อใช้ค้นหาจุดบกพร่องและสาเหตุของความบกพร่องในการเรียนวิชาคณิตศาสตร์ เรื่อง ภาคตัดกรวย สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4 กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 จำนวน 1,398 คน ของปีการศึกษา 2541 ของโรงเรียนสังกัดกรมสามัญศึกษา กลุ่มที่ 5 ในกรุงเทพมหานครซึ่งเลือกมาโดยการสุ่มแบบ

แบ่งชั้น และมีวิธีการสร้างแบบทดสอบดังนี้ สร้างแบบทดสอบสำรวจชนิดให้ให้เติมคำตอบและแสดงวิธีทำ และหาความเที่ยงตรงเชิงเนื้อหา โดยผู้เชี่ยวชาญ 5 ท่านหลังจากนั้นนำไปสำรวจจุดบกพร่องในการตอบผิดของนักเรียนและรวบรวมคำตอบผิดและสร้างเป็นแบบทดสอบวินิจฉัยแบบเลือกตอบ 5 ตัวเลือก โดยนำคำตอบที่นักเรียนส่วนใหญ่ตอบผิด ในแบบทดสอบสำรวจมาสร้างเป็นตัวलग แล้วนำแบบทดสอบไปสอบ 3 ครั้ง ซึ่งการสอบครั้งที่ 1 เพื่อวิเคราะห์ข้อสอบรายข้อ คัดเลือกและปรับปรุงข้อสอบ ทดสอบครั้งที่ 2 เพื่อวิเคราะห์ข้อสอบรายข้อ และคัดเลือกข้อสอบ ทดสอบครั้งที่ 3 เพื่อหาคุณภาพของแบบทดสอบแต่ละฉบับ

ผลการศึกษาพบว่า แบบทดสอบวินิจฉัยทั้ง 6 ฉบับ ที่วัดเกี่ยวกับการเลื่อนแกนทางขนาน วงกลม พาราโบลา วงรี ไฮเพอร์โบลา และชนิดของกราฟ มีค่าของแบบทดสอบตั้งแต่ 0.5407-0.8793 และค่าอำนาจจำแนกของข้อสอบตั้งแต่ 0.2101-0.7715 ส่วนค่าความเชื่อมั่นในแบบทดสอบในแต่ละฉบับ ซึ่งคำนวณโดยใช้สูตรไบโนเมียลมีค่า 0.7435 , 0.8053 , 0.7749 , 0.7557 , 0.7558 และ 0.8678 ตามลำดับ สำหรับคะแนนจุดตัดของแบบทดสอบมีค่า 9 , 10 , 11 , 10 , 13 และ 13 ตามลำดับ

สุภาพร ละอองวิจิตร (2543 : 64-66) ได้ทำการวินิจฉัยเพื่อพัฒนาแบบทดสอบวินิจฉัยวิชาคณิตศาสตร์ เรื่อง ฟังก์ชัน โดยใช้ทฤษฎีการตอบสนองข้อสอบ แบบทดสอบแบ่งเป็น 6 โดเมนซึ่งแยกย่อยได้เป็น 19 จุดประสงค์ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ปีการศึกษา 2542 จำนวน 1,110 คน แบ่งออกเป็น 2 กลุ่ม โดยกลุ่มที่ 1 จำนวน 515 คน ใช้สำหรับพัฒนาวิเคราะห์ข้อสอบโดยทฤษฎีการตอบสนองข้อสอบ กลุ่มที่ใช้สำหรับทดสอบแบบทดสอบวินิจฉัยเพื่อหาเกณฑ์ปกติ จำนวน 595 คน วิเคราะห์ข้อสอบโดยใช้โปรแกรม Bilog Version 3.04 ตรวจสอบคุณภาพของแบบ ทดสอบทางด้านความเที่ยงตามแนวอิงเกณฑ์ด้วยสูตรแบบ Binomial รวมทั้งความแม่นยำในการประมาณค่าความสามารถที่แท้จริง ซึ่งหา ได้ Test Information Function ส่วนด้านความตรงได้พัฒนาแบบทดสอบให้มีความตรงเชิงเนื้อหาตามวิธีของโรวิเนลลี และแฮมเบิลตัน หาคะแนนจุดตัดโดยวิธีโลจิสติก และหาเกณฑ์ปกติในรูปเปอร์เซ็นต์ไทล์ (Percentile Norms) และคะแนนที่ปกติ (Normalized T score)

การวิจัยครั้งนี้ได้แบบทดสอบวินิจฉัยที่มีคุณภาพดังนี้

1. มีความตรงเชิงเนื้อหาโดยมีดัชนีความสอดคล้องในช่วง ระหว่าง 0.67-1.00
2. มีคุณสมบัติการวัดเพียงคุณลักษณะเดียว
3. มีค่าอำนาจจำแนก ค่าความยาก และค่าการเดา โดยมีค่าอำนาจจำแนกตั้งแต่ 0.449 ถึง 7.959 ค่าความยากตั้งแต่ 0.795 ถึง 1.898 และค่าการเดาตั้งแต่ 0.085 ถึง 0.466
4. มีค่าความเที่ยงตรงแนวอิงเกณฑ์ของโดเมนที่ 1 ถึงโดเมนที่ 6 และทั้งฉบับเป็น

0.949,0.140,0.798,0.705,0.776,0.797 และ 0.956 ตามลำดับ สำหรับค่าความแม่นยำในการประมาณค่าความสามารถที่แท้จริง พบว่าแบบทดสอบมีประสิทธิภาพสูงสุดเมื่อใช้วัดผู้สอบที่มีความสามารถปานกลางและค่อนข้างสูง คือมีค่า θ อยู่ในช่วง -0.3 ถึง 2.2

5. คะแนนจุดตัดโดยวิธีโลจิสติกของโดเมนที่1 ถึงโดเมนที่6 และทั้งฉบับเป็น 21,14,3,3,5 และ 38 ตามลำดับ

แบบทดสอบที่พัฒนาขึ้นมานี้มีคุณภาพดี เหมาะสำหรับการตรวจสอบความรู้ของนักเรียนการสอนสิ้นสุดลงว่า นักเรียนมีข้อบกพร่องในเรื่องใด เพื่อจะได้เป็นแนวทางในการปรับปรุงแก้ไข หรือ ทำการซ่อมเสริม และเป็นแนวทางในการจัดการเรียนการสอนเพื่อจัดข้อบกพร่องให้สอดคล้องกับข้อบกพร่องต่อไป

จรรยาพร หนูทอง (2545 : บทคัดย่อ) การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อสร้างแบบทดสอบวินิจฉัยวิชาคณิตศาสตร์ เรื่องพหุนาม สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 (ม. 3) ให้มีคุณภาพสูง แบบทดสอบที่สร้างขึ้นมีจำนวน 6 ฉบับคือ แบบทดสอบความรู้เบื้องต้น เกี่ยวกับเอกนาม แบบทดสอบการบวกและการลบเอกนาม แบบทดสอบการคูณพหุนามและการแบบทดสอบการหารพหุนามเพื่อใช้ค้นหาจุดบกพร่องในการเรียน กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 (ม.3) ปีการศึกษา 2544 ของโรงเรียนสังกัดกรมสามัญศึกษา ในจังหวัดปัตตานี จำนวน 1,649 คน ซึ่งเลือกโดยวิธีการสุ่มแบบแบ่งชั้น สำหรับการสร้างแบบทดสอบ มีวิธีดำเนินการสร้างดังนี้ คือ สร้างแบบทดสอบเพื่อสำรวจแบบแสดงวิธีทำและเติมคำตอบ โดยเขียนข้อสอบให้สอดคล้องกับจุดประสงค์เชิงพฤติกรรม นำไปทดสอบเพื่อสำรวจจุดบกพร่อง และรวบรวมคำตอบผิด แล้วดัดแปลงเป็นแบบทดสอบวินิจฉัยแบบเลือกตอบชนิดสี่ตัวเลือก โดยนำคำตอบที่นักเรียนส่วนมากตอบผิดในแบบทดสอบเพื่อสำรวจมาเป็นตัวลวง นำไปทดสอบสามครั้ง คือ ทดสอบครั้งที่หนึ่งและครั้งที่สอง เพื่อคัดเลือกและปรับปรุงข้อสอบ และทดสอบครั้งที่สาม เพื่อหาคุณภาพของแบบทดสอบ

ผลการวิจัยพบว่า

แบบทดสอบวินิจฉัยฉบับที่ 1 มีค่าความยากตั้งแต่ 0.37-0.78 ค่าอำนาจจำแนกตั้งแต่ 0.03-0.62 ค่าความเชื่อมั่น 0.8796

แบบทดสอบวินิจฉัยฉบับที่ 2 มีค่าความยากตั้งแต่ 0.29-0.65 ค่าอำนาจจำแนกตั้งแต่ 0.10-0.66 ค่าความเชื่อมั่น 0.8754

แบบทดสอบวินิจฉัยฉบับที่ 3 มีค่าความยากตั้งแต่ 0.34-0.68 ค่าอำนาจจำแนกตั้งแต่ 0.07-0.73 ค่าความเชื่อมั่น 0.9026

แบบทดสอบวินิจฉัยฉบับที่ 4 มีค่าความยากตั้งแต่ 0.30-0.71 ค่าอำนาจจำแนก ตั้งแต่ 0.14-0.53 ค่าความเชื่อมั่น 0.8471

แบบทดสอบวินิจฉัยฉบับที่ 5 มีค่าความยากตั้งแต่ 0.27-0.82 ค่าอำนาจจำแนก ตั้งแต่ 0.14-0.56 ค่าความเชื่อมั่น 0.8420

แบบทดสอบวินิจฉัยฉบับที่ 6 มีค่าความยากตั้งแต่ 0.30-0.61 ค่าอำนาจจำแนก ตั้งแต่ 0.07-0.64 ค่าความเชื่อมั่น 0.8526

ค่าความตรงเชิงเนื้อหาโดยให้ผู้เชี่ยวชาญทางด้านเนื้อหาเป็นผู้พิจารณาตรวจสอบ ปรากฏว่าแบบทดสอบวินิจฉัยทั้งหกฉบับสามารถวัดเรื่องพหุนามได้จริง

บอยเดน (Boyden, 1970 : 1504-A) ได้สร้างแบบทดสอบวินิจฉัยเรื่องจุดบกพร่องในการแก้โจทย์ปัญหาเลขคณิต (Verbal Arithmetic Problem Solving) สำหรับนักเรียนเกรด 5 โดยใช้กลุ่มตัวอย่าง จำนวน 993 คน เริ่มจากการสร้างแบบทดสอบสำรวจแบบตอบอิสระเพื่อหาจุดบกพร่อง ผลของการสำรวจพบจุดบกพร่องในลักษณะต่าง ๆ 12 ลักษณะ จากนั้นสร้างเป็นแบบทดสอบวินิจฉัยแบบเลือกตอบโดยได้คำตอบผิดจากแบบทดสอบสำรวจมาใช้เป็นตัวตรวจวิเคราะห์เพื่อหาคุณภาพของแบบทดสอบหาค่าความเที่ยงโดยใช้สูตรคูเดอร์-ริชาร์ดสัน 20 และค่าอำนาจจำแนกโดยใช้สูตรสหสัมพันธ์แบบพอยท์ไบซีเรียล ปรากฏว่าแบบทดสอบสำรวจมีค่าความเที่ยงตั้งแต่ 0.727-0.850 ค่าอำนาจจำแนกตั้งแต่ 0.334-0.629 สรุปได้ว่าแบบทดสอบที่สร้างขึ้นสามารถค้นหาจุดบกพร่องในการเรียนเป็นรายบุคคลและทั้งชั้นได้ และเป็นประโยชน์ในการสอนซ่อมเสริมเป็นอย่างมาก

เอลลิส (Ellis, 1972 : 2234-A) ได้ศึกษาจุดบกพร่องในการเรียน เรื่องเลขจำนวนเต็มของนักเรียนชั้นประถมศึกษา มีจุดมุ่งหมายเพื่อค้นหาข้อผิดพลาดที่เกิดขึ้นบ่อย ๆ เกี่ยวกับการคำนวณเลขและปรับปรุงแบบทดสอบวินิจฉัย โดยศึกษากับนักเรียนเกรด 6 โรงเรียนลิฟวิงสโตนปารีส (Living Stone Parish) มลรัฐหลุยเซียนา (Louisiana) จำนวน 690 คนในปีการศึกษา 1970 นักเรียนแต่ละคนได้รับการทดสอบ ด้วยแบบทดสอบค้นหาจุดบกพร่องเรื่องเลขจำนวนเต็ม แล้วจัดนักเรียนเป็น 3 กลุ่ม คือกลุ่มที่ตอบถูกทั้งหมด กลุ่มที่ทำถูกวิธีแต่คำตอบผิดและกลุ่มที่ทำผิดทั้งวิธีทำและคำตอบ หลังจากนั้นกลุ่มซึ่งทำถูกวิธีแต่คำตอบผิดได้รับการทดสอบด้วยแบบทดสอบวินิจฉัย ผลจากการศึกษาพบว่ามีจุดบกพร่องในการบวก ร้อยละ 17 มีจุดบกพร่องในการคูณด้วยเลขหลักเดียว ร้อยละ 14 และมีจุดบกพร่องในการคูณด้วยเลขสองหลักร้อยละ 60

โบว์แมน (Bowman, 1976 : 7260-A) ได้สร้างแบบทดสอบวินิจฉัยเกี่ยวกับคณิตศาสตร์เบื้องต้น (A Basic Mathematics Diagnostic Instrument) เพื่อค้นหาจุดเด่นและจุดบกพร่องของนักเรียนในเรื่อง การบวก ลบ คูณ หาร การแก้โจทย์ปัญหา และพีชคณิตศาสตร์เบื้องต้น

แบบทดสอบชุดนี้สร้างขึ้นเพื่อใช้ในวิทยาลัยที่มีโครงการช่วยเหลือนักเรียนที่มีพื้นฐานความรู้ทางคณิตศาสตร์ต่ำและทำการทดสอบเป็นกลุ่ม แต่นำผลจากคำตอบของนักเรียนแต่ละคนมาพิจารณาว่านักเรียนมีจุดเด่นและจุดบกพร่องในเนื้อหาใด มักผิดพลาดในลักษณะใด ข้อมูลที่ได้จะถูกบันทึกเป็นเส้นภาพ (Profile) เพื่อความสะดวกในการตีความหมายผลงานของนักเรียน กลุ่มตัวอย่างที่ใช้ศึกษาครั้งนี้มีจำนวน 435 คน ผลปรากฏว่าแบบทดสอบชุดนี้มีประโยชน์ในการสอนซ่อมเสริมนักเรียนเป็นรายบุคคลได้อย่างเหมาะสม

จิน (Jean, 1978 : 4636-A) ได้ศึกษาจุดบกพร่องในการเรียนวิชาคณิตศาสตร์เรื่องการบวก (Addition Process) ของนักเรียนเกรด 3 และเกรด A โดยใช้แบบทดสอบวินิจฉัยขั้นหาจุดบกพร่องและทำการสอนซ่อมเสริมจุดบกพร่องนั้น ผลจากการศึกษาพบว่า นักเรียนที่บกพร่องในการเรียนคณิตศาสตร์เกี่ยวกับตัวเลขเป็นเพราะขาดทักษะพื้นฐานเกี่ยวกับระบบจำนวนส่วนนักเรียนที่ได้รับการสอนซ่อมเสริมได้คะแนนเพิ่มขึ้นจากการทำแบบทดสอบหลังการสอนซ่อมเสริมแล้วมากกว่ากลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

วอล์คเกอร์ (Walker, 1983 : 981-A) ได้ศึกษาผลกระทบจากวิธีการบริหารการสอนที่มีต่อผลการทำแบบทดสอบวินิจฉัยการคณิตศาสตร์ที่เป็นแบบมาตรฐานเพื่อพิจารณาว่าวิธีการบริหารการสอนจะมีต่อการทำแบบทดสอบวินิจฉัยการเรียนที่เป็นแบบมาตรฐานหรือไม่โดยใช้กลุ่มตัวอย่างที่สุ่มมาจากนักเรียนในระดับชั้น 5-8 จากโรงเรียนมัธยมศึกษาตอนต้นที่ตั้งอยู่ในซานเมื่องและชนบทจำนวน 822 คน แบบเป็นกลุ่มใหญ่และกลุ่มเล็ก แล้วใช้วิธีการบริหารการสอน 2 แบบ คือ จัดสอบโดยตรงที่มีรูปแบบการจัดเป็นพิธีการ และจัดสอบทางอ้อม ที่ไม่เป็นพิธีการ แล้วทดสอบด้วยแบบทดสอบที่ประยุกต์มาจากแบบทดสอบวินิจฉัยการเรียนคณิตศาสตร์ของสแตนฟอร์ด (SDMT) โดยจัดสอบโดยตรงและจัดสอบทางอ้อมกับนักเรียนที่เป็นกลุ่มใหญ่และกลุ่มเล็ก แล้วนำค่าเฉลี่ยของคะแนนในแต่ละกลุ่มมาวิเคราะห์ความแปรปรวน 3 ทาง (Three-way-ANOVA) โดยใช้ตัวแปรเน้นการบริหารสอบ 2 แบบ กลุ่มที่ได้รับการสอบกลุ่มใหญ่และกลุ่มเล็ก และเพศชาย - หญิงของนักเรียนที่เป็นกลุ่มตัวอย่าง ผลวิเคราะห์ปรากฏว่า ค่าเฉลี่ยของคะแนนผลสัมฤทธิ์ของกลุ่มที่จัดสอบทางอ้อมสูงกว่ากลุ่มที่จัดสอบโดยตรงที่ระดับความมีนัยสำคัญ 0.05 ในระดับชั้น 5 และระดับชั้น 7 และค่าเฉลี่ยของคะแนนนักเรียนชายสูงกว่านักเรียนหญิงในระดับ 8 ที่ระดับนัยสำคัญ 0.01 แต่ไม่พบปฏิสัมพันธ์สองทางหรือสามทางของตัวแปรทั้งหมดว่ามีนัยสำคัญทางสถิติที่ระดับใดเลย

ไนท์ (Knight, 1984 : 499-A) ได้ทำการศึกษาผลการสอบเชิงวินิจฉัยที่มีผลต่อผลสัมฤทธิ์ ในวิชาคณิตศาสตร์ของนักเรียนระดับประถมศึกษา จุดมุ่งหมายหลักของการศึกษาค้นคว้าครั้งนี้คือ การตรวจสอบผลของการทดสอบผลสัมฤทธิ์เชิงวินิจฉัยที่มีต่อผลสัมฤทธิ์ที่เกิดตามมาภายหลังและยังมุ่งที่จะสร้างแบบทดสอบวินิจฉัยโดยใช้แหล่งข้อมูลนอกชั้นเรียน เพื่อพิจารณาว่าแบบทดสอบที่สร้างขึ้น จากข้อเสนอแนะต่างๆ ที่ได้จากการวินิจฉัยที่ผ่านมาจะเป็นที่ยอมรับของครูหรือไม่ จุดประสงค์ ประการสุดท้ายเพื่อการศึกษาว่าครูจะยอมให้นำข้อมูลต่างๆ ของนักเรียนออกมาใช้เพื่อวิเคราะห์ และสรุปผลมากน้อยเพียงใด กลุ่มตัวอย่างเป็นโรงเรียน 30 แห่ง และครู 120 คนซึ่งสุ่มจากโรงเรียน ที่มีชั้นประถม 4-6 ในเขต Eadtern Ontario และแบ่งออกเป็น 3 กลุ่ม กลุ่มแรกคือ กลุ่มทดสอบ และวินิจฉัย ได้รับแบบทดสอบและวินิจฉัยซึ่งสร้างขึ้นตามหลักสูตรของท้องถิ่นและดำเนินการให้ ครูทดสอบโดยใช้เวลาที่สะดวก เมื่อครูให้คะแนนขั้นต้นแล้ว ให้ส่งแบบทดสอบไปยังส่วนรวม เพื่อสรุปผลด้วยคอมพิวเตอร์ ซึ่งจะส่งสถิติผลสัมฤทธิ์รวมทั้งสรุปผลด้านทักษะไปให้โรงเรียน ภายในเวลาประมาณ 2 วัน กลุ่มที่ 2 คือ กลุ่มทดสอบแต่เพียงอย่างเดียวได้รับแบบทดสอบ เช่นเดียวกัน แต่ไม่มีการส่งผลสรุปกลับไปให้โรงเรียน กลุ่มที่ 3 คือ กลุ่มควบคุมไม่ได้ทำอะไรเลย คงปล่อยให้ดำเนินการตามปกติ หลังจากนั้นได้ส่งแบบสอบถามเพื่อสำรวจความสนใจและความพึงพอใจของครูและมีการทดสอบอีกครั้ง โดยใช้ชั้นนักเรียนทุกคนในทุกกลุ่ม ผลการศึกษาแสดงว่า (1) ครูผู้สอนเห็นว่าแบบทดสอบวินิจฉัยที่สร้างขึ้นในการศึกษาค้นคว้าครั้งนี้มีประโยชน์ (2) การศึกษาค้นคว้านี้ไม่มีปัญหาเกี่ยวกับตัวครูเนื่องจากครูผู้สอนมีความเต็มใจใช้แบบทดสอบ และ (3) กลุ่มแรกมีผลสัมฤทธิ์ที่ได้จากการทดสอบครั้งหลังสูงกว่าอีกสองกลุ่มอย่างน้อยมีนัยสำคัญใน 2 ระดับชั้นจาก 3 ระดับชั้นที่ใช้ในการศึกษาค้นคว้า

คอนสแตนติน (Constantine, 1985 : 3621-A) ได้ศึกษาความเชื่อมั่นของการแบ่งกลุ่มตามความสามารถ โดยใช้จุดประสงค์ด้วยแบบทดสอบวินิจฉัยในวิชาคณิตศาสตร์ การหาความเชื่อมั่นของการตัดสินแบ่งกลุ่มความสามารถ ทำโดยการทดสอบด้วยแบบทดสอบคู่ขนานที่สร้างขึ้น จากจุดประสงค์ของการสำรวจเชิงวินิจฉัยในวิชาคณิตศาสตร์ จำนวนผู้เข้าทดสอบและจำนวนข้อคำถามที่อาศัยข้อมูลที่ได้จากสำนักพิมพ์ จุดประสงค์แต่ละข้อในจำนวน 37 ข้อนั้นจะมีคำถาม 10 แบบ โดยมีจำนวนข้อคำถาม (เหตุบังเอิญ) 1 ข้อ ถึง 8 ข้อ (ข้อเสนอแนะต่ำสุดของ Popham) และ 20 ข้อ (ข้อเสนอแนะสูงสุดของ Popham) การวิเคราะห์ข้อมูลใช้ดัชนี P สัมประสิทธิ์แคมป์ พบว่าค่าเฉลี่ยที่ได้จากแบบทดสอบแบบต่างๆ ที่มีข้อคำถามแต่ละจุดประสงค์มีค่าแตกต่างกัน และการเปรียบเทียบรายคู่ด้วยการทดสอบ Multiple rangtest ของ Cucan ผลการศึกษาไม่ตรงกับสมมติฐานที่คาดหวังไว้ คือ พบว่าจุดประสงค์ DMP ที่มีข้อคำถามข้อเดียวสามารถใช้ตัดสินแบ่งกลุ่มตามความสามารถได้อย่างมีความเชื่อมั่นในทุกๆ ระดับและจุดประสงค์ที่มีข้อคำถามเพียง 10 ข้อแต่

อย่างไร จุดประสงค์ที่มีข้อคำถามหลายข้อและที่มี 4 ข้อ มีความเชื่อมั่นในการแบ่งกลุ่มได้สูงกว่าที่มีเพียงข้อเดียวและจุดประสงค์ที่มีข้อคำถาม 10 ข้อ ไม่ได้ช่วยให้ความเชื่อมั่นได้สูงกว่าจุดประสงค์ที่มีหลายข้อและที่มีเพียง 4 ข้อ แต่อย่างไร

วิลสัน (Wilson, 1988 : 55-A) ได้สร้างแบบทดสอบวินิจฉัยตนเองและชุดฝึกทักษะพื้นฐานทางคณิตศาสตร์ในการศึกษาวิชาชีพ วัตถุประสงค์ในการวิจัยเพื่อกำหนดคุณลักษณะของนักศึกษาในสาขาอาชีพในโรงเรียนอาชีวศึกษาของรัฐ ในรัฐเคนตักกี (Kentucky) และพัฒนาเครื่องมือคณิตศาสตร์ และชุดการฝึกในการช่วยให้นักศึกษาเกิดความเชี่ยวชาญทักษะพื้นฐานที่จำเป็นในวิชาคณิตศาสตร์ โดยได้สร้างคลังข้อสอบขึ้น เพื่อวัดทักษะและนำไปใช้กับนักเรียนโรงเรียนรัฐบาลประมาณ 500 คน จากระดับ 4 ถึงระดับ 8 ข้อคำถามแต่ละตอนในคลังข้อสอบต้องมีค่าเท่ากับเชิงสถิติใช้โปรแกรมคอมพิวเตอร์ภาษาเบสิก ในการสุ่มเลือกและจำแนกเครื่องมือเชิงวินิจฉัย และโปรแกรมสำหรับการฝึกเกี่ยวข้องกับทักษะพื้นฐานทางคณิตศาสตร์ มีความสัมพันธ์กับเครื่องมือวินิจฉัยที่สุ่มได้จำแนกเป็น 3 แบบ และนำไปทดสอบกับนักศึกษาจำนวน 100 คน จำนวน 20 โรงเรียน เพื่อหาเกณฑ์ปกติและความเที่ยงตรงของข้อสอบ เครื่องมือเชิงวินิจฉัยแสดงให้เห็นค่าความสัมพันธ์ในเชิงบวกสูงมาก ค่าความเชื่อมั่นโดยเฉลี่ย .95 จากการใช้แบบทดสอบทั้ง 3 แบบ หาค่าเกณฑ์ปกติไม่มีความแตกต่างระหว่างค่ามัธยฐานเลขคณิตของคะแนนดิบของแบบทดสอบทั้ง 3 แบบ พบว่าคะแนนเปอร์เซ็นต์ที่ 67 เป็นเกณฑ์ปกติ

สรุปได้ว่า ปัญหาการสร้างเครื่องมือเชิงวินิจฉัยวิชาคณิตศาสตร์ และชุดการฝึกทักษะโดยคอมพิวเตอร์ มีความสามารถเพิ่มความเชี่ยวชาญเชิงทักษะพื้นฐานทางคณิตศาสตร์และช่วยให้นักศึกษาอาชีวศึกษาประสบผลสำเร็จในการสอบเข้าเรียนต่อในโรงเรียนอาชีวศึกษา

จากผลงานที่เกี่ยวข้องกับแบบทดสอบวินิจฉัยที่กล่าวมาแล้ว สามารถค้นจุดบกพร่องในการเรียนของนักเรียนทั้งที่เป็นรายบุคคลและทั้งชั้นได้ เป็นประโยชน์ในการจัดการสอนซ่อมเสริมและในการปรับปรุงด้านการเรียนการสอนในชั้นได้เป็นอย่างมาก

งานวิจัยที่เกี่ยวข้องกับการใช้ทฤษฎีการตอบสนองข้อสอบในการพัฒนาแบบทดสอบ

งานวิจัยของ ประดิษฐ์ เรื่องตระกูล (2529 : บทคัดย่อ) ได้ศึกษาการใช้โมเดลโลจิสติกเพื่อพัฒนาแบบสอบวินิจฉัยวิชาคณิตศาสตร์สำหรับนักเรียนเรียนชั้นมัธยมศึกษาปีที่ 1 โดยใช้กลุ่มตัวอย่างจำนวน 1,913 คน ผลวิจัยปรากฏว่า ค่าอำนาจจำแนก (a) มีค่าอยู่ในช่วงตั้งแต่ 0.19036 ถึง 2.00000 ค่าความยาก (b) มีค่าอยู่ในช่วงตั้งแต่ -3.75288 ถึง 1.90366 ค่าการเดา (c) มีค่าอยู่ในช่วงตั้งแต่ 0.00 ถึง 0.5 พบว่าแบบทดสอบนี้จะมีประสิทธิภาพสูงสุดเมื่อใช้ทดสอบกับผู้เรียนที่ระดับความสามารถปานกลาง

สกว สันติเทวกุล (2540 : บทคัดย่อ) ได้ศึกษาการพัฒนาแบบทดสอบวินิจฉัยวิชาฟิสิกส์ ว.021 โดยใช้ทฤษฎีการตอบสนองข้อสอบ โดยใช้กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 จำนวน 994 คน สามารถพัฒนาได้แบบทดสอบที่มีค่าพารามิเตอร์ดังนี้ ค่าอำนาจจำแนก (a) มีค่าอยู่ในช่วงตั้งแต่ 0.30 ถึง 0.27 ค่าความยาก (b) มีค่าอยู่ในช่วงตั้งแต่ -2.75 ถึง 3.00 ค่าการเดา (c) มีค่าอยู่ในช่วงตั้งแต่ 0.03 ถึง 0.30 และค่าฟังก์ชันสารสนเทศสูงสุดที่ 19.529 ที่ระดับความสามารถ 0.1 นั่นคือ แบบทดสอบนี้จะเหมาะสมหรือมีประสิทธิภาพสูงสุดเมื่อนำไปทดสอบกับผู้เรียนในระดับความสามารถปานกลาง งานวิจัยที่พัฒนาแบบทดสอบวินิจฉัยทางการเรียนอีกงานคือ

งานวิจัยของ สุภาพร ละอองวิจิตร (2543 : บทคัดย่อ) ที่ศึกษาการพัฒนาแบบทดสอบวินิจฉัยทางการเรียนวิชาคณิตศาสตร์ เรื่อง ฟังก์ชัน โดยใช้ทฤษฎีการตอบสนองข้อสอบ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4 ซึ่งใช้กลุ่มตัวอย่างในการวิเคราะห์หาคุณภาพแบบทดสอบ 515 คน ผลวิจัยปรากฏว่า ค่าอำนาจจำแนก (a) มีค่าอยู่ในช่วงตั้งแต่ 0.449 ถึง 7.959 ค่าความยาก (b) มีค่าอยู่ในช่วงตั้งแต่ -0.795 ถึง 1.898 ค่าการเดา (c) มีค่าอยู่ในช่วงตั้งแต่ 0.085 ถึง 0.466 และพบว่าแบบทดสอบนี้จะมีประสิทธิภาพสูงสุดเมื่อใช้ทดสอบกับผู้เรียนที่ระดับความสามารถปานกลางและค่อนข้างสูง

ส่วนงานวิจัยของ บรรหาร จิตหวัง (2543 : บทคัดย่อ) ได้ศึกษาการพัฒนาแบบสอบความคิดรวบยอดทางคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 5 โดยวิเคราะห์แบบสอบด้วยทฤษฎีการตอบสนองข้อสอบ โดยใช้โมเดลโลจิสติก 3 พารามิเตอร์ มีผลวิจัยปรากฏว่า ค่าอำนาจจำแนก (a) มีค่าอยู่ในช่วงตั้งแต่ 0.558 ถึง 3.064 ค่าความยาก (b) มีค่าอยู่ในช่วงตั้งแต่ -0.097 ถึง 2.822 ค่าการเดา (c) มีค่าอยู่ในช่วงตั้งแต่ 0.105 ถึง 0.346 และพบว่าแบบทดสอบนี้จะมีประสิทธิภาพสูงสุดเมื่อใช้ทดสอบกับผู้เรียนที่ระดับความสามารถค่อนข้างสูง

งานวิจัยของ ปิยวดี คงช่วย (2544 : บทคัดย่อ) ได้ศึกษาการพัฒนาแบบทดสอบความคิด วิเคราะห์ญาณ โดยใช้ทฤษฎีการตอบสนองข้อสอบ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 พบว่า ค่าอำนาจจำแนก (a) มีค่าอยู่ในช่วงตั้งแต่ 0.69 ถึง 1.34 ค่าความยาก (b) มีค่าอยู่ในช่วงตั้งแต่ 0.52 ถึง 1.80 ค่าการเดา (c) มีค่าอยู่ในช่วงตั้งแต่ 0.19 ถึง 0.29 และค่าฟังก์ชันสารสนเทศของ แบบทดสอบสูงมากที่ระดับความสามารถ 1.1-1.6 นั่นคือ แบบทดสอบนี้มีความเหมาะสมเมื่อใช้ ทดสอบกับนักเรียนที่มีความสามารถระดับสูง

งานวิจัยของ สิริภพ กาฬสุวรรณ (2546 : บทคัดย่อ) ได้ศึกษาการพัฒนาแบบทดสอบวัด ทักษะกระบวนการทางวิทยาศาสตร์ โดยใช้ทฤษฎีการตอบสนองข้อสอบ สำหรับนักเรียนชั้น ประถมศึกษาปีที่ 6 โดยใช้กลุ่มตัวอย่างจำนวน 1,366 คน ผลวิจัยปรากฏว่า ค่าอำนาจจำแนก (a) มี ค่าอยู่ในช่วงตั้งแต่ 0.34601 ถึง 1.989 ค่าความยาก (b) มีค่าอยู่ในช่วงตั้งแต่ -0.931 ถึง 1.921 ค่า การเดา (c) มีค่าอยู่ในช่วงตั้งแต่ -0.101 ถึง 0.260 นั่นคือ แบบทดสอบนี้มีความเหมาะสมเมื่อใช้ ทดสอบกับนักเรียนที่มีความสามารถระดับปานกลางถึงค่อนข้างสูง

จากงานวิจัยดังกล่าวพอจะสรุปได้ว่า ทฤษฎีการตอบสนองข้อสอบนอกจากจะใช้ในการ วิเคราะห์คุณภาพของแบบทดสอบได้แล้ว ยังสามารถวิเคราะห์ได้ว่าแบบทดสอบนั้นมีความ เหมาะสมเมื่อใช้ทำการทดสอบกับกลุ่มผู้สอบที่ระดับความสามารถใด อันจะช่วยทำให้การใช้ แบบทดสอบนั้น ๆ มีประสิทธิภาพ