

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การพัฒนาแบบทดสอบวินิจฉัยรายวิชาภาษาอังกฤษระดับชั้นประถมศึกษาปีที่ 4 โดยใช้
ทฤษฎีการตอบสนองข้อสอบ ผู้วิจัยได้ศึกษาและรวบรวมเอกสารในประเด็นที่เกี่ยวข้องกับการ
ศึกษาและได้นำเสนอในประเด็นต่อไปนี้

1. หลักสูตรภาษาอังกฤษระดับประถมศึกษา พ.ศ. 2539
2. แบบทดสอบวินิจฉัย
3. เกณฑ์ปกติ (Norms)
4. คะแนนที่ปกติ (Normalized T – score)
5. ทฤษฎีการตอบสนองข้อสอบ
6. งานวิจัยที่เกี่ยวข้อง

หลักสูตรภาษาอังกฤษระดับประถมศึกษา พ.ศ. 2539

กรมวิชาการ กระทรวงศึกษาธิการ (2540 : 8-17) ได้กำหนดรายละเอียดต่าง ๆ ของ
หลักสูตรภาษาอังกฤษระดับประถมศึกษา พ.ศ. 2539 ไว้ดังนี้ ในโลกยุคข้อมูลข่าวสาร
ภาษาต่างประเทศมีความสำคัญและจำเป็นต่อการสื่อสารเป็นอย่างยิ่ง ภาษาอังกฤษเป็นภาษาที่ทั่ว
โลกใช้สื่อสารเป็นภาษาสากล กระทรวงศึกษาธิการตระหนักถึงความสำคัญและเห็นความจำเป็น
ของการปรับปรุง หลักสูตร และกระบวนการเรียนการสอนภาษาอังกฤษ ทั้งนี้เพื่อพัฒนาความ
สามารถทางภาษาอังกฤษของผู้เรียนในการติดต่อส่งสารและรับสารและให้มีพื้นฐานภาษาอังกฤษ
เพียงพอที่จะนำไปใช้ในการศึกษาต่อ หรือในการทำงานอาชีพของตนจึงได้กำหนดแนวทางปรับ
ปรุงหลักสูตรภาษาอังกฤษโดยมุ่งเน้นให้ผู้เรียนเกิดความสามารถในการสื่อสารด้วยภาษาอังกฤษ
คือสามารถใช้ภาษาในการเข้าสู่สังคมและวัฒนธรรม (Socio – Cultural Functions) และสามารถ
สื่อความโดยใช้ภาษาอย่างถูกต้องตามหลักภาษา และเหมาะสมกับสถานการณ์การใช้ (Cognitive
Linguistic Functions)

หลักสูตรภาษาอังกฤษระดับประถมศึกษา จัดเป็นภาษาอังกฤษระดับต้น แบ่งเป็น 3
ระดับ คือ ภาษาอังกฤษระดับเตรียมความพร้อม สอนในชั้นประถมศึกษาปีที่ 1-2 ภาษาอังกฤษ
ระดับอ่านออกเขียนได้ สอนในชั้นประถมศึกษาปีที่ 3-4 และภาษาอังกฤษระดับมาตรฐานพื้นฐาน

ตอนต้น สอนในชั้นประถมศึกษาปีที่ 5-6 แนวการจัดการเรียนการสอนเน้นให้ผู้เรียนเป็นศูนย์กลาง จัดสถานการณ์และบรรยากาศการเรียนการสอนให้ผู้เรียน ได้ใช้ภาษาอังกฤษให้มากที่สุด โดยใช้สื่อ อุปกรณ์การเรียนการสอนที่หลากหลาย

ในระดับเตรียมความพร้อม เน้นความสามารถด้านการฟัง พูด โดยการจัดกิจกรรมที่ทำให้ผู้เรียนเกิดความสุข สนุกสนาน คำนึงเกี่ยวกับภาษาอังกฤษ ส่วนในระดับอ่านออกเขียนได้ จัดกิจกรรมที่ผสมผสานการอ่าน เขียน และสะกดคำ เพื่อเป็นพื้นฐานการเรียนภาษาอังกฤษ สำหรับระดับมาตรฐานตอนต้น ซึ่งมุ่งพัฒนาความสามารถด้านการสื่อสารด้วยการฟัง พูด อ่าน และเขียน ที่ถูกต้องตามหลักภาษาและเหมาะสมกับสถานการณ์การใช้

การวัดและประเมินผลของผู้เรียนในระดับเตรียมความพร้อมและระดับอ่านออกเขียนได้มุ่งประเมินภาคปฏิบัติ และเจตคติต่อการเรียนภาษาอังกฤษ ในระดับมาตรฐานพื้นฐานตอนต้น เน้นการวัดความสามารถโดยรวมการใช้ภาษา ใช้การวัดและประเมินผลทั้งด้านความรู้ การปฏิบัติ และเจตคติ ประเมินความก้าวหน้าในการใช้ภาษาด้วยการสังเกต การตรวจสอบ กระบวนการทำงาน และผลของการทำงานด้วยเครื่องมือวัดและประเมินผลที่หลากหลาย

1. วัตถุประสงค์ของหลักสูตรภาษาอังกฤษระดับประถมศึกษา

วัตถุประสงค์ของการเรียนการสอนภาษาอังกฤษในระดับประถมศึกษาซึ่งเป็นการเรียนภาษาอังกฤษในระดับต้น มีดังนี้

1.1 เพื่อให้มีความสามารถในการใช้ภาษาเพื่อการสื่อสารได้อย่างถูกต้องตามหลักภาษาและเหมาะสมกับวัฒนธรรม

1.2 เพื่อให้มีความรู้ความสามารถในการใช้ภาษาอังกฤษในระดับต้นอย่างเพียงพอเพื่อการศึกษาต่อหรือนำไปใช้ในการประกอบอาชีพตามความจำเป็น

1.3 เพื่อให้มีความสามารถในการใช้ภาษาเพื่อการสื่อสารในสถานการณ์ง่าย ๆ ที่เป็นสถานการณ์จำลองหรือสถานการณ์จริง

1.4 เพื่อให้มีความรู้และความเข้าใจวัฒนธรรมภาษาอังกฤษ

1.5 เพื่อให้มีนิสัยใฝ่หาความรู้เพิ่มเติมด้วยการอ่านหนังสือภาษาอังกฤษและใช้สื่อภาษาอังกฤษอื่น ๆ

2. โครงสร้างหลักสูตรภาษาอังกฤษระดับประถมศึกษา และอัตราเวลาเรียน

โครงสร้างหลักสูตรภาษาอังกฤษระดับประถมศึกษา ได้มีการกำหนดให้จัดการเรียนการสอนวิชาภาษาอังกฤษในกลุ่มประสบการณ์พิเศษ การจัดการเรียนการสอนวิชาภาษาอังกฤษในระดับประถมศึกษาเป็นภาษาอังกฤษระดับต้น (Beginner Level) จำแนกได้เป็น 3 ระดับ ดังนี้

2.1 ภาษาอังกฤษระดับเตรียมความพร้อม (Preparatory Level) กำหนดให้เรียนในชั้นประถมศึกษาปีที่ 1 ภาคเรียนที่ 2 และในชั้นประถมศึกษาปีที่ 2 ตลอดปีการศึกษา

2.2 ภาษาอังกฤษระดับอ่านออกเขียนได้ (Literacy Level) กำหนดให้เรียนในชั้นประถมศึกษาปีที่ 3 และชั้นประถมศึกษาปีที่ 4 ตลอดปีการศึกษา

2.3 ภาษาอังกฤษระดับมาตรฐานพื้นฐานตอนต้น (Beginner Fundamental Level) ประกอบด้วยภาษาอังกฤษหลัก 1-4 กำหนดให้เรียนในชั้นประถมศึกษาปีที่ 5 และชั้นประถมศึกษาปีที่ 6 ตลอดปีการศึกษา

การกำหนดอัตราเวลาเรียนภาษาอังกฤษ ในระดับประถมศึกษาปีที่ 1-4 ควรกำหนดให้เรียนไม่น้อยกว่า 6 คาบ (120 นาที) ต่อสัปดาห์ และควรกำหนดให้เรียนที่ละน้อยแต่กระจายไปตลอดสัปดาห์ สำหรับนักเรียนประถมศึกษาปีที่ 5-6 ควรให้เรียนไม่น้อยกว่า 15 คาบ (300 นาที) ต่อสัปดาห์ และควรกำหนดให้เรียนทุกวัน

3. แนวการจัดการเรียนการสอนภาษาอังกฤษระดับประถมศึกษา

แนวการจัดการเรียนการสอนภาษาอังกฤษตามหลักสูตรนี้ มุ่งให้ผู้เรียนมีความสามารถในการสื่อสารควบคู่กัน 2 ด้าน ได้แก่ ความสามารถในการใช้ภาษาเพื่อเข้าสู่สังคมและวัฒนธรรม และความสามารถในการใช้ภาษาเพื่อสื่อความได้อย่างถูกต้องตามหลักภาษาและเหมาะสมกับสถานการณ์ โดยมีแนวหลักดังนี้

3.1 จัดการเรียนการสอนโดยเน้นผู้เรียนเป็นศูนย์กลาง

3.2 จัดการเรียนการสอนภาษาด้วยกิจกรรมที่มีความหมายและหลากหลาย ฝึกการสื่อสารในสถานการณ์ต่าง ๆ เพื่อให้ผู้เรียนสามารถใช้ภาษาในสถานการณ์จริง

3.3 จัดให้ผู้เรียนได้ฝึกฝนการสื่อสารด้วยการฟังและการพูดในระดับเตรียมความพร้อมเพิ่มเติมการฝึกฝนการสื่อสารด้วยการอ่าน การเขียน และการสะกดคำในระดับอ่านออกเขียนได้ และฝึกฝนการส่งสารและรับสารด้วยการฟัง พูด อ่าน และเขียน ในการเรียนภาษาอังกฤษระดับมาตรฐานพื้นฐานตอนต้น

4. กระบวนการวัดและประเมินผลภาษาอังกฤษระดับประถมศึกษา

กระบวนการวัดและประเมินผลเน้นการวัดความสามารถในการสื่อสารด้วยภาษาอังกฤษทั้งในด้านการใช้ภาษาเพื่อเข้าสู่สังคมและวัฒนธรรม และการใช้ภาษาเพื่อสื่อความได้อย่างถูกต้องตามหลักภาษา และเหมาะสมกับสถานการณ์ วัดทั้งความสามารถในการสื่อสารและความรู้ทางภาษา ผลของการวัดและการประเมินผล จะนำมาใช้ในการพัฒนาผู้เรียนอย่างต่อเนื่อง ระดับเตรียมความพร้อม วัดและประเมินผลโดยเน้นการประเมินพฤติกรรมการเรียนของผู้เรียนจากการเข้าร่วมกิจกรรม และความสามารถในการสื่อสารด้วยการฟังและการพูด ระดับอ่านออกเขียนได้ วัดและประเมินจากความสามารถในการใช้ภาษาในด้านความเข้าใจ การฟัง พูด อ่าน เขียน และการสะกดคำ ส่วนระดับมาตรฐานพื้นฐานตอนต้น วัดความรู้ทางภาษาและความสามารถในการใช้ภาษา โดยประเมินให้ครบทั้ง 3 ลักษณะ คือ ประเมินผลก่อนเรียน ประเมินผลระหว่างเรียน และประเมินผลภายหลังเรียน ประเมินทั้งภาคปฏิบัติ ภาคความรู้ และเจตคติ โดยเครื่องมือวัดและการประเมินผลที่หลากหลาย มีการวัดระดับมาตรฐานขั้นพื้นฐานทางภาษาของผู้เรียน โดยใช้การทดสอบมาตรฐานความรู้ความสามารถ (Proficiency Test) สำหรับผู้จะสำเร็จการศึกษา

5. สื่อการเรียนการสอนภาษาอังกฤษระดับประถมศึกษา

สื่อการเรียนการสอนที่ดีมีหลายประเภท แต่ละประเภทมีคุณสมบัติเฉพาะที่สามารถนำมาใช้ประกอบการเรียนการสอนได้อย่างน่าสนใจ สื่อเหล่านี้จะช่วยให้การเรียนการสอนภาษาอังกฤษพัฒนาได้อย่างรวดเร็ว นอกเหนือจากหนังสือเรียน แบบฝึกหัด คู่มือครู และแถบบันทึกเสียง ประกอบบทเรียนแล้ว สื่อภาษาอังกฤษที่พบเห็นในชีวิตประจำวัน (Authentic Materials) เช่น หนังสือพิมพ์ แบบฟอร์มต่าง ๆ จดหมาย แผ่นภาพ ป้ายโฆษณา และอื่น ๆ นอกจากนี้สื่อของจริงที่พบเห็นรอบ ๆ ตัว ทั้งในและนอกห้องเรียนก็สามารถนำมาใช้สอนภาษาอังกฤษในระดับประถมศึกษาได้เป็นอย่างดี ทั้งนี้หากโรงเรียนมีความพร้อมก็อาจพิจารณาใช้สื่อที่เป็นเทคโนโลยีทันสมัย เช่น คอมพิวเตอร์ช่วยสอนภาษา วิดิทัศน์ ศูนย์การเรียนรู้ด้วยตนเอง (Self Access Learning Center) สื่อทางไกล และอื่น ๆ เป็นเครื่องช่วยให้ผู้เรียนสามารถพัฒนาความสามารถในการใช้ภาษาได้อย่างรวดเร็วยิ่งขึ้น อุปกรณ์และสื่อดังกล่าวนี้ ผู้สอนสามารถเลือกใช้หรือสร้างขึ้นให้เหมาะสมกับ สถานการณ์การเรียนการสอน สภาพห้องเรียน ระดับและพื้นฐานทางภาษาของผู้เรียน โดยจะต้องแสวงหาความรู้เกี่ยวกับสื่อและพิจารณาเลือกใช้สื่อจนสามารถนำสื่อแต่ละชนิดมาประกอบการเรียนการสอนได้อย่างมีประสิทธิภาพและเกิดประโยชน์สูงสุด

การเรียนการสอนภาษาอังกฤษในระดับประถมศึกษา จะต้องเน้นการใช้ภาษาอังกฤษ เป็นสื่อในการปฏิสัมพันธ์ระหว่างผู้สอนกับผู้เรียน และระหว่างผู้เรียนกับผู้เรียน โดยยึดเป็นแนวปฏิบัติที่สำคัญ อันจะก่อให้เกิดความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารได้อย่างแท้จริง

6. เนื้อหาของหลักสูตรภาษาอังกฤษระดับประถมศึกษาปีที่ 3-4 ภาษาอังกฤษอ่านออกเขียนได้ (Literacy English)

6.1 แนวคิดในการจัดการเรียนการสอนภาษาอังกฤษอ่านออกเขียนได้

การเรียนการสอนภาษาอังกฤษระดับอ่านออกเขียนได้ เป็นการเรียนการสอนที่ผู้เรียน เริ่มเรียนรู้ระบบการเขียน และการสะกดคำภาษาอังกฤษที่ฟัง – พูด ได้แล้วในระดับเตรียมความพร้อม ในระดับนี้ผู้เรียนจะเกิดความสามารถในการสื่อสารทั้งด้วยการฟัง พูด อ่าน เขียน โดยการใช้ปฏิบัติ กิจกรรมต่าง ๆ ที่ทำให้เกิดความสุข สนุกสนาน เริ่มพัฒนาความสามารถในการอ่านและการเขียนรวมทั้งฝึกหัดคัดลายมือ และเขียนตอบคำถาม โดยใช้ภาษาในระดับพื้นฐาน ในการเข้าสู่สังคมและวัฒนธรรมและในการสื่อความได้ถูกต้องตามหลักภาษา เรียนรู้คำนามเกี่ยวกับคน สัตว์ สิ่งของที่พบเห็นในชีวิตประจำวันทั่ว ๆ ไป และคำกริยาแสดงลักษณะและอาการของสิ่งเหล่านี้โดยขยาย วงศัพท์เพิ่มขึ้นตามความเหมาะสม ฝึกใช้หนังสือ พจนานุกรม ตลอดจนสื่อที่พบเห็นรอบ ๆ ตัวเป็นเครื่องช่วยในการเรียนรู้ และส่งเสริมให้มีเจตคติที่ดีต่อการเรียนภาษาอังกฤษ เห็นประโยชน์ และคุณค่าของภาษาอังกฤษ

6.2 จุดประสงค์การเรียนรู้ภาษาอังกฤษอ่านออกเขียนได้

6.2.1 เพื่อให้สามารถใช้ภาษาเพื่อการสื่อสารด้วยการฟัง พูด อ่าน และเขียนอย่างถูกต้อง ชัดเจน และเหมาะสมกับระดับที่เรียน

6.2.2 เพื่อให้สามารถฟัง พูด และสนทนาเป็นภาษาอังกฤษในสถานการณ์ง่าย ๆ

6.2.3 เพื่อให้สามารถอ่าน เขียน และสะกดคำที่เป็นตัวอักษรตัวพิมพ์ ใช้เครื่องหมายวรรคตอนถูกต้อง

6.2.4 เพื่อให้สามารถใช้พจนานุกรมเป็นเครื่องมือในการเพิ่มพูนความรู้ด้านคำศัพท์

6.2.5 เพื่อให้มีเจตคติที่ดีต่อการเรียนภาษาอังกฤษ

6.3 ขอบข่ายเนื้อหาภาษาอังกฤษ

หลักสูตรวิชาภาษาอังกฤษ พุทธศักราช 2539 นี้ เป็นหลักสูตรที่เน้นการพัฒนาความสามารถในการสื่อสารด้วยภาษาอังกฤษของผู้เรียนเป็นหลัก โดยให้มีมาตรฐานพื้นฐานเดียวกันและต่อเนื่องกันตลอดแนวตั้งแต่ระดับประถมศึกษาถึงมัธยมศึกษา ขอบข่ายเนื้อหาที่กำหนดไว้ในหลัก

สูตรเป็นเพียงมาตรฐานขั้นพื้นฐาน โดยกำหนดจุดประสงค์การเรียนรู้เพื่อพัฒนาความสามารถในการใช้ภาษาเพื่อเข้าสู่สังคมและวัฒนธรรม และพัฒนาความสามารถในการใช้ภาษาเพื่อสื่อความได้อย่างถูกต้องตามหลักภาษา และเหมาะสมกับสถานการณ์การใช้ มีการกำหนดโครงสร้างทางไวยากรณ์ และคลังคำสำหรับผู้เรียนในแต่ละระดับ การกำหนดมาตรฐานขั้นพื้นฐานนี้ เพื่อให้สามารถจัดการเรียนการสอนให้ผู้เรียนพัฒนาความสามารถในการใช้ภาษาได้มาตรฐานในระดับใกล้เคียงกัน ในการปฏิบัติผู้ใช้หลักสูตรสามารถเพิ่มเติมคัดแปลงและปรับปรุงเนื้อหาต่าง ๆ ของหลักสูตรให้เหมาะสมกับผู้เรียนและท้องถิ่น

6.4 โครงสร้างทางไวยากรณ์

เป็นประโยคสั้น ๆ ง่าย ๆ เกี่ยวกับเรื่องการทักทาย การกล่าวลา การแนะนำตนเอง การแนะนำผู้อื่น การขอบคุณ พுகขอโทษ พுகแทรกอย่างสุภาพ และพுகขออนุญาต

6.5 คลังคำ

เป็นคำที่ปรากฏในโครงสร้างไวยากรณ์ในระดับนี้เป็นคำศัพท์เกี่ยวกับสิ่งที่อยู่ใกล้ตัว และสภาพแวดล้อมของผู้เรียน เกี่ยวกับตนเอง ครอบครัว โรงเรียน ซึ่งเป็นคำนาม สรรพนาม คำกริยา คำคุณศัพท์ คำวิเศษณ์ คำบุพบท คำอุทาน ในวงคำศัพท์ประมาณ 240-300 คำ นอกเหนือจากคำศัพท์ที่ผู้เรียนเคยเรียนมาแล้วในระดับประถมศึกษาปีที่ 1-2 รวมคำศัพท์สะสมในระดับนี้ประมาณ 360-450 คำ ฯลฯ

เนื้อหาที่นำมาสร้างแบบทดสอบวินิจฉัย ส่วนของโครงสร้างไวยากรณ์ในเรื่อง ประโยคบอกเล่า ประโยคคำถาม ประโยคคำตอบ คลังคำในเรื่อง คำศัพท์ที่เกี่ยวกับสิ่งใกล้ตัว วัฒนธรรมและสิ่งแวดล้อมของผู้เรียนเกี่ยวกับตนเอง ครอบครัว โรงเรียน สถานที่ วัน เดือน ปี ในวงคำศัพท์ที่กำหนด

แบบทดสอบวินิจฉัย

1. ความหมายของแบบทดสอบวินิจฉัย

แบบทดสอบวินิจฉัย เป็นลักษณะของแบบทดสอบแบบหนึ่งที่ใช้สำหรับค้นหา จุดบกพร่อง หรือปัญหาต่าง ๆ ในการเรียนของนักเรียน นักการศึกษาหลายท่านได้ให้ความหมายของแบบทดสอบวินิจฉัยไว้ดังนี้

เทอร์นไคค์ และฮาเกน (Thorndike and Hagen, 1977 : 646) ได้ให้ความหมายของแบบทดสอบวินิจฉัยว่า คือ แบบทดสอบที่รวบรวมปัญหาและสาเหตุที่ทำให้เกิดความบกพร่องในการเรียนวิชาต่าง ๆ ไว้ในแบบทดสอบเพื่อเป็นแนวทางในการจัดหาวิธีการสอนซ่อมเสริมที่ตรงจุด และเป็นการช่วยปรับปรุงความรอบรู้ (Mastery) ของนักเรียนให้เพิ่มขึ้นอีกด้วย

บราวน์ (Brown, 1976 : 255) ได้กล่าวถึงแบบทดสอบวินิจฉัยว่า เป็นแบบทดสอบที่ใช้ค้นหาข้อบกพร่องของนักเรียนเป็นรายบุคคล โดยมุ่งที่จะทำการสอนซ่อมเสริมและให้การแนะแนวซึ่งจะทำให้ชี้ให้เห็นถึงจุดอ่อนหรือจุดบกพร่องของนักเรียนเป็นรายบุคคลในแต่ละส่วนย่อย ๆ ของแบบทดสอบนั้น

พัมฟรีย์ (Pumfrey, 1976 : 14-15) ได้ให้ความหมายของแบบทดสอบวินิจฉัยทักษะในการอ่านว่าเป็นแบบทดสอบที่ชี้ให้เห็นถึงจุดบกพร่องในทักษะการอ่านของนักเรียนได้อย่างถูกต้อง ทำให้สามารถสรุปสาเหตุต่าง ๆ ที่มีผลต่อการอ่านได้ผลจากการใช้แบบทดสอบจะทำให้เข้าใจกระบวนการอ่านและสามารถปรับปรุง แก้ไข ป้องกัน และลดปัญหาทางการอ่านของนักเรียนเป็นรายบุคคลได้ถูกต้อง และตรงจุด

สิงห์ (Singha, 1974 : 200-201) กล่าวถึงแบบทดสอบวินิจฉัยว่ามีจุดมุ่งหมายเพื่อค้นหาข้อบกพร่องของนักเรียนเพื่อที่จะทำการซ่อมเสริม แบบทดสอบประเภทนี้การสุ่มเนื้อหาจึงจำเป็นต้องละเอียดมาก

อีเบล (Ebel, 1965 : 449) ได้ให้คำจำกัดความของแบบทดสอบวินิจฉัยเป็นแบบทดสอบที่สร้างขึ้นเพื่อใช้ค้นหาข้อบกพร่องหรือความไม่ประสบความสำเร็จในการเรียนของนักเรียน เช่น การอ่าน และเลขคณิต เป็นต้น แบบทดสอบวินิจฉัยมุ่งสนใจคะแนนจากคำตอบของนักเรียนใน แต่ละข้อหรือกลุ่มเล็ก ๆ ของข้อสอบที่มีความคล้ายคลึงกัน

อดัมส์ และทอร์เกอร์สัน (Adams and Torgerson, 1964 : 39-40) กล่าวว่าแบบทดสอบวินิจฉัยเป็นแบบทดสอบที่ใช้เพื่อชี้ให้เห็นจุดบกพร่องและสาเหตุของความบกพร่องแบบทดสอบวินิจฉัยไม่ให้ความสำคัญในด้านคะแนนรวม แต่จะเน้นรูปแบบของคำตอบเป็นสำคัญ ผลจากการทำข้อสอบจะทำให้เห็นว่า นักเรียนคนใดมีปัญหา หรือมีความบกพร่องในการเรียนและบกพร่องในเรื่องใด

สุเทพ สันติวรานนท์ (2533 : 69) ได้สรุปความหมายของแบบทดสอบวินิจฉัยว่าแบบทดสอบวินิจฉัย คือ แบบทดสอบที่สร้างขึ้นเพื่อชี้ให้เห็นจุดบกพร่องที่เกิดขึ้นในการเรียนเนื้อหานั้น ๆ

พวงรัตน์ ทวีรัตน์ (2530 : 27) ให้ความหมายของแบบทดสอบวินิจฉัยไว้ว่าแบบทดสอบวินิจฉัยเป็นแบบทดสอบที่ใช้ค้นหาข้อบกพร่อง และสาเหตุของการบกพร่องในการเรียนรู้เป็นเรื่อง ๆ ไป คือ หลังจากจบการเรียนการสอนในหน่วยหนึ่ง ๆ จะทำการสอบนักเรียนด้วยแบบทดสอบย่อย ถ้าพบว่า นักเรียนยังไม่สามารถสอบผ่านจุดประสงค์บางจุดประสงค์ก็ใช้แบบทดสอบวินิจฉัยมาทำการสอบอีกครั้งหนึ่ง เพื่อค้นหาว่านักเรียนยังไม่เข้าใจหรือบกพร่องในเรื่องใด และหาสาเหตุของการบกพร่อง

อนันต์ ศรีโสภา (2525 : 5) กล่าวว่าแบบทดสอบวินิจฉัยเป็นแบบทดสอบที่มีวัตถุประสงค์ เพื่อค้นหาสาเหตุของความบกพร่องและปัญหาต่าง ๆ ในการเรียนของนักเรียน แบบทดสอบวินิจฉัยจะมีจำนวนข้อสอบมากในแต่ละเนื้อหาวิชาที่เราต้องการทดสอบ

บุญชม ศรีสะอาด (2523 :10) ได้ให้ความหมายของแบบทดสอบวินิจฉัยว่า คือแบบทดสอบที่สร้างขึ้นเพื่อชี้ให้เห็นถึงจุดบกพร่อง จุดที่เป็นปัญหาหรืออุปสรรคในการเรียนเรื่องหนึ่ง ๆ ของนักเรียน

สมศักดิ์ สินธุระเวชญ์ (2522 : 1) ได้ให้ความหมายของแบบทดสอบวินิจฉัยไว้ว่าเป็นแบบทดสอบเพื่อวัดจุดอ่อนหรือจุดบกพร่องในแต่ละวิชา ผลของการสอบที่ได้จากแบบทดสอบวินิจฉัยจะนำไปสู่การแก้ไขข้อบกพร่องต่าง ๆ และค้นหาสาเหตุของความบกพร่องนั้นได้

ชวาล แพร์ตกุล (2518 : 317) ได้กล่าวถึงแบบทดสอบวินิจฉัยข้อบกพร่องในการเรียนเป็นแบบทดสอบเพื่อแก้ไข และส่งเสริมการเรียนของเด็ก ตลอดจนปรับปรุงการสอนของครูให้เกิดประสิทธิภาพยิ่งขึ้น

จากนิยามทั้งหมดที่กล่าวมาสรุปได้ว่า แบบทดสอบวินิจฉัยเป็นแบบทดสอบที่ใช้ค้นหาความบกพร่องทางการเรียนของนักเรียนเป็นรายบุคคล ผลการตอบแบบทดสอบสามารถบอกได้ว่านักเรียนบกพร่องในทักษะใดจุดใด และนำผลการวินิจฉัยมาทำการแก้ไข และส่งเสริมการเรียนของนักเรียนได้ถูกต้องและตรงจุด ตลอดจนปรับปรุงการสอนของครูให้มีประสิทธิภาพยิ่งขึ้น

2. ลักษณะของแบบทดสอบวินิจฉัย

แนวคิดเกี่ยวกับการสร้างแบบทดสอบวินิจฉัย เพื่อค้นหาความบกพร่องทางการเรียนของนักเรียน ได้มีนักการศึกษาทั้งในและต่างประเทศได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัย (Diagnostic Test) ไว้หลายท่านซึ่งพอจะรวบรวมได้ดังนี้

กรอนลันด์ (Gronlund, 1990 : 139) ได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยดังนี้

1. เป็นแบบทดสอบที่มีวัตถุประสงค์เพื่อหาความบกพร่องทางการเรียน
2. ความบกพร่องที่ได้จากการวัด เป็นความบกพร่องเฉพาะอย่างของบุคคล
3. แบบทดสอบมีค่าความยากอยู่ในระดับต่ำ
4. เป็นแบบทดสอบที่ใช้ประเมินระหว่างเรียน
5. ผลที่ได้จากการทดสอบ จะเป็นข้อมูลในการพิจารณาซ่อมเสริม

เมห์เรนส์ และเลห์มานน์ (Mehrens and Lehmann, 1975 : 462-464) กล่าวถึงลักษณะของแบบทดสอบวินิจฉัย ไว้ดังนี้

1. แบบทดสอบวินิจฉัยไม่คำนึงถึงคะแนนการสอบเพียงอย่างเดียว แต่จะพิจารณาถึงรายละเอียดต่าง ๆ จากผลงานของนักเรียนประกอบด้วยเพื่อเป็นแนวทางในการจัดสอนซ่อมเสริม
 2. แบบทดสอบวินิจฉัยจะสร้างเกณฑ์ปกติ (Norms) ในกรณีที่ต้องการจะแสดงว่าโดยทั่วไปนักเรียนมีความสามารถ อยู่ในระดับใดของกลุ่มและไม่มีเกณฑ์ปกติ (Norms) ในกรณีที่เราถือว่าเกณฑ์ปกติ (Norms) ได้มาจากข้อสอบมาตรฐานอื่น ๆ ซึ่งเป็นเกณฑ์ปกติระดับชาติอยู่แล้ว
 3. แบบทดสอบวินิจฉัยจะเป็นแบบทดสอบมาตรฐาน ในกรณีที่เครื่องมือที่ใช้นั้นถูกใช้ภายใต้เงื่อนไขเดียวกัน และการให้คะแนนมีความเป็นปรนัย
 4. แบบทดสอบวินิจฉัยอาจให้เกณฑ์ปกติเปอร์เซ็นต์ไทล์ (Percentile Norms) เกณฑ์ปกติเทียบชั้น (Grade Equivalent Norms) ก็ได้ตามความเหมาะสม
 5. แบบทดสอบวินิจฉัยจะใช้เฉพาะกับนักเรียนที่มีข้อบกพร่องทางการเรียนซึ่งจะต้องใช้เวลามากในการดำเนินการสอบ การตรวจและการตีความหมายของคะแนน
 6. แบบทดสอบวินิจฉัยสร้างยากกว่าแบบทดสอบวัดผลสัมฤทธิ์อื่น ๆ เพราะนอกจากจะต้องการคำตอบของนักเรียนแล้วยังต้องทำให้สามารถรู้ได้ว่านักเรียนมีข้อบกพร่องในด้านใด
- บลูม (Bloom, 1971 : 91-92) กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยว่า
1. เป็นแบบทดสอบเพื่อหาข้อบกพร่องของนักเรียนเกี่ยวกับทักษะพื้นฐาน เพื่อหา ระดับความรู้ของนักเรียน นอกจากนี้ยังเป็นประโยชน์ต่อครูในด้านปรับปรุงวิธีสอนและเพื่อหาสาเหตุของความบกพร่องทางการเรียนของนักเรียน เพื่อทำการสอนซ่อมเสริมในจุดบกพร่องที่เกิดขึ้นซ้ำ ๆ กันของนักเรียนจำนวนมาก
 2. ใช้เมื่อนักเรียนได้เรียนบทเรียนแต่ละบทเสร็จสิ้นลงแล้ว
 3. ใช้ประเมินผลได้ทั้งพฤติกรรมด้านพุทธิพิสัย (Cognitive Domain) ด้านจิตพิสัย (Affective Domain) และด้านทักษะพิสัย (Psychomotor Domain)
 4. แบบทดสอบวินิจฉัยประกอบด้วยกลุ่มข้อสอบจำนวนมาก ๆ และเป็นข้อสอบที่ง่าย โดยแต่ละข้อมีค่าความยากตั้งแต่ 0.65 ขึ้นไป
 5. การประเมินผลคะแนนจากแบบทดสอบอาจใช้ทั้งแบบอิงกลุ่มและแบบอิงเกณฑ์
 6. วิธีการรายงานคะแนนจะอยู่ในรูปเส้นภาพ (Profile) ของคะแนนของนักเรียนแต่ละคนในแต่ละทักษะย่อย
- เพน (Payne, 1968 : 167) กล่าวว่าแบบทดสอบวินิจฉัยควรมีลักษณะดังนี้
1. สอดคล้องกับจุดมุ่งหมายของหลักสูตรและจุดมุ่งหมายของการสอน
 2. ประกอบด้วยข้อสอบซึ่งเกิดจากการวิเคราะห์รายละเอียดขององค์ประกอบที่เกี่ยวข้อง

กับการเรียนรู้เรื่องนั้น ๆ

3. ชี้แนวทางในการแก้ไขข้อบกพร่องว่าควรแก้ ณ จุดใด
4. ครอบคลุมลำดับขั้นการเรียนรู้ในเรื่องนั้น ๆ

บุญชม ศรีสะอาด (2523 : 9-11) กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยไว้ดังนี้

1. มุ่งวัดเป็นเรื่อง ๆ หรือด้าน ๆ ไป ต้องอาศัยทักษะย่อยหลายทักษะก็อาจแบ่งเป็นแบบทดสอบย่อยวัดทักษะย่อยนั้น ๆ
2. มีคะแนนของแต่ละด้าน แต่ละตอน เพราะมุ่งค้นหาจุดบกพร่องในแต่ละด้าน ดังนั้นคะแนนรวมของแต่ละคนจะไม่เป็นประโยชน์ในกรณีนี้
3. มีข้อสอบหลายข้อที่วัดทักษะเดียวกัน ซึ่งจะช่วยให้เพิ่มโอกาสทำผิดพลาดได้มากขึ้น อันจะช่วยให้สามารถจำแนกนักเรียนที่มีความบกพร่องในการเรียนเรื่องนั้น ๆ ได้อย่างเพียงพอ นั่นคือ ชี้ให้เห็นถึงจุดบกพร่องที่แท้จริงได้อย่างชัดเจน
4. มักเป็นข้อสอบระดมพลัง (Power Test)
5. การสร้างแบบทดสอบ จะสร้างจากรากฐานของการวิเคราะห์ทักษะเฉพาะที่ส่งผลให้นักเรียนเรียนสำเร็จ และจากการศึกษาข้อผิดพลาด หรือความบกพร่องที่มักเกิดขึ้นกับนักเรียน
6. ความเป็นมาตรฐานของแบบทดสอบอยู่ในรูปที่ว่าเครื่องมือที่ใช้ดำเนินการสอบอยู่ภายใต้สภาพการณ์เดียวกันและการให้คะแนนมีความเป็นปรนัย

สมศักดิ์ สินธุระเวชญ์ (2522 : 1) กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยว่า เป็นแบบทดสอบเพื่อวัดจุดอ่อนหรือจุดบกพร่องในแต่ละวิชา ผลของการสอบที่ได้จากแบบทดสอบวินิจฉัยจะนำไปสู่การแก้ไขข้อบกพร่องต่าง ๆ และการค้นหาสาเหตุของการบกพร่องนั้น

ชวาล แพรัตกุล (2518 : 5-6) กล่าวว่าแบบทดสอบวินิจฉัย เป็นแบบทดสอบที่นิยมแยกข้อสอบของแต่ละวิชาออกเป็นฉบับย่อย ๆ โดยมีเป้าหมายที่จะวัดความรู้ความสามารถของนักเรียนเป็นด้าน ๆ ไป เพื่อให้ครอบคลุมทั้งเนื้อหาและพฤติกรรมที่กำหนดไว้ในหลักสูตร คุณสมบัติของแบบทดสอบย่อย ๆ เหล่านี้ คือ สามารถช่วยให้ ครูผู้สอนวินิจฉัยนักเรียน เป็นรายลักษณะได้ว่า ใครมีสมรรถภาพเด่น หรือด้อยในด้านใดบ้างจะได้ช่วยแก้ไขให้ตรงจุดยิ่งขึ้น

ซึ่งสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ (2539 : 2-3) ได้กล่าวถึงระดับของการวินิจฉัยว่ามี 3 ระดับ คือ

1. ระดับทั่วไป (General Level) เป็นการวินิจฉัยอย่างหยาบ เพราะเป็นขั้นตอนการสำรวจ เพื่อหาระดับความสามารถทั่ว ๆ ไปของนักเรียนแบบทดสอบที่จะใช้ในระดับนี้ ถ้าหากในต่างประเทศมักใช้แบบทดสอบมาตรฐาน แต่ในเมืองไทยใช้แบบทดสอบวัดผลสัมฤทธิ์ที่ใช้สอบ

วัดปลายภาคเรียน หรือปลายปีก็สามารถนำมาใช้ได้ ทั้งนี้ในการตรวจให้คะแนนมิใช่ดูที่คะแนนรวม แต่จะพิจารณาคะแนนเป็นรายสมรรถภาพในแต่ละสมรรถภาพ หรือคะแนนในแต่ละโดเมน (Domain) หรือแต่ละโดเมนย่อย (Sub Domain) ว่านักเรียนไม่บรรลุผลการเรียนในโดเมนย่อยใดบ้าง ก็คน

2. ระดับเฉพาะ (Specific Level) เป็นระดับที่ต้องการทราบว่านักเรียนมีข้อบกพร่องในเรื่องใด ณ จุดใด เป็นการวัดความสามารถเฉพาะเจาะจงลงไปในเรื่องใดเรื่องหนึ่ง เช่น การบวก การลบ และในการบวกยังแตกแยกย่อยลงไปอีกว่าบวกจำนวนเต็ม บวกเศษส่วน บวกทศนิยม เป็นต้น แบบทดสอบที่ใช้วัดในระดับนี้คือ แบบทดสอบวินิจฉัย ในต่างประเทศมีแบบทดสอบประเภทนี้ให้เลือกใช้มากมาย แต่ในเมืองไทยเกือบพูดได้ว่าไม่มีเลย ยกเว้นงานวิจัยของนักศึกษาศรีวิชัยของมหาวิทยาลัยต่าง ๆ พอมีบ้างแต่ก็มีน้อยมาก แบบทดสอบในลักษณะนี้ครูผู้สอนที่มีความตั้งใจ สนใจก็สามารถสร้าง และพัฒนาให้เป็นแบบทดสอบวินิจฉัยที่มีคุณภาพได้ ทั้งนี้ต้องมีความรู้ความเข้าใจในหลักการวิธีการสร้าง รู้ลักษณะของแบบทดสอบ และกระบวนการสร้างแบบทดสอบ ตลอดจนการแปลผลของคะแนน

3. ระดับละเอียด (Intensive Level) เป็นการวินิจฉัยอย่างละเอียดลึกซึ้งซึ่งเป็นการหาข้อมูลหลาย ๆ ด้าน หลาย ๆ แห่ง ทั้งนี้ มิใช่จะใช้แบบทดสอบเพียงอย่างเดียว การหาข้อมูลอาจใช้วิธีการสังเกต การสัมภาษณ์ ข้อมูลไม่ได้ใช้เฉพาะผลสัมฤทธิ์จากการตอบแบบทดสอบ แต่อาจวัดเจตคติ บุคลิกภาพ สุขภาพ หรืออื่น ๆ แหล่งข้อมูลมิใช่ได้แก่นักเรียนเพียงอย่างเดียว อาจต้องสอบถามจากผู้ปกครองครูที่สอนวิชาอื่น ๆ เพื่อนสนิท ผู้วินิจฉัยไม่ใช่เพียงแต่ครูผู้สอนวิชานั้น ๆ อาจประกอบด้วยผู้เชี่ยวชาญทางด้านวัดผล นักจิตวิทยาหรือครูแนะแนวด้วยก็ได้

จากลักษณะของแบบทดสอบวินิจฉัยที่กล่าวมา ต่างได้อธิบายลักษณะของแบบทดสอบวินิจฉัยไปในทางเดียวกัน ดังนี้

1. แบบทดสอบวินิจฉัยเป็นแบบทดสอบที่ใช้สำหรับค้นหาข้อบกพร่องและสาเหตุของความบกพร่องทางการเรียนในแต่ละเรื่องแต่ละเนื้อหา
2. เป็นแบบทดสอบที่มุ่งวัดในเนื้อหาที่เป็นเรื่อง ๆ หรือเป็นด้าน ๆ ไป ซึ่งสอดคล้องและครอบคลุมเนื้อหา จุดมุ่งหมายหลักสูตรและจุดประสงค์เชิงพฤติกรรมของเนื้อหาที่จะวินิจฉัย
3. เป็นข้อสอบที่จัดเรียงข้อสอบตามเนื้อหาโดยไม่คำนึงถึงความยากง่าย ต้องมีจำนวนข้อที่มากพอ
4. แบบทดสอบต้องมีความเที่ยงตรงเชิงเนื้อหา
5. เป็นแบบทดสอบที่ค่อนข้างง่าย
6. ความเป็นมาตรฐานของแบบทดสอบอยู่ที่การดำเนินการสอบและการให้คะแนน

3. ความตรงของแบบทดสอบวินิจฉัย

นักการศึกษาหลายท่านได้อธิบาย ความตรงของแบบทดสอบวินิจฉัยซึ่งมีความสำคัญสำหรับแบบทดสอบชนิดนี้ไว้ดังนี้

สิงห์ (Singha, 1974 : 200-201) ได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัย ที่เกี่ยวกับความตรงของแบบทดสอบไว้ว่า แบบทดสอบวินิจฉัยนี้จะต้องสุ่มเนื้อหาให้ละเอียดมากและความตรงเชิงเนื้อหาที่มีความจำเป็นมากกว่าแบบทดสอบชนิดอื่น ๆ

เพน (Payne, 1968 : 167) ได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยที่เกี่ยวกับความตรงของแบบทดสอบไว้ว่า แบบทดสอบวินิจฉัยจะต้องสร้างให้สอดคล้องกับจุดมุ่งหมายของหลักสูตร และจุดประสงค์ของการสอน ประกอบด้วยข้อสอบซึ่งเกิดจากการวิเคราะห์รายละเอียดขององค์ประกอบที่เกี่ยวกับการเรียนรู้เนื้อหานั้น ๆ

อาห์แมน และกล็อก (Ahmann and Glock, 1967 : 364-365) ได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัย ที่เกี่ยวกับความตรงของแบบทดสอบไว้ว่า แบบทดสอบวินิจฉัยเน้นความตรงเชิงเนื้อหา (Content Validity) เป็นสำคัญ

สุเทพ สันติวรานนท์ (2533 : 69-70) ได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัย ที่เกี่ยวกับความตรงของแบบทดสอบไว้ว่า แบบทดสอบวินิจฉัยนั้นต้องสร้างให้สอดคล้องกับจุดประสงค์ของหลักสูตรและจุดประสงค์การสอน มีการวิเคราะห์และสุ่มเนื้อหาอย่างรอบคอบเน้นความตรงเชิงเนื้อหา (Content Validity) เป็นหลัก

ชวาล แพรัตกุล (2518 : 5-6) ได้กล่าวถึงลักษณะของแบบทดสอบวินิจฉัย ที่เกี่ยวกับความตรงของแบบทดสอบไว้ว่า แบบทดสอบวินิจฉัยนิยมที่จะแยกข้อสอบแต่ละวิชาออกเป็นฉบับย่อย ๆ หลายฉบับ โดยมีเป้าหมายที่จะวัดความรู้ความสามารถของนักเรียนเป็นขั้น ๆ ไป สอบเพื่อให้ครอบคลุมเนื้อหา และพฤติกรรมที่สำคัญตามที่กำหนดไว้ในหลักสูตร

สรุปได้ว่าความตรงของแบบทดสอบวินิจฉัยจะต้องอาศัยการสร้างข้อสอบให้สอดคล้องกับจุดมุ่งหมายของหลักสูตรและจุดประสงค์การเรียนรู้ ครอบคลุมเนื้อหาวิชา เพราะความตรงเชิงเนื้อหา (Content Validity) มีความสำคัญมากสำหรับแบบทดสอบวินิจฉัย

4. เทคนิคและวิธีการสร้างแบบทดสอบวินิจฉัย

การสร้างแบบทดสอบแต่ละประเภทนั้น มีวิธีการและเทคนิคที่แตกต่างกันไปขึ้นอยู่กับว่าจุดมุ่งหมายในการนำแบบทดสอบนั้นไปใช้อะไร แบบทดสอบวินิจฉัยก็เช่นกันต้องมีเทคนิคและวิธีการสร้างที่แตกต่างไปจากแบบทดสอบประเภทอื่น ๆ จำเป็นจะต้องศึกษาหลักเกณฑ์

ในการสร้างแบบทดสอบให้เข้าใจแล้วจึงจะเริ่มดำเนินการสร้างแบบทดสอบ สำหรับขั้นตอนในการสร้างนั้น นักการศึกษาได้เสนอแนะไว้ดังนี้

ทอร์นไดค์ และฮาเกน (Thorndike and Hagen, 1977 : 269-271) ได้เสนอแนะขั้นตอนในการสร้างแบบทดสอบวินิจฉัยว่ามี 2 ขั้น คือ

1. วิเคราะห์ทักษะหรือเนื้อหาวิชาที่ต้องการทดสอบออกเป็นทักษะหรือองค์ประกอบย่อย ๆ
2. สร้าง และปรับปรุงแบบทดสอบที่ใช้วัดทักษะย่อย ๆ เหล่านั้น เพื่อให้สามารถค้นหาจุดบกพร่องในทักษะย่อย ๆ นั้นได้

สิงห์ (Singha, 1974 : 200-204) ยังกล่าวถึงการสร้างแบบทดสอบวินิจฉัยว่ามีลักษณะสำคัญ ดังนี้

1. ในกรณีการสร้างแบบทดสอบปรนัยชนิดเลือกตอบหรือแบบตอบสั้นๆ ควรมีจำนวนข้อไม่น้อยกว่าสามข้อในแต่ละเนื้อหาย่อย
2. ไม่จำเป็นต้องสร้างตารางวิเคราะห์หลักสูตร ทั้งนี้เพราะไม่ต้องการหาความสัมพันธ์ระหว่างเนื้อหา และวิธีการ
3. แบบทดสอบวินิจฉัยจะเรียงข้อสอบตามเนื้อหา คือ เอาข้อความที่อยู่ในเนื้อหาเดียวกันเข้าไว้ด้วยกัน โดยไม่จำเป็นต้องคำนึงถึงความยาก
4. แบบทดสอบวินิจฉัยอาจสร้างโดยแบบทดสอบมาตรฐาน (Standardized Test) หรือแบบทดสอบที่ครูสร้างขึ้น (Teacher-made Test) แต่แบบทดสอบที่ครูสร้างขึ้นมักจะคุ้มค่ากว่า เพราะประหยัดเวลาและกำลังงานมากกว่า เมื่อเทียบกับแบบทดสอบมาตรฐาน

ลินด์ควิสต์ (Lindquist, 1966 : 37-38) ได้กล่าวถึงเกณฑ์การสร้างแบบทดสอบวินิจฉัยเพื่อให้มีคุณภาพยิ่งขึ้นไว้ดังนี้

1. แบบทดสอบวินิจฉัยต้องสัมพันธ์กับหลักสูตร และมีความชัดเจนในวัตถุประสงค์ที่ต้องการทดสอบ
2. คำถามในแต่ละแบบทดสอบต้องสร้างให้สามารถวัดได้ตรงตามจุดประสงค์ที่ต้องการทดสอบ
3. ทำการวิเคราะห์ข้อสอบอย่างละเอียด โดยอาศัยการทดลอง และอุปกรณ์หรือความไม่เข้าใจในการเรียนเป็นหลัก
4. แบบทดสอบต้องสามารถวัดพฤติกรรมทางสมองของนักเรียนได้อย่างเพียงพอ และต้องใช้ค้นหาจุดบกพร่องทางการเรียนได้

5. แบบทดสอบต้องเสนอแนะจุดบกพร่องในแต่ละองค์ประกอบทางการเรียนที่ทำการวัดได้อย่างถูกต้อง

6. แบบทดสอบต้องครอบคลุมกฎเกณฑ์ทางการเรียนรู้อย่างทั่วถึง

7. แบบทดสอบต้องสามารถทดสอบความบกพร่องทางการเรียนที่ผ่านมาได้ และสามารถสืบหาความบกพร่องนั้นจากเนื้อหาแต่ละตอนที่ทำการสอบได้

8. ความก้าวหน้าของเด็กแต่ละคน จะแสดงให้เห็นได้จากคำตอบที่แบบทดสอบทำการวัด

นอลล์ (Noll, 1957 : 363) ได้กล่าวถึงลำดับขั้นในการสร้างแบบทดสอบวินิจฉัยไว้ดังนี้

1. การสร้างแบบทดสอบวินิจฉัยจะต้องมีการวิเคราะห์กฎ หลักการ ความรู้ หรือทักษะ ซึ่งเกี่ยวข้องกับสิ่งที่ต้องการวัดอย่างละเอียด

2. แบบทดสอบวินิจฉัยที่ดีจะต้องสร้างให้ครอบคลุมกฎและหลักการต่าง ๆ

3. ข้อสอบควรจัดเรียงเป็นกลุ่ม เพื่อทำให้ง่ายขึ้นเวลาวิเคราะห์คำตอบและวินิจฉัยสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ (2539 : 11)

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

- ค่าความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม (IOC)

- ค่าความลำเอียง (Bias)

9. ทดลองสอบ หาค่าสถิติ ปรับปรุงคุณภาพข้อสอบ

10. จัดแบบทดสอบ ทดลองสอบ หาคคุณภาพของแบบทดสอบ

11. เขียนคู่มือการสร้าง และพัฒนาแบบทดสอบ คู่มือการใช้แบบทดสอบ การแปลความหมายของคะแนน และคู่มือในการวินิจฉัย

สุเทพ สันติวรานนท์ (2533 : 71) ได้สรุปขั้นตอนในการสร้างแบบทดสอบวินิจฉัยในเชิงปฏิบัติ ดังนี้

1. กำหนดจุดมุ่งหมายและวางแผนในการดำเนินการสร้างแบบทดสอบ
2. วิเคราะห์ทักษะที่จำเป็นและเนื้อหาวิชาการอย่างละเอียด แล้วแบ่งออกเป็นองค์ประกอบย่อย ๆ
3. เขียนจุดประสงค์เชิงพฤติกรรม (Behavioral Objective) ให้ครอบคลุม เนื้อหาที่กำหนด
4. เขียนข้อสอบให้สอดคล้องกับจุดประสงค์เชิงพฤติกรรม ในข้อสอบจะกำหนดให้นักเรียนหาคำตอบและสาเหตุของการเลือกตอบ ซึ่งในขั้นนี้ถือเป็นการสร้างแบบทดสอบ เพื่อสำรวจหาสาเหตุของการเลือกตอบ
5. นำไปสอบกับนักเรียนในกลุ่มที่ได้เรียนเนื้อหานั้นผ่านมาแล้ววิเคราะห์คำตอบและหาสาเหตุของการไม่สัมฤทธิ์ผลตามจุดประสงค์เชิงพฤติกรรมจากแบบทดสอบเพื่อสำรวจทั้งนี้เพื่อนำผลการวิเคราะห์มากำหนดสร้างตัวเลือกของข้อสอบวินิจฉัยต่อไป
6. เขียนข้อสอบ โดยตัวเลือกร่างจากสาเหตุของการเลือกตอบของนักเรียน
7. นำข้อสอบในขั้น 6 มารวมเป็นฉบับแบบทดสอบวินิจฉัย แล้วนำไปทดลองใช้และพัฒนาปรับปรุงคุณภาพให้ดีขึ้น
8. เขียนคู่มือในการใช้แบบทดสอบ และกำหนดแนวทางที่เหมาะสม เพื่อให้สามารถบ่งชี้ถึงความบกพร่องในแต่ละทักษะนั้นได้

ส่วนสมศักดิ์ สินธุระเวชญ์ (2522 : 1-2) อธิบายถึงลำดับขั้นตอนการสร้างแบบทดสอบวินิจฉัยไว้ดังนี้

1. แยกแยะหรือวิเคราะห์ทักษะใหญ่ออกเป็นทักษะย่อย เช่น ทักษะในการอ่านอาจแยกเป็นความเข้าใจในการอ่านศัพท์ การแยกคำออกเป็นพยางค์ การแยกแยะเสียงต่าง ๆ การอ่านคำควบกล้ำ เป็นต้น

2. การสร้างคำหรือข้อคำถามแต่ละทักษะย่อย ๆ เหล่านั้น ไม่จำเป็นต้องยากแต่ควรมีจำนวนมาก ๆ ข้อ

จากขั้นตอนในการสร้างแบบทดสอบวินิจฉัยที่กล่าวมา ผู้วิจัยสรุปเป็นขั้นตอนการสร้างได้ดังนี้

1. ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบทดสอบวินิจฉัย หลักสูตร คู่มือคู่มือหนังสือประกอบการเรียนในรายวิชาที่จะสร้าง
2. ตั้งจุดมุ่งหมายและวางแผนในการสร้างแบบทดสอบวินิจฉัย

3. วิเคราะห์เนื้อหาวิชาที่ต้องการทดสอบออกเป็นองค์ประกอบย่อย ๆ ให้ชัดเจนและเขียนจุดประสงค์เชิงพฤติกรรมตามองค์ประกอบย่อยนั้น

4. เขียนข้อสอบตามจุดประสงค์ที่ต้องการวัดพฤติกรรมที่กำหนดไว้
5. ตรวจสอบความเที่ยงตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ
6. นำแบบทดสอบไปทดลองใช้ (Try out)
7. ปรับปรุง จัดทำคั่นฉบับแบบทดสอบ นำไปทดสอบ
8. กำหนดเกณฑ์ขั้นต่ำสำหรับการวินิจฉัย
9. หาเกณฑ์ปกติของแบบทดสอบวินิจฉัย
10. เขียนคู่มือในการใช้แบบทดสอบวินิจฉัย

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 นักเรียนแต่ละห้องเรียน ซึ่งแบ่งคละระหว่างเด็กเก่ง ปานกลาง และอ่อน การสุ่มแบบนี้มีลักษณะการสุ่มแบบแบ่งกลุ่ม (Cluster Random Sampling) คือสุ่มแบบแบ่งกลุ่มจากประชากรกลุ่มย่อย ทั้ง 3 วิธีนี้ใช้ในการสุ่มเพื่อสร้างเกณฑ์ปกติมากที่สุด

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

1.3 มีความทันสมัย เกณฑ์ปกตินั้นขึ้นอยู่กับความสามารถของประชากรกลุ่มนั้น การพัฒนาคนมีอยู่ตลอดเวลา เทคโนโลยี สภาพแวดล้อม อาหารการกิน สิ่งเหล่านี้ จะช่วยให้คนเก่งหรืออ่อนได้ ดังนั้นเกณฑ์ปกติที่เคยศึกษาไว้นานแล้วหลายปี อาจมีความผิดพลาดจากความเป็นจริง จำเป็นต้องสร้างขึ้นใหม่ให้ทันสมัย โดยทั่วไปแล้วเกณฑ์ปกติควรเปลี่ยนทุก ๆ 5 ปี

2. ชนิดของเกณฑ์ปกติ

ถ่วน สายยศ และ อังคณา สายยศ (2543 : 315-317) แบ่งชนิดของเกณฑ์ปกติเป็น 2 ลักษณะใหญ่ ๆ

2.1 เกณฑ์ปกติตามลักษณะของประชากร

2.1.1 เกณฑ์ปกติระดับชาติ (National norms) การสร้างเกณฑ์ปกติระดับชาตินั้นใช้ประชากรที่นิยามไว้มากมายทั่วประเทศ เช่น หาเกณฑ์ปกติของวิชาเลขคณิตระดับชั้นประถมศึกษาปีที่ 4 ระดับชาติ ก็ต้องสอบนักเรียนชั้นประถมศึกษาปีที่ 4 ทั่วประเทศ หรือสุ่มตัวอย่างให้ครอบคลุมทั่วประเทศ จำนวนนักเรียนที่จะต้องสอบจึงมีมากมาย เพื่อให้รู้ว่าสร้างเมื่อปี พ.ศ. ไດ ก็ต้องกำหนดวัน เดือน ปีการสร้างไว้ด้วย เพื่อคนใช้เกณฑ์ปกติจะได้รู้ว่าทันสมัยหรือไม่

2.1.2 เกณฑ์ปกติระดับท้องถิ่น (Local norms) เป็นการสร้างเกณฑ์ปกติระดับเล็กลงมาเช่น ระดับจังหวัด หรือระดับอำเภอ การสร้างเกณฑ์ปกติระดับนี้ค่าใช้จ่ายจะน้อยลงและเป็นประโยชน์ในการเปรียบเทียบคะแนนของผู้สอบกับคนทั้งจังหวัดหรืออำเภอ ในการจัดการศึกษาบางครั้งจังหวัดแต่ละจังหวัด อาจเน้นเนื้อหาบางวิชาไม่เหมือนกัน โดยเฉพาะทางด้านวิชาชีพ บางจังหวัดเน้นเกษตร บางจังหวัดเน้นอุตสาหกรรม บางจังหวัดเน้นการประมง วิชาที่มีการเน้นแตกต่างกัน การสร้างเกณฑ์ปกติระดับท้องถิ่นจะมีประโยชน์มาก แต่วิชาพื้นฐานอื่น ๆ ก็สามารถหาเกณฑ์ปกติระดับท้องถิ่นได้เหมือนกัน เพื่อประโยชน์ในการเปรียบเทียบความสามารถในวิชาการของ นักเรียนคนหนึ่งกับคนทั้งจังหวัดหรืออำเภอว่าเด็กคนนั้นสอบแล้วจะอยู่ในระดับใด เก่งหรืออ่อนกว่าคนอื่นเพียงใด จะได้หาทางปรับปรุง แก้ไขทัน ถ้าไม่มีการเปรียบเทียบกับไม่สามารถพัฒนาได้ถูกต้อง

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

2.2 แบ่งเกณฑ์ปกติตามลักษณะของการใช้สถิติเปรียบเทียบ

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

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

2.2.3 เกณฑ์ปกติสเตนิน (Stanines norms) คะแนนแบบนี้เป็นคะแนนมาตรฐาน ชนิดหนึ่ง แต่มีค่าเพียง 9 ตัว (Standard nine points) ค่าตั้งแต่ 1 ถึง 9 คะแนนเฉลี่ยอยู่ที่คะแนน 5 มีคะแนนเบี่ยงเบนมาตรฐานประมาณ 2 คะแนน วิธีการหามักจะเทียบจากเปอร์เซ็นต์ของความถี่ที่คะแนนเรียงตามค่าจะสะดวกกว่า

2.2.4 เกณฑ์ปกติตามอายุ (Age norms) แบบทดสอบมาตรฐานบางอย่างหาเกณฑ์ปกติตามอายุ เพื่อดูพัฒนาการในเรื่องเดียวกันว่า อายุต่างกันจะมีพัฒนาการอย่างไร โดยมากจะเป็นแบบทดสอบวัดเชาว์ปัญญาและความถนัดจะหาเกณฑ์ปกติโดยวิธีนี้ ส่วนแบบทดสอบวัดผลสัมฤทธิ์จะหาเฉพาะแบบทดสอบที่เป็นวิชาพื้นฐาน เช่น ภาษา และคณิตศาสตร์ โดยเนื้อหาวิชาจะต้องไม่มีผลด้วย ภาษาเช่นคำศัพท์ สามารถหาได้ตั้งแต่อายุ 5 ปี ถึง 20 ปี ความสามารถในการวกลบ คูณ หาร ก็สามารถหาได้ในช่วงอายุดังกล่าวเหมือนกัน ทั้งนี้เพื่อจะดูว่าคำศัพท์ที่กำหนดไว้จำนวนหนึ่งนั้น ถ้านักเรียนคนหนึ่งอายุ 10 ปี สอบได้จำนวนหนึ่ง ลองไปเทียบกับเกณฑ์ปกติดูว่า

น่าจะเป็นความสามารถคำศัพท์เท่าอายุเท่าไรแน่ อาจจะเท่ากับ เด็กอายุ 8 ปี 10 ปี หรือ 15 ปี ก็ต้องเปรียบเทียบดูเกณฑ์แบบนี้วัดผลสัมฤทธิ์ใช้น้อยมากแต่จะทำได้เปรียบเทียบกับก็เป็นประโยชน์สำหรับการศึกษาได้

2.2.5 เกณฑ์ปกติตามระดับชั้น (Grade norms) เป็นการหาเกณฑ์ปกติตามระดับชั้นว่าคะแนนเท่าไรควรอยู่ระดับชั้นไหนจึงจะเหมาะสม แบบทดสอบที่จะทำเกณฑ์ปกติชนิดนี้ได้ก็ต้องเป็นเนื้อหาเดียวกัน ดังนั้นการวัดที่มีเนื้อหาแตกต่างกันตามระดับชั้นไม่สามารถที่จะหาเกณฑ์ปกติแบบนี้ได้ เพราะไม่สามารถจะอธิบายการแปลผลเปรียบเทียบได้ ดังนั้นวิชาที่ทำจึงเป็นวิชาพื้นฐาน ดังกล่าวแล้ว ในการสร้างเกณฑ์ปกติอายุตนเอง เช่น คำศัพท์ คณิตศาสตร์เบื้องต้น แบบทดสอบก็ต้องออกความรู้ความสามารถที่กว้าง ๆ เช่น คำศัพท์ก็ให้ครอบคลุมตั้งแต่ชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 แล้วศึกษาว่าระดับชั้นประถมปีที่ 1 จะได้กี่คะแนน ปีที่ 2 จะได้กี่คะแนน ไปเรื่อย ๆ จนถึงชั้นมัธยมปีที่ 6 จะได้กี่คะแนน โดยมากแต่ละระดับชั้นก็จะเป็นช่วงถือว่าการแจกแจงของคะแนนจะซ้อนทับกันไปเป็นระยะไป แต่เมื่อสร้างเสร็จแล้ว ถ้าเด็กคนหนึ่งมาสอบแบบทดสอบฉบับนี้ได้คะแนน 20 คะแนน และกำลังเรียนชั้นมัธยมปีที่ 2 แต่เทียบแล้วเท่ากับนักเรียนชั้นประถมปีที่ 6 จะได้นำไปพัฒนาต่อ

คะแนนที่ปกติ (Normalized T-score)

ภัทรา นิคมานนท์ (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 ($T = 10Z + 50$) เพราะไม่ต้องคำนวณหาค่าเฉลี่ย (\bar{X}) ความเบี่ยงเบนมาตรฐาน (S) คะแนนซี (Z) เราสามารถแปลงคะแนนได้ครั้งละจำนวนมากไม่เหมือนคะแนนที่แนวเส้น ซึ่งต้องแปลทีละจำนวน

บุญเรียง ขจรศิลป์ (2527 : 112-114) กล่าวว่า คะแนนมาตรฐาน T ปกติ (Normalized T – Score) คือคะแนนมาตรฐานที่ได้จากข้อมูล ที่มีการกระจายเป็นโค้งปกติ หรือโดยการปรับให้เป็นโค้งปกติ การเปลี่ยนคะแนนดิบ ให้เป็นคะแนนมาตรฐาน T โดยใช้สูตร $T = 50 + 10Z$ นั้นเป็นการเปลี่ยนในลักษณะ ที่เรียกว่า Linear transformation ซึ่งการเปลี่ยนแปลงคะแนนแบบนี้ไม่

ทำให้ลักษณะของการกระจายของคะแนนเดิมเปลี่ยนไป เมื่อเปลี่ยนเป็น T-score แล้ว ลักษณะของการกระจายจะยังคงเหมือนกับการกระจายของคะแนนดิบ แต่การเปลี่ยนคะแนนดิบ ให้เป็นคะแนนมาตรฐานปกตินั้น เป็นการเปลี่ยนในลักษณะที่เรียกว่า Area transformation ซึ่งการเปลี่ยนแปลงคะแนนแบบนี้ เมื่อเปลี่ยนมาเป็น คะแนน T ปกติแล้ว จะทำให้การกระจายมีลักษณะเป็นโค้งปกติ

สำหรับงานวิจัยครั้งนี้ได้สร้างเกณฑ์ปกติระดับท้องถิ่น (Local norms) ซึ่งเป็นเกณฑ์ปกติระดับจังหวัด เพื่ออธิบายถึง ข้อเท็จจริงทางสถิติที่บรรยายการแจกแจงของคะแนนของกลุ่มตัวอย่าง ซึ่งได้จากแบบทดสอบวินิจฉัย และเป็นคะแนนตัวแทนที่บอกระดับการสอบของผู้สอบว่าอยู่ในระดับใดของกลุ่มประชากร โดยแสดงลงในตารางซึ่งบอกความสัมพันธ์ระหว่างคะแนนดิบกับคะแนนที่ปกติ (Normalized T – Score) ที่แปลงมาจากคะแนนดิบ

ทฤษฎีการตอบสนองข้อสอบ

1. หลักการของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีที่อธิบายความสัมพันธ์ระหว่างลักษณะ หรือความสามารถที่อยู่ภายในตัวบุคคลกับพฤติกรรมการตอบสนองข้อสอบของบุคคลนั้น โดยทฤษฎีนี้มีความเชื่อว่าพฤติกรรมการตอบสนองข้อสอบของผู้สอบจะถูกกำหนดโดยคุณลักษณะ (Trait) หรือความสามารถ (Ability) ที่มีอยู่ภายในตัวบุคคลซึ่งไม่สามารถสังเกตได้โดยตรง ความสัมพันธ์ดังกล่าวแสดงออกมาในรูปแบบทางคณิตศาสตร์ที่กำหนดด้วยฟังก์ชันทางคณิตศาสตร์ (Item Characteristic Function) หรือเรียกว่าโค้งลักษณะข้อสอบ (Item Characteristic Curve) หรือฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) ซึ่งกำหนดได้หลายรูปแบบ แต่ทั้งนี้ จะใช้ รูปแบบของทฤษฎีการตอบสนองข้อสอบได้ จะขึ้นอยู่กับข้อมูลที่ได้จากการทดสอบว่าเป็นไปตามข้อตกลงเบื้องต้นหรือไม่ (Lord and Novick, 1968 : 360)

2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

การนำทฤษฎีการตอบสนองข้อสอบมาใช้ในการวิเคราะห์ข้อสอบ จะต้องคำนึงถึงลักษณะของแบบทดสอบให้เป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ดังนี้

2.1 แบบทดสอบที่มีมิติเดียว (Unidimension Test) หมายความว่าข้อสอบแต่ละข้อในแบบทดสอบจะต้องวัดความสามารถหรือคุณลักษณะเดียวกัน หรือมีความเป็นเอกพันธ์กัน

2.2 ข้อสอบแต่ละข้อจะต้องเป็นอิสระจากกัน (Local Independence) หมายความว่า การตอบข้อสอบข้อใดข้อหนึ่งจะไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ

2.3 โอกาสที่ผู้เข้าสอบจะตอบข้อสอบข้อนั้นถูก เป็นไปตามโค้งลักษณะข้อสอบ (Item Characteristic Curve ; ICC) ของแต่ละโมเดลที่ใช้ ไม่ขึ้นกับการแจกแจงความสามารถของกลุ่มตัวอย่าง

3. พารามิเตอร์ของทฤษฎีการตอบสนองข้อสอบ

แนวคิดของทฤษฎีการตอบสนองข้อสอบสามารถแบ่งพารามิเตอร์ออกเป็น 2 ชนิด (สุพรรณ สุกมลสันต์, 2538 : 98) ดังนี้

3.1 พารามิเตอร์ของข้อสอบ (Item Parameter) ได้แก่

3.1.1 ค่าอำนาจจำแนก (a) หมายถึง ความสามารถของข้อสอบที่แยกเด็กออกเป็น 2 พวก คือพวกตอบถูกกับพวกตอบผิด หมายถึง ค่าที่เป็นสัดส่วนโดยตรงกับความชันของโค้งคุณลักษณะของข้อสอบ ณ จุดเปลี่ยน โค้งมีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติมีค่าตั้งแต่ 0 ถึง 2 เพราะค่า a ที่เป็นลบแสดงว่าข้อสอบไม่ดี ใช้ไม่ได้ต้องตัดทิ้ง ค่า 0 แสดงว่าข้อสอบไม่มีค่าอำนาจจำแนก ค่า +2 แสดงว่าข้อสอบมีค่าอำนาจจำแนกสูง ในการคัดเลือกข้อสอบจะเลือก ข้อสอบที่มีค่า a ตั้งแต่ 0.5 ขึ้นไป

3.1.2 ค่าความยาก (b) หมายถึง ค่าที่แสดงถึงระดับความสามารถของผู้สอบ (θ) ที่จุดโค้งลักษณะข้อสอบมีความชันมากที่สุด มีค่าตั้งแต่ $-\infty$ ถึง ∞ แต่ในทางปฏิบัติมักจะมีค่าอยู่ระหว่าง -3 ถึง +3 ค่า -3 แสดงว่าข้อสอบนั้นง่ายมาก และค่า +3 แสดงว่าข้อสอบนั้นยากมาก

3.1.3 ค่าสัมประสิทธิ์การเดา (c) หมายถึง ความน่าจะเป็นของบุคคลหนึ่งที่มีความสามารถต่ำสุด จะมีความสามารถตอบข้อสอบนั้นได้ถูกต้อง เป็นค่าที่แสดงถึงโอกาสการตอบข้อสอบถูกโดยไม่มีความรู้ในเรื่องนั้น ๆ มีค่าจาก 0 ถึง 1 จะคัดเลือกเอาข้อสอบที่มีค่า c ต่ำกว่า 0.3 ลงมา

3.2 พารามิเตอร์ของผู้สอบ (Examinees Parameter) ได้แก่

ระดับความสามารถของผู้สอบ (θ) หมายถึง ศักยภาพของผู้สอบที่ประมาณได้จากการทำข้อสอบ ทฤษฎีการตอบสนองข้อสอบส่วนใหญ่จะมีค่าอยู่ระหว่าง -3 ถึง +3 ค่า -3 แสดงว่ามีความสามารถต่ำ และค่า +3 แสดงว่ามีความสามารถสูง

4. รูปแบบของโมเดลโลจิสติก

เบิร์นบอม (Birnbaum) เป็นผู้ที่มิบบทบาทสำคัญในการพัฒนารูปแบบของโมเดลโลจิสติก (Hambleton and Swaminathan, 1985 : 36) ซึ่งโมเดลชนิดนี้สามารถแสดงความ

สัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบถูกกับระดับความสามารถ ในรูปของฟังก์ชันการแจกแจงสะสมแบบโลจิสติก (Logistic Cumulation Function) มีโค้งลักษณะข้อสอบเป็นรูปตัว S สามารถแบ่งได้เป็น 3 รูปแบบ ตามจำนวนพารามิเตอร์ของโค้งลักษณะเฉพาะของข้อสอบ ดังนี้

4.1 รูปแบบโลจิสติกที่มี 1 พารามิเตอร์ (One – Parameter Logistic Model)

เบิร์ตบอม (Birnbaum) พัฒนารูปแบบโลจิสติกนี้ขึ้นมาในปี 1968 ซึ่งบังเอิญตรงกับรูปแบบของราส์ช (Rasch Model) เป็นโมเดลที่อธิบายข้อสอบด้วยค่าพารามิเตอร์เพียงตัวเดียว คือค่าความยากโดยเชื่อว่าโอกาสที่ผู้สอบจะทำข้อสอบได้ถูกหรือไม่ ขึ้นอยู่กับระดับความสามารถของตนเองกับค่าความยากของข้อสอบ ดังนั้นจึงถือว่าค่าการเดาเป็นศูนย์ และค่าอำนาจจำแนกของข้อสอบจะคงที่ที่หนึ่ง เขียนเป็นสมการได้ดังนี้ (Birnbaum, 1968 quoted in Hambleton, Swaminathan and Rogers, 1991 :12-14)

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

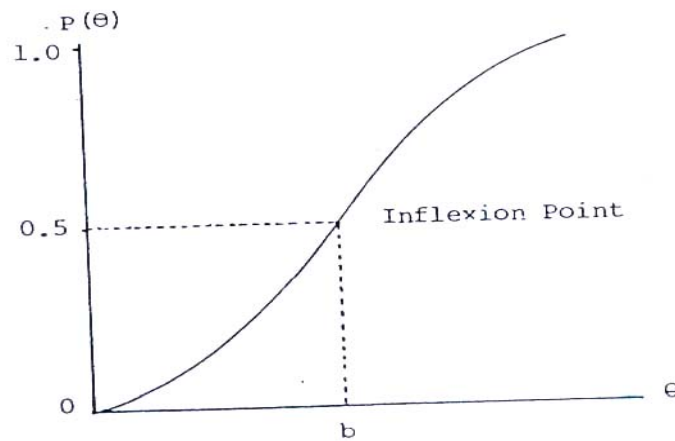
เมื่อ $P_i(\theta)$ คือ โอกาสที่ผู้มีความสามารถ θ จะทำข้อสอบข้อที่ i ได้ถูกต้อง

θ คือ ระดับความสามารถที่แท้จริงของผู้สอบ

b_i คือ ค่าความยากของข้อสอบข้อที่ i

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818

ภาพประกอบ 1 ความหมายของค่า Item Parameter (1 พารามิเตอร์) ของข้อสอบ



ที่มา : ประดิษฐ์ เรื่องตระกูล, 2529 : 19

4.2 รูปแบบโลจิสติกที่มี 2 พารามิเตอร์ (Two – Parameter Logistic Model)

เบียร์นบอม (Birnbaum) ได้พัฒนาโมเดลนี้ขึ้นมาจาก Normal Ogive Model ซึ่งรูปของโมเดลใหม่ก็มีรูปแบบไม่แตกต่างจากเดิมมากนัก จะเปลี่ยนแต่เพียงการคิดคำนวณเท่านั้น และกำหนดให้ทุกข้อไม่มีการเดาเกิดขึ้น คือ ค่า c_i มีค่าเป็นศูนย์ กล่าวคือ ผู้สอบที่มีความสามารถต่ำสุด ไม่มีโอกาสที่จะทำข้อสอบถูกในข้อสอบที่มีค่าความยากสูง ซึ่งเบียร์นบอม ได้เสนอรูปแบบของสมการดังนี้ (Birnbaum, 1968 quoted in Hambleton, Swaminathan and Rogers, 1991 :14-16)

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

เมื่อ $P_i(\theta)$ คือ โอกาสที่ผู้มีความสามารถ θ จะทำข้อสอบข้อที่ i ได้ถูกต้อง

θ คือ ระดับความสามารถที่แท้จริงของผู้สอบ

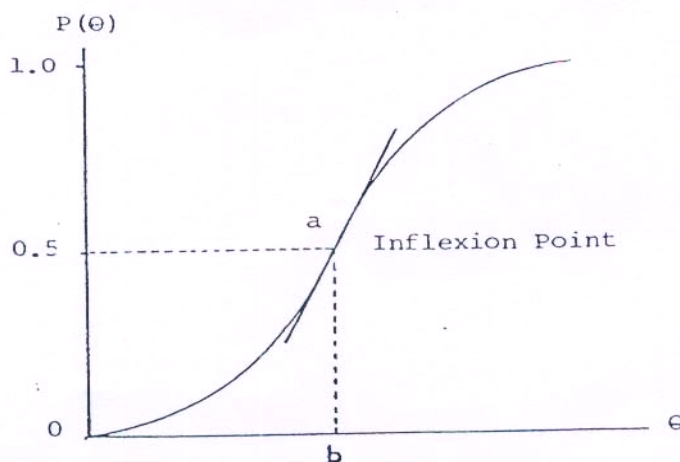
a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i คือ ค่าความยากของข้อสอบข้อที่ i

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818

D คือ Scaling factor มีค่าเท่ากับ 1.7

ภาพประกอบ 2 ความหมายของค่า Item Parameter (2 พารามิเตอร์) ของข้อสอบ



ที่มา : ประดิษฐ์ เรื่องตระกูล, 2529 : 20

4.3. รูปแบบโลจิสติกที่มี 3 พารามิเตอร์ (Three – parameter Logistic Model)

เป็นโมเดลที่พัฒนามาจาก Two-Parameter Logistic Model เพื่อให้เหมาะกับแบบสอบที่มีอิทธิพลจากการเดาเข้ามาแฝงอยู่ด้วย และเป็นโค้งลักษณะข้อสอบที่แสดงถึงลักษณะข้อสอบที่มีค่าพารามิเตอร์ของข้อสอบ 3 ตัว ซึ่งเบิร์นบอมได้เสนอรูปแบบของสมการดังนี้ (Birnbbaum, 1968 quoted in Hambleton, Swaminathan and Rogers, 1991 : 17)

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

เมื่อ $P_i(\theta)$ คือ โอกาสที่ผู้มีความสามารถ θ จะทำข้อสอบข้อที่ i ได้ถูกต้อง

θ คือ ระดับความสามารถที่แท้จริงของผู้สอบ

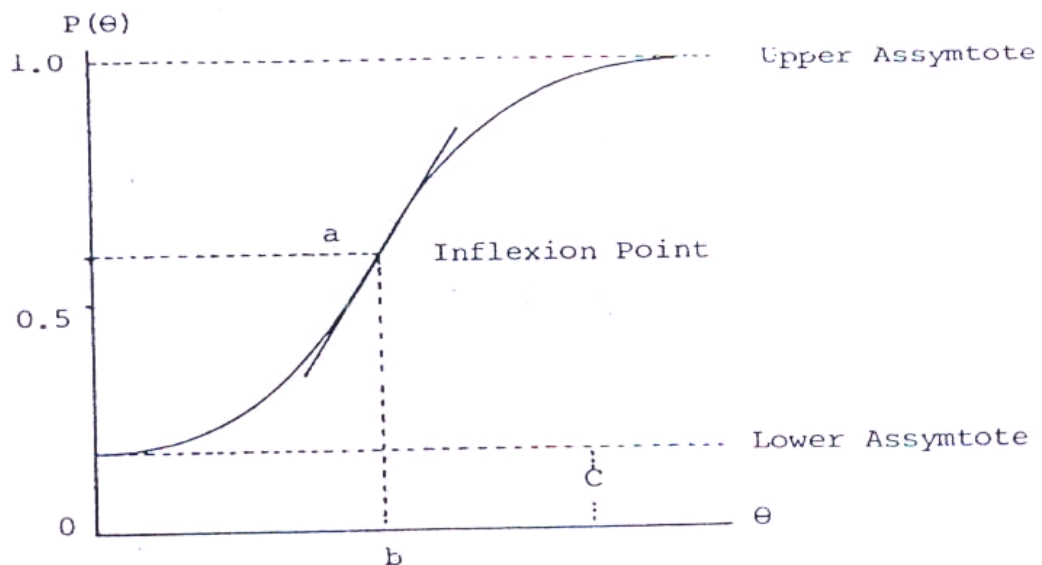
a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i คือ ค่าความยากของข้อสอบข้อที่ i

c_i คือ ค่าการเดาของข้อสอบข้อที่ i

- e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818
 D คือ Scaling factor มีค่าเท่ากับ 1.7

ภาพประกอบ 3 ความหมายของค่า Item Parameter (3 พารามิเตอร์) ของข้อสอบ



ที่มา : ประดิษฐ์ เรื่องตระกูล, 2529 : 20

5. ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบทดสอบ

ในการประมาณค่าความสามารถของผู้สอบด้วยวิธีแมกซิมัมไลกิลิฮูด (Maximum Likelihood) นั้นความแน่นอนของการประมาณค่าความสามารถแสดงได้ในเทอมของค่าฟังก์ชันสารสนเทศ (Information Function) โดยที่ในทฤษฎีการทดสอบดั้งเดิมนั้นเราศึกษาเรื่องความเที่ยง (Reliability) ของคะแนนและความคลาดเคลื่อนมาตรฐานของการวัด (Standard Error of Measurement) ซึ่งค่าที่ได้จะแปรเปลี่ยนไปตามกลุ่มผู้สอบอันเป็นจุดอ่อนประการหนึ่งของการศึกษา แต่ในทฤษฎีการตอบสนองข้อสอบจะศึกษาถึง ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function) แทนการหาค่าความเที่ยง (Hambleton, 1979 : 64)

สำหรับโมเดลโลจิสติก 3 พารามิเตอร์สามารถหาค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information Functions) และค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Item Information Functions) ในแต่ละระดับความสามารถ θ ได้ดังนี้

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

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

เมื่อ $I_i(\theta)$ คือ ค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information Functions)

$P_i'(\theta)$ คือ ความชันของ ICC ที่ระดับความสามารถ θ

$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบมีความสามารถ θ ตอบข้อสอบข้อที่ i ถูก

$Q_i(\theta)$ คือ $1 - P_i(\theta)$

หรือสามารถหาฟังก์ชันสารสนเทศของสมการได้จากสมการดังนี้

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

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

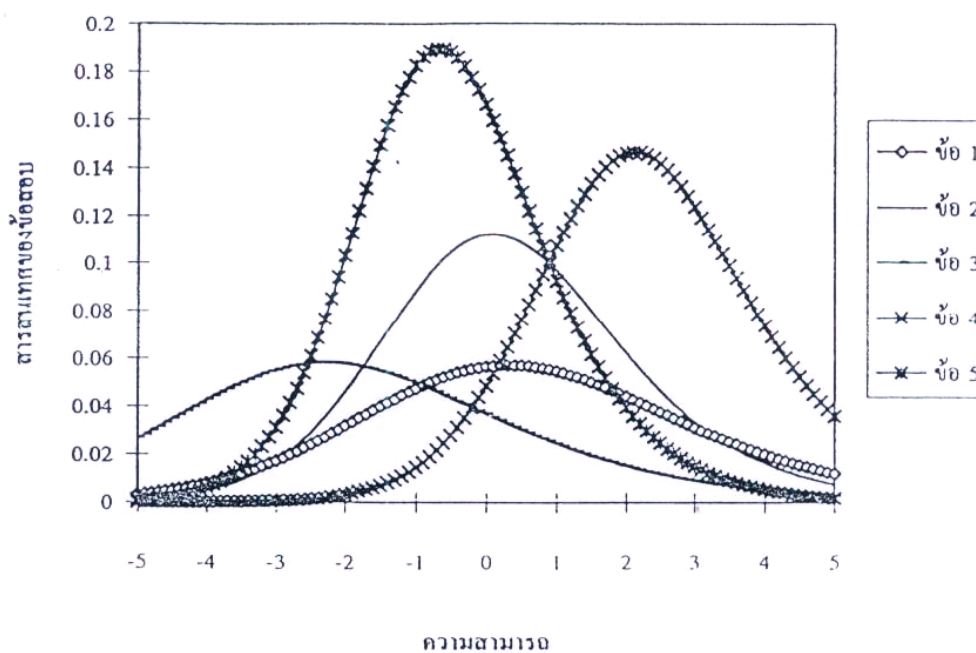
$$I(\theta) = \sum_{i=1}^n I_i(\theta)$$

เมื่อ $I(\theta)$ คือ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function)

n คือ จำนวนข้อสอบในแบบทดสอบ

ค่าฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อจะขึ้นอยู่กับโค้งลักษณะของข้อสอบ (ICC) และความแปรปรวนของการตอบข้อสอบถูกในแต่ละข้อของแต่ละระดับความสามารถ และยังความชันของโค้งลักษณะของข้อสอบมีค่ามาก ๆ ประกอบกับความแปรปรวนมีค่าน้อย ๆ โค้งฟังก์ชันสารสนเทศของข้อสอบที่ระดับความสามารถนั้นจะยิ่งสูงขึ้น สำหรับโค้งฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function) ที่มีค่าสูงสุด ณ ระดับความสามารถใดก็จะจำแนกระดับความสามารถของผู้สอบได้ดี ณ ระดับความสามารถนั้น (Hambleton, 1979 : 66) รายละเอียดแสดงในภาพประกอบ 4 และภาพประกอบ 5

ภาพประกอบ 4 สารสนเทศของข้อสอบจำนวน 5 ข้อ

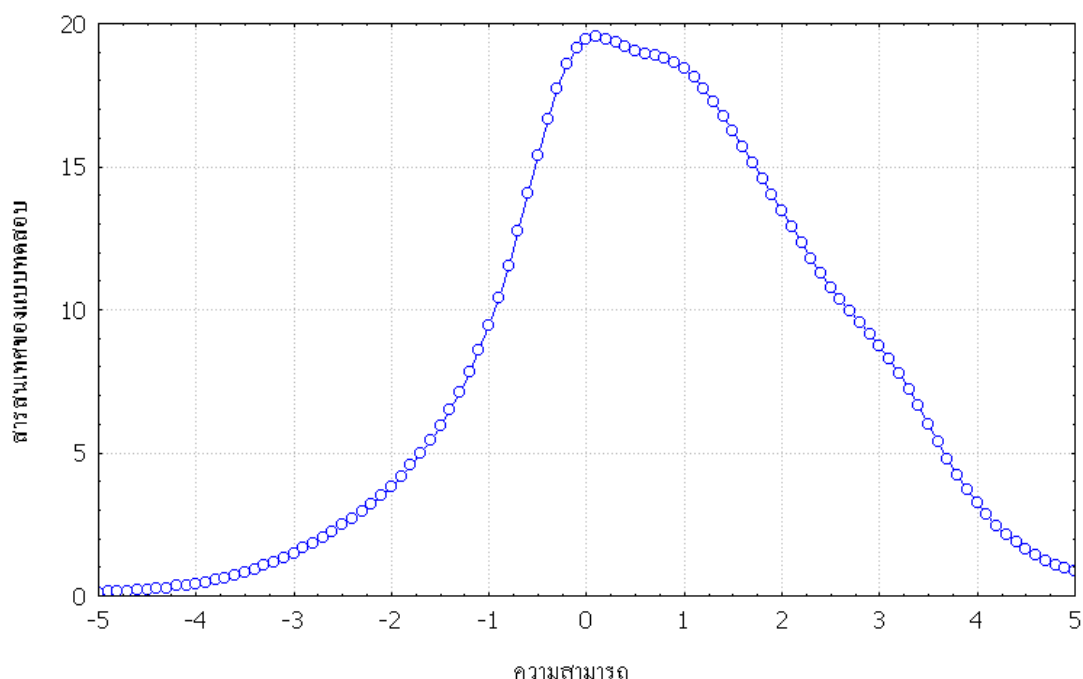


ที่มา: สกว สันติเทวกุล, 2540 :70

ภาพประกอบ 4 แสดงลักษณะของข้อสอบที่เหมาะสมกับนักเรียนที่มีระดับความสามารถต่างกัน คือข้อสอบเหมาะที่จะนำไปใช้ทดสอบกับผู้ที่มีระดับความสามารถปานกลาง 2 ข้อ คือ ข้อ 1 ซึ่งให้สารสนเทศสูงสุดกับผู้สอบที่มีความสามารถ 0.322 และ ข้อ 2 ให้สารสนเทศสูงสุดกับผู้สอบที่มีความสามารถ 0.092 ข้อ 3 และข้อ 5 เหมาะที่จะนำไปใช้ทดสอบกับผู้ที่มีระดับ

ความสามารถต่ำ เพราะมีค่าสารสนเทศสูงสุดกับผู้สอบที่มีความสามารถ -2.304 และ -0.658
 ส่วนข้อ 4 เหมาะที่จะนำไปใช้ทดสอบกับผู้ที่มีระดับความสามารถสูง เพราะให้ค่าสารสนเทศสูงสุด
 กับผู้สอบที่มีความสามารถ 2.120

ภาพประกอบ 5 สารสนเทศของแบบทดสอบ



ที่มา : สกาว สันติเทวกุล, 2540 : 61

ภาพประกอบ 5 แสดงโค้งสารสนเทศของแบบทดสอบมีค่าสูงสุดเท่ากับ 19.526 ที่
 ระดับความสามารถ 0.1 สามารถอธิบายได้ว่าแบบทดสอบฉบับนี้เหมาะที่จะนำไปใช้สอบกับผู้สอบ
 ที่มีระดับความสามารถปานกลาง

ในการสอบทุกครั้งย่อมมีความคลาดเคลื่อนเกิดขึ้น ความคลาดเคลื่อนมาตรฐานในการ
 ประมาณค่า (Standard Error of Estimate : $SE(\theta)$) คือส่วนเบี่ยงเบนมาตรฐานในการประมาณค่า

ความสามารถ (θ) ซึ่งมีความสัมพันธ์กับค่าฟังก์ชันสารสนเทศของแบบทดสอบสามารถคำนวณได้จากสมการ (Hambleton, Swaminathan and Rogers, 1991 : 94)

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

เมื่อ $SE(\theta)$ คือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ
ณ ระดับความสามารถ (θ)

ดังนั้นประโยชน์ที่ได้จากประเด็นข้างต้นก็คือ ถ้ามีกลุ่มของข้อสอบอยู่ชุดหนึ่งที่สามารถทราบโค้งฟังก์ชันสารสนเทศ (Item Information Function) ของแต่ละข้อ เราก็จะสามารถสร้างแบบสอบฉบับหนึ่งให้มี ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function) ณ ระดับหนึ่งของความสามารถตามที่เราต้องการได้ และนั่นหมายถึงว่าเราสามารถสร้างฉบับแบบทดสอบให้เป็นไปตามจุดมุ่งหมายของการสอบได้ เช่น ถ้าต้องการได้แบบสอบคัดเลือก ก็ต้องเลือกใช้ข้อสอบที่มีความสูงสุดของโค้งที่ระดับความสามารถสูง ๆ ซึ่งก็คือ เลือกข้อสอบที่จะให้โค้งฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function) สูงที่ระดับความสามารถสูง ๆ เป็นต้น

6. การประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ

ปัจจุบันได้มีการนำเอาทฤษฎีการตอบสนองข้อสอบโดยเฉพาะโมเดลของราสช์ (Rasch Model) มาใช้อย่างกว้างขวางในการวัดผลและการทดสอบโดยทั่วไป ซึ่งไรต์ (Wright , 1980 : 194-196 อ้างถึงใน ประดิษฐ์ เรื่องตระกูล, 2529 : 26) ได้รวบรวมประโยชน์การนำโมเดลของราสช์ (Rasch Model) มาใช้ไว้มากมาย แต่อย่างไรก็ตามในงานที่โมเดลของราสช์ สามารถทำได้นั้นโมเดลโลจิสติก 3 พารามิเตอร์ย่อมทำได้และอาจทำได้ดีกว่าในบางกรณี ดังนั้นการนำโมเดลโลจิสติก 3 พารามิเตอร์ไปใช้ประโยชน์จึงพอสรุปได้ดังนี้

6.1 ใช้ในการสร้างคลังข้อสอบ (Item Bank) เนื่องจากข้อสอบที่วิเคราะห์แล้วค่าพารามิเตอร์มีลักษณะคงที่ ดังนั้น ข้อสอบเหล่านี้จึงนำมาใช้สร้างข้อสอบชุดใหม่ตามเกณฑ์ที่ต้องการวัด

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

- 6.2.1 ค่าสถิติและค่าพารามิเตอร์ต่างขึ้นอยู่กับสภาพกลุ่มตัวอย่างที่ทำการสอบ
- 6.2.2 การเปรียบเทียบความสามารถหรือคุณลักษณะ (Trait) ใด ๆ ในแต่ละบุคคลจะเปรียบเทียบกันได้ก็ต่อเมื่อต้องสอบด้วยแบบสอบฉบับเดียวกัน
- 6.2.3 ความคลาดเคลื่อนมาตรฐานในการวัดของแต่ละคนเท่ากัน ซึ่งลอร์ด และ โนวิก (Lord and Novick, 1968) ได้แสดงให้เห็นว่าความคลาดเคลื่อนมาตรฐานในการวัดสำหรับคะแนนปานกลางจะมีขนาดเล็กกว่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนสูงหรือคะแนนต่ำ
- 6.3 ใช้ในการกำหนดเกณฑ์ของระดับความสามารถของแบบทดสอบอิงเกณฑ์ ผลของการวิเคราะห์ข้อสอบถูกดัดแปลงให้เป็นคะแนนความสามารถของแต่ละกลุ่มบุคคลจึงสามารถเปรียบเทียบกับคะแนนความสามารถซึ่งเป็นเกณฑ์คงที่ได้ ทำให้เราทราบได้ว่า ระดับความสามารถอิงเกณฑ์แต่ละชุด (Minimum Mastery Level) ของข้อสอบแต่ละข้อควรเป็นอย่างไร
- 6.4 ใช้วินิจฉัยความสามารถของผู้สอบ (Diagnostic) ในกรณีที่ ใ้คงคุณลักษณะของข้อสอบ (ICC) ของข้อสอบไม่เหมาะสมกับโครงสร้างของโมเดล แสดงว่าบางสิ่งบางอย่างผิดปกติในตัวผู้สอบที่เราควรสนใจแก้ไข
- 6.5 ใช้ในการค้นหาความเป็นอคติของข้อสอบ (Item Bias) เมื่อเกิดเหตุดังกล่าวในข้อ 6.3 แสดงว่า มีความสามารถอื่นแฝงเข้ามาในความสามารถที่ต้องการวัด เราก็สามารถที่จะทำการตรวจสอบความลำเอียงของข้อสอบได้
- 6.6 ใช้ในการวัดระดับความสามารถของแต่ละบุคคล (Tailoring Test) เราอาจสุ่มข้อสอบที่วิเคราะห์แล้ว และมีความยากแยกตามลำดับมาสอบระดับความสามารถของแต่ละบุคคลได้
- 6.7 ใช้ในการจัดชั้นเรียน (Grade – Placement Tailoring) ค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์ข้อสอบอาจใช้ในการจัดชั้นเรียนที่เหมาะสมกับระดับความสามารถของผู้เรียนเป็นกลุ่ม ๆ ได้
- 6.8 ใช้ในการสร้างแบบทดสอบที่ดีที่สุด (Best Test Design) ผลจากการวิเคราะห์ข้อสอบโดยทฤษฎีคุณลักษณะภายใน (Latent Trait Models) สามารถนำไปใช้ในการสร้างข้อสอบที่มีคุณลักษณะต่าง ๆ ตามที่ต้องการได้
- 6.9 ใช้ในการเปรียบเทียบคะแนนต่างชุด (Equating Score) ในแบบสอบที่วิเคราะห์แล้ว 2 ชุดที่ต่างกัน แต่วัดในสิ่งเดียวกัน จะสามารถนำคะแนนของผู้สอบในแบบสอบฉบับหนึ่งเปรียบเทียบกับคะแนนของผู้สอบในแบบสอบอีกฉบับหนึ่งได้ ทั้งนี้เพราะคะแนนแต่ละชุดของทฤษฎีการตอบสนองข้อสอบ นั้นถูกแปลงเป็นคะแนนมาตรฐานที่สามารถเปรียบเทียบกันได้

6.10 ใช้หาค่าการเดา (Guessing) ในทฤษฎีการทดสอบดั้งเดิมได้มีการใช้สูตรแก้การเดา

แต่ในทางปฏิบัติจริงแล้วถ้าผู้สอบมีเวลาเพียงพอ มีตัวเลือกที่มีประสิทธิภาพ และผู้สอบได้แสดงความสามารถอย่างเต็มที่ การเดาแบบสุ่มน่าจะเกิดขึ้นได้น้อยมาก ถ้าเป็นเช่นนี้สูตรแก้การเดาจึงเป็นเสมือนการลงโทษผู้สอบผู้นั้น สำหรับในทฤษฎีการตอบสนองข้อสอบจะไม่ใช้สูตรแก้การเดา แต่จะใช้การพิจารณาค่าการเดา ซึ่งหาได้จากการนำโมเดลทางคณิตศาสตร์เข้ามาช่วยสนับสนุนทำให้หลักการพิจารณาค่าการเดามีความสมเหตุสมผลยิ่งขึ้น

7. สรุปแนวคิดทฤษฎีการตอบสนองข้อสอบ

สาระสำคัญของทฤษฎีการตอบสนองข้อสอบ ก็คือ ค่าคุณลักษณะของข้อสอบ (a, b และ c) เป็นค่าพารามิเตอร์ประจำข้อสอบ นั่นคือค่าคุณลักษณะข้อสอบจะเป็นค่าคงที่ ไม่ว่าจะนำไปสอบกับผู้สอบกลุ่มใด ต่างกับทฤษฎีการทดสอบดั้งเดิม (Classical Test Theory) ซึ่งความยากของข้อสอบขึ้นอยู่กับกลุ่มนักเรียนที่เรา นำข้อสอบนั้นไปสอบ ถ้ากลุ่มนักเรียนนั้นเก่ง ข้อสอบนั้นก็จะถูกวิเคราะห์ว่าข้อสอบนั้นง่าย เพราะคนส่วนใหญ่ในกลุ่มทำได้ ในทำนองเดียวกันถ้า นำข้อสอบข้อเดียวกันมาสอบกลุ่มนักเรียนที่อ่อน ข้อสอบข้อนั้นจะถูกวิเคราะห์ออกมาว่ายาก เพราะคนส่วนใหญ่ในกลุ่มทำไม่ได้ แต่สำหรับทฤษฎีการตอบสนองข้อสอบนั้น ถ้า นำไปสอบกับนักเรียนอ่อนก็จะได้ความสัมพันธ์ของสมการในตำแหน่งความสามารถต่ำ แต่ถ้าไปสอบกับนักเรียนเก่งก็จะได้ความสัมพันธ์ของสมการในตำแหน่งความสามารถสูง ซึ่งเป็นโค้งความสัมพันธ์อันเดียวกัน

ค่าพารามิเตอร์ของข้อสอบในทฤษฎีการตอบสนองข้อสอบแต่ละตัวมีความหมายดังนี้

θ : คือ ระดับความสามารถของผู้สอบในสเกลคะแนนมาตรฐาน อยู่ระหว่าง -3 ถึง +3 โดยประมาณ ผู้ที่มีความสามารถปานกลางจะมีค่า θ อยู่ระหว่าง -1.0 ถึง +1.0 ถ้าค่า θ มากกว่า 1.0 แสดงว่า ค่าความสามารถค่อนข้างสูงถ้า (น้อยกว่า -1.0 แสดงว่า ค่าความสามารถค่อนข้างต่ำ)

a : คือ ค่าอำนาจจำแนก (Item Discriminating) ของสมการที่จับบน โค้งที่ตรงกับค่าโอกาสที่จะทำข้อสอบถูกเท่ากับ 0.5 (ค่าเกณฑ์ตั้งที่ 0.5) เป็นค่าพารามิเตอร์ที่แสดงถึงความสามารถของข้อสอบในการจำแนกบุคคลที่มีความสามารถต่างไปออกได้มากน้อยเพียงใด ถ้าค่าความชันสูงคนที่มีความสามารถต่างกันเพียงเล็กน้อยก็ส่งผลให้โอกาสที่ทำข้อสอบถูกต่างกันมาก โดยปกติค่าควรจะมากกว่า 0.3 ขึ้นไป

$a > 0.5$	การจำแนกดี
$0.3 < a < 0.5$	การจำแนกปานกลาง
$a < 0.3$	การจำแนกไม่ดี

b : คือ ค่าความยาก (Item Difficulty) คือตำแหน่งบนสเกลของค่าความสามารถ (แกนนอน) ที่ตรงกับ ค่าโอกาสที่จะทำข้อสอบถูกต้องเท่ากับ 0.5 นั่นคือค่าความยากก็คือ ค่าความสามารถของผู้สอบที่มีโอกาสทำข้อสอบข้อนั้นถูกร้อยละ 50 ค่าความยากเป็นค่าที่บอกถึงความเหมาะสมในการที่จะเอาข้อสอบข้อนี้ไปใช้ ถ้าค่าความยากสูงกว่า 1.0 ก็ควรนำไปใช้กับผู้มีความสามารถสูง (กลุ่มเก่ง) เช่น นำไปใช้ในการสอบแข่งขัน ถ้าค่าความยากต่ำกว่า -1.0 ก็ควรนำไปใช้กับผู้ที่มีความสามารถต่ำ (กลุ่มอ่อน) เช่น นำไปใช้กับนักเรียนอ่อนเพื่อวินิจฉัยข้อบกพร่องในการเรียนวิชานั้น ส่วนข้อที่มีค่า b อยู่ระหว่าง -1.0 ถึง 1.0 เหมาะสำหรับไปใช้กับนักเรียนที่มีความสามารถปานกลาง เช่น ใช้ในการเรียนการสอนในชั้นปกติ หรือเรียกว่า วัตถุประสงค์

c : คือ ค่าการเดา (Item Guessing) เป็นค่าจุดตัดบนแกนตั้ง (ค่าโอกาสที่จะทำข้อสอบข้อนั้นถูก) เพราะเป็นโอกาสที่จะทำข้อสอบนั้นถูกโดยไม่มีความสามารถมีค่าที่เป็นไปได้อยู่ระหว่าง 0 ถึง 1 ถ้าข้อสอบข้อนั้นไม่สามารถที่จะตอบถูกด้วยการเดาแล้ว $c_i = 0$ และปกติควรจะ มีค่าน้อยกว่า 0.3

$c < 0.2$	ตัวเลือกดีมาก
$0.2 < c < 0.3$	ตัวเลือกดี
$c > 0.3$	ตัวเลือกเดาง่าย

งานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมงานวิจัยที่เกี่ยวข้องกับแบบทดสอบวินิจฉัยและทฤษฎีการตอบสนองข้อสอบ ดังนี้

สุภาพร ละอองวิจิตร (2543 : บทคัดย่อ) ได้พัฒนาแบบทดสอบวินิจฉัยทางการเรียน วิชาคณิตศาสตร์เรื่องฟังก์ชัน โดยใช้ทฤษฎีการตอบสนองข้อสอบ สำหรับนักเรียนชั้นมัธยมศึกษา ปีที่ 4 วิเคราะห์ข้อสอบโดยใช้โปรแกรม Bilog Version 3.04 ตรวจสอบคุณภาพของแบบทดสอบ ด้วยความเชื่อมั่นแนวอิงเกณฑ์ด้วยสูตรแบบ Binomial รวมทั้งความแม่นยำในการประมาณค่า ความสามารถที่แท้จริง ซึ่งหาได้จาก Test Information Function ส่วนด้านความตรงได้พัฒนา แบบทดสอบให้มีความตรงเชิงเนื้อหาตามวิธีของ โรวินลลีและแฮมเบิลตัน หากจะเน้นจุดตัด โดยวิธีโลจิสติก และหาเกณฑ์ปกติในรูปเปอร์เซ็นต์ไทล์ (Percentile Norms) และคะแนนที่ปกติ

(Normalized T-score) ได้แบบทดสอบที่มีคุณภาพดังนี้ มีความตรงเชิงเนื้อหา โดยมีดัชนีความสอดคล้องอยู่ในช่วงระหว่าง 0.67-1.00 มีคุณสมบัติการวัดเพียงคุณลักษณะเดียว มีค่าอำนาจจำแนกตั้งแต่ 0.449 ถึง 7.959 ค่าความยากตั้งแต่ -0.795 ถึง 1.898 และค่าการเดาตั้งแต่ 0.085 ถึง 0.466 มีค่าความเที่ยงตรงตามแนวอิงเกณฑ์ของโดเมนที่ 1 ถึงโดเมนที่ 6 และทั้งฉบับเป็น 0.949, -0.140, 0.798, 0.705, 0.776, 0.797 และ 0.956 ตามลำดับ สำหรับความแม่นยำในการประมาณค่าที่แท้จริง พบว่าแบบทดสอบมีประสิทธิภาพสูงสุดเมื่อใช้วัดผู้สอบที่มีความสามารถปานกลางและค่อนข้างสูง คือ ค่า θ อยู่ในช่วง -0.3 ถึง 2.2 คะแนนจุดตัดโดยวิธีโลจิสติกของโดเมนที่ 1 ถึงโดเมนที่ 6 และทั้งฉบับเป็น 21, 14, 3, 3, 5 และ 38 ตามลำดับ

จินตนา สีนกั้ง (2542 : ออนไลน์) ศึกษาเรื่องการพัฒนาแบบทดสอบวินิจฉัยทักษะพื้นฐานของความเข้าใจในการอ่านภาษาอังกฤษ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ในจังหวัดสกลนคร โดยใช้แบบทดสอบเลือกตอบ 4 ตัวเลือก จำนวน 3 ฉบับ ประกอบด้วย ฉบับที่ 1 คำนาม คำสรรพนามและคำที่แสดงความเป็นเจ้าของ ฉบับที่ 2 คำกริยา คำศัพท์ และฉบับที่ 3 คำบุพบท การตั้งคำถาม สรุปผลการวิจัยพบว่า แบบทดสอบวินิจฉัยฉบับที่ 1 มีค่าความยากตั้งแต่ .22-.81 ค่าอำนาจจำแนก ตั้งแต่ .23-.86 ค่าความเชื่อมั่นเท่ากับ .772 แบบทดสอบวินิจฉัยฉบับที่ 2 มีค่าความยากตั้งแต่ .20-.63 ค่าอำนาจจำแนกตั้งแต่ .30-.82 ค่าความเชื่อมั่นเท่ากับ .900 แบบทดสอบวินิจฉัยฉบับที่ 3 มีค่าความยากตั้งแต่ .38-.75 ค่าอำนาจจำแนกตั้งแต่ .35-.80 ค่าความเชื่อมั่นเท่ากับ .915 ค่าความเที่ยงตรงเชิงเนื้อหาให้โดยผู้เชี่ยวชาญทางด้านเนื้อหาเป็นผู้พิจารณาตรวจสอบปรากฏว่า แบบทดสอบวินิจฉัยทั้งสามฉบับสามารถวัดเรื่องทักษะพื้นฐานของความเข้าใจในการอ่านภาษาอังกฤษได้จริง

พรรณพกา มณีวัง (2542 : บทคัดย่อ) ศึกษาเรื่องการใช้แบบทดสอบวินิจฉัยวิชาภาษาอังกฤษชั้นประถมศึกษาปีที่ 6 เพื่อศึกษาข้อบกพร่องในการใช้ภาษาอังกฤษของนักเรียนที่ได้รับการทดสอบโดยแบบทดสอบวินิจฉัย และเพื่อสร้างแบบทดสอบวินิจฉัยวิชาภาษาอังกฤษสำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 โดยใช้กลุ่มตัวอย่างที่เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 จากโรงเรียนในอำเภอค้อยสะแก จังหวัดเชียงใหม่ จำนวน 346 คน พบว่า การศึกษาข้อบกพร่องของนักเรียนเกี่ยวกับการใช้ภาษาอังกฤษได้แก่ นักเรียนขาดทักษะในการหาตัวชี้แนะในประโยคร้อยละ 26 และนักเรียนตอบคำถามนำนามคิดโดยเดาคความหมายของคำรอบข้างมาแปลแทนร้อยละ 23 ความยากของแบบทดสอบมีค่าตั้งแต่ .65 ถึง .86 ค่าอำนาจจำแนกของข้อสอบมีค่าตั้งแต่ .16 ถึง .65

ความเชื่อมั่นของแบบทดสอบซึ่งคำนวณจากสูตร K-R 20 แล้วปรับแก้ให้ถูกต้องโดยใช้สูตร ลิฟวิงตัน มีค่า .889 และแบบทดสอบมีค่าความเที่ยงเชิงเนื้อหาโดยผู้เชี่ยวชาญเป็นผู้พิจารณา

สมจิตร หามาลี (2540 :ออนไลน์) ได้สร้างแบบทดสอบวินิจฉัยความเข้าใจในการอ่าน ภาษาอังกฤษ ระดับชั้นมัธยมศึกษาปีที่ 3 โดยมีจุดมุ่งหมายเพื่อสร้างแบบทดสอบวินิจฉัยความเข้าใจ ในการอ่านภาษาอังกฤษ และหาสาเหตุของความบกพร่องความเข้าใจในการอ่านภาษาอังกฤษ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า มี 2 กลุ่ม กลุ่มแรกเป็น นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 810 คน ใช้ในการดำเนินการสร้างข้อสอบวินิจฉัย อีกกลุ่ม เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 900 คน ใช้เพื่อตรวจสอบความบกพร่องทางการอ่าน ภาษาอังกฤษ การสร้างแบบทดสอบเริ่มจากการให้เติมคำ แล้วนำข้อบกพร่องในการตอบผิดของนัก เรียนมาเป็นตัวลง สำหรับแบบทดสอบวินิจฉัยแบบเลือกตอบ 4 ตัวเลือก แล้วนำไปทดสอบ 3 ครั้ง เพื่อพัฒนาคุณภาพของแบบทดสอบแต่ละฉบับ และครั้งที่ 4 เพื่อวิเคราะห์ความบกพร่องของกลุ่ม ตัวอย่าง ผลการศึกษาพบว่า แบบทดสอบวินิจฉัยทั้ง 5 ฉบับ มีค่าความยากของข้อสอบตั้งแต่ .65- .93 ค่าอำนาจจำแนกของข้อสอบตั้งแต่ .02- .64 ค่าความเชื่อมั่นของแบบทดสอบวินิจฉัยทั้ง 5 ฉบับ ซึ่งคำนวณโดยใช้สูตร KR-20 และปรับแก้โดยวิธีการค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์มีค่า ตั้งแต่ .6535 - .8459 สำหรับความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบให้ผู้เชี่ยวชาญพิจารณา ตรวจสอบ ปรากฏว่าแบบทดสอบทั้ง 5 ฉบับ มีค่าสอดคล้องกับจุดประสงค์การเรียนรู้ ความเข้าใจ ในการอ่านภาษาอังกฤษอยู่ในเกณฑ์ดี ข้อบกพร่องของกลุ่มตัวอย่างเรียงลำดับจากมากไปหาน้อย ดังนี้ คำศัพท์ ส่วนประกอบของประโยค คำคุณศัพท์ คำกริยา คำสรรพนาม

นิตยา พลดงนอก (2540 : ออนไลน์) ได้พัฒนาแบบวัดความเข้าใจในการอ่านภาษาไทย ของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยมีวัตถุประสงค์เพื่อสร้างและพัฒนาแบบวัดความเข้าใจ ในการอ่านภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยอิงทฤษฎีการตอบสนองข้อสอบ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2539 จำนวน 1,831 คน ผู้วิจัยได้สร้าง แบบสอบเลือกตอบจำนวน 43 ข้อ ขึ้นทดลองใช้และวิเคราะห์ข้อมูล ปรากฏผลดังนี้ แบบวัดมีค่า อำนาจจำแนกอยู่ระหว่าง 0.23 ถึง 2.26 ค่าความยากอยู่ในช่วง -2.00 ถึง 3.45 และค่าโอกาสการ เคาข้อสอบถูกอยู่ในช่วงระหว่าง 0.07 ถึง 0.27 สำหรับค่าฟังก์ชันสารสนเทศของแบบวัดตามทฤษฎี การตอบสนองข้อสอบ พบว่า แบบวัดจะมีประสิทธิภาพสูงสุดเมื่อใช้ทดสอบกับผู้สอบที่มีความ สามารถระดับ 0.60

สกาวิ สันติเทวกุล (2540 : บทคัดย่อ) ได้สร้างแบบทดสอบวินิจฉัยวิชาฟิสิกส์ ว 021 ใช้ทฤษฎีการตอบสนองข้อสอบ กลุ่มตัวอย่างที่ใช้ในการพัฒนาและหาคุณภาพเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ภาคเรียนที่ 2 ปีการศึกษา 2539 แผนการเรียนวิทยาศาสตร์และคณิตศาสตร์ของโรงเรียนในสังกัดกรมสามัญศึกษาเขตการศึกษา 2 และ 3 จำนวน 1,761 คน ที่แบ่งออกเป็น 2 กลุ่ม โดยกลุ่มที่ 1 จำนวน 994 คน ใช้สำหรับพัฒนาคัดเลือกกลุ่มข้อสอบ กลุ่มตัวอย่างนี้ได้มาโดยวิธีการสุ่มแบบกลุ่ม ส่วนกลุ่มที่ 2 จำนวน 767 คน ใช้สำหรับสร้างเกณฑ์ปกติ กลุ่มตัวอย่างนี้ได้มาโดยสุ่มแบบแบ่งชั้น พบว่ามีจำนวนข้อสอบ จำนวน 81 ข้อที่มีคุณภาพตามเกณฑ์ที่กำหนดไว้โดยมีค่าอำนาจจำแนกตั้งแต่ 0.300 ถึง 2.070 ค่าความยากตั้งแต่ -2.750 ถึง 3.000 และค่าการเดาตั้งแต่ 0.030 ถึง 0.300 และการจัดเป็นแบบทดสอบวินิจฉัยที่หาคะแนนจุดตัดตามจุดประสงค์ โดยวิธีโลจิสติก และหาความเที่ยงจากฟังก์ชันสารสนเทศของแบบทดสอบ โดยมีค่าสูงสุดเท่ากับ 0.9526 ที่ระดับความสามารถ 0.1 ซึ่งแสดงว่าเหมาะสมกับผู้สอบที่มีความสามารถปานกลาง

อำไพ วรรณเกษม (2539 : ออนไลน์) ศึกษาเรื่องการสร้างแบบทดสอบเฉพาะกลุ่มแบบซีเอที มีวัตถุประสงค์เพื่อสร้างแบบทดสอบเฉพาะบุคคล แบบ ซีเอที รูปแบบแยกทางแปรผัน โดยใช้กลยุทธ์แบบเบย์ และหาคุณภาพของแบบทดสอบที่สร้างขึ้น กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 จำนวน 2,810 คน วิธีการศึกษา ประกอบด้วย การทดสอบครั้งที่ 1 เป็นการทดลองกลุ่มข้อสอบ เพื่อนำผลมาปรับปรุงคุณภาพของข้อสอบ การทดสอบครั้งที่ 2 เพื่อประมาณค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบรูปแบบ 3 พารามิเตอร์ การทดสอบครั้งที่ 3 เป็นการทดลองการใช้แบบทดสอบเฉพาะบุคคลโดยใช้เครื่องคอมพิวเตอร์ในการทดสอบ และการทดสอบครั้งที่ 4 เป็นการทดสอบเพื่อประมาณค่าความสามารถของผู้สอบโดยวิธีเบเซียนอัปเดต

เพ็ญพรรณ กริ่งไกร (2538 : ออนไลน์) ศึกษาเรื่องการพัฒนาแบบสอบความคิดรวบยอดทางคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 2 มีวัตถุประสงค์เพื่อสร้างและวิเคราะห์แบบสอบความคิดรวบยอดทางคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 2 โดยอิงทฤษฎีการตอบสนองข้อสอบ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 2 จำนวน 793 คน แบบสอบเลือกตอบจำนวน 53 ข้อ ได้ถูกสร้างขึ้น นำไปทดลองใช้และวิเคราะห์ปรากฏผลดังนี้ แบบสอบมีค่าอำนาจจำแนกอยู่ระหว่าง 0.30 ถึง 0.79 ค่าความยากอยู่ในช่วงระหว่าง -3.93 ถึง 4.00 และค่าโอกาสการเดาข้อสอบถูกอยู่ในช่วงระหว่าง 0.11 ถึง 0.31 สำหรับความเที่ยงตามแนวทฤษฎีการ

ตอบสนองข้อสอบ พบว่าแบบสอบจะมีประสิทธิภาพสูงสุดเมื่อใช้วัดกับผู้สอบที่มีความสามารถระดับ 3.00

ศรินทิพย์ แก้วมawangศ์ (2532 : ออนไลน์) ศึกษาเรื่องการใช้โมเดลโลจิสติกเพื่อพัฒนาแบบสอบอิงโดเมน มีวัตถุประสงค์เพื่อสร้างแบบสอบอิงโดเมนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 3 เรื่อง พหุนามและใช้โมเดลโลจิสติก 3 พารามิเตอร์ในการวิเคราะห์แบบสอบที่สร้างขึ้นเพื่อหาค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก ค่าความยาก และค่าการเดา และคัดเลือกข้อสอบตามเกณฑ์ที่เหมาะสมของค่าพารามิเตอร์ พร้อมทั้งตรวจสอบคุณภาพของแบบสอบตามแนวอิงโดเมนและทฤษฎี IRT นอกจากนี้ยังมุ่งใช้โมเดลโลจิสติก 3 พารามิเตอร์ เพื่อกำหนดคะแนนเกณฑ์ของแบบสอบ ตลอดจนเพื่อการประมาณค่าความสามารถและคะแนนโดเมน สำหรับใช้เป็นตารางเทียบคะแนนของผู้สอบด้วย แบบสอบอิงโดเมนที่ได้เป็นชนิดเลือกตอบ 4 ตัวเลือก จำนวน 50 ข้อ โดยแบ่งเป็น 4 โดเมน ตามพฤติกรรมในขอบเขตเนื้อหาเรื่องพหุนาม ข้อสอบเขียนขึ้นตามลักษณะเฉพาะของข้อสอบ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2531 จำนวน 2,145 คน ซึ่งจากการวิจัยครั้งนี้ได้แบบสอบอิงโดเมนที่มีคุณภาพดังนี้

1. มีคุณสมบัติการวัดเพียงคุณลักษณะเดียวและความเป็นอิสระในการตอบข้อสอบ
2. ค่าพารามิเตอร์ของข้อสอบอยู่ในเกณฑ์เหมาะสม โดยมีค่าอำนาจจำแนกตั้งแต่ 0.82293 ถึง 2.00 มีค่าเฉลี่ยของค่าอำนาจจำแนกเป็น 1.27235 มีค่าความยากตั้งแต่ -0.779607 ถึง 2.32667 ค่าเฉลี่ยของความยากเป็น 0.79895 และมีค่าการเดาตั้งแต่ 0.03228 ถึง 0.29393 ค่าเฉลี่ยของการเดาเป็น 0.18271
3. มีความตรงเชิงเนื้อหา โดยมีค่าเฉลี่ยของคะแนนการตัดสินใจสอดคล้องผ่านเกณฑ์ที่ยอมรับคือ 0.5 มีความตรงเชิงโครงสร้าง คือ พบว่าข้อสอบในโดเมนเดียวกันถูกจัดไว้ในตัวประกอบเดียวกันเป็นส่วนใหญ่ และมีความตรงร่วมสมัย คือ มีค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันเป็น 0.643 ที่ระดับนัยสำคัญ .001
4. มีความเที่ยงตามแนวอิงโดเมนที่ 1 ถึง 4 และทั้งฉบับเป็น 0.8884 0.8485 0.8443

0.7700 และ 0.9428ตามลำดับ สำหรับความเที่ยงตามแนว IRT พบว่า ได้ค่าฟังก์ชันสารสนเทศแบบสอบสูงมากในช่วงของค่า θ ระหว่าง 0.00 ถึง 1.8 ซึ่งแสดงว่า แบบสอบมีความเหมาะสมหรือมีประสิทธิภาพสูงสุดเมื่อใช้ทดสอบกับนักเรียนที่มีระดับความสามารถปานกลางและค่อนข้างสูง

5. ได้คะแนนเกณฑ์การตัดสินความสามารถขั้นต่ำที่ยอมรับได้ของแบบสอบใน โดเมนที่ 1-4 และทั้งฉบับ ซึ่งอยู่บนสเกลความสามารถเป็น 1.48 0.58 0.37 1.04 และ 0.93 ตามลำดับ

ประดิษฐ์ เรื่องตระกูล (2529 : บทคัดย่อ) ได้ใช้โมเดลโลจิสติกพัฒนาแบบทดสอบ วินิจฉัยวิชาคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2527 จำนวน 1,913 คน แบบทดสอบกำหนดมวลประสพการณ์ 6 มวล ประสพการณ์ ซึ่งแยกย่อยได้เป็น 13 จุดประสงค์การเรียนรู้ แบบทดสอบวินิจฉัยเป็นแบบ ทดสอบประเภทเลือกตอบ 4 ตัวเลือก จำนวน 60 ข้อ ผลการศึกษาพบว่า แบบทดสอบวินิจฉัยมี คุณสมบัติการวัดเพียงคุณลักษณะเดียว มีค่าอำนาจจำแนก ค่าความยากและค่าการเดา เท่ากับ 0.19036 ถึง 2.00000 , -3.75288 ถึง 1.90366 และ 0.0 ถึง 0.5 ตามลำดับ มีค่าความเที่ยง ตามแนวอิงเกณฑ์ของโดเมนย่อยที่ 1 ถึงโดเมนย่อยที่ 6 และทั้งฉบับเป็น .8352 , .7353 , .6483 , .8729 , .7655 , .6055 และ .9126 ตามลำดับ สำหรับความเที่ยงตามแนวทฤษฎี IRT พบว่า แบบทดสอบจะมีประสิทธิภาพสูงสุดเมื่อวัดกับผู้สอบที่มีระดับความสามารถปานกลาง คือ มีค่า 0 อยู่ในช่วง -1.5 ถึง 1.5 มีค่าความตรงเชิงเนื้อหาและความตรงเชิงวินิจฉัย โดยมีค่าเฉลี่ยผ่าน เกณฑ์ที่ยอมรับคือ 2.5 มีความตรงร่วมสมัยและความตรงเชิงพยากรณ์ คือมีค่าสัมประสิทธิ์สห สัมพันธ์เป็น .578 และ .665 ที่ระดับนัยสำคัญ .001 ตามลำดับ และมีความตรงเชิงทฤษฎี

เดือนเพ็ญ หว่านณรงค์ (2528 : ออนไลน์) ได้สร้างแบบทดสอบวินิจฉัยทักษะพื้นฐาน ของความเข้าใจในการอ่านภาษาอังกฤษ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีจุดมุ่งหมายเพื่อ สร้างแบบทดสอบวินิจฉัยทักษะพื้นฐานของความเข้าใจในการอ่านภาษาอังกฤษที่มีคุณภาพ จำนวน 7 ฉบับ คือ แบบทดสอบฉบับคำกริยา แบบทดสอบฉบับคำบุพบท แบบทดสอบฉบับคำสรรพนาม และคำที่แสดงความเป็นเจ้าของ แบบทดสอบฉบับคำคุณศัพท์และคำวิเศษณ์ แบบทดสอบฉบับ คำสันธาน แบบทดสอบฉบับการตั้งคำถาม และแบบทดสอบฉบับคำศัพท์ เพื่อใช้ค้นหาจุดบกพร่อง และสาเหตุของความบกพร่องในทักษะพื้นฐานของความเข้าใจในการอ่านภาษาอังกฤษ สำหรับ นักเรียนที่สำเร็จชั้นมัธยมศึกษาปีที่ 2 และจะเริ่มเรียนทักษะความเข้าใจในการอ่านภาษาอังกฤษใน ชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2538 ของโรงเรียนสังกัดกรมสามัญศึกษา ในจังหวัด สมุทรปราการ จำนวน 777 คน ซึ่งเลือกมาโดยการสุ่มแบบแบ่งชั้น การสร้างแบบทดสอบวินิจฉัยมี วิธีดำเนินการดังนี้ สร้างแบบทดสอบสำรวจแบบเติมคำโดยเขียนข้อสอบให้สอดคล้องกับ

จุดประสงค์เชิงพฤติกรรม เพื่อสำรวจหาข้อบกพร่องในการตอบผิดของนักเรียน นำข้อบกพร่องนั้น มากำหนดเป็นลักษณะตัวเลือกในการเขียนลักษณะเฉพาะของข้อสอบ สร้างแบบทดสอบวินิจฉัย ตามลักษณะเฉพาะของข้อสอบ ซึ่งเป็นแบบทดสอบแบบเลือกตอบ ชนิด 5 ตัวเลือก ผลการศึกษา พบว่า แบบทดสอบวินิจฉัยทั้ง 7 ฉบับ มีค่าความยากของข้อสอบตั้งแต่ .50 ถึง .95 ค่าอำนาจจำแนกของข้อสอบตั้งแต่ .09 ถึง 1.00 ค่าความเชื่อมั่นของแบบทดสอบมีค่าเรียงลำดับดังนี้ .9443 .5368 .8596 .7891 .7984 .8080 และ .7918 สำหรับความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบหาโดยให้ผู้เชี่ยวชาญทางเนื้อหาพิจารณาความสอดคล้องระหว่างข้อสอบรายข้อกับจุดประสงค์เชิงพฤติกรรม และนำผลการพิจารณามาคำนวณหาค่าดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิง พฤติกรรม ซึ่งมีค่าตั้งแต่ .50 ถึง 1.00 แสดงว่าแบบทดสอบทั้ง 7 ฉบับมีความเที่ยงตรงเชิงเนื้อหาจริง

จากงานวิจัยที่เกี่ยวข้องกับแบบทดสอบวินิจฉัยและทฤษฎีการตอบสนองข้อสอบดังกล่าวทำให้ได้แบบทดสอบวินิจฉัยที่มีคุณภาพ สำหรับวินิจฉัยการเรียนของนักเรียนในหลายรายวิชา ทั้งระดับประถมศึกษาและระดับมัธยมศึกษา แต่รายวิชาภาษาอังกฤษระดับชั้นประถมศึกษาปีที่ 4 ซึ่งเป็นภาษาอังกฤษระดับอ่านออกเขียนได้ยังไม่มีการสร้างเป็นแบบทดสอบวินิจฉัย ผู้วิจัยจึงสนใจที่จะสร้างแบบทดสอบวินิจฉัยรายวิชานี้ เพื่อนำแบบทดสอบไปใช้ประโยชน์ต่อไป