

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานที่เกี่ยวข้องกับการวิจัยในการสร้างแบบทดสอบอิงเกณฑ์ กลุ่มสาระการเรียนรู้ภาษาไทย เรื่อง ชนิดของคำ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ครั้งนี้ผู้วิจัยจะได้นำเสนอ ในหัวข้อต่อไปนี้

1. หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 กลุ่มสาระการเรียนรู้ภาษาไทย
 - 1.1 สาระและมาตรฐานการเรียนรู้
 - 1.2 คุณภาพผู้เรียน
 - 1.3 ตัวชี้วัดและสาระการเรียนรู้แกนกลาง สาระที่ 4 หลักการใช้ภาษาไทย
 - 1.4 การวัดและประเมินผลการเรียนรู้
2. การทดสอบแบบอิงเกณฑ์
 - 2.1 จุดมุ่งหมายของการทดสอบอิงเกณฑ์
 - 2.2 ความสำคัญของการทดสอบอิงเกณฑ์
3. แบบทดสอบอิงเกณฑ์
 - 3.1 ความหมายของแบบทดสอบอิงเกณฑ์
 - 3.2 ความเป็นมาของแบบทดสอบอิงเกณฑ์
 - 3.3 ลักษณะของแบบทดสอบอิงเกณฑ์
 - 3.4 หลักการสร้างแบบทดสอบอิงเกณฑ์
 - 3.5 การหาคุณภาพของแบบทดสอบอิงเกณฑ์
 - 3.6 การกำหนดคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์
4. การสร้างข้อสอบแบบเลือกตอบและรูปแบบของข้อสอบแบบเลือกตอบ
5. งานวิจัยที่เกี่ยวข้อง
 - 5.1 งานวิจัยในประเทศ
 - 5.2 งานวิจัยต่างประเทศ

หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 กลุ่มสาระการเรียนรู้ภาษาไทย

การจัดการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ. 2552 : 1 - 4) ได้กล่าวถึงการจัดการเรียนรู้กลุ่มสาระการเรียนรู้ภาษาไทยไว้ ดังนี้

ภาษาไทยเป็นเอกลักษณ์ของชาติ เป็นสมบัติทางวัฒนธรรมอันก่อให้เกิดความเป็นเอกภาพและเสริมสร้างบุคลิกภาพของคนในชาติให้มีความเป็นไทย เป็นเครื่องมือในการติดต่อสื่อสารเพื่อสร้างความเข้าใจและความสัมพันธ์ที่ดีต่อกัน ทำให้สามารถประกอบกิจธุระ การงาน และดำรงชีวิตร่วมกันในสังคมประชาธิปไตยได้อย่างสันติสุข และเป็นเครื่องมือในการแสวงหาความรู้ประสบการณ์จากแหล่งข้อมูลสารสนเทศต่าง ๆ เพื่อพัฒนาความรู้ กระบวนการคิดวิเคราะห์ วิจัยและสร้างสรรค์ให้ทันต่อการเปลี่ยนแปลงทางสังคม และ ความก้าวหน้าทางวิทยาศาสตร์เทคโนโลยี ตลอดจนนำไปใช้ในการพัฒนาอาชีพให้มีความ มั่นคงทางเศรษฐกิจ นอกจากนี้ยังเป็นสื่อแสดงภูมิปัญญาของบรรพบุรุษด้านวัฒนธรรม ประเพณี สุนทรียภาพ เป็นสมบัติล้ำค่าควรแก่การเรียนรู้ อนุรักษ์ และสืบสานให้คงอยู่ คู่ชาติ ไทยตลอดไป

หลักสูตรกลุ่มสาระการเรียนรู้ภาษาไทยเป็นหลักสูตรที่เน้นผู้เรียนเป็นสำคัญ ผู้เรียน ได้เรียนรู้จากประสบการณ์จริง ให้ทำได้ ทำเป็น รักการอ่านและใฝ่เรียนรู้ สามารถเรียนรู้ได้ อย่างต่อเนื่องตลอดชีวิต ลักษณะของหลักสูตรเป็นหลักสูตรสมรรถฐาน (Standard-Based Curriculum) ที่เน้นด้านความรู้ ทักษะและกระบวนการ ด้านคุณธรรม จริยธรรม ค่านิยม และ ความสามารถของผู้เรียนมากกว่าการเน้นเนื้อหา กลุ่มสาระการเรียนรู้ภาษาไทยจึงเป็นกลุ่ม สาระที่สำคัญสำหรับเป็นหลักในการจัดการเรียนการสอนตามหลักสูตรการศึกษาขั้นพื้นฐาน เพื่อสร้างพื้นฐานความคิด เพราะภาษาเป็นสื่อของความคิด พัฒนาสติปัญญา กระบวนการคิด ความคิดสร้างสรรค์ คิววิพากษ์วิจารณ์ คิววิเคราะห์ คิดตัดสินใจแก้ปัญหาและวินิจฉัยอย่างมี เหตุผล เป็นเครื่องมือสำคัญในการเรียนรู้ และศึกษากลุ่มสาระการเรียนรู้ต่าง ๆ เป็นวิชาทักษะ ที่ต้องฝึกฝนจนเกิดความชำนาญในการใช้ภาษาเพื่อการสื่อสารสาระ เป็นแก่นความรู้ทางภาษา ที่ครูผู้สอนต้องจัดให้เหมาะสมกับผู้เรียนและสภาพแวดล้อมในท้องถิ่น ภาษาไทยเป็นทักษะที่ ต้องฝึกฝนจนเกิดความชำนาญในการใช้ภาษา เพื่อการสื่อสารการเรียนรู้อย่างมีประสิทธิภาพ และเพื่อนำไปใช้ในชีวิตจริง

1. สาระและมาตรฐานการเรียนรู้

สาระที่ 1 การอ่าน

มาตรฐาน ท 1.1 ใช้กระบวนการอ่านสร้างความรู้และความคิดเพื่อนำไปใช้ตัดสินใจ แก้ปัญหาในการดำเนินชีวิต และมีนิสัยรักการอ่าน

สาระที่ 2 การเขียน

มาตรฐาน ท 2.1 ใช้กระบวนการเขียนสื่อสาร เขียนเรียงความ ย่อความ และเขียนเรื่องราวในรูปแบบต่าง ๆ เขียนรายงานข้อมูลสารสนเทศและรายงานการศึกษาค้นคว้าอย่างมีประสิทธิภาพ

สาระที่ 3 การฟัง การดู การพูด

มาตรฐาน ท 3.1 สามารถเลือกฟังและดูอย่างมีวิจารณญาณ และพูดแสดงความรู้ ความคิด และความรู้สึกในโอกาสต่าง ๆ อย่างมีวิจารณญาณและสร้างสรรค์

สาระที่ 4 หลักการใช้ภาษาไทย

มาตรฐาน ท 4.1 เข้าใจธรรมชาติของภาษาและหลักภาษาไทยการเปลี่ยนแปลงของภาษา และพลังของภาษา ภูมิปัญญาทางภาษา และรักษาภาษาไทยไว้เป็นสมบัติของชาติ

สาระที่ 5 วรรณคดีและวรรณกรรม

มาตรฐาน ท 5.1 เข้าใจและแสดงความคิดเห็น วิเคราะห์วรรณคดีและวรรณกรรม ไทยอย่างเห็นคุณค่าและนำมาประยุกต์ใช้ในชีวิตจริง

2. คุณภาพผู้เรียน

เมื่อจบชั้นประถมศึกษาปีที่ 6 แล้วนักเรียนต้องมีคุณภาพในแต่ละสาระ ดังนี้

2.1 อ่านออกเสียงบทร้อยแก้วและบทร้อยกรองเป็นทำนองเสนาะได้ ถูกต้อง อธิบายความหมายโดยตรงและความหมายโดยนัยของคำ ประโยค ข้อความ สำนวนโวหารจากเรื่องที่อ่าน เข้าใจคำแนะนำ คำอธิบายในคู่มือต่าง ๆ แยกแยะข้อคิดเห็นและข้อเท็จจริง รวมทั้งจับใจความสำคัญของเรื่องที่อ่าน และนำความรู้ความคิดจากเรื่องที่อ่านไปตัดสินใจแก้ปัญหาในการดำเนินชีวิตได้ มีมารยาทและมีนิสัยรักการอ่าน และเห็นคุณค่าสิ่งที่อ่าน

2.2 มีทักษะในการคัดลายมือตัวบรรจงเต็มบรรทัดและครึ่งบรรทัด เขียนสะกดคำ แต่งประโยค และเขียนข้อความ ตลอดจนเขียนสื่อสาร โดยใช้ถ้อยคำชัดเจนเหมาะสม

ใช้แผนภาพ โครงเรื่อง และแผนภาพความคิดเพื่อพัฒนางานเขียนเขียนเรียงความ ข้อความ
จดหมายส่วนตัว กรอกแบบรายการต่าง ๆ เขียนแสดงความรู้สึกรู้สึกและความคิดเห็น
เขียนเรื่องตามจินตนาการอย่างสร้างสรรค์ และมีมารยาทในการเขียน

2.3 พูดยุคแสดงความรู้ ความคิดเกี่ยวกับเรื่องที่ฟังและดู เล่าเรื่องย่อหรือสรุป
จากเรื่องที่ฟังและดู ตั้งคำถาม ตอบคำถามจากเรื่องที่ฟังและดู รวมทั้งประเมินความน่า
เชื่อถือจากการฟังและดู โฆษณาอย่างมีเหตุผล พูดยุคตามลำดับขั้นตอนเรื่องต่าง ๆ อย่าง
ชัดเจนพูดรายงานหรือ ประเด็นค้นคว้าจากการฟัง การดู การสนทนา และพูดยุคโน้มน้าวได้
อย่างมีเหตุผลรวมทั้งมีมารยาทในการดูและพูด

2.4 สะกดคำและเข้าใจความหมายของคำ จำนวน คำพ้องเพยและสุภาษิต
รู้และเข้าใจชนิดและหน้าที่ของคำในประโยค ชนิดของประโยค และคำภาษาต่างประเทศ
ในภาษาไทย ใช้คำราชาศัพท์และคำสุภาพได้อย่างเหมาะสม แต่งประโยค แต่งบท
ร้อยกรองประเภทกลอนสี่ กลอนสุภาพ และกาพย์ยานี

2.5 เข้าใจและเห็นคุณค่าวรรณคดีและวรรณกรรมที่อ่าน เล่านิทานพื้นบ้าน
ร้องเพลงพื้นบ้านของท้องถิ่น นำข้อคิดเห็นจากเรื่องที่อ่านไปประยุกต์ใช้ในชีวิตจริง และ
ท่องจำบท อาขยานตามที่กำหนดได้

3. ตัวชี้วัดและสาระการเรียนรู้แกนกลาง สาระที่ 4 หลักการใช้ภาษาไทย

มาตรฐาน ท 4.1 เข้าใจธรรมชาติของภาษาและหลักภาษาไทย การเปลี่ยนแปลงของ
ภาษาและพลังของภาษา ภูมิปัญญาทางภาษา และรักษาภาษาไทยไว้เป็นสมบัติของชาติ

ตารางที่ 1 ตัวชี้วัดและสาระการเรียนรู้แกนกลาง สาระที่ 4 หลักการใช้ภาษาไทย

ชั้น	ตัวชี้วัด	สาระการเรียนรู้แกนกลาง
ป.6	วิเคราะห์ชนิดและหน้าที่ของคำใน ประโยค	ชนิดของคำ - คำนาม - คำสรรพนาม - คำกริยา - คำวิเศษณ์ - คำบุพบท - คำเชื่อม - คำอุทาน

ชั้น	ตัวชี้วัด	สาระการเรียนรู้แกนกลาง
ป.6	ใช้คำได้เหมาะสมกับกาลเทศะและบุคคล	คำราชาศัพท์ ระดับภาษา ภาษาถิ่น
	3. รวบรวมและบอกความหมายของคำภาษาต่างประเทศที่ใช้ในภาษาไทย	คำที่มาจากภาษาต่างประเทศ
	4. ระบุลักษณะของประโยค	กลุ่มคำหรือวลี ประโยคสามัญ ประโยครวม ประโยคซ้อน
	5. แต่งบทหรือยกรอง	กลอนสุภาพ
	วิเคราะห์และเปรียบเทียบสำนวนที่ เป็นคำพังเพยและสุภาษิต	สำนวนที่เป็นคำพังเพย และสุภาษิต

4. การวัดและประเมินผลการเรียนรู้

การวัดและการประเมินผลการเรียนภาษาไทยเป็นงานที่ยากซึ่งต้องทำความเข้าใจที่ถูกต้องเกี่ยวกับการพัฒนาทางภาษา ดังนั้นผู้ปฏิบัติหน้าที่วัดผลการเรียนรู้ด้านภาษาจำเป็นต้องเข้าใจหลักการของการเรียนรู้ภาษา เพื่อเป็นพื้นฐานการดำเนินงาน ดังนี้

4.1 ทักษะทางภาษาทั้งการฟัง การดู การพูด การอ่าน และการเขียนมีความสำคัญเท่า ๆ กัน ทักษะเหล่านี้จะบูรณาการกัน ในการเรียนการสอนจะไม่แยกฝึกทักษะทีละอย่างจะต้องฝึกทักษะไปพร้อม ๆ กัน และทักษะทางภาษาทักษะหนึ่งจะส่งผลต่อการพัฒนาทักษะทางภาษาอื่น ๆ ด้วย

4.2 ผู้เรียนต้องได้รับการพัฒนาความสามารถทางภาษาพร้อมกับการพัฒนาความคิดเพราะภาษาเป็นสื่อของความคิด ผู้ที่มีทักษะความสามารถในการใช้ภาษา มีประมวลคำมากจะช่วยให้ผู้เรียนมีความสามารถในการคิดด้วย ขณะเดียวกันการเรียนภาษาจะเรียนร่วมกับผู้อื่น มีการติดต่อสื่อสาร ใช้ภาษาในการติดต่อกับเพื่อนกับครูจึงเป็นการฝึกทักษะทางสังคมด้วยเมื่อผู้เรียนได้ใช้ภาษาในสถานการณ์จริงทั้งในบริบททางวิชาการในห้องเรียนและในชุมชนจะทำให้ผู้เรียนได้ใช้ภาษาและได้ฝึกทักษะทางสังคมในสถานการณ์จริง

4.3 ผู้เรียนต้องเรียนรู้การใช้ภาษาพูดและภาษาเขียนอย่างถูกต้อง ด้วยการฝึกการใช้ภาษา มิใช่เรียนรู้กฎเกณฑ์ทางภาษาแต่เพียงอย่างเดียว การเรียนภาษาจะต้องเรียนรู้

ไวยากรณ์ หรือหลักภาษาการสะกดคำ การใช้เครื่องหมายวรรคตอน และนำความรู้ดังกล่าวไปใช้ในการฝึกฝนการเขียนและพัฒนาทักษะทางภาษาของตน

4.4 ผู้เรียนทุกคนจะได้รับการพัฒนาทักษะทางภาษาเท่ากัน แต่การพัฒนาทางภาษา จะไม่เท่ากัน และวิธีการเรียนรู้จะต่างกัน

4.5 ภาษากับวัฒนธรรมมีความสัมพันธ์กันอย่างใกล้ชิด หลักสูตรจะต้องให้ความสำคัญและใช้ความเคารพและเห็นคุณค่าของเชื้อชาติ จัดกิจกรรมภูมิหลังของภาษาและการใช้ภาษาถิ่นของผู้เรียน และช่วยให้ผู้เรียนพัฒนาภาษาไทยของตน และพัฒนาความรู้ลึกที่เกี่ยวกับภาษาไทยและกระตุ้นให้ผู้เรียนสามารถเรียนภาษาไทยด้วยความสุข

4.6 ภาษาไทยเป็นเครื่องมือของการเรียนรู้ และทุกกลุ่มสาระการเรียนรู้จะต้องใช้ภาษาไทยเป็นเครื่องมือการสื่อสารและการแสวงหาความรู้ การเรียนทุกกลุ่มสาระการเรียนรู้จะใช้ภาษาในการคิดวิเคราะห์ การคิดสร้างสรรค์ การอภิปราย การเขียนรายงาน การเขียนโครงการ การตอบคำถาม การตอบข้อทดสอบ ดังนั้นครูทุกคนไม่ว่าจะสอนวิชาใดก็ตามจะต้องใช้ภาษาที่เป็นแบบแผนเป็นตัวอย่างที่ดีแก่นักเรียน และต้องสอนการใช้ภาษาแก่ผู้เรียนด้วยเสมอ (สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ. 2552 : 30)

วิธีการวัดผลที่สามารถนำมาประเมินผลการเรียนรู้ เพื่อนำผลมาปรับปรุงพัฒนาการเรียนรู้ของผู้เรียนมีหลากหลาย ดังต่อไปนี้

1. การให้ตอบแบบทดสอบ ทั้งในลักษณะที่เป็นแบบเลือกคำตอบได้แก่ ข้อสอบแบบเลือกตอบ ถูก - ผิด จับคู่ และข้อสอบชนิดให้ผู้สอบสร้างคำตอบ ได้แก่ เติมข้อความในช่องว่าง คำตอบสั้น ๆ เป็นประโยค เป็นข้อความ แผนภูมิ
2. การดูจากผลงาน เช่น เรียงความ รายงานการวิจัย บันทึกประจำวัน รายงานการทดลอง บทละคร บทร้อยกรอง แฟ้มสะสมงาน เป็นต้น
3. การปฏิบัติ โดยผู้สอนสามารถสังเกตการณ์นำทักษะและความรู้ไปใช้โดยตรง ในสถานการณ์ที่ให้ปฏิบัติจริง วิธีนี้ถูกนำไปใช้อย่างกว้างขวางในการประเมินการปฏิบัติที่มีระเบียบข้อบังคับ เช่น คนตรี ร้องเพลง พลศึกษา การโต้วาที การกล่าวสุนทรพจน์ ละครเวที
4. การประเมินการ วิธีนี้จะให้ข้อมูลเกี่ยวกับวิธีการเรียนรู้ กระบวนการคิดของผู้เรียนมากกว่าดูผลงานหรือการปฏิบัติ ซึ่งจะทำให้ทราบกระบวนการคิดของผู้เรียนได้ โดยครู

เป็นผู้สังเกตวิธีการคิดของผู้เรียน วิธีเหมาะในการประเมินด้านคุณธรรม จริยธรรม และลักษณะนิสัย (กรมวิชาการ. 2545 : 172-176)

การวัดและประเมินผลการเรียนรู้ของผู้เรียนตั้งอยู่บนหลักการพื้นฐานสองประการคือ การประเมินเพื่อพัฒนาผู้เรียนและเพื่อตัดสินผลการเรียน ในการพัฒนาคุณภาพการเรียนรู้ของผู้เรียนให้ประสบผลสำเร็จนั้น ผู้เรียนจะต้องได้รับการพัฒนาและประเมินตามตัวชี้วัดเพื่อให้บรรลุตามมาตรฐานการเรียนรู้ สะท้อนสมรรถนะสำคัญ และคุณลักษณะอันพึงประสงค์ของผู้เรียนซึ่งเป็นเป้าหมายหลักในการวัดและประเมินผลการเรียนรู้ในทุกระดับไม่ว่าจะเป็นระดับชั้นเรียน ระดับสถานศึกษา ระดับเขตพื้นที่การศึกษา และระดับชาติ การวัดและประเมินผลการเรียนรู้เป็นกระบวนการพัฒนาคุณภาพผู้เรียน โดยใช้ผลการประเมินเป็นข้อมูลและสารสนเทศที่แสดงพัฒนาการความก้าวหน้า และความสำเร็จทางการเรียนของผู้เรียน ตลอดจนข้อมูลที่เป็นประโยชน์ต่อการส่งเสริมให้ผู้เรียนเกิดการพัฒนาและเรียนรู้อย่างเต็มตามศักยภาพ

การวัดและประเมินผลการเรียนรู้ แบ่งออกเป็น 4 ระดับ ได้แก่ ระดับชั้นเรียน ระดับสถานศึกษา ระดับเขตพื้นที่การศึกษา และระดับชาติ มีรายละเอียด ดังนี้

1. การประเมินระดับชั้นเรียน เป็นการวัดและประเมินผลที่อยู่ในกระบวนการจัดการเรียนรู้ ผู้สอนดำเนินการเป็นปกติและสม่ำเสมอ ในการจัดการเรียนการสอน ใช้เทคนิคการประเมิน อย่างหลากหลาย เช่น การซักถาม การสังเกต การตรวจการบ้าน การประเมินโครงการ การประเมินชิ้นงาน/ ภาระงาน แฟ้มสะสมงาน การใช้แบบทดสอบ ฯลฯ โดยผู้สอนเป็นผู้ประเมินเอง หรือ เปิดโอกาสให้ผู้เรียนประเมินตนเอง เพื่อนประเมินเพื่อน ผู้ปกครองร่วมประเมินในกรณีที่ไม่ผ่านตัวชี้วัดให้มีการสอนซ่อมเสริม

การประเมินระดับชั้นเรียนเป็นการตรวจสอบว่า ผู้เรียนมีพัฒนาการความก้าวหน้าในการเรียนรู้ อันเป็นผลมาจากการจัดกิจกรรมการเรียนการสอนหรือไม่และมากน้อยเพียงใด มีสิ่งที่จะต้องได้รับการพัฒนาปรับปรุงและส่งเสริมในด้านใด นอกจากนี้ยังเป็นข้อมูลให้ผู้สอนใช้ปรับปรุงการเรียนการสอนของตนด้วย ทั้งนี้โดยสอดคล้องกับมาตรฐานการเรียนรู้และตัวชี้วัด

2. การประเมินระดับสถานศึกษา เป็นการประเมินที่สถานศึกษาดำเนินการเพื่อตัดสิน ผลการเรียนรู้ของผู้เรียนเป็นรายปี/รายภาค ผลการประเมินการอ่าน คิดวิเคราะห์และเขียน คุณลักษณะอันพึงประสงค์ และกิจกรรมพัฒนาผู้เรียน นอกจากนี้เพื่อให้ได้ข้อมูลเกี่ยวกับ การจัดการศึกษาของสถานศึกษา ว่าส่งผลต่อการเรียนรู้ของผู้เรียนตามเป้าหมาย

หรือไม่ ผู้เรียนมี จุดพัฒนาในด้านใด รวมทั้งสามารถนำผลการเรียนของผู้เรียนในสถานศึกษา เปรียบเทียบกับเกณฑ์ระดับชาติ ผลการประเมินระดับสถานศึกษาจะเป็นข้อมูลและสารสนเทศ เพื่อการปรับปรุงนโยบาย หลักสูตร โครงการ หรือวิธีการจัดการเรียนการสอน ตลอดจนเพื่อ การจัดทำแผนพัฒนาคุณภาพการศึกษาของสถานศึกษาตามแนวทางการประกันคุณภาพ การศึกษา และการรายงานผลการจัดการศึกษาต่อคณะกรรมการสถานศึกษา สำนักงานเขต พื้นที่การศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน ผู้ปกครองและชุมชน

3. การประเมินระดับเขตพื้นที่การศึกษา เป็นการประเมินคุณภาพผู้เรียนใน ระดับเขตพื้นที่การศึกษาตามมาตรฐานการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน เพื่อใช้เป็นข้อมูลพื้นฐานในการพัฒนาคุณภาพการศึกษาของเขตพื้นที่การศึกษา ตามภาระความ รับผิดชอบ สามารถดำเนินการโดยประเมินคุณภาพผลสัมฤทธิ์ของผู้เรียน ด้วยข้อสอบ มาตรฐานที่จัดทำและดำเนินการ โดยเขตพื้นที่การศึกษา หรือด้วยความร่วมมือกับหน่วยงาน ต้นสังกัดในการดำเนินการจัดสอบ นอกจากนี้ยังได้จากการตรวจสอบทบทวนข้อมูลจากการ ประเมินระดับสถานศึกษา ในเขตพื้นที่การศึกษา

4. การประเมินระดับชาติ เป็นการประเมินคุณภาพผู้เรียนในระดับชาติ ตาม มาตรฐานการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน สถานศึกษาต้องจัดให้ผู้เรียน ทุกคนที่เรียน ในชั้นประถมศึกษาปีที่ 3 ชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และชั้น มัธยมศึกษาปีที่ 6 เข้ารับการประเมินผลจากการประเมินใช้เป็นข้อมูลในการเทียบเคียงคุณภาพ การศึกษาในระดับ ต่าง ๆ เพื่อนำไปใช้ในการวางแผนยกระดับคุณภาพการจัดการศึกษา ตลอดจนเป็นข้อมูลสนับสนุนการตัดสินใจในระดับนโยบายของประเทศสถานศึกษาในฐานะ ผู้รับผิดชอบจัดการศึกษา จะต้องจัดทำระเบียบว่าด้วยการวัดและประเมินผลการเรียนของ สถานศึกษาให้สอดคล้อง และเป็นไปตามหลักเกณฑ์และแนวปฏิบัติที่เป็นข้อกำหนดของ หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน เพื่อให้บุคลากรที่เกี่ยวข้องทุกฝ่าย ถือปฏิบัติร่วมกัน (กรมวิชาการ. 2551 : 28)

จากที่กล่าวมาแล้วข้างต้น สรุปได้ว่าการวัดผลและประเมินผล เป็นองค์ประกอบที่ สำคัญของการจัดการเรียนการสอน เพื่อเป็นการตรวจสอบว่าการเรียนรู้ของนักเรียนเป็นไป อย่างมีประสิทธิภาพและบรรลุจุดมุ่งหมายของหลักสูตรหรือไม่ นอกจากนี้ผลจากการวัดและ ประเมินจะเป็นข้อมูลย้อนกลับให้ครูผู้สอนนำไปพิจารณาหาวิธีแก้ไขจุดอ่อนและข้อบกพร่อง ต่าง ๆ ต่อไป

การทดสอบแบบทดสอบอิงเกณฑ์

1. จุดมุ่งหมายของการทดสอบอิงเกณฑ์

ไพศาล หวังวานิช (2535 : 187 - 188) กล่าวถึงจุดมุ่งหมายของการทดสอบอิงเกณฑ์ไว้ว่า จุดมุ่งหมายของการทดสอบอิงเกณฑ์คือ มุ่งสอบวัดเพื่อตรวจสอบการเรียนรู้ของผู้เรียนว่าบรรลุถึงเป้าหมายหรือถึงเกณฑ์ที่ตั้งไว้หรือไม่ การสอบจึงมุ่งเน้นการวินิจฉัยความสามารถของผู้เรียนเพื่อใช้ผลสำหรับการปรับปรุงการเรียนการสอน การสอบตามจุดมุ่งหมายดังกล่าวจึงเป็นการตอบคำถามว่าใครเรียน ได้ถึงเกณฑ์ที่ตั้งไว้ และใครควรต้องแก้ไขการเรียนหรือควรได้รับการปรับปรุงซ่อมเสริมการเรียนบ้าง คำถามเหล่านี้สามารถตอบได้โดยการนำคะแนนของผู้เรียนไปเทียบกับเกณฑ์ที่ตั้งไว้ โดยยอมรับว่าคะแนนคือสิ่งที่แทนประสิทธิภาพในการเรียนของแต่ละบุคคล

บุญชม ศรีสะอาด (2528 : 1) กล่าวถึงวัตถุประสงค์ของการประเมินผลแบบอิงเกณฑ์ว่า การประเมินผลแบบนี้ มีวัตถุประสงค์เพื่อวินิจฉัยว่าผู้เรียนบรรลุจุดประสงค์ที่วางไว้หรือไม่ ผู้เรียน คนใดยังไม่บรรลุวัตถุประสงค์ที่วางไว้ก็จะได้หาทางปรับปรุงแก้ไข เช่น จัดให้มีการซ่อมเสริมจะเห็นได้ว่าการประเมินผลแบบอิงเกณฑ์ เป็นผลจากความเชื่อในทฤษฎีการเรียนรู้เพื่อรอบรู้ (Learning for Mastery)

กัญญามาส สุดจริง (2545 : 13) ได้กล่าวถึงจุดมุ่งหมายของแบบทดสอบอิงเกณฑ์ไว้ว่า เพื่อตรวจสอบว่านักเรียนมีความสามารถผ่านเกณฑ์ที่กำหนดไว้หรือไม่ เพื่อนำข้อมูลไปปรับปรุง การเรียนการสอนต่อไป

ไพโรจน์ ใจดี (2546 : 11) ได้ให้จุดมุ่งหมายของแบบทดสอบแบบอิงเกณฑ์ไว้ว่า เพื่อการมุ่งตรวจสอบความรู้ของผู้เรียน โดยเทียบกับเกณฑ์มาตรฐานว่าผู้เรียนมีความรู้ผ่านเกณฑ์หรือไม่ เพื่อหาแนวทางพัฒนารูปแบบการเรียนการสอนให้มีประสิทธิภาพมากขึ้น

ไพรัตน์ จันทร โคตร (2547 : 14) ได้ให้จุดมุ่งหมายของแบบทดสอบอิงเกณฑ์ไว้ว่า คือ การทดสอบที่มุ่งให้ปรับปรุงการเรียนการสอน เพื่อตอบคำถามว่า “เด็กทำอะไรได้บ้าง” และเพื่อเปรียบเทียบความสามารถของเด็กกับเกณฑ์ที่มุ่งหวัง

เขวาคี วิบูลย์ศรี (2549 : 36) ได้กล่าวถึงจุดมุ่งหมายของแบบทดสอบอิงเกณฑ์ไว้ว่า เพื่อมุ่งวัดความสามารถในขั้นรอบรู้ของผู้เรียนตามเกณฑ์มาตรฐานที่กำหนด และมุ่งวัดระดับความสามารถในขั้นเรียนรู้ของผู้เรียนที่สูงกว่าขั้นรอบรู้ เพื่อบ่งชี้ถึงระดับของความก้าวหน้าที่ผู้เรียนได้พัฒนาขึ้นมาใหม่จากการเรียนการสอน

ลำดับความ เหล็กกล้า (2550 : 34) ได้ให้จุดมุ่งหมายของการทดสอบอิงเกณฑ์ไว้ว่า เพื่อปรับปรุงการเรียนการสอนในการตัดสินผลการเรียนการสอน โดยเทียบกับเกณฑ์การประเมินและเพื่อพัฒนาการเรียนการสอนให้มีประสิทธิภาพมากขึ้น

จากที่กล่าวมาสรุปได้ว่า จุดมุ่งหมายของแบบทดสอบอิงเกณฑ์คือ ใช้วัดผล การเรียนรู้ของผู้เรียน โดยใช้คะแนนจุดตัดหรือคะแนนเกณฑ์ตัดสินความรู้ และความไม่รอบรู้ของผู้เรียน เพื่อใช้เป็นข้อมูลในการปรับปรุงการเรียนการสอนหรือสอนซ่อมเสริมต่อไป

2. ความสำคัญของการทดสอบอิงเกณฑ์

ภัทรา นิคมานนท์ (2525 : 7) กล่าวถึงความสำคัญของแบบทดสอบอิงเกณฑ์ว่าเป็นเครื่องมือในการสอบและการวัดผลแบบอิงเกณฑ์ ซึ่งเป็นกระบวนการควบคุมไปกับการเรียนการสอนเป็นระยะ ๆ เป็นการสอบวัดเพื่อปรับปรุงการเรียนการสอน เมื่อผู้เรียนไม่สามารถทำข้อสอบได้ถึงเกณฑ์ที่กำหนดไว้ต้องมีการเรียนซ่อมเสริมในเนื้อหานั้น จนกว่าจะผ่านถึงเกณฑ์จึงสามารถเรียนหน่วยต่อไปได้ การสอบแบบอิงเกณฑ์จึงเหมาะสมกับการสอนในห้องเรียน

บุญเชิด ภิญโญนนตพงษ์ (2527 : 19) กล่าวว่าแบบทดสอบอิงเกณฑ์จะมีประโยชน์ที่สุดในการทำการตัดสินเกี่ยวกับความก้าวหน้า หรือการเรียนรู้ของนักเรียนภายใต้หลักสูตรใดหลักสูตรหนึ่ง ดังนั้นแบบทดสอบอิงเกณฑ์จึงเหมาะสมสำหรับใช้ในการตัดสินใจว่านักเรียนจำเป็นต้องเรียนซ่อมเสริมในหน่วยการเรียนนั้นหรือไม่ หรือใช้ในการตัดสินใจว่านักเรียนจำเป็นต้องเรียนต่อไปได้หรือไม่

ดังนั้น สรุปได้ว่า การทดสอบแบบอิงเกณฑ์และแบบทดสอบอิงเกณฑ์ที่มีความสำคัญต่อการเรียนการสอน คือ

1. ทำให้ทราบความสามารถของผู้เรียน ได้อย่างแจ่มชัด
2. สามารถบ่งชี้ประสิทธิภาพของการเรียนการสอนได้เป็นอย่างดี
3. เป็นแบบทดสอบที่เหมาะสมอย่างยิ่งสำหรับการทดสอบเพื่อปรับปรุง

การเรียน

แบบทดสอบอิงเกณฑ์

1. ความหมายของแบบทดสอบอิงเกณฑ์

นักการศึกษาหลายท่านได้ให้ความหมายของแบบทดสอบ ไว้ดังนี้

Millman (1974 : 327) ได้ให้ความหมายของแบบทดสอบอิงเกณฑ์ว่า

แบบทดสอบ อิงเกณฑ์เป็นแบบทดสอบที่ใช้วัดสภาพที่เป็นปัจจุบันของนักเรียน โดยมีการสุ่มตัวอย่างข้อสอบเป็นตัวแทนของประชากรข้อสอบในจุดมุ่งหมายนั้นทั้งหมด หรือเพื่อใช้ในการเป็นตัวแทนของกลุ่มพฤติกรรมทั้งหมด (Domain) ที่แสดงถึงคุณสมบัติตามจุดประสงค์นั้น ข้อสอบแบบนี้บางทีเรียกว่า (Domain- Referenced Test)

โปแฟม และ ฮูเซค (Popham and Husek. 1976 : 139) ให้ความหมายของแบบทดสอบอิงเกณฑ์ว่า เป็นแบบทดสอบสำหรับตรวจสอบผลสัมฤทธิ์ของบุคคลในการแสดงออกซึ่งพฤติกรรมต่าง ๆ แล้วเปรียบเทียบกับเกณฑ์ที่กำหนดไว้

Kubisqyn and Borich (1990 : 22) ได้ให้ความหมายของแบบทดสอบอิงเกณฑ์ไว้ว่า คือแบบทดสอบที่บอกระดับความสามารถของนักเรียน บอกความรอบรู้ในทักษะที่จะวัดหรือไม่เพียงใด โดยนำผลการทดสอบไปเทียบกับเกณฑ์มาตรฐาน

วิราพร พงศ์อาจารย์ (2542 : 55) ได้ให้ความหมายของแบบทดสอบอิงเกณฑ์ไว้ว่า คือแบบทดสอบที่สร้างเพื่อวัดพฤติกรรมหลักทั้งหมดของแต่ละเนื้อหา และกำหนดคะแนนเกณฑ์ไว้ล่วงหน้าว่าผู้สอบจะต้องมีความสามารถหรือทำข้อสอบได้ในระดับใด จึงจะยอมรับว่ามีความรู้ความสามารถบรรลุตามจุดประสงค์นั้น ๆ

กัญญาภาศ สุคจริง (2545 : 10) ได้ให้ความหมายของแบบทดสอบอิงเกณฑ์ไว้ว่า คือ แบบทดสอบที่สร้างขึ้นตามจุดประสงค์เชิงพฤติกรรมได้มาจากการวิเคราะห์เนื้อหาและจุดมุ่งหมายของหลักสูตร เพื่อใช้วัดความสามารถของผู้เรียนว่ามีความสามารถผ่านเกณฑ์ที่กำหนดไว้หรือไม่

ไพโรจน์ ใจดี (2546 : 9) ได้ให้ความหมายของแบบทดสอบอิงเกณฑ์ไว้ว่า คือแบบทดสอบที่สร้างขึ้นตามจุดประสงค์เชิงพฤติกรรมของรายวิชาหลักสูตร เพื่อใช้ทดสอบประเมินผลการเรียนการสอน และคะแนนที่ได้ของผู้สอบแต่ละคน ไปเปรียบเทียบกับเกณฑ์ที่กำหนดไว้ เพื่อตัดสินว่า ผู้สอบคนใดเป็นผู้รอบรู้หรือไม่รอบรู้

ไพรัตน์ จันทรโคตร (2547 : 11) ได้ให้ความหมายของแบบทดสอบอิงเกณฑ์ไว้ว่า คือแบบทดสอบที่สร้างขึ้นตามจุดมุ่งหมายเชิงพฤติกรรม ซึ่งได้จากการวิเคราะห์เนื้อหาและจุดมุ่งหมายของหลักสูตร คะแนนที่เด็กทำแบบทดสอบได้มาแปลความหมายเปรียบเทียบกับ

คะแนนเกณฑ์ที่กำหนดไว้แล้วว่ารอบรู้ในเนื้อหาอันเมื่อผ่านเกณฑ์ และไม่รอบรู้ในเนื้อหาอันเมื่อไม่ผ่านเกณฑ์

ถ้าควน เหล็กกล้า (2550 : 32) ได้ให้ความหมายของแบบทดสอบอิงเกณฑ์ไว้ว่า คือ แบบทดสอบที่สร้างขึ้นตามจุดประสงค์เชิงพฤติกรรมที่กำหนดไว้สำหรับการเรียนการสอน และคะแนนที่ได้ของผู้สอบแต่ละคนจะเปรียบเทียบกับเกณฑ์ขั้นต่ำที่กำหนดไว้ ตัดสินว่า ผู้สอบคนใดเป็นผู้รอบรู้ หรือไม่รอบรู้

จากที่กล่าวมาสรุปได้ว่า ความหมายของแบบทดสอบอิงเกณฑ์ คือ ชุดของคำถามที่สร้างขึ้นตามตัวชี้วัด และนำคะแนนที่ได้ของแต่ละคนมาเปรียบเทียบกับคะแนนเกณฑ์ขั้นต่ำที่กำหนด เพื่อตัดสินว่าผู้สอบคนใดเป็นผู้รอบรู้หรือไม่รอบรู้

2. ความเป็นมาของแบบทดสอบอิงเกณฑ์

ในปี ค.ศ. 1962 Glaser and Klaus เป็นผู้ใช้คำว่า การวัดผลแบบอิงเกณฑ์เป็นครั้งแรก และในปี 1963 Glaser เป็นผู้นำเอาความคิดในการวัดผลแบบอิงเกณฑ์มาใช้วัดผลสัมฤทธิ์ทางการเรียนเป็นครั้งแรกเช่นกัน โดยเขากล่าวว่าการวัดผลแบบอิงเกณฑ์นั้นขึ้นอยู่กับเกณฑ์มาตรฐานที่สมบูรณ์ที่จะบอกระดับคุณภาพของสิ่งที่วัดได้ (สำเร็จ บุญเรืองรัตน์. 2527: 4)

ส่วนในประเทศไทยได้มีผู้กล่าวถึงการวัดผลแบบอิงเกณฑ์ในราว พ.ศ. 2517 (ศ.วาสนา ประวาลพฤษย์. 2517 : 80 - 85) และได้มีการตีพิมพ์บทความด้านนี้ลงในวารสาร การวัดผล โดยเฉพาะวารสารการวัดผลการศึกษาฉบับเดือนมกราคม - เมษายน 2523 ได้ลงบทความทั้งเล่มว่าด้วยการวัดผลแบบอิงเกณฑ์ หลังจากนั้นกลุ่มผู้ที่สนใจด้านการวัดผลแบบอิงเกณฑ์ได้ค้นคว้าเพิ่มเติมตลอดมา ในที่สุดก็นำเอาการวัดผลแบบอิงเกณฑ์บางส่วนเข้าบรรจุในระเบียบกระทรวงศึกษาธิการว่าด้วยการประเมินผลการเรียนตามหลักสูตรประถมศึกษา 2521 (ฉบับปรับปรุง พ.ศ. 2533) พ.ศ. 2533 ตลอดจนระเบียบกระทรวงศึกษาธิการว่าด้วยการประเมินผล การเรียนตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 และมัธยมศึกษาตอนปลาย พุทธศักราช 2524 จนถึงปัจจุบัน

ในประเทศสหรัฐอเมริกาได้มีการเปลี่ยนแปลงระบบการเรียนการสอนในช่วงทศวรรษ 1960 - 1970 ที่เน้นการเรียนการสอนเป็นรายบุคคลและการเรียนเพื่อรอบรู้ จึงจำเป็นต้องใช้การวัดเพื่อรายงานความก้าวหน้าของนักเรียนเป็นรายบุคคล เพื่อประกอบการพิจารณาให้นักเรียนเลื่อนไปเรียนหน่วยเรียนต่อไป ต่อมาปี ค.ศ. 1963 แกลเซอร์ ได้นำเอาความคิดในการวัดผลแบบอิงเกณฑ์มาใช้ในการวัดผลสัมฤทธิ์ทางการเรียนเป็นครั้งแรก โดยเขากล่าวว่าการวัดผลแบบอิงเกณฑ์นั้นขึ้นอยู่กับเกณฑ์มาตรฐานที่สมบูรณ์ที่จะบอกระดับคุณภาพ

ของสิ่งที่วัดได้ (ถ้าควน เหล็กกล้า. 2550 : 32 - 33)

3. ลักษณะของแบบทดสอบอิงเกณฑ์

บุญชม ศรีสะอาด (2528 : 2) ได้กล่าวถึงลักษณะของแบบทดสอบอิงเกณฑ์ แทรกไว้ในลักษณะของการประเมินผลแบบอิงเกณฑ์ สรุปได้ดังนี้ แบบทดสอบอิงเกณฑ์เป็นแบบทดสอบที่สร้างเพื่อใช้วัดตามจุดประสงค์ของการเรียนการสอน โดยไม่นึกถึงความยากแต่เน้นสิ่งสำคัญ คือ เขียนคำถามตามเนื้อหาและตามจุดประสงค์ที่กำหนดไว้ และคำนึงถึงความเที่ยงตรงเชิงเนื้อหาเป็นสำคัญ การวิเคราะห์ข้อสอบใช้เกณฑ์ภายนอก เช่นกลุ่มคนตัวกับไม่ใช่คนตัว และคะแนนจะแปลออกมาในรูปของความรอบรู้ไม่รอบรู้

จากนักวัดผลได้กล่าวถึงลักษณะสำคัญของแบบทดสอบอิงเกณฑ์ไว้ว่า พอสรุปได้ว่า แบบทดสอบอิงเกณฑ์เป็นแบบทดสอบที่สร้างเพื่อใช้วัดตามจุดประสงค์ของการเรียนการสอน ใช้สำหรับวัดความรู้ความสามารถของนักเรียนแต่ละคนว่าถึงเกณฑ์ขั้นต่ำตามที่เราคาดหวังไว้หรือไม่

4. หลักการสร้างแบบทดสอบอิงเกณฑ์

หลักการสร้างแบบทดสอบอิงเกณฑ์มีนักวัดผลหลายท่านได้กล่าวไว้ ดังนี้ Stanley and Hopkins (1972 : 171) ได้ให้ขั้นตอนการสร้างแบบทดสอบอิงเกณฑ์ไว้ดังนี้

1. ขั้นวางแผนสร้างแบบทดสอบ ให้เลือกลักษณะของข้อสอบเป็นปรนัย
อัตรัย เต็มคำ หรือเลือกตอบก็ได้

2. ขั้นเตรียมงานเขียนข้อสอบและลงมือเขียนข้อสอบ

3. ขั้นทดลอง

4. ขั้นประเมินผลการทดลอง

Hambleton (1978 : 10) ได้กล่าวถึงหลักการสร้างข้อสอบแบบอิงเกณฑ์ไว้ ดังนี้

1. เลือกจุดประสงค์เตรียมกำหนดรายละเอียดของข้อสอบ เกี่ยวกับเวลา
จำนวนข้อ ขอบเขต คำศัพท์ การให้คะแนน

2. เขียนข้อสอบให้สอดคล้องกับจุดประสงค์

3. ตรวจสอบข้อคำถามขั้นต้น

4. กำหนดความเที่ยงตรงตามเนื้อหาข้อคำถาม โดยการตรวจสอบของ

ผู้เชี่ยวชาญ และวิเคราะห์จากการตอบของผู้สอบ

5. ตรวจสอบพิจารณาข้อคำถามอีกครั้ง

Gronlund (1993 : 8 - 11) ได้กล่าวถึงหลักการสร้างแบบทดสอบอิงเกณฑ์ไว้ ดังนี้

1. นิยามพฤติกรรม หรือผลการเรียนรู้ที่ต้องการวัด กำหนดในรูปแบบของ จุดประสงค์การเรียนรู้
2. สร้างแบบทดสอบให้ครอบคลุมผลการเรียนรู้ทั้งในระดับความรู้ ความจำ ความเข้าใจ การนำไปใช้ และระดับที่ซับซ้อนมากขึ้น
3. เขียนข้อสอบตามตัวชี้วัด และขอบเขตของผลการเรียนรู้ที่กำหนดไว้
4. สร้างข้อสอบให้สอดคล้องกับการวัดพฤติกรรม หรือผลการเรียนรู้ที่กำหนดไว้

5. สร้างข้อสอบตามแผน หรือวัตถุประสงค์ที่กำหนดไว้

6. แบบทดสอบที่สร้างการตรวจให้คะแนน ไม่มีความคาดเคลื่อนจากการวัด

พวงรัตน์ ทวีรัตน์ (2530 ก : 13) การสร้างข้อสอบแบบอิงเกณฑ์นั้น ก่อนลงมือเขียนข้อสอบจะต้องมีการวางแผนในขั้นตอนการวางแผนนี้ มีสิ่งที่จะต้องพิจารณาเพิ่มเติม นอกเหนือจากพิจารณาจุดมุ่งหมายของข้อสอบและเนื้อหาที่จะสร้างข้อสอบ คือ ต้องพิจารณาว่าจะสร้างข้อสอบโดยยึดจุดประสงค์ (Objective Based) คือ เขียนตามจุดประสงค์เชิงพฤติกรรม หรือสร้างข้อสอบโดยยึดกลุ่มพฤติกรรม (Domain Based) จึงจะเหมาะสม สำหรับการทดสอบ โดยใช้ข้อสอบที่สร้าง โดยยึดกลุ่มพฤติกรรมเป็นหลัก (Domain Referenced Test) นี้ถือว่าเมื่อสอบแล้วสามารถอ้างอิงได้ว่าบุคคลมีความสามารถในระดับใด หรือมีความรอบรู้ขนาดเท่าใด เมื่อเทียบจากประชากรพฤติกรรมทั้งหมด การสอบวัด โดยยึดพฤติกรรมเป็นหลัก (Domain Reference Test) นี้ถือว่าข้อสอบที่ใช้สอบเป็นกลุ่มตัวอย่างของพฤติกรรมที่กำหนดจากประชากรพฤติกรรม

บุญชม ศรีสะอาด (2535 : 85) ได้กล่าวถึงขั้นตอนในการสร้างแบบทดสอบอิงเกณฑ์ สรุปได้ดังนี้

1. วิเคราะห์จุดประสงค์ เนื้อหาวิชา

ขั้นแรกจะต้องทำการวิเคราะห์ดูว่ามีหัวข้อเนื้อหาใดบ้างที่ต้องการให้ผู้เรียนเกิดการเรียนรู้และที่จะต้องวัดแต่ละหัวข้อเหล่านั้น ต้องการให้ผู้เรียนเกิดพฤติกรรม หรือสมรรถภาพอะไร กำหนดออกมาให้ชัดเจน

2. กำหนดพฤติกรรมย่อยที่จะออกข้อสอบ

จากขั้นแรก พิจารณาต่อไปว่าจะวัดพฤติกรรมย่อยอะไรบ้าง อย่างละกี่ข้อ พฤติกรรมย่อยดังกล่าวคือจุดประสงค์เชิงพฤติกรรมนั่นเอง เมื่อกำหนดจำนวนข้อที่ต้องการจริงเสร็จแล้ว ต่อมาพิจารณาว่าจะต้องออกข้อสอบเกินไว้กี่ข้อ ควรออกเกินไว้ไม่ต่ำกว่า 25% ทั้งนี้ เนื่องจากหลังจากที่นำไปทดลองใช้ และวิเคราะห์หาคุณภาพของข้อสอบรายข้อแล้ว จะตัดข้อที่มีคุณภาพ ไม่เข้าเกณฑ์ออก ข้อสอบที่เหลือจะได้ไม่น้อยกว่าจำนวนที่ต้องการจริง

3. กำหนดรูปแบบของข้อคำถามและศึกษาวิธีเขียนข้อสอบ

ขั้นตอนนี้ จะเป็นการตัดสินใจว่าจะใช้คำถามรูปแบบใด และศึกษาวิธีการเขียนข้อสอบ เช่น ศึกษาหลักในการเขียนข้อคำถามแบบนั้น ๆ ศึกษาวิธีเขียนข้อสอบเพื่อวัดจุดประสงค์ประเภทต่าง ๆ ศึกษาเทคโนโลยีในการเขียนข้อสอบ เพื่อที่จะได้นำมาใช้ในการเขียนข้อสอบ ของตน

4. เขียนข้อสอบ

ลงมือเขียนข้อสอบตามจุดประสงค์เชิงพฤติกรรม ตามตารางที่ได้ กำหนดจำนวนข้อสอบของแต่ละจุดประสงค์เชิงพฤติกรรมไว้

5. ตรวจสอบข้อสอบ

นำข้อสอบที่ได้เขียนไว้แล้วในขั้นตอน 4 มาพิจารณาทบทวนอีกครั้งหนึ่ง โดยพิจารณาความถูกต้องตามหลักวิชา แต่ละข้อวัดพฤติกรรมย่อยหรือจุดประสงค์เชิงพฤติกรรมที่ต้องการหรือไม่ ภาษาที่ใช้เขียนมีความชัดเจนเข้าใจง่ายหรือไม่ ตัวถูกตัวลวงเหมาะสมเข้าเกณฑ์หรือไม่ ทำการปรับปรุงให้เหมาะสมยิ่งขึ้น

6. ให้ผู้เชี่ยวชาญพิจารณาความเที่ยงตรงตามเนื้อหา

นำจุดประสงค์เชิงพฤติกรรมและข้อสอบที่วัดแต่ละจุดประสงค์เชิงพฤติกรรมไปให้ผู้เชี่ยวชาญด้านการวัดผล และด้านเนื้อหาจำนวนไม่ต่ำกว่า 3 คน พิจารณาว่าข้อสอบแต่ละข้อวัดตามจุดประสงค์ที่ระบุไว้นั้นหรือไม่

7. พิมพ์แบบทดสอบฉบับทดลอง

นำข้อสอบทั้งหมดที่ผ่านการพิจารณาว่าเหมาะสมเข้าเกณฑ์ในขั้น 6 มาพิมพ์เป็นแบบทดสอบมีคำชี้แจงเกี่ยวกับแบบทดสอบวิธีตอบ จัดวางรูปแบบการพิมพ์ให้เหมาะสม

8. ทดลองใช้ วิเคราะห์คุณภาพและปรับปรุง

นำเอาแบบทดสอบไปทดลองสอบกับกลุ่มที่คล้ายกับกลุ่มตัวอย่างจริง

จำนวนประมาณ 40 คนหรือมากกว่า โดยสอบในช่วงแรกของการเรียนวิชานั้น เรียกว่า การสอบก่อนเรียน และนำแบบทดสอบเดิมมาสอบกับกลุ่มตัวอย่างเดิมอีกครั้งหนึ่งหลังจากที่ เรียนวิชานั้นจบแล้ว เรียกว่าการสอบหลังเรียน นำเอาผลการสอบสองครั้งมาวิเคราะห์หาค่า อำนาจจำแนก ของข้อสอบรายข้อ โดยใช้วิธีการวิเคราะห์ตามแบบอิงเกณฑ์ กัดเลือกข้อสอบที่มี อำนาจจำแนกเข้าเกณฑ์ตามจำนวนที่ต้องการ หาค่าความเชื่อมั่นแบบอิงเกณฑ์

9. พิมพ์แบบทดสอบฉบับจริง

นำข้อสอบที่มีค่าอำนาจจำแนกเข้าเกณฑ์ จากผลการวิเคราะห์ใน ขั้นตอนที่ 8 มาพิมพ์เป็นแบบทดสอบฉบับจริงต่อไป โดยเน้นรูปแบบการพิมพ์ที่ประณีต มีความถูกต้อง มีคำชี้แจงที่ละเอียดแจ่มชัด ผู้อ่านเข้าใจง่าย

วิราพร พงศ์อาจารย์ (2542 : 56) ได้กล่าวถึงหลักการสร้างแบบทดสอบไว้ดังนี้

1. กำหนดขอบเขตของการวัด โดยวิเคราะห์เนื้อหาและจุดมุ่งหมายของ หลักสูตร แยกเป็นหัวข้อเรื่องและพฤติกรรมหลักที่ต้องการวัด
2. เขียนจุดประสงค์ให้เฉพาะเจาะจง ก็นำพฤติกรรมหลักมาเขียนเป็น พฤติกรรมย่อยหลาย ๆ ข้อ เพื่อให้เฉพาะเจาะจงยิ่งขึ้น โดยเขียนจุดประสงค์เชิงพฤติกรรมเพื่อใช้ เป็นแนวทางในการเขียนข้อสอบ
3. เขียนข้อสอบ เป็นขั้นที่นำจุดประสงค์เชิงพฤติกรรม ลักษณะเฉพาะของ ข้อสอบ มาสร้างข้อคำถามให้สอดคล้องกับสิ่งที่ต้องการวัดและมีลักษณะของข้อสอบที่ดี ซึ่ง ควรให้จำนวนข้อคำถามในแต่ละจุดประสงค์มีมากกว่าจำนวนที่ต้องการจริงประมาณ 25 %
4. ตรวจสอบความเที่ยงตรงข้อสอบเป็นรายข้อ โดยตรวจสอบว่าคำถามแต่ละ ข้อ มีความตรงเชิงเนื้อหาและโครงสร้างหรือไม่ การตรวจสอบความเที่ยงตรงเชิงเนื้อหา โดย ผู้เชี่ยวชาญว่าข้อคำถามแต่ละข้อมีความสอดคล้องกับจุดประสงค์และเนื้อหาวิชาที่ต้องการวัด หรือไม่

5. ตรวจสอบความเที่ยงตรงเชิงโครงสร้าง เป็นการตรวจสอบว่า ข้อคำถาม สามารถจำแนกการรอบรู้หรือไม่รอบรู้ได้หรือไม่ โดยการวิเคราะห์หาค่าอำนาจจำแนก ระหว่างกลุ่มที่รอบรู้กับกลุ่มที่ไม่รอบรู้

6. กัดเลือกข้อสอบ โดยพิจารณาตัดสินว่า มีข้อสอบข้อใดบ้างที่มีคุณภาพ ตามเกณฑ์ที่กำหนด กล่าวคือมีความเที่ยงตรงเชิงเนื้อหา (สอดคล้องกับจุดประสงค์ที่จะวัด) ตั้งแต่ 50 % ขึ้นไป และมีความเที่ยงตรงเชิงโครงสร้าง (อำนาจจำแนกระหว่างกลุ่มที่รอบรู้และไม่รอบรู้) ตั้งแต่ 20 % ขึ้นไป

7. กำหนดความยาวและเกณฑ์ เป็นการนำข้อคำถามที่คัดเลือกแล้วมารวมกันเป็นฉบับ โดยกำหนดจำนวนข้อคำถามและเกณฑ์การบรรจุจุดประสงค์ที่จะเป็นมาตรฐานการปฏิบัติขั้นต่ำสุดที่ยอมรับว่า ผู้สอบรอบรู้ในเนื้อหานั้นได้เหมาะสม โดยต้องพิจารณาองค์ประกอบหลายอย่าง เพื่อให้เกิดความคลาดเคลื่อนน้อยที่สุด

8. ตรวจสอบคุณภาพทั้งฉบับว่า มีความเชื่อมั่นเพียงใด

9. นำแบบทดสอบที่สร้างไปใช้ และแปลความหมายคะแนนของผู้สอบเป็นรายบุคคล โดยเปรียบเทียบกับเกณฑ์หรือคะแนนจุดตัดที่กำหนดไว้

มหาวิทยาลัยสุโขทัยธรรมมาธิราช (2548 : 51 - 60) ได้ให้หลักการสร้างแบบทดสอบอิงเกณฑ์ไว้ว่า การสร้างข้อสอบวัดผลสัมฤทธิ์ทางการเรียนในระดับประถมศึกษาตามหลักสูตรที่แบ่งเป็นช่วงชั้นนั้น ข้อสอบที่สร้างจะมีลักษณะเป็นข้อสอบแบบรอบรู้ (Mastery Test) หรือเป็นข้อสอบแบบอิงเกณฑ์ ครูผู้สร้างจะต้องยึดว่ากิจกรรมการเรียนที่จัดขึ้นต้องมุ่งส่งเสริมและพัฒนาผู้เรียน ทุกคนให้ประสบความสำเร็จในการเรียน ถึงแม้ว่าผู้เรียนจะมีความสามารถที่แตกต่างกัน แต่ผู้เรียนทุกคนก็ควรได้รับการส่งเสริมให้พัฒนาเต็มตามศักยภาพสูงสุดที่มีอยู่ โดยอาจใช้เวลาที่แตกต่างกัน ดังนั้นการทดสอบแบบอิงเกณฑ์ที่กำหนดขึ้น จึงอยู่ที่การกำหนดคะแนนเกณฑ์เป็นสำคัญ สำหรับขบวนการสร้างมี 3 ขั้นตอน ดังนี้

1. วางแผนสร้างข้อสอบ

1.1 กำหนดวัตถุประสงค์ของการสอบให้ชัดเจน

1.2 ศึกษาหลักสูตร ทฤษฎี หลักการและงานวิจัยที่เกี่ยวข้องกับการสร้าง

แบบทดสอบ อิงเกณฑ์

1.3 วิเคราะห์เนื้อหาสาระวิชาและพฤติกรรมที่จะสอบวัด โดยวิเคราะห์เนื้อหาสาระที่จะออกข้อสอบเป็นหน่วยการเรียนรู้ สร้างตารางที่แสดงค่าตัวเลขระหว่างเนื้อหาสาระ และพฤติกรรมการเรียนรู้ด้านพุทธิพิสัย ได้แก่ ความรู้ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า

1.4 จัดทำตารางวิเคราะห์วัตถุประสงค์เชิงพฤติกรรม และแผนผังสร้าง

ข้อสอบ

1.5 กำหนดจำนวนข้อของข้อสอบ คะแนนทดสอบ ระยะเวลาทำการทดสอบ การดำเนินการทดสอบ วิธีการตรวจให้คะแนน การแปลความหมายคะแนน

2. การดำเนินการสร้างข้อสอบ

ดำเนินการสร้างข้อสอบให้สอดคล้องกับเนื้อหาสาระ จุดประสงค์เชิง

พฤติกรรม จัดทำต้นร่างแบบทดสอบ ตรวจสอบคุณภาพก่อนนำไปใช้ ปรับปรุงแก้ไขข้อสอบที่ยังคลุมเครือไม่ชัดเจน

3. การตรวจสอบคุณภาพข้อสอบก่อนนำไปใช้

ตรวจสอบคุณภาพข้อสอบ โดยการตรวจสอบความเที่ยงตรงเชิงเนื้อหา หรือหากำคัญนี้ความสอดคล้องของเนื้อหา ความยาก ค่าอำนาจจำแนก ความเชื่อมั่น จัดพิมพ์แบบทดสอบ และจัดทำคู่มือการใช้แบบทดสอบ

เขาวดี วิบูลย์ศรี (2549 : 178 - 186) ได้ให้หลักการสร้างแบบทดสอบไว้ 4 ขั้นตอน ดังนี้

1. กำหนดวัตถุประสงค์ทั่วไปให้อยู่ในรูปวัตถุประสงค์เชิงพฤติกรรมเป็นข้อ ๆ ให้สอดคล้องกับเนื้อหาสาระ

2. กำหนดโครงเรื่องของเนื้อหาสาระที่จะสอบให้ครบถ้วน

3. เตรียมตารางเฉพาะหรือผังของแบบสอบ เพื่อแสดงน้ำหนักของ

เนื้อหาวิชาแต่ละส่วน และพฤติกรรมต่าง ๆ ที่ต้องการ

4. สร้างข้อกระทงทั้งหมดที่ต้องการตามสัดส่วนของน้ำหนักที่ระบุไว้

อนุวัติ คุณแก้ว (2549 ก : 167 - 169) ได้ให้หลักการสร้างแบบทดสอบไว้ 3 ขั้นตอน คือ

1. การวางแผนสร้างข้อสอบ

1.1 ศึกษาวิธีสร้างแบบทดสอบจากเอกสารและงานวิจัย

1.2 กำหนดจุดมุ่งหมายของการใช้แบบทดสอบ

1.3 กำหนดเนื้อหา และพฤติกรรมที่ต้องการวัดในตารางวิเคราะห์

หลักสูตร ผู้สร้างต้องกำหนดขอบเขตเนื้อหา มาตรฐานการเรียนรู้ สาระการเรียนรู้ พฤติกรรมที่จะวัดด้าน พุทธิพิสัย ได้แก่ ความรู้ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า

1.4 กำหนดลักษณะของข้อสอบ เป็นแบบอิงเกณฑ์หรืออิงกลุ่ม เป็นแบบปรนัยหรืออัตนัย เป็นทั้งแบบปรนัยและอัตนัยรวมกันก็ได้ กำหนดความยาวของข้อสอบ หรือจำนวนข้อสอบและคะแนน ระยะเวลาในการทำแบบทดสอบ วิธีดำเนินการสอบ วิธีตรวจให้คะแนน

2. การลงมือสร้างข้อสอบ

2.1 สร้างข้อสอบ ผู้สร้างข้อสอบตามรายละเอียดในตารางวิเคราะห์หลักสูตรตามลักษณะของข้อสอบ คำนึงถึงความยากของแบบทดสอบ ระยะเวลาที่ใช้สอบ คะแนน และการตรวจให้คะแนน

2.2 ตรวจทานข้อสอบ ผู้สร้างต้องตรวจทานข้อสอบ ให้ข้อสอบมีความถูกต้อง ครบถ้วน ตามรายละเอียดที่กำหนดไว้ในตารางวิเคราะห์หลักสูตร แล้วจัดพิมพ์

3. การตรวจสอบคุณภาพข้อสอบก่อนนำไปใช้

3.1 นำแบบทดสอบที่สร้าง ไปให้ผู้เชี่ยวชาญด้านเนื้อหา และด้านวัดผลการศึกษาจำนวน 3 - 5 ท่าน ตรวจสอบเที่ยงตรงด้านเนื้อหา (Content Validity) โดยผู้เชี่ยวชาญพิจารณาความสอดคล้องของข้อสอบกับจุดประสงค์การเรียนรู้ หรือเนื้อหา ตามตารางวิเคราะห์หลักสูตร โดยใช้เกณฑ์การประเมิน ดังนี้

- +1 หมายถึง แน่ใจว่าข้อสอบวัดจุดประสงค์ข้อนั้น
- 0 หมายถึง ไม่แน่ใจว่าข้อสอบวัดจุดประสงค์ข้อนั้น
- 1 หมายถึง แน่ใจว่าข้อสอบไม่วัดจุดประสงค์ข้อนั้น

นำข้อมูลที่ได้ หากค่าความสอดคล้อง (IOC) คัดเลือกข้อสอบที่มีค่า IOC ตั้งแต่ .50 ขึ้นไป จัดพิมพ์เป็นแบบทดสอบฉบับใหม่

3.2 ทดลองสอบ นำแบบทดสอบที่ปรับปรุงแก้ไขแล้ว ไปทดลองสอบ (Try Out) กับนักเรียนที่มีลักษณะคล้ายคลึง หรือนักเรียนที่เพิ่งเรียนในเรื่องนั้น ๆ จำนวนตั้งแต่ 30 คนขึ้นไป

3.3 วิเคราะห์หาคุณภาพข้อสอบ นำผลการสอบมาวิเคราะห์หาค่าความยาก และค่าอำนาจจำแนกเป็นรายข้อ แล้วคัดเลือกข้อสอบที่ใช้ได้ คือ มีค่าความยากง่าย (P) ระหว่าง .20 - .80 และมีค่าอำนาจจำแนก (r) ตั้งแต่ .20 ขึ้นไป ให้นำข้อสอบที่คัดเลือกแล้ว จัดพิมพ์เป็นแบบทดสอบฉบับใหม่ นำไปทดลองสอบกับนักเรียนที่มีลักษณะคล้ายคลึง หรือนักเรียนที่เพิ่งเคยเรียนเรื่องนั้น ๆ จำนวนตั้งแต่ 30 คนขึ้นไป เพื่อหาความเชื่อมั่น

3.4 จัดพิมพ์เป็นฉบับจริง เพื่อนำไปใช้กับกลุ่มเป้าหมายต่อไป

พิชิต ฤทธิ์จรูญ (2550 : 97) ได้ให้หลักการสร้างแบบทดสอบอิงเกณฑ์ไว้ ดังนี้

1. วิเคราะห์หลักสูตรและสร้างตารางวิเคราะห์หลักสูตร
2. กำหนดจุดประสงค์การเรียนรู้
3. กำหนดชนิดของข้อสอบและศึกษาวิธีสร้างตารางวิเคราะห์หลักสูตร

วิธีเขียนข้อสอบ

4. เขียนข้อสอบตามรายละเอียดที่กำหนดไว้ในตารางวิเคราะห์หลักสูตรให้สอดคล้องกับจุดประสงค์การเรียนรู้

5. ตรวจสอบข้อสอบให้มีความถูกต้องตามหลักวิชา ก่อนที่จะจัดพิมพ์และนำไปใช้

6. จัดพิมพ์แบบทดสอบฉบับทดลอง มีคำชี้แจงวิธีตอบแบบทดสอบ (Direction) และจัดวางรูปแบบการพิมพ์ให้เหมาะสม

7. ทดลองสอบและวิเคราะห์ข้อสอบ เพื่อหาคุณภาพก่อนนำไปใช้จริงกับกลุ่มที่มีลักษณะคล้ายคลึงกับกลุ่มที่ต้องการสอบจริง แล้วนำผลการสอบมาวิเคราะห์และปรับปรุงข้อสอบให้มีความเหมาะสม

8. จัดทำแบบทดสอบฉบับจริงเพื่อนำไปใช้กับกลุ่มเป้าหมายต่อไป

จากหลักการสร้างแบบทดสอบอิงเกณฑ์ที่กล่าวมา ผู้วิจัยได้สร้างแบบทดสอบอิงเกณฑ์ สรุปได้ดังนี้คือ ศึกษาแนวคิดทฤษฎี เอกสารงานวิจัยที่เกี่ยวข้อง ศึกษาสาระการเรียนรู้แกนกลางและ มาตรฐานการเรียนรู้ สร้างตารางวิเคราะห์หลักสูตรเพื่อกำหนดน้ำหนักของสาระการเรียนรู้แกนกลางให้อยู่ในรูปตัวชี้วัดเป็นสัดส่วนร้อยละ และคำนวณหาจำนวนข้อสอบแต่ละมาตรฐานการเรียนรู้ เช่น ความรู้ความจำ ความเข้าใจ การนำไปใช้ นำแบบทดสอบไปให้ผู้เชี่ยวชาญตรวจสอบความเที่ยงตรงเชิงเนื้อหา จัดพิมพ์ข้อสอบฉบับทดลอง ทดลองสอบ นำผลมาวิเคราะห์และปรับปรุงคุณภาพด้วยวิธีการทางสถิติ จัดพิมพ์แบบทดสอบฉบับจริงนำไปใช้ในการวัดผลและประเมินผลต่อไป

5. การหาคุณภาพของแบบทดสอบอิงเกณฑ์

นักวิชาการหลายท่านได้กล่าวถึงการวิเคราะห์คุณภาพของแบบทดสอบอิงเกณฑ์ไว้ดังนี้

บุญชม ศรีสะอาด นิภา ศรีไพโรจน์ และนุชนา ทองทวี (2528 : 126) กล่าวว่า เนื่องจากการประเมินผลแบบอิงเกณฑ์เป็นการทดสอบซึ่งเปรียบเทียบความสามารถของผู้เรียนกับเกณฑ์ที่เป็นมาตรฐานกับเกณฑ์ที่แท้จริงว่าอยู่ในระดับถึงมาตรฐานที่ยอมรับได้หรือไม่การวิเคราะห์ข้อสอบเพื่อการประเมินผลแบบอิงเกณฑ์จึงเน้นความเที่ยงตรงเชิงเนื้อหาและความเที่ยงตรงเชิงโครงสร้าง กล่าวคือข้อสอบทุกข้อต้องสามารถวัดพฤติกรรมที่ต้องการจะวัดตามจุดประสงค์ และคะแนนสอบสามารถบ่งชี้ความรู้ – ไม่รอบรู้ได้

การวิเคราะห์แบบทดสอบอิงเกณฑ์ ที่จะนำมาวิเคราะห์ต้องเป็นแบบทดสอบแบบเลือกตอบที่มีวิธีให้คะแนน 1 คะแนน สำหรับคำตอบที่ถูกต้อง และ 0 คะแนน สำหรับคำตอบที่ผิด (บุญชม ศรีสะอาด. 2535 : 61)

พิมพา สุวรรณฤทธิ์ (2542 : 185-187) ได้กล่าวถึงการหาคุณภาพของแบบทดสอบอิงเกณฑ์ไว้ว่า พิจารณาได้จากความแตกต่างของผลการสอบก่อนสอน และหลังสอน ข้อสอบใดที่มีนักเรียนตอบถูกหลังสอนมากกว่าก่อนสอน แสดงว่าข้อสอบข้อนั้นสามารถวัดความก้าวหน้าในการสอนได้ ค่าความแตกต่างระหว่างผลการสอบหลังสอนกับผลการสอบก่อนสอนยิ่งมากเท่าไรยิ่งดี เพราะแสดงให้เห็นถึงประสิทธิภาพของการสอน ซึ่งเรียกว่า ค่าความไวในการวัดผล ที่ได้รับการสอน (Sensitivity) เขียนเป็นสูตรได้ดังนี้ (พิมพา สุวรรณฤทธิ์. 2542 : 185)

$$S = \frac{A - B}{N}$$

เมื่อ	S	แทน	ค่าความไวในการวัดผลที่ได้รับการสอน
	A	แทน	จำนวนคนที่ตอบถูกหลังการสอน
	B	แทน	จำนวนคนที่ตอบถูกต้องก่อนการสอน
	N	แทน	จำนวนคนที่เข้าสอบทั้งหมด

บรรดล สุขปีติ (2542 : 191- 205) ได้กล่าวถึงการหาคุณภาพของแบบทดสอบอิงเกณฑ์ไว้ว่าจะมุ่งเน้นความเที่ยงตรง 2 ลักษณะ คือ

1. ตรวจสอบความสอดคล้องของข้อสอบกับพฤติกรรมที่ต้องการวัดตามที่ระบุไว้ในจุดประสงค์การเรียนรู้ มีวิธีตรวจสอบ 3 วิธี ตามแนวคิดของโรวินลดีและแฮมเบิลตัน

1.1 ใช้ดัชนีความสอดคล้องระหว่างข้อคำถามกับจุดประสงค์ โดยใช้สูตรดังนี้ (บรรดล สุขปีติ. 2542 : 192)

$$IOC = \frac{\Sigma R}{N}$$

เมื่อ	IOC	แทน	ค่าความสอดคล้องระหว่างข้อคำถามกับจุดประสงค์
	ΣR	แทน	ผลรวมของคะแนนของผู้เชี่ยวชาญแต่ละคน
	N	แทน	จำนวนผู้เชี่ยวชาญที่ได้ตรวจสอบ
	ค่า IOC	ที่ยอมรับได้คือ	มีค่าตั้งแต่ 0.50 ขึ้นไป

1.2 ใช้ดัชนีความสอดคล้องระหว่างข้อคำถามกับจุดประสงค์ โดยค่า \bar{X} วิธีนี้มีลักษณะเช่นเดียวกับวิธีที่ 1 คือนำข้อคำถามและจุดประสงค์การเรียนรู้ไปให้ผู้เชี่ยวชาญหลาย ๆ คนตรวจสอบ จะต่างกันตรงการกำหนดน้ำหนักคะแนน ซึ่งจะใช้มาตราส่วนประเมินค่า 5 ช่วงน้ำหนัก เพื่อประเมินความมั่นใจในการวัดพฤติกรรมตามจุดประสงค์ ดังนี้

มีความมั่นใจมากที่สุด	ให้	4
มีความมั่นใจมาก	ให้	3
มีความมั่นใจปานกลาง	ให้	2
มีความมั่นใจน้อย	ให้	1
ไม่มีความมั่นใจเลย	ให้	0

เกณฑ์ในการแปลความหมาย ค่า \bar{X} ที่ยอมรับได้คือ 2.50 ข้อคำถามที่ยอมรับมีค่าความเบี่ยงเบนมาตรฐาน (S.D) ไม่เกิน 1.00

1.3 ใช้ดัชนีการจับคู่ระหว่างข้อคำถามกับจุดประสงค์ วิธีนี้ให้ผู้เชี่ยวชาญพิจารณาเช่นเดียวกับวิธีที่ 1 และวิธีที่ 2 แต่จะแตกต่างกันในวิธีการตรวจสอบ กล่าวคือจะนำจุดประสงค์การเรียนรู้ชุดหนึ่ง และข้อคำถามในแบบทดสอบ ซึ่งวัดพฤติกรรมตามจุดประสงค์ไปให้ผู้เชี่ยวชาญพิจารณาจับคู่กัน นำผลการจับคู่มาพิจารณาร่วมกันว่ามีความสอดคล้องกันในระดับใด ข้อคำถามที่ผู้เชี่ยวชาญมีความเห็นสอดคล้องกันมากยอมรับว่าเป็นข้อคำถามที่วัดพฤติกรรมตามจุดประสงค์การเรียนรู้ นั้น แต่ถ้าข้อคำถามใดที่ผู้เชี่ยวชาญมีความเห็นสอดคล้องกันน้อยจะไม่ยอมรับ สำหรับสถิติที่ใช้ในการตรวจสอบคือ ค่าไค - สแควร์ (χ^2) เกณฑ์ในการพิจารณา ข้อคำถามที่มีเปอร์เซ็นต์ของความสอดคล้องตั้งแต่ 80 ขึ้นไป ก็ยอมรับว่าข้อคำถามข้อนั้นวัดพฤติกรรมตรงตามจุดประสงค์การเรียนรู้ ส่วนข้อคำถามที่มีเปอร์เซ็นต์ของความสอดคล้องน้อยกว่า 80 ไม่ยอมรับ

2. ตรวจสอบความสัมพันธ์สอดคล้องระหว่างผู้ที่รอบรู้หรือไม่รอบรู้

การตรวจสอบความเที่ยงตรงของแบบทดสอบ โดยวิธีนี้จะอาศัยผลการตอบของกลุ่มตัวอย่างที่แบ่งเป็น 2 กลุ่ม คือ กลุ่มผู้รอบรู้และกลุ่มผู้ไม่รอบรู้เป็นหลักในการพิจารณา ดังนี้

2.1 ใช้คุลยพินิจของครูผู้สอน วินิจฉัยว่าใครคือผู้รอบรู้ ใครคือผู้ไม่รอบรู้

2.2 ใช้สภาพการเรียนของนักเรียน นักเรียนที่ได้รับการเรียนการสอน

ตามจุดประสงค์เป็นอย่างดีมาแล้วจะถือว่าเป็นผู้รอบรู้ และนักเรียนที่ยังไม่ได้รับการเรียน การสอนตามจุดประสงค์ดังกล่าวเป็นผู้ไม่รอบรู้ วิธีนี้จะมีสมมติฐานว่าการเรียนการสอนคือเงิน เชื่อมั่นได้ว่า ผู้ที่ผ่านการเรียนการสอนนั้นแล้วจะมีความรู้ความสามารถตามจุดประสงค์ที่กำหนด

2.3 ใช้คะแนนรวมจากแบบทดสอบฉบับนั้น โดยกำหนดคะแนนเกณฑ์ ขั้นต่ำจำแนกผู้รอบรู้ (ผ่านเกณฑ์) กับผู้ไม่รอบรู้ (ไม่ผ่านเกณฑ์) การตรวจสอบวิธีนี้คำนวณได้ 2 วิธี คือ

2.3.1 วิธีการหาค่าอำนาจจำแนกของเบรนนอน ใช้สูตรดังนี้

(บรรดล สุขปิติ. 2542 : 198)

$$B = (U/N_1) - (L/N_2)$$

เมื่อ B แทน ค่าอำนาจจำแนกที่จะใช้พิจารณาค่าความเที่ยงตรง

N_1 แทน จำนวนนักเรียนในกลุ่มผู้รอบรู้

N_2 แทน จำนวนนักเรียนในกลุ่มผู้ไม่รอบรู้

U แทน จำนวนนักเรียนในกลุ่มผู้รอบรู้ (N_1) ตอบถูก

L แทน จำนวนนักเรียนในกลุ่มผู้ไม่รอบรู้ (N_2) ตอบถูก

ข้อคำถามที่ยอมรับว่านำไปใช้ได้อย่างมีความเที่ยงตรงจะต้องมีค่าอำนาจ จำแนกของเบรนนอนตั้งแต่ .20 ขึ้นไป

2.3.2 วิธีหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างการตอบถูกหรือผิด กับการแบ่งกลุ่มผู้รอบรู้หรือยังไม่รู้ กับการแจกแจงแล้วจัดลงในตาราง 2 X 2 และใช้สถิติ คือ สัมประสิทธิ์สหสัมพันธ์ แบบพี (ϕ) หรือดัชนีความแม่นยำของคาร์เวอ์ (Carver)

เยวดี วิบูลย์ศรี (2549 : 156 - 157) ได้กล่าวถึงการหาคุณภาพของแบบทดสอบไว้ ว่า ความยากง่ายของแบบทดสอบอิงเกณฑ์จะพิจารณาจากระดับความยากง่ายของการเรียนรู้ การตัดข้อที่ยากเกินไปหรือง่ายเกินไปตามแนวคิดของการวิเคราะห์แบบทดสอบอิงกลุ่มนั้น ทำให้ขาดข้อที่จะให้สาระความรู้ที่มีประโยชน์ แบบทดสอบอิงเกณฑ์ควรมีประสิทธิภาพในการวัดผลของการจัดการเรียนการสอนโดยตรง ถ้าครูใช้แบบสอบชุดเดียวกันทดสอบนักเรียน ก่อนการสอนและหลังการสอนในเนื้อหานั้นเสร็จสิ้นลงแล้ว แบบทดสอบอิงเกณฑ์ที่มี ประสิทธิภาพอยู่บนสมมติฐานที่ว่ากระบวนการเรียนการสอน ส่งผลโดยตรงต่อการเรียนรู้ของ นักเรียน มีผู้ตอบถูกเมื่อการเรียนการสอนสิ้นสุดมากกว่าผู้สอบถูกก่อนที่จะมีการเรียนการสอน จากแนวคิดดังกล่าว ดัชนีความไวของผลการเรียนการสอน (Index of Sensitivity) จึงเกิดขึ้น

คำนวณ ได้ดังสูตร

$$S = \frac{R_A - R_B}{T}$$

เมื่อ S แทน ดัชนีความไวของผลการเรียนการสอน

R_A แทน จำนวนนักเรียนที่ตอบถูกหลังการสอน

R_B แทน จำนวนนักเรียนที่ตอบถูกก่อนการสอน

T แทน จำนวนนักเรียนทั้งหมดที่ตอบข้อสอบ

ดัชนีความไวนี้จะมีค่าอยู่ระหว่าง +1 ถึง -1 ข้อสอบที่มีประสิทธิภาพจะมีค่าดัชนีความไวเป็นบวกและเข้าใกล้ 1 มากเท่าใดข้อสอบข้อนั้นยิ่งมีประสิทธิภาพมากยิ่งขึ้น ค่าดัชนีที่มีค่าต่ำหรือมีค่าเป็นลบ แสดงว่าข้อสอบข้อนั้นไม่มีประสิทธิภาพควรตัดออก ซึ่งการหาดัชนีความไว จะบ่งบอกความมีประสิทธิภาพของข้อสอบ หรือความไม่มีประสิทธิภาพของข้อสอบ ความมีประสิทธิภาพของการสอน หรือความไม่มีประสิทธิภาพของการสอนได้

อนูวัติ คุณแก้ว (2549 : 238 - 246) ได้กล่าวถึงการหาคุณภาพของแบบทดสอบ
อิงเกณฑ์ ไว้ดังนี้

1. หาความเที่ยงตรงเชิงเนื้อหา (Content Validity)

หาความเที่ยงตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ พิจารณาว่าข้อสอบหรือข้อคำถามแต่ละข้อว่าวัดเนื้อหาตรงจุดประสงค์การเรียนรู้หรือไม่ ซึ่งใช้เกณฑ์การประเมิน ดังนี้

ให้คะแนน +1 หมายถึง แน่ใจว่าข้อสอบวัดจุดประสงค์/เนื้อหานั้น

ให้คะแนน 0 หมายถึง ไม่แน่ใจว่าข้อสอบวัดจุดประสงค์/เนื้อหานั้น

ให้คะแนน -1 หมายถึง แน่ใจว่าข้อสอบไม่วัดจุดประสงค์/เนื้อหานั้น

แล้วนำข้อมูลที่ได้จากการพิจารณาของผู้เชี่ยวชาญ หาค่าความสอดคล้องระหว่างข้อสอบแต่ละข้อกับจุดประสงค์ หรือ เนื้อหา (Index of Item Objective Congruence หรือ IOC) โดยใช้สูตรของ Rovinelli and Hambelton (อนูวัติ คุณแก้ว. 2549 : 238) ดังนี้

$$IOC = \frac{\sum R}{N}$$

เมื่อ $\sum R$ แทน ผลรวมของคะแนนการพิจารณาของผู้เชี่ยวชาญ

N แทน จำนวนผู้เชี่ยวชาญ

ข้อสอบที่ใช้ได้ คือ ข้อสอบที่มีค่า IOC ตั้งแต่ 0.50 ขึ้นไป

2. หาความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) เป็นการหาว่านักเรียนที่มีความรอบรู้จะทำข้อสอบข้อนั้นได้ นักเรียนที่ไม่มีความรอบรู้จะทำข้อสอบข้อนั้นไม่ได้ ซึ่งสามารถหาได้โดยวิธีของคาร์เวอร์ (Carver) และหาค่าสหสัมพันธ์แบบพี (Phi-correlation) มีนักวิชาการหลายท่าน อาทิ วาโร เฟ็งสวัสต์ (2542 : 46) ได้กล่าวว่า การหาความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) หมายถึง คุณลักษณะของแบบทดสอบที่สามารถวัดคุณลักษณะ หรือพฤติกรรมที่ต้องการวัดได้อย่างถูกต้อง เช่น ต้องการวัดเจตคติ ลักษณะของคำถามที่ใช้ควรเป็นพฤติกรรมที่เกี่ยวกับเจตคติ ซึ่งสอดคล้องกับบรรดล สุขปิติ (2542 : 156) ได้กล่าวว่าความเที่ยงตรงเชิงโครงสร้าง โดยส่วนใหญ่จะใช้ตรวจสอบกับแบบทดสอบที่วัดทางบุคลิกภาพ ด้านความถนัดหรือด้านเจตคติ ซึ่งแตกต่างกับความเที่ยงตรงเชิงเนื้อหาใช้ในการตรวจสอบแบบทดสอบวัดผลสัมฤทธิ์

3. การหาดัชนีความไวของข้อสอบ (Sensitivity Index) ที่เสนอโดยคริสปิน และเฟลด์สตูเซน (Krtspin and Feldlusion) มีสูตรดังนี้ (อนุวัติ ภูณแก้ว. 2549 : 244)

$$S = P_{\text{post}} - P_{\text{pre}}$$

เมื่อ S แทน ดัชนีความไวของข้อสอบ
 P_{post} แทน ความยากหลังเรียน
 P_{pre} แทน ความยากก่อนเรียน

4. การหาค่าอำนาจจำแนก โดยวิธีของเบรนนเนน (Brennan) จะแบ่งผู้เรียนเป็น 2 กลุ่ม คือ กลุ่มนักเรียนที่สอบผ่าน กับกลุ่มนักเรียนที่สอบไม่ผ่าน โดยนำข้อสอบไปทดสอบหลังเรียนและนำมาเทียบกับเกณฑ์การผ่าน หรือคะแนนจุดตัดของแบบทดสอบ โดยใช้สูตรของเบรนนเนน ดังนี้ (อนุวัติ ภูณแก้ว. 2549 : 245)

$$B = \frac{U}{n_1} - \frac{L}{n_2}$$

เมื่อ B แทน ดัชนีค่าอำนาจจำแนกของแบบทดสอบอิงเกณฑ์
 U แทน จำนวนผู้สอบที่ตอบถูกของกลุ่มผ่านเกณฑ์
 L แทน จำนวนผู้สอบที่ถูกต้องของกลุ่มไม่ผ่านเกณฑ์
 n_1 แทน จำนวนผู้สอบที่ผ่านเกณฑ์
 n_2 แทน จำนวนผู้สอบที่ไม่ผ่านเกณฑ์

การแปลดัชนีอำนาจจำแนก มีดังนี้

ค่าอำนาจจำแนก	ความหมาย
1.00	จำแนกผู้รอบรู้ - ไม่รอบรู้ได้ถูกต้องทุกคน
50 - 0.90	จำแนกผู้รอบรู้ - ไม่รอบรู้ได้ถูกต้องเป็นส่วนใหญ่
0.20 - 0.49	จำแนกผู้รอบรู้ - ไม่รอบรู้ได้ถูกต้องเป็นบางส่วน
0.00 - 0.19	จำแนกผู้รอบรู้ - ไม่รอบรู้ได้ถูกต้องน้อยมากหรือไม่ถูกเลย
เป็นลบ	จำแนกผู้รอบรู้ - ไม่รอบรู้ผิดพลาด ตรงข้ามกับความจริง

พิชิต ฤทธิจรูญ (2550 : 143 - 146) ได้กล่าวถึงการหาคุณภาพของแบบทดสอบอิงเกณฑ์ไว้ดังนี้

1. ความแตกต่างของการสอบก่อนสอนและการสอบหลังสอน หรือ คำนีความไวของการวัดผล (Sensitivity Index) หรือ คำนีความไวของการวัดผลจากการสอน (Sensitivity to Instructional Effects) ซึ่งมีวิธีการดังนี้

1.1 นำแบบทดสอบไปทดสอบกับผู้สอบกลุ่มหนึ่งทั้งก่อนสอนและหลังสอนด้วยข้อสอบชุดเดิม

1.2 ตรวจสอบกระดาษคำตอบของผู้สอบที่สอบก่อนสอนและหลังสอน

1.3 หาค่าความยากของแต่ละข้อโดยสูตรเช่นเดียวกับการวิเคราะห์

ข้อสอบแบบอิงกลุ่ม คือ (พิชิต ฤทธิจรูญ. 2550 : 143)

$$\text{ความยาก (P)} = \frac{\text{จำนวนคนที่ตอบถูกในแต่ละข้อ}}{\text{จำนวนคนทั้งหมดที่ทำข้อสอบในแต่ละข้อ}}$$

ดังนั้นข้อสอบแต่ละข้อจะมีค่า P จำนวน 2 ค่า คือ ความยากก่อนสอน (P_{pre}) และความยากหลังสอน (P_{post})

1.4 พิจารณาค่า P ก่อนสอน ดังนี้

1.4.1 ค่า $P \leq .30$ แสดงว่า มีผู้ตอบถูกไม่เกิน 30 % ซึ่งเป็นค่าที่เป็นไปตามคาดหวังว่า ก่อนสอนไม่ควรจะมีผู้ตอบถูกเลย หรือตอบถูกไม่ควรมากกว่านี้

1.4.2 ค่า $P > .30$ แสดงว่า เป็นข้อสอบที่ควรปรับปรุง หากค่า P สูงมาก แสดงว่าผู้เรียนมีความรู้ตั้งแต่ยังไม่ได้อ่าน ซึ่งจำเป็นต้องตัดข้อนั้นทิ้งโดยตัด

จุดประสงค์ข้อนี้ทั้งเพราะนักเรียนรู้แล้ว หรือปรับข้อสอบให้ยากขึ้น โดยปรับจุดประสงค์ข้อนี้ให้ยากขึ้น

1.5 หาดัชนีความไวในการวัดผล ที่เสนอโดยคริสปินและเฟลด์ลูซอน (Krtspin and Feldluson) (พิชิต ฤทธิ์จรูญ. 2550 : 143) ดังนี้

$$S = P_{\text{post}} - P_{\text{pre}}$$

เมื่อ S แทน ดัชนีความไวของการวัดผล
 P_{post} แทน ความยากหลังเรียน
 P_{pre} แทน ความยากก่อนสอน

2. วิธีการแบ่งผู้เรียนออกเป็น 2 กลุ่ม คือ กลุ่มที่สอบผ่านเกณฑ์ และกลุ่มที่สอบไม่ผ่านเกณฑ์ วิธีนี้คำนวณจากผลการทดสอบเพียงครั้งเดียว โดยเอาแบบทดสอบไปสอบผู้เรียนหลังจากผู้เรียนเรียนรู้แล้ว แล้วนำข้อมูลมาคำนวณจากสัดส่วนของผู้เรียนที่ทำข้อสอบถูกของกลุ่มที่สอบผ่านเกณฑ์ กับกลุ่มที่สอบไม่ผ่านเกณฑ์ ซึ่งจะต้องรู้คะแนนเกณฑ์หรือคะแนนจุดตัดของแบบทดสอบก่อน โดยใช้สูตรของ Brennan หรือ Discrimination Index B ดังนี้ (พิชิต ฤทธิ์จรูญ. 2550 : 145)

$$B = \frac{U}{n_1} - \frac{L}{n_2}$$

เมื่อ B คือ ดัชนีค่าอำนาจจำแนกของข้อสอบแบบอิงเกณฑ์
 U คือ จำนวนผู้เรียนที่ตอบถูกของกลุ่มที่สอบผ่านเกณฑ์
 L คือ จำนวนผู้เรียนที่ตอบถูกของกลุ่มที่สอบไม่ผ่านเกณฑ์
 N_1 คือ จำนวนผู้เรียนที่สอบผ่านเกณฑ์
 N_2 คือ จำนวนผู้เรียนที่สอบไม่ผ่านเกณฑ์

การหาความยากของแบบทดสอบอิงเกณฑ์

พวงรัตน์ ทวีรัตน์ (2531 : 15) ค่าความยากของแบบทดสอบอิงเกณฑ์ จะพิจารณาจากจุดประสงค์เชิงพฤติกรรมที่ตั้งขึ้น ทั้งนี้เพราะจุดประสงค์ที่ตั้งขึ้นจะเป็นตัวกำหนดความยากของแบบทดสอบดังกล่าวคือ ถ้าจุดประสงค์นั้นวัดพฤติกรรมขั้นสูงก็จะมี ความยากมากกว่าจุดประสงค์ที่วัดพฤติกรรมขั้นต่ำ ดังนั้นความยากของข้อทดสอบอิงเกณฑ์จึงขึ้นอยู่กับพฤติกรรมที่ทำการวัด

บุญชม ศรีสะอาด (2535 : 87) ได้เสนอแนวคิดที่ว่า ความยากของข้อสอบอิงเกณฑ์ ตามวิธีของคริสปินและเฟลด์ลิวเซน (Krtspin and Feldlusion) เป็นค่าแสดงถึงร้อยละหรือสัดส่วนของผู้ที่ตอบข้อสอบนั้นถูก หรือที่เลือกคำตอบนั้น เขียนแทนด้วยสัญลักษณ์ p ระดับความยาก มีค่าตั้งแต่ 0 ถึง 100 หรือ .00 หรือ 1.00 (กรณีใช้ระบบสัดส่วน) ค่าของความยาก หรือ p ที่อยู่ในเกณฑ์เหมาะสมควรอยู่ระหว่าง 0.20 ถึง 0.80 ซึ่งไม่ยากเกินไปหรือง่ายเกินไป สำหรับการหาค่าระดับความยาก จำนวน โดยใช้สูตรดังนี้

$$P = \frac{R}{N}$$

เมื่อ P แทน ระดับความยาก

R แทน จำนวนผู้ตอบถูกทั้งหมด

N แทน จำนวนคนในกลุ่มสูง และกลุ่มต่ำ

ล้วน สายยศ และอังคณา สายยศ (2539 : 196) ได้เสนอแนวคิดที่ว่าข้อสอบอิงเกณฑ์ในแต่ละข้อต้องมีค่าความยากน้อยกว่า .40 ก่อนที่นักเรียนจะได้รับการสอนและเมื่อนักเรียนได้รับการสอนแล้วข้อสอบแต่ละข้อจะต้องมีค่าความยากมากกว่า .75 ทั้งนี้เป็นเพราะว่าข้อสอบอิงเกณฑ์เน้นที่คุณภาพในการสอนของครู กล่าวคือ ถ้าครูยังไม่ได้สอนเนื้อหา นั้นข้อสอบความยาก คือ มีค่า p ต่ำกว่า .40 แต่ถ้าครูทำการสอนแล้ว และครูสอนดีนักเรียนควรจะเรียนรู้ในเนื้อหา นั้น คือควรจะทำข้อสอบนั้น ได้ ซึ่งข้อสอบควรง่ายคือ มีค่า p มากกว่า .75

พิจิต ฤทธิจรูญ (2551 :138) กล่าวถึงความยาก ดังนี้ ความยาก (Difficulty) เป็นคุณสมบัติของข้อสอบที่บอกให้ทราบว่าข้อสอบนั้นมีคนตอบถูกมากหรือน้อย ถ้ามีคนตอบถูกมากข้อสอบข้อนั้นก็ง่าย ถ้ามีคนตอบถูกน้อยข้อสอบข้อนั้นก็ยาก ถ้ามีคนตอบถูกบ้างผิดบ้างหรือมีคนตอบถูกปานกลางข้อสอบข้อนั้นก็มีความยากปานกลาง ข้อสอบที่ดีมีความยากพอเหมาะมีคนตอบถูกไม่ต่ำกว่า 20 คนและไม่เกิน 80 คนจากผู้สอบ 100 คน ค่าความยากหาได้โดยการนำจำนวนนักเรียนที่ตอบถูกหารด้วยจำนวนคนที่ตอบถูกทั้งหมด

ไพศาล วรคำ (2552 : 287) ได้เสนอแนวคิดที่ว่า ความยากของข้อสอบ (Item Difficulty) เป็นคุณลักษณะประจำตัวของข้อสอบแต่ละข้อ ที่บ่งบอกถึงโอกาสที่กลุ่มตัวอย่างจะตอบข้อนั้น ได้ถูก ดังนั้นความยากของข้อสอบจึงพิจารณาได้จากจำนวนผู้ตอบข้อนั้นถูก ถ้ามีจำนวนผู้ตอบถูกมากแสดงว่าข้อสอบนั้นง่าย หรือมีค่าดัชนีความยาก (Item Difficulty Index : p) สูง ถ้ามีจำนวนผู้ตอบถูกน้อยแสดงว่าข้อสอบนั้นยากหรือมีค่าดัชนีความยากต่ำ ข้อสอบที่มีความยากเหมาะสมจะมีดัชนีความยากอยู่ระหว่าง 0.20 - 0.80 ส่วนในการสอบแบบอิงเกณฑ์

นั้น ต้องพิจารณาความรอบรู้ (ผ่านเกณฑ์) หรือไม่รอบรู้ (ไม่ผ่านเกณฑ์) การหาดัชนีความยากในการสอบแบบอิงเกณฑ์จึงเป็นการหาเพื่อให้ทราบระดับความยากเท่านั้น ซึ่งถ้ามีการหาดัชนีความยากในการสอบแบบอิงเกณฑ์ ก็มักจะหาทั้งดัชนีความยากก่อนเรียนและดัชนีความยากหลังเรียน

จากข้อเสนอแนะของนักการศึกษาหลายท่าน จะเห็นว่าความยากของแบบทดสอบ อิงเกณฑ์เป็นสิ่งที่ถูกกำหนดขึ้นจากตัวชี้วัด หรือจากการสอนของครู และการศึกษาค้นคว้าในครั้งนี้ ผู้วิจัยหาความยากของแบบทดสอบอิงเกณฑ์ โดยวิธีของครีสปิน และเฟลด์สูลเซน (Krtspin and Feldluson) (บุญชม ศรีสะอาด. 2535 : 87)

การหาอำนาจจำแนกของแบบทดสอบอิงเกณฑ์

ถ้วน สายยศ และอังคณา สายยศ (2539 : 197) เห็นว่า ค่าอำนาจจำแนกของข้อสอบ อิงเกณฑ์นั้นจะเป็นค่าอำนาจจำแนกระหว่างกลุ่มที่ยังไม่ได้รับการเรียนรู้หรือกลุ่มที่ยังไม่รู้ (Non Master) กับกลุ่มที่ได้รับการเรียนรู้แล้วหรือที่รู้แล้ว (Master) ข้อสอบอิงเกณฑ์ไม่เน้นที่อำนาจจำแนก เนื่องจากแบบทดสอบอิงเกณฑ์จะใช้ในการวัดผลที่ใช้ในการเรียนการสอนแบบ มีระบบ (Systematic Instruction) เช่นการเรียนการสอนรอบรู้ซึ่งการสอนนี้จะเน้นที่คุณภาพการสอนของครู ถ้าครูสอนดีเด็กทุกคนจะเรียนรู้หมด นั่นคือมีคะแนนเต็มหรือใกล้เคียงเต็มทุกคนเมื่อหาค่าอำนาจจำแนกจะมีค่าเป็น 0 หรือเข้าใกล้ 0 ดังนั้นค่าอำนาจจำแนกของข้อสอบอิงเกณฑ์ ควรมีค่าเท่ากับหรือมากกว่า 0 การคำนวณหาค่าอำนาจจำแนกของแบบทดสอบอิงเกณฑ์ มีอยู่ 2 วิธี ดังนี้

วิธีที่ 1 การคำนวณจากผลการทดสอบซ้ำกัน ใช้สูตร

$$PPDI = \frac{R_{pos}}{n} - \frac{R_{pre}}{n}$$

เมื่อ PPDI แทน ดัชนีค่าอำนาจจำแนกของข้อสอบอิงเกณฑ์

R_{pos} แทน จำนวนนักเรียนที่ทำข้อสอบถูกหลังได้รับการสอน

R_{pre} แทน จำนวนนักเรียนที่ทำข้อสอบถูกก่อนได้รับการสอน

n แทน จำนวนนักเรียนที่ทำข้อสอบทั้งหมด

ถ้าดัชนีค่าอำนาจจำแนกมีค่าเป็นบวกแสดงว่าข้อสอบสามารถจำแนกระหว่างกลุ่ม ที่เรียนรู้แล้วกับกลุ่มที่ยังไม่ได้เรียนรู้

ถ้าดัชนีค่าอำนาจจำแนกมีค่าเป็นลบแสดงว่าข้อสอบมีลักษณะข้อสอบถามได้ไม่ชัดเจน ทำให้นักเรียนที่ยังไม่ได้เรียนตอบถูกมากกว่านักเรียนที่เรียนแล้ว ข้อสอบยากมาก

นักเรียนจึงมีวิธีการเดาตอบทั้งก่อนเรียนและหลังเรียน และการสอนของครูแต่ละชั้นตอน
 สักสนจนนักเรียนไม่สามารถเรียนรู้ได้

วิธีการที่ 2 การคำนวณผลจากการทดสอบเพียงครั้งเดียววิธีนี้อาแบบทดสอบ
 ไปสอบนักเรียนหลังจากนักเรียนเรียนรู้แล้ว แล้วนำข้อมูลมาคำนวณจากสัดส่วนของนักเรียนที่
 ทำข้อสอบถูกของกลุ่มที่สอบผ่านเกณฑ์กับกลุ่มที่ยังไม่ผ่านเกณฑ์ การคำนวณวิธีนี้ต้องรู้
 คะแนนเกณฑ์ หรือคะแนนจุดตัดของแบบทดสอบก่อนแล้ว ใช้สูตรของ Brennan ดังนี้

$$B = \frac{U}{n_1} - \frac{L}{n_2}$$

เมื่อ	B	แทน	ดัชนีค่าอำนาจจำแนกของแบบทดสอบอิงเกณฑ์
	U	แทน	จำนวนนักเรียนที่ทำข้อสอบถูกของกลุ่มที่สอบผ่านเกณฑ์
	L	แทน	จำนวนนักเรียนที่ทำข้อสอบถูกของกลุ่มที่สอบไม่ผ่านเกณฑ์
	n_1	แทน	จำนวนนักเรียนที่สอบผ่านเกณฑ์
	n_2	แทน	จำนวนนักเรียนที่สอบไม่ผ่านเกณฑ์

บุญชม ศรีสะอาด (2543 : 53) ได้เสนอแนวคิดเห็นว่า อำนาจจำแนกของ
 ข้อสอบอิงเกณฑ์ตามวิธีของ Brennan หมายถึง ประสิทธิภาพในการจำแนกผู้สอบออกเป็นผู้
 ครอบรู้หรือสอบผ่าน กับผู้ที่ไม่ครอบรู้หรือสอบไม่ผ่าน เขียนแทนด้วยสัญลักษณ์ “B” จำนวน
 โดยใช้สูตร ดังนี้

$$B = \frac{U}{n_1} - \frac{L}{n_2}$$

เมื่อ	B	แทน	อำนาจจำแนก
	U	แทน	จำนวนผู้ครอบรู้หรือสอบผ่านเกณฑ์ที่ตอบถูก
	L	แทน	จำนวนผู้ไม่ครอบรู้หรือสอบไม่ผ่านที่ตอบถูก
	n_1	แทน	จำนวนผู้ครอบรู้หรือสอบผ่านเกณฑ์
	n_2	แทน	จำนวนผู้ไม่ครอบรู้หรือสอบไม่ผ่านเกณฑ์

ข้อสอบที่มีอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป เป็นข้อสอบที่เข้าเกณฑ์
 สามารถนำไปใช้ได้

สุรวาท ทองบุ (2550 : 102) การหาอำนาจจำแนกจากผลการสอบสองครั้ง
 (ก่อนเรียนและหลังเรียน) เป็นวิธีหาอำนาจจำแนกของข้อสอบที่เสนอ โดยคริสปินและเฟลด์ฮูเซน

(Krtspin and Feldluson) หรือเรียกว่าดัชนี S (S-Index หรือ Sensitivity Index) โดยใช้สูตรดังนี้

$$S = \frac{R_{pos} - R_{pre}}{N}$$

เมื่อ S แทน	อำนาจจำแนกของข้อสอบ
R_{pos} แทน	จำนวนคนที่ตอบถูกหลังสอน
R_{pre} แทน	จำนวนคนที่ตอบถูกก่อนสอน
N แทน	จำนวนผู้เข้าสอบทั้งหมด

ศิริชัย กาญจนวาสี (2552 : 225) เสนอแนวคิดว่า อำนาจจำแนก

(Discrimination) หมายถึงความสามารถของข้อสอบในการจำแนกความแตกต่างระหว่างผู้สอบที่มีผลสัมฤทธิ์ต่างกัน เช่น จำแนกคนเก่งออกจากคนอ่อนได้ โดยถือว่าคนเก่งมีความสามารถควรทำข้อสอบนั้นได้ ส่วน ผู้อ่อนหรือไม่มีความสามารถไม่ควรทำข้อสอบนั้นได้ อำนาจจำแนกของข้อสอบจะมีค่าตั้งแต่ -1 ถึง +1 แต่อำนาจจำแนกที่ดีจะต้องมีค่าเป็นบวกตั้งแต่ 0.20 ขึ้นไป

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยจะหาอำนาจจำแนกของข้อสอบอิงเกณฑ์ที่สร้างขึ้นตามวิธีของคริสปินและเฟลด์ลูเซน (Krtspin and Feldluson) (สุรวาท ทองบุ, 2550 : 102) โดยการหาดัชนี S เพราะเป็นการพิจารณาจากผลต่างของคะแนนก่อนเรียนและหลังเรียน ซึ่งสอดคล้องกับการปรับปรุงการเรียนเพื่อรอบรู้ที่ใช้แบบทดสอบอิงเกณฑ์เป็นเครื่องมือ

การหาความเที่ยงตรงของข้อสอบอิงเกณฑ์

การหาความเที่ยงตรงของเครื่องมือ เป็นการตรวจสอบว่า เครื่องมือนั้นสามารถวัดได้ตรงตามเกณฑ์ที่ต้องการหรือไม่ เป็นคุณภาพที่จะทำให้ผลการวัดที่ได้สามารถแทนคุณลักษณะที่ต้องการจะวัดได้มากน้อยเพียงใด ได้มีนักการศึกษาได้เสนอแนวความคิดเกี่ยวกับความเที่ยงตรงของข้อสอบอิงเกณฑ์ไว้ดังนี้

เบอร์ก (Berk, 1976 : 4) กล่าวถึงความหมายของความเที่ยงตรงของแบบทดสอบอิงเกณฑ์ว่า หมายถึง ความสอดคล้องระหว่างพฤติกรรมที่วัดได้โดยรายข้อต่าง ๆ ในแบบทดสอบกับจุดมุ่งหมายที่แบบทดสอบถูกกำหนดให้วัด

บุญชม ศรีสะอาด นิภาศิริไพโรจน์ และนุชวมา ทองทวี (2528 : 192) ได้ให้ความหมายของความเที่ยงตรงเชิงโครงสร้างว่า หมายถึงคุณภาพของแบบทดสอบที่วัดโครงสร้างหรือคุณลักษณะ (Trait) ตามทฤษฎีซึ่งเป็นการวัดสภาพปัจจุบันมากกว่าการวัดสภาพอดีตหรืออนาคต ดังนั้น แบบทดสอบที่มีความเที่ยงตรงเชิงโครงสร้างก็คือ แบบทดสอบ

ที่สามารถวัดได้ตรงตามพฤติกรรมที่ต้องการนั่นเอง

นิภา ศรีไพโรจน์ (2528 : 1) ได้เสนอวิธีหาความเที่ยงตรงเชิงเนื้อหาตามวิธีของ โรวินเนลลี และแฮมเบลตัน ในเชิงปฏิบัติ ดังนี้

1. นำจุดประสงค์เชิงพฤติกรรมและข้อสอบที่วัดจุดประสงค์นั้น ๆ ไปให้ผู้เชี่ยวชาญเนื้อหาวิชา หรือครูผู้สอนแต่ละคนพิจารณาว่าข้อสอบเหมาะสมกับจุดประสงค์เชิงพฤติกรรมมากน้อยเพียงใด โดยอาศัยมาตราส่วนประมาณค่า ดังนี้

- 4 หมายถึง ข้อสอบเหมาะสมกับจุดประสงค์เชิงพฤติกรรมมากที่สุด
- 3 หมายถึง ข้อสอบเหมาะสมกับจุดประสงค์เชิงพฤติกรรมมาก
- 2 หมายถึง ข้อสอบเหมาะสมกับจุดประสงค์เชิงพฤติกรรมปานกลาง
- 1 หมายถึง ข้อสอบเหมาะสมกับจุดประสงค์เชิงพฤติกรรมน้อย
- 0 หมายถึง ข้อสอบเหมาะสมน้อยที่สุดหรือไม่เหมาะสมเลยกับ

จุดประสงค์เชิงพฤติกรรม

2. บันทึกผลการลงความคิดเห็นของผู้เชี่ยวชาญเนื้อหาวิชา หรือครูผู้สอนแต่ละคน เป็นรายข้อ แล้วคำนวณหาค่าเฉลี่ย (\bar{X}) และค่าความเบี่ยงเบนมาตรฐาน (S.D.)

3. แปลความหมายค่าเฉลี่ยของคะแนนความคิดเห็น และค่าของส่วนเบี่ยงเบนมาตรฐาน

4. คัดเลือกข้อสอบที่มีค่าเฉลี่ย (\bar{X}) ตั้งแต่ 2.5 - 4.00 และค่าของส่วนเบี่ยงเบนมาตรฐาน (S.D.) ไม่เกิน 1.0

ล้วน สายยศ และอังคณา สายยศ (2539 : 246 - 255) ได้ให้ความหมายของแบบทดสอบ อิงเกณฑ์เกี่ยวกับความเที่ยงตรงเชิงเนื้อหาว่า เป็นเครื่องมือที่สามารถวัดได้ตามเนื้อหา หรือวัดได้ตามจุดประสงค์ที่ต้องการวัด และความเที่ยงตรงตามโครงสร้างหมายถึงแบบทดสอบอิงเกณฑ์ ที่สามารถวัดได้ตามลักษณะหรือตามทฤษฎีต่าง ๆ ของโครงสร้างนั้น สำหรับวิธีคำนวณค่า ความเที่ยงตรงของแบบทดสอบอิงเกณฑ์ มีวิธีหาได้ดังนี้

1. ความเที่ยงตรงเชิงเนื้อหาเป็นความเที่ยงตรงที่ให้ผู้เชี่ยวชาญพิจารณาว่าข้อสอบแต่ละข้อนั้นวัดได้ตรงตามจุดประสงค์นั้นหรือไม่ โดยใช้สูตรของ Rovinelli and Hamblton ซึ่งคำนวณจากค่าดัชนีความสอดคล้อง (IOC) ดังนี้

$$IOC = \frac{\sum R}{N}$$

เมื่อ	IOC	แทน	ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์
	$\sum R$	แทน	สรุปคะแนนความคิดเห็นของผู้เชี่ยวชาญเนื้อหาวิชา
	N	แทน	จำนวนผู้เชี่ยวชาญ

2. ความเที่ยงตรงตามเกณฑ์สัมพัทธ์ หมายถึง คุณภาพของเครื่องมือที่เอาผลการวัดแบบทดสอบไปหาความสัมพันธ์กับเกณฑ์ที่ต้องการ จำแนกเป็น 2 ชนิด ดังนี้

2.1 ความเที่ยงตรงเชิงสภาพ หมายถึงแบบทดสอบอิงเกณฑ์ที่สามารถวัดความรู้ (Master) และไม่รู้ (Non – Master) ของนักเรียนในการทดสอบในแต่ละจุดประสงค์ว่าตรงกับสภาพจริงหรือไม่ ถ้าตรงกับสภาพจริงก็แสดงว่ามีความเที่ยงตรงสูง

2.2 ความเที่ยงตรงเชิงพยากรณ์ หมายถึง ความเที่ยงตรงที่ได้จากการเอาผลการวัดของแบบทดสอบไปพยากรณ์ผลความสำเร็จในอนาคต ซึ่งความเที่ยงตรงเชิงพยากรณ์ ของแบบทดสอบอิงเกณฑ์ เป็นการหาค่าความเที่ยงตรงในลักษณะค่าความสอดคล้องของความรู้กับไม่รู้

3. ความเที่ยงตรงตาม โครงสร้าง หมายถึง คุณภาพของเครื่องมือที่สามารถวัดได้ ตรงตามลักษณะหรือตามทฤษฎีต่าง ๆ ของโครงสร้างนั้นหรือวัดได้ครอบคลุมตามลักษณะของโครงสร้างของแบบทดสอบอิงเกณฑ์

บุญชม ศรีสะอาด (2543 : 65 – 66) ได้เสนอแนวคิดเกี่ยวกับการหาความเที่ยงตรงของแบบทดสอบอิงเกณฑ์ว่า ควรพิจารณาความเที่ยงตรง ดังนี้

1. ความเที่ยงตรงเชิงเนื้อหา (Content Validity) การสร้างแบบทดสอบจะต้องสร้างตามจุดประสงค์เชิงพฤติกรรม วิธีการตรวจสอบว่าแต่ละข้อมีความเที่ยงตรงเชิงเนื้อหาหรือไม่โดยนำเอาจุดประสงค์เชิงพฤติกรรมและข้อสอบไปให้ผู้เชี่ยวชาญตัดสิน การวิเคราะห์หาความเที่ยงตรงตามเนื้อหา และความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม โดยการนำไปให้ผู้เชี่ยวชาญพิจารณาให้คะแนนความสอดคล้องของข้อสอบแต่ละข้อกับจุดประสงค์ เชิงพฤติกรรม แล้วนำมาหาค่าดัชนีความสอดคล้อง (IOC) ตามวิธีของ Rovinelli and Hambelton (บุญชม ศรีสะอาด. 2543 : 65) ดังนี้

$$IOC = \frac{\sum R}{N}$$

เมื่อ	IOC	แทน	ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์
	$\sum R$	แทน	สรุปคะแนนความคิดเห็นของผู้เชี่ยวชาญเนื้อหาวิชา
	N	แทน	จำนวนผู้เชี่ยวชาญ

- การกำหนดคะแนนของผู้เชี่ยวชาญอาจจะเป็น +1 หรือ 0 หรือ -1 ดังนี้
- +1 = แน่ใจว่าข้อสอบข้อนั้นวัดจุดประสงค์เชิงพฤติกรรมที่ระบุไว้จริง
 - 0 = ไม่แน่ใจว่าข้อสอบข้อนั้นวัดจุดประสงค์เชิงพฤติกรรมที่ระบุไว้
 - 1 = แน่ใจว่าข้อสอบข้อนั้นไม่ได้วัดจุดประสงค์เชิงพฤติกรรมที่ระบุไว้
- ค่าดัชนีความสอดคล้องที่ยอมรับได้ต้องมีค่าตั้งแต่ 0.50 ขึ้นไป

2. ความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) หมายถึง

ประสิทธิภาพของแบบทดสอบในการจำแนก ผู้ที่เรียนแล้วกับผู้ที่ยังไม่ได้เรียน โดยถือแนวคิดที่ว่า ผู้ที่เรียนแล้วน่าจะสอบผ่าน ผู้ที่ยังไม่ได้เรียนน่าจะสอบไม่ผ่าน จากแนวคิดดังกล่าว แบบทดสอบฉบับใดที่นำไปสอบแล้วพบว่าผู้ที่เรียนเรื่องนั้นแล้วสอบผ่านจำนวนมาก และผู้ที่ยังไม่ได้เรียนเรื่องนั้นสอบผ่านจำนวนน้อยก็จะมีค่าความเที่ยงตรงสูง ถ้าทุกคนที่เรียนแล้วสอบผ่านหมด และทุกคนที่ยังไม่ได้เรียนสอบไม่ผ่าน แบบทดสอบฉบับนี้จะมีค่าความเที่ยงตรงอย่างสมบูรณ์ ค่าความเที่ยงตรงจะเท่ากับ 1.00 การหาความเที่ยงตรงเชิงโครงสร้างตามวิธีของคาร์เวอร์ (Carver) (บุญชม ศรีสะอาด. 2543 : 66) มีสูตร ดังนี้

$$R_c = \frac{a + c}{n}$$

เมื่อ R_c	หมายถึง	ค่าความเที่ยงตรง
a	หมายถึง	จำนวนที่ผู้เรียนแล้วสอบผ่าน
c	หมายถึง	จำนวนผู้ที่ยังไม่ได้เรียนและสอบไม่ผ่าน
n	หมายถึง	จำนวนผู้สอบทั้งหมด

(จำนวนผู้ที่ยังไม่ได้เรียนรวมกับจำนวนผู้เรียนแล้ว)

ศิริชัย กาญจนวาสี (2552 : 118) กล่าวถึงความเที่ยงตรงเชิงโครงสร้างว่า ในกรณีที่ต้องการแปรผลคะแนนสอบ เพื่อสรุปอ้างอิงถึงลักษณะทั่วไปของบุคคลที่สนใจอาจเป็นลักษณะทางจิตวิทยา เช่น เชาวน์ปัญญา ความคิดสร้างสรรค์ ทักษะคิด ความสามารถ หรือคุณภาพด้านต่าง ๆ ของบุคคลลักษณะเหล่านี้ ถือว่าเป็นโครงสร้างความคิดหรือภาวะสันนิษฐาน (Construct) ซึ่งเป็นลักษณะภายในและเป็นนามธรรมไม่สามารถสังเกตได้โดยตรง เมื่อนักวัดผลทำการวัดลักษณะภายในใดก็ตามแสดงให้เห็นถึงความเชื่อมโยงบนพื้นฐานของข้อตกลงเบื้องต้นว่า ลักษณะนั้นมีอยู่จริง ลักษณะนั้นมีความแตกต่างจากลักษณะอื่น ๆ และลักษณะนั้นมีอิทธิพลต่อพฤติกรรมที่ปรากฏ หรือคะแนนที่ได้จากการสอบ แบบทดสอบที่นำมาใช้วัดลักษณะเหล่านี้จึงต้องมีความเที่ยงตรง เชิงโครงสร้าง

จากการศึกษาเกี่ยวกับความเที่ยงตรงของแบบทดสอบ ผู้วิจัยได้คำนวณหาความเที่ยงตรงเชิงเนื้อหาตามวิธีของ โรวินลลีและแฮมเบลตัน (Rovinelli and Hambelton) (บุญชม ศรีสะอาด. 2543 : 65) หาความเที่ยงตรงเชิงโครงสร้างตามวิธีของคาร์เวอร์ (Carver) (บุญชม ศรีสะอาด. 2543 : 66)

การกำหนดคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์

เกณฑ์การสอบหรือคะแนนจุดตัด (Cut – off scores) มีชื่อเรียกต่าง ๆ กัน เช่น เกณฑ์ (Criteria) คะแนนมาตรฐาน (Standard score) คะแนนจุดผ่าน (Passing score) ระดับความรู้ (Master level) ความสามารถขั้นต่ำ (Minimal competence) โดยมีนักการศึกษา กล่าวไว้ ดังนี้

บุญชม ศรีสะอาด (2538 : 148) ได้ให้ความหมายของคะแนนจุดตัดหรือเกณฑ์ว่าเป็นค่าคะแนนที่น้อยที่สุดที่จะยอมรับว่าผู้สอบแต่ละคนเป็นผู้รอบรู้ในเนื้อหาเรื่องนั้น ผู้ที่ทำข้อสอบได้คะแนนรวมเท่ากับจุดตัดหรือสูงกว่าถือว่าเป็นผู้รอบรู้ ในเนื้อหานั้น ถ้าได้คะแนนต่ำกว่าจุดตัด ถือว่าเป็นผู้ไม่รอบรู้

ล้วน สายยศ และอังคณา สายยศ (2539 : 267) ได้ให้ความหมายของคะแนนจุดตัดว่าเป็นคะแนนที่ใช้สำหรับเป็นเกณฑ์ในการนำผลการสอบ ไปเปรียบเทียบกับว่านักเรียนมีความรอบรู้ (Master) หรือไม่รอบรู้ (Monmaster) และการกำหนดจุดตัดของแต่ละครั้งจะทำให้เกิดการตัดสินผู้สอบออกเป็น 4 ลักษณะ ดังตารางที่ 2

ตารางที่ 2 การตัดสินผู้สอบของแบบทดสอบอิงเกณฑ์

สภาพจริง (True Status)	รอบรู้	ผลการสอบ	
		ไม่รอบรู้	รอบรู้
	รอบรู้	ความคลาดเคลื่อนแบบไม่ยอมรับ	การตัดสินที่ถูกต้อง
	ไม่รอบรู้	การตัดสินที่ถูกต้อง	ความคลาดเคลื่อนแบบยอมรับ

จากตารางที่ 2 คะแนนจุดตัดจะมีลักษณะที่เกิดจากการกำหนดคะแนน 4 ลักษณะด้วยกัน ดังนี้

1. ลักษณะที่เกิดจากการทำแบบทดสอบอิงเกณฑ์แล้วตัดสินว่านักเรียนสอบไม่ผ่านทั้ง ๆ ที่ตามสภาพจริงแล้วเป็นผู้ที่มีความรู้ จึงเกิดความคลาดเคลื่อนขึ้น เรียกว่า

ความคลาดเคลื่อนแบบไม่ยอมรับ (Error of rejection) หรือความผิดพลาดแบบลบ (False negative) คือเป็นผู้ที่ไม่รอบรู้แบบไม่จริง

2. ลักษณะที่เกิดจากการทำแบบทดสอบอิงเกณฑ์แล้วตัดสินว่านักเรียนสอบผ่านทั้ง ๆ ที่สภาพจริงแล้วเป็นผู้ไม่มีความรู้ จึงเกิดความคลาดเคลื่อนขึ้น เรียกว่า ความคลาดเคลื่อนแบบยอมรับ (Error of acceptance) หรือความผิดพลาดแบบบวก (False positive) คือเป็นผู้รอบรู้จริง

3. ลักษณะที่เกิดจากการทำแบบทดสอบอิงเกณฑ์แล้วตัดสินว่านักเรียนสอบผ่านและสภาพจริงเป็นผู้มีความรู้ แสดงว่าการตัดสินผลการสอบครั้งนี้ถูกต้อง (Correct Decision) ไม่มีความคลาดเคลื่อนเกิดขึ้น

4. ลักษณะที่เกิดจากการทำแบบทดสอบอิงเกณฑ์แล้วตัดสินว่านักเรียนสอบไม่ผ่านทั้ง ๆ ที่สภาพจริงเป็นผู้ไม่มีความรู้ แสดงว่าการตัดสินผลการสอบครั้งนี้ถูกต้อง (Correct Decision) ไม่มีความคลาดเคลื่อนเกิดขึ้น

ในการกำหนดจุดตัดที่เหมาะสมที่สุดคือ จุดที่ทำให้เกิดผลการตัดสินแบบที่ 1 กับแบบที่ 2 (ความผิดพลาดแบบลบกับความผิดพลาดแบบบวก) มีค่าน้อยที่สุดหรือมีค่าเป็นศูนย์นั้นก็คือต้องการตัดสินที่ถูกต้องแบบที่ 3 และแบบที่ 4

การกำหนดเกณฑ์การสอบหรือคะแนนจุดตัด (Standard – setting method) (ล้วน สายยศ และอังคณา สายยศ. 2539 : 268 - 295) ได้กล่าวถึงวิธีกำหนดคะแนนจุดตัดไว้หลายวิธี ดังนี้

1. วิธีกำหนดเกณฑ์การสอบหรือคะแนนจุดตัด โดยวิธีการพิจารณาคัดเลือกจากเนื้อหาวิชาและข้อสอบแต่ละข้อแล้วคำนวณหาจุดตัด โดยมีผู้เสนอไว้หลายท่าน เช่น

1.1 วิธีของ Nedelsky (ล้วน สายยศ และอังคณา สายยศ. 2539 : 268-295) เป็นวิธี ที่กำหนดคะแนนจุดตัดของแบบทดสอบเลือกตอบโดยวิธีดังนี้

1.1.1 ให้ผู้เชี่ยวชาญในเนื้อหาวิชาเป็นผู้พิจารณาตัวเลือกของแบบทดสอบ เลือกตอบแต่ละข้อว่าตัวเลือกข้อใด ที่คิดว่านักเรียนที่มีความสามารถต่ำที่สุด (Lowest student) จะไม่เลือกตอบ

1.1.2 นำตัวเลือกที่เหลือมาหาค่าความน่าจะเป็นที่นักเรียนจะเลือกตอบ

1.1.3 คำนวณผลรวมของความน่าจะเป็นของแต่ละข้อของแบบทดสอบใช้สัญลักษณ์ว่า M

1.1.4 เอาค่า M ของผู้เชี่ยวชาญแต่ละคนมาคำนวณค่าเฉลี่ยใช้สัญลักษณ์ φ_M และหาค่าคะแนนเฉลี่ยเบี่ยงเบนมาตรฐานของ M ใช้สัญลักษณ์ φ_M แล้วคำนวณหาคะแนนจุดตัดจากสูตร

$$\text{คะแนนจุดตัด } C_x = \mu_M + K\sigma_M$$

เมื่อ K คือค่าคงตัวที่มีค่า -1 , 0 , 1 และ 2 เมื่อให้นักเรียนที่มีความรู้ต่ำสุดมีโอกาสตก 16% , 50% , 84% , และ 98% ตามลำดับ ซึ่งกำหนดโดยผู้เชี่ยวชาญเนื้อหาวิชา โดยทั่วไป K มีค่าอยู่ระหว่าง 0.5 ถึง 1.0

1.2 วิธีของ Angoff (ล้วน สายยศ และอังคณา สายยศ. 2539 : 270) เป็นวิธีที่กำหนดคะแนนจุดตัดของผู้เชี่ยวชาญในการสอนวิชานั้น โดยพิจารณาข้อสอบแต่ละข้อว่าผู้ที่มีความรู้มีค่าความน่าจะเป็น (โอกาสที่จะตอบถูก) ในการตอบถูกข้อนั้นอย่างน้อยเท่าไร แล้วหาค่าเฉลี่ยของผู้เชี่ยวชาญเหล่านั้นเป็นคะแนนจุดตัด

1.3 วิธีของ Ebel's technique (ล้วน สายยศ และอังคณา สายยศ. 2539 : 271) วิธีนี้เป็นการใช้การพิจารณาจากลักษณะความยากง่าย และความเกี่ยวข้องในเนื้อหาของแบบทดสอบอิงเกณฑ์เป็นหลักการพิจารณาความสำเร็จที่คาดหวังไว้ในข้อสอบ ซึ่งสรุปเป็นขั้นตอนในการหาคะแนนจุดตัด ดังนี้

1.3.1 นำข้อสอบมาให้ผู้เชี่ยวชาญพิจารณาลักษณะข้อสอบ เช่น ความจำเป็น ความสำคัญ การยอมรับและยังเป็นปัญหา

1.3.2 กำหนดระดับความยากง่าย ซึ่งมีอยู่ 3 ระดับ คือ ง่าย ปานกลาง ยาก

1.3.3 กำหนดเปอร์เซ็นต์การสอบผ่านตามลักษณะของข้อสอบและความยากง่ายของข้อสอบ

1.3.4 นำแบบทดสอบแต่ละข้อมาแจกแจงลักษณะของสิ่งที่เกี่ยวข้องกับเนื้อหา

1.3.5 เอาความถี่ของแต่ละข้อคูณกับเปอร์เซ็นต์ความสำเร็จที่คาดหวัง

1.3.6 คำนวณหาคะแนนจุดตัดโดยใช้สูตร

$$\text{คะแนนจุดตัด} = \frac{\text{ผลรวมทั้งหมดของผลคูณระหว่างจำนวนข้อกับความสำเร็จที่คาดหวังไว้}}{\text{ผลรวมจำนวนข้อของผู้เชี่ยวชาญทั้งหมด}}$$

2. วิธีการกำหนดคะแนนจุดตัดโดยวิธีเชิงประจักษ์ (Empirical Methods) วิธีนี้เป็นวิธีการกำหนดคะแนนจุดตัดด้วยคะแนนจากการทดสอบของนักเรียน ซึ่งมีอยู่หลายวิธีด้วยกัน เช่น วิธีของ Livingston วิธีทฤษฎีการตัดสินใจของ Glass วิธีของ Huynh วิธีหาความคลาดเคลื่อนที่เกิดจากการเดาตอบและการสุ่มข้อสอบ

2.1 วิธีทฤษฎีการตัดสินใจ (Decision-Theoretic Approach) ของ Glass เป็นวิธีการที่แบ่งนักเรียนออกเป็น 2 กลุ่ม โดยอาศัยเกณฑ์ภายนอก ซึ่งจะเป็นผลการเรียนโดยปกติของนักเรียนหรือผลสำเร็จในการทำงานแล้วแบ่งเป็นกลุ่มผู้ผ่านเกณฑ์ภายนอก (Pass) และกลุ่มไม่ผ่านเกณฑ์ภายนอก (Fail) ในแต่ละกลุ่มเมื่อทำแบบทดสอบอิงเกณฑ์ที่ต้องการหาคะแนนจุดตัดนั้นแล้วมีจำนวนคนที่ผ่านและไม่ผ่านเกณฑ์ที่กำหนดขึ้นในแต่ละจุดของคะแนนเกณฑ์เท่าไร ดังแสดงในตารางที่ 3 ต่อไปนี้

ตารางที่ 3 คะแนนเกณฑ์ที่กำหนดในแบบทดสอบอิงเกณฑ์โดยอาศัยเกณฑ์ภายนอก

คะแนนเกณฑ์ที่กำหนด ในแบบทดสอบอิง เกณฑ์	เกณฑ์แบบทดสอบ	เกณฑ์ภายนอก	
		ผ่าน	ไม่ผ่าน
	ไม่ผ่าน	P_A	P_B
	ผ่าน	P_C	P_D

จากตารางที่ 3 กำหนดให้

P_A แทน สัดส่วนของคนที่ไม่ผ่านการทดสอบแบบอิงเกณฑ์แต่ผ่านเกณฑ์ภายนอก

P_B แทน สัดส่วนของคนที่ไม่ผ่านทั้งการทดสอบแบบอิงเกณฑ์และเกณฑ์ภายนอก

P_C แทน สัดส่วนของคนที่ผ่านการทดสอบแบบอิงเกณฑ์ และเกณฑ์ภายนอก

P_D แทน สัดส่วนของคนที่ผ่านการทดสอบแบบอิงเกณฑ์ แต่ไม่ผ่านเกณฑ์ภายนอก

คะแนนจุดตัดของแบบทดสอบอิงเกณฑ์สามารถเปลี่ยนได้หลายค่า แต่เกณฑ์ภายนอกกำหนดนั้นจะไม่มีเปลี่ยนแปลง ดังนั้นคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ก็คือฟังก์ชันของคะแนนจุดตัด $f(C_x)$ มีค่าน้อยที่สุดจากสูตร ดังนี้

$$f(C_x) = \frac{P_A P_D}{P_B + P_C}$$

ในการคำนวณคะแนนจุดตัดด้วยสมการดังกล่าว มีโอกาสที่จะจำแนกผู้สอบผิดทางลบ (False negative : α) กับจำแนกผู้สอบผิดทางบวก (False positive : β) มีค่าเท่ากัน ถ้า

พิสูจน์ได้ว่าโอกาสที่จำแนกผู้สอบผิดทางลบกับจำแนกผู้สอบผิดทางบวกมีค่าไม่เท่ากัน จะต้องคำนวณคะแนนจุดตัดจากค่าฟังก์ชันที่ปรับแก้แล้วคือ

$$f(C_x) = \frac{\alpha P_A \beta P_D}{P_B + P_C}$$

โดยกำหนดให้ค่าโอกาสที่จำแนกผิดทางลบ คือ α และโอกาสที่จำแนกผิดทางบวกคือ β มีค่าอยู่ระหว่าง 0 ถึง 1 และจะมีค่าเท่าไรหรือนั้นขึ้นอยู่กับผู้ประเมินผลการสอบจะต้องคำนึงถึงความสำคัญสองประการ คือ

1. นักเรียนสอบผ่านเกณฑ์แบบทดสอบอิงเกณฑ์ แต่สอบไม่ผ่านภายนอก หรือสอบตก หรือเรียนไม่สำเร็จควรให้ความสำคัญเท่าไร เป็นตัวกำหนด α

2. นักเรียนที่สอบไม่ผ่านเกณฑ์ของแบบทดสอบอิงเกณฑ์ แต่ไม่สามารถสอบผ่านเกณฑ์ภายนอกหรือสามารถเรียนสำเร็จควรให้ความสำคัญเท่าไร เป็นตัวกำหนด β โดยทั่วไปแล้วในทางปฏิบัติการคำนวณหาคะแนนจุดตัดโดยวิธีทฤษฎีการตัดสินใจนี้มักจะให้ค่าการจำแนกผิดทางลบ α กับค่าการจำแนกผิดทางบวก β มีค่าเท่ากัน

2.2 โดยวิธีของ Huynh (ถ้วน สายยศ และอังคณา สายยศ. 2539 : 280) เป็นวิธีที่กำหนดคะแนนจุดตัดโดยการจำแนกออกเป็นสองกลุ่ม คือ กลุ่มที่เรียนแล้วกับกลุ่มที่ยังไม่ได้เรียน สำหรับกลุ่มที่เรียนแล้วหมายถึง กลุ่มที่สอบภายหลังจากได้รับการสอน และกลุ่มที่ยังไม่ได้เรียนหมายถึง กลุ่มที่สอบก่อนได้รับการสอนจากนั้นจะคำนวณหาค่า P ของคะแนนแต่ละตัวที่กำหนดเป็นเกณฑ์ของแบบทดสอบอิงเกณฑ์ ซึ่งค่า P จะแปรผันไปตามคะแนนที่กำหนดคะแนนจุดตัด จะเป็นคะแนนที่กำหนดค่า P ได้ค่าสูงสุด โดยคำนวณจากสูตร

$$P = \frac{1}{2} \left[\frac{n_{11}}{n_1} + \frac{n_{00}}{n_0} \right]$$

เมื่อ	n_{11}	แทน	จำนวนนักเรียนที่สอบผ่านเกณฑ์หลังเรียน
	n_{00}	แทน	จำนวนนักเรียนที่สอบไม่ผ่านก่อนเรียน
	n_1	แทน	จำนวนนักเรียนที่สอบทั้งหมดหลังเรียน
	n_0	แทน	จำนวนนักเรียนที่สอบทั้งหมดก่อนเรียน

3. วิธีกำหนดคะแนนจุดตัดแบบวิธีผสม วิธีนี้เป็นการกำหนดคะแนนจุดตัดที่มีทั้งวิธีพิจารณาคุณยพินิจและวิธีเชิงประจักษ์ (Judgment – Empirical) ซึ่งมีอยู่หลายวิธีด้วยกัน เช่น วิธีกลุ่มตรงข้ามของ Livingston วิธีของเบอร์ก (Berk) วิธีของ Bayesian วิธีของ

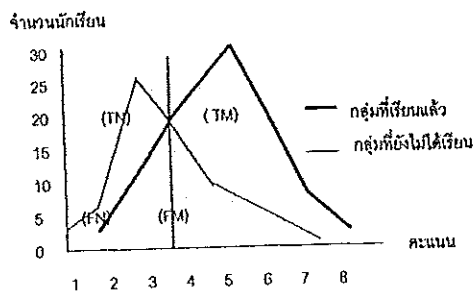
แฮมเบิลตัน (Hambleton)

3.1 วิธีของเบอร์ก (Berk) (ล้วน สายยศ และอังคณา สายยศ. 2539 : 285) เป็นวิธีที่พิจารณาคะแนนเกณฑ์ของแบบทดสอบก่อน โดยนำเอาคะแนนแจกแจงความถี่ระหว่างกลุ่ม ที่ยังไม่ได้เรียนและกลุ่มที่เรียนเนื้อหาแล้ว มาสร้างภาพเพื่อพิจารณาจุดตัดของกราฟแล้ว กำหนดเกณฑ์แบบทดสอบเป็นเกณฑ์พยากรณ์ การกำหนดเกณฑ์คือการกำหนดจุดตัดของคะแนนที่แบ่งผู้เรียนออกเป็น 2 พวก คือ กลุ่มที่ได้รับการสอนให้เป็นพวกรอบรู้ (Master) กลุ่มที่ไม่ได้ รับการสอนเป็นพวกไม่รอบรู้ (Non Master) หลังจากให้นักเรียนทั้ง 2 กลุ่ม ทำแบบทดสอบแล้วพิจารณาการกระจายของคะแนน 2 กลุ่ม จะคาบเกี่ยวกับจุดที่ฟังก์ชันทั้งสองตัดกัน คือ คะแนนพยากรณ์ที่จะแบ่งการเรียนรู้เป็น 4 กลุ่ม ดังตารางที่ 4

ตารางที่ 4 การกำหนดเกณฑ์แบบทดสอบอิงเกณฑ์เป็นเกณฑ์พยากรณ์

คะแนนพยากรณ์		การจำแนกเกณฑ์	
		กลุ่มที่เรียน	กลุ่มที่ยังไม่เรียน
รอบรู้		รอบรู้จริง (TM)	รอบรู้ไม่จริง (FM)
ไม่รอบรู้		ไม่รอบรู้ไม่จริง (FN)	ไม่รอบรู้จริง (TN)

คะแนนจากการทดสอบก่อนได้รับการสอนกับหลังได้รับการสอน นำมาแจกแจงเป็นโค้งจะได้ลักษณะดังภาพที่ 1 ดังนี้



ภาพที่ 1 คะแนนจุดตัดที่ได้จากจุดตัดของการกระจายในกลุ่มที่เรียนแล้วกับกลุ่มที่ยังไม่ได้เรียน

ที่มา : (Berk. 1976 : 6)

คะแนนจุดตัดแต่ละคะแนนนำมาหาค่าความน่าจะเป็นของการตัดสินใจที่ต้อง จากสูตร $P(TM) + P(TN)$ และความคลาดเคลื่อนในการจำแนกผิดประเภทที่ 1 (ผู้ที่เรียนแล้วแต่ สอบไม่ผ่านเกณฑ์) คือ $P(FN)$ และความคลาดเคลื่อนในการจำแนกผิดประเภทที่ 2 (ผู้ที่ยังไม่ เรียนแต่สอบผ่านเกณฑ์) คือ $P(TN)$

$$\begin{aligned} \text{เมื่อ } P(TM) &= TM/(M+N) \\ P(TN) &= TN/(M+N) \\ P(FM) &= FM/(M+N) \\ P(FN) &= FN/(M+N) \\ N &\text{ แทน จำนวนนักเรียนทั้งหมดที่ยังไม่เรียน} \\ M &\text{ แทน จำนวนนักเรียนทั้งหมดที่เรียนแล้ว} \end{aligned}$$

คะแนนจุดตัดแต่ละคะแนนสามารถตรวจสอบความเที่ยงตรงโดยใช้สูตร สัมประสิทธิ์ความเที่ยงตรงของเกณฑ์ เพื่อเลือกค่าสัมประสิทธิ์สูงสุดของความน่าจะเป็นใน การตัดสินใจของการเลือกคะแนนจุดตัดดังกล่าวมาเป็นเกณฑ์ สูตรการหาค่าความเที่ยงตรง ของเกณฑ์ คือ

$$\phi_{vc} = \frac{P(TM) - BR(SR)}{BR(1 - BR)SR(1 - SR)}$$

$$\begin{aligned} \text{เมื่อ } \phi_{vc} &\text{ แทน ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์} \\ BR &\text{ แทน ความน่าจะเป็นของผู้รอบรู้ในประชากร} \\ SR &\text{ แทน ความน่าจะเป็นของการพยากรณ์ผู้รอบรู้ในประชากร} \\ BR &= P(FM) + P(TM) \\ SR &= P(TM) + P(FM) \end{aligned}$$

จากวิธีการกำหนดคะแนนจุดตัดที่กล่าวมาแล้ว พอสรุปได้ว่า คะแนนจุดตัดคือ คะแนนที่น้อยที่สุดที่ใช้ในการตัดสินใจเรียนว่าเป็นผู้รอบรู้หรือไม่รอบรู้ ในการวิจัยครั้งนี้ผู้วิจัย ได้เลือกวิธีการกำหนดจุดตัดโดยวิธีของเบอร์ก (Berk) (ลิวัน สายยศ และอังคณา สายยศ. 2539 : 285) เพราะเป็นวิธีที่เหมาะสมกับแบบทดสอบที่มีจำนวนข้อไม่มากนัก จึงเหมาะสมกับ แบบทดสอบที่สร้างขึ้น

การหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์

ค่าความเชื่อมั่นของแบบทดสอบเป็นองค์ประกอบที่สำคัญของการสร้างแบบทดสอบเพราะว่าค่าความเชื่อมั่นเป็นดัชนีบ่งชี้ว่าแบบทดสอบนั้นมีคุณภาพหรือไม่ ซึ่งได้มีนักการศึกษาหลายท่านได้เสนอแนวคิดเกี่ยวกับค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ได้ดังนี้

บุญชม ศรีสะอาด (2543 : 68) ได้กล่าวถึง การหาค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ (Reliability) มีหลายวิธีได้แก่

1. วิธีของคาร์เวอร์ (Carver) ความเชื่อมั่นตามวิธีนี้ หมายถึงความสอดคล้องในการจำแนกผู้รอบรู้หรือผู้สอบผ่านกับผู้ไม่รอบรู้หรือผู้สอบไม่ผ่าน คำนวณหาค่าความเชื่อมั่น จากสูตร

$$r_{cc} = \frac{a + c}{N}$$

เมื่อ r_{cc} หมายถึง ค่าความเชื่อมั่น

กรณีใช้แบบทดสอบคู่ขนาน

a หมายถึง จำนวนผู้สอบผ่านทั้งสองฉบับ

c หมายถึง จำนวนผู้สอบไม่ผ่านทั้งสองฉบับ

กรณีใช้วิธีสอบกลุ่มเดิม 2 ครั้ง

a หมายถึง จำนวนผู้สอบผ่านทั้งสองครั้ง

c หมายถึง จำนวนผู้สอบไม่ผ่านทั้งสองครั้ง

N หมายถึง จำนวนผู้สอบทั้งหมด

2. วิธีของโลเวทท์ (Lovett) เป็นวิธีหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์จากผลสอบครั้งเดียว (ใช้คะแนนหลังเรียน) โดยใช้สูตร ดังนี้

$$r_{cc} = 1 - \frac{k \sum x_i - \sum x_i^2}{(k-1) \sum (x_i - c)^2}$$

เมื่อ r_{cc} แทน ความเชื่อมั่นของแบบทดสอบ

k แทน จำนวนข้อสอบ

x_i แทน คะแนนของแต่ละคน

C แทน คะแนนเกณฑ์หรือจุดตัดของแบบทดสอบ

แบบทดสอบที่เข้าเกณฑ์ มีค่าความเชื่อมั่นของแบบทดสอบตั้งแต่ 0.5 ขึ้นไป

3. วิธีของคูเดอร์ - ริชาร์ดสัน (Kuder-Richardson - 21) ใช้สูตร KR - 21 มีสูตรดังนี้

$$r_{11} = \frac{N}{n-1} \left\{ \frac{1 - \bar{x}(n - \bar{x})}{n\sigma_x^2} \right\}$$

โดยที่ R_{11} คือ ค่าความเชื่อมั่นของเครื่องมือวัด

n คือ จำนวนข้อในเครื่องมือวัด

$n\sigma_x^2$ คือ ความแปรปรวนของคะแนนทั้งหมด

\bar{X} คือ ค่าเฉลี่ยของคะแนนทั้งหมด

เกณฑ์ความเชื่อมั่นที่ยอมรับได้ จะมีค่าตั้งแต่ 0.75 ขึ้นไป

ลิวิน สายยศ และอังคณา สายยศ (2539 : 232) ได้ให้ความหมายความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ว่าเป็นผลของคะแนนที่สอบได้ มีความคงที่ในการจำแนกเป็นผู้รอบรู้หรือไม่รอบรู้ในเรื่องที่สอบ สำหรับวิธีการหาค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์สามารถทำได้หลายวิธี ดังนี้ (ลิวิน สายยศ และอังคณา สายยศ. 2539 : 232 - 244)

1. ความเชื่อมั่นแบบหาความคงที่ของความรอบรู้ (Stability Reliability) เป็นการหาค่าความเชื่อมั่น โดยการนำแบบทดสอบอิงเกณฑ์มาสอบซ้ำ 2 ครั้ง โดยใช้สูตรของชรอก และคอนสแคร์ตี ดังนี้

$$\Phi = \frac{AD - BC}{\sqrt{(AB) + (CD) + (AC) + (BD)}}$$

เมื่อ A แทน จำนวนผู้สอบผ่านก่อนเรียนและหลังเรียน

B แทน จำนวนผู้สอบผ่านหลังเรียน

C แทน จำนวนผู้สอบผ่านก่อนเรียนและสอบไม่ผ่านหลังเรียน

D แทน จำนวนผู้สอบไม่ผ่านก่อนเรียนและหลังเรียน

2. ความเชื่อมั่นแบบความสอดคล้องในการตัดสินใจ (Decision Consistency Reliability) เป็นการหาความสอดคล้องระหว่างการทดสอบ 2 ครั้งจากแบบทดสอบครั้งเดียวหรือแบบทดสอบที่คู่ขนานกัน 2 ฉบับ โดยใช้สูตรของ Carver และ Hambleton and Novick ในการคำนวณหาค่าสัมประสิทธิ์ความสอดคล้อง (Agreement Coefficient) ดังนี้

วิธีที่ 1

$$P = \frac{A+D}{N}$$

- เมื่อ P แทน ค่าสัมประสิทธิ์ความสอดคล้อง
 A แทน จำนวนผู้ที่ผ่านเกณฑ์จากการทดสอบทั้ง 2 ครั้ง
 D แทน จำนวนผู้ที่ไม่ผ่านเกณฑ์จากการทดสอบทั้ง 2 ครั้ง

วิธีที่ 2 Hambleton and Novick ได้เสนอสูตรคำนวณดังนี้

$$P = P_{11} + P_{00}$$

- เมื่อ P แทน สัดส่วนของความสอดคล้องในการตัดสินใจเพื่อจำแนก ผู้รอบรู้
 P_{11} แทน สัดส่วนของผู้ถูกตัดสินว่ารอบรู้ตรงกันทั้งสองฉบับหรือสองครั้ง
 P_{00} แทน สัดส่วนของผู้ถูกตัดสินว่าไม่รอบรู้ตรงกันทั้งสองฉบับ หรือสองครั้ง

วิธีที่ 3 Swaminathan, Hambleton and Algina ได้เสนอสูตร Kappa ของ Cohon ในการคำนวณหาค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ ดังนี้

$$K = \frac{P - P_c}{1 - P_c}$$

- เมื่อ K แทน ค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์
 P แทน สัดส่วนของความสอดคล้องในการตัดสินใจความรอบรู้
 ที่ได้จากการสอบซ้ำหรือการทดสอบ 2 ครั้ง
 P_c แทน สัดส่วนของความสอดคล้องที่คาดหวังโดยบังเอิญ

3. การคำนวณหาค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ จากการ
 ทดสอบเพียงครั้งเดียว

วิธีที่ 1 การหาค่าความเชื่อถือได้จากสูตรของ Livingston

$$r_c = \frac{\sigma^2(KR.20) + (\mu - KC)^2}{\sigma^2 + (\mu - KC)^2}$$

- เมื่อ r_{cc} แทน ความเชื่อถือได้ของแบบทดสอบอิงเกณฑ์
 σ^2 แทน คะแนนความแปรปรวนของแบบทดสอบอิงเกณฑ์
 K แทน จำนวนข้อของแบบทดสอบ
 C แทน สัดส่วนของเกณฑ์ที่ผ่าน

- μ แทน คะแนนเฉลี่ยของแบบทดสอบอิงเกณฑ์
 KR_{20} แทน ความเชื่อถือได้ของแบบทดสอบที่หาจากสูตร KR_{20}

วิธีที่ 2 การหาค่าความเชื่อมั่นจากสูตรไบโนเมียล (Binomial Formula)

ของ Lovett

$$r_{cc} = 1 - \frac{K \sum X_i - \sum X_i^2}{\{(K-1) \sum (X_i - c)^2\}}$$

เมื่อ r_{cc} = แทน ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ K แทนจำนวน
ข้อสอบ

$\sum X_i$ แทน ผลรวมของคะแนนทั้งหมด

$\sum X_i^2$ แทน ผลรวมทั้งหมดของคะแนนแต่ละตัวยกกำลังสอง

X_i แทน คะแนนของนักเรียนแต่ละคน

C แทน คะแนนจุดตัดของแบบทดสอบอิงเกณฑ์

วิธีที่ 3 การหาค่าความเชื่อมั่น โดยใช้วิธีวิเคราะห์ความแปรปรวนของ

Hoyt's Anova Procedure

$$r_{tt} = 1 - \frac{MS_E}{MS_p}$$

เมื่อ MS_E แทน คะแนนความแปรปรวนของความคลาดเคลื่อน (Error)

MS_p แทน คะแนนความแปรปรวนระหว่างคน (Between People)

วิธีที่ 4 การหาค่าความเชื่อมั่น โดยวิธีแบ่งครึ่งแบบทดสอบโดยใช้สูตร
ของ Spearman-Brown แล้วใช้สูตรปรับแก้ของ Angoff ดังนี้

$$r_{cc} = \frac{r_{12} \sigma^2}{(\sigma_1 + r_{12})(\sigma_1 + r_{12})}$$

เมื่อ σ^2 แทน ความแปรปรวนของข้อสอบทั้งฉบับ

σ_1 แทน ความเบี่ยงเบนของข้อสอบข้อคู่

σ_2 แทน ความเบี่ยงเบนของข้อสอบข้อคี่

r_{12} แทน สหสัมพันธ์ระหว่างข้อสอบคู่กับข้อคี่

วิธีที่ 5 การหาค่าความเชื่อถือได้ โดยใช้สูตรของ Harris ดังนี้

$$r_{cc} = \frac{SS_b}{SS_b + SS_w}$$

เมื่อ SS_b แทน ผลรวมกำลังสองระหว่างกลุ่ม
 SS_w แทน ผลรวมกำลังสองภายในกลุ่ม

4. การหาความเชื่อมั่นที่คำนึงจุดประสงค์เชิงพฤติกรรมของแบบทดสอบอิงเกณฑ์ Raju ได้เสนอสูตร Comgeneric ดังนี้

$$r_{cc} = \frac{\sigma^2 + (\mu - c)^2}{[\sigma^2 + (\mu - c)^2]1 - \sum \lambda_i^2}$$

เมื่อ σ^2 แทน ความแปรปรวนของข้อสอบทั้งฉบับ
 μ แทน คะแนนเฉลี่ยของแบบทดสอบอิงเกณฑ์
 μ_1 แทน คะแนนเฉลี่ยในแต่ละจุดประสงค์
 λ_i แทน $\frac{K_i}{\sum K_i}$ เมื่อ K_i คือจำนวนข้อสอบในแต่ละจุดประสงค์
 เชิงพฤติกรรม

λ_i แทน คะแนนจุดตัดของแต่ละจุดประสงค์

สมนึก ภักดิ์ทิษณี (2551 : 229) กล่าวถึงการหาค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ว่า ความเชื่อมั่นชนิดที่เป็นการตรวจหาความสอดคล้องของคะแนนแต่ละคนที่แปรปรวนไปจากคะแนนจุดตัด ตามวิธีของโลเวทท์ (Lovett Method) วิธีนี้นำแบบทดสอบอิงเกณฑ์ฉบับเดียว ไปทดสอบนักเรียนกลุ่มเดียว เพียงครั้งเดียว สามารถนำผลมาวิเคราะห์หาความเชื่อมั่นได้จากสูตร ดังนี้

$$r_{cc} = 1 - \frac{k \sum x_i - \sum x_i^2}{(k-1) \sum (x_i - c)^2}$$

เมื่อ r_{cc} แทน ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์
 k แทน จำนวนข้อของแบบทดสอบทั้งฉบับ
 x_i แทน คะแนนสอบของนักเรียนแต่ละคน
 C แทน คะแนนจุดตัด

เขาวดี ราชชัยกุล วิบูลย์ศรี (2552 : 88) กล่าวถึงความเชื่อมั่น ตรงกับภาษาอังกฤษว่า “Reliability” ซึ่งหมายถึง “Stability and Consistency” ของคะแนนสอบจึงเป็นที่เข้าใจของกลุ่ม นักวัดผลคนไทยว่า Reliability นั้นหมายถึง ระดับความคงที่หรือความคงเส้นคงวาของคะแนนสอบจากการทดสอบเรื่องเดียวกันในเวลาใดก็ตาม อย่างไรก็ตาม ใ้การใช้คำนั้นก็อาจใช้คำที่ต่างกันไปเช่นความเชื่อมั่น ความเที่ยง เป็นต้น

ไพศาล วรคำ (2552 : 267 - 268) ให้ความหมายของความเชื่อมั่นว่า หมายถึง ความคงที่ของผลที่ได้จากการวัดด้วยเครื่องมือชุดใดชุดหนึ่งในการวัดหลายๆ ครั้ง การหาความเชื่อมั่นของแบบวัดพัฒนามาจากนิยาม คือเป็นความสัมพันธ์กันระหว่างค่าการวัดหลายๆ ครั้ง แต่ด้วยเหตุที่คุณลักษณะที่ต้องการวัดของบุคคลนั้นมักจะมีการเปลี่ยนแปลงเสมอเมื่อเวลาผ่านไปจึงได้มีการพัฒนาวิธีการหาความเชื่อมั่นของแบบวัดขึ้นมาอีกหลายวิธีภายใต้แนวคิดหลัก 3 แนวคิด คือ

1. การวัดความคงที่ซึ่งจะเป็นการวัดความคงที่ของผลการวัดหลายๆ ครั้ง
2. การวัดความสมมูลกัน เป็นการวัดด้วยแบบวัดที่คู่ขนานกันเพื่อเทียบการวัดซ้ำ
3. การวัดความสอดคล้องภายใน ซึ่งเป็นการพิจารณาความเชื่อมั่นจากการวัดเพียงครั้งเดียวแล้วหาความสอดคล้องของผลการวัดภายในแบบวัดนั้น

จากการศึกษาข้อคิดเห็นของนักการศึกษาที่กล่าวมาข้างต้น จะเห็นว่าการหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์มีวิธีการหาได้หลายวิธี ซึ่งในการสร้างแบบทดสอบอิงเกณฑ์ เรื่อง ชนิดของคำ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ในครั้งนี้ผู้วิจัยหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์จากผลสอบครั้งเดียว โดยวิธีของโลเวทท์ (Lovett Method) สมนึก ภัททิยธนี (2551 : 229)

การสร้างข้อสอบแบบเลือกตอบและรูปแบบของข้อสอบแบบเลือกตอบ

การสร้างแบบทดสอบอิงเกณฑ์ สิ่งที่จะต้องคำนึงถึงอีกคือ รูปแบบของข้อสอบแบบเลือกตอบและการเขียนข้อสอบแบบเลือกตอบ เพื่อให้ได้แบบทดสอบที่มีคุณภาพดียิ่งขึ้น ซึ่งผู้วิจัยแยกกล่าวรายละเอียด ดังต่อไปนี้ (บุญเชิด ภิญ โญอนันตพงษ์. 2526 : 122)

ข้อสอบแบบเลือกตอบ (Multiple choice test) ลักษณะข้อสอบ จะประกอบด้วย โจทย์หรือข้อความที่เป็นประโยคที่สมบูรณ์เป็นตัวคำถาม (Stem) เพื่อวัดความรู้ความสามารถ

และตัวเลือกตั้งแต่ 3 ตัวเลือกขึ้นไปอีก 1 ชุด รวมเป็น 1 ข้อ ในตัวเลือกนั้นจะมีทั้งคำตอบถูก (Key) และคำตอบผิด (Distracter) ที่เป็นตัวลวงมาให้นักเรียนพิจารณา คำถามควรเป็นประโยค คำถามที่สมบูรณ์ เพื่อจะช่วยให้มีความชัดเจน และเข้าใจง่ายกว่าประโยคบอกเล่า

1. เน้นจุดที่เป็นคำถามให้ชัดเจน เพื่อให้เกิดความเป็นปรนัย
2. หลีกเลี่ยงคำถามที่เป็นประโยคปฏิเสธ โดยเฉพาะปฏิเสธซ้อน แต่ถ้า

จำเป็นต้องใช้ประโยคปฏิเสธควรขีดเส้นใต้ให้ชัดเจน

3. การถามคำถามจะต้องสั้น กระชับและได้ใจความ ไม่ควรใช้คำฟุ่มเฟือย
4. ถามในสิ่งที่มีประโยชน์ เพราะจะช่วยให้เด็กได้เรียนรู้สิ่งที่ตรง เป็นการปลูกฝังค่านิยมที่พึงประสงค์ กล่าวคือ ถ้าสิ่งใดเป็นสิ่งที่ดีสังคมยอมรับ ให้ถามในทางที่ดี แต่ถ้าสิ่งใดไม่ดีสังคมไม่ยอมรับ ให้ถามในทางที่ไม่ดี

5. ถามในสิ่งที่สามารถหาข้อยุติได้ตามหลักวิชา เพื่อให้เด็กได้ใช้ความคิด ไม่ถามในสิ่งที่เป็นความเชื่อ

6. ควรถามพฤติกรรมที่ต้องใช้ความคิด และควรหลีกเลี่ยงการถามความจำจากตำรา

7. ใช้ภาษาให้เหมาะสมกับวัยของเด็ก เพราะถ้าใช้ภาษาที่ยากเกินไป เด็กก็จะไม่สามารถที่จะเข้าใจในความหมาย จึงไม่สามารถทำข้อสอบนั้นได้

8. ควรใช้คำถามที่ยั่วให้เด็กขวนคิด และบางครั้งคำถามหรือตัวเลือกอาจเป็นรูปภาพ สัญลักษณ์ สถานการณ์ เพื่อยั่วให้เด็กอยากทำ

9. ตัวเลือกควรสั้น กระชับ และมีความหมาย คำฟุ่มเฟือยตัดทิ้ง

10. ตัวเลือกแต่ละตัวควรมีความยาวเท่า ๆ กัน ถ้าตัวเลือกลาวไม่เท่ากัน

ควรเรียงจากสั้นไปหายาว แต่ทั้งนี้ถ้าเป็น วัน เดือน พ.ศ. ตัวเลขหรือจำนวน ให้เรียงจากน้อยไปหามาก

11. หลีกเลี่ยงการเขียนตัวเลือกซ้ำซ้อนกัน หรือมีความหมายเหมือนกัน เพราะจะทำให้ตัวเลือกมีคุณค่าลดน้อยลงตัวเลือกต้องเป็นเอกพันธ์กัน (Homogeneity) หรือไม่มีโครงสร้างสอดคล้องกัน

12. ควรระมัดระวังการใช้ตัวเลือกประเภท ปลายเปิดและปลายปิด ซึ่งได้แก่ ถูกทุกข้อ ไม่มีข้อถูก ผิดทุกข้อ สรุปแน่นอนไม่ได้ เป็นต้น

13. หลีกเลี่ยงการแนะคำตอบ เพราะเพียงแต่เด็กสังเกตก็สามารถหาคำตอบได้ โดยไม่ได้ใช้ความคิดเลย

14. ไม่ควรถามในสิ่งที่เด็กคล่องปาก
15. ควรกระจายตัวเลือกที่เป็นตัวถูกให้อยู่ในตำแหน่งที่ต่าง ๆ กัน และแต่ละตัวเลือกควรมีโอกาสเป็นตัวถูกในจำนวนที่เท่า ๆ กัน
16. หลีกเลี่ยงการเขียนตัวถูกที่ฟ้องเสียงหรือฟ้องความหมายกับตัวคำถาม เพราะจะเป็นการแนะนำคำตอบ

ข้อดีของข้อสอบแบบเลือกตอบ

1. วัดพฤติกรรมทางการศึกษาได้หลายด้าน ตั้งแต่ความรู้ความจำไปจนถึงการประมาณค่าเป็นข้อสอบที่ตรวจให้คะแนนง่าย ถูกต้อง รวดเร็ว และมีความเป็นปรนัย
2. สามารถควบคุมความยากง่ายของข้อสอบได้
3. เป็นข้อสอบที่ครูสามารถวินิจฉัยสาเหตุแห่งการทำข้อสอบผิด ว่าเนื่องจากสาเหตุอะไรบ้าง โดยพิจารณาจากตัวเลือกต่าง ๆ จากกระดาษคำตอบ
4. มีความเชื่อมั่นสูง เพราะมีจำนวนข้อสอบมาก และตอบถูกโดยการเดา มีน้อย
5. สามารถใช้สัญลักษณ์ รูปภาพหรือกราฟมาเขียนข้อสอบได้

ข้อจำกัดของข้อสอบแบบเลือกตอบ

1. สร้างข้อสอบให้ทำได้ยาก และใช้เวลาในการสร้างนาน
2. ไม่เหมาะที่จะวัดความคิดริเริ่มสร้างสรรค์ การเสนอแนวคิด หรือทักษะ

การเขียน

3. สิ้นเปลืองค่าใช้จ่ายสูงในการสร้างข้อสอบ

รูปแบบของข้อสอบแบบเลือกตอบ

ข้อสอบแบบเลือกตอบจะมีคุณภาพมากหรือน้อย มักเกิดจากการเขียนตัวคำถามและตัวเลือก นอกจากนี้ควรเขียนคำถามที่วัดพฤติกรรมขั้นสูง ๆ โดยใช้รูปแบบคำถามที่เหมาะสม (บุญเชิด ภิญ โยอนันตพงษ์. 2526 : 122)

การสร้างแบบทดสอบแบบเลือกตอบมีหลักการและข้อเสนอแนะ ดังนี้

ตัวคำถาม (Stem)

1. ตัวคำถามมีความหมายสมบูรณ์ในตัวเอง และถามปัญหาอย่างเฉพาะเจาะจง
2. ตัวคำถามแต่ละข้อควรเขียนให้สั้น ชัดเจน และใช้ภาษาที่เข้าใจง่าย
3. ตัวคำถามควรสร้างด้วยประโยคบอกเล่า หลีกเลี่ยงการใช้ประโยคปฏิเสธ หากจำเป็นให้ขีดเส้นใต้ประโยคปฏิเสธนั้น แต่ถ้าเป็นประโยคปฏิเสธซ้อนไม่ควรใช้เด็ดขาด เพราะประโยคปฏิเสธจะอ่านเข้าใจยากกว่าประโยคบอกเล่า

4. ตัวคำถามแต่ละข้อจะต้องถามปัญหาเพียงประเด็นเดียว

5. ตัวคำถามแต่ละข้อต้องเป็นอิสระกัน อย่าให้เกี่ยวข้องกัน เพื่อป้องกันการแนะนำ

ตัวเลือกถูกให้ข้ออื่น

6. ตัวคำถามควรเขียนด้วยภาษาง่าย ๆ ศัพท์เทคนิคไม่จำเป็นก็ไม่ควรใช้ หากใช้ควรใช้ภาษาอังกฤษกำกับหรือแปลความหมายไว้ด้วย

ตัวเลือก (Alternative)

1. ตัวเลือกทุกตัวในข้อเดียวกันจะต้องมีความเป็นเอกพันธ์

2. ตัวเลือกทุกตัวควรมีโอกาสถูกพอ ๆ กัน

3. ตัวเลือกทุกตัวควรมีความยากง่ายพอ ๆ กัน

4. ตัวเลือกในแต่ละข้อควรเรียงตามหลักและเหตุผล หรือเรียงอย่างเป็นระบบ

5. ตัวเลือกในแต่ละข้อต้องมีความถูกเพียงตัวเดียว

6. ภาษาที่ใช้ในตัวเลือกไม่ควรตรงกับตัวคำถาม ถ้าต้องมีจะต้องนำไปรวมไว้

ในตัวคำถาม

7. ตัวเลือกที่ว่า ทุกข้อข้างต้นถูกหมด ทุกข้อข้างต้นผิดหมด และไม่มีข้อใดถูกเลยไม่ควรนำมาใช้ เว้นแต่กรณีใช้วัดความสามารถทางคณิตศาสตร์

รูปแบบการเขียนข้อสอบแบบเลือกตอบ

1. แบบคำถามโคดหรือคำถามเดี่ยว (Single Question)

2. แบบตัวเลือกคงที่ (Constant Choices)

3. แบบกำหนดสถานการณ์ (Situation test)

ข้อดีของข้อสอบแบบเลือกตอบ

1. ใช้วัดผลการเรียนรู้ระดับสูง ๆ และที่ซับซ้อนได้

2. ตรวจให้คะแนนง่ายสะดวกและรวดเร็ว

3. มีประสิทธิภาพการวัดได้ดีกว่าข้อสอบแบบอื่น ๆ

4. มีโอกาสการเดาน้อยกว่าข้อสอบปรนัยแบบอื่น ๆ

5. วัดครอบคลุมเนื้อหาได้มาก จึงมีความตรงตามเนื้อหาสูง

6. มีความเที่ยงธรรมในการนำไปใช้วัดผลการเรียนรู้สูง เพราะข้อสอบแต่ละข้อ

จะมีความเป็นปรนัยมาก

7. เหมาะสมสำหรับเก็บไว้ใช้ได้อีก เพราะสามารถนำไปวิเคราะห์รายข้อเพื่อตรวจสอบคุณภาพ ถ้ามีคุณภาพก็เก็บไว้ใช้ในโอกาสอื่นต่อไปได้ ถ้าไม่มีคุณภาพก็สามารถนำไปปรับปรุงแก้ไข แล้วเก็บไว้ใช้ได้เช่นเดียวกัน

ข้อจำกัดของข้อสอบแบบเลือกตอบ

1. สร้างให้มีคุณภาพดี ๆ สร้างได้ยาก ต้องใช้ผู้มีความรู้และทักษะในการสร้างข้อสอบมาพอสมควร

2. เสียเวลาและแรงงานในการสร้างมาก เมื่อเทียบกับข้อสอบประเภทอื่น ๆ

3. สิ้นเปลืองเวลาและการดาษในการพิมพ์ข้อสอบมาก

4. ถ้าข้อสอบยากมาก ๆ นักเรียนจะตอบด้วยการเดามากขึ้น

5. ไม่เหมาะที่จะใช้วัดความคิดริเริ่มและความคิดสร้างสรรค์

ไพศาล วรคำ (2552 : 227) ได้กล่าวถึงหลักการสร้างแบบทดสอบเลือกตอบไว้ว่า

1. ควรตั้งคำถามด้วยประโยคคำถามที่สมบูรณ์ เขียนด้วยภาษาที่มีความชัดเจนมากที่สุดเท่าที่จะทำได้

2. ประเด็นคำถามต้องมีความเฉพาะเจาะจง คำถามแต่ละข้อจะวัดตามวัตถุประสงค์ เพียงข้อใดข้อหนึ่งเท่านั้น และถ้าวัตถุประสงค์นั้นมีรายละเอียดมากก็ควรแยกถามตามประเด็น ย่อย ๆ

3. คำถามและตัวเลือกต้องไม่มีลักษณะชี้แนะคำตอบ หรือคำถามในข้อหนึ่งเป็นคำตอบของอีกข้อหนึ่ง หรือตัวเลือกถูกในข้อหนึ่งเป็นแนวทางในการเลือกตัวเลือกในข้อถัดไป

4. ควรหลีกเลี่ยงคำถามเชิงนิเสธ โดยเฉพาะอย่างยิ่งคำถามเชิงนิเสธซ้อนเพราะจะทำให้ผู้ตอบเกิดความสับสน แต่หากไม่สามารถหลีกเลี่ยงได้ ก็ควรพิมพ์ตัวหนาหรือขีดเส้นใต้คำที่แสดงนิเสธให้เห็นเด่นชัด

5. ควรปรับตัวเลือกทุกตัวให้มีความยาวพอ ๆ กัน หรือถ้าไม่สามารถปรับให้ความยาวใกล้เคียงกันได้ ก็ควรเรียงตัวเลือกจากสั้นไปหายาว หรือจากยาวไปหาสั้นอย่างเป็นระบบ

6. ถ้าตัวเลือกเป็นตัวเลขหรือตัวอักษร ควรเรียงตัวเลือกตามลำดับค่าหรือตามลำดับอักษรอย่างเป็นระบบ

7. ตัวเลือกแต่ละตัวควรเป็นอิสระจากกัน

8. ไม่ควรใช้ภาษาฟุ่มเฟือยโดยไม่จำเป็นทั้งข้อความและตัวเลือก การเขียนข้อสอบควรเลือกใช้คำที่มีความหมายตรง และชัดเจนมากที่สุดเพียงไม่กี่คำที่สามารถสื่อความหมายตามที่ต้องการ แต่ต้องไม่ก่อให้เกิดความกำกวมขึ้นกับข้อสอบ
 9. ตัวถูกและตัวลวงไม่ควรให้ถูกหรือผิดชัดเจนมากเกินไป ซึ่งจะทำให้ผู้ตอบเดาคำตอบที่ถูกต้องได้
 10. ตัวเลือกในแต่ละข้อควรมีความเป็นเอกพันธ์ หรือเป็นตัวเลือกที่อยู่ในเรื่องเดียวกันหรือคล้ายคลึงกัน
 11. การใช้ตัวเลือกปลายปิดและปลายเปิดควรเลือกใช้อย่างมีเหตุผล เช่น ถูกหมดทุกข้อ สรุบนั่นเองไม่ได้ หรือไม่มีข้อใดถูก
 12. ควรจัดวางข้อถูกให้กระจายอย่างสุ่ม ไม่ควรจัดวางอย่างเป็นระบบเด็ดขาด เพราะจะทำให้ผู้ตอบสามารถเดาคำตอบได้
 13. ในแบบทดสอบชุดหนึ่ง ๆ ควรมีจำนวนตัวเลือกที่คงที่ การกำหนดจำนวนตัวเลือกโดยทั่วไปจะกำหนดให้เหมาะสมกับระดับของผู้ตอบ
 14. หลีกเลี่ยงสิ่งที่ทำให้ข้อสอบยากขึ้นโดยไม่เกี่ยวข้องกับจุดมุ่งหมายของการวัด
 15. ตรวจสอบข้อสอบทั้งหมดอีกครั้งก่อนนำไปใช้ ทั้งในเรื่องของภาษา ระดับความยาก และตัวเลือกเพื่อให้แน่ใจว่ามีความถูกต้องเหมาะสม
- จากการศึกษาหลักการสร้างแบบทดสอบของนักการศึกษาที่กล่าวมาข้างต้น ในครั้งนี้ผู้วิจัยได้สร้างแบบทดสอบอิงเกณฑ์ชนิดเลือกตอบ เนื่องจากแบบทดสอบชนิดเลือกตอบมีประสิทธิภาพในการวัดได้ดีกว่าข้อสอบแบบอื่น ๆ วัดครอบคลุมเนื้อหาได้มาก จึงมีความตรงเชิงเนื้อหาสูง ไพศาล วรคำ (2552 : 229)

งานวิจัยที่เกี่ยวข้อง

1. งานวิจัยในประเทศ

เจริญชัย จงนุเคราะห์ (2540 : 85) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องสมการและการแก้สมการ ชั้นประถมศึกษาปีที่ 6 กลุ่มตัวอย่างได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2536 ดำเนินงานการประถมศึกษาอำเภอร่องคำ จังหวัดกาฬสินธุ์ จำนวน 200 คน แบบทดสอบที่สร้างขึ้นมี 2 ฉบับ ชนิด 5 ตัวเลือก ผลการวิจัยสรุปได้

ดังนี้ แบบทดสอบฉบับที่ 1 เรื่องความรู้เบื้องต้นเกี่ยวกับสมการจำนวน 15 ข้อ ค่าดัชนี S ตั้งแต่ .20 - .39 ค่าความเที่ยงตรงเชิงโครงสร้างเท่ากับ .70 ค่าความเชื่อมั่นและคะแนนเกณฑ์เป็น 11 คะแนน แบบทดสอบฉบับที่ 2 เรื่องการแก้สมการ จำนวน 25 ข้อ ค่าดัชนี S ตั้งแต่ .03 - .38 ค่าความเที่ยงตรงเชิงโครงสร้างเท่ากับ .69 ค่าความเชื่อมั่นเท่ากับ .87 และคะแนนเกณฑ์ เป็น 15 คะแนน

เสาวนิต ร่มศรี (2546 : 103) ได้สร้างแบบทดสอบอิงเกณฑ์ วิชาคณิตศาสตร์ เรื่องเมตริกซ์ ชั้นมัธยมศึกษาปีที่ 5 หากคุณภาพและหาคะแนนเกณฑ์ที่เหมาะสมของแบบทดสอบที่สร้างขึ้นกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 ภาคเรียนที่ 1 ปีการศึกษา 2540 ของโรงเรียนสังกัดกรมสามัญศึกษา อำเภอดุสิต กรุงเทพมหานคร จำนวน 176 คน โดยวิธีการสุ่มอย่างง่าย แบบทดสอบที่สร้างขึ้นมี 8 ฉบับ แต่ละฉบับมี 5 ตัวเลือก แต่ละฉบับมีจำนวนข้อ ดังนี้ 15,10,10,12,20,15,15 และ 10 ข้อตามลำดับ ปรากฏว่ามีค่าอำนาจจำแนกตั้งแต่ .22 - .66 ค่าความเที่ยงตรงเชิงโครงสร้าง ตั้งแต่ .725 - .893 ค่าความเชื่อมั่นตั้งแต่ .750 - .915 และคะแนนเกณฑ์แต่ละฉบับคิดเป็นร้อยละ 53.33,50,50,40,40,40,53.33 และ 40 ตามลำดับ

เกียรติพงษ์ พิสงปรากร (2546 : 94) ได้พัฒนาแบบทดสอบอิงเกณฑ์ วิชาคณิตศาสตร์ เรื่อง ระบบจำนวนเต็ม กลุ่มตัวอย่างที่ใช้ศึกษาเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนสังกัดสำนักงานการประถมศึกษาอำเภอภูสิงห์ จังหวัดศรีสะเกษ จำนวน 212 คน ได้มาโดยวิธีการสุ่มอย่างง่าย ผลการวิจัยพบว่า 1) แบบทดสอบอิงเกณฑ์ เรื่อง ระบบจำนวนเต็ม ฉบับที่ 1 จำนวน 40 ข้อ มีค่าอำนาจจำแนกระหว่าง .31 ถึง .56 ค่าความแม่นยำเชิงโครงสร้างเท่ากับ .94 ค่าความเชื่อถือได้ เท่ากับ .88 และคะแนนจุดตัดเท่ากับ 13 คะแนน คิดเป็นร้อยละ 47.50 2) แบบทดสอบ อิงเกณฑ์ เรื่อง ระบบจำนวนเต็ม ฉบับที่ 2 จำนวน 30 ข้อ มีค่าอำนาจจำแนกระหว่าง .24 ถึง .52 ค่าความแม่นยำเชิงโครงสร้างเท่ากับ .89 ค่าความเชื่อถือได้ เท่ากับ .85 และคะแนนจุดตัดเท่ากับ 13 คะแนน คิดเป็นร้อยละ 43.33

ไพโรจน์ ใจดี (2546 : 87) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่อง เศษส่วน ชั้นประถมศึกษาปีที่ 5 ปีการศึกษา 2545 แบบทดสอบที่สร้างขึ้นมี 2 ฉบับ ฉบับที่ 1 มีค่าความเที่ยงตรงเชิงเนื้อหา ตั้งแต่ .83 ถึง 1.00 ค่าอำนาจจำแนกตั้งแต่ .20 ถึง .77 ค่าความยาก ตั้งแต่ .51 ถึง .92 ค่าความเที่ยงตรงเชิงโครงสร้างเท่ากับ .67 และค่าความเชื่อมั่นเท่ากับ .91 และมีคะแนนจุดตัดเท่ากับ 23 คะแนน คิดเป็นร้อยละ 50 ฉบับที่ 2 มีค่าความเที่ยงตรงเชิงเนื้อหา ตั้งแต่ .83 ถึง 1.00 ค่าอำนาจจำแนกตั้งแต่ .20 ถึง .85 ค่าความยากตั้งแต่ .50 ถึง .86 ค่าความ

เที่ยงตรงเชิงโครงสร้างเท่ากับ .79 และค่าความเชื่อมั่นเท่ากับ .89 และมีคะแนนจุดตัดเท่ากับ 20 คะแนน คิดเป็นร้อยละ 50

จินตนา แสงประเสริฐสุข (2553 : 52) ได้สร้างแบบทดสอบอิงเกณฑ์ กลุ่มสาระการเรียนรู้ภาษาไทย สาระหลักการใช้ภาษา ชั้นประถมศึกษาปีที่ 2 สร้างและหาคุณภาพ หาคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 2 สังกัดสำนักงานเขตพื้นที่การศึกษาเพชรบูรณ์เขต 3 จังหวัดเพชรบูรณ์ ปีการศึกษา 2551 จำนวน 345 คน โดยการสุ่มแบบหลายชั้นตอน ผลการวิจัยสรุปได้ดังนี้ แบบทดสอบอิงเกณฑ์ กลุ่มสาระการเรียนรู้ภาษาไทย สาระหลักการใช้ภาษาชั้นประถมศึกษาปีที่ 2 สร้างได้จำนวน 9 ฉบับ ๆ ละ 15 ข้อ มีคุณภาพ ดังนี้ ความเที่ยงตรงเชิงเนื้อหา ตั้งแต่ .80 ถึง 1.00 อำนาจจำแนก ตั้งแต่ .20 ถึง .82 และความเชื่อมั่น ตั้งแต่ .76 ถึง .89 หาคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ มีคะแนนจุดตัดเท่ากับ 7 ถึง 8 คะแนนจากคะแนนเต็ม 15 คะแนน หรือคิดเป็นร้อยละ 46.67 ถึง 53.33 ของคะแนนเต็ม

2. งานวิจัยต่างประเทศ

Hsu (1971 : 87-95) ได้ศึกษาค่าอำนาจจำแนกของแบบทดสอบอิงเกณฑ์ โดยหาค่าความแตกต่างระหว่างสัดส่วนของผู้ตอบถูกในกลุ่มผู้รอบรู้ กับสัดส่วนของผู้ตอบถูกในกลุ่มผู้ไม่รอบรู้ (D %) หากค่า Fti - Coefficient (C) แล้วนำค่านีมาเปรียบเทียบกับสัมประสิทธิ์สหสัมพันธ์ Point-biserial (rpbis) ของข้อสอบกับคะแนนจากแบบทดสอบสามแบบคือ แบบแรก เป็นการทดสอบก่อนเรียน - หลังเรียน แบบที่สอง เป็นการดำเนินการสอบกับนักเรียนระดับชั้นต่าง ๆ กันในแบบทดสอบสองรูปแบบ แบบที่ 3 เป็นการนำข้อสอบมาเข้ากลุ่มตามดัชนีความยากสามระดับ พบว่าค่าสถิติทั้งสามแบบคือ D % C และ rpbis มีความสัมพันธ์กันสูง (Haladyna. 1974 : 94)

Haladyna (1974 : 93 - 99) ได้ศึกษาผลการใช้กลุ่มตัวอย่างต่างกัน ต่อลักษณะของข้อสอบอิงเกณฑ์ พบว่าความแปรปรวนของคะแนนจะเพิ่มขึ้นตามประสิทธิภาพการสอบ

Crehan (1974 : 255 - 256) ได้ศึกษาแบบทดสอบอิงเกณฑ์โดยใช้วิธีการคัดเลือกข้อสอบ 6 วิธี คือวิธีของ Cox and Vagas วิธีของ Brennan วิธีเรียงตามค่าสัมประสิทธิ์สหสัมพันธ์แบบ Point Biserial และวิธีเรียงตามสุ่มกลุ่มตัวอย่างแต่ละกลุ่มใช้ตั้งแต่ 7 - 30 คน โดยใช้แบบทดสอบคู่ขนาน พบว่าวิธีการเลือกข้อสอบแบบอิงเกณฑ์ทั้ง 6 วิธีนี้ให้ผลต่อค่าความเชื่อถือได้ไม่แตกต่างกัน ค่าความแม่นยำตรงได้เลือกเกณฑ์ที่ให้ค่าสูงสุด โดยใช้จุด

เดียวกันทั้งสองฉบับปรากฏว่าวิธีการของ Cox and Vagas และ Brennan ให้ค่าความแม่นยำสูงสุด

Swaminathan, Hambleton and Algina (1974 : 87 - 88) ได้ทำการศึกษาวิธีกำหนดเกณฑ์ของแบบทดสอบอิงเกณฑ์ โดยใช้ทฤษฎีการตัดสินใจของ Bayes (Bayesian Decision – Theoretic Procedure) ทำการศึกษากับนักเรียน 25 คน ใช้แบบทดสอบอิงเกณฑ์จำนวน 10 ข้อ กำหนดค่าค่าสูญเสียที่น้อยที่สุดของเกณฑ์โดยกำหนดค่าแอลฟา (α) เท่ากับ 1 และค่าเบต้า (β) เท่ากับ 2 พบว่าเมื่อแบ่งผู้สอบออกเป็น 2 ระดับ และกำหนดเกณฑ์เป็น .9 ใน 10 ข้อ เป็นผู้รอบรู้ และเมื่อแบ่งผู้สอบออกเป็น 3 ระดับ และกำหนดเกณฑ์เป็น .6 และ .8 ผู้เข้าสอบทำได้ 7 - 9 ข้อ ต้องสอนบททวนใหม่อีกเล็กน้อย ส่วนผู้ที่ได้คะแนนต่ำกว่า 7 ข้อ เป็นผู้ไม่รอบรู้ต้องสอนบททวนใหม่ทั้งหมด

Berk (1976 : 5 - 7) ได้ศึกษาวิธีการเลือกจุดตัดของคะแนนในแบบทดสอบอิงเกณฑ์ โดยใช้กลุ่มตัวอย่างจำนวน 100 คน พบว่า เกณฑ์ที่เหมาะสมตัดสินได้ดีที่สุดคือ .5

Subkoviak (1976 : 265 - 276) ได้ทำการศึกษาพบว่า เมื่อเกณฑ์มีค่าต่ำสุดและสูงสุด ความเชื่อมั่นจะมีค่าสูงสุด และเมื่อเกณฑ์อยู่ในระดับกลาง ๆ ความเชื่อมั่นจะมีค่าต่ำสุด

Homer (1977 : 4139 - A) ทำการศึกษาใช้แบบทดสอบอิงเกณฑ์และแบบทดสอบอิงกลุ่มวิชาคณิตศาสตร์กับนักเรียนระดับ 4 จำนวน 133 คน ในการทดลองสองทาง พบว่าสัมประสิทธิ์ระหว่างตัวแปรของแบบทดสอบทั้งสองชนิดคือ .87 ไม่มีนัยสำคัญของความแตกต่างระหว่างแบบทดสอบทั้งสองชนิด

Ganapole (1978 : 274 - A) ได้ศึกษาการสร้างแบบทดสอบอิงเกณฑ์ วัดความสามารถพื้นฐาน ในการอ่านของนักเรียนระดับมัธยมศึกษา โดยเลือกสถานการณ์ที่ตั้งขึ้นจากมวลประชากรในเรื่องนี้ ซึ่งจะตัดสินความรู้ในการอ่านที่สามารถประเมินได้ และนิยามจำกัดของพฤติกรรมที่จะวัด โดยแบบทดสอบความเชื่อมั่นของแบบทดสอบได้จากการสอบซ้ำและใช้แบบทดสอบคู่ขนาน ความเที่ยงตรงได้จากการเลือกข้อสอบจากโดเมน

James (1980 : 575) ได้ศึกษาประโยชน์ที่ได้รับจากการใช้ข้อสอบอิงเกณฑ์นักเรียนที่เรียนคณิตศาสตร์ เกรด 9 พบว่า ความสำเร็จของการเรียนอยู่ที่การสอบผ่านคะแนนเกณฑ์หรือคะแนนจุดตัด 2 ใน 3 ของข้อสอบ จุดตัดแต่ละ CRT และคะแนนรวม CRT มีความสัมพันธ์กับเนื้อหาแต่ละหน่วยของวิชา

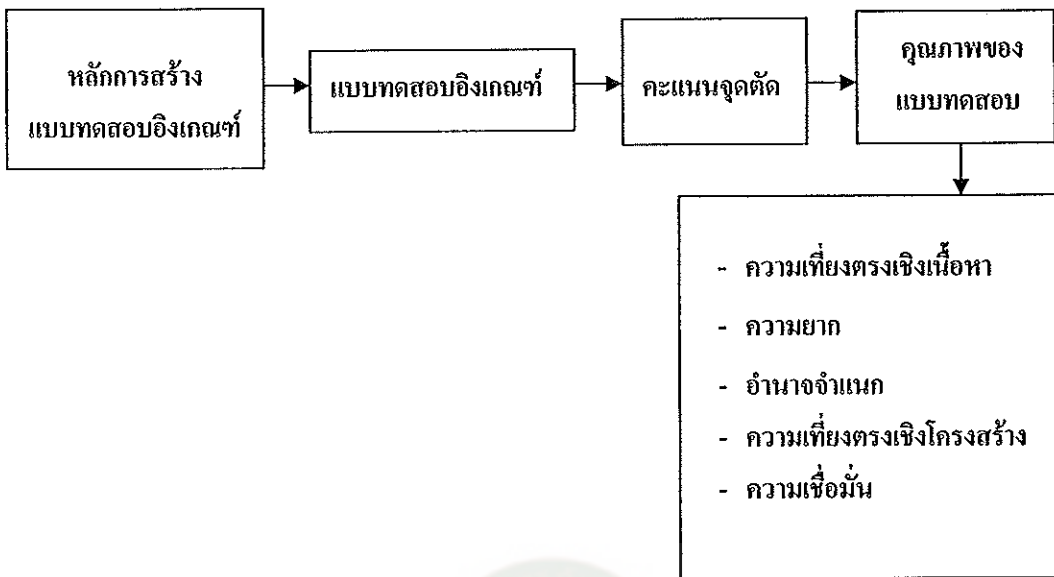
Isbell Celia Ann Berg (1986 : 158) ได้ศึกษาความขัดแย้งของแบบทดสอบอิงเกณฑ์คณิตศาสตร์ 2 ชุด ซึ่งใช้ในกลุ่มรอบรู้ที่มีทักษะเหมือนกัน จากกลุ่มตัวอย่าง จำนวน 221

คน พบว่า แบบทดสอบอิงเกณฑ์ (CRTM) สามารถนำไปใช้ประเมินประสิทธิภาพของ แบบทดสอบฟลอริดา (SSAT) โดยการวิเคราะห์คะแนนของนักเรียนที่เรียนคณิตศาสตร์เกรด 8 จากแบบทดสอบ 2 ชุด ได้คะแนนเปรียบเทียบกัน ดังนี้ คะแนนดิบของแบบทดสอบอิงเกณฑ์ (CRTM) ไม่สามารถทำนายคะแนนดิบแบบทดสอบฟลอริดาได้, กลุ่มรอบรู้และไม่รอบรู้ของ แบบทดสอบอิงเกณฑ์ ไม่ได้ประกอบขึ้นจากการแบ่งชั้นของแบบทดสอบฟลอริดา โดยทั่วไป แบบทดสอบอิงเกณฑ์ ต้องการค่าความเชื่อมั่นมากกว่าแบบทดสอบฟลอริดา และใช้ 60 % เป็น คะแนนจุดตัด แยกเป็นกลุ่มรอบรู้และไม่รอบรู้ ทำให้เกิดความคลาดเคลื่อนน้อยกว่า 80 % ของ คะแนนจุดตัด

จากการศึกษาค้นคว้าที่ได้กล่าวไว้แล้วข้างต้น สรุปได้ว่านักวิจัยทั้งภายในประเทศ และต่างประเทศ ได้ให้ความสนใจต่อการวัดผลแบบอิงเกณฑ์ในรูปแบบต่าง ๆ มากมาย แต่ จากการสำรวจข้อมูลเบื้องต้นปรากฏว่าการสร้างแบบทดสอบอิงเกณฑ์ในกลุ่มสาระการเรียนรู้ ภาษาไทย เพื่อการต่อยอดความรู้มีจำนวนน้อยมาก ผู้วิจัยจึงมีความสนใจที่จะศึกษาการสร้าง แบบทดสอบอิงเกณฑ์ เรื่อง ชนิดของคำ ชั้นประถมศึกษาปีที่ 6 ขึ้น ดังนั้นจึงได้ทำการศึกษา ค้นคว้าเกี่ยวกับการสร้างแบบทดสอบอิงเกณฑ์ในรูปแบบต่าง ๆ จากกลุ่มสาระอื่น ซึ่งนับว่ามีความสำคัญอันที่จะเป็นประโยชน์ในการวัดผลประเมินผลการเรียน และปรับปรุงการเรียนการสอนให้มีประสิทธิภาพและสอดคล้องกับเป้าหมายของหลักสูตรได้เป็นอย่างดี

กรอบดำเนินการวิจัย

ผู้วิจัยได้สรุปเป็นกรอบดำเนินการวิจัยครั้งนี้ โดยมีวัตถุประสงค์เพื่อสร้าง แบบทดสอบ อิงเกณฑ์ เรื่อง ชนิดของคำ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่มีคุณภาพ ดังภาพที่ 2



ภาพที่ 2 กรอบดำเนินการวิจัย