

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการวิจัยเรื่อง การพัฒนาแบบทดสอบอิงเกณฑ์วัดผลสัมฤทธิ์ การใช้โปรแกรม Microsoft Word สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น ผู้วิจัยได้เสนอรายละเอียดของเอกสารและงานวิจัยที่เกี่ยวข้องต่าง ๆ ตามลำดับดังนี้

1. แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน
2. การประเมินผลและแบบทดสอบอิงเกณฑ์
3. คะแนนจุดตัด
4. การวิเคราะห์คุณภาพของแบบทดสอบอิงเกณฑ์
5. งานวิจัยที่เกี่ยวข้อง

1. แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

ความหมายของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

แบบทดสอบ หมายถึง ชุดของคำถามหรือกลุ่มงานที่ใช้เป็นสื่อในการวัดคุณสมบัติใดคุณสมบัติหนึ่งของบุคคล โดยที่บุคคลนั้นเมื่อได้รับสิ่งเร้าจะแสดงพฤติกรรมสนองที่สามารถวัดได้สังเกตได้ ซึ่งจะนำไปสู่การแปลความหมายของผลได้ จึงอาจกล่าวได้ว่าแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน หมายถึง ชุดของคำถามที่มุ่งวัดความรู้ความสามารถ ทักษะและสมรรถภาพทางสมองด้านต่าง ๆ ของผู้เรียนหลังเกิดการเรียนรู้

คุณลักษณะที่ดีของแบบทดสอบ

ในการสอบวัดครั้งหนึ่ง ๆ นั้น ข้อสอบจำเป็นจะต้องให้มีลักษณะเป็นตัวแทนที่ดีของเนื้อหาและองค์ความรู้ต่าง ๆ ของผู้เรียน เพราะการใช้ข้อสอบวัดความรู้ความสามารถของผู้เรียนนั้น เป็นเพียงตัวแทนของเนื้อหาที่ผู้เรียนได้เรียนไปแล้วเท่านั้น ข้อสอบจึงต้องเป็นตัวแทนที่ดี เพราะผลการสอบวัดเป็นการสรุปถึงความสามารถโดยส่วนรวมของผู้เรียนในวิชานั้น ๆ

พวงรัตน์ ทวีรัตน์ (2530 : 41-45) ได้กล่าวถึงลักษณะของแบบทดสอบที่ดีดังนี้

1. มีคุณภาพด้านความเที่ยงตรง (Validity) หมายถึง แบบทดสอบสามารถวัดได้ตรงตามจุดประสงค์ที่ต้องการ ความเที่ยงตรงแบ่งเป็น 4 ลักษณะ คือ

1.1 ความเที่ยงตรงตามเนื้อหา (Content Validity) หมายถึงแบบทดสอบที่สามารถวัดได้สอดคล้องและครอบคลุมเนื้อหาที่ต้องการจะวัด ข้อสอบที่มีความเที่ยงตรงตามเนื้อหาหมายความว่า ข้อสอบนั้นประกอบด้วยข้อคำถามที่ถามเนื้อหาได้ตรงตามที่ระบุไว้ในตารางวิเคราะห์หลักสูตรอย่างครบถ้วน

1.2 ความเที่ยงตรงตามโครงสร้าง (Construct Validity) หมายถึง คุณสมบัติของแบบทดสอบที่สามารถวัดได้สอดคล้องตรงตามที่กำหนดไว้ในทฤษฎี ในกรณีที่เป็นข้อสอบ หมายถึงข้อสอบที่สร้างได้ครอบคลุมพฤติกรรม ตามที่วิเคราะห์ได้ในตารางวิเคราะห์หลักสูตรของรายวิชานั้น ๆ ถ้าพิจารณาจากจุดมุ่งหมายเชิงพฤติกรรมก็คือ ดูจากพฤติกรรมที่คาดหวังและเกณฑ์

1.3 ความเที่ยงตรงตามสภาพ (Concurrent Validity) หมายถึง คุณสมบัติของแบบทดสอบที่ให้ผลการวัดสอดคล้องกับสภาพที่เป็นจริงในขณะนั้นของผู้สอบ ซึ่งดูได้จากการสังเกตหรือการสอบภาคปฏิบัติ

1.4 ความเที่ยงตรงเชิงพยากรณ์ (Predictive Validity) หมายถึง คุณสมบัติของแบบทดสอบที่ช่วยให้สามารถทำนายผลในอนาคตได้ถูกต้อง แบบทดสอบที่มีความเที่ยงตรงเชิงพยากรณ์สูงคือ มีคะแนนจากแบบทดสอบทั้งฉบับนั้นสอดคล้องกับคะแนนผลการเรียนในอนาคต

การที่จะสร้างแบบทดสอบให้มีความเที่ยงตรง คือ แบบทดสอบนั้นจะต้องถามให้ครอบคลุม (Comprehensive) หลักสูตรที่กำหนดไว้ ซึ่งมีลักษณะดังนี้

1. ถามทุกเรื่อง ทุกเนื้อหาที่มีในหลักสูตร
2. ถามพฤติกรรมการเรียนรู้ครบถ้วนตามจุดมุ่งหมายของหลักสูตร
3. ถามแต่ละเนื้อหาและพฤติกรรมอย่างได้สัดส่วนกัน พฤติกรรมใดมีความสำคัญมากเน้นมาก ก็ควรถามมากข้อ ถ้าสำคัญน้อยก็ถามข้อน้อย

2. มีคุณภาพด้านความเชื่อมั่น (Reliability) หมายถึง แบบทดสอบที่สามารถให้ผลการวัดได้คงที่ ไม่ว่าจะนำเครื่องมือขึ้นไปสอบวัดกี่ครั้งก็ตาม แบบทดสอบที่มีความเชื่อมั่น หมายถึง แบบทดสอบที่ให้ผลการวัดในแต่ละครั้งสอดคล้องกัน เช่น ในการสอบวัด 2 ครั้ง คนที่ได้คะแนนสูงในครั้งแรกจะได้คะแนนสูงในครั้งที่สอง คนที่ได้คะแนนต่ำในครั้งแรกก็จะได้คะแนนต่ำในครั้งที่สองเช่นกัน

การสร้างแบบทดสอบให้มีความเชื่อมั่นสูงก็คือ ข้อคำถามของแบบทดสอบนั้นจะต้องถามในสิ่งที่ควรถาม คือ ถามพฤติกรรมขั้นสูงและมีจำนวนมากพอที่จะครอบคลุมเนื้อหาในวิชานั้น ๆ

3. ความเป็นปรนัย (Objectivity) คือมีลักษณะ 3 ประการ ได้แก่

3.1 คำถามมีความแจ่มแจ้งชัดเจน

3.2 การตรวจให้คะแนนชัดเจนทำให้ผู้ตรวจไม่ว่าใครก็ตามตรวจให้คะแนนได้ตรงกัน

3.3 การแปลความหมายของคะแนนตรงกัน กล่าวคือ คะแนนที่ได้ บอกลักษณะภาพของผู้สอบได้ตรงกัน

แบบทดสอบปรนัย หรืออัตนัย เช่น ข้อสอบความเรียงสามารถสร้างให้มีคุณลักษณะ ทั้ง 3 ประการดังกล่าว แบบทดสอบนั้นก็จะเป็นปรนัยได้เท่าเทียมกัน

4. มีการถามลึก (Searching) หมายถึง ไม่ถามเพียงแค่พฤติกรรมขั้นความรู้ความจำ ตามตำรา หรือถามตามที่ครูสอน แต่ต้องพยายามถามพฤติกรรมขั้นสูงกว่าความรู้ความจำ ได้แก่ ถามพฤติกรรมความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า แต่ถ้าจำเป็นต้องถามความรู้ความจำ ก็ควรถามสิ่งที่เป็นความคิดรวบยอด ถ้าข้อคำถามสามารถวัดพฤติกรรมขั้นสูงได้มากเท่าใด แบบทดสอบนั้นก็จะมีคุณค่ามากขึ้นเท่านั้น เพราะสามารถนำผลการสอบมาใช้ในการพัฒนาสมรรถภาพทางสมองของผู้เรียนให้ก้าวหน้ากว่าเดิมได้ดี

5. มีความยุติธรรม (Fair) หมายถึง ข้อคำถามของแบบทดสอบนั้นจะต้องไม่มีช่องทางแนะให้เด็กฉลาดใช้ไหวพริบในการเอาได้ถูก ไม่เปิดโอกาสให้คนเกียจคร้านที่ดูตำราลอก ๆ ตอบได้ คือ ต้องเป็นแบบทดสอบที่ไม่ลำเอียงติดกลุ่มหนึ่งกลุ่มใดโดยเฉพาะ การที่จะให้แบบทดสอบมีความยุติธรรมหรือความเสมอภาคได้ ข้อสอบนั้นจะต้องถามให้ครอบคลุมเนื้อหาและพฤติกรรมทุกประเภทของวิชานั้น ๆ

6. มีลักษณะช่วยเป็นเยี่ยงอย่างในทางดี (Exemplary) หมายถึงแบบทดสอบนั้นจะต้องประกอบด้วยคำถามที่จะสร้างแบบอย่างที่ดีให้แก่ผู้เรียน ไม่ควรถามสิ่งที่เป็นตัวอย่างที่ไม่เหมาะสมไม่ควรปฏิบัติตาม เพราะในช่วงเวลาของการสอบนั้น ผู้สอบมีโอกาสที่จะเรียนรู้จากข้อสอบได้จึงควรถามแต่สิ่งที่จะนำไปเป็นแบบอย่างที่ดีงามจึงจะเป็นการดี เช่น

คำถาม 1 “สิ่งใดที่สอบได้โดยไม่ผิดกฎหมาย” (บุหรี กัญชา ผีน)

คำถาม 2 “การสอบบุหรีให้โทษอย่างไร”

คำถาม 1 เป็นคำถามที่ไม่ควรถาม ควรเลี่ยงไปถามคำถาม 2 จะเหมาะสมกว่า
เป็นต้น

7. มีอำนาจจำแนก (Discrimination) หมายถึง แบบทดสอบนั้นจะต้องประกอบด้วยคำถามที่สามารถจำแนกผู้สอบออกเป็นประเภท ๆ ได้ทุกระดับอย่างถี่ถ้วน ตั้งแต่อ่อนสุดจนถึงเก่งสุด

8. มีความยากพอเหมาะ (Difficulty) หมายถึงแบบทดสอบนั้นจะต้องไม่ยากเกินไปและง่ายเกินไป ผลการสอบโดยเฉลี่ยควรเท่ากับหรือสูงกว่า 50 % ของคะแนนเต็มเล็กน้อย นั่นคือข้อสอบที่ยากมากถือว่าไม่มีประโยชน์ เพราะไม่สามารถเร้าผู้สอบให้แสดงคุณลักษณะที่ต้องการวัดออกมาได้ เพราะคนเก่งก็ยังไม่สามารถทำได้ ในทำนองเดียวกันแบบสอบที่ง่ายมากก็ถือว่าไม่มีประโยชน์ เพราะทั้งคนเก่งคนอ่อนสามารถทำได้เหมือนกันหมด ทำให้ไม่มีอำนาจจำแนก ดังนั้นแบบทดสอบจึงควรมีความยากพอเหมาะในแต่ละข้อคำถามและโดยส่วนรวมทั้งฉบับ

9. มีความเฉพาะเจาะจง (Definite) หมายถึง แบบทดสอบนั้นจะต้องประกอบด้วยคำถามที่มีความชัดเจน ไม่คลุมเครือจนผู้สอบตีความหมายไปคนละอย่าง คำถามประเภททวนสองแง่สองมุมไม่ควรใช้คำถามนั้น แต่จะต้องให้ผู้สอบเข้าใจอย่างแจ่มแจ้งว่าต้องการถามในแง่มุมใดเพื่อผู้สอบที่มีความสามารถในเรื่องนั้นอย่างแท้จริงจะต้องตอบได้ถูก

10. มีประสิทธิภาพ (Efficiency) หมายถึง แบบทดสอบนั้นจะต้องให้ผลการวัดที่เที่ยงตรงและเชื่อถือได้มากที่สุดในขณะที่ใช้เวลา แรงงาน และเงินทุนในการสร้างอย่างประหยัดที่สุด การสร้างข้อสอบให้มีประสิทธิภาพควรคำนึงถึงในเรื่องต่อไปนี้

10.1 ลักษณะคำถาม ควรเป็นคำถามที่ถามพฤติกรรมขั้นสูงให้มากข้อ ตลอดจนถามแต่สิ่งที่มีความสำคัญที่จะเป็นตัวแทนของมวลความรู้ในวิชานั้น ๆ

10.2 ความเหมาะสมของจำนวนข้อกับเวลา แบบทดสอบนั้นไม่ควรให้มีจำนวนข้อมากไป ควรมีจำนวนข้อพอเหมาะแต่มีความครอบคลุมในเนื้อหาของวิชานั้น ๆ และเวลาที่ให้ทำก็เหมาะสมไม่มากจนเกินไป

10.3 ความถูกต้องเรียบร้อยของตัวข้อสอบ คือเป็นแบบทดสอบที่พิมพ์ถูกต้องชัดเจนไม่มีหน้าว่าง ซึ่งสิ่งเหล่านี้ถ้ามีความบกพร่องจะมีผลทำให้แบบทดสอบขาดประสิทธิภาพได้

กระบวนการสร้างแบบทดสอบ

ในการสร้างแบบทดสอบให้มีประสิทธิภาพนั้น พวงรัตน์ ทวีรัตน์ (2530 : 47-56)

ได้กล่าวถึงขั้นตอนในการสร้างแบบทดสอบไว้ว่ามี 6 ขั้นตอนที่สำคัญคือ

1. ขั้นวางแผน
2. ขั้นเตรียมงาน
3. ขั้นลงมือปฏิบัติ
4. ขั้นประเมินหรือตรวจสอบคุณภาพ
5. ขั้นจัดพิมพ์
6. การจัดทำคู่มือการใช้

1. ขั้นวางแผน

โดยทั่วไปในการสร้างแบบทดสอบจะต้องมีการวางแผนล่วงหน้าเสมอ ซึ่งในขั้นของการวางแผนนั้นควรปฏิบัติดังนี้

1.1 กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบทุกครั้งจะต้องกำหนดจุดมุ่งหมายให้ชัดเจน และแน่นอนว่าเพื่อวัตถุประสงค์ใด เช่น

- เพื่อวัดความเจริญงอกงามในการเรียนรู้ของนักเรียนในแต่ละเนื้อหาหรือบทเรียน
- เพื่อหาข้อบกพร่องในการเรียนคือการสอนเพื่อวินิจฉัย
- เพื่อนำผลมาใช้ในการลำดับความสามารถของนักเรียน เป็นต้น

การทราบจุดมุ่งหมายในการสอบวัดจะทำให้ทราบถึงลักษณะของข้อสอบที่จะนำไปใช้ เช่น

- ในกรณีที่ต้องตรวจสอบความรู้เดิมหรือความรู้พื้นฐานของผู้เรียน ข้อสอบที่สร้างก็ควรออกเฉพาะเนื้อหาที่เป็นพื้นฐานจำเป็นของวิชานั้นเท่านั้น และไม่ควรเป็นข้อสอบที่ยากเกินไป อาจมีความยากอยู่ประมาณ 60-70 % เป็นต้น
- ในกรณีที่ต้องการค้นหาข้อบกพร่องทางการเรียนของผู้เรียน แบบทดสอบที่สร้างจะแบ่งออกตามเนื้อหาเป็นตอน ๆ ซึ่งแต่ละตอนจะมีจำนวนข้อคำถามมากข้อเพื่อวัดอย่างละเอียดและควรเป็นแบบทดสอบที่ค่อนข้างง่าย
- ในกรณีที่เป็นข้อสอบคัดเลือก ก็ควรเป็นข้อสอบที่ค่อนข้างยาก เป็นต้น

1.2 กำหนดเนื้อหาและพฤติกรรมที่ต้องการวัด ขั้นนี้เป็นการกำหนดรายละเอียดของหัวข้อเนื้อหา และพฤติกรรมที่ต้องการวัด ซึ่งถ้าสามารถกำหนดขอบข่ายของเนื้อหาและพฤติกรรมที่จะออกข้อสอบได้เหมาะสมก็จะช่วยให้ได้ข้อสอบที่มีความเที่ยงตรง การที่จะสามารถกำหนดขอบเขตของเนื้อหาและพฤติกรรมที่ต้องการวัดได้เป็นอย่างดีนั้นต้องอาศัยการศึกษาหลักสูตรและการสร้างตารางวิเคราะห์หลักสูตร

1.3 กำหนดชนิดและรูปแบบข้อสอบ เนื่องจากการสอบวัดหรือประเมินผลผู้เรียนนั้นสามารถใช้เครื่องมือได้หลายชนิด ซึ่งแต่ละชนิดแต่ละรูปแบบก็เหมาะที่จะใช้วัดพฤติกรรมที่แตกต่างกัน ในการสอบวัดจึงต้องระมัดระวังในการเลือกใช้ชนิดและรูปแบบของแบบทดสอบให้เหมาะสม

1.4 กำหนดส่วนประกอบอื่น ๆ ที่จำเป็นในการออกข้อสอบและในการสอบ ได้แก่ การกำหนดเวลาในการออกข้อสอบ บุคลากรในการออกข้อสอบ เวลาในการสอบ วิธีการสอบ กระดาษคำตอบวิธีการตรวจให้คะแนน วิธีตรวจสอบคุณภาพของข้อสอบ และวิธีรายงานผลการสอบ เป็นต้น

2. ขั้นเตรียมงาน

ขั้นนี้เป็นการเตรียมสิ่งของต่าง ๆ ที่จะต้องใช้และเอื้ออำนวยต่อการสร้างแบบทดสอบ ได้แก่

2.1 หลักสูตร หนังสือแบบเรียน แผนการสอน คู่มือครู

2.2 ทำการวิเคราะห์หลักสูตรรายวิชาที่จะออกข้อสอบ พร้อมกับบันทึกผลในตารางวิเคราะห์หลักสูตร

2.3 อุปกรณ์ต่าง ๆ สำหรับการพิมพ์ การอัดสำเนา

2.4 กระดาษคำตอบและอื่น ๆ

3. ขั้นลงมือปฏิบัติ

ขั้นนี้เป็นขั้นลงมือเขียนข้อสอบซึ่งมีหลักปฏิบัติดังนี้

3.1 ผู้เขียนข้อสอบควรจะทราบคุณลักษณะของแบบทดสอบที่ดี คุณลักษณะที่จำเป็นสำหรับผู้เขียนข้อสอบและเทคนิคการเขียนข้อสอบชนิดต่าง ๆ เทคนิคการเขียนข้อสอบวัดพฤติกรรมด้านความรู้และความคิด

3.2 การเขียนข้อสอบจะต้องยึดผลวิเคราะห์ในตารางวิเคราะห์หลักสูตรเป็นหลัก

3.3 หากมีปัญหาด้านเนื้อหาหรือด้านเทคนิควิธีการเขียนข้อสอบ หากเขียนในรูปแบบของคณะกรรมการก็ควรมีการประชุมตกลงกัน หากเขียนข้อสอบคนเดียวก็ควรปรึกษาผู้รู้และผู้เชี่ยวชาญในการเขียนข้อสอบหรือนักวัดผล

4. ขั้นประเมินหรือตรวจสอบคุณภาพ

ขั้นนี้มีวัตถุประสงค์เพื่อนำผลไปปรับปรุงข้อสอบ ซึ่งแยกเป็นขั้น ๆ ดังนี้

4.1 ขั้นประเมินเบื้องต้น เป็นขั้นของการวิจารณ์ข้อสอบ โดยอาจจะให้บุคคลอื่นหรือผู้เชี่ยวชาญเป็นผู้พิจารณา ซึ่งเป็นสิ่งจำเป็น โดยเฉพาะข้อสอบแบบปรนัยเลือกตอบ เป็นการพิจารณาปรับปรุงข้อคำถามหรือตัวเลือกให้ดีขึ้น การวิจารณ์ข้อสอบเป็นการพิจารณาในเรื่องต่อไปนี้

1. ข้อคำถามวัดในสิ่งที่ต้องการวัดหรือไม่
2. ข้อคำถามมีคำตอบถูกที่แน่นอนเพียงคำตอบเดียวหรือไม่
3. ข้อคำถามชัดเจนเข้าใจตรงกันหรือไม่
4. ข้อคำถามใช้ภาษารัดกุมเหมาะสมกับระดับชั้นของนักเรียนหรือไม่
5. ในกรณีที่เป็นข้อสอบเลือกตอบ ควรพิจารณาสิ่งต่อไปนี้ด้วย
 - ตัวลวงเหมาะสมหรือไม่
 - การเรียงคำถามเรียงถูกต้องตามหลักหรือไม่ เช่น เรียงตามลำดับเนื้อหา เรียงจากง่ายไปหายาก เป็นต้น
 - การเรียงตัวเลือกในแต่ละข้อเรียงเหมาะสมสวยงาม หรือไม่

ข้อสอบที่ผ่านการพิจารณา ผู้เขียนข้อสอบจะต้องนำไปปรับปรุงแก้ไข เพื่อให้ได้ข้อสอบที่มีคุณภาพในเบื้องต้นก่อนจะนำไปทดลองจริง

4.2 ขั้นตรวจสอบคุณภาพหลังการทดลอง ข้อสอบที่ผ่านการพิจารณาและนำไปปรับปรุงแก้ไขเรียบร้อยแล้วจะนำไปจัดพิมพ์เพื่อนำไปทดลอง (Try out) และเมื่อนำไปทดลองแล้วจะต้องนำมาตรวจคะแนนและทำการตรวจสอบคุณภาพ ซึ่งเป็นการพิจารณาในเรื่องต่อไปนี้

1. ความยากง่าย (Difficulty)
2. อำนาจจำแนก (Discrimination)

การวิเคราะห์ข้อสอบ (Item analysis) ผลการตรวจสอบจะทำให้ทราบว่าข้อสอบใดยากหรือง่ายเกินไป ข้อสอบข้อใดจำแนกเด็กเก่งเด็กอ่อนออกจากกันได้หรือไม่ และทำให้ทราบว่าข้อสอบข้อใดตัวเล็กลึกหรือไม่ดี ซึ่งจะช่วยให้ปรับปรุงข้อสอบได้ตรงจุด

3. ค่าสถิติพื้นฐานของข้อสอบ ได้แก่ คะแนนเฉลี่ยและความแปรปรวน คะแนนเฉลี่ยจะทำให้ทราบว่า ข้อสอบทั้งฉบับมีความยากง่ายอยู่ในระดับใด ส่วนค่าความแปรปรวนจะทำให้ทราบว่ากลุ่มที่ใช้ในการทดลองข้อสอบนั้นมีความสามารถแตกต่างกันหรือกระจายกันมากน้อยเพียงใด ซึ่งจะเป็นโยบายในการปรับปรุงข้อสอบ

ในการตรวจสอบคุณภาพหลังการทดลอง อาจมีการตรวจสอบมากกว่า 1 ครั้ง หากผลการทดลองครั้งแรก ผลการวิเคราะห์รายข้อและรายตัวเลือกไม่ได้ตามเกณฑ์ที่กำหนดไว้ ก็จะต้องนำข้อสอบนั้นมาปรับปรุงและนำไปทดลองใหม่จนกว่าจะได้คุณภาพตามเกณฑ์ที่กำหนด

4.3 ขั้นตรวจสอบขั้นสุดท้าย ข้อสอบที่ผ่านการปรับปรุงแก้ไขจนได้คุณภาพของข้อสอบในเรื่องของความยากและอำนาจจำแนกตามเกณฑ์ที่กำหนดไว้เรียบร้อยแล้ว จะต้องนำข้อสอบนั้นไปทดสอบกับคนกลุ่มใหญ่ที่ครอบคลุมความสามารถทั้งเก่ง ปานกลาง และอ่อนกระจายกัน แล้วทำการประเมินคุณภาพครั้งสุดท้ายและสร้างเกณฑ์ปกติ การประเมินคุณภาพครั้งสุดท้ายเป็นการหาค่าต่อไปนี้

1. ความเชื่อมั่นและความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบทั้งฉบับ
2. ค่าสถิติพื้นฐานทั้งฉบับ

5. ขั้นจัดพิมพ์

การจัดพิมพ์กระทำภายหลังการประเมินคุณภาพของข้อสอบแบ่งเป็นขั้นย่อย ๆ ดังนี้

5.1 การจัดพิมพ์ขั้นต้น กระทำภายหลังการวิจารณ์ข้อสอบและปรับปรุงแก้ไขขั้นต้นเพื่อที่จะนำไปทดลองใช้ (Try out) ครั้งแรก ในการเตรียมต้นฉบับเพื่อจัดพิมพ์งานที่ต้องกระทำคือ

1. การเรียงข้อสอบอาจจะเรียงลำดับตามเนื้อหาหรือเรียงจากง่ายไปยาก
2. การวางรูปแบบข้อสอบให้พอเหมาะกับหน้ากระดาษ
3. การเขียนคำชี้แจงในการดำเนินการสอบรวมถึงการระบุจำนวนข้อสอบ เวลาในการทำ

และการกรอกหัวกระดาษคำตอบ

4. ความชัดเจนในการพิมพ์และความสะอาดเรียบร้อย
5. การพิสูจน์อักษรต้องกระทำอย่างรอบคอบอย่าให้มีที่ผิด
6. การอัดสำเนาต้องชัดเจนและสะอาด

5.2 การจัดพิมพ์ภายหลังการทดลอง กระทำภายหลังการนำข้อสอบไปทดลองในแต่ละครั้ง และได้หาคุณภาพของข้อสอบและปรับปรุงข้อสอบเป็นที่เรียบร้อยแล้ว จะต้องนำข้อสอบนั้นมาทำต้นฉบับเพื่อจัดพิมพ์ใหม่ ในขั้นนี้จะต้องพิจารณาเรื่องต่อไปนี้

1. จัดเรียงข้อสอบใหม่ โดยเรียงลำดับเนื้อหาและในแต่ละเนื้อหาเรียงตามผลการวิเคราะห์จากข้อง่ายไปหาข้อยาก (สำหรับข้อสอบปรนัยชนิดเลือกตอบ)
2. เรียงตัวเล็กรหัสข้อมาก่อน อาจเรียงจากข้อความสั้นไปหาข้อความยาว (สำหรับข้อสอบปรนัยชนิดเลือกตอบ)
3. ตรวจสอบสำนวนภาษาให้เหมาะสม
4. ตรวจสอบความชัดเจนในการพิมพ์และความสะอาดเรียบร้อย
5. ทำการพิสูจน์อักษรอย่างรอบคอบ ให้มีความถูกต้อง
6. การอัดสำเนาต้องชัดเจนและสะอาด

5.3 การจัดพิมพ์แบบทดสอบฉบับสมบูรณ์ เป็นการพิมพ์ครั้งสุดท้ายหลังจากการหาคุณภาพและปรับปรุงแบบทดสอบเป็นที่เรียบร้อยแล้ว ในขั้นนี้ควรพิจารณาเพิ่มจาก 2 ชั้นแรก ดังนี้

1. ตรวจสอบจำนวนข้อสอบให้ครบถ้วนตามกำหนด
2. สำนวนการเรียงข้อสอบใหม่
3. ตรวจสอบการจัดวางข้อความรูปภาพให้เหมาะสมกับหน้ากระดาษ
4. ตรวจสอบคำชี้แจงการทำข้อสอบอีกครั้งให้มีหัวข้อต่าง ๆ ครบถ้วน ได้แก่ ชื่อวิชาที่สอบ วิธีการสอบ วิธีคิดข้อสอบ ซึ่งควรมีตัวอย่างประกอบกรอกหวักระดาษคำตอบ จำนวนข้อและเวลาที่ให้ทำ

6. การจัดทำคู่มือการใช้

การเขียนคู่มือการใช้ (Test Manual) ประกอบด้วยกรให้รายละเอียดของหัวข้อสำคัญต่อไปนี้

1. วัตถุประสงค์ของแบบทดสอบ
2. วิธีการใช้และวิธีดำเนินการสอบ
3. วิธีการตรวจให้คะแนน
4. การแปลความหมายคะแนนเกณฑ์ปกติ

จุดมุ่งหมายของการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

ในการวัดผลสัมฤทธิ์ทางการเรียนนั้น นิยมวัดกันใน 2 องค์ประกอบคือ วัดตามจุดมุ่งหมายและตามลักษณะเนื้อหาวิชาเรียน ซึ่งมีรายละเอียดดังนี้

1. การวัดด้านการปฏิบัติ เป็นการตรวจสอบความรู้ความสามารถของผู้เรียน โดยให้ลงมือปฏิบัติจริงปรากฏเป็นผลงานออกมาให้วัดและสังเกตได้ การประเมินผลจะพิจารณาที่กระบวนการปฏิบัติ และผลงาน เช่น งานการช่าง การประกอบอาหาร พลศึกษา ศิลปศึกษา ลักษณะของข้อสอบที่ใช้วัดด้านการปฏิบัติเป็นข้อสอบที่เรียกว่า แบบวัดทักษะปฏิบัติ

2. การวัดด้านเนื้อหา เป็นการตรวจสอบความรู้ความสามารถของผู้เรียนในด้านเนื้อหาวิชา และพฤติกรรมความรู้ความสามารถอื่น ๆ ซึ่งเป็นผลมาจากการเรียนการสอน สามารถวัดได้โดยการสอบปากเปล่าและให้เขียนตอบ

2.1 การสอบปากเปล่า เป็นการสอบเพื่อพิจารณาผลเฉพาะอย่างที่เกิดขึ้นเองแต่ละบุคคล เช่น การสอบฟัง พูด และอ่าน ซึ่งต้องการพิจารณาการใช้ถ้อยคำ การแสดงความคิดเห็น ตลอดจนบุคลิกภาพของผู้สอบ การสอบปากเปล่าสามารถสอบวัดได้ละเอียดลึกซึ้ง โดยสามารถเพิ่มเติมและเปลี่ยนแปลงคำถามได้ตลอดเวลา

2.2 การสอบแบบให้เขียนตอบ เป็นการสอบวัดที่ให้ผู้เข้าสอบเขียนตอบเป็นลายลักษณ์อักษรซึ่งมีรูปแบบการเขียนตอบอยู่ 2 รูปแบบ คือ

2.2.1 แบบจำกัดคำตอบ เป็นการสอบวัดที่กำหนดคำตอบมาให้ตอบ ซึ่งได้แก่การสอบวัดโดยใช้ข้อสอบปรนัยนั่นเอง

2.2.2 แบบไม่จำกัดคำตอบ เป็นการสอบวัดที่ผู้เข้าสอบสามารถเขียนคำตอบได้โดยเสรี ซึ่งได้แก่ การสอบวัดโดยใช้ข้อสอบอัตนัย

การวัดผลสัมฤทธิ์ด้านเนื้อหาโดยการเขียนตอบที่นิยมใช้กันอย่างแพร่หลาย เครื่องมือที่ใช้ในการสอบวัดนั้นเรียกว่า “แบบทดสอบวัดผลสัมฤทธิ์” ดังนั้น แบบทดสอบวัดผลสัมฤทธิ์จึงเป็นเครื่องมืออย่างหนึ่งที่ใช้ประเมินผลการเรียนการสอน โดยเฉพาะวัดสมรรถภาพทางสมองเป็นสำคัญ จุดมุ่งหมายของการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนสรุปได้ดังนี้

1. เพื่อจัดตำแหน่ง ผู้เรียน จุดมุ่งหมายของการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วัดและประเมินผลผู้เรียน เพื่อจัดตำแหน่งนั้น จะดำเนินการก่อนจะมีการเรียนการสอนเกิดขึ้น เพื่อจัดผู้เรียนออกเป็นกลุ่ม ๆ ตามขีดความสามารถ กล่าวคือ ผู้เรียนที่มีความสามารถใกล้เคียงกันไว้ด้วยกัน จัดการเรียนการสอนได้เหมาะสมและสอดคล้องกับผู้เรียนแต่ละกลุ่มในทาง

การศึกษา จุดมุ่งหมายการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนเพื่อจัดตำแหน่งนิยามใช้กันใน 2 ลักษณะ คือ เพื่อการจัดจำแนก เช่น แยกผู้เรียนออกเป็นประเภทตามระดับคะแนนและเพื่อการคัดเลือก

2. เพื่อตรวจสอบความก้าวหน้าของผู้เรียน จุดมุ่งหมายของการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วัดและประเมินผู้เรียน เพื่อตรวจสอบความก้าวหน้าของผู้เรียนนั้น เพื่อช่วยให้ผู้เรียนมีการพัฒนา ความเจริญทั้งทางด้านความรู้ความสามารถ ลักษณะนิสัยและทักษะต่าง ๆ ให้เป็นไปตามจุดมุ่งหมายที่กำหนดไว้

ดังนั้น การตรวจสอบความก้าวหน้า นอกจากจะทำให้ผู้เรียนทราบศักยภาพของตนเอง ขณะนั้น ๆ ยังสามารถใช้เป็นแนวทางให้ผู้เรียนพัฒนาพฤติกรรมต่าง ๆ ของตนเองให้เป็นไปตามที่พึงประสงค์ได้ นั่นก็คือ การใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่มีจุดมุ่งหมายเพื่อการเปรียบเทียบ

3. เพื่อปรับปรุงการเรียนการสอน จุดมุ่งหมายของการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วัดและประเมินผลผู้เรียน เพื่อการปรับปรุงการเรียนการสอนควรทำอย่างต่อเนื่อง อาจใช้เมื่อสิ้นสุดการสอนในแต่ละหน่วยย่อย ซึ่งเมื่อพบว่าผู้เรียนคนใดไม่มีความสามารถไม่ผ่านเกณฑ์ของแต่ละจุดประสงค์การเรียนรู้ ผู้สอนก็ควรจะได้ศึกษาว่าผู้เรียนมีข้อบกพร่องหรือจุดอ่อนในเรื่องใด จะได้ทำการแก้ไขข้อบกพร่องนั้น จากนั้น จึงประเมินผลอีกครั้งหนึ่ง

สำหรับจุดมุ่งหมายของการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน เพื่อการปรับปรุงการเรียนการสอน อาจจำแนกตามระยะเวลาของการวัดและการประเมินได้ 3 ระยะ ดังนี้

1. การวัดและการประเมินผลก่อนการเรียนการสอน
2. การวัดและการประเมินผลระหว่างการเรียนการสอน
3. การวัดและการประเมินผลเมื่อสิ้นสุดการเรียนการสอน

4. เพื่อให้คำปรึกษาและแนะแนว จุดมุ่งหมายของการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนเพื่อช่วยให้ครูแนะแนวมีข้อมูลที่เป็นความสามารถของผู้เรียนในด้านการเรียน ซึ่งเป็นประโยชน์สำหรับการวางแผนการศึกษาต่อและการประกอบอาชีพ โดยจะส่งผลให้ผู้เรียนประสบความสำเร็จได้ นั่นก็คือ การใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่มีจุดมุ่งหมายเพื่อการพยากรณ์ นั่นเอง

5. เพื่อสรุปผลการเรียน จุดมุ่งหมายของการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน การสอน ดำเนินการเมื่อสิ้นสุดการเรียนการสอน เพื่อสรุปผลการเรียนรู้ของผู้เรียนทั้งระบบและตัดสินผลคะแนนได้-ตก ลักษณะของข้อสอบควรมีตั้งแต่ง่ายไปจนถึงยาก โดยต้องวัดให้ครอบคลุมทุกจุดประสงค์การเรียนรู้และทุกเนื้อหาของหลักสูตร

2. การประเมินผลและแบบทดสอบอิงเกณฑ์

การประเมินผลแบบอิงเกณฑ์ มีพื้นฐานมาจากทฤษฎีด้านพฤติกรรมของสกินเนอร์ (B.F. Skinner) แห่งมหาวิทยาลัยฮาร์วาร์ด ที่มีความเชื่อว่า การศึกษาเป็นสิ่งสากลสำหรับนักเรียนทุกคน ทุกคนสามารถเรียนได้ และหน้าที่ของครูก็คือ ช่วยให้นักเรียนจำนวนมากได้เรียนรู้ ทฤษฎีนี้เชื่อว่า นักเรียนทุกคนสามารถเรียนบทเรียนใด ๆ ก็ได้ หากให้เวลาและโอกาสเขามากพอ (เบญจวรรณ ผ่องแผ้ว , 2529 : 3)

ปอปแฮมและฮูเซก (Popham and Husek , 1969 : 1) กล่าวว่า การประเมินผลแบบอิงเกณฑ์เป็นการตรวจสอบผลสัมฤทธิ์ของบุคคลในการกระทำพฤติกรรมต่าง ๆ โดยเปรียบเทียบกับเกณฑ์ที่กำหนดไว้

อนันต์ ศรีโสภณ (2525 : 31) กล่าวว่า การวัดแบบอิงเกณฑ์เป็นการตีความหมายของคะแนนผลการสอบของนักเรียนแต่ละคน โดยการนำไปเปรียบเทียบกับเกณฑ์ ซึ่งได้กำหนดไว้ล่วงหน้า

บุญธรรม กิจปรีดาวิสุทธิ (2535 : 31) กล่าวถึง การวัดผลแบบอิงเกณฑ์ (Criterion Referenced Measurement) ว่าเป็นการวัดที่เปรียบเทียบผลกับเกณฑ์มาตรฐานที่กำหนดไว้ โดยมุ่งที่จะประเมินว่า ผู้เรียนมีความรู้หรือไม่มีความรู้อะไรบ้าง

แบบทดสอบอิงเกณฑ์

ความหมายของแบบทดสอบอิงเกณฑ์

ล้วน สายยศ และอังคณา สายยศ (2539 : 20) กล่าวว่า แบบทดสอบอิงเกณฑ์ (Criterion-Referenced Test) เป็นแบบทดสอบที่สร้างขึ้นโดยยึดหลักการเรียนเพื่อความรอบรู้ สอบเพื่อปรับปรุงการเรียนการสอนเป็นหลัก เนื้อหาในแบบทดสอบมักเป็นวงแคบ และเน้นความครอบคลุมเนื้อหาและจุดประสงค์อย่างมาก แบบทดสอบแบบนี้มุ่งสอบ เพื่อพัฒนาความรู้ความสามารถ นั่นคือ คนใดสอบไม่ผ่านเกณฑ์ที่กำหนดก็ร่ำเรียนใหม่ แล้วสอบซ่อมเสริมด้วยแบบทดสอบเดิม ถ้าเกิดการเรียนรู้เพิ่มขึ้นก็จะสามารถสอบผ่านเกณฑ์ได้

อาจสรุปความหมายของแบบทดสอบอิงเกณฑ์ ได้ดังนี้

1. เป็นแบบทดสอบที่สร้างขึ้นตามจุดประสงค์การเรียนรู้ ซึ่งได้มีการกำหนดหรือนิยามไว้ อย่างชัดเจน เพื่อให้สามารถวัดความรู้ความสามารถได้ตรงตามความคาดหวังที่กำหนดไว้ใน จุดประสงค์การเรียนรู้ นั้น ๆ

2. มีการกำหนดเกณฑ์หรือมาตรฐานขั้นต่ำ เพื่อใช้ตัดสินว่า ผู้เรียนมีความรู้ความสามารถ หรือปฏิบัติในเรื่องนั้น หรือที่เรียกว่า รอบรู้ (Master) ตามเจตนาของจุดประสงค์การเรียนรู้แล้ว หรือไม่

จากคำจำกัดความของการประเมินผลแบบอิงเกณฑ์ และแบบทดสอบอิงเกณฑ์ที่กล่าวมาแล้วข้างต้น จะเห็นได้ว่าการประเมินผลแบบอิงเกณฑ์ ต้องอาศัยเกณฑ์ในการแปลความหมาย คะแนนผลการสอบของนักเรียน และการสร้างแบบทดสอบอิงเกณฑ์จะต้องคำนึงถึงจุดประสงค์ การเรียนรู้ เพื่อให้สามารถแปลความหมายของคะแนนผลการสอบได้ว่า ผู้เรียนมีความรอบรู้ใน จุดประสงค์ที่กำหนดไว้มากน้อยเพียงใด หลังจากที่มีการเรียนการสอนไปแล้ว

ความเป็นมาของแบบทดสอบอิงเกณฑ์

แนวความคิดเกี่ยวกับการวัดผลแบบอิงเกณฑ์นั้นมีมาตั้งแต่ ค.ศ. 1951 เมื่อ ฟลานาแกน (Flanagan) ได้อธิบายคำว่า “เกณฑ์ปกติ” และ “มาตรฐาน” นอกจากนี้ ฟลานาแกน ยังได้อธิบาย ถึงการทดสอบแบบอิงเกณฑ์ในลักษณะของผู้เรียนจะแปลหนังสือโดยคิด ไม่เกิน 3 แห่ง (สมศักดิ์ สนิทระเวชญ์, 2525 : 47 อ้างถึงในเสาวณิต ร่มศรี, 2541 : 7)

เกลเซอร์ (Glaser, 1963 : 519-521) เป็นบุคคลแรกที่ใช้คำว่า “การวัดผลแบบอิงเกณฑ์” (Criterion-Referenced Measurement) นอกจากนี้ อีเบล (Ebel) กล่าวว่า ในการทดสอบใดก็ตาม จะมีความหมายสมบูรณ์ก็ต่อเมื่อมีการอ้างอิงถึงเนื้อหาอย่างแจ่มชัด ซึ่งคะแนนจะแปลได้ความว่า อยู่ในระดับใดของเนื้อหาที่ทำการสอบวัด หรือที่หาเรียนรู้เป็นร้อยละเท่าใดของเนื้อหาทั้งหมด จะเห็นได้ว่า แนวคิดเหล่านี้คือส่วนที่เรียกว่า การวัดผลแบบอิงเกณฑ์ นั่นเอง (โกวิท ประวาลพุกษ์, 2523 : 17-18)

ส่วนในประเทศไทย การวัดผลแบบอิงเกณฑ์เริ่มมีการกล่าวถึงเมื่อประมาณปี พ.ศ. 2517 (ส. วาสนา ประวาลพุกษ์, 2517 : 80-85 อ้างถึงใน เสาวณิต ร่มศรี, 2541) และได้มีการตีพิมพ์ บทความลงในวารสารวัดผลการศึกษาฉบับเดือน มกราคม – เมษายน 2523 ได้ลงบทความที่กล่าวว่าด้วยเรื่องของการวัดผลแบบอิงเกณฑ์ หลังจากนั้นกลุ่มผู้ที่สนใจในด้านการวัดผลแบบอิงเกณฑ์ ได้ค้นคว้าเพิ่มเติมตลอดเวลา ในที่สุดมีการนำเอาการวัดผลแบบอิงเกณฑ์บางส่วนเข้าบรรจุใน

ระเบียบการประเมินผลหลักสูตรประถมศึกษา 2521 ตลอดจนหลักสูตรมัธยมศึกษาตอนต้น 2521 และตอนปลาย 2524 จนถึงปัจจุบัน และในเวลาใกล้เคียงกันนั้น การเขียนข้อสอบโดยยึดจุดประสงค์เชิงพฤติกรรม ได้เผยแพร่ไปสู่ครูผู้สอนทุกระดับ การเขียนข้อสอบแนวนี้ถือว่าเป็นข้อสอบแบบอิงเกณฑ์ชนิดหนึ่ง (บุญเชิด ภิญโญนนตพงษ์, 2527 : 9)

ลักษณะของแบบทดสอบอิงเกณฑ์

ได้มีผู้อธิบายลักษณะของแบบทดสอบอิงเกณฑ์ไว้ดังนี้

โกวิท ประวาลพุกษ์ (2523 : 18-19) ได้สรุปลักษณะของแบบทดสอบอิงเกณฑ์ไว้ดังนี้

1. การยึดจุดประสงค์เป็นหลักหรือจะเรียกว่าอิงจุดประสงค์เชิงพฤติกรรมเหล่านั้นเป็นเกณฑ์เพื่อจะบ่งชี้ว่าใครผ่านเกณฑ์ที่กำหนดไว้ในจุดประสงค์แล้วหรือยัง
 2. การขยายจุดประสงค์ให้ครอบคลุมประชากรของพฤติกรรมที่เรียกว่า โดเมน แบบทดสอบในลักษณะนี้ เรียกว่า Domain-Referenced test ซึ่งเมื่อสอบวัดแล้วสามารถอ้างอิงได้ว่า บุคคลมีความสามารถในระดับใด ตามคะแนนมาตรฐานหรือมีความรู้เท่าใด
- สงบ ลักษณะ (2523 : 36-37) ได้กล่าวถึงลักษณะสำคัญของแบบทดสอบอิงเกณฑ์ว่า
1. เป็นแบบทดสอบที่สร้างขึ้นจากพื้นฐานของจุดประสงค์ หรือสิ่งที่ต้องการให้ผู้สอบนั้นปฏิบัติได้ที่ได้รับนิยามไว้อย่างชัดเจน
 2. รายชื่อของแบบทดสอบ จะต้องวัดจุดประสงค์หรือสิ่งที่ต้องการให้ผู้สอบนั้นปฏิบัติได้
 3. คะแนนที่ได้ควรแปลความหมายว่า ผู้สอบสามารถบรรลุตามเกณฑ์ที่กำหนดไว้หรือไม่

อนันต์ ศรีโสภณ (2525 : 190) กล่าวว่า แบบทดสอบอิงเกณฑ์ใช้สำหรับวัดความรู้ความสามารถของนักเรียนแต่ละคน ว่าถึงเกณฑ์ขั้นต่ำตามที่เราคาดหวังไว้หรือไม่ บางทีก็เรียกแบบทดสอบนี้อีกชื่อหนึ่งว่า Mastery test แบบทดสอบนี้จึงเป็นการวัดความรู้และทักษะต่าง ๆ ในตัวนักเรียนว่าถึงเกณฑ์ที่กำหนดไว้หรือไม่ โดยไม่คำนึงถึงความแตกต่างระหว่างนักเรียน

ลำเรียง บุญเรืองรัตน์ (2527 : 83) กล่าวถึงลักษณะของแบบทดสอบที่สร้างขึ้นมาวัดผลสัมฤทธิ์ทางการเรียนตามหลักของการวัดผลแบบอิงเกณฑ์นั้น มักจะสร้างเพื่อใช้กับการเรียนการสอนเพื่อรอบรู้ (Mastery Learning) ซึ่งหลักการนี้บลูม (Bloom) เป็นผู้เสนอไว้ โดยมีแนวคิดดังนี้

1. นักเรียนทุกคนสามารถเรียนรู้สิ่งต่าง ๆ ตามที่สอนสั่งกันในโรงเรียนนั้น ได้อย่างเท่าเทียมกัน ถ้าจัดการเรียนการสอนให้เหมาะสม

2. นักเรียนทุกคนสามารถเรียนรู้ได้ผลสัมฤทธิ์เท่ากัน ก็ต่อเมื่อนักเรียนประเภท “โง่-ช้า” ได้รับเวลาในการเรียนอย่างเพียงพอและได้รับการช่วยเหลือในเรื่องการเรียนอย่างดีจากครูผู้สอน และเพื่อนนักเรียนที่มีพื้นเพทางการศึกษามาอย่างดี

บุญชม ศรีสะอาด (2528 : 2) ได้กล่าวถึงลักษณะของแบบทดสอบอิงเกณฑ์แรกไว้ในลักษณะของการประเมินผลแบบอิงเกณฑ์ สรุปได้ว่า แบบทดสอบอิงเกณฑ์เป็นแบบทดสอบที่สร้างขึ้นมาเพื่อใช้วัดตามจุดประสงค์ที่กำหนดไว้ และคำนึงถึงความเที่ยงตรงตามเนื้อหาเป็นสำคัญ การวิเคราะห์ข้อสอบใช้เกณฑ์ภายนอก เช่น กลุ่มคนผิว กับไม่ใช่คนผิว และคะแนนจะแปลออกมาในรูปของความรอบรู้ ไม่รอบรู้

สมนึก ภัททิยธนี (2533 : 238) ได้สรุปลักษณะของแบบทดสอบอิงเกณฑ์ไว้ว่า

1. เป็นแบบทดสอบที่ใช้เกณฑ์ภายนอกในการประเมิน
2. มีประโยชน์ในการประเมินการเรียนการสอน
3. การวิเคราะห์ข้อสอบต้องคำนึงถึงความเที่ยงตรงตามเนื้อหา และความเที่ยงตรงตามโครงสร้างเป็นหลัก ส่วนค่าความเชื่อมั่น ความยากง่าย อำนาจจำแนก ไม่มีผลกระทบต่อโครงสร้างแบบทดสอบ แต่เน้นการวัดที่จุดประสงค์เป็นสำคัญ
4. ไม่มุ่งวัดนักเรียนกับนักเรียนคนอื่น
5. แบบทดสอบอิงเกณฑ์เหมาะสำหรับใช้ในวงแคบ ๆ กับการประเมินผลย่อยเพื่อนำผลไปปรับปรุงการเรียนการสอน
6. หาเกณฑ์การประเมินต้องอาศัยทฤษฎีในการสร้าง

สรุปได้ว่า แบบทดสอบอิงเกณฑ์มีลักษณะเป็นแบบทดสอบที่วัดผลสัมฤทธิ์ที่สร้างข้อคำถามสอดคล้องกับจุดมุ่งหมายเชิงพฤติกรรมตามเนื้อหาเฉพาะเจาะจง การแปลความหมายของคะแนน ได้จากการนำคะแนนที่เด็กสอบได้ไปเทียบกับเกณฑ์ที่กำหนดไว้ แบบทดสอบอิงเกณฑ์จะมีประโยชน์อย่างยิ่งในการนำไปใช้เป็นแนวทางในการปรับปรุงการเรียนการสอน

การสร้างแบบทดสอบอิงเกณฑ์

นักวัดผลสัมฤทธิ์การและนักวัดผลการศึกษา ได้เสนอแนะแนวทางในการสร้างแบบทดสอบอิงเกณฑ์ไว้ดังนี้

แอร์ราเซียน และเมดัส (Aivasian and Meduas , 1972 : 8) แบ่งขั้นตอนของการสร้างแบบทดสอบอิงเกณฑ์เป็น 4 ขั้นตอน ดังนี้

1. เขียนสมรรถภาพที่จะวัดให้ชัดเจน โดยเขียนในรูปของจุดประสงค์เชิงพฤติกรรมที่สามารถสังเกตได้ และสอบวัดได้
2. กำหนดการรอบรู้ในสมรรถภาพนั้น
3. คัดสถานการณ์เพื่อให้นักเรียนได้แสดงถึงความสามารถที่ต้องการวัดแล้วเขียนข้อสอบตามสถานการณ์นั้น
4. ตัดสินความรอบรู้ของนักเรียน โดยเทียบกับมาตรฐานการปฏิบัติขั้นต่ำสุดตามจุดประสงค์

กรอนลันด์ (Gronlund , 1978 อ้างถึงใน พานิช ศรีงาม , 2530 : 12-13) ได้เสนอแนะแนวทางการสร้างแบบทดสอบอิงเกณฑ์ไว้ดังนี้

1. แบบทดสอบอิงเกณฑ์จะต้องมีการกำหนดขอบเขตของงานเอาไว้อย่างชัดเจนเพราะแบบทดสอบต่าง ๆ จะนำมาใช้ให้ประสบผลสำเร็จได้นั้นจะต้องบอกถึงความเจาะจงของผลการสอบที่ผู้สอบกระทำได้
2. แบบทดสอบอิงเกณฑ์จะต้องมีความชัดเจนในการกำหนดจุดประสงค์ของการเรียนการสอน ในสิ่งที่ต้องการให้ผู้เรียนแสดงออกมาในรูปของพฤติกรรมที่ปฏิบัติได้จริงซึ่งเมื่อดูจากผลการสอบ ก็จะทำให้ทราบทันทีว่านักเรียนสามารถทำอะไรได้บ้าง จากงานที่มอบหมายให้
3. แบบทดสอบอิงเกณฑ์จะต้องมีการกำหนดเกณฑ์ในการปฏิบัติเอาไว้อย่างชัดเจน เพราะการกำหนดเกณฑ์ปฏิบัติมีผลต่อการนำแบบทดสอบอิงเกณฑ์ไปใช้ในห้องเรียน เพื่อที่จะแสดงว่าผู้เรียนมีความรู้ในเนื้อหานั้น ๆ แล้ว
4. แบบทดสอบอิงเกณฑ์ต้องสามารถวัดความเป็นตัวแทนของเนื้อหาในการวัดได้อย่างเพียงพอ โดยสามารถทำได้ดังนี้
 - 4.1 แบ่งเนื้อหาในการเรียนออกเป็นหน่วยย่อย ๆ
 - 4.2 กำหนดโดเมน (Domain) ของเนื้อหาให้มีความชัดเจน
 - 4.3 กำหนดเนื้อหาให้มีความเฉพาะเจาะจงกับกลุ่มผู้เรียน

5. แบบทดสอบอิงเกณฑ์จะต้องสร้างข้อสอบแต่ละข้อให้ตรงกับจุดประสงค์ที่ต้องการวัด เพราะจะสามารถทำให้แปลความหมายของการกระทำของผู้เรียนได้ตามโดเมนของเนื้อหา

6. แบบทดสอบอิงเกณฑ์จะต้องมีการรายงานผลของคะแนนที่แสดงถึงการปฏิบัติได้จริงของผู้เรียนในขอบเขตของงาน (Task) ที่กำหนดเอาไว้อย่างชัดเจน

บุญเชิด ภิญโญนนตพงษ์ (2527 : 38-44) ได้สรุปขั้นตอนการสร้างแบบทดสอบอิงเกณฑ์ไว้ดังนี้

ขั้นที่ 1 วิเคราะห์เนื้อหา และจุดประสงค์ เมื่อต้องการทดสอบเนื้อหาวัตถุประสงค์ใด สิ่งที่ต้องทำประการแรกก็คือ ทำการวิเคราะห์เนื้อหาวัตถุประสงค์ให้เป็นหัวข้อใหญ่ ๆ และในแต่ละหัวข้อต้องทำการวิเคราะห์จุดประสงค์หรือพฤติกรรมที่ต้องการให้เกิดในหัวข้อนั้น ๆ ซึ่งทำให้ได้พฤติกรรมหลักที่ต้องการวัด หรือจุดประสงค์หลัก

ขั้นที่ 2 แปลงจุดประสงค์หลักให้เฉพาะเจาะจง โดยแตกพฤติกรรมหลักให้เป็นจุดประสงค์เชิงพฤติกรรมย่อย ๆ

ขั้นที่ 3 เขียนข้อสอบหรือผลิตข้อสอบตามจุดประสงค์เชิงพฤติกรรม

ขั้นที่ 4 ตรวจสอบข้อสอบ เป็นการตรวจสอบข้อสอบเป็นรายข้อ หลังจากเขียนข้อสอบเสร็จ เพื่อแก้ไขปรับแต่งข้อสอบให้มีความเหมาะสมของภาษาที่ใช้ ว่ามีความชัดเจนรัดกุม ถูกต้อง และความถูกต้องของเนื้อหาวัตถุประสงค์ที่ต้องการทดสอบ

ขั้นที่ 5 วิเคราะห์ข้อสอบ เป็นการหาคุณภาพของข้อสอบด้านความตรงเชิงเนื้อหาและความตรงเชิงโครงสร้างเพื่อคัดเลือกข้อสอบที่มีคุณภาพมารวมฉบับ การหาคุณภาพด้านความตรงเชิงเนื้อหาเป็นการตรวจสอบว่าข้อสอบมีความสอดคล้องกับเนื้อหาวัตถุประสงค์ที่ต้องการวัดหรือไม่โดยใช้จุดประสงค์เชิงพฤติกรรมเป็นเกณฑ์เทียบ การหาคุณภาพด้านความตรงเชิงโครงสร้างเป็นการตรวจสอบว่าข้อสอบสามารถจำแนกการรอบรู้ ไม่รอบรู้ ได้หรือไม่ โดยนำคะแนนสอบของกลุ่มที่รอบรู้กับกลุ่มที่ไม่รอบรู้มาเปรียบเทียบกัน ดังนั้นการหาความตรงเชิงโครงสร้างจึงต้องนำข้อสอบที่มีความตรงเชิงเนื้อหาไปทดสอบกับนักเรียนแล้วนำคะแนนผลการสอบมาวิเคราะห์ดัชนีอำนาจจำแนก

ขั้นที่ 6 คัดเลือกข้อสอบ กำหนดความยาว และคะแนนจุดตัดของแบบทดสอบ การคัดเลือกข้อสอบรวมฉบับเป็นการพิจารณาโดยอาศัยเกณฑ์คุณภาพด้านความตรงเชิงเนื้อหา และความตรงเชิงโครงสร้าง

ขั้นที่ 7 วิเคราะห์แบบทดสอบ เป็นการหาคุณภาพด้านความเชื่อมั่นและความตรงของแบบทดสอบหลังจากที่คัดเลือกข้อสอบมารวมเป็นฉบับแล้ว ส่วนใหญ่การวิเคราะห์แบบทดสอบอิงเกณฑ์ มักทำการตรวจสอบความเชื่อมั่นเพียงอย่างเดียว เพราะความตรงของแบบทดสอบนั้นมักพิจารณาความตรงรายข้อ

อังคณา สายยศ (ม.ป.ป. : 3-8 อ้างถึงใน ประพนอม พันธุ์ไสว , 2530 : 15-16) ได้กล่าวถึงขั้นตอนในการสร้างแบบทดสอบอิงเกณฑ์ ไว้ดังนี้

1. กำหนดจุดมุ่งหมายในการสอน เป็นขั้นที่จะต้องพิจารณาว่าผู้สอนต้องการจะให้นักเรียน เรียนแล้วเกิดพฤติกรรมอะไร ซึ่งผู้สร้างข้อสอบจะรู้ได้จากจุดมุ่งหมายของการสอน
2. สร้างตารางกำหนดรายละเอียด (Table of Specification) นำจุดมุ่งหมายที่ได้จากข้อ 1 มาวิเคราะห์ให้สัมพันธ์กับเนื้อหาที่จะสอน เพื่อที่จะดูรายละเอียดของจุดมุ่งหมายของการสอน
3. เขียนจุดประสงค์เชิงพฤติกรรมหรือกำหนดรายละเอียดของโดเมน ในการเขียนจุดประสงค์เชิงพฤติกรรมนี้จะเขียนจากพฤติกรรมหลักในตารางวิเคราะห์หลักสูตร โดยนำเอาเนื้อหาและพฤติกรรมหลักมาเขียนเป็นจุดประสงค์เชิงพฤติกรรม
4. การเขียนข้อคำถาม นำจุดประสงค์เชิงพฤติกรรมที่เขียนในขั้นที่ 3 มาเขียนเป็นข้อสอบ
5. การพิจารณาข้อคำถาม ขั้นนี้เป็นขั้นที่เอาข้อคำถามที่เขียนตามจุดประสงค์เชิงพฤติกรรมนั้นมาพิจารณาดังนี้
 - 5.1 พิจารณาว่าข้อคำถามแต่ละข้อเขียนสอดคล้องกับจุดประสงค์เชิงพฤติกรรมหรือไม่ โดยให้ผู้เชี่ยวชาญทางด้านวัดผลเป็นผู้พิจารณา
 - 5.2 พิจารณาความครอบคลุมของข้อคำถามในแต่ละจุดประสงค์เชิงพฤติกรรมคือให้ผู้เชี่ยวชาญทางด้านเนื้อหาพิจารณาว่า จำนวนข้อคำถามในแต่ละจุดประสงค์เชิงพฤติกรรมนั้นมีจำนวนเพียงพอหรือไม่
6. คัดเลือกข้อคำถาม ขั้นนี้จะต้องคัดเลือกข้อคำถามที่มีความสอดคล้องและครอบคลุมจุดประสงค์เชิงพฤติกรรมเท่านั้น ข้อที่ไม่สอดคล้องจะตัดออก
7. หาคุณภาพของแบบทดสอบ ขั้นนี้จะต้องนำแบบทดสอบไปทดลองสอบกับนักเรียน แล้วนำมาหาคุณภาพทางด้านความยากง่าย อำนาจจำแนก ความเชื่อมั่น และความตรงของแบบทดสอบต่อไป

จุดมุ่งหมายของแบบทดสอบอิงเกณฑ์

สมศักดิ์ สิ้นธุระเวช (2521 : 74) กล่าวถึงจุดมุ่งหมายของแบบทดสอบอิงเกณฑ์ว่ามีจุดมุ่งหมายอย่างน้อยที่สุด 4 ประการ คือ

1. เพื่อใช้เป็นเครื่องมือในการประเมินจุดประสงค์
2. เพื่อใช้ในการปรับปรุงการเรียนการสอน โดยดูความก้าวหน้าของเด็ก หากพบเด็กคนใดบกพร่องหรือล่าช้าในเรื่องใดก็จะให้หาทางช่วยเหลือ
3. เพื่อใช้เป็นเครื่องมือในการพิจารณารูปแบบของการสอน เช่น จุดมุ่งหมายของการสอน วิธีดำเนินการสอน
4. เพื่อใช้ในการพิจารณาว่าเมื่อนักเรียนได้เรียนจบในแต่ละหน่วยการเรียนแล้วนักเรียนรอบรู้ตามเกณฑ์ที่ตั้งไว้หรือไม่

ไพศาล หวังพานิช (2526 : 187-188) กล่าวว่าจุดมุ่งหมายสำคัญของการวัดผลแบบอิงเกณฑ์ คือ มุ่งสอบวัดเพื่อตรวจสอบระดับการเรียนรู้ของผู้เรียนว่าบรรลุถึงเป้าหมายหรือเกณฑ์ที่ตั้งไว้หรือไม่ การสอบจึงมุ่งเน้นเรื่องการวินิจฉัยความสามารถของผู้เรียน เพื่อใช้ผลสำหรับการปรับปรุงการเรียนการสอน การสอนตามจุดมุ่งหมายดังกล่าวจึงเป็นการตอบคำถามว่าใครเรียนได้ถึงเกณฑ์ที่ตั้งไว้และใครควรต้องแก้ไขการเรียน หรือควรได้รับการปรับปรุงซ่อมเสริมการเรียนบ้าง คำถามเหล่านี้สามารถตอบได้โดยการนำคะแนนของผู้เรียนไปเทียบกับเกณฑ์ที่ตั้งไว้ โดยยอมรับว่าคะแนนคือสิ่งที่แทนประสิทธิภาพในการเรียนของแต่ละบุคคล

อารมณั์ เพชรชื่น (2527 : 23) กล่าวว่าจุดมุ่งหมายของการทดสอบแบบอิงเกณฑ์ คือ เพื่อวัด เพื่อตรวจสอบว่า ผู้เรียนรู้ตามจุดประสงค์หรือไม่ เพื่อวินิจฉัยประเมินการเรียนการสอน และปรับปรุงการเรียนการสอน

บุญชม ศรีสะอาด (2528 : 1) กล่าวถึงวัตถุประสงค์ของการประเมินผลแบบอิงเกณฑ์ว่าการประเมินผลแบบนี้ มีวัตถุประสงค์เพื่อวินิจฉัยผู้เรียนว่าบรรลุจุดประสงค์ที่วางไว้หรือไม่ ผู้เรียนคนใดยังไม่บรรลุจุดประสงค์ที่วางไว้ ก็จะได้หาทางปรับปรุงแก้ไข เช่น จัดให้มีการเรียนซ่อมเสริม เป็นต้น จะเห็นได้ว่า การประเมินผลแบบอิงเกณฑ์ เป็นผลจากความเชื่อในทฤษฎีการเรียนรู้เพื่อรอบรู้ (Learning for Mastery)

พวงรัตน์ ทวีรัตน์ (2530 : 32) กล่าวว่าจุดมุ่งหมายของการทดสอบอิงเกณฑ์ คือเป็นการตรวจสอบดูว่าใครเรียนได้ถึงเกณฑ์และใครยังเรียนได้ไม่ถึงเกณฑ์ และถ้าใครยังไม่ถึงเกณฑ์ควรได้รับการปรับปรุงแก้ไขต่อไป เช่น อาจให้มีการซ่อมเสริม

สรุปได้ว่า จุดมุ่งหมายของแบบทดสอบอิงเกณฑ์คือ เพื่อตรวจสอบว่าผู้เรียนมีความสามารถผ่านเกณฑ์ที่กำหนดไว้หรือไม่ เพื่อนำข้อมูลไปปรับปรุงการเรียนการสอนต่อไป

3. คะแนนจุดตัด

ความหมายของคะแนนจุดตัด

นักการศึกษาและนักวัดผล เรียก คะแนนจุดตัด (Cut-off Score) ในชื่อต่าง ๆ กัน เช่น เกณฑ์ (Criteria) มาตรฐาน (Standard) คะแนนผ่าน (Passing Score) ระดับความรอบรู้ (Mastery Level) หรือ ความสามารถต่ำสุด (Minimal Competence) โดยให้ความหมายไว้ต่าง ๆ กัน ดังนี้

เบิร์ก (Berk, 1986 : 138 อ้างถึงใน อรอนงค์ บริเพชร, 2543 : 15) กล่าวว่า คะแนนจุดตัด หมายถึง จุดที่ใช้แบ่งคะแนนออกเป็น 2 ส่วน ส่วนหนึ่งเป็นคะแนนของผู้สอบที่จัดว่าเป็นผู้รอบรู้ อีกส่วนหนึ่งเป็นคะแนนของผู้สอบที่จัดว่าเป็นผู้ไม่รอบรู้ เนื่องจากคะแนนจุดตัดเป็นค่าที่ได้จากการสังเกต เพราะฉะนั้นการกำหนดคะแนนจุดตัดจึงย่อมมีความคลาดเคลื่อนมาตรฐาน (Standard Error) สำหรับความคลาดเคลื่อนนี้ เบิร์ก แบ่งเป็น 2 ประเภท คือ

1. ผู้สอบที่มีความรอบรู้อย่างแท้จริง แต่ถูกจัดประเภทไม่มีความรอบรู้ ความคลาดเคลื่อนนี้เกิดจากการกำหนดคะแนนจุดตัดสูงเกินไป จึงทำให้ผู้สอบที่มีความรอบรู้ไม่ผ่านเกณฑ์ที่กำหนด

2. ผู้สอบที่ไม่มีความรอบรู้อย่างแท้จริงถูกจัดประเภทมีความรอบรู้ ความคลาดเคลื่อนนี้เกิดจากการกำหนดคะแนนจุดตัดต่ำเกินไป จึงทำให้ผู้สอบที่ไม่มีความรอบรู้ผ่านเกณฑ์ที่กำหนด

แฮมเบิลตัน (Hambleton, 1978 : 279 อ้างถึงใน อรอนงค์ บริเพชร, 2543 : 15) กล่าวว่า มาตรฐาน หมายถึง คะแนนที่ได้จากการสอบที่ใช้ในการแยกผู้สอบออกเป็น 2 กลุ่ม ตามระดับความสามารถที่แตกต่างกัน คือ เป็นกลุ่มที่รอบรู้ และกลุ่มที่ไม่รอบรู้

อังคณา สายยศ (2525 : 70) ให้ความหมายของคะแนนจุดตัดว่า หมายถึง คะแนนที่น้อยที่สุดที่นักเรียนจะต้องทำได้ในการที่จะได้รับการตัดสินให้เป็นผู้รอบรู้

บุญเชิด ภิญโญนนตพงษ์ (2527 : 110) กล่าวว่า เกณฑ์ หรือ มาตรฐาน หมายถึง คะแนนจุดตัดของแบบทดสอบที่ใช้แบ่งผู้เรียนออกเป็นผู้สอบได้สอบตก

ดำรง ศิริเจริญ (2529 : 139) กล่าวว่า คะแนนจุดตัด หมายถึง คะแนนที่น้อยที่สุดที่นักเรียนจะต้องทำได้ในการที่จะได้รับการตัดสินให้เป็นผู้รอบรู้ คะแนนเกณฑ์อาจจะอยู่ในรูปของจำนวนข้อที่ผู้เรียนทำถูกในแต่ละจุดประสงค์

กล่าวโดยสรุปแล้ว คะแนนจุดตัด คือ คะแนนที่เป็นเกณฑ์ต่ำสุดที่ใช้ตัดสินให้ผู้สอบเป็นผู้รอบรู้หรือไม่รอบรู้

จากความหมายของคะแนนจุดตัดดังกล่าวข้างต้น จะเห็นว่า การกำหนดคะแนนจุดตัดเป็นสิ่งจำเป็น ซึ่งสืบเนื่องมาจากการวัดผลแบบอิงเกณฑ์จะต้องทำการประเมินให้สอดคล้องกับวัตถุประสงค์การเรียนการสอน (Objective – Based Instruction Program) ที่กำหนดขึ้น หากพิจารณาในเชิงทฤษฎีจะเห็นว่า ในการที่จะปรับปรุงคุณภาพของการเรียนการสอน จะต้องเริ่มจาก

1. การกำหนดวัตถุประสงค์ของหลักสูตร
2. การจัดการเรียนการสอนให้สอดคล้องกับวัตถุประสงค์
3. การประเมินผลการเรียนการสอน

และจากการศึกษาพบว่า การจัดการเรียนการสอนตามวัตถุประสงค์ที่กำหนดจะทำให้ผลสัมฤทธิ์ทางการเรียนของผู้เรียนสูงกว่าการจัดการเรียนการสอนที่ไม่เป็นไปตามวัตถุประสงค์ เพราะฉะนั้นในการที่จะประเมินผลสัมฤทธิ์ทางการเรียนจึงมีความจำเป็นที่จะต้องกำหนดคะแนนจุดตัด

แนวทางในการกำหนดคะแนนจุดตัด

เบิร์ก (Berk , 1976 อ้างถึงใน เขวดี วิบูลย์ศรี , 2540 : 51-53) ได้เสนอกรอบในการจัดประเภทแนวคิดของการกำหนดเกณฑ์มาตรฐานไว้ 3 เกณฑ์ ดังนี้

เกณฑ์ที่ 1 แนวคิดของการกำหนดเกณฑ์มาตรฐาน โดยอิงธรรมชาติของการเรียนรู้ในวัดผลสัมฤทธิ์ต่าง ๆ ซึ่งสามารถแบ่งออกเป็น 2 สภาวะ คือ

1. สภาวะตามธรรมชาติของการเรียนรู้ที่สามารถแยกออกจากกันอย่างชัดเจน โดยแบ่งได้เป็น 2 ลักษณะ คือ การเรียนรู้ได้ทั้งหมดกับการไม่ได้เรียนรู้เลย ธรรมชาติเช่นนี้มีความหมายว่า ในการเรียนรู้บางวัดผลสัมฤทธิ์นั้น ถ้าผู้เรียนสามารถเรียนรู้ก็จะสามารถเข้าใจได้ทั้งหมดในคราวเดียว แต่ถ้ายังไม่เรียนรู้ก็จะไม่สามารถเข้าใจอะไรได้เลยในวัดผลสัมฤทธิ์นั้น ภาวะดังกล่าวจะปรากฏออกมาในรูปของคะแนน คือ ทำคะแนนจากแบบทดสอบในเนื้อหานั้น ๆ ได้ทั้งหมด หรือทำคะแนนจากแบบสอบในเนื้อหานั้น ๆ ไม่ได้เลย ซึ่งจะแยกผู้สอบว่าเป็นผู้รอบรู้หรือไม่รอบรู้ออกจากกันตามคุณลักษณะที่ปรากฏเป็นเกณฑ์มาตรฐาน

2. สภาวะตามธรรมชาติของการเรียนรู้อย่างต่อเนื่อง ในลักษณะที่ค่อย ๆ เป็นค่อย ๆ ไป กล่าวคือ ค่อย ๆ สะสมพอกพูนขึ้นทีละเล็กทีละน้อย และมากขึ้นเรื่อย ๆ ตามลำดับ ลักษณะธรรมชาติเช่นนี้ จะปรากฏออกมาในรูปของคะแนน คือสัดส่วนของคะแนนที่ทำได้ของผู้สอบจะสะท้อนให้เห็นถึงระดับการสะสมของความสามารถในขอบเขตเนื้อหาหนึ่ง ๆ ดังนั้นในการตัดสินใจเป็นผู้รอบรู้หรือไม่รอบรู้ก็จะพิจารณาได้จากระดับความสามารถช่วงใดช่วงหนึ่งของความต่อเนื่องในการเรียนรู้เป็นเกณฑ์มาตรฐาน

เกณฑ์ที่ 2 แนวคิดของการกำหนดเกณฑ์มาตรฐาน โดยพิจารณาถึงองค์ประกอบพื้นฐานที่เป็นหลักในการได้มาซึ่งเกณฑ์มาตรฐานนั้น ๆ องค์ประกอบพื้นฐานตามเกณฑ์ข้อนี้มี 2 ส่วนหลัก คือ

1. ส่วนที่เป็นการตัดสินใจของผู้เชี่ยวชาญ (Judgmental) ส่วนนี้ หมายความว่า มาตรการที่จะนำมากำหนดเป็นเกณฑ์มาตรฐานนั้น จะยึดเอาผลการพิจารณาจากข้อสรุปของผู้เชี่ยวชาญเฉพาะเนื้อหาในแต่ละสาขาวิชาผลสัมฤทธิ์เป็นหลัก
2. ส่วนที่ข้อมูลเชิงประจักษ์ (Empirical) ส่วนนี้หมายความว่า มาตรการที่จะนำมากำหนดเป็นเกณฑ์มาตรฐานนั้นจะยึดเอาข้อมูลเชิงประจักษ์ที่ได้จากผลของการนำแบบทดสอบไปทดลองใช้ก่อนเป็นหลัก

จากองค์ประกอบพื้นฐานทั้งสองส่วนดังกล่าว ยังก่อให้เกิดรูปแบบของการผสมระหว่างองค์ประกอบทั้งสอง ทำให้ได้แนวทางใหม่ในการกำหนดเกณฑ์มาตรฐานอีก 3 แนวทางคือ

1. แนวทางที่ใช้ข้อสรุปของผู้เชี่ยวชาญเพียงอย่างเดียว
2. แนวทางที่ใช้ข้อมูลเชิงประจักษ์เป็นส่วนสำคัญ และใช้การตัดสินใจของผู้เชี่ยวชาญเป็นส่วนประกอบ
3. แนวทางที่ใช้การตัดสินใจของผู้เชี่ยวชาญเป็นส่วนสำคัญ และใช้ข้อมูลเชิงประจักษ์เป็นส่วนประกอบ

เกณฑ์ที่ 3 แนวคิดของการกำหนดเกณฑ์มาตรฐาน โดยการพิจารณาว่าเป็นแนวคิดใดระหว่าง 2 แนวทางต่อไปนี้

1. แนวคิดของการกำหนดค่าของเกณฑ์มาตรฐาน (Setting Standards)
2. แนวคิดของการพยายามปรับเปลี่ยนเพื่อหาเกณฑ์มาตรฐานที่เหมาะสม (Standards)

ข้อควรคำนึงในการกำหนดคะแนนจุดตัด

มิลแมน (Millman, 1975 อ้างถึงในจรัสศรี ทองชุมนุม , 2540 : 15-16) ได้ให้คำแนะนำว่าในการกำหนดคะแนนจุดตัดนั้น ควรคำนึงถึงปัจจัยต่าง ๆ ดังต่อไปนี้

1. ความสามารถ (Performance) ของผู้เรียนอื่น

การกำหนดคะแนนจุดตัดนั้นต้องอาศัยผลการสอบของผู้เรียนอื่นด้วย จึงต้องคำนึงถึงด้วยว่า โดยทั่วไปความสามารถของผู้สอบที่ใช้เพื่อกำหนดคะแนนจุดตัดเป็นอย่างไร อาจใช้ผู้สอบที่มีความสามารถเป็นบรรทัดฐานในการกำหนดคะแนนจุดตัดที่ปฏิบัติได้ เช่น นักทดสอบอาจจะกำหนดว่าคะแนนจุดตัดควรเป็นอย่างน้อย 10 % ของคะแนนของผู้ที่มีความสามารถในระดับรอบรู้ (Master) เป็นต้น

2. ความยากของแบบทดสอบ

ควรใช้ดุลยพินิจในการตัดสินใจว่า จุดประสงค์ต่าง ๆ ที่จะทำการทดสอบนั้น ผู้สอบที่มีความสามารถปานกลางนั้นควรจะได้ถูกต้องสักเท่าใด จึงจะอยู่ในระดับที่พอใจ

3. ผลกระทบที่ตามมา

จุดมุ่งหมายที่สำคัญของการกำหนดคะแนนจุดตัด คือ การที่ผู้เรียนมีความรู้ความสามารถในระดับต่ำที่สุดเพื่อให้เพียงพอที่จะศึกษาต่อในระดับที่สูงกว่าได้ แต่บางครั้งก็ต้องคำนึงด้วยว่า ถ้าคะแนนจุดตัดสูงเกินไปจะทำให้ผู้เรียนไม่ผ่านโดยไม่จำเป็น แต่ถ้าหากว่าต่ำเกินไปก็จะทำให้มาตรฐานการศึกษาต่ำด้วย หรือทำให้ผู้เรียนเกิดปัญหาในการเรียนต่อไป

ดังนั้นจึงควรหาคะแนนจุดตัดที่เหมาะสมมาใช้ในการตัดสินใจโดยยึดความสำคัญของเนื้อหาหรือจุดประสงค์เป็นหลักว่า เรื่องใดจำเป็นต้องให้ผู้เรียนสอบผ่าน และเรื่องใดที่จำเป็นน้อยก็ให้พิจารณาตัดได้

4. ผลทางจิตวิทยาและการลงทุนทางการศึกษา

การที่ให้ผู้เรียนสอบตกอยู่เสมอ ๆ ย่อมเป็นผลเสียต่อจิตใจของผู้สอบ และเป็นผลเสียต่อการลงทุน ดังนั้นจึงควรพิจารณาว่าจะแก้ปัญหาเหล่านี้ได้อย่างไร เพราะว่า “ความสามารถ ความเข้าใจ และความซบซึ้ง เป็นประสบการณ์ที่คนโดยมากมีอยู่” ไม่ใช่ว่ามีอยู่ทั้งหมดหรือไม่มีเลย แต่ว่าต่างคนก็มีอยู่เล็กน้อยแตกต่างกัน

5. ความคลาดเคลื่อนจากการเดาและการสุ่มตัวอย่างข้อสอบ

การที่ผู้สอบสอบผ่านหรือไม่ผ่าน อาจเกิดจากความคลาดเคลื่อนในการเดาก็ได้ ถ้ารูปแบบของข้อสอบเป็นแบบชี้แนะหรือเปิดโอกาสให้นักเรียนตอบถูกได้โดยการเดา เราอาจแก้ไขได้โดยการกำหนดคะแนนจุดตัดให้สูงขึ้น ส่วนความคลาดเคลื่อนในการสุ่มข้อสอบเป็นความลำเอียง

ที่เนื่องมาจากการไม่พิจารณาชนิดของคำถามและเนื้อหาในโดเมน ทำให้การเลือกตัวอย่างไม่ครอบคลุมวัตถุประสงค์ที่ตั้งไว้ ซึ่งอาจดูได้จากดัชนีความสัมพันธ์ของเนื้อหา

บุญเชิด ภิญโญอนันตพงษ์ (2527 : 114) กล่าวว่า การกำหนดคะแนนจุดตัดต้องคำนึงถึงสิ่งต่อไปนี้

1. ประเภทของแบบทดสอบที่ใช้ว่าเป็นแบบเติมคำ ถูกผิด หรือเลือกตอบหลายตัวเลือก
2. จำนวนข้อสอบ เพราะจำนวนข้อน้อยมีแนวโน้มต้องกำหนดคะแนนจุดตัดให้สูง แต่ถ้าจำนวนข้อมาก มีแนวโน้มต้องกำหนดคะแนนจุดตัดให้ต่ำ
3. พิจารณาจากการใช้แบบทดสอบเก่า ๆ สังเกตพฤติกรรม หรือจากผลการวิจัยอื่น ๆ ประกอบ

วิธีการกำหนดคะแนนจุดตัด

ในที่นี้จะกล่าวถึงเฉพาะวิธีการกำหนดคะแนนจุดตัดที่ได้รับความนิยมสำหรับการนำไปประยุกต์ใช้ ดังนี้

1. การกำหนดคะแนนจุดตัดโดยใช้คุณลักษณะของผู้เชี่ยวชาญ การกำหนดคะแนนจุดตัดในแนวนี้นี้มีผู้เสนอไว้หลายวิธี (บุญเชิด ภิญโญอนันตพงษ์ , 2527 : 123-125 ; Glass , 1978 : 246-248) เช่น

1.1 วิธีการกำหนดคะแนนจุดตัดโดยใช้ผลที่ได้จากการปฏิบัติของคนอื่น ๆ เป็นเกณฑ์ (Performance of other as a Criterion) ระดับเกณฑ์หรือคะแนนจุดตัดนี้สร้างขึ้นโดยอ้างอิงพารามิเตอร์ของประชากรผู้เข้าสอบ ดังนั้น คะแนนจุดตัดอาจกำหนดได้โดยใช้ค่ามัธยฐาน (Median) ของคะแนนจากแบบทดสอบหรือคะแนนที่สอดคล้องกับจำนวนเปอร์เซ็นต์ของผู้สอบผ่านเกณฑ์ซึ่งระบุไว้ล่วงหน้า โดยอาศัยผู้เชี่ยวชาญเนื้อหาตัดสินสัมฤทธิ์พิจารณาแบบทดสอบ แล้วให้ระบุจำนวนเปอร์เซ็นต์ของผู้สอบผ่านว่าควรเป็นเท่าไรก่อน ซึ่งการกำหนดวิธีนี้เป็นอ้างอิงนอร์ม (Norm) อย่างแท้จริง นักทฤษฎีเกี่ยวกับอิงเกณฑ์จึงไม่เห็นด้วยเนื่องจากยังเป็นการอิงกลุ่มนั่นเอง

1.2 นับถอยหลังจาก 100 % (Counting Backwards from 100 %) ในการสร้างแบบทดสอบโดยพยายามให้ข้อสอบทุก ๆ ข้อสอดคล้องกับวัตถุประสงค์การเรียนรู้แล้ว ผู้สอบทุกคนควรทำแบบทดสอบนั้นได้ทุกข้อ คือ ระดับปฏิบัติจะเป็น 100 % แต่ความเป็นจริงแล้วทำไม่ได้ จึงต้องใช้วิธีลดระดับคะแนนจุดตัดลงมาสำหรับความไม่สมบูรณ์และความคลาดเคลื่อนต่าง ๆ จาก 100 % เหลือ 95 % หรือ 80 % หรืออื่น ๆ เกณฑ์ที่ตั้งในลักษณะนี้เกิดปัญหาในการกำหนดกันตามอำเภอใจ

1.3 การพิจารณาจากความสามารถต่ำสุด (Judging Minimal Competence) การกำหนดคะแนนจุดตัดด้วยวิธีนี้ อาศัยแนวคิดของเกลเซอร์ที่เสนอไว้ว่า ความสามารถจะต่อเนื่องกันจากไม่มีเลขจนถึงปฏิบัติได้โดยสมบูรณ์ทั้งหมด และคะแนนจุดตัดจะแบ่งผู้ที่มีความสามารถและไม่มีความสามารถออกจากกัน ดังนั้นหาผู้ที่มีความสามารถต่ำสุดที่จะยอมรับว่าผ่านได้ ก็จะทำให้ได้ผลคะแนนจุดตัดที่เหมาะสม สำหรับวิธีการนี้มีผู้ศึกษาและได้เสนอเทคนิคในการหาคะแนนความสามารถต่ำสุดไว้หลายคนด้วยกัน เช่น

1.3.1 วิธีของนีเดลสกี (Nedelsky , 1954)

วิธีนี้เป็นการหาคะแนนจุดตัด โดยอาศัยความน่าจะเป็นของการเดาตอบถูก จากแบบทดสอบเลือกตอบหลายตัวเลือก คะแนนจุดตัดจะกำหนดจากคะแนนสอบผ่านขั้นต่ำของนักเรียนที่มีสมรรถภาพขั้นต่ำสุด (กลุ่มคาบเส้น) ซึ่งคำนวณได้จากสูตร ดังนี้

$$D = M_{FD} + K S_{FD}$$

เมื่อ	D	แทน	คะแนนสอบผ่านขั้นต่ำ (คะแนนจุดตัด)
	M_{FD}	แทน	ค่าเฉลี่ยคะแนนการเดาตอบถูกทั้งหมด หาได้จากผลรวมของความน่าจะเป็นในการเดาตอบถูก
	S_{FD}	แทน	ส่วนเบี่ยงเบนมาตรฐานของคะแนนการเดาตอบถูก (หรือความน่าจะเป็นในการเดาตอบถูก)
	K	แทน	ค่าคงที่ ถ้าตัวเลือกผิดเด่นชัดจะใช้ค่า 0.5 ถ้าตัวเลือกแตกต่างกันเล็กน้อยจะใช้ค่า 1.0 ถ้าตัวเลือกคล้ายคลึงกันมากจะใช้ค่า 1.5

วิธีการหาคะแนนจุดตัด มีลำดับขั้นดังนี้

1. นำข้อสอบไปให้ครูประจำวัดผลสัมฤทธิ์กลุ่มหนึ่งพิจารณา
2. ให้กาเครื่องหมายทับตัวเลือกของข้อสอบ ซึ่งคิดว่านักเรียนที่มีสมรรถภาพขั้นต่ำสุดรู้ว่าไม่ใช่ตัวเลือกถูก อาจมีมากกว่า 1 ตัวเลือก
3. ให้นำจำนวนตัวเลือกที่เหลือมาเขียนเป็นเศษส่วนกลับ เช่นมีจำนวนตัวเลือกเหลืออยู่ 3 ตัวเลือก ก็เขียนเป็น $1/3$ ซึ่งก็คือ คะแนนการเดาตอบถูก

4. เมื่อครูพิจารณาข้อสอบไปได้จำนวนหนึ่งแล้ว ควรหยุดพิจารณาชั่วคราว เพื่อนำผลที่พิจารณาแล้วมาประชุมพิจารณาร่วมกัน เพื่อเปรียบเทียบและอภิปรายถึงการใช้มาตรฐานในการพิจารณาว่าตรงกันหรือไม่ และในขั้นนี้ให้พิจารณากำหนดค่าของ K ไว้เลยว่าควรใช้เท่าไร หลังจากนั้นแล้วก็ให้ครูพิจารณาข้อสอบข้ออื่น ๆ ต่อไปอย่างอิสระ
5. หาค่าเฉลี่ยของคะแนนการเดาตอบถูกของข้อสอบแต่ละข้อ (ผลลัพธ์ที่ได้ เรียกว่า “ความน่าจะเป็นในการเดาตอบถูก”)
6. หาผลรวม และส่วนเบี่ยงเบนมาตรฐานของ “ความน่าจะเป็นในการเดาตอบถูก” (ซึ่งคำนวณจากข้อสอบทุกข้อในแบบทดสอบนั้น) แล้วนำไปแทนค่าในสูตร เพื่อหาค่า D ซึ่งจะได้คะแนนจุดตัดของแบบทดสอบ

1.3.2 วิธีของอีเบล (Ebel , 1972)

วิธีนี้คำนึงถึงจุดอ่อนของคะแนนจุดตัดที่ใช้ในการกำหนดเปอร์เซ็นต์ของคะแนนสอบทั้งหมด ว่าอาจมีข้อบกพร่อง เนื่องจากข้อสอบนั้นง่าย หรือยากเกินไป หรืออำนาจจำแนกน้อยกว่าที่ผู้สร้างตั้งใจไว้ การที่จะสอบผ่านหรือไม่อาจเนื่องมาจากตัวข้อสอบมากกว่าระดับความสามารถที่แท้จริงของผู้สอบก็ได้ อีเบลจึงเสนอวิธีแก้ไขโดยการให้วิเคราะห์เนื้อหาของข้อสอบในแง่ของความเกี่ยวข้อง (Relevance) และความยากง่าย (Difficulty) ของข้อสอบแต่ละข้อเทียบกับเปอร์เซ็นต์ความคาดหวังซึ่งเป็นปริมาณที่คาดว่าผู้ที่มีความสามารถต่ำสุดที่จะสอบผ่านควรทำได้โดยได้จำแนกไว้ ดังตาราง

ตาราง 1 แสดง เปอร์เซนต์การสอบผ่านที่คาดหวัง จากลักษณะข้อสอบ และระดับความยากของข้อสอบ

ลักษณะข้อสอบ	ระดับความยาก		
	ง่าย	ปานกลาง	ยาก
จำเป็น (essential).....	100%	-	-
สำคัญ (important).....	90%	70%	-
ยอมรับได้ (acceptable).....	80%	60%	40%
ไม่แน่ใจ (questionable).....	70%	50%	30%

วิธีการหาคะแนนจุดตัดมีลำดับขั้นดังนี้

1. นำข้อสอบที่ต้องการวิเคราะห์ไปให้ครูประจำวัดผลสัมฤทธิ์พิจารณาว่าข้อสอบมีลักษณะตรงกับปัญหามากน้อยเพียงใด มีระดับความยากง่ายเพียงใด
2. จำแนกข้อสอบตามลักษณะและความยาก โดยแจกแจงเป็นความถี่รวม
3. กำหนดเปอร์เซ็นต์การสอบผ่านตามลักษณะและความยากของข้อสอบ
4. คูณความถี่จำนวนข้อกับเปอร์เซ็นต์การสอบผ่านที่คาดหวังและรวมผลคูณทั้งหมด
5. หาค่าเปอร์เซ็นต์การสอบผ่านที่คาดหวังโดยเฉลี่ย
6. กำหนดคะแนนจุดตัดจากค่าเปอร์เซ็นต์การสอบผ่านที่คาดหวังโดยเฉลี่ยจากขั้นที่ 5

1.3.3 วิธีของแองกอฟ (Angoff, 1971)

วิธีนี้จะอาศัยหลักการของความน่าจะเป็นที่นักเรียนซึ่งมีสมรรถภาพขั้นต่ำสุดที่จะยอมรับได้ตอบข้อสอบถูก โดยนำไปให้ผู้เชี่ยวชาญเนื้อหาหรือครูผู้สอนพิจารณา ค่าความน่าจะเป็นในการตอบข้อสอบถูก จะนำมาใช้เป็นคะแนนจุดตัด วิธีการหาคะแนนจุดตัดมีลำดับขั้นดังนี้

1. นำข้อสอบทั้งหมดไปให้ครูผู้สอนกลุ่มหนึ่งพิจารณาเนื้อหาข้อสอบและความยาก
2. ให้ครูผู้สอนพิจารณาต่อไปว่านักเรียนที่มีความสามารถขั้นต่ำสุดตามเนื้อหา ข้อสอบจะมีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกเป็นเท่าไร
3. นำค่าความน่าจะเป็นในการตอบข้อสอบถูกที่ครูผู้สอนแต่ละคนพิจารณาไว้มาหาค่าเฉลี่ยของความน่าจะเป็น
4. กำหนดคะแนนจุดตัดจากค่าเฉลี่ยในขั้นที่ 3 โดยการรวมค่าเฉลี่ยของความน่าจะเป็นจากข้อสอบทุกข้อ จะได้คะแนนจุดตัดของแบบทดสอบฉบับนั้น

2. การกำหนดคะแนนจุดตัดโดยใช้การทดลองสอบ หรือใช้การทดสอบร่วมกับการใช้ดุลยพินิจ การกำหนดคะแนนจุดตัด 2 แนวทางที่กล่าวถึงพร้อมกันนี้ ต้องนำแบบทดสอบไปทดลองสอบกับนักเรียน แล้วพิจารณาเลือกคะแนนจุดตัดที่เหมาะสม ซึ่งส่วนใหญ่จะใช้หลักสถิติเป็นเครื่องมือ ในการกำหนดคะแนนจุดตัด มีผู้เสนอไว้หลายวิธี (บุญเชิด ภิญโญอนันตพงษ์ , 2527 : 131-147 ; จรัสศรี ทองชุมนุม , 2540 : 21-24) เช่น

2.1 วิธีกำหนดคะแนนจุดตัดที่แท้จริงของบุญเชิด ภิญโญอนันตพงษ์ หรือวิธี กำหนดคะแนนจุดตัดจากความคลาดเคลื่อนของการเดาตอบ

วิธีนี้ได้คำนึงถึงทฤษฎีการทดสอบที่ว่า คะแนนย่อมมีความคลาดเคลื่อนอยู่ ด้วย ดังนั้น คะแนนจุดตัดจึงควรพิจารณาขอบเขตของความคลาดเคลื่อนประกอบด้วย ซึ่ง ความคลาดเคลื่อนอาจเกิดจากการเดาตอบถูก คะแนนที่สอบได้จึงมีค่าสูงกว่าความสามารถแท้จริง อยู่เล็กน้อย วิธีการอย่างหนึ่งก็คือ จะต้องหาคะแนนจุดตัดที่กำจัดความคลาดเคลื่อนจากการเดา ตอบถูกออกไปเสียก่อน จึงจำเป็นต้องนำแบบทดสอบไปสอบกับนักเรียน แล้วใช้สูตรปรับแก้ คะแนนเดาเพื่อหาคะแนนจุดตัดที่แท้จริง บุญเชิด ภิญโญอนันตพงษ์ ได้เสนอสูตรไว้ โดย คัดแปลงแนวคิดของกิลฟอร์ด (Guilford) ที่ใช้ในการปรับแก้คะแนนเดาตอบ ดังนี้

$$P' = \frac{R_i - \frac{W_i}{A-1}}{N}$$

เมื่อ	P'	แทน	สัดส่วนของผู้ตอบถูกจริงหรือผู้รู้คำตอบ
	R_i	แทน	จำนวนผู้ตอบถูกในข้อนั้น
	W_i	แทน	จำนวนผู้ตอบผิดในข้อนั้น
	A	แทน	จำนวนตัวเลือกของแบบทดสอบ
	N	แทน	จำนวนผู้สอบทั้งหมด

นำสัดส่วนของผู้ตอบถูกจริง (P') ในแต่ละข้อมารวมกัน จะได้คะแนนจุดตัด (C) ของแบบทดสอบ นั่นคือ

$$C = \sum P'$$

2.2 วิธีการหาคะแนนจุดตัดโดยเพิ่มคะแนนเกณฑ์อื่น ๆ (Bootstrapping on other Criterion Score) โดยอาศัยการระบุจากเกณฑ์ภายนอกกำหนดว่า สำเร็จหรือรอบรู้อย่างชัดเจน เช่น อาจใช้เกณฑ์ว่า ผู้ได้รับประกาศนียบัตรเป็นเกณฑ์ที่จะจำแนกผู้มีความสำเร็จการศึกษากับผู้ยังมีความรู้ไม่ครบถ้วนเพียงพอที่จะสำเร็จการศึกษา การหาคะแนนจุดตัดจะพิจารณาจากคะแนน จุดตัดที่สามารถแบ่งคนได้สอดคล้องตามเกณฑ์ภายนอก โดยการปรับคะแนนจุดตัดที่ละคะแนนจนกว่าจะได้ผลสอดคล้องกับเกณฑ์ภายนอกมากที่สุด

2.3 วิธีการหาคะแนนจุดตัดโดยวิธีของเบิร์ก (Berk , 1976)

เบิร์กได้หาคะแนนจุดตัดโดยประยุกต์มาจากวิธีการเพิ่มคะแนนเกณฑ์อื่น ๆ ซึ่งเบิร์กกล่าวว่า การกำหนดเกณฑ์คือการกำหนดจุดตัดของคะแนนที่แบ่งผู้เรียนออกเป็นสองพวกคือพวกที่ได้รับการสอนให้เป็นพวกที่รอบรู้ (Master) พวกที่ไม่ได้รับการสอนให้เป็นพวกไม่รอบรู้ (Non-Master) หลังจากให้นักเรียนทั้งสองกลุ่มทำแบบทดสอบแล้ว พิจารณาการกระจายของคะแนนสองกลุ่มที่คาบเกี่ยวกัน จุดที่ฟังก์ชันทั้งสองตัดกัน คือคะแนนพยากรณ์ที่จะแบ่งการเรียนรู้อันเป็นสี่พวก ดังนี้

		เกณฑ์	
		ได้รับการสอน	ไม่ได้รับการสอน
คะแนนพยากรณ์	รอบรู้	รอบรู้จริง (TM)	รอบรู้ไม่จริง(FM)
	ไม่รอบรู้	ไม่รอบรู้ไม่จริง(FN)	ไม่รอบรู้จริง(TN)

คะแนนจุดตัดนี้ทำหน้าที่เป็นคะแนนพยากรณ์ โดยคะแนนที่ทำให้ค่าความน่าจะเป็นในการตัดสินใจถูกต้อง คือ ค่า $P(TM) + P(TN)$ มีค่าสูงสุด หรือทำให้ค่าความน่าจะเป็นในการตัดสินใจผิด คือ $P(FM) + P(FN)$ มีค่าต่ำสุด ณ จุดคะแนนนั้น ก็จะเป็นคะแนนจุดตัดที่เหมาะสมที่สุด ในการคำนวณจะเลื่อนคะแนนพยากรณ์ไปเรื่อย ๆ จุดคะแนนหนึ่งที่มีค่า $P(TM) + P(TN)$ สูงสุด และค่า $P(FM) + P(FN)$ ต่ำสุด ก็จะเป็นคะแนนจุดตัดที่เหมาะสม

$$\text{เมื่อ } P(TM) = TM/(M+N)$$

$$P(TN) = TN/(M+N)$$

$$P(FN) = FN/(M+N)$$

$$P(FM) = FM/(M+N)$$

M แทน จำนวนนักเรียนทั้งหมดที่ได้รับการสอนแล้ว

N แทน จำนวนนักเรียนทั้งหมดที่ยังไม่ได้รับการสอน

ในการพิจารณาคะแนนแต่ละคะแนน เพื่อจะกำหนดให้เป็นคะแนนจุดตัดที่เหมาะสมนั้น สามารถใช้การตรวจสอบความตรง เพื่อเลือกคะแนนที่ให้ค่าสัมประสิทธิ์ความตรงสูงที่สุดมาเป็นคะแนนจุดตัดได้ โดยใช้สูตร

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

เมื่อ	ϕ_{vc}	แทน	ค่าสัมประสิทธิ์ความตรงของคะแนนจุดตัด
	BR	แทน	ค่าความน่าจะเป็นของผู้รอบรู้ในประชากร ดังนั้น $BR = P(FN) + P(TM)$
	SR	แทน	ค่าความน่าจะเป็นของการพยากรณ์ผู้รอบรู้ในประชากร ดังนั้น $SR = P(FM) + P(TM)$

วิธีการหาคะแนนจุดตัด มีลำดับขั้น ดังนี้

1. กำหนดกลุ่มรอบรู้และไม่รอบรู้ โดยใช้การสอนเป็นเกณฑ์ จำแนกกลุ่มที่ได้รับการสอน และกลุ่มที่ไม่ได้รับการสอนไว้สำหรับเป็นเกณฑ์เทียบ
2. นำแบบทดสอบอิงเกณฑ์ไปสอบกับกลุ่มตัวอย่างนักเรียนที่เลือกไว้ทั้งหมด
3. จำแนกนักเรียนเป็น 4 กลุ่ม ในแต่ละคะแนนจุดตัด ซึ่งเกิดจากคะแนนจุดตัดกับเกณฑ์ภายนอกจำแนกร่วมกัน
4. หาค่าความน่าจะเป็นของแต่ละเหตุการณ์ ในแต่ละคะแนนแล้วคำนวณความน่าจะเป็นของการตัดสินถูก ความน่าจะเป็นของการตัดสินผิด และสัมประสิทธิ์ความแม่นยำตรงของคะแนนจุดตัดทุก ๆ คะแนน
5. กำหนดคะแนนจุดตัดที่เหมาะสมจากคะแนนจุดตัดที่สมนัยกับการตัดสินถูกที่มีค่าสูงสุด หรือความน่าจะเป็นของการตัดสินที่มีค่าต่ำสุด หรือสัมประสิทธิ์ความแม่นยำตรงของคะแนนจุดตัดที่มีค่าสูงสุด

2.4 ทฤษฎีการตัดสินใจของเกลส (Glass , 2978 : 243-257)

เป็นวิธีที่อาศัยความสัมพันธ์ของคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์กับเกณฑ์ภายนอกที่สอดคล้องซึ่งกำหนดไว้ก่อน โดยการสร้างฟังก์ชันของคะแนนจุดตัด ซึ่งได้จากการใช้เกณฑ์ภายนอกจำแนกคนเป็นสองกลุ่ม เช่น ใช้เกณฑ์แบ่งคนเป็นกลุ่มที่จบการศึกษา กับกลุ่มที่ไม่จบการศึกษา สัดส่วนของคนทั้งสองกลุ่มแทนด้วย P_E และ $1 - P_E$ ตามลำดับ แล้วนำแบบทดสอบอิงเกณฑ์ไปทดสอบกับคนทั้งสองกลุ่มและกำหนดคะแนนจุดตัด (C_X) ขึ้นมาเพื่อแบ่งคนที่สอบผ่านและสอบไม่ผ่านเกณฑ์ จะทำให้เกิดการจำแนกร่วมกันทั้งหมด 4 กลุ่ม ดังนี้

		เกณฑ์ภายนอก	
		ผ่าน	ไม่ผ่าน
แบบทดสอบอิงเกณฑ์	ไม่ผ่าน	P_A	P_B
	ผ่าน	P_C	P_D

เมื่อ	P_A	แทน	สัดส่วนจำนวนคนที่ไม่ผ่านเกณฑ์ของแบบทดสอบแต่ผ่านเกณฑ์ภายนอก เป็นการจำแนกผิดพลาด
	P_B	แทน	สัดส่วนจำนวนคนที่ไม่ผ่านเกณฑ์ของแบบทดสอบและไม่ผ่านเกณฑ์ภายนอก เป็นการจำแนกถูก
	P_C	แทน	สัดส่วนจำนวนคนที่ผ่านเกณฑ์ของแบบทดสอบและผ่านเกณฑ์ภายนอก เป็นการจำแนกถูก
	P_D	แทน	สัดส่วนจำนวนคนที่ผ่านเกณฑ์ของแบบทดสอบแต่ไม่ผ่านเกณฑ์ภายนอก เป็นการจำแนกผิดพลาด

ซึ่งจะได้ฟังก์ชันของคะแนนจุดตัด $f(C_X)$ ดังนี้

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

ค่าของ C_X ที่ทำให้ $f(C_X)$ มีค่าน้อยที่สุดจะถูกเลือกเป็นคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ การใช้ฟังก์ชันมีข้อตกลงว่า การจำแนกผิดทางบวก (P_D) และการจำแนกผิดทางลบ (P_A) มีความสำคัญเท่ากัน ถ้าให้ความสำคัญของการจำแนกผิดทั้งสองแบบไม่เท่ากัน จะทำให้ค่าของฟังก์ชันแปรเปลี่ยนไปมาก ดังนั้นจึงต้องมีการปรับแก้โดยกำหนดค่าถ่วงขึ้นมาสองค่าคือ α เป็นค่าความสำคัญของการจำแนกผิดทางลบ และ β เป็นค่าความสำคัญของการจำแนกผิดทางบวก ดังนั้นจะได้ฟังก์ชันที่เหมาะสมดังนี้

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

ค่า α และ β นี้จะกำหนดขึ้นจากการใช้ดุลยพินิจพิจารณาว่า มีความสำคัญมากน้อยเท่าไร ซึ่งมีพิสัยความสำคัญตั้งแต่ 0 ถึง 1

วิธีการหาคะแนนจุดตัด มีลำดับขั้น ดังนี้

1. เลือกเกณฑ์ภายนอกที่สอดคล้องกับแบบทดสอบอิงเกณฑ์เพื่อจำแนกนักเรียนเป็น 2 กลุ่ม
2. นำแบบทดสอบอิงเกณฑ์ไปสอบกับกลุ่มตัวอย่างนักเรียนที่เลือกไว้ทั้งหมด
3. จำแนกนักเรียนเป็น 4 กลุ่ม ในแต่ละคะแนนจุดตัด ซึ่งคิดจากคะแนนจุดตัดกับเกณฑ์ภายนอกจำแนกร่วมกัน
4. หาค่าฟังก์ชันของคะแนน แต่ละคะแนนจากสมการ 1 หรือ สมการ 2 ขึ้นอยู่กับการกำหนดความสำคัญของการจำแนกผิดทั้งสองแบบ
5. กำหนดคะแนนจุดตัดที่เหมาะสมจากคะแนนจุดตัดที่สมนัยกับฟังก์ชันของคะแนนที่มีค่าต่ำสุด

เกณฑ์ในการประเมินวิธีกำหนดคะแนนจุดตัด

Berk (1986 อ้างถึงใน วิจารณ์ ศรีบุตรตา , 2541 : 25-26) เสนอเกณฑ์การประเมินวิธีกำหนดคะแนนจุดตัดไว้ดังนี้

1. ความพอเพียงตามหลักการวัดผลสัมฤทธิ์ (Technical adequacy) เกณฑ์เกี่ยวกับความพอเพียงตามหลักวิชาการได้มาจาก 3 แหล่งข้อมูล คือ

1.1 คุณลักษณะของวิธีกำหนดมาตรฐานที่ตัดสินโดยผู้เชี่ยวชาญที่มีหลักฐานทางด้านการวิจัยและการวิเคราะห์เชิงเหตุผล

1.2 เป็นมาตรฐานที่ตรงกับปัญหา และได้รับคำแนะนำจากมาตรฐานการทดสอบด้านการศึกษาและจิตวิทยา (Standard for Educational and Psychological Testing)

1.3 การตัดสินใจที่ว่าด้วยประเด็นเกี่ยวกับการกำหนดมาตรฐาน (Standard Setting)

จาก 3 แหล่งข้อมูลนี้นำมากำหนดเป็นเกณฑ์การประเมินวิธีกำหนดจุดตัดได้ 6 เกณฑ์ ดังนี้

1.3.1 เป็นวิธีให้สารสนเทศเกี่ยวกับการแบ่งกลุ่มที่เหมาะสม โดยวิธีที่ใช้ควรให้จุดตัดที่แบ่งกลุ่มผู้สอบออกเป็นผู้รอบรู้กับไม่รอบรู้ ผ่านกับไม่ผ่าน หรือมีความสามารถกับไม่มีความสามารถ

1.3.2 ควรเป็นวิธีที่คำนึงถึงการปฏิบัติของข้อสอบและคุณภาพของข้อสอบ โดยเป็นวิธีที่มีความไวต่อระดับความยากของข้อสอบที่ต่างกัน ตลอดจนคุณลักษณะอื่น ๆ ที่เกี่ยวกับข้อสอบ เช่น อำนาจจำแนกของข้อสอบระหว่างกลุ่มที่รอบรู้กับไม่รอบรู้

1.3.3 เป็นวิธีที่ไวต่อการเรียนการสอน สอดคล้องกับการเรียนการสอน เพราะถ้าแบบสอบ วัดได้ไม่ตรงตามวัตถุประสงค์การเรียนการสอน คะแนนจุดตัดที่ได้ก็จะไม่มีความตรงและไม่เหมาะสม

1.3.4 เป็นวิธีที่ใช้สถิติอย่างเหมาะสมและถูกต้องตามวัตถุประสงค์ ตลอดจนการแปลความหมายได้ถูกต้อง

1.3.5 เป็นวิธีที่สามารถระบุมาตรฐานที่แท้จริงให้เห็นความแตกต่างระหว่างมาตรฐานบนสเกลของคะแนนจริง (มาตรฐานที่แท้จริง) กับสเกลของคะแนนที่สังเกตได้

1.3.6 เป็นวิธีที่มีหลักฐานเกี่ยวกับความตรงในการตัดสินจุดตัด

2. เกณฑ์เกี่ยวกับการนำไปปฏิบัติได้จริง (Practicability) เกณฑ์ความเพียงพอด้านวิชาการทั้ง 6 ข้อ ยังไม่เพียงพอสำหรับการเลือกวิธีกำหนดมาตรฐาน จึงเพิ่มเกณฑ์เกี่ยวกับการนำไปปฏิบัติได้จริง ซึ่งว่าด้วยเรื่องการนำไปใช้ การคำนวณ และการตีความหมาย ดังนี้

2.1 เป็นวิธีที่ง่ายต่อการนำไปปฏิบัติ มีความชัดเจน เข้าใจได้ง่าย มีขั้นตอนในการปฏิบัติที่ไม่ยุ่งยาก ซับซ้อน ไม่เสียเวลา และเป็นระบบ

2.2 เป็นวิธีที่ใช้สถิติในการคำนวณที่ง่าย หรือสามารถใช้โปรแกรมสำเร็จรูปทางสถิติได้ทั้งไมโครคอมพิวเตอร์ และเมนเฟรม

2.3 เป็นวิธีที่แปลความหมายได้ง่าย และสามารถอธิบายให้ผู้ที่เกี่ยวข้อง เช่น นักเรียน ครู ผู้ปกครอง หรือผู้บริหารโรงเรียน เข้าใจได้

2.4 เป็นวิธีที่น่าเชื่อถือ บางครั้งวิธีที่ใช้สถิติที่ยุ่งยากก็ไม่ใช่วิธีที่น่าเชื่อถือเสมอไป เพราะอาจทำให้เกิดความสับสนได้มากกว่าวิธีที่ใช้การคำนวณอย่างง่าย ๆ

ข้อเสนอแนะเกี่ยวกับคะแนนจุดตัด

แฮมเบิลตัน (Hambleton, 1978) ให้ข้อเสนอแนะเกี่ยวกับการกำหนดคะแนนจุดตัดดังต่อไปนี้

1. การพิจารณาคะแนนจุดตัดควรทำโดยกลุ่มคนที่ทำงานร่วมกันหลาย ๆ กลุ่ม เช่น ครู ผู้บริหาร ผู้ปกครอง ผู้เชี่ยวชาญด้านหลักสูตร และนักเรียน สำหรับจำนวนคนในแต่ละกลุ่มขึ้นกับความสำคัญของแบบทดสอบที่นำมาพิจารณา และจำนวนของตารางวิเคราะห์โดเมนของข้อสอบ

2. ผู้ตัดสินควรนำตารางวิเคราะห์โดเมนของข้อสอบมาอภิปรายร่วมกันด้วย

3. การใช้วิธีของอibel หรือ นิเคลสกี ควรทดลองกับกลุ่มตัวอย่าง หลาย ๆ กลุ่มเพื่อตัดสินและแก้ปัญหาเกี่ยวกับความแตกต่างระหว่างกลุ่ม

4. ตารางวิเคราะห์โดเมนข้อสอบที่มีความซับซ้อน หรือมีความสำคัญมาก ควรใช้เวลาในการพิจารณามากขึ้น และควรกำหนดจุดตัดให้สูงกว่า

5. พยายามทำให้เกิดความมั่นใจว่าผู้ตัดสินมีความรู้เกี่ยวกับการใช้แบบสอบ และรู้ว่าจะใช้กับนักเรียนกลุ่มใด

6. เมื่อใช้ผู้ตัดสินตั้งแต่ 2 กลุ่มหรือมากกว่า ควรศึกษาความคงที่ในการกำหนดน้ำหนัก และความแตกต่างในการกำหนดน้ำหนักของผู้ตัดสินและอภิปรายเพื่อหาการตัดสินที่สอดคล้องกัน

7. นำข้อมูลเกี่ยวกับการสอบครั้งก่อนมาพิจารณาเพื่อกำหนดจุดตัด

8. เปรียบเทียบสภาพการรอบรู้ของผู้สอบที่ยังไม่ได้เรียนและเรียนมาแล้ว หากมีนักเรียนจำนวนมากได้รับการจัดประเภทผิดต้องพิจารณาคะแนนจุดตัดอีกครั้ง

9. ถ้าหลักสูตรมีการเปลี่ยนแปลง ต้องมีการตรวจสอบคะแนนจุดตัดใหม่ให้เหมาะสม

4. การวิเคราะห์คุณภาพของแบบทดสอบอิงเกณฑ์

1. ค่าความยาก (Difficulty)

การประเมินตามแนวคิดอิงเกณฑ์ เป็นการทดสอบเพื่อเปรียบเทียบความสามารถของผู้เรียนกับเกณฑ์ที่เป็นมาตรฐานว่าอยู่ในระดับถึงมาตรฐานที่ยอมรับหรือไม่ การตรวจสอบคุณภาพของแบบทดสอบตามแนวคิดนี้มีวิธีการหาความยากของข้อสอบเช่นเดียวกับแนวคิดแบบอิงกลุ่ม เพียงแต่ค่าความยากนั้นไม่ได้ถือว่าข้อสอบที่ยากหรือง่ายเป็นข้อสอบที่ไม่ดี แต่จะเป็นการวัดจุดประสงค์เป็นสำคัญ ดังนั้น ข้อสอบที่วัดตรงตามจุดประสงค์และเป็นข้อสอบที่ง่ายหรือยาก ก็ถือว่าเป็นข้อสอบที่ดี

ค่าความยากของข้อสอบแบบอิงเกณฑ์จึงไม่ได้นำมาชี้ถึงคุณภาพ และไม่ได้นำมาเป็นเกณฑ์สำคัญในการตัดข้อสอบ ซึ่งการหาค่าความยากจะใช้สูตรดังนี้

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

เมื่อ	P	แทน	ค่าความยากของข้อสอบ
	R	แทน	จำนวนคนตอบถูก
	N	แทน	จำนวนคนทั้งหมด

2. อำนาจจำแนก (Discrimination)

ค่าอำนาจจำแนก หมายถึง ประสิทธิภาพในการจำแนกระดับความสามารถของผู้เรียนแล้ว (กลุ่มรอบรู้) กับกลุ่มที่ยังไม่เรียน (กลุ่มไม่รอบรู้) การวิเคราะห์ข้อสอบเป็นรายข้อแบบอิงเกณฑ์ จะมุ่งหาค่าอำนาจจำแนกของข้อสอบ โดยถือว่าข้อสอบอิงเกณฑ์ที่ดี ควรมีค่าอำนาจจำแนกดี (สมศักดิ์ สินธุระเวชย์, 2522 : 11-13) ซึ่งในที่นี้จะนำเสนอวิธีการหาค่าอำนาจจำแนก 2 วิธี ดังนี้

1. การหาค่าอำนาจจำแนกตามวิธีของคริสปีน และเฟลด์ลูเซน (Kryspin and Feldlusion) คริสปีน และเฟลด์ลูเซน ได้เสนอการหาค่าอำนาจจำแนกที่เรียกว่า ดัชนี S (index of sensitivity) หรือดัชนีความไวในการวัด ซึ่งมีสูตรดังนี้

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

เมื่อ	S	แทน	ค่าอำนาจจำแนกของข้อสอบ
	R_A	แทน	จำนวนคนตอบถูกหลังสอน
	R_B	แทน	จำนวนคนตอบถูกก่อนสอน
	T	แทน	จำนวนคนที่เข้าสอบทั้งสองครั้ง

การแปลความหมายค่า S

การพิจารณาคุณภาพของข้อสอบในด้านความไว พิจารณาตามระดับค่า S ดังนี้

ค่า S	ความหมาย
1.00	เป็นข้อสอบที่ดี เป็นไปตามทฤษฎี
.80 ถึง .99	เป็นข้อสอบที่ดี หาได้ในเชิงปฏิบัติ
.30 ถึง .79	เป็นข้อสอบที่พอใช้ได้
.00 ถึง .29	เป็นข้อสอบที่ไม่ดี ควรตัดทิ้ง
-1.00 ถึง .00	เป็นข้อสอบที่ใช้ไม่ได้ ควรตัดทิ้ง

ในการพิจารณาค่าอำนาจจำแนก (S) ถ้าค่า S เป็นบวกใกล้ +1.00 หมายถึง การเรียนการสอนบรรลุตามเป้าหมาย คือ ก่อนเรียนผู้เรียนไม่มีความรู้ หลังจากเรียนแล้ว ปรากฏว่ามีความรู้ตามจุดประสงค์ที่ตั้งไว้ แต่ถ้าค่า S เป็นลบใกล้ -1.00 หมายถึง ก่อนเรียน ผู้เรียนมีความรู้ แต่เมื่อเรียนจบเนื้อหาแล้ว ปรากฏว่าผู้เรียนกลับไม่มีความรู้เลย

2. การหาค่าอำนาจจำแนกตามวิธีของเบรนนัน (Brennan)

เบรนนัน ได้เสนอสูตรในการหาค่าอำนาจจำแนกของข้อสอบแล้วตั้งชื่อ เป็น ดัชนีบี (Discrimination index B) การหาค่าอำนาจจำแนกวิธีนี้จะสอบครั้งเดียวจากกลุ่มตัวอย่างเดียว แล้วแบ่งออกเป็น 2 กลุ่ม คือ กลุ่มผู้ที่สอบได้คะแนนผ่านเกณฑ์ และกลุ่มผู้ที่สอบได้คะแนนไม่ผ่านเกณฑ์ มีสูตร ดังนี้

$$B = \frac{U}{N_1} - \frac{L}{N_2}$$

เมื่อ	B	แทน	ค่าอำนาจจำแนกของข้อสอบ
	U	แทน	จำนวนคนทำข้อสอบข้อนั้นถูกของกลุ่มที่ผ่านเกณฑ์
	L	แทน	จำนวนคนทำข้อสอบข้อนั้นถูกของกลุ่มที่ไม่ผ่านเกณฑ์
	N ₁	แทน	จำนวนคนที่สอบผ่านเกณฑ์
	N ₂	แทน	จำนวนคนที่สอบไม่ผ่านเกณฑ์

การแปลความหมายค่าดัชนีบี (B-index) เป็นดังนี้

ค่า (B-index)	ความหมายว่าข้อสอบนั้นสามารถ
+1.00	บ่งชี้ผู้รอบรู้-ไม่รอบรู้ ได้ถูกต้องทุกคน
.50 ถึง .99	บ่งชี้ผู้รอบรู้-ไม่รอบรู้ ได้ถูกต้องเป็นส่วนใหญ่
.20 ถึง .49	บ่งชี้ผู้รอบรู้-ไม่รอบรู้ ได้ถูกต้องเป็นบางส่วน
.00 ถึง .19	บ่งชี้ผู้รอบรู้-ไม่รอบรู้ ได้ถูกต้องน้อยมาก หรือไม่ถูกต้อง
ติดลบ	บ่งชี้ผู้รอบรู้-ไม่รอบรู้ ผิดพลาด หรือตรงข้ามกับความจริง

3. ความตรงของแบบทดสอบอิงเกณฑ์

ความตรง (Validity) เป็นคุณภาพของแบบทดสอบที่หมายถึง แบบทดสอบที่สามารถวัดได้ตรงตามลักษณะหรือจุดประสงค์ที่ต้องการจะวัด จำแนกตามคุณลักษณะหรือจุดประสงค์ที่ต้องการวัดได้ 3 ประเภทใหญ่ ๆ (ล้วน สายยศ และอังคณา สายยศ, 2539 : 246) คือ

1. ความตรงเชิงเนื้อหา (Content Validity)
2. ความตรงเชิงเกณฑ์สัมพันธ์ (Criterion-Related Validity)
3. ความตรงเชิงโครงสร้าง (Construct Validity)

ความตรงเชิงเนื้อหา หมายถึง เครื่องมือที่สามารถวัดได้ตามเนื้อหาที่ต้องการจะวัด และการพิจารณาความตรงชนิดนี้ จะใช้การวิเคราะห์อย่างมีเหตุผล (rational analysis) จำแนกออกเป็น 2 ชนิด คือ ความตรงเชิงเหตุผล (Logical Validity) และความตรงเชิงพินิจ (Face Validity)

ความตรงเชิงเกณฑ์สัมพันธ์ หมายถึง คุณภาพของเครื่องมือที่เอาผลของแบบทดสอบไปหาความสัมพันธ์กับเกณฑ์ที่ต้องการ จำแนกออกเป็น 2 ชนิด คือความตรงเชิงสภาพ (Concurrent Validity) หมายถึง ความตรงที่เอาผลการวัดของแบบทดสอบที่สร้างขึ้นไปหาความสัมพันธ์กับเกณฑ์ในสภาพปัจจุบัน และความตรงเชิงพยากรณ์ (Predictive Validity) หมายถึง ความตรงที่เอาผลการวัดของแบบทดสอบที่สร้างขึ้นไปหาความสัมพันธ์กับเกณฑ์ในอนาคต เพื่อที่จะเอาผลการสอบไปพยากรณ์ผลความสำเร็จในอนาคต

ความตรงเชิงโครงสร้าง หมายถึง คุณภาพของเครื่องมือที่สามารถวัดได้ตรงตามลักษณะหรือตามทฤษฎีต่าง ๆ ของโครงสร้างนั้น หรือวัดได้ครอบคลุมตามลักษณะของโครงสร้างของแบบทดสอบมาตรฐาน

วิธีการตรวจสอบและประมาณค่าความตรง

1. ความตรงเชิงเนื้อหา (Content Validity)

โรวินลลี และแฮมเบิลตัน (Rovinelli and Hambleton, 1978 อ้างถึงใน บุญเชิด ภิญโญอนันตพงษ์, 2527 : 68-74) ได้เสนอแนวทางในการตรวจสอบความตรงเชิงเนื้อหา โดยเกี่ยวข้องกับการตัดสินของผู้เชี่ยวชาญในด้านเนื้อหาเกี่ยวกับข้อสอบ คือ การหาดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ (The Index of Item Objective Congruence : IOC) วิธีการนี้ใช้ผลการตัดสินของผู้เชี่ยวชาญในด้านในข้อสอบแต่ละข้อว่าสอดคล้องกับจุดประสงค์ที่กำหนดไว้หรือไม่ โดยกำหนดวิธีการให้คะแนนผลการตัดสิน ไว้ว่า

- +1 เมื่อแน่ใจว่าข้อสอบวัดจุดประสงค์ที่กำหนดให้
 0 เมื่อไม่แน่ใจว่าข้อสอบวัดจุดประสงค์ที่กำหนดให้หรือไม่
 -1 เมื่อแน่ใจว่าข้อสอบนั้นไม่ได้วัดจุดประสงค์ที่กำหนดให้

นำคะแนนที่ได้จากการตัดสินมาคำนวณดัชนีความสอดคล้องโดยใช้สูตร

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

เมื่อ	IOC	แทน	ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ เชิงพฤติกรรม
	$\sum R$	แทน	ผลรวมของคะแนนความคิดเห็นของผู้เชี่ยวชาญ ทั้งหมด
	N	แทน	จำนวนผู้เชี่ยวชาญทั้งหมด

2. ความตรงเชิงเกณฑ์สัมพันธ์ (Criterion-Related Validity)

ครอกเกอร์ และแอลจิงนา (Crocker and Algina, 1986 : 224) กล่าวถึงแนวทางทั่วไปที่สามารถใช้ในการศึกษาความตรงเชิงเกณฑ์สัมพันธ์ ที่สำคัญรวม 4 ขั้นตอน คือ

1. กำหนดกรอบพฤติกรรมของตัวแปรเกณฑ์ และวิธีการวัดที่เหมาะสม
2. กำหนดกลุ่มตัวอย่างให้เป็นตัวแทนของประชากรผู้สอบที่เราจะนำแบบทดสอบไปใช้
3. ดำเนินการสอบ และทำการบันทึกคะแนนของผู้สอบแต่ละคน
4. เมื่อกำหนดกรอบข้อมูลของตัวแปรเกณฑ์ได้แล้ว จึงทำการวัดกับผู้สอบแต่ละคน
5. ประเมินค่าความสัมพันธ์ระหว่างคะแนนที่ได้จากแบบทดสอบ กับผลการวัดตัวแปรเกณฑ์

2.1 ความตรงเชิงสภาพ

ความตรงเชิงสภาพของแบบทดสอบอิงเกณฑ์ หมายถึง แบบทดสอบอิงเกณฑ์ที่สามารถวัดความรู้ (Master) และไม่รอบรู้ (Nonmaster) ของนักเรียนในการทดสอบ (Test Status) ในแต่ละจุดประสงค์ ว่าตรงกับสถานภาพความรู้จริง (Known Status) หรือไม่ ถ้าตรงกับสถานภาพความรู้จริง ก็แสดงว่ามีความตรงสูง คำนวณได้จากค่า “ ϕ ” ซึ่งมีค่าอยู่ระหว่าง -1 ถึง +1 ถ้าเข้าใกล้ +1 แสดงว่ามีความตรงสูง และค่าใกล้ 0 หรือ -1 แสดงว่ามีความตรงต่ำ หรือไม่ตรง (ล้วน สายยศ และอังคณา สายยศ, 2539 : 255)

2.2 ความตรงเชิงพยากรณ์

ความตรงเชิงพยากรณ์ของแบบทดสอบอิงเกณฑ์ หมายถึง แบบทดสอบอิงเกณฑ์ที่สามารถวัดความรู้ (Master) และไม่รอบรู้ (Nonmaster) ของนักเรียนในการทดสอบ (Test Status) ในแต่ละจุดประสงค์ ว่าสอดคล้องกับสถานภาพในอนาคต (Future Status) หรือไม่ ถ้าสอดคล้องกับสถานภาพในอนาคตจริง ก็แสดงว่ามีความตรงสูง

3. ความตรงเชิงโครงสร้าง (Construct Validity)

3.1 วิธีของคาร์เวอร์ (Carver)

คาร์เวอร์ (Carver, 1970 อ้างถึงใน สำริง บุญเรืองรัตน์, 2527 : 67) ได้เสนอวิธีการคำนวณค่าความตรงของแบบทดสอบอิงเกณฑ์จากข้อมูลของกลุ่มผู้เรียนก่อนสอนและหลังจากได้รับการสอน ดังนี้

		ก่อนสอน	หลังสอน
คะแนนจุดตัด	ผ่าน	b	a
	ไม่ผ่าน	c	d

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

เมื่อ	r_{ck}	แทน	ค่าความเที่ยงตรงของแบบทดสอบอิงเกณฑ์
	a	แทน	จำนวนนักเรียนที่สอบผ่านหลังจากที่เรียนแล้ว
	b	แทน	จำนวนนักเรียนที่สอบผ่านโดยที่ยังไม่ได้เรียน
	c	แทน	จำนวนนักเรียนที่สอบไม่ผ่าน โดยที่ยังไม่ได้เรียน
	d	แทน	จำนวนนักเรียนที่สอบไม่ผ่านหลังจากที่เรียนแล้ว
	N	แทน	จำนวนนักเรียนทั้งหมด (a+b+c+d)

เนื่องจากค่าความตรงจะมากหรือน้อยขึ้นอยู่กับจำนวนคนที่สอบไม่ผ่านก่อนสอน (c) และคนที่สอบผ่านหลังสอน (a) ดังนั้นค่าความตรงนี้จะขึ้นอยู่กับประสิทธิภาพในการสอนเป็นสำคัญ

3.2 วิธีของมิลแมน (Millman)

มิลแมน (Millman , 1974 : 370-371) เสนอวิธีหาความตรงเชิงโครงสร้าง โดยคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์แบบฟี (Phi – Correlation Coefficient) ดังนี้

$$\phi = \frac{ac - bd}{\sqrt{(a + b)(b + c)(c + d)(a + d)}}$$

เมื่อ	ϕ	แทน	ค่าความตรงเชิงโครงสร้างของแบบทดสอบอิงเกณฑ์
	a	แทน	จำนวนผู้สอบผ่านหลังสอน
	b	แทน	จำนวนผู้สอบผ่านก่อนสอน
	c	แทน	จำนวนผู้สอบไม่ผ่านก่อนสอน
	d	แทน	จำนวนผู้สอบไม่ผ่านหลังสอน

4. ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์

กลุ่มนักวัดผลหลายท่าน เช่น แฮมเบิลตัน และคณะ (Hambleton and other , 1980 quoter Berk , 1980 : 329) ได้แบ่งกลุ่มและนิยามความหมายของความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ไว้ 3 กลุ่ม ดังนี้ คือ

1. ความเชื่อมั่นในการตัดสินใจจำแนกความรู้ (Reliability of Master Classification Decision) เป็นการจำแนกความสอดคล้องในการตัดสินใจจำแนกความรู้ ไม่รอบรู้ จากผลการสอบซ้ำด้วยแบบทดสอบฉบับเดียวหรือแบบทดสอบคู่ขนาน ดัชนีที่วิเคราะห์เป็นดัชนีฟังก์ชันการสูญเสียเนื่องจากการกำหนดคะแนนจุดตัดล่วงหน้า
2. ความเชื่อมั่นของคะแนนของแบบทดสอบอิงเกณฑ์ (Reliability of Criterion Referenced Test Score) เป็นความสอดคล้องของคะแนนแต่ละคนที่เบี่ยงเบนจากคะแนนจุดตัดด้วยแบบทดสอบคู่ขนานหรือคู่ขนานแบบสุ่ม ดัชนีที่วิเคราะห์เป็นดัชนีฟังก์ชันการสูญเสียเนื่องจากความคลาดเคลื่อนในการยกกำลังสอง
3. ความเชื่อมั่นในการประมาณคะแนนมวลความรู้ (Reliability of Domain Score Estimates) เป็นการหาความสอดคล้องของคะแนนสอบแต่ละคนจากแบบทดสอบคู่ขนานสองฉบับ สถิติที่ใช้เป็นการประมาณค่าคะแนนมวลความรู้หรือสัดส่วนของการเดาตอบถูกในประชากรข้อสอบแต่ไม่เกี่ยวข้องกับมาตรฐานของความรู้

จากความหมายดังกล่าว จะเห็นได้ว่า คะแนนจุดตัดเป็นตัวประกอบที่สำคัญในการหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ โดยเฉพาะความเชื่อมั่นในการตัดสินใจจำแนกความรู้ และความเชื่อมั่นของคะแนนของแบบทดสอบอิงเกณฑ์

วิธีประมาณค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์

1. การประมาณค่าความเชื่อมั่นในการตัดสินใจจำแนกผู้รอบรู้ (Reliability of Mastery Classification Decision) เป็นการหาความสอดคล้องในการตัดสินใจจำแนกผู้รอบรู้ ไม่รอบรู้ โดยการสอบซ้ำด้วยแบบทดสอบฉบับเดียวกัน หรือแบบทดสอบคู่ขนาน ซึ่งมีผู้เสนอไว้หลายวิธี เช่น

1.1 วิธีของแฮมเบิลตัน และ โนวิก

แฮมเบิลตัน และ โนวิก (Hambleton and Novick, 1973 : 159-170) ได้เสนอสูตรในการประมาณค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์โดยใช้ผลรวมของสัดส่วนความสอดคล้องในการตัดสินใจจำแนกผู้รอบรู้ จากการสอบด้วยแบบทดสอบคู่ขนานหรือแบบทดสอบฉบับเดียวกัน

สองครั้ง สัดส่วนของความสอดคล้องในการตัดสินใจ ก็คือ สัดส่วนของการตัดสินใจว่ารอบรู้ตรงกัน ทั้งสองฉบับ หรือสองครั้ง และสัดส่วนของการตัดสินใจว่าไม่รอบรู้ตรงกันทั้งสองฉบับหรือสองครั้ง โดยเขียนเป็นสูตร ได้ดังนี้

$$P_0 = P_{11} + P_{12}$$

เมื่อ	P_0	แทน	สัดส่วนของความสอดคล้องในการตัดสินใจจำแนกผู้รอบรู้
	P_{11}	แทน	สัดส่วนของผู้ตัดสินใจว่ารอบรู้ตรงกันทั้งสองฉบับหรือสองครั้ง
	P_{12}	แทน	สัดส่วนของผู้ตัดสินใจว่าไม่รอบรู้ตรงกันทั้งสองฉบับหรือสองครั้ง

1.2 วิธีของคาร์เวอร์

คาร์เวอร์ (Carver, 1970 อ้างถึงในปราชญ์ ทองคำ , 2539 : 219) ได้เสนอวิธีการประมาณค่าความเชื่อมั่นของแบบทดสอบ โดยการคำนวณสัดส่วนของความสอดคล้องในการตัดสินใจ ความรอบรู้ของผู้สอบแต่ละคนจากแบบทดสอบคู่ขนาน หรือแบบทดสอบฉบับเดียวกัน 2 ครั้ง ตามแนวคิดนี้สามารถนำมาใช้ประมาณค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ได้ โดยการนำแบบทดสอบคู่ขนานสองฉบับไปสอบกับนักเรียนกลุ่มเดียวกัน แล้วนำผลการสอบไปแจกแจงลงในตาราง วิธีนี้ได้ผลเท่ากับวิธีของแฮมเบิลตัน และ โนวิด เพราะเป็นค่าสัดส่วนเดียวกัน

		แบบทดสอบ ข (หรือครั้งที่ 2)	
		ผ่าน	ไม่ผ่าน
แบบทดสอบ ก (หรือครั้งที่ 1)	ผ่าน	b	a
	ไม่ผ่าน	c	d

$$P_0 = \frac{b + d}{a + b + c + d}$$

เมื่อ	P_0	แทน	สัดส่วนของความสอดคล้องในการตัดสินใจความรอบรู้หรือค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์
-------	-------	-----	--

- a แทน สัดส่วนจำนวนคนที่ผ่านในครั้งที่ 1 แต่ไม่ผ่านในครั้งที่ 2
 b แทน สัดส่วนจำนวนคนที่ผ่านในครั้งที่ 1 และผ่านในครั้งที่ 2
 c แทน สัดส่วนจำนวนคนที่ไม่ผ่านในครั้งที่ 1 แต่ผ่านในครั้งที่ 2
 d แทน สัดส่วนจำนวนคนที่ไม่ผ่านในครั้งที่ 1 และไม่ผ่านในครั้งที่ 2

1.3 วิธีของสวามินาธาน , แฮมเบิลตัน และอัลจินา (Swaminathan , Hambleton and Algina)

วิธีนี้เป็นการหาความเชื่อมั่นโดยใช้นิยามของความเชื่อมั่นในรูปเดียวกันกับวิธีของคาร์เวอร์ และแฮมเบิลตัน และโนวิก แต่จะละเอียดกว่า 2 วิธีแรก เพราะทำการหักความสอดคล้องที่เกิดขึ้นโดยบังเอิญซึ่งอาจจะปนอยู่กับความสอดคล้องจริง อันเป็นเหตุให้ค่าความเชื่อมั่นที่คำนวณได้สูงกว่าที่เป็นจริง

วิธีนี้อาจนำแบบทดสอบฉบับเดียวไปสอบซ้ำกับกลุ่มเดิม หรืออาจนำแบบทดสอบอิงเกณฑ์คู่ขนาน 2 ฉบับ ไปทดสอบกับกลุ่มเดียวกันก็ได้ แล้วจัดลงตาราง ดังนี้ (บุญชม ศรีสะอาด, 2528 : 84-85)

สอบครั้งที่ 2 (หรือ ฉบับ ข.)

		สอบผ่าน	สอบไม่ผ่าน	รวม
สอบครั้งที่ 1 (หรือ ฉบับ ก.)	สอบผ่าน	a	b	a+b
	สอบไม่ผ่าน	d	c	c+d
	รวม	a+d	c+b	N

$$K = (P_o - P_e) / (1 - P_e)$$

- เมื่อ K แทน ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์
 P_o แทน สัดส่วนของความสอดคล้องในการจำแนกว่าเป็นผู้รอบรู้และผู้ไม่รอบรู้ = $(a+c) / N$
 P_e แทน สัดส่วนความสอดคล้องที่คาดหวังไว้
 $= (a+b)(a+d) + (b+c)(c+d) / N^2$

2. การประมาณค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์โดยพิจารณาจากความเชื่อมั่นของคะแนนจากแบบทดสอบอิงเกณฑ์ (Reliability of Criterion Referenced Test Scores) วิธีนี้เป็นการหาดัชนีความสอดคล้องระหว่างคะแนนของผู้สอบแต่ละคนที่เบี่ยงเบนไปจากคะแนนจุดตัดซึ่งมีผู้เสนอไว้หลายท่าน (บุญเชิด ภิญโญอนันตพงษ์, 2527 : 189-197 ; Lovett, 1978 : 239-251) เช่น

2.1 วิธีของลิวิงสตัน

ลิวิงสตัน (Livingston, 1972) ได้เสนอสูตรการคำนวณค่าความเชื่อมั่นโดยอาศัยนิยามความแปรปรวนของคะแนนจริงและคะแนนที่สอบได้ จากการหาค่าเบี่ยงเบนกำลังสองของคะแนนและค่าที่เบี่ยงเบนออกจากคะแนนจุดตัด ดังนี้

$$r_{cc} = \frac{S^2 r_{xx} + (\bar{X} - C)^2}{S^2 + (\bar{X} - C)^2}$$

เมื่อ	r_{cc}	แทน	ความเชื่อมั่นตามวิธีของลิวิงสตัน
	S^2, \bar{X}	แทน	ความแปรปรวนและค่าเฉลี่ยของคะแนนที่สอบได้
	r_{xx}	แทน	ค่าความเชื่อมั่นแบบอิงกลุ่ม อาจหาได้จากสูตร K-R 20 หรือ K-R 21
	C	แทน	คะแนนจุดตัด

เมื่อคะแนนจุดตัดเท่ากับคะแนนเต็ม ค่าความเชื่อมั่นนี้จะเท่ากับความเชื่อมั่นแบบอิงกลุ่มซึ่งจะมีค่าต่ำสุด ถ้าคะแนนทั้งหมดตกอยู่ที่คะแนนจุดตัดหรือคะแนนของทุกคนเท่ากัน กล่าวคือ ไม่มีความแปรปรวนของคะแนน จะไม่มีการนิยามดัชนีความเชื่อมั่นของลิวิงสตัน และระยะทางระหว่างคะแนนตัวหนึ่งกับคะแนนจุดตัดมากเท่าใด ค่าคะแนนจริงจะตกอยู่ในด้านเดียวกันกับคะแนนจุดตัดนั้น ดังนั้น ดัชนีความเชื่อมั่นของลิวิงสตัน มีความไวต่อระยะทางระหว่างคะแนนจุดตัดกับค่าเฉลี่ย อย่างไรก็ตามวิธีการนี้ใช้ได้ดีกับการกระจายของคะแนนที่มีฐานนิยมเดียว

แฮร์ริส (Harris, 1972 : 27) ได้ชี้ให้เห็นว่าวิธีการของลิวิงสตันให้ค่าความเชื่อมั่นที่สูงกว่าแบบเดิม แต่ทั้งสองแบบก็ให้ค่าความแปรปรวนของความคลาดเคลื่อนเท่ากัน ดังนั้น สัมประสิทธิ์ที่มีค่าสูงกว่าไม่ได้แสดงถึงความน่าเชื่อถือของการกำหนดว่าคะแนนจริงจะตกอยู่ได้หรือเหนือคะแนนจุดตัด

2.2 วิธีของเบรนนันและเคน

เบรนนันและเคน (Brennan and Kane, 1977) ได้เสนอสูตรการหาค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ ที่ได้มาจากประชากรผู้สอบและข้อสอบที่สุ่มมาจากประชากรข้อสอบ ซึ่งดัชนีความเชื่อถือได้สามารถอธิบายได้ด้วยทฤษฎีสรุปอ้างอิง เป็นการวิเคราะห์ส่วนย่อยของความแปรปรวนที่เป็นผลมาจากข้อสอบ รวมกับปฏิสัมพันธ์ของผู้สอบและข้อสอบ ถ้าการให้คะแนนเป็นตอบถูก 1 ตอบผิด 0 สูตรความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ คือ

$$M(c) = 1 - \frac{1}{n-1} \left[\frac{\bar{X}_{ij}(1 - \bar{X}_{ij}) - (S_{x_i}^2)}{(\bar{X}_{ij} - C) + (S_{x_i}^2)} \right]$$

เมื่อ	$M(c)$	แทน	ดัชนีความสามารถเชื่อถือได้
	\bar{X}_{ij}	แทน	ค่าเฉลี่ยคะแนนรวมทั้งหมด
	$S_{x_i}^2$	แทน	ค่าความแปรปรวนของค่าเฉลี่ยคะแนนของรายบุคคล
	C	แทน	คะแนนจุดตัดของแบบทดสอบ

2.3 โลเวทท์ (Lovett) ได้เสนอสูตรหาค่าความเชื่อมั่นของคะแนนแบบทดสอบอิงเกณฑ์ โดยการนำแบบทดสอบฉบับเดียว ไปทดสอบกับนักเรียนกลุ่มเดียว เพียงครั้งเดียว โดยมีสูตรในการคำนวณ 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

5. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการสร้างและพัฒนาแบบทดสอบ

ปัญญาลักษณ์ สุวรรณรัตน์ (2545 : บทคัดย่อ) ได้ทำการวิจัยเรื่อง การสร้างแบบทดสอบ วัดความสามารถทางคณิตศาสตร์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 5 ซึ่งมีจุดประสงค์เพื่อ สร้างและพัฒนาแบบทดสอบวัดความสามารถทางคณิตศาสตร์ สำหรับนักเรียนชั้นประถมศึกษา ปีที่ 5 รวมทั้งสร้างเกณฑ์ปกติของแบบทดสอบในรูปคะแนนที่ปกติ กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนสังกัดสำนักงาน การประถมศึกษาจังหวัดสงขลา จำนวน 1,170 คน ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น โดยใช้ขนาดของโรงเรียนเป็นชั้น และมีโรงเรียนเป็นหน่วยการสุ่ม

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดความสามารถทางคณิตศาสตร์ สำหรับ นักเรียนชั้นประถมศึกษาปีที่ 5 จำนวน 4 ด้าน คือ แบบทดสอบวัดความสามารถด้านความรู้ ความเข้าใจ จำนวน 35 ข้อ แบบทดสอบวัดความสามารถด้านทักษะการคิดคำนวณ จำนวน 30 ข้อ แบบทดสอบวัดความสามารถด้านกระบวนการทางคณิตศาสตร์ จำนวน 35 ข้อ และ แบบทดสอบวัดความสามารถด้านการแก้โจทย์ปัญหา จำนวน 35 ข้อ จากผลการวิจัย พบว่า

1. แบบทดสอบวัดความสามารถด้านความรู้ความเข้าใจ มีค่าความยาก ตั้งแต่ .33 - .73 ค่าอำนาจจำแนก ตั้งแต่ .26 - .64 ค่าความเชื่อมั่น .858 และมีค่าคะแนนที่ปกติ ตั้งแต่ T18 – T76

2. แบบทดสอบวัดความสามารถด้านทักษะการคิดคำนวณมีค่าความยาก ตั้งแต่ .26 - .77 ค่าอำนาจจำแนก ตั้งแต่ .26 - .75 ค่าความเชื่อมั่น .889 และมีค่าคะแนนที่ปกติ ตั้งแต่ T26 – T76

3. แบบทดสอบวัดความสามารถด้านกระบวนการทางคณิตศาสตร์มีค่าความยาก ตั้งแต่ .32 - .76 ค่าอำนาจจำแนก ตั้งแต่ .26 - .72 ค่าความเชื่อมั่น .824 และมีค่าคะแนนที่ปกติ ตั้งแต่ T18 – T82

4. แบบทดสอบวัดความสามารถด้านการแก้โจทย์ปัญหามีค่าความยาก ตั้งแต่ .25 - .75 ค่าอำนาจจำแนก ตั้งแต่ .25 - .62 ค่าความเชื่อมั่น .808 และมีค่าคะแนนที่ปกติ ตั้งแต่ T18 – T78

บุญประกอบ จายคำย (2545 : บทคัดย่อ) ได้ทำการวิจัยเรื่อง การสร้างแบบวัดภาคปฏิบัติ สำหรับโปรแกรมประมวลผลคำไมโครซอฟต์เวิร์ด ซึ่งมีจุดประสงค์เพื่อ 1) สร้างแบบวัด ภาคปฏิบัติสำหรับโปรแกรมประมวลผลคำไมโครซอฟต์เวิร์ด 2) เพื่อสร้างเกณฑ์ในการให้ คะแนนในการวัดภาคปฏิบัติ 3) เพื่อหาคุณภาพของแบบวัดภาคปฏิบัติที่ผู้วิจัยสร้างขึ้น

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ภาคเรียนที่ 1 ปี การศึกษา 2545 ของโรงเรียนสาธิตมหาวิทยาลัยสงขลานครินทร์ จำนวน 60 คน แบ่งออกเป็น 2 กลุ่ม กลุ่มละ 30 คน โดยกลุ่มตัวอย่างดังกล่าวเลือกมาโดยวิธีสุ่มอย่างง่าย (Sample random sampling) กลุ่มตัวอย่างกลุ่มที่ 1 ใช้ทดสอบเพื่อหาเวลาที่นักเรียนใช้ในการทำแบบวัด หากความถูกต้องชัดเจนของใบงาน ส่วนกลุ่มที่ 2 ใช้ทดสอบเพื่อหาค่าความเชื่อมั่นของแบบวัด และความเชื่อมั่นระหว่างผู้ประเมิน 2 คน

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ แบบวัดภาคปฏิบัติสำหรับโปรแกรมประมวลผลคำ ไมโครซอฟต์เวิร์ด ซึ่งประกอบด้วยใบงานสำหรับให้นักเรียนปฏิบัติ แบบประเมินและเกณฑ์ การให้คะแนน โดยแบบวัดดังกล่าวได้ผ่านการประเมินคุณภาพขั้นต้นโดยผู้เชี่ยวชาญ จำนวน 5 คน ซึ่งมีค่าดัชนีความสอดคล้องระหว่างข้อปฏิบัติกับจุดประสงค์เชิงพฤติกรรม ตั้งแต่ 0.80 – 1.00

การวิจัยครั้งนี้ได้แบบวัดภาคปฏิบัติ ดังนี้

1. แบบวัดที่สร้างขึ้นในการวิจัยครั้งนี้ คือ แบบวัดภาคปฏิบัติสำหรับโปรแกรม ประมวลผลคำไมโครซอฟต์เวิร์ด มีจำนวน 3 ฉบับ ซึ่งประกอบด้วย แบบวัดการจัดการข้อมูล แบบวัดการสร้างตาราง และแบบวัดการทำจดหมายเวียน ซองจดหมายและป้ายผนึก

แบบวัดการจัดการข้อมูลมี 2 ฉบับย่อยคือ แบบวัดการจัดการข้อมูล 1 และแบบวัด การจัดการข้อมูล 2

แบบวัดการทำจดหมายเวียน ซองจดหมายและป้ายผนึก มี 3 ฉบับย่อย คือ แบบวัด การทำจดหมายเวียน แบบวัดการทำซองจดหมาย และแบบวัดการทำป้ายผนึก

2. ผลจากการสร้างเกณฑ์ในการให้คะแนนในการวัดภาคปฏิบัติ ผู้วิจัยได้สร้างเกณฑ์การ ให้คะแนนของแบบวัดแต่ละข้อเป็น 4 ระดับ คือ 0, 1, 2 และ 3 คะแนน ซึ่งการสร้างเกณฑ์การ ให้คะแนนได้ผ่านการพิจารณาของผู้เชี่ยวชาญ โดยมีค่าดัชนีความสอดคล้องระหว่างข้อปฏิบัติ กับเกณฑ์การให้คะแนน ตั้งแต่ 0.80 – 1.00 และมีการกำหนดเกณฑ์การผ่านแบบวัดแต่ละฉบับ โดยผู้เชี่ยวชาญ ได้เกณฑ์การผ่านแบบวัดที่ระดับร้อยละ 80

3. การหาคุณภาพของแบบวัดภาคปฏิบัติ โดยการหาค่าความเชื่อมั่นของแบบวัดซึ่งคำนวณโดยสูตรสัมประสิทธิ์การสรูปอ้างอิง และหาค่าความเชื่อมั่นระหว่างผู้ประเมิน 2 คน คำนวณโดยใช้การประมาณค่าจากสูตรสัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson Product Moment) ผลจากการหาคุณภาพของแบบวัดปรากฏดังนี้

3.1 การหาความเชื่อมั่นของแบบวัด ปรากฏดังนี้ แบบวัดการจัดการข้อมูล 1 มีค่าความเชื่อมั่นเท่ากับ 0.928 แบบวัดการจัดการข้อมูล 2 มีค่าความเชื่อมั่นเท่ากับ 0.987 แบบวัดการสร้างตาราง มีค่าความเชื่อมั่นเท่ากับ 0.987 แบบวัดการทำจดหมายเวียน มีค่าความเชื่อมั่นเท่ากับ 0.952 แบบวัดการทำซองจดหมาย มีค่าความเชื่อมั่นเท่ากับ 0.941 และแบบวัดการทำป้ายผนึก มีค่าความเชื่อมั่นเท่ากับ 0.931

3.2 การหาความเชื่อมั่นระหว่างผู้ประเมิน 2 คน ปรากฏดังนี้ แบบวัดการจัดการข้อมูล 1 มีค่าความเชื่อมั่นเท่ากับ 0.928 แบบวัดการจัดการข้อมูล 2 มีค่าความเชื่อมั่นเท่ากับ 0.987 แบบวัดการสร้างตาราง มีค่าความเชื่อมั่นเท่ากับ 0.988 แบบวัดการทำจดหมายเวียน มีค่าความเชื่อมั่นเท่ากับ 0.952 แบบวัดการทำซองจดหมาย มีค่าความเชื่อมั่นเท่ากับ 0.937 และแบบวัดการทำป้ายผนึก มีค่าความเชื่อมั่นเท่ากับ 0.936

งานวิจัยที่เกี่ยวข้องกับการสร้างและพัฒนาแบบทดสอบอิงเกณฑ์

ทิพย์รัตน์ กิจเฉลา (2543 : บทคัดย่อ) พัฒนาแบบทดสอบอิงเกณฑ์วัดผลสัมฤทธิ์คณิตศาสตร์ เรื่องสมบัติของจำนวนนับ ชั้นมัธยมศึกษาปีที่ 1 โดยหาค่าความเที่ยงตรง ค่าความยาก และค่าอำนาจจำแนก หาค่าความเชื่อมั่น และกำหนดคะแนนจุดตัดที่เหมาะสมของแบบทดสอบ โดยศึกษาจากนักเรียนจำนวน 36 คน ชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2543 ของวิทยาลัยนาฏศิลปสุพรรณบุรี ผลการศึกษาพบว่า แบบทดสอบอิงเกณฑ์ที่สร้างขึ้นทั้งสองฉบับ มีความเที่ยงตรงตามเนื้อหาจากการพิจารณาของผู้เชี่ยวชาญ 5 ท่าน มีค่าเท่ากับ 0.80 ถึง 1.00 และ 0.60 ถึง 1.00 โดยมีค่าความยากก่อนเรียนเท่ากับ 0.12 ถึง 0.39, 0.22 ถึง 0.39 และหลังเรียนมีค่าเท่ากับ 0.78 ถึง 0.94, 0.78 ถึง 0.97 ตามลำดับ มีค่าอำนาจจำแนกโดยใช้สูตร Cox และ Vargas เท่ากับ 0.48 ถึง 0.77, 0.50 ถึง 0.78 มีคะแนนจุดตัดโดยใช้สูตรของแกลสมีค่าเท่ากับ 52%, 53% ค่าความเชื่อมั่นโดยใช้สูตรไบโนเมียลของโลเวท มีค่าเท่ากับ 0.73 และ 0.74 ตามลำดับ

สิรินธร สังขรัตน์ (2543 : บทคัดย่อ) ทำการวิจัยเรื่องการสร้างแบบทดสอบอิงเกณฑ์ วัดผลสัมฤทธิ์วิทยาศาสตร์ เรื่องชีวิตสัตว์ ชั้นมัธยมศึกษาปีที่ 1 การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อสร้างแบบทดสอบอิงเกณฑ์วัดผลสัมฤทธิ์วิทยาศาสตร์ เรื่องชีวิตสัตว์ ชั้นมัธยมศึกษาปีที่ 1 โดยหาค่าความยาก และอำนาจจำแนกรายข้อ กำหนดคะแนนจุดตัดที่เหมาะสมหาค่าความเชื่อมั่น และหลักฐานความเที่ยงตรงของแบบ ทดสอบ โดยศึกษาจากนักเรียนจำนวน 60 คน ชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2542 ภาคเรียนที่ 2 ของโรงเรียนวรราชาทินัดดามาตุวิทยา และโรงเรียนเตรียมอุดมศึกษา น้อมเกล้า นนทบุรี จำนวนเท่า ๆ กัน ผลการศึกษาพบว่า แบบทดสอบอิงเกณฑ์ที่สร้างขึ้น มีค่าหลักฐานความเที่ยงตรง เชิงเนื้อหาจากการพิจารณาของผู้เชี่ยวชาญ 5 ท่าน โดยการหาค่าเฉลี่ย และส่วน เบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 2.6 ถึง 3.8 และ 0.44 ถึง 1.00 ตามลำดับ ค่าความ ยากของกลุ่มรอบรู้มีค่าเท่ากับ 0.27 ถึง 1.00 และค่าความยากของกลุ่มไม่รอบรู้มีค่าเท่ากับ 0.03 ถึง 0.57 ค่าอำนาจจำแนกหาโดยสูตรของเบรนนอนมีค่าเท่ากับ 0.20 ถึง 0.80 หาคะแนนจุดตัดโดยสูตรของเกลสมีค่าเท่ากับ 56% (28 ใน 50 ข้อ) หาค่า ความเชื่อมั่นของแบบทดสอบโดยสูตรของบุญเชิด ภิญ โญอนันตพงษ์ มีค่าเท่ากับ 0.9474 และคำนวณหาหลักฐานความเที่ยงตรงเชิง โครงสร้างโดยใช้สูตรของคาร์เวอร์ มีค่า เท่ากับ 0.8677

อุทิศย์ อนุพรบริบูรณ์ (2543 : บทคัดย่อ) ทำการวิจัยเรื่อง การสร้างแบบทดสอบอิงเกณฑ์ วัดผลสัมฤทธิ์คณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อสร้างแบบทดสอบอิงเกณฑ์วัดผลสัมฤทธิ์คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่องสมการและการแก้สมการ กลุ่มตัวอย่างที่ใช้ในการศึกษา ครั้งนี้คือนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2543 โรงเรียนวัดศิริไอยสวรรค์ เขตบางพลัด กรุงเทพมหานคร จำนวน 30 คน แบบทดสอบประกอบด้วยแบบทดสอบย่อย 2 ฉบับ มีการหาคุณภาพ โดยการหาค่าความเที่ยงตรงเชิงเนื้อหา ระหว่างจุดประสงค์เชิง พฤติกรรมกับเนื้อหาโดยใช้ค่าดัชนีความสอดคล้อง การตัดสินใจของผู้เชี่ยวชาญ และ ค่า ความเที่ยงตรงระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรมโดยใช้ ค่าดัชนีความเหมาะสม ตามวิธีของโรวินลลี และ แฮมเบลตัน ค่าความยากง่ายก่อนสอน ค่าความยากง่ายหลังสอน ค่าอำนาจจำแนกโดยใช้ค่าดัชนีความไว (ดัชนี S) ค่าคะแนนจุดตัดโดยวิธีของแองกอฟ และ ค่าความเชื่อมั่นโดยวิธีของลิวินสตัน สรุปผลการศึกษาค้นคว้าได้ดังนี้ แบบทดสอบฉบับที่ 1 ประกอบด้วยเนื้อหาเรื่องรูปแบบสมการ และการแก้สมการ เป็น แบบทดสอบเลือกตอบ 4 ตัวเลือก จำนวน 30 ข้อ ค่าความเที่ยงตรงระหว่างจุดประสงค์เชิง พฤติกรรมกับเนื้อหา มีค่าอยู่ระหว่าง 0.67 - 1.00 ค่าความเที่ยงตรงระหว่างข้อสอบกับ จุดประสงค์เชิงพฤติกรรมมีค่าอยู่ระหว่าง 2.50 - 4.00 ค่าความยากง่ายก่อนสอนมีค่าอยู่ ระหว่าง 0.13 - 0.40

ค่าความยากง่ายหลังสอนมีค่าอยู่ระหว่าง 0.77 - 1.00 ค่าอำนาจ จำแนกมีค่าอยู่ระหว่าง 0.43 - 0.80 ค่าคะแนนจุดตัดมีค่าเท่ากับ 24 คะแนน ซึ่งมีค่า เท่ากับร้อยละ 80 ของคะแนนเต็ม ค่าความเชื่อมั่นมีค่าเท่ากับ 0.888 แบบทดสอบฉบับที่ 2 ประกอบด้วยเนื้อหาเรื่อง โจทย์ปัญหา สมการ เป็นแบบทดสอบแบบ เลือกตอบ 4 ตัวเลือก จำนวน 30 ข้อ ค่าความเที่ยงตรงระหว่าง จุดประสงค์เชิงพฤติกรรมกับ เนื้อหามีค่าอยู่ระหว่าง 0.67 - 1.00 ค่าความเที่ยงตรงระหว่างข้อสอบ กับจุดประสงค์เชิง พฤติกรรม มีค่าอยู่ระหว่าง 2.50 - 4.00 ค่าความยากง่ายก่อนสอนมีค่าอยู่ระหว่าง 0.43 - 0.74 ค่าคะแนนจุดตัด มีค่าเท่ากับ 24 คะแนน ซึ่งมีค่าเท่ากับร้อยละ 80 ของ คะแนนเต็ม ค่าความเชื่อมั่นมีค่าเท่ากับ 0.919

วีระ ประเสริฐ (2538 : บทคัดย่อ) ทำการวิจัยเรื่อง การสร้างแบบทดสอบอิงเกณฑ์ วัดทักษะการคิดคำนวณชั้นประถมศึกษาปีที่ 6 การศึกษาครั้งนี้ มีจุดมุ่งหมายเพื่อสร้าง แบบทดสอบ อิงเกณฑ์วัดทักษะการคิดคำนวณเรื่องการบวก การลบ การคูณ การหาร จำนวนนับ เศษส่วน และทศนิยม ในชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับ ปรับปรุง พ.ศ.2533) พร้อมทั้งหาคุณภาพและคะแนนเกณฑ์ ที่เหมาะสมของแบบทดสอบ กลุ่มตัวอย่างที่ใช้ในการศึกษา เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2537 ของ โรงเรียน ในสังกัดสำนักงานการประถมศึกษาจังหวัดพิษณุโลก จำนวน 247 คน ซึ่งได้มาโดยการสุ่มแบบ หลายขั้นตอน ผลการศึกษาได้แบบทดสอบจำนวน 15 ฉบับๆ ละ 10 ข้อ ซึ่งแบบทดสอบทุกฉบับ มีค่าความตรงเชิงเนื้อหาเท่ากับ 1.00 ค่าความยากง่ายข้ออยู่ระหว่าง .23 ถึง .78 และค่าเฉลี่ย ของค่า ความยากง่ายฉบับอยู่ระหว่าง .30 ถึง .71 ค่าอำนาจ จำแนกรายข้ออยู่ระหว่าง .34 ถึง .85 และค่าเฉลี่ย ของ ค่าอำนาจจำแนกรายฉบับอยู่ระหว่าง .60 ถึง .69 ค่าความ ตรงเชิงโครงสร้างอยู่ระหว่าง .85 ถึง 1.00 ค่าความเที่ยง อยู่ระหว่าง .78 ถึง .91 คะแนนเกณฑ์อยู่ระหว่างร้อยละ 30 ถึง 80 แบบทดสอบที่สร้างขึ้นเป็นแบบทดสอบที่มีคุณภาพสามารถ นำไปใช้วัดผล และ ประเมินผล ทักษะการคิดคำนวณ สำหรับ นักเรียนชั้นประถมศึกษาปีที่ 6 ของโรงเรียนในสังกัด สำนักงาน การประถมศึกษาจังหวัดพิษณุโลกได้

สุรพล ฝืนศิริ (2534 : บทคัดย่อ) ทำการวิจัยเรื่อง การสร้างแบบทดสอบอิงเกณฑ์ วัดผลสัมฤทธิ์วิทยาศาสตร์ ชั้น ปวช.1 เรื่องการเคลื่อนที่ มีจุดมุ่งหมาย เพื่อสร้างแบบทดสอบ อิงเกณฑ์ วัดผลสัมฤทธิ์วิทยาศาสตร์ ชั้น ปวช.1 เรื่องการเคลื่อนที่ หาเกณฑ์ที่เหมาะสมของ แบบทดสอบที่สร้างขึ้น วิธีการวิจัย กลุ่มตัวอย่าง สุ่มอย่างง่ายจากนักเรียนชั้น ปวช.1 ภาคเรียนที่ 1

ปีการศึกษา 2534 วิทยาลัยเทคนิคมหาสารคาม 280 คน แบบทดสอบมี 36 ข้อ มีค่าอำนาจจำแนก อยู่ระหว่าง 0.15 ถึง 0.61 ค่าความเที่ยงตรงเชิงโครงสร้างของแต่ละจุดประสงค์อยู่ระหว่าง 0.54 ถึง 0.76 ทั้งฉบับเท่ากับ 0.78 และค่าความเชื่อมั่นของแต่ละจุดประสงค์อยู่ระหว่าง 0.97 ถึง 1.00 ทั้งฉบับเท่ากับ 0.50 มีคะแนนเกณฑ์ที่เหมาะสมเท่ากับ 17 คะแนน หรือร้อยละ 47.22 ของคะแนนเต็ม

วิระพงษ์ ศรีชาติ (2535 : บทคัดย่อ) ทำการวิจัยเรื่อง การสร้างแบบทดสอบอิงเกณฑ์ วัดผลสัมฤทธิ์ภาษาอังกฤษด้านทักษะการอ่าน เพื่อความเข้าใจ ชั้นมัธยมศึกษาปีที่ 3 มีจุดมุ่งหมาย เพื่อสร้างแบบทดสอบอิงเกณฑ์วัดผลสัมฤทธิ์ภาษาอังกฤษด้านทักษะการอ่าน เพื่อความเข้าใจ ชั้นมัธยมศึกษาปีที่ 3 หาคะแนนเกณฑ์ที่เหมาะสมของแบบทดสอบอิงเกณฑ์ที่สร้างขึ้น และหา คุณภาพ ของแบบทดสอบ อิงเกณฑ์ที่สร้างขึ้น วิธีการวิจัย กลุ่มตัวอย่าง ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2534 ของ โรงเรียน มัธยมศึกษาในอำเภอสว่างแดนดิน จังหวัดสกลนคร ซึ่งสุ่มมาโดยวิธีการสุ่มแบบหลายขั้นตอน ได้นักเรียน 131 คน จาก 4 โรงเรียน แบ่งเป็น 2 กลุ่ม คือ กลุ่มแรก 68 คน เพื่อคัดเลือก และปรับปรุงข้อสอบ และกลุ่มหลังจำนวน 63 คน เพื่อหาคุณภาพของแบบทดสอบ แบบทดสอบ ที่สร้างขึ้นเป็นแบบทดสอบอิงเกณฑ์วัดผลสัมฤทธิ์ภาษาอังกฤษด้านทักษะการอ่าน เพื่อความเข้าใจ ชั้น มัธยมศึกษาปีที่ 3 จำนวน 4 ฉบับ เป็นแบบทดสอบเลือกตอบชนิด 5 ตัวเลือก ซึ่งได้ดำเนินการ ทดสอบ 2 ครั้ง เพื่อคัดเลือกและปรับปรุงข้อสอบ ทำการวิเคราะห์คุณภาพ

ผลการวิจัย ฉบับที่ 1 แบบทดสอบ เรื่อง การจับใจความ จำนวน 21 ข้อ มีค่าอำนาจจำแนก (ค่าดัชนี S) อยู่ในช่วง .206-.698 ค่าความเที่ยงตรงตามโครงสร้างเท่ากับ .968 ค่าความเชื่อมั่น เท่ากับ .864 และคะแนนเกณฑ์เป็น 38 เปอร์เซนต์ของคะแนนเต็ม ฉบับที่ 2 แบบทดสอบ เรื่อง การสรุปความ จำนวน 12 ข้อ มีค่าอำนาจจำแนก (ค่าดัชนี S) อยู่ในช่วง .270-.825 ค่าความเที่ยงตรงตามโครงสร้างเท่ากับ .968 ค่าความเชื่อมั่นเท่ากับ .721 และคะแนนเกณฑ์เป็น 58 เปอร์เซนต์ของคะแนนเต็ม ฉบับที่ 3 แบบทดสอบ เรื่อง การตีความ จำนวน 15 ข้อ มีค่าอำนาจจำแนก (ค่าดัชนี S) อยู่ในช่วง .222-.619 ค่าความเที่ยงตรงตามโครงสร้างเท่ากับ .968 ค่าความเชื่อมั่นเท่ากับ .417 และคะแนนเกณฑ์เป็น 53 เปอร์เซนต์ของคะแนนเต็ม ฉบับที่ 4 แบบทดสอบ เรื่อง การสรุปความ เมื่อกำหนดเวลาในการอ่าน จำนวน 12 ข้อ มีค่าอำนาจจำแนก (ค่าดัชนี S) อยู่ในช่วง .26 - .460 ค่าความเที่ยงตรงตามโครงสร้างเท่ากับ .873 ค่าความเชื่อมั่นเท่ากับ .570 และคะแนนเกณฑ์เป็น 58 เปอร์เซนต์ของคะแนนเต็ม

เสาวณิต ร่มศรี (2541 : บทคัดย่อ) ทำการวิจัยเรื่อง การสร้างแบบทดสอบอิงเกณฑ์ วัดผลสัมฤทธิ์คณิตศาสตร์ เรื่องเมตริกซ์ ชั้นมัธยมศึกษาปีที่ 5 การวิจัยครั้งนี้มีความมุ่งหมายเพื่อสร้างแบบทดสอบอิงเกณฑ์ วัดผลสัมฤทธิ์คณิตศาสตร์ เรื่องเมตริกซ์ ชั้นมัธยมศึกษาปีที่ 5 หาคุณภาพและ หาคะแนนเกณฑ์ที่เหมาะสมของแบบทดสอบ ที่สร้างขึ้น กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 ภาคเรียนที่ 1 ปีการศึกษา 2540 ของโรงเรียนสังกัดกรมสามัญศึกษา ในอำเภอจตุรพักตรพิมาน จังหวัดร้อยเอ็ด จำนวน 176 คน โดยวิธีการสุ่มอย่างง่าย แบบทดสอบที่สร้างขึ้นมี 8 ฉบับ แต่ละข้อมี 5 ตัวเลือก ผลการวิเคราะห์คุณภาพของแบบทดสอบแต่ละฉบับปรากฏดังนี้แบบทดสอบฉบับที่ 1 เรื่องสัญลักษณ์ของเมตริกซ์ จำนวน 15 ข้อ มีค่าอำนาจ จำแนก ระหว่าง .25 ถึง .66 ค่าความเที่ยงตรงเชิงโครงสร้าง เท่ากับ .893 ค่าความเชื่อมั่น เท่ากับ .870 และคะแนนเกณฑ์ 8 คะแนน คิดเป็นร้อยละ 53.33แบบทดสอบฉบับที่ 2 เรื่องการเท่ากันของเมตริกซ์ จำนวน 10 ข้อ มีค่าอำนาจ จำแนกระหว่าง .27 ถึง .64 ค่าความเที่ยงตรงตามโครงสร้าง เท่ากับ .848 ค่าความ เชื่อมั่น เท่ากับ .772 และคะแนนเกณฑ์ 5 คะแนน คิดเป็นร้อยละ 50 แบบทดสอบฉบับที่ 3 เรื่องการบวกเมตริกซ์ จำนวน 10 ข้อ มีค่าอำนาจจำแนก ระหว่าง .28 ถึง .56 ค่าความเที่ยงตรงตามโครงสร้าง เท่ากับ .826 ค่าความเชื่อมั่น เท่ากับ .857 และคะแนนเกณฑ์ 5 คะแนน คิดเป็นร้อยละ 50 แบบทดสอบฉบับที่ 4 เรื่องการคูณเมตริกซ์ด้วยจำนวนจริง จำนวน 20 ข้อ มีค่า อำนาจจำแนกระหว่าง .22 ถึง .51 ค่าความเที่ยงตรงตามโครงสร้าง เท่ากับ .831 ค่า ความเชื่อมั่น เท่ากับ .816 และคะแนนเกณฑ์ 10 คะแนน คิดเป็นร้อยละ 50แบบทดสอบฉบับที่ 5 เรื่องการคูณเมตริกซ์ด้วยเมตริกซ์ จำนวน 10 ข้อ มีค่า อำนาจจำแนกระหว่าง .27 ถึง .49 ค่าความ เที่ยงตรงตาม โครงสร้าง เท่ากับ .792 ค่าความเชื่อมั่น เท่ากับ .886 และคะแนนเกณฑ์ 4 คะแนน คิดเป็นร้อยละ 40แบบทดสอบฉบับที่ 6 เรื่องดีเทอร์มิแนนต์ จำนวน 15 ข้อ มีค่าอำนาจจำแนก ระหว่าง .31 ถึง .71 ค่าความเที่ยงตรงตามโครงสร้าง เท่ากับ .882 ค่าความเชื่อมั่น เท่ากับ .915 และคะแนนเกณฑ์ 6 คะแนน คิดเป็นร้อยละ 40แบบทดสอบฉบับที่ 7 เรื่องการหาอินเวอร์สการคูณของเมตริกซ์ จำนวน 15 ข้อ มีค่าอำนาจจำแนกระหว่าง .27 ถึง .63 ค่าความเที่ยงตรงตามโครงสร้าง เท่ากับ .882 ค่าความเชื่อมั่น เท่ากับ .750 และคะแนนเกณฑ์ 8 คะแนน คิดเป็นร้อยละ 53.33 แบบทดสอบ ฉบับ ที่ 8 เรื่องการใช้เมตริกซ์แก้ระบบสมการเชิงเส้น จำนวน 10 ข้อ มีค่าอำนาจ จำแนกระหว่าง .28 ถึง .56 ค่าความเที่ยงตรงตามโครงสร้าง เท่ากับ .860 ค่าความเชื่อมั่น เท่ากับ .859 และคะแนนเกณฑ์ 4 คะแนน คิดเป็นร้อยละ 40