บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การศึกษาวิจัยครั้งนี้เป็นการศึกษาคะแนนจุดตัด และความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ เมื่อใช้วิธีหาคะแนนจุดตัด วิธีหาความเชื่อมั่นด้วยวิธีต่างกัน การจัดเรียงข้อสอบด้วยวิธีต่างกัน ผู้วิจัยขอเสนอเอกสารและงานวิจัยที่เกี่ยวข้องตามลำดับหัวข้อดังนี้

1. ความหมายของการวัดผลแบบอิงเกณฑ์
2. การสร้างแบบทดสอบอิงเกณฑ์
3. ความหมายของคะแนนจุดตัด
4. วิธีการกำหนดคะแนนจุดตัด
5. ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์
6. การเรียงข้อคำถามในแบบทดสอบ
7. งานวิจัยที่เกี่ยวข้อง

ความหมายของการวัดผลแบบอิงเกณฑ์

นักการศึกษาและนักวัดผลได้ให้ความหมายของการวัดผลแบบอิงเกณฑ์ไว้ดังนี้

อังคณา สายยศ (2533 : 1) กล่าวว่า การวัดผลแบบอิงเกณฑ์เป็นการวัดผลสัมฤทธิ์ทางการเรียนโดยการเอาผลการสอบของนักเรียนแต่ละคนไปเทียบกับเกณฑ์ในจุดมุ่งหมายของการสอน

อนันต์ ศรีโสภา (2525 : 190) กล่าวว่า แบบทดสอบอิงเกณฑ์ใช้สำหรับวัดความรู้ความสามารถของนักเรียนแต่ละคนว่า ถึงเกณฑ์ขั้นต่ำที่คาดหวังไว้หรือไม่ บางทีเรียกแบบทดสอบนี้ว่าแบบทดสอบวัดความรู้ (Mastery Test) แบบทดสอบชนิดนี้เน้นการวัดความรู้ และทักษะต่าง ๆ ในตัวนักเรียนมีถึงเกณฑ์ที่กำหนดไว้หรือไม่ โดยไม่คำนึงถึงความแตกต่างของนักเรียน

บุญเชิด ภิญโญอนันตพงษ์ (2527 : 4) กล่าวว่า การวัดผลแบบอิงเกณฑ์เป็นการวัดผลซึ่งแปลความหมายของผลการวัด โดยการนำเอาผลการวัดของผู้สอบไปเทียบกับมาตรฐานที่แท้จริงว่าอยู่ในระดับใด ถึงมาตรฐานที่ยอมรับได้หรือไม่

สงบ ลักษณะ (2523 : 1-2) ได้กล่าวถึงลักษณะของแบบทดสอบอิงเกณฑ์ว่า คือ การทดสอบที่สามารถนำคะแนนมาแปลความหมายได้ว่า ผู้สอบมีผลสัมฤทธิ์ในการกระทำพฤติกรรม

อะไรได้บ้าง ลักษณะของแบบทดสอบอิงเกณฑ์ คือ

1. เป็นแบบทดสอบที่สร้างบนพื้นฐานของจุดประสงค์ หรือสิ่งที่ต้องการให้สอบปฏิบัติที่ได้รับการนิยามไว้อย่างชัดเจน

2. รายข้อของแบบทดสอบจะต้องวัดจุดประสงค์

3. คะแนนที่ได้แปลความหมายว่า ผู้สอบสามารถบรรลุตามเกณฑ์ที่กำหนดไว้หรือไม่

กรอนลันด์ (Gronlund. 1988 : 117) กล่าวว่า การวัดผลแบบอิงเกณฑ์ เป็นการวัดความรู้หรือขีดความสามารถขั้นต่ำสุด ซึ่งแปลความหมายได้จากคะแนนจุดตัดหรือมาตรฐานที่กำหนดไว้ล่วงหน้า

เกลเซอร์ (Glaser. 1963 : 519) กล่าวว่า การวัดผลแบบอิงเกณฑ์เป็นการวัดผลสัมฤทธิ์ที่ยึดหลักมาตรฐานของผู้สอบอย่างสมบูรณ์ เป็นการตรวจสอบว่าผู้สอบมีความรอบรู้หรือไม่ เมื่อเทียบกับเกณฑ์ที่กำหนดหรือคะแนนจุดตัดของความรอบรู้ซึ่งไม่ขึ้นอยู่กับผลการสอบของบุคคลอื่น

จากความหมายและลักษณะของแบบทดสอบอิงเกณฑ์ สรุปได้ว่า แบบทดสอบอิงเกณฑ์เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่สร้างขึ้นตามจุดประสงค์การสอน และมีการแปลความหมายคะแนน โดยนำคะแนนของผู้สอบแต่ละคนไปเทียบกับคะแนนจุดตัด ซึ่งเป็นเกณฑ์ตัดสินว่าผู้สอบมีความรอบรู้หรือไม่รอบรู้

<u>การสร้างแบบทดสอบอิงเกณฑ์</u>

นักวิชาการและนักวัดผลได้เสนอแนะแนวทางในการสร้างแบบทดสอบอิงเกณฑ์ไว้ดังนี้

อังคณา สายยศ (2526 : 25-29) ได้กล่าวถึงขั้นตอนในการสร้างแบบทดสอบอิงเกณฑ์ไว้ดังนี้

1. วิเคราะห์หลักสูตรหรือวิเคราะห์งาน นำหลักสูตรในวิชาที่ต้องการจะเขียนข้อสอบมาทำการวิเคราะห์โดยนำเนื้อหาย่อย ๆ กับพฤติกรรมมาหาความสัมพันธ์กัน

2. เขียนจุดมุ่งหมายเชิงพฤติกรรม จากตารางวิเคราะห์หลักสูตรโดยนำพฤติกรรมหลักและเนื้อหามาเขียนเป็นจุดมุ่งหมายเชิงพฤติกรรม

3. เขียนคำถามหรือข้อสอบ เมื่อมีจุดมุ่งหมายเชิงพฤติกรรมแล้ว ก็ดำเนินการเขียนข้อสอบให้สอดคล้องกัน จำนวนข้อสอบในแต่ละพฤติกรรมนั้นควรจะออกประมาณ 5-10

ข้อ ซึ่งขึ้นอยู่กับผู้สอนหรือผู้ออกข้อสอบเป็นประการสำคัญ โดยยึดว่าจุดมุ่งหมายในเนื้อหานั้น ควรจะออกกี่ข้อจึงจะเป็นเครื่องบ่งชี้ว่า เด็กได้เรียนรู้หรือรอบรู้แล้ว

โกวิท ประวาลพฤกษ์ (2522 : 124-149) ได้เสนอแนะวิธีการสร้างแบบทดสอบอิงเกณฑ์ไว้ดังนี้

1. กำหนดจุดมุ่งหมายทั่วไปของการเรียนรู้ในเนื้อหาตอนหนึ่ง ถ้ามีตารางวิเคราะห์หลักสูตรอยู่แล้วก็นำเอามาใช้ได้เลย พฤติกรรมที่พึงประสงค์จะครอบคลุมไปถึงพฤติกรรมชั้นสูง ๆ ด้วย

2. นำจุดมุ่งหมายทั่วไปมาเขียนเป็นจุดประสงค์เชิงพฤติกรรม

3. จัดสร้างข้อสอบตามจุดประสงค์เชิงพฤติกรรมไว้มาก ๆ เพื่อนำเอาไปทดสอบกับนักเรียนได้ตลอดเวลา

บุญเชิด ภิญโญอนันตพงศ์ (2527 : 49) ได้สรุปขั้นตอนของการสร้างแบบทดสอบอิงเกณฑ์ไว้ดังนี้

ขั้นที่ 1 วิเคราะห์เนื้อหาและจุดประสงค์

ขั้นที่ 2 แปลงจุดประสงค์หลักให้เป็นจุดประสงค์เฉพาะเจาะจง

ขั้นที่ 3 เขียนข้อสอบหรือผลิตข้อสอบ

ขั้นที่ 4 ตรวจทานข้อสอบ

ขั้นที่ 5 วิเคราะห์ข้อสอบ

ขั้นที่ 6 คัดเลือกข้อสอบ กำหนดความยาวและคะแนนจุดตัดของแบบทดสอบ

ขั้นที่ 7 วิเคราะห์แบบทดสอบ

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

1. เขียนสมรรถภาพที่จะวัดให้ชัดเจน โดยเขียนในรูปของจุดประสงค์เชิงพฤติกรรมที่สามารถสังเกตได้ และสอบวัดได้

2. กำหนดการรอบรู้ในสมรรถภาพนั้น

3. คิดสถานการณ์ เพื่อให้นักเรียนได้แสดงถึงความสามารถที่ต้องการวัดแล้วเขียนข้อสอบตามสถานการณ์นั้น

4. ตัดสินความรอบรู้ของนักเรียน โดยเทียบกับมาตรฐานการปฏิบัติขั้นต่ำสุดตามจุดประสงค์

จากที่กล่าวมาทั้งหมด สรุปขั้นตอนการสร้างแบบทดสอบอิงเกณฑ์ ได้ดังนี้

1. วิเคราะห์หลักสูตรหรือวิเคราะห์รายละเอียด โดยสร้างตารางวิเคราะห์หลักสูตร เพื่อหาความสัมพันธ์ระหว่างเนื้อหาย่อย ๆ กับพฤติกรรมหลัก

2. เขียนจุดมุ่งหมายเชิงพฤติกรรม จากตารางวิเคราะห์หลักสูตร โดยเขียนให้สอดคล้องกับเนื้อหา และพฤติกรรมหลัก

3. เขียนข้อสอบจากจุดมุ่งหมายเชิงพฤติกรรม

4. พิจารณาความสอดคล้องของข้อสอบกับจุดประสงค์เชิงพฤติกรรม โดยให้ผู้เชี่ยวชาญในเนื้อหาเป็นผู้พิจารณา

5. คัดเลือกข้อสอบที่สอดคล้อง และครอบคลุมจุดประสงค์เชิงพฤติกรรม

6. หาคุณภาพของแบบทดสอบ ในด้านความยากง่าย อำนาจจำแนก ความเชื่อมั่น และความตรงของแบบทดสอบ

<u>ความหมายของคะแนนจุดตัด</u>

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

อังคณา สายยศ (2525 : 70) ให้ความหมายของคะแนนจุดตัดว่า หมายถึง คะแนนที่น้อยที่สุดที่นักเรียนจะต้องทำได้ในการที่จะได้รับการตัดสินให้เป็นผู้รอบรู้

บุญเชิด ภิญโญอนันตพงษ์ (2527 : 110) กล่าวว่า เกณฑ์หรือมาตรฐาน หมายถึง คะแนนจุดตัดของแบบทดสอบที่ใช้แบ่งผู้เรียนออกเป็นผู้สอบได้สอบตก

ดำรง ศิริเจริญ (2529 : 139) กล่าวถึง คะแนนจุดตัดว่าหมายถึงคะแนนที่น้อยที่สุดที่นักเรียนจะต้องทำได้ในการที่จะได้รับการตัดสินให้ผู้รอบรู้ คะแนนเกณฑ์อาจจะอยู่ในรูปของจำนวนข้อที่ผู้เรียนทำถูกในแต่ละจุดประสงค์

แฮมเบิลตัน (Hambleton, 1978 : 279) กล่าวว่า มาตรฐาน (Standard) หมายถึง คะแนนที่ได้จากการสอบที่ใช้ในการแยกผู้สอบออกเป็น 2 กลุ่ม ตามระดับความสามารถที่แตกต่างกัน คือ เป็นกลุ่มที่รอบรู้ (Masters) และกลุ่มที่ไม่รอบรู้ (Non-Masters)

เบอร์ก (Berk, 1986 : 138) กล่าวว่า คะแนนจุดตัด (Cut-off-Score) หมายถึง จุดที่ใช้แบ่งคะแนนออกเป็น 2 ส่วน ส่วนหนึ่งเป็นคะแนนของผู้สอบที่จัดว่าเป็นผู้รอบรู้ อีกส่วนหนึ่งเป็นคะแนนของผู้สอบที่จัดว่าเป็นผู้ไม่รอบรู้ เนื่องจากคะแนนจุดตัดเป็นค่าที่ได้จากการสังเกต เพราะฉะนั้นการกำหนดคะแนนจุดตัด จึงย่อมมีความคลาดเคลื่อนมาตรฐาน (Standard Error) สำหรับความคลาดเคลื่อนนี้ เบอร์กแบ่งออกเป็น 2 ประเภท คือ

1. ผู้สอบที่มีความรอบรู้อย่างแท้จริง แต่ถูกจัดประเภทไม่มีความรอบรู้ ความคลาดเคลื่อนนี้เกิดจากการกำหนดคะแนนจุดตัดสูงเกินไป จึงทำให้ผู้สอบที่มีความรอบรู้สอบไม่ผ่านเกณฑ์ที่กำหนด

2. ผู้สอบที่ไม่มีความรอบรู้อย่างแท้จริงถูกจัดประเภทมีความรอบรู้ ความคลาดเคลื่อนนี้เกิดจากการกำหนดคะแนนจุดตัดต่ำเกินไป จึงทำให้ผู้สอบที่ไม่มีความรอบรู้ผ่านเกณฑ์ที่กำหนด

กล่าวโดยสรุปแล้ว คะแนนจุดตัด คือ คะแนนที่เป็นเกณฑ์ต่ำสุดที่ใช้ตัดสินให้ผู้สอบเป็นผู้รอบรู้หรือไม่รอบรู้

จากความหมายของคะแนนจุดตัดดังกล่าวข้างต้น จะเห็นว่าการกำหนดคะแนนจุดตัดเป็นสิ่งจำเป็น ซึ่งสืบเนื่องมาจากการวัดผลแบบอิงเกณฑ์จะต้องทำการประเมินให้สอดคล้องกับวัตถุประสงค์การเรียนการสอน (Objective-Based Instructional Programs) ที่กำหนดขึ้น หากพิจารณาในเชิงทฤษฎีจะเห็นว่า ในการที่จะปรับปรุงคุณภาพของการเรียนการสอน จะต้องเริ่มจาก 1) การกำหนดวัตถุประสงค์ของหลักสูตร 2) การจัดการเรียนการสอนให้สอดคล้องกับวัตถุประสงค์ 3) การประเมินผลการเรียนการสอน และจากการศึกษาพบว่า การจัดการเรียนการสอนตามวัตถุประสงค์ที่กำหนดจะทำให้ผลสัมฤทธิ์ทางการเรียนของผู้เรียนสูงกว่าการจัดการเรียนการสอนที่ไม่เป็นไปตามวัตถุประสงค์ เพราะฉะนั้นในการที่จะประเมินผลสัมฤทธิ์ทางการเรียนจึงมีความจำเป็นที่จะต้องกำหนดคะแนนจุดตัด (Hambleton, 1978 : 280-281)

วิธีการกำหนดคะแนนจุดตัด

การกำหนดคะแนนจุดตัดได้แบ่งออกเป็นหลายลักษณะตามแนวคิดของนักการศึกษาและนักวัดผลแต่ละท่านที่ได้เสนอไว้ดังนี้

มิลล์แมน (Millman. 1973 อ้างถึงใน อุดม ชูลีวรรณ, 2539 : 19-20) ได้สรุปแนวคิดของการกำหนดเกณฑ์มาตรฐานไว้ว่า ควรคำนึงถึงสิ่งต่าง ๆ ต่อไปนี้

1. ระดับความสามารถในการทำแบบทดสอบของคนอื่น ๆ (Performance of Other) โดยเฉพาะผลการสอบของกลุ่มที่ผ่านการยอมรับแล้ว ย่อมนำมาใช้เป็นเกณฑ์ในการพิจารณากำหนดมาตรฐานได้

2. เนื้อหาสาระของข้อสอบในแบบทดสอบ (Item Content) เนื้อหาของข้อสอบเป็นส่วนสำคัญส่วนหนึ่งในการที่จะพิจารณากำหนดมาตรฐาน ข้อสอบที่วัดเนื้อหาที่มีความสำคัญต้องให้ผู้สอบทำได้ในสัดส่วนที่สูง แต่ในส่วนของเนื้อหาที่ไม่สำคัญนัก หรือมีความยากเกินไปเกณฑ์มาตรฐานก็ควรลดต่ำลง

3. ผลของการศึกษาที่จะตามมาจากการกำหนดเกณฑ์มาตรฐาน (Educational Consequences) ถ้าหากว่า กำหนดเกณฑ์มาตรฐานต่ำเกินไปอาจมีผลเสียคือ ทำให้ผู้เรียนไม่สามารถเรียนบทเรียนที่สูงต่อไปได้ แต่ถ้ากำหนดมาตรฐานสูงเกินไป จะมีผลทำให้มีผู้ไม่รอบรู้ตกค้างเป็นจำนวนมาก

4. ผลทางจิตวิทยาและสภาพทางการเงิน (Psychological and Financial Cost) การกำหนดเกณฑ์มาตรฐาน ควรจะได้คำนึงถึงผลทางจิตใจที่เกิดขึ้นกับผู้สอบ เช่น แรงจูงใจ ความเบื่อหน่าย การทำลายอัตมโนทัศน์ และรวมถึงค่าใช้จ่ายในการสอนซ่อมเสริมให้กับผู้ที่ไม่ผ่านเกณฑ์

5. ความคลาดเคลื่อนในการเดา และการสุ่มข้อสอบ (Error due to chance and Item Sampling) การกำหนดมาตรฐานควรคำนึงถึงโอกาสที่ผู้สอบจะเดาถูก รวมทั้งลักษณะการสุ่มข้อสอบจากประชากรข้อสอบทั้งหมดในขอบเขตเนื้อหาที่กำหนดขึ้นด้วย

แกลส (Glass, 1978 : 243-257) ได้แบ่งวิธีการกำหนดคะแนนจุดตัด ออกเป็น 6 วิธี คือ

1. การใช้คะแนนของคนอื่น ๆ (Performance of Others)
2. การใช้วิธีนับถอยหลังจาก 100% (Counting Backwards from 100%)
3. การปรับคะแนนเกณฑ์อื่น ๆ (Bootstrapping on Other Criterion Scores)
4. การตัดสินใจโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบ (Judging Minimal Competence)
5. การใช้ทฤษฎีการตัดสินใจ (Decision-Theoretic Approaches)
6. การใช้วิธีการวิจัยเชิงปฏิบัติการ (Operations Research Methods)

1. การใช้คะแนนของคนอื่น ๆ

การกำหนดจุดตัดวิธีนี้ จะอาศัยหลักการกำหนดคะแนนจุดตัดที่สอดคล้องกับคะแนนเปอร์เซ็นต์ของผู้สอบผ่านเกณฑ์ซึ่งระบุไว้ล่วงหน้าโดยอาศัยผู้เชี่ยวชาญเนื้อหาวิชา แล้วให้ระบุเปอร์เซ็นต์ของผู้สอบผ่านว่าควรเป็นเท่าไรก่อน แล้วหาคะแนนที่สอดคล้องกับเปอร์เซ็นต์นั้น

2. การใช้วิธีนับถอยหลังจาก 100%

วิธีการนี้จะอาศัยหลักการเดียวกับการกำหนดค่าเกณฑ์หรือระดับมาตรฐานของจุดประสงค์เชิงพฤติกรรม คือ ค่าเกณฑ์ที่ต้องกำหนดตามความสำคัญของจุดประสงค์ ถ้าจุดประสงค์ใดมีความสำคัญมากกว่า เกณฑ์ที่ต้องการ ต้องเป็น 100% ถ้าจุดประสงค์ใดมีความสำคัญน้อยลงมา ค่าเกณฑ์ที่ต้องการก็ลดต่ำลงมาจาก 100% ฉะนั้นเกณฑ์ที่ตั้งไว้ 100% อาจลดลงมาเป็น 95%, 90% หรือ 80%

3. การปรับคะแนนเกณฑ์อื่น ๆ

เป็นการกำหนดคะแนนจุดตัด โดยอาศัยเกณฑ์ภายนอกเป็นตัวเปรียบเทียบ เกณฑ์ภายนอกนี้จะต้องเป็นที่ยอมรับทั่วไป และประจักษ์ที่จะชี้บอก “ความรอบรู้” หรือ “ความสำเร็จ”

4. การตัดสินใจโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบ

เป็นวิธีตัดสินที่ขึ้นอยู่กับข้อคำถามในแบบทดสอบวิธีนี้ แกลส (Glass, 1976 : 247-251) เรียกว่า เป็นวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบ เป็นวิธีการที่อาศัยแนวคิดเกี่ยวกับผู้สอบที่มีความรู้ และทักษะอยู่ในระดับคาบเส้น (Borderline) ระหว่างกลุ่มที่ได้คะแนนสูงกับกลุ่มที่ได้คะแนนต่ำ คะแนนของผู้สอบที่มีความรู้และทักษะอยู่ในระดับคาบเส้นก็จะเป็นคะแนนจุดตัดการตัดสินใจโดยอาศัยวิธีการนี้ จะต้องพิจารณาจากข้อคำถามในแบบทดสอบวิธีการนี้สามารถกำหนดคะแนนจุดตัดได้ ทั้งก่อนและหลังจากการนำแบบทดสอบไปทำการดำเนินการสอบแล้ว หากกล่าวโดยสรุปจะเห็นว่า กระบวนการวิธีนี้จะมุ่งพิจารณาที่เนื้อหาของแบบทดสอบเป็นสำคัญ เพราะฉะนั้นผู้ตัดสินหรือผู้เชี่ยวชาญก็จะต้องพิจารณาว่า ผู้สอบที่อยู่ในระดับคาบเส้นจะสามารถตอบข้อคำถามแต่ละข้อในแบบทดสอบได้อย่างไร

การคำนวณคะแนนจุดตัด โดยวิธีการนี้มีขั้นตอนพื้นฐานดังนี้

1. เลือกผู้ตัดสิน หรือผู้เชี่ยวชาญ
2. กำหนดความรู้ และทักษะของผู้สอบที่อยู่ในระดับคาบเส้น

3. ฝึกฝนวิธีการตัดสินในแต่ละวิธีการที่เลือกใช้

4. เก็บรวบรวมผลการตัดสิน

5. นำผลการตัดสินมารวมกัน เพื่อกำหนดคะแนนจุดตัด

สำหรับวิธีการนี้มีผู้ศึกษาและได้เสนอเทคนิคในการหาคะแนนความสามารถต่ำสุดไว้หลายคนด้วยกัน เช่น

1. **วิธีของนีเดลสกี** (Nedelsky, 1954 quoted in Glass, 1978 : 246-247) วิธีนี้เป็นการหาคะแนนผ่านต่ำสุด (Munimum Passing Scores) ของแบบทดสอบปรนัยชนิดเลือกตอบที่ เรียกว่า เทคนิคการหาคะแนนการเดาระหว่างนักเรียนเกรด F และ D (F-D Guess Score Techniques) ซึ่งจะแบ่งนักเรียนออกเป็น 2 พวก คือ พวกที่อยู่เหนือจุดแบ่งเป็นพวกที่สอบผ่านบริเวณเส้นแบ่ง (Borderline) คือ นักเรียนที่อยู่ระหว่างเกรด F และ D (F-D Students) เขาได้บัญญัติศัพท์ขึ้นใช้ในการคำนวณดังนี้

- ตัวลวงที่นักเรียนระดับต่ำสุดของพวกที่ยอมให้สอบผ่าน (D-Students) จะไม่เลือกเพราะรู้ว่าผิดและดึงดูดความสนใจของพวกที่สอบไม่ผ่าน (F-Students) เรียกว่า “คำตอบที่ผิด” (F-Response)

- นักเรียนที่มีความรู้เพียงพอที่จะไม่เลือกคำตอบที่ผิด แต่เลือกตัวลวงที่เหลือด้วยการสุ่ม เรียกว่า “นักเรียนที่อยู่ระหว่างเกรด F และ D” (F-D Students) ซึ่งคือผู้ที่มีความรู้อยู่ตรงเส้นแบ่งระหว่างเกรด F และ D

- คะแนนเฉลี่ยของนักเรียนที่อยู่ระหว่างเกรด F และ D ทั้งหมดเรียกว่า “คะแนนการเดาระหว่างนักเรียนเกรด F และ D” (F-D Guess Score) เขียนแทนด้วย M_{FD} ซึ่งจะมีค่าเท่ากับผลรวมของส่วนกลับของจำนวนตัวเลือกที่เหลือจากการคัดตัวเลือกที่มั่นใจว่าผิดออก

- ค่าส่วนเบี่ยงเบนมาตรฐานของความน่าจะเป็นในการเดาตอบถูก เขียนแทนด้วย S_{FD}

วิธีการดังกล่าวนี้ ก่อนนำแบบทดสอบไปให้นักเรียนทำ ต้องนำแบบทดสอบไปให้ครูผู้สอนวิชานั้น ๆ พิจารณาข้อสอบแต่ละข้อว่า ตัวเลือกใดที่นักเรียนระดับต่ำสุดที่เรายอมให้ผ่านบอกได้ว่าผิด แล้วเขียนเศษส่วนกลับของจำนวนตัวเลือกที่เหลืออยู่ไว้ทางซ้ายมือของข้อสอบ เช่น ถ้าครูเลือกตัวลวงที่นักเรียนระดับต่ำสุดที่เรายอมให้ผ่านบอกได้ว่าผิดได้ 1 ข้อ จาก 5 ข้อ ก็ให้

เขียนเป็น $\frac{1}{4}$ เป็นต้น แล้วนำเศษส่วนเหล่านี้ของข้อสอบทุกข้อมารวมกันได้เป็น M_{FD} เพื่อนำไปคำนวณหาค่าคะแนนผ่านต่ำสุดจากสูตร

$$\text{คะแนนผ่านต่ำสุด (C)} = M_{FD} + kS_{FD}$$

เมื่อ k เป็นค่าคงที่ซึ่งกำหนดขึ้นจากการพิจารณาหลาย ๆ ครั้ง โดยครั้งแรกครูผู้สอนหลาย ๆ คนจะมาตกลงกันเพื่อกำหนดค่า k ขึ้นทดลองใช้ก่อน ซึ่งเขากล่าวไว้ว่า คะแนนผ่านต่ำสุดจะมีมาตรฐานพอเมื่อนักเรียนที่อยู่ระหว่างเกรด F และ D ส่วนใหญ่สอบตกและจากการศึกษา เขาพบว่า ถ้าให้ k เป็น -1, 0, 1, 2 จะทำให้พวกที่อยู่ระหว่างเกรด F และ D สอบตก 16%, 50%, 84%, และ 98% ตามลำดับ การพิจารณาเลือกค่า k ครั้งสุดท้ายจะทำหลังจากที่ให้ผู้สอนหลาย ๆ คน เลือกคำตอบที่นักเรียนระดับต่ำสุดที่ผ่านรู้ว่าผิดแล้ว แต่การกำหนดค่ามาตรฐานสัมบูรณ์ (Absolute Standard) นั้น ควรกำหนดค่า k ขึ้นมาก่อนจะคำนวณค่า M_{FD} และก่อนที่จะรู้คะแนนการสอบของนักเรียน

คำตอบที่นักเรียนระดับต่ำสุดที่ผ่านรู้ว่าผิด จะมีส่วนทำให้การกำหนดค่า k เปลี่ยนแปลงไปได้ เนื่องจากการสอบแต่ละครั้ง คำตอบที่ผิดจะมีทั้งคำตอบที่ผิดอย่างชัดเจนและผิดปานกลาง ซึ่งผู้ที่จะไม่เลือกจะต้องมีความรู้ถึงขั้นนั้นแล้ว ในการพิจารณาเปลี่ยนแปลงค่า k นั้น ไม่ควรเปลี่ยนมากไปกว่า 0.5 จากค่า k ที่กำหนดขึ้นเพื่อทดลองใช้ การตัดสินทางทฤษฎีมักไม่ปรากฏว่าค่า k จะสูงถึง 2 และสำหรับแบบทดสอบถ้ากำหนดให้ค่า k เท่ากับ 0 อาจจะต่ำเกินไป จึงมีข้อเสนอแนะว่าควรเริ่มใช้ที่ k เท่ากับ 1 ก่อน ดังนั้นช่วงของค่า k จะเปลี่ยนไประหว่าง 0.5 ถึง 1.5 ซึ่งการกำหนดค่า k นี้ บุญเชิด ภิญโญอนันตพงษ์ (2527 : 124) มีข้อเสนอแนะว่า ถ้าตัวเลือกผิดเด่นชัดจะใช้ค่า 0.5 ถ้าตัวเลือกแตกต่างกันเล็กน้อยจะใช้ค่า 1.0 ถ้าตัวเลือกคล้ายคลึงกันมากจะใช้ค่า 1.5

2. **วิธีของอีเบล** (Ebel, 1972 quoted in Glass, 1978 : 247-248) อีเบลมีความเห็นว่าคะแนนผ่าน (Passing Scores) ที่กำหนดเป็นเปอร์เซ็นต์ของคะแนนสอบทั้งหมดอาจมีข้อบกพร่องเนื่องจากข้อสอบนั้นง่ายเกินไปหรือยากเกินไป หรืออำนาจจำแนกน้อยกว่าที่ผู้สร้างตั้งใจไว้ การที่จะสอบผ่านหรือไม่อาจเนื่องมาจากตัวข้อสอบมากกว่าระดับความสามารถที่แท้จริงของผู้สอบก็ได้ เขาจึงเสนอวิธีแก้ไขโดยการให้วิเคราะห์เนื้อหาของข้อสอบในแง่ของความเกี่ยวข้อง (Relevance) และความยาก (Difficulty) ของข้อสอบแต่ละข้อ เทียบกับเปอร์เซ็นต์

ความคาดหวัง ซึ่งเป็นปริมาณที่คาดว่าผู้ที่มีความสามารถต่ำสุดที่จะสอบผ่านควรทำได้ ซึ่งอีเบลได้จำแนกไว้แล้ว ดังตาราง

ตาราง 1 ค่าเปอร์เซ็นต์ความคาดหวังในแต่ละระดับความยากของแต่ละประเภทของความเกี่ยวข้อง

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

วิธีการหาคะแนนจุดตัด มีลำดับขั้นดังนี้

1) นำข้อสอบที่ต้องการวิเคราะห์ไปให้ครูประจำวิชาพิจารณาว่า ข้อสอบมีลักษณะตรงกับปัญหามากน้อยเพียงใด มีระดับความยากง่ายเพียงใด

2) จำแนกข้อสอบตามลักษณะและความยาก โดยแจกแจงเป็นความถี่รวม

3) กำหนดเปอร์เซ็นต์การสอบผ่านตามลักษณะข้อสอบและความยากของข้อสอบ

4) คูณความถี่จำนวนข้อกับเปอร์เซ็นต์การสอบผ่านที่คาดหวังและรวมผลคูณทั้งหมด

5) หาค่าเปอร์เซ็นต์การสอบผ่านที่คาดหวังโดยเฉลี่ย

6) กำหนดคะแนนจุดตัดจากค่าเปอร์เซ็นต์การสอบผ่านที่คาดหวังโดยเฉลี่ยจากขั้นที่ 5

วิธีการหาคะแนนจุดตัดของอีเบล ก็คือ นำจำนวนข้อสอบในแต่ละลักษณะคูณกับค่าร้อยละที่คาดหวัง แล้วนำผลคูณที่ได้มาหาค่าเฉลี่ย ค่าเฉลี่ยที่ได้ก็จะเป็นคะแนนจุดตัด หรือคะแนนที่ผู้มีความสามารถต่ำสุดจะสอบผ่าน ดังตัวอย่างในตาราง 2

ตาราง 2 ตัวอย่างของการคำนวณตามเทคนิคของอีเบล

ลักษณะของคำถาม	จำนวนของข้อคำถาม	ร้อยละที่คาดหวัง	จำนวนข้อคำถามคูณร้อยละที่คาดหวัง
มีความจำเป็น	94	100	9,400
มีความสำคัญ			
ง่าย	106	90	9,540
ปานกลาง	153	70	10,710
เป็นที่ยอมรับได้			
ง่าย	24	80	1,920
ปานกลาง	49	60	2,940
ยาก	52	40	2,080
ไม่แน่ใจ			
ง่าย	4	70	280
ปานกลาง	11	50	550
ยาก	7	30	210
รวม	500		37,130

จากตาราง 2 จะได้คะแนนจุดตัดเท่ากับ $\frac{37,130}{500} = 74.26$ หรือ 74 %

3. **วิธีของแองกอฟฟ์** (Angoff, 1971 : 514 : 515) วิธีนี้จะอาศัยหลักการของความน่าจะเป็นที่นักเรียนซึ่งมีสมรรถภาพขั้นต่ำสุดที่จะยอมรับได้ตอบข้อสอบถูก โดยนำไปให้ผู้เชี่ยวชาญเนื้อหาวิชาหรือครูผู้สอนพิจารณา ค่าความน่าจะเป็นในการตอบข้อสอบถูกจะนำมาใช้เป็นคะแนนจุดตัด วิธีการหาคะแนนจุดตัดมีลำดับขั้นดังนี้

1) นำข้อสอบทั้งหมดไปให้ครูผู้สอนกลุ่มหนึ่ง พิจารณาเนื้อหาข้อสอบและความยาก

2) ให้ครูผู้สอนพิจารณาต่อไปว่า นักเรียนที่มีความสามารถขั้นต่ำสุดตามเนื้อหาข้อสอบจะมีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกเป็นเท่าไร

3) นำค่าความน่าจะเป็นในการตอบข้อสอบถูกที่ครูผู้สอนแต่ละคนพิจารณาไว้มาหาค่าเปอร์เซ็นต์เฉลี่ยของความน่าจะเป็น

4) กำหนดคะแนนจุดตัดจากค่าเปอร์เซ็นต์เฉลี่ยในขั้นที่ 3

ส่วนในการคำนวณหาคะแนนจุดตัดสามารถทำได้ 3 วิธี (Zieky, 1987 : 5-6) คือ วิธีที่ 1 ใช้คะแนนเฉลี่ย (Mean) วิธีที่ 2 ใช้ค่ามัธยฐาน (Median) วิธีที่ 3 ใช้ทั้งค่าเฉลี่ยและค่ามัธยฐานร่วมกัน โดยตัดค่าสูงสุดและค่าต่ำสุดออก วิธีการดังกล่าวนี้เรียกว่า Trimmed Mean ดังตัวอย่างในตาราง 3 สมมติว่า มีข้อคำถามอยู่ 10 ข้อ และให้ผู้เชี่ยวชาญ 5 คน ตัดสินได้ผลดังนี้

ตาราง 3 ตัวอย่างผลการตัดสินโดยใช้เทคนิคแองกอฟฟ์

ข้อคำถาม	ผู้ตัดสิน				
	A	B	C	D	E
1	.95	.90	1.00	.85	.90
2	.80	.80	1.00	.80	.85
3	.90	.85	1.00	.80	.85
4	.60	.65	.95	.70	.75
5	.75	.70	.90	.75	.75
6	.40	.60	.90	.65	.65
7	.50	.65	.90	.60	.70
8	.25	.30	.85	.45	.55
9	.25	.25	.80	.30	.50
10	.40	.30	.75	.25	.45
รวม	5.80	6.00	9.05	6.15	6.90

ค่าเฉลี่ยเท่ากับ 6.78

ค่ามัธยฐานเท่ากับ 6.15

Trimmed Mean เท่ากับ 6.35 (โดยตัดค่าสูงสุดและค่าต่ำสุดออก)

เทคนิคของแองกอฟฟ์เป็นเทคนิคที่ง่ายและใช้ได้รวดเร็ว (Livingston and Zieky, 1982 : 54) หากกลุ่มผู้เชี่ยวชาญพิจารณากำหนดคะแนนจุดตัดที่มีความเชี่ยวชาญในเนื้อหาที่จะนำมาในการเขียนข้อสอบ (Norcini, Shea and Kanya, 1988 : 57) เนื่องจากเทคนิคนี้ผู้เชี่ยวชาญจะต้องพิจารณาถึงโอกาสของผู้สอบที่มีความสามารถต่ำสุดหรือผู้ที่อยู่ในระดับคาบเส้นที่จะตอบข้อสอบในแต่ละข้อคำถามได้ถูกต้อง เพราะฉะนั้นก่อนอื่นผู้เชี่ยวชาญจะต้องให้คำนิยามหรือกำหนดความสามารถต่ำสุดของผู้สอบก่อน โดยผู้เชี่ยวชาญแต่ละคนจะต้องคาดคะเนความยากในแต่ละข้อคำถามของแบบทดสอบแล้วนำผลที่ได้จากการคาดคะเนของผู้เชี่ยวชาญมาหาค่าเฉลี่ย ค่าที่ได้ก็จะเป็นคะแนนจุดตัด (Mills, 1983 : 284)

ส่วนเทคนิคของอีเบล เป็นเทคนิคที่ผู้เชี่ยวชาญจะต้องพิจารณาถึงความยาก และความสำคัญของข้อคำถามประกอบกัน ซึ่งเทคนิคนี้จะทำให้ผู้เชี่ยวชาญต้องใช้เวลาในการพิจารณามากและคงจะไม่เหมาะสมกับแบบทดสอบแบบตอบสั้น (Zieky, 1987 : 16) นอกจากนี้แบบทดสอบของอีเบลไม่เหมาะกับแบบทดสอบที่มีความยาวมาก ๆ เนื่องจากผู้เชี่ยวชาญจะต้องเสียเวลาในการจัดประเภทของแบบทดสอบไปตามค่าความยาก และความสำคัญ (Cross, Impara, Frary and Jacger, 1984 : 114)

สำหรับเทคนิคของนีเดลสกี เป็นเทคนิคที่เหมาะกับแบบทดสอบเลือกตอบเท่านั้น (Zieky, 1987 : 11) การกำหนดคะแนนจุดตัดโดยใช้เทคนิคนี้จะต้องพิจารณาถึงความยากของข้อคำถามเป็นรายตัวเลือกว่า ตัวเลือกใดที่ผู้สอบที่มีความสามารถต่ำสุด หรือผู้สอบในระดับคาบเส้นสามารถทราบได้ว่าเป็นตัวเลือกที่ผิด (Plake and Milican, 1989 : 46 ; Cross, Impara, Frary and Jacger, 1984 : 114) ซึ่งความเป็นจริงแล้ว ผู้เชี่ยวชาญก็จะไม่มีโอกาสรู้ได้ (Zieky, 1987 : 16)

5. **การใช้ทฤษฎีการตัดสินใจ** (Decision-Theoretic Approaches)

การกำหนดจุดตัดวิธีนี้อาศัยผลการสอบมาใช้ประกอบการพิจารณาตัดสินคะแนนจุดตัดที่เหมาะสมซึ่งมีอยู่ด้วยกันหลายวิธี บางวิธีจะใช้การนิยามความรอบรู้ด้วยคะแนนสอบหรือคะแนนดิบ เช่น วิธีของแกลส (Glass) และวิธีของเบอร์ก (Berk) (บุญเชิด ภิญโญอนันตพงษ์, 2526 : 184)

การกำหนดคะแนนจุดตัดตามวิธีของแกลส (Glass, 1978 : 243 - 257) เป็นวิธีที่อาศัยความสัมพันธ์ของคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์กับเกณฑ์ภายนอกที่

สอดคล้องซึ่งกำหนดไว้ก่อน โดยการสร้างฟังก์ชันของคะแนนจุดตัด ซึ่งได้จากการใช้เกณฑ์ภายนอกจำแนกคนเป็นสองกลุ่ม เช่น ใช้เกณฑ์แบ่งคนเป็นกลุ่มที่จบการศึกษากับกลุ่มที่ไม่จบการศึกษา ผู้ที่ได้รับจ้าง และไม่ได้รับจ้าง เป็นต้น สัดส่วนของของคนทั้งสองกลุ่ม แทนด้วย P_B และ $1\text{-}P_B$ ตามลำดับ แล้วให้คนทั้งสองกลุ่มนี้ทำแบบทดสอบอิงเกณฑ์และกำหนดคะแนนจุดตัด (C_X) ขึ้นมาใช้แบ่งคนออกเป็นผู้สอบผ่านและไม่ผ่านเกณฑ์ จะทำให้เกิดการจำแนกคนออกเป็น 4 กลุ่ม ดังตาราง

		เกณฑ์ภายนอก	
		ผ่าน	ไม่ผ่าน
เกณฑ์ของแบบทดสอบ	ไม่ผ่าน	P_A	P_B
	ผ่าน	P_C	P_D

เมื่อ P_A แทน สัดส่วนของจำนวนคนที่ไม่ผ่านเกณฑ์ของแบบทดสอบ แต่ผ่านเกณฑ์ภายนอก เป็นการจำแนกผิดทางลบ

P_B แทน สัดส่วนของจำนวนคนที่ไม่ผ่านเกณฑ์ของแบบทดสอบ และไม่ผ่านเกณฑ์ภายนอก เป็นการจำแนกถูก

P_C แทน สัดส่วนของจำนวนคนที่ผ่านเกณฑ์ของแบบทดสอบและผ่านเกณฑ์ภายนอก เป็นการจำแนกถูก

P_D แทน สัดส่วนของจำนวนคนที่ผ่านเกณฑ์ของแบบทดสอบ แต่ไม่ผ่านเกณฑ์ภายนอก เป็นการจำแนกผิดทางบวก

การกำหนดเกณฑ์ภายนอกนั้นจะไม่ทำให้เปลี่ยนแปลง แต่คะแนนจุดตัดของแบบทดสอบอิงเกณฑ์จะแปรผันไปได้หลายค่า แล้วแต่การกำหนดซึ่งจะทำให้สัดส่วนของ P_A, P_B, P_Cและ P_D แปรผันตามไปด้วย คะแนนจุดตัดของแบบทดสอบแบบอิงเกณฑ์ก็คือ ค่าฟังก์ชั่นของคะแนนเกณฑ์ C_X ที่ทำให้ $f(C_X)$ มีค่าน้อยที่สุด

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

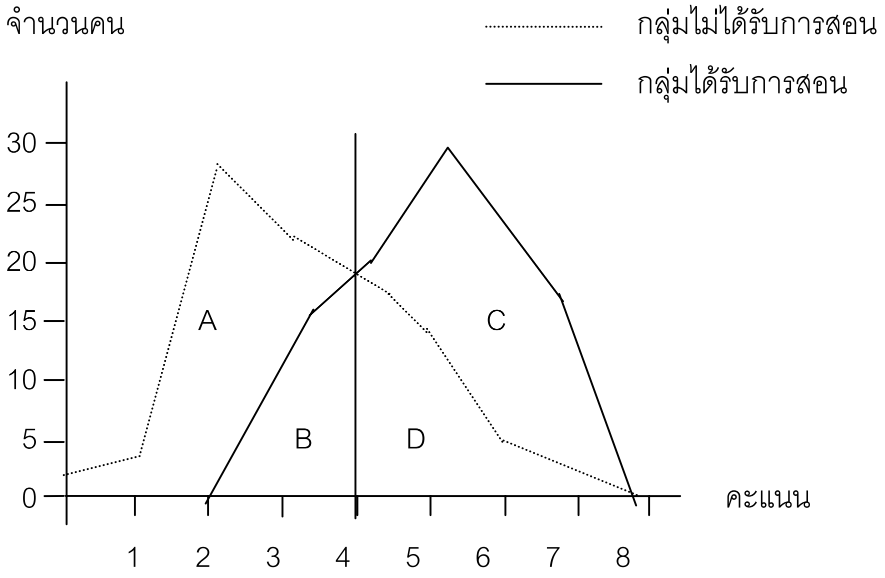
ค่าของ C_X ที่ทำให้ $f(C_X)$ มีค่าน้อยที่สุดจะถูกเลือกเป็นคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ การใช้ฟังก์ชันมีข้อตกลงว่า การจำแนกผิดทางบวก (P_D) และการจำแนกผิดทางลบ (P_A) มีความสำคัญเท่ากัน ถ้าให้ความสำคัญของการจำแนกผิดทั้งสองแบบไม่เท่ากันจะทำให้ค่าของฟังก์ชันแปรเปลี่ยนไปมาก ดังนั้นจึงต้องมีการปรับแก้โดยกำหนดค่าถ่วงขึ้นมาสองค่า α เป็นค่าความสำคัญของการจำแนกผิดทางลบ และ β เป็นค่าความสำคัญของการจำแนกผิดทางบวก จะได้ฟังก์ชั่นต่ำสุดของคะแนนเกณฑ์ดังนี้

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

การกำหนดคะแนนจุดตัดตามวิธีของเบอร์ก (Berk, 1976 : 4-9) เบอร์กได้หาคะแนนจุดตัด โดยประยุกต์มาจากวิธีการเพิ่มคะแนนเกณฑ์อื่น ๆ ซึ่งเบอร์กกล่าวว่า การกำหนดเกณฑ์คือการกำหนดจุดตัดของคะแนนที่แบ่งผู้เรียนออกเป็นสองพวก คือ พวกที่ได้รับการสอนให้เป็นพวกที่รอบรู้ (Master) พวกที่ไม่ได้รับการสอนเป็นพวกไม่รอบรู้ (Non-Master) หลังจากให้นักเรียนทั้งสองกลุ่มทำแบบทดสอบแล้ว พิจารณาการกระจายของคะแนนสองกลุ่มจะคาบเกี่ยวกัน จุดที่ฟังก์ชันทั้งสองตัดกัน คือ คะแนนพยากรณ์ที่จะแบ่งการเรียนรู้เป็นสี่พวกดังนี้

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

คะแนนจากการทดสอบของกลุ่มได้รับการสอนกับกลุ่มไม่ได้รับการสอน นำมาแจกแจงเป็นโค้งจะได้ลักษณะดังภาพข้างล่าง ถ้าแบบทดสอบมีอยู่ 8 ข้อ สอบกับนักเรียน 100 คน



A แทน กลุ่มไม่รอบรู้จริง

B แทน กลุ่มไม่รอบรู้ไม่จริง

C แทน กลุ่มรอบรู้จริง

D แทน กลุ่มรอบรู้ไม่จริง

ภาพประกอบ 1 เส้นโค้งการกระจายของคะแนนระหว่างกลุ่มที่ได้รับการสอนกับกลุ่มไม่ได้รับการสอน

คะแนนจุดตัดนี้เป็นคะแนนพยากรณ์ นำมาหาค่าคะแนนเกณฑ์ที่ให้ค่าความน่าจะเป็นในการตัดสินอย่างถูกต้อง คือ ค่า P(TM)+P(TN) สูงสุด หรือให้ค่าความน่าจะเป็นในการตัดสินผิด คือ ค่า P(FM) + P(FN) ต่ำสุด ณ จุดคะแนนนั้นจะเป็นคะแนนจุดตัดที่เหมาะสมที่สุด ในการหาจะเลื่อนค่าคะแนนพยากรณ์ไปเรื่อย ๆ จุดคะแนนหนึ่งที่มีค่า P(TM) + P(TN) สูงสุด และค่า P(FM) + P(FN) ต่ำสุดเป็นคะแนนจุดตัดที่เหมาะสม

เมื่อ $P(TM) = \frac{TM}{M+N}$

$P(TN) = \frac{TN}{M+N}$

$$P(FM) = \frac{FM}{M+N}$$

$$P(FN) = \frac{FN}{M+N}$$

N แทน จำนวนนักเรียนทั้งหมดที่ยังไม่ได้เรียน

M แทน จำนวนนักเรียนทั้งหมดที่เรียนแล้ว

คะแนนจุดตัดแต่ละคะแนนที่หาออกมาได้สามารถตรวจสอบความเที่ยงตรงโดยใช้สัมประสิทธิ์ความเที่ยงตรงของเกณฑ์ เพื่อเลือกค่าสัมประสิทธิ์ที่สูงที่สุดของความน่าจะเป็นในการตัดสินถูกของแต่ละคะแนนจุดตัดมาเป็นคะแนนเกณฑ์ สูตรการหาความเที่ยงตรงของเกณฑ์ได้จากสูตรดังนี้

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

เมื่อ ϕ_{vc} แทน ค่าสัมประสิทธิ์ความเที่ยงตรงของคะแนนเกณฑ์

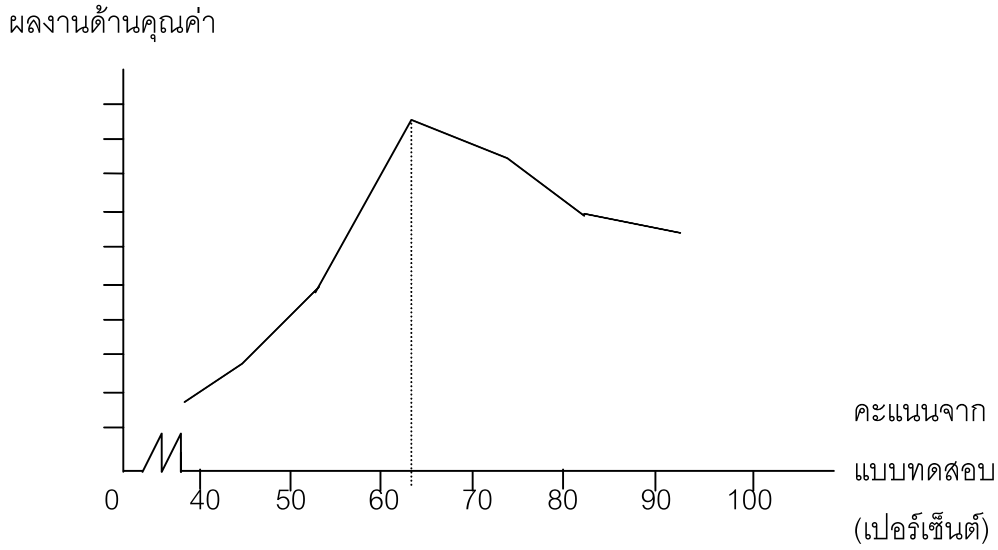
BR แทน ค่าความน่าจะเป็นของผู้รอบรู้ในประชากร = P(FN)+P(TM)

SR แทน ค่าความน่าจะเป็นของการพยากรณ์ผู้รอบรู้ในประชากร = P(FM)+P(TM)

6. **การใช้วิธีการวิจัยเชิงปฏิบัติการ** (Operations Research Methods) วิธีนี้จะต้องอาศัยผลด้านคุณค่า (Valued Outcome) อย่างใดอย่างหนึ่งมาช่วยในการพิจารณาเกณฑ์ โดยวัดผลด้านคุณค่าของผู้ที่ได้คะแนนต่าง ๆ กันในการทดสอบด้วยแบบทดสอบอิงเกณฑ์ แล้วใช้คะแนนของผู้ที่มีผลด้านคุณค่านั้นสูงสุดมาเป็นคะแนนจุดตัด โดยพิจารณาจากลักษณะกราฟแสดงความสัมพันธ์ระหว่าง ผลที่ได้จากการสอบด้วยแบบทดสอบอิงเกณฑ์และจากการวัดผลด้านคุณค่านั้น คะแนนตรงจุดสูงสุดของกราฟจะถือเป็นคะแนนจุดตัดซึ่ง บลอค (Block, 1972 : 183-190)

บลอคกำหนดคะแนนจุดตัดของแบบทดสอบให้มีระดับต่าง ๆ กัน เช่น 40%, 45%, 60%, …, 90%, 100 จากนั้นทำการสอนกลุ่มต่าง ๆ ที่มีความสามารถเท่าเทียมกัน จนกระทั่งนักเรียนบรรลุตามเกณฑ์รอบรู้ที่กำหนดโดยทำการสอบกับแบบทดสอบอิงเกณฑ์ แล้ววัดผลด้านคุณค่า

อย่างใดอย่างหนึ่งที่สัมพันธ์กับคะแนนจากแบบทดสอบอิงเกณฑ์ เช่น ความคงทนของการเรียนรู้ ความสำเร็จในการทำงาน ความสนใจหรือทัศนคติต่อวิชานั้น เป็นต้น นำคะแนนที่ได้จากการทดสอบและจากการวัดคุณค่ามาเขียนกราฟแสดงความสัมพันธ์เพื่อพิจารณาค่าของคะแนนจุดตัด (C_X) โดยพิจารณาจากการแจกแจงคะแนนจุดตัดที่สมนัยกับจุดโค้งกลับของกราฟ ดังภาพ



ภาพประกอบ 2 กราฟแสดงความสัมพันธ์ระหว่างคะแนนด้านคุณค่ากับคะแนนจากแบบทดสอบ

จุดโค้งกลับหรือส่วนโค้งที่สูงที่สุดของกราฟตรงกับคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ จากภาพคะแนนจุดตัดที่เหมาะสม คือ 70%

<u>ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์</u>

กลุ่มนักวัดผลหลายท่าน เช่น แฮมเบิลตัน และคณะ (Hambleton and others, 1980 quoted Berk, 1980 : 329) ได้แบ่งกลุ่มและนิยามความหมายของความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ไว้ 3 กลุ่ม ดังนี้คือ

1. ความเชื่อมั่นในการตัดสินจำแนกความรอบรู้ (Reliability of Master Classification Decision) เป็นการจำแนกความสอดคล้องในการตัดสินความรอบรู้ ไม่รอบรู้จากผลการสอบซ้ำด้วยแบบทดสอบฉบับเดียวหรือแบบทดสอบคู่ขนาน ดัชนีที่วิเคราะห์เป็นดัชนีฟังก์ชั่นการสูญเสียเนื่องจากการกำหนดคะแนนจุดตัดล่วงหน้า (Threshold Loss Function Indices)

2. ความเชื่อมั่นของคะแนนของแบบทดสอบอิงเกณฑ์ (Reliability of Criterion Referenced Test Score) เป็นความสอดคล้องของคะแนนแต่ละคนที่เบี่ยงเบนจากคะแนนจุดตัดด้วยแบบทดสอบคู่ขนานหรือคู่ขนานแบบสุ่ม (Parallel or Randomly Parallel Test Forms) ดัชนีที่วิเคราะห์เป็นดัชนีฟังก์ชั่นการสูญเสียเนื่องจากความคลาดเคลื่อนในการยกกำลังสอง (Squared Error Loss Function Indices)

3. ความเชื่อมั่นในการกะประมาณคะแนนมวลความรู้ (Reliability of Domain Score Estimates) เป็นการหาความสอดคล้องของคะแนนสอบแต่ละคนจากแบบทดสอบ คู่ขนานสองฉบับ สถิติที่ใช้เป็นการประมาณค่าคะแนนมวลความรู้หรือสัดส่วนของการตอบถูกในประชากรข้อสอบ แต่ไม่เกี่ยวข้องกับมาตรฐานของความรอบรู้

จากความหมายดังกล่าว จะเห็นได้ว่าคะแนนจุดตัดเป็นตัวประกอบที่สำคัญในการหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ โดยเฉพาะความเชื่อมั่นในการตัดสินจำแนกความรอบรู้และความเชื่อมั่นของคะแนนของแบบทดสอบอิงเกณฑ์

<u>วิธีหาค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์</u>

วิธีการหาความเชื่อมั่นของแบบทดสอบมีหลายวิธี ได้แก่

1. การหาความเชื่อมั่นในการตัดสินจำแนกความรอบรู้ (Reliability of Mastery Classification Decision) เป็นการหาดัชนีความสอดคล้องในการตัดสินใจจำแนกรอบรู้ ไม่รอบรู้โดยการสอบซ้ำด้วยแบบทดสอบฉบับเดียวหรือแบบทดสอบคู่ขนาน ซึ่งมีผู้เสนอไว้หลายวิธีดังนี้

1.1 วิธีของแฮมเบิลตันและโนวิค

วิธีของแฮมเบิลตันและโนวิค (Hamble and Novick, 1973 : 159-170) ได้เสนอสูตรในการประมาณค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์โดยใช้ผลรวมของสัดส่วนความสอดคล้องในการตัดสินจำแนกผู้รอบรู้จากการสอบด้วยแบบทดสอบคู่ขนานหรือแบบทดสอบฉบับเดียวกัน 2 ครั้ง สัดส่วนของความสอดคล้องในการตัดสินก็คือ สัดส่วนของการตัดสินว่ารอบรู้ตรงกันทั้งสองฉบับหรือสองครั้งและสัดส่วนของการตัดสินว่าไม่รอบรู้ตรงกันทั้งสองฉบับหรือสองครั้ง เขียนเป็นสูตรได้ดังนี้

$$P_0 = P_{11}+P_{22}$$

เมื่อ	P_0	แทน	สัดส่วนของความสอดคล้องในการตัดสินจำแนกผู้รอบรู้
	P_{11}	แทน	สัดส่วนของผู้ถูกตัดสินว่ารอบรู้ตรงกันทั้งสองฉบับหรือสองครั้ง
	P_{22}	แทน	สัดส่วนของผู้ถูกตัดสินว่าไม่รอบรู้ตรงกันทั้งสองฉบับหรือสองครั้ง

1.2 วิธีของสวามินาธาน แฮมเบิลตัน และอัลจินา

วิธีของสวามินาธาน แฮมเบิลตัน และอัลจินา (Swaminathan, Hambleton and Algina. 1974 : 263-267) เสนอสูตรการหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์จากผลการสอบสองครั้ง โดยอาศัยแนวความคิดของแฮมเบิลตันและโนวิคที่ว่า ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์สะท้อนถึงระดับที่ผู้เรียนถูกกำหนดความคงเส้นคงวาในสภาวะการเรียนรู้เดียวกันจากผลการสอบสองครั้งของแบบทดสอบคู่ขนาน ซึ่งวัดได้โดยดัชนีของความสอดคล้องในการจำแนกความรอบรู้-ไม่รอบรู้ ระหว่างการสอบสองครั้ง ที่มีการปรับแก้ความสอดคล้อง โดยบังเอิญและอาศัยสัมประสิทธิ์แคปป้า (K) ของโคเฮน (Cohen. 1960) ซึ่งมีสูตรดังนี้

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

เมื่อ	K	แทน	ค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ซึ่งได้หักค่าความสอดคล้องโดยบังเอิญออกแล้ว
	P_O	แทน	สัดส่วนของความสอดคล้องในการตัดสินใจจำแนกผู้รอบรู้ นั่นคือ $P_O = P_{11}+P_{22}$
	P_C	แทน	สัดส่วนของความสอดคล้องที่คาดหวังซึ่งอาจเกิดขึ้นโดยบังเอิญ นั่นคือ $P_C = (P_{.1})(P1.)+(P_{.2})(P_{2.})$

เมื่อ

		สอบครั้งที่ 2 (ฉบับ ข)		
		รอบรู้	ไม่รอบรู้	สัดส่วนแยก
สอบครั้งที่ 1 (ฉบับ ก)	รอบรู้	P_{11}	P_{12}	$P_{1.}$
	ไม่รอบรู้	P_{21}	P_{22}	$P_{2.}$
	สัดส่วนแยก	$P_{.1}$	$P_{.2}$	

P_{11} แทน สัดส่วนของผู้สอบที่ถูกตัดสินว่ารอบรู้ตรงกันทั้ง 2 ครั้ง หรือ 2 ฉบับ

P_{12} แทน สัดส่วนของผู้สอบที่ถูกตัดสินว่ารอบรู้ในการสอบครั้งที่ 1 (ฉบับ ก) แต่ไม่รอบรู้ในการสอบครั้งที่ 2 (ฉบับ ข)

P_{21} แทน สัดส่วนของผู้สอบที่ถูกตัดสินว่ารอบรู้ในการสอบครั้งที่ 2 (ฉบับ ข) แต่ไม่รอบรู้ในการสอบครั้งที่ 1 (ฉบับ ก)

P_{22} แทน สัดส่วนของผู้สอบที่ถูกตัดสินว่าไม่รอบรู้ตรงกันทั้ง 2 ครั้งหรือ 2 ฉบับ

$P_{1.}$ แทน สัดส่วนแยกของผู้ที่รอบรู้ในการสอบครั้งที่ 1 (ฉบับ ก)

$P_{2.}$ แทน สัดส่วนแยกของผู้ที่ไม่รอบรู้ในการสอบครั้งที่ 1 (ฉบับ ก)

$P_{.1}$ แทน สัดส่วนแยกของผู้ที่รอบรู้ในการสอบครั้งที่ 2 (ฉบับ ข)

$P_{.2}$ แทน สัดส่วนแยกของผู้ที่ไม่รอบรู้ในการสอบครั้งที่ 2 (ฉบับ ข)

1.3 วิธีของคาร์เวอร์

คาร์เวอร์ (Caver, 1970 อ้างถึงใน ปราณี ทองคำ, 2539 : 219) ได้เสนอสูตรการหาค่าความเชื่อมั่นของแบบทดสอบโดยการเปรียบเทียบสัดส่วนของความสอดคล้องในการตัดสินความรอบรู้ของผู้สอบแต่ละคนจากแบบทดสอบคู่ขนานหรือแบบทดสอบฉบับเดียวกัน 2 ครั้ง ตามแนวคิดนี้สามารถนำมาใช้ในการประมาณค่าความเที่ยงของแบบทดสอบอิงเกณฑ์ได้ โดยการนำแบบทดสอบคู่ขนานสองฉบับไปสอบกับนักเรียนกลุ่มเดียวกัน แล้วนำผลการสอบไปแจกแจงลงในตาราง วิธีนี้ได้ผลเท่ากับวิธีของแฮมเบิลตันและโนวิค เพราะลักษณะสัดส่วนนั้นเป็นค่าสัดส่วนเดียวกัน

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

สูตร

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

เมื่อ P_o แทน สัดส่วนของความสอดคล้องในการตัดสินความรอบรู้หรือค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์

1.4 วิธีของสับโคเวียค

สับโคเวียค (Subkoviak, 1976 : 265-276) เป็นการหาความเชื่อมั่นในการสอบเพียงครั้งเดียว ด้วยแบบทดสอบคู่ขนานสองฉบับ ซึ่งมีคะแนนเป็น X และ X' โดยที่คะแนนทั้งสองเป็นอิสระกัน และมีการแจกแจงแบบทวินาม เหมือนกันตลอด (Identically Binomial) แต่ค่าสถิติที่นำมาคำนวณมาจากคะแนน X' นิยามของสัมประสิทธิ์ความสอดคล้องสำหรับบุคคล i ว่าเป็นความน่าจะเป็นที่บุคคลนั้นถูกกำหนดให้เป็นผู้รอบรู้เหมือนกัน อันเนื่องมาจากผลการสอบ แบบทดสอบคู่ขนาน เพื่อกำหนดคะแนนเกณฑ์ C โดยเขียนเป็นสมการสัมประสิทธิ์ความสอดคล้อง คือ

$$P_C = \frac{\sum_{i=1}^{N} P_C(i)}{N}$$

เมื่อ P_C แทน ความน่าจะเป็นของการตัดสิน

N แทน จำนวนผู้สอบทั้งหมด

$P_C(i)$ แทน สัมประสิทธิ์ของความสอดคล้อง ของคนที่ i เมื่อกำหนดคะแนนจุดตัดเท่ากับ C

$$P_C(i) = [P(X_i \geq C)]^2 + [1- P(X_i \geq C)]^2$$

โดยที่

$$P(X_i \geq C) = \sum_{X_i=0}^{n} \binom{n}{X_i} P_i^{X_i} (1-P_i)^{n-X_i}$$

เมื่อ P_i แทน ความน่าจะเป็นที่แท้จริงในการตอบข้อคำถามถูกของคนที่ i ค่า ของ P_i สามารถหาได้ ดังนี้

1. ถ้าแบบทดสอบมีจำนวนข้อสอบมากกว่า 40 ข้อ

$$\hat{P}_i = \frac{X_i}{n}$$

เมื่อ X_i แทน จำนวนข้อที่ตอบถูก

n แทน จำนวนข้อสอบทั้งหมดในแบบทดสอบ

2. ถ้าแบบทดสอบมีจำนวนข้อสอบน้อย

$$\hat{P}_i = \alpha_{21}\left(\frac{X_i}{n}\right) + (1-\alpha_{21})\frac{\bar{X}}{n}$$

เมื่อ α_{21} แทน สัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบโดยวิธี Kuder Richardson Formula 21

$\bar{X}$ แทน คะแนนเฉลี่ยที่ได้จากการสอบ

สับโคเวียค พบว่า ถ้าข้อมูลจากแบบทดสอบคู่ขนาน 2 ชุด มีความสัมพันธ์กัน ทางบวก ค่าความเชื่อมั่นจะมีค่าต่ำเกินไป แต่ถ้ามีความสัมพันธ์กันทางลบ ค่าที่ได้จะมีค่าสูงเกินไป และการกระจายแบบทวินามของคะแนนจากข้อสอบที่ผลการตอบแต่ละข้อ เป็นอิสระต่อกันนั้น พบว่า เมื่อคะแนนเกณฑ์ C มีค่าต่ำสุดและสูงสุด ค่าความเชื่อมั่นจะสูงสุด และเมื่อเกณฑ์อยู่กลาง ๆ ค่าความเชื่อมั่นจะมีค่าต่ำสุด

ข้อบกพร่องของวิธีนี้ คือ (Hambleton and others, 1978 : 23) การประมาณค่าความเชื่อมั่น มีความเป็นไปได้ที่จะสูงกว่าความเป็นจริง เนื่องจากได้รวมเอาความสอดคล้องโดยบังเอิญเข้าไว้ (P_o ไม่มีการปรับแก้) นอกจากนี้ในการใช้แบบทดสอบคู่ขนาน เพื่อหาสัมประสิทธิ์ความสอดคล้อง ในการกำหนดความรอบรู้ของนักเรียนแต่ละคน และของกลุ่มนั้นเป็นการปฏิบัติที่ยากที่จะให้แบบทดสอบทั้งสองมีข้อคำถาม (Item) ที่มีความยากง่ายเท่ากันทุกข้อ

1.5 วิธีของฮวิน

ฮวิน (Huynh, 1976 : 253-263) ได้เสนอสูตรการหาความเชื่อมั่นจากการสอบเพียงครั้งเดียว แต่สมมติขึ้นเป็นสองฉบับ โดยสมมติคะแนนที่ได้จากแบบทดสอบอิงมวลความรู้ที่คู่ขนานกันสองฉบับ คือ X และ Y ที่ได้จากการสุ่มจากประชากรข้อสอบเดียวกัน โดยที่ X เป็นคะแนนที่วัดได้จากแบบทดสอบ และ Y เป็นคะแนนที่ได้จากแบบทดสอบที่สมมติขึ้น X และ Y มีการแจกแจงแบบ Beta – Binomial ค่าฟังก์ชันนี้เป็นค่ามีสมมาตร ดังนั้น F(X,Y) = F(Y,X) จากรูปแบบจำลองทางคณิตศาสตร์นี้ ฮวิน ได้เสนอวิธีการในการคำนวณค่าสัมประสิทธิ์แคปป้าจากแบบทดสอบอิงเกณฑ์ จำนวน n ข้อ ดังต่อไปนี้

เมื่อคะแนนจุดตัดมีค่าใกล้ n

$$K = \frac{(P_{11} - P_1^2)}{(P_1 - P_1^2)}$$

และเมื่อคะแนนจุดตัดมีค่าใกล้ 0

$$K = \frac{(P_{00} - P_0^2)}{(P_0 - P_0^2)}$$

เมื่อ P_{11} แทน จำนวนของผู้สอบผ่านในการสอบทั้งสองครั้ง

P_1 แทน จำนวนของผู้สอบผ่านครั้งเดียว

P_{00} แทน จำนวนของผู้สอบไม่ผ่านในการสอบทั้งสองครั้ง

P_0 แทน จำนวนของผู้สอบไม่ผ่านในการสอบครั้งหนึ่ง

2. การหาค่าความเชื่อมั่นโดยพิจารณาจากความเชื่อมั่นของคะแนนจากแบบทดสอบ

อิงเกณฑ์ (Reliability of Criterion Test scores) วิธีนี้เป็นการหาดัชนีความสอดคล้องระหว่างคะแนนของผู้สอบแต่ละคนที่เบี่ยงเบนไปจากคะแนนจุดตัด ซึ่งมีผู้เสนอไว้ดังนี้

2.1 วิธีของเบรนนอน และเคน

เบรนนอน และเคน (Brennan and Kane, 1977 : 277-279) ได้เสนอสูตรการหาความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ เป็นดัชนีความสามารถเชื่อถือได้ (Index of Dependability) ที่ได้มาจากประชากรผู้สอบและข้อสอบที่สุ่มมาจากประชากรข้อสอบ ซึ่งดัชนีความเชื่อถือได้สามารถอธิบายได้ด้วยทฤษฎีสรุปอ้างอิง เป็นการวิเคราะห์ส่วนย่อยของความแปรปรวนที่เป็นผลมาจากข้อสอบรวมกับปฏิสัมพันธ์ของผู้สอบและข้อสอบ ถ้าการให้คะแนนตอบถูก 1 ตอบผิด 0 สูตรความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ คือ

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

เมื่อ $M(c)$ แทน ดัชนีความสามารถเชื่อถือได้

$\overline{X}_{ij}$ แทน ค่าเฉลี่ยคะแนนรวมทั้งหมด

$S^2(\overline{X}_i)$ แทน ค่าความแปรปรวนของค่าเฉลี่ยคะแนนรายบุคคล

C แทน คะแนนจุดตัดของแบบทดสอบ

2.2 วิธีของลิวิงสตัน

ลิวิงสตัน (Livingston, 1972 : 12-26) เสนอสูตรการหาความเชื่อมั่น โดยอาศัยแนวความคิดของการหาค่าความเชื่อมั่นแบบอิงกลุ่ม ซึ่งนิยามความแปรปรวนของคะแนนเป็น Mean Square Deviation ของคะแนนจุดตัดแทนค่าเฉลี่ยของกลุ่ม และนิยามความเชื่อมั่นว่า เป็นอัตราส่วนของความแปรปรวนของคะแนนจริงต่อความแปรปรวนของคะแนนที่สอบได้ตามทฤษฎีแบบดั้งเดิม ซึ่งมีสูตรในการหาความเชื่อมั่น ดังนี้

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

เมื่อ r_{cc} แทน ความเชื่อมั่นตามวิธีของลิวิงสตัน

$\bar{X}$ แทน ค่าเฉลี่ยของคะแนนที่สอบได้

S^2 แทน ความแปรปรวนของคะแนนที่สอบได้

r_{xx} แทน ค่าความเชื่อมั่นแบบอิงกลุ่มซึ่งอาจหาได้จากสูตร KR-20 หรือ KR-21

C แทน คะแนนจุดตัด

เมื่อคะแนนจุดตัดเท่ากับคะแนนเต็ม ค่าความเชื่อมั่นนี้จะเท่ากับความเชื่อมั่นแบบอิงกลุ่ม ซึ่งจะมีค่าต่ำสุด ถ้าคะแนนทั้งหมดตกอยู่ที่คะแนนจุดตัดหรือคะแนนของทุกคนเท่ากัน กล่าวคือ ไม่มีความแปรปรวนของคะแนน จะไม่มีการนิยามดัชนีความเชื่อมั่นของลิวิงสตัน และระยะทางระหว่างคะแนนตัวหนึ่งกับคะแนนจุดตัดมากเท่าใด ค่าคะแนนจริงจะตกอยู่ในด้านเดียวกันกับคะแนนจุดตัดนั้น ดังนั้น ดัชนีความเชื่อมั่นของลิวิงสตันมีความไวต่อระยะทางระหว่างคะแนนจุดตัดกับค่าเฉลี่ย อย่างไรก็ตามวิธีการนี้ใช้ได้ดีกับการกระจายของคะแนนที่มีฐานนิยมเดียว

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

2.3 โลเวทท์ (Lovett) ได้เสนอสูตรหาค่าความเชื่อมั่นของคะแนนแบบทดสอบอิงเกณฑ์ไว้หลายสูตรดังนี้

2.3.1 วิธีหาความเชื่อมั่นจากการแจงแจงแบบทวินาม โลเวทท์ (Lovett) ได้เสนอสูตรการคำนวณความเชื่อมั่นโดยอาศัยการแจกแจงแบบทวินาม เพราะถือว่า เมื่อผู้สอบแต่ละคนตอบข้อสอบ ความน่าจะเป็นของกาตอบข้อสอบถูกมีค่าคงที่เท่ากันหมดทุกข้อ และข้อสอบทุกข้อจะให้คะแนนตอบถูกได้ 1 ตอบผิดได้ 0 โดยมีสูตรในการคำนวณดังนี้

$$r_{cc}(\text{Binomial}) = 1 - \frac{n\sum X_i - \sum X_i^2}{(n-1)\sum (X_i - C)^2}$$

เมื่อ r_{cc}(Binomial) แทน ค่าความเชื่อมั่นของคะแนนแบบทดสอบอิงเกณฑ์

X_I แทน คะแนนของแต่ละคน

N แทน จำนวนข้อสอบ

C แทน คะแนนจุดตัดของแบบทดสอบ

2.3.2 วิธีหาความเชื่อมั่นจากการวิเคราะห์ความแปรปรวนของคะแนนแบบ 0-1 โลเวทท์ (Lovett) ได้เสนอสูตรในการหาค่าความเชื่อมั่นของคะแนนแบบทดสอบอิงเกณฑ์ โดยอาศัยส่วนประกอบของการวิเคราะห์ความแปรปรวนแบบ 0-1 ที่แปรปรวนไปจากคะแนนจุดตัดของข้อสอบ โดยเสนอสูตรไว้ดังนี้

$$r_{cc}\ (\text{ANOVA Corrected}) = 1 - \frac{N(n-1)}{N(n-1)-2} \cdot \frac{MS_E}{MS_P}$$

$$MS_P = \frac{n\sum\left[\frac{X_i}{n} - C\right]^2}{N}$$

$$MS_E = \frac{\sum\sum(X_{ij} - C)^2 - N\sum\left[\frac{X_j}{N} - \bar{X}_{ij}\right]^2 - n\sum\left[\frac{X_i}{n} - C\right]^2}{N(n-1)(n-1)}$$

เมื่อ MS_P แทน ค่าเฉลี่ยของผลบวกของคะแนนเบี่ยงเบนยกกำลังสองระหว่างบุคคล

MS_E แทน ค่าเฉลี่ยของผลบวกของคะแนนเบี่ยงเบนยกกำลังสองของความคลาดเคลื่อน

N แทน จำนวนนักเรียนที่เข้าสอบ

n แทน จำนวนข้อของแบบทดสอบ

C แทน คะแนนจุดตัด

<u>การเรียงข้อคำถามในแบบทดสอบ</u>

การเรียงข้อคำถามในแบบทดสอบได้มีนักวิชาการ และนักวัดผลได้เสนอแนวคิดไว้ดังนี้

ชวาล แพรัตกุล (2516 : 193 อ้างถึงใน อุดม ชูลีวรรณ, 2539 : 55) กล่าวว่า ข้อสอบแต่ละฉบับควรเรียงลำดับคำถามเริ่มจากง่ายไปหายากเสมอ ไม่ควรเรียงตามความยากง่ายของบทเรียน

อนันต์ ศรีโสภา (2525 : 141) กล่าวว่า การจัดเรียงข้อสอบที่ดี ควรจัดเรียงจากข้อง่ายไปหาข้อยาก

ลินควิสท์ (Lindquist, 1955 : 179) กล่าวว่า ในการเรียงข้อคำถาม ไม่จำเป็นต้องจัดเรียงลำดับคำถามตามความยากง่าย ถ้าคำถามเหล่านั้นเป็นเนื้อหาเดียวกัน และมีความยากง่ายพอ ๆ กัน แต่ถ้าคำถามเหล่านั้นมีเนื้อหาต่างกันด้วย แยกข้อคำถามออกเป็นฉบับ ๆ และแยกเวลาในการสอบของแต่ละฉบับ

สแตนเลย์ และฮอฟกิ้นส์ (Stanley and Hopkins, 1972 : 191) กล่าวว่า การเรียงลำดับข้อคำถามควรจะเริ่มต้นจากข้อที่ง่ายที่สุด และให้ข้อยากที่สุดไว้ตอนท้ายของแบบทดสอบจะเป็นไปตามหลักจิตวิทยา เพราะนักเรียนพบข้อยาก ๆ ในตอนต้นของแบบทดสอบอาจทำให้เกิดความท้อใจได้ โดยเฉพาะกลุ่มนักเรียนที่มีความสามารถทางการเรียนระดับปานกลางและต่ำ

ทราเวอร์ (Travers, 1950 : 127-128) ได้เสนอแนะการจัดเรียงข้อคำถามไว้ 4 วิธี ดังนี้

1. เรียงตามลำดับความยาก (Arrangement in order Difficulty) ทำได้โดยเรียงลำดับข้อคำถามตามความยากของข้อคำถาม ซึ่งพิจารณาจากค่าสถิติในการเรียงข้อคำถามแบบนี้ ทำให้ผู้สอบมีกำลังใจ เมื่อพบปัญหาที่ยากขึ้น และมีข้อเสียคือ เนื้อหาและพฤติกรรมที่ต้องการวัดไม่อยู่รวมเป็นพวกเดียวกัน ทำให้ผู้สอบต้องเปลี่ยนความคิดอย่างรวดเร็ว และผู้สอบอาจหมดกำลังใจที่จะทำแบบทดสอบ เพราะเมื่อเริ่มพบข้อยากแล้ว ข้อต่อไปของแบบทดสอบน่าจะยากยิ่งขึ้นไปอีก

2. เรียงลำดับความยากแบบหมุนเวียน (Arrangement in cyclic order of Difficulty) การเรียงข้อคำถามในลักษณะนี้เพื่อหลีกเลี่ยงข้อเสียของการเรียงลำดับข้อคำถามตามระดับความยาก และเพื่อกระตุ้นให้ผู้สอบอ่านข้อสอบทุกข้อ เพราะผู้สอบรู้ว่า ถ้าอ่านต่อไปอีกระยะหนึ่ง จะพบข้อง่ายอีก การจัดเรียงแบบนี้จึงมีประโยชน์ แต่ผู้สอบจะต้องเปลี่ยนความคิดในแต่ละเรื่องอย่างรวดเร็ว

3. เรียงลำดับตามกลุ่มเนื้อหาวิชา (Arrangement According to Subject Matter Area) ข้อดีของวิธีนี้ คือ ผู้สอบจะคิดปัญหาในเนื้อหาเดียวกันก่อนที่จะเปลี่ยนไปคิดเนื้อหาอื่น ๆ

4. เรียงตามจุดมุ่งหมายของการวัด (Arrangement According to the Goals Measured) แบบทดสอบบางชุด ผู้จัดทำข้อสอบจะจัดเรียงเป็นกลุ่มตามพฤติกรรมการวัดประเภทเดียวกัน เช่น แบบทดสอบโค-ออฟเพเรทิฟ (Co-operative Tests) จะจัดข้อคำถามที่วัดความจำเกี่ยวกับคำศัพท์ และความคิดรวบยอด (Terms and Concept) ไว้ด้วยกัน ส่วนการวัดความเข้าใจแยกไว้อีกพวกหนึ่ง ข้อดีของการจัดเรียงวิธีนี้ คือ ครูสามารถตรวจสอบว่าวัดตรงตามจุดมุ่งหมายที่ต้องการวัดหรือไม่

จากการศึกษารูปแบบการจัดเรียงข้อคำถามในแบบทดสอบ สรุปได้ว่า การจัดเรียงแบบทดสอบอาจเรียงตามลำดับความยากง่าย เรียงตามเนื้อหาและเรียงแบบสุ่ม ดังนั้น ผู้วิจัยจึงสนใจนำวิธีการเรียงข้อคำถามในรูปแบบการเรียงลำดับตามเนื้อหา และเรียงลำดับตามความยากมาใช้ในแบบทดสอบอิงเกณฑ์ เพื่อศึกษาเปรียบเทียบความเชื่อมั่นในการจำแนกความรอบรู้

<u>งานวิจัยที่เกี่ยวข้อง</u>

งานวิจัยที่เกี่ยวข้องกับการหาคะแนนจุดตัด

เบฮูเนียค (Behuniak, 1981 อ้างถึงใน จรัสศรี ทองชุมนุม, 2540 : 43) ได้ศึกษาวิธีการหาคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์และการอ่านกับนักเรียนระดับ 9 ตามวิธีของนีเดลสกี และแองกอฟฟ์ โดยใช้ผู้เชี่ยวชาญทางด้านเนื้อหาจำนวน 27 คน จากการศึกษาผลของการตัดสินของผู้เชี่ยวชาญและผลสัมฤทธิ์ทางการเรียน พบว่า

1) วิธีกำหนดคะแนนจุดตัดต่างกันจะทำให้คะแนนจุดตัดที่ได้ต่างกัน

2) วิธีกำหนดคะแนนจุดตัดวิธีเดียวกันแต่ใช้กลุ่มตัวอย่างต่างกันจะทำให้คะแนนจุดตัดที่ได้ต่างกัน

3) วิธีกำหนดคะแนนจุดตัดตามวิธีของนีเดลสกีและแองกอฟฟ์ ให้คะแนนจุดตัดเพิ่มขึ้นหรือลดลงเท่า ๆ กัน

แวน เดอร์ ลินเด็น (Van der Linden, 1982 อ้างถึงใน สุรชัย มีชาญ, 2540 : 31) ใช้ทฤษฎีการตอบสนองข้อสอบ ในการศึกษาความไม่คงเส้นคงวาภายในผู้ตัดสิน โดยวัดความคลาดเคลื่อนในการกำหนดคะแนนจุดตัด วิธีของแองกอฟฟ์ และนีเดลสกี ของผู้ตัดสินแต่

ละคน โดยใช้แบบทดสอบวิชาฟิสิกส์ เรื่อง แรงและการเคลื่อนที่ พบว่า ในการพิจารณาของผู้ตัดสินโดยวิธีของนีเดลสกี มีความคลาดเคลื่อนโดยเฉลี่ยสูงกว่า วิธีของแองกอฟฟ์ และวิธีของนีเดลสกี มีพิสัยของความคลาดเคลื่อนระหว่างคนที่มีความคลาดเคลื่อนสูงสุดกับคนที่มีความคลาดเคลื่อนต่ำสุด มากกว่าวิธีของแองกอฟฟ์

เครเฮน (Crehen, 1974 : 255 - 262) ได้ศึกษาแบบทดสอบอิงเกณฑ์ ใช้วิธีการคัดเลือกข้อสอบ 6 วิธี คือ วิธีการของคอกซ์-วากัส (Cox-vargas) วิธีของเบรนนัน (Brennan) วิธีการเรียงตามค่าสัมประสิทธิ์สหสัมพันธ์แบบ Point Biserial และวิธีเรียงตามสุ่มกลุ่มตัวอย่าง แต่ละกลุ่มใช้ระหว่าง 7-30 คน โดยใช้แบบทดสอบคู่ขนาน พบว่า วิธีการเลือกข้อสอบแบบอิงเกณฑ์ทั้ง 6 วิธีนี้ ให้ผลต่อความเชื่อมั่นไม่แตกต่างกัน ค่าความเที่ยงตรงได้เลือกเกณฑ์ที่ให้ค่าสูงสุดโดยใช้จุดเดียวกันทั้งสองฉบับ ปรากฏว่า วิธีการของคอกซ์-วาร์กัล และเบรนนัน ให้ค่าสูงสุด

อัจฉริยา ปราบอริพ่าย (2532 : 101-106) ได้ศึกษาเปรียบเทียบคะแนนจุดตัดของแบบทดสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 และค่าดัชนีความคงที่ของการตัดสินคะแนนจุดตัด ซึ่งตัดสินโดยผู้เชี่ยวชาญที่ไม่มีสารสนเทศ และผู้เชี่ยวชาญที่มีสารสนเทศ โดยใช้วิธีของแองกอฟฟ์ กลุ่มตัวอย่างผู้เชี่ยวชาญ เป็นครูผู้มีประสบการณ์สอนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ในเขตกรุงเทพมหานคร ปีการศึกษา 2531 จำนวน 96 คน ซึ่งแบ่งเป็น 2 กลุ่ม กลุ่มละ 48 คน ให้กลุ่มหนึ่งตัดสินคะแนนจุดตัด โดยไม่มีสารสนเทศอีกกลุ่มตัดสินคะแนนจุดตัดโดยมีสารสนเทศ (อันได้แก่ ค่าอำนาจจำแนก ค่าความยาก ค่าการเดา ความถี่ของคะแนนสอบสัดส่วนของผู้สอบที่ตอบข้อสอบแต่ละข้อได้ถูกและความน่าจะเป็นที่จะตอบข้อสอบแต่ละข้อได้ถูก) ส่วนกลุ่มตัวอย่างนักเรียน คือนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2531 ในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษาในเขตกรุงเทพมหานคร จำนวน 2,341 คน ซึ่งใช้ในการสอบแบบทดสอบเพื่อหาค่าพารามิเตอร์ของข้อสอบ (อันได้แก่ ค่าอำนาจจำแนก ค่าความยาก และค่าการเดา) ผลการวิจัย พบว่า คะแนนจุดตัดที่ใช้เป็นเกณฑ์ตัดสินความสามารถขั้นต่ำที่ยอมรับได้ของผู้สอบ ซึ่งอยู่บนสเกลของคะแนนโดเมนหรือคะแนนจริงที่ได้จากการตัดสินของผู้เชี่ยวชาญที่ไม่มีสารสนเทศ และผู้เชี่ยวชาญที่มีสารสนเทศ เท่ากับ 0.68 และ 0.53 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนในการหาดัชนีความคงที่ของการตัดสินคะแนนจุดตัด ซึ่งตัดสินโดยผู้เชี่ยวชาญที่ไม่มีสารสนเทศและผู้เชี่ยวชาญที่มีสารสนเทศนั้น ผู้วิจัยได้แปลงคะแนนจุดตัดในสเกลของคะแนนโดเมนให้เป็นระดับความสามารถ โดยอาศัยโค้งคุณลักษณะแบบทดสอบปรากฏว่า ดัชนีความคงที่ของการตัดสินเท่ากับ 0.73 และ 0.76 ตามลำดับ ซึ่งแตกต่างกันอย่างไม่มี

นัยสำคัญทางสถิติ

วุฒิคุณ เสาวรัตน์ (2523 : 161) ได้ศึกษาคะแนนเกณฑ์ที่เหมาะสมของแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องภาคตัดกรวย ในระดับชั้นมัธยมศึกษาปีที่ 4 จำนวน 11 ฉบับ โดยใช้ทฤษฎีการตัดสินใจทั้งสามวิธี คือ วิธีของเบอร์ก วิธีของเบส์ และวิธีของแกลส ปรากฏว่าวิธีที่ทำให้คะแนนเกณฑ์มีค่าสูงสุด คือ วิธีของเบส์ วิธีที่ทำให้คะแนนเกณฑ์สูงรองลงมา คือ วิธีของแกลส และวิธีที่ทำให้คะแนนเกณฑ์มีค่าต่ำสุดคือวิธีของเบอร์ก

สุทธิวรรณ พีรศักดิ์ (2529 : 76-77) ได้ศึกษาคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ชนิดเลือกตอบซ้อนวิชาวิทยาศาสตร์ เรื่อง สู่อวกาศ มีแบบทดสอบแบ่งตามเนื้อหาย่อย 4 ฉบับ และใช้วิธีการหาคะแนนจุดตัดตามวิธีของเบอร์ก วิธีของแกลส และวิธีของ ฮวิน ปรากฏว่าในแบบทดสอบฉบับเดียวกันจะได้คะแนนที่เหมาะสมเป็นจุดตัดเดียวกันและความเชื่อมั่นของแบบทดสอบฉบับเดียวกันแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

กาญจนา วัธนสุนทร (2521 : 75-76) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่อง สมการ ชั้นมัธยมศึกษาปีที่ 1 จำนวน 40 ข้อ แยกเป็น 4 ฉบับ ฉบับละ 10 ข้อ และได้พบว่า คะแนนเกณฑ์จะมีผลต่อค่าความเที่ยงตรงและความเชื่อมั่นของแบบทดสอบ คือ เมื่อกำหนดคะแนนเกณฑ์สูงค่าความเที่ยงตรงและความเชื่อมั่นจะลดลง

ชวลิต โพธิ์นคร (2528 : 93-100) ได้เปรียบเทียบผลของการกำหนดจุดตัดของแบบทดสอบอิงเกณฑ์ชนิดเลือกตอบโดยวิธีของเบอร์ก วิธีประยุกต์ราชค์โมเดล และวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุด ผลการวิจัยพบว่า คะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ที่คำนวณโดยวิธีประยุกต์ราชค์โมเดลในกลุ่มที่มีความสามารถสูงแตกต่างไปจากกลุ่มปานกลางและกลุ่มต่ำ ส่วนกลุ่มปานกลางและกลุ่มต่ำได้คะแนนจุดตัดเท่ากัน นอกจากนั้นยังพบอีกว่าคะแนนจุดตัดที่คำนวณได้โดยวิธีของเบอร์ก มีแนวโน้มจะทำให้ค่าความตรงเชิงโครงสร้างของแบบทดสอบมีค่าสูงที่สุดแต่มีแนวโน้มที่จะแปรเปลี่ยนไปตามระดับความสามารถของกลุ่มตัวอย่าง และคะแนนจุดตัดที่คำนวณโดยวิธีประยุกต์ราชค์โมเดลนั้นมีแนวโน้มไม่แปรเปลี่ยนไปตามระดับความสามารถของกลุ่มตัวอย่าง

ศุภลักษณ์ สินธนา (2531 : 60-63) ได้ศึกษาเปรียบเทียบผลการใช้คะแนนจุดตัดโดยการประยุกต์โลจิสติกโมเดล แบบพารามิเตอร์ 1,2 และ 3 ตัว ในแบบทดสอบอิงเกณฑ์ วิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 ผลการวิจัยพบว่า คะแนนจุดตัดทั้ง 3 แบบที่ได้จากการใช้พารามิเตอร์ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ และจำนวนคนรอบรู้หรือผ่านคะแนนจุดตัดที่ได้จากการใช้พารามิเตอร์ทั้ง 3 แบบ ก็ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

งานวิจัยที่เกี่ยวข้องกับการจัดเรียงลำดับข้อคำถาม

แมค นีคอล (Mac Nical, 1956 : 1-21 อ้างถึงใน โสภณา เหลืองวิลาวัณย์. 2524 : 12) พบว่า การเรียงลำดับข้อคำถามจากข้อยากไปหาข้อง่าย จะทำให้ข้อสอบยากขึ้นกว่าการเรียงลำดับจากข้อง่ายไปหาข้อยากอย่างมีนัยสำคัญทางสถิติ

แฟลงเกอร์ เมลตอน และไมเออส์ (Flangher, 1966 : 29-30 อ้างถึงใน ศิริภรณ์ ศิริใจชิงกุล, 2533 : 24-25) ได้ศึกษาการเรียงลำดับคำถาม 4 วิธี โดยใช้แบบทดสอบ CEEB Scholastic Aptitude Test (SAT) ด้านภาษา 20 ข้อ และด้านคณิตศาสตร์ 30 ข้อ ซึ่งจัดแบ่งข้อสอบเป็นตอน ๆ ตอนละ 5 ข้อ นำมาจัดเรียงใหม่ 4 วิธี วิธีที่ 1 จัดเรียงตามมาตรฐานเดิม คือ เรียงเป็นตอน ๆ ตอนละ 5 ข้อ คือ เรียงจากง่ายไปหายาก วิธีที่ 2 จัดเรียงภายในแต่ละตอนใหม่ ส่วนระหว่างตอนยังเรียงเหมือนมาตรฐานเดิม วิธีที่ 3 จัดเรียงระหว่างตอนใหม่ ส่วนการเรียงภายในตอนแต่ละตอนเหมือนเดิม วิธีที่ 4 จัดเรียงใหม่ทั้งภายในแต่ละตอนและระหว่างตอน ซึ่งผลการวิจัยพบว่า การจัดเรียงข้อสอบด้วยวิธีต่างกัน มีผลทำให้คะแนนจากแบบทดสอบด้านภาษาแตกต่างกัน แต่สำหรับแบบทดสอบด้านคณิตศาสตร์ ไม่มีผลแตกต่างกัน

โสภณา เหลืองวิลาวัณย์ (2524 : 59-61) ได้ศึกษาผลของการเรียงลำดับข้อสอบและการได้รู้ค่าความยากของข้อสอบ ของแบบทดสอบที่มีต่อความยาก ความตรง และเที่ยงของแบบทดสอบผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ โดยเรียงลำดับข้อคำถามต่างกัน 4 แบบ คือ เรียงจากข้อง่ายไปหาข้อยากโดยไม่คำนึงถึงเนื้อหา เรียงตามลำดับเนื้อหาของหัวข้อที่เรียน เรียงจากข้อง่ายไปหาข้อยากตามลำดับเนื้อหาหัวข้อที่เรียนและเรียงลำดับข้อสอบแบบสุ่มโดยไม่คำนึงถึงเนื้อหาหรือความยากของข้อสอบ พบว่าการเรียงลำดับข้อถามทั้ง 4 แบบ ไม่มีผลต่อค่าความยาก ความตรง และความเที่ยง

ศิริภรณ์ ศิริใจชิงกุล (2533 : บทคัดย่อ) ศึกษาเปรียบเทียบคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ชนิดเลือกตอบ ที่เรียงลำดับข้อคำถามด้วยวิธีต่างกันในวิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 4 โดยคำนวณจากวิธีของเบอร์กและวิธีของแกลส ซึ่งการจัดเรียงข้อสอบแบ่งออกเป็น 4 แบบ คือ เรียงจากข้อง่ายไปหาข้อยาก เรียงจากข้อยากไปหาข้อง่าย เรียงตามหัวข้อเนื้อหาวิชา และเรียงจากข้อง่ายไปหาข้อยากในแต่ละหัวข้อเนื้อหาวิชา พบว่า คะแนนจุดตัดที่คำนวณโดยวิธีของเบอร์กและวิธีของแกลส ของแบบทดสอบที่เรียงลำดับข้อคำถามจากข้อยากไปหาข้อง่าย แตกต่างไปจากแบบทดสอบที่เรียงลำดับข้อคำถามด้วยวิธีอื่น และคะแนนจุดตัดของ

แบบทดสอบที่เรียงลำดับข้อคำถามต่างกัน โดยพิจารณาทีละแบบระหว่างวิธีของเบอร์กและวิธีของแกลส ได้คะแนนจุดตัดที่เท่ากัน

จากเอกสารและงานวิจัยที่เกี่ยวข้องทั้งในประเทศ และต่างประเทศที่กล่าวมาทั้งหมด จะเห็นได้ว่าความเชื่อมั่นและการกำหนดคะแนนจุดตัด หรือเกณฑ์มาตรฐาน มีความสำคัญต่อการวัดผลแบบอิงเกณฑ์ ซึ่งการหาความเชื่อมั่นและการกำหนดคะแนนจุดตัดยังต้องคำนึงถึงข้อมูลในหลาย ๆ ด้าน เพื่อให้การพิจารณาตัดสินผลการสอบมีความคลาดเคลื่อนเกิดขึ้นน้อยที่สุด ดังนั้นผู้วิจัยมีความสนใจศึกษาผลของการจัดเรียงข้อสอบที่มีรูปแบบต่างกันและการกำหนดคะแนนจุดตัดด้วยวิธีต่างกัน จะส่งผลต่อความเชื่อมั่นของแบบทดสอบอิงเกณฑ์อย่างไร