

## เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยเรื่องความเที่ยงและความตรงของแบบสอบอิงเกณฑ์ชนิดเลือกตอบและตอบสั้นเมื่อใช้กับนักเรียนที่มีความสามารถทางการเรียนแตกต่างกัน ผู้วิจัยได้ศึกษาเอกสารเกี่ยวกับการวัดผลแบบอิงเกณฑ์ การสร้างแบบสอบอิงเกณฑ์ ความเที่ยงและความตรงของแบบสอบอิงเกณฑ์ การกำหนดคะแนนจุดตัดของแบบสอบอิงเกณฑ์ แบบสอบปรนัยชนิดเลือกตอบและตอบสั้น และงานวิจัยที่เกี่ยวข้องกับการกำหนดคะแนนจุดตัดของแบบสอบอิงเกณฑ์ ความเที่ยงและความตรงของแบบสอบเลือกตอบและแบบสอบตอบสั้น โดยเสนอตามลำดับต่อไปนี้

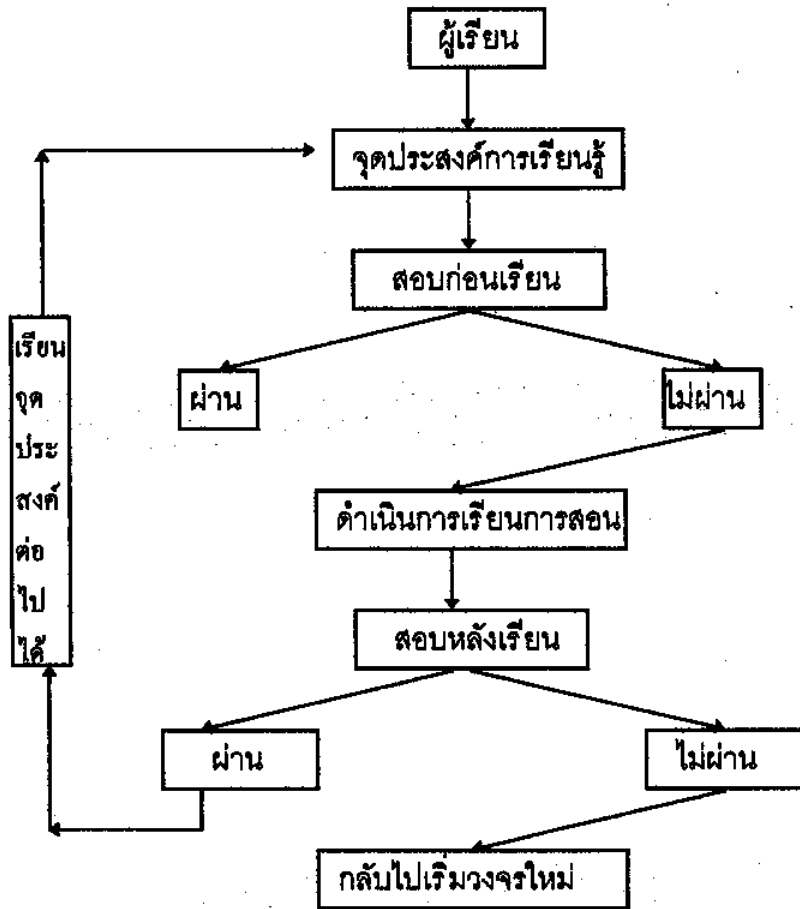
### การวัดผลแบบอิงเกณฑ์

#### 1. แนวคิดของการวัดผลแบบอิงเกณฑ์

การวัดผลแบบอิงเกณฑ์ (Criterion-Referenced Measurement) เป็นการวัดที่เป็นผลสืบเนื่องมาจากทฤษฎีการเรียนรู้เพื่อรอบรู้ (Mastery Learning) ของบลูม (Bloom) ซึ่งกล่าวว่า การเรียนรู้ทั้งหลาย ควรเป็นการเรียนรู้เพื่อรอบรู้ในเนื้อหาต่าง ๆ ระดับที่เรียกว่านักเรียนเป็นผู้รอบรู้แล้วก็ควรเป็น 80 หรือ 90 เปอร์เซ็นต์ของเนื้อหาทั้งหมดที่สอน (โกวิท ประวาลพุกษ์ และ ส.วาสนา ประวาลพุกษ์, 2518 : 22) การวัดผลในระบบอิงเกณฑ์จึงเป็นการเปรียบเทียบผลการวัดกับเกณฑ์ทางการศึกษา เกณฑ์ทางการศึกษาในที่นี้ หมายถึง กลุ่มของพฤติกรรมที่คาดหวัง หรือกลุ่มของการปฏิบัติ (Performance Domain) และกลุ่มของพฤติกรรมนี้มีลักษณะเหมือนประชากรของพฤติกรรมย่อยๆ กลุ่มหนึ่งที่มีสมบัติร่วมกันบางประการ (อเนก เพ็ชรอนุกุลบุตร, 2527 : 551)

อูสเตอร์รอฟ (Oosterhof, 1990 : 202) ได้กล่าวถึงการวัดผลแบบอิงเกณฑ์ว่า คะแนนจะถูกแปลความหมายในเทอมของพฤติกรรมที่ได้จากการสอบเป็นเกณฑ์ในการแปลความหมายในการวัดไม่จำเป็นต้องมีรายละเอียดมากนักแต่ต้องรัดกุม

การจัดการเรียนการสอนที่สอดคล้องกับแนวการวัดแบบอิงเกณฑ์นั้น มิลล์แมน (Millman, 1979 quoted in Popham, 1975 : 52) ได้เสนอรูปแบบในการจัดการเรียนการสอนเพื่อช่วยเหลือนักเรียนเป็นรายบุคคลไว้ดังนี้



ภาพ 1 รูปแบบการจัดการเรียนการสอนของมิลล์แมน

จากรูปแบบนี้แสดงให้เห็นว่าการดำเนินการเรียนการสอน ได้เน้นการช่วยเหลือนักเรียนรายบุคคลเป็นลำดับขั้น คือเมื่อนักเรียนไม่สามารถผ่านเกณฑ์ขั้นแรกได้ ก็จะต้องไปเริ่มต้นเรียนใหม่จนกว่าจะสามารถเรียนได้

เอนก เพียรอนุกุลบุตร (2527 : 553-555) ได้กล่าวถึงการเรียนรู้ที่เป็นรากฐานของการวัดและประเมินผลแบบอิงเกณฑ์ ในทัศนะของการเรียนเพื่อรอบรู้เห็นว่า ผู้เรียนจะต้องเรียนรู้ทั้งหมดในสิ่งที่เรียน โดยต้องมีความสามารถหรือสมบัติถึงระดับที่คาดหวังไว้ หรือถึงมาตรฐานของเกณฑ์ที่กำหนดไว้ เนื่องจากแนวคิดเกี่ยวกับการเรียนเพื่อรอบรู้มีความคิดพื้นฐานว่า คนปกติธรรมดาสามารถเรียนรู้เรื่องราวทั่ว ๆ ไปได้แต่อัตราเร็วต่างกันและด้วยวิธีการต่างกันและการเรียนรู้เรื่องใดเรื่องหนึ่งต้องเป็นการเรียนรู้จนถึงมาตรฐานของเกณฑ์ที่กำหนดไว้ในจุดมุ่งหมาย ดังนั้นหลักการสำคัญของแนวคิดนี้จึงมีลักษณะดังนี้

1) นักเรียนแต่ละคนมีความแตกต่างกัน ทั้งในแง่ความถนัดและความสามารถในการเข้าใจเรื่องราว ดังนั้น "อัตราเร็ว" ในการเรียนของแต่ละคนจะไม่เท่ากัน นักเรียนเกือบทุกคนจะสามารถเรียนรู้ จนถึงระดับ "รู้รอบ" ได้ในระดับเดียวกัน ถ้าให้เวลาแก่พวกเขาอย่างเพียงพอ และใช้วิธีการเรียน วิธีสอน และอุปกรณ์ให้สอดคล้องกับลักษณะเฉพาะตัวของเขา

2) เนื่องจากผู้เรียนมีความแตกต่างกันในแง่ต่าง ๆ อีกมาก เช่น ในด้านเจตคติ แรงจูงใจใฝ่สัมฤทธิ์ แบบของการคิด ฯลฯ ดังนั้นการสอนทุกคนด้วยวิธีเดียวกัน ใช้วัสดุอุปกรณ์ชนิดเดียวกันย่อมจะไม่มีผลดีที่สุดสำหรับทุกคน บางคนเรียนได้ไม่ดี อาจเป็นผลจากวิธีสอน วัสดุอุปกรณ์ วิธีเรียนที่มอบให้ ไม่เหมาะสมแก่เขาก็ได้ ดังนั้นจึงควรจัดแบบของวิธีสอนเนื้อหา อุปกรณ์ให้เหมาะสมกับนักเรียน เพื่อให้แต่ละคนได้พัฒนาถึงระดับสูงสุดตามวิธีการที่เหมาะสมกับผู้เรียน

3) กิจกรรมการเรียนการสอนของแต่ละบุคคลมีการเปลี่ยนแปลงเข้าสู่ "การเรียนรู้รอบ" ทั้งสิ้นและการตัดสินใจว่า "ใครเรียนรู้รอบแล้ว" จะต้องพิจารณาพฤติกรรมของบุคคลว่าสอดคล้องกับมาตรฐานที่ตั้งไว้หรือไม่

ดังนั้น การเรียนการสอนที่สอดคล้องหรือเป็นไปตามหลักการเรียนเพื่อรอบรู้จึงเป็น "การสอนแบบรายคน" (Individualized Instruction) และการประเมินผลจากการเรียนการสอน ต้องเป็นแบบที่เปรียบเทียบพฤติกรรมกับจุดมุ่งหมายหรือเกณฑ์ ซึ่งเป็นการวัดแบบอิงเกณฑ์

เนื่องจากการการเรียนเพื่อรอบรู้มีหลักความคิดพื้นฐานอยู่ที่จัดสภาพการเรียนให้สอดคล้องกับแต่ละบุคคลซึ่งต้องใช้เวลาต่างกันและใช้วิธีที่เหมาะสม และต้องกำหนดเกณฑ์ให้ชัดเจน แนวทางการเรียนรู้เพื่อรอบรู้มีดังนี้

1) กำหนดจุดมุ่งหมายในการเรียนการสอนให้ชัดเจน โดยกำหนดเป็นจุดมุ่งหมายที่เฉพาะเจาะจง ในรูปแบบของจุดมุ่งหมายเชิงพฤติกรรม แล้วกำหนดมาตรฐานของการปฏิบัติที่จะบรรลุจุดมุ่งหมายเชิงพฤติกรรมนั้น

2) ใช้แนวการสอนแบบการสอนรายบุคคล (Individualized Instruction) โดยคำนึงถึงความแตกต่างรายบุคคล

3) ใช้การประเมินผลการเรียนการสอน 2 แบบ คือ ประเมินผลย่อย (Formative Evaluation) เป็นการประเมินครั้งย่อย ๆ ตามจุดมุ่งหมายเชิงพฤติกรรมครั้งละไม่กี่ข้อ อาจจะประเมินเมื่อจบสิ้นการเรียนแต่ละหน่วยย่อยเพื่อพิจารณาว่าใครบรรลุแล้ว ใครยังไม่บรรลุ เพื่อนำไปสู่การตัดสินใจ เช่น การสอนซ่อมเสริม หรือการเรียนในหน่วยต่อไปซึ่งเรียงอันดับต่อเนื่องกันได้ จึงใช้การสอบย่อยนี้ สำหรับวินิจฉัยผลการเรียนได้ดี

จากที่กล่าวมาทั้งหมดจึงสรุปได้ว่า การวัดผลแบบอิงเกณฑ์ เป็นการวัดเพื่อที่จะให้ทราบว่าผู้เรียนได้เรียนรู้ตามจุดประสงค์ที่ต้องการหรือได้เรียนรู้ถึงเกณฑ์ที่กำหนดแล้วหรือยัง โดยใช้แบบสอบที่สร้างขึ้นมาอย่างเฉพาะเจาะจงสอดคล้องกับจุดประสงค์เชิงพฤติกรรมของเนื้อหาที่เรียน และจะต้องมีการกำหนดเกณฑ์ที่เหมาะสม

## 2. ลักษณะของแบบสอบอิงเกณฑ์

แบบสอบอิงเกณฑ์ ใช้สำหรับการวัดความรู้ความสามารถของนักเรียนแต่ละคนว่าถึงเกณฑ์ที่คาดหวังไว้หรือไม่ บางทีเรียกแบบสอบนี้ว่า Master Test แบบสอบนี้จึงเป็นการวัดความรู้และทักษะต่าง ๆ ในตัวนักเรียนว่ามีถึงเกณฑ์ที่กำหนดไว้หรือไม่ โดยไม่คำนึงถึงความแตกต่างระหว่างนักเรียน (อนันต์ ศรีโสภา, 2525 : 190)

เกรเซอร์และนิทโก (Glaser and Nitko, 1971 quoted in Hambleton and Novick, 1973 : 160) ได้กล่าวถึงแบบสอบอิงเกณฑ์ว่าเป็นแบบสอบที่ตั้งใจสร้างขึ้น เพื่อให้ได้ผลของการวัดที่แปลความหมายได้โดยตรงตามมาตรฐานของความสามารถที่กำหนดไว้

พอปแฮม (Popham, 1978 : 91) ได้กล่าวถึงแบบสอบอิงเกณฑ์ว่า เป็นแบบสอบที่สร้างขึ้นเพื่อแสดงถึงผลของการปฏิบัติของผู้สอบแต่ละคนอย่างชัดเจนโดยไม่มีการคำนึงถึงหรือนำไปเปรียบเทียบกับผลการปฏิบัติของผู้สอบคนอื่น ๆ

สงบ ลักษณะ (2523 : 36-37) ได้กล่าวถึงลักษณะของแบบสอบแบบอิงเกณฑ์ไว้ 3 ประการดังนี้

- 1) เป็นแบบสอบที่สร้างจากพื้นฐานของจุดประสงค์ หรือสิ่งที่ต้องการให้สอบปฏิบัติที่ได้รับการนิยามไว้อย่างชัดเจน
- 2) รายชื่อของแบบสอบจะต้องวัดจุดประสงค์ หรือสิ่งที่ต้องการให้ผู้สอบนั้นปฏิบัติได้
- 3) คะแนนที่ได้ควรแปลความหมายว่า ผู้สอบสามารถบรรลุตามเกณฑ์ที่กำหนดไว้หรือไม่

โกวิท ประวาลพุกฤษ (2523 : 18-19) ได้สรุปลักษณะของแบบสอบอิงเกณฑ์ว่ามี 2 ประเภทคือ

- 1) การยึดจุดประสงค์เป็นหลัก หรือเรียกว่าอิงจุดประสงค์ กล่าวคือ ในการเรียนการสอนมีการกำหนดจุดประสงค์เชิงพฤติกรรม และการสอบวัดก็ยึดเอาจุดประสงค์เหล่านั้นเป็นเกณฑ์เพื่อจะชี้บ่งว่าใครผ่านเกณฑ์ที่กำหนดไว้ในจุดประสงค์แล้วหรือยัง

2) การขยายจุดประสงค์ให้ครอบคลุมประชากรของพฤติกรรมที่เรียกว่า โดเมน ข้อสอบต่าง ๆ ได้มาจากการสุ่มตัวอย่างของพฤติกรรมที่กำหนดจากโดเมน แบบสอบในลักษณะนี้เรียกว่า Domain-Referenced Test ซึ่งเมื่อสอบแล้วสามารถอ้างอิงได้ว่าบุคคลมีความสามารถในระดับใดตามคะแนนมาตรฐานเนื้อหาหรือมีระดับความรู้ขนาดไหน

จากที่กล่าวมาทั้งหมดสามารถสรุปได้ว่าแบบสอบอิงเกณฑ์มีลักษณะที่สำคัญคือ เป็นแบบสอบที่สร้างขึ้นมาจากพื้นฐานของจุดประสงค์ของการเรียนรู้ เพื่อให้ทราบว่าผู้เรียนได้เรียนรู้บรรลุจุดประสงค์ที่กำหนดไว้แล้วหรือไม่ โดยจะไม่มีให้นำผลการสอบของแต่ละคนมาเปรียบเทียบกัน

### การสร้างแบบสอบอิงเกณฑ์

แบบสอบอิงเกณฑ์เป็นแบบสอบที่เหมาะสมสำหรับใช้ในการประเมินผลการเรียนในระดับชั้นเรียน การสร้างแบบสอบอิงเกณฑ์ มีความเกี่ยวข้องโดยตรงกับการแปลความหมายของคะแนนจากการปฏิบัติได้จริงของผู้เรียนว่า มีความรู้ในจุดมุ่งหมายที่กำหนดไว้มากน้อยเพียงใด ที่ผ่านการเรียนการสอนไปแล้ว ดังนั้นแนวการสร้างแบบสอบอิงเกณฑ์จึงควรกำหนดไว้อย่างชัดเจน

สแตนเลย์ และฮอปคินส์ (Stanley and Hopkins, 1972 : 171) ได้เสนอหลักทั่ว ๆ ไปของการสร้างแบบสอบอิงเกณฑ์ว่าประกอบด้วยขั้นตอนที่สำคัญ 4 ขั้นตอนคือ

- 1) ขั้นการวางแผนสร้างแบบสอบ
- 2) ขั้นการเตรียมการเขียนข้อสอบและลงมือเขียนข้อสอบ
- 3) ขั้นการทดลอง
- 4) ขั้นการประเมินผลการทดลอง

รอย และฮาลาดินา (Roid and Haladyna, 1980 : 293-295) ได้เสนอขั้นตอนในการสร้างของแบบสอบวัดผลสัมฤทธิ์ ไว้ 5 ขั้นตอนดังนี้

1) มโนทัศน์แนวความคิด เหตุผลในการเรียนการสอน หรือการทดสอบ ในขั้นนี้ผู้สร้างเครื่องมือ หรือผู้สอนจะต้องกำหนดพฤติกรรมที่ต้องการให้เกิดขึ้นในตัวผู้เรียนหลังจากได้รับการเรียนการสอนแล้ว และจะต้องสัมพันธ์กับโดเมนของเนื้อหา ซึ่งอาจจะหาได้โดยการวิเคราะห์เนื้อหากลุ่มของงานหรือวิเคราะห์งาน

2) เป็นขั้นที่เกี่ยวข้องกับเนื้อหาของการสอน ซึ่งแบ่งออกเป็น

2.1) กำหนดจุดประสงค์ในการเรียนการสอน พฤติกรรมของนักเรียน ที่เป็นผลของการสอน

2.2) กำหนดลักษณะเฉพาะของโดเมนเนื้อหา อันสอดคล้องกับมโนทัศน์ที่มุ่งให้เกิดแก่ผู้เรียน

3) เป็นการพัฒนาข้อสอบ เป็นวิธีการหนึ่งของเทคนิคการเขียนข้อสอบ ซึ่งข้อสอบที่สร้างนั้นควรให้เป็นไปตามจักรวาลหรือโดเมนของข้อสอบ อันจะเป็นตัวแทนของการเรียนการสอนที่ต้องการ

4) เป็นการทบทวนข้อสอบ ปรับปรุงข้อสอบ ในขั้นนี้หากมีข้อคำถามที่บกพร่องต้องปรับปรุงให้เรียบร้อยก่อนนำไปใช้ทดสอบต่อไป

5) คัดเลือกข้อสอบ ปัจจุบันนิยมคัดเลือกข้อสอบ โดยการสุ่มตัวอย่างข้อสอบจากโดเมนข้อสอบ

สมศักดิ์ สินธุระเวชญ์ (2521 : 79-80) ได้กล่าวถึงการสร้างแบบสอบอิงเกณฑ์ว่ามีลำดับขั้นตอนในการสร้างดังนี้

1) เขียนจุดมุ่งหมายให้ชัดเจน

2) เตรียม Content Outline ที่จะออกข้อสอบ

3) สร้างข้อคำถามให้สอดคล้องกับ Content Outline โดยสอบเฉพาะส่วนที่สำคัญ ๆ

4) ตรวจสอบแบบสอบที่สร้างว่าสอดคล้องตามจุดมุ่งหมายหรือไม่

บุญเชิด ภิญญอนันตพงษ์ (2527 : 38-45) ได้เสนอขั้นตอนการสร้างแบบสอบอิงเกณฑ์ โดยดัดแปลงมาจากแนวคิดของ รอยด์ และฮาลาไดนา (Roid and Haladyna, 1980 : 294-295) และเฟรมเมอร์ (Fremmer, 1974) ไว้ 7 ขั้นตอนคือ

ขั้นที่ 1 วิเคราะห์เนื้อหาและจุดประสงค์

ขั้นที่ 2 แปลงจุดประสงค์หลักให้เฉพาะเจาะจง

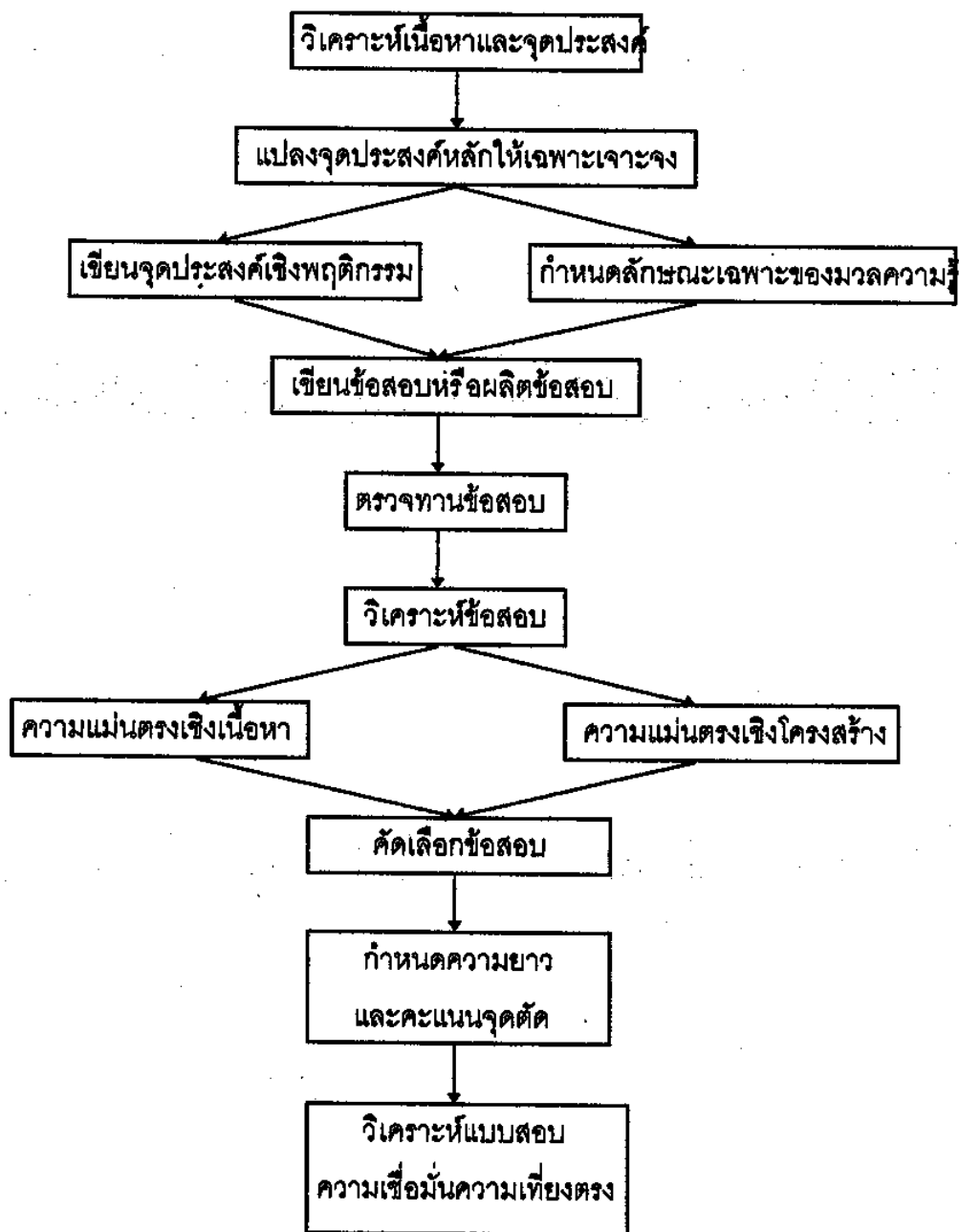
ขั้นที่ 3 เขียนข้อสอบหรือผลิตข้อสอบ

ขั้นที่ 4 ตรวจสอบข้อสอบ

ขั้นที่ 5 วิเคราะห์ข้อสอบ

ขั้นที่ 6 คัดเลือกข้อสอบ กำหนดความยาวและคะแนนจุดตัดของแบบสอบ

ขั้นที่ 7 วิเคราะห์แบบสอบ



ภาพ 2 แสดงขั้นตอนการสร้างแบบสอบอิงเกณฑ์

### การวิเคราะห์เนื้อหาและจุดประสงค์

วิเคราะห์เนื้อหาและจุดประสงค์ เป็นการวิเคราะห์เนื้อหาวิชาที่ต้องการวัดเพื่อให้ได้จุดประสงค์หรือพฤติกรรมที่ต้องการให้เกิดในหัวข้อเนื้อหานั้น ๆ อาจต้องใช้วิธีการวิเคราะห์กิจกรรมการเรียนรู้ (Task Analysis) โดยวิเคราะห์พฤติกรรมตามลำดับของการเรียนรู้ ซึ่งจะทำการทราบพฤติกรรมย่อย ๆ มากมาย ทั้งพฤติกรรมปลายทาง ต้นทาง หรือพฤติกรรมระหว่างทาง

ซึ่งกลุ่มพฤติกรรมปลายทางของแต่ละหัวข้อเนื้อหาวิชา คือพฤติกรรมหลักหรือจุดประสงค์หลัก ที่ต้องการวัด

การแปลงจุดประสงค์หลักให้เฉพาะเจาะจง

จุดประสงค์หลักที่วิเคราะห์ได้จะต้องนำมาแปลงให้เฉพาะเจาะจง โดยนำจุดประสงค์หลักหรือพฤติกรรมหลักมาแตกเป็นพฤติกรรมย่อย ๆ เมื่อได้พฤติกรรมย่อย ๆ ของแต่ละหัวข้อ แล้วนำพฤติกรรมเหล่านั้นมาเขียนเป็นจุดประสงค์เชิงพฤติกรรม หรือเขียนเป็นลักษณะเฉพาะของมวลความรู้

การเขียนข้อสอบหรือผลิตข้อสอบ

เมื่อแปลงจุดประสงค์หลักเป็นจุดประสงค์เชิงพฤติกรรมแล้ว ก็ทำการเขียนข้อสอบตามจุดประสงค์เชิงพฤติกรรมนั้น หากแปลงจุดประสงค์หลักเป็นลักษณะเฉพาะของมวลความรู้ สำหรับเขียนข้อสอบแล้วก็สามารถผลิตข้อสอบจากกฎเกณฑ์ต่าง ๆ ที่ระบุไว้ตามลักษณะเฉพาะที่ต้องการ

การตรวจทานข้อสอบ

การตรวจทานข้อสอบ เป็นการสำรวจตรวจสอบข้อสอบเป็นรายข้อหลังจากเขียนข้อสอบเสร็จแล้ว เพื่อแก้ไขปรับแต่งข้อสอบให้มีความเหมาะสมกับกลุ่มตัวอย่างนักเรียนที่จะสอบ และให้มีความถูกต้องในหลักวิชาทั้งด้านเนื้อหาและพฤติกรรมที่ต้องการทดสอบ

การวิเคราะห์ข้อสอบ

การวิเคราะห์ข้อสอบ เป็นการประเมินคุณภาพของข้อสอบด้านความตรงเชิงเนื้อหา และความตรงเชิงโครงสร้าง เพื่อคัดเลือกข้อสอบที่มีคุณภาพมารวมฉบับ หรือเก็บไว้ในคลังข้อสอบการตรวจสอบความตรงของแบบข้อสอบอิงเกณฑ์ เป็นการประเมินว่าข้อสอบนั้นสามารถวัดพฤติกรรมได้ตรงตามที่ระบุไว้ในจุดประสงค์หรือไม่ และสามารถบ่งชี้ผู้ที่มีความรอบรู้ ได้ถูกต้องหรือไม่

ความตรงเชิงเนื้อหา (Content Validity)

ความตรงเชิงเนื้อหาของข้อสอบอิงเกณฑ์ หมายถึง ความสอดคล้องระหว่างพฤติกรรมที่ข้อสอบวัดกับพฤติกรรมที่ระบุไว้ในจุดประสงค์เชิงพฤติกรรม หรือข้อสอบสามารถวัดพฤติกรรมได้ตรงตามที่ระบุไว้ในจุดประสงค์ (บุญเชิด ภิญโญอนันตพงษ์, 2527 : 67)

การวิเคราะห์ความตรงเชิงเนื้อหาของข้อสอบอิงเกณฑ์ได้มีผู้เสนอไว้หลายวิธีด้วยกัน ดังนี้

โรวินเนลลี และแฮมเบิลตัน (Rovinelli and Hambleton) ได้เสนอการวิเคราะห์ความตรงเชิงเนื้อหาโดยอาศัยดุลพินิจของผู้เชี่ยวชาญไว้ 3 วิธีดังนี้ (บุญเชิด ภิญโญอนันตพงษ์, 2527 : 68)



1) การวิเคราะห์ความตรงเชิงเนื้อหาโดยใช้ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์

คำนวณจากสูตรดังต่อไปนี้

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

เมื่อ IOC คือ ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์  
 $\Sigma R$  คือ ผลรวมคะแนนความคิดเห็นของผู้เชี่ยวชาญเนื้อหาวิชาทั้งหมด  
 $N$  คือ จำนวนผู้เชี่ยวชาญเนื้อหาวิชา

โดยจะนำจุดประสงค์เชิงพฤติกรรม และข้อสอบที่วัดจุดประสงค์นั้น ๆ ไปให้ผู้เชี่ยวชาญเนื้อหาวิชา (อย่างน้อย 3 คน) แต่ละคนพิจารณาลงความคิดเห็นว่าข้อสอบแต่ละข้อวัดจุดประสงค์เชิงพฤติกรรมที่ต้องการวัดหรือไม่ โดยกำหนดคะแนนความคิดเห็นไว้ดังนี้

+1 = แน่ใจว่าข้อสอบวัดจุดประสงค์ข้อนั้น  
 0 = ไม่แน่ใจว่าข้อสอบวัดจุดประสงค์ข้อนั้นหรือไม่  
 -1 = แน่ใจว่าข้อสอบไม่วัดจุดประสงค์ข้อนั้น

จากนั้นนำความคิดเห็นของผู้เชี่ยวชาญเนื้อหาวิชา มาคำนวณตามสูตรโดยใช้เกณฑ์ขั้นต่ำที่สามารถยอมรับได้ เป็น 5

2) การวิเคราะห์ความตรงตามเนื้อหาโดยใช้ดัชนีความเหมาะสมระหว่างข้อสอบกับจุดประสงค์ คำนวณได้จากสูตรค่าเฉลี่ยของคะแนนและค่าความแปรปรวนของคะแนน

$$\bar{X} = \frac{\Sigma X}{N}$$

$$S = \sqrt{[N\Sigma X^2 - (\Sigma X)^2] / N[N-1]}$$

เมื่อ  $\bar{X}$  คือ ค่าเฉลี่ยคะแนนความคิดเห็น  
 $S$  คือ ค่าความเบี่ยงเบนมาตรฐานของคะแนนความคิดเห็น  
 $X$  คือ คะแนนความคิดเห็นของผู้เชี่ยวชาญเนื้อหาวิชา  
 $N$  คือ จำนวนผู้เชี่ยวชาญเนื้อหาวิชา

ค่าเฉลี่ยคะแนนความคิดเห็นจะเป็นดัชนีบ่งชี้ความเหมาะสม และค่าความแปรปรวนของคะแนนความคิดเห็นจะเป็นดัชนี บ่งชี้มติความคิดเห็นว่าเป็นเอกฉันท์หรือขัดแย้งกัน โดยจะ

นำจุดประสงค์เชิงพฤติกรรม และข้อสอบที่วัดจุดประสงค์นั้น ๆ ไปให้ผู้เชี่ยวชาญเนื้อหาวิชา (อย่างน้อย 3 คน) พิจารณาลงความคิดเห็นว่าข้อสอบแต่ละข้อวัดจุดประสงค์เชิงพฤติกรรม ที่ต้องการวัดมากน้อยเพียงไร โดยอาศัยมาตราส่วนประเมินค่า

3) การวิเคราะห์ความตรงเชิงเนื้อหาโดยใช้ดัชนีการจับคู่ระหว่างข้อสอบกับจุดประสงค์ วิธีนี้จะนำข้อสอบและจุดประสงค์เชิงพฤติกรรมชุดหนึ่ง ไปให้ผู้เชี่ยวชาญจับคู่ว่าข้อสอบข้อใด วัดจุดประสงค์ใด จากนั้นนำผลมาแจกแจงลงในตารางที่จำแนกตามข้อสอบและจุดประสงค์ แล้วคำนวณค่าไคสแควร์ ( $\chi^2$ ) ทดสอบความเป็นอิสระแก่กันของข้อมูล ถ้ามีนัยสำคัญแสดงว่า ข้อสอบและจุดประสงค์มีความสัมพันธ์กัน หรือจะใช้วิธีกำหนดเกณฑ์ของเปอร์เซ็นต์ความถี่ โดยใช้เกณฑ์ตั้งแต่ 80% ขึ้นไป

#### ความตรงเชิงโครงสร้าง (Construct Validity)

ข้อสอบที่มีความตรงเชิงโครงสร้าง หมายถึง ข้อสอบเมื่อนำไปสอบกับนักเรียน กลุ่มตัวอย่างแล้ว นักเรียนที่มีความรอบรู้ตอบข้อสอบข้อนั้นถูก ส่วนนักเรียนที่ยังไม่รอบรู้ ตอบข้อสอบข้อนั้นผิด ดังนั้นการวิเคราะห์ความตรงเชิงโครงสร้างของข้อสอบอิงเกณฑ์ ต้องอาศัยเทคนิคการทดลอง หรืออาศัยข้อมูลที่ประจักษ์ชัด ซึ่งได้แก่การหาค่าพารามิเตอร์ ของข้อสอบสองค่า คือ ค่าความยากและค่าอำนาจจำแนกของข้อสอบ ซึ่งจะเป็นดัชนีบ่งชี้ ความตรงเชิงโครงสร้าง (บุญเชิด ภิญโญอนันตพงษ์, 2527 : 80)

#### ค่าความยากของข้อสอบ

เมโย (Mayo) มีความเห็นว่าเนื่องจากข้อสอบสร้างให้สอดคล้องกับจุดมุ่งหมาย เชิงพฤติกรรมซึ่งนักเรียนทุกคนจะต้องผ่านจุดมุ่งหมายนี้ ค่าความยากของข้อสอบจึงควรมีค่าสูง และบลูม (Bloom) กล่าวว่า ค่าความยากของแบบสอบอิงเกณฑ์มิได้ถือเป็นเรื่องสำคัญ สิ่งสำคัญ คือเขียนคำถามให้สอดคล้องกับจุดประสงค์เชิงพฤติกรรม (Kibler and others, 1974 อ้างถึงใน สำเร็จ บุญเรืองรัตน์, 2528 : 65)

#### ค่าอำนาจจำแนกของข้อสอบ

ฮู (T. Hsu, 1971 อ้างถึงใน สมบุญ ภูวนล, 2525 : 327) ได้กล่าวถึงข้อสอบวัดผล แบบอิงเกณฑ์ที่ดีว่า ควรมีอำนาจจำแนกสูง ระหว่างหลังจากเรียนแล้ว กับก่อนการเรียน กล่าวคือ การสอบก่อนเรียน นักเรียนจะตอบข้อสอบได้ไม่มาก แต่การสอบหลังจากได้เรียนแล้ว นักเรียนตอบข้อสอบได้มาก (การสอบก่อนและหลังเรียนใช้แบบสอบฉบับเดียวกัน และสอบกับ นักเรียนกลุ่มเดียวกัน)

วิธีการประมาณค่าอำนาจจำแนกมีผู้เสนอไว้ดังนี้  
 วิธีของเครเฮน (Crehan, 1974 : 257) โดยใช้สูตรดังนี้

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

- เมื่อ C เป็นดัชนีความตรงของแบบสอบ  
 U เป็นจำนวนผู้ตอบข้อสอบถูกในกลุ่ม  $N_1$   
 L เป็นจำนวนผู้ตอบข้อสอบถูกในกลุ่ม  $N_2$   
 $N_1$  เป็นจำนวนผู้สอบในกลุ่มที่ได้รับการสอน  
 $N_2$  เป็นจำนวนผู้สอบที่ไม่ได้รับการสอน

วิธีของคาร์เวอร์ (Carver, 1970 อ้างถึงใน บุญเว็ด วิทยุอนันตพงษ์, 2527 : 86)  
 ได้เสนอสูตรโดยอาศัยความสัมพันธ์ของการนิยามกลุ่ม ผู้รอบรู้-ไม่รอบรู้ จากกลุ่มผู้เรียน  
 ที่ยังไม่ได้รับการสอน และได้รับการสอนแล้ว มีสูตรการคำนวณดังนี้

$$\text{ดัชนีความตรง} = (a+c)/(a+b+c+d)$$

เมื่อ a,b,c, และ d เป็นจำนวนความถี่ร่วมซึ่งนิยามดังนี้  
 กลุ่มนักเรียน

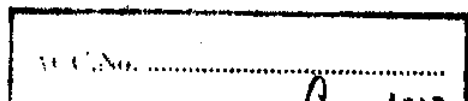
	ไม่ได้รับการสอน	ได้รับการสอน
ตอบถูก	b	a
ตอบผิด	c	d

การหาดัชนีอำนาจจำแนกปี ตามสูตรของเบรนแนน (Brennan)

เบรนแนน (Brennan, 1972 อ้างถึงใน สำเร็จ บุญเรืองรัตน์, 2528 : 65)

ได้หาค่าอำนาจจำแนกของข้อสอบแบบอิงเกณฑ์โดยใช้หลักการจำแนกกลุ่มสูง-กลุ่มต่ำ โดยใช้  
 การสอบเพียงครั้งเดียวกับกลุ่มตัวอย่างกลุ่มเดียว โดยใช้สูตร

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



- เมื่อ  $B$  คือ จำนวนจำแนกของข้อสอบ  
 $U$  คือ จำนวนคนที่ทำข้อสอบข้อนั้นถูกของกลุ่มที่ผ่านเกณฑ์  
 $L$  คือ จำนวนคนที่ทำข้อนั้นถูกของกลุ่มที่ไม่ผ่านเกณฑ์  
 $N_1$  คือ จำนวนคนที่สอบผ่านเกณฑ์  
 $N_2$  คือ จำนวนคนที่สอบไม่ผ่านเกณฑ์

การหาค่าดัชนี  $S$  (Sensitivity Index)

คริสปิน และเฟลด์ฮูเซน (Kryspin and Feldhusen, 1974 อ้างถึงใน สมศักดิ์ สินธุระเวชญ์, 2521 : 83-85) เสนอดัชนี  $S$  เรียกว่า Sensitivity Index เพื่อวิเคราะห์ข้อสอบอิงเกณฑ์ตามสูตรดังนี้

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

- เมื่อ  $R_{pos}$  คือ จำนวนนักเรียนที่ตอบถูกหลังการสอนผ่านไป  
 $R_{pre}$  คือ จำนวนนักเรียนที่ตอบถูกก่อนการสอน  
 $T$  คือ จำนวนนักเรียนที่ตอบข้อสอบทั้งสองครั้ง

ข้อสอบอิงเกณฑ์ที่ต้องการควรมีค่า  $S$  อยู่ระหว่าง .00 - 1.00 ถ้าค่า  $S$  ติดลบ แสดงว่าข้อสอบข้อนั้นง่ายมากจนทุกคนทำถูกก่อนการสอน หรือไม่ก็ข้อสอบยากมากจนนักเรียนใช้วิธีการเดาทั้งก่อนสอนและหลังสอน หรือไม่ก็การสอนของครูสับสนสร้างความรู้ความเข้าใจให้นักเรียนผิด และถ้าค่า  $S$  มีค่าเป็นบวกสูงแสดงว่าข้อสอบข้อนั้นสามารถจำแนกการสอนได้อย่างมีประสิทธิภาพ

การหาค่าความตรงเชิงโครงสร้างทุกวิธีจะอาศัยการแปรเปลี่ยนหรือการกระจายของคะแนนจากกลุ่มรอบรู้-ไม่รอบรู้มาเปรียบเทียบกัน ค่าสถิติอำนาจจำแนกของข้อสอบก็คือผลต่างของความยากของข้อสอบก่อนการสอนและหลังสอน การวิเคราะห์ความยากของข้อสอบจึงไม่จำเป็นจะต้องนำมาวิเคราะห์แยกต่างหากเหมือนการวิเคราะห์แบบอิงกลุ่ม (บุญเชิด ภิญญอนันตพงษ์, 2527 : 82) ดังนั้นผู้วิจัยจึงหาค่าความตรงเชิงโครงสร้างของข้อสอบด้วยการหาค่าอำนาจจำแนกเพียงอย่างเดียว โดยใช้วิธีหาค่าดัชนี  $S$  (Sensitivity Index) เนื่องจากเป็นวิธีที่ไม่ยุ่งยากในการคำนวณและการแปลความหมาย

### การคัดเลือกข้อสอบ

การคัดเลือกข้อสอบรวบรวมฉบับเป็นการพิจารณาตัดสินว่าข้อสอบใดบ้างอยู่ในเกณฑ์คุณภาพด้านความตรงเชิงเนื้อหาและความตรงเชิงโครงสร้าง เพื่อรวบรวมข้อสอบให้เป็นฉบับการเลือกข้อสอบอิงเกณฑ์มารวบรวมใช้เกณฑ์คุณภาพความตรงเชิงเนื้อหา 50% - 100% และความตรงเชิงโครงสร้างตั้งแต่ 20%-100% (บุญเชิด ภิฏญโญนนตพงษ์, 2527 : 43)

### การกำหนดความยาวและคะแนนจุดตัดของแบบสอบ

เมื่อคัดเลือกข้อสอบที่มีคุณภาพได้จำนวนมากพอสมควรแล้ว จึงนำมารวบรวมฉบับซึ่งจะต้องพิจารณาตัดสินว่าจะใช้ข้อสอบจำนวนกี่ข้อและจะใช้จุดตัดเป็นเท่าไร ขึ้นอยู่กับผู้สร้างจะใช้วิธีการใดในการตัดสินใจ

สำหรับการกำหนดความยาวของแบบสอบอิงเกณฑ์นั้น เบอร์ก (Berk, 1980 : 36-37) ได้สำรวจและรวบรวมผลงานวิจัยที่เกี่ยวข้องและได้สรุปถึงองค์ประกอบที่สำคัญในการกำหนดความยาวของแบบสอบอิงเกณฑ์ว่ามี 4 องค์ประกอบ คือ

- 1) ประเภทของการตัดสินผลการสอบ
- 2) ความสำคัญของจุดประสงค์ที่ใช้สอนตลอดจนการเน้นหนักในการสอน
- 3) จำนวนจุดประสงค์
- 4) ความจำเป็นในทางปฏิบัติ

ซึ่งองค์ประกอบทั้ง 4 นี้สามารถนำมาใช้เป็นเกณฑ์การกำหนดจำนวนข้อสอบของแบบสอบอิงเกณฑ์ได้ และจากการศึกษาถึงความคลาดเคลื่อนประเภทที่ 1 (Type I Error) คือ ผู้รอบรู้แต่ถูกตัดสินให้เป็นผู้ไม่รอบรู้ (False Non-mastery) และความคลาดเคลื่อนประเภทที่ 2 (Type II Error) คือผู้ไม่รอบรู้แต่ถูกตัดสินให้เป็นผู้รอบรู้ (False Mastery) ทำให้เบอร์ก (Berk) สรุปว่าถ้าใช้ ข้อสอบจำนวน 15 ข้อ ในแต่ละจุดประสงค์ทำการสอบและส่งผลย้อนกลับทันทีที่สอบเสร็จ จะทำให้การตัดสินมีความถูกต้องแม่นยำ แต่ถ้าใช้ 2-3 ข้อ การตัดสินจะเกิดความคลาดเคลื่อนมาก ดังนั้นเบอร์กจึงได้เสนอให้กำหนดจำนวนข้อสอบสำหรับการประเมินความก้าวหน้า (Formative Evaluation) หรือการประเมินเพื่อจัดตำแหน่ง (Placement) ระหว่าง 5-10 ข้อ แต่สำหรับการประเมินผลสรุปรวม (Summative Evaluation) หรือการตัดสินแบบถาวร (Permanent Decision) จำเป็นจะต้องใช้ข้อสอบมากขึ้นเป็นสองเท่า คือประมาณ 10-20 ข้อ

## การวิเคราะห์แบบสอบ

การวิเคราะห์แบบสอบ เป็นการประเมินคุณภาพด้านความเที่ยงและความตรงของ แบบสอบทั้งฉบับ เมื่อตรวจสอบได้ค่าคุณภาพของแบบสอบในระดับที่พอใจ ก็นำแบบสอบไปจัดพิมพ์เป็นรูปเล่ม

### ความเที่ยงและความตรงของแบบสอบอิงเกณฑ์

#### 1. ความเที่ยงของแบบสอบอิงเกณฑ์

ความเที่ยง (Reliability) เป็นคุณภาพของแบบสอบทั้งฉบับ เมื่อแบบสอบฉบับนั้นมีคุณภาพเป็นรายชื่ออยู่ในเกณฑ์แล้ว ความเที่ยงเป็นดัชนีที่บ่งชี้ว่าแบบสอบนั้นมีคุณภาพหรือไม่ในการสร้างแบบสอบทุกครั้งจึงจำเป็นต้องตรวจหาความเที่ยงของแบบสอบ

แฮมเบิลตันและคณะ (Hambleton et al., 1798 : 15 - 23 quoted in Berk, 1980 : 323-346) ได้แบ่งกลุ่มวิธีการหาความเที่ยงออกเป็น 3 กลุ่ม พร้อมทั้งนิยามความหมายของแต่ละกลุ่มไว้ดังนี้

1) การประมาณค่าความเที่ยงของแบบสอบ โดยพิจารณาจากความคงที่ในการตัดสินใจจำแนกความรอบรู้ (Reliability of Mastery Classification Decision) ซึ่งเป็นการตรวจหาสัมประสิทธิ์ความสอดคล้องในการตัดสินใจ ความรอบรู้-ไม่รอบรู้จากการสอบซ้ำด้วยแบบสอบฟอร์มเดียว หรือฟอร์มคู่ขนาน ซึ่งได้แก่วิธีต่าง ๆ ดังนี้

1.1) วิธีของคาร์เวอร์ (Carver, 1970)

1.2) วิธีของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973)

1.3) วิธีของสวามินาทาน, แฮมเบิลตัน และอัลจิโน (Swamiratan, Hambleton and Algina, 1974)

1.4) วิธีของสับโคเวียก (Subkoviak, 1976 )

1.5) วิธีของฮวีน (Huynh, 1976)

2) การประมาณค่าความเที่ยงของแบบสอบ โดยพิจารณาจากความเที่ยงของคะแนนแบบสอบอิงเกณฑ์ (Reliability of Criterion-Referenced Test Score) ซึ่งเป็นวิธีการหาสัมประสิทธิ์ความสอดคล้องของคะแนนแต่ละคนที่เขียนเบนไปจากคะแนนจุดตัดจากการสอบด้วยแบบสอบคู่ขนานสองฟอร์ม โดยคำนึงถึงความแปรปรวนของความคลาดเคลื่อน ซึ่งได้จากการวิเคราะห์องค์ประกอบของความแปรปรวน ซึ่งแบ่งเป็น 2 ชนิด คือ

2.1) ความแปรปรวนที่เกิดจากปฏิสัมพันธ์ระหว่างบุคคลกับข้อสอบ (Person X Item Interaction) ซึ่งทำให้เกิดมีวิธีหาความเที่ยงของลิวิงตัน (Livingston)

2.2) ความแปรปรวนที่เกิดจาก ผลของข้อสอบรวมกับปฏิสัมพันธ์ระหว่างบุคคลกับข้อสอบ (Item Plus Person X Item Interaction) ซึ่งทำให้เกิดมีวิธีหาความเชื่อมั่นของ เบรนนอนและเคน (Brennan and Kane)

3) การประมาณค่าความเที่ยงของแบบสอบ โดยพิจารณาจากความเที่ยงของการประมาณค่าคะแนนมวลความรู้ (Reliability of Domain Score Estimates) เป็นการประมาณค่าความสอดคล้อง ของคะแนนผู้สอบแต่ละคนจากแบบสอบคู่ขนานสองฟอร์ม การประมาณค่าความเที่ยงของแบบสอบด้วยวิธีนี้ ไม่ต้องใช้คะแนนจุดตัดในการจำแนกความรอบรู้ แต่จะพิจารณาจากความคลาดเคลื่อนมาตรฐาน 4 แบบ

3.1) ความคลาดเคลื่อนมาตรฐานมี 4 แบบ ดังนี้ คือ

3.1.1) ความคลาดเคลื่อนมาตรฐานของสัดส่วนการตอบถูกแต่ละคน (E(P))

3.1.2) ความคลาดเคลื่อนมาตรฐานในการวัดของผู้สอบแต่ละคน (S.E.(X<sub>ij</sub>))

3.1.3) ความคลาดเคลื่อนมาตรฐานในการวัดเฉลี่ยทั้งกลุ่มจากความคลาดเคลื่อนมาตรฐานของสัดส่วนการตอบถูก ( $\sigma(E)$ )

3.1.4) ความคลาดเคลื่อนมาตรฐานในการวัดเฉลี่ยทั้งกลุ่มจากความคลาดเคลื่อนมาตรฐานในการวัดของผู้สอบแต่ละคน ( $\sigma(\Delta)$ )

3.2) การประมาณค่าความเที่ยงตามรูปแบบนี้ มีวิธีการอยู่ 4 วิธี คือ

3.2.1) วิธีของเบอร์ก (Berk, 1980)

3.2.2) วิธีของลอร์ด (Lord, 1959)

3.2.3) วิธีของลอร์ดและโนวิก (Lord and Novick, 1968)

3.2.4) วิธีของเบรนนอน (Brennan, 1980)

จากนิยามของความหมาย และวิธีการในการคำนวณทั้ง 3 กลุ่ม การประมาณค่าความเที่ยงโดยพิจารณาจากความคงที่ในการตัดสินใจจำแนกความรอบรู้เป็นนิยามที่สอดคล้องกับแนวคิดหรือหลักของการวัดแบบอิงเกณฑ์ ที่ว่า การวัดผลแบบอิงเกณฑ์เป็นการวัดเพื่อมุ่งประเมินความสามารถของบุคคลโดยเทียบกับเกณฑ์ที่กำหนดขึ้นเพื่อบอกให้ทราบว่าผู้ใดรอบรู้หรือไม่รอบรู้ ดังนั้นจึงเลือกเอาวิธีการประมาณค่าความเที่ยงในลักษณะข้อ 1 ในการหาความเที่ยงของแบบสอบอิงเกณฑ์ที่ผู้วิจัยสร้างขึ้น และขอเสนอการประมาณค่าความเที่ยงตามนิยามในกลุ่มนี้เท่านั้น

### 1.1 การหาค่าความเที่ยงตามวิธีของคาร์เวอร์ (Carver)

คาร์เวอร์ (Carver, 1970 quoted in Crehan, 1974 : 256) ได้เสนอสูตรในการหาค่าความเที่ยงของแบบสอบอิงเกณฑ์โดย นำแบบสอบคู่ขนาน 2 ฉบับ ไปสอบกับนักเรียนกลุ่มเดียวกันแล้วนำผลการสอบไปแจกแจงลงในตารางดังนี้

ฟอร์ม ข.

		ไม่ผ่าน	ผ่าน
ฟอร์ม ก.	ผ่าน	b	a
	ไม่ผ่าน	c	d

$$\hat{P}_o = (a + c)/(a+b+c+d)$$

เมื่อ  $\hat{P}_o$  คือ สัดส่วนในความสอดคล้องของการตัดสินใจรอบรู้

### 1.2 การหาความเที่ยงตามวิธีของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973)

การหาความเที่ยงตามวิธีของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973 : 159-170) เป็นวิธีที่นำผลการสอบ ที่สอบด้วยฟอร์มเดียวกันสองครั้ง หรือฟอร์มคู่ขนานสองฉบับ มาแจกแจง และหาสัดส่วนความสอดคล้องในการตัดสินใจรอบรู้จากผลการสอบทั้งสองครั้ง หรือสองฉบับดังนี้

ฟอร์ม ข. (หรือสอบครั้งที่ 2)

		ผ่าน	ไม่ผ่าน
ฟอร์ม ก. (สอบครั้งที่ 1)	ผ่าน	$P_{11}$	
	ไม่ผ่าน		$P_{22}$

สัดส่วนของความสอดคล้องในการตัดสินใจคือ สัดส่วนของการตัดสินใจว่า รอบรู้ตรงกัน ทั้งสองฉบับ หรือสองครั้ง และสัดส่วนของการตัดสินใจว่า ไม่รอบรู้ตรงกันทั้งสองฉบับหรือสองครั้ง ซึ่งสามารถเขียนเป็นสูตรได้ดังนี้

$$P_o = \sum_{i=1}^2 P_{ii} = P_{11} + P_{22}$$



- เมื่อ  $P_o$  แทน สัดส่วนความสอดคล้องในการตัดสินใจจำแนกผู้รอบรู้  
 $P_{11}$  แทน สัดส่วนของผู้ถูกตัดสินว่ารอบรู้ตรงกันทั้งสองฉบับ หรือ 2 ครั้ง  
 $P_{22}$  แทน สัดส่วนของผู้ถูกตัดสินว่าไม่รอบรู้ตรงกันทั้งสองฉบับหรือ 2 ครั้ง

1.3 การหาความเที่ยงตามวิธีของสวามินาทาน, แฮมเบิลตัน และอัลจิงนา (Swaminatan, Hambleton and Algina, 1974 )

สวามินาทาน, แฮมเบิลตัน และอัลจิงนา (Swaminatan, Hambleton and Algina, 1974 : 263 - 267) ได้เสนอสูตรสำหรับการประมาณค่าความเที่ยงของแบบสอบอิงเกณฑ์ โดยใช้แนวคิดของแฮมเบิลตันและโนวิก (Hambleton and Novick) แต่มีการปรับแก้โดยการหักเอาค่าความสอดคล้องโดยบังเอิญออก ดังนี้

$$K = [P_o - P_c] / [1 - P_c]$$

เมื่อ  $K$  แทน ค่าความเที่ยงของแบบสอบอิงเกณฑ์ซึ่งหักความสอดคล้องโดยบังเอิญออกแล้ว

$P_o$  แทน สัดส่วนความสอดคล้องในการตัดสินว่ารอบรู้จากการสอบซ้ำ

$P_c$  แทน สัดส่วนความสอดคล้องที่คาดหวังซึ่งอาจเกิดขึ้นโดยบังเอิญ

$$P_o = \sum_{i=1}^2 P_{ii}$$

$$P_c = \sum_{i,j=1}^2 P_i \cdot P_j$$

ในการคำนวณจะนำข้อมูลที่นำมาแจกแจงลงในตารางดังต่อไปนี้

สอบครั้งที่ 2 (ฟอร์ม ข.)

		รอบรู้	ไม่รอบรู้	สัดส่วนแยก
สอบครั้งที่ 1	ผ่าน	$P_{11}$	$P_{12}$	$P_{1.}$
(ฟอร์ม ก.)	ไม่ผ่าน	$P_{21}$	$P_{22}$	$P_{2.}$
	สัดส่วนแยก	$P_{.1}$	$P_{.2}$	

#### 1.4 การหาความเที่ยงตามวิธีของสับโคเวียก (Subkoviak, 1976)

สับโคเวียก (Subkoviak, 1976 : 265-276) ได้เสนอวิธีการประมาณค่าความเที่ยงของแบบสอบอิงเกณฑ์จากการสอบเพียงครั้งเดียว โดยอาศัยหลักการหาสัมประสิทธิ์ความสอดคล้องในการกำหนดความรู้ของนักเรียนแต่ละคนและของกลุ่มที่ได้จากการสอบคู่ขนาน 2 ฉบับ ซึ่งมีคะแนนเป็น  $X$  และ  $X'$  แต่ค่าสถิติต่าง ๆ ที่ต้องคำนวณมาจากคะแนน  $X'$  นั้น สับโคเวียกได้ดัดแปลงสูตรให้สามารถคำนวณมาจากคะแนน  $X$  จึงสามารถหาค่าความเที่ยงได้จากการสอบเพียงครั้งเดียว โดยอาศัยแบบจำลองทางคณิตศาสตร์ดังนี้

1) สัมประสิทธิ์ความสอดคล้องของคนที่  $i$  คือ ความน่าจะเป็นของคนที่ถูกกำหนดให้เป็นผู้รอบรู้ เหมือนกันจากแบบสอบคู่ขนาน  $X$  และ  $X'$  และเมื่อนิยามคะแนนเกณฑ์  $c$  ก็จะระบุความรู้ของผู้สอบเป็น 2 กลุ่ม ( $k = 2$ ) แต่ถ้านิยามคะแนนเกณฑ์หลายค่า เป็น  $c_1, c_2, c_3, \dots, c_{k-1}$  ก็จะระบุความรู้ของผู้สอบออกมากกว่า 2 กลุ่ม ( $k > 2$ )

1.1) การกำหนดให้ผู้สอบมีความรู้สอดคล้องกัน ระหว่างแบบสอบ  $X$  และ  $X'$  สามารถทำได้ 2 วิธี โดยใช้ คะแนน  $c$  เป็นเกณฑ์ดังนี้

1.1.1)  $X_i \geq c$  และ  $X'_i \geq c$  แสดงความสอดคล้องในการตัดสินใจว่ารอบรู้/รอบรู้ เหมือนกันจากแบบสอบ 2 ฉบับ

1.1.2)  $X_i < c$  และ  $X'_i < c$  แสดงความสอดคล้องในการตัดสินใจว่าไม่รอบรู้/ไม่รอบรู้ เหมือนกันจากแบบสอบ 2 ฉบับ

1.2) ในทำนองเดียวกันความไม่สอดคล้องที่เกิดจากการตัดสินใจความรู้ อาจแสดงได้ 2 ทาง ดังนี้

1.2.1)  $X_i \geq c$  แต่  $X'_i < c$

1.2.2)  $X_i < c$  แต่  $X'_i \geq c$

ดังนั้นสัมประสิทธิ์ของความสอดคล้อง  $P_o^{(i)}$  สำหรับคนที่  $i$  เมื่อให้คะแนนเทียบกับ  $c$  จึงเขียนสมการได้ดังนี้

$$P_o^{(i)} = P(X_i \geq c, X'_i \geq c) + P(X_i < c, X'_i < c)$$

เมื่อ  $P(X_i \geq c, X'_i \geq c)$  เป็นความน่าจะเป็นของการตัดสินใจว่า รอบรู้/รอบรู้

$P(X_i < c, X'_i < c)$  เป็นความน่าจะเป็นของการตัดสินใจว่าไม่รอบรู้/ไม่รอบรู้

2) สัมประสิทธิ์ความสอดคล้อง สำหรับกลุ่ม  $P_o$  จำนวน  $N$  คน คือค่าเฉลี่ยของสัมประสิทธิ์ของความสอดคล้องรายบุคคล เมื่อกำหนดคะแนนเกณฑ์เท่ากับ  $c$  เขียนเป็นสมการได้ดังนี้

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

เมื่อ  $P_c$  แทน ค่าความน่าจะเป็นของการตัดสินใจสอดคล้องกันทั้งกลุ่ม

$\sum_{i=1}^N p_c^{(i)}$  แทน ผลรวมของความน่าจะเป็นของการตัดสินใจที่สอดคล้องกัน

สำหรับแต่ละบุคคลซึ่งถ่วงน้ำหนักโดยค่าความน่าจะเป็นที่เกิดขึ้นในกลุ่ม

2.1) ข้อตกลงในการประมาณค่าสัมประสิทธิ์ของแต่ละบุคคลและของกลุ่มมี

2 ข้อ คือ

2.1.1) คะแนน  $X_i$  และ  $X'_i$  ของแต่ละบุคคลที่  $i$  ต้องมีการแจกแจง

เป็นอิสระแสดงว่าประสบการณ์จากแบบสอบ  $X$  จะไม่ส่งผลต่อการสอบด้วยแบบสอบ และ  $X'$  ซึ่งสามารถเขียนสมการใหม่ได้ดังนี้

$$P_c^{(i)} = P(X_i \geq c) P(X'_i \geq c) + P(X_i < c) P(X'_i < c)$$

2.1.2) การแจกแจงของคะแนน  $X$  และ  $X'$  สำหรับแต่ละบุคคลที่กำหนด

ไว้จะต้องแจกแจงแบบทวินาม ซึ่งข้อสอบแต่ละข้อมีการให้คะแนนเป็นแบบ ถูกได้ 1 ผิดได้ 0 ผลการตอบแต่ละข้อต้องไม่ส่งผลต่อกัน และแต่ละข้อจะต้องมีโอกาสในการตอบถูกเหมือนกัน ซึ่งทำให้เขียนสมการใหม่ได้ดังนี้

$$\begin{aligned} P_c^{(i)} &= [P(X_i \geq c)]^2 + [P(X_i < c)]^2 \\ &= [P(X_i \geq c)]^2 + [1 - P(X_i \geq c)]^2 \end{aligned}$$

$$\text{เมื่อ } P(X_i \geq c) = \sum_{X_i} \binom{n}{X_i} P_i^{X_i} (1 - p_i)^{n - X_i}$$

ค่า  $P_i$  แทน ค่าความน่าจะเป็นที่แท้จริงของบุคคล  $i$  ที่ตอบข้อสอบถูกซึ่งสามารถประมาณมาจากคะแนนที่สอบได้  $X_i$  จากแบบสอบฉบับเดียว มีวิธีการหาดังนี้

2.1.2.1) ในกรณีที่มีจำนวนข้อมาก ( $n > 40$ ) หากค่า  $P_i$  ได้ดังนี้

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

เมื่อ  $X_i$  คือ จำนวนข้อที่ตอบถูก

$n$  คือ จำนวนข้อสอบทั้งหมดในแบบสอบ

2.1.2.2) ในกรณีที่มีจำนวนข้อน้อยและมีการแจกแจงแบบ

ฐานนิยมเดียว หาค่า  $P_i$  ได้ดังนี้

$$\hat{P}_i = \alpha_{21/x} [X_i/n] + [1-\alpha_{21/x}][M_x/n]$$

เมื่อ	$\alpha_{21/x}$	แทน	ค่าความเที่ยงแบบคูเดอริชาร์ดสัน สูตร 21
	$M_x$	แทน	ค่าเฉลี่ย
	$n$	แทน	จำนวนข้อทั้งหมดของแบบสอบ

1.5 การหาความเที่ยงตามวิธีของฮวิน (Huynh, 1976)

ฮวิน (Huynh, 1976 : 253-262) ได้เสนอวิธีการประมาณค่าความเที่ยงในรูปของสัมประสิทธิ์แคปป่า ( $K$ ) โดยอาศัยแนวคิดของการสอบซ้ำด้วยแบบสอบคู่ขนาน 2 ฟอร์ม ประมาณค่าความเที่ยงจากการสอบเพียงครั้งเดียว โดยมีข้อสอบภายในคู่ขนานกันแบบสุ่ม มีวิธีการคำนวณดังนี้

เมื่อแบบสอบมีจำนวนข้อน้อย

1) คำนวณค่า  $\bar{X}$ ,  $S^2$  และ  $\alpha_{21}$  จากคะแนนที่สอบได้

$$\text{โดย } \alpha_{21} = \frac{n}{n-1} \left[ 1 - \frac{\mu(n-\mu)}{n\sigma^2} \right]$$

2) ประมาณค่า  $\alpha$  และ  $\beta$  จากสูตร

$$\alpha = [-1 + 1/\alpha_{21}] \mu$$

$$\beta = -\alpha + \alpha_{21} - n$$

3) คำนวณค่า  $K$  โดยแยกเป็น 2 กรณี คือ

$$K = [P_{11} - P_1^2] / [P_1 - P_1^2]$$

เมื่อคะแนนเกณฑ์  $C$  มีค่าน้อยหรือใกล้ๆ 0

$$K = [p_{00} - P_0^2] / [P_0 - P_0^2]$$

เมื่อคะแนนเกณฑ์  $C$  มีค่ามากใกล้ๆ  $n$

$$\text{เมื่อ } P_0 = f(0) + \dots + f(c-1)$$

$$P_{00} = f(0,0) + f(0,1) + \dots + f(c-1, c-1)$$

$$P_{11} = \sum_{X,Y=C}^n f(X,Y)$$

$$= f(n,n) + f(n,n-1) + f(n-1,n) + \dots + f(c,c)$$

$$\begin{aligned}
P_1 &= f(n)+\dots+f(c) \\
f(0,0) &= \prod_{i=1}^{2n} \frac{2n+\beta-i}{2n+\alpha+\beta-i} = f(0) \cdot \prod_{i=1}^n \frac{2n+\beta+i}{2n+\alpha+\beta+i} \\
f(X+1,Y) &= f(X,Y) \cdot \frac{(n-X)(\alpha+X+Y)}{(X+1)(2n+\beta-X-Y-1)} \\
f(0) &= \prod_{i=1}^n \frac{n+\beta-i}{n+\alpha+\beta-i} \\
f(X+1) &= f(X) \cdot \frac{X(n+\beta-X)}{(n-X+1)(\alpha+X-1)}, X=1, \dots, n-1 \\
f(n) &= \prod_{i=1}^n \frac{n+\alpha-i}{n+\alpha+\beta-1} \\
f(n,n) &= \prod_{i=1}^{2n} \frac{2n+\alpha-i}{2n+\alpha+\beta-i} = f(n) \cdot \prod_{i=1}^n \frac{2n+\alpha-i}{2n+\alpha+\beta-i} \\
f(X-1) &= f(X) \cdot \frac{X(n+\beta-X)}{(n-X+1)(\alpha+X-1)}, X=1, \dots, n. \\
f(X-1,Y) &= f(X,Y) \cdot \frac{X(2n+\beta-i)}{(n-X+1)(\alpha+X+Y-1)}
\end{aligned}$$

สำหรับการวิจัยนี้ ผู้วิจัยได้ใช้วิธีการประมาณค่าความเที่ยงของแบบสอบ

โดยวิธีของสวามินาทาน, แฮมเบิลตัน และอัลจินา เนื่องจาก เป็นวิธีที่คำนวณง่าย และจากการศึกษาของสับโคเวียค (Subkoviak, 1978 : 111-115 อ้างถึงใน พานิช ศรีงาม, 2530 : 19) ซึ่งได้เปรียบเทียบวิธีการประมาณค่าความเที่ยงของแบบสอบอิงเกณฑ์ในรูปของความสอดคล้องในการตัดสินใจจำแนกผู้รอบรู้ 4 วิธี 1) วิธีของสวามินาทาน, แฮมเบิลตัน และอัลจินา (Swaminathan, hambleton and Algina Procedure) 2) วิธีของมาร์แชล และแฮร์เทล (Marshall-Harctel Procedure) 3) วิธีของสับโคเวียค (SubKoviak Procedure) 4) วิธีของฮวีน (Huynh procedure) พบว่าวิธีของสวามินาทาน, แฮมเบิลตัน และอัลจินา ให้ผลการประมาณที่ไม่ลำเอียง

## 2. ความตรงของแบบสอบอิงเกณฑ์

ความตรงของแบบสอบอิงเกณฑ์เป็น คุณภาพของแบบสอบ ในการที่จะจำแนกผู้ที่เรียนแล้วและผู้ที่ยังไม่ได้เรียนออกจากกัน นั่นคือนักเรียนที่เรียนแล้วจะต้องทำข้อสอบอิงเกณฑ์ได้มากกว่านักเรียนที่ยังไม่ได้เรียนที่ หรือนักเรียนหลังเรียนจะต้องทำข้อสอบได้มากกว่าก่อนเรียน วิธีการหาค่าความตรงของแบบสอบอิงเกณฑ์ ได้มีผู้เสนอวิธีไว้ดังนี้

2.1 คาร์เวอร์ (Caver, 1970 quoted in Crehan, 1974 : 256) ได้เสนอการคำนวณค่าความตรงของแบบสอบอิงเกณฑ์จากข้อมูลของกลุ่มผู้เรียนแล้วและยังไม่ได้เรียนดังนี้

	กลุ่มที่ไม่ได้รับการสอน	กลุ่มที่ได้รับการสอน
ผ่านเกณฑ์	b	a
ไม่ผ่านเกณฑ์	c	d

$$\Gamma_{ck} = (a + c) / (a + b + c + d)$$

เมื่อ  $\Gamma_{ck}$  แทน ค่าดัชนีความตรงของแบบสอบอิงเกณฑ์

คาร์เวอร์ (Carver) กล่าวว่า ค่าความตรงจะมากหรือน้อยขึ้นอยู่กับจำนวนคนที่สอบไม่ผ่านก่อนเรียน (c) และคนที่สอบผ่านหลังเรียน (a) ดังนั้นค่าความตรงนี้จะขึ้นอยู่กับประสิทธิภาพการสอนเป็นสำคัญ

2.2 มิลล์แมน (Millman, 1974 อ้างถึงใน เอนก เพียรอนุกุลบุตร, 2527 : 565) ได้เสนอวิธีการหาค่าความตรง ด้วยการคำนวณค่าสหสัมพันธ์แบบพี ( $\Phi$ ) มีสูตรดังนี้

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

เมื่อ a	คือ	จำนวนผู้สอบผ่านหลังเรียน
b	คือ	จำนวนผู้สอบผ่านก่อนเรียน
c	คือ	จำนวนผู้สอบไม่ผ่านหลังเรียน
d	คือ	จำนวนผู้สอบไม่ผ่านก่อนเรียน

สำหรับการวิจัยนี้ผู้วิจัยได้ใช้วิธีของคาร์เวอร์ (Carver) ในการหาค่าความตรงของแบบสอบเพราะเป็นวิธีที่สอดคล้องกับการเรียนการสอน คือ หลังเรียนนักเรียนควรที่จะสอบผ่านมากกว่าก่อนเรียน หรือ หลังเรียนนักเรียนควรทำข้อสอบได้มากกว่าก่อนเรียน

### การกำหนดคะแนนจุดตัดของแบบสอบอิงเกณฑ์

คำว่าคะแนนจุดตัด บางครั้งเรียกว่า เกณฑ์ (Criteria) คะแนนมาตรฐาน (Standard) คะแนนจุดผ่าน (Passing Score) ระดับความรอบรู้ (Mastery Level) ความสามารถขั้นต่ำสุด (Minimal Competence) หรือคะแนนจุดแบ่งของแบบสอบอิงเกณฑ์ ซึ่งความหมายของคำเหล่านี้ทั้งหมดก็คือ คะแนนที่น้อยที่สุดที่นักเรียนจะต้องทำได้ในการที่จะได้รับการตัดสินให้เป็นผู้รอบรู้

ในการประเมินผลแบบอิงเกณฑ์ การกำหนดคะแนนเกณฑ์หรือคะแนนจุดตัดเป็นสิ่งสำคัญมาก ถ้ากำหนดคะแนนจุดตัดต่ำไปอาจจะเกิดปัญหาในกลุ่มที่สอบผ่านแต่ยังไม่รอบรู้ แต่ถ้ากำหนดคะแนนจุดตัดสูงเกินไปก็จะเกิดปัญหามีผู้รอบรู้แต่สอบไม่ผ่าน ซึ่งจะทำให้เกิดความคลาดเคลื่อนจำแนกเป็น 2 ประเภท คือ (อังกฤษ สายยศ, 2525 : 70)

- 1) เกิดกลุ่มที่สอบไม่ผ่านแต่รอบรู้ เป็นการผิดพลาดที่เรียกว่า  $\beta$ -error ซึ่งเกิดจากการกำหนดคะแนนจุดตัดสูงเกินไป ทำให้เด็กสอบตกมาก ทำให้เสียเวลาและกำลังใจ สิ้นเปลืองแรงงานในการที่จะต้องสอบซ้ำ
- 2) เกิดกลุ่มที่สอบผ่านแต่ไม่รู้จักจริง เป็นการผิดพลาดที่เรียกว่า  $\alpha$ -error ซึ่งเกิดจากการที่กำหนดคะแนนจุดตัดต่ำเกินไป ทำให้เด็กที่มีคุณภาพไม่มาตรฐานสอบผ่านได้อาจมีผลเสีย เช่น เด็กจะมีความสามารถไม่พอในการเรียนจุดมุ่งหมายใหม่ หรือเรียนในระดับสูงต่อไป หรืออาจเป็นพลเมืองที่ไม่มีคุณภาพดีพอ

## 1. ประโยชน์ของการกำหนดคะแนนจุดตัดของแบบสอบอิงเกณฑ์

สุพัฒน์ สุกมลสันต์ (2530 : 71) ได้กล่าวถึงประโยชน์ของการกำหนดคะแนนจุดตัดไว้ดังนี้

- 1) เป็นการควบคุมมาตรฐานของการเรียนการสอนให้เป็นไปตามหลักสูตรอย่างแท้จริง คือ ช่วยให้เราสามารถแยกการเรียนรู้ของนักเรียนออกเป็นผู้รอบรู้ และผู้ไม่รอบรู้ (Master and Non-Master) ออกจากกันได้โดยตรง
- 2) ช่วยในการตรวจสอบว่า เด็กคนใดมีความสามารถขั้นต้น (Prerequisite) ครบถ้วน และสามารถเรียนต่อไปได้ตามลำดับการเรียนรู้ (learning Hierarchy) ได้

## 2. วิธีการกำหนดคะแนนจุดตัดของแบบสอบอิงเกณฑ์

ในการกำหนดคะแนนจุดตัดหรือเกณฑ์ของแบบสอบอิงเกณฑ์ มีผู้เสนอให้หลายวิธี ดังนี้คือ

2.1 กำหนดคะแนนจุดตัดโดยใช้ผลที่ได้จากการปฏิบัติของคนอื่น ๆ เป็นเกณฑ์ (Performance of Others as a Criterion) คะแนนจุดตัดนี้สร้างขึ้นโดยอ้างอิงพารามิเตอร์ ของประชากรผู้เข้าสอบ ดังนั้น คะแนนจุดตัดอาจกำหนดขึ้นโดยใช้ค่ามัธยฐาน (Median) ของคะแนนจากแบบสอบ นักทฤษฎีเกี่ยวกับอิงเกณฑ์หลายท่านไม่เห็นด้วยกับวิธีนี้ เนื่องจากเป็นการอิงกลุ่มอยู่นั่นเอง

มิลส์แมน ได้เสนอวิธีการหาคะแนนจุดตัดจากผลการสอบของคนอื่นๆ ไว้ 2 วิธีคือ (บุญเจิด ภิญญอนันตพงษ์, 2527 : 118)

1) อาศัยหลักการกำหนดคะแนนจุดตัดที่สอดคล้องกับจำนวนเปอร์เซ็นต์ของผู้สอบผ่านเกณฑ์ ซึ่งระบุไว้ล่วงหน้า โดยอาศัยผู้เชี่ยวชาญเนื้อหาวิชา หรือครูประจำชั้นพิจารณาแบบสอบแล้วให้ระบุจำนวนเปอร์เซ็นต์ของผู้สอบผ่านว่าควรเป็นเท่าไรก่อน แล้วจึงหาคะแนนที่สอดคล้องกับเปอร์เซ็นต์นั้น

2) อาศัยหลักของการหาคะแนนดิบที่สอดคล้องกับตำแหน่งของเปอร์เซนไทล์ ที่เลือกมาจากผลการสอบของนักเรียนที่มีความรอบรู้ ตามเนื้อหาวิชาที่จะสอบมาก่อนแล้ว โดยให้ครูผู้สอนเลือกตำแหน่งเปอร์เซนไทล์ที่เหมาะสมขึ้นก่อน

## 2.2 กำหนดจุดตัดโดยนับถอยหลังจาก 100% (Counting Backwards from 100%)

วิธีนี้จะอาศัยหลักการกำหนดค่าเกณฑ์ หรือระดับมาตรฐานของการกำหนดมาตรฐานของการกำหนดมาตรฐานของจุดประสงค์ คือ กำหนดค่าเกณฑ์ตามความสำคัญของจุดประสงค์ ถ้าจุดประสงค์ใดมีความสำคัญมากค่าเกณฑ์ที่ต้องการ ต้องเป็น 100% ถ้าจุดประสงค์ใดมีความสำคัญน้อยกว่า ค่าเกณฑ์ที่ต้องการก็จะลดต่ำลงจาก 100% (สุพัฒน์ สุกมลสันต์, 2530 : 57)

2.3 กำหนดคะแนนจุดตัดโดยเพิ่มคะแนนเกณฑ์อื่น ๆ (Bootstrapping on other Criterion Scores) โดยอาศัยการระบุจากเกณฑ์ภายนอกกำหนดว่า "สำเร็จ" หรือ "รอบรู้" อย่างชัดเจน เช่น อาจใช้เกณฑ์ว่า ผู้ได้รับประกาศนียบัตรในอาชีพต่าง ๆ เช่น ตัดเสื้อผ้า ตัดผม นักสถิติ เป็นต้น ถือว่าเป็นผู้มีความสามารถ ส่วนผู้ไม่ได้รับประกาศนียบัตรถือเป็นผู้ไม่มีความสามารถ การหาคะแนนจุดตัดจะพิจารณาการกระจายของคะแนนที่ได้จากการทำแบบสอบของคนเหล่านี้ แล้ว ดูว่าคะแนนเกณฑ์ใดที่สามารถแยกผู้มีความสามารถจากผู้ไม่มีความสามารถ ได้สอดคล้องกับเกณฑ์ภายนอกมากที่สุด ข้อจำกัดอยู่ที่ว่าเกณฑ์ภายนอกที่จะนำมาใช้มีมาตรฐานเพียงพอหรือไม่ (สุพัฒน์ สุกมลสันต์, 2530 : 71)

2.4 การกำหนดคะแนนจุดตัดจากความคลาดเคลื่อนในการเดาตอบและการสุ่มข้อสอบ (บุญเจิด ภิญญอนันตพงษ์, 2527 : 133) วิธีการนี้คำนึงถึงทฤษฎี การทดสอบที่ว่า คะแนนสอบย่อมมีความคลาดเคลื่อนอยู่ด้วย ดังนั้นคะแนนจุดตัดจึงพิจารณาขอบเขตของความคลาดเคลื่อนประกอบด้วย หรือต้องปรับแก้คะแนนโดยกำจัดความคลาดเคลื่อนออกไปเสียก่อน ดังนั้นคะแนนจุดตัดที่กำหนดขึ้นมาจึงจำเป็นต้องปรับแก้ โดยบวกความคลาดเคลื่อนที่อาจเกิดขึ้นจากการเดาตอบถูกไว้ด้วย



## 2.5 การกำหนดคะแนนจุดตัดจากสมรรถภาพขั้นต่ำ

วิธีนี้จะให้ผู้เชี่ยวชาญศึกษาแบบสอบแล้วระบุว่านักเรียนที่มีสมรรถภาพขั้นต่ำจะต้องมีคะแนนเท่ากับเท่าไร ได้แก่เทคนิคของนีเดลสกี (Nedelsky, 1957) อีเบล (Ebel, 1972) และ แองกอฟ (Angoff, 1971) (บุญเชิด ภิญโญนนตพงษ์, 2527 : 123-129)

เทคนิคของนีเดลสกี (Nedelsky, 1957)

เทคนิคนี้อาศัยการพิจารณาว่า นักเรียนที่มีสมรรถภาพต่ำสุดต้องมีคะแนนสอบผ่านขั้นต่ำเป็นเท่าไร โดยอาศัยความน่าจะเป็นของการเดาตอบถูกจากแบบตอบเลือกตอบหลายตัวเลือก คำนวณจากสูตร

	D	=	$M_{FD} + KS_{FD}$
เมื่อ	D	แทน	คะแนนสอบผ่านขั้นต่ำ
	$M_{FD}$	แทน	ค่าเฉลี่ยคะแนนการเดาถูกหาจากผลรวมความน่าจะเป็นในการตอบถูก
	$S_{FD}$	แทน	ค่าเบี่ยงเบนมาตรฐานของคะแนนเดาตอบถูก
	K	แทน	ค่าคงที่ถ้าตัวเลือกผิดเด่นชัดจะใช้ค่า 0.5 ถ้าตัวเลือกแตกต่างกันเล็กน้อยจะใช้ค่า 1.0 ถ้าตัวเลือกคล้ายกันมากจะใช้ค่า 1.5

เทคนิคของอีเบล (Ebel, 1972)

มีวิธีการโดย ให้ครูประจำวิชาวิเคราะห์ข้อสอบเทียบกับเปอร์เซ็นต์การสอบผ่านที่คาดหวัง จากลักษณะของข้อสอบ และความยากของข้อสอบ ซึ่งอีเบลได้จำแนกไว้แล้วคือ ลักษณะของข้อสอบที่ตรงกับประเด็นปัญหาจำแนกเป็น 4 ระดับ คือ ระดับที่ตรงกับปัญหาและจำเป็นมาก ระดับที่ตรงปัญหาและสำคัญ ระดับที่ยอมรับว่าตรงกับปัญหา และระดับที่ไม่แน่ใจว่าตรงกับปัญหา ค่าความยากของข้อสอบจำแนกเป็น 3 ระดับ คือ ง่าย ปานกลางและยาก จากนั้นคำนวณหาเปอร์เซ็นต์การสอบผ่านที่คาดหวังโดยเฉลี่ย เป็นคะแนนจุดตัดของแบบสอบ

เทคนิคของแองกอฟ (Angoff, 1971)

มีวิธีการคือ ให้ผู้เชี่ยวชาญหรือครูผู้สอนพิจารณาค่าความน่าจะเป็นในการตอบข้อสอบถูกเพื่อใช้เป็นคะแนนจุดตัด

2.6 กำหนดคะแนนเกณฑ์โดยวิธีวิจัยเชิงปฏิบัติการ (Operation Research Method) หรือวิธีศึกษาผลที่ตามมา โดยอาศัยความสัมพันธ์ระหว่างคะแนนในแบบสอบอิงเกณฑ์กับผลการปฏิบัติงาน ช่วงคะแนนใดที่ปรากฏผลงานให้ค่าสูงสุด จะถือว่าเป็นคะแนนเกณฑ์โดยดูจากกราฟ

วิธีวิจัยปฏิบัติการนี้ บล็อก (Block) ได้นำมาประยุกต์ใช้โดยผู้คนที่มีความสามารถใกล้เคียงกัน จัดเป็นกลุ่ม ๆ ทำการสอบกลุ่มต่าง ๆ จนมีผลสัมฤทธิ์จากการทำแบบสอบอิงเกณฑ์ในระดับต่าง ๆ กัน เช่น 10%, 15%, 20%, 95%, 100% แล้ววัดผลคุณค่าอย่างใดอย่างหนึ่งที่สัมพันธ์กับคะแนนจากแบบสอบอิงเกณฑ์ของแต่ละกลุ่ม ที่มีผลสัมฤทธิ์ระดับต่าง ๆ กัน นำผลที่ได้มาเขียนกราฟแสดงความสัมพันธ์เพื่อพิจารณาค่าของคะแนนเกณฑ์ ( $C_x$ ) (สุพัฒน์ สุขมลสันต์, 2530 : 57)

2.7 กำหนดคะแนนจุดตัดโดยใช้ทฤษฎีการตัดสินใจ (Decision Theoretic Approaches) วิธีนี้จะแบ่งคนออกเป็น 2 กลุ่ม โดยอาศัยเกณฑ์ภายนอกบางอย่างที่เกี่ยวข้องกับสิ่งที่ศึกษา แล้วให้บุคคลทั้ง 2 กลุ่ม ทำแบบสอบอิงเกณฑ์ แล้วกำหนดคะแนนจุดตัดขึ้นมาใช้ บุคคลในแต่ละกลุ่มจะถูกแบ่งออกเป็นผู้สอบผ่านและไม่ผ่าน เมื่อใช้เกณฑ์ภายนอกมาตัดสินจะแบ่งคนออกเป็น 4 พวก ดังตาราง

		เกณฑ์ภายนอก	
		ผ่าน	ไม่ผ่าน
แบบสอบอิงเกณฑ์	ไม่ผ่าน	PA	PB
	ผ่าน	PC	PD

เมื่อ PA คือ สัดส่วนของคนที่ไม่ผ่านแบบสอบอิงเกณฑ์แต่ผ่านเกณฑ์ภายนอก (False Negative)

PD คือ สัดส่วนของคนที่ผ่านมาแบบสอบอิงเกณฑ์แต่ไม่ผ่านเกณฑ์ภายนอก (False Positive)

PB คือ สัดส่วนของคนที่ไม่ผ่านทั้งแบบสอบอิงเกณฑ์และเกณฑ์ภายนอก

PC คือ สัดส่วนของคนที่ผ่านมาทั้งแบบสอบอิงเกณฑ์และเกณฑ์ภายนอก

เกณฑ์ภายนอกที่กำหนดขึ้นนั้นจะไม่เปลี่ยนแปลง แต่คะแนนจุดตัดในแบบสอบอิงเกณฑ์ นั้นจะแปรผันไปได้หลายค่าแล้วแต่กำหนด คะแนนจุดตัดของแบบสอบอิงเกณฑ์คือค่าฟังก์ชันของคะแนนเกณฑ์  $F(C_x)$  มีค่าน้อยที่สุด

$$F(C_x) = \frac{(PA + PD)}{(PB + PC)}$$

## 2.8 กำหนดคะแนนจุดตัดโดยวิธีของเบอร์ก (Berk)

เบอร์ก (Berk, 1976 : 4-9) ได้หาคะแนนจุดตัดของแบบสอบอิงเกณฑ์ โดยการประยุกต์แนวคิดมาจากวิธีการเพิ่มคะแนนเกณฑ์อื่น ๆ เบอร์กได้ใช้การสอนเป็นเกณฑ์ภายนอก แบ่งผู้เรียนออกเป็นสองพวก คือ พวกที่ได้รับการสอนเป็นพวกที่รอบรู้ (Master) และพวกที่ไม่ได้รับการสอนเป็นพวกที่ไม่รอบรู้ (Non-Master) จากนั้นให้นักเรียนทั้งสองกลุ่มทำแบบสอบ แล้วพิจารณาการกระจายของคะแนนทั้งสองกลุ่มที่จะคาบเกี่ยวกัน จุดที่ฟังก์ชันทั้งสองตัดกัน คือ คะแนนพยากรณ์ที่จะแบ่งการเรียนรู้เป็นสี่พวกดังนี้

		เกณฑ์ภายนอก	
		ได้รับการสอน	ไม่ได้รับการสอน
คะแนนพยากรณ์	รอบรู้	รอบรู้จริง (TM)	รอบรู้ไม่จริง (FM)
	ไม่รอบรู้	ไม่รอบรู้ไม่จริง (FN)	ไม่รอบรู้จริง (TN)
		M	N

พวกรอบรู้จริง (True Master : TM) คือนักเรียนที่ได้รับการสอนซึ่งได้คะแนนเท่ากับหรือมากกว่าคะแนนจุดตัด

พวกรอบรู้ไม่จริง (False Master : FM) คือนักเรียนที่ไม่ได้รับการสอนซึ่งได้คะแนนเท่ากับหรือมากกว่าคะแนนจุดตัด

พวกไม่รอบรู้จริง (True Nonmaster : TN) คือนักเรียนที่ไม่ได้รับการสอนซึ่งได้คะแนนต่ำกว่าคะแนนจุดตัด

พวกไม่รอบรู้ไม่จริง (False Nonmaster : FN) คือนักเรียนที่ได้รับการสอนซึ่งได้คะแนนต่ำกว่าคะแนนจุดตัด

คะแนนจุดตัดจะเป็นคะแนนพยากรณ์ ที่ให้ความน่าจะเป็นในการตัดสินใจถูกสูงสุด คือ ค่า  $P(TM)+P(TN)$  สูงสุดหรือให้ค่าความน่าจะเป็น  $P(FM)+P(FN)$  ต่ำสุด ณ จุดคะแนนนั้นก็จะเป็นคะแนนจุดตัดที่เหมาะสมที่สุด ในการหาจะเลื่อนค่าคะแนนพยากรณ์ไปเรื่อย ๆ จนกว่าจะได้คุณสมบัติที่กล่าวมา ก็จะได้คะแนนจุดตัดที่เหมาะสม

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

N	แทน	จำนวนนักเรียนทั้งหมดที่ยังไม่ได้เรียน
M	แทน	จำนวนนักเรียนทั้งหมดที่เรียนแล้ว

คะแนนจุดตัดแต่ละคะแนนที่คำนวณได้สามารถตรวจสอบความเที่ยงได้โดยใช้สัมประสิทธิ์ความเที่ยงของเกณฑ์ เพื่อเลือกคะแนนจุดตัดที่มีค่าสัมประสิทธิ์ความเที่ยงของเกณฑ์สูงที่สุดมาเป็นคะแนนจุดตัดของแบบสอบ หากความเที่ยงของเกณฑ์ได้จากสูตรดังนี้

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

เมื่อ	$\phi_{vo}$	แทน	ค่าสัมประสิทธิ์ความเที่ยงของคะแนนพยากรณ์
	BR	แทน	ค่าความน่าจะเป็นของผู้รอบรู้ของประชากร = P(FN) + P(TM)
	SR	แทน	ค่าความน่าจะเป็นของการพยากรณ์ผู้รอบรู้ของประชากร = P(FM) + P(TM)

## 2.9 การกำหนดจุดตัดโดยใช้ทฤษฎีการตัดสินใจของเบส์ (Bayesian Decision Procedure)

วิธีนี้ตัดสินจัดประเภทผู้สอบโดยใช้คะแนนโดเมน ( $\pi$ ) ของผู้สอบมาเปรียบเทียบกับคะแนนจุดตัด ( $\pi_0$ ) ที่กำหนดขึ้นมาใช้ ผู้สอบที่มีคะแนนโดเมนมากกว่าหรือเท่ากับคะแนนจุดตัดจะจัดเป็นผู้รอบรู้ ส่วนผู้สอบที่มีคะแนนโดเมนน้อยกว่าจุดตัด จัดเป็นผู้ไม่รอบรู้ ในทางปฏิบัติไม่สามารถหาคะแนนโดเมนที่แท้จริงได้ เพราะไม่สามารถนำข้อคำถามในโดเมนมาใช้ได้ทั้งหมดจึงหาค่าโดเมนได้อย่างประมาณเท่านั้น สำหรับวิธีการหาค่าคะแนนโดเมนต้องอาศัยความรู้เดิมเกี่ยวกับตัวผู้สอบ ซึ่งอาจได้จากการสอบครั้งก่อนๆ จากสมุดรายงานของนักเรียน เป็นต้น ในการตัดสินนอกจากอาศัยความรู้เดิมแล้ว ยังอาศัยค่าความสูญเสียที่เกี่ยวข้องกับความคลาดเคลื่อนเนื่องจากการยอมรับที่ผิดและการปฏิเสธที่ผิดด้วย (อังคณา สายยศ, 2525 : 77)

## 2.10 กำหนดคะแนนจุดตัดโดยวิธีของแฟร์เนอร์ (Fhaner's procedure)

การกำหนดคะแนนจุดตัดตามวิธีของแฟร์เนอร์ จะอาศัยคะแนนมวลความรู้จากการแจกแจงทวินามที่มีการกระจายแบบปกติ โดยพิจารณาจาก

$$\text{Prop} [X > c | \pi] < \alpha \text{ สำหรับทุกค่า } \pi < \pi_0$$

$$\text{Prop} [X > c | \pi] < \beta \text{ สำหรับทุกค่า } \pi < \pi_0$$

เมื่อ	X	แทน	คะแนนมวลความรู้ของแต่ละคนที่ประมาณได้
	$\alpha$ และ $\beta$	แทน	ระดับการเสี่ยงที่ยอมรับได้ ซึ่งมีค่ามากที่สุด

ในกรณีที่มีจำนวนข้อของแบบสอบมีน้อยมาก เมื่อเทียบกับจำนวนข้อในประชากรแล้ว ค่าของ  $\alpha$  และ  $\beta$  จะไม่เป็นที่ยอมรับ แฟร์เนอร์ได้เสนอให้ใช้ค่า  $\pi_1$  และ  $\pi_2$  แทนซึ่งเป็นค่าความคลาดเคลื่อนในการตัดสินใจผิด คือ ตัดสิน  $\pi > \pi_0$  เมื่อแท้จริงแล้ว  $\pi_1 < \pi < \pi_0$  และตัดสิน  $\pi < \pi_0$  เมื่อแท้จริงแล้ว  $\pi_0 < \pi < \pi_2$  เมื่อกำหนดค่า  $\pi_1$  และ  $\pi_2$  แล้ว การแจกแจงทวินามที่เป็นการกระจายแบบปกติสามารถนำมาใช้กำหนดคะแนนจุดตัด

วิธีของแฟร์เนอร์ มีปัญหาที่ การแจกแจงทวินามจะไม่เหมาะสมกับแบบจำลองของการสุ่มข้อสอบ ที่ผู้สอบทุกคนต้องสอบจากข้อสอบที่สุ่มมาเหมือนกันหมด แต่วิธีนี้เป็นวิธีที่ง่ายกว่า วิธีหาคะแนนมวลความรู้วิธีอื่น ๆ (บุญเชิด ภิญญอนันตพงษ์, 2527 : 138-139)

### 2.11 การกำหนดจุดตัดโดยอาศัยทฤษฎีความสามารถแฝง (Latent Trait Theory)

นักวัดผลบางท่านได้พยายามหาจุดตัดของคะแนนแต่ละข้อโดยใช้แนวคิดจากการวิเคราะห์ข้อทดสอบจาก “ทฤษฎีความสามารถแฝง (Latent Trait Theory)” มาใช้โดยอาศัยจุดที่ทำให้ค่าอำนาจจำแนก (a) สูงที่สุด หรือคือจุดที่ ICC (Item Characteristic Curve) ขั้นที่สุด จุดตัดนี้บอกถึงระดับความสามารถของผู้สอบข้อสอบแต่ละข้อ ซึ่งเรียกกันว่า  $\theta_c$  ผู้ที่มี Latent score สูงกว่า  $\theta_c$  ถือว่า Master และผู้ที่ได้คะแนน Latent score ต่ำกว่า  $\theta_c$  ถือว่า Non-master สิ่งที่ยืนยัน แต่วิธีดังกล่าวนี้ยังไม่แพร่หลายเพราะการคำนวณหาค่า parameter จาก Latent Trait Model ต้องอาศัยเครื่องคอมพิวเตอร์ (สุพัฒน์ สุขมลสันต์, 2530 : 69)

### 3. ข้อคำนึงในการกำหนดคะแนนจุดตัด

มิลล์แมน (Millman, 1974 : 206 อ้างอิงใน สุพัฒน์ สุขมลสันต์, 2530 : 69) ได้ให้คำแนะนำว่า ในการกำหนดจุดตัดนั้น ควรคำนึงถึงปัจจัยต่าง ๆ ดังต่อไปนี้

1) คำนึงถึงความสามารถจริง (Performance) ของผู้เรียนอื่นด้วย

การกำหนดจุดตัดนั้นต้องอาศัยผลของการสอบของผู้เรียนอื่นด้วย ดังนั้นนักทดสอบต้องคำนึงด้วยว่า โดยทั่วไปความสามารถของผู้สอบที่ใช้เพื่อกำหนดจุดตัดนั้นเป็นอย่างไร ผู้รับการสอบที่มีความสามารถอาจใช้เป็นบรรทัดฐานในการกำหนดจุดตัดที่ปฏิบัติได้ เช่น นักทดสอบอาจกำหนดว่าจุดตัดควรเป็นอย่างน้อย 10% ของคะแนนโดยเฉลี่ยของผู้ที่มีความสามารถในระดับ Master เป็นต้น

2) คำนึงถึงความยากของแบบสอบ

นักทดสอบควรใช้ความพินิจพิจารณาตัดสินว่า ปัญหาต่าง ๆ หรือจุดมุ่งหมาย ต่าง ๆ ที่นำมาทำการทดสอบนั้น ผู้สอบที่มีความสามารถปานกลางนั้นควรจะได้ถูกต้องสักเท่าใด

จึงจะอยู่ในระดับที่พอใจ

### 3) คำนิยามถึงผลกระทบที่ตามมา

จุดมุ่งหมายที่สำคัญในการกำหนดจุดตัด คือ การที่ผู้เรียนมีความรู้ความสามารถในระดับต่ำที่สุดเพื่อให้เพียงพอที่จะศึกษาต่อในระดับสูงกว่าได้ แต่บางครั้งก็ต้องคำนึงด้วยว่า ถ้าจุดตัดสูงเกินไปจะทำให้ผู้เรียนไม่ผ่านโดยไม่จำเป็น แต่ถ้าหากว่าต่ำเกินไปก็จะทำให้มาตรฐานการศึกษาต่ำด้วย หรือทำให้ผู้เรียนเกิดปัญหาในการเรียนต่อไป ดังนั้นจึงควรหาจุดตัดที่เหมาะสมที่สุดมาใช้ในการตัดสินใจ โดยยึดความสำคัญของเนื้อหาหรือจุดประสงค์เป็นหลักว่า เรื่องใดจำเป็นต้องให้ผู้เรียนสอบผ่าน และเรื่องใดที่จำเป็นน้อยให้พิจารณาตัดได้ถ้าจำเป็น

### 4) คำนิยามถึงผลทางจิตวิทยาและการลงทุนทางการศึกษา

การที่ให้ผู้เรียนสอบตกอยู่เสมอ ๆ ย่อมเป็นผลเสียต่อจิตใจของผู้สอบ และเป็นผลเสียต่อการลงทุน ดังนั้น นักทดสอบควรพิจารณาว่าจะแก้ปัญหาเหล่านี้ได้อย่างไรเพราะว่า "ความสามารถ ความเข้าใจ และความซบซึ้งเป็นประสบการณ์ที่คนโดยมากมีอยู่" ไม่ใช่ว่ามีอยู่ทั้งหมดหรือไม่มีเลย แต่ว่าต่างคนก็มีอยู่เล็กน้อยแตกต่างกัน

### 5) คำนิยามถึงความผิดพลาดในการเดาและการสุ่มตัวอย่างข้อสอบ

การที่ผู้สอบ สอบผ่าน หรือไม่ผ่าน อาจเกิดจากความผิดพลาดจากการเดาก็ได้ หรือเพราะว่าการสุ่มตัวอย่างของข้อสอบมาได้ไม่ดี ดังนั้น นักทดสอบต้องคำนึงถึงด้วยว่าข้อสอบชวนให้เดาได้ง่ายหรือไม่ ให้เวลาเพียงพอหรือไม่ และสุ่มมาเป็นตัวแทนของประชากรหรือไม่ ซึ่งกรณีหลังนี้อาจดูได้จากค่าดัชนีความสัมพันธ์ของเนื้อหา (Index of Congruence) ก็ได้

## 4. ข้อเสนอแนะในการกำหนดคะแนนจุดตัด

ได้มีผู้ให้ข้อเสนอในการกำหนดคะแนนจุดตัดไว้ดังนี้

1) การตัดสินใจคะแนนจุดตัดควรใช้กลุ่มผู้มีประสบการณ์ ช่วยกันพิจารณากำหนดความสำคัญของแต่ละจุดประสงค์ และกำหนดจุดตัดที่เหมาะสมของแต่ละจุดประสงค์ตามความสำคัญของมันผลที่ได้คือ แต่ละจุดประสงค์มีจุดตัดต่างกัน

2) กลุ่มผู้ที่จะตัดสินคะแนนจุดตัดนี้ควรมีตั้งแต่สองกลุ่มขึ้นไป เพื่อพิจารณาคะแนนจุดตัดและศึกษาความคงที่และความแตกต่างในการกำหนดน้ำหนักของผู้ตัดสินก็จะได้มีการอภิปรายกันเพื่อหาความสอดคล้องในการตัดสินใจ

3) การหาจุดตัดที่ดีนั้น ต้องแน่ใจว่าข้อสอบทุกข้อมีคุณภาพและวัตถุประสงค์เชิงพฤติกรรมที่กำหนดได้จริง คะแนนหรือจุดตัดจะไม่มี ความหมายหากไม่คำนึงถึงการออกข้อสอบให้ตรงวัตถุประสงค์

4) เนื้อหาและพฤติกรรมที่มีการกำหนดลำดับขั้นต่อเนื่องจากระดับพื้นฐานไปหา ระดับสูงในการวัดผลวิชาเหล่านี้ควรจะกำหนดคะแนนจุดตัดให้สูง เพื่อจะได้มั่นใจว่าผู้ที่สอบผ่าน เป็นผู้ ที่เรียนรู้ตามจุดประสงค์ในระดับต้นเพียงพอที่จะเป็นพื้นฐานในการเรียนระดับสูงขึ้นไปแล้ว ในวิชาเดียวกันสำหรับกลุ่มผู้เรียนที่มีจุดมุ่งหมายในการเรียนต่อต่างกัน อาจจะใช้จุดตัดต่างกันได้

5) ข้อสอบหนึ่งฉบับอาจวัดหลายจุดประสงค์ แต่ละจุดประสงค์อาจออกข้อสอบ ได้หลายข้อ การหาจุดตัดจุดเดียวของคะแนนรวมทั้งฉบับอาจไม่เหมาะสม แม้ว่าบังเอิญทุก จุดประสงค์มีความสำคัญเท่ากัน และแต่ละจุดประสงค์มีจำนวนข้อสอบเท่า ๆ กันก็ตาม เพราะ หน้าที่ของครูต้องตรวจทีละจุดประสงค์ วิธีที่ดีอย่างหนึ่งคือ อาจต้องมีการจัดน้ำหนักของ จุดประสงค์ตามความสำคัญต่อการนำไปใช้แล้วหาจุดตัดของคะแนนของข้อสอบในแต่ละ จุดประสงค์

6) ควรมีการตรวจสอบปรับจุดตัดตั้งไว้อยู่เสมอ ๆ โดยนำข้อมูลที่ผ่าน ๆ มา ของการ ที่ใช้เป็นเกณฑ์ตัดสินผ่าน-ไม่ผ่าน เพื่อไม่ให้จุดตัดต่ำเกินไปจนขาดมาตรฐาน และเพื่อไม่ให้จุดตัดนั้น สูงเกินไปจนเด็กสอบตกมากๆ ทั้งๆ ที่มีความ สามารถ การติดตามผลเด็กที่เรียนผ่านไปแล้วก็ เป็นข้อมูลอีกทางหนึ่งของการปรับปรุงจุดตัดได้

ในการกำหนดคะแนนจุดตัดของแบบสอบอิงเกณฑ์ แนวคิดหนึ่งได้แบ่งวิธีการกำหนด ให้ 2 ลักษณะใหญ่ ๆ คือ

- 1) กำหนดคะแนนจุดตัดอิงเกณฑ์ โดยใช้ทฤษฎีตัดสินใจของมนุษย์อย่างเดียว
- 2) กำหนดคะแนนจุดตัดอิงเกณฑ์ โดยใช้ทฤษฎีตัดสินใจของมนุษย์และวิธีทางสถิติ มาประกอบกัน

ปัจจุบันการกำหนดคะแนนจุดตัดอิงเกณฑ์โดยวิธีหลัง เป็นที่นิยมและเชื่อถือมากกว่า วิธีแรกเนื่องจากเป็นวิธีที่มีเหตุผลและมีความเป็นปรนัยมากกว่า (สุพัตนะ สุกมลสันต์, 2530 : 73) ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยจึงได้เลือกใช้วิธีการกำหนดคะแนนจุดตัดของเบอร์ก ซึ่งเป็นวิธีที่ง่าย ครูสามารถที่จะนำไปใช้ได้จริง (อังคณา สายยศ, 2525 : 77) และจากการวิจัยของ ขวลิขิต โพธิ์นคร (2528) วุฒิมุณ เสาวรัตน์ (2526) และสุโชติ สันตติวงศ์ไชย (2526) พบว่า การกำหนด คะแนนจุดตัดวิธีของเบอร์ก เป็นวิธีที่มีแนวโน้มให้ค่า ความตรงสูงสุด

## แบบสอบปรนัยชนิดเลือกตอบและชนิดตอบสั้น

แบบสอบสามารถแบ่งได้เป็น 2 ประเภท ตามรูปแบบของคำถามและคำตอบ คือ แบบสอบอัตนัย (Subjective Test) เป็นแบบสอบที่ให้ผู้ตอบได้ตอบยาว ๆ แสดงความคิดเห็นอย่างเต็มที่ภายในเวลาที่กำหนด และแบบสอบปรนัย (Objective Test) เป็นแบบสอบที่มุ่งให้ผู้สอบตอบสั้น ๆ หรือเพียงเขียนเครื่องหมาย สัญลักษณ์ใด ๆ เท่านั้น

### 1. แบบสอบปรนัย

แบบสอบปรนัย หมายถึง แบบสอบที่ประกอบด้วยข้อสอบที่มีความเป็นปรนัยสูง (Objective) ความเป็นปรนัยของข้อสอบได้แก่คุณสมบัติ 3 ประการคือ (ปราณี ทองคำ, 2539 : 31)

- 1) คำถามมีความหมายชัดเจน ทุกคนอ่านแล้วเข้าใจตรงกัน ไม่กำกวม
- 2) ตรวจให้คะแนนได้ตรงกัน ข้อสอบที่มีความเป็นปรนัยไม่ว่าใครจะเป็นผู้ตรวจ

ข้อสอบก็จะให้คะแนนตรงกัน

- 3) แปลความหมายคะแนนได้ตรงกัน

ข้อสอบปรนัยที่ใช้กันทั่วไปมีหลายรูปแบบ แบ่งได้เป็น 4 ชนิดคือ (ภัทรา นิคมานนท์, 2538 : 72 ; ปราณี ทองคำ, 2539 : 31)

- 1) ชนิดตอบสั้น หรือเติมคำ ( Short-answer หรือ Completion)
- 2) ชนิดให้เลือกทางใดทางหนึ่ง เช่น ถูก-ผิด ใช่-ไม่ใช่ (True-false Item)
- 3) ชนิดให้จับคู่คำถามกับคำตอบที่ถูกต้อง (Matching Item)
- 4) ชนิดให้เลือกข้อถูกที่สุด (Multiple Choice Item)

### 2. แบบสอบเลือกตอบ (Multiple Choices)

แบบสอบชนิดเลือกตอบ เป็นแบบสอบปรนัยที่มีการใช้กันอย่างแพร่หลาย ข้อสอบชนิดเลือกตอบประกอบด้วย 2 ส่วน คือ

1). ตอนนำหรือตัวคำถาม (Stem) ทำหน้าที่เสนอปัญหาหรือคำถามเป็นสิ่งเร้าเพื่อให้ ผู้สอบตอบสนอง

2) ตัวเลือก (Choices หรือ Option ) เป็นสิ่งที่ผู้สอบจะต้องเลือกเป็นคำตอบ ซึ่งแบ่งออกเป็น 2 ประเภท คือ

- 2.1) ตัวถูก (Correct Choices)
- 2.2) ตัวลวง (Decoys หรือ Distracters)



## รูปแบบของแบบสอบเลือกตอบ

นักวัดผลหลายท่านได้จำแนกรูปแบบของแบบสอบเลือกตอบไว้ในลักษณะเดียวกัน เป็น 3 รูปแบบใหญ่ ๆ คือ

1) คำถามเลือกตอบประเภทคำถามโดด ๆ (Single Item) ลักษณะคำถามจะถามเรื่องเดียวโดยเฉพาะไม่เกี่ยวกับข้ออื่น คำถามและตัวเลือกจะจบสมบูรณ์ในข้อนั้น ๆ การตอบแต่ละข้อจะอิสระไม่เกี่ยวข้องกับข้ออื่น คำถามชนิดนี้จะประกอบไปด้วยโจทย์ หรือข้อความที่ต้องการจะถามเพื่อวัดความรู้ความสามารถเรื่องหนึ่งกับตัวเลือกตั้งแต่ 3 ตัวเลือก หรือมากกว่าอีก 1 ชุดรวมกันเป็น 1 ข้อ

2) คำถามเลือกตอบประเภทตัวเลือกคงที่ (Constant Choice) เป็นแบบสอบที่ตัวเลือก 1 ชุดใช้ได้กับคำถามหลาย ๆ ข้อ รูปแบบนี้เกิดจากตัวเลือกแต่ละข้อคำถามโดดๆ ซ้ำกันอยู่บ่อยๆ ดังนั้นเพื่อให้คำถามและตัวเลือกมีประสิทธิภาพขึ้น จึงเอาตัวเลือกที่ซ้ำมาเป็นตัวเลือกคงที่ แล้วเขียนคำถามเป็นข้อ ๆ ไปโดยให้เลือกตอบตามตัวเลือกที่กำหนดให้

### 3) คำถามเลือกตอบประเภทสถานการณ์ (Situational Test)

เป็นรูปแบบที่มีการเสนอสิ่งเร้า เช่น สถานการณ์จำลอง ข้อความหรือภาพสิ่งของหรือเรื่องราวต่าง ๆ อย่างใดอย่างหนึ่งมาเร้าให้เด็กเกิดความคิดก่อน แล้วจากนั้นก็ตั้งคำถามวัดเฉพาะในสถานการณ์เท่านั้น จะอาศัยส่วนภายนอกมาตอบถูกไม่ได้

มิเชลล์ (Michell, 1950 : 160) ได้กล่าวถึงแบบสอบชนิดเลือกตอบว่า เป็นแบบสอบที่มีคุณค่ามากที่สุดในบรรดาแบบสอบทั้งหลาย เพราะสามารถถามให้นักเรียนตีความ พินิจพิเคราะห์ เลือกและนำสิ่งที่เรียนมาใช้ให้เป็นประโยชน์ และสามารถวัดความเข้าใจ การตัดสินใจ การวินิจฉัยเหตุผล และวัดความจำของนักเรียนได้อย่างกว้างขวางกว่าการใช้แบบสอบชนิดเติมคำ และในทำนองเดียวกันกรอนลันด์และลินน์ (Gronlund and Linn, 1990 : 166) ได้กล่าวว่าแบบสอบชนิดเลือกตอบเป็นแบบสอบปรนัย ที่มีประโยชน์และสามารถนำไปใช้วัดผลได้อย่างกว้างขวางมากที่สุด เพราะนอกจากจะใช้สอบวัดความรู้ ความจำที่เป็นผลการเรียนรู้ได้ง่าย ๆ ได้ดีกว่าแบบอื่น ๆ แล้วยังสามารถวัดผลการเรียนรู้ ที่เป็นความรู้ ความเข้าใจ และการประยุกต์ใช้มีความยืดหยุ่นและมีประสิทธิภาพในการวัดสูงกว่าแบบสอบชนิดอื่น แต่การสร้างแบบสอบชนิดเลือกตอบที่มีคุณภาพดีนั้นยาก ต้องอาศัยเวลาประสบการณ์ ทักษะมาก และปัญหาที่พบบ่อยที่สุดก็คือการเดา เนื่องจากมีคำตอบไว้ให้เลือก ผู้ตอบสามารถเลือกคำตอบโดยที่ไม่มีความรู้ในเรื่องนั้นเลย ก็มีโอกาที่จะตอบถูก (อนันต์ ศรีโสภณ, 2524 : 129)

### 3. แบบสอบชนิดตอบสั้น (Short-answer)

แบบสอบชนิดตอบสั้นเป็นแบบสอบที่ประกอบด้วยข้อสอบที่กำหนดให้ผู้ตอบตอบสั้น ๆ ด้วยการตอบเป็นคำ วลี ตัวเลข หรือสัญลักษณ์ทางคณิตศาสตร์

ข้อสอบชนิดตอบสั้นสามารถแบ่งได้เป็น 3 แบบดังนี้ (Mehrens and Lehmann, 1984 : 133)

1) แบบตั้งคำถามโดยตรง หรือแบบข้อความสมบูรณ์ เป็นข้อคำถามที่ใช้เป็นประโยคที่มีเนื้อความสมบูรณ์

2) แบบข้อความไม่สมบูรณ์ ข้อคำถามเป็นประโยคที่ไม่สมบูรณ์

3) แบบความสัมพันธ์ เป็นข้อคำถามที่ตั้งด้วยประโยคหลัก และตามด้วยข้อความย่อย ๆ แล้วให้หาคำตอบเติมให้มีความสัมพันธ์ตรงกับคำหรือข้อความย่อย ๆ นั้น

อูสเตอร์ฮอฟ (Oosterhof, 1990 : 87) ได้กล่าวถึงแบบสอบชนิดตอบสั้นว่าเป็นแบบสอบที่ใช้กันมากที่สุดในห้องเรียน กรอนลันด์และลินน์ (Gronlund and Linn, 1990 : 145) ได้กล่าวว่าข้อสอบชนิดตอบสั้นว่าเป็นแบบสอบที่สร้างได้ง่ายที่สุด เนื่องจากเกี่ยวข้องกับตัวอย่างผลของการเรียนที่วัดบ่อย ๆ เป็นความคาดหวังสำหรับการแก้ปัญหาผลการวัดในวิชาคณิตศาสตร์และวิทยาศาสตร์ ข้อสอบชนิดตอบสั้นจะใช้ในการวัดเกี่ยวกับการระลึกได้ (Recall) ของข้อมูลข่าวสารในความจำโดยเฉพาะมากที่สุด และแบบสอบตอบสั้นจะมีประสิทธิภาพมากที่สุดเมื่อใช้วัดผลของการเรียนที่เฉพาะเจาะจง

แบบสอบชนิดตอบสั้น เป็นแบบสอบที่มีประโยชน์เป็นอย่างยิ่งในการวัดผลสาขาวิทยาศาสตร์ และคณิตศาสตร์ เพราะสามารถวัดความรู้เกี่ยวกับข้อเท็จจริง คำศัพท์ กฎเกณฑ์ ความสามารถที่จะหาคำตอบโดยการคำนวณ ทักษะในการใช้สัญลักษณ์ สมดุลย์ สมการเคมี ความสามารถที่จะตีความหมายของข้อมูล วัดกระบวนการทางปัญญาขั้นสูง และความสามารถที่จะแปลความหมายจากรูปภาพหรือแผนภูมิที่จะกำหนดให้ (ภัทรา นิมานนท์, 2538 : 75) และมีความรู้สึกเชื่อมั่นว่าการให้นักเรียนหาคำตอบมาใส่จะมีความเที่ยงตรงกว่าให้นักเรียนเลือกคำตอบที่ถูกต้องจากตัวเลือกต่าง ๆ ที่กำหนดให้ (อนันต์ ศรีโสภณ, 2524 : 114)

อีเบล (Ebel, 1979 : 166-167) ได้กล่าวถึงแบบสอบชนิดเลือกตอบและชนิดตอบสั้นว่าแบบสอบทั้งสองมีความสัมพันธ์กันสูงมาก เมื่อแบบสอบทั้งสองวัดในสิ่งเดียวกันและเป็นแบบสอบคู่ขนาน (Parallel Form) กัน คือถ้านักเรียนสามารถเขียนคำตอบได้ถูกต้องก็น่าที่จะจำแนกตัวเลือกที่ถูกต้องได้เช่นกัน อย่างไรก็ตามการจำแนกตัวเลือกน่าจะง่ายกว่าการสร้างคำตอบขึ้นมาเพราะโดยรูปแบบของแบบสอบชนิดเลือกตอบมีการเสนอสิ่งเร้าและการตอบสนอง

ให้เสรีจอมงายกว่าการที่จะตอบถูกโดยรู้แต่สิ่งเร้าหรือคำถามเท่านั้น ซึ่งผู้สอบต้องใช้ความสามารถในการระลึกได้ (Recall) จึงทำให้ยากที่จะเดาคำตอบได้ถูก

### งานวิจัยที่เกี่ยวข้อง

#### 1. งานวิจัยที่เกี่ยวข้องกับการกำหนดจุดตัดของแบบสอบอิงเกณฑ์

งานวิจัยที่เกี่ยวกับการกำหนดคะแนนจุดตัดของแบบสอบอิงเกณฑ์ได้มีผู้สนใจและศึกษากันมากทั้งต่างประเทศและในประเทศ เนื่องจากคะแนนจุดตัดเป็นเป็นปัจจัยที่สำคัญอย่างหนึ่งที่จะส่งผลต่อความเที่ยงและความตรงของแบบสอบอิงเกณฑ์ เช่น

กาญจนา วัฒนสุนทร (2521 : 106-111) ได้ศึกษาทดลองกำหนดคะแนนจุดตัดแบบสอบอิงเกณฑ์วิชาคณิตศาสตร์เรื่องสมการชั้นมัธยมศึกษาปีที่ 1 จำนวน 4 ฉบับ ๆ ละ 10 ข้อ ใช้กลุ่มตัวอย่าง 60 คน โดยกำหนดเกณฑ์ไว้ 60, 70 และ 80 เปอร์เซนต์ตามลำดับ พบว่าเกณฑ์ 60 เปอร์เซนต์ให้ค่าความตรงเชิงโครงสร้างของคาร์เวอร์ และค่าความเที่ยงของลิฟวิงตันสูงสุด

สมถวิล วิจิตรวรรณ (2524 : 115-120) ได้ใช้วิธีการกำหนดคะแนนจุดตัดตามวิธีนับถอยหลังจาก 100% และวิธีของเบอร์ก ในการคำนวณคะแนนจุดตัดของแบบสอบอิงเกณฑ์วิชาคณิตศาสตร์ จำนวน 4 ฉบับ ฉบับละ 10 ข้อ พบว่าคะแนนจุดตัดที่ได้มีค่าระหว่าง 50% ถึง 80%

วุฒิคุณ เสาวรัตน์ (2526 : 288) ได้ศึกษาการสร้างแบบสอบอิงเกณฑ์วิชาคณิตศาสตร์เรื่องภาคตัดกรวย ชั้นมัธยมศึกษาปีที่ 4 จำนวน 11 ฉบับ กำหนดคะแนนด้วยวิธีของเบส์ เบอร์ก และแกลส ปรากฏว่าวิธีของเบอร์ก เป็นวิธีที่ให้คะแนนเกณฑ์เหมาะสมที่สุดคือ 9 ฉบับ รองลงมาคือวิธีของแกลส 5 ฉบับ และวิธีของเบย์ 3 ฉบับ

สุชาติ สันตติวงศ์ไชย (2526 : 175) ได้สร้างแบบสอบอิงเกณฑ์วิชาคณิตศาสตร์เรื่อง สมการควอดราติก ชั้นมัธยมศึกษาปีที่ 3 จำนวน 6 ฉบับ หากคะแนนจุดตัดด้วยวิธีของเบอร์ก วิธีการใช้ทฤษฎีการตัดสินใจ พบว่าแบบสอบฉบับที่ 1, 2, 4 และ 6 ได้คะแนนจุดตัดด้วยวิธีของเบอร์ก สำหรับแบบสอบฉบับที่ 3 และ 4 คะแนนเกณฑ์ที่เหมาะสมได้จากทฤษฎีการตัดสินใจ

รังสรรค์ มณีเล็ก (2527 : 74-75) ได้เปรียบเทียบค่าความเที่ยงของแบบสอบที่ได้จากการกำหนดคะแนนจุดตัด 4 วิธีคือวิธีนับถอยหลังจาก 100% วิธีนิตลสกี วิธีของเบอร์ก และวิธีของเบย์ พบว่าวิธีนับถอยหลังจาก 100% ให้คะแนนจุดตัดสูงสุด แต่วิธีนิตลสกีให้คะแนนจุดตัดต่ำสุด และความเที่ยงที่ได้จากการกำหนดตามวิธีดังกล่าวไม่แตกต่างกัน

ชวลิต โพธิ์นคร (2528) ได้ศึกษาเปรียบเทียบผลของการกำหนดคะแนนจุดตัดแบบสอบอิงเกณฑ์ชนิดเลือกตอบ โดยวิธีของเบอร์ก วิธีประยุกต์ราชดีโมเดล และวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุด พบว่าค่าความตรงเชิงโครงสร้างของแบบสอบที่กำหนดจุดตัดโดยวิธีของเบอร์กมีแนวโน้มให้ค่าสูงสุด รองลงมาได้แก่วิธีประยุกต์ราชดีโมเดล และวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุด และเมื่อวิเคราะห์ตามกลุ่มความสามารถระดับสูง ปานกลาง และต่ำ จุดตัดที่ได้จากวิธีของเบอร์ก จะแตกต่างกัน โดยกลุ่มสูงมีจุดตัดสูงสุด รองลงมาได้แก่กลุ่มปานกลางและกลุ่มต่ำตามลำดับ

พานิช ศรีงาม (2530) ได้เปรียบเทียบคุณภาพของแบบสอบอิงเกณฑ์โดยใช้วิธีการกำหนดจุดตัดด้วยวิธีการประยุกต์รูปแบบของราสส์ วิธีการใช้ทฤษฎีการตัดสินใจของแกลส และวิธีนับถอยหลัง พบว่า วิธีกำหนดจุดตัดที่ให้ค่าจุดตัดสูงสุดได้แก่วิธีนับถอยหลัง รองลงมาคือวิธีการประยุกต์รูปแบบของราสส์ และต่ำสุดคือวิธีการใช้ทฤษฎีการตัดสินใจของแกลส ความตรงของแบบสอบอิงเกณฑ์ที่ได้จากวิธีการกำหนดจุดตัดด้วย วิธีการประยุกต์รูปแบบของราสส์ วิธีการใช้ทฤษฎีการตัดสินใจของแกลส และวิธีนับถอยหลัง แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 แต่ความเที่ยงของแบบสอบอิงเกณฑ์ที่ได้จากวิธีการกำหนดจุดตัดด้วยวิธีทั้ง 3 ไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

เบญเนียค (Behuniak, 1981 : 3998-A-3999-A) ได้ศึกษาวิธีหาจุดตัดของแบบสอบอิงเกณฑ์วิชาคณิตศาสตร์ และการอ่านของนักเรียนระดับ 9 ตามวิธีของนีเดลสกีและวิธีของแองกอฟ โดยใช้ผู้เชี่ยวชาญทางด้านเนื้อหาจำนวน 27 คน พบว่า

- 1) การใช้วิธีการกำหนดจุดตัดที่ต่างกันจะทำให้ได้จุดตัดที่ต่างกัน
- 2) การใช้กลุ่มตัวอย่างที่ต่างกันกำหนดจุดตัดวิธีเดียวกันจะทำให้ได้จุดตัดต่างกัน
- 3) การคำนวณจุดตัดตามวิธีของแองกอฟ และวิธีของนีเดลสกี จะได้จุดตัดที่มีความแปรเปลี่ยนไปเท่า ๆ กัน

ฮาราซิม (Harasym, 1981 : 725-727 ) ได้ทำการเปรียบเทียบวิธีการกำหนดจุดตัดของแบบสอบอิงเกณฑ์ โดยใช้วิธีของแองกอฟ และนีเดลสกี กับกลุ่มตัวอย่างซึ่งเป็นนักศึกษาแพทย์ของมหาวิทยาลัยคัลการี จำนวน 212 คน จากการศึกษาพบว่าจุดตัดที่ได้จากวิธีของนีเดลสกีมีค่าต่ำกว่าวิธีของแองกอฟ และพบว่าวิธีของนีเดลสกี ให้ผลในการจำแนกได้เป็นที่น่าพอใจกว่าวิธีของแองกอฟ

ซาร์ราน ( Zahran, 1982 : 3131-A) ได้ศึกษาเกี่ยวกับตัวแปรต่าง ๆ ที่ส่งผลต่อการกำหนดคะแนนจุดตัดตามวิธีของนิเคลสกี กลุ่มตัวอย่างเป็นนักศึกษาในมหาวิทยาลัยไอไฮโอ จำนวน 145 คน ใช้ครูผู้สอนตัดสินเกณฑ์ จำนวน 10 คน โดยเป็นครูที่มีประสบการณ์ทางการสอน 5 คน และครูที่ไม่มีประสบการณ์ทางการสอนมาเลย 5 คน จากการศึกษาพบว่าครูที่มีประสบการณ์ กำหนดคะแนนจุดตัดต่ำกว่าครูที่ไม่มีประสบการณ์ และรูปแบบข้อสอบมีผลต่อการกำหนดคะแนนจุดตัด

อราสมิธ ( Arrasmith , 1987 : 3400-A) ได้ศึกษาวิธีการกำหนดจุดตัดของแบบสอบอิงเกณฑ์ โดยใช้วิธีการพิจารณาเนื้อหาข้อสอบกับการตัดสินตัวบุคคลที่เข้าสอบผลการเปรียบเทียบ 2 วิธีนี้ พบว่าให้ผลไม่แตกต่างกัน

สำหรับการวิจัยนี้ ผู้วิจัยได้นำเอาวิธีการกำหนดจุดตัดของเบอร์ก (Berk) มาใช้ในการกำหนดคะแนนจุดตัด เพราะจากการศึกษาเปรียบเทียบของแบบสอบที่ได้จากการกำหนดจุดตัดที่แตกต่างกันของชวลิต โพธิ์นคร (2528) พบว่าความเที่ยงที่คำนวณจากคะแนนจุดตัดของเบอร์กมีค่าสูงสุด และจากการศึกษาของวุฒิคุณ เสาวรัตน์ (2526) และสุโชติ จันตติวงศ์ไชย (2536) พบว่าการกำหนดจุดตัดด้วยวิธีของเบอร์กเป็นวิธีที่คะแนนจุดตัดเหมาะสมที่สุด

## 2. งานวิจัยที่เกี่ยวข้องกับแบบสอบเลือกตอบและตอบสั้น

ไพบูลย์ จิตต์โต (2514) ได้ศึกษาเกี่ยวกับการใช้แบบสอบชนิดเลือกตอบชนิดต่าง ๆ และแบบเติมคำ วัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ของนักเรียนชั้นประถมศึกษาปีที่ 5 และปีที่ 7 ของโรงเรียนในสังกัดเทศบาลนครนนบุรี พบว่าแบบสอบชนิดเติมคำให้ค่าความเที่ยง ความตรง อำนวยจำแนก และค่าความยากมาตรฐานสูงสุด

สมหวัง พิธิยานุวัฒน์ และเพ็ญศิริ ต่านชนะ (2524 : 32-45) ได้เปรียบเทียบคุณภาพของแบบสอบอิงกลุ่มชนิดเลือกตอบกับชนิดตอบสั้นโดยใช้กลุ่มตัวอย่าง นิสิตปริญญาโทชั้นปีที่ 1 ที่มีพื้นฐานความรู้ทางคณิตศาสตร์ 30 คน และนิสิตที่ไม่มีพื้นฐานความรู้ทางคณิตศาสตร์ 58 คน ใช้แบบสอบ 2 ฉบับ ฉบับละ 25 ข้อ ผลการวิจัยพบว่าความแปรปรวนของคะแนนจากแบบสอบทั้ง 2 ประเภทไม่แตกต่างกันในกลุ่มที่มีพื้นฐานความรู้ทางคณิตศาสตร์ แต่ความแปรปรวนของคะแนนจากแบบสอบชนิดตอบสั้นในกลุ่มที่มีพื้นฐานความรู้ทางคณิตศาสตร์สูงกว่ากลุ่มตัวอย่างที่ไม่มีพื้นฐานความรู้ทางคณิตศาสตร์ แบบสอบชนิดเลือกตอบง่ายกว่าแบบสอบชนิดตอบสั้น โดยไม่ขึ้นกับกลุ่มตัวอย่าง ความเที่ยงของแบบสอบทั้งสองประเภทไม่แตกต่างกันในกลุ่มตัวอย่างที่ไม่มี

พื่นความรู้ทางคณิตศาสตร์ แต่กลุ่มตัวอย่างที่มีพื่นความรู้ทางคณิตศาสตร์ พบว่าความพียงของแบบสอบชนิดตอบสั้นสูงกว่าชนิดเลือกตอบอย่างมีนัยสำคัญ และแบบสอบทั้งสองประเภทนี้สามารถจำแนกกลุ่มตัวอย่างได้ตามทฤษฎี

กรรองทอง เทพศิริอำนวนย (2525) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบวิชาคณิตศาสตร์ชนิดเลือกตอบกับชนิดเติมคำ ผลการวิจัยพบว่าแบบสอบชนิดเลือกตอบง่ายกว่าแบบสอบชนิดเติมคำ และมีความพียงน้อยกว่า ค่าอำนาจจำแนกก็มีแนวโน้มจะน้อยกว่าเช่นกัน แต่ความตรงร่วมสมัยที่ใช้คะแนนจากแบบสอบมาตรฐานเป็นเกณฑ์ของแบบสอบทั้ง 2 ชนิดไม่แตกต่างกัน แต่เมื่อใช้คะแนนสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ ประจำภาคเรียนที่ 2 ปี 2523 เป็นเกณฑ์ปรากฏว่า แบบสอบเลือกตอบมีความตรงร่วมสมัยน้อยกว่าแบบสอบชนิดเติมคำ

สุรินทร์ แผงจันทิก (2528) ได้ศึกษาเปรียบเทียบคะแนนจุดตัด และความพียงของแบบสอบอิงเกณฑ์ระหว่างแบบสอบประเภทเลือกตอบกับตอบสั้น ในวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 โดยใช้ทฤษฎีการตัดสินใจของแกลลในการกำหนดคะแนนจุดตัด และหาความตรงโดยการวิเคราะห์ความแปรปรวน (ANOVA) ของโลเวตต์ ใช้กลุ่มตัวอย่าง 260 คน โดยใช้แบบสอบฉบับละ 10 ข้อ ผลการวิจัยพบว่า แบบสอบอิงเกณฑ์ชนิดเลือกตอบและตอบสั้นมีคะแนนจุดตัดเป็น 5 และ 2 คะแนนตามลำดับ และมีความแตกต่างกัน 30 เปอร์เซนต์ สัดส่วนจำนวนผู้รอบรู้และไม่รอบรู้ ระหว่างแบบสอบทั้งสองประเภท ไม่แตกต่างกัน ส่วนค่าความพียงของแบบสอบอิงเกณฑ์ประเภทเลือกตอบต่ำกว่าประเภทตอบสั้นอย่างมีนัยสำคัญที่ระดับ .01

ประนอม พันธุ์ไสว (2530) ได้ศึกษาคะแนนจุดตัดของแบบสอบอิงเกณฑ์ชนิดเลือกตอบ ตอบสั้น และถูกผิด วิชาคณิตศาสตร์ โดยใช้ทฤษฎีการตัดสินใจของเบส์ และวิธีของแกลล ใช้แบบสอบ 3 ฉบับๆ ละ 25 ข้อ พบว่าวิธีของเบส์ให้คะแนนจุดตัดแบบสอบเลือกตอบเท่ากับ 13 หรือ 52 เปอร์เซนต์ และแบบสอบตอบสั้นเท่ากับ 8 หรือ 32 เปอร์เซนต์ ส่วนวิธีของแกลล ให้คะแนนจุดตัดชนิดเลือกตอบเท่ากับ 12 หรือ 48 เปอร์เซนต์ ตอบสั้นเท่ากับ 8.5 หรือ 34 เปอร์เซนต์

อุบล แสงเพ็ง (2535) ได้ศึกษาเปรียบเทียบความคลาดเคลื่อนของแบบสอบอิงเกณฑ์ระหว่างแบบสอบเลือกตอบและตอบสั้น ในการจำแนกระดับความรู้ของนักเรียน โดยใช้แบบสอบฉบับละ 30 ข้อ วิเคราะห์ความคลาดเคลื่อนที่ความยาว 8, 11, 14, 17, 20, และ 23 ข้อ โดยใช้เกณฑ์ตัดสินจากเทคนิคการวิเคราะห์ตามลำดับขั้นเทียบกับผลการวิเคราะห์เต็มฉบับ โดยใช้เกณฑ์ผ่านร้อยละ 55 และตกร้อยละ 45 ผลการวิเคราะห์พบว่า

แบบสอบเลือกตอบและตอบสั้นที่ความยาวเดียวกัน เกิดความคลาดเคลื่อนที่ปรากฏและความคลาดเคลื่อนที่ปรากฏชัดเจนไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01

ลินด์สแลนด์ (Rinsland, 1938 : 295-299 อ้างถึงใน กรองทอง เทพศิริอำนาย, 2525 : 18) ได้รวบรวมเอกสารการทดลอง และศึกษาเกี่ยวกับการใช้แบบสอบชนิดต่าง ๆ และได้สรุปว่าแบบสอบปรนัยทั้งหมด ยกเว้นชนิดถูก-ผิด มีค่าความเที่ยงพอ ๆ กับแบบสอบอัตนัย หรือน้อยกว่าเล็กน้อย และชนิดเติมคำหรือตอบสั้น จะมีความตรงสูงสุด และแบบสอบปรนัยที่ต่างชนิดกัน ใช้เวลาในการทดสอบเท่ากัน มีค่าความเที่ยงเท่ากัน แต่การที่ค่าความเที่ยงต่างกันนั้นอาจเป็นเพราะการใช้คำหรือภาษาในคำถามที่ต่างกันมากกว่าจะเป็นที่ชนิดของแบบสอบ

แน็ปป์ (Knapp, 1969 อ้างถึงใน สุรินทร์ แผงจันทิก , 2530 : 63) ได้ศึกษาเปรียบเทียบความเที่ยงของแบบสอบชนิดเลือกตอบและแบบเติมคำ (Open-end) ในวิชาคณิตศาสตร์ โดยทดลอง กับนักเรียนเกรด 5 จำนวน 83 คน ใช้ข้อสอบฉบับละ 25 ข้อ คำนวณค่าความเที่ยงโดยไม่ใช้ค่าสหสัมพันธ์ (Correlaionless Approach) เขานิยามความเที่ยงเป็นรายข้อภายในความเชื่อที่ว่า แบบสอบในแต่ละข้อจะเป็น 1 เสมอ ถ้านักเรียนตอบถูก เขาใช้การบริหารการสอบ 2 ครั้ง ห่างกัน 1 สัปดาห์ ผลการวิจัยพบว่าบางข้อชนิดเลือกตอบ สูงกว่าบางข้อต่ำกว่าแบบเติมคำ เมื่อพิจารณาทั้ง 25 ข้อแล้ว พบว่า ชนิดเลือกตอบมีความเที่ยง .79 แบบเติมคำมีความเที่ยง .78 เขาสรุปว่าเรื่องนี้ไม่ทราบที่เกิดจากความลำเอียงมากนักอย่างไร ในการประมาณค่าความเที่ยงและสรุปว่า เรื่องความเที่ยงเป็นเรื่องที่ความซับซ้อนมาก

เกย์ (Gay, 1980) ได้ศึกษาเปรียบเทียบผลของแบบสอบชนิดเลือกตอบและตอบสั้นต่อความจำ โดยใช้กลุ่มตัวอย่างที่เป็นนักเรียนที่ลงทะเบียนเรียนในวิชาสถิติเบื้องต้น จำนวน 24 คน แบ่งเป็น 2 กลุ่ม ๆ ละ 14 คน โดยกลุ่มที่ 1 จะได้รับแบบสอบเลือกตอบในระหว่างเทอม และอีกกลุ่มหนึ่งจะได้รับแบบสอบชนิดตอบสั้น แล้วทั้ง 2 กลุ่มจะได้รับแบบสอบทั้งสองแบบในการสอบปลายเทอม ผลการศึกษาพบว่า กลุ่มตัวอย่างที่ได้รับแบบสอบเลือกตอบและตอบสั้นในการเรียนระหว่างเทอมสามารถทำแบบสอบเลือกตอบได้ไม่แตกต่างกัน แต่สำหรับแบบสอบชนิดตอบสั้นกลุ่มตัวอย่างที่ได้รับแบบสอบตอบสั้นในระหว่างเทอมสามารถทำได้ดีกว่ากลุ่มที่ได้รับแบบสอบชนิดเลือกตอบ

เฮย์นีย์ (Haynie, 1994) ได้ศึกษาผลของแบบสอบเลือกตอบและตอบสั้นต่อความคงทนในการเรียนรู้ ใช้กลุ่มตัวอย่างระดับต่ำกว่าปริญญาตรี จำนวน 187 คน โดยให้ข้อมูลจากจุลสาร

แก่กลุ่มตัวอย่าง แล้วนำแบบสอบถามเลือกตอบและตอบสั้นมาสอบกับกลุ่มตัวอย่างกลุ่มละ 64 คน และ 3 สัปดาห์ต่อมาได้นำแบบสอบถามมาสอบกับกลุ่มตัวอย่างทั้งหมดอีกครั้งหนึ่ง ผลการศึกษาพบว่า กลุ่มที่ได้รับการสอบก่อนทั้งสองกลุ่มจะมีความจำที่ดีกว่ากลุ่มที่ไม่ได้รับการสอบมาก่อน และกลุ่มที่ได้รับแบบสอบถามเลือกตอบมาก่อนจะมีผลต่อความจำดีกว่ากลุ่มที่ได้รับแบบสอบถามตอบสั้น

การวิจัยที่เกี่ยวข้องกับแบบสอบถามเลือกตอบและชนิดตอบสั้นนั้น จะเห็นได้ว่าส่วนใหญ่จะศึกษากับแบบสอบถามอิงกลุ่ม และใช้กับวิชาคณิตศาสตร์ สำหรับแบบสอบถามอิงเกณฑ์นั้นยังมีผู้ศึกษากันน้อย ผู้วิจัยจึงได้นำ แบบสอบถามทั้งสองมาศึกษาถึงความเที่ยงและความตรง ในลักษณะของแบบสอบถามอิงเกณฑ์ และศึกษาในกลุ่มตัวอย่างที่มีผลสัมฤทธิ์ทางการเรียนที่แตกต่างกัน