บทที่ 2

วรรณคดีที่เกี่ยวข้อง

การศึกษาเปรียบเทียบคุณภาพของการกำหนดจุดตัดด้วยวิธีของวิลคอกซ์กับวิธีของแองกอฟโดยการประยุกต์ทฤษฎีการตอบสนองข้อสอบสำหรับแบบสอบอิงโดเมนประเภทเลือกตอบและตอบสั้นผู้วิจัยได้แบ่งแนวคิดและทฤษฎีที่เกี่ยวข้องเป็น 5 ตอนคือ

ตอนที่ 1 แนวคิดเกี่ยวกับทฤษฎีการตอบสนองข้อสอบ

ตอนที่ 2 แนวคิดการกำหนดจุดตัดด้วยวิธีของวิลคอกซ์และวิธีของแองกอฟโดยการประยุกต์ทฤษฎีการตอบสนองข้อสอบและความหมายของคะแนนโดเมน

ตอนที่ 3 ความเที่ยงและความตรงของการกำหนดจุดตัด

ตอนที่ 4 แนวคิดเกี่ยวกับแบบสอบแบบเลือกตอบกับแบบตอบสั้น

ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 ทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการวัดผลแนวใหม่คือ ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) หรือทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) หรือ Item Characteristic หรือ Modern Theoryเป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างคุณลักษณะภายในหรือความสามารถที่มีอยู่ภายในตัวบุคคลกับพฤติกรรมการตอบสนองข้อสอบของบุคคลนั้นว่ามีโอกาสตอบข้อสอบได้ถูกต้องมากน้อยเพียงใด ทฤษฎีนี้มีความเชื่อพื้นฐานว่า พฤติกรรมการตอบสนองต่อข้อสอบของผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายใน(trait) หรือความสามารถ(ability)ที่อยู่ในตัวบุคคล ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง
(ศิริชัย กาญจนวาสี, 2534) ซึ่งทฤษฎีการตอบสนองข้อสอบมีข้อตกลงเบื้องต้นที่สำคัญดังนี้

1. แบบสอบวัดคุณลักษณะเด่นเพียงลักษณะเดียว(Unidimensionality : One trait) หรือแบบสอบมีความเป็นเอกพันธ์ การตรวจสอบข้อตกลงเบื้องต้นข้อนี้ทำได้โดยการวิเคราะห์ตัวประกอบ

2. การตอบข้อสอบเป็นไปอย่างอิสระ(Independent : Local independent) ประกอบด้วย

2.1 ความเป็นอิสระระหว่างข้อสอบ หมายความว่า การตอบข้อสอบข้อหนึ่งไม่มีผลต่อการตอบสนองข้อสอบข้ออื่น ๆ ในแบบสอบ

2.2 ความเป็นอิสระระหว่างผู้สอบ หมายความว่า ผลการตอบสนองข้อสอบของผู้สอบคนหนึ่งจะต้องไม่มีผลต่อการตอบสนองข้อสอบของผู้สอบคนอื่น

3.โค้งคุณลักษณะข้อสอบสามารถใช้อธิบายพฤติกรรมการตอบสนองข้อสอบ(Item Characteristic Curve : Item Response Model) ความสัมพันธ์ระหว่างระดับความสามารถผู้ตอบ(θ_i)กับโอกาสที่จะตอบข้อสอบถูก[$P_i(\theta)$)] ความสามารถดังกล่าวแสดงได้ด้วยโค้งคุณลักษณะข้อสอบ(ICC)

4. ข้อสอบที่ใช้ต้องไม่เป็นข้อสอบประเภทความเร็ว ผู้สอบทุกคนควรมีโอกาสทำข้อสอบทุกข้อ เพื่อให้คะแนนรวมจากการสอบประกอบกับลักษณะของข้อสอบเป็นตัวประมาณค่าความสามารถที่แท้จริง โดยไม่มีข้อจำกัดเกี่ยวกับเวลาที่ใช้ในการสอบ

สำหรับโมเดลทฤษฎีการตอบสนองข้อสอบ พัฒนาโดย Lawley ในปี 1943 โดยใช้โมเดลทางคณิตศาสตร์เป็นครั้งแรกคือ Normal Ogive Model ซึ่งต่อมาได้มีการพัฒนาโมเดลเพื่อความเหมาะสมโดยลอร์ด (Load, 1952) แต่เนื่องจากโมเดลดังกล่าวมีความยุ่งยากในการประมาณค่าพารามิเตอร์เบิร์นบอม (Birnbaum, 1968) จึงพัฒนาโมเดลใหม่คือ โมเดลโลจิสติก (Logistic Model) ซึ่งให้ความละดวกในการประมาณค่าพารามิเตอร์ เพราะโมเดลให้โค้งที่สามารถปรับให้มีความใกล้เคียงกับโค้งปกติ(Normal Ogive Model)ได้ง่าย ประกอบกับ Normal Ogive มีความยุ่งยากในการคิดคำนวณมากกว่าดังนั้นการวิเคราะห์แบบสอบจึงได้มีการนำฟังก์ชันของโลจิสติคมาใช้กันมาก

โมเดลการตอบสนองข้อสอบที่นิยมใช้กันมี 3 รูปแบบคือ 1,2 และ 3 พารามิเตอร์โดยที่โมเดลโลจิสติก 3 พารามิเตอร์เป็นที่นิยมมากกว่าโมเดลแบบ 1 และ 2 พารามิเตอร์ เพราะสามารถวิเคราะห์ค่าพารามิเตอร์ของข้อสอบได้ถึง 3 ค่า ได้แก่ ค่าความยาก(b) ค่าอำนาจจำแนก(a) และค่าการเดา (c) อย่างไรก็ดี โมเดลนี้ควรใช้กับกลุ่มตัวอย่างขนาดใหญ่ ประมาณ 1,000 คนขึ้นไป และข้อสอบที่ใช้ในการวิเคราะห์ควรมีจำนวนมากข้อด้วย

ลักษณะฟังก์ชันโลจิสติค (Logistic Function) ที่มีพารามิเตอร์ 3 ตัว แสดงความสัมพันธ์ดังนี้

$$P_i(\theta) = c_i + \frac{1-c_i}{1+e^{-1.7(\theta - b_i)}}$$

เมื่อ

$P_i(\theta)$ คือ ความน่าจะเป็นที่จะตอบข้อที่ i ถูกต้อง

1.7 คือ Scaling Factor ซึ่งเป็นค่าเพื่อปรับให้ Logistic Function

e คือ ค่าคงที่ ซึ่งมีค่าประมาณ 2.71828

b_i คือ ค่าความยากของข้อสอบข้อที่ i หมายถึง ดัชนีของข้อสอบที่แสดงถึงระดับความยากง่ายของข้อสอบ และแสดงถึงระดับความสามารถของบุคคลที่มีความน่าจะเป็นของการตอบข้อสอบข้อนั้นถูก b = (1+c)/2 ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่า b ระหว่าง -2.5 ถึง +2.5 ถ้าค่า b= -2.5 แสดงว่าข้อสอบง่ายมาก ถ้า b=+2.5 แสดงว่าข้อสอบยากมาก

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i หมายถึง ดัชนีที่แสดงถึงความสามารถของข้อสอบที่จะจำแนกกลุ่มผู้สอบ ระหว่างผู้ที่มีระดับความสามารถต่ำกว่าระดับความยาก และผู้ที่อยู่สูงกว่าความยาก โดยถือค่าโอกาสที่จะตอบถูกเป็นเกณฑ์ ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่า a อยู่ระหว่าง 0.5 ถึง +2.5 ถ้า a= 0.5 แสดงว่าข้อสอบมีอำนาจจำแนกต่ำ ถ้า a= +2.5 แสดงว่าข้อสอบมีค่าอำนาจจำแนกสูงถ้า a เป็นลบแสดงว่าข้อสอบไม่ดี จำแนกไม่ได้

c_i คือค่าการเดาของข้อสอบข้อที่ i หมายถึง ดัชนีของข้อสอบที่แสดงถึงว่าความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำมากจะมีโอกาสตอบข้อสอบข้อนั้นถูก มีค่าเป็นไปได้อยู่ระหว่าง 0 ถึง 1 ถ้าข้อสอบข้อนั้นไม่สามารถที่จะตอบถูกด้วยการเดาแล้ว c = 0โดยทั่วไปนิยมใช้ข้อสอบมีค่า c อยู่ระหว่าง 0 ถึง 0.3

θ_p คือค่าความสามารถของผู้สอบคนที่ p ซึ่งประมาณได้จากโมเดลตามทฤษฎีการตอบสนองข้อสอบและปรับให้เป็นคะแนนที่มีค่าเฉลี่ยเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ค่า θ มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัตินิยมอยู่ในช่วง --3 ถึง +3

รูปแบบการตอบสนองข้อสอบ (Item Response Model)

รูปแบบทางคณิตศาสตร์ที่ใช้ในการวิเคราะห์ข้อสอบ อาศัยฟังก์ชันการกระจายรูปแบบโลจิสติก (logistic distribution function) ซึ่งแบ่งออกเป็น 4 รูปแบบ ตามจำนวนพารามิเตอร์ของโค้งลักษณะข้อสอบ

1. รูปแบบโลจิสติกที่มี 1 พารามิเตอร์ (One-Parameter Logistic Model) รูปแบบนี้ถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากัน (a_i = 1.00) และไม่มีค่าการเดา (c_i = 0) คำนึงถึงเฉพาะค่าความยากของข้อสอบ (b_i) ดังสมการ

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}}$$

หรือ $$P_i(\theta) = \left[1 + e^{-D(\theta - b_i)}\right]^{-1}$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ θ จะตอบข้อสอบข้อ i ได้ถูก

b_i คือ ค่าความยากของข้อสอบข้อ i ที่แสดงถึงระดับความสามารถที่จุดโค้งลักษณะข้อสอบมีความชันที่สุด ในทางปฏิบัติจะมีค่าอยู่ระหว่าง -2.5 ถึง +2 .5

D คือ ค่า Scaling factor = 1.7

e คือ ค่าคงที่ซึ่งมีค่าประมาณ 2.71828

2. รูปแบบโลจิสติกที่มี 2 พารามิเตอร์ (Two-Parameter Logistic Model) รูปแบบนี้ถือว่าข้อสอบทุกข้อไม่มีค่าการเดา คำนึงถึงค่าความยากและค่าอำนาจจำแนกของข้อสอบเท่านั้น ดังสมการ

$$P_i(\theta) = \frac{e^{D a_i(\theta - b_i)}}{1 + e^{D a_i(\theta - b_i)}}$$

$$P_i(\theta) = \left[1 + e^{-D a_i(\theta - b_i)}\right]^{-1}$$

เมื่อ a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อ i ที่เป็นสัดส่วนโดยตรงกับความชันของโค้งลักษณะข้อสอบ ซึ่งมีค่าได้ตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติจะใช้ค่าที่อยู่ระหว่าง 0.5 ถึง +2.5

3. รูปแบบโลจิสติกที่มี 3 พารามิเตอร์ (Three-Parameter Logistic Model) รูปแบบนี้คำนึงถึงค่าความยาก ค่าอำนาจจำแนกและค่าการเดา ดังสมการ

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

หรือ $$P_i(\theta) = c_i + (1 - c_i)\left[1 + e^{-Da_i(\theta - b_i)}\right]^{-1}$$

เมื่อ c_i คือ ความน่าจะเป็นที่ผู้สอบซึ่งมีความสามารถต่ำมากจะตอบข้อสอบข้อ i ได้ถูก ในทางปฏิบัติจะเลือกข้อสอบที่มีค่าการเดา < 0.3

4. รูปแบบโลจิสติกที่มี 4 พารามิเตอร์ (Four-Parameter Logistic Model) รูปแบบนี้คำนึงถึงความสะเพร่าในการตอบข้อสอบด้วย นอกเหนือจาก 3 พารามิเตอร์ที่กล่าวมาแล้ว โดยเพิ่มพารามิเตอร์ 1 ตัวคือ γ_i ซึ่งมีค่าต่ำกว่า 1 เล็กน้อย ซึ่งเป็นค่าที่ผู้สอบที่มีความสามารถสูงมีความสะเพร่าตอบข้อสอบไม่ถูกแต่รูปแบบนี้ยังไม่สามารถนำมาใช้ในทางปฏิบัติได้

$$P_i(\theta) = c_i(\gamma_i - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

$$P_i(\theta) = c_i + (\gamma_i - c_i)\left[1 + e^{-Da_i(\theta - b_i)}\right]^{-1}$$

การหาคุณภาพของแบบสอบตามแนวทฤษฎีการตอบสนองข้อสอบ
(Item Response Theory : IRT)

การหาคุณภาพของแบบสอบตามแนวทฤษฎีการตอบสนองข้อสอบ จะพิจารณาจาก **ฟังก์ชันสารสนเทศของแบบสอบ** (Test Information Function : TIF : $I(\theta)$) ซึ่งเป็นดัชนีชี้ให้เห็นคุณภาพในการวัดของแบบสอบ โดยแบบสอบจะมีประสิทธิภาพสูงสุดที่ระดับความสามารถที่ให้ฟังก์ชันสารสนเทศแบบสอบสูงสุดซึ่งฟังก์ชันสารสนเทศของแบบสอบ (TIF : $I(\theta)$) หาได้จากผลรวมของฟังก์ชันสารสนเทศของข้อสอบ(Item Information Function : $I\{\theta, u_i\}$) โดยที่

$$I\{\theta, u_i\} = \frac{D^2 a_i^2 (1 - c_i)}{\left[c_i + e^{Da_i(\theta - b_i)}\right]\left[c_i + e^{-Da_i(\theta - b_i)}\right]}$$

เขียนให้อยู่ในรูปอย่างง่ายคือ

$$I\{\theta, u_i\} = D^2 a_i^2 (1 - c_i) / (c_i + e^{DL})(c_i + e^{-DL})^2$$

เมื่อ $D = 1.7$, $L = a_i(\theta - b_i)$

หรือ

$$I\{\theta, u_i\} = \frac{P_i'^2(\theta)}{P_i(\theta)Q_i(\theta)}$$

เมื่อ $P_i'(\theta)$ = อนุพันธ์ที่ 1 (First Derivative) ของ $P_i(\theta)$

$$P_i'(\theta) = \frac{Da_i(1-c_i)}{e^{Da_i(\theta-b_i)}+2+e^{-Da_i(\theta-b_i)}}$$

โดยที่ $Q_i(\theta)$ = $1 - P_i(\theta)$

ฟังก์ชันสารสนเทศของแบบสอบหาได้ดังนี้

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

หรือ

$$I(\theta) = \sum_{i=1}^{n} P_i'^2(\theta) / P_i(\theta)Q(\theta)$$

พบว่าค่าฟังก์ชันสารสนเทศของแบบสอบขึ้นอยู่กับความชันของโค้งลักษณะแบบสอบ และความแปรปรวนของการตอบข้อสอบถูกของแต่ละข้อในแต่ละระดับความสามารถ θ ถ้าความชันมีค่ามากและความแปรปรวนมีค่าน้อย ทำให้โค้งสารสนเทศของข้อสอบ ณ ระดับความสามารถนั้น ๆ จะมีค่าสูง ค่าฟังก์ชันสารสนเทศของแบบสอบในแต่ละฉบับจะมีได้หลายค่าแตกต่างกันไปตามระดับความสามารถ θ และ ณ ระดับความสามารถใดที่ให้ค่าฟังก์ชันสารสนเทศของแบบสอบสูงที่สุดแสดงว่าแบบสอบจะมีคุณภาพดีในการประมาณค่า ณ ระดับความสามารถนั้น ๆ และเมื่อแบบสอบมีจำนวนมากข้อ จะทำให้ได้ค่าฟังก์ชันสารสนเทศสูงขึ้นตามด้วย (Hambleton and Swaminathan 1985 : 107)

วอร์ม (Worm, 1978 :76-77) ได้หาความสัมพันธ์ของความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard Error of Estimate : SEE) กับฟังก์ชันสารสนเทศของแบบสอบ โดยกล่าวว่าถ้าการประมาณค่าความสามารถ (θ) เป็นการประมาณค่าความน่าจะเป็นสูงสุด (Maximum Likelihood of Estimate) แล้ว แต่ละค่าความสามารถ θ ก็สามารถหาความคลาดเคลื่อนมาตรฐานได้จากฟังก์ชันสารสนเทศของแบบสอบ I (θ) ดังนี้

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

เนื่องจากค่า I (θ) แปรผันไปตามมาตรวัดความสามารถ (θ) จึงเกิดค่าความคลาดเคลื่อนมาตรฐานของการประมาณ(SEE) ขึ้นยิ่ง I (θ) มีค่ามาก SEE ก็ยิ่งมีค่าน้อย และSEE จะมีค่าน้อยที่จุดตัด วอร์มได้หาความสัมพันธ์ของค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานของการประมาณ $(\overline{see})$ กับค่าความเที่ยง (r) ดังนี้

$$r = 1 - \overline{see}^2$$

ซึ่งความสัมพันธ์นี้ให้ความหมายได้ชัดเจน เมื่อความเที่ยงได้จากการกระจายความสามารถ (θ) ถ้าผู้สอบทั้งหมดกระจายบนมาตราวัดความสามารถที่มีสารสนเทศ (Information) สูงแล้วค่าความเที่ยงย่อมจะสูงกว่าความสามารถที่ให้สารสนเทศต่ำ

ตอนที่ 2 การกำหนดจุดตัดด้วยวิธีของวิลคอกซ์กับวิธีของแองกอฟโดยการประยุกต์ทฤษฎีการตอบสนองข้อสอบ และความหมายของคะแนนโดเมน

แฮมเบิลตันและกรุจเตอร์(Hambleton and Gruijter 1983 : 355 – 366) ได้ประยุกต์ทฤษฎีการตอบสนองตอบข้อสอบกับการคัดเลือกข้อสอบอิงเกณฑ์และได้คำนวณคะแนนจุดตัดของแบบสอบอิงเกณฑ์ด้วยวิธีวิลคอกซ์โดยการประยุกต์ทฤษฎีการตอบสนองข้อสอบ

<u>การกำหนดคะแนนจุดตัดตามวิธีของวิลคอกซ์</u> มีขั้นตอนดังต่อไปนี้

ก. หาคะแนนโดเมนบนทุกค่ามาตราวัดความสามารถ จากการสุ่มข้อกระทงในโดเมนแต่ละจุดประสงค์ให้มีจำนวนข้อที่ผู้เชี่ยวชาญกำหนดน้ำหนักความสำคัญของเนื้อหา แล้วหาคะแนนโดเมนโดยใช้สถิติของทฤษฎีการตอบสนองข้อสอบดังนี้

$$\pi = \frac{1}{m} \sum_{i=1}^{m} P_i(\theta)$$

π คือ คะแนนโดเมน

m คือ จำนวนข้อในแบบสอบ

$P_i(\theta)$ คือ ความน่าจะเป็นที่จะตอบข้อที่ i ถูกต้องของผู้ที่มีความสามารถ(θ)

โดยที่

$$P_i(\theta) = c_i + \frac{1-c_i}{1+e^{-1.7(\theta - b_i)}} \quad \text{เมื่อ } \theta = -3,-2.9,\ldots,3$$

ข. การคัดเลือกข้อกระทงที่ใช้หาคะแนนจุดตัด โดยพิจารณาจากค่า θ, π ที่ได้จากข้อ ก. นำมาหาค่า θ เมื่อกำหนดคะแนนโดเมนเป็น .5 ซึ่งเป็นเกณฑ์ขั้นต่ำในการผ่าน การกำหนดคะแนนโดเมนดังกล่าวเพื่อให้สอดคล้องกับแนวทางการประเมินผลการเรียนตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2533

ค. การหาคะแนนจุดตัด มีขั้นตอนดังนี้

การกำหนดจุดตัดโดยการประยุกต์ของวิลคอกซ์(Wilcox, 1976) มีขั้นตอนดังต่อไปนี้

1.สร้างข้อสอบไว้จำนวนมาก

2.หาค่าพารามิเตอร์ของข้อสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ

3.ตรวจสอบว่าข้อมูลที่ได้สอดคล้องกับโค้งลักษณะข้อสอบ(Item Characteristic Curve) ตามทฤษฎีการตอบสนองข้อสอบหรือไม่

4.สุ่มข้อสอบให้ได้จำนวนข้อที่ต้องการตามสัดส่วนความสำคัญของแต่ละจุดประสงค์

5.กำหนดคะแนนโดเมน(π)ที่เป็นจุดตัดที่ต้องการวัด

6.แปลงคะแนนโดเมน(π)ให้เป็นค่าความสามารถ(θ)โดยใช้สมการซึ่งแสดงความสัมพันธ์ระหว่างคะแนนโดเมน(π)กับความสามารถ(θ)คือ

$$\pi = \frac{1}{m} \sum_{i=1}^{m} P_i(\theta)$$

π คือ คะแนนโดเมน

m คือ จำนวนข้อในกลุ่มตัวอย่างของข้อสอบ

$P_i(\theta)$ คือ ความน่าจะเป็นที่ตอบข้อที่ i ถูกต้องของผู้ที่มีความสามารถ θ โดยที่ θ = -3,-2.9,-2.8,...,3

7.เลือกข้อสอบที่จะให้ค่าใกล้ค่า θ_0 โดยพิจารณาความสัมพันธ์จากสมการดังนี้

$$\theta_i^* = b_i + \frac{1}{Da_i} \log_e 0.5(1 + \sqrt{1 + 8c_i})$$

D = 1.7

a คือ ค่าอำนาจจำแนกที่ได้จากการวิเคราะห์ข้อสอบ

b คือ ค่าความยาก

c คือ ค่าการเดา

θ_i^* คือ ความสามารถที่ให้ค่า Item Information Curve $[I_i(\theta_0)]$ สูงสุด

โดยที่ $I(\theta_0)$ เป็นดัชนีบอกคุณภาพของข้อสอบ ถ้า $I\{\theta,u\}$ค่าสูงมากเพียงใดก็แสดงว่าข้อสอบข้อนั้นมีคุณภาพในการชี้บอกผู้รอบรู้และไม่รอบรู้มากที่ $\theta = \theta_0$ หรือใกล้เคียง

8.แปลงค่า θ_0 ให้เป็น π_0 โดยใช้โค้งลักษณะแบบสอบที่มี m ข้อ

9.คำนวณคะแนนจุดตัด พิจารณาค่าที่เป็นจำนวนเต็มจากความสัมพันธ์ดังนี้

$$\text{คะแนนจุดตัด} = \sum_{i=1}^{m} P_i(\theta)$$

แนวคิดการกำหนดจุดตัดด้วยวิธีแองกอฟโดยการประยุกต์ทฤษฎีการตอบสนองข้อสอบ

เคน (Kane 1987 : 333 – 345) ได้รวบรวมวิธีการแปลงคะแนนจุดตัดที่ได้จากวิธีการกำหนดคะแนนจุดตัดโดยใช้ดุลพินิจของผู้เชี่ยวชาญตัดสิน ให้เป็นระดับความสามารถขั้นต่ำสุดที่ยอมให้ผ่าน (θ^*) โดยมีข้อตกลงว่า จะต้องทราบค่าโค้งลักษณะข้อสอบ ($P_i(\theta)$) ซึ่งเป็นหลักของทฤษฎีการตอบสนองข้อสอบในการประมาณความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ θ ต่างๆ จะสามารถตอบข้อสอบได้ถูก ซึ่งมีอยู่ 3 วิธีคือ

<u>วิธีที่ 1</u> ประมาณค่า θ^* ของข้อสอบแต่ละข้อโดยอาศัยโค้งลักษณะข้อสอบรายข้อ แล้วจึงคำนวณหา θ^* เฉลี่ยของข้อสอบทุกข้อในแบบสอบ มี 3 ขั้นตอน ดังนี้

1. คำนวณหาค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบ 1 ได้ถูกที่ใช้ดุลพินิจของผู้เชี่ยวชาญตัดสิน (M_{iR})

2. ให้ M_{iR} เป็นค่าประมาณของความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถขั้นต่ำสุดที่ยอมให้ผ่าน สามารถตอบข้อสอบได้ถูกในรูปแบบโลจิสติก ซึ่งแทนได้ด้วยโค้งลักษณะข้อสอบที่ระดับความสามารถ θ^*_{iR} ดังสมการ

$$M_{iR} = P_i(\theta^*_{iR})$$

3. ประมาณค่าความแปรปรวนของ θ^*_{iR} ในข้อ i

กำหนดให้ $P_i(\theta^*_{iR}) = \frac{\Delta M_{iR}}{\Delta \theta_{iR}}$ (6)

เป็นความชันของโค้งลักษณะข้อสอบ ข้อ i ที่ความสามารถ (θ^*_{iR})

แทนด้วยความเบี่ยงเบนมาตรฐาน

$$\frac{P_i'(\theta_{iR}^*)}{\sigma_i(M_{iR})} \cong \frac{1}{\sigma_i(\theta_{iR}^*)} \quad (7)$$

$$\frac{[P_i'(\theta_{iR}^*)]^2}{\sigma^2{}_i(M_{iR})} \cong \frac{1}{\sigma_i^2(\theta_{iR}^*)} \quad (8)$$

4. ประมาณค่า θ^* จากค่าเฉลี่ยของ θ^*_{iR}

$$\theta^* = \frac{1}{n}\sum_i \theta_{iR}^* \quad (9)$$

ซึ่งมีความแปรปรวน $\sigma^2(\theta^*) = \frac{1}{n^2}\sum_i \sigma_i^2(\theta_{iR}^*)$ (10)

โดยมีข้อตกลงเบื้องต้นว่าการตัดสินความน่าจะเป็นในการตอบถูกของข้อสอบแต่ละข้อเป็นอิสระกัน แต่การที่ผู้ตัดสินคนหนึ่งต้องตัดสินข้อสอบทุกข้อในแบบสอบอาจทำให้ผลการตัดสินข้อสอบรายข้อไม่เป็นอิสระกัน จึงแทนสมการที่ (9) ด้วยการประมาณแบบถ่วงน้ำหนักดังนี้

$$\theta_w^* = \sum_i w_i \theta_{iR}^* \quad (11)$$

เมื่อ $\sum_i w_i = 1$ (12)

และมีความแปรปรวนเป็น

$$\sigma^2(\theta_w^*) = \sum_i w_i^2 \sigma_i^2(\theta_{iR}^*) \quad (13)$$

ค่าของ W_i ที่จะทำให้ความแปรปรวนในสมการที่ (13) มีค่าน้อยที่สุดและมีคุณสมบัติดังในสมการที่ (12) ทำได้โดยการใช้ Lagrange multipliers

$$w_i = \frac{1}{\sigma_i^2(\theta_{iR}^*)} \Big/ \sum_i \frac{1}{\sigma_i^2(\theta_{iR}^*)} \quad (14)$$

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การประมาณแบบถ่วงน้ำหนักนี้จะช่วยลดความคลาดเคลื่อนมาตรฐานของ $\theta^{\cdot}$ ข้อสอบที่ผู้ตัดสิน ตัดสินคล้าย ๆ กัน จะถูกถ่วงน้ำหนักมาก ส่วนข้อที่ผู้ตัดสิน ตัดสินต่างกันออกไปก็จะถูกถ่วงน้ำหนักน้อย

<u>วิธีที่ 2</u> เป็นวิธีที่ใช้กันมากที่สุด เป็นการประมาณ $\theta^{\cdot}$ โดยอาศัยโค้งลักษณะแบบสอบ มี 3 ขั้นตอน คือ

1. คำนวณหาค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อ i ถูก ที่ตัดสินโดยผู้ตัดสินทั้งหมด (M_{iR})

2. คำนวณหาคะแนนจุดตัดของแบบสอบซึ่งอยู่บนสเกลของคะแนนจริงจากผลรวมของ M_{iR} ในแต่ละข้อของแบบสอบ

3. ใช้โค้งลักษณะแบบสอบ (test characticsistic curve) หาค่าระดับความสามารถ θ ซึ่งสัมพันธ์กับคะแนนจริงในขั้นที่ 2 จากสมการ

$$\sum M_{iR} = \sum P_i(\theta^{\cdot}) \qquad (15)$$

$$\sum_i [M_{iR} - P_i(\theta^{\cdot})] = 0 \qquad (16)$$

จากสมการที่ (5) $\sum [P_i(\theta^{\cdot}_{iR}) - P_i(\theta^{\cdot})] = 0$ (17)

แต่ $\theta^{\cdot}_{iR} \rightarrow \theta^{\cdot}$ ความสัมพันธ์สามารถประมาณได้ด้วย linear function (Lord 1980 : 69)

$$\sum P_i(\theta^{\cdot}_{iR})(\theta^{\cdot}_{iR} - \theta^{\cdot})] \cong 0 \qquad (18)$$

$$\theta^{\cdot} \cong \frac{1}{\sum P'_i(\theta^{\cdot}_{iR})} \sum P'_i(\theta^{\cdot}_{iR})\theta^{\cdot}_{iR} \qquad (19)$$

จากสมการที่ (7) และ (19)

$$\theta^{\cdot} \cong \frac{1}{\sum \sigma_i(M_{iR})/\sigma_i(\theta^{\cdot}_{iR})} \sum_i \frac{\sigma(M_{iR})}{\sigma_i(\theta^{\cdot}_{iR})} \theta^{\cdot}_{iR} \qquad (20)$$

การประมาณ $\theta^{\cdot}$ จากสมการที่ 20 ยังไม่ใช่ค่าที่ดีที่สุด สามารถปรับปรุงได้โดยการถ่วงน้ำหนักด้วย $1 / \sigma_i(M_{iR})$ ในสมการที่ 17 จะได้

$$\theta^{\cdot} \cong \frac{1}{\sum_i [P'_i(\theta^{\cdot}_{iR})/\sigma_i(M_{iR})]} \sum_i \frac{P'_i(\theta^{\cdot}_{iR})}{\sigma_i(M_{iR})} \theta^{\cdot}_{iR} \qquad (21)$$

จากสมการที่ 7 จะได้ว่า

$$\theta^{*}_{w} \cong \frac{1}{\sum_{i}\left[1/\sigma_{i}\left(\theta^{*}_{iR}\right)\right]}\sum_{i}\frac{1}{\sigma_{i}\left(\theta^{*}_{iR}\right)}\theta^{*}_{iR} \qquad (22)$$

<u>วิธีที่ 3</u> เป็นการประมาณ θ^{*} ที่เหมาะสมที่สุดระหว่างการใช้ดุลพินิจในการตัดสินความน่าจะเป็นในการตอบข้อสอบได้ถูก (M_{ir}) กับโค้งลักษณะข้อสอบ มี 3 ขั้นตอน คือ

1. คำนวณหาค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อ i ถูก ที่ตัดสินโดยผู้ตัดสินทั้งหมด (M_{iR})

2. คำนวณหาความแปรปรวนของค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบ i ถูก ที่ตัดสินโดยผู้ตัดสินทั้งหมด σ_{i}^{2} (M_{iR})

3. คำนวณหา θ^{*} ซึ่งมีผลบวกของความเบี่ยงเบนกำลังสองระหว่าง M_{iR} และ P_{i} (θ^{*}) มีค่าน้อยที่สุดจากสมการ

$$\sum_{i}\frac{\left(M_{iR}-P_{i}\left(\theta^{*}\right)\right)^{2}}{\sigma^{2}\left(M_{iR}\right)} = \sum_{i}\frac{\left[P_{i}\left(\theta^{*}_{iR}\right)-P_{i}\left(\theta^{*}\right)\right]^{2}}{\sigma_{i}^{2}\left(M_{iR}\right)} \qquad (23)$$

แต่ θ^{*}_{iR} มีค่าใกล้เคียงกับ θ^{*} และสมการที่ 18 จะได้

$$\sum_{i}\frac{\left(M_{iR}-P_{i}\left(\theta^{*}\right)\right)^{2}}{\sigma^{2}\left(M_{iR}\right)} \cong \sum_{i}\frac{1}{\sigma_{i}^{2}\left(M_{iR}\right)}\left[P_{i}'\left(\theta^{*}_{iR}\right)\left(\theta^{*}_{iR}-\theta^{*}\right)\right]^{2}$$

$$\cong \sum_{i}\frac{P_{i}'^{2}\left(\theta^{*}_{iR}\right)}{\sigma_{i}^{2}\left(M_{iR}\right)}\left(\theta^{*}_{iR}-\theta^{*}\right)^{2} \qquad (24)$$

กำหนดให้สมการที่ 24 เท่ากับ 0 จะได้

$$\sum_{i}\frac{P_{i}'^{2}\left(\theta^{*}_{iR}\right)}{\sigma_{i}^{2}\left(M_{iR}\right)}\left(\theta^{*}_{iR}-\theta^{*}\right) \cong 0 \qquad (25)$$

เมื่อแก้สมการแล้วหาค่า θ^{*} จะได้

$$\theta^{*} \cong \frac{1}{\sum_{i}\left[P_{i}'^{2}\left(\theta^{*}_{iR}\right)/\sigma_{i}^{2}\left(M_{iR}\right)\right]}\sum_{i}\frac{P_{i}'^{2}\left(\theta^{*}_{iR}\right)}{\sigma_{i}^{2}\left(M_{iR}\right)}\theta^{*}_{iR} \qquad (26)$$

จากสมการที่ 8 จะได้ว่า

$$\theta^{\cdot} \cong \frac{1}{\sum_{i}\left[1/\sigma_i^2(\theta_{iR}^{\cdot})\right]} \sum_{i} \frac{1}{\sigma_i^2(\theta_{iR}^{\cdot})}\theta_{iR}^{\cdot} \qquad (27)$$

สรุปวิธีการประมาณ $\theta^{\cdot}$ ของทั้ง 3 วิธี จะเห็นว่า วิธีที่ 1 และ 3 จะอาศัยผลการสอบของข้อสอบแต่ละข้อประกอบกับการใช้ดุลพินิจของผู้เชี่ยวชาญ ส่วนวิธีที่ 2 จะหาคะแนนจุดตัดจากดุลพินิจของผู้เชี่ยวชาญก่อนแล้วจึงอาศัยโค้งลักษณะแบบสอบเพื่อแปลงคะแนนจุดตัดในสเกลของคะแนนโดเมนให้อยู่บนสเกลของความสามารถซึ่งเป็นวิธีที่นิยมกัน ดังนั้นในการวิจัยครั้งนี้ ผู้วิจัยจึงกำหนดจุดตัดด้วยวิธีที่ 2 เพราะต้องการคะแนนจุดตัดจากการตัดสินของผู้เชี่ยวชาญแล้วจึงแปลงให้เป็นระดับความสามารถ $\theta^{\cdot}$

<u>การกำหนดคะแนนจุดตัดตามวิธีของแองกอฟ(Angoff 1971 :656)</u> มีขั้นตอนดังต่อไปนี้

1. นำผลการพิจารณาของผู้เชี่ยวชาญการสอนจำนวน 10 ท่านที่ได้ตัดสินความน่าจะเป็นที่นักเรียนซึ่งมีระดับความสามารถขั้นต่ำสุดที่จะยอมรับได้ตอบข้อสอบแต่ละข้อถูกมาเฉลี่ยรายข้อและรายโดเมนตามลำดับ จะได้เป็นคะแนนเกณฑ์(π_0) ของแต่ละโดเมน ซึ่งอยู่บนสเกลของคะแนนจริงหรือคะแนนโดเมน(π)

2.แปลงคะแนนโดเมนที่เป็นคะแนนเกณฑ์(π_0) ของแต่ละโดเมนซึ่งอยู่บนสเกลของคะแนนโดเมนให้อยู่บนสเกลของระดับความสามารถ ($\theta^{\cdot}$) โดยอาศัยโค้งลักษณะแบบสอบ(Test Characteristic Curve) เนื่องจากโค้งลักษณะแบบสอบได้จากกความสัมพันธ์ของคะแนนโดเมน(π) กับระดับความสามารถ(θ) ซึ่งได้จากสูตร

$$\pi = \frac{1}{m} \sum_{i=1}^{m} P_i(\theta)$$

และเนื่องจาก $\sum M_{iR} = \sum P_i(\theta^{\cdot})$

เมื่อ M_{iR} คือ ค่าเฉลี่ยของค่าการตัดสินระดับผ่านต่ำสุด(minimum pass level) ของผู้เชี่ยวชาญ k คน บนข้อสอบแต่ละข้อ

$P_i(\theta^{\cdot})$ คือ ค่าความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถขั้นต่ำสุดที่ยอมรับว่าผ่านจะตอบข้อสอบข้อนั้นได้ถูกต้อง

จากความสัมพันธ์ดังกล่าว ดังนั้นจึงสามารถแปลงคะแนนเกณฑ์ที่อยู่ในรูปคะแนนโดเมน(π_0) ให้อยู่ในรูปของระดับความสามารถขั้นต่ำสุดที่ยอมรับว่าผ่านหรือคะแนนจุดตัด ได้จากสูตรดังต่อไปนี้

$$\theta^{*} \cong \frac{1}{\sum_{i}\left[P_i'\left(\theta^{*}_{iR}\right)/\sigma_i\left(M_{iR}\right)\right]} \sum_{i} \frac{P_i'\left(\theta^{*}_{iR}\right)}{\sigma_i\left(M_{iR}\right)} \theta^{*}_{iR}$$

เมื่อ $P_i(\theta_{iR}^{*})$ คือ ค่า slope ของโค้งลักษณะข้อสอบสำหรับข้อสอบข้อที่ i ที่ระดับความสามารถขั้นต่ำสุดที่ยอมรับว่าผ่าน θ_{iR}^{*} โดยที่

$$P_i'\left(\theta_{iR}^{*}\right) = \frac{Da_i\left(1-c_i\right)}{e^{Da_i\left(\theta_{iR}^{*}-b_i\right)} + 2 + e^{-Da_i\left(\theta_{iR}^{*}-b_i\right)}}$$

และ $\sigma_i(M_{iR})$ คือ ค่าความคลาดเคลื่อนมาตรฐานของค่าการตัดสินระดับผ่านต่ำสุดโดยผู้เชี่ยวชาญ k คน บนข้อสอบข้อ iโดยที่

$$\sigma_i\left(M_{iR}\right) = \sqrt{\frac{\sum_{r=1}^{k}\left(M_{ir}-M_{iR}\right)^2}{k-1}}$$

M_{ir} คือ ค่าการตัดสินระดับผ่านต่ำสุดของข้อสอบข้อ iโดยผู้เชี่ยวชาญคนที่ r

M_{iR} คือ ค่าเฉลี่ยของค่าการตัดสินระดับผ่านต่ำสุดของข้อสอบข้อ I โดยผู้เชี่ยวชาญ k คน

ความหมายของคะแนนโดเมน

บอค ทิสเซน และไซโมวสกี (Bock R.D. , David Tissen & Michele F. Zimowski 1997,197 – 221) ได้ทำการศึกษาการใช้ทฤษฎีการตอบสนองข้อสอบ(Item Response Theory : IRT) ในการประมาณคะแนนโดเมนพบว่าการประมาณคะแนนโดเมนโดยใช้ทฤษฎีการตอบสนองข้อสอบมีความถูกต้องกว่าการประมาณคะแนนโดเมนโดยใช้ทฤษฎีแบบดั้งเดิม (Classical Test Theory : CTT)

คะแนนโดเมน(π) หมายถึง สัดส่วนของจำนวนข้อที่สอบถูกต่อจำนวนข้อสอบทั้งหมด เช่น 1.00 หมายถึง ผู้สอบตอบแบบสอบถูก 100 % จะอ้างอิงไปสู่ปริมาณ 100 % ของปริเขตของการสอบวัดนั้น ซึ่งคะแนนโดเมนนี้จะเป็นตัวคาดคะเนระดับผลสัมฤทธิ์ทางการเรียนที่ปราศจากอคติ (Unbias estimator) และมีความพอเพียง (sufficient statistic) ก็ต่อเมื่อต้องมีการระบุเขตของข้อสอบที่เป็นไปได้ในแต่ละโดเมนไว้ล่วงหน้า และข้อสอบในแบบสอบอิงโดเมนต้องเป็นกลุ่มตัวอย่างแบบสุ่มของข้อสอบในโดเมนนั้น โดยที่แบบสอบนั้นมีลักษณะเป็นเอกพันธ์ (Homogeneity) และแบบสอบอิงโดเมน เป็นแบบสอบที่มีคุณสมบัติในเชิงเอกพันธ์สูงซึ่งสามารถใช้คะแนนรวมเป็นตัวประมาณค่าผลสัมฤทธิ์ทางการเรียนหรือลักษณะ (Trait) ที่มุ่งวัดได้ (สมหวัง พิธิยานุวัฒน์ 2529 : 18-35)ซึ่งค่าพารามิเตอร์ของข้อสอบหลายข้อที่อยู่ในโดเมนเดียวกัน หรือค่าพารามิเตอร์ของข้อสอบชุดหนึ่งที่คิดว่าเป็นตัวแทนที่ดีของโดเมนจะทำให้สามารถหาความสัมพันธ์ระหว่างคะแนนโดเมนในเชิงทฤษฎีการตอบสนองข้อสอบ คำนวณจากเปอร์เซ็นต์ของผลรวมของโอกาสตอบถูกรายข้อของข้อสอบในแต่ละค่าความสามารถ (θ) ซึ่งคะแนนโดเมนคิดจาก

$$\pi = \frac{1}{m} \sum_{i=1}^{m} P_i(\theta)$$

เมื่อ π คือ คะแนนโดเมน

m คือ จำนวนข้อในกลุ่มตัวอย่างข้อสอบ

$P_i(\theta)$ คือ โอกาสที่ผู้สอบซึ่งมีความสามารถตรงระดับความสามารถ (θ) จะทำข้อสอบข้อที่ i ได้

ตอนที่ 3 การประมาณค่าความเที่ยงและความตรงของการกำหนดจุดตัด

ความตรงของการกำหนดจุดตัดด้วยวิธีวิลคอกซ์กับวิธีของแองกอฟโดยการประยุกต์ทฤษฎีการตอบสนองข้อสอบ สำหรับแบบสอบอิงโดเมนประเภทเลือกตอบและตอบสั้นจะพิจารณาจาก ความตรงของแบบสอบอิงเกณฑ์ โดยฟิซแพทริก (Fitxpatrick, 1981) ได้เสนอว่าความตรงควรมี 3 ประเภทคือ ความตรงเชิงเนื้อหา(Content Validity) ความตรงเชิงทฤษฎี(Construct Validity) ความตรงในการตัดสิน(Decicion Validity)

ความตรงเชิงเนื้อหา(Content Validity) หมายถึง การตัดสินเกี่ยวกับความสอดคล้องกันระหว่างเนื้อหาของข้อสอบกับจุดประสงค์ที่มุ่งวัด ความตรงเชิงเนื้อหาจะไม่ผันแปรไปตามเวลาหรือกลุ่มตัวอย่าง

ความตรงเชิงทฤษฎี(Construct Validity) หมายถึง ความสามารถของข้อสอบในการวัดสิ่งที่ต้องการจะวัดได้ถูกต้องตรงตามทฤษฎีที่กำหนดไว้ อาจหาความตรงได้จากการทดลองเชิงประจักษ์กับกลุ่มตัวอย่างที่มีความรอบรู้ โดยกลุ่มที่มีความรอบรู้จะตอบข้อสอบข้อนั้นถูกต้องและกลุ่มตัวอย่างที่ไม่รอบรู้จะตอบข้อสอบข้อนั้นไม่ถูกต้อง

ความตรงในการตัดสิน(Decision Validity) หมายถึง ความถูกต้องในการตัดสินใจจำแนกผู้สอบออกเป็นผู้รอบรู้และไม่รอบรู้ได้ตรงตามจุดมุ่งหมายในการวัด

จะเห็นได้ว่าความตรงเชิงทฤษฎีและความตรงในการตัดสินใจมีความใกล้เคียงกันมาก ดังนั้นการวิจัยครั้งนี้จึงหาความตรงเชิงเนื้อหาและความตรงในการตัดสินเท่านั้น

การประมาณค่าความตรง

1.ความตรงเชิงเนื้อหา(Content Validity) คำนวณโดยการหาดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์(IOC) ดังนี้

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

IOC คือ ดัชนีความสอดคล้องระหว่างข้อกระทงกับจุดประสงค์

R คือ ผลรวมของคะแนนความคิดเห็นของผู้เชี่ยวชาญเนื้อหาวิชาทั้งหมด

N คือ จำนวนผู้เชี่ยวชาญเนื้อหาวิชาทั้งหมด

2.ความตรงในการตัดสิน (Decision Validity) ของการกำหนดจุดตัด คำนวณได้โดยวิธีของคาร์เวอร์ (Caver, 1974) ซึ่งเป็นการคำนวณดัชนีความตรงของแบบสอบจากข้อมูลของกลุ่มผู้เรียนและยังไม่ได้เรียนดังนี้

ดัชนีความตรง = (a + c) / (a + b + c + d)

เมื่อ a คือ กลุ่มที่ผ่านเกณฑ์หลังจากได้รับการเรียนการสอน

b คือ กลุ่มที่ผ่านเกณฑ์โดยไม่ได้รับการเรียนการสอน

c คือ กลุ่มที่ไม่ผ่านเกณฑ์และไม่ได้รับการเรียนการสอน

d คือ กลุ่มที่ไม่ผ่านเกณฑ์หลังจากได้รับการเรียนการสอน

การประมาณค่าความเที่ยงของการกำหนดจุดตัดจะพิจารณาจากความเที่ยงของแบบสอบอิงโดเมนซึ่งในการวิจัยครั้งนี้ใช้แนวคิดจากการประมาณความเที่ยงของแบบสอบอิงเกณฑ์เบอร์ก(Berk, 1980 : 127) ได้สรุปแนวคิดเกี่ยวกับความเที่ยงของแบบสอบอิงเกณฑ์ไว้ 3 มโนทัศน์ดังนี้

1.คือความคงที่ในการตัดสินจำแนกผู้รอบรู้ และไม่รอบรู้โดยการสอบซ้ำ ด้วยแบบฟอร์มเดียวกันหรือแบบสอบคู่ขนาน

2.คือความคงที่ของความแตกต่างกำลังสองของคะแนนแต่ละคนที่เบี่ยงเบนไปจากคะแนนจุดตัดโดยการสอบด้วยแบบสอบคู่ขนาน

3.คือความคงที่ของคะแนนแต่ละคนที่แตกต่างกันจากการสอบด้วยแบบสอบคู่ขนาน

จากข้อสรุปมโนทัศน์ความเที่ยงของแบบสอบอิงเกณฑ์ทั้ง 3 ข้อข้างต้น มโนทัศน์ที่เกี่ยวข้องกับจุดตัดอย่างชัดเจนได้แก่มโนทัศน์ในข้อที่ 1 และ 2 ส่วนมโนทัศน์ที่ 3 ไม่ได้กล่าวถึงจุดตัด และในการวิจัยครั้งนี้ผู้วิจัยได้เลือกการหาค่าความเที่ยงโดยการสอบซ้ำ ด้วยแบบฟอร์มเดียวกันดังนั้นผู้วิจัยจึงใช้การประมาณค่าความเที่ยงตามมโนทัศน์ที่ 1 เท่านั้น

การประมาณค่าความเที่ยงของแบบสอบอิงเกณฑ์โดยพิจารณาจากความคงที่ในการตัดสินจำแนกผู้รอบรู้และไม่รอบรู้โดยการสอบซ้ำ ด้วยแบบฟอร์มเดียวกัน มีสูตรคำนวณดังนี้

1.วิธีของแฮมเบิลตัน และโนวิค(Hambleton and Novick,1973)ได้เสนอการประมาณค่าความเที่ยงโดยใช้ผลรวมของสัดส่วนความสอดคล้องในการตัดสินใจ จำแนกผู้รอบรู้จากการสอบด้วยแบบสอบฉบับเดียวกันสองครั้ง หรือจากการสอบด้วยแบบสอบคู่ขนาน ดังต่อไปนี้

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

เมื่อ P_0 คือ ความเที่ยงของแบบสอบอิงเกณฑ์

P_{11} คือ สัดส่วนของผู้ถูกตัดสินว่ารอบรู้ตรงกันทั้งสองครั้ง

P_{22} คือ สัดส่วนของผู้ถูกตัดสินไม่รอบรู้ตรงกันทั้งสองครั้ง

2.การประมาณค่าความเที่ยงของแบบสอบอิงเกณฑ์โดยวิธีของสวามินาทาน แฮมเบิลตัน และอัลจินา(Swaminathan,Hambleton and Angina,1974) มีการหาค่าความเที่ยงคล้าย ๆ กับวิธีของแฮมเบิลตันและโนวิค(Hambleton and Novick,1973) แต่มีการปรับแก้โดยการหักค่าความสอดคล้องโดยบังเอิญออกไปในการวิจัยครั้งนี้ได้ใช้การประมาณค่าความเที่ยงของแบบสอบอิงโดเมนตามวิธีของสวามินาทาน แฮมเบิลตัน และอัลจินา(Swaminathan,Hambleton and Angina,1974) ดังนี้

$$K = \frac{(P_0 - P_c)}{(1 - P_c)}$$

เมื่อ K คือ ค่าความเที่ยงของแบบสอบอิงโดเมนเมื่อหักความสอดคล้องโดยบังเอิญออก

P_0 คือ สัดส่วนความสอดคล้องในการตัดสินจำแนกผู้รอบรู้

P_c คือ สัดส่วนของความสอดคล้องที่คาดหวังซึ่งผลเกิดขึ้นโดยบังเอิญ

ซึ่งค่าของ P_0 และ P_c หาได้จากการคำนวณดังนี้

$$P_0 = \sum_{i=1}^{2} P_{ii}$$

$$P_c = \sum_{i,j=1}^{2} P_{i.} P_{.i}$$

โดยเมื่อนำผลการสอบมาแจกแจงลงในตารางจะได้ดังนี้

ทดสอบด้วยแบบสอบฉบับที่ 2

ทดสอบด้วยแบบสอบฉบับที่ 1		รอบรู้	ไม่รอบรู้	สัดส่วนคัดแยก
	รอบรู้	P_{11}	P_{12}	$P_{1.}$
	ไม่รอบรู้	P_{21}	P_{22}	$P_{2.}$
	สัดส่วนคัดแยก	$P_{.1}$	$P_{.2}$	

เมื่อ P_{11} คือ สัดส่วนของผู้สอบที่ถูกตัดสินว่ารอบรู้ตรงกัน 2 ฉบับ หรือ 2 ครั้ง

P_{12} คือ สัดส่วนของผู้สอบที่ถูกตัดสินว่ารอบรู้ในการสอบด้วยแบบสอบฉบับที่ 1 แต่ยังไม่รอบรู้ในการสอบด้วยแบบสอบฉบับที่ 2

P_{21} คือ สัดส่วนของผู้สอบที่ถูกตัดสินว่าไม่รอบรู้ในการสอบด้วยแบบสอบฉบับที่ 1 แต่รอบรู้ในการสอบด้วยแบบสอบฉบับที่ 2

P_{22} คือ สัดส่วนของผู้สอบที่ถูกตัดสินว่าไม่รอบรู้ในการสอบด้วยแบบสอบฉบับที่ 1 และแบบสอบฉบับที่ 2

$P_{1.}$ คือ สัดส่วนแยกของผู้สอบที่รอบรู้ในการสอบด้วยแบบสอบฉบับที่ 1

$P_{2.}$ คือ สัดส่วนแยกของผู้สอบที่ไม่รอบรู้ในการสอบด้วยแบบสอบฉบับที่ 1

$P_{.1}$ คือ สัดส่วนแยกของผู้สอบที่รอบรู้ในการสอบด้วยแบบสอบฉบับที่ 2

$P_{.2}$ คือ สัดส่วนแยกของผู้สอบที่ไม่รอบรู้ในการสอบด้วยแบบสอบฉบับที่ 2

จากการวิจัยในครั้งนี้ต้องการเปรียบเทียบคุณภาพของการกำหนดจุดตัดด้วยวิธีวิลคอกซ์กับวิธีแองกอฟโดยการประยุกต์ทฤษฎีการตอบสนองข้อสอบ ซึ่งจุดตัดของการกำหนดจุดตัดด้วยวิธีวิลคอกซ์คือจำนวนเต็มที่ได้จากผลรวมค่าความน่าจะเป็นในการทำข้อสอบแต่ละข้อถูกต้อง ส่วนจุดตัดของการกำหนดจุดตัดด้วยวิธีแองกอฟคือระดับความสามารถ(θ) ขั้นต่ำสุดที่จะยอมรับว่าผ่านได้ ซึ่งจุดตัดทั้ง 2 วิธีมีความต่างกันการที่จะนำมาเปรียบเทียบว่าจุดตัดใดดีกว่ากันจะต้องมาพิจารณาที่ความตรงและความเที่ยงของการกำหนดจุดตัด ซึ่งจากการวิจัยในครั้งนี้ได้เลือกใช้ความตรงจากวิธีของคาร์เวอร์ (Caver, 1974) เนื่องจากเป็นวิธีที่หาได้จากความสามารถของจุดตัดขอบแบบสอบที่คัดแยกผู้สอบออกเป็นผู้รอบรู้(Mastery)หรือไม่รอบรู้(Non-mastery) โดยใช้วิธีทดลองเชิงประจักษ์กับกลุ่มตัวอย่างที่มีความรอบรู้หรือได้รับการเรียนการสอนแล้วจะสามารถทำข้อสอบได้ถูกต้อง ส่วนกลุ่มตัวอย่างที่ไม่รอบรู้ หรือไม่ได้รับการเรียนการสอนจะตอบข้อสอบข้อนั้น ๆ ผิด เมื่อนักเรียนได้เรียนรู้จบเนื้อหาแล้วจะเป็นผู้รอบรู้และนักเรียนที่ยังไม่ได้เรียนถือว่าเป็นผู้ไม่รอบรู้ เมื่อทำการทดสอบแล้วความสามารถของจุดตัดของแบบสอบที่จะจำแนกได้ว่าผู้สอบเป็นผู้รอบรู้หรือไม่รอบรู้แสดงว่าแบบสอบมีความตรงของการกำหนดจุดตัดนั่นเองและความเที่ยงคำนวณจากวิธีของสวามินาทาน แฮมเบิลตัน และอัลจินา(Swaminathan,Hambleton and Angina,1974) เนื่องจากมีความสอดคล้องกับมโนทัศน์ในด้านความเที่ยงของเบอร์ก (Berk,R.A.1980:127) คือมีความคงที่ในการตัดสินจำแนกผู้รอบรู้ และไม่รอบรู้โดยการสอบซ้ำด้วยแบบฟอร์มเดียวกันหรือแบบสอบคู่ขนาน ซึ่งมีความชัดเจนและเหมาะสมกับคะแนนจุดตัดที่คำนวณด้วยวิธีวิลคอกซ์กับวิธีแองกอฟโดยการประยุกต์ทฤษฎีการตอบสนองข้อสอบ

ตอนที่ 4 แนวคิดแบบสอบแบบเลือกตอบกับตอบสั้น

แบบสอบแบบเลือกตอบ(Multiple Choice Test) ประกอบด้วย 2 ส่วนใหญ่ ๆ คือ

1.ตัวคำถาม (stem) หรือ ส่วนนำ หรือตัวปัญหา ส่วนนี้มีหน้าที่เสนอปัญหาและวางกรอบแห่งปัญหา(frame of reference)

2.ตัวเลือกตอบหรือ ข้อเลือก(option or alternative) ซึ่งตัวเลือกนี้แบ่งได้เป็น 2 ประเภท คือ ตัวที่เป็นคำตอบ(answer or corrector) และตัวลวง (distracter)

ข้อเสียของแบบสอบแบบเลือกตอบ

1.แบบสอบเลือกตอบสร้างยาก ผู้ออกข้อสอบจะต้องเขียนตัวเลือกไม่น้อยกว่า 3 ตัวเลือก ในแต่ละข้อตัวลวงนั้นจะต้องมีทีท่าว่าจะถูกแต่แท้จริงแล้วผิด หากการเขียนตัวลวงไม่ดี จะทำให้ง่ายต่อการเดาคำคำตอบ

2.เมื่อมีการกำหนดคำตอบมาให้เลือก จึงทำให้ข้อสอบแบบนี้ไม่เหมาะที่จะวัดความคิดริเริ่มสร้างสรรค์ของนักเรียน

3.ถ้านำข้อสอบแบบเลือกตอบใช้กับวิชาคณิตศาสตร์และวิทยาศาสตร์ที่วัดทักษะด้านการคำนวณ นักเรียนอาจหาคำตอบได้จากการย้อนหลังทำให้ข้อสอบมิได้วัดตรงตามจุดมุ่งหมาย

นอกจากนี้การให้นักเรียนทำข้อสอบแบบเลือกตอบมากเกินไป ทำให้นักเรียนขาดความสามารถ และทักษะในการเขียนเพื่อสื่อความหมาย

ข้อดีของแบบสอบแบบเลือกตอบ

1.ใช้วัดความสามารถได้ทุกด้าน เริ่มจากพื้นฐานคือความรู้ความจำ ไปจนถึงขั้นสูง เช่น การวิเคราะห์ การสังเคราะห์ และการประเมินผล

2.มีความเป็นปรนัยในการตรวจให้คะแนน

3.เป็นข้อสอบที่ดีกว่าแบบสอบตอบสั้นหรือเติมคำในแง่ทำให้ปัญหาเรื่องความกำกวมหมดไป เพราะมีคำตอบให้เลือก และเลือกข้อที่ดีที่สุดหรือถูกที่สุด

4.ข้อสอบแบบนี้มีตัวลวง ทำให้สามารถวินิจฉัยได้ว่า นักเรียนมีความบกพร่องหรือไม่เข้าใจในวิชาที่เรียนอย่างไรบ้าง

คำแนะนำในการเขียนข้อสอบแบบเลือกตอบ(วิเชียร เกตุสิงห์:2520,สมบุญ ภูนวล: 2525. เสริมศักดิ์ วิศาลาภรณ์: 2522,อุทุมพร จามรมาน: 2535,อัญชิกา ดิษเจริญ: 2535)

1.ตัวปัญหาหรือตัวคำถามต้องมีความชัดเจน อ่านแล้วเข้าใจว่าถามอะไร

2.พยายามทำให้ตัวปัญหาเป็นข้อความยาว ๆ เพื่อที่จะทำให้ตัวเลือกเป็นข้อความสั้น ๆ ช่วยให้นักเรียนประหยัดเวลาในการอ่าน ทำให้นักเรียนมีเวลาคิดพิจารณามากขึ้น

3.ข้อคำถามควรหลีกเลี่ยงข้อความที่ไม่เกี่ยวข้องกับเรื่องที่ต้องการถาม

4.พยายามใช้คำถามวัดความสามารถขั้นสูง ควรเขียนสถานการณ์ใหม่ หรือเขียนปัญหาใหม่ให้น่าสนใจ อย่าคัดลอกจากหนังสือโดยตรง

5.พยายามหลีกเลี่ยงคำถามที่เป็นประโยคปฏิเสธ โดยเฉพาะปฏิเสธซ้อนปฏิเสธ ถ้าจำเป็นต้องใช้ประโยคปฏิเสธ ควรเน้นให้เห็นอย่างชัดเจนโดยขีดเส้นใต้คำว่า **ไม่** หรือพิมพ์เป็นตัวหนา

6.พยายามให้แต่ละข้อเป็นอิสระแก่กัน อย่าให้ตัวปัญหาคาบเกี่ยวกัน เช่นถ้าตอบข้อ 2 ได้จะต้องตอบข้อ 1 ได้ก่อน

7.ตัวเลือกตอบควรมีลักษณะเป็นเอกพันธ์(homogeneous)

8.ตัวเลือกแต่ละข้อควรยาวเท่ากัน ตัวลวงต้องไม่ผิดอย่างเด่นชัด ขณะที่ตัวเลือกที่ถูกต้องไม่ชี้ชัดจนผู้สอบเดาคำตอบได้

9.การใช้ตัวเลือก "ไม่มีคำตอบถูก" หรือ "ถูกทุกข้อ" ควรใช้อย่างระมัดระวังและเลือกใช้ให้เหมาะสมกับเนื้อหาวิชาและข้อคำถาม

10.ตัวเลือกในแต่ละข้อคำถามจะต้องมีคำตอบที่ถูกที่สุดเพียงคำตอบเดียว

11.ตัวเลือกตอบนั้น ไม่ควรจะน้อยกว่า 4 ตัว เพราะถ้าเป็น 3 ตัวเลือก หากตัวเลือกหนึ่งขาดคุณภาพจะทำให้ข้อสอบแบบเลือกตอบไม่ต่างจากข้อสอบแบบถูกผิด อย่างไรก็ดีจำนวนตัวเลือกนั้นต้องขึ้นอยู่กับอายุของผู้สอบ เนื้อหาวิชา และเวลาในการทำข้อสอบด้วย

12.ตัวเลือกควรเรียงลำดับตามเกณฑ์ใดเกณฑ์หนึ่ง เช่น เรียงลำดับเหตุการณ์ เรียงจากมากไปน้อย หรือเรียงจากน้อยไปมาก

13.การเขียนตัวเลือกไม่ควรให้เหลื่อมล้ำกัน และภาษาที่ใช้ควรเป็นภาษาที่ง่ายชัดเจนและเข้าใจได้ง่ายไม่ควรใช้คำศัพท์ที่ยากเกินวัยของผู้ทำข้อสอบ

14.การสร้างข้อสอบควรใช้ฟอร์มข้อสอบที่มีคุณภาพ

15.ตำแหน่งของคำตอบที่ถูกต้องควรอยู่ในลักษณะสุ่ม กระจายไป ไม่ควรเรียงกันอย่างเป็นระบบหรืออยู่ซ้ำ ๆ กัน

แบบสอบแบบตอบสั้น(Short Answer Test) เป็นแบบสอบที่ต้องการให้ผู้สอบหาคำตอบเอง และเติมคำตอบในช่องว่างที่กำหนดให้ ข้อสอบแบบตอบสั้นแบ่งได้เป็น 2 อย่าง คือ

1.ตัวปัญหาเป็นรูปคำถามและต้องการคำตอบแต่เพียงสั้น ๆ

2.ตัวปัญหาเป็นข้อความไม่สมบูรณ์ มีข้อความเว้นว่างไว้ ต้องการให้ผู้ตอบเติมข้อความหรือคำที่เว้นว่างไว้ให้ได้ความสมบูรณ์

ข้อสอบแบบตอบสั้น สามารถใช้วัดผลการเรียนรู้ ได้หลายอย่าง ได้แก่ วัดความรู้เกี่ยวกับคำศัพท์ วัดความรู้เกี่ยวกับข้อเท็จจริงเฉพาะ วัดความรู้เกี่ยวกับกฎเกณฑ์ วัดความรู้เกี่ยวกับวิธีการหรือกระบวนการ วัดความสามารถในการแก้ปัญหา และวัดทักษะในการคำนวณ

ข้อเสียของข้อสอบแบบตอบสั้น

1.ถ้าสร้างข้อสอบไม่ดีจะวัดได้แค่ระดับความรู้ความจำ ซึ่งเป็นการวัดระดับขั้นต้นเท่านั้น

2.เนื่องจากข้อสอบประเภทนี้ต้องการคำตอบสั้น ๆ หรือคำตอบเพียงคำตอบเดียวดังนั้น การวัดกระบวนการขั้นสูงเช่น การวัดการวิเคราะห์ การสังเคราะห์ และการประเมินผลนั้นจะทำได้ยากมาก

3.ถ้าเขียนคำถามไม่ดีอาจทำให้มีหลายคำตอบ ทำให้เกิดความลำบากในการตรวจให้คะแนน

4.ข้อสอบประเภทนี้ตรวจยากและใช้เวลานานกว่าข้อสอบปรนัยอื่น ๆ เช่น ถูกผิด หรือแบบเลือกตอบ

ข้อดีของแบบสอบประเภทตอบสั้น

1.ข้อสอบประเภทนี้วัดความรู้ความจำได้เป็นอย่างดี

2.สร้างง่ายกว่าแบบสอบแบบเลือกตอบ หรือแบบจับคู่ เพราะไม่ต้องสร้างตัวเลือกหรือตัวลวง ทำให้ประหยัดเวลาในการสร้างด้วย

3.นักเรียนไม่สามารถเดาคำตอบได้

4.ข้อสอบเป็นข้อสอบเติมคำสั้น ๆ ทำให้ใช้เวลาในการทำข้อสอบน้อยจึงสามารถออกข้อสอบได้มากข้อทำให้ครอบคลุมเนื้อหาวิชาได้กว้างขวางมาก

นอกจากนี้นักวัดผลหลายท่านเช่น Stanley,Gronlund,Hedges,Mashall and Hales ต่างมีความเห็นตรงกันว่าแบบสอบแบบเติมคำสั้น ๆ มีประโยชน์ในการวัดผลวิชาวิทยาศาสตร์และคณิตศาสตร์ เพราะสามารถวัดความรู้เกี่ยวกับข้อเท็จจริง กฎเกณฑ์ ความสามารถที่จะหาคำตอบโดยการคำนวณ ทักษะต่าง ๆในการใช้สัญลักษณ์ และวัดกระบวนการขั้นสูงได้ด้วย(สมบูญ ภู่นวน, 2525 : 108 -109)

คำแนะนำในการเขียนข้อสอบแบบตอบสั้น

1.เขียนคำสั่งให้ชัดเจนว่าต้องการให้นักเรียนตอบอย่างไร ตอบที่ไหน

2.พยายามเขียนปัญหาหรือคำถาม โดยให้มีคำตอบเพียงอย่างเดียวหรือเพียงคำตอบเดียว

3.ไม่ควรลอกข้อความจากหนังสือแล้วตัดข้อความบางส่วนออก

4.ตัวปัญหาน่าจะเป็นคำถามมากกว่าเป็นข้อความที่ไม่สมบูรณ์

5.ถ้าคำตอบเป็นจำนวนที่มีหน่วย ควรระบุหน่วยที่ต้องการตอบด้วย

6.ช่องว่างที่ให้เติมควรอยู่ตอนท้ายมากกว่าตอนต้นและอย่าเว้นช่องว่างไว้หลายช่องจนไม่ทราบว่าโจทย์ต้องการอะไรและถ้าจำเป็นต้องมีหลายช่อง ช่องว่างเหล่านั้นควรมีช่วงยาวเท่ากัน

7.ช่องว่างสำหรับเติมคำตอบนั้นควรอยู่ในที่ที่ทำให้ตรวจได้ง่าย เช่นทางขวามือหมดทุกข้อ

ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง

ศรินทิพย์ แก้วมหาวงศ์ (2531) ได้พัฒนาแบบสอบอิงโดเมน โดยใช้โมเดลโลจิสติค 3 พารามิเตอร์กับวิธีแองกอฟในการกำหนดจุดตัดของแบบสอบ โดยแบ่งแบบสอบออกเป็น 4 โดเมน พบว่าจุดตัดของแบบสอบหรือคะแนนความสามารถ (θ) ซึ่งเป็นเกณฑ์ในการผ่านการเรียนรู้ตามโดเมนที่ 1,2,3,4 และทั้งฉบับเป็น 1.48 , 0.59 , 0.37 , 1.04 และ 0.93 ตามลำดับ เมื่อคิดเป็นคะแนนโดเมนของแบบสอบทั้งฉบับมีค่าเป็น 0.6

มาลี จิตติวุฒิการ(2532) ได้ศึกษาผลของการหาคะแนนจุดตัด คะแนนโดเมน และความยาวของแบบสอบที่มีต่อความน่าจะเป็นในการจำแนกความรอบรู้และความเที่ยงในการตัดสินใจ โดยการหาคะแนนจุดตัดของคะแนนโดเมนได้ประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบกับวิธีวิลคอกซ์พบว่าคะแนนจุดตัดเป็น 60% ของความยาวแบบสอบ หรือมีคะแนนโดเมนเป็น 0.6

อัจฉริยา ปราบอริพ่าย (2531) ได้ศึกษาผลของการใช้สารสนเทศแก่ผู้เชี่ยวชาญต่อการกำหนดจุดตัดของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ พบว่าคะแนนจุดตัดที่ใช้เป็นเกณฑ์ตัดสินความสามารถขั้นต่ำที่ยอมรับได้ของผู้สอบซึ่งอยู่บนสเกลของคะแนนโดเมนที่ได้จากการตัดสินของผู้เชี่ยวชาญที่ไม่มีสารสนเทศ และผู้เชี่ยวชาญที่มีสารสนเทศเท่ากับ 0.66 และ 0.53 ส่วนค่าดัชนีความคงที่ของการตัดสินคะแนนจุดตัดซึ่งตัดสินโดยผู้เชี่ยวชาญที่ไม่มีสารสนเทศและผู้เชี่ยวชาญที่มีสารสนเทศเท่ากับ 0.78 และ 0.76 ตามลำดับซึ่งไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

จากงานวิจัยข้างต้นพบว่าคะแนนโดเมนที่ได้จากการวิจัยมีค่าอยู่ในช่วง 0.5 ถึง 0.7 และคะแนนโดเมนที่ได้จากการวิจัยของอัจฉริยา ปราบอริพ่ายเป็นคะแนนโดเมนที่ได้จากวิธีการหาจุดตัดของแองกอฟ เช่นเดียวกับงานวิจัยของศรินทิพย์ แก้วมหาวงศ์ ซึ่งคะแนนโดเมนที่ได้จะได้จากการประยุกต์ทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ในขณะที่คะแนนจุดตัดที่ได้จากงานวิจัยของมาลี จิตติวุฒิการ ได้จากวิธีวิลคอกซ์โดยการประยุกต์ทฤษฎีการตอบสนองข้อสอบเช่นเดียวกัน

พานิช ศรีงาม (2530) ได้ศึกษาเปรียบเทียบคุณภาพแบบสอบอิงเกณฑ์โดยวิธีกำหนดจุดตัดด้วยวิธีประยุกต์รูปแบบของราสช์ วิธีตัดสินใจของแกลส และวิธีการนับถอยหลัง พบว่าการกำหนดจุดตัดที่ให้ค่าจุดตัดสูงสุดได้แก่วิธีนับถอยหลัง รองลงมาคือวิธีประยุกต์รูปแบบของราสช์ และต่ำสุดคือ วิธีการใช้ทฤษฎีการตัดสินใจของแกลส ค่าความเที่ยงของแบบสอบอิงเกณฑ์ที่ได้จากการกำหนดจุดตัดด้วยวิธีประยุกต์รูปแบบของราสช์ วิธีการใช้ทฤษฎีการตัดสินใจของแกลส และวิธีนับถอยหลัง ไม่แตกต่างกันเมื่อความเที่ยงดังกล่าวถูกแปลงให้อยู่ในรูปของคะแนนฟิชเชอร์ซี ส่วนความตรงของแบบสอบอิงเกณฑ์ที่ได้จากการกำหนดจุดตัดด้วยวิธีประยุกต์รูปแบบของราสช์ วิธีการใช้ทฤษฎีการตัดสินใจของแกลสและวิธีนับถอยหลัง แตกต่างกันเมื่อความตรงดังกล่าวถูกแปลงให้อยู่ในรูปของคะแนนฟิชเชอร์ซี งานวิจัยครั้งนี้ได้เริ่มมีการนำทฤษฎีการตอบสนองข้อสอบมาใช้เป็นโมเดลโลจิสติก 1 พารามิเตอร์ พบว่าวิธีการกำหนดจุดตัดโดยการประยุกต์รูปแบบของราสช์ (IRT) มีค่าความตรงสูงกว่าวิธีอื่น ๆ ส่วนความเที่ยงมีค่าไม่แตกต่างจากวิธีอื่น

รังสรรค์ มณีเล็ก (2527) ได้ศึกษาเปรียบเทียบความเที่ยงของแบบสอบอิงเกณฑ์โดยใช้การกำหนดจุดตัดต่างกัน 4 วิธีคือ วิธีนับถอยหลัง 100% (Counting backward from 100%) วิธีของนีเดลสกี(Nedelsky) วิธีของเบอร์ก(Berk) วิธีของเบส์(Baysian) พบว่าวิธีกำหนดจุดตัดสูงสุดและต่ำสุดได้แก่วิธีนับถอยหลัง 100% (Counting backward from 100%)และวิธีของนีเดลสกี(Nedelsky)ตามลำดับส่วนความเที่ยงของแบบสอบอิงเกณฑ์ที่ได้จากการกำหนดจุดตัดที่แตกต่างกันจะไม่แตกต่างกันที่ระดับ .05 เมื่อความเที่ยงดังกล่าวแปลงให้อยู่ในรูปคะแนนฟิชเชอร์ซี (Fisher – Z)

สุรินทร์ แพ่งจันทึก (2528) ได้ศึกษาเกี่ยวกับแบบสอบอิงเกณฑ์ประเภทเลือกตอบและตอบสั้น โดยได้ทำการวิจัยเปรียบเทียบคะแนนจุดตัด และความเที่ยงของแบบสอบอิงเกณฑ์ระหว่างแบบสอบประเภทเลือกตอบ และตอบสั้นจากการใช้ทฤษฎีการตัดสินใจของแกลส พบว่าแบบสอบเลือกตอบและตอบสั้นมีคะแนนจุดตัดแตกต่างกัน 30 % การจำแนกความรอบรู้ของแบบสอบทั้งสอบไม่แตกต่างกัน ส่วนค่าความเที่ยงของแบบสอบประเภทเลือกตอบน้อยความความเที่ยงของแบบสอบประเภทตอบสั้น

อุบล แสงเพ็ง (2534) ได้ศึกษาเปรียบเทียบความคลาดเคลื่อนของแบบสอบอิงเกณฑ์ระหว่างแบบสอบประเภทเลือกตอบและตอบสั้นในการจำแนกระดับความรอบรู้ของนักเรียน พบว่าแบบสอบเลือกตอบและตอบสั้นที่มีความยาวเดียวกัน เกิดความคลาดเคลื่อนที่ปรากฏที่ชัดเจน ไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01

จากการวิจัยของสุรินทร์ แพ่งจันทึก และอุบล แสงเพ็ง เป็นการศึกษาเปรียบเทียบเกี่ยวกับคุณภาพของแบบสอบเลือกตอบและตอบสั้นสรุปได้ว่าแบบสอบเลือกตอบจะมีคะแนนจุดตัดต่ำกว่าคะแนนจุดตัดของแบบสอบตอบสั้น และค่าความเที่ยงของแบบสอบประเภทตอบสั้นมีค่าสูงกว่าค่าความเที่ยงของแบบสอบเลือกตอบ ส่วนการจำแนกความรอบรู้ของนักเรียนของแบบสอบทั้ง 2 ประเภทนี้มีค่าไม่แตกต่างกัน

จากงานวิจัยข้างต้นพบว่าเป็นการศึกษาเกี่ยวกับการสร้างแบบสอบอิงเกณฑ์และแบบสอบอิงโดเมน แล้วศึกษาถึงการกำหนดจุดตัดด้วยวิธีต่าง ๆ โดยแบ่งเป็นการกำหนดจุดตัดที่ยังคงใช้ทฤษฎีแบบดั้งเดิม(Classical Test Theory : CTT) ซึ่งได้แก่ การกำหนดจุดตัดโดยวิธีนับถอยหลัง วิธีการตัดสินใจของแกลส วิธีของนีเดลสกี วิธีของเบอร์ก และวิธีของเบส์ ซึ่งจากงานวิจัยพบว่าความเที่ยงของการกำหนดจุดตัดเหล่านี้ไม่แตกต่างกัน เมื่อเริ่มมีการพัฒนานำทฤษฎีการตอบสนองข้อสอบ(Item Response Theory : IRT)มาใช้ในการกำหนดจุดตัดเริ่มจากการกำหนดจุดตัดโดยใช้รูปแบบของราสช์ (IRT 1 พารามิเตอร์) และเมื่อมีการเปรียบเทียบความตรงและความเที่ยงของการกำหนดจุดตัดด้วยรูปแบบของราสช์กับการกำหนดจุดตัดแบบทฤษฎีดั้งเดิมพบว่าการกำหนดจุดตัดด้วยรูปแบบของราสช์มีความตรงมากกว่าส่วนความเที่ยงไม่แตกต่างกัน แสดงว่าการกำหนดจุดตัดโดยใช้ทฤษฎีการตอบสนองข้อสอบน่าจะดีการกำหนดจุดตัดโดยใช้ทฤษฎีแบบดั้งเดิม ต่อมามีการนำทฤษฎีการตอบสนองข้อสอบ(IRT 3 พารามิเตอร์) มาใช้ในการกำหนดจุดตัดซึ่งได้แก่การกำหนดจุดตัดของวิลคอกซ์และแองกอฟ จึงทำให้ผู้วิจัยสนใจในการเปรียบเทียบคุณภาพการกำหนดจุดตัดด้วย 2 วิธีดังกล่าว ส่วนแบบสอบเลือกตอบกับตอบสั้นซึ่งเป็นแบบสอบที่มีข้อดีแตกต่างกันไปจากการวิจัยพบว่าแบบสอบทั้ง 2 ประเภทนี้จำแนกความรอบรู้ของนักเรียนไม่แตกต่างกัน นอกจากนี้ความคลาดเคลื่อนในการจำแนกความรอบรู้ของนักเรียนไม่แตกต่างกันเมื่อแบบสอบทั้ง 2 มีความยาวเท่ากัน ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยจึงไม่ได้เปรียบเทียบคุณภาพของแบบสอบเลือกตอบและตอบสั้นแต่สนใจเปรียบเทียบคุณภาพของการกำหนดจุดตัดสำหรับแบบสอบประเภทเลือกตอบและตอบสั้นแทนเพื่อประโยชน์ของการนำจุดตัดไปใช้จะใช้ได้ทั้งแบบสอบประเภทเลือกตอบและตอบสั้น

I 18394115