

วรรณคดีที่เกี่ยวข้อง

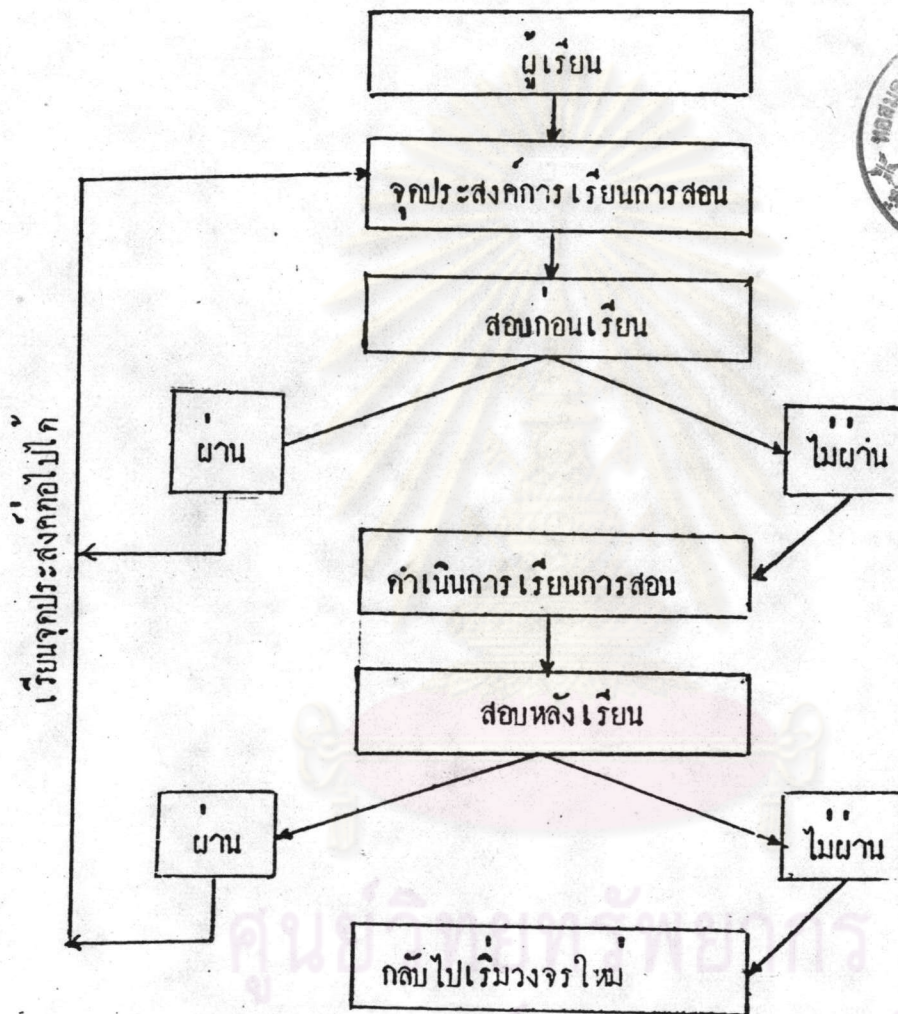
ในบทนี้จะกล่าวถึงลักษณะการวัดแบบอิงเกณฑ์ การหาค่าความเที่ยง การหาค่าความตรง การกำหนดจุดตัดของแบบสอบอิงเกณฑ์ และรูปแบบของราศี พร้อมทั้งวรรณคดีที่เกี่ยวข้องใน 5 ประเด็น ดังต่อไปนี้

1. ลักษณะการวัดแบบอิงเกณฑ์
2. ความเที่ยงของแบบสอบอิงเกณฑ์และงานวิจัยที่เกี่ยวข้อง
3. ความตรงของแบบสอบอิงเกณฑ์และงานวิจัยที่เกี่ยวข้อง
4. การกำหนดจุดตัดของแบบสอบอิงเกณฑ์และงานวิจัยที่เกี่ยวข้อง
5. รูปแบบของราศีและงานวิจัยที่เกี่ยวข้อง

1. ลักษณะการวัดแบบอิงเกณฑ์

จุดเริ่มของการวัดผลทางการศึกษานั้น มีจุดมุ่งหมายเพื่อนำผลสอบของแต่ละบุคคลไปเปรียบเทียบกันว่าแต่ละคนมีผลการเรียนแตกต่างกันอย่างไรบ้าง และอยู่ในระดับใด ซึ่งในเวลาต่อมาความสำคัญในด้านการวินิจฉัยผลการเรียน การอธิบายถึงงานต่าง ๆ ที่ผู้สอบสามารถกระทำได้ รวมทั้งผลสัมฤทธิ์ของการจัดโปรแกรมการศึกษาก็ได้รับความสนใจมากขึ้นเป็นลำดับ จึงมีความจำเป็นที่ผู้ทำหน้าที่เกี่ยวข้องกับเรื่องดังกล่าวนี้ จะต้องหารูปแบบรวมทั้งเทคโนโลยีของการทดสอบมาใช้เพื่อให้เกิดประโยชน์มากขึ้น (Haretel, 1981) นักวิจัยกลุ่มแรก ที่ให้ความสนใจต่อการพัฒนาเทคโนโลยีของการทดสอบ คือ แกลสเซอร์ (Glasser, 1963) โปแพมและฮูเสค (Popham and Husek, 1969) จุดมุ่งหมายในการพัฒนาเทคโนโลยีของการทดสอบขึ้นมา ก็เพื่อต้องการแปลความหมายของคะแนนที่ได้จากการสอบให้มีความชัดเจนมากยิ่งขึ้นกว่าที่พามา แบบสอบที่นักวิจัยกลุ่มนี้สร้างขึ้นนั้น คือแบบสอบอิงเกณฑ์ (Criterion Referenced test) จากนั้นอีกไม่นานก็ให้งานเขียน และงานวิจัยที่เกี่ยวข้องกับการวัดแบบอิงเกณฑ์มากกว่า 600 ชิ้น ที่กล่าวถึงการพัฒนาและการสร้างแบบสอบในแง่มุมต่าง ๆ (Hambleton, 1980 : 422)

การจัดการเรียนการสอนที่สอดคล้องกับแนวการวัดแบบอิง เกณฑ์ มิลล์แมน (Millman, 1979) ได้เสนอรูปแบบในการจัดการเรียนการสอนเพื่อช่วยเหลือนักเรียน เป็นรายบุคคลไว้ดังนี้



จากรูปแบบนี้จะเห็นได้ว่า การดำเนินการจัดการเรียนการสอน ได้เน้นการช่วยเหลือ นักเรียน รายบุคคลเป็นลำดับขั้น คือ เมื่อนักเรียนไม่สามารถผ่านเกณฑ์ขั้นแรกไปก็กลับไป เริ่มต้นใหม่จนกว่าจะสามารถผ่านเกณฑ์ได้

การสร้างแบบสอบอิง เกณฑ์ความ เกี่ยวข้อง โดยตรงกับการแปลความหมายของคะแนน จากการปฏิบัติได้จริงของผู้เรียนว่ามีความรู้ในจุดมุ่งหมายที่กำหนดไว้มากน้อยเพียงใด หลังจาก

ที่ผ่านการเรียนการสอนไปแล้ว ทั้งนี้แนวการสร้างแบบสออิง เกณฑ์ควรกำหนดเอาไว้
อย่างชัดเจน ซึ่ง กรอนลันด์ (Gronlund, 1978 : 3-6) ได้เสนอแนะไว้ดังนี้

1. แบบสออิง เกณฑ์จะต้องมีการกำหนดขอบเขตของงานเอาไว้อย่างชัดเจน เพราะ
แบบสอต่าง ๆ จะนำมาใช้ให้ประสพผลสำเร็จได้นั้นจะต้องบอกถึงความเจาะจงของผลสอ ที่
ผู้สอกระทำได้ การกำหนดความเฉพาะเจาะจงของ เนื้อหาถือว่าเป็นข้อจำกัดอย่างหนึ่ง
ในการนำแบบสออิง เกณฑ์ไปใช้ ปัญหาที่สามารถแก้ไขได้หากร่วมมือกันหลาย ๆ ฝ่าย

2. แบบสออิง เกณฑ์จะต้องมีความชัดเจนในการกำหนดจุดประสงค์ของการ เรียน
การสอน ในสิ่งที่ต้องการให้ผู้เรียนได้แสดงออกมาในรูปของพฤติกรรมที่ปฏิบัติได้จริง ซึ่ง เมื่อ
ดูจากผลการสอก็จะทำให้ทราบทันทีว่านักเรียนสามารถทำอะไร ได้บ้างจากงานที่มอบหมายให้

3. แบบสออิง เกณฑ์จะต้องมีการกำหนดเกณฑ์ในการปฏิบัติเอาไว้อย่างชัดเจน
เพราะการกำหนดเกณฑ์การปฏิบัติมีผลต่อการนำแบบสออิง เกณฑ์มาใช้ในห้องเรียน ซึ่งใน
ขณะนี้ยังถือว่าเป็นอุปสรรคสำคัญในการนำแบบสอมาใช้ จนกว่าจะมีงานวิจัยหรือมีทฤษฎี
ที่สามารถนำมาใช้ให้เป็นที่ยอมรับกันได้ แต่อย่างไรก็ตามแม้ว่าการกำหนดเกณฑ์การปฏิบัติ
ยังไม่สามารถเป็นที่ยอมรับได้ก็ตามแต่ก็จำเป็นจะต้องมีการกำหนดเกณฑ์ขึ้นมาใช้เพื่อที่จะแสดง
ว่าอย่างน้อยที่สุดผู้เรียนมีความรู้ในเนื้อหานั้น ๆ แล้ว

4. แบบสออิง เกณฑ์ต้องสามารถวัดความเป็นตัวแทนของ เนื้อหาในการวัดได้อย่าง
เพียงพอ โดยสามารถวัดความเป็นตัวแทนของ เนื้อหาของความรู้ได้ดังนี้

4.1 แบ่ง เนื้อหาในการเรียนออกเป็นหน่วยย่อย ๆ

4.2 กำหนดโดเมน (Domain) ของ เนื้อหาให้มีความชัดเจน

4.3 กำหนดเนื้อหาให้มีความเฉพาะเจาะจงกับกลุ่มผู้สอ

5. แบบสออิง เกณฑ์ต้องสร้างข้อสอแต่ละข้อให้ตรงกับจุดประสงค์ที่ของการจะวัด
เพราะจะสามารถทำให้การแปลความหมายของการกระทำของผู้เรียนได้ตามโดเมนของ เนื้อหา

6. แบบสออิง เกณฑ์ต้องมีการรายงานผลของคะแนนที่แสดงถึงการปฏิบัติได้จริง
ของผู้เรียนในของ เขตของงาน (task) ที่กำหนดเอาไว้อย่างชัดเจน

การนำแบบสอยอิง เกณฑ์ไปใช้กับการจัดการเรียนการสอนเราสามารถนำไปใช้
ประโยชน์ได้ในหลาย ๆ ด้านดังนี้คือ (Gronlund, 1978 : 6)

1. ใช้วัดความรู้และทักษะก่อนการเรียนในหน่วยการเรียนรู้หนึ่ง ๆ ว่าผู้เรียน
มีความรู้ในเรื่องนั้นมากน้อยเพียงใด (pretest)
2. ใช้วัดความก้าวหน้าและวัดความรู้ในหน่วยการเรียนรู้หนึ่ง ๆ (Formative test)
3. ใช้วัดเพื่อวิเคราะห์ปัญหาที่เกิดขึ้นในขณะที่การเรียนการสอนกำลังดำเนินอยู่
(Diagnostic test)
4. ใช้วัดเพื่อสรุปผลการเรียนในหน่วยการเรียนรู้หนึ่ง ๆ (Summative test)

2. ความเที่ยงของแบบสอยอิง เกณฑ์และงานวิจัยที่เกี่ยวข้อง

ก. ความเที่ยงของแบบสอยอิง เกณฑ์

ความเที่ยงของแบบสอยอิง เกณฑ์เป็นคตินี้ตัวหนึ่งที่สามารถบ่งบอกถึงคุณภาพของ
แบบสอยอิง เกณฑ์ได้เป็นอย่างดี ทั้งนี้ในการสร้างแบบสอยทุกครั้งจึงจำเป็นต้องมีการตรวจสอบ
หาความเที่ยงของแบบสอยเพื่อคตินี้ตัวหนึ่งในการนำมาประกอบการตัดสินใจเพื่อนำแบบสอยมาใช้

เบอร์ก (Berk, 1980 : 127) ได้สรุปโน้ตสันของความเที่ยงของแบบสอย
อิง เกณฑ์ไว้ 3 มโนทัศน์ ดังนี้

1. คือความคงที่ในการตัดสินใจจำแนกผู้รอบรู้ ไม่รอบรู้โดยการสอบซ้ำ ด้วย
แบบสอยฟอร์มเดียวกันหรือแบบสอยคู่ขนาน
2. คือความคงที่ของความแตกต่างกำลังสองของคะแนนแต่ละคนที่เบี่ยงเบนไปจาก
คะแนนจุดตัดโดยการสอบด้วยแบบสอยคู่ขนาน
3. คือความคงที่ของคะแนนแต่ละคนที่แตกต่างกันจากการสอบด้วยแบบสอยคู่ขนาน
จากข้อสรุปโน้ตสันความเที่ยงของแบบสอยอิง เกณฑ์ทั้ง 3 รูปแบบ เราสามารถเลือก
เอานิยามใดนิยามหนึ่งในการประมาณค่าความเที่ยงก็ได้ แต่ตามมโนทัศน์ที่ 1 และ 2 น่าจะนำ
มาใช้ได้เหมาะสมกว่า เพราะมีจุดตัดเข้าไปเกี่ยวข้องอย่างชัดเจน ส่วนตามมโนทัศน์ที่ 3 นั้นไม่ได้

กล่าวถึงจุดที่เอาไว้ ผู้วิจัยจึงนำเสนอการประมาณค่าความเที่ยงของแบบสอบถามเชิงทัศนคติ
ตามมโนทัศน์ที่ 1 และที่ 2 เท่านั้น

ข. เทคนิคการประมาณค่าความเที่ยงของแบบสอบถามเชิงทัศนคติ

ตั้งแต่ได้เริ่มมีการนำการวัดแบบสอบถามมาใช้ ก็มีกลุ่มนักวิจัยหลาย เช่น
แฮมเบิลตันและคณะ (Hambleton, et al. 1980) ลินน์ (Linn, 1970) และ
มิลล์แมน (Millman, 1979) ได้ทำการสรุปสูตรที่ใช้ในการคำนวณค่าความเที่ยงของ
แบบสอบถามมโนทัศน์ที่ 1 และ 2 ดังนี้

1. การประมาณค่าความเที่ยงโดยพิจารณาจากความคงที่ในการตัดสินใจจำแนกความ
รอบรู้จากการสอบซ้ำด้วยแบบสอบถามฟอร์มเดียวกัน ซึ่งมีผู้เสนอสูตรวิธีต่าง ๆ ดังนี้

1.1 วิธีของแฮมเบิลตัน และโนวิก (Hambleton and Novick, 1973) ซึ่งได้เสนอสูตรการประมาณค่าความเที่ยงโดยใช้ผลรวมของสัดส่วนความสอดคล้องในการ
ตัดสินใจ จำแนกผู้รอบรู้จากการสอบด้วยแบบสอบถามเดียวกันสองครั้งหรือจากการสอบ
ด้วยแบบสอบถามคนานาครั้ง

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

เมื่อ	P_0	คือ ความเที่ยงของแบบสอบถาม
	P_{11}	คือ สัดส่วนของผู้ถูกตัดสินว่ารอบรู้ทั้งทั้งสองครั้ง
	P_{22}	คือ สัดส่วนของผู้ถูกตัดสินว่าไม่รอบรู้ทั้งทั้งสองครั้ง

1.2 วิธีของสวามินาธาน แฮมเบิลตัน และอัลจินา (Swaminathan, Hambleton and Algina, 1974) ซึ่งได้เสนอสูตรสำหรับการประมาณค่าความเที่ยงของ
แบบสอบถามคล้าย ๆ กับวิธีของแฮมเบิลตัน และโนวิก (Hambleton and Novick) แต่
แก้วิธีนี้มีการปรับแก้โดยการหักเอาค่าความสอดคล้องโดยบังเอิญออกไปดังนี้

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

เมื่อ K คือ ค่าความเที่ยงของแบบสอบถาม เกณฑ์ไทยหักเอา
ความสอกลอง โคมบัง ไขว้ออก

P_{\bullet} คือ สัดส่วนความสอกลอง ในการตัดสินจำแนกผู้รอบรู้

P_c คือ สัดส่วนของความสอกลอง ที่คาดหวังซึ่งผลเกิดขึ้น โคมบัง ไขว้

วิธีหา

$$P_{\bullet} = \sum_{i=1}^2 P_{i1}$$

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

เพื่อให้การคำนวณสะดวกขึ้นนำผลการ สอมาแจกแจงลงในตารางต่อไปนี้

	รอบรู้	ไม่รอบรู้	สัดส่วนแยก
รอบรู้	P_{11}	P_{12}	$P_{1\bullet}$
ไม่รอบรู้	P_{21}	P_{22}	$P_{2\bullet}$
สัดส่วนแยก	$P_{\bullet 1}$	$P_{\bullet 2}$	

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

P_{12} คือ สัดส่วนของผู้สอบที่ถูกตัดสินว่ารอบรู้ในฉบับ ก. แต่ไม่รอบรู้
ในฉบับ ข. หรือรอบรู้ในครั้งที่ 1 แต่ไม่รอบรู้ในครั้งที่ 2

P_{21} คือ สัดส่วนของผู้สอบที่ถูกตัดสินว่าไม่รอบรู้ในฉบับ ก. แต่รอบรู้
ในฉบับ ข. หรือรอบรู้ในครั้งที่ 2 แต่ไม่รอบรู้ในครั้งที่ 1

- P_{22} คือ สัดส่วนของผู้สอบที่ถูกตัดสินว่าไม่รอบรู้ครั้งที่ 1 หรือครั้งที่ 2
 $P_{1.}$ คือ สัดส่วนแยกของผู้สอบที่รอบรู้ฉบับ ก. หรือครั้งที่ 1
 $P_{2.}$ คือ สัดส่วนแยกของผู้สอบที่ไม่รอบรู้ฉบับ ก. หรือครั้งที่ 1
 $P_{.1}$ คือ สัดส่วนแยกของผู้สอบที่รอบรู้ฉบับ ข. หรือครั้งที่ 2
 $P_{.2}$ คือ สัดส่วนแยกของผู้สอบที่ไม่รอบรู้ฉบับ ข. หรือจากการสอบครั้งที่ 2

1.3 วิธีของสับโคเวียค (Subkoviak, 1976)

สับโคเวียค (Subkoviak) ได้เสนอวิธีการที่ประยุกต์มาจากการสอบ 2 ครั้ง โดยการสอบเพียงครั้งเดียวซึ่งแปรระดับการรอบรู้ของผู้สอบออกเป็น 2 ระดับ โดยใช้สูตรดังนี้

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

$$\text{โดยที่ } P_c(i) = P(X_1 \geq C) \cdot P(X'_1 \geq C) + P(X_1 < C) \cdot P(X'_1 < C)$$

- เมื่อ X_1 คือ คะแนนที่ได้จากแบบสอบฉบับที่ 1
 X'_1 คือ คะแนนที่ได้จากแบบสอบฉบับที่ 2 ซึ่งเป็นคู่ขนานกับฉบับที่ 1
 C คือ จุดตัด

วิธีของสับโคเวียคมีข้อตกลงเบื้องต้น 2 ข้อคือ

1. คะแนนจากแบบสอบฉบับที่ 1 (X_1) และจากแบบสอบฉบับที่ 2 (X'_1) ของผู้สอบแต่ละคนมีการแจกแจงที่เป็นอิสระต่อกัน หมายความว่าประสิทธิภาพในการทำแบบสอบฉบับที่ 1 ของผู้สอบไม่มีผลต่อการทำแบบสอบฉบับที่ 2
2. การแจกแจงของคะแนนจากการสอบแบบสอบ X_1 และ X'_1 ของผู้สอบจะต้องเป็นแบบไบโนเมียล ซึ่งมีเงื่อนไขดังนี้
 - ก. ข้อสอบแต่ละข้อมีการให้คะแนนแบบ 0, 1
 - ข. ผลของการตอบข้อสอบข้อหนึ่ง ไม่มีผลต่อข้ออื่น ๆ
 - ค. โอกาสในการทำข้อสอบถูกแต่ละข้อมีความคงที่ในทุกข้อ



จากข้อตกลงเบื้องต้นจะเห็นว่า

$$\begin{aligned}
 P_c(1) &= [P(X_1 > c)]^2 + [P(X_1 < c)]^2 \\
 &= [P(X_1 > c)]^2 + [1 - P(X_1 > c)]^2
 \end{aligned}$$

เมื่อ

$$P(X_1 > c) = \sum_{x_i=c}^n \binom{n}{x_i} P_1^{x_i} (1-p)^{n-x_i}$$

P_1 คือ ความน่าจะเป็นที่แท้จริงของค่าทอยในข้อสอบที่ถูกทอยสำหรับ
 ผู้สอบคนที่ i ซึ่งคำนวณได้จากคะแนนสอบโดยใช้สูตร

$$P_1 = \Sigma X_i / n$$

1.4 วิธีของเฮวิน (Huynh, 1976)

เฮวิน ได้เสนอวิธีการประมาณค่าความเที่ยงในรูปของสัมประสิทธิ์
 แคปป่า ซึ่งมีสูตรดังนี้

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

เมื่อ

$$\begin{aligned}
 P_{11} &= \sum_{x,y=c}^n f(x,y) \\
 &= f(n,n) + f(n,n-1) + f(n-k,n) + \dots + f(e,c) \\
 P_1 &= \sum_{x=c}^n f(x) \\
 &= f(n) + \dots + f(c)
 \end{aligned}$$

1.5 วิธีของมาร์แชลและแฮร์เทล (Marshall and Haretel, 1976)

วิธีนี้จะมีกระบวนการคล้าย ๆ กับวิธีของเฮวีน และวิธีของ สับโคเวียค คือใช้ข้อมูลจากการสอบเพียงครั้งเดียวโดยมีข้อตกลงว่า ถ้าให้ผู้สอบได้ทำการ สอบซ้ำแล้ว การแจกแจงของคะแนนของผู้สอบจะเป็นแบบ ไบโนเมียล

2. การประมาณค่าความเที่ยง โดยพิจารณาจากคะแนนของแต่ละคนที่เบี่ยงเบน ไปจากจุดตัด จากการสอบด้วยแบบสอบคู่ขนาน ซึ่งมีผู้นำเสนอไว้ดังนี้

2.1 วิธีของลิฟวิงสตัน (Livingston, 1972)

ลิฟวิงสตัน ได้เสนอวิธีการประมาณค่าความเที่ยงโดยกำหนดว่าจุดประสงค์ ของแบบสอบอิง เกณฑ์คือแยกคะแนนโคเมนที่ได้จากการประมาณของผู้สอบแต่ละคนออกจาก คะแนนจุดตัด จึงควรใช้การหาค่าเบี่ยงเบนของคะแนนโคเมนที่ได้จากการประมาณ และคะแนน โคเมนของผู้สอบออกจากคะแนนจุดตัดมากกว่าคะแนนซึ่ง เป็นค่ามัธยิม เลขคณิตของคะแนนโคเมน ซึ่ง เป็นวิธีที่ทฤษฎีคลาสสิกใช้ สูตรในการหาความเที่ยงมีดังนี้

$$(\sigma_i, \sigma) = \frac{\sigma^2(\pi) + (\pi - \pi_0)^2}{\sigma^2(\pi) + (\pi - \pi)^2}$$

เมื่อ π คือ คะแนนโคเมนที่ได้จากการประมาณ

π คือ คะแนนโคเมน

π_0 คือ มัธยิม เลขคณิตของคะแนนโคเมน

$\sigma(\pi)$ คือ ความแปรปรวนของคะแนนโคเมนที่ได้จากการประมาณ ออกจากคะแนนจุดตัด

$\sigma(\pi)$ คือ ความแปรปรวนของคะแนนโคเมนออกจากคะแนนจุดตัด

π_0 คือ คะแนนจุดตัด

ก. งานวิจัยที่เกี่ยวข้องกับความเที่ยงของแบบสอบอิง เกณฑ์

การศึกษาค่าความเที่ยงของแบบสอบอิง เกณฑ์ในต่างประเทศและในประเทศไทย
ให้มี ผู้ศึกษาไว้ในลักษณะต่าง ๆ กันดังนี้

สับโคเวียค (Subkoviak, 1978:111-115) ได้เปรียบเทียบวิธีการประมาณค่า
ความเที่ยงของแบบสอบอิง เกณฑ์ในรูปของความสอดคล้องในการตัดสินใจจำแนกผู้รอบรู้

4 วิธีด้วยกันคือ

1. วิธีของสวามินาธาน แฮมเบิลตัน และอัลจินา (Swaminathan,
Hambleton and Algina Procedure)
2. วิธีของมาร์แชล และแฮร์เทล (Marshall-Haertel
Procedure)
3. วิธีของสับโคเวียค (Subkoviak Procedure)
4. วิธีของเฮวน์ (Huynh Procedure)

ข้อมูลได้จากนักเรียน 1,586 คน โดยสุ่มมาเป็นกลุ่มตัวอย่างแบบแทนที่ 1,500 คน
แบ่งเป็น 50 กลุ่ม กลุ่มละ 30 คน เครื่องมือที่ใช้คือแบบสอบคุณานยาว 10, 30 และ 50 ข้อ
ซึ่งแบบสอบทั้งกล่าวเป็นแบบสอบความถนัดเชิงวิชาการ (Scholastic Aptitude Test)
ผลการเปรียบเทียบปรากฏว่า วิธีของสวามินาธานและคณะ (Swaminathan, et al.)
คำนวณง่ายและให้ผลการประมาณที่ไม่ค่อยลำเอียง แต่ต้องใช้ข้อมูลจากการสอบ 2 ครั้ง และ
การประมาณค่าจะเกิดความคลาดเคลื่อนมากถ้าใช้กลุ่มตัวอย่างขนาดเล็ก สำหรับวิธีของเฮวน์
(Huynh) สับโคเวียค (Subkoviak) มาร์แชลและแฮร์เทล (Marshall
and Haertel) ซึ่งใช้ข้อมูลจากการสอบเพียงครั้งเดียวจะทำให้ความคลาดเคลื่อนในการ
ประมาณน้อย แต่วิธีการประมาณค่าจะยุ่งยาก ซับซ้อน และผลการประมาณจะลำเอียงมาก
ถ้าแบบสอบที่ใช้นั้นสั้นมาก ๆ

คริสเตนเซน (Chirstensen, 1981 : 3963-A) ได้ทำการศึกษาวิธีการ
นำแบบสอบอิง เกณฑ์ไปใช้กับการจัดการศึกษาแบบมีจุดประสงค์เป็นหลัก (Goal-Based
Educational Management system) กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียน
ระดับ 5 จำนวน 407 คน เครื่องมือที่ใช้ในการวิจัยเป็นแบบสอบวัดทักษะการอ่านจำนวน

96 ข้อ โดยวัดใน 4 ด้าน คือ การออกเสียง ไวยากรณ์ คำศัพท์ 56 ข้อ และวัดความเข้าใจ จำนวน 40 ข้อ จากการศึกษาพบว่าค่าความเที่ยงของแบบสอบอิง เกณฑ์ทุกฉบับมีค่ามากกว่า .85 ขึ้นไป

จากงานวิจัยของต่างประเทศจะศึกษาค่าความเที่ยงของแบบสอบอิง เกณฑ์ในลักษณะของการเปรียบเทียบค่าความเที่ยงที่ได้จากการใช้วิธีที่แตกต่างกัน สำหรับในประเทศไทยนั้น จะศึกษาในลักษณะของการประยุกต์เอาวิธีการประมาณค่าความเที่ยงวิธีต่าง ๆ มาใช้กับแบบสอบอิง เกณฑ์ ซึ่งส่วนใหญ่จะเป็นวิชาทาง คณิตศาสตร์ ในระดับประถมศึกษา และมัธยมศึกษา เช่น กาญจนาวรรณสุนทร (2521) ได้สร้างแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ เรื่องสมการ ชั้นมัธยมศึกษาปีที่ 1 จำนวน 40 ข้อ แยกเป็นฉบับย่อย ๆ 4 ฉบับ ๆ ละ 10 ข้อ ใช้กลุ่มตัวอย่างจำนวน 60 คน โดยกำหนดเกณฑ์ผ่านไว้ 60, 70 และ 80 เปอร์เซนต์ตามลำดับ จากการศึกษาพบว่า เกณฑ์ 60 เปอร์เซนต์ ให้ค่าความเที่ยงตามวิธีของลิฟวิงตัน (Livington) มีค่าสูงที่สุด ไพฑูรย์ เวทการ (2524) ได้สร้างแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ เรื่องอัตราส่วน ชั้นมัธยมศึกษาปีที่ 1 ใช้กลุ่มตัวอย่าง 132 คน แบบสอบมี 5 ฉบับย่อยยาวฉบับละ 10 ข้อ ประมาณค่าความเที่ยงโดยวิธีของเฮวัน (Huynh) พบว่าค่าความเที่ยงของแบบสอบมีค่าระหว่าง .05 - .31 สมถวิล วิจิตรวรรณ (2524) ได้ใช้วิธีการประมาณค่าความเที่ยงของแบบสอบอิง เกณฑ์โดยวิธีของสับโคเวียค (Subkoviak) กับแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ เรื่องการหาร ชั้นประถมศึกษาปีที่ 2 โดยใช้กลุ่มตัวอย่าง 124 คน พบว่าค่าความเที่ยงของแบบสอบมีค่าระหว่าง .61 - .78 บุญเลิศ คำหอม (2525) ได้ใช้วิธีการประมาณค่าความเที่ยงของแบบสอบอิง เกณฑ์โดยวิธีของสวามินาธานและคณะ (Swaminathan, et al.) หาค่าความเที่ยงของแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ เรื่องสมการ และอนุสมการ ชั้นมัธยมศึกษาปีที่ 2 ใช้กับกลุ่มตัวอย่างจำนวน 120 คน แบบสอบแบ่งออกเป็น 4 ฉบับย่อย ยาวฉบับละ 20 ข้อ พบว่าค่าความเที่ยงของแบบสอบที่ใหม่ค่าอยู่ระหว่าง .43 - .71 รังสรรค์ มณีเล็ก (2528) ได้เปรียบเทียบค่าความเที่ยงของแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ เรื่องเศษส่วน กับนักเรียนชั้นประถมศึกษาปีที่ 6 โดยวิธีของเฮวัน (Huynh) ซึ่งใช้วิธีการกำหนดจุดตัดที่ต่างกัน 4 วิธี ผลการศึกษาพบว่าค่าความเที่ยงที่ได้ไม่แตกต่างกัน สุรินทร์ แห่งจันทัก (2528) ได้เปรียบเทียบค่าความเที่ยงของแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ เรื่องควอราติก

ชั้นมัธยมศึกษาปีที่ 3 ระหว่างแบบสอบประเภทเลือกตอบและประเภทคอบสั้น ประมาณค่าความเที่ยงโดยใช้วิธีของโลเวตต์ (Lovett) พบว่าแบบสอบประเภทเลือกตอบมีค่าความเที่ยงน้อยกว่าแบบคอบสั้น

จากการวิจัยเกี่ยวกับเรื่องความเที่ยงของแบบสอบอิง เกณฑ์ของต่างประเทศ และในประเทศไทยจะเห็นได้ว่าเริ่มมีการศึกษาและให้ความสำคัญมากขึ้นเป็นลำดับ และจากผลการศึกษาของสับโคเวียค (Subkoviak, 1978 :111-115) พบว่าวิธีการหาค่าความเที่ยงซึ่งเสนอโดย สวามินาธานและคณะ (Swaminathan, et al.) คำนวณได้ง่ายให้ผลการประมาณค่าที่ไม่ลำเอียง กว้างเหตุผลทั้งกล่าวนี้ผู้วิจัยจึงใช้วิธีของสวามินาธานและคณะ มาใช้ประมาณค่าความเที่ยงในการวิจัยครั้งนี้

3. ความทรงของแบบสอบอิง เกณฑ์และงานวิจัยที่เกี่ยวข้อง

ก. ความทรงของแบบสอบอิง เกณฑ์

การหาค่าความทรงของแบบสอบอิง เกณฑ์แบร์ก (Berk, 1980 : 87) ได้ให้ความเห็นว่า ความทรงของแบบสอบอิง เกณฑ์จะท้อมี 2 ประเภทคือ ความทรงเชิงเนื้อหา (Content Validity) และความทรงเชิงทฤษฎี (Construct Validity) ส่วน ฟิทซ์แพทริก (Fitzpatrick, 1981: 2091A) ได้ให้ความเห็นว่าความทรงของแบบสอบอิง เกณฑ์ ควรประกอบด้วย 3 ประเภทด้วยกัน โดยมีความทรงในการตัดสินใจ (Decision Validity) เพิ่มเข้ามาอีก 1 ประเภท ซึ่งความหมายของความทรงแต่ละประเภทมีดังนี้

ความทรงเชิงเนื้อหา (Content Validity) หมายถึงการตัดสินใจเกี่ยวกับความสอดคล้องกันระหว่าง เนื้อหาของข้อสอบกับจุดประสงค์ที่มุ่งวัด ความทรงเชิง เนื้อหานี้จะไม่เปลี่ยนแปลงไปตามเวลาหรือกลุ่มตัวอย่างที่ใช้ (Berk, 1980 : 84)

ความทรงเชิงทฤษฎี (Construct Validity) เป็นความทรงที่แสดงถึงความสามารถของข้อสอบในการวัดสิ่งที่ต้องการจะวัดได้ถูกต้องตรงตามทฤษฎีที่กำหนดไว้ การหาค่าความทรงชนิดนี้ อาจคำนวณได้โดยใช้วิธีการทดลอง เชิงประจักษ์กับกลุ่มตัวอย่างที่มีความรอบรู้ โดยที่กลุ่มที่มีความรอบรู้ควรจะทำข้อสอบข้อนั้นได้ถูกต้องและกลุ่มตัวอย่างที่ไม่รอบรู้ควรจะทำข้อสอบข้อนั้นผิด (Berk, 1980 : 94)

ความทรงในการตัดสินใจ (Decision Validity) เป็นความทรงที่แสดงถึง
ความถูกต้องของการตัดสินใจจำแนกผู้สอบออกเป็นผู้รอบรู้ - ไม่รอบรู้ได้ตรงตามจุดมุ่งหมาย
ในการวัด (Fitzpatrick, 1981 : 2091-A)

ข. การประมาณค่าความทรงของแบบสอบอิงเกณฑ์

1. ความทรงเชิงเนื้อหา (Content Validity)

การหาค่าความทรงเชิงเนื้อหา เบอร์ก (Berk, 1980 : 87) ได้เสนอ
แนวทางในการหาไว้ โดยเกี่ยวข้องกับการตัดสินใจของผู้เชี่ยวชาญในค่านี้ออกข้อสอบ
ซึ่งได้เสนอไว้ 3 วิธีการดังนี้

1.1 หากดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ (The index
of item objective Congruence) วิธีการนี้ใช้ผลการตัดสินใจของผู้เชี่ยวชาญในข้อสอบ
แต่ละข้อว่าสอดคล้องหรือใช้จุดประสงค์ที่กำหนดไว้หรือไม่ โดยกำหนดวิธีการให้คะแนนผล
การตัดสินใจว่า

กำหนดให้คะแนนเป็น	+ 1	เมื่อแน่ใจว่าข้อสอบวัตถุประสงค์ที่กำหนดให้
	0	เมื่อไม่แน่ใจว่าข้อสอบนั้นวัตถุประสงค์ที่กำหนดให้
	- 1	เมื่อแน่ใจว่าข้อสอบนั้นไม่ใช่วัตถุประสงค์ที่
		กำหนดให้

จากผลการตัดสินใจและวิธีการให้คะแนนดังกล่าวจะนำมาคำนวณหาดัชนีความ
สอดคล้องโดยใช้สูตร (Berk, 1980 : 89)

$$I_{ik} = \frac{(N-1) \sum_{j=1}^n x_{ijk} - \sum_{i=1}^N \sum_{j=1}^n x_{ijk} + \sum_{j=1}^n x_{ijk}}{2(N-1) n}$$

เมื่อ	I_{ik}	คือ ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ สำหรับข้อสอบที่ k และจุดประสงค์ที่ i
	N	คือ จำนวนจุดประสงค์เชิงพฤติกรรม (i=1,2,...,N)

n คือ จำนวนผู้เชี่ยวชาญ

X_{ijk} คือ ผลการตัดสิน (+1, 0, -1) สำหรับข้อสอบข้อที่ k ว่าวัตถุประสงคที่ i โดยผู้เชี่ยวชาญที่ j หรือไม่

การกำหนดค่าข้อสอบนั้น ๆ วัตถุประสงคั้น ๆ จริงหรือไม่ มีผู้เสนอค่าทัศนคติความสอดคล้องไว้หลายค่า เช่น .75 (Berk, 1980 : 89) หรือ .50 (บุญเชิด วิทยุอนันต์พงษ์ 2527 : 69) ถ้าหากข้อสอบมีค่าทัศนคติความสอดคล้องตามเกณฑ์ที่กำหนดไว้จะได้รับการคัดเลือกไว้ ส่วนข้อสอบที่เหลือจะกักทิ้งหรือนำมาปรับปรุงต่อไป

นอกจากสูตร (1) แล้ว ยังมีสูตรทั้งายและสะดวกต่อการคำนวณ คือ

$$IOC = \sum R/N$$

- เมื่อ IOC คือ ทัศนคติความสอดคล้องระหว่างข้อสอบกับจุดประสงค์
- R คือ ผลรวมคะแนนความถี่เห็นของผู้เชี่ยวชาญทั้งหมด
- N คือ จำนวนผู้เชี่ยวชาญทางด้านเนื้อหา

1.2 หากความเหมาะสมระหว่างข้อสอบกับจุดประสงค์โดยใช้มาตราประมาณค่า (Rating Scale Procedure) วิธีนี้ผู้เชี่ยวชาญในทางด้านเนื้อหาจะตรวจสอบความเหมาะสมระหว่างข้อสอบกับจุดประสงค์จากมาตราประมาณค่าที่สร้างขึ้น

1.3 ใช้วิธีการจับคู่ระหว่างข้อสอบกับจุดประสงค์ (Matching Task Procedure) วิธีการนี้ผู้เชี่ยวชาญในทางด้านเนื้อหาจะเก็บรายละเอียดของข้อสอบกับจุดประสงค์ที่จะวัด แล้วให้ตัดสินว่าข้อสอบวัดได้ตามจุดประสงค์หรือไม่จากนั้นนำผลการตัดสินของผู้เชี่ยวชาญมาสร้างตารางการแจกแจงความถี่ แล้วนำมาทดสอบความเป็นอิสระ ไคสแควร์ (The Chi-Square test for independence) (Berk, 1980 : 92) ถ้าผลการทดสอบพบว่ามีความสำคัญก็แสดงว่าข้อสอบและจุดประสงค์มีความสัมพันธ์กัน หรือวัดได้จริงตามจุดประสงค์นั้น

2. ความตรงเชิงทฤษฎี (Construct Validity)

การหาค่าความตรงเชิงทฤษฎีมีผู้เสนอวิธีการไว้ดังนี้

2.1 วิธีของเกรเฮน (Crehen อ้างถึงใน บุญเชิด วิทยุอนันต์พงษ์,

2527 : 85) ซึ่งมีดังนี้

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

เมื่อ	c	เป็นค่านีความตรงของแบบสอย
	U	เป็นจำนวนผู้ขอสอบถูกในกลุ่ม N_1
	L	เป็นจำนวนผู้ขอสอบถูกในกลุ่ม N_2
	N_1	เป็นจำนวนผู้สอยในกลุ่มที่ได้รับการสอน
	N_2	เป็นจำนวนผู้สอยที่ไม่ได้รับการสอน

2.2 วิธีของคาร์เวอร์ (Carver quoted in Crehen, 1974 : 256)

ซึ่งได้เสนอสูตรในการคำนวณค่านีความตรงของแบบสอยอิง เกณฑ์จากข้อมูลของกลุ่มผู้เรียนและยังไม่ได้เรียนค่านี

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

เมื่อ	a	คือกลุ่มที่ผ่านเกณฑ์หลังจากได้รับการเรียนการสอน
	b	คือกลุ่มที่ผ่านเกณฑ์โดยไม่ได้รับการเรียนการสอน
	c	คือกลุ่มที่ไม่ผ่านเกณฑ์และไม่ได้รับการเรียนการสอน
	d	คือกลุ่มที่ไม่ผ่านเกณฑ์หลังจากได้รับการเรียนการสอน

โดยการแจกแจงผู้สอยลงในตารางดังต่อไปนี้

	กลุ่มที่ไม่ได้รับการเรียนการสอน	กลุ่มที่ได้รับการเรียนการสอน
ผ่านเกณฑ์	b	a
ไม่ผ่านเกณฑ์	c	d

วิธีของคาร์เวอร์นี้อาจคำนวณเป็นแบบสหสัมพันธ์ที่ (ϕ - Coefficient) ได้

ค่านี

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

- เมื่อ
- a คือ จำนวนผู้ผ่านหลังเรียน
 - b คือ จำนวนผู้ผ่านก่อนเรียน
 - c คือ จำนวนผู้ไม่ผ่านก่อนเรียน
 - d คือ จำนวนผู้ไม่ผ่านหลังเรียน

2.3 วิเคราะห์ตัวประกอบ (Factor analysis) วิเคราะห์ความสัมพันธ์ระหว่างข้อสอบด้วยกันเป็นรายคู่เป็นพื้นฐานสำคัญ โดยใช้ผู้สอบที่สามารถแบ่งได้เป็น 2 กลุ่ม คือกลุ่มที่สอบและกลุ่มที่ไม่สอบ ค่าสหสัมพันธ์ระหว่างข้อสอบแต่ละคู่จะนำมาวิเคราะห์ตัวประกอบ ซึ่งจะทำให้ได้กลุ่มข้อสอบที่รวมกันเป็นตัวประกอบหรือตัวประกอบประสมของข้อสอบกลุ่มนั้นถูกสร้างขึ้นมา เพื่อจะวัด ข้อสอบที่ไม่ได้อยู่ในตัวประกอบนี้ จะถือว่าไม่มีความตรงเชิงทฤษฎีควรมานำไปปรับปรุงให้คุณภาพหรือทิ้งไป (Berk, 1980 :96)

3. ความตรงในการตัดสินใจ (Decision Validity)

การหาความตรงในการตัดสินใจนั้น มีลักษณะคล้ายคลึงกับการหาความตรงในเชิงทฤษฎี แต่จะแตกต่างกันเพียงเล็กน้อยในการแบ่งกลุ่มผู้สอบ ซึ่งแทนที่จะแบ่งผู้สอบออกเป็น ผู้ได้รับการสอนและไม่ได้รับการสอน ก็แบ่งเป็นผู้สอบและไม่สอบแทน

จากความหมายและแนวคิดของการหาความตรงเชิงทฤษฎีและความตรงในการตัดสินใจของแบบสอบอิงเกณฑ์ทั้งกล่าวนี้ จะเห็นได้ว่าแนวคิดพื้นฐานและสมโนทัศน์เกี่ยวกับทฤษฎีของการวัดแบบอิงเกณฑ์ว่า เครื่องมือที่ใช้ควรเป็นเครื่องมือที่สามารถจะแยกผู้เรียนแล้ว และผู้ที่ยังไม่ได้เรียนออกจากกันได้ ดังนั้นความหมายของความตรงเชิงทฤษฎีและความตรงในการตัดสินใจ ตามที่ผู้เสนอไว้ดังกล่าวนี้จึงใกล้เคียงกันมาก ในการวิจัยครั้งนี้ผู้วิจัยจึงได้หาความตรงใน 2 ด้านในความหมายเดียวกัน คือ ความตรงในการตัดสินใจผู้สอบ ทั้งนี้เพราะการหาค่าความตรงทั้ง 2 ด้านดังกล่าวนี้ผู้เสนอสูตรในการคำนวณเอาไว้เป็นสูตรเดียวกัน ซึ่งในการวิจัยครั้งนี้ได้ใช้วิธีของ คาร์เวอร มาคำนวณหาความตรง เพราะเป็นวิธีที่ง่าย สะดวก และไม่ยุ่งยากต่อการคำนวณและการแปลความหมาย

ก. งานวิจัยที่เกี่ยวข้องกับความทรงของแบบสอบอิง เกณฑ์

การศึกษาค่าความทรงของแบบสอบอิง เกณฑ์นั้นยังไม่แพร่หลายเหมือนกับการหาค่าความเที่ยง ซึ่งในต่างประเทศและในประเทศไทยก็มีผู้ศึกษากันไว้บ้าง ดังนี้

ลอควูด (Lockwood, 1977 : 2207-A) ได้ศึกษาวิธีการพัฒนาและประยุกต์เอาวิธีการใช้แบบสอบอิง เกณฑ์ใหม่ความทรงต่อการนำไปใช้จัดการเรียนการสอน วิธีศึกษาได้ทำการเปรียบเทียบคะแนนการรับรู้ของนักเรียน โดยการใช้ผลการสอบของนักเรียน กับผลการทัศนโยครุ กลุ่มตัวอย่างที่ใช้ในการศึกษาคั้งนี้คือ ครู และนักเรียนจำนวน 115 คน โดยใช้ข้อสอบ 53 ข้อ ซึ่งเป็นเนื้อหาวิชาเรขาคณิต กำหนดเกณฑ์ผ่านไว้ที่ 80 เปอร์เซ็นต์ ซึ่งในการศึกษาได้ทั้งสมมติฐานในการวิจัยไว้ 2 ข้อ คือ

1. การจำแนกผู้เรียนออกเป็นผู้รอบรู้ระหว่างวิธีการทัศนโยครุกับผลการสอบของนักเรียนจะไม่มี ความคงที่
2. สัดส่วนการทำข้อสอบ ใ้ของนักเรียนในแต่ละข้อจะไม่แตกต่างกันทั้งก่อนเรียน และหลังเรียน

จากการศึกษาพบว่าปฏิเสธสมมติฐานข้อที่ 1 เพราะมีความคงที่ของการจำแนกถึง 44 จาก 53 ข้อ และปฏิเสธสมมติฐานข้อที่ 2 เพราะว่าสัดส่วนของนักเรียนมีความแตกต่างกัน 45 ข้อ จาก 53 ข้อ นอกจากนี้ยังพบว่าข้อสอบ 4 ข้อ เป็นอิสระจากทั้งการทัศนโยครุ และการสอบ และไม่มี ความไวต่อการเรียนการสอนอีกด้วย จากการศึกษาในเรื่องนี้จึง เป็นการแสดงให้เห็นถึงการทัศนโยครุเรียนว่าเป็นผู้รอบรู้ โดยดูจากผลการทัศนโยครุว่ามีความแตกต่างกันอย่างไร ซึ่งเป็นการศึกษาถึงความทรงในการทัศนโยครุผู้นั้นเอง

กาญจนา วันสุนทร (2521) ทดลองกำหนดจุดตัดของแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ เรื่องสมการ ฐานมัธยมศึกษาปีที่ 1 จำนวน 40 ข้อ แยกเป็นฉบับย่อย ๆ 4 ฉบับ ฉบับละ 10 ข้อ โดยใช้กลุ่มตัวอย่าง 60 คน กำหนดเกณฑ์ไว้ 60, 70 และ 80 เปอร์เซ็นต์ ตามลำดับ จากการศึกษาพบว่าเกณฑ์ 60 เปอร์เซ็นต์ ให้ค่าความทรงตามวิธีของคาร์เวอ (Carver) มีค่าสูงที่สุด

ขมภู จันทอมรพร (2522) ได้ปรับปรุงแบบสอบของ กาญจนา วัฒนสุนทร ที่สร้างไว้ แล้วคำนวณจุดตัดโดยวิธีของ เบส ใช้กลุ่มตัวอย่างจำนวน 80 คน คัดเลือกจุดตัดโดย ใช้ค่าความตรงของการเวอ์ ใช้จุดตัดของแบบสอบฉบับที่ 1,2,3 และ 4 เท่ากับ 6,5,5 และ 5 ตามลำดับ

บุญเลิศ คำหอม (2525) ได้ใช้วิธีการกำหนดจุดตัดแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ เรื่องสมการและอสมการ ชั้นมัธยมศึกษาปีที่ 2 จำนวน 4 ฉบับ ๆ ละ 20 ข้อ โดยใช้วิธีจาก เกณฑ์เฉลี่ยของครูผู้สอนและวิธีของ เบอร์ก ใช้จุดตัดซึ่งทำให้ค่าความตรงตามวิธีของการเวอ์ มีค่าสูงสุกเท่ากับ 50, 60, 60 และ 50 เปอร์เซนต์ตามลำดับ

4. การกำหนดจุดตัดของแบบสอบอิง เกณฑ์และงานวิจัยที่เกี่ยวข้อง

ก. ทฤษฎีและแนวคิด เกี่ยวกับการกำหนดจุดตัด

ในการวัดผลสัมฤทธิ์ทางการเรียน แกลเซอร์ (Glaser, 1969) ได้กล่าวว่า ผลสัมฤทธิ์ที่ได้จากการเรียนการสอนจะวัดผลเป็นที่น่าพอใจหรือไม่นั้นจะสามารถประเมินได้ด้วยการวัดแบบอิง เกณฑ์ ซึ่งการแปลความหมายของการวัดแบบอิง เกณฑ์นั้นน่าจะแนบไปเปรียบเทียบกับ เกณฑ์ที่กำหนดขึ้น เกณฑ์ที่กำหนดขึ้นนี้อาจจะกำหนดที่จุดใด ๆ ของการสอบก็ได้ ซึ่งจะเป็น จุดตัด (Cut-off) ระหว่างความสามารถและความไม่สามารถของบุคคล คะแนนที่จุดนี้ จะใช้ในการตัดสินใจจำแนกผู้สอบออกเป็น ผู้รอบรู้ (Master) และ ไม่รอบรู้ (Non-master) หรือแยกผู้สอบออกเป็น ผู้สอบผ่านหรือสอบตก (Pass or Fail) มีนักวิชาการทางด้าน การวัด และประเมินผลได้กล่าวถึงคะแนนเกณฑ์หรือจุดตัดไว้หลายท่านเช่น โทแพม และฮูเสค (Popham and Husek, 1969) กล่าวถึงการวัดแบบอิง เกณฑ์ไว้ว่า เป็นการวัดที่ ค้นหาสถานภาพของแต่ละบุคคลโดยอาศัย เกณฑ์หรือมาตรฐานการปฏิบัติ (Standard setting Performance) ส่วน เบอร์ก (Berk, 1980) ใช้คำว่ามาตรฐาน (Standard) ในความหมายว่าเป็นคะแนนที่คาดหวังว่าผู้สอบจะต้องทำคะแนนให้ได้ถึง และถ้าผู้สอบทำคะแนนได้ไม่ถึงจะต้อง ใ้รับการฝึกฝนจนสามารถทำได้ อย่างสมบูรณ์หรือ เกือบสมบูรณ์

จะเห็นได้ว่าจุดตัด ซึ่งบางที่เรียกว่าเกณฑ์ (Criteria) คะแนนมาตรฐาน (Standard) คะแนนจุดผ่าน (Passing score) ระดับความรอบรู้ (Mastery Level) ความสามารถขั้นต่ำสุด (Minimal Competencies) หรือคะแนนจุดแบ่งของ

แบบสอบอิง เกณฑ์ ซึ่งความหมายเหล่านี้คือความหมายอันเกี่ยวกันในแง่ของคะแนนที่น้อยที่สุดที่นักเรียนจะต้องทำได้ในการที่จะได้รับการตัดสินให้เป็นผู้ชรูนั้นเอง (อังคณา สายยศ, 2525 : 70) ในการกำหนดจุดตัดหรือการกำหนดเกณฑ์จึงเป็นสิ่งที่สำคัญมาก เพราะถ้าการกำหนดจุดตัดไม่ถูกต้องแล้ว จะทำให้เกิดความคลาดเคลื่อนในการวัดขึ้นได้ใน 2 ลักษณะ ดังต่อไปนี้คือ (Berk, 1986 : 138 - 239)

1. ความคลาดเคลื่อนที่เกิดจากการจำแนกผู้เรียนที่ชรูแล้วให้เป็นผู้ที่ไม่ชรูหรือเป็นผู้ที่ไม่มีความสามารถ (False non-master/incompetent)
2. ความคลาดเคลื่อนที่เกิดจากการจำแนกผู้เรียนที่ไม่ชรูให้เป็นผู้ที่ชรูหรือเป็นผู้ที่มีความสามารถ (False master/Competent)

ถึงอย่างไรก็ตามแม้ว่าจะได้มีการพยายามลดความคลาดเคลื่อนทั้งสองอย่างดังกล่าวถึงขั้นก็ตาม แต่ก็มีสถานการณ์บางอย่างที่อาจทำให้เกิดความคลาดเคลื่อนขึ้นได้ โดยเฉพาะอย่างยิ่งการเกิดความคลาดเคลื่อนในข้อที่ 2 จะมีผลเสียแก่ผู้เรียนมากกว่าความคลาดเคลื่อนในข้อที่ 1

ข. วิธีกำหนดจุดตัดของแบบสอบอิง เกณฑ์

การกำหนดจุดตัดของแบบสอบอิง เกณฑ์ให้มีผู้ศึกษาวิธีการต่าง ๆ เอาไว้อย่างกว้างขวาง ในขั้นผู้วิจัยจะนำเสนอเฉพาะวิธีการที่สำคัญ ๆ และนำเสนอใจดังนี้

แกลส (Glass, 1978: 237-259) ได้รวบรวมวิธีการหาคะแนนจุดตัดโดยพิจารณาจากสิ่งต่าง ๆ ต่อไปนี้

1. ไซคาเปอร์เซนต์ไทลของคะแนนที่สอบได้จากกลุ่มตัวอย่างนักเรียนที่มีความชรู
2. เมื่อหลังจาก 100 เปอร์เซนต์ วิธีนี้จะให้ผู้เชี่ยวชาญทางด้านเนื้อหาวิชาหรือครูผู้สอนพิจารณาความสำคัญของจุดประสงค์และกำหนดการอยละความสำคัญ
3. กำหนดจุดตัดจากสมรรถภาพขั้นต่ำสุดโดยให้ผู้เชี่ยวชาญในเนื้อหาวิชาหรือครูประจำวิชาศึกษาข้อสอบแล้วระบุว่าคะแนนเท่าใ้จึงจะยอมรับว่าเป็นผู้ชรู
4. ปรับคะแนนตามเกณฑ์อื่น ๆ วิธีนี้ใช้เกณฑ์ภายนอกแบ่งเป็นกลุ่มผู้ชรูและไม่ชรู โดยนำแบบสอบไปสอบกับทั้ง 2 กลุ่ม แล้วหาคะแนนจุดตัดที่มีความสอดคล้องกับเกณฑ์ภายนอก

5. ใช้ผลการวิจัยเชิงปฏิบัติ วิธีนี้เป็นการกำหนดจุดศึกษาจากผลการทดลองที่
 ตามมาโดยอาศัยโครงการแจกแจงทางคณิตศาสตร์หรือกราฟช่วยในการตัดสินใจความสัมพันธ์ของ
 คะแนนจุดศึกษาแบบสลับกับเกณฑ์ภายนอกที่กำหนดขึ้น

6. ใช้ทฤษฎีการตัดสินใจ โดยอาศัยเกณฑ์ภายนอกอย่างใดอย่างหนึ่งที่เกี่ยวข้องกับ
 กับสิ่งที่จะศึกษา แบ่งกลุ่มคนเป็น 2 กลุ่ม แล้วให้ทั้ง 2 กลุ่มทำแบบสลับอิงเกณฑ์ แล้วกำหนด
 จุดศึกษามาใช้

เบอร์ก (Berk, 1980 : 101) ได้เสนอวิธีการพื้นฐานเพื่อใช้ในการพิจารณา
 กำหนดจุดศึกษาใช้ดังนี้

1. พิจารณาจากเนื้อหาของข้อสอบ
2. พิจารณาจากการเคาะและการสัมผัสข้อสอบมาใช้
3. ใช้ข้อมูลเชิงประจักษ์จากกลุ่มที่รอบรู้และไม่รอบรู้
4. พิจารณาจากผลที่ตามมาในทางการศึกษา
5. ใช้ทฤษฎีการตัดสินใจ
6. ใช้เกณฑ์การวัดภายนอก

เบอร์ก (Berk, 1986 : 137-171) ได้รวบรวมวิธีการกำหนดจุดศึกษาของแบบสลับ
 อิงเกณฑ์เอาไว้ซึ่งมีทั้งหมด 38 วิธี โดยการรวบรวมไว้จำแนกการกำหนดจุดศึกษาออกเป็น 2
 โมเดล คือ State Model และ Continuum Model การจำแนกตามแบบ State
 Model นั้นมีแนวคิดมาจากการตัดสินใจความสามารถของบุคคลซึ่งแบ่งได้เพียง 2 อย่างเท่านั้น
 คือ การมีความสามารถและการไม่มีความสามารถ จุดศึกษาจำแนกตามวิธีนี้มีได้เพียงค่าเดียว
 เท่านั้น ส่วนการจำแนกตามแบบ Continuum Model มีแนวคิดจากการพิจารณาความรู้ หรือ
 มีความสามารถของบุคคลนั้นสามารถแบ่งได้หลายระดับต่อเนื่องกันไป การกำหนดจุดศึกษา
 วิธีนี้มีได้หลายค่า ซึ่งรายละเอียดของการกำหนดจุดศึกษาตามที่ เบอร์ก (Berk) ได้รวบรวม
 ไว้ให้นำเสนอในตารางที่ 1 ดังต่อไปนี้

ตารางที่ 1 แสดงการจำแนกการกำหนดคุณลักษณะตามแบบ State Model และ Continuum Model ซึ่งมีทั้งหมด 38 วิธี ซึ่งรวบรวมโดยเบอร์ก (Berk)

State	Continuum			
	Judgmental (Setting standards)	Judgmental-empirical (Setting standards)	Empirical-judgmental	
Empirical-judgmental (Adjusting standards)			(Setting standards)	(Adjusting standards)
Berk (1984, chap. 6)	Adjusted/modified M-C Angoff (Bernknopf, Curry, & Bashaw, 1979)	Absolute-relative compromise I (Beuk, 1982, 1984)	Borderline group (Livingston & Zieky, 1982)	Huynh (1976)
Bergan, Cancelli, & Luiien (1980)	Angoff (1971)	Absolute-relative compromise II (Hofstee, 1977, 1983)	Contrasting groups (Livingston & Zieky, 1982)	Kriewall (1972)
Emrick (1971)	Angoff-Nedelsky combination (Reid, 1984)	Angoff-contrasting groups plus composite (Shepard, 1980b, 1984)	Criterion groups (Berk, 1976, 1984, chap. 6, in press)	Livingston (1975)
Knapp (1977)	Difficulty-importance estimate (Cangelosi, 1984)	Informed judgment (Yalow & Popham, 1983)	Educational consequences (Block, 1972)	Livingston (1980)
Macready & Dayton (1977, 1980)	Difficulty-relevance Ebel (Ebel, 1979, pp. 339-340)	Iterative Angoff (Saunders & Mappus, 1984)	Norm-referenced criterion (Garcia-Quintana & Mappus, 1980)	Novick & Lewis (1974)
Roudabush (1974)	Difficulty-taxonomy Ebel (Skakun & Kling, 1980)	Iterative two-choice Angoff (Jaeger, 1978, 1982)		van der Linden & Mellenbergh (1977)
Wilcox (1977a, 1977b)	Item specifications (Mills & Barr, 1983)	Modified Angoff-empirical (McLean & Halpin, 1984)		Wilcox (1979)
Wilcox (1981)	Modified M-C Angoff (ETS, 1976)			
	Nedelsky (1954)			
	Relevance-taxonomy Ebel (Skakun & Kling, 1980)			
	Two-choice Angoff (Nassif, 1978)			

จากตารางที่ 1 จะพบว่า 79 เปอร์เซ็นต์ของวิธีการกำหนดคุณลักษณะทั้งหมดจะจัดอยู่ในประเภท Continuum Model และ 3 ใน 4 ของวิธีการกำหนดคุณลักษณะทั้งหมด จะเป็นวิธีการตัดสิน (Judgmental) (Berk, 1986 : 141)



ค. งานวิจัยที่เกี่ยวข้องกับจุดตัดของแบบสอบอิง เกณฑ์

ในต่างประเทศใหม่การศึกษาค้นคว้าวิจัยเกี่ยวกับการกำหนดจุดตัดของแบบสอบอิง เกณฑ์เป็นเวลานานแล้ว จึงมีงานวิจัยเกี่ยวกับการกำหนดจุดตัดมีจำนวนหลายเรื่อง ซึ่งได้ศึกษาในขั้นของการนำมาประยุกต์ใช้และการศึกษาในเชิงทฤษฎี ผู้วิจัยขอเสนอผลงานวิจัยที่เกี่ยวข้องพอสังเขปดังนี้

เบฮูเนียค (Behuniak, 1981 : 3998-A) ได้ศึกษาวิธีการหาจุดตัดของแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ และการอ่านของนักเรียนระดับ 9 ตามวิธีของนีเคลสกีและวิธีของแองกอฟ โดยให้ผู้เชี่ยวชาญทางคณิตศาสตร์จำนวน 27 คน พบว่า

1. การใช้วิธีการกำหนดจุดตัดที่ต่างกันจะทำให้ได้จุดตัดที่ต่างกัน
2. การใช้กลุ่มตัวอย่างที่ต่างกันกำหนดจุดตัดวิธีเดียวกันจะทำให้ได้จุดตัดที่ต่างกัน
3. การกำหนดจุดตัดตามวิธีของแองกอฟ และวิธีของนีเคลสกี จะให้จุดตัดที่มีความแปรเปลี่ยนไปเท่า ๆ กัน
4. กลุ่มที่มีการเปลี่ยนแปลง และเนื้อหาสัมพันธ์กัน เป็นองค์ประกอบที่เกี่ยวกับการกำหนดจุดตัด

5. ความสัมพันธ์ระหว่างประสิทธิภาพของแบบสอบ และการจำแนกกับผลสัมฤทธิ์พื้นฐานของนักเรียนไม่คำนึงถึงวิธีการที่ใช้ในการกำหนดจุดตัด

6. การจำแนกผลการปฏิบัติของนักเรียนออกเป็น 3 กลุ่ม คือ กลุ่มเชี่ยวชาญ กลุ่มที่ยังไม่แน่ใจว่าเชี่ยวชาญ และกลุ่มที่ไม่เชี่ยวชาญ จะช่วยลดความคลาดเคลื่อนในการจำแนกผิดได้ดีกว่าการจำแนกนักเรียนเป็น 2 กลุ่ม คือ กลุ่มเชี่ยวชาญ และไม่เชี่ยวชาญ

ฮาราสิม (Harasym, 1981 : 725 - 727) ได้ทำการเปรียบเทียบวิธีการกำหนดจุดตัดของแบบสอบอิง เกณฑ์ โดยใช้วิธีของแองกอฟ และนีเคลสกี กับกลุ่มตัวอย่างซึ่งเป็นนักศึกษาแพทย์ของมหาวิทยาลัยคัลการี จำนวน 212 คน จากการศึกษาพบว่าจุดตัดที่ได้จากวิธีของนีเคลสกี มีค่าต่ำกว่าวิธีของแองกอฟ และพบว่าวิธีของนีเคลสกี ให้ผลในการจำแนกได้เป็นที่น่าพอใจกว่าวิธีของแองกอฟ

แวนเคอลินเดน (Vanderlinden, 1982 : 205-307) ให้นำเอาทฤษฎี
คุณลักษณะภายใน (Latent trait Theory) แบบ 3 พารามิเตอร์ มาใช้กำหนด
จุดตัดตามวิธีของนีเคลสกี และแองกอฟ เพื่อพิจารณาความไม่คงที่ภายในของการตัดสิน โดย
ให้ผู้ตัดสินแต่ละคนตัดสินตามวิธีของนีเคลสกี 8 คน และวิธีแองกอฟ 9 คน นำผลการตัดสิน
ของแต่ละคนมาเปลี่ยนเป็นค่าความคลาดเคลื่อนสมบูรณ์ (Absolute errors)
ปรากฏว่าค่าเฉลี่ยของความคลาดเคลื่อนสมบูรณ์ตามวิธีของแองกอฟต่ำกว่าวิธีของนีเคลสกี

ครอส อิมพารา ฟรารี และเจเกอร์ (Cross, Impara, Frary and
Jaeger, 1984: 113-123) ได้ทำการเปรียบเทียบวิธีกำหนดเกณฑ์ตามระดับค่า สุก 3 วิธี
คือ วิธีของแองกอฟ นีเคลสกี และเจเกอร์ เพื่อนำผลที่ได้ไปกำหนดเป็นเกณฑ์ของแบบสอบ
ใน NTE (National Teacher Examination) ซึ่งมีทั้งหมด 18 ชุด วิชาที่ใช้ในการ
ศึกษา คือ คณิตศาสตร์และการศึกษาเบื้องต้น กลุ่มตัวอย่างที่ให้พิจารณาตัดสินความสามารถต่ำสุด
คือ ครูที่ทำการสอนในระดับวิทยาลัยครูมาแล้ว ไม่น่ากว่า 2 ปี จำนวน 30 คน โดยแบ่งออก
เป็น 2 กลุ่ม กลุ่มละ 15 คน โดยให้กลุ่มแรกพิจารณาตัดสินวิชาคณิตศาสตร์ กลุ่มที่สองให้
พิจารณาวิชาการศึกษาเบื้องต้น ในการตัดสินจะให้ผู้ตัดสินให้พิจารณาการตัดสิน 3 ครั้ง แล้ว
นำวิธีการ รายวิชา จำนวนครั้ง มาเปรียบเทียบกันโดยใช้ ANOVA พบว่า

1. วิธีการและจำนวนครั้งที่กำหนดเกณฑ์ จะมีค่าแตกต่างกัน
2. การกำหนดเกณฑ์ตามวิธีของแองกอฟ มีความคงที่สูงที่สุด ส่วนวิธีของนีเคลสกี
จะให้ค่าคงที่ต่ำสุด
3. การประมาณค่าความเที่ยงของแบบสอบโดยใช้วิธีของอีเบล (Ebel, 1951) พบว่า
วิธีของแองกอฟ ให้ค่าความเที่ยงสูงที่สุด คือ .53 - .82 ส่วนวิธีของนีเคลสกีให้ค่าความเที่ยง
ต่ำสุดคือ ระหว่าง .38 - .53
4. ผลของการศึกษาในเรื่องนี้ จะใ้วิธีการซึ่งจะนำไปใช้ในการกำหนดจุดตัดของ
แบบสอบใน NTE จำนวน 18 ชุด คือวิธีของแองกอฟ เพราะใ้ผลมีความคงที่และเชื่อถือได้
กว่าวิธีอื่น ๆ

อราสมิท (Arrasmith, 1986 : 3400-A) ได้ศึกษาวิธีการกำหนดจุดตัด
ของแบบสอบอิง เกณฑ์โดยใช้วิธีการพิจารณาเนื้อหาของข้อสอบกับการตัดสินตัวบุคคลที่เข้าสอบ
ผลการเปรียบเทียบ 2 วิธีนี้พบว่าให้ผลไม่แตกต่างกัน

การศึกษาการกำหนดจุดตัดของแบบสอบอิง เกณฑ์ในประเทศไทย จะอยู่ในรูปของ
การนำเอาวิธีการต่าง ๆ มาประยุกต์ใช้กับเนื้อหาวิชาที่เป็นแบบสอบอิง เกณฑ์ ซึ่งส่วนใหญ่
จะเป็นวิชาทางด้านการคำนวณและคณิตศาสตร์ ทั้งนี้

สมหวัง พิธิยานุวัฒน์ (2520) ได้นำเอาวิธีการกำหนดจุดตัดของเบอร์ก
(Berk) มาใช้กำหนดจุดตัดของแบบสอบอิง เกณฑ์วิชาสถิติ เรื่องสหสัมพันธ์ ยาว 14 ข้อ
โดยใช้กลุ่มตัวอย่าง 18 คน พบว่าจุดตัดขึ้นต่ำอย่างสูงเท่ากับ 7.5 จุดตัดขึ้นสูงอย่างต่ำ
เท่ากับ 9.8

ไพฑูริย์ เวทการ (2524) ได้สร้างแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ เรื่อง
อัตราส่วน ชั้นมัธยมศึกษาปีที่ 1 โดยใช้กลุ่มตัวอย่างจำนวน 132 คน แบบสอบที่สร้างขึ้น
ได้แบ่งเป็น 5 ฉบับย่อย ฉบับละ 10 ข้อ กำหนดจุดตัดโดยใช้ทฤษฎีการตัดสินใจ ได้จุดตัด
มีค่าเท่ากับ 5 และ 6

บุญเลิศ คำหอม (2525) ได้ใช้วิธีการกำหนดจุดตัดแบบสอบอิง เกณฑ์ วิชา
คณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 2 วิธีคือ วิธีเบ็ดเตล็ดหลัง กับวิธีของเบอร์ก (Berk)
โดยใช้สร้างแบบสอบเป็น 4 ฉบับย่อย ยาวฉบับละ 20 ข้อ พบว่าจุดตัดที่ให้มีค่า
เท่ากับ 50 และ 60 เปอร์เซนต์ ตามลำดับ

วิฑูริย์ เสาวรัตน์ (2526) ได้ศึกษาเปรียบเทียบวิธีการกำหนดจุดตัดแบบสอบ
อิง เกณฑ์วิชาคณิตศาสตร์ เรื่องภาคทศกรวย ชั้นมัธยมศึกษาปีที่ 4 ใน 3 วิธี คือ วิธีของเบตต์
วิธีของแกลส และวิธีของเบอร์ก จากการศึกษาการกำหนดจุดตัดของแบบสอบย่อยทั้ง 11 ฉบับ
พบว่า จุดตัดที่ได้ตามวิธีของเบตต์มีค่าอยู่ระหว่าง 60 - 80 เปอร์เซนต์ วิธีของแกลสอยู่ระหว่าง
50 - 80 เปอร์เซนต์ และวิธีของเบอร์กอยู่ระหว่าง 40 - 80 เปอร์เซนต์

ชาวลิก โทษินคร (2528) ได้เปรียบเทียบผลของการกำหนดจุดตัดของแบบสอบ
อิง เกณฑ์ชนิดเลือกตอบโดยวิธีของเบอร์ก วิธีประยุกต์ราสโมเดล และวิธีกำหนดเกณฑ์ผ่าน
ระดับต่ำสุด โดยใช้ให้นักเรียนชั้นประถมศึกษาปีที่ 4 ปีการศึกษา 2527 ในโรงเรียนสังกัด

สำนักงานการประถมศึกษาจังหวัดนครสวรรค์ จำนวน 325 คน เครื่องมือที่ใช้ในการศึกษาคือ แบบสอบถามวัดความถนัดทางการเรียน จำนวน 2 ฉบับ เพื่อวัดระดับความสามารถของกลุ่มตัวอย่าง แบบสอบถาม เกณฑ์กลุ่มทักษะ วิชาคณิตศาสตร์ เพื่อวัดผลสัมฤทธิ์ของกลุ่มรอบรู้ ผลการศึกษาพบว่า

1. คะแนนผลสัมฤทธิ์ของกลุ่มรอบรู้ที่ผ่านจุดตัดที่กำหนดโดยวิธีของ เบอร์ก วิชประยุกต์ ราสซโมเกล และวิธีกำหนดเกณฑ์ตามระดับต่ำสุด แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และพบว่าค่าเฉลี่ยของกลุ่มรอบรู้ที่ผ่านจุดตัดโดยวิธีประยุกตราสซโมเกล ทำควากลุ่มที่ผ่านจุดตัดโดยวิธีกำหนดเกณฑ์ตามระดับต่ำสุดอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2. ค่าความตรงเชิงทฤษฎีของแบบสอบถาม เกณฑ์ที่ได้จากการกำหนดจุดตัดโดยวิธีของ เบอร์ก มีแนวโน้มให้ค่าสูง ร่องลงมาได้แก่วิธีประยุกตราสซโมเกล และวิธีกำหนดเกณฑ์ตามระดับต่ำสุดตามลำดับ

3. จุดตัดของแบบสอบถาม เกณฑ์ที่คำนวณโดยวิธีของ เบอร์ก ในกลุ่มตัวอย่างที่มีความสามารถสูง ปานกลาง และต่ำ พบว่าได้คะแนนจุดตัดที่แตกต่างกัน และมีแนวโน้มว่ากลุ่มสูงมีจุดตัดสูงที่สุด ร่องลงมาได้แก่กลุ่มปานกลาง และกลุ่มต่ำตามลำดับ

4. จุดตัดแบบสอบถาม เกณฑ์ที่คำนวณโดยวิธีประยุกตราสซโมเกล ในกลุ่มตัวอย่างที่มีความสามารถสูง ปานกลาง และต่ำ พบว่าในกลุ่มสูงได้คะแนนจุดตัดแตกต่างไปจากกลุ่มปานกลาง และกลุ่มต่ำ ส่วนระหว่างกลุ่มปานกลาง และกลุ่มต่ำ ได้คะแนนจุดตัดเท่ากัน แต่เมื่อคำนวณเปอร์เซ็นต์ของคะแนนจุดตัดพบว่าเท่ากับ 50 เปอร์เซ็นต์ของคะแนนทั้งหมด เมื่อคิดเฉพาะข้อที่เป็นไปตามข้อตกลงของราสซโมเกล

5. รูปแบบของราสซและทฤษฎีที่เกี่ยวข้อง

เนื่องจากทฤษฎีการวัดผลการศึกษาที่ใช้กันอยู่ในปัจจุบันนี้ คือทฤษฎีการวัดตามแนวเดิม (Classical Test Theory) ยังมีปัญหาในการวัดบางประการที่ไม่สามารถแก้ไขให้เป็นที่น่าพอใจได้ เช่น การวัดที่ไม่สามารถให้ผลตรงกับความสามารถที่แท้จริงของผู้สอบ เพราะคะแนนที่ได้ยังมีความคลาดเคลื่อนปนอยู่ด้วย กุญแจหนึ่งมีนักทดสอบทางด้านจิตวิทยาหลายท่านได้พยายามพัฒนาทฤษฎีการทดสอบแนวใหม่ขึ้นมาเพื่อแก้ไขปัญหการวัดตามแนวเดิมที่ไม่สามารถแก้ไขได้

ทฤษฎีใหม่ที่พัฒนาขึ้นและกำลังเป็นที่น่าสนใจ คือ ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) ซึ่งได้นำมาอธิบายลักษณะของข้อสอบแต่ละข้อคหพารามิเตอร์ของข้อสอบนั้น ๆ และอธิบายลักษณะของผู้สอบ ค่าพารามิเตอร์เหล่านี้สามารถนำไปอธิบายหรือทำนายโอกาสในการทำข้อสอบของผู้สอบ ถึงแม้ว่าผู้สอบจะมีลักษณะไม่คล้ายคลึงกับกลุ่มตัวอย่างที่ได้ค่าพารามิเตอร์นั้น ๆ ก็ตาม

บุคคลที่เสนอแนวคิดและหลักการเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ คือ Ferguson และ Lawley (Lord and Novick, 1968 : 369) ซึ่งออกมาในปี ค.ศ. 1952 Lord ได้เสนอทฤษฎีในรูปโค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) Lord ได้กล่าววาพฤติกรรมกรรมการตอบข้อสอบแต่ละข้อของผู้สอบสามารถอธิบายได้ด้วยรูปแบบทางคณิตศาสตร์ที่เรียกว่า Normal Ogive Model แต่เนื่องจากการคำนวณความรูปร่างแบบความยุ่งยาก ต่อมา Birnbaum จึงได้เสนอรูปแบบโลจิสติก (Logistic) ขึ้นมาเพื่อแก้ปัญหาคำนวณความยุ่งยากในการคำนวณ และสามารถนำมาใช้คำนวณโดยคอมพิวเตอร์ได้เป็นอย่างดีในปัจจุบันนี้

ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

1. แบบสอบจะคงมีมิติเดียว (Unidimension Test) หมายถึงข้อสอบแต่ละข้อจะคงวัดความสามารถหรือลักษณะเดียวกัน หรือข้อสอบมีความเป็นเอกพันธ์กัน
2. ข้อสอบแต่ละข้อจะคงเป็นอิสระจากกัน หมายถึงการตอบข้อสอบข้อหนึ่ง ๆ จะไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ
3. โอกาสที่ผู้สอบจะตอบข้อสอบใดถูกต้องขึ้นอยู่กับโค้งลักษณะข้อสอบแต่ละรูปแบบที่ใช้ไม่ขึ้นกับการแจกแจงความสามารถของกลุ่มตัวอย่าง (Lord and Novick, 1968 : 359)

รูปแบบต่าง ๆ ในทฤษฎีการตอบสนองข้อสอบ

ในการนำเสนอรูปแบบต่าง ๆ ของทฤษฎีการตอบสนองข้อสอบจะนำเสนอเพียงรูปแบบโลจิสติกเท่านั้น เพราะเป็นรูปแบบที่ใช้กันอยู่ในปัจจุบันนี้ ซึ่งรูปแบบนี้ได้แบ่งเป็น

รูปแบบย่อย ๆ 3 รูปแบบ คือ (Hambleton and Cook, 1977 : 81 - 82)

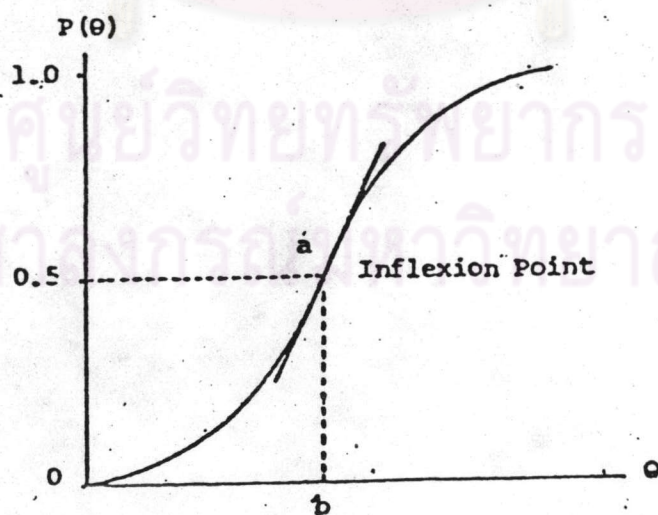
1. รูปแบบที่ใช้พารามิเตอร์ 2 ตัว (Two-parameters Logistic Model)

รูปแบบนี้ เบิร์นบวม (Birnbaum) ได้พัฒนาขึ้นมาในปี ค.ศ. 1968 ซึ่งมีค่าพารามิเตอร์ 2 ตัวคือ ค่าความยาก และค่าอำนาจจำแนก มีสมการดังนี้

$$P_g(\theta) = \frac{e^{Da_g(\theta - b_g)}}{1 + e^{Da_g(\theta - b_g)}} ; g = 1, 2, \dots, N$$

- เมื่อ $P_g(\theta)$ คือ ความน่าจะเป็นในการตอบข้อสอบของผู้สอบด้วยความสามารถที่จะตอบข้อสอบข้อที่ g ถูก
- a_g คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ g
- b_g คือ ค่าความยากของข้อสอบข้อที่ g
- D คือ Scaling factor มีค่า 1.702
- e คือ ค่าคงที่ค่าเท่ากับ 2.7182

แผนภาพที่ 1 แสดงความหมายของค่า Item parameter (2 พารามิเตอร์) ของข้อสอบ



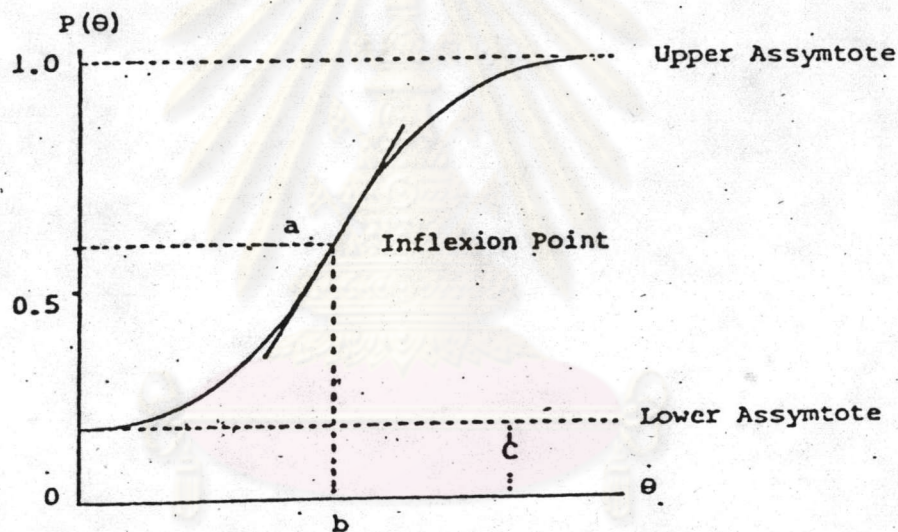
2. รูปแบบที่ใช้พารามิเตอร์ 3 ตัว (Three parameters Logistic Model)

รูปแบบ 3 พารามิเตอร์สามารถเขียนได้จากรูปแบบ 2 พารามิเตอร์โดยเพิ่มพารามิเตอร์
ตัวที่ 3 คือ ค่าการเกา ซึ่งมีสูตรการคำนวณ ดังนี้

$$P_g(\theta) = c_g + (1 - c_g) \frac{e^{Da_g(\theta - b_g)}}{1 + e^{Da_g(\theta - b_g)}} ; g = 1, 2, \dots, N$$

เมื่อ c_g คือ ค่าการเกาของข้อสอบข้อที่ g

แผนภาพที่ 2 แสดงความหมายของค่า Item parameter (3 พารามิเตอร์) ของข้อสอบ

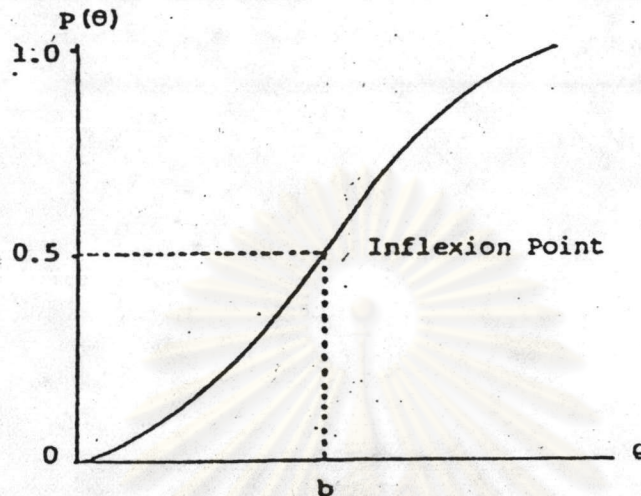


3. รูปแบบที่ใช้พารามิเตอร์ 1 ตัว (One-parameter Logistic Model)

รูปแบบนี้ ราสซ (Rasch) ได้พัฒนาทฤษฎีการทดสอบของข้อสอบโดยที่ฟังก์ชันของโมเดลนั้น
สามารถอธิบายได้ด้วยพารามิเตอร์ของข้อสอบเพียงตัวเดียวคือ ค่าความยาก ซึ่งตรงกับรูปแบบ
ที่ใช้พารามิเตอร์ 1 ตัว ที่ เบิร์นบอม (Birnbaum) ได้พัฒนาขึ้น สามารถเขียนสมการได้
ดังนี้

$$P_g(\theta) = \frac{e^{Da_g(\theta - b_g)}}{1 + e^{Da_g(\theta - b_g)}} ; g = 1, 2, \dots, N$$

แผนภาพที่ 3 แสดงความหมายของค่า Item parameter (1 พารามิเตอร์) ของข้อสอบ หรือรูปแบบของรายช



รูปแบบของรายช เป็นรูปแบบที่มีความซับซ้อนน้อยที่สุด มีพารามิเตอร์แสดง ลักษณะของข้อสอบเพียงพารามิเตอร์เดียว คือ พารามิเตอร์ความยาก

ข้อตกลงเบื้องต้นในการใช้รูปแบบของรายช

ในการนำรูปแบบของรายชมาใช้ มีข้อตกลงเบื้องต้นดังต่อไปนี้ (Hambleton, 1977 : 77 - 78)

1. แบบสอบทั้งฉบับต้องวัดคุณลักษณะเดียวกัน (Unidimension Latent Space) หรือความสามารถเดียวกัน (Single ability) กล่าวคือข้อสอบจะต้องเป็นเอกพันธ์กัน หรือวัดในเนื้อหาเดียวกัน ในการที่จะตัดสินใจว่าข้อสอบมีลักษณะดังกล่าวหรือไม่ นั้น อาจจะใช้วิธีวิเคราะห์ทั่วประกอบ

2. ข้อสอบจะต้องมีความเป็นอิสระของค่าแทน (Local independent)

ใน 2 ลักษณะคือ

2.1 มีความเป็นอิสระทางสถิติ (Statistically independent)

กล่าวคือข้อสอบแต่ละข้อเป็นอิสระไม่เกี่ยวข้องกัน แต่ละข้อวัดความสามารถไม่ซ้ำกันเลย ทั้งนี้ค่าทศของแต่ละข้อของแต่ละคนเป็นอิสระกัน แต่เมื่อรวมกันแล้วจะวัดคุณลักษณะเดียวกัน เท่านั้น

2.2 มีความเป็นอิสระของตำแหน่ง (Uncorrelated independent)

กล่าวคือ ข้อสอบแต่ละข้อจะปรากฏอยู่ในตำแหน่งใดของข้อสอบก็ได้ ไม่มีผลต่อการสอบ

3. คะแนนเป็นแบบทอยลูกเต๋า 1 และทอยดิกให้ 0
4. ความเร็วในการทำข้อสอบ ไม่มีผลต่อโอกาสในการทอยลูกเต๋า
5. โอกาสในการทอยลูกเต๋านั้นขึ้นอยู่กับความสัมพัทธ์ของความสามารถของผู้สอบกับความยากของข้อสอบเท่านั้น

ความใกล้เคียงของรูปแบบของราสช์

ใ้มีการศึกษารูปแบบของราสช์ในรูปแบบต่าง ๆ เพื่อศึกษาความแกร่ง ปรากฏว่า รูปแบบของราสช์ มีข้อใกล้เคียงกว่ารูปแบบอื่น ๆ ดังนี้

1. ถ้าข้อมูลเหมาะสมกับรูปแบบของราสช์แล้ว อาจใช้กลุ่มตัวอย่างเพียง 100 คน ก็เพียงพอในการวิเคราะห์ค่าพารามิเตอร์ (Wright, 1977 : 219)
2. การกระจายความสามารถของกลุ่มตัวอย่าง ไม่จำเป็นต้องเป็นโค้งปกติ และไม่จำเป็นต้องมาจากการสุ่มก็ได้ (Wright, 1979 : 20)
3. ข้อสอบที่ใช้ไม่จำเป็นจะต้องเป็นแบบทอยลูกเต๋า 1 และทอยดิกได้ 0 ก็ได้
4. ข้อสอบแต่ละข้อ ไม่จำเป็นต้องมีค่าอำนาจจำแนกเท่ากันก็ได้ และไม่ต้องการค่าถึงเรื่องการแตกก็ได้ (Wright and Panchapakesan, 1969 : 25)
5. รูปแบบของราสช์มีความซับซ้อนน้อยที่สุด เมื่อเทียบกับรูปแบบอื่น ๆ ของทฤษฎีการทดสอบข้อสอบ จึงมีการนำมาประยุกต์ใช้ในด้านต่าง ๆ มากกว่ารูปแบบอื่น ๆ
6. สามารถใช้กับการสอบทั่ว ๆ ไปได้ การรวมคะแนนก็สามารถรวมได้เลย เพราะการกำหนดน้ำหนักไม่ซับซ้อนเหมือนรูปแบบอื่น ๆ (Wright, 1977 : 102)

การประมาณค่าพารามิเตอร์ตามรูปแบบของราสช์

การประมาณค่าพารามิเตอร์ในรูปแบบของราสช์ เพื่อวิเคราะห์ข้อสอบ สามารถประมาณค่าได้หลายวิธี คือ (Wright afterword in Rasch 1980 : 188 - 193)

1. The LOG Method เป็นวิธีการประมาณค่าพารามิเตอร์ในรูปของความสำเร็จ ซึ่ง เป็นวิธีการที่เหมาะสมกับแบบสอยที่มีอำนาจจำแนกเท่ากับ 1 และใช้จำนวนในการประมาณค่าพารามิเตอร์เป็นจำนวนมาก

2. The PAIR Method เป็นวิธีการประมาณค่าพารามิเตอร์ โดยใช้ข้อสอบจับคู่กัน เช่น 1 - 2, 1 - 3, 1 - 4, เป็นต้น เอาจำนวนคนที่พยายามตอบ 2 ข้อ แยกทำถูกเพียงข้อเดียวมาวิเคราะห์เพื่อหา Sample free คำนึง จึงเหมาะที่จะใช้ประมาณค่าแบบสอยที่มีจำนวนข้อไม่มาก

3. The FCON Method เป็นวิธีประมาณค่าพารามิเตอร์ โดยใช้ข้อสอบทุกข้อและทุกคนมาวิเคราะห์ เหมาะสำหรับแบบสอยที่มีข้อสอบน้อยกว่า 30 ข้อ ถ้ามักกว่าจะทำให้การประมาณค่ามีความคลาดเคลื่อน แต่ภายหลังสามารถใช้กับแบบสอยยาว 60-70 ข้อได้

4. The UCON Method เป็นวิธีประมาณค่าพารามิเตอร์ด้วยวิธี Unconditional Maximumlikelihood Procedure ใช้กับแบบสอยที่มีการกระจายของความสามารถและความยากของแบบสอยเป็นการกระจายแบบปกติ (Normal distribution) และใช้กับแบบสอยที่มีตั้งแต่ 25 ข้อขึ้นไป โดยมากจะใช้กับแบบสอยที่มี 1,000 ข้อขึ้นไป

5. The PROX Method เป็นวิธีประมาณค่าพารามิเตอร์ด้วยวิธีของ Cohen's Approximation ใช้กับแบบสอยที่มีการกระจายของความสามารถของคน และความยากของแบบสอยเป็นลักษณะเบ้ หรือมีแนวโน้มเป็นปกติ โดยมาก PROX และ UCON จะให้ผลเหมือนกัน จะต่างกันที่ค่าความคลาดเคลื่อนมาตรฐานเพียง เล็กน้อย

6. The UFORM Method ใช้กับแบบสอยที่มีการกระจายความยากของข้อสอบเป็นแบบ Uniform สำหรับวิธี PROX สามารถประมาณค่าพารามิเตอร์ด้วยการใช้สูตรคำนวณหรือประมาณค่าโดยใช้เครื่องคอมพิวเตอร์ โดยโปรแกรมสำเร็จรูป BICAL หรืออาจใช้โปรแกรม LOGIST (Wood and Lord, 1976, Wingersky and Lord, 1976) โดยกำหนดเงื่อนไขให้ค่าอำนาจจำแนกเท่ากับ 1 และค่าการคาดเป็น 0

การนำรูปแบบของราสซ์ไปประยุกต์ใช้ในด้านต่าง ๆ

ไรท์ (Wright, 1980) ได้รวบรวมเกี่ยวกับประโยชน์ของการนำรูปแบบของราสซ์ไปประยุกต์ใช้ในหลาย ๆ ด้าน ดังต่อไปนี้ (อวยพร วิทยาลัยกาญจน, 2526 : 21)

1. นำมาวิเคราะห์ข้อสอบ (Item Analysis) เพื่อแก้ปัญหาข้อสอบของการวัดตามแนวเดิม
2. สร้างคลังข้อสอบ (Item Bank) เนื่องจากรูปแบบของราสซ์สามารถนำมาวิเคราะห์ข้อสอบที่ให้ค่าพารามิเตอร์ มีความคงที่ ฉะนั้นจึงสามารถสร้างไว้เพื่อเลือกไปใช้ข้อสอบได้ตามต้องการ
3. สร้างแบบสอบที่ดีที่สุด (Best Test Design) ผลจากการวิเคราะห์สามารถนำไปใช้ออกข้อสอบที่มีคุณลักษณะต่าง ๆ ได้ตามต้องการ
4. นำมาใช้กับการสอบแบบเจาะจง (Tailor Testing) การทดลองแบบมุ่งที่จะเลือกข้อสอบที่มีระดับความยากเหมาะสมกับระดับความสามารถของผู้สอบแต่ละคน การสอบครั้งหนึ่ง ๆ ผู้สอบแต่ละคนไม่จำเป็นต้องตอบข้อสอบเหมือนกันทุกข้อทั้งฉบับ แต่ขึ้นอยู่กับผลการตอบข้อสอบในแบบสอบของแต่ละคน
5. ตรวจสอบความลำเอียงของข้อสอบ (Item Bias) การตรวจสอบความลำเอียง อาจจะเปรียบเทียบในกลุ่มผู้สอบในหลาย ๆ ด้าน เช่น ในเมือง-นอกเมือง ผู้ชาย-ผู้สาว แล้วพิจารณาความลำเอียงของข้อสอบที่เกิดขึ้น
6. วินิจฉัยความสามารถของผู้สอบ (Individual Diagnosis) ในกรณีที่ Item Characteristic Curve : ICC ของข้อสอบไม่เหมาะสมกับโครงสร้างแบบ แสดงว่าอาจจะมีส่วนบางอย่างผิดปกติในตัวผู้สอบ

นอกเหนือจากการนำเอารูปแบบของราสซ์ มาประยุกต์ใช้ในด้านต่าง ๆ ดังที่กล่าวไปแล้ว ก็ยังมีผู้สนใจในการนำเอาไปประยุกต์ใช้ในแง่อื่น ๆ ที่เกี่ยวข้องกับการวัดผลทางการศึกษาอยู่ตลอดเวลาอันเป็นการพัฒนาปรับเปลี่ยนให้เกิดประโยชน์มากที่สุดต่อไปในอนาคต



งานวิจัยที่เกี่ยวข้องกับรูปแบบของราสช์

รูปแบบของราสช์ ได้รับการพัฒนาขึ้น ตั้งแต่ปี ค.ศ. 1960 โดย ราสช์ (Rasch) ซึ่งเป็นชาวเดนมาร์ก จากช่วง 20 ปีที่ผ่านมา จึงได้มีการนำเอารูปแบบของราสช์มาประยุกต์ใช้ หรือทดสอบในประเด็นต่าง ๆ มากมายพอสมควร ผู้วิจัยจึงขอเสนอผลงานวิจัยที่เกี่ยวข้อง ซึ่งเป็นผลงานในต่างประเทศ และในประเทศไทยที่น่าสนใจต่อไปนี้

ไวท์ลีย์ และ เดวิส (Whitely and dewis, 1974) วิเคราะห์แบบสอบอุปมาอุปไมยทางภาษา จำนวน 66 ข้อ โดยใช้รูปแบบของราสช์ พบว่าข้อสอบไม่เหมาะสมกับรูปแบบประมาณ 30 - 40 เปอร์เซ็นต์

ทินส์เลย์ และ เดวิส (Tinley and dawis, 1975) ได้ศึกษาความคงที่ของค่าความยากของข้อสอบที่วิเคราะห์ด้วยรูปแบบของราสช์ โดยใช้แบบสอบอุปมาอุปไมยที่วัดใน 4 ด้านคือ ด้านการใช้คำ ด้านตัวเลข ด้านรูปภาพ และด้านสัญลักษณ์ ซึ่งจำนวนข้อสอบมี 66, 60, 50 และ 40 ข้อ ตามลำดับ นำแบบสอบไปสอบกับกลุ่มตัวอย่างและนำผลการสอบมาวิเคราะห์หาค่าความยากโดยสอบ 10 ครั้ง แล้วนำค่าความยากของแต่ละครั้งมาหาค่าความสัมพันธ์ ผลปรากฏว่าในแบบสอบทั้ง 4 ฉบับนี้ ถ้าจำนวนข้อสอบมากกว่า 30 ข้อขึ้นไป ค่าความยากจะมีความสัมพันธ์กันสูงมาก และถ้าข้อสอบที่มีโอกาสของการตอบถูกมีค่าต่ำ และข้อสอบที่ไม่เหมาะสมกับรูปแบบของราสช์ จะทำให้ค่าความยากมีความสัมพันธ์กันสูงขึ้นอีก

เรนซ์ และ บาร์ชอว์ (Rantz and barshaw, 1975) ได้ศึกษาการใช้ขนาดของกลุ่มตัวอย่างที่จะนำมาวิเคราะห์ด้วยรูปแบบของราสช์ พบว่าขนาดของกลุ่มตัวอย่างระหว่าง 500 - 1,000 คน จะทำให้การประมาณค่าพารามิเตอร์ของรูปแบบของราสช์มีความคงที่เพิ่มขึ้นอย่างรวดเร็ว และจะค่อย ๆ เพิ่มขึ้นเมื่อกลุ่มตัวอย่างที่ใช้ประมาณค่าพารามิเตอร์มีจำนวน 2,000 คน และ 4,000 คน

ฟอร์ซีส และคณะ (Forsyth, et al. 1981) ได้ศึกษาความคงที่และความยากของข้อสอบโดยใช้แบบสอบ Iowa test of Educational Development โดยแบ่งออกเป็น 4 ฉบับย่อย ซึ่งมีจำนวนข้อสอบ 81, 54, 73 และ 60 ข้อตามลำดับ แล้วจัดแบบสอบแต่ละฉบับออกเป็น 3 ส่วนโดยส่วนที่ 1 ง่ายกว่าส่วนที่ 2 และส่วนที่ 2 ง่ายกว่าส่วนที่ 3

จากนั้นคัดแปลงทั้ง 3 ส่วนให้เป็น 2 ระดับ คือ ระดับ 1 และ ระดับ 2 ในระดับที่ 1 รวมเอาส่วนที่ 1 และส่วนที่ 2 เข้าด้วยกัน ระดับที่ 2 รวมเอาส่วนที่ 2 และ 3 เข้าด้วยกัน แล้วนำแบบสอบระดับที่ 1 ไปสอบกับนักเรียนเกรด 9, 10 จำนวน 927 คน ระดับที่ 2 นำไปสอบกับนักเรียน เกรด 10, 11 จำนวน 899 คน และเกรด 12 จำนวน 650 คน สุกท้ายนักเรียนทุกคนต้องสอบในส่วนที่ 2 แล้วนำผลการสอบไปวิเคราะห์ด้วยรูปแบบของราสส์ หากค่าความยากของข้อสอบแต่ละข้อ แล้วนำค่าความยากของแบบสอบจากนักเรียน เกรด 9 ไปหาความสัมพันธ์กับนักเรียนเกรดอื่น ๆ พบว่า ค่าความยากมีความสัมพันธ์กันสูงมาก แสดงว่าการเรียงลำดับความยากของข้อสอบในแต่ละฉบับ ต่าง ๆ กันจะให้ผลไม่ต่างกัน และค่าความยากของข้อสอบมีความคงที่สูง

สมิท (Smith, 1982:3574-A) ได้ศึกษาเกี่ยวกับสิ่งรบกวนในการวัดด้วยรูปแบบของราสส์ โดยเปรียบเทียบความเหมาะสมกับตัวประมาณค่า 2 แบบคือ BIWEIGHT และ AMPJACK พบว่าวิธีของราสส์ให้ค่าที่เหมาะสมกว่าตัวประมาณค่าทั้ง 2 ตัว และจากการวิเคราะห์ความเหมาะสมของผู้สอบพบว่า ค่ารบกวนในการวัด (ϵ) ส่งผลต่อผู้สอบอยู่ในช่วงระหว่าง 18 - 34 เปอร์เซ็นต์ นอกจากนี้ยังพบว่าเพศไม่มีผลต่อความลำเอียงของข้อสอบ (Item Bias) จากการคำนวณพบว่าทั้งเพศชายและหญิง ประมาณ 40 เปอร์เซ็นต์ มีความสามารถอยู่ในระดับสูง

แลช (Lash, 1983:1429-A) ได้ศึกษาเกี่ยวกับการใช้รูปแบบของราสส์ใช้แบบสอบที่วัดหลาย ๆ ด้าน (Multidimensional Item) โดยแปลงรูปให้เป็นการวัดเพียงมิติเดียว (Unidimensional Item) หลาย ๆ กลุ่มวิธีการที่ใช้ประมาณค่ามันใช้คะแนนจริงจากการสอบ จากการศึกษาพบว่าค่าพารามิเตอร์ที่ใช้ของการวัดแต่ละกลุ่มมีการกระจายที่ไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบ หรือแบบสอบ และจากการเปรียบเทียบความแตกต่างของพารามิเตอร์ของข้อสอบระหว่างการวัดแต่ละด้าน พบว่าระดับความสามารถของกลุ่มมีความเป็นอิสระ แต่ไม่พบในการกระจายของข้อสอบ

อวยพร วิบูลย์กาญจน์ (2526) ได้เปรียบเทียบผลการวิเคราะห์แบบสอ
 อุปมาอุปไมยกับยาคคลาสิคอลล กับ ราชสโมเคล โดยใช้กลุ่มตัวอย่าง 1,884 คน ซึ่งเป็นนักเรียน
 ที่จบชั้นประถมศึกษาปีที่ 6 ในปีการศึกษา 2525 ที่สมัครสอบคัดเลือกเข้าเรียนต่อในชั้น
 มัธยมศึกษาปีที่ 1 เครื่องมือที่ใช้เป็นแบบสอมาตรฐาน วิชาวิทยาศาสตร์ และแบบสอ
 อุปมาอุปไมยของสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ
 ประสานมิตร วิเคราะห์หอสอด้วยคอมพิวเตอร์โปรแกรมไมคาล และโปรแกรมวิเคราะห์
 หอสอ (Item Analysis) ทดสอบความแตกต่างระหว่างจำนวนหอสอ โดยใช้ค่าสถิติซี
 (Z - test) แล้วพบว่าวิบูลย์กาญจน์ได้ 34 ข้อ และราชสโมเคลได้ 26 ข้อ หากค่าความ
 เทียงโดยใช้สูตร K-R 20 ได้ค่าความเทียง .85 และแบบราชสโมเคลได้ .739 หากความทรง
 เียงสภาพแบบคลาสิคอลลได้ .49 และแบบราชสโมเคลได้ .46

สมพร บุญอิม (2529) ได้ศึกษาความคงที่ของค่าพารามิเตอร์ ความยาก
 ในการวิเคราะห์หอสอด้วยราชสโมเคล กับประชากรที่มีความสามารถต่างกัน 3 กลุ่ม
 โดยมีความโง่และความแปรปรวนเท่ากัน การแจกแจงของแต่ละกลุ่ม คือ 1) แจกแจง
 เบ้บวกลสำหรับกลุ่มที่มีความสามารถต่ำ 2) แจกแจงแบบปกติ สำหรับกลุ่มที่มีความสามารถ
 ปานกลาง 3) แจกแจงแบบเบ้บวกล สำหรับกลุ่มที่มีความสามารถสูง โดยใช้เทคนิคมอนติคาร์โล
 ซิมูเลชัน จำลองการทดลองในเครื่องคอมพิวเตอร์ ศึกษาทั้งกลุ่มตัวอย่างกลุ่มละ 500
 หอสอ จำนวน 30 ข้อ ทดลองซ้ำในแต่ละสถานการณ์ 100 ครั้ง พบว่า

1. พารามิเตอร์ความยากของหอสอมีความคงที่ในหอสอที่มีความยาก ไม่สูงหรือ
 ต่ำมาก กลุ่มที่มีความสามารถต่ำ และกลุ่มที่มีความสามารถสูง พารามิเตอร์ความยากมี
 ความคงที่ 19 ข้อ กลุ่มที่มีความสามารถปานกลางมีความคงที่ 22 ข้อ

2. การกระจายของค่าความคลาดเคลื่อนมาตรฐานของค่าความยากของหอสอ
 เมื่อผู้สอมีความสามารถต่ำ และผู้สอที่มีความสามารถสูง มีแนวโน้มไม่เป็นสมมาตร เมื่อ
 ผู้สอมีความสามารถปานกลาง มีการกระจายเป็นสมมาตร

3. ค่าความคลาดเคลื่อนมาตรฐานของค่าความยากของหอสอทั้ง 3 กลุ่ม ไม่แตกต่าง
 กันอย่างมีนัยสำคัญ

บัญชี แสนทวิ (2530) ไทประยุกต์เอารูปแบบของราสสมาใช้ในการออกแบบ
 โคงแสดงสารสนเทศ (Information) ของแบบสอผลสัมฤทธิ์ เพื่อใช้ทวระกัความ
 สามารถ เครื่องมือที่ใช้เป็นแบบสอผลสัมฤทธิ์ วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6
 จำนวน 2 ฌบ ฌบละ 60 ฌบ นำไปสอกับกลุ่มทวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6
 จำนวน 1,429 คน นำผลการสอมาวิเคราะห์ทวโปรแกรมไบคาล (BICAL))
 เพื่อหาค่าพารามิเตอร์ของซอสอ และพารามิเตอร์ความสามารถ แล้วทวการเทียบมาทรา
 (Eguating) ในแบบสอทั้ง 2 ฌบ จักกลุ่มความสามารถออกเป็น 3 กลุ่ม คือ
 กลุ่มความสามารถสูง ปานกลาง ท่ำ แล้วออกแบบ แบบสอให้เหมาะสมกับระกัความสามารถ
 ทั้ง 3 กลุ่ม ซึ่งมีผลการวิจัย ดังนี้

1. เมื่อใช้แบบสอที่วัดความสามารถต่าง ๆ กัน ทั้งกลุ่มทวค่าความ
 คลาคเคลื่อนมาทราฐานในการประมาณค่าความสามารถ แตกต่างกันอย่างไม่มีนัยสำคัญ
 ที่ระกั .05
2. เมื่อใช้แบบสอที่วัดความสามารถต่าง ๆ กัน 3 ฌบ กับกลุ่มที่มีความสามารถ
 ท่ำ และสูง ทวว่า ค่าความคลาคเคลื่อนมาทราฐานในการประมาณค่าความสามารถ แตกต่าง
 กันอย่างมีนัยสำคัญที่ระกั .01 และไม่มีวแตกต่างกันอย่างมีนัยสำคัญที่ระกั .05 เมื่อ
 กลุ่มผู้สอที่มีความสามารถปานกลางทอสอกับแบบสอทั้ง 3 ฌบ ดังกล่าวนี้
3. เมื่อใช้แบบสอที่วัดความสามารถต่าง ๆ กัน 3 ฌบ กับกลุ่มที่มีความสามารถท่ำ
 ทวว่า ค่าความเที่ยงของแบบสอที่ใดแตกต่างกันอย่างมีนัยสำคัญที่ระกั .01 ระหว่างแบบสอ
 ที่เหมาะสมกับความสามารถท่ำและปานกลาง และแบบสอที่เหมาะสมกับความสามารถท่ำและ
 สูง ส่วนค่าความเที่ยงระหว่างแบบสอที่เหมาะสมกับความสามารถปานกลางและสูง แตกต่าง
 กันอย่างไม่มีนัยสำคัญที่ระกั .05
4. เมื่อใช้แบบสอ ที่วัดความสามารถต่าง ๆ กัน 3 ฌบ กับกลุ่มที่มีความ
 สามารถปานกลาง ทวว่าค่าความเที่ยงของแบบสอแตกต่างกันอย่างมีนัยสำคัญที่ระกั
 .01 ระหว่างแบบสอที่เหมาะสมกับความสามารถปานกลางและสูง ส่วนค่าความเที่ยง
 ระหว่างแบบสอที่เหมาะสมกับความสามารถปานกลางและท่ำ และแบบสอที่เหมาะสม
 กับความสามารถสูง และท่ำแตกต่างกันอย่างมีนัยสำคัญที่ระกั .05

5. เมื่อใช้แบบสอบที่วัดความสามารถต่าง ๆ กัน 3 ฉบับ กับกลุ่มที่มีความสามารถสูง พบว่าค่าความเที่ยงของแบบสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ระหว่างแบบสอบที่เหมาะสมกับความสามารถต่ำและปานกลาง และแบบสอบที่เหมาะสมกับความสามารถต่ำและสูง แตกต่างกันอย่างไม่มีนัยสำคัญที่ระดับ .05



ศูนย์วิทยพัชร์พยากร
จุฬาลงกรณ์มหาวิทยาลัย