

วรรณคดีที่เกี่ยวข้อง

ผู้วิจัยขอนำเสนอวรรณคดีที่เกี่ยวข้องกับการวิจัยนี้ โดยแบ่งออกเป็น 3 ตอน

ดังนี้

ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับคุณภาพของแบบสอบถาม : ระบุถึงความยาก
อ่านง่าย และความเที่ยงของแบบสอบถาม

ระบุถึงความยากของข้อกระทง

1. นิยามระบุถึงความยากของข้อกระทง

ระบุถึงความยากของข้อกระทง หมายถึง สัดส่วนของผู้เข้าสอบที่ตอบ
ข้อกระทงนั้นได้ถูกต้อง หรือคือความน่าจะเป็นของการตอบข้อกระทงนั้นได้ถูกต้อง โดย
นิยาม ระบุถึงความยากของข้อกระทงจะมีพิสัยอยู่ระหว่าง 0 ถึง 1.00 กรณีที่
ระบุถึงความยากของข้อกระทงมีค่าเข้าใกล้ 0 หรือมีค่าเท่ากับ 0 ก็แสดงว่า
ข้อกระทงนั้นยาก มีผู้เข้าสอบน้อยคนหรือไม่มีใครตอบถูกเลยในข้อนั้น และเมื่อระบุ
ความยากของข้อกระทงมีค่าเข้าใกล้ 1.00 หรือมีค่าเท่ากับ 1.00 แสดงว่า
ข้อกระทงนั้นง่าย มีผู้เข้าสอบจำนวนมากตอบถูกหรือตอบถูกทุกคน

2. ความสำคัญของระบุถึงความยากต่อการสร้างแบบสอบถาม

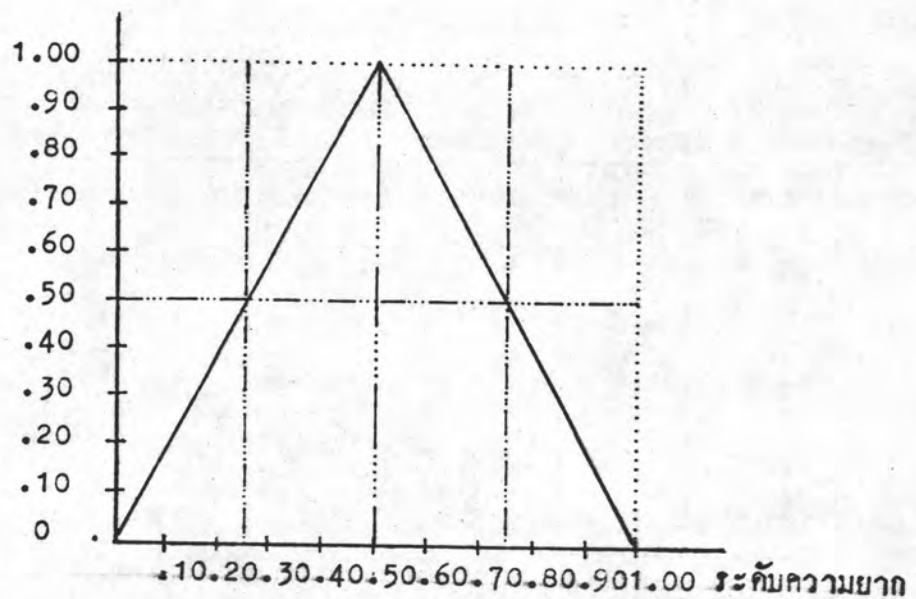
ระบุถึงความยากของข้อกระทงมีความสำคัญต่อการสร้างแบบสอบถาม
เหตุผลที่สำคัญ 2 ประการคือ ประการแรก ระบุถึงความยากของข้อกระทงมีอิทธิพลต่อ
ลักษณะการแจกแจงของคะแนนสอบ ถ้าระบุถึงความยากของข้อกระทงเบี่ยงเบนห่างจาก
ค่า .50 ไปทางใดทางหนึ่งก็ทำให้การแจกแจงของคะแนนมีลักษณะเบ้ ซึ่งโดย
ทั่วไปต้องการให้คะแนนมีการแจกแจงเป็นโค้งปกติ และมีค่าความเบี่ยงเบนมาตรฐาน
มากที่สุดเท่าที่จะทำได้ การทำให้ระบุถึงความยากเฉลี่ยของแบบสอบถามเท่ากับ .50 จะ
ทำให้มีการจำแนกผู้เข้าสอบได้มากที่สุด ประการที่สอง ระบุถึงความยากของข้อกระทง
มีอิทธิพลต่อค่าความเที่ยงของแบบสอบถาม เนื่องจากสูตรสัมประสิทธิ์อัลฟาของครอนบาค

และสุกรกุเคอร์ - ริชาร์ดสันที่ 20 จะให้ค่าความเที่ยงสูงสุดก็ต่อเมื่อ สหสัมพันธ์ระหว่าง
ข้อระงมมีค่าสูง ทั่วยเหตุนี้ค่าเฉลี่ยของระดับความยากที่เขียนเบนออกห่างจาก .50
ไปทางใดทางหนึ่งจะทำให้สหสัมพันธ์ระหว่างข้อระงมมีค่าต่ำ ดังนั้นถ้าต้องการให้ค่า
ความเที่ยงสูง ก็ควรทำให้ระดับความยากเฉลี่ยของแบบสอบมีค่าใกล้ .50
(อนันต์ ศรีโสภา 2525 : 166)

3. ความสัมพันธ์ระหว่างระดับความยากกับอำนาจจำแนกของข้อระงม
และระดับความยากที่เหมาะสม

ฮอบกินส์และแสตนเลย์ (Hopkins and Stanley 1981 : 270)
ได้แสดงความสัมพันธ์ระหว่างระดับความยากกับศักยภาพของอำนาจจำแนกของข้อระงม
(Potential Item Discrimination) ทั่วยกราฟดังนี้

ค่าอำนาจจำแนกสูงสุด



จากกราฟ เมื่อผู้เข้าสอบจำนวนครึ่งหนึ่งสามารถตอบข้อระงมได้ถูกต้อง
ระดับความยากของข้อระงมย่อมมีค่าเท่ากับ .50 แล้วเป็นไปได้ที่ค่าอำนาจจำแนก
ของข้อระงมจะมีค่าเท่ากับ 1.00 ข้อระงมที่ง่ายเกินไปหรือยากเกินไปจะให้ข้อมูล
ของการจำแนกความแตกต่างของผู้เข้าสอบได้น้อย ข้อระงมที่มีระดับความยากอยู่ใน
ช่วงกลาง (.25 ถึง .75) ทุกข้อระงมนั้นอาจจะมีค่าอำนาจจำแนกสูง

เมื่อพิจารณาลักษณะของเนื้อหาของแบบสอบถาม แบบสอบถามซึ่งมีเนื้อหาเป็นเอกพจน์ กล่าวคือมีค่าเฉลี่ยของสหสัมพันธ์แบบไบซีเรียลเท่ากับ .60 ถึง .70 แบบสอบถามเช่นนี้ควรประกอบด้วยข้อกระทงที่มีพิสัยของระดับความยากอยู่ระหว่าง .84 ถึง .16 และแบบสอบถามที่มีค่าเฉลี่ยของสหสัมพันธ์แบบไบซีเรียลเท่ากับ .30 ถึง .40 ซึ่งถือว่าเป็นแบบสอบถามที่มีเนื้อหาค่อนข้างเป็นวิวิธพันธ์ ข้อกระทงในแบบสอบถามเช่นนี้ควรมีพิสัยของระดับความยากอยู่ระหว่าง .60 ถึง .40 (Henrysson 1971 : 153) ในเรื่องนี้ อัลเลนและเยน (Allen and Yen 1979 : 121) ได้เสนอว่าโดยทั่วไปถ้าต้องการให้เกิดการจำแนกสูงสุดระหว่างผู้เข้าสอบที่มีระดับความสามารถแตกต่างกัน ข้อกระทงในแบบสอบถามนั้นควรมีพิสัยของระดับความยากระหว่าง .30 ถึง .70

เมื่อมีการนำเอาการตอบถูกโดยบังเอิญซึ่งผู้เข้าสอบตอบอย่างเผลอสุ่มเข้ามาเกี่ยวข้องกับขั้วแล้ว ระดับความยากที่เหมาะสมของข้อกระทงเลือกตอบควรมีค่าสูงค่า $\left[\frac{1}{2} (\text{สัดส่วนของการตอบถูกโดยการเผลอสุ่ม} + 1.00) \right]$ อีกประมาณ .10 (Lord 1953 อ้างจาก อนันต์ ศรีโสภณ 2525 : 192) ดังนั้นข้อกระทง 5 ตัวเลือก คาระดับความยากที่เหมาะสมควรมีค่าประมาณ .70

ตารางที่ 1 ระดับความยากที่เหมาะสมของข้อกระทงเลือกตอบตามข้อเสนอของลอร์ด

จำนวนตัวเลือก	ระดับความยากที่เหมาะสม
0 (คอบสั้น , ความเรียง)	.50
2	.85
3	.77
4	.74
5	.69

โดยหลักการทางทฤษฎีแล้ว ข้อกระทงที่ดีควรมีระดับความยากเป็น .50 เพราะจะเป็นข้อกระทงที่มีโอกาสจะทำให้เกิดการจำแนกสูงสุด และทำให้ได้ความเที่ยง

สูงสุด แต่อย่างไรก็ตามในทางปฏิบัติทางการศึกษายังคงนิยมที่จะจัดให้แบบสอบมีทั้งข้อกระทงที่ง่ายหรือค่อนข้างง่ายจำนวนหนึ่ง และในขณะที่เดียวกันก็มีข้อกระทงที่ยากหรือค่อนข้างยากอีกจำนวนหนึ่ง แต่ข้อกระทงที่มีค่าระดับความยากใกล้เคียง .50 อยู่เป็นจำนวนมาก ข้อกระทงที่ง่ายจะช่วยให้ผู้เข้าสอบที่อ่อนมีกำลังใจทำข้อกระทงที่ตนพอทำได้ และข้อกระทงที่ยากจะเป็นการสร้างแรงจูงใจให้กับผู้เข้าสอบที่มีความสามารถสูง (เยาวดี วิบูลย์ศรี 2526 : 173)

4. ระดับความยากของข้อกระทงที่มีค่าคะแนนได้หลายค่า

การวิจัยนี้ค้นคว้าวิธีการสรรหาเซทย่อยของค่าคอมมาไซ้ เป็นผลให้คะแนนของข้อกระทงมีค่าได้หลายค่า ผู้วิจัยจึงค้นคว้าสูตรการหาระดับความยากของข้อกระทงซึ่ง ไวท์เนย์และซาเบอร์ส (Whitney and Sabers 1971 : 88 - 92 cited in Scannell and Tracy 1975 : 223) ได้เสนอไว้มาใช้ สูตรเป็นดังนี้

$$P = \frac{\sum X_H + \sum X_L - N(X_{min})}{N(X_{max} - X_{min})} \dots\dots\dots (5)$$

เมื่อ	p	แทน	ระดับความยากของข้อกระทง
	$\sum X_H$	แทน	ผลบวกของคะแนนในกลุ่มสูงของข้อกระทงนั้น
	$\sum X_L$	แทน	ผลบวกของคะแนนในกลุ่มต่ำของข้อกระทงนั้น
	X_{min}	แทน	คะแนนต่ำสุดที่เป็นไปได้ในข้อกระทงนั้น
	X_{max}	แทน	คะแนนสูงสุดที่เป็นไปได้ในข้อกระทงนั้น
	N	แทน	จำนวนผู้เข้าสอบในกลุ่มสูงรวมกับ กลุ่มต่ำ

5. องค์ประกอบที่มีอิทธิพลต่อค่าระดับความยากของแบบสอบ

แคมเบล (Cambell 1961 : 899 - 913) ได้กล่าวถึงองค์ประกอบที่มีอิทธิพลต่อค่าระดับความยากของแบบสอบไว้ดังนี้

5.1 องค์ประกอบภายใน (Intrinsic Factors) ได้แก่

5.1.1 เนื้อหาของข้อกระทง อันได้แก่ ความซับซ้อน ความเป็นนามธรรม และความแปลกใหม่

5.1.2 ลักษณะโครงสร้างของข้อกระทง เป็นวิธีการแสดง
ออกของเนื้อหาที่ข้อกระทงนั้นวัด

5.2 องค์ประกอบภายนอก (Extrinsic Factors) ได้แก่

5.2.1 ความไม่คุ้นเคยข้อเนื้อหา เป็นการถามในเรื่องที่
ผู้เข้าสอบไม่เคยพบมาก่อน

5.2.2 ความสัมพันธ์กันของข้อกระทง เป็นความสัมพันธ์ของ
ข้อกระทงนั้นกับข้อกระทงอื่นในแบบสอบเดียวกัน

5.2.3 ทิวแปรค่านบุคลิกภาพ ได้แก่ สภาพร่างกาย
ลักษณะนิสัย และความตั้งใจของผู้เข้าสอบ

นอกจากนี้ ธรรมชาติของเนื้อหา พฤติกรรมที่ก่อการวัด และสิ่งที่ทำให้เกิด
ความซับซ้อนอื่น ๆ เช่น ภาษาที่ใช้ในข้อกระทง รูปแบบของคำถาม และคำชี้แจงต่าง ๆ
ยังเป็นองค์ประกอบที่มีอิทธิพลต่อระดับความยากของแบบสอบด้วย

อำนาจจำแนกของข้อกระทง

1. นิยามของอำนาจจำแนกของข้อกระทง

อำนาจจำแนกของข้อกระทง หมายถึง ความสามารถของข้อกระทงที่จะ
จำแนกหรือแยกผู้เข้าสอบออกได้ตามระดับความสามารถ เช่น จำแนกคนเก่ง คนอ่อน
คนที่มีความถนัด ออกจากกันได้ (เยาวดี วิบูลย์ศรี 2526 : 174)

2. การเลือกใช้เกณฑ์ในการวิเคราะห์ข้อกระทง

ตามหลักการแล้ว การที่จะสรุปว่าข้อกระทงที่มีค่าอำนาจจำแนกสูงเป็น
ข้อกระทงที่ดีนั้น จำเป็นต้องตั้งข้อตกลงว่าแบบสอบนั้นประกอบด้วยข้อกระทงที่ดี กังนั้น
เกณฑ์ที่ใช้ในการคัดเลือกกลุ่มที่มีผลสัมฤทธิ์สูง และกลุ่มที่มีผลสัมฤทธิ์ต่ำเพื่อใช้วิเคราะห์
ข้อกระทงของแบบสอบวัดผลสัมฤทธิ์ทางการเรียนจึงมักใช้เกณฑ์ภายใน (Internal
Criterion) แทนที่จะใช้เกณฑ์ภายนอก (External Criterion) เกณฑ์ภายใน
กึ่งกลางนี้ก็คือ คะแนนรวมของแบบสอบทั้งฉบับนั่นเอง โดยถือค่าผู้เข้าสอบที่ได้คะแนน
รวมสูง จักอยู่ในกลุ่มสูง และผู้เข้าสอบที่ได้คะแนนรวมต่ำเป็นกลุ่มต่ำ การใช้คะแนนรวม
เป็นเกณฑ์ในการคัดเลือกกลุ่มนี้มีข้อดี 2 ประการคือ ประการแรก ได้แก่ความสอดคล้อง

กับซึ่คความรู้ ความสามารถและทักษะของผู้สร้างแบบสอบ คะแนนที่ได้จากแบบสอบที่เขาสร้างจะสอดคล้องกับสิ่งที่เขาต้องการวัดได้ก็ว่าการวัดด้วยเครื่องมือหรือแบบสอบอื่น ๆ ประการที่สอง ได้แก่ความสะดวก คะแนนรวมของแบบสอบให้ความสะดวกต่อการนำไปใช้ประโยชน์มากก่า (Ebel 1972 : 392)

3. การประเมินคุณภาพของข้อระหง โดยอาศัยค่าอำนาจจำแนกของข้อระหง

การพิจารณาคุณภาพของข้อระหงนั้น ข้อระหงที่ดีควรเป็นข้อระหงที่มีคนในกลุ่มสูงทำถูกเป็นส่วนมาก ขณะที่คนในกลุ่มต่ำทำถูกเป็นส่วนน้อย ซึ่งคะแนนของข้อระหงนั้นจะมีความสัมพันธ์เชิงบวกกับคะแนนรวมของแบบสอบ อำนาจจำแนกมีค่าได้ตั้งแต่ 1.00 ถึง - 1.00 ถ้าผู้เข้าสอบในกลุ่มสูงตอบถูกมากกว่าผู้เข้าสอบในกลุ่มต่ำ ข้อระหงนั้นจะมีค่าอำนาจจำแนกเป็นบวก และยิ่งค่าอำนาจจำแนกเข้าใกล้ 1.00 ข้อระหงนั้นจะสามารถจำแนกได้มากขึ้น ข้อระหงที่มีค่าอำนาจจำแนกเป็นศูนย์จะไม่สามารถจำแนกนักเรียนได้ เกิดจากการที่ผู้เข้าสอบกลุ่มสูงกับกลุ่มต่ำตอบข้อระหงนั้นได้ถูกเท่ากัน ส่วนข้อระหงที่มีค่าอำนาจจำแนกเป็นลบ แสดงว่าผู้เข้าสอบกลุ่มต่ำตอบข้อระหงนั้นได้ถูกต้องมากกว่าผู้เข้าสอบกลุ่มสูง จึงเป็นข้อระหงที่ไม่สามารถวัดคุณลักษณะหรือความสามารถตามที่ต้องการได้ การประเมินคุณภาพของข้อระหงที่ครูสร้างขึ้น โดยอาศัยค่าอำนาจจำแนกนั้น อีเบล (Ebel 1972 : 399) ได้ให้หลักเกณฑ์ไว้ดังนี้

ตารางที่ 2 เกณฑ์ในการประเมินคุณภาพของข้อระหงโดยใช้ค่าอำนาจจำแนก

ค่าอำนาจจำแนก	การประเมิน
.40 ขึ้นไป	ดีมาก
.30 ถึง .39	ดีพอสมควร อาจต้องปรับปรุงบ้าง
.20 ถึง .29	พอใช้ได้ แต่ต้องปรับปรุง
ต่ำกว่า .19	ไม่ดี ควรตัดทิ้งหรือแก้ไขใหม่



4. การหาค่าอำนาจจำแนกของข้อระหวง

วิธีการหาค่าอำนาจจำแนกของข้อระหวงสามารถกระทำไ้หลายวิธี ซึ่งการที่จะเลือกใช้วิธีใดขึ้นอยู่กับ ลักษณะข้อมูล ข้อตกลงเบื้องต้นของแต่ละวิธี การประหยัคแรงงานและเวลา ตลอดจนค่าคอกที่คองการ เช่น ถ้าคองการทราบว่ข้อระหวงสามารถวัดคุณลักษณะที่ไ้เป็นเกณฑ์ไ้เพียงใด ก็ไ้ใช้สหสัมพันธ์แบบไบซีเรียลหรือแบบเทคราคอริก ถ้าคองการทราบว่ข้อระหวงมีอำนาจเชิงทำนายอย่างไร ก็ควรไ้ใช้สหสัมพันธ์แบบพอยท์ไบซีเรียล เป็นต้น (Guilford and Fruchter 1978 : 462) ในที่นี้จะไ้กล่าวว่วิธีการหาค่าอำนาจจำแนกมากล่่าวบางวิธีดังนี้

- 4.1 วิธีใช้กลุ่มปลายสุด ("D" the Extreme Group Method)
- 4.2 วิธีหาค่าสหสัมพันธ์ (The Correlation Approach)
 - 4.2.1 สหสัมพันธ์แบบเทคราคอริก
 - 4.2.2 สหสัมพันธ์แบบพอยท์ไบซีเรียล
 - 4.2.3 สหสัมพันธ์แบบไบซีเรียล
 - 4.2.4 สหสัมพันธ์แบบพี
- 4.3 วิธีใช้ตารางสำเร็จรูป
- 4.4 วิธีทดสอบค่าที

ในที่นี้ขอ กล่าวเฉพาะการหาค่าอำนาจจำแนกโดยการหาค่าสหสัมพันธ์แบบไบซีเรียล ซึ่งเป็นวิธีหาค่าอำนาจจำแนกในการวิจัยนี้ การหาค่าอำนาจจำแนกวิธีนี้เป็นการหาค่าประมาณของสหสัมพันธ์แบบเพียร์สัน จากตัวแปรทั้งสองซึ่งไ้แก่ คะแนนรายข้อ และคะแนนรวม มีสูตรดังนี้ (Guilford and Fruchter 1978 : 305)

$$r_{bis} = \frac{(\bar{X}_p - \bar{X}_q) \cdot pq}{s_t \cdot y} \dots\dots\dots (6)$$

เมื่อ	r_{bis}	แทน	ค่าอำนาจจำแนกของข้อระหวง
	\bar{X}_p	แทน	ค่าเฉลี่ยของคะแนนรวมของผู้ตอบถูก
	\bar{X}_q	แทน	ค่าเฉลี่ยของคะแนนรวมของผู้ตอบผิด
	s_t	แทน	ส่วนเบี่ยงเบนมาตรฐานของคะแนนรวมของผู้เข้าสอบทุกคน



- p แทน สัดส่วนของผู้เข้าสอบที่ตอบข้อกระทงนั้นถูก
q แทน สัดส่วนของผู้เข้าสอบที่ตอบข้อกระทงนั้นผิด
y แทน ออกิเนศของโค้งปกติซึ่งแบ่งพื้นที่ใต้โค้งออกเป็น p และ q

การใช้สหสัมพันธ์แบบไมซีเรียลมีข้อตกลงเบื้องต้นว่า ตัวแปรทั้งสองซึ่งได้แก่ คะแนนของข้อกระทง และคะแนนรวม มีลักษณะการแจกแจงแบบต่อเนื่องและเป็นปกติ โดยที่โค้งจักรจะทำให้ตัวแปรหนึ่งในที่มีโค้งแล้ว คะแนนของข้อกระทง มีค่าแปรผันได้เพียง 2 ค่า เช่น ใต้ - ตก หรือ ผ่าน - ไม่ผ่าน เป็นต้น

5. องค์ประกอบที่มีอิทธิพลต่อค่าอำนาจจำแนกของแบบสอบ

องค์ประกอบที่มีอิทธิพลต่อค่าอำนาจจำแนกของแบบสอบ มีดังนี้

5.1 คุณภาพของตัวเลือก และจำนวนตัวเลือก ข้อกระทงที่สร้างตัวเลือก ใต้ดี มีความใกล้เคียงกับตัวเลือก จะทำให้ค่าอำนาจจำแนกของข้อกระทงสูงขึ้น (Ebel 1965 : 359) เนื่องจากข้อกระทงเช่นนี้ ผู้เข้าสอบที่มีความรู้จริงเท่านั้นที่สามารถตอบข้อกระทงได้ถูกต้อง แต่ผู้เข้าสอบที่ไม่มีความรู้จริงจะเลือกตัวเลือกที่ใกล้เคียงกับคำตอบ ถูกกระจายออกไปทุกตัวเลือก และโดยเหตุผลทางทฤษฎีแล้ว แบบสอบที่มีจำนวนตัวเลือก มากก็ย่อมมีความสามารถในการวัดมากขึ้น แต่จากการศึกษาเชิงประจักษ์ คอสติน (Costin 1970 : 353 - 358) ได้พบว่า แบบสอบเลือกตอบ 3 ตัวเลือกทำให้ค่าอำนาจจำแนกของแบบสอบสูงสุด

5.2 ความยากของข้อกระทง การจำแนกผลสัมฤทธิ์ระหว่างนักเรียน ที่เข้าสอบขึ้นอยู่กับระดับความยากของข้อกระทง กล่าวคือ ข้อกระทงที่มีระดับความยาก .50 จะมีโอกาสทำให้ค่าอำนาจจำแนกของข้อกระทงสูงสุด เมื่อนักเรียนในกลุ่มสูง ทุกคนตอบได้ถูกต้อง ในขณะที่นักเรียนในกลุ่มต่ำทุกคนตอบผิด ซึ่งหากข้อกระทงมีค่าระดับ ความยากสูงกว่าหรือต่ำกว่านี้ ค่าอำนาจจำแนกของข้อกระทงจะต่ำลง (Gronlund 1976 : 270) หมายความว่า ข้อกระทงที่ยากหรือง่ายเกินไป จะเป็นข้อกระทงที่มี ค่าอำนาจจำแนกต่ำลงนั่นเอง

5.3 ลักษณะของเนื้อหาวิชาในแบบสอบ เมื่อเนื้อหาของแบบสอบมีลักษณะเป็นเอกพันธ์ คือวัดในลักษณะนิสัยหรือความสามารถเชิงว ในการจัดกลุ่มสูง - กลุ่มต่ำ ก็สามารถใช้คะแนนรวมเป็นเกณฑ์ได้ การได้ค่าอำนาจจำแนกเป็นบวกก็หมายความว่า ข้อทรงนั้นสามารถวัดในสิ่งที่ต้องการวัดได้ แต่ถ้าแบบสอบไม่มีความเป็นเอกพันธ์ แม้ว่าจะได้ค่าอำนาจจำแนกเป็นบวกก็มีได้หมายความว่าข้อทรงนั้นวัดในสิ่งที่ต้องการวัด (อนันต์ ศรีโสภา 2525 : 190)

ความเที่ยงของแบบสอบ

1. นิยามความเที่ยง

ความเที่ยงเป็นสิ่งสำคัญในการประเมินคุณภาพของแบบสอบหรือมาตรการวัดทางจิตวิทยา ถ้าแบบสอบนั้นสามารถวัดคุณลักษณะหรือพฤติกรรมได้อย่างคงที่แน่นอนแล้ว ผู้เข้าสอบคนเดียวกัน จะได้คะแนนเหมือนเดิมในการทดสอบซ้ำด้วยแบบสอบชุดเดิม สำหรับนิยามของความเที่ยงที่ชัดเจนนั้นนักวัดผลหลายท่านได้ให้นิยามไว้ เป็นต้นว่า เมห์เรนส์และเลห์แมน (Mehrens and Lehmann 1984 : 267) ได้ให้ความหมายว่าคือ รัศมีของความคงที่ระหว่างการวัดสิ่งเดียวกันสองครั้ง อนาตาสี (Anatasi 1976 : 103) กล่าวว่าความเที่ยง หมายถึง ความคงที่ของคะแนนที่ได้จากการสอบของผู้เข้าสอบกลุ่มเดียวกันสองครั้ง ด้วยแบบสอบเดิมในเวลาต่างกัน หรือด้วยแบบสอบต่างชุดที่เทียบเท่ากัน หรือภายใต้สภาพการสอบที่แตกต่างกัน ส่วนอีเบล (Ebel 1972 : 410) ได้ให้ความหมายเชิงปฏิบัติการว่า สัมประสิทธิ์ความเที่ยงของคะแนนชุดหนึ่ง ที่ได้จากผู้เข้าสอบกลุ่มเดียวกัน คือ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนชุดนั้น กับคะแนนแบบสอบเทียบเท่าอีกชุดหนึ่ง ซึ่งได้มาอย่างอิสระจากผู้เข้าสอบกลุ่มเดียวกัน

2. ทฤษฎีของความเที่ยง

แนวความคิดง่าย ๆ ที่เกี่ยวกับความเที่ยงคือ คะแนนที่ได้จากการวัดใด ๆ สามารถจำแนกออกเป็น 2 ส่วนที่สำคัญคือ คะแนนจริง (True Score) กับคะแนนความคลาดเคลื่อน (Error Score) ซึ่งสามารถเขียนเป็นสมการพื้นฐานได้ดังนี้ (Mehrens and Lehmann 1984 : 269)

$$X = T + E \dots\dots\dots (7)$$

เมื่อ X แทน คะแนนผลการสอบ
 T แทน คะแนนจริง
 E แทน คะแนนความคลาดเคลื่อน

คะแนนจริง หมายถึง คะแนนที่ผู้เข้าสอบทำได้ จากการวัดด้วยเครื่องมือที่มีคุณภาพสมบูรณ์ ปราศจากความคลาดเคลื่อนหรือหมายถึง คะแนนเฉลี่ยของบุคคลที่ได้จากการทำแบบสอบถาม หรือแบบสอบถามจำนวนหลาย ๆ ครั้ง (Brown 1976 : 50) จากทฤษฎีความเที่ยง ความคลาดเคลื่อนที่เกิดขึ้นอย่างมีระบบจะไม่มีผลกระทบต่อความเที่ยง ดังนั้นคะแนนจริงจะถูกพิจารณาว่าเป็นส่วนที่คงที่ หรือเป็นส่วนของคะแนนผลการสอบที่ไม่เปลี่ยนแปลง

คะแนนความคลาดเคลื่อน หมายถึง ค่าความผิดพลาดที่เกิดขึ้นจากการวัด เป็นความคลาดเคลื่อนที่เกิดขึ้นอย่างสุ่ม (Random Error) ตามทฤษฎีความคลาดเคลื่อนที่เกิดขึ้นอย่างสุ่มนี้อาจเป็นไปได้ทั้งทางบวกและทางลบ ซึ่งจะหักล้างกันหมดไป คะแนนความคลาดเคลื่อนจึงมีลักษณะการแจกแจงเป็นปกติ ซึ่งทำให้มีค่าเฉลี่ยเท่ากับศูนย์ และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับหนึ่ง เมื่อคะแนนความคลาดเคลื่อนเกิดขึ้นอย่างสุ่มจึงไม่มีความสัมพันธ์กับคะแนนจริง ทำให้ได้ความสัมพันธ์สหสัมพันธ์ระหว่างคะแนนความคลาดเคลื่อน กับคะแนนจริงมีค่าเท่ากับศูนย์ จากข้อตกลงดังกล่าวสามารถเขียนเป็นสมการเชิงความแปรปรวนได้ดังนี้

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2 + 2 \text{cov}(T , E) \dots\dots (8)$$

เพราะว่า $r_{T,E} = \frac{\text{cov} (T , E)}{\sigma_T \cdot \sigma_E} \dots\dots\dots (9)$

และ
จะได้ว่า

$$r_{T,E} = 0 \quad (\text{T กับ E ไม่สัมพันธ์กัน})$$

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2 \quad \dots\dots\dots (10)$$

เมื่อ

σ_X^2 แทน ความแปรปรวนของคะแนนผลการสอบ

σ_T^2 แทน ความแปรปรวนของคะแนนจริง

σ_E^2 แทน ความแปรปรวนของคะแนนความคลาดเคลื่อน

เนื่องจากความเที่ยง คือ อัตราส่วนระหว่างความแปรปรวนของคะแนนจริง กับความแปรปรวนของคะแนนผลการสอบ จึงเขียนสมการ ได้ดังนี้

$$r_{tt} = \frac{\sigma_T^2}{\sigma_X^2} \quad \dots\dots\dots (11)$$

เมื่อ r_{tt} แทน ความเที่ยงของแบบสอบ

จากสมการที่ (10) และ (11) จะได้ว่า

$$r_{tt} = 1 - \frac{\sigma_E^2}{\sigma_X^2} \quad \dots\dots\dots (12)$$

จากสมการที่ (12) ความเที่ยงจะมีค่าสูงขึ้นเมื่อความแปรปรวนของคะแนนความคลาดเคลื่อนลดลง โดยกำหนดให้คะแนนผลการสอบมีค่าความแปรปรวนคงที่ หรือมีค่าเพิ่มขึ้น (Guilford and Fruchter 1978 : 438)

3. วิธีประมาณค่าความเที่ยงของแบบสอบ

การประมาณค่าความเที่ยงของแบบสอบกระทำได้หลายวิธี แต่ละวิธีทำให้เกิดความคลาดเคลื่อนต่างกัน การที่จะเลือกวิธีใดขึ้นอยู่กับจุดมุ่งหมายและลักษณะของคะแนนที่ได้จากการสอบ วิธีการทั่วไปที่ใช้รายงานค่าความเที่ยงในคู่มือแบบสอบทั้งหลาย มี 5 แบบดังนี้ (Mehrens and Lehmann 1984 : 271 - 278)

- 3.1 แบบวัดความคงที่ (Measures of Stability)
- 3.2 แบบวัดความเทียบเท่า (Measures of Equivalence)
- 3.3 แบบวัดความเทียบเท่าและความคงที่ (Measure of Equivalence and Stability)
- 3.4 แบบวัดความคงที่ภายใน (Measures of Internal Consistency)
 - 3.4.1 การแบ่งครึ่งแบบส้อม (Split - Half)
 - 3.4.2 การประมาณค่าด้วยสูตรคูเคอร์ - ริชาร์ดสัน (Kuder - Richardson Estimates)
 - 3.4.3 สัมประสิทธิ์อัลฟาของครอนบาค (Cronbach 's Alpha Coefficient)
 - 3.4.4 วิธีการวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt 's Analysis of Variance Procedure)
- 3.5 ความเที่ยงของผู้ให้คะแนน (Score Reliability)

ในที่นี้ขอกล่าวถึงเฉพาะวิธีการประมาณค่าความเที่ยงแบบการวัดความคงที่ภายในโดยวิธีหาค่าสัมประสิทธิ์อัลฟาของครอนบาค ซึ่งเป็นวิธีที่ใช้ในการประมาณค่าความเที่ยงในการวิจัยนี้

วิธีประมาณค่าความเที่ยงของแบบส้อมโดยการหาค่าสัมประสิทธิ์อัลฟาของครอนบาคนี้พัฒนามาจากสูตร คูเคอร์ - ริชาร์ดสันที่ 20 โดยครอนบาคในปี ค.ศ. 1951 เป็นการหาค่าเฉลี่ยของสหสัมพันธ์แบบแบ่งครึ่งแบบส้อมหลาย ๆ วิธีเท่าที่จะเป็นไปได้ ในทางปฏิบัติไม่จำเป็นที่จะต้องคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์จากทุกกรณีของการแบ่งครึ่งแบบส้อม เพราะสัมประสิทธิ์อัลฟาสามารถคำนวณหาได้จาก ความแปรปรวนของคะแนนสอบในแต่ละข้อกระทง และความแปรปรวนของคะแนนรวม ดังสมการ

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^2 \sigma_i^2}{\sigma_x^2} \right) \dots \dots \dots (13)$$

เมื่อ	α	แทน	ค่าความเที่ยงของแบบสอบ
	n	แทน	จำนวนข้อกระทงในแบบสอบ
	$\sum_{i=1}^2$	แทน	ผลรวมของความแปรปรวนของคะแนนในแต่ละข้อกระทง
	σ_x^2	แทน	ความแปรปรวนของคะแนนรวม

สัมประสิทธิ์อัลฟาจะถือว่าข้อกระทงแต่ละข้อเป็นแบบสอบเล็ก ๆ (Mini Test) ถ้าข้อกระทงในแบบสอบทุกข้อวัดในสิ่งเดียวกัน ค่าสหสัมพันธ์ระหว่างข้อกระทงจะสูง ค่า α จะใกล้ 1.00 แต่ถ้าข้อกระทงเหล่านั้นวัดในสิ่งที่ต่างกัน ค่าสหสัมพันธ์ระหว่างข้อกระทงบางส่วนจะต่ำ ดังนั้นค่า α จะลดลงจาก 1.00 วิธีการประมาณค่านี้เป็นที่นิยม เพราะสามารถใช้กับแบบสอบต่าง ๆ ที่มีค่าคะแนนเฉพาะ (Unique Value) ในแต่ละข้อต่างกันไปได้ (Mehrens and Lehmann 1984 : 277 ; เขาวที วิมลยศรี 2526 : 64)

4. องค์ประกอบที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบ

องค์ประกอบหลายอย่างที่สามารถทำให้ค่าความเที่ยงของแบบสอบสูงขึ้นหรือต่ำลงได้ องค์ประกอบเหล่านั้นได้แก่ (Mehrens and Lehmann 1984 : 278 - 281 ; Gronlund 1976 : 117 - 123)

4.1 ความยาวของแบบสอบ ในทางทฤษฎีการเพิ่มจำนวนข้อกระทงที่มีคุณภาพเท่าเทียมกับข้อกระทงที่มีอยู่เดิมในแบบสอบ ประกอบกับการควบคุมคุณภาพอื่น ๆ ที่จะมีผลต่อค่าความเที่ยงของแบบสอบให้คงที่แล้ว แบบสอบที่มีจำนวนข้อกระทงมากกว่าย่อมมีค่าความเที่ยงสูงกว่า แบบสอบที่มีจำนวนข้อกระทงน้อยกว่า ทั้งนี้ด้วยเหตุผล 2 ประการคือ ประการแรก แบบสอบที่ยาวกว่าจะสามารถวัดมวลความรู้ หรือคุณลักษณะได้คล้ายคลึงกับประชากรของความรู้มากขึ้น ประการที่สอง เมื่อแบบสอบยาวขึ้นทำให้โอกาสของการเตาถูกโดยสุ่มลดลง ความสัมพันธ์ระหว่างความยาวของแบบสอบกับค่าความเที่ยงคำนวณได้โดยอาศัยสูตรสเปียร์แมน - บราวน์

4.2 เวลาในการทำแบบสอบ แบบสอบที่กำหนดเวลาไม่เหมาะสมจะมีผลต่อค่าความเที่ยงของแบบสอบ ถ้ากำหนดเวลาให้น้อยเกินไป ผู้เข้าสอบทำไม่ทัน เช่นแบบสอบความเร็ว (Speed Test) ผู้เข้าสอบอาจใช้วิธีเดาสุ่มจากข้อที่ทำไม่ทัน

ซึ่งจะทำให้การกระจายของคะแนนแตกต่างจากแบบสอบที่ให้เวลาพอเหมาะ แต่ถ้าให้เวลาสอบมากเกินไปจะทำให้คนเรียนเก่งและคนเรียนอ่อน ทำคะแนนในแบบสอบนั้นได้ใกล้เคียงกัน การกระจายของคะแนนจะมีน้อยก็มิใช่ผลต่อค่าความเที่ยงของแบบสอบเช่นกัน ดังนั้นจึงไม่ควรใช้วิธีการวัดความคงที่ภายในกับการประมาณค่าความเที่ยงของแบบสอบ ความเร็ว ควรหาควยวิธีวัดความคงที่ หรือวิธีวัดควยแบบสอบที่เท่าเทียมกัน เพราะค่าความเที่ยงที่ประมาณโดยการวัดความคงที่ภายในสูงกว่าที่ควรจะเป็น

4.3 ความเป็นเอกพันธ์ของกลุ่มผู้เข้าสอบ กลุ่มผู้เข้าสอบที่มีความเป็นเอกพันธ์ จะทำให้ค่าความเที่ยงของแบบสอบต่ำกว่ากลุ่มผู้เข้าสอบที่มีความสามารถเป็นวิวิธพันธ์ ควยเหตุผลที่ว่ากลุ่มผู้เข้าสอบที่มีความสามารถเป็นวิวิธพันธ์มีพิสัยของคะแนนผลการสอบมากกว่า ทำให้คะแนนสอบมีการกระจายมากกว่า จากสมการที่ (12) เมื่อมีการเปลี่ยนกลุ่มผู้เข้าสอบ σ_E^2 ซึ่งเป็นความแปรปรวนของคะแนนความคลาดเคลื่อนควรมีค่าคงเดิม แต่ σ_X^2 จะเพิ่มมากขึ้นสำหรับกลุ่มวิวิธพันธ์ จึงทำให้ค่าความเที่ยงของแบบสอบมีค่าสูงขึ้น

4.4 ความยากของข้อกระทง ทั้งที่กล่าวมาแล้วในเรื่องของระดับความยากของข้อกระทง แบบสอบที่ยากมากผู้เข้าสอบทำผิดเกือบทุกคน และแบบสอบที่ง่ายมากผู้เข้าสอบทำได้เกือบทุกคน ทำให้ความแปรปรวนของคะแนนผลการสอบมีค่าน้อย มีผลให้ค่าความเที่ยงของแบบสอบต่ำลง แบบสอบที่มีความยากเฉลี่ยใกล้เคียง .50 จะมีโอกาสทำให้ค่าความเที่ยงสูงสุด

4.5 ความเป็นปรนัยของการให้คะแนน แบบสอบที่มีความเป็นปรนัยในการให้คะแนนสูง ความคลาดเคลื่อนที่เกิดขึ้นมีน้อยเป็นผลให้คะแนนที่ได้ใกล้เคียงกับคะแนนจริงมากขึ้น ความเที่ยงของแบบสอบจะสูงขึ้นควย

4.6 วิธีประมาณค่าความเที่ยง วิธีประมาณค่าความเที่ยงของแบบสอบสามารถกระทำได้หลายวิธี แต่ละวิธีจะทำให้เกิดความคลาดเคลื่อนต่างกัน และเหมาะกับจุดมุ่งหมายเฉพาะในการตีความหมายของค่าความเที่ยง ค่าความเที่ยงที่ประมาณโดยวิธีแบ่งครึ่งแบบสอบมักให้ค่าสูงสุด ค่าความเที่ยงที่ประมาณจากการวัดความคงที่มักให้ค่าปานกลาง และค่าความเที่ยงที่ประมาณควยสุทร คูเคอร์ - ริชาร์ดสันมักให้ค่า

ความเที่ยงค่า

นอกจากนี้ ความคลาดเคลื่อนในการประมาณค่าความเที่ยงอาจเกิดขึ้นจากการบริหารแบบสอบ เช่น การจับเวลาสอบ การให้คำชี้แจง ความชัดเจนของคำสั่ง การป้องกันการทุจริต สภาพแวดล้อมของห้องสอบ หรือองค์ประกอบที่เกี่ยวกับตัวผู้เข้าสอบเอง เช่น ความคุ้นเคยกับเนื้อหาที่สอบ ความหลักแหลมในการสอบ (Testwiseness) ความวิตกกังวลการเจ็บป่วย การเตา หรือองค์ประกอบที่เกี่ยวกับลักษณะของข้อกระทงในแบบสอบ เช่น ความคาบเกี่ยวกันของแต่ละข้อกระทงในแบบสอบนั้น เป็นต้น

ตอนที่ 2 ผลงานวิจัยเกี่ยวกับการกำหนดจำนวนตัวเลือกที่เหมาะสมของข้อกระทงแบบเลือกตอบ

แบบสอบเลือกตอบได้รับการพิสูจน์ว่าเป็นแบบสอบที่ปฏิบัติได้อย่างกว้างขวาง จึงเป็นที่ยอมรับและใช้กันทั่ว ๆ ไป (Anatasi 1976 : 162 ; Ebel 1972 : 187) โดยเฉพาะในการสอบความถนัดหรือการสอบวัดผลสัมฤทธิ์ทางการเรียน แต่ข้อสงสัยประการหนึ่งที่มีขึ้นเกี่ยวกับการสร้างและการพัฒนาแบบสอบเลือกตอบก็คือ ข้อกระทงชนิดเลือกตอบควรมีจำนวนตัวเลือกเท่าไรจึงจะทำให้แบบสอบมีคุณภาพสูงสุด ในเรื่องนี้ นักวัดผลและประเมินผลการศึกษาหลายท่าน (Ahmann and Glock 1971 ; Micheels and Karnes 1950 ; Thorndike and Hagen 1961 ; Torgerson and Adams 1954) ได้สนับสนุนให้ใช้ข้อกระทง 4 หรือ 5 ตัวเลือก ขณะที่เมห์เร็นส์ และเลห์แมน (Mehrens and Lehmann 1984 : 162) ได้เสนอว่า โดยทั่วไปควรใช้ข้อกระทงชนิด 3 ถึง 5 ตัวเลือก

นอกจากข้อเสนอแนะดังกล่าวแล้ว นักวัดผลการศึกษาจำนวนหนึ่งได้ทำการศึกษาเชิงทฤษฎี และทำการวิจัยเชิงประจักษ์ เพื่อหาค่าคอมเกี่ยวกับตัวเลือกที่เหมาะสมของข้อกระทงแบบเลือกตอบที่ทำให้แบบสอบมีคุณภาพสูงสุด โดย ซิมเมอร์แมนและฮัมพรีย์ (Zimmerman and Humphreys 1953 : 460 - 461) ได้ศึกษาถึงผลของการกำจัดตัวเลือกที่มีค่าอำนาจจำแนกต่ำสุดออกจากข้อกระทง ที่มีค่าความเที่ยงของแบบสอบพบว่า แบบสอบเลือกตอบ 2 ตัวเลือก แบบสอบเลือกตอบ 3 ตัวเลือก แบบสอบเลือกตอบ 4 ตัวเลือก ซึ่งเกิดจากการกำจัดตัวเลือกที่มีค่าอำนาจจำแนกต่ำสุดออกจากข้อกระทง ชนิด



5 ตัวเลือกที่ละตัวความลำบาก ทำให้ความเที่ยงของแบบสอบมีค่าสูงขึ้น ทั้งยังทำให้เวลาในการตอบแต่ละข้อกระทงน้อยลงอีกด้วย ผลการศึกษานี้สอดคล้องกับการศึกษาของ วิลเลียมและอีเบล (Williams and Ebel 1957 : 63 - 65) ซึ่งได้นำแบบสอบค่าศัพท์ ของสถาบันพัฒนาแบบสอบทางการศึกษาแห่งรัฐไอโอวา มาสร้างเป็นแบบสอบชุด 3 ตัวเลือกและชุด 2 ตัวเลือก โดยการคัดตัวดวงที่มีค่าอำนาจจำแนกค่าสูงออกจากข้อกระทงเดิมซึ่งมี 4 ตัวเลือกที่ละตัวความลำบาก แล้วจึงนำแบบสอบทั้ง 3 ชุดไปสอบกับนักศึกษาจำนวน 230 คน ศึกษากำหนดเวลาของการสอบเท่ากันทั้ง 3 ชุดได้พบว่า แบบสอบชุด 2 ตัวเลือกและชุด 3 ตัวเลือกให้ค่าความเที่ยงสูงกว่าแบบสอบชุด 4 ตัวเลือกอย่างไม่มีนัยสำคัญ

ในปีคริสต์ศักราช 1964 ทเวอส์กี (Tversky 1964 : 386 - 391) ได้ศึกษาหาจำนวนตัวเลือกที่เหมาะสมของข้อกระทง โดยการพิสูจน์ทางคณิตศาสตร์ภายใต้ข้อตกลงเบื้องต้นที่กำหนดให้จำนวนตัวเลือกในแบบสอบชุดต่าง ๆ ที่นำมาเปรียบเทียบกันนั้นมีค่าคงที่ ได้พบว่า เมื่อจำนวนตัวเลือกของข้อกระทงมีค่า $2.718\dots$ หรือเป็น 3 ตัวเลือก (กรณีที่มีค่าเป็นจำนวนเต็ม) จะทำให้แบบสอบนั้นมีคุณภาพสูงสุด คุณภาพของแบบสอบตามนัยที่ ทเวอส์กี กล่าวถึงได้แก่

1. กำลังของแบบสอบ (Power of Test) หมายถึง ความสามารถในการป้องกันการใช้การเดาของข้อกระทงในแบบสอบ มีค่าเท่ากับ $1 - \alpha$ ค่าโอกาสของการตอบถูกโดยการเดาสุ่ม
2. ความจุแห่งการจำแนกของแบบสอบ (Discrimination Capacity of Test) หมายถึง จำนวนรูปแบบเฉพาะของค่าตอบที่เป็นไปได้ทั้งหมด ซึ่งมีค่าเท่ากับ a^n เมื่อ a แทนจำนวนตัวเลือกของข้อกระทงและ n แทนจำนวนข้อกระทงในแบบสอบนั้น
3. ทัศนของความไม่แน่นอน (Uncertainty Index) หมายถึง ข้อมูลที่ได้รับเพิ่มจากการใช้แบบสอบนั้น ซึ่งมีค่าสอดคล้องกับค่าความจุแห่งการจำแนกของแบบสอบนั้น

การศึกษาของ ทเวอรัสกี ได้สร้างรูปแบบใหม่ที่นำเสนอใจสำหรับการศึกษา
ปัญหานี้ โดยเขาได้อาศัยข้อตกลงเบื้องต้นของความเป็นสัดส่วน (Assumption of
Proportionalty) กล่าวคือ กำหนดให้ผลคูณของจำนวนตัวเลือกในข้อกระทงกับ
จำนวนข้อกระทงในแบบสอบมีค่าคงที่ ดังนั้นภายในเวลาสอบที่เท่ากัน แบบสอบที่นำมา
ศึกษาเปรียบเทียบกันจะสามารถบริหารภายในเวลาสอบได้ด้วยการทำให้แบบสอบแต่ละ
ชุดยาวขึ้นหรือสั้นลงด้วยการเปลี่ยนแปลงค่าจำนวนตัวเลือกของข้อกระทงกับจำนวนของ
ข้อกระทงในแบบสอบอย่างเป็นสัดส่วนกัน

เมื่อกำหนดให้	T	แทน	เวลาสอบ (หั่งฉับ)
	n	แทน	จำนวนข้อกระทงในแบบสอบ
	a	แทน	จำนวนตัวเลือกของข้อกระทง ซึ่งกำหนดให้ เท่ากันทุกข้อ
	t	แทน	เวลาที่ใช้อ่านและพิจารณาตัวเลือกแต่ละตัว โดย ถือว่าตัวเลือกแต่ละตัวมีแรงดึงดูดใจต่อผู้เข้าสอบ เท่ากัน ดังนั้นผู้เข้าสอบจะใช้เวลากับแต่ละตัวเลือก เท่า ๆ กันทุกตัวเลือก

จากข้อตกลงเบื้องต้น ทำให้ได้สมการในรูปของเวลาดังนี้

$$T = nat$$

ต่อมา อีเบล (Ebel 1969 : 565 - 570) นำมาเรื่องการเอาเข้ามา
เกี่ยวข้องกับการศึกษาปัญหาเดียวกันนี้ โดยได้ใช้ข้อตกลงเบื้องต้น 5 ประการดังนี้

1. คะแนนเฉลี่ยที่ได้จากแบบสอบเลือกตอบที่ t มีค่าเท่ากับ $\frac{1}{2}$ ของคะแนน
สูงสุดบวกกับคะแนนคาดหวังที่เกิดจากการเอาสุ่ม
2. ค่าเบี่ยงเบนมาตรฐานของคะแนนสอบที่ได้จากแบบสอบเลือกตอบที่มีค่า
เท่ากับ $\frac{1}{6}$ ของผลต่างระหว่างคะแนนสูงสุดที่เป็นไปได้ กับคะแนนคาดหวังที่เกิดจาก
การเอาสุ่ม
3. สัมประสิทธิ์ความเที่ยงหาจากสูตร คูเคอร์ - ริชาร์ดสันที่ 21

4. คะแนนสูงสุดที่เป็นไปได้ของแบบสอบถาม มีค่าเท่ากับ จำนวนข้อกระทง
ในแบบสอบถามนั้น
5. คะแนนค่าคาดหวังที่เกิดจากการคาดสุ่ม มีค่าเท่ากับ ผลหารของจำนวน
ข้อกระทงในแบบสอบถามกับจำนวนตัวเลือกของข้อกระทง
จากข้อตกลงเบื้องต้นนี้ทำให้ได้สมการดังนี้

$$r = \frac{n}{n-1} \cdot \left[1 - \frac{9(a+1)}{n(a-1)} \right] \dots\dots (14)$$

เมื่อ r แทน ความเที่ยงของแบบสอบถาม
 n แทน จำนวนข้อกระทงในแบบสอบถาม
 a แทน จำนวนตัวเลือกของข้อกระทง

ผลการศึกษานี้ อีเบล พบว่า เมื่อแบบสอบถามมีความยาวมากพอและเป็นแบบสอบถาม
ที่ประกอบด้วยข้อกระทงที่มีคุณภาพแล้ว ค่าความเที่ยงจะเพิ่มขึ้นอย่างมาก เมื่อตัวเลือก
ของข้อกระทงเพิ่มจาก 2 เป็น 3 ตัวเลือก และค่าความเที่ยงของแบบสอบถามจะเพิ่มขึ้น
อีกเล็กน้อย เมื่อจำนวนตัวเลือกของข้อกระทงเพิ่มขึ้นเป็น 4 ตัวเลือกและมากกว่า 4
ตัวเลือก การศึกษาของอีเบลนี้ให้ผลแตกต่างจากการศึกษาของ ซิมเมอร์แมนและฮัมฟรี
(Zimmerman and Humphrey 1953 : 460 - 461) ผลนี้อาจเกิดขึ้นจาก
วิธีการที่ใช้ตัดตัวเลือกออกจากข้อกระทงแตกต่างกัน ขณะที่อีเบลใช้วิธีการสุ่มตัดตัวเลือก
ออกจากข้อกระทง และถือว่าแต่ละตัวเลือกมีแรงดึงดูดใจต่อผู้เข้าสอบเท่ากัน ส่วน
ซิมเมอร์แมนและฮัมฟรีใช้วิธีตัดตัวเลือกที่มีค่าอำนาจจำแนกต่ำสุดออกจากข้อกระทง
นอกจากนี้การศึกษาของอีเบลยังได้เอาส่วนของการคาดสุ่มเข้ามาใช้อีกด้วย

ท้ายความต้องการที่จะยืนยันข้อพิสูจน์ทางคณิตศาสตร์ของ ทเวอรัสกี โดยการ
หาหลักฐานเชิงประจักษ์ คอสติน (Costin 1970 : 353 - 358) จึงได้ใช้
แบบสอบถามวัดผลสัมฤทธิ์ทางการเรียนชนิดเลือกตอบ 4 ตัวเลือกในวิชาจิตวิทยาเบื้องต้น
โดยแบ่งเนื้อหาออกเป็น 4 เรื่อง กำหนดให้ข้อกระทงในแต่ละเรื่องมีจำนวนเป็นสัดส่วน
โดยตรงกับจำนวนชั่วโมงเรียนของเรื่องนั้น ๆ หลังจากได้ข้อกระทงในแต่ละเรื่อง

ความที่ทองการแล้วจึงสุ่มข้อกระทงจำนวนครึ่งหนึ่งจากแต่ละเรื่องออกมาสร้างเป็นข้อกระทงชนิด 3 ตัวเลือกด้วยการคัดตัวดวงออกหนึ่งตัวอย่างสุ่ม ทำให้แต่ละเรื่องมีข้อกระทง 4 ตัวเลือกครึ่งหนึ่งและ 3 ตัวเลือกครึ่งหนึ่ง นำแบบสอบที่ได้ไปสอบกับนักศึกษา 270 คนที่เรียนรายวิชานี้ ได้พบว่า ค่าระดับความยากและอำนาจจำแนกของแบบสอบชุด 3 ตัวเลือกมีค่าสูงกว่าแบบสอบชุด 4 ตัวเลือกในทุก ๆ เรื่อง ส่วนความเที่ยงของแบบสอบนั้น แบบสอบชุด 3 ตัวเลือกให้ค่าสูงกว่าแบบสอบชุด 4 ตัวเลือกในทุกเรื่องเช่นกัน ในเรื่องนี้ คอสติน ได้ให้ข้อเสนอเพิ่มเติมกันว่า โดยทั่วไปผู้เข้าสอบจะสามารถทำแบบสอบชุด 3 ตัวเลือกเสร็จเร็วกว่าแบบสอบชุด 4 ตัวเลือก ทั้งนี้ภายในเวลาที่กำหนดให้เท่ากัน ถ้าเลือกใช้แบบสอบชุด 3 ตัวเลือก ก็จะทำให้สามารถเพิ่มข้อกระทงเข้าไปในแบบสอบนั้นได้อีกจำนวนหนึ่ง ทำให้แบบสอบนั้นวัดเนื้อหา หรือพฤติกรรมที่ทองการวัดได้ครอบคลุมกว่า ทั้งยังเป็นการลดความตรากตรำและการสิ้นเปลืองเวลาในการศึกษาตัวดวงที่คิดเพื่อเพิ่มเข้าไปในแต่ละข้อกระทงอีกด้วย

รามอสและสเทิร์น (Ramos and Stern 1973 : 305 - 310)

ทำการวิจัยเชิงประจักษ์เพื่อศึกษาความสอดคล้องของคุณภาพแบบสอบ เมื่อมีการเปลี่ยนแปลงจำนวนตัวเลือกในข้อกระทงชนิดเลือกตอบ เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลเป็นแบบสอบวิชาภาษาฝรั่งเศสจำนวน 2 ฉบับ (F_1 และ F_2) และแบบสอบวิชาภาษาสเปนจำนวน 2 ฉบับ (S_1 และ S_2) แบ่งแบบสอบแต่ละฉบับออกเป็น 2 ตอนโดยทำให้ตอนที่ 1 ของ F_1 และ S_1 ประกอบด้วยข้อกระทง 5 ตัวเลือก ตอนที่ 2 ของ F_1 และ S_1 ประกอบด้วยข้อกระทง 4 ตัวเลือก และทำให้ตอนที่ 1 ของ F_2 และ S_2 ประกอบด้วยข้อกระทง 4 ตัวเลือก (ข้อกระทงเดียวกันกับตอนที่ 1 ของ F_1 และ S_1) ตอนที่ 2 ของ F_2 และ S_2 ประกอบด้วยข้อกระทง 5 ตัวเลือก การคัดตัวดวงออกจากข้อกระทง 5 ตัวเลือกให้เป็นข้อกระทง 4 ตัวเลือกกระทำโดยการคัดตัวดวงที่มีคนเลือกน้อยที่สุดออก แล้วนำแบบสอบทั้งสองฉบับไปสอบกับนักศึกษาจำนวน 2,423 คน ให้แต่ละคนทำแบบสอบฉบับหนึ่งโดยวิธีสุ่มแจกแบบแบ่งชั้น นำกระดาษคำตอบที่ได้มาตรวจให้คะแนนโดยใช้สูตรแก้การเคา

$$X = R - \frac{W}{a}$$

เมื่อ X แทนคะแนนแก้การเคาแล้ว R แทนจำนวนข้อที่ตอบถูก W แทนจำนวนข้อที่ตอบผิด และ a แทนจำนวนตัวเลือกของข้อกระทง

ผลการศึกษาพบว่า คะแนนเฉลี่ยจากแบบสอบถามชุด 5 ตัวเลือก มีค่าสูงกว่าคะแนนเฉลี่ยที่ได้จากแบบสอบถามชุด 4 ตัวเลือกเล็กน้อย ค่าอำนาจจำแนกและระดับความยากของแบบสอบถามชุด 5 ตัวเลือกสูงกว่าแบบสอบถามชุด 4 ตัวเลือกอย่างไม่มีนัยสำคัญในทางสถิติ แต่ความเที่ยงของแบบสอบถามชุด 5 ตัวเลือกสูงกว่าความเที่ยงของแบบสอบถามชุด 4 ตัวเลือกอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ในปีคริสต์ศักราช 1975 กริเออร์ (Grier 1975 : 109 - 112) ได้ทำการสืบสวนปัญหาและพบว่า แบบสอบ 3 ตัวเลือกให้ค่าความเที่ยงสูงสุด และแบบสอบ 2 ตัวเลือกให้ค่าความเที่ยงรองลงมา กริเออร์เข้าถึงข้อสรุปนี้ได้โดยอาศัยการประมาณค่าความเที่ยงด้วยสูตร คูเคอร์ - ริชาร์ดสันที่ 21 ซึ่งกริเออร์ได้ใช้ข้อตกลงเบื้องต้นเช่นเกี่ยวกับการศึกษาของอีเบล (Ebel 1969 : 565 - 570) พร้อมทั้งข้อตกลงเบื้องต้นของความเป็นสัดส่วน คือ กำหนดให้จำนวนตัวเลือกทั้งหมดที่มีอยู่ในแต่ละแบบสอบมีค่าคงที่ ตามแนวความคิดที่ ทเวอ์สกี เสนอไว้ กริเออร์ได้แสดงให้เห็นว่า การศึกษาของ อีเบล จะเป็นจริงก็ต่อเมื่อ แบบสอบที่นำมาศึกษานั้นต้องมีจำนวนข้อกระทงมากกว่า 18 ข้อ นอกจากนั้นเขายังได้ให้ข้อเสนอเพิ่มเติมด้วยว่าข้อสรุปที่ว่า ข้อกระทง 3 ตัวเลือกทำให้ความเที่ยงของแบบสอบมีค่าสูงสุดจะใกล้เคียงกับความเป็นจริงมากขึ้น เมื่อข้อกระทงในแบบสอบนั้นมีค่าตามสั้นมาก จนทำให้เวลาสอบส่วนใหญ่ถูกใช้ไปกับการอ่านและพิจารณาตัวเลือก แต่กริเออร์เห็นว่าโดยการปฏิบัติที่เป็นอยู่ทั่วไป เวลาของการสอบส่วนใหญ่ถูกใช้ไปกับการอ่านและพิจารณาตัวคำถาม

ในปีต่อมา กริเออร์ (Grier 1976 : 91 - 97) ได้ทำการสืบสวนปัญหาอื่น โดยในครั้งนี้ได้พิจารณาให้เวลาสอบของแต่ละข้อกระทงแยกออกเป็น 2 ส่วน คือ ส่วนหนึ่งเป็นเวลาที่ใช้อ่านและพิจารณาตัวคำถาม (t_1) ส่วนที่สองเป็นเวลาที่ใช้อ่านและพิจารณาตัวเลือกแต่ละตัว (t_2) ในการศึกษาครั้งนี้เขาได้ใช้ข้อตกลงเบื้องต้นของความเป็นสัดส่วน คือ กำหนดให้จำนวนตัวเลือกในแบบสอบกับจำนวนข้อกระทงในแบบสอบแปรเปลี่ยนกันอย่างเป็นสัดส่วน และมีค่าคงที่สำหรับแบบสอบที่นำมาเปรียบเทียบกัน ซึ่งสามารถเขียนอยู่ในรูปของสมการเวลาดังนี้



$$T' = nt' + \sum_{i=1}^m a_i t$$

กรีเออร์เชื่อว่าโดยทั่ว ๆ ไปเวลาที่ผู้เข้าสอบใช้ไปกับตัวคำถามมีค่ามากกว่าเวลาที่ใช้ไปกับตัวเลือก ดังนั้นกรีเออร์จึงได้ให้ข้อสรุปว่า สำหรับแบบสอบตัว ๆ ไป ข้อกระทงที่มีตัวเลือกมากกว่า 3 ตัวเลือกให้ค่าความเที่ยงสูงขึ้นตามลำดับ จากแนวคิดในการศึกษาปัญหานี้ กรีเออร์ได้เสนอให้ส่วนของเวลาที่ใช้ไปกับตัวคำถาม เข้ามามีส่วนในการกำหนดขนาดของแบบสอบด้วย จึงทำให้ขนาดของแบบสอบที่เท่ากันตามแนวความคิดของกรีเออร์ต่างไปจากแนวความคิดของ ทเวอร์สกี กล่าวคือเดิม ทเวอร์สกีได้กำหนดให้ผลคูณระหว่างจำนวนข้อกระทงกับจำนวนตัวเลือกของข้อกระทงมีค่าคงที่ เช่นเมื่อกำหนดให้ขนาดของแบบสอบที่นำมาเปรียบเทียบกันเท่ากับ 60 แล้วแบบสอบเลือกตอบ 5 ตัวเลือกจะประกอบด้วย 12 ข้อกระทง ($5 \times 12 = 60$) แบบสอบเลือกตอบ 4 ตัวเลือกจะประกอบด้วย 15 ข้อกระทง ($4 \times 15 = 60$) แบบสอบเลือกตอบ 3 ตัวเลือกจะประกอบด้วย 20 ข้อกระทง ($3 \times 20 = 60$) แต่ขนาดของแบบสอบที่เท่ากันตามแนวคิดของกรีเออร์จะทำให้ผลคูณของจำนวนข้อกระทงกับค่า ($a + m$) ของแบบสอบทุกชุดที่นำมาเปรียบเทียบกันมีค่าเท่ากัน จากที่กล่าวมาแล้วจะเห็นได้ว่า ข้อตกลงเบื้องต้นของความเป็นสัดส่วนตามแนวคิดของกรีเออร์นี้ มีความครอบคลุมต่อการศึกษานี้มากกว่าข้อตกลงเบื้องต้นของความเป็นสัดส่วนตามแนวคิดของทเวอร์สกี กล่าวคือ กรณีที่พิจารณาให้เวลาที่ใช้อ่านและพิจารณาตัวคำถามมีค่าน้อยมากแล้วรูปสมการที่ (3) $T' = nt' + nat$ ก็จะได้เป็น $T = nat$ ซึ่งก็คือรูปสมการที่ทเวอร์สกีเสนอไว้นั่นเอง

ลอร์ด (Lord 1980 : 106 - 113) ได้สรุปวิธีการที่ใช้กับการศึกษาปัญหาจำนวนตัวเลือกของข้อกระทงที่เหมาะสมไว้เป็น 4 วิธีดังนี้

1. วิธีการทางคณิตศาสตร์ (Mathematical Approach) เป็นวิธีที่ทเวอร์สกีได้ใช้ในการศึกษานี้ ดังนั้นแล้วข้างต้น
2. วิธีการของกรีเออร์ (Grier's Approach) เป็นวิธีที่กรีเออร์ใช้ศึกษานี้ในปี ค.ศ. 1975 ดังกล่าวแล้วข้างต้น

3. วิธีการที่ใช้ทฤษฎีแบบสอบประเพณีนิยม (Classical Test Theory Approach) เป็นการศึกษาของ ลอร์ด ที่กระทำภายใต้ข้อตกลงเบื้องต้น 2 ประการ คือ ประการแรกเมื่อผู้เข้าสอบไม่มีความรู้ในคำถามนั้น ผู้เข้าสอบจะตอบข้อนั้นด้วยการเดาสุ่ม ประการที่สองข้อกระทงในแบบสอบชุดนั้นมีความเท่าเทียมกันและตัวเลือกทุกตัวมีแรงดึงดูดใจเท่ากัน ผลการศึกษาพบว่า ความเที่ยงของแบบสอบจะมีค่าสูงสุดเมื่อ

$$a = 1 + \frac{1}{\sqrt{(1-r)p}} \dots\dots\dots (15)$$

เมื่อ	a	แทน	จำนวนตัวเลือกของข้อกระทงที่จะทำได้แบบสอบมีความเที่ยงสูงสุด
	p	แทน	ระดับความยากของข้อกระทง
	r	แทน	สหสัมพันธ์แบบ Product - Moment ระหว่างข้อกระทงที่เทียบเท่ากัน

จากสมการที่ (15) ลอร์ดได้พบว่าสำหรับแบบสอบที่มีระดับความยากปานกลาง ($p = .50$) และมีสหสัมพันธ์ระหว่างข้ออยู่ในช่วง .10 ถึง .70 แล้ว ข้อกระทงชนิด 3 ตัวเลือกจะทำให้แบบสอบมีความเที่ยงสูงสุด ซึ่งสอดคล้องกับการศึกษาของ ทเวอส์กี (Tversky 1964) และ กริเออร์ (Grier 1975)

4. วิธีการที่ใช้ทฤษฎีการตอบสนองข้อกระทง (Item Response Theory Approach) ลอร์ดได้นำเอาทฤษฎีการตอบสนองข้อกระทงมาประยุกต์ใช้กับการศึกษานี้โดยอาศัยข้อมูลการสอบคัดเลือกเข้าศึกษาต่อที่ได้จากแบบสอบวัดความถนัดบุคคลาคีฬ กำนันให้ค่าโอกาสการตอบถูกโดยเดาสุ่มเป็นอิสระจากพารามิเตอร์อื่น ๆ ให้พารามิเตอร์ c_i ของแบบสอบชุด 5 ตัวเลือก 4 ตัวเลือก 3 ตัวเลือก และ 2 ตัวเลือกมีค่าเป็น .200 , .250 , .333 และ .500 ตามลำดับ ขณะที่พิจารณาให้พารามิเตอร์ a_i และ b_i มีค่าคงที่ ผลการศึกษาพบว่า แบบสอบชุด 2 ตัวเลือกที่ประกอบด้วยข้อกระทงจำนวนมาก เหมาะกับผู้เข้าสอบที่มีความสามารถสูง ผู้เข้าสอบที่มีความสามารถต่ำควรใช้แบบสอบชุด 5 ตัวเลือกที่มีจำนวนข้อน้อย



และผู้ใช้สอบที่มีความสามารถปานกลาง ควรใช้แบบสอบชุด 3 ตัวเลือก

วิธีการหั่งสี่ที่ ลอร์ดไครวบรวมมาเสนอไว้ทั้งอยู่บนข้อตกลงเบื้องต้นแห่งความเป็นลักษณะความแนวคิดของทเวอร์สกี คือ เวลาสอบของแบบสอบแต่ละชุดมีค่าเท่ากัน แบบสอบชุดต่าง ๆ สามารถทำให้สั้นลงหรือยาวขึ้น ด้วยการเพิ่มหรือลดจำนวนตัวเลือกอย่างเป็นทางการเป็นสัดส่วนกับจำนวนข้อกระทงในแบบสอบ ซึ่งต่อมา บุคเศสคูและนีโว (Budescu and Nevo 1985 : 183 - 196) ได้สืบสวนหาความตรงเชิงประจักษ์ของข้อตกลงเบื้องต้นนี้ โดยใช้แบบสอบ 3 ฉบับคือ แบบสอบคำศัพท์ 60 ข้อ แบบสอบเหตุผลเชิงคณิตศาสตร์ 25 ข้อ และแบบสอบความเข้าใจในการอ่าน 24 ข้อ ทำการคัดตัวดวงของแบบสอบแต่ละฉบับด้วยวิธีการที่ต่างกัน 2 แบบคือ แบบแรกคัดตัวดวงที่มีแรงดึงดูดใจสูงสุด (มีผู้ใช้สอบเลือกมากที่สุด) ออกจากข้อกระทงที่ละตัวแบบที่สอง คัดตัวดวงที่มีแรงดึงดูดใจน้อยที่สุดออกจากข้อกระทงที่ละตัวจนได้แบบสอบชุด 4 , 3 และ 2 ตัวเลือก กลุ่มตัวอย่างที่ใช้เป็นผู้สมัครสอบคัดเลือกเข้าศึกษาต่อในมหาวิทยาลัยไฮฟา ประจำปีการศึกษา 1983 จัดแบ่งผู้ใช้สอบแบบสุ่มในการแจกแบบสอบ ผลการศึกษาพบว่า กระบวนการคัดตัวดวงมีผลต่อการกำหนดจำนวนตัวเลือกในแบบสอบคำศัพท์ แต่ไม่พบปฏิสัมพันธ์ในแบบสอบเหตุผลเชิงคณิตศาสตร์ และแบบสอบวัดความเข้าใจในการอ่าน ผลอันสำคัญก็คือพบว่า ข้อตกลงเบื้องต้นของความเป็นลักษณะของทเวอร์สกีเป็นไปอย่างไม่สมเหตุผล (not valid) ดังนั้นจำนวนตัวเลือกในข้อกระทงที่ทำให้แบบสอบมีความเที่ยงสูงสุดจึงไม่ควรเป็น 3 ตัวเลือก ด้วยจำข้อตกลงเบื้องต้นที่ใช้ในการศึกษานี้ได้รับการพิสูจน์แล้วว่าไม่สมเหตุผล การที่นักสร้างและพัฒนาแบบสอบคัดสินใจให้ข้อกระทงที่เขาสร้างมีจำนวนตัวเลือกมากกว่า 3 ตัวเลือกจึงเป็นการกระทำที่ถูกต้องแล้ว

การศึกษาเปรียบเทียบคุณภาพของแบบสอบที่มีจำนวนตัวเลือกของข้อกระทงแตกต่างกันในประเทศไทยนั้น งานวิจัย 4 ฉบับได้ศึกษาโดยใช้ทฤษฎีแบบสอบประเพณีนิยม ส่วนงานวิจัยอีกฉบับหนึ่งได้ใช้ทฤษฎีการตอบสนองข้อกระทง งานวิจัยเหล่านี้ได้ใช้วิธีการศึกษาที่แตกต่างกันออกไป ทั้งในส่วนของข้อตกลงเบื้องต้นที่จะกำหนดจำนวนข้อกระทงให้กับแบบสอบแต่ละชุด การกำหนดเวลาสอบ วิธีการคัดตัวดวง ถึงผลการศึกษาที่ได้นำเสนอถึงต่อไปนี้

วารุณี ปัทมวิชัย (2513 : 1 - 98) ทำการศึกษาเปรียบเทียบคุณภาพของแบบสอบในค่าน ความเที่ยง อ่านาจจำแนก และระดับความยาก โดยใช้แบบสอบวัดความถนัดทางการเรียนประเภทอุปมา - อุปไมย ซึ่งสร้างโดยโครงการวิจัยเลือกสรร คณะวิชาวิจัยการศึกษา วิทยาลัยวิชาการศึกษาประสานมิตร เป็นแบบสอบเลือกตอบ 5 ตัวเลือกจำนวน 48 ข้อ ใช้วิธีการสุ่มตัดตัวดวงออกจากข้อกระทงทีละตัวทำให้ได้แบบสอบ 3 ฉบับคือ แบบสอบ 5 ตัวเลือก แบบสอบ 4 ตัวเลือก และแบบสอบ 3 ตัวเลือก แบบสอบทั้งสามฉบับมีจำนวนข้อกระทงเท่ากัน ตัวคำถามข้อเดียวกันเหมือนกัน แต่ต่างกันเฉพาะจำนวนตัวเลือก กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 7 โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย แจกแบ่งกลุ่มตัวอย่างนี้ออกเป็น 3 กลุ่ม ๆ ละ 60 คน นำแบบสอบไปสอบแบบเว้นช่วงละ 1 เดือนจำนวน 3 ครั้ง ผลการวิจัยพบว่า แบบสอบทั้งสามฉบับมีค่าความเที่ยงต่างกัน แบบสอบ 5 ตัวเลือกมีความเที่ยงสูงสุด รองลงไปคือแบบสอบ 4 ตัวเลือกและแบบสอบ 3 ตัวเลือก โดยที่แบบสอบ 5 ตัวเลือกและแบบสอบ 4 ตัวเลือกมีค่าไม่แตกต่างกันในทางสถิติ ส่วนอ่านาจจำแนกและระดับความยากของแบบสอบทั้งสามฉบับมีค่าไม่ต่างกันทางสถิติ

มาลี ชิคสวน (2514 : 1 - 70) ได้ศึกษาเปรียบเทียบค่าความเที่ยง ความตรง อ่านาจจำแนก ระดับความยาก และการกระจายของคะแนนของแบบสอบเลือกตอบที่มีจำนวนตัวเลือกของข้อกระทงแตกต่างกัน โดยใช้แบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ซึ่งเป็นแบบสอบเลือกตอบ 5 ตัวเลือกจำนวน 50 ข้อนำไปทดลองใช้แล้วทำการตัดตัวดวงออกจากข้อกระทง 5 ตัวเลือกอย่างสุ่มทีละตัว ทำให้ได้แบบสอบ 4 ตัวเลือก และ 3 ตัวเลือก ซึ่งแบบสอบทั้งสามฉบับมีจำนวนข้อกระทงเท่ากัน คำถามเหมือนกับข้อข้อ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 390 คนแบ่งออกเป็น 3 กลุ่มโดยให้แต่ละกลุ่มประกอบด้วยนักเรียนที่มีระดับการเรียนเท่าเทียมกัน ผลการวิจัยพบว่า

1. แบบสอบ 5 ตัวเลือกให้ค่าความเที่ยงสูงสุด แล้วลดค่าความลำคัมเมื่อจำนวนตัวเลือกของข้อกระทงลดลงเป็น 4 และ 3 ตัวเลือก โดยแบบสอบ 5 ตัวเลือกและแบบสอบ 3 ตัวเลือกมีค่าความเที่ยงแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

2. แบบสอบทั้งสามฉบับ มีค่าอำนาจจำแนก และค่าความตรงไม่ต่างกัน ในทางสถิติ

3. แบบสอบทั้งสามฉบับ มีระดับความยากแตกต่างกัน โดยแบบสอบ 5 ตัวเลือกมีค่าระดับความยากสูงสุด และระดับความยากของแบบสอบทั้งสามแตกต่างกัน อย่างมีนัยสำคัญที่ระดับ .01

4. คะแนนเฉลี่ยของแบบสอบ 3 ตัวเลือกมีค่าสูงสุด และคะแนนเฉลี่ยของแบบสอบทั้งสามแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01

ครูพ หาดูตระกูล (2519 : 1 - 70) ได้ศึกษาเปรียบเทียบค่าความตรง ความเที่ยง และอำนาจจำแนกของแบบสอบเลือกตอบที่มีจำนวนตัวเลือกในข้อกระทง ต่างกัน ผู้วิจัยได้สร้างแบบสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 7 โดยได้แยกแบบสอบออกเป็น 3 ชุดคือ แบบสอบวัดทักษะทางคณิตศาสตร์ 5 ตัวเลือกจำนวน 30 ข้อ แบบสอบวัดการแก้ปัญหาทางคณิตศาสตร์ 5 ตัวเลือกจำนวน 30 ข้อ และแบบสอบวัดเหตุผลทางคณิตศาสตร์ 5 ตัวเลือกจำนวน 30 ข้อ นำแบบสอบทั้งสามชุดไปทดลองสอบ แล้วตัดตัวลวงที่มีค่าอำนาจจำแนกต่ำสุด ออกทีละตัวตามลำดับจนได้แบบสอบ 4 ตัวเลือก , 3 ตัวเลือก และ 2 ตัวเลือก จึงได้แบบสอบทั้งหมด 12 ชุด นำไปสอบกับกลุ่มตัวอย่าง ผลการวิจัยพบว่า

1. ความเที่ยงของแบบสอบ แบบสอบทักษะทางคณิตศาสตร์ 4 ตัวเลือก มีค่าความเที่ยงสูงกว่าแบบสอบวัดทักษะทางคณิตศาสตร์ 2 ตัวเลือก 3 ตัวเลือก และ 5 ตัวเลือกอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนแบบสอบวัดเหตุผลทางคณิตศาสตร์ 3 ตัวเลือก 4 ตัวเลือก และ 5 ตัวเลือกมีความเที่ยงสูงกว่าแบบสอบ 2 ตัวเลือกอย่างมีนัยสำคัญที่ระดับ .01 แต่แบบสอบวัดการแก้ปัญหาทางคณิตศาสตร์ ที่มีจำนวนตัวเลือกต่างกันมีความเที่ยงไม่ต่างกันทางสถิติ

2. ความตรงของแบบสอบ แบบสอบวัดทักษะทางคณิตศาสตร์และแบบสอบวัดเหตุผลทางคณิตศาสตร์ชุด 2 ตัวเลือกมีความตรงต่ำกว่าแบบสอบชุด 3 ตัวเลือก 4 ตัวเลือก และ 5 ตัวเลือกอย่างมีนัยสำคัญที่ระดับ .01 ในขณะที่แบบสอบวัดการแก้ปัญหาทางคณิตศาสตร์มีความตรงไม่ต่างกันทางสถิติ

3. ค่าอำนาจจำแนกของแบบสอบ ค่าอำนาจจำแนกของแบบสอบทุกชุด
ไม่มีความแตกต่างกันในทางสถิติ

ชนะ ทานะวงศ์ (2521 : 1 - 86) ได้ศึกษาเปรียบเทียบคุณภาพของ
แบบสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ระดับชั้นประถมศึกษา
ปีที่ 4 ในด้านความเที่ยง และความตรง ผู้วิจัยได้สร้างแบบสอบชนิดเลือกตอบ 5
ตัวเลือกจำนวน 3 ฉบับคือ แบบสอบวัดความเข้าใจทางวิทยาศาสตร์ แบบสอบวัด
ทักษะทางวิทยาศาสตร์ แบบสอบวัดปฏิบัติการทางวิทยาศาสตร์ แล้วนำแบบสอบนี้ไป
ทดลองสอบกับนักเรียนที่มีลักษณะคล้ายกับกลุ่มตัวอย่าง วิเคราะห์ผล ทำการตัดตัวลอง
ที่มีค่าอำนาจจำแนกค่าสูงออกจากข้อกระทงที่ละตัวทำให้ได้แบบสอบทั้งหมด 12 ชุด
การกำหนดจำนวนข้อกระทงในแต่ละแบบสอบได้ใช้ข้อตกลงเบื้องต้นของความเป็นสัดส่วน
ตามที่ทเวอร์สกีเสนอไว้ แต่ละฉบับจึงถูกกำหนดให้เป็น 4 ชุดคือ แบบสอบ 2 ตัวเลือก
30 ข้อ แบบสอบ 3 ตัวเลือก 20 ข้อ แบบสอบ 4 ตัวเลือก 15 ข้อ แบบสอบ 5
ตัวเลือก 12 ข้อ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 4 จำนวน 1,284 คน
ผลการวิจัยพบว่า

1. ความเที่ยงของแบบสอบ แบบสอบวัดความเข้าใจชุด 2 ตัวเลือก
ให้ค่าความเที่ยงสูงสุด แบบสอบ 2 ตัวเลือกและ 3 ตัวเลือกให้ค่าความเที่ยงสูงกว่า
แบบสอบ 4 และ 5 ตัวเลือกอย่างมีนัยสำคัญที่ระดับ .05 แบบสอบวัดทักษะที่มีจำนวน
ตัวเลือกต่างกันทุกชุด ให้ค่าความเที่ยงไม่ต่างกันทางสถิติ สำหรับแบบสอบวัด
ปฏิบัติการชุด 4 ตัวเลือกมีค่าความเที่ยงสูงกว่าชุด 2 ตัวเลือกอย่างมีนัยสำคัญที่ระดับ
.05 และแบบสอบ 2 ตัวเลือกมีค่าความเที่ยงสูงกว่าแบบสอบ 3 ตัวเลือกและแบบสอบ
5 ตัวเลือก อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2. ความตรงของแบบสอบ แบบสอบวัดความเข้าใจชุด 5 ตัวเลือกมี
ความตรงสูงกว่าแบบสอบชุด 2 ตัวเลือก 3 ตัวเลือก และ 4 ตัวเลือก อย่างมี
นัยสำคัญทางสถิติที่ระดับ .05 แบบสอบวัดทักษะชุด 2 ตัวเลือกให้ค่าความตรงสูงกว่า
แบบสอบ 3 ตัวเลือก 4 ตัวเลือก และ 5 ตัวเลือกอย่างมีนัยสำคัญทางสถิติที่ระดับ
.05 สำหรับแบบสอบวัดปฏิบัติการชุด 3 ตัวเลือกให้ค่าความตรงสูงกว่าแบบสอบ
4 ตัวเลือก และ 5 ตัวเลือกอย่างมีนัยสำคัญที่ระดับ .05

ธีระศักดิ์ อูร์จันานท์ (2530 : 1 - 109) ทำการศึกษาเปรียบเทียบ
คุณภาพของแบบสอบเลือกตอบที่มีจำนวนตัวเลือกของข้อกระทงต่างกัน โดยใช้ทฤษฎี
การตอบสนองข้อกระทง แบบสอบที่นำมาเปรียบเทียบกันเป็นแบบสอบความถนัดประเภท
ความสามารถในการใช้เหตุผลเชิงภาษาแบบอุปมา - อุปไมยชนิดเลือกตอบจำนวน 3
ฉบับ แต่ละฉบับมีตัวคำถามเดียวกัน มีจำนวนข้อกระทงเท่ากันแต่มีจำนวนตัวเลือกของ
ข้อกระทงต่างกัน คือ แบบสอบ 5 ตัวเลือก แบบสอบ 4 ตัวเลือก และแบบสอบ 3
ตัวเลือก ในการคัดตัวลองออกจากข้อกระทงชนิด 5 ตัวเลือกนั้นใช้วิธีคัดเลือกอย่างสุ่ม นำ
แบบสอบทั้งสามชุดไปสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 2,960 คน โดยแบ่ง
นักเรียนออกเป็น 3 กลุ่มตามระดับความสามารถซึ่งได้แก่ กลุ่มที่มีระดับความสามารถสูง
กลุ่มที่มีระดับความสามารถปานกลาง และกลุ่มที่มีระดับความสามารถต่ำ ผลการวิจัย
พบว่า ในกลุ่มที่มีระดับความสามารถสูงแบบสอบเลือกตอบ 5 ตัวเลือกมีคุณภาพสูงสุด
รองลงมาคือ แบบสอบ 4 ตัวเลือกและ 3 ตัวเลือกตามลำดับ ในกลุ่มที่มีระดับ
ความสามารถปานกลาง แบบสอบ 3 ตัวเลือกมีคุณภาพสูงสุด รองลงมาคือแบบสอบ 5
ตัวเลือก และ 4 ตัวเลือก สำหรับกลุ่มที่มีระดับความสามารถต่ำ แบบสอบ 4 ตัวเลือก
มีคุณภาพสูงสุด รองลงมาคือแบบสอบ 5 ตัวเลือกและ 3 ตัวเลือกตามลำดับ และเมื่อ
พิจารณาตามกรอบของทฤษฎีแบบสอบประเพณีนิยมแล้ว ความเที่ยงของแบบสอบทั้งสาม
ชุดไม่แตกต่างกันในทางสถิติ

ตอนที่ 3 ผลงานวิจัยเกี่ยวกับการตรวจสอบความมั่นใจของการตอบข้อกระทงแบบ
เลือกตอบ

ปัญหาหนึ่งของการใช้แบบสอบชนิดเลือกตอบที่ได้รับความสนใจคือ
กลไกเวลาหลายสิบปีที่ยานมากก็คือ การเคา โดยรูปแบบของข้อกระทงชนิดเลือกตอบ
จะกำหนดทางเลือกในการตัดสินใจไว้ให้กับผู้เข้าสอบแล้ว เมื่อผู้เข้าสอบจำเป็นต้อง
ตอบคำถามนี้ซึ่งที่ไม่ทราบคำตอบที่แท้จริงผู้เข้าสอบจะตอบข้อกระทงนั้นด้วยการ เคา
จากหลักความน่าจะเป็น ถ้าผู้เข้าสอบตอบข้อกระทงนั้นอย่างเคาสุ่มจากทุกตัวเลือก
โอกาสของการตอบถูกจะเป็นหนึ่งในจำนวนตัวเลือกของข้อกระทงนั้น แต่ถ้าวผู้เข้าสอบ
พอมีความรู้ในคำถามนั้นอยู่บ้าง จนทำให้เขาสามารถขจัดตัวลองบางตัวออกไป แล้วจึง



เกาสุ่มจากตัวเลือกที่เหลืออยู่ โอกาสของการตอบข้อกระทงนั้นถูกโดยบังเอิญในแต่ละข้อจึงอาจแตกต่างกัน การตอบโดยการเดานี้ทำให้คะแนนที่ได้เบี่ยงเบนไปจากคะแนนจริง จึงทำให้การตีความหมายของคะแนน และการแจกแจงคะแนนผลการสอบผิดพลาดไปด้วย

ความพยายามที่จะลดความคลาดเคลื่อนที่เกิดจากการเดา นักวิจัยหลายคนได้ศึกษาหาวิธีการเพื่อขจัดโอกาสที่ไม่เท่าเทียมกันของการได้คะแนนจากการเดา และขจัดโอกาสที่ผู้เข้าสอบจะได้คะแนนเพราะการเดาออกไป ในบรรดาวิธีการทั้งหลาย วิธีการหนึ่งที่ถูกนำมาใช้ เพื่อที่จะให้ได้รับข้อมูลจากการตอบแบบสอมละเอียดยิ่งขึ้นก็คือวิธีการที่เรียกว่า การตรวจสอบความมั่นใจ (Confidence Testing) เป็นวิธีการสอบที่กำหนดให้นักคะแนนโดยทางตรงหรือโดยทางอ้อมให้แก่ค่าตอบ ในลักษณะที่จะสะท้อนให้เห็นความเชื่อมั่นของผู้เข้าสอบว่า ตัวเลือกที่เขาได้เลือกนั้นจะเป็นตัวถูก การตรวจสอบความมั่นใจมีวิธีการแตกต่างกันหลายวิธี

(Echternacht 1972 : 218)

เมื่อกระบวนการตรวจสอบความมั่นใจถูกนำมาใช้ในคอนตันศวรรษนั้น มีวัตถุประสงค์เพื่อใช้เป็นวิธีการหนึ่งในการทำให้ได้ข้อมูลเพิ่มขึ้นจากการใช้แบบสอมชนิดเลือกตอบ ในทศวรรษที่ 1930 กระบวนการนี้ได้กลายมาเป็นวิธีการหนึ่งที่จะทำให้การเดามีผลน้อยที่สุดต่อการสอมควยแบบสอมถูก - ผิด

เฮฟเนอร์ (Hevner 1932 : 359 - 362) ได้รายงานผลการศึกษาว่า เมื่อใช้การตรวจสอบความมั่นใจในการบริหารแบบสอม เป็นผลให้ความเที่ยงของแบบสอมถูก - ผิด มีค่าสูงขึ้น โดยเขาใช้ระบบการให้คะแนนที่ต่างกัน 4 แบบ ในการตรวจสอบความแตกต่างกันของความเที่ยง คือ การให้คะแนนที่คิดจาก

1. จำนวนข้อที่ตอบถูก
2. จำนวนข้อที่ตอบถูก - จำนวนข้อที่ตอบผิด
3. คะแนนถ่วงน้ำหนักของข้อที่ตอบถูก คือ เมื่อตอบถูกควยความมั่นใจมากได้ 3 คะแนน ถ้าตอบถูกมีความมั่นใจปานกลางได้ 2 คะแนน และเมื่อตอบถูกควยความมั่นใจน้อยจะได้ 1 คะแนน

4. คะแนนดวงน้ำหนักของข้อที่ตอบถูก - คะแนนดวงน้ำหนักของข้อที่ตอบผิด

เฮฟเนอร์ได้พบว่า วิธีการให้คะแนนแบบดวงน้ำหนักของข้อที่ตอบถูก ทำให้ความเที่ยงของแบบสอบถามมีค่าสูงสุด นอกจากนั้นเฮฟเนอร์ยังได้ให้ข้อสังเกตด้วยว่า กลุ่มตัวอย่างของการทดลองนี้มีความเต็มใจในการแสดงระดับความมั่นใจ โดยเฉพาะอย่างยิ่ง คนที่ไม่มั่นใจในการตอบ ทั้งเขายังคิดว่าควรเก็บวิธีการให้คะแนนเป็นความลับไม่ให้ผู้เข้าสอบทราบ ซึ่งเป็นการไม่เปิดโอกาสให้ผู้เข้าสอบที่ไม่ซื่อสัตย์ได้คะแนนเพิ่มขึ้น

โซเคอร์ควิสต์ (Soderquist 1936 : 290 - 292) ได้ทำการศึกษาที่คล้ายกันนี้ แต่ระบบการให้คะแนนที่โซเคอร์ควิสต์ใช้ค่อนข้างแตกต่างจากของเฮฟเนอร์ โดยระบบการให้คะแนนของโซเคอร์ควิสต์ ผู้เข้าสอบได้คะแนนของการตอบถูกต่างกัน ความระดับความมั่นใจที่ไคร้ระบุ คะแนนจึงกลายเป็น 4 , 3 , 2 และ 1 ส่วนข้อที่ตอบผิดจะกำหนดน้ำหนักคะแนนเป็น 2 เท่าคือ 8 , 6 , 4 และ 2 การสอบนี้ใช้แบบสอบถาม - ผิด และนักเรียนได้ทราบระบบคะแนนที่ใช้ การให้คะแนนใช้ทั้งคะแนนดวงน้ำหนักของข้อที่ตอบถูก - คะแนนดวงน้ำหนักของข้อที่ตอบผิด และจำนวนข้อที่ตอบถูก - จำนวนข้อที่ตอบผิด เขาได้พบว่า การใช้ระบบการให้คะแนนแบบดวงน้ำหนัก ทำให้ค่าความเที่ยงของแบบสอบถามสูงกว่าการให้คะแนนแบบไม่ดวงน้ำหนัก ซึ่งเป็นข้อค้นพบที่แตกต่างไปจากเฮฟเนอร์ ความแตกต่างนี้อาจเนื่องมาจากการที่กลุ่มตัวอย่างของโซเคอร์ควิสต์ได้ทราบระบบการให้คะแนนในขณะที่กลุ่มตัวอย่างของเฮฟเนอร์ไม่ทราบระบบการให้คะแนน กล่าวได้ว่าองค์ประกอบทางบุคลิกภาพอาจเป็นส่วนที่ทำให้ผลการวิจัยทั้งสองนี้แตกต่างกัน

สไวน์ฟอร์ด (Swineford 1938 : 295 - 300) ได้ศึกษาตัวแปรทางบุคลิกภาพ โดยบริหารแบบสอบถาม - ผิดตามคำสั่งที่โซเคอร์ควิสต์เคยใช้มาก่อน ทำให้สามารถหาคะแนนที่เขาเองเรียกว่า คะแนนการเสี่ยง (Gambling Score) จากค่าตอบที่บ่งความมั่นใจนี้ ต่อมาในปีคริสต์ศักราช 1941 สไวน์ฟอร์ดได้ใช้แบบสอบถามอีก 4 ฉบับเพื่อวัดแนวโน้มของการเสี่ยงได้ค้นพบว่า เด็กชายมีแนวโน้มของการเสี่ยงสูงกว่าเด็กหญิงอย่างมีนัยสำคัญ ทั้งเด็กชายและเด็กหญิงมีแนวโน้มของการเสี่ยงในเนื้อหาวิชาที่ไม่คุ้นมากกว่าในเนื้อหาวิชาที่คุ้น และคะแนนการเสี่ยงมีแนวโน้มเป็น

อิสระจากคะแนนของแบบสอบวัดผลสัมฤทธิ์

เกรสเซล และ ชมิค (Dressel and Schmid 1953 : 574 - 595)
ได้ทำการวิจัยโดยใช้แบบสอบรูปแบบต่าง ๆ กันใช้ระบบการให้คะแนนที่ต่างกัน เพื่อ
พยายามปรับปรุงอำนาจจำแนกของแบบสอบเลือกตอบที่ไม่จำกัดเวลาของการสอบ
แบบสอบทั้ง 4 แบบที่ใช้ได้แก่

1. แบบสอบเลือกตอบเสรี (Free - Choice Test) เป็นแบบสอบ
เลือกตอบชนิด 5 ตัวเลือกที่ผู้เข้าสอบสามารถเลือกคำตอบใดก็ได้ โดยให้
คำตอบที่เลือกนั้นมีคำตอบถูกรวมอยู่ด้วย
2. แบบสอบแสดงระดับความมั่นใจ (A Degree of Certainty
Test) เป็นแบบสอบเลือกตอบที่ผู้เข้าสอบเลือกตอบ 1 ตัวเลือก พร้อมทั้งบ่งระดับ
ความมั่นใจของความถูกต้องของคำตอบที่เลือกนั้น
3. แบบสอบที่มีคำตอบถูกหลายตัวเลือก (A Multiple - Answer
Test) เป็นแบบสอบเลือกตอบที่มีคำตอบถูกหลายตัว ให้ผู้เข้าสอบเลือกตัวเลือกทุกตัวที่
เป็นคำตอบนั้น
4. แบบสอบที่มีคำตอบถูกสองคำตอบ (A Two - Answer Test)
เป็นแบบสอบเลือกตอบ 5 ตัวเลือกที่รู้แน่นอนว่ามีคำตอบ 2 ตัวเลือก

ในแต่ละกรณีผู้เข้าสอบจะต้องเข้าใจถึงชนิดของแบบสอบที่ตนทำ เนื่องจาก
แบบสอบสองแบบแรกเท่านั้นที่เกี่ยวข้องกับการตรวจสอบความมั่นใจ และเฉพาะแบบแรก
เท่านั้นที่เกี่ยวข้องกับการวิจัยนี้ จึงขอแสดงวิธีการให้คะแนนเฉพาะแบบแรกเท่านั้น
ดังตารางที่ 3

ตารางที่ 3 วิธีการให้คะแนนของแบบสอบถามเลือกตอบเสรีตามวิธีของเกรสเชล และ
ชมิค

จำนวนตัวเลือกในคำตอบ	คะแนน	
	ถ้ามีตัวเลือกรวมอยู่ด้วย	ถ้าไม่มีตัวเลือกรวมอยู่ด้วย
1	4	- 1
2	3	- 2
3	2	- 3
4	1	- 4
5	0	-

เกรสเชล และ ชมิค ได้พบว่าเมื่อใช้คะแนนที่ได้จากแบบสอบถามเป็นเกณฑ์
นักเรียนกลุ่มเก่งตอบแบบสอบถามเลือกตอบเสรีด้วยจำนวนตัวเลือกน้อยกว่านักเรียนในกลุ่ม
ปานกลางและนักเรียนในกลุ่มอ่อนอย่างมีนัยสำคัญ ในขณะที่แบบสอบถามแสดงความมั่นใจ
ไม่สามารถแสดงให้เห็นลักษณะการตอบที่แตกต่างกันของนักเรียนในกลุ่มสูง ปานกลาง
และอ่อน ได้ดีเท่าวิธีแรกเพราะเมื่อพิจารณาคำตอบแล้วพบว่า ผู้เข้าสอบตอบข้อกระทง
ที่มีความยากปานกลางและยากมากด้วยระดับความมั่นใจเดียวกัน และยังสรุปด้วยว่า
องค์ประกอบที่วัดได้ด้วยแบบสอบถามเลือกตอบเสรีนั้นแตกต่างจากที่วัดได้ด้วยแบบสอบถาม
แสดงความมั่นใจ

ต่อมา กิบมอนส์ โอลคิน และโซเบล (Gibbons , Olkin and Sebel
1979 : 259 - 270) ได้เสนอวิธีการตอบและตรวจให้คะแนนแบบสอบถามเลือกตอบ
ที่คล้ายกับวิธีการให้คะแนนของแบบสอบถามเลือกตอบเสรีซึ่งเกรสเชล และ ชมิค ได้เสนอ
ไว้ กิบมอนส์ และคณะเรียกวิธีการนี้ว่า " วิธีการสรรหาเซตย่อยของคำตอบ "
(Subset Selection Technique) วิธีการนี้ผู้เข้าสอบจะเลือกเซตย่อยของ
คำตอบใดก็ได้เซตหนึ่ง โดยที่เขาต้องมั่นใจว่าในเซตย่อยที่เลือกมานั้นจะรวมตัวเลือก

อยู่ด้วยตัวหนึ่ง ขนาดของเซตย่อยอาจประกอบด้วยจำนวนตัวเลือกตั้งแต่ 0 ตัวเลือก คือผู้เข้าสอบไม่ประสงค์จะเลือกตัวเลือกใดเลย ทั่วยังไม่มีความรู้ในข้อนั้นและไม่ต้องการเดา จนถึงมีขนาดเท่ากับจำนวนตัวเลือกของข้อกระทงนั้น เช่น ถ้าข้อกระทงนั้นมีตัวเลือกเป็น ก ข ค เซตย่อยของคำตอบที่เป็นไปได้ทั้งหมดคือ - (เซตว่าง) , ก , ข , ค , ก ข , ก ค , ข ค และเซต ก ข ค สำหรับการให้คะแนนพิจารณาจากขนาดของเซตย่อย และในเซตย่อยนั้นมีตัวถูกรวมอยู่ด้วยหรือไม่ คือเมื่อเซตย่อยที่เลือกมานั้นมีตัวถูกรวมอยู่ด้วยข้อนั้นจะได้คะแนนเท่ากับ จำนวนตัวเลือกของข้อกระทง - จำนวนตัวเลือกในเซตย่อย แต่ถ้าในเซตย่อยที่เลือกมานั้นไม่มีตัวถูกรวมอยู่ด้วยคะแนนที่ได้จะเป็น ค่าลบที่เท่ากับจำนวนตัวเลือกในเซตย่อยนั้น ซึ่งการคิดคะแนนของแบบสอบทั้งสามชุดที่ใช้ในการวิจัยนี้เป็น กิ่งการวางที่ 4

ตารางที่ 4 การคิดคะแนนตามวิธีการสรรหาเซตย่อยของคำตอบ ของแบบสอบทั้งสามชุด

จำนวนตัวเลือก ในเซตย่อย	คะแนนเมื่อตอบถูกของแบบสอบ			คะแนนเมื่อตอบผิดของแบบสอบ		
	3ตัวเลือก	4ตัวเลือก	5ตัวเลือก	3ตัวเลือก	4ตัวเลือก	5ตัวเลือก
0	0	0	0	0	0	0
1	2	3	4	- 1	- 1	- 1
2	1	2	3	- 2	- 2	- 2
3	0	1	2	0	- 3	- 3
4	-	0	1	-	0	- 4
5	-	-	0	-	-	0

การคิดคะแนนตามวิธีการสรรหาเซตย่อยของคำตอบนี้ ยึดหลักการว่า คะแนนคาดหวัง (Expected Score) ที่เกิดขึ้นจากการเดาสุ่มยอมเท่ากับศูนย์ เช่นในกรณีข้อกระทงมี 5 ตัวเลือก ถ้าผู้เข้าสอบตอบอย่างเดาสุ่มทั่วยเซตย่อยของคำตอบที่มี

2 ตัวเลือก ความน่าจะเป็นของการตอบถูกเป็น $\frac{2}{5}$ ขณะที่ความน่าจะเป็นของการตอบผิดเป็น $\frac{3}{5}$ กรณีนี้ถ้าเขทย่อยของค่าคอมที่เลือกนั้นมีตัวถูกรวมอยู่ด้วยจะได้คะแนน 3 คะแนน (5 - 2) แต่ถาเขทย่อยของค่าคอมนั้นไม่มีตัวถูกรวมอยู่ด้วยจะได้คะแนนเป็น - 2 คะแนน กึ่งนั้น

$$\begin{aligned} \text{คะแนนคาดหวังของการเกาสุ่ม} &= 3 \left(\frac{2}{5} \right) - 2 \left(\frac{3}{5} \right) \\ &= 0 \end{aligned}$$

วิธีการสรรหาเขทย่อยของค่าคอมมีข้อดีคือ ช่วยควบคุมการเกา โดยทำให้คะแนนความคาดหวังของการเกาสุ่มมีค่าเป็นศูนย์ ผู้เข้าสอบจึงไม่ได้รับประโยชน์จากการตอบโดยการเกาสุ่ม และวิธีนี้ยังช่วยส่งเสริมให้ผู้เข้าสอบได้มีโอกาสใช้ส่วนของความรู้ที่มีอยู่บ้างในการตอบข้อกระทงนั้น ผู้เข้าสอบที่มีความรู้จริง หรือมีความมั่นใจมากก็จะเลือกเขทย่อยที่ประกอบด้วยตัวเลือกเพียงตัวเลือกเดียว ขณะที่ผู้เข้าสอบที่มีความรู้อยู่บ้าง แต่ไม่ทราบว่าตัวเลือกใดเป็นตัวถูก เขาอาจจะช้จักตัวดวงบางตัวออกไป ทำให้เขทย่อยของค่าคอมที่เลือกมีขนาดใหญ่มากกว่าผู้เข้าสอบที่รู้จริง ขนาดของเขทย่อยที่ผู้เข้าสอบเลือกจึงแสดงระดับความมั่นใจของการตอบ และเมื่อพิจารณาในแง่ของบุคลิกภาพแล้ว วิธีการนี้ช่วยให้ผู้เข้าสอบสามารถตอบได้ทุกข้อกระทงในแบบสอบ โดยไม่ช้กับลักษณะการเสี่ยงของบุคคลบางคนในการตอบแบบสอบเลือกตอบอีกด้วย

จาราคาท และ ซาแวก (Jaradat and Sawaged 1986 : 369 - 376) ได้ทำการศึกษาเปรียบเทียบผลที่เกิดจาก การตอบและการตรวจให้คะแนนด้วยวิธีการสรรหาเขทย่อยของค่าคอม วิธีการตรวจให้คะแนนแบบ 0 - 1 และวิธีการให้คะแนนแบบใช้สูตรแก่การเกา โดยเขาได้ใช้คะแนนผลการเรียนในภาคเรียนแรก และคะแนนจากแบบสอบวัดลักษณะการเสี่ยงที่เขาสร้างขึ้น จำแนกนักศึกษาที่ใช้เป็นกลุ่มตัวอย่าง จำนวน 160 คนออกเป็น 4 กลุ่ม ๆ ละ เท่า ๆ กันคือ กลุ่มผลสัมฤทธิ์ทางการเรียนสูง - ลักษณะการเสี่ยงสูง , กลุ่มผลสัมฤทธิ์ทางการเรียนสูง - ลักษณะการเรียนต่ำ , กลุ่มผลสัมฤทธิ์ทางการเรียนต่ำ - ลักษณะการเสี่ยงสูง และกลุ่มผลสัมฤทธิ์ทางการเรียนต่ำ - ลักษณะการเสี่ยงต่ำ เครื่องมือที่ใช้เป็นแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์แบบเลือกตอบ 4 ตัวเลือกจำนวน 60 ข้อ

จัดแบ่งออกเป็น 3 ชุด ๆ ละ 20 ข้อแต่ละชุดมีลักษณะเทียบเท่ากันในด้านของเนื้อหา ระดับความยาก และระดับพฤติกรรมที่วัด การบริหารแบบสอนใช้วิธีสุ่มคำชี้แจงแบบใดแบบหนึ่งให้กับแบบสอนแต่ละชุด จึงทำให้แบบสอนทั้งสามชุดมีคำชี้แจงที่ต่างกัน ผลของการศึกษาพบว่า สัมประสิทธิ์ความตรงของแบบสอนที่ใช้วิธีการสรรหาเซทย่อยของค่าคอมมีค่าสูงกว่าสัมประสิทธิ์ความตรงของแบบสอนที่ใช้วิธีการทอมและตรวจให้คะแนนแบบแก่การเคา และแบบ 0 - 1 อย่างมีนัยสำคัญในทุกกลุ่ม ยกเว้นในกลุ่มผลสัมฤทธิ์ทางการเรียนสูง - ลักษณะการเสี่ยงต่ำ และในกลุ่มผลสัมฤทธิ์ทางการเรียนต่ำ - ลักษณะการเสี่ยงต่ำ ซึ่งได้ค่าต่ำกว่าแบบสอนที่ใช้วิธีการตรวจและให้คะแนนแบบแก่การเคา และในกลุ่มผลสัมฤทธิ์ทางการเรียนสูง - ลักษณะการเสี่ยงต่ำ กับกลุ่มผลสัมฤทธิ์ทางการเรียนต่ำ - ลักษณะการเสี่ยงสูง สัมประสิทธิ์ความตรงของแบบสอนที่ใช้วิธีการสรรหาเซทย่อยของค่าคอมมีค่าต่ำกว่าวิธีการทอมและตรวจให้คะแนนแบบ 0 - 1 ส่วนความเที่ยงของแบบสอนที่ใช้วิธีการสรรหาเซทย่อยของค่าคอมมีค่าสูงกว่าแบบสอนที่ใช้วิธีการทอมและตรวจให้คะแนนแบบ 0 - 1 ในทุกกลุ่ม

ส่วนในประเทศไทย มีงานวิจัยอย่างมากที่ศึกษาผลของการใช้วิธีการทอมและตรวจให้คะแนนเพื่อแก้ไขความคลาดเคลื่อนที่เกิดจากการเคา แต่ที่เกี่ยวข้องโดยตรงกับวิธีการสรรหาเซทย่อยของค่าคอมนั้นมีเพียงงานวิจัยของ ทวี ทองคำ (2526 : 1 - 91) เท่านั้นซึ่งได้ทำการศึกษาเปรียบเทียบค่าความเที่ยง ความตรง และค่าอำนาจจำแนกของแบบสอนเลือกทอมที่ใช้วิธีการทอมและการตรวจให้คะแนนต่างกัน 3 วิธีคือ วิธี 0 - 1 วิธีของสำราญ มีแจ้ง และวิธีของกิมบอนส์และคณะ* กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นประกาศนียบัตรวิชาชีพปีที่ 1 จำนวน 180 คน แบ่งกลุ่มตัวอย่างออกเป็น 3 กลุ่มด้วยวิธีสุ่มอย่างมีระบบ แล้วทำการเก็บรวบรวมข้อมูลด้วยแบบสอนวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์พหุคูณขัณกรม 2 ซึ่งผู้วิจัยสร้างขึ้นเอง ผลการวิจัยพบว่า ค่าความเที่ยง ความตรง และค่าอำนาจจำแนกของแบบสอนที่ใช้วิธีการทอมและการตรวจให้คะแนนต่างกันทั้งสามวิธีนั้น ไม่มีความแตกต่างกันในทางสถิติ

* วิธีการสรรหาเซทย่อยของค่าคอม

ผู้วิจัยได้ให้ข้อเสนอเพิ่มเติมที่น่าสนใจคือ กลุ่มตัวอย่างถึงร้อยละ 75 ที่เลือกเพียง
ตัวเลือกเดียวในข้อกระทงหนึ่ง ๆ เมื่อใช้วิธีการทอมและทรวจให้คะแนนตามวิธีของ
สําราชู มีแจ้ และวิธีของกิฆมอนส์และคณะ ซึ่งไม่เป็นไปตามวัตถุประสงค์ของทั้ง
สองวิธี ที่ต้องการวัดส่วนของความรู้ให้ละเอียดกว่าวิธี 0 - 1 และถือว่าเป็นความ
จำกัดของการวิจัยนี้