



บทที่ 2

บรรณคดีที่เกี่ยวข้อง

ผู้วิจัยได้แบ่งบรรณคดีที่เกี่ยวข้องกับการวิจัยเรื่องนี้ออกเป็น 2 ตอนคือ ตอนที่ 1 เป็นมโนทัศน์เบื้องต้นเกี่ยวกับคุณภาพของแบบสอบ คือ ค่าความเที่ยง ค่าความตรง และ ค่าอำนาจจำแนก ตลอดจนความคลาดเคลื่อนมาตรฐานในการวัด ส่วนตอนที่ 2 จะเป็น บรรณคดีและงานวิจัยที่เกี่ยวข้องกับการใช้คำสั่งและวิธีการให้คะแนนที่ต่างกัน

ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับคุณภาพของแบบสอบ

ความเที่ยงของแบบสอบ

1. ได้มีผู้ให้ความหมายของความเที่ยงไว้หลายท่าน เช่น
ความเที่ยง หมายถึงความคงที่ของผลที่ได้รับจากการใช้เครื่องมือชนิดเดียวกัน สอบมากกว่า 1 ครั้ง (Gronlund 1968: 163)
ความเที่ยง หมายถึงระดับ (degree) ของความคงที่ ระหว่างการวัดสิ่งเดียวกัน 2 ครั้ง (Mehrens Lehmann 1973: 88)
ความเที่ยง หมายถึงความคงที่ของคะแนนที่ได้จากการวัด (Gronlund 1976: 105)
ความเที่ยง หมายถึงความคงที่ของคะแนนที่ได้จากการสอบบุคคลกลุ่มเดียวกัน ด้วยแบบสอบชุดเดียวกัน หรือสอบด้วยแบบสอบคนละชุดที่เทียบเท่ากัน หรือภายใต้สภาพการสอบที่ต่างกัน (อนาสตาซี 2519: 73)
จากความหมายดังกล่าวข้างต้นสรุปได้ว่า ความเที่ยงหมายถึงความคงที่ของคะแนนที่ได้จากการวัดด้วยเครื่องมือชุดเดียวกันหรือเทียบเท่ากันในโอกาสที่ต่างกัน

ทฤษฎีของความเที่ยง (Theory of Reliability)

การวัดใดๆก็ตามย่อมจะต้องมีความคลาดเคลื่อนเกิดขึ้นเสมอ โดยเฉพาะอย่างยิ่งการวัดในทางจิตวิทยาด้วยแล้วจะมีค่าความคลาดเคลื่อนมากกว่าการวัดทางกายภาพ เช่นในการทดสอบนักเรียน คะแนนที่ได้จากการทดสอบจะมีความคลาดเคลื่อนจากสาเหตุต่างๆได้ เช่น การคุมสอบ การตรวจให้คะแนน เวลาในการสอบ ตลอดจนความคลาดเคลื่อนที่เกิดจากผู้สอบเอง ซึ่งต่างก็เป็นสาเหตุให้คะแนนที่ทดสอบได้เปลี่ยนแปลงไปจากความเป็นจริงที่ควรจะเป็น ดังนั้นคะแนนที่ทดสอบได้ของนักเรียนจะประกอบด้วยคะแนน 2 ส่วนคือ ส่วนที่เป็นตัวแทนสมรรถภาพที่แท้จริงของนักเรียนหรือเรียกว่าคะแนนจริง (True Score) ซึ่งจะคงที่ตลอดเวลาไม่ว่าจะใช้เครื่องมือใดๆที่มีคุณสมบัติในการวัดสมรรถภาพอันเดียวกันวัดอีกส่วนหนึ่งคือคะแนนความคลาดเคลื่อน (Error Score) ที่เกิดจากปัจจัยต่างๆดังกล่าวแล้ว ซึ่งเป็นเหตุให้นักเรียนตอบถูกทั้งๆที่ตัวเองไม่รู้ หรือตอบผิดทั้งๆที่ตัวเองรู้ และอาจกล่าวได้ว่าเราไม่มีวิธีการใดๆไปตรวจสอบเรื่องนี้ได้

จากความจริงดังกล่าวมาแล้วอาจแสดงความสัมพันธ์ระหว่างคะแนนส่วนต่างๆได้ดังนี้ (กิลฟอร์ด ฟรังก์เตอร์ 2526: 553-555)

$$X_t = X_{\infty} + X_e$$

เมื่อ X_t แทนคะแนนที่ได้จากการวัด

X_{∞} แทนส่วนที่เป็นคะแนนจริง

X_e แทนส่วนที่เป็นคะแนนความคลาดเคลื่อน

จากความสัมพันธ์จะเห็นว่าไม่สามารถแก้สมการหาค่า X_{∞} ได้ไม่ว่าจะทำการทดสอบกี่ครั้งก็ตาม เพราะไม่สามารถทราบค่า X_e แต่เราก็คือว่าคะแนนความคลาดเคลื่อนเกิดขึ้นอย่างอิสระและแบบสุ่มโดยมีขั้วนิยมเลขคณิตเป็นศูนย์กลาง ไม่มีสหสัมพันธ์กับส่วนที่เป็นคะแนนจริง ไม่มีสหสัมพันธ์กับคะแนนความคลาดเคลื่อนในการวัดอื่น จึงสามารถเขียนสมการเชิงความแปรปรวนได้ดังนี้

$$\sigma_t^2 = \sigma_{\omega}^2 + \sigma_e^2$$

เมื่อ σ_t^2 แทนความแปรปรวนของคะแนนผลการสอบ

σ_{ω}^2 แทนความแปรปรวนของคะแนนจริง

σ_e^2 แทนความแปรปรวนของคะแนนความคลาดเคลื่อน

ความเที่ยง หมายถึงอัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่วัดได้ จึงสามารถเขียนเป็นสมการได้เป็น

$$r_{tt} = \frac{\sigma_{\omega}^2}{\sigma_t^2}$$

$$r_{tt} = 1 - \frac{\sigma_e^2}{\sigma_t^2}$$

จะเห็นว่าค่าความเที่ยงจะสูงขึ้นเมื่อความแปรปรวนของความคลาดเคลื่อนลดลง โดยกำหนดความแปรปรวนของคะแนนผลการสอบคงที่หรือเพิ่มขึ้น

วิธีประมาณค่าความเที่ยงของแบบสอบ

โดยทั่วไปมีวิธีประมาณค่าความเที่ยงของแบบสอบ 4 แบบ (Mehrens Lehmann 1973: 109) คือ

1. แบบวัดซ้ำ (Measure of Stability)
2. แบบใช้ข้อสอบคล้าย (Measure of Equivalence)
3. แบบใช้ข้อสอบคล้ายและวัดซ้ำ (Measure of Equivalence and Stability)
4. แบบวัดความคงที่ภายใน (Measure of Internal Consistency)
 - 4.1 แบบแบ่งครึ่งข้อสอบ (Split - Half)
 - 4.2 แบบ กูดเคอร์ ริชาร์ดสัน (Kuder - Richardson Estimates)
 - 4.3 แบบวิเคราะห์ความแปรปรวนของ ข้อสอบ (Hoyt's Analysis of Variance)

ในที่นี้จะขอกล่าวถึงเฉพาะวิธีการประมาณค่าความเที่ยงแบบวิเคราะห์ความแปรปรวนของ ฮอยท์ ซึ่งเป็นวิธีประมาณที่ใช้ในการวิจัยครั้งนี้

วิธีประมาณค่าความเที่ยงแบบวิเคราะห์ความแปรปรวนของ ฮอยท์

การประมาณค่าความเที่ยงโดยวิธีวิเคราะห์ความแปรปรวนของฮอยท์ เป็นการประมาณค่าความเที่ยงชนิดความคงที่ภายในซึ่ง ซี เจ ฮอยท์ เป็นผู้พัฒนาสูตรขึ้นเมื่อ ค.ศ. 1941 (Mehrens Lehmann 1975 : 47-48) โดยถือว่าคะแนนของคนที่คนหนึ่งจากแบบสอบถามแบ่งออกได้เป็น 4 องค์ประกอบคือ

1. องค์ประกอบที่รวมอยู่ในทุกคนและทุกข้อกระทง เป็นค่าคงที่ในทุกคนและทุกข้อ
2. องค์ประกอบที่มีลักษณะเฉพาะตัวในแต่ละข้อกระทง
3. องค์ประกอบที่เกี่ยวข้องกับบุคคลแต่ละคน
4. องค์ประกอบที่เกิดจากความคลาดเคลื่อน

นอกจากองค์ประกอบทั้ง 4 แล้ว องค์ประกอบที่เกิดจากความคลาดเคลื่อนมีการกระจายแบบปกติ และแต่ละข้อกระทงเป็นอิสระจากกันซึ่งสามารถประมาณค่าความเที่ยงได้ตามลำดับขั้นดังนี้

นักเรียนคนที่	ข้อสอบข้อที่						คะแนนรวม
	1	2	3	4	...	n	
1							t_1
2							t_2
3							t_3
⋮							⋮
⋮							⋮
k							t_k
	p_1	p_2	p_3	p_4	...	p_n	$\sum p_i = \sum t_i$

เมื่อ n แทนจำนวนข้อสอบ
 k แทนจำนวนนักเรียน
 t_i แทนคะแนนของนักเรียนแต่ละคน
 p_i แทนคะแนนของแต่ละข้อ

1. หา Sum of Squares "among students"

$$SS_{st} = \frac{1}{n} \sum t_i^2 - \frac{(\sum t_i)^2}{nk}$$

2. หา Sum of Squares "among items"

$$SS_{it} = \frac{1}{k} \sum p_i^2 - \frac{(\sum t_i)^2}{nk}$$

3. หา Total Sum of Squares

$$SS_{tol} = \frac{(\sum t_i)(nk - \sum t_i)}{nk}$$

4. หา Sum of Squares "remainder"

$$SS_{rem} = SS_{tol} - SS_{st} - SS_{it}$$

5. หา $MS_{st} = SS_{st} / (k - 1)$

6. หา $MS_{rem} = SS_{rem} / (n-1)(k-1)$

7. หา $r_{tt} = 1 - MS_{rem} / MS_{st}$

องค์ประกอบที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบ

นอกจากวิธีการที่ใช้ในการประมาณค่าความเที่ยงจะมีอิทธิพลต่อค่าความเที่ยงของแบบสอบแล้ว ยังมีองค์ประกอบอื่นๆที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบด้วยคือ (Mehrens Lehmann 1973: 100-103)

1. ความยาวของแบบสอบ (Test Length) การเพิ่มข้อสอบที่มีคุณภาพเท่าเทียมกับข้อสอบเดิม จะทำให้แบบสอบมีค่าความเที่ยงสูงขึ้น นั่นคือแบบสอบที่มีจำนวน

ข้อมากจะมีความเที่ยงสูงกว่าแบบสอบที่มีจำนวนข้อน้อย สูตรที่แสดงให้เห็นความสัมพันธ์ระหว่างความยาวของแบบสอบกับความเที่ยงคือสูตรของ สเปียร์แมน-บราวน์ (Spearman-Brown Formula)

2. ความเป็นเอกพันธ์ของกลุ่มผู้รับการทดสอบ (Group Homogeneity) กลุ่มที่มีความเป็นเอกพันธ์จะมีหีสัยของคะแนนความสามารถของกลุ่มน้อยกว่ากลุ่มที่มีความเป็นวิวิธพันธ์ (Group Heterogeneous) ซึ่งจะหาค่าความเที่ยงของแบบสอบของกลุ่มที่มีความเป็นเอกพันธ์ต่ำกว่ากลุ่มที่มีความเป็นวิวิธพันธ์

3. ความยากของข้อสอบ (Item Difficulty) เนื่องจากค่าความเที่ยงนั้นประมาณจากความแปรปรวนของคะแนนผลการสอบ ถ้าข้อสอบยากเกินไปทุกคนทำผิดหมด หรือข้อสอบง่ายเกินไปทุกคนทำถูกหมด ก็จะทำให้หีสัยของคะแนนลดลง ค่าความแปรปรวนของคะแนนก็จะน้อย ซึ่งจะหาค่าความเที่ยงลดลงด้วย ดังนั้นถ้าค่าความยากของข้อสอบทำให้คะแนนกระจายมากก็จะหาค่าความเที่ยงของแบบสอบมีค่าสูงขึ้นด้วย

4. ความเร็ว (Speed) การประมาณค่าความเที่ยงของแบบสอบความเร็ว (Speed Test) ด้วยวิธีวัดความถี่ที่ภายในเป็นวิธีไม่เหมาะสม เพราะแบบสอบชนิดนี้เป็นแบบสอบที่ง่าย แต่ให้เวลาน้อย ซึ่งนักเรียนไม่สามารถทำได้ครบทุกข้อ และจะหาค่าความเที่ยงสูงกว่าที่ควรจะเป็น

5. ความเป็นปรนัย (Objectivity) ถ้าแบบสอบมีความเป็นปรนัยสูง ค่าความเที่ยงของแบบสอบก็จะสูงขึ้นด้วย

นอกจากนี้ การบริหารแบบสอบ เช่น การจับเวลา การให้คำชี้แจง การป้องกันการทำทุจริต เป็นต้น หรือองค์ประกอบที่เกี่ยวกับตัวผู้สอบเอง เช่น แรงจูงใจ ความวิตกกังวล ประสบการณ์ในการสอบ การเดา เป็นต้น สิ่งเหล่านี้อาจเป็นแหล่งความคลาดเคลื่อนที่ทำให้ค่าความเที่ยงของแบบสอบเปลี่ยนแปลงได้

ความตรงของแบบสอบ

ได้มีผู้ให้ความหมายของความตรงของแบบสอบไว้หลายท่าน เช่น
ความตรง หมายถึงความถูกต้องแน่นอนในการวัดสิ่งที่ต้องการวัด (Lindquist 1942: 213)

ความตรงของแบบสอบ หมายถึงค่าสหสัมพันธ์ระหว่างแบบสอบกับเกณฑ์ (Gullikson 1950: 88)

ความตรง หมายถึงคุณสมบัติของแบบสอบที่สามารถวัดในสิ่งที่ต้องการวัด (Ebel 1965: 213)

ความตรงของแบบสอบ หมายถึงขอบเขตของผลการประเมินที่สอดคล้องกับ
จุดประสงค์นั้น (Gronlund 1976: 79)

ความตรงของแบบสอบ หมายถึงความสามารถที่แบบสอบนั้นบรรลุถึงจุดมุ่งหมาย
อย่างใดอย่างหนึ่ง (Mehrens Lehmann 1973: 109)

จากความหมายของความตรงดังกล่าว สรุปได้ว่า ความตรงของแบบสอบหมายถึง
ถึงคุณสมบัติของแบบสอบที่สามารถวัดในสิ่งที่แบบสอบนั้นต้องการวัดได้ถูกต้องแน่นอน

ชนิดของความตรง

ความตรงของแบบสอบมีหลายชนิด ซึ่งแต่ละชนิดแบ่งตามธรรมชาติและจุดมุ่ง
หมายของการวัด แต่โดยทั่วไปแล้วแบ่งได้เป็น 3 ชนิด (Brown 1976: 99) คือ

1. ความตรงตามเนื้อหา (Content Validity)
2. ความตรงตามเกณฑ์สัมพันธ์ (Criterion Related Validity)
3. ความตรงตามคุณลักษณะ (Construct Validity)

การวิจัยครั้งนี้เกี่ยวข้องกับความตรงตามสภาพ ซึ่งเป็นแบบหนึ่งของความตรง
ตามเกณฑ์สัมพันธ์ มีรายละเอียดดังนี้

ความตรงตามสภาพ (Concurrent Validity)

ความตรงตามสภาพหมายถึงความสามารถของแบบสอบที่จะบ่งบอกสิ่งที่วัดได้ ถูกต้องตามสภาพที่แท้จริงขณะนั้น (บุญเชิด ภิญโญอนันตพงษ์ 2520: 295) โดยอาศัยความสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบกับคะแนนเกณฑ์สัมพันธ์ ซึ่งกำหนดขึ้นในขณะนั้น เช่น นำแบบสอบไปวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ถ้าต้องการทราบว่าแบบสอบนั้นสามารถอธิบายผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ของนักเรียนได้ดีเพียงใด หรือแบบสอบนั้นมีความตรงตามสภาพหรือไม่ ก็นำคะแนนที่ได้จากการสอบนั้นไปหาความสัมพันธ์กับคะแนนเกณฑ์ที่กำหนดขึ้น เช่น อาจจะเป็นคะแนนผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ในภาคเรียนก่อน ถ้ามีความสัมพันธ์กันสูงก็แสดงว่าแบบสอบนั้นมีความตรงตามสภาพ นอกจากนี้ยังสามารถหาความตรงตามสภาพของแบบสอบได้โดยการเปรียบเทียบกลุ่มที่แบ่งขึ้นตามเกณฑ์ (Brown 1976: 116) โดยดูว่าคะแนนเฉลี่ยที่ได้จากแบบสอบของกลุ่มแตกต่างกันหรือไม่ ถ้าแตกต่างกันก็แสดงว่าแบบสอบนั้นมีความตรงตามสภาพ

องค์ประกอบที่มีอิทธิพลต่อค่าความตรงของแบบสอบ

กรอนสันด์ (Gronlund 1976: 98-102) ได้กล่าวถึงองค์ประกอบที่มีอิทธิพลต่อค่าความตรงของแบบสอบไว้ดังนี้

1. องค์ประกอบภายในของแบบสอบได้แก่
 - 1.1 ความไม่ชัดเจนของคำชี้แจง คำชี้แจงที่ไม่ชัดเจนเกี่ยวกับวิธีตอบข้อสอบ การบันทึกคำตอบ เป็นต้น ทำให้ค่าความตรงลดลง
 - 1.2 โครงสร้างของคำถามและประโยคยากเกินไป ทำให้ให้นักเรียนสับสน แทนที่จะวัดได้ตรงกับวัตถุประสงค์ที่ต้องการ ก็จะกลายเป็นการวัดความสามารถทางภาษา
 - 1.3 ระดับความยากที่ไม่เหมาะสมของข้อสอบทำให้ไม่สามารถจำแนกนักเรียนได้ จะทำให้ความตรงลดลง
 - 1.4 ข้อสอบที่สร้างไม่ดีบางข้อเป็นการแนะนำคำตอบให้นักเรียน
 - 1.5 ความกำกวมของภาษาที่ใช้ในแบบสอบทำให้ให้นักเรียนสับสน

- 1.6 ความไม่เหมาะสมของแบบสอบที่จะวัดสิ่งที่ต้องการวัด
- 1.7 แบบสอบสั้นเกินไปไม่สามารถวัดได้ครอบคลุมพฤติกรรมที่ต้องการวัด
- 1.8 การเรียงข้อสอบไม่เหมาะสม ข้อสอบที่เรียงจากยากไปหา่ง่ายทำให้นักเรียนใช้เวลาคิดหาคำตอบนาน อาจจะไม่มีความเข้าใจข้อสอบที่อยู่ที่ตอนหลัง การเรียงข้อสอบแบบนี้มีผลต่อความตรงโดยที่มีผลต่อแรงจูงใจในการทำแบบสอบของนักเรียน
- 1.9 การเรียงคำตอบอย่างมีระบบจะช่วยให้เด็กเรียนเดาคำตอบได้ง่าย
2. เนื้อหาและวิธีดำเนินการสอน การสอนในเนื้อหาที่ตรงกับเนื้อหาที่ต้องการวัดในข้อสอบจะทำให้เครื่องมือขาดความตรงในการวัด ขบวนการทางสมองที่ซับซ้อนกว่า
3. การบริหารแบบสอบและการให้คะแนน เช่น เวลาในการสอบน้อยเกินไป การชี้แนะนักเรียนบางคนในการตอบ การให้คะแนนที่ไม่คงที่ เหล่านี้ทำให้ความตรงของแบบสอบลดลงได้
4. องค์ประกอบในการตอบแบบสอบของนักเรียนเอง รูปแบบการตอบแบบสอบของนักเรียนแต่ละคนแตกต่างกันออกไป ซึ่งทำให้ผลการสอบไม่ตรงกับความสามารถที่แท้จริง ตลอดจนความวิตกกังวล ความเปลี่ยนแปลงทางอารมณ์ เป็นต้น ทำให้ค่าความตรงของแบบสอบลดลง
5. ธรรมชาติของกลุ่มและเกณฑ์ที่ใช้ ตามปกติความตรงของแบบสอบจะมีค่าสูงเฉพาะในบางกลุ่ม เนื่องจากในการสอบวัดนั้นมียุทธวิธีผลจากองค์ประกอบอื่น เช่น อายุ เพศ ระดับความสามารถ ภูมิหลังทางการศึกษา และภูมิหลังทางวัฒนธรรม เป็นต้น ดังนั้นในการหาความตรงควรคำนึงถึงกลุ่มและเกณฑ์ที่ใช้ให้เหมาะสมด้วย

อำนาจจำแนกของแบบสอบ

อำนาจจำแนก หมายถึง ประสิทธิภาพของข้อคำถามในการแบ่งเด็กออกเป็นกลุ่มเก่ง กลุ่มอ่อน (สุทัศน์ สักขะวิลาศ 2526: 14) อำนาจจำแนกมีค่าตั้งแต่ -1.00 ถึง 1.00 ข้อสอบที่มีค่าอำนาจจำแนกเข้าใกล้ -1.00 หรือ 1.00 แสดงว่ามีอำนาจจำแนกสูง แต่ข้อสอบที่มีค่าตรงตามวัตถุประสงค์ของการสอบคือข้อสอบที่มีค่าอำนาจ

จำแนกเป็นบวก ซึ่งหมายความว่านักเรียนในกลุ่มสูงตอบข้อสอบได้ถูกต้องมากกว่าในกลุ่มต่ำ และแสดงว่าข้อสอบสามารถจำแนกไปในทางเดียวกับคะแนนรวมที่ได้จากการสอบ และข้อสอบที่มีค่าอำนาจจำแนกเป็น 0 ก็ถือว่าจำแนกความแตกต่างระหว่างคน 2 กลุ่มไม่ได้เลย

ในการประเมินคุณภาพของข้อสอบโดยอาศัยค่าอำนาจจำแนกควรถือเกณฑ์ดังนี้ (Ebel 1965 : 364)

ค่าอำนาจจำแนก	คุณภาพของข้อสอบ
.40 ขึ้นไป	ดีมาก
.30 ถึง .39	ดี
.20 ถึง .29	ใช้ได้
ต่ำกว่า .20	ไม่ดีควรตัดทิ้ง

อำนาจจำแนกของข้อสอบเป็นสิ่งสำคัญอย่างหนึ่งที่ต้องคำนึงถึงในการประเมินคุณภาพของข้อสอบควบคู่ไปกับความยากของข้อสอบ เนื่องจากข้อสอบที่มีความยากเท่ากันยังแตกต่างกันในด้านค่าอำนาจจำแนกได้ด้วย ข้อสอบที่มีความยาก .50 อาจจะทำให้ค่าอำนาจจำแนกสูงสุด หรืออาจจะไม่มีเลยก็ได้ ฉะนั้นข้อสอบที่มีความยาก .50 มีโอกาสทำให้ค่าความเที่ยงของแบบสอบต่ำได้เมื่อข้อสอบนั้นมีค่าอำนาจจำแนกเป็น 0 แต่ข้อสอบที่มีค่าอำนาจจำแนกสูงจะมีผลทำให้ค่าความเที่ยงของแบบสอบสูงด้วยอย่างแน่นอน เนื่องจากค่าอำนาจจำแนกของข้อสอบเกี่ยวข้องกับส่วนเบี่ยงเบนมาตรฐานของคะแนนซึ่งเป็นองค์ประกอบสำคัญที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบ (Ebel 1965 : 302)

องค์ประกอบที่มีอิทธิพลต่อค่าอำนาจจำแนกของแบบสอบ

องค์ประกอบที่สำคัญที่มีอิทธิพลต่อค่าอำนาจจำแนกของแบบสอบได้แก่

1. คุณภาพของตัวลวงและจำนวนตัวเลือก ข้อสอบที่สร้างตัวลวงได้ก็มีความใกล้เคียงกับตัวเลือกที่เป็นคำตอบถูกหลายๆตัว จะทำให้จำนวนตัวเลือกเพิ่มขึ้น ซึ่งจะทำ

ให้อำนาจจำแนกของข้อสอบสูงขึ้นด้วย (Ebel 1965: 165) เนื่องจากผู้ที่มีความรู้จริงเท่านั้นจึงจะสามารถทำข้อสอบได้ถูกต้อง ส่วนผู้ที่ไม่มีความรู้จริงจะเลือกตอบตัวกลางที่ใกล้เคียงกับคำตอบถูก

2. ความยากของข้อสอบ การจำแนกความแตกต่างของผลสัมฤทธิ์ระหว่างกลุ่มขึ้นอยู่กับความยากของข้อสอบ (Gronlund 1976: 154) กล่าวคือ ข้อสอบที่มีความยาก .50 มีโอกาสทำให้อำนาจจำแนกของข้อสอบสูงสุดได้ เมื่อนักเรียนในกลุ่มสูงทุกคนตอบข้อสอบได้ถูกต้อง ในขณะที่นักเรียนในกลุ่มต่ำทั้งหมดตอบข้อสอบผิด และหากข้อสอบมีค่าความยากสูงหรือต่ำกว่า .50 ค่าอำนาจจำแนกของข้อสอบจะต่ำลง (Gronlund 1976: 270) ซึ่งหมายความว่าข้อสอบที่ยากหรือง่ายเกินไปจะเป็นข้อสอบที่มีอำนาจจำแนกต่ำนั่นเอง

ความคลาดเคลื่อนมาตรฐานของการวัด (Standard Error of Measurement)

ความคลาดเคลื่อนมาตรฐานของการวัด เป็นสถิติค่าหนึ่งที่ใช้ในการประมาณค่าคะแนนจริงของบุคคลจากคะแนนที่ได้จากการวัด (obtained score) การคำนวณค่าความคลาดเคลื่อนมาตรฐานของการวัดต้องอาศัยข้อตกลงเบื้องต้นที่กำหนดว่า คะแนนความคลาดเคลื่อนเป็นคะแนนที่เกิดขึ้นโดยสุ่ม (random scores) มีการแจกแจงเป็นโค้งปกติ โดยมีค่าเฉลี่ยเท่ากับศูนย์ เป็นอิสระจากคะแนนจริง การกระจายของคะแนนความคลาดเคลื่อนของแบบสอบคู่ขนานในลักษณะเท่ากัน ซึ่งสมการที่ใช้ในการคำนวณหาพัฒนาจากสมการของการหาความเที่ยงของการวัดได้ดังนี้

$$r_{tt} = 1 - s_e^2 / s_t^2$$

$$s_e^2 = s_t^2 (1 - r_{tt})$$

$$s_e = s_t \sqrt{1 - r_{tt}}$$

เมื่อ s_e แทนความคลาดเคลื่อนมาตรฐานของการวัด

s_t แทนส่วนเบี่ยงเบนมาตรฐานของคะแนนที่ได้จากการวัด

r_{tt} แทนค่าความเที่ยงของเครื่องมือวัด

สมมุติว่าค่าความเที่ยงของแบบสอบเท่ากับ 0.84 ส่วนเบี่ยงเบนมาตรฐานของคะแนนที่ได้จากการวัดเท่ากับ 10 เมื่อแทนค่าในสมการจะได้ค่าความคลาดเคลื่อนมาตรฐานในการวัดเท่ากับ 4.0 ทำให้สามารถแปลความหมายในลักษณะของการประมาณคะแนนจริงของบุคคลในอุดมคติได้คือ

คะแนนจริง 40 จะมีการกระจายของคะแนนจากการวัดอยู่ในอันตรภาคความเชื่อมั่นดังนี้

40 ± 4 ค่ายระดับความเชื่อมั่น 68 %

40 ± 7.84 ค่ายระดับความเชื่อมั่น 95 %

อย่างไรก็ตามในทางปฏิบัติไม่สามารถรู้คะแนนจริงของบุคคลได้ จึงทำได้เพียงการประมาณว่าคะแนนจริงของคะแนนที่ได้จากการวัดจะอยู่ในอันตรภาคระหว่างค่าใดถึงค่าใด ค่ายระดับความเชื่อมั่นเท่าใด จากตัวอย่างข้างต้นนักเรียนที่สอบได้คะแนน 50 คะแนน จะหมายความว่า คะแนนจริงของเขา จะมีค่าอยู่ระหว่าง $50 \pm 1.96(4.0)$ หรือตั้งแต่ 42.16 ถึง 57.84 ค่ายระดับความเชื่อมั่น 95 %

ตอนที่ 2 วรรณคดีและงานวิจัยที่เกี่ยวกับการใช้คำสั่งและวิธีการให้คะแนนที่ต่างกัน

แบบสอบชนิดเลือกตอบเป็นแบบสอบที่ผู้ตอบจะต้องเลือกคำตอบที่ถูกจากคำตอบหลายคำตอบที่ผู้ออกข้อสอบกำหนดให้ ซึ่งเมื่อเปรียบเทียบกับข้อสอบอัตนัยแล้ว ข้อสอบชนิดเลือกตอบสามารถวัดได้ครอบคลุมเนื้อหามากกว่า แต่ข้อสอบชนิดเลือกตอบยังมีปัญหาอยู่ก็คือ เปิดโอกาสให้นักเรียนเดาคำตอบจากที่กำหนดให้ได้ ซึ่งมีโอกาสจะเดาถูกอยู่บ้าง จึงทำให้นักวัดผลการศึกษานานท่านพยายามที่จะหาวิธีกำจัดการเดานี้ และได้ทำการวิจัยเพื่อหาวิธีที่จะแก้การเดา ซึ่งมีผู้ทำการศึกษาไว้ดังนี้

กูมบ์ส (Coombs) เสนอให้มีวิธีตอบโดยให้เลือกคำตอบที่ผิดแทนการเลือกคำตอบที่ถูก กูมบ์ส ให้ความเห็นว่า การเลือกคำตอบที่ผิดซึ่งมีอยู่หลายตัวเลือกในข้อคำถามหนึ่งๆนั้นผู้ตอบต้องใช้ส่วนของความรู้ย่อมมาพิจารณามากกว่าการเลือกคำตอบถูกเพียง

คำตอบเดียว ความรู้ที่วัดได้จะสมบูรณ์มากกว่า เพราะผู้ตอบรู้ว่าสิ่งใดผิดสิ่งใดถูก วิธีนี้สามารถวัดส่วนของความรู้ได้ละเอียดกว่าวิธีธรรมดา สำหรับการตรวจให้คะแนนนั้น กูมบ์ส์ได้ให้คะแนนดังนี้ ถ้านักเรียนทำเครื่องหมายตรงกับตัวเลือกที่ผิดจริง ให้คะแนนตัวเลือกละ 1 คะแนน ถ้านักเรียนทำเครื่องหมายตรงตัวเลือกที่ถูก ตัวเลือกนั้นจะได้ $1-k$ คะแนน (k คือจำนวนตัวเลือกทั้งหมดในแต่ละข้อ) เพราะฉะนั้นถ้ามี 5 ตัวเลือกคะแนนที่ได้จะอยู่ระหว่าง -4 ถึง 4 ซึ่งกูมบ์ส์ถือว่าเป็นวิธีที่ดีกว่าการให้คะแนนแบบ 0-1 (โกวิท ประมวลพฤษ์ และสมศักดิ์ สันธูระเวชญ์ 2523: 220)

ในปี ค.ศ. 1956 กูมบ์ส์ มิลฮอลแลนด์ และโวมเมอร์ (Coombs, Milholland and Womer 1956: 13-37) ได้ทดลองศึกษาเปรียบเทียบวิธีที่เสนอแนะกับวิธีธรรมดาโดยใช้แบบสอบ 3 ฉบับ ฉบับ 40 ข้อ คือ

1. แบบสอบที่เกี่ยวกับคำศัพท์
2. แบบสอบที่เกี่ยวกับความรู้ของนักขับรถ
3. แบบสอบที่เกี่ยวกับการใช้สายตาในค่านิยมสัมพันธ์

การตรวจให้คะแนนมี 3 วิธีด้วยกันคือ วิธี 0-1 วิธีที่กูมบ์ส์เสนอแนะไว้ และวิธี 0-1 กับวิธีของกูมบ์ส์รวมกัน กลุ่มตัวอย่างที่ใช้ในการศึกษาคั้งนี้เป็นนักเรียนชั้นปีที่ 3 และปีที่ 4 ของแจคสัน ไฮสกูล (Jackson High School) รัฐมิชิแกน จำนวน 855 คน แบ่งเป็น 3 กลุ่ม แต่ละกลุ่มมีวิธีการตอบและการตรวจให้คะแนนที่แตกต่างกันไปกลุ่มละวิธี ผลปรากฏว่า

1. ค่าความเที่ยงของแบบสอบจากวิธีการตอบและการตรวจให้คะแนนตามที่กูมบ์ส์เสนอแนะสูงกว่าวิธี 0-1
2. ค่าความตรงของแบบสอบแตกต่างกันอย่างไม่มีนัยสำคัญ
3. มาตรฐานความมั่นใจในการทำแบบสอบ วิธีที่กูมบ์ส์เสนอแนะมีมากกว่าวิธี 0-1 สามารถป้องกันการเดาได้มากกว่า
4. ทัศนคติของกลุ่มตัวอย่าง ชอบวิธีที่กูมบ์ส์เสนอแนะมากกว่า โดยให้เหตุผลว่าเป็นวิธีที่ยุติธรรม และง่ายกว่าวิธี 0-1

วิธีที่ดูมั่วส์เสนอมือข้อดีคือ ทำให้พิสัยของคะแนนกว้างขึ้น สามารถวัดความรู้ของนักเรียนได้ละเอียดกว่าวิธี 0-1 และยังทราบความรู้ที่ผิดของนักเรียนได้อีกด้วย แต่ยังมีข้อบกพร่องในเรื่องโอกาสของการได้คะแนนจากการรู้ตัวเลือกที่ผิดและถูกต้องต่างกัน เช่น ในข้อสอบที่มี 5 ตัวเลือก นักเรียนคนหนึ่งรู้ว่าตัวเลือกใดผิด ตัวเลือกใดถูก สามารถทำเครื่องหมายลงบนตัวเลือกที่ผิดได้อย่างถูกต้อง และได้คะแนน 4 คะแนน ส่วนนักเรียนอีกคนหนึ่งไม่รู้ว่าตัวเลือกใดผิด แต่รู้ว่าตัวเลือกใดถูก เขาจึงทำเครื่องหมายลงบนตัวเลือกอื่นๆที่ไม่ใช่ตัวเลือกที่ถูก เขาได้คะแนน 4 คะแนนเท่ากับนักเรียนคนแรก ทั้งๆที่ความจริงแล้วนักเรียนทั้งสองคนมีความรู้ไม่เท่ากัน

ต่อมา เดวิส และฟิฟเฟอร์ (Davis Fifer 1959: 159-169) ได้ทดลองศึกษาเกี่ยวกับค่าความเที่ยงและค่าความตรงของแบบสอบวัดความถนัดและแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ โดยกำหนดค่าน้ำหนักคะแนนของแต่ละตัวเลือกไม่เท่ากัน เปรียบเทียบกับวิธี 0-1 เดวิส และฟิฟเฟอร์ มีความคิดว่า การคิดใจหยาบคณิตศาสตร์นั้นตัวเลือกแต่ละตัวเลือกมีความยากง่ายไม่เท่ากัน และใช้สมรรถภาพสมองในการคิดไม่เท่ากันด้วย จึงกำหนดคะแนนของแต่ละตัวเลือกไม่เท่ากัน กลุ่มตัวอย่างที่ใช้เป็นนักเรียนการบินที่ Lackland Air Force Base จำนวน 65 คน ผลการศึกษพบว่า

1. การตรวจให้คะแนนโดยกำหนดคะแนนของแต่ละตัวเลือกไม่เท่ากันนั้นให้ค่าความเที่ยงสูงกว่าการตรวจโดยวิธี 0-1 อย่างมีนัยสำคัญทางสถิติ
2. การตรวจให้คะแนนโดยกำหนดคะแนนของแต่ละตัวเลือกไม่เท่ากันนั้นให้ค่าความตรงสูงกว่าการตรวจให้คะแนนโดยวิธี 0-1 แต่ไม่มีนัยสำคัญทางสถิติ

จะเห็นว่า การกำหนดให้คะแนนแต่ละตัวเลือกไม่เท่ากันนั้นเป็นวิธีหนึ่งที่จะช่วยให้คะแนนที่ได้จากแบบสอบมีค่าความเที่ยงสูงขึ้น แต่มีข้อยุ่งยากของวิธีนี้ก็คือ การกำหนดค่าน้ำหนักให้แต่ละตัวเลือก ข้อสอบที่จะนำมากำหนดให้ค่าน้ำหนักคะแนนของแต่ละตัวเลือกให้แตกต่างกันนั้นต้องเป็นข้อสอบที่สร้างขึ้นโดยผู้ชำนาญ และใช้ผู้เชี่ยวชาญหลายท่านใน

การกำหนดค่าน้ำหนักคะแนนของแต่ละตัวเลือก แล้วนำมาหาค่าเฉลี่ยว่าจะให้คะแนนแต่ละตัวเลือกเป็นเท่าไร ซึ่งเป็นวิธีที่ยู่ยากในทางปฏิบัติ ไม่เหมาะที่จะนำมาใช้ในแบบสอบที่ครูสร้างขึ้น

แพทเนอิก และเทราป (Patnaic Traub 1973: 281-285) ได้ทดลองศึกษาวิธีตอบและการตรวจให้คะแนนที่คล้ายกับวิธีของ เดวิส และฟิฟเฟอร์ โดยกำหนดคะแนนของตัวเลือกถูกในแต่ละข้อให้ต่างกัน เปรียบเทียบกับวิธีตอบและให้คะแนนวิธี 0-1 โดยใช้แบบสอบ Dominion Group Test of Learning Capacity (DLC) จำนวน 60 ข้อให้เวลาในการทำแบบสอบ 20 นาที ทดสอบกับกลุ่มตัวอย่างที่เป็นนักเรียนเกรด 10-13 จำนวน 1054 คน ในการกำหนดค่าน้ำหนักคะแนนของตัวเลือกที่ถูกใช้การตัดสินใจของครูและผู้เชี่ยวชาญรวม 61 คน แล้วหาค่าเฉลี่ยออกมาเป็นคะแนนในแต่ละข้อ ผลการศึกษาครั้งนี้พบว่า สอดคล้องกับผลการวิจัยของ เดวิส กล่าวคือ ค่าความเที่ยงของแบบสอบที่คำนวณหาโดยวิธีแบ่งครึ่งข้อสอบ และตรวจให้คะแนนตามวิธี 0-1 เป็น 0.881 และแบบให้คะแนนข้อถูกแต่ละข้อไม่เท่ากันเป็น 0.915 ส่วนค่าความตรงเชิงพยากรณ์ ซึ่งคำนวณหาจากค่าสหสัมพันธ์ระหว่างคะแนนของแบบสอบกับคะแนนผลสัมฤทธิ์ทางการเรียนเฉลี่ยพบว่า วิธีตรวจแบบให้ค่าน้ำหนักคะแนนแต่ละข้อไม่เท่ากันมีความตรงน้อยกว่าวิธีตอบและตรวจให้คะแนนตามวิธี 0-1 จะเห็นว่า การกำหนดค่าน้ำหนักคะแนนแต่ละข้อไม่เท่ากันจะช่วยทำให้ค่าความเที่ยงของแบบสอบสูงขึ้นกว่าการตรวจให้คะแนนตามวิธี 0-1 แต่เป็นการยากที่จะนำมาปฏิบัติจริง เพราะต้องใช้ผู้เชี่ยวชาญจำนวนมากในการกำหนดค่าน้ำหนักคะแนน จึงมีผู้คิดวิธีตอบและตรวจให้คะแนนวิธีอื่นอีก โดยการให้ผู้ตอบบอกความมั่นใจในการตอบข้อคำถามเหล่านั้นด้วย ผู้ที่ทำการศึกษาคือ ฟิฟ และบรันซา

ฟิฟ และบรันซา (Pugh Brunza 1975: 73-78) ได้เสนอความคิดให้มีวิธีตอบโดยบอกความมั่นใจในการตอบและนําน้ำหนักความมั่นใจในการตอบมาเป็นส่วนในการให้คะแนนด้วย ฟิฟ และบรันซา ได้ศึกษาหาค่าความเที่ยงของแบบสอบที่มีวิธีตอบและการตรวจให้คะแนน 2 วิธีคือ วิธีตอบแบบธรรมดา และวิธีตอบโดยให้ผู้ตอบบอกระดับ

ความมั่นใจในการตอบข้อความเหล่านั้นด้วย โดยใช้แบบสอบถามชนิดเลือกตอบ เป็นแบบสอบถามวัดด้านสติปัญญาชื่อ I.R.E. Intelligence Scale จำนวน 48 ข้อแบ่งเป็น 2 ส่วนๆละ 24 ข้อ ทดลองกับกลุ่มตัวอย่างที่เป็นนักศึกษามหาวิทยาลัยอินเดียน่า จำนวน 84 คน โดยแบ่งเป็น 3 กลุ่มๆละ 28 คนให้แต่ละกลุ่มได้รับแบบสอบถามที่แตกต่างกันคือ กลุ่มแรกได้รับแบบสอบถามชุด เอ ซึ่งมีวิธีตอบแบบธรรมดาทั้งสองส่วน กลุ่มที่สองได้รับแบบสอบถามชุด บี ซึ่งมีวิธีตอบแบบธรรมดาในส่วนที่หนึ่งและวิธีตอบโดยบอกความมั่นใจในการตอบในส่วนที่สอง กลุ่มที่สามได้รับแบบสอบถามชุด ซี ซึ่งมีวิธีตอบโดยบอกความมั่นใจในการตอบทั้งสองส่วน ผลการศึกษาพบว่า ค่าความเที่ยงของแบบสอบถามซึ่งหาโดยวิธีวิเคราะห์ความแปรปรวนของแบบสอบถามที่มีวิธีตอบโดยบอกความมั่นใจในการตอบสูงกว่าของแบบสอบถามที่ตอบแบบธรรมดา ถือเป็น 0.85 และ 0.57 ตามลำดับ

เกี่ยวกับเรื่องการตอบโดยให้บอกระดับความมั่นใจนี้ อาบู ซาอ์ฟ (Abu-Sayf 1975: 160-162) ได้ทำการศึกษาโดยให้นักเรียนระดับมัธยม 152 คน ตอบแบบสอบถามชนิดเลือกตอบซึ่งมี 4 ตัวเลือก วิชาภาษาอังกฤษ จำนวน 60 ข้อ ทดลองโดยให้นักเรียนเลือกตอบตัวเลือกที่ถูกที่สุดเพียงตัวเลือกเดียวและบอกระดับความมั่นใจในการตอบว่าตอบด้วยความมั่นใจ ไม่ค่อยมั่นใจ หรือ เคา แล้วเปรียบเทียบเปอร์เซ็นต์ในการตอบถูกหรือผิด ผลปรากฏว่า จากคำตอบทั้งหมดที่นักเรียนบอกว่าตอบด้วยความมั่นใจ 3491 ข้อ มีนักเรียนตอบผิด 846 ข้อ หรือ 24.23 % คำตอบที่นักเรียนตอบโดยไม่ค่อยมั่นใจ 2648 ข้อ มีข้อผิด 1683 ข้อ หรือ 61.92 % และคำตอบที่นักเรียนบอกว่าตอบโดยการเคา 2782 ข้อ มีนักเรียนตอบผิด 2045 ข้อหรือ 73.31 % จะเห็นว่าการตอบโดยการเคานั้นมีโอกาสตอบถูกน้อยมาก

ต่อมา อาบู ซาอ์ฟ และไดมอนด์ (Abu-Sayf Diamond 1976: 62-65) ทำการศึกษาเกี่ยวกับค่าความเที่ยงและความตรงของแบบสอบถาม เมื่อให้นักเรียนตอบแบบสอบถามชนิดเลือกตอบโดยบอกความมั่นใจในการตอบจากข้อมูลที่ อาบู ซาอ์ฟ ได้ศึกษาไว้แล้ว จากกลุ่มตัวอย่าง 152 คน ใช้แบบสอบถามภาษาอังกฤษชนิดเลือกตอบ 4 ตัวเลือก จำนวน

60 ข้อ โดยการตรวจให้คะแนน 3 วิธี คือ ให้คะแนนเฉพาะข้อถูกและตอบอย่างมั่นใจ ให้คะแนนข้อที่ถูกและไม่ได้คำตอบ และให้คะแนนข้อที่ตอบถูกโดยไม่คำนึงว่ามีความมั่นใจ ในการตอบหรือไม่เพียงใด แล้วหาค่าความเที่ยงของแบบสอบโดยใช้ α -Coefficient ผลปรากฏว่า การให้คะแนนเฉพาะข้อที่ตอบถูกและมีความมั่นใจได้ค่าความเที่ยงสูงสุด คือ 0.898 ส่วนการให้คะแนนข้อที่ตอบถูกโดยไม่ได้คำตอบได้ค่าความเที่ยงสูงกว่าการให้คะแนนทุกข้อที่ตอบถูก คือได้ค่าความเที่ยง 0.871 และ 0.774 ตามลำดับ ส่วนค่าความตรงนั้น วิธีการให้คะแนนข้อที่ตอบถูกและไม่ได้คำตอบมีค่าความตรงสูงสุด คือ 0.260 ส่วนการให้คะแนนเฉพาะที่ตอบถูกและมีความมั่นใจได้ค่าความตรงสูงกว่าการให้คะแนนทุกข้อที่ตอบถูก คือ 0.236 และ 0.219 ตามลำดับ

กิบบอนส์และคณะ (Gibbons and Others 1979: 261) ได้เสนอแนะวิธีตอบและวิธีการตรวจให้คะแนนในแต่ละข้อ โดยให้ผู้ตอบเลือกตอบตัวเลือกที่ถูกได้ตั้งแต่ 1 ถึง $k - 1$ ตัวจากความมั่นใจ เมื่อ k คือจำนวนตัวเลือกทั้งหมดในแต่ละข้อ กล่าวคือ ถ้ามั่นใจว่าตัวเลือกใดถูกแน่ก็ให้ตอบเพียงตัวเลือกเดียว ถ้าไม่มั่นใจหลายตัวเลือกก็ให้ตอบทุกตัวเลือกที่ไม่มั่นใจ ส่วนวิธีการตรวจให้คะแนน ยึดถือหลักว่า คะแนนคาดหวัง (Expected Score) ที่ได้จากการเดาสุ่มจะต้องเป็นศูนย์ ในกรณีที่ข้อสอบมี 5 ตัวเลือก และผู้ตอบตอบอย่างสุ่ม 2 ตัวเลือก ความน่าจะเป็นที่เขาจะตอบถูกเป็น $2/5$ และความน่าจะเป็นที่เขาจะตอบผิดเป็น $3/5$ ถ้าต้องการให้คะแนนคาดหวังเป็นศูนย์ เมื่อเขาตอบถูกจะได้คะแนน 3 คะแนน แต่ถ้าผิดจะได้คะแนน -2 คะแนน

จะเห็นว่าการตอบและวิธีการให้คะแนนตามวิธีของกิบบอนส์และคณะ สามารถวัดคนที่มีความรู้น้อยๆ ไปจนถึงคนที่มีความรู้จริงในเรื่องนั้นๆ ทั้งยังสามารถแก้การเดาสุ่มได้อีกด้วย และคนที่มีความสามารถเท่ากันมีโอกาสได้คะแนนเท่ากัน

การศึกษาเกี่ยวกับเรื่องการตอบและวิธีการตรวจให้คะแนนในประเทศไทย ได้มีนักวัดผลการศึกษาหลายท่านได้ทำการศึกษาไว้ดังนี้

นันทนา เมื่อน้อง (2510: 30) ได้ให้ความเห็นสนับสนุนการตอบโดยให้บอกความมั่นใจในการตอบ โดยกล่าวว่า การให้นักเรียนตอบแบบสอบชนิดเลือกตอบ เมื่อนักเรียนตอบถูกผู้ตรวจไม่สามารถทราบได้เลยว่านักเรียนตอบถูกเพราะมีความรู้จริงหรือตอบถูกเพราะการเดา ฉะนั้นเพื่อที่จะวัดความรู้จริงของนักเรียนให้แน่นอนและป้องกันการเดา จึงเสนอให้มีวิธีตอบโดยบอกความมั่นใจในการตอบให้แตกต่างไปจากวิธีตอบแบบธรรมดาที่ใช้กันอยู่ในปัจจุบันนี้ กล่าวคือ นอกจากผู้ตอบจะต้องเลือกตัวเลือกเป็นคำตอบแล้วจะต้องตอบด้วยว่า เลือกตอบด้วยความมั่นใจว่าถูกจริง หรือไม่มั่นใจในคำตอบเท่าใดนัก กระดาษคำตอบจะต้องเปลี่ยนใหม่ เพื่อให้การตอบแต่ละตัวเลือกมีช่องให้ผู้ตอบบอกระดับความมั่นใจในการตอบด้วย การตรวจให้คะแนนก็ต้องให้คะแนนแตกต่างจากธรรมดาโดยให้น้ำหนักความมั่นใจมาเกี่ยวข้องด้วย เช่น ให้คะแนนเป็น 5, 4, 3, 2, 1 เมื่อตอบถูกและบอกว่ามั่นใจที่สุด มั่นใจมาก ก่อนข้างมั่นใจ ไม่ค่อยมั่นใจ และไม่มั่นใจเลย ตามลำดับ ถ้านักเรียนตอบผิด ก็จะได้คะแนนเป็น -5, -4, -3, -2, -1 เมื่อบอกว่า มั่นใจที่สุด มั่นใจมาก ก่อนข้างมั่นใจ ไม่ค่อยมั่นใจ และไม่มั่นใจเลย ตามลำดับ หรือจะให้คะแนนเป็นอย่างอื่นที่แตกต่างไปจากนี้ก็ได้ แต่ต้องให้น้ำหนักของความมั่นใจมาเกี่ยวข้องด้วย

อนันต์ ศรีโสภา (2516: 13-19) ได้เสนอวิธีการตอบวิธีใหม่โดยให้ตอบทุกตัวเลือก การตอบวิธีนี้จะต้องพิจารณาตอบทุกตัวเลือกว่า ตัวเลือกใดถูก ตัวเลือกใดผิด ซึ่งจะทำให้เป็นการวัดความรู้ในข้อคำถามนั้นๆ ได้สมบูรณ์ยิ่งขึ้นกว่าการเลือกคำตอบที่ถูกเพียงคำตอบเดียว และยังสามารถทราบความรู้ที่ผิดของนักเรียนได้ดีกว่าวิธีตอบแบบธรรมดาอีกด้วย การให้คะแนนก็ให้เป็นรายตัวเลือก ถ้านักเรียนทำเครื่องหมายใดตรงกับสภาพความเป็นจริงว่าตัวเลือกนั้นถูกหรือผิด จะได้คะแนนตัวเลือกละ 1 คะแนน ถ้าทำเครื่องหมายตรงข้ามกับความเป็นจริง จะได้คะแนนตัวเลือกละ -1 คะแนน ซึ่งวิธีนี้จะทำให้คะแนนที่ได้จากการวัดถูกต้องและเหมาะสมกับความรู้ของนักเรียนมากขึ้น

อรรณพ ศักดิ์เจริญรัตน์ (2517: 23-26) ได้ศึกษาเปรียบเทียบค่าความเที่ยงและเปอร์เซ็นต์การเดาของแบบสอบชนิดเลือกตอบที่มีวิธีการตอบและการตรวจให้คะแนน

คามวิธี 0-1 คามวิธีของกูมบส์ และคามวิธีของอนันต์ ศรีโสภา การศึกษาครั้งนี้ ใช้แบบ
สอบวิชาภาษาไทย ระดับประกาศนียบัตรวิชาการศึกษา ทดลองกับนักศึกษาจำนวน 214 คน
แบ่งออกเป็น 3 กลุ่ม แต่ละกลุ่มได้รับแบบสอบที่มีวิธีการตอบและการให้คะแนนแตกต่างกัน
3 วิธีดังกล่าว ผลการศึกษาครั้งนี้พบว่า ค่าความเที่ยงของแบบสอบที่มีวิธีตอบและการตรวจ
ให้คะแนนตามวิธีของอนันต์ ศรีโสภา สูงกว่าค่าความเที่ยงของแบบสอบที่ตอบและตรวจให้
คะแนนตามวิธีของกูมบส์ และวิธีตอบแบบธรรมดาให้คะแนน 0-1 ให้ค่าความเที่ยงค่าสุด
เปอร์เซ็นต์การเดาของการตอบตามวิธีของอนันต์ ศรีโสภา ต่ำกว่าเปอร์เซ็นต์การเดา
ในการตอบตามวิธีของกูมบส์ แต่สูงกว่าเปอร์เซ็นต์การเดาในการตอบตามวิธี 0-1

กาญจนา ศิริวัฒนพงษ์ (2520: 68-72) ได้ศึกษาเปรียบเทียบค่าความเที่ยง
ค่าความตรง ค่าอำนาจจำแนก และเปอร์เซ็นต์การเดาของแบบสอบชนิดเลือกตอบที่มีวิธี
การตอบและตรวจให้คะแนนที่ต่างกัน 3 วิธีคือ ตอบแบบธรรมดา วิธีของกูมบส์ และวิธี
ของอนันต์ ศรีโสภา โดยใช้แบบสอบวิชาวิทยาศาสตร์ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยม
ศึกษาปีที่ 3 ในจังหวัดปทุมธานี จำนวน 810 คน ผลปรากฏว่า การตอบและการตรวจ
ให้คะแนนตามวิธีของอนันต์ ศรีโสภา ให้ค่าความเที่ยง และค่าความตรงของแบบสอบ
สูงกว่าวิธีให้ตอบแบบธรรมดา และวิธีของกูมบส์ วิธีธรรมดาให้ค่าความเที่ยงสูงกว่า
วิธีของกูมบส์ แต่ค่าความตรงไม่แตกต่างกัน ส่วนค่าอำนาจจำแนกของแบบสอบจากการ
ตอบและการตรวจให้คะแนนแต่ละวิธีไม่ต่างกันอย่างมีนัยสำคัญทางสถิติ ในด้านการเดา
ผลปรากฏว่า วิธีตอบและการตรวจให้คะแนนตามวิธีของอนันต์ ศรีโสภามีการเดาน้อยกว่า
วิธีตอบแบบธรรมดาและวิธีของกูมบส์ วิธีของกูมบส์ มีการเดาสูงกว่าวิธีอื่นๆ

ธีรศักดิ์ อินทรมาศย์ (2520: 64-69) ได้ศึกษาเปรียบเทียบค่าความเที่ยง
ค่าความตรง และปริมาณการเดาของแบบสอบชนิดเลือกตอบที่มีวิธีการตอบแตกต่างกัน
3 วิธีคือ วิธีตอบแบบธรรมดา วิธีตอบโดยบอกความมั่นใจในการตอบ และวิธีตอบทุก
ตัวเลือก โดยใช้แบบสอบวิชาวิทยาศาสตร์ ทดสอบกลุ่มตัวอย่างที่เป็นนักเรียนชั้นประถม
ปีที่ 7 ที่มีระดับความสามารถในวิชาวิทยาศาสตร์แตกต่างกัน 3 กลุ่มซึ่งเป็นนักเรียนของ

โรงเรียนเมืองสมุทรสงครามจำนวน 216 คน ผลการศึกษาพบว่า ค่าความเที่ยงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในกลุ่มนักเรียนที่มีความสามารถในวิชาวิทยาศาสตร์ระดับสูงและปานกลาง และแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในกลุ่มนักเรียนที่มีความสามารถในวิชาวิทยาศาสตร์ระดับต่ำ สำหรับค่าความตรงของแบบสอบที่มีวิธีตอบแตกต่างกันพบว่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติในทุกระดับความสามารถในวิชาวิทยาศาสตร์ของนักเรียน ในด้านปริมาณการเดาพบว่าแบบสอบที่ตอบแบบธรรมชาติมีปริมาณการเดาสูงสุด รองลงมาคือแบบสอบที่ตอบโดยบอกระดับความมั่นใจในการตอบ และแบบสอบที่ตอบทุกตัวเลือกมีปริมาณการเดาต่ำสุด และจากการศึกษาเรื่องปริมาณการเดาพบว่า นักเรียนที่มีความสามารถในวิชาวิทยาศาสตร์ระดับต่ำจะตอบแบบสอบโดยการเดามากกว่านักเรียนที่มีความสามารถในวิชาวิทยาศาสตร์ระดับสูงในทุกวิธีการตอบ

เพ็ญศรี สว่างเนตร (2520: 41-42) ได้ศึกษาเปรียบเทียบค่าความเที่ยงของแบบสอบชนิดเลือกตอบเมื่อใช้วิธีการให้คะแนนที่ต่างกัน 4 วิธี คือ วิธี 0-1 วิธีลดคะแนนข้อที่ตอบผิด วิธีเพิ่มคะแนนข้อที่เว้น และวิธีทดสอบความมั่นใจ โดยทดสอบนักศึกษาระดับประกาศนียบัตรวิชาชีพการศึกษาระดับชั้นสูงจำนวน 218 คน ผลปรากฏว่า ค่าความเที่ยงของแบบสอบที่ให้คะแนนทดสอบความมั่นใจสูงกว่าค่าความเที่ยงของแบบสอบที่ให้คะแนนตามวิธี 0-1 ตามวิธีลดคะแนนข้อที่ตอบผิด และตามวิธีเพิ่มคะแนนให้ข้อที่เว้น ส่วนค่าความเที่ยงระหว่างการให้คะแนนตามวิธี 0-1 ตามวิธีลดคะแนนข้อที่ตอบผิด และตามวิธีเพิ่มคะแนนข้อที่เว้น ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

สำราญ มีแจ้ง (2525: 55-61) ได้เสนอแนะวิธีการตอบและการตรวจให้คะแนนแบบสอบชนิดเลือกตอบที่มี 5 ตัวเลือก โดยให้นักเรียนตอบได้ตั้งแต่ 1 ถึง 3 ตัวเลือกตามความมั่นใจ ถ้ามั่นใจตัวเลือกใดก็ตอบตัวเลือกนั้น ถ้าไม่มั่นใจ 2 หรือ 3 ตัวเลือกก็ให้ตอบทั้ง 2 หรือ 3 ตัวเลือกนั้น ถ้าตอบตัวเลือกเดียวแล้วถูกจะให้คะแนน 4 คะแนน ถ้าตอบ 2 ตัวเลือกแล้วถูกจะให้คะแนน 2 คะแนน ถ้าตอบ 3 ตัวเลือกแล้วถูกจะให้คะแนน 1 คะแนน ทุกกรณีถ้าผิดจะให้คะแนน 0 คะแนน สำราญ มีแจ้ง ได้ศึกษาเปรียบเทียบ ค่า

ความเที่ยง ค่าความตรง และค่าอำนาจจำแนกของแบบสอบชนิดเลือกตอบที่ใช้คำสั่งและการให้คะแนนตามวิธี 0-1 ตามวิธีของกุ่มบัส ตามวิธีของอนันต์ ศรีโสภา และตามวิธีของสำราญ มีแจ้ง โดยใช้แบบสอบชนิดเลือกตอบ 5 ตัวเลือกวิชาคณิตศาสตร์ จำนวน 60 ข้อ ไปทดสอบนักเรียนชั้นมัธยมศึกษาปีที่ 4 ของโรงเรียนสุรศักดิ์มนตรี และโรงเรียนวัดน้อยอิน จำนวน 4 กลุ่มๆละ 60 คน แต่ละกลุ่มได้รับแบบสอบที่ใช้คำสั่งและการตรวจให้คะแนนกลุ่มละ 1 วิธี ผลปรากฏว่า ค่าความเที่ยงของแบบสอบที่ใช้คำสั่งและการให้คะแนนตามวิธีของ สำราญ มีแจ้ง สูงกว่าค่าความเที่ยงของแบบสอบที่ใช้คำสั่งและการให้คะแนนตามวิธี 0-1 และค่าความเที่ยงของแบบสอบที่ใช้คำสั่งและการให้คะแนนตามวิธีของกุ่มบัส สูงกว่าค่าความเที่ยงของแบบสอบที่ใช้คำสั่งและการให้คะแนนตามวิธี 0-1 ค่าความตรงของแบบสอบที่ใช้คำสั่งและการให้คะแนนตามวิธีของ สำราญ มีแจ้ง สูงกว่าค่าความตรงของแบบสอบที่ใช้คำสั่งและการให้คะแนนตามวิธี 0-1 และค่าความตรงของแบบสอบที่ใช้คำสั่งและการให้คะแนนตามวิธีของ อนันต์ ศรีโสภา สูงกว่าค่าความตรงของแบบสอบที่ใช้คำสั่งและการให้คะแนนตามวิธี 0-1 ส่วนค่าอำนาจจำแนกของแบบสอบที่ใช้คำสั่งและการให้คะแนนทั้ง 4 วิธีไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย