# บทที่ 2

## วรรณคดีที่เกี่ยวข้อง

ผู้วิจัยได้แบ่งวรรณคดีที่เกี่ยวข้องออกเป็น 2 ตอน คือ ตอนที่ 1 เป็นมโนทัศน์เบื้องต้นเกี่ยวกับค่าความเที่ยง ความตรงของแบบสอบ ตอนที่ 2 เป็นงานวิจัยเกี่ยวกับการเปรียบเทียบค่าความเที่ยง ความตรงของแบบสอบเลือกตอบที่มีวิธีการให้คะแนนความมั่นใจกับวิธีการให้คะแนนวิธีอื่น ๆ ทั้งในประเทศและต่างประเทศ

### ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับค่าความเที่ยง ความตรงของแบบสอบ

เนื่องจากค่าความเที่ยง ความตรงเป็นคุณสมบัติสำคัญ ซึ่งจะบ่งชี้ว่าแบบสอบนั้น ๆ มีคุณภาพดีไม่ดีเพียงใด ในขณะเดียวกันมีทฤษฎีและงานวิจัยอีกมากที่สนับสนุนว่าคุณภาพของแบบสอบจะแตกต่างกันเมื่อกำหนดเทคนิคการให้คะแนนที่แตกต่างกัน (สุพัฒน์ สุกมลสันต์ 2529:1) แต่ในความเป็นจริงเทคนิคการให้คะแนนที่ใช้ก็คือ การให้คะแนนวิธีธรรมดา ซึ่งกำหนดน้ำหนักคะแนน 1 ให้กับข้อที่ตอบถูก ส่วนข้อที่ผิดหรือเว้นว่างให้คะแนนเป็น 0 การหาคะแนนรวมสำหรับผู้เข้าสอบแต่ละคน จึงนับจากจำนวนข้อกระทงที่ตอบถูก ดังสมการ

$$X = R$$

เมื่อ X แทนคะแนนดิบของแบบสอบทั้งฉบับ

R แทนจำนวนข้อที่ตอบถูก

Harold Gullikson (Harold Gullikson 1965 อ้างถึงในสุพัฒน์ สุกมลสันต์ 2529:1) ยอมรับว่าวิธีการให้คะแนนวิธีธรรมดามีความสะดวกและเหมาะสมถ้าผู้สอบตอบข้อกระทงทุกข้อติดต่อกันไป จำนวนข้อกระทงที่เว้นข้ามไปก็จะเป็นศูนย์หรือมีน้อยจนไม่จำเป็นต้องนำมาพิจารณา แต่ในความเป็นจริงการเรียงข้อกระทงของแบบสอบมักไม่เรียงลำดับตามความยากง่ายและผู้สอบไม่ได้ทำข้อกระทงอย่างต่อเนื่อง จึงมีผลทำให้ค่าความเที่ยง ความตรงของแบบสอบลดลง จึงทำให้นักวัดผลหลายท่านพยายามศึกษาและนำเทคนิคการให้คะแนนวิธีต่าง ๆ มาใช้ ซึ่งผู้วิจัยจะไม่กล่าวถึงในที่นี้ แต่จะขอกล่าวถึงมโนทัศน์เบื้องต้นเกี่ยวกับความเที่ยง ความตรง เพื่อให้เข้าใจในเรื่องนี้ก่อน

## ความเที่ยงของแบบสอบ

มีผู้ให้นิยามของความเที่ยงไว้หลายท่านได้แก่

ความเที่ยงหมายถึงความคงเส้นคงวาของชุดคะแนนจากแบบสอบที่วัดในสิ่งที่ต้องการวัด (Ebel 1956 : 310)

ความเที่ยงหมายถึงความคงเส้นคงวาของการวัด (Grounlund 1968:105)

ความเที่ยงหมายถึงระดับ (Degree) ของความคงที่ระหว่างการวัดสิ่งเดียวกัน 2 ครั้ง (Mehrens and Lehmann 1984:267)

ความเที่ยงหมายถึงความคงที่ของคะแนนที่ได้จากการทดสอบคนกลุ่มเดียวกัน 2 ครั้งด้วยแบบสอบเดิมในเวลาต่างกัน หรือด้วยแบบสอบคนละชุดที่มีการเทียบเท่ากัน หรือภายใต้สภาพการสอบที่ต่างกัน ( อนาสตาซี 2519 : 73 )

จากนิยามความเที่ยงดังกล่าวข้างต้น จะเห็นว่าความเที่ยงมีความสำคัญต่อสถานการณ์การทดสอบ 2 ประการคือ 1) จากนิยามของความเที่ยงจะบ่งชี้ถึงระดับความสามารถที่กระทำได้ของผู้สอบ 2) ค่าที่คำนวณได้จากความเที่ยง (Reliabitity Coefficient ) จะช่วยให้สามารถประมาณคะแนนจริง (True Score) ของผู้สอบ ซึ่งแสดงถึงความสามารถที่แท้จริงของผู้สอบ

## ทฤษฎีของความเที่ยง ( Theory of Reliability )

การอธิบายทฤษฎีความเที่ยงจำเป็นต้องกล่าวถึงทฤษฎีการทดสอบโดยเริ่มจากคะแนนผลการสอบ (Observed Score) ซึ่งประกอบด้วยส่วนที่เป็นคะแนนจริง (True Score) และส่วนที่เป็นคะแนนความคลาดเคลื่อน (Error Score) ซึ่งเขียนเป็นสมการได้ดังนี้

$$X = T + E$$

เมื่อ $X$ = คะแนนจากผลการสอบ

$T$ = คะแนนจริง

$E$ = คะแนนความคลาดเคลื่อน

คะแนนจริง (True Score) หมายถึงคะแนนที่ผู้สอบพึงได้รับจากการวัดด้วยเครื่องมือที่มีคุณลักษณะสมบูรณ์ปราศจากความคลาดเคลื่อน หรือหมายถึงคะแนนเฉลี่ยของผู้สอบซึ่งได้จากการทำแบบสอบเดิมหลาย ๆ ครั้ง โดยมีข้อตกลงว่าไม่มีอิทธิพลจากการฝึกฝน ความเมื่อยล้า และการเรียนรู้ในการทดสอบซ้ำ (Brown 1976:50)

คะแนนความคลาดเคลื่อน (Error Score) หมายถึงค่าความผิดพลาดที่เกิดขึ้นจากการวัดซึ่งเป็นความคลาดเคลื่อนที่เกิดขึ้นในลักษณะสุ่ม (Random Score) เช่นความคลาดเคลื่อนที่เกิดขึ้นจากตัวแบบสอบ การบริหารการสอบ การให้คะแนน และสิ่งต่าง ๆ ที่เกิดจากตัวผู้เข้าสอบเป็นต้น โดยทฤษฎีความคลาดเคลื่อนที่เกิดขึ้นนี้เป็นไปได้ทั้งทางบวกและทางลบซึ่งจะหักล้างกันหมดไป และความคลาดเคลื่อนที่เกิดขึ้นนี้ไม่มีความสัมพันธ์กับคะแนนจริง นั่นคือสัมประสิทธิ์สหพันธ์ (Correlation) ระหว่างคะแนนจริงกับคะแนนคลาดเคลื่อนมีค่าเท่ากับ 0 ดังนั้นความแปรปรวนของคะแนนที่วัดได้สามารถเขียนได้ดังสมการ (Mehrens and Lehmann 1984 : 269) ดังนี้

$$S_x^2 = S_t^2 + S_e^2 \text{ ----------(1)}$$

เมื่อ $S_x^2$ = ความแปรปรวนของคะแนนจากผลการสอบ

$S_t^2$ = ความแปรปรวนของคะแนนจริง

$S_e^2$ = ความแปรปรวนของคะแนนความคลาดเคลื่อน

โดยทฤษฎีความเที่ยงหมายถึง อัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนจากผลการวัด (Mehrens and Lehmann 1984:269) จึงสามารถเขียนสมการได้เป็น

$$r_{tt} = 1 - \frac{S_t^2}{S_x^2}$$

เมื่อ $r_{tt}$ = ค่าความเที่ยง

จากสมการค่าความเที่ยงบอกให้ทราบถึงสัดส่วนของความแปรปรวนของคะแนนจริงกับคะแนนที่ได้จากผลการสอบ และจากสมการที่ (1) สามารถเขียนสมการที่ (2) ได้เป็น

$$r_{tt} = 1 - \frac{S_e^2}{S_x^2} \text{ ---------- (2)}$$

ซึ่งถือเป็นสมการพื้นฐานสำหรับการประมาณค่าความเที่ยง

<u>วิธีประมาณค่าความเที่ยง</u>

วิธีประมาณค่าความเที่ยงมีหลายวิธี แต่ละวิธีทำให้เกิดความคลาดเคลื่อนแตกต่างกัน การจะเลือกใช้วิธีใดนั้นขึ้นอยู่กับความมุ่งหมาย ลักษณะของคะแนน และชนิดของแบบสอบ (Guilford 1978:414) โดยทั่วไปวิธีประมาณค่าความเที่ยงมี 4 แบบ (Mehrens and Lehmann 1984:272) คือ

1. แบบสอบซ้ำ (Measure of Stability)

2. แบบใช้แบบสอบคู่ขนาน (Measure of Equivalence)

3. แบบใช้แบบสอบคู่ขนานและสอบซ้ำ (Measure of Equivalence and stability)

4. แบบวัดความคงที่ภายใน(Measure of Internal Consistency) ซึ่งแบ่งออกเป็น

4.1 แบบแบ่งครึ่งจำนวนข้อสอบ (Split-half)

4.2 แบบคูเดอร์ริชาดสัน (Kuder-Richardson Estimation)

4.3 แบบสัมประสิทธิ์อัลฟา (Alpha Coefficient)

4.4 แบบวิเคราะห์ความแปรปรวนของฮอยด์ (Hoyts Analysis of Variance Procedure )

ในที่นี้จะกล่าวเฉพาะความเที่ยงแบบสัมประสิทธิ์อัลฟา ซึ่งเป็นวิธีประมาณค่าความเที่ยงที่ใช้ในการวิจัยครั้งนี้

<u>ความเที่ยงแบบสัมประสิทธิ์อัลฟา</u>

ในปี ค.ศ. 1951 Cronbach (Cronbach 1951) ได้พัฒนาวิธีการประมาณค่าความเที่ยงภายในของแบบสอบจากวิธีของคูเดอร์ริชาดสันในกรณีที่คะแนนของข้อกระทงไม่สามารถจะให้คะแนนข้อถูกเท่ากับ 1 และข้อผิดเท่ากับ 0 (dichotomous) ได้ โดยมีหลักการคล้ายวิธีแบ่งครึ่งจำนวนข้อสอบต่างกันที่ว่าวิธีอัลฟาจะให้ข้อกระทงแต่ละข้อเป็นแบบสอบเล็ก ๆ (mini test ) คือแบบสอบชุดหนึ่งมี 5 ข้อกระทงก็เสมือนว่าเป็นแบบสอบเล็ก ๆ 5 ชุดและถ้าข้อกระทงทุกข้อวัดในสิ่งเดียวกัน ค่าสัมประสิทธิ์ระหว่างข้อกระทงก็จะสูงใกล้เคียง 1.00 ตรงกันข้ามถ้าข้อกระทงต่าง ๆ วัดในสิ่งที่ต่างกัน ค่าสัมประสิทธิ์ระหว่างข้อกระทงบางส่วนจะต่ำ ค่าก็จะลดลงจาก 1.00

วิธีสัมประสิทธิ์อัลฟาเป็นที่นิยมใช้เพราะสามารถใช้กับแบบสอบประเภทต่าง ๆ แม้จะมีค่าคะแนนเฉพาะในแต่ละข้อที่ต่างกันก็ได้

สำหรับสูตรที่ใช้คำนวณคือ

$$\text{Alpha } (\alpha) = \frac{K}{K-1}\left(1 - \frac{\sum \text{Var}(i)}{\text{Var}(x)}\right)$$

เมื่อ K คือจำนวนข้อกระทงในแบบสอบ

$\sum \text{Var}(i)$ คือผลรวมจากความแปรปรวนของการกระจายของคะแนนในข้อกระทงต่าง ๆ

$\text{Var}(x)$ คือความแปรปรวนของการกระจายของคะแนนสอบทั้งหมด

ค่าอัลฟาสูงหมายความว่าแบบสอบฉบับนั้นมีความคงที่ภายในคืออยู่ในเกณฑ์ที่แสดงว่าแบบสอบทั้งฉบับวัดในคุณลักษณะเดียวกัน ค่าสหพันธ์ระหว่างข้อกระทงโดยเฉลี่ยสูงซึ่งแสดงว่าความสามารถจากผลการสอบข้อกระทงใดก็ตามจะพยากรณ์ถึงความสามารถในการสอบข้อกระทงอื่น ๆ ได้อยู่ในเกณฑ์ดี

<u>องค์ประกอบที่มีอิทธิพลต่อความเที่ยง</u>

องค์ประกอบที่มีอิทธิพลต่อความเที่ยงของแบบสอบมีหลายองค์ประกอบ (Mehrens and Lehmann 1984 : 278-281) ดังนี้คือ

1. ความยาวของแบบสอบ (Test Length) การเพิ่มจำนวนข้อสอบซึ่งมีคุณภาพดีเท่าเทียมกับข้อสอบเดิมจะทำให้แบบสอบมีค่าความเที่ยงสูงขึ้น นั่นคือแบบสอบที่มีจำนวนข้อมากจะมีค่าความเที่ยงสูงกว่าแบบสอบที่มีจำนวนข้อน้อย สูตรที่แสดงให้เห็นความสัมพันธ์ระหว่างความยาวของแบบสอบกับความเที่ยงคือสูตรของสเปียรแมน-บราวน์ (Spearman-Brown Formular)

2. ความเป็นเอกพันธ์ของกลุ่มผู้สอบ (Group Homogeneity) กลุ่มผู้สอบที่มีความเป็นเอกพันธ์จะมีพิสัยของคะแนนความสามารถของกลุ่มผู้สอบน้อยกว่ากลุ่มที่มีความเป็นวิวิธพันธ์ (Group Heterogeneity) ซึ่งจะทำให้ค่าความเที่ยงของกลุ่มที่มีความเป็นเอกพันธ์ต่ำกว่ากลุ่มที่มีความเป็นวิวิธพันธ์

014019

3. ความยากของข้อกระทง(Item Difficulty) เนื่องจากค่า ค่าความเที่ยงประมาณจากความแปรปรวนของคะแนนผลการสอบ ข้อกระทงยากเกินไปทุกคนทำผิดหมดหรือข้อกระทงง่ายเกินไปทุกคนทำถูกหมดก็จะทำให้พิสัยของคะแนนลดลง ค่าความแปรปรวนของคะแนนมีน้อยจะทำให้ค่าความเที่ยงลดลงด้วย ดังนั้นค่าความยากของข้อกระทงซึ่งทำให้คะแนนมีการกระจายมากจะทำให้ค่าความเที่ยงสูง

4. ความเร็ว(Speed) การประมาณค่าความเที่ยงของแบบสอบความเร็ว (Speed test) ด้วยวิธีวัดความคงที่ภายในเป็นวิธีที่ไม่เหมาะสม เพราะแบบสอบชนิดนี้เป็นแบบสอบที่ง่ายแต่ใช้เวลาทำน้อย ซึ่งผู้สอบไม่สามารถทำได้ครบทุกข้อจึงทำให้ค่าความเที่ยงสูงกว่าที่ควรจะเป็น

5. ความเป็นปรนัย (Objective) ถ้าแบบสอบมีความเป็นปรนัยสูง ค่าความเที่ยงของแบบสอบก็สูงขึ้นด้วย

นอกจากปัจจัยดังกล่าวข้างต้นยังมีองค์ประกอบเกี่ยวกับการบริหารการสอบเช่น การจับเวลา การให้คำชี้แจง การป้องกันการทุจริตเป็นต้น หรือองค์ประกอบเกี่ยวกับตัวผู้สอบเองเช่น แรงจูงใจ ความวิตกกังวล ประสบการณ์ในการสอบ การเดาเป็นต้น สิ่งเหล่านี้อาจเป็นแหล่งความคลาดเคลื่อนที่ทำให้ค่าความเที่ยงของแบบสอบเปลี่ยนแปลงได้

<u>ความตรงของแบบสอบ</u>

มีผู้ให้นิยามของความตรงไว้หลายท่านได้แก่

ความตรงของแบบสอบคือ ค่าสหสัมพันธ์ระหว่างแบบสอบกับเกณฑ์ ( Gullikson 1950 : 88 )

ความตรงของแบบสอบคือขอบเขตของผลการประเมินที่สอดคล้องกับจุดประสงค์ (Grounlund 1976 : 76)

ความตรงของแบบสอบคือความถูกต้องแม่นยำของคะแนนในการพยากรณ์ที่จำเพาะเจาะจง (Stanley 1978 : 101)

ความตรงของแบบสอบคือความสามารถวัดสิ่งที่ต้องการวัดได้ถูกต้องแม่นยำหรือความสามารถในการให้ความหมายในสิ่งที่วัดได้อย่างไม่ผิดพลาด(Linquist 1942: อ้างถึงในเยาวดี วิบูลย์ศรี 2528 : 69 )

จากความหมายของความตรงที่กล่าวมาข้างต้นสรุปได้ 2 ประการคือ

1. ความตรงคือค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบกับคะแนนเกณฑ์

2. ความตรงคือความสามารถของแบบสอบที่วัดในสิ่งที่ต้องการวัดหรือวัดได้สอดคล้องกับจุดประสงค์ คะแนนที่ได้จากแบบสอบที่มีค่าความตรงสูงสามารถบอกถึงสภาพที่แท้จริงและพยากรณ์ได้ถูกต้องแม่นยำ

<u>ประเภทของความตรง</u>

สมาคมจิตวิทยาอเมริกัน(American Psychological Association) ได้จำแนกความตรงออกเป็น 3 ประเภทใหญ่ ๆ (mehrens and Lehmann 1984: 289) คือ

1. ความตรงตามเนื้อหา(Content Validity)

2. ความตรงตามเกณฑ์สัมพันธ์ (Criterior-related Validity)

3. ความตรงตามทฤษฎี (Construct Validity)

<u>ความตรงตามเนื้อหา</u> (Content Validity)

ความตรงตามเนื้อหาจะเกี่ยวข้องกับเนื้อหาสาระของข้อกระทงในแบบสอบ ซึ่งข้อกระทงนั้นต้องเป็นตัวแทนของเนื้อหาสาระและครอบคลุมขอบเขตของพฤติกรรมที่ต้องการวัด ความตรงตามเนื้อหามีความสำคัญอย่างยิ่งในการสอบวัดผลสัมฤทธิ์ทางการเรียน การตัดสินว่าแบบสอบนั้นมีความตรงตามเนื้อหามากน้อยเพียงใดจะต้องพิจารณาทั้งหัวข้อเนื้อหาวิชาและชนิดของพฤติกรรมที่วัดไปพร้อม ๆ กัน

การหาค่าความตรงตามเนื้อหาไม่สามารถคำนวณเป็นตัวเลขได้ โดยทั่วไปจะพิจารณาจากข้อกระทงในแบบสอบเปรียบเทียบกับตารางวิเคราะห์หลักสูตร (Table of Specification) แล้วประมาณออกมาว่ามีความตรงตามเนื้อหามากน้อยเพียงใด ผู้ที่จะตัดสินความตรงตามเนื้อหาได้ดีก็คือผู้เชี่ยวชาญในสาขานั้น ๆ ( Mehrens and Lehmann 1984 : 290 )

<u>ความตรงตามเกณฑ์สัมพันธ์</u> (Criterior-Related Validity)

ความตรงตามเกณฑ์สัมพันธ์เป็นเรื่องที่เกี่ยวกับการศึกษาความสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบกับคะแนนการวัดจากเกณฑ์ภายนอกที่เป็นอิสระ (Mehreas and Lehmann 1984 : 291)

เกณฑ์ (Criterior) คือมาตรฐานในการตัดสิน การพัฒนาแบบสอบเกณฑ์คือ คุณลักษณะหรือกลุ่มของคุณลักษณะที่ใช้เป็นพื้นฐานสำหรับตัดสินความตรงของแบบสอบหรือกระบวนการวัดอื่น ๆ (Eble 1965 : 446)

ความตรงตามเกณฑ์สัมพันธ์จำแนกได้เป็น 2 ชนิด (เยาวดี วิบูลย์ศรี 2528 : 73 - 77)

1. ความตรงเชิงทำนาย (Predictive Validity)
2. ความตรงร่วมสมัย (Concurrent Validity)

ในที่นี้จะกล่าวเฉพาะความตรงร่วมสมัยซึ่งใช้ในงานวิจัยครั้งนี้

<u>ความตรงร่วมสมัย</u> ( Concurrent Validity)

ความตรงร่วมสมัยหมายถึงความตรงของแบบสอบที่จะบ่งบอกสิ่งที่วัดได้ถูกต้องตามสภาพที่แท้จริงในปัจจุบัน โดยอาศัยความสัมพันธ์ระหว่างคะแนนของแบบสอบกับคะแนนเกณฑ์ซึ่งกำหนดขึ้นในขณะนั้นเช่น เกณฑ์สัมพันธ์ของแบบวัดทักษะในการคำนวณก็อาจใช้วิธีการสังเกตการทำแบบฝึกหัดเลยในชั่วโมงสอนเป็นเวลา 2 เดือน ถ้าคะแนนที่นักเรียนได้จากแบบสอบให้ผลสอดคล้องกับเกณฑ์สัมพันธ์แสดงว่าแบบสอบนี้สามารถวัดสภาพปัจจุบันของนักเรียนได้

กล่าวโดยสรุปความตรงร่วมสมัยนั้นใช้กับแบบสอบสำหรับวินิจฉัยสภาพที่เป็นอยู่ในปัจจุบัน และช่วงระยะเวลาของการเก็บข้อมูลที่เป็นเกณฑ์เป็นช่วงเวลาเดียวกับคะแนนการสอบ

การวัดเกณฑ์

Mehrens and Lehmann(1984:293)กล่าวว่าการพิจารณาเกณฑ์นั้นจะต้องพิจารณาในสิ่งต่อไปนี้

1. ความเกี่ยวข้อง (Relevance) สิ่งที่จะใช้เป็นเกณฑ์จะต้องมีความเกี่ยวข้องกับเรื่องที่ศึกษา นั่นคือเกณฑ์นั้นจะต้องสะท้อนถึงคุณลักษณะที่สำคัญของเรื่องที่จะศึกษา โดยทั่วไปการพิจารณาความเกี่ยวข้องของเกณฑ์นี้อยู่กับการตัดสินคุณค่าไม่ได้วัดออกมาเป็นตัวเลข

2. ความเที่ยง (Reliability) สิ่งที่จะใช้เป็นเกณฑ์จะต้องมีความเที่ยงพอควร ทั้งนี้เพราะค่าความตรงของแบบสอบขึ้นอยู่กับค่าสหสัมพันธ์ระหว่างแบบสอบกับเกณฑ์ ซึ่งทฤษฎีความสัมพันธ์กล่าวไว้ว่าความสัมพันธ์สูงสุดระหว่าง 2 ตัวแปรจะเท่ากับหรือน้อยกว่ารากที่สองของผลคูณระหว่างความเที่ยงของตัวแปรทั้งสองนั้นนั่นอีก

$$r_{xy} = \sqrt{(r_{xx})(r_{yy})}$$

เมื่อ $r_{xy}$ = ค่าสหสัมพันธ์ระหว่างตัวทำนาย (X) กับเกณฑ์ (Y)

$r_{xx}$ = ค่าความเที่ยงของแบบสอบที่เป็นตัวทำนาย

$r_{yy}$ = ค่าความเที่ยงของเกณฑ์

จะเห็นได้ว่าค่าความตรงตามเกณฑ์สัมพันธ์ ($r_{xy}$) มีค่ามากน้อยขึ้นอยู่กับความเที่ยงของแบบสอบเอง ($r_{xx}$) และความเที่ยงของเกณฑ์ ($r_{yy}$)

3. ความเป็นอิสระจากอคติ (Free from Bias ) เกณฑ์ที่จะใช้จะต้องเป็นอิสระต้องแน่ใจว่าคะแนนจากแบบสอบที่จะหาความตรงนั้นไม่มีอิทธิพลต่อสถานภาพใด ๆ ของเกณฑ์สัมพันธ์ที่ใช้เทียบ นั่นคือต้องไม่ให้เป็นเกณฑ์มลทิน(Criteria Contamination ตัวอย่างเช่นนำแบบสอบความถนัดทางวิชาการไปสอบกับนักเรียนตอนต้นปีพยากรณ์ผลการเรียนของนักเรียนปลายปี ถ้าปรากฏว่าความรู้ที่ครูได้รับจากการสอบครั้งนี้ไปส่งเสริมให้ประเมินคะแนนการสอบปลายปีของนักเรียนบางคนสูงขึ้นหรือต่ำลง ลักษณะเช่นนี้แสดงว่าคะแนนของเกณฑ์ไม่มีความอิสระหรือเป็นเกณฑ์มลทิน ซึ่งอิทธิพลเช่นนี้ทำให้ค่าความตรงตามเกณฑ์สัมพันธ์ของแบบสอบสูงขึ้น โดยมิใช่เกิดจากคุณภาพของแบบสอบจริง ๆ การป้องกันความคลาดเคลื่อนเช่นนี้อาจทำได้โดยหลีกเลี่ยงไม่ให้ผู้ประเมินทราบคะแนนการสอบที่ใช้พยากรณ์

<u>วิธีประมาณค่าความตรงตามเกณฑ์สัมพันธ์</u>

การประมาณค่าความตรงตามเกณฑ์สัมพันธ์ สามารถวิเคราะห์ได้จากสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบสอบกับเกณฑ์ ดังนั้นสูตรที่ใช้ในการคำนวณโดยทั่วไปนิยมใช้สูตรสัมประสิทธิ์สหสัมพันธ์แบบเพียรสัน (Pearsons Product Moment Correlation Coefficient) (Guilford 1978 : 83)

<u>ความตรงตามทฤษฎี</u> (Construct Validity)

ความตรงตามทฤษฎี หมายถึง ความสัมพันธ์ระหว่างคะแนนที่ได้รับจากแบบสอบกับคะแนนที่ได้จากการวัดอื่น ๆ ซึ่งต่างก็วัดในทฤษฎีเดียวกันตามที่กำหนด (Leona Tyler 1965 อ้างถึงในเยาวดี วิบูลย์ศรี 2528 : 78)

ความตรงตามทฤษฎี หมายถึงระดับของการอ้างอิงคุณลักษณะตามทฤษฎีทางจิตวิทยาจากคะแนนของแบบสอบที่สร้างขึ้น (Mehrens and Lehmann 1984:294)

กล่าวโดยสรุปความตรงตามทฤษฎี หมายถึง ความสามารถของแบบสอบที่สามารถวัดทฤษฎีหรือคุณลักษณะของพฤติกรรมที่ได้อธิบายไว้หรือตามสมมติฐานที่ตั้งไว้หรือภาวะสันนิฐานที่ตั้งไว้ ดังนั้นค่าความตรงตามทฤษฎีจะเป็นตัวบ่งชี้ความสัมพันธ์ระหว่างทฤษฎีกับความสามารถที่แสดงออกได้จริงตามที่กำหนดไว้ในแบบลอบ ถ้าแบบสอบใดวัดคุณลักษณะที่ต้องการจะศึกษาจริงแล้วแบบสอบนั้นจะต้องมีความตรงตามทฤษฎีสูง

<u>วิธีการประมาณค่าความตรงตามทฤษฎี</u>

โดยทั่วไปวิธีการประมาณค่าความตรงตามทฤษฎีที่นิยมใช้มี 3 วิธี (เยาวดี วิบูลย์ศรี 2528 : 81-82) คือ

1. Intratest Method
2. Intertest Method
3. Factor Analysis

<u>องค์ประกอบที่มีอิทธิพลต่อความตรง</u>

องค์ประกอบที่ทำให้ค่าความตรงของแบบสอบลดลง (Grounlund 1976:98-102)

1. องค์ประกอบภายในตัวแบบสอบ ได้แก่

1.1 คำสั่งที่ไม่ชัดเจน มีผลทำให้นักเรียนปฏิบัติตามคำสั่งผิดพลาดได้

1.2 ข้อความในประโยคยาวมากเกินไป ทำให้ผู้สอบเกิดความสับสน ผู้ที่ตอบถูกต่อผู้ที่มีความสามารถในการอ่านสูง ดังนั้นแบบสอบนี้จึงวัดความเข้าใจในการอ่านของผู้สอบแทนที่จะวัดพฤติกรรมที่ต้องการ

1.3 ระดับความยากของข้อกระทงไม่เหมาะสม ข้อกระทงที่ยากหรือง่ายเกินไปไม่สามารถจำแนกผู้สอบได้

1.4 การสร้างข้อกระทงไม่รัดกุม ข้อกระทงที่สร้างโดยไม่ระมัดระวังอาจชี้แนะคำตอบแก่ผู้สอบได้

1.5 ภาษาที่ใช้ในข้อกระทงกำกวม ทำให้ผู้สอบตีความผิดและสับสน

1.6 ประเภทของแบบสอบที่ใช้ไม่เหมาะสมจะใช้วัดพฤติกรรมที่กำหนด

1.7 แบบสอบสั้นเกินไป ไม่สามารถที่จะถามให้ครอบคลุมวัตถุประสงค์ของการวัด

1.8 การจัดเรียงข้อกระทงไม่เหมาะสม ข้อกระทงควรเรียงจากง่ายไปยากเพราะถ้าเรียงจากยากไปง่ายผู้สอบอาจใช้เวลาคิดนานจนกระทั่งไม่มีเวลาในการทำข้อง่ายทำให้การวัดไม่ตรงตามความสามารถแท้จริง

1.9 การเรียงคำตอบอย่างเป็นระบบ ทำให้ผู้สอบสามารถเดาคำตอบที่ถูกได้ในข้อกระทงข้อหลัง ๆ

นอกจากนี้ค่าความเที่ยงของแบบสอบมีส่วนสำคัญที่จะกำหนดค่าความตรงของแบบสอบได้ (Brown 1971:99) ดังสมการ

$$r_{xy} \leqslant \sqrt{r_{xx}}$$

เมื่อ $r_{xy}$ ค่าความตรงของแบบสอบ

$r_{xx}$ ค่าความเที่ยงของแบบสอบ

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จะเห็นได้ว่าค่าความตรงของแบบสอบมีค่าไม่เกินรากที่สองของค่าความเที่ยงของแบบสอบ แสดงว่าค่าความตรงถูกจำกัดโดยค่าความเที่ยงของตัวมันเอง

2. การบริหารแบบสอบและการให้คะแนน ทำให้ความตรงของแบบสอบลดลงได้ เช่นให้เวลาในการสอบน้อยเกินไป การให้คะแนนที่ขาดความเชื่อถือเช่นกรณีแบบสอบอัตนัย ถ้าเป็นแบบสอบมาตรฐาน มักพบในกรณีผู้ดำเนินการสอบไม่ปฏิบัติตามคู่มือ เช่นไม่จับเวลาตามที่กำหนด ตรวจให้คะแนนผิด การปฏิบัติเหล่านี้มีผลให้ความตรงของแบบสอบลดลง

3. องค์ประกอบในตัวผู้สอบ เช่นสภาพจิตใจไม่ปกติมีความวิตกกังวล ขาดแรงจูงใจ การเดาคำตอบของผู้สอบ สิ่งหล่านี้ทำให้ค่าความตรงของแบบสอบลดลง

4. ลักษณะของกลุ่มและเกณฑ์ ความตรงมักเป็นของเฉพาะกลุ่ม เนื่องจากสิ่งที่แบบสอบวัดได้นั้นมักได้รับอิทธิพลของตัวแปรต่าง ๆ เช่น อายุ เพศ ระดับความสามารถ ภูมิหลังของการศึกษาและวัฒนธรรมของผู้สอบ ดังนั้นในการพิจารณาหาความตรงของแบบสอบที่เป็นมาตรฐาน ควรคำนึงถึงลักษณะเฉพาะของกลุ่มที่ใช้หาความตรง นอกจากนั้นต้องพิจารณาเกณฑ์ที่ใช้หาความตรงของแบบสอบด้วย

<u>ตอนที่ 2 งานวิจัยที่เกี่ยวข้อง</u>

งานวิจัยที่เกี่ยวข้องกับการวิจัยนี้จะเสนอการวิจัยเกี่ยวกับการเปรียบเทียบค่าความเที่ยง ความตรงของแบบสอบเลือกตอบที่มีวิธีการให้คะแนนความมั่นใจกับวิธีการอื่น ๆ ทั้งในประเทศและต่างประเทศ

Pressey (Pressey 1950:417-447, quoted in Whetton and Childs 1981:336) ได้พัฒนาแบบสอบเลือกตอบให้สามารถบอกคะแนนและตัวถูกให้ทราบทันทีในขณะทำแบบสอบ วิธีตอบแบบใหม่ต่อมาได้เรียกว่าวิธีตอบจนถูก(Answer-until-correct หรือ AUC :) วิธีการนี้ทำให้พิสัยของคะแนนของแต่ละข้อคำถามเพิ่มขึ้นและทำให้ส่วนของความรู้ต่อเนื่องกันไป แต่วิธีการนี้มีปัญหาคือต้องทำกระดาษคำตอบขึ้นเป็นพิเศษจึงยุ่งยากสิ้นเปลือง

Coombs(Coombs 1950:13-37 อ้างถึงในอนันต์ ศรีโสภา 2516:41-46) ได้เสนอให้เปลี่ยนวิธีในการตอบและการให้คะแนนเสียใหม่ โดยสั่งให้ทำเครื่องหมายอย่างใดอย่างหนึ่งลงบนตัวเลือกที่ถูกต้องเช่นให้วงกลม ซึ่งตัวเลือกที่ถูกต้องนี้อาจเลือกได้มากกว่าหนึ่งตัวหรือไม่เลือกเลยก็ได้ และทำเครื่องหมายบนตัวเลือกที่คิดว่าผิดด้วย โดยใช้เครื่องหมายที่ไม่ซ้ำกันกับข้อถูก เช่นอาจให้เครื่องหมายกากบาทก็ได้ วิธีการนี้ต้องการกำจัด

ตัวเลือกที่ผู้สอบคิดว่าไม่ถูกออกไป สมมติว่าแต่ละข้อมี K ตัวเลือกและมีตัวเลือกหนึ่งเป็นคำตอบที่ถูกต้องดังนั้นจึงมีตัวลวง (distractors) K-1 ตัว การให้คะแนน จะให้ +1 คะแนนเมื่อผู้สอบทำเครื่องหมายกากบาทลงบนตัวลวงและให้ -1 คะแนนเมื่อผู้สอบทำเครื่องหมายวงกลมบนตัวลวง ดังนั้นพิสัยของคะแนนจะอยู่ระหว่าง1-K ถึง K-1 คะแนน

วิธีของดูมบ์สมีข้อดีคือพิสัยของคะแนนกว้างขึ้น สามารถจำแนกความรู้ของผู้สอบได้ละเอียดกว่าวิธีธรรมดาและยังทราบความรู้ที่ผิดของผู้สอบด้วย แต่วิธีนี้มีข้อบกพร่องในเรื่องโอกาสของการได้คะแนนเช่นแบบสอบที่มี 5 ตัวเลือก ผู้สอบคนหนึ่งรู้ว่าตัวเลือกใดผิดตัวเลือกใดถูกเขาสามารถทำเครื่องหมายลงบนตัวลวงได้ถูกต้องและได้ 4 คะแนน ส่วนผู้สอบอีกคนไม่รู้ว่าตัวเลือกใดผิดแต่รู้ว่าตัวเลือกใดถูกก็สามารถทำเครื่องหมายลงบนตัวเลือกอื่น ๆ ที่ไม่ใช่ตัวถูกและได้ 4 คะแนนเท่าคนแรกทั้งที่ผู้สอบทั้งสองมีความรู้ไม่เท่ากัน

สำหรับวิธีให้คะแนนความมั่นใจ (Confidence Weighting Procedures) วิธีนี้ผู้สอบจะต้องบอกถึงระดับความมั่นใจของตัวเลือกที่เขาเลือกว่าถูกต้องเพียงไร วิธีการบอกความมั่นใจนี้มีหลายวิธี เริ่มใช้ครั้งแรกในระหว่างปี ค.ศ. 1930-1939 ในฐานะที่เป็นวิธีในการเพิ่มข้อมูลให้กับการทดสอบและกำจัดผลที่จะเกิดขึ้นจากการเดาของแบบสอบประเภทถูกผิด(Echternacht 1972 : 218)

Dressel และ Schmid(Dressel and Schmid 1953:574-595, quoted in Echternacht 1972:217-234)ให้คะแนนความมั่นใจกับงานวิจัยของเขา ถึงแม้จะมีวัตถุประสงค์ที่จะศึกษาระดับความมั่นใจในการตอบข้อสอบ เมื่อใช้คำสั่งชี้แจงแบบให้เดาและไม่ให้เดาแต่ประเด็นสำคัญของการทดสอบความมั่นใจที่เขาทำก็คือเริ่มใช้กับแบบสอบเลือกตอบซึ่งแตกต่างกับที่ผู้วิจัยคนก่อน ๆ ทำ และมีการแบ่งระดับความมั่นใจให้ละเอียดขึ้น โดยการแบ่งออกเป็น 5 ระดับด้วยกัน

Dressel และ Schmid(Dressel and Schmid 1953:574-595,quoted quoted in Echternacht 1972:217-234) ทำการวิจัยโดยใช้ข้อคำถามและระบุการให้คะแนนแบบต่าง ๆ เพื่อปรับปรุงคุณภาพของแบบสอบเลือกตอบ แบบสอบเลือกตอบที่ใช้มี 4 แบบด้วยกัน 1) แบบสอบเลือกตอบเสรี (Free-choice test) เป็นแบบสอบที่ผู้สอบจะเลือกคำตอบกี่ข้อก็ได้ตามที่เขาคิดว่าถูกต้อง

2) แบบสอบบอกความมั่นใจ (Degree-of certainly test ) เป็นแบบสอบที่ผู้สอบเลือกคำตอบได้หนึ่งข้อและบอกว่าตนเองแน่ใจในคำตอบนั้นอยู่ในระดับใด โดยแบ่งความมั่นใจเป็น 4 ระดับ 3) แบบมีคำตอบถูกมากกว่าหนึ่ง (Multiple answer test) เป็นแบบสอบที่มีจำนวนคำตอบถูกในแต่ละข้อไม่จำกัด ผู้สอบจะต้องทำเครื่องหมายทุกข้อที่คิดว่าถูก 4) แบบมีคำตอบถูก 2 คำตอบ(Two-answer Test) จากตัวเลือก 5 ตัวเลือกในแต่ละข้อ

วิธีการให้คะแนนแบบที่ 1 แบบเลือกตอบเสรี

| จำนวนคำตอบที่ทำเครื่องหมาย | ตอบถูกได้คะแนน | ตอบผิดได้แนน |
|---|---|---|
| 1 | 4 | -1 |
| 2 | 3 | -2 |
| 3 | 2 | -3 |
| 4 | 1 | -4 |
| 5 | 0 | |

วิธีให้คะแนนแบบที่ 2 แบบบอกระดับความมั่นใจ

| ระดับความมั่นใจ | ตอบถูกได้คะแนน | ตอบผิดได้คะแนน |
|---|---|---|
| แน่ใจ | 4 | -4 |
| ค่อนข้างแน่ใจ | 3 | -3 |
| เดาอย่างมีเหตุผล | 2 | -2 |
| ไม่แน่ใจ | 1 | -1 |

Dressel และ Schmid พบว่าการตอบข้อกระทงของเด็กเก่ง ซึ่งพิจารณาจากคะแนนการสอบแตกต่างกันนักเรียนประเภทปานกลาง และอ่อนอย่างมีนัยสำคัญ เมื่อใช้แบบสอบเลือกตอบเสรี คือกลุ่มเก่งจะทำเครื่องหมายคำตอบบนตัวเลือกต่าง ๆ จำนวนน้อยกว่ากลุ่มอ่อน ทั้งข้อกระทงที่มีระดับความยากสูง ปานกลาง และต่ำ ในทางตรงข้ามแบบสอบแบบบอกระดับความมั่นใจไม่ได้จำแนกผู้สอบกลุ่มเก่ง ปานกลาง และอ่อนได้ดีเท่าวิธีแรก เพราะเมื่อพิจารณาการตอบข้อกระทงแล้วพบว่าผู้สอบตอบข้อกระทงที่มีความยากปานกลางและยากมากด้วยระดับความมั่นใจเดียวกัน Dressel และ Schmid ได้สรุปว่าองค์ประกอบของความมั่นใจที่วัดโดยแบบสอบชนิดเลือกตอบเสรีแตกต่างจากที่วัดได้โดยแบบสอบที่บอกระดับความมั่นใจ

Shih-Sung Wen(Shih-Sung Wen 1975:197-199)ศึกษาสหสัมพันธ์ระหว่างคะแนนที่นับจากผลรวมของจำนวนข้อที่ตอบถูกกับคะแนนความมั่นใจ พบว่ามีสหสัมพันธ์กันสูงเท่ากับ 0.97 ซึ่งในทางปฏิบัติแล้วคะแนนที่ได้จากทั้งสองวิธีนี้สามารถนำมาจัดอันดับ ( rank order ) ได้เหมือนกันแต่เขาให้ข้อคิดว่าถึงแม้สหสัมพันธ์ระหว่างคะแนนที่ได้รับจากสองวิธีนี้สูง แต่ก็ไม่ได้เป็นข้ออ้างที่จะปฏิเสธวิธีให้คะแนนความมั่นใจ เพราะถึงอย่างไรวิธีการให้คะแนนความมั่นใจก็ให้รายละเอียดหรือข้อมูลในการตอบข้อกระทงที่มากกว่า

Pughและ Brunza(Pugh and Brunza 1975:76-78) ทำการศึกษาหาค่าความเที่ยงของแบบสอบที่มีวิธีการตอบและตรวจให้คะแนน 2 วิธี คือวิธีตอบแบบธรรมดาและวิธีตอบโดยให้คะแนนความมั่นใจในการเลือกคำตอบของข้อกระทง โดยใช้แบบสอบเลือกตอบเป็นแบบวัดสติปัญญา ชื่อ I.E.R. Intelligent Scale จำนวน 48 ข้อ แบ่งออกเป็น 3 กลุ่ม ๆ ละ 28 คน ให้แต่ละกลุ่มได้รับแบบสอบที่แตกต่างกันดังนี้

กลุ่มแรกได้รับแบบสอบชุด เอ ซึ่งมีวิธีตอบแบบธรรมดาทั้งสองส่วน

กลุ่มที่สองได้รับแบบสอบชุด บี ซึ่งมีวิธีตอบแบบธรรมดาในส่วนที่หนึ่ง และวิธีตอบแบบให้คะแนนความมั่นใจในส่วนที่สอง

กลุ่มที่สามได้รับแบบสอบชุด ซี ซึ่งมีวิธีตอบโดยบอกความมั่นใจในการตอบทั้งสองส่วน

เขาได้หาความเที่ยงของแบบสอบโดยวิธีวิเคราะห์ความแปรปรวนปรากฏว่าแบบสอบที่มีวิธีตอบโดยบอกความมั่นใจมีความเที่ยงสูงกว่าแบบสอบที่มีวิธีตอบแบบธรรมดาคือ 0.85 และ 0.57 ตามลำดับ

Abu Sayf (Abu-Sayf 1975:160-161) ได้ทำการศึกษาเกี่ยวกับการตอบโดยบอกระดับความมั่นใจ โดยให้นักเรียนมัธยมศึกษาจำนวน 152 คน ตอบแบบสอบปรนัยเลือกตอบชนิด 4 ตัวเลือกวิชาภาษาอังกฤษจำนวน 60 ข้อ โดยให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุดเพียงคำตอบเดียวและบอกระดับความมั่นใจในการตอบว่าตอบด้วยความมั่นใจ ไม่ค่อยมั่นใจหรือเดา แล้วเปรียบเทียบเปอร์เซ็นต์ในการตอบถูกหรือผิด ผลการศึกษาปรากฏว่าจากคำตอบทั้งหมดที่นักเรียนบอกว่าตอบด้วยความมั่นใจ 3491 ข้อ มีนักเรียนตอบผิด 846 ข้อ หรือ 24.23% คำตอบที่นักเรียนบอกว่าตอบโดย

ไม่ค่อยมั่นใจ 2648 ข้อ มีผิด 1683 ข้อ หรือ 61.92% และคำตอบที่นักเรียนบอกว่าตอบโดยการเดา 2782 มีนักเรียนตอบผิด 2045 ข้อ หรือ 73.31%

จะเห็นว่าการตอบโดยอาศัยการเดานั้นมีโอกาสถูกน้อยมาก การเปิดโอกาสให้นักเรียนตอบแบบสอบโดยอาศัยการเดาด้วยไม่ได้ช่วยให้นักเรียนได้คะแนนเพิ่มขึ้นมากนัก

Abu-Sayf และ Diamond(Abu-Sayf and Diamond 1976:62-63) ทำการศึกษาเกี่ยวกับค่าความเที่ยงและค่าความตรงของแบบสอบ เมื่อให้นักเรียนตอบแบบสอบปรนัยชนิดเลือกตอบโดยบอกระดับความมั่นใจในการตอบ จากข้อมูลที่ Abu-Sayf ได้ทดลองไว้แล้วจากกลุ่มตัวอย่าง 152 คน ใช้แบบสอบวิชาภาษาอังกฤษแบบปรนัยชนิดเลือกตอบ 4 ตัวเลือกจำนวน 60 ข้อ โดยการตรวจให้คะแนน 3 วิธีคือให้คะแนนเฉพาะข้อถูกและตอบด้วยความมั่นใจ ให้คะแนนข้อที่ถูกและไม่ได้เดา และให้คะแนนทุกข้อที่ตอบถูกโดยไม่คำนึงว่ามีความมั่นใจในการตอบหรือไม่เพียงใด แล้วหาค่าความเที่ยงของแบบสอบจากสัมประสิทธิ์อัลฟา ผลปรากฏว่าการให้คะแนนเฉพาะข้อที่ตอบถูกและตอบด้วยความมั่นใจมีค่าความเที่ยงสูงสุดคือ 0.898 ส่วนการให้คะแนนข้อที่ตอบถูกโดยไม่ได้เดามีค่าความเที่ยงสูงกว่าการให้คะแนนทุกข้อที่ตอบถูกคือมีค่าความเที่ยงเป็น 0.871 และ 0.774 ตามลำดับ

ส่วนค่าความตรงนั้นวิธีการให้คะแนนข้อที่ตอบถูกและไม่ได้เดามีค่าความตรงสูงสุด คือ 0.260 ส่วนการให้คะแนนเฉพาะข้อที่ตอบถูกด้วยความมั่นใจมีค่าความตรงสูงกว่าการให้คะแนนทุกข้อที่ตอบถูกคือ 0.236 และ 0.219 ตามลำดับ

Hanna (Hanna 1975 : 145 ) ศึกษาเกี่ยวกับคุณภาพของแบบสอบที่มีวิธีการให้คะแนนความมั่นใจ โดยใช้แบบสอบเลือกตอบชนิด 4 ตัวเลือกวิชาวิทยาศาสตร์จำนวน 2 ฉบับ ซึ่งยาว 10 ข้อและ 30 ข้อ ทดสอบกับนักศึกษาในระดับบัณฑิตศึกษาจำนวน 85 คน ผลการศึกษาพบว่า ค่าความเที่ยงของแบบสอบที่คำนวณโดยวิธีแบ่งครึ่งของแบบสอบที่มีวิธีให้คะแนนความมั่นใจ มีค่าความเที่ยงสูงกว่าค่าความเที่ยงของแบบสอบที่ให้คะแนนวิธีธรรมดาเล็กน้อย แต่ค่าความตรงเชิงพยากรณ์ ซึ่งใช้คะแนนสอบการตีความและการเขียนรายงานเป็นเกณฑ์ของแบบสอบที่มีวิธีให้คะแนนความมั่นใจมีค่าต่ำกว่าวิธีให้คะแนนวิธีธรรมดา

Hakstian และ Kansup (Hakstian and Kansup 1975 : 231) ทำการศึกษาผลการให้คะแนนความมั่นใจ และการให้คะแนนโดยวิธีการกำจัดตัวเลือก โดยใช้นักเรียนระดับ grade 9 จำนวน 1,023 คน เป็นกลุ่มตัวอย่าง ผลการศึกษาพบว่า ค่าความเที่ยงของแบบสอบที่ได้จากการให้คะแนนความมั่นใจและการให้คะแนนโดยวิธีการกำจัดตัวเลือกมีค่าสูงกว่าวิธีธรรมดาเล็กน้อย แต่ค่าความตรงเชิงพยากรณ์ของแบบสอบที่มีวิธีการให้คะแนนทั้ง 2 วิธีดังกล่าวมีค่าต่ำกว่าวิธีให้คะแนนวิธีธรรมดา

Hanna (Hanna 1977 : 643 ) ทำการศึกษาวิธีการให้คะแนนความมั่นใจอีกครั้งโดยใช้แบบสอบเลือกตอบวิชาวิทยาศาสตร์และวิชาคณิตศาสตร์ จำนวน 2 ฉบับ มีความยาวฉบับละ 18 ข้อ และใช้นักเรียน grade 5 และ 6 จำนวน 1,341 คน เป็นกลุ่มตัวอย่าง ผลการศึกษาปรากฏว่า มีความสอดคล้องกับผลการศึกษาในปี 1974 คือ แบบสอบมีค่าความเที่ยงสูงกว่าค่าความเที่ยงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยมอย่างมีนัยสำคัญ แต่ให้ค่าความตรงไม่แตกต่างจากค่าความตรงที่คำนวณจากคะแนนที่ให้โดยวิธีประเพณีนิยม

เพ็ญศรี สว่างเนตร (2520:41-42) ศึกษาค่าความเที่ยงของแบบสอบเลือกตอบด้วยเทคนิคการให้คะแนนต่างกัน เครื่องมือที่ใช้เป็นแบบสอบชนิด 5 ตัวเลือกซึ่งเป็นแบบสอบวัดผลสัมฤทธิ์วิชาภาษาอังกฤษ นำไปสอบกับนักศึกษาระดับประกาศนียบัตรวิชาการศึกษาชั้นสูงที่เรียนวิชาภาษาอังกฤษเป็นวิชาเอกและโทจำนวน 216 คน พร้อมกระบวนการทดสอบความมั่นใจจากเทคนิคการให้คะแนน 4 วิธีคือ วิธีธรรมดา วิธีให้คะแนนความมั่นใจ วิธีลดคะแนนข้อที่ผิด วิธีเพิ่มคะแนนข้อที่เว้น คำนวณค่าความเที่ยงของแบบสอบโดยวิธีวิเคราะห์ความแปรปรวนของฮอยต์ (Hoyt's analysis of variance) ได้ข้อค้นพบว่าค่าความเที่ยง ภายใต้เทคนิคการให้คะแนนความมั่นใจสูงกว่าค่าความเที่ยงที่ได้จากวิธีธรรมดาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 คือมีค่าความเที่ยง .61 และ .55 ตามลำดับ ค่าความเที่ยงโดยวิธีลดคะแนนข้อที่ผิดเพิ่มขึ้นกว่าวิธีธรรมดาแต่ไม่มีนัยสำคัญทางสถิติ ส่วนวิธีการให้คะแนนข้อที่เว้นไม่ทำให้ค่าความเที่ยงของแบบสอบเพิ่มขึ้น

ชีรศักดิ์ อินทรมาตย์ (2520:70-73) ศึกษาผลของวิธีตอบแบบสอบปรนัยชนิดเลือกตอบที่มีวิธีการให้คะแนน 3 วิธีคือวิธีธรรมดา วิธีให้คะแนนความมั่นใจ และวิธีตอบทุกตัวเลือกที่มีต่อค่าความเที่ยง ความตรงและปริมาณการเดาของแบบสอบวิชาวิทยาศาสตร์กับกลุ่มตัวอย่างที่เป็นนักเรียนชั้นประถมปีที่ 7 ผลปรากฏว่าแบบสอบที่มีวิธีการให้คะแนนความมั่นใจมีค่าความเที่ยงสูงกว่าวิธีการตอบแบบอื่น ๆ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนวิธีตอบทุกตัวเลือกให้ค่าความเที่ยงสูงกว่าวิธีธรรมดาอย่างไม่มีนัยสำคัญทางสถิติ ด้านความตรงของแบบสอบพบว่าวิธีการให้คะแนนทั้งสามวิธีมีค่าความตรงของแบบสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ สำหรับปริมาณการเดาพบว่าแบบสอบที่มีวิธีการให้คะแนนแตกต่างกันทั้งสามวิธีมีปริมาณการเดาแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แบบสอบที่ให้ตอบทุกตัวเลือกมีปริมาณการเดาต่ำสุด และนักเรียนที่มีความสามารถในระดับต่ำจะตอบแบบลอบโดยการเดามากกว่านักเรียนที่มีความสามารถในระดับสูงทั้งสามวิธี

เกริกชัย ฮวบเจริญ (2525:36-37) ศึกษาเปรียบเทียบค่าความเที่ยงและค่าคะแนนการเดาของแบบสอบเลือกตอบที่มีวิธีการให้คะแนน 3 วิธีคือการให้คะแนนวิธีธรรมดาวิธีให้คะแนนความมั่นใจ และวิธีตอบและตรวจให้คะแนนแบบใหม่โดยใช้กระดาษคำตอบชนิดพิเศษ เมื่อตอบแล้วผู้สอบจะทราบทันทีว่าตนทำถูกหรือผิดในข้อนั้น ๆ และถ้าทำผิดผู้สอบสามารถตัดสินใจอย่างเสรีว่าจะทำข้อเดิมต่อไปหรือไม่หรือจะทำจนกระทั่งถูก ผลการศึกษาพบว่าแบบสอบเลือกตอบเมื่อใช้วิธีตอบและตรวจให้คะแนนด้วยวิธีใหม่มีค่าความเที่ยงเป็น .894 เมื่อตอบและตรวจให้คะแนนความมั่นใจมีค่าความเที่ยงเป็น .885 เมื่อตอบและตรวจให้คะแนนวิธีธรรมดามีค่าความเที่ยงเป็น .716 ผลการทดสอบความแตกต่างของค่าความเที่ยงพบว่าแบบสอบที่มีวิธีการตอบและตรวจให้คะแนนแบบใหม่และแบบวิธีให้คะแนนความมั่นใจมีค่าความเที่ยงไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ สำหรับค่าคะแนนการเดาพบว่าเมื่อใช้วิธีตอบและตรวจให้คะแนนวิธีใหม่และวิธีให้คะแนนความมั่นใจให้ค่าการเดาของผู้สอบน้อยกว่าวิธีตอบแบบธรรมดาอย่างมีนัยสำคัญทางสถิติ แต่วิธีตอบและตรวจให้คะแนนแบบใหม่และวิธีให้คะแนนความมั่นใจให้ค่าการเดาของผู้สอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

จริยา จงนาบุรักษ์ (2527:46-47) ศึกษาถึงความสามารถในการเรียนรู้วิชาภาษาอังกฤษของนักเรียนมัธยมศึกษาปีที่ 5 ที่ได้รับการสอบย่อยและการให้ข้อมูลย้อนกลับจากแบบสอบเลือกตอบที่มีวิธีการให้คะแนนวิธีธรรมดาและการให้คะแนนความมั่นใจในกลุ่มผู้สอบที่มีความสามารถในการเรียนรู้วิชาภาษาอังกฤษต่างกัน ผลการวิจัยพบว่าผู้สอบที่ได้รับการสอบย่อยด้วยแบบสอบเลือกตอบวิธีธรรมดาและวิธีการให้คะแนนความมั่นใจมีความสามารถในการเรียนรู้ภาษาอังกฤษสูงกว่าผู้สอบที่ไม่ได้รับการสอบย่อยอย่างมีนัยสำคัญทางสถิติ ในกรณีไม่แยกระดับความสามารถของผู้สอบพบว่าผู้สอบที่มีการสอบย่อยด้วยแบบสอบเลือกตอบวิธีธรรมดาและวิธีให้คะแนนความมั่นใจมีความสามารถในการเรียนรู้วิชาภาษาอังกฤษไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ แต่เมื่อแยกระดับความสามารถทางการเรียนภาษาอังกฤษพบว่าผู้สอบที่มีระดับความสามารถทางการเรียนวิชาภาษาอังกฤษต่ำที่ได้รับการสอบย่อยด้วยแบบสอบธรรมดาและวิธีให้คะแนนความมั่นใจมีความสามารถในการเรียนรู้ภาษาอังกฤษแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ส่วนผู้ที่มีความสามารถทางการเรียนภาษาอังกฤษสูงและปานกลางมีความสามารถในการเรียนรู้ภาษาอังกฤษไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย