

## บทที่ 2

### วรรณคดีที่เกี่ยวข้อง

ผู้วิจัยได้แบ่งวรรณคดีที่เกี่ยวข้องออกเป็น 2 ตอน คือ ตอนที่ 1 เป็นมโนทัศน์เบื้องต้นเกี่ยวกับชนิดของแบบสอบ แบบสอบแบบเลือกตอบ แบบสอบแบบถูกผิด คุณภาพของแบบสอบ ความเที่ยง ความตรง ความยาก และอำนาจจำแนก ส่วนตอนที่ 2 เป็นงานวิจัยเกี่ยวกับการเปรียบเทียบค่าความเที่ยง ความยาก ความตรง และอำนาจจำแนกของแบบสอบแบบเลือกตอบ กับแบบสอบแบบถูกผิดทั้งในประเทศและต่างประเทศ

ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับ ชนิดของแบบสอบ แบบสอบแบบเลือกตอบ แบบสอบแบบถูกผิด คุณภาพของแบบสอบ ความเที่ยง ความตรง ความยาก และอำนาจจำแนก

#### ชนิดของแบบสอบ

การแบ่งชนิดของแบบสอบโดยใช้วิธีตอบเป็นเกณฑ์ แบ่งออกได้เป็น 3 ชนิด

(เอนก เพียรอนุกุลบุตร 2524: 148) คือ

1. แบบสอบความเรียง (Essay Test) หรือแบบสอบอัตนัย หรือบรรยาย มีลักษณะที่สำคัญ คือ ผู้ตอบต้องเขียนตอบยาว ๆ มีน้อยข้อ

2. แบบสอบปรนัย (Objective Test) เป็นชนิดที่ผู้ตอบตอบเพียงสั้น ๆ หรือเพียงเขียนเครื่องหมายสัญลักษณ์ใด ๆ ตอบเท่านั้น มีหลายรูปแบบคือ

ก. ชนิดเลือกตอบ (Multiple Choices Type) ให้เลือกข้อถูกที่สุด

ข. ชนิดเติมคำ (Completion Type) ให้เติมคำหรือข้อความให้ถูกต้อง

ค. ชนิดตอบสั้น ๆ (Short Response Type) ให้ตอบคำถามสั้น ๆ

ง. ชนิดจับคู่ (Matching Type) จับคู่คำตอบกับคำถามให้ถูกต้อง

จ. ชนิดถูกผิด (True-False Type) ให้เลือกตอบทางใดทางหนึ่ง เช่น

ถูกหรือผิด จริงหรือไม่จริง ใช่หรือไม่ใช่

3. แบบสอบให้ปฏิบัติ (Performance test) เป็นแบบสอบที่ใช้วัดงานภาคปฏิบัติ ส่วนใหญ่เป็นแบบให้ปฏิบัติ ให้จัด ให้ทำงานที่กำหนดให้ เช่น แบบสอบภาคปฏิบัติพลานามัย ศิลปปฏิบัติ เป็นต้น

#### แบบสอบแบบ เลือกตอบ



เป็นข้อสอบที่ประกอบด้วยส่วนใหญ่ ๆ 2 ส่วนคือ

1. ตัวคำถามนำ หรือตัวปัญหา (Stem) ซึ่งเขียนได้ 3 แบบ
  - ก. เขียนเป็นรูปคำถาม
  - ข. เขียนเป็นข้อความที่ไม่สมบูรณ์
  - ค. เขียนเป็นข้อความที่สมบูรณ์
2. ตัวเลือกหรือตัวคำตอบ (Options) แยกเป็น 2 ประเภทคือ
  - ก. คำตอบ (Keyed Response)
  - ข. ตัวลวงหรือ ตัวหลอกล่อ (Distracter)

ธอร์นไดค์ (Thorndike 1971 : 113-119) ได้กล่าวถึงคุณสมบัติของแบบสอบชนิด เลือกตอบที่ดี ไว้ดังนี้

1. เขียนตัวคำถามให้เป็นรูปคำถามที่สมบูรณ์
2. คำใดที่มีอยู่ในตัวเลือกซ้ำ ๆ กัน ควรดึงขึ้นมาไว้ในตัวคำถาม
3. หลีกเลี่ยงคำถามที่เป็นปฏิเสธ
4. ตัวเลือกที่ใช้ควรเหมาะสม สมเหตุสมผล และสอดคล้องกัน
5. ตัวเลือกทุกตัวต้องสอดคล้องกับคำถาม
6. ให้ตัวลวงทุกตัว ดึงดูดความสนใจของผู้ตอบที่ไม่รู้จริง
7. หลีกเลี่ยงการใช้ตัวเลือกที่เป็นศัพท์เทคนิคที่สูงเกินไป
8. หลีกเลี่ยงตัวเลือกที่เหลื่อมล้ำ (Overlap) กัน
9. ไม่ควรใช้คำว่า "ข้างบนทั้งหมด" และ "ไม่ใช่ข้างบนทั้งหมด"
10. เรียงลำดับตัวเลือก ตามลำดับตรรกวิทยา (Logical Order)
11. ถ้าคำถามเกี่ยวกับคำจำกัดความของคำ ควรอธิบายความหมายไว้ที่ตัวคำถาม

แล้วให้ตัวเลือก เป็นคำศัพท์

## 12. ไม่เขียนตัวเลือก เป็นลักษณะคล้ายแบบถูกผิด

อนันต์ ศรีโสภา (2520: 130-134) ได้เสนอแนะหลักการเขียนข้อสอบชนิด เลือกตอบไว้ดังนี้

1. ในตัวคำถามควรมีใจความที่สำคัญบรรจุ เพื่อให้ นักเรียนได้เกิดแนวความคิดก่อนที่จะอ่านถึงตัวเลือกต่าง ๆ โดยทั่วไปมักจะทำข้อความในตัวปัญหา มากกว่าข้อความในตัวเลือกแต่ละตัว

2. ควรหลีกเลี่ยงการใช้คำที่ซ้ำกันในพวกตัวเลือกต่าง ๆ
3. เมื่อเราใช้ข้อความที่ไม่สมบูรณ์ ส่วนที่ขาดนั้นควรจะเป็นตัวเลือกต่าง ๆ
4. ควรจัดเรียงตัวเลือกต่าง ๆ ให้ง่ายและเป็นระเบียบ
5. ควรหลีกเลี่ยงการใช้คำศัพท์ที่ยากเกินไป
6. ตัวเลือกทุกตัวควรจะเป็นไปได้ และมีความเป็นเอกพันธ์ ตัวลวงที่จัดว่ามีประสิทธิภาพดี จะต้องเป็นตัวลวงที่สามารถลวงนักเรียนที่มีความสามารถผิด ๆ ได้ดี
7. ควรหลีกเลี่ยงการทำให้คำตอบที่ถูกต้องยาวกว่าตัวลวงอื่น ๆ
8. ควรหลีกเลี่ยงสิ่งทีบอกร่องรอยในการตอบที่ถูกต้องอย่างอื่น ๆ
9. ควรหลีกเลี่ยงตัวเลือกที่ว่าไม่ทราบเสีย
10. ข้อสอบแต่ละข้อควรมีตัวเลือกที่ถูกต้องหรือถูกที่สุด เพียงตัวเดียว
11. ควรหลีกเลี่ยงการใช้ตัวเลือก "ถูกทั้งหมด"
12. ควรใช้ "ไม่มีข้อถูก" เป็นตัวเลือกในบางสถานการณ์
13. ควรให้มีจำนวนตัวเลือก 4 หรือ 5 ตัวเลือก
14. ควรหลีกเลี่ยงการใช้ตัวเลือกที่ข้อความเกี่ยวข้องกัน
15. ควรสร้างข้อสอบให้สามารถวัดขอบวนการคิดทางสมองในระดับสูง

### แบบสอบแบบถูกผิด

เป็นแบบสอบซึ่งจำกัดการตอบสนองของผู้สอบโดยบังคับให้ผู้สอบ เลือกตอบ เพียงอย่างใดอย่างหนึ่งจากสองสิ่งที่กำหนดให้ เช่น ถูก-ผิด, ใช่-ไม่ใช่, เห็นด้วย-ไม่เห็นด้วย, ข้อเท็จจริง-ความเห็น เหมือนกัน-ต่างกัน (เสริมศักดิ์ วิชาลาภรณ์ และ เอนกกุล กริแสง 2522: 93)

เยาวตี วิบูลย์ศรี (2526: 112-114) ได้กล่าวถึงลักษณะข้อทดสอบแบบถูกผิดที่ดีไว้ดังนี้

1. ข้อความที่ว่าถูกหรือผิด ควรประกอบด้วยเรื่องสำคัญ ที่จะถามเพียงเรื่องเดียว
2. ข้อความที่ถาม จะเป็นจริงหรือไม่จริง ถูกหรือผิดนั้นควรเป็นไปตามเหตุผล

ข้อเท็จจริงที่ทุกคนเห็นตรงกัน ไม่ควรเป็นข้อความที่เป็นความคิดเห็น หรือการวิเคราะห์เหตุการณ์ทั่วไป

3. คำหรือภาษาที่ใช้ในข้อความควรจะชัดเจนและสั้น หลีกเลี่ยงการขยายความจนยืดยาว

4. ใช้ข้อความที่เป็นคำเชิงปฏิเสธที่จำเป็นเท่านั้น และเมื่อใช้ก็ควรจะชัดเจนได้คำเชิงปฏิเสธให้เห็นชัดเจน

5. ควรหลีกเลี่ยงการใช้ข้อความที่เป็นปฏิเสธซ้อนปฏิเสธ

6. หลีกเลี่ยงการใช้คำว่า "ทุก ๆ" "ทั้งหมด" "ทั้งสิ้น" "เสมอ" "ปกติ"

หรือคำในลักษณะเดียวกัน ทั้งนี้เพราะผู้ตอบอาจจะใช้ความรู้บางส่วนมาช่วย เดาหาคำตอบที่ถูกต้องได้

7. ไม่ควรคัดลอกข้อความจากส่วนใดส่วนหนึ่งในตำราโดยตรง

8. การเขียนข้อความเพื่อนำมาใช้เป็นข้อทดสอบถูกผิดนั้นควรเขียนข้อความถูกต้องแล้วจึงเปลี่ยน เป็นข้อความที่ผิด

9. การจัดเรียงตำแหน่งของข้อทดสอบที่ถูกหรือผิดนั้น ควรจะเรียงอยู่ในลักษณะสุ่มหรือเรียงอย่างไม่เป็นระบบ

เสริมศักดิ์ วิศาลาภรณ์ และเอนกกุล กริแสง (2522: 97-100) ได้ให้คำแนะนำในการสร้างข้อสอบแบบถูกผิด ไว้ดังนี้

1. พยายามใช้ข้อความธรรมดา และชัดเจน ที่ทุกคนสามารถเข้าใจได้อย่างเดียวกัน และตรงกัน

2. พยายามให้ข้อมูลและขยายข้อความพอที่จะตัดสินได้ว่าข้อความนั้นถูกหรือผิด

3. การถูกหรือผิดของข้อสอบนั้นให้เป็นสากลคือผู้รู้มีความเห็นตรงกัน

4. พยายามให้ข้อสอบแต่ละข้อมีความคิดเดียว พยายามเลี่ยงประโยคที่ซับซ้อนและมีหลายความคิด

5. พยายามใช้ข้อความที่แสดงปริมาณมากกว่าข้อความที่แสดงคุณภาพ เพราะข้อสอบที่แสดงคุณภาพตัดสินได้ลำบากกว่าถูกหรือผิด

6. สิ่งที่กำหนดว่าผิดหรือถูกนั้น ควรเป็นส่วนที่สำคัญของข้อความ และ เกี่ยวข้องกับข้อความที่ถาม
7. ไม่ควรให้มีคำบางคำ เป็น เครื่องชี้คำตอบหรือช่วยให้คำตอบถูกหรือผิด เด่นชัดขึ้น
8. ถ้าไม่จำเป็นอย่าลอกข้อความจากหนังสือ เพราะจะทำให้ นักเรียนใช้ความจำอย่างเดียว
9. พยายามใช้ภาษาที่ง่าย ตรงไปตรงมา อย่าใช้คำถามที่ยกย่อนจนเกินไป
10. พยายามหลีกเลี่ยงคำถามที่กำกวม
11. ถ้าไม่จำเป็นอย่าใช้ประโยคที่มีรูปปฏิเสธ
12. คำถามที่อ้างถึงความ เห็น กฎเกณฑ์ หรือข้อตกลงใดที่ยังหาข้อยุติไม่ได้ ควรระบุต้นคิดหรือแหล่งที่มาด้วย
13. พยายามหลีกเลี่ยงข้อความที่เป็นคำสั่ง เพราะคำสั่งนั้นบอกไม่ได้ว่าถูกหรือผิด
14. พยายามเลี่ยงข้อความที่กว้างเกินไป ควรเขียนให้รัดกุมแต่ชัดเจน
15. พยายามให้ข้อสอบแต่ละข้อเป็นอิสระแก่กัน อย่าให้ข้อสอบข้อใดข้อหนึ่งเป็น เครื่องแนะคำตอบข้ออื่น
16. พยายามให้ข้อสอบแต่ละข้อยาวเท่ากัน อย่าให้ความยาวหรือสั้นของข้อสอบเป็น เครื่องแนะคำตอบ
17. พยายามเลี่ยงแบบฉบับของคำตอบ
18. จำนวนข้อผิดและข้อถูกควรใกล้เคียงกัน หรืออาจอยู่ระหว่าง 40%-60%

#### คุณภาพของแบบสอบ

เนื่องจากผลการสอบมีความสำคัญยิ่งต่อผู้สอบจึงมีความจำเป็นที่ผู้สร้างแบบสอบจะต้องทราบถึงคุณลักษณะของแบบสอบที่ดี อันจะเป็นแนวทางในการสร้างแบบสอบ และใช้เป็นเกณฑ์ในการวินิจฉัยว่า แบบสอบที่สร้างขึ้นนั้นมีคุณภาพดี-ไม่ดีอย่างไร เพื่อจะได้ปรับปรุงแก้ไขให้เป็นแบบสอบที่มีประสิทธิภาพยิ่งขึ้น

วิเชียร เกตุสิงห์ (2517: 27-32) ได้กล่าวถึงลักษณะของแบบสอบที่ดีไว้ 10 ประการ คือ

1. ความตรง (Validity) หมายถึง คุณลักษณะของข้อสอบที่สามารถวัดได้ตรง  
จุดมุ่งหมายที่ตั้งไว้
2. ความเที่ยง (Reliability) แบบสอบที่มีความเที่ยงหมายถึง แบบสอบที่สามารถ  
วัดได้แน่นอนไม่เปลี่ยนแปลงไปเปลี่ยนมา การวัดครั้งแรกเป็นอย่างใด เมื่อวัดซ้ำอีกครั้งหนึ่งหรือหลาย ๆ  
ครั้งก็ตาม ผลการวัดก็ยังคงเหมือนเดิม
3. อำนาจจำแนก (Discrimination) คือลักษณะของแบบสอบที่สามารถจะแยก  
หรือจำแนกเด็กเก่ง-อ่อนได้
4. ความเป็นปรนัย (Objectivity) ข้อสอบที่จะเป็นปรนัยได้ประกอบด้วยคุณสมบัติ  
3 ประการ คือ 1) มีความชัดเจนในความหมายของคำถาม 2) มีความคงที่ในการตรวจให้คะแนน  
3) มีความแจ่มชัดในการแปลความหมายของคะแนน
5. ความมีประสิทธิภาพ (Efficiency) หมายถึงลักษณะข้อสอบที่มีคุณสมบัติแสดงถึง  
การประหยัดเศรษฐกิจ เช่น ลงทุนน้อย ราคาถูก ฯลฯ แต่ใช้เวลาสอบน้อยดังนี้ เป็นต้น
6. การวัดอย่างลึกซึ้ง (Searching) หมายถึงลักษณะของข้อสอบที่ถามครอบคลุม  
พฤติกรรมหลาย ๆ ด้าน ไม่ใช่วัดแต่ความรู้ความจำอย่างเดียว
7. ข้อสอบที่ดีต้องยุติธรรม (Fair) ไม่ถามหลายแง่หลายมุมข้อเดียวพยายามถาม  
เพียงอย่างเดียว
8. ข้อสอบที่ดีต้องมีความยากง่าย (Difficulty) พอเหมาะ
10. ข้อสอบที่ดีต้องมีการกระตุ้น ยุหน่ย (Exemplary) มีการถามล่อ โดยจัดเอาข้อสอบ  
ง่าย ๆ ไว้ในตอนแรก ๆ แล้วจึงค่อย ๆ ถามให้ยากขึ้นตามลำดับ

ในการวิจัยครั้งนี้ผู้วิจัยได้เลือกคุณลักษณะที่ดีของแบบสอบมาเพื่อพิจารณาและเปรียบเทียบ  
คุณภาพระหว่างแบบสอบ 3 ชุด รวม 4 ด้านคือ ความเที่ยง ความตรง ความยาก และอำนาจ  
จำแนก เนื่องจากสามารถคำนวณเป็นค่าตัวเลข และใช้เทคนิคทางสถิติทดสอบความแตกต่างได้  
ดังจะกล่าวรายละเอียดต่อไปนี้

#### ความเที่ยงของแบบสอบ

ได้มีผู้ให้ความหมายของความเที่ยงไว้มากมาย เช่น

ความเที่ยงหมายถึง ความคงเส้นคงวาของการวัด (Gronlund 1968: 105)

ความเที่ยงหมายถึง ระดับ (Degree) ของความคงที่ระหว่างการวัดสิ่งเดียวกัน 2 ครั้ง (Mehrens and Lehmann 1973: 88)

ความเที่ยงหมายถึง ความคงเส้นคงวาของชุดคะแนนจากแบบสอบที่วัดในสิ่งที่ต้องการวัด (Ebel 1965: 310)

ความเที่ยงหมายถึง ความคงที่ของคะแนนที่ได้จากการทดสอบคนกลุ่มเดียวกัน 2 ครั้ง ด้วยแบบสอบเดิมใน เวลาต่างกัน หรือ ด้วยข้อสอบคนละชุดที่มีข้อเทียบ เท่ากันหรือภายใต้สภาพ การสอบที่ต่างกัน (อนาสตาซี 2519: 73)

จากนิยามความเที่ยงดังกล่าวข้างต้น ก็พอสรุปได้ว่า ความเที่ยงหมายถึง ความคงที่ของคะแนนที่ได้จากการวัดด้วย เครื่องมือชุด เดียวกันในโอกาสที่ต่างกัน

#### ทฤษฎีของความเที่ยง (Theory of Reliability)

การอธิบายทฤษฎีความเที่ยงโดย เริ่มจากคะแนนสอบที่นักเรียนทำได้ (Observed Scores) ซึ่งประกอบด้วยส่วนที่เป็นคะแนนจริง (True Scores) และส่วนที่เป็นคะแนนที่เกิดจากความคลาดเคลื่อน (Error Scores) ซึ่งเขียนเป็นสมการได้ดังนี้

$$X = T + E$$

เมื่อ X = คะแนนที่นักเรียนทำได้

T = คะแนนจริง

E = คะแนนที่เกิดจากความคลาดเคลื่อน

คะแนนจริง (True Scores) หมายถึงคะแนนที่ผู้สอบพึงได้รับจากการวัดด้วย เครื่องมือ ที่มีคุณลักษณะสมบูรณ์ปราศจากความคลาดเคลื่อน หรือหมายถึงคะแนน เฉลี่ยของผู้สอบซึ่งได้จากการ ทำแบบสอบ เดิมหลาย ๆ ครั้ง โดยมีข้อตกลงว่า ไม่มีอิทธิพลจากการฝึกฝน ความเมื่อยล้า และ การเรียนรู้ในการทดสอบซ้ำ (Brown 1976: 50)

คะแนนที่เกิดจากความคลาดเคลื่อน (Error Scores) หมายถึง ค่าความผิดพลาด ที่เกิดขึ้นจากการวัด ซึ่งเป็นความคลาดเคลื่อนที่เกิดขึ้นในลักษณะสุ่ม (Random Error) เช่น ความคลาดเคลื่อนที่เกิดขึ้นจากตัวแบบสอบ การบริหารการสอบ การให้คะแนน และสิ่งต่าง ๆ ที่เกิดจากตัวผู้เข้าสอบ เป็นต้น โดยทฤษฎีความคลาดเคลื่อนที่เกิดขึ้นในลักษณะสุ่มนี้ อาจเป็นไปได้



ทั้งทางบวกและทางลบ ซึ่งจะหักล้างกันหมดไป และถ้าความคลาดเคลื่อนเกิดขึ้นในลักษณะสุ่มแล้ว ก็จะไม่มีความสัมพันธ์กับคะแนนจริง นั่นคือสัมประสิทธิ์สหสัมพันธ์ (Correlation) ระหว่างคะแนนจริงกับคะแนนความคลาดเคลื่อนมีค่าเท่ากับ 0 ดังนั้นความแปรปรวนของคะแนนที่วัดได้สามารถเขียนได้ดังสมการ (Mehrens and Lehman 1973: 91) ดังนี้

$$S_X^2 = S_t^2 + S_e^2$$

เมื่อ  $S_X^2 =$  ความแปรปรวนของคะแนนที่วัดได้

$S_t^2 =$  ความแปรปรวนของคะแนนจริง

$S_e^2 =$  ความแปรปรวนของคะแนนความคลาดเคลื่อน



โดยทฤษฎี ความเที่ยง หมายถึงอัตราส่วนระหว่างความแปรปรวนของคะแนนจริง กับความแปรปรวนของคะแนนที่วัดได้ (Brown 1976: 51) จึงสามารถเขียนสมการได้เป็น

$$r_{tt} = \frac{S_t^2}{S_X^2} \quad \text{--- (1)}$$

เมื่อ  $r_{tt} =$  ค่าความเที่ยง

จากสมการค่าความเที่ยงจึงบอกให้เราทราบถึงความแปรปรวนของคะแนนจริง ว่าเป็นสัดส่วนเท่าใด เมื่อเทียบกับความแปรปรวนของคะแนนที่วัดได้ และจาก

$S_t^2 = S_X^2 - S_e^2$  ดังนั้น สมการ (1) อาจเขียนได้เป็น

$$r_{tt} = 1 - \frac{S_e^2}{S_X^2}$$

ซึ่งเป็นสมการพื้นฐานสำหรับการประมาณค่าความเที่ยง

#### วิธีประมาณค่าความเที่ยง

วิธีประมาณค่าความเที่ยงสามารถทำได้หลายวิธี แต่ละวิธีทำให้เกิดความคลาดเคลื่อนแตกต่างกัน การจะเลือกใช้วิธีใดนั้นขึ้นอยู่กับ ความมุ่งหมาย ลักษณะของคะแนน และชนิดของ



แบบสอบ (Guilford 1978: 414) โดยทั่วไปวิธีประมาณค่าความเที่ยงมี 4 แบบ (Mehrens and Lehmann 1973: 109) คือ

1. แบบสอบซ้ำ (Measure of Stability)
2. แบบใช้ข้อสอบคล้ายกัน (Measure of Equivalence)
3. แบบใช้ข้อสอบคล้ายกันและสอบซ้ำ (Measure of Equivalence and Stability)
4. แบบวัดความคงที่ภายใน (Measure of internal consistency)

ซึ่งแบ่งออกเป็น

- 4.1 แบบแบ่งครึ่งจำนวนข้อสอบ (Split-Half)
- 4.2 แบบคูเดอริชาร์ดสัน (Kuder-Richardson Estimates)
- 4.3 แบบวิเคราะห์ความแปรปรวนของชอยท์ (Hoyts Analysis of Variance Procedure)

ในที่นี้จะกล่าวเฉพาะความเที่ยงแบบคูเดอริชาร์ดสัน ซึ่งเป็นวิธีประมาณค่าความเที่ยงที่ใช้ในการวิจัยครั้งนี้

#### ความเที่ยงแบบคูเดอริชาร์ดสัน

ในปี ค.ศ. 1937 คูเดอริชาร์ดสัน (Kuder and Richardson) ได้คิดหาวิธีสำหรับหาความคงที่ภายในโดยไม่ต้องแบ่งครึ่งข้อสอบ เขาได้เสนอสูตรสำหรับประมาณค่าความเที่ยง ซึ่งเป็นที่รู้จักและนิยมใช้กันอย่างแพร่หลาย คือ สูตร คูเดอริชาร์ดสัน 20 (Kuder-Richardson 20) (Brown 1976: 78)

$$r_{tt} = \left( \frac{k}{k-1} \right) \left( \frac{S_t^2 - \sum p_i q_i}{S_t^2} \right)$$

เมื่อ $r_{tt}$	แทน	ค่าความเที่ยงของแบบสอบ
$k$	แทน	จำนวนข้อในแบบสอบ
$p_i$	แทน	สัดส่วนของผู้ตอบถูกในแต่ละข้อ
$q_i$	แทน	สัดส่วนของผู้ตอบผิดในแต่ละข้อ
$S_t^2$	แทน	ความแปรปรวนของคะแนนทั้งหมด

การหาค่าความเที่ยงแบบคูเคอร์-ริชาร์ดสัน เป็นการหาความสอดคล้องกันระหว่างข้อสอบซึ่งดำเนินการสอบครั้งเดียว และใช้แบบสอบชุดเดียว ความสอดคล้องกันระหว่างข้อนี้ได้รับอิทธิพลจากแหล่งความแปรปรวนคลาดเคลื่อน 2 แหล่ง (อนาสตาซี 2519: 84) คือ

1. เนื้อหาที่สุ่ม

2. ความเป็นวิวิธพันธ์ของพฤติกรรมที่สุ่ม ยิ่งข้อสอบมีความเป็นเอกพจน์ ความสอดคล้องระหว่างข้อก็ยิ่งสูง (Stanley 1978: 126)

ค่าความเที่ยงแบบคูเคอร์ริชาร์ดสัน มีค่าประมาณได้กับค่าเฉลี่ยของสัมประสิทธิ์ความเที่ยงที่ได้จากการแบ่งครึ่งแบบสอบด้วยวิธีต่าง ๆ กัน (อนาสตาซี 2519: 86)

ข้อตกลงเบื้องต้นของสูตรคูเคอร์-ริชาร์ดสัน 20 คือ

1. การตรวจให้คะแนนแต่ละข้อ ให้ 1 คะแนนเมื่อตอบถูก และให้ 0 คะแนนเมื่อตอบผิด (Gronlund 1968: 112)

2. ข้อกระทงในแบบสอบ จะต้องมีลักษณะ เป็น เอกพจน์คือวัดคุณลักษณะเดียวกัน (Ebel 1965 : 318)

สูตรคูเคอร์-ริชาร์ดสัน 20 นี้ เหมาะที่จะหาความเที่ยงของแบบสอบความสามารถ (Power test) เท่านั้น ไม่เหมาะที่จะหาค่าความเที่ยงของแบบสอบความเร็ว (Speed Test) เพราะค่า  $p$  และ  $q$  ของแต่ละข้อ จะต้องเป็นค่าที่ได้จากการที่ผู้สอบทุกคนมีโอกาสทำข้อสอบข้อนั้นแล้วซึ่งเป็นไปไม่ได้สำหรับแบบสอบความเร็ว เพราะผู้สอบทุกคนไม่มีโอกาสทำถึงข้อสุดท้าย (Brown 1976: 326)

นอกจากสูตร KR-20 แล้ว คูเคอร์ และริชาร์ดสัน ยังได้เสนอสูตร คูเคอร์-ริชาร์ดสัน 21 (KR-21) ซึ่งสูตรนี้มีข้อตกลงว่า ข้อสอบแต่ละข้อมีความยากเท่ากัน หรือกำหนดให้ค่า  $p$  คงที่ ซึ่งเป็นสูตรที่ง่ายและสะดวกในการคำนวณ แต่ค่าความเที่ยงที่ได้จากสูตรนี้จะต่ำกว่าสูตร KR-20 เนื่องจากว่าโดยทั่ว ๆ ไปข้อสอบแต่ละข้อจะมีความยากแตกต่างกัน ซึ่งไม่เป็นไปตามข้อตกลง

#### องค์ประกอบที่มีอิทธิพลต่อความเที่ยง

นอกจากวิธีที่ใช้ในการประมาณค่าความเที่ยงจะมีอิทธิพลต่อค่าความเที่ยงของแบบสอบแล้วยังมีองค์ประกอบอื่น ๆ (Mehrens and Lehmann 1973: 100-103) ต่อไปนี้ที่มี

อิทธิพลต่อความเที่ยงของแบบสอบด้วย

1. ความยาวของแบบสอบ (Test Length) การเพิ่มจำนวนข้อสอบที่มีคุณภาพดีเท่าเทียมกับข้อสอบเดิม จะทำให้แบบสอบมีค่าความเที่ยงสูงขึ้น นั่นคือแบบสอบที่มีจำนวนข้อมาก จะมีความเที่ยงสูงกว่าแบบสอบที่มีจำนวนข้อน้อย สูตรที่แสดงให้เห็นความสัมพันธ์ระหว่างความยาวของแบบสอบกับความเที่ยงคือ สูตรของ สเปียร์แมน-บราวน์ (Spearman-Brown Formula)

2. ความเป็นเอกพันธ์ของกลุ่มผู้รับการทดสอบ (Group Homogeneity)

กลุ่มที่มีความเป็นเอกพันธ์ จะมีพิสัยของคะแนนความสามารถของกลุ่มน้อยกว่ากลุ่มที่มีความเป็นวิวิธพันธ์ (Group Heterogeneous) ซึ่งจะทำให้ค่าความเที่ยงของกลุ่มที่มีความเป็นเอกพันธ์ต่ำกว่ากลุ่มที่มีความเป็นวิวิธพันธ์

3. ความยากของข้อสอบ (Item Difficulty) เนื่องจากค่าความเที่ยงนั้น ประมาทจากความแปรปรวนของคะแนนผลการสอบ ถ้าข้อสอบยากเกินไป ทุกคนทำผิดหมด หรือข้อสอบง่ายเกินไปทุกคนทำถูกหมด ก็จะทำให้พิสัยของคะแนนลดลง ค่าความแปรปรวนของคะแนนมีน้อย ซึ่งจะทำให้ค่าความเที่ยงลดลงด้วย ดังนั้นถ้าค่าความยากของข้อสอบทำให้คะแนนกระจายมาก ๆ ก็จะทำให้ความเที่ยงมีค่าสูงขึ้นด้วย

4. ความเร็ว (Speed) การประมาณค่าความเที่ยงของแบบสอบความเร็ว (Speed Test) ด้วยวิธีวัดความคงที่ภายในเป็นวิธีไม่เหมาะสม เพราะแบบสอบชนิดนี้เป็นแบบสอบที่ง่ายแต่ใช้เวลาทำน้อย ซึ่งนักเรียนไม่สามารถทำได้ครบทุกข้อ และจะทำให้ค่าความเที่ยงสูงกว่าที่ควรจะเป็นจริง

5. ความเป็นปรนัย (Objective) ถ้าแบบสอบมีความเป็นปรนัยสูง ค่าความเที่ยงของแบบสอบก็สูงขึ้นด้วย

นอกจากนี้ การบริหารการสอบ เช่น การจับเวลา การให้คำชี้แจง การป้องกัน การทุจริต เป็นต้น หรือองค์ประกอบในด้านเกี่ยวกับตัวผู้สอบเอง เช่น แรงจูงใจ ความวิตกกังวล ประสบการณ์ในการสอบ การเดา เป็นต้น สิ่งเหล่านี้อาจเป็นแหล่งความคลาดเคลื่อนที่ทำให้ค่าความเที่ยงของแบบสอบเปลี่ยนแปลงได้

#### ความตรงของแบบสอบ

มีผู้ให้ความหมายของความตรง (Validity) ไว้ต่าง ๆ กันดังนี้

ความตรงของแบบสอบคือค่าสหสัมพันธ์ระหว่างแบบสอบกับเกณฑ์ (Gulliksen 1950: 88)

ความตรงของแบบสอบคือความสามารถวัดในสิ่งที่ต้องการวัดได้ถูกต้องแม่นยำหรือคือความสามารถในการให้ความหมายในสิ่งที่วัดได้อย่างไม่ผิดพลาด (Linguist 1942: 213 อ้างถึงใน ไสภมา เหลืองวิลาวัณย์ 2524: 19)

ความตรงของแบบสอบคือความถูกต้องแม่นยำของคะแนนในการพยากรณ์ที่จำเพาะเจาะจง (Stanley 1978: 101)

ความตรงของแบบสอบคือขอบเขตของผลการประเมินที่สอดคล้องกับจุดประสงค์ (Gronlund 1976: 79)

จากความหมายต่าง ๆ ข้างต้นสามารถสรุปความหมายของความตรงได้ 2 ประการ ดังนี้

1. หมายถึง ค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบกับคะแนน เกณฑ์
2. หมายถึง ความสามารถของแบบสอบที่วัดได้ในสิ่งที่ต้องการวัดหรือวัดได้สอดคล้องกับวัตถุประสงค์ และคะแนนที่ได้จากแบบสอบที่มีความตรงสูงสามารถบอกถึงสภาพที่แท้จริง และพยากรณ์ได้ถูกต้องแม่นยำ

#### ชนิดของความตรง

จากการให้ความหมายของความตรงที่แตกต่างกันทำให้จำแนกความตรงออกได้เป็นหลายชนิดและต่าง ๆ กันไป สมาคมจิตวิทยาอเมริกัน (American Psychological Association) ได้จำแนกความตรงออกเป็น 3 ชนิดใหญ่ ๆ (Mehrens and Lehmann 1973: 109-110) คือ

1. ความตรงตามเนื้อหา (Content Validity)
2. ความตรงตามเกณฑ์สัมพันธ์ (Criterion-related Validity)
3. ความตรงตามทฤษฎี (Construct Validity)

ในที่นี้จะกล่าวถึงความตรงตามเนื้อหาและความตรงตาม เกณฑ์สัมพันธ์ซึ่งเกี่ยวข้องกับ การวิจัยนี้ เท่านั้น

### ความตรงตามเนื้อหา (Content Validity)

ความตรงตามเนื้อหาจะ เกี่ยวข้องกับ เนื้อหาสาระข้อกระทงในแบบสอบ ซึ่งข้อกระทง นั้นต้องเป็นตัวแทนของเนื้อหาสาระ และครอบคลุมขอบ เขตของพฤติกรรมที่ต้องการวัด ความตรง ตาม เนื้อหา มีความสำคัญอย่างยิ่งในการสอบวัดผลสัมฤทธิ์ทางการ เรียน การที่จะตัดสินว่าแบบสอบ นั้นมีความตรงตาม เนื้อหามากน้อย เพียงใด จะต้องพิจารณาทั้งหัวข้อ เนื้อหาวิชาและชนิดของ พฤติกรรมที่ต้องการวัดไปพร้อม ๆ กัน

อย่างไรก็ตาม เราไม่สามารถแสดงค่าความตรงตาม เนื้อหาออกมา เป็นตัวเลขได้ โดยทั่วไปมักจะพิจารณาข้อกระทงในแบบสอบ โดยเทียบกับตารางวิเคราะห์หลักสูตร (Table of Specifications) แล้วประมาณออกมาว่ามีความตรงตาม เนื้อหามากน้อยเพียงใด ผู้ที่จะ ตัดสินความตรงตาม เนื้อหาได้ดีคือผู้เชี่ยวชาญทางสาขาวิชานั้น ๆ (Mehrens and Lehmann 1973: 111)

### ความตรงตาม เกณฑ์สัมพันธ์ (Criterion-Related Validity)

ความตรงตาม เกณฑ์สัมพันธ์ เป็น เรื่องที่เกี่ยวกับการศึกษาความสัมพันธ์ระหว่าง คะแนนที่ได้จากแบบสอบกับคะแนนการวัดจาก เกณฑ์ภายนอกที่เป็นอิสระ (Mehrens and Lehmann 1973: 112)

เกณฑ์ (Criterion) คือมาตรฐานในการตัดสิน ในการพัฒนาแบบสอบ เกณฑ์คือ คุณลักษณะหรือกลุ่มของคุณลักษณะที่ใช้เป็นพื้นฐานสำหรับตัดสินความตรงของแบบสอบหรือกระบวน การวัดอื่น ๆ (Ebel 1965: 448)

ความตรงตาม เกณฑ์สัมพันธ์จำแนกได้เป็น 2 ชนิด (บุญเชิด ธิญโญนันตพงษ์ 2521: 337) คือ

1. ความตรงร่วมสมัย (Concurrent Validity) หมายถึง ความตรงของ แบบสอบที่จะบ่งบอกสิ่งที่วัดได้ถูกต้องตามสภาพที่แท้จริงในปัจจุบัน โดยอาศัยความสัมพันธ์ระหว่าง คะแนนของแบบสอบ กับคะแนนเกณฑ์ซึ่งกำหนดขึ้นในขณะนั้น เช่น เกณฑ์สัมพันธ์ของแบบวัดทักษะ ในการคำนวณ ก็อาจใช้วิธีการสังเกตการทำแบบฝึกหัดเลยในชั่วโมงสอนเป็นเวลา 2 เดือน

ถ้าคะแนนที่นักเรียนได้จากแบบสอบให้ผลสอดคล้องกับ เกณฑ์สัมพันธ์ แสดงว่าแบบสอบนี้สามารถวัดสภาพปัจจุบันของนักเรียนได้

2. ความตรงเชิงทำนาย (Predictive Validity) หมายถึงความสามารถของแบบสอบที่จะบ่งบอกผลที่วัดในขณะนั้นได้ถูกต้องตามสภาพที่แท้จริงในอนาคต โดยอาศัยความสัมพันธ์ระหว่างคะแนนของแบบสอบกับคะแนน เกณฑ์สัมพันธ์ ซึ่งจะปรากฏในอนาคต เช่น เกณฑ์สัมพันธ์ของแบบสอบความถนัดทางวิชาการเพื่อทำนายผลการเรียน ก็อาจใช้คะแนนเฉลี่ยสะสมปีสุดท้าย เป็น เกณฑ์สัมพันธ์ ซึ่งการคำนวณหาความตรง เชิงพยากรณ์จะต้องอาศัย เวลา รอคอย เพราะคะแนนของแบบสอบกับ เกณฑ์สัมพันธ์ได้มาคนละ เวลา กัน ความตรงเชิงทำนายเหมาะสำหรับที่จะใช้กับแบบสอบความถนัดชนิดต่าง ๆ และแบบสอบอื่นที่ต้องการวัด เพื่อทำนายผล

ข้อแตกต่างระหว่างความตรงร่วมสมัยกับความตรงเชิงทำนาย ขึ้นอยู่กับวัตถุประสงค์ของการสอบ ความตรงร่วมสมัยใช้กับแบบสอบสำหรับวินิจฉัยสภาพที่เป็นอยู่ ส่วนความตรงเชิงทำนายใช้กับแบบสอบสำหรับพยากรณ์เหตุการณ์ในอนาคต (อนาสตาซี 2519: 105) นอกจากนี้ เม เรนส์และ เลแมนน์ (Mehrens and Lehmann 1973: 112) ได้กล่าวอีกว่า ข้อแตกต่างระหว่างความตรง 2 ชนิดนี้ คือช่วงระยะเวลาของการเก็บข้อมูลที่เป็นเกณฑ์ ถ้าเก็บรวบรวมข้อมูลที่เป็น เกณฑ์กับคะแนนการสอบใน เวลาเดียวกันจะเป็นความตรงร่วมสมัย แต่ถ้าเก็บคะแนนที่เป็น เกณฑ์ล่าออกไปหรือหลังการเก็บคะแนน จะเป็นความตรงเชิงทำนาย

#### การวัด เกณฑ์

การที่จะได้ข้อมูลของ เกณฑ์ที่เหมาะสมและถูกต้องนั้นทำได้ยาก การวัด เกณฑ์ก็ เหมือนกับการวัดสิ่งอื่น ๆ คือผู้วัดจะต้องทราบคุณลักษณะของสิ่งที่จะ เป็น เกณฑ์อย่างดี เสียก่อน การพิจารณา เกณฑ์นั้น เม เรนส์และ เลแมนน์ (Mehrens and and Lehmann) เสนอว่าจะต้องพิจารณาในสิ่งต่อไปนี้

1. ความเกี่ยวข้อง (Relevance) สิ่งที่จะใช้เป็น เกณฑ์จะต้องมีความเกี่ยวข้องกับ เรื่องที่ศึกษา นั่นคือ เกณฑ์นั้นจะต้องสะท้อนถึงคุณลักษณะที่สำคัญของ เรื่องที่จะศึกษา โดยทั่วไปการพิจารณาความ เกี่ยวข้องของ เกณฑ์นี้ขึ้นอยู่กับ การตัดสินใจคุณค่าไม่ได้วัดออกมาเป็นตัวเลข



2. ความเที่ยง (Reliability) เกณฑ์นั้นจะต้องมีความเที่ยงพอสมควร เนื่องจากความเที่ยงของแบบสอบมีอิทธิพลต่อค่าสหสัมพันธ์ระหว่างแบบสอบกับ เกณฑ์ ซึ่งทฤษฎีความสัมพันธ์ กล่าวไว้ว่า ความสัมพันธ์สูงสุดระหว่าง 2 ตัวแปร จะเท่ากับหรือน้อยกว่ารากที่สองของผลคูณระหว่างความเที่ยงของตัวแปรทั้งสองนั้น นั่นคือ

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

เมื่อ  $r_{xy}$  = ค่าสหสัมพันธ์ระหว่างตัวทำนาย (X) กับ เกณฑ์ (y)

$r_{xx}$  = ค่าความเที่ยงของแบบสอบที่เป็นตัวทำนาย

$r_{yy}$  = ค่าความเที่ยงของ เกณฑ์

จะเห็นได้ว่า ค่าความตรงตาม เกณฑ์สัมพันธ์ ( $r_{xy}$ ) มีค่ามากหรือน้อยขึ้นอยู่กับความเที่ยงของแบบสอบเอง ( $r_{xx}$ ) และความเที่ยงของ เกณฑ์ ( $r_{yy}$ )

3. ความเป็นอิสระจากอคติ (Free from Bias) เกณฑ์ที่จะใช้จะต้องเป็นอิสระ ต้องแน่ใจว่าจะคะแนนจากแบบสอบที่จะหาความตรงนั้น ไม่มีอิทธิพลต่อสถานภาพใด ๆ ของเกณฑ์สัมพันธ์ที่ใช้เทียบ นั่นคือต้องไม่ให้เป็นเกณฑ์มลทิน (Criteria Contamination) ตัวอย่างเช่น นำแบบสอบความถนัดทางวิชาการไปสอบกับนักเรียนตอนต้นปีพยากรณ์ผลการเรียนของนักเรียนตอนปลายปี ถ้าปรากฏว่าความรู้ที่ครูได้รับจากการสอบครั้งนี้ไปส่งเสริมให้ประเมินคะแนนการสอบปลายปีของนักเรียนบางคนสูงขึ้นหรือต่ำลง ลักษณะเช่นนี้แสดงว่าคะแนนของเกณฑ์ไม่มีความอิสระ หรือเป็น เกณฑ์มลทิน ซึ่งอิทธิพลเช่นนี้ทำให้ค่าความตรงตาม เกณฑ์สัมพันธ์ของแบบสอบสูงขึ้น โดยมีสาเหตุจากคุณภาพของแบบสอบจริง ๆ การป้องกันความคลาดเคลื่อนเช่นนี้อาจทำได้โดยหลีกเลี่ยงไม่ให้ผู้ประเมินทราบคะแนนการสอบที่ใช้พยากรณ์ (Mehrens and Lehmann 1973: 113-114)

#### วิธีประมาณค่าความตรงตาม เกณฑ์สัมพันธ์

การประมาณค่าความตรงตาม เกณฑ์สัมพันธ์ สามารถวิเคราะห์ได้จากสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบสอบกับ เกณฑ์ ดังนั้นสูตรที่ใช้จึง เป็นสูตรคำนวณสัมประสิทธิ์สหสัมพันธ์แบบต่าง ๆ โดยทั่วไปนิยมใช้สูตรสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson's Product Moment Correlation coefficient) (Guilford 1976: 140-141)



### องค์ประกอบที่มีอิทธิพลต่อความตรง

องค์ประกอบที่ทำให้ค่าความตรงของแบบสอบลดลง (Gronlund 1976: 98-102)

#### 1. องค์ประกอบในแบบสอบ ได้แก่

1.1 คำสั่งไม่ชัดเจน คำสั่งไม่ชัดเจนจะทำให้นักเรียนปฏิบัติตามคำสั่งผิดพลาดได้

1.2 การใช้คำและประโยคที่ยากเกินไป การใช้คำและประโยคที่ยากเกินไป

จะทำให้นักเรียนเกิดความสับสนในการสอบ ผลที่ได้กลายเป็นนักเรียนที่ทำข้อสอบได้ถูกต้อง คือผู้ที่มีความสามารถในการอ่านสูง ดังนั้นแบบสอบนี้จึงวัดความเข้าใจในการอ่านของนักเรียน แทนที่จะวัดพฤติกรรมที่ต้องการ

1.3 ระดับความยากของข้อกระทงไม่เหมาะสม ข้อกระทงที่ยากหรือง่ายเกินไปไม่สามารถจำแนกนักเรียนได้

1.4 การสร้างข้อกระทงไม่รัดกุม ข้อสอบที่สร้างอย่างไม่ระมัดระวังอาจชี้แนะคำตอบแก่นักเรียน

1.5 ภาษาในข้อกระทงกำกวม ภาษาในข้อกระทงที่กำกวมจะทำให้นักเรียนตีความผิดและสับสน

1.6 แบบสอบที่ใช้ไม่เหมาะสมจะใช้วัดพฤติกรรมที่ต้องการวัด

1.7 แบบสอบสั้นเกินไป แบบสอบที่สั้นเกินไปจะถามได้ไม่ครอบคลุมวัตถุประสงค์ของการวัด

1.8 การจัดเรียงข้อกระทงไม่เหมาะสม ข้อกระทงควรเรียงจากง่ายไปหายาก ถ้าเรียงจากยากก่อนนักเรียนอาจใช้เวลาคิดนานจนกระทั่งไม่มีเวลาทำข้อง่าย ๆ ซึ่งอยู่ตอนท้ายของแบบสอบ จึงวัดได้ไม่ตรงตามความสามารถที่แท้จริง

1.9 คำตอบที่ถูกเรียงกันอย่างเป็นระบบ การมีระบบของคำตอบที่ถูกต้อง ทำให้นักเรียนสามารถเดาคำตอบข้อหลัง ๆ ได้

นอกจากนั้น ค่าความเที่ยงของแบบสอบก็มีส่วนสำคัญที่จะกำหนดค่าความตรงของแบบสอบ (Brown 1976: 99) ดังสมการ



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

เมื่อ  $r_{xy}$  = ค่าความตรงของแบบสอบ

$r_{xx}$  = ค่าความเที่ยงของแบบสอบ

จะเห็นได้ว่าค่าความตรงของแบบสอบมีค่าไม่เกินรากกำลังที่สองของค่าความเที่ยงของแบบสอบนั้น นั่นคือ ค่าความตรงของแบบสอบจะถูกจำกัดโดยค่าความเที่ยงของมันเอง

2. การบริหารแบบสอบและการให้คะแนน ทำให้ความตรงของแบบสอบลดลงได้ เช่น ให้ความเวลาในการสอบน้อยเกินไป การอธิบายข้อสอบให้แก่ นักเรียนบางคนที่ยังสงสัยหรือขอร้องให้ช่วย การทุจริตของนักเรียน การให้คะแนนที่ขาดความเชื่อถือได้ เมื่อใช้ข้อสอบอัตโนมัติ สิ่งเหล่านี้ทำให้ความตรงของแบบสอบที่ครูสร้างลดลง ในกรณีของแบบสอบมาตรฐาน ถ้าผู้ดำเนินการสอบไม่ปฏิบัติตามคำสั่ง เช่น ไม่จับเวลาตามที่กำหนด ตรวจให้คะแนนผิด หรือให้ความช่วยเหลือเด็กบางคน การปฏิบัติเหล่านี้ทำให้ความตรงของแบบสอบลดลง เช่นเดียวกัน

3. องค์ประกอบในการตอบของนักเรียน นักเรียนบางคนอาจมีสภาพจิตใจไม่ปกติ เช่น มีความวิตกกังวลในการสอบ จึงไม่สามารถทำแบบสอบได้เต็มตามความสามารถหรือไม่มีแรงบันดาลใจให้ทำแบบสอบ สิ่งเหล่านี้ก็มีส่วนทำให้ความตรงของแบบสอบต่ำ นอกจากนี้ อาจเกี่ยวกับลักษณะการตอบของนักเรียน เช่น นักเรียนบางคนเลือกตอบถูกเสมอในกรณีไม่แน่ใจ เมื่อทำแบบสอบแบบถูกผิด ในขณะที่บางคนเลือกตอบผิดเสมอเมื่อไม่แน่ใจ และถ้าการเลือก โดยการเดาสุ่มนี้บังเอิญไปตรงกับคำตอบที่ถูกต้องของข้อกระทงเป็นส่วนใหญ่ คะแนนที่ได้จะไม่ตรงกับความสามารถที่แท้จริงของนักเรียน ทำให้ความตรงของแบบสอบต่ำ

4. ลักษณะของกลุ่มและ เกณฑ์ ความตรงมักจะ เป็นของเฉพาะกลุ่ม เนื่องจากสิ่งที่เป็นแบบสอบวัดได้นั้นมักได้รับอิทธิพลของตัวแปรต่าง ๆ เช่น อายุ เพศ ระดับความสามารถ ภูมิหลังของการศึกษา และวัฒนธรรมของผู้ตอบ ดังนั้นในการพิจารณาหาความตรงของแบบสอบที่เป็นมาตรฐาน ควรคำนึงถึงลักษณะเฉพาะของกลุ่มที่ใช้หาความตรง นอกจากนั้นต้องพิจารณาเกณฑ์ที่ใช้หาความตรงของแบบสอบด้วย

ความยากของข้อสอบแต่ละข้อขึ้นอยู่กับสัดส่วนหรือ เปอร์เซนต์ของผู้สอบที่ตอบข้อสอบนั้นถูกต้อง (อนาสตาซี 2519 : 161) ถ้าเปอร์เซนต์ของผู้ตอบถูกมีมาก ก็แสดงว่าข้อสอบนั้นง่าย ค่าความยากของข้อสอบไม่ได้เป็นเครื่องชี้ว่าข้อสอบนั้นดีหรือ เลว มัน เป็นเพียงแต่ระบุว่าข้อสอบข้อนั้นยากหรือง่ายเท่านั้น อย่างไรก็ตามดัชนีความยากของข้อสอบก็ยังสะท้อนให้เห็นความสามารถของกลุ่มผู้เข้ารับการทดสอบด้วย (Ebel 1965 : 360) เหตุผลสำคัญในการหาความยากของข้อสอบก็เพื่อว่าจะคัดเลือกข้อสอบที่มีความยากไม่เหมาะสมออก ข้อสอบที่ไม่มีใครทำถูกเลยกับข้อที่ทุก ๆ คนทำได้หมดนั้น หากเป็นแบบสอบที่ต้องการจำแนกคนที่มีความสามารถแตกต่างกันออกจากกันแล้ว ควรจะตัดออก เพราะข้อสอบทั้งสองประเภทนี้ไม่สามารถแยกความแตกต่างของผู้สอบได้ และข้อสอบแบบนี้ไม่ได้มีผลต่อความแปรปรวนของคะแนนสอบ จึงไม่มีผลต่อค่าความเที่ยงและความตรงของแบบสอบนี้ (อนาสตาซี 2519 : 161-162) ในการวัดแบบอิงเกณฑ์ (Criterion-referenced) โดยทั่วไปข้อสอบจะมีค่าความยากสูง นักเรียนส่วนมากจะได้คะแนนสูงซึ่งจะช่วยให้ทราบว่านักเรียนคนใดบ้างที่มีความรู้และทักษะไม่ถึงเกณฑ์ที่เราคาดหวัง (อนันต์ ศรีโสภณ 2519: 257) ส่วนในการวัดแบบอิงกลุ่ม (Norm-Referenced) นั้น เพื่อให้ความจำแนกระหว่างบุคคลมีค่าสูงสุดจึง เป็นการดีที่สุดที่จะ เลือกข้อสอบที่มีความยากประมาณ .50 ซึ่งจะทำให้ได้คะแนนรวมที่มีการกระจายมากที่สุด ซึ่งจะมีผลทำให้ค่าสัมประสิทธิ์ความเที่ยงมีค่าสูงสุดด้วย (อนาสตาซี 2519: 161-162) นั้นแสดงว่าค่าความยากย่อมส่งผลถึงความเที่ยงและความตรงของแบบสอบด้วย

อย่างไรก็ตามค่าความยากนี้มีความสัมพันธ์กับค่าอำนาจจำแนก ค่าความยากที่ทำให้ค่าอำนาจจำแนกมากที่สุดคือ ค่าความยากเท่ากับ .05 แต่ข้อพึงระลึกอยู่เสมออีกคือ การที่ข้อสอบมีค่าความยากเท่ากับ .05 อาจจะทำให้ข้อสอบนั้นมีค่าอำนาจจำแนกสูงสุดหรือไม่มีเลยก็ได้ เพราะถ้าครึ่งหนึ่งของจำนวนคนในกลุ่มสูงตอบถูก และครึ่งหนึ่งของจำนวนคนในกลุ่มต่ำตอบถูก ก็ทำให้ค่าอำนาจจำแนกเท่ากับ 0 (อนันต์ ศรีโสภณ 2519: 257) นั้นแสดงว่าค่าความยากเท่ากับ .50 นั้นอาจทำให้ข้อสอบมีอำนาจจำแนกสูงสุดหรือไม่มีอำนาจจำแนกเลยก็ได้

#### องค์ประกอบที่มีอิทธิพลต่อความยากของแบบสอบ

องค์ประกอบที่มีอิทธิพลต่อความยากของแบบสอบแบ่งออกเป็น 2 ประเภท (Campbell 1961 : 899-913, อ้างถึงในนิรมล บุญตะรัตน์ 2525 : 18-19) คือ

1. องค์ประกอบภายใน (Intrinsic Factors) ประกอบด้วย

- ก. เนื้อหาของข้อสอบแต่ละข้อ (Item Content) ในด้านความซับซ้อน ความเป็นนามธรรม และความแปลกใหม่
- ข. ลักษณะโครงสร้างของข้อสอบ (Item Structure) หมายถึงวิธีการ ในการแสดงออกซึ่งเนื้อหา

2. องค์ประกอบภายนอก (Extrinsic Factors) ประกอบด้วย

- ก. ความไม่คุ้นเคยต่อเนื้อหา (Unfamiliarity) คืออยู่นอกเหนือ ประสบการณ์ของผู้ตอบ
- ข. สิ่งที่สัมพันธ์กับข้อสอบ (Item Content) เช่น ข้อสอบข้อที่อยู่ใกล้เคียงกัน
- ค. ตัวแปรด้านบุคลิกภาพ ได้แก่ สภาพร่างกาย ลักษณะนิสัย และความตั้งใจ ของผู้สอบ

นอกจากนี้องค์ประกอบที่อาจจะมีผลต่อค่าความยากของแบบสอบก็คือ ธรรมชาติของ เนื้อหา ชนิดของพฤติกรรมที่ต้องการวัดจากผู้สอบ และตัวการอื่น ๆ ที่สลับซับซ้อนมองเห็นไม่เด่นชัด เช่น ลักษณะของภาษาที่ใช้ รูปแบบของคำถาม ตลอดจนคำชี้แจงและอื่น ๆ ซึ่งสิ่งเหล่านี้ผู้สร้าง แบบสอบควรจะสนใจนำเข้ามาพิจารณาในการสร้างแบบสอบด้วย

อำนาจจำแนกของแบบสอบ

อำนาจจำแนก หมายถึง ประสิทธิภาพของข้อคำถามในการแบ่งเด็กออกเป็นกลุ่ม คนเก่งและอ่อน อำนาจจำแนกมีค่าตั้งแต่  $-1.00$  ถึง  $1.00$  ข้อสอบที่มีค่าอำนาจจำแนก เข้าใกล้  $-1.00$  หรือ  $1.00$  แสดงว่ามีอำนาจจำแนกสูง แต่ข้อสอบที่มีคุณค่าตรงตามวัตถุประสงค์ของ การสอบคือ ข้อสอบที่มีค่าอำนาจจำแนก เป็นบวก ซึ่งหมายความว่านักเรียนในกลุ่มสูงตอบข้อสอบ ได้ถูกต้องมากกว่านักเรียนในกลุ่มต่ำ และแสดงว่าข้อสอบสามารถจำแนกไปในทางเดียวกับคะแนน รวมที่ได้จากแบบสอบ ส่วนข้อสอบที่มีค่าอำนาจจำแนก เป็นทางลบถือว่าจำแนกผิดทิศทาง ไม่ตรง ตามวัตถุประสงค์ที่ต้องการ และข้อสอบที่มีอำนาจจำแนก เป็น  $0$  ถือว่าจำแนกความแตกต่าง ระหว่างคน 2 กลุ่มไม่ได้เลย

ในการประเมินคุณภาพของข้อสอบโดยอาศัยค่าอำนาจจำแนกควรคือ เกณฑ์ดังนี้ (Ebel 1965: 364)

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

อำนาจจำแนกของข้อสอบ เป็นสิ่งสำคัญอีกอย่างหนึ่งที่ควรคำนึงถึงในการประเมินคุณภาพของข้อสอบควบคู่ไปกับความยากของข้อสอบ เนื่องจากข้อสอบที่มีความยากเท่ากัน ยังแตกต่างกันในด้านอำนาจจำแนกได้ด้วย ดังที่ได้กล่าวแล้วว่าข้อสอบที่มีความยาก .50 อาจจะให้ค่าอำนาจจำแนกสูงสุดหรืออาจจะไม่มีเลยก็ได้ ฉะนั้นข้อสอบที่มีความยาก .50 มีโอกาสทำให้ความเที่ยงของแบบสอบต่ำได้ เมื่อข้อสอบนั้นมีอำนาจจำแนกเป็น 0 แต่ข้อสอบที่มีค่าอำนาจจำแนกสูงจะมีผลทำให้ความเที่ยงของแบบสอบสูงด้วยอย่างแน่นอน เนื่องจากค่าอำนาจจำแนกของข้อสอบ เกี่ยวข้องกับส่วน เบี่ยงเบนมาตรฐานของคะแนนซึ่งเป็นองค์ประกอบสำคัญที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบ (Ebel 1965: 302)

#### องค์ประกอบที่มีอิทธิพลต่ออำนาจจำแนกของแบบสอบ

องค์ประกอบที่สำคัญที่มีอิทธิพลต่ออำนาจจำแนกของแบบสอบได้แก่

1. คุณภาพของตัวเลือก และจำนวนตัวเลือก ข้อสอบที่สร้างตัวเลือกได้ดีมีความใกล้เคียงกับตัวเลือกที่เป็นคำตอบถูกหลาย ๆ ตัวจะทำให้จำนวนตัวเลือกเพิ่มขึ้น ซึ่งจะให้อำนาจจำแนกของข้อสอบสูงขึ้นด้วย (Ebel 1965: 165) เนื่องจากผู้ที่มีความรู้จริงเท่านั้นสามารถทำข้อสอบได้ถูกต้อง แต่ผู้ที่ไม่มีความรู้จริงจะเลือกตอบตัวเลือกที่มีความใกล้เคียงกับคำตอบถูกกระจายกันออกไปทุกตัวเลือก

2. ความยากของข้อสอบ การจำแนกความแตกต่างของผลสัมฤทธิ์ระหว่างกลุ่มขึ้นอยู่กับความยากของข้อสอบ (Gronlund 1976: 154) คือข้อสอบที่มีความยาก .50 มีโอกาสให้ค่าอำนาจจำแนกสูงสุดได้ เมื่อนักเรียนในกลุ่มสูงทุกคนตอบข้อสอบได้ถูกต้อง ในขณะที่นักเรียนในกลุ่มต่ำทั้งหมดตอบข้อสอบผิด และหากข้อสอบมีค่าความยากสูงหรือต่ำกว่านี้ ค่าอำนาจจำแนกของข้อสอบจะต่ำลง ซึ่งหมายความว่าข้อสอบที่ง่ายหรือยากเกินไปจะเป็นข้อสอบที่มีค่าอำนาจจำแนกต่ำนั่นเอง

## ตอนที่ 2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการวิจัย เรื่องนี้จะ เสนอการวิจัย เกี่ยวกับการเปรียบเทียบ  
คุณภาพของแบบสอบแบบ เลือกตอบและแบบสอบแบบถูกผิด ทั้งต่างประเทศและในประเทศ โดย  
แบ่งเป็น 2 ตอน คือ ผลการวิจัยที่พบว่าคุณภาพของแบบสอบทั้งสองไม่แตกต่างกัน กับพวกที่พบว่า  
คุณภาพของแบบสอบทั้งสองแตกต่างกัน

ผลการวิจัยที่พบว่าคุณภาพของแบบสอบทั้งสองไม่แตกต่างกัน

รินส์แลนด์ (Rinsland 1938: 295-299 อ้างถึงใน กรองทอง เทพศิริอำนวยการ  
2525: 18) ซึ่งได้รวบรวมผลการทดลองที่เกี่ยวกับแบบสอบต่าง ๆ ได้สรุปไว้ว่าแบบสอบปรนัย  
ทั้งหมดยกเว้นแบบสอบแบบถูกผิด มีค่าความตรงพอ ๆ กันกับแบบสอบแบบอัตนัยหรือน้อยกว่า  
เล็กน้อย ส่วนในด้านความเที่ยงปรากฏว่า แบบสอบปรนัยทุกชนิดให้ค่าความเที่ยงพอ ๆ กันเมื่อ  
ใช้เวลาในการสอบเท่ากัน แต่การที่จะมีค่าความเที่ยงต่างกันนั้นอาจ เป็นเพราะการใช้คำหรือ  
ภาษาที่แตกต่างกันมากกว่าจะเป็น เพราะรูปแบบของคำถามหรือข้อสอบ

อีเบล (Ebel, quoted in Frisbie 1973: 297) ได้ศึกษาเปรียบเทียบค่าความ  
เที่ยงและความตรงของแบบสอบแบบถูกผิดและแบบ เลือกตอบ โดยใช้อัตราส่วนจำนวนข้อของ  
แบบสอบแบบถูกผิดต่อแบบสอบแบบ เลือกตอบที่นักเรียนทำได้ในเวลาเท่ากัน มีค่าเท่ากับ 2:1  
เป็นตัวปรับค่าความเที่ยงของแบบสอบแบบถูกผิดที่ใช้เวลาในการสอบ เท่ากับแบบสอบแบบ  
เลือกตอบ ผลการศึกษาพบว่า ค่าความเที่ยงของแบบสอบแบบ เลือกตอบและแบบสอบแบบถูกผิด  
ไม่แตกต่างกัน และแบบสอบแบบ เลือกตอบและแบบสอบแบบถูกผิดสามารถวัดได้ในสิ่งเดียวกัน

เออร์วิน ฮิล เบิร์น และแลนด์แมน (Irvin, Hahpern and Landman  
1980: 51-58) ได้ประเมินผลสัมฤทธิ์เด็กที่ เรียนซ้ำด้วยแบบสอบมาตรฐานแบบถูกผิดและแบบ  
เลือกตอบด้วยแบบสอบ 3 ฉบับ ที่เกี่ยวกับสังคมและอาชีพ คือแบบสอบแบบถูกผิด แบบสอบแบบ  
เลือกตอบชนิด 3 ตัวเลือก และแบบสอบเลือกตอบชนิด 2 ตัวเลือกโดยสร้างจากแบบสอบ 3  
ตัวเลือกโดยตัดตัวลงที่มีอำนาจจำแนกต่ำสุดของข้อสอบออกข้อละ 1 ตัว กลุ่มตัวอย่างที่ใช้ใน  
การวิจัยเป็นนักเรียนมิตปกติเกรด 9-12 จำนวน 425 คน ซึ่งแบ่งเป็น 3 กลุ่ม แต่ละกลุ่ม  
ให้ทำแบบสอบฉบับใดฉบับหนึ่งใน 3 ฉบับ ผลการวิจัยพบว่าค่าความเที่ยงและความตรงของ



แบบสอบแบบ เลือกตอบทั้งสองฉบับและแบบสอบแบบถูกผิดมีค่าสูงและใกล้เคียงกัน

กรีน (Green 1978: 42-44) ศึกษาเปรียบเทียบค่าความเที่ยงและความตรงของแบบสอบแบบ เลือกตอบกับแบบสอบแบบถูกผิดโดยใช้แบบสอบสัมฤทธิ์ผลทางการเรียนวิชาทดสอบและการวัดผลที่ครูสร้างขึ้น 2 ฉบับ ฉบับที่ 1 เป็นแบบสอบแบบ เลือกตอบ 3 ตัวเลือก จำนวน 40 ข้อ ข้อกระทงของแบบสอบแบบ เลือกตอบแต่ละข้อมีลักษณะที่ตัวคำถาม (Stem) อยู่รวมกับตัวเลือก (Option) ฉบับที่ 2 เป็นแบบสอบแบบถูกผิดซึ่งแปลงมาจากแบบสอบแบบ เลือกตอบ โดยแปลงข้อกระทงแบบสอบแบบ เลือกตอบ 1 ข้อ ไปเป็นแบบสอบแบบถูกผิด 3 ข้อ ซึ่งเป็นข้อถูก 1 ข้อ ข้อผิด 2 ข้อ ซึ่งแปลงได้ทั้งหมด 120 ข้อ กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้เป็นนิสิตปริญญาตรีจำนวน 50 คน แบ่งเป็น 2 กลุ่ม ๆ ละ 25 คน ให้แต่ละกลุ่มทำแบบสอบฉบับใดฉบับหนึ่ง เมื่อหมดเวลา 12 นาทีแรกของการสอบให้นิสิตบันทึกจำนวนข้อที่ทำไว้ได้ ผลการศึกษาพบว่าอัตราส่วนของจำนวนข้อกระทงแบบถูกผิดต่อแบบสอบแบบ เลือกตอบที่นิสิตทำได้ในเวลาเท่ากันมีค่าเท่ากับ 1.4 : 1 ค่าความเที่ยงของแบบสอบแบบถูกผิดกับแบบสอบแบบ เลือกตอบ หลังจากปรับโดยการเพิ่มความยาวของแบบสอบแบบ เลือกตอบ เป็น 1.25 เท่าของจำนวนข้อเดิม เพื่อให้ใช้เวลาเท่ากับแบบสอบแบบถูกผิดมีค่าไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ค่าความตรงไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ผลการวิจัยที่พบว่าคุณภาพของแบบสอบแบบ เลือกตอบและแบบสอบแบบถูกผิด แตกต่างกัน มีดังนี้

ฟรีสบี (Frisbie 1973: 297-304) ได้ศึกษาค่าความเที่ยงและความตรงของแบบสอบแบบ เลือกตอบและแบบสอบแบบถูกผิด วิชาสังคมศึกษา (Social Science) และวิชาวิทยาศาสตร์ธรรมชาติ (Natural Science) โดยสร้างเป็นแบบสอบแบบ เลือกตอบก่อนแล้วแปลงเป็นแบบสอบแบบถูกผิด ทั้งหมดมี 8 ฟอรัม กลุ่มตัวอย่างใช้นักเรียนจำนวน 1,018 คน ให้นักเรียนทำแบบสอบคนละ 1 ฟอรัม โดยแบ่งเวลาเป็น 2 ช่วง ๆ แรกทำ 8 นาที แล้วให้นักเรียนบันทึกจำนวนข้อที่ทำได้ หลังจากนั้นให้ทำต่ออีก 50 นาที ผลการศึกษาพบว่าอัตราส่วนของข้อกระทงของแบบสอบแบบถูกผิดต่อแบบสอบแบบ เลือกตอบที่นักเรียนทำได้ในเวลาเท่ากันของวิชาวิทยาศาสตร์ธรรมชาติ เป็น 1.44 : 1 และวิชาสังคมศึกษาเป็น 1.53 : 1 ส่วนค่าความเที่ยงของแบบสอบแบบ เลือกตอบและแบบสอบแบบถูกผิดหลังจากปรับให้เวลาในการสอบ



เท่ากันแตกต่างกันอย่างมีนัยสำคัญทางสถิติ และแบบสอบทั้งสองมีแนวโน้มที่จะวัดในสิ่งเดียวกัน

ต่อมาในปี ค.ศ. 1974 ฟรีสบี (Frisbie 1974: 885-892) ได้ศึกษาอิทธิพลของข้อกระทงแบบ เลือกตอบและแบบถูกผิดที่มีต่อค่าความเที่ยงอีกครั้งหนึ่ง โดยมีวัตถุประสงค์เพื่อยืนยันกับผลงานวิจัยครั้งแรก กลุ่มตัวอย่างเป็นนักเรียน จำนวน 529 คน แบบสอบที่ใช้ศึกษาคือแบบสอบสัมฤทธิ์ผลที่ใช้ในการศึกษาคั้งก่อน แต่จัดให้เหลือ 4 ฟอรัม โดยแต่ละฟอรัม ข้อกระทงของแบบสอบแบบถูกผิดมีความยาวเป็น 1.5 เท่าของแบบสอบแบบเลือกตอบ วิธีเก็บรวบรวมข้อมูล เหมือนครั้งแรกทุกประการ ปรากฏว่าผลการศึกษาคั้งนี้สอดคล้องกับผลการศึกษาคั้งแรกทุกประการ

ในปีเดียวกันอูสเตอร์ฮอฟและกลาสแนท (Oosterhof and Glasnapp 1974: 62-64) ได้ศึกษาเปรียบเทียบความเที่ยงและความยากของแบบสอบแบบ เลือกตอบและแบบสอบแบบถูกผิด โดยใช้แบบสอบสัมฤทธิ์ผลที่ครูสร้างขึ้นมา 2 ฉบับ ฉบับที่ 1 เป็นแบบสอบแบบเลือกตอบ 4 ตัวเลือก จำนวน 40 ข้อ ฉบับที่ 2 เป็นแบบสอบแบบถูกผิดที่แปลงมาจากฉบับที่ 1 ด้วยวิธีรวมข้อคำถามกับตัวเลือกที่เป็นตัวถูก เป็นข้อกระทง 1 ข้อ และรวมข้อคำถามกับตัวลวงที่มีอำนาจจำแนกสูงสุด เป็นข้อกระทง 1 ข้อ นำข้อกระทงทั้ง 80 ข้อมาแยกเป็น 2 ส่วน โดยแยกข้อกระทงแบบถูกผิดทีละคู่ไปอยู่คนละส่วนโดยวิธีสุ่มแล้วจัดลำดับของข้อกระทงใน แต่ละส่วนด้วยวิธีสุ่มเช่นเดียวกัน กลุ่มตัวอย่างของการวิจัย คือ นิสิตปริญญาตรี จำนวน 101 คน แบ่งเป็น 2 กลุ่ม ๆ แรกมี 51 คน ให้ทำแบบสอบฉบับที่ 1 ก่อน แล้วค่อยทำฉบับที่ 2 ส่วนกลุ่มที่ 2 มี 50 คน ให้ทำแบบสอบฉบับที่ 2 ก่อนแล้วค่อยทำฉบับที่ 1 ในการทำแบบสอบให้บันทึกเวลาที่ใช้ทำเมื่อเสร็จ 40 ข้อ 80 ข้อ และ 120 ข้อ ผลการศึกษาได้ข้อสรุปว่า อัตราส่วนจำนวนข้อกระทงแบบสอบแบบถูกผิดต่อแบบสอบแบบ เลือกตอบที่กลุ่มตัวอย่างทำได้ในเวลาเท่ากัน เป็น 1.73 : 1 ค่าความเที่ยงของแบบสอบแบบถูกผิดหลังจากปรับแล้วมีค่าต่ำกว่าความเที่ยงของแบบสอบแบบ เลือกตอบ และถ้าจะให้แบบสอบแบบถูกผิดมีค่าความเที่ยงเท่ากับแบบสอบแบบ เลือกตอบแล้วแบบสอบแบบถูกผิดจะต้องมีความยาวเป็น 3.75 เท่าของแบบสอบแบบ เลือกตอบ สำหรับค่าความยากของแบบสอบแบบ เลือกตอบหลังจากปรับารเดาแล้ว ปรากฏว่าแบบสอบแบบ เลือกตอบง่ายที่สุด ส่วนแบบสอบแบบถูกผิดที่มีข้อถูกทั้งหมดจะง่ายกว่าแบบสอบแบบถูกผิดที่มีข้อผิดทั้งหมด

สุธรรม จันทร์หอม (2513: 92-95) ศึกษาเปรียบเทียบการใช้แบบสอบแบบถูกผิด แบบ เลือกตอบและแบบเติมคำ วัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 7 กลุ่มตัวอย่างใช้นักเรียนจำนวน 1,350 คน เครื่องมือที่ใช้คือแบบสอบวิชาคณิตศาสตร์ ซึ่งแยก เป็น 3 ฉบับคือ คณิตศาสตร์ทักษะ คณิตศาสตร์ปัญหา และคณิตศาสตร์ความเข้าใจ เก็บรวบรวม ข้อมูลโดยให้นักเรียนทำแบบสอบคนละฉบับโดยแต่ละคนได้รับรูปแบบของข้อสอบไม่ซ้ำกัน การ ศึกษาครั้งนี้ยึดจำนวนข้อของแบบสอบทั้งสามเท่ากันและวัด เนื้อหาเดียวกัน ผลการศึกษาพบว่า ค่าความเที่ยงของแบบสอบทั้งสามแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ค่าความตรงของ แบบสอบวิชาคณิตศาสตร์ความเข้าใจแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ส่วนค่าความยากของ แบบสอบทั้งสามแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 โดยแบบสอบแบบถูกผิดเป็นแบบสอบที่ ง่ายที่สุด แบบสอบแบบ เลือกตอบมีความยากปานกลาง และแบบสอบแบบเติมคำยากที่สุด

ศุภชัย เอื้ออิสระวิมล (2519: 46-48) ศึกษาผลการใช้แบบสอบแบบ เลือกตอบ แบบเติมคำ และแบบถูกผิด วัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 7 กลุ่มตัวอย่างจำนวน 900 คน เครื่องมือที่ใช้ศึกษา คือแบบสอบวิชาวิทยาศาสตร์ที่ผู้วิจัยสร้าง ขึ้นเอง จำนวน 6 ฉบับ โดยแบบสอบแต่ละประเภทมี 2 ฉบับ คือ ฉบับ ก ฉบับ ข ผลการศึกษาพบว่า ค่าความยากของแบบสอบทั้ง 3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดย แบบสอบเติมคำยากที่สุด แบบสอบแบบ เลือกตอบยากปานกลาง และแบบสอบแบบถูกผิดง่ายที่สุด ค่าอำนาจจำแนกของแบบสอบทั้งสามแตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยแบบสอบเติมคำ ให้ค่าอำนาจจำแนกสูงสุด แบบ เลือกตอบให้ค่าอำนาจจำแนกปานกลาง และแบบสอบถูกผิด ให้ค่าอำนาจจำแนกต่ำสุด ค่าความเที่ยงของแบบสอบทั้งสามแตกต่างกันอย่างมีนัยสำคัญ โดย แบบสอบแบบเติมคำให้ค่าความเที่ยงสูงสุด รองลงมาคือแบบ เลือกตอบและแบบถูกผิดให้ค่า ความเที่ยงต่ำสุด ส่วนค่าความตรงของแบบสอบวิทยาศาสตร์ ฉบับ ข ไม่พบความแตกต่าง แต่ฉบับ ก แตกต่างอย่างมีนัยสำคัญทางสถิติ โดยแบบสอบแบบเติมคำให้ค่าความตรงสูงสุด และ แบบสอบถูกผิดให้ค่าความตรงต่ำสุด

อุจน์ มณีคำ (2522: 60-63) ได้ศึกษาเปรียบเทียบความเที่ยงและความตรง ของแบบสอบสัมฤทธิ์ผล แบบ เลือกตอบ และแบบถูกผิดวิชาวิทยาศาสตร์ที่ผู้วิจัยสร้างขึ้น 3 ฉบับ คือ แบบสอบแบบ เลือกตอบ แบบสอบถูกผิดที่แปลงมาจากแบบสอบ เลือกตอบโดยพิจารณา ค่าอำนาจจำแนกของตัวลอง และอีกฉบับหนึ่งแปลงโดยวิธีสุ่มให้เป็นข้อผิดและข้อถูก กลุ่มตัวอย่าง



ใช้นักเรียนจำนวน 570 คน เก็บรวบรวมข้อมูลโดยให้นักเรียนแต่ละคนตอบแบบสอบถามเพียงฉบับเดียวเท่านั้น ผลการวิจัยพบว่า อัตราส่วนของจำนวนข้อกระทงแบบถูกผิดที่แปลงมาจากแบบสอบถามแบบเลือกตอบโดยพิจารณาอำนาจจำแนกของตัวलगต่อแบบเลือกตอบ และอัตราส่วนของแบบสอบถามแบบถูกผิดที่แปลงโดยการสุ่มต่อแบบสอบถามแบบเลือกตอบมีค่าเป็น 1.41 : 1 และ 1.5 : 1 ตามลำดับ ค่าความเที่ยงของแบบสอบถามเลือกตอบกับแบบถูกผิด หลังจากปรับให้ใช้เวลาในการสอบ เท่ากันแตกต่างกันอย่างมีนัยสำคัญทางสถิติ สำหรับค่าความตรงของแบบสอบถามแบบเลือกตอบกับแบบถูกผิดที่แปลงโดยการพิจารณาค่าอำนาจจำแนกของตัวलगแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แต่ความตรงของแบบสอบถามแบบเลือกตอบและแบบสอบถามถูกผิดที่แปลงโดยการสุ่ม แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

นิหาพร จุลมนตรี (2524: 54-56) ศึกษาความยากและความเที่ยงของแบบสอบถามแบบถูกผิดและแบบสอบถามแบบเลือกตอบที่มีความตรงเทียบเคียงกัน กลุ่มตัวอย่างใช้นักเรียนจำนวน 522 คน เครื่องมือที่ใช้คือแบบสอบถามวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 4 ซึ่งผู้วิจัยสร้างขึ้นเอง จำนวน 3 ฉบับ เป็นแบบสอบถามเลือกตอบ 1 ฉบับ จำนวน 40 ข้อ แบบสอบถามถูกผิด 2 ฉบับ โดยการแปลงมาจากแบบสอบถามแบบเลือกตอบข้อต่อข้อ ซึ่งมีข้อผิดประมาณ 50 % ของทั้งหมด 1 ฉบับ และมีข้อผิดประมาณ 67 % ของทั้งหมดอีก 1 ฉบับ เก็บข้อมูลโดยให้นักเรียนทำแบบสอบถามคนละฉบับ ผลการวิจัยพบว่าอัตราส่วนของจำนวนข้อกระทงของแบบสอบถามแบบเลือกตอบต่อแบบถูกผิดที่มีข้อผิดประมาณ 50 % และแบบสอบถามแบบเลือกตอบต่อแบบสอบถามแบบถูกผิดที่มีข้อผิดประมาณ 67 % ที่นักเรียนทำได้ในเวลาเท่ากัน มีค่าประมาณ 1 : 1.16 และ 1 : 1.13 ตามลำดับ ค่าความเที่ยงของแบบสอบถามแบบเลือกตอบสูงกว่าแบบสอบถามแบบถูกผิดทั้ง 2 ฉบับ อย่างมีนัยสำคัญทางสถิติ สำหรับของแบบสอบถามแบบถูกผิดทั้งสองฉบับแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าความยากของแบบสอบถามแบบเลือกตอบต่างจากแบบสอบถามแบบถูกผิดทั้งสองฉบับอย่างมีนัยสำคัญทางสถิติ แต่สำหรับแบบสอบถามแบบถูกผิดทั้งสองฉบับแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ