



## บทที่ 2

### วรรณคดีที่เกี่ยวข้อง

วรรณคดีที่เกี่ยวข้องจะแบ่งออกเป็น 2 ตอน ตอนที่ 1 เป็นมโนทัศน์เบื้องต้นเกี่ยวกับแบบสอบเลือกตอบทั่วไปเลือกสรรมาและทั่วไปเลือกข้อ ความเที่ยง ความตรง ความยาก และอำนาจจำแนกของแบบสอบ ตอนที่ 2 เป็นงานวิจัยที่เกี่ยวข้อง ซึ่งผู้วิจัยจะได้นำมาถึงรายละเอียดอีกต่อไป

ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับแบบสอบเลือกตอบทั่วไปเลือกสรรมาและทั่วไปเลือกข้อ ความเที่ยง ความตรง ความยาก และอำนาจจำแนกของแบบสอบ

#### แบบสอบเลือกตอบทั่วไปเลือกสรรมาและทั่วไปเลือกข้อ

แบบสอบเลือกตอบเป็นแบบสอบที่ได้รับความนิยมมาก เพราะมีส่วนดีหลายประการ คือ ไม่ทำให้เกิดความกำกวมของคำถาม สามารถถามได้ครอบคลุมเนื้อหา วัเคราะห์กับความรู้ชั้นต่าง ๆ ได้อย่างกว้างขวาง การตรวจให้คะแนนได้ผลคงที่ ประหยัดเวลาและแรงงานในการตรวจ นอกจากนั้นยังสามารถวิเคราะห์ได้ว่าข้อใดถูกหรือไม่ก็ได้ อีกด้วย (ชวาล แพทย์กุล 2516: 164) ในการวัดผลสัมฤทธิ์นั้นสิ่งสำคัญประการหนึ่งคือ การวัดผลการเรียนรู้ในระดับความรู้-ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินผล (Gronlund 1976: 270) จะเห็นได้ว่าแบบสอบเลือกตอบสามารถวัดวัเคราะห์กับความรู้เหล่านี้ได้ดี และได้มีการพัฒนาวิธีการเขียนข้อระหวงแบบเลือกตอบให้มีข้อความและทั่วไปเลือกแตกต่างกันออกไป เพื่อใช้ในการวัดผลการเรียนรู้ในระดับความรู้ชั้นต่าง ๆ ได้เหมาะสม

#### แบบสอบเลือกตอบทั่วไปเลือกสรรมา

ข้อระหวงแบบเลือกตอบทั่วไปเลือกสรรมา ประกอบด้วย 2 ส่วน คือ ส่วนที่เป็นคำถาม และส่วนที่เป็นทั่วไปเลือก

คำถามของแบบสอบถามเลือกตอบสามารถเขียนได้ 2 แบบ คือ คำถามที่ให้ผู้สอบเลือกตัวเลือกที่ถูกจากตัวเลือกอื่น ๆ ที่ผิด และคำถามที่ให้ผู้สอบเลือกตัวเลือกที่ถูกที่สุดจากตัวเลือกอื่น ๆ ที่ถูกต่อน้อยกว่า ซึ่งข้อกระทงที่ใช้คำถามประเภทนี้จะยากกว่าข้อกระทงที่ใช้คำถามที่ให้ผู้สอบเลือกตัวเลือกที่ถูกจากตัวเลือกอื่น ๆ ที่ผิด (Gronlund 1978: 190)

การเขียนตัวเลือกของแบบสอบถามเลือกตอบสามารถเขียนได้หลายแบบ ทั้งนี้ต้องให้สอดคล้องกับคำถามด้วย คำถามที่ให้ผู้สอบเลือกตัวเลือกที่ถูกจากตัวเลือกอื่น ๆ ที่ผิดจะใช้ได้กับตัวเลือกทุกรูปแบบ แต่คำถามที่ให้ผู้สอบเลือกตัวเลือกที่ดีที่สุดจากตัวเลือกอื่น ๆ ที่ถูกต่อน้อยกว่าจะใช้ได้เหมาะสมในบางสถานการณ์เท่านั้น กล่าวคือ ใช้ได้กับตัวเลือกธรรมชาติทั่วไป แต่ไม่ควรใช้กับตัวเลือกปลายเปิด "ไม่มีข้อใดถูก" และตัวเลือกปลายปิด "ถูกทุกข้อ" (Thorndike 1971: 117) เนื่องจากผู้สอบจะไม่เลือกตัวเลือกแบบปลายเปิดเมื่อเห็นว่าตัวเลือกใดตัวหนึ่งถูกต้อง หรือผู้สอบจะเลือกตัวเลือกแบบปิดเมื่อพบว่าตัวเลือกทุกตัวถูกต้องเพียงแต่ถูกมากหรือน้อยกว่ากันเท่านั้น จากที่กล่าวมานี้จะเห็นได้ว่าเมื่อใช้คำถามที่มีความซับซ้อนมากขึ้น จะทำให้การเขียนตัวเลือกมีข้อจำกัดมากขึ้น

### แบบสอบถามเลือกตอบตัวเลือกซ้อน

#### การเขียนข้อกระทงแบบเลือกตอบตัวเลือกซ้อน

ข้อกระทงแบบเลือกตอบตัวเลือกซ้อนประกอบด้วย 3 ส่วน คือ ส่วนที่เป็นคำถาม ส่วนที่เป็นข้อความหรือสถานการณ์ และส่วนที่เป็นตัวเลือก วิธีการเขียนข้อกระทงแบบเลือกตอบตัวเลือกซ้อนกระทำโดยเขียนข้อความเช่นเดียวกับข้อกระทงแบบเลือกตอบตัวเลือกธรรมชาติ เขียนข้อความหรือสถานการณ์เช่นเดียวกับตัวเลือกของข้อกระทงแบบเลือกตอบตัวเลือกธรรมชาติ แล้วเขียนตัวเลือกจากข้อความหรือสถานการณ์เหล่านั้น

การเขียนตัวเลือกของข้อกระทงแบบเลือกตอบตัวเลือกซ้อนนี้เกิดจากการกำหนดข้อความหรือสถานการณ์มาก่อน แล้วเขียนตัวเลือกจากข้อความหรือสถานการณ์เหล่านั้น ด้วยวิธีการนี้จะทำให้สามารถเขียนตัวเลือกที่เป็นไปได้หลายตัวเลือก เช่น เมื่อกำหนดข้อความหรือสถานการณ์มาให้ 3 ข้อ จะสามารถเขียนตัวเลือกได้ 8 ตัวเลือก ทั้งนี้คือ ข้อ 1, ข้อ 2, ข้อ 3, ข้อ 1 และข้อ 2, ข้อ 1 และข้อ 3, ข้อ 2 และข้อ 3, ข้อ 1 ข้อ 2 และข้อ 3, และไม่มีข้อใดถูกต้อง ในทำนองเดียวกันเมื่อกำหนดข้อความหรือสถานการณ์มา 4 หรือ 5 ข้อ จะสามารถเขียนตัวเลือกได้ 16 หรือ 32 ตัวเลือกตามลำดับ

ซึ่งจะทำให้ผู้สร้างข้อสอบสามารถเลือกเฉพาะตัวเลือกที่น่าสนใจไว้ในข้อกระทงเท่านั้น ข้อที่ควรพิจารณาในการเลือกตัวเลือกใดไว้ในข้อกระทง ควรเป็นตัวเลือกที่ทำให้ข้อของ ข้อความหรือสถานการณ์ (คือ ข้อ 1, ข้อ 2, \_\_\_) ปรากฏในตัวเลือกทั้งหมดด้วยจำนวนครั้งที่เท่ากัน (Hubbard 1978; Whitley 1979 อ้างถึงใน Kolstad and others 1983: 45) ทั้งนี้เพราะข้อของข้อความหรือสถานการณ์ที่ปรากฏในตัวเลือกหลายครั้งจะมีผลต่อการ เคาของผู้สอบ ทั้งตัวอย่างข้อกระทงที่แสดงให้เห็นข้างล่างนี้

#### อาหารชนิดใดมีคอเรสเตอรอลสูง

1. เนื้อ
  2. ไก่
  3. ไข่
  4. กุ้งฝอย
  5. ข้าว
- ก. ข้อ 5  
ข. ข้อ 2 และ 4  
ค. ข้อ 1, 3 และ 4  
ง. ข้อ 1, 2, 3 และ 5  
จ. ข้อ 1, 2, 3, 4 และ 5

ข้อกระทงในตัวอย่างนี้จะกำหนดข้อความมา 5 ข้อความ แล้วเขียนตัวเลือก 5 ตัวเลือกจากข้อความหรือสถานการณ์ที่กำหนดมาให้ นั่น โดยข้อความแต่ละข้อจะปรากฏในตัวเลือกทั้งหมด 3 ครั้งเท่ากัน คือ ข้อ 1 จะปรากฏในตัวเลือก ก, ง, จ ข้อ 2 จะปรากฏในตัวเลือก ข, ง, จ ข้อ 3 จะปรากฏในตัวเลือก ค, ง, จ ข้อ 4 จะปรากฏในตัวเลือก ข, ค, จ และข้อ 5 จะปรากฏในตัวเลือก ก, ง, จ

การเขียนข้อกระทงแบบเลือกตอบตัวเลือกข้อ 5 ตัวเลือก จำนวนครั้งที่ข้อของข้อความหรือสถานการณ์ปรากฏในตัวเลือกทั้งหมด 3 ครั้งเท่ากันทุกข้อจะทำให้แบบสอบมีค่าความเที่ยงสูงสุด (Tversky 1964; Lord 1977 อ้างถึงใน Kolstad and Other 1983: 45) ในการสร้างแบบสอบวัดผลสัมฤทธิ์ควรระลึกถึงหลักเกณฑ์ข้อนี้ แต่ไม่ควรยึดเป็นกฎเกณฑ์ตายตัว เพราะต้องพิจารณาความเป็นไปได้ของเนื้อหาที่ออกข้อสอบ

ถ้ายมีฉะนั้นตัวเลือกบางตัวจะเป็นตัวเลือกที่ผิดชัดเจนและไม่มีผู้ใดเลือกตัวเลือกนั้น

เนื่องจากการใช้ข้อระงแบบเลือกตอบที่ใช้ตัวเลือกแบบปลายเปิด "ถูกทุกข้อ" จะทำให้เกิดความคลาดเคลื่อนในการวัดได้ เพราะนักเรียนที่มีความชำนาญในการทำแบบสอบจะสามารถเดาคำตอบได้ ทั้งตัวอย่างข้อระงต่อไปนี้ เมื่อนักเรียนทราบว่าตัวเลือก 2 ตัวเลือกถูกต้อง นักเรียนจะเลือกตอบตัวเลือก จ. ซึ่งเป็นตัวเลือกที่ถูกต้อง โดยไม่จำเป็นต้องทราบว่าตัวเลือกทุกตัวถูกต้อง

นักมีลักษณะดังข้อใด

- ก. เป็นสัทวเลือกคุณ
- ข. หายใจกวยปอก
- ค. หัวใจมี 4 ห้อง
- ง. จักอยู่ในประเภทสัทวปีก
- จ. ถูกทุกข้อ

ความคลาดเคลื่อนในการวัดที่เกิดจากการใช้ข้อระงแบบเลือกตอบที่ใช้ตัวเลือกแบบปลายเปิด จะทำให้หมดไปได้โดยการเขียนข้อระงให้อยู่ในรูปแบบของข้อระงแบบเลือกตอบตัวเลือกซอน (Copperud 1979: 56)

เวทเทน (Weiten 1982: 47) ได้กล่าวถึงข้อจำกัดของแบบสอบเลือกตอบตัวเลือกซอนไว้ดังนี้

1. การเขียนข้อระงแบบเลือกตอบตัวเลือกซอนไม่เป็นไปตามกฎเกณฑ์การเขียนข้อระงแบบเลือกตอบที่ว่าควรให้สั้นและกระชับ ซึ่งจะมีผลต่อการตอบถูกหรือผิดของผู้สอบกวย
2. ความยาวและความซับซ้อนของข้อระงมีผลให้ผู้ที่มิทักษะการอ่านและการให้เหตุผลจะตอบแบบสอบเลือกตอบตัวเลือกซอนได้กีกกว่า ซึ่งจะเป็นการเพิ่มความแปรปรวนคลาดเคลื่อน (Error variance)
3. ความสามารถในการทำแบบสอบเลือกตอบตัวเลือกซอนขึ้นอยู่กับความชำนาญในการทำแบบสอบ (Test-wiseness) และทักษะในการทำแบบสอบ (Test-taking skill) ของผู้สอบแต่ละคน

### หลักการและเหตุผล

ข้อควรระวังแบบเลือกตอบตัวเลือกซ้อนสร้างขึ้นเพื่อให้สามารถสร้างข้อสอบวัดระดับความรู้ชั้นต่าง ๆ ได้อย่างเหมาะสม ทั้งนี้เนื่องจากมีนักวิจัยหลายคนได้ให้ความเห็นเกี่ยวกับแบบสอบเลือกตอบตัวเลือกซ้อนไว้ดังนี้ คือ อีเบล (Ebel 1972: 103) กล่าวว่า ในการวัดระดับความรู้-ความจำ ไม่ควรใช้แบบสอบเลือกตอบตัวเลือกซ้อน ควรใช้แบบสอบเลือกตอบตัวเลือกธรรมดา หรือแบบสอบถูกผิด เวสแมน (Wesman อ้างถึงใน Thorndike 1971: 97-98) มีความเห็นว่าแบบสอบเลือกตอบตัวเลือกซ้อนควรใช้วัดความรู้ ความสามารถที่ซับซ้อน เช่นการรวบรวมสิ่งต่าง ๆ เพื่อหาข้อสรุป การประเมินสถานการณ์หรือมโนทัศน์ เป็นต้น ทอร์นไคค์ และเฮเกน (Thorndike and Hagen 1977: 240) มีความเห็นว่าแบบสอบเลือกตอบตัวเลือกซ้อนใช้ได้อย่างมีประสิทธิภาพ เกี่ยวกับการนำความรู้หลักการ ไปประยุกต์ใช้ซึ่งคงพิจารณาเป็นกรณีไป ในทำนองเดียวกับกรอนลันด์ (Gronlund 1977: 63-64) ที่มีความเห็นว่าแบบสอบเลือกตอบตัวเลือกซ้อนควรใช้ในสถานการณ์ที่เหมาะสม เช่น การประยุกต์ความจริง กฎ หลักการ มโนทัศน์ออกมาเป็นผลการเรียนรู้ เป็นต้น

จากความเห็นของนักวิจัยที่กล่าวมานี้ แสดงให้เห็นว่า แบบสอบเลือกตอบตัวเลือกซ้อนเหมาะสำหรับวัดผลการเรียนรู้ในระดับความรู้ขั้นสูงขึ้นไป

### ความเที่ยงของแบบสอบ

ความเที่ยง หมายถึง ระดับ (Degree) ของความคงที่ระหว่างการวัดสิ่งเดียวกัน 2 ครั้ง (Mehrens and Lehmann 1978: 88)

### ทฤษฎีความเที่ยง (Theory of Reliability)

ทฤษฎีความเที่ยงของคะแนนอธิบายไว้ดังนี้คือ คะแนนที่วัดได้จะประกอบด้วย คะแนน 2 ส่วน คือ คะแนนที่ได้จากสมรรถภาพที่แท้จริง เรียกว่า คะแนนจริง (True Score) และคะแนนความคลาดเคลื่อน (Error Score) ซึ่งคะแนนความคลาดเคลื่อนนี้เมื่อเกิดขึ้นจะมีผลทำให้คะแนนที่วัดได้ผิดไปจากคะแนนจริง ดังนั้นจึงสามารถเขียนเป็นสมการไว้ดังนี้ (Mehrens and Lehmann 1978: 90)



$$X = T + E$$

เมื่อ  $X$  แทนคะแนนที่วัดได้

$T$  แทนคะแนนจริง

$E$  แทนคะแนนความคลาดเคลื่อน

คะแนนจริง (True Score) หมายถึง คะแนนที่ผู้สอบทำได้จากการวัด  
ด้วยเครื่องมือที่มีคุณภาพสมบูรณ์ปราศจากความคลาดเคลื่อน หรือหมายถึงคะแนนเฉลี่ยของ  
บุคคลที่ได้จากการทำแบบสอบถาม หรือแบบสอบถามหลาย ๆ ครั้ง (Brown 1970: 50)

คะแนนความคลาดเคลื่อน (Error Score) หมายถึง ค่าความผิดพลาด  
ที่เกิดจากการวัด ซึ่งเป็นความคลาดเคลื่อนที่เกิดขึ้นโดยสุ่ม (Random Error) คะแนน  
ความคลาดเคลื่อนนี้อาจเกิดขึ้นจากสาเหตุหลายอย่าง เช่น เกิดจากตัวแบบสอบ การให้  
คะแนน การดำเนินการสอบ หรือเกิดจากตัวผู้เข้าสอบเอง โดยทฤษฎีความคลาดเคลื่อนที่  
เกิดขึ้นนี้อาจเป็นได้ทั้งทางบวกและทางลบ ซึ่งจะหักล้างกันหมดไป นั่นคือค่าเฉลี่ยของความ  
คลาดเคลื่อนจะเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 คะแนนความคลาดเคลื่อน  
นี้จะไม่มีความสัมพันธ์กับคะแนนจริง นั่นคือสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจริงกับคะแนน  
ความคลาดเคลื่อนจะมีค่าเท่ากับ 0 ดังนั้นความแปรปรวนของคะแนนที่วัดได้สามารถเขียน  
เป็นสมการได้ดังนี้ (Mehrens and Lehmann 1978: 91)

$$s_x^2 = s_t^2 + s_e^2$$

เมื่อ  $s_x^2$  แทนความแปรปรวนของคะแนนที่วัดได้

$s_t^2$  แทนความแปรปรวนของคะแนนจริง

$s_e^2$  แทนความแปรปรวนของคะแนนความคลาดเคลื่อน

เนื่องจากความเที่ยงหมายถึง อัตราส่วนระหว่างความแปรปรวนของ  
คะแนนจริงกับความแปรปรวนของคะแนนที่วัดได้ (Brown 1970: 51) จึงสามารถเขียน  
สมการได้เป็น

$$r_{tt} = \frac{s_t^2}{s_x^2}$$

เมื่อ  $r_{tt}$  แทนค่าความเที่ยง

เนื่องจาก 
$$s_t^2 = s_x^2 + s_e^2$$

ดังนั้น 
$$r_{tt} = 1 - \frac{s_e^2}{s_x^2}$$

จะเห็นว่าค่าความเที่ยงจะสูงขึ้นเมื่อความแปรปรวนคลาดเคลื่อนลดลง โดยกำหนดให้คะแนนผลการสอบมีค่าคงที่หรือเพิ่มขึ้น และค่าความเที่ยงจะมีค่าสูงสุดเท่ากับ 1

### วิธีประมาณค่าความเที่ยง

การประมาณค่าความเที่ยงสามารถทำได้หลายวิธี เมื่อต้องการใช้วิธีใดของพิจารณาถึงสถานการณ์การทดสอบและลักษณะของคะแนนที่ได้จากผลการสอบนั้น วิธีการประมาณค่าความเที่ยงโดยทั่วไปมี 4 แบบ (Mehrens and Lehmann 1978: 94) คือ

1. แบบสอบซ้ำ (Measure of Stability)
2. แบบไขข้อสอบคล้ายกัน (Measure of Equivalence)
3. แบบไขข้อสอบคล้ายกันและสอบซ้ำ (Measure of Equivalence and Stability)
4. แบบวัดความคงที่ภายใน (Measure of Internal Consistency)

วิธีการนี้แบ่งย่อยออกได้เป็น

- 4.1 แบบแบ่งครึ่งข้อสอบ (Split-half)
- 4.2 แบบคูเคอร์ริชาร์ดสัน (Kuder-Richardson Estimates)
- 4.3 แบบวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt's Analysis of Variance Procedure)

ในที่นี้จะขอกล่าวถึงเฉพาะวิธีประมาณค่าความเที่ยงแบบคูเคอร์ ริชาร์ดสัน ซึ่งเป็นวิธีการประมาณค่าความเที่ยงในการวิจัยครั้งนี้

### วิธีประมาณค่าความเที่ยงแบบคูเคอร์ริชาร์ดสัน

การประมาณค่าความเที่ยงแบบคูเคอร์ ริชาร์ดสัน เป็นการประมาณค่าความเที่ยงที่คูเคอร์และริชาร์ดสันคิดขึ้น (Kuder and Richardson) ได้เสนอไว้ในปี

ค.ศ. 1937 มีหลายสูตร ซึ่งเป็นการประมาณค่าความเที่ยงแบบความคงที่ภายใน สูตรที่นิยมใช้กันมากที่สุดคือ คูเคอร์ ริชาร์ดสัน 20 (Ebel 1965: 318)

$$r = \frac{k}{k-1} \left[ 1 - \frac{\sum pq}{62} \right]$$

เมื่อ  $r$  แทนค่าความเที่ยงของแบบสอบถาม

$k$  แทนจำนวนข้อในแบบสอบถาม

$p$  แทนสัดส่วนผู้ตอบถูกในแต่ละข้อ

$q$  แทนสัดส่วนผู้ตอบผิดในแต่ละข้อ

$62$  แทนความแปรปรวนของคะแนนทั้งหมด

การประมาณค่าความเที่ยงแบบคูเคอร์ ริชาร์ดสันนี้ เป็นการหาความสอดคล้องกันระหว่างข้อสอบ ซึ่งมีการดำเนินการสอบครั้งเดียวและใช้แบบสอบถามเดียว ความสอดคล้องกันระหว่างข้อสอบนี้ ได้รับอิทธิพลจากแหล่งความแปรปรวนคลาดเคลื่อน 2 แหล่ง คือ

1. เนื้อหาที่สุ่ม

2. ความเป็นวิวิธพันธ์ของพฤติกรรมที่สุ่ม ยิ่งข้อสอบมีความเป็นเอกพันธ์มาก ความสอดคล้องกันระหว่างข้อก็ยิ่งสูงขึ้น

ข้อตกลงเบื้องต้นของสูตรคูเคอร์ ริชาร์ดสัน 20 คือ

1. การตรวจให้คะแนนในแต่ละข้อเป็นแบบ 0-1 คือ เมื่อตอบถูกจะให้ 1 คะแนน เมื่อตอบผิดจะให้ 0 คะแนน

2. ข้อสอบในแบบสอบถามจะต้องมีลักษณะเป็นเอกพันธ์ คือ วัดคุณลักษณะเดียวกัน

เนื่องจากสูตรคูเคอร์ ริชาร์ดสัน 20 ไม่สะดวกในการคำนวณด้วยมือ คูเคอร์และริชาร์ดสัน (Kuder and Richardson) จึงได้เสนอสูตรที่สามารถคำนวณได้ง่ายกว่า คือ สูตรคูเคอร์ ริชาร์ดสัน 21 (Kuder Richardson 21) โดยสูตรนี้มีข้อตกลงเบื้องต้นว่า ข้อสอบแต่ละข้อมีค่าความยากโดยเฉลี่ยเท่ากัน หรือกำหนดค่า  $p$  คงที่ ซึ่งจะเห็นได้ว่าโดยทั่วไปแล้วข้อสอบแต่ละข้อในแบบสอบถามจะมีค่าความยากแตกต่างกัน ซึ่งไม่เป็นไปตามข้อตกลงดังกล่าว จึงทำให้ค่าความเที่ยงที่ประมาณได้จากสูตรคูเคอร์ ริชาร์ดสัน 21 น้อยกว่าค่าความเที่ยงที่ประมาณได้จากสูตรคูเคอร์ ริชาร์ดสัน 20



## องค์ประกอบที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบ

นอกจากวิธีการประมาณค่าความเที่ยงของแบบสอบจะมีอิทธิพลต่อค่าความเที่ยงของแบบสอบแล้ว ยังมีองค์ประกอบอื่น ๆ ที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบด้วย (Mehrens and Lehmann 1973: 100-103) คือ

1. ความยาวของแบบสอบ (Test Length) การเพิ่มข้อสอบที่มีคุณภาพเท่าเทียมกับข้อสอบที่มีอยู่เดิม จะทำให้แบบสอบมีค่าความเที่ยงสูงขึ้น นั่นคือแบบสอบที่มีจำนวนข้อมากจะมีความเที่ยงสูงกว่าแบบสอบที่มีจำนวนข้อน้อย ความสัมพันธ์ระหว่างความยาวและความเที่ยงของแบบสอบพิจารณาได้จากสูตรของสเปียร์แมน-บราวน์ (Spearman - Brown)

2. ความเร็ว (Speed) การประมาณค่าความเที่ยงของแบบสอบความเร็ว (Speed Test) ไม่ควรใช้วิธีการวัดความคงที่ภายใน ทั้งนี้เพราะแบบสอบชนิดนี้เป็นแบบสอบที่ง่ายแต่ใช้เวลาทำการสอบน้อย จะทำให้ผู้สอบไม่สามารถทำได้ครบทุกข้อ ซึ่งจะมีผลทำให้ค่าความเที่ยงของแบบสอบสูงกว่าที่ควรจะเป็น การประมาณค่าความเที่ยงแบบวัดความคงที่ภายในควรรีใช้เมื่อผู้สอบทุกคนมีโอกาสทำข้อนั้นแล้ว

3. ความเป็นเอกพันธ์ของกลุ่ม (Group Homogeneity) กลุ่มที่มีความสามารถเป็นเอกพันธ์ จะมีพิสัยของคะแนนความสามารถน้อย ทำให้มีการกระจายของคะแนนน้อย สำหรับกลุ่มที่มีความสามารถเป็นวิวิธพันธ์ (Group Heterogeneous) จะมีพิสัยของคะแนนความสามารถมาก ทำให้มีการกระจายของคะแนนมากกว่า ซึ่งจะมีผลให้ค่าความเที่ยงของแบบสอบจากกลุ่มที่เป็นเอกพันธ์ต่ำกว่าความเที่ยงของแบบสอบจากกลุ่มที่เป็นวิวิธพันธ์

4. ความยากของข้อกระทง (Difficulty of items) เนื่องจากค่าความเที่ยงของแบบสอบนั้นประมาณจากความแปรปรวนของคะแนนผลการสอบ ถ้าข้อสอบยากเกินไปจะทำให้ทุกคนทำผิดหมด หรือถ้าข้อสอบง่ายเกินไปจะทำให้ทุกคนทำถูกหมด จะทำให้ความแปรปรวนของคะแนนมีน้อย จะมีผลให้ค่าความเที่ยงลดลงด้วย ถ้าความยากของข้อสอบทำให้คะแนนกระจายมาก จะทำให้ค่าความเที่ยงของแบบสอบสูงขึ้น

5. ความเป็นปรนัย (Objectivity) แบบสอบที่มีความเป็นปรนัยสูง จะทำให้ค่าความเที่ยงของแบบสอบสูงด้วย

นอกจากนี้ความคลาดเคลื่อนของการประมาณค่าความเที่ยงอาจเกิดจากการบริหารการสอบ เช่น การจับเวลา การให้คำชี้แจง การป้องกันการทุจริต หรือองค์ประกอบเกี่ยวกับตัวผู้สอบเอง เช่น แรงจูงใจ ความวิตกกังวล ประสิทธิภาพในการสอบ การเกา เป็นต้น อาจมีผลทำให้ค่าความเที่ยงของแบบสอบเปลี่ยนแปลงได้

### ความทรงของแบบสอบ

ความทรง หมายถึง คุณสมบัติของแบบสอบที่สามารถวัดในสิ่งที่ต้องการวัด (Ebel 1965: 377)

### ชนิดของความทรง

การแบ่งประเภทของความทรงมีหลายวิธี ทั้งนี้ขึ้นอยู่กับจุดมุ่งหมายของการวัด โดยทั่วไปการวัดทางการศึกษาและจิตวิทยาแบ่งความทรงออกเป็น 3 ประเภท (Mehrens and Lehmann 1973: 109)

1. ความตรงตามเนื้อเรื่อง (Content Validity)
2. ความตรงตามเกณฑ์สัมพันธ์ (Criterion Related Validity)
3. ความตรงตามทฤษฎี (Construct Validity)

ในที่นี้จะกล่าวถึงความตรงตามเนื้อเรื่อง และความตรงตามเกณฑ์สัมพันธ์ ซึ่งเกี่ยวข้องกับการวิจัยครั้งนี้เท่านั้น

### ความตรงตามเนื้อเรื่อง

ความตรงตามเนื้อเรื่องของแบบสอบ หมายถึง ความสามารถของแบบสอบที่จะวัดสถานการณ์ซึ่ง เป็นตัวแทนของ เนื้อเรื่องในขอบเขตที่ต้องการวัดได้ (เขาวดี วิบูลย์ศรี 2526: 71)

การที่จะตัดสินว่าแบบสอบใดมีความตรงตามเนื้อเรื่องต้องพิจารณาถึงรายละเอียดของเนื้อหาวิชา และชนิดของพฤติกรรมที่ต้องการวัดไปพร้อมกัน โดยทั่วไปจะพิจารณาข้อกระทงในแบบสอบเทียบกับแผนผังการสร้าขอสอบแล้วประมาณออกมาว่าข้อสอบนั้น มีความตรงตามเรื่องมากน้อยเพียงใด ซึ่งผู้ที่ตัดสินความตรงตามเนื้อเรื่องได้ก็ ต้อง เป็นผู้เชี่ยวชาญในสาขาวิชานั้น ๆ (Mehrens and Lehmann 1978: 111) ในการสร้างแบบสอบผลสัมฤทธิ์ต้องให้มีความตรงตามเนื้อเรื่อง เป็นสำคัญ

### ความตรงตามเกณฑ์สัมพัทธ์

ความตรงตามเกณฑ์สัมพัทธ์ของแบบสอบแสดงถึงความสามารถของแบบสอบในการทำนายพฤติกรรมของบุคคลในสถานการณ์เฉพาะ หรือแสดงถึงความสัมพันธ์ระหว่างคะแนนที่ผู้สอบกระทำได้จากแบบสอบกับเกณฑ์ภายนอก (External Criteria) ที่เป็นอิสระ โดยทั่วไปแล้วความตรงตามเกณฑ์สัมพัทธ์จะใช้เมื่อมีความต้องการที่จะประมาณจุดยืนที่น่าจะเป็นไปได้ของผู้สอบบนเกณฑ์ที่อ้างอิงถึง (เขาวที วิบูลย์ศรี 2526: 73)

ความตรงตามเกณฑ์สัมพัทธ์จำแนกได้เป็น 2 ชนิด คือ

1. ความตรงร่วมสมัย (Concurrent Validity) หมายถึง ความตรงของแบบสอบที่สามารถบอกให้เรารู้ว่าสิ่งที่วัดถูกต้องตามสภาพที่แท้จริงในปัจจุบัน โดยการหาความสัมพันธ์ระหว่างคะแนนของแบบสอบกับคะแนนเกณฑ์ที่กำหนดขึ้นในเวลาเดียวกัน หรือใกล้เคียงกัน เช่น ต้องการทราบความตรงร่วมสมัยของแบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ กับตัวแปรเกณฑ์อื่น ๆ ก็อาจหาได้จาก การหาความสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบ กับ การสังเกตพฤติกรรมกระบวนการทางวิทยาศาสตร์ในขณะนักเรียนทำการทดลองในห้องเรียน ถ้าคะแนนที่นักเรียนทำได้จากแบบสอบได้สอดคล้องคล่องกับความสามารถที่แสดงออกจริงซึ่งใช้เป็นเกณฑ์ แสดงว่าแบบสอบนั้นมีความตรงร่วมสมัย

2. ความตรงเชิงทำนาย (Predictive Validity) หมายถึง ความตรงของแบบสอบที่จะบอกสิ่งที่วัดได้ถูกต้องตามสภาพที่เป็นจริงในอนาคต โดยหาความสัมพันธ์ระหว่างคะแนนของแบบสอบกับคะแนนเกณฑ์ที่จะเกิดขึ้นในอนาคต เช่น ใช้แบบสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ทำนายผลสัมฤทธิ์ในการเรียนวิชาวิทยาศาสตร์ซึ่งอาจใช้คะแนนผลการสอบวิชาวิทยาศาสตร์ปลายปีเป็นเกณฑ์ การหาความตรงชนิดนี้ต้องอาศัยเวลา เพราะต้องหาความสัมพันธ์ของคะแนนที่ได้จากแบบสอบกับคะแนนที่เกิดขึ้นภายหลัง ความตรงเชิงทำนายนี้จะมีในแบบสอบที่ใช้ทำนายผล

ความถูกต้องของความตรงตามเกณฑ์สัมพัทธ์นั้นขึ้นอยู่กับการวัดเกณฑ์ ผู้วัดจะต้องทราบลักษณะของสิ่งที่ใช้เป็นเกณฑ์เป็นอย่างดี โดยพิจารณาจากสิ่งต่อไปนี้

(Mehrens and Lehmann 1978: 112-114) คือ

1. ความเกี่ยวข้อง (Relevance) การที่เราจะเลือกสิ่งใดมาเป็นเกณฑ์นั้น สิ่งนั้นควรมีความเกี่ยวข้อง หรือสะท้อนถึงคุณลักษณะที่สำคัญของเรื่องที่จะศึกษา โดยทั่วไปการพิจารณาความเกี่ยวข้องของเกณฑ์ขึ้นอยู่กับการตัดสินคุณค่า ไม่สามารถวัดออกมาเป็นตัวเลขได้

2. ความเที่ยง (Reliability) เนื่องจากความเที่ยงของแบบสอบถามมีอิทธิพลต่อค่าสหสัมพันธ์ระหว่างแบบสอบถามกับเกณฑ์ จากทฤษฎีความสัมพันธ์กล่าวว่า ความสัมพันธ์ระหว่างตัวแปร 2 ตัวแปร จะเท่ากับ หรือน้อยกว่ารากที่สองของผลคูณระหว่างค่าความเที่ยงของตัวแปรทั้งสองนั้น นั่นคือ

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

เมื่อ  $r_{xy}$  แทนค่าสหสัมพันธ์ระหว่างตัวทำนาย (x) กับ  
เกณฑ์ (y)

$r_{xx}$  แทนค่าความเที่ยงของแบบสอบถามที่เป็นตัวทำนาย

$r_{yy}$  แทนค่าความเที่ยงของเกณฑ์

จะเห็นว่าค่าความเที่ยงของ เกณฑ์มีอิทธิพลต่อความตรงตามเกณฑ์สัมพันธ์พอ ๆ กับค่าความเที่ยงของตัวทำนาย

3. ความเป็นอิสระจากอคติ (Free From bias) เกณฑ์ที่ใช้ควรเป็นอิสระ และไม่มีอิทธิพลจากการทดสอบที่นำมาใช้เป็นตัวทำนาย เช่น ครูนำแบบสอบถามความถนัดทางวิชาการมาทดสอบกับนักเรียนตอนต้นปี เพื่อพยากรณ์ผลการเรียนของนักเรียนปลายปี แล้วครูนำความรู้จากผลการสอบครั้งนี้ไปมีผลต่อการประเมินผลการเรียนปลายปี คือ ถ้านักเรียนคนใดใคร่คะแนนแบบสอบถามความถนัดค่าจะประเมินผลการเรียนให้นักเรียนผู้นั้นต่ำไปควย ซึ่งจะทำให้ตัวทำนายมีอิทธิพลต่อคะแนนเกณฑ์ อันจะเป็นผลให้ความสัมพันธ์ระหว่างแบบสอบถามกับเกณฑ์สูงขึ้น การป้องกันความคลาดเคลื่อนในลักษณะนี้ทำได้โดยอย่าให้ผู้ที่ทำหน้าที่ประเมินผลการเรียนทราบคะแนนสอบที่ใช้เป็นตัวทำนาย

#### วิธีการประมาณค่าความตรงตามเกณฑ์สัมพันธ์

การประมาณค่าความตรงตามเกณฑ์สัมพันธ์สามารถหาได้จากสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบสอบถามกับเกณฑ์ สูตรที่ใช้คำนวณจึงเป็นสูตรสหสัมพันธ์

แบบต่าง ๆ โดยทั่วไปนิยมใช้สูตรสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson's Product Moment Correlation Coefficient)

องค์ประกอบที่มีอิทธิพลต่อค่าความตรงของแบบสอบ

องค์ประกอบที่มีอิทธิพลต่อค่าความตรงของแบบสอบมีดังนี้ (Gronlund 1976: 98-102) คือ

1. องค์ประกอบภายในแบบสอบ มีดังต่อไปนี้
  - 1.1 ความไม่ชัดเจนของคำชี้แจงในการทำแบบสอบ ทำให้นักเรียนปฏิบัติตามคำสั่งผิดพลาด
  - 1.2 การใช้คำหรือประโยคยากเกินไป ทำให้นักเรียนไม่เข้าใจวัตถุประสงค์ที่ต้องการวัดจึงตอบข้อสอบผิดพลาดไป แม้จะมีความรู้ในเรื่องเหล่านั้นก็ตาม ดังนั้นจึงเป็นการวัดความสามารถทางภาษามากกว่าวัดความรู้ในเนื้อหาที่เรียน
  - 1.3 ระดับความยากของข้อระงไม่เหมาะสม ทำให้จำแนกความสามารถของนักเรียนไม่ได้
  - 1.4 การสร้างข้อสอบไม่ระมัดระวัง จะทำให้เป็นการแนะนำคำตอบนักเรียนได้
  - 1.5 ความกำกวมของภาษาที่ใช้ทำให้นักเรียนตีความผิดไป
  - 1.6 ใช้ประเภทของแบบสอบไม่เหมาะสมกับพฤติกรรมที่ต้องการวัด
  - 1.7 แบบสอบมีจำนวนข้อระงน้อยเกินไป ทำให้ไม่สามารถวัดได้ครอบคลุมวัตถุประสงค์ที่ต้องการวัด
  - 1.8 จัดเรียงข้อระงไม่เหมาะสม เช่น เรียงข้อระงจากยากไปหาง่ายจะทำให้นักเรียนท้อใจเวลาทำแบบสอบนานจนบางคนอาจไม่มีเวลาทำข้อสอบง่าย ที่อยู่ตอนท้ายของแบบสอบ ควรจัดเรียงข้อระงจากง่ายไปหายาก ซึ่งจะมีผลต่อแรงจูงใจในการทำแบบสอบด้วย
  - 1.9 คำตอบที่ถูกต้องในแบบสอบมีการจัดเรียงอย่างเป็นระบบ จะทำให้นักเรียนเดาคำตอบได้
2. เนื้อหาและวิธีดำเนินการสอน ในการออกข้อสอบต้องออกให้ตรงกับเนื้อหาที่สอน นอกจากนั้นยังต้องคำนึงถึงระดับความรู้ที่ต้องการวัดด้วย เพราะถ้าต้องการ



วัตถุประสงค์การเรียนรู้ในระดับความรู้ชั้นสูงแล้วถามในเนื้อหาที่เรียนมาแล้วจะไม่ใช่เป็นการวัดผล การเรียนรู้ในระดับความรู้ชั้นสูง จะเป็นเพียงการวัดในระดับความรู้ความจำเท่านั้น

3. การบริหารแบบสอมและการโหดแน่น การบริหารแบบสอมที่ไม่เป็นมาตรฐาน เช่น ให้อายุขัยเกินไป การอธิบายข้อสอบให้นักเรียนบางคนที่ยังรอคอยให้ช่วย การทุจริตของนักเรียนจะมีผลให้ความตรงของแบบสอมลดลง ในทำนองเดียวกันวิธีการโหดแน่นในแบบสอมอักษณิศาไม่เป็นไปตามมาตรฐานเดียวกัน จะทำให้ความตรงของแบบสอมลดลงได้

4. องค์ประกอบเกี่ยวกับตัวนักเรียนในขณะตอบแบบสอม ในการทดสอบแต่ละครั้งนักเรียนจะมีความพร้อมแตกต่างกัน ทั้งนี้เนื่องจากการมีความวิตกกังวล การมีอารมณ์เปลี่ยนแปลงง่าย การมีแรงจูงใจในการทำแบบสอม การมีความสามารถในการเดา สิ่งเหล่านี้จะทำให้นักเรียนสามารถตอบแบบสอมได้แตกต่างจากความสามารถที่แท้จริงของนักเรียน

5. ชรรถนคติของกลุ่มและเกณฑ์ที่ใช้ โดยทั่วไปแบบสอมจะใช้ได้เหมาะสมกับผู้อยู่ในกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ ทั้งนี้เพราะบุคคลจะมีความแตกต่างกันในด้านอายุ เพศ ความสามารถ ภูมิหลังทางการศึกษาและวัฒนธรรม ซึ่งจะทำให้คนในแต่ละกลุ่มสามารถตอบแบบสอมได้แตกต่างกัน ดังนั้นการใช้แบบสอมควรคำนึงถึงลักษณะเฉพาะของกลุ่มที่ใช้หาความตรง นอกจากนั้นยังต้องคำนึงถึงเกณฑ์ที่ใช้หาความตรงของแบบสอมด้วย

#### ความยากของแบบสอม

ความยากของข้อสอบ หมายถึง อัตราส่วนของจำนวนคนทั้งหมดที่ตอบข้อสอบแต่ละข้อได้ถูกต้อง (Hopkins and Stanley 1981: 271)

จากนิยามของค่าความยากจะเห็นว่าค่าความยากสามารถคำนวณได้จากจำนวนนักเรียนที่ตอบข้อสอบนั้นถูกต้องหารด้วยจำนวนนักเรียนที่ตอบข้อสอบข้อนั้นทั้งหมด ดังนั้นค่าความยากของข้อสอบจึงมีค่ามากที่สุดเท่ากับ 1 และน้อยที่สุดเท่ากับ 0 ถ้าค่าความยากสูงหรือมีค่าใกล้เคียง 1 แสดงว่าข้อสอบนั้นง่ายมีนักเรียนจำนวนมากที่สามารถตอบถูก แต่ถ้าระดับค่าความยากต่ำหรือมีค่าใกล้เคียง 0 แสดงว่าข้อสอบนั้นยากมีนักเรียนน้อยคนที่สามารถตอบถูก การพิจารณาระดับความยากที่เหมาะสมนั้นต้องทราบจุดมุ่งหมายของการวัด ในการวัดผลแบบอิงกลุ่ม (Norm Reference) นั้น ค่าความยากของข้อสอบที่เหมาะสมคือ .50

ทั้งนี้เพราะทำให้มีการกระจายของคะแนนมากที่สุด และจำแนกความสามารถของนักเรียนได้มากที่สุด แต่การจะสร้างข้อสอบในแบบสอบใหม่ค่าความยากพอเหมาะทุกข้อนั้นทำได้ยาก ทั้งนี้จึงให้ใช้ข้อสอบที่มีค่าความยากอยู่ในช่วง .20 ถึง .80 ก็นับว่าพอใช้ได้ (ชวาล แพทย์กุล 2509: 316) ส่วนข้อสอบที่มีค่าความยากต่ำกว่า .20 หรือเกินกว่า .80 จะเป็นข้อสอบที่ยากหรือง่ายเกินไป ข้อสอบเหล่านี้จำแนกความสามารถของนักเรียนไม่ได้ หรือจำแนกได้น้อย สำหรับการวัดผลอิงเกณฑ์ (Criterion Reference) ส่วนใหญ่จะเป็นข้อสอบที่มีค่าความยากสูงซึ่งจะเป็นข้อสอบที่ง่าย เพราะสิ่งที่ครูนำมาออกข้อสอบนั้นเป็นเนื้อหาวิชาในหลักสูตร และเป็นสิ่งที่ครูสอนมาแล้วนักเรียนส่วนใหญ่จึงทำข้อสอบได้ อย่างไรก็ตามข้อสอบที่ตรงวัตถุประสงค์การเรียนการสอน ถ้าเป็นข้อสอบที่มีค่าความยากสูงมากหรือต่ำมาก ก็นับว่าเป็นข้อสอบที่ใช้ได้เช่นเดียวกัน

ค่าความยากมีความสัมพันธ์กับค่าอำนาจจำแนก กล่าวคือ เมื่อความยากมีค่าเท่ากับ .50 จะทำให้ค่าอำนาจจำแนกมีค่าสูงสุด แต่การที่ข้อสอบมีค่าความยากเท่ากับ .50 ไม่ก็ทำให้ค่าอำนาจจำแนกสูงสุดเสมอไป บางครั้งค่าอำนาจจำแนกอาจจะต่ำหรือไม่มีเลยก็ได้ เพราะขึ้นอยู่กับจำนวนผู้ตอบถูกในกลุ่มสูงและกลุ่มต่ำ ถ้าผู้ตอบถูกในกลุ่มสูงและกลุ่มต่ำเท่ากันจะทำให้ค่าอำนาจจำแนกเท่ากับ 0

#### องค์ประกอบที่มีอิทธิพลต่อค่าความยากของแบบสอบ

แคมเบล (Campbell 1961: 899-913, อ้างถึงใน นิรมล บุญตะวันตก 2524: 18-19) ได้กล่าวถึงองค์ประกอบที่มีอิทธิพลต่อค่าความยากของแบบสอบไว้ดังนี้

1. องค์ประกอบภายใน (Intrinsic Factors) ได้แก่
  - ก. เนื้อหาของข้อสอบแต่ละข้อ (Item Content) มีความซับซ้อน ความเป็นนามธรรม และความแปลกใหม่
  - ข. ลักษณะโครงสร้างของข้อสอบ (Item Structure) เป็นวิธีการแสดงออกของเนื้อหาที่ข้อสอบนั้นวัด
2. องค์ประกอบภายนอก (Extrinsic Factors) ได้แก่สิ่งต่อไปนี้
  - ก. ความไม่คุ้นเคยต่อเนื้อหา (Unfamiliarity) เป็นการถามในเนื้อหาที่ผู้สอบไม่เคยพบมาก่อน

ข. ความสัมพันธ์กับข้อสอบ (Item Context) เป็นความสัมพันธ์ของข้อสอบข้อนั้น ๆ กับข้อสอบข้ออื่นที่อยู่ในแบบสอบเดียวกัน

ค. ตัวแปรทางบุคลิกภาพ ได้แก่ สภาพร่างกาย ลักษณะนิสัย และความตั้งใจของผู้สอบ

นอกจากนี้ธรรมชาติของเนื้อหา พฤติกรรมที่ต่องการวัด และสิ่งที่ทำให้เกิดความซับซ้อนอื่น ๆ เช่น ภาษาที่ใช้ในข้อสอบ รูปแบบของคำถาม และคำชี้แจงต่าง ๆ ยังเป็นองค์ประกอบที่อาจมีอิทธิพลต่อความยากของแบบสอบได้ด้วย

### อำนาจจำแนก

อำนาจจำแนก คือ อำนาจที่ข้อสอบแต่ละข้อสามารถจำแนกนักเรียนได้ตามความสามารถ ซึ่งจะทำได้หากสามารถแบ่งนักเรียนออกเป็นกลุ่มที่มีความสามารถต่าง ๆ ได้ เช่น กลุ่มเก่ง กลุ่มอ่อน เป็นต้น (ประคอง กรรณสูต 2525: 24) ในการพิจารณาว่านักเรียนคนใดเก่งหรืออ่อน ถ้าเป็นแบบสอบผลสัมฤทธิ์ เขามักจะใช้คะแนนรวมของแบบสอบทั้งหมดเป็นเกณฑ์ โดยนักเรียนที่คะแนนรวมสูงถือว่าเป็นพวกเก่ง ส่วนนักเรียนที่คะแนนรวมต่ำถือว่าเป็นพวกอ่อน

การพิจารณาว่าข้อสอบข้อใดมีค่าอำนาจจำแนกดีจะเป็นข้อที่กลุ่มเก่งทำถูกและกลุ่มอ่อนทำผิดเป็นส่วนใหญ่ วิธีการหาค่าอำนาจจำแนกจะหาความสัมพันธ์ระหว่างคะแนนของข้อสอบข้อนั้น ๆ กับคะแนนรวมของทุกข้อ การคำนวณค่าอำนาจจำแนกจึงใช้การหาสัมประสิทธิ์สหสัมพันธ์แบบคอง ๆ ทั้งนี้ค่าอำนาจจำแนกจึงมีค่าได้ตั้งแต่ 1.00 ถึง -1.00 โดยข้อสอบที่มีค่าอำนาจจำแนกเป็นบวกจะจำแนกความสามารถของนักเรียนได้นั้นคือกลุ่มเก่งจะตอบถูกมากกว่ากลุ่มอ่อน ยิ่งค่าอำนาจจำแนกเข้าใกล้ 1 มาก จะสามารถจำแนกได้มากขึ้น ข้อสอบที่มีค่าอำนาจจำแนกเป็นศูนย์จะจำแนกความสามารถของนักเรียนไม่ได้ นั่นคือนักเรียนกลุ่มเก่งและกลุ่มอ่อนจะตอบได้เท่ากัน ส่วนข้อสอบที่มีค่าอำนาจจำแนกเป็นอนันต์นักเรียนกลุ่มอ่อนจะตอบถูกมากกว่ากลุ่มเก่ง สำหรับการประเมินคุณภาพของข้อสอบโดยอาศัยค่าอำนาจจำแนกนั้นใช้เกณฑ์ในการพิจารณาดังนี้ (Ebel 1965: 364)

ค่าอำนาจจำแนก	การประเมินคุณภาพ
.40 ขึ้นไป	ดีมาก
.30 ถึง .39	ดี
.20 ถึง .29	พอใช้ควรมีการปรับปรุงตัวเลือกบางตัว
.19 ลงไป	ไม่ดี ควรตัดทิ้ง หรือแก้ไขใหม่

ในการประเมินคุณภาพของข้อสอบ ควรพิจารณาว่าค่าอำนาจจำแนกควบคู่ไปกับค่าความยากด้วย เนื่องจากข้อสอบที่มีค่าความยากเท่า ๆ กัน อาจมีความแตกต่างกันในด้านค่าอำนาจจำแนกด้วย ทั้งนี้ถ้าวมาแลวว่าข้อสอบที่มีค่าความยาก 0.50 อาจมีค่าอำนาจจำแนกสูงที่สุด หรือไม่มีเลยก็ได้ ทั้งนี้ข้อสอบที่มีค่าความยาก .50 มีโอกาสที่จะทำหาค่าความเที่ยงของแบบสอบต่ำได้ เมื่อข้อสอบนั้นมีค่าอำนาจจำแนกเป็นศูนย์ แต่ข้อสอบที่มีค่าอำนาจจำแนกสูงจะทำให้ค่าความเที่ยงของแบบสอบสูงด้วยอย่างแน่นอน เนื่องจากค่าอำนาจจำแนกเกี่ยวข้องกับส่วนเบี่ยงเบนมาตรฐานของคะแนนซึ่งเป็นองค์ประกอบสำคัญที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบ (Ebel 1965: 302)

กรอนลินด์ (Gronlund 1976: 270) ได้กล่าวถึงค่าอำนาจจำแนกของแบบสอบวัดผลสัมฤทธิ์ไว้ดังนี้คือ ข้อสอบที่มีค่าอำนาจจำแนกต่ำไม่ใช่ว่าจะ เป็นข้อสอบที่ไม่ดีเสมอไป ทั้งนี้เพราะแบบสอบวัดผลสัมฤทธิ์จะมีการวัดผลการเรียนรู้ในระดั้มความรู้-ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมิน ถ้าการวัดในระดั้มความรู้ใดมีจำนวนข้อสอบน้อยจะทำให้ข้อสอบที่วัดในระดั้มนั้นมีแนวโน้มที่จะมีค่าอำนาจจำแนกต่ำกว่าข้อสอบที่ใช้วัดในระดั้มความรู้ที่มีจำนวนข้อสอบมากกว่า ตัวอย่างเช่นแบบสอบประกอบควยข้อสอบที่วัดในระดั้มความรู้-ความจำ 40 ข้อ วัดในระดั้มความเข้าใจ 20 ข้อ จะทำให้ค่าต่ำกว่าข้อสอบที่วัดในระดั้มความเข้าใจมีค่าอำนาจจำแนกต่ำกว่าข้อสอบที่วัดในระดั้มความรู้-ความจำ เพราะข้อสอบที่วัดในระดั้มความเข้าใจเป็นตัวแทนของคะแนนจากแบบสอบทั้งหมดมีค่าน้อยกว่า และการวัดในระดั้มความเข้าใจมีความสัมพันธ์กันน้อยกับการวัดในระดั้มความรู้-ความจำ การที่ข้อสอบมีค่าอำนาจจำแนกต่ำเป็นการแสดงให้เห็นว่าข้อสอบนั้นวัดในสิ่งที่ต่างไปจากข้อสอบส่วนใหญ่ในแบบสอบ การที่ตัดข้อสอบที่มีค่าอำนาจจำแนกต่ำออกไปจากแบบสอบจะทำให้ข้อสอบในแบบสอบนั้นวัดในระดั้มความรู้-ความจำมากขึ้น นอกจากนั้นยังทำให้ความตรงของแบบสอบที่วัดในระดั้มความเข้าใจลดลง โดยทั่วไปในแบบสอบวัดผลสัมฤทธิ์



นั้นจะมีจำนวนข้อสอบที่ใช้วัดผลการเรียนรู้ในความรู้ระดับสูงน้อยกว่าข้อสอบที่ใช้วัดผลการเรียนรู้ในความรู้ระดับต่ำ จึงทำให้ข้อสอบที่ใช้วัดผลการเรียนรู้ในความรู้ระดับสูงมีค่าอำนาจจำแนกต่ำกว่าข้อสอบที่ใช้วัดผลการเรียนรู้ในระดับความรู้ที่ต่ำกว่า

### องค์ประกอบที่มีอิทธิพลต่อค่าอำนาจจำแนกของแบบสอบ

องค์ประกอบสำคัญที่มีอิทธิพลต่อค่าอำนาจจำแนกของแบบสอบมีดังนี้

1. คุณภาพของตัวเลือกและจำนวนตัวเลือก ข้อสอบที่สร้างตัวเลือกทุกตัวไว้ที่มีความใกล้เคียงกับตัวเลือกที่เป็นคำตอบที่ถูกจะทำให้ค่าอำนาจจำแนกของข้อสอบสูงขึ้น (Ebel 1965: 359) เนื่องจากข้อสอบแบบนี้ นักเรียนที่มีความรู้จริงเท่านั้นที่สามารถตอบข้อสอบได้ถูกต้อง แต่ นักเรียนที่ไม่มีความรู้จริงจะเลือกตัวเลือกที่ใกล้เคียงกับคำตอบที่ถูกกระจายออกไปทุกตัวเลือก และแบบสอบเลือกตอบที่มีจำนวนตัวเลือกมากจะทำให้มีความสามารถในการตรวจมากขึ้นซึ่งแบบสอบมาตรฐานทั่วไปจะมี 4-5 ตัวเลือก (ชาวอแพริทกุล 2516: 178)

2. ความยากของข้อสอบ การจำแนกผลสัมฤทธิ์ระหว่างกลุ่มนักเรียนขึ้นอยู่กับความยากของข้อสอบ กล่าวคือ ข้อสอบที่มีค่าความยาก .50 มีโอกาสทำให้ค่าอำนาจจำแนกสูงสุดได้ เพราะทำให้ให้นักเรียนกลุ่มสูงทุกคนตอบได้ถูกต้อง ในขณะที่นักเรียนในกลุ่มต่ำทั้งหมดตอบผิด ซึ่งหากข้อสอบมีค่าความยากสูงหรือต่ำกว่านี้ ค่าอำนาจจำแนกของข้อสอบจะต่ำลง (Gronlund 1976: 270) ซึ่งหมายความว่าข้อสอบที่ยากหรือง่ายเกินไป จะเป็นข้อสอบที่มีค่าอำนาจจำแนกต่ำลงนั่นเอง

### ตอนที่ 2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการวิจัยเรื่องนี้จะเสนอการวิจัยเกี่ยวกับการเปรียบเทียบคุณภาพของแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบขรรคมคาและแบบสอบเลือกตอบที่ใช้ตัวเลือกซ้อน และการวิจัยเกี่ยวกับการใช้แบบสอบเลือกตอบที่มีตัวเลือกต่างกันวัดระดับความรู้ชั้นต่าง ๆ ตามการจำแนกลำดับชั้นการเรียนรู้ของบลูม (Bloom) ตามลำดับ

งานวิจัยที่เกี่ยวข้องกับการเปรียบเทียบคุณภาพแบบสอบเลือกตอบตัวเลือกขรรคมคาและตัวเลือกซ้อนเท่าที่ค้นพบมีเรื่อง เกี่ยวคือ



เวเทิน (Weiten 1982: 46-50) ได้ศึกษาประสิทธิภาพสัมพัทธ์ของแบบสอบเลือกตอบทั่วไปเลือกขรรรรมคา และแบบสอบเลือกตอบทั่วไปเลือกชอน กลุ่มตัวอย่างเป็นนักศึกษาปริญญาตรีที่ลงทะเบียนเรียนวิชาจิตวิทยาทั่วไป จำนวน 47 คน เครื่องมือที่ใช้เป็นแบบสอบเลือกตอบทั่วไปเลือกขรรรรมคาและแบบสอบเลือกตอบทั่วไปเลือกชอน โดยชอกระหงแบบเลือกตอบทั่วไปเลือกชอนคัดแปลงมาจากชอกระหงแบบเลือกตอบทั่วไปเลือกขรรรรมคา มีจำนวนชอกระหง 40 ชอ ในแต่ละแบบสอบ ในการทดสอบใช้แบบสอบ 2 พอร์ม ซึ่งมีการจัดเรียงชอกระหงดังนี้

พอร์ม A แบบสอบย่อยแบบเลือกตอบทั่วไปเลือกขรรรรมคา 20 ชอ และแบบสอบย่อยแบบเลือกตอบทั่วไปเลือกชอน 20 ชอ

พอร์ม B แบบสอบย่อยเลือกตอบทั่วไปเลือกชอน 20 ชอ และแบบสอบย่อยแบบเลือกตอบทั่วไปเลือกขรรรรมคา 20 ชอ

ผลการวิจัยปรากฏว่า แบบสอบเลือกตอบทั่วไปเลือกชอนยากกว่าแบบสอบเลือกตอบทั่วไปเลือกขรรรรมคา ค่าอำนาจจำแนกของแบบสอบเลือกตอบทั่วไปเลือกขรรรรมคาสูงกว่าแบบสอบเลือกตอบทั่วไปเลือกชอน การเปรียบเทียบค่าความเที่ยงได้เปรียบเทียบ 2 คู่ เนื่องจากมี 2 แบบสอบย่อย โดยค่าความเที่ยงของแบบสอบย่อยแบบเลือกตอบทั่วไปเลือกขรรรรมคามีค่าความเที่ยงสูงกว่าแบบสอบย่อยแบบเลือกตอบทั่วไปเลือกชอนในการเปรียบเทียบครั้งแรก ส่วนการเปรียบเทียบอีก 1 คู่ แบบสอบย่อยแบบเลือกตอบที่มีตัวเลือกทั้ง 2 แบบ มีค่าความเที่ยงไม่แตกต่างกัน สำหรับการเปรียบเทียบค่าความตรงได้เปรียบเทียบ 2 คู่ เช่นเดียวกัน โดยค่าความตรงของแบบสอบเลือกตอบที่มีตัวเลือกทั้ง 2 แบบ มีค่าความตรงไม่แตกต่างกันทั้ง 2 คู่ นั่นคือ ค่าความตรงของแบบสอบเลือกตอบทั่วไปเลือกขรรรรมคาและตัวเลือกชอนไม่แตกต่างกัน

จากผลการวิจัยที่กล่าวมาแล้วจะเห็นว่าแบบสอบเลือกตอบทั่วไปเลือกขรรรรมคายากกว่าแบบสอบเลือกตอบทั่วไปเลือกชอน มีค่าอำนาจจำแนกสูงกว่าแบบสอบเลือกตอบทั่วไปเลือกชอน และมีแนวโน้มว่าจะมีค่าความเที่ยงสูงกว่าแบบสอบเลือกตอบทั่วไปเลือกชอน ส่วนค่าความตรงของแบบสอบเลือกตอบทั่วไปเลือกขรรรรมคา และตัวเลือกชอนไม่แตกต่างกัน

เนื่องจากลักษณะชอกระหงแบบเลือกตอบทั่วไปเลือกชอนมีลักษณะของตัวเลือกบางตัวเป็นตัวเลือกผสม ได้แก่ "ถูกทั้ง ข้อ ก และ ข้อ ข" หรือ "ถูกทั้ง ข้อ ก และ ข้อ ค"

เช่นกัน ผู้วิจัยจึงเสนองานวิจัยเกี่ยวกับแบบสอบถามเลือกตอบที่ใช้ตัวเลือกผสมดังนี้

ฮิวส์ และ ทริมเบิล (Hughes and Trimble 1965: 117-125)

ได้ศึกษาถึงคุณภาพของแบบสอบถามเลือกตอบที่มีตัวเลือกซับซ้อนแตกต่างกันในวิชาจิตวิทยาทั่วไป กลุ่มตัวอย่าง เป็นนักศึกษาระดับปริญญาตรีชั้นปีที่ 1 จำนวน 63 คน แบ่งออกเป็น 3 กลุ่ม กลุ่มที่ 1 ให้ทำแบบสอบถามเลือกตอบที่ใช้ตัวเลือกธรรมดา กลุ่มที่ 2 ทำแบบสอบถามที่ใช้ตัวเลือกแบบปลายเปิดและแบบปลายปิด กลุ่มที่ 3 ทำแบบสอบถามเลือกตอบที่ใช้ตัวเลือกผสม ซึ่งแบบสอบถามแต่ละฉบับประกอบด้วยข้อสอบจำนวน 55 ข้อ แบ่งออกเป็น 3 ส่วน ส่วนแรกเป็นข้อสอบความคลุมเครือ เป็นข้อสอบที่ใช้ตัวเลือกธรรมดาเหมือนกันทั้ง 3 ฉบับ ส่วนที่สองเป็นข้อสอบที่ใช้เปรียบเทียบซึ่งเป็นข้อสอบที่ใช้ตัวเลือกสุดท้ายเป็นตัวเลือกแบบต่าง ๆ ส่วนที่สามเป็นข้อสอบที่ใช้ตัวเลือกแบบต่าง ๆ เป็นคำตอบที่ถูกต้อง โดยมีการจัดเรียงข้อสอบในแบบสอบถามแต่ละฉบับให้ข้อสอบทุก ๆ 10 ข้อ ใน 50 ข้อแรก ประกอบด้วยข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกธรรมดา 3 ข้อ ข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกสุดท้ายเป็นตัวเลือกแบบต่าง ๆ 6 ข้อ และข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกแบบต่าง ๆ เป็นคำตอบที่ถูกต้อง 1 ข้อ สำหรับข้อสอบ 5 ข้อสุดท้ายเป็นข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกสุดท้ายเป็นตัวเลือกแบบต่าง ๆ ซึ่งจะทำให้จำนวนข้อสอบในส่วนที่หนึ่ง ซึ่งเป็นข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดามี 15 ข้อ ข้อสอบในส่วนที่สอง ซึ่งเป็นข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกสุดท้ายเป็นตัวเลือกแบบต่าง ๆ 35 ข้อ และข้อสอบในส่วนที่สามซึ่งเป็นข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกแบบต่าง ๆ เป็นคำตอบที่ถูกต้อง 5 ข้อ ผลการวิจัยปรากฏว่า แบบสอบถามเลือกตอบที่ใช้ตัวเลือกซับซ้อนทั้ง 2 ฉบับ ยากกว่าแบบสอบถามที่ใช้ตัวเลือกแบบธรรมดาอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าอำนาจจำแนกของแบบสอบถามทั้ง 3 ฉบับไม่แตกต่างกันเมื่อเปรียบเทียบคะแนนเฉลี่ยของนักเรียนทั้ง 3 กลุ่ม ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยคะแนนผลการสอบของนักเรียนกลุ่มที่ทำแบบสอบถามเลือกตอบตัวเลือกธรรมดาสูงกว่าคะแนนของนักเรียนกลุ่มที่ทำแบบสอบถามที่ใช้ตัวเลือกซับซ้อนทั้งสองกลุ่ม

ทอมป์สัน (Tompson 1975: 135-141) ได้ศึกษาในทำนอง

เกี่ยวกับฮิวส์และทริมเบิล โดยศึกษากับนักศึกษาในมหาวิทยาลัยอินเดียนา จำนวน 2000 คน ในวิชาการประกอบอาชีพเป็นพนักงานขาย (Salesman) ผลการวิจัยปรากฏว่า แบบของตัวเลือกมีผลต่อค่าความยากของแบบสอบถามโดยแบบสอบถามที่ใช้ตัวเลือกแบบผสมยากที่สุด รองลงมา

เป็นแบบสอบที่ใช้ตัวเลือกแบบปลายเปิดและแบบปลายปิด แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดาที่ง่ายที่สุด สำหรับอำนาจจำแนกของแบบสอบที่ใช้ตัวเลือกแตกต่างกันทุกชนิดไม่แตกต่างกัน เมื่อพิจารณาประสิทธิภาพของตัวลวง ปรากฏว่านักเรียนเลือกตัวลวงแบบผสมมากที่สุด รองลงมาเป็นตัวลวงแบบปลายเปิด ตัวลวงแบบปลายปิด ส่วนตัวลวงแบบธรรมดา นักเรียนเลือกน้อยที่สุด

นวน้อย แต่บรรพกุล (2520: 55-56) ได้ศึกษาสมรรถนะในการวัดของแบบสอบเลือกตอบที่ใช้ตัวเลือกซับซ้อนแตกต่างกัน ในวิชาภาษาไทย ระดับชั้นประถมศึกษาปีที่ 7 โดยศึกษากับแบบเลือกตอบที่มีตัวเลือกแตกต่างกัน 4 แบบ คือ ตัวเลือกแบบธรรมดาแบบปลายเปิด "ผิดทุกข้อ" แบบปลายปิด "ถูกทุกข้อ" แบบแบบผสมซึ่งได้แก่ "ถูกทั้ง ข้อ ก และข้อ ข" หรือ "ถูกทั้งข้อ ก และข้อ ค" เป็นต้น แบบสอบที่ใช้ตัวเลือกแต่ละแบบยังแบ่งย่อยตามเนื้อหาเป็น 3 ฉบับ คือ แบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน แบบสอบแต่ละฉบับประกอบด้วย 2 ส่วน ส่วนแรกเป็นข้อสอบที่ใช้เปรียบเทียบจริง จำนวน 25 ข้อ ซึ่งเป็นข้อสอบแบบเลือกตอบที่ใช้ตัวเลือกธรรมดาหรือตัวเลือกซับซ้อนแบบใดแบบหนึ่งเป็นตัวลวงสุดท้าย อีกส่วนหนึ่งเป็นข้อสอบควบคุมซึ่งเป็นข้อสอบแบบเลือกตอบตัวเลือกธรรมดา 5 ตัวเลือก จำนวน 10 ข้อ ซึ่งปรับปรุงให้เป็นข้อสอบที่ง่าย ๆ เพื่อไม่ให้มีผลต่อข้อสอบที่ใช้เปรียบเทียบจริง ผลการวิจัยปรากฏว่า

1. ค่าความเที่ยงของแบบสอบที่ใช้ตัวเลือกต่างกันทั้ง 4 ชุด ไม่แตกต่างกันทั้งในแบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน

2. ค่าความตรงของแบบสอบเลือกตอบที่ใช้ตัวเลือกต่างกันทั้ง 4 ชุด เป็นดังนี้

2.1 แบบสอบสะกดคำ แบบสอบเลือกตอบที่ใช้ตัวเลือกแตกต่างกันทั้ง 4 ชุด มีค่าความตรงไม่แตกต่างกัน

2.2 แบบสอบหลักภาษา แบบสอบเลือกตอบที่ใช้ตัวเลือกธรรมดา มีค่าความตรงสูงกว่าแบบสอบเลือกที่ใช้ตัวเลือกแบบปลายปิด และแบบผสมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบปลายเปิดมีค่าความตรงสูงกว่าแบบสอบเลือกตอบที่ใช้ตัวเลือกผสมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนการเปรียบเทียบค่าความตรงของแบบสอบเลือกตอบที่ใช้ตัวเลือกต่างกันคู่อื่น ๆ ไม่แตกต่างกัน

2.3 แบบสอบถามความเข้าใจในการอ่าน แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดาที่มีค่าความตรงสูงกว่าแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบปลายเปิด และแบบผสมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนการเปรียบเทียบค่าความตรงของแบบสอบเลือกตอบที่ใช้ตัวเลือกต่าง ๆ กันอื่น ๆ ไม่แตกต่างกัน

3. ค่าอำนาจจำแนกของแบบสอบเลือกตอบที่ใช้ตัวเลือกต่างกันทั้ง 4 ชุด ไม่แตกต่างกัน ทั้งในแบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน

4. ค่าความยากมาตรฐานของแบบสอบที่ใช้ตัวเลือกต่างกันทั้ง 4 ชุด แตกต่างกันอย่างเฉพาะในแบบสอบถามความเข้าใจในการอ่าน โดยที่แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบผสมมีค่าความยากมาตรฐานสูงสุด สำหรับแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดา แบบปลายเปิด และแบบปลายปิด ไม่แตกต่างกัน

5. การศึกษาเกี่ยวกับการเคา โดยที่แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบผสมมีเปอร์เซ็นต์การเคาน้อยที่สุด ทั้งในแบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน

6. คะแนนผลการสอบของนักเรียนที่ทำแบบสอบที่ใช้ตัวเลือกต่างกันทั้ง 4 ชุด มีความแตกต่างกัน โดยเฉพาะอย่างยิ่ง คะแนนผลการสอบของนักเรียนที่ทำแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบผสมมีค่าต่ำสุด ทั้งในแบบสอบสะกดคำ หลักภาษา และความเข้าใจในการอ่าน

จากผลการศึกษาที่กล่าวมาแล้วจะเห็นได้ว่า แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบผสมยากกว่าแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดาส่วนค่าอำนาจจำแนก และค่าความเที่ยงของแบบสอบเลือกตอบที่ใช้ตัวเลือกทั้ง 2 แบบไม่แตกต่างกัน สำหรับค่าความตรงแบบสอบเลือกตอบตัวเลือกธรรมดาที่มีค่าความตรงสูงกว่าแบบสอบเลือกตอบตัวเลือกผสม

งานวิจัยเกี่ยวกับการเปรียบเทียบคุณภาพของแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดาและแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบต่าง ๆ ในการวัดระดับความรู้ชั้นต่าง ๆ ตามการจกคำกับชั้นการเรียนรูของ Bloom) เท่าที่ค้นพบมีเรื่องเกี่ยวกับ

ทริป (Tripp 1983: 3886-A) ได้ศึกษาเปรียบเทียบค่าความยากและค่าอำนาจจำแนกของแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดา (Multiple Choice)



และแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบซับซ้อน (Complex Multiple Choice) ในการวัดระดับพุทธิสัย (Cognitive Domain) แตกต่างกันในวิชาชีพด้านสุขภาพพื้นฐาน (Basic health profession subject) และได้เปรียบเทียบในเนื้อหาวิชาที่แตกต่างกัน 3 เนื้อหา กลุ่มตัวอย่าง เป็นนักศึกษาพยาบาลชั้นปีที่ 4 ในรัฐมิสซูรี และแคนซัส 15 โรงเรียน จำนวน 515 คน เครื่องมือที่ใช้เป็นแบบสอบวัดผลสัมฤทธิ์ 2 ฉบับ คือ แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดา และตัวเลือกแบบซับซ้อน ฉบับละ 18 ข้อ นักเรียนที่ได้รับเลือกเป็นกลุ่มตัวอย่างจะได้รับแบบสอบคนละฉบับโดยการสุ่มการวิเคราะห์ความแตกต่างของค่าความยาก ใช้การวิเคราะห์ความแปรปรวนรวม (ANCOVA) โดยใช้การแนะนำด้านเภสัชกรรม (Pharmacology instruction) เป็นตัวแปรร่วม (Covariate) ผลการวิจัยปรากฏว่าแบบสอบเลือกตอบที่ใช้ตัวเลือกซับซ้อนยากกว่าแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดาในการวัดระดับความรู้ความจำทั้ง 3 เนื้อหา ในการวัดระดับความรู้ชั้นสูง คือ ระดับการสังเคราะห์ และการประเมินผล ค่าความยากของแบบสอบเลือกตอบที่ใช้ตัวเลือกทั้ง 2 แบบ ไม่แตกต่างกันทั้ง 3 เนื้อหา สำหรับค่าอำนาจจำแนกของแบบสอบเลือกตอบตัวเลือกแบบธรรมดาและตัวเลือกแบบซับซ้อนไม่แตกต่างกัน

จะเห็นว่าในการวัดความรู้ระดับต่ำ อันได้แก่ การวัดในระดับความรู้-ความจำ แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบซับซ้อนยากกว่าแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดา และในการวัดความรู้ระดับสูง อันได้แก่ การวัดในระดับการสังเคราะห์ และการประเมินผล แบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดาและแบบซับซ้อนมีค่าความยากไม่แตกต่างกัน สำหรับค่าอำนาจจำแนกของแบบสอบเลือกตอบที่ใช้ตัวเลือกทั้ง 2 แบบ ไม่แตกต่างกันทั้งในการวัดความรู้ระดับต่ำ และระดับสูง

การศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบธรรมดาและแบบสอบเลือกตอบที่ใช้ตัวเลือกแบบต่าง ๆ ในการวัดระดับความรู้ชั้นต่าง ๆ ตามการจำแนกลำดับชั้นการเรียนรู้ของบลูม (Bloom) ในประเทศไทยยังมีได้มีผู้ใดศึกษาไว้ แต่มีการศึกษาเกี่ยวกับผลการใช้แบบสอบเลือกตอบ และแบบสอบอักษณัยคอมสัน์วัดระดับความรู้ชั้นต่าง ๆ ตามการจำแนกลำดับชั้นการเรียนรู้ของบลูม (Bloom) ทั้งนี้



รุจิร ภูสาระ (2514: 115-118) ได้ศึกษาผลการใช้แบบทดสอบแบบ  
เลือกตอบและแบบอัตนัยตอบสั้นวัดระดับความรู้ชั้นต่าง ๆ ในวิชาภาษาไทย ชั้นประถมศึกษา  
ปีที่ 7 กลุ่มตัวอย่างจำนวน 465 คน เครื่องมือที่ใช้เป็นแบบสอบ 3 ชนิด คือ แบบ  
เลือกตอบ แบบอัตนัยตอบสั้น และแบบผสมระหว่างแบบเลือกตอบและแบบอัตนัยตอบสั้น  
วัดระดับความรู้ชั้นต่าง ๆ ตามการจัดลำดับชั้นการเรียนรู้ของบลูม (Bloom) จำนวน  
120 ข้อ ซึ่งในแบบสอบนี้วัดระดับความรู้-ความจำ 43 ข้อ ความเข้าใจ 28 ข้อ  
การนำไปใช้ 9 ข้อ การวิเคราะห์ 27 ข้อ การสังเคราะห์ 4 ข้อ และการประเมินผล  
9 ข้อ ผลการวิจัยเป็นดังนี้

1. คะแนนเฉลี่ยจากการตอบแบบสอบ 3 ชนิดแตกต่างกันในการวัด  
ความรู้ทุกระดับ โดยคะแนนเฉลี่ยจากแบบสอบแบบเลือกตอบมีค่าสูงสุด รองลงมาเป็น  
แบบผสม และแบบอัตนัยตอบสั้น ตามลำดับ
2. ค่าความยากของแบบสอบ การวัดในระดับความรู้-ความจำแบบอัตนัย  
ตอบสั้นให้ค่าความยากสูงสุด ส่วนแบบเลือกตอบและแบบผสมมีค่าพอกัน การวัดในระดับความ  
เข้าใจ การนำไปใช้ และการประเมินผล แบบสอบทั้ง 3 ชนิดมีค่าความยากไม่แตกต่างกัน  
สำหรับการวัดในระดับการวิเคราะห์ และการสังเคราะห์แบบอัตนัยตอบสั้นยากที่สุด แบบ  
เลือกตอบและแบบผสมยากพอกัน
3. ค่าอำนาจจำแนกของแบบสอบที่ได้จากการวัดความรู้ทุกระดับ  
แบบสอบทั้ง 3 ชนิดให้ค่าอำนาจจำแนกไม่แตกต่างกัน
4. ค่าความเที่ยงของแบบสอบที่ใช้วัดระดับความรู้ความจำ แบบสอบทั้ง  
3 ชนิดมีค่าความเที่ยงไม่แตกต่างกัน สำหรับการวัดในระดับความเข้าใจ การวิเคราะห์  
และการประเมินผล แบบเลือกตอบให้ค่าความเที่ยงสูงสุด รองลงมาเป็นแบบอัตนัยตอบสั้น  
และแบบผสม ตามลำดับ สำหรับการวัดในระดับการนำไปใช้ และการสังเคราะห์ แบบอัตนัย  
ตอบสั้นให้ค่าความเที่ยงสูงสุด รองลงมาเป็นแบบเลือกตอบ และแบบผสม จะเห็นได้ว่า  
แบบผสมส่วนมากจะให้ความเที่ยงต่ำกว่าแบบอื่น ๆ
5. ค่าความตรงของแบบสอบที่ใช้วัดระดับความรู้ ความจำ ความเข้าใจ  
การนำไปใช้ และการประเมินผล แบบสอบทั้ง 3 ชนิดมีค่าความตรงไม่แตกต่างกัน แต่การ  
วัดในระดับการวิเคราะห์ แบบผสมให้ค่าความตรงสูงสุด รองลงมาเป็นแบบเลือกตอบและ  
แบบอัตนัยตอบสั้น ตามลำดับ สำหรับการวัดในระดับการสังเคราะห์นั้น แบบอัตนัยตอบสั้นให้

ค่าความตรงสูงสุด แบบเลือกตอบ และแบบผสมมีค่าความตรงพอกัน

6. นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่างกันจะตอบแบบสอบทั้ง 3 ชนิด ไค้คะแนนเฉลี่ยแตกต่างกัน นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูงจะตอบแบบสอบเลือกตอบ และแบบผสมไค้คะแนนเฉลี่ยไม่แตกต่างกัน แต่จะสูงกว่าแบบอักษัยตอบสั้น นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่ำจะตอบแบบสอบเลือกตอบไค้ที่่ที่สุด รองลงมาคือ แบบผสม และแบบอักษัยตอบสั้น ตามลำดับ

7. นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง มีสมรรถภาพสมองทุกด้าน สูงกว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่ำ

ในปีเดียวกันนั้น นภาพร อมรเลิศสินไทย (2514: 158-162) ไค้ศึกษา ในทำนองเดียวกับ รุจิรี ภูสาระ กล่าวคือ ไค้ศึกษาผลการใช้แบบทดสอบแบบเลือกตอบและแบบอักษัยตอบสั้นวัดระดับความรู้ชั้นต่าง ๆ ในวิชาสังคมศึกษา ชั้นประถมศึกษาปีที่ 7 กลุ่มตัวอย่างจำนวน 558 คน เครื่องมือเป็นแบบสอบ 3 ชนิด คือ แบบเลือกตอบ แบบอักษัยตอบสั้น และแบบผสมระหว่างแบบเลือกตอบและแบบอักษัยตอบสั้น วัดระดับความรู้ชั้นต่าง ๆ ตามการจัดลำดับชั้นการเรียนรู้ของบลูม (Bloom) จำนวน 120 ข้อ ซึ่งในแบบสอบนี้วัดระดับความรู้-ความจำ 39 ข้อ ความเข้าใจ 19 ข้อ การนำไปใช้ 22 ข้อ การวิเคราะห์ 23 ข้อ การสังเคราะห์ 8 ข้อ และการประเมินผล 9 ข้อ การวิจัยปรากฏผลดังนี้

1. คะแนนเฉลี่ยจากการตอบแบบสอบทั้ง 3 ชนิด แตกต่างกันในการวัดความรู้ทุกระดับ โดยคะแนนเฉลี่ยจากแบบสอบแบบเลือกตอบมีค่าสูงสุด รองลงมาเป็นแบบผสม และแบบอักษัยตอบสั้น ตามลำดับ

2. ค่าความยากของแบบสอบ ในการวัดระดับความรู้ ความจำ แบบอักษัยตอบสั้น และแบบผสมยากพอกัน แต่ทั้ง 2 แบบยากกว่าแบบเลือกตอบ ในการวัดระดับความรู้ เข้าใจและการนำไปใช้ การวิเคราะห์ และการประเมินผลแบบอักษัยตอบสั้นยากที่สุด แบบเลือกตอบและแบบผสมยากพอกัน สำหรับการวัดในระดับการสังเคราะห์ แบบสอบทั้ง 3 ชนิดยากไม่แตกต่างกัน

3. ค่าอำนาจจำแนกของแบบสอบ ในการวัดความรู้ทุกระดับ แบบสอบทั้ง 3 ชนิดมีค่าอำนาจจำแนกไม่แตกต่างกัน

4. ค่าความเที่ยงของแบบสอบ ในการวัดระดับความรู้-ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ และการสังเคราะห์ แบบสอบทั้ง 3 ชนิดมีค่าความเที่ยงไม่แตกต่างกัน สำหรับการวัดในระ�คัมการประเมินผลแบบอัตนัยทอบสนั้ให้ค่าความเที่ยงสูงสุด ส่วนแบบเลือกทอบและแบบผสมมีค่าความเที่ยงทอบกัน

5. ค่าความตรงของแบบสอบ ในการวัดความรู้ทุกระ�คัม แบบสอบทั้ง 3 ชนิดมีค่าความตรงไม่แตกต่างกัน

6. นักเรียนกลุ่มที่มีผลสัมฤทธิ์ทางการเรียนสูงและต่ำจะทอบแบบสอบทั้ง 3 ชนิดไคคะแนนเฉลี่ยแตกต่างกัน โดยทอบแบบเลือกทอบไคคะแนนเฉลี่ยสูงสุด รองลงมาเป็นแบบผสม และแบบอัตนัยทอบสนั้ไคคะแนนเฉลี่ยต่ำสุด โดยที่นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูงไคคะแนนเฉลี่ยสูงกว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่ำ

จากผลการศึกษาที่ไคกล่าวมาแล้วจะเห็นไคว่าเมื่อใช้แบบสอบทางชนิดกัน วัดระดับความรู้เกี่ยวกัน จะทำให้ค่าความยาก ค่าความเที่ยง และค่าความตรงของแบบสอบแตกต่างกัน แต่ไม่มีผลทอบคอบอำนาจจำแนกของแบบสอบ

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย