



วรรณคดีที่เกี่ยวข้อง

ผู้วิจัยได้แบ่งวรรณคดีที่เกี่ยวข้องออกเป็น 2 ตอน คือ ตอนที่ 1 ว่าด้วยมโนทัศน์เบื้องต้นเกี่ยวกับความยากของข้อสอบ ความเที่ยงและความตรงของแบบสอบ ส่วนตอนที่ 2 ว่าด้วยงานวิจัยที่เกี่ยวข้องกับลักษณะ โครงสร้างข้อกระทงและการจัดตำแหน่งตัวถูก

ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับความยากของข้อสอบ ความเที่ยงและความตรงของแบบสอบ ความยากของข้อสอบหรือข้อกระทง

ในการวิจัยนี้ ผู้วิจัยมุ่งศึกษาระแนเฉลี่ยจากแบบสอบซึ่งใช้แทนความยากของแบบสอบทั้งฉบับเป็นสำคัญ แต่เนื่องจากความยากของแบบสอบทั้งฉบับขึ้นอยู่กับความยากของข้อกระทงแต่ละข้อ ดังเช่น อีเบล (Ebel) กล่าวว่า ระแนเฉลี่ยของแบบสอบยอมได้จากความยากเฉลี่ยของข้อสอบทั้งหมดที่เป็นส่วนประกอบในแบบสอบนั้น¹ ดังนั้นจะกล่าวถึงความยากของข้อกระทงพลั้ง เชปคังค่อไปนี้

ความยากของข้อสอบแต่ละข้อขึ้นอยู่กับสัดส่วนหรือเปอร์เซ็นต์ของผู้ตอบข้อสอบนั้นถูก² ยิ่งเปอร์เซ็นต์ของผู้ตอบถูกมีค่าสูง ข้อสอบนั้นก็ยิ่งง่าย อย่างไรก็ตาม คำนีความยากของข้อสอบไม่เพียงแต่บ่งชี้ถึงคุณลักษณะของข้อสอบแต่เพียงอย่างเดียว ยังสะท้อนให้เห็นความสามารถของกลุ่มผู้รับการทดสอบด้วย³ ดังนั้น การกล่าวถึงความยากของข้อสอบ ควรระบุ

¹Robert L. Ebel, Measuring Educational Achievement, p.300.

²แอน อนุสาคาชี, การตรวจสอบเชิงจิตวิทยา, หน้า 161.

³Robert L. Ebel, Measuring Educational Achievement, p.360.

ถึงกลุ่มเฉพาะนั้น ๆ ด้วย

การกำหนดความยากของข้อสอบในระหว่างการเขียนข้อสอบนั้น ผู้เขียนควรคำนึงถึง 1) ความมุ่งหมายของการทดสอบ 2) ระดับความสามารถของผู้รับการทดสอบ 3) อายุหรือระดับการศึกษาของผู้รับการทดสอบ¹ อย่างไรก็ตามข้อสอบที่นำไปใช้ทดสอบแล้วเท่านั้นจึงจะวัดความยากให้ประจักษ์ได้ อนาสตาซี ได้ให้เหตุผลสำคัญในการวัดความยากของข้อสอบว่า เพื่อจะคัดข้อสอบที่มีระดับความยากไม่เหมาะสมออก ข้อสอบที่ไม่มีใครทำถูกกับข้อสอบที่ทุก ๆ คนทำได้หมดจะเป็นน้ำหนักถ่วงในแบบสอบนั้น เพราะข้อสอบทั้งสองประเภทนี้ไม่ได้ให้ความรู้อะไร เกี่ยวกับความแตกต่างระหว่างบุคคล และเนื่องจากข้อสอบแบบนี้ไม่ได้มีผลต่อความแปรปรวนของคะแนนแบบสอบ จึงไม่ได้ให้อะไรในเรื่องความตรงและความเที่ยงของแบบสอบ ในการวัดแบบอิงกลุ่มนั้น เพื่อให้ความจำแนกระหว่างบุคคลมีค่าสูงสุด จึงเป็นการดีที่สุดที่จะเลือกข้อสอบที่มีระดับความยากประมาณ .50 การเลือกเช่นนี้จะทำให้ได้คะแนนรวมที่มีการกระจายกว้างขวางที่สุด และค่าสัมประสิทธิ์ความเที่ยงจะมีค่าสูงสุดในกรณีนี้ด้วย² จะเห็นได้ว่า ความยากง่ายของข้อสอบย่อมส่งผลถึงความเที่ยงและความตรงของแบบสอบด้วย

องค์ประกอบที่มีอิทธิพลต่อความยากของข้อสอบ

แคมเบล (Campbell) ได้แบ่งองค์ประกอบที่มีอิทธิพลต่อความยากของข้อสอบออกเป็น 2 ประเภทใหญ่ ๆ คือ

1. องค์ประกอบภายใน (Intrinsic Factors) ได้แก่

¹William A. Mehrens, and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology 2d ed. (New York : Holt, Rinehart and Winston, 1978), p.191.

²แอน อนาสตาซี, การตรวจสอบเชิงจิตวิทยา, หน้า 161 - 162.

ก. เนื้อหาของข้อกระทง (Item Content) ในด้านความซับซ้อน
ความเป็นนามธรรมและความแปลกใหม่

ข. ลักษณะโครงสร้างข้อกระทง (Item Structure) หมายถึง วิธีการ
ในการแสดงออกซึ่งเนื้อหานั้น

2. องค์ประกอบภายนอก (Extrinsic Factors) ได้แก่

ก. ความไม่คุ้นเคยต่อเนื้อหา (Unfamiliarity) คืออยู่นอกเหนือ
ประสบการณ์ของผู้ตอบ

ข. สิ่งที่มีสัมพันธ์กับข้อกระทง (Item Context) เช่น ข้อกระทงที่อยู่
ใกล้เคียงกัน

ค. ตัวแปรส่วนบุคคลสภาพ เช่น สภาพร่างกาย ลักษณะนิสัย และความ
ตั้งใจของผู้ตอบ เป็นต้น¹

วอน (Vaughn) มีความเห็นที่คล้ายคลึงกัน เขากล่าวว่า องค์ประกอบสำคัญที่มี
ผลต่อความยากของข้อสอบก็คือ ธรรมชาติของเนื้อหาและชนิดของพฤติกรรมที่ทดลองการจาก
ผู้ตอบ นอกจากนั้นยังมีตัวการอื่น ๆ ที่สลับซับซ้อนและอาจไม่เด่นชัด เช่น ลักษณะของภาษา
ที่ใช้ รูปแบบคำถาม ตลอดจนคำชี้แจง สิ่งเหล่านี้ควรต้องได้รับความสนใจจากผู้สร้างแบบสอบ
ด้วย²

ศูนย์วิจัยทรัพยากร

จุฬาลงกรณ์มหาวิทยาลัย

¹ Alison C. Campbell, "Some Determinants of the Difficulty of Non-Verbal Classification Item," pp.899-913.

² K.W.Vaughn, "Planning The Objective Test," in Educational Measurement, p.174. Edited by E.F. Lindquist. (Washington D.C. : American Council on Education, 1961).

ความเที่ยงของแบบสอบ

นิยามของความเที่ยง

แอน ฮนาสคาซี ได้ให้นิยามว่า "ความเที่ยง หมายถึง ความคงที่ของคะแนนที่ได้จากการทดสอบคนกลุ่มเดียวกันสองครั้ง ด้วยแบบสอบ เดิมในเวลาที่ต่างกันหรือทดสอบคนกลุ่มเดียวกันด้วยข้อสอบต่างชุดที่มีข้อสอบเทียบเท่ากัน หรือภายใต้สภาพการสอบที่ต่างกัน"¹

อีเบล (Ebel) กล่าวว่า "สัมประสิทธิ์ความเที่ยงของคะแนนชุดหนึ่งของคนกลุ่มหนึ่งก็คือสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนชุดนั้นกับคะแนนอีกชุดหนึ่ง จากแบบสอบที่เทียบเท่ากับแบบสอบเดิม และได้มาอย่างอิสระ จากคนกลุ่มเดิม"²

เมเรนส์ และ เลแมน (Mehrens and Lehmann) ได้ให้คำจำกัดความสั้น ๆ ว่า "ความเที่ยง หมายถึง ดีกรี (Degree) ของความคงที่ระหว่างการวัดสิ่งเดียวกัน 2 ครั้ง"³

ทฤษฎีของความเที่ยง

การอธิบายทฤษฎีของความเที่ยงนั้น เริ่มต้นจากคะแนนที่วัดได้ (Observed Scores) ซึ่งคะแนนที่วัดได้นี้ประกอบด้วยคะแนนจริง (True Scores) กับคะแนนความคลาดเคลื่อน (Error Scores) จึงเขียนเป็นสมการ⁴ ได้ดังนี้

¹แอน ฮนาสคาซี, การตรวจสอบเชิงจิตวิทยา, หน้า 78.

²Robert L. Ebel, Measuring Educational Achievement, p.311.

³William A. Mehrens, and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology, p.88.

⁴Ibid., p.90.

$$X = T + E$$

เมื่อ X = คะแนนที่วัดได้

T = คะแนนจริง

E = คะแนนความคลาดเคลื่อน

คะแนนจริง หมายถึง คะแนนที่บุคคลพึงจะได้รับถ้าเครื่องมือที่ใช้วัดนั้นมีคุณลักษณะที่สมบูรณ์ปราศจากความคลาดเคลื่อน หรือ หมายถึง คะแนนเฉลี่ยของบุคคลซึ่งได้จากการทำแบบสอบเดิมหลาย ๆ ครั้ง โดยมีข้อสมมุติว่า ปราศจากการเรียนรู้ ความเมื่อยล้า และอิทธิพลจากการฝึกฝนในการทดสอบซ้ำ¹

คะแนนความคลาดเคลื่อนได้จากความคลาดเคลื่อนที่เกิดขึ้น โดยบังเอิญหรือไม่เป็นระบบ กล่าวได้ว่าเป็นตัวแปรที่ไม่เกี่ยวข้องกับจุดมุ่งหมายของการทดสอบ และทำให้เกิดความไม่คงที่ในการวัด อาจเกิดขึ้นเนื่องจากตัวแบบสอบเอง การบริหารแบบสอบ การให้คะแนน และสิ่งต่าง ๆ ที่เกิดจากตัวผู้สอบ เป็นต้น ตามทฤษฎีแล้ว ความคลาดเคลื่อนที่เกิดขึ้นอย่างไม่เป็นระบบเหล่านี้จะมีค่าทั้งบวกและลบ ซึ่งจะหักล้างกันเองหมดไป ดังนั้นค่าเฉลี่ยของความคลาดเคลื่อนจะเท่ากับศูนย์ และความคลาดเคลื่อนเหล่านี้จะไม่มีความสัมพันธ์กับคะแนนจริง อาจเขียนเป็นสมการ² ในรูปของความแปรปรวนได้ดังนี้

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

เมื่อ s_x^2 = ความแปรปรวนของคะแนนที่วัดได้

s_t^2 = ความแปรปรวนของคะแนนจริง

¹Frederick G. Brown, Principles of Educational and Psychological Testing (Illinois : The Dryden Press, 1970), p.50.

²William A. Mehrens, and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology, p.91.

s_e^2 = ความแปรปรวนของคะแนนความคลาดเคลื่อน

ในทางทฤษฎี ความเที่ยงก็คืออัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่วัดได้¹ นั่นคือ

$$r_{tt} = \frac{s_t^2}{s_x^2} \dots \dots \dots (1)$$

เมื่อ r_{tt} = ค่าความเที่ยง

ค่าความเที่ยงจึงบอกให้เราทราบถึงความแปรปรวนของคะแนนจริงว่าเป็นสัดส่วนเท่าใดเมื่อเทียบกับความแปรปรวนของคะแนนที่วัดได้ และเนื่องจาก $s_t^2 = s_x^2 - s_e^2$ สมการ (1) จึงนิยมเขียนใหม่ได้เป็น²

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

ซึ่งเป็นสมการที่ใช้เป็นรากฐานในการประมาณค่าความเที่ยงจากคะแนนผลการสอบ³

แบบของความเที่ยง

006410

มีวิธีต่าง ๆ ในการคำนวณหาค่าความเที่ยง แต่ละวิธีทำให้เกิดความคลาดเคลื่อนไม่เหมือนกัน การที่จะเลือกใช้วิธีใดนั้นขึ้นอยู่กับความมุ่งหมายและลักษณะของคะแนนที่ได้จากการทดสอบนั้น วิธีหาค่าความเที่ยงโดยทั่วไปมี 4 แบบ⁴ คือ

¹Ibid.

²Ibid.

³Frederick G. Brown, Principles of Educational and Psychological Testing, p.57.

⁴อนันต์ ศรีโสภากา, การวัดและประเมินผลการศึกษา (กรุงเทพฯ : ไทยวัฒนาพานิช, 2520), หน้า 49.

1. แบบสอบซ้ำ (Measure of Stability)
2. แบบใช้ข้อสอบคล้ายกัน (Measures of Equivalence)
3. แบบใช้ข้อสอบคล้ายกันและสอบซ้ำ (Measures of Equivalence and Stability)
4. แบบวัดความคงที่ภายใน (Measures of Internal Consistency)
 - 4.1 แบบแบ่งครึ่งจำนวนข้อสอบ (Split-Half)
 - 4.2 แบบคูเคอร์-ริชาร์ดสัน (Kuder-Richardson Estimates)
 - 4.3 แบบสัมประสิทธิ์อัลฟา (Coefficient Alpha)
 - 4.4 แบบการวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt's Analysis of Variance Procedure)

ในที่นี้จะกล่าวถึงเฉพาะความเที่ยงแบบคูเคอร์-ริชาร์ดสัน ซึ่งเป็นวิธีที่ใช้ในการวิจัยนี้เท่านั้น

ความเที่ยงแบบคูเคอร์-ริชาร์ดสัน

เมื่อปีค.ศ. 1937 คูเคอร์ และ ริชาร์ดสัน ได้คิดวิธีสำหรับหาความคงที่ภายใน โดยไม่ต้องแบ่งครึ่งแบบสอบ สมการที่มีชื่อเสียงและใช้กันอย่างกว้างขวางก็คือ สูตรคูเคอร์-ริชาร์ดสัน 20¹ (Kuder-Richardson Formula 20)

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

เมื่อ r_{tt} = ค่าความเที่ยงของแบบสอบ
 k = จำนวนข้อในแบบสอบ
 s_t^2 = ความแปรปรวนของคะแนนทั้งหมด

¹ Frederick G. Brown, Principles of Educational and Psychological Testing, p.78.

$$P_i = \text{สัดส่วนของผู้ตอบถูกในแต่ละข้อ}$$

$$q_i = 1 - p_i = \text{สัดส่วนของผู้ตอบผิดในแต่ละข้อ}$$

การหาความเที่ยงแบบคูเคอร์-ริชาร์คสันนี้ดำเนินการสอบครั้งเดียว ด้วยแบบสอบชุดเดียว ความเที่ยงแบบนี้ขึ้นอยู่กับความสม่ำเสมอของผู้รับการทดสอบข้อสอบทุกข้อในแบบสอบ ความสอดคล้องกันระหว่างข้อได้รัับอิทธิพลจากแหล่งความแปรปรวนคลาดเคลื่อน 2 ชนิด คือ 1) เนื้อหาที่สุ่มมา 2) ความเป็นวิวิธพันธ์ของพฤติกรรมที่สุ่มมา ยิ่งข้อสอบมีความเป็นเอกพันธ์มากขึ้น ความสอดคล้องกันระหว่างข้อก็ยิ่งสูงขึ้น¹

ความเที่ยงที่ได้จากสูตรนี้มีค่าประมาณได้กับ

1. ค่าสหสัมพันธ์เฉลี่ยระหว่าง k ข้อกระทงในแบบสอบ
2. ค่าสัมประสิทธิ์ความเที่ยงของข้อกระทงที่เป็นตัวแทนในแบบสอบ
3. ค่าเฉลี่ยของสัมประสิทธิ์ความเที่ยงที่คำนวณจากสูตรสเปียร์แมน บราวน์ (Spearman-Brown Formula) ซึ่งแบ่งครึ่งแบบสอบด้วยวิธีต่าง ๆ²

ข้อตกลงเบื้องต้นของสูตรคูเคอร์-ริชาร์คสัน 20 ก็คือ

1. การให้คะแนนสำหรับแต่ละข้อกระทงเป็น 1 เมื่อตอบถูก และเป็น 0 เมื่อตอบผิด³
2. ข้อกระทงในแบบสอบต้องวัดในเรื่องเดียวกัน⁴

¹ แอน อนาสตาซี, การวัดเชิงจิตวิทยา, หน้า 84.

² Julian C. Stanley, and Kenneth D. Hopkins, Educational and Psychological Measurement and Evaluation (New Delhi : Prentice-Hall of Indian Private Limited, 1978), p.126.

³ Robert L. Ebel, Measuring Educational Achievement, p.326.

⁴ Frederick G. Brown, Principles of Educational and Psychological Testing, p.78.

นอกจากนี้สูตรนี้เหมาะสำหรับใช้คำนวณค่าความเที่ยงของแบบสอบถาม
สามารถ (Power Test) เท่านั้น แต่ไม่เหมาะสำหรับหาค่าความเที่ยงของแบบสอบ
ความเร็ว (Speed Test) เพราะค่า p_1 และ q_1 ของแต่ละข้อจะคำนวณได้ก็ต่อเมื่อทุกคน
ได้มีโอกาสทำข้อสอบข้อนั้นแล้ว¹ ซึ่งข้อมเป็นไปไดยากสำหรับแบบสอบถามเร็ว

องค์ประกอบที่มีอิทธิพลต่อความเที่ยง

นอกจากวิธีการที่ใช้ในการคำนวณหาค่าความเที่ยงแล้ว เมเร็นส์ และ เลแมน
(Mehrens and Lehmann) ได้กล่าวถึงองค์ประกอบอื่น ๆ ที่มีผลต่อความเที่ยงของแบบ
สอบ ได้แก่

1. ความยาวของแบบสอบ การเพิ่มจำนวนข้อสอบที่มีคุณภาพดีเท่าเทียมกับ
ข้อสอบเดิมที่มีอยู่จะทำให้แบบสอบมีความเที่ยงมากขึ้น นั่นคือแบบสอบที่ยาวจะมีความเที่ยง
มากกว่าแบบสอบที่สั้น
2. ความเป็นเอกพันธ์ของกลุ่มผู้รับการทดสอบ (Group Homogeneity)
ถ้ากลุ่มผู้รับการทดสอบมีความสามารถแตกต่างกันน้อย จะทำให้ความเที่ยงของแบบสอบมีค่า
ค่า แต่ถ้านักคิดในกลุ่มมีความสามารถต่างกันมาก ทำให้พิสัยของคะแนนมีค่ามาก ก็จะทำให้ค่า
ความเที่ยงสูงขึ้นด้วย
3. ความยากง่ายของข้อสอบ ความยากของข้อกระทงและของแบบสอบ
ทั้งฉบับย่อมมีผลต่อความเที่ยงของแบบสอบ เนื่องจากการประมาณค่าความเที่ยงของแบบสอบ
ขึ้นอยู่กับความแปรปรวนของคะแนนสอบ ถ้าข้อสอบยากเกินไปหรือง่ายเกินไปจะทำให้พิสัย
ของคะแนนลดลง ซึ่งจะทำให้ค่าความเที่ยงลดลงด้วย อาจกล่าวได้ว่าค่าความเที่ยงจะสูงได้
เมื่อข้อสอบมีระดับความยากที่ทำให้คะแนนมีการกระจายอย่างมา ก ๆ นั่นเอง
4. ความเร็ว (Speed) แบบสอบความเร็วมักไม่เหมาะในการหา
ความเที่ยง โดยการวัดความคงที่ภายใน เพราะจะให้ค่าความเที่ยงสูงกว่าที่เป็นจริง

¹Ibid., p.79.

5. ความเป็นปรนัย (Objectivity) ถ้าทำให้คะแนนมีความเป็นปรนัยสูง จะช่วยให้ความเที่ยงของแบบสอบมีค่าสูงขึ้น¹

นอกจากนั้น การบริหารแบบสอบ เช่น การจับเวลา การให้คำชี้แจง การป้องกัน การทุจริต เป็นต้น หรือองค์ประกอบในตัวผู้สอบเอง เช่น แรงจูงใจ ความวิตกกังวล ประสิทธิภาพในการสอบ การเคา เป็นต้น สิ่งเหล่านี้อาจเป็นแหล่งความคลาดเคลื่อนที่ทำให้ค่าความเที่ยงของแบบสอบลดลงได้

ความตรงของแบบสอบ

นิยามของความตรง

มีผู้ให้นิยามของความตรงไว้มากมาย ซึ่งส่วนใหญ่มีความหมายคล้ายคลึงกัน เช่น ความตรง หมายถึง ความแม่นยำของแบบวัดที่สามารถวัดค่าได้ในสิ่งที่ต้องการให้มันวัด²

ความตรงของการวัดจะบอกให้ทราบว่า การวัดนั้นได้ทำหน้าที่ตามต้องการได้ก็เพียงไร นั่นคือ หมายถึง ระดับที่การวัดนั้นสามารถบรรลุถึงความวัตถุประสงค์ที่กำหนดไว้³

ความตรงของแบบสอบ หมายถึง ความสามารถที่แบบสอบนั้นบรรลุถึงจุดมุ่งหมายอย่างใดอย่างหนึ่ง⁴

¹William A. Mehrens, and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology, pp.100-103.

²David Magnusson, Test Theory, Translated by Hunter Mabon (Massachusetts : Addison Wesley, 1967), p.128.

³Julian C. Stanley, and Kenneth D.Hopkins, Educational and Psychological Measurement and Evaluation, p.101.

⁴William A. Mehrens, and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology, p.109.

แนบ 1

แบบสอบจะมีความตรงมากถ้ามันสามารถชี้วัดความใค้อย่างสมบูรณ์และ

แบบสอบที่มีความตรงสูง หมายถึง แบบสอบนั้นสามารถวัดในสิ่งที่เรา
ต้องการวัดใค้อย่างถูกต้องมาก²

แบบของความตรง

แบบของความตรงที่ใช้กัน โดยทั่วไปในการวัดทางการศึกษาและจิตวิทยามี 3

ชนิด³ คือ

1. ความตรงตามเนื้อหา (Content Validity)
 2. ความตรงตามเกณฑ์สัมพันธ์ (Criterion Related Validity)
 3. ความตรงเชิงโครงสร้าง (Construct Validity)
- กรอนลันด์ (Gronlund) ได้สรุปความหมายโดยทั่วไปของความตรงทั้ง 3

ชนิดใค้ดังนี้

1. ความตรงตามเนื้อหา

ความหมาย : แบบสอบวัดเนื้อหาวิชาและพฤติกรรมที่กำหนดใค้ใค้ใค้
เพียงใค้

วิธีการ : เปรียบเทียบเนื้อหาในแบบสอบกับเนื้อหาวิชาและ
พฤติกรรมใค้ต้องการวัดทั้งหมด

¹Lee J. Cronbach, "Validity" in Encyclopedia of Educational Research Edited by C.W. Harris (New York : The Macmillan Company, 1960), Quoted in Robert L. Ebel, Measuring Educational Achievement, p.379.

²อนันต์ ศรีโสภณ, การวัดและประเมินผลการศึกษา, หน้า 69.

³Norman E. Gronlund, Measurement and Evaluation in Teaching, p.81.

2. ความตรงตามเกณฑ์สัมพันธ

ความหมาย : ผลที่ได้จากแบบสอบถามสามารถทำนายอนาคตได้ก็เพียงใจ
หรือผลที่ได้จากแบบสอบถามใช้ประมาณการกระทำ
ในปัจจุบันที่วัดได้โดยใช้สิ่งอื่นได้ก็เพียงใจ

วิธีการ : เปรียบเทียบคะแนนจากแบบสอบถามกับการกระทำที่เกิดขึ้น
ขึ้นในเวลาต่อมา (สำหรับการทำนาย) หรือเปรียบ
เทียบกับการกระทำอย่างอื่นที่วัดได้ในเวลาใกล้เคียง
กัน (สำหรับประมาณสถานภาพปัจจุบัน)

3. ความตรงเชิงโครงสร้าง

ความหมาย : ผลที่ได้จากแบบสอบถามอธิบายในเชิงจิตวิทยาได้
ก็เพียงใจ

วิธีการ : ใช้การทดลองเพื่อพิจารณาว่ามีองค์ประกอบใดบ้างที่มี
อิทธิพลต่อคะแนนที่ได้จากการสอบ¹

ในที่นี้จะกล่าวถึง เฉพาะความตรงตามเกณฑ์สัมพันธ ซึ่ง เกี่ยวข้องกับการวิจัยนี้

เท่านั้น

ความตรงตามเกณฑ์สัมพันธ

ความตรงชนิดนี้แสดงถึงขนาดของความสัมพันธ์ ระหว่างผลที่ได้จากการทำแบบสอบถาม
กับผลการกระทำที่วัดได้อีกอย่างหนึ่ง² หรืออาจกล่าวได้ว่า ความตรงชนิดนี้เกี่ยวข้องกับการ
ศึกษาความสัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์³

¹ Ibid., p.82.

² Ibid., p.84.

³ William A. Mehrens, and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology, p.112.

เกณฑ์ (Criteria) หมายถึง มาตรฐานในการตัดสิน ในการพัฒนาแบบสอบ หมายถึง คุณลักษณะหรือกลุ่มของคุณลักษณะ (Combination of Characteristics) ที่ใช้เป็นรากฐานในการตัดสินความตรงของแบบสอบหรือขบวนการวัดอื่น ๆ เกณฑ์สำหรับความตรงของคะแนนจากแบบสอบความถนัดโดยทั่วไปก็คือผลสัมฤทธิ์ทางการศึกษาที่วัดได้จากแบบสอบผลสัมฤทธิ์ที่มีคุณภาพดี¹

ความตรงตามเกณฑ์สัมพันธ์ แบ่งออกเป็น 2 แบบ คือ

1. ความตรงเชิงทำนาย (Predictive Validity) หมายถึง ความสัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์ที่วัดได้ในภายหลัง² ความตรงชนิดนี้ใช้กับแบบสอบที่ทำนายผลบางอย่างเมื่อเวลาผ่านไประยะหนึ่ง เช่น แบบสอบสำหรับการแนะนำการศึกษา หรืออาชีพ การคัดเลือกและการจำแนกบุคคลเพื่อการฝึกฝนงานบางประเภท ข้อมูลที่ใช้เป็นเกณฑ์นั้นรวมถึงการวัดผลบางอย่าง เช่น เกรดหลังจากเรียนแล้ว หรือการตัดสินหลังจากการฝึกฝนหรือการว่าจ้างผ่านไประยะหนึ่ง ความตรงเชิงทำนายคำนวณได้เป็นค่าสัมประสิทธิ์ความตรง ซึ่งเป็นค่าสหสัมพันธ์ระหว่างคะแนนสอบกับข้อมูล เกณฑ์

2. ความตรงร่วมสมัย (Concurrent Validity) หมายถึงความสัมพันธ์ระหว่างคะแนนสอบกับการปฏิบัติที่ได้รับ การยอมรับว่าเป็นเกณฑ์ในขณะนั้นสำหรับตัวแปรที่แบบสอบนั้นต้องการวัด³ ความตรงชนิดนี้ใช้กับแบบสอบที่วินิจฉัยสถานภาพที่เป็นอยู่ การวัดตัวแปรที่เป็นเกณฑ์อยู่ในระยะเวลาเดียวกับ การทดสอบ เหตุผลในการสร้างเครื่องมือวัดตัวแปรที่มีข้อมูลแล้ว ก็เพื่อประหยัดเวลาและค่าใช้จ่าย แต่สามารถให้ผล เช่น เกี่ยวกับการวัดเกณฑ์ ความตรงร่วมสมัย แสดงได้เป็นค่าสัมประสิทธิ์ความตรง เช่นเดียวกับความตรงเชิงทำนาย

¹Robert L. Ebel, Measuring Educational Achievement, p.448.

²Ibid., p.380.

³Ibid.

ความตรงเชิงทำนายกับความตรงร่วมสมัยต่างกัน 2 ประเด็นคือ

1. ช่วงระยะเวลาของการเก็บข้อมูลที่เป็นเกณฑ์ เกณฑ์ของความตรงเชิงทำนายจะเกิดขึ้นภายหลังจากการเก็บคะแนนการสอบ ส่วนความตรงร่วมสมัยจะเก็บรวบรวมข้อมูลที่ เป็นเกณฑ์และคะแนนการสอบในเวลาเดียวกัน

2. ความมุ่งหมายของการทดสอบ ความตรงเชิงทำนายเป็นการนำคะแนนทดสอบมาพยากรณ์ความสามารถหรือการกระทำในอนาคต ส่วนความตรงร่วมสมัยเป็นการนำคะแนนทดสอบเทียบกับความสามารถหรือความสำเร็จในขณะนั้น

ในการศึกษาหาความตรงตามเกณฑ์สัมพัทธ์นั้น สิ่งสำคัญที่สุดก็คือการวัดเกณฑ์ที่ถูกต้องและเหมาะสม ในการพิจารณาว่าเกณฑ์ควรคำนึงถึงสิ่งต่อไปนี้

1. ความเกี่ยวข้องหรือความตรง (Relevance or Validity) สิ่งที่จะใช้ เป็นเกณฑ์นั้นต้องมีความเกี่ยวข้องและสะท้อนถึงคุณลักษณะที่สำคัญของความมุ่งหมายของเรื่องที่จะศึกษา การพิจารณาความเกี่ยวข้องของ เกณฑ์นั้น โดยทั่วไปขึ้นอยู่กับ การตัดสินใจคุณค่าไม่ได้วัดออกมาเป็นตัวเลข

2. ความเที่ยง (Reliability) เกณฑ์ที่จะวัดนั้นควรมีความเที่ยงหรือความคงที่ในการวัด เนื่องจากความเที่ยงของแบบสอบมีอิทธิพลต่อค่าสหสัมพันธ์ระหว่างแบบสอบกับเกณฑ์ จากทฤษฎีความสัมพันธ์กล่าวว่า ความสัมพันธ์ระหว่าง 2 ตัวแปรย่อมมีค่าเท่ากับหรือน้อยกว่ารากที่สองของผลคูณระหว่างค่าความเที่ยงของตัวแปรทั้งสอง หรือ

$$r_{XY} \leq \sqrt{(r_{XX})(r_{YY})}$$

เมื่อ r_{XY} = ค่าสหสัมพันธ์ระหว่างตัวพยากรณ์ (X) กับเกณฑ์ (Y)
 r_{XX} = ค่าความเที่ยงของแบบสอบที่เป็นตัวพยากรณ์
 r_{YY} = ค่าความเที่ยงของเกณฑ์

ดังนั้น ค่าความเที่ยงของ เกณฑ์จึงมีอิทธิพลต่อค่าความตรงตามเกณฑ์สัมพันธ์
มากพอ ๆ กับค่าความเที่ยงของตัวพยากรณ์

3. ความเป็นอิสระจากอคติหรือมลทิน (Freedom from Bias or Contamination) เกณฑ์มลทินจะเกิดขึ้น เมื่อคะแนนเกณฑ์ได้รับอิทธิพลจากการทดสอบที่นำมาใช้เป็นตัวพยากรณ์ ตัวอย่าง เช่น ครูผู้สอนนำแบบสอบถามความถนัดทางวิชาการไปทดสอบนักเรียนในคอนต้นปี เพื่อพยากรณ์ผลการเรียนของนักเรียนในคอนปลายปี ถ้าครูพบว่านักเรียนบางคนได้คะแนนความถนัดทางวิชาการต่ำมากก็จะประเมินผลการเรียนหรือให้อันดับคะแนนแก่นักเรียนผู้นั้นต่ำไปด้วย นั่นก็คือ การรู้คะแนนตัวพยากรณ์มีอิทธิพลต่อคะแนนเกณฑ์ อิทธิพลทำนองนี้ทำให้สหสัมพันธ์ระหว่างคะแนนสอบกับ เกณฑ์สูงขึ้นด้วยการเสริมแต่ง ปัญหาดังกล่าวอาจป้องกันได้โดยไม่ยอมให้ผู้มีหน้าที่ตีค่าเกณฑ์ (Rater) รู้คะแนนสอบที่เป็นตัวพยากรณ์

การประมาณค่าความตรงตามเกณฑ์สัมพันธ์โดยทั่วไปนิยมใช้ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน¹ (Pearson's Product Moment Correlation Coefficient)

องค์ประกอบที่มีอิทธิพลต่อความตรง

1. องค์ประกอบในตัวแบบสอบ กรอนลันด์ ได้กล่าวถึงองค์ประกอบภายในตัวแบบสอบที่มีผลทำให้ความตรงต่ำ ได้แก่

1.1 คำชี้แจงไม่ชัดเจน

1.2 ใช้คำยากเกินไป วัดการอ่านมากกว่าอย่างอื่น

1.3 ระดับความยากง่ายของข้อสอบไม่เหมาะสม ในแบบสอบอิงกลุ่ม ข้อสอบที่ยากเกินไปหรือง่ายเกินไปไม่สามารถจำแนกคนได้ ทำให้ความตรงต่ำด้วย

¹J.P. Guilford, Fundamental Statistics in Psychology and Education 3d ed. (Tokyo : Insatsu Printing Co., 1956), pp.140-141.

1.4 การสร้างข้อสอบไม่รัดกุม มีการชี้แนะคำตอบ

1.5 คำพูดกำกวม ทำให้เกิดความสับสน

1.6 แบบสอบสั้นเกินไป

1.7 การเรียงลำดับข้อไม่เหมาะสม

1.8 มีคำตอบที่เป็นระบบแบบแผน ทำให้นักเรียนเดาคำตอบได้¹

นอกจากนั้น ค่าความเที่ยงของแบบสอบก็มีส่วนสำคัญที่จะกำหนดค่าความตรงของแบบสอบ จากสมการ²

$$r_{XY} \leq \sqrt{r_{XX}}$$

เมื่อ r_{XY} = ค่าความตรงของแบบสอบ

r_{XX} = ค่าความเที่ยงของแบบสอบ

จะเห็นว่าค่าความตรงของแบบสอบมีค่าไม่เกินรากที่สองของค่าความเที่ยงของแบบสอบนั้น นั้นย่อมหมายความว่า ค่าความตรงของแบบสอบจะถูกจำกัด โดยค่าความเที่ยงของมันเอง อย่างไรก็ตาม ความเที่ยง เป็นสิ่งที่จำเป็นแต่ก็ยังไม่เพียงพอสำหรับความตรงของแบบสอบ³

2. การบริหารแบบสอบและการให้คะแนน ในการบริหารแบบสอบมาตรฐาน ถ้าไม่ดำเนินการสอบตามคำชี้แจงอย่างเคร่งครัดแล้ว อาจทำให้ความตรงของแบบสอบลดลง นอกจากนั้นการให้คะแนนที่ไม่เป็นปรนัยมีส่วนทำให้แบบสอบมีความตรงต่ำด้วย

¹Norman E. Gronlund, Measurement and Evaluation in Teaching, pp.98-99.

²Frederick G. Brown, Principles of Educational and Psychological Testing, p.99.

³Julian C. Stanley, and Kenneth D. Hopkins, Educational and Psychological Measurement and Evaluation, p.280.

3. องค์ประกอบในตัวนักเรียน ในบาง โอกาสนักเรียนอาจถูกรบกวนสมาธิในการทำแบบทดสอบหรือมีอาการไม่ปกติ มีความวิตกกังวล กลัวและเครียดเกินไปจนไม่สามารถตอบแบบสอบได้อย่างปกติ สิ่งเหล่านี้ย่อมมีผลทำให้ความตรงของแบบสอบลดลง อนึ่ง การเคาะคำตอบของนักเรียนก็เป็นสิ่งที่ทำลายความตรงของแบบสอบด้วย

4. ลักษณะของกลุ่มและเกณฑ์ที่ใช้ โดยทั่วไปแล้ว ค่าความตรงของแบบสอบเหมาะสำหรับอ้างอิงถึงกลุ่มเฉพาะ เนื่องจากสิ่งที่แบบสอบวัดนั้นมักได้รับอิทธิพลจากตัวแปรต่างๆ เช่น อายุ เพศ ระดับความสามารถ ภูมิหลังทางการศึกษา และวัฒนธรรมของผู้ตอบ เป็นต้น ดังนั้น ในการพิจารณาค่าความตรงของแบบสอบที่เป็นมาตรฐาน ควรคำนึงถึงลักษณะของกลุ่มที่ใช้หาความตรง (Validation Group) นั้นด้วย นอกจากนี้ ลักษณะของเกณฑ์ที่ใช้ก็มีอิทธิพลต่อค่าความตรงของแบบสอบดังที่กล่าวมาแล้ว

ตอนที่ 2 งานวิจัยที่เกี่ยวข้องกับลักษณะ โครงสร้าง ข้อระหวง และการจัดตำแหน่งตัวถูก

งานวิจัยที่ศึกษาผลของลักษณะ โครงสร้าง ข้อระหวงต่อความยากของข้อสอบแบบอุปมาอุปไมยนั้นยังมีความเห็นแบ่งออกเป็น 2 ฝ่าย คือ ฝ่ายแรกได้ข้อสรุปว่า การเปลี่ยนลักษณะ โครงสร้าง ข้อระหวงจะทำให้ความยากของข้อสอบอุปมาอุปไมยเปลี่ยนแปลงไปด้วย ส่วนฝ่ายหลังกลับพบว่า การเปลี่ยนลักษณะ โครงสร้าง ข้อระหวงไม่มีผลต่อความยากของข้อสอบอุปมาอุปไมยแต่อย่างใด

แคมเบล¹ (Campbell) มีความคิดเห็นจัดอยู่ในฝ่ายแรก โดยเขาได้เสนอหลักฐานจากผลงานของ เฮม² (Heim) ที่ได้ศึกษาเกี่ยวกับอิทธิพลของลักษณะ โครงสร้าง ข้อระหวงต่อระดับความยากของข้อสอบอุปมาอุปไมย และพบว่าข้อสอบอุปมาอุปไมยที่มีลักษณะ โครงสร้าง

¹Alison C. Campbell, "Some Determinants of Difficulty of Non-Verbal Classification Items," pp.899-913.

²A.W. Heim, The Appraisal of Intelligence (London : Methuen and Company, 1954), cited by Alison C. Campbell, "Some Determinants of Difficulty of Non-Verbal Classification Items," pp.899-913.

แบบเว้นตัวหลังของคู่หลัง (ก : ข → ค : ___) จะมีความยากแตกต่างจากแบบเว้นตัวหน้าของคู่หลัง (ก : ข → ___ : ง) หรือแบบเว้นตัวหลังของคู่หน้า (ก : ___ → ค : ง) หรือแบบเว้นตัวหน้าของคู่หน้า (___ : ข → ค : ง) และ แคมเบลเองก็ทำการศึกษาองค์ประกอบที่กำหนดความยากของข้อสอบจัดจำพวกที่เป็นรูปภาพ โดยนำแบบสอบที่แตกต่างกัน 8 ฉบับ ไปทดสอบกับเด็กนักเรียนที่มีอายุระหว่าง 11 - 12.5 ปี จำนวน 693 คน ผลการวิจัยได้ข้อสรุปที่สำคัญดังนี้

ก. องค์ประกอบภายในที่สำคัญที่สุดคือความยากของข้อสอบจัดจำพวกรูปภาพนั้น ได้แก่คุณสมบัติของรูปภาพในคำต่อไปนี้ เรียงจากง่ายไปยาก คือ รูปร่าง ความลึก การแลเงา ตำแหน่ง และขนาด

ข. ถ้ารูปภาพมีคุณสมบัติแตกต่างกันหลายคำจะทำให้ข้อสอบยากขึ้น

ค. การเปลี่ยนลักษณะ โครงสร้างภายในของข้อกระทงจากแบบกระจาย (Dispersed Form) ไปสู่แบบรวม (Concentrated Form) จะทำให้ข้อสอบยากขึ้น

ง. ลักษณะของรูปภาพที่ใช้ในการยกตัวอย่าง มีผลต่อความยากของข้อสอบที่มีรูปภาพลักษณะทำนองเดียวกันนั้น¹

จากผลการวิจัยดังกล่าว ทำให้ แคมเบล สนับสนุนข้อค้นพบของ เอม ที่ว่าการเปลี่ยนลักษณะ โครงสร้างภายในของข้อกระทงจะทำให้เกิดผลต่อความยากของข้อสอบ สำหรับฝ่ายที่ได้ผลการวิจัยแตกต่างไปจากฝ่ายแรกนั้น ได้แก่ ซิลเวอร์ สไตน์ และ แมคเลน (Silverstein and McLain) ซึ่งได้ศึกษาผลจากการเปลี่ยนลักษณะ โครงสร้างภายในข้อกระทงต่อระดับความยากของข้อสอบแบบอุปมาอุปไมยเชิงภาษา โดยใช้แบบสอบ 4 ฟอรัม ที่มีลักษณะ โครงสร้างของตัวคำถามต่างกัน 4 แบบ คือ ก : ข → ค : ___ , ข : ก → ___ : ค, ค : ___ → ก : ข และ ___ : ค → ข : ก กลุ่มตัวอย่างเป็นนิสิตปริญญาตรีภาควิชาจิตวิทยา จำนวน 160 คน จากผลการวิเคราะห์ข้อมูลไม่ปรากฏ

¹Ibid.

ความแตกต่างระหว่างคะแนนจากแบบสอบทั้ง 4 φόρμ จึงทำให้ได้ข้อสรุปว่า การเปลี่ยนแปลงลักษณะ โครงสร้างภายในของข้อสอบอุปมาอุปไมยเชิงภาษาโดยคงคำตอบเดิมไว้ไม่มีผลต่อความยากของข้อสอบ¹

แคมเบล ได้โต้ตอบผลงานวิจัยดังกล่าวของ ซิลเวอร์สไตน์ และ แมคเลนว่าอาจจะสรุปได้สำหรับการเปลี่ยนแปลงลักษณะ โครงสร้างข้อสอบอุปมาอุปไมยเฉพาะ 4 แบบนี้เท่านั้น สำหรับการเปลี่ยนแปลง โครงสร้างภายในของข้อสอบอุปมาอุปไมยในลักษณะอื่นก็อาจทำให้ความยากของข้อสอบเปลี่ยนแปลงไปได้²

วอลเตอร์ เจ แมคนามารา และ อีลิส เวสแมน (Walter J. McNamara and Ellis Weitzman) ได้ทำการศึกษาถึงผลของตำแหน่งตัวถูกต่อความยากของข้อสอบแบบเลือกตอบ โดยใช้ข้อสอบที่มี 5 ตัวเลือก จำนวน 4,774 ข้อ และข้อสอบที่มี 4 ตัวเลือก จำนวน 3,918 ข้อ ทั้งหมดเป็นเนื้อหาในวิชาคณิตศาสตร์ ฟิสิกส์ หลักการบิน วิชาการบิน และวิชาเครื่องยนต์การบิน กลุ่มตัวอย่างเป็นนักเรียนจาก โรงเรียนเตรียมทหารอากาศ 20 โรงเรียน ผลการวิเคราะห์ข้อมูลได้ข้อสรุปที่สำคัญคือ

ก. ตำแหน่งตัวถูกในข้อสอบเลือกตอบแบบ 4 ตัวเลือก หรือ 5 ตัวเลือก มีผลต่อระดับความยากของข้อสอบอยู่บ้าง

ข. ข้อสอบที่มี 4 ตัวเลือก ความยากจะเพิ่มขึ้นตามลำดับที่ของตัวถูก คือจากตำแหน่งแรกจนถึงตำแหน่งที่ 3 จะยากที่สุด และเมื่อถึงตำแหน่งที่ 4 ความยากจะลดลงใกล้เคียงกับตำแหน่งแรก

ค. ข้อสอบที่มี 5 ตัวเลือก ถ้าจัดตัวถูกไว้ตำแหน่งที่ 2, 3 จะง่ายที่สุด

¹A.B. Silverstein, and R.E. McLain, "Note on the Internal Structure of Verbal Analogy Items," p.434.

²Alison C. Campbell, "Internal Structure of Items, Item Difficulty, and Solution Process : Rejoinder to Silverstein and McLain," Psychological Report 13 (December 1963) : 753-754.

ตำแหน่งที่ 1 จะยากมากกว่า และตำแหน่งที่ 4 จะยากที่สุด ส่วนตำแหน่งที่ 5 มีความยากใกล้เคียงกับตำแหน่งที่ 1

ง. เนื่องจากผลอันเนื่องมาจากตำแหน่งตัวถูกต่อความยากของข้อสอบนั้นมีน้อยมากและจะสังเกตเห็นได้เพียงเล็กน้อยก็ต่อเมื่อใช้กลุ่มตัวอย่างและข้อสอบจำนวนมาก ดังนั้นปัญหานี้จึงมีความสำคัญเมื่อใช้ข้อสอบจำนวนมาก ๆ กับกลุ่มตัวอย่างขนาดใหญ่เท่านั้น¹

อย่างไรก็ตาม ปัญหานี้ก็ยังมี การวิจัยที่ ได้ผลแตกต่างออกไป เช่น ผลการวิจัยของ อาเธอร์ มาร์คัส (Arthur Marcus) ที่ ได้ทำการศึกษาค้นคว้าของตำแหน่งตัวถูกต่อระดับความยากของข้อสอบแบบเลือกตอบ กลุ่มตัวอย่าง เป็นนักศึกษาของมหาวิทยาลัยแมสซาชูเซต (University of Massachusetts) จำนวน 434 คน ซึ่งถูกแบ่งออกเป็น 4 กลุ่ม โดยใช้คะแนนสอบกลางเทอมเป็นเกณฑ์ในการแบ่งกลุ่ม เครื่องมือที่ใช้คือแบบสอบวิชาจิตวิทยา จำนวน 100 ข้อ นำมาดัดแปลงให้เป็นแบบสอบที่ต่างกัน 4 ฟอรัม โดยจัดให้ข้อเดียวกันในแต่ละฟอรัมมีตำแหน่งตัวถูกแตกต่างกัน ผลการวิเคราะห์ข้อมูลปรากฏว่า คะแนนเฉลี่ยของนักเรียนทั้ง 4 กลุ่ม แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ระดับความยากของแต่ละข้อ-กระทง เมื่อตัวถูกอยู่ที่ตำแหน่งต่างกันไม่ปรากฏว่าแตกต่างกัน²

ต่อมา เมอส์ ซี เอส และ รีน วี เดวิส (Merle C. Ace and Rene V. Davis) ได้ทำการศึกษาอิทธิพลของลักษณะ โครงสร้างข้อกระทงและตำแหน่งตัวถูกต่อระดับความยากของข้อสอบแบบอุปมาอุปไมย โดยนำเอาข้อสอบอุปมาอุปไมย จำนวน 30 ข้อ ที่สุ่มเลือกมาได้ มาดัดแปลงเป็นข้อสอบที่มีตำแหน่งช่องว่างในตัวคำถามต่างกัน 4 แบบ คือ ก : ข → ค : __ , ข : ก → __ : ค, ค : __ → ก : ข และ __ : ค → ข : ก

¹Walter J. McNamara and Ellis Weitzman, "The Effect of Choice Placement on the Difficulty of Multiple-Choice Questions," Journal of Educational Psychology 36 (February 1945) : 103 - 113.

²Arthur Marcus, "The Effect of Correct Response Location on the Difficulty Level of Multiple-Choice Questions," pp.43-51

และตำแหน่งตัวถูกของแต่ละข้อก็เปลี่ยนไปอย่างมีระบบในจำนวน 5 ตัวเลือก ทำให้ได้แบบสอบรูปมาอุปไมยที่มีตำแหน่งช่องว่างในตัวคำถามและตำแหน่งตัวถูกแตกต่างกัน (4×5) 20 พอร์ม แบบสอบเหล่านี้ได้นำไปทดสอบนิสิตมหาวิทยาลัย ชั้นปีที่ 2 จำนวน 221 คน ผลจากการวิเคราะห์รายข้อสรุปได้ว่า

ก. ตำแหน่งช่องว่างในตัวคำถาม ไม่มีอิทธิพลต่อระดับความยากของข้อสอบแบบอุปมาอุปไมย

ข. ตำแหน่งตัวถูกอาจเป็นตัวการที่มีอิทธิพลต่อความยากของข้อสอบและมีแนวโน้มที่ว่า ถ้าจัดตัวถูกไว้ที่ตำแหน่งสุดท้ายคือตำแหน่งที่ 5 ของตัวเลือกจะทำให้ข้อสอบยากที่สุด และรองลงมาคือตำแหน่งที่ 3 ปรากฏการณ์นี้อาจเป็นลักษณะเฉพาะของข้อสอบแบบอุปมาอุปไมยที่ต่างไปจากข้อสอบเลือกตอบอื่น ๆ

ค. ปฏิกริยาร่วมระหว่างตำแหน่งช่องว่างในตัวคำถามและตำแหน่งตัวถูก อาจมีผลต่อระดับความยากของข้อสอบอุปมาอุปไมย

ง. ความยากของข้อสอบที่เปลี่ยนไปเนื่องจากการเปลี่ยนตำแหน่งตัวถูกนั้นมักเกิดขึ้นกับข้อที่ค่อนข้างง่าย ($P > .50$) มากกว่าข้อที่ค่อนข้างยาก ($P < .50$)¹

จอห์น ซี เจสเซล และ วอลเตอร์ แอล ซุลลินส์ (John C. Jessell and Walter L. Sullins) ได้ทำการศึกษาลักษณะของการจัดเรียงลำดับตัวถูกที่คะแนนและความเที่ยงของแบบสอบ โดยนำเอาข้อสอบวิชาวิทยาศาสตร์ชีวภาพ จำนวน 60 ข้อ แบบ 4 ตัวเลือก มาดัดแปลงเป็นแบบสอบที่ต่างกัน 7 ฉบับคือ

ฉบับที่ 1 ให้ตัวถูกอยู่ที่แต่ละตำแหน่งตัวเลือกด้วยสัดส่วนเท่า ๆ กัน และไม่ให้ข้อที่มีตัวถูกอยู่ตำแหน่งเดียวกันอยู่ติดต่อกันเกินกว่า 2 ข้อ

ฉบับที่ 2 จัดให้ 7 ข้อแรก ติดต่อกัน มีตัวถูกอยู่ตำแหน่งที่ 2 ข้อที่เหลือให้เหมือนกับในฉบับที่ 1

¹Merle C. Ace, and Rene V. Dawis, "Item Structure as a Determinant of Item Difficulty in Verbal Analogies," pp. 143-149.

- ฉบับที่ 3 จักให้ 7 ข้อกลาง (ข้อ 27 - 33) คัดลอกกัน มีตัวถูกอยู่ตำแหน่ง
ที่ 2 ข้ออื่น ๆ ที่เหลือจักเหมือนฉบับที่ 1
- ฉบับที่ 4 จักให้ 7 ข้อสุดท้ายมีตัวถูกอยู่ตำแหน่งที่ 2 ข้ออื่น ๆ ที่เหลือจัก
เหมือนฉบับที่ 1
- ฉบับที่ 5 จักให้ 14 ข้อแรก คัดลอกกัน มีตัวถูกอยู่ตำแหน่งที่ 2 ข้ออื่น ๆ
เหมือนในฉบับที่ 1
- ฉบับที่ 6 จักให้ 14 ข้อกลางคัดลอกกัน มีตัวถูกอยู่ตำแหน่งที่ 2 ข้ออื่น ๆ จัก
เหมือนในฉบับที่ 1
- ฉบับที่ 7 จักให้ 14 ข้อสุดท้ายคัดลอกกัน มีตัวถูกอยู่ตำแหน่งที่ 2 ข้ออื่น ๆ
จักเหมือนในฉบับที่ 1

เมื่อนำเอาแบบสอบทั้ง 7 ฉบับนี้ไปทดสอบนิสิตชั้นปีที่ 1 มหาวิทยาลัยอินเดียนนา (Indiana State University) จำนวน 454 คน ปรากฏผลว่า คะแนนเฉลี่ยและค่าความเที่ยงของแบบสอบทั้ง 7 ฉบับ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ¹

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

¹John C. Jessell, and Walter L. Sullins, "The Effect of Keyed Response: Sequencing of Multiple Choice Items on Performance and Reliability," Journal of Educational Measurement 12 (Spring, 1975): 45 - 48.