



วรรณคดีที่เกี่ยวข้อง

วรรณคดีที่เกี่ยวข้องจะแบ่งเป็น 2 ตอน คือ ตอนที่ 1 ว่าด้วยสังข์กับเบ็องตัน  
เกี่ยวกับความยากของข้อสอบ ความยาก ความเที่ยงและความตรงของแบบสอบ ส่วนตอน  
ที่ 2 เป็นวรรณคดี และงานวิจัยที่เกี่ยวข้องกับการเรียงลำดับข้อสอบและการไถ่ราคาความ  
ยากของข้อสอบ

ตอนที่ 1 สังข์กับเบ็องตันเกี่ยวกับความยากของข้อสอบ ความยาก ความเที่ยงความตรง  
ของแบบสอบ

ในการวิจัยนี้ ผู้วิจัยมีจุดมุ่งหมายที่จะศึกษาว่าถ้าผู้สอบไถ่ราคาความยากของข้อสอบ  
แต่ละข้อ แล้วจะมีผลอย่างไรกับคะแนนเฉลี่ยจากแบบสอบ ซึ่งชี้แทนความยากของแบบสอบ  
ทั้งหมด เพราะว่า "คะแนนเฉลี่ยของแบบสอบย่อมได้จากความยากเฉลี่ยของข้อสอบทั้งหมด  
ที่เป็นส่วนประกอบในแบบสอบนั้น"<sup>1</sup>

ความยากของข้อสอบแต่ละข้อขึ้นอยู่กับสัดส่วนหรือเปอร์เซ็นต์ของผู้สอบที่ตอบข้อ  
สอบข้อนั้นถูก ถ้าเปอร์เซ็นต์ของผู้ตอบถูกมีมาก ก็แสดงว่าข้อสอบนั้นง่าย ค่าความยากของ  
ข้อสอบไม่ได้เป็นเครื่องชี้ว่าข้อสอบนั้นดีหรือเลว มันเป็นเพียงแต่ระบุว่าข้อสอบข้อนั้นยาก  
หรือง่ายเท่านั้น "อย่างไรก็ตามดัชนีความยากของข้อสอบก็ยังสะท้อนให้เห็นความสามารถ  
ของกลุ่มผู้เข้ารับการทดสอบด้วย"<sup>2</sup> เหตุผลสำคัญในการหาความยากของข้อสอบก็เพื่อว่า

---

<sup>1</sup>Robert L. Ebel, Measuring Educational Achievement,  
(New Jersey: Prentice-Hall, 1965), p.300.

<sup>2</sup>Ibid., p.360.

จะคัดเลือกว่าข้อสอบที่มีความยากไม่เหมาะสมออก ข้อสอบที่ไม่มีใครทำถูกเลยกับข้อที่ทุก ๆ คนทำได้หมดนั้นหากเป็นแบบสอบที่ต้องการจำแนกคนที่มีความสามารถแตกต่างกัน ออกจากกันแล้วควรจะคัดออก เพราะข้อสอบทั้งสองประเภทนี้ไม่สามารถแยกความแตกต่างของผู้สอบได้ และข้อสอบแบบนี้ไม่ได้มีผลต่อความแปรปรวนของคะแนนสอบ จึงไม่มีผลต่อค่าความเที่ยง และความตรงของแบบสอบนัก<sup>1</sup> ในการวัดแบบอิงเกณฑ์ (Criterion reference) โดยทั่วไปข้อสอบจะมีค่าความยากสูง นักเรียนส่วนมากจะได้คะแนนสูง ซึ่งจะช่วยให้ทราบว่า มีนักเรียนคนใดบ้างที่มีความรู้และทักษะไม่ถึงเกณฑ์ที่เราคาดหวัง<sup>2</sup> ส่วนในการวัดแบบอิงกลุ่ม (Norm reference) นั้น เพื่อให้ความจำแนกระหว่างบุคคลมีค่าสูงสุด จึงเป็นการดีที่สุดที่จะเลือกข้อสอบที่มีความยากประมาณ .50 ซึ่งจะทำให้ได้คะแนนรวมที่มีการกระจายมากที่สุด ซึ่งจะมีผลทำให้ค่าสัมประสิทธิ์ความเที่ยงมีค่าสูงสุดด้วย<sup>3</sup> นั้นแสดงว่าค่าความยากยอมส่งผลถึงความเที่ยงและความตรงของแบบสอบด้วย

อย่างไรก็ดีค่าความยากนี้มีความสัมพันธ์กับค่าอำนาจจำแนก ค่าความยากที่ทำให้ค่าอำนาจจำแนกมากที่สุดคือ ค่าความยากเท่ากับ .50 แต่พึงระลึกอยู่เสมอว่า การที่ข้อสอบมีค่าความยากเท่ากับ .50 อาจจะทำให้ข้อสอบนั้นมีค่าอำนาจจำแนกสูงสุด หรือไม่มีเลยก็ได้ เพราะถ้าครึ่งหนึ่งของจำนวนคนในกลุ่มสูงตอบถูก และครึ่งหนึ่งของจำนวนคนในกลุ่มต่ำตอบถูก ก็ทำให้ค่าอำนาจจำแนกเท่ากับ 0<sup>4</sup> นั้นแสดงว่าค่าความยากเท่ากับ .50 นั้นอาจทำให้ข้อสอบมีอำนาจจำแนกสูงสุดหรือไม่มีอำนาจจำแนกเลยก็ได้

<sup>1</sup>แอน อมาตยาชี, การตรวจสอบเชิงจิตวิทยา, แปลโดย ประชุมสุข อาชวอร่ารุง และคณะ, (กรุงเทพฯ: ไทยวัฒนาพานิช, 2519), หน้า 161-162.

<sup>2</sup>อนันต์ ศรีโสภาก, ทฤษฎีการวัดและการทดสอบ (กรุงเทพฯ: คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2519), หน้า 257.

<sup>3</sup>แอน อมาตยาชี, การตรวจสอบเชิงจิตวิทยา, หน้า 161-162.

<sup>4</sup>อนันต์ ศรีโสภาก, ทฤษฎีการวัดและการทดสอบ, หน้า 257.

## องค์ประกอบที่มีอิทธิพลต่อความยากของแบบสอบ

แคมเบล (Campbell)<sup>1</sup> ได้แบ่งองค์ประกอบที่มีอิทธิพลต่อความยากของแบบสอบออกเป็น 2 ประเภท คือ

1. องค์ประกอบภายใน (Intrinsic Factors) ประกอบด้วย
  - ก. เนื้อหาของข้อสอบแต่ละข้อ (Item Content) ในด้านความซับซ้อน ความเป็นนามธรรม และความแปลกใหม่
  - ข. ลักษณะโครงสร้างของข้อสอบ (Item Structure) หมายถึงวิธีการในการแสดงออกซึ่งเนื้อหา
2. องค์ประกอบภายนอก (Extrinsic Factors) ประกอบด้วย
  - ก. ความไม่คุ้นเคยต่อเนื้อหา (Unfamiliarity) คืออยู่นอกเหนือประสบการณ์ของผู้สอบ
  - ข. สิ่งที่มีสัมพันธ์กับข้อสอบ (Item Context) เช่น ข้อสอบข้อที่อยู่ใกล้เคียงกัน
  - ค. ตัวแปรด้านบุคลิกภาพ ได้แก่ สภาพร่างกาย ลักษณะนิสัย และความตั้งใจของผู้สอบ

นอกจากนี้องค์ประกอบที่อาจจะมีผลต่อค่าความยากของแบบสอบก็คือ ธรรมชาติของเนื้อหา ชนิดของพฤติกรรมที่ต้องกวาดจากผู้สอบ และตัวการอื่น ๆ ที่สัมพันธ์กันมองเห็นไม่เด่นชัด เช่น ลักษณะของภาษาที่ใช้ รูปแบบของคำถาม ตลอดจนคำชี้แจงและอื่น ๆ ซึ่งสิ่งเหล่านี้ผู้สร้างแบบสอบควรจะสนใจนำเข้ามาพิจารณาในการสร้างแบบสอบด้วย<sup>2</sup>

<sup>1</sup>Alison C. Campbell, "Some Determinants of the Difficulty of Non-Verbal Classification Items" Educational and Psychological Measurement 21 (Winter 1961): 899-913

<sup>2</sup>K. W. Vaughn, "Planning the Objective Test" Educational Measurement, p. 174. Edited by H. I. Lindquist. (Washington D.C.: American Council on Education, 1961), p. 174.

## ความเที่ยงของแบบสอบ

นิยามของความเที่ยง

ได้มีผู้ให้ความหมายของคำว่า ความเที่ยงมากมาย เช่น

ความเที่ยง หมายถึง ความคงที่ของคะแนนที่ได้จากการวัด<sup>1</sup>

ความเที่ยง หมายถึง ความคงที่ (Consistency) ของคะแนนที่ได้จากการทดสอบบุคคล กลุ่มเดียวกันด้วยแบบสอบชุดเดียวกันในเวลาที่แตกต่างกัน หรือสอบด้วยแบบสอบคนละชุดที่เทียบเท่ากัน หรือภายใต้สภาพการสอบที่ต่างกัน<sup>2</sup>

ความเที่ยง หมายถึง ระดับ (Degree) ของความคงที่ระหว่างการวัดสิ่งเดียวกัน 2 ครั้ง<sup>3</sup>

จากนิยามความเที่ยงดังกล่าวข้างต้น ก็จะสรุปได้ว่า ความเที่ยง หมายถึง ความคงที่ของคะแนนที่ได้จากการวัด ด้วยเครื่องมือชุดเดียวกันในโอกาสที่ต่างกัน

ทฤษฎีของความเที่ยง

การอธิบายทฤษฎีความเที่ยงโดยเริ่มจากคะแนนสอบที่นักเรียนทำได้ (Observed Scores) ซึ่งคะแนนที่นักเรียนทำได้นี้ประกอบไปด้วยส่วนที่เป็นคะแนนจริง (True Scores) และส่วนที่เป็นคะแนนที่เกิดจากความคลาดเคลื่อน (Error Score) ซึ่งเขียนเป็นสมการได้ดังนี้

$$X = T + E$$

<sup>1</sup> Norman E. Gronlund, Measurement and Evaluation in Teaching, 3d ed. (New York: Macmillan, 1976), p.105.

<sup>2</sup> แอน อนาสตาซี, การตรวจสอบเชิงจิตวิทยา, หน้า 73.

<sup>3</sup> William A. Mehren & Ervin J. Lehman, Measurement and Evaluation in Education and Psychology, p.88.

<sup>4</sup> Ibid., p.90.

เมื่อ X แทน คะแนนที่นักเรียนทำได้

T แทน คะแนนจริง

E แทน คะแนนที่เกิดจากความคลาดเคลื่อน

คะแนนจริง (True Score) หมายถึง คะแนนที่ผู้สอบควรจะได้รับถ้าเครื่องมือที่ใช้นั้นมีลักษณะสมบูรณ์ปราศจากความคลาดเคลื่อน หรือหมายถึง คะแนนเฉลี่ยของผู้สอบซึ่งได้จากการทำแบบสอบเดิมหลาย ๆ ครั้ง โดยมีข้อตกลงว่าไม่มีอิทธิพลจากการฝึกฝน ความเมื่อยล้า และการเรียนรู้ในการทดสอบซ้ำ

คะแนนที่เกิดจากความคลาดเคลื่อน (Error Scores) หมายถึง ค่าความผิดพลาดที่เกิดขึ้นจากการวัด ซึ่งเป็นความคลาดเคลื่อนที่เกิดขึ้นในลักษณะสุ่ม (Random Error) เช่น ความคลาดเคลื่อนที่เกิดจากตัวแบบสอบ การบริหารการสอบ การให้คะแนน หรือเกิดจากตัวผู้สอบเอง โดยทฤษฎีความคลาดเคลื่อนที่เกิดขึ้นในลักษณะสุ่มนี้ อาจเป็นได้ทั้งทางบวกและในทางลบ ซึ่งความคลาดเคลื่อนเหล่านี้จะหักล้างกันหมดไป การแจกแจงของคะแนนความคลาดเคลื่อนเป็นการแจกแจงแบบปกติ และถ้าความคลาดเคลื่อนเกิดขึ้นในลักษณะสุ่มแล้ว ก็จะไม่มีความสัมพันธ์กับคะแนนจริง นั่นคือ สัมประสิทธิ์สหสัมพันธ์ (Correlation) ระหว่างคะแนนจริงกับคะแนนความคลาดเคลื่อนมีค่าเท่ากับ  $0^2$  ดังนั้นความแปรปรวนของคะแนนที่วัดได้ สามารถเขียนเป็นสมการ<sup>3</sup> ดังนี้

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

<sup>1</sup>Frederick G. Brown, Principle of Educational and Psychological Testing, 2d ed. (New York: Holt, Rinehart and Winston, 1976), p.53.

<sup>2</sup>อนันต์ ศรีโสภา, การวัดและการประเมินผลการศึกษา (กรุงเทพฯ: ไทยวัฒนาพานิช, 2520), หน้า 44.

<sup>3</sup>William A. Mohrens and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology, p. 91.

เมื่อ  $s_x^2$  แทน ความแปรปรวนของคะแนนสอบที่วัดได้  
 $s_t^2$  แทน ความแปรปรวนของคะแนนจริง  
 $s_e^2$  แทน ความแปรปรวนของคะแนนความคลาดเคลื่อน

โดยทฤษฎี ความเที่ยงก็คือ อัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่วัดได้ นั่นคือ

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

เมื่อ  $r_{tt}$  แทน ค่าความเที่ยง

ดังนั้น ค่าความเที่ยงจะบอกให้ทราบถึงความแปรปรวนของคะแนนจริงว่าเป็นสัดส่วนเท่าใด เมื่อเทียบกับความแปรปรวนของคะแนนที่วัดได้ และจาก  $s_t^2 = s_x^2 - s_e^2$  ดังนั้นสมการ (1) อาจเขียนได้เป็น

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

ซึ่งเป็นสมการพื้นฐานสำหรับ ค่าสัมประสิทธิ์สหสัมพันธ์ของความเที่ยง<sup>1</sup>

### วิธีประมาณค่าความเที่ยง

วิธีประมาณค่าความเที่ยงสามารถทำได้หลายวิธี แต่ละวิธีทำให้เกิดความคลาดเคลื่อนแตกต่างกัน "ไม่สามารถบอกได้ว่าวิธีใดดีที่สุด เพราะจะใช้วิธีไหนขึ้นอยู่กับวัตถุประสงค์ของผู้ทำ ต้องพิจารณาข้อมูลที่ได้อีก และต้องดูว่าเป็นแบบสอบชนิดใด หรือการวัดแบบใด"<sup>2</sup>

<sup>1</sup>Joy Paul Guilford and Benjamin Fruchter, Fundamental Statistics in Psychology and Education, 6d ed. (Tokyo : Tosho, 1978), p. 410.

<sup>2</sup>Ibid, p.414.

โดยทั่วไปวิธีประมาณค่าทำได้ 4 แบบ<sup>1</sup> คือ

1. แบบวัดซ้ำ (Measure of Stability)
2. แบบใช้ข้อสอบคล้าย (Measure of Equivalence)
3. แบบใช้ข้อสอบคล้ายและวัดซ้ำ (Measure of Equivalence and Stability)
4. แบบวัดความคงที่ภายใน (Measure of internal Consistency)
  - 4.1 แบบแบ่งครึ่งข้อสอบ (Split-Half)
  - 4.2 แบบคูเคอร์ ริชาร์ดสัน (Kuder-Richardson Estimates)
  - 4.3 แบบวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyts Analysis of Variance Procedure)

ในที่นี้จะกล่าวเฉพาะความเที่ยงแบบคูเคอร์-ริชาร์ดสัน ซึ่งเป็นวิธีประมาณค่าความเที่ยงที่ใช้ในการวิจัยครั้งนี้

003994

### ความเที่ยงแบบคูเคอร์-ริชาร์ดสัน

ในปี 1937 คูเคอร์ และริชาร์ดสัน (Kuder & Richardson) ได้คิดหาวิธีสำหรับหาความคงที่ภายในโดยไม่ต้องแบ่งครึ่งข้อสอบ เขาได้เสนอสูตรสำหรับประมาณค่าความเที่ยง ซึ่งเป็นที่รู้จักและนิยมใช้อย่างแพร่หลาย คือ สูตรคูเคอร์-ริชาร์ดสัน 20<sup>2</sup> (Kuder-Richardson 20)<sup>2</sup>

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

<sup>1</sup> William A. Mohrens and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology, p. 109.

<sup>2</sup> Frederick G. Brown, Principles of Educational and Psychological Testing, p. 78.



เมื่อ	$r_{tt}$	แทน	ความเที่ยงของแบบสอบ
	$k$	แทน	จำนวนข้อในแบบสอบ
	$s_t^2$	แทน	ความแปรปรวนของคะแนนทั้งหมด
	$p_i$	แทน	สัดส่วนของผู้ตอบถูกในแต่ละข้อ
	$q_i$	แทน	สัดส่วนของผู้ตอบผิดในแต่ละข้อ = $1 - p_i$

การหาความเที่ยงแบบ คูเคอร์-ริชาร์ดสัน นี้เป็นการหาความสอดคล้องกันระหว่างข้อสอบ ซึ่งดำเนินการสอบครั้งเดียว และใช้แบบสอบชุดเดียว ความสอดคล้องกันระหว่างข้อนี้ ใ้ได้รับอิทธิพลจากแหล่งความแปรปรวนคลาดเคลื่อน 2 แหล่งคือ<sup>1</sup>

1. เนื้อหาที่สุ่ม
2. ความเป็นวิวิธพันธุ์ของพฤติกรรมที่สุ่ม ยิงข้อสอบมีความเป็นเอกพันธ์มากขึ้น ความสอดคล้องกันระหว่างข้อสอบก็ยิ่งสูง

ค่าความเที่ยงแบบคูเคอร์-ริชาร์ดสัน มีค่าประมาณได้กับ

1. ค่าสหสัมพันธ์เฉลี่ยระหว่าง  $k$  ข้อกระทงในแบบสอบ
2. ค่าสัมประสิทธิ์ความเที่ยงของข้อกระทงที่เป็นตัวแทนในแบบสอบ
3. ค่าเฉลี่ยของสัมประสิทธิ์ความเที่ยงที่ประมาณค่าจากสูตร สเปียร์แมน บราวน์

(Spearman-Brown Formula) ซึ่งแบ่งครึ่งข้อสอบ (Split half) ด้วยวิธีต่าง ๆ<sup>2</sup>

ข้อตกลงเบื้องต้นของสูตร คูเคอร์-ริชาร์ดสัน 20

1. การตรวจให้คะแนนข้อสอบแต่ละข้อเป็นแบบ 0-1 คือตอบถูกให้ 1 คะแนน

<sup>1</sup> แอน อนาคตาชี, การตรวจสอบเชิงจิตวิทยา, หน้า 84.

<sup>2</sup> Julian C. Stanley & Kenneth D. Hopkins, Educational and Psychological Measurement and Evaluation, p. 126.



ตอบผิดให้ 0 คะแนน<sup>1</sup>

2. ข้อสอบแต่ละข้อในแบบสอบจะต้องมีลักษณะเป็นเอกพันธ์คือ วัดในเรื่องเดียวกัน<sup>2</sup>

สูตรคูเคอร์-ริชาร์ดสัน เหมาะสำหรับการประมาณค่าความเที่ยงของแบบสอบความสามารถ (Power Test) เท่านั้น ไม่เหมาะสำหรับแบบสอบความเร็ว (Speed Test) เพราะค่า  $p_i$  และ  $q_i$  ของแต่ละข้อ จะหาได้ก็ต่อเมื่อทุกคนทำข้อนั้นแล้ว<sup>3</sup> ซึ่งเป็นไปไม่ได้ ที่ทุกคนจะทำถึงข้อสุดท้ายในข้อสอบแบบนี้

สูตรคูเคอร์-ริชาร์ดสัน 20 ไม่สะดวกในการคำนวณ คูเคอร์-ริชาร์ดสัน ได้เสนอสูตรในการคำนวณที่ง่ายและสะดวกกว่า ได้แก่ สูตรคูเคอร์-ริชาร์ดสัน 21 ซึ่งสูตรนี้มีข้อตกลงว่าข้อสอบแต่ละข้อมีความยากเท่ากัน หรือกำหนดให้ค่า  $p$  คงที่ ซึ่งค่าที่ได้จากสูตร K-R 21 นี้ จะต่ำกว่าสูตร K-R 20 เนื่องจากว่าโดยทั่วไป ข้อสอบแต่ละข้อจะมีความยากแตกต่างกัน ซึ่งไม่เป็นไปตามข้อตกลง

#### องค์ประกอบที่มีอิทธิพลต่อความเที่ยง

นอกจากวิธีการที่ใช้ประมาณค่าความเที่ยงจะมีอิทธิพลต่อค่าความเที่ยงของแบบสอบแล้ว ยังมีองค์ประกอบอื่น ๆ<sup>4</sup> ต่อไปนี้มีอิทธิพลต่อความเที่ยงของแบบสอบด้วย

<sup>1</sup> Robert L. Lbel, Measuring Educational Achievement, p.326.

<sup>2</sup> Frederick G. Brown, Principle of Educational and Psychological Testing, p. 78.

<sup>3</sup> Ibid, p.79.

<sup>4</sup> William A. Mehrens, and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology, pp.100-103.

1. ความยาวของแบบสอบ (Test Length) การเพิ่มจำนวนข้อสอบที่มีคุณภาพก็เท่าเทียมกับข้อสอบเดิม จะทำให้แบบสอบมีค่าความเที่ยงสูงขึ้น นั่นคือแบบสอบที่มีจำนวนข้อมากจะมีความเที่ยงสูงกว่าแบบสอบที่มีจำนวนข้อน้อย สูตรที่แสดงให้เห็นความสัมพันธ์ระหว่างความยาวของแบบสอบกับความเที่ยงคือ สูตรของสเปียร์แมน บราว (Spearman Brown)

2. ความเป็นเอกพันธ์ของกลุ่มผู้รับการทดสอบ (Group Homogeneity) กลุ่มที่มีความเป็นเอกพันธ์ จะมีผลของคะแนนความสามารถของกลุ่มน้อยกว่ากลุ่มที่มีความเป็นวิวิธพันธ์ (Group Heterogeneous) ซึ่งจะทำให้ค่าความเที่ยงของกลุ่มที่มีความเป็นเอกพันธ์ต่ำกว่ากลุ่มที่มีความเป็นวิวิธพันธ์

3. ความยากของข้อสอบ (Item Difficulty) ความยากของข้อสอบ มีอิทธิพลต่อค่าความเที่ยงของแบบสอบ เพราะว่าค่าความเที่ยงนั้นประมาณจากความแปรปรวนของคะแนนผลการสอบ ถ้าข้อสอบยากเกินไปทุกคนทำผิดหมด หรือข้อสอบง่ายเกินไปทุกคนทำถูกหมดก็จะทำให้ผลของคะแนนลดลง ค่าความแปรปรวนระหว่างคะแนนมีน้อย ซึ่งจะทำให้ค่าความเที่ยงลดลงด้วย ดังนั้นถ้าความยากของข้อสอบทำให้คะแนนมีความกระจายมาก ๆ ก็จะทำให้ความเที่ยงมีค่าสูงขึ้นด้วย

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

5. ความเป็นปรนัย (Objective) ถ้าแบบสอบมีความเป็นปรนัยสูง ค่าความเที่ยงของแบบสอบก็สูงขึ้นด้วย

นอกจากนั้นการบริหารแบบสอบ เช่น การจับเวลา การให้คำชี้แจง การป้องกัน การทุจริต เป็นต้น หรือองค์ประกอบในค่านเกี่ยวกับตัวผู้สอบเอง เช่น แรงจูงใจ ความวิตกกังวล ประสบการณ์ในการสอบ การเคา เป็นต้น สิ่งเหล่านี้อาจเป็นแหล่งความคลาดเคลื่อนที่ทำให้ค่าความเที่ยงของแบบสอบเปลี่ยนแปลงได้



## ความตรงของแบบสอบ

นิยามของความตรง

ได้มีผู้ให้คำนิยามความตรงไว้มากมาย เช่น

ความตรง หมายถึง ขอบเขตของผลการประเมินที่สอดคล้องกับจุดประสงค์นั้น<sup>1</sup>

ความตรงของแบบสอบ หมายถึง ความสามารถที่แบบสอบนั้นบรรลุถึงจุดมุ่งหมายอย่างใดอย่างหนึ่ง<sup>2</sup>

ความตรงของแบบสอบ คือ ความสามารถวัดในสิ่งที่ต้องการวัดได้ถูกต้องแม่นยำ หรือ คือความสามารถในการให้ความหมายในสิ่งที่วัดได้อย่างไม่ผิดพลาด<sup>3</sup>

ความตรงของแบบสอบ คือ ค่าสหสัมพันธ์ระหว่างแบบสอบกับเกณฑ์<sup>4</sup>

จากนิยามความตรงดังกล่าวข้างต้น ก็จะสรุปได้ว่าความตรง หมายถึง ความสามารถของเครื่องมือที่จะใช้วัดในสิ่งที่จะวัดได้แม่นยำ สอดคล้องกับวัตถุประสงค์ของการวัด หรือ คือค่าสหสัมพันธ์ระหว่างแบบสอบกับเกณฑ์

## ชนิดของความตรง

ความตรงของแบบสอบมีหลายชนิด ซึ่งแต่ละชนิดแบ่งตามธรรมชาติและจุดมุ่งหมายของการวัด โดยทั่ว ๆ ไปในการวัดทางการศึกษา และทางจิตวิทยาจำแนกออกได้เป็น 3

<sup>1</sup>Norman E. Gronlund, Measurement and Evaluation in Teaching, p.79.

<sup>2</sup>William A. Mohrens & Irvin J. Lehman, Measurement and Evaluation in Education and Psychology, p. 109.

<sup>3</sup>E.E. Linn, A First Course in Statistics (Boston : Houghton Mifflin, 1942), p.213.

<sup>4</sup>H. Gulliksen, Theory of Mental Tests, (New York : John Wiley & Son, 1950), p.88.

ชนิด<sup>1</sup> คือ

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

ความตรงทั้ง 3 ชนิดนี้มีความแตกต่างกันคือ<sup>2</sup>

1. ความตรงตามเนื้อหา นั้นเราดูว่า แบบสอบวัดเนื้อหาวิชาและพฤติกรรมที่กำหนดไว้ได้คือเพียงใด การหาความตรงชนิดนี้ใช้วิธีเปรียบเทียบเนื้อหาในแบบสอบกับเนื้อหาวิชาและพฤติกรรมที่ต้องการวัด

2. ความตรงตามเกณฑ์สัมพันธ์นั้นเราจะดูว่าผลจากแบบสอบสามารถทำนายได้ตรงเพียงใด หรือผลที่ได้จากแบบสอบสามารถใช้ประมาณการกระทำในปัจจุบันที่วัดได้ โดยใช้แบบสอบอื่นได้ตรงเพียงใด การนำความตรงชนิดนี้ใช้การเปรียบเทียบคะแนนจากแบบสอบกับการกระทำที่เกิดขึ้นในอนาคต (สำหรับการทำนาย) หรือเปรียบเทียบกับ การกระทำอย่างอื่นที่วัดได้ในเวลาใกล้เคียงกัน (สำหรับประมาณสถานะภาพปัจจุบัน)

3. ความตรงตามคุณลักษณะนั้นเราดูว่า ผลที่ได้จากแบบสอบสามารถอธิบายในเชิงจิตวิทยาได้คือเพียงใด การหาความตรงชนิดนี้ใช้วิธีการทดลองเพื่อพิจารณาว่า มีองค์ประกอบใบบ้างที่มีอิทธิพลต่อคะแนนที่ได้จากการสอบ

ในที่นี้จะกล่าวถึงเฉพาะความตรงตามเนื้อหา และความตรงเชิงเกณฑ์สัมพันธ์ ซึ่งเกี่ยวข้องกับการวิจัยนี้

---

<sup>1</sup> Norman E. Gronlund, Measurement and Evaluation in Teaching, p. 81.

<sup>2</sup> Ibid., p. 82.

### ความตรงตามเนื้อหา (Content Validity)

ความตรงตามเนื้อหาวิชา คือความตรงที่เกี่ยวกับการสุ่ม (Sampling) เนื้อหาวิชาที่ใช้ในการทดสอบ โดยเฉพาะอย่างยิ่งการทดสอบวัดผลสัมฤทธิ์ทางการเรียน (Achievement Test) มักจะเน้นในเรื่องความตรงในเนื้อหาวิชา การที่จะตัดสินว่าแบบสอบมีความตรงในเนื้อหาวิชามากน้อยเพียงใด มักจะพิจารณาทั้งหัวข้อเนื้อหาวิชา (Subject Matter) และชนิดของพฤติกรรมที่เราต้องการจะวัดไปพร้อมกัน เราจะกำหนดสองสิ่งนี้ในตารางวิเคราะห์หลักสูตร หรือผังการสร้างข้อสอบ (Table of Specification or Test Blueprint) ของแบบสอบวัดผลสัมฤทธิ์ทางการเรียนซึ่งเป็นตาราง 2 มิติ (Two Way Dimensions) ที่จะช่วยให้การสุ่มหัวข้อเนื้อหาวิชาและชนิดของพฤติกรรมที่จะนำมาสร้างเป็นข้อสอบ เป็นตัวแทนที่ดีของการถามเนื้อหาวิชาและพฤติกรรมที่มีอยู่ทั้งหมด<sup>1</sup>

เราไม่สามารถจะแสดงความตรงในเนื้อหาวิชาของแบบสอบออกมาเป็นตัวเลขได้ โดยทั่วไปมักจะพิจารณาข้อสอบในแบบสอบ โดยเทียบกับผังการสร้างข้อสอบ แล้วประมาณออกมาว่ามีความตรงตามเนื้อหา มากน้อยเพียงใด ผู้ที่จะตัดสินความตรงตามเนื้อหาได้ดีที่สุดก็คือ ผู้เชี่ยวชาญในสาขาวิชานั้น ๆ<sup>2</sup> สำหรับแบบสอบผลสัมฤทธิ์ในชั้นเรียน ผู้ที่ตัดสินความตรงตามเนื้อหาได้ดีที่สุดก็คือ ผู้สอนในวิชาที่จะทดสอบนั้น

### ความตรงตามเกณฑ์สัมพันธ

ความตรงตามเกณฑ์สัมพันธ หาได้โดยการเปรียบเทียบคะแนนจากแบบทดสอบกับตัวแปรที่เป็นตัวเกณฑ์ภายนอกหนึ่งตัวหรือมากกว่า ซึ่งเป็นคุณลักษณะที่กล่าวถึงการเปรียบเทียบ อาจเป็นไปในรูปตารางคาดคะเน (expectency tables) หรือค่าสหสัมพันธระหว่าง

<sup>1</sup>อนันต์ ศรีโสภณ, การวัดและประเมินผลการศึกษา, หน้า 69-70.

<sup>2</sup>William A. Mehren and Irvin J. Lehman, Measurement and Evaluation in Education and Psychology, p. 112.

คะแนนทดสอบกับคะแนนตัวเกณฑ์<sup>1</sup> ความตรงตามเกณฑ์สัมพันธ์แบ่งได้เป็น 2 แบบคือ<sup>2</sup>

1. ความตรงเชิงทำนาย (Predictive Validity) เป็นการหาความสัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์ในอนาคตที่เกิดขึ้น ความตรงชนิดนี้ใช้แบบสอบเพื่อทำนายผลบางอย่างที่จะเกิดขึ้นในเวลาต่อมา เช่น แบบสอบความถนัดจะทำนายผลการเรียนในตอนหลัง การหาความตรงชนิดนี้ต้องอาศัยเวลา เพราะต้องหาความสัมพันธ์ระหว่างคะแนนสอบความถนัดกับเกณฑ์ ซึ่งเป็นผลการเรียนที่จะเกิดภายหลัง เหมาะสำหรับใช้กับแบบสอบเพื่อทำนายผล

2. ความตรงร่วมสมัย (Concurrent Validity) เป็นการหาความสัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์ที่เกิดขึ้นในระยะเวลาใกล้เคียงกัน ความตรงชนิดนี้ใช้แบบสอบเพื่อจะดูสถานการณ์ที่เป็นอยู่ในปัจจุบัน เช่น การต้องการทราบผลสัมฤทธิ์ในการเรียนในปัจจุบันของเด็กในวิชาคณิตศาสตร์ ว่าถูกต้องเพียงใดกับสถานการณ์ปัจจุบัน โดยการหาความสัมพันธ์ระหว่างคะแนนคณิตศาสตร์ กับเกณฑ์ที่เกิดในเวลาเดียวกันหรือใกล้เคียงกันว่าสอดคล้องกันเพียงใด ความตรงร่วมสมัยใช้เพื่อดูว่าแบบสอบที่สร้างขึ้นใหม่นั้นสามารถใช้ได้คือเพียงใด เมื่อเทียบกับเกณฑ์

ความแตกต่างระหว่างความตรงเชิงทำนาย กับความตรงร่วมสมัยนั้นพิจารณาได้

2 ประการคือ<sup>3</sup>

1. ช่วงระยะเวลาของการเก็บข้อมูลที่เป็น เกณฑ์กับคะแนนสอบในเวลาเดียวกันหรือใกล้เคียงกัน ก็จะเป็นความตรงร่วมสมัย แต่ถาการเก็บข้อมูลที่เป็น เกณฑ์หลังจากการสอบผ่านไปช่วงระยะเวลาหนึ่งของการสอบแล้ว ก็จะเป็นความตรงเชิงทำนาย

<sup>1</sup>American Psychological Association, Standard for Educational and Psychological Test and Manuals, (Washington, D.C., 1956), p.13.

<sup>2</sup>อนันต์ ศรีโสภณ, การวัดและประเมินผลการศึกษา, หน้า 70.

<sup>3</sup>เรื่องเดียวกัน, หน้า 70.

2. ความมุ่งหมายของการสอบ ความตรงเชิงทำนายเป็นการใช้คะแนนสอบเพื่อทำนายความสามารถหรือการกระทำในอนาคต ส่วนความตรงร่วมสมัยเป็นการนำคะแนนสอบมาเทียบกับความสามารถหรือความสำเร็จในปัจจุบัน

ในการหาความตรงตามเกณฑ์สัมพัทธ์นั้น สิ่งที่สำคัญที่สุดคือ การวัดเกณฑ์ที่ถูกต้องและเหมาะสมในการพิจารณาว่าวัดเกณฑ์นั้นควรคำนึงถึงสิ่งต่อไปนี้<sup>2</sup>

1. ความเกี่ยวข้องหรือความตรง (Relevant or Validity) สิ่งที่จะใช้เป็นเกณฑ์นั้นควรมีความเกี่ยวข้องหรือสะท้อนถึงคุณลักษณะที่สำคัญของความมุ่งหมายของเรื่องที่จะศึกษา โดยทั่วไปจะพิจารณาความเกี่ยวข้องกับการวัดเกณฑ์จากการตัดสินใจว่าจะควรจะใช้อะไรเป็นเกณฑ์

2. ความเที่ยง (Reliability) เนื่องจากความเที่ยงของแบบสอบมีอิทธิพลต่อความสัมพันธ์ระหว่างแบบสอบกับเกณฑ์ จากทฤษฎีความสัมพันธ์กล่าวว่า ความสัมพันธ์ระหว่างตัวแปร 2 ตัวย่อมมีค่าเท่ากับหรือน้อยกว่ารากที่สองของผลคูณระหว่างค่าความเที่ยงของตัวแปรทั้งสองนั้นคือ

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

เมื่อ  $r_{xy}$  แทน สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวทำนาย (X) กับเกณฑ์ (Y)  
 $r_{xx}$  แทน ค่าความเที่ยงของแบบสอบที่เป็นตัวทำนาย  
 $r_{yy}$  แทน ค่าความเที่ยงของเกณฑ์

ดังนั้นจะเห็นได้ว่าค่าความเที่ยงของเกณฑ์มีอิทธิพลต่อค่าความตรงร่วมสมัยพอ ๆ กับค่าความเที่ยงของตัวทำนาย

3. ความเป็นอิสระจากอคติหรือมลทิน (Freedom from Bias or

---

<sup>1</sup>William A. Lehren and Irvin J. Lehman, Measurement and Evaluation in Education and Psychology, pp. 113-144.

Contamination) คือไม่ให้คะแนนเกณฑ์ได้รับอิทธิพลจากแบบสอบที่นำมาใช้เป็นตัวทำนาย เช่น ครูนำแบบสอบความถนัดทางวิชาการมาทดสอบเด็กคนหนึ่ง ถ้าครูพบว่าเด็กคนนี้ได้คะแนนความถนัดต่ำ ก็จะประเมินผลการเรียนให้เด็กผู้นั้นต่ำไปด้วย นั่นก็คือการรู้คะแนนตัวทำนายมีอิทธิพลต่อคะแนนเกณฑ์ อิทธิพลเช่นนี้จะมีผลทำให้สัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์สูงขึ้นโดยการเสแสร้ง ปัญหานี้อาจแก้ไขได้โดยไม่ให้ผู้มีหน้าที่ตีค่าเกณฑ์ (rater) รู้คะแนนสอบที่เป็นตัวทำนาย

### องค์ประกอบที่มีอิทธิพลต่อความตรง



องค์ประกอบต่าง ๆ ที่มีอิทธิพลต่อความตรงคือ<sup>1</sup>

1. องค์ประกอบภายในแบบสอบได้แก่

1.1 คำชี้แจง ถ้าคำชี้แจงไม่ชัดเจน อาจทำให้นักเรียนทำผิดพลาด

1.2 การใช้คำและประโยคที่ยากเกินไป ทำให้นักเรียนเกิดความสับสน

ผลจะกลายเป็นนักเรียนที่ทำข้อสอบได้ถูกต้อง คือผู้ที่มีความสามารถในการอ่านสูง แบบสอบนี้จึงกลายเป็นแบบสอบวัดความเข้าใจในการอ่านแทนที่วัดพฤติกรรมที่ต้องการ

1.3 ระดับความยากของข้อสอบไม่เหมาะสม ข้อสอบที่ง่ายเกินไป หรือยากเกินไป ไม่สามารถจำแนกนักเรียนได้

1.4 การสร้างข้อสอบไม่รัดกุม ข้อสอบบางข้อเป็นการชี้แนะคำตอบให้นักเรียน

1.5 ภาษาที่กำกวม ทำให้นักเรียนตีความผิด

1.6 แบบสอบที่ใช้ไม่เหมาะสมกับพฤติกรรมที่ต้องการวัด

1.7 แบบสอบสั้นเกินไป ตามได้ไม่ครอบคลุมวัตถุประสงค์ของการวัด

1.8 การจัดเรียงข้อสอบไม่เหมาะสม ข้อสอบควรเรียงจากง่ายไปหายาก

<sup>1</sup>Norman E. Gronlund, Measurement and Evaluation in Teaching, pp. 98-102.



ถ้าเรียงจากยากก่อนนักเรียนอาจใช้เวลาศึกษานาน จนไม่มีเวลาทำข้อง่าย ๆ ที่อยู่ตอนท้าย ๆ ของแบบสอบ จึงไม่สามารถวัดความสามารถที่แท้จริง

1.9 คำตอบที่ถูกเรียงอย่างเป็นระบบ ทำให้นักเรียนเดาคำตอบได้

นอกจากนั้นค่าความเที่ยงของแบบสอบก็เป็นสิ่งที่จะกำหนดค่าความตรงของแบบสอบจากสมการ<sup>1</sup>

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

เมื่อ  $r_{xy}$  แทน ค่าความตรงของแบบสอบ

$r_{xx}$  แทน ค่าความเที่ยงของแบบสอบ

จะเห็นว่าค่าความตรงมีได้ไม่เกินรากกำลังที่สองของค่าความเที่ยงของแบบสอบ นั่นคือค่าความตรงของแบบสอบจะถูกจำกัดโดยค่าความเที่ยงของแบบสอบ

2. การบริหารแบบสอบและการให้คะแนน การบริหารแบบสอบและการให้คะแนนมีผลทำให้ความตรงเปลี่ยนแปลง เช่น การไม่ดำเนินการตามคำชี้แจง การให้คะแนนที่ไม่เป็นปรนัย การให้ความช่วยเหลือเด็กในการตอบ จะทำให้ค่าความตรงลดต่ำลงได้

3. องค์ประกอบในตัวนักเรียน ซึ่งได้แก่ ความวิตกกังวล ความเครียด จนไม่สามารถทำแบบสอบได้เต็มความสามารถ. นอกจากนั้นการ เดาคำตอบของนักเรียนก็ทำให้คะแนนที่ได้ไม่ตรงกับความสามารถที่แท้จริงของนักเรียน ทำให้ความตรงของแบบสอบเปลี่ยนแปลง

4. ลักษณะของกลุ่มและเกณฑ์ที่ใช้ โดยทั่วไปแล้ว ค่าความตรงของแบบสอบเหมาะสมสำหรับอ้างอิงเฉพาะกลุ่ม เพราะเนื่องจากสิ่งที่มีแบบสอบวัดนั้นมักได้รับอิทธิพลของตัวแปรต่าง ๆ เช่น อายุ เพศ ระดับความสามารถ ภูมิหลังของการศึกษา และวัฒนธรรมของผู้ตอบ ดังนั้นในการพิจารณาค่าความตรงของแบบสอบที่เป็นมาตรฐานควรคำนึงถึงลักษณะเฉพาะของ

---

<sup>1</sup>Frederick G. Brown, Principle of Education and Psychological Testing. p. 99.

กลุ่มที่ใช้หาความตรงนั้นด้วย นอกจากนั้นต้องพิจารณาเกณฑ์ที่ใช้หาความตรงของแบบสอบด้วย

นอกจากนี้ระยะเวลาระหว่างการวัดตัวทำนายและเกณฑ์ จำนวนข้อในแบบสอบทำนาย ยังมีอิทธิพลต่อค่าความตรงเชิงทำนายด้วย เพราะในค่านเวลาการเว้นระยะเวลากการวัดตัวทำนาย และเกณฑ์ทางกันมากก็ย่อมเปิดโอกาสให้ตัวแปรอื่น ๆ เข้ามามีอิทธิพลต่อตัวแปรทั้งสองมากขึ้น เป็นผลให้ค่าสหสัมพันธ์ที่ได้อ่า แต่ถ้าว้นระยะเวลากการวัดตัวแปรทั้งสองให้สั้นเข้ามาก็จะทำให้ค่าสหสัมพันธ์ที่คำนวณได้สูงขึ้น ซึ่งก็อาจจะเกิดจากอิทธิพลของการจำ (Memory effect) ก็ได้ ส่วนในค่านจำนวนข้อในแบบสอบ การเพิ่มจำนวนข้อในแบบสอบย่อมทำให้ค่าความเที่ยงและค่าความตรงเพิ่มขึ้น เพราะว่าค่าความตรงเพิ่มขึ้นสูงสุด (Maximum) เมื่อเท่ากับรากที่สองของค่าความเที่ยง อย่างไรก็ตามก็จำนวนข้อสอบที่เพิ่มเข้าไปนั้นจะต้องมีลักษณะเช่นเดียวกับข้อสอบที่มีอยู่เดิม<sup>1</sup> นั้นแสดงว่า ระยะเวลาระหว่างการวัดตัวทำนายกับเกณฑ์ และจำนวนข้อในแบบสอบทำนาย มีอิทธิพลต่อค่าความตรงเชิงทำนาย

## ตอนที่ 2 วรรณคดี และงานวิจัยที่เกี่ยวข้อง

ในเรื่องการเรียงลำดับข้อสอบนั้น ทราเวอร์ส (Travers) ได้เสนอแนะวิธีเรียงลำดับข้อสอบไว้ 4 วิธี<sup>2</sup> คือ

1. การเรียงลำดับตามความยากของข้อสอบ (Arrangement in Order of Difficulty) วิธีนี้เป็นการเรียงลำดับข้อสอบจากข้อง่ายไปหาข้อยาก ประโยชน์ส่วนใหญ่ของการเรียงวิธีนี้ก็คือนักเรียนจะได้พบข้อสอบข้อง่าย ๆ ก่อน ทำให้ไม่เกิดความ

<sup>1</sup>อนันต์ ศรีโสภณ, ทฤษฎีการวัดและการทดสอบ หน้า 63.

<sup>2</sup>Robert H.W. Travers, How to Make Achievement Test, (New York: Odyssey Press, 1950), pp.127-128.

เห็นภัยหนายมากนักเมื่อได้พบข้อยาก ๆ ต่อมา แต่การเรียงวิธีนี้มีข้อเสีย 2 ประการคือ ประการแรกการเรียงข้อสอบจากง่ายไปยากนั้นทำให้ข้อสอบที่มีเนื้อหาเดียวกัน ต้องกระจายกันออกไปตามลำดับความยากง่าย ซึ่งจะมีผลทำให้ความคิดของนักเรียนเปลี่ยนแปลงอยู่เสมอ ประการที่สอง การเรียงข้อสอบวิธีนี้ ถ้านักเรียนพบข้อที่ทำไม่ได้ในตอนแรก ๆ ก็จะทำให้เกิดความท้อใจที่จะทำข้อสอบข้อต่อไป เนื่องจากนักเรียนทราบว่าข้อสอบจะยิ่งยากขึ้นไปอีก

2. การเรียงลำดับความยากแบบหมุนเวียน (Arrangement in Cyclic Order of Difficulty) เป็นการเรียงแบบเรียงจากง่ายไปหาข้อยากในกลุ่มหมุนเวียนสลับกันไปโดยไม่คำนึงถึงเนื้อหา ตัวอย่างเช่น มีข้อสอบ 40 ข้อ แบ่งเป็น 4 กลุ่ม ในแต่ละกลุ่มก็เรียงจากง่ายไปยาก การเรียงแบบนี้มีข้อดีคือเป็นการกระตุ้นให้เด็กอ่านข้อสอบให้หมดทุกข้อ เพราะนักเรียนรู้ว่าถ้าทำข้อสอบไประยะหนึ่งแล้วก็จะพบข้อง่าย ๆ อีก แต่ก็มีจุดอ่อนที่นักเรียนจะต้องเปลี่ยนแปลงความคิดในการแก้ปัญหาเร็วเกินไป

3. การเรียงลำดับตามกลุ่มเนื้อหาวิชา (Arrangement According to Subject Matter Area) การเรียงแบบนี้เป็นการเรียงโดยรวมข้อสอบที่มีเนื้อหาเดียวกันเข้าไว้ด้วยกัน แล้วจัดเรียงข้อสอบในแต่ละเนื้อหาวิชาจากง่ายไปยาก ซึ่งสามารถช่วยให้นักเรียนได้ใช้ความคิดอย่างแน่วแน่ในการแก้ปัญหาในหนึ่งเนื้อหา ก่อนที่จะเปลี่ยนไปคิดในเนื้อหาอื่น ๆ

4. การเรียงลำดับตามจุดมุ่งหมายของการวัด (Arrangement According to the Goals Measured) ในการเรียงข้อสอบ ผู้จัดทำข้อสอบบางคนนิยมเรียงลำดับข้อ โดยรวมพฤติกรรมที่ต้องการจะวัดประเภทเดียวกัน เช่น ในแบบสอบโค-ออปเรทีฟ (Co-Operative Test) การเรียงลำดับข้อที่วัดเกี่ยวกับคำศัพท์ และสังกัป (Term and Concepts) เข้าไว้ในกลุ่มเดียวกัน ส่วนการวัดในด้านความเข้าใจนั้นก็แยกไว้เป็นอีกพวกหนึ่ง การเรียงลำดับข้อสอบวิธีนี้มีประโยชน์สำหรับครูในการตรวจสอบการวัดว่าตรงจุดมุ่งหมายที่ต้องการหรือไม่ แต่มีข้อเสีย คือไม่สามารถจะรวมเนื้อหาเดียวกัน หรือที่คล้ายกันเข้าไว้ด้วยกันได้

ลิงควิสต์ (Lindquist)<sup>1</sup> ก็มีความเห็นในการเรียงข้อสอบคล้ายคลึงกับข้อเสนอแนะของทราเวอร์ส โดยได้เสนอแนะวิธีจัดกลุ่มข้อสอบเข้าไว้ด้วยกันว่าถ้าข้อสอบเหล่านั้นอยู่ในเนื้อหาเดียวกัน และมีความยากง่ายพอ ๆ กัน แล้วก็อาจสลับข้อกันได้ ไม่จำเป็นต้องเรียงลำดับกันตามความยากง่าย แต่ถ้าวางข้อสอบเหล่านั้นอยู่ในเนื้อหาที่ต่างกัน และให้คะแนนแต่ละข้อเท่ากัน ก็จะต้องแยกกลุ่มข้อสอบออกเป็นฉบับย่อย ๆ (Subtest) และแยกเวลาในการทดสอบในแต่ละฉบับด้วย การเรียงลำดับข้อสอบในแต่ละฉบับนั้น อาจขึ้นอยู่กับความสัมพันธ์ของเนื้อหาวิชาคือ อาจให้สอบเนื้อหาทั่ว ๆ ไปก่อน แล้วจึงค่อยสอบเนื้อหาที่จำเพาะเจาะจง

แตซวาล<sup>2</sup> กลับมีความเห็นว่าสำหรับข้อสอบในแต่ละฉบับ ควรเรียงคำถามเรียงจากง่ายไปหายากเสมอ ไม่ควรเรียงตามความยากง่ายของบทเรียน และสแตนเลย์ กับ ฮอปกินส์ (Stanley and Hopkins)<sup>3</sup> ก็มีความเห็นตรงกันว่า การเรียงลำดับข้อสอบเป็นเรื่องสำคัญ ควรจะเริ่มต้นจากข้อที่ง่ายที่สุดไปหาข้อที่ยากที่สุด เพราะถ้านักเรียนพบข้อยาก ๆ ไปก่อนแรกแล้ว อาจจะทำให้เกิดความท้อใจ โดยเฉพาะในกลุ่มนักเรียนที่มีความสามารถในการเรียนระดับปานกลางและระดับต่ำ

สำหรับงานวิจัยที่เกี่ยวข้องกับเรื่องนี้ได้มีผู้ศึกษาไว้ดังนี้

แมคนิคอลล (Mac Nical)<sup>4</sup> ได้เปรียบเทียบลักษณะการเรียงลำดับข้อสอบของ

<sup>1</sup>Everet F. Lindquist, Educational Measurement (Washington: American Council on Education, 1966), p.179.

<sup>2</sup>ซวาล แพร์ติกุล, เทคนิคการวัดผล (พระนคร : วัฒนาพานิช, 2516) หน้า 6.

<sup>3</sup>Julian C. Stanley and Kenneth D. Hopkins, Educational and Psychological Evaluation, p.191.

<sup>4</sup>Katherin Mac Nical, "Varying Order of Item Difficulty in an Unspeed Verbal Test," Research Bulletin (Educational Testing Service, 1956), pp.1-21.

แบบสอบ ซึ่งแต่เดิมเรียงจากง่ายไปหายาก ไปเป็นการเรียงแบบสุ่ม (A Random Arrangement) ก็มีการเรียงจากข้อยากไปหาข้อง่าย ซึ่งปรากฏผลว่า การเรียงลำดับข้อสอบจากยากไปหาง่าย จะทำให้ข้อสอบยากขึ้นกว่าเดิมอย่างมีนัยสำคัญทางสถิติ

คลอสเนอร์ และเกลแมน (Klosner and Gellman)<sup>1</sup> ได้ศึกษาเปรียบเทียบการเรียงลำดับข้อสอบ 3 แบบคือ เรียงตามหัวข้อเรื่อง เรียงจากง่ายไปยาก และเรียงจากง่ายไปยากในแต่ละหัวข้อเรื่อง กลุ่มตัวอย่างเป็นนิสิตที่เรียนวิชาวัดผลเบื้องต้น ได้ผลปรากฏว่า โดยทั่ว ๆ ไป การเรียงลำดับข้อสอบทั้ง 3 วิธี ไม่มีผลต่อการสอบของนักเรียน แต่สำหรับในกลุ่มนักเรียนที่มีความสามารถต่ำแล้ว การเรียงข้อคำถามของแบบสอบมีผลทำให้ผลการสอบเปลี่ยนไปควย และค่าความเที่ยงของแบบสอบที่เรียงลำดับต่างกันมีค่าแตกต่างกัน โดยการเรียงจากง่ายไปยากในแต่ละหัวข้อเรื่องมีค่าความเที่ยงต่ำกว่าวิธีอื่น ๆ

ฟแลงเกอร์, เมลตัน และไมเออร์ส (Flanher Felton and Myers)<sup>2</sup> ได้ศึกษาการเรียงลำดับข้อสอบ 4 วิธี โดยใช้แบบสอบ ซี.อี.อี.บี. สกอลราสติก แอปติจูด (CEEB Scholastic Aptitude Test (SAT)) ด้านภาษาและคณิตศาสตร์ ซึ่งเดิมจัดแบ่งข้อสอบไว้เป็นตอน ๆ ตอนหนึ่งมี 5 ข้อ นำมาจัดเรียงใหม่เป็น 4 แบบคือ

1. เรียงตามแบบเดิมคือเรียงเป็นตอน ๆ จากง่ายไปยาก ตอนหนึ่งมี 5 ข้อ เรียงจากง่ายไปหายากในแต่ละตอน

<sup>1</sup>Naomi Certner Klosner and Estelle Klithick Gullmann, "The Effect of Item Arrangement on Classroom Test Performance : Implication for Content Validity," Educational and Psychological Measurement, 33 (1973), pp.413-418.

<sup>2</sup>Ned A. Flanher, Arthur W. Melton and Louise Kifer Myers. "A Study of Effect of Item Rearrangement," Research Bulletin, (Educational Testing Service, 1966), pp.1-21.

2. เรียงภายในแต่ละตอนใหม่ เป็นแบบสุ่ม ส่วนการเรียงระหว่างตอนเหมือนมาตรฐานเดิม

3. เรียงระหว่างตอนใหม่ เป็นแบบสุ่ม ส่วนการเรียงภายในแต่ละตอนเหมือนมาตรฐานเดิม

4. จัดเรียงใหม่ทั้งภายในแต่ละตอนและระหว่างตอน

โดยทดลองกับนักเรียนระดับวิทยาลัย ผลปรากฏว่า การเรียงด้วยวิธีต่างกันมีผลทำให้คะแนนจากแบบสอบถามภาษาต่างกันโดยวิธีเรียงแบบที่ 1 ให้คะแนนเฉลี่ยสูงกว่าแบบอื่น แต่ไม่มีผลต่อคะแนนจากแบบสอบถามคณิตศาสตร์ แสดงว่าการจัดเรียงข้อสอบต่างวิธีกันจะทำให้ข้อสอบยากและง่ายแตกต่างกันในบางวิชา ซึ่งสอดคล้องกับผลการวิจัยของคลิงเก (Kleinke)<sup>1</sup> ซึ่งทดลองกับเด็กเกรด 4 ที่พบว่า การเรียงลำดับข้อสอบจากง่ายไปยากมีผลทำให้คะแนนสอบของนักเรียนสูงกว่าการเรียงข้อสอบแบบสุ่ม

ส่วนการวิจัยในประเทศไทยก็ได้มีผู้วิจัยไว้บ้างดังนี้

ศศิธร สุวรรณสุข<sup>2</sup> ได้ศึกษาการเรียงลำดับข้อสอบโดยเรียงจากข้อง่ายไปหาข้อยาก เรียงจากข้อยากไปหาข้อง่าย เรียงตามเนื้อหาวิชา และเรียงจากข้อง่ายไปยากในแต่ละหัวเรื่อง โดยใช้แบบสอบถามคณิตศาสตร์ กลุ่มตัวอย่างเป็นเด็กชั้นประถมศึกษาปีที่ 7 ผลปรากฏว่าการเรียงลำดับข้อสอบแตกต่างกันทำให้ความเที่ยงของแบบสอบถามทักษะทางคณิตศาสตร์แตกต่างกันอย่างมีนัยสำคัญโดยการเรียงแบบเรียงจากข้อง่ายไปยากในแต่ละหัวเรื่องมีค่าความ

<sup>1</sup>David J. Kleinke, "Item Order, Response Location and Examinee Sex and Handedness and Performance on a Multiple-Choice Test," Journal of Educational Research, 4 (March-April 1980) : 225-228.

<sup>2</sup>ศศิธร สุวรรณสุข, "การศึกษาค่าความเที่ยงตรง ค่าความเชื่อมั่นของแบบสอบที่เรียงลำดับคำถามต่างกันในวิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 7" (วิทยานิพนธ์การศึกษามหาบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2519), หน้า 69-72.

เพียงต่ำกว่าการเรียงแบบอื่น ๆ แต่ในแบบทดสอบวัดการแก้ปัญหาและการหาเหตุผลทาง  
 คณิตศาสตร์ การเรียงลำดับข้อต่างกันไม่มีผลทำให้ค่าความเที่ยงแตกต่างกันแต่อย่างใด และ  
 ไม่มีผลต่อค่าความตรงในทุกแบบสอบ แกลลิฟร พุ่มแก้ว<sup>1</sup> ซึ่งศึกษาผลของวิธีเรียงข้อสอบ  
 4 วิธี คือเรียงจากช้อยไปยาก จากช้อยากไปง่าย เรียงตามเนื้อเรื่อง เรียงแบบสุ่ม  
 กลับ พบว่าการเรียงข้อสอบไม่มีผลต่อค่าความเที่ยง ความตรง แต่มีผลต่อคะแนนสอบผล  
 สัมฤทธิ์ทางการเรียน โดยแบบสอบที่มีการเรียงตามเนื้อเรื่องให้คะแนนเฉลี่ยต่ำสุด ส่วน  
 วิธีอื่น ๆ ให้ค่าเฉลี่ยไม่แตกต่างกัน ซึ่งทั้งขัดแย้งและสอดคล้องกับการวิจัยของ แนน้อย  
 เพชรรัตน์<sup>2</sup> ที่พบว่าการเรียงลำดับข้อสอบในแบบสอบขบวนการเรียนรู้ทางวิทยาศาสตร์ มี  
 ผลต่อค่าความเที่ยง โดยการเรียงแบบสุ่มให้ค่าความเที่ยงสูงกว่าการเรียงแบบอื่น แต่ไม่  
 มีผลต่อค่าความเที่ยง ในแบบสอบวิชาวิทยาศาสตร์ แต่พบว่าไม่ว่าจะเรียงแบบใดก็ไม่มีผล  
 ต่อคะแนนสอบของเด็ก

จากเอกสารและงานวิจัยเกี่ยวกับการเรียงลำดับข้อสอบดังกล่าวมาแล้วจะเห็น  
 ว่าผลของการศึกษายังแยกเป็น 2 ฝ่าย ฝ่ายหนึ่งพบว่าการเรียงลำดับข้อสอบมีผลต่อค่า  
 ความเที่ยง ความตรง และคะแนนสอบ ส่วนอีกฝ่ายหนึ่งพบว่าไม่มีผลต่อค่าความเที่ยง  
 ความตรง และคะแนนสอบ

ส่วนผลของการให้ผู้สอบรู้ค่าความยากนั้น ได้มีผู้ทำการศึกษามาไว้แล้วที่คณะมี 2

<sup>1</sup>อิริฟร พุ่มแก้ว, "ผลของการเรียงลำดับข้อทดสอบต่อค่าความเที่ยง ความตรง  
 ของแบบทดสอบ และผลสัมฤทธิ์ทางการเรียน" (วิทยานิพนธ์ครุศาสตรมหาบัณฑิต ภาควิชา  
 ศึกษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2522), หน้า 37-40.

<sup>2</sup>แนน้อย เพชรรัตน์, "ผลของการเรียงลำดับข้อคำถามด้วยวิธีต่าง ๆ กันที่มี  
 ต่อคุณภาพของแบบสอบวิชาวิทยาศาสตร์ และทักษะขบวนการทางวิทยาศาสตร์" หน้า 36-40.

ทาน คือ ฮัค (Huck)<sup>1</sup> ได้ทำการศึกษาโดยใช้กลุ่มตัวอย่างเป็นนิสิตบัณฑิตศึกษาระดับปีที่ 1 คณะศึกษาศาสตร์ ที่เรียนวิชาสถิติ แบ่งกลุ่มตัวอย่างเป็น 2 กลุ่ม กลุ่มหนึ่งบอกค่าความยากของข้อสอบ อีกกลุ่มหนึ่งไม่บอกค่าความยากของข้อสอบ ผลปรากฏว่ากลุ่มที่ใ้รู้ค่าความยากของข้อสอบได้คะแนนสูงกว่ากลุ่มที่ไม่รู้ค่าความยากของข้อสอบอย่างมีนัยสำคัญทางสถิติ ต่อมา คูบิสซึน (Kubiszyn)<sup>2</sup> ก็ได้ศึกษาในเรื่องนี้เช่นเดียวกัน โดยใช้ทดลองกับนิสิตระดับปริญญาตรีชั้นปีที่ 1 ที่เรียนวิชาการวัดและประเมินผล โดยแบ่งเป็น 2 กลุ่มเช่นกัน ก็ได้ผลเช่นเดียวกัน ฮัค คือกลุ่มที่ใ้รู้ค่าความยากของข้อสอบได้คะแนนสูงกว่ากลุ่มที่ไม่รู้ค่าความยากของข้อสอบอย่างมีนัยสำคัญทางสถิติ ซึ่งจะเห็นว่า การที่ให้ผู้สอบใ้รู้ค่าความยากของข้อสอบนั้นทำให้คะแนนสอบเปลี่ยนแปลงไปซึ่งมีผลต่อค่าความยาก

---

<sup>1</sup>Schuyler W. Huck, "Test performance Under the Condition of Known Item Difficulty, Journal of Educational Measurement, pp. 53-58.

<sup>2</sup>Thomas Walter Kubiszyn, "The Effect of Knowledge of Item Difficulty, IQ and Test Anxiety on Classroom Test Performance in Undergraduate Females" Dissertation Abstracts International, p. 1362-A.