

บทที่ 6

คุณลักษณะที่ดีของแบบทดสอบ

เนื้อหา

เนื้อหาที่จะกล่าวถึงในบทนี้มีดังนี้

- 6.1 ความเชื่อมั่น (Reliability)
- 6.2 ความเที่ยงตรง (Validity)
- 6.3 ความเป็นปรนัย (Objectivity)
- 6.4 ความยากง่าย (Difficulty)
- 6.5 อำนาจจำแนก (Discrimination)
- 6.6 ความยุติธรรม (Fairness)
- 6.7 ประโยชน์ใช้สอย (Usability)

วัตถุประสงค์

เมื่อผู้อ่านได้ศึกษารายละเอียดของเนื้อหาบทที่ 6 แล้ว ควรจะมีความสามารถดังนี้

1. บอกคุณสมบัติที่ดีของแบบทดสอบได้
2. แปลความหมายค่าความเชื่อมั่นของแบบทดสอบได้
3. วิเคราะห์ความแตกต่างระหว่างความเที่ยงตรงตามสภาพและความเที่ยงตรงตามพยากรณ์ได้
5. ตรวจสอบได้ว่าแบบทดสอบฉบับใดเป็นแบบทดสอบที่มีคุณภาพดี

ความนำ

ในการสร้างแบบทดสอบนั้น ครูผู้สร้างข้อสอบทุกคนควรจะได้ศึกษาและรู้จักลักษณะที่ดีของแบบทดสอบ ทั้งนี้เพื่อนำมาใช้เป็นเกณฑ์ในการตัดสินคุณภาพข้อสอบ ทำให้สามารถตัดสินได้ว่าข้อสอบฉบับใดเป็นข้อสอบที่ดีควรนำมาใช้ ฉบับใดเป็นข้อสอบที่ไม่ดี ซึ่งในการประเมินผลแบบทดสอบดังกล่าวจะต้องประเมินโดยยึดหลักลักษณะของแบบทดสอบที่ดีเป็นสำคัญ

แบบทดสอบที่ดี จะต้องประกอบด้วยลักษณะที่สำคัญ 7 ประการ คือ

1. มีความเชื่อมั่น
2. มีความเที่ยงตรง

3. มีความเป็นปรนัย
4. มีความยากง่าย
5. มีอำนาจจำแนก
6. มีความยุติธรรม
7. มีประโยชน์ใช้สอย

6.1 ความเชื่อมั่น (Reliability)*

ความเชื่อมั่นของแบบทดสอบคือ ค่าความคงที่ของคะแนนที่ได้จากการนำแบบทดสอบฉบับหนึ่งไปทดสอบกับนักเรียนกลุ่มเดียวกัน 2 ครั้งหรือหลายครั้ง ความเชื่อมั่นของแบบทดสอบจึงเป็นเรื่องที่เกี่ยวข้องกับความคงเส้นคงวาของคะแนนที่ได้จากการทดสอบแบบทดสอบฉบับนั้นหลาย ๆ ครั้ง แบบทดสอบที่มีค่าความเชื่อมั่นสูง หมายความว่าคะแนนที่ได้จากการทดสอบแบบทดสอบฉบับนั้นเป็นที่เชื่อถือได้ ไม่ว่าผู้สอบจะสอบข้อสอบฉบับนั้นก็ครั้งก็ตาม เขาก็จะได้คะแนนเหมือนเดิมทุกครั้งที่ไป หรือถึงแม้คะแนนจะไม่คงเดิม อาจจะได้สูงหรือต่ำไปบ้างเล็กน้อย (เนื่องจากเกิดความคลาดเคลื่อนในการทดสอบ เช่น ในการสอบครั้งหลังนั้น อาจจะมีจัดสภาพห้องสอบที่ไม่เหมือนเดิม หรือผู้เข้าสอบเกิดการเรียนรู้เพิ่มขึ้น) ก็ตาม แต่ขอให้อันดับที่ในการสอบแต่ละครั้งเหมือนเดิม ก็ถือว่าแบบทดสอบฉบับนั้นมีความเชื่อมั่นสูงได้เช่นเดียวกัน ความเชื่อมั่นของแบบทดสอบใด ๆ มีค่าอยู่ระหว่าง .00 ถึง 1.00

แบบทดสอบฉบับใดก็ตาม ถ้ามีความเชื่อมั่นเป็น 1.00 หมายความว่า แบบทดสอบฉบับนี้มีความเชื่อมั่นสูงสุด นั่นคือคะแนนที่ได้จากแบบทดสอบฉบับนี้ที่เชื่อถือได้ ไว้ใจได้ไม่ว่าจะนำไปทดสอบกับผู้สอบคนเดิมก็ครั้งก็หนก็ตาม เขาก็จะได้คะแนนหรืออันดับที่เหมือนเดิมทุกที่ไป

ส่วนแบบทดสอบที่มีความเชื่อมั่นเป็น 0.00 (ศูนย์) หรือใกล้เคียงศูนย์ แสดงว่าแบบทดสอบฉบับนั้นขาดความเชื่อมั่น คะแนนที่ได้จากการทดสอบแบบทดสอบฉบับนั้นเชื่อถือไม่ได้ สอบครั้งแรกอาจจะได้คะแนนอย่างหนึ่ง แต่พอสอบครั้งต่อมากลับได้คะแนนอย่างหนึ่ง

ตามปกติแล้ว ปริมาณความเชื่อมั่นของแบบทดสอบควรจะสูงต่ำปานใด ขึ้นอยู่กับจุดมุ่งหมายของการนำแบบทดสอบไปใช้ ถ้าเราต้องการจะใช้ผลการสอบเพื่อเป็นการทบทวนเนื้อหาวิชาที่ครูสอน เราก็อาจจะใช้แบบทดสอบที่ยังไม่รู้ค่าความเชื่อมั่นก็ได้ เพราะการตัดสินใจของเราขึ้นอยู่กับผลการสอบของผู้เข้าสอบทั้งกลุ่ม เราจะไม่เพ่งเล็งผลการสอบของผู้เข้าสอบแต่ละคนมากนักถึงจะตัดสินใจผิดพลาดไปบ้างก็คงไม่เกิดผลเสียหายร้ายแรงเท่าใด ผลเสียก็จะมี

* อ่านรายละเอียดเกี่ยวกับวิธีหาและสิ่งที่มีอิทธิพลต่อค่าความเชื่อมั่นในบทถัดไป

แต่เพียงว่าผู้เรียนอาจจะขาดเนื้อหาบางตอนที่จำเป็นต้องทำการทบทวนไป แต่ถ้าเราจะใช้ผลการสอบไปจัดอันดับความสามารถของนักเรียนในกลุ่ม ผลจากการวัดก็จะต้องมีความเชื่อมั่นสูงสุด โดยทั่ว ๆ ไปแล้ว แบบทดสอบที่ครูสร้างขึ้นมาเพื่อใช้ในห้องเรียนควรมีค่าความเชื่อมั่นตั้งแต่ .60 ขึ้นไป แต่สำหรับแบบทดสอบมาตรฐานนั้น จะต้องมีความเชื่อมั่นตั้งแต่ .80 ขึ้นไป จนถึง 1.00 จึงจะจัดว่าเป็นแบบทดสอบที่ดี

6.2 ความเที่ยงตรง (Validity)*

แบบทดสอบที่ดีต้องมีความเที่ยงตรง แบบทดสอบที่มีความเที่ยงตรงสูง คือ แบบทดสอบที่สามารถวัดได้ตรงตามวัตถุประสงค์ที่ผู้ออกข้อสอบต้องการจะวัด ความเที่ยงตรงของแบบทดสอบแบ่งออกเป็นชนิดใหญ่ ๆ ได้ 3 ชนิด คือ

- 6.2.1 ความเที่ยงตรงตามเนื้อหา (Content Validity)
- 6.2.2 ความเที่ยงตรงตามเกณฑ์ (Criterion - related Validity)
- 6.2.3 ความเที่ยงตรงตามโครงสร้าง (Construct Validity)

6.2.1 ความเที่ยงตรงตามเนื้อหา

การหาความเที่ยงตรงตามเนื้อหา มีจุดประสงค์เพื่อจะศึกษาดูว่าแบบทดสอบนั้นสามารถวัดเนื้อหาวิชาตามที่กำหนดไว้ในหลักสูตรมากน้อยเพียงใด ดังนั้น โดยความหมายทั่ว ๆ ไปแล้ว ความเที่ยงตรงตามเนื้อหาหมายถึงความสามารถของแบบทดสอบฉบับหนึ่งที่จะมีข้อความคำถามเป็นตัวแทนที่แท้จริงของเนื้อหาวิชาซึ่งกำหนดไว้ในหลักสูตร แบบทดสอบที่ควรมีความเที่ยงตรงตามเนื้อหาวิชาสูงนั้น ควรเป็นแบบทดสอบประเภทวัดผลสัมฤทธิ์ทางการเรียน

6.2.2 ความเที่ยงตรงตามเกณฑ์

การหาความเที่ยงตรงตามเกณฑ์ของแบบทดสอบ มีจุดประสงค์ที่จะดูว่าคะแนนที่ได้จากการทดสอบแบบทดสอบฉบับนั้นสามารถนำมาใช้พยากรณ์ผลการเรียนภายภาคหน้า หรือนำมาคาดคะเนผลการเรียนในปัจจุบันนี้ได้ดีเพียงใด ความเที่ยงตรงตามเกณฑ์มีอยู่ 2 ชนิด คือ **ความเที่ยงตรงตามพยากรณ์ (Predictive Validity)** กับ **ความเที่ยงตรงตามสภาพ (Concurrent Validity)** การหาความเที่ยงตรงทั้ง 2 ชนิดนี้ เป็นการหาความสัมพันธ์ระหว่างคะแนนของแบบทดสอบที่จะหาความเที่ยงตรงกับตัวเกณฑ์ (Criteria) ในการหาความเที่ยงตรงตามพยากรณ์นั้น ตัวเกณฑ์ที่จะนำมาใช้เปรียบเทียบคือ คะแนนที่ได้จากผลการเรียน ซึ่งจะต้องติดตามไปในอนาคต ส่วนการหาความเที่ยงตรงตามสภาพนั้น ตัวเกณฑ์จะได้จากการที่ครูจัดอันดับความ

*ดูรายละเอียดในบทถัดไป

สามารถที่นักเรียนแสดงออกในห้องเรียนให้ครูสังเกตเห็นได้ ถ้าข้อสอบมีความเที่ยงตรงตามสภาพสูงก็แสดงว่าคะแนนที่ได้จากแบบทดสอบนั้นสอดคล้องไปในทางเดียวกันกับอันดับที่ครูจัด นั่นคือ ถ้าเขาได้คะแนนจากแบบทดสอบสูง ก็จะต้องได้อันดับดีด้วย โดยปกติแล้วความเที่ยงตรงทั้ง 2 ชนิดนี้ มักเรียกรวมเป็นชื่อเดียวกันว่า Criterion - Related Validity ทั้งนี้เนื่องจากความเที่ยงตรงทั้ง 2 ชนิดนี้มีหลักการและความหมายเป็นแบบเดียวกัน แต่จะต่างกันตรงระยะเวลาที่จะได้ตัวเกณฑ์มาเปรียบเทียบ

6.2.3 ความเที่ยงตรงตามโครงสร้าง

ความเที่ยงตรง 2 แบบที่กล่าวมาข้างต้น เกี่ยวข้องโดยตรงกับการใช้ ผลจากการทดสอบ ความเที่ยงตรงตามเนื้อหาจะบอกให้ทราบว่าผลที่ได้จากการทดสอบจะเป็นตัวแทนความสัมฤทธิ์ผลในแต่ละวิชาได้ดีเพียงไร ความเที่ยงตรงตามเกณฑ์จะบ่งให้ทราบว่า จะพยากรณ์หรือคาดคะเน ผลการเรียนรู้ได้ถูกต้องเพียงไร แต่ความเที่ยงตรงตามโครงสร้างจะบ่งถึงคุณลักษณะบางประการทางจิตวิทยา เช่น สมรรถภาพสมองทางด้านต่าง ๆ เช่น สมรรถภาพทางด้านความมีเหตุผล ความสามารถในการทำทฤษฎีไปใช้ ความสามารถในการวิเคราะห์ หรือพฤติกรรมอื่น ๆ ที่ต้องการให้นักเรียนมีหลังจากที่ครูสอนวิชานั้น ๆ ไปแล้ว ดังนั้น ข้อสอบวัดผลสัมฤทธิ์ที่มีความเที่ยงตรงตามโครงสร้าง จะเป็นข้อสอบที่วัดสมรรถภาพสมองด้านต่าง ๆ ที่กำหนดไว้ในตารางวิเคราะห์หลักสูตรด้านพฤติกรรม นั้นเอง

จึงอาจสรุปได้ว่า ความเที่ยงตรงของแบบทดสอบหมายถึง ความเที่ยงตรงในการวัดหรือหมายถึงปริมาณของการวัด ว่าใกล้เคียงกับจุดมุ่งหมายที่ต้องการจะวัดหรือไม่ ดังนั้น แบบทดสอบที่มีความเที่ยงตรง คือ แบบทดสอบที่สามารถวัดได้ตรงตามสิ่งที่ต้องการจะวัด แบบทดสอบจะมีความเที่ยงตรงได้นั้น สิ่งสำคัญประการแรกคือ แบบทดสอบนั้นต้องมีความเชื่อมั่น แต่แบบทดสอบที่มีความเชื่อมั่นไม่จำเป็นต้องมีความเที่ยงตรงตามไปด้วย ตัวอย่างเช่น แบบทดสอบฉบับหนึ่งผู้สร้างต้องการวัดความรู้เกี่ยวกับวิชาชีววิทยา แต่ไปออกคำถามเกี่ยวกับเนื้อหาวิชาเคมีส่วนใหญ่ แสดงว่าแบบทดสอบฉบับนี้ไม่สามารถวัดในสิ่งที่ต้องการจะวัดจึงขาดความเที่ยงตรง แต่ไม่ว่าจะนำแบบทดสอบฉบับนี้ไปสอบกับนายวินิจ สักกี่ครั้งก็ตามเขาก็จะได้คะแนนเท่าเดิมทุกที ก็แสดงว่าแบบทดสอบฉบับนี้มีความเชื่อมั่น นี่เป็นตัวอย่างที่แสดงให้เห็นว่าแบบทดสอบที่มีความเชื่อมั่นไม่จำเป็นต้องมีความเที่ยงตรงเสมอไป

การหาความเที่ยงตรงของแบบทดสอบนั้น จะต้องมีเกณฑ์ (Criteria) สำหรับตัดสินแบบทดสอบที่มีความเที่ยงตรงสูงนั้นอาจเกิดจาก

- 1) แบบทดสอบที่มีคุณภาพดี สัมพันธ์กับเกณฑ์ที่ดี หรือ
- 2) แบบทดสอบที่มีคุณภาพเลว สัมพันธ์กับเกณฑ์ที่ไม่ดี แต่ถ้าอย่างใดอย่างหนึ่งมีคุณภาพตรงกันข้ามแล้ว แบบทดสอบก็จะมีค่าเที่ยงตรงต่ำ ดังนั้นในการพิจารณาค่าความเที่ยงตรงของแบบทดสอบ จึงต้องพิจารณาค่าคุณภาพของเกณฑ์ด้วย

6.3 ความเป็นปรนัย (Objectivity)

ก่อนอื่นต้องขอทำความเข้าใจเสียก่อนว่า แบบทดสอบที่มีความเป็นปรนัยนั้นหมายถึงแบบทดสอบชนิดใด เพราะเรื่องนี้ครู - อาจารย์ส่วนใหญ่ยังเข้าใจสับสนกันอยู่มาก แบบทดสอบที่มีความเป็นปรนัยนั้นจะต้องมีคุณสมบัติ 3 อย่าง คือ

- ก. ผู้สอบอ่านใจก็แล้วต้องเข้าใจคำถามตรงกัน
- ข. การตรวจกระดาษคำตอบนั้น ไม่ว่าจะให้ใครตรวจก็ต้องให้คะแนนเท่ากันทุกคน
- ค. ต้องสามารถนำคะแนนมาแปลความหมายเป็นรูปเดียวแบบกันได้

แบบทดสอบใดก็ตามถ้ามีคุณสมบัติครบ 3 ประการนี้จึงจะเรียกว่าเป็นแบบทดสอบที่มีความเป็นปรนัย โปรดอย่านึกว่าแบบทดสอบที่มีความเป็นปรนัยหมายถึง ข้อสอบแบบถูก-ผิด จับคู่ เติมคำ หรือเลือกตอบเลยทีเดียว เพราะนั่นเป็นแต่เพียงวิธีการ หรือขบวนการเขียนข้อคำถาม ซึ่งจะนำข้อสอบไปสู่ความเป็นปรนัยเท่านั้น

จะเห็นว่าข้อสอบแบบเขียนบรรยายนั้นมีโอกาสที่จะเป็นปรนัยได้ยาก เพราะมักจะขาดคุณสมบัติในข้อ ข. คือถ้าให้ครู 2 คนตรวจ แล้วทิ้งไว้ซักพักหนึ่งแล้วมาตรวจใหม่ก็มีแนวโน้มว่าจะให้คะแนนเปลี่ยนแปลงไปจากการตรวจครั้งแรก แม้ในคำถามชนิดจำกัดคำตอบแบบถูก-ผิด หรือเติมคำ ถ้าสร้างไม่ดีก็อาจจะขาดความเป็นปรนัยได้ เช่น คำถามกำกวมอ่านแล้วไม่เข้าใจว่าผู้ถามต้องการให้ตอบในแนวใด หรือแม้ในการแปลความหมายคะแนนที่กำหนดไว้ว่าผู้สอบจะต้องสอบได้ตั้งแต่ 50% ขึ้นไปจึงจะนับว่าสอบได้โดยไม่คำนึงถึงความยากง่ายของข้อสอบเลย ก็เรียกว่าไม่มีความเป็นปรนัย จะเห็นว่าแบบทดสอบจะมีความเป็นปรนัยนั้นไม่ใช่ของง่าย ๆ จะต้องระวังทั้งตัวคำถามและการแปลความหมายคะแนนควบคู่กันไปด้วย

6.4 ความยากง่าย (Difficulty)

แบบทดสอบที่ดีต้องมีความยากง่ายพอเหมาะ เรื่องนี้มีความสำคัญมาก เพราะคะแนนที่นักเรียนสอบได้นั้นขึ้นอยู่กับความยากง่ายของข้อสอบเป็นสำคัญ มิได้ขึ้นอยู่กับจำนวนข้อคำถามหรือจำนวนเวลาที่กำหนดให้ทำการสอบมากนัก ลงข้อสอบยากแล้วต่อให้มีข้อสอบเป็นร้อย ๆ พัน ๆ ข้อ นักเรียนก็ได้คะแนนต่ำอยู่นั่นเอง เพราะทำข้อสอบไม่ได้ ข้อสอบที่ดีนั้นจะต้องมีคำถาม

ที่ผู้สอบตอบถูกบ้างผิดบ้างสลับกันไป ฉะนั้น ข้อสอบที่ยากที่สุดหรือง่ายที่สุดจึงไม่มีประโยชน์ เพราะถ้าข้อสอบยากที่สุด ผู้สอบทุกคนก็จะได้คะแนนศูนย์ หรือถ้าข้อสอบง่ายที่สุด ผู้สอบก็จะได้คะแนนเต็มเท่ากันทุกคน เราเลยไม่สามารถแยกได้ว่าใครเก่งกว่ากัน

ข้อสอบที่ใช้ในการสอบรวม หรือสอบประจำภาคเรียน ควรจะมีอัตราส่วนความยากของข้อคำถามเป็นอัตราส่วน 1 : 2 : 1 คือมีข้อสอบที่ง่าย ๆ 1 ส่วน ข้อสอบที่ยาก ๆ 1 ส่วน และข้อสอบที่ไม่ง่ายไปไม่ยากไปอีก 2 ส่วน ดังนั้น ถ้าข้อสอบฉบับหนึ่งมี 100 ข้อ ก็จะมีข้อสอบง่าย ๆ ประมาณ 25 ข้อ ข้อสอบยากง่ายปานกลางประมาณ 50 ข้อ และมีข้อสอบยาก ๆ อีกประมาณ 25 ข้อ จึงจะดี ความยากง่ายของข้อคำถามเราใช้ตัว p เป็นสัญลักษณ์แทน ถ้า p เท่ากับ .50 หมายถึงข้อสอบมีความยากง่ายพอเหมาะดี แต่ถ้า p น้อยกว่า .50 แปลว่าข้อสอบยาก p มากกว่า .50 แปลว่าข้อสอบง่าย เราอาจจะท่องเป็นสูตรว่า “น้อยยากมากง่าย” ก็ได้

การเรียงข้อคำถามในแบบทดสอบนั้น โดยทั่ว ๆ ไปแล้วนิยมเรียงจากข้อง่ายไปหาข้อยาก คือเอาข้อง่ายเป็นข้อ 1 แล้วค่อย ๆ ยากขึ้นไปตามลำดับ (วิธีการหาความยากง่ายของข้อสอบและการคัดเลือกค่าความยากง่ายของข้อสอบให้ศึกษารายละเอียดในบทต่อไป) อย่างไรก็ตามก็ยังมีวิธีเรียงข้อคำถามที่มีหลายวิธีอาจจะสรุปได้เป็นวิธีใหญ่ ๆ ได้ 4 วิธีคือ

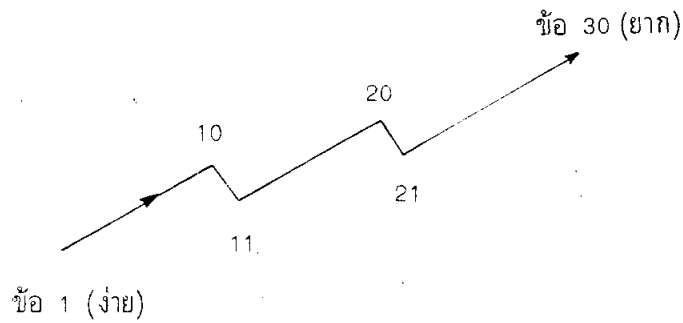
1. จัดเรียงตามลำดับความยากง่าย วิธีนี้เป็นการจัดเรียงข้อแบบธรรมดาทั่ว ๆ ไป คือเรียงจากข้อง่ายไปหายาก ประโยชน์ส่วนใหญ่ของการเรียงข้อสอบวิธีนี้ก็คือ ผู้สอบจะได้พบข้อสอบง่าย ๆ ก่อนทำให้เกิดกำลังใจ จึงไม่เกิดความเหนื่อยมากนักเมื่อมาพบข้อสอบยาก ๆ แต่การเรียงข้อสอบแบบนี้ก็มีข้อเสีย 2 ประการ คือ

ก. การเรียงข้อสอบแบบนี้ ทำให้ข้อสอบที่มีเนื้อหาเดียวกันต้องกระจายออกไปตามลำดับความยากง่าย ซึ่งมีผลทำให้ดวงความคิดของผู้เข้าสอบต้องเปลี่ยนแปลงอยู่เสมอ เช่น เดียวคิดเกี่ยวกับเลขบวก ก็ต้องเปลี่ยนมาคิดเลขคูณ แล้วไปคิดเลขลบแล้วกลับมาคิดเลขคูณอีก ดังนั้นเป็นต้น

ข. การเรียงข้อสอบแบบนี้ มีผลเสียตรงที่ถ้านักเรียนพบข้อยาก ๆ เสียแล้วในตอนต้น ๆ ก็จะทำให้เกิดความท้อใจที่จะทำข้อต่อไป เนื่องจากนักเรียนรู้ว่าข้อสอบจะยิ่งยากขึ้นไปอีก

2. จัดเรียงความยากง่ายแบบไซคลิกออร์เดอร์ (Cyclic Order)

วิธีเรียงข้อสอบแบบนี้ต่างจากแบบที่ 1 ตรงที่แทนที่จะเรียงข้อสอบจากง่ายไปหายากที่เดียวก็จะมีกการลดความยากง่าย เป็นขั้น ๆ บันได เช่น จากข้อ 1-10 เรียงจากง่ายไปหายาก พอข้อ 11-14 แทนที่จะยากขึ้นไปอีก กลับลดความยากลงมาเท่ากับข้อ 6-10 พอถึงข้อ 15 ก็จะยากกว่าข้อ 10 ดังรูป



การเรียงข้อสอบแบบนี้ดีตรงที่ จะเป็นการกระตุ้นให้ผู้ทำข้อสอบให้หมดทุกข้อ เพราะเขารู้ว่าทำข้อสอบในระยะหนึ่งก็พบข้อง่าย ๆ อีก วิธีนี้ถึงจะดี แต่ก็ยังเสียตรงที่ผู้สอบยังต้องเปลี่ยนดวงความคิดในการแก้ปัญหาเร็วเกินไป

3. จัดเรียงตามกลุ่มของเนื้อหาวิชา การเรียงข้อสอบโดยวิธีนี้เป็นการเรียงโดยรวมข้อสอบที่มีเนื้อหาเดียวกันไว้ด้วยกัน แล้วจัดเรียงข้อตามลำดับความยากง่าย ซึ่งวิธีนี้จะช่วยให้ผู้สอบใช้ความคิดในการแก้ปัญหาหนึ่งก่อน ที่จะเปลี่ยนไปคิดเนื้อหาอื่น

4. จัดเรียงตามจุดมุ่งหมายของการวัด ในการจัดเรียงข้อสอบผู้สร้างแบบทดสอบบางคนนิยมจัดเรียงข้อโดยรวมพฤติกรรมที่ต้องการวัดประเภทเดียวกันเข้าไว้ด้วยกัน เช่น รวมข้อคำถามเกี่ยวกับความรู้ในเรื่องไว้พวกหนึ่ง การเรียงข้อสอบด้วยวิธีนี้มีประโยชน์สำหรับครูมาก ในการที่จะรู้ว่าสามารถวัดได้ตรงตามจุดมุ่งหมายที่ต้องการหรือไม่ แต่มีข้อเสียคือ บางทีครูไม่สามารถออกข้อสอบได้ตรงตามพฤติกรรมที่ต้องการจะวัด

ผู้เขียนได้ทำการวิจัยเกี่ยวกับการเรียงข้อสอบ ของเด็กไทยในระดับชั้น ป.7 (เดือนใจ เศรษฐสุกก, 2511) พบว่าสำหรับนักเรียนที่เรียนดีนั้น จะทำข้อสอบได้คะแนนดีที่สุด ถ้าเขาให้สอบจากวิชาที่ง่ายไปหาวิชาที่ยากและภายในแต่ละวิชาเรียงข้อจากง่ายไปหายาก ส่วนนักเรียนกลุ่มที่มีความสามารถปานกลางนั้น จะเรียงข้อสอบแบบไหนก็ไม่สำคัญ เพราะจะได้คะแนนเท่ากันทุกแบบแต่สำหรับนักเรียนกลุ่มอ่อนนั้น ชอบทำข้อสอบประเภทฉบับหนึ่งมีหลาย ๆ วิชา วิชาละ 5-10 ข้อ และภายในแต่ละวิชานั้น เรียงจากข้อง่ายไปหายากส่วนถ้าเรียงแบบนี้ละก็เขาจะได้คะแนนดีกว่าเรียงแบบอื่น ๆ แต่ถ้าดูโดยส่วนรวมแล้ว การเรียงข้อสอบแบบให้สอบวิชาที่ง่ายไปหาวิชาที่ยาก และภายในแต่ละวิชาเรียงจากง่ายไปหายากนั้นเป็นวิธีที่ดีที่สุดสำหรับนักเรียนชั้นประถมศึกษาปีที่ 7 เนื่องจากผลจากการวิจัยส่วนใหญ่ ระบุว่า การจัดลำดับข้อสอบด้วยวิธีที่แตกต่างกัน ส่งผลให้คะแนนที่ได้จากการทดสอบแตกต่างกัน (Mollenkopf, 1950; Macnical 1956; Flaughar, Melton, and Mayers, 1968; Hambleton and Traub, 1974) ครูจึงควรให้ความสนใจและระมัดระวังว่าสำหรับเด็กชั้นประถมควรจะเรียงข้อสอบจากง่ายไปหายากจึงจะให้ผลดีที่สุด

6.5 อำนาจจำแนก (Discrimination)

แบบทดสอบที่ดีต้องมีอำนาจจำแนกสูง หมายความว่าแบบทดสอบนั้นต้องมีคุณสมบัติในการแจกแจงผู้สอบออกเป็นประเภท ๆ ตามความสามารถของเขาได้ ตั้งแต่เก่งสุดจนถึงอ่อนสุด แบบทดสอบที่มีอำนาจจำแนกสูงคือ แบบทดสอบที่เด็กเก่งเท่านั้นที่จะตอบถูก ส่วนเด็กอ่อนตอบที่ไร้มิตทุกที่ ถึงจะเดาก็ตอบไม่ถูก ข้อสอบใดสามารถเดาได้ถูกข้อสอบนั้นก็ไม่มีอำนาจจำแนก จึงเป็นข้อสอบที่ดีไม่ได้ ดังนั้นที่มีเสียงพูดกันว่าข้อสอบปรนัยไม่ดีเพราะเดาได้ เปลว ๆ บางทีก็พูดกันถึงขนาดว่าให้ลิงขีดตอบก็ได้ นั่นไม่จริง เพราะข้อสอบปรนัยที่เขียนโดยครูที่มีความรู้ทางการวัดผลนั้น ท่านไม่สามารถเดาได้ถูกเลย ข้อสอบฉบับไหนที่เดาได้ก็แปลว่าเป็นข้อสอบที่ไม่ดี ควรเก็บเข้าลิ้นชักได้แล้ว ไม่ควรเอาออกมาใช้อีก ดังนั้น อำนาจจำแนกจึงเป็นคุณสมบัติที่สำคัญอีกประการหนึ่งของแบบทดสอบ เราใช้ตัว r เป็นสัญลักษณ์แทนค่าอำนาจจำแนกของข้อสอบ ข้อสอบที่มีค่า ตั้งแต่ .20 ถึง 1.00 ขึ้นไปจึงจะถือว่าเป็นข้อสอบที่ใช้ได้ ส่วนวิธีการหาค่าอำนาจจำแนกของแบบทดสอบนั้น จะหาได้จากทวิเคราะห์ข้อสอบ ซึ่งจะได้อีกในบทต่อไป

6.6 ความยุติธรรม (Fairness)

แบบทดสอบที่มีความยุติธรรมก็คือ แบบทดสอบที่ให้ความเสมอภาคแก่ผู้เข้าสอบทุกคน ไม่เปิดโอกาสให้คนเก่งเก่งข้อสอบได้ หรือเด็กอ่อนเดาได้ คือไม่ลำเอียงสำหรับเด็กกลุ่มหนึ่งกลุ่มใดโดยเฉพาะ (เช่น เด็กอาจจะแตกต่างกันในด้านเชื้อชาติ ศาสนา หรือวัฒนธรรม) ดังนั้นแบบทดสอบที่ยุติธรรมจะต้องออกวัดให้ครอบคลุมหลักสูตร โดยออกมาก ๆ ข้อและสอบหลาย ๆ ครั้งจึงจะดี ต้องถามให้ละเอียดทุกแห่งทุกมุม ดังนั้น ข้อสอบดี ๆ ในสมัยปัจจุบันจึงนิยมใช้คำถามปรนัยแบบเลือกตอบ เพราะจะได้มีโอกาสถามได้มาก ๆ ข้อ ในช่วงระยะเวลาสอบอันจำกัด

คาถาสำหรับช่วยให้การเกิดความยุติธรรมคือ “สอบหลาย ๆ อย่าง อย่างละหลาย ๆ ที่ ที่ละมาก ๆ ข้อ”

6.7 ประโยชน์ใช้สอย (Usability)

ในการประเมินผลแบบทดสอบนั้น สิ่งสำคัญที่จะต้องคำนึงถึงอีกประการหนึ่งก็คือ แบบทดสอบนี้มีประโยชน์ในด้านใช้สอยเพียงใด แบบทดสอบที่ดีนั้นจะต้องมีประโยชน์ดังนี้ คือ

1. ง่ายต่อการดำเนินการสอบ

การดำเนินการสอบที่ถูกต้องเป็นสิ่งจำเป็นสำหรับการสอบ ดังนั้น แบบทดสอบที่ดีจะต้องง่ายต่อการดำเนินการสอบ เพื่อสนองจุดประสงค์อันนี้ ในแบบทดสอบทุกฉบับจะต้อง

มีค่าชี้แจงสำหรับดำเนินการสอบ และกำหนดเวลาที่ใช้ในการสอบให้แน่นอน เพื่อให้ผู้ดำเนินการสอบสามารถดำเนินการได้ตามนั้น การดำเนินการสอบที่ผิดพลาดจะทำให้ความเที่ยงตรงและความเชื่อมั่นของข้อสอบเปลี่ยนไป ซึ่งจะมีผลโดยตรงต่อคะแนนที่ได้จากการทดสอบ

2. ใช้เวลาพอเหมาะในการทดสอบ

เวลาที่ใช้ในการทดสอบควรกำหนดให้พอเหมาะ โดยปกติแล้วเรามีความนิยมที่จะใช้ข้อสอบสั้น ๆ แต่เนื่องจากค่าความเชื่อมั่นจะเกี่ยวข้องโดยตรงกับความยาวของแบบทดสอบ ถ้าเราพยายามจะลดเวลาในการทดสอบให้น้อยลงกว่าปกติก็จะทำให้ค่าความเชื่อมั่นของคะแนนเปลี่ยนแปลงไปด้วย ดังนั้น การกำหนดเวลาที่มากเกินไปหรือน้อยเกินไปก็จะมีผลทำให้ความเชื่อมั่นของข้อสอบเปลี่ยนแปลงไปในทางที่ไม่ดีทั้งสิ้น

3. ง่ายต่อการให้คะแนน

ปัญหาที่สำคัญประการหนึ่งของครูก็คือการตรวจให้คะแนนผลการสอบ ในสมัยก่อนนั้นครูต้องเสียเวลามากต่อการตรวจข้อสอบ และจะยิ่งยากขึ้นถ้าเป็นแบบทดสอบประเภทให้เขียนตอบแบบบรรยาย ในปัจจุบันนี้การทดสอบได้พัฒนาขึ้นเพื่อให้ง่ายต่อการให้คะแนนผลการสอบโดยยึดหลัก ดังนี้

- 1) มีแนวโน้มที่จะสร้างข้อสอบปรนัยให้เป็นข้อสอบมาตรฐาน
- 2) ปรับปรุงวิธีการให้คะแนนให้สะดวกขึ้นโดยสร้างกุญแจ (key) สำหรับคำตอบถูก
- 3) ใช้กระดาษคำตอบแยกต่างหากจากตัวข้อสอบ
- 4) ใช้เครื่องจักรตรวจ

4. ง่ายต่อการแปลผลและการนำไปใช้

ในการวิเคราะห์ขั้นสุดท้ายนั้น โปรแกรมการทดสอบจะดีหรือไม่ขึ้นอยู่กับผลของแบบทดสอบ ถ้าผู้ใช้สามารถแปลผลได้อย่างถูกต้องและนำไปใช้ได้โดยมีประสิทธิภาพแล้วก็จะเป็นการเพิ่มคุณภาพทางการศึกษาได้อีกทางหนึ่ง แต่ถ้าการแปลผลผิดพลาด และไม่สามารถนำผลจากการสอบไปใช้ให้เป็นประโยชน์ได้หรือได้แต่มีคุณค่าเล็กน้อย ก็จะก่อให้เกิดอันตรายแก่กลุ่มหรือบุคคลที่ใช้แบบทดสอบนั้น

สรุปสาระสำคัญ

ในการพิจารณาว่าแบบทดสอบฉบับหนึ่งจะเป็นแบบทดสอบที่มีคุณภาพดีเพียงใด ให้ตรวจสอบจากคุณสมบัติที่สำคัญของแบบทดสอบดังนี้

1. เชื่อมั่น เป็นข้อสอบที่สามารถให้คะแนนได้คงที่หรือไม่
2. เที่ยงตรง เป็นคำถามที่สามารถวัดสิ่งที่ต้องการวัดได้อย่างถูกต้องครบถ้วนตามความมุ่งหมายที่เราต้องการหรือไม่ และเที่ยงตรงทางด้านไหนบ้าง

3. ประนัย เป็นข้อสอบที่มีคำถามชัดเจน ให้คะแนนถูกต้อง และคะแนนนั้นสามารถนำมาแปลความหมายได้อย่างแจ่มชัดหรือไม่
4. ยากพอเหมาะ เป็นข้อสอบที่มีความยากง่ายพอเหมาะกับระดับความสามารถของผู้เรียนหรือไม่
5. อำนาจจำแนก เป็นคำถามที่สามารถแยกผู้เรียนออกเป็นกลุ่ม ๆ ตามระดับความสามารถของเขาหรือไม่
6. ยุติธรรม เป็นคำถามที่ไม่เปิดโอกาสให้เด็กฉลาดใช้ไหวพริบในการเดาได้ถูก ไม่เปิดโอกาสให้เด็กขี้เกียจตอบได้ และไม่ช่วยให้เด็กโชคดีถึงข้อสอบได้ หรือไม่
7. ประโยชน์ใช้สอย เป็นแบบทดสอบที่ง่ายต่อการดำเนินการสอบ ใช้เวลาพอเหมาะในการทดสอบ ง่ายต่อการให้คะแนนและง่ายต่อการแปลผลและการนำไปใช้หรือไม่

คำถามท้ายบทที่ 6

1. แบบทดสอบที่ดีต้องมีคุณสมบัติอย่างไรบ้าง
2. แบบทดสอบฉบับหนึ่งมีค่าความเชื่อมั่นใกล้เคียงศูนย์ หมายความว่าแบบทดสอบฉบับนี้มีคุณภาพอย่างไร
3. ความเที่ยงตรงตามสภาพ และความเที่ยงตรงตามพยากรณ์แตกต่างกันอย่างไร
4. ท่านจะมีวิธีการอย่างไร จึงจะทำให้แบบทดสอบของท่านเป็นแบบทดสอบที่มีความยุติธรรม
5. ถ้าพิจารณาในด้านประโยชน์ใช้สอย เป็นสำคัญแล้ว แบบทดสอบที่ท่าน สร้างขึ้นมาควรมีลักษณะอย่างไร

บรรณานุกรม

ชวาล แพร์ตกุล. เทคนิคการวัดผล วัฒนาพานิช 2516.

เตือนใจ เศรษฐสุตโก. การศึกษาถึงการจัดระเบียบข้อทดสอบด้วยวิธีต่างๆ ที่จะส่งผลต่อความสามารถในการสอบของนักเรียนชั้นประถมศึกษาปีที่ 7 ปรินซ์นิพนธ์ 2513.

Ebel, R.L., *Essentials of Educational Measurement*, Prentice-Hall, Inc. 1979.

Mehrens, W.A., and I.J. Lehmann, *Measurement and Evaluation in Education and Psychology*, Holt, Rinehart and Winston, 1978.