

บทที่ 8

หลักการสร้างข้อสอบปรนัย

ในการสร้างข้อสอบปรนัยนั้น ถ้าต้องการจะให้ได้ข้อสอบที่มีคุณภาพดีสมกับจะนำมาใช้ เป็นเครื่องมือตัดสินอนาคตของคนแล้วผู้สร้างข้อสอบควรจะมีความรู้เกี่ยวกับหลักสำคัญในการ สร้างข้อสอบปรนัยว่าควรจะมีขั้นตอนดำเนินการสร้างอย่างไร จะนั้นเนื้อหาของบทที่ 8 นี้ จะ กล่าวถึงรายละเอียดเกี่ยวกับขั้นตอนดำเนินการสร้างข้อสอบปรนัย ดังนี้

เนื้อหา

- 8.1 กำหนดจุดมุ่งหมายของการสร้างข้อสอบปรนัย
- 8.2 กำหนดลักษณะของข้อสอบ
- 8.3 กำหนดเนื้อหาของข้อสอบ
- 8.4 หมายกำหนดการเกี่ยวกับการสร้างข้อสอบ

วัตถุประสงค์

เมื่อผู้อ่านศึกษารายละเอียดของเนื้อหาบทที่ 8 แล้ว ควรจะมีความสามารถดังนี้

- 1. บอกขั้นตอนสำคัญในการดำเนินการสร้างข้อสอบปรนัยได้
- 2. เขียนจุดมุ่งหมายของการสร้างข้อสอบปรนัยได้
- 3. วิเคราะห์หลักสำคัญที่ใช้ในการกำหนดลักษณะของข้อสอบได้
- 4. บอกวิธีการกำหนดเนื้อหาของข้อสอบให้สอดคล้องกับจุดประสงค์ของการสอนได้
- 5. เขียนหมายกำหนดการเกี่ยวกับการสร้างข้อสอบได้

8.1 กำหนดจุดมุ่งหมายของการสร้างข้อสอบปรนัย

ในการสร้างแบบทดสอบแต่ละชุด จำเป็นต้องกำหนดเป้าหมายให้ชัดเจน เพราะแบบ ทดสอบแต่ละชุดที่สร้างขึ้นเพื่อจุดมุ่งหมายต่างกันก็หมายที่จะใช้ในงานต่างกัน ถ้านำข้อสอบที่มี จุดมุ่งหมายอย่างหนึ่งไปสอบเพื่อจุดมุ่งหมายอีกอย่างหนึ่ง ผลที่ได้มา ก็ย่อมไม่สอดคล้องกับสิ่งที่ ต้องการวัด จุดมุ่งหมายทั่วไปของการสอบแยกพิจารณาดังนี้

8.1.1 สิ่งที่ต้องการวัด เป็นการกำหนดเนื้อหาของแบบทดสอบว่าจะวัดวิชาใด ในแง่ มุมใด เช่น จะวัดวิชาคณิตศาสตร์ จะวัดตามจุดมุ่งหมายใหญ่ข้อใด จะวัดทักษะ หรือความเข้าใจ หรือการแก้ปัญหา หรือจะวัดรวม ๆ คicumกันไป หรือจะวัดในแง่ผลสัมฤทธิ์หรือความพร้อม หรือ ความถนัด ๆ ลฯ ซึ่งต้องกำหนดไว้ให้ชัดเจน ส่วนใหญ่เราข้ามขั้นนี้กันไป เพราะเรามีจุดมุ่งหมาย อุยกิไจแล้วว่าจะวัดอะไร ແง່ได แต่ต้องทำในรูปแบบแผนงานสร้างข้อสอบที่ดีแล้ว ต้องระบุให้

ชัดเจน เพราะถ้ามีผู้เขียนข้อสอบคนใดนำเอาแผนนี้ไปใช้จะได้ไม่ต้องไปเดาใจผู้วางแผนว่าเขากิตไว้ในใจอย่างไร ซึ่งยังเป็นแนวทางให้แยกย่อยจุดมุ่งหมายนี้ลงเป็นจุดมุ่งหมายเฉพาะที่สามารถนำไปใช้แก่ผู้เขียนข้อสอบเท่านั้นได้อย่างมีผลตีและมีประสิทธิภาพสูงอีกด้วย

8.1.2 ผู้ที่ต้องการวัด ต้องกำหนดไว้ในจุดมุ่งหมายว่าข้อสอบที่เราสร้างขึ้นจะนำไปสู่
กับใคร ใคร ที่wanนี้มีลักษณะเบื้องหลังเบื้องหน้า ความเป็นมาความเป็นไปอย่างไร เช่น เป็นเด็ก
ชั้นม.1 ก็ต้องศึกษาความเป็นของเด็กว่า มีอายุประมาณ 12-14 ปี เดยเรียนและไม่เคยเรียนวิชาอะไรบ้าง
อะไรทำนองนี้ จะช่วยให้ตัดสินใจได้ว่าจะออกข้อสอบภาษาอังกฤษให้มีลักษณะเช่นไร จะให้
ตอบคำถามยาวๆ ได้ไหม? หรือให้มีข้อเลือกมากๆ มีมากข้อใหม่ๆ “คร” ดังกล่าวยังกิน
ความไปถึงกลุ่มที่จะสามารถใช้แบบทดสอบนั้นด้วย ว่าจะใช้กับคนกลุ่มเล็กหรือกลุ่มใหญ่ จะเป็น
เพียงข้อสอบวัดผลสัมฤทธิ์ในโรงเรียนหรือจะเป็นข้อสอบมาตรฐานระดับท้องถิ่น ระดับชาติ
ซึ่งจะเกี่ยวพันถึงลักษณะของแบบทดสอบไปอีกหลายเรื่อง หรืออาจพิจารณาว่าจะสอบวิชานี้กับ
เด็กกลุ่มนี้ในฐานะใด ฐานะที่เป็นวิชาเลือก หรือวิชาบังคับ วิชาเอกหรือวิชาโท ซึ่งจะไปเกี่ยวพัน
ไปถึงเนื้อหาและจุดมุ่งหมายของหลักสูตรว่า ควรออก凸ให้ ยก่ง่ายลึกซึ้งเพียงไร นั่นคือการ
ต้องรู้ว่าสอบ “คร” เป็นการช่วยให้รู้ภัยหลังของผู้จะใช้ข้อสอบชุดนี้นั่นเอง

8.1.3 จุดมุ่งหมายของการสอบ ขั้นนี้เป็นการกำหนดว่าเราจะใช้แบบทดสอบนี้ในงานใด หรือจะใช้ผลการสอบไปใช้ในการตัดสินใจเรื่องใด เช่น เพื่อวินิจฉัย เพื่อจัดตำแหน่ง เพื่อเลื่อนชั้น เพื่อดูสถานภาพปัจจุบันของเด็ก เพื่อทำนายผลภายหน้า ๆ ฯลฯ สิ่งเหล่านี้จะชี้แนะนำว่า ข้อสอบควรจะมีลักษณะอย่างไร เสมือนเข้มทิศบ่งทิศทางชาวเรือ จะนั้น ถ้าหากว่าตั้งจุดมุ่งหมายไว้ว่า จะนำผลการสอบไปวินิจฉัยผลการเรียนในปัจจุบัน เนื้อหาของข้อสอบก็ต้องพุงไปที่คันหน้าจุดอ่อน จุดเด่นของเด็กแต่ละคน จึงต้องสร้างข้อสอบให้มีลำดับขั้นของเนื้อหาและพฤติกรรม คลุมจุดมุ่งหมายของหลักสูตรทั้งหมดอย่างที่ถ้วนกว่าข้อสอบมาตรฐาน จะได้รู้ว่า เก่ง—อ่อน ตรงไหนครับแก่ไขที่จุดไหนจึงจะดี หรือว่าจะวินิจฉัยเฉพาะจุดใดจุดหนึ่ง ก็มุ่งเนื้อหาและพฤติกรรม ที่จะวัดไปที่จุดนั้น เช่น ถ้าครูเห็นว่าเด็กซึ้นนี้จะอ่อนการแยกแฝดเตอร์ ก็วัดลำดับขั้นและวิธีการ แยกแฝดเตอร์กันอย่างที่ถ้วนโดยไม่ต้องคำนึงว่าข้อสอบจะยากสักเพียงไร ถ้าจะสอบเพื่อพยากรณ์ ว่าเด็กคนนี้จะเรียนต่อสาขานี้ให้ไวไหม จะไปได้สวยไหม เนื้อข้อสอบก็ต้องโน้มไปทางที่คาดหวัง ว่าจะพยากรณ์ผลการเรียนในอนาคตได้ เช่น วิชาภาษาไทยแก่เหตุผล ก็ต้องโน้มวัดทางเหตุผลให้เข้มข้น หรือใช้ข้อสอบชนิดวัดเหตุผลกันตรง ๆ เช่น แบบทดสอบความสนใจด้านเหตุผล หรือถ้าจะสอบเพื่อจัดตำแหน่งตัดสินผลในวิชานี้ก็ต้องเบนเข็มไปอีกทางหนึ่ง โดยมุ่งไปที่จุดมุ่งหมายของ หลักสูตรในวิชานั้น ทั้งทางด้านเนื้อหาและพฤติกรรม กล่าวคือให้คลุมทั้งเนื้อหาและพฤติกรรม ว่าที่ได้รับเรียนเขียนอ่านผ่านมา 1 ปี 1 ภาคเรียนนั้น บรรลุจุดมุ่งหมายหลักสูตรแล้วหรือยัง เช่น

วิทยาศาสตร์ มุ่งให้เกิดเจตคติทางวิทยาศาสตร์ ให้คิดอย่างมีเหตุผล รู้จักใช้วิธีการทางวิทยาศาสตร์ ถ้าเรียนวิทยาศาสตร์แล้วยังคงเชื่อเรื่องกฎ屁ปิศาจอำนาจศักดิ์สิทธิ์อิทธิฤทธิ์ต่าง ๆ ก็ได้เช่นว่า ยังไม่บรรลุเป้าหมายแน่ ๆ

โดยสรุป จุดมุ่งหมายของการสร้างแบบทดสอบมุ่งไปที่จะให้ได้คำตอบจากคำตามว่า สอบอะไร—สอบใคร—สอบไปทำไม่ นั้นเอง โดยที่ว่าไปผู้ออกข้อสอบมักขึ้นนี้เหตุด้วยตอบ คำถามนี้ไว้แล้วในใจ แต่ถ้ากำหนดไว้ก็จะช่วยให้คิด เกิดแนวคิดเกี่ยวกับการสร้างแบบทดสอบ ข้อสอบได้อีกมากmany จะอยู่ตัวอย่างของการตั้งจุดมุ่งหมายของการสร้างข้อสอบวิชาภาษาอังกฤษ ดังนี้

“กำหนดจุดมุ่งหมายของการทดสอบภาษาอังกฤษของแบบทดสอบชุดนี้ว่า ภาษาอังกฤษประกอบด้วยศัพท์ ความสามารถในการแปลโครง ความรู้เกี่ยวกับไวยากรณ์ ข้อมูลเกี่ยวกับประเทคโนโลยีและวัฒนธรรมอังกฤษ แบบทดสอบฉบับนี้ใช้กับนักเรียนชั้นมัธยมศึกษาตอนต้นทั่วประเทศไทยที่เรียนอยู่ในชั้น ม.3 คะแนนที่ได้จะใช้เป็นส่วนประกอบส่วนหนึ่งสำหรับให้เกรด”

จะเห็นว่าจุดมุ่งหมายนี้กำหนดว่า “สอบอะไร” สอบภาษาอังกฤษซึ่งมีเนื้อหาและพุทธิกรรมดังกล่าว “สอบใคร” นักเรียน ม.3 “สอบทำไม่” สอบเพื่อเอาผลไปใช้กำหนดเกรดผลการเรียนนั้นเอง

8.2 กำหนดลักษณะของข้อสอบ

ถ้าเราหยิบข้อสอบฉบับหนึ่ง (ที่ดี ๆ) ขึ้นมาพิจารณาจากปกหน้าไปจนถึงปกหลังจะพบว่า มีคำชี้แจงการทำข้อสอบตัวข้อสอบซึ่งมีแบบต่าง ๆ กันออกไป ตัวข้อสอบก็วัดเนื้อหาและพุทธิกรรมต่าง ๆ กันออกไป ยกบ้างง่ายบ้าง ลักษณะเหล่านี้แหลกเป็นลักษณะรวม ๆ ของข้อสอบฉบับนั้น ถ้าหากเราพิจารณาให้ละเอียดจะพบว่าบางลักษณะสังเกตเห็นได้เลย บางลักษณะมองไม่เห็นด้วยตา เช่น ข้อสอบฉบับนี้ยาก—ง่ายเพียงใด มีการวัดพุทธิกรรมหรือสมรรถภาพใดบ้าง ต้องดีเปลี่ยนหมายจากข้อสอบแต่ละข้อออกแบบให้ก่อให้เกิดก่อให้เกิด ลักษณะของข้อสอบที่ควรคำนึงถึง มีดังนี้

- 8.2.1 การเลือกแบบของข้อสอบปรนัย
- 8.2.2 การกำหนดความยาวของข้อสอบ
- 8.2.3 การกำหนดช่วงเวลาที่ใช้ในการทดสอบ
- 8.2.4 การกำหนดจำนวนข้อที่จะใช้ในการทดลองข้อสอบ
- 8.2.5 กำหนดความยากของข้อสอบ
- 8.2.6 กำหนดวิธีดำเนินการสอบ

8.2.1 การเลือกแบบของข้อสอบปรนัย ข้อสอบที่มีความเป็นปรนัย (Objectivity) มีลักษณะ 3 ประการดังกล่าวไว้ในทั้งๆ แล้ว ปัญหาของเรานะนี้คือการเลือกแบบ (Form Type) ของข้อสอบแบบปรนัยว่าจะเลือกใช้อย่างใดดี การที่จะเลือกได้ถูกต้องกับจุดมุ่งหมาย ผู้ออกข้อสอบต้องรู้ว่าแต่ละแบบมีลักษณะอย่างไร ดีเสียด้วยที่ตรงไหน เสียก่อนเป็นประการแรก รูปแบบต่างๆ ของการเขียนข้อสอบปรนัยมีดังนี้

1) แบบตอบคำสั้นๆ (Short Response) ข้อสอบแบบนี้จะมีคำสั้นๆ ให้ตอบสั้นๆ ผู้ตอบมีอิสระเสรีมากกว่าแบบอื่นๆ คำตอบที่ได้ออกมา ก็อยู่ในวงกว้างกว่าไม่เฉพาะเจาะจง เมื่อตายตัวมากันนัก แต่ก็ยังมีทิศทางว่าคำตอบที่ถูกควรจะเป็นอย่างไร เช่น

- ความเป็นปรนัยของข้อสอบมีอะไรบ้าง
- สำคัญด้านใดแล้วจะเกิดอะไรตามมา
- จะเขียนวงจรการเกิดของฝีเสืออย่างคร่าวๆ

2) แบบจับคู่ (Matching) แบบนี้จะกำหนดคำหรือข้อความมาให้ 2 ตอน โดยมากมักแยกเป็นสองทางซ้ายมือกับขวาเมื่อ แล้วให้ผู้สอบจับคู่คำหรือข้อความนั้นให้ถูกต้อง ตัวอย่าง จงนำตัวอักษรหน้าข้อความทางขวาไปใส่ใน (...) หน้าข้อความทางซ้ายมือให้มีความสัมพันธ์กันถูกต้อง

- | | |
|----------------|----------------------------|
| () 1. ค้างคาว | ก. สัตว์เลี้ยงคลาน |
| () 2. ปลาดาว | ข. สัตว์ไม่มีกระดูกสันหลัง |
| () 3. ปลาแพ | ค. สัตว์เลี้ยงลูกด้วยนม |
| () 4. จิงห์ด | ง. สัตว์พากแมลง |
| () 5. จิงเหلن | จ. สัตว์ครึ่งบกครึ่งน้ำ |

ข้อสอบแบบนี้ที่จริงแล้วก็คือแบบเลือกตอบนั่นเอง เช่น ข้อ 1. มีข้อเลือก 5 ข้อ ทางขวาเมื่อ ส่วนข้อ 2. มีข้อเลือก 5 ข้อ หรือน้อยกว่า จุดอ่อนของแบบนี้ก็คือตัวเลือกจะลดลงทุกที ยิ่งข้อท้ายๆ แบบไม่มีตัวเลือก ทำให้เปิดโอกาสให้การเดามาก มีอัตราผลมากในข้อท้ายๆ แต่ก็หมายความว่าจะเกี่ยวกับความจำในแบบของความสัมพันธ์

3) แบบถูกผิด (True False) ข้อสอบแบบนี้จะกำหนดข้อความมาให้ แล้วให้ตัดสินว่าถูกหรือผิด ข้อเสียมีอยู่ตรงที่เปิดโอกาสให้เดาได้มาก เพราะแต่ละข้อมือโอกาสเดาถูกถึง 50% แต่ก็เป็นข้อสอบที่สร้างง่าย

ตัวอย่าง : จงเขียนเครื่องหมาย ✓ หน้าข้อที่ทำนหนึ่นว่าถูกต้อง และ ✗ หน้าข้อที่ทำนหนึ่นว่าผิด1. การทำไร่เลื่อนloy เป็นการทำลายป่าไม้โดยไม่เจตนา

-2. น้ำมันคือสินค้าออกที่สำคัญของประเทศอินโดนีเซีย
-3. การอนุมาติอาหารมีประโยชน์ในด้านการประยุกต์รายจ่ายมากที่สุด
-4. พิชเชอร์ชูกิจหมายถึงพิชที่สามารถเก็บไว้ได้นาน
-5. ปัญหาสำคัญในการทำงานคือการขาดแคลนน้ำ

4) แบบเติมคำหรือข้อความ (Completion) ข้อสอบแบบนี้จะให้ข้อความมาแต่เว้น
บางส่วนไว้ให้ผู้สอบเติมให้ถูกต้อง การเติมอาจให้เติมไปในตัวข้อสอบหรือจัดที่ไว้ต่างหากหรือ
เติมในกระดาษคำตอบก็ได้

ตัวอย่าง :

1. ถ้าท่านไม่มีอาหารประเภทเนื้อสัตว์ ท่านควรรับประทาน..... ทดแทน
2. $3 + (5 \times 7) =$
3. พระเด่นค เป็นผู้ทรงคุณสมบัติในด้าน.....

5) แบบเลือกตอบ (Multiple Choice) ข้อสอบแบบนี้แต่ละข้อจะประกอบด้วย 2
ส่วนคือ ส่วนที่เป็นคำถามเรียกว่า Stem และส่วนที่ให้เป็นตัวเลือกเรียกว่า Options หรือ Choices
หรือ Responses คำสั่งจะเป็นแบบให้เลือกข้อใดข้อหนึ่งจาก Choices ที่ให้มา

ตัวอย่าง : จงเลือกคำตอบที่ถูกที่สุดเพียงคำตอบเดียว

คุณสมบัติที่ดีเด่นของคำถามแบบเลือกตอบคือข้อใด

- | | |
|------------------|------------------------------|
| 1. ตอบได้ง่าย | 2. ตรวจง่าย |
| 3. วัดสมองได้ลึก | 4. นักเรียนเข้าใจคำถามตรงกัน |
| 5. ขัดโภกประสงค์ | |

คำถามแบบเลือกตอบนี้ลดการเดาไปได้มาก ลดความก้าวหน้าของคำถามได้มาก เหมาะ
สำหรับใช้ในรายวิชาที่ต้องการเรียนหรือสมรรถภาพสมองได้ลึก มีความเชื่อมั่นสูงให้
คะแนนหลากหลาย แต่ก็สร้างให้ได้ยาก ต้องอาศัยเวลาประสบการณ์ และความสามารถเฉพาะตัว
ของผู้เขียนด้วย

หลักการพิจารณาเลือกแบบของข้อสอบ

ในการพิจารณาว่าควรจะเลือกออกแบบใด ขึ้นอยู่กับเงื่อนไขต่อไปนี้ ดังนี้

1) ความเหมาะสมระหว่างแบบของข้อสอบกับเนื้อหาวิชา การที่จะใช้แบบของข้อสอบ
ปรนัยแบบใดจะต้องพิจารณาตามความเหมาะสม สมกับแบบทดสอบแต่ละฉบับ เพราแบบของ
ข้อสอบปรนัยแต่ละแบบมีจุดเด่นด้อยต่างกัน เช่น ข้อสอบแบบเติมคำมีความเป็นปรนัยน้อยกว่า
แบบอื่น แต่ก็เหมาะสมสำหรับวัดทางด้านความรู้ความจำ (Knowledge) หรือข้อเท็จจริง (fact) ของ

วิชาต่างๆ แบบจับคู่ (Matching) เหมาะสำหรับวัดความรู้ความจำในแง่ของความสัมพันธ์ เช่น เนื้อหาที่จะตอบคำถามได้ว่า “ครัวกับใคร ใครที่ไหน, อะไร ที่ไหน” แต่ไม่เหมาะสมกับการตอบคำถามว่า “ทำไม” พูดง่าย ๆ ก็คือ วัดเหตุผลได้ไม่ค่อยดี ซึ่งทั้งสองแบบดังกล่าวหากแก่การวิเคราะห์เพื่อปรับปรุงข้อสอบฉบับนั้นหรือข้อนั้น ๆ ส่วนแบบเลือกตอบใช้ได้กว้างขวาง คือวัดได้ทุกแบบทุกเนื้อหา ทุกระดับสมรรถภาพหรือพฤติกรรม ซ้ายยังเหมาะสมแก่การวิเคราะห์ปรับปรุงให้ดีได้อีก กล่าวโดยสรุปก็คือ จะใช้แบบข้อสอบปรนัยแบบใดก็ตามขึ้นอยู่กับความเหมาะสมกับเนื้อหาและจุดมุ่งหมายของหลักสูตรที่ข้อสอบฉบับนั้นดังใจจะวัดเป็นสำคัญ

2) ความเหมาะสมกับสภาพการสร้างแบบทดสอบ เนื่องจากข้อสอบมีแบบที่ยาว สั้น กินเนื้อที่มากน้อยต่างกัน จึงต้องคำนึงถึงทุน วัสดุที่จะใช้ เครื่องมือที่จะใช้ผลิตข้อสอบด้วย เช่น กระดาษ เครื่องพิมพ์ ถ้าไม่มีกระดาษพอเพียงการจะใช้ข้อสอบแบบเลือกตอบที่กินที่หลายหน้า ก็ไม่เหมาะสม

3) ความเหมาะสมกับระบบตรวจสอบ เกณฑ์ในการพิจารณาว่าควรใช้แบบของข้อสอบแบบใดที่สำคัญประการหนึ่งคือ การให้คะแนน ข้อสอบปรนัยมีหัวใจสำคัญอยู่ที่ความเป็นปรนัยของคะแนน คือไม่ว่าใครมาตรวจ ตรวจเมื่อไร ก็ได้คะแนนคงตัวไม่เปลี่ยนไปมา ข้อนี้นับว่า แบบเลือกตอบได้เปรียบข้อสอบปรนัยอื่น ๆ กล่าวคือให้คะแนนที่มีความเป็นปรนัยกว่าแบบเติมคำ ไม่ว่าจะเหมาะสมในการตรวจด้วยมือหรือด้วยเครื่อง

4) ความเหมาะสมกับคุณสมบัติสำคัญของแบบทดสอบ ซึ่งได้แก่ค่าความเชื่อมั่น (Reliability) และค่าความเที่ยงตรง (Validity) คุณสมบัติ 2 ประการนี้ นับว่าเป็นเกณฑ์ (Criteria) ที่สำคัญมาก แบบของข้อสอบแบบปรนัยที่กล่าวมาแล้วให้ค่าความเชื่อมั่นและความเที่ยงตรงต่างกัน เช่น แบบเลือกตอบให้ค่าความเชื่อถือได้ของคะแนนตีกาว่าแบบถูกผิดและจับคู่ตรงที่ความคลัดเคลื่อนจากการเดาเข้ามายืนกับพัฒนาการจริง (True Score) น้อยกว่าและให้ค่าความเที่ยงตรงตีกาว่าแบบอื่นก็ตรงที่สามารถวัดได้ครอบคลุมเนื้อหา เนื่องด้วยวัดได้ทุกจุดของเนื้อหา เมื่ออภิมากข้อเข้า แลงยังเหมาะสมที่จะวัดคุณภาพสมองที่สูงกว่าความรู้ความจำได้อีกด้วย เช่น การนำไปใช้ ความเข้าใจ การวิเคราะห์ สังเคราะห์ ประเมินค่า ซึ่งเป็นจุดมุ่งหมายของหลักสูตร จึงนับว่าให้ค่าความเที่ยงตรงดีทั้งด้าน โครงสร้าง และ เนื้อหา

5) ความเหมาะสมกับการดำเนินการสอบ ในกรณีการสอบนั้นเรามุ่งหวังจะให้ແ่ความยุติธรรมให้ทั่วทุกตัวคนประการหนึ่ง ให้ความสะดวกมากที่สุดแก่ผู้เข้าสอบอีกประการหนึ่ง ดังนั้น จึงควรเลือกแบบของข้อสอบที่จะทำให้เกิดปัญหาต่าง ๆ แก่ผู้เข้าสอบน้อยที่สุด เช่น การไม่เข้าใจคำสั่งพอเพียง การสับสนเรื่องคำศัพท์ ยิ่งเป็นข้อสอบที่ใช้สอบกับคนจำนวนมากสอบหลายแห่งหลายที่พร้อม ๆ กัน ผู้ดำเนินการสอบไม่ได้รับการอบรมพอเพียงแล้ว ผู้วางแผนสร้าง

ข้อสอบยิ่งจะต้องพยายามพิจารณาแบบของข้อสอบที่จะใช้มากเป็นพิเศษ เพื่อลดปัญหาในการดำเนินการสอบให้เหลือน้อยที่สุด

กล่าวโดยสรุปแล้ว เป็นหน้าที่และความรับผิดชอบโดยตรงของฝ่ายทดสอบที่จะพยายามหาทางตีแผ่ผลลัพธ์ที่ทางด้านต่าง ๆ ของผู้เข้าสอบออกมามากที่สุดเท่าที่จะมากได้ ดังนั้นการดำเนินการสอบจึงต้องไม่ทำให้ผู้สอบถูกกรุณด้วยปัญหาอื่นใดที่จะมีผลต่อการทำข้อสอบ นั้นคือพยายามให้เกิดปัญหาต่าง ๆ น้อยที่สุด ผู้สอบคนหนึ่งลูกขึ้นสามเพียงประโภคเดียวก็มีผลต่อการทำข้อสอบของทุกคนทั้งห้อง เช่น รบกวนสมาชิก ทำให้ชุนใจ ฯลฯ ถ้าเกิดปัญหานี้แม้แต่เพียง กิงหอยนับไปได้ว่าเป็นข้อกพร่องของผู้วางแผนออกแบบข้อสอบ หรือฝ่ายทดสอบทั้งสิ้น จะเห็นได้ว่า คุณภาพของการดำเนินการสอบว่าจะดีเลิศ—เลวยิ่งเพียงใด ข้อนี้ยังคงแบบของข้อสอบมิใช่น้อย ผู้วางแผนจึงต้องพิจารณาให้เหมาะสมว่าควรจะใช้แบบของข้อสอบแบบใดจึงจะดี

นอกจากเกณฑ์ดังกล่าวแล้วยังมีปัญหาที่ถูกเกี่ยวกับมากอีกประการหนึ่งก็คือ ในข้อสอบฉบับเดียวกันควรจะมีแบบข้อสอบหลายแบบหรือไม่ ฝ่ายที่ต้องการให้มีหลายแบบก็กล่าวอ้างด้วยเหตุผลสำคัญ 2 ประการ คือ

1) ถ้ามีหลายแบบในฉบับเดียวกันแล้ว จะทำให้ผู้สอบสนใจกว่าแบบเดียวตรงที่ไม่น่าเบื่อ ได้เปลี่ยนบรรยากาศไปเสียบ้าง ไม่ซ้ำซากอยู่แบบเดียว

2) แบบของข้อสอบแต่ละหัวข้อแต่ละตอน จะเหมาะสมกับข้อสอบแต่ละแบบเท่านั้น จึงเห็นควรให้มีการยืดหยุ่นในการใช้แบบของข้อสอบในแต่ละตอน

แต่อย่างไรก็ตามส่วนใหญ่จะเห็นพ้องต้องกันว่า การยุกความสนใจของผู้สอบไว้กับข้อสอบนั้น คุณภาพของข้อสอบต่างหากที่มีอิทธิพลมากกว่าลักษณะภายนอก (หลายแบบ) ที่ลงตัว และแบบทดสอบบางแบบ เช่น แบบเลือกตอบสามารถใช้ได้กับเนื้อหาและจุดมุ่งหมายของหลักสูตรทุกตอนและทุกระดับ ก็ยอมจะแนะนำสมกว่าที่จะใช้มากแบบ จึงอาจสรุปได้ว่า ถ้าจะใช้มากกว่า 1 แบบ ก็ไม่ควรมีมากแบบ จนไปมีอิทธิพลต่อการดำเนินการสอบประสิทธิภาพของการตอบข้อสอบ อันจะมีผลต่อความเที่ยงตรงและความเชื่อมั่นของแบบทดสอบ ซึ่งเป็นเรื่องใหญ่กว่า กล่าวคือถ้าข้อสอบฉบับใด เดียวให้เดิมคำเดียวให้ตอบสั้น ๆ เดียวให้จับคู่ สลับกันเป็นโกลาหลแล้ว ก็จะมีปัญหาต่อการตรวจให้คะแนน ผู้สอบก็สับสน จึงสรุปได้วายิ่งน้อยแบบยิ่งดี

8.2.2 การกำหนดความยาวของแบบทดสอบ

ข้อทดสอบแต่ละข้อที่รวมกันขึ้นเป็นฉบับนั้น โดยแท้จริงแล้วมีสภาพเป็นเพียง “กลุ่มตัวอย่าง” ของผลการศึกษาเล่าเรียนในกระบวนการวิชานั้น ๆ ซึ่งเป็น “ประชากร” ของความรู้ การที่เราสอบกันทุกวันนี้จึงเป็นการสอบ “ตัวอย่าง” ของความรู้ ถ้าสอบ “ประชากร” ของเนื้อหา

แล้วอาจต้องใช้เวลาสอบเกินกว่า 1 วัน เหมือนสอบจอหงวน ถึงกับต้องส่งข้าวส่งน้ำกัน และก็ต้องสอบกันหลายพันข้อ จึงจะวัดได้ทั่ววิเคราะห์ความรู้ในเรื่องนั้น ๆ จริงหรือไม่ แต่การทำเช่นนั้นจะทำให้เสียเงินเสียแรงงานและเวลาไม่คุ้มค่าที่ทุ่มไป จึงต้องเลือกเอาตัวแทนของความรู้มาสอบ มาวัดกันอย่างมากที่ไม่เกิน 200 ข้อ ใช้เวลา 1 ชั่วโมงหรือ 2 ชั่วโมงก็เสร็จ ปัญหาจึงมีอยู่ว่า กลุ่มตัวอย่างความรู้ที่ว่าความมีจำนวนเท่าไรจะเหมาะสม จึงจะเป็นตัวแทนประชากรความรู้ทั้งสิ้นทั้งปวงได้ดี หลักสำคัญมีว่าจะใช้สอบมากขึ้น้อยข้อย่างไรไม่สำคัญ ขอเพียงให้ข้อสอบมีความเชื่อมั่นและความเที่ยงตรงสูงก็ใช้ได้ เราจึงต้องมาพิจารณาว่าความเชื่อมั่นและความเที่ยงตรงของข้อสอบ มีความเกี่ยวข้องสัมพันธ์กับความยาว—สั้น ของข้อสอบอย่างไร ดังนี้

1) ข้อสอบที่มีความเที่ยงตรงทางเนื้อหา หมายถึงต้องถามวัดครอบคลุมเนื้อหาทั้งหมด ในทางปฏิบัติจึงเน้นและเลือกเอาเฉพาะหัวข้อสำคัญที่คาดว่าจะเป็นตัวแทนประชากรในด้านเนื้อหาได้ ส่วนในทางจุดมุ่งหมายของการศึกษาที่หมายถึง คุณพุทธิกรรมที่มุ่งจะให้เกิดขึ้นจากการฝึกการผนតัวมันให้อาชญาดังกล่าว เราถ้าเลือกตัวแทนพุทธิกรรมมาอีก จึงเห็นได้ว่าความยาว ความสั้นของข้อสอบหนึ่งจะขึ้นอยู่กับ “ตัวอย่าง” ที่สำคัญ ๆ ทางด้านเนื้อหาและพุทธิกรรมที่เราตัดสินใจคิดว่ามัน “สำคัญ” และเป็น “ตัวแทน” ได้ ซึ่งสามารถดูได้จากการวิเคราะห์หลักสูตรที่กำหนดน้ำหนักของเนื้อหาและพุทธิกรรมไว้

2) ความเชื่อมั่นของข้อสอบ ถ้าหากคุณสมบัติอื่นเหมือนกันทุกประการแล้ว ข้อสอบฉบับที่มีจำนวนข้อมาก ๆ จะมีความเชื่อมั่นสูงกว่าฉบับที่มีข้อน้อย ๆ คือให้คะแนนเชื่อมั่นได้สูงกว่าหรือใกล้เคียงกับคะแนนจริง (True Score) มากกว่า จึงเห็นได้ว่าการกำหนดความยาวของแบบทดสอบต้องคำนึงถึงปัญหานี้มากด้วย เหตุฉะนี้ข้อสอบปรนัยส่วนใหญ่จึงมีมากข้อ เพราะถ้ามีน้อยข้อค่าความเชื่อมั่นจะต่ำ คะแนนที่ได้มาจะเชื่อถือไม่ได้

เนื่องจากจำนวนข้อมากหรือน้อยเกี่ยวข้องโดยตรงกับเวลาที่ให้ทำ ถ้ามากข้อขึ้น เวลาที่ต้องมากขึ้นตามไปด้วย ถ้ามีเวลาจำกัด เช่น 1—2—3 ชั่วโมง ก็ต้องกำหนดให้จำนวนข้อสอบ พอเหมาะสมพอควรกับเวลา โดยสภาพความเป็นจริงแล้วไม่มีครรภ์เวลาที่เหมาะสมของข้อสอบ แต่ละข้อควรเป็นเท่าไร เพราะขึ้นอยู่กับลักษณะเนื้อหา ขบวนการคิด ความยาก ฯลฯ มากมายจนนับไม่ถ้วน แต่ก็มีหลักพิจารณาอย่าง ๆ ว่าในการสอบแต่ละครั้งควรจะมีผู้ทำเรื่องทุกข้อประมาณ 90% จากจุดนี้เองเราถ้าสามารถประเมินจำนวนข้อและเวลาได้เหมาะสม หลังจากได้นำแบบทดสอบไปทดลองเสียก่อน (ถ้าเป็นข้อสอบที่จะพัฒนาเป็นข้อสอบมาตรฐานหรือข้อสอบมีคุณภาพสูงทั่ว ๆ ไป)

โดยทางปฏิบัติแล้วเรากำหนดจำนวนข้อมากก่อน แล้วค่อยมาประเมินเวลาที่เหมาะสม ซึ่งมักจะกำหนดเวลาให้มากไว้ เพราะถือว่าเราจะตัดสินว่าเขามีความรู้ปานได้ก็ต้องให้เขาได้ทำ

ข้อสอบจนหมดทุกข้อเสียก่อน ถ้าเวลาน้อยเกินไปทำไม่ทัน ได้คะแนนมาน้อยจะว่าเขามีความรู้น้อย ก็ไม่ได้ เพราะถ้าหากให้เวลามากเขากำหนดจำนวนข้อสอบได้มากกว่านั้น

ในการสร้างแบบทดสอบที่ดี การกำหนดความยาวของข้อสอบและเวลาที่เหมาะสม จะได้จากข้อมูลที่ทำการทดลองข้อสอบ (Pretest) นี้เอง โดยกำหนดเวลาและจำนวนข้อไปก่อน โดยอาศัยประสบการณ์ แล้วให้จดเวลาที่นักเรียน 90% ทำข้อสอบนั้นเสร็จเวลาไว้ (เช่น 60 ข้อ คน 90% ใช้เวลา 60 นาที แสดงว่าใช้เวลาประมาณข้อละ 1 นาที) จึงมาจะประมาณว่าถ้าสอบจริงมีข้อสอบกี่ข้อ จะใช้เวลาสักเท่าไร หรือมีเวลาอยู่จำกัด 1 ชั่วโมง จะใช้ข้อสอบสักกี่ข้อจึงจะดี

8.2.3 การกำหนดช่วงเวลาที่ใช้ในการทดสอบ

ในหัวข้อก่อนได้กล่าวถึงความสัมพันธ์ระหว่างจำนวนข้อกับจำนวนเวลาที่ใช้ให้เหมาะสม ทำไมจึงต้องให้เหมาะสม ถ้าไม่เหมาะสมแล้วจะเกิดอะไรขึ้น เช่น ให้เวลาน้อยเกินไป ทำได้นิดเดียวหมดเวลา แสดงว่าผู้สอบได้ทำข้อสอบเพียงบางส่วนของข้อสอบ หรือได้ทดสอบบางส่วนของความรู้ ตอนที่ทำไม่ทันก็คือตอนที่ไม่ได้ทดสอบ จะพบว่าเมื่อทำไม่ครบทั้งฉบับ แบบทุกคนก็จะกลับเป็นข้อสอบฉบับนั้น (แค่ตรงที่ทำไปถึง) ไม่มีความเที่ยงตรงตามเนื้อหาและพฤติกรรม เมื่อเวลาไม่พอ ก็คงใช้การเสี่ยงทายเดา เอา ทำให้คะแนนที่ได้เชื่อถือไม่ได้ นี้เป็นเหตุผลที่ว่า “ทำไม่จึงต้องใช้เวลาที่เหมาะสม” ถ้าหากให้เวลามากเกินไปจะมีผลเช่นไร ถ้าหากข้อสอบไม่ยากเกิน เป็นภารกิจแก่ตอบได้มาก คนเก่งน้อยกว่าจะพาลตอบได้มากไปด้วย เหตุด้วยมีเวลาเหลือเพือที่จะพิจารณาอย่างถี่ถ้วนทำให้มีผลต่ออำนาจจำแนกของข้อสอบว่าจำแนกคนออกเป็นลำดับเก่ง-อ่อน ได้ไม่ละเอียด คะแนนก็ถือความหมายลงไป จุดนี้เองเป็นที่มาของแบบทดสอบชนิด Speeded Test ที่ว่าถ้าให้เวลาน้อย ๆ ทำข้อสอบที่ไม่ยากนัก คนเก่งกว่าจะทำได้มากกว่าและถูกต้องมากกว่า จำแนกคนเป็นอันดับลดหลั่นได้ละเอียดละเอียดกว่าให้เวลามาก ๆ กล่าวได้ว่า “เวลาที่เหมาะสม” ก็คือ “เวลาที่จะทำให้ข้อสอบฉบับนั้นได้ค่าความเชื่อมั่นและความเที่ยงตรงสูงสุด” ฉะนั้น ความยาวของข้อสอบจึงขึ้นอยู่กับ

ก. จำนวนเนื้อหาและพฤติกรรมสำคัญที่จะวัด

ข. เวลาที่จะใช้ในการตอบข้อสอบฉบับนั้น

8.2.4 กำหนดจำนวนข้อที่จะใช้ในการทดลองสอบ

จำนวนข้อสอบที่ใช้ในการทดลองสอบ (Pretest) ตามปกติจะต้องสร้างไว้ให้มากกว่าที่จะสร้างเป็นฉบับจริง เพราะเมื่อนำไปทดลองสอบแล้ววิเคราะห์ จะปรากฏว่าจะมีข้อสอบอยู่จำนวนหนึ่งที่ใช้ไม่ได้ต้องยก去 ขึ้นนำไปใช้ก็รังแต่จะทำให้ข้อสอบชุดนั้นแล้ว จำนวนข้อสอบที่ใช้ไม่ได้นี้จะมีมากน้อยต่างกันไปขึ้นอยู่กับความรู้และประสบการณ์ของผู้สร้างเป็นสำคัญ

บางที่ก็อยู่ในช่วง 10–20% บางที่ก็มากถึง 50%

8.2.5. กำหนดความยากของข้อสอบ

การจะกำหนดความยากของข้อสอบอย่างไร ขึ้นอยู่กับจุดมุ่งหมายของการทดสอบ เป็นสำคัญ เช่น ถ้าสอบเพื่อประเมินผลรวมกิจกรรมใช้ข้อสอบที่มีความยากง่ายปานกลาง ถ้าสอบเพื่อการคัดเลือกที่ต้องใช้ข้อสอบที่ค่อนข้างยาก เป็นต้น (ดูรายละเอียดวิธีการคัดเลือกข้อสอบเพื่อจุดประสงค์ต่าง ๆ ในบทต่อไป)

8.2.6. กำหนดวิธีดำเนินการสอบ

ในการวางแผนการสร้างข้อสอบจะต้องกำหนดวิธีดำเนินการสอบไว้ด้วย เพื่อให้ผู้นำข้อสอบไปใช้สามารถใช้ได้ถูกต้อง ดำเนินการสอบเหมือน ๆ กัน เช่น จะเลือกเอาเป็นชนิดติดไว้ที่ปากข้อสอบ หรือทำเป็นคู่มือต่างหาก โดยแนะนำอย่างละเอียดว่า จะทำอะไร อย่างไร เมื่อไร ซึ่งจะประกอบด้วยการซึ่งแบบของข้อสอบ เวลา จำนวนข้อ ตัวอย่างวิธีการตอบชนิดและลักษณะของกระดาษคำตอบ และยังต้องกำหนดให้แน่ชัดลงไปว่าจะพิมพ์ข้อสอบแบบใด มีรูปแบบอย่างไร

8.3 กำหนดเนื้อหาของข้อสอบ

เนื้อหาของข้อสอบในที่นี้หมายถึงเนื้อวิชา (Subject Matter) ที่กำหนดไว้ในหลักสูตร รวมทั้งจุดมุ่งหมาย (Objective) ของการเรียนการสอนที่เรามุ่งหวังให้ผู้เรียนเกิดสมรรถภาพอุปกรณ์ที่เราสามารถจะรับได้หรือเรียกอีกอย่างว่า “แบบของพฤติกรรม” ที่มุ่งหวังจะให้เกิดขึ้นหลังจากได้เรียนได้สอนในเนื้อหาใดเนื้อหานั่นไปแล้ว ดังนั้นในการเตรียมเนื้อหาของแบบทดสอบ จึงต้องมีเนื้อหาวิชา คู่กับจุดมุ่งหมายอันเป็นพฤติกรรมที่มุ่งหวัง และต้องมีการให้นำหนักหัวข้อดังกล่าวให้เห็นว่าสำคัญมากสำคัญน้อยแตกต่างกันออกไป เช่น เรียนเรื่องลมมรสุมเอเชีย จะมีหัวข้อแตกต่างกันออกไปตั้งแต่ สถานที่เกิด สาเหตุการเกิด ผลกระทบมรสุมเอเชีย ฯลฯ หัวข้อเหล่านี้จะมีอันดับความสำคัญต่างกันออกไป เพื่อจะนำออกเป็นข้อสอบ จำนวนข้อสอบก็ต้องไม่เท่ากันในแต่ละหัวข้อ และในแต่ละหัวข้อยังมีจุดมุ่งหมายจะให้เกิดพฤติกรรมต่าง ๆ ไม่เหมือนกัน มีความสำคัญมากน้อยต่างกันไปอีก เช่น สาเหตุการเกิดลมมรสุมเอเชีย เรายังจะต้องเน้นสมรรถภาพด้าน “ความเข้าใจ” มากกว่า “ความจำ” เป็นแนว การเตรียมโครงร่างของเนื้อหาของแบบทดสอบจึงต้องจัดทำแบบตาราง 2 ทาง บ่งช่องเนื้อหาและพฤติกรรม (ดูรายละเอียดของตารางวิเคราะห์หลักสูตร ในบทที่ 3) จึงเห็นว่าการกำหนดเนื้อหาของแบบทดสอบก็คือการกำหนดเนื้อหาและจุดมุ่งหมายของการเรียนการสอน ว่าจะสอบอะไร อย่างละเอียด นั่นเอง ถ้าหากมีตารางวิเคราะห์หลักสูตรอยู่แล้วก็ยิ่งสะดวกมากขึ้น เพียงแต่นำหนักที่กำหนดในตารางวิเคราะห์หลักสูตรมาปรับให้เข้า

กับจำนวนข้อของแบบทดสอบก็เป็นอันใช่ได้ จะขอยกตัวอย่างตารางวิเคราะห์ของวิชาวิทยาศาสตร์ ที่จะนำมาใช้ออกแบบข้อสอบ 100 ข้อ ดังนี้

เนื้อหา	จุดมุ่งหมาย				รวม
	ความรู้ ความจำ	ความ เข้าใจ	การ นำไปใช้	การ วิเคราะห์	
1. ระบบสุริยะ	4	1		1	6
2. หิน แร่	4	3	1		8
3. การเปลี่ยนแปลงผิวโลก	3	4	2	1	10
4. การจัดประগเขตัวร์	2	5	4		11
5. พืช	3	3	5	1	12
6. ประชากร วิถีชนเผ่า	2	4	1	5	12
7. น้ำ	2	3	5	2	12
8. พลังงานไฟฟ้า	2	8	4	2	16
9. ทรัพยากรธรรมชาติ	3	4	5	1	13
รวม	25	35	27	13	100

จากตารางวิเคราะห์ข้างบนนี้ จะเห็นว่าถ้ามีข้อสอบ 100 ข้อ จะถามเน้นมากที่สุดในเรื่อง ความเข้าใจ และการนำไปใช้ ความรู้ความจำ และการวิเคราะห์ น้อยลงเป็นลำดับ ถ้าตารางนี้ ปรับมาจากการวิเคราะห์หลักสูตรที่ใช้เป็น แนวทาง ใน การสอนก็แสดงว่า แต่ละหัวข้อมีเน้น ความสำคัญต่างกัน เรื่องพลังงานไฟฟ้ามากที่สุด ระบบสุริยะน้อยที่สุด และในหัวข้อทั้งปวงก็ เน้นให้เกิดพฤติกรรมแบบต่างกันออกไป

จากที่กล่าวมาข้างต้นจะเห็นว่า คุณภาพข้อสอบจะวัดได้ลึกได้ดีนี้มีสัดส่วนพอเหมาะสมพอ ควรหรือไม่ แบบทดสอบชุดนี้จะมีโครงสร้างที่ดีหรือไม่ซึ่นอยู่กับตารางวิเคราะห์หลักสูตรอยู่ ไม่น้อย ไม่ว่าผู้ออกแบบข้อสอบเก่งกาจสามารถปานได้ตามที่ได้มีตารางวิเคราะห์นี้ไว้จะสมมูลน คนตาบอดคลำทาง ข้อสอบสร้างมาจะมีลักษณะอย่างไรก็ไม่สามารถพบออกได้แน่นอน

ในการนี้ที่จะสร้างแบบทดสอบที่ไม่ได้มีหลักสูตรกำหนดไว้แน่นอนตามทั่ว ถ้าเราต้องการ จะวัดสมรรถภาพอย่างใดอย่างหนึ่ง ก็จำเป็นต้องรู้ข้อมูลข่ายของสมรรถภาพนั้น หรือรู้หัวข้อของ

งานนั้นว่าต้องการความรู้ความสามารถในด้านใดบ้าง นี่คือต้องแยกแยะวิเคราะห์งานนั้นออกมาเป็นส่วนย่อย ๆ เป็นหัวข้อ การวิเคราะห์นี้อาจใช้การอ้างอิงหลายอย่าง เช่น ผู้เชี่ยวชาญในสาขานั้น คู่มือ ตำรา ฯลฯ เพื่อจะให้ได้มาซึ่ง “หัวข้อของข้อสอบ” ตัวอย่าง สมมุติว่าผู้สร้างแบบทดสอบวางแผนจะสร้างแบบทดสอบความสามารถในการใช้เครื่องยนต์ ก็พยายามจะหาองค์ประกอบที่ส่งผลต่อการใช้เครื่องยนต์ จึงรวบรวมหัวข้อจากคู่มือการใช้เครื่องยนต์มากที่สุดเท่าที่จะมากได้ จับเป็นกลุ่ม ๆ และไปปรึกษากับผู้เชี่ยวชาญว่าหัวข้อเหล่านั้นคุณสมรถภาพในการใช้เครื่องยนต์หรือไม่ และจึงปรับปรุง สมมุติว่าได้หัวข้อตามตัวอย่าง จึงจะมาถึงขั้นที่จะกำหนดน้ำหนัก หรือให้ความสำคัญแต่ละหัวข้อเพื่อจะได้ใช้จำนวนข้อสอบให้เป็นสัดเป็นส่วนกัน จึงสังหัวข้อทั้ง 6 นี้ไปให้ผู้เชี่ยวชาญในสาขาเครื่องยนต์หลาย ๆ คนให้น้ำหนักกว่า ถ้าคิดน้ำหนัก 100% แล้วละก็ ควรจะให้หัวข้อใดก็เบอร์เข็น สมมุติว่าส่งให้ผู้เชี่ยวชาญ 20 คน ก็จะได้มาก 20 แบบ แล้วจึงหาค่าเฉลี่ยหรือหมายฐานของเบอร์เข็น ของแต่ละหัวข้อเป็นน้ำหนักที่จะใช้ในการสร้างข้อสอบ จึงได้ทั้งหัวข้อและน้ำหนักของแต่ละหัวข้อ

หัวข้อ	X ของ %
1. ศัพท์เฉพาะ	20
2. เครื่องมือ	10
3. ระบบหล่อลิ่น	10
4. โครงสร้าง	20
5. การทำงาน	25
6. ระบบไฟฟ้า	15
รวม	<u>100</u>

8.4 หมายกำหนดการเกี่ยวกับการสร้างข้อสอบ

เพื่อจะให้แผนดำเนินไปอย่างราบรื่น ควรต้องวางแผนหมายกำหนดการไว้ให้แน่นอนและยึดหยุ่นได้ ว่าจะทำอะไร เมื่อไร งานแต่ละขั้นตอนจะใช้เวลาสักเท่าไร จะเริ่มงานขั้นต่อไปได้เมื่อไร จึงขออธิบายถึงอันดับขั้นและลักษณะของงานแต่ละลำดับขั้นไว้เพียงสั้นๆเพื่อจะเป็นแนวทางในการพิจารณาว่าจะทำอะไรเมื่อไร ดังนี้

1. ขั้นวางแผนมุ่งหมาย ถ้าต้องทำเป็นแบบทดสอบชนิดใด ทำเป็นกลุ่มในรูปของคณะกรรมการ ก็อาจพิจารณาร่วมกันว่าจะสอบอะไร สอบใคร เพื่ออะไร ซึ่งขั้นนี้จะใช้เวลาเพียงเล็กน้อย อาจนำไปปรับขั้นอื่น ๆ ได้

2. ขั้นกำหนดตัวผู้ออกแบบข้อสอบ ผู้ออกแบบข้อสอบໄต่ดีจะต้องเป็นผู้เชี่ยวชาญแขนงวิชาที่ออกแบบข้อสอบผ่านการอบรมหรือมีประสบการณ์เกี่ยวกับการออกแบบข้อสอบมาพอสมควร เนื่องจากคุณประสงค์ของการสอนการสอบได้เป็นอย่างดี ถ้าจะทำเป็นงานใหญ่ควรมีเจ้าหน้าที่ฝ่ายเทคนิค คือนักวัดผลร่วมอยู่ด้วยอย่างน้อย 1 คน เพื่อจะวิเคราะห์ในช่วงแรกในด้านเทคนิค เช่น แบบของข้อสอบลักษณะภายนอกที่จะมีผลต่อความยากอ่านใจจำแนก และยังความมีผู้เชี่ยวชาญในสาขาวิชานั้นพิจารณาความถูกต้องในด้านเนื้อหาด้วย ฯลฯ

3. ขั้นเขียนข้อสอบ ขั้นนี้กินเวลานาน ควรให้เวลาเพียงพอ เพราะข้อสอบบางแบบออกให้ดีได้ยาก ขั้นนี้จะเริ่มตั้งแต่พิจารณาเลือกแบบข้อสอบ เขียนข้อสอบ สอบทานข้อสอบ บันทึกรายข้อ โดยอาศัยประสบการณ์และความรู้ในการทดสอบ เป็นแนวทางในการวิเคราะห์

4. พิจารณาจัดฉบับ ขั้นนี้เริ่มต้นจากการพิจารณาความยากอย่างคร่าวๆ แล้วจัดซุด จนพิมพ์เป็นฉบับ กำหนดเวลาและวิธีการดำเนินการสอบ

5. ทดลองสอบ (Pretest) ถ้าไม่สร้างเป็นข้อสอบถาวรอาจข้ามขั้น 6-7 "ไปเสียก" ได้ การทดลองสอบเริ่มตั้งแต่การเลือกกลุ่มตัวอย่างซึ่งจะเป็นตัวแทนของประชากรกลุ่มที่ตั้งใจจะใช้ข้อสอบฉบับนั้นทดลองสอบ ตรวจ วิเคราะห์หาค่าความยากง่าย ค่าอำนาจจำแนก ถ้าเป็นข้อสอบมาตรฐาน จะทำขั้นนี้ซ้ำแล้วซ้ำอีกจนได้ค่าสถิติที่คงตัวจนพอใจ แล้วจึงออกแบบทั้งหมด แล้วออกแบบคู่มือการใช้การดำเนินการสอบ

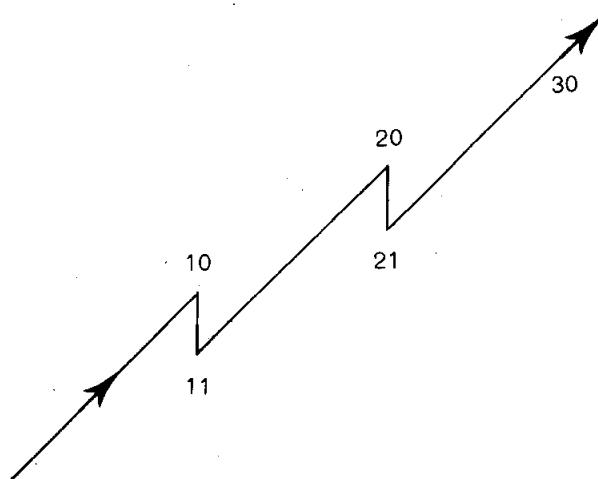
6. การจัดฉบับจริง เริ่มต้นจากการเลือกข้อเสียทั้ง พิจารณาเรียงข้อเสียใหม่ตามเกณฑ์ที่ยึด เช่น เรียงจากข้อง่ายสุดไปทางข้อยากสุด การเลือกข้อใดไว้พิจารณาจากค่าความยากง่าย และอำนาจจำแนกเป็นเกณฑ์

ในการเรียงข้อคำถามในแบบทดสอบนั้น โดยทั่ว ๆ ไปแล้วนิยมเรียงจากข้อง่ายไปทางข้อยาก คือเอื้อข้อง่ายเป็นข้อ 1 แล้วค่อย ๆ ยากขึ้นไปตามลำดับ (วิธีการหาความยากง่ายของข้อสอบให้ดูบกัสดไป) อย่างไรก็ดีวิธีเรียงข้อคำถามก็มีหลายวิธี อาจสรุปได้เป็นวิธีใหญ่ ๆ ได้ 4 วิธี คือ ก. จัดเรียงตามลำดับความยากง่าย วิธีนี้เป็นการจัดเรียงข้อแบบธรรมชาติทั่ว ๆ ไป คือเรียงจากข้อง่ายไปทางยาก ประโยชน์ส่วนใหญ่ของการเรียงข้อสอบวิธีนี้ก็คือผู้สอบจะได้พบข้อสอบง่าย ๆ ก่อน ทำให้เกิดกำลังใจจึงไม่เกิดความเหนื่อยล้ามากนักเมื่อมาพบข้อสอบยาก ๆ แต่การเรียงข้อสอบแบบนี้มีข้อเสีย 2 ประการ คือ

1) การเรียงข้อสอบแบบนี้ทำให้ข้อสอบที่มีเนื้อหาเดียวกันต้องกระจายออกไปตามลำดับความยากง่าย ซึ่งมีผลทำให้ดูความคิดของผู้เข้าสอบต้องเปลี่ยนแปลงอยู่เสมอ เช่นเดียวก็คือเกี่ยวกับเลขบวก ก็ต้องเปลี่ยนมาคิดเลขบวก แล้วไปคิดเลขลบ แล้วกับบวกมาคิดเลขบวกอีก ซึ่งนี้เป็นดัน

2) การเรียงข้อสอบแบบนี้มีผลเสียตรงที่ตัวนักเรียนพบข้อยาก ๆ เสียแล้วในตอนต้น ๆ ก็จะทำให้เกิดความท้อใจที่จะทำข้อต่อไป เนื่องจากนักเรียนรู้ว่าข้อสอบจะยิ่งยากขึ้นไปอีก

ข. จัดเรียงความยากง่ายแบบไซคลิกออร์เดอร์ (Cyclic Order) วิธีเรียงข้อสอบแบบนี้ต่างจากแบบที่ 1 ตรงที่แทนที่จะเรียงข้อสอบจากง่ายไปหางาน พอข้อ 11–14 แทนที่จะยกขึ้นไปอีกกลับลดความยากลงมาเท่ากับข้อ 6–10 พอถึงข้อ 15 ก็จะยากกว่าข้อ 10 ดังรูป



ข้อ 1 (ง่าย)

การเรียงข้อสอบแบบนี้ดีตรงที่จะเป็นการกระตุนให้ผู้สอบทำข้อสอบให้หมดทุกข้อ เพราะเขารู้ว่าถ้าทำข้อสอบในระยะหนึ่งก็พบข้อง่าย ๆ อีก วิธีนี้ถึงจะดี แต่ก็ยังมีข้อเสียอยู่ตรงที่ผู้สอบยังต้องเปลี่ยนความคิดในการแก้ปัญหาเรื่องเกินไป

ค. จัดเรียงตามกลุ่มของเนื้อหาวิชา การเรียงข้อสอบโดยวิธีนี้เป็นการเรียงโดยรวมข้อสอบที่มีเนื้อหาเดียวกันเข้าไว้ด้วยกัน และจัดเรียงข้อตามลำดับความยากง่าย ซึ่งวิธีนี้จะช่วยให้ผู้สอบใช้ความคิดแก้ปัญหาหนึ่งก่อนที่จะเปลี่ยนไปคิดในเนื้อหาอื่น

ง. จัดเรียงตามจุดมุ่งหมายของการวัด ในการจัดเรียงข้อสอบผู้สร้างแบบทดสอบบางคนนิยมจัดเรียงข้อโดยรวมพุทธิกรรมที่ต้องการจะวัดประเภทเดียวกันเข้าไว้ด้วยกัน เช่น รวมข้อคำานที่ถูกต้องกับความรู้ในเนื้อเรื่องไว้พากหนึ่ง การเรียงข้อสอบด้วยวิธีนี้มีประโยชน์สำหรับครูมากในการที่จะรู้ว่าสามารถวัดได้ตรงตามจุดมุ่งหมายที่ต้องการวัดหรือไม่ แต่มีข้อเสียคือบางทีครูไม่สามารถตอกข้อสอบได้ตรงตามพุทธิกรรมที่ต้องการจะวัด

ผู้เขียนได้ทำการวิจัยเกี่ยวกับการเรียงข้อสอบของเต็กไทยในระดับชั้น ป.7 (ເຕືອນໄຈ ເຄຣະຊູສັກໂກ : ປຣິຢູ່ຢານີພັນົງ) พบรວบรวมนักเรียนที่เรียนดีนั้น จะทำข้อสอบได้คะแนนดีที่สุด

ถ้าให้เข้าสอบจากวิชาที่ง่ายไปหาวิชาที่ยาก และภายในแต่ละวิชาเรียงข้อจากง่ายไปหางาก ส่วนนักเรียนกลุ่มที่มีความสามารถปานกลางนั้น จะเรียงข้อสอบแบบไหนก็ไม่สำคัญ เพราะจะได้คะแนนเท่ากันทุกแบบ แต่สำหรับนักเรียนกลุ่มอ่อนน้ำนั้น ชอบทำข้อสอบประเภทบับหนึ่งมีหลาย ๆ วิชา วิชาละ 5–10 ข้อ และภายในแต่ละวิชานั้นก็เรียงจากข้อง่ายไปหางาก ถ้าเรียงแบบนี้เข้าจะได้คะแนนดีกว่าเรียงแบบอื่น ๆ แต่ถ้าดูโดยส่วนรวมแล้ว การเรียงข้อสอบแบบให้สอบวิชาที่ง่ายไปหาวิชาที่ยาก และภายในแต่ละวิชาเรียงข้อจากง่ายไปหางากนั้น เป็นวิธีที่ดีที่สุดสำหรับนักเรียนชั้นประถมศึกษาปีที่ 7 ซึ่งผลจากการวิจัยนี้ควรที่ผู้บริหารโรงเรียนจะได้นำไปพิจารณาในการจัดตารางสอบด้วยจึงจะดี

7. การพิมพ์ข้อสอบ ต้องกำหนดระยะเวลาที่แน่นอนว่าจะส่งต้นฉบับได้เมื่อไร มีรูปแบบอย่างไร โดยปกติแล้วการเลือกรูปแบบของข้อสอบจะพิจารณาจากการให้ความสะดวกแก่ผู้เข้าสอบ เป็นประการสำคัญที่สุด ส่วนเรื่องการประยัดและความคงทนถาวรเมื่อมีความสำคัญรองลงมา

สรุปสาระสำคัญ

สาระสำคัญของเนื้อหาที่อธิบายไว้ในบทนี้อาจสรุปได้เป็นข้อ ๆ ดังนี้

1. ในการกำหนดจุดมุ่งหมายของการสอบจะช่วยทำให้ทราบว่าเราจะทดสอบอะไร สอบกับใครและสอบไปทำไม
2. การพิจารณาว่าควรจะเลือกออกข้อสอบแบบใด ขึ้นอยู่กับความเหมาะสมระหว่างแบบของข้อสอบกับเนื้อหาวิชา ความเหมาะสมสมกับการดำเนินการสอบ และประโยชน์ใช้สอยเป็นสำคัญ
3. แบบทดสอบที่ประกอบด้วยข้อสอบจำนวนมาก ๆ ข้อมูลแนวโน้มที่จะมีความเชื่อมั่นและความเที่ยงตรงสูงกว่าข้อสอบที่มีจำนวนน้อย ๆ ข้อ
4. การกำหนดความยาวของข้อสอบและเวลาที่เหมาะสมจะได้จากการนำข้อสอบไปทำ การทดลองสอบ
5. เวลาที่เหมาะสมสำหรับใช้ในการทดสอบคือเวลาที่จะทำให้ข้อสอบฉบับนั้นมีความเชื่อมั่นและความเที่ยงตรงสูงสุด
6. แบบทดสอบจะมีความยากง่ายระดับใดขึ้นอยู่กับจุดมุ่งหมายของการนำผลการสอบไปใช้
7. เนื้อหาที่กำหนดไว้สำหรับออกข้อสอบต้องสอดคล้องกับจุดประสงค์ของการสอน
8. 在การเรียงข้อสอบนั้น ควรเรียงจากเนื้อหาง่าย ๆ ไปหาเนื้อหาท้าย ๆ และในเนื้อหาแต่ละบทควรเรียงจากข้อง่ายไปหาข้อยาก

คำความท้ายบทที่ 8

1. การสร้างข้อสอบปรนัยมีขั้นตอนการสร้างอย่างไร
2. การกำหนดจุดมุ่งหมายของการสร้างข้อสอบปรนัย มีประโยชน์อย่างไร จอธิบาย
3. ให้ท่านบอกเกณฑ์ที่ใช้ในการพิจารณาแบบของข้อสอบว่ามีเกณฑ์อะไรบ้าง พร้อมทั้งยกตัวอย่าง ประกอบคำอธิบาย
4. ให้ท่านอธิบายความหมายของประโยชน์ข้างล่างนี้
 - 4.1 เวลาที่เหมาะสมสำหรับใช้ในการทดสอบ คือ เวลาที่จะทำให้ข้อสอบฉบับนี้มีความเชื่อมั่น และความเที่ยงตรงสูงสุด
 - 4.2 แบบทดสอบจะมีความยากง่ายระดับใดขึ้นอยู่กับจุดมุ่งหมายของการนำผลการสอนไปใช้

บรรณานุกรม

ชาล แพรตต์กุล. เทคนิคการวัดผล วัฒนาพาณิช 2516.

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

Gronlund, N.E. **Preparing Criterion - Referenced Tests for Classroom Instruction**.
The Macmillan Company, New York, 1973.

Weidemann, C.C. "Written Examination Procedures," **Phi Delta Kappan**, 16, October 1933.