

บทที่ 5

เทคนิคการพัฒนาแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

เนื้อหา

เนื้อหาที่จะกล่าวถึงในบทที่ 5 มีรายละเอียด ดังนี้

- 5.1 กำหนดจุดมุ่งหมายของการสร้างแบบทดสอบ
- 5.2 กำหนดลักษณะของแบบทดสอบ
- 5.3 กำหนดเนื้อหาของแบบทดสอบ
- 5.4 หมายกำหนดการเกี่ยวกับการสร้างแบบทดสอบ

5.1 กำหนดจุดมุ่งหมายของการสร้างแบบทดสอบ

ในการสร้างแบบทดสอบแต่ละชุด จำเป็นต้องกำหนดเป้าหมายให้ชัดเจน เพราะแบบทดสอบแต่ละชุดที่สร้างขึ้นเพื่อจุดมุ่งหมายต่างกันก็เหมาะที่จะใช้ในงานต่างกัน ถ้านำข้อสอบที่มีจุดมุ่งหมายอย่างหนึ่งไปสอบเพื่อจุดมุ่งหมายอีกอย่างหนึ่ง ผลที่ได้มาก็ย่อมไม่สอดคล้องกับสิ่งที่ต้องการวัด จุดมุ่งหมายทั่วไปของการสอบแยกพิจารณา ดังนี้

1. **สิ่งที่ต้องการวัด** เป็นการกำหนดเนื้อหาของแบบทดสอบว่าจะวัดวิชาใด ในแง่ มุมใด เช่น จะวัดวิชาคณิตศาสตร์ จะวัดตามจุดมุ่งหมายใหญ่ข้อใด จะวัดทักษะ หรือความเข้าใจ หรือการแก้ปัญหา หรือจะวัดรวมๆ คลุมกันไป หรือจะวัดในแง่ผลสัมฤทธิ์หรือความพร้อม หรือ ความถนัด ฯลฯ ซึ่งต้องกำหนดไว้ให้ชัดเจน ส่วนใหญ่เราข้ามขั้นนี้กันไป เพราะเรามีจุดมุ่งหมาย อยู่ในใจแล้วว่า จะวัดอะไร แง่ใด แต่ถ้าต้องทำในรูปวางแผนงานสร้างสอบที่ดีแล้ว ต้องระบุให้ชัดเจน เพราะถ้ามีผู้เขียนข้อสอบคนใดนำเอาแผนนี้ไปใช้จะได้ไม่ต้องไปเดาใจผู้วางแผนว่าเขาคิดไว้ในใจ อย่างไร ซ้ำยังเป็นแนวทางให้แยกย่อยจุดมุ่งหมายนี้ลงเป็นจุดมุ่งหมายเฉพาะที่สามารถนำทาง ให้แก่ผู้เขียนข้อสอบเขียนได้อย่างมีผลดีและมีประสิทธิภาพสูงอีกด้วย

2. **ผู้ที่ต้องการวัด** ต้องกำหนดไว้ในจุดมุ่งหมายว่าข้อสอบที่เราสร้างขึ้นจะนำไปสอบ กับใคร ใครที่ว่านั้นมีลักษณะเบื้องหลังเบื้องหน้า ความเป็นมาความเป็นไปอย่างไร เช่น เป็นเด็ก ชั้นอนุบาล 1 ก็ต้องตีความเบื้องหลังว่า มีอายุประมาณ 3-4 ปี เคยเรียนและไม่เคยเรียนสิ่งใดมาบ้าง อะไรทำนองนี้ จะช่วยให้ตัดสินใจได้ว่าจะออกข้อสอบภาษาอังกฤษให้มีลักษณะเช่นไร จะให้ตอบ คำถามยาวๆ ได้ไหม? หรือให้มีข้อเลือกมาก ๆ มีมากข้อไหม ฯลฯ “ใคร” ดังกล่าวยังกินความไป

ถึงกลุ่มที่จะสามารถใช้แบบทดสอบนั้นด้วยว่าจะใช้กับคนกลุ่มเล็กหรือกลุ่มใหญ่ จะเป็นเพียงข้อสอบ วัดผลสัมฤทธิ์ในโรงเรียนหรือจะเป็นข้อสอบมาตรฐานระดับท้องถิ่น ระดับชาติ ซึ่งจะเกี่ยวพันถึง ลักษณะของแบบทดสอบไปอีกหลายเรื่อง ซึ่งจะไปเกี่ยวพันไปถึงเนื้อหาและจุดมุ่งหมายของหลักสูตร ว่า ควรออกแค่ไหน ยากง่ายลึกซึ้งเพียงไร นั่นคือการต้องรู้ว่าสอบ “ใคร” เป็นการช่วยให้รู้ภูมิหลัง ของผู้จะใช้ข้อสอบชุดนี้นั่นเอง

3. จุดมุ่งหมายของการสอบ ชั้นนี้เป็นการกำหนดว่าเราจะใช้แบบทดสอบนี้ในงานใด หรือจะใช้ผลการสอบไปใช้ในการตัดสินใจเรื่องใด เช่น เพื่อวินิจฉัย เพื่อจัดตำแหน่ง เพื่อเลื่อนชั้น เพื่อดูสถานภาพปัจจุบันของเด็ก เพื่อทำนายผลภายหน้า ฯลฯ สิ่งเหล่านี้จะชี้แนะว่า ข้อสอบควรจะมีลักษณะอย่างไร เหมือนเข็มทิศบ่งทิศทางชาวเรือ ฉะนั้น ถ้าหากว่าตั้งจุดมุ่งหมายไปว่า จะนำผลการสอบไปวินิจฉัยผลการเรียนในปัจจุบัน เนื้อหาของข้อสอบก็ต้องพุ่งไปที่ค้นหาจุดอ่อน จุดเก่งของเด็กแต่ละกลุ่มแต่ละคน จึงต้องสร้างข้อสอบให้มีลำดับขั้นของเนื้อหาและพฤติกรรมที่จะวัดไปที่จุด นั้น ถ้าจะสอบเพื่อพยากรณ์ว่าเด็กคนนี้จะเรียนต่อชั้นประถมศึกษาปีที่ 1 ได้หรือไม่ เนื้อข้อสอบก็ต้องโน้มไปทางที่คาดหวังว่าจะพยากรณ์ผลการเรียนในอนาคตได้

โดยสรุป จุดมุ่งหมายของการสร้างแบบทดสอบมุ่งไปที่จะให้คำตอบจากคำถาม ว่าสอบอะไร – สอบใคร – สอบไปทำไม นั่นเอง โดยทั่วไปผู้ออกข้อสอบมักข้ามขั้นนี้เหตุด้วยตอบ คำถามนี้ไว้แล้วในใจ แต่ถ้ากำหนดไว้ก็จะช่วยให้คิด เกิดแนวคิดเกี่ยวกับการสร้างแบบทดสอบข้อสอบ ได้อีกมากมาย จะขอยกตัวอย่างของการตั้งจุดมุ่งหมายของการสร้างข้อสอบวิชาภาษาอังกฤษ ดังนี้ “กำหนดจุดมุ่งหมายของการทดสอบภาษาอังกฤษของแบบทดสอบชุดนี้ว่า ภาษาอังกฤษประกอบด้วย ศัพท์ ข้อมูลเกี่ยวกับประเทศอังกฤษและวัฒนธรรมอังกฤษ แบบทดสอบฉบับนี้ใช้กับนักเรียนชั้น อนุบาล 3”

จะเห็นว่าจุดมุ่งหมายนี้กำหนดว่า “สอบอะไร” สอบภาษาอังกฤษซึ่งมีเนื้อหาและ พฤติกรรมดังกล่าว “สอบใคร” นักเรียนชั้นอนุบาล 3 “สอบทำไม” สอบเพื่อเอาผลไปใช้ ประเมินผลการเรียนนั่นเอง

5.2 กำหนดลักษณะของแบบทดสอบ

ถ้าเราหยิบข้อสอบฉบับหนึ่ง (ที่ดี) ขึ้นมาพิจารณาจากปกหน้าไปจนถึงปกหลังจะพบว่า มีคำชี้แจงการทำข้อสอบตัวข้อสอบซึ่งมีแบบต่าง ๆ กันออกไป ตัวข้อสอบก็วัดเนื้อหาและพฤติกรรมต่าง ๆ กันออกไป ยากบ้างง่ายบ้าง ลักษณะเหล่านี้แหละเป็นลักษณะรวม ๆ ของข้อสอบฉบับนั้น ถ้าหากเราพิจารณาให้ละเอียดจะพบว่าบางลักษณะสังเกตเห็นได้เลย บางลักษณะมองไม่เห็นด้วยตา เช่น ข้อสอบฉบับนี้ยาก-ง่ายเพียงใด มีการวัดพฤติกรรมหรือสมรรถภาพใดบ้าง ต้องแปลความหมายจากข้อสอบแต่ละข้อออกมาอีกทอดหนึ่ง ลักษณะของข้อสอบที่ควรคำนึงถึงมีดังนี้

1. การเลือกแบบของข้อสอบปรนัย
2. การกำหนดความยาวของข้อสอบ
3. การกำหนดช่วงเวลาที่ใช้ในการทดสอบ
4. การกำหนดจำนวนข้อที่จะใช้ในการทดลองข้อสอบ
5. กำหนดความยากของข้อสอบ
6. กำหนดวิธีดำเนินการสอบ

1. การเลือกแบบของข้อสอบปรนัย

ข้อสอบที่มีความเป็นปรนัย (Objectivity) มีลักษณะ 3 ประการดังกล่าวไว้ในบทแรก ๆ แล้ว ปัญหาของเรานั้นคือการเลือกแบบ (Form Type) ของข้อสอบว่าจะใช้ข้อสอบอัตนัยหรือปรนัย ถ้าใช้ข้อสอบแบบปรนัยว่าจะเลือกใช้อย่างใดดี การที่จะเลือกได้ถูกต้องกับจุดมุ่งหมายผู้ออกข้อสอบต้องรู้ว่าแต่ละแบบมีลักษณะอย่างไร ดีเลวเด่นด้อยที่ตรงไหนเสียก่อนเป็นประการแรก รูปแบบต่าง ๆ ของการเขียนข้อสอบปรนัยมีดังนี้

1) **แบบตอบคำถามสั้น ๆ (Short Response)** ข้อสอบนี้จะมีคำถามสั้น ๆ ให้ตอบสั้น ๆ ผู้ตอบมีอิสระเสรีมากกว่าแบบอื่น ๆ คำตอบที่ได้ออกมาก็อยู่ในวงกว้างกว่าไม่เฉพาะเจาะจง ไม่ตายตัวมากนัก แต่ก็ยังมีทิศทางว่าคำตอบที่ถูกต้องควรจะเป็นอย่างไร เช่น

- ความเป็นปรนัยของข้อสอบมีอะไรบ้าง
- ถ้าเกิดฟ้าแลบแล้วจะเกิดอะไรตามมา
- จงเขียนวงจรการเกิดของผีเสื้ออย่างคร่าว ๆ

2) **แบบจับคู่ (Matching)** แบบนี้จะกำหนดคำหรือข้อความมาให้ 2 ตอน โดยมากมักแยกเป็นแถวทางซ้ายมือกับขวามือ แล้วให้ผู้สอบจับคู่คำหรือข้อความนั้นให้ถูกต้อง

ตัวอย่าง : จงนำตัวอักษรหน้าข้อความทางขวามือมาใส่ใน (.....) หน้าข้อความทางซ้ายมือให้มีความสัมพันธ์กันถูกต้อง

- | | |
|----------------|-------------------------|
| () 1. กบ | ก. สัตว์เลี้ยงลูก |
| () 2. ผีเสื้อ | ข. สัตว์เลี้ยงลูกด้วยนม |
| () 3. วัว | ค. สัตว์พวกแมลง |
| () 4. งู | ง. สัตว์ครึ่งบกครึ่งน้ำ |

ข้อสอบแบบนี้ที่จริงแล้วก็คือแบบเลือกตอบนั่นเอง จุดอ่อนของแบบนี้ก็คือตัวเลือกจะลดลงทุกที ยิ่งข้อท้ายๆ แทบไม่มีตัวเลือก ทำให้เปิดโอกาสให้การเดาตามมามีอิทธิพลมากในข้อท้ายๆ แต่ก็เหมาะที่จะใช้วัดเกี่ยวกับความจำในแง่ของความสัมพันธ์

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

ตัวอย่าง : จงเขียนเครื่องหมายถูก (✓) หน้าข้อที่ท่านเห็นว่าถูกต้อง และผิด (X) หน้าข้อที่ท่านเห็นว่าผิด

- 1. ค้างคาวชอบออกหากินในเวลากลางคืน
- 2. เดือนที่ลงท้ายด้วย “คม” มี 30 วัน
- 3. ยุงลายนำเชื้อไข้เลือดออกมาสู่คน
- 4. โองน้ำทำมาจากดิน

4) **แบบเติมคำหรือข้อความ (Completion)** ข้อสอบแบบนี้จะให้ข้อความมาแต่เว้นบางส่วนไว้ให้ผู้สอบเติมให้ถูกต้อง การเติมอาจให้เติมไปในตัวข้อสอบหรือจัดที่ไว้ต่างหากหรือเติมในกระดาษคำตอบก็ได้

ตัวอย่าง :

1. ถ้าท่านไม่มีอาหารประเภทเนื้อสัตว์ท่านควรรับประทาน.....ทดแทน
2. $3 + 5 = \dots\dots\dots$

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

ตัวอย่าง : จงเลือกคำตอบที่ถูกต้องที่สุดเพียงคำตอบเดียว

คุณสมบัติที่ดีเด่นของคำถามแบบเลือกตอบคือข้อใด

1. ตอบได้ง่าย
2. ตรวจง่าย
3. วัดสมองได้ลึก
4. นักเรียนเข้าใจคำถามตรงกัน
5. ขจัดโอกาสที่จะเดาได้ถูกต้อง

คำถามแบบเลือกตอบนี้ลดการเดาไปได้มาก ลดความกำกวมของคำถามได้มาก เหมาะสำหรับใช้วิเคราะห์ วัดระดับของการเรียนหรือสมรรถภาพสมองได้ลึก มีค่าความเชื่อมั่นสูง ให้คะแนนสะดวกแต่ก็สร้างให้ตีได้ยาก ต้องอาศัยเวลาประสบการณ์ และความสามารถเฉพาะตัวของผู้เขียนด้วย

หลักการพิจารณาเลือกแบบข้อสอบ

ในการพิจารณาว่าควรเลือกข้อสอบแบบใดขึ้นอยู่กับเงื่อนไขต่าง ๆ ดังนี้

1) **ความเหมาะสมระหว่างแบบของข้อสอบกับเนื้อหาวิชา** การที่จะใช้แบบของข้อสอบปรนัยแบบใดจะต้องพิจารณาตามความเหมาะสมกับแบบทดสอบแต่ละฉบับ เพราะแบบของข้อสอบปรนัยแต่ละแบบมีจุดเด่นด้อยต่างกัน เช่น ข้อสอบแบบเติมคำมีความเป็นปรนัยน้อยกว่าแบบอื่น แต่ก็เหมาะสำหรับวัดทางด้านความรู้ความจำ (Knowledge) หรือข้อเท็จจริง (Fact) ของวิชาต่าง ๆ ข้อสอบแบบจับคู่ (Matching) เหมาะสำหรับวัดความรู้ความจำในแง่ความสัมพันธ์ เช่น เนื้อหาที่จะตอบคำถามได้ว่า “ใครกับใคร ใครที่ไหน, อะไร ที่ไหน” แต่ไม่เหมาะกับการตอบคำถามว่า “ทำไม” พุดง่าย ๆ ก็คือ วัดเหตุผลได้ไม่ค่อยดี ซึ่งทั้งสองแบบดังกล่าวยากแก่การวิเคราะห์ เพื่อ ปรับปรุงข้อสอบฉบับนั้นหรือข้อนั้น ๆ ส่วนแบบเลือกตอบใช้ได้กว้างขวาง คือวัดได้ทุกแบบทุกเนื้อหา ทุกระดับสมรรถภาพหรือพฤติกรรม ซ้ำยังเหมาะแก่การวิเคราะห์ปรับปรุงให้ดีขึ้นได้อีก กล่าวโดยสรุปก็คือ จะใช้แบบข้อสอบปรนัยแบบใดก็ตามขึ้นอยู่กับความเหมาะสมกับเนื้อหาและจุดมุ่งหมายของหลักสูตรที่ข้อสอบฉบับนั้นตั้งใจจะวัดเป็นสำคัญ

2) **ความเหมาะสมกับสภาพการสร้างแบบทดสอบ** เนื่องจากข้อสอบมีแบบที่ยาวสั้นกินเนื้อที่มากน้อยต่างกัน จึงต้องคำนึงถึงทุน วัสดุที่จะใช้ เครื่องมือที่จะใช้ผลิตข้อสอบด้วย เช่น กระดาษ เครื่องพิมพ์ ถ้าไม่มีกระดาษพอเพียงการใช้ข้อสอบแบบเลือกตอบที่กินที่หลายหน้าก็ไม่เหมาะสม

3) **ความเหมาะสมกับระบบตรวจข้อสอบ** เกณฑ์ในการพิจารณาว่าควรใช้แบบของข้อสอบแบบใดที่สำคัญประการหนึ่ง คือ การให้คะแนน ข้อสอบปรนัยมีหัวใจสำคัญอยู่ที่ความ

เป็นปรนัยของคะแนน คือไม่ว่าใครมาตรวจ ตรวจเมื่อไร ก็ได้คะแนนคงตัวไม่แปรเปลี่ยนไปมา ข้อนี้ นับว่าแบบเลือกตอบได้เปรียบข้อสอบปรนัยอื่น ๆ กล่าวคือให้คะแนนที่มีความเป็นปรนัยกว่าแบบเติมคำไม่ว่าจะเหมาะสมในการตรวจด้วยมือหรือด้วยเครื่อง

4) ความเหมาะสมกับคุณสมบัติสำคัญของแบบทดสอบ ซึ่งได้แก่ค่าความเชื่อมั่น (Reliability) และค่าความเที่ยงตรง (Validity) คุณสมบัติ 2 ประการนี้ นับว่าเป็นเกณฑ์ (Criteria) ที่สำคัญมาก แบบของข้อสอบแบบปรนัยที่กล่าวมาแล้วให้ค่าความเชื่อมั่นและความเที่ยงตรงต่างกัน เช่น แบบเลือกตอบให้ค่าความเชื่อถือได้ของคะแนนดีกว่าแบบถูกผิดและจับคู่ตรงที่ความคลาดเคลื่อนจากการเดาเข้ามามีบทบาทพัวพันกับคะแนนจริง (True Score) น้อยกว่าและให้ค่าความเที่ยงตรงดีกว่าแบบอื่นก็ตรงที่สามารถวัดได้ครอบคลุมเนื้อหา เนื่องด้วยวัดได้ทุกจุดของเนื้อหาเมื่อออกมากข้อเข้า และยังเหมาะที่จะวัดคุณภาพสมองที่สูงกว่าความรู้ความจำได้อีกด้วย เช่น การนำไปใช้ความเข้าใจ การวิเคราะห์ สังเคราะห์ ประเมินค่า ซึ่งเป็นจุดมุ่งหมายของหลักสูตรจึงนับว่าให้ค่าความเที่ยงตรงดีทั้งด้านโครงสร้าง และเนื้อหา

5) ความเหมาะสมกับการดำเนินการสอบ ในการดำเนินการสอบนั้นเรามุ่งหวังจะให้แค่ความยุติธรรมให้ทั่วทุกตัวคนประการหนึ่ง ให้ความสะดวกมากที่สุดแก่ผู้เข้าสอบอีกประการหนึ่ง ดังนั้นจึงควรเลือกแบบของข้อสอบที่จะทำให้เกิดปัญหาต่าง ๆ แก่ผู้เข้าสอบน้อยที่สุด เช่น การไม่เข้าใจคำสั่งพอเพียง การสับสนเรื่องคำชี้แจง ยังเป็นข้อสอบที่ใช้สอบกับคนจำนวนมาก สอบหลายแห่งหลายที่พร้อม ๆ กัน ถ้าผู้ดำเนินการสอบไม่ได้รับการอบรมพอเพียงแล้ว ผู้วางแผนสร้างข้อสอบยิ่งจะต้องพยายามพิจารณาแบบของข้อสอบที่จะให้มากเป็นพิเศษ เพื่อลดปัญหาในการดำเนินการสอบให้เหลือน้อยที่สุด

กล่าวโดยสรุปแล้ว เป็นหน้าที่และความรับผิดชอบโดยตรงของฝ่ายทดสอบที่จะพยายามหาทางตีแผ่ผลสัมฤทธิ์ทางด้านต่าง ๆ ของผู้เข้าสอบออกมามากที่สุดเท่าที่จะทำได้ ดังนั้นการดำเนินการสอบจึงต้องไม่ทำให้ผู้สอบถูกรบกวนด้วยปัญหาอื่นใดที่จะมีผลต่อการทำข้อสอบ นั่นคือพยายามให้เกิดปัญหาต่าง ๆ น้อยที่สุด ผู้สอบคนหนึ่งลุกขึ้นถามเพียงประโยคเดียวก็มีผลต่อการทำข้อสอบของทุกคนทั้งห้อง เช่น รบกวนสมาธิ ทำให้ขุ่นใจ ฯลฯ ถ้าเกิดปัญหาขึ้นแม้แต่เพียงกึ่งก้อยนับได้ว่าเป็นข้อบกพร่องของผู้วางแผนออกข้อสอบ หรือฝ่ายทดสอบทั้งสิ้น จะเห็นได้ว่าคุณภาพของการดำเนินการสอบว่าจะดีเลิศ-เลวยิ่งเพียงใด ขึ้นอยู่กับแบบของข้อสอบมิใช่น้อย ผู้วางแผนจึงต้องพิจารณาให้เหมาะสมว่าควรจะใช้แบบของข้อสอบแบบใดจึงจะดี

นอกจากเกณฑ์ดังกล่าวแล้วยังมีปัญหาที่ถกเถียงกันมากอีกประการหนึ่งก็คือ ในข้อสอบฉบับเดียวกันควรมีแบบข้อสอบหลายแบบหรือไม่ ฝ่ายที่ต้องการให้มีหลายแบบก็กล่าวอ้างด้วยเหตุผลสำคัญ 2 ประการ คือ

1) ถ้ามีหลายแบบในฉบับเดียวกันแล้ว จะทำให้ผู้สอบสนใจดีกว่าแบบเดียวตรงที่ไม่น่าเบื่อ ได้เปลี่ยนบรรยากาศไปเสียบ้าง ไม่ซ้ำซากอยู่แบบเดียว

2) แบบของข้อสอบแต่ละหัวข้อแต่ละตอน จะเหมาะกับข้อสอบแต่ละแบบเท่านั้น จึงเห็นควรให้มีการยืดหยุ่นในการใช้แบบของข้อสอบในแต่ละตอน

แต่อย่างไรก็ตามส่วนใหญ่จะเห็นพ้องต้องกันว่า การผูกความสนใจของผู้สอบไว้กับข้อสอบนั้น คุณภาพของข้อสอบต่างหากที่มีอิทธิพลมากกว่าลักษณะภายนอก (หลายแบบ) ที่ลวงตา และแบบทดสอบบางแบบ เช่น แบบเลือกตอบสามารถใช้ได้กับเนื้อหาและจุดมุ่งหมายของหลักสูตรทุกตอนและทุกระดับ ก็ย่อมจะเหมาะสมกว่าที่ใช้มากแบบ จึงอาจสรุปได้ว่าถ้าจะใช้มากกว่า 1 แบบ ก็ไม่ควรมีมากแบบจนไปมีอิทธิพลต่อการดำเนินการสอบ และประสิทธิภาพของการตอบข้อสอบ อันจะมีผลต่อความเที่ยงตรงและความเชื่อมั่นของแบบทดสอบ ซึ่งเป็นเรื่องใหญ่กว่า กล่าวคือถ้าข้อสอบฉบับใด เดี่ยวให้เติมคำ ให้ตอบสั้น ๆ เดี่ยวให้จับคู่สลับกันเป็นโกลาหลแล้ว ก็จะมีปัญหาต่อการตรวจให้คะแนน ผู้สอบก็สับสน จึงสรุปได้ว่ายิ่งน้อยแบบยิ่งดี

2. การกำหนดความยาวของแบบทดสอบ

ข้อทดสอบแต่ละข้อที่รวมกันขึ้นเป็นฉบับนั้น โดยแท้จริงแล้วมีสภาพเป็นเพียง “กลุ่มตัวอย่าง” ของผลการศึกษาล่าเรียนในกระบวนวิชานั้น ๆ ซึ่งเป็น “ประชากร” ของความรู้ การที่เราสอบกันทุกวันนี้จึงเป็นการสอบ “ตัวอย่าง” ของความรู้ ถ้าสอบ “ประชากร” ของเนื้อหาแล้วอาจต้องใช้เวลาสอบเกินกว่า 1 วัน เหมือนสอบจอหงวน ถึงกับต้องส่งข้าวส่งน้ำกัน และก็ต้องสอบกันหลายพันข้อ จึงจะวัดได้ทั่วว่ามีความรู้ในเรื่องนั้น ๆ จริงหรือไม่ แต่การทำเช่นนั้นจะทำให้เสียเงินเสียแรงงานและเวลาไม่คุ้มค่าที่ทุ่มไป จึงต้องเลือกเอาตัวแทนของความรู้มาสอบมาวัดกันอย่างมากก็ไม่เกิน 200 ข้อ ใช้เวลา 1 ชั่วโมง หรือ 2 ชั่วโมงก็เสร็จ ปัญหาจึงมีอยู่ว่า กลุ่มตัวอย่างความรู้ที่ว่าควรมีจำนวนเท่าไรจึงจะเหมาะ จึงจะเป็นตัวแทนประชากรความรู้ทั้งสิ้นทั้งปวงได้ดี หลักสำคัญมีว่าจะใช้สอบมากน้อยข้ออย่างไรไม่สำคัญ ขอเพียงได้ข้อสอบมีความเชื่อมั่นและความเที่ยงตรงสูงก็ใช้ได้ เราจึงต้องมาพิจารณาว่าความเชื่อมั่นและความเที่ยงตรงของข้อสอบมีความเกี่ยวข้องสัมพันธ์กับความยาว-สั้น ของข้อสอบอย่างไร ดังนี้

1) **ข้อสอบที่มีความเที่ยงตรงทางเนื้อหา** หมายถึงต้องถามวัดครอบคลุมเนื้อหาทั้งหมด ในทางปฏิบัติจึงเน้นและเลือกเอาเฉพาะหัวข้อสำคัญที่คาดว่าจะเป็นตัวแทนประชากรในด้านเนื้อหาได้ ส่วนในทางจุดมุ่งหมายของการศึกษาก็หมายถึงคลุมพฤติกรรมที่มุ่งจะให้เกิดขึ้นจากการฝึกฝนด้านเนื้อหาวิชาดังกล่าว เราก็เลือกตัวแทนพฤติกรรมมาถาม จึงเห็นได้ว่าความยาวสั้นของข้อสอบจะขึ้นอยู่กับ “ตัวอย่าง” ที่สำคัญๆ ทางด้านเนื้อหาและพฤติกรรมที่เราตัดสินใจคิดว่ามัน “สำคัญ” และเป็น “ตัวแทน” ได้ ซึ่งสามารถดูได้จากตารางวิเคราะห์หลักสูตรที่กำหนดน้ำหนักของเนื้อหาและพฤติกรรมไว้

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

เนื่องจากจำนวนข้อมากหรือน้อยเกี่ยวข้องกับโดยตรงกับเวลาที่ให้ทำ ถ้ามากข้อขึ้นเวลาก็ต้องมากขึ้นตามไปด้วย ถ้ามีเวลาจำกัด เช่น 1-2-3 ชั่วโมง ก็ต้องกำหนดให้จำนวนข้อสอบพอเหมาะพอควรกับเวลา โดยสภาพความเป็นจริงแล้วไม่มีใครรู้ว่าเวลาที่เหมาะสมของข้อสอบแต่ละข้อควรเป็นเท่าไร เพราะขึ้นอยู่กับลักษณะเนื้อหา ขบวนการคิด ความยาก ฯลฯ มากมายจนนับไม่ถ้วน แต่ก็มีหลักพิจารณากลาง ๆ ว่าในการสอบแต่ละครั้งควรมีผู้ทำเสร็จทุกข้อประมาณ 90% จากจุดนี้เองเราก็สามารถกะประมาณจำนวนข้อและเวลาได้เหมาะสม หลังจากได้นำแบบทดสอบไปทดลองเสียก่อน (ถ้าเป็นข้อสอบที่จะพัฒนาเป็นข้อสอบมาตรฐานหรือข้อสอบมีคุณภาพสูงทั่ว ๆ ไป)

โดยทางปฏิบัติแล้วเรากำหนดจำนวนข้อมาก่อน แล้วค่อยมาประมาณเวลาที่เหมาะสมซึ่งมักจะกำหนดเวลาให้มากไว้ เพราะถือว่าเราจะตัดสินใจว่าเรามีความรู้ปานใดก็ต้องให้เขาได้ทำข้อสอบจนหมดทุกข้อเสียก่อน ถ้าเวลาน้อยเกินไปทำไม่ทัน ได้คะแนนมาน้อยจะว่าเรามีความรู้น้อยก็ไม่ได้ เพราะถ้าหากให้เวลามากเขาจะทำข้อสอบได้มากกว่านั้น

ในการสร้างแบบทดสอบที่ดี การกำหนดความยาวของข้อสอบและเวลาที่เหมาะสมจะได้จากข้อมูลที่ทำกรทดลองข้อสอบ (Pretest) นี้เอง โดยกำหนดเวลาและจำนวนข้อไปก่อนโดยอาศัยประสบการณ์แล้วให้จดเวลาที่นักเรียน 90% ทำข้อสอบนั้นเสร็จเอาไว้ (เช่น 60 ข้อ คน 90% ใช้เวลา 60 นาที แสดงว่าใช้เวลาประมาณข้อละ 1 นาที) จึงมากะประมาณว่าถ้าสอบจริงมีข้อสอบกี่ข้อ จะใช้เวลาสักเท่าไร หรือมีเวลาอยู่จำกัด 1 ชั่วโมง จะใช้ข้อสอบสักกี่ข้อจึงจะดี

3. การกำหนดช่วงเวลาที่ใช้ในการทดสอบ

ในหัวข้อก่อนได้กล่าวถึง ความสัมพันธ์ระหว่างจำนวนข้อกับจำนวนเวลาที่ใช้ให้เหมาะสม ทำไมจึงต้องให้เหมาะสม ถ้าไม่เหมาะสมแล้วจะเกิดอะไรขึ้น เช่น ให้นเวลาน้อยเกินไปทำได้นิดเดียวหมดเวลา แสดงว่าผู้สอบได้ทำข้อสอบเพียงบางส่วนของข้อสอบ หรือได้ทดสอบบางส่วนของความรู้ตอนที่ทำไม่ทันก็คือตอนที่ไม่ได้ทดสอบ จะพบว่าเมื่อทำไม่ครบทั้งฉบับแทบทุกคนก็จะกลายเป็นข้อสอบฉบับนั้น (แค่ตรงที่ทำไปถึง) ไม่มีความเที่ยงตรงตามเนื้อหาและพฤติกรรม เมื่อเวลาไม่พอก็ก่อให้เกิดการเสียหายเดาเอา ทำให้คะแนนที่ได้เชื่อถือไม่ได้ นี่เป็นเหตุผลที่ว่า “ทำไมจึงต้องใช้เวลาที่เหมาะสม” ถ้าหากให้เวลามากเกินไปจะมีผลเช่นไร ถ้าหากข้อสอบไม่ยากเกินไปนักคนเก่งตอบได้มาก คนเก่งน้อยกว่าจะพลาดตอบได้มากไปด้วย เหตุด้วยมีเวลาเหลือเพื่อที่จะพิจารณาอย่างถี่ถ้วนทำให้มีผลต่ออำนาจจำแนกของข้อสอบว่าจำแนกคนออกเป็นลำดับ เก่ง-อ่อน ได้ไม่ละเอียด คะแนนก็ด้อยความหมายลงไป จุดนี้เองเป็นที่มาของแบบทดสอบชนิด Speeded Test ที่ว่าถ้าให้เวลาน้อย ๆ ทำข้อสอบที่ไม่ยากนัก คนเก่งกว่าจะทำได้มากกว่าและถูกต้องมากกว่า จำแนกคนเป็นอันดับลดหลั่นได้ละเอียดละออดีกว่าให้เวลามาก ๆ กล่าวได้ว่า “เวลาที่เหมาะสม” ก็คือ “เวลาที่จะทำให้ข้อสอบฉบับนั้นได้ค่าความเชื่อมั่นและความเที่ยงตรงสูงสุด” ฉะนั้น ความยาวของข้อสอบจึงขึ้นอยู่กับ

ก. จำนวนเนื้อหาและพฤติกรรมสำคัญที่จะวัด

ข. เวลาที่จะใช้ในการตอบข้อสอบฉบับนั้น

4. กำหนดจำนวนข้อที่จะใช้ในการทดลองสอบ

จำนวนข้อสอบที่ใช้ในการทดลองสอบ (Pretest) ตามปกติจะต้องสร้างไว้ให้มากกว่าที่จะสร้างเป็นฉบับจริง เพราะเมื่อนำไปทดลองสอบแล้ววิเคราะห์ จะปรากฏว่าจะมีข้อสอบอยู่จำนวนหนึ่งที่ใช้ไม่ได้ต้องโยนทิ้งตะกร้าไป ชี้นำไปใช้ก็รังแต่จะทำให้ข้อสอบชุดนั้นเลวลง จำนวนข้อสอบที่ใช้ไม่ได้นี้มีมากน้อยต่างกันไปขึ้นอยู่กับความรู้และประสบการณ์ของผู้สร้างเป็นสำคัญบางทีก็อยู่ในช่วง 10-20% บางทีก็มีมากถึง 50%

5. กำหนดความยากของข้อสอบ

การจะกำหนดความยากของข้อสอบอย่างไร ขึ้นอยู่กับจุดมุ่งหมายของการทดสอบเป็นสำคัญ เช่น ถ้าสอบเพื่อประเมินผลรวมก็ควรใช้ข้อสอบที่มีความยากง่ายปานกลาง ถ้าสอบเพื่อการคัดเลือกก็ต้องใช้ข้อสอบที่ค่อนข้างยาก เป็นต้น

5.3 กำหนดเนื้อหาของแบบทดสอบ

การสอบว่าจะสอบอะไร อย่างไร อย่างไร นั้นเองถ้าหากมีตารางวิเคราะห์หลักสูตรมาปรับให้เข้ากับจำนวนข้อของแบบทดสอบก็เป็นอันใช้ได้ จะขอยกตัวอย่างตารางวิเคราะห์ของวิชาความรู้รอบตัวที่จะนำมาใช้ออกข้อสอบ 100 ข้อ ดังนี้

เนื้อหา	จุดมุ่งหมาย				รวม
	ความรู้ ความจำ	ความ เข้าใจ	การ นำไปใช้	การ วิเคราะห์	
1. สัตว์เลี้ยง	4	1		1	6
2. ของใช้ในครัว	4	3	1		8
3. ครอบครัวยุคนี้	3	4	2	1	10
4. การจัดประเภทสัตว์	2	5	4		11
5. พืช	3	3	5	1	12
6. ดอกไม้	2	4	1	5	12
7. ยานพาหนะ	2	3	5	2	12
8. ที่อยู่อาศัย	2	8	4	2	16
9. ทรัพยากรธรรมชาติ	3	4	5	1	13
รวม	25	35	27	13	100

จากตารางวิเคราะห์ข้างบนนี้ จะเห็นว่าถ้ามีข้อสอบ 100 ข้อ จะถามเน้นมากที่สุดในเรื่องความเข้าใจและการนำไปใช้ ความรู้ความจำ และการวิเคราะห์น้อยลงเป็นลำดับ ถ้าตารางนี้ปรับมาจากตารางวิเคราะห์หลักสูตรที่ใช้เป็นแนวทางในการสอนก็แสดงว่า แต่ละหัวข้อมีเน้นความสำคัญต่างกัน และในหัวข้อทั้งปวงก็เน้นให้เกิดพฤติกรรมแบบต่างกันไป

จากที่กล่าวมาข้างต้นจะเห็นว่า คุณภาพข้อสอบจะวัดได้ลึกได้ตื้นมีสัดส่วนพอเหมาะพอควรหรือไม่ แบบทดสอบชุดนั้นจะมีโครงสร้างที่ดีหรือไม่ขึ้นอยู่กับตารางวิเคราะห์หลักสูตรอยู่ไม่น้อย ไม่ว่าผู้ออกข้อสอบเก่งกล้าสามารถปานใดก็ตามถ้าไม่ได้มีตารางวิเคราะห์นี้ก็เหมือนคนตาบอดคลำช้าง ข้อสอบสร้างมาจะมีลักษณะอย่างไรก็ไม่สามารถบอกได้แน่นอน

5.4 หมายกำหนดการเกี่ยวกับการสร้างแบบทดสอบ

เพื่อให้แผนดำเนินไปอย่างราบรื่น ควรต้องวางหมายกำหนดการไว้ในแน่นอนและยืดหยุ่นได้ว่าจะทำอะไร เมื่อไร งานแต่ละขั้นแต่ละตอนจะใช้เวลาสักเท่าไร จะเริ่มงานขั้นต่อไปได้เมื่อไร จึงขออธิบายถึงอันดับขั้นและลักษณะของงานแต่ละลำดับขั้นไว้เพียงสังเขปเพื่อจะเป็นแนวในการพิจารณาว่าจะทำอะไรเมื่อไรดังนี้

คำถามท้ายบทที่ 5

1. การสร้างข้อสอบปรนัยมีขั้นตอนการสร้างอย่างไร
2. การกำหนดจุดมุ่งหมายของการสร้างข้อสอบปรนัย มีประโยชน์อย่างไร จงอธิบาย
3. ให้ท่านบอกเกณฑ์ที่ใช้ในการพิจารณาแบบของข้อสอบว่ามีเกณฑ์อะไรบ้าง พร้อมทั้งยกตัวอย่างประกอบคำอธิบาย
4. ให้ท่านอธิบายความหมายของประโยคข้างล่างนี้
 - 4.1 เวลาที่เหมาะสมสำหรับการทดสอบ คือ เวลาที่จะทำให้ข้อสอบฉบับนี้มีความเชื่อมั่นและความเที่ยงตรงสูงสุด
 - 4.2 แบบทดสอบจะมีความยากง่ายระดับใดขึ้นอยู่กับจุดมุ่งหมายของการนำผล การสอนไปใช้