



ความเป็นมาของปัญหา

ภาระหน้าที่ผู้เป็นครูจะต้องปฏิบัตินั้นมีหลายด้าน แต่ที่สำคัญและเป็นงานหลักที่ครูจะต้องปฏิบัติเป็นประจำมีอยู่ 2 ด้าน คือ การสอน และการทดสอบ เพื่อวัดผลการสอนของครู และวัดผลการเรียนของนักเรียนว่า การสอนของครูมีข้อดีหรือข้อเสียอย่างไร และนักเรียนมีความรู้หรือเจตคติออกตามความมุ่งหมายในการสอนหรือไม่ เพียงใด

1. ความหมายของการทดสอบ

ผู้เชี่ยวชาญทางด้านการวัดผลหลายท่านได้อธิบายความหมายของการทดสอบไว้หลายด้านด้วยกัน แต่ผู้เขียนจะขอยกมากล่าวพอสังเขป ดังนี้

Glenn M. Blair กล่าวว่า "การทดสอบเป็นการสอนชนิดหนึ่ง"¹

Donald D. Paige กล่าวว่า "การทดสอบเป็นเครื่องมือที่ก่อให้เกิดการเรียนรู้ได้คือพอ ๆ กับเป็นเครื่องมือสำหรับการประเมินผล"²

Robert L. Thorndike กล่าวถึงความสำคัญของการทดสอบที่มีต่อการเรียนรู้ว่า "แบบทดสอบที่สร้างอย่างดีและใช้อย่างได้ผล สามารถใช้เป็นแรงจูงใจในการสร้างนิสัยในการเรียนที่ดี แก่ข้อผิดพลาดใหญ่ๆ และเป็นแนวทางที่จะนำไปสู่ผลสัมฤทธิ์

¹Glenn M. Blair, R. Stewart Jones, and Ray H. Simpson, Educational Psychology (2d ed.; New York: The Macmillan Co., 1962), p. 272.

²Donald D. Paige, "Learning While Testing," Journal of Educational Research, 95 (Feb., 1966), pp. 276-277.

ในการเรียนรู้ได้ แต่ในทางตรงกันข้ามถ้าแบบทดสอบนั้นสร้างขึ้นอย่างไม่ดี ไม่มีหลักเกณฑ์ ก็ย่อมไม่สัมฤทธิ์ผลตามความปรารถนา ขบวนการทดสอบสามารถควบคุมประสิทธิภาพของ ขบวนการเรียนรู้ บางทีอาจกล่าวได้ว่า คือว่าเครื่องมือการสอนชนิดใด ๆ เสียอีก"³

ดร. ขวาล แพร์ตกุล กล่าวว่า "การวัดผลหรือการทดสอบเป็นขบวนการที่จะให้ค่าซึ่งจำนวนหรือปริมาณ เพื่อจะนำไปสู่การประเมินค่า ดีราคาผลสัมฤทธิ์ของการ ศึกษาทั้งครูและเด็ก"⁴

ตามที่ได้อธิบายมาแล้วจะเห็นได้ว่า การทดสอบที่คือนั้นเป็นทั้งการสอน เครื่องมือ ที่ก่อให้เกิดการเรียนรู้ แรงจูงใจในการสร้างนิสัยในการเรียนที่ดี ตลอดจนเป็นเครื่องมือในการประเมินผลการสอนการเรียนของทั้งครูและเด็ก จากการศึกษาเราจะได้ทราบข้อ บกพร่องต่าง ๆ ของการเรียนการสอนและจะได้แก้ไขข้อบกพร่องนั้น เพื่อนำไปสู่ผลสัมฤทธิ์ ในการเรียนรู้ในที่สุด

2. ประเภทของแบบทดสอบ

การทดสอบ เป็นวิชาที่มีหลายสาขาและมีแบบทดสอบหลายประเภทด้วยกัน เช่น แบบทดสอบทางจิตวิทยา แบบทดสอบทางการศึกษา แบบทดสอบการใช้วัยระต่าง ๆ ของ ร่างกายในวิชาพลานามัย เป็นต้น แต่ในที่นี้จะกล่าวถึงแบบทดสอบประเภทที่ใช้ในโรงเรียน ซึ่งเกี่ยวกับหน้าที่ของครูโดยตรงแต่เพียงประเภทเดียว

แบบทดสอบในโรงเรียน แบ่งออกเป็นแบบทดสอบ 3 ประเภทใหญ่ ๆ คือ

- 2.1 แบบทดสอบผลสัมฤทธิ์ (Achievement Test) แบ่งออกเป็น
 - ก. แบบทดสอบที่ครูสร้างขึ้นเอง (Teacher - made test)
 - ข. แบบทดสอบมาตรฐาน (Standardized test)

³Robert L. Thorndike, and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education (3d ed.; New York: Wiley International. Edition, 1969), p. 27.

⁴ขวาล แพร์ตกุล, เทคนิคการวัดผล (พิมพ์ครั้งที่ 5 ฉบับปรับปรุง; พระนคร: โรงพิมพ์วัฒนาพานิช, 2516), หน้า 21.

2.2 แบบทดสอบความถนัด (Aptitude test) แบ่งออกเป็น

- ก. ความถนัดทางการเรียน (Scholastic Aptitude test)
- ข. ความถนัดจำเพาะ (Specific Aptitude test)

2.3 แบบทดสอบบุคคลกับสังคม (Personal - social test)

หรือการปรับคน แบ่งเป็น

- ก. ทศินคติ (Attitude)
- ข. ความสนใจ (Interest) และอื่น ๆ⁵

แม้ว่าแบบทดสอบที่ใช้ในโรงเรียนจะแบ่งออกเป็นหลายประเภทด้วยกัน แต่เนื่องจากขอบข่ายของการวิจัยเรื่องนี้ครอบคลุมเฉพาะแบบทดสอบผลสัมฤทธิ์ (Achievement test) จึงจะไม่กล่าวถึงแบบทดสอบชนิดอื่น จะกล่าวแต่เฉพาะแบบทดสอบผลสัมฤทธิ์เท่านั้น

3. แบบทดสอบผลสัมฤทธิ์

มีผู้ให้คำจำกัดความของแบบทดสอบผลสัมฤทธิ์ไว้มากมายหลายความหมาย ตั้งแต่ความหมายที่กว้างมาก กล่าวคือ หมายถึงแบบทดสอบที่ใช้วัดผลตั้งแต่ชั้นอนุบาลขึ้นมา จนถึงชั้นอุดมศึกษาและโรงเรียนอาชีพต่าง ๆ หรือหมายถึงแบบทดสอบที่ใช้วัดความพร้อมในการเรียน แบบทดสอบเพื่อวิเคราะห์ขอบกพร่อง หรือเพื่อทำนายผลการเรียนในอนาคตก็ได้ ตามที่ W.A. Mehrens และ I.J. Lehmann ได้กล่าวไว้ว่า

Achievement tests run the gamut from readiness and diagnostic tests to prognostic content-oriented tests (be they single subject matter tests, single survey tests, or survey batteries), and from pre-school to graduate and professional school.⁶

⁵ เรื่องเดียวกัน, หน้า 111.

⁶ William A. Mehrens, and Irvin J. Lehmann, Standardized Tests in Education (New York: Holt, Rinehart and Winston, Inc., 1969), p. 187.

นอกจากนี้ Robert M.W. Travers ยังกล่าวว่า แบบทดสอบผลสัมฤทธิ์ นั้น ในสมัยโบราณใช้วัดความรู้ทางค่านเนื้อหาวิชาแต่เพียงอย่างเดียว แต่ในปัจจุบันนี้คำว่า ผลสัมฤทธิ์ครอบคลุมถึงการวัดความสนใจ ทศนคติ ความรู้จักคุณค่า ตลอดจนการปรับตัวทางสังคมด้วย

... the kinds of achievement tests which psychologists have built have been subject - matter tests. However, the term achievement has a much broader meaning to the teacher, and refers to the learning of interests, attitudes, appreciations, and social adjustments as well as to the learning of facts.⁷

ส่วน ดร. ขวาล แพทย์กุล ให้ความหมายของแบบทดสอบผลสัมฤทธิ์ไว้ดังนี้

แบบทดสอบผลสัมฤทธิ์ หมายถึงแบบทดสอบที่วัดความรู้ ทักษะ และสมรรถภาพสมองคนต่าง ๆ ที่เด็กได้รับจาก ประสบการณ์ทั้งปวง ทั้งจากโรงเรียนและทางบ้าน ยกเว้น การวัดทางร่างกาย ความถนัด และบุคลิกนิสัยสังคม หรือ พุทธิคุณ ๆ ก็คือ แบบทดสอบประเภทมุ่งที่จะวัดว่าเด็ก สามารถนำความรู้ไปใช้ในสถานการณ์ต่าง ๆ ไหมมากน้อยและเฉลว คลองเพียงใด วัดว่าเด็กรู้มาแล้วเท่าใด อันเป็นเรื่องราว ของอดีตกาล⁸

ส่วนความหมายของแบบทดสอบผลสัมฤทธิ์ที่แคบลงนั้น หมายถึงแบบทดสอบ ที่ต้องการวัดว่า ผู้เรียนได้เรียนรู้เนื้อหาวิชาหรือทักษะเฉพาะด้านใดด้านหนึ่งจากการสอน เพียงใด ดังที่ D.J. Brown ได้กล่าวไว้ว่า

⁷Robert M.W. Travers, How to Make Achievement Tests

(New York: The Odyssey Press, 1950), p. 2.

⁸ขวาล แพทย์กุล, เรื่องเดิม, หน้า 112.

An achievement test is a test that measures the extent to which a person has 'achieved' something -- acquired certain information or mastered certain skill,⁹ usually as a result of specific instruction

หรือหมายถึงการวัดสัมฤทธิ์ผลจากการเรียนวิชาใดวิชาหนึ่งที่ผ่านมาแล้ว "An achievement test is designed to indicate degree of success in some past learning activity"¹⁰

ส่วน M.W. Richardson และ J.M. Stalnaker สองผู้นำคนสำคัญในการพัฒนาแบบทดสอบผลสัมฤทธิ์แบบใหม่ กล่าวว่า

... an achievement examination need not have direct pedagogical value in itself ... The purpose of achievement examining is essentially measurement for certifying academic credit ... An achievement examination is good or valid if it simply does a good job of measuring what it purports to measure¹¹



จุดมุ่งหมายที่สำคัญของการทดสอบผลสัมฤทธิ์นั้น คือการกำหนดผลการเรียนทางวิชาการ แบบทดสอบผลสัมฤทธิ์ที่ดี คือแบบทดสอบที่สามารถวัดสิ่งที่ต้องการจะวัดได้อย่างถูกต้อง

⁹Donald J. Brown, Appraisal Procedure in the Secondary Schools (Englewood Cliffs; New Jersey: Prentice-Hall, Inc., 1970), p. 153.

¹⁰Norman E. Gronlund, Measurement and Evaluation in Teaching (5th ed., New York: The Macmillan Company, 1967), p. 10.

¹¹Robert L. Ebel, Essentials of Educational Measurement (Englewood Cliffs; New Jersey: Prentice-Hall, Inc., 1972), p. 43, quoting M.W. Richardson and J.M. Stalnaker, "Comments on Achievement Examinations," Journal of Educational Research, Vol. 28 (1935), pp. 425-32.

จะเห็นได้ว่า แบบทดสอบผลสัมฤทธิ์ไม่มีความหมายตั้งแต่กว้างมากลงมาถึงจำเพาะเจาะจง อาจจะเป็นแบบทดสอบที่สร้างขึ้นเพื่อวัดผลสัมฤทธิ์ทางการเรียนวิชาใดวิชาหนึ่ง หรือวัดความพร้อมของผู้เรียนเพื่อจะสอนเรื่องอื่นต่อไป หรือเพื่อวิเคราะห์ข้อบกพร่องของการเรียนของเด็กหรือการสอนของครู ในวงการศึกษายุคใหม่ ทุกสาขาวิชาได้ หันมายอมรับแล้วแต่วัตถุประสงค์ของผู้สร้างว่าต้องการจะใช้แบบทดสอบเพื่ออะไร ก็จะสร้างแบบทดสอบตามวัตถุประสงค์นั้น ๆ หมายความว่า แบบทดสอบผลสัมฤทธิ์นี้จะเป็นแบบใดแบบหนึ่งที่ถูกกล่าวมาแล้วได้ทั้งสิ้นตามวัตถุประสงค์เฉพาะของผู้สร้างแบบทดสอบเป็นคราว ๆ ไป

4. ลำดับขั้นในการเตรียมการสร้างแบบทดสอบผลสัมฤทธิ์

เนื่องจากการวิจัยเรื่องนี้เกี่ยวกับการสร้างแบบทดสอบผลสัมฤทธิ์ ฉะนั้นจึงจะกล่าวถึงวิธีการเตรียมการสร้างแบบทดสอบชนิดนี้ตามหลักมาตรฐานพอสังเขป เพราะเป็นเรื่องที่สำคัญที่สุดที่ผู้สร้างแบบทดสอบผลสัมฤทธิ์จะต้องเรียนรู้เสียก่อน เพื่อที่จะได้ปฏิบัติตามได้ถูกต้อง

4.1 Robert M.W. Travers ได้กำหนดการเตรียมการสร้างแบบทดสอบผลสัมฤทธิ์ไว้ดังนี้

1. ระบุวัตถุประสงค์ทั่วไปของการศึกษาวิชานั้น โดยมากมักจะตั้งตามหลักสูตรซึ่งมักจะบ่งไว้อย่างกว้าง ๆ ผู้เขียนแนะนำว่า การเขียนวัตถุประสงค์ทั่วไปที่ดีควรมีจำนวน 8 - 15 ข้อ

2. ตั้งวัตถุประสงค์เฉพาะที่ต้องการจะวัดสัมฤทธิ์ผลของบุคคลหนึ่ง ๆ ในการที่จะนำสิ่งที่ได้เรียนไปประยุกต์ใช้อย่างกว้างขวาง หรือตั้งเกณฑ์ในการวัดขึ้นนั่นเอง

3. จัดลำดับความสำคัญมากน้อยของวัตถุประสงค์ที่จะวัด โดยให้คลุมเนื้อหาทั้งหมด

4. เขียนโครงร่างของเนื้อหาที่จะทดสอบ ควรเขียนเป็นบท เพื่อจะได้เขียนให้ได้ตามวัตถุประสงค์ที่จะวัดในแต่ละบทนั้น

5. สร้างตารางวิเคราะห์ (test blueprint หรือ table of specifications) เป็นตารางที่แสดงความสัมพันธ์ระหว่างวัตถุประสงค์กับเนื้อหาที่จะ

ทดสอบ การวางวิเคราะห์นั้นจะทำให้ผู้สร้างข้อสอบมั่นใจได้ว่า แบบทดสอบของเขาจะวัดเนื้อหาในหลักสูตรได้อย่างสมมูลย์กับวัตถุประสงค์ที่ตั้งไว้

6. กำหนดสิ่งที่อาจเป็นปัญหาสำหรับผู้เข้าสอบ หลังจากทำข้อสอบนั้นแล้ว
7. บ่งถึงองค์ประกอบอื่น ๆ ที่อาจมีผลกระทบสะท้อนต่อขบวนการทดสอบ
8. เลือกเครื่องมือที่จะใช้วัดว่าจะใช้ข้อสอบประเภทใด เช่น แลสถูกผิด

แบบเลือกตอบ หรือแบบอัตนัย¹²

4.2 William J. Micheels และ M. Ray Karnes ได้กำหนดลำดับขั้นในการสร้างข้อทดสอบไว้ว่า

1. สำนวความมุ่งหมายและบันทึกพฤติกรรมจากความมุ่งหมายนั้น
2. สำนวเนื้อหาวิชาที่สอนตามความมุ่งหมายนั้น
3. ให้คำจำกัดความพฤติกรรมที่สำรวได้จากความมุ่งหมาย และเลือกเฉพาะพฤติกรรมที่เด็กสามารถปฏิบัติได้จริง ๆ
4. สร้างตารางวิเคราะห์หลักสูตร และวัตถุประสงค์ที่จะทำการทดสอบ
5. สร้างคำถามวัดพฤติกรรมนั้น ๆ
6. ถ้าข้อสอบนั้นเป็นตอน ๆ ก็เอาตอนต่าง ๆ เหล่านี้มารวมกันเป็น

ชุดเดียว

7. เขียนคำสั่งและคำชี้แจงแต่ละตอนให้ชัดเจน
8. ตรวจสอบกรรอกอีกครั้งหนึ่ง
9. ใ้ผู้มีความรู้เรื่องการสร้างข้อทดสอบวิพากษ์วิจารณ์
10. ทำเฉลยไว้ให้เรียบร้อย
11. ตรวจสอบทวนหลาย ๆ ครั้ง

¹²Travers, op. cit., pp. 14-29.

12. นำข้อสอบไปทดสอบ
13. วิเคราะห์ข้อสอบหลังจากการสอบแล้ว¹³

วิธีการสร้างแบบทดสอบผลสัมฤทธิ์ตามที่ผู้เชี่ยวชาญการทดสอบได้กล่าวไว้แล้วมีหลักสำคัญที่คล้ายคลึงกัน กล่าวคือ ประการแรกผู้สร้างจะต้องตั้งวัตถุประสงค์ทั่วไป และวัตถุประสงค์เฉพาะของการทดสอบ ต่อจากนั้นจึงเขียนโครงร่างของเนื้อหาวิชาที่จะทดสอบและแบ่งเนื้อหาวิชานั้นโดยการสร้างตารางวิเคราะห์ (Table of Specifications) ตารางวิเคราะห์นี้ประกอบด้วยวัตถุประสงค์กับเนื้อหาต่าง ๆ ที่ต้องการจะวัด การสร้างตารางวิเคราะห์รายการนี้เป็นการหาความสัมพันธ์ระหว่างจุดประสงค์ในการทดสอบ เนื้อหาวิชา และจำนวนข้อคำถามของแบบทดสอบ ตารางนี้จะแสดงให้เห็นว่า ครูควรจะออกข้อสอบตามเนื้อหาแต่ละเรื่องสักกี่ข้อ ซึ่งทั้งนี้ขึ้นอยู่กับวัตถุประสงค์ของแบบทดสอบนั้นว่าต้องการวัดเนื้อหาหนักไปด้านใด รวมทั้งเวลาที่สอนในแต่ละเรื่องนั้นด้วย

Marshall และ Hales กล่าวว่า วัตถุประสงค์ของการสอนวิชาหนึ่ง ๆ นั้น แจกแจงได้ยาก แต่การแจกแจงวัตถุประสงค์แบบที่มีผู้นิยมใช้กันมากที่สุดคือ การแจกแจงของ Bloom และเพื่อนร่วมงาน ใน Taxonomy of Educational Objectives ซึ่งแบ่งวัตถุประสงค์ในการสอนไว้เป็น knowledge, comprehension, application, analysis, synthesis และ evaluation (ดูรายละเอียดในหน้า 12 - 14) โดยอธิบายว่า ในการสร้างข้อสอบอาจสรุปได้เป็น 3 รายการใหญ่ ๆ คือ ด้านความรู้ (knowledge) ความเข้าใจ (understanding) และการนำไปใช้ (application) ก็ได้ ส่วนอัตราส่วนในการวัดวัตถุประสงค์ทั้ง 3 ด้านนี้ ในแบบทดสอบครูอาจแบ่งวัดทางด้านความรู้

¹³William J. Micheels & M. Ray Karnes, Measuring Educational Achievement (New York: McGraw-Hill Book Company, Inc., 1950), pp. 126-129.

(knowledge) 30 % ความเข้าใจ (understanding) 40 % และการนำไปใช้ (application) 30 %¹⁴

ต่อไปจะได้อธิบายถึงการคำนวณน้ำหนักของเนื้อหาแต่ละด้านที่จะนำมาออกในแบบทดสอบตามอัตราส่วนระหว่างจำนวนเวลา (คาบ) ที่สอนในชั้น และวัตถุประสงค์ที่ตั้งไว้ตามหลักของ Marshall & Hales ดังแสดงไว้ในตารางที่ 1

ตารางที่ 1 แสดงตารางวิเคราะห์รายการในการออกข้อสอบวิชา "Rhythm and Pitch"

Weight	Content	Level of Objectives			Number of items by content area
		Knowledge 30 %	Understanding 40 %	Application 30 %	
10 %	Identifying notes by names	1*	1	1	3
35 %	Combining rhythms to form measures	3	4	3	10
35 %	Placing pitches on music staff	3	5	3	11
20 %	Combining pitches and rhythms	2	2	2	6
Nubmer of items by cognitive classification		9	12	9	30

* Number of items in cell = row weight x column weight x number of test items (for Area 1, Knowledge) = 0.30 x 0.10 x 30 = 0.9 items.

¹⁴Jon Clark Marshall, and Loyde Wesley Hales, Essentials of Testing (Reading; Massachusetts : Addison - Wesley Publishing Company, 1972), p. 8.

เกณฑ์ที่จะช่วยครูในการคำนวณน้ำหนักของเนื้อหาวิชาในข้อสอบมีดังนี้

1) จำนวนเวลาที่สอนเนื้อหาวิชาค่านั้น ๆ ในห้องเรียน (class time devoted to each area) สมมุติว่า เนื้อหาทางค่าน "Identifying notes" ใช้เวลาสอน $\frac{1}{2}$ คาบ ในจำนวน 5 คาบ ของการสอนเรื่อง "Rhythm and Pitch" อาจคือน้ำหนักของเนื้อหาค่านนี้จะออกในข้อสอบได้ดังนี้

$$\frac{.5 \text{ คาบ}}{5 \text{ คาบ}} = .10 \text{ หรือ } 10 \%$$

นอกจากนี้ สมมุติว่า หัวข้อเรื่อง "Combining Rhythms" ใช้เวลาสอน $1\frac{3}{4}$ คาบ น้ำหนักของเรื่องนี้ในข้อสอบจะเป็น

$$\frac{1.75 \text{ คาบ}}{5 \text{ คาบ}} = .35 \text{ หรือ } 35 \%$$

ส่วนเกณฑ์อื่น ๆ ที่อาจนำมาใช้ได้คือ

2) จำนวนเวลาเรียนทั้งหมดสำหรับวิชานั้นของนักเรียน (total student time) และ

3) ระดับซึ่งเนื้อหาวิชาค่านั้น ๆ สัมพันธ์กับการเรียนในอนาคต (degree to which content areas relate to future learnings)

น้ำหนักในเนื้อหาหนึ่ง ๆ (area weights) จะต้องเท่ากับน้ำหนักรวมของวัตถุประสงค์ที่อยู่ภายในเนื้อหานั้น ๆ โดย

ก. น้ำหนักนั้นจะต้องบ่งถึงวัตถุประสงค์ในลำดับขั้นนั้น ๆ

ข. จะต้องบวกน้ำหนักทั้งหมดเพื่อแสดงน้ำหนักรวมของลำดับขั้นนั้น ๆ

สมมุติว่า ใ้คั้งวัตถุประสงค์ทางค่านความรู้ (knowledge) ไว้ 6 ข้อ และแต่ละข้อมีน้ำหนัก 5 % ดังนั้นน้ำหนักทั้งหมดของเนื้อหาค่านความรู้จะเท่ากับ 30 %

(= 6×5) น้ำหนักของวัตถุประสงค์ในการสอนก็จะหาได้โดยวิธีเดียวกับน้ำหนักค่านเนื้อหาวิชา (ตามที่ไดกล่าวมาแล้วกำหนดให้น้ำหนักวัตถุประสงค์ของการวัดค่านความรู้ = 30 % ค่านความเข้าใจ 40 % และค่านการนำไปใช้ 30 %)

หลังจากที่ได้กำหนดน้ำหนักที่ทองการของเนื้อหาแต่ละส่วน ๆ และกำหนดน้ำหนักของวัตถุประสงค์แต่ละด้านตามทฤษฎีปัญญา (cognitive domain) แล้ว สามารถหาน้ำหนักของแต่ละ cell ได้ น้ำหนักนี้คิดได้จากการคูณจำนวนเปอร์เซ็นต์ของแต่ละแถว กับจำนวนข้อในคอลัมน์ที่ตรงกัน ส่วนจำนวนข้อสอบในแต่ละ cell หนึ่ง ๆ นั้น คิดได้โดยการคูณ cell weight กับจำนวนข้อทั้งหมดที่คิดจะสร้างในแบบทดสอบ อย่างไรก็ตาม ครูไม่ยวบยิดหลักเกณฑ์มากนัก เพราะความสำคัญของระดับต่าง ๆ ของกิจกรรมทางสติปัญญาอาจแตกต่างกันไปตามเนื้อหาวิชาได้¹⁵

หลังจากที่ได้สร้างตารางวิเคราะห์รายการแล้ว ผู้สร้างจะต้องเลือกอีกว่าจะใช้แบบทดสอบประเภทใด เช่น แบบปรนัย แบบถูกผิด แบบเลือกตอบ หรือแบบอัตนัย เพื่อให้เหมาะสมกับวิชาหรือจุดมุ่งหมายที่ใช้ทดสอบ ต่อจากนั้นจะต้องเลือกรูปแบบ (format) ของแบบทดสอบ เขียนคำสั่ง คำชี้แจงให้ชัดเจน และกำหนดสิ่งทีอาจเป็นปัญหาสำหรับผู้เข้าสอบ จัดเวลาของการทำข้อสอบให้เพียงพอ ตรวจแก้ไขข้อบกพร่อง ทำเฉลยให้เรียบร้อย แล้วนำข้อสอบไปให้ผู้เชี่ยวชาญตรวจหรือแก้ไข ต่อจากนั้นจึงนำข้อสอบไปทดลองสอบ การทดลองสอบ (pre-test หรือ trail test) นี้ควรทำสัก 2 ครั้งเป็นอย่างน้อย ก่อนที่จะนำไปทดสอบกับกลุ่มตัวอย่างประชากรจริง เพื่อจะได้วิเคราะห์ข้อสอบ แก้ไขข้อบกพร่องให้ที่เสียก่อน สำหรับการวิเคราะห์ข้อสอบนั้น ปัจจุบัน Chung-Teh Fan ได้สร้างตารางวิเคราะห์ข้อสอบ (Item Analysis Table)¹⁶ บอกค่าความยาก และค่าอำนาจจำแนกของข้อสอบไว้อย่างสมบูรณ์ ซึ่งเป็นการอำนวยความสะดวกในการวิเคราะห์ข้อสอบขึ้นมา เมื่อได้แก้ไขข้อสอบ 2 ครั้งแล้วจึงนำข้อสอบไปทดสอบกับกลุ่มตัวอย่างประชากรจริงเป็นครั้งที่ 3 หลังจากการทดสอบทั้ง 3 ครั้งนี้ ข้อสอบนั้นอาจจะเป็นข้อสอบผลสัมฤทธิ์ที่เข้าชั้น

¹⁵ Ibid., pp. 11-13.

¹⁶ Chung - Teh Fan, ตารางวิเคราะห์ข้อสอบ Item Analysis Table พิมพ์ในประเทศไทย โดยได้รับอนุญาตจาก E.T.S. สหรัฐอเมริกา (กรุงเทพฯ: โรงพิมพ์วัฒนาพานิช, 2514.)

มาตรฐาน มีความเที่ยงตรง (validity) ความเชื่อถือได้ (reliability) และ
การนำไปใช้ได้ (usability) ซึ่งเป็นคุณสมบัติที่สำคัญ 3 ประการของแบบทดสอบที่ดี¹⁷
 จากลำดับขั้นในการสร้างแบบทดสอบที่กล่าวมาแล้วจะเห็นได้ว่า ในการสร้างแบบ
 ทดสอบใด ๆ ก็ตาม ผู้สร้างจะต้องคำนึงถึงจุดประสงค์ทั่วไปของการศึกษา และจุดประสงค์
 เฉพาะที่ตั้งขึ้น เพื่อต้องการจะทดสอบในวิชานั้น ๆ ควบคู่กันไป ดังนั้น ต่อไปจึงจะขอยก
 จุดประสงค์ทั่วไปของการศึกษา (Taxonomy of Educational Objectives) ตามที่
 Benjamin S. Bloom และเพื่อนร่วมงานได้จำแนกไว้เป็น 3 ด้าน (domains) ใหญ่ ๆ
 ด้วยกัน และเป็นที่ยอมรับกันทั่วไปในวงการศึกษามา เพื่อเป็นหลักแก่ครูและผู้สร้างแบบทดสอบ
 ที่จะคำนึงถึงในการสอน หรือการทดสอบต่อไป

5. Taxonomy of Educational Objectives แบ่งออกเป็น 3 domains

5.1 Cognitive domain เกี่ยวกับพุทธิปัญญาหรือความรู้ (intellectual outcomes)

5.2. Affective domain เกี่ยวกับความสนใจ ทศนคติ และค่านิยมต่าง ๆ

5.3 Psychomotor domain เกี่ยวกับทักษะประเภทที่ต้องใช้ช่วยยวះในร่างกายในการเรียน (Motor skills)

พุทธิปัญญาหรือความรู้ใน Cognitive domain นั้น ได้แบ่งออกเป็นชั้นใหญ่ ๆ 2 ชั้น (classes) คือความรู้และความสามารถในทางพุทธิปัญญากับทักษะ ชั้นเหล่านี้ยังแบ่งออกเป็นชั้นย่อย ๆ อีก 6 ชั้น ดังรายละเอียดต่อไปนี้

¹⁷Gronlund, op. cit., p. 59.

Cognitive Domain

- 1.00 Knowledge (Remembering Previously Learned Material)
 - 1.10 Knowledge of specifics.
 - 1.11 Knowledge of terms.
 - 1.12 Knowledge of specific facts.
 - 1.20 Knowledge of ways and means of dealing with specifics.
 - 1.21 Knowledge of conventions.
 - 1.22 Knowledge of trends and sequences.
 - 1.23 Knowledge of classifications and categories.
 - 1.24 Knowledge of criteria.
 - 1.25 Knowledge of methodology.
 - 1.30 Knowledge of the universals and abstractions in a field.
 - 1.31 Knowledge of principles and generalizations.
 - 1.32 Knowledge of theories and structures Intellectual Abilities and Skills.
- 2.00 Comprehension (Grasping the meaning of material)
 - 2.10 Translation (Converting from one form to another)
 - 2.20 Interpretation (Explaining or summarizing material)
 - 2.30 Extrapolation (Extending the meaning beyond the data)
- 3.00 Application (Using information in concrete situations)
- 4.00 Analysis (Breaking down material into its parts)
 - 4.10 Analysis of elements (Identifying the parts)
 - 4.20 Analysis of relationships (Identifying the relationship)

- 4.30 Analysis of organizational principles (Identifying the way the parts are organized)
- 5.00 Synthesis (Putting parts together into a whole)
- 5.10 Production of a unique communication.
- 5.20 Production of a plan or proposed set of operations.
- 5.30 Derivation of a set of abstract relations.
- 6.00 Evaluation (Judging the value of a thing for a given purpose using definite criteria)
- 6.10 Judgements in terms of internal evidence.
- 6.20 Judgements in terms of external criteria.¹⁸

จะเห็นได้ว่า การแบ่งชั้นต่าง ๆ นั้นเรียงลำดับผลการเรียนรู้ของนักเรียนจากง่ายไปหายาก จากสิ่งที่เป็นรูปธรรมไปหาสิ่งที่เป็นนามธรรม ซึ่งเป็นการจำแนกพฤติกรรมของนักเรียนตามธรรมชาติ

วัตถุประสงค์ของการสอนด้านพุทธิปัญญาหรือความรู้ (Cognitive domain) นี้ มีประโยชน์มากในการสร้างแบบทดสอบผลสัมฤทธิ์ (Achievement Test) เพราะเน้นหนักในคำความหมายของเนื้อหาวิชา คือบอกให้ทราบว่า ความรู้ที่เด็กควรรู้ ควรมี นั้นได้แก่อะไรบ้าง และจำแนกผลการเรียนรู้ของนักเรียนในรูปของพฤติกรรมของนักเรียนด้วย แต่ไม่จำเป็นว่าทุกวิชาจะต้องมีจุดประสงค์ของการสอนครบตามรายการต่าง ๆ ที่จำแนกไว้แล้วนั้นเสมอไป การจำแนกนั้นเป็นเพียงการจำแนกที่มีความสัมพันธ์กับผลการเรียนรู้ของนักเรียนอย่างกลาง ๆ เท่านั้น ตัวครูผู้สอนเป็นผู้ซึ่งจะต้องคิดว่า ผลการเรียนรู้อะไรที่จะช่วยปรับปรุงการสอน การทดสอบ และการประเมินผลได้บ้างในเรื่องแล้วนี้จึงควรคิดต่อไป 19

¹⁸ Benjamin S. Bloom (ed.) Taxonomy of Educational Objectives, Handbook I : The Cognitive Domain (New York: David McKay Co., 1956), pp. 201-207.

14 อนันต์ ศรีโรจน์, ภาควิชาคณิตศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ (กรุงเทพฯ: กองส่งเสริมและ
พัฒนาการศึกษา, 2515), น. 11

ปรับปรุงการสอน การทดสอบ และการเน้นหนักในแต่ละเรื่องเหล่านั้น ควรทำอย่างไร¹⁹

6. สำหรับจุดมุ่งหมายในการสอนภาษาที่สองหรือภาษาต่างประเทศ และจุดมุ่งหมายเชิงพฤติกรรมนั้น Rebecca M. Valette เห็นว่าครอบคลุมทั้ง 3 ด้าน ดังที่ไคกล่าวไว้ดังนี้

The behavioral objectives in second-language learning cross all three domains of the taxonomy : cognitive, psychomotor, and affective. The presence of the affective component is typical of all subject-matter areas of the curriculum. Languages differ from other academic subjects, however, because of the heavy psychomotor component presents in the early stages of learning.²⁰



นอกจากนี้ Valette ยังไคกล่าวถึงวัตถุประสงค์ของการสอนภาษาอังกฤษเป็นภาษาที่สองตามแนวความคิดใหม่ ซึ่งแตกต่างไปจากวัตถุประสงค์เดิมเมื่อ 20 ปีก่อน เขากล่าวว่าในปัจจุบันนี้การสอนภาษาอังกฤษเป็นภาษาที่สองนั้นเน้นทางด้านการสอนภาษาพูดและภาษาเขียนเท่า ๆ กัน นอกจากนี้ยังเน้นในค่านิยม อารยธรรม (รวมประวัติศาสตร์และศิลปะ) ขบวนการเรียนรู้ทางพุทธิปัญญา และการสร้างนิสัยต่าง ๆ มิได้เน้นทางด้านการพูดและภาษาพูดแต่อย่างเดียวเหมือนอย่างแต่ก่อน

The objectives of second-language instruction are entering a period of stabilization. Over the past two decades, emphasis in modern

001148

¹⁹อนันต์ ศรีโสภะ, รายงานการพัฒนาการทดสอบมัธยมศึกษา (กรุงเทพฯ: กองส่งเสริมและวัดผลการศึกษา, 2515), หน้า 11.

²⁰Rebecca M. Valette, "Evaluation of Learning in A Second Language," In B.S. Bloom and Others (Ed.), Handbook on Formative and Summative Evaluation of Student Learning (New York: McGraw-Hill Book Company, 1971), p. 823.

languages has been placed on the spoken language and the development of proper speech habits. Now a balance is being sought between the spoken language and the written language, culture and civilization (including history and the arts), cognitive processes, and acquisition of habits.²¹

7. สำหรับวัตถุประสงค์ของการทดสอบทางคำภาษานั้น David P. Harris กล่าวว่า มีดังต่อไปนี้

7.1 เพื่อสำรวจความพร้อมของผู้เรียนเพื่อที่จะจัดโครงการสอนได้ถูกต้อง (To determine readiness for instructional program)

7.2 เพื่อจัดผู้เรียนแต่ละคนเข้าชั้นเรียนที่ถูกต้องตามระดับความสามารถของเขา (To classify or place individuals in appropriate language classes)

7.3 เพื่อวินิจฉัยเรื่องที่ยุ่เรียนแต่ละคนเรียนได้ดี หรือเรื่องที่เป็นจุดอ่อนของยู่เรียน (To diagnose the individual's specific strengths and weaknesses)

7.4 เพื่อวัดความถนัดในการเรียนของยู่เรียน (To measure aptitude for learning)

7.5 เพื่อวัดประสิทธิผลของการสอน (To evaluate the effectiveness of instruction)

7.6 เพื่อวัดสัมฤทธิ์ผลในการเรียนของนักเรียนว่าตรงตามวัตถุประสงค์ของการสอนหรือไม่ (To measure the extent of student achievement of the instructional goals)²²

²¹Ibid., p. 817.

²²David P. Harris, Testing English as a Second Language (New York: McGraw-Hill Book Company, 1969), pp. 2-4.

จากวัตถุประสงค์ของการทดสอบทางศัพทวิทยาตามที่กล่าวมาแล้ว Harris กล่าวว่า อาจสรุปเป็นการทดสอบได้ 3 ประเภท คือ

1. Proficiency test คือ การทดสอบเพื่อวัดความสามารถของผู้เรียน ที่มีอยู่เดิมว่ามีมากน้อยเพียงใด หรือเพื่อสำรวจความพร้อมของผู้เรียน เพื่อจะได้จัดเข้าชั้นเรียนที่ถูกต้อง

2. Aptitude test คือ การทดสอบเพื่อวัดความถนัดทางการเรียนของผู้เรียน

3. Achievement test คือ การทดสอบเพื่อสำรวจผลสัมฤทธิ์การเรียนของผู้เรียน และการสอนของครู²³

8. องค์ประกอบ (Components) ที่สำคัญในการเรียนภาษา

เมื่อได้กล่าวถึงวัตถุประสงค์ทั่วไปของการศึกษา วัตถุประสงค์ของการสอนภาษาต่างประเทศ และวัตถุประสงค์ของการทดสอบทางภาษาแล้ว ต่อไปจะกล่าวถึงองค์ประกอบที่สำคัญในการเรียนภาษา ซึ่งนักภาษาทั่วไปมีความเห็นพ้องกันว่า มีอยู่ 3 อย่างด้วยกัน คือ เสียง (Sound system หรือ Phonology) คำศัพท์ (Vocabulary) และโครงสร้างไวยากรณ์ (Structure)

แต่ David P. Harris ได้มีความเห็นว่า เพื่อที่จะให้มององค์ประกอบในการเรียนภาษาครบตามทักษะทั้ง 4 ของการเรียน คือ ฟัง พูด อ่าน เขียน แล้วอาจเพิ่มองค์ประกอบให้เป็น 4 อย่างได้ โดยเพิ่มด้าน Rate and general fluency เข้าไป

²³Ibid., pp. 4-5.

ตารางที่ 2 แสดงทักษะและองค์ประกอบของการเรียนภาษา (Chart of the Language Skills and Their Components)²⁴

Component	Language Skills			
	Listening	Speaking	Reading	Writing
Phonology/orthography				
Structure				
Vocabulary				
Rate and general fluency				

จากตารางนี้จะเห็นได้ว่า องค์ประกอบทั้ง 4 มีความสำคัญเท่าเทียมกันในการเรียนภาษา นักภาษาหลายท่านได้ให้ความเห็นสนับสนุนไว้ดังต่อไปนี้

A person has "learned" a foreign language when he has thus first, within a limited vocabulary mastered the sound system (that is, when he can understand the stream of speech and achieve an understandable production of it) and has, second, made the structural devices (that is, the basic arrangements of utterances) matters of automatic habits.²⁵

²⁴Harris, op. cit., p. 11.

²⁵Charles C. Fries, Teaching and Learning English as a Foreign Language (Ann Arbor: The University of Michigan Press, 1967), p. 1.

"...knowing a language means, first of all, knowing its phonology and grammatical structure."²⁶

...หากนักเรียนรู้ต้องการจะเข้าใจภาษาอังกฤษ และให้ผู้อื่นเข้าใจตนด้วยแล้ว ก็ไม่มีหนทางหลีกเลี่ยงที่จะไม่เรียนรู้ลักษณะไวยากรณ์เหล่านั้นให้จริง และสามารถเข้าใจได้อย่างไม่ตองคิด ในเมื่อคนอยู่ในสภาพการเผชิญไขภาษาอังกฤษจริง ๆ²⁷

เกรทราวแล้วว่า แนวโน้มของวัตถุประสงค์ของการสอนภาษาอังกฤษเป็นภาษาที่สองในปัจจุบันนี้ได้หันมาเน้นคำภาษาเขียนและการเขียนพอ ๆ กับภาษาพูด ดังนั้นผู้เรียนจึงจำต้องมีความรู้ทางด้านโครงสร้างทางไวยากรณ์ (grammatical structures) ของภาษาอังกฤษให้เพียงพอที่จะใช้พูดหรือเขียนอย่างถูกต้องและอย่างไม่ตองคิด จึงจะเห็นได้ว่าโครงสร้างทางไวยากรณ์มีความสำคัญสำหรับนักเรียนผู้เรียนภาษาอังกฤษเป็นภาษาที่สองมาก และกระทรวงศึกษาธิการก็ได้ระบุเรื่องนี้ไว้อย่างชัดเจนทั้งในหลักสูตรและระเบียบวิธีวัดผลวิชาภาษาอังกฤษของทุกระดับชั้นเรียนที่จัดให้มีการเรียนการสอนวิชาภาษาอังกฤษ

9. ในหลักสูตรมัธยมศึกษาตอนต้น โรงเรียนมัธยมแบบประสม ได้อธิบายรายการสอนวิชาภาษาอังกฤษ 32²⁸ ของชั้นมัธยมศึกษาปีที่ 3 ไว้ว่า

²⁶John B. Carroll, "Research on Teaching Foreign Language," Handbook of Research on Teaching (Chicago: Rand McNally and Company, 1967), p. 1072.

²⁷ไพรัช แวน ไชยกุล, แนววิธีสอนภาษาอังกฤษตามหลักภาษาศาสตร์สำหรับประเทศไทย คมคาย จงเจริญสุข (ผู้แปล) (พระนคร: โรงพิมพ์สมาคมนงเคราะห์ศาสตร์แห่งประเทศไทย, 2505), หน้า 72.

²⁸ภาษาอังกฤษ 32 หมายถึงวิชาภาษาอังกฤษที่จัดสำหรับนักเรียนที่จะศึกษาต่อในชั้นมัธยมศึกษาตอนปลาย โดยให้เรียนสัปดาห์ละ 5 คาบ (periods) และมุ่งให้นักเรียนฝึกทักษะทั้ง 4 ด้าน คือ การฟัง การพูด การอ่าน และการเขียน.

รายวิชานี้จัดสำหรับนักเรียนที่จะศึกษาต่อในชั้นมัธยมศึกษาตอนปลาย โดยมุ่งให้นักเรียนได้รับความรู้ทางภาษาและฝึกฝนทักษะ การฟัง พูด และอ่าน ท่อจากที่โคเรียนมาแล้ว ให้นักเรียนสามารถใช้ภาษาภายในวงศัพท และไวยากรณที่วางไวโคยอย่างคูลองแคลลวและแมนย่า ในการสนธิเรียงความนักเรียนจะตองโครับการฝึกทังการพูดปากเปลวและเขียนธิเรียงธิเรียงในธิระยะแรก ๆ แผลวการฝึกเขียนของนักเรียนจะอยู่ในวงศัพทและแบบสรทงที่วางไว แต่ครุควธิจะชวธิให้นักเรียนคูลอย ๆ ฝึกใช้ความคิดของคองคองเอง และใช้ศัพทและไวยากรณที่กวว่งชวธิขึ้นตามล่ำคัธิ รยธิวิชานทองจู้คัธิให้นักเรียนมีคูลวธิธิเพิ่มขึ้นเกยวคัธิกับวัฒนธรรมของชนชาติที่พูดภาษาอังกฤเป็นภาษาประจำชาติ²⁹

10. ส่วนระเบียบการวัดผลวิชาภาษาอังกฤษในระดับชั้นมัธยมศึกษาตอนต้น ได้กำหนดให้มีการวัดผลการเรียนภาษาอังกฤษ ดังนี้

การวัดผลของแต่ละหมวดหรือรายวิชาประจำปีการศึกษา หรือประจำภาคเรียน

ก. สำหรับวิชาที่นักเรียนต้องเรียนตลอดปี ให้ทำการวัดผลจากงานระหว่างปี และการสอบปลายปี

ข. สำหรับวิชาที่นักเรียนไม่ต้องเรียนตลอดปี ให้ทำการวัดผลจากงานระหว่างภาคเรียน และการสอบปลายภาคเรียน เมื่อจบหลักสูตรที่กำหนดไว้สำหรับวิชานั้น ๆ

ค. คะแนนงานระหว่างปี หรือคะแนนระหว่างภาคเรียนนั้น ให้พิจารณาจากการทดสอบทั้งภาคความรู้ ภาคปฏิบัติ งานที่มอบหมายให้นักเรียนทำ และจากการสังเกตของครูในค่านเจตคติและความสนใจของนักเรียน การแบ่งสัดส่วนของคะแนนระหว่างปีตลอดทั้งปี หรือคะแนนงานระหว่างภาคเรียนตลอดระยะเวลาที่หลักสูตรกำหนดไว้ขึ้นอยู่กับ

²⁹ กรมสามัญศึกษา กระทรวงศึกษาธิการ, คำอธิบายรายวิชาหลักสูตรชั้นมัธยมศึกษาตอนต้นและตอนปลายโรงเรียนมัธยมแบบประสม (กรุงเทพฯ: สำนักงานโครงการพัฒนาการศึกษาโรงเรียนมัธยมแบบประสม, 2514), หน้า 4.

ลักษณะของหมวดหรือรายวิชา และคุณสมบัติของโรงเรียน

การทดสอบ และกิจกรรม อย่างน้อยควรทำ 5 ครั้งต่อเทอม โดยแบ่งเป็นการสอบปากเปล่า 1 ครั้ง การสอบข้อเขียน 2 ครั้ง และกิจกรรมต่าง ๆ 2 ครั้ง³⁰

สำหรับระเบียบการวัดผลของโรงเรียนมัธยมแบบประสมนี้ อาจสรุปได้ว่าแบ่งการวัดผลออกเป็น 2 ประเภทใหญ่ ๆ ด้วยกัน ตามที่ B.S. Bloom ได้กล่าวว่าการวัดผลในปัจจุบันแบ่งออกเป็น 2 ประเภทเรียกว่า Formative และ Summative Evaluation

Formative evaluation (first used by Scriven in 1967) is the use of systematic evaluation in the process of curriculum construction, teaching and learning for the purpose of improving any of these three processes.

Summative evaluation is the type of evaluation used at the end of a term, course or program for purpose of grading, certification, evaluation of progress, or research on the effectiveness of a curriculum, course of study, or educational plan.³¹

จะเห็นว่า Formative Evaluation เป็นการวัดผลในระหว่างการเรียนการสอนเมื่อจบตอนย่อย ๆ แต่ละตอน อาจจบบทเรียนหรือไม่จบบทเรียนก็ได้ แต่เป็นการวัดผลเมื่อสอนจบตอนที่ครูเห็นว่าในช่วงนี้อาจมีอะไรที่นักเรียนสงสัยหรือไม่เข้าใจทำการทดสอบเสียครั้งหนึ่ง

³⁰กรมสามัญศึกษา กระทรวงศึกษาธิการ, คู่มือการวัดผลชั้นมัธยมศึกษาตอนต้น โรงเรียนมัธยมแบบประสม (กรุงเทพฯ: สำนักงาน กมส., 2516), หน้า 2.

³¹Benjamin S. Bloom, J. Thomas Hastings, and George F. Madaus, (eds.), Handbook on Formative and Summative Evaluation of Student Learning (New York: McGraw-Hill Book Company, 1971), pp. 117-118.

ส่วน Summative Evaluation นั้น เป็นการวัดผลรวมหลังจากทดสอบแต่ละตอนย่อย ๆ แล้ว เพื่อประเมินว่านักเรียนมีความรู้ความเข้าใจเรื่องที่เรียนไปทั้งหมดเพียงใด มีทักษะ ทักษะเปลี่ยนแปลงไปในทิศทางที่เราต้องการหรือไม่ ซึ่งจะเทียบได้กับการสอบปลายภาคเรียน หรือปลายปี

ผลจากการวัดผลทั้ง 2 ชั้นนี้จะนำมารวมกันเป็นข้อมูลที่ใช้ในการประเมินผล หรือการให้ระดับคะแนนของการวัดผลแผนใหม่ ซึ่งได้แก่คะแนนจากการทดสอบย่อยเมื่อเรียนจบเรื่องหนึ่ง ๆ ในระหว่างภาคเรียน กับคะแนนสอบปลายภาคหรือสอบปลายปี โรงเรียนมัธยมแบบประสมก็ใช้วิธีการวัดผลแบบนี้ด้วยเช่นเดียวกัน การทดสอบโดยวิธีนี้เป็นการให้ความสำคัญธรรมชาติการเรียนรู้ดีกว่าการสอบเพียงปลายภาคหรือปลายปีแต่เพียงอย่างเดียว ดังนั้นในการทดสอบการเรียนภาษาก็ควรที่จะทดสอบทั้ง Formative และ Summative Evaluation ด้วย

การวัดผลประเภทนี้กำลังเป็นที่นิยมใช้กันทั่วไป แม้แต่ในวงการภาษา ผู้เชี่ยวชาญด้านการทดสอบทางด้านการสอนภาษาอังกฤษเป็นภาษาที่สอง เช่น Rebecca M. Valette ก็ได้ใช้วิธีการทดสอบแบบ Formative Evaluation ในการสรุปการวัดผลย่อยระหว่างการสอน ตามทักษะต่าง ๆ ของการเรียนภาษาอังกฤษเป็นภาษาที่สอง ไว้เป็นตารางดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3 แสดงตารางวิเคราะห์รายการในการวัดผลการสอนภาษาที่ 2 แบบ Formation Evaluation
 (Table of Specification for Formation Evaluation in Second-Language
 Instruction)³²

	A	B	C	D	E	F
	Knowledge of vocabulary and phonology	Knowledge of morphology	Knowledge of syntax and phonological pattern	Skill in manipulate dialogues and drills	Listening and reading comprehen- sion	Speaking and writing abili- ties
Spoken Language						
Written Language						

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

³²Valette, op. cit., p. 841.



11. ปัญหาและจุดอ่อนของการทดสอบของไทย

เมื่อปี พ.ศ. 2505 สำนักงานทดสอบทางการศึกษาและจิตวิทยาได้ทำการสำรวจข้อคำถามและวิธีสอบต่าง ๆ ของนักเรียนหลายระดับ ตั้งแต่ชั้นประถมจนถึงอุดมศึกษา และตั้งแต่การสอบคัดเลือกจนถึงการสอบเลื่อนชั้นปลายปีการศึกษา พบว่า การสอบของเรา ยังมีจุดอ่อนที่สำคัญหลายประการ คือ

1. ยังใช้การสอบไม่คุ้มค่า เพราะการสอบของเรานักถึงแต่ค่าเพียง 2 คำ คือ ได้ กับ ตก เท่านั้น โดยไม่ได้ใช้ผลการสอบไปเป็นข้อมูลเตือนใจครูให้ย้อนระลึกถึงการสอนที่ผ่านมาแล้วว่าดีหรือยัง หรือนำไปเตือนนักเรียนให้รู้ตัวว่าเขาควรจะแก้ไขตนเองอย่างไรจึงจะเรียนได้ดีขึ้นกว่าเดิม สาเหตุในเรื่องนี้เกิดจากผู้สอบยังไม่ทราบถึงเทคนิคและวิธีการใหม่ ๆ ของการวัดผลว่าจะต้องทำอย่างไรจึงจะถูกต้อง

2. ข้อสอบยังวัดไม่ครอบคลุม แต่ละวิชายังคงเน้นถามแต่เพียงบางเนื้อหาหรือบางพฤติกรรมเท่านั้น หรือบางที่วัดไม่ตรงกับความมุ่งหมายของวิชานั้นไปเสียเลย เพราะผู้ออกข้อสอบยังไม่ทราบหลักการวัดผลว่าวิชานั้น ๆ ควรจะสอบวัดอะไรบ้าง และควรวัดด้วยวิธีใดนั่นเอง

3. คำถามยังมีคุณภาพต่ำ คือจะถามแต่ในค่านความจำอย่างเข้มข้น มีผู้น้อยนักที่จะนำเอาเนื้อหาและข้อเท็จจริงเหล่านั้นมาพลิกแพลงถามให้นักเรียนตีความ ขยายความ และวิเคราะห์ความหมาย หรือถามให้นำความรู้นั้นไปใช้ในสถานการณ์อื่นที่ผิดแผกไปจากตำรา

4. การแปลผลสอบยังคลาดเคลื่อน โดยทั่วไปยังคงใช้วิธีนำคะแนนที่เด็กสอบได้มารวมกันโดยตรงว่าใครได้ถึง 50 % หรือไม่ แล้วก็ตัดสินให้สอบได้หรือตกไปตามนั้น ซึ่งเป็นวิธีการที่โบราณ ล้าสมัยและมีจุดอ่อนหลายประการ ถึงแม้จะแปลงคะแนนเหล่านั้นให้เป็นเปอร์เซ็นต์ก็ยังคงไร้ความหมายอยู่อีกมาก เพราะคะแนนเต็มของแต่ละวิชาไม่เท่ากัน และยังไม่มีการรับรองไว้ว่า 1 คะแนนของวิชาเลข กับ 1 คะแนนของวิชาอ่านไทย จะมีคุณภาพและขนาดใดเท่ากันหรือเป็นหน่วยเดียวกันจนสามารถนำคะแนนเหล่านั้นมารวมกันได้โดยตรง ปัญหานี้แก้ไขได้โดยการแปลงตัว เลขคะแนนที่นักเรียนสอบได้จากทุกวิชาให้เป็นคะแนนมาตรฐาน ในรูปของคะแนน x ปกติ ตามหลักการสถิติเสียก่อน จึงนำมาเปรียบเทียบกันและรวมกันได้ โดยตรงอย่างมีความหมาย

5. การสอบคัดเลือกยังไม่สมบูรณ์ เพราะใช้การสอบข้อเขียนเป็นพื้น และมุ่งวัดเนื้อหาวิชาเป็นส่วนใหญ่ การสอบโดยวิธีนี้มิได้รับรองว่าผู้ที่สอบผ่านเข้าไปได้นั้นจะมี รัศมีเขาวนั้ปัญญาหรือสมรรถวิสัยของการเรียนรู้อย่างที่ในปัจจุบันเรียกว่า ความถนัดทาง การเรียน สูงพอที่จะเรียนวิทยาการนั้น ๆ ได้สำเร็จและโดยสะดวกเสมอไป การสอบ คัดเลือกที่คั้งนี้จึงควรมีการทดสอบด้วยแบบทดสอบความถนัดทางการเรียน (Scholastic Aptitude Test) ด้วย³³

นอกจากนี้ ดร. ขวาล ยังได้กล่าวไว้ในหนังสือเทคนิคการวัดผลที่พิมพ์ครั้งแรกปี 2504 จนกระทั่งถึงฉบับปรับปรุง พิมพ์ครั้งที่ 5 ในปี 2516 ว่า วิทยาการด้านการวัดผล ของไทยยังล้าหลังหลักสูตรและวิธีสอนอยู่มาก มากจนไม่ทราบว่าจะเทียบกะอะไรก็ ขนาด คนละมุมโลกก็ยังน้อยไป หรือสรุปว่า การสอนเก่งมาก แต่สอนไปแล้วได้ผลอย่างไรไม่ทราบ เพราะวัดไม่เป็น³⁴

จากปัญหาและจุดอ่อนของการทดสอบของเราดังกล่าวแล้ว รวมทั้งความสำคัญของ โครงสร้างทางไวยากรณ์ในการเรียนภาษาอังกฤษ เป็นภาษาที่สองหรือเป็นภาษาต่างประเทศ ทำให้ผู้เขียนมีความสนใจที่จะสร้างแบบทดสอบเพื่อวัดผลสัมฤทธิ์ในการใช้โครงสร้างไวยากรณ์ ภาษาอังกฤษของนักเรียนในระดับมัธยมศึกษาปีที่ 3 ซึ่งผู้เขียนกำลังสอนอยู่ในโรงเรียนที่ใช้ หลักสูตรโรงเรียนมัธยมแบบประสมนี้ เพื่อจะได้แบบทดสอบที่มีคุณภาพดี เป็นที่เชื่อถือได้ตาม หลักวิชาทางสถิติและการวัดผล อันจะเป็นแนวทางในการปรับปรุงแบบทดสอบที่ครูสร้างขึ้น ให้มีมาตรฐานที่ขึ้นต่อไป

³³สำนักงานทดสอบทางการศึกษาและจิตวิทยา วิทยาลัยวิชาการศึกษา ประสานมิตร, รายงานความก้าวหน้าของโครงการสร้างแบบทดสอบมาตรฐานชั้นประถมศึกษาปีที่ 7 ของกระทรวง ศึกษาธิการ ฉบับที่ 1 (กรุงเทพฯ โรงพิมพ์สำนักงานทดสอบฯ, 2513), หน้า 9-16.

³⁴ขวาล แพทย์กุล, เรื่องเดิม, หน้า 9.

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาระเบียบวิธีในการสร้างแบบทดสอบผลสัมฤทธิ์ทางการใช้โครงสร้างไวยากรณ์อังกฤษ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในโรงเรียนที่ใช้หลักสูตรโรงเรียนมัธยมแบบประสมในกรุงเทพมหานคร
2. เพื่อสร้างและปรับปรุงแบบทดสอบนี้ให้เป็นแบบทดสอบที่ดี มีค่าความยากง่ายพอเหมาะ มีค่าอำนาจจำแนกดี มีค่าความเชื่อถือได้ และค่าความเที่ยงตรงสูง

ขอบเขตของการวิจัย

1. การสร้างแบบทดสอบผลสัมฤทธิ์ไวยากรณ์อังกฤษครั้งนี้ ยึดถือเนื้อหาโครงสร้างไวยากรณ์ในหนังสือ Intermediate English Course for Thailand, Book III³⁵ เป็นหลัก
2. วิเคราะห์ข้อสอบ (Item Analysis) เพื่อหาอำนาจจำแนก และระดับความยาก (Discriminating Power and Difficulty Index)
3. หาความเชื่อถือได้ของแบบทดสอบ (Reliability of the Test)
4. ตัวอย่างประชากรที่ใช้ในการวิจัย เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนที่ใช้หลักสูตรโรงเรียนมัธยมแบบประสมในกรุงเทพมหานคร 4 โรงเรียน
5. โรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง คือ โรงเรียนสตรีศรีสุริโยทัย โรงเรียนบางกะปิ โรงเรียนสุวรรณภคารมวิทยาคม และโรงเรียนวัดรางบัว
6. การวิจัยนี้มีได้พิจารณาในเรื่องอายุ เพศ ความสามารถในการเรียนของนักเรียนทุกวิชา ตลอดจนอาชีพของบิดามารดา และสิ่งแวดล้อมทางบ้าน

³⁵กรมวิชาการ, Intermediate English Course for Thailand, Book III (พิมพ์ครั้งที่ 3; กรุงเทพฯ: โรงพิมพ์คุรุสภาลาดพร้าว, 2515.)

ความจำกัดของการวิจัย

คะแนนที่นักเรียนได้รับ อาจจะไม่มาจากความสามารถในการใช้โครงสร้างไวยากรณ์อังกฤษโดยแท้จริงของนักเรียน เนื่องจากในขณะที่ทำการสอบอาจมีสภาพการบางอย่างเข้ามามีอิทธิพลทำให้ผลการสอบคลาดเคลื่อนจากความเป็นจริง เช่น ความคุ้นเคย ความกังวลใจ ความไม่สบาย ความเกียจคร้าน อันอาจทำให้เกิดความคลาดเคลื่อนเมื่อนำผลการทดสอบมาวิเคราะห์

ประโยชน์ที่จะได้รับจากการวิจัย

1. ครูผู้สอนภาษาอังกฤษหรือวิชาอื่น ๆ สามารถใช้เป็นแนวทางในการสร้างแบบทดสอบตามหลักวิธีการวัดผลอย่างถูกต้อง เพื่อที่จะสร้างแบบทดสอบให้ใกล้เคียงกับมาตรฐานมากที่สุดเท่าที่จะทำได้
2. เพื่อเสนอแนะวิธีการสร้างแบบทดสอบของครูให้ได้มาตรฐานดียิ่งขึ้น

นิยามของคำต่าง ๆ ที่ใช้ในการวิจัย

1. โครงสร้างไวยากรณ์อังกฤษ หมายถึง ระเบียบวิธีการเรียบเรียงคำในประโยค ตลอดจนการเรียบเรียงส่วนประกอบของคำในภาษาอังกฤษ "... The patterns of arrangement of words in sentences, and the patterns of arrangement of parts of words into words are its grammatical structures"³⁶
2. แบบทดสอบ หมายถึงแบบทดสอบผลสัมฤทธิ์ไวยากรณ์อังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ซึ่งผู้วิจัยได้สร้างขึ้นเพื่อใช้เป็นเครื่องมือในการวิจัย

³⁶ Robert Lado, Language Testing (Briston: Western Printing Services, Ltd., 1961), p. 142.

3. ความเที่ยงตรงของแบบทดสอบ (Validity) หมายความว่า แบบทดสอบนั้นสามารถวัดได้ตรงตามที่ตั้งความมุ่งหมายไว้

4. ความเชื่อถือได้ของแบบทดสอบ (Reliability) หมายความว่า แบบทดสอบนี้มีความแน่นอน ความคงเส้นคงวาในการวัด เมื่อนำไปทดสอบอีกก็จะได้ยังคงรักษาค่าแห่งของผู้เข้าทดสอบไว้ได้เหมือนเดิม (คือผู้ทดสอบที่มีความสามารถในระดัใดก็ยังคงได้คะแนนในระดัเดียวกันกับการทดสอบครั้งก่อน) แบบทดสอบที่ดีควรมีค่าความเชื่อถือได้สูง



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย