



เอกสารและการวิจัยที่เกี่ยวข้อง

ถ้าจะให้การศึกษาเป็นไปอย่างมีประสิทธิภาพแล้ว บุคคลที่ทำหน้าที่สอนควรจะ สามารถที่จะแสวงหารายละเอียดเกี่ยวกับนักเรียนให้มากที่สุด¹ การวัดผลจะเป็นวิธีการ ที่ทำให้ครูและนักเรียนทราบว่า หลักสูตรที่วางไว้ดำเนินไปตามจุดมุ่งหมายหรือไม่ กิจกรรม และประสบการณ์ต่าง ๆ ที่นักเรียนได้รับ มีคุณค่ากับการเรียนการสอนหรือไม่ และนักเรียน มีการเปลี่ยนแปลงพฤติกรรมอย่างไร²

การวัดผลที่ใช้กันมากที่สุดในโรงเรียนได้แก่ การวัดผลในด้านการสอบข้อเขียน ซึ่งใช้แบบสอบเป็นเครื่องมือในการทดสอบ ทั้งนี้ คุณค่าของการสอบจึงขึ้นอยู่กับคุณภาพของ แบบสอบโดยตรงที่จะช่วยให้การสอบที่มีคุณค่าหรือไร้คุณค่า แบบสอบจึงเป็นตัวจักรกลอันสำคัญ ของขบวนการประเมินผล³

การสร้างแบบสอบจำเป็นต้องมีหลักการ (Principles) ที่จะช่วยให้ผู้ออกแบบสอบ สามารถสร้างแบบสอบที่มีคุณภาพดี เพราะแบบสอบจะต้องมีเนื้อหาครอบคลุมความสามารถของ นักเรียนที่ได้เรียนมาเป็นเวลาหลายเดือนหรือหลายปี แต่ใช้เวลาทำแบบสอบเพียงสองหรือ สามชั่วโมง

การสร้างแบบสอบในด้านภาษาก็เช่น เกี่ยวกับการสร้างแบบสอบวิชาอื่น ๆ ที่จะต้อง คำนึงถึงองค์ประกอบสำคัญ ๆ องค์ประกอบเหล่านี้ได้แก่จุดมุ่งหมายของการทดสอบ หลักการ

¹C.M. Lindvull, Testing and Evaluation: An Introduction (New York: Harcourt, Brace & World Inc., 1961), p. 1.

²รองรัตน์ อิศรภักดี, "ข้อคิดเห็นเกี่ยวกับการวัดผลการเรียน," วิทยาคารย ฉบับที่ 3 (พฤษภาคม, 2509), 256 - 67.

³เควิก เค อีแนน, การประเมินผลการสอนในประเทศไทย ดร. ชาวล แพร์ตกุล แปล. (พระนคร : ไทยแบบเรียน, 25097, หน้า 16.

สร้างแบบสอบ การดำเนินการสอบ ตลอดจนการตีความของผลการสอบ

การกำหนดจุดมุ่งหมายของการทดสอบเป็นองค์ประกอบสำคัญในการสร้างแบบสอบ ผู้สร้างแบบสอบจะต้องวางวัตถุประสงค์ของการทดสอบที่เฉพาะแน่นอนลงไป การมีวัตถุประสงค์เฉพาะและแน่นอนดังกล่าวนี้ Robert Cooper กล่าวว่า จะช่วยให้ผู้ออกแบบสอบสามารถกำหนดเนื้อหาของแบบทดสอบ และทราบแน่นอนว่าจะนำผลการทดสอบมาใช้ทำอะไร⁴

เดวิด แฮร์ริส⁵ (David Harris) ได้ให้แนวทางเกี่ยวกับการตั้งวัตถุประสงค์ในการทดสอบภาษาออกเป็น 6 ประการ ที่สำคัญ ๆ คือ เพื่อวัดความพร้อมของนักเรียนในการเรียนตามโปรแกรมการเรียนการสอนที่โรงเรียนจัดขึ้น เพื่อจัดแบ่ง (Classify) ให้นักเรียนเข้าชั้นเรียนตามความสามารถในปัจจุบันที่นักเรียนมีอยู่ (Proficiency) เพื่อสำรวจข้อบกพร่องและจุดเด่น (Weaknesses and Strengths) ทางด้านภาษาของนักเรียน เพื่อวัดความสามารถในการเรียนภาษาของนักเรียน (Aptitude for Learning) เพื่อวัดผลสัมฤทธิ์ (Achievement) ของนักเรียนในด้านความก้าวหน้าในการเรียนรู้และเพื่อวัดประสิทธิภาพของการสอน

วิลเลียม แมคคีย์⁶ (William Mackey) ได้สรุปความสำคัญของการตั้งวัตถุประสงค์ในการทดสอบว่าจะช่วยให้ผู้สร้างแบบสอบสามารถสร้างแบบสอบทางภาษาออกเป็น 4 ประเภทใหญ่ ได้แก่ แบบทดสอบ "Aptitude Tests" ซึ่งใช้ในการทดสอบความสามารถในด้าน

⁴Harold B. Allen (ed.), Teaching English as a Second Language (New York: McGraw - Hill International Book Company, 1965), pp. 330 - 32.

⁵David Harris, Testing English as a Foreign Language (New York: McGraw - Hill Book Company, 1969), p. 2 - 3.

⁶William Mackey, Language Teaching Analysis (London: Longman, Green & Co., Ltd., 1966), pp. 404 - 05.

การเรียนภาษา แบบทดสอบ "Proficiency Tests" ซึ่งใช้ในการทดสอบความสามารถทางภาษาในปัจจุบันของนักเรียน แบบทดสอบ "Achievement Tests" เพื่อทดสอบความก้าวหน้าของนักเรียนในการเรียนรู้ และแบบทดสอบ "Disgnostic Tests" เพื่อทดสอบว่านักเรียนมีจุดเด่นและจุดบกพร่องที่จะต้องแก้ไขอะไรบ้าง

องค์ประกอบที่สำคัญประการที่สองในการสร้างแบบสอบทางภาษาได้แก่ หลักการสร้างข้อสอบ (Constructing the Tests) ซึ่งประกอบด้วยขบวนการสำคัญเป็นขั้น ๆ คือ การวางแผนการสอบ การเตรียมข้อสอบแต่ละข้อ ตลอดจนคำสั่ง การตรวจทานข้อสอบที่ออกแล้ว การนำข้อสอบไปทดลองสอบและวิเคราะห์ผลการสอบ การเลือกข้อสอบที่ดีเพื่อนำไปทดสอบจริง การจัดพิมพ์ข้อสอบสำหรับการสอบ และการดำเนินการสอบ

ในด้านการวางแผนการสอบในขั้นแรกได้แก่การพิจารณาอย่างละเอียดว่าการสอนภาษา ในการเรียนภาคหนึ่งหรือในปีหนึ่ง มีวัตถุประสงค์โดยทั่วไปอย่างไร มีอะไรที่ต้องการให้นักเรียนเรียนรู้ เพื่อจะได้แจ่มแจ้งจุดประสงค์เฉพาะในการทดสอบได้ถูกต้อง⁷ เช่น จะทดสอบในคำนำใจ คำเสียงและการเขียน (Phonology and Orthology) คำโครงสร้าง (Structure) คำศัพท์ (Vocabulary) หรือในด้านความชำนาญและความคล่องแคล่ว กล่าวคือ มีวัตถุประสงค์ในการทดสอบทักษะทางภาษาอะไรบ้าง และทดสอบส่วนประกอบของภาษา (Elements) อะไรบ้าง นอกจากนั้นยังจะต้องพิจารณาจุดมุ่งหมายเฉพาะของการทดสอบเฉพาะด้าน และส่วนประกอบของภาษา เมื่อได้กำหนดแบบ (Design) ของการออกแบบสอบเป็นที่แน่นอนแล้ว ก็จะต้องคำนึงถึงเวลาและอัตราความเร็ว ซึ่งทั้งเวลาและอัตราความเร็วนี้ขึ้นอยู่กับชนิดของแบบสอบหรือวัตถุประสงค์ของแบบสอบ เช่นแบบสอบที่ต้องการจะวัดความสามารถทางภาษาในปัจจุบัน (Proficiency Tests) ก็อาจจะกำหนดเวลาให้นักเรียนส่วนใหญ่สามารถทำได้ทันเวลา⁸

⁷C.M. Linavall, op. cit., p. 11.

⁸Harris, op. cit., p. 110.

การวางแผนการสอบนอกจากจะต้องกำหนดวัตถุประสงค์และกำหนดแบบ ดังได้กล่าวมาแล้ว สิ่งที่สำคัญต่อมาที่ผู้สร้างแบบสอบจะต้องคำนึงถึงคือ การเลือกเนื้อหาออกข้อสอบ เนื้อหาจะต้องขึ้นอยู่กับแบบของการทดสอบ (Design) เช่น ถ้าแบบสอบเป็นแบบสอบ "Proficiency Tests" เนื้อหาที่จะใช้เป็นหลักในการออกข้อสอบก็ไม่จำเป็นจะต้องเป็นเนื้อหาที่เกี่ยวข้องกับหลักสูตรใดหลักสูตรหนึ่ง แต่ถ้าแบบสอบเป็นแบบสอบ "Achievement Tests" การออกข้อสอบก็จะออกข้อสอบเฉพาะเนื้อหาที่นักเรียนได้เรียนไปแล้วเท่านั้น⁹

ในการเลือกเอาเนื้อหาทดสอบนี้ (Sample of the English Language) เนื่องจากเนื้อหาที่จะนำมาทดสอบนั้น ไม่สามารถจะนำมาออกข้อสอบได้อย่างละเอียดทุกเรื่อง ทุกตอน ดังนั้นผู้สร้างแบบสอบจะต้องคำนึงถึงค่านี้นี้ให้มากและ จะต้องจำกัดขอบข่ายของเนื้อเรื่อง ออกมาอย่างรอบคอบ¹⁰ ในการเลือกเนื้อหาทดสอบนี้ เดวิด แฮร์ริส¹¹ (David Harris) กล่าวว่า ควรพิจารณาในแง่ที่เนื้อหาอะไรที่ยากในการเรียนรู้ ของนักเรียนส่วนใหญ่ โครงสร้างอะไรที่หลักสูตร เน้นหนักและ จะต้องใช้ติดต่อกันไป คำศัพท์ใดที่จะเป็นประโยชน์แก่นักเรียนในการนำไปใช้ต่อไป ในด้านเนื้อหาของภาษานี้ เอจิ้น ไบรร์¹² (Eugene Briere) กล่าวว่า ควรจะเลือกเนื้อหาที่เราคาดหมายหรือยอมรับว่า (Assumption) มีสิ่งต่าง ๆ เฉพาะหรือมีเนื้อหาเฉพาะอย่าง (Specific Things) ในภาษา ซึ่งเมื่อรู้สิ่งเฉพาะดังกล่าวแล้ว ก็จะทำให้สามารถรู้ภาษาใดภาษาหนึ่งได้มาเป็นส่วนสำคัญในการออกสอบ

⁹Rebecca Valette, Modern Language Testing (New York: Harcourt, Brace & World Inc., 1967), pp. 111 - 115.

¹⁰Allen (ed.), op.cit., pp. 313 - 21.

¹¹Harris, op. cit., p. 112.

¹²Eugene Briere, "Are We Really Measuring Proficiency With Our Foreign Language Tests," Testing in Second Language Teaching: New Dimensions ATESOL - Ireland. Edited by Maureen cancano O'Brien. (99 St. Stephen's Green Dublin 2), p. 68.

ซึ่งจำนวนศัพท์และโครงสร้างที่เราคาดหมายทั้งกล่าวนี้ออจีน ไบรร์ (Eugene Briere) เรียกว่า "Discrete Point" สามารถจะเลือกออกมาโดยการวิเคราะห์ภาษาที่นักเรียนเรียนอยู่นั้น หรือกับการ เปรียบเทียบภาษาแม่กับภาษาอังกฤษ (Contrative Analysis) ก็ได้

องค์ประกอบสำคัญอีกประการหนึ่งในการสร้างแบบสอบได้แก่ การเตรียมข้อสอบ แต่ละข้อตลอดจนเตรียมคำสั่ง (Preparing the Test Items and Directions) ข้อสอบควร จะมีจำนวนมากกว่าจำนวนข้อที่จะนำไปทดสอบจริง เพื่อจะได้คัดข้อสอบข้อที่ไม่เหมาะสม และเพื่อไว้สำหรับการนำข้อสอบไปทดสอบก่อน (Pre-Test) ซึ่งจะต้องคัดทิ้งเหลือแต่ข้อสอบที่ดีเท่านั้น การกะเนื้อควรถูกกะเนื้อไว้ประมาณ 30 - 40 เปอร์เซ็นต์ ส่วนการเขียนข้อสอบแต่ละข้อควร จะแยกเขียนเป็นข้อๆ เพื่อสะดวกในการ เพิ่มหรือคัดทิ้งและคำคอมที่ผู้ควร เขียนไว้ด้านหลังของแต่ละข้อ เพื่อลดอิทธิพลและความลำเอียงเมื่อกลับมาอ่านพบทวนข้อสอบอีกครั้งหนึ่ง

ส่วนการเขียนคำสั่ง คำสั่งควร สั้น เขียนและใช้ภาษาที่ง่ายแก่การเข้าใจ และควรมีตัวอย่างให้นักเรียนหัดทำ ซึ่งช่วยให้นักเรียนที่เรียนช้าที่สุดสามารถจะเข้าใจวิธีการทำข้อสอบได้อย่างถูกต้อง นอกจากนั้นควร จะบอกเวลาที่กำหนดให้ทำข้อสอบ และการให้คะแนนว่าอนุมัติให้นักเรียนเขาได้หรือไม่¹³

การตรวจทานข้อสอบ (Reviewing the Items) อาจจะทำโดยการ เก็บข้อสอบที่ออกแล้วไว้สองสามวัน หลังจากนั้นก็ได้ครูสอนภาษาอังกฤษด้วยกัน ช่วยกันพิจารณาข้อสอบแต่ละข้ออีกครั้งหนึ่ง ถ้าเป็นที่น่าพอใจก็อาจจะเขียนคำว่า "O.K." ไว้ด้านหลังของแบบสอบก็ได้¹⁴

¹³Harris, op. cit., p. 110.

¹⁴Ibid., p. 111.



การนำข้อสอบไปทดลองสอบ (Pretesting the Material) กับกลุ่มตัวอย่าง ที่มีคุณสมบัติเหมือนหรือใกล้เคียงกับประชากรจริงที่จะทำการทดสอบ จะช่วยให้ผู้สร้างแบบสอบ ทราบว่าค่าสิ่งของข้อทดสอบ เวลาที่ใช้ในการสอบเหมาะสมกับนักเรียนเพียงใด และควรนำ ผลของการทดลองสอบมาวิเคราะห์¹⁵ เพราะการวิเคราะห์ข้อสอบแต่ละข้อจะทำให้ผู้สร้าง แบบสอบเห็นสภาพความจริงของแบบสอบ ช่วยผู้สร้างแบบสอบในการปรับปรุงข้อสอบตลอดจน ตัวเลือกให้มีประสิทธิภาพ¹⁶ อีกทั้งช่วยให้ทราบว่าข้อสอบแต่ละข้อมีความยากเพียงใด มี อำนาจจำแนกของข้อสอบดีหรือไม่ ในการวิเคราะห์ความยากและอำนาจจำแนกแต่ละข้อ กระทำเป็นขั้น ๆ คือ ขั้นแรกแยกกระดาษคำตอบของนักเรียนที่ใดคะแนนรวมสูงสุดลงมา จำนวน 27 เปอร์เซ็นต์ ของผู้เข้าสอบทั้งหมด และอีก 27 เปอร์เซ็นต์ของผู้ที่ใดคะแนน ต่ำสุด ขั้นที่สองได้แก่การชี้ครอคะแนนของตัวเลือกทุกตัว เพื่อหาค่าความยากง่าย (Item Difficulty) และค่าอำนาจจำแนก สำหรับการเก็บข้อมูลในการวิเคราะห์ (Recording Analysis Data) ควรจะใช้กระดาษเก็บข้อมูล ค่าความยากและค่าอำนาจจำแนกของแต่ละข้อ¹⁷

องค์ประกอบในการสร้างแบบสอบอีกประการหนึ่งได้แก่การรวบรวมข้อสอบในการ ทดสอบจริง (Assembling the Final Form) การเลือกข้อสอบขึ้นอยู่กับความต้องการ ของผู้ออกแบบสอบว่าจะให้ข้อสอบมีความยากง่ายอยู่ในระดับใด โดยทั่วไปข้อสอบทางภาษา ควรจะมีความยากในระดับที่มีค่าเฉลี่ยประมาณ 60 ถึง 70 เปอร์เซ็นต์ เพื่อให้เป็นการท้าทาย ความสามารถของนักเรียนในการทำแบบสอบ ส่วนการจัดลำดับตัวเลือกควรจะให้ตัวเลือกทุกตัว มีโอกาสปรากฏเท่า ๆ กัน เช่น ถ้าข้อสอบจำนวน 100 ข้อ ตัวถูก ก ข ค และ ง ควรจะ

¹⁵ Ibid., p. 112.

¹⁶ Anne Anatasi, Psychological Testing (New York: The McMillan Company, 1968), p. 158.

¹⁷ Harris, op. cit., p. 115.

ปรากฏเป็นจำนวน 25 ครั้ง เท่า ๆ กัน ข้อสำคัญอีกประการหนึ่ง ตัวเลือกที่ให้ปรากฏนี้ ควรระวังในการจัดลำดับไม่ให้เป็นกระสวน (Pattern) ที่นักเรียนสังเกตหรือคาดได้¹⁸ นอกจากนี้ข้อสอบแต่ละข้อที่เลือกแล้วควรเรียงลำดับจากข้อที่ง่ายไปหาข้อที่ยากตามลำดับ¹⁹

องค์ประกอบอีกประการหนึ่ง ได้แก่การจัดพิมพ์แบบสอบสำหรับสอบจริง (Reproducing the Test) ซึ่ง ดร. ชวาล แพทย์กุล กล่าวว่า มีความสำคัญมากไม่ยิ่งหย่อนไปกว่าตัวข้อสอบ อาจมีผลต่อการสอบหรือคะแนนของนักเรียน²⁰ การจัดพิมพ์ข้อสอบดังกล่าว ควรชัดเจนและไม่มีข้อผิดพลาด ควรจะเว้นช่องว่างแต่ละบรรทัดให้สามารถอ่านได้ง่ายและข้อกระทงคำถามควรอยู่ในหน้าเดียวกับตัวเลือก ถ้าข้อสอบยังไม่จบ ควรมีคำว่า "พลิก" และถ้าเป็นข้อสอบที่จับเวลาให้ทำเป็นตอน ๆ ไป ควรมีคำว่า "หยุด" นอกจากนี้ ควรมีใบปะหน้าบอกรายละเอียดเกี่ยวกับจุดมุ่งหมายโดยทั่วไปของการทดสอบ วิธีทำแบบสอบ วิธีการแก้ไขคำตอบ รายละเอียดเกี่ยวกับการเก็บ²¹

องค์ประกอบที่สำคัญอีกประการหนึ่งได้แก่การดำเนินการสอบ การดำเนินการสอบ มีความสำคัญมาก เพราะสามารถมีอิทธิพลต่อค่าความเชื่อถือ (Reliability) และค่าความเที่ยงตรง (Validity) ของแบบสอบ การดำเนินการสอบจะต้องกระทำอย่างระมัดระวัง และทำอย่างดีที่สุดในการจัดสภาพการสอบให้เหมือนกันหรืออย่างน้อยก็ใกล้เคียงกัน การที่จะช่วยให้การดำเนินการสอบดำเนินไปด้วยดี การเตรียมห้องสอบที่เหมาะสม การตรวจนับข้อสอบ

¹⁸ Ibid., p. 116.

¹⁹ H.H. Remmers, A Practical Introduction to Measurement and Evaluation (New York: Harper & Brothers Publishers, 1965), p. 194.

²⁰ ชวาล แพทย์กุล, เทคนิคการวัดผล (พิมพ์ครั้งที่ 3, พระนคร : ไทยวัฒนาพานิช, 2508), หน้า 76.

²¹ Harris, op. cit., p. 118.

การจัดที่นั่งสอบ การอ่านข้อสอบทุกฉบับก่อนการคุมสอบ การปฏิบัติระหว่างคุมสอบ การจับเวลา การตรวจตรา ตลอดจนการตรวจนับจำนวนผู้เข้าสอบ ก็มีความสำคัญที่จะช่วยให้ได้แบบสอบที่ดี²²

องค์ประกอบประการสุดท้าย ในการสร้างแบบสอบก็คือประโยชน์ที่จะได้รับจากแบบสอบ กล่าวคือ จะทองสามารถนำผลของการทดสอบมาตีความเพื่อประเมินค่าหรือผลของพฤติกรรมของนักเรียน การตีความจากผลของการทดสอบอาจจะตีความ เป็นประเภทใหญ่ได้สองประเภท คือ Norm - Referenced Interpretation และ Criterion - Referenced Interpretation²³

การตีความประเภท Norm - Reference Interpretation มักจะเป็นการตีความจากแบบสอบมาตรฐานเพื่อวัดความก้าวหน้าหรือผลสัมฤทธิ์ของนักเรียน การตีความประเภทนี้ มีประโยชน์ในการคาดคะเนความสำเร็จของนักเรียนแต่ละบุคคลและของกลุ่ม ตลอดจนของนักเรียนกลุ่มใหญ่ ๆ เพื่อชี้และนำเอาประเด็นปัญหาการศึกษา การตีความประเภทนี้ คะแนนของแต่ละบุคคล และคะแนนของกลุ่มควรจะเป็นมาตรฐาน (Norm) ในรูปของอายุ ระดับชั้น หญิง ชาย วรรณะ หรือแม้กระทั่งประเทศ ซึ่งขึ้นอยู่กับผู้ทดสอบ ท้องการจะนำคะแนนไปเปรียบเทียบกับมาตรฐานชนิดใด โดยทั่วไปจะเปรียบเทียบออกมาเป็นเกรทหรือเรียงลำดับเปอร์เซ็นต์ (Percentile Ranks) สำหรับแบบทดสอบบางอัน เช่น แบบสอบ Metropolitan Achievement Test, 1970 Form F, G และ H ก็จะมีมาตรฐาน (Norm) ของข้อสอบแต่ละข้อในคานความยากง่ายด้วย

²² Robert L. Thorndike (ed.), Educational Measurement

(Washington: American Council on Education, 1971), pp. 188 - 92.

²³ Norman Grolund, Measurement and Evaluation in Teaching

(New York: The Macmillan Company, 1965), p. 109.

ส่วนการตีความประเภท Criterion - Referenced Interpretation หรือที่เรียกว่า Content - Referenced Interpretation ไม่ต้องการเปรียบเทียบแต่ละบุคคลกับกลุ่ม แต่ต้องการจะวัดพฤติกรรม (Performance) ของแต่ละบุคคลว่าจะประสบความสำเร็จ หรือล้มเหลวในการเรียนรู้ตามที่ตั้งจุดมุ่งหมายไว้หรือไม่ และจะช่วยครูในการวัดผลการเรียนการสอน โดยเฉพาะการเรียนการสอนที่สอนตามความแตกต่างระหว่างบุคคล การตีความประเภทนี้จะช่วยให้ครูทราบแน่ชัดว่า ทักษะใดนักเรียนควรจะได้รับการฝึกฝนหรือเรียนเพิ่มเติมจนเกิดความชำนาญ (Mastery) การตีความการสอนประเภทนี้ผลของพฤติกรรมของนักเรียนแต่ละคนจะไม่ถูกนำไปเปรียบเทียบกับมาตรฐานหรือนักเรียนคนอื่น ๆ แต่จะนำผลการสอบของนักเรียนมาเน้นในแง่ที่ว่าเนื้อหาหรือทักษะใดที่นักเรียนยังทำไม่ได้ตามที่วิชาหรือหลักสูตร ได้วางไว้ ทักษะใดที่นักเรียนยังทำไม่ได้ครูก็จะสอนซ้ำหรือปรับปรุงจนกว่านักเรียนจะทำได้ การตีความประเภทนี้ติดกับประเภทแรกที่ทักษะบางอย่างของนักเรียนอาจจะทำไม่ได้ แต่ไม่ต้องแก้ไขอีก ส่วนประเภทหลังนี้จะชี้ออกมาให้เห็นว่า ข้อสอบข้อใดนักเรียนทำผิด เพื่อนำไปสอนซ้ำให้นักเรียนทำได้ หรือถ้านักเรียนทำไม่ได้เลยก็อาจจะต้องปรับปรุงแก้ไขวิธีการสอนหรือความมุ่งหมายให้เหมาะสมต่อไป²⁴

การสร้างแบบทดสอบการอ่าน

ในการเรียนภาษาสิ่งทีนักเรียนเรียนก็คือภาษา ดังนั้นการทดสอบก็ต้องเน้นการทดสอบ ความสามารถทางภาษาของนักเรียน²⁵ ในด้านทักษะการอ่านก็เช่นกัน การทดสอบความสามารถในการอ่านของนักเรียน ก็ต้องทดสอบด้วยการให้นักเรียนอ่าน แล้วนำผลการทดสอบมาประเมินผล

²⁴ National Council of Teachers of English, Common Sense and Testing in English (Illinois: National Council of Teachers of English, 1975), p. 12.

²⁵ Robert L. Lado, Language Testing (London: Longmans Green and Co., Ltd., 1962), p. 20.



การสร้างแบบสอบการอ่านก็เช่นเดียวกับการสร้างแบบสอบโดยทั่วไป ในทาง การศึกษา ที่จะต้องกำหนดแน่ชัดลงไปว่า การสอนอ่านในระดับนั้น ๆ หรือในหลักสูตรที่จะ ทำการทดสอบนั้น เน้น และมีจุดมุ่งหมายให้นักเรียนเกิดพฤติกรรมในการอ่านอย่างไร ลึกซึ้ง แค่ไหน²⁶ โรเบิร์ต แอล ลาดอ²⁷ (Robert L. Lado) ได้แบ่งการทดสอบการอ่านออก เป็น 3 ระดับ ได้แก่ ระดับก่อนการอ่าน (Pre - reading) ได้แก่ความสามารถในการ จำแนกและรู้จักสัญลักษณ์ของภาษา ซึ่งได้แก่ ตัวอักษร ระดับที่สองได้แก่ ความสามารถในการ อ่านเป็นประโยค ๆ ได้อย่างเข้าใจ และระดับที่สามได้แก่ ความสามารถในการอ่านเอาความ จากข้อความที่ยกมาเป็นตอน ๆ (Passages)

สำหรับการอ่านข้อความที่เป็นตอน ๆ (Passages) นี้ รีเบคคา เอ็ม แวลีท²⁸ (Rebecca M. Valette) กล่าวว่า เป็นแบบของการทดสอบการอ่านที่ใช้ได้ผลดี และแพร่หลาย ทั้งกับเจ้าของภาษาและกับผู้ที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ เดวิด แฮร์ริส²⁹ (David Harris) กล่าวว่า การทดสอบการอ่านโดยการให้อ่านข้อความที่ยกมาเป็นตอน ๆ นี้ แบบสอบจะประกอบด้วยข้อความหลายข้อความ และตัวเลือก (Multiple Choice Comprehension Items) หลาย ๆ ข้อ เพื่อทดสอบว่า นักเรียนมีความเข้าใจในการอ่านข้อความ ที่นำมาทดสอบเพียงใด ซึ่งมีหลักการต่าง ๆ ที่จะต้องคำนึงถึงในเรื่องแนวการ เขียนความยาว เนื้อหาวิชา และภาษาที่ใช้ ของข้อความที่นำมาทดสอบ เช่นเกี่ยวกับการเขียนตัวเลือกและ คำคำถาม

²⁶Rebecca M. Valette, Directions in Foreign Language Testing (New York: ERIC Clearinghouse on the Teaching of Foreign Languages and of English in Higher Education), p. 12.

²⁷Lado, op. cit., pp. 232 - 38.

²⁸Valette, op. cit., p. 109.

²⁹Harris, op. cit., pp. 60 - 1.

ในการนำข้อความมาทดสอบนั้น ควรจะเลือกข้อความที่คาดคะเนว่าเป็นจุดมุ่งหมายของขบวนการเรียนการสอน เช่น การเรียนการสอนมีจุดมุ่งหมายในเรื่องให้นักเรียนรู้จักแนวการเขียน (Style) ในแบบต่าง ๆ ข้อความที่เลือกมาก็คงจะเป็นข้อความที่สนองจุดมุ่งหมายในเรื่องนี้ ดังนั้นการเลือกข้อความก็จะคงคำนึงถึงจุดมุ่งหมายเฉพาะของการเรียนการสอนให้สอดคล้องกับการทดสอบด้วย

ในค่านความยาวของข้อความที่ยกมาให้ให้นักเรียนอ่าน ไม่มีข้อกำหนดตายตัวแน่นอนว่าจะคงมีความยาวเป็นจำนวนมากน้อยก็ค่า แต่อย่างไรก็ตามข้อความที่ยกมานั้นควรมีความยาวที่เหมาะสม และบรรจุเนื้อหาเพียงพอที่จะตั้งคำถามประมาณ 6 หรือ 7 คำถามได้

ในค่านเนื้อหาวิชา (Subject Matter) ปกติแล้ว จุดมุ่งหมายของการทดสอบจะเป็นตัวกำหนดและควบคุมเนื้อหาของข้อความที่จะนำไปทดสอบ เช่น ในการทดสอบความสามารถในการอ่าน เพื่อประเมินผลสัมฤทธิ์ของนักเรียนในภาคการเรียนหนึ่ง เนื้อหาวิชา ก็ควรจะสามารถครอบคลุมเนื้อหาที่สำคัญ หรือในการรับนักศึกษาเข้าศึกษาต่อในมหาวิทยาลัย เนื้อหาของข้อความต่าง ๆ ก็ควรจะมีความกว้างขวางหลายอย่าง นอกจากนี้เนื้อหาของข้อความในแต่ละตอนควรมีความสมบูรณ์ ชัดเจน และไม่ขาดตอน มีความสมบูรณ์เพียงพอ เนื้อหาก็คงกล่าวนี้ควรจะเป็นเนื้อหาในค่านความรู้โดยทั่วไป ไม่เป็นเนื้อหาเฉพาะที่จะต้องใช้ความรู้เฉพาะในค่านใดค่านหนึ่ง แต่ก็ไม่ได้หมายความว่า ถึงแม้จะไม่อ่านข้อความก็สามารถตอบคำถามได้อย่างถูกต้อง

ในค่านภาษาของข้อความควรคำนึงถึงภาษาในข้อความนั้นว่า มีระดับความยากง่ายในค่านโครงสร้าง และคำศัพท์ที่เหมาะสมกับระดับชั้นหรือเหมาะสมกับความสามารถของนักเรียนด้วย

ในค่านการเขียนตัวคำถาม หรือที่เรียกว่า "Stem" ของคำถามแต่ละข้อนั้น ควรจะมีรายละเอียดที่เพียงพอในการแนะ หรือช่วยให้นักเรียนเข้าใจในการที่จะนำรายละเอียดในตัวคำถามไปตอบหรือทำในตัวเลือกได้ ตัวอย่างเช่น ตัวคำถามที่มีคำว่า John เพียงคำเดียว ก็เป็นตัวคำถามที่ไม่เหมาะสม เพราะขาดรายละเอียดที่จำเป็น นอกจากนั้นการเขียนคำถามทุกคำถามควรจะเขียนอย่างระมัดระวัง เพื่อทดสอบว่านักเรียนได้อ่านข้อความที่ยกมาได้อย่างละเอียดและเข้าใจหรือไม่ คำถามที่ถามจะต้องเป็นคำถามที่นักเรียนจะต้อง

อ่านข้อความแล้วจึงจะตอบได้ นอกจากนั้นคำถามจะต้องใช้ภาษาง่าย ๆ ชัดเจนและเหมาะสม
กับระดับความรู้ของผู้เข้าสอบทั้งในคำนำศัพท์และโครงสร้าง

ในคานการเขียนตัวเลือก ควรคำนึงถึงการใช้ภาษาเช่นกัน ภาษาที่ใช้ควรจะง่าย
ในการตีความและง่ายในการอ่าน นอกจากนั้นตัวเลือกที่ถูกต้องไม่ควร เป็นคำตอบที่ยกออกมา
เหมือนทุกตัวอักษร จากข้อความที่อ่าน³⁰

สำหรับการวิจัยที่เกี่ยวข้องกับการสร้างแบบสอบความสามารถในการอ่านนี้ ผู้วิจัย
ขอแยกกล่าวเป็น การวิจัยในต่างประเทศ และการวิจัยในประเทศไทย

การวิจัยในต่างประเทศ

การสร้างแบบสอบเพื่อวัดความสามารถของนักเรียนที่เรียนภาษาอังกฤษเป็นภาษา
ที่สองได้แก่ แบบสอบที่สร้างขึ้นโดยสถาบันสอนภาษาอังกฤษของมหาวิทยาลัยมิชิแกน แบบสอบ
นี้มีชื่อว่า "Michigan Test of English Language Proficiency" ใช้ชื่อย่อว่า
"MTELP" แบบสอบนี้มีจุดมุ่งหมายในการทดสอบคือ วัดระดับความสามารถทางภาษาใน
ปัจจุบันของนักศึกษาว่าจะมีพื้นฐานความรู้พอเพียงในการ เรียนต่อในชั้นวิทยาลัยหรือมหาวิทยาลัย
หรือไม่ และสามารถจะเรียนรู้ต่อไปในอนาคตมากน้อยเพียงใด

แบบสอบนี้ประกอบด้วยแบบสอบทักษะทางด้านต่าง ๆ ได้แก่ ด้านไวยากรณ์
(Structure) จำนวน 40 ข้อ ด้านคำศัพท์ (Vocabulary) จำนวน 40 ข้อ และด้าน
ความเข้าใจในการอ่าน (Reading Comprehension) จำนวน 20 ข้อ เนื้อหาที่ใช้ใน
การสร้างแบบสอบนี้ใช้เนื้อหาที่เป็นปัญหาสำหรับชาวต่างประเทศแต่ไม่เป็นปัญหาสำหรับ
เจ้าของภาษา แบบสอบนี้กำหนดเวลาให้ทำรวม 75 นาที ข้อสอบแต่ละด้านไม่ได้แบ่งเวลา
แยกออกจากกัน แต่ให้ทำรวมกันไป นักเรียนที่มีความสามารถทางด้านภาษาสูง ก็จะทำแบบ
สอบได้ครบและทันเวลา ส่วนนักเรียนที่มีความสามารถทางด้านภาษาคำ ถ้าอาจทำข้อสอบไม่
ทัน และอาจไม่ทันได้ทำแบบสอบทางด้านกรอาน

³⁰ Ibid., pp. 60 - 2.

ในด้านการศึกษาทดสอบความสามารถในการอ่าน แบบสอบนี้ประกอบด้วยข้อความจำนวน 4 ข้อความ ซึ่งมีความยาว 139 - 217 คำ และคำถามจำนวน 5 ข้อ จอห์น แครโรล³¹ (John Carroll) กล่าวว่า แบบสอบนี้นอกจากจะวัดความสามารถในการอ่านของผู้เข้าสอบแล้ว ยังอาจจะวัดความสามารถทางคำศัพท์ปัญญาในด้านภาษา (Intelligent Verbal Ability) ของผู้เข้าสอบอีกด้วย

จอห์น แครโรล (John Carroll) ยังกล่าวอีกว่า แบบสอบชุดนี้ (MTELP) ควรจะใช้ข้อมูลอื่น ๆ ประกอบในการพิจารณาจัดนักศึกษาเข้าเรียน เพราะเป็นแบบสอบที่ไม่สามารถจะทำนายความสำเร็จทางการเรียนในอนาคตของนักศึกษาได้ อีกประการหนึ่งการที่นักเรียนสามารถจะทำคะแนนได้ดีในการสอบที่เป็นข้อเขียน ก็ไม่ได้หมายความว่า นักศึกษาที่เข้าเรียนในสถาบันที่ใช้ภาษาอังกฤษจะประสบความสำเร็จ และมีความสามารถในการพูดได้ก็เท่ากับการเขียน ดังนั้น ในการใช้ผลจากแบบสอบนี้เป็นข้อมูลในการคาดคะเนความสำเร็จในอนาคตของนักศึกษา ก็เป็นอันตรายอย่างยิ่ง

ถึงแม้ว่า จอห์น แครโรล (John Carroll) จะให้ความเห็นดังกล่าว แบบสอบนี้ (MTELP) จะไม่สามารถเป็นสิ่งที่คาดคะเนความสำเร็จในการเรียนของนักเรียนก็ตาม การวิจัยของ อลิซาเบท คัชแมน แฮมลิน³² (Elizabeth Cushman Hamlin) แห่งมหาวิทยาลัยออริกอน (University of Oregon) ได้ชี้ให้เห็นว่าคะแนนจากการสอบแบบสอบนี้มีส่วนสัมพันธ์กับหน่วยกิตที่นักเรียนทำได้

³¹Oscar K. Buros (ed.), The Sixth Mental Measurement Yearbook (New Jersey: The Gryphon Press Highland Park, 1965), pp. 638 - 60.

³²Elizabeth C. Hamlin, "An Analysis of the Relationships between the English Language Proficiency Scores for Entering Foreign Graduate Students and their Academic Achievement in an Advanced Degree Program," Dissertation Abstract, 33 (November, 1972), pp. 2125 - 26 A.

อลิซาเบท คัชแมน แฮมลิน (Elizabeth Cushman Hamlin) ได้ทำการวิจัยในปี ค.ศ. 1972 เพื่อศึกษาว่า ผู้ที่มีความสามารถทางภาษา จากผู้ที่ทำแบบทดสอบชุดนี้ (MTELP) จะมีความสัมพันธ์กับความสามารถในด้านกร เรียนมากกว่าผู้ที่ทำคะแนนใดต่ำกว่าหรือไม่ โดยได้สุ่มตัวอย่างนักศึกษาระดับปริญญาโท ที่มหาวิทยาลัยออริกอน จำนวน 213 คน ในจำนวนนี้เป็นนักศึกษาที่ได้คะแนนจากการสอบแบบทดสอบชุดนี้ ที่ได้คะแนนมากกว่า 80 คะแนน จำนวน 118 คน และผู้ที่ได้คะแนนต่ำกว่า 79 คะแนน จำนวน 95 คน นักศึกษาทั้งหมดคงกล่าวถึงการศึกษาอยู่ในมหาวิทยาลัยแห่งนี้ตั้งแต่ปี ค.ศ. 1966 ถึงปี ค.ศ. 1970 ผู้วิจัยพบว่า ความสามารถทางภาษา (English Proficiency) มีความสัมพันธ์กับหน่วยกิตที่นักศึกษาเรียนได้และนักศึกษาที่ทำคะแนนในแบบทดสอบชุดนี้ที่ได้คะแนนมากกว่า 80 คะแนน จะมีความสามารถทำหน่วยกิต ได้แตกต่างจากผู้ที่ได้คะแนนต่ำกว่า 79 คะแนน อย่างมีนัยสำคัญ

แบบทดสอบภาษาอีกชุดหนึ่งสำหรับนักเรียนต่างประเทศ ได้แก่แบบทดสอบ "The Test of English as a Foreign Language" (TOEFL) แบบสอบนี้สร้างขึ้นเพื่อทดสอบทักษะทางภาษาสำหรับนักศึกษาที่จะเข้าศึกษาต่อในวิทยาลัยหรือมหาวิทยาลัยในประเทศสหรัฐอเมริกา

แบบทดสอบชุดนี้ (TOEFL) ประกอบด้วยแบบสอบย่อย (Sub tests) ในด้านทักษะความเข้าใจในการฟัง (Listening Comprehension) โครงสร้าง (Structure) คำศัพท์ (Vocabulary) ความเข้าใจในการอ่าน (Reading Comprehension) และการเขียน (Writing)

แบบทดสอบในด้านกรอ่านประกอบด้วยข้อความต่าง ๆ และคำถามที่เกี่ยวกับข้อความที่อ่านนั้น Clinton I. Chase กล่าวว่า การออกแบบสอบวัดความสามารถในการอ่านนี้ นักเรียนที่ฉลาดจะอ่านคำถามแล้วเลือกอ่านเฉพาะข้อความที่ใ้หม่าก็จะสามารถทำแบบสอบได้ ซึ่งลักษณะการเลือกอ่าน (Skim) นี้ ถึงแม้จะเป็นทักษะที่จำเป็นในการค้นคว้าเมื่อนักเรียนเรียนอยู่ในมหาวิทยาลัยก็ตาม ก็ไม่ใช่เป็นทักษะที่แท้จริงของการอ่าน

ในด้านความเชื่อถือได้ของแบบสอบในด้านกรอ่านมีค่าเท่ากับ .84 ค่าความเชื่อถือโดยเฉลี่ยของทักษะต่าง ๆ ทั้งสี่ด้านเท่ากับ .97 แต่อย่างไรก็ตาม Clinton I. Chase กล่าวว่า เนื่องจากแบบสอบนี้ไม่ได้ใช้เวลา และอัตราความเร็วเข้ามาเป็นองค์ประกอบ -

ในการหาค่าความเชื่อถือได้ของแบบสอบชุดนี้ จึงไม่สามารถใช้เป็นเกณฑ์ในการคัดเลือกผู้เข้าสอบอย่างแท้จริง

นอกจากนั้น การที่จะใช้แบบสอบนี้เพื่อพิจารณาความสามารถในอนาคต (Predictor of - later criterion performance) หรือชี้ความสำเร็จในด้านการศึกษาในอนาคตนั้น ยังไม่สามารถจะทำได้ เพราะไคค่าความเที่ยงตรง (Validity) คำนวณไม่สามารถจะใช้เป็นมาตรฐานในการทำนายความสำเร็จในด้านการศึกษาของนักศึกษาได้ เช่น ค่าความเที่ยงที่มหาวิทยาลัยแคลิฟอร์เนีย (University of California) เท่ากับ .17 และ .26 ที่มหาวิทยาลัยวอชิงตัน (University of Washington) คลินตัน ไอ เซส³³ (Clinton I. Chase) กล่าวว่า การใช้ข้อมูลจากผลการสอบของแบบสอบชุดนี้ ในการคาดคะเนความสำเร็จด้านการศึกษาในอนาคตเป็นอันครายอย่างยิ่ง สถาบันที่ใช้ข้อมูลจากแบบสอบชุดนี้ควรจะต้องใช้ข้อมูลอื่น ๆ ประกอบด้วย

ยอร์ช โดมิโน³⁴ (George Domino) เป็นอีกผู้หนึ่งที่ทำให้เห็นว่า การที่สถาบันต่าง ๆ ในสหรัฐอเมริกาใช้แบบสอบนี้ ในการคาดคะเนความสามารถในการเรียนของนักศึกษานั้น เป็นการไม่ถูกต้อง เพราะความสามารถในด้านภาษาไม่ได้เป็นส่วนประกอบเพียงส่วนเดียวที่จำเป็นต่อผลการเรียน ยอร์ช โดมิโน (George Domino) กล่าวว่า ถ้าสถาบันต่าง ๆ เพียงแต่ต้องการทราบว่า ผู้สมัครมีพื้นฐานความรู้ภาษาอังกฤษเท่าใด การทดสอบในคำศัพท์ (Vocabulary) แต่เพียงด้านเดียวก็น่าจะเป็นการเพียงพอ

เชย์ และ เฮลกา รุท³⁵ (Shay และ Helga Ruth) แห่งมหาวิทยาลัยเวสต์เวอร์จิเนีย (West Virginia University) ได้ทำการวิจัยแบบสอบ (TOEFL) นี้

³³ Buross (ed.), op.cit., p. 670.

³⁴ Buross (ed.), op.cit., p. 671.

³⁵ Shay Helga Ruth, "Affect of Foreign Student Language Proficiency on Academic Performance" Dissertation Abstracts, 36 (October, 1975), pp. 183 A - 84 A.

เพื่อศึกษาว่าแบบสอบดังกล่าวนี้ จะสามารถเป็นตัวแทนความสำเร็จในด้านกรเรียนหรือไม่ โดยใ้กลุ่มตัวอย่างนักศึกษาระดับปริญญาโทชาวต่างประเทศในมหาวิทยาลัยเวสต์เวอร์จิเนีย จำนวน 174 คน ที่เข้าในมหาวิทยาลัยแห่งนี้ในปี ค.ศ. 1964 และปี ค.ศ. 1974 ผู้วิจัยพบว่า คะแนนเฉลี่ยของนักศึกษาที่เรียนไม่มีความสัมพันธ์กับคะแนนที่ได้จากข้อทดสอบ (TOEFL) นี้ ดังนั้นผู้วิจัยจึงสรุปผลของการวิจัยว่า คะแนนในแบบสอบ (TOEFL) เป็นคะแนนที่บอกว่ นักศึกษามีสัมฤทธิ์ผล (Proficiency) ทางภาษาในปัจจุบันเท่านั้น แต่จะไม่เป็นสิ่งที่ชี้ว่า นักศึกษาจะประสบผลสำเร็จทางด้านการเรียนเพียงใด

นอกจากแบบสอบเพื่อทดสอบความสามารถในการอ่านดังกล่าวแล้ว ยังมีผู้สร้างแบบสอบในการวัดความสามารถในการอ่านอื่น ๆ อีก ในปี ค.ศ. 1969 ซีดริค บี. จีเวลล์³⁶ (Cedric B. Jewell) ได้สร้างแบบสอบความสามารถในการอ่านของนักเรียนชั้นมัธยมตอนปลาย เรียกแบบสอบนี้ว่า "The High School Reading Comprehension" เพื่อทดสอบความสามารถในการอ่านด้านความเข้าใจความหมายตามตัวอักษรที่อ่าน (Literal Meaning) ด้านความสามารถในการติดตามคำสั่ง (Follow Direction) ด้านความเข้าใจความหมายซ่อนเร้น (Implicit or latent meaning) ด้านการสรุปหรือคาดคะเน (Inference of Prediction) ด้านความทรงจำ (Appreciation) ด้านความเข้าใจวัตถุประสงค์ของผู้แต่ง ด้านความเข้าใจอารมณ์ของผู้แต่ง (Sensing mood) และด้านแนวการเขียนของผู้แต่ง เช่นการอุปมาอุปไมย

ผู้สร้างแบบสอบได้เลือกข้อความที่มีความยาวตั้งแต่ 100 - 200 คำ จากหนังสือแบบเรียนต่าง ๆ ได้แก่ แบบเรียนวิชาภาษาอังกฤษ คณิตศาสตร์ สังคมศึกษา และวิทยาศาสตร์ เป็นจำนวน 12 ข้อความ และคำถามในการวัดความสามารถในการอ่านข้อความละ 8 ข้อ

³⁶William T. Fagan, Chales R. Cooper and Julie M. Jensen, Measures for Research and Evaluation in the English Language Arts (Illinois: ERIC Clearing house on Reading and Communication Skills and National Council of Teacher of English, 1975), pp. 20 - 1.

ซึ่งนำไปทดลอง (Pilot) ในเดือนพฤษภาคม ค.ศ. 1967 กับนักเรียนในเกรด 10 และ เกรด 12 ระยะที่สองได้พัฒนาแบบสอบถามเป็นคำถามจำนวน 138 ข้อ หลังจากที่ได้มีการอภิปรายกับครูที่มีประสบการณ์ กับผู้เชี่ยวชาญในด้านการอ่าน และผู้ให้คำแนะนำในด้านลักษณะ (Format) ของแบบสอบถาม ผู้วิจัยเหลือข้อสอบไว้จำนวน 81 ข้อ หลังจากนั้นผู้วิจัยได้นำแบบสอบถามไปใช้กับนักเรียนในเกรด 10 จำนวน 92 คน นักเรียนในเกรด 11 จำนวน 74 คน และนักเรียนในเกรด 12 จำนวน 94 คน ข้อทดสอบซึ่งเสร็จสมบูรณ์แล้วเป็นข้อสอบที่ได้จากการวิเคราะห์ความยากง่าย (Item Analysis) และได้จากความเห็นของคณะกรรมการ ซึ่งให้คะแนนความเห็นของข้อสอบแต่ละข้อ โดยใช้ "five point scale" ข้อทดสอบที่เลือกแล้ว มีความยากอยู่ในระหว่าง 20 - 80 เปอร์เซนต์ และมีค่าความเที่ยง (Item Validity) ไม่น้อยกว่า .20 รวมข้อทดสอบในแบบสอบถามที่ได้จำนวน 44 ข้อ

ผู้สร้างแบบสอบถามได้นำแบบสอบถามไปใช้กับตัวอย่างประชากรในโรงเรียนต่าง ๆ ในเกรด 10 และ เกรด 12 จำนวน 1052 คน ที่มีความสามารถแตกต่างกัน ได้ค่าต่าง ๆ ดังนี้คือ นักเรียนในเกรด 10 ได้ค่าเฉลี่ย (Mean) เท่ากับ 19.95 ค่าส่วนเบี่ยงเบนมาตรฐาน (S.D.) เท่ากับ 7.42 และค่าความคลาดเคลื่อน (S.E.) เท่ากับ 3.07 ส่วนในเกรด 11 ได้ค่าต่าง ๆ ตามลำดับดังนี้คือ 20.91, 7.57 และ 3.09 และในเกรด 12 ได้ค่าต่าง ๆ เท่ากับ 24.60, 7.66 และ 3.09 ตามลำดับ

ส่วนค่าความเที่ยงตรงของแต่ละข้อ (Item Validity) ในเกรด 10 ได้เท่ากับ .35 ในเกรด 11 ได้เท่ากับ .35 และในเกรด 12 เท่ากับ .36 ค่าความเชื่อถือได้ในเกรด 10 เท่ากับ .83 เกรด 11 เท่ากับ .84 และเกรด 12 เท่ากับ .85

ในปี ค.ศ. 1970 รุท แมรี จาร์วิส³⁷ (Ruth Mary Jarvis) แห่งมหาวิทยาลัยบอสตัน ได้สร้างแบบสอบถามความสามารถในการอ่านขึ้น เพื่อวัดความรู้ว่า นักเรียนในระดับเกรด 2

³⁷Ruth Mary Jarvis, "Construction and Evaluation for a Critical Reading and Thinking Measure for Grade Two and Three," Dissertation Abstracts, 31 (June, 1971), pp. 6469 A - 70 A.

และเกรด 3 มีความสามารถในการอ่านและคิดอย่างวิพากษ์วิจารณ์ (Critical Reading and Thinking) เพียงไร แบบสอบนี้แยกทักษะในด้านกรอ่านเป็นด้าน ๆ คือ ด้านความจริง (Fact) ด้านความสัมพันธ์ (Relevancy) ด้านการสรุปความเห็น (Inference) และด้านการสรุปหลักเกณฑ์ (Generalization)

แบบสอบในการอ่านอย่างวิพากษ์วิจารณ์ (Critical Reading) ได้นำไปทดลองใช้ (Tryout) กับนักเรียนในเกรด 1 เกรด 2 และเกรด 3 ส่วนข้อสอบชุดที่สองซึ่งเป็นการทดสอบความคิดอย่างวิพากษ์วิจารณ์ ได้นำไปทดลองใช้กับนักเรียนในเกรด 2 และเกรด 3

แบบสอบที่นำไปใช้กับกลุ่มตัวอย่างในการทดสอบครั้งหลัง (Post Test) ผู้วิจัยได้นำแบบสอบไปใช้กับนักเรียนในเกรด 2 และเกรด 3 ผู้วิจัยได้วิเคราะห์แบบสอบและได้ค่าความเชื่อถือของแบบสอบในระดับเกรด 2 เท่ากับ .74 สำหรับในเกรด 3 ได้ค่าความเชื่อถือเท่ากับ .87 ที่ระดับ .01

ในปี ค.ศ. 1975 จอห์นสัน และ บาร์บารา แอน สตับส์³⁸ (Johnson and Barbara Anne Stubbs) ได้สร้างแบบสอบวัดความเข้าใจในการอ่านจากข้อความที่ผู้วิจัยสร้างขึ้น เพื่อใช้ข้อความดังกล่าวประกอบการอ่านในวิชาสังคมศึกษาในระดับมัธยมศึกษาตอนต้น และมัธยมศึกษาตอนกลาง

ผู้วิจัยได้สร้างข้อความ (Reading Passages) โดยเลือกเรื่อง (Topics) ต่าง ๆ ที่คิดว่าเหมาะสมกับระดับชั้นของนักเรียน โดยใช้ข้อความจำนวน 10 ข้อความ และใช้ขบวนการของ Cloze - Procedure ซึ่งมีความยากตามระดับชั้น มีความยากสูงกว่าระดับชั้น และมีความยากต่ำกว่าระดับชั้น

³⁸ Johnson and Barbara Stubbs, "The Developments and Validation of ten Passages Written for Informal Studies Reading Inventories for Use in Secondary Schools," Dissertation Abstracts, 36 (November, 1975), pp. 2737 A - 38 A.

ผู้สร้างแบบสอบได้นำแบบสอบไปทดลองใช้ (Pre Test) กับนักเรียนจำนวน 476 คน ผลของการวิจัยปรากฏว่า ข้อความต่าง ๆ ที่ผู้ทดสอบสร้างขึ้นนั้นในจำนวน 10 ข้อความ (Passages) มีข้อความอยู่ถึง 9 ข้อความที่จะต้องเขียนและปรับปรุงใหม่ เพื่อให้ข้อความทั้ง 9 ข้อความดังกล่าวง่ายเข้า ผู้วิจัยสรุปว่า จากผลที่นักเรียนอ่านข้อความต่าง ๆ นั้น ชี้ให้เห็นว่า หนังสือแบบเรียนที่ครูให้นักเรียนอ่านมีความยากเหนือระดับชั้นของนักเรียน ตั้งแต่ 2 ถึง 6 ระดับ เพราะผลของการวิจัยมีข้อความอยู่ข้อความเดียวที่นักเรียนอ่านได้นักเรียนจำนวน 70 เปอร์เซ็นต์ อ่านข้อความทั้ง 9 ข้อความ อย่างยากลำบาก (Frustration Level)

การวิจัยในประเทศไทย

งานวิจัยเกี่ยวกับการสร้างแบบสอบความสามารถในการอ่านซึ่งผู้วิจัยจะศึกษา ผู้วิจัยคิดว่า จะเป็นประโยชน์ในด้านการเรียนการสอนภาษาอังกฤษเนื่องจาก การอ่านเป็นทักษะที่สำคัญและงานวิจัย เกี่ยวกับการสร้างแบบสอบในค่านนี้ในประเทศไทย ยังมีผู้ทำการวิจัยน้อยมาก ส่วนใหญ่จะเป็นงานวิจัยที่เกี่ยวกับการสร้างแบบสอบในค่านอื่น ซึ่งพอสรุปได้ดังนี้

ในปี พ.ศ. 2507 ไชศรี วรชกวิศา³⁹ ได้ทำการวิจัยเรื่อง "การสร้างแบบทดสอบความพร้อมในการอ่านและศึกษาความพร้อมในการอ่านของนักเรียนชั้นประถมศึกษาปีที่ 1 บางโรงเรียนในภาคการศึกษา 1" โดยผู้วิจัยได้สร้างแบบสอบความพร้อมเฉพาะด้านความสามารถในการแยกสิ่งต่าง ๆ ด้วยคำ ประกอบด้วยแบบสอบ 4 ค่าน คือ ค่านคำศัพท์ ค่านการรู้จักความแตกต่าง ค่านการรู้จักความเหมือน และค่านการเชื่อมโยงความคิด ได้ค่าความเชื่อถือเท่ากับ .92 และค่าความเที่ยงเท่ากับ .66

³⁹ไชศรี วรชกวิศา, "การสร้างแบบทดสอบความพร้อมในการอ่านและศึกษาความพร้อมในการอ่านของนักเรียนชั้นประถมศึกษาปีที่ 1 บางโรงเรียนในภาคการศึกษา 1" (ปริญญาานิพนธ์การศึกษามหาบัณฑิต แขนงวิชาจิตวิทยาพัฒนาการ วิทยาลัยวิชาการศึกษา ประสานมิตร, 2507).

ในปี พ.ศ. 2510 จิตต์โส มฤงฉิน⁴⁰ ได้ทำการวิจัย เรื่อง "การสร้างแบบทดสอบความพร้อมที่จะอ่านด้านความสามารถในการแยกเสียงที่ได้ยิน สำหรับนักเรียนชั้นประถมศึกษาปีที่ 1" โดยผู้วิจัยได้สร้างแบบสอบ 3 ด้าน คือ ด้านการฟังเสียงพยัญชนะ ด้านการฟังเสียงสระ และด้านการฟังเสียงวรรณยุกต์ แบบทดสอบที่สร้างขึ้นนี้ มีความเชื่อถือเท่ากับ .85 ค่าความเที่ยงตรงเท่ากับ .62

ในปี พ.ศ. 2513 พรณี ชูทัย⁴¹ ได้วิจัยเรื่อง "การสร้างและทดลองใช้แบบทดสอบความพร้อมในการอ่านด้านความสามารถแยกเสียงที่ได้ยินสำหรับนักเรียนชั้นประถมศึกษาปีที่ 1" แบบทดสอบที่สร้างขึ้น มี 4 ด้าน คือ ด้านความสามารถในการแยกเสียงที่ได้ยิน ด้านการรู้จักพยัญชนะ ด้านการรู้จักสระและวรรณยุกต์ และด้านอัตราการเปลี่ยนคำ ได้ค่าความเชื่อถือเท่ากับ .81 และค่าความเที่ยงตรงเท่ากับ .63

ในด้านการสร้างแบบทดสอบความพร้อมนี้ ในปี พ.ศ. 2517 ชัยยุทธ บุญยสวัสดิ์⁴² ได้วิจัยเรื่อง "การสร้างแบบทดสอบความพร้อมในการอ่านด้านความสามารถจำแนกด้วยคำสำหรับเด็กก่อนเข้าเรียน" ผู้วิจัยได้สร้างแบบทดสอบรวม 5 ด้าน คือ ด้านความพร้อมในเรื่องศัพท์ และความเข้าใจศัพท์ ซึ่งได้ค่าความเชื่อถือ .57 ด้านการรู้จักหาสิ่งที่หายไป

⁴⁰จิตต์โส มฤงฉิน, "การสร้างแบบทดสอบความพร้อมที่จะอ่านด้านความสามารถในการแยกเสียงที่ได้ยิน สำหรับนักเรียนชั้นประถมศึกษาปีที่ 1" (ปริญญาานิพนธ์การศึกษามหาบัณฑิต แผนกวิชาจิตวิทยาการศึกษา วิทยาลัยวิชาการศึกษา ประสานมิตร, 2510).

⁴¹พรณี ชูทัย, "การสร้างและทดลองใช้แบบทดสอบความพร้อมในการอ่านด้านความสามารถแยกเสียงที่ได้ยิน สำหรับนักเรียนชั้นประถมศึกษาปีที่ 1" (ปริญญาานิพนธ์การศึกษามหาบัณฑิต แผนกจิตวิทยา วิชาการศึกษา วิทยาลัยวิชาการศึกษา ประสานมิตร, 2513).

⁴²ชัยยุทธ บุญยสวัสดิ์, "การสร้างแบบทดสอบความพร้อมในการอ่านด้านความสามารถจำแนกด้วยคำ สำหรับเด็กก่อนเข้าเรียน" (วิทยานิพนธ์ปริญญาามหาบัณฑิต แผนกประถมศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2517).

มีค่าความเชื่อถือเท่ากับ .61 ค่าการรู้จักแยกความเหมือน ค่าความเชื่อถือเท่ากับ .80
 ค่าการรู้จักแยกความแตกต่าง ค่าความเชื่อถือเท่ากับ .85 และค่าการเชื่อมโยงความคิด
 ได้ค่าความเชื่อถือเท่ากับ .79 และค่าความเชื่อถือโดยเฉลี่ยทั้ง 5 ค่า มีค่าเท่ากับ .80

งานวิจัยที่เกี่ยวข้องกับการสร้างแบบทดสอบความสามารถในการสะกดคำ คือ งานวิจัย
 ของ สมประสงค์ สถาปิตานนท์⁴³ ทำการวิจัยเรื่อง "การสร้างแบบทดสอบเขียนสะกดคำ
 ประถมปีที่ 1 และศึกษาผลสัมฤทธิ์ในการเขียนสะกดคำของนักเรียนใน 11 โรงเรียน
 ในภาคการศึกษา 1" ซึ่งได้ทำการวิจัยในปี พ.ศ. 2507 ผู้วิจัยได้สร้างแบบทดสอบ
 เขียนสะกดคำ ซึ่งประกอบด้วยรูปภาพ 90 ภาพ เป็นรูปภาพคำกริยาจำนวน 40 ภาพ และ
 เป็นรูปภาพคำนามจำนวน 50 ภาพ เมื่อได้นำไปทดลองแล้วได้แบบทดสอบจำนวน 60 ข้อ
 มีค่าความเชื่อถือเท่ากับ .9763 ค่าความเที่ยงตรงสำหรับโรงเรียน ก. เท่ากับ .579
 สำหรับโรงเรียน ข. ค่าความเที่ยงตรงเท่ากับ .775 ค่าความยากเฉลี่ยเท่ากับ .5535
 ในด้านการศึกษาผลสัมฤทธิ์จากนักเรียน 11 โรงเรียนนี้ ผู้วิจัยพบว่า การแจกแจงของ
 คะแนนไม่เป็นโค้งปกติ ซึ่งหมายความว่า อาจแบ่งนักเรียนได้เป็น 2 กลุ่ม คือกลุ่มที่มีความ
 สามารถสูงในการสะกดคำ และกลุ่มที่ไม่มีความสามารถในการสะกดคำ ส่วนในด้านความ
 แตกต่างของผลสัมฤทธิ์ เด็กหญิงมีผลสัมฤทธิ์ในการสะกดคำสูงกว่าเด็กชาย เด็กที่วัยต่างกัน
 มีผลสัมฤทธิ์ในการเขียนสะกดคำใกล้เคียงกัน และเด็กชั้นประถมศึกษาปีที่ 2 มีผลสัมฤทธิ์สูงกว่า
 นักเรียนที่สอบตกซ้ำชั้นประถมศึกษาปีที่ 1

การวิจัยเกี่ยวกับการสร้างแบบทดสอบอีกเรื่องหนึ่งคือ การวิจัยเรื่อง "การสร้าง
 แบบทดสอบความถนัดในการเขียนสำหรับนักเรียนที่จบชั้นประถมศึกษาปีที่ 5 และความถนัดใน

⁴³สมประสงค์ สถาปิตานนท์, "การสร้างแบบทดสอบเขียนสะกดคำ ประถมปีที่ 1
 และศึกษาผลสัมฤทธิ์ในการเขียนสะกดคำของนักเรียนใน 11 โรงเรียน ในภาคการศึกษา 1"
 (ปริญาานิพนธ์การศึกษามหาบัณฑิต วิทยาลัยวิชาการศึกษา ประสานมิตร, 2507).

การเรียนของนักเรียนชั้นประถมศึกษาปีที่ 5" ซึ่ง สหส สุขเวช⁴⁴ ได้วิจัยในปี พ.ศ. 2509 ผู้วิจัยได้สร้างแบบทดสอบความถนัด 9 แบบ คือ การเลือกคำที่มีความหมายผิดพวกไปจากคำอื่น ๆ ที่กำหนดให้ การจัดเรียงคำโคลง ๆ หลายคำให้เป็นประโยคที่ถูกต้อง การเติมตัวเลขในอนุกรมตัวเลขที่กำหนดให้ การทบทวนกระบวนการพื้นฐานทางเลขคณิต โดยการเติมตัวเลข ความรู้และความจำเกี่ยวกับสนเทศ (Information) การเลือกรูปภาพอื่นที่มีคุณลักษณะร่วมกัน การเติมเครื่องหมายลงในตารางให้สอดคล้องกับความสัมพันธ์ของเครื่องหมายที่กำหนดให้ และการเลือกคู่คำที่มีความหมายสอดคล้องกับคู่คำที่กำหนดให้ แบบทดสอบทั้ง 9 แบบนี้ ผลของการวิจัยปรากฏว่า ค่าความยากของแบบทดสอบทั้ง 83 ข้อ มีความยากตั้งแต่ .14 ถึง .78 ค่าการจำแนกเท่ากับ .31 ถึง .72 ค่าความเชื่อถือใจของแบบทดสอบเท่ากับ .936 ค่าความเที่ยงในการทำนายผลสัมฤทธิ์วิชาเลขคณิต ความเข้าใจในการอ่าน การเขียนสะกดคำ วิทยาศาสตร์ และคะแนนรวมเป็นอัตราส่วนร้อยละ จากการสอบไล่ เท่ากับ .77, .58, .56, .61 และ .72 ตามลำดับ ผู้วิจัยพบว่า นักเรียนชายมีความถนัดในการเรียนสูงกว่านักเรียนหญิง นักเรียนที่มีอายุต่างกัน มีความถนัดในการเรียนไม่ต่างกัน นักเรียนที่มาจากครอบครัวที่มีคามารดาประกอบอาชีพค้าขาย มีความถนัดในการเรียนสูงกว่านักเรียนที่มาจากครอบครัวที่มีคามารดาบริหารราชการ และอาชีพที่ให้บริการ

สำหรับงานวิจัยที่เกี่ยวกับการสร้างแบบสอบอีกเรื่องหนึ่งคือ งานวิจัยของ นฤมล ภัทรภาณุ⁴⁵ ได้วิจัยเรื่อง "การสร้างแบบทดสอบผลสัมฤทธิ์ไวยากรณ์อังกฤษสำหรับนักเรียน

⁴⁴ สหส สุขเวช, "การสร้างแบบทดสอบความถนัดในการเรียนสำหรับนักเรียนที่จบชั้นประถมศึกษาปีที่ 5 และความถนัดในการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 5" (ปริญญาานิพนธ์ การศึกษามหาบัณฑิต วิทยาลัยวิชาการศึกษา ประสานมิตร, 2509).

⁴⁵ นฤมล ภัทรภาณุ, "การสร้างแบบทดสอบผลสัมฤทธิ์ไวยากรณ์อังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่สาม ในโรงเรียนมัธยมแบบประสมในกรุงเทพมหานคร" (วิทยานิพนธ์ ปริญญามหาบัณฑิต แผนกมัธยมศึกษา มัธยมศึกษาวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2517).

ชั้นมัธยมศึกษาปีที่สาม ในโรงเรียนมัธยมแบบประสมในกรุงเทพมหานคร" ซึ่งได้ทำการวิจัย ในปี พ.ศ. 2516 โดยได้สร้างแบบสอบผลสัมฤทธิ์ไวยากรณ์อังกฤษ สำหรับนักเรียนที่ใช้แบบเรียน Intermediate English for Thailand เล่มสาม และจากหนังสือ Quick Reference to Intermediate English Course จำนวน 120 ข้อ ได้แบบสอบ 106 ข้อ ซึ่งมีระดับความยากระหว่าง .25 ถึง .79 ค่าอำนาจจำแนกระหว่าง .22 ถึง .70 และค่าความเชื่อถือได้เท่ากับ .94

สำหรับงานวิจัยที่เกี่ยวกับการสร้างแบบสอบการอ่านโดยตรง เป็นการสร้างแบบทดสอบในโครงการการ เรียนการสอนตามความแตกต่างแต่ละบุคคลของสถาบันศูนย์สอนภาษาอังกฤษ

แบบทดสอบการอ่านที่สร้างขึ้นนี้ สถาบันศูนย์สอนภาษาอังกฤษได้สร้างขึ้นสำหรับใช้ทดสอบความสามารถในการอ่านของนักเรียนชั้นประถมศึกษาปีที่ 5 และปีที่ 6 ที่ใช้แบบเรียน JILAP (Jacaranda Individualized Language Arts Program) และแบบเรียน The Oxford Course English for Thailand เล่มหนึ่ง สำหรับชั้นประถมศึกษาปีที่ 5 และสำหรับชั้นประถมศึกษาปีที่ 6 โดยให้อ่านข้อความและตามคำถามแบบเลือกตอบ

แบบทดสอบการอ่านสำหรับวัดความเข้าใจในการอ่านของนักเรียนชั้นประถมศึกษาปีที่ 5 มีสองฉบับ ฉบับที่ใช้เนื้อหาจากแบบเรียน JILAP ในชุดสีน้ำตาล ซึ่งมีจำนวน 15 ข้อ ได้ค่าเฉลี่ยเท่ากับ 7.36 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 3.29 และค่าความเชื่อถือได้เท่ากับ .70 ส่วนฉบับที่ใช้เนื้อหาจากแบบเรียน The Oxford English Course for Thailand เล่มหนึ่ง ได้ค่าเฉลี่ยเท่ากับ 6.59 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 2.62 และค่าความเชื่อถือได้เท่ากับ .49

สำหรับแบบทดสอบการอ่านสำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 มีสองฉบับด้วยกันคือ ฉบับที่ใช้เนื้อหาจากแบบเรียน JILAP ชุดสีน้ำตาล จำนวนข้อสอบ 15 ข้อ ได้ค่าเฉลี่ยเท่ากับ 9.72 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 3.45 และค่าความเชื่อถือได้เท่ากับ .76 ส่วนฉบับที่ใช้เนื้อหาจากแบบเรียน The Oxford English Course for Thailand เล่มหนึ่ง จำนวน 15 ข้อ ได้ค่าเฉลี่ยเท่ากับ 6.57 ค่าส่วนเบี่ยงเบนมาตรฐาน 2.62 และค่าความเชื่อถือได้เท่ากับ .49