

การสร้างและพัฒนาแบบทดสอบภาษาอังกฤษเพื่อการสื่อสารในการทำงาน Test Design and Development of Test of English for Communication in the Workplace (TEC-W)

จิรดา วุฒธยากร^{1*} และ กุมภการ สวัสดิ์โกมล²

Jirada Wudthayagorn^{1*} and Kumphan Sawaddikomon²

Received: October 30, 2023; Revised: November 28, 2023; Accepted: December 07, 2023

บทคัดย่อ

สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. มีหน้าที่และบทบาทสำคัญประการหนึ่งในการสร้างและพัฒนาแบบทดสอบระดับชาติ เพื่อใช้วัดและประเมินความสามารถในการใช้ภาษาอังกฤษของผู้สอบชาวไทยโดยเน้นนักเรียนในระบบการศึกษาภาคบังคับและการศึกษาขั้นพื้นฐาน นอกจากนี้ สทศ. ยังได้ตระหนักถึงหน้าที่และบทบาทในการสร้างแบบทดสอบระดับชาติโดยเน้นผู้สอบไทยในวัยทำงาน ดังนั้น สทศ. จึงได้ริเริ่มสร้างและพัฒนาแบบทดสอบภาษาอังกฤษเพื่อการสื่อสารในการทำงาน เพื่อใช้วัดความสามารถในการใช้ภาษาอังกฤษสำหรับกลุ่มบุคคลดังกล่าวที่เหมาะสมกับบริบทไทย และพัฒนาแบบทดสอบให้ได้รับการยอมรับในระดับนานาชาติ แบบทดสอบนี้เริ่มสร้างใน พ.ศ. 2561 และพัฒนาต่อเนื่องเรื่อยมา บทความนี้บรรยายและอธิบายที่มาของแบบทดสอบการสร้างและพัฒนาแบบทดสอบภาษาอังกฤษเพื่อการสื่อสารในการทำงาน การแปลผลคะแนนเทียบกับ Common European Framework of Reference (CEFR) อรรถประโยชน์ของแบบทดสอบตลอดจนทิศทางการวัดและประเมินผลภาษาอังกฤษในระดับมหภาคของประเทศไทย

คำสำคัญ: การสร้างแบบทดสอบ การพัฒนาแบบทดสอบภาษาอังกฤษเพื่อการสื่อสารในการทำงาน

¹ รองศาสตราจารย์ ดร. ผู้อำนวยการสถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย

¹ Assistant Professor., Director of Chulalongkorn University Language Institute

² นักวิชาการวิจัยและพัฒนา สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)

² Researcher and developer., National Institute of Educational Testing Service (Public Organization)

*Corresponding Author e-mail: Jirada.W@chula.ac.th

Abstract

One of the significant roles and responsibilities of the National Institute of Educational Testing Service (NIETS) is to design and to develop a national test aimed for the English language proficiency assessment of Thai people, especially in compulsory and fundamental education system. In addition, NIETS realized our national competent agent of creating and improving the English testing for working age people. Hence, NIETS initiated and proved the English for Communication in the Workplace (TEC-W) to be served as national testing that is compatible with the Thai context as well as the international recognition. TEC-W was launched in 2018 and has been continuously progressed. This article subjected to describe and to elucidate the background of the test, including design and development, score comparing of TEC-W to Common European Framework of Reference (CEFR), test utilities, measuring and evaluated directions of English language skills in the large scale level of Thailand.

Keywords: Test design, Test development, Test of English for Communication in the Workplace

บทนำ

ภาษาอังกฤษเป็นเครื่องมือที่สำคัญในการพัฒนาความสามารถและยกระดับอาชีพของคนทำงานในปัจจุบัน เนื่องจากการใช้ภาษาอังกฤษในการทำงานทั้งในอุตสาหกรรมต่าง ๆ และแวดวงธุรกิจ ไม่ว่าจะเป็น การท่องเที่ยว การศึกษา ธุรกิจระหว่างประเทศ หรืองานวิจัย องค์กรในประเทศไทยมีผู้ที่มีความสามารถใช้ภาษาอังกฤษในระดับดีจำนวนน้อย ในขณะที่ค่าเฉลี่ยสัดส่วนการใช้ภาษาอังกฤษสำหรับทักษะการฟัง พูด อ่าน เขียน ในการทำงานในปัจจุบันมีแนวโน้มสูงขึ้นเป็นสัดส่วน 20-80% แตกต่างกันไปตามรูปแบบงาน ดังนั้น หากองค์กรในประเทศไทยสามารถพัฒนาทักษะภาษาอังกฤษได้ก็จะสามารถยกระดับการแข่งขันได้ในระดับสากล ภาษาอังกฤษเป็นภาษาต่างประเทศที่ไม่ว่าภาครัฐและเอกชนในประเทศไทยเลือกใช้เป็นเครื่องมือในการสื่อสารลำดับต้น ๆ สถาบันการศึกษาตั้งแต่การศึกษาภาคบังคับ การศึกษาขั้นพื้นฐาน การอุดมศึกษา กำหนดให้ภาษาอังกฤษเป็นวิชาหนึ่งในหลักสูตรที่นักเรียนและนักศึกษาจำเป็นต้องเรียนและต้องสอบ เมื่อเขาเหล่านั้นเข้าสู่ตลาดแรงงาน นายจ้างมักกำหนดให้ความสามารถในการใช้ภาษาอังกฤษเป็นเกณฑ์หนึ่งในการคัดเลือกผู้สมัครเข้าทำงาน ชัยญนิษฐ์ อินทร (2563) กล่าวว่า การสื่อสารภาษาอังกฤษนับเป็นเครื่องมือที่สำคัญ ไม่เพียงแต่เพื่อการประชาสัมพันธ์หรือเสริมสร้างภาพลักษณ์ที่ดี ก่อให้เกิดความเชื่อมั่นและศรัทธาต่อองค์กรจากผู้มีส่วนเกี่ยวข้อง แต่การสื่อสารภาษาอังกฤษที่มีประสิทธิภาพยังก่อให้เกิดความร่วมมือ ประสานการทำงาน มีความเข้าใจในเป้าหมายหรือทิศทางการทำงานทั้งภายในและระหว่างหน่วยงานได้เป็นอย่างดี Ufiera (2016) ยังกล่าวว่าความสามารถในการใช้ภาษาอังกฤษมีความสัมพันธ์กับรายได้อย่างมีนัยสำคัญ ความสามารถในการใช้ภาษาอังกฤษช่วยให้บุคคลนั้นเข้าถึงเทคโนโลยีต่าง ๆ ที่มีข้อแนะนำหรือวิธีการใช้งานสารสนเทศต่าง ๆ ที่เป็นภาษาอังกฤษได้ง่ายและมีประสิทธิภาพมากขึ้น ความรู้ความสามารถในการใช้ภาษาอังกฤษเพื่อการเจรจาต่อรองในธุรกรรมต่าง ๆ ทำให้ต้นทุนการค้าลดลง โอกาสในการนำเข้าส่งออกมีมากขึ้น สิ่งต่าง ๆ เหล่านี้นำมาซึ่งรายได้ที่มากขึ้น สถาบันสอนภาษา NOVA ONLINE (2023) ได้กล่าวถึงความสำคัญของภาษาอังกฤษในด้านธุรกิจและการทำงาน และด้านการศึกษา เพื่อสร้างโอกาสในการทำงานระหว่างประเทศ สร้างเสถียรภาพในอาชีพ สร้างความได้เปรียบในการเข้าถึงข้อมูล และสร้างโอกาสสื่อสารแบบไร้พรมแดน เข้าถึงข้อมูลและแหล่งความรู้ ใช้ในการศึกษาต่อ ใช้ในการเรียนรู้เกี่ยวกับวัฒนธรรมและความหลากหลาย

ต่อเนื่องกับฉันทศน์นี้คือเรื่องเครื่องมือที่ใช้ในการวัดและประเมินผลความสามารถในการใช้ภาษาอังกฤษ เครื่องมือที่มักนึกถึงคือแบบทดสอบมาตรฐานทางภาษาอังกฤษ ซึ่งแบบทดสอบที่นิยมกันทั่วไปมักเป็นแบบทดสอบที่ได้รับการยอมรับในระดับนานาชาติ เช่น TOEFL iBT (Test of English as a Foreign Language Internet-based Test) ที่มีวัตถุประสงค์เพื่อวัดความสามารถของผู้เรียนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศว่า จะสามารถศึกษาเล่าเรียนและใช้ชีวิตในสิ่งแวดล้อม (เน้นวิชาการ) ที่เป็นภาษาอังกฤษ (แบบอเมริกัน) ได้หรือไม่ และมากน้อยเพียงใด เช่นเดียวกับแบบทดสอบ IELTS (International English Language Testing System) ที่มีวัตถุประสงค์ในการวัดเช่นเดียวกันกับ TOEFL iBT แต่ลักษณะของ

แบบทดสอบ (test task) ต่างกัน ผู้สร้างและพัฒนาคนละกลุ่มกัน หรือแบบทดสอบ TOEIC (Test of English for International Communication) ที่ใช้วัดความสามารถในการใช้ภาษาอังกฤษในการสื่อสารระดับนานาชาติที่เน้นการทำงานและการใช้ภาษาในชีวิตประจำวัน อย่างไรก็ตาม แบบทดสอบเหล่านี้มักมีราคาสูง และเนื้อหาของแบบทดสอบก็ไม่ได้ออกแบบมาให้สอดคล้องกลมกลืนกับบริบทของสังคมและวัฒนธรรมไทย หรือกล่าวอีกอย่างหนึ่งคือ ตัวข้อสอบ (test item) ของแบบทดสอบต่างชาติที่ได้รับความนิยมอาจมีเนื้อหาที่ไม่ครอบคลุมบริบทของไทยและเป็นสิ่งที่ผู้สอบควรรู้ สอดคล้องกับ Dimova et al. (2020) ที่เล็งเห็นความสำคัญของแบบทดสอบที่ควรออกแบบและพัฒนาให้เหมาะสมกับบริบทเฉพาะของตน (Specific local context) เพื่อให้มีโครงสร้าง (Construct) เหมาะสมกับการใช้งานในบริบท

สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. มีพันธกิจหลักคือ การบริหารจัดการและดำเนินการเกี่ยวกับการศึกษา วิจัย พัฒนา และให้บริการทางการประเมินผลทางการศึกษาและทดสอบทางการศึกษาอย่างต่อเนื่อง รวมทั้งเป็นศูนย์กลางความร่วมมือด้านการทดสอบทางการศึกษาในระดับชาติและระดับนานาชาติ ได้ตระหนักถึงหน้าที่และบทบาทในการสร้างและพัฒนาแบบทดสอบระดับชาติเพื่อใช้วัดและประเมินความสามารถในการใช้ภาษาอังกฤษของผู้สอบโดยเน้นผู้สอบวัยทำงาน ทั้งนี้ สทศ. ได้เริ่มทำแบบทดสอบใหม่ขึ้นมาเพิ่มเติมจากแบบทดสอบระดับชาติที่ใช้ในขอบเขตของการจัดการศึกษาภาคบังคับและการศึกษาขั้นพื้นฐานที่ทำมาก่อนหน้านั้นแล้ว เช่น O-NET (Ordinary National Educational Test) โดยแบบทดสอบที่สร้างขึ้นใหม่ คือ Test of English for Communication in the Workplace (TEC-W) ที่มีเนื้อหาเกี่ยวข้องกับการใช้งานในที่ทำงานและในชีวิตประจำวันทั่วไป

วัตถุประสงค์

บทความนี้มีวัตถุประสงค์เพื่อบรรยายและอธิบายการสร้างและพัฒนาแบบทดสอบภาษาอังกฤษเพื่อการสื่อสารในการทำงาน Test of English for Communication-Workplace (TEC-W) พร้อมทั้งนำเสนอทิศทางการพัฒนาและใช้แบบทดสอบนี้ในระดับมหภาค (Large-scale) โดยหวังว่า แบบทดสอบนี้จะเป็นแบบทดสอบระดับชาติฉบับหนึ่งที่สามารถใช้วัดความสามารถในการใช้ภาษาอังกฤษของพลเมืองไทยในวัยทำงานได้อย่างเหมาะสมกับบริบทของประเทศไทยและได้รับการยอมรับในระดับนานาชาติ

บทความนี้ประกอบด้วยจุดเริ่มต้นของการสร้างแบบทดสอบ วิธีการสร้างและพัฒนา การแปลผลคะแนนเทียบกับ CEFR และบทสรุปตอนท้ายของผู้เขียนที่กล่าวถึงสรรพประโยชน์ของแบบทดสอบและทิศทางการพัฒนาแบบทดสอบนี้

จุดเริ่มต้นของการสร้าง Test of English for Communication-Workplace (TEC-W)

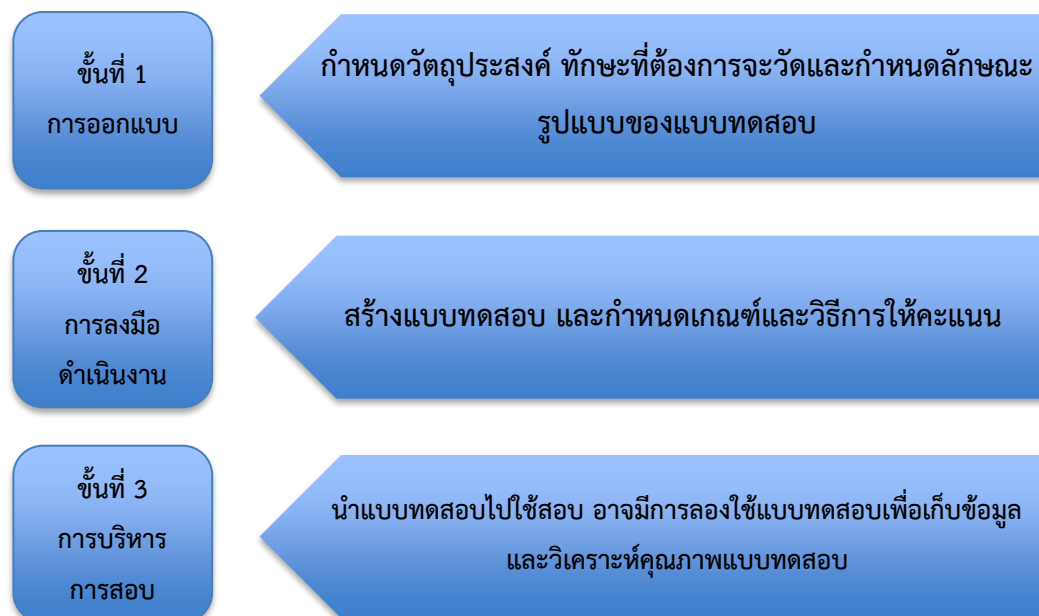
ช่วง พ.ศ. 2561 สทศ. ได้จัดประชุมหารือกับผู้แทนสภาอุตสาหกรรมแห่งประเทศไทย ผู้แทนสมาคมโรงแรมไทย ผู้แทนจากมหาวิทยาลัยและกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม เพื่อขอความเห็นเรื่องระดับความสามารถในการใช้ภาษาอังกฤษและการทดสอบภาษาอังกฤษของคนในวัยทำงานใน

ภาคอุตสาหกรรม การโรงแรม ซึ่งผู้แทนต่าง ๆ ได้ให้ความเห็นว่า ความสามารถในการใช้ภาษาอังกฤษเป็นดัชนีชี้ความสามารถในการแข่งขัน โดยทั่วไปพนักงานมักจะสอบข้อสอบของต่างชาติ เช่น TOEIC เพื่อสมัครเข้าทำงาน แต่หลังจากนั้นนายจ้างมีความประสงค์ให้พนักงานหลาย ๆ คนได้มีโอกาสทดสอบวัดความสามารถในการใช้ภาษาอังกฤษเป็นระยะ ๆ เพื่อพัฒนาตนเอง นำไปสู่การมอบหมายงานที่เหมาะสม และหากมีข้อสอบที่ใช้แทนข้อสอบของต่างชาติได้ก็จะเป็นการช่วยประหยัดเงินตราได้มาก ผู้แทนสมาคมโรงแรมไทยได้เสนอว่า อุตสาหกรรมโรงแรมและการท่องเที่ยวต้องการพนักงานที่มีความสามารถในการใช้ภาษาอังกฤษที่หลากหลาย เช่น ผู้จัดการต้องใช้ภาษาอังกฤษได้ดีกว่าพนักงานต้อนรับ พนักงานต้อนรับต้องใช้ภาษาอังกฤษได้ดีกว่าแม่บ้านทำความสะอาด ผู้แทนสภาอุตสาหกรรม กล่าวว่า พนักงานที่สอบภาษาอังกฤษได้ในระดับที่นายจ้างกำหนด จะมีเงินพิเศษเพิ่มมากขึ้นตามคะแนนที่ได้ ดังนั้นจากการประชุมหารือกับผู้แทนต่าง ๆ เห็นพ้องต้องกันว่าแบบทดสอบนี้น่าจะออกแบบมาให้ใช้วัดระดับความสามารถได้หลายระดับ (Multi-level test) นอกจากนี้ได้เสนอว่า เนื่องจากประเทศไทยมีนโยบายนำ Common European Framework of Reference (CEFR) มาใช้เป็นกรอบมาตรฐานในการเรียนการสอนและการทดสอบ (สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน 2557, คณะกรรมการอุดมศึกษา 2559) เห็นควรนำ CEFR มาเป็นแนวทางในการสร้างและพัฒนาแบบทดสอบ ตลอดจนการแปลผลคะแนนด้วย

หลังจากนั้น สทศ. ได้ตั้งคณะทำงานกำหนดกรอบการสร้างและพัฒนาเครื่องมือวัดความสามารถทางภาษาอังกฤษโดยมีผู้ทรงคุณวุฒิทางด้านภาษา และด้านการสร้างและพัฒนาเครื่องมือวัด และ สทศ. จำนวน 12 คน คณะทำงานนี้ได้ประชุมกันหลายครั้งเพื่อกำหนดองค์ประกอบเชิงโครงสร้างของแบบทดสอบ (Test Constructs) และการกำหนดรายละเอียดของแบบทดสอบ (Test Specification) รวมถึงการตั้งชื่อแบบทดสอบนี้ด้วย ในช่วงแรกคณะทำงานฯ ตั้งชื่อแบบทดสอบนี้ว่า Test of English Communication for Career: TEC-C แต่ภายหลังการอภิปรายในคณะทำงานฯ พบว่า ชื่อของแบบทดสอบเดิมไม่สอดคล้องกับวัตถุประสงค์ของโครงการที่ต้องการสร้างแบบทดสอบให้สอดคล้องกับพลวัตทางการศึกษา ดังนั้นคณะทำงานฯ จึงเสนอให้เปลี่ยนชื่อแบบทดสอบจาก Test of English Communication for Career: TEC-C เป็น Test of English Communication in the Workplace: TEC-W เพื่อให้สอดคล้องกับวัตถุประสงค์ของการทดสอบ บริบท และการนำผลคะแนนไปใช้ ซึ่งผู้ที่คิดค้นชื่อแบบทดสอบใหม่นี้คือ ศาสตราจารย์กิตติคุณ ดร. อัจฉรา วงศ์โสธร คณะทำงานฯ ยังได้วางแผนให้แบบทดสอบวัดได้ 4 ทักษะ คือ ทักษะการฟัง พูด อ่าน และเขียน และให้มีลักษณะเป็นการทดสอบในสภาพการณ์จริง (Actual Test) โดยเฉพาะในทักษะการพูด (พูดจริง) และทักษะการเขียน (เขียนจริง)

วิธีการสร้างและพัฒนา

Bachman and Palmer (1996) ได้เสนอขั้นตอนการสร้างแบบทดสอบไว้ 3 ขั้นตอน คือ



ภาพที่ 1 ขั้นตอนการสร้างแบบทดสอบ

ขั้นที่ 1 การออกแบบ (Design) เป็นขั้นตอนที่จะกำหนดวัตถุประสงค์ ทักษะที่ต้องการจะวัดและกำหนดลักษณะรูปแบบของแบบทดสอบ

วัตถุประสงค์ของ TEC-W คือการวัดความสามารถในการใช้ภาษาอังกฤษเพื่อการทำงานในบริบทของประเทศไทย โดยวัดทั้ง 4 ทักษะ คือ ฟัง พูด อ่าน และเขียน ซึ่งสามารถจัดทักษะที่กล่าวมาเป็นกลุ่มใหญ่ ๆ 2 กลุ่ม คือ ทักษะการรับสาร (ฟัง อ่าน) และทักษะการส่งสาร (พูด เขียน) โดยได้กำหนดลักษณะรูปแบบของแบบทดสอบทักษะการรับสารเป็นรูปแบบตัวเลือก 4 ตัว และกำหนดลักษณะรูปแบบของแบบทดสอบทักษะการส่งสารเป็นรูปแบบการพูดจริงเขียนจริงโดยกำหนดหัวข้อให้ผู้สอบได้พูดจริงและเขียนจริง นอกจากนี้ คณะทำงานฯ ได้ข้อสรุปว่า ระดับของภาษาที่คนในวัยทำงานสามารถนำไปใช้นั้นควรอยู่ในช่วงระหว่าง CEFR A2 ถึง CEFR B2 สรุปได้ว่า ความสามารถในระดับ A2 เป็นความสามารถในการใช้ภาษาในระดับประโยคเป็นเรื่องใกล้ตัว ความสามารถในระดับ B1 เป็นความสามารถในการใช้ภาษาในระดับกลาง ๆ ในชีวิตประจำวันทั่วไป ส่วนความสามารถในระดับ B2 เป็นความสามารถในการใช้ภาษาในระดับกลางค่อนข้างสูง สามารถใช้ภาษาในการสื่อสารได้อย่างมีประสิทธิภาพมากขึ้นทั้งในหัวข้อที่เป็นรูปธรรม และนามธรรมและสามารถใช้ในหน้าที่การงานได้ (Council of Europe, 2001)

คณะทำงานฯ เห็นว่า แบบทดสอบควรมีจำนวนข้อพอเหมาะไม่ยาวเกินไปหรือไม่สั้นเกินไป โดยคำนึงถึงหลักการหนึ่งของอรรถประโยชน์ของแบบทดสอบได้ด้านการนำไปปฏิบัติได้ (Practicality) ตามคำแนะนำของ Bachman and Palmer (1996) หมายรวมถึงเวลาการสอบ เวลาในการตรวจ

คณะกรรมการเห็นว่าแบบทดสอบจำนวน 100 ข้อ มีความเหมาะสมที่จะใช้จำแนกความสามารถของผู้สอบ ในช่วง CEFR A2, B1 และ B2 ได้พอเหมาะและใช้เวลาสอบ 2 ชั่วโมง (อย่างไรก็ตาม ด้วยข้อจำกัดเรื่อง การบริหารการสอบและการตรวจให้คะแนนในข้อสอบที่ใช้วัดทักษะการรับสาร ในรอบของการทดลองใช้จึง กำหนดให้แบบทดสอบทั้งชุดมีรูปแบบตัวเลือก 4 ตัวเลือกทั้งหมด)

ขั้นที่ 2 การลงมือดำเนินงาน (Operationalization) เป็นขั้นตอนการสร้างแบบทดสอบ และกำหนด เกณฑ์และวิธีการให้คะแนน

ในขั้นนี้ สทศ. ได้ตั้งคณะทำงานสร้าง พัฒนา และกลั่นกรองเครื่องมือวัดความสามารถทาง ภาษาอังกฤษ โดยได้เชิญผู้เชี่ยวชาญที่มีประสบการณ์ด้านการเรียนการสอนภาษาอังกฤษในระดับอุดมศึกษา และมีประสบการณ์การออกข้อสอบมาเป็นผู้ออกข้อสอบ 4 ด้าน คือ ด้านการฟัง จำนวน 13 คน ด้านการพูด จำนวน 13 คน ด้านการอ่าน จำนวน 16 คน และด้านการเขียน จำนวน 14 คน มีผู้กลั่นกรองข้อสอบ จำนวน 4 คน และผู้เชี่ยวชาญกลั่นกรองด้านภาษา ซึ่งเจ้าของภาษามาเป็นผู้ตรวจภาษาและทำหน้าที่พิสูจน์อักษร แบบทดสอบด้วย

ขั้นที่ 3 การบริหารการสอบ (Administration) เป็นขั้นตอนที่นำแบบทดสอบไปใช้สอบ อาจมีการ ลองใช้แบบทดสอบเพื่อเก็บข้อมูลและวิเคราะห์คุณภาพแบบทดสอบด้วย

ใน พ.ศ. 2562 มีการนำแบบทดสอบไปทดลองกับกลุ่มตัวอย่างที่หลากหลายและคาดว่าจะจะเป็นกลุ่มที่ กำลังจะเข้าสู่ตลาดแรงงาน เช่น กลุ่มนักเรียนระดับมัธยมศึกษาปีที่ 6 นักศึกษาระดับประกาศนียบัตรวิชาชีพ ปีที่ 3 นักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง ปีที่ 2 และนักศึกษาระดับปริญญาตรี (วิชาเอก ภาษาอังกฤษและวิชาอื่น ๆ) รวมจำนวนทั้งสิ้น 1,090 คน แล้วนำผลการทดสอบมาวิเคราะห์คุณภาพข้อสอบ รายข้อ (Item analysis) ร่องรอยการตอบของกลุ่มตัวอย่างนำมาวิเคราะห์ตัวข้อสอบรายข้อและภาพรวมเป็น ของแบบทดสอบทั้งฉบับ นำมาสู่การพัฒนาข้อสอบและเป็นแนวทางในการเขียนข้อสอบต่อไปและนอกจากนี้ สทศ. ยังได้กำหนดให้กลุ่มตัวอย่างเหล่านี้ได้สอบ TOEIC เพื่อนำผลคะแนนที่ได้มาหาความสัมพันธ์กับ TEC-W พบว่า คะแนนผลการทดสอบ TEC-W มีความสัมพันธ์กับคะแนน TOEIC ภาพรวมอยู่ในระดับสูง

ภายหลังเมื่อมีสถานการณ์การแพร่เชื้อโรคระบาดของเชื้อไวรัสโคโรนา การพัฒนาแบบทดสอบนี้จึง หยุดชะงักไปชั่วคราว ต่อมาใน พ.ศ. 2564 เมื่อสถานการณ์การแพร่ระบาดของเชือดังกล่าวดีขึ้น สทศ. ได้จัดให้ มีการทดลองใช้แบบทดสอบที่ปรับปรุงแล้วอีกครั้งหนึ่ง โดยทำการทดสอบกับกลุ่มตัวอย่างนิสิตและนักศึกษา ชั้นปีสุดท้ายของมหาวิทยาลัย 7 แห่ง เข้ารับการทดสอบด้วยความสมัครใจ จากมหาวิทยาลัยของรัฐ 5 แห่ง จำนวน 467 คน และมหาวิทยาลัยเอกชน 2 แห่ง จำนวน 91 คน รวมจำนวน 558 คน แบบทดสอบที่สร้าง แล้วควรได้รับการปรับปรุงและทดลองใช้อย่างน้อย 2 รอบ เป็นไปตามแนวทางการสร้างและพัฒนา แบบทดสอบมาตรฐานตามที่หน่วยงานการทดสอบต่าง ๆ แนะนำไว้

จิรดา วุฒยากร และคณะ (2565) รายงานว่า ผลการวิเคราะห์ความยากของข้อสอบ (Difficult) และอำนาจจำแนก (Discrimination Power) การฟัง การพูด การอ่าน และการเขียน จำนวนอย่างละ 25 ข้อ พบว่า ส่วนใหญ่เป็นข้อสอบที่มีคุณภาพดี คือ มีความยากระหว่าง 0.20–0.80 และมีค่าอำนาจจำแนกตั้งแต่

0.20 ขึ้นไป ทั้งนี้ในแต่ละทักษะก็ยังมีข้อสอบจำนวนไม่มากนักที่ยังมีคุณภาพด้านความยาก และ/หรือค่าอำนาจจำแนกไม่เป็นไปตามเกณฑ์และต้องปรับปรุงตัวลอง (ถือว่าเป็นสัดส่วนที่น้อยเมื่อเทียบกับจำนวนข้อสอบทั้งฉบับ ทั้งนี้ คณะทำงานสร้างข้อสอบได้ดำเนินการปรับปรุงแก้ไขข้อสอบดังกล่าวเพื่อจัดเก็บข้อสอบที่มีคุณภาพดีไว้ในคลังข้อสอบต่อไป ส่วนผลการวิเคราะห์คุณภาพความตรงตามสภาพ (Concurrent validity) ด้วยการหาค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson's correlation coefficient) ระหว่างแบบทดสอบ TEC-W กับแบบทดสอบ TOEIC นั้นพบว่า ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันด้านการอ่านเท่ากับ 0.887 ด้านการฟังเท่ากับ 0.903 และทั้งฉบับเท่ากับ 0.937 แสดงว่ามีค่าความตรง (Validity) ในระดับสูงถึงสูงมาก ส่วนการวิเคราะห์ค่าความเที่ยงด้วยค่าสัมประสิทธิ์แอลฟาของข้อสอบส่วนต่าง ๆ ได้แก่ การฟัง การพูด การอ่าน และการเขียน มีค่าเท่ากับ 0.875, 0.779, 0.855 และ 0.811 ตามลำดับ ในภาพรวมแบบทดสอบนี้มีค่าความเที่ยงระดับค่อนข้างสูง ส่วนผลการวิเคราะห์องค์ประกอบเชิงยืนยันของแบบทดสอบทั้ง 4 ทักษะ พบว่า โมเดลการวัดมีความสอดคล้องกับข้อมูลเชิงประจักษ์ทุกทักษะ โดยมีค่าดัชนี GFI, CFI, AGFI อยู่ระหว่าง 0.90-0.95 และดัชนี SRMR และ RMSEA ต่ำกว่า 0.05 ทุกทักษะ จึงยืนยันได้ว่าทฤษฎีการและข้อมูลเชิงประจักษ์ในการวัดทักษะทั้ง 4 ทักษะ มีความสอดคล้องในระดับดี สรุปว่า แบบทดสอบที่สร้างขึ้นมีความตรงกลมกลืนกับโมเดลโครงสร้างองค์ประกอบของการวัดความตรงตามภาวะสันนิฐาน (Construct Validity) นอกจากนี้ จิรดา วุฒยากร และคณะ (2565) ยังได้วิเคราะห์ความลำเอียงของข้อสอบที่มีต่อเพศชายและเพศหญิงโดยใช้วิธีการ Differentiation Item Functioning (DIF) พบว่า มีข้อสอบที่ผู้ตอบเพศชายและเพศหญิงมีความได้เปรียบเสียเปรียบกันอยู่จำนวน 9 ข้อ (คิดเป็นร้อยละ 9) โดยหัวข้อ (Topic) ที่เพศชายได้เปรียบ เช่น เรื่องที่เกี่ยวข้องกับกีฬา เทคโนโลยี การใช้เครื่องมือหรือเครื่องยนต์กลไกต่าง ๆ ส่วนหัวข้อที่เพศหญิงได้เปรียบ เช่น เรื่องเกี่ยวกับอาหารการกิน สิ่งแวดล้อม การเลือกซื้อสินค้า และการท่องเที่ยว เป็นต้น

ดังนั้นจึงสรุปได้ว่า สทศ. ได้สร้างและพัฒนาแบบทดสอบนี้ตามหลักการและทฤษฎี มีงานวิจัยในเรื่องที่เกี่ยวข้องกับการวิเคราะห์คุณภาพข้อสอบ การตรวจสอบความตรง ความเที่ยง การวิเคราะห์ความลำเอียงของแบบทดสอบ เพื่อเป็นการยืนยันว่าแบบทดสอบนี้มีมาตรฐานและมีความน่าเชื่อถือ นอกจากนี้สารสนเทศต่าง ๆ ที่ได้จากการตรวจสอบคุณภาพข้อสอบด้านความยากง่าย อำนาจจำแนก และวิเคราะห์ค่าความตรง ความเที่ยง และความลำเอียงของแบบทดสอบสามารถนำมาปรับปรุงข้อสอบแต่ละข้อเพื่อพัฒนาเป็นคลังข้อสอบและใช้เป็นแนวทางในการเขียนแบบทดสอบชุดต่อ ๆ ไปด้วย

การแปลผลคะแนน TEC-W กับ Common European Framework of Reference (CEFR)

สืบเนื่องจากนโยบายการใช้ CEFR (Council of Europe 2001) ในระบบการศึกษาไทยมาตั้งแต่ พ.ศ. 2557 (คณะกรรมการการศึกษาขั้นพื้นฐาน, 2557) ทำให้ประชาคมเริ่มรับรู้ถึงระดับความสามารถในการใช้ภาษาอังกฤษและคำบรรยายคุณลักษณะของ CEFR สทศ. จึงมีโครงการวิจัยสร้างคะแนนจุดตัด (Cut score)

ของแบบทดสอบนี้เทียบกับ CEFR ด้วยวิธีการตั้งเกณฑ์ผ่านหรือการสร้างมาตรฐานคะแนนจุดตัด (standard setting หรือ cut off score)

Cizek and Bunch (2006) กล่าวว่า Standard setting (หรือการสร้างมาตรฐานคะแนนจุดตัด) เป็นระเบียบวิธีวิจัยหนึ่งที่ช่วยให้ผู้พัฒนาแบบทดสอบได้สร้างคะแนนจุดตัดเทียบกับมาตรฐานหรือเกณฑ์การประเมินที่กำหนดขึ้น และได้รวบรวมเทคนิควิธีการไว้หลากหลายวิธี ในการแปลผลคะแนน TEC-W กับ CEFR นี้คือ ระเบียบวิธีวิจัยที่ให้ผู้เชี่ยวชาญจำนวน 10-15 คน ลงความเห็นว่าคุณสมบัติที่มีความสามารถต่ำสุด (Borderline candidate) ในระดับ CEFR ที่กำหนดขึ้นจะทำข้อสอบแต่ละข้อได้หรือไม่ ซึ่งการสร้างคะแนนจุดตัด TEC-W เทียบกับ CEFR นั้น ผู้เชี่ยวชาญจะต้องลงความเห็นว่าคุณสมบัติข้อที่ 1 ถึงข้อที่ 100 ผู้สอบที่มีความสามารถต่ำสุดในระดับ CEFR A2, B1 และ B2 จะทำได้หรือไม่ เรียกวิธีนี้ว่า Yes/No Ankoff ผู้เชี่ยวชาญในระดับ CEFR A2, B1 และ B2 (ระดับละ 3 รอบ) แล้วนำมาหาค่าเฉลี่ยว่า ผู้สอบต้องทำแบบทดสอบได้กี่ข้อ จึงจัดว่าเป็นผู้ที่มีความสามารถในการใช้ภาษาในระดับ CEFR A2, B1 และ B2 ความคิดเห็นที่ได้เป็นความคิดเห็นที่มีความเป็นอัตนัย (Subjective agreement) เป็นความคิดเห็นที่เกิดจากการพิจารณาข้อสอบรายข้อ โดยมีสารสนเทศเรื่องความยากง่ายรายข้อเป็นข้อมูลที่จะช่วยในการตัดสินใจ

การแปลผลคะแนน TEC-W กับ CEFR สรุปได้ดังตารางนี้

คะแนน TEC-W (เต็ม 100 คะแนน)	CEFR Level
85-100	B2
60-84	B1
< 60	A2

ผู้สอบที่ได้คะแนนน้อยกว่า 60 จัดเป็นผู้สอบที่มีความสามารถในระดับ CEFR A2 ผู้สอบที่ได้คะแนน 60-84 จัดเป็นผู้สอบที่มีความสามารถในระดับ CEFR B1 และผู้สอบที่ได้คะแนน 85-100 จัดเป็นผู้สอบที่มีความสามารถในระดับ CEFR B2

สาเหตุที่แบบทดสอบไม่ได้สร้างให้วัดและประเมินผลลงไปถึง CEFR A1 เพราะ CEFR ระดับนี้เป็นระดับต่ำสุด เป็นความสามารถในการใช้ภาษาในระดับวลี รูปแบบภาษาเป็นรูปแบบที่คุ้นเคยติดปาก ซึ่งนักเรียน นิสิต นักศึกษาที่กำลังเข้าสู่ตลาดแรงงานควรมีความสามารถเรื่องนี้มาก่อนแล้ว ทั้งนี้ กระทรวงศึกษาธิการได้กำหนดนโยบายว่า CEFR A1 ควรเป็นระดับที่นักเรียนชั้นประถมศึกษาปีที่ 6 สามารถใช้งานได้ ประกอบกับงานวิจัยของ Wudthayagorn et al. (manuscript for publication) พบว่า นักเรียนชั้นประถมศึกษาปีที่ 6 ส่วนใหญ่มีความสามารถในการใช้ภาษาอังกฤษในระดับ A1 แล้ว แบบทดสอบนี้จึงเริ่มวัดและประเมินที่ระดับ CEFR A2 และหากผู้สอบได้คะแนนต่ำกว่าคะแนนจุดตัดในระดับนี้ ก็ให้แปลผลคะแนนเป็น Below A2 (ซึ่งอาจจะเป็น A1 หรือ Pre-A1 ก็ได้)

อีกประการหนึ่ง สาเหตุที่แบบทดสอบนี้สร้างให้วัดและประเมินผลสูงสุดที่ประมาณ CEFR B2 เพราะ CEFR ระดับนี้เป็นระดับของผู้สอบที่ต้องการเข้าสู่ตลาดแรงงานสามารถแสดงศักยภาพในการใช้ภาษาอังกฤษเพื่อประกอบหน้าที่การงานได้ นอกจากนี้แบบทดสอบที่มีจำนวนเพียง 100 ข้อ ในรูปแบบ 4 ตัวเลือก อาจจะไม่เพียงพอต่อการสรุปอ้างอิงความสามารถในการใช้ภาษาอังกฤษได้ถึงระดับ CEFR C1 อันจะเป็นการประมาณค่าความสามารถของผู้สอบเกินจริง (overestimate delete) แต่ถ้าหากผู้สอบสามารถทำคะแนนมากกว่าคะแนนจุดตัดของ CEFR B2 อาจจะอนุมานได้ว่า ผู้สอบมีความสามารถในการใช้ภาษาสูงกว่าระดับ CEFR B2 (ซึ่งอาจจะเป็น C1 หรือ C2 ก็ได้)

อรรถประโยชน์ของ TEC-W

Bachman and Palmer (1996) กล่าวว่า แบบทดสอบที่มีอรรถประโยชน์สามารถนำไปใช้ได้ อย่างมั่นใจควรเป็นแบบทดสอบที่มีความตรง (Validity) ความเที่ยง (Reliability) วัดตรงตามสภาพจริง (Authenticity) นำไปปฏิบัติใช้ได้หรือบริหารการสอบได้ (Practicality) และผลกระทบของแบบทดสอบที่นำไปสู่การปรับการเรียนเปลี่ยนการสอน (Washback) ผู้สอบสามารถฝึกฝนการใช้ภาษาให้ทำแบบทดสอบได้ และการฝึกฝนนี้เป็นการสร้างศักยภาพในการใช้ภาษาด้วย

จากที่กล่าวมาในตอนต้น อาจกล่าวได้ว่า TEC-W มีอรรถประโยชน์เพราะคุณภาพในภาพรวมของข้อสอบรายข้ออยู่ในระดับดี มีความตรงคู่ขนาน (จากการวิเคราะห์ Concurrent validity) มีความตรงเชิงโครงสร้าง (จากการวิเคราะห์องค์ประกอบเชิงยืนยัน) มีค่าความเที่ยงที่ดี ส่วนเรื่องความตรงตามสภาพจริงนั้น อาจจะยังไม่สมบูรณ์เพราะมีแต่การฟังจริงและอ่านจริง ส่วนการเขียนและการพูดยังเป็นการวัดทางอ้อมด้วยรูปแบบตัวเลือก 4 ตัวเลือก (อย่างไรก็ตาม สทศ. มีโครงการที่จะพัฒนาเรื่องการทดสอบพูดจริงและเขียนจริงต่อไป) ส่วนเรื่องการนำ TEC-W ไปปฏิบัติใช้ได้หรือบริหารการสอบได้นั้น ทำได้ดีเพราะเป็นแบบทดสอบ 4 ตัวเลือกโดยใช้กระดาษคำตอบ จึงจัดสอบ ตรวจข้อสอบ และรายงานผลการสอบได้ง่าย ประการสุดท้ายคือผลกระทบของแบบทดสอบที่นำไปสู่การปรับการเรียนเปลี่ยนการสอนนั้น TEC-W ทำได้ดีในเรื่องการฟัง และการอ่าน หากต้องการให้ผู้สอบได้ฝึกฝนทักษะการพูดหรือทักษะการเขียนและให้ผู้สอนได้สอนทักษะนั้นโดยการปฏิบัติจริง TEC-W ต้องเปลี่ยนแบบทดสอบในส่วนการพูดและการเขียนเป็น Actual test (ซึ่งเป็นแผนการที่ได้ริเริ่มไว้ตั้งแต่ต้นแล้ว มีการทดลองใช้ไปบ้างแล้ว แต่ติดขัดเรื่องการฝึกผู้ตรวจทั้ง 2 ทักษะนี้)

กล่าวโดยสรุป TEC-W มีอรรถประโยชน์ต่อองค์ประกอบเรื่องการนำไปปฏิบัติใช้ได้หรือบริหารการสอบได้ และผลกระทบของแบบทดสอบที่นำไปสู่การปรับการเรียนเปลี่ยนการสอน (Washback) ยังไม่โดดเด่นเท่าองค์ประกอบอื่น แต่ถ้าสามารถพัฒนา TEC-W ให้สามารถวัดและประเมินการพูดจริงและเขียนจริงไม่ผ่านรูปแบบการเลือกตอบแบบ 4 ตัวเลือก ก็จะทำให้องค์ประกอบ 2 ตัวนี้มีความโดดเด่นเพิ่มมากขึ้น

ทิศทางของการทดสอบ TEC-W ในประเทศไทย

สทศ. ได้เริ่มงานสำคัญในการสร้างและพัฒนาแบบทดสอบ TEC-W ที่ใช้ในระดับมหภาค เหมาะแก่ผู้ที่ต้องการทราบความสามารถในการใช้ภาษาอังกฤษเพื่อการทำงาน เหมาะกับบริบทและพื้นที่ มีการวิเคราะห์วิจัยที่ทำให้เชื่อมั่นได้ว่าแบบทดสอบนี้มีมาตรฐานน่าเชื่อถือ เนื่องจากแบบทดสอบได้สร้างและพัฒนาโครงสร้างตามภาวะสันนิษฐานทางทฤษฎีที่สอดคล้องกลมกลืนกับบริบทของประเทศไทย จัดเป็นข้อสอบที่เหมาะสมกับพื้นที่ (Local test) มีการทดลองใช้ มีการตรวจสอบคุณภาพรายข้อ มีการวิเคราะห์ค่าความตรงแบบต่าง ๆ ค่าความเที่ยง ความลำเอียงของแบบทดสอบ มีการสร้างคะแนนจุดตัดและแปลผลเทียบเคียงกับ CEFR

ขั้นตอนต่อไปคือการพัฒนารูปแบบการสอบทักษะการพูดและทักษะการเขียนที่ผู้สอบจะได้พูดจริงและเขียนจริง พร้อมทั้งฝึกอบรมผู้ประเมินทักษะทั้ง 2 นี้ และในขณะเดียวกัน สทศ. จะได้เร่งทำคลังข้อสอบและพัฒนาระบบการสอบด้วยคอมพิวเตอร์ พร้อมทั้งโฆษณา ประชาสัมพันธ์ และเผยแพร่ให้แบบทดสอบนี้เป็นที่รู้จักยอมรับทั้งระดับชาติและนานาชาติ ผ่านการตีพิมพ์ นำเสนอผลงานวิจัยในวารสารวิชาการการประชุมวิชาการต่าง ๆ ทั้งในระดับประเทศ ภูมิภาค และนานาชาติต่อไป

ข้อเสนอแนะในทางปฏิบัติ

1. ข้อสอบที่มีคุณภาพไม่เป็นไปตามเกณฑ์ คณะทำงานสร้าง พัฒนา และกลั่นกรองเครื่องมือวัดความสามารถทางภาษาอังกฤษจะต้องดำเนินการปรับปรุงให้มีคุณภาพตามเกณฑ์ที่กำหนด และสามารถเก็บไว้ในคลังข้อสอบ
2. สทศ. จะได้เร่งดำเนินการทดสอบนำร่องข้อสอบที่วัดทักษะการพูดจริงและการเขียนจริงกับกลุ่มตัวอย่าง พร้อมทั้งฝึกอบรมผู้ตรวจข้อสอบพูดจริงและเขียนจริง

ข้อเสนอแนะเพื่อการวิจัย

1. ประเด็นวิจัยที่สามารถนำไปสู่การพัฒนาข้อสอบได้มีหลายประเด็น เช่น การศึกษาเรื่องความแตกต่างของผู้สอบ (Individual difference) เช่น ศาสนา (Religion) ถิ่นที่อยู่อาศัย (Geographical location) เพศ (Gender) ว่ามีผลต่ออคติในการตอบหรือไม่
2. ประเด็นวิจัยที่สามารถนำไปสู่การพัฒนาผู้เรียนได้ คือ การวิจัยเรื่อง Washback หรือผลกระทบของข้อสอบที่มีต่อผู้เรียนว่า ผู้เรียนสามารถนำองค์ความรู้จากการสอบไปฝึกฝน พัฒนาการเรียนภาษาอังกฤษ การใช้ภาษาอังกฤษของตนเองได้หรือไม่ อย่างไร

รายการอ้างอิง

ภาษาไทย

- สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน. (2557). *แนวปฏิบัติตามประกาศกระทรวงศึกษาธิการ เรื่อง นโยบายการปฏิรูปการเรียนการสอนภาษาอังกฤษ*. [เอกสารที่ไม่มีการตีพิมพ์]. กระทรวงศึกษาธิการ
- คณะกรรมการการอุดมศึกษา. (2559). *นโยบายการยกระดับมาตรฐานภาษาอังกฤษในสถาบันอุดมศึกษา*. [เอกสารที่ไม่มีการตีพิมพ์]. สำนักงานปลัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม.
- จิรดา วุฒพยากรและคณะ. (2566). *รายงานการวิจัยฉบับสมบูรณ์การวิเคราะห์คุณภาพแบบทดสอบ ภาษาอังกฤษเพื่อการสื่อสารในการทำงาน (Test of English for Communication-Workplace)*. สถาบันทดสอบการศึกษาแห่งชาติ (องค์การมหาชน)
- ชัยญนิษฐ์ อินทร. (2563). *ปัญหาและความจำเป็นด้านการพัฒนาทักษะการสื่อสารภาษาอังกฤษในการทำงาน ของพนักงานธนาคารพาณิชย์แห่งหนึ่งในเขตเมืองพัทยา*. มหาวิทยาลัยบูรพา.
- สถาบันสอนภาษา NOVA ONLINE. (2023). *ความสำคัญของการเรียนภาษาอังกฤษในยุคปัจจุบัน*. <https://novaonline-th.com/blog/importance-of-learning-english/#aioseo->

ภาษาอังกฤษ

- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford University Press.
- Cizek, G.J., & Bunch, M.B. (2006). *Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests*. Sage Publication.
- Council of Europe. (2001). *Common European Framework of Reference for Learning, teaching, assessment*. <https://assets.cambridge.org/052180/3136/sample/0521803136ws.pdf>
- Dimova, S., Yan, X. & Ginther, A. (2020). *Local Language Testing: Design, Implementation, and Development*. <https://doi.org/10.4324/9780429492242>
- Ufiera, A. (2016). *The Impact of English Language Skills on National Income: A Cross-National Comparison*. <https://www.fdic.gov/analysis/cfr/researchers/ufier/ufier-english-wp.pdf>