

การเทียบผลคะแนนแบบทดสอบสมิทธิภาพภาษาอังกฤษ
แห่งมหาวิทยาลัยศรีนครินทรวิโรฒกับกรอบมาตรฐานอ้างอิงสากลของยุโรป

Mapping Srinakharinwirot University - Standardized English Test
(SWU-SET) onto the Common European Framework of Reference (CEFR)

ชินจิตต์ อธิวรกุล¹ และจिरดา วุฒฒยากร^{2*}
Chuenjit Athiworakun¹ and Jirada Wudthayagorn^{2*}

¹ศูนย์ภาษาและบริการวิชาการ วิทยาลัยนานาชาติเพื่อศึกษาความยั่งยืน มหาวิทยาลัยศรีนครินทรวิโรฒ

²สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย

¹Language and Academic Services Centre, College of Sustainability Studies, Srinakharinwirot University

²Chulalongkorn University Language Institute, Chulalongkorn University

ABSTRACT

This study used standard setting to map SWU-SET scores onto the CEFR by employing Modified Angoff method. Fourteen experts judged each test item in the Srinakharinwirot University - Standardized English Test (SWU-SET) against the Common European Framework of Reference for Languages (CEFR) in order to recommend cut-off scores covering A2, B1, and B2 levels. The results showed that, based on 100 test items in the SWU-SET, test takers should achieve at A2 level if the scores range from 22-49, at B1 level from 50-77, and at B2 from 78 or above.

ARTICLE INFO

Article history:

Received 29 December 2017

Received in revised form

30 March 2018

Accepted 9 May 2018

Available online

24 December 2018

Keywords:

Mapping (การเทียบผลคะแนน),
SWU-SET (แบบทดสอบ
สมิทธิภาพภาษาอังกฤษแห่ง
มหาวิทยาลัยศรีนครินทรวิโรฒ),
CEFR (กรอบมาตรฐานอ้างอิง
สากลของยุโรป)

* ผู้เขียนที่ให้การติดต่อ

E-mail address: jirada.w@chula.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเทียบคะแนนของแบบทดสอบสมิทธิภาพภาษาอังกฤษแห่งมหาวิทยาลัยศรีนครินทรวิโรฒ (Srinakharinwirot University - Standardized English Test หรือ SWU-SET) กับกรอบมาตรฐานอ้างอิงสากลของยุโรป (Common European Framework of Reference for Languages หรือ CEFR) โดยใช้วิธี Modified Angoff method โดยให้ผู้เชี่ยวชาญ 14 คน ลงความเห็นว่ามีสัดส่วนคะแนนในแบบทดสอบ SWU-SET ได้กี่เปอร์เซ็นต์จึงจะอยู่ในระดับ A2, B1, และ B2 ผลปรากฏว่า จากจำนวนข้อสอบ 100 ข้อ ความสามารถระดับ A2 นั้นอยู่ระหว่าง 22-49 คะแนน ความสามารถระดับ B1 อยู่ระหว่าง 50-77 คะแนน และความสามารถระดับ B2 นั้นมีสัดส่วนทำได้ 78 คะแนนเป็นต้นไป

บทนำ

การพัฒนาความรู้ความสามารถในการใช้ภาษาอังกฤษของบัณฑิตเป็นพันธกิจหลักหนึ่งของ มหาวิทยาลัยศรีนครินทรวิโรฒ สิ่งหนึ่งที่มหาวิทยาลัยได้ให้ความสำคัญคือการพัฒนาแบบทดสอบ สมิทธิภาพภาษาอังกฤษเพื่อใช้วัดและประเมินสมิทธิภาพในการใช้ภาษาอังกฤษของนิสิต พันธกิจนี้ ได้ดำเนินการมาระยะหนึ่งก่อนที่สำนักงานคณะกรรมการการอุดมศึกษา (Office of the Higher Education Commission, 2016) จะประกาศนโยบายยกระดับมาตรฐานภาษาอังกฤษในระดับอุดมศึกษา ในปีการศึกษา 2559 โดยมีสาระสำคัญหลักให้ผู้เรียนมีความรู้ความสามารถใช้ภาษาอังกฤษเป็นเครื่องมือศึกษาค้นคว้า องค์ความรู้และก้าวทันโลก รองรับการเข้าสู่ประชาคมอาเซียน ตลอดจนการเพิ่มขีดความสามารถในการ แข่งขันของประเทศ นโยบายดังกล่าวมีประเด็นหลักย่อยที่สำคัญ 5 ประการ คือ

“1. ให้สถาบันอุดมศึกษากำหนดนโยบายและเป้าหมายการยกระดับมาตรฐานภาษาอังกฤษ ใน สถาบันอุดมศึกษาในทุกหลักสูตรและทุกระดับการศึกษา เพื่อเป็นแนวทางในการพัฒนาทักษะ ความสามารถการใช้ภาษาอังกฤษของนิสิตนักศึกษาให้เป็นบัณฑิตที่มีความพร้อมทั้งวิชาการ วิชาชีพ และ ทักษะการสื่อสารภาษาอังกฤษในระดับที่ใช้งานได้ (Working Knowledge)

2. ให้สถาบันอุดมศึกษาจัดทำแผนเพื่อดำเนินการให้เป็นไปตามนโยบายและเป้าหมายโดยมี ตัวชี้วัดและมีการประเมินผลที่ชัดเจน

3. ให้สถาบันอุดมศึกษาพิจารณาปรับปรุงการจัดการเรียนการสอนในรายวิชาภาษาอังกฤษ โดย มุ่งผลสัมฤทธิ์ตามเป้าหมายที่กำหนด

4. ให้สถาบันอุดมศึกษาพิจารณาการจัดกิจกรรมนอกหลักสูตร กระบวนการ สื่อ และ/หรือ สิ่งแวดล้อม ที่จะเปิดโอกาสและเสริมแรงจูงใจให้นิสิตนักศึกษาสามารถพัฒนาทักษะการใช้ภาษาอังกฤษ ได้ด้วยตนเอง

5. ให้สถาบันอุดมศึกษาพิจารณาจัดให้นิสิตนักศึกษาทุกคน ทดสอบความรู้ภาษาอังกฤษตาม แบบทดสอบมาตรฐานระดับอุดมศึกษาที่สถาบันสร้างขึ้น หรือที่เห็นสมควรจะนำมาใช้วัดสมิทธิภาพ ทางภาษาอังกฤษ (English Proficiency) โดยสามารถเทียบเคียงผลกับกรอบมาตรฐานอ้างอิงสากลของ ยุโรป (Common European Framework of Reference for Languages หรือ CEFR) หรือมาตรฐานอื่น เพื่อให้ทราบระดับความสามารถของนิสิตนักศึกษาแต่ละคน และสถาบันอุดมศึกษาอาจพิจารณานำ

ผลการทดสอบความรู้ทางภาษาอังกฤษบันทึกไว้ในใบรับรองผลการศึกษาหรือจัดทำเป็นประกาศนียบัตร โดยเริ่มตั้งแต่ปีการศึกษา 2559 เป็นต้นไป”

ศูนย์ภาษาและบริการวิชาการ วิทยาลัยนานาชาติเพื่อศึกษาความยั่งยืน ได้รับมอบหมายจาก มหาวิทยาลัยให้ดำเนินการพัฒนาแบบทดสอบสมรรถภาพภาษาอังกฤษแห่งมหาวิทยาลัยศรีนครินทรวิโรฒ (Srinakharinwirot University - Standardized English Test หรือ SWU-SET) เพื่อใช้วัดสมรรถภาพในการใช้ภาษาอังกฤษของนิสิต และเพื่อให้ผลคะแนนของการวัดด้วยแบบทดสอบนี้สามารถเทียบเคียงกับกรอบมาตรฐานฯ ได้ ผู้วิจัยจึงได้ดำเนินการศึกษาการเทียบผลคะแนนแบบทดสอบ SWU-SET กับกรอบมาตรฐานฯ ด้วยวิธีการสร้างมาตรฐาน (Standard Setting) เพื่อสร้างคะแนนจุดตัด (Cut-off Score) ที่ผู้เชี่ยวชาญเห็นว่าเหมาะสม

คำถามการวิจัย

1. เมื่อเทียบแบบทดสอบ SWU-SET กับกรอบมาตรฐานอ้างอิงสากลของยุโรป นิสิตที่มีความสามารถทางภาษาค่ำสุดในระดับ A2 จะสามารถทำแบบทดสอบฯ ได้กี่เปอร์เซ็นต์
2. เมื่อเทียบแบบทดสอบ SWU-SET กับกรอบมาตรฐานอ้างอิงสากลของยุโรป นิสิตที่มีความสามารถทางภาษาค่ำสุดในระดับ B1 จะสามารถทำแบบทดสอบฯ ได้กี่เปอร์เซ็นต์
3. เมื่อเทียบแบบทดสอบ SWU-SET กับกรอบมาตรฐานอ้างอิงสากลของยุโรป นิสิตที่มีความสามารถทางภาษาค่ำสุดในระดับ B2 จะสามารถทำแบบทดสอบฯ ได้กี่เปอร์เซ็นต์

ขอบเขตการวิจัย

การวิจัยนี้เทียบผลคะแนนของแบบทดสอบ SWU-SET กับกรอบมาตรฐานฯ ที่ระดับ A2, B1 และ B2 ด้วยเหตุผลเบื้องต้น 2 ประการคือ

1. แบบทดสอบอยู่ในระยะแรกของการพัฒนา และมุ่งเน้นการวัดทักษะรับสารและการใช้ภาษา ประกอบกับประสบการณ์ของคณาจารย์ที่ได้สอนนิสิตมาเป็นเวลานาน จึงได้ข้อสรุปว่าแบบทดสอบนี้น่าจะเทียบเคียงได้ตั้งแต่ระดับ A2, B1, และ B2
2. ระดับ A1 อาจจะเป็นระดับที่ต่ำเกินไปเพราะสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (Office of the Basic Education Commission, 2014) ได้กำหนดให้นักเรียนที่จบชั้นระดับมัธยมศึกษาปีที่ 6 ต้องมีความสามารถทางภาษาอังกฤษขั้นต่ำระดับ B1 การเริ่มต้นวัดตั้งแต่ระดับ A1 จึงเป็นการวัดที่ต่ำกว่าความคาดหวังของนโยบายของคณะกรรมการการศึกษาขั้นพื้นฐานจนเกินไป อีกทั้ง ตัวแบบทดสอบก็ไม่ได้ง่ายมากสำหรับระดับ A1 และที่ไม่ระบุว่าเป็น C1 หรือ C2 เพราะตัวข้อสอบไม่ได้วัดทั้งสี่ทักษะและไม่ได้วัดทักษะบูรณาการที่เหมาะสมกับการได้ระดับ C1 หรือ C2 จึงระบุผลสูงสุดที่ B2 อย่างไรก็ตาม ค่าคะแนนที่ได้อาจจะมีความหมายโดยนัยว่านิสิตอาจมีความสามารถที่ระดับ B2 หรือสูงกว่านั้น

ประโยชน์ที่คาดว่าจะได้รับ

งานวิจัยนี้จะทำให้การตีค่าคะแนนแบบทดสอบ SWU-SET มีความหมายว่า คะแนนที่นิสิตสอบได้หมายความว่านิสิตสามารถใช้ภาษาอังกฤษทำอะไรได้บ้าง โดยการเทียบผลคะแนนกับกรอบมาตรฐานฯ เป็นไปตามที่กระทรวงศึกษาธิการได้กำหนดไว้

นอกจากนี้ ยังเป็นการพัฒนาแบบทดสอบให้เป็นมาตรฐาน เพราะการหาคะแนนจุดตัดด้วยวิธีการสร้างมาตรฐาน (Standard Setting) เป็นกระบวนการสำคัญกระบวนการหนึ่งในการพัฒนาแบบทดสอบ
อนึ่ง การเทียบผลคะแนนแบบทดสอบนี้ซึ่งเป็นแบบทดสอบระดับอุดมศึกษากับกรอบมาตรฐานฯ จัดว่าเป็นการเทียบคะแนนแบบทดสอบชุดแรกของประเทศไทย ดังนั้น งานวิจัยนี้น่าจะเป็นต้นแบบสำคัญในการสร้างจุดตัดคะแนนให้กับแบบทดสอบชุดอื่น ๆ ต่อไป

วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

1. กรอบมาตรฐานอ้างอิงสากลของยุโรป (Common European Framework of Reference for Languages)

กรอบมาตรฐานอ้างอิงสากลของยุโรป (Common European Framework of Reference for Languages หรือ CEFR) จัดทำขึ้นโดยสหภาพยุโรป (Council of Europe, 2001) เพื่อสะท้อนความสามารถทางภาษาของผู้ใช้ภาษาด้วยมาตรฐานเดียวกัน ผู้ใช้ภาษาสามารถนำกรอบมาตรฐานดังกล่าวเป็นข้อมูลอ้างอิงในการตีความความสามารถของผู้ใช้ภาษาในระดับการใช้งานได้อย่างเข้าใจมากยิ่งขึ้น คำบรรยายคุณลักษณะ (Descriptors) ในภาพรวมของกรอบมาตรฐานอ้างอิงสากลของยุโรป (Global Scales) ดังตารางที่ 1

ตารางที่ 1 คำบรรยายคุณลักษณะในภาพรวมของกรอบมาตรฐานอ้างอิงสากลของยุโรป

ระดับกรอบมาตรฐาน ฯ	คำบรรยายคุณลักษณะในภาพรวมของกรอบมาตรฐานอ้างอิงสากลของยุโรป
A1	<ul style="list-style-type: none"> - สามารถเข้าใจและใช้วลีที่คุ้นเคยในชีวิตประจำวันและวลีพื้นฐานที่ตอบสนองความต้องการที่เป็นรูปธรรม - สามารถแนะนำตนเองและผู้อื่นโดยการถามและตอบเกี่ยวกับรายละเอียดส่วนตัว เช่น อาศัยอยู่ที่ไหน รู้จักใครบ้าง และมีสิ่งของใดบ้าง - สามารถมีปฏิสัมพันธ์แบบง่าย ๆ โดยผู้อื่นพูดช้า ๆ ชัด ๆ และพร้อมที่จะโต้ตอบเมื่อได้รับความช่วยเหลือจากคู่สนทนา
A2	<ul style="list-style-type: none"> - สามารถเข้าใจประโยคและวลีที่ซับซ้อนและเกี่ยวข้องกับเรื่องใกล้ตัว (เช่น เรื่องข้อมูลส่วนตัว และครอบครัว การซื้อของ ภูมิศาสตร์ท้องถิ่น การจ้างงาน) - สามารถสื่อสารอย่างง่ายในเรื่องที่เป็นปกติวิสัย โดยใช้การแลกเปลี่ยนข้อมูลง่าย ๆ ตรงไปตรงมา เกี่ยวกับเรื่องที่คุ้นเคย และที่ทำเป็นประจำ - สามารถบรรยายโดยใช้คำง่าย ๆ ที่เกี่ยวข้องกับภูมิหลังของตน เรื่องแวดล้อมใกล้ตัว และเรื่องที่เกี่ยวข้องกับความต้องการเร่งด่วน

ตารางที่ 1 คำบรรยายคุณลักษณะในภาพรวมของกรอบมาตรฐานอ้างอิงสากลของยุโรป (ต่อ)

ระดับกรอบมาตรฐาน ฯ	คำบรรยายคุณลักษณะในภาพรวมของกรอบมาตรฐานอ้างอิงสากลของยุโรป
B1	<ul style="list-style-type: none"> - สามารถเข้าใจประเด็นหลักจากสิ่งที่ได้ฟัง ได้อ่าน จากภาษามาตรฐานที่ชัดเจน เนื้อหาที่คุ้นเคยในการทำงานโรงเรียน การใช้เวลาว่าง และอื่น ๆ - สามารถรับมือกับสถานการณ์ต่าง ๆ ที่มักเกิดขึ้นระหว่างการเดินทางท่องเที่ยวในสถานที่ที่ผู้คนใช้ภาษาอังกฤษ - สามารถพูดหรือเขียนข้อความง่าย ๆ ติดต่อกันในหัวข้อที่คุ้นเคยและเป็นเรื่องที่ตนสนใจ - สามารถบรรยายประสบการณ์ เหตุการณ์ ความใฝ่ฝัน สามารถให้เหตุผลสั้น ๆ และอธิบายความคิดเห็นและแผนการได้
B2	<ul style="list-style-type: none"> - สามารถเข้าใจความสำคัญของข้อความที่ซับซ้อน ทั้งหัวข้อที่เป็นรูปธรรมและนามธรรม รวมทั้งสามารถอภิปรายเชิงเทคนิคในเรื่องที่ตนเชี่ยวชาญได้ - สามารถมีปฏิสัมพันธ์ที่คล่องแคล่วและเป็นธรรมชาติ ทำให้เกิดการปฏิสัมพันธ์ที่เป็นประจำสม่ำเสมอกับเจ้าของภาษานั้น ราบรื่น ไม่คิดขัด - สามารถพูดหรือเขียนข้อความที่ชัดเจนและมีรายละเอียดในหัวข้อที่หลากหลาย และสามารถอธิบายมุมมองในประเด็นหัวข้อที่กำหนดให้ รวมถึงทางเลือกของข้อดีข้อเสียที่หลากหลาย
C1	<ul style="list-style-type: none"> - สามารถเข้าใจข้อความที่ยาวขึ้นและลึกซึ้งขึ้น และสามารถรับรู้ความหมายโดยนัย - สามารถนำเสนอตนเองได้อย่างคล่องแคล่ว เป็นธรรมชาติ ไม่มีสะดุดหยุดหาคำศัพท์ วลี คำพูด - สามารถใช้ภาษาได้อย่างยืดหยุ่น มีประสิทธิภาพ เพื่อวัตถุประสงค์ทางสังคม วิชาการ และการประกอบอาชีพ - สามารถพูดหรือเขียนข้อความในหัวข้อที่ซับซ้อนได้ชัดเจน ถูกต้อง ตามโครงสร้างไวยากรณ์ และมีรายละเอียด สามารถพูดหรือเขียนด้วยรูปแบบต่าง ๆ ใช้คำสันธานเชื่อม
C2	<ul style="list-style-type: none"> - สามารถเข้าใจทุกสิ่งทุกอย่างที่ได้ยินได้ฟังอย่างมีประสิทธิภาพ - สามารถสรุปข้อมูลจากแหล่งข้อมูลการพูดและการเขียนที่หลากหลายและสร้างประเด็นโต้แย้งขึ้นมาใหม่และนำเสนอได้อย่างต่อเนื่อง - สามารถนำเสนอตนเองได้อย่างเป็นธรรมชาติ คล่องแคล่ว แม่นยำ และแยกแยะรายละเอียดของความสามารถโดยเฉพาะอย่างยิ่งในสถานการณ์ที่ซับซ้อน

หมายเหตุ: แปลโดย Wudthayagorn (2016)

อย่างไรก็ตาม กรอบมาตรฐานฯ สร้างขึ้นเพื่อเป็นแนวทางในการเรียน การสอน และการทดสอบ นอกจาก คำบรรยายคุณลักษณะในภาพรวมแล้วยังมีคำบรรยายปลีกย่อยสำหรับทักษะการฟัง การพูด การอ่าน และการเขียนด้วย (ดูรายละเอียดเพิ่มเติมได้ที่ Council of Europe (2001, 2003))

2. การสร้างมาตรฐานเพื่อหาคะแนนจุดตัด

การเทียบผลคะแนนแบบทดสอบกับกรอบมาตรฐานต่าง ๆ ทำได้ด้วยวิธีการสร้างมาตรฐานเพื่อหาคะแนนจุดตัด ซึ่งมีเทคนิคต่าง ๆ มากกว่า 60 วิธี (Cizek, 2012) เช่น Angoff, Bookmark และ Body of

Work ในแต่ละวิธีจะมีขอบเขตในการใช้งานต่างกัน แบบทดสอบที่เป็นตัวเลือกมักจะนิยมใช้วิธีการ Angoff และ Bookmark ความแตกต่างระหว่างสองวิธีการนี้ คือ วิธีการ Angoff อาจไม่จำเป็นต้องมีค่าคุณภาพข้อสอบรายชื่อ ส่วนวิธีการ Bookmark จำเป็นต้องมีค่าคุณภาพข้อสอบรายชื่อ แบบทดสอบที่วัดทักษะการส่งสาร (การพูดและการเขียน) มักจะนิยมใช้วิธีการ Body of Work วิธีการดังกล่าวอาศัยความคิดเห็นของผู้เชี่ยวชาญจำนวน 10-15 คน ในการลงความเห็นร่วมกันว่าผู้สอบในระดับที่ระบุจะสามารถทำข้อสอบข้อนั้น ๆ ได้หรือไม่และได้จำนวนกี่ข้อ ผู้เชี่ยวชาญมักจะลงความเห็น 2 ถึง 3 รอบ เพื่อให้ได้มติที่เป็นเอกฉันท์ การเลือกวิธีการต่าง ๆ ขึ้นอยู่กับลักษณะของแบบทดสอบ จำนวนข้อสอบ ตลอดจนผลการวิเคราะห์คุณภาพรายชื่อ ดังนั้น การเลือกวิธีการต่าง ๆ ให้เหมาะสมนั้น ผู้วิจัยต้องทำการศึกษาให้รอบคอบและเหมาะสมกับลักษณะของแบบทดสอบ

การสร้างมาตรฐานเพื่อหาคะแนนจุดตัดใช้ความคิดเห็นของผู้เชี่ยวชาญเป็นหลัก ผู้เชี่ยวชาญต้องมีความรู้ ความชำนาญ ประสบการณ์ที่เกี่ยวข้อง เช่น การเรียนการสอนภาษาอังกฤษ การออกแบบการทดสอบ คำบรรยายมาตรฐานคุณลักษณะในกรอบมาตรฐานอ้างอิงสากลของยุโรป เพื่อลงความเห็นในเรื่องการกำหนดคะแนนจุดตัด ซึ่งความคิดเห็นของผู้เชี่ยวชาญเป็นความคิดเห็นเชิงอัตนัย (Subjective Agreement) และมีผู้ดำเนินการในการสร้างคะแนนมาตรฐานอำนวยความสะดวกให้การลงความเห็นของผู้เชี่ยวชาญเป็นไปด้วยดี

3. การเทียบคะแนนข้อสอบมาตรฐานกับกรอบมาตรฐานอ้างอิงสากลของยุโรป

กรอบมาตรฐานฯ และการสร้างมาตรฐานเพื่อหาคะแนนจุดตัดที่กล่าวมาข้างต้นสามารถเชื่อมโยงกันได้ ผลของการเชื่อมโยงนั้นจะทำให้คะแนนมีความหมายว่าผู้สอบสามารถใช้ภาษาทำอะไรได้บ้าง ดังนั้น นักวิจัยหลายท่านจึงทำการเทียบคะแนนข้อสอบมาตรฐานต่าง ๆ กับกรอบมาตรฐานอ้างอิงสากลของยุโรป ดังนี้

แบบทดสอบ TOEFL ITP (Institute Test Preparation) กับ กรอบมาตรฐานอ้างอิงสากลของยุโรป

Tannenbaum และ Baron (2011) ได้ทำการเทียบผลคะแนนแบบทดสอบ TOEFL ITP กับ กรอบมาตรฐานฯ ด้วยวิธีการ Modified Angoff โดยให้ผู้เชี่ยวชาญจำนวน 16 คน จากชาติต่าง ๆ 14 ชาติ เป็นเพศหญิง 11 คน เพศชาย 7 คน ที่มีประสบการณ์ในการสอนภาษาอังกฤษ คู่กันเคยกับผู้เรียนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ คู่กันเคยกับแบบทดสอบ TOEFL และกรอบมาตรฐานฯ แล้วทำการพิจารณาข้อสอบทีละข้อเทียบกับระดับความสามารถทางภาษาตามกรอบมาตรฐานฯ ในระดับ A2, B1, และ B2 เป็นจำนวนทั้งสิ้น 3 รอบ

วิธีการ Modified Angoff คือการพิจารณาว่าผู้เรียนที่มีความสามารถต่ำสุดในแต่ละระดับมีโอกาที่จะสามารถทำข้อสอบข้อนั้น ๆ ได้กี่เปอร์เซ็นต์ เช่น

- ผู้เรียนที่มีความสามารถต่ำสุดในระดับ A2 มีโอกาที่จะทำข้อสอบข้อที่ i ได้ถูกต้องกี่เปอร์เซ็นต์
- ผู้เรียนที่มีความสามารถต่ำสุดในระดับ B1 มีโอกาที่จะทำข้อสอบข้อที่ i ได้ถูกต้องกี่เปอร์เซ็นต์
- ผู้เรียนที่มีความสามารถต่ำสุดในระดับ B2 มีโอกาที่จะทำข้อสอบข้อที่ i ได้ถูกต้องกี่เปอร์เซ็นต์

หลังจากนั้น นำคะแนนที่ได้แต่ละรอบมาหาค่าเฉลี่ย ส่วนการหาคะแนนจุดตัดในระดับ C1 Tannenbaum และ Baron (2011) ใช้วิธีประมาณการจากการเปรียบเทียบคะแนนกับข้อสอบ TOEFL PBT (Paper-based Test) และ TOEFL CBT (Computer-based Test) ผลการวิจัยพบว่า ผู้ที่สอบได้คะแนนในระดับต่าง ๆ สามารถเทียบเคียงกับ กรอบมาตรฐานฯ ได้ดังนี้

ตารางที่ 2 ผลการเทียบคะแนน TOEFL ITP กับกรอบมาตรฐานอ้างอิงสากลของยุโรป

การฟัง	โครงสร้างและการเขียน	การอ่าน	คะแนนรวม	กรอบมาตรฐาน ฯ
64	64	63	627	C1
54	53	56	543	B2
47	43	48	460	B1
38	32	31	337	A2

จากการเทียบคะแนน TOEFL ITP กับกรอบมาตรฐานฯ พบว่าผู้ที่สอบที่ได้คะแนน TOEFL ITP 337 คะแนน จะมีความสามารถอยู่ในระดับ A2 หรือทำข้อสอบการฟัง โครงสร้างและการเขียน และข้อสอบอ่านได้ 38, 32, และ 31 ข้อตามลำดับ ผู้สอบที่ได้คะแนน TOEFL ITP 460 คะแนน จะมีความสามารถอยู่ในระดับ B1 หรือทำข้อสอบการฟัง โครงสร้างและการเขียน และข้อสอบอ่านได้ 47, 43, และ 48 ข้อตามลำดับ ผู้สอบที่ได้คะแนน TOEFL ITP 543 คะแนน จะมีความสามารถอยู่ในระดับ B2 หรือทำข้อสอบการฟัง โครงสร้างและการเขียน และข้อสอบอ่านได้ 54, 53, และ 56 ข้อตามลำดับ ผู้สอบที่ได้คะแนน TOEFL ITP 627 คะแนน จะมีความสามารถอยู่ในระดับ C1 หรือทำข้อสอบการฟัง โครงสร้างและการเขียน และข้อสอบอ่านได้ 64, 64, และ 63 ข้อตามลำดับ

แบบทดสอบ TOEFL iBT กับ กรอบมาตรฐานอ้างอิงสากลของยุโรป

Tannenbaum และ Wylie (2008) ได้ดำเนินการเทียบผลคะแนนข้อสอบ TOEFL iBT กับ กรอบมาตรฐานฯ ด้วยวิธี Modified Angoff กับข้อสอบปรนัย (การฟังและการอ่าน) และวิธี Constructed-Response กับข้อสอบอัตนัย (การเขียนและการพูด) โดยมีผู้เชี่ยวชาญ 23 คน จากประเทศต่าง ๆ 16 ประเทศ เป็นเพศหญิง 13 คน เพศชาย 10 คน พิจารณาข้อสอบเพื่อเทียบกับกรอบมาตรฐานฯ ตั้งแต่ระดับ A1 ถึง C2 ผลการวิจัยพบว่า ผู้ที่สอบได้คะแนนในระดับต่าง ๆ สามารถเทียบเคียงกับกรอบมาตรฐานฯ ได้ดังนี้

ตารางที่ 3 ผลการเทียบคะแนน TOEFL iBT กับกรอบมาตรฐานอ้างอิงสากลของยุโรป

การฟัง	การพูด	การอ่าน	การเขียน	คะแนนรวม	กรอบมาตรฐาน ๑
-	-	29	-	-	C2
26	28	28	28	110	C1
21	23	22	21	87	B2
13	19	8	17	57	B1
-	13	-	11	24	A2
-	8	-	-	8	A1

จากตารางจะเห็นได้ว่าผู้ที่มีความสามารถระดับ C1 ต้องได้คะแนน 110 คะแนน (จากคะแนนเต็ม 120 คะแนน) ส่วนผู้มีความสามารถระดับ B2, B1, A2, A1 ต้องได้คะแนน 87, 57, 24, และ 8 คะแนน ตามลำดับ ส่วนระดับ C2 นั้น ไม่ได้ระบุคะแนนรวมเพราะมีเฉพาะข้อสอบการอ่านเท่านั้นที่สามารถวัดความสามารถในการอ่านได้ ถ้าผู้เข้าสอบทำคะแนนการอ่านได้ 29 คะแนน (จากคะแนนเต็ม 30 คะแนน) จัดว่าผู้เข้าสอบคนนั้นมีความสามารถในการอ่านในระดับ C2 ส่วนข้อสอบส่วนอื่น ๆ ได้แก่ ข้อสอบ การฟัง การเขียน การพูด ไม่สามารถวัดได้ถึงในระดับนี้

ดังนั้นในการเทียบผลคะแนนแบบทดสอบกับกรอบมาตรฐาน๑ นักวิจัยต้องพิจารณาให้ถ่วงถ้เกี่ยวกับระดับที่ควรจะต้องเทียบกับเกณฑ์มาตรฐาน เทคนิควิธีที่จะใช้ ตลอดจนการคัดเลือกผู้เชี่ยวชาญที่จะให้ความคิดเห็นเรื่องคะแนนจุดตัด

แบบทดสอบ SWU-SET (Srinakharinwirot University - Standardized English Test) และวิธีการสร้างมาตรฐาน

1. SWU-SET

แบบทดสอบ SWU-SET (Srinakharinwirot University - Standardized English Test) สร้างขึ้นโดยคณาจารย์ศูนย์ภาษาและบริการวิชาการ โดยมี Test Specification ที่ครอบคลุมบริบทสำคัญที่นักศึกษาต้องใช้ภาษาอังกฤษ 3 บริบท ได้แก่ บริบททั่วไป (General Setting) บริบทชีวิตมหาวิทยาลัย (Campus Life) และบริบทวิชาการ (Academic Setting)

แบบทดสอบแบ่งออกเป็น 5 ตอน ตอนละ 20 ข้อ รวมจำนวนทั้งหมด 100 ข้อ ประกอบด้วย การฟัง (Listening) คำศัพท์ (Vocabulary) การใช้ภาษา (Usage and Functional Language) โครงสร้างภาษา (Structure) และการอ่าน (Reading)

ทีมงานวิจัยศูนย์ภาษาและบริการวิชาการได้กำหนดวางระดับช่วงชั้นของแบบทดสอบไว้ที่ A2, B1, B2 ซึ่งการกำหนดช่วงชั้นนี้เกิดจากประสบการณ์สอนของคณาจารย์ศูนย์ภาษาและบริการวิชาการ จำนวน 14 คน ที่สอนนิสิตมหาวิทยาลัยศรีนครินทรวิโรฒทุกชั้นปีจากทุกคณะและคณาจารย์๑ มีความเห็นพ้องต้องกันว่าช่วงชั้นความสามารถของนิสิตอยู่ระหว่าง A2 ถึง B2 มีวิธีการดังนี้

1.1 ทีมงานวิจัยได้ทดลองแบบทดสอบ SWU-SET กับนิสิตมหาวิทยาลัยศรีนครินทรวิโรฒ จำนวน 350 คน จากคณะทันตแพทย์ สังคมศาสตร์ และศิลปกรรมศาสตร์ (แบ่งตามระดับความสามารถทางภาษาอังกฤษที่ได้จากคะแนนสอบ O-NET วิชาภาษาอังกฤษที่จัดโดยส่วนกลาง) และนำมาวิเคราะห์คุณภาพแบบทดสอบ หลังจากนั้น ประชุมทีมงานวิจัยเพื่อปรับปรุงข้อสอบที่มีปัญหาเรื่องค่าความยาก และค่าอำนาจจำแนก แบบทดสอบ

1.2 หลังจากที่ได้แบบทดสอบที่ปรับปรุงแล้ว จึงดำเนินการเทียบคะแนนแบบทดสอบ SWU-SET กับกรอบมาตรฐานอ้างอิงสากลของยุโรป

2. ข้อตกลงเบื้องต้น

แบบทดสอบชุดที่ใช้ทำการเทียบคะแนนกับกรอบมาตรฐานฯ เป็นแบบทดสอบที่สร้างขึ้นตาม Test Specification และมีทีมผู้สร้างแบบทดสอบช่วยกันตรวจสอบว่าเป็นไปตาม Test Specification มากน้อยเพียงใด แบบทดสอบทุกชุดจะได้รับการวิเคราะห์ค่าคุณภาพรายข้อหลังจากผ่านการใช้งานแล้ว ซึ่งเป็นกระบวนการหนึ่งของการพัฒนาแบบทดสอบมาตรฐาน แบบทดสอบที่สร้างขึ้นในฉบับต่อมาก็ผ่านกระบวนการสร้างและควบคุมคุณภาพเช่นเดียวกับแบบทดสอบฉบับนี้ ดังนั้น จึงอนุมานได้ว่าแบบทดสอบทุกชุดไม่มีความแตกต่างกัน

3. วิธีการสร้างมาตรฐาน

3.1 ผู้เชี่ยวชาญ (Panelist)

ผู้เชี่ยวชาญในการสร้างมาตรฐานประกอบด้วย คณาจารย์ประจำของศูนย์ภาษาและบริการวิชาการ วิทยาลัยนานาชาติเพื่อศึกษาความยั่งยืน มหาวิทยาลัยศรีนครินทรวิโรฒ ผู้ที่มีประสบการณ์การสอนและการออกแบบทดสอบภาษาอังกฤษในระดับอุดมศึกษาตั้งแต่ 5 ปี ขึ้นไป รวมทั้งสิ้น 14 คน เป็นเพศหญิง 10 คน และเพศชาย 4 คน ผู้เชี่ยวชาญมีภูมิหลังทางการศึกษาในระดับปริญญาโทและปริญญาเอกในทิศทางเดียวกัน ได้แก่ การสอนภาษาอังกฤษ ภาษาศาสตร์ และการสื่อสารระหว่างวัฒนธรรม ปัจจุบัน ผู้เชี่ยวชาญทุกคนรับผิดชอบสอนวิชาเดียวกัน คือ วิชาภาษาอังกฤษพื้นฐาน ให้นิสิตทุกคณะระหว่างชั้นปีที่ 1 และ 2 ของมหาวิทยาลัยศรีนครินทรวิโรฒ

3.2 ขั้นเตรียมความพร้อม (Orientation)

ผู้ดำเนินการฯ (นักวิจัยคนที่สอง) เปิดโอกาสให้ผู้เชี่ยวชาญแสดงความคิดเห็น แลกเปลี่ยนประสบการณ์ในการพัฒนาแบบทดสอบ การทดลองใช้แบบทดสอบ และการปรับปรุงแบบทดสอบ โดยใช้แบบทดสอบ SWU-SET ด้วย

3.3 ขั้นการสร้างความชัดเจนเกี่ยวกับวิธีการและขั้นตอนการเทียบคะแนน (Training Session) ในการเทียบคะแนนแบบทดสอบกับกรอบมาตรฐานอ้างอิงสากลของยุโรป ผู้ดำเนินการฯ ใช้วิธีการ Modified Angoff เพราะเหมาะสมกับแบบทดสอบที่เป็นแบบตัวเลือก (Multiple Choice) และกำหนดจุดตัดเพียง 3 จุด คือ A2, B1, และ B2 (ถ้าหากเป็นแบบทดสอบลักษณะอื่นและมีจุดตัดมากกว่านี้วิธีการนี้อาจจะไม่เหมาะสม)

ผู้ดำเนินการฯ เริ่มอธิบายคำบรรยายคุณลักษณะในระดับต่าง ๆ ของกรอบมาตรฐานฯ เพื่อให้ผู้เชี่ยวชาญเห็นภาพรวมความสามารถของนิสิต และเปิดโอกาสให้มีการอภิปรายแลกเปลี่ยนประสบการณ์การสอนของคณาจารย์ฯ ที่เคยสอนนิสิตคณะต่าง ๆ ว่ามีความสามารถในการใช้ภาษาอังกฤษแตกต่างกัน ทั้งนี้เพื่อให้เข้าใจธรรมชาติและความสามารถของนิสิตและมองเห็นภาพของผู้ที่มีความสามารถต่ำสุดของระดับนั้น (Just Qualified Candidate) ตรงกัน คำถามสำคัญในการดำเนินงานเพื่อเทียบคะแนนแบบทดสอบด้วยวิธีการ Modified Angoff กับระดับ A2, B1 และ B2 คือ

- นิสิตที่มีความสามารถต่ำสุดในระดับ A2 มีโอกาสที่จะทำข้อสอบข้อที่ i ได้ถูกต้องกี่เปอร์เซ็นต์
- นิสิตที่มีความสามารถต่ำสุดในระดับ B1 มีโอกาสที่จะทำข้อสอบข้อที่ i ได้ถูกต้องกี่เปอร์เซ็นต์
- นิสิตที่มีความสามารถต่ำสุดในระดับ B2 มีโอกาสที่จะทำข้อสอบข้อที่ i ได้ถูกต้องกี่เปอร์เซ็นต์

ผู้เชี่ยวชาญระบุเปอร์เซ็นต์ที่นิสิตที่มีความสามารถต่ำสุดในระดับนั้น ๆ จะตอบถูกเพิ่มขึ้นทีละ 5 เปอร์เซ็นต์ จาก 0, 5, 10, 15, 20, 25, ไปเรื่อยๆ จนถึง 100 เปอร์เซ็นต์ พร้อมทั้งบันทึกลงในแบบฟอร์มที่กำหนดให้

3.4 ขั้นวิธีการสร้างมาตรฐานเพื่อหาคะแนนจุดตัด รอบที่ 1

3.4.1 การเทียบคะแนนระดับ A2

ผู้ดำเนินการฯ เริ่มต้นให้ผู้เชี่ยวชาญ พิจารณาว่า “นิสิตที่มีความสามารถต่ำสุดในระดับ A2 มีโอกาสที่จะทำข้อสอบข้อที่ i ได้ถูกต้องกี่เปอร์เซ็นต์” โดยพิจารณาข้อสอบทีละข้อ เริ่มจากข้อที่ 1 ไปจนถึงข้อที่ 100 เช่น ผู้เชี่ยวชาญคนที่ i พิจารณาแบบทดสอบในระดับ A2 จะสามารถทำข้อนี้ถูกกี่เปอร์เซ็นต์ โดยบันทึกค่าเปอร์เซ็นต์ลงในแบบฟอร์ม ดังต่อไปนี้

ตารางที่ 4 เปอร์เซ็นต์ที่จะตอบถูกของนิสิตที่มีความสามารถต่ำสุดระดับ A2

ข้อที่	เปอร์เซ็นต์ที่จะตอบถูก
1	
2	
3	
.	
.	
.	
98	
99	
100	

ยกตัวอย่างเช่น

“Her illness was a **valid** reason for her absence yesterday. I would not penalize her.”

A. mysterious B. justified C. powerful D. sufficient

ผู้เชี่ยวชาญทั้ง 14 คน ต่างคนต่างพิจารณาข้อสอบข้อดังกล่าวว่า นิสิตที่มีความสามารถต่ำสุดในระดับ A2 มีโอกาสที่จะทำข้อสอบข้อนี้ได้ถูกต้องกี่เปอร์เซ็นต์ ผู้เชี่ยวชาญทำการพิจารณาเช่นนี้กับข้อสอบทุกข้อจากข้อที่ 1 ถึง 100 เป็นจำนวนเปอร์เซ็นต์ต่าง ๆ กัน ผู้เชี่ยวชาญบางคนอาจคิดว่า นิสิตที่มีความสามารถต่ำสุดในระดับ A2 มีโอกาสตอบข้อนี้ถูก 5 เปอร์เซ็นต์ อีกคนอาจจะคิดว่า 10 เปอร์เซ็นต์ เป็นต้น ในการพิจารณารายข้อ ผู้เชี่ยวชาญจะบันทึกเปอร์เซ็นต์ในแบบฟอร์มที่กำหนดให้

3.4.2 การเทียบคะแนนระดับ B2

ต่อจากนั้น ผู้ดำเนินการฯ ได้เข้ามาเทียบคะแนนระดับ B2 เพื่อให้ผู้เชี่ยวชาญได้เห็นภาพความสามารถภาษาอังกฤษที่แตกต่างชัดเจนจาก A2 และ B1 โดยในระดับ B2 นี้มีคำถามสำคัญว่า นิสิตที่มีความสามารถต่ำสุดในระดับ B2 มีโอกาสที่จะทำข้อสอบข้อที่ i ได้ถูกต้องกี่เปอร์เซ็นต์ โดยพิจารณาข้อสอบทีละข้อ เริ่มจากข้อที่ 1 ไปจนถึงข้อที่ 100

ยกตัวอย่างเช่น ในข้อสอบข้อเดียวกันกับที่กล่าวมาข้างต้น

“Her illness was a **valid** reason for her absence yesterday. I would not penalize her.”

A. mysterious B. justified C. powerful D. sufficient

ผู้เชี่ยวชาญจะลงความเห็นว่าคุณมีความสามารถต่ำสุดในระดับ B2 จะสามารถทำข้อนี้ถูกกี่เปอร์เซ็นต์ โดยบันทึกค่าเปอร์เซ็นต์ลงในแบบฟอร์ม ดังต่อไปนี้

ตารางที่ 5 เปอร์เซ็นต์ที่จะตอบถูกของนิสิตที่มีความสามารถต่ำสุดระดับ B2

ข้อที่	เปอร์เซ็นต์ที่จะตอบถูก
1	
2	
3	
.	
.	
.	
98	
99	
100	

3.4.3 การเทียบคะแนนระดับ B1

ผู้ดำเนินการฯ ได้กลับมาเทียบคะแนนระดับ B1 โดยให้คำแนะนำกับผู้เชี่ยวชาญ ว่าระดับคะแนนของ B1 ควรจะเป็นคะแนนที่อยู่ระหว่างระดับ A2 และ B2 แต่ทั้งนี้ยังคงให้ยึดคำถามสำคัญว่า นิสิตที่มีความสามารถต่ำสุดในระดับ A2 มีโอกาสที่จะทำข้อสอบข้อที่ i ได้ถูกต้องกี่เปอร์เซ็นต์ โดยพิจารณาข้อสอบทีละข้อ เริ่มจากข้อที่ 1 ไปจนถึงข้อที่ 100 เช่นเดียวกันโดยบันทึกค่าเปอร์เซ็นต์ลงในแบบฟอร์ม ดังต่อไปนี้

ตารางที่ 6 เปอร์เซนต์ที่จะตอบถูกของนิสิตที่มีความสามารถต่ำสุดระดับ B1

ข้อที่	เปอร์เซนต์ที่จะตอบถูก
1	
2	
3	
.	
.	
.	
98	
99	
100	

เมื่อเสร็จสิ้นการพิจารณาแบบทดสอบเทียบเคียงกับกรอบมาตรฐานฯ ทั้งสามระดับในรอบที่ 1 แล้วผู้ดำเนินการฯ ได้แสดงค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานคะแนนจุดตัดของแต่ละระดับ

3.5 ขั้นตอนวิธีการสร้างมาตรฐานเพื่อหาคะแนนจุดตัด รอบที่ 2

ในการดำเนินการเทียบคะแนนในรอบที่ 2 มีขั้นตอนเหมือนกับการเทียบคะแนนในรอบที่ 1 แต่ผู้ดำเนินการฯ ได้ให้ข้อมูลสารสนเทศเพิ่มเติม 2 เรื่อง คือ 1. ค่าของคะแนนจุดตัดในรอบที่ 1 และ 2. ค่าคุณภาพของข้อสอบรายข้อ ได้แก่ ค่าความยากง่าย และค่าอำนาจจำแนก เพื่อพิจารณาเพิ่มเติม เมื่อผู้เชี่ยวชาญได้รับข้อมูลสารสนเทศเพิ่มเติม ผู้เชี่ยวชาญสามารถตัดสินใจใหม่ได้ว่าเปอร์เซนต์ที่นิสิตที่มีความสามารถต่ำสุดในระดับนั้น ๆ จะตอบถูกเป็นเท่าไร เช่น ในรอบแรกอาจจะให้เปอร์เซนต์สูง แต่เมื่อได้ข้อมูลเพิ่มเติมในรอบนี้ว่าข้อสอบข้อนั้นมีค่าความยากมาก มีนิสิตตอบถูกน้อยมาก ผู้เชี่ยวชาญอาจปรับลดระดับเปอร์เซนต์การตอบถูกลงมาได้ ในทางกลับกัน ถ้าข้อสอบใดมีค่าความยากไม่มากนัก มีนิสิตตอบถูกจำนวนมาก ผู้เชี่ยวชาญอาจเพิ่มระดับเปอร์เซนต์การตอบถูกขึ้นได้

3.6 ขั้นตอนวิธีการสร้างมาตรฐานเพื่อหาคะแนนจุดตัด รอบที่ 3

ผู้ดำเนินการฯ ได้ให้ผู้เชี่ยวชาญ พิจารณาถึงความเหมาะสมของคะแนนจุดตัดในระดับ A2, B1, และ B2 พร้อมเปิดโอกาสให้ผู้เชี่ยวชาญ อภิปรายแลกเปลี่ยนถึงความเหมาะสมของคะแนนในแต่ละช่วง

ระดับว่ามีความเป็นไปได้และเหมาะสมกับการใช้วัดและประเมินผลความสามารถภาษาอังกฤษของนิสิตมหาวิทยาลัยศรีนครินทรวิโรฒ ให้ผู้เชี่ยวชาญได้ปรับค่าคะแนนจุดตัดรอบสุดท้าย

ดังนั้นจะเห็นว่า การสร้างมาตรฐานเพื่อหาคะแนนจุดตัดในรอบที่ 1 ประเด็นสำคัญคือผู้เชี่ยวชาญทั้ง 14 คนต้องใช้วิจารณญาณของตนเองในการตัดสินผล ในรอบที่ 2 ผู้เชี่ยวชาญจะได้ข้อมูลสารสนเทศของข้อสอบรายข้อเพิ่มเพื่อช่วยในการตัดสินใจ ส่วนในรอบที่ 3 เปิดโอกาสให้ผู้เชี่ยวชาญอภิปรายแลกเปลี่ยนถึงความเหมาะสมของคะแนนในแต่ละระดับ ซึ่งการตัดสินผลรอบที่ 3 จะเป็นรอบที่ผู้เชี่ยวชาญมีความเห็นใกล้เคียงกันมากที่สุดเกี่ยวกับจำนวนข้อที่ผู้เข้าสอบแต่ละระดับสามารถทำได้

กระบวนการในข้อ 3.2 ถึง 3.6 ใช้เวลาทั้งหมด 3 วันเต็มเพื่อทำการสร้างคะแนนจุดตัดทั้ง 3 รอบ หลังจากนั้น ผู้เชี่ยวชาญได้ทำการประเมินเกี่ยวกับความเข้าใจในวิธีการ ความมั่นใจในการลงความเห็น ตลอดจนความพึงพอใจในการเข้าร่วมการสร้างเกณฑ์มาตรฐานเพื่อหาคะแนนจุดตัด และผลการประเมินพบว่าผู้เชี่ยวชาญมีความพึงพอใจกับกระบวนการเตรียมความพร้อมเป็นอย่างมาก มีความมั่นใจในการสร้างคะแนนจุดตัด สามารถใช้ประสบการณ์และความชำนาญในการสอนมาช่วยในการลงความเห็นและมีความมั่นใจในการกำหนดผลคะแนนแบบทดสอบเมื่อเทียบกับกรอบมาตรฐาน ฯ

ผลการเทียบคะแนนแบบทดสอบ SWU-SET กับกรอบมาตรฐานอ้างอิงสากลของยุโรป

หลังการเทียบคะแนนแบบทดสอบ SWU-SET กับกรอบมาตรฐานฯ ได้ผลการเทียบคะแนนดังนี้

ตารางที่ 7 การเทียบคะแนนแบบทดสอบ SWU-SET ระดับ A2 รอบที่ 1, 2, และ 3

ตอน	รอบที่ 1		รอบที่ 2		รอบที่ 3	
	ค่าเฉลี่ย	ค่าส่วนเบี่ยงเบนมาตรฐาน	ค่าเฉลี่ย	ค่าส่วนเบี่ยงเบนมาตรฐาน	ค่าเฉลี่ย	ค่าส่วนเบี่ยงเบนมาตรฐาน
การฟัง	37.09	5.72	40.79	4.95	40.06	2.79
คำศัพท์	26.30	5.91	26.07	3.44	26.07	3.43
การใช้ภาษา	20.36	6.33	19.66	4.10	19.66	4.10
โครงสร้างภาษา	21.18	5.84	16.41	4.09	16.41	4.09
การอ่าน	11.73	6.71	8.86	4.49	8.86	4.49
รวม	23.18	6.10	22.16	4.21	22.01	3.78

จากผลการเทียบคะแนนแบบทดสอบ SWU-SET ระดับ A2 รอบที่ 1 พบว่าค่าเฉลี่ยของคะแนนจุดตัดอยู่ที่ 23.18 คะแนน และค่าส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 6.10 คะแนน รอบที่ 2 พบว่าค่าเฉลี่ยของคะแนนจุดตัด และค่าส่วนเบี่ยงเบนมาตรฐานลดลง โดยมีค่าเฉลี่ยรอบที่ 2 อยู่ที่ 22.16 คะแนน และรอบที่ 3 อยู่ที่ 22.01 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐานรอบที่ 2 อยู่ที่ 4.21 คะแนน และรอบที่ 3 อยู่ที่ 3.78 คะแนน ตามลำดับ

ในภาพรวม ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานลดลงจากรอบที่ 1 มารอบที่ 2 และรอบที่ 3 ทั้งนี้เนื่องจากผู้เชี่ยวชาญมีความเห็นที่สอดคล้องกันมากยิ่งขึ้นจากการอภิปรายแลกเปลี่ยนประกอบกับข้อมูลสารสนเทศเกี่ยวกับคุณภาพข้อสอบรายข้อที่ผู้ดำเนินการฯ ให้เพิ่มเติม

ตารางที่ 8 การเทียบคะแนนแบบทดสอบ SWU-SET ระดับ B2 รอบที่ 1, 2, และ 3

ตอน	รอบที่ 1		รอบที่ 2		รอบที่ 3	
	ค่าเฉลี่ย	ค่าส่วนเบี่ยงเบนมาตรฐาน	ค่าเฉลี่ย	ค่าส่วนเบี่ยงเบนมาตรฐาน	ค่าเฉลี่ย	ค่าส่วนเบี่ยงเบนมาตรฐาน
การฟัง	85.05	3.05	81.74	4.46	81.74	4.46
คำศัพท์	84.29	9.88	74.81	3.88	74.27	3.88
การใช้ภาษา	87.35	3.51	84.30	3.28	84.30	3.28
โครงสร้างภาษา	82.12	3.25	73.67	4.76	73.08	4.76
การอ่าน	82.47	3.18	74.52	3.45	73.08	3.45
รวม	84.22	4.57	77.73	3.97	77.52	3.97

จากผลการเทียบคะแนนแบบทดสอบระดับ B2 รอบที่ 1 พบว่าค่าเฉลี่ยของคะแนนจุดตัดอยู่ที่ 84.22 คะแนน และค่าส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 4.57 คะแนน รอบที่ 2 และรอบที่ 3 พบว่าค่าเฉลี่ยของคะแนนจุดตัดลดลง โดยมีค่าเฉลี่ย 77.73 คะแนน และ 77.52 คะแนน ตามลำดับ และค่าส่วนเบี่ยงเบนมาตรฐานลดลงจากรอบที่ 1 มารอบที่ 2 และครั้งที่ในรอบที่ 2 มารอบที่ 3 (4.57, 3.97 และ 3.97 ตามลำดับ)

ในภาพรวม ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานลดลงในแต่ละรอบ เกิดจากความเห็นของผู้เชี่ยวชาญที่เป็นไปในทางเดียวกันมากขึ้นเช่นเดียวกับการเทียบคะแนนในระดับ A2

ส่วนการเทียบคะแนนในระดับ B1 ในภาพรวมพบว่า ค่าเฉลี่ยของรอบที่ 1 คือ 53.70 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 5.34 คะแนน รอบที่ 2 ค่าเฉลี่ย 49.95 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 4.09 คะแนน และรอบที่ 3 ค่าเฉลี่ย 49.77 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 3.83 คะแนน ดังนั้น ค่าคะแนนจุดตัดรอบที่ 1 ถึง 3 ในภาพรวมมีค่าลดลง ส่วนค่าส่วนเบี่ยงเบนมาตรฐานก็ลดลงเช่นเดียวกัน โดยสรุปการเทียบคะแนนระดับ B1 อยู่ที่ประมาณ 50 คะแนน ซึ่งเป็นคะแนนที่ผู้เชี่ยวชาญเห็นพ้องต้องกันและคิดว่าเป็นคะแนนจุดตัดที่เหมาะสมสำหรับระดับ B1

สรุปผลการเทียบคะแนนแบบทดสอบ SWU-SET กับกรอบมาตรฐานอ้างอิงสากลของยุโรป

จากผลการเทียบคะแนนแบบทดสอบ SWU-SET กับกรอบมาตรฐานฯ ได้ค่าคะแนนจุดตัด ดังนี้

ตารางที่ 9 ผลการเทียบคะแนนแบบทดสอบ SWU-SET กับกรอบมาตรฐานอ้างอิงสากลของยุโรป

คะแนน SWU-SET	กรอบมาตรฐานฯ
78-100	B2
50-77	B1
22-49	A2

ตารางที่ 9 แสดงให้เห็นถึงค่าคะแนน SWU-SET โดยเทียบกับกรอบมาตรฐานฯ ทำให้ค่าคะแนนนั้น มีความหมายมากยิ่งขึ้นว่าผู้เข้าสอบสามารถใช้ภาษาอังกฤษทำอะไรได้บ้าง เมื่อตีความตามคำบรรยาย คุณลักษณะตามกรอบมาตรฐานฯ เช่น ผู้ที่ได้คะแนน SWU-SET ระหว่าง 22-49 คะแนนสามารถเข้าใจ ประโยคและวลีที่ใช้อยู่และเกี่ยวข้องกับเรื่องใกล้ตัว (เช่น เรื่องข้อมูลส่วนตัวและครอบครัว การซื้อของ ภูมิศาสตร์ท้องถิ่น การจ้างงาน) สามารถสื่อสารอย่างง่ายในเรื่องที่เป็นปกติวิสัย โดยใช้การแลกเปลี่ยน ข้อมูลง่าย ๆ ตรงไปตรงมา เกี่ยวข้องกับเรื่องที่คุ้นเคย และที่ทำเป็นประจำ สามารถบรรยายโดยใช้คำง่าย ๆ ที่เกี่ยวข้องกับภูมิหลังของตน เรื่องแวดล้อมใกล้ตัว และเรื่องที่เกี่ยวข้องกับความต้องการเร่งด่วน เป็นต้น

ทั้งนี้ผู้ที่ได้คะแนนต่ำกว่า 22 คะแนนจัดว่ามีความสามารถต่ำกว่าระดับ A2 นอกจากนี้ผลการเทียบ คะแนนครั้งนี้ไม่ได้เทียบกับระดับ C1 หรือ C2 จึงคาดการณ์ว่าผู้ที่ได้คะแนน 78 ขึ้นไปมีความสามารถ ภาษาอังกฤษที่ระดับ B2 หรืออาจจะสูงกวานั้น ทั้งนี้ต้องใช้ข้อมูลอื่น ๆ ประกอบการพิจารณา เช่น ความสามารถในการพูดหรือการเขียน เป็นต้น

นอกจากนี้ ศูนย์ภาษาและบริการวิชาการได้นำผลการเทียบคะแนนในตารางที่ 9 ไปทดลองใช้ พิจารณาความสามารถทางภาษาอังกฤษของผู้เข้าสอบที่เป็นนิสิตชั้นปีที่ 1 และ 2 จำนวน 271 คน พบว่า มีผู้เข้าสอบที่ได้ระดับ A2 55.63 เปอร์เซ็นต์ ได้ระดับ B1 39.84 เปอร์เซ็นต์ และระดับ B2 4.51 เปอร์เซ็นต์ จึงสรุปได้ว่านิสิตชั้นปีที่ 1 และ 2 มีความสามารถภาษาอังกฤษอยู่ในระดับ A2 รองลงมาคือระดับ B1 และ B2 (หรืออาจจะสูงกวานั้น) ตามลำดับ

ข้อจำกัดในการวิจัย

1. แบบทดสอบ SWU-SET มีเพียงการวัดทักษะการรับสาร (Receptive Skill) ส่วนการวัดทักษะ การส่งสาร (Productive Skill) อยู่ในขั้นตอนการพัฒนา เมื่อพัฒนาเสร็จสิ้นแล้วจะดำเนินการเทียบ คะแนนกับกรอบมาตรฐานฯ ต่อไป

2. การเทียบผลคะแนนกับกรอบมาตรฐานฯ ขึ้นอยู่กับการเห็นพ้องต้องกันของผู้เชี่ยวชาญ ซึ่งมีความ เป็นอัตนัย (Subjective Agreement) ดังนั้น ความเห็นดังกล่าวเกิดขึ้นเฉพาะบริบทของมหาวิทยาลัย ศรีนครินทรวิโรฒที่ใช้แบบทดสอบ SWU-SET ในการวัดสมิทธิภาพภาษาอังกฤษ ไม่สามารถนำไปใช้ อ้างอิงกับบริบทอื่นที่มีความแตกต่างกัน

References

- Cizek, G. J. (2012). An introduction to contemporary standard setting. In G. J. Cizek (Ed.), **Setting performance standards: Foundations, methods, and innovations** (2nd ed., pp.467-484). New York: Routledge.
- Council of Europe (2001). **Common European Framework of Reference for Languages: Learning, Teaching, Assessment**. Cambridge, Cambridge University Press.
- Council of Europe. (2003). **Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)** [On-line]. Available at: <http://rm.coe.int/1680667a2d>
- Office of the Basic Education Commission. (2014). **Ministry of Education’s Guidance on Teaching and Learning English Language Reform**. [On-line]. Available at: [file:///C:/Users/LC-09/Downloads/5705140101%20\(3\).pdf](file:///C:/Users/LC-09/Downloads/5705140101%20(3).pdf)
- Office of the Higher Education Commission. (2016). **The Enhancing Standard of English Proficiency in Higher Educational Institutes Policy**. [On-line]. Available at: http://www.mua.go.th/users/bhes/front_home/Data%20Bhes_2559/04052559.pdf
- Shin, S. Y. and Lidster, R. (2014). Evaluating different standard-setting methods in a ESL placement test context. **Language Testing**, 34(3): 357-386.
- Tannenbaum, R. J. and Baron, P. A. (2011). Mapping TOEIC® ITP Scores Onto the Common European Framework of Reference. **Research Memorandum ETS RM—11-33**. Princeton, NJ: Educational Testing Service.
- Tannenbaum, R. J. and Wylie, E. C. (2008). **Linking English language test scores onto the Common European Framework of Reference: An application of standard setting methodology**. Princeton, NJ: Educational Testing Service.
- Wudthayagorn, J. (2016). **Mapping CU-TEP (Chulalongkorn University Test of English Proficiency) to Common European Framework of Reference (CEFR)**. Research Report. Learning Innovation Center. Chulalongkorn University.