

การวิเคราะห์ผลการทดสอบและแนวทางเพื่อยกระดับคุณภาพแบบทดสอบความสามารถทาง ภาษาอังกฤษ มหาวิทยาลัยศรีปทุม (SPU-TEP)

ไพบุณย์ สุขวิชิต บาร์^{1,*}, ถิรนนท์ วงศ์วิวัฒน์², Oscar Jr. Malicad³, สุบิน ยุระรัช⁴

^{1,2,3}คณะศิลปศาสตร์ มหาวิทยาลัยศรีปทุม

⁴ศูนย์ส่งเสริมการวิจัยและการประกันคุณภาพการศึกษา มหาวิทยาลัยศรีปทุม

Received: 28 June 2025

Revised: 18 August 2025

Accepted: 1 September 2025

บทคัดย่อ

บทความนี้มีวัตถุประสงค์ 3 ประการ คือ (1) เพื่อวิเคราะห์คะแนนผลการทดสอบความสามารถทางภาษาอังกฤษของนักศึกษามหาวิทยาลัยศรีปทุม (2) เพื่อวิเคราะห์ความยากง่ายและอำนาจจำแนกของแบบทดสอบความสามารถทางภาษาอังกฤษ มหาวิทยาลัยศรีปทุม (SPU-TEP) และ (3) เพื่อนำเสนอแนวทางการปรับปรุงข้อสอบวัดระดับความสามารถทางภาษาอังกฤษของนักศึกษาระดับอุดมศึกษา กลุ่มตัวอย่าง คือ นักศึกษาระดับปริญญาตรี มหาวิทยาลัยศรีปทุม จำนวนผู้เข้าสอบจริง 220 คน คณะผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลด้วยตนเอง วิเคราะห์ข้อมูลโดยใช้สถิติพื้นฐาน ค่าความยากง่าย ค่าอำนาจจำแนก และค่าความเชื่อมั่น ผลการวิจัยพบว่า (1) ผลการทดสอบของนักศึกษา จำนวน 220 คน ที่ทำข้อสอบวัดระดับความสามารถทางภาษาอังกฤษ (SPU-TEP) จำนวน 100 ข้อ พบว่า ค่าเฉลี่ยผลการทดสอบวัดระดับความสามารถทางภาษาอังกฤษ (SPU-TEP) อยู่ในเกณฑ์ระดับปานกลาง ($M = 68.90$, $SD = 13.01$) และผลการวิเคราะห์ระดับความสามารถทางภาษาอังกฤษเทียบกับ CEFR พบว่า ส่วนใหญ่อยู่ในระดับ B2 (ร้อยละ 42.80) รองลงมา คือ ระดับ B1 (ร้อยละ 37.40) ระดับ A2 (ร้อยละ 6.58) และระดับ A1 (ร้อยละ 0.82) ตามลำดับ (2) ค่าความยากง่าย (p) ของข้อสอบ SPU-TEP ที่วัดทักษะการฟังมีค่าอยู่ระหว่าง 0.08-0.98 ส่วนค่าความยากง่ายของข้อสอบ SPU-TEP ที่วัดทักษะการอ่านมีค่าอยู่ระหว่าง 0.27-0.95 สำหรับค่าอำนาจจำแนก (D) ของข้อสอบ SPU-TEP ที่วัดทักษะการฟังมีค่าอยู่ระหว่าง -0.10-0.72 ส่วนค่าอำนาจจำแนกของข้อสอบ SPU-TEP ที่วัดทักษะการอ่านมีค่าอยู่ระหว่าง -0.11-0.71 (3) แนวทางการปรับปรุงข้อสอบวัดระดับความสามารถทางภาษาอังกฤษ ได้แก่ การปรับปรุงโครงสร้างและเนื้อหาของข้อสอบ การเพิ่มรูปแบบการฟังที่หลากหลายสำหรับข้อสอบวัดทักษะการฟัง และการเพิ่มข้อคำถามที่ต้องใช้การวิเคราะห์และสรุปความมากขึ้นสำหรับข้อสอบวัดทักษะการอ่าน

คำสำคัญ: ข้อสอบ แบบทดสอบความสามารถทางภาษาอังกฤษ อุดมศึกษา

*ผู้ประสานงานหลัก; อีเมล: piboon.su@spu.ac.th

An Analysis of Test Result and Guidelines for Enhancing the Quality of the Sripatum University Test of English Proficiency (SPU-TEP)

Piboon Sukvijit Barr^{1,*}, Tiranun Wongwiwat², Oscar Jr. Malicad³, Subin Yurarach⁴

^{1,2,3} School of Liberal Arts, Sripatum University, Thailand

⁴ Research Facilitation and Development Center, Sripatum University, Thailand

Received: 28 June 2025

Revised: 18 August 2025

Accepted: 1 September 2025

Abstract

This article consists of three main objectives: (1) to analyze the English proficiency test scores of Sripatum University students, (2) to examine the difficulty levels and discrimination indices of the Sripatum University Test of English Proficiency (SPU-TEP), and (3) to propose approaches for improving the English proficiency test for higher education students. The sample consisted of 220 undergraduate students from Sripatum University, who took the test. Data were collected by the research team and analyzed using basic statistics, item difficulty (p), item discrimination (D), and reliability values. The research findings revealed that (1) The test results of 220 students who took the English proficiency test (SPU-TEP), which consisted of 100 items, revealed that the average score was at a moderate level ($M = 68.90$, $SD = 13.01$). When comparing the English proficiency levels with the CEFR, it was found that the majority of students were at the B2 level (42.80%), followed by B1 (37.40%), A2 (6.58%), and A1 (0.82%), respectively. (2) The item difficulty (p) of the SPU-TEP listening section ranged from 0.08 to 0.98, while that of the reading section ranged from 0.27 to 0.95. The item discrimination (D) for the listening section ranged from -0.10 to 0.72, and for the reading section, from -0.11 to 0.71. (3) The suggested approaches for improving the English proficiency test include revising the test structure and content, incorporating more diverse listening formats in the listening section, and adding more analytical and inferential items in the reading section.

Keywords: Test Item, English Proficiency Test, Higher Education

*Corresponding Author; Email: piboon.su@spu.ac.th

ความเป็นมาและความสำคัญของปัญหา

การพัฒนาความสามารถทางภาษาอังกฤษของนักศึกษาระดับอุดมศึกษาเป็นหนึ่งในประเด็นสำคัญที่ได้รับ ความสนใจอย่างต่อเนื่องจากหน่วยงานด้านการศึกษา ทั้งในระดับนโยบายของภาครัฐและในระดับสถาบันการศึกษา ต่างๆ ทั่วประเทศ เนื่องจากภาษาอังกฤษถือเป็นเครื่องมือสำคัญในการสื่อสาร การแสวงหาความรู้ และการทำงานใน สังคมโลกยุคใหม่ (Kirkpatrick, 2012) การผลิตบัณฑิตที่มีทักษะภาษาอังกฤษที่เพียงพอจึงเป็นปัจจัยสำคัญที่ส่งผลต่อ คุณภาพของแรงงานและการยกระดับขีดความสามารถในการแข่งขันของประเทศ (Office of the Higher Education Commission: OHEC, 2019)

ในปี พ.ศ. 2567 คณะกรรมการมาตรฐานการอุดมศึกษา (กมอ.) ได้ประกาศนโยบายเรื่องการยกระดับ มาตรฐานภาษาอังกฤษของนักศึกษาในระดับอุดมศึกษา โดยเน้นย้ำให้นักศึกษาทุกคนต้องผ่านเกณฑ์การวัดระดับ ความสามารถทางภาษาอังกฤษก่อนสำเร็จการศึกษา (Office of the Higher Education Standards Committee, 2024) นโยบายนี้มีเป้าหมายเพื่อผลักดันให้ทุกมหาวิทยาลัยในประเทศไทยพัฒนาระบบการเรียนการสอนและ การประเมินผลทางภาษาอังกฤษให้มีประสิทธิภาพ และเป็นไปตามกรอบอ้างอิงความสามารถทางภาษาของยุโรป (Common European Framework of Reference for Languages: CEFR) ซึ่งเป็นมาตรฐานสากลในการวัดทักษะ ภาษา (Council of Europe, 2020)

มหาวิทยาลัยศรีปทุม ในฐานะสถาบันอุดมศึกษาชั้นนำแห่งหนึ่ง ได้ตอบสนองต่อนโยบายดังกล่าวอย่าง ชัดเจน โดยประกาศแนวนโยบายในการยกระดับทักษะภาษาอังกฤษของนักศึกษาระดับปริญญาตรีอย่างจริงจัง ทั้งนี้ เพื่อเตรียมความพร้อมให้นักศึกษาสามารถประกอบอาชีพในบริบทสากลได้อย่างมีประสิทธิภาพ (Sripatum University, 2024) ในการดำเนินการตามนโยบายดังกล่าว มหาวิทยาลัยได้จัดตั้ง “คณะกรรมการขับเคลื่อน การยกระดับมาตรฐานภาษาอังกฤษของนักศึกษา” ซึ่งในการประชุมครั้งที่ 3 เมื่อวันที่ 18 มกราคม พ.ศ. 2567 ได้มี มติให้ยกเลิกประกาศแนวทางเดิม (ปก.มศป.050/2562) และกำหนดแนวปฏิบัติใหม่ในการวัดระดับความสามารถทาง ภาษาอังกฤษของนักศึกษา โดยเน้นให้การวัดผลมีมาตรฐาน มีความน่าเชื่อถือ และสามารถใช้อย่างเหมาะสมเพื่อ การพัฒนาศักยภาพนักศึกษาได้อย่างแท้จริง (English Language Standards Enhancement Steering Committee, 2024)

จากนโยบายใหม่นี้ คณะศิลปศาสตร์ มหาวิทยาลัยศรีปทุม จึงได้รับมอบหมายให้ดำเนินการพัฒนา แบบทดสอบความสามารถทางภาษาอังกฤษของนักศึกษาในชื่อ Sripatum University Test of English Proficiency (SPU-TEP) เพื่อใช้เป็นเครื่องมือหลักในการประเมินความสามารถด้านการฟังและการอ่านของนักศึกษาระดับปริญญาตรี โดยในการออกแบบข้อสอบดังกล่าว คณะผู้จัดทำได้ร่วมมือกับผู้เชี่ยวชาญจากสถาบันภาษาที่มีประสบการณ์ในการสอน ภาษาและการวัดผลทางภาษาตามมาตรฐานสากล เพื่อให้แน่ใจว่าแบบทดสอบจะมีคุณภาพสูง มีความเที่ยงตรง (validity) และความเชื่อถือได้ (reliability) ตามหลักทฤษฎีการวัดผลทางจิตวิทยาการศึกษา (Brown, 2014; Fulcher & Davidson, 2007)

โดยกระบวนการในการพัฒนาข้อสอบความสามารถทางภาษาอังกฤษ (SPU-TEP) พัฒนาโดยทีมอาจารย์ ผู้สอนภาษาอังกฤษที่มีประสบการณ์การสอน ซึ่งทีมอาจารย์พัฒนาข้อสอบนี้ได้ผ่านการอบรมการสร้างข้อสอบ และ ทำการแบ่งกันสร้างข้อสอบตามกรอบ CEFR ในสองทักษะคือทักษะการฟัง และทักษะการอ่าน เพื่อออกแบบ

ข้อคำถามให้สอดคล้องกับระดับความสามารถในการใช้ภาษาอังกฤษของผู้เรียนแต่ละระดับ และแบ่งระดับความยาก-ง่ายของข้อสอบรายข้อตามสัดส่วนที่เหมาะสม และผ่านการประเมินคุณภาพข้อสอบโดยผู้ทรงคุณวุฒิ

การพัฒนาแบบทดสอบที่มีประสิทธิภาพไม่เพียงแต่จะช่วยให้สามารถประเมินทักษะของนักศึกษาได้อย่างแม่นยำ แต่ยังมีส่วนสำคัญในการออกแบบแผนการเรียนรู้เฉพาะบุคคล (Personalized learning) การจัดกลุ่มเรียนเสริม และการกำหนดมาตรการเพื่อสนับสนุนการพัฒนาภาษาอังกฤษอย่างต่อเนื่อง ซึ่งจะเป็นกลไกสำคัญในการขับเคลื่อนคุณภาพการศึกษาของมหาวิทยาลัยศรีปทุมให้สอดคล้องกับบริบทโลกในศตวรรษที่ 21

วัตถุประสงค์ของการวิจัย

1. เพื่อวิเคราะห์คะแนนผลการทดสอบความสามารถทางภาษาอังกฤษ มหาวิทยาลัยศรีปทุม
2. เพื่อวิเคราะห์คะแนนผลการทดสอบความสามารถทางภาษาอังกฤษของนักศึกษามหาวิทยาลัยศรีปทุม
3. เพื่อนำเสนอแนวทางการปรับปรุงข้อสอบวัดระดับความสามารถทางภาษาอังกฤษของนักศึกษาระดับอุดมศึกษา

วิธีดำเนินการวิจัย

แบบแผนการวิจัย

การวิจัยเรื่องนี้เป็นการศึกษาสถาบัน (Institutional Research) โดยดำเนินการทดลองใช้ (pilot test) กับกลุ่มนักศึกษาที่ลงทะเบียนในรายวิชา EBC323 - ภาษาอังกฤษสำหรับงานบริการลูกค้า (English for Customer Services) ในภาคการศึกษา 1/2566

ประชากรและตัวอย่าง

ประชากร คือ นักศึกษามหาวิทยาลัยศรีปทุม ปีการศึกษา 2566 จำนวน 654 คน

ตัวอย่าง คือ นักศึกษามหาวิทยาลัยศรีปทุม จำนวน 243 คน จาก 6 คณะ ได้แก่ (1) วิทยาลัยการท่องเที่ยวและการบริการ (2) คณะศิลปศาสตร์ (3) คณะดิจิทัลมีเดีย (4) วิทยาลัยนานาชาติ (5) คณะวิศวกรรมศาสตร์ และ (6) วิทยาลัยการบินและคมนาคม การเลือกตัวอย่าง ใช้วิธีการเลือกแบบเจาะจง (Purposive Selection) จากนักศึกษาชั้นปีที่ 2 และ 4 ที่ลงทะเบียนรายวิชาภาษาอังกฤษ EBC323 – ภาษาอังกฤษสำหรับงานบริการลูกค้า (English for Customer Services) ในภาคการศึกษา 1/2566

เครื่องมือวิจัย

เครื่องมือที่ใช้ในการวิจัย มี 2 ฉบับ ได้แก่ (1) แบบทดสอบความสามารถทางภาษาอังกฤษ มหาวิทยาลัยศรีปทุม (Sripatum University Test of English Proficiency Test: SPU-TEP) และ (2) แบบบันทึกคะแนนผลการสอบ การตรวจสอบคุณภาพของเครื่องมือวิจัยดำเนินการทั้งความตรงและความเที่ยงเพื่อให้ข้อสอบและแบบสอบมีคุณภาพและสามารถนำไปใช้ได้ (Tabachnick and Fidell, 2019) โดยแบบทดสอบประกอบไปด้วยข้อสอบจำนวน 100 ข้อ แบ่งออกเป็น 2 ส่วน คือ การฟัง และการอ่าน พัฒนาโดยคณะอาจารย์ที่สอนภาษาอังกฤษสื่อสารธุรกิจ มหาวิทยาลัยศรีปทุม ที่มีความสามารถทางด้านภาษาและประสบการณ์ในการสอนภายใต้คำชี้แนะ การตรวจสอบคุณภาพข้อสอบความตรงตามเนื้อหา (Content Validity) ความเที่ยง (Reliability) ความยากง่ายของข้อสอบ โดยผู้เชี่ยวชาญ

ทางด้านการพัฒนาข้อสอบที่มาจากสถาบันภาษา ที่มีศูนย์สอบทางด้านภาษาอังกฤษที่เป็นที่ยอมรับระดับประเทศ ด้านความตรงเชิงเนื้อหา (Content Validity) ใช้การวิเคราะห์ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับ วัตถุประสงค์ (Index of Objective-Congruence: IOC) พบว่า มีค่าอยู่ระหว่าง 0.60-1.00 ส่วนผลการวิเคราะห์ ค่าความเที่ยง (Reliability) ของแบบทดสอบ ใช้สูตร KR20 พบว่า ได้ค่าเท่ากับ 0.87

การเก็บรวบรวมข้อมูล

คณะผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลทั้งจากแบบทดสอบความสามารถทางภาษาอังกฤษและแบบ บันทึกรายชื่อ โดยได้นำแบบทดสอบที่พัฒนาขึ้นจำนวน 1 ชุด รวม 100 ข้อ ไปทดลองใช้ครั้งที่ 1 เพื่อวิเคราะห์คุณภาพ ข้อสอบและปรับปรุงแก้ไข ก่อนนำไปใช้ในการทดลองครั้งที่ 2 โดยจัดให้นักศึกษาเข้าทำข้อสอบผ่านระบบ Google Form ซึ่งช่วยอำนวยความสะดวกในการดึงข้อมูลเพื่อการวิเคราะห์ ทั้งนี้ การสอบได้จัดขึ้นในห้องเรียนภายใต้ การกำหนดเวลาที่ชัดเจน และผลการสอบถูกนับเป็นส่วนหนึ่งของกิจกรรมการเรียนรู้ในรายวิชา เพื่อสร้างแรงจูงใจให้ นักศึกษามีความตั้งใจในการทำข้อสอบอย่างจริงจัง หลังจากนั้น คณะผู้วิจัยได้ดำเนินการประมวลผลและวิเคราะห์ ข้อมูลเชิงสถิติ พร้อมทั้งจัดประชุมหารือร่วมกับผู้ทรงคุณวุฒิเพื่อรับข้อเสนอแนะในการปรับปรุงและพัฒนา แบบทดสอบ ทั้งในชุดที่ใช้อยู่และการสร้างแบบทดสอบในชุดต่อไปในอนาคต

การวิเคราะห์ข้อมูล

ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลโดยใช้สถิติเชิงบรรยาย (Descriptive Statistics) โดยแบ่งผลการวิเคราะห์ ข้อมูลออกเป็น 2 ส่วน ได้แก่ (1) ผลการวิเคราะห์คะแนนการสอบวัดระดับความสามารถทางภาษาอังกฤษ มหาวิทยาลัยศรีปทุม (SPU-TEP) และ (2) การวิเคราะห์ค่าความยากของข้อสอบ (Difficulty Index) และค่าอำนาจ จำแนกของข้อสอบ (Discrimination Index)

สำหรับการวิเคราะห์ค่าความยากง่ายของข้อสอบ การวิเคราะห์ค่าความยากของข้อสอบ (P) ถูกวิเคราะห์ เพื่อให้แน่ใจว่าข้อสอบมีความเหมาะสม ไม่ง่ายหรือยากเกินไป โดยใช้เกณฑ์วัดระดับของค่าความยากง่ายของ โดยกำหนดเกณฑ์ ดังนี้

ตารางที่ 1 เกณฑ์การแปลความหมายค่าความยากง่ายของข้อสอบ (Brown, 2014)

ค่า P	เกณฑ์การแปลความหมาย
$P > 0.80$	ข้อสอบง่ายมาก
$0.61 \leq P \leq 0.80$	ข้อสอบค่อนข้างง่าย
$0.41 \leq P \leq 0.60$	ข้อสอบยากง่ายปานกลาง
$0.21 \leq P \leq 0.40$	ข้อสอบค่อนข้างยาก
$P < 0.20$	ข้อสอบยากมาก

การวิเคราะห์อำนาจจำแนกของข้อสอบ การวิเคราะห์ค่าอำนาจจำแนก (D) ใช้ในการวัดว่าข้อสอบสามารถ แยกความแตกต่างระหว่างนักศึกษาที่มีความสามารถสูงและต่ำได้ดีเพียงใด โดยค่าความจำแนกที่ดีควรมีค่ามากกว่า 0.20 โดยกำหนดเกณฑ์ ดังนี้

ตารางที่ 2 เกณฑ์การแปลความหมายค่าอำนาจจำแนกของข้อสอบ (Ebel & Frisbie, 1991)

ค่า D	เกณฑ์การแปลความหมาย
$D > 0.40$	อำนาจจำแนกดีมาก
$0.30 \leq P \leq 0.39$	อำนาจจำแนกดี
$0.20 \leq P \leq 0.29$	อำนาจจำแนกน้อย (ต้องปรับปรุง)
$P < 0.19$	อำนาจจำแนกไม่ดี (ตัดทิ้ง)

สำหรับหลักการเทียบคะแนนสอบของ SPU-TEP กับระดับ CEFR คะแนน นั้น ผู้พัฒนาการทดสอบ SPU-TEP กำหนดช่วงคะแนนที่สอดคล้องกับแต่ละระดับ CEFR ที่สะท้อนผลความสามารถทางภาษาอังกฤษของนักศึกษาในสองทักษะ ได้แก่ การฟังและการอ่าน โดยการวิเคราะห์ผลการสอบในอดีต ค่าแนะนำจากผู้ทรงคุณวุฒิ และเปรียบเทียบกับตัวชี้วัดของ CEFR (Sripatum University, 2025)

ตารางที่ 3 เกณฑ์การแปลความหมายคะแนนสอบของ SPU-TEP กับ CEFR

SPU-TEP	CEFR	เกณฑ์การแปลความหมาย
0-30	A1	ระดับเริ่มต้น (สามารถใช้ประโยคง่ายๆ ได้)
31-50	A2	ระดับพื้นฐาน (สามารถเข้าใจประโยคและสำนวนที่ใช้บ่อยๆ)
51-70	B1	ระดับปานกลาง (สามารถรับมือกับสถานการณ์ที่อาจเกิดขึ้น)
71-90	B2	ระดับกลางสูง (สามารถสื่อสารได้อย่างมีประสิทธิภาพในบริบทที่หลากหลาย และมีความซับซ้อน)
91-100	C1	ระดับชำนาญการ (สามารถสื่อสารในการอภิปรายที่ซับซ้อนได้หลากหลายหัวข้อ มีทักษะการใช้ภาษาที่ลื่นไหล แม่นยำ และมีประสิทธิภาพสูง)

ผลการวิจัย

1. ผลการวิเคราะห์ข้อสอบวัดระดับความสามารถทางภาษาอังกฤษที่พัฒนาขึ้นโดยทีมอาจารย์ของคณะศิลปศาสตร์ มหาวิทยาลัยศรีปทุม กรุงเทพมหานคร

ตารางที่ 4 ผลการวิเคราะห์คะแนนผลการทดสอบวัดระดับความสามารถทางภาษาอังกฤษ (SPU-TEP)

(n=220)

คะแนนผลการทดสอบ	M	SD	แปลผล
คะแนนผลการทดสอบ SPU-TEP: ทักษะการฟัง	32.67	6.88	-
คะแนนผลการทดสอบ SPU-TEP: ทักษะการอ่าน	36.24	7.71	-
คะแนนรวม	68.90	13.01	ปานกลาง

จากตารางพบว่า ผลการทดสอบของนักศึกษาที่ทำข้อสอบวัดระดับความสามารถทางภาษาอังกฤษ (SPU-TEP) จำนวน 100 ข้อ จำนวน 220 คน พบว่า ค่าเฉลี่ยผลการทดสอบวัดระดับความสามารถทางภาษาอังกฤษ (SPU-TEP) อยู่ในเกณฑ์ระดับปานกลาง ($M = 68.90$, $SD = 13.01$) ทั้งนี้ ค่าเฉลี่ยผลการทดสอบทักษะการฟัง เท่ากับ 32.67 ($M = 32.67$, $SD = 6.88$) ส่วนค่าเฉลี่ยผลการทดสอบทักษะการอ่านเท่ากับ 36.24 ($M = 36.24$, $SD = 7.71$)

ตารางที่ 5 ผลการวิเคราะห์ระดับความสามารถทางภาษาอังกฤษเทียบกับ CEFR (n=220)

ระดับ CEFR	ความหมาย	จำนวน	ร้อยละ
A1	0-30 (Basic User)	2	0.82
A2	31-50 (Basic User)	16	6.58
B1	51-70 (Independent User)	91	37.40
B2	71-90 (Independent User)	104	42.80
C1	91-100 (Advanced User)	7	2.88
รวม		220	100.00

จากตารางที่ 2 ผลการวิเคราะห์พบว่าผลคะแนนของนักศึกษาเมื่อทำการเทียบกับระดับ The Common European Framework of Reference for Languages (CEFR) ส่วนใหญ่อยู่ในระดับ B2 (ร้อยละ 42.80) รองลงมา คือ ระดับ B1 (ร้อยละ 37.40) ระดับ A2 (ร้อยละ 6.58) และระดับ A1 (ร้อยละ 0.82) ตามลำดับ

2. ผลการวิเคราะห์ระดับความยากง่ายและอำนาจจำแนกของแบบทดสอบความสามารถทางภาษาอังกฤษ (SPU-TEP)

ตารางที่ 6 ผลการวิเคราะห์ระดับความยากง่ายและอำนาจจำแนกของแบบทดสอบความสามารถทางภาษาอังกฤษ ของมหาวิทยาลัยศรีปทุม: แบบทดสอบการฟัง (Listening Test)

Test Item	Difficulty (P)	Interpretation	P_T	P_L	Discrimination (D)	Interpretation
1	0.85	ง่ายมาก	0.96	0.66	0.30	อำนาจจำแนกดี
2	0.13	ยากมาก	0.11	0.21	-0.10	อำนาจจำแนกไม่ดี
3	0.80	ค่อนข้างง่าย	0.98	0.49	0.49	อำนาจจำแนกดีมาก
4	0.73	ค่อนข้างง่าย	0.96	0.36	0.6	อำนาจจำแนกดีมาก
5	0.92	ง่ายมาก	1	0.74	0.26	อำนาจจำแนกน้อย
6	0.89	ง่ายมาก	1	0.72	0.28	อำนาจจำแนกน้อย
7	0.65	ค่อนข้างง่าย	0.89	0.38	0.51	อำนาจจำแนกดีมาก
8	0.46	ยากง่ายปานกลาง	0.87	0.15	0.72	อำนาจจำแนกดีมาก
9	0.32	ค่อนข้างยาก	0.7	0.11	0.59	อำนาจจำแนกดีมาก
10	0.79	ค่อนข้างง่าย	0.96	0.66	0.30	อำนาจจำแนกดี
11	0.89	ง่ายมาก	0.98	0.68	0.30	อำนาจจำแนกดี

Test Item	Difficulty (P)	Interpretation	P_T	P_L	Discrimination (D)	Interpretation
12	0.89	ง่ายมาก	0.98	0.7	0.28	อำนาจจำแนกน้อย
13	0.86	ง่ายมาก	0.96	0.6	0.36	อำนาจจำแนกดี
14	0.84	ง่ายมาก	0.96	0.62	0.34	อำนาจจำแนกดี
15	0.98	ง่ายมาก	1	0.91	0.09	อำนาจจำแนกไม่ดี
16	0.95	ง่ายมาก	1	0.83	0.17	อำนาจจำแนกไม่ดี
17	0.93	ง่ายมาก	1	0.74	0.26	อำนาจจำแนกน้อย
18	0.93	ง่ายมาก	1	0.74	0.26	อำนาจจำแนกน้อย
19	0.88	ง่ายมาก	0.98	0.6	0.38	อำนาจจำแนกดี
20	0.91	ง่ายมาก	1	0.72	0.28	อำนาจจำแนกน้อย
21	0.35	ค่อนข้างยาก	0.59	0.19	0.4	อำนาจจำแนกดีมาก
22	0.75	ค่อนข้างง่าย	0.89	0.49	0.4	อำนาจจำแนกดีมาก
23	0.39	ค่อนข้างยาก	0.7	0.15	0.55	อำนาจจำแนกดีมาก
24	0.30	ค่อนข้างยาก	0.43	0.32	0.11	อำนาจจำแนกไม่ดี
25	0.55	ยากง่ายปานกลาง	0.85	0.28	0.57	อำนาจจำแนกดีมาก
26	0.49	ยากง่ายปานกลาง	0.67	0.17	0.5	อำนาจจำแนกดีมาก
27	0.75	ค่อนข้างง่าย	0.96	0.32	0.64	อำนาจจำแนกดีมาก
28	0.43	ยากง่ายปานกลาง	0.56	0.38	0.18	อำนาจจำแนกไม่ดี
29	0.71	ค่อนข้างง่าย	0.87	0.45	0.42	อำนาจจำแนกดีมาก
30	0.38	ค่อนข้างยาก	0.83	0.13	0.7	อำนาจจำแนกดีมาก
31	0.76	ค่อนข้างง่าย	1	0.57	0.43	อำนาจจำแนกดีมาก
32	0.18	ยากมาก	0.31	0.19	0.12	อำนาจจำแนกไม่ดี
33	0.83	ง่ายมาก	0.96	0.6	0.36	อำนาจจำแนกดี
34	0.22	ค่อนข้างยาก	0.3	0.23	0.07	อำนาจจำแนกไม่ดี
35	0.42	ยากง่ายปานกลาง	0.78	0.19	0.59	อำนาจจำแนกดีมาก
36	0.85	ง่ายมาก	0.98	0.53	0.45	อำนาจจำแนกดีมาก
37	0.63	ค่อนข้างง่าย	0.94	0.32	0.62	อำนาจจำแนกดีมาก
38	0.75	ค่อนข้างง่าย	0.96	0.43	0.53	อำนาจจำแนกดีมาก
39	0.91	ง่ายมาก	0.98	0.68	0.30	อำนาจจำแนกดี
40	0.88	ง่ายมาก	1	0.49	0.51	อำนาจจำแนกดีมาก
41	0.08	ยากมาก	0.09	0.11	-0.02	อำนาจจำแนกไม่ดี
42	0.66	ค่อนข้างง่าย	0.89	0.47	0.42	อำนาจจำแนกดีมาก
43	0.93	ง่ายมาก	1	0.81	0.19	อำนาจจำแนกไม่ดี
44	0.79	ค่อนข้างง่าย	0.93	0.53	0.40	อำนาจจำแนกดีมาก

Test Item	Difficulty (P)	Interpretation	P _T	P _L	Discrimination (D)	Interpretation
45	0.41	ยากง่ายปานกลาง	0.63	0.32	0.31	อำนาจจำแนกดี
46	0.76	ค่อนข้างง่าย	0.85	0.62	0.23	อำนาจจำแนกน้อย
47	0.81	ง่ายมาก	0.98	0.55	0.43	อำนาจจำแนกดีมาก
48	0.31	ค่อนข้างยาก	0.59	0.23	0.36	อำนาจจำแนกดี
49	0.55	ยากง่ายปานกลาง	0.85	0.4	0.45	อำนาจจำแนกดีมาก
50	0.27	ค่อนข้างยาก	0.5	0.19	0.31	อำนาจจำแนกดี

โดยสรุป ผลการวิเคราะห์ค่าความยากง่ายของข้อสอบของแบบทดสอบการฟัง (Listening Test) พบว่าแบบทดสอบปรนัย มีจำนวนข้อทั้งหมด 50 ข้อ เมื่อพิจารณาคุณภาพรายข้อ สามารถแบ่งออกได้เป็น 5 กลุ่ม คือ (1) ข้อสอบง่ายมาก มีจำนวน 19 ข้อ คิดเป็นร้อยละ 38 (2) ข้อสอบค่อนข้างง่าย มีจำนวน 13 ข้อ คิดเป็นร้อยละ 26 (3) ข้อสอบยากง่ายปานกลาง มีจำนวน 7 ข้อ คิดเป็นร้อยละ 14 (4) ข้อสอบค่อนข้างยาก จำนวน 8 ข้อ คิดเป็นร้อยละ 16 และ (5) ข้อสอบยากมาก จำนวน 3 ข้อ คิดเป็นร้อยละ 6 โดยค่าความยากง่ายของข้อสอบ (P) มีค่าอยู่ระหว่าง 0.08-0.98 ทั้งนี้ ผลการวิเคราะห์ค่าความยากง่ายของข้อสอบที่มีค่าอยู่ระหว่าง 0.21 - 0.80 (ระดับที่เหมาะสมและยอมรับได้) จะถูกนำมาเก็บรวบรวมไว้ในคลังข้อสอบ (Test Bank) ต่อไป

ผลการวิเคราะห์อำนาจจำแนกของข้อสอบปรนัย จำนวน 50 ข้อ พบว่า ข้อสอบมีค่าอำนาจจำแนกอยู่ในระดับดีมาก จำนวน 23 ข้อ คิดเป็นร้อยละ 46 มีค่าอำนาจจำแนกอยู่ในระดับดี จำนวน 11 ข้อ คิดเป็นร้อยละ 22 มีค่าอำนาจจำแนกอยู่ในระดับน้อย จำนวน 7 ข้อ คิดเป็นร้อยละ 14 และมีข้อที่ควรปรับปรุงหรือตัดทิ้ง จำนวน 9 ข้อ คิดเป็นร้อยละ 18 ทั้งนี้ ค่าอำนาจจำแนกของข้อสอบ (D) มีค่าอยู่ระหว่าง -0.10 ถึง 0.72 ทั้งนี้ ผลการวิเคราะห์ค่าอำนาจจำแนกของข้อสอบที่พบว่ามีค่ามากกว่า 0.20 (ระดับที่เหมาะสมและยอมรับได้) จะถูกนำมาเก็บรวบรวมไว้ในคลังข้อสอบ (Test Bank) ต่อไป

ตารางที่ 7 ผลการวิเคราะห์ระดับความยากง่ายและอำนาจจำแนกของแบบทดสอบความสามารถทางภาษาอังกฤษของมหาวิทยาลัยศรีปทุม: แบบทดสอบการอ่าน (Reading Test)

Test Item	Difficulty (P)	Interpretation	P _T	P _L	Discrimination (D)	Interpretation
1	0.71	ค่อนข้างง่าย	0.95	0.46	0.49	อำนาจจำแนกดีมาก
2	0.60	ยากง่ายปานกลาง	0.86	0.43	0.43	อำนาจจำแนกดีมาก
3	0.90	ง่ายมาก	1	0.7	0.3	อำนาจจำแนกดี
4	0.27	ค่อนข้างยาก	0.51	0.13	0.38	อำนาจจำแนกดี
5	0.95	ง่ายมาก	0.96	0.85	0.11	อำนาจจำแนกไม่ดี
6	0.92	ง่ายมาก	1	0.78	0.22	อำนาจจำแนกน้อย
7	0.75	ค่อนข้างง่าย	0.96	0.37	0.59	อำนาจจำแนกดีมาก
8	0.85	ง่ายมาก	0.95	0.65	0.3	อำนาจจำแนกดี
9	0.71	ค่อนข้างง่าย	0.84	0.5	0.34	อำนาจจำแนกดี

วารสารนวัตกรรมการจัดการ ปีที่ 10 ฉบับที่ 2 (กรกฎาคม-ธันวาคม 2568)

Test Item	Difficulty (P)	Interpretation	P _T	P _L	Discrimination (D)	Interpretation
10	0.78	ค่อนข้างง่าย	0.96	0.57	0.39	อำนาจจำแนกดี
11	0.78	ค่อนข้างง่าย	0.89	0.5	0.39	อำนาจจำแนกดี
12	0.92	ง่ายมาก	1	0.8	0.20	อำนาจจำแนกน้อย
13	0.54	ยากง่ายปานกลาง	0.74	0.28	0.46	อำนาจจำแนกดีมาก
14	0.94	ง่ายมาก	0.96	0.85	0.11	อำนาจจำแนกไม่ดี
15	0.73	ค่อนข้างง่าย	0.95	0.43	0.52	อำนาจจำแนกดีมาก
16	0.95	ง่ายมาก	0.98	0.85	0.13	อำนาจจำแนกไม่ดี
17	0.69	ค่อนข้างง่าย	0.7	0.5	0.20	อำนาจจำแนกน้อย
18	0.73	ค่อนข้างง่าย	0.82	0.46	0.36	อำนาจจำแนกดี
19	0.74	ค่อนข้างง่าย	0.91	0.5	0.41	อำนาจจำแนกดีมาก
20	0.94	ง่ายมาก	1	0.8	0.20	อำนาจจำแนกน้อย
21	0.91	ง่ายมาก	1	0.7	0.30	อำนาจจำแนกดี
22	0.73	ค่อนข้างง่าย	0.89	0.46	0.43	อำนาจจำแนกดีมาก
23	0.84	ง่ายมาก	0.98	0.5	0.48	อำนาจจำแนกดีมาก
24	0.86	ง่ายมาก	0.86	0.72	0.14	อำนาจจำแนกไม่ดี
25	0.56	ยากง่ายปานกลาง	0.91	0.24	0.67	อำนาจจำแนกดีมาก
26	0.80	ค่อนข้างง่าย	0.95	0.65	0.30	อำนาจจำแนกดี
27	0.84	ง่ายมาก	0.98	0.5	0.48	อำนาจจำแนกดีมาก
28	0.93	ง่ายมาก	1	0.8	0.20	อำนาจจำแนกพบน้อย
29	0.80	ค่อนข้างง่าย	0.96	0.43	0.53	อำนาจจำแนกดีมาก
30	0.80	ค่อนข้างง่าย	1	0.43	0.57	อำนาจจำแนกดีมาก
31	0.48	ยากง่ายปานกลาง	0.86	0.15	0.71	อำนาจจำแนกดีมาก
32	0.77	ค่อนข้างง่าย	1	0.39	0.61	อำนาจจำแนกดีมาก
33	0.58	ยากง่ายปานกลาง	0.89	0.28	0.61	อำนาจจำแนกดีมาก
34	0.53	ยากง่ายปานกลาง	0.91	0.2	0.71	อำนาจจำแนกดีมาก
35	0.49	ยากง่ายปานกลาง	0.51	0.35	0.16	อำนาจจำแนกไม่ดี
36	0.58	ยากง่ายปานกลาง	0.77	0.39	0.38	อำนาจจำแนกดี
37	0.81	ง่ายมาก	0.95	0.52	0.43	อำนาจจำแนกดีมาก
38	0.44	ยากง่ายปานกลาง	0.79	0.09	0.70	อำนาจจำแนกดีมาก
39	0.65	ค่อนข้างง่าย	0.95	0.35	0.60	อำนาจจำแนกดีมาก
40	0.59	ยากง่ายปานกลาง	0.89	0.2	0.69	อำนาจจำแนกดีมาก
41	0.80	ค่อนข้างง่าย	0.95	0.48	0.47	อำนาจจำแนกดีมาก
42	0.55	ยากง่ายปานกลาง	0.79	0.28	0.51	อำนาจจำแนกดีมาก
43	0.74	ค่อนข้างง่าย	0.98	0.43	0.55	อำนาจจำแนกดีมาก

Test Item	Difficulty (P)	Interpretation	P _T	P _L	Discrimination (D)	Interpretation
44	0.48	ยากง่ายปานกลาง	0.74	0.22	0.52	อำนาจจำแนกดีมาก
45	0.68	ค่อนข้างง่าย	0.91	0.2	0.71	อำนาจจำแนกดีมาก
46	0.85	ง่ายมาก	0.96	0.63	0.33	อำนาจจำแนกดี
47	0.82	ง่ายมาก	0.95	0.78	0.17	อำนาจจำแนกไม่ดี
48	0.77	ค่อนข้างง่าย	0.95	0.48	0.47	อำนาจจำแนกดีมาก
49	0.77	ค่อนข้างง่าย	0.96	0.54	0.42	อำนาจจำแนกดีมาก
50	0.31	ค่อนข้างยาก	0.6	0.11	0.49	อำนาจจำแนกดีมาก

โดยสรุป ผลการวิเคราะห์ค่าความยากง่ายของข้อสอบของแบบทดสอบการฟัง (Listening Test) พบว่า เมื่อพิจารณาคุณภาพเป็นรายข้อ สามารถแบ่งออกได้เป็น 4 กลุ่ม ได้แก่ (1) ข้อสอบง่ายมาก มีจำนวน 16 ข้อ คิดเป็นร้อยละ 32 (2) ข้อสอบค่อนข้างง่าย มีจำนวน 20 ข้อ คิดเป็นร้อยละ 40 (3) ข้อสอบยากง่ายปานกลาง มีจำนวน 12 ข้อ คิดเป็นร้อยละ 24 และ (4) ข้อสอบค่อนข้างยาก จำนวน 2 ข้อ คิดเป็นร้อยละ 4 โดยค่าความยากง่ายของข้อสอบอยู่ในช่วง 0.27 ถึง 0.95 โดยมีค่าความยากง่ายของข้อสอบมีค่าเฉลี่ยอยู่ในระดับค่อนข้างง่าย ($p = .72$) ทั้งนี้ ผลการวิเคราะห์ค่าความยากง่ายของข้อสอบที่อยู่ระหว่าง 0.21 ถึง 0.80 ซึ่งอยู่ในระดับที่เหมาะสมและยอมรับได้ จะถูกนำมาเก็บรวบรวมไว้ในคลังข้อสอบ (Test Bank) ต่อไป

การวิเคราะห์อำนาจจำแนกของข้อสอบ ผลการวิเคราะห์ พบว่า ข้อสอบปรนัยจำนวน 50 ข้อ มีค่าอำนาจจำแนกดีมาก จำนวน 28 ข้อ คิดเป็นร้อยละ 56 ค่าอำนาจจำแนกดี จำนวน 11 ข้อ คิดเป็นร้อยละ 22 ค่าอำนาจจำแนกน้อย จำนวน 5 ข้อ คิดเป็นร้อยละ 10 และข้อที่ควรปรับปรุงหรือตัดทิ้ง จำนวน 6 ข้อ คิดเป็นร้อยละ 12 ทั้งนี้ ค่าอำนาจจำแนกของข้อสอบมีค่าอยู่ระหว่าง 0.11 ถึง 0.71 ทั้งนี้ ผลการวิเคราะห์ค่าอำนาจจำแนกที่มีค่ามากกว่า 0.20 (ระดับที่เหมาะสมและยอมรับได้) จะถูกนำมาเก็บรวบรวมไว้ในคลังข้อสอบ (Test Bank) ต่อไป

โดยสรุป ผลการวิเคราะห์ข้อสอบวัดระดับความสามารถทางภาษาอังกฤษของนักศึกษาระดับอุดมศึกษา สามารถใช้เป็นแนวทางในการพัฒนาคุณภาพข้อสอบให้ดียิ่งขึ้น ทั้งในด้านการปรับปรุงคำถามและตัวเลือกเพื่อให้เหมาะสมกับระดับความยากง่าย การปรับปรุงหรือตัดข้อสอบเพื่อเพิ่มค่าอำนาจจำแนก การปรับโครงสร้างและเนื้อหา ตลอดจนการเพิ่มรูปแบบการฟังและคำถามที่เน้นการวิเคราะห์และสรุปความ นอกจากนี้ ควรมีการอบรมอาจารย์ผู้ออกข้อสอบอย่างต่อเนื่องให้สามารถจัดทำข้อสอบตามกรอบ CEFR เพื่อยกระดับคุณภาพของข้อสอบโดยรวม

อภิปรายผล

1. ผลการวิเคราะห์คะแนนแบบทดสอบความสามารถทางภาษาอังกฤษ มหาวิทยาลัยศรีปทุม (SPU-TEP) ของนักศึกษาจากนักศึกษาชั้นปีที่ 2 และ 4 ที่ลงทะเบียนรายวิชาภาษาอังกฤษ EBC323 – ภาษาอังกฤษสำหรับงานบริการลูกค้า (English for Customer Services) ในภาคการศึกษา 1/2566 ซึ่งแบ่งออกเป็น 2 ทักษะ ได้แก่ การฟัง (Listening) และการอ่าน (Reading) พบว่าคะแนนเฉลี่ยของทักษะการฟังอยู่ในระดับปานกลาง ($M = 32.7$, $SD = 6.79$) และการอ่านก็อยู่ในระดับปานกลางเช่นกัน ($M = 36.00$, $SD = 7.90$) โดยเมื่อเทียบผลคะแนนรวมกับกรอบ CEFR พบว่า นักศึกษาอยู่ในระดับ B2 (42.8%), B1 (37.4%), A2 (6.58%), C1 (2.88%), และ A1 (1.23%) ข้อมูลนี้สะท้อน

ให้เห็นว่ากลุ่มนักศึกษาส่วนใหญ่ยังไม่สามารถผ่านเกณฑ์ระดับ B2 ได้ในสัดส่วนเกินครึ่ง ทั้งที่กลุ่มตัวอย่างส่วนใหญ่มีประสบการณ์การเรียนรู้ภาษาอังกฤษมาหลายรายวิชา ซึ่งสอดคล้องกับแนวคิดของ Weir (2005) และ Alderson (2005) ที่ชี้ว่าการประเมินความสามารถทางภาษาควรสะท้อนผลลัพธ์ของการเรียนรู้ตามระดับที่เหมาะสม หากผลคะแนนโดยรวมยังไม่ถึงระดับเป้าหมาย ย่อมสะท้อนถึงข้อจำกัดเชิงหลักสูตร การจัดการเรียนการสอน หรือวิธีการวัดผล อย่างไรก็ตาม การกำหนดช่วงคะแนนตัดสินผลสอบ SPU-TEP ที่ใช้เทียบกับ CEFR อาจส่งผลต่อความตรงเชิงโครงสร้างของการตีความคะแนน (Construct validity) และอาจทำให้ผลสอบสะท้อนความสามารถทางภาษาอังกฤษของนักศึกษาได้ไม่สมบูรณ์ ทั้งนี้ ต้องคำนึงถึงแง่มุมด้านผลลัพธ์ของความตรงเชิงโครงสร้าง โดยเฉพาะการกำหนดช่วงคะแนนตัดสิน (Cut-off scores) ซึ่งไม่เที่ยงตรง (Invalidity) อาจเกิดจากทั้งกระบวนการกำหนดคะแนนและผลลัพธ์ที่ตามมา (Messick, 1980; Messick, 1989)

2. ผลการวิเคราะห์ความยากง่ายและอำนาจจำแนกของแบบทดสอบความสามารถทางภาษาอังกฤษที่พัฒนาโดยทีมอาจารย์คณะศิลปศาสตร์ มหาวิทยาลัยศรีปทุม การวิเคราะห์เชิงเทคนิคของข้อสอบรายข้อโดยใช้ค่าความยากง่าย (p) และค่าอำนาจจำแนก (Discrimination Index: D) พบว่า ทักษะการฟัง ค่า p อยู่ระหว่าง 0.08–0.98 และ D อยู่ระหว่าง -0.10–0.72 ส่วนทักษะการอ่าน ค่า p อยู่ระหว่าง 0.27–0.95 และ D อยู่ระหว่าง 0.11–0.71 ทั้งนี้ ข้อสอบทั้ง 100 ข้อถูกจัดกลุ่มคุณภาพออกเป็น 3 กลุ่ม ได้แก่ (1) ข้อสอบที่ได้คุณภาพ: การฟัง 34 ข้อ (68%) และการอ่าน 39 ข้อ (78%) โดยมี D อยู่ระหว่าง 0.30–0.72 ซึ่งอยู่ในระดับ “ดี” ตามเกณฑ์ของ Ebel และ Frisbie (1991) (2) ข้อสอบที่ควรปรับปรุง: การฟัง 7 ข้อ (14%) และการอ่าน 5 ข้อ (10%) มี D อยู่ระหว่าง 0.20–0.28 ซึ่งถือว่าพอใช้ได้ แต่ควรปรับปรุงเพื่อเพิ่มความแม่นยำในการวัด และ (3) ข้อสอบที่ควรตัดทิ้ง: การฟัง 9 ข้อ (18%) และการอ่าน 6 ข้อ (12%) ที่มี D ต่ำกว่า 0.20 หรือติดลบ ซึ่งแสดงว่าข้อสอบไม่สามารถแยกความแตกต่างของความสามารถผู้เรียนได้อย่างมีประสิทธิภาพ (Haladyna & Rodriguez, 2013) ค่าเฉลี่ยโดยรวมของข้อสอบฟัง ($p = 0.654$, $D = 0.364$) และอ่าน ($p = 0.723$, $D = 0.417$) อยู่ในเกณฑ์ “ค่อนข้างง่าย” แต่มีค่าอำนาจจำแนกอยู่ในระดับ “ดีถึงดีมาก” ซึ่งถือว่าเป็นเกณฑ์ที่ยอมรับได้ (Nitko, 2001) และสามารถนำไปจัดเก็บในคลังข้อสอบเพื่อใช้ซ้ำหรือปรับใช้ในการประเมินผลในอนาคต จากที่ทีมผู้พัฒนาแบบทดสอบทำการวิเคราะห์ข้อสอบรายข้อ พบว่าข้อสอบที่ไม่ผ่านเกณฑ์เกิดจาก 4 ปัจจัยหลัก คือ (1) ข้อคำถามเน้นเพียงความจำ ความเข้าใจ ไม่ได้เน้นการคิดวิเคราะห์และการสรุปความมากพอ (2) ภาษาที่ยากและง่ายเกินไป และ (3) ตัวเลือกที่ยังไม่สมดุล

3. แนวทางการปรับปรุงข้อสอบวัดระดับความสามารถทางภาษาอังกฤษของนักศึกษา ระดับอุดมศึกษา จากผลการวิจัยดังกล่าว คณะผู้วิจัยเสนอแนวทางการพัฒนาข้อสอบ SPU-TEP ได้แก่ (1) พัฒนาโครงสร้างข้อสอบและเนื้อหาให้หลากหลายมากขึ้น โดยเฉพาะในส่วนของการฟัง ควรมีบทสนทนา บทพูดจากแหล่งต่างๆ ที่สอดคล้องกับระดับภาษา CEFR เพื่อเพิ่มความสามารถในการวัดทักษะในบริบทที่หลากหลาย (Council of Europe, 2001) (2) เพิ่มข้อคำถามที่เน้นการคิดวิเคราะห์และการสรุปความ ตามลำดับขั้นใน Bloom’s Taxonomy (Krathwohl, 2002) ซึ่งช่วยยกระดับข้อสอบจากเพียงการทดสอบความจำไปสู่การวัดความเข้าใจและการประยุกต์ใช้ และ (3) จัดอบรมอาจารย์อย่างสม่ำเสมอ เพื่อพัฒนาแนวทางการออกข้อสอบที่อิงตามมาตรฐาน CEFR โดยเน้นหลักการออกข้อสอบที่มีคุณภาพสูง มีความชัดเจนของตัวเลือก และมีประสิทธิภาพในการวัดผลตามวัตถุประสงค์การเรียนรู้ (Haladyna & Downing, 1989) ข้อเสนอเหล่านี้มีเป้าหมายเพื่อยกระดับคุณภาพการวัดผลทางภาษาของนักศึกษาระดับอุดมศึกษาอย่างยั่งยืน และเป็นพื้นฐานสำคัญในการพัฒนาทั้งคลังข้อสอบ การประเมินผล และคุณภาพการจัดการเรียนการสอน

ของมหาวิทยาลัย นอกจากนี้ จากการวิเคราะห์พบว่าข้อสอบ SPU-TEP ที่พัฒนาขึ้นมีคุณภาพในระดับที่น่าพึงพอใจ ทั้งในแง่ของค่าความยากและค่าความจำแนก ข้อสอบสามารถใช้เป็นเครื่องมือวัดความสามารถทางภาษาอังกฤษของนักศึกษาได้อย่างมีประสิทธิภาพ อย่างไรก็ตาม ควรมีการปรับปรุงบางส่วน เช่น การเพิ่มจำนวนข้อสอบที่มีค่าความจำแนกสูงกว่า 0.40 เพื่อเพิ่มคุณภาพของแบบทดสอบ (Kline, 2015)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

- 1.1 จัดเก็บข้อสอบที่มีคุณภาพไว้ในคลังข้อสอบ (Test Bank) เพื่อใช้ซ้ำและพัฒนาต่อยอดในการสอบครั้งถัดไป โดยเน้นข้อที่มีค่าอำนาจจำแนกสูง ($D > 0.30$)
- 1.2 ออกแบบกิจกรรมการเรียนรู้เสริม สำหรับนักศึกษาที่ได้คะแนนต่ำกว่าเกณฑ์ B2 ตาม CEFR โดยใช้ผลคะแนนสอบเป็นเครื่องมือวางแผนการสอนรายบุคคล (Personalized learning)
- 1.3 อบรมอาจารย์ผู้สร้างข้อสอบอย่างต่อเนื่อง ในการเขียนข้อสอบที่สอดคล้องกับกรอบ CEFR และหลักการสร้างข้อสอบมาตรฐาน เพื่อยกระดับคุณภาพข้อสอบในระยะยาว

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

- 2.1 เปรียบเทียบผลสอบ SPU-TEP กับข้อสอบมาตรฐานสากล (เช่น TOEIC, IELTS) เพื่อประเมินความสอดคล้องและความน่าเชื่อถือของข้อสอบที่พัฒนาเอง
- 2.2 ขยายกลุ่มตัวอย่างไปยังนักศึกษาต่างคณะหรือวิทยาเขตอื่น เพื่อวิเคราะห์ความแตกต่างเชิงสถิติในทักษะภาษาอังกฤษของนักศึกษาต่างกลุ่ม
- 2.3 ศึกษาความสัมพันธ์ระหว่างผลสอบ SPU-TEP กับผลสัมฤทธิ์ทางวิชาการ เพื่อประเมิน Predictive Validity ของข้อสอบต่อการเรียนรู้ในรายวิชาต่างๆ

เอกสารอ้างอิง

- Alderson, J. C. (2005). *Diagnosing Foreign Language Proficiency: The Interface Between Learning and Assessment*. Continuum.
- Brown, J. D. (2014). *Testing in Language Programs*. McGraw-Hill.
- Commission on Higher Education Standards. (2024). *Announcement of the Commission on Higher Education Standards: Policy on Raising English Proficiency Standards in Higher Education Institutions*. Ministry of Higher Education, Science, Research, and Innovation (MHESI), Thailand. (in Thai)
- Committee for Advancing English Proficiency Standards. (2024). *Guidelines for Raising English Proficiency Standards of Students*. Bangkok: Sripatum University. (in Thai)
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. [Online]. Retrieved from: <https://www.coe.int/en/web/common-european-framework-reference-languages>

- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of Educational Measurement*. (5th ed.). Prentice Hall.
- English Language Standards Enhancement Steering Committee. (2024). *Minutes of the 3rd meeting held on January 18, 2024*. Sripatum University. (in Thai)
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.
- Haladyna, T. M., & Downing, S. M. (1989). Validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education*, 2(1), 51–78.
- Kline, P. (2015). *A Handbook of Test Construction: Introduction to Psychometric Design*. Routledge.
- Kirkpatrick, A. (2012). *English as an international language in Asia: Implications for language education*. Springer.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41(4), 212–218.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012–1027.
- Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York, NY: Macmillan.
- Nitko, A. J. (2001). *Educational Assessment of Students*. 3rd ed. Merrill/Prentice Hall.
- Office of the Higher Education Standards Committee. (2024). *Announcement of the Higher Education Standards Committee on the policy for enhancing English language standards in higher education institutions*. Ministry of Higher Education, Science, Research and Innovation. (in Thai)
- Office of the Higher Education Commission. (2019). *National English proficiency development policy*. Bangkok: OHEC.
- Sripatum University. (2025). *Draft Mapping of SPU-TEP English Proficiency Test Results to the CEFR Framework*. Bangkok: School of Liberal Arts, Sripatum University.
- Sripatum University. (2024). *Strategic plan for student development of Sripatum University (2024–2028)*. Sripatum University. (in Thai)
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics*. 7th ed. Pearson.
- Weir, C. J. (2005). *Language Testing and Validation: An Evidence-Based Approach*. Palgrave Macmillan.