

**การพัฒนาแบบทดสอบวัดระดับภาษาอังกฤษ
สำหรับนิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยา**
**A Development of English Proficiency Test
for Undergraduate Students of the University of Phayao**

ชวนพิศ ศรีวิชัย^{1*} ชนาภา ดวงไฟ¹ นริศา ไพเจริญ¹ และ สนิท ยืนศักดิ์¹
Chuanpit Sriwichai^{1*}, Chanapha Duangfai¹, Narisa Paicharoen and
Sanit Yuensak¹

¹ภาควิชาภาษาอังกฤษ คณะศิลปศาสตร์ มหาวิทยาลัยพะเยา จังหวัดพะเยา 56000

¹ English Department, School of Liberal Arts, University of Phayao, Phayao 56000

*Corresponding author: chuanpit.sr@up.ac.th

Received: August 10, 2022; **Revised:** June 16, 2023; **Accepted:** June 26, 2023

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและหาคุณภาพของแบบทดสอบวัดระดับภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยาตามกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรป ระดับ B 1 แบบทดสอบที่พัฒนาขึ้นเป็นแบบเลือกตอบจำนวน 4 ชุด ชุดละ 60 ข้อ เนื้อหาแบ่งเป็น 4 ส่วน ได้แก่ คำศัพท์ การอ่าน ไวยากรณ์และการเขียน และบทสนทนา ส่วนละ 15 ข้อ กลุ่มตัวอย่างคือนิสิตระดับปริญญาตรีชั้นปีที่ 3 จำนวน 227 คน ที่ศึกษาอยู่ใน 18 คณะของมหาวิทยาลัยพะเยา ทีมผู้วิจัยหาคุณภาพแบบทดสอบโดยใช้ดัชนีความสอดคล้องให้ผู้เชี่ยวชาญสาขาภาษาอังกฤษและการวัดประเมินผลการศึกษาจำนวน 3 คน ให้ประเมินความเที่ยงตรงเชิงเนื้อหา จากนั้นนำแบบทดสอบไปทดลองใช้กับกลุ่มตัวอย่างและนำไปวิเคราะห์ค่าความยากง่ายและค่าอำนาจจำแนกโดยใช้เทคนิค 27 % ของจุดเดฟาน เกณฑ์การพิจารณาคือแบบทดสอบข้อที่มีค่าความ

ยากง่าย (P) อยู่ระหว่าง 0.20 ถึง 0.80 และมีค่าอำนาจจำแนก (r) ตั้งแต่ 0.20 ขึ้นไปจะคัดเลือกไว้ จากนั้นวิเคราะห์ค่าความเชื่อมั่นโดยใช้สูตร KR-20

ผลการศึกษาพบว่า ค่าดัชนีความสอดคล้องรายข้อของแบบทดสอบอยู่ระหว่าง 0.33 ถึง 1.00 และรวมรายชุดอยู่ระหว่าง 0.92 ถึง 0.93 การวิเคราะห์ค่าความยากง่ายและค่าอำนาจจำแนกพบว่าแบบทดสอบชุดที่ 1 มีข้อที่ผ่านเกณฑ์ 49 ข้อ ชุดที่ 2 ผ่านเกณฑ์ 48 ข้อ ชุดที่ 3 ผ่านเกณฑ์ 49 ข้อ และชุดที่ 4 ผ่านเกณฑ์ 50 ข้อ และความเชื่อมั่นของแบบทดสอบแต่ละชุดอยู่ระหว่าง 0.79 ถึง 0.87 ซึ่งอยู่ในเกณฑ์ที่ดี

คำสำคัญ: แบบทดสอบวัดระดับภาษาอังกฤษ นิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยา

Abstract

This study aimed to develop English proficiency tests for undergraduate students of the University of Phayao based on the Common European Framework of Reference for Languages B 1 level and to evaluate the quality of the developed tests. There were 4 sets of the tests had been developed. Each set of the test consisted of 60 items, and it was divided into 4 parts including vocabulary, reading, grammar and writing, and conversation. There were 15 items in each part. The sample of this study were 227 third-year students from 18 faculties of the University of Phayao. The tests were verified content validity by 3 experts in the fields of English language and Education Evaluation and Assessment using Index of Item– Objective Congruence (IOC). Then the tests were tried out with the sample to analyze difficulty and discrimination based on 27 % technique of Chung–Teh–Fan. The criteria for index of difficulty and discrimination was the difficulty should be between 0.20 and 0.80 and discrimination should be at least 0.20. Next, the tests were analyzed reliability using KR-20 formula.

The findings showed that overall, the Index of item objective congruence (IOC) of the tests was between 0.33 – 1.00. The IOC of each test was between 0.92–0.93. In terms of the standardized quality of the difficulties and the discrimination of the test, there were 49 qualified items in the first test, 48 qualified items in the second test, 49 qualified items in the third test, and 50 qualified items in the fourth test

respectively. The reliability of each test was between 0.79 and 0.87. This indicated a good quality of the tests.

Keywords: English Proficiency Test, Undergraduate Students, University of Phayao

บทนำ

ปัจจุบันนี้ การปฏิสัมพันธ์กับผู้คนจากหลายประเทศโดยใช้ภาษาอังกฤษเป็นสิ่งที่หลีกเลี่ยงได้ยาก (Praimahaniyom and Chairinkom, 2020: 18) ภาษาอังกฤษจึงมีความสำคัญเป็นอย่างยิ่ง ประเทศไทยเป็นประเทศหนึ่งที่เล็งเห็นถึงความสำคัญดังกล่าว อย่างไรก็ตาม การเรียนภาษาอังกฤษในประเทศไทยยังไม่ประสบความสำเร็จเนื่องจากผู้เรียนส่วนใหญ่ขาดโอกาสในการเรียนรู้และฝึกภาษาอังกฤษในสภาพแวดล้อมที่ใช้ภาษาอังกฤษ (Wachiraya, 2003 อ้างถึงใน Lakhmja and Tachom, 2022: 61) ด้วยเหตุนี้ประเทศไทยจึงมีนโยบายในการปฏิรูปการศึกษา เพื่อพัฒนาให้ผู้เรียนมีศักยภาพในการเรียนรู้ พัฒนาดตนเองได้เต็มศักยภาพ โดยเฉพาะภาษาอังกฤษในทุกหลักสูตร เพื่อรองรับการเข้าสู่ประชาคมอาเซียน และเพิ่มขีดความสามารถในการแข่งขันของประเทศไทยในเวทีโลก สำนักงานคณะกรรมการการอุดมศึกษา จึงได้ออกประกาศ เรื่อง นโยบายการยกระดับมาตรฐานภาษาอังกฤษในสถาบันอุดมศึกษา พ.ศ. 2559 โดยระบุให้สถาบันอุดมศึกษากำหนดนโยบายในการพัฒนาทักษะความสามารถในการใช้ภาษาอังกฤษของนิสิตนักศึกษาให้เป็นบัณฑิตที่มีความพร้อมทั้งวิชาการ วิชาชีพ และทักษะการสื่อสารภาษาอังกฤษในระดับที่ใช้งานได้ (Working Knowledge) โดยการพัฒนาดังกล่าวจะต้องมีตัวชี้วัด และการประเมินผลที่ชัดเจน นอกจากนี้ยังกำหนดให้นิสิตนักศึกษาทุกคนต้องทดสอบความรู้ภาษาอังกฤษตามแบบทดสอบมาตรฐานระดับอุดมศึกษาที่สถาบันสร้างขึ้น หรือที่เห็นสมควรจะนำมาใช้วัดสมรรถภาพทางภาษาอังกฤษ (English Proficiency) (สำนักงานคณะกรรมการการอุดมศึกษา, 2559)

ในปี 2562 มหาวิทยาลัยพะเยาจึงได้ประกาศนโยบายในการยกระดับมาตรฐานภาษาอังกฤษของนิสิตในระดับปริญญาตรีซึ่งสอดคล้องกับประกาศของสำนักงานคณะกรรมการการอุดมศึกษา เพื่อให้ผู้เรียนมีความรู้ความสามารถในการใช้ภาษาอังกฤษสำหรับการสื่อสาร และสำหรับใช้เป็นเครื่องมือค้นคว้าองค์ความรู้ที่เป็นสากล โดยประกาศมหาวิทยาลัยพะเยา เรื่อง นโยบายการยกระดับมาตรฐานภาษาอังกฤษสำหรับนิสิตมหาวิทยาลัยพะเยา พ.ศ. 2562 (มหาวิทยาลัยพะเยา, 2562) ได้กำหนดให้ทุกหลักสูตรที่เปิด

การเรียนการสอนในมหาวิทยาลัยพะเยาพัฒนาปรับปรุงการเรียนการสอนเพื่อพัฒนาทักษะภาษาอังกฤษ รวมถึงการจัดทดสอบวัดความรู้ภาษาอังกฤษให้นิสิตทุกคนก่อนสำเร็จการศึกษา นอกจากนี้ยังได้กำหนดไว้ในประกาศมหาวิทยาลัยพะเยาเรื่อง มาตรฐานการทดสอบวัดระดับความรู้ภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี พ.ศ. 2562 (มหาวิทยาลัยพะเยา, 2562) ว่านิสิตที่มีความประสงค์ที่จะยื่นขอสำเร็จการศึกษาจะต้องมีผลคะแนนจากการทดสอบ TOEFL IBT TOEFL ITP IELTS หรือ การสอบวัดระดับที่จัดขึ้นโดยศูนย์ภาษา คณะศิลปศาสตร์ มหาวิทยาลัยพะเยาที่สามารถเทียบเคียงกับกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรป (Common European Framework of Reference for Languages: CEFR) ในระดับไม่ต่ำกว่า B1 ทั้งนี้ตามกรอบมาตรฐานดังกล่าว ผู้เรียนที่มีทักษะภาษาอังกฤษระดับ B1 ต้องสามารถเข้าใจประเด็นหลักของข้อความที่ถูกสื่อสารด้วยภาษา มาตรฐานอย่างชัดเจน รวมทั้งเข้าใจเนื้อหาที่คุ้นเคยในบริบทของการทำงาน โรงเรียน การใช้ เวลาว่าง และอื่น ๆ ตลอดจนสามารถรับมือกับสถานการณ์ต่าง ๆ ที่มักเกิดขึ้นระหว่างการเดินทางท่องเที่ยวในสถานที่ที่ผู้คนใช้ภาษาอังกฤษ และสามารถพูดหรือเขียนข้อความง่าย ๆ ติดต่อกันในหัวข้อที่คุ้นเคยและเป็นเรื่องที่ตนสนใจ นอกจากนั้นผู้เรียนที่มีทักษะภาษาอังกฤษในระดับ B1 ต้องสามารถบรรยายประสบการณ์ เหตุการณ์ ความใฝ่ฝัน สามารถให้เหตุผลสั้น ๆ และอธิบายความคิดเห็นและแผนการได้ (Council of Europe, n.d.)

ด้วยเหตุผลข้างต้น ทีมวิจัยจึงเล็งเห็นความสำคัญที่จะต้องพัฒนาแบบทดสอบวัดระดับภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยาที่เทียบเคียงมาตรฐานตามกรอบ CEFR หรือกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรป ในระดับ B1 เพื่อให้สอดคล้องตามประกาศมหาวิทยาลัยพะเยา เรื่อง มาตรฐานการทดสอบวัดระดับความรู้ภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี พ.ศ. 2562 ตลอดจนเพื่อให้แบบทดสอบดังกล่าวมีประสิทธิภาพและสามารถใช้เป็นเกณฑ์ในการระบุระดับความรู้ภาษาอังกฤษของนิสิต ทั้งยังสร้างความเชื่อมั่นให้แก่ผู้ใช้บัณฑิต หรือผู้จ้างงานด้วย

วัตถุประสงค์

1. เพื่อพัฒนาแบบทดสอบวัดระดับภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยาตามกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรประดับ B1

2. เพื่อวิเคราะห์คุณภาพแบบทดสอบวัดระดับภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยา ตามกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรประดับ B1 ที่พัฒนาขึ้น

ข้อจำกัดในการวิจัย

ข้อจำกัดในการวิจัยครั้งนี้คือการเข้าถึงกลุ่มตัวอย่างในขั้นตอนการนำแบบทดสอบไปทดลองใช้เนื่องจากช่วงที่ทดลองใช้แบบทดสอบเป็นช่วงระบาคของโคโรนาไวรัส มหาวิทยาลัยพะเยาได้ประกาศให้จัดการเรียนการสอนออนไลน์เพื่อความปลอดภัยของผู้เรียนยกเว้นกรณีเป็นรายวิชาที่ต้องฝึกปฏิบัติและจำเป็นต้องจัดการเรียนการสอนในห้องเรียน ทีมผู้วิจัยจึงขออาสาสมัครกลุ่มตัวอย่างที่สามารถเข้าร่วมการทดลองทำแบบทดสอบในพื้นที่มหาวิทยาลัยได้โดยมีกลุ่มตัวอย่างเข้าร่วมจำนวน 227 คน ซึ่งเป็นสัดส่วนค่อนข้างน้อย เมื่อเทียบกับจำนวนประชากรและขนาดกลุ่มตัวอย่างจากคณะต่าง ๆ อาจไม่เพียงพอที่จะเป็นตัวแทนที่ดีของประชากรจากแต่ละคณะ

วิธีการศึกษา

1. รูปแบบการวิจัย

งานวิจัยนี้ใช้รูปแบบการวิจัยและพัฒนาที่เน้นการออกแบบพัฒนา เดชกุล มัทวานุกุล (ม.ป.ป.) ได้อธิบายว่าระเบียบวิธีวิจัยรูปแบบนี้มุ่งเน้นการสร้างนวัตกรรมหรือประดิษฐ์กรรม รูปแบบการวิจัยคือ

$$R_1 \rightarrow D_1 \rightarrow R_2 \rightarrow D_2$$

รูปแบบการวิจัยข้างต้นหมายถึงการวิจัย 4 ขั้นตอน ได้แก่

R_1 = ขั้นตอนที่ 1 การวิจัย (Research)

D_1 = ขั้นตอนที่ 2 การพัฒนา (Development)

R_2 = ขั้นตอนที่ 3 การวิจัย (Research)

D_2 = ขั้นตอนที่ 4 การพัฒนา (Development)

2. ประชากรและกลุ่มตัวอย่าง

2.1 ประชากร

ประชากรในการวิจัยครั้งนี้คือนิสิตชั้นปีที่ 3 ซึ่งเข้าศึกษาในมหาวิทยาลัยพะเยาปี พ.ศ. 2561 จำนวนประมาณ 4,000 คน ที่มผู้วิจัยเลือกกลุ่มประชากรดังกล่าว เนื่องจากแบบทดสอบที่พัฒนาขึ้น จะนำไปใช้ ทดสอบนิสิตที่เข้าศึกษาในมหาวิทยาลัยตั้งแต่ปีพ.ศ. 2562 เป็นต้นไปตามประกาศนโยบายการยกระดับมาตรฐาน ภาษาอังกฤษ สำหรับนิสิตมหาวิทยาลัยพะเยา พ.ศ. 2562 ประชากรที่นำแบบทดสอบไปทดลองใช้เพื่อหาคุณภาพ จึงควรมีลักษณะใกล้เคียงกัน แต่เป็นคนละกลุ่มกับนิสิตที่จะนำแบบทดสอบไปใช้จริง

2.2 กลุ่มตัวอย่าง

กลุ่มตัวอย่าง ได้แก่ นิสิตจากคณะต่าง ๆ ดังรายชื่อในตารางด้านล่างนี้ รวมจำนวน 227 คน ที่มผู้วิจัยสุ่มตัวอย่างแบบการจัดกลุ่ม (Cluster Random Sampling) ตามคณะที่นิสิตเข้าศึกษา

ตารางที่ 1 รายชื่อคณะและจำนวนกลุ่มตัวอย่าง

คณะ	จำนวน
คณะเกษตรศาสตร์และทรัพยากรธรรมชาติ	7 คน
คณะเทคโนโลยีสารสนเทศและการสื่อสาร	26 คน
คณะนิติศาสตร์	5 คน
คณะพยาบาลศาสตร์	7 คน
คณะแพทยศาสตร์	51 คน
คณะเภสัชศาสตร์	8 คน
คณะบริหารธุรกิจและนิเทศศาสตร์	2 คน
คณะวิทยาศาสตร์	3 คน
คณะวิทยาศาสตร์การแพทย์	5 คน
คณะวิศวกรรมศาสตร์	5 คน
คณะศิลปศาสตร์	18 คน
คณะสถาปัตยกรรมศาสตร์และศิลปกรรม	19 คน
คณะสหเวชศาสตร์	4 คน
คณะพลังงานและสิ่งแวดล้อม	2 คน

คณะ	จำนวน
คณะทันตแพทยศาสตร์	2 คน
คณะรัฐศาสตร์และสังคมศาสตร์	3 คน
วิทยาลัยการศึกษา	46 คน
คณะสาธารณสุขศาสตร์	14 คน
รวม	227 คน

3. เครื่องมือวิจัย

3.1 แบบทดสอบฉบับร่าง

เครื่องมือวิจัยครั้งนี้คือแบบทดสอบวัดระดับภาษาอังกฤษของนิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยา ฉบับร่างซึ่งพัฒนาขึ้นตามกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรป (CEFR) ระดับ B1 แบบเลือกตอบ (Multiple Choices) 4 ตัวเลือก จำนวน 4 ชุด แต่ละชุดประกอบด้วย 4 ส่วนได้แก่

ส่วนที่ 1 คำศัพท์	จำนวน 15 ข้อ
ส่วนที่ 2 การอ่าน	จำนวน 15 ข้อ
ส่วนที่ 3 ไวยากรณ์และการเขียน	จำนวน 15 ข้อ
ส่วนที่ 4 บทสนทนา	จำนวน 15 ข้อ
รวม	จำนวน 60 ข้อ

เนื้อหาของแบบทดสอบครอบคลุมคุณลักษณะหลักได้แก่ 1) ความสามารถเข้าใจประเด็นสำคัญจากสิ่งที่ได้อ่านจากเนื้อหาที่คุ้นเคยเกี่ยวกับการทำงาน การศึกษา การใช้เวลาว่างและอื่น ๆ 2) การรับมือกับสถานการณ์ต่าง ๆ ที่อาจเกิดขึ้นได้ระหว่างการท่องเที่ยวในที่ที่ใช้ภาษาอังกฤษในการสื่อสาร 3) ความสามารถในการติดต่อสื่อสารกันในหัวข้อที่ตนเอง ชื่นชอบและสนใจ 4) การบรรยายประสบการณ์ เหตุการณ์ ความฝัน การอธิบายเหตุผลสั้น ๆ รวมทั้งความคิดเห็นและ แผนการที่วางไว้

3.2 แบบประเมินค่าดัชนีความสอดคล้อง หรือ IOC สำหรับผู้เชี่ยวชาญ

แบบประเมินค่าดัชนีความสอดคล้อง (IOC) สำหรับผู้เชี่ยวชาญแบ่งออกเป็นสองส่วน ส่วนที่หนึ่งเป็นเกณฑ์การให้คะแนนสำหรับผู้เชี่ยวชาญ โดยมีเกณฑ์ดังต่อไปนี้

+1 = แน่ใจว่าข้อสอบข้อนั้นมีความสอดคล้องกับวัตถุประสงค์

0 = ไม่แน่ใจว่าข้อสอบข้อนั้นมีความสอดคล้องกับวัตถุประสงค์หรือไม่

-1 = แน่ใจว่าข้อสอบข้อนั้นไม่มีความสอดคล้องกับวัตถุประสงค์

ส่วนที่ส่งให้ผู้เชี่ยวชาญแสดงความคิดเห็นหรือให้คำแนะนำเพิ่มเติมเพื่อปรับปรุงแบบทดสอบ

4. ขั้นตอนการทำวิจัย

งานวิจัยนี้ดำเนินการในช่วงวันที่ 3 กุมภาพันธ์ 2563 – 31 มีนาคม 2564 ขั้นตอนการวิจัยแบ่งออกเป็น 4 ขั้นตอนตามรูปแบบการวิจัยและพัฒนา

4.1 ขั้นตอนที่ 1 การวิจัย (R_1)

1. กำหนดระดับความรู้ภาษาอังกฤษที่ต้องการวัดโดยอ้างอิงจากประกาศมหาวิทยาลัยพะเยา เรื่องนโยบายการยกระดับมาตรฐานภาษาอังกฤษสำหรับนิสิตมหาวิทยาลัยพะเยา พ.ศ. 2562 และ เรื่อง มาตรฐานการทดสอบวัดความรู้ภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี พ.ศ. 2562

2. ศึกษารายละเอียดและข้อกำหนดทักษะภาษาอังกฤษตามกรอบมาตรฐานการประเมิน ความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรป หรือ CEFR (Common European Framework of Reference) ระดับ B1 ตามที่ระบุไว้ในประกาศของมหาวิทยาลัยพะเยา

4.2 ขั้นตอนที่ 2 การพัฒนา (D_1)

1. กำหนดโครงสร้างแบบทดสอบตามข้อกำหนดทักษะภาษาอังกฤษตามกรอบมาตรฐานการ

ประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรป หรือ CEFR (Common European Framework of Reference) ระดับ B1

2. สร้างแบบทดสอบตามโครงสร้างที่กำหนด

3. ตรวจสอบคุณภาพเครื่องมือด้านความเที่ยงตรงของเนื้อหา โดยผู้เชี่ยวชาญ 3 คน ซึ่งมีความเชี่ยวชาญด้านภาษาอังกฤษ จำนวน 2 คน และการวัดและประเมินผล จำนวน 1 คน โดยใช้ดัชนีความสอดคล้อง (IOC)

4. คัดเลือกและปรับปรุงแบบทดสอบตามคำแนะนำของผู้เชี่ยวชาญ

4.3 ขั้นตอนที่ 3 การวิจัย (R₂)

1. นำแบบทดสอบที่ผ่านการพิจารณาจากผู้เชี่ยวชาญและปรับปรุงตามคำแนะนำแล้วไปทดลองใช้กับกลุ่มตัวอย่างจำนวน 227 คน

4.4 ขั้นตอนที่ 4 การพัฒนา (D₂)

1. วิเคราะห์หาความยากง่าย ค่าอำนาจจำแนก ค่าความเชื่อมั่นของแบบทดสอบ
2. คัดเลือกและปรับปรุงแบบทดสอบ

5. การวิเคราะห์ข้อมูล

5.1 การวิเคราะห์ความเที่ยงตรงเชิงเนื้อหา

วิเคราะห์คะแนนความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบที่ได้จากการประเมินดัชนีความสอดคล้อง (Index of Item-Objective Congruence – IOC) ของผู้เชี่ยวชาญ 3 คน โดยใช้สูตร (Rovinelli and Hambleton, 1977: 49–60 อ้างถึงใน สมชาย วรภิเษมกุล, 2554)

$$IOC = \frac{\sum R}{N}$$

เกณฑ์การพิจารณาค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับจุดประสงค์คือแบบทดสอบข้อที่มี IOC ตั้งแต่ 0.5 ขึ้นไปจะผ่านเกณฑ์และคัดเลือกไว้ ส่วนแบบทดสอบข้อที่มีค่า IOC ต่ำกว่า 0.5 จะนำไปแก้ไขปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญ

5.2 การวิเคราะห์แบบทดสอบรายข้อโดยใช้ค่าความยากง่ายและค่าอำนาจจำแนก

งานวิจัยนี้ได้วิเคราะห์ค่าความยากง่ายและค่าอำนาจจำแนกโดยใช้เทคนิค 27% ของจุดเพน เนื่องจากจำนวนกลุ่มตัวอย่างที่ทดลองทำแบบทดสอบมีจำนวนมากกว่า 100 คน (N = 227) แบบทดสอบ

ข้อที่ได้รับคัดเลือกไว้ต้องผ่านเกณฑ์ค่าความยากง่ายและค่าอำนาจจำแนกทั้งสองเกณฑ์ แบบทดสอบข้อที่ผ่านเกณฑ์ใดเกณฑ์หนึ่ง หรือไม่ผ่านทั้งสองเกณฑ์จะถูกตัดทิ้ง เกณฑ์พิจารณาแบบทดสอบรายข้อคือต้องมีค่าความยากง่ายอยู่ระหว่าง 0.20–0.80 และมีค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป โดยการแปลผลค่าความยากง่าย (สมชาย วรภิเษมกุล, 2554) มีดังนี้

0.81 – 1.00 เป็นข้อสอบที่ง่ายมาก ควรตัดทิ้ง หรือนำไปปรับปรุง

0.60 – 0.80 เป็นข้อสอบที่ค่อนข้างง่าย ใช้ได้ดี

0.40 – 0.59 เป็นข้อสอบที่ความยากง่ายปานกลางดีมาก

0.20 – 0.39 เป็นข้อสอบที่ค่อนข้างยาก ใช้ได้ดี

0.00 – 0.19 เป็นข้อสอบที่ยากมาก ควรตัดทิ้งหรือนำไปปรับปรุง

เกณฑ์การแปลค่าอำนาจจำแนก (ล้วน สายยศและอังคณา สายยศ, 2543) มีดังนี้

ค่าอำนาจจำแนก (r) .20 ถึง 1.00 = จำแนกได้

ค่าอำนาจจำแนก (r) -.19 ถึง +.19 = จำแนกไม่ได้

ค่าอำนาจจำแนก (r) -.20 ถึง -1.00 = จำแนกกลับ (ใช้ไม่ได้)

6. การวิเคราะห์ค่าความเชื่อมั่น

ทีมผู้วิจัยวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบโดยใช้วิธีของคูเดอร์ริชาร์ดสัน 20 (KR-20) เนื่องจากเป็นแบบทดสอบที่พัฒนาเป็นชนิดที่ให้คะแนน 0 หรือ 1 เท่านั้น และแต่ละข้อมีความยากง่ายไม่เท่ากันและมีแบบทดสอบมากกว่า 20 ข้อ (สมชาย วรจิเกษมกุล, 2554)

ผลการศึกษา

1. ผลการพัฒนาแบบทดสอบวัดระดับภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยา ตามกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรประดับ B1

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบทดสอบวัดระดับภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรีมหาวิทยาลัยพะเยาตามกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรป หรือ CEFR ระดับ B1 ทีมผู้วิจัยได้พัฒนาแบบทดสอบจำนวน 4 ชุด โดยแต่ละชุดประกอบด้วยเนื้อหา ดังนี้

ส่วนที่ 1 คำศัพท์	จำนวน 15 ข้อ
ส่วนที่ 2 การอ่าน	จำนวน 15 ข้อ
ส่วนที่ 3 ไวยากรณ์และการเขียน	จำนวน 15 ข้อ
ส่วนที่ 4 บทสนทนา	จำนวน 15 ข้อ

รวมชุดละ จำนวน 60 ข้อ

2. ผลการวิเคราะห์คุณภาพแบบทดสอบวัดระดับภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยา ตามกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพ ยุโรประดับ B1 ที่พัฒนาขึ้น

งานวิจัยนี้ได้วิเคราะห์คุณภาพแบบทดสอบด้วย 3 วิธีการ ได้แก่ การวิเคราะห์ความเที่ยงตรงเชิงเนื้อหา การวิเคราะห์ค่าความยากง่ายและค่าอำนาจจำแนก และการวิเคราะห์ค่าความเชื่อมั่น

ผลการวิเคราะห์คุณภาพแบบทดสอบมีรายละเอียดดังต่อไปนี้

1. ผลการวิเคราะห์ความเที่ยงตรงเชิงเนื้อหา

ตารางที่ 2 ผลการวิเคราะห์ค่าความเที่ยงตรงเชิงเนื้อหา

ชุดข้อสอบ	ช่วงค่าดัชนีความสอดคล้องระหว่างข้อคำถามและเนื้อหา ระดับ B1 (IOC) รายข้อ	จำนวนข้อที่ผ่านเกณฑ์ (≥ .50)	จำนวนข้อที่ไม่ผ่านเกณฑ์ (ต่ำกว่า < .50)	ค่าดัชนีความสอดคล้องระหว่างข้อคำถามและเนื้อหา ระดับ B1 (IOC) รวมทั้งชุด
ชุดที่ 1	0.33 – 1.00	58 ข้อ (96.67 %)	2 ข้อ (ข้อ 8 และ 26)	0.92
ชุดที่ 2	0.33 – 1.00	57 ข้อ (95 %)	3 ข้อ (ข้อ 6 7 และ 22)	0.93
ชุดที่ 3	0.33 – 1.00	58 ข้อ (96.67 %)	2 ข้อ (ข้อ 12 และ 30)	0.92
ชุดที่ 4	0.33 – 1.00	56 ข้อ (93.33 %)	4 ข้อ (ข้อ 24 26 29 และ 33)	0.93

ตารางที่ 1 นำเสนอข้อมูลการวิเคราะห์ความเที่ยงตรงเชิงเนื้อหาโดยใช้ดัชนีความสอดคล้อง (IOC) โดยผู้เชี่ยวชาญจำนวน 3 คน ได้แก่ ผู้เชี่ยวชาญทางด้านภาษาอังกฤษจำนวน 2 คน และด้านวัดผลและประเมินผลการศึกษาจำนวน 1 คน การพิจารณาคือแบบทดสอบต้องมีค่าดัชนีความสอดคล้องตั้งแต่ 0.50 ขึ้นไปจึงจะผ่านเกณฑ์ และข้อที่มีค่าดัชนีความสอดคล้องต่ำกว่า 0.50 จะต้องนำมาปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญจากตารางที่ 1 พบว่าระดับค่าดัชนีความสอดคล้องรายข้อของข้อสอบทั้ง 4 ชุดอยู่ระหว่าง 0.33 ถึง 1.00 และค่าดัชนีความสอดคล้องรวมทั้งชุดของแบบทดสอบทั้ง 4 ชุดอยู่ระหว่าง 0.92 ถึง 0.93 โดยข้อสอบชุดที่ 1 มีแบบทดสอบที่ผ่านเกณฑ์จำนวน 58 ข้อ (ร้อยละ 96.67) ไม่ผ่านเกณฑ์จำนวน 2 ข้อ คือข้อ 8 และ ข้อ 26 ค่าดัชนีความสอดคล้องรวมเท่ากับ 0.92 สำหรับแบบทดสอบชุดที่ 2 ผ่านเกณฑ์ 57 ข้อ (ร้อยละ 95) มีข้อที่ไม่ผ่านเกณฑ์จำนวน 3 ข้อ ได้แก่ ข้อ 6 ข้อ 7 และ ข้อ 22 ค่าดัชนีความสอดคล้องรวมทั้งชุดเท่ากับ 0.93 ในส่วนของแบบทดสอบชุดที่ 3 มีข้อสอบที่ผ่านเกณฑ์ 58 ข้อ (ร้อยละ 96.67) และพบข้อที่มีค่าความสอดคล้องต่ำกว่าเกณฑ์จำนวน 2 ข้อ ได้แก่ข้อ 12 และ ข้อ 30 ค่าดัชนีรวมคือ 0.92 และแบบทดสอบชุดที่ 4 มีข้อสอบผ่านเกณฑ์ 56 ข้อ (ร้อยละ 93.33) และพบข้อที่ไม่ผ่านเกณฑ์จำนวน 4 ข้อ ได้แก่ ข้อ 24 ข้อ 26 ข้อ 29 และ ข้อ 33 ค่าดัชนีความสอดคล้องเท่ากับ 0.93 กล่าวโดยสรุปคือค่าดัชนีความสอดคล้องรายข้อของแบบทดสอบส่วนใหญ่ผ่านเกณฑ์ที่กำหนดคิดเป็นร้อยละ 93.33 – 96.67 และค่าดัชนีความสอดคล้องรวมผ่านเกณฑ์ทั้ง 4 ชุด

2. ผลการวิเคราะห์ค่าความยากง่ายและค่าอำนาจจำแนก

ทีมผู้วิจัยได้วิเคราะห์ดัชนีค่าความยากง่ายและค่าอำนาจจำแนกโดยใช้เทคนิค 27 % ของจุ่งเตฟาน (Chung–The Fan) เนื่องจากจำนวนผู้ทำแบบทดสอบเกิน 100 คน (N = 228 คน) (ลิวัน สายยศ และอังคณา สายยศ, 2538) เกณฑ์การพิจารณาคือแบบทดสอบที่มีค่าความยากง่าย (P) อยู่ระหว่าง 0.20 ถึง 0.80 และควรมีค่าอำนาจจำแนก (r) ตั้งแต่ 0.20 ขึ้นไปจะถูกคัดเลือกไว้ (สมชาย วรภิเกษมกุล, 2554) ส่วนแบบทดสอบที่ไม่ผ่านเกณฑ์จะถูกคัดทิ้ง

2.1 ผลการวิเคราะห์แบบทดสอบรายข้อจำแนกตามค่าความยากง่ายและค่าอำนาจจำแนก

ตารางที่ 3 สรุปผลการวิเคราะห์แบบทดสอบด้วยวิธีคำนวณจากตารางของจุงเตฟาน
จำแนกตามดัชนีความยากง่ายและค่าอำนาจจำแนก

ข้อสอบ ชุดที่	ค่า ความยากง่าย (P)	ค่าอำนาจจำแนก (r)	ผลการพิจารณา	
			จำนวนข้อที่ คัดเลือกไว้	จำนวนข้อที่ ตัดทิ้ง
1	0.19 – 0.76	-0.05 ถึง 0.72	50 ข้อ (83.33 %)	10 ข้อ (ข้อ 8 12 14 16 21 22 24 27 28 และ 30) (16.67 %)
2	0.15 – 0.77	-0.09 ถึง 0.67	47 ข้อ (78.33 %)	13 ข้อ (ข้อ 2 3 11 12 14 17 18 21 24 25 28 29 และ 44) (21.67 %)
3	0.12 – 0.80	-0.27 ถึง 0.66	48 ข้อ (80 %)	12 ข้อ (ข้อ 4 8 12 13 17 19 25 34 35 54 58 และ 60) (20 %)
4	0.19 – 0.80	-0.07 ถึง 0.81	50 ข้อ (83.33 %)	10 ข้อ (ข้อ 3 6 8 9 12 15 20 23 33 และ 39) (16.67 %)

จากตารางที่ 3 ผลการวิเคราะห์แบบทดสอบรายข้อด้วยวิธีคำนวณจากตารางของ
จุงเตฟาน หรือเทคนิค 27% ของแบบทดสอบทั้ง 4 ชุด แต่ละชุดมีข้อคำถาม 60 ข้อ จากกลุ่ม
ตัวอย่างจำนวน 227 คนพบว่า ข้อสอบชุดที่ 1 มีค่าความยากง่าย (P) อยู่ระหว่าง 0.19 ถึง 0.76
และค่าอำนาจจำแนก (r) มีค่าอยู่ระหว่าง -0.05 ถึง 0.72 ทั้งนี้ข้อสอบที่ผ่านเกณฑ์จะต้องมี
ค่าความยากง่ายอยู่ระหว่าง 0.20 ถึง 0.80 และค่าอำนาจจำแนก ตั้งแต่ 0.20 ขึ้นไป และ
จากการวิเคราะห์พบว่าแบบทดสอบชุดที่ 1 ที่ผ่านเกณฑ์และทำการคัดเลือกไว้มีจำนวน 50 ข้อ

(ร้อยละ 83.33) ส่วนข้อที่ไม่ผ่านเกณฑ์และทำการตัดทิ้ง มีจำนวน 10 ข้อ (ร้อยละ 16.67) ได้แก่ ข้อ 8 12 14 16 21 22 24 27 28 และ 30

สำหรับแบบทดสอบชุดที่ 2 ข้อสอบมีค่าความยากง่าย (P) อยู่ระหว่าง 0.15 ถึง 0.77 และค่าอำนาจจำแนก (r) มีค่าอยู่ระหว่าง -0.09 ถึง 0.67 จากการวิเคราะห์แบบทดสอบทำให้ได้ข้อที่ผ่านเกณฑ์และทำการคัดเลือกไว้จำนวน 47 ข้อ (ร้อยละ 78.33) ส่วนข้อสอบที่ไม่ผ่านเกณฑ์และทำการตัดทิ้ง มีจำนวน 13 ข้อ (ร้อยละ 21.67) ได้แก่ข้อ 2 3 11 12 14 17 18 21 24 25 28 29 และข้อ 44

การวิเคราะห์แบบทดสอบชุดที่ 3 พบว่า แบบทดสอบมีค่าความยากง่าย (P) อยู่ระหว่าง 0.12 ถึง 0.80 และค่าอำนาจจำแนก (r) อยู่ระหว่าง -0.27 ถึง 0.66 มีแบบทดสอบข้อที่ผ่านเกณฑ์และทำการคัดเลือกไว้จำนวน 48 ข้อ (ร้อยละ 80) ส่วนข้อที่ไม่ผ่านเกณฑ์และทำการตัดทิ้งมีจำนวน 12 ข้อ (ร้อยละ 20) ได้แก่ข้อ 4 8 12 13 17 19 25 34 35 54 58 และ 60

แบบทดสอบชุดที่ 4 มีค่าความยากง่าย (P) อยู่ระหว่าง 0.19 ถึง 0.80 และค่าอำนาจจำแนก (r) มีค่าอยู่ระหว่าง -0.07 ถึง 0.81 และจากการวิเคราะห์ทำให้ได้ข้อที่ผ่านเกณฑ์และทำการคัดเลือกไว้ มีจำนวน 50 ข้อ (ร้อยละ 83.33) ส่วนแบบทดสอบที่ไม่ผ่านเกณฑ์และทำการตัดทิ้ง มีจำนวน 10 ข้อ (ร้อยละ 16.67) ได้แก่ข้อ 3 6 8 9 12 15 20 23 33 และ 39

3. ผลการวิเคราะห์ค่าความเชื่อมั่น

ทีมผู้วิจัยได้วิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบโดยใช้สูตร KR-20 และผลการวิเคราะห์มีดังต่อไปนี้

ตารางที่ 4 ผลการวิเคราะห์ค่าความเชื่อมั่น

ชุดแบบทดสอบ	ค่าความเชื่อมั่น
ชุดที่ 1	0.87
ชุดที่ 2	0.79
ชุดที่ 3	0.81
ชุดที่ 4	0.86

จากตารางที่ 4 พบว่าค่าความเชื่อมั่นของแบบทดสอบอยู่ระหว่าง 0.79 ถึง 0.87 โดยค่าความเชื่อมั่นของแบบทดสอบชุดที่ 1 เท่ากับ 0.87 ชุดที่ 2 เท่ากับ 0.79 ชุดที่ 3 เท่ากับ 0.81 และชุดที่ 4 เท่ากับ 0.86 จากผลการวิเคราะห์พบว่าแบบทดสอบผ่านเกณฑ์และมีความเชื่อมั่นทั้ง 4 ชุด เนื่องจากหากเป็นแบบทดสอบที่พัฒนาขึ้นควรมีความเชื่อมั่น 0.70 ขึ้นไป (Burns, Grove & Sutherland, 2017, 583)

อภิปรายและสรุปผล

อภิปรายผล

1. ค่าดัชนีความสอดคล้องรายข้อของข้อสอบทั้ง 4 ชุดอยู่ระหว่าง 0.33 ถึง 1.00 และค่าดัชนีความสอดคล้องรวมทั้งชุดของข้อสอบทั้ง 4 ชุดอยู่ระหว่าง 0.92 ถึง 0.93 ค่าดัชนีความสอดคล้องรายข้อของแบบทดสอบส่วนใหญ่ผ่านเกณฑ์ที่กำหนด นั่นคือมีค่าดัชนีตั้งแต่ 0.50 ขึ้นไป และค่าดัชนีความสอดคล้องรวมผ่านเกณฑ์ทุกชุด สำหรับข้อที่มีค่าดัชนีความสอดคล้องต่ำกว่า 0.5 ผู้วิจัยได้นำไปปรับปรุง ซึ่งเป็นไปตามสมชาย วรภิเกษมกุล (2554: 269) ที่ระบุว่าแบบทดสอบข้อที่มีค่าดัชนีความสอดคล้อง 0.5 ขึ้นไปสามารถคัดเลือกไว้ใช้ได้และข้อที่มีค่าดัชนีความสอดคล้องต่ำกว่า 0.5 ควรปรับปรุงหรือตัดทิ้ง

2. ค่าความยากง่ายและค่าอำนาจจำแนกของแบบทดสอบได้รับการวิเคราะห์โดยใช้เทคนิค 27 % ของจุงเตฟาน (Chung–The Fan) เนื่องจากจำนวนกลุ่มตัวอย่างที่ทดลองทำแบบทดสอบมีจำนวน 227 คน เกณฑ์การพิจารณาแบบทดสอบคือจะคัดเลือกข้อที่มีค่าความยากง่าย (P) อยู่ระหว่าง 0.20 ถึง 0.80 และมีค่าอำนาจจำแนก (r) ตั้งแต่ 0.20 ขึ้นไปไว้ และตัดข้อที่ไม่ผ่านเกณฑ์ทิ้ง จากการวิเคราะห์พบว่าแบบทดสอบชุดที่ 1 มีข้อที่ผ่านเกณฑ์และได้รับการคัดเลือกจำนวน 50 ข้อ แบบทดสอบชุดที่ 2 มีข้อที่ผ่านเกณฑ์จำนวน 47 ข้อ แบบทดสอบชุดที่ 3 มีข้อที่ผ่านเกณฑ์จำนวน 48 ข้อ และแบบทดสอบชุดที่ 4 มีข้อที่ผ่านเกณฑ์จำนวน 50 ข้อ เกณฑ์การพิจารณาความยากง่ายในการศึกษาครั้งนี้สอดคล้องกับสิรินธร ลินจินตวงศ์ (2547) ที่ได้ระบุว่าแบบทดสอบที่ดีต้องมีระดับความยากที่เหมาะสม กล่าวคือมีข้อที่ยาก ค่อนข้างยาก ง่าย และค่อนข้างง่ายประกอบกัน แบบทดสอบที่ดีควรมีระดับความยากตั้งแต่ 0.2 – 0.8 และค่าอำนาจจำแนกรายข้อของแบบทดสอบที่ดีควรเท่ากับหรือมากกว่า .20 ($r \geq 0.20$) หรือ .20 – 1.00

3. จากการวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบโดยใช้สูตร KR-20 พบว่า ค่าความเชื่อมั่นของแบบทดสอบอยู่ระหว่าง 0.79 ถึง 0.87 โดยแบบทดสอบชุดที่ 1 เท่ากับ 0.89 ชุดที่ 2 เท่ากับ 0.79 ชุดที่ 3 เท่ากับ 0.81 และชุดที่ 4 เท่ากับ 0.86 ดังนั้นแบบทดสอบจึงมีความเชื่อมั่นทั้ง 4 ชุดเนื่องจากหากเป็นแบบทดสอบที่พัฒนาขึ้นจึงควรมีค่าความเชื่อมั่นตั้งแต่ 0.70 ขึ้นไป (Burns, Grove & Sutherland, 2017, p. 583) เกณฑ์การพิจารณานี้สอดคล้องกับสิรินธร ลินจินดาวงค์ (2547) และสมชาย วรภิเษมกุล (2554) ซึ่งแนะนำว่า หากเป็นแบบทดสอบมาตรฐานทั่ว ๆ ไปควรมีความเชื่อมั่นเท่ากับ 0.8 แต่ถ้าเป็นแบบทดสอบที่สร้างและพัฒนาขึ้นควรมีความเชื่อมั่นอย่างน้อย 0.70 ขึ้นไป

สรุปผล

ทีมผู้วิจัยได้พัฒนาแบบทดสอบวัดระดับภาษาอังกฤษสำหรับนิสิตระดับปริญญาตรี มหาวิทยาลัยพะเยาตามกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรป หรือ CEFR ระดับ B1 จำนวน 4 ชุด แต่ละชุดแบ่งออกเป็น 4 ส่วน ได้แก่ คำศัพท์จำนวน 15 ข้อ การอ่านจำนวน 15 ข้อ ไวยากรณ์และการเขียนจำนวน 15 ข้อ และ บทสนทนาจำนวน 15 ข้อ รวมจำนวน 60 ข้อ จากนั้นได้วิเคราะห์คุณภาพแบบทดสอบโดยให้ผู้เชี่ยวชาญ 3 ท่าน แบ่งเป็นผู้เชี่ยวชาญในสาขาภาษาอังกฤษ 2 ท่าน และสาขาวัดและประเมินผลการศึกษา 1 ท่าน ประเมินความเหมาะสมโดยภาพรวมของแบบทดสอบ ในด้านเนื้อหา ความถูกต้องทางภาษา การจัดรูปแบบของแบบทดสอบ จำนวนข้อกับเวลาที่กำหนดให้ทำแบบทดสอบรวมทั้งความชัดเจนของคำสั่งในแต่ละส่วนของแบบทดสอบ นอกจากนี้ผู้เชี่ยวชาญทั้ง 3 ท่านได้ประเมินความเที่ยงตรงเชิงเนื้อหาโดยใช้ดัชนีความสอดคล้อง (IOC)

หลังจากได้รับการประเมินและปรับปรุงแบบทดสอบตามคำแนะนำของผู้เชี่ยวชาญแล้ว ทีมผู้วิจัยได้นำแบบทดสอบไปทดลองกับกลุ่มตัวอย่างซึ่งเป็นนิสิตจากสาขาวิชาต่าง ๆ ของมหาวิทยาลัยพะเยาจำนวน 227 คน เพื่อนำผลคะแนนไปวิเคราะห์หาค่าความยากง่าย และค่าอำนาจจำแนก รวมถึงค่าความเชื่อมั่นของแบบทดสอบ เพื่อคัดเลือกข้อที่มีคุณภาพไว้ และตัดแบบทดสอบข้อที่ไม่มีคุณภาพทิ้ง

ผลการวิเคราะห์คุณภาพแบบทดสอบสามารถสรุปได้ดังนี้

1. ความเที่ยงตรงเชิงเนื้อหา

ระดับค่าดัชนีความสอดคล้องรายข้อของข้อสอบทั้ง 4 ชุดอยู่ระหว่าง 0.33 ถึง 1.00 และค่าดัชนีความสอดคล้องรวมทั้งชุดของข้อสอบทั้ง 4 ชุดอยู่ระหว่าง 0.92 ถึง 0.93 ข้อสอบชุดที่ 1 มีแบบทดสอบที่ผ่านเกณฑ์จำนวน 58 ข้อ ค่าดัชนีความสอดคล้องรวมทั้งชุดเท่ากับ 0.92 แบบทดสอบชุดที่ 2 มีข้อที่ผ่านเกณฑ์จำนวน 57 ข้อ ค่าดัชนีความสอดคล้องรวมทั้งชุดเท่ากับ 0.93 แบบทดสอบชุดที่ 3 มีข้อที่ผ่านเกณฑ์จำนวน 58 ข้อ ค่าดัชนีรวมคือ 0.92 และแบบทดสอบชุดที่ 4 มีข้อที่ผ่านเกณฑ์จำนวน 56 ค่าดัชนีความสอดคล้องเท่ากับ 0.93 ดังนั้นค่าดัชนีความสอดคล้องรายข้อของแบบทดสอบส่วนใหญ่ผ่านเกณฑ์ที่กำหนด และค่าดัชนีความสอดคล้องรวมผ่านเกณฑ์ทุกชุด

2. ค่าความยากง่ายและค่าอำนาจจำแนก

จากการวิเคราะห์ดัชนีค่าความยากง่ายและค่าอำนาจจำแนกโดยใช้เทคนิค 27% ของจุงเตพาน (Chung-The Fan) โดยมีเกณฑ์การพิจารณาคือแบบทดสอบต้องมีค่าความยากง่าย (P) อยู่ระหว่าง 0.20 ถึง 0.80 และมีค่าอำนาจจำแนก (r) ตั้งแต่ 0.20 ขึ้นไปจะถูกคัดเลือกไว้ ส่วนแบบทดสอบที่ไม่ผ่านเกณฑ์จะถูกตัดทิ้งพบว่าแบบทดสอบชุดที่ 1 มีข้อที่ผ่านเกณฑ์และได้รับการคัดเลือกจำนวน 50 ข้อ ส่วนถูกตัดทิ้งจำนวน 10 ข้อ แบบทดสอบชุดที่ 2 มีข้อที่ผ่านเกณฑ์จำนวน 47 ข้อ และถูกตัดทิ้งจำนวน 13 ข้อ แบบทดสอบชุดที่ 3 มีข้อที่ผ่านเกณฑ์จำนวน 48 ข้อและถูกตัดทิ้งจำนวน 12 ข้อ ส่วนแบบทดสอบชุดที่ 4 มีข้อที่ผ่านเกณฑ์จำนวน 50 ข้อและถูกตัดทิ้งจำนวน 10 ข้อ

3. ค่าความเชื่อมั่น

การวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบโดยใช้สูตร KR-20 พบว่าค่าความเชื่อมั่นของแบบทดสอบอยู่ระหว่าง 0.79 ถึง 0.87 โดยแบบทดสอบชุดที่ 1 เท่ากับ 0.87 ชุดที่ 2 เท่ากับ 0.79 ชุดที่ 3 เท่ากับ 0.81 และชุดที่ 4 เท่ากับ 0.86 ดังนั้นแบบทดสอบจึงมีความเชื่อมั่นทั้ง 4 ชุด

ข้อเสนอแนะ

ข้อเสนอแนะสำหรับการนำผลการวิจัยไปใช้

1. ในการนำแบบทดสอบที่พัฒนาขึ้นไปใช้ ควรมีการอธิบายลักษณะแบบทดสอบให้แก่ผู้เข้ารับการทดสอบอย่างชัดเจน เช่น แบบทดสอบประกอบไปด้วยการทดสอบความรู้และทักษะด้านใดบ้าง จำนวนข้อของแบบทดสอบและเวลาที่กำหนดให้ รวมทั้งควรแนะนำให้ผู้เข้ารับการทดสอบอ่านคำอธิบายและคำสั่งในแบบทดสอบอย่างละเอียด

2. ควรเพิ่มแบบทดสอบในส่วนของทักษะการฟังและพูดเพื่อให้สามารถวัดความสามารถทางภาษาอังกฤษของผู้เข้ารับการทดสอบได้ครบทุกด้าน

ข้อเสนอแนะสำหรับการทำวิจัยครั้งต่อไป

1. ควรเปรียบเทียบคะแนนแบบทดสอบที่พัฒนาขึ้นกับกรอบมาตรฐานการประเมินความสามารถทางภาษาจากประเทศในกลุ่มสหภาพยุโรป หรือ CEFR ระดับ B1 โดยใช้วิธี Modified Angoff เพื่อศึกษาว่าผู้ทำแบบทดสอบต้องทำแบบทดสอบได้กี่เปอร์เซ็นต์จึงจะถือว่าอยู่ในระดับ B1

2. ควรศึกษาความคิดเห็นของผู้เข้ารับการทดสอบเพื่อนำข้อมูลไปปรับปรุงแบบทดสอบ

เอกสารอ้างอิง

จิตติรัตน์ แสงเลิศอุทัย. (2558). เครื่องมือที่ใช้ในการวิจัย (Research Instrument). **วารสารบัณฑิตศึกษามหาวิทยาลัยราชภัฏสกลนคร**, 12(58), 13–24.

เดชกุล มัทวานุกุล. (ม.ป.ป.). **การวิจัยและพัฒนา (Research and Development)**.

สืบค้น 15 มีนาคม 2564, จาก <http://www.curriculum-instruction.com/index.php/journal>.

สำนักงานคณะกรรมการการอุดมศึกษา. (2559, 12 เมษายน). **ประกาศคณะกรรมการการอุดมศึกษา เรื่อง นโยบายการยก ระดับมาตรฐานภาษาอังกฤษในสถาบันอุดมศึกษา**. กองยกระดับคุณภาพการจัดการศึกษาระดับอุดมศึกษา, จาก http://www.dqe.mhesi.go.th/front_home/Data%20Bhes_2559/04052559.pdf.

มหาวิทยาลัยพะเยา. (2562, 18 พฤษภาคม). ประกาศมหาวิทยาลัยพะเยา เรื่อง

นโยบายการยกระดับมาตรฐานภาษาอังกฤษ สำหรับนิสิตมหาวิทยาลัยพะเยา

พ.ศ.2562. ระบบบริการการศึกษา มหาวิทยาลัยพะเยา, จาก

[https://does.up.ac.th/public/storage/announce-](https://does.up.ac.th/public/storage/announce-detail/May2020/uiVEBilyia1JOynUfeZt.pdf)

[detail/May2020/uiVEBilyia1JOynUfeZt.pdf](https://does.up.ac.th/public/storage/announce-detail/May2020/uiVEBilyia1JOynUfeZt.pdf).

มหาวิทยาลัยพะเยา. (2562, 18 พฤษภาคม). ประกาศมหาวิทยาลัยพะเยา เรื่อง

มาตรฐานการทดสอบวัดระดับความรู้ภาษาอังกฤษสำหรับนิสิตระดับปริญญา

ตรี พ.ศ.2562. ระบบบริการการศึกษา มหาวิทยาลัยพะเยา, จาก

[https://does.up.ac.th/public/storage/announce-](https://does.up.ac.th/public/storage/announce-detail/March2021/TyWUNe1fZOgLyZvQFSf.pdf)

[detail/March2021/TyWUNe1fZOgLyZvQFSf.pdf](https://does.up.ac.th/public/storage/announce-detail/March2021/TyWUNe1fZOgLyZvQFSf.pdf).

ชื่นจิตต์ อธิวรกุล และจิรดา วุฒผยากร. (2561). การเทียบผลคะแนนแบบทดสอบสมิทธิ

ภาพภาษาอังกฤษแห่งมหาวิทยาลัยศรีนครินทรวิโรฒกับกรอบมาตรฐานอ้างอิง

สากลของยุโรป. *วารสารเทคโนโลยีสุรนารี*, 12(2), 69–84.

ณัฏฐกรณ์ หลาวทอง. (2547). รายงานการวิจัยเรื่อง การพัฒนาคลังข้อสอบวิชา 2702303

การวัดและการประเมินผลทางการศึกษาด้วยโปรแกรมคอมพิวเตอร์. กรุงเทพฯ:

จุฬาลงกรณ์มหาวิทยาลัย.

ล้วน สายยศและอังคณา สายยศ. (2543). **เทคนิคการวัดผลการเรียนรู้.** กรุงเทพฯ:

สุวีริยาสาส์น.

_____. (2538). **เทคนิคการวิจัยทางการศึกษา** (พิมพ์ครั้งที่ 5). กรุงเทพฯ: สุวีริยาสาส์น.

สุกัญญา ทองนาค, ศิริชัย กาญจนวาสี และสุพจน์ เกิดสุวรรณ. (2558). การพัฒนา

แบบทดสอบวัดสมรรถนะนักศึกษาครุศาสตร์/ศึกษาศาสตร์ตามมาตรฐานวิชาชีพครู

ที่มีการตรวจให้คะแนนแบบพหุภาค. *วารสารวิจัย มสส*, 9(1), 169–187.

สังวรรณ จัตตะโทกและคณะ. (2561). คุณภาพข้อสอบระดับชาติ (โอเน็ต) และแนวทาง

การพัฒนาคุณภาพการทดสอบ. *วารสารสุโขทัยธรรมาธิราช*, 31(2), 110–123.

สิรินธร สิ้นจินดาวงศ์. (2547). วิธีการวิเคราะห์ข้อสอบ. *วารสารศรีปทุมปริทัศน์*,

4(1), 21–33.

สมชาย วรภิเกษมสกุล. (2554). **ระเบียบวิธีการวิจัยทางพฤติกรรมศาสตร์และ**

สังคมศาสตร์. อุตรธานี: มหาวิทยาลัยราชภัฏอุตรธานี.

อลิสา วาณิชดี, สุพิมพ์ ศรีพันธุ์วรสกุล และนรินทร์ทิพย์ ทองศร. (2551). **การพัฒนาแบบทดสอบวัดความสามารถด้านการใช้ภาษาอังกฤษ โดยใช้สื่ออิเล็กทรอนิกส์ในระบบการศึกษาทางไกล**. กรุงเทพฯ: มหาวิทยาลัยสุโขทัยธรรมาธิราช.

Bachman, L. F. (1990). **Fundamental considerations in Language Testing**. Oxford: Oxford University Press.

Burns, N., Grove, S.K. & Sutherland, S. (2017). **Burn and Grove's the Practice of Nursing Research: Appraisal, synthesis, and generation of evidence** (8th Edition). Missouri: Elsevier.

Council of Europe. (n.d.) **Global scale – Table 1 (CEFR 3.3): Common Reference levels**. Council of Europe, Form <https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale>.

Lakhamja, R. & Tachom, K. (2022). Implementing Triangulation Teaching Techniques to Enhance English Speaking Skills for Thai Undergraduate Students. **Journal of Humanities and Social Sciences University of Phayao**, 10(2), 60–75.

Praimahaniyom, T. & Chairinkom, B. (2020). Developing GROUPS Model for Learning English Communication Skills: Kamphaeng Phet Rajabhat University, Maesot. **Journal of Humanities and Social Sciences University of Phayao**, 8(2), 16–42.