

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหาวิจัย

ตามที่รัฐมนตรีว่าการกระทรวงศึกษาธิการมีนโยบายยกระดับมาตรฐานภาษาอังกฤษในทุกหลักสูตรเพื่อพัฒนาการเรียนการสอนภาษาอังกฤษ ให้ผู้เรียนมีความรู้ความสามารถใช้ภาษาอังกฤษเป็นเครื่องมือศึกษาค้นคว้าองค์ความรู้ที่เป็นสากลและก้าวทันโลก สำนักงานคณะกรรมการการอุดมศึกษาแห่งชาติ กระทรวงศึกษาธิการ ปัจจุบัน คือ สำนักงานปลัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม จึงได้ออกประกาศคณะกรรมการการอุดมศึกษา เรื่อง นโยบายการยกระดับมาตรฐานภาษาอังกฤษในสถาบันอุดมศึกษา โดยให้สถาบันอุดมศึกษากำหนดนโยบายและเป้าหมายการยกระดับมาตรฐานภาษาอังกฤษในสถาบันอุดมศึกษา ในทุกหลักสูตรและทุกระดับการศึกษา เพื่อเป็นแนวทางในการพัฒนาทักษะการใช้ภาษาอังกฤษของนิสิตนักศึกษาให้เป็นบัณฑิตที่มีความพร้อมทางวิชาการ วิชาชีพ และทักษะการสื่อสารภาษาอังกฤษในระดับที่ใช้งานได้ (working knowledge) รวมทั้งมีการจัดทำแผนเพื่อดำเนินการให้เป็นไปตามนโยบายและเป้าหมายที่มีตัวชี้วัดและการประเมินผลที่ชัดเจน (สำนักงานคณะกรรมการการอุดมศึกษา, 2559)

แนวทางสำคัญของนโยบายนี้ คือ กำหนดให้นิสิตนักศึกษาทุกคนต้องเข้ารับการทดสอบความรู้ภาษาอังกฤษผ่านแบบทดสอบมาตรฐานที่พัฒนาโดยสถาบันเอง หรือแบบทดสอบอื่น ๆ ที่สามารถเทียบเคียงผลกับกรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป (Common European Framework of Reference for Languages: CEFR) หรือมาตรฐานอื่น ๆ เพื่อประเมินระดับความรู้ภาษาอังกฤษของนิสิตนักศึกษา (สำนักงานคณะกรรมการการอุดมศึกษา, 2559)

จากนโยบายดังกล่าว สถาบันอุดมศึกษาที่มีการพัฒนาแบบทดสอบวัดความรู้ภาษาอังกฤษ เช่น CU-TEP TU-GET (CBT) KU-EPT ได้มีการเทียบเคียงคะแนนสอบกับกรอบ CEFR นอกจากนี้ สถาบันอุดมศึกษาหลายแห่ง เช่น มหาวิทยาลัยเทคโนโลยีสุรนารี มหาวิทยาลัยศรีนครินทรวิโรฒ มหาวิทยาลัยเชียงใหม่ มหาวิทยาลัยสุโขทัยธรรมาธิราช ได้พัฒนาแบบทดสอบวัดความรู้ภาษาอังกฤษขึ้นมาให้สอดคล้องกับกรอบ CEFR เช่นกัน

ในส่วนของมหาวิทยาลัยสุโขทัยธรรมาธิราช เป็นมหาวิทยาลัยหนึ่งที่อยู่ภายใต้การกำกับดูแลของสำนักงานปลัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม ก็ได้ตอบสนองนโยบายการยกระดับมาตรฐานภาษาอังกฤษในสถาบันอุดมศึกษา โดยกำหนดแนวทางในการยกระดับมาตรฐานภาษาอังกฤษของนักศึกษาทั้งระดับปริญญาตรี และระดับบัณฑิตศึกษา สำหรับนักศึกษาระดับ

บัณฑิตศึกษาต้องสอบผ่านความรู้ภาษาอังกฤษ และต้องมีคะแนนทดสอบความรู้ภาษาอังกฤษตามเกณฑ์ใดเกณฑ์หนึ่งที่มหาวิทยาลัยกำหนด ส่วนนักศึกษาระดับปริญญาตรีที่สมัครเข้าศึกษาตั้งแต่ภาคต้น ปีการศึกษา 2560 เป็นต้นไป ทุกคนจะต้องผ่านการสอบวัดความรู้ภาษาอังกฤษเป็นเงื่อนไขหนึ่งในการสำเร็จการศึกษา หรือนักศึกษาสามารถนำผลการสอบภาษาอังกฤษจากศูนย์ทดสอบภาษาที่ได้รับมาตรฐานเทียบเคียง CEFR เช่น TOEFL IELTS CU-TEP TU-GET หรือสถาบันการศึกษาอื่นมายื่นต่อมหาวิทยาลัยภายในวัน เวลาที่กำหนด ซึ่งผลการสอบไม่จำกัดเกณฑ์มาตรฐาน/คะแนน/ระดับผลการสอบ ขึ้นอยู่กับดุลยพินิจของคณะกรรมการ หรือคณะทำงานยกระดับมาตรฐานภาษาอังกฤษของมหาวิทยาลัย ในกรณีที่นักศึกษาระดับบัณฑิตศึกษา และนักศึกษาระดับปริญญาตรีไม่มีผลการสอบวัดความรู้ภาษาอังกฤษ นักศึกษาสามารถเข้ารับการสอบวัดความรู้ภาษาอังกฤษด้วยแบบทดสอบที่มหาวิทยาลัยสุโขทัยธรรมมาธิราชพัฒนาขึ้นได้ โดยสามารถสมัครสอบได้ตามช่วงเวลาที่มหาวิทยาลัยกำหนด

แบบทดสอบวัดระดับความรู้ภาษาอังกฤษที่มหาวิทยาลัยสุโขทัยธรรมมาธิราชได้พัฒนาขึ้นเรียกว่า Sukhothai Thammathirat Open University English Proficiency Test (STOU-EPT) ซึ่งลักษณะแบบทดสอบที่พัฒนาขึ้นนั้นได้ออกแบบตามกรอบ CEFR ครอบคลุม 3 ทักษะ ได้แก่ (1) การฟัง (listening) มีข้อสอบทั้งหมด 25 ข้อ ประกอบด้วย ข้อสอบที่มีลักษณะแบบเลือกตอบ 3 ตัวเลือก โดยให้ผู้สอบฟังบทสนทนาสั้น ๆ แล้วตอบคำถาม จำนวน 8 ข้อ ส่วนข้อสอบที่ให้ผู้สอบฟังเสียงแล้วตอบคำถาม มี 3 ลักษณะ ได้แก่ แบบจับคู่ จำนวน 6 ข้อ แบบเติมคำ จำนวน 8 ข้อ และแบบถูก-ผิด จำนวน 5 ข้อ (2) โครงสร้างภาษา (structure) มีข้อสอบทั้งหมด 35 ข้อ มีลักษณะเป็นแบบเลือกตอบ 4 ตัวเลือก และ (3) การอ่าน (reading) มีข้อสอบทั้งหมด 40 ข้อ มีลักษณะเป็นแบบเลือกตอบ 4 ตัวเลือก สำหรับเกณฑ์การแปลความหมายคะแนนเมื่อเทียบระดับความสามารถตามกรอบ CEFR แต่ละทักษะ มีเกณฑ์การแปลความหมายคะแนน 3 ระดับ ได้แก่ การฟัง ระดับ A2 (1-8 คะแนน) ระดับ B1 (9-20 คะแนน) และระดับ B2 (21-25 คะแนน) โครงสร้างภาษา ระดับ A2 (1-10 คะแนน) ระดับ B1 (11-23 คะแนน) และระดับ B2 (24-35 คะแนน) และการอ่าน ระดับ A2 (1-7 คะแนน) ระดับ B1 (8-21 คะแนน) และระดับ B2 (22-40 คะแนน) โดยผู้พัฒนาแบบทดสอบและผู้กำหนดเกณฑ์การแปลความหมายคะแนนแบบทดสอบ STOU-EPT เป็นคณาจารย์แขนงวิชาภาษาอังกฤษ สาขาวิชาศิลปศาสตร์ มหาวิทยาลัยสุโขทัยธรรมมาธิราช ส่วนสำนักทะเบียนและวัดผลได้พัฒนารูปแบบการสอบวัดความรู้ภาษาอังกฤษ STOU-EPT จากการสอบในรูปแบบกระดาษเป็นการสอบผ่านระบบอิเล็กทรอนิกส์ เรียกว่า Sukhothai Thammathirat Open University English Proficiency Test (E-Testing) หรือ STOU-EPT (E-Testing) รวมทั้งดำเนินการจัดสอบวัดความรู้ภาษาอังกฤษ STOU-EPT ให้กับนักศึกษามหาวิทยาลัยสุโขทัยธรรมมาธิราชทั้งในระดับบัณฑิตศึกษา

ระดับปริญญาตรี ผู้เรียนโครงการสัมฤทธิ์บัตร และผู้สนใจทั่วไป (ศจี จิระโร และสุภมาส อังศุโชติ, 2563)

การสอบวัดความรู้ภาษาอังกฤษ STOU-EPT ผ่านระบบอิเล็กทรอนิกส์ เริ่มดำเนินการจัดสอบมาตั้งแต่ปี พ.ศ. 2560 จนถึงปัจจุบัน ทำให้ข้อสอบที่จัดเก็บไว้ในคลังข้อสอบถูกนำมาใช้ซ้ำเป็นระยะเวลานาน อย่างไรก็ตาม หลังการนำข้อสอบไปใช้จริง ยังไม่เคยมีการตรวจสอบคุณภาพข้อสอบในเชิงสถิติ เช่น ความยาก อำนาจจำแนกของข้อสอบแต่ละข้ออย่างเป็นระบบ ต่อมาในปี พ.ศ. 2563 ศจี จิระโร และ สุภมาส อังศุโชติ (2563) ได้ศึกษาเกี่ยวกับการตรวจสอบคุณภาพข้อสอบวัดสมิทิภาพในการใช้ภาษาอังกฤษของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ตามกรอบการอ้างอิงทางภาษาของสหภาพยุโรป ผลการศึกษา พบว่า มหาวิทยาลัยสุโขทัยธรรมมาธิราช มีกระบวนการพัฒนาแบบทดสอบโดยส่งผู้แทนคณะกรรมการออกข้อสอบเข้ารับการฝึกอบรมเกี่ยวกับการจัดทำแบบทดสอบมาตรฐานภาษาอังกฤษ ณ สถาบันที่มีความเชี่ยวชาญในประเทศสหราชอาณาจักร หลังจากนั้นจึงได้พัฒนาลักษณะเฉพาะของแบบทดสอบ (test specifications) ให้สอดคล้องกับกรอบ CEFR โดยมีผู้เชี่ยวชาญที่ทำหน้าที่เป็นวิทยากรฝึกอบรมให้ข้อเสนอแนะ และตรวจสอบความเหมาะสมอย่างรอบด้าน ในส่วนของโครงสร้างข้อสอบ ได้ผ่านการพิจารณาและให้ความเห็นชอบจากผู้เชี่ยวชาญ เพื่อให้มั่นใจว่าแบบทดสอบมีความสอดคล้องกับระดับความสามารถตามกรอบ CEFR อย่างแท้จริง นอกจากนี้มหาวิทยาลัยยังได้ดำเนินการตรวจสอบคุณภาพของข้อสอบในเชิงเนื้อหา (content validity) และความสอดคล้องกับกรอบ CEFR ก่อนนำไปใช้สอบจริง โดยอาศัยความร่วมมือจากนักวิชาการด้านการวัดและประเมินผลทางภาษาจากประเทศสหราชอาณาจักร รวมถึงการใช้เครื่องมือทางเทคโนโลยี ได้แก่ โปรแกรม Lexitutor เพื่อวิเคราะห์ความเหมาะสมของระดับภาษา และโปรแกรม AWL list เพื่อตรวจสอบความเหมาะสมของการใช้คำศัพท์เชิงวิชาการให้สอดคล้องกับระดับความสามารถที่ต้องการวัด ผลการศึกษา พบว่า แบบทดสอบฉบับที่ 1, 2, 3, และ 4 มีข้อสอบสูงกว่าระดับมาตรฐาน CEFR คิดเป็น 44%, 61%, 59%, และ 25% ตามลำดับ และต่ำกว่าระดับมาตรฐาน CEFR คิดเป็น 2%, 3%, 1%, 1% ตามลำดับ การตรวจสอบความยากของข้อสอบ (difficulty) ด้วยทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) พบว่า ข้อสอบส่วนใหญ่ในแบบทดสอบทั้ง 4 ฉบับ มีค่าความยากที่เหมาะสม ( $p = .20-.80$ ) การคำนวณค่าสัมประสิทธิ์ความเที่ยง (reliability) ด้วย KR-20 พบว่า การวิเคราะห์ KR-20 บ่งชี้ว่า ความน่าเชื่อถือของแบบทดสอบทั้ง 4 ชุด มีความแตกต่างกันอย่างมาก (0.35–0.89) โดยเฉพาะอย่างยิ่งทักษะการฟังของแบบทดสอบฉบับที่ 3 และฉบับที่ 4 แสดงให้เห็นความน่าเชื่อถือที่ต่ำ (.35 และ .53) ซึ่งให้เห็นข้อจำกัดในด้านคุณภาพของข้อสอบและความสามารถในการจำแนกข้อสอบ นอกจากนี้ การวิเคราะห์ความสัมพันธ์ระดับ CEFR ของ STOU-EPT กับ IELTS และ SWU-SET ด้วย Kendall's tau-b ในกลุ่มอาสาสมัคร 28 คน พบว่าระดับ B1 และระดับ B2 จากการสอบ STOU-EPT มีความสอดคล้องสูงกับ IELTS และ SWU-SET

ขณะที่ระดับ A2 มีความสอดคล้องต่ำ โดย STOU-EPT มีแนวโน้มที่จะประเมินทักษะในระดับที่สูงกว่าความสามารถที่แท้จริงของผู้สอบ โดยเฉพาะอย่างยิ่ง ช่วงคะแนนที่กว้างเกินไปสำหรับระดับ B1 อาจลดทอนความแม่นยำในการจำแนกประเภท เนื่องจากเพิ่มโอกาสในการจำแนกผู้สอบผิดพลาดบริเวณจุดตัดคะแนน ความเสี่ยงในการจำแนกผิดพลาดนี้ส่งผลกระทบต่ออย่างมากทั้งการตีความผลคะแนนและความยุติธรรมของการตัดสินใจที่อ้างอิงจากคะแนนสอบ โดยเฉพาะอย่างยิ่งเมื่อเชื่อมโยงกับผลลัพธ์ที่มีความสำคัญสูง เช่น การสำเร็จการศึกษา การจ้างงาน

ต่อมาในปี พ.ศ. 2567 ศุภรัตน์ อิงชาติเจริญ (2567) ได้ตรวจสอบคุณสมบัติทางจิตมิติของข้อสอบ STOU-EPT โดยใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ผลการศึกษา พบว่า ข้อสอบบางข้อมีค่าความยาก (b) และค่าอำนาจจำแนก (a) ไม่เป็นไปตามเกณฑ์ นอกจากนี้ ยังพบข้อสอบบางข้อที่มีปัญหา DIF (Differential Item Functioning) ซึ่งหมายถึงข้อสอบทำหน้าที่แตกต่างกันระหว่างกลุ่มผู้สอบที่มีคุณลักษณะแตกต่างกัน เช่น เพศ ภูมิภาค ประสิทธิภาพในการสอบ

จากข้อมูลดังกล่าวชี้ให้เห็นว่า การพัฒนาแบบทดสอบ STOU-EPT ที่ใช้อยู่ปัจจุบันยังมีช่องว่างสำคัญที่ควรได้รับการพัฒนาให้สมบูรณ์ในสองด้านหลัก ได้แก่ (1) การตรวจสอบความตรงของแบบทดสอบ (validity) และ (2) การกำหนดคะแนนจุดตัด (cut scores) แม้แบบทดสอบ STOU-EPT พัฒนาขึ้นโดยยึดตามกรอบ CEFR แต่กระบวนการตรวจสอบความสอดคล้องระหว่างข้อสอบกับกรอบ CEFR ยังไม่สมบูรณ์ โดยเฉพาะในด้านความตรง ซึ่งเป็นสิ่งจำเป็นต่อการยืนยันว่าแบบทดสอบสามารถวัดระดับความสามารถของผู้สอบได้อย่างถูกต้องตามมาตรฐานที่กำหนด นอกจากนี้ แบบทดสอบยังขาดกระบวนการกำหนดคะแนนจุดตัดอย่างเป็นระบบ ซึ่งควรดำเนินการด้วยวิธีมาตรฐานที่ได้รับการยอมรับ เช่น วิธีของแองกอฟท์ที่ได้รับการปรับปรุง แทนการใช้ดุลยพินิจของผู้พัฒนาแบบทดสอบ การขาดกระบวนการมาตรฐานอาจทำให้การจำแนกระดับความสามารถของผู้สอบคลาดเคลื่อน โดยเฉพาะระดับ B1 ซึ่งมีช่วงคะแนนกว้างเกินไป จึงเสี่ยงต่อความแม่นยำในการแปลผลคะแนน ยิ่งไปกว่านั้น ผลการศึกษาของ ศุภรัตน์ อิงชาติเจริญ (2567) พบว่า ข้อสอบบางข้อในแบบทดสอบ STOU-EPT มีปัญหา DIF ซึ่งบ่งชี้ว่า ข้อสอบบางข้ออาจทำหน้าที่แตกต่างกันระหว่างกลุ่มผู้สอบที่มีคุณลักษณะแตกต่างกัน เช่น เพศ ภูมิภาค หรือ ประสิทธิภาพในการสอบ ส่งผลให้เกิดความลำเอียง (bias) และความไม่เป็นธรรมในการวัดระดับความสามารถทางภาษาอังกฤษของผู้สอบบางกลุ่ม นอกจากนี้ พบว่า ค่าพารามิเตอร์ของข้อสอบบางข้อ เช่น ค่า a และ b ตามทฤษฎีการตอบสนองข้อสอบ อาจส่งผลต่อความแม่นยำในการแปลผลคะแนน โดยเฉพาะในระดับ B1 ที่มีช่วงคะแนนกว้าง และมีความหลากหลายในระดับความสามารถของผู้สอบ ซึ่งอาจทำให้การจำแนกระดับความสามารถขาดความชัดเจน และไม่สะท้อนความสามารถของผู้สอบอย่างแท้จริง

การการแปลผลคะแนนที่คลาดเคลื่อนหรือไม่สะท้อนความสามารถที่แท้จริงของผู้สอบอาจส่งผลกระทบต่อในวงกว้างได้ เนื่องจากผู้สอบวัดความรู้ภาษาอังกฤษ STOU-EPT ส่วนใหญ่มีเป้าหมายสำคัญที่นำผลการสอบไปใช้เป็นหลักฐานอ้างอิงระดับความรู้ภาษาอังกฤษสำหรับสมัครเข้าศึกษาต่อขอสำเร็จการศึกษา ฯลฯ โดยเฉพาะการสมัครเข้าศึกษาต่อระดับบัณฑิตศึกษา มีความจำเป็นอย่างยิ่งที่ผู้เข้าศึกษาจะต้องมีความรู้ความสามารถในการใช้ภาษาอังกฤษในระดับที่เหมาะสมตามข้อกำหนดของสถาบันอุดมศึกษานั้น ๆ เนื่องจากต้องศึกษาค้นคว้าหาความรู้จากตำรา เอกสาร และ/หรืองานวิจัยที่เป็นภาษาอังกฤษ เพื่อประกอบการเรียนการสอน และต่อยอดองค์ความรู้ในระดับสากลด้วยตัวเอง รวมทั้งสามารถใช้ภาษาอังกฤษสื่อสารทางวิชาการ วิชาชีพได้ในระดับหนึ่ง ดังนั้น ผลการสอบจึงควรอธิบายระดับความรู้ภาษาอังกฤษของผู้สอบอย่างแท้จริง หากกำหนดเกณฑ์การแปลผลคะแนนไม่เหมาะสม เช่น คะแนนต่ำหรือสูงกว่าความสามารถจริงในการใช้ภาษาของผู้สอบในแต่ละระดับความสามารถ อาจส่งผลกระทบต่อผู้สอบได้ โดยผลกระทบที่อาจเกิดขึ้นคือ คะแนนที่กำหนดไว้ทำให้คนผ่านเกณฑ์ทั้ง ๆ ที่เขาไม่ควรจะสอบผ่าน ทางกลับกันทำให้คนไม่ผ่านเกณฑ์ทั้ง ๆ ที่เขาควรจะสอบผ่าน อันเนื่องมาจากคลาดเคลื่อนจากการกำหนด เกณฑ์การแปลผล

จากการศึกษาค่าสถิติข้อสอบ STOU-EPT ที่ใช้สอบมาแล้ว พบว่า มีข้อสอบบางข้อยากเกินไป และจำแนกความสามารถของผู้สอบได้ค่อนข้างน้อย สำนักทะเบียนและวัดผลร่วมกับคณาจารย์แขนงวิชาภาษาอังกฤษ สาขาวิชาศิลปศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช ปรับปรุงข้อสอบที่อยู่ในคลัง 400 ข้อ และพัฒนาข้อสอบใหม่เพิ่มอีก 400 ข้อ เพื่อพัฒนาข้อสอบ STOU-EPT ให้มีคุณภาพ และเพิ่มปริมาณข้อสอบในคลังให้เพียงพอสำหรับการจัดฉบับแบบทดสอบ รองรับนักศึกษาของมหาวิทยาลัยสุโขทัยธรรมาธิราช รวมถึงผู้สนใจทั่วไปซึ่งมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่องในแต่ละปี อย่างไรก็ตาม ปัจจุบันการพัฒนาและใช้ข้อสอบ STOU-EPT ยังขาดกระบวนการตรวจสอบความตรงอย่างเป็นระบบเพื่อยืนยันว่าแบบทดสอบสามารถวัดความรู้ภาษาอังกฤษของผู้สอบได้ถูกต้องและแม่นยำ อีกทั้งการแปลผลคะแนน ไม่ได้อิงกับแนวทางมาตรฐานระดับสากลซึ่งอาจส่งผลให้เกิดความคลาดเคลื่อนในการตีความผลสอบ ด้วยเหตุนี้ การตรวจสอบความตรงของข้อสอบ STOU-EPT จึงเป็นสิ่งจำเป็น เพื่อให้แน่ใจว่าแบบทดสอบวัดความสามารถของผู้สอบได้สอดคล้องกับกรอบ CEFR และรองรับการใช้งานในบริบทสากลได้อย่างมีประสิทธิภาพ

การปรับปรุงและพัฒนาข้อสอบใหม่ STOU-EPT ให้มีคุณภาพมีมาตรฐานเป็นที่ยอมรับ จำเป็นต้องมีกระบวนการตรวจสอบคุณภาพของข้อสอบก่อนนำไปใช้ โดยการตรวจสอบความตรงของแบบทดสอบ STOU-EPT ว่าสอดคล้องกับกรอบ CEFR หรือไม่ และมีการกำหนดมาตรฐาน (standards setting) เพื่อกำหนดเกณฑ์ในการตัดสินความสามารถของผู้สอบ เรียกว่า คะแนนจุดตัด

ด้วยเทคนิควิธีที่เหมาะสม และน่าเชื่อถือ อันจะทำให้การตีความหมายผลการสอบมีมาตรฐาน และสามารถใช้อ้างอิงได้อย่างถูกต้อง

การกำหนดมาตรฐาน เป็นสิ่งสำคัญสำหรับกระบวนการพัฒนาแบบทดสอบที่ดีที่สุด ซึ่งต้องดำเนินการร่วมกับกระบวนการอื่น ๆ เพื่อให้มั่นใจว่าผลการสอบมีประโยชน์ และสามารถอธิบายข้อโต้แย้งต่าง ๆ ได้ (Bejar, 2008) การกำหนดมาตรฐาน เป็นการดำเนินการเพื่อกำหนดคะแนนจุดตัดที่มีความเหมาะสมและมีความน่าเชื่อถือ สำหรับใช้เป็นเกณฑ์ในการแบ่งคะแนนออกเป็น 2 กลุ่ม หรือมากกว่า 2 กลุ่ม ที่สะท้อนให้เห็นถึงความสามารถที่แตกต่างกัน วิธีการกำหนดจุดตัดมีหลากหลายวิธี เช่น วิธี Angoff วิธี Bookmark วิธี Ebel วิธี Nedelsky โดยวิธีการกำหนดคะแนนจุดตัดแบบทดสอบแบบเลือกตอบ มักนิยมใช้วิธีการของแองกอฟ (Angoff) วิธีการนี้ถูกพัฒนาขึ้นตั้งแต่ ปี ค.ศ. 1970 โดย William H. Angoff และมีนักวิชาการหลายท่านได้ดัดแปลงวิธีการกำหนดคะแนนจุดตัดของแองกอฟ และแยกออกเป็นขั้นตอนต่าง ๆ เพื่อให้เหมาะสมกับบริบทที่หลากหลาย (Plake & Cizek, 2012) ภายใต้อ้างอิง แองกอฟใช่/ไม่ใช่ (Yes/No Angoff) แองกอฟที่ได้รับการปรับปรุง (Modified Angoff) และแองกอฟปรับขยาย (Extended Angoff) นอกจากนี้ จุดเด่นของการกำหนดคะแนนจุดตัดของแองกอฟ คือ มีขั้นตอนดำเนินการที่ง่าย และรวดเร็ว เปิดโอกาสให้ผู้เชี่ยวชาญสามารถทบทวนผลกำหนดคะแนนจุดตัดได้ ประมาณ 3 รอบ ตลอดจนผู้เชี่ยวชาญสามารถอภิปรายและแลกเปลี่ยนความคิดเห็นระหว่างผู้เชี่ยวชาญในแต่ละรอบได้ จึงทำให้การกำหนดคะแนนจุดตัดสอดคล้องกันมากขึ้น

การเทียบคะแนนแบบทดสอบวัดความรู้ภาษาอังกฤษกับกรอบ CEFR นิยมใช้วิธีแองกอฟที่ได้รับการปรับปรุง (Modified Angoff Method) โดยอาศัยผู้เชี่ยวชาญพิจารณาความสามารถของผู้สอบในแต่ละระดับ เช่น Tannenbaum and Baron (2011) ใช้วิธีนี้เทียบคะแนน TOEFL ITP กับ CEFR โดยใช้ผู้เชี่ยวชาญ 16 คน กำหนดคะแนนจุดตัดของข้อสอบแต่ละข้อในระดับ A2 B1 B2 และ C1 มี 3 ทักษะ ได้แก่ การฟัง โครงสร้างและการเขียน และการอ่าน ในทำนองเดียวกัน Baron and Papageorgiou (2016) ใช้วิธีแองกอฟที่ได้รับการปรับปรุง กำหนดคะแนนจุดตัด TOEFL Junior Standard และ TOEFL Junior Comprehensive เพื่อจำแนกผู้เรียนในระดับมัธยมศึกษาที่ต้องการ ESL Support และ No ESL Support มี 3 ทักษะ ได้แก่ ความเข้าใจในการฟัง ความเข้าใจรูปแบบและความหมายของภาษา และความเข้าใจในการอ่าน นอกจากนี้ Papageorgiou et al. (2019) ได้เทียบคะแนน TOEFL iBT กับ China's Standards of English Language Ability (CSE) ซึ่งเป็นมาตรฐานที่ใช้ในประเทศจีน แบ่งเป็น 9 ระดับความสามารถ ตั้งแต่ขั้นพื้นฐานจนถึงขั้นสูง และใช้วิธีแองกอฟที่ได้รับการปรับปรุง กำหนดคะแนนจุดตัด 3 ทักษะ ได้แก่ การฟัง การพูด และการเขียน สำหรับประเทศไทย ชื่นจิตต์ อธิวรกุล และ จิรดา วุฑฒยากร (2561) ได้เทียบคะแนนแบบทดสอบ SWU-SET กับ CEFR ในระดับ A2 B1 และ B2 แต่ละระดับ มี 5 ทักษะ ได้แก่ การฟัง คำศัพท์ การใช้ภาษา โครงสร้างภาษา และการอ่าน จากงานวิจัยข้างต้น ชี้ให้เห็นว่า การเทียบคะแนนแบบทดสอบวัด

ความสามารถทางภาษาอังกฤษกับ CEFR นิยมใช้วิธีแองกอฟที่ได้รับการปรับปรุง โดยมีการกำหนดคะแนนจุดตัดแบบแยกตามทักษะ ได้แก่ การฟัง การพูด การอ่าน และการเขียน อย่างไรก็ตาม ไม่พบงานวิจัยที่ใช้วิธีแองกอฟที่ได้รับการปรับปรุงกำหนดคะแนนจุดตัดแบบรวมทุกทักษะ ซึ่งเป็นประเด็นที่ต้องศึกษาเพิ่มเติมเพื่อพัฒนาเกณฑ์ที่สะท้อนความสามารถทางภาษาของผู้สอบได้อย่างครอบคลุม

จากความสำคัญและปัญหาที่นำเสนอไปแล้วนั้น จึงมีความสำคัญและความจำเป็นที่ต้องเพิ่มเติมกระบวนการพัฒนาแบบทดสอบ STOU-EPT ให้มีคุณภาพและมีมาตรฐานเป็นที่ยอมรับ โดยการศึกษาความสอดคล้องระหว่างแบบทดสอบ STOU-EPT กับกรอบ CEFR ด้วยวิธีของเว็บซึ่งเป็นวิธีที่ได้รับการยอมรับอย่างแพร่หลายในการประเมินความสอดคล้องระหว่างแบบทดสอบกับมาตรฐานการเรียนรู้ ซึ่งครอบคลุมด้านเนื้อหา ความลึกของความรู้ ความกว้างของความรู้ ความสมดุล และแหล่งที่มาของความท้าทาย (Webb, 1997; Martone & Sireci, 2009) ทั้งนี้ วิธีดังกล่าวได้นำไปใช้ในงานวิจัยที่หลากหลาย เช่น Roach and Elliott (2004), Roach, Elliott, and Webb (2005), Roach, Wixson, and Talapatra (2010), Forte (2017), Traynor et al. (2020) ซึ่งต่างยืนยันถึงประสิทธิภาพวิธีของเว็บในการประเมินความสอดคล้องของแบบทดสอบอย่างเป็นระบบและเป็นธรรม รวมทั้งเทียบเคียงคะแนนแบบทดสอบ STOU-EPT กับกรอบ CEFR โดยใช้การกำหนดมาตรฐาน เพื่อกำหนดคะแนนจุดตัดที่สามารถจำแนกผู้สอบว่ารอบรู้ และไม่รอบรู้ตามกรอบ CEFR ตลอดจนลดความคลาดเคลื่อนในการแบ่งคะแนนจุดตัดให้น้อยที่สุด Berk (1986) กล่าวว่า จุดแบ่งคะแนนจุดตัด มี 2 ส่วน ได้แก่ ส่วนที่เป็นคะแนนของผู้สอบที่จัดว่าเป็นผู้รอบรู้ และไม่รอบรู้ เนื่องจากคะแนนจุดตัดเป็นค่าที่ได้จากการสังเกต ฉะนั้น การกำหนดคะแนนจุดตัดจึงมีความคลาดเคลื่อนมาตรฐานเสมอ ความคลาดเคลื่อนที่เกิดขึ้น คือ การกำหนดคะแนนจุดตัดสูงเกินไปทำให้ผู้สอบที่มีความรอบรู้ที่แท้จริง ถูกจัดประเภทไม่มีความรอบรู้ และกำหนดคะแนนจุดตัดต่ำเกินไปทำให้ผู้สอบที่ไม่มีความรอบรู้อย่างแท้จริง ถูกจัดประเภทมีความรอบรู้ ดังนั้น เพื่อให้การกำหนดคะแนนจุดตัดมีความถูกต้อง เหมาะสม และลดความคลาดเคลื่อนในการแบ่งคะแนนจุดตัดให้น้อยที่สุด รวมทั้งทำให้คะแนนผลการสอบมีความน่าเชื่อถือและเป็นที่ยอมรับในวงกว้าง และผู้สอบสามารถนำผลการสอบไปใช้เป็นหลักฐานเพื่อประกอบการเข้าศึกษาต่อ การสำเร็จการศึกษา ฯลฯ ผู้วิจัยจึงสนใจเทียบเคียงคะแนนแบบทดสอบ STOU-EPT กับกรอบ CEFR โดยใช้วิธีการของแองกอฟที่ได้รับการปรับปรุง ซึ่งเป็นวิธีที่ใช้ง่าย ได้รับความนิยมในการกำหนดมาตรฐานเพื่อหาคะแนนจุดตัดอย่างกว้างขวาง และเหมาะสมกับลักษณะของข้อสอบ STOU-EPT ที่ให้คะแนนแบบ 2 ค่า (ตอบถูกให้ 1 ตอบผิดให้ 0) และการกำหนดคะแนนจุดตัดในการวิจัยครั้งนี้ มี 2 แบบ ได้แก่ (1) คะแนนจุดตัดรวม (ไม่แยกทักษะ) และ (2) คะแนนจุดตัดแยกตามทักษะ ได้แก่ การฟัง โครงสร้างภาษา และการอ่าน ซึ่งการกำหนดคะแนนจุดตัดแยกตามทักษะ นับว่ามีประโยชน์ต่อผู้สอบอย่างมาก เนื่องจากผู้สอบได้ทราบจุดเด่นหรือจุดที่ควรพัฒนาแต่ละทักษะ รวมทั้ง สถาบันอุดมศึกษาสามารถนำผลการสอบไปใช้ประโยชน์ในการ

พัฒนาการเรียนการสอนภาษาอังกฤษให้กับผู้สอบเพื่อพัฒนาผู้สอบให้มีความรู้ทางภาษาอังกฤษตามเกณฑ์ที่กำหนดได้

## 1.2 วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ของการวิจัย ดังนี้

1. เพื่อศึกษาความสอดคล้องระหว่างแบบทดสอบวัดระดับความรู้ภาษาอังกฤษของมหาวิทยาลัยสุโขทัยธรรมิกราชกับกรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรปด้วยวิธีของเว็บ
2. เพื่อเทียบเคียงคะแนนแบบทดสอบวัดระดับความรู้ภาษาอังกฤษของมหาวิทยาลัยสุโขทัยธรรมิกราชกับกรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรปด้วยวิธีการของแองกอฟที่ได้รับการปรับปรุง

## 1.3 ขอบเขตของการวิจัย

การวิจัยนี้มีขอบเขตของการวิจัย ดังนี้

1. การวิจัยนี้ผู้วิจัยศึกษาความสอดคล้อง และเทียบเคียงคะแนนแบบทดสอบ STOU-EPT กับกรอบ CEFR ใน 3 ทักษะ ได้แก่ การฟัง โครงสร้างภาษา และการอ่าน แต่ละทักษะแบ่งออกเป็น 3 ระดับความสามารถตามกรอบ CEFR ได้แก่ ระดับ A2 B1 และ B2
2. การศึกษาความสอดคล้องในครั้งนี้ มุ่งประเมินความสอดคล้องด้านเนื้อหา (content focus) ระหว่างแบบทดสอบ STOU-EPT กับกรอบ CEFR ด้วยวิธีของเว็บ (Webb, 1997) โดยพิจารณาจาก 3 เกณฑ์หลัก ได้แก่ (1) ความสอดคล้องของเนื้อหา (categorical occurrence) (2) ความสอดคล้องด้านความลึกของความรู้ (depth of knowledge: DOK) แบ่งออกเป็น 4 ระดับ ได้แก่ ระดับ 1 ความจำ (recall) ระดับ 2 การมีทักษะและมโนทัศน์ (skills and concepts) ระดับ 3 การคิดเชิงกลยุทธ์ (strategic thinking) และระดับ 4 การคิดหลายขั้นตอน (extended thinking) และ (3) ความสอดคล้องด้านความสมดุล (balance of representation) อย่างไรก็ตาม 2 เกณฑ์ที่ไม่ได้นำมาใช้ ได้แก่ ความกว้างของความรู้ (range of knowledge correspondence) เนื่องจากแบบทดสอบ STOU-EPT พัฒนาขึ้นโดยอิงกับกรอบ CEFR ซึ่งมีการกำหนดขอบเขตของเนื้อหาไว้อย่างชัดเจน ดังนั้น จึงไม่จำเป็นต้องศึกษาความสอดคล้องในมิตินี้ ส่วนแหล่งที่มาของความท้าทาย (source of challenge) เป็นเกณฑ์ที่ใช้ตรวจสอบว่าข้อสอบมีความยากเกินความจำเป็นอันเนื่องมาจากปัจจัยอื่นที่ไม่เกี่ยวข้องกับทักษะที่ต้องการวัด เช่น อคติทางวัฒนธรรม หรือความรู้เฉพาะก็ไม่นำมาใช้เช่นกัน เนื่องจากแบบทดสอบ STOU-EPT พัฒนาขึ้นโดยอิงกับกรอบ CEFR โดยตรงทำให้น้ำหนักเนื้อหา โครงสร้าง และองค์ประกอบต่าง ๆ ของข้อสอบสอดคล้องกับกรอบ CEFR อยู่แล้ว ซึ่งช่วยลดโอกาสที่ข้อสอบจะเกิดอคติที่ไม่เกี่ยวข้องกับการวัดความสามารถทางภาษา



3. แบบทดสอบ STOU-EPT ชุดที่ใช้ในการศึกษาความสอดคล้อง และการเทียบเคียงคะแนน กับกรอบ CEFR เป็นแบบทดสอบ STOU-EPT ที่พัฒนาขึ้นใหม่ ในปี พ.ศ. 2566 ตามลักษณะเฉพาะ ของแบบทดสอบ (test specifications) ในแต่ละทักษะตามกรอบ CEFR ระดับ A2 B1 และ B2

#### 1.4 นิยามเชิงปฏิบัติการ

แบบทดสอบวัดระดับความรู้ภาษาอังกฤษของมหาวิทยาลัยสุโขทัยธรรมมาธิราช (STOU-EPT) หมายถึง แบบทดสอบที่ใช้วัดความสามารถทางภาษาอังกฤษที่มหาวิทยาลัยสุโขทัยธรรมมาธิราช พัฒนาขึ้น เพื่อใช้วัดระดับความสามารถทางภาษาอังกฤษของผู้เรียนโครงการสัมฤทธิ์บัตร นักศึกษา ระดับปริญญาตรี และนักศึกษาระดับบัณฑิตศึกษาของมหาวิทยาลัยสุโขทัยธรรมมาธิราช รวมทั้งบุคคล ทั่วไปที่ต้องการสอบวัดระดับความสามารถทางภาษาอังกฤษของตนเอง โดยครอบคลุม 3 ทักษะ ได้แก่ การฟัง โครงสร้างภาษา และการอ่าน แต่ละทักษะถูกออกแบบตามกรอบ CEFR ในระดับ A2 B1 และ B2

กรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป (The Common European Framework of Reference for Languages: CEFR) หมายถึง เนื้อหาหรือมาตรฐานที่ใช้ประเมิน ความสามารถทางภาษาอังกฤษ แบ่งออกเป็น 3 ระดับ ได้แก่ A2 B1 และ B2 ซึ่งมาตรฐานนี้ให้ ผู้เชี่ยวชาญใช้สำหรับประเมินความสอดคล้องตามวิธีของเว็บ

การประเมินความสอดคล้องด้านเนื้อหา (content focus) หมายถึง กระบวนการวิเคราะห์เพื่อ ตรวจสอบว่าแบบทดสอบวัดระดับความรู้ภาษาอังกฤษของมหาวิทยาลัยสุโขทัยธรรมมาธิราช (STOU-EPT) กับกรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป (CEFR) มีความตรงกับมาตรฐาน CEFR หรือไม่ โดยพิจารณาความสอดคล้องของเนื้อหา (categorical occurrence) ควบคู่กับความลึกของ ความรู้ (depth of knowledge: DOK) และความสมดุล (balance of representation) เพื่อให้มั่นใจ ว่าแบบสอบ STOU-EPT วัดความสามารถทางภาษาอังกฤษของผู้สอบได้ตรงและยุติธรรม

ความสอดคล้องของเนื้อหา หมายถึง การประเมินเนื้อหาข้อสอบ STOU-EPT แต่ละข้อ ว่าสอดคล้องกับกรอบมาตรฐาน CEFR ระดับใด โดยใช้การพิจารณาของผู้เชี่ยวชาญ ซึ่งประเมินว่า ข้อสอบ STOU-EPT แต่ละข้อตรงกับตัวชี้วัดของกรอบมาตรฐาน CEFR ระดับ A2 B1 หรือ B2

ความสอดคล้องด้านความลึกของความรู้ หมายถึง การประเมินเนื้อหาข้อสอบ STOU-EPT แต่ละข้อว่าวัดความซับซ้อนของการคิดของผู้สอบที่จะทำข้อสอบให้ถูกหรือผ่านระดับใด โดยใช้การ พิจารณาของผู้เชี่ยวชาญ ซึ่งระดับความซับซ้อนของการคิด แบ่งออกเป็น 4 ระดับ ได้แก่ ความจำ การมีทักษะและมโนทัศน์ การคิดเชิงกลยุทธ์ และการคิดหลายขั้นตอน โดยข้อสอบแต่ละข้อสามารถ วัดระดับความซับซ้อนของการคิดได้เพียงระดับเดียว

ความสอดคล้องด้านความสมดุล หมายถึง การประเมินว่าข้อสอบ STOU-EPT ในทักษะ การฟัง โครงสร้างภาษา และการอ่าน มีกระจายตัวของข้อสอบอย่างเท่าเทียมตามมาตรฐาน CEFR

ระดับ A2 B1 และ B2 หรือทำให้ทราบว่ามาตรฐานใดที่ออกข้อสอบมากกว่ามาตรฐานอื่น ๆ ถ้ามี แสดงว่า แบบทดสอบ STOU-EPT วัดมาตรฐานใดมาตรฐานหนึ่งมากกว่ามาตรฐานอื่น ๆ ซึ่งถือว่าไม่สมดุล การประเมินความสอดคล้องด้านความสมดุล ไม่ใช้การตัดสินโดยผู้เชี่ยวชาญ แต่ใช้การคำนวณดัชนีความสมดุล (balance index) หรือ Q

วิธีการกำหนดมาตรฐาน (standard setting) หมายถึง กระบวนการกำหนดมาตรฐานเพื่อแบ่งความสามารถของผู้สอบที่สอบแบบทดสอบ STOU-EPT ออกเป็น 3 ระดับ ได้แก่ ระดับ A2 B1 และ B2 โดยวิธีการกำหนดมาตรฐานเพื่อหาคะแนนจุดตัดในการวิจัยครั้งนี้ใช้วิธีการของแองกอฟท์ที่ได้รับการปรับปรุง

คะแนนจุดตัด (cut score) หมายถึง จุดคะแนนที่แบ่งผู้สอบที่สอบแบบทดสอบ STOU-EPT ตามความสามารถของผู้สอบ โดยคะแนนจุดตัด แบ่งออกเป็น 2 จุด ได้แก่ จุดที่ 1 คือ จุดคะแนนที่แบ่งความสามารถของผู้สอบที่มีความสามารถระดับ B1 ออกจากผู้สอบที่มีความสามารถระดับ A2 และจุดที่ 2 คือ จุดที่แบ่งความสามารถของผู้สอบที่มีความสามารถระดับ B1 ออกจากผู้สอบที่มีความสามารถระดับ B2

## 1.5 ประโยชน์ที่ได้รับ

ประโยชน์ที่ได้รับจากการวิจัย มีดังนี้

1. แบบทดสอบ STOU-EPT เป็นไปตามกรอบ CEFR และมีคะแนนจุดตัดที่เหมาะสม และมีความคลาดเคลื่อนน้อยที่สุด อันทำให้ผลการสอบน่าเชื่อถือ สามารถอธิบายระดับความสามารถทางภาษาอังกฤษของผู้สอบอย่างแท้จริง ผู้สอบสามารถนำผลการสอบไปใช้เป็นหลักฐาน เพื่อประกอบการสมัครเข้าศึกษาต่อ ขอสำเร็จการศึกษา หรือวัดระดับความสามารถทางภาษาอังกฤษของตนเองว่าตรงกับระดับใดของกรอบ CEFR นอกจากนั้น สถาบันอุดมศึกษาสามารถใช้ผลสอบในการพัฒนาการเรียนการสอนภาษาอังกฤษ โดยเป็นข้อมูลเพื่อปรับปรุงหลักสูตร และแนวทางการสอนให้เหมาะสมกับระดับความสามารถของผู้เรียน รวมทั้งใช้เป็นเครื่องมือวัดความก้าวหน้าทางภาษาของผู้เรียนในระยะยาว

2. ยกระดับการสอบวัดความสามารถทางภาษาอังกฤษของมหาวิทยาลัยสุโขทัยธรรมมาธิราช ให้เป็นสถานที่ให้บริการในการจัดสอบวัดความสามารถทางภาษาอังกฤษที่เทียบเคียงกับกรอบ CEFR ให้กับนักศึกษามหาวิทยาลัยสุโขทัยธรรมมาธิราช และผู้สนใจทั่วไปที่ต้องการทราบระดับความสามารถทางภาษาอังกฤษของตนเองเพื่อเป็นแนวทางในการพัฒนาความรู้ทางภาษาอังกฤษ นอกจากนี้ การใช้กรอบ CEFR เป็นมาตรฐานกลาง ทำให้สามารถเทียบเคียงผลการสอบกับมหาวิทยาลัยอื่น ๆ ที่ใช้มาตรฐานเดียวกันได้มากขึ้น ส่งเสริมความเป็นสากลของแบบทดสอบ STOU-EPT และทำให้ผู้สอบสามารถใช้ผลสอบเป็นข้อมูลอ้างอิงที่ได้รับการยอมรับในวงกว้าง